Helena Jin · Cesar Sciammarella
Sanichiro Yoshida · Luciano Lamberti *Editors*

Advancement of Optical Methods in Experimental Mechanics, Volume 3

Proceedings of the 2014 Annual Conference on Experimental and Applied Mechanics

SEM

Springer

Conference Proceedings of the Society for Experimental Mechanics Series

Helena Jin • Cesar Sciammarella • Sanichiro Yoshida • Luciano Lamberti
Editors

Advancement of Optical Methods in Experimental Mechanics, Volume 3

Proceedings of the 2014 Annual Conference on Experimental and Applied Mechanics

Editors
Helena Jin
Sandia National Laboratories
Livermore, CA, USA

Cesar Sciammarella
Illinois Institute of Technology
Chicago, IL, USA

Sanichiro Yoshida
Southeastern Louisiana University
Hammond, LA, USA

Luciano Lamberti
Politecnico di Bari
Bari, Italy

Preface

Advancement of Optical Methods in Experimental Mechanics, Volume 3: Proceedings of the 2014 Annual Conference on Experimental and Applied Mechanics represents one of eight volumes of technical papers presented at the SEM 2014 SEM Annual Conference & Exposition on Experimental and Applied Mechanics organized by the Society for Experimental Mechanics held in Greenville, SC, June 2–5, 2014. The complete Proceedings also includes volumes on: *Dynamic Behavior of Materials; Challenges In Mechanics of Time-Dependent Materials; Mechanics of Biological Systems and Materials; MEMS and Nanotechnology; Composite, Hybrid, and Multifunctional Materials; Fracture, Fatigue, Failure and Damage Evolution; and Experimental and Applied Mechanics*.

Each collection presents early findings from experimental and computational investigations on an important area within Experimental Mechanics, Optical Methods being one of these areas.

With the advancement in imaging instrumentation, lighting resources, computational power, and data storage, optical methods have gained wide applications across the experimental mechanics society during the past decades. These methods have been applied for measurements over a wide range of spatial domain and temporal resolution. Optical methods have utilized a full range of wavelengths from X-Ray to visible lights and infrared. They have been developed not only to make two-dimensional and three-dimensional deformation measurements on the surface but also to make volumetric measurements throughout the interior of a material body.

Livermore, CA, USA   Helena Jin
Chicago, IL, USA   Cesar Sciammarella
Hammond, LA, USA   Sanichiro Yoshida
Bari, Italy   Luciano Lamberti

Contents

Contents

Chapter 1
The Kinematics of Crystalline Arrays at the Subnanometric Level

C.A. Sciammarella, F.M. Sciammarella, and L. Lamberti

Abstract This article presents additional developments in the analysis of displacement fields and strains around edge dislocations via electron microscopy. The goal of this work, which extends earlier work of the authors, is to provide additional information about the connection between the Continuum Mechanics model of the events taking place in a nanometric size region of a crystal and experimentally observed geometrical changes of the crystalline array.

The implications of the understanding of the connection of Continuum Mechanics and actual distortions of the crystalline array are vast. The development of new semiconductor electronic devices based on nanotubes and other semiconductors in circuits that operate at high temperature, high-power or high radiation will benefit by this input. The new technique of manufacturing, 3D printing, can also benefit from the better understanding of the process of formation of crystalline arrays and the defects that this process generates.

Keywords Kinematics of crystalline arrays • Sub-nanometer observations • Edge dislocations • High resolution transmission electron microscopy • Image analysis

1.1 Introduction

Properties of materials are observable variables that are determined in the macro world. These properties are a consequence of interactions that occur at the atomic level and hence are ruled by Quantum Mechanics. Between these two descriptions of matter there is a big gap and one of the endeavors of the science of Mechanics of Materials is to connect these two worlds. Human thought supported by experimental tools has provided the path way to provide solutions of this extremely difficult scientific question. The search of the solution between structure of matter and macroscopic properties can be traced back to early 19th with Cauchy and his attempt to relate Continuum Mechanics with the atomic structure of matter that resulted in the Cauchy-Born rule [1, 2]. A fundamental help in untangling this complex puzzle has been the advances in the experimental technology of observation of the behavior of matter at the atomic and subatomic levels. The developments of X-rays diffraction, electron microscopy, neutron diffraction, and other methods to observe condensed matter have provided the needed information to understand the structure of solid matter. Crystallography is the science that provides the understanding of the atomic arrangements in solid matter. X-rays give information by analyzing the interaction of electromagnetic waves (photons) with the valence electrons of atoms and ions; electron microscopy provides information on

C.A. Sciammarella (✉)
Department of Mechanical, Materials and Aerospace Engineering, Illinois Institute of Technology, 10 SW 32nd Street, 60616 Chicago, IL, USA

Department of Mechanical Engineering, College of Engineering and Engineering Technology, Northern Illinois University, 590 Garden Road, 60115 DeKalb, IL, USA
e-mail: sciammarella@iit.edu

F.M. Sciammarella
Department of Mechanical Engineering, College of Engineering and Engineering Technology, Northern Illinois University, 590 Garden Road, 60115 DeKalb, IL, USA

L. Lamberti
Dipartimento Meccanica, Matematica e Management, Politecnico di Bari, Viale Japigia 182, 70126 Bari, Italy

the spatial arrangement of matter through the interaction of electrons, charged particles, with atoms and ions, describing the total charge distributions. Neutrons interact with atomic nuclei through the phenomenon of diffraction.

By combining X-rays, electron microscopy and neutron diffraction, details of the structure of the organization of matter have been investigated and the arrangement of atoms and ions in complex structures revealed. These breakthroughs in technology provided the necessary tools to describe the structure of matter from the geometric point of view. These experimental observations have also enabled the mechanics of deformations to be the subject of theoretical and experimental investigations at different scales. As a result, the theory of Continuum Mechanics has been extended from the field of macroscopic bodies to the structure of crystalline matter. A large amount of work has been done to relate the internal local forces in crystalline structures expressed by a suitable stress tensor definition to the deformations expressed by a compatible strain tensor definition. These efforts have resulted in those aspects of the mathematical theory of dislocations closely related to the theory of elasticity. A far more difficult task has been the effort to connect the mathematical theory of plasticity with events observed at the matter structure level.

A key aspect in understanding the plastic behavior of matter was the introduction of the concept of defects in crystalline materials, dislocations and point defects. Completely independent of the analysis of defects in crystalline structures, in the 1920s, the concept of a dislocation was introduced by Volterra [3], and later analyzed by Somigliana [4] and Love [5]. Volterra defined the basic type of dislocations that today are utilized in fracture mechanics. These abstract concepts originated in the analysis of the uniqueness of displacements fields came back to crystallography as a result of the need to explain discrepancies between theoretical and experimental prediction of crystal strength (see, for example, Taylor [6], Orowan [7], Polanyi [8] and Nabarro [9]). Starting with these initial developments a new discipline came into existence, the mathematical theory of dislocations [10].

The first direct experimental observation of the presence of dislocations in crystalline structures took place in the 1960s with the use of X-rays. This visualization came about through X-ray photogrammetry. In the 1970s the use of the electron microscope provided another tool to look at the structure of crystalline matter. This article presents additional developments in the analysis of displacements fields and strains around edge dislocations via electron microscopy extending earlier work of the authors [11–13]. As the kinematics of crystal structures based on finite deformations is a basic step in the mathematical theory of dislocations, this study investigates the connection between the Continuum Mechanics model of the events taking place in a nanometric size region of a crystal and experimentally observed geometrical changes of the crystalline array.

The implications of the understanding of the connection of Continuum Mechanics and actual distortions of the crystalline array are vast. The development of new semiconductor electronic devices based on nanotubes and other semiconductors in circuits that operate at high temperatures, high-power or high radiation will benefit by this input. In the new manufacturing technology, 3D printing, laser beams are used to melt powders of the material that is being formed into a given geometrical shape. The energy balance between laser beam power and energy required to form a given geometrical shape is a very important subject of research. This energy balance must produce a material structure with an optimum arrangement of atoms and ions for the particular purpose that the part is fabricated for. Hence, advances in the understanding of the mathematical theory of dislocations will be of great importance in the technology of 3D printing.

1.2 High Resolution Transmission Electron Microscope Images

For the full understanding of the observed images it is necessary to go over a brief explanation of the workings of an electron microscope (see the schematic shown in Fig. 1.1). The images obtained in high resolution transmission electron microscope (HRTEM) have a process of formation that is similar to that of an interference microscope. There is an illuminating beam of coherent electrons (polarized) that is focused on an object, an aperture diaphragm and a focusing lens. Using Fourier optics terminology, the focusing lens creates the diffraction pattern of the object in its back focal plane, and then through the diaphragm it performs an inverse FT. The direct beams through the specimen interfere with the diffracted beams and produce an image of the object. The incident wave front of the electron beam, ideally a plane wave front, interacts with the structure of the object.

The resolution of the microscope from the point of view of the image formation can be expressed in terms of the classic Rayleigh criterion. The smallest distance that can be observed is given by the equation

$$\delta = \frac{0.61\lambda}{n_r \sin \alpha} \tag{1.1}$$

Fig. 1.1 Schematic of high
resolution transmission
electron microscope
(HRTEM)
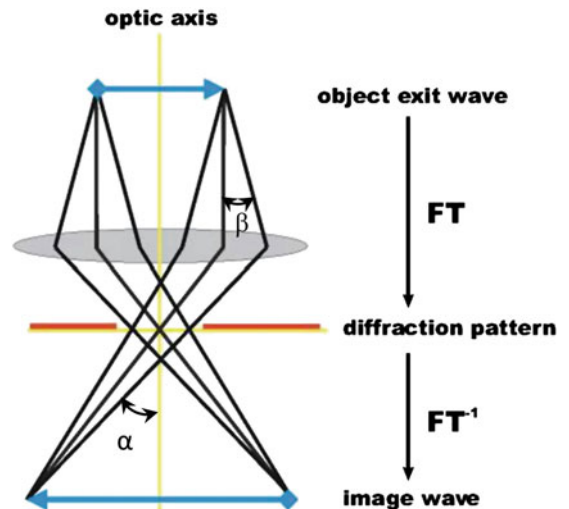
where λ is the wavelength of the radiation, n_r the index of refraction of the medium between the imaging lens and the image plane, α the semi-angle of the imaging lens. The wave length associated with electrons can be approximately computed as

$$\lambda = \frac{1.22}{\sqrt{E_g}} \qquad (1.2)$$

In Eq. (1.2), the wavelength of the wave associated with the electron propagation is given in nm and E_g is the energy imparted to the wave expressed in eV. For a HRTEM the power of the electron beams could be for example 400 K$_{eV}$. Applying Eq. (1.2), one obtains for example $\lambda = 0.00193$ nm. Since atomic radii are of the order of 0.1 nm, the wavelength of the electron beam of a HRTEM is well within the resolution of the atomic radii.

It should be noted that Eq. (1.1) applies to the case of incoherent illumination while the HRTEM image has a high degree of coherence. Furthermore, the resolution of the microscope is affected by lens aberrations: the quality of magnetic lenses is poor and may affect the information retrieved from the image.

Information detection can be represented by the conventional mathematics of amplitude and phase modulation. The emerging wavefronts are given by a phasor (see Chap. 7.4.6 of [14]):

$$\mathbf{E}(\mathbf{r}, t) = E(\mathbf{r}, t)\, e^{i\phi(r,t)} \qquad (1.3)$$

In Eq. (1.3), bold letters indicate vectors, E($\mathbf{r}$,t) is the amplitude of the wave associated with the electron propagation and ϕ the corresponding phase, $\mathbf{r}$ is the position vector and t is the time.

As stated before, magnetic lenses have aberrations (e.g. astigmatism, spherical, chromatic) that are included in the information retrieved by the wave front. The information is captured as an intensity distribution detected in modern HRTEM by a raster sensor similar to electronic cameras. Image resolution is affected by the sensor similarly to what occurs in electronic cameras [15]. It is possible to correct the experimentally determined intensity via software by utilizing the concept of contrast transfer function, the equivalent of the optical transfer function used in visible optics.

In order to understand the source of intensity distribution observed in electron microscopy, it is necessary to briefly analyze the interaction of matter with an electronic beam: that is, the diffraction phenomenon of electrons by matter, in this particular case by crystalline matter. This is an extremely complex subject and deals with the duality of particles and waves, it can only be rigorously explained in quantum mechanics terms. However, in the particular case being considered it can be summarized in few lines. When an electronic wave impinges in the cross-section of an atom different interactions take place. Of interest in our case is the stimulated emission of waves of the same frequency of the incident radiation. This interaction gives rise to waves of amplitudes proportional to the electronic density of the atom or ionic structure. This electronic density, since we are dealing with images captured on a plane, is a two dimensional density. That is, if we are observing an image plane the maximum observed intensity will take place in the region where the atomic nucleus is located and then the observed distribution can be utilized as a tool to define the position of the atom in 2D.

In order to locate atoms in 3D the different operation modes of the electron microscope can be used taking advantage of other types of interactions between electronic waves and observed atoms. Absorption is a very important mechanism: electronic waves can only penetrate matter in a dimensional range of few nm to 100 nm. Because the absorption coefficient increases with the energy of the electronic wave, if one uses 400 KeV the specimens must have a thickness of few nm.

To complete the understanding of the observed images, it is necessary to analyze the process of image formation in some more detail. One is observing a specimen of a certain thickness, d_{ob}, and this thickness is imaged at a given plane. It is necessary to find out what it is the relationship of the observed image and d_{ob}. We can go back to basic concepts of classical geometrical optics: depth of field and depth of focus. The depth of field is measured in the object space and tells us how much of d_{ob} is in focus, that is how much the object can be displaced from a given position back and forth without loosing the focus in the image plane. The depth of focus refers to the image plane and tells us how much back and forth from the focal plane one can move the detector plane without an apparent loss of focus. The depth of field and the depth of focus depend on the illumination aperture, the inclination of the beam entering the lens (angle β in Fig. 1.1), and the inclination of the beam forming the image (angle α in Fig. 1.1). In HRTEM, paraxial beams are utilized and the angles are small enough that all the features contained in d_{0b} are practically in focus [16].

Another important concept that will help the interpretation of the observed images is the concept of resolution. In the present case, resolution is the minimum distance that can be measured in the image plane [16]. As we are going to see by utilizing pixel interpolation in the image plane this distance can be reduced.

1.3 The 4HSiC Crystal

Before we can proceed further it is necessary to provide additional information on the analyzed image. In earlier publications of the present authors [11–13], an image of a 4HSiC crystal was analyzed [17]. That particular image was selected because it is supported by relevant information. It was also of interest to study the basic unit of a hexagonal polytype of SiC whose schematic is shown in Fig. 1.2.

Two inches wafers were grown by vapor deposition oriented along the plane (0001) within $\pm 0.5°$ of SiC crystal. KOH etching at 500 °C or 510 °C was done on the Si (0001) faces of the wafers. The etch pits were observed with Nomarski interferometry while X-rays were utilized to observe the orientation of each pit array.

The most common polytypes require four and six Si-C bi-layers, respectively, to define the unit cell repeating distance along the c-axis [0001] directions. Figure 1.2 shows the sequence ACAB of the 4HSiC crystal; this sequence is relative because the radii of the atomic regions corresponding to each atom all have the same magnitude. Hence, relative positions result from the possible geometrical arrangements of spheres. Figure 1.2a shows the arrangement denoted by the letter A and assuming a Si atom as starting the sequence. The stacking arrangement is outlined in Fig. 1.2b and a perspective of the three initial layers is also illustrated in Fig. 1.2c.

Figure 1.3a shows an optical micrograph of the etch pit bands—(0001) face—of a 4HSiC wafer. The bands run along directions $\langle 11\bar{2}0 \rangle$ and reveal the presence of threading edge dislocations. The theoretical structure of the hexagonal polytype of silicon carbide with $a_1 = a_2 = a_3$ is shown in Fig. 1.3b.

Figure 1.4 illustrates the process of formation of the dislocations seen in Fig. 1.3. This type of threading edge dislocation originated from the cooling of the SiC crystal which then caused a misoriented crystal to grow (Fig. 1.3). The large amount
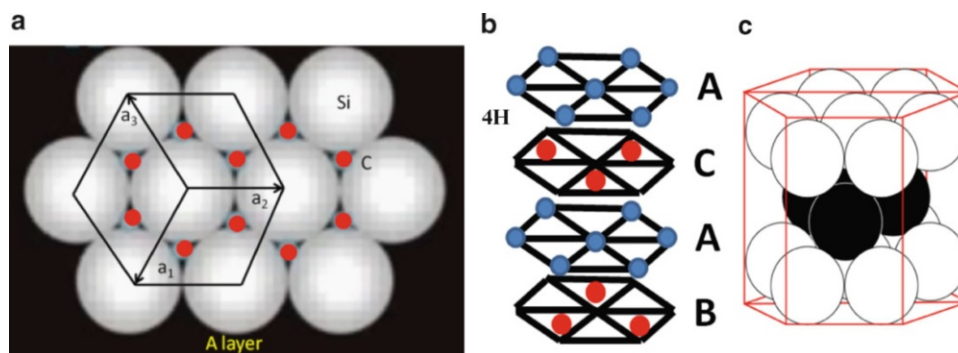
Fig. 1.2 Staking sequence ACAB of the 4HSiC crystal. The *red dots* in (**a**) indicate the position of the following atom layer shown in (**b**); Perspective of the three initial layers is shown in (**c**)

Fig. 1.3 (**a**) Optical micrograph of etch pit bands—(0001) face—of a 4HSiC wafer with bands along directions $\langle 11\bar{2}0 \rangle$; (**b**) Theoretical structure of the hexagonal polytype of silicon carbide with $a_1 = a_2 = a_3$
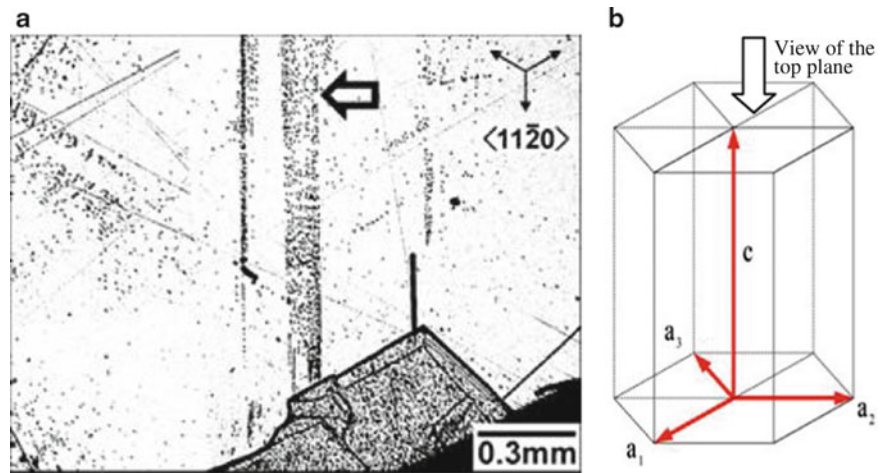
Fig. 1.4 Gliding of dislocations in the slip system $\langle 11\bar{2}0 \rangle \{\bar{1}100\}$ during post-growth cooling: (**a**) Optical micrograph; (**b**) electron microscope image of the red circled region; (**c**) Theoretical SiC structure with $a_1 = a_2 = a_3$ and Burgers vector of dislocation oriented as a_3; (**d**) HRTEM pattern of the crystal (*dark dots* indicate the extra-atomic planes)
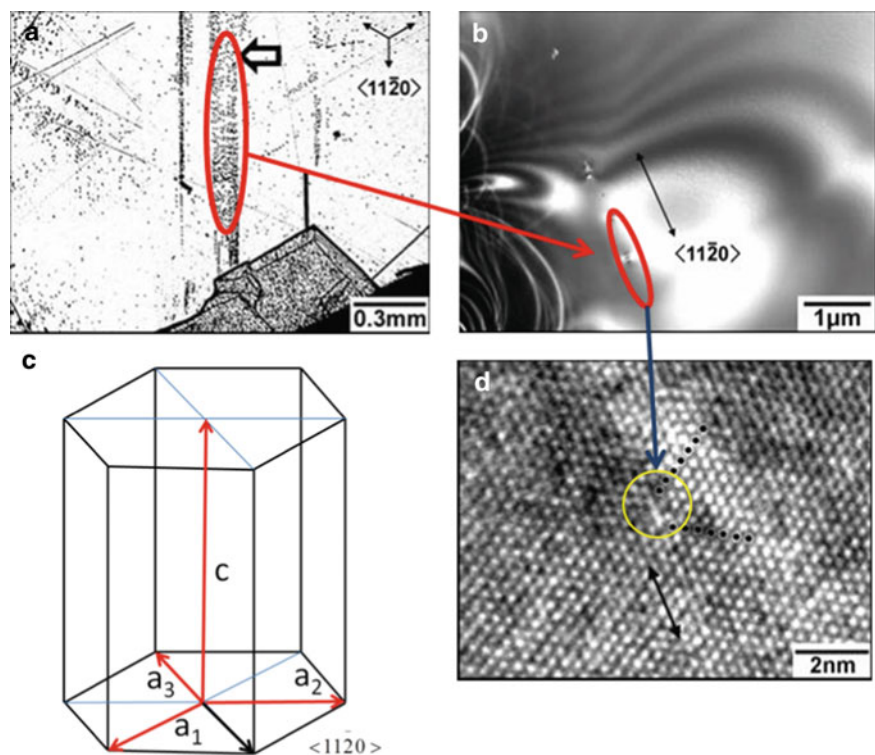
of energy that concentrated in that region led to the formation of dislocations (slip bands) that caused the plane to slide as illustrated in Fig. 1.4. As a result of this dislocation movement two extra half planes—indicated by the black dots in Fig. 1.4—were generated to reduce the energy in the unit cell thus allowing stability of the crystal structure to be achieved.

Figure 1.4 shows together the optical micrograph (see Fig. 1.4a) with a electron microscope image of an etch pitch (see Fig. 1.4b) and a HRTEM image around a threading dislocation in the array (see Fig. 1.4d). The Burgers vectors of the edge dislocations are of the type a/3 $\langle 11\bar{2}0 \rangle$ (see Fig. 1.4c) where a is the 4HSiC lattice parameter.

1.4 HRTEM Image Analysis

The HRTEM image encompasses areas where the basic cell of the crystal is highly distorted and thus provides information suitable to perform the analysis of deformations at different scales and relate macro continuum mechanics to deformations in the micro- and nano-scales. The first step of analysis is to interpret the images produced by HRTEM. Image formation is the
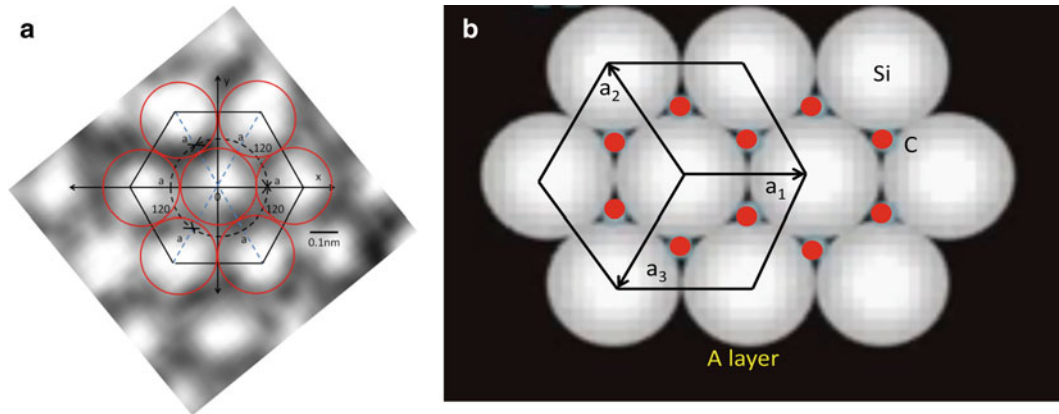
Fig. 1.5 Elementary cell of the hexagonal packing identified in the HRTEM pattern resembling the theoretical structure represented in (**b**)

result of complex processes involving subject matters of great theoretical complexity and difficulty. The very thin specimen investigated in this study is the order of 10 nm and hence contains a large number of atomic planes, more than 30. In Sect. 1.2, it was pointed out that usually the depth of focus is large enough that the whole depth of the specimen is in focus [16]: that is, features at different depths are shown in the image. To evaluate the image content, we can analyze the actual image of a hexagonal elementary cell shown in Fig. 1.5a. The cell is extracted away from the region circled in yellow in Fig. 1.4c. Figure 1.5b recalls the theoretical atomic structure explained in Fig. 1.2.

Different aspects of the image formation can be evaluated. It is easy to notice the similitude between the experimentally observed image and the theoretical structure of Fig. 1.2a. It is possible to conclude that the observed pattern is the image of an atomic layer. First it is necessary to explain the process followed to put together this image. From the original image (Fig. 1.4c) a small square region containing the elementary cell was cropped. A bicubical spline interpolation of the pixels was applied resulting in the scale shown in Fig. 1.5. The red circles in the figure have the radius of the spheres r = a/2 = 0.3073/2 = 0.1537 nm, that corresponds to the lattice parameter a = 0.3073 nm at 300 K. With the above parameter the corresponding hexagon connecting the center of the spheres has been drawn.

The question of the minimum distance that can be measured in the image can be now addressed. The image of the crystal was recorded with a HRTEM JEOL 4000 EX-TEM operating at 400 KeV [18]. According to Eq. (1.2) the wavelength of the electron wave is 0.00193 nm that is of the order of 0.002 nm. Since the lattice parameter a is 0.3073 nm the ratio a/λ is 159. According with the criteria of optical microscopy, with this wavelength one should be able to resolve $\lambda/2$: that is, we should be able to measure 0.001 nm. We have seen that this is not feasible in electron microscopy and the resolution defined as the fine detail of an image is much lower than the theoretical value. However, in electron microscopy another resolution concept is utilized [16]: the point to point distance, this is the concept that applies to the measurements that we perform in the analyzed image. The HRTEM JEOL 4000EX-TEM is rated [18] as having a maximum classical resolution of 0.14 nm. The original of the picture has a pixel value S = 0.0368 nm/pixel. The sub-image has a pixel value S = 0.0011 nm/pixel that is a gain in resolution of 33 times. The bicubical spline interpolation has yielded a pixel resolution of the order of 0.001 nm which implies a 140 increase of sensitivity to measure distances with respect to the rated minimum classical resolution of 0.14 nm.

1.5   Analysis of the Kinematics of the Distorted Region

Figure 1.6a shows the positions of crystallographic axes in the electron microscope patterns. The red dots show the extra atomic rows. Figure 1.6b shows an enlargement of the dislocation region showing the position of two edge dislocations generated by the added extra atomic planes. The analysis of the kinematics of the elementary cells is based on the frame work and limitations that came from the idealization of 3D array of atoms as a superposition of identical 2D array of atoms stacked in the vertical direction. Since this stack is not identical in depth due to the sequence shown in Fig. 1.2 based in the covalent bonding of Si atoms and C atoms, the analysis is restricted to a two-dimensional array.

Figure 1.7 represents the FFT of the HRTEM pattern of the crystal shown in Fig. 1.4c. The direction f_x corresponds to the family of crystallographic directions $\mathbf{a_1} \equiv [1000]$ (See Fig. 1.6a). The scale of Fig. 1.7 is $S_f = (0.0111)$ 1/nm per pixel.

Fig. 1.6 (**a**) Crystallographic directions $\langle 11\overline{2}0 \rangle \equiv \mathbf{a}_3$; (**b**) Schematic view of the position of atoms in the neighborhood of the edge dislocations (the meaning of the *yellow dots* will be discussed later on the paper)
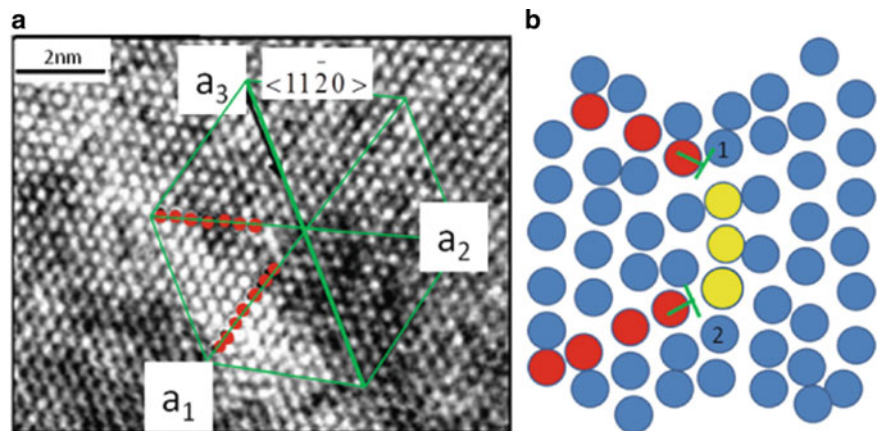
Fig. 1.7 (**a**) FFT of the HRTEM pattern shown in Fig. 1.6a; (**b**) Reference system associated with the FFT
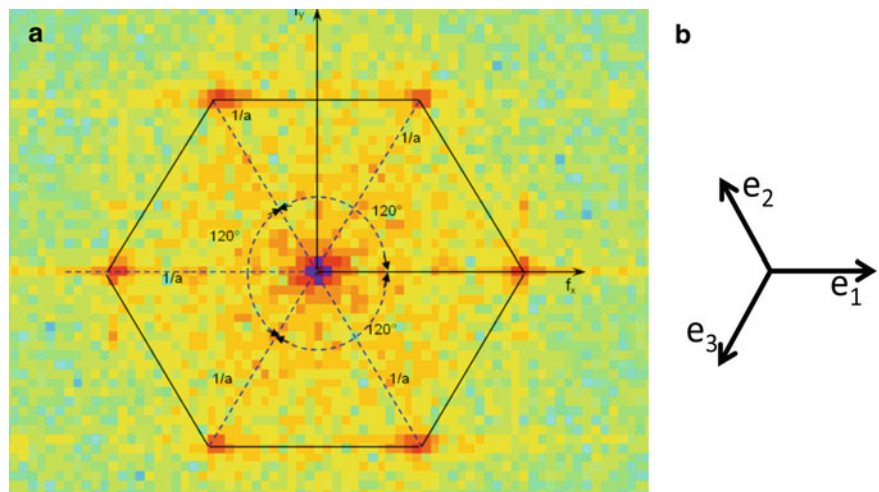
Figure 1.7 provides a statistical average of the changes of the elementary cells in the region under analysis. The hexagon represented in Fig. 1.7 corresponds to the average size of **a** at 300 K. The departures of the positions of the corners of the hexagon with respect to the average position indicated in Fig. 1.7 reflect the distortions of the elementary cells caused by the presence of the dislocations. Figure 1.2a (i.e. Fig. 1.5b) represent the elementary undeformed cell that is characterized for this particular crystal by the constancy of the atomic vector **a**. This vector reflects the spatial configuration of the atomic bonds that bind together the atoms. Since atoms are in continuous oscillatory motion caused by the temperature T, **a** is a statistical average in time. Figure 1.7 shows that upon the deformation caused by the presence of the dislocations the values of **a** are modified. In the 2D approach we can analyze the changes of the elementary cell as superposition of different changes in the dimensions and orientations of **a**.

Figure 1.8 shows the changes of the elementary cell that can be superimposed to get the different configurations statistically observed in Fig. 1.7. All the main diagonals of the hexagon show increases and decreases of the value of the modulus of **a**. Increasing frequencies correspond to shortening of a_1, modulus of **a** in the direction [1000]. Lower frequencies correspond to elongations of a_1. Let a_{1r} be the value of a_1 for 300 K and $a_1(+f)$ the value of a_1 for increasing frequencies: it holds $a_1(+f) < a_{1r}$. Conversely, if $a_1(-f)$ is the value of a_1 for decreasing frequencies, it holds $a_1(-f) > a_{1r}$. The corresponding distribution of a_{1r} is asymmetric with wider ranges for the $a_1(-f)$ than for the $a_1(+f)$. This distribution reflects the asymmetry of the potential function that determines the forces binding atoms. From the position of equilibrium a_{1r}, the repulsion forces between atoms have a much steeper gradient that the attraction forces between atoms as they move apart. This asymmetry is very important in the analysis of the kinematics of the elementary cell. In solid mechanics and for small deformations, compression and tension displacements are considered symmetric. This is not the case in the deformation of the elementary cell. It is possible to see also in Fig. 1.7 that the diameters of the hexagon experience rigid body rotations. Because the size of the region under observation is very small, rigid body rotations also are very small.
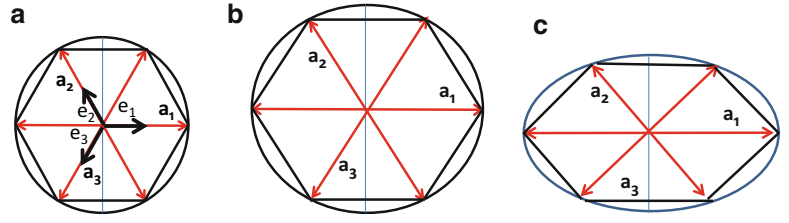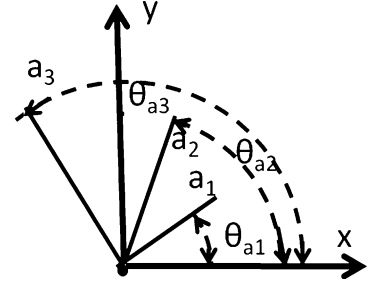
Figure 1.8 shows the basic system of reference with the versors $\mathbf{a_i}$ ($i = 1, 2, 3$). The displacements of the atomic positions in the deformed configuration in the hexagon can be expressed as:

$$\mathbf{u} = u_{a1}\mathbf{a}_1 + u_{a2}\mathbf{a}_2 + u_{a3}\mathbf{a}_3 \tag{1.4}$$

If the displacement components along three versors are available, the displacement derivatives with respect to these directions that enter in any definition of strain tensor adopted in the analysis are related with the Cartesian components of the displacement derivatives. Looking at Fig. 1.9, the following equations can be written:

$$\begin{cases} \dfrac{\partial u_{a1}}{\partial a_1} = \dfrac{\partial u}{\partial x}\cos^2\theta_{a1} + \dfrac{\partial v}{\partial y}\sin^2\theta_{a1} + \left[\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right]\cos\theta_{a1}\sin\theta_{a1} \\[2mm] \dfrac{\partial u_{a2}}{\partial a_2} = \dfrac{\partial u}{\partial x}\cos^2\theta_{a2} + \dfrac{\partial v}{\partial y}\sin^2\theta_{a2} + \left[\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right]\cos\theta_{a2}\sin\theta_{a2} \\[2mm] \dfrac{\partial u_{a3}}{\partial a_3} = \dfrac{\partial u}{\partial x}\cos^2\theta_{a3} + \dfrac{\partial v}{\partial y}\sin^2\theta_{a3} + \left[\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right]\cos\theta_{a3}\sin\theta_{a3} \end{cases} \tag{1.5}$$

From the above equations, it is possible to determine the principal derivatives with respect to coordinate system X-Y and the orientation of the principal directions with respect to these axes:

$$\begin{cases} \dfrac{\partial u_1}{\partial x} = \dfrac{\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y}}{2} + \sqrt{\left[\dfrac{\dfrac{\partial u}{\partial x} - \dfrac{\partial v}{\partial y}}{2}\right]^2 + \left[\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right]^2} \\[6mm] \dfrac{\partial u_2}{\partial y} = \dfrac{\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y}}{2} - \sqrt{\left[\dfrac{\dfrac{\partial u}{\partial x} - \dfrac{\partial v}{\partial y}}{2}\right]^2 + \left[\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right]^2} \end{cases} \tag{1.6}$$
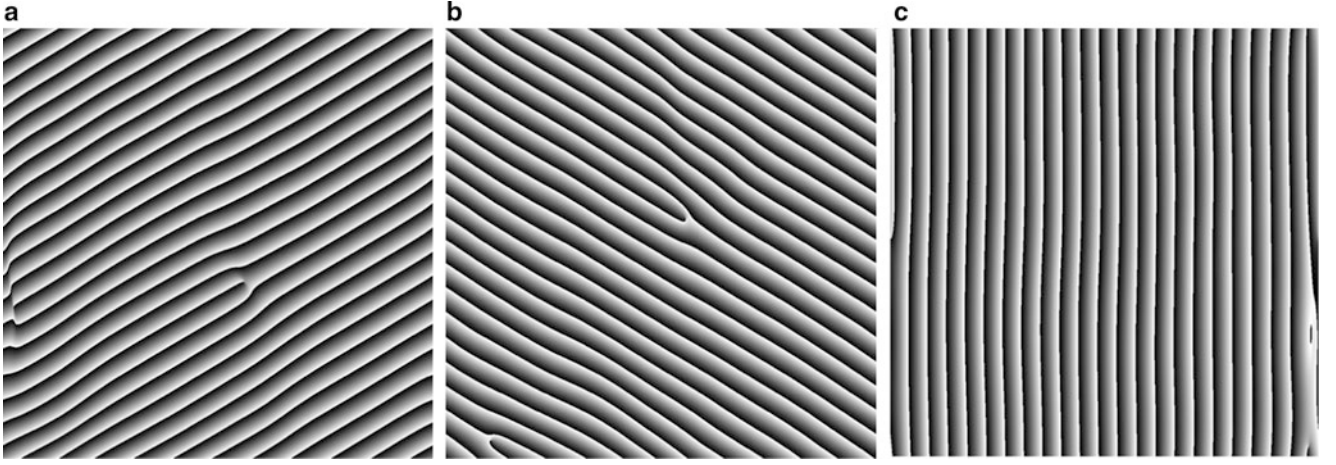
Fig. 1.10 Phases of the harmonics that modulate the structure of the crystal: (**a**) 240° in FFT space; (**b**) 120° in FFT space; (**c**) 0° in FFT space, respectively correspond to directions $\mathbf{e}_2$, $\mathbf{e}_3$ and $\mathbf{e}_1$

$$\tan 2\boldsymbol{\theta} = -\frac{\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}}{\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}} \tag{1.7}$$

The above quantities can be determined experimentally to apply the preceding equations. It is important to realize that the above relationships provide an Eulerian description of the deformed crystal and the adopted system of reference is attached to the structure of the crystal. In the derivation of Eqs. (1.5, 1.6, and 1.7) it was assumed that, upon deformations, the directions $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ change orientations with respect to the initial configuration. The above equations can be applied to the deformed shapes of the elementary cells located in different positions in Figs. 1.4c and 1.6a.

There is an alternative analysis procedure based in extending Continuum Mechanics to the field under observation. Figure 1.7a shows the FFT of the crystal structure and the labeling of the corresponding axes in the physical space. These axes correspond to the undeformed crystal orientations, with axis $\mathbf{e}_1$ oriented along the crystallographic family of directions $\mathbf{a}_1 \equiv \, <1000>$. We can introduce the concept of digital moiré (see Chap. 13.8 of [14]) and through this procedure provide a tool to analyze the kinematics of the hexagonal crystal at the level of the elementary cell. Symbolically, the deformation of the elementary cell can be represented as a function $F(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,t)$ and the gradient of this function represented by $\nabla F(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3, t)$ provides the derivatives that characterize the state of deformation (see Fig. 1.8). From the inverse FFT of Fig. 1.7 we can get three systems of fringes at orientation 120° to each other. We have the equivalent of a three elements rosette with arms at 0°, 120° and 240°. The correspondence between rosette orientations and directions (orthogonal to fringe orientation) in the FFT space is also shown in the figure. Calling $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ the axes of the rosette as shown in Fig. 1.8a and following a similar argument to that utilized for deriving Eqs. (1.6 and 1.7), principal derivatives can be obtained as a function of the derivatives measured along the rosette axes:

$$\frac{\partial u_1}{\partial x_1}, \frac{\partial u_2}{\partial x_{2p}} = \frac{1}{3}\left(\frac{\partial u_{e1}}{\partial e_1} + \frac{\partial u_{e2}}{\partial e_2} + \frac{\partial u_{e3}}{\partial e_3}\right) \pm \frac{\sqrt{2}}{3}\left[\sqrt{\left(\frac{\partial u_{e1}}{\partial e_1} - \frac{\partial u_{e2}}{\partial e_2}\right)^2 + \left(\frac{\partial u_{e2}}{\partial e_2} - \frac{\partial u_{e3}}{\partial e_3}\right)^2 + \left(\frac{\partial u_{e3}}{\partial e_3} - \frac{\partial u_{e1}}{\partial e_1}\right)^2}\right] \tag{1.8}$$

where: x_1 and x_2 are the coordinate axes; $e_1 \equiv x_1$ and the x_2-axis is orthogonal to x_1.

The principal directions of the derivatives are obtained from the following relationship:

$$\boldsymbol{\theta} = \frac{1}{2}\text{arctg}\left\{\frac{\sqrt{3}\left(\frac{\partial u_{e3}}{\partial e_3} - \frac{\partial u_{e2}}{\partial e_2}\right)}{2\frac{\partial u_{e1}}{\partial e_1} - \left(\frac{\partial u_{e2}}{\partial e_2} + \frac{\partial u_{e3}}{\partial e_3}\right)}\right\} \tag{1.9}$$

From Eqs. (1.8 and 1.9) one can compute principal derivatives and obtain all the necessary information on the elementary cell that has been adopted. Figure 1.10 shows the phase of the fringes obtained filtering the diffraction pattern around the directions defining $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,t_o)$ where time t_o indicates the fact that what the microscope shows is a time average of the

events that took place in the observed thin slice. HRTEM patterns were processed with the Holo Moiré Strain Analyzer image processing software developed by C.A. Sciammarella and his coworkers [20].

The phase patterns show phase dislocations (see Chap. 10.5.1 of [14]). Moiré patterns provide the Eulerian description of the deformed elementary cell and again the versors are oriented in coincidence with crystallographic directions highlighted in Fig. 1.8a. The preceding developments in the kinematics of the elementary cell can be considered as an extension of the conceptual idea contained in the so called 1st Cauchy–Born rule. It directly connects the Continuum Mechanics kinematics of solids to the movement of atoms in the hexagonal crystal in a two-dimensional context.

1.6 Analysis of Elementary Cells

To further shed light on the proposed approach to the crystal kinematics, it is interesting to survey some regions of the HRTEM pattern shown in Figs. 1.4c and 1.6a. The scale of Fig. 1.11a is $S_c = 0.000900$ nm/pixel, the axis $\mathbf{a_1}$ is oriented as the crystallographic direction [1000]. In Fig. 1.11a, the green lines represent the limits of the region of influence of each atom. These outlines are no longer circles but ellipses, they fit better the elementary cell configuration and the elementary cell is contained in an ellipse that is very closely similar to the shapes of the green ellipses that define the atoms positions of equilibrium. Principal strains can be computed: $\varepsilon_1^e = 0.0501$, $\varepsilon_2^e = -0.0332$, the principal direction ε_1^e lays along the crystallographic direction [1000] while the other principal direction ε_2^e is perpendicular to it. These results are compatible with the frequencies of the diffraction pattern shown in Fig. 1.7.

Figure 1.11b shows an elementary cell located in the compression region. The scale of Fig. 1.11b is $S_c = 0.0009789$ nm/pixel. The principal strain $\varepsilon_1^e = -0.02550$ and coincides with the direction [1000], $\varepsilon_2^e = -0.05240$ and is perpendicular to this direction. The above shown results are also compatible with the frequencies of the diffraction pattern shown in Fig. 1.7.

Figure 1.12 shows the region of the dislocation and the positions of the atoms in this region. The red dots represent the atoms of the extra-rows. The yellow lines indicate the direction of the extra rows that are shown also as red dots in Fig. 1.6. The edge dislocations are indicated with the green symbols ⊥, the vector B indicates the direction of the Burgers vector of the edge dislocations.

Two basic hexagonal cells are indicated in the vicinity of the extra atomic rows (Fig. 1.12). The dislocations interrupt the basic cell arrangement corresponding to the equilibrium condition shown in Fig. 1.5b. The red circles indicate the intersections of the plane of the image with the spherical regions that correspond to the equilibrium conditions (Fig. 1.5b). The intensity distributions within these circles are functions of the electron density with the maxima in correspondence with the nuclei of the atoms. The atoms positions are outlined by the corresponding circles with the exception of the atoms
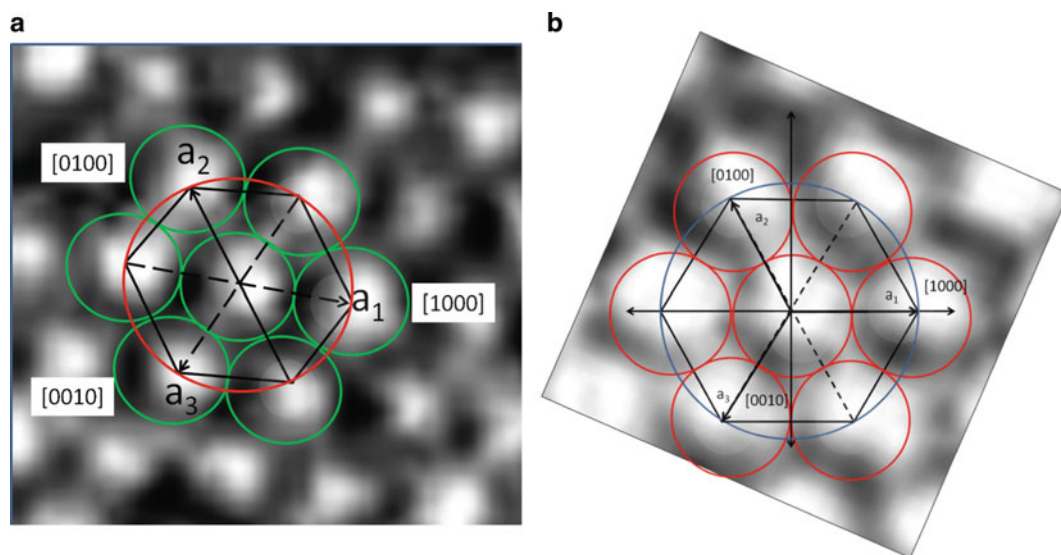
Fig. 1.11 (a) Elementary cell in the tension region of Fig. 1.6a; (b) Elementary cell in the compression region
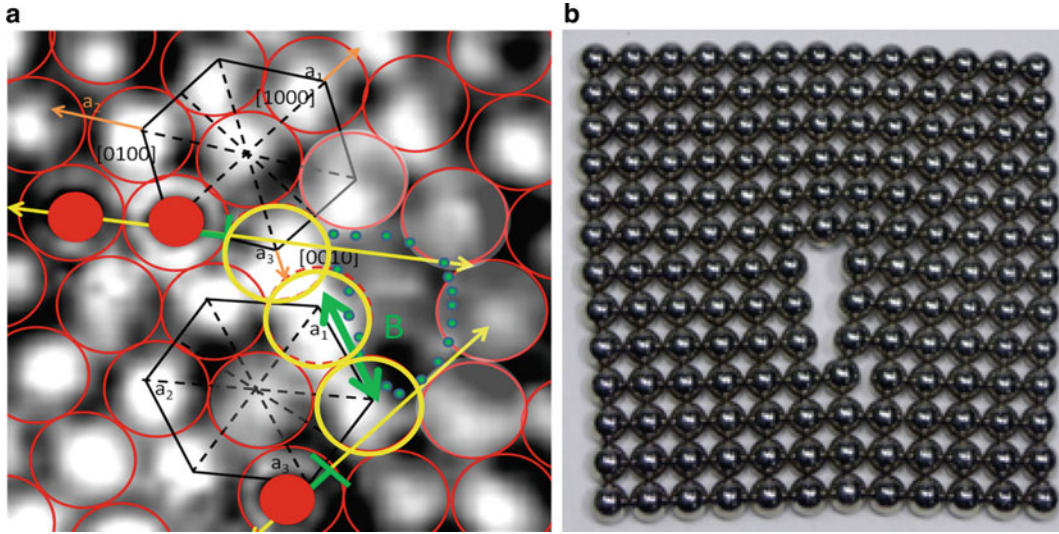
Fig. 1.12 (**a**) Atoms location in the crystal region hosting the dislocation; (**b**) Simulation of the presence of a dislocation using magnetic spheres [19]

just after the extra-rows end indicated by yellow circles in Fig. 1.12a and outlined by yellow dots in Fig. 1.6b. These images show high intensity with little variation like blurred images. Just behind this atoms row there is a region of very low intensity indicating a low density of electrons. This region interrupts the sequence of the hexagonal patterns corresponding to the equilibrium positions of the atoms and is outlined with green dots. This region is the beginning of a cavity that located in the intersection of the yellow lines defining the orientation of the extra-rows. The next row of atoms, close to the yellow lines, shows a deformation in the direction of the Burger vector, $\langle 11\bar{2}0 \rangle$, of approximately +13 %. This quantity is also compatible with the frequency pattern of Fig. 1.7.

Figure 1.12b shows the packing of rigid magnetic spheres; it is possible to see that the model of rigid spheres, although an interesting visualization, does not provide the actual field of atoms seen before that can be approximated by deformable ellipsoids. In the theory of dislocations itZ is assumed that a sufficient large pile up of dislocation can transform the outlined area into an actual physical cavity, the starter of a crack. The cavity formed between the atoms has an interesting property that makes its presence detectable. Because of the low electronic density if positrons are sent to this cavity a pair formed by an electron and a positron is created, this pair is not stable and decays into two γ rays. Using a γ-rays detector it is possible to count the number of events and this number can be used to measure the number of cavities present in a specimen [21].

1.7 Extension of Continuum Kinematics at Inter-atomic Spaces

As mentioned before, three systems of fringes at orientation 120° to each other can be obtained from the FFT shown in Fig. 1.7: this is the equivalent of a three arm rosette. The correspondence between rosette orientations and directions (orthogonal to fringe orientation) in the FFT space is also shown in Fig. 1.7. From Eqs. (1.8 and 1.9) we can get the derivatives that can be input into a definition of strain tensor. For example, the components of the Almansi's non-linear strain tensor referred to the Eulerian coordinates can be computed from these derivatives. For the coordinate system shown in Fig. 1.6a with $\langle 1000 \rangle \equiv x_1$ and x_2 in the perpendicular direction, it can be obtained the Jacobian tensor [22]:

$$[\mathbf{J}] = [\mathbf{I}] - [\mathbf{F}]^{-1} = \begin{bmatrix} \dfrac{\partial \mathbf{u}_1}{\partial x_1} & \dfrac{\partial \mathbf{u}_1}{\partial x_2} \\ \dfrac{\partial \mathbf{u}_2}{\partial x_1} & \dfrac{\partial \mathbf{u}_2}{\partial x_2} \end{bmatrix} \tag{1.10}$$
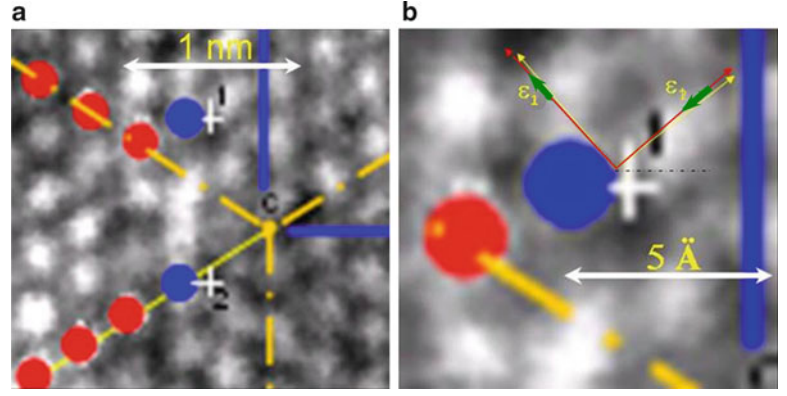
Having known the components of the tensor [J], one can compute the components of the Almansi's strain tensor as follows:

$$\varepsilon_{11}^{A} = \frac{\partial u_1}{\partial x_1} - \frac{1}{2}\left[\left(\frac{\partial u_1}{\partial x_1}\right)^2 + \left(\frac{\partial u_2}{\partial x_1}\right)^2\right] \tag{1.11}$$

$$\varepsilon_{22}^{A} = \frac{\partial u_2}{\partial x_2} - \frac{1}{2}\left[\left(\frac{\partial u_1}{\partial x_2}\right)^2 + \left(\frac{\partial u_2}{\partial x_2}\right)^2\right] \tag{1.12}$$

$$\varepsilon_{12}^{A} = \frac{1}{2}\left[\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right] - \left[\frac{\partial u_1}{\partial x_1}\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\frac{\partial u_2}{\partial x_2}\right] \tag{1.13}$$

The above components are referred to the initial undeformed system of reference. Since the final coordinate system has a rigid body rotation, one can correct the derivatives for this rigid body rotation. In this case the corrections are negligible.

In Fig. 1.6b, there are indicated the atoms 1 and 2 in whose neighborhood strain tensor components have been computed. Figure 1.13a shows again the same region of Fig. 1.6b while Fig. 1.13b shows the components of the Almansi's strain tensor. The value of the maximum strain is $e_{11}{}^{A} = 0.5886$ and it is almost in the direction of $\langle 11\overline{2}0\rangle$, the rigid body rotation is 1.6°; the minimum strain is $e_{22}{}^{A} = -0.3593$ and it is almost perpendicular to $\langle 11\overline{2}0\rangle$.

The fringe analysis provides important information that is being retrieved from the 2D image produced by the electron microscope. The kinematics of the continuum has been extended to the inter-atomic space by the utilization of digital moiré. The presence of fringe dislocations in the fringe patterns in the same position as physical dislocations indicates the place where there are discontinuities in the displacement field (see Chap. 10.5.1 of [14]). The analyzed region is located inside the atomic field. Strains in this region must be compressive in one direction because in this region extra atomic planes have been introduced during the crystal growth process.

Figures 1.14a, b provide the displacement derivatives in the indicated directions; these derivatives show a spatial distribution that is very similar to the pattern corresponding to the solution of the theory of elasticity. The blue lines of equal values of the derivatives correspond to the area subjected to compression while the red lines correspond to the area subjected to tension. The yellow lines correspond to the transition from tensile to compressive stresses. These two regions have two extreme values almost equidistant from the transition line. It was also observed that at point "C" of Fig. 1.13a the lattice constant is very close to the value for the perfect crystal (a=0.3073 nm). This indicates the passage from compression to tension in the array, a neutral region. As it has been said before, this neutral region is present in the elasticity solution of an edge dislocation and separates the region where the extra planes are located (compression region) from the region that is expanded (region that does not contain extra planes), Fig. 1.14c.

The moiré method provides the Eulerian components of the strain tensor: that is, the strains at the location in the deformed configuration. Through the above analysis it is possible to have an idea of the magnitude of the deformation caused by dislocations generated during the growth stage of a 4HSiC crystal in the region where the cooling rates have altered the regularity of the crystalline array.
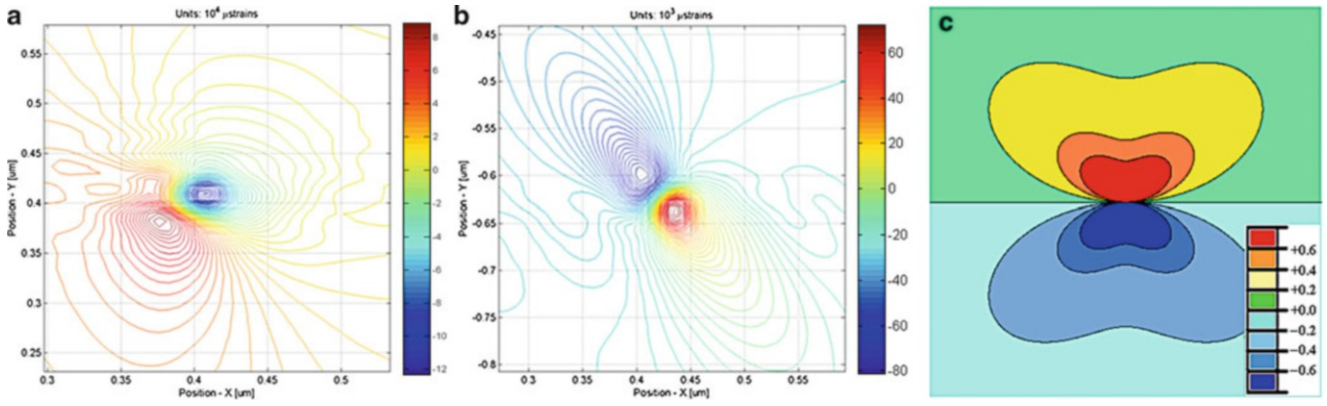
Fig. 1.14 (**a**) Displacement derivative $\partial ue_2/\partial e_2$ in the direction e_2; (**b**) Displacement derivative $\partial ue_3/\partial e_3$ in the direction e_3; (**c**) Theory of elasticity solution strains in the direction parallel to the line of separation between tension and compression [19]

1.8 Discussion and Conclusions

The subject of the dynamics at the level of the atomic structure is very complicated because it forms part of the many bodies interaction of physics. This is a very ancient problem that started in Greek times in the astronomy field and today forms part of quantum mechanics. In quantum systems, the repeated interactions between particles create quantum correlations. The wave function describing such a system includes an extraordinary amount of information creating a very complex system that makes its solution extremely difficult to tackle. In this paper part of the general problem is researched: the kinematics involving basic geometric relationships. In order to arrive to conclusions on the kinematics of atomic motions, the approach followed in the article was based on the experimental observation of dislocations on an HRTEM image pattern of a 4HSiC crystal. Image formation is itself a subject of great complexity and the image was analyzed by means of state-of-the-art image processing algorithms. A generalization of the Cauchy-Born rule was utilized in order to obtain a realistic description of the kinematic of atom motion.

A fundamental question to be answered is the connection between the phenomenological descriptions of the motion of solids by Continuum Mechanics and the basic geometrical changes at the level of crystalline organization. The answer to this question can be approached in many different ways. In this study, the basic approach was based on an ancient platonic principle: geometry is the underpinning of the observable world. By geometry it is meant the organization of matter at the crystalline level. The starting point is Fig. 1.5, that contains the arrangement of atoms as rigid spheres. This problem is known in the Geometry literature as the "cannon balls problem" and originated in the stacking of cannon balls. Atoms are arranged as spheres in geometrical patterns that represents the equilibrium condition.

In the next step, the FFT of the image of the atomic positions yields a hexagon in the frequency space that it is a statistical description of all the hexagonal cell arrangements in the observation plane. From the FFT it is possible to extract (Fig. 1.8) basic distortions of the fundamental cell that represent basic steps in the description of the deformation of the continuum. It is possible to see that the points of the distorted hexagon are enveloped by ellipses. It is well known the relationship of the ellipse geometry and the second order tensor describing deformation in the 2D continuum. Hence, it follows that one can connect the deformations of the elementary cell with the [J] tensor defined in Eq. (1.10). Then one can apply to the elementary cell the rosette analysis utilized in extensometry.

The next step is to verify if in the different regions of the image where compression or tension prevails the derived relationship may fit the experimentally observed distortions of the elementary cell. It has been established a bridge between the macro-world and the kinematic of atomic motions. The differential geometry description of the deformation of the continuum contains the mathematics required to explain the observed motions of atoms in 2D context within what we can call the continuum phase of deformations. However, what happens if there are discontinuities? The observation of the region containing the dislocations in the analyzed image provides some answers with respect to this eventuality. The regular arrangement of the elementary cells is disrupted. The row of atoms immediately after the extra rows end appear blurred in the image; the reason for this fact is quite difficult to explain. Figure 1.12b shows the model of rigid spheres with a edge dislocation: it is possible to see that a large cavity is formed extended about 3.5 atomic diameters. The image of Fig. 1.12a shows that this cavity exists but it is about one atomic diameter in depth and it is not an early crack but a region where still some electron density can be observed. Evidently, atoms are not rigid spheres but are deformable and their field of

influence changes shape and can become ellipsoids under the influence of stresses as depicted in Fig. 1.11. Under the influence of large distortions the atoms take configurations of equilibrium that are precursors of cavities. These regions can be detected experimentally through gamma rays spectroscopy.

Another important question is the following. How far we can extend the continuum model? Digital moiré provides some answers, fringe dislocations indicating of the discontinuity in the field of displacements appear at the location of the physical discontinuity. Figure 1.14 extends the field of deformations as measured by the displacement derivatives inside the atomic field, where the image intensity is related to electronic density. The obtained pattern shows an interesting result: the region where light intensity is maximum that corresponds to the location of the atomic nucleus has low deformations and deformations increase with the distance from the nucleus. Further work is required to establish the validity of these conclusions.

In conclusion, Continuum Mechanics is a more powerful tool than originally conceived. The reason for this fact lays in the geometrical basis of the organization of matter that underlines the observable behavior in the macro world. Although the derived conclusions have the narrow basis of a single image, careful observations of the geometrical changes seem to support the idea that the continuum can be modeled by interacting atoms and kinematics is represented by the distance vectors of these atoms. In this research work one more step was taken by extending interaction kinematics to the atomic space.

From the obtained results it can be concluded that HRTEM technology has arrived to a point that with the help of image analysis algorithms can provide information that allow observing the world at the sub-nanometric level controlling the behavior of matter at the macro world.

References

1. Cauchy AL. De la pression ou tension dans un système des points matériels. Sur le équilibre et le mouvement d'un système de points matériels sollicités par des forces d'attraction ou de répulsion mutuelle. Sur l'équilibre et le mouvement intérieur des corps considérés comme masses continues. In: Cauchy Augustin-Louis, Ouvres Complétés, Tomes 20–21. Gauthier-Villars et fils, Paris, France, 1882–1974
2. Born M (1915) Dynamik der kristallgitter. BiblioLife, USA, 2011
3. Volterra V (1907) Sur l'équilibre des corps élastiques multiplement connexes. Ann EC Norm 3(24):401–517
4. Somigliana C (1914,1915) Sulla teoria delle distorsioni elastiche. Rendiconti dell'Accademia dei Lincei Vol. 23, pp 463–472 (Note I, 1914); Vol 24, pp 655–666 (Note II, 1915)
5. Love AEH (2011) A treatise on the mathematical theory of elasticity, 4th edn. Dover, New York
6. Taylor GI (1934) Mechanism of plastic deformation of crystals: Part I - Theoretical, Part II - Comparison with observations. Proc. Royal Soc A145 (855):362–404
7. Orowan E (1934) Zur Kristallpastizität III. Z. Phys 89:634–659
8. Polanyi MZ (1934) Über eine art gitterstörung, die einen kristall plastisch machen könnte. Z. Phys 89:660–664
9. Nabarro FRN (1987) Theory of crystal dislocations. Dover, New York
10. Lardner RW (1974) Mathematical theory of dislocations and fracture. University of Toronto Press, Toronto, ON
11. Sciammarella CA, Lamberti L, Sciammarella FM (2006) Measurement of deformations at the sub-nanometric level. In: Proceedings of the NANOMEC06 symposium on materials science and materials mechanics at the nanoscale, 19–23 November 2006, Bari, Italy
12. Sciammarella F, Sciammarella CA, Lamberti L (2007) Processing of a HRTEM image pattern to analyze an edge dislocation. In: Gdoutos EE (ed) Experimental analysis of nano and engineering materials and structures. Springer, The Netherlands
13. Sciammarella F, Sciammarella CA, Lamberti L (2007) Experimental nanomechanics: a look at edge dislocations in 4HSiC crystals. In: Proceedings of the materials science and technology conference and exhibition (MS&T '07), Detroit, MI, September 2007
14. Sciammarella CA, Sciammarella FM (2012) Experimental mechanics of solids. Wiley, Chichester
15. Boreman GD (2001) Modulation transfer function in optical and electro-optical systems. SPIE Press, Bellingham, WA
16. Williams DR, Barry Carter C (2009) Transmission electron microscopy, a textbook for materials science, 2nd edn. Springer, New York
17. Ha SY, Nuhfer NT, De Graef M, Rohrer GS, Skowronski M (2000) Origin of threading dislocation arrays in SiC boules grown by PVT. In: Carter CH, Devaty RP, Rohrer GS (eds) Proceedings of the international conference on SiC and related materials. Materials science forum 338–342, Trans Tech Publications, Switzerland, pp 477–480
18. JEOL USA. Transmission electron microscopes. http://www.jeolusa.com/
19. Subramaniam A, Balani K (2013) Materials Science and Engineering (MSE). Chapter 5b: Crystal imperfections dislocations. Indian Institute of Technology, Kanpur, India. http://home.iitk.ac.in/~anandh/E-book/. Last accessed 05 Mar 2014
20. General Stress Optics Inc. Holo-Moiré Strain Analyzer Software HoloStrain, Version 2.0. General Stress Optics Inc., Chicago, IL. http://www.stressoptics.com
21. Sciammarella CA (1998) Positron annihilation spectroscopy for life assessment of super alloys. In: Viswathnathan R (ed) Proceedings for the NDE for damage assessment workshop, EPRI TR-110291
22. Sciammarella CA, Lamberti L (2014) Basic models supporting experimental mechanics of deformations, geometrical representations, connections among different techniques. To appear in Meccanica, Special Issue on "Experimental Solid Mechanics: in honour of Professor Emmanuel Gdoutos". Early View 8 January 2014. http://link.springer.com/article/10.1007%2Fs11012-013-9867-8

Chapter 2
Comprehensive Theory of Deformation

Sanichiro Yoshida

Abstract Linear elasticity is considered as a special case of the field theory of deformation and fracture capable of all stages of deformation based on the same theoretical basis. With the appropriate expression for the term representing the longitudinal force in the field equations, it has been shown that the equation of motion derived from the field equations reduces to the exact form of elastic wave equations known in the conventional continuum mechanics.

Keywords Field theory of deformation • Local symmetry • Comprehensive description of deformation and fracture • Continuum mechanics • Equation of motion of deformation

2.1 Introduction

Fracture of solid is initiated at the atomistic level. It starts with an irreversible atomic rearrangement such as dislocation, and develops to the final, macroscopic fracture through different stages of deformation occurring at different scale-levels. Dislocations exist in a material whose macroscopic behavior is elastic. Plastic and elastic deformations coexist near a residual stress. All these well-known facts indicate it necessary to develop a theory capable of describing the process of deformation on the same foundation across all the scale levels and different macroscopic regimes, i.e., the elastic, plastic and pre-fracture stages.

A recent field theory of deformation and fracture [1–3] has such capability. The basic idea behind this theory is that in any stage of deformation, the material locally obeys the law of linear elasticity. Alternatively, it can be said that at any stage of deformation, it is possible to find local regions where the law of linear elasticity is applicable. This is justified by the fact that the inter-atomic potential can be approximated by a quadratic function near the equilibrium regardless of the stage of deformation. This means that if we go down to the atomistic level, the displacement of an atom from its equilibrium position can be described by elastic force. We call such a local region where the law of linear elasticity is applicable as the deformation structural element. Within each deformation structural element, we can define a local frame that acts as principal axes to describe the local linear elasticity. The exception is when a chain of atoms has dislocation or other defects because with those defects the force acting on atoms is not proportional to the displacement from the atom's equilibrium position. In that case, the force law does not take the form of linear elasticity. Since the law of linear elasticity is applicable only within the deformation structural element, it is not possible to describe the law of elasticity for a given deformation structural element using the local frame of other deformation structural element. This indicates that the introduction of multiple deformation structural elements becomes necessary when dislocations and other defects exist. It also indicates that under the situation where the material has a number of defects and therefore the use of multiple deformation structural elements is necessary, the dynamics associated with the existence of those defects can be derived from the interrelationship among the multiple deformation structural elements.

This is where the concept of local invariance comes to the picture. According to the principle of local invariance, when each individual deformation structural element undergoes linear elastic deformation, they need to know elastic deformation occurring in other deformation structural elements. The situation is analogous to the dynamics of electrons orbiting

S. Yoshida (✉)
Department of Chemistry and Physics, Southeastern Louisiana University, SLU 10878, Hammond, LA 70402, USA
e-mail: syoshida@selu.edu; sanichiro.yoshida@selu.edu

around nuclei. While the dynamics of electrons orbiting around the same nucleus can be described by the corresponding Schrödinger equation, dynamics of electrons orbiting different nuclei cannot. However, those electrons still feel each other's motion through the electric force. This situation can be described as saying that the electromagnetic field is the gauge field to make quantum dynamics locally symmetric. The corresponding field equations are the well-known Maxwell's equation of electrodynamics.

In the dynamics of deformation, the field equations can be derived in the same fashion. The derivation of the field equations of deformation dynamics is reported in detail elsewhere [1–3]. From the physical meaning discussed in the preceding paragraph, the field equations describe the interrelationship among different deformation structural elements. It is important to note that the dynamics describing multiple deformation structural elements does not necessarily represent plasticity or irreversible deformation. Rather, it describes curvilinear elasticity. Plasticity or irreversible dynamics is introduced when a certain term in the field equation represents energy dissipation. This term behaves as a source of deformation. Following the convention of electrodynamics, the term is called the deformation charge [1]. A previous paper considers that the elastic, plastic and fracturing stage can be comprehensively described by using different forms of the charge. In the elastic regime, the charge represents elastic force proportional to the stretch, or in the case of a three-dimensional picture, elastic force proportional to volume expansion. Based on the formalism described herein, the field equations should be able to describe linear elastic deformation as the limit case where all deformation structural elements can be handled by a single coordinate system. Furthermore, in this limit, the theory should reduce to the conventional continuum mechanics. In ref. 3, we discussed this issue to some extent. The aim of this paper is to prove that the field equations reduce to continuum mechanical equations and discuss the physical meaning of related issues.

2.2   Field Equation and Equation of Motion

The field equations of deformation dynamics can be given in the following form [1].

$$\nabla \cdot \boldsymbol{v} = j_0 \tag{2.1}$$

$$\nabla \times \boldsymbol{v} = \frac{\partial \boldsymbol{\omega}}{\partial t} \tag{2.2}$$

$$\nabla \times \boldsymbol{\omega} = -\frac{1}{c^2} \frac{\partial \boldsymbol{v}}{\partial t} - \boldsymbol{j} \tag{2.3}$$

$$\nabla \cdot \boldsymbol{\omega} = 0 \tag{2.4}$$

Here $\boldsymbol{v}$ is the rate of the displacement, $\boldsymbol{\omega}$ is the rotation, c is the phase velocity, j_0 and $\boldsymbol{j}$ are the temporal and spatial component of charge of symmetry [4]. As clear from the form, the field equations yield wave equations that describe the spatiotemporal behavior of points in a material under deformation.

Rearranging the terms of Eq. (2.3), the equation of motion can be given as follows.

$$\frac{1}{c^2} \frac{\partial \boldsymbol{v}}{\partial t} = -\nabla \times \boldsymbol{\omega} - \boldsymbol{j} \tag{2.5}$$

In wave dynamics, the phase is the parameter that determines the state of the physical quantity of interest as a function of time and space. The phase velocity connects the time dependence and space dependence. Using f as representing the spatiotemporal function of the physical quantity of interest φ, the wave solution can be written in the following form.

$$\varphi = f(ct \pm r)$$

The phase velocity "translates" the time dependence to space dependence and vice versa.

$$\frac{\partial \varphi}{\partial t} = cf', \quad \frac{\partial \varphi}{\partial r} = f'$$

So, in a one dimensional case,

$$\frac{\partial}{\partial t} = c\frac{\partial}{\partial x} \tag{2.6}$$

In a three dimensional case, this can be extended to

$$\frac{\partial}{\partial t} = c\nabla \tag{2.7}$$

In a mechanical system in general, the phase velocity is the square root of the ratio of the elastic constant to the density. Using κ to denote the general elastic constant, we can rewrite Eq. (2.5) as follows.

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} = -\kappa\nabla \times \boldsymbol{\omega} - \kappa\boldsymbol{j} \tag{2.8}$$

Here the term $\kappa\boldsymbol{j}$ represents the longitudinal effect on the change in velocity, as opposed to $\nabla \times \boldsymbol{\omega}$ represents the transverse effect via the curl operation. As clear from the dimension of the terms, this is in fact force density representing the longitudinal effect. The elastic and plastic regime can be commonly described by the form of Eq. (2.8) with the use of a different expression for $\kappa\boldsymbol{j}$.

2.3   Elastic Deformation

In the elastic regime, $\kappa\boldsymbol{j}$ represents longitudinal, elastic (restoring) force. Below, we will examine it for different degrees of freedom.

2.3.1   Three-Dimensional Compressive Elastic Wave

Take divergence of Eq. (2.8).

$$\rho\frac{\partial (\nabla \cdot \boldsymbol{v})}{\partial t} = -\kappa\nabla \cdot (\nabla \times \boldsymbol{\omega}) - \kappa\nabla \cdot \boldsymbol{j} = -\nabla \cdot (\kappa\boldsymbol{j}) \tag{2.9}$$

In Going through the last equal sign, the mathematical identity $\nabla \cdot (\nabla \times \boldsymbol{\omega})$ (divergence of curl is zero) is used. Replacing velocity $\boldsymbol{v}$ with displacement $\boldsymbol{\zeta}$, Eq. (2.8) can further be rewritten as follows.

$$\rho\frac{\partial^2 (\nabla \cdot \boldsymbol{\zeta})}{\partial t^2} = -\nabla \cdot (\kappa\boldsymbol{j}) \tag{2.10}$$

Using Eqs. (2.7), (2.10) can be put as follows.

$$\rho c^2 \nabla \cdot \nabla(\nabla \cdot \boldsymbol{\zeta}) = -\nabla \cdot (\kappa\boldsymbol{j}) \tag{2.11}$$

Comparing both hand-sides of Eq. (2.11), we can express $\kappa \boldsymbol{j}$ as

$$\kappa \boldsymbol{j} = -\rho c^2 \nabla (\nabla \cdot \boldsymbol{\zeta}) \tag{2.12}$$

From continuum mechanics, we know that the phase velocity of a compressive wave is given by

$$c^2 = \frac{\lambda + 2G}{\rho} \tag{2.13}$$

Substituting Eq. (2.13) into Eq. (2.12), we obtain

$$\kappa \boldsymbol{j} = -(\lambda + 2G)\nabla (\nabla \cdot \boldsymbol{\zeta}) \tag{2.14}$$

Substitution of Eq. (2.14) into Eq. (2.10) leads to

$$\rho \frac{\partial^2 (\nabla \cdot \boldsymbol{\zeta})}{\partial t^2} = (\lambda + 2G)\nabla \cdot \nabla (\nabla \cdot \boldsymbol{\zeta}) \tag{2.15}$$

Using $\nabla \cdot \nabla (\nabla \cdot \boldsymbol{\zeta}) = \nabla^2 (\nabla \cdot \boldsymbol{\zeta})$, we can rewrite Eq. (2.15) as

$$\rho \frac{\partial^2 (\nabla \cdot \boldsymbol{\zeta})}{\partial t^2} = (\lambda + 2G)\nabla^2 (\nabla \cdot \boldsymbol{\zeta}) \tag{2.16}$$

Equation (2.16) is the well-known equation of elastic compressive (volume) wave.
   Alternatively, substitution of Eq. (2.14) into Eq. (2.8) leads to

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} = -\kappa \nabla \times \boldsymbol{\omega} + (\lambda + 2G)\nabla (\nabla \cdot \boldsymbol{\zeta}) \tag{2.17}$$

In this context, the elastic constant appearing in the first term on the right hand-side is shear modulus as it is the coefficient multiplied to the curl of rotation, which represents shear deformation.

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} = -G \nabla \times \boldsymbol{\omega} + (\lambda + 2G)\nabla (\nabla \cdot \boldsymbol{\zeta}) \tag{2.18}$$

Using field Eq. (2.2) and the mathematical identity $\nabla \times \nabla \times \boldsymbol{\zeta} = \nabla (\nabla \cdot \boldsymbol{\zeta}) - \nabla^2 \boldsymbol{\zeta}$ we can put Eq. (2.18) as

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} = G \nabla^2 \boldsymbol{\zeta} + (\lambda + G)\nabla (\nabla \cdot \boldsymbol{\zeta}) \tag{2.19}$$

2.3.2   One-Dimensional Linear Elastic

In this case, the phase velocity is given with the Young's modulus as

$$c^2 = \frac{E}{\rho} \tag{2.20}$$

Take curl of the second field Eq. (2.2) and temporal derivative of the third field Eq. (2.3) to obtain the following equations.

$$\nabla \times (\nabla \times \boldsymbol{v}) = \frac{\partial (\nabla \times \boldsymbol{\omega})}{\partial t} \tag{2.21}$$

$$\frac{\partial(\nabla \times \boldsymbol{\omega})}{\partial t} = -\frac{1}{c^2}\frac{\partial^2 \boldsymbol{v}}{\partial t^2} - \frac{\partial \boldsymbol{j}}{\partial t} \tag{2.22}$$

With the use of the mathematical identity $\nabla \times (\nabla \times \boldsymbol{v}) = \nabla(\nabla \cdot \boldsymbol{v}) - \nabla^2\boldsymbol{v}$, this leads to

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{v}}{\partial t^2} = \nabla^2\boldsymbol{v} - \frac{\partial \boldsymbol{j}}{\partial t} - \nabla(\nabla \cdot \boldsymbol{v}) \tag{2.23}$$

Integrate Eq. (2.23) with respect to time and rewrite it as

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{\zeta}}{\partial t^2} = \nabla^2\boldsymbol{\zeta} - \boldsymbol{j} - \nabla(\nabla \cdot \boldsymbol{\zeta}) \tag{2.24}$$

Substitute the above phase velocity (2.20) into Eq. (2.24) to obtain

$$\rho\frac{\partial^2 \boldsymbol{\zeta}}{\partial t^2} = E\nabla^2\boldsymbol{\zeta} - E[\boldsymbol{j} + \nabla(\nabla \cdot \boldsymbol{\zeta})] \tag{2.25}$$

Notice that if we put

$$\boldsymbol{j} = -\nabla(\nabla \cdot \boldsymbol{\zeta}) \tag{2.26}$$

Equation (2.25) becomes the equation of a longitudinal elastic wave as known well in continuum mechanics.

$$\rho\frac{\partial^2 \boldsymbol{\zeta}}{\partial t^2} = E\nabla^2\boldsymbol{\zeta} \tag{2.27}$$

The physical meaning of the right-hand side of Eq. (2.26) is volume expansion. In the context of wave in a rod, this is the one-dimensional volume expansion. If we chose the coordinate so that the wave travels along the z-axis,

$$\rho\frac{\partial^2 \boldsymbol{\zeta}}{\partial t^2} = E\frac{\partial^2 \boldsymbol{\zeta}}{\partial z^2} \tag{2.28}$$

Equation (2.26) is the condition of $\boldsymbol{j}$ to represent a one-dimensional wave (a wave in a rod).

2.3.3 Rotation Wave

Take temporal derivative of the second field Eq. (2.2) and curl of the third field Eq. (2.3) to obtain the following equations.

$$\frac{\partial}{\partial t}(\nabla \times \boldsymbol{v}) = \frac{\partial^2 \boldsymbol{\omega}}{\partial t^2} \tag{2.29}$$

$$\nabla \times (\nabla \times \boldsymbol{\omega}) = -\frac{1}{c^2}\frac{\partial(\nabla \times \boldsymbol{v})}{\partial t} - \frac{\partial}{\partial t}(\nabla \times \boldsymbol{j}) \tag{2.30}$$

Thus, we can eliminate $\boldsymbol{v}$ to obtain

$$\frac{\partial^2 \boldsymbol{\omega}}{\partial t^2} = -\nabla \times (\nabla \times \boldsymbol{\omega}) - \frac{\partial}{\partial t}(\nabla \times \boldsymbol{j}) \tag{2.31}$$

Using field Eq. (2.4) and mathematical identity $\nabla \times (\nabla \times \boldsymbol{\omega}) = \nabla(\nabla \cdot \boldsymbol{\omega}) - \nabla^2 \boldsymbol{\omega}$

$$\frac{\partial^2 \boldsymbol{\omega}}{\partial t^2} = c^2 \left[\nabla^2 \boldsymbol{\omega} - \frac{\partial}{\partial t}(\nabla \times \boldsymbol{j}) \right] \tag{2.32}$$

In this case, apparently the elastic constant is the shear modulus.

$$c^2 = \frac{G}{\rho} \tag{2.33}$$

In addition, in the elastic case $\boldsymbol{j}$ is proportional to $\nabla(\nabla \cdot \boldsymbol{\zeta})$. From the mathematical identity "curl of gradient is zero", the second term of Eq. (2.32) is zero. So,

$$\rho \frac{\partial^2 \boldsymbol{\omega}}{\partial t^2} = G \nabla^2 \boldsymbol{\omega} \tag{2.34}$$

Equation (2.34) is known as the equation of rotational (deformation) wave in continuum mechanics.

2.4 Summary

Previously derived field equations that govern spatiotemporal behavior of the displacement field of solid material under deformation are discussed. Specifically, linear elastic deformation is considered as a limiting case of the formalism. The discussion indicates that the linear elasticity is characterized as the case where the longitudinal effect is given as a gradient of the displacement field. Physically, this means that in the elastic law of dynamics the longitudinal force is proportional to the gradient of displacement. Moreover, the longitudinal effect is rotation-less as the mathematical identity guarantees that curl of gradient is zero. With this condition, the equation of motion derived from the present field theory reduces to the exact form of the wave equation whose solution is known in the conventional continuum mechanics as the compressive wave or rotation wave.

References

1. Yoshida S (2010) Physical meaning of physical-mesomechanical formulation of deformation and fracture. AIP Conf Proc 1301:146–155
2. Yoshida S (2011) Scale-independent approach to deformation and fracture of solid-state materials. J Strain Anal 46:380–388
3. Yoshida S (2013) Scale independent approach to deformation and fracture. In: Imaging methods for novel materials and challenging applications, Conference proceedings of the Society for Experimental Mechanics series, vol 3. pp 1–7
4. Aitchson IJR, Hey AJG (1989) Gauge theories in particle physics. IOP Publishing, Briston

Chapter 3
Topography of Rough Dielectric Surfaces Utilizing Evanescent Illumination

F.M. Sciammarella, C.A. Sciammarella, and L. Lamberti

Abstract The authors utilize optical evanescent fields to analyze the topography of metallic and non metallic surfaces. The methodology initiated with the phenomenon of planar surface waves produced by surface plasmon polaritons. By direct experimental observations in 2009 the method was extended to ceramic surfaces in the micron and sub-micron range. Since the ceramics are dielectric materials the plasmon polariton model cannot explain the observed phenomena. For almost a century researchers have analyzed surface electromagnetic waves observed in planar interfaces that involve metallic surfaces, or metallic surfaces and dielectric media. These studies resulted in the theory of surface-plasmon waves and surface-plasmon-polariton waves. Additional planar surface waves are the so called Dyakonov waves, Tamm waves, and Dyakonov–Tamm waves. These waves were originally theoretically derived by M.I. Dyakonov about 25 years ago and were observed for the first time in 2009. The Dyakonov–Tamm waves are generated in the interface of two dielectric materials with periodic internal structures.

Keywords Evanescent illumination • Surface electromagnetic waves • Surface topography of metallic and ceramic materials

3.1 Introduction

Atoms inside a medium are connected to other atoms through inter-atomic potentials that imply sharing electronic orbits. When atoms are at a surface they lack their counterparts and special electronic band configurations arise and create surface states. The possible surface states depend on the atomic structure of the material: metals, semi-conductors, and dielectric materials give rise to different surface states. The Russian physicist I. Tamm analyzed surface states originating at the boundary of crystals and vacuum through solutions of the Schrödinger equation [1]. If different materials are in close contact, more complex states can be experimentally observed. In 1988, the Russian physicist M.I. Dyakonov [2] theoretically predicted the existence of surface waves for specified boundary conditions. Surface states that can produce surface waves have generated a great deal of scientific interest and for possible technical applications. Lahktakia and Polo [3] extended the concept of the surface waves defining the Dyakonov-Tamm waves. They analyzed the interface of a homogeneous dielectric material and a dielectric chiral thin sculptured film that is anisotropic and periodically inhomogeneous. While the classical Dyakonov waves have a very limited spectrum of angle propagation in the second medium, the

F.M. Sciammarella (✉)
Department of Mechanical Engineering, College of Engineering and Engineering Technology,
Northern Illinois University, 590 Garden Road, 60115 DeKalb, IL, USA
e-mail: sciammarella@niu.edu

C.A. Sciammarella
Department of Mechanical Engineering, College of Engineering and Engineering Technology,
Northern Illinois University, 590 Garden Road, 60115 DeKalb, IL, USA

Department of Mechanical, Materials and Aerospace Engineering, Illinois Institute of Technology,
10 SW 32nd Street, 60616 Chicago, IL, USA

L. Lamberti
Dipartimento Meccanica, Matematica e Management, Politecnico di Bari, Viale Japigia 182, 70126 Bari, Italy

Dyakonov-Tamm waves have a much larger angular spectrum of waves. A number of publications have resulted from theoretical analysis of the Dyakonov-Tamm waves [4–8].

In the field of optics, planar waves propagating in the interface of two media have been predicted and observed for a long time under the denomination evanescent waves. The authors have analyzed surface waves as evanescent waves [9–21] and took advantage of the properties of these waves in a number of scientific and technical applications.

3.2 Evanescent Wave's Solution of Maxwell Equations

Let us consider the Maxwell equations for planar propagating waves, the vectorial Helmholtz equation. By taking an exponential function as a solution of the Maxwell equations of the form $\exp(i\omega t)$ one gets the following vector equation:

$$\left(\nabla^2 + \omega^2 \mu_0 \varepsilon_0\right) \left\{ \begin{array}{c} \vec{E} \\ \vec{B} \end{array} \right\} = 0 \tag{3.1}$$

The electric and the magnetic field vectors $\vec{E}$ and $\vec{B}$ satisfy Eq. (3.1). The notation ∇^2 indicates the Laplace operator, $\omega = 2\pi f = 2\pi/T$ is the temporal angular period, ε_o and μ_o are the permittivity and permeability of vacuum, respectively. By introducing the argument $\exp\left(\vec{k} \bullet \vec{r} - i\omega t\right)$, the periodicity in space is included in the wave solution. The quantity k is the wave number $2\pi/\lambda$. For a plane wave, the position vector $\vec{r}$ is characterized by the equation $\vec{r} = r\left(\alpha \vec{i} + \beta \vec{j} + \gamma \vec{k}\right)$, where α, β and γ are the direction cosines of the considered vector. A general solution of Eq. (3.1) is possible by assuming that α, β and γ are complex quantities that have real and imaginary components. By defining:

$$\begin{cases} \alpha = a_1 + ia_2 \\ \beta = b_1 + ib_2 \\ \gamma = c_1 + ic_2 \end{cases} \tag{3.2}$$

the dotted $\vec{k} \bullet \vec{r}$ product gives:

$$\vec{k} \bullet \vec{r} = k[-(a_2 x + b_2 y + c_2 z) + i(a_1 x + b_1 y + c_1 z)] \tag{3.3}$$

Let us apply the above derivations to the problem of total reflection (Fig. 3.1). For simplicity, the analysis is reduced to a two dimensional problem by analyzing the process in the incidence plane of a propagating beam from medium 1 to medium 2 such that $n_1 > n_2$, where n_1 and n_2 are the indices of refraction of the two media. Considering the total internal reflection condition, introducing the refraction equation and calling θ_i the angle of incidence and θ_t the angle of the transmitted light one arrives to the following conclusions:

1. There is a surface wave that propagates in the X-direction as a plane wave.
2. The planes of constant phase and constant amplitude are orthogonal to each other.
3. The electromagnetic field penetrates the second medium but decays very rapidly in the Z-direction.
4. If the electric vector is contained in the plane of incidence, the evanescent field electric vector becomes elliptically polarized in the plane of incidence (p-polarization).
5. If the electric vector is orthogonal to the plane of incidence (s-polarization), the polarization does not experience changes.
6. If one computes the Poynting vector of the field on the upper half of the plane, one concludes that no energy is transmitted to the second medium.
7. Computing the modulus of the component k_z of the $\vec{k}$ vector (see Fig. 3.1b) one arrives to the conclusion that this component oscillates with spatial frequencies $2\pi/\lambda_{ez}$ much higher than the spatial oscillation of all ordinary electromagnetic waves, λ_{ez} is the wavelength of a super oscillatory field. Berry [22] showed that such a field is a super-oscillatory field that through scattering in a medium can be transformed into an actual near field. This approach leads to the near-field microscopic detection that is well documented in the literature.
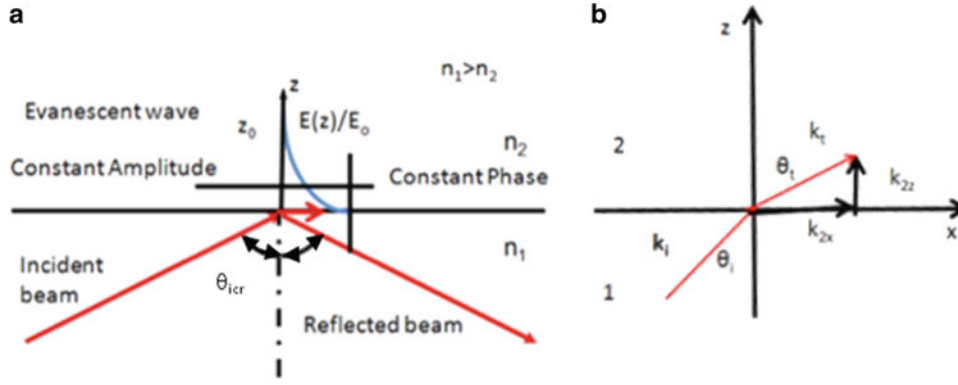
Fig. 3.1 Characteristics of evanescent field: (**a**) Orthogonality of the planes of constant phase and constant amplitude and exponential decay of the vector amplitude in the Z-direction (**b**) Resultant $\mathbf{k_t}$ vector penetrating in the second medium

3.3 Improvement of the Optical Resolution of an Optical Instrument via Evanescent Illumination

Scanning probe devices are currently utilized to collect local information on the electromagnetic field lying near a surface of different geometrical configuration. Over the past decade, optical experiments based on near field observations have been developed via scanning probe devices. These experiments demonstrated the actual realization of optical resolutions well beyond the classical $\lambda/2$ "Rayleigh limit". In this work, the analysis is on how the electromagnetic field interacts with nano-sized objects producing propagating light waves capable of carrying information generated in the near field up to the far field where this information is retrieved. This phenomenon involves in the near field what is called "confined electromagnetic fields" since the near field electromagnetic fields exists in nano-objects that are smaller than the wavelength of the illuminating light.

Reference [23] addresses the relationship between the field generated by the evanescent illumination in the case of an array of objects similar to the one depicted in Fig. 3.2a, a prismatic nano-crystal lying on a microscopic slide. Figure 3.2b shows the field resulting from transverse-magnetic TM polarization for such an array, the k_x component is a super-oscillatory component. The objects represented in Fig. 3.2b are 25 nm tall and the field is represented at 30 nm from the plane X.Y the field resembles the geometry of the objects. However, the transverse-electric TE polarization provides patterns that indicate the edges of the protrusions. The results given in [23] indicate the increasing resolution of the field in following the shape of the objects, as the objects become sub-wavelength in dimensions and this field has its source of energy in the electro-magnetic field created by the evanescent waves.

Having summarized the basic properties of the evanescent waves, and concluded from [22] that the presence of a dielectric object in the evanescent field produces a scattering field that when the dimensions of the object are sub-wavelength [23] resembles the object's geometry, it is necessary to look further to the relationship existing between evanescent waves and super-resolution. The presence of super-oscillations in the near field was postulated by Toraldo di Francia [24, 25]. A formal approach relating evanescent fields and super-resolution is due to Vigoureux [26]. It is interesting to look back to this argument for the implications that it has in the retrieval of information from the geometry of the near field. The starting point of the argument is the uncertainty principle of FT. From the uncertainty principle of FT [27] applied in the x-direction, and applied to the pair of complementary quantities position and wave vector component, Δx and Δk_x respectively, Vigoreux shows that the uncertainty principle leads to the inequality:

$$\Delta x \Delta k_x > 2\pi \tag{3.4}$$

From the above inequality, one can conclude that to obtain the smallest spatial resolution Δx the interval Δk_x must be as large as possible. The classical solution of Maxwell equation for propagating waves in vacuum (approximately in air) provides a limit. This limit can be computed by observing that the limit imposed by the condition of total reflection from a material of index of refraction n with respect to vacuum gives values of $k_x = \omega/c$, c the speed of light, $+ \omega/c = -2\pi/\lambda, +2\pi/\lambda$, resulting in a $\Delta k_x = 4\pi/\lambda$. Replacing this value in Eq. (3.4) one obtains $\Delta x = \lambda/2$, which brings the Rayleigh limit.
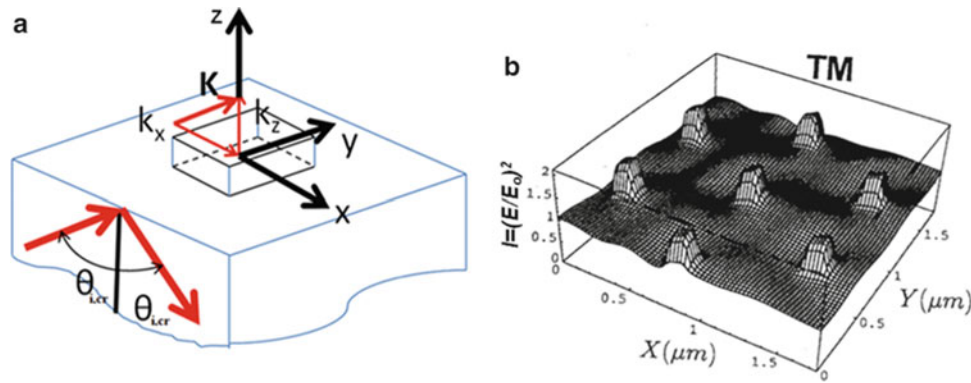
Fig. 3.2 (**a**) Example of one of the observed objects in [15], a prismatic nano-crystal lying on a microscope slide. The **k** vector of the propagating wave front and its components are indicated. Object dimensions are sub-wavelength, (order of magnitude close to λ/10); (**b**) Electric field intensity lines in the neighborhood of array of objects similar to that depicted in (**a**)

If we now consider evanescent waves as sources of electromagnetic energy capable to excite a medium and create super-oscillating fields, the values Δk_x are no longer limited to the above mentioned interval. When the diffracted waves corresponding to the plane wave fronts with real direction cosines arrive to the limit λ/2, the solutions of the evanescent plane wave fronts with complex direction cosines add new frequencies to the spectrum. All the above derivations have been applied to the observation of nano-objects [9–11, 13, 15, 16, 18–21].

3.4 Application of Surface Waves to the Study of Surface Roughness Topography

The same basic principles that were just outlined can be applied to the study of the topography of surfaces and deflection of surfaces. As previously stated, when a plane wave front impinges the surface separating two media such that the index of refraction of medium 1 (e.g. glass) is higher than the index of refraction of medium 2 (e.g. air), at the limit angle total reflection takes place (see Fig. 3.1). Under these circumstances a very interesting phenomenon occurs. At the interface (glass-air) evanescent waves are produced. At the same time, scattered waves emanate from the medium 1 (glass). If a third medium, a conducting material, for example a metal like copper, is close enough to interact with this field the energy from this field interacts with the metal's surface free electrons to generate plasmons (dark orange area in Fig. 3.3) that by decaying cause the metal surface to emit light.

Between the copper and the glass surface there is an optical cavity or optical resonator that produces standing waves. The electromagnetic field confined in the cavity is reflected multiple times inside the cavity producing standing waves for certain resonance frequencies that depend on the geometry of cavity. The standing wave patterns generated are called modes: each mode is characterized by a frequency f_n, where the subscript n is an integer. Optical cavities are characterized by the quality factor, or Q factor. The Q factor is a dimensionless parameter that defines the resonator's bandwidth relatively to its center frequency. Low values of Q indicate high losses of energy in the cavity and a wide bandwidth. High values of Q indicate a low rate of energy loss in the cavity with respect to the stored energy of the oscillator and a narrow bandwidth. In general, Q is defined in terms of the ratio of the energy stored in the resonator to the energy being lost in one cycle:

$$Q = 2\pi \frac{\text{Energy Stored}}{\text{Energy loss per cycle}} \qquad (3.5)$$

It can be proven that the maximum energy is concentrated in the direction of the normal to the face of the glass medium.

Fig. 3.3 Generation of the evanescent field and surface plasmon resonance in the cavity between a glass plate and a rough copper surface
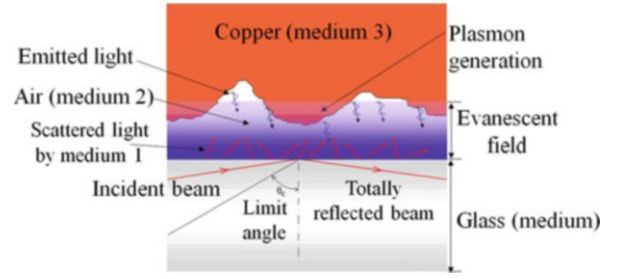
3.5 Generation of a Wide Spectrum of Frequencies Due to Interaction Between Evanescent Waves and Rough Surfaces

In the classical Fabry-Perot cavity analysis the two interacting surfaces are mirror-like surfaces. In the present case one of the surfaces is mirror-like (the glass) while the other (the metal) is an optically rough surface. Therefore, many different spatial frequencies can be observed experimentally. At this point it is important to describe the phenomenon leading to the generation of the emitted light with different spatial frequencies on rough metallic surfaces. A rough surface can be thought of as the superposition of many gratings of different periodicities. Kretschmann analyzed this problem in the following fashion [28]. A rough surface can be defined through the following statistical correlation function:

$$G(x, y) = \frac{1}{A} \int_A z\left(x', y'\right) z\left(x' - x, y' - y\right) dx' dy' \tag{3.6}$$

where $z(x,y)$ is the Monge's representation of the surface height and A is the area of integration. Under the assumption that the height distribution is a random function (this is customary for the analysis of random surfaces), a Gaussian distribution can be utilized. The correlation function becomes:

$$G(x, y) = R_q^2 \exp\left(-\frac{r^2}{\sigma_i^2}\right) \tag{3.7}$$

where: R_q^2 is the root-mean-square value of the surface heights assumed to be random variables with a correlation length σ_i; r is the distance from the generic point $P(x,y)$ of the object surface. From the Fourier transform (FT) of Eq. (3.7) the spectrum s of spatial frequencies present in the surface can be obtained.

From the point of view of plasmon excitation, one can prove that, in order to excite a plasmon resonance, it is necessary that the exciting frequency coincides with a frequency in the Fermi's electromagnetic state. Hence, the larger is the spectrum of the frequencies the greater will be the amount of energy available for producing coupling of plasmons within the metallic surface. The spatial frequency spectrum is described by the following equation:

$$s\left(\|k\|_{surf}\right) = \frac{1}{4\pi} \sigma_i^2 R_q^2 \exp\left[-\frac{\sigma_i^2 \|k\|_{surf}^2}{4}\right] \tag{3.8}$$

The above equation shows that the spectrum of light emitted by the surface consists of multiple wave vectors that are related directly with surface topography properties. Each wave vector corresponds to an equivalent pitch p_{gr} defined as:

$$\|k\|_{surf} = \frac{2\pi}{p_{gr}}. \tag{3.9}$$

If the surface has only one Fourier component of roughness (i.e. the surface profile is sinusoidal), then the s function is discrete and exists only at $\|k\|_{surf} = 2\pi/p_{gr}$ where p_{gr} is the pitch of the equivalent sinusoidal grating. However, most surfaces of practical interest have a definite structure and cannot be considered random surfaces. Surfaces of technical interests

manufactured industrially present a periodic structure. For this reason finished surfaces are more similar to a deterministic diffraction grating than to a random grating.

Any plasmon propagating on a rough surface with the appropriate $\mathbf{k}_{surf}$ can generate the emission of a photon [29]. Since $\mathbf{k}_{surf}$ can be a random quantity, even if the light has a defined direction it is possible to generate plasmons in all directions. This phenomenon was verified experimentally by Teng and Stern [30].

3.6   Analysis of Topography of Rough Surfaces

Section 3.4 discussed the interaction of evanescent waves within the glass-air-metal interfaces. The material presented provides insight into how it is possible to make an experimental observation of a surface's topography based on the ability of evanescent illumination interacting with the surface to create plasmons that upon decay can emit light. Since the experimental setup has a double interface (glass-air and air-metal) a cavity is created: this cavity acts as a passive optical resonator and hence resonances can be observed from the emitted light. Figure 3.4a illustrates the scattering (diffraction effect) of the evanescent field entering in a dielectric medium. The process that provides information in the Z-direction of the events taken place in the X-Y plane can be summarized as follows. It is necessary to have a mechanism that transforms the changes taking place in the X-Y plane into observable quantities in the Z-direction. The process of light scattering [31] achieves this objective. The observer sees the projection of the electric vector $\vec{E}$ of the propagating electromagnetic field in the direction perpendicular to the observation vector.

The original experimental set up utilized to perform surface topography measurements on metallic specimens [17] is shown in Fig. 3.5. An incident laser beam is totally reflected at the glass-air interface and returned by the mirror shown
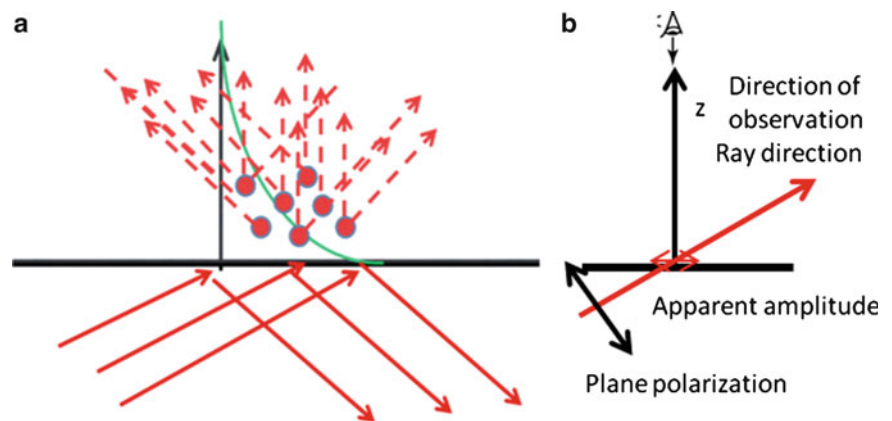
Fig. 3.4 (**a**) Scattered light produced by the evanescent illumination; (**b**) Observation takes place in the direction perpendicular to plane separating the two media
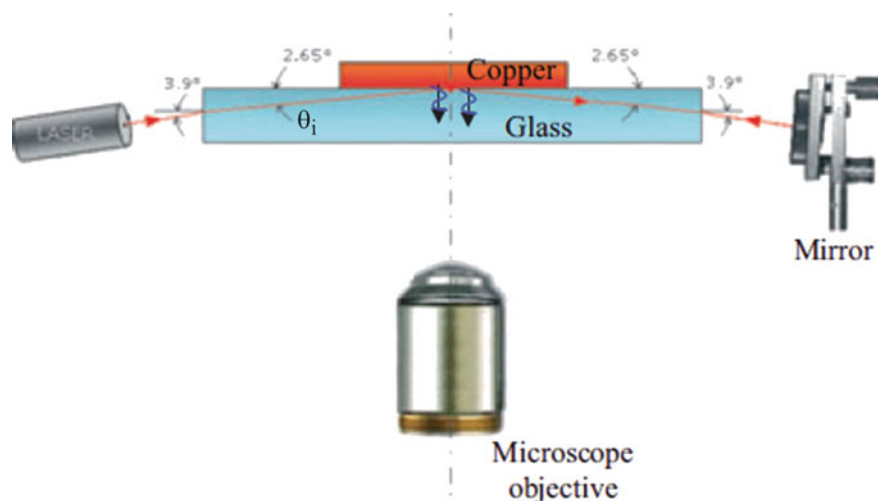
Fig. 3.5 Experimental setup for the surface plasmon resonance analysis of metallic specimens
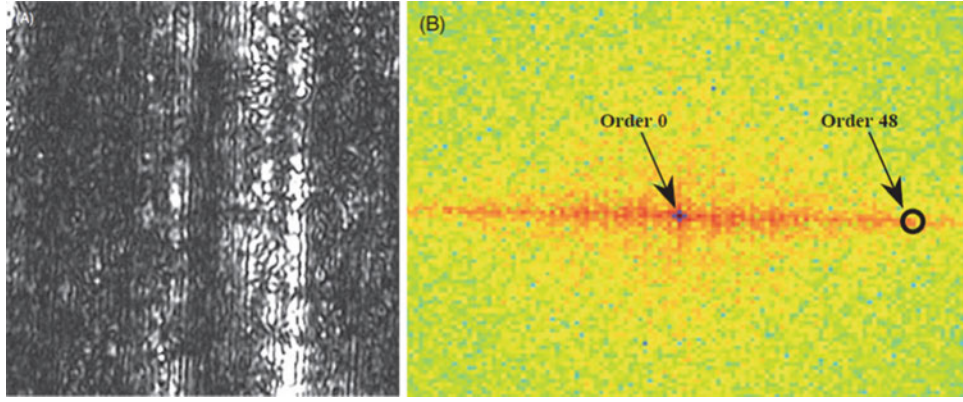
Fig. 3.6 (**a**) Enlarged portion of the image of the copper plate surface; interference fringes on the metal surface are clearly visible; (**b**) FT of the image (225 × 225 μm)
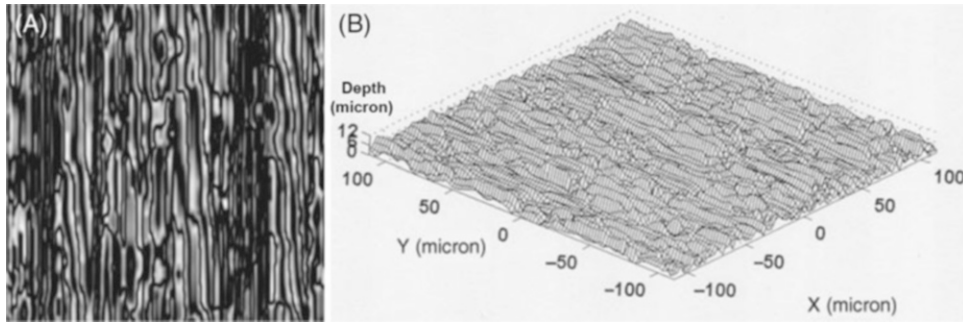
Fig. 3.7 (**a**) *Top view* of the depth information in gray levels of a region of 225 × 225 μm; (**b**) 3D view of the same region (the same scale for all the three coordinates was taken)

in the figure. The metal-air-glass interface is observed with an inverted optical microscope and an image is recorded by a CCD camera.

Figure 3.6a shows the recorded image of a copper surface. A system of fringes can be seen in this image. This system of fringes is localized on the surface since the focused image captures the surface features as well as the fringes. The spatial frequency of the fringes (pitch) can be determined by analyzing the FT of the image (see Fig. 3.6b). In the FT, it is possible to identify the fundamental harmonic corresponding to the fringe's image. From the analysis of the fringes it was concluded that the fundamental harmonic has a pitch of 4.69 μm.

The wave vector of the emitted field corresponds to a spatial frequency of order 48, that translated to fringe spacing in the physical space, gives a value of 4.69 μm (i.e. from the size of the region of 225 μm, one obtains 225/48 = 4.69 μm). In [17], it is shown that the diffracted order of the equivalent grating resulting from the electromagnetic resonance is:

$$p_{egr} = \frac{\lambda}{2\sin\theta_i} \tag{3.10}$$

where θ_i is the angle made by the illumination beam with the glass surface (see Fig. 3.5) and the factor 2 is due to the double illumination introduced by the mirror (see Chap. 18.10.1 of [32]). By entering $p_{egr} = 4.69$ μm into Eq. (3.10), a value of $\theta_i = 3.87°$ is obtained. This result is in very good agreement with the initial experimentally set value of 3.9°.

Figure 3.7b shows the surface profile obtained from the fringe pattern of Fig. 3.7a. The profile of the surface in terms of the surface depth h can be obtained from:

$$h = \frac{np_{egr}}{2\sin(90° - \theta_i)} \tag{3.11}$$

where n is the fringe order corresponding to the fact that an equivalent grating of pitch p_{egr} is modulated by the surface. This relationship is similar to the equation of shadow-projection moiré except for the factor 2 that comes from the double illumination effect. Equation (3.11) corresponds to Eq. (26) of [17] if the cosine of θ_i is replaced by the sine of the complementary angle. In the next section the physical meaning of this angle will be further clarified.

The derivation of Eq. (3.11) corresponds to the case of rough surfaces where the light is diffused by the surface. However, the method could also be applied to mirror surfaces where light is reflected in place of being diffused. Ebbeni dealt with the extension of the equivalent to the shadow moiré applied to reflecting surfaces in place of diffusing surfaces [33] by utilizing the Fresnel-Kirchhoff integral for the grating-based contouring of very large curvature radii reflecting surfaces. Sciammarella and Combel [34] also analyzed this problem. While in [33] only the first harmonic was considered, in [34], following the method introduced by Guild [35], all the harmonics that emerge with the same angle of inclination are accounted for thus obtaining multiple interference fringes. Hence, as shown in [32], Eq. (3.11) will apply to mirror surfaces. In this case the factor 2 will be due to the reflection of light and not to the double illumination corresponding to the optical circuit of Fig. 3.5.

In order to verify the reliability of the experimental data gathered with the above described model, other independent measurements were carried out with a Surtronic S25 profilometer (Taylor Hobson Ltd.). The ten-point height of the surface S_z was chosen as the reference parameter [36, 37]. Ten transversal sections (i.e. in the direction orthogonal to grooves) and ten longitudinal sections (i.e. in the direction of the furrows present in the surface) were analyzed and the corresponding ten-point heights were computed. For the profilometer measurements, two sections of length 8 mm were analyzed both in the longitudinal (i.e. parallel to the furrows) and transversal direction (i.e., perpendicular to the furrows). Optical and mechanical measurements were in good agreement. The test of significance for small samples revealed that both optical and mechanical data belong to the same statistical population within 95 % level of confidence.

In the second set of experimental tests, gratings of different pitches were utilized in order to obtain the surface topography of standard samples. The values of the roughness average parameter R_a were determined [36, 37]. Surface roughness measurements carried out for the different calibrated samples were again in good agreement with profilometer measurements.

3.7 Alternative Set Up for Surfaces Topography Determination

In all the different applications of evanescent illumination the presence of multiple plane wavefronts is required to observe the features of interest in the different cases. Gratings are optical elements that provide the generation of multiple plane wavefronts. For this reason a new interferometric configuration was designed for the topographic measurement of surfaces introducing a grating in the set up.

In [38], the specific details of the utilized specimen are given, a NIST standard of roughness with certified values. Figure 3.8 shows the interface between the glass surface where a grating is printed and the metallic surface of the standard. The surface of the standard consists of a saw tooth profile of nominal pitch $L_t = 100$ μm and nominal depth $h_t = 6$ μm with deviations that are included in the certification of the particular standard. The grating is formed by chromium strips deposited on optical glass. The face of the plate that carries the deposited chromium is supported on the standard surface.
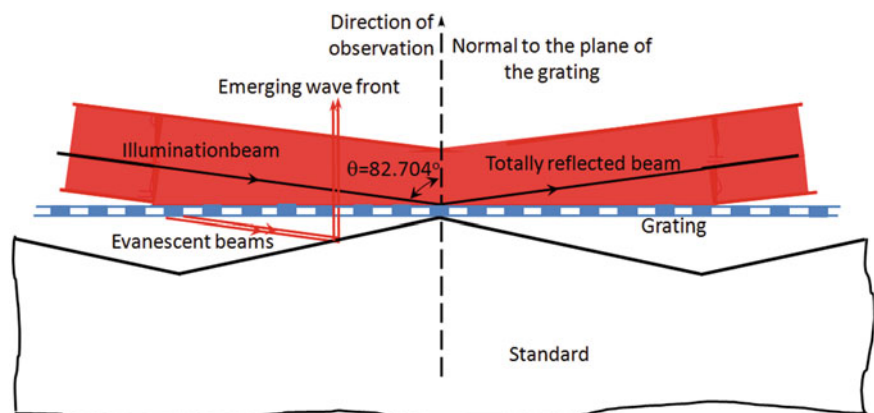
Fig. 3.8 Experimental setup for surface topography analysis including a grating
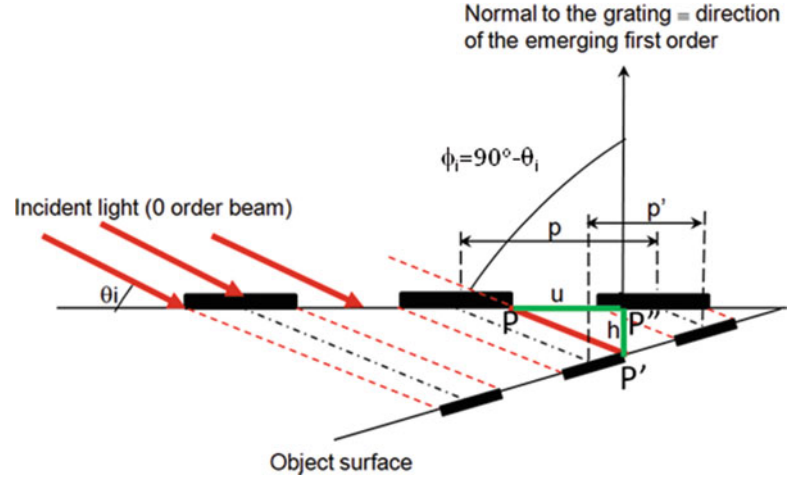
The resultant R_a (average depth) is 3 μm. This standard is used to calibrate devices based on the use of stylus probes. Figure 3.8 provides the schematics of the illumination and observation beams. The inclination of the illumination beam is greater than the critical reflection angle and therefore the light is totally reflected at the glass-air interface. However, as it was indicated previously, the electromagnetic field penetrates the cavity formed by the standard's surface and the plate that supports the grating. Schematically, Fig. 3.9 shows the trajectory of the photons that enter the cavity and according to the preservation of momentum, continue their trajectory. Photons carrying the surface topography information, as it is has been explained in Sect. 3.5, emerge along the direction approximately perpendicular to the grating surface according to the corresponding diffraction orders of the grating.

Figure 3.9 is a schematic representation to explain the contouring process model although from the theoretical point of view of optics the process is very complex if one considers all the steps required for a full explanation. A ray optics approach can be utilized to obtain the relationship between the depth information and the diffraction orders produced by the grating. The final result agrees with derivation of the classical equation of shadow moiré.

There is a difference to point out between the derivation that will be introduced and the classical deduction of the shadow moiré equation. The wavefronts that illuminate the grating in the classical analysis come directly from the light source. In the present derivation, the corresponding wavefronts come from the light source via evanescent illumination and the scattering light that is involved in transforming the evanescent waves in actual propagating waves. This fact introduces a $\pi/2$ rotation of the illumination beam.

Figure 3.9 shows the beam arriving at the grating face with an inclination θ_i with respect to the glass surface plane. The point observed at P'' has resulted from a displacement its original position P by the amount:

$$u = h \operatorname{tg}(90^\circ - \theta_i) \tag{3.12}$$

From the above equation utilizing the moiré relationship (see Chap. 15.2 of [32]), it follows:

$$h = \frac{\phi}{2\pi} \frac{p_j}{\operatorname{tg}(90^\circ - \theta_i)} \tag{3.13}$$

where ϕ is the phase and p_j is the pitch of the projected grating. Since $p_j = p\cos(90^\circ - \theta_i)$ it follows:

$$h = \frac{\phi}{2\pi} \frac{p}{\sin(90^\circ - \theta_i)} \tag{3.14}$$

The $\phi_i = (90^\circ - \theta_i)$ angle is complementary of the angle of incidence with respect to the normal to the plane of the grating. Equation (3.14) can be rewritten as:
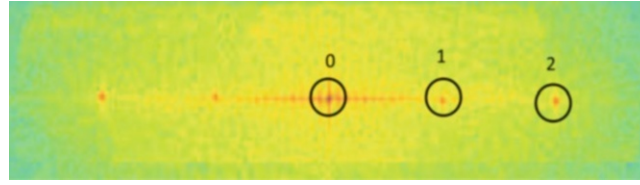
Fig. 3.10 FT pattern of the 5 μm pitch grating imaged by the CCD. The first harmonic (corresponding to the pitch of 5 μm) and the second harmonic (corresponding to the pitch of 2.5 μm) are visible in the FT spectrum

$$h = \frac{\phi}{2\pi}\frac{p}{\sin\phi_i} \tag{3.15}$$

Then one regains the equation of shadow moiré when the illumination source is at infinity and the observation direction is also at infinity and normal to the surface plane. Equation (3.15) is in agreement with Eq. (3.11) except for the factor 2 that comes from the double illumination that was utilized in the derivation of Eq. (3.11). The reader should pay attention to the multiple meaning of Eq. (3.11) since it will also apply to the contouring of a reflecting surface. In previous publications of the authors this point was not fully clarified.

The use of evanescent illumination makes it possible to utilize illumination beams that cannot be used in the classical shadow moiré because of the limitation imposed by the total reflection angle. Let us look at some examples of the utilization of Eq. (3.15). A grating of pitch p = 5 μm is utilized and the illuminating laser is a diode laser of $\lambda = 635$ nm. For the angle of diffraction of the first order it holds $\sin\theta_1 = \lambda/p = 0.635/5 = 0.127$: this corresponds to an angle of 7.296°. Since $\phi_1 = (90° - \theta_1)$, it follows $\phi_1 = 82.704°$. Consequently, sensitivity is S = 5/sin(82.704°) = 5.041 μm.

For the second order, the diffraction angle with respect to the 0 order can be computed as $\sin\theta_2 = \lambda/(p/2) = 0.635/2.5 = 0.254$: this corresponds to an angle of 14.7143°. Since $\phi_2 = (90° - \theta_2)$, it follows $\phi_2 = 75.286°$. Consequently, sensitivity is S = 2.5/sin(75.286°) = 2.585 μm. Figure 3.10 shows the corresponding FT pattern.

3.8 Analysis of the Observed Image

Figure 3.11 shows two images of the standard with the superimposed 5 μm pitch grating. Figure 3.11a corresponds to white light while Fig. 3.11b corresponds to coherent illumination. It is interesting to point out that, while the white light illumination that is normal to the plane of the grating shows peaks and valleys but does not separate them, the evanescent illumination shows higher intensities in the region of contact with the grating. The evanescent field dies out in depth within a fraction of the wavelength of light. By measuring the horizontal size (i.e. in the direction normal to fringes) of the bright region, the value of 11.81 μm is obtained. This means that the local depth at the end of the bright region is 11.81*tan[h_t/(L_t/2)], that gives 1.42 μm. The corresponding depth of penetration of the evanescent field can be computed by following equation from [39]:

$$z = \frac{\lambda}{2\pi\sqrt{\frac{\sin^2\theta_i}{n_{12}^2} - 1}} \tag{3.16}$$

where z is the depth of penetration of the evanescent field, and n_{12} is the relative index of refraction of the two media in contact, glass and air. Equation (3.16) gives z = 89.6 nm. Consequently, the size of the region of illumination caused by the evanescent field is 750 nm.

Figure 3.12 shows in the insert the distribution of light intensity in the image as a function of the coordinates of the field of view. It can be seen that the bright illumination corresponds to a very small region of the saw tooth surface. Figure 3.12 shows the intensity distribution in the brightest areas of the image. The area corresponding to the evanescent illumination depth of penetration obtained from Eq. (3.16) is a very small region. At about 3.5 times the theoretical distance computed with Eq. (3.16), the intensity begins to decay reaching a minimum of about 20 % of the maximum intensity in the central valley. The insert included in Fig. 3.12 shows the intensity distribution corresponding to three peaks.
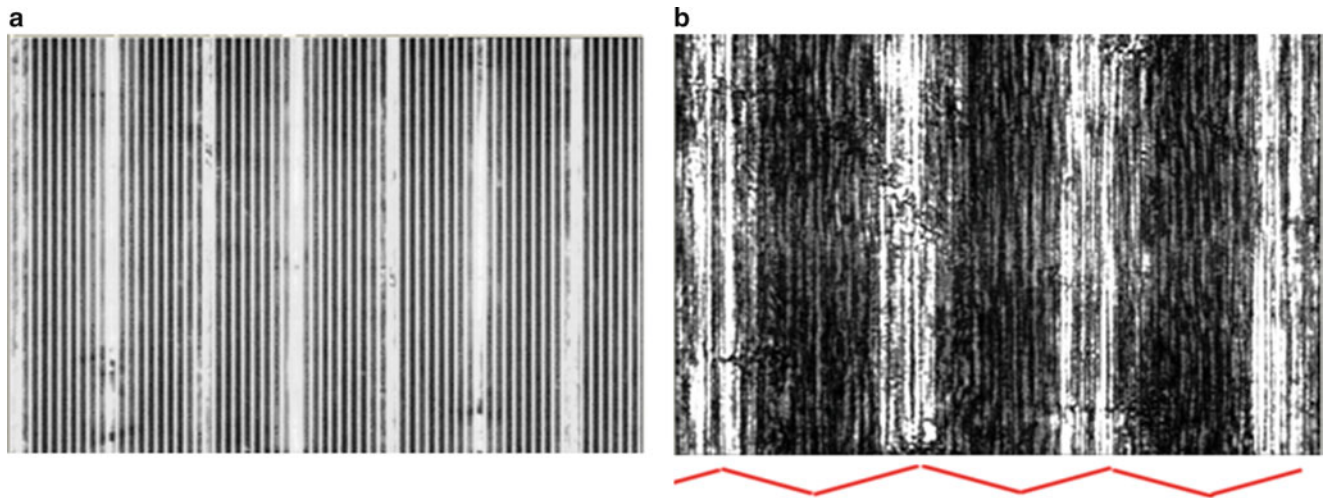
Fig. 3.11 (**a**) White light image of the HQC226 standard (5 μm grating superimposed on the specimen); (**b**) Coherent illumination image (5 μm grating superimposed on the standard). The profile of the standard is sketched below (**b**)
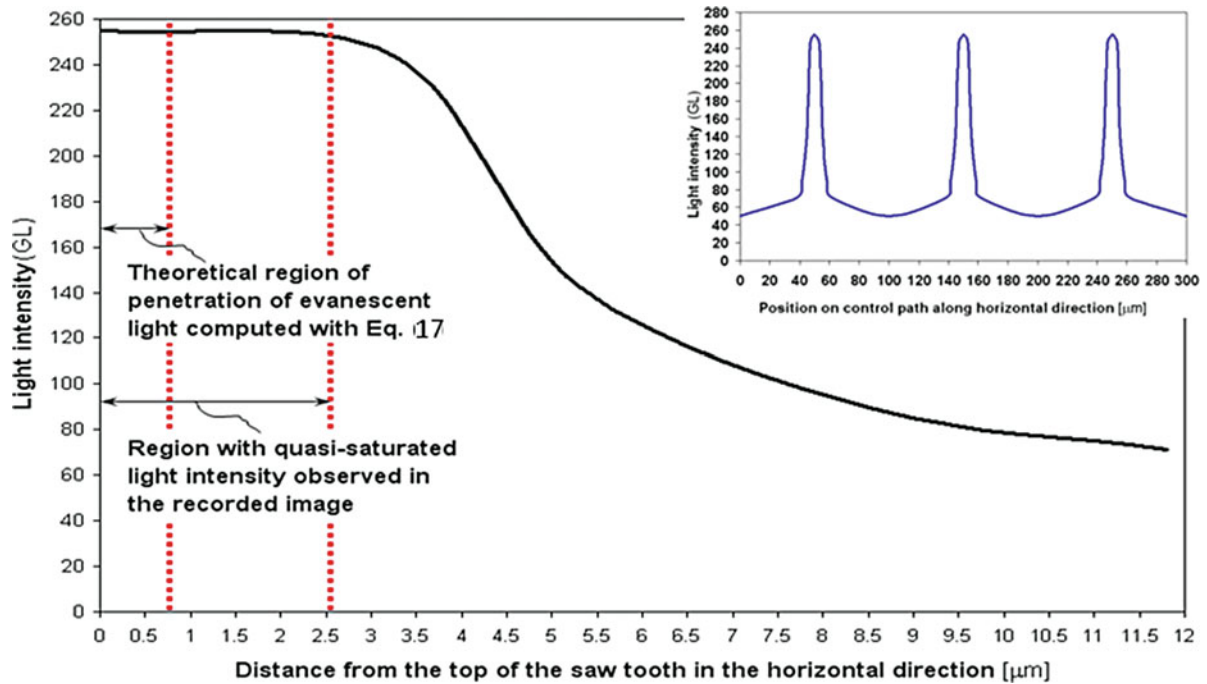
Fig. 3.12 Intensity distribution received by CCD camera sensor observing in the direction normal to the plane of the grating

The direct influence of the evanescent field is very small, however it is evident that in the created Fabry-Perot cavity there is enough light intensity to provide information on the depth of the standard and this intensity corresponds to the light scattered by the grating surface. This is a very important observation because extends the applicability of the evanescent illumination to greater depths than the actual depth of penetration, in the current case about 67 times.

Although the image of the standard surface (Fig. 3.11) is noisy, the signal carrying the depth information is present in the image as it can be seen in Fig. 3.10 and can be recovered through filtering.

3.9 Information Retrieval from the Recorded Image

The depth information contained in the image of Fig. 3.11 is recovered by the process of filtering of the diffraction orders shown in Fig. 3.10. The information is encoded in frequency changes of the carriers. There is an important feature that is necessary to point out. As it is shown in Fig. 3.13, the profile is formed by plane surfaces whose shape is similar to saw teeth. The frequency change of the carriers are the same whether the planes go up or down. Hence, it is necessary to develop a special data processing method that can provide the required information.

Illuminating the grating with coherent illumination (see Fig. 3.14) will generate a large number of diffraction orders. This means that the surface under analysis will be illuminated by a number of plane wave fronts. A number of diffraction orders are generated that depend on the properties of the grating. Although the theoretical computation of these orders is extremely difficult, they can be experimentally recorded as it is shown in Fig. 3.10. The increasing orders contain the same information but with higher sensitivity because they correspond to higher frequencies.

Figure 3.15 shows a magnified view of the orders 1 and 2 of Fig. 3.10. In these orders, it is encoded the information on the profile of the standard under analysis. By properly filtering orders 1 and 2, the modulating frequencies f_{mf1} and f_{mf2} due to the profile of the standard teeth can be obtained. Because the surfaces of the standard are plane, these frequencies are constants.

The corresponding moiré equation is:

$$I_{mc} = I_o + I_1 \cos\left[\frac{2\pi}{p_{cn}}x + \Psi_n(x,y)\right] \tag{3.17}$$

Fig. 3.13 Saw tooth profile: Lp length of the profile, depth 6 μm, and average depth, definition of Ra roughness of the standard
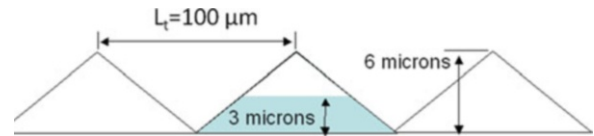
Fig. 3.14 The zero order beam impinging on the grating will produce the different diffraction orders of the grating
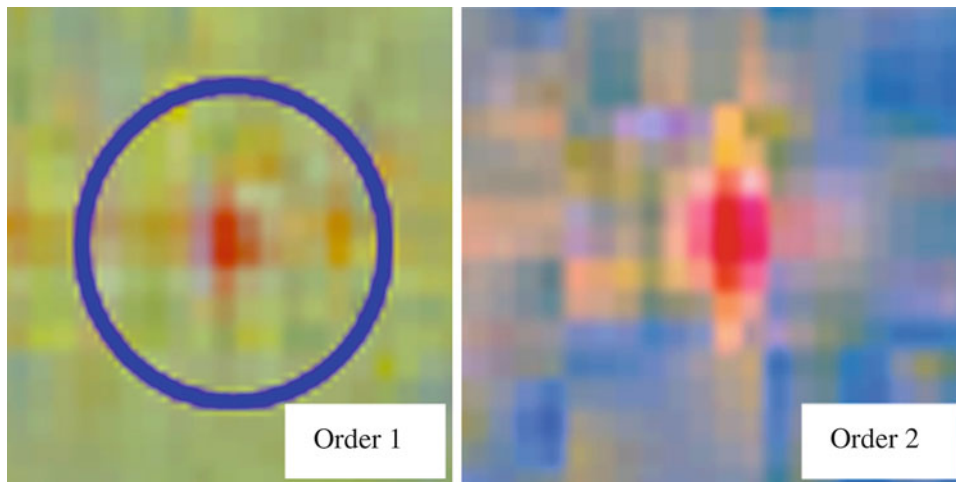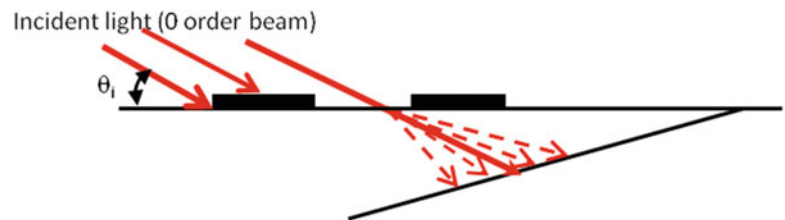
Fig. 3.15 Magnified view of the orders 1 and 2 of the FT shown in Fig. 3.10
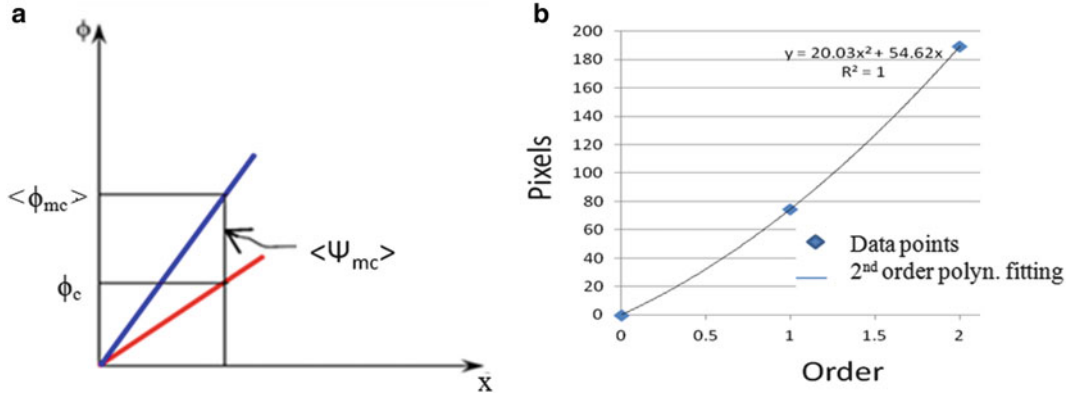
Fig. 3.16 (a) Phases of carrier fringes ϕ_c and modulated carrier ϕ_{mc}; (b) 2nd order polynomial data fitting corresponding to first and second harmonics

In Eq. (3.17), p_{cn} refers to the pitch while the subscript n indicates the particular analyzed order n (projected grating). Furthermore,

$$\Psi_n(x,y) = 2\pi f_{mfn}(y)x \tag{3.18}$$

In Eq. (3.18), Ψ_n is the modulating function and f_{mfn} is the frequency of the modulating function for a given order n (in this particular case, actual observed orders 1 and 2). As stated before, these frequencies f_{mfn} are constant and moiré does not discriminate between up and down. Hence the modulation function is a linearly increasing function of x as indicated in the preceding equation. The corresponding pitch p_{mf} is defined as one over the frequency of the modulation function.

In Fig. 3.16a, the phase corresponding to the carrier (projected grating) and the phase of the modulated carrier are plotted as a function of the coordinate x. In Fig. 3.16b, the pixel values of the two orders are fitted with a second order polynomial to take advantage of the presence of more than one order. The field of view is computed from the white light image of the grating as follows:

$$L_w = n_g p \tag{3.19}$$

The utilized microscope is corrected from chromatic aberrations: n_g is the order obtained from Fig. 3.11a while p is the pitch of the utilized grating.

Consequently, the number of teeth in the field of view is:

$$N_t = \frac{L_w}{L_t} \tag{3.20}$$

where L_t is the quantity defined in Fig. 3.13 and L_w is the field of view defined in Eq. (3.19).

We can compute N_t by utilizing the total phase changes of phase for x = L_w,

$$N_t(y) = \frac{1}{2} \frac{\frac{<\Psi_{mc}(L_w,y)>}{2\pi} p_{mc}}{\frac{\varphi_c}{2\pi}p} = \frac{p_{mc} < \Psi_{mc}(L_w,y) >}{p\,\varphi_c} \tag{3.21}$$

The coordinate y is introduced because the region under observation is an area where changes in L_t can occur.

Once N_t is determined, L_t and h can be determined as follows:

$$L_t = \frac{L_w}{N_t} \tag{3.22}$$

$$h_t(y) = \left(\frac{<\Psi(y)>}{2\pi} \right) p_{mf}(y) \frac{L_w}{2N_t} \tag{3.23}$$

The final values of L_t and h_t are the averages of all the sections computed at different y's.

3.10 Obtained Results

The value of R_a evaluated, accordingly to ANSI B46.1, from the experimental data gathered with the present advanced digital moiré contouring technique falls in the 3.0175 ÷ 3.0784 μm range certified by NIST. The average measured pitch is 101.24 μm with a standard deviation of ±0.322 μm. The average measured depth is 6.078 μm: the average value of Ra is hence 3.039 μm, well within the range of NIST's measurements. The difference between the value of roughness measured optically and the average value of roughness indicated by NIST is only 0.231 % (i.e. 3.039 μm vs. 3.032 μm).

Finally, a square area has been extracted from the image and resized to 2,048 × 2,048 pixels in order to precisely reconstruct the profile of one saw tooth. Figure 3.17a shows the tooth profile and Fig. 3.17b shows the reconstructed 3D shape of the tooth. The local height of the tooth profile measured in this region is 6.0645 μm, practically the same as the nominal height of 6 μm. The local length of the tooth is 101.3 μm, very close to nominal length of 100 μm. The average R_a is 3.0505 μm and oscillates between 3.0175 μm and 3.0785 μm. By extracting different profiles it is possible to make an estimate of the average surface finish of the standard. The average depth thus determined is 6.035 ± 0.1367 μm. Therefore, the finish of surface standard can be estimated as 0.1367 μm/0.635 μm, that is about λ/5. The utilization of the evanescent illumination together with data retrieval and data processing techniques provides high accuracy contouring information of a surface.

3.11 Surface Waves Properties Applied to Contouring of Interfaces Other Than Glass-Air-Metal: Topography of Rough Ceramic Surfaces

All the above presented developments are based in the surface plasmon polaritons, resulting from coupling of surface plasmons with light, a phenomenon that is associated with metallic surfaces. This was the basis of the developments introduced in [12, 14, 17]. However, the related phenomena imply a strong light-matter interaction that is not limited to the Fermi-electron layers present in metals. In spite of the apparent differences between the subject matter of the current paper and the subject matter of authors' papers [9–11, 13, 15, 16, 19–21] dealing with interaction at interfaces between dielectric materials both research efforts have a common basis: interactions of electromagnetic waves and matter at interfaces including cavities where electromagnetic resonances can take place. In both cases, light emission through scattering processes of the type analyzed by Berry [22] plays a prominent role in the formation of images that can be utilized for metrological purposes.
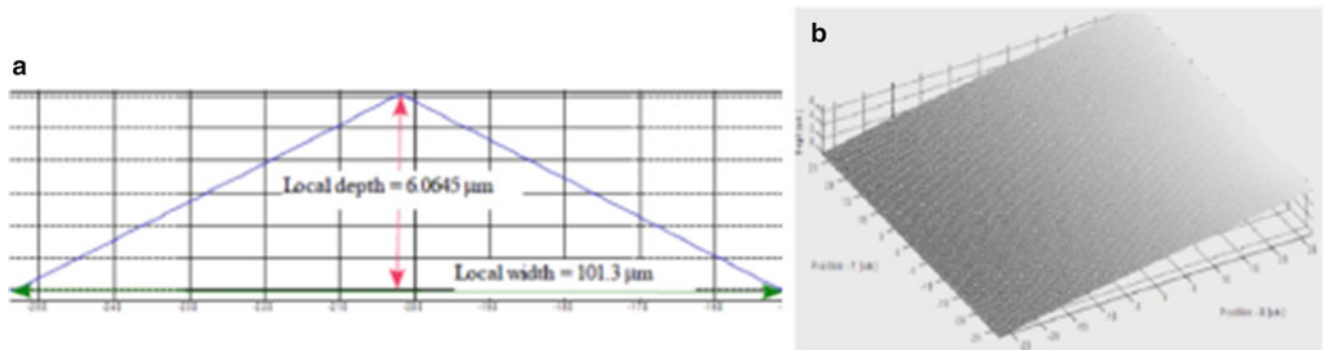
Fig. 3.17 (**a**) Detail of one tooth; (**b**) 3D MATLAB representation of the reconstructed surface of half of one tooth

Fig. 3.18 Weibull strength σ vs. Weibull surface roughness for all surface conditions
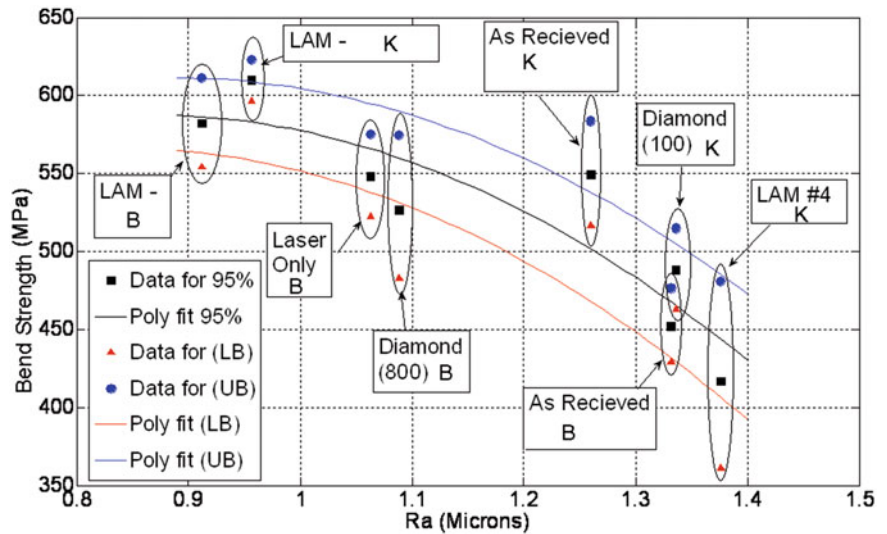
On the basis of these considerations the same set-up described in preceding sections was utilized to measure the different finishes of Si_3Ni_4 ceramics [40, 41]. Ceramics are known to have very high coefficient of friction, excellent compression strength and resistance to corrosion especially at elevated temperatures. These qualities make them great candidates for a variety of engine components as well as for bearings. There are two major factors, however, that currently limits the use of ceramics for these types of applications. The first has to do with control of surface defects: due to their brittle nature, a reduction or elimination of flaws that can prove critical in the failure of ceramics is required. The reduction of flaws is connected with the surface finish, surfaces are generally the areas of higher stresses and the surface finish plays a paramount role as a source of surface defects. The second factor has to do with cost. The manufacturing costs to diamond grind can be as much as 70–90 % of the total component cost for complex components. The wide application of advanced ceramics has been restrained largely because of the difficulty and high cost associated with shaping these hard and brittle materials into products. Laser assisted machining (LAM) of ceramics is proving to be a viable alternative to conventional manufacturing methods in terms of cost. Machining time can be cut down drastically and with proper arrangement other more complex machining operations may be possible. One of the main issues is ensuring that the surface finish of these ceramics be good enough or better than the conventional diamond grind process a well established industrial procedure. The method described in the preceding sections provides a very powerful tool to the study of the surface topography of ceramics due to the sub-micron accuracy that can be achieved and the relative simplicity of the set up that is required to make the observations. Due to their brittleness important defects that characterize their strength are in the micron and sub-micron range.

The same setup described in Sect. 3.6 was utilized to determine the roughness of Si_3Ni_4 ceramics. The same angle $\phi_i = 82.704°$, but the grating frequency was increased: p = 2.5 µm resulting in a sensitivity S = 2.52 µm. The room temperature flexural strength of Si_3Ni_4 ceramics resulting from two different processes of fabrication called materials A and B were investigated for different surface conditions. The strengths were measured with special specimens subject to four-point bending following the method described in [42]. The flexural strengths where correlated with surfaces finishes; the surface finishes were divided in the following four categories: as received, diamond grind (800 Grit), LAM turned, laser glazing. A total of 63 specimens were tested, a total of 81 measurements of roughness were performed per each different condition and a total of 324 measurements were carried out.

The correlation plotted in Fig. 3.18 indicates that as the Ra values increase the strength decreases. It can also be seen that variations of Ra of few hundred nanometers have a sizable effect on the strength. Figure 3.18 shows that from Ra = 900 nm to Ra = 1,400 nm a considerable change of the strength occurs. Further investigations will be carried out to see the factors that can explain the reason for the observed correlation. The obtained results support the conclusion that the interferometer designed on the basis of evanescent illumination is a very powerful tool that can be used for the determination of surface roughness values for any kind of material. The fact that this technique can be implemented with a far field microscope opens the possibilities to many other kinds of analysis with a simple experimental set up.
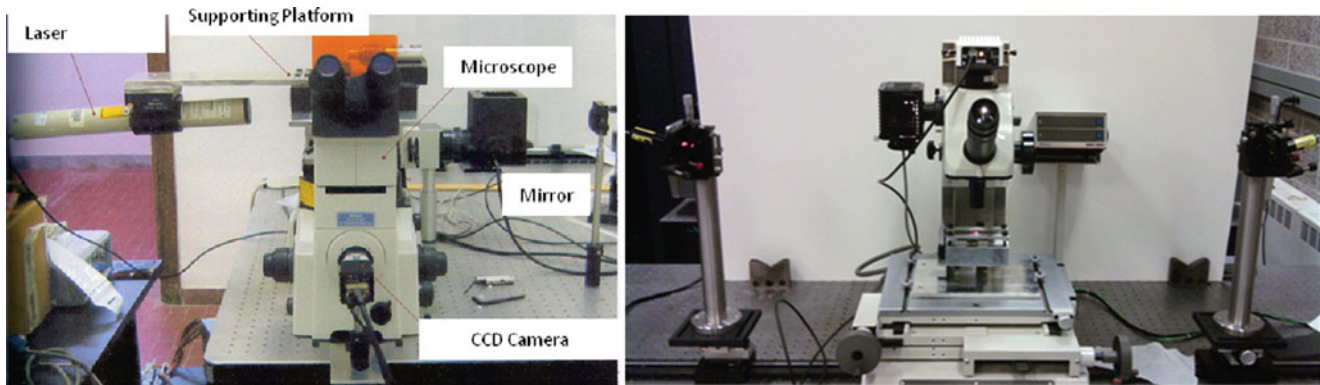
Fig. 3.19 Double illumination set ups

3.12 Optical Set-Ups

Figure 3.19 shows the implementation of two optical set-ups with double illumination. The one to the left of Fig. 3.19 utilizes a laser and a mirror to reflect the laser beam. The utilized microscope is a Nikon Eclipse TE200U inverted microscope with a CF160 infinity objective. The double illumination was achieved through reflection on a mirror. The set up to the right utilizes a Mitutoyo metrological microscope. Two procedures were utilized depending on the roughness of the observed surfaces. If Ra is larger than the sensitivity of the set up, the standard processing with fringe unwrapping was utilized. If the roughness is less than the sensitivity two images are taken separately, left and right, and the corresponding phases are subtracted. Since the phase difference between the patterns is less than a fringe order, one is dealing with fractional orders and the subtraction gives the surface profile directly without unwrapping. All the image patterns were processed with the Holo-Strain Analyzer^{TM} (HMSA) image processing software developed by C.A. Sciammarella and his collaborators [43].

3.13 Summary and Conclusions

This paper presented an approach to Experimental Mechanics applied to rough surfaces that reach the nano-scale based on the utilization of phenomena associated with evanescent illumination. Although it may appear to be a separate development of previous work of the authors in nano-mechanics it is very closely related. The main difference is the type of materials that is applied to, non transparent materials were the light reflection prevails over the transmission of light in the observed objects. The proposed methodology provides the means of measuring topography of micro and nano-sized roughness related parameters of surfaces. The extensive application of the interferometer based on evanescent illumination to the study of roughness in ceramics is briefly described in Sect. 3.11. The large amount of performed measurements, yields statistically reliable and meaningful values of parameters that characterize the properties of rough ceramics surfaces with accuracies in the range of nanometers and statistically related fracture mechanics properties of ceramics.

In the introduction section, a number of publications on surface waves closely related to material properties of the involved surfaces offer new avenues to apply Experimental Mechanics to the study of surface properties other than metrological. This may give a deeper insight in problems of great technological importance particularly in the field of additive manufacturing, a topic of great actuality. A great deal of theoretical and technological developments are required to materialize these potential capabilities of surface waves. However, it can be said that the basis for these developments are already available. The results presented in this paper provide a very powerful approach to the study of problems in the nano-field using light but these studies can be extended to other forms of radiation that, supplementing optical observations, can provide clues on the mechanics of fracture and on processes that can improve the capabilities of materials made with modern methods of manufacturing.

In view of the results presented in the paper, it can be concluded that surface waves play a unique role in the measurements of dimensions, shapes, and properties of surfaces. At the same time surface waves provide information at the nano-scale. Surface waves are the key to making Experimental Mechanics methodologies suitable for future nano-

engineering applications. Remarkably, nanometer resolutions can be achieved by using experimental setups that include conventional optical microscopes. This last point is critical as it now makes the possibility for anyone to research this topic without the need of very expensive equipment.

References

1. Tamm I (1991) Selected papers. In: Bolotovsky B, Frenkel V.Ya (eds), Springer, Berlin
2. Dyakonov (1988) New type of electromagnetic wave propagating at an interface. Zh Eksp Teor Fiz 94:119–123
3. Lakhtakia A, Polo JA (2007) Dyakonov-Tamm wave at the planar interface of a chiral sculptured thin film and an isotropic dielectric material. J Eur Opt Soc 2(22): Article No. 07021
4. Agarwal K, Polo JA, Lakhtakia A (2009) Theory of Dyakonov–Tamm waves at the planar interface of a sculptured nematic thin film and an isotropic dielectric material. J Opt A Pure Appl Opt 11(7): Article No. 074003
5. Gao J, Lakhtakia A, Lei M (2011) Synoptic view of Dyakonov–Tamm waves localized to the planar interface of two chiral sculptured thin films. SPIE J Nanophotonics 5(1): Article No. 051502
6. Faryad M, Lakhtakia A (2013) Prism-coupled excitation of Dyakonov–Tamm waves. Opt Commun 294:192–197
7. Maab H, Faryad M, Lakhtakia A (2013) Prism-coupled excitation of multiple Tamm waves. J Mod Opt 60(5):355–358
8. Pulsifier DP, Faryad M, Lakhtakia A (2013) Observation of the Dyakonov-Tamm wave. Phys Rev Lett 111(24): Article No. 243902
9. Sciammarella CA, Lamberti L (2006) Optical detection of information at the sub-wavelength level. In: Proceedings of the NANOMEC06 symposium on materials science and materials mechanics at the nanoscale, November 2006, Bari, Italy
10. Sciammarella CA, Lamberti L (2007) Determination of the shape of objects in the range 1/20 of the wavelength of light. In: Proceedings of 2007 SEM annual conference on experimental and applied mechanics, Springfield, MA, June 2007
11. Sciammarella CA, Lamberti L (2007) Observation of fundamental variables of optical techniques in the nanometric range. In: Gdoutos EE (ed) Experimental analysis of nano and engineering materials and structures. Springer, The Netherlands
12. Sciammarella CA, Lamberti L, Demelio G, Di Cuonzo A, Boccaccio A (2008) Application of plasmons to the determination of surface profile and contact stress distribution. In: Proceedings of XXXVIII meeting of the Italian Association of stress analysis. Rome, Italy, September 2008
13. Sciammarella CA, Lamberti L (2008) The equivalent of fourier holography at the nano-range. In: Proceedings of the 2008 SEM XI international congress on experimental and applied mechanics, Orlando, FL, June 2008
14. Sciammarella CA, Lamberti L, Sciammarella FM, Demelio G, Dicuonzo A, Boccaccio A (2009) High accuracy micro scale measurements using a conventional far field microscope. In: Proceedings of the 2009 SEM annual conference on experimental and applied mechanics, Albuquerque, NM, June 2009
15. Sciammarella CA, Lamberti L, Sciammarella FM (2009) The equivalent of Fourier holography at the nanoscale. Exp Mech 49:747–773
16. Sciammarella CA, Lamberti L, Sciammarella FM (2010) Light generation at the nano scale, key to interferometry at the nano scale. In: Proulx T (ed) Conference proceedings of the society for experimental mechanics series; experimental and applied mechanics, vol 6. Springer, New York, pp 103–115
17. Sciammarella CA, Lamberti L, Sciammarella FM, Demelio G, Di Cuonzo A, Boccaccio A (2010) Application of plasmons to the determination of surface profile and contact strain distribution. Strain 46(4):307–323
18. Sciammarella CA, Sciammarella FM, Lamberti L (2011) Experimental mechanics in nano-engineering. In: Gdoutos EE, Kounadis AN (eds) Recent advances in mechanics. Springer, The Netherlands, pp 275–312
19. Sciammarella CA, Lamberti L, Sciammarella FM (2011) Chapter 9: Optical holography reconstruction of nano-objects. In: Rosen J (ed) Holography, research and technologies. INTECH, Rijeka, pp 191–216
20. Sciammarella CA, Lamberti L, Sciammarella FM (2013) Chapter 11: Holography at the nano level with visible light wavelengths. In: Mihaylova E (ed) Holography—basic principles and contemporary applications. INTECH, Rijeka, pp 243–281
21. Sciammarella FM, Sciammarella CA, Lamberti L (2013) Chapter 17: Nano-holographic interferometry for in vivo observations. In: Shaked NT, Zalevskey Z, Satterwhite LL (eds) Biomedical optical phase microscopy and nanoscopy. Elsevier, The Netherlands, pp 353–385
22. Berry MV (1994) Evanescent and real waves in quantum billiards and Gaussian beams. J Phys A Math Gen 27:391–398
23. Girard C, Dereux A, Martin OJF, Devel M (1995) Generation of optical standing waves around mesoscopic surface structures: scattering and light confinement. Phys Rev B 52(4):2889–2898
24. Toraldo di Francia G (1952) Super-gain antennas and optical resolving power. Nuovo Cimento 9(S3):426–435
25. Toraldo di Francia G (1958) La diffrazione della luce. Edizioni Scientifiche Einaudi, Torino
26. Vigoureux JM (2003) De l'onde évanescente de Fresnel au champ proche optique. Annales de la Fondation Luis de Broglie 28(3–4):525–548
27. Papoulis A (1962) The Fourier integral and its applications. McGraw-Hill, New York
28. Kretschmann E (1974) Die Bestimmung der Oberflächenrauhigkeit dünner schichten durch Messung der Winkelabhängigkeit der Streustrahlung von Oberflächen Plasma Schwingungen. Opt Commun 10:353–356
29. Heitmann D (1977) Radiative decay of surface plasmons excited by fast electrons on periodically modulated silver surfaces. J Phys C Solid State Phys 10:397–405
30. Teng YY, Stern EA (1967) Plasmon radiation from metal gratings. Phys Rev Lett 19:511–514
31. van de Hulst HC (1981) Light scattering by small particles, 2nd edn. Courier Dover Publications, Mineola, NY
32. Sciammarella CA, Sciammarella FM (2012) Experimental mechanics of solids. Wiley, Chichester
33. Ebbeni J (1966) Etude du phénomène de moirure par réflexion d'un réseau plan sur une surface gauchie et de son application en analyse des contraintes et des déformations. VDI Experimentelle Spannung Analyse Berichte 102:75–81
34. Sciammarella CA, Combel O (1995) Interferometric reflection moiré. In: Pryputniewicz RJ et al (eds) Materials characterization interferometry VII, vol 2545, Proceedings of SPIE. International Society for Optical Engineering, Bellingham, WA, pp 72–85

35. Guild J (1956) The interference systems of crossed diffraction gratings. Clarendon, Oxford
36. Stout KF, Blunt L (2000) Three-dimensional surface topography, 2nd edn. Penton Press, London
37. International Organization for Standardization (1997) Surface texture: profile method—terms, definitions and surface texture parameters. ISO Specification 4287
38. Sciammarella CA, Lamberti L, Sciammarella FM (2010) High accuracy optical measurements of surface topography. In: Proulx T (ed) Conference proceedings of the society for experimental mechanics series; experimental and applied mechanics, vol 6. Springer, New York, pp 1–7
39. Born M, Wolf E (1999) Principles of optics, VIIth edn. Cambridge University Press, Cambridge
40. Sciammarella FM, Sciammarella CA, Lamberti L, Burra V (2010) Industrial finishes of ceramic surfaces at the micro-level and its influence on strength. In: Proulx T (ed) Conference proceedings of the society for experimental mechanics series; experimental and applied mechanics, vol 6. Springer, New York, pp 9–16
41. Sciammarella FM, Matusky MJ (2011) Correlation between mechanical strength and surface conditions of laser assisted machined silicon nitride. In: Proulx T (ed) Conference proceedings of the society for experimental mechanics series; experimental and applied mechanics, vol 5. Springer, New York, pp 187–197
42. Quinn GD (2012) Fractography of glasses and ceramics VI: ceramic transactions, vol 230. Wiley-American Ceramic Society, USA
43. General Stress Optics Inc. Holo-Moiré Strain Analyzer Software HoloStrain, Version 2.0. General Stress Optics Inc, Chicago, IL. http://www.stressoptics.com

Chapter 4
Strain Assessment in Cracked Sheet Metals by Optical Grid Method

M. Sasso, G. Chiappini, M. Rossi, and D. Amodio

Abstract The present work is an extension of the optical grid method for the experimental investigation of deformation in stamped sheet metals. The classical method presents difficulty or is inapplicable where the deformation has resulted in the tearing of the sheet. In these cases, the measurement result is not available right where the most interesting data are expected.

In this work, authors try to overcome this limitation by using an xFem-like approach, in which the classical polynomial functions, that approximate the displacement field of an element, are supported by additional "enrichment" functions that allow to describe discontinuities or singularities of the displacement field. The proposed method extends the measure up to the free edge of the crack. The capabilities of this technique were assessed by two virtual experiments where the deformed shapes were assumed as samples; the former consists in a FEM simulated bulge test and permitted to evaluate the method error in presence of through crack; the latter reproduces a plausible crack opening and was used to evaluate the method behavior in presence of a crack tip.

Keywords Optical grid method • xFem • Sheet metal • Stamping • Crack

4.1 Introduction

The work represents an attempt of improving the optical method named "grid method" usually adopted to evaluate the strain fields in stamped sheet metals.

The optical grid method consists in applying a regular pattern of markers onto a surface before it is deformed by some forming process. Sometimes, especially in the early or manual versions of the method, ink circles are used, that deforms and turns into ellipses giving a straight visual evaluation of the principal strains; measuring the axis lengths provide a quantitative measure as well. However, the pattern most often consists of evenly spaced dots or intersecting lines, with each dots or line cross representing the markers. This technique is well suited to be automated and, as such, adopted in commercial system.

In experimental mechanics nowadays, more appealing techniques are present for the full field measurement of displacements and strains; in particular the digital image correlation (DIC) is by far the most employed method both in laboratory [1, 2] and industrial environments [3]. Its advantages mainly lies in the ease of surface preparation with a relatively high accuracy and resolution.

Comparatively, the grid method appears quite outmoded and obsolete from certain point of view; nevertheless, still exist applications where DIC cannot be employed, such as the sheet metal stamping and, actually, all the processes where the material cannot be observed during the deformation course.

This work originates from the need of measuring the strain field in a stamped part for automotive application (see Fig. 4.1); in such a situation, the only applicable technique is the grid method since the object cannot be framed during the process and no image correlation can be carried out. Typically, a stereoscopic camera system is used to compute the 3D coordinates of

M. Sasso (✉) • G. Chiappini • M. Rossi • D. Amodio

Dipartimento di Ingegneria Industriale e Scienze Matematiche, Universita Politecnica delle Marche,
Via Brecce Bianche, 60131 Ancona, Italy
e-mail: m.sasso@univpm.it

Fig. 4.1 Cracked stamped part with grid pattern
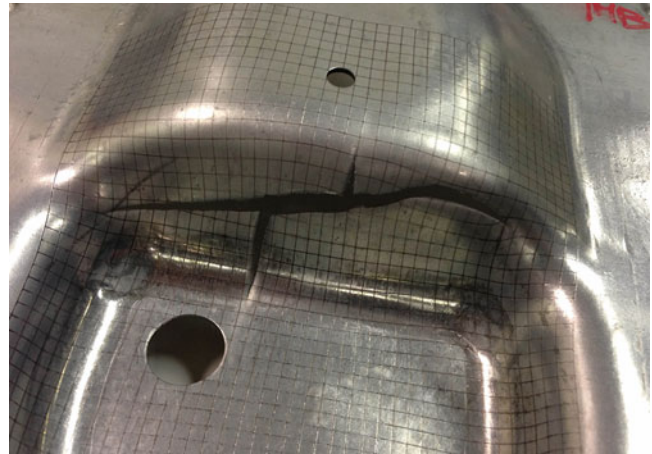
the markers onto the deformed surface; then, by knowledge of their positions in the undeformed configuration, actual strains in the framework of large displacement can be computed [4, 5].

Considering the stamped part in Fig. 4.1, the large crack observed in the central region actually prevents from measuring the strains in the most revealing region, where a comparison with FEM analyses would provide a useful insight for the process engineering.

In the following of the paper, a possible method for extending the measure of the strain by optical grid method up to the crack edge is numerically assessed.

4.2 Optical Grid Method with Enriched Shape Functions

The main steps of this technique can be described as follows: a grid of known pitch is impressed on the sheet blank before stamping; the grid, after being deformed during the forming process, is acquired by means of two high resolution cameras which were previously calibrated in a common reference frame. The grid intersections are recognized by digital image analysis, while a stereoscopic algorithm is used to computed their 3D coordinates X in an absolute reference frame.

Clearly, the elements of the grid that are cut by the crack in Fig. 4.1, cannot be processed anymore with the simple formulation explained above; indeed, the nodes lying at opposite sides of the crack experimented large displacements which are not related to the internal deformation of the element, since the continuity of the internal displacement is not satisfied.

In this work the authors propose a method, here named xGrid, to extend the measurement, with some hypothesis and approximations, to the free edges of the crack. For doing this, the typical bilinear shape functions that are used in quadrangular elements must be modified. We borrowed the enrichment shape functions used in the xFem formulation [6] for taking into account the discontinuity in the displacement field.

This approach, which has been already adopted with some success to DIC [7], is more challenging in the optical grid method, since the mesh cannot be chosen in anyway (it is given by the measured points) and only the deformed and undeformed configurations are known, without information of intermediate steps.

The problem of the internal displacements within a FE mesh, where one or more elements are cut by a crack, is expressed by:

$$X(U,V) = \sum_{i \in I} X_i \phi_i(U,V) + \sum_{j \in J} b_j \phi_j(U,V) H(U,V) \sum_{k \in K} \phi_k(U,V) \left[\sum_{l=1}^{4} c_k^l F_l(U,V) \right] \tag{4.1}$$

where I is the set of nodes belonging to regular elements, J is the set of nodes of the elements that are completely cut by a crack, K is the set of nodes of the elements containing crack tips; $\phi(U, V)$ are the bilinear shape functions representing the partition of unity within the element domain in natural coordinates U and $V \cdot H$ is a discontinuous or jump function, given for example by

Fig. 4.2 Enrichment configuration

$+$ I, regular grid

$\bigcirc$ J, cut elements

$\Diamond$ K, crack tip

a

b   e1 by xGrid with strain strain energy min.

c   e1 by xGrid with strain diff.
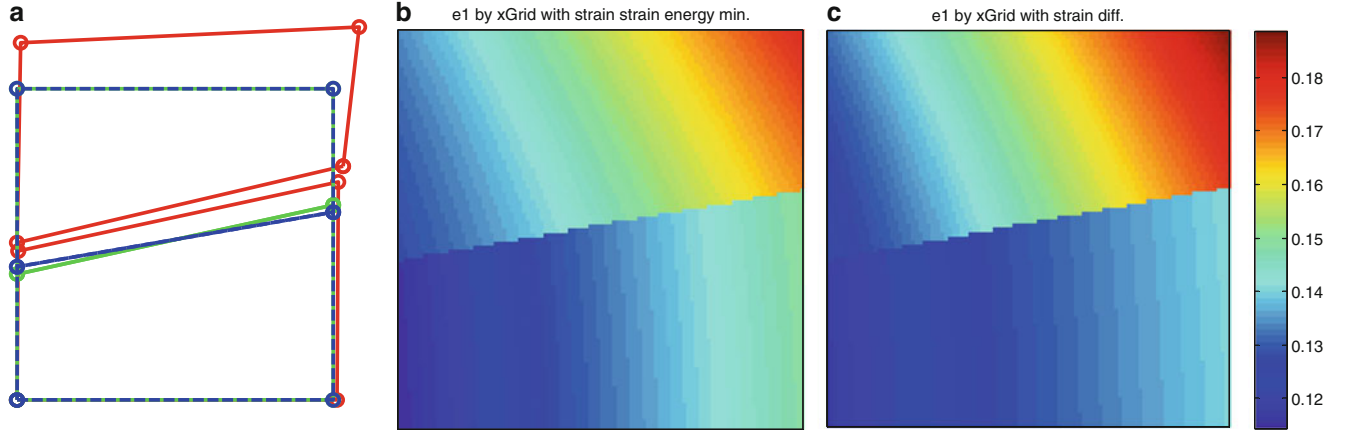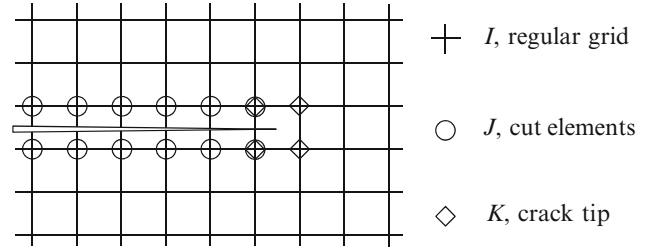
Fig. 4.3 Single square element with through crack

$$H(U,V) = \begin{cases} +1 & for\ 1^{st} subelement \\ -1 & for\ 2^{nd} subelement \end{cases} \tag{4.2}$$

The functions F_l are enrichment shape functions that contain the field discontinuity that one desires to take into account. In [6], four functions are suggested for approximating the displacement field close the crack tip:

$$\{F_l(U,V)\} = \left\{ \sqrt{r}\sin\frac{\theta}{2}, \sqrt{r}\cos\frac{\theta}{2}, \sqrt{r}\sin\frac{\theta}{2}\sin\theta, \sqrt{r}\cos\frac{\theta}{2}\sin\theta \right\} \tag{4.3}$$

where r and θ are the local polar coordinates with respect to the crack tip. The nodes sets are illustrated in Fig. 4.2.

Then, accordingly to the continuum mechanics theory [8] for large strain, the strain gradient tensors can be computed:

$$F^0_{pq} = \frac{\partial X^0_p}{\partial U_q} \text{ for the undeformed configuration} \tag{4.4a}$$

$$F_{pq} = \frac{\partial X_p}{\partial U_q} \text{ for the deformed configuration} \tag{4.4b}$$

The logarithmic principal strains on a tangent plane are computed solving the eigenvalue problem:

$$\varepsilon_{1,2} = \log\sqrt{eig\left[\left(F^{0T}F_0\right)^{-1}F^T F\right]} \tag{4.5}$$

Considering for example a single square element of unit side length, subjected to membranal displacements, a possible deformed shape, with an internal through crack, is the red one shown in Fig. 4.3a.

The four corners of the element could be regarded to as the markers of a regular grid impressed onto the planar sheet metal before the stamping; in this way, the coordinates of the four corners are known in both undeformed and deformed (e.g. by photogrammetry) configurations. On the contrary, the coordinates of the mid-side nodes cut by the crack still can be

measured in the deformed configuration (finding for example the intersection between the grid lines and the crack edge), but are not known in the undeformed configuration. As such, their initial positions are to be guessed, and then optimized according to some minimization criterion.

The optimization criteria may vary from application to application; here two similar approaches have been adopted and compared: the former involves the minimization of the global strain energy, the latter consists in minimizing the strain difference (in a tensorial sense) between the conjugate boundaries of the two sub-elements. Obviously their results depends on the nodal displacements, however, preliminary virtual tests showed that they are likely to provide quite similar results, as shown in Fig. 4.3b, c.

What is important is that both methods permit to compute the strain field up to the free edges of the through crack; this already represents an upgrading of the classical grid method where the elements cut by the crack are inevitably omitted from the strain computation. In order to assess the capability of the proposed method, two virtual experiments were conducted, the first regarding the measurement of strain in elements that are totally cut by the crack, the second dealing with the elements containing the crack tip.

4.3   First Virtual Experiment

In the first virtual experiment, a FE model was used to simulate a sheet metal that is deformed by a spherical bulge. The sheet metal is a square of 100 mm size, and 1 mm thickness, whereas the rigid punch has a radius of 20 mm and is moved upward of 25 mm after the initial contact occurring at the bottom surface of the sheet metal. The sheet metal was assumed to have an elastic modulus of 200 GPa, a plastic tangent modulus of 1 GPa, and a yield stress of 200 MPa.

The sheet metal has been modelled with shell 43 elements, 0.5 mm in size, by means of Ansys® software, exploiting the double symmetry with respect to x and y axes (Fig. 4.4a). The nodes at the external sides were fixed, while the nodes on the internal side obeyed the symmetry conditions, exception made for the nodes from the centre to the x = 24 mm coordinate, that was constrained in the y direction by means of uniaxial non-linear springs; the springs have an almost infinite stiffness up to 205 N, then their stiffness vanishes. In this way it was possible to reproduce a plausible fracture initiation and propagation, with accumulation of plastic strain at the crack edges, without the need for element kill (which would correspond to material loss). The nodal results of this simulation computed at the middle of the shell thickness, see Fig. 4.4b, were used as a virtual object to be acquired by means of optical grid method, both the classical one and the extended method here proposed.

The first principal strain distribution, as computed by the fine meshed FE simulation, mapped onto the undeformed configuration, is reported in Fig. 4.5, where the symmetry has been expanded to show the entire map.

Now, if the FEM shape is considered as real, the nodal displacement data can be "sub-sampled" thus to simulate what would happen in the optical grid method, where the markers are evenly spaced in the undeformed configuration; here a regular grid of 4 mm pitch is considered.
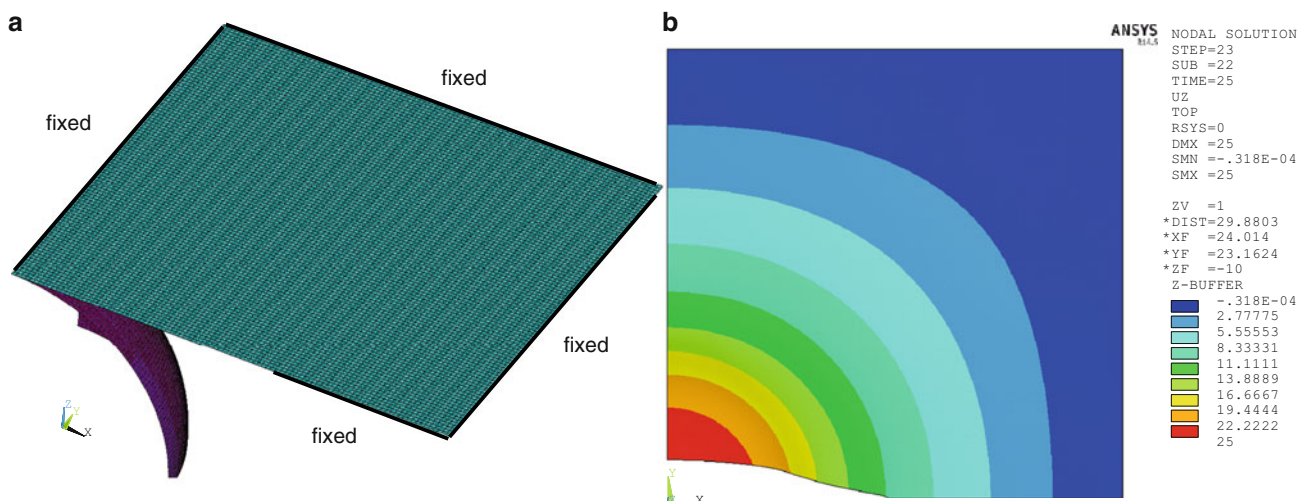
Fig. 4.4 FE model (**a**) mesh and constraints, (**b**) resulting out-of-plane displacements

Fig. 4.5 "Real" strain
distribution map as computed
by FEM with fine mesh
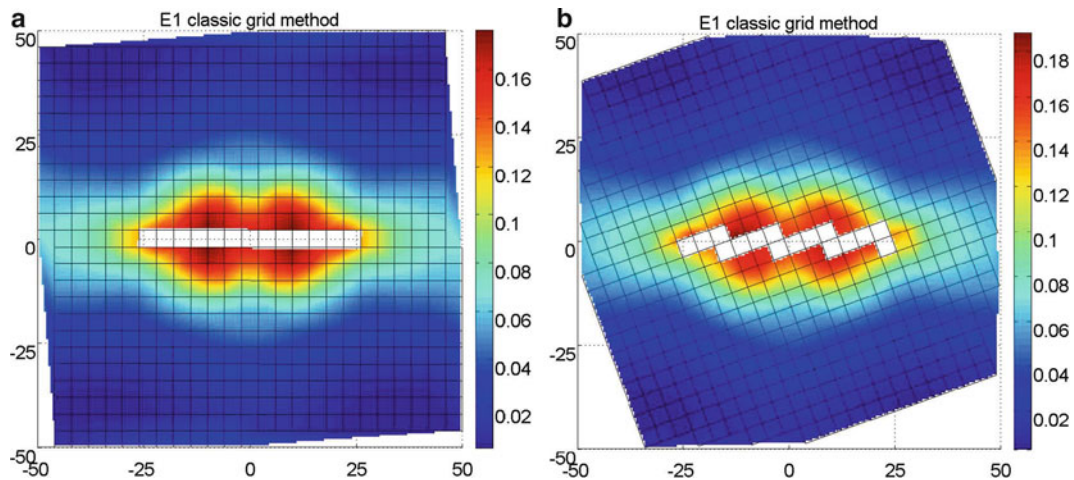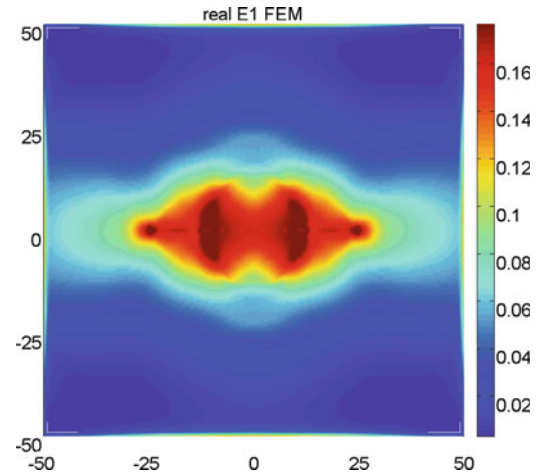
Fig. 4.6 ε_1 distribution map for a simulated grid method of pitch 4 mm (**a**) crack parallel, (**b**) 20° rotated

This sub-sampling results inevitably in a sort low-pass filter of the real data, with the stronger gradient and peak values being smoothed. Figure 4.6 show example of results, in terms of the first principal strain, with the grid parallel to the crack (a) and with a rotation of 20° and 3 mm offset in Y direction (b). The virtual markers grid is superimposed for clarity. The missing areas correspond to the elements that are cut by the grid. It is observed that the grid method, because of the pitch size, provides quite accurate results where the strain gradient are small, but the error increases when the strain concentrates, with an underestimation of 2–3 % in the yellow areas; at the crack tip the error is even greater but this region is not considered in this part of the work.

The extended optical grid method has been applied to the subsampled nodal data, including the enrichment shape functions given by Eq. (4.2) and adopting the strain difference minimization criterion. In this way, it was possible to assess the strain even in the elements that are cut in two parts by the crack. The corresponding first principal strain maps are reported in Fig. 4.7. Now, only the two elements containing the crack tips are omitted from the computation, with the newly filled areas appearing to provide acceptable results.

Since all the data are mapped onto the undeformed configuration, it is possible to compute the error by performing a simple point-by-point subtraction of these data from the real ones given in Fig. 4.5. Moreover, the error due to subsampling depends on the position (offset) of the grid with respect to the crack; for this reason, the grid has been offset in X and Y direction in several positions, obtaining the maximum error distribution shown in Fig. 4.8.

One observes that the extension of the measure to the elements with the through crack determines a widening of the area with larger error, but the absolute value of the error doesn't increase significantly.
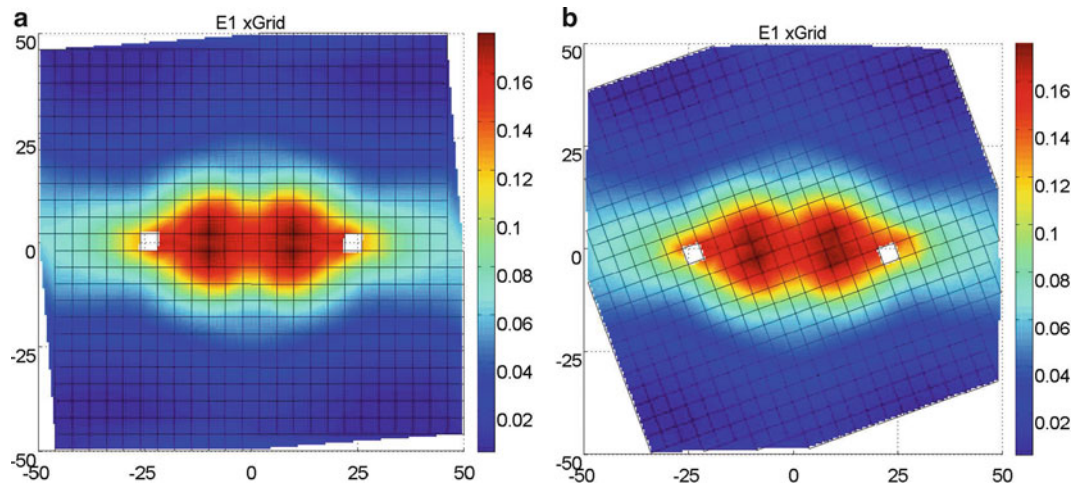
Fig 4.7 ε_1 distribution map for the simulated extended grid method of pitch 4 mm (**a**) crack parallel, (**b**) 20° rotated
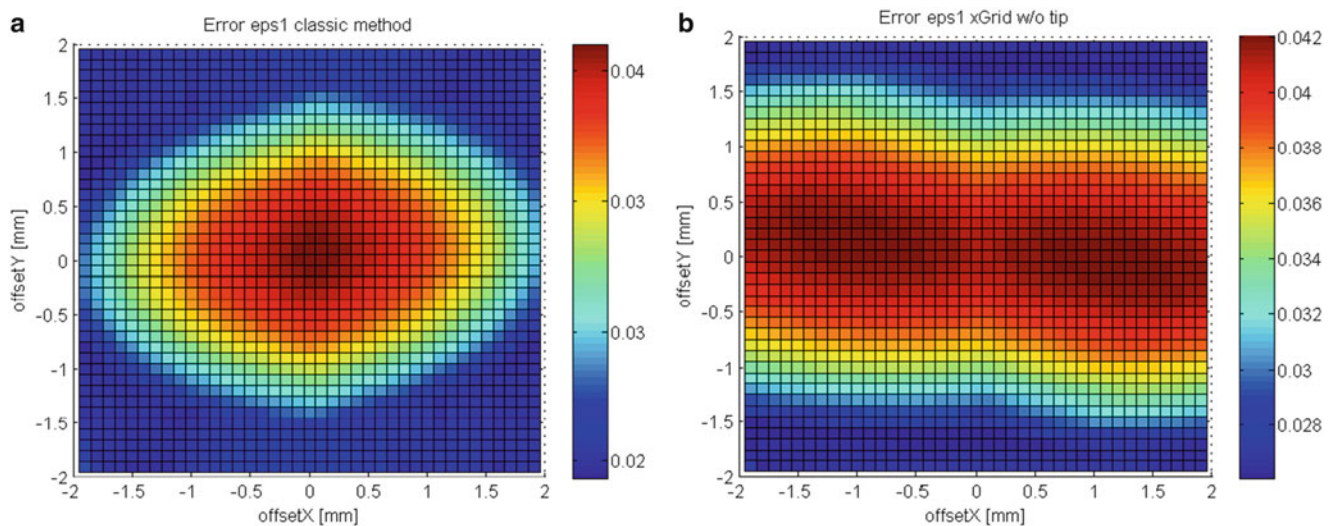
Fig. 4.8 ε_1 maximum error in the strain maps varying the grid offset (**a**) classic grid method, (**b**) extended grid method

4.4 Second Virtual Experiment

The second virtual experiment aims to evaluate the error committed when trying to extend the measure of strain to the elements containing the crack tip. Again the FE method was used; now a quadrilateral of 8 mm side is meshed with Plane 42 elements, 0.1 mm in size, with an initial crack ranging from the middle left side to centre of the square. Plausible in-plane displacements were assigned to the external sides of the quadrilateral, determining the crack opening shown in Fig. 4.9; again, the nodal data were assumed as real data and the red square represents a typical 4 nodes element with 4 mm side that can be acquired by the optical grid method. The distribution of the first principal strain, as computed by fine meshed FEM model, is reported in Fig. 4.10a. Conversely, Fig. 4.10b shows the strain map computed by the 4 nodes element having centroid at the crack tip (yellow square in Fig. 4.9) and using the enriched shape functions given by Eq. (4.3).

Analogously to error analysis shown in the previous section, the 4 nodes element can be offset in several positions, leading to different values of maximum error with respect to the "real" data. The distribution of maximum error as function of the offsets in X and Y directions is shown in Fig. 4.11. It must be noted that the error represents the maximum difference between the real strain map (Fig. 4.10a) and the strain map of the 4 nodes element, avoiding from the computation a region within 0.2 mm from the crack tip, where the strain field is singular.

Fig. 4.9 Second virtual experiment
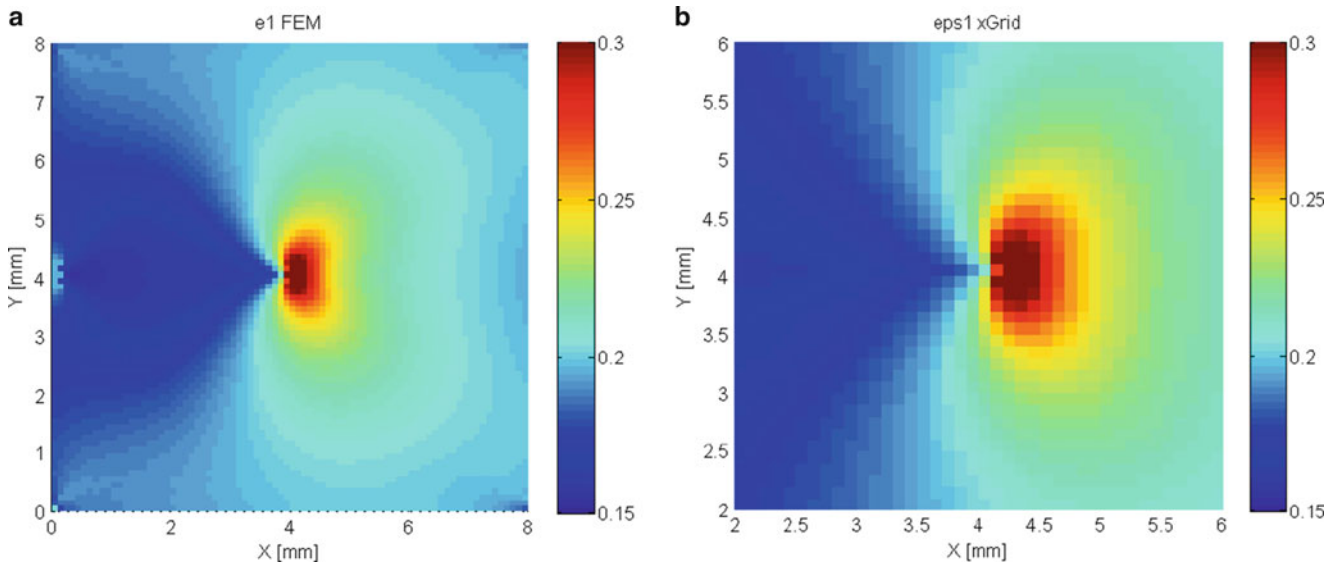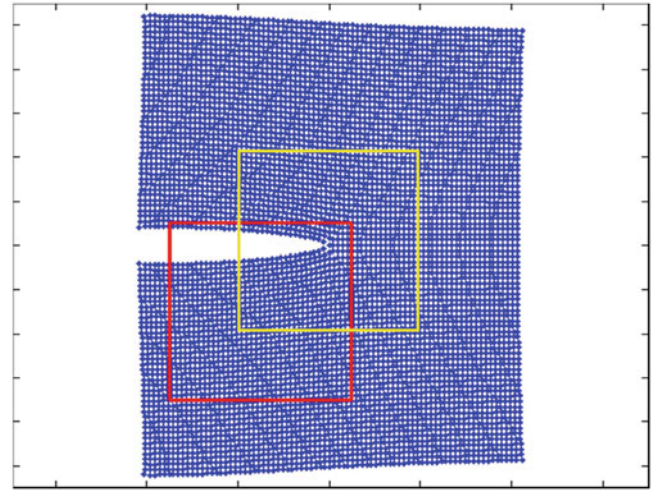
Fig. 4.10 ε_1 distribution within the entire quadrilateral (**a**) and the 4 nodes element with xGrid (**b**)
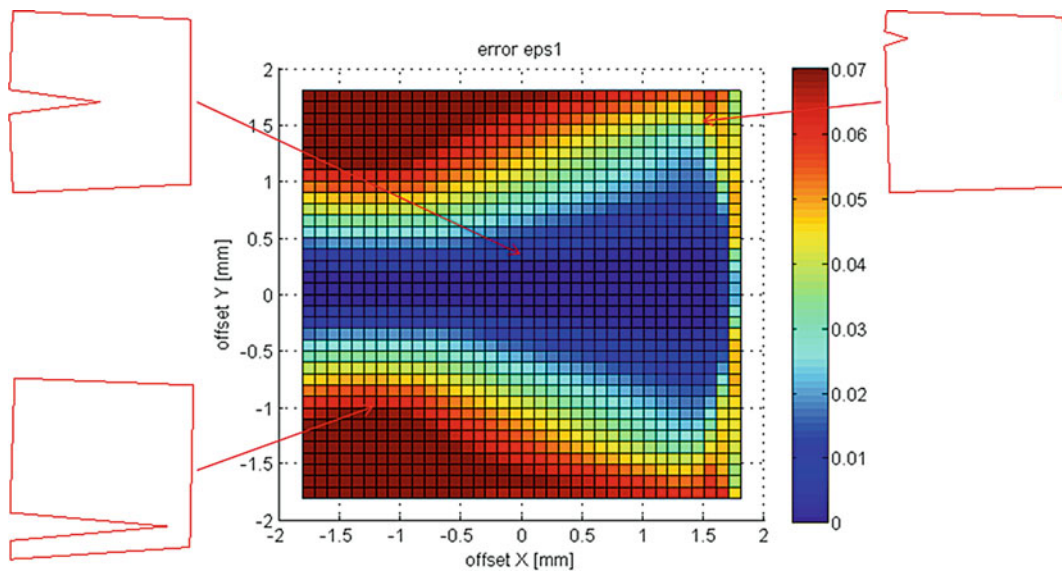
Fig. 4.11 ε_1 maximum error in the strain maps including the crack tip and varying the grid offset

It is observed that the error committed is quite large when the crack tip occurs close to the nodes, and especially when the crack cut almost completely the element in two parts; on the contrary, when the crack tip occurs in central region of the element or it is close to a node but without cutting too much of the element, the error is reduced, becoming acceptable for a pretty large area.

4.5 Conclusions

The xFem approach, based on the utilization of enriched shape functions for taking into account the discontinuities in the displacement fields, has been implemented in a simulated grid method procedure, aiming to extract information on the strain levels up to crack edges. The first results of this extended grid method appears encouraging for further study. In particular, the elements with a through crack provided good matching with the source FEM data, highlighting errors that are of the same magnitude of the unavoidable low-pass effect due to grid subsampling. As may be expected in advance, some difficulties are present if one try to extend the measure to elements containing the crack tips, especially when the crack determines a large distortion of the element itself; however, when the crack tip occurs at the central region of the element, errors are very small.

References

1. Palmieri G, Sasso M, Chiappini G, Amodio D (2011) Virtual fields method on planar tension tests for hyperelastic materials characterisation. Strain 47(2):196–209
2. G. Chiappini, M. Sasso, M. Rossi (2012) Comparison of contouring techniques applied to sheet metal testing. In: Proceedings of SEM 2012 annual conference, vol 3, pp 173–182
3. Sutton MA, Orteu JJ, Shreier HW (2009) Image correlation for shape, motion and deformation measurement: basic concepts, theory and applications. Springer, New York
4. Sasso M, Callegari M, Amodio D (2008) Incremental forming: an integrated robotized cell for production and quality control. Meccanica 43:153–163
5. M. Sasso, D. Amodio (2006) Development of a biaxial stretching machine for rubbers by optical methods. Proceedings of SEM 2006 annual conference, vol 3, pp 1161–1171
6. Nicholas M, Dolbow J, Belytschko T (1999) A finite element method for crack growth without remeshing. Int J Numer Methods Eng 46:131–150
7. Rethore J, Hild F, Roux S (2008) Extended digital image correlation with crack shape optimization. Int J Numer Methods Eng 73:248–272
8. Chung TJ (1988) Continuum mechanics. Prentice Hall International, Englewood Cliffs

Chapter 5
A Preliminary Investigation on the Mechanical Behavior of Umbilical Cord With Moiré Techniques

A. Boccaccio, R. Brunelli, L. Lamberti, M. Papi, T. Parasassi, M. De Spirito, and C. Pappalettere

Abstract The umbilical cord is a peculiar and complex structure, about 50–60 cm in length and 1–2 cm in diameter, that is essentially composed of three vessels, i.e. the umbilical vein and two umbilical arteries, arranged in coils around the vein surrounded by a great amount of support tissue, the Wharton's Jelly (WJ) that binds and encases the umbilical vessels.

WJ is a mucoid connective tissue (5 % cells, 95 % extracellular matrix) described as a three-dimensional spongy network of interlacing collagen fibers and small woven bundles of glycoprotein microfibrils with an interdispersed soluble phase composed by hydrophylic hyaluronans and proteoglycans. WJ grants the protection of the umbilical vessels against compressive forces due to fetal movements and uterine contractions and is very important to guarantee venous and arterial umbilical blood flows.

WJ response to mechanical loading is not well understood; another unsolved problem concerns WJ putative contribution to store and release the energy of the cardiac cycle, therefore in maintaining the anterograde flow in the cord arteries.

This article presents a preliminary study on the mechanical behavior of umbilical cord. For that purpose, an optical set up based on intrinsic moiré will be developed. Slices cut in the transverse directions of the cord will be submitted to equibiaxial tests and specimen deformations will be monitored in real time with moiré by printing a grating on the cord slice. In this way, it will be possible to gather information on the mechanical anisotropy of the cord.

Keywords Umbilical cord • Intrinsic moiré • Biomaterials • Mechanical characterization • Anisotropy

5.1 Introduction

Among experimental techniques used for characterizing biomaterials, optical tehcniques play a role of crucial relevance. They do not require a direct contact with the sample to be investigated (unlike Atomic Force Spectroscopy [1–4]), are full-field techniques and provide large amounts of data, are much less expensive than tactile mechanical systems [5, 6]. The issues related to the characterization of biomaterials not only refer to the choice of the experimental technique more suited for the specific scopes, but also to the handling of the biological samples and to the design of the loading frame required to perform the test. Moiré techniques are commonly used to characterize biological membranes [7, 8].

The umbilical cord is a complex and fascinating structure that connects the fetus to the placenta and encases the umbilical vessels. The response of its tissues to mechanical loading due to fetal movements and uterine contractions is not

A. Boccaccio • L. Lamberti • C. Pappalettere
Dipartimento di Meccanica, Matematica e Management, Politecnico di Bari, 70126 Bari, Italy

R. Brunelli (✉)
Dipartimento di Scienze Ginecologico-Ostetriche e Scienze Urologiche, Università "La Sapienza", 00161 Roma, Italy
e-mail: roberto.brunelli@uniroma1.it

M. Papi • M. De Spirito
Istituto di Fisica, Università Cattolica del Sacro Cuore, 00168 Roma, Italy

T. Parasassi
Istituto di Farmacologia Traslazionale, Consiglio Nazionale delle Ricerche, 00133 Roma, Italy

well understood. Different studies are reported in the literature on the mechanical characterization of the umbilical cord. Pennati [9] subjected samples of umbilical vein and Wharton's Jelly to uniaxial tensile and stress-relaxation tests. Li et al. [10] determined the biomechanical properties of different segments of human umbilical cord vein. Hamedani et al. [11] compared the mechanical properties of saphenous and umbilical veins. The interest in the mechanical characterization of the umbilical cord derives from the fact, already well known, that a strict correlation exists between the health of the neonate and the mechanical behavior of the cord. An umbilical cord not sufficiently stiff does not allow to provide the blood to the fetus at the correct values of pressure and flow.

In this study the intrinsic moiré technique is utilized to study the in-plane displacement field of 2 mm thick slices of umbilical cord subject to equibiaxial load. The properties of isotropy/anisotropy of the tested are then assessed.

5.2 Experimental Tests

Ten 2 mm thick slices of fresh umbilical cord were cut in the transverse directions of the cord and submitted to equibiaxial tension. An ad hoc metallic frame was designed and realized to perform the experimental tests (see Fig. 5.1).

Twelve nylon wires were fixed on the samples through fishing hooks. On the other extremity, three calibrated weights (20 g, 40 g and 60 g, for each wire) were suspended in succession. Pulleys with a 10 mm diameter were utilized to change the direction of application of the load. By using a timbre, a grating of points with a 500 μm pitch was printed on the samples while a Dalsa CCD camera (1600 × 1360 pixels, 2/3″), equipped with a Fujinon HF9HA-1B objective was utilized to monitor the deformation of the samples. High precision stages were utilized to align the optical axis of the CCD camera with the center of the samples as well as to focus the image of the sample acquired by the camera.

For each of the acquired images, the two-dimensional FFT was computed. The frequency at which the maximum of the energy in the power spectrum is associated, was identified. Band-pass filters were utilized to select the significant harmonics of the signal; inverse FFT was hence applied to return to the physical space. The phase of the pattern was calculated via in-quadrature filtering. For each point of the sample, the difference between the phase computed in the absence of the load and the phase computed with the load applied is determined. By performing the phase unwrapping operation, it is possible to determine the actual values of in-plane displacements experienced by the sample. The sensitivity of the optical set-up is the pitch of the grating printed on the specimen.

Fig. 5.1 Experimental apparatus including the loading frame utilized to perform the equibiaxial tests on the umbilical cord sections

5.3   Experimental Results

Figure 5.2 shows the grating printed on a sample for different levels of applied load ranging from 0 to 60 g. It is interesting to observe how the application of the load produces a modulation of the dots of the grating printed on the sample surface. Based on this modulation, we can determine the in-plane displacement field of the slice along X and Y-directions.

Figure 5.3 shows the phase map computed in the reference configuration (i.e. absence of load) and that computed (for the horizontal displacement u_x) for a load of 20 g.

Figure 5.4 shows the displacement maps measured along directions X and Y, for a load of 20 g applied to all wires. The difference between the spatial domain in which the displacement u_x and u_y was computed derives from the fact that, in order to remove singular regions, different masks were utilized in processing the data.

Similar displacement maps can be obtained for each level of load applied in the experiments. The diagrams in Fig. 5.5 show the distribution of displacements u_x and u_y that were experimentally measured along the control paths X_1–X_2 and Y_1–Y_2 illustrated in the figure.

5.4   Discussion and Conclusions

Intrinsic moiré was utilized to analyze the structural response of 2 mm thick slices of umbilical cord. In particular, the technique was utilized to monitor the in-plane displacement field of slices subject to equibiaxial test. For that purpose, an ad hoc loading frame was designed and realized.

Experimental tests showed that the mechanical behavior of the umbilical cord slices is highly anisotropic. The displacement field along the X direction is significantly different from that measured along the Y direction. The strong anisotropy of the tested samples can be justified with the anisotropic arrangement of the collagen fibers within the Wharton's Jelly as well as with the anisotropic disposal of the vessels within the sample.
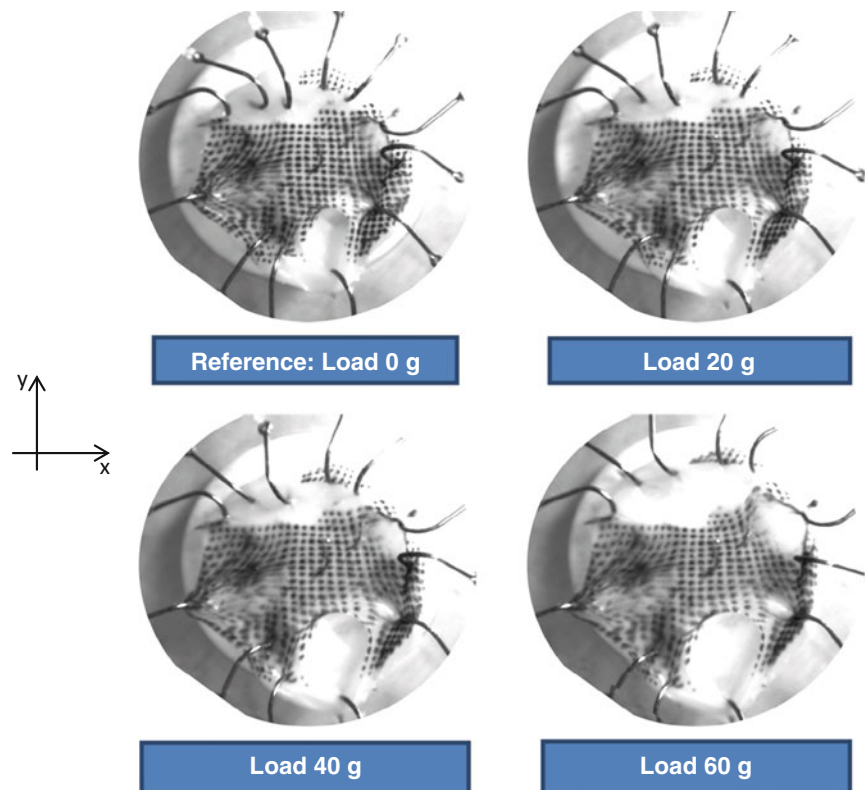
Fig. 5.2 Modulation of the dot grating printed on a slice of umbilical cord subjected to different load levels
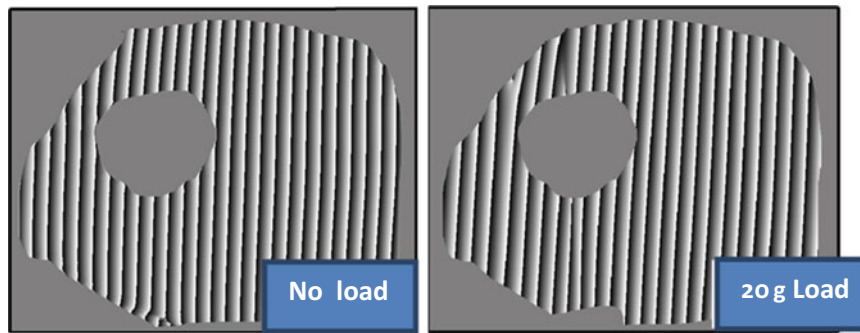
Fig. 5.3 Phase maps computed for the undeformed configuration and in presence of a 20 g load
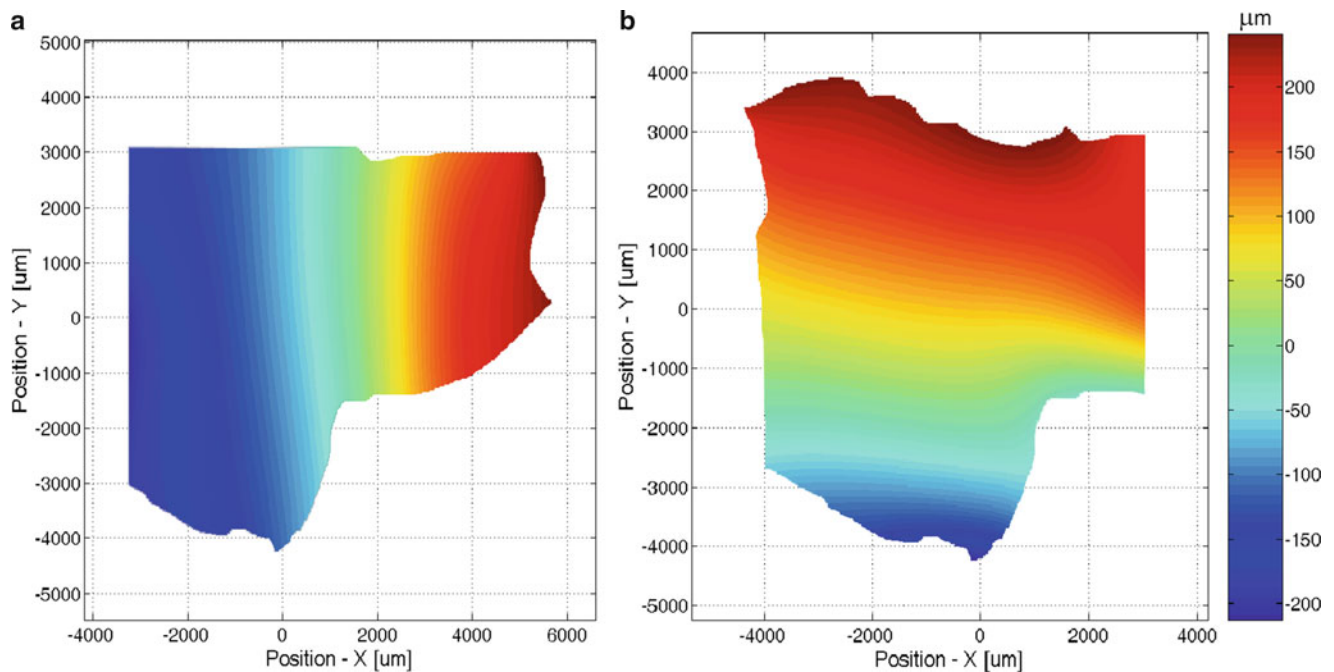
Fig. 5.4 In plane displacement computed along the X (a) and the Y (b) directions in the sample for a load of 20 g

The preliminary investigation described in this paper has pointed out some critical aspects. First, special cares should be taken in positioning samples in the testing apparatus because samples are very small and difficult to handle. Second, it is very difficult to realize a perfectly equibiaxial test condition. In fact, as the load is being applied, micro-movements of the sample may occur that produce fictitious displacements thus altering the acquired displacement field. The third important critical point is represented by the difficulty of printing a grating on the sample. In spite of the care taken in properly drying the sample before to press the timbre against it, different regions of the sample remained without any grating printed on them. This could be due to the fact that the cut of the cord slices produced highly rough surfaces.

In conclusion, the study conducted demonstrated the applicability of the intrinsic moiré for the analysis of the in-plane displacement field of umbilical cord slices. The tests performed showed that the mechanical behavior of umbilical cord is highly anisotropic. Ad hoc additional frames will be developed in the future to better control the application of the loads and then to better monitor the resulting displacement field.
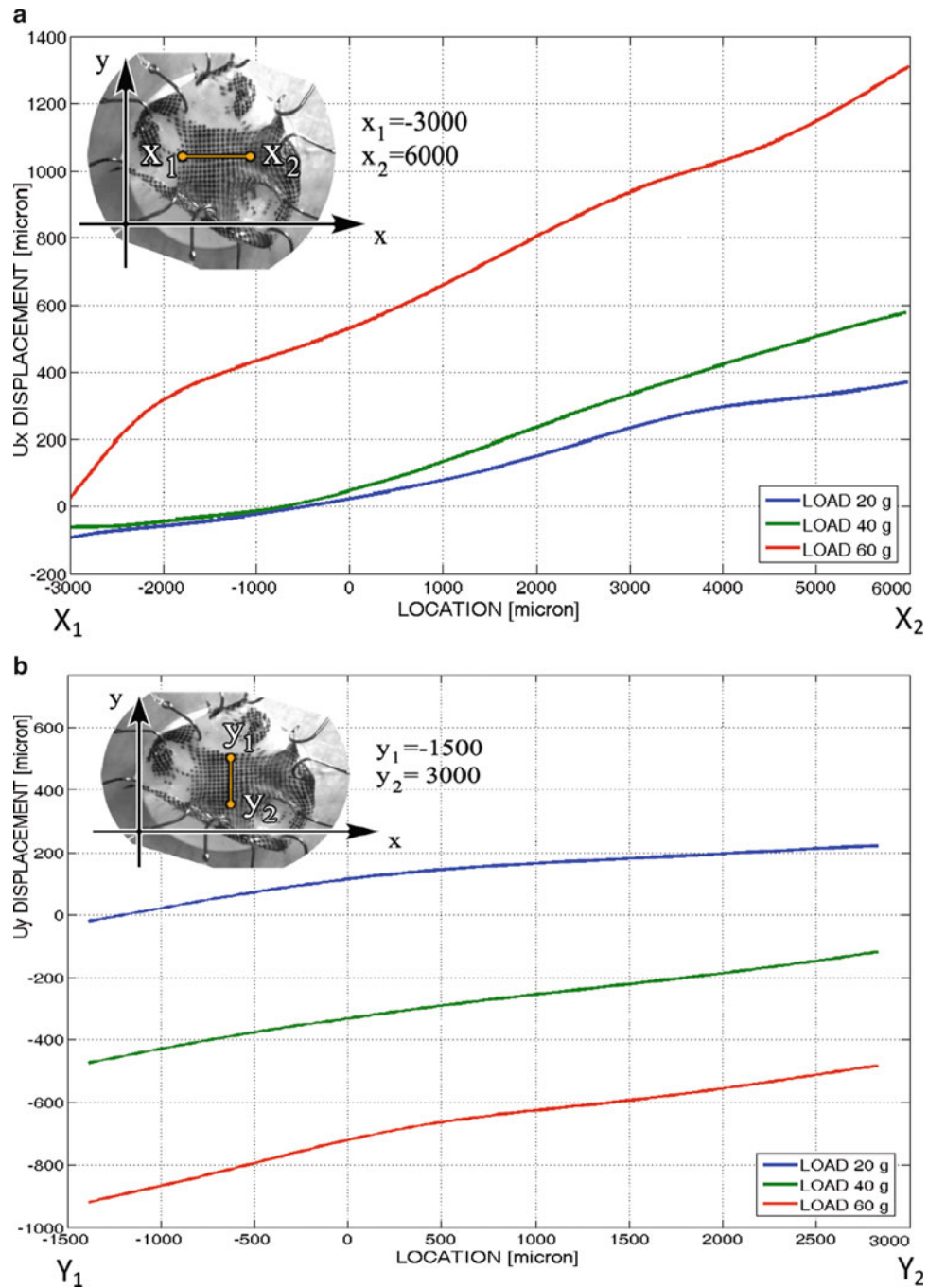
Fig. 5.5 In-plane displacement computed along the indicated paths in the X (a) and the Y (b) directions for different values of applied load

References

1. Boccaccio A, Frassanito MC, Lamberti L, Brunelli R, Maulucci G, Monaci M, Papi M, Pappalettere C, Parasassi T, Sylla L, Ursini F, De Spirito M (2012) Nanoscale characterization of the biomechanical hardening of bovine zona Pellucida. J R Soc Interface 9:2871–2882
2. Boccaccio A, Lamberti L, Papi M, De Spirito M, Douet C, Goudet G, Pappalettere C (2014) A hybrid characterization framework to determine the visco-hyperelastic properties of a porcine zona pellucida. Interface Focus 4: Article No. 20130066
3. Boccaccio A, Papi M, De Spirito M, Lamberti L, Pappalettere C (2013) Effect of the residual stress on soft sample nanoindentation. Appl Phys Lett 102: Article No. 133704.

4. Frassanito MC, Lamberti L, Boccaccio A, Pappalettere C (2011) Discussion on hybrid approach to determination of cell elastic properties. In: Proulx T (ed) Conference proceedings of the society for experimental mechanics series, optical measurements, modelling, and metrology, vol 5. Springer, New York, pp 119–124

5. Sciammarella CA, Lamberti L, Boccaccio A, Sciammarella FM (2011) High precision contouring with the projection moiré and related methods, a review. Strain 47(s2):43–64

6. Sciammarella CA, Lamberti L, Boccaccio A (2008) High precision contouring utilizing the moiré method. In: Proceedings of the 11th SEM international congress and exhibition on experimental and applied mechanics, vol 3, Orlando, FL, June 2008, pp 1495–1510

7. Cosola E, Genovese K, Lamberti L, Pappalettere C (2008) Mechanical characterization of biological membranes with moiré techniques and multi-point simulated annealing. Exp Mech 48:465–478

8. Cosola E, Genovese K, Lamberti L, Pappalettere C (2008) A general framework for identification of hyper-elastic membranes with moiré techniques and multi-point simulated annealing. Int J Solids Struct 45:6074–6099

9. Pennati G (2001) Biomechanical properties of the human umbilical cord. Biorheology 38:355–366

10. Li WC, Ruan XZ, Zhang HM, Zeng YJ (2006) Biomechanical properties of different segments of human umbilical cord vein and its value for clinical application. Journal of Biomedical Materials Research. Part B. Appl Biomater 76:93–97

11. Hamedani BA, Navidbakhsh M, Tafti HA (2012) Comparison between mechanical properties of human saphenous vein and umbilical vein. Biomed Eng Online 11:59

Chapter 6
Study on the Visco-Hyperelastic Behavior of the Zona Pellucida

A. Boccaccio, L. Lamberti, M. Papi, C. Douet, G. Goudet, M. De Spirito, and C. Pappalettere

Abstract   The zona pellucida (ZP) is a specialized extracellular matrix surrounding the developing oocyte. This thick matrix consists of different types of glycoprotein, which have different roles in fertilization. Nowadays several techniques are developed and refined to establish the ZP mechanical response. The assumption at the basis of these methods is that the ZP behaves like an elastic body, dissipative forces are neglected, and thus the Young modulus value remains unaffected by probe dynamics. On the contrary dissipative force are strongly regulated by the slippage of ZP chains past one another whereas the absolute reaction force value is mainly due to the architecture of the ZP structure (number of cross-links and distances between knots). Elastic deflection is then due to the ability of each chain to stretch, whereas viscous flow is caused by the sliding of the molecules over one another. Therefore viscous reaction forces generated by the ZP have to be considered one of the main player in regulating the sperm transit but their peculiar behavior along the ZP structure is still poorly understood. In this context, for the first time, we developed and verified a visco-hyperelastic model able to reproduce the ZP reaction force stressed at different probe rate.

Keywords   Porcine zona pellucida • Visco-hyperelasticity • Atomic force microscopy • Prony series • Finite element analysis • Nonlinear optimization

6.1   Introduction

The zona pellucida (ZP) is the spherical extracellular membrane surrounding mammalian oocyte at fertilization. The ZP is mainly composed of sulfated glycoproteins accomplishing, each, specific functions; glycoproteins are assembled into a three-dimensional network organized according to a specific nano-architecture [1–3]. The principal tasks of the ZP are (1) to regulate binding of sperm to ovulated eggs during fertilization; (2) to prevent polispermy after fertilization; (3) to protect early embryos while they traverse the female reproductive tract. Penetration of ZP membrane by spermatozoa is the crucial step in oocyte fertilization and any inability of spermatozoa to penetrate ZP inevitably leads to infertility [4–6]. A deep knowledge of ZP mechanical properties is hence at the basis of any study on the sperm-ZP interaction.

Techniques recently utilized to measure ZP mechanical properties include micro tactile sensors made of piezoelectric material [7, 8], micropipette aspiration [9, 10], micropipette indentation [11], vision-based nanoforce estimation [12] and Atomic Force Microscopy (AFM) [13–17]. While other techniques can characterize the mechanical behaviour of the entire oocyte as they induce macroscopic deformations of the whole cell, AFM allows nanoscale investigations to be carried out in specific regions of the cell specimen thus detecting at very high resolution any source of inhomogeneity and/or anisotropy in mechanical response. Furthermore, AFM can operate in liquid environment, that is in physiological conditions reproduced in vitro.

A. Boccaccio • L. Lamberti • C. Pappalettere
Dipartimento di Meccanica, Matematica e Management, Politecnico di Bari, Bari 70126, Italy

M. Papi (✉) • M. De Spirito
Istituto di Fisica, Università Cattolica del Sacro Cuore, Roma 00168, Italy
e-mail: m.papi@rm.unicatt.it

C. Douet • G. Goudet
UMR 85, Physiologie de la Reproduction et des Comportements, INRA-CNRS-Université de Tours,
IFCE, 37380 Nouzilly, France

AFM was utilized to evaluate mechanical properties of immature, mature and fertilized bovine oocyte ZPs [13, 14], also trying to assess the effect of ZP structure inhomogeneity across membrane thickness [15, 16], and the role played by residual stresses that develop in the tested ZP membrane (that is originally spherical shaped) when specimens are spread out on a rigid substrate for nanoindentation measurements [17].

In spite of the fact that the structural response of polymeric networks like those contained in the ZP membrane is strongly affected by viscous/dissipative effects due to the relative slippage of polymeric chains [18, 19], the typical assumption made in the determination of ZP mechanical properties via AFM is that the cell membrane behaves as a linearly elastic or hyperelastic material and viscous forces are negligible. However, viscous reaction forces generated by ZP should be considered one of the main factors regulating the sperm transit.

If viscous behaviour is neglected, the elastic modulus should actually be considered as an "apparent elastic modulus" because it does not account for nanoindentation rate. A recent study carried out by some of the present authors [20] has demonstrated that viscous forces play a role of crucial importance in the mechanical response of porcine and equine zonae pellucidae submitted to AFM nanoindentation. From the physiologic point of view, such a result indicates that for spermatozoa with excessive motility the reaction force of the ZP can be so large that the membrane may turn in impenetrable barrier.

However, peculiar behaviour of viscous forces along the ZP structure remains poorly understood. In order to overcome this limitation, this study will present for the first time in literature a visco-hyperelastic model able to reproduce the ZP reaction forces stressed by nanoindentation at different probe rates. This approach can provide some insight on the sperm-ZP interactions occurring in the fertilization process.

In a previous study conducted by some of the present authors [16], a hybrid mechanical characterization procedure combining AFM nanoindentation measurements, nonlinear finite element analysis and nonlinear optimization was developed to determine the ZP hyperelastic parameters of immature, mature and fertilized bovine oocytes. This approach was successfully utilized to characterize both inner and outer sides of ZP specimens extracted from the fertilized oocyte. In the present study, the characterization framework is further developed in order to include viscous effects. AFM measurements are conducted on the ZP of immature porcine oocytes at different indentation rate. By combining experimental data with a transient FE model reproducing the nanoindentation measurement and carrying out nonlinear optimization, it is possible to determine the visco-hyperelastic parameters of the porcine zona pellucida. Remarkably, the proposed approach allows hyperelastic behaviour and viscous effects to be clearly separated. Furthermore, it is possible to evaluate the error made in the determination of ZP mechanical properties if viscous effects were not considered.

6.2 Materials and Methods

6.2.1 Preparation of ZP Samples

Porcine immature cumulus-oocyte complexes (COCs) were collected from slaughtered gilts. Ovaries were obtained in local slaughterhouses, immediately after females were killed, and transported to the laboratory within 2 h in 0.9 % NaCl at 37 °C. COCs were collected by aspiration of the follicles with a 16-gauge needle mounted on an aspiration pump and recovered under a stereomicroscope. COCs were washed in Dulbecco's phosphate buffered saline solution (DPBS, Dulbecco A, Paris, France). Before each experiment, porcine and equine COCs were flushed in and out of a glass pipette in DPBS to remove cumulus cells. The zona pellucida was then mechanically removed from the oocytes, washed in DPBS and water, and deposited onto a polylisine coated slide.

6.2.2 Atomic Force Spectroscopy Measurements

In order to evaluate changes in the ZP apparent Young modulus caused by different indentation rates, mechanical response of ZP specimens was analyzed by means of an atomic force microscope (NanowizardII, JPK) combined with an optical microscope (Axio Observer, Zeiss). The AFM probe consisted in an ultrasharp silicon nitride cantilever of calibrated force constant, with a tip radius of less than 10 nm (CSC16, MikroMash).

All the samples were kept in the buffer at the constant temperature of 37 °C throughout data acquisition. In order to carry out mechanical measurements, the exposed AFM tip was lowered onto the ZP surface at different rates, 0.5, 2, 4, 6, 8 and 10 μm/s, respectively: this range of indentation rate is comparable with the typical speed of spermatozoa observed during "in vitro"

Fig. 6.1 Finite element
model utilized to determine
the visco-hyperelastic
properties of the ZP
of porcine oocytes submitted
to nanoindentation
measurements
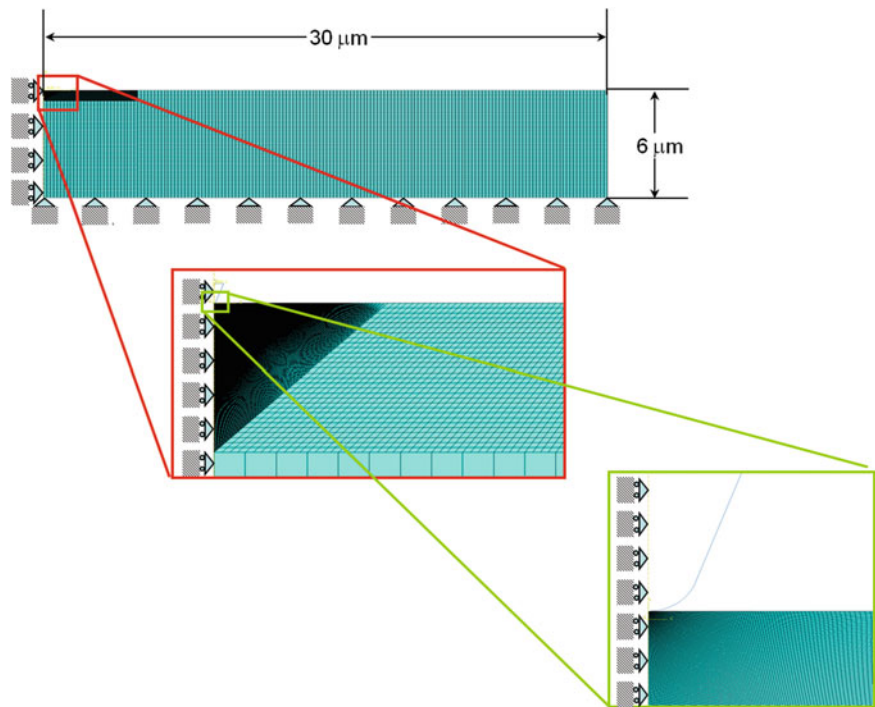
fertilization [21]. Following contact, the AFM tip exerts a force against the ZP, proportionate to cantilever deflection. The cantilever deflection Δ was recorded as a function of the piezoelectric translator position z. The reaction force of the membrane F was determined from the Hooke's law $F = K_C \Delta$ and the indentation δ was calculated as $\delta = z - \Delta$. The cantilever spring constant K_C, nominally 0.01 N/m, was computed for each measurement by thermal calibration and the half opening angle of tip apex $\alpha = 20°$ was accurately determined by means of scanning electron microscopy.

In order to obtain statistically significant results, for each indentation rate, force-indentation curves were recorded at 50 different points of each analysed sample and were repeated on 10 oocytes.

6.2.3   Finite Element Modelling of the Nanoindentation Measurements on Porcine ZPs

The finite element model of the ZP membrane and the AFM indenter was built by utilizing the general purpose finite element software ABAQUS® Version 6.12 (Dassault Systèmes, France). Given the symmetry of the problem, an axisymmetric FE model was developed. Since the ZP membrane is much softer than the AFM probe, the indenter was modelled as a rigid blunt cone with tip radius of 10 nm and angle of aperture of 40°. The ZP membrane was instead modelled as an incompressible visco-hyperelastic slab, 30 μm long and 6 μm thick (the value of thickness is consistent with data reported in [22] for porcine oocytes). The finite element model simulating the AFM nanoindentation process is shown in Fig. 6.1.

The four-node hybrid axisymmetric CAX4H element available in the ABAQUS library was utilized to model the ZP slab. The mesh of the ZP membrane was properly refined in the contact region (see details in Fig. 6.1). The very small dimensions of the AFM indenter with respect to those of the ZP membrane along with the necessity of having small enough and regularly shaped elements in the contact region (this must be done to facilitate convergence of FE analysis) have led to divide the ZP slab domain in regions of different mesh density. The average element size is 0.06 nm in the contact region and about 180 nm in regions far from the AFM indenter. By utilizing these element sizes and the above described partition strategy, a good compromise between convergence of nonlinear analysis and computation time could be reached.

The nodes of the ZP bottom edge were clamped while symmetry constraints preventing only horizontal displacements were applied to the nodes of the left edge of the FE model. A vertical displacement of $\delta = 100$ nm was imposed to the AFM indenter. The contact between the indenter and the membrane was assumed to be frictionless and the "hard contact" option (i.e. surfaces come into contact when their gap reduces to zero; any contact pressure can be transmitted through

surfaces in contact) available in ABAQUS was selected. Finite element analysis accounted for geometric nonlinearity (i.e. large deformations) and time-dependent behaviour (i.e. viscous effects).

6.2.4 Visco-Hyperelastic Constitutive Model of the Porcine ZP

The porcine ZP membrane investigated in the study was modelled as a visco-hyperelastic material following the Arruda-Boyce constitutive law [23]. The AB hyperelastic model, originally developed to reproduce the mechanical behaviour of polymer chain networks, was previously utilized in literature to analyze biospecimens including filamentous collagen networks [24, 25], and has recently been found to describe very well the mechanical response of bovine ZPs at different stages of the fertilization process [16]. The Arruda-Boyce constitutive model requires that two constants are given in input to ABAQUS: the shear modulus μ_{8chain} and the distensibility λ_L. The strain energy function can be expressed as (6.1):

$$W = \mu_{8chain}\left[\frac{1}{2}\left(\overline{I}_1 - 3\right) + \frac{2}{20\lambda_L^2}\left(\overline{I}_1^2 - 9\right) + \frac{33}{1,050\lambda_L^4}\left(\overline{I}_1^3 - 27\right) + \frac{76}{7,000\lambda_L^6}\left(\overline{I}_1^4 - 81\right) + \frac{519}{673,750\lambda_L^8}\left(\overline{I}_1^5 - 243\right)\right]. \quad (6.1)$$

where $\overline{I}_1 = \text{tr}[C]$ is the first strain invariant and [C] is the Cauchy-Green strain tensor. The Young modulus of the material can be determined as $E = 2(1 + \nu)\mu_{8chain}$ where ν is the Poisson ratio. By setting $\nu = 0.5$ to account for material incompressibility, it follows $E = 63 \cdot \mu_{8chain}$.

The viscous behaviour of the ZP membrane (i.e. the rate-dependent part of mechanical response) is defined by a N-terms Prony series expansion of the dimensionless relaxation modulus. The effective relaxation modulus $\mu_{8chain}^R(t)$ is obtained as the product of the instantaneous elastic modulus (in the case of the Arruda-Boyce hyperelastic constitutive model, the shear modulus μ_{8chain}) by the dimensionless relaxation function. That is:

$$\mu_{8chain}^R(t) = \mu_{8chain} \cdot \left(1 - \sum_{k=1}^{N} g_k \cdot \left(1 - e^{-\frac{t}{\tau_k}}\right)\right) \quad (6.2)$$

where g_k is the kth Prony constant ($k = 1, 2, \ldots, N$) and τ_k is the corresponding relaxation time constant, respectively.

The number of terms included in the Prony expansion was limited to 1 to minimize the computational cost of the finite element analyses entailed by the ZP characterization process. This choice was supported by the fact that preliminary numerical tests carried out for the highest indentation rate considered in this study revealed that Prony expansion is largely dominated by the first term of the series.

The Prony series expansion (6.2) was not implemented for the distensibility parameter λ_L basically for two reasons. In the first place, the structural response of ZP membranes is mostly driven by the shear modulus (see the findings of [16] for the bovine oocytes). Second, the viscous response of a polymeric network like that included in the ZP membrane depends on the relative slippage of filaments [18, 19] rather than on material distensibility which instead determines the ability of filaments to stretch and hence affects the overall stiffness of the material. The latter assumption sounds very logical as distensibility does not change much through the material.

Time dependent finite element analyses were carried out for each indentation rate v_i (i.e. 0.5, 2, 4, 6, 8 and 10 μm/s, respectively) considered in the experimental tests. The total duration of the indentation process t_{step_i}, corresponding to the analysis time specified in ABAQUS, hence is $t_{step\_i} = \delta/v_i$. For each indentation rate, the displacement given to the AFM indenter is ramped over the time step t_{step_i}.

6.2.5 Optimization-Based Algorithm for Extracting ZP Visco-Hyperelastic Properties

The visco-hyperelastic properties of the ZP membrane were determined with the optimization-based identification algorithm schematized in Fig. 6.2. A similar approach was utilized in other mechanical characterization studies of nonlinear materials at the micro and nanoscale [26–28, 16].

The identification algorithm, coded in the MATLAB® (The Mathworks Inc., Austin, TX, USA) software environment, minimizes the difference between nanoindentation data and finite element analyses via nonlinear optimization. The flow chart shown in Fig. 6.2 includes several steps. Initial values are assigned to the four unknown material parameters $\mu_{8chain}°$, $\lambda_L°$, $g_1°$ and $\tau_1°$ included as optimization variables (Step 1). An input file containing data on model geometry and material properties is
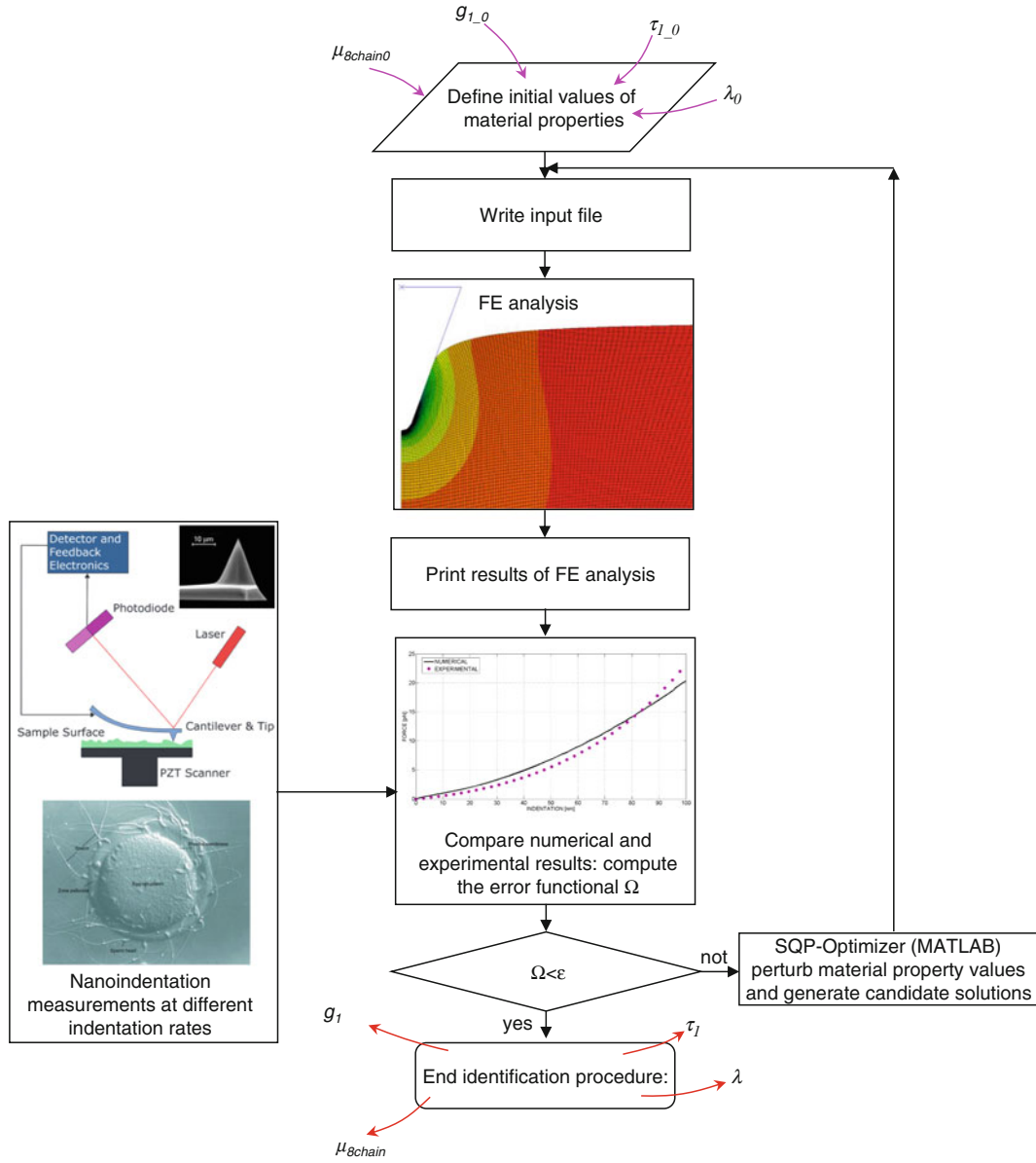
Fig. 6.2 Flow chart of the optimization-based algorithm implemented in MATLAB to identify ZP's mechanical properties

prepared (Step 2) and given in input to the ABAQUS finite element solver (Step 3). ABAQUS saves the results of FE analysis in another file (Step 4) and the computed force-indentation curve is compared with its counterpart measured experimentally (in the schematic of Fig. 6.2, AFM measurements carried out at a given indentation rate correspond to Step 0) in order to evaluate the difference between numerical results and experimental data (Step 5). This leads to formulate an optimization problem including the unknown material properties as design variables where the objective is to minimize the error functional Ω:

$$
\begin{cases}
\text{Min} \left[\Omega(\mu_{8\text{chain}}, \lambda_L, g_1, \tau_1) = \sqrt{\dfrac{1}{N_{\text{CNT}}} \sum_{j=1}^{N_{\text{CNT}}} \left(\dfrac{F_{\text{FEM}}{}^j - \overline{F^j}}{\overline{F^j}} \right)^2} \right] \\
0.1\,\text{kPa} \leq \mu_{8\text{chain}} \leq 100\,\text{kPa} \\
1 \leq \lambda_L \leq 10 \\
0.001 \leq g_1 \leq 1 \\
10^{-5} \leq \tau_1 \leq 10^{-1}
\end{cases}
\tag{6.3}
$$

As mentioned above, there are four unknown material properties: two hyperelastic constants μ_{8chain} and λ_L, and two viscous parameters g_1 and τ_1. The bounds of the shear modulus μ_{8chain} were chosen so as to cover the range of variation of the apparent Young modulus of porcine oocytes reported in [20] for the same range of indentation rate (i.e. the apparent Young modulus varied from 30 to 160 kPa, which corresponds to have the shear modulus varying between 5 and 26.7 kPa), and to account for the fact that values of elastic moduli determined under the assumption of hyperelastic behaviour may be up to one or two orders of magnitude smaller than those determined with the classical Hertzian model which assumes instead linear elasticity (see, for example, [16, 29]). In general, material property bounds were set large enough to increase design freedom and rapidly converge to global optimum.

In Eq. (6.3), $F_{FEM}{}^j$ and $\overline{F^j}$, respectively, are the indentation force values for the jth load step computed by ABAQUS and those measured experimentally by AFM. The number of control points N_{CNT} corresponds to the number of load steps executed for completing nonlinear FE analysis.

The computed error functional Ω is compared with a pre-defined convergence limit ε_{CONV} set as 0.0001 (i.e. 0.1 %) (Step 6). If it occurs $\Omega < \varepsilon_{CONV}$, the identification process terminates and material properties are listed in output (Step 8). Conversely, if it occurs $\Omega > \varepsilon_{CONV}$, material parameters μ_{8chain}, λ_L, g_1 and τ_1 are perturbed in the subsequent design cycles until convergence (Step 7).

The above described identification process was carried out to determine the corresponding material properties μ_{8chain}, λ_L, g_1 and τ_1 for each indentation rate. The inverse problem stated in (6.3) was solved with the Sequential Quadratic Programming (SQP) method, a globally convergent gradient-based optimization algorithm, coded in the MATLAB® Version 7.0 commercial software. For each indentation rate, optimization runs were started from five different sets of material properties randomly generated. This multi-start optimization strategy together with the large range of variability chosen for material parameters allowed to cover the whole search space and increased the probability of finding the global optimum. Serial optimization runs were carried out for each test case to avoid premature convergence. This process ended as soon as relative variations of error functional $|(\Omega_K - \Omega_{K-1})/\Omega_{K-1}|$ and design vector $\|\mathbf{X}_K - \mathbf{X}_{K-1}\|/\|\mathbf{X}_{K-1}\|$ between the last two serial runs became smaller than 0.0001 (i.e. 0.1 %).

It should be noted that the common hypothesis behind most of the studies reported in literature is that the viscous forces involved in the nanoindentation process are negligible. Further optimization runs were hence carried out in order to evaluate the error on hyperelastic material properties that would be made if viscous effects were neglected. For that purpose, only the hyperelastic parameters μ_{8chain} and λ_L were identified, without including any Prony term, and time-independent hyperelastic finite element analyses were executed. The "static" values of μ_{8chain} and λ_L minimizing the error functional (6.3) for each force-indentation curve relative to a given indentation rate were then compared with the corresponding "dynamic" values determined by including viscous effects. This allowed to assess the limit of validity as well as the practical implications of the assumption of time-independent behaviour usually followed in literature.

6.3   Results

Table 6.1 presents the values of visco-hyperelastic properties determined with the hybrid mechanical characterization procedure described in Sect. 6.2 for the $0.5 \div 10$ μm/s indentation rate range considered in this study. The table reports also the average values of identified properties and the corresponding standard deviations over the indentation rate range as well as the values of the correlation coefficient R^2 between AFM data and force-indentation curves reconstructed numerically for each indentation rate.

It can be seen that values of visco-hyperelastic properties were rather insensitive to indentation rate. In fact, the ratio of the standard deviation on identified parameters to the average material property values was 9.06 %, 0.845 %, 3.32 % and 8.87 %, respectively, for shear modulus, distensibility, Prony constant and relaxation time. Such a level of statistical dispersion on identified material properties is due to the fact that the target force-indentation curve taken for each indentation rate was the average of the corresponding force–indentation curves recorded experimentally in 50 different points of 10 samples. In fact, the standard deviation of experimental data with respect to the target data used in the characterization process was on average about 10 %, hence consistent with the values of standard deviation found on identified material properties.

Figure 6.3 compares the average force-indentation curves measured experimentally with those reconstructed numerically by giving in input to ABAQUS the visco-hyperelastic properties identified via nonlinear optimization. The overall quality of data fitting was always good over the entire range of indentation rates: in fact, the correlation coefficient ranged between 0.9 and 0.975 with an average value of 0.939. Interestingly, no trend of variation of the data correlation coefficient with respect to the indentation rate could be found.

Table 6.1 Visco-hyperelastic properties of the porcine zona pellucida determined with the hybrid characterization procedure

Indentation rate (μm/s)	μ_{8chain} [kPa]	λ_L	g_1	τ_1 [s]	R^2
0.5	1.406	3.001	0.9031	0.00109	0.975
2	1.686	2.962	0.9742	0.00134	0.900
4	1.692	2.960	0.9797	0.00136	0.925
6	1.835	2.927	0.9819	0.00140	0.920
8	1.816	2.941	0.9848	0.00140	0.950
10	1.694	2.959	0.9849	0.00137	0.963
Average ± STD	1.688 ± 0.153	2.958 ± 0.0250	0.9681 ± 0.0321	0.00133 ± 0.000118	

The values listed in the table were obtained by taking the force-indentation curves averaged over AFM measurements done for each indentation rate as targets
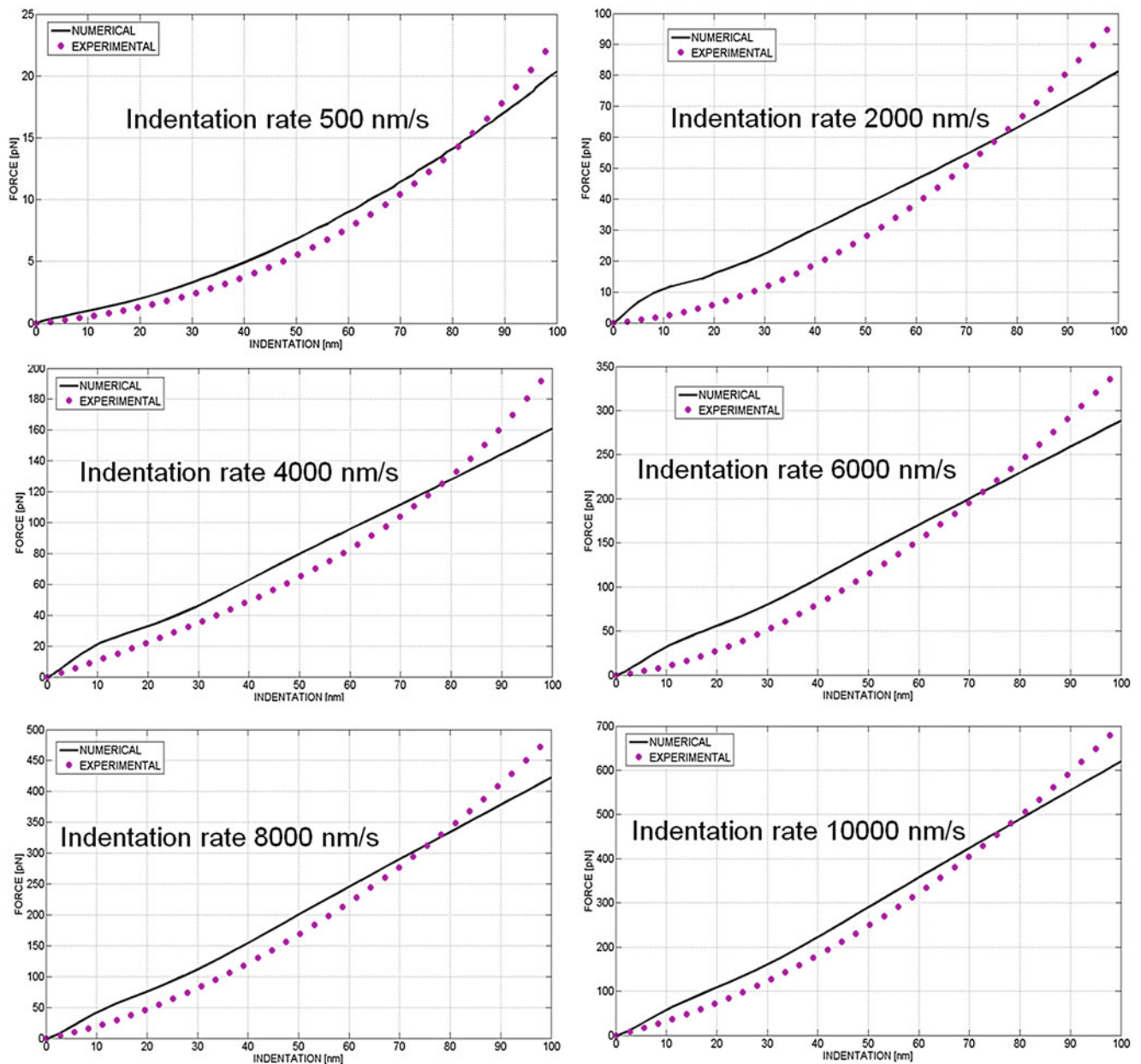

Fig. 6.3 Force Comparison of force-indentation curves experimentally measured by AFM with their counterpart "optimized" by ABAQUS for the different indentation rates

The time-independent hyperelastic parameters determined for the very low indentation rate of 0.5 μm/s without considering viscous effects are μ_{8chain} = 1.415 kPa and λ_L = 2.988, practically the same as the corresponding values determined with the visco-hyperelastic model, μ_{8chain} = 1.406 kPa and λ_L = 3.001.

6.4 Conclusions

This study presented the first visco-hyperelastic model ever developed in literature to describe the viscous response of the porcine zona pellucida. The purpose of the study was to analyze the variation of viscous reaction forces exerted by the porcine ZP under different conditions of sperm motility. Since these forces are one of the main factors regulating the sperm transit, their detailed knowledge must be the start point of any study on the sperm-ZP interaction. In fact, even small variations of spermatozoa velocity may lead to significant changes in the ZP reaction force turning the ZP shell into an impenetrable barrier and leading to fertility impairments. Variations of sperm motility were simulated by carrying out AFM nanoindentation tests at different rates between 0.5 μm/s and 10 μm/s, the typical range adopted in in vitro fertilization. The force exerted by the AFM probe on the ZP membrane simulates the interaction between the spermatozoa and the ZP itself.

Although visco-elastic behaviour of human and animal cells including oocytes was extensively analyzed in literature (see, for example, [30–35, 11, 12, 20]), no systematic study was ever carried out on the visco-hyperelastic behaviour of ZP membranes. However, hyperelasticity is the best constitutive model for assessing the mechanical response of ZP in the different stages of fertilization process [16], much better than the Hertzian model usually considered in literature. Nonlinear elasticity and more specifically hyperelasticity should always be considered in modelling soft materials made of polymeric chain networks such as the ZP membrane. Very recent studies have correctly pointed out the viscous behaviour of mouse [11, 12], porcine and equine ZPs [20] but provided just qualitative information on material properties that were determined via standard data fitting based on the classical Hertzian model or more sophisticate contact mechanics models.

This study attempted to overcome the above mentioned limitation trying to separate the time-independent part of the ZP's mechanical response, which is driven by the ZP polymeric network structure (i.e. number of cross-links and distances between knots), from the viscous part (i.e. time-dependent), which is driven instead by the slippage of ZP polymeric chains one past another. The visco-hyperelastic properties of the porcine ZP membrane were determined by combining atomic force microscopy measurements and optimization-based inverse finite element analysis. The hyperelastic behaviour of the ZP membrane was described by the Arruda-Boyce constitutive model while the viscous behaviour (i.e. the rate-dependent part of material behaviour) was described by means of a Prony series expansion of the dimensionless relaxation modulus.

A gradient-based optimization algorithm minimizing the difference between the force-indentation curve measured experimentally by AFM and its counterpart computed by ABAQUS from the visco-hyperelastic model allowed the hyperelastic material parameters, μ_{8chain} and λ_L, and the viscous terms, g_1 and τ_1, to be determined for each indentation rate. Remarkably, all material properties changed by less than 10 % (see Table 6.1) in spite of the fact that indentation rate changed by 20 times (i.e. from 0.5 to 10 μm/s). This was somehow expected because the visco-hyperelastic parameters describing the "static" and "dynamic" parts of the ZP response are material properties inherent to the investigated material and should be independent of the velocity at which spermatozoa impact on the ZP membrane. However, such an aspect is not always correctly assessed in literature where most of the studies on cell visco-elasticity assume a given impact velocity. The standard deviation observed on material properties actually derives from the fact that a very large amount of experimental data were gathered and average force-indentation curves were taken as target data in the optimization-based identification process.

Correlation between experimental data and force-indentation curves computed by ABAQUS for the material parameter values determined via optimization was good for each indentation rate, ranging between 0.9 and 0.975 (see Table 6.1 and Fig. 6.3). Since no clear trend of variation between the correlation coefficient R^2 and the indentation rate could be established, it appears that the proposed identification framework is robust and able to correctly describe the real mechanical behaviour of the porcine ZP.

Errors made in the determination of material properties if viscous terms were neglected can be evaluated as well with the present approach. Hyperelastic properties obtained with the "static" hyperelastic model without considering viscous effects (i.e. by including only the shear modulus and the distensibility as design variables) were very close to those obtained from the visco-hyperelastic model for the lowest indentation rate of 0.5 μm/s: 1.415 vs. 1.406 kPa and λ_L = 2.988 vs. 3.001 for, respectively, shear modulus and distensibility. This sounds logical because at low indentation rates the structural response of ZP is mostly driven by the elastic deformations that depend on the structure of the ZP's polymeric network (i.e. number of cross-links and distances between knots) and the ability of polymeric chains to stretch. As the indentation rate increases, viscous effects dominate and neglecting them leads to significant errors (more than one order of magnitude, as it can be seen

from Table 6.1) on the level of material stiffness: in the present case, the apparent Young modulus determined without considering viscous effects ranged from ~4 to ~94 kPa (those values are in good agreement with the findings of [20]) vs. the 5.06 ± 0.459 kPa values found with the visco-hyperelastic model.

An interesting issue emerging from this study is the following. The correlation coefficient R^2 evaluated for the "static" material properties was on average 0.97, hence higher than the 0.939 value obtained in the "dynamic" case where visco-elastic effects were considered. However, the elastic modulus determined in the static case represents only an "apparent" stiffness because the real stiffness is an intrinsic property of the material that, in the case of a polymeric network, depends on the network structure and on the ability of chains to deform independently one from the other. Such characteristics are not altered by the loading velocity. Another implication would be that from the point of view of the instrumentation utilized to characterize biomaterials, the fact that the structural response of the ZP membrane is rate dependent, leads to conclude that is practically impossible to compare the mechanical properties derived from different measurement instruments that work at different speeds. This statement is totally incorrect as material properties cannot depend on testing equipments and conditions.

A number of assumptions and simplifications were made in the present study. First, residual stresses developing in the ZP membrane specimens spread out on the rigid substrate for indentation were not considered. However, since the AFM tip utilized in the experimental tests is very sharp, residual stresses do not affect significantly the determination of material properties [17]. Second, a three-dimensional model of the ZP membrane and the indenter would have better reproduced the real nanoindentation process. However, axisymmetric finite element models are very common in literature and allow a more detailed description of the contact problem between the AFM tip and the ZP membrane. Third, the ZP-indenter contact was assumed to be frictionless. However, since the influence of friction in the extraction of elastic properties becomes significant if the indentation depth reaches half of sample thickness and that the indentation depth simulated in the FE model represents only 1.6 % of the total sample thickness, one can reasonably conclude that the error committed in neglecting friction is very small.

References

1. Gwatkin RBL (1977) Fertilization mechanisms in man and mammals. Plenum, New York, NY
2. Yanagimachi R (1994) Mammalian fertilization. In: Knobil E, Neill JD (eds) Physiology of reproduction, vol 1, 2nd edn. Raven Press, New York, NY, pp 189–318
3. Florman HM, Ducibella T (2006) Fertilization in mammals. In: Neill JD (ed) Physiology of reproduction, vol 1. Elsevier, San Diego, CA, pp 55–112
4. Greve JM, Wassarman PM (1985) Mouse egg extracellular coat is a matrix of interconnected filaments possessing a structural repeat. J Mol Biol 181:253–264. doi:10.1016/0022-2836(85)90089-0
5. Wassarman PM (1988) Zona pellucida glycoproteins. Annu Rev Biochem 57:415–442. doi:10.1146/annurev.bi.57.070188.002215
6. Green DPL (1997) Three dimensional structure of the zona pellucida. Rev Reprod 2:147–156. doi:10.1530/ror.0.0020147
7. Murayama Y, Constantinou CE, Omata S (2004) Micromechanical sensing platform for the characterization of the elastic properties of the ovum via uniaxial measurement. J Biomech 37:67–72. doi:10.1016/S0021-9290(03)00242-2
8. Murayama Y, Yoshida M, Mizuno J, Nakamura H, Inoue S, Watanabe Y, Akaishi K, Inui H, Constantinou CE, Omata S (2008) Elasticity measurement of zona pellucid using a micro tactile sensor to evaluate embryo quality. J Mamm Ova Res 25:8–16. doi:10.1274/jmor.25.8
9. Khalilian M, Navidbakhsh M, Valojerdi MR, Chizari M, Yazdi PE (2010) Estimating Young's modulus of zona pellucida by micropipette aspiration in combination with theoretical models of ovum. J R Soc Interface 7:687–694. doi:10.1098/rsif.2009.0380
10. Khalilian M, Navidbakhsh M, Valojerdi MR, Chizari M, Yazdi PE (2011) Alteration in the mechanical properties of human ovum zona pellucida following fertilization: experimental and analytical studies. Exp Mech 51:175–182. doi:10.1007/s11340-010-9357-z
11. Liu X, Shi J, Zong Z, Wan KT, Sun Y (2012) Elastic and viscoelastic characterization of mouse oocytes using micropipette indentation. Ann Biomed Eng 40:2122–2130. doi:10.1007/s10439-012-0595-3
12. Kim J, Kim J (2013) Viscoleastic characterization of mouse zona pellucida. IEEE Trans Biomed Eng 60:569–575. doi:10.1109/TBME.2012.2230444
13. Papi M, Sylla L, Parasassi T, Brunelli R, Monaci M, Maulucci G, Missori M, Arcovito G, Ursini F, De Spirito M (2009) Evidence of elastic to plastic transition in the zona pellucida of oocytes using atomic force spectroscopy. Appl Phys Lett, 94 Paper No. 153902 (3 p). doi:10.1063/1.3107265
14. Papi M, Brunelli R, Sylla L, Parasassi T, Monaci M, Maulucci G, Missori M, Arcovito G, Ursini F, De Spirito M (2010) Mechanical properties of zona pellucida hardening. Eur Biophys J 39:987–992. doi:10.1007/s00249-009-0468-3
15. Papi M, Brunelli R, Familiari G, Frassanito MC, Lamberti L, Maulucci G, Monaci M, Pappalettere C, Parasassi T, Relucenti M, Ursini F, Sylla L, De Spirito M (2012) Whole-depth change in bovine zona pellucida biomechanics after fertilization: how relevant in hindering polyspermy? PLoS One 7: Paper No. e45696 (7 p). doi:10.1371/journal.pone.0045696
16. Boccaccio A, Frassanito MC, Lamberti L, Brunelli R, Maulucci G, Monaci M, Papi M, Pappalettere C, Parasassi T, Sylla L, Ursini F, De Spirito M (2012) Nanoscale characterization of the biomechanical hardening of bovine zona pellucida. J R Soc Interface 9:2871–2882. doi:10.1098/rsif.2012.0269

17. Boccaccio A, Papi M, De Spirito M, Lamberti L, Pappalettere C (2013) Effect of the residual stress on soft sample nanoindentation. Appl Phys Lett 102: Art. No. 133704 (4 p). doi:10.1063/1.4801428

18. Claessens MMAE, Tharmann R, Kroy K, Bausc AR (2006) Microstructure and viscoelasticity of confined semiflexible polymer networks. Nat Phys 2:186–189. doi:10.1038/nphys241

19. Papi M, Arcovito G, De Spirito M, Vassalli M, Tiribilli B (2006) Fluid viscosity determination by means of uncalibrated atomic force microscopy cantilevers. Appl Phys Lett 88: Art. No. 194102 (3 p). doi:10.1063/1.2200588

20. Papi M, Maiorana A, Douet C, Maulucci G, Parasassi T, Brunelli R, Goudet G, De Spirito M (2013) Viscous forces are predominant in the zona pellucida mechanical resistance. Appl Phys Lett 102: Art. No. 043703 (4 p). doi:10.1063/1.4789503

21. Sallam HN, Fathy E, Sallam A, Agameya AF, Farrag A (2003) Sperm velocity and morphology, female characteristics, and the hypo-osmotic swelling test as predictors of fertilization potential: Experience from the IVF model. Int J Fertil Womens Med 48:88–95

22. Funahashi H, Ekwall H, Rodriguez-Martinez H (2000) Zona reaction in porcine oocytes fertilized in vivo and in vitro as seen with scanning electron microscopy. Biol Reprod 63:1437–1442. doi:10.1095/biolreprod63.5.1437

23. Arruda EM, Boyce MC (1993) A three dimensional constitutive model for the large stretch behavior of rubber elastic materials. J Mech Phys Solids 41:389–412. doi:10.1016/0022-5096(93)90013-6

24. Bischoff JE, Arruda EM, Grosh K (2000) Finite element modeling of human skin using an isotropic, nonlinear elastic constitutive model. J Biomech 33:645–652. doi:10.1016/S0021-9290(00)00018-X

25. Palmer JS, Boyce MC (2008) Constitutive modeling of the stress–strain behavior of F-actin filament networks. Acta Biomater 4:597–612. doi:10.1016/j.actbio.2007.12.007

26. Cosola E, Genovese K, Lamberti L, Pappalettere C (2008) Mechanical characterization of biological membranes with moiré techniques and multi-point simulated annealing. Exp Mech 48:465–478. doi:10.1007/s11340-008-9135-3

27. Cosola E, Genovese K, Lamberti L, Pappalettere C (2008) A general framework for identification of hyper-elastic membranes with moiré techniques and multi-point simulated annealing. Int J Solids Struct 45:6074–6099. doi:10.1016/j.ijsolstr.2008.07.019

28. Liu K, Van Landingham MR, Ovaert TC (2009) Mechanical characterization of soft viscoelastic gels via indentation and optimization-based inverse finite element analysis. J Mech Behav Biomed Mater 2:355–363. doi:10.1016/j.jmbbm.2008.12.001

29. Lin DC, Shreiber DI, Dimitriadis EK, Horkay F (2009) Spherical indentation of soft matter beyond the Hertzian regime: numerical and experimental validation of hyperelastic models. Biomech Model Mechanobiol 8:345–358. doi:10.1007/s10237-008-0139-9

30. Guilak F, Tedrow JR, Burgkart R (2000) Viscoelastic properties of the cell nucleus. Biochem Biophys Res Commun 269:781–786. doi:10.1006/bbrc.2000.2360

31. Mathur AB, Collinsworth AM, Reichert WM, Kraus WE, Truskey GA (2001) Endothelial, cardiac muscle and skeletal muscle exhibit different viscous and elastic properties as determined by atomic force microscopy. J Biomech 34:1545–1553. doi:10.1016/S0021-9290(01)00149-X

32. Trepat X, Grabulosa M, Puig F, Maksym GN, Navajas D, Farré R (2004) Viscoelasticity of human alveolar epithelial cells subjected to stretch. Am J Physiol Lung Cell Mol Physiol 287:1025–1034. doi:10.1152/ajplung.00077.2004

33. Puig-de-Morales-Marinkovic M, Turner KT, Butler JP, Fredberg JJ, Suresh S (2007) Viscoelasticity of the human red blood cell. Am J Physiol Cell Physiol 293:597–605. doi:10.1152/ajpcell.00562.2006

34. Lubarda VA, Marzani A (2009) Viscoelastic response of thin membranes with application to red blood cells. Acta Mech 202:1–16. doi:10.1007/s007-008-0005-y

35. Pravincumar P, Bader DL, Knight MM (2012) Viscoelastic cell mechanics and actin remodelling are dependent on the rate of applied pressure. PLoS ONE 7: Paper No. e43938. doi:10.1371/journal.pone.0043938

Chapter 7
Data Processing Techniques to Analyze Large 3-D Deformations of Cardiac Cycles

C.A. Sciammarella, L. Lamberti, and A. Boccaccio

Abstract Quantification of 3-D deformations of human organs plays an important role in the understanding phenomena that have an impact in medical diagnosis and treatment of diseases. One important example is the mechanics of heart functions. Comparing normal deformation patterns of the cardiac cycle in healthy and diseased individuals can be a diagnostic tool that provides early and accurate indications of the onset of heart diseases.

The tagging technique is an experimental mechanics method that makes it possible to utilize the extensive literature existing on the analysis of deformations utilizing the digital moiré method for accurate and fast quantification of the heart 3-D kinematics. MRI tagging is an imaging technique used in medicine to visualize the structures of tissues of the human body in detail. MRI uses of the phenomenon of nuclear magnetic resonance to image tissues by exciting the nuclei of atoms in the tissue. Because of the different chemical composition of the tissues it can provide details that cannot be visible with CT Scans. By modulating magnetization it is possible to inscribe lattice-patterns in the tissue volume. These lattices are fixed to the under laying tissues for periods of time long enough to follow a cardiac cycle.

The objective of this paper is to outline image processing techniques that can be utilized to decode the displacements and strains taking into consideration that one is dealing with large 3-D deformations that form a time sequence of images. These techniques are based on fundamental principles that have been developed in the field of digital moiré.

Keywords MRI • MRI tagging • Digital moiré • Large deformations • Deformation of cardiac tissues

7.1 Introduction

In [1, 2] the subject of the 3-D Experimental Mechanics analysis of opaque media utilizing techniques that measure displacements has been addressed. The basic derivations of continuum mechanics mathematical models were reviewed and the particular modus operandi of the currently utilized methods—moiré, holography, speckle methods, digital image correlation (DIC) and the HARP method—were outlined indicating the common basis of all of these techniques and their differences. This paper is devoted to one of these techniques, the HARP method. This method was originally introduced in the bioengineering community [3–32] and very little connection was established between the extensive developments in the Experimental Mechanics community and the more biologically focused application of the HARP method.

The Experimental Mechanics application of the MRI method has very interesting aspects for the community of experimental mechanicians because of its implications as a generalization of well established methods in a completely different domain with a diverse physical basis. It is not a simple endeavor to provide a picture of the phenomenon of

C.A. Sciammarella (✉)
Department of Mechanical, Materials and Aerospace Engineering, Illinois Institute of Technology,
10 SW 32nd Street, 60616 Chicago, IL, USA

Department of Mechanical Engineering, College of Engineering and Engineering Technology,
Northern Illinois University, 590 Garden Road, 60115 DeKalb, IL, USA
e-mail: sciammarella@iit.edu

L. Lamberti • A. Boccaccio
Dipartimento Meccanica, Matematica e Management, Politecnico di Bari, Viale Japigia 182, 70126 Bari, Italy

magnetic resonance (MR), its application to image generation, the introduction of the tagging process and finally the analysis of displacements and deformations of organs of the human body. In this paper, a comprehensive presentation of the overall subject is developed.

7.2 Fundamentals of MRI Images Generation

The MRI method to visualize the internal structure of human organs is not unlike the usual methods based on the utilization of visible light to get information on displacements and their derivatives in transparent media. Both methodologies are based on the electromagnetic spectrum: that is, they have the same foundations, solution of the Maxwell equation of propagation of the electromagnetic fields. In the case of optical methods with visible light, the electric vector $\mathbf{E}$ is the vector utilized to obtain displacement information and the presence of the magnetic vector $\mathbf{B}$ is ignored in most cases. In the MRI methodology, the vector $\mathbf{B}$ plays the role of the illuminating vector since it can penetrate the human body different tissues and generate signals that can reflect the state of magnetization of a voxel located inside a volume. The electrical and the magnetic fields can be represented by a mathematical entity, a phasor:

$$\overrightarrow{\overrightarrow{\mathbf{B}}}(\mathbf{x}, t) = B(\mathbf{x}, t)e^{i\boldsymbol{\Theta}(\mathbf{x}, t)} \tag{7.1}$$

Equation (7.1) represents a propagating wavefront in space and time. $\overrightarrow{\overrightarrow{\mathbf{B}}}(\mathbf{x}, t)$ is the field amplitude at a given point and at a given time in the 3-D space, $\mathbf{x}$ is the position vector, $B(\mathbf{x}, t)$ is the amplitude of the phasor, t is the time, $\boldsymbol{\Theta}(\boldsymbol{x}, t)$ is an angular variable (phase).

The phase $\boldsymbol{\Theta}(x, t)$ can be expressed as:

$$\boldsymbol{\Theta}(x, t) = \mathbf{k} \bullet \mathbf{x} - \omega t + \alpha \tag{7.2}$$

where: $\mathbf{k}$ is the wave vector, $\mathbf{x}$ is the position vector, ω is the angular frequency, α and initial phase that depends on the selected initial conditions. In the case of the visible spectrum, the angular frequency term is ignored because the frequency of the signal is so high that in general the times required for signal sensing are very large to detect this part of the signal.

Equation (7.1) represents the signal that needs to be detected. Hence, we need an illumination field so to speak: in this case, it is a strong magnetic field where the object to be studied is immersed. This magnetic field propagates plane wavefronts; this implies a very uniform field in the region of observation. The interaction of the magnetic field with human tissues depends on a phenomenon, the magnetic resonance of the protons in the nuclei of the hydrogen contained in the tissues. The response depends on the density of the present nuclei and hence will depend on the chemical composition: this makes possible to separate different types of tissues. To collect information, it is necessary to detect weak changes in the uniform magnetic field existing in the volume of the illuminated body. To achieve this objective, signals must be sent to interrogate the state of magnetization at given points of the analyzed volume. These signals generate electromagnetic waves that upon recording and processing provide local information of the magnetic state of the volume. In order to get information concerning human tissues, we need to produce an image that provides features of the observed tissues as levels of gray as it is done usually in conventional images produced with a camera and a lens. Except that in this case there is not going to be a lens or an array sensor of an electronic camera to record information in pixels and voxels.

In the present case, the illumination is a magnetic field and the distinctive features come from small changes experienced by the magnetic field that have to be sensed and transformed in levels of gray at a given location. This last operation is the interface between the system generating and processing the information and the human observer that has to make inferences on the obtained information. Furthermore, if we want to measure displacements we must create a reference system embedded in the tissues and that experiences the same changes in time as the tissue due to the applied deformation between two time intervals. Let us consider these two aspects separately beginning with the generation of an image of the organ under observation.

7.3 The Image Generation

We adopt a Cartesian system of coordinates (X, Y, Z) and to simplify the analysis, as it is usual in MRI literature, in-plane images are given by the coordinates (X, Y) while depth information is given by the coordinate Z (see Fig. 7.1).

The information is given as a level of gray in voxels that cover the analyzed volume. It is necessary to locate the voxel position given by the coordinates (X, Y, Z). Therefore, we need to perform separate measurements that allow us to locate the

Fig. 7.1 Voxels containing
level of gray information
and coordinates system
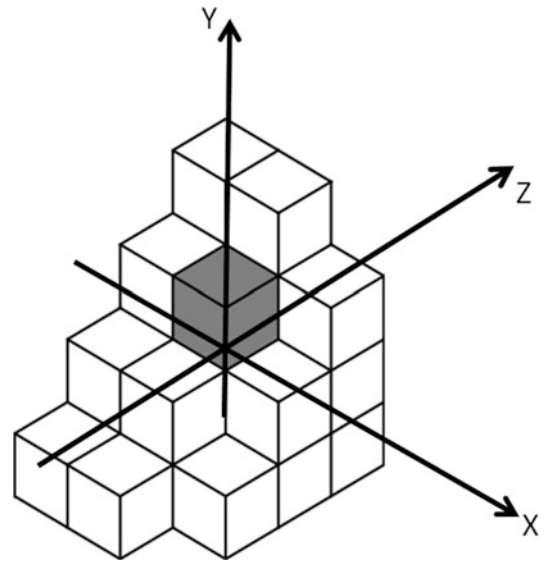
corresponding voxel in the volume: this implies three unknowns. Furthermore, the intensity of the magnetic field is defined by a vector that has three separate components. This introduces another three unknowns, and hence a three dimensional image implies at least another three unknowns for a total of six unknowns. One can utilize interpolation algorithms to interpolate gray levels between voxels and display images. To understand the process of image formation it is necessary to review the physics of magnetic resonance (MR).

7.4 Basic Physical Law of Magnetic Resonance

As mentioned before the effect of the magnetic field on the protons of the nuclei of hydrogen contained in the tissues is the experimental tool to visualize features of the observed organs. The protons of the hydrogen molecules in the nuclei of the tissues have two properties defined by their quantum mechanics description (1) a spin or intrinsic angular momentum, (Fig. 7.2a); (2) associated with the spin, there is a magnetic dipole moment of the same direction of the spin (Fig. 7.2b). The spin is a vector but not in the classical sense; the spin experiences changes according to quantum mechanics rules.

Another important property is that nucleons like electrons fill orbitals. Consequently, the total resultant spin of the nucleus cancels out; number of neutrons and protons even. If the number of nucleons is odd (this occur when one is dealing with an isotope of a given element) there is a resultant spin in the nucleus. Only those elements that have abundant isotopes can be detected because the odd nucleon numbers are only present in isotopes.

If an external magnetic field is introduced, different spin states are produced in the protons that depend on potential energy levels. The external magnetic field produces the polarization of the entire population of protons; they align the spin axis in the direction of the applied magnetic field. If an additional magnetic signal is introduced, changes of the polarization state take place between two alternative states, up and down. The up and down refers to a vertical magnetic field B_o. The spins of the protons will align with the applied field. There are two possibilities: they can point in the same direction of the field (up), or they can be opposite in direction (down). The up position corresponds to a higher energy level, the down position to a lower energy level. If the signal is switched off, the return of the population of protons from the upper to a lower energy level produces a measurable emission of electromagnetic waves. The energy drop provides the energy to propagating waves that carry information concerning the state of magnetization of a given voxel; this phenomenon is called nuclear magnetic resonance (NMR) and this is the base of the MRI technique.

The torque introduced by the magnetic field is given by the equation,

$$\frac{d}{dt}\boldsymbol{\mu}_s = -\boldsymbol{\gamma}\mathbf{B}_o \times \boldsymbol{\mu}_s \tag{7.3}$$

where: $\boldsymbol{\mu}_s$ is the magnetic moment, $\mathbf{B}_o$ is the externally applied magnetic field, γ is called the gyromagnetic constant, the symbol $\times$ indicates the vector product. $\mathbf{B}_o$ is given in Tesla and for hydrogen $\gamma = 42.58$ MHz/T, the bold letters represent vectors.

Fig. 7.2 Spinning proton
and precession of the spinning
proton
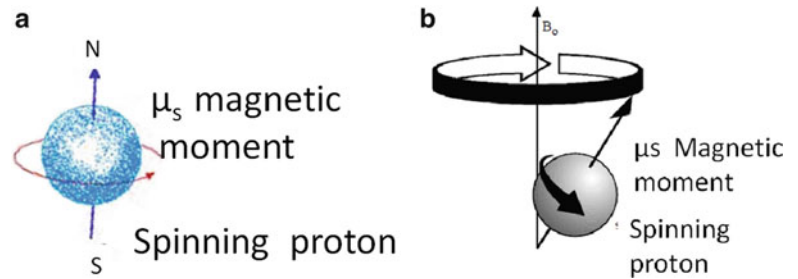
In the transition from the up to the down position the spin vector becomes inclined with respect to the applied direction of the magnetic field. This process has a mathematical similitude with the precession of a top in the gravitational field: as the energy of the top decays, it starts to have a precession motion.

Projecting the component of the proton spin into a plane perpendicular to the magnetic field, this component rotates with an angular frequency,

$$v_{Lo} = \gamma |\mathbf{B}_o| \tag{7.4}$$

where the bar indicates the magnitude of the magnetic field vector.

This angular frequency is called Larmor frequency. Equation (7.4) is a fundamental equation in MRI because, as stated previously, the process of decaying energy is related to the emission of a signal that provides a measure of the magnetization state of a voxel. It is possible to see that the Larmor frequency is proportional to the magnitude of the vector $\mathbf{B}_o$. This process is similar to the stimulated emission phenomenon in optics: the main difference is that while in the stimulated emission the orbital electrons are involved, in the magnetic resonance the atomic nucleus is the source of the emitted radiation. There are an active medium (the nuclei of the tissues that will generate an electromagnetic radiation), a pumping source (an applied magnetic field) and a resonator (the cavity that contains the magnetic field).

7.5 Energy Levels and Transitions

As shown in Fig. 7.2, the proton acts as a small magnet and there are two possible arrangements of the poles with respect to the applied magnetic field (see Fig. 7.3). The applied magnetic field is symbolically represented by the two poles of a magnet. The two possible arrangements are N-N-S-S or N-S-N-S. These arrangements depend on the energy levels. The following equation relates the energy to the Larmor frequency defined by Eq. (7.4) that is a fundamental parameter in the nuclear phenomenon.

The energy involved in the transition of states is given by:

$$E = h v_L \tag{7.5}$$

Where h is the Planck's constant, $h = 6.626 \times 10^{-14}$ Js.

Figure 7.4 visualizes the quantum energy change involved in flipping over two states of spin of the proton. The angular frequency associated with the flip over is the Larmor frequency which is proportional to the applied magnetic field amplitude $\mathbf{B}_o$ as indicated by Eq. (7.4).

7.6 Statistic of a Voxel Magnetization

MR observations are statistical and within a voxel there is an extremely large number of molecules: hence, an ensemble average is detected. The Bolztmann statistics applies to this particular case; at a given absolute temperature T the ratio of the number of protons in the high energy level N-S shown in Fig. 7.4 (called N^+) to the protons in the low energy level S-N (called N^-) is given by:

Fig. 7.3 Possible
arrangements of
the proton poles
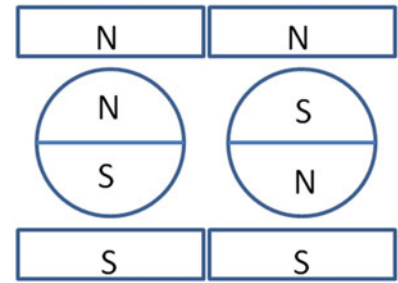
Fig. 7.4 Energy levels of the
magnetic potential energy
of the proton in the transition
of energy levels
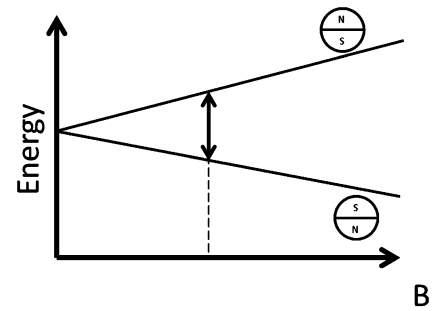
Fig. 7.5 Magnetic moments
produced by a strong
magnetic field **B**
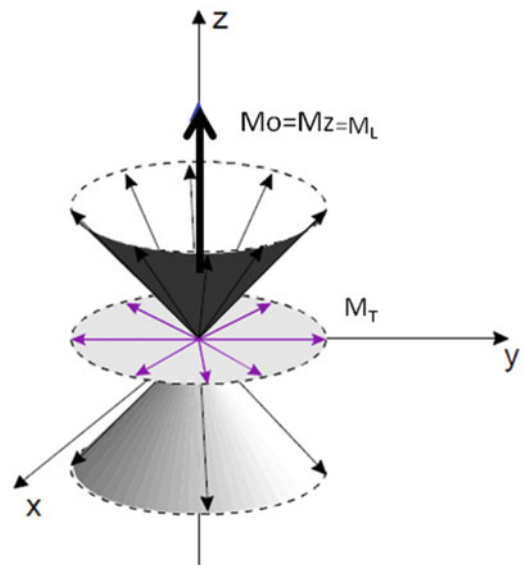
$$\frac{N^-}{N^+} = e^{-\left[E/kT\right]} \tag{7.6}$$

where k = 1.3806488 × 10^{-23} J/K is the Boltzmann constant. We call these two states of energy E_1 and E_2. A transition between these two states of energy will provide the energy to generate electromagnetic photons that characterizes MR.

Taking the Z-axis as reference, one sum up all the individual magnetic moments in a voxel, $M_z = \sum \mu$, and define a resultant moment for the voxel. At room temperature and in absence of a pumping radiation, the Boltzmann equation predicts an equilibrium state, $M_o = M_z = 0$. This means that there is not an observable resultant magnetization moment. If a pumping magnetic field of low intensity **B** is applied (Fig. 7.5), there is going to be a longitudinal resultant moment M_L and transverse components of the magnetic moment M_T. However, due to the statistical effect that we have mentioned the observable $M_L = M_T = 0$. The pumping energy must be strong enough to produce observable M_o, M_L, M_T. For magnetic fields, **B** must of the order of **B** > 3 T. That is, according to the Boltzman's Eq. (7.6), the number of protons in the up position energy E_1 will predominate over the number of protons in the down position E_2. The transition between the two states provides an energy E = $E_1 - E_2$ given by Eq. (7.5). This circumstance implies that by removing the pumping magnetic field propagating wavefronts will be generated that carry the excess

Fig. 7.6 Behavior of the longitudinal moment $\mathbf{M_o}$ and the transverse moment $\mathbf{M_T}$ with respect to time t
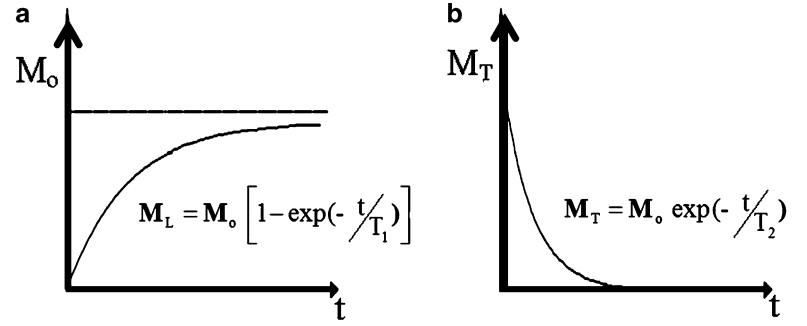
energy. These wavefronts contain information concerning the state of magnetization of the different tissues since their chemical composition will be different and consequently their responses will change.

The above description provides the basic understanding of MR. However, in order to comprehend the whole process of detection, we need additional information on the generated signals and how these signals can be manipulated in the process of detection to define the magnetization at a given voxel. In the excited state, the total magnetic moment of a voxel will lie along the magnetic field $\mathbf{B_o}$: under this condition there is not transversal $\mathbf{M_T}$. Once a pumping magnetic field is applied a transversal magnetization moment $\mathbf{M_T}$ appears (see Fig. 7.5). The precession of the protons spinning axis takes place. If the pumping field is removed, the spinning axes return to their original direction. The pumping process takes a time period called relaxation time T_1. This process follows an exponential function. The initial conditions are: at time t = 0 $M_o = M_z = 0$. The final condition is: at time t = T_1, $M_0 = M_z = M_L$.

This process is governed by Eq. (7.7), Fig. 7.6a.

$$\mathbf{M_L} = \mathbf{M_o}\left[1 - e^{-t/T1}\right] \tag{7.7}$$

The initial condition of the decaying process is: at t = 0, $\mathbf{M_T} = M_0$; the final condition is: at t = T_2, $M_T = 0$. This process is governed by the following equation:

$$\mathbf{M_T} = \mathbf{M_o}\left[1 - e^{-t/T2}\right] \tag{7.8}$$

This behavior is illustrated in Fig. 7.6b. The result of this analysis tells us that the emitted signals upon decay are damped sinusoidal signals. This is an important piece of information in the detection process.

The value of T_1 is larger than T_2. The values of T_1 and T_2 depend not only on the nuclear particles behavior but also on the chemical composition of the observed tissues. The two processes take place together and the transverse magnetization will be rotating at the Larmor frequency.

7.7 The Pumping System

We need now to analyze the spatial effect of the pumping process since it plays an essential role in MR. In Fig. 7.7 we assume that the magnetic field $\mathbf{B_o}$ is in the direction of the Z-axis, the pumping signal is oriented in the direction of the axis of the coil (Y-axis), $\mathbf{B_1}$ is the pumping magnetic field produced by the coil, a periodic function of the time. The magnetized volume is schematically represented by a cylinder and the electromagnetic signal is symbolized by a sinusoidal signal. The frequency of the pumping system is made to coincide with the Larmor frequency. Dynamical events must be described. For that purpose, we will resort to a rotating frame of reference that has the Larmor angular frequency. The changing field in the Y-direction can be viewed as to circularly polarized fields with opposite rotations in the XY plane (Fig. 7.7b). The reference frame rotates with the Larmor angular speed with respect to the spatial coordinate system. With respect to this frame one of the two vectors, the one that has the same sense of rotation as the Larmor angular vector will be stationary with respect to the rotating frame. We can assume that the total voxel magnetization has a precession around the X-axis as shown in Fig. 7.8 at the same time that the axis rotates with the Larmor angular speed.

Calling the period of the applied sinusoidal field τ, the resulting magnetization vector $\mathbf{M_o}$ that is observed along the direction of the applied field $\mathbf{B_o}$ will rotate around $\mathbf{B_1}$ (see Fig. 7.8) of an amount that depends on τ according to the equation

Fig. 7.7 (**a**) Radio frequency excitation to pump the magnetic resonance. Symbolic representation of the pumping coil (the selection of the Y-axis is arbitrary); (**b**) Magnetic field in the XY plane (the selection of the X-axis is arbitrary)
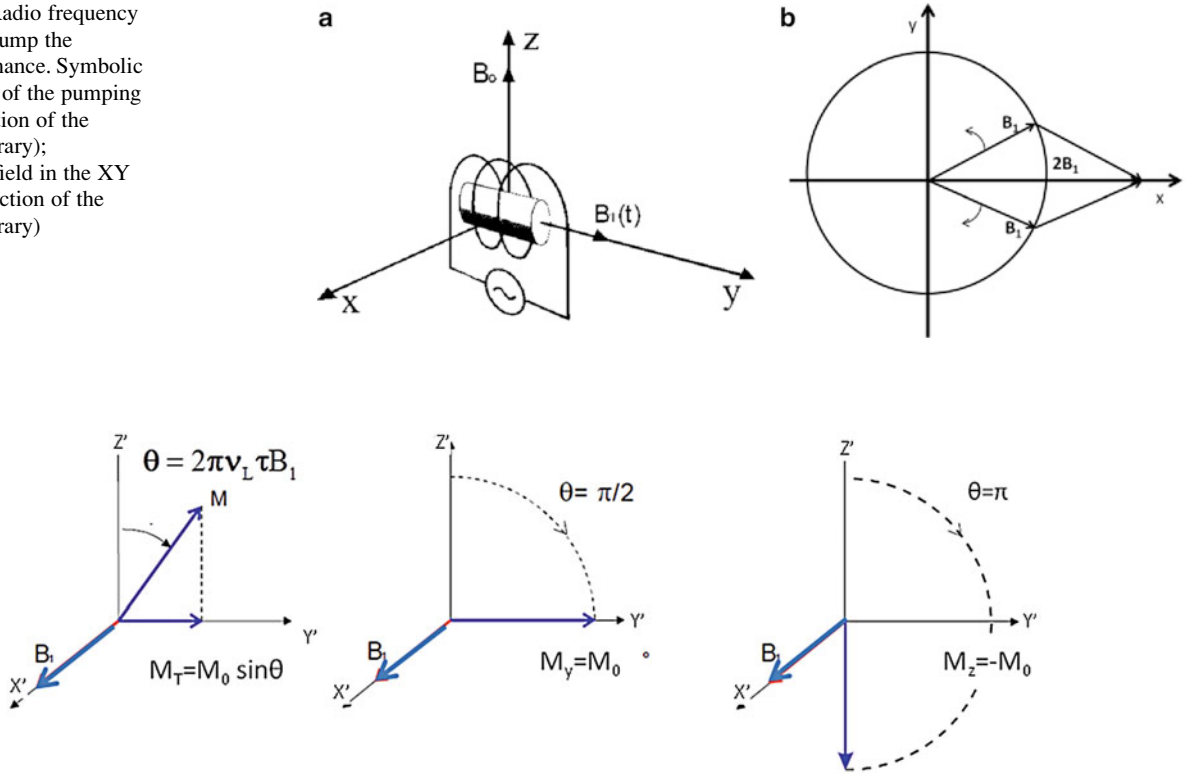
Fig. 7.8 Rotation of the resulting magnetization vector depending on the magnitude of the applied magnetic field. The primes in the coordinate axes indicate the rotating framework

$$\theta = 2\pi\nu_L\tau B_1 \tag{7.9}$$

where the angle θ is expressed in radians, according the relationship, 2π rad/s $=1$ Hz $= 1/$s.

Figure 7.8 shows that the $\mathbf{M_T}$ vector generated by applying a magnetic pulse can rotate the transversal moment by different angles including the complete reversal of $\mathbf{M_z}$. In the MRI literature one refers to a $\pi/2$ pulse that changes the vector from one axis to the perpendicular axis or a π pulse that produces the reversal of the magnetization $\mathbf{M_o}$ from $+Z$ to $-Z$. One must remember that the flipping of the magnetization takes place at the same time that the coordinate system rotates.

There is an additional effect that has to be considered due to two factors (1) different levels of magnetization induced by inhomogeneous fields in the Z-direction; (2) different molecular interactions, in adjacent voxels, the transversal magnetization has different Larmor frequencies, and hence the phases Θ change for the different voxels.

By adopting a reference Larmor frequency ν_{Lo} some voxels will have higher frequencies while other voxels will have lower frequencies because of the changes in magnetization. For example, let us assume that for a voxel of coordinate y the Larmor frequency is $\nu_L(y)$ and it is higher than the reference ν_{Lo}. The positive frequency difference can be defined:

$$\Delta\nu = \nu_L(y) - \nu_{Lo} \tag{7.10}$$

The total voxel behavior of the resulting magnetic moments is given by the Bloch equation:

$$\frac{d\mathbf{M(x)}}{dt} = \gamma\mathbf{M(x)} \times \mathbf{B(x)} - \frac{\mathbf{M_T(x)}}{T_2^*} - \frac{\mathbf{M_L(x)} - \mathbf{M_o(x)}}{T_1} \tag{7.11}$$

The above differential equation tells us that the state of magnetization of a voxel is a function of space and time, that depends on the applied magnetic field and on the mechanisms of relaxation mentioned before. In the equation, the symbol T_2^* was introduced. The constant T_2 is related to the decay of $\mathbf{M_T}$. The decay of $\mathbf{M_T}$ depends on two effects (1) the pure molecular effect caused by the chemical composition of the tissues; (2) the inhomogeneity of the field $\mathbf{B_o}$.

The relationship between T_2 and T_2^* is:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2inhom}}$$

(7.12)

7.8 Voxel Magnetization Measurement

We have gained some more understanding of the signal and its possible positions in space as a function of the time. We have analyzed the effect of an added electromagnetic signal to the state of magnetization of the different voxels. We now need to look to detection mechanisms capable of registering the changes caused by the pumping signal that is introduced by a sinusoidal electromagnetic field.

There is a signal that can be recorded and reflects the state of magnetization of a given voxel. According to Farady's law, a change of the magnetic field in time generates an electromagnetic force on a coil. In Fig. 7.9, there are shown two coils: one with axis in the X-direction that records the rotating field in the plane XY as shown in Fig. 7.8; the other with axis in the Y-direction, the pumping coil. The electromotive force is a damped sinusoidal signal according to Eq. (7.8).

To this point we have described the main features of the physics involved in MRI imaging of tissues. Going back to Eqs. (7.1) and (7.2) that represent the propagating magnetic field we need now to detect the state of magnetization of a voxel through the signals coming from the voxel according to the process described in the preceding paragraphs. We need to utilize the knowledge of the magnetic resonance phenomenon to make inference about the amplitude and the temporal phase of the signal. As stated before, in Optics the temporal phase is not easily available due to its extremely high frequency. In MR, the temporal frequency is in the range of radio frequencies and hence is available for detection with standard detection procedures. In the next section, procedures to encode pixel information in MRI and determine voxel location and state of magnetization will be outlined.

7.9 Basic Imaging Principle

In order to detect the magnetic field at a given voxel, the voxel should experience a unique magnetic field that should provide the voxel position in the analyzed volume. This detection is not feasible if the magnetic field is uniform in the volume under analysis. In MRI technology, determination of voxel position is achieved by introducing magnetic gradients that modify the uniform magnetic field created by the main magnet producing the magnetic "illumination" mentioned at the introduction. To simplify the description let us look to the magnetic gradient in one direction, for example the X-direction (Fig. 7.10): B_{ox} indicates the value of the magnetic field at the reference center of coordinates; 2L is referred in the MRI literature as the field of view or region where the initial field has a high degree of homogeneity that secures a uniform field, the presence of gradients in the original field will cause changes in the magnetic signal.
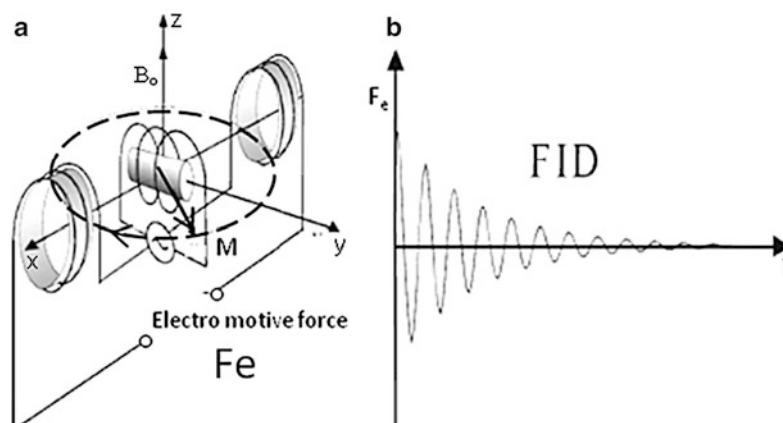
Fig. 7.9 (**a**) Coil with axis in the X-direction to sense the rotating electromagnetic field caused by a pulse applied by the coil with axis in the Y-direction and producing a rotating field in the XY plane; (**b**) Symbolic representation of the pumping system and the signal recording system [The word FID stands for free induction decay of the transverse magnetization, Eq. (7.8)]

Fig. 7.10 Magnetic gradient in the X-direction
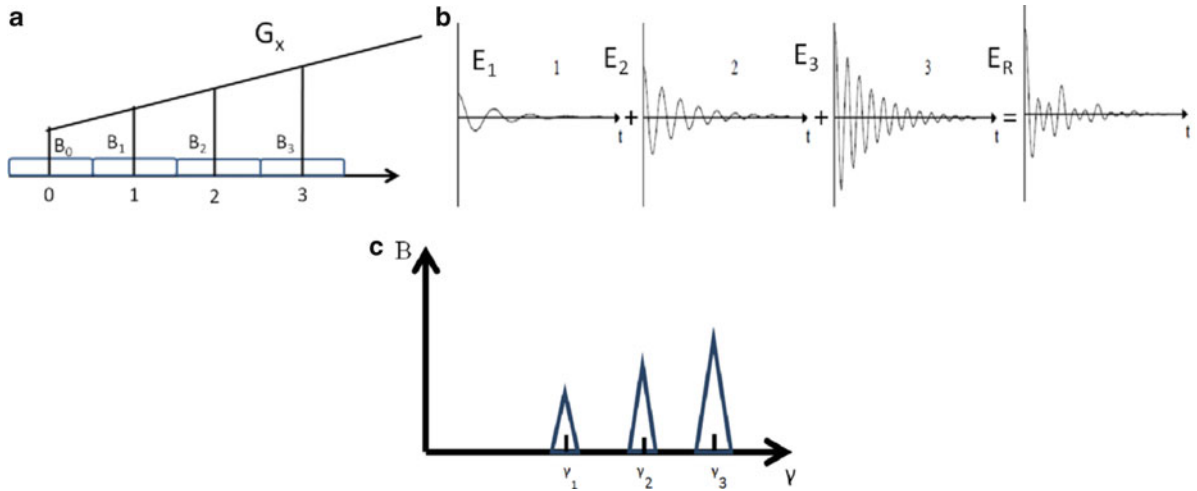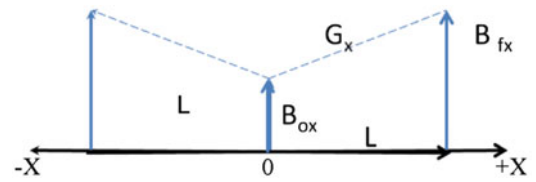
Fig. 7.11 (**a**) Voxels and the different fields acting at each voxel; (**b**) Signals corresponding to different voxels: as the intensity of the magnetic field increases the Larmor frequency increases and the E_R signal is detected by the receiving coil; (**c**) Fourier transform of the E_R signal of (**b**)

The detection of $\mathbf{M}(\mathbf{x})$ is based on time frequency detection. The term to be utilized is the phase term contained in Eq. (7.2): the angular frequency of the propagating field. In this particular case, the reference frequency is the Larmor frequency ν_{Lo} defined by Eq. (7.4). Then, at a given voxel of coordinate x, the Larmor frequency will be, if we include the presence of a gradient G_x,

$$\nu(x) = \gamma(B_o + G_x x) \tag{7.13}$$

From Eq. (7.13), it follows:

$$x = \frac{\nu_x - \nu_o}{\gamma G_x} \tag{7.14}$$

A similar equation can be applied to the other coordinates. The detection of position is then obtained through frequency analysis of the resultant ER signal (see Fig. 7.11) carried out with a digital/software Fourier spectrum analyzer. Consequently, the position in space of a given voxel depends on the difference of the imposed Larmor frequency due to the reference field $\mathbf{B_o}$ and the actual Larmor frequency at the considered voxel.

Figure 7.11 illustrates the process mentioned above. In particular, it shows the frequency composition of the signal E_R and the amplitude of the field in a certain X-scale that can be established through calibration. The position of each voxel is known with respect to the X-axis: x_1, x_2 and x_3. The corresponding amplitudes B_1, B_2 and B_3, proportional to the number of nuclei in the corresponding voxels, are known as well. In principle, all the elements that are necessary to produce an image of the magnetic field are available. However, there is a large number of additional effects that it is necessary to understand to grasp the power of the MR technique.

7.10 Extension of the MRI Encoding Procedure to the 3-D Case

The simple procedure outlined in one dimension cannot directly be extended to a volume. The obtained results are based on the fact that there is a unique field $\mathbf{B}(\mathbf{x})$ with different values at each point of the X-axis. However, it is not possible to have a field $\mathbf{B}(\mathbf{x}, \mathbf{y})$ with simultaneous different amplitudes at a given point of coordinates (x, y) in view of the continuity of the field

that responds to Eq. (7.11). This equation has two main independent variables, the space and the time. The solution of the problem resides in taking advantage of both variables in combinations that will make it possible to obtain the six unknowns required to characterize the state of magnetization of a given tissue and produce an image. It is clear that the relaxation mechanisms will play a fundamental role in the corresponding derivations.

It is possible to approach the solution of the problem in different ways: one of many possible solutions will be outlined in the article. The local magnetic field is $\mathbf{B}(\mathbf{x}, t)$ and it will be required to apply gradients $G_i(\mathbf{x}, t) = \partial \mathbf{B}_z / \partial x_i$ ($i = 1, 2, 3$). In the preceding expression of the gradients, the convention of coordinates shown in Fig. 7.1 has been followed: the Z-coordinate is the direction of the applied field and we are looking at sections in the XY plane.

Equation (7.4) can be generalized to obtain a more general expression of the magnetic field:

$$\mathbf{B}(\mathbf{x}, t) = \mathbf{B}_o + \mathbf{x} \bullet \mathbf{G}(t) \tag{7.15}$$

Then the Larmor frequency is:

$$\nu(\mathbf{x}, t) = \nu_{Lo} + \gamma [\mathbf{x} \bullet \mathbf{G}(t)] \tag{7.16}$$

which depends of the position $\mathbf{x}$ and the time t. The phase of the magnetization vector, Eq. (7.2), can be expressed as:

$$\theta_{T_{ot}}(\mathbf{x}, t) = 2\pi \int_0^{t_o} \nu(\mathbf{x}, t) dt = 2\pi \left[\nu_{Lo} t_o + \gamma \int_0^{t_o} [\mathbf{x} \bullet \mathbf{G}(t)] dt \right] \tag{7.17}$$

Considering the second term of Eq. (7.15) and only the component of the phase in excess of the phase generated by the Larmor frequency, it can be written:

$$\theta_R(\mathbf{x}, t) = 2\pi \left[\gamma \int_0^{t_o} [\mathbf{x} \bullet \mathbf{G}(t)] dt \right] \tag{7.18}$$

The state of magnetization is encoded in this phase: should this information be retrieved, inferences about the magnetization of a given voxel can be made.

7.11 Phase Encoding

In Sect. 7.9, it was outlined a procedure to retrieve the X-coordinates of a voxel utilizing a spatial gradient G_x. In the preceding sections the more general case of gradients in space and time was addressed and it was shown that if we take a rotating frame of reference with the angular speed of the Larmor reference frequency, the phase accumulated by the effect of the applied gradient provides information that is related to the voxel location. This additional information can be utilized to locate voxels in the volume of the observed object.

Let us consider three neighbor voxels, the magnetization has been turned along the X-axis as shown in Fig. 7.8. Since the three neighbor pixels in the YZ plane have the same chemical composition, in a uniform magnetic field they will rotate at the same Larmor frequency. They will start spinning at the same time: setting the X-axis as a reference, the angle θ formed with respect to the X-axis will be the same (see Fig. 7.12a).

If a gradient G_y in the Y-direction is applied, the spin vectors of the different voxels will have a precession around the X-axis at the frequency

$$\nu(y) = \left(\nu_o + \gamma G_y y \right) \tag{7.19}$$

If the pulse is on, the vectors of each voxel will have a different Larmor frequency. This pulse, because it changes the field, makes the different vectors rotate with different frequencies: hence, their relative phases will change. If the gradient in the Y-direction is removed, then the same pulse originally in the X-direction, is applied again. The Larmor frequency returns to its original value, the vectors rotate at the same speeds but because of the effect of G_y they have different phases (see Fig. 7.12b).

Fig. 7.12 Phase encoding in neighbor voxels (**a**) Initial state (uniform magnetic field); (**b**) Perturbed state (G_y gradient)
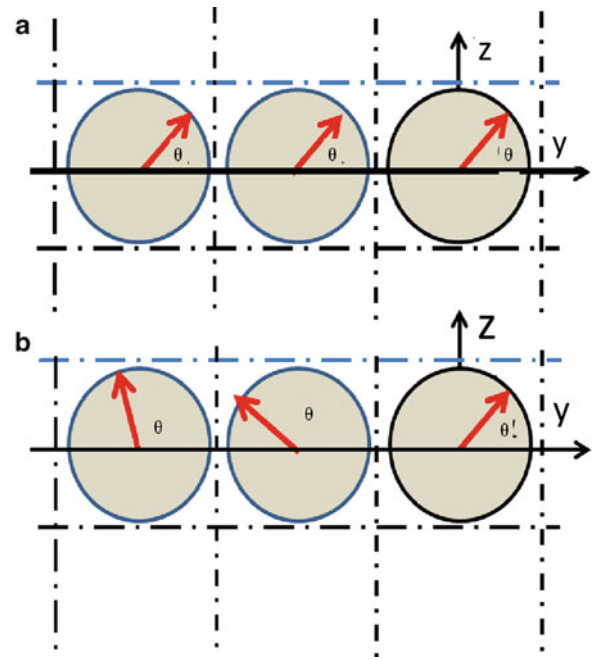
7.12 Three Dimensional Information Encoding

The next step is to accomplish the recording of the phases. We can now analyze the problem in three dimensions. In Fig. 7.13, there are indicated different voxels in a given volume. The vertical lines indicate the uniform field $\mathbf{B}_o$ applied by the main magnet. Reference axes are chosen following the same convention utilized in Fig. 7.1. Figure 7.13 also shows the different applied gradients to obtain information from the whole volume. $\mathbf{G}_S$ is called the slice selection gradient: this gradient is applied in the Z-direction to select slices in the Z-direction; $\mathbf{G}_\upsilon = \mathbf{G}_x$ is applied in the X-direction and is utilized to locate the position of a voxel in the X-direction according to the process described in Sect. 7.9; $\mathbf{G}_\theta = \mathbf{G}_y$ is utilized to encode the voxel Y-coordinate. S_i is the pumping signal that starts the process, S_R is the recorded signal and the time lines represent the sequence in time of the different steps needed to produce an image. The sequence outlined on the right of Fig. 7.13 is the customary way utilized in MRI to describe the complete process of obtaining the magnetization of a voxel to produce an image.

Section 9 described the process required to find the coordinate X; in that case there was a straightforward relationship between the gradient and the X-location of the voxel. Section 7.10 described the basic steps to measure the precession phase angle θ. In the next section, it will be explained the procedure to measure the phase θ. This technique requires a process that generates an echo magnetic signal.

7.13 Echo Magnetic Signal

Figure 7.14a shows the sequence of pulses required to produce the spin echo phenomenon. As shown in the time-space diagram, a $\pi/2$ pulse is applied. The magnetization vector of direction +Z is rotated to the +Y′-axis (the primal indicating the rotating frame of reference). At a period of time τ, a π pulse is applied and this pulse rotates of 180° from the +Y′ to the −Y′; this last operation generates the echo signal. The transversal signal decays with the time $T_2^* > T_2$. Figure 7.14b shows the echo pulse. For the sake of brevity, the derivation of the formation of the echo pulse is omitted.

Fig. 7.13 Determination of the state of magnetization of different voxels in a given volume represented by a cube
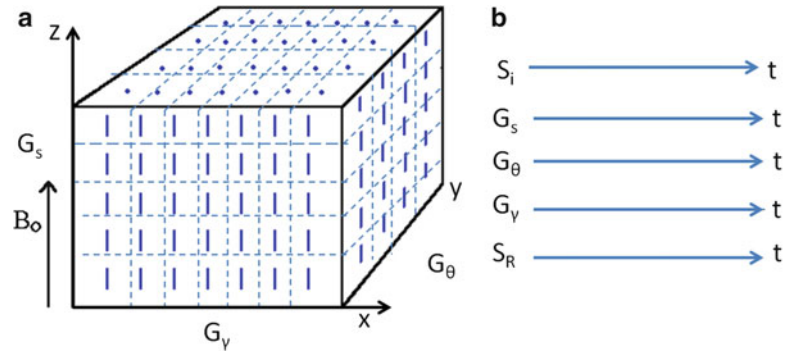
Fig. 7.14 Sequence of pulses required to produce an echo magnetic signal
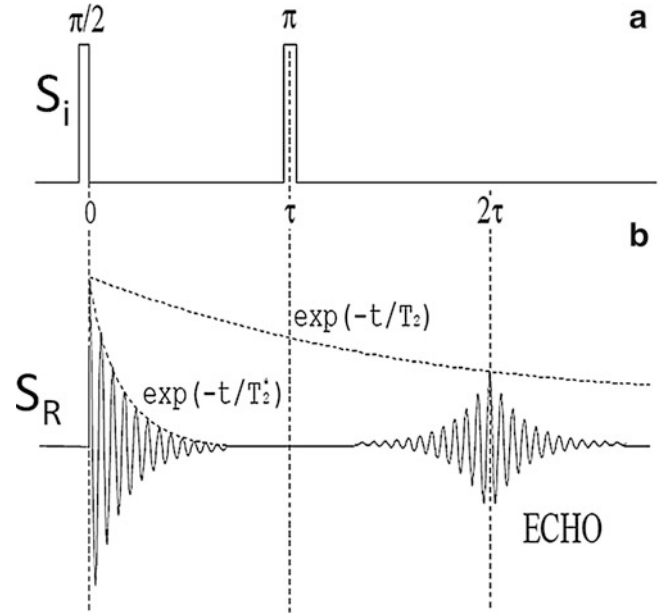
7.14 Location of a Voxel in the XY Plane

Figure 7.15 shows the outline of the pulses required to determine the position of a voxel in the XY plane. It is possible to follow the evolution of the frequencies in the frequency plane obtained by applying the FT to the collected voltages. At time $t = 0$, the $\pi/2$ pulse is applied and the gradients G_v and G_θ are applied, represented in the frequency plane by the point A of coordinates $(\nu_{x1}, \nu_{\theta 1})$, see Fig. 7.16. A pulse π is applied at $t = \tau$, and, because of the flipping, the point A' of coordinates $(\nu_{x1}, -\nu_{\theta 1})$ is obtained. At the time $t = t_b$, it starts the readout period where only G_v is applied and one obtains the point B'' $(\nu_{x1}, -\nu_{\theta 2})$.

The initial phase is $\theta_i = \text{arctg} \dfrac{\nu_{\theta 1}}{\nu_{x1}}$ while the final phase is $\theta_f = \text{arctg} \dfrac{\nu_{\theta 2}}{\nu_{x1}}$. The corresponding phase difference is:

$$\theta_R = \theta_f - \theta_i \tag{7.20}$$

Recalling Eq. (7.13), it can be written for the Y-direction:

$$\nu(y) = \nu_\theta \gamma (B_o + G_y y) \tag{7.21}$$

It follows:

$$y = \frac{\nu_y - \nu_o}{\gamma G_y} \tag{7.22}$$

In this way one obtains the Y-coordinate of a voxel.

Fig. 7.15 Sequence of pulses to locate the position of a voxel in the XY plane
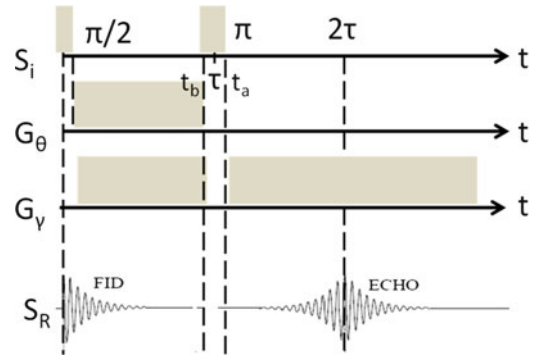
Fig. 7.16 Detection of the phase of the magnetization vector in the frequency plane
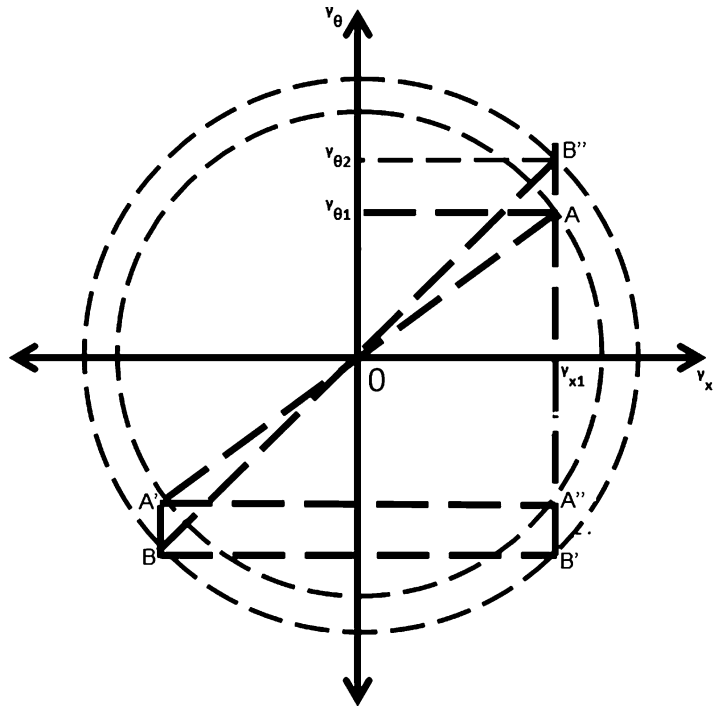
Every voxel will contribute to the overall voltage by an amount,

$$S_v(t) = M_T(x, y) e^{-i\left[-y\gamma G_y \tau - x\gamma G_x \tau + x\gamma G_x (t-\tau)\right]} \qquad (7.23)$$

Hence, the contribution of all voxels in one XY slice will be:

$$S_T(t) = \int_{x, y} M_T(x, y) e^{-i\left[-y\gamma G_y \tau - x\gamma G_x \tau + x\gamma G_x (t-\tau)\right]} dxdy \qquad (7.24)$$

The intensity of the field then is a function of the time. The maximum of the quantity obtained from Eq. (7.24) occurs for $t = 2\tau$ as shown in Fig. 7.15.

The sequence that has been described is repeated many times with a repetition period T_R. Repetition periods could be 128 or 256 to get the data necessary to reconstruct an image; every time that a time sequence is repeated the magnitude of G_x and G_θ are changed in equal steps.

The preceding sections have introduced the basics of any MRI method. The signals recorded following the pulses set up in time are analyzed in the frequency space (FT). The time during which the recordings can be made are limited by the damped nature of the observed phenomena. The recorded $S_R(t)$ signal depends on the decay times, the FID last between 50 and 200 ms after the pulses are applied. The time T_R depends on T_1 that ranges from 500 to 2000 ms.

From the presented developments the images produced are recorded under the form of a FT that yields at the end of the process the images. In the MRI literature, time frequency space is referred as the **k** space with components k_x that records the X-coordinate of a slice XY, and k_y that records the Y-coordinate through the phase θ. The symbol k plays a similar role to the **k**-vector in diffraction limited optics, $|\mathbf{k}| = 2\pi/\lambda$.

The corresponding equations are:

$$
\begin{cases}
k_x = \displaystyle\int_0^{t_{rec}} \gamma G_x(t)\,dt = k_{xrec} \\[4mm]
k_y = \displaystyle\int_0^{t_{yrec}} \gamma G_y(t)\,dt = k_{yrec}
\end{cases}
\tag{7.25}
$$

In the preceding equations, we have replaced the subscripts ν by x and θ by y. The above defined spatial frequencies define the resolution of the obtained images when the information in time is transformed into a space frequency to generate an image by means of software that converts the information of time frequencies into spatial voxels.

7.15 Selection of a Slice, Spatial Resolution, Signal Intensity

The derivations in this section are general in the sense that they apply to the limits of the dimensions of a voxel, that is the resolution that can be achieved in selecting a volume. In order to help visualize the different steps involved, the Z-coordinate will be considered. Figure 7.17 illustrates the slicing of the object (a human body) in the Z-direction.

The recorded signal by the sensing coil is proportional to:

$$
S(t) \propto \int_{-z}^{+z} M_o(z) e^{i\gamma(B_o + zG_z)t} e^{-\left(t/T_2^*\right)}\,dz
\tag{7.26}
$$

Since the signal is zero outside the body volume, the integral can be extended to infinity. Utilizing the notation of Eq. (7.25), the signal is proportional to the following quantity:

$$
S[(k(t)] \propto e^{-\frac{k(t)}{\gamma G_z T_2^*}} \left[\int_{-\infty}^{+\infty} M_o(z) e^{i[k(t)\,z]}\,dz \right]
\tag{7.27}
$$

The $S[(k(t)]$ term is the product of two functions: $P[k(t)] = e^{-\frac{k(t)}{\gamma G_z T_2^*}}$ and $Q[k(t)] = \int_{-\infty}^{+\infty} M_o(z) e^{i[k(t)\,z]}\,dz$. Therefore, we can write:

$$
S(t) \propto P[k(t)] \cdot Q[k(t)]
\tag{7.28}
$$

Where the "·" symbol indicates the product of the two functions.
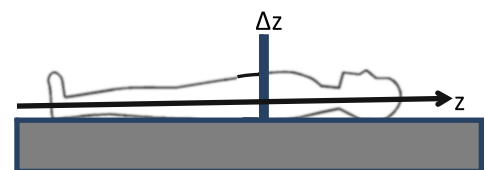
Fig. 7.17 Slicing in the Z-direction

Applying the convolution theorem of the FT in the Z-coordinate, it follows:

$$FT\{S[(k(t)]\} = L_f(z) \otimes M_o(z) \tag{7.29}$$

where L_f is the normalized Lorentzian function with a maximum equal to one

$$L_f(z) = \frac{1}{1 + \left(\gamma G_z T_2^*\right)^2 z^2} \tag{7.30}$$

The half width is,

$$\Delta z = \frac{1}{\gamma G_z T_2^*} \tag{7.31}$$

The quantity Δz gives the minimum possible size of a point in the image created by a MRI system. Hence, this quantity defines the resolution of the system in the same way that the Airy's ring defines the resolution of an optical system.

For the FT, following the relationship between pixel size and the maximum spatial frequency in an image, it holds $\delta z = 1/k_{rec}$. Recalling Eq. (7.25) it can be written:

$$\delta z = \frac{1}{\gamma G_z t_{rec}} \tag{7.32}$$

where t_{rec} is the time corresponding to the recording of the signal. Utilizing Rayleigh's criterion of resolution of two points in an image $\delta z < \Delta Z$, the following condition is reached:

$$\frac{1}{T_2^*} < \frac{1}{t_{rec}} \Rightarrow T_2^* > t_{rec} \tag{7.33}$$

Equation (7.33) tells us that to improve resolution one has to make T_2^* as large as possible by, for example, optimizing the homogeneity of the field produced by the main magnet. We can see that the actual resolution improves as the gradient increases and if we make the half width of the Lorentzian very small then:

$$S[(k(t)] = \int\limits_{-\infty}^{+\infty} M_o(z) e^{i[k(t)\,z]} \, dz \tag{7.34}$$

It appears that the magnetic image of the observed object is the FT of the acquired magnetic resonance signal S[k(t)].

The presented developments deal with the formation of an image of a tissue with magnetic "illumination". We need now to see how geometrical changes of the tissues can be observed and Continuum Mechanics can be applied to human organs. This is achieved through a process called tagging, basically to introduce a volumetric system of reference similar to the process of printing a grating in the moiré method. This can be done by manipulating the state of magnetization of the voxels, introducing reference planes where the resulting intensities change to produce periodic signals. This operation has an effect in the quality of the images produced but with adequate software corrections can be introduced.

7.16   Tagging Procedure

The basic idea of tagging in MRI is to create families of orthogonal planes that form a 3-D spatial reference system that is embedded on a tissue and hence can follow the motion of the tissue upon deformation. Figure 7.18 shows the events that take place during the tagging of the voxels of a tissue.

In Fig. 7.18a, it is represented the initial magnetization of the volume applied in the Z-direction. In Fig. 7.18b, it is represented the direction of the magnetization vector along the X-axis as a result of the application of the $\pi/2$ pulse at $t = t_a$
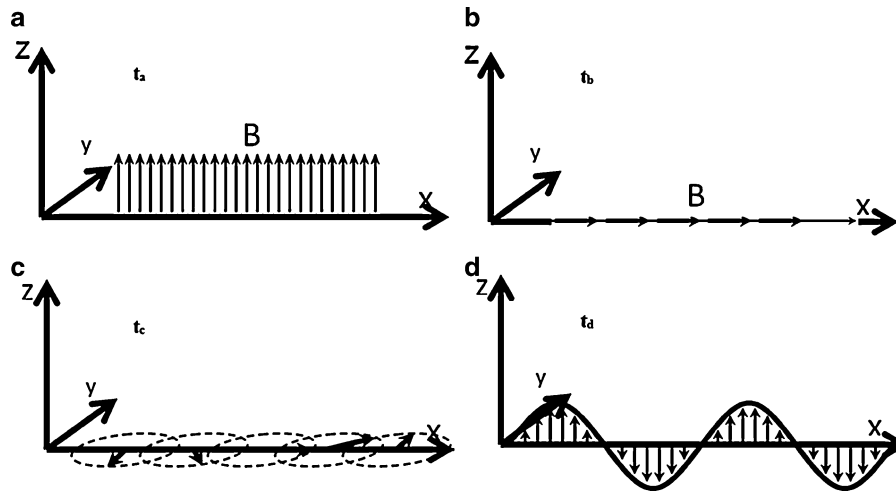
Fig. 7.18 Schematic representation of the tagging process (events take place at times t_a, t_b, t_c and t_d)

Fig. 7.19 Tagging of cardiac tissue (**a**) One directional system of lines; (**b**) System of orthogonal lines
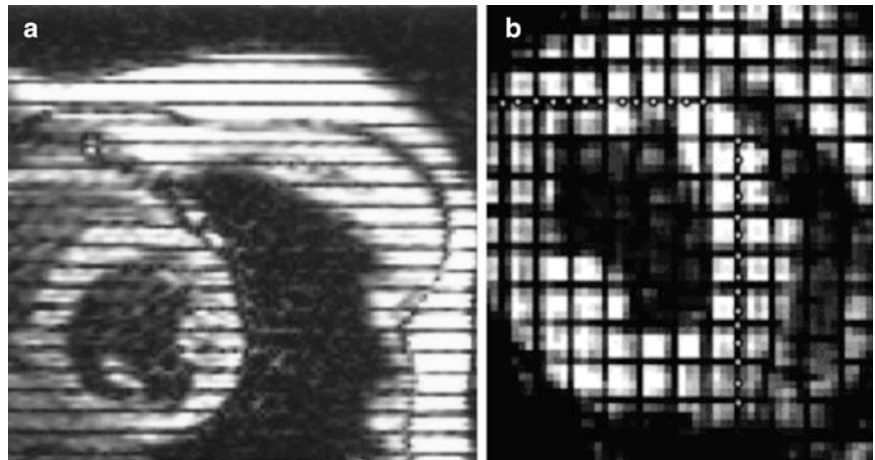
that lasts for the interval $\Delta t = t_b - t_a$. At $t = t_b$ a pulse G_y is applied. Figure 7.18c shows the rotation of the vector M_T as a consequence of the effect of G_y similarly to what is shown in Fig. 7.12.

At the time $t = t_c$, the gradient is interrupted and another pulse $\pi/2$ is applied during the interval $\Delta t = t_d - t_c$; the solution of Eq. (7.11) indicates that the magnetization M_o flips back to the Z-axis (Fig. 7.18d) and, as a consequence of the applied pulse, a sinusoidal oscillation up and down takes place along the X-axis. Although this oscillation is shown as having constant amplitude, in reality it is a damped oscillation and eventually will fade away. Then one is faced with the problem that the raster coordinate system is temporary and the measurements can only be made while the reference raster is available. Improvements can be introduced by additional manipulation of the magnetization and the corresponding software developments. To simplify representation, a sinusoidal oscillation is shown in Fig. 7.18d. However, since the tagging lines are thin lines (see Fig. 7.19), a large number of frequencies contribute to the formation of the tagging pattern.

Figure 7.19a shows a tagged cardiac tissue; the tagging takes place at the beginning of the recording cycle of the heart sequence that is being analyzed. This is the zero reference from which successive images will follow the heart muscle motion. The tagging shows the heart muscle and separates the blood in the heart cavity from the muscle that is the target of the observation. Repeating the operations and utilizing the Y-axis one can get another set of planes in the Y-direction (see Fig. 7.19b), thus creating a grid reference system equivalent to a moiré reference system that can be utilized for a complete analysis of the displacements and deformation in the different XY slices of a volume. Since the moiré method has already achieved a great deal of development, the main issues that remain to be discussed are those connected with the MRI side of the technique.

The tagging sequence produces an orthogonal system of lines of pitch $p_x = p_y = p_k$ that is determined by the maximum resolution that can be obtained from the MRI generation of signals. That is, the pitch limit is given by Eq. (7.32):

$$p_k = \frac{1}{\gamma G_z t_{rec}} = \frac{1}{k_{rec}} \tag{7.35}$$

From Eq. (7.32), it can also be seen that the gradient has to be increased to get better resolutions. The Nyquist condition requires that the pitch p_k is:

$$p_k \le 2\Delta x_i = \frac{2}{\gamma G_{x_i} T_2^*} \quad (i = 1, 2, 3) \tag{7.36}$$

where the expression has been generalized for all the three coordinates x, y, z by utilizing the indicial notation.

In order to increase the resolution, it is necessary to increase the magnitude of the gradient. The value T_2^* can be improved by improving the homogeneity of the main magnet field thus reducing the damping contribution of the field local variations.

There is another aspect that needs to be considered and it has to do with the observation of the configuration of the successive images of the heart motion as a dynamic event. In order to obtain an image that is a correct configuration at a time t of the cardiac cycle, the Whittaker-Shannon theorem must be satisfied. Considering the projected displacements in the moiré method, for example, calling *u* the displacement in the X-direction between two consecutive frames and applying the Nyquist condition, it follows:

$$u(x, t_{rec}) < \frac{2}{k_{rec}} \tag{7.37}$$

The selected tag distance and the recording times should be adjusted to satisfy Eqs. (7.36) and (7.37).

In planning a tagging protocol one should be aware of the limitations imposed by T_1 and T_2. In the case of heart, one further restriction is the motion created by breathing cycle requiring the person under study to hold their breath.

7.17 MRI Basic Hardware

Figure 7.20 provides a schematic representation of the principal components that are present in a MRI set up. The main magnet is a very powerful magnet that generates the magnetic field $\mathbf{B_o}$. The gradient coils produce gradients in the X, Y and Z-directions. The radio frequency coils produce the magnetic fields B_1 necessary to rotate the spins by $\pi/2$ or π. The RF coils also detect the signal emitted by the different voxels. The patient lies on a computer controlled table that positions the body with accuracy of the order of a mm.
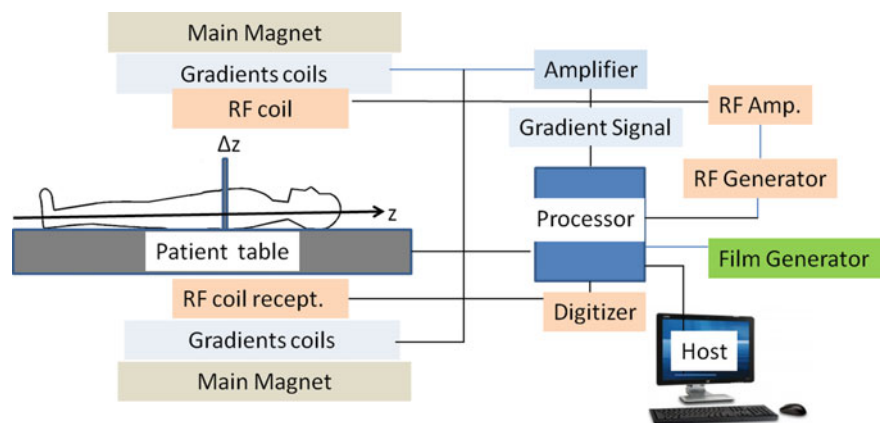
Fig. 7.20 Schematic representation of a MRI system

All the required operations to generate MRI images are performed by the processor that creates the outputs and receives the inputs of the RF coil system. The radio frequency generator produces a sinusoidal signal and the radio frequency programmer gives the signal a desired shape and the amplifier provides the voltages required by the pulse. The processor shapes the gradient signals and the corresponding amplifier provides the required voltages. An array processor is integrated in the main processor and can produce high speed FT's. The operator communicates with the processor through a host computer or a console. The device is then a complex system and what can be achieved depends on the particular design selected by the manufacturer of the system.

The intrinsic variables of the produced signals are: T_1, T_2, T_2^* and the spin density ρ. The main properties of the instrument are: the repetition time T_R, the echo time T_E, the rotation angle θ, the inversion time T_I which is the time between the inversion pulse and the corresponding sample pulse.

7.18 Experimental Determination of 3-D Deformations

In the preceding sections a basic summary of the method to generate coordinate planes in a volume of a human organ, for example a human heart, was presented. Now it is important to analyze the kinematics of 3-D deformation fields. Figure 7.21 gives the fundamental elements involved in a 3-D deformation analysis. The basic parameter to define displacements is the pitch p_o, distance between the coordinate planes, and the corresponding spatial frequency in the frequency space $k_o = 1/p_o$ (see Fig. 7.21a).

Figure 7.21b provides the descriptive geometry representation of a given volume, showing the intersections of the reference planes with the coordinate planes. These intersections form a deterministic signal, a grid that provides the information required for applying the optical methods developed to obtain displacements and the derivatives of the displacements. Figure 7.21b shows also the FT peaks that represent the corresponding families of intersections in the three coordinate planes of the frequency space assuming that the signals correspond to cosinusoidal grids.

It is necessary to clarify an important aspect of the moiré method applied to large deformations. In the present context, the generalized concept of digital moiré (see Chap. 13.8.1 of [33]) is utilized. We are operating with carrier gratings generated by the tagging process and not with the traditional moiré effect, moiré fringes resulting from the superposition effect of a grating and the deformed shape of the grating. The tagged grating is a carrier grating and one can give a parameter or order with respect to a reference coordinate system [2]. Since we are dealing with relative displacements, the zero order can be selected in the way that best fits the problem under analysis. With this procedure [2] every point in the image has an order that can be followed in the successive images no matter how large are the applied deformations [34]. It is necessary to point out that the case illustrated in Fig. 7.22 is a particular case that preserves the axis of symmetry of the deformed body. The more general case will be analyzed later on in this paper.

Figure 7.22a is a descriptive geometry representation of a given object to be analyzed. There are two reference systems (1) a global coordinates system of coordinates X_i ($i = 1, 2, 3$); (2) a system of coordinates attached to the deformed body x_i, $i = 1, 2, 3$. This system provides the orders of the different grid lines that will provide the Eulerian coordinates of the points of the deformed body.
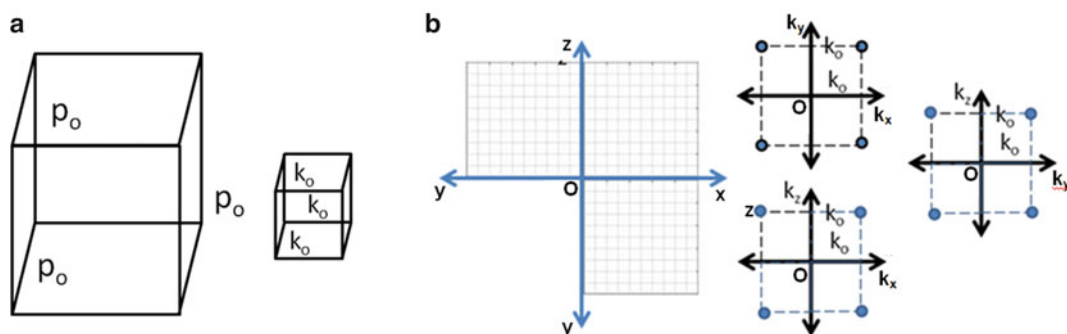
Fig. 7.21 Basic geometrical elements to describe deformations in the 3-D space (**a**) Definition of pitch and spatial frequency; (**b**) Descriptive geometry representation of a given volume and FT representation of frequency peaks of the intersections of families of reference planes
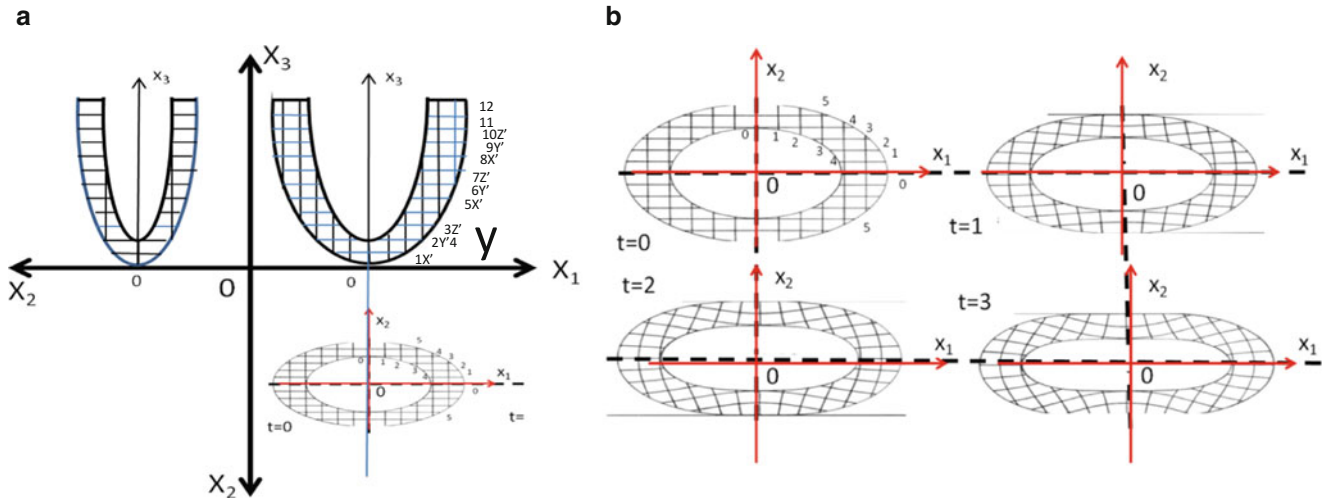
Fig. 7.22 (**a**) Descriptive geometry representation of a 3-D body whose volume has been tagged with an orthogonal system of carrier planes represented in 3-D by orthogonal lines; (**b**) Deformed shape in time of a cross-section

A point P in the initial configuration is identified by its coordinates X_{P1}, X_{P2} and X_{P3} with respect to a global system of coordinates X_i. The objective of the deformation analysis is to follow the trajectories of the points of the body and provide the corresponding strains as a function of a parameter that grows monotonically, for example the time.

7.19 Tracking Point Displacements in 3-D

Since we are looking at large deformations, there will be considerable changes on the geometry of the analyzed specimen. Hence, the Eulerian description of the continuum is utilized. There are several procedures that can be followed to track a point located inside a 3-D body or at the boundary of the body. At a given point of the continuum the displacement of the point is encoded in gray levels. The position and the displacement can be connected to the notion of fringe order n, where n is a real number (see Chaps. 13.2 and 13.3 of [33]). One has to remember that coordinate positions can be transformed into angular variables through the concept of phase. The actual value of the phase at a given point depends on the selection of the marker corresponding to the order zero, which provides the reference from which the orders are determined. Since we are referring to relative displacements, the zero order location is arbitrary and is selected by the operator analyzing the deformation process. The usefulness of the transformation of distances into angles is based on an important relationship that separates the levels of gray at a given point into two variables through a procedure that is applied in digital signal processing. The signal generated by the recording system can be represented by a phasor, similarly to what was discussed in the case of Eq. (7.1):

$$\vec{\vec{I}}\,(\mathbf{x},t) = I(\mathbf{x},t)e^{i\phi(\mathbf{x},t)} \tag{7.38}$$

where: $\vec{\vec{I}}\,(\mathbf{x},t)$ represents a gray level generated by the recording system at a given voxel and a given time t; $I(\mathbf{x},t)$ is the amplitude of the phasor in gray levels; $\mathbf{x}$ is the Eulerian position vector of a voxel at the time t; $\phi(\mathbf{x}, t)$ is an angular variable that encodes displacement information. Since $I(\mathbf{x}, t)$ depends on many factors that are independent of the motion, following the approach of the vast majority of optical methods, the amplitude of the phasor is disregarded and the displacement is related to the phase trough the relationship:

$$u(\mathbf{x},t) = \frac{\phi(\mathbf{x},t)}{2\pi}p \tag{7.39}$$

In 3-D, the following vector equation is satisfied by the displacement vector (Fig. 7.23a):

$$\mathbf{u}(\mathbf{x},\Delta t) = \mathbf{u}_1(\mathbf{x},\Delta t) + \mathbf{u}_2(\mathbf{x},t_m) + \mathbf{u}_3(\mathbf{x},\Delta t) \tag{7.40}$$
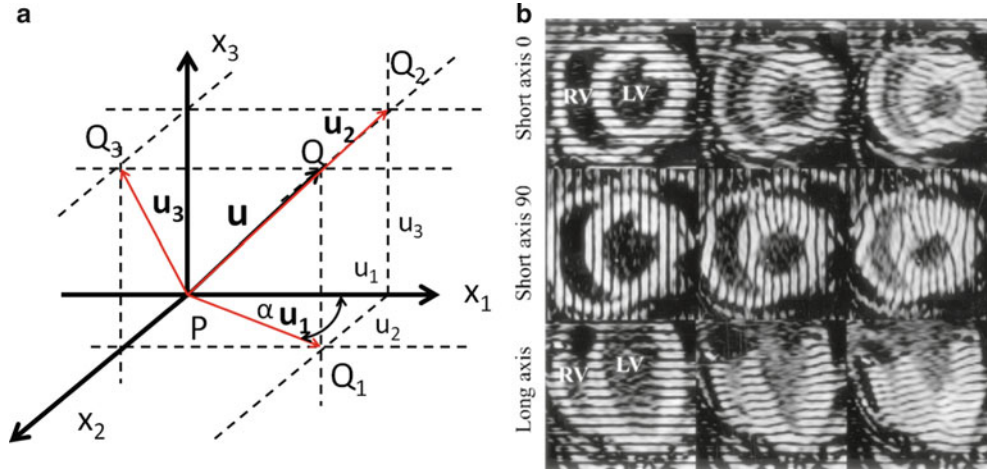
Fig. 7.23 (a) Relative displacement between two points at times t_m and t_{m+1}; (b) Tags required to find the three components of the displacement vector of cardiac tissue (MR images are taken from [31])

where the $\mathbf{u}_i(\mathbf{x}, t_m)$ are the components of the displacement vector in each of the reference planes at some reference time interval $\Delta t = t_{m+1} - t_m$. These component vectors satisfy the following equations:

$$
\begin{cases}
\mathbf{u}_1(x, \Delta t) = u_1(x_1, x, \Delta t)\, \mathbf{e}_1 + u_2(x_1, x, 2\Delta t)\, \mathbf{e}_2 \\
\mathbf{u}_2(x, \Delta t) = u_1(x_1, x_2, \Delta t)\mathbf{e}_1 + u_3(x_1, x_3, \Delta t)\, \mathbf{e}_3 \\
\mathbf{u}_3(x, \Delta t) = u_2(x_2, x_3, \Delta t)\, \mathbf{e}_2 + u(x_2, x_3, \Delta t)\, \mathbf{e}_3
\end{cases}
\tag{7.41}
$$

where the $\mathbf{e}_i$ vector are the versors of the coordinate axes.

The components u_i are given by:

$$
\begin{cases}
u_1(x_1, x_2, \Delta t) = \dfrac{\phi_1(x_1, x_2, \Delta t)}{2\pi} p = n_1(x_1, x_2, \Delta t)\, p \\[2mm]
u_2(x_1, x_2, \Delta t) = \dfrac{\phi_2(x_1, x, \Delta t)}{2\pi} p = n_2(x_1, x_2, \Delta t)\, p \\[2mm]
u_3(x_2, x_3, {}^3\Delta t) = \dfrac{\phi_3(x_2, x_3, \Delta t)}{2\pi} p = n_3(x_2, x_3, \Delta t)\, p
\end{cases}
\tag{7.42}
$$

where the n_i are the fringe orders. The above equations have a direct interpretation in Fig. 7.23b (MR images are taken from [31]). The volume under analysis is tagged in two orthogonal directions in the plane x_1-x_2 and in the horizontal direction along the x_3 axis. From this system of lines it is possible to get the three vector components of Eqs. (7.40), (7.41) and (7.42) and track the motions of the points in the three projection planes. All the different strain tensors in the Eulerian description can be obtained from the tensor usually called [J] in Continuum Mechanics [2]

$$
[J] =
\begin{bmatrix}
\dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \dfrac{\partial u_1}{\partial x_3} \\[3mm]
\dfrac{\partial u_2}{\partial x1} & \dfrac{\partial u_2}{\partial x2} & \dfrac{\partial u_2}{\partial x_3} \\[3mm]
\dfrac{\partial u_3}{\partial x_1} & \dfrac{\partial u_3}{\partial x2} & \dfrac{\partial u_3}{\partial x_3}
\end{bmatrix}
\tag{7.43}
$$

These derivatives can be determined directly from the wrapped phase (see Chap. 13.8.2.1 of [33]) obtained from the displacements components in the planes x_1-x_3, x_1-x_3, x_2-x_3.

7.20 Image Processing

With the information gathered in the different planes one can get a full description of the 3-D kinematics. In [2], the different families of lines that define the kinematics of the continuum in 2-D have been described. These families can be generalized to three dimensions thus providing a complete picture of the large deformations as a function of a reference parameter, for example the time t. One of the main purposes of the MRI technique is to provide a tool of diagnosis of heart diseases. The values of the principal secondary strains at individual planes provide enough comparison basis for the purposes of diagnosis.

As said before, image generation depends on the particular device that one has as at disposal and this depends on the particular brand of the instrument. The post-processing is the subject of main interest to us in this paper. In [2], the different techniques to retrieve displacement information from images where this information is given as levels of gray in pixels (2-D) or voxels (3D) are presented. The HARP method is a version of digital moiré that operates directly on the carrier signal, hence utilizes spatial frequency filtering to get harmonic components. In place of utilizing projected displacements the HARP method adopts the same assumption as the digital image correlation (DIC) and searches for the vector displacement in place of the projections as the moiré does instead. An important aspect of HARP is that material points are tracked from one image to a subsequent image. A number of points are used as reference points. Since large deformations and rigid body rotations are involved, certain directions and reference circles are utilized to position the reference points. If the conditions postulated by Eq. (7.37) are satisfied, and the changes experienced by the observed body are such that the condition of preservation of the axis of symmetry—if there is one system or the centroidal axis of the cross-sections—is satisfied within the limits of resolution of the system, the condition of self reference can be utilized to analyze subsequent images. The general case is more complex and more general transformations will be involved. If condition (7.37) is satisfied one has to relate each image to the subsequent image and general equations of transformation of coordinates must be applied to locate points in successive images.

In all these cases in which point tracking is not necessary, one can utilize global rigid body translations and rotations to match successive images in the plane. In each image one can obtain the relative displacements and the Eulerian strains [2]. By writing software to transform Eulerian description to Lagrange description, values corresponding to material points can be computed. It is necessary to remember that the nonlinear strain tensors are invariant upon rigid body motions and that displacements can be obtained by strain integration [2].

7.21 Comparison of the Results of HARP and Digital Moiré

To compare results obtained with the cine HARP method and the Holo Moiré Strain Analyzer (Holostrain) program [35], images presented in [31, 32] are analyzed. Due to the lack of a reference image pattern of Fig. 1 of [32] the first pattern of the above mentioned sequence of patterns is utilized as reference, Fig. 7.24a; this image is very close to the initiation of the cycle. To calibrate the pixel size (about 700 μm) the base line red circle was measured and referred to standard values in the literature. The data processing was done in the frequency plane utilizing the Holostrain program. Figure 7.24b shows the phase maps of reference and modulated patterns along coordinate directions X (Fig. 7.24b, c, respectively) and Y (Fig. 7.24d, e, respectively).

By subtracting the phase of the deformed configuration from the phase of the reference configuration considered for each displacement direction, the phase patterns shown in Fig. 7.25a, b (respectively, for u and v-displacements) were obtained. Each phase difference pattern was modulated by adding a carrier including 64 lines and then in-quadrature filtered. Fringe extension and masking operations (see Chap. 10.6 of [33]) allowed patterns to be processed smoothly. Phase maps were finally unwrapped to obtain the displacement maps u and v, respectively, in coordinate directions X and Y (see Fig. 7.25c, d, respectively, for displacement components u and v).

The total displacement $\mathbf{u}_r$ was computed as follows:

$$\|\mathbf{u}_r(\mathbf{x}, t_m)\| = \sqrt{u^2(x, y, t_m) + v^2(x, y, t_m)} \tag{7.44}$$

The corresponding color map of the isokinetic pattern lines of equal displacement is plotted in Fig. 7.26.

Eulerian strains were computed using the simplified Almansi strain tensor. For e_x^E, it is [34]:

$$e_x^E = \frac{\partial u}{\partial x} - \frac{1}{2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2\right] \tag{7.45}$$
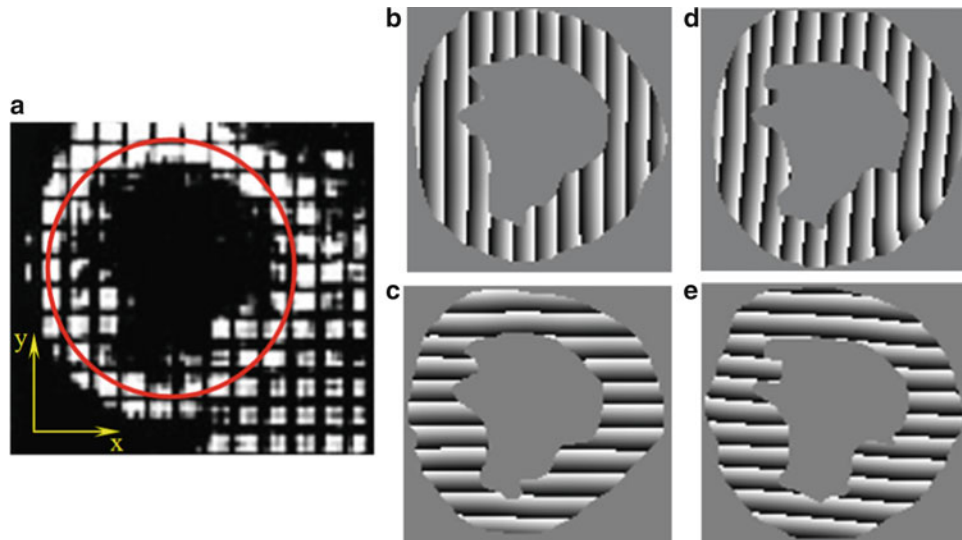
Fig. 7.24 (**a**) Tagged image of the lower ventricle short-axis [32] with indication of reference system and heart baseline utilized in strain determinations; (**b**) Phase of the reference image and (**c**) modulated image involved in the determination of u-displacements; (**d**) Phase of the reference image and (**e**) modulated image involved in the determination of v-displacements

Fig. 7.25 Unwrapped phase maps determined by Holostrain for u (**a**) and v-displacement components (**b**); displacement components u (**c**) and v (**d**), respectively, computed in the coordinate directions x and y
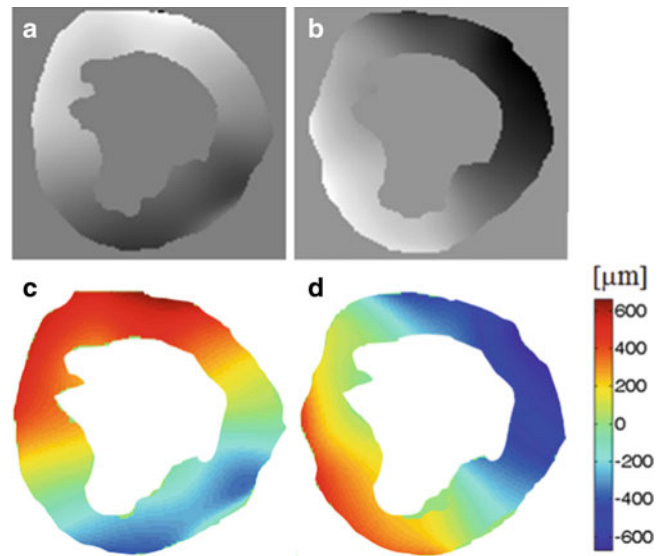
Fig. 7.26 Isokinetic map of relative displacement vector in the cardiac wall
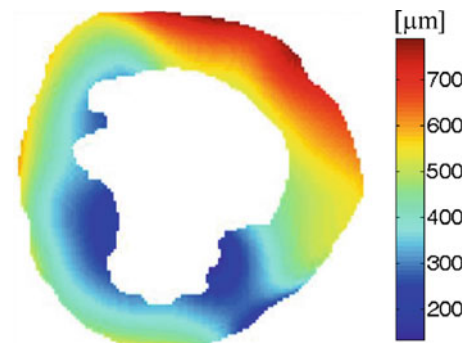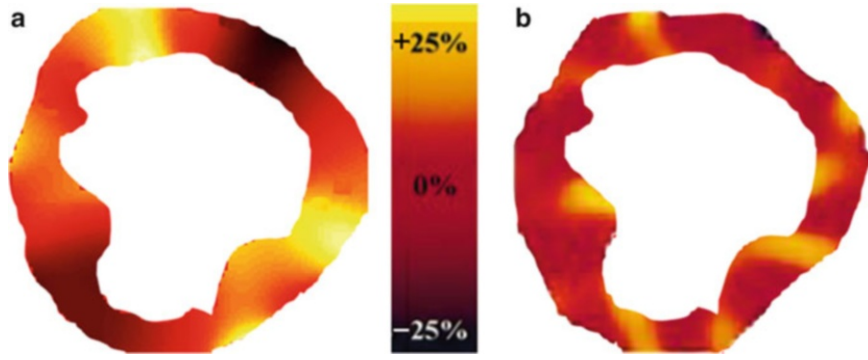
Fig. 7.27 (**a**) Strains of cardiac wall obtained in the present study; (**b**) Strains as given in [32]. *Yellow* indicates positive strains (stretching), *black* indicates negative strains (shortening). The scale of the strains is the same scale given in [32]. Only regions where the signal could be recovered accurately enough are compared in the figure
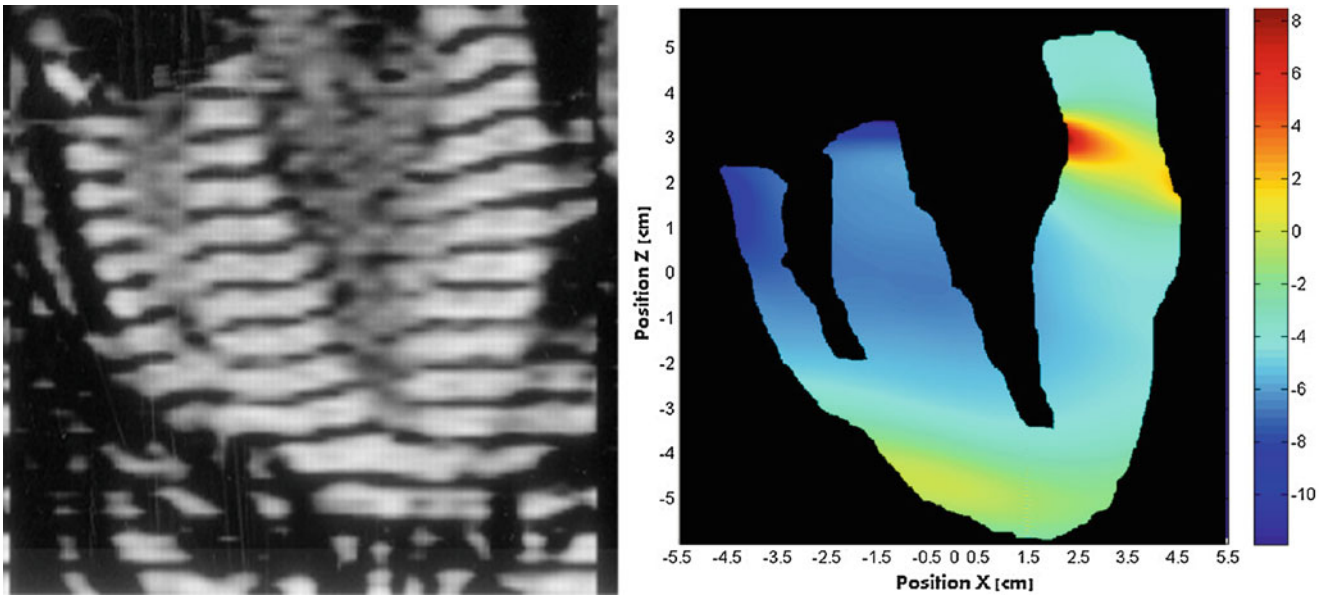
Fig. 7.28 Right and left ventricles strains (%) in the Z-direction (i.e. long axis of Fig. 7.23b)

Similar expressions were used for e_y^E and for the shear strain e_{xy}^E [34]:

$$e_y^E = \frac{\partial v}{\partial y} - \frac{1}{2}\left[\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right] \quad e_{xy}^E = \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) - \frac{1}{2}\left[\frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\frac{\partial v}{\partial y}\right] \tag{7.46}$$

Figure 7.27 compares the Eulerian strains in the circumferential direction computed with the above described procedure (see Fig. 7.27a) with the cine-HARP results shown in [32] (see Fig. 7.27b).

The HARP derivatives are computed using finite differences, while the Holostrain derivatives are obtained from actual differentiation in the frequency space. In spite of the difference of the procedures utilized the agreement is satisfactory. It is possible to see that in the comparison areas orders of magnitude and signs of the deformations are consistent. Since the images represent contraction of the heart, most of the heart wall must be in compression, with sections in the diametrical direction that should experience traction.

Finally a cross section of a left ventricle during the heart contraction is analyzed in Fig. 7.28 (MR image is taken from [31]). There is no HARP analysis available for this particular picture. The strain distribution shows practically an almost uniform contraction (about 10 %) in the direction of the long axis with small regions in tension.

7.22   Discussion and Conclusions

The whole development of tagging and the HARP methodology are extremely important contributions to the 3-D analysis of deformations. It applies to media that are transparent to magnetic fields and contain isotopes that can support the phenomenon of MR. This phenomenon has a similitude with the stimulated emission phenomenon in optics. The process of creating images from magnetic field information is in itself an extraordinary scientific and technological achievement. The addition of the tagging process to the image generation in media that are transparent to magnetic fields provides an invaluable tool to Experimental Mechanics that has been constrained for a long time in its main developments to the analysis of 2-D fields or 3-D optically transparent media. These limitations are due to the lack of methodologies to get volumetric information without limitations of volume size as it occurs with other types of electromagnetic fields such as X-rays based on the interaction of electromagnetic photons of high frequency with valence electrons, or the electron microscopy based on the interaction of electrons (charged particle) with atoms. In both cases adsorption of the utilized radiation limits the volumes that can be observed.

From the point of view of the kinematics of the continuum the existing methods that are based on displacement and displacement derivatives information determination provide the necessary tools to tackle 3-D problems for small or large deformations.

It is also true that the MR technology is based not only in notable scientific developments but also in a very sophisticated technology both in hardware and software leading to complex and costly equipment.

References

1. Sciammarella CA, Lamberti L, Sciammarella FM, Boccaccio A (2014) The kinematics and dynamics of 3-D displacement fields. In: Jin H, Sciammarella CA, Yoshida S, Lamberti L (eds) Advancement of optical methods in experimental mechanics. Springer, New York, NY, pp 43–67, Chapter 7
2. Sciammarella CA, Lamberti L (2014) Basic models supporting experimental mechanics of deformations, geometrical representations, connections among different techniques. To appear in Meccanica, Special Issue on "Experimental Solid Mechanics: in honour of Professor Emmanuel Gdoutos". Early View 8 January 2014. http://link.springer.com/article/10.1007%2Fs11012-013-9867-8
3. Axel L, Dougherty L (1989) Heart wall motion: improved method of spatial modulation of magnetization for MR imaging. Radiology 172(2):349–350
4. Axel L, Dougherty L (1989) MR imaging of motion with spatial modulation of magnetization. Radiology 171:841–845
5. McVeigh E (1997) MRI of myocardial function: motion tracking techniques. Magn Reson Med 14(2):137–150
6. Fischer SE, McKinnon GC, Scheidegger MB, Prins W, Meier D, Boesiger P (1994) True myocardial motion tracking. Magn Reson Med 31(4):401–413
7. Doyle M, Walsh EG, Foster RE, Pohost GM (1997) Common k-space acquisition: a method to improve myocardial grid-tag contrast. Magn Reson Med 37(5):754–763
8. Zhang S, Douglas M, Yaroslavsky L, Summers R, Dilsizen V, Fananapazir L, Bacharach S (1996) A Fourier based algorithm for tracking spam tags in gated magnetic resonance cardiac images. Med Phys 23(8):1359–1369
9. Liao JR, Pauly JM, Brosnan TJ, Pelc NJ (1997) Reduction of motion artifacts in cine MRI using variable-density spiral trajectory. Magn Reson Med 37(4):569–575
10. Guttman MA, Prince JL, McVeigh ER (1994) Tag and contour detection in tagged MR images of the left ventricle. IEEE Trans Med Imaging 13(1):74–88
11. Kumar S, Goldgof D (1993) Automatic tracking of SPAMM grid and the estimation of deformation parameters from cardiac MR images. IEEE Trans Med Imaging 13(1):122–132
12. O'Dell WG, Moore CC, Hunter WC, Zerhouni EA, McVeigh ER (1995) Three-dimensional myocardial deformations: Calculations with displacement field fitting of tagged MR images. Radiology 195:829–835
13. Denney TS, Prince JL (1995) Reconstruction of 3-D left ventricular motion from planar tagged cardiac MR images: an estimation theoretic approach. IEEE Trans Med Imaging 14(4):625–635
14. Radeva P, Amini A, Huang J (1997) Deformable B-solids and implicit snakes for localization and tracking of SPAMM MRI-data. Comput Vis Image Underst 66(2):163–178
15. Denney TS (1997) Identification of myocardial tags in tagged MR images without prior knowledge of myocardial contours. In: Inform. Processing in medical imaging—Lecture Notes in Computer Science 1230:327–340
16. Prince JL, McVeigh ER (1992) Motion estimation from tagged MR image sequences. IEEE Trans Med Imaging 11(2):238–249
17. Gupta SN, Prince JL (1995) On variable brightness optical flow for tagged MRI. Technical Report 95-13, JHU/ECE
18. Gupta SN, Prince JL, Androutsellis-Theotokis S (1997) Bandpass optical flow for tagged MR imaging. In: Proceedings of the IEEE international conference on image processing, vol 3, Santa Barbara, CA, October 1997, pp 364–367
19. Pelc NJ, Herfkens RJ, Shimakawa A, Enzmann DR (1991) Phase contrast cine magnetic resonance imaging. Magn Reson Q 7(4):229–254
20. Pelc NJ, Drangova M, Pelc LR, Zhu Y, Noll DC, Bowman BS, Herfkens RJ (1995) Tracking of cyclic motion with phase-contrast cine MR velocity data. J Magn Reson Imaging 5(3):339–345

21. Constable RT, Rath KM, Sinusas AJ, Gore JC (1994) Development and evaluation of tracking algorithms for cardiac wall motion analysis using phase velocity MR imaging. Magn Reson Med 32(1):33–42
22. Meyer FG, Constable RT, Sinusas AJ, Duncan JS (1996) Tracking myocardial deformation using phase contrast MR velocity fields: a stochastic approach. IEEE Trans Med Imaging 15(4):453–465
23. Wedeen VJ (1992) Magnetic resonance imaging of myocardial kinematics. Technique to detect, localize and quantify the strain rates of the active human myocardium. Magn Reson Med 27(1):52–67
24. Robson MD, Constable RT (1996) Three-dimensional strain-rate imaging. Magn Reson Med 36(4):537–546
25. Perman WH, Creswell LL, Wyers SG, Moulton MJ, Pasque MK (1995) Hybrid DANTE and phase-contrast imaging technique for measurement of three-dimensional myocardial wall motion. J Magn Reson Imaging 5(1):101–106
26. Osman NF, Kerwin WS, McVeigh ER, Prince JL (1999) Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging. Magn Reson Med 42(6):1048–1060
27. Zerhouni EA, Parish DM, Rogers WJ, Yang A, Shapiro EP (1988) Human heart: tagging with MR imaging—a method for noninvasive assessment of myocardial motion. Radiology 169(1):59–63
28. Nguyen TD, Reeves SJ, Denney TS (1998) New magnetic resonance tagging technique for directly measuring the strain tensor of the in vivo human heart. In: Proceedings of IEEE conference on image processing, Chicago, IL, October 1998
29. Moon-Ho Song S, Napel S, Pelc NJ, Glover GH (1995) Phase unwrapping of MR phase images using Poisson equation. IEEE Trans Image Process 4(5):667–676
30. Waks E, Prince JL, Douglas AS (1996) Cardiac motion simulator for tagged MRI. In: Proceedings of workshop on mathematical methods in biomedical image analysis, San Francisco, CA, pp 182–191
31. Wyman BT, Hunter WC, Prinzen FW, McVeigh ER (1999) Mapping propagation of mechanical activation in the paced heart with MRI tagging. Am J Physiol 276:H881–H891
32. Garot J, Bluemke DA, Osman NF, Rochitte CE, McVeigh ER, Zerhouni EA, Prince JL, Lima JAC (2000) Fast determination of regional myocardial strain fields from tagged cardiac images using harmonic phase MRI. Circulation 101:981–988
33. Sciammarella CA, Sciammarella FM (2012) Experimental mechanics of solids. Wiley, Chichester
34. Durelli AJ, Parks VJ (1970) Moiré analysis of strain. Prentice-Hall, Englewood Cliffs, NJ
35. General Stress Optics Inc. Holo-Moiré Strain Analyzer Software HoloStrain Version 2.0. General Stress Optics Inc., Chicago, IL, http://www.stressoptics.com

Chapter 8
Bi-Directional Displacement Measurement by Speckle Interferometry Immune to Random Vibration

Shuichi Arikawa and Satoru Yoneyama

Abstract A bi-directional in-plane displacement measurement by speckle interferometry using an optimum image extraction under environmental disturbances as random vibrations is investigated. A compact speckle interferometer which has 4 laser beams for bi-directional sensitivity is constructed on a tripod. The polarizations of laser beams of each sensitive direction are arranged orthogonal and diffuse reflected laser beams from an aluminum target can be separated by a polarizing beam splitter in front of digital cameras. The bi-directional in-plane displacement measurement is made possible by a simultaneous capturing of the speckle images for each sensitive direction. An in-plane rotation of the target is measured by the proposed method in a presence of the random vibration. A lot of images are captured at the initial and the rotated state. Then, optimum images which can make interference fringes are extracted. A phase analysis by random phase-stepping method using extracted images is performed. As a result, the analyzed phase difference maps show a good agreement with the given rotation angle. Therefore, it is validated that the bi-directional displacement measurement by speckle interferometry under environmental disturbance is possible.

Keywords Deformation measurement • Speckle interferometry • Optimum image extraction • Environmental disturbance • Polarization

8.1 Introduction

Speckle interferometry, also called electronic speckle pattern interferometry [1, 2] is a most sensitive measurement technique for full-field deformation measurements of objects surfaces. Speckle interferometry has been used for various deformation measurements [3–11]. The high sensitiveness of deformation measurements by speckle interferometry provides some non-destructive tests [12–14]. However, it is difficult to use interferometric techniques for the measurements without vibration isolators in the presence of the environmental disturbance coming from the ground or the floor, because of the sensitiveness. As a solution to this problem, it can be considered that the application of high-speed or dynamic speckle interferometry [6–9] is possible. In this situation, it is required that the exposure time is short enough to freeze the object. When the object speed is high, it has been known that the introduction of a high-power laser is required, to gain the sufficient intensity during the exposure time. Additionally, it leads relatively high costs and requirements of some safety equipments. These difficulties impede applications of speckle interferometry to various fields, such as pressure vessels in power plants.

Static deformation measurements by speckle interferometry under the environmental disturbance without a vibration isolator have been investigated [15, 16]. The measurement has been made possible by an optimum image extraction. In this technique, a lot of speckle images, hundreds or thousands, are captured at both states of the before and the after deformation, and optimum images which can make interference fringes are extracted before making interference fringes or a phase analysis. The image extraction is based on an evaluation of speckle pattern contrasts in the time axis value [15]. Since the significant reduction of the exposure time is not required for image recording, a middle-class-power laser can be available for the interferometer. Hence a compact interferometer which can set on a tripod can be constructed. Random phase-stepping

S. Arikawa (✉) • S. Yoneyama
Department of Mechanical Engineering, Aoyama Gakuin University, 5-10-1 Fuchinobe, Chuo-ku, Sagamihara, Kanagawa 252-5258, Japan
e-mail: arikawa@me.aoyama.ac.jp

method [17] is effective for this technique using some extracted images for each state [16]. Therefore, static deformation measurements by speckle interferometry under the environmental disturbance have been possible by these studies. However, only single directional measurements have been studied. So, an investigation of a bi-directional in-plane displacement measurement is required to improve the usability of the technique. Additionally, a simultaneous image capturing is suitable to ensure the reliability of the bi-directional measurement, because there is some possibility that displacement fields measured at different moment are not the same, even if the object condition is static.

In this study, a bi-directional in-plane displacement measurement by speckle interferometry using an optimum image extraction under environmental disturbances as random vibrations is investigated. A compact speckle interferometer which includes a middle-class-power laser is constructed on a tripod. Polarizations of laser beams of each sensitive direction are arranged orthogonal, and diffuse reflected laser beams from a target are separated by a polarizing beam splitter in front of digital cameras. Therefore, the bi-directional in-plane displacement measurement is made possible by a simultaneous capturing of the speckle images for each sensitive direction by the two cameras. An in-plane rigid body rotation of an object is measured by the proposed method under the random vibration to validate the method. A lot of speckle images are captured at the initial and the rotated state, and optimum images which can make interference fringes are extracted before a phase analysis. Random phase-stepping method using extracted images is applied for the phase analysis. As a result of the evaluation of the measured rotation angle by the proposed method, the measured rotation angle shows a good agreement with the given rotation angle. Therefore, it is expected that deformation measurements by speckle interferometry are available for various environments.

8.2 Speckle Interferometry with Optimum Image Extraction

Speckle interferometry uses laser speckle patterns as shown Fig. 8.1. When a coherent laser beam illuminates a rough surface, the speckle pattern is formed on an image plane in a camera by random reflection of a coherent laser beam from the surface. Figure 8.2a shows a typical interferometer for in-plane displacement measurement called dual-beam interferometer. When expanded two laser beams separated from a single light source illuminate a rough surface of an object, the spackle image are captured by a digital camera placed in front of the surface. If a point of the surface move in horizontal axis, the intensity of corresponding point of the speckle pattern fluctuates by changing the phase difference of the two beams. Interference fringe patterns can be obtained by an image subtraction of speckle images captured at an initial and a deformed state. Then, each speckle displacement must be enough smaller than the average speckle diameter. The average speckle diameter d_s is calculated as

$$d_s = 2.44\lambda F \tag{8.1}$$

where λ is the wavelength of the laser beam and F is the F-number of the camera lens. As shown in Fig. 8.2b, when the displacement becomes larger than the diameter, the overlap region of the initial and the displaced speckles cannot be obtained. Then, the phase difference in the overlap region cannot be calculated. Furthermore, it is desirable that the displacement is to be less than 1/4 of the diameter for obtaining high contrast interference fringes [5]. In a certain condition,
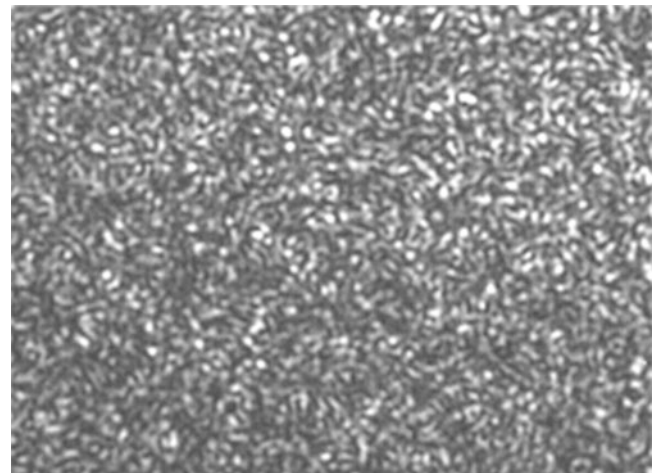
Fig. 8.1 Typical laser speckle pattern

Fig. 8.2 Schematic figure of speckle interferometry (**a**) dual-beam interferometer, (**b**) relation of speckle displacement and the overlap region
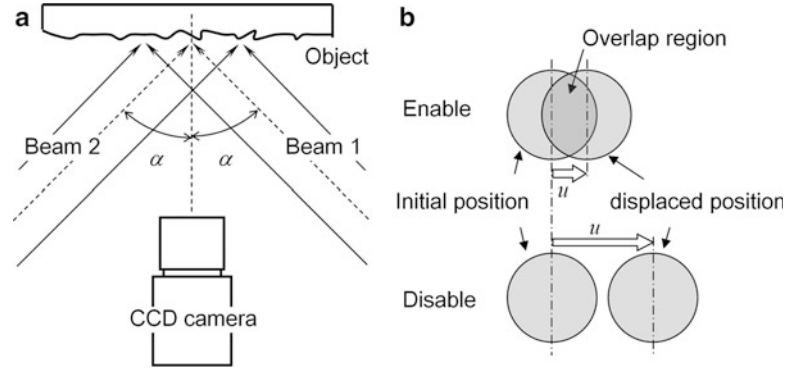
if the displacement becomes larger than the diameter, the interference fringe can be obtained [18, 19] by a speckle position correction using digital image correlation. Intensities of spackle pattern, I, is written as

$$I = A + B \cos \varphi \tag{8.2}$$

where A is the average intensity, B is the modulation factor, and φ is the phase difference of the two beams. The intensity of the subtracted image I_{sub} is calculated as

$$
\begin{aligned}
I_{\text{sub}} &= |I_{\text{d}} - I_{\text{u}}| \\
&= |2B \sin (\varphi + \varDelta\varphi/2) \sin (\varDelta\varphi/2)|
\end{aligned}
\tag{8.3}
$$

where I_{u} and I_{d} are the initial and the deformed state intensity, and the $\varDelta\varphi$ is the phase change of the deformed state from the initial state. The relation between the phase change $\varDelta\varphi$ and the displacement component u which lies on sensing direction is given as

$$\varDelta\varphi = (2\pi/\lambda)\, u\, 2 \sin \alpha \tag{8.4}$$

where α is the incident angle of the beams as shown in Fig. 8.2a.

When speckle images are captured under random vibrations, there is a possibility that the static deformation is measured by speckle interferometry, even if the exposure time is not so short. If large number images of an initial and a deformed state are captured, some optimum image sets of the initial and the deformed state which can make interference fringes are included in them. However, the calculation of all patterns of these image sets is not realistic because of the large quantity of data and the calculation time. Therefore, the extraction of the optimum image set is required. The extraction is based on an evaluation of speckle pattern contrasts in a time axis. For some number of pixels in each image, the sum of the absolute deviations of the intensity around the time average E_i are calculated as [15]

$$E_i = \sum_{j=1}^{n} \left| I_{ij} - \overline{I_j} \right| \tag{8.5}$$

where i and j are the image index of time series and the pixel index on each image respectively. I_{ij} is the intensity value of the image number i at the pixel number j, and $\bar{I}_j$ is the time average value at the pixel j. Then, the largest sum E_{max} may indicate the optimum image. Phase analysis is possible by subtraction-addition method [20] or random phase-stepping method [17]. In this technique, random phase-stepping method is effective using some extracted images for each state.

8.3 Bi-Directional Measurement

For the bi-directional measurement with the simultaneous image capturing of each direction, polarizations of laser beams of each sensitive direction are arranged orthogonal, and diffuse reflected laser beams from a target are separated by a polarizing beam splitter in front of digital cameras. The schematic figure of the interferometer is shown as Fig. 8.3. An object surface is

Fig. 8.3 Schematic figure
of bi-directional sensitive
interferometer
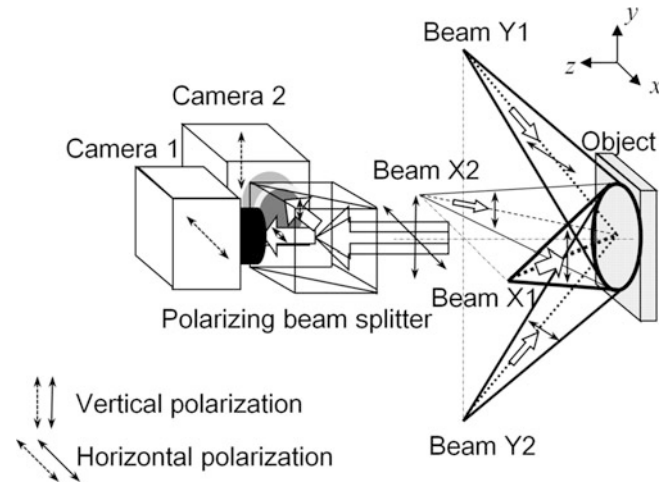
illuminated by four expanded laser beams, horizontal X1 and X2, and vertical Y1 and Y2. Polarizations of laser beams of each sensitive direction are arranged orthogonal. The polarization of the horizontal beams, X1, X2, is set to vertical, and the polarization of the vertical beams, Y1, Y2, is set to horizontal. Then diffuse reflected laser beams from a target are separated by a polarizing beam splitter in front of digital cameras, and each separated component of the horizontal and the vertical polarization is captured by camera 1 and camera 2 respectively. In general, it is known that polarizations are changed by reflections, because of changing the amplitude and the phase of the laser light. However, the effect of the changing polarization is small in some materials. Thus, the bi-directional measurement with the simultaneous image capturing of each direction is possible for such materials.

8.4 Experiment

A compact speckle interferometer which can measure bi-directional in-plane displacements is constructed on a tripod. Incident angles of laser beams are set to $11°$ for each direction. A constant wave 300 mW DPSS (diode pumped solid state) laser (KLASTECH, Scherzo) with the wavelength of 532 nm is used. The laser power is reduced by a ND filter. The monochrome CCD cameras (Allied Vision Technologies, Guppy pro F125B) have a resolution of 1292×964 pixels, cell size of $3.75 \ \mu m \times 3.75 \ \mu m$ and a maximum frame rate of 31 frames/s. The F-number of the camera lens is set to 4. The exposure time is set to 1.7 ms. An in-plane rigid body rotation target made from an aluminum plate is used for an observation target. The surface of the aluminum plate is polished by an emery paper of #1000. The target is set on a vibration isolator in front of the compact interferometer attached on the tripod standing on the floor. An environmental disturbance as a vibration condition of the experimental space is measured in advance. As the result, the vibration is random and has $2 \sim 20$ Hz in the frequency and a few μm in the amplitude. In this condition, a rotation of the target is measured by proposed method. 100 speckle images for each camera are captured at the initial and the rotated state. The rotation angle is $0.00880°$.

8.5 Results and Discussion

Captured speckle images with various conditions are shown in Fig. 8.4a–f. The area size in the image is 39.8 mm $\times$ 31.9 mm (720×576 pixels). The speckle images captured by camera 1 and 2 with bi-directional illumination, horizontal X1, X2 and vertical Y1, Y2, are shown in Fig. 8.4a, b. In the image by camera 1, we arranged that the vertical component of the illumination is captured. Also, camera 2 is arranged to capture the horizontal component of the illumination. The speckle

Fig. 8.4 Captured speckle images (**a**) by camera 1 with bi-directional illumination, (**b**) by camera 2 with bi-directional illumination, (**c**) by camera 1 with horizontal illumination, (**d**) by camera 2 with horizontal illumination, (**e**) by camera 1 with vertical illumination, (**f**) by camera 2 with vertical illumination
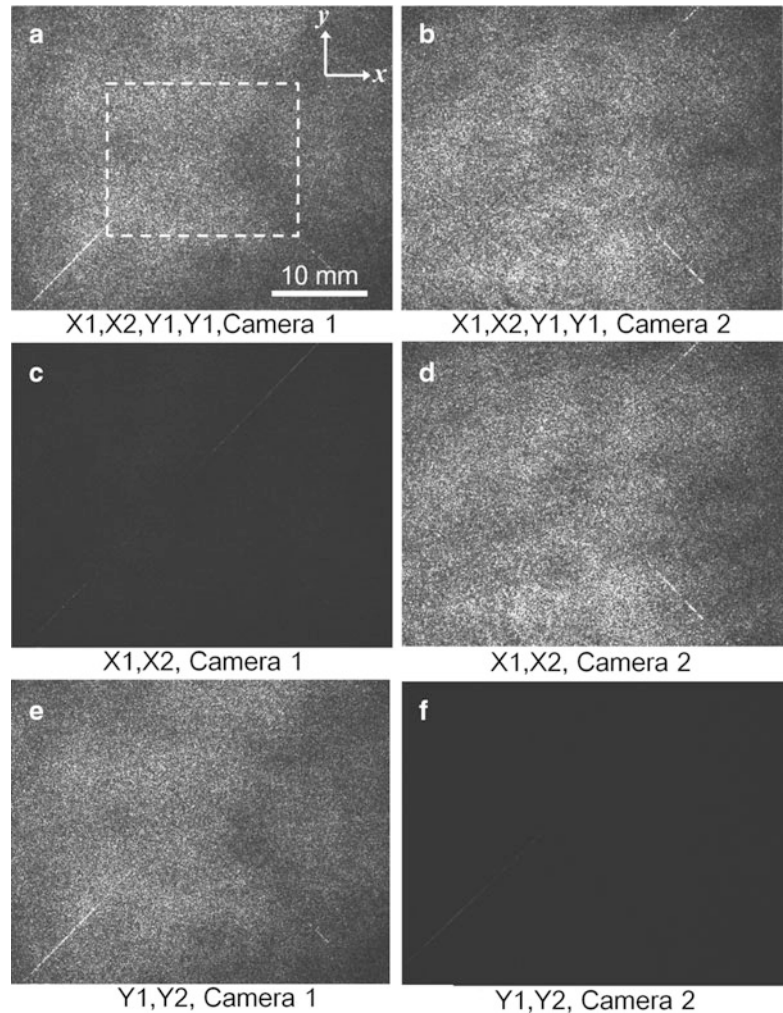
images captured by camera 1 and 2 with only horizontal illumination, X1, X2, are shown in Fig. 8.4c, d. The speckle image by camera 1 is dark and that by camera 2 is bright. The speckle images with only vertical illumination, Y1, Y2, are shown in Fig. 8.4e, f. The speckle image by camera 1 is bright and that by camera 2 is dark. Therefore, the change of the laser polarization by the reflection at the target surface is small in this condition.

The phase analysis is performed using 4 extracted speckle images for each direction. The analyzed phase difference maps of each direction are shown in Fig. 8.5a, b. The analyzed area is 19.9 mm × 15.9 mm (360 × 288 pixels) of the center of the captured image as shown the white dashed line in Fig. 8.4a. In this result, a spatial filtering, local least squares is used to remove some noises in the analyzed results. The region size of local least squares is 21 × 21 pixels. Since mathematical sign of the result is unknown in random phase-stepping method, the sign is decided by the given rotation angle. These phase difference maps slightly fluctuate. However, it shows comparatively good results in speckle interferometry measurements. The rotation angle calculated from the average phase gradients is 0.00877°. It is good agreement with the given angle of 0.00880°. Since the resolution of the rotation target is less than 0.00044° and the range of the experimental error is less than 5 %, the measurement result shows a high measurement accuracy. Therefore, it is validated that the bi-directional displacement measurement by speckle interferometry under environmental disturbance is possible.

In this technique, it is expected that the vibration condition, the exposure time, number of capturing images, F-number of the camera lens, the polarization change and the laser power affect the measurement result. Additionally, reductions of the noise and the fluctuation of phase values are important to improve the spatial resolution of the measurement. Therefore, quantitative investigations of the effects are required for the practical use of this technique.

Fig. 8.5 Phase difference maps (**a**) vertical component, (**b**) horizontal component

Vertical component   Horizontal component
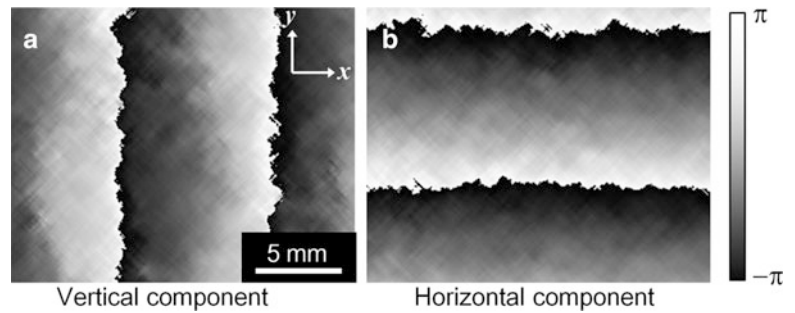
8.6 Conclusion

A bi-directional in-plane displacement measurement by speckle interferometry using an optimum image extraction under environmental disturbances as random vibrations is investigated. A compact speckle interferometer which has 4 laser beams for bi-directional sensitivity is constructed on a tripod. The polarizations of laser beams of each sensitive direction are arranged orthogonal and diffuse reflected laser beams from an aluminum target can be separated by a polarizing beam splitter in front of digital cameras. The polarization change of the reflected laser from the aluminum target is very small. An in-plane rotation of the target is measured by the proposed method in a presence of a random vibration. A phase analysis by random phase-stepping method using extracted images from a lot of images captured at the initial and the rotated state is performed. As a result, the analyzed phase difference maps show a good agreement with the given rotation angle. Therefore, it is validated that the bi-directional displacement measurement by speckle interferometry under environmental disturbance is possible. It is expected that deformation measurements by speckle interferometry is available for various environments.

References

1. Chiang FP (1989) Speckle metrology, ASM handbook, vol 17, Nondestructive evaluation and quality control. ASM International, Novelty, OH, pp 432–437
2. Sirohi RS (2002) Speckle interferometry. Contemp Phys 43(3):161–180
3. Vial-Edwards C, Lira I, Martinez A, Münzenmayer M (2001) Electronic speckle pattern interferometry analysis of tensile test of semihard copper sheets. Exp Mech 41(1):58–62
4. Hinsch KD, Gülker G, Kelmers H (2007) Checkup for aging artwork—optical tools to monitor mechanical behavior. Opt Lasers Eng 45:578–588
5. Arikawa S, Yoneyama S (2011) A simple method for detecting a plastic deformation region formed by local loading using electronic speckle pattern interferometry (in Japanese). Trans Jpn Soc Mech Eng Ser A 77(775):383–390
6. Madjarova V, Toyooka S, Widiastuti R, Kadono H (2002) Dynamic ESPI with subtraction-addition method for obtaining the phase. Opt Commun 212:35–43
7. Madjarova V, Kadono H, Toyooka S (2003) Dynamic electronic speckle pattern interferometry (DESPI) phase analyses with temporal Hilbert transform. Opt Express 11:617–623
8. Ikeda T, Ichinose K, Gomi K, Yoshida S (2007) Study of dynamic fracture toughness measuring method by electronic spackle pattern interferometry (in Japanese). In: Proceedings of the mechanical engineering congress, 2011 Japan of the Japan Society of Mechanical Engineers, No. 7-1, pp 57–58
9. Arikawa S, Gaffney JA, Gomi K, Ichinose K, Ikeda T, Mita T, Rourks RL, Schneider C, Yoshida S (2007) Application of electronic speckle pattern interferometry to high-speed phenomena. J Mater Test Res Assoc Jpn 52(3):176–184
10. Ma C, Huang C (1998) Vibration characteristics for piezoelectric cylinders using amplitude-fluctuation electronic speckle pattern interferometry. AIAA J 36(12):2262–2268
11. Chen C, Huang C, Chen Y (2009) Vibration analysis and measurement for piezoceramic rectangular plates in resonance. J Sound Vib 326:251–262
12. Silva Gomes JF, Monteiro JM, Vaz MAP (2000) NDI of interface in coating systems using digital interferometry. Mech Mater 32:837–843
13. Gryzagoridis J, Findeis D, Tait RB (2005) Residual stress determination and defect detection using electronic speckle pattern interferometry. Insight Non-Destr Test Cond Monit 47(2):91–94
14. Ambu R, Aymerich F, Ginesu F, Priolo O (2006) Assessment of NDT interferometric techniques for impact damage detection in composite laminates. Compos Sci Technol 66:199–205
15. Arikawa S, Nakaya Y, Yoneyama S (2012) Electronic speckle pattern interferometry with optimum image extraction for deformation measurement under environmental disturbance. J Solid Mech Mater Eng 6(6):634–644

16. Arikawa S, Ashizawa K, Yoneyama S (2012) Deformation measurement under environmental disturbance using electronic speckle pattern interferometry with optimum image extraction. In: Proceedings of the ISEM-ACEM-SEM-7th ISEM'12, USB, pp 1–6
17. Wang Z, Han B (2004) Advanced iterative algorithm for phase extraction of randomly phase-shifted interferograms. Opt Lett 29(14):1671–1673
18. Bingleman LW, Schajer GS (2011) DIC-based surface motion correction for ESPI measurements. Exp Mech 51(7):1207–1216
19. Arikawa S, Yoneyama S (2011) Correcting the effect of rigid body displacement in speckle interferometry (in Japanese). J Jpn Soc Exp Mech 11(3):195–200
20. Yoshida S, Suprapedi A, Widiastuti R, Triastuti ET, Kusnowo A (1995) Phase evaluation for electronic speckle pattern interferometry deformation analyses. Opt Lett 20:755–757

Chapter 9
Design of a Double-Illumination ESPI System for the Measurement of Very Slow Motions

C. Casavola, G. Pappalettera, and C. Pappalettere

Abstract Necessity appears in some contexts to measure velocity of objects moving at very low speed. In this work a double illumination speckle interferometer was used in order to determine the velocity of a plate moved by a PZT stage at velocity from 7 to 21 nm/s. When the time history for each pixel intensity is recorded oscillations in the signal appear due to phase variation connected to the movement of the object. By measuring the frequency of the oscillations of the autocorrelation function of the signal intensity and by knowing the geometrical and physical parameters of the set-up it is possible to calculate the velocity of the objects. Results show that the set up adopted in this work is able to measure very slow motion down to 7 nm/s with small percentage error and rather quickly.

Keywords ESPI • Slow movements • Optical methods • Speckle • Autocorrelation function

9.1 Introduction

There exist a number of events that occur at very low speed. For example the average speed of scatterers in painting during the drying process is of about 10 nm/s [1]; the hair growth rate can be estimated to be about 5 nm/s [2]; the average velocity of scattering particles in apples is about 400 nm/s [3] and so on. In this context a technique having the ability to measure very low velocities in a relatively easy way, with a great level of accuracy and in a reduced amount of time would constitute a valuable tool in many applications. Several approaches have been developed in order to measure such a kind of motions. For example the hair growth rate can be measured by means of the intake of specific marking molecules [2]. A quite recently developed approach to this kind of problem is based on the so called dynamic speckle. It is well known that whenever a coherent beam of light is reflected by an optically rough surface a random distribution of intensity is produced [4–14]. A motion of the surface causes fluctuation in the speckle pattern intensity [4], that is to say that speckles become dynamic and contain information about the velocity of the scattering surface. Analogues considerations can be done for the case when a living tissue is illuminated by coherent light [5] in this case dynamic speckle pattern is due to scattering by internal moving particles [3]. In the cited works the used measurement configuration is quite simple consisting essentially in an illuminating laser source with eventually some polarizing or collimating optics and a CCD camera system with imaging optics or alternatively a photo-detector with a spatial filtering element [4]. This work proposes the realization of a double illumination interferometer for the analysis of dynamic speckle. This configuration allows having the sensitivity vector of the system along a definite in-plane direction so that it is possible to quantitatively measure the relative component of the velocity of the object. Motion of the object, in fact, changes the optical path of the light to a given pixel of the camera; this phase variation cause oscillations in the recorded intensity whose frequency is directly connected with the velocity. In this paper results from analysis of dynamic speckle from a metal diffusive plate moved by a PZT stage are presented.

C. Casavola (✉) • G. Pappalettera • C. Pappalettere
Dip. di Meccanica, Matematica e Management, Politecnico di Bari, viale Japigia 182, 70126 Bari, Italy
e-mail: caterina.casavola@poliba.it

9.2   Materials and Methods

Apparatus for electronic speckle pattern interferometry based on the double illumination Lendeertz interferometer was built (Fig. 9.1) [4].A 17 mW He-Ne Laser ($\lambda = 632.8$ nm) was used as a source of coherent illumination. Laser light was spatially filtered and then collimated by a plano-convex lens. Half of the beam directed towards the sample which is placed at an angle $\theta_1 = 45°$ with respect to the direction of propagation of the beam. A second half of the beamis reflected by the mirror M_1 towards the surface of the sample so that the incidence angle is $\theta_2 = \theta_1 = \theta = 45°$. An 8 bit FireWire CCD camera was placed perpendicularly to the sample in order to record the speckle images. The camera was equipped with an objective lens having a field of view FOV = 21 mm. The specimen was clamped over a piezoelectric translation stage. The motion of the PZT stage, the settings of the camera and the acquisition of the images are controlled by dedicated software.

The overall speckle pattern recorded by the camera at the time t_0 is given by:

$$I(t_0) = I_{01} + I_{02} + 2\sqrt{I_{01}I_{02}} \, \cos \gamma \tag{9.1}$$

Being I_{01} and I_{02} the intensities of the two beams and γ is the relative phase difference. As a consequence of the motion of the surface of the sample the recorded intensity at $t_0 + \Delta t$ changes into:

$$I(t_0 + \Delta t) = I_{01} + I_{02} + 2\sqrt{I_{01}I_{02}} \, \cos (\gamma + \Delta\varphi(\Delta t)) \tag{9.2}$$

Being $\Delta\varphi$ an extra phase difference connected to the motion of the surface.

In the configuration of Fig. 9.1 $\Delta\varphi(\Delta t)$ is a function of the $u(\Delta t)$ component of the displacement of the sample along the x direction according the formula:

$$\Delta\varphi(\Delta t) = \frac{4\pi}{\lambda} u(\Delta t) \sin \vartheta \tag{9.3}$$

If the sample is moved at uniform velocity it can be inferred that $\Delta\varphi$ linearly increases with time so that the time history of intensity, that is to say the evolution in time, at each pixel has a sinusoidal-like behavior.

The displacement per cycle S_c can be deduced by (9.3) to be:

$$S_C = \frac{\lambda}{2 \sin \vartheta} \tag{9.4}$$

So that for the shown set-up is $S_c \approx 447$ nm.
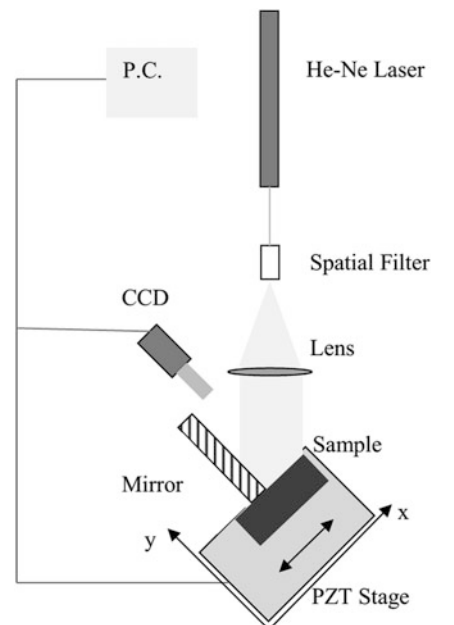
Fig. 9.1 Schematic of the set-up used for the measurement of velocity of motion of a sample by the dynamic speckle technique

The approach proposed in this paper aims to measure the motion velocity and is based on the analysis of the time history series for a given pixel [15–17] that is to say by studying the evolution of the intensity associated to a given pixel as a consequence of the motion of the surface. A higher velocity of the sample will result in a higher frequency of the time history signal. If the frequency f, expressed in cycles per frame, is measured then the velocity of the sample can be calculated as:

$$v = \frac{f S_f}{\delta t} \tag{9.5}$$

In order to measure the frequency f of the signal an algorithm in LabView$^®$ was implemented. This algorithm calculates power spectrum of the input signal and extract the frequency component having the highest amplitude.

The autocorrelation function was also analyzed. This is a common tool in dynamic speckle analysis [15, 18]. It can be defined as follows:

$$A(t) = \frac{1}{\langle I^2 \rangle} \int\limits_{-\infty}^{+\infty} I(\tau) I(\tau - t) d\tau \tag{9.6}$$

It can be easily shown that if the function I is periodic with period T the autocorrelation function is periodic with the same period [19].This leads to the possibility to extract the information on the frequency from the autocorrelation function too instead than from the signal itself as it will be shown in the next section.

9.3 Results

The time histories were calculated both for the case of steady object and for the case of surface moving at difference velocity. Three different velocities were tested in this experiment namely 7, 14 and 21 nm/s. Results corresponding to an arbitrarily chosen row are shown in Fig. 9.2.

A single row in the time histories represents the intensity evolution for a given pixel whose typical behavior is reported below for a velocity of the sample v = 14 nm/s

Starting from the intensity signal (Fig. 9.3) the autocorrelation function can be calculated as it is shown in Fig. 9.4.

Once the intensity for each pixel is analyzed and the relative frequency is calculated the velocity has been calculated by using (9.5). Due to the fact that the whole sample moves uniformly along the x direction each pixel contains the information about the velocity of the sample. Three lines were selected and for each of them the time history of ten pixels was analyzed.
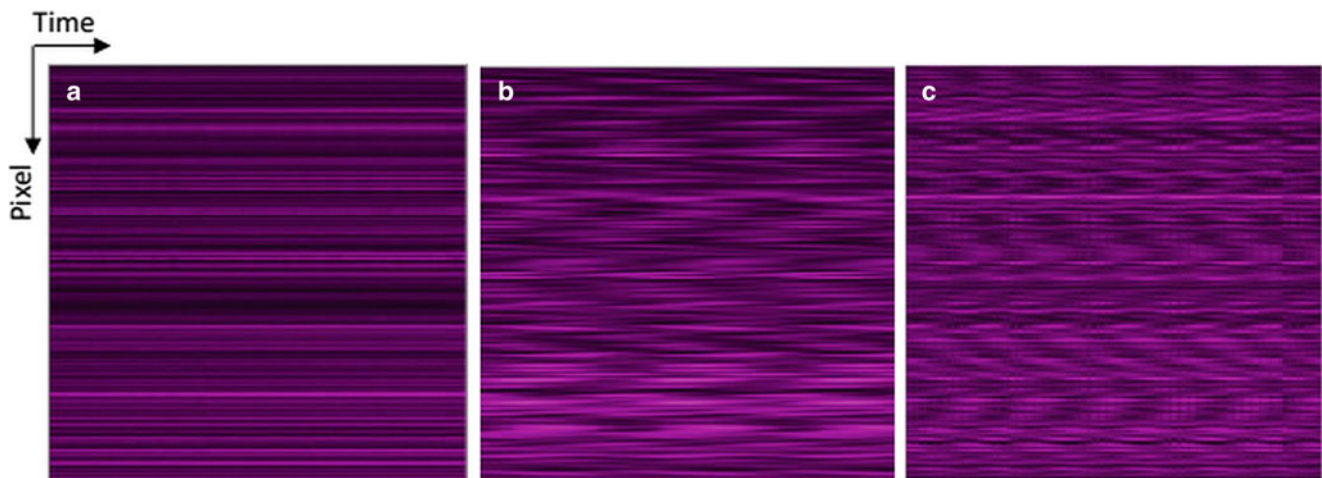
Fig. 9.2 Time history of a line of the speckle pattern for a still object (**a**), an object moving at v = 7 nm/s (**b**) and an object moving a 21 nm/s (**c**). The plots refer to 512 recorded images acquired at a frame rate of 3.75 fps
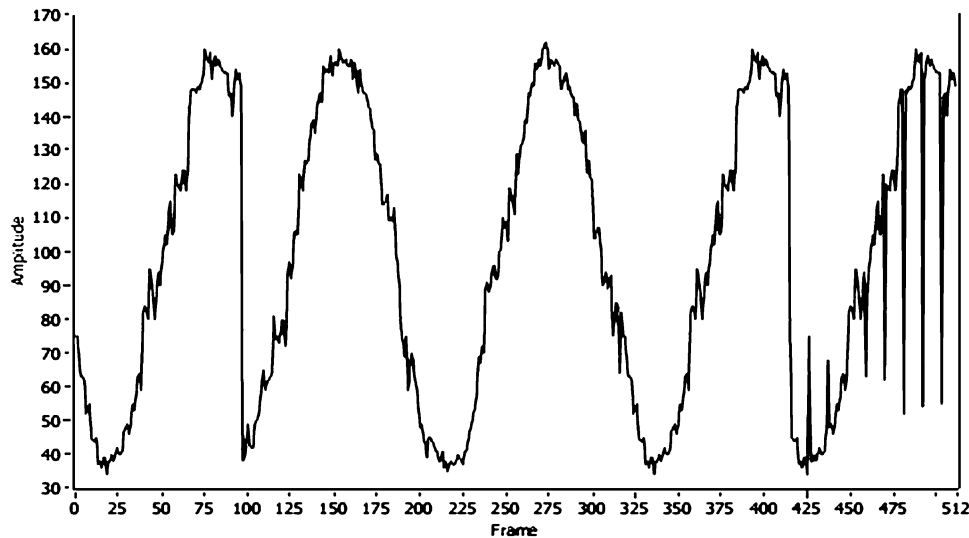
Fig. 9.3 Plot of the evolution of the intensity recorded at a given pixel as a consequence of the motion of the sample. The plot refers to a velocity v = 14 nm/s

Fig. 9.4 Plot of the autocorrelation function calculated for the intensity signal reported in Fig. 9.3. The plot refers to a velocity v = 14 nm/s
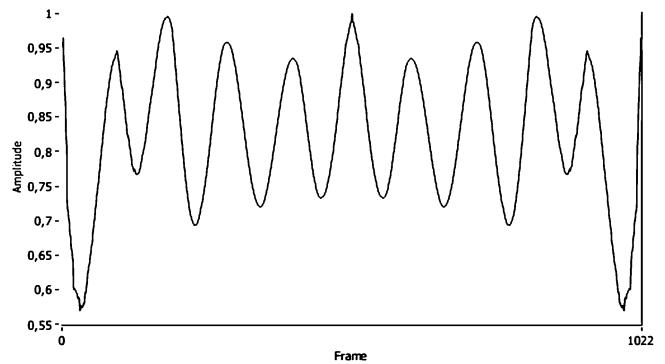
In order to calculate v and the results for v = 7 nm/s and 14 nm/s are reported in Figs. 9.5 and 9.6. The value of the velocities reported in the figures was calculated by extracting the frequency of the autocorrelation signal for a given pixel and then by applying (9.6).

9.4 Discussion

The time histories reported in Fig. 9.2 show qualitatively the appearance of a periodic structure connected with the movement of the object. For a steady object (Fig. 9.2a) the intensity for each pixel stays almost constant; for moving object (Fig. 9.2b, c) the intensity shows to oscillate with a frequency that gets higher when velocity gets higher. The evolution of the intensity with time is quantitatively reported in Fig. 9.3 where is also possible to observe that presence of noise contaminates the carrier signal containing the information about the velocity of the sample. The level of noise can be moderate, as it appears to be in the first 420 frames in Fig. 9.3. External disturbances introduced by motions and vibration can generate high level of noise as it appears, for example, in the last eighty frames in Fig. 9.3. Presence of noise can severely affect the correct evaluation of the signal frequency. Benefits in this sense are connected with the use the autocorrelation function because it effectively works as a filter. Figures 9.3 and 9.4 in fact show that the autocorrelation signal is strongly less affected by noise. If a band-pass filter is used, the problems related to presence of noise are reduced. However an optimized sizing of the filter would require an a-priori knowledge of the expected frequency value. Measurement results shown in Figs. 9.5 and 9.6 demonstrate that the system is quite accurate to distinguish difference in velocity equal to 7 nm/s. Velocity was evaluated by

Fig. 9.5 Plot of the measured velocity in correspondence of three different lines near the upper part, the middle part and the lower part of the sample. The measures reported refers to different pixel along the chosen line. The expected velocity was v = 7 nm/s. The indicated average value has been calculated over 30 measurements
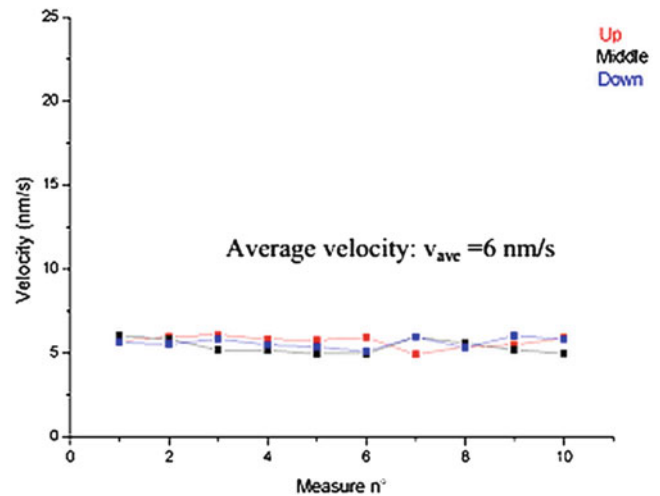
Fig. 9.6 Plot of the measured velocity in correspondence of three different lines near the upper part, the middle part and the lower part of the sample. The measures reported refers to different pixel along the chosen line. The expected velocity was v = 14 nm/s. The indicated average value has been calculated over 30 measurements
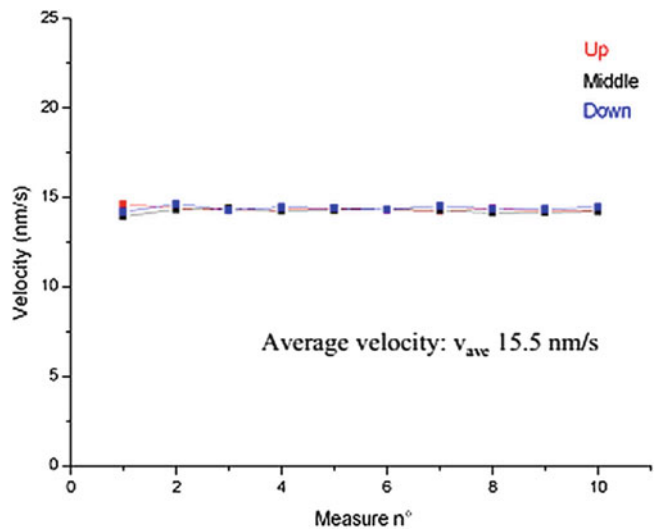
the autocorrelation function in three different locations in the upper, in the lower and in the middle part of the sample leading to the same result due the uniform motion of the sample. Difference between theoretical and measured value was 6 % for $v_{theoretical} = 7$ nm/s, 10 % for $v_{theoretical} = 14$ nm/s and 7 % for $v_{theoretical} = 21$ nm/s.

9.5   Conclusions

In the present paper a set-up and a methodology to measure the velocity of an object moving at few nm per second was shown. The system is based on the evaluation of the phase oscillation of an interference speckle pattern due to the rigid motion of the sample. The methods essentially rely on the measurement of the frequency of the phase oscillation; this is done indirectly by measuring the frequency of the autocorrelation function of the intensity signal for a given pixel. Autocorrelation works like a filter and allows reducing noise disturbance. The system is able to correctly measure velocity with an error of 10 % in a time of 2 min and 16 s. An important aspect that should be highlighted is the rather quick velocity of this set up in giving the results. Improvement of the system could be done in the direction of optimizing the noise filtering process and to reduce the dimension of the system, this point could be afforded by using fiber optics to generate speckle pattern.

References

1. Rabal HJ, Arizaga R, Cap NL, Grumel E, Trivi M (2003) Numerical model for dynamic speckle: an approach using the movement of the scatterers. J Opt A Pure Appl Opt 5(5):S381
2. Norio M, Toshihiko U (1992) Analysis of ofloxacin in hair as a measure of hair growth and as a time marker for hair analysis. Ther Drug Monit 14(6):525–528
3. Trivi M (2009) Dynamic speckle: origin and features. In: Rabal HJ, Braga RA (eds) Dynamic laser speckle and application. CRC, Broken Sound Parkway
4. Cloud G (1998) Optical methods of engineering analysis. Cambridge University Press, Cambridge
5. Barile C, Casavola C, Pappalettera G, Pappalettere C (2011) Residual stress measurement by electronic pattern interferometry: a study of the influence of the geometrical parameters. Integritetivekkonstrukcija 11(3):177–182
6. Barile C, Casavola C, Pappalettera G, Pappalettere C (2011) Mechanical characterization of SLM specimens with speckle interferometry and numerical optimization. Exp Appl Mech 6:837–843
7. Barile C, Casavola C, Pappalettera G, Pappalettere C (2013) Hybrid characterization of laminated wood with ESPI and optimization methods. In: Imaging methods for novel materials and challenging applications, vol 2. Conference proceedings of the society for experimental mechanics, Indianapolis
8. Casavola C, Barile C, Pappalettera G, Pappalettere C (2012) Residual stress measurement by electronic speckle pattern interferometry: a study of the influence of analysis parameters. Integritetivekkonstrukcija 12(3):159–163
9. Barile C, Casavola C, Pappalettera G, Pappalettere C (2014) Analysis of the effects of process parameters in residual stress measurements on titanium plates by HDM/ESPI. Measurement 48:220–227
10. Casavola C, Lamberti L, Moramarco V, Pappalettera G, Pappalettere C (2013) Experimental analysis of thermo-mechanical behavior of electronic components with speckle interferometry. Strain 49(6):497–506
11. Casavola C, Campanelli SL, Pappalettere C (2008) Experimental analysis of residual stresses in the selective laser melting process. Society for Experimental Mechanics—11th international congress and exhibition on experimental and applied mechanics, pp 1479–1486
12. Kamshilin AA, Miridonov SV, Sidorov IS, Semenov DV, Nippolainen E (2009) Statistics of dynamic speckles in applications to distance measurements. Opt Rev 16(2):160–166
13. Fomin N, Fuentes C, Saulnier JB, Tuhault JL (2001) Tissue blood flux monitoring by laser speckle photography. Laser Phys 11(4):525–529
14. Cloud GL (2001) Apparatus for electronic speckle pattern interferometry. USA Patent Apparatus for electronic speckle pattern interferometry, February 2001
15. Bergkvist A (1997) Biospeckle-based study of the line profile of light scattered in strawberries, Sweden
16. Arizaga R, Trivi M, Rabal H (1999) Speckle time evolution characterization by the co-occurence matrix analysis. Opt Laser Technol 31:163–169
17. Braga RA, Silva WS, Sàfadi T, Nobre CMB (2008) Time history speckle pattern under statistical view. Opt Commun 281:2443–2448
18. Pajuelo M, Baldwin G, Rabal H, Cap N, Arizaga R, Trivi M (2003) Bio-speckle assessment of bruising in fruits. Opt Lasers Eng 40:13–24
19. Hobbies BK, Roth BJ (2007) Intermediate physics for medicine and biology, IVth edn. Springer, New York

Chapter 10
Multiplexed Holography for Single-Shot Three-Dimensional Shape and Displacement Measurements

Morteza Khaleghi, Cosme Furlong, Jeremie Guignard, Ivo Dobrev, Jeffrey Tao Cheng, and John J. Rosowski

Abstract Holographic interferometric methods for measuring 3D displacement fields require at least three individual sensitivity vectors. Methods based on multiple directions of illumination have limited application when studying transient phenomena, including, but not limited to, measurements of biological tissues that have temporally-varying responses, such as the eardrum or Tympanic Membrane (TM). Therefore, to measure 3D displacements in such applications, all the measurements have to be done concomitantly. In this paper, we propose a new multiple illumination direction approach to measure 3D displacements from a single-shot hologram that contains displacement information from three sensitivity vectors. In our approach, the hologram of an object of interest is recorded with three simultaneous incoherently-superimposed pairs of reference and object beams, such that the modulation image corresponding to each illumination direction is reconstructed at a particular position on the image. Incoherent superposition of the beams is implemented by using three different laser diodes. Because of the differences in the position of each reference beam and wavelength of each pair of beams, the reconstruction distance and magnification of each sensitivity vector are different. We, therefore, developed and implemented a registration algorithm to accurately translate individual views into a single global coordinate system. Representative results will include measurements of shape and sound-induced 3D displacements of the TM.

Keywords 3D Displacement measurements • Digital holography • Image registration • Multiplexed hologram • Tympanic membrane

10.1 Introduction

The eardrum or Tympanic Membrane (TM) transfers acoustic energy from the ear canal (at the external ear) into mechanical motions of the ossicles (at the middle ear). The acousto-mechanical-transformer behavior of the TM is determined by its shape and mechanical properties. For a better understanding of hearing mechanisms, full-field-of-view techniques are required to quantify shape, and nanometer-scale sound-induced displacement of the TM in 3D [1–3]. Typical 3D displacement measurement techniques require repeating the measurement from at least three sensitivity directions, which might not be applicable in some cases with temporally-varying nature (e.g., biomaterials). Due to time-varying nature of the TM,

M. Khaleghi (✉) • I. Dobrev
Mechanical Engineering Department, Center for Holographic Studies and Laser micro-mechaTronics (CHSLT),
Worcester Polytechnic Institute, Worcester, MA 01609, USA
e-mail: mkm@wpi.edu

C. Furlong
Mechanical Engineering Department, Center for Holographic Studies and Laser micro-mechaTronics (CHSLT),
Worcester Polytechnic Institute, Worcester, MA 01609, USA

Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02114, USA

J. Guignard • J.T. Cheng • J.J. Rosowski
Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02114, USA

a novel, unique method based on multiplexed holography is developed to minimize the measurements time. In this method, the hologram of an object of interest is recorded with three simultaneous incoherently-superimposed pairs of reference and object beams, such that the modulation image corresponding to each illumination direction is reconstructed at a particular position on the image. In addition, due to the differences in the position of each reference beam and wavelength of each pair of beams, the reconstruction distance and magnification of each sensitivity vector are different. We, therefore, developed and implemented a registration algorithm to accurately translate individual views into a single global coordinate system. The registration method uses Phase-Only Correlation (POC) [4] and a swept cut-off filtering to improve robustness. In this paper, three images obtained from three sensitivity vectors in a multiplexed hologram are registered and, consequently, 3D data are extracted.

10.2 Methods

10.2.1 Single-Frame Multiplexed Holography

In 3D displacement measurements using multiplexed holography, three incoherently-superimposed pairs of object and reference beams are recorded at the CCD detector, simultaneously. The object of interest is illuminated concurrently with three object beams from different directions to provide three sensitivity vectors that are required for 3D displacement measurements [5]. The reference beams provide spatial carrier frequency on the CCD to enable the capabilities of single-frame phase extraction [6]. The total intensity recorded on the CCD detector, $I(x, y)$, due to incoherent superposition of three pairs of reference-object beams [7] can be shown by

$$I(x,y) = \sum_{n=1}^{3} I_n(x,y) = \sum_{n=1}^{3} |R_n(x,y) + O_n(x,y)|^2, \tag{10.1}$$

where x and y are coordinates at the CCD plane, R_n and O_n are reference and object waves, respectively, with n = 1,2,3. Equation (10.1) can be expanded to Eq. (10.2) using basics of interferometry

$$I(x,y) = \sum_{n=1}^{3} a_n(x,y) + c_n(x,y)\exp\left[2\pi i\left(f_{n\xi}x + f_{n\eta}y\right)\right] + c_n^*(x,y)\exp\left[-2\pi i\left(f_{n\xi}x + f_{n\eta}y\right)\right], \tag{10.2}$$

where $a_n(x, y) = R_n^2(x, y) + O_n^2(x, y)$, and $c_n(x, y) = O_n(x, y)R_n(x, y)exp[i\phi_n(x, y)]$.

The two-dimensional Fourier Transform of Eq. (10.2) can be written as

$$\mathrm{FT}\{I\} = \mathrm{H}(\xi,\eta) = \sum_{n=1}^{3} A_n(\xi,\eta) + C_n\left(\xi - f_{n\xi0}, \eta - f_{n\eta0}\right) + C_n^*\left(\xi + f_{n\xi0}, \eta + f_{n\eta0}\right), \tag{10.3}$$

where uppercase letters denote the Fourier-transformed components. The terms C_n and C_n^* describe the complex distribution of the twin images components that mathematically are described as complex conjugates of one another, each containing the required phase information $\phi_n(x, y)$. As shown in Fig. 10.1, these terms are shifted in the Fourier domain and can be separated from one other by appropriate calculation and setting of carrier frequencies $f_{n\xi0}$ and $f_{n\eta0}$. These spatial carrier frequencies can be adjusted by varying the position of each reference beam with respect to the CCD.

Additionally, the DC component can be mathematically removed from the multiplexed hologram by recognizing that the DC component in the Fourier domain is due to the low-frequency spatial variation in the hologram, and it can thus be isolated by applying a high-pass filter. This is computationally efficiently achieved by subtracting the local average of the interferogram obtained by convolution of the interferogram with a low-pass kernel. The optical phase of each component of the multiplexed hologram is obtained by

$$\emptyset_n(x,y) = \frac{\mathrm{Im}[c_n(x,y)]}{\mathrm{Re}[c_n(x,y)]}, \quad n = 1, 2, 3. \tag{10.4}$$

Fig. 10.1 Numerically-reconstructed multiplexed holograms showing the objects and twin images (**a**) schematic spectrum of hologram consisting of three incoherently-superimposed interference patterns; (**b**) representative example of multiplexed hologram with two pairs of interference which can be used for in-plane measurements; and (**c**) representative example of multiplexed hologram with three pairs of interference, which can be used for 3D displacement measurements. In (**b**) and (**c**) the object of interest is a latex membrane which is acoustically excited by tone of 1.5 and 2.6 kHz
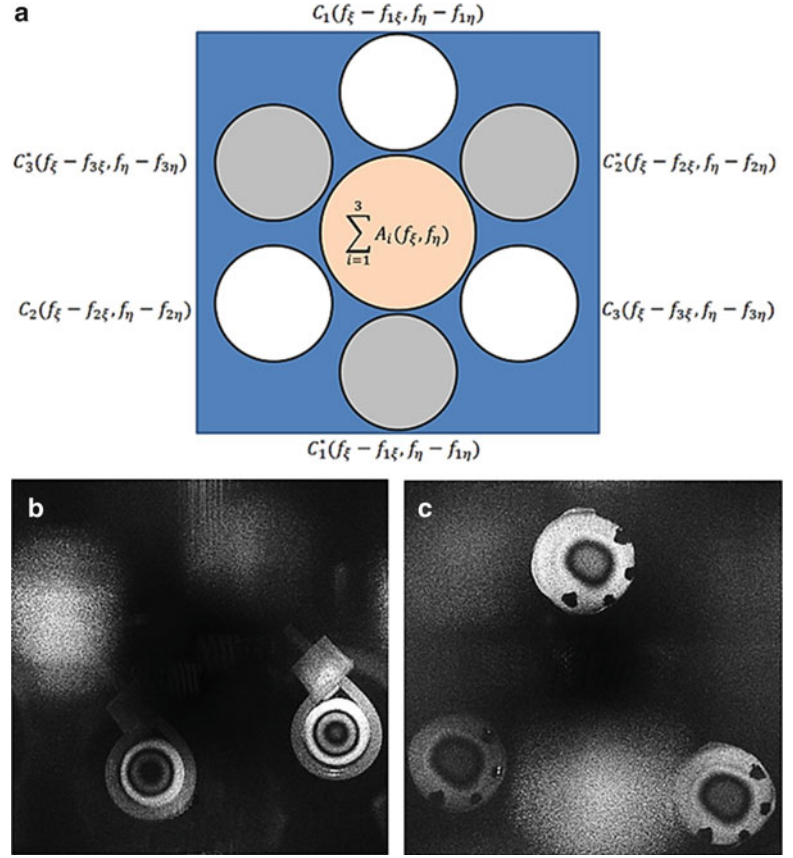
10.2.2   Experimental System for 3D Displacement Measurements with Multiplexed Holography

In our experimental system, incoherent superposition of the beams is implemented by using three different laser diodes. Three near Infra-Red (IR) external cavity laser sources with central wavelengths of 779.8, 780.2, and 780.6 nm are used. As shown in Fig. 10.2, the laser lights are coupled to the fibers and then get split into reference and object beams. The reference beams are directly illuminated to the CCD by a beamsplitter (BS) and the object beams illuminate the sample concurrently, from three different directions in order to provide three sensitivity vectors for 3D displacement measurements. In each Laser Delivery sections (LD1 to LD3), an Acousto-Optic Modulator (AOM) is used to switch the corresponding laser light on/off with a frequency of 40 MHz to enable the system with having stroboscopic measurements capability. For clarity, the optoelectronic components are shown for only one of the LDs.

Because of the differences in the position of each reference beam and wavelength of each pair of beams, the position, reconstruction distance and magnification of each image corresponding to each sensitivity vector are different. Therefore, for quantification of the shape and 3D displacement, image registration algorithms are required to register each image into one global coordinate system.

10.2.3   Registration of Multiplexed Holograms

In multiplexed holography with slightly different wavelengths of light sources, if the distances from the reference beams with respect to the camera are nearly similar for all three beams, a pure translation can be assumed at the position of the object in the reconstruction plane (as shown in Fig. 10.1b, c). Therefore, estimating translational displacements of images with high accuracy is important. Image registration using Phase-Only Correlation (POC) algorithms [4] can be used to determine the relative translation between two images. In this section, a high-accuracy translation estimation technique using the POC function is used to find the translation of one image with respect to the other. If $f_c(x_1, x_2)$ is considered as the
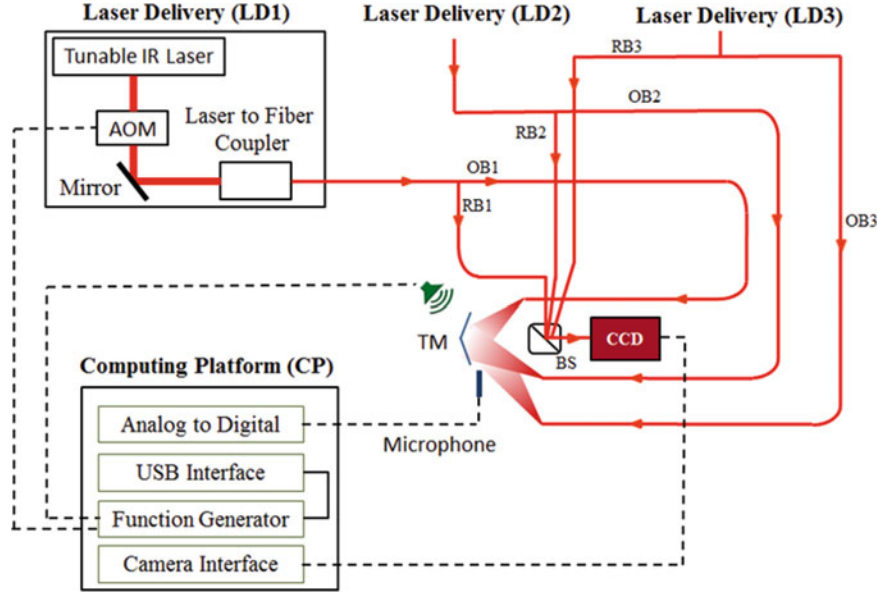
Fig. 10.2 Developed experimental system for simultaneous 3D displacement measurements with multiplexed holography

continuous-domain, original image, the displaced images has a δ_1 and δ_2 subpixel displacement of $f_c(x_1, x_2)$ along x_1 and x_2 directions, respectively. So, the displaced image can be represented as $f_c(x_1 - \delta_1, x_2 - \delta_2)$. Let $f(n_1, n_2)$ and $g(n_1, n_2)$ be the spatially sampled images of $f_c(x_1, x_2)$ and $f_c(x_1 - \delta_1, x_2 - \delta_2)$, defined as

$$f(n_1, n_2) = f_c(x_1, x_2)| \quad x_1 = n_1 T_1, x_2 = n_2 T_2, \tag{10.5}$$

$$g(n_1, n_2) = f_c(x_1 - \delta_1, x_2 - \delta_2)| \quad x_1 = n_1 T_1, x_2 = n_2 T_2, \tag{10.6}$$

where T_1 and T_2 are the spatial sampling intervals. For simplicity, $T_1 = T_2 = 1$. The POC function, $r(n_1, n_2)$, between $f(n_1, n_2)$ and $g(n_1, n_2)$ can be obtained by

$$r(n_1, n_2) = \frac{\alpha}{N_1 N_2} \frac{\sin\{\pi(n_1 + \delta_1)\}}{\sin\{\frac{\pi}{N_1}(n_1 + \delta_1)\}} \frac{\sin\{\pi(n_2 + \delta_2)\}}{\sin\{\frac{\pi}{N_2}(n_2 + \delta_2)\}}, \tag{10.7}$$

where $\alpha = 1$. Equation (10.7) represents the shape of the POC function. The peak position of the POC function corresponds to the translation between the two images [4]. The directions of illumination being different for the three beams, slight differences in the shape of the object's image have to be accounted for. In addition, the high resolution of the camera results in high-frequency noise. The POC algorithm was repeated with a swept cut-off low pass filtering in the frequency domain and the most frequent peak of the POC function was considered.

10.3 Results

In our experiments, the wavelengths of the laser sources are slightly different. As shown in Fig. 10.3, each component in the multiplexed hologram is reconstructed with the corresponding wavelength. Images are registered on one global coordinate system and the shape of the object is obtained by two-wavelengths digital holographic contouring [8]. In the future, similar registration algorithms will be used to register multiplexed holograms to characterize 3D sound-induced displacements of the TM from a single-shot.

Fig. 10.3 Shape measurements of a membrane with multiplexed holography (**a**) numerically-reconstructed hologram imported to the registration algorithms; (**b**) phase map; and (**c**) corresponding shape
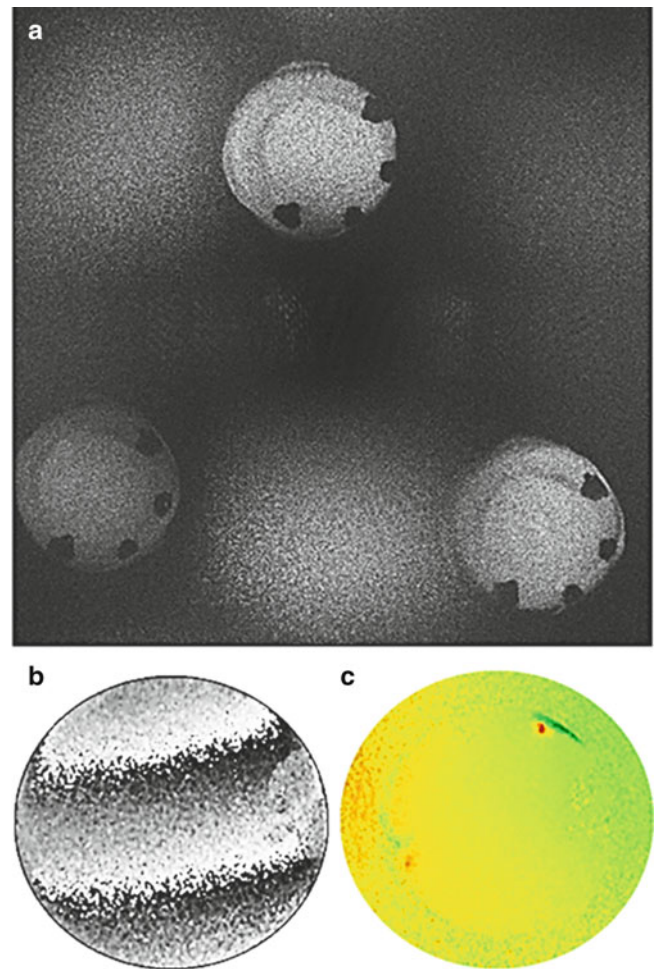
10.4 Conclusions

Due to time-varying nature of the TM, a novel, unique method for 3D displacement measurements based on multiplexed holography is developed that minimizes the measurements time. In our approach, the hologram of an object of interest is recorded with three simultaneous incoherently-superimposed pairs of reference and object beams, such that the modulation image corresponding to each illumination direction is reconstructed at a particular position of the image. An image registration algorithm based on the shift theorem of the Fourier Transform is implemented to register the images with sub-pixel accuracy. The same registration protocol will be used to register images of the multiplexed holograms of the acoustically-induced vibrations of TM in order to obtain shape and 3D displacement components of it from a single-shot hologram.

References

1. Khaleghi M, Dobrev I, Harrington E, Furlong C, Rosowski JJ (2014) Study of the transient response of Tympanic Membranes under acoustic excitation. Proceedings of SEM 2013, In Mechanics of biological systems and materials, vol 4, Springer, pp 1–9
2. Rosowski JJ, Dobrev I, Khaleghi M, Lu W, Cheng JT, Harrington E, Furlong C (2012) Measurements of three-dimensional shape and sound-induced motion of the chinchilla tympanic membrane. Hear Res 301:44–52

3. Khaleghi M, Lu W, Dobrev I, Cheng JT, Furlong C, Rosowski JJ (2013) Digital holographic measurements of shape and three-dimensional sound-induced displacements of Tympanic Membrane. Opt Eng 52(10):101916
4. Takita K, Sasaki Y, Higuchi T, Kobayashi K (2003) High-accuracy subpixel image registration based on phase-only correlation. IEICE Trans Fundam Electron Commun Comput Sci 86(8):1925–1934
5. Kreis T (2005) Handbook of holographic interferometry: optical and digital methods. Wiley-VCH, Weinheim, pp 243–255
6. Takeda M, Ina H, Kobayashi S (1982) Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. JOSA 72(1):156–160
7. Schedin S, Pedrini G, Tiziani HJ, Santoyo FM (1999) Simultaneous three-dimensional dynamic deformation measurements with pulsed digital holography. Appl Opt 38(34):7056–7062
8. Furlong C, Pryputniewicz RJ (2000) Absolute shape measurements using high-resolution optoelectronic holography methods. Opt Eng 39(1):216–223

Chapter 11
Observation of Grain-Size Effect in Serration of Aluminum Alloy

Tomohiro Sasaki, Tatsuya Nakamura, and Sanichiro Yoshida

Abstract Serration phenomena of AA5083 alloys with various crystal grain sizes during tensile tests were visualized using Electronic Speckle Pattern Interferometry. The behavior of deformation band during the plastic deformation associated with the serration was investigated by a dynamic observation of the images of displacement contours. The characteristic in the serrated curve varied depending on the grain size of specimen and the applied total strain at which the deformation band appeared. The occurrence and propagation processes of the deformation band observed in the displacement contours were investigated being compared with the serrated curve in the load. The correlation between the band propagation and the serrated curve, and the difference in the behavior of band propagation by the grain size were discussed based on the variations in the magnitude of the load drop and the velocity of the moving deformation band.

Keywords Speckle interferometry • Aluminum alloy • Deformation measurement • Dislocation • Serration

11.1 Introduction

The stress-strain curve of many aluminum alloys during the tensile test exhibits sharp stress drops caused by the discontinuous plastic flow. The magnitude and the temporal variation of this phenomenon, termed "serration", depends on various factors including tensile conditions, metallographical factors such as solute concentration and heat treatment condition. The origin of serration has been experimentally studied by many researchers [1–14]. It is generally accepted that the serrated flow is attributed to dynamic strain aging due to interaction between mobile dislocations and solute atoms in the alloy [1]. At the same time, the serrated flow is also associated with macroscopic phenomena such as the appearance of a localized band-like strain (deformation band, DB) and the propagation of DB. However, the correlation between the DB and the microscopic event mentioned above is still not completely understood at present. In our preliminary work [15], we investigated the spatiotemporal aspects of the DBs during the serration by a dynamic observation using Electronic Speckle Pattern Interferometry (ESPI), and the serrated flow has been found to be closely related to the location where the DB appears and the behavior of band propagation. Assuming that the occurrence and the propagation of DB are based on the slip of crystal grains due to the dislocations, these behaviors of DB may be strongly affected by the grain boundaries which prevent the dislocation movement, or crystal orientations. So far, however, comparatively few detail researches have been reported on the grain size effect in the serration focusing on the macroscopic propagation of DB [7, 16].

In the present study, Al-Mg binary alloys with different grain sizes are prepared, and the deformation behavior associated with the serrated flow during tensile tests has been visualized using ESPI. The effect of grain size on the occurrence and the propagation processes of DB have been investigated comparing with the load-displacement curve where the deformation band occurs.

T. Sasaki (✉) • T. Nakamura
Department of Engineering, Niigata University, Ikarashininocho 8050, Nishi-ku, Niigata-shi, Niigata 950-2181, Japan
e-mail: tomodx@eng.niigata-u.ac.jp

S. Yoshida
Department of Chemistry and Physics, Southeastern Louisiana University, SLU 10878, Hammond, LA 70402, USA
e-mail: syoshida@selu.edu

11.2 Experimental Procedure

11.2.1 Specimen

The material used in this study was Al-4.44mass%Mg binary alloy (industrial Aluminum alloy, AA5083) of 2 mm in thickness. Tensile specimen with 10 mm in gauge width and 25 mm gauge length was cut from the aluminum alloy sheet in rolled direction by electronic discharge machining. Crystal grain size of the specimen was controlled by strain-annealing method as follows. The specimen was stretched to a specified strain (pre-strain) using a tensile machine, followed by annealing at 530 °C in air. The pre-strain and annealing conditions are shown in Table 11.1. Figure 11.1 shows the macro and microstructure of specimens obtained. The mean grain sizes were 50 μm for the small grain (Fig. 11.1a), 2.5 mm for the medium grain (Fig. 11.1b), and 8 mm (Fig. 11.1c), respectively. Since the thickness of specimen used in this study was 3 mm, the large grain size can be regarded as almost single crystal.

11.2.2 Tensile Condition and Optical Set-Up for the ESPI

The specimen was clamped at both ends and stretched under displacement controlled tensile test at room temperature. The cross-head speed of the tensile machine was constant at 1.7×10^{-2} mm/s. Deformation process during the tensile test was observed using ESPI. Optical setup used in this study was the same as a previous paper shown in Fig. 11.2 [15]. Two optical configurations were arranged to vertical direction (*y-axis*) and horizontal (*x-axis*) direction to the tensile machine. Each optical configuration consists of a dual beam ESPI setup described as follows. The laser beam (semiconductor lasers with the wave length of 660 nm) was expanded by an optical lens and split into two paths by a cube beam splitter. The two beams illuminate the surface of specimen in the same incidence via two reflective mirrors. The angle of incidence used in this study was 36°. A superposed speckle pattern resulting from the two beams appeared on the surface. When the displacement occurs on the surface in the parallel direction to the optical configuration, optical intensity received by each pixel of the CCD changes due to the difference in optical path length. The horizontal configuration which the optical path is shown by solid line in Fig. 11.1 was sensitive to in-plane displacement component along only *x-axis*. Similarly, the vertical configuration shown by dashed line was sensitive to displacement along *y-axis*. The speckle pattern during the tensile test was captured by the CCD camera at a constant frame rate of 15 fps. The light sources on the two optical configurations were alternatively turned on, being synchronized with the CCD camera at its frame capture. Thus, alternate images formed by the vertical and

Table 11.1 Condition of strain-annealing method

Grain size	Pre-strain	Anneal condition
Small	0	530 °C for 24 h
Medium	0.05	530 °C for 4 h
Large	0.03	530 °C for 96 h
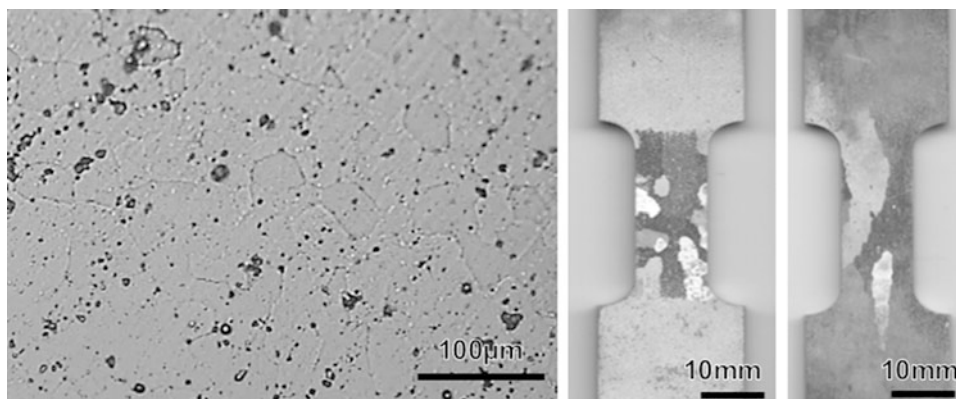
Fig. 11.1 Crystal grain size of specimen used in this study (**a**) small grain, (**b**) medium grain, and (**c**) large grain

horizontal configurations were stored into a computer memory. Each difference in the intensity of speckles in a certain time interval was calculated by subtracting each frame of image from a frame captured at a later frame. Consequently, the displacement distributions to *x-axis* and *y-axis* direction were individually obtained as contours.

11.3   Result and Discussion

11.3.1   Serration Observed in the Specimen with Small Grain

Serration has been qualitatively classified into some types on the basis of feature observed in the S-S curve by many researchers [16–23]. According to the classification by these researchers, Type A typically appears at large grain sizes, low solute concentrations in the specimen (Mg in the Al-Mg alloy), and high tensile rates. This serration is characterized by the small amplitude of load pulsation in relatively long period. In contrast, Type B appears at small grain sizes, high solute concentrations, and medium tensile rates, and is characterized by the larger amplitude of load pulsation compared to Type A. Figure 11.3a shows a load—displacement curve for the small grain specimen. The serration with the load pulsation was
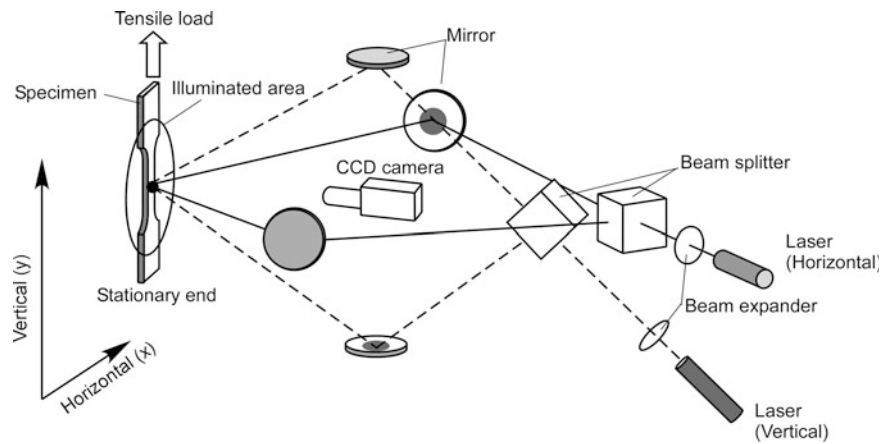
Fig. 11.2   Experimental arrangement [15]
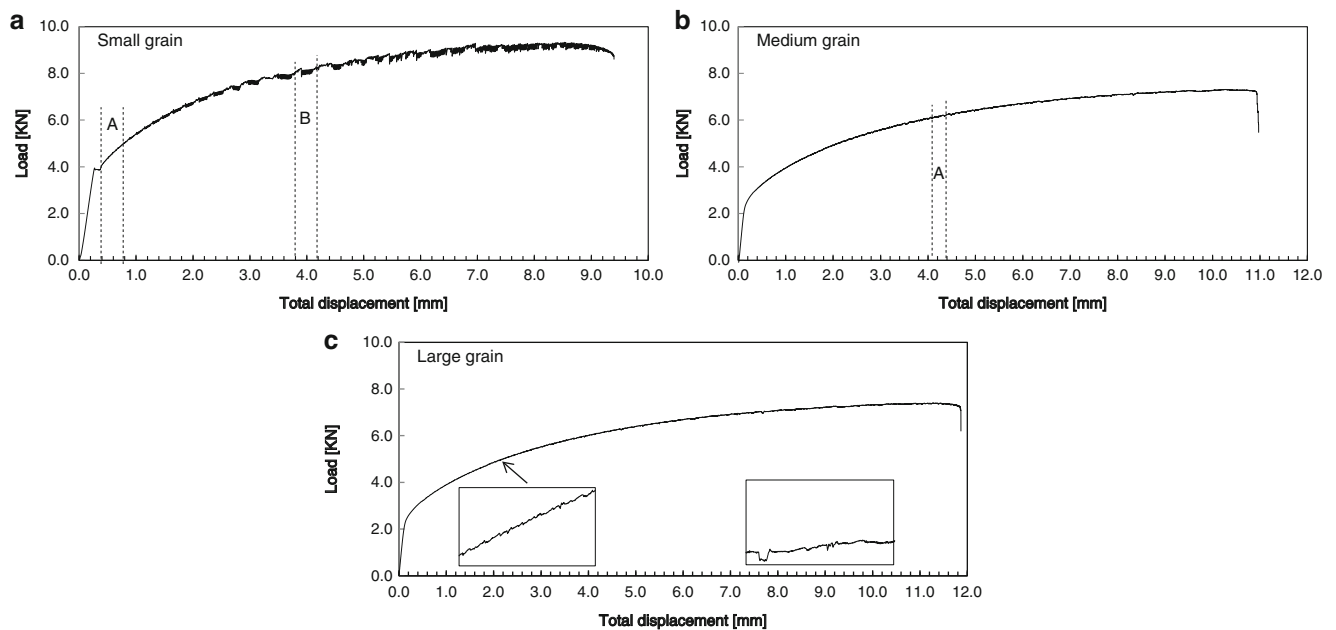
Fig. 11.3   Load—displacement curves of the specimens with (**a**) small grain, (**b**) medium grain, and (**c**) large grain
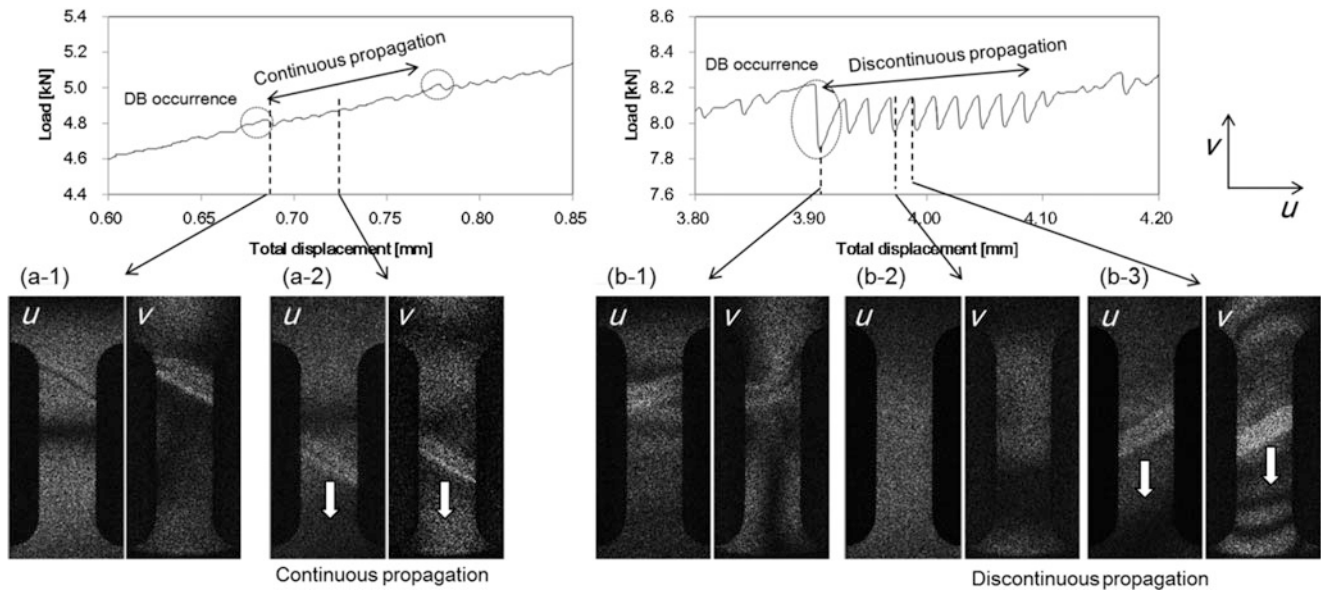
Fig. 11.4 Fringe pattern showing activity of DB observed in specimen with small grain. (**a**) and (**b**) are the serrated flows observed in the small total displacement, and the large total displacement, respectively. The displacement range of upper graph are indicated in Fig. 11.3a by "A" and "B"

observed immediately after the yielding. The pulsation amplitude tended to increase with an increase of the total displacement. From the above classification, the serration type observed in the earlier stage (small total displacement) agrees with type A, and it gradually changes to Type B as the total displacement increases. Furthermore, in the previous work [15], we pointed out that both types of serrations arise from an identical propagation mechanism. Typical changes of fringe pattern in the serrations corresponding to the conventional Type A and B are given in Fig. 11.4. It is found that the serrated curve consists of relatively large load drops in a long period and subsequent load pulsation between the large load drops. From the viewpoint of the behaviors of DB occurrence and propagation, the serrated flow can be explained as follows. A DB appears at the large load drop with the rapid fringe concentration (Fig. 11.4a-1, or b-1). The occurrence of DB is usually one shoulder of specimen because of stress concentration to the neck, and the DB subsequently starts to move along the gauge length toward the other shoulder with the load pulsation. The behavior of DB propagation in the small grain depended on the total displacement where the DB appears. When the load pulsation cannot be observed as shown in Fig. 11.4a, the DB shows a continuous propagation. As the pulsation amplitude increases with the total displacement, the propagation becomes discontinuous. The DB disappears at each load drop in the load pulsation (Fig. 11.4b-2), and a new DB appears beside the previously disappeared DB (Fig. 11.4b-2). In this manner, the DB shows an apparent movement. This phenomenon has been termed "discontinuous propagation" in the literatures [16–23]. Next propagation begins again with a large load drop. Therefore, it can be stated that the serrated curve is related to the two phenomena of DB occurrence and propagation. The variation in amplitude of load pulsation is understood as reflecting the difference between the strain rate in DB associated with dislocation movement and the total strain rate applying to the specimen.

11.3.2 Serration Observed in the Specimens with Medium Grain and Large Grain

Figures 11.3b, c show the Load-displacement curves for the medium grain and the large grain specimens. The load amplitude had a trend to decrease serration as the grain size becomes large. The ESPI fringe patterns and the corresponding load for the medium grain are given in Fig. 11.5. The DB appeared with a load drop similarly to that in the small grain (Fig. 11.5a-1). However, it was often observed that the DB did not show the propagation behavior as described above, and the other DBs occurred at random locations. Thus, the behavior of DB in the medium and large grain specimens can be interpreted as two processes of the random occurrence, and the propagation behavior similarly to that in small grain size (Fig. 11.5a-2–a-4). It is considered that the occurrence of DB is initiated by the local deformation caused by slipping at a grain with a favored orientation, and the grain slip propagates in the direction of gauge width of specimen, forming the
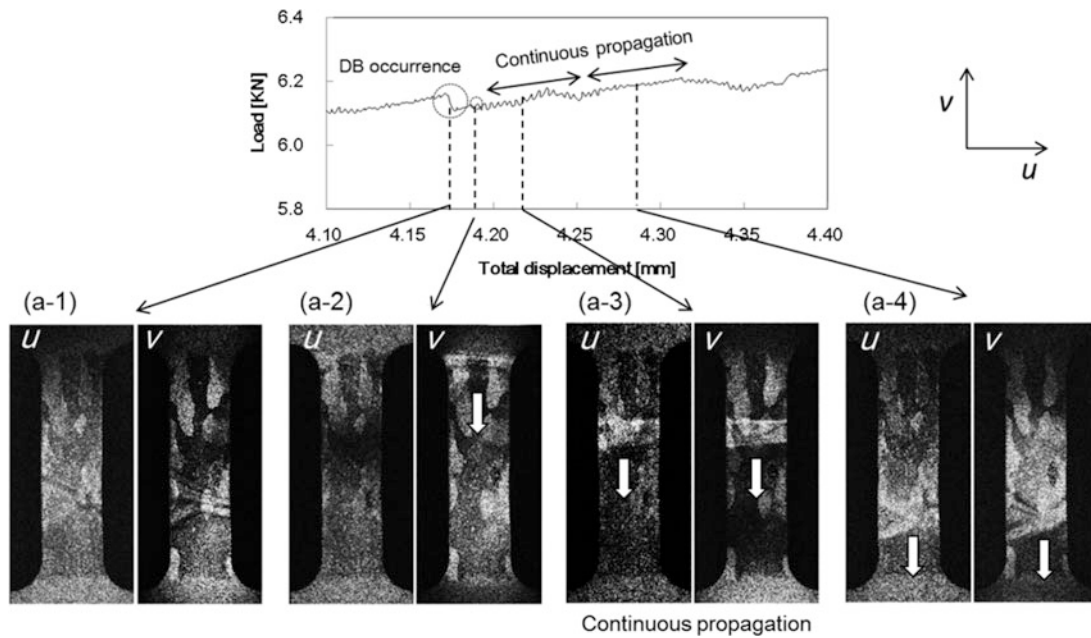
Fig. 11.5 Fringe pattern showing activity of DB observed in specimen with small grain. The displacement range of *upper graph* is indicated in Fig. 11.3b by "A"
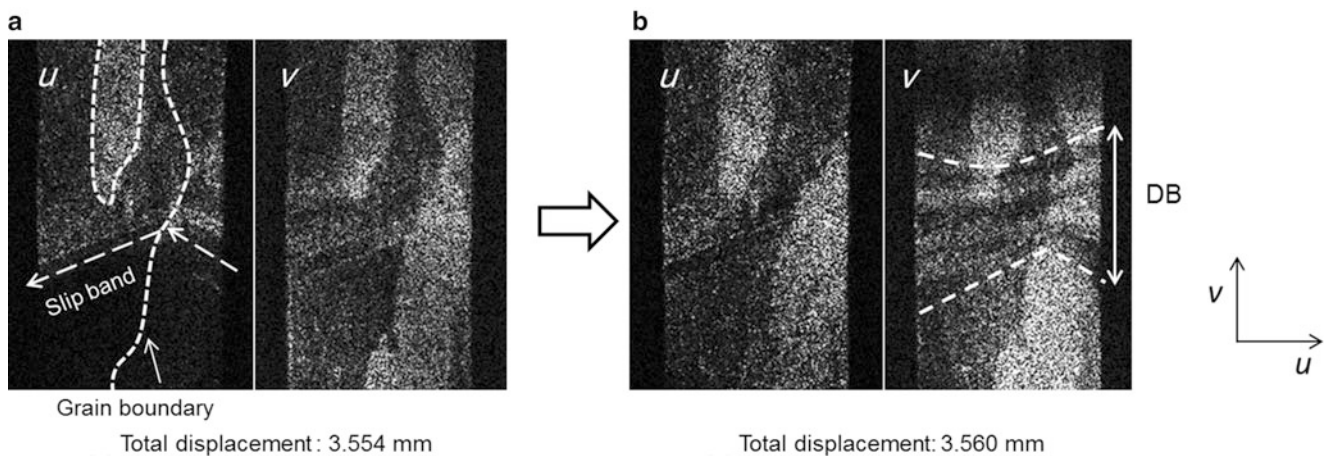
Fig. 11.6 Occurrence of DB observed in the specimen with large grain size

macroscopic DB. Figure 11.6 shows a typical development of DB in the large grain. Slip bands propagating across the grain boundary are observed in Fig. 11.6a. In addition the small interval fringe against the tensile direction as shown in Fig. 11.6b indicates that the stress concentration occurred at the grain boundary with the slip propagation. In poly-crystalline metals, the occurrence of DB is thought to be attained by such strain concentration due to dislocation pile-ups to the grain boundaries. Thus, the amplitudes of load drop and pulsation in the serrated curve should be correlated to the formation process of DB. Figure 11.7 shows relationship between the amplitude of pulsation and the total displacement for each grain size. The pulsation amplitudes of the medium and large grain are relatively small compared with that in the small grain, and have no dependence on the total displacement. In addition, Fig. 11.8 shows the relationship between the pulsation amplitude and reloading time. The amplitude shows approximately linear relationship against the reloading time, and the reloading time of the larger grain is shorter than in the smaller grain. These facts imply that DB in the larger grain occurs easily because the grain boundaries do not prevent the dislocation movement. Hence, it is concluded that the small pulsation amplitude in the large grain mainly responsible for the difference in formation process of DB.

Fig. 11.7 Variation of pulsation amplitude with total displacement
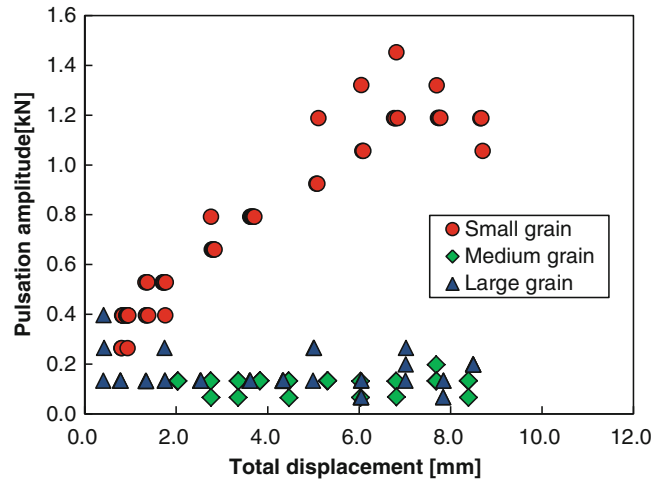
Fig. 11.8 Pulsation amplitude vs. reloading time
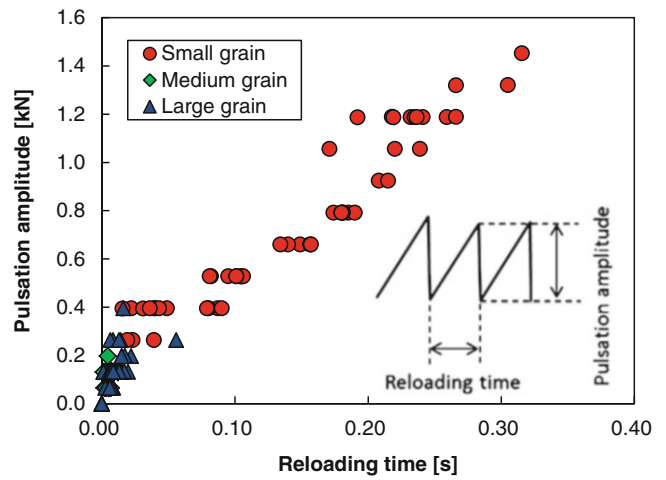
Fig. 11.9 Variation of propagation velocity of DB with total displacement
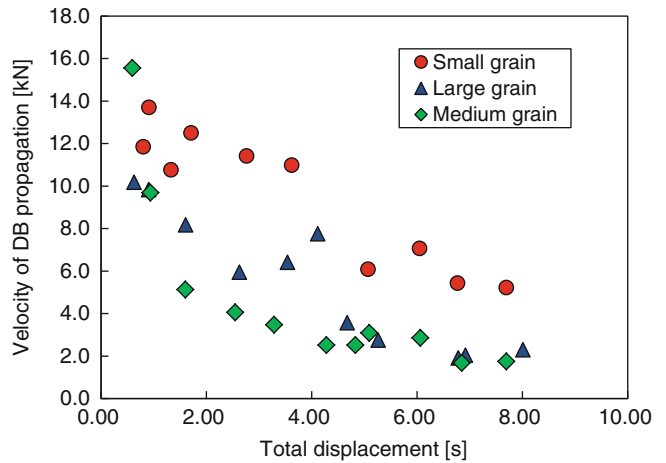
On the other hand, the propagation behavior of DB along the gauge length of specimen is also thought to be related to the pile-up mechanism of dislocations. In the small grain, the occurrence of DB should result in stress concentration around the DB, hence inducing an occurrence of new DB beside the DB. Figure 11.9 shows the velocity of DB showing the propagation behavior along the gauge length. The velocity of DB is smaller in the specimen with the larger grain, and decreases as the

total displacement increase. The high velocity of DB propagation in the small grain can be explained by its occurrence of DB accompanied by higher stress concentration. In contrast, the lower velocity and the random occurrence of DB in the larger grains may arise from the fact that the difference in grain orientation predominantly affects the occurrence of DB, rather than the stress concentration at the grain boundary.

11.4   Summary

Serration phenomena during tensile tests of Al-Mg alloy with various crystal grain sizes have been observed using ESPI. Spatiotemporal characteristics of the DB have been discussed by the ESPI fringe pattern being compared with the serrated curve, and the grain size effect on the serrated curve has been discussed. The serrated curve is related to at least two phenomena: the occurrence, and propagation of the DB. The amplitude of serrated curve decreased as the grain size became large. It was found that the occurrence of DB is easier in the specimen with larger grain size because of the low amount of grain boundary which prevents the slip propagation in the stage of DB initiation. On the other hand, in the larger grain size, it was observed that the velocity of DB propagation was smaller, and the DB occurs at random locations. From the viewpoint of grain slip due to dislocation movement, these results have been explained by the difference in the degree of stress concentration to the grain boundary.

References

1. Cottrell AH (1953) A note on the PorteVin-Le Chatelier effect. Philos Mag 44:829
2. Sleeswyk AW (1958) Slow strain-hardening of ingot iron. Acta Metall 6:598
3. Ham RK, Jaffrey D (1967) Dislocation multiplication, vacancy accumulation, and the onset of jerky flow during forward and reverse 1 strain in Cu-3.2 at.% Sn. Philos Mag 15:247
4. Worthington PJ, Brindley BJ (1969) Serrated yielding in substitutional alloys. Philos Mag 19:1175
5. McCormick PG (1971) The Portevin-Le Chatelier effect in an Al-Mg-Si alloy. Acta Metall 19:463
6. Korbel A, Zasadzinski J, Sieklucka Z (1976) A new approach to the Portevin-LeCatelier effect. Acta Metall 24:919
7. Brindley BJ, Worthington PJ (1981) Serrated yielding in aluminium-3 % magnesium. Acta Metall 17:1357
8. Korbel A, Dybiec H (1981) The problem of the negative strain-rate sensitivity of metals under the portevin-lechatelier deformation conditions. Acta Metall 29:89
9. Kubin LP, Chihab K, Estrin Y (1988) The rate dependence of the portevin-Le chatelier effect. Acta Metall 36:2707
10. Woo KD, Kim SW, Lou TP (2002) Mechanical properties and deformation behavior of Al-4Mg-0.4Sc-1.5Mn alloy at room temperature. Mater Sci Eng A 334:257
11. Tian B (2003) Ageing effect on serrated flow in Al-Mg alloys. Mater Sci Eng A349:272
12. Wen W, Morris JG (2003) An investigation of serrated yielding in 5000 series aluminum alloys. Mater Sci Eng A 354:279
13. Ziania L, Boudrahema S, Ait-Amokhtara H, Mehennia M, Kedjarb B (2012) Unstable plastic flow in the Al-2%Mg alloy, effect of annealing process. Mater Sci Eng A 536:239
14. Wen W, Zhao Y (2005) The effect of Mg precipitation on the mechanical properties of 5xxx aluminum alloys. Mater Sci Eng A 392:136
15. Nakamura T, Sasaki T, Yoshida S (2013) Analysis of Portevin-Le Chatelier effect of Al-Mg alloy by Electronic Speckle Pattern Interferometry. In: Proceedings of SEM 2013 annual conference, June 2013, Chicago, IL, p 109
16. Ikeno S, Uetani Y, Tada S (1984) Effect of grain size on serrated yield in Al-7 mass Mg alloys. J Jpn Inst Metals 48:1163
17. Chihab K, Estrin Y, Kubin LP, Vergnol J (1987) The kinetics of the Potevin-Le Chatelier bands in an Al-5at%Mg alloy. Scr Metall 21:203
18. Robinson JM, Shaw MP (1992) The influence of specimen geometry on the Portevin-Le Chatelier effect in an Al-Mg alloy. Mater Sci Eng A159:159–165
19. Lebyodkin M, Brechet Y, Estrin Y, Kubin L (1996) Statistical behavior and strain localization patterns in the Portevin-Le Chatelier effect. Acta Mater 44:4531–4541
20. Chatterjee A, Sarkar A, Barat P, Mukherjee P, Gayathri N (2009) Character of the deformation bands in the (A+B) regime of the Portecin-Le Chatelier effect in Al=2.5%Mg alloy. Mater Sci Eng A 508:156–60
21. Ziegenbein A, Hähner P, Neuhäuser H (2000) Correlation of temporal instabilities and spatial localization during Potevin-Le Chatelier deformation of Cu-10at.% Al and Cu-15at.% Al. Comput Mater Sci 19:27–34
22. Reed JM, Walter ME (2003) Observations of serration characteristics and acoustic emission during serrated flow of an Al-Mg alloy. Mater Sci Eng A359:1–10
23. Ranca N, Wagner D (2008) Experimental study by pyrometry of Portevin-Le Chatelier plastic instabilities-Type A to type B transition. Mater Sci Eng A 474:188–196

Chapter 12
Opto-Acoustic Technique to Investigate Interface of Thin-Film Systems

Sanichiro Yoshida, David Didie, Daniel Didie, Sushovit Adhikari, and Ik-Keun Park

Abstract The adhesion strength of nano-scale thin-film systems is evaluated with an opto-acoustic technique. The thin-film specimens are oscillated with an acoustic transducer at a range of audible frequency and the resultant film-surface displacement is detected with an optical interferometer. For each film material, a pair of specimens is prepared; one is coated on a silicon substrate after the surface is treated with plasma bombardment, and the other is coated on an identical silicon substrate without a treatment. For comparison, a bare silicon specimen of the same dimension is tested. All coated specimen show greater film-surface displacement than the bare silicon specimen in the entire frequency range, confirming that the detected oscillation represents the differential displacement at the interface. In some cases, the specimen coated on the plasma treated substrate shows greater film-surface oscillation than the one coated on the untreated substrate. The frequency dependence of the oscillation indicates resonance-like behavior at a frequency orders-of-magnitude lower than the resonance estimated from the bonding strength.

Keywords Thin-film systems • Opto-acoustic technique • Adhesion strength • Metal-silicon interface • Michelson interferometer

12.1 Introduction

Thin film coating is used in a wide variety of engineering systems ranging from semiconductor products to artificial joints, for various purposes such as provision of better lubrication, prevention from erosion, and enhancement of wear resistance. The recent trend is that the film thickness is reduced for better performance of the system. In some cases, the film thickness is a few tens of nanometers for an application where the film surface is exposed to vigorous rubbing motion. With these advancements in applications, the quality control of the film adhesion strength and endurance has become more important than ever. Although the mass percentage of the film itself is low, often a minute damage on the coating ends the life of the entire system. As a good example, slight degradation of the coating of an artificial knee joint can not only shorten the life of the joint but also cause excruciating pain and force replacement of the entire joint.

Prevailing techniques employed for evaluation of adhesion strength can be classified into destructive methods and non-destructive methods. The former includes the pull-off test [1], the bending test [2], the nano-scratch test [3], and the nano-indention test [4]. These methods measure the ultimate strength of coating adhesion. The main stream of the techniques classified as the nondestructive methods is the ultrasonic technology [5–7]. These techniques either detect abnormality in the interface such as partial delamination and defects, or measure the elastic constant of the interface through examination of changes in the acoustic velocity.

Apparently, the destructive methods provide accurate and useful information regarding the ultimate adhesion-strength. As for the evaluation of endurance, however, the information provided by these methods is not necessarily accurate. It is rather rare that external force is applied to the film-surface at a level comparable to the ultimate stress. Usually, the coating is

---

S. Yoshida (✉) • D. Didie • D. Didie • S. Adhikari • I.-K. Park
Department of Chemistry and Physics, Southeastern Louisiana University, SLU 10878, Hammond, LA 70402, USA

Seoul National University of Science and Technology, Seoul, Republic of Korea
e-mail: syoshida@selu.edu; sanichiro.yoshida@selu.edu

degraded under normal use of the system where the applied force is even lower than the yield stress. It is likely that the coating is damaged through repetitive, differential motion of the film to the substrate due to those low-level stresses. Thus, it is important to assess dynamic behaviors of the interface. The ultrasonic methods are advantageous in this respect because they are able to evaluate the elastic constant of the interface. However, the selection of the operating frequency is not straightforward. As will be discussed in the following sections, currently prevailing ultrasonic methods use high frequency ranging hundreds MHz or higher. As the specimen thickness decreases, the wavelength must be shortened to excite an acoustic wave inside the specimen. Since the wave velocity is fixed for a given material, the reduction in the wavelength forces the frequency to increase. In addition, short wavelength provides better spatial resolution. The recent trend in the acoustic technology seems to incline towards further increase in the operation frequency. However, this trend may not work well for evaluation of the adhesion endurance. Usually, elastic materials fail more easily if they experience disturbance at or near the resonance of the elastic property. It is possible that the elasticity of the film-substrate interface is represented by a number of different harmonic dynamics each having different resonant frequency as indicated by a number of molecular-level investigations [8]. Thus, the use of high frequency may work on some of the bonding mechanisms, but not necessarily on the most important bonding. In addition, if the operating frequency is far from the main resonance of the interface, the interaction of the interface with the acoustic wave is low, lowering the sensitivity of the detection. The stiffness found in literature of macroscopic research [2, 9] varies in a wide range. It is hard to predict dynamic behaviors of the interface.

Considering the above-issues, we came up with an idea of using a simple optical interferometer to evaluate the adhesion via analysis of harmonic response of the interface in a range of audible frequency. The specimen was integrated in a Michelson interferometer as one of the end mirrors with the film side facing the beam splitter, and was oscillated with an acoustic transducer from the rear. The resultant oscillatory motion of the film-surface was detected as the corresponding optical path difference behind the beam splitter. In audible frequency, both the film and substrate behave as a rigid body (because of the high stiffness the acoustic wavelength is much longer than their thickness), the film-surface oscillation represents the harmonic response of the interface to the transducer's oscillation. Previously, we used this apparatus to test titanium-coated silicon specimens to study the effect of pre-coating plasma treatment on the silicon surface. These tests clearly differentiated the effect of the pre-coating plasma treatment in terms of different amplitude in the film-surface oscillation and its frequency dependence. In addition, the tests showed resonance-like behaviors in the observed frequency-dependence of the film surface oscillation [10]. To further examine the resonance-like behavior, we changed the mass of the film by adding an extra layer of platinum and gold to titanium. Consequently, we observed that the frequency of the resonance-like spectrum shifted in accordance with the change in the mass. However, some mysteries remained; those resonance-like frequencies correspond to the stiffness an order of magnitude lower than the lowest stiffness found in literatures. In some cases, the pre-coating surface-treated specimens showed greater oscillation amplitude than the untreated case. The question was that at this low frequency range whether the surface treatment could lower the interface stiffness or we observe some other phenomenon irrelevant to the interface properties. To clarify these questions, we tested bare silicon with the same setup, and compared the resultant surface oscillation with the coated case. The aim of this paper is to report these recent advancements and discuss the results of analysis.

12.2 Theoretical Background

12.2.1 Elasticity of Film-Substrate Interface

The elasticity of the film-substrate interface is caused by the restoring force between the atoms involved in the bonding. The inter-atomic force for each bonding can be modeled by spring force, as the potential is approximately quadratic near the equilibrium; hence, the force as the first-order spatial derivative of the potential energy is proportional to the displacement from the equilibrium. There are a number of different modes of bonding each associated with a certain degree of freedom in the atomic motion such as the stretching, and axial rotational mode. Thus, the entire interface can be modeled by a complex combination of these individual springs, each of which has its own spring constant that is different from one another. Molecular dynamic computations [3] indicate many of these modes have substantial contribution to the overall potential representing the interfacial bonding. When external force acts on the film at a certain frequency, those springs whose resonant frequency is close to the frequency of the external force is selectively excited, as the transfer function is greater than other springs. From this viewpoint, the elasticity of the interface can be modeled by a spring with a frequency-dependent spring constant. Figure 12.1 illustrates the situation schematically.
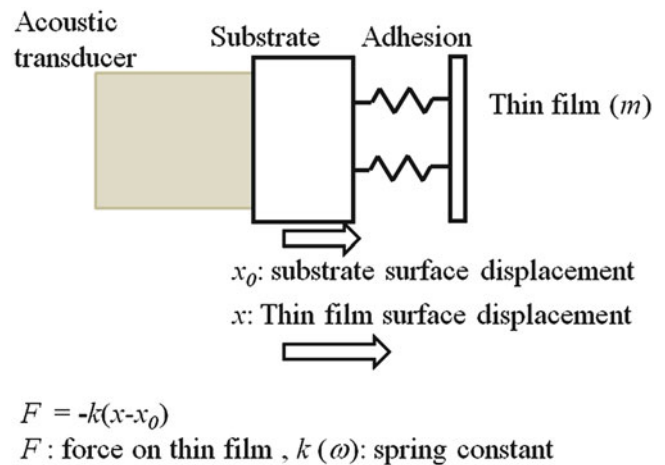
As far as the adhesion-strengthening technology is concerned, the chemistry of the interface has not been fully understood. It is empirically known that if the substrate surface is treated prior to the coating with plasma bombardment, the ultimate strength of the film-substrate adhesion is enhanced. There are two prevailing theories to explain the effect of this type of treatment. The first claims that the treatment changes the film-substrate chemical bonding as the plasma species knock off contaminations that weaken the bonding. The second claims that the plasma bombardment increases the area of the interface and thereby enhances the bonding force.

12.2.2   Acoustic Frequency of Conventional Ultrasonic Techniques

Conventional ultrasonic techniques excite acoustic wave in thin-film systems and evaluate the adhesion strength through detection of the film-surface displacement [11] or the acoustic velocity [6, 7]. These techniques have limitations from two viewpoints. First, the operation frequency is likely to be orders of magnitude higher than the resonance frequency of the interface. Consequently, the output signal tends to be small, and some extra measure is necessary to increase the signal-to-noise ratio. Second, for the reason discussed in the preceding section, the ultrasonic signal selectively detects the interface dynamics whose resonance is close to the acoustic frequency. This lowers the detection sensitivity in practical applications where the dominant frequency of the disturbance to the subject thin-film system is much lower than the acoustic frequency. As an example, the dominant frequency of vertical force on human knees under jogging exercise is in the range of tens of Hz or lower. Therefore, to evaluate the adhesion strength of thin-film coating on artificial joints, the use of ultrasonic technique may be inappropriate, depending on the frequency response of the interface. Below, we discuss these points using specimens used in our experiments as an example.

When an acoustic wave propagates in an elastic medium, the phase velocity v is determined by the Young's modulus E and the density ρ as

$$v = \sqrt{E/\rho}$$

For a given acoustic frequency f, the wavelength $\lambda = v/f$ is

$$\lambda = \frac{\sqrt{E/\rho}}{f}$$

The Young's modulus of materials used for typical thin-film systems (e.g., silicon for the substrate and metals for the film) is on the order of hundreds GPa, or typically 1.8×10^{11} (N/m^2). The density ranges from 2.33 (g/cm^3) for silicon to 21.5 (g/cm^3) for platinum. So, the acoustic phase velocity is in the range of 3 (km/s) to 9 (km/s). If the acoustic frequency is 100 (MHz), for instance, the wavelength becomes 30 (μm) to 90 (μm). If the film and the substrate are 100 (nm) and 100 (μm) thick, respectively, the wavelength is two to three orders of magnitudes longer than the film thickness and

Fig. 12.2 Spectra of acoustic
signal from a knee
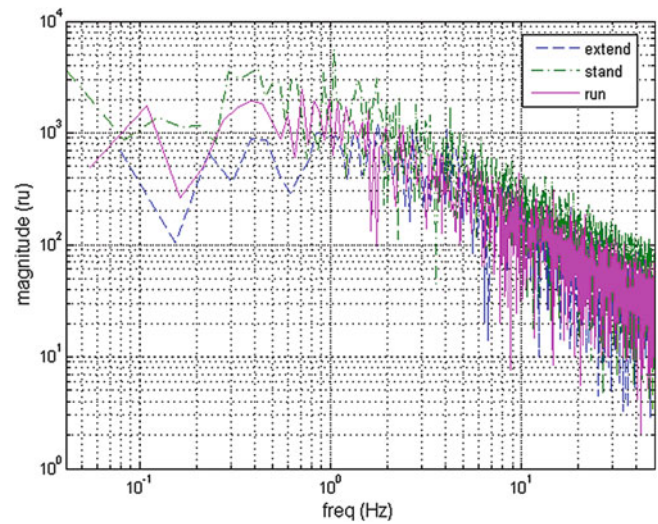
Fig. 12.3 Ground force
to knee acoustic signal
transfer function
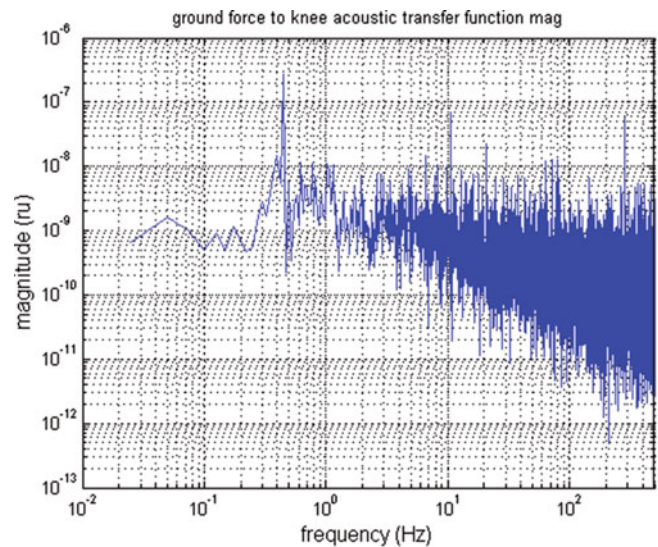
comparable to the substrate thickness. In order to excite an acoustic wave in the substrate, it is necessary to increase the frequency by a factor of two to three. To excite a wave in the film, the operating frequency ought to be in hundreds of GHz.

On the other hand, the elastic constant of the interface is considered much lower. The loading characteristics of the 100 (μm) diameter columnar SU-8 photoresist specimen Ishiyama [9] tested indicates a shear modulus on the order of 20 (kN/m). Assume that the Young's modulus of this case is on the same order of magnitude, say a factor of two or so higher and be 50 (kN/m). Since, the area of 100 (μm) diameter is 7.85×10^{-9} (m^2), and the mass of platinum of 100 (nm) is 1.69×10^{-11} (kg), the resonant frequency of the interface when 100 nm of platinum film is coated will be $(1/(2\pi))\sqrt{(50 \times 10^3)/(1.69 \times 10^{-11})}$=8.7 (MHz), which is orders of magnitude lower than the acoustic frequency.

12.2.3 Near Resonance Behavior Knee Data

As an example of realistic case where a thin-film system is exposed to low frequency rubbing motion, consider the application to artificial knee joints. This is one of the emergent fields of application where the quality control is difficult. According to a study of 33 PCA (Porous-coated anatomic) inserts, the group noted that 53 % PCA's showed severe delamination within 4 years of use [12]. Figure 12.2 are Fourier spectra of acoustic signal detected by an acoustic sensor on the femur side of the knee when the subject extends one leg from the knee, stands up, and runs on a force plate. Figure 12.3 is

the transfer function from the force to the acoustic signal for the running case in Fig. 12.2, where the force signal is detected by the force plate simultaneously with the acoustic signal. Note that dominant frequency components are less than 1 Hz in Fig. 12.2, and that the transfer function's largest peak is at 0.4 Hz. The subject runs at the pace faster than 8 km/h. These results indicate that the friction that the knee joint experiences in this test is mostly slower than 1 Hz.

12.3 Experimental

Table 12.1 summarizes the specimen used in this study. Each film material was coated on a silicon substrate ([1,0,0] plane) prepared with the pre-coating plasma treatment (the treated specimen) and on a silicon substrate of the same size without a pre-coating surface treatment (the untreated specimen). The films were initially coated on the silicon substrate as a disk of 75 mm in diameter and cut into an approximately 5 mm × 5 mm square specimen.

With the setup shown in Fig. 12.4, we acquired a number of data for all types of the specimen shown in Table 12.1. As control data, we repeated exactly the same procedure using a bare silicon specimen of the same size. The data acquisition procedure and the working principle of the setup to evaluate the amplitude of the film surface oscillation are described elsewhere [10, 13]. In short, the procedure is as follows:

(1) Align the beam splitter and the two end mirrors (one of the mirror is a specimen's coated side) so that the laser beams of the respective interferometric paths completely overlap each other on the screen placed behind the beam splitter.
(2) Tilt one of the end mirrors slightly so that the optical path difference between the two arms are displayed as parallel interferometric fringes (dark stripes). The inserts in Fig. 12.4 show the slightly tilted optical paths and a sample fringe pattern.
(3) Turn on the acoustic transducer to oscillate the specimen at a frequency in the range of 5–30 kHz. Since the image acquisition rate (limited by the CCD camera's image-capturing rate of 30 frames per second) is much lower than the acoustic frequency, the fluctuation of the relative optical-path length caused by the film-surface oscillation lowers the fringe visibility. (The oscillation shifts the dark stripe locations at a rate three orders-of-magnitude higher than the image-capturing rate.)
(4) Each captured fringe pattern is taken into computer memory where the spatial intensity variation is Fourier transformed to a spectrum of spatial frequency. The spatial periodicity of the dark stripes appears as the spectral peak of the lowest

Table 12.1 Specimen properties (specimen size: 5 mm × 5 mm)

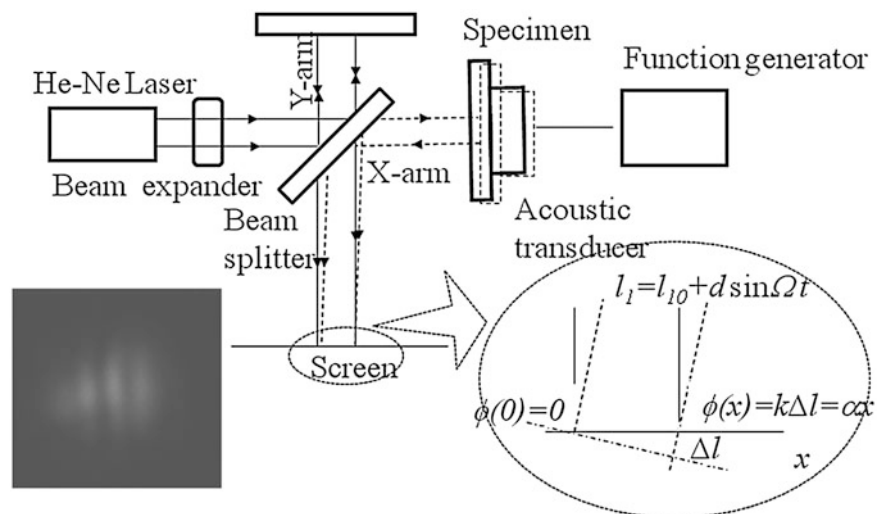|	Ti	Ti-Au	Ti-Pt
Thickness (nm)	75	10/100	10/100
Density (kg/m^3)	4,510	4,510/19,030	4,510/21,450
Film total mass (kg)	8.4×10^{-9}	5.0×10^{-8}	5.5×10^{-8}

Fig. 12.4 Experimental arrangement

frequency. As the fringe visibility reduces due to the film-surface oscillation, this spectral peak is broadened because the spatial periodicity deviates from the most purely sinusoidal pattern obtained when the film-surface is still (the relative optical path length does not change over time). Thus, the change in the fringe visibility is quantified through the relative peak height compared with the maximum peak height (corresponding to the transducer off case).

12.4 Results and Discussions

Figure 12.5 plots the frequency dependence of the relative peak height [defined in procedure (4) above] for the Ti-Au specimens and the bare silicon specimen for comparison (Table 12.1). The horizontal axis is the frequency of the acoustic transducer used to oscillate the specimen. To check the reproducibility of the data, we (a) reexamined the same untreated specimen (#1) as our previous studies [10, 13] and tested another untreated Ti-Au specimen (#2) cut from the same parent disk. Both the untreated and treated specimens showed lower relative peak values than the bare silicon, as expected. The coated specimens show similar features to the previous studies in that the untreated specimen shows lower peak values in the entire frequency range tested and that the uncoated #1 specimen shows valleys around 11 and 22 kHz. In addition, like the previous study, the two valleys are similar in the shape indicating the possibility that the latter can represents the second harmonic of the former. The untreated #2 specimen shows valleys around 12 and 25 kHz, which are not too different from the specimen #1. However, in this case the valley at the higher frequency was greater (deeper) than the lower frequency, which denies the possibility that the higher frequency valley represents the second harmonic of the lower frequency valley. The source of the discrepancy observed between the two untreated specimens is unclear at this time. A possibility is that since the two specimens were cut from different part of the parent disk, the state of adhesion was indeed different from each other for some location-dependent factors such as residual stresses induced during the coating process.

Figure 12.6 compares the relative peak height of the Ti-Au, Ti-Pt and Ti specimens, each compared with the bare silicon case. All the coated specimens, regardless pre-coat-treated or not, commonly show lower relative peak height than the bare silicon case. This indicates that observed reduction in the Fourier spectrum peak is caused by the coating, representing the elastic property of the film-substrate interface. In the case of Ti-Pt specimen, the treated specimen shows much lower relative peak height in a wide range of 9–12 kHz. This behavior is completely different from the Ti-Au specimens in spite of the similarity in the film mass. The reason for the difference is unclear at this point. One interesting point of argument is that this frequency range is a about 2–2.5 times lower than the frequency of the valley observed in the Ti-treated specimen, and this ratio is close to the square root of the mass ratio of the Ti-Pt film to the Ti film (see Table 12.1). Assuming that the elasticity of the interface of the Ti and Ti-Pt is the same, this indicates the possibility that these valley frequencies represent resonance like behavior of the interface (i.e., the valley frequency is proportional to the square of the ratio of the spring constant to the mass).
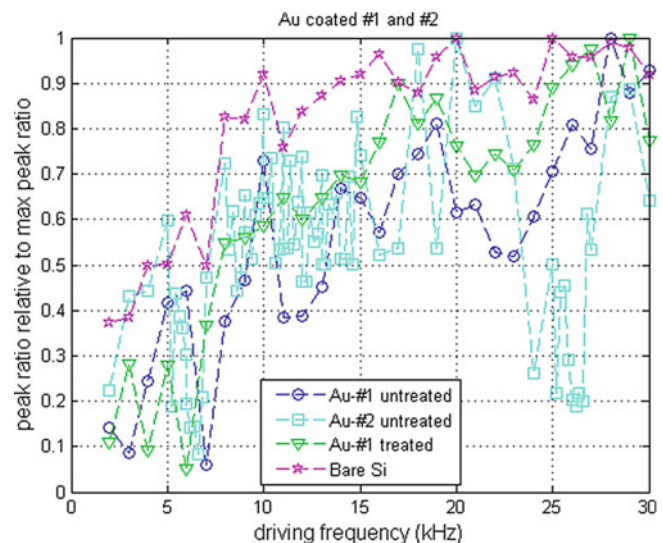
Fig. 12.5 Relative peak height of Ti-Au specimens

Fig. 12.6 Comparison of relative Fourier spectrum peak height of Ti-Au (*top*), Ti-Pt (*middle*) and Ti (*bottom*) specimen along with Bare Silicon specimen
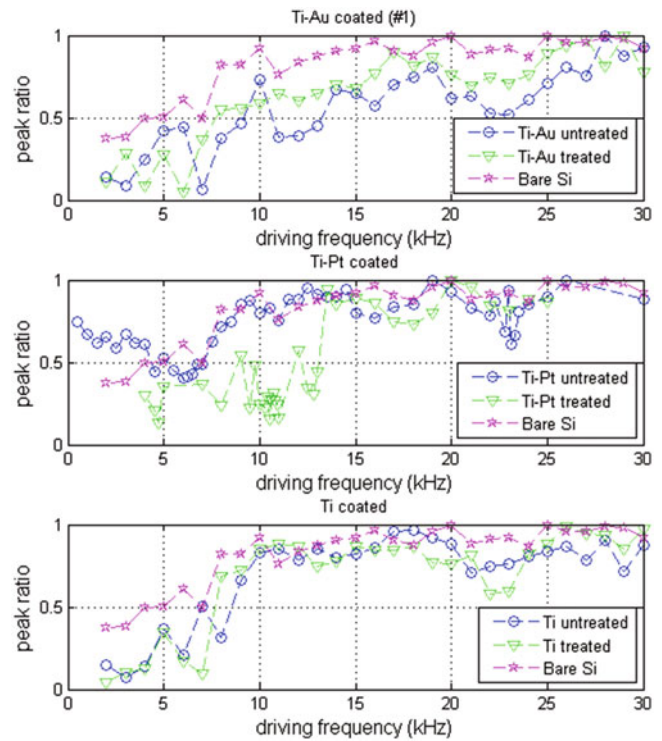
Fig. 12.7 The Fourier spectrum peak height of Ti-Au (*top*), Ti-Pt (*middle*) and Ti (*bottom*) specimen relative to Bare Silicon specimen
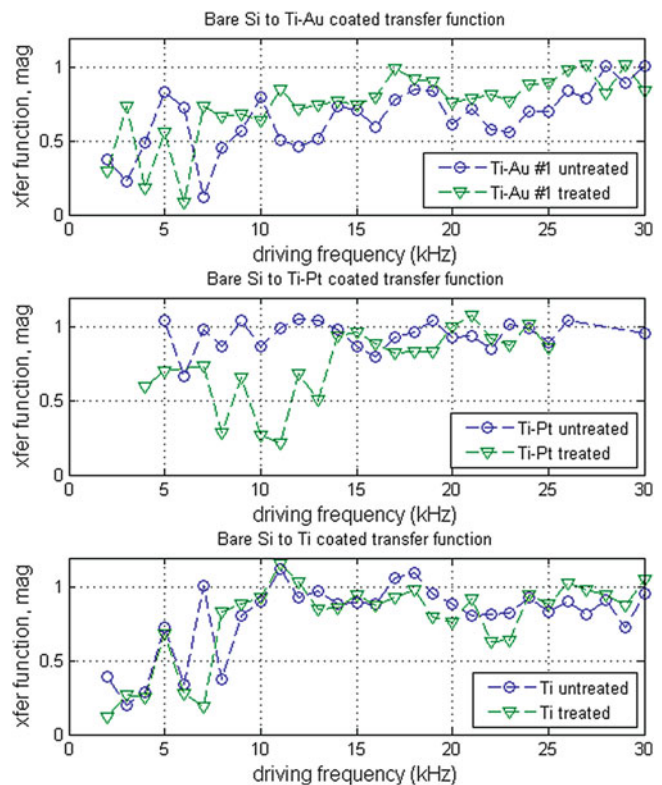
The ratio of the peak height of the coated specimen to that of the bare silicon provides us a quantity similar to the transfer function of the substrate displacement to the film displacement. In Fig. 12.7, this transfer function-like value is plotted for all the coated specimens in the same order as Fig. 12.6. It is seen that in the case of Ti-Pt and Ti specimens, the treated specimens show lower peak height than the untreated ones in a wide frequency range. This tendency is completely different from the Ti-Au case.

Fig. 12.8 Stiffness of interface and corresponding resonance estimated for various cases
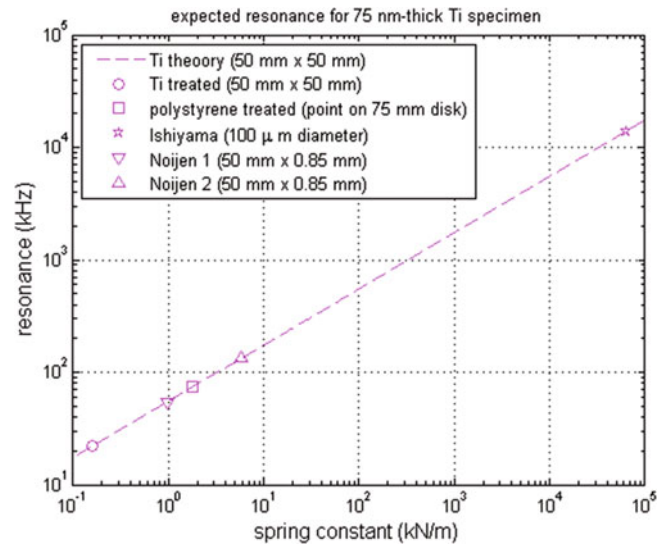
Based on the interpretation that the valleys observed in Fig. 12.7 represent resonance-like behavior of the interface (the film material is connected to the substrate via the spring-like interface), it is possible to estimate the stiffness (the spring constant) of the interface from the valley frequency and the mass. Figure 12.8 shows the relationship between the spring constant and the resonant frequency for various cases assuming that the 75 nm thick 5 mm $\times$ 5 mm titanium film is connected to the spring. The dashed line is a theoretical value based on that the resonant angular frequency is equal to the square root of the ratio of the spring constant to the mass. The circle is the above shown experimental result of this study. The square is our previous study [14] in which we estimated the resonance frequency of the interface between the same type of silicon substrate as this study and a polystyrene film based on a single acoustic frequency to oscillate the specimen and Doppler Vibrometry to detect the film-surface displacement. The triangles are the spring constant measured by Noijen et al. [2] with the four-point support method, and the pentagon is the above-mentioned Ishiyama's [9] measurement on the elastic constant of the columnar photoresist adhered to the silicon substrate. Note that all the spring constants are normalized to the cross-sectional area of 5 mm $\times$ 5 mm. Since for all the cases it is assumed that the same titanium film is connected to the respective spring and the relationship between the spring constant and the resonant frequency is plotted, all the data point should be on the theoretical line. The purpose of this plot is to compare the range of resonant frequency among the cases.

Figure 12.8 indicates that the resonance frequency observed in this study is in the same order of magnitude as that estimated from the Noijen's four-point support analysis [2] and our previous study on the polystyrene film system. However, it is three orders of magnitude lower than the adhesion strength measurement made by Ishiyama [9] using a columnar specimen. The source of such a huge difference in the resonance frequency is unclear at this point. It should be noted that the cross-sectional area of the interface for the Ishiyama case (100 μm in diameter) is four-order of magnitude smaller than this study (5 mm $\times$ 5 mm) or the effective area of the Noijen's experiment (50 mm $\times$ 0.85 mm). A possible explanation is that the low resonant frequency, i.e., the corresponding low spring constant of the interface, is due to imperfection of the bonding. As the cross-sectional area is increased, the chance of imperfect bonding increases. It is possible that residual stresses [15] are induced during the coating process, and they cause plastic deformation generating dislocations.

Kim et al. [7] analyzed the adhesion of the identical Ti-Pt specimen as this study using an acoustic scanning microscope using the so-called V(z) analysis method [6]. In this method, a Surface Acoustic Wave (SAW) is excited near the interface. From the interference of the SAW and a longitudinal acoustic wave reflected from the substrate, the SAW's phase velocity is estimated. Since the SAW's phase velocity is related to the elasticity of the interface, the adhesion strength can be evaluated from the deviation of the measured SAW velocity from a reference value. They find the SAW velocity from the treated Ti-Pt specimen (4,323.99 m/s) is 1.5 % higher than the untreated specimen (4,259.55 m/s), and conclude the effectiveness of the surface treatment. These phase velocity is in the same order as the bulk titanium (6,100 m/s) or platinum (3,260 m/s), indicating that the elastic property of the interface is essentially represented by the Young's modulus of the same order as the bulk cases. While direct comparison of the spring constant and the Young's modulus is not easy as the thickness of the interface (the length of the spring) is unknown, it is clear that the spring constant corresponding to these SAW velocities is several orders of magnitude higher than those discussed above in the context of the valleys observed in Fig. 12.7. A possible explanation regarding this huge difference is that the SAW is sensitive to the atomic bonding at the surface itself and the interferometer is sensitive to the imperfection of the bonding. When a SAW travels near an interface with imperfect

bonding (atomic scale delmaination), substantial part of the wave energy will be lost by reflection. It will be interesting to measure the transmitted SAW energy along with the SAW velocity, and assess the energy loss from the viewpoint of bonding imperfection.

12.5 Conclusions

Comparison with the bare silicon specimen confirms our previous conclusion that the valleys observed in the frequency dependence of the relative peak height of the Fourier spectrum of film surface oscillation is indeed due to the film-substrate interface. The mystery regarding the resonance-like behavior observed at orders of magnitude lower frequency than that derived from the direct measurement of the interface stiffness remains.

References

1. Turunen MPK, Marjamaki P, Paajanen M, Lahtinen J, Kivilahti JK (2004) Pull-off test in the assessment of adhesion at printed wiring board metallization/epoxy interface. Microelectron Reliab 44:993–1007
2. Noijenh SPM, van der Sluis O, Timmermans PHM (2012) An extensive investigation of the four point bending test for interface characterization. 13th international conference on thermal, mechanical and multi-physics simulation and experiments in microelectronics and microsystems, EuroSimE 2012
3. Laugier MT (1984) An energy approach to the adhesion of coatings using the scratch test. Thin Solid Films 117(4):243–249
4. Bunett PJ, Rickerby DS (1988) The scratch adhesion test: an elastic-plastic indentation analysis. Thin Solid Films 157(2):233–244
5. Weglein RD (1980) Acoustic microscopy applied to SAW dispersion and film thickness measurement. IEEE Trans Sonics SU-27(2):82
6. Atalar A (1978) An angular-spectrum approach to contrast in reflection acoustic microscopy. Appl Phys Lett 49:1530–1539
7. Kim JN, Tutwiler R, Kwak DR, Park I, Miyasakaa C (2013) Multilayer transfer matrix characterization of complex materials with scanning acoustic microscopy. In: Proceedings of SPIE 8694, Nondestructive characterization for composite materials, aerospace engineering, civil infrastructure, and homeland security 2013, 86941O
8. Liao Q, Fu J, Jin X (1999) Single-chain polystyrene particles adsorbed on the silicon surface: a molecular dynamics simulation. Langmuir 15:7795–7801
9. Ishiyama C, Tasaki T, Mark Chan T-F, Sone M (2012) Effects of specimen dimensions on adhesive shear strength between a microsized SU-8 column and a silicon substrate. Jpn J Appl Phys 51: 06FL19-1-6
10. Didie D, Didie D, Ghimire B, Kabza K, Adhikari S, Yoshida S, Miyaska C, Park I-K (2014) Optical interferometry for evaluation of adhesion strength of thin-film systems. Advancement of optical methods in experimental mechanics, vol 3, Conference proceedings of the society for experimental mechanics series, pp 259–266
11. Telschow KL, Deason VA, Cottle DL, Larson JD III (2003) Full-field imaging of gigahertz film bulk acoustic resonator motion. IEEE Trans Ultrason 50:1279–1285
12. McTighe T, Clarke I (2009) Failure mechanisms on total knee arthroplasty. Joint Implant Surgery & Research Foundation, Chagrin Falls, OH
13. Yoshida S, Adhikari S, Gomi K, Shrestha R, Huggett D, Miyasaka C, Park I (2012) Opto-acoustic technique to evaluate adhesion strength of thin-film systems. AIP Adv 2(2):022126-1–022126-7
14. Basnet M, Yoshida S, Tittmann BR, Kalkan AK, Miyasaka C (2010) Quantitative nondestructive evaluation for adhesive strength at an interface of a thin film system with opto-acoustic techniques. 47th annual technical meeting of society of engineering science, Oct 4–6, Iowa State University, Ames, IA
15. Volinsky AA, Moody NR, Gerberich WW (2002) Interfacial toughness measurements for thin films on substrates. Acta Mater 50:441–466

Chapter 13
Analysis of Fatigue of Metals by Electronic Speckle Pattern Interferometry

Shun Hasegawa, Tomohiro Sasaki, Sanichiro Yoshida, and Seth L. Hebert

Abstract Electronic Speckle Pattern Interferometry (ESPI) was applied to analysis of the fatigue of metals. The deformation behavior of stainless steel sheet specimen during high-cyclic fatigue tests was dynamically observed by ESPI. The effect of fatigue progress on the deformation behaviors was discussed by evaluating the change in the ESPI fringe pattern focusing on the fatigue stages of crack initiation and crack propagation. The deformation behavior observed in the fatigue test was changed due to the occurrence of crack, and this phenomenon was observed as the fringe concentration around the tip of crack. In addition, static tensile tests with ESPI were performed for specimens which were previously subjected to the fatigue cycle within the fatigue life. The effect of fatigue damage appeared in the deformation behavior near the yielding; the strain concentration behavior caused by local plastic deformation changed depending on the fatigue stage. It was found that the previous cyclic load promotes the local plastic deformation in the static tensile test.

Keywords Fatigue • Deformation behavior • Speckle interferometry • Analysis • Prediction

13.1 Introduction

It is known that fatigue process of metals mainly consists of two stages of crack initiation and crack propagation [1]. The micro-crack is initiated on the surface of metals resulting from local plastic deformation by the cyclic loading in the fatigue process. The subsequent crack propagation begins when the micro-crack grows a critical size. A number of fatigue estimation methods using x-ray, ultrasonic wave and acoustic emission have been proposed [2, 3]. These methods generally aim to inspect the fatigue damage by detecting the crack at the stage of crack propagation. In this study, a different approach of the fatigue inspection using Electro Speckle Pattern Interferometry (ESPI) [4] is proposed focusing on a change in deformation behavior by the fatigue damage. ESPI is a technique capable of measuring displacement fields of a solid-state object, and has been developed as a non-destructive evaluation of deformation and fracture of materials. Our previous works [5–7] have demonstrated that the deformation behavior of metal previously loaded to a certain level shows different behavior from a non-preloaded metal, and the load hysteresis can be diagnosed by applying re-load test with ESPI. The change in deformation behavior at the re-load test is caused by local plastic deformation in the preloading. Therefore, it is expected that cyclic loading in the fatigue process also affects deformation behavior on the basis of local plastic deformation associated with the crack initiation, and the state of fatigue progress can be estimated at the earlier stage.

In the present study, we have performed the following two experiments with a dynamic observation using ESPI; high cycle fatigue tests in order to investigate the deformation behavior in the stage of crack propagation, and static tensile tests after pre-cyclic fatigue tests in order to evaluate fatigue damage on specimens at stage of crack initiation. The effect of fatigue progress on the deformation behaviors has been discussed.

S. Hasegawa • T. Sasaki (✉)
Department of Engineering, Niigata University (Japan), Ikarashininocho 8050, Nishi-ku, Niigata-shi, Niigata 9502181, Japan
e-mail: tomodx@eng.niigata-u.ac.jp

S. Yoshida • S.L. Hebert
Department of Chemistry and Physics, Southeastern Louisiana University, SLU 10878, Hammond, LA 70402, USA
e-mail: syoshida@selu.edu

13.2 Experimental Details

13.2.1 *Specimen and Optical Setup of ESPI*

The material used in this study was a 0.1 mm thick cold rolled stainless steel sheet (18Cr-8Ni steel, AISI304). The steel sheet was cut into specimens for fatigue tests and tensile tests in the roll direction with an electrical discharge machine. The specimen geometry is shown in Fig. 13.1. The yield load of the specimen obtained from a tensile test was 0.23 kN.

Deformation processes in the fatigue test and the uniaxial tensile test were observed by ESPI. Figure 13.2 shows the optical configuration which has sensitivity to the longitudinal displacement along tensile direction (Y-axis). A dual beam ESPI setup was arranged parallel to the tensile direction. The light source was a continuous wave generated by semiconductor laser, and the wave length was 660 nm. The laser beam was split into two paths by a beam splitter, and the two beams illuminated the specimen under deformation at an incidence of 38°. When the specimen was stretched and the longitudinal displacement occurred on the surface, the intensity of the speckle field was recorded into a computer. The change in the intensity of speckle field was calculated numerically by subtracting each frame of image from a frame captured at a later frame. Consequently, the displacement contour representing the longitudinal displacement was
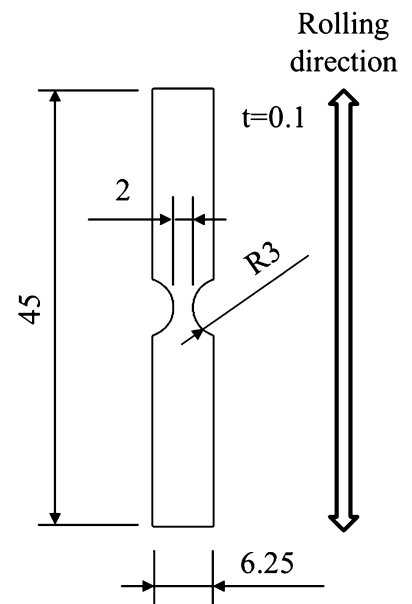
Fig. 13.1 Specimen used for fatigue test
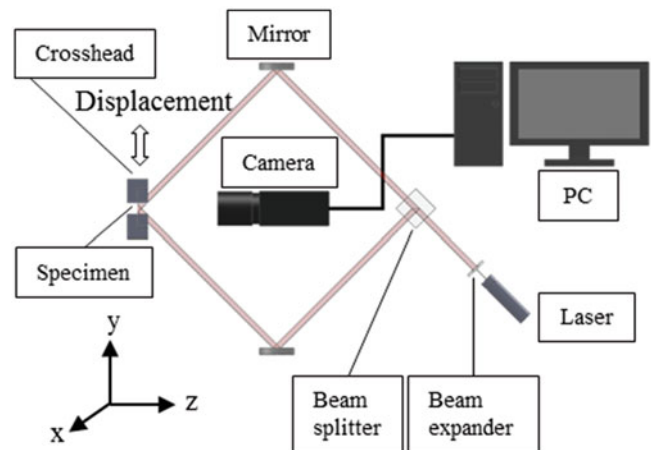
Fig. 13.2 Experimental arrangement

Fig. 13.3 Shape of cyclic load (*solid line*) and capturing timing of images (*solid circles*)
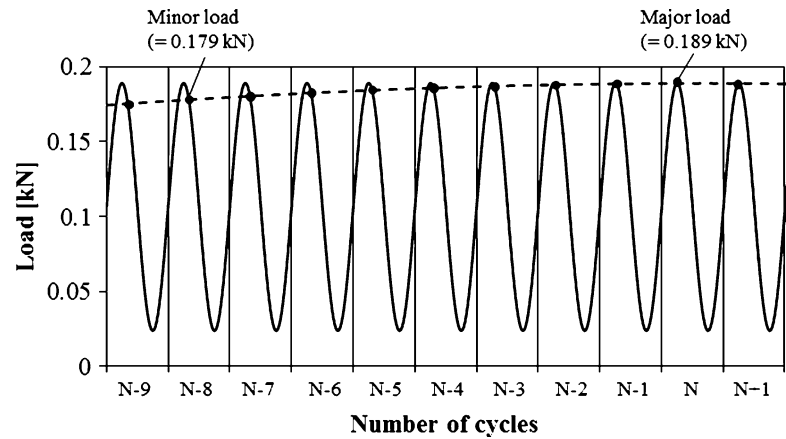
obtained as the fringe pattern. Each interval of the fringe contours represents equal values of the in-plane displacement components. Relationship between the relative number of contour, *n* and the displacement component along the *v,* is given as a following equation.

$$n = \frac{2}{\lambda} v \sin \theta \qquad (13.1)$$

where λ is the wave length of the light source and θ is the angle of incidence of the laser beam. The coordinates in the pixel value of black fringes appeared were measured in the subtracted image, and the fringe number was indexed from the stationary end of the specimen.

13.2.2 Experiment 1: Observation of Displacement Field in Fatigue Test

Load controlled fatigue test until fracture was performed observing the fringe pattern. Figure 13.3 shows the shape of applied load as solid line and capturing timing of images as solid circle. The minimum load was constant at 0.024 kN, and the maximum load was 0.189 kN. The frequency of the load cycle was 10 Hz. To measure the change in the displacement field in real time, the sub-sampling principle [8] was applied to the ESPI observation in the fatigue test. In this method, a slightly lower frequency than the frequency of the actual measurement signal was used for sampling. The CCD camera took the image at the frame rate of 10.1 frame/s, which is slightly higher than the frequency of load cycle. Consequently, in each frame the load at the moment of capture varies slightly. As shown in Fig. 13.3, the broken line connecting these capturing points behaves like decelerated load. The frequency corresponds to 0.101 Hz from the formula of sub-sampling principle.

The observation area was 3.5 mm × 5.0 mm area around the neck part of specimen. Fringe patterns represent the displacement fields resulting from the difference in the applying load of 0.179 kN and 0.189 kN. The time interval between the minor load and the major load included eight cycles (0.8 s). The major load corresponds to the maximum load in a certain cycle (N-th cycle), and the minor load corresponds to eight cycles before the N-th cycle (N-8) as shown in Fig. 13.3. The micro-crack generated by the fatigue test was measured with an optical microscopy in order to investigate the relationship between the fringe pattern and the crack length.

13.2.3 Experiment 2: Observation of Deformation Behavior of Fatigued Specimen in Static Tensile Test

In this experiment, the cyclic load was applied to the specimen within the fatigue life, and the static tensile test was subsequently conducted for the pre-fatigued specimen. The applied load range was identical to that in the experiment 1, and the frequency was 40 Hz. The number of cycles at the fracture of specimen used in this study was about 34,000 under the fatigue condition. Thus, the number of cycle prepared were 500, 3,000, 10,000 and 30,000 respectively. These specimens are

called N-cycle specimen (N is the number of cycle). The static tensile test was conducted at a constant crosshead speed of 1.7×10^{-3} mm/s, and the deformation behavior was observed with ESPI. The frame rate of the CCD camera in this experiment was at 15 fps (frames per second). The deformation behavior of N-cycle specimens was compared with that in the non-fatigued specimen. The time interval to obtain the fringe patterns was a constant at 6.6 s (100 frames in CCD capture). The displacement values at measurement points on the fringe were obtained by Eq. (13.1). The values at the other coordinates in the image were computed by bilinear interpolation. Normal strain, ε_{yy}, was calculated by $\partial v / \partial y$. Finally, the strain rate $d\varepsilon_{yy}/dt$ at each pixel was calculated [7].

13.3 Result and Discussion

13.3.1 Experiment 1

Figure 13.4 shows change in the ESPI fringe pattern during the fatigue test. In this specimen, the fatigue failure occurred at N = 39,933. In Fig. 13.4a corresponding to N = 38,000, two fringes can be observed almost symmetrically to the vertical center of the specimen. This indicates that the neck of specimen was stretched uniformly on both sides. At the N < 38,000, observed fringe patterns were similar to that observed in Fig. 13.4a. At N = 38,500 corresponding to 95 % of the fatigue life, the fringes began to tilt to one side of the neck of the specimen as shown in Fig. 13.4b. The fringes concentrate drastically with the increase of fatigue cycles at N > 38,500 as shown in Fig. 13.4c, d. Figure 13.5 shows the original images corresponding to fringe patterns of Fig. 13.4. There was almost no change in the N < 38,500 as shown in Fig. 13.4a, b, while in the N > 38,500 (Fig. 13.5c, d) a large crack was clearly observed in the neck of the specimen. Through comparison of Fig. 13.4c with Fig. 13.5c, it is found that fringes concentrate near the crack tip at N > 39,500. On the other hand, although a crack was invisible in the original image, the slight tilt in fringes was confirmed at N = 38,500 (Fig. 13.4b). To investigate the relationship between the microstructure and the fringe concentration like Fig. 13.4b, the fatigue test was performed for the other specimen. The fatigue test was stopped when the fringes began to tilt to one side of the neck. Figure 13.6 shows the resultant fringe pattern and the optical micrograph near the neck part where the fringes concentrated. The optical micrograph indicates that the specimen has a micro-crack of about 200 μm. Therefore, it is concluded that the fringe concentration observed in the Fig. 13.6b is attributed to the change in deformation behavior by the crack initiation, and ESPI can detect the micro-order crack.

13.3.2 Experiment 2

Figure 13.7 shows the load-displacement curves during the static tensile tests of 0-cycles, 10,000-cycles and 30,000-cycles specimen, and the typical fringe patterns observed at certain load levels; 75 %, 95 % 100 % of the yield load. The load-displacement curves showed the clear yield points, and the difference in the deformation behavior was confirmed in the
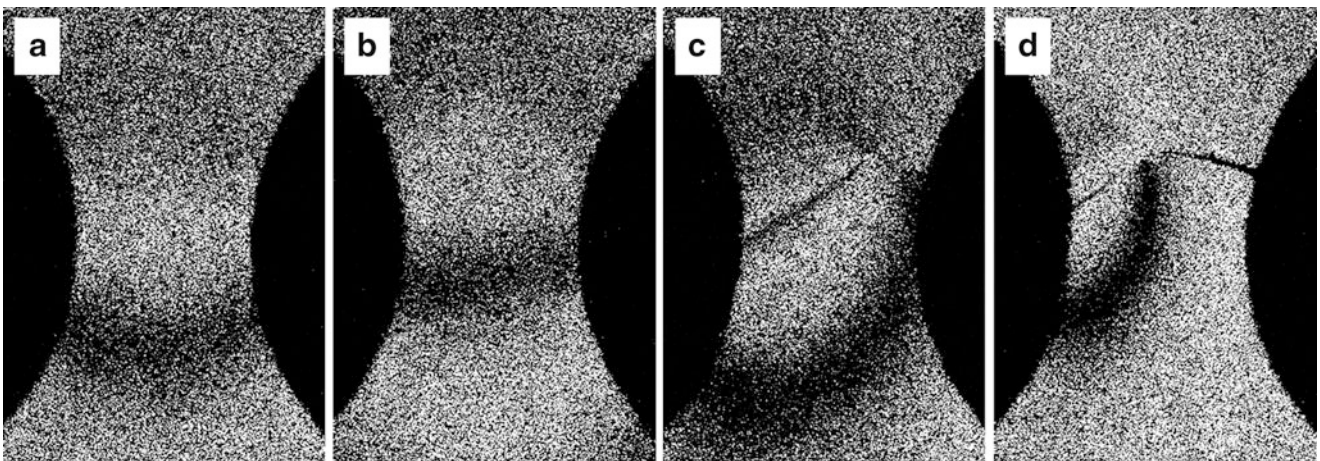
Fig. 13.4 Fringe pattern at the fatigue cycle N of (**a**) 38,000, (**b**) 38,500, (**c**) 39,500 and (**d**) 39,880. The fatigue life was 39,933 cycles
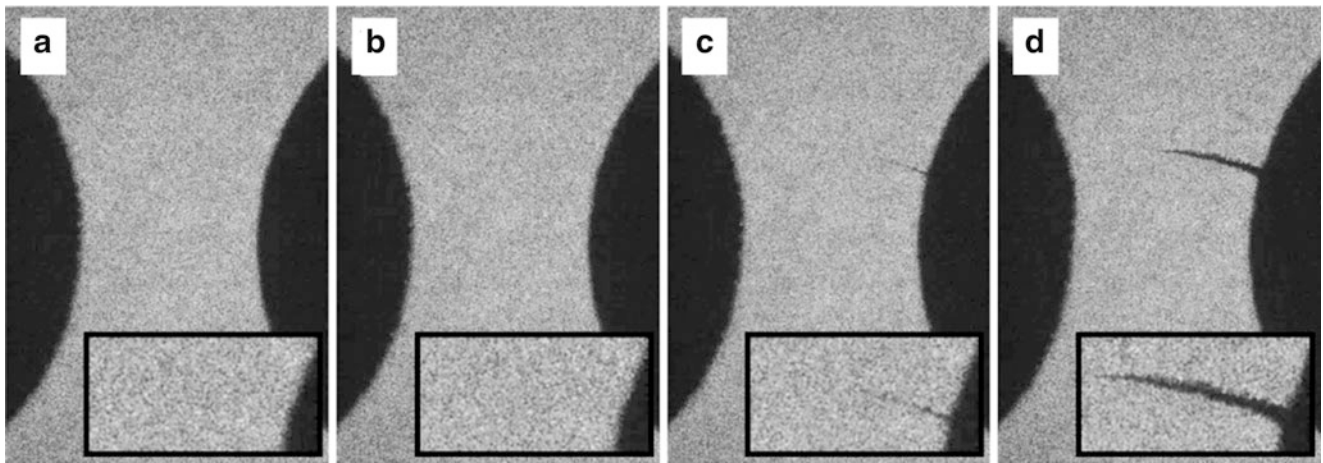
Fig. 13.5 Original images for the fringe patterns in Fig. 13.3: (**a**) N = 38,000, (**b**) N = 38,500, (**c**) N = 39,500 and (**d**) N = 39,880

Fig. 13.6 Fringe pattern in the stage of crack initiation and the crack observed in the neck of specimen. (**a**) was observed at N = 18,200, and (**b**) is the optical micrograph near the fringe concentration area
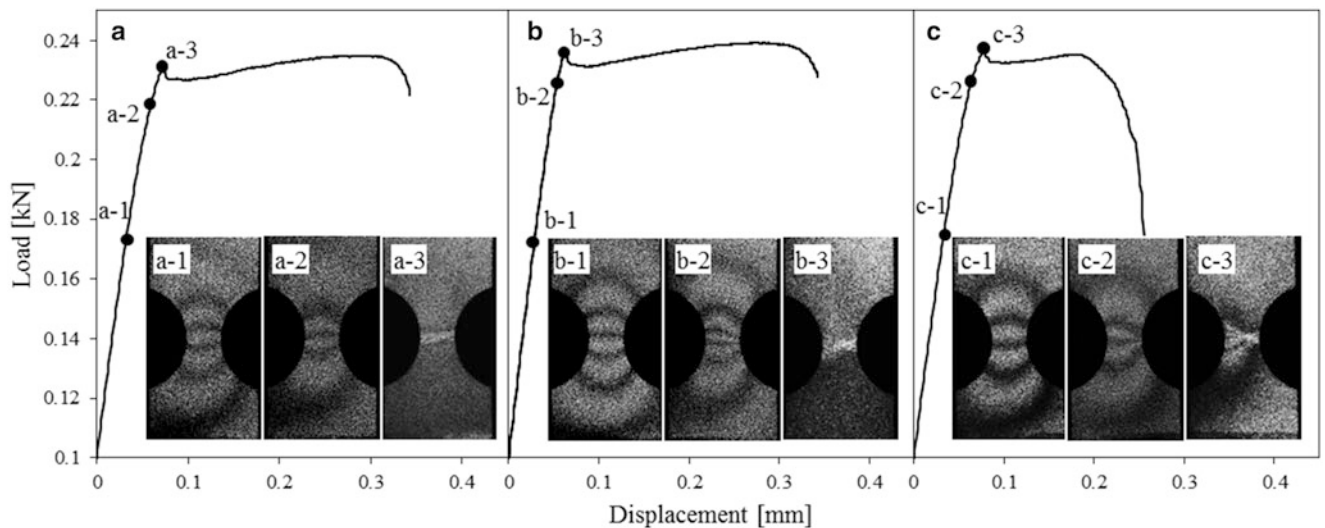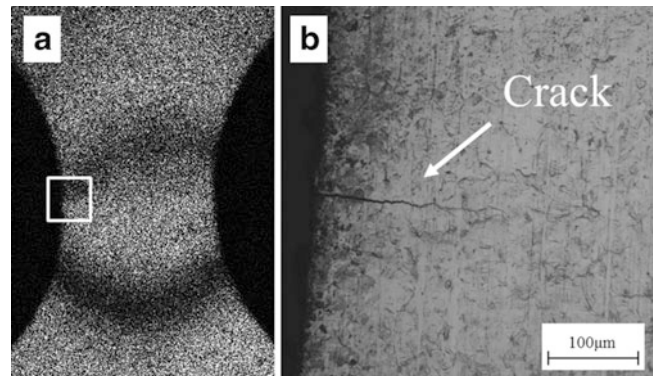
Fig. 13.7 Load-displacement curve and fringe patterns before the yielding. (**a**) 0-cycles specimen. (**b**) 10,000-cycles specimen. (**c**) 30,000-cycles specimen

fringe patterns observed before the yielding. Fringe pattern at 75 % of yield load of the 0-cycles specimen is shown in Fig. 13.7. At the lower load level, fringes distributed uniformly within the observed area (Fig. 13.7a-1). When the load reached near the yield load (95 % of yield load, Fig. 13.7a-2), fringe intervals became narrow at the center of the specimen,
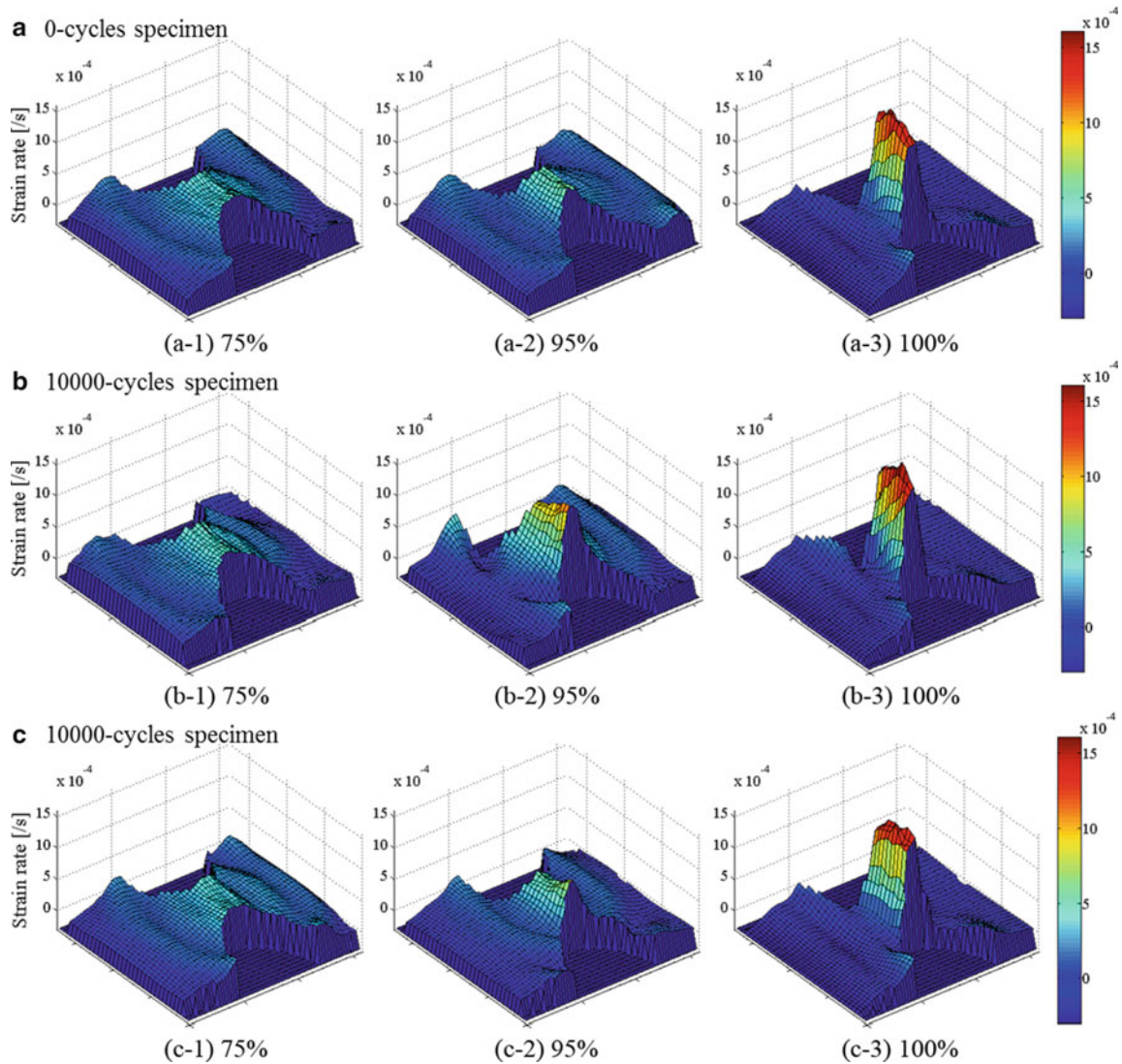
Fig. 13.8 Strain rate map obtained from fringe patterns of (**a**) 0-cycles, (**b**) 10,000-cycles and (**c**) 30,000-cycles specimen for the load/yield load of 75 %, 95 % and 100 %

i.e. the fringes started to concentrate to the neck of specimen. The fringes progressively concentrated as the applied load increases, and form a white band at the yield point as shown in Fig. 13.7a-3. This white band appeared as a result of concentration of fringes to an extremely small region. Fringe patterns of the 10,000-cycles specimen and the 30,000-cycles specimen showed a similar change in fringe concentration at the lower load level (Fig. 13.5b-1, c-1), while the fringe interval at the 95 % of the yield load (Fig. 13.7b-2, c-2) tended to become narrow, although the same load level was applied to the specimens.

The distributions of strain rate in the tensile direction, $d\varepsilon_{yy}/dt$ obtained from the fringe patterns in Fig. 13.7 are shown in Fig. 13.8. The number of fringe in the white band was assumed to be 7 because this number of fringes was observed at the lower load. $d\varepsilon_{yy}/dt$ is nearly constant on the entire area of specimen as shown in Fig. 13.8a-1. The strain rate increases as the applied load increases. It is found that the 10,000-cycles specimen (Fig. 13.8b-2) shows the higher strain concentration than the other specimens (Fig. 13.6a-2, c-2) at the beginning of local plastic deformation. At the yield point, all specimens show almost the same distribution of $d\varepsilon_{yy}/dt$ as shown in Fig. 13.8a-3, b-3, and c-3. The higher strain concentration in 10,000-cycles specimen suggests that pre-fatigue changes the strain concentration behavior observed near the yield load. Thus, change in the strain rate with the applied load was evaluated. Figure 13.9 shows the variation of strain rate in static tensile tests for specimens subjected to cyclic loads of 0, 500, 3,000, 10,000 and 30,000 cycles. The mean strain rate in 0.5 mm × 2.0 mm area around the neck part is plotted against standardized load, load/yield load in the graph. Except for the

Fig. 13.9 Mean strain rate plotted against the load yield load
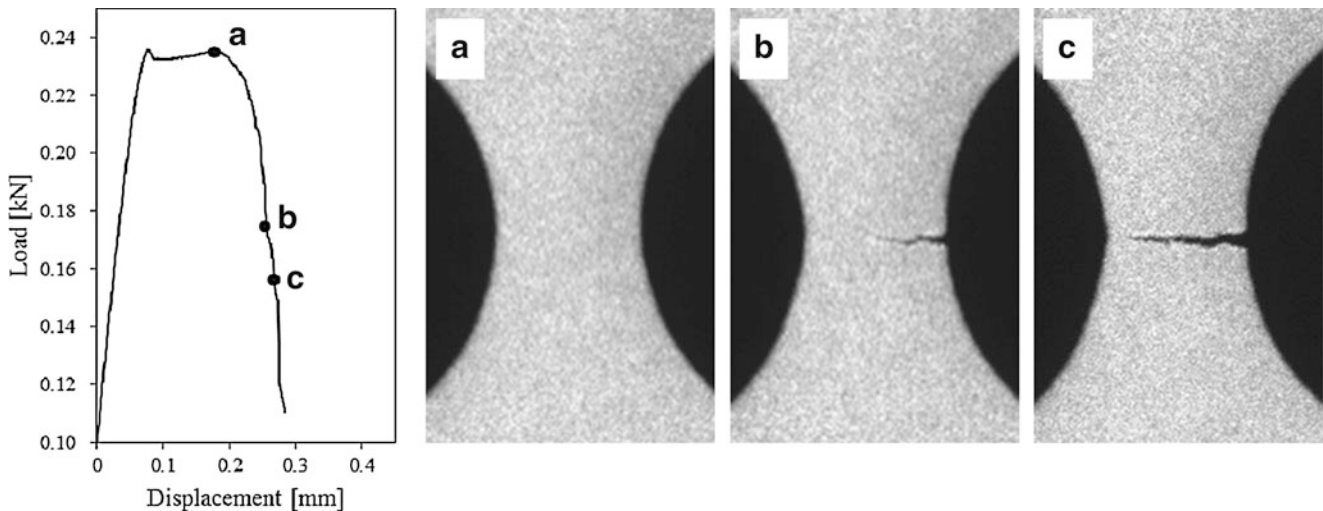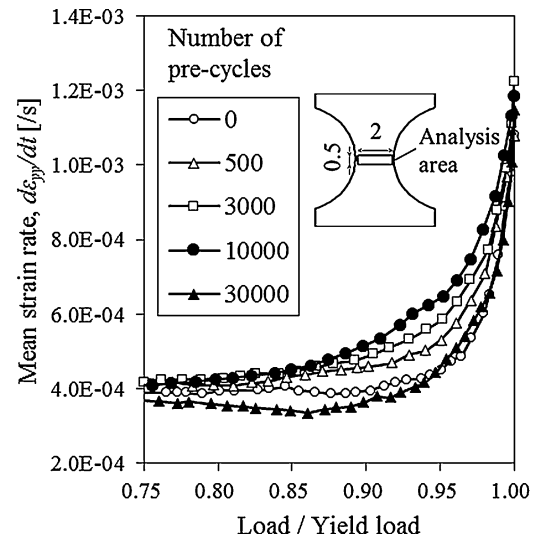
Fig. 13.10 Fracture behavior of 30,000-cycles specimen in tensile test

30,000-cycles specimen, the mean strain rate shows nearly constant at the load/yield load below 80 %, then it starts to increase and reaches up to about 1.2×10^{-3} at the yield load. The start of increase in the mean strain rate appears at lower load levels as the number of pre-cyclic load increases. This fact indicates that the strain concentration to the neck of specimen is promoted by the pre-cyclic load. On the other hand, the mean stain rate in 30,000-cycles specimen decreases at the load/yield load below about 85 %, and it starts to increase at a higher load level than that in the other specimens. The 30,000-cycles specimen showed lower elongation in the static tensile test until the fracture, and the fracture occurred being accompanied by a crack propagation as shown in Fig. 13.10a–c. It is surmised that the 30,000-cycles specimen had a micro crack before the tensile test, and the characteristic behavior of mean strain rate observed in the 30,000-cycles specimen is due to the existence of microscopic crack.

In order to investigate the relationship between increase of mean strain rate and micro structure, the static-tensile test and observation using ESPI were performed for specimens with mirror surfaces. The specimen was prepared by polishing with emery papers and diamond slurry. When the fringes started to concentrate to the neck, the tensile test was stopped, followed by unloading. Figure 13.11 shows the optical micrograph observed before and after the tensile test. Slip bands can be observed in the specimen after the tensile test (Fig. 13.11b). This result indicates that concentration of fringes is responsible for the propagation of local plastic deformation. It is concluded that the difference in deformation behavior observed in the static tensile test is caused by the local plastic deformation in the metals.

Fig. 13.11 Optical micrograph of (**a**) specimen as polished and (**b**) tensiled specimen
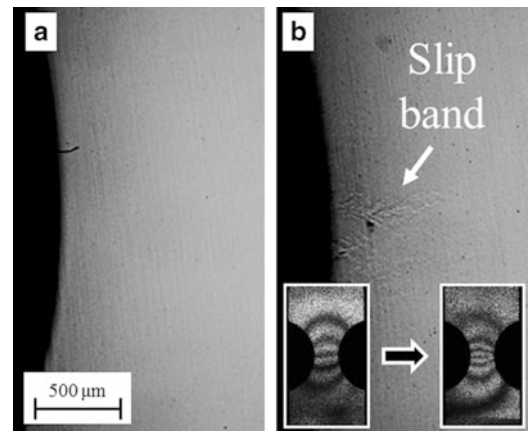
13.4 Summary

Deformation behaviors of stainless steel sheets in the fatigue test and the static tensile test for the specimen which is previously subjected to cyclic load conditions were observed by ESPI. In the fatigue test, the fringes started to concentrate to near the tip of crack at the stage of crack propagation. Micro-cracks caused by the fatigue with the size of 200 μm were able to be detected by observing of the fringe concentration behavior. In the static tensile test, similar fringe concentration behavior was observed near the yield load relating to occurrence of local plastic deformation. It has been shown that cyclic load promotes the fringes concentration, and the deformation behavior in the static tensile test changes depending on the number of cyclic load conditions. In the number of cycles with about 90 % of fatigue failure life, the concentration of fringes showed a different behavior due to the microscopic crack initiation.

References

1. Suresh S (1998) Fatigue of Materials. 2nd edn. Baihukan, in Japanese, Chapter 1, Cambridge University Press, p 1
2. Honda T (2007) Non-destructive inspections for the aging degradation of machinery and structures. J Reliab Eng Assoc Japan Reliab 29(6):350–357, in Japanese
3. Moorth V, Jayakumar T, Raj B (1996) Influence of microstructure on acoustic emission behavior during stage 2 fatigue crack growth in solution annealed, thermally aged and weld specimens of AISI type 316 stainless steel. Mater Sci Eng A 212:273–280
4. Rastogi PK (1993) Speckle metrology. In: Sirohi RS (ed), Chapter 3 Measurement of Static Surface Displacements, Dericatives of Displacements, and Three-dimensional Surface Shapes - Examples of Applications to Non-destructive Testing, Marcel Dekker, New York, p 41
5. Yoshida S, Sasaki T, Gaffiney GA (2009) Dynamics of plastic deformation based on a field theory. Proceedings of the 2009 SEM annual conference and exposition on experimental and applied mechanics
6. Sasaki T, Yoshida S (2012) Revealing load hysteresis based on electronic speckle pattern interferometry and physical mesomechanics. Phys Mesomech 15(1–2):47
7. Sasaki T, Suzuki H, Watanabe T, Yoshida S (2012) Effect of loading history of aluminum on fringe pattern in electronic speckle pattern interferometry. In: Proceedings of the 2012 SEM annual conference and exposition on experimental and applied mechanics
8. Vanlanduit S, Vanherzeele J, Longo R, Guillaume P (2009) A digital image correlation method for fatigue test experiments. Opt Lasers Eng 47:378

Chapter 14
Simultaneous Application of Acoustic and Optical Techniques to Nondestructive Evaluation

Ik-Keun Park, Sanichiro Yoshida, David Didie, Haesung Park, Daniel Didie, and Saugat Ghimire

Abstract Applications of acoustic and optical techniques to nondestructive evaluation of solid-state materials are considered. An acoustic wave is sensitive to materials' elastic behavior and an optical wave is sensitive to displacement. Combination of both techniques allows one to probe the constitutive relations and related behaviors of materials. Basic properties of sound and light from the viewpoint of nondestructive evaluations are reviewed. Our ongoing researches in this direction are introduced and preliminary results are discussed.

Keywords Nondestructive evaluation • Optical interferometry • Scanning acoustic microscopy • Residual stress analysis • Thin-film adhesion evaluation

14.1 Introduction

A number of acoustic and optical techniques have been developed and are applied to nondestructive evaluation of solid-state materials. Recent advancement of technologies in the respective areas allows us to analyze deformation, dynamics, material properties and defects at various scale levels even in the atomistic level. While pursuing further development of respective techniques is important, simultaneous application of both types of techniques has a number of advantages. Light and sound posses fundamentally different properties as a wave. By utilizing them at the same time, we can probe various material properties that cannot be probed by an optical or acoustic method alone.

Essentially, acoustic techniques probe material properties through analysis of the propagation characteristics of acoustic waves. As propagating through the material being inspected, the probing acoustic beam changes its amplitude and phase depending on the acoustic impedance of the material. These changes are monitored as reflected or transmitted acoustic power, or phase retardation detected with an interferometric arrangement. Acoustic waves interact with materials through lattice vibrations. Therefore, the output signal represents the material's elastic property. On the other hand, an optical signal probes the contour or displacement of points of the material under inspection. From the viewpoint of the constitutive relation of materials, acoustic methods are sensitive to the force or stress, whereas optical methods are to displacement or strain. Application of acoustic and optical techniques is able to probe the constitutive property of material, without relying on the constitutive constants.

The idea of combining acoustic and optical methods for nondestructive evaluation is not new. Stegemean [1] reviews various optical techniques that probe the amplitude of Surface Acoustic Waves (SAW's) utilizing the diffraction pattern generated by the SAW. Burov et al. [2] uses Raman-Nath diffraction of a laser flux generated by two counter-propagating SAW and successfully measures the frequency dependence of the SAW amplitude on the order of 10-nm in an acoustic frequency range of 30 MHz. These works were mostly conducted in the 1970s. Since then various optical technologies such

I.-K. Park • H. Park
Department of Mechanical and Automotive Engineering, Seoul National University of Science and Technology, 232 Gongneung-ro, Nowon-gu, Seoul 139-743, South Korea
e-mail: ikpark@seoultech.ac.kr

S. Yoshida (✉) • D. Didie • D. Didie • S. Ghimire
Department of Chemistry and Physics, Southeastern Louisiana University, SLU 10878, Hammond, LA 70402, USA
e-mail: syoshida@selu.edu

as fast imaging systems, e.g., a digital camera with a frame rate higher than one million frames per second have been developed. It is time to revisit these or similar acousto-optical schemes to utilize the advantages of light and sound.

Based on the above ideas, we have recently started a project HAI (Hybrid Acoustic Imaging). The aim of this paper is to report on our preliminary results from the project and discuss various ideas behind the simultaneous applications of acoustic and optics. We hope our report is beneficial to both the optical and acoustic communities.

14.2   Contrasts of Optical and Acoustic Fields

Light waves and acoustic waves compensate each other in numerous respects. In most applications, the phase velocity and frequency are different by several orders of magnitude. Light can travel in vacuum but not most materials. Sound cannot travel in vacuum but travels through most materials. Light interacts with charged particles and sound with lattices. The aim of this section is to reconsider all those differences between the two types of waves from the viewpoint of efficient use of them for nondestructive evaluation of materials.

14.2.1   Temporal Resolution

In usual usages, the optical frequency is orders of magnitude higher than the acoustic frequency. This means that when a light is used to detect the response of a material to an acoustic excitation, such as the generation of surface acoustic waves, the optical signal by nature has a sampling rate much higher than the dominant frequency of the dynamics. In addition, normally the coherence in light waves (all one needs is to use a laser source) is much higher than sound waves. Using interference of coherent light beams, one can detect material's displacement at much higher level than an acoustic method alone.

14.2.2   Spatial Resolution

The wavelength determines the spatial resolution of diagnosis. The attenuation of an acoustic wave in solids increases quadratically with the frequency. The wavelength of a sound wave at 1 GHz is typically of the order of micro-meter. The wavelength of light can easily go into the sub-micrometer range. With the use of phase (as opposed to amplitude), analysis beyond the diffraction limit is also possible [3]. Acoustic excitation with optical detection has potential to increase the spatial resolution as well.

14.2.3   Propagation Property

In applications to nondestructive evaluation, the material is in most cases not transparent to visible light. Even X-ray diffraction can analyze within limited depth from the surface. Contrary, sound can propagate through most materials. To characterize the properties inside a material, sound is much more powerful than light. On the other hand, light can propagate in vacuum. For remote sensing or testing in a hazardous environment, optical methods are more advantageous to read out the signal.

14.2.4   Interactions with Materials

Sound waves propagate through solids via lattice vibration. Thus, it is sensitive to the material's elasticity. Light can measure displacement or even strain directly. Combination of an optical interferometry and acoustic wave excitation can provide one with the constitutive characteristic of materials without relying on a material constant (such as the elastic constant to evaluate the stress from strain or the compliance from the stress to strain).

14.3 Preliminary Studies and Future Plan

14.3.1 Residual Stress Analysis

One interesting application of acoustic and optical methods is analysis of welding-induced residual stresses with the use of Scanning Acoustic Microscopy (SAM) [4, 5] and Electronic Speckle Pattern Interferometry (ESPI) [6]. Acoustoelasticity [7–11] states that the elastic constant of materials is changed by a stress as it alters the inter-atomic distances. Compressive stresses increase the elastic constant and tensile stresses decrease. SAM can detect changes in elastic constants via measurement of acoustic velocity whose square is proportional to the ratio of the elastic constant to the density. Thus, in principle, SAM can be used to evaluate residual stress due to welding. However, in reality, there are some issues. First, determination of a residual stress based on a SAM measurement at a single point can mislead the overall effect of residual stress on the work. As Fig. 14.1 illustrates schematically, SAM measures the phase velocity of a surface acoustic wave (SAW) by focusing the acoustic beam on the specimen surface. The area of the focused beam is typically a few mm in diameter. Residual stresses in a welded work are normally distributed near the weld with a complicated spatial pattern where compressive and tensile stresses alternate. Our study indicates that the spatial periodicity of this alternating pattern of compressive and tensile residual stress varies on a case-by-case basis depending on the welding conditions and materials. Therefore, if a given SAM analysis indicates that the residual stress on a certain point of the specimen is compressive because the measured phase velocity is higher than the nominal value, it is possible that the phase velocity at a neighboring point is lower than the nominal value. Indeed, a previous SAM measurement (T. Sasaki and M. Usui, private communication, 2014) indicates that the SAW velocity can vary as much as a few % by moving the point of analysis near the weld line for about a half of the acoustic beam size. To evaluate the pattern of compressive and tensile residual stress accurately, it is necessary to scan over the specimen with an increment of the focused acoustic beam size or smaller. This is unrealistic.

Second, the residual stress based on the conventional acousticelasticity can be greatly overestimated. Residual stresses are locked in the material. In most cases, permanent deformation of the surrounding area is likely the main mechanism of locking the elastic deformation (reversible stretch or compression) at the location of SAM measurement. Conventional acousticoelasticity ascribes the change in the elastic constant observed as a change in the SAW velocity from the nominal value to the third-order elastic constant, assuming that all the atoms of the area uniformly change their inter-atomic distance. If a large portion of the area is plastically deformed, the evaluation of residual stress based on this principle is erroneous because the reduction in the elastic constant is not caused by the uniform displacement of all the atoms of region from the equilibrium position. It is possible that those elastically bonded atoms are actually still at the equilibrium and the apparent elastic constant of the area is reduced because the number of these "elastic" atoms is small. In that case, there is no elastic residual stress in the area of measurement. From the viewpoint of practical engineering, elastic and plastic residual stresses affect the welded work differently. If elastic, the residual stress exerts force on neighboring materials continuously, affecting the deformation and even shortening the life of the work. On the other hand, a plastic residual stress exerts no force on the
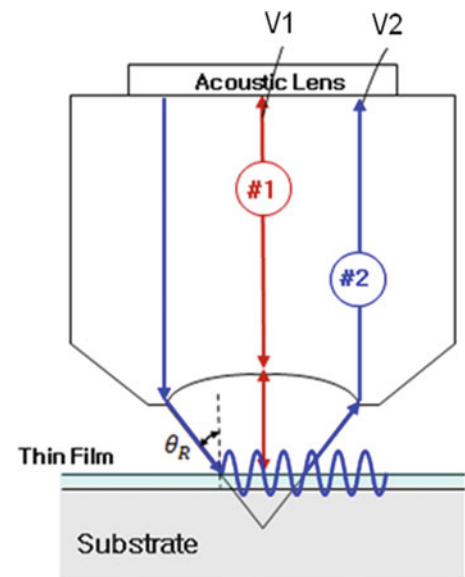

Fig. 14.1 Scanning Acoustic Microscope used for residual stress analysis and thin-film analysis. For residual stress analysis, specimen was placed where "Thin Film" is placed in above figure. Phase difference between path #1 and #2 changes as specimen is moved toward acoustic lens

Fig. 14.2 (**a**) Welded specimen and (**b**) SAW velocity near weld line. The *weld line* is center of the specimen running vertically in above figure. Nominal phase velocity was measured 7 mm away from weld line. Other SAW velocity data were taken along a line parallel to and 3 mm away from weld, 2 mm and 3 mm from center
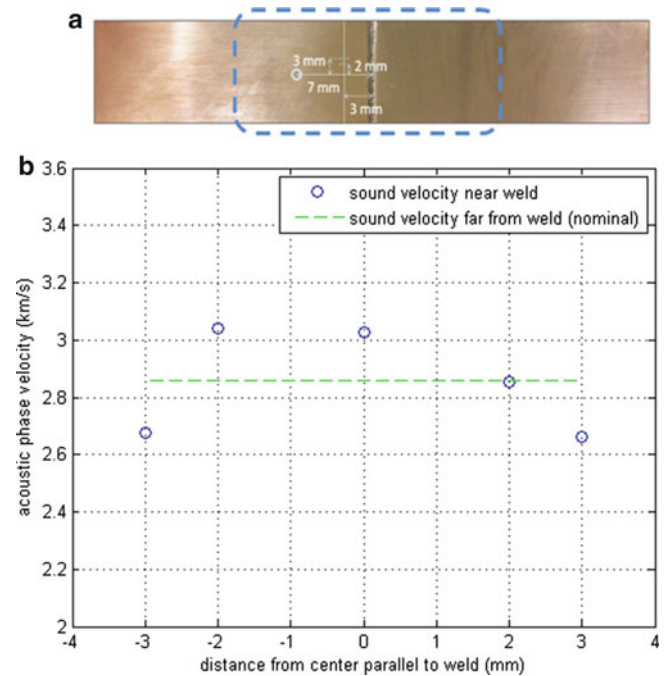

surrounding materials. From this viewpoint, it is less harmful to the welded work than an elastic residual stress. However, if the plastic deformation due to the residual stress is sever, that particular point can be the weakness that leads to a failure of the work. It is important to distinguish elastic residual stresses from plastic residual stresses. Since acousticelasticity is sensitive to only force, there is no way to distinguish a given reduction in the elastic constant is due to elastic or plastic deformation. Independent measurement is necessary to determine if the deformation is elastic or plastic.

Application of ESPI solves the above issues. By applying an external tensile load to a welded specimen and measuring the resultant displacement with an in-plane sensitive ESPI setup, one can evaluate the resultant strain as a whole-field image of the specimen. The observed strain reflect the elastic constant at all the point on the specimen; if the elastic constant is low, the strain is large. This solves the first issue of point-to-point variation. By keeping the external load at a level substantially lower than the residual stress, it is possible to make the external load not cause additional plastic deformation to the specimen. By applying the same external tensile load after releasing the specimen from the initial load and analyzing the reproducibility, it is possible to determine whether the observed strain is plastic or elastic at each point on the specimen. If the strain is not reproduced in the second loading at a certain area of the specimen, it is likely that the specimen has been plastically deformed by the welding.

With this idea in mind, we conducted the following experiment on butt-welded specimens. We measured residual stresses near the weld line without applying an external load to the specimen using a typical SAM setup (Fig. 14.1). We applied a tensile load to another butt-welded specimen and a non-welded specimen of the same material and dimensions at a minimum stress level (much lower than typical residual stress) and measured the resultant strain as a whole image. Figure 14.2 shows the phase velocity measured with the SAM near the weld line parallel to it. The dashed line indicates the SAW velocity measured at a certain distance away from the weld where the effect of residual stress is negligible. Figure 14.2 indicates that the phase velocity is altered as much as ±10 %. Since the elastic constant is proportional to the square root of the phase velocity, this indicates that the residual stress alters the elastic constant of this specimen accordingly, and the change must be reflected on the strain data obtained with the ESPI.

We conducted an ESPI experiment on a welded specimen shown in Fig. 14.2 (100 mm long × 20 mm wide × 0.4 mm thick butt-welded at the middle point along the 100 mm length or 50 mm away from the end). The specimen was attached to a tensile machine for a tensile load applied perpendicular to the weld (the specimen was gripped by the tensile machine at the two 20 mm-wide ends). A conventional, dual-beam ESPI setup sensitive to in-plane displacement along the tensile axis was arranged in front of the specimen. Some initial tensile load was applied until the specimen stopped slipping inside the tensile machine's grips. The first interferometric image was taken at that point with a digital camera. After a load increment of 10 N, or an average tensile stress of 1.25 MPa, the second interferometric image was taken. A fringe pattern was formed by subtracting the second image from the first image. (In the actual experiment, a carrier fringes of known phase was added to the second image to make the fringe analysis easier.) For comparison, the same procedure was taken for a non-welded

Fig. 14.3 Relative compliance observed with ESPI
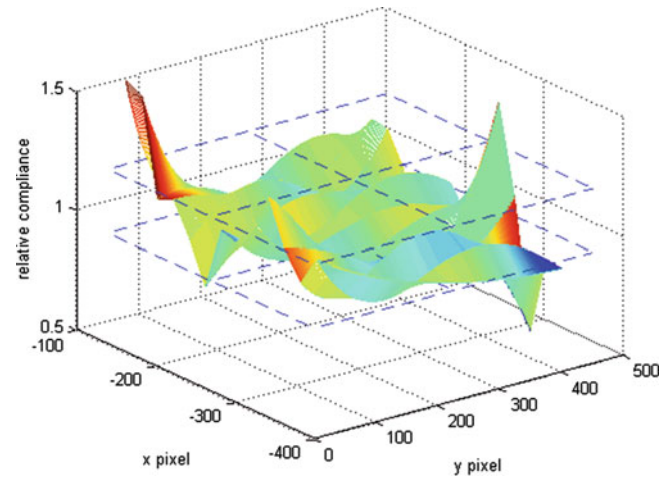
specimen of the same dimension (100 mm long × 20 mm wide × 0.4 mm thick). This procedure generated fringe patterns typically consisting of seven to eight dark fringes. From the wavelength of the laser source (632.8 nm) and the angle of incidence (24.8°) of the laser light onto the specimen, the displacement on each fringe (dark stripe) was calculated. The displacement between fringes was interpolated based on the displacement data calculated on the fringes. In this fashion, a displacement matrix was formed with the two-dimensional pixel arrays (480 × 640) of the digital camera as the grid points. Normal strain along the tensile axis was evaluated by numerically differentiate the displacement data at each grid point (the strain matrix).

Since the same load was applied to the welded and non-welded specimens, comparison of the strain matrix of the welded and non-welded cases represents the difference in the compliance. Figure 14.3 shows the relative compliance (the ratio of the strain measured with the welded specimen to that with the non-welded specimen). The change in the acoustic velocity from the nominal value measured by SAM (Fig. 14.2) can be used to evaluate the range of elastic modulus due to the residual stress. The ±10 % change of the acoustic velocity compared with the nominal value (the value without the effect of residual stress) can be translated into the range of elastic modulus as $\sqrt{0.9}$=0.95 to $\sqrt{1.01}$=1.05, or that of the compliance as $1/\sqrt{1.01}$=0.95 to $1/\sqrt{0.9}$=1.05. This range of compliance is indicated in Fig. 14.3 with dashed lines. It is seen that the relative compliance based on the ESPI measurement is within the range of the compliance evaluated by SAM.

14.3.2 Evaluation of Thin-Film Adhesion Strength

Another area of application is the evaluation of adhesion strength of a thin-film to the substrate. The prevailing nondestructive method to evaluate the adhesion strength of thin-film systems is ultrasonic techniques. The recent trend is that the operation acoustic frequency is increased for higher spatial resolutions. Recently, Kim et al. [12] applied the above SAM technique to a system of platinum-titanium films coated on a silicon substrate (100 nm thick platinum on 10 nm thick titanium coated on a [1, 0, 0] plane of silicon). They tested a pair of specimens; one coated on the silicon substrate with a pre-coating surface treatment (oxygen-plasma bombardment) and the other without a pre-coating treatment. From the 1.5 % difference in the measured SAW velocity, they conclude that the elastic constant of the interface is higher for the surface-treated specimen than the untreated specimen.

While these acoustic techniques are well developed and reliable, the instrument sensitivity is low for elastic dynamics whose resonant frequency is much lower than the operation acoustic frequency. (The transfer function of the acoustic vibration to the film displacement associated with those dynamics is expected to be very small because the acoustic frequency is far away on the high-frequency side of the resonance.) Consequently, it is possible that they overlook elastic dynamics in the interface associated with uniform but much weaker chemical bonding.

Considering the above issue, we came up with an idea of using much lower acoustic frequency to oscillate the thin-film specimen and detect the resultant film surface oscillation with the use of an optical interferometer. We demonstrated the proof of principle and reported a mathematical method to estimate the amplitude of the film surface oscillation [13].

Fig. 14.4 Optical interferometer to evaluate thin-film adhesion
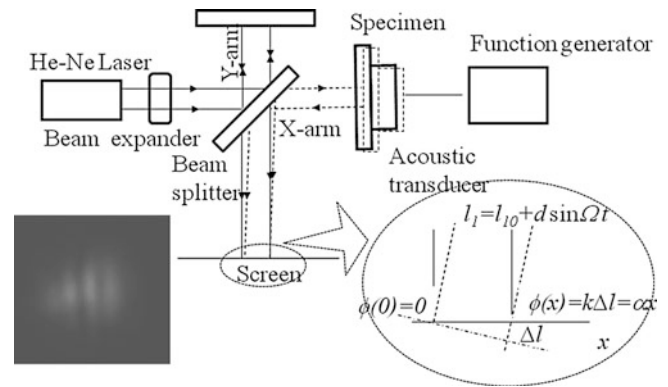
Fig. 14.5 Fourier spectrum peak ratio to bare silicon case
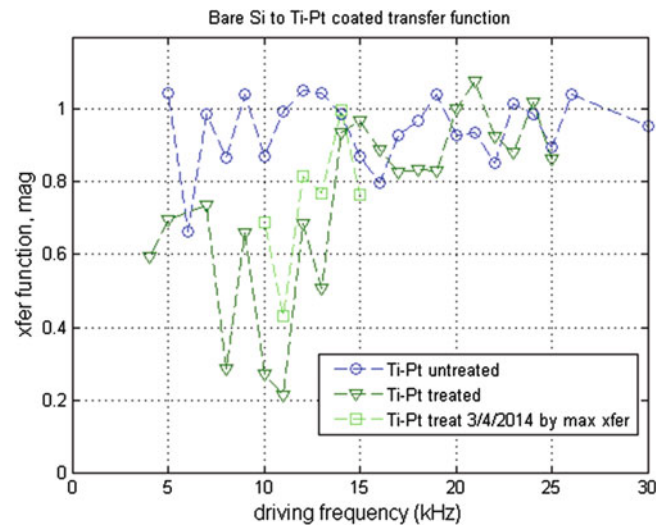
Figure 14.4 illustrates the working principle of this method. A Michelson interferometer is configured with the thin-film specimen as one of the end-mirrors with the film side inward (facing the beam splitter). The thin-film specimen is oscillated with an acoustic transducer from the rear surface of the substrate in a range of audible frequency. The end-mirror comprised of the specimen is slightly tilted so that the optical path difference between the two interferometer-arms is visualized as equidistant, parallel fringes on a screen placed behind the beam splitter. As the acoustic frequency is swept, the film-surface oscillates at various amplitudes in accordance with the substrate-surface to film-surface transfer function. Since the sampling rate (frame rate) of the imaging system is much lower than the audible frequency, the visibility of the fringe system is reduced with an increase in the amplitude of the film surface oscillation (the fringes become blurry). The reduction in the fringe visibility can be quantitatively evaluated from the change in the Fourier spectrum peak associated with the spatial frequency of the fringe system from the case when the film surface is still (no acoustic oscillation). Because of the high elastic constant and small thickness, the silicon substrate or the films oscillate as rigid bodies. (They are much thinner than the acoustic wavelength determined by the phase velocity, which is proportional to the square root of the elastic constant, and the operation acoustic frequency). Thus, the reduction in the fringe visibility, if any, is due to the interface whose elasticity is lower than the substrate or film materials.

We used the interferometer shown in Fig. 14.4 to test the surface-treated and untreated, platinum-titanium thin-film specimens identical to the ones used by Kim et al. [12]. For comparison, a bare silicon specimen of the same size was tested with the same procedure as the thin-film specimens. Figure 14.5 plots the ratio of the Fourier spectrum peak obtained with the thin-film specimens to that obtained with the bare silicon specimen. This ratio represents the oscillation amplitude of the film-surface with reference to that of the silicon surface. Thus, it can be considered as the transfer function of the interface (how much of silicon-surface displacement is transferred to the film-surface). Note that Fig. 14.5 is different from usual Bode plots of transfer function. Since the greater oscillation-amplitude of the film surface causes lower Fourier spectrum peak via lower fringe visibility, the higher the transfer function, the lower the plot appears in Fig. 14.5.

Fig. 14.6 Out-of-plane sensitive ESPI and standing SAW
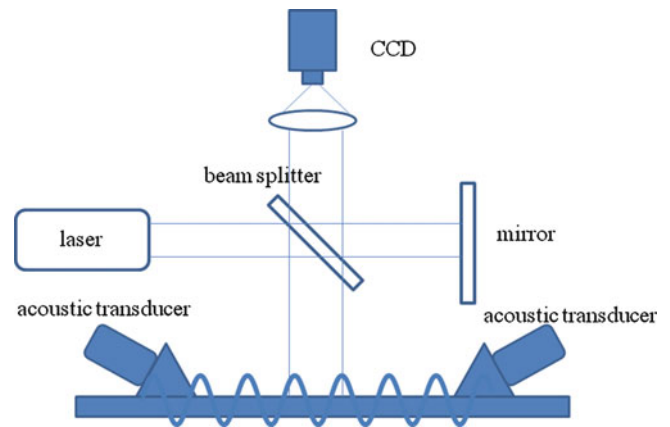
In Fig. 14.5, the circle plots are the data for the untreated specimen and the triangles are for the surface-treated specimen. Based on the above argument, the valley seen in Fig. 14.5 in the frequency range of 5–15 kHz for the treated specimen can be interpreted as representing resonance-like behavior of the interface with a central frequency around 10 kHz. The circle and triangle plots were obtained from independent measurements for the untreated specimen, the treated specimen and the bare silicon specimen. It is possible that the Fourier spectrum peak is affected by the laser light's condition, such as the intensity fluctuation and temporal change of the polarization. To remove this ambiguity, we recently modified the interferometer so that two independent interferometric beams can be formed inside the same interferometer, each reflected off its own specimen. The two interferometric beams are originated from the same laser source, so they are subject to the same fluctuation in the intensity and polarization. The square plots shown in Fig. 14.5 were obtained with this dual-beam configuration, using the surface-treated platinum-titanium thin-film specimen and the bare silicon specimen for each interferometric beam, respectively. Although the dual-beam experiment data are available for a small portion of the frequency-range at this time, the results from the dual-beam configuration appears to be very similar to the previous result (the triangle plots). This agreement indicates the reproducibility of the experiment, and thereby strengthens the possibility that the observed valley indeed represents a resonance-behavior (Fig. 14.5).

Apparently, the resonance-like behavior observed in the treated specimen in Fig. 14.5 is different from the elastic dynamics observed by Kim et al. in [12]. Based on the observed phase velocity of the order of 4,000 m/s and the mass of the film that the interface holds, the expected resonance is of the order of 10–100 GHz. (It depends on the thickness of the interface and these numbers are based on the assumed thickness of 0.1–1 nm). This frequency range is six to seven orders of magnitude higher than the above-mentioned central frequency of 10 kHz. We do not have a clear explanation regarding the underlying mechanism of this resonance-like behavior at such a low frequency range at this time. One possible explanation regarding why this dynamic was not detected by the SAW data in [12] is that this oscillatory dynamics with the central frequency of 10 kHz is not responsive to the operation acoustic frequency of 400 MHz for the above-mentioned reason (the driving frequency is so high that the transfer function is low). More investigation on this mysterious behavior is currently conducted. The importance is that some phenomenon that is not visible by the SAM setup is being observed by the interferometric setup with the use of much lower acoustic frequency.

14.3.3   Other Possibilities

It is desirable if the SAW oscillation is visualized as a whole field. Fig. 14.6 illustrates a possible scheme to this end. The idea is similar to that of [2]; generate a standing surface acoustic wave on the specimen surface and read out the resultant surface displacement with an optical interferometer. The difference from [2] is that the optical interferometer monitors the surface displacement as a whole-field image. The laser beam illuminating the specimen surface forms speckles on the image. By subtracting an image taken with the transducer on from an image taken with the transducer off, the displacement can be characterized as an interferometric fringe pattern. Dark fringes should be formed on the specimen where the acoustic standing wave has nodes. It is expected that abnormality on the elastic property of the specimen material change the pattern of standing wave and the change can be analyzed via fringe pattern analysis.

14.4 Summary

The idea of simultaneous application of acoustic and optical techniques to nondestructive evaluation is discussed. Our recent experiments based on the idea are presented. Results from these experiments indicate that some observations are based on dynamics that cannot be observed by the acoustic or optical technique alone.

References

1. Stegemean GI (1976) Optical probing of surfaced waves and surface wave devices. IEEE Trans Son Ultrason SU-23(1):33–63
2. Burov JI, Thanh NC, Anatasova NV (1979) Optical method for measuring the amplitude response of surface acoustic wave devices. Appl Phys 20:51–54
3. Sciammarella CA, Lamberti L, Sciammarella FM, Demelio G, Dicuonzo A, Boccaccio A (2010) Application of plasmons to the determination of surface profile and contact stress distribution. Strain 46(4):307–323
4. Atalar A, Quate CF, Wickramasinghe HK (1977) Phase imaging in reflection with the acoustic microscope. Appl Phys Lett 31:791–793
5. Ono Y, Kushibiki J-i (2000) Experimental study of construction mechanism of V (z) curves obtained by line-focus beam acoustic microscopy. IEEE Trans Ultrason Ferroelectr Freq Control 47(4):1042–1050
6. Butters JN, Leendertz JA (1971) Speckle pattern and holographic techniques in engineering metrology. Opt Laser Technol 3:26–30
7. Hughes DS, Kelly JL (1953) Second-order elastic deformation of solids. Phys Rev 92:1145–1149
8. Qu J, Liu G (1998) Effects of residual stress on guided waves in layered media. Rev Prog Quant Nondestruct Eval 17:1635–1642
9. Duquennoy M, Ouaftouh M, Ourak M, Jenot F, Xu WJ (2002) Theoretical and experimental determination of the natural and initial acoustoelastic coefficients. AIP conference proceedings; annual review; 14th, review of progress in quantitative nondestructive evaluation, pp 1696–1702
10. Duquennoy M, Ouaftouh M, Ourak M, Jenot F (2002) Theoretical determination of Rayleigh wave acoustoelastic coefficients: comparison with experimental values. Ultrasonics 39(8):575–583
11. Dorfi HR, Busby HR, Janssen M (1996) Acoustoelasticity: ultrasonic stress field reconstruction. Exp Mech 36:325–332
12. Kima JN, Tutwiler R, Kwak DR, Park I, Miyasaka C (2013) Multilayer transfer matrix characterization of complex materials with scanning acoustic microscopy. In: Proceedings of SPIE 8694, nondestructive characterization for composite materials, aerospace engineering, civil infrastructure, and homeland security 2013, 869410
13. Yoshida S, Adhikari S, Gomi K, Shrestha R, Huggett D, Miyasaka C, Park I (2012) Opto-acoustic technique to evaluate adhesion strength of thin-film systems. AIP Adv 2(2):022126-1–022126-7

Chapter 15
Stress Analysis on Welded Specimen with Multiple Methods

Sanichiro Yoshida, Tomohiro Sasaki, Sean Craft, Masaru Usui, Jeremy Haase, Tyler Becker, and Ik-Keun Park

Abstract Effects of the thermal stress due to welding on the mechanical property of the welded work are analyzed with multiple methods. Strain gauges and an acoustic microscope are used to measure the residual stress, and electronic speckle pattern interferometry (ESPI) is used to analyze the response of the welded wok to external force. A tensile load is applied to butt-welded, thin-plate steel specimens, and the resultant strain field is analyzed with the ESPI. Comparison is made with the case of a non-welded specimen of the same material and dimensions. The analyses indicate that the residual stress due to welding makes the normal strain due to the external tensile load asymmetric. The asymmetry enhances shear and rotational modes of deformation, generating stress concentration at a point away from the weld where the residual stress is substantially negligible.

Keywords Residual stress • Electronic speckle-pattern interferometry • Acoustoelasticity • Welding • Strain analysis

15.1 Introduction

The stress field near a weld is complicated. Thermal load during the weld generates a permanent stress locked in the material. This type of stress is known as the residual stress and is harmful to the welded work as it accelerates deformation and even leads to a failure of the work. The mechanism of residual stress is generally explained as follows. During the heating phase, thermally expanded regions stretch the surrounding material. Consequently, the surrounding regions experience tensile stress. During the cooling phase, the previously expanded regions get stretched as they shrink whereas the surrounding regions get compressively stressed (because the surrounding area does not shrink). Thus, the welded work tends to have tensile residual stresses near the weld and compressive residual stresses in the other regions. However, in reality, the pattern of residuals stress is not as simple as this. It depends on various factors such as the shape of the work, the thermal conductivity of the material and the welding conditions.

A number of theoretical and experimental studies are reported on residual stresses due to welding [1–3]. These studies mostly focus on the analysis, removal or prevention of the residual stresses. As a stress hidden in the material, detection of residual stress is quite challenging and these types of studies are extremely important. What is also important, however, is to understand the response of materials with residual stresses when an external force is applied. It is most likely that the welded work is used under the influence of a certain external force; as an example, the stress due to the inner pressure on the welded-wall of a high-pressure vessel. It should be noted that simple superposition of residual stresses and the stress due to an external load at the point of residually stressed area is not sufficient to understand the behavior of the welded work. Normally, the action point of the external load is distanced from the point of residual stress. When the external force acts on

S. Yoshida (✉) • S. Craft • J. Haase • T. Becker
Department of Chemistry and Physics, Southeastern Louisiana University, SLU 10878, Hammond, LA 70402, USA
e-mail: syoshida@selu.edu; sanichiro.yoshida@selu.edu

T. Sasaki • M. Usui
Department of Mechanical Engineering, Niigata University, IkarashinNinocho 8050, Nishi-ku, Niigata-shi, Niigata 950-2181, Japan

I.-K. Park
Department of Mechanical and Aotomotive Engineering, Seoul National University of Science and Technology,
232 Gongneung-ro, Nowon-gu, Seoul 139-743, South Korea

the work, the force is transferred from the action point to the entire work in accordance with the elastic property of each point. Residual stresses alter the local elastic constant of the material; as widely known, a compressive stress due to the thermal load from a welding torch can easily cause permanent plastic strain. Therefore, the response to the external force of the entire work containing residual stresses is substantially different from that without residual stresses.

The aim of this study is to explore this exact question of "how do the residual stresses alter the response of welded work to external force?" To this end, a low-level tensile load is applied to a butt-welded thin-plate specimen and its response is compared with the response of a non-welded specimen of the same dimension and material. An electronic speckle pattern interferometer (ESPI) is used to analyze the specimen's behavior as a full-field image. A strain gauges and a scanning acoustic microscope are used to analyze the residual stress. The results indicate that the change in the elastic constant due to the residual stress near the weld alters the pattern of overall deformation, making it less symmetric around the central line of the specimen (both parallel and perpendicular to the tensile axis). This generates strain concentration at a point away from the weld line, causing more bodily rotation in response to the applied tensile load.

15.2 Experimental

The materials tested in this study were cold-rolled carbon steel plates. Two types of welded specimens were prepared. For the first type (specimen 1), two plates of 90 mm (wide) × 50 mm (long) × 0.4 mm (thick) were butt-welded into a plate of 90 mm × 100 mm × 0.4 mm, and cut into 20 mm wide × 100 mm long specimens (Fig. 15.1a). The second type (specimen 2) was prepared in a similar fashion but the final dimension was 40 mm wide × 150 mm long × 3 mm thick (Fig. 15.1b). For each type, a non-welded specimen of the same dimension was prepared as the control.

The specimen was attached to a tensile machine for application of a tensile load. An ESPI setup sensitive to in-plane displacement was arranged in front of the tensile machine (Fig. 15.2). Fringe patterns were formed by subtracting the image taken before the application of the tensile load from the image taken after the application of the load. The phase associated with the displacement was evaluated by assigning the order to each fringe and interpolating the phase between fringes. A care was taken to apply as small load as possible to avoid plastic deformation caused by the tensile load. For comparison, a non-welded specimen of the same material and dimension was tested in the same fashion.

A challenge in the application of the ESPI as stated above to analysis of residual stresses due to welding is that the phase analysis on the formed fringes is not easy. Often the welded specimen exhibits bodily rotation, which yields fringes parallel to the sensitivity vector of the ESPI interferometer. When the interferometer is sensitive to horizontal displacement, for example, the fringe pattern is approximately equidistant, horizontally parallel, as Fig. 15.3a indicates. This makes difficult to evaluate the normal strain, because the variation of the fringe order along the axis is small (in the case of Fig. 15.3a, the change in the fringe order in going horizontally from the left end of the image to the right end is two or less). As a solution to this issue, a system of carrier fringes was introduced by rotating the wedge placed after the beam expander for the left interferometric arm. The wedge was rotated after the tensile load was applied. The rotation of the wedge provided a constant phase variation along the tensile axis on the specimen as Fig. 15.2b indicates. The provision of the carrier fringe altered the fringe pattern from Fig. 15.3a to Fig. 15.3c. Now that the horizontal variation in the fringe order was high, the phase could be

Fig. 15.1 Butt-welded specimen 1 (**a**) and specimen 2 (**b**). The *box with dashed lines* indicates approximate area of view. For specimen 2 strain was measured with strain gauges at 12 points; 5 mm and 20 mm away from the weld line along the tensile axis, 16 mm above/below the central line parallel to the tensile axis. *Circles* indicate these locations for the *right side* of the weld line only. Strain gauge measurements were made on the *left side* of the weld line at the locations symmetric to these six points. The circle and distances inserted in (**a**) indicate the locations where acoustic velocity is previously measured in a welded specimen similar to this study

Fig. 15.2 Experimental arrangement. The optical wedge is to introduce carrier fringes
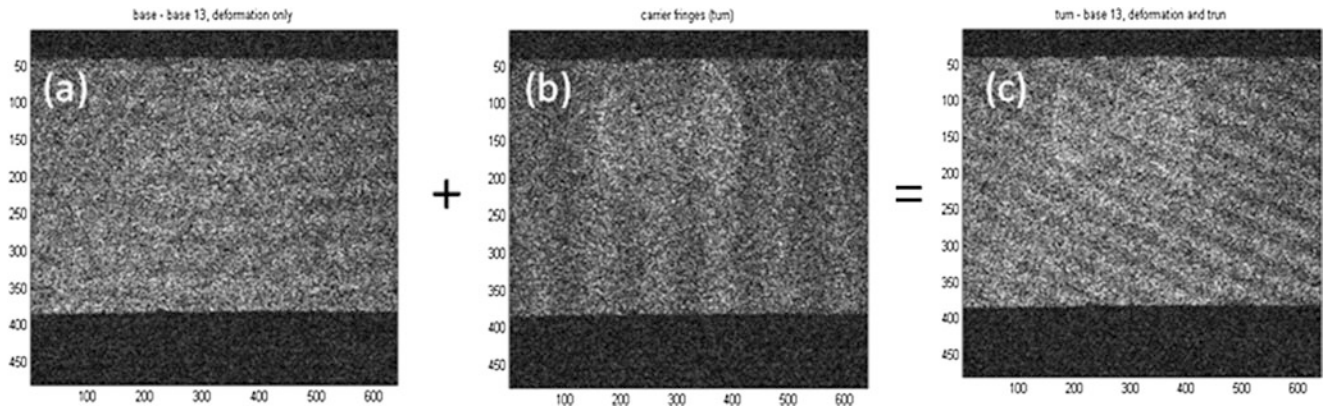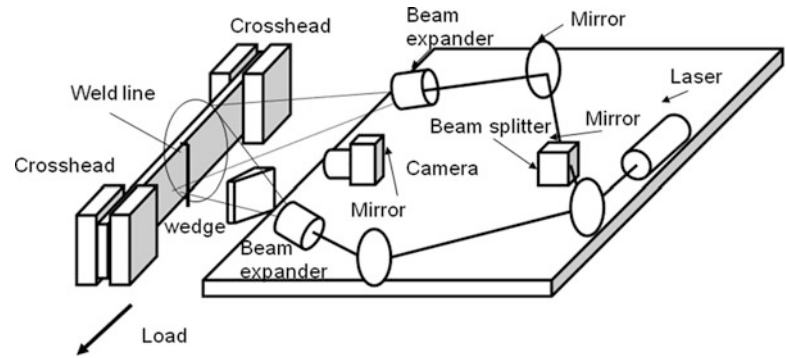
Fig. 15.3 Typical fringe patterns. (**a**) deformation only; (**b**) carrier fringes; (**c**) deformation plus carrier fringes

evaluated with a reasonable accuracy by interpolation. The true phase change was evaluated by subtracting the known phase change due to the wedge rotation (Fig. 15.3b) from the total phase variation (Fig. 15.3c).

For specimen 2, residual strain was measured with a strain gauge. After the tensile experiment was completed, the specimen was detached from the tensile machine, and a pair of strain gauges oriented along the tensile axis and perpendicular to it was attached at 12 reference points near the weld as indicated in Fig. 15.1b. After strain was measured at each point for both x and y directions, these points were isolated by cutting as a square of approximately 9 mm $\times$ 9 mm around the point. After the isolation, the strains were measured again. The difference of the strain gauge reading before and after the isolation was interpreted as the residual strain. Each measurement was repeated five times and the average of the five measurements was employed.

15.3   Results and Discussion

Table 15.1 summarizes the tensile experiments conducted for the two types of specimens. The same load was applied to the non-welded and welded specimens. "Average strain" is calculated from the displacement of the tensile machine's crosshead and the gauge length of the specimen (the former divided by the latter). The average stress is calculated from the tensile machine's load reading (the difference between the load cell reading when the fist ESPI image is taken and that when the second ESPI image is taken). The maximum stress (x) and (y) are evaluated from the maximum values of the strain gauge's reading at the 12 measurement points (Fig. 15.1b).

Figure 15.4 are normal strain (ε_{yy}) along the tensile axis (y) obtained from the fringe analysis on specimen 1 for the non-welded (left) and welded (right) case. The box drawn near the center of the y-axis indicates the location of the weld. The strain levels are close to the average strains shown in Table 15.1.

Comparison of the non-welded and welded cases in Fig. 15.4 indicates the following features. Generally, the strain of the non-welded specimen is much smoother than that of the welded specimen. The strain field of the welded specimen appears more or less symmetric along the y-axis (parallel to the tensile axis) about the weld line. It appears that the normal strain vigorously fluctuates along the x-axis (perpendicular to the tensile axis) than y-axis over the all region, not necessarily near

Table 15.1 Summary of tensile experiment on Specimen 1 and 2

	Specimen 1		Specimen 2			
	Average strain (cross head reading)	Average stress (load cell) (MPa)	Average strain (cross head reading)	Average stress (load cell) (MPa)	Max stress (y) (strain gauge) (MPa)	Max stress (x) (strain gauge) (MPa)
Non-welded	1.00×10^{-4}	1.25	6.29×10^{-4}	15.9	–	–
Welded	2.00×10^{-4}	1.25	6.29×10^{-4}	17.2	168.7	−224.96
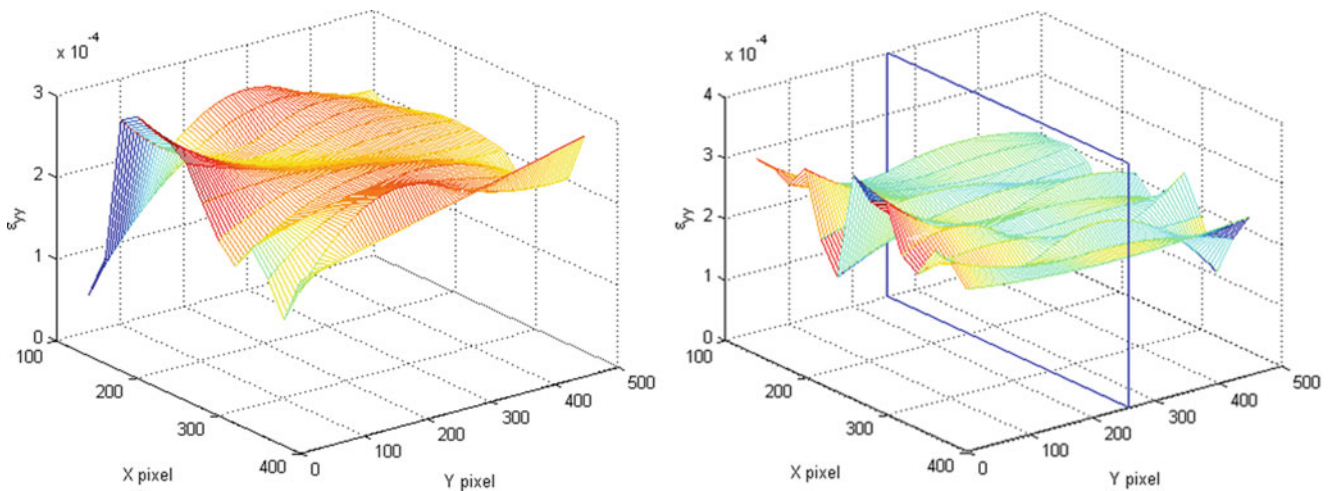
Fig. 15.4 Normal strain observed in non-welded (*left*) and welded (*right*) specimen 1

the weld line. This is explained by considering that the residual stress due to welding is mainly caused by the cooling and heating effects of the welding laser beam as it sweeps along the weld line. It is consistent with Karlsson's observation [3] that the residual stress due to weld is greater parallel to the weld than perpendicular to it. The high spatial frequency in the fluctuation of the strain along the x-axis indicates that the heating and cooling effect associated with the passage of the welding laser beam is rather fast. This spatial frequency may depend on the speed of the welding, which is an interesting subject for future study. The strain fluctuations along the x- and y-axis, i.e. parallel and perpendicular to the weld line, are more explicitly analyzed in Figs. 15.5 and 15.6 where the strain of the welded specimen is compared with that of the non-welded specimen for several lines perpendicular and parallel to the weld line.

It is well known that the elastic constant of materials changes under stress. The reason for the change is that a stress alters the inter-atomic distance and that the inter-atomic potential is not symmetric except the vicinity of the equilibrium. In most materials, the inter-atomic potential is steeper on the short-distance side of the equilibrium than the long-distance side, whereas it is approximately a quadratic function of the inter-atomic distance near the equilibrium. Acoustoelasticity utilizes this fact [4–6]. When an elastic medium is under a compressive stress, the inter-atomic distance is decreased and therefore the elastic constant, the change in the slope of the inter-atomic potential, becomes higher than around the equilibrium. On the contrary, when an elastic medium is under a tensile stress, the elastic constant becomes lower than the equilibrium. Consequently, the acoustic phase-velocity, which is proportional to the square root of the elastic constant, increases/decreases under a compressive/tensile stress. A previous study (C. Miyasaka, private communication, 2013) shows that the acoustic velocity varies ±6 % along a line parallel to the weld line in a SPCC specimen similar to this study, as shown in Fig. 15.7. Here the horizontal axis indicates the distance from the center of the specimen along a line parallel to the weld line and 3 mm away from it. The same study indicates that the acoustic velocity measured at 7 mm away from the weld line is not affected by the welding. The points of measurement are indicated in Fig. 15.1a.

Although the non-welded and welded data are based on two independent ESPI measurements, the applied tensile load is identical. Therefore, the strain data of the non-welded and welded specimen shown in Fig. 15.4 are proportional to the compliance, and a comparison between the two cases should indicate the relative compliance reflecting the effect of welding. Figure 15.8 plots the strain data of the welded specimen divided by that of the non-welded specimen (the relative compliance plot). Also plotted in this figure is the maximum and minimum relative compliance calculated from the 6 % variation of the acoustic velocity data shown in Fig. 15.7. It is seen that the relative compliance plot fits between the maximum and minimum
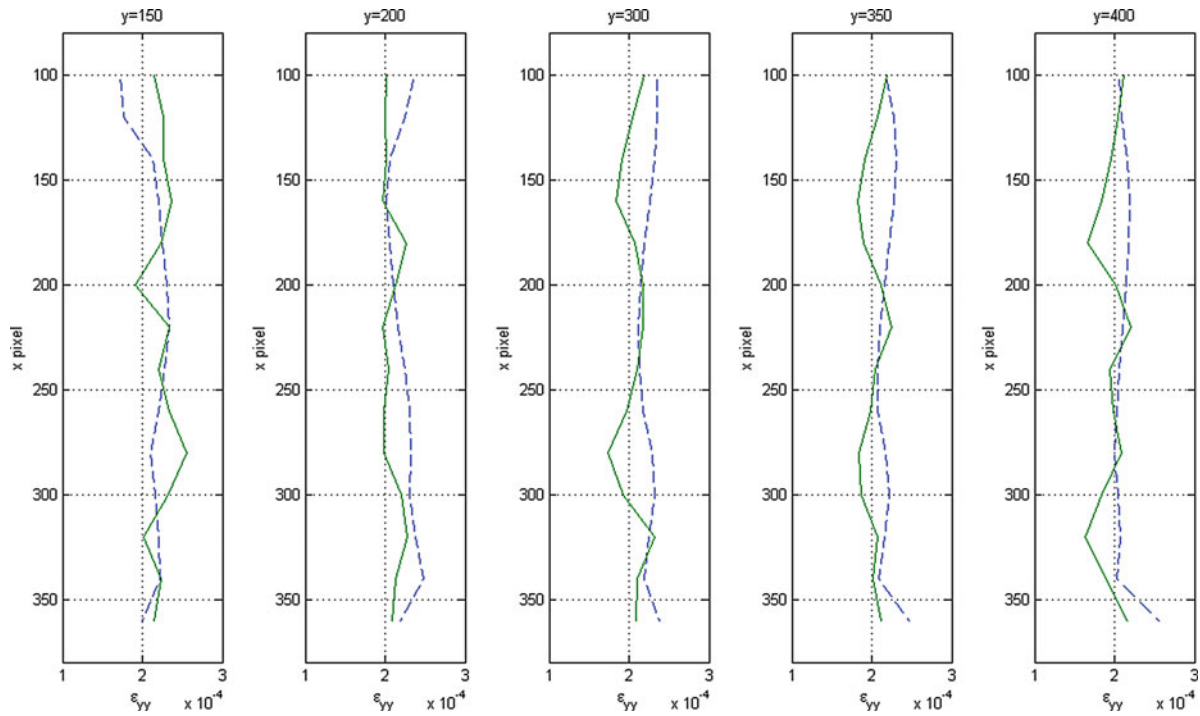
Fig. 15.5 Dependence of ε_{yy} on x (parallel to weld line). The solid line is for the welded specimen and the dashed line is for the non-welded specimen

Fig. 15.6 Dependence of ε_{yy} on y (perpendicular to weld line). The solid line is for the welded specimen and the dashed line is for the non-welded specimen
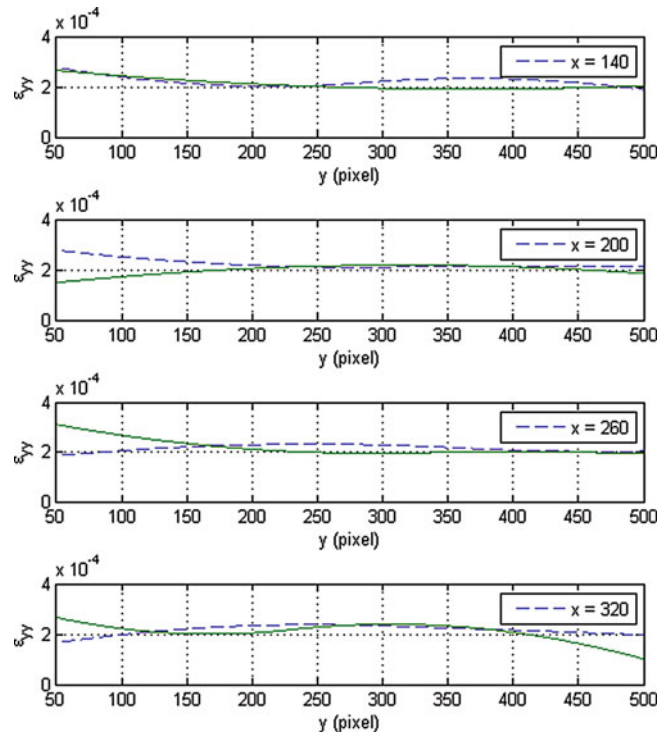
Fig. 15.7 Acoustic velocity measured near weld
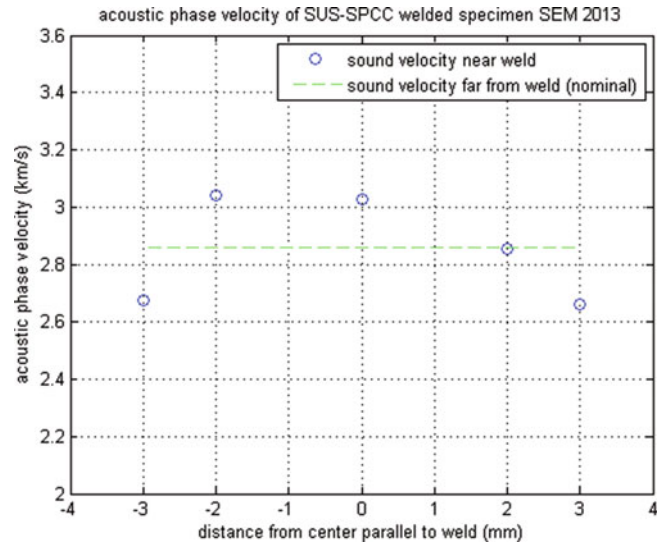
Fig. 15.8 Relative compliance (welded/ non-welded)
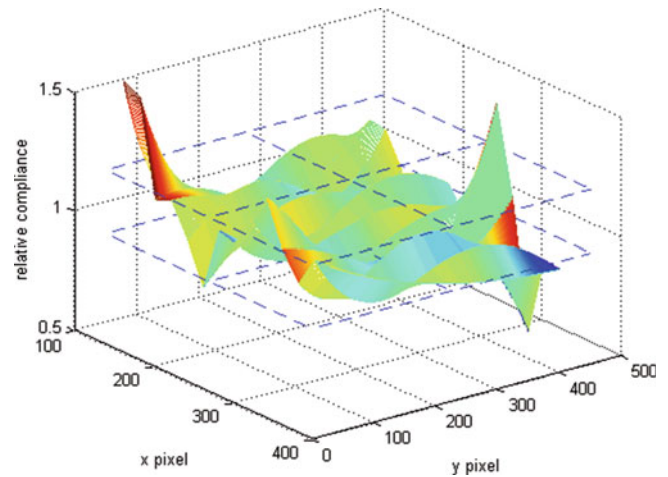
relative compliance based on the acoustic velocity measurement, indicating the consistency of the ESPI measurement and acoustoelastic measurement.

Usually, the change in the elastic constant is ascribed to the non-linear elasticity of the inter-atomic potential. The nonlinearity is expressed via the third-order term of the potential [5, 6], where the first order term represents a reference potential and the second-order term the Young's modulus. The Young's modulus and the third order term of steel are typically 200 GPa and -500 GPa, respectively [6].

$$\frac{d\sigma}{d\varepsilon} = C^{(2)} + C^{(3)}\varepsilon = 200 - 500\varepsilon \tag{15.1}$$

Here σ is the stress, $C^{(2)}$ is the Young's modulus, $C^{(3)}$ is the third order term of the elastic constant and ε is the strain. The relative elastic constant (the reciprocal of the relative compliance) can be expressed as

$$\left(\frac{d\sigma}{d\varepsilon}\right) \bigg/ \left(\frac{d\sigma}{d\varepsilon}\right)_0 = (200 - 500\varepsilon)/200 = 1 - \frac{5}{2}\varepsilon \tag{15.2}$$

If the observed change in the phase velocity of 6 % is due to the third order term of elastic constant, the value of Eq. (15.2) must be in the range of 1.06^2–0.94^2. From Eq. (15.2), the corresponding strain is -0.0494 (compressive) to 0.0464 (tensile). Since the yield strain of steel is typically 0.2 % $= 0.002$ or lower, these numbers are much higher to explain the observation by the nonlinear elasticity. Thus, the observed change in the compliance must be attributed to plastic deformation.
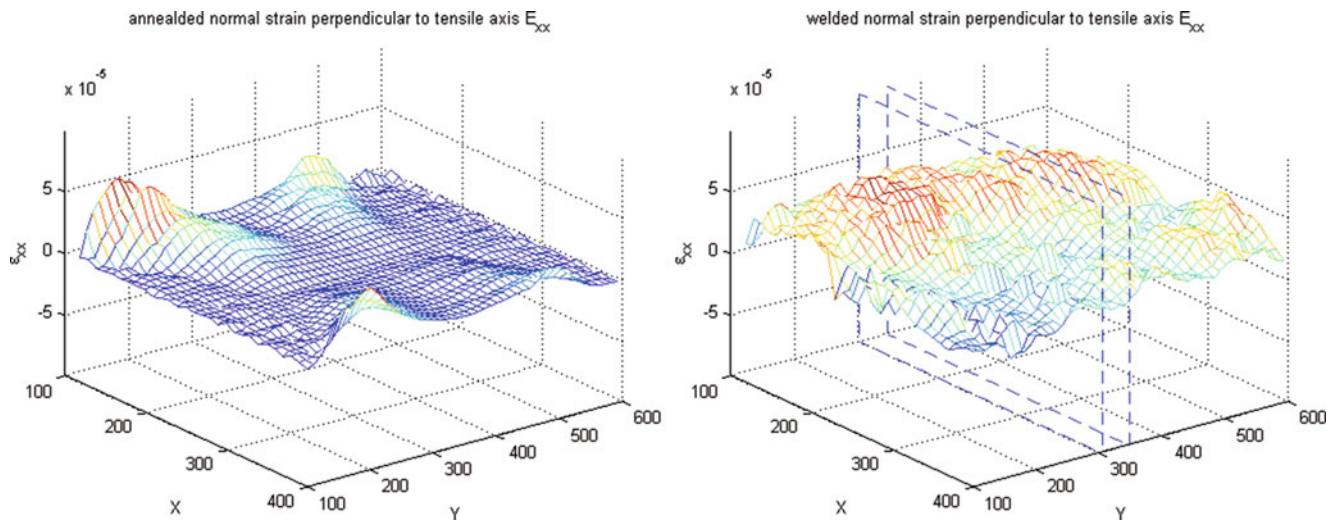
Fig. 15.9 Strain observed in non-welded (*left*) and welded (*right*) specimen 2

It should be noted that these strain values (-0.0494 and 0.0464) are two orders of magnitude higher than the corresponding value of the maximum residual stress measured with the strain gauge; 168.7 MPa/200 GPa $= 8.4 \times 10^{-4}$ (See Table 15.1. Here the residual strain is evaluated as the ratio of the measured residual stress 168.7 MPa to the Young's modulus 200 GPa). This is understandable because the strain gauge is sensitive only to the elastic portion of the residual strain while the acoustic wave velocity is determined by the overall elasticity of the region that can be partially plastically deformed.

Figure 15.9 shows strain plots for the non-welded and welded cases of specimen 2. For specimen 2, the ESPI setup was sensitive to both x (perpendicular to the tensile axis) and y (parallel to the tensile axis) components of the in-plane displacement. In this case, the effect of the welding is most prominent in the normal strain ε_{xx}. Like the case of specimen 1, the strain observed in the welded specimen shows vigorous fluctuations.

Figure 15.10 compares the y-dependence of ε_{xx} and ε_{yy} measured by the strain gauge and ε_{xx}, ε_{yy}, ε_{xy} (shear strain) and ω_z (rotation). Here dashed lines are the non-welded cases and solid lines are the welded cases. The strain gauge data show much higher correlation with shear and rotation than the normal strains. This supports the proposition that the effect of residual stress is not a simple superposition of the residual stress and the external stress but instead more overall response of the material.

The overall response of the specimen as a whole to the external load can be seen more clearly in Fig. 15.11 where contour plots of all strains and rotation are compared between the non-welded cases (left column) and the welded cases (right column). The normal strains (the first two rows from the top) indicate that the welding causes the deformation due to the tensile load to be less symmetric. Under tensile loading of a thin plate with no transverse constraint, the deformation is essentially symmetric. In a specimen of isotropic material with no initial stress concentration, the normal strain parallel to the tensile axis should be uniform, and the normal strain perpendicular to the tensile axis should be symmetric about the central line on the specimen parallel to the tensile axis. The top two contour plots on the left column certainly indicate these features for the non-welded case. [The ε_{yy} observed near the two horizontal ends (corresponding to the two ends gripped by the tensile machine) appears lower than the horizontal middle section. This is considered to be due to the lateral constraint by the tensile machine's grip.] On the other hand, the normal strains in the welded specimen do not show these feature. The ε_{xx} plot does not show the symmetry about the central line parallel to the tensile axis; instead, it exhibits some degree of symmetry about the diagonal line running from the top left corner to the bottom right corner of the specimen. The ε_{yy} plots does not indicate the uniform strain in the middle section but exhibits strain concentration near the bottom of the specimen about halfway to the right from the weld line. The contour plots of the shear strain and rotation are consistent with the above observations. The shear strain contours in the welded specimen shows a high value near the above-mentioned strain concentration. The rotation contours in the welded specimen exhibits much higher clockwise rotation in the middle section of the specimen. The displacement field plots shown in Fig. 15.12 are consistent with the observation. The displacement vectors of the welded case are similar to the non-welded case going from the right end through the middle of the right side of the weld. However, starting around this region toward the left end of the specimen, the displacement vectors of the welded specimen clearly indicate the tendency that the material rotates clockwise as a whole.
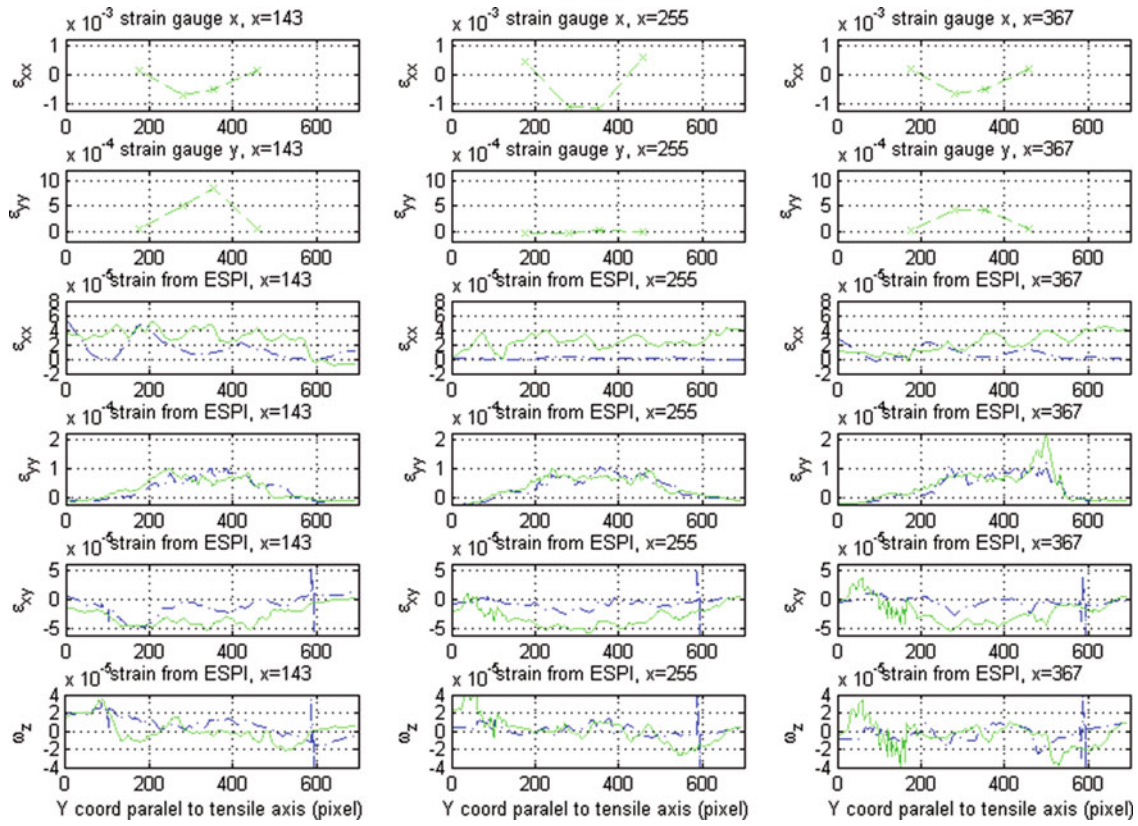
Fig. 15.10 Comparison of strain measured with the strain gauge and strain and rotation measured by ESPI

Fig. 15.11 Contour plots of strain and rotation
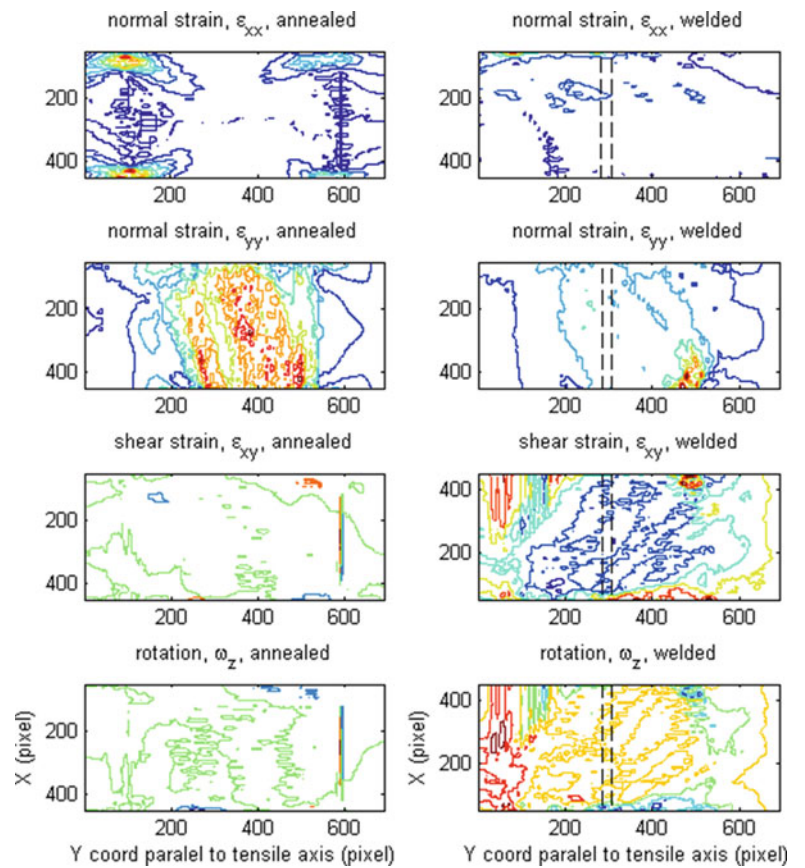
Fig. 15.12 Displacment vector field observed in the non-welded specimen (*top*) and in the welded specimen (*bottom*)
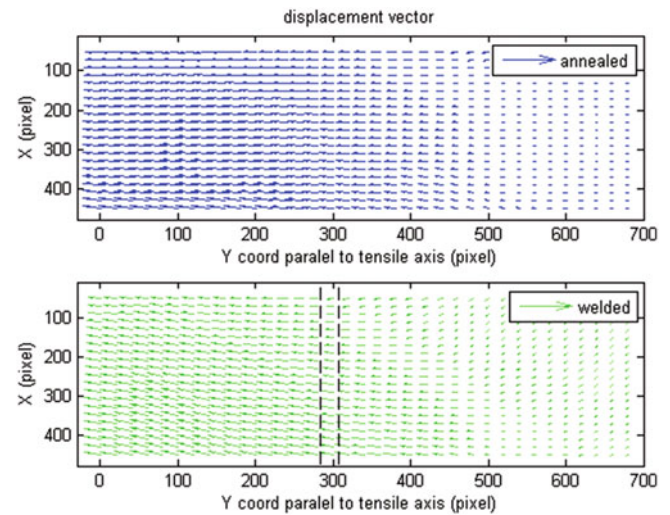
From the viewpoint of the residual stress, the above observations can be explained as follows. In Fig. 15.10, the strain gauge data indicates high compressive ε_{xx} near the weld ($y \approx 300$ pixel), and very low ε_{yy} near the weld around the vertical center (middle column). The high compressive ε_{xx} near the weld indicates that the stiffness parallel to the weld line is increased in this region. This hinders the specimen from experiencing the natural deformation, when the external load is applied, that the central portion contracts due to the Poisson's effect and the tensile deformation along the tensile axis is uniform. In accordance with the Poisson's effect, the material must somehow have compressive strain perpendicular to the tensile axis. The easiest location for the specimen to have compressive ε_{xx} strain is where the residual ε_{xx} turns the sign from negative (compressive) to positive (tensile). In the present case, this happens to be near $y \approx 200$ and $y \approx 420$ pixel (see top row of Fig. 15.10). Since between these two locations in y the material has difficulty in contraction, the stress is concentrated around $y \approx 200$ and $y \approx 420$ pixel. This explains the high shear strain between $y \approx 200$ and $y \approx 420$ pixel (third row, right column in Fig. 15.11 contour plots). Interestingly, the rotation changes the sign near $y \approx 100$ (bottom right plots in Fig. 15.10). The opposite rotations on the opposite sides of $y \approx 100$ indicate that the total angular momentum must remain zero, as the tensile machine does not exert torque. Note that the strain gauge and acoustic microscopic measurement indicate that the residual stress is substantially negligible at the point of the stress concentration.

15.4   Conclusions

In summary, this analysis indicates that the effect of residual stress due to welding is not necessarily limited near the weld where the residual stress itself is high. The change in the constitutive property of the material due to the residual stress alters the response of the material as a whole when an external load is applied. The findings resulting from this study can be summarized as follows:

1. The thermal load due to the welding has greater effect parallel to the weld than perpendicular as reported in ref 3.
2. The high compressive residual stress near the weld hinders natural compressive deformation due to the Poisson effect when an external tensile is applied.
3. The above effect can cause a stress concentration at a location some distance away from the weld where the residual stress is negligible.

References

1. Bate SK, Green D, Buttle D (1997) A review of residual stress distributions in welded joints for the defect assessment of offshore structures. Health and Safety Executive—offshore technology report
2. Colegrove P, Ikeagu C, Thistlethwaite A, Williams S, Nagy T, Suder W, Steuwer A, Pirling T (2009) The welding process impact on residual stress and distortion. Sci Technol Weld Join 14(8):717–725
3. Karlsson L (2005) Residual stresses due to welding of a nozzle to a pressure vessel. Master's dissertation, Division of Solid Mechanics, Lund University, Sweden
4. Every AG, Sachse W (1990) Determination of the elastic constants of anisotropic solids from acoustic-wave group-velocity measurements. Phys Rev B Condens Matter 42:8196–8205
5. Muir DD (2009) One-sided ultrasonic determination of third order elastic constants using angle-beam acoustoelasticity measurements. Ph.D. thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology
6. Stobbe DM (2005) Acoustoelasticity in 7075-T651 aluminum and dependence of third order elastic constants on fatigue damage. Master's thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology

Chapter 16
Sound Attenuation for Dogs Barking Using of Transfer Function Method

Shuichi Sakamoto, Takatsune Narumi, Yuichi Toyoshima, Nobuaki Murayama, Toru Miyairi, and Akira Hoshino

Abstract This paper reports a natural ventilation silencer for an animal cage having considerable sound insertion loss. Many veterinary hospitals face a serious problem that hospitalized animals are stressed from the barks of other dogs in the hospital. To address this problem, a solution that barely affects caged dogs' health is required. We attempted to use a silencer with sound insertion loss by attaching it to an opening side of the cage. We constructed the tested prototype silencers by using the resonance of a rectangular room and a nonwoven fabric. We measured the acoustic characteristics of the silencers that were attached to the real scale model of the animal cage. We also measured their insertion loss by considering the estimated frequency characteristics of the dogs' barking sound. The result shows that proposed silencers have more than 10 dB insertion loss; therefore, they can maintain respiratory environment of a caged dogs along with good visibility.

Keywords Noise control • Sound insertion loss • Resonance of rectangular room • Cell-type silencer • Nonwoven fabric

16.1   Introduction

This study is related to the use of natural ventilation silencers for the howling and barking (hereafter referred to as "barking") of dogs. Pets play an important role in advanced nations because the nuclear families and low birth are getting spread there. The number of complaints and problems caused by the noise created by barking dogs is increasing in Japan; it represents the major component of noises in living spaces [1]. Therefore, it is necessary to sort some countermeasure. Furthermore, dogs under treatment in the hospital are housed in connecting cages; barking of one dog can cause others to bark as well, creating stress in the other animals in the hospital. To remedy this situation, one method has being considered: attaching a sound insulating board to the opening of the cages and utilizing the forced ventilation. However, using of sound boards and forced ventilation is imperfect way because this method creates a number of issues, including problems such as hindrance in communicating with animals, noise associated with ventilation intake and output, noise from fans within cages, cost, energy consumption, and the risks of malfunction and power outages. Collectively, these problems make this solution unfeasible.

In this paper, we report the evaluation results of sound attenuation of silencer for animal-cage for using dog's barking model. We focus on cages with an opening on one side, which are commonly used in veterinary hospitals. This method can also be applied to indoor and outdoor animal cages in general households. According to opinions provided from veterinary hospitals, following conditions are required for a silencer: Effective silencing, no burden on the breathing of the animal, two-sided visibility, a thickness that will not inhibit veterinary treatment, and lightweight in order to avoid adding weight to the cage door and burdening workers. To satisfy these conditions, we created a prototype silencer based on resonance within a rectangular chamber divided into cells, adding non-woven sheets to the interior, tail pipes, and coaxial side branch tube silencers to the open end. We then evaluated the sound attenuation performance.

S. Sakamoto (✉) • T. Narumi
Faculty of Engineering, Department of Mechanical and Production Engineering, Niigata University,
Ikarashi-2-no-cho 8050, Nishi-ku, Niigata 950-2181, Japan
e-mail: sakamoto@eng.niigata-u.ac.jp

Y. Toyoshima • N. Murayama • T. Miyairi • A. Hoshino
Graduate School of Science and Technology, Niigata University, Ikarashi-2-no-cho 8050, Nishi-ku, Niigata 950-2181, Japan

16.2 Experimental Method

16.2.1 Experimental Equipment

The structure of the measurement apparatus is shown in Fig. 16.1. Simulated sound generated by Computer 1 is emitted through a speaker placed in a real-size model of a small dog cage. The sound wave was measured by using a microphone placed outside of the cage. Then, it was recorded and analyzed by using Computer 2.

The inner dimensions of the cage were 570 mm W × 570 mm H for the opening, with a depth of 700 mm. On the basis of our observations of barking dogs in a veterinary hospital, we placed the speaker in the center of the cage opening at a depth of 200 mm. The microphone was placed above the central axis of the cage at a position 1 m away from the opening on the outside. Thus, the distance from the front of the speaker to the microphone was 1.2 m. The experiment was conducted in a semi-anechoic room.

16.2.2 Frequency Characteristics and Sound Level of Dog Barking

According to survey results related to dog barking [2], the barks of all breeds have high-level frequency components from 1,000 to 2,000 Hz, and sound levels were in the range of approximately 80–90 dB at a distance of 5 m from the dog.

In this experiment, we set the distance from the speaker sound source to microphone is 1.2 m, and recording measurements were performed in this condition. Because the distance changed to approximately one-fourth of 5 m (5 m to 1.2 m), if we consider that the attenuation with distance for a point sound source is −6 dB/dd, the sound level is predicted to increase to 12 dB. Thus, at a distance of 1.2 m, we estimated the sound level of the dog to be 100 dB.

Figure 16.2 shows the dog barking model referencing the frequency characteristics of dogs barking. The dog barking model adds frequency weight to flat noise (pink noise) in the graph where the frequency axis is logarithmic. For the added frequency weight, at the frequencies less than 1 kHz and greater than 2 kHz, the attenuation characteristic was assumed to be approximately 12 dB/oct. The sound level (overall value) was set to 100 dB.
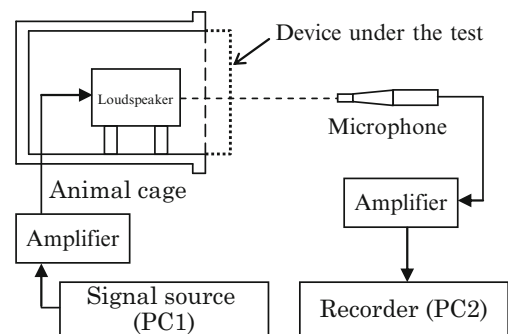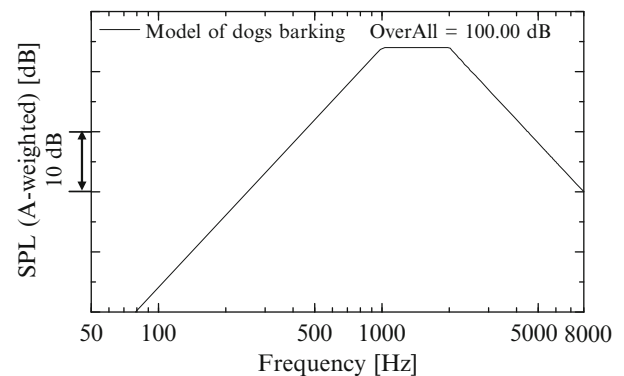

Fig. 16.1 Experimental setup

Fig. 16.2 Model of dogs barking (12 dB/oct BPF, logarithmic scale)

Fig. 16.3 Model of dogs
barking and measured
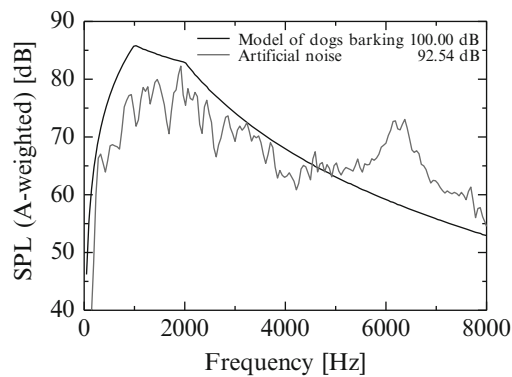artificial noise (linear scale,
24 kHz/512 line)

Figure 16.3 shows the dog barking model, shown in logarithmic scale on the frequency axis in Fig. 16.2, converted to a linear scale. According to this result, the dog barking model included the 3 dB/oct decreasing frequency characteristic of pink noise. Figure 16.3 includes frequencies up to 8 kHz; however, in linear frequency analysis with 24 kHz/512 line resolution, as was the case with the measurement apparatus used, we set the level values for each frequency such that the overall value of the dog barking model was 100 dB.

Because the dog bark is difficult to reproduce, we used simulated sound output from the speaker. The results for simulated sound output from the speaker and measured with the microphone are shown in Fig. 16.3. These results show that the speaker frequency characteristics are in superposition with the simulated sound measured. Using the difference between the frequency characteristics of the simulated sound and the dog barking model shown in Fig. 16.3, we equalized the measurement results acquired after the silencer was attached. This approach allowed us to obtain measurement results equivalent to the dog barking model, simulated sound. Hereafter, we refer to this sound attenuation of the dog barking model simply as "sound attenuation."

16.3   Silencer Prototype and Experiment

16.3.1   Basic Type of the Silencer (Resonance in Rectangular Chambers with Cylindrical Openings)

The silencer with rectangular chambers arranged in an array configuration with openings on both sides is shown in Fig. 16.4. Permitting communication like touch and provide food to animals in the cage, the diameter of the opening is set to a size through which a human hand can pass to allow.

When the aperture ratio is set to 25 % with a 4×4 arrangement of 16 silencers and thickness of material is 4 mm, an internal length of one square section of the cell ($l_y = l_z$) is 137.5 mm and the diameter of the opening is 80 mm. This diameter is enough large to accommodate all, but the largest Japanese hands.

The results of experiments conducted with this basic-type silencer are shown in Fig. 16.5. In this figure, sound attenuation of approximately 4 dB is obtained. Because the internal length of the diagonal line observed from in front of the cell is 194 mm, plane wave in the thickness direction of the silencer is formed up to approximately 1 kHz. Thus, for frequencies up to approximately 1 kHz, we can obtain expansion silencing effect by suddenly shrinking or expanding the cross-sectional area in front and behind the opening. The attenuating peak frequency of expansion silencer is known as the function of length of silencer ($l_x = 92$ mm). In this case, the first attenuating peak is calculated approximately 920 Hz, and we can also see attenuation around 900 Hz in Fig. 16.7. In addition, multiple rectangular chamber resonance modes probably occur at frequencies higher than 1 kHz. Each first Resonant frequencies of direction of l_y and l_z mode and diagonal line mode of l_y and l_z are calculated 1,250 Hz and 1,770 Hz, respectively. We can also see a sound attenuation effect due to energy dispersion caused by these resonances in Fig. 16.5.

16.3.2   Attachment of Sheets of Non-woven Fabric

We conducted experiments with a silencer that linked the openings of both ends of the basic silencer discussed in the previous section with cylindrical non-woven fabric sheets. Figure 16.5 shows these measurement results. Through the back

Fig. 16.4 Outlook of the rectangular room silencer (basic-type silencer with double openings, $l_x = 92$ mm)
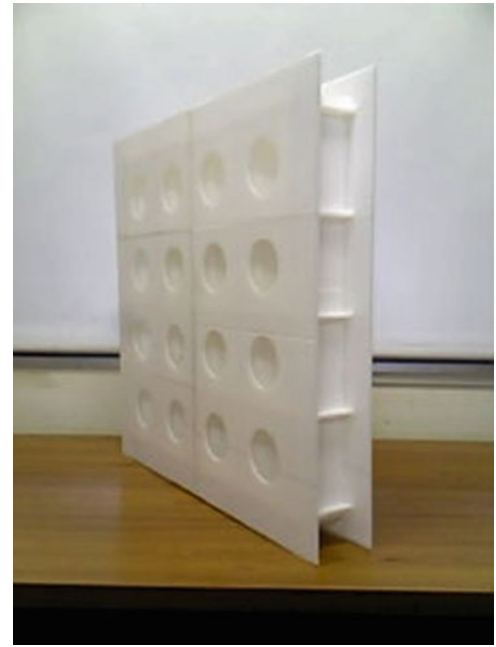
Fig. 16.5 Equalized attenuation of basic type (with or without nonwoven fabric)
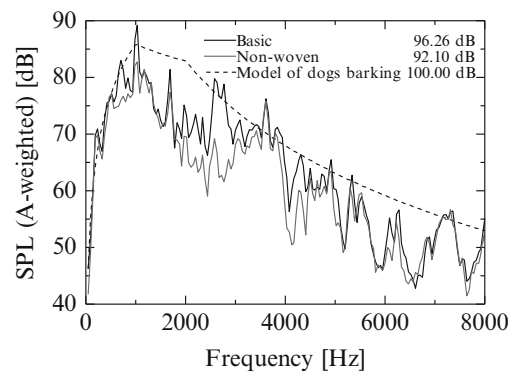
air space effect of the non-woven fabric [3], sound attenuation increased in comparison with that achieved with the basic silencer in a wide frequency range of 1,000–3,000 Hz. We achieved sound attenuation of approximately 8 dB. Compared with the attenuation achieved with the basic silencer, attached the non-woven fabric sheets increased sound attenuation by approximately 4 dB, making this an effective countermeasure.

16.3.3 Addition of a Tail Pipe to the Opening

In this section, we present the results of silencing by throttle effect using tail-pipe.

A photograph of a tail pipe added to the opening of the silencer is shown in Fig. 16.6. The tail pipe was constructed from a poly (vinyl chloride) pipe with an internal diameter of 83 mm, a thickness of 3.0 mm, and a length of 50 mm. A plane wave in the tail pipe with an internal diameter of 83 mm is formed at frequency up to approximately 2 kHz. Therefore, we can treat tail pipe as an acoustic tube in this frequency range. Using Formula (16.1), we can obtain the attenuating peak frequency of the throttle effect by the tail pipe as both ends. Here, we calculated the open-end correction length using $E = 0.4$ for orifice [4] through Formula (16.2).

Fig. 16.6 Outlook of tail tube attached to basic type
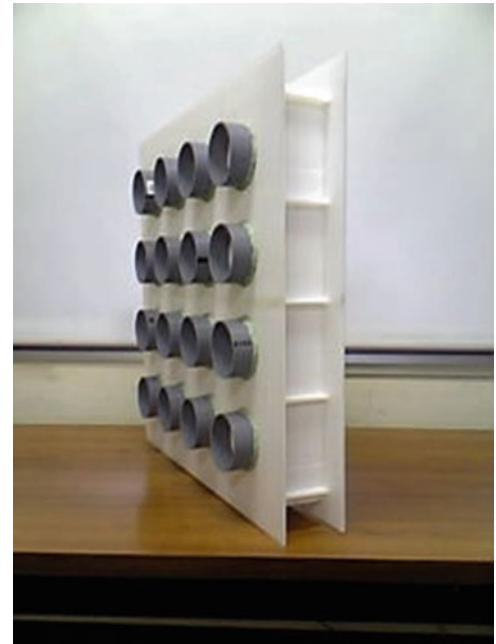
Fig. 16.7 Equalized attenuation with or without tail tube
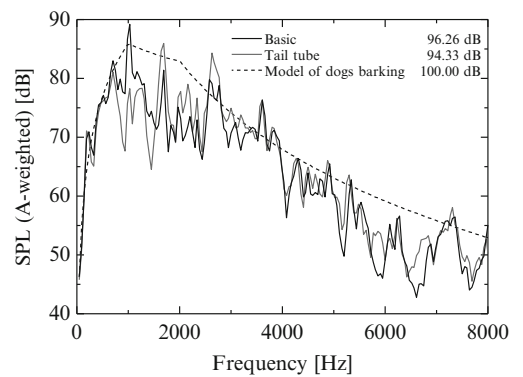
$$f = \frac{mc}{4(l_t + \Delta l_1 + \Delta l_2)} \quad m = 1, 3, 5, \ldots \tag{16.1}$$

$$\Delta l = aE \tag{16.2}$$

l_t: Length of tail tube (m)
a: Inner diameter 0.04 m
E: Open end correction
Δl_1, Δl_2: Open end correction length for each end (m)

According to Formula (16.2), open-end correction length for both ends is 16 mm. Thus the length of an effective tail pipe [$l_t + \Delta l_1 + \Delta l_2$ in Formula (16.1)] can be estimated from this total to be 82 mm. According to Formula (16.1), the first attenuating peak frequency of the tail pipe is estimated to be approximately 1,050 Hz. Besides, the third attenuating peak does not occur because a plane wave does not form within a tail pipe at that frequency.

Measurement results are shown in Figs. 16.7 (basic configuration + tail pipe) and 16.8 (basic configuration + tail pipe + non-woven fabric sheet). In all cases, sound attenuation was great at frequency around 1,000 Hz. These results are well matched to the previously predicted attenuating peak frequency. Furthermore, in both case, the addition of the tail pipe was observed to increase sound attenuation by approximately 2 dB, irrespective of the presence of the non-woven fabric.

Fig. 16.8 Equalized attenuation with or without tail tube (with nonwoven fabric)
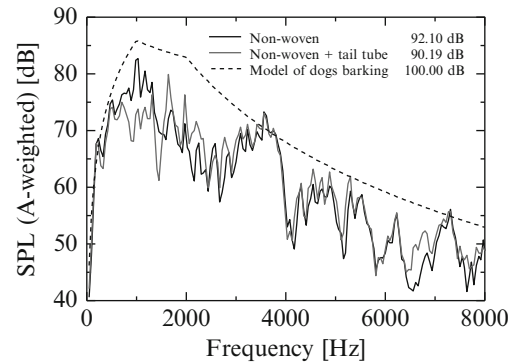
Fig. 16.9 Partial outlook of surrounding tail tube with grating
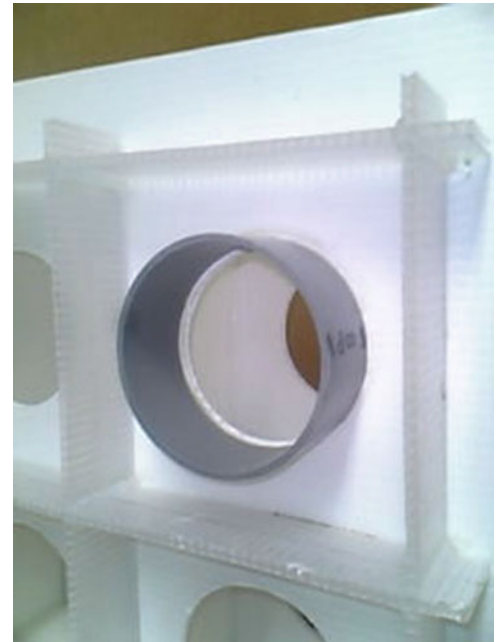
16.3.4 Addition of Side Branch Tube Silencers on the Same Axis to the Opening

Here, we discuss about a silencer that has the tail pipe and grating of the same length as the tail pipe discussed in the previous section. The donut-shaped cross-section pipe fitted along the same axis between the tail pipe and grating as shown in Fig. 16.9 is open on one end and closed on the opposite end. This arrangement produces the same effect as a side branch silencer installed on the same axis as the tail-pipe opening.

Figure 16.10 shows the tail pipe and grating expressed as an electrical equivalent circuit. The open end of the tail pipe expressed with a four-terminal network transfer matrix is terminated with radiation impedance at the open end and incident impedance of the side branch tube. At the resonance frequency of the length of the side branch tube (here, the depth of the grating), the impedance is smaller at the open end. Accordingly, we can expect a sound absorption effect.

The measurement results are shown in Fig. 16.11. Centered on frequency of 1,200–1,400 Hz, sound attenuation was improved in the range of 1–2 kHz. Accordingly, the results show that we matched the previous predictions for the sound attenuation frequency.

Through the addition of the grating, the decreasing of the peak frequency of 1,700 Hz contributed to sound attenuation for dog barking model. And we obtained sound attenuation of 1.2 dB in comparison with that achieved in the previous section. We achieved sound attenuation of 11.01 dB for the dog barking model, which exceeds our initial target of 10 dB.

Fig. 16.10 Equivalent circuit for open end of tail tube (terminated in two impedances)
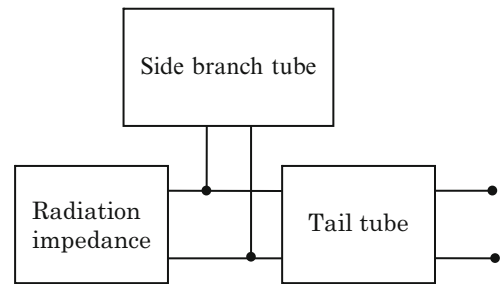
Fig. 16.11 Equalized attenuation with nonwoven fabric and tail tube (with or without grating)
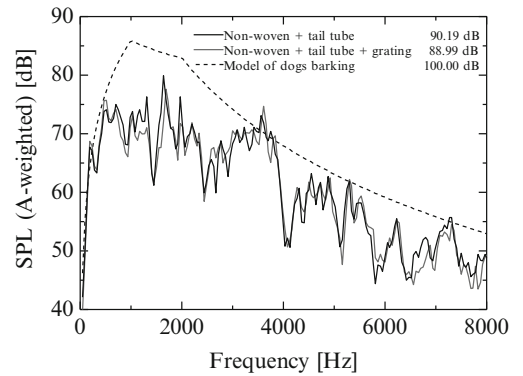
Table 16.1 Attenuation under various conditions (with or without nonwoven fabric)

Equalized attenuation (dB)		
	Without nonwoven fabric	With nonwoven fabric
Rectangular room only (basic type)	3.74	7.90
Basic with tail tube	5.67	9.81
Basic with tail tube and grating	–	11.01

16.3.5 Summarizing Sound Attenuation for the Dog Barking Model

Table 16.1 shows the results obtained for sound attenuation for the dog barking model. The addition of non-woven fabric sheets improved sound attenuation by approximately 2 dB, irrespective of the presence of a tail pipe. Also, the addition of non-woven fabric sheets to either the basic silencer or the silencer with a tail pipe attached resulted in an improvement in sound attenuation by approximately 4 dB.

16.4 Conclusions

We conducted evaluation and experiments with silencers for application to cage openings to alleviate the noise of dogs barking, and we obtained the following results:

1. Silencers based on resonance of rectangular chamber achieved sound attenuation in a sufficiently wide frequency range owing to resonance in a variety of modes.
2. Silencers based on rectangular chamber resonance resulted in greater sound attenuation effects across a wide frequency range by the adding non-woven fabric sheets.
3. The addition of a tail pipe to the opening of the rectangular chamber resulted in sound attenuation effects according to the throttle effect.

4. Regarding radiation impedance, the adding grating to the opening was same as connecting concentric side branch tubes in parallel for open end, and resulted sound attenuation effects through the resonance of a tube with one closed side according to the depth of the grating.
5. By adopting a combination of the previously described silencer structure, we fulfilled the requirements for an aperture ratio that is physiologically sustainable for dogs, allows good visibility into the cage, allows hand access, and is light in weight, with a thickness of 150 mm. We simultaneously achieved sound attenuation results of greater than 10 dB for the dog barking model.

References

1. Hashimoto N (2007) Questionnaire on neighborhood noise dispute to person in charge of noise problem in municipal office. AIJ J Technol Des 13(25):171–176
2. Ministry of environment (Japan) (2004) The 3rd investigative committee of the state of protection management of animal (Document 2 occurrence situations of harms and troublesome problems), Government of Japan, pp 19–21. http://www.env.go.jp/nature/dobutsu/aigo/2_data/arikata/h16_03/mat03.pdf
3. Sakamoto S, Miyairi T, Hoshino A, Kourakata I, Sato T (2013) Attenuation characteristics of double orifice silencer jointed by permeable membrane (measurement and theoretical analysis of sound transmission loss). Trans Jpn Soc Mech Eng Ser C 79(806):3474–3486
4. Bolt HR, Labate S, Ingård U (1949) The acoustic reactance of small circular orifices. J Acoust Soc Am 21(2):94–97
5. Coppola LC, Enns MR, Grandin T (2006) Noise in the animal shelter environment: Building design and the effects of daily noise exposure. J Appl Animal Welfare Sci 9(1):1–7

Chapter 17
On the Use of Regularized DVC to Analyze Strain Localization

Thibault Taillandier-Thomas, Thilo Morgeneyer, Stéphane Roux, and François Hild

Abstract For an in-depth understanding of the failure of structural materials the study of the deformation mechanisms in the bulk is fundamental. In situ synchrotron radiation computed laminography provides 3D images of thin plates subsequently processed by digital volume correlation to measure displacement and strain fields by using the natural contrast of the material. Difficulties arise from the lack of data, which is intrinsic to laminography and leads to several artifacts, and the weak absorption contrast in the 3D image texture of the studied material. To lower uncertainty levels and to have a better mechanical admissibility of the measured displacement fields, a regularized digital volume correlation procedure is introduced and applied to analyze localized displacement and strain fields.

Keywords Artifacts • Digital volume correlation • Laminography • Regularization • Strain localization

17.1 Introduction

Thin sheet structures are widely used in the transportation industry but their failure behavior is not well understood. The three stages of the ductile fracture, namely, void nucleation, growth and coalescence, are well-established but the flat to slant transition of a crack at low triaxialities (Fig. 17.1) is not clear, especially the interaction between damage and strain localization. With 3D imaging combined with full-field measurements techniques, numerous experimental data can be extracted. Synchrotron X-ray computed laminography (akin to tomography) is a 3D imaging technique at the micrometer scale for sheet-like samples, at the cost of additional noise due to the lack of information.

Naturally contrasted volumes can be used in digital volume correlation (DVC) to measure 3D displacement and deduce 3D strain fields. DVC is an extension of 2D digital image correlation to 3D situations [1]. Due to the imaging technique, DVC only relies on the natural contrast of the material, which generally cannot be enhanced. Combined with high noise, DVC cannot necessarily deal with such volumes. A previous feasibility study with a global approach to DVC has been carried out on a similar aluminum alloy [2]. A standard strategy to reach low levels of uncertainty is to use bigger elements, yet at the expense of poor spatial resolution. In the present case, the strain field is localized, which limits the size of the element to properly capture the kinematic fields. Another strategy consists of regularizing the DVC procedure [3, 4] to lower the noise sensitivity and uncertainty. Last, the artifacts may also be filtered at the risk of removing real microstructural markers useful for image registration. The aim of the present paper is to show that such strategies are viable when dealing with very difficult textures.

T. Taillandier-Thomas (✉) • S. Roux • F. Hild
LMT-Cachan, ENS Cachan/CNRS/PRES UniverSud Paris, 61 avenue du Président Wilson, 94235 Cachan Cedex, France
e-mail: thibault.taillandier-thomas@lmt.ens-cachan.fr; stephane.roux@lmt.ens-cachan.fr; francois.hild@lmt.ens-cachan.fr

T. Morgeneyer
Mines ParisTech, Centre des Matériaux, CNRS UMR 7633, BP 87, 91003 Evry Cedex, France
e-mail: thilo.morgeneyer@mines-paristech.fr

Fig. 17.1 Flat to slant transition of an aluminum alloy sheet loaded under mode I
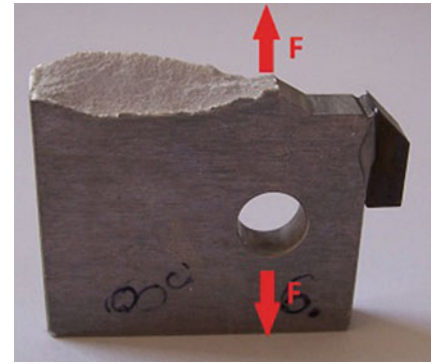
17.2 Experiment

17.2.1 Laminography

Synchrotron radiation computed tomography is a 3D non-destructive technique to study objects extended in one direction and thin in the other two. By tilting the sample by an angle $\theta < 90°$ with respect to the beam direction a region of interest can be scanned (Fig. 17.2). A filtered-back projection algorithm is then used to reconstruct the volume.

The imaging has been performed at beamline ID19 of the European Synchrotron Radiation Facility (Grenoble, France) with a monochromatic beam of 25 keV, a 65° angle and 1,500 scans. The reconstructed volumes have a size of $2,040 \times 2,040 \times 2,040$ voxels, with a voxel size of 0.7 μm.

17.2.2 Material and Experimental Setup

The material used for this study is commercial Al-Cu alloy (AA 2198) in T8 condition. The flat and notched specimen shown in Fig. 17.3a has a size of 60 mm $\times$ 70 mm $\times$ 1 mm, and a notch radius of 0.17 mm (Fig. 17.3b). The measurement will be reported between steps 0 and 1, and steps 0 and 4 as shown in Fig. 17.3d.

The load is applied by turning the two screws of the loading device (Fig. 17.3a). The scanned volume allows us to measure the crack mouth opening displacement (Fig. 17.3b, c). The Region of Interest (ROI) is located 1,000 voxels far away from the notch, whose mid-section cuts are shown in Fig. 17.4a, b. The texture of the material is poor as shown by the gray level histogram in Fig. 17.4c. A volume fraction of 0.4 % of inclusions is observed in the present case. The presence of ring and reconstruction artifacts can be noticed in Fig. 17.4 and will make DVC difficult because it should follow the motion of microstructural features and not the various artifacts.

17.2.3 Regularized Digital Volume Correlation

DVC consists of registering two reconstructed volumes assuming gray level conservation. A global approach is used herein, which is based on a mesh made of C8 elements. The regularization strategy consists of adding some mechanical constrains with a second functional based on the equilibrium gap method, which requires the displacement field to be a solution to an elastic problem with known body forces. A weight is given to this functional, thereby defining the regularization length, which acts as a cut-off wavelength of a low pass filter. The linear system to be solved reads [5]

$$([M] + [N])\{\delta u\} = \{b\} - [N]\{u\} \tag{17.1}$$

where $[M]$ is the DVC matrix, $\{\delta u\}$ the correction to vector $\{u\}$ gathering all unknown degrees of freedom, and $\{b\}$ the DVC vector that needs to decrease so that convergence is achieved (*i.e.* the corrections $\{\delta u\}$ become vanishingly small) and [N] the regularization matrix

Fig. 17.2 Schematic view of computed laminography setup with the ESRF parallel beam line
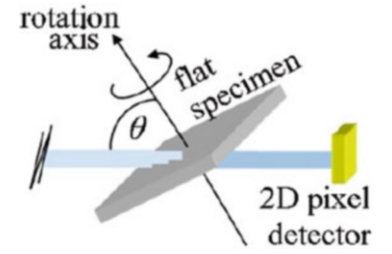
Fig. 17.3 Experimental setup (**a**) and reconstructed 3D laminography data around the notch tip in the reference configuration (**b**) and deformed configuration (**c**). The test is characterized by the crack mouth opening displacement (**d**)
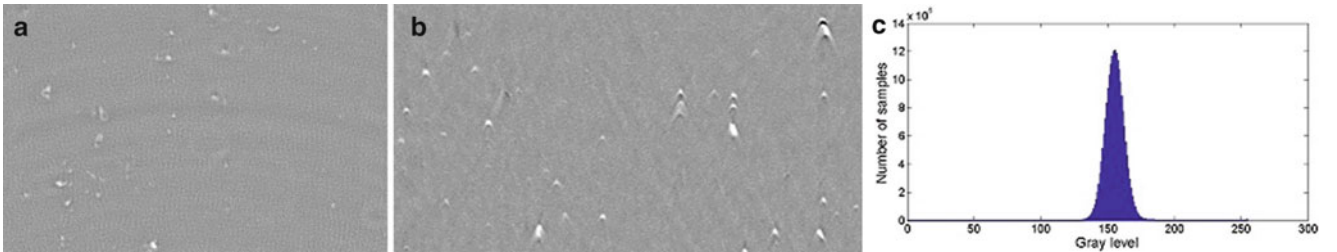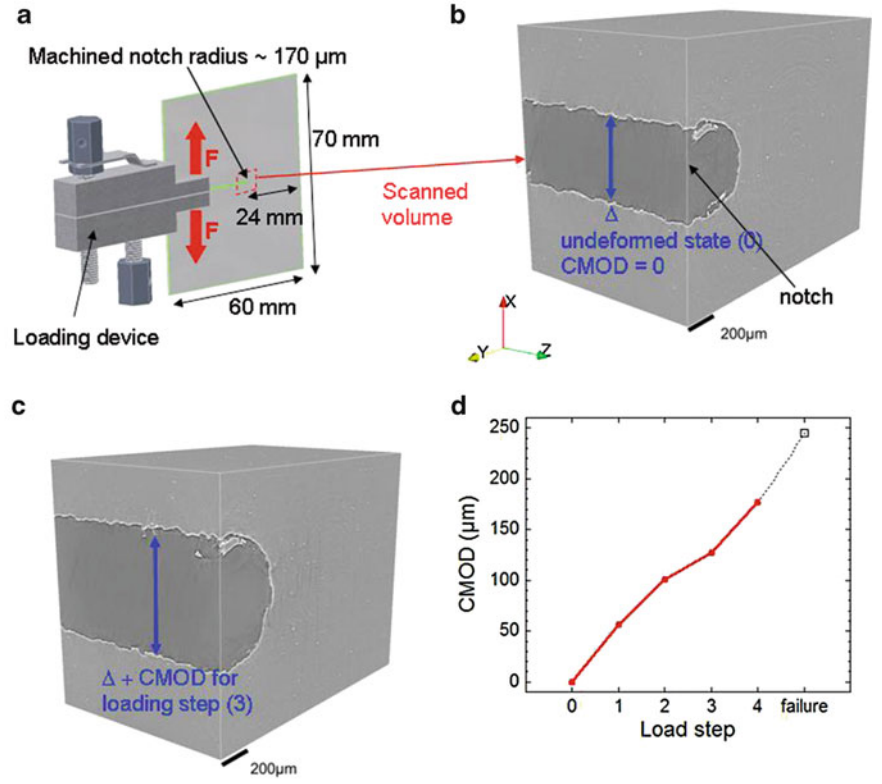
Fig. 17.4 Mid-section cuts (**a**, **b**) of the studied ROI and corresponding gray level histogram (**c**)

$$[N] = \omega_m \frac{\{v\}^t[M]\{v\}}{\{v\}^t[K]^t[K]\{v\}}[K]^t[K] + \omega_b \frac{\{v\}^t[M]\{v\}}{\{v\}^t[L]^t[L]\{v\}}[L]^t[L] \qquad (17.2)$$

with ω_m and ω_b the weights put on the mechanical and boundary regularization, $[K]$ and $[L]$ the matrices associated with these two regularizations and $\{v\}$ a displacement field used for the normalization procedure [5].

17.3 Results and Discussion

17.3.1 Resolution Analysis

The resolution analysis consists of correlating the first two volumes. Very small displacements (i.e., rigid body motions) have occurred between the two acquisitions. However, because of various artifacts, these two volumes are not identical but essentially corrupted by noise. Different displacement fields have been used for the regularization procedure and more information is provided in [5]. However, as shown in Fig. 17.5, for a given element size the uncertainty is lower if more weight is given to the regularization part when 16-voxel elements are chosen.

17.3.2 Digital Volume Correlation on Raw Scans

The chosen ROI has a size of $448 \times 288 \times 544$ voxels. Only the strain fields measured between the reference and early (a) and late (b) loading stages are shown in Fig. 17.6 when 16-voxel elements are used. Strain localization is already captured at the first loading step.

17.3.3 Regularized Digital Volume Correlation on Raw Scans

In the following, the same analysisis carried out with the regularized code for the early stage of loading. The results are shown in Fig. 17.7.

The effect of using a larger regularization length is to lower the uncertainty, see Fig. 17.5 b, and to smoothen the strain field (Fig. 17.7) while still capturing the localization band.

17.3.4 Regularized Digital Volume Correlation on Filtered Scans

Ring and reconstruction artifacts shown in Fig. 17.3b do not bear any mechanical information since they do not rotate along with the inclusions or are fixed in the center of the scanned volume. A numerical filter can be used to erase them while keeping the microstructure safe as shown in Fig. 17.8. The consequence of such a procedure is that the histogram of gray level is even narrower than that shown in Fig. 17.4. The three phases of the material, namely, the matrix, the inclusions and damage are still observed. However, the microstructure is extremely poor, and regularization is needed for DVC. Unregularized DVC does not converge in the present case.
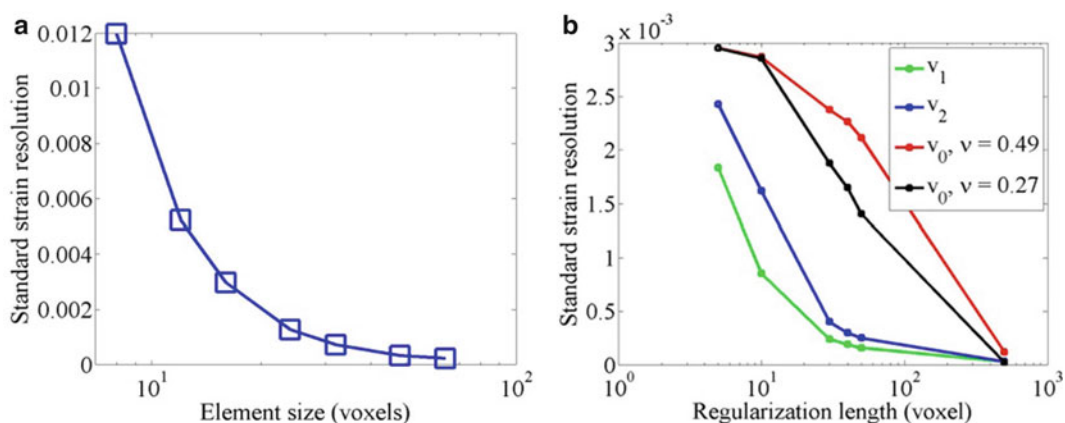
Fig. 17.5 Von Mises' equivalent strain uncertainty measured for different element sizes with no regularization (**a**) and for an element size of 16 voxels and different normalization and regularization lengths (**b**)
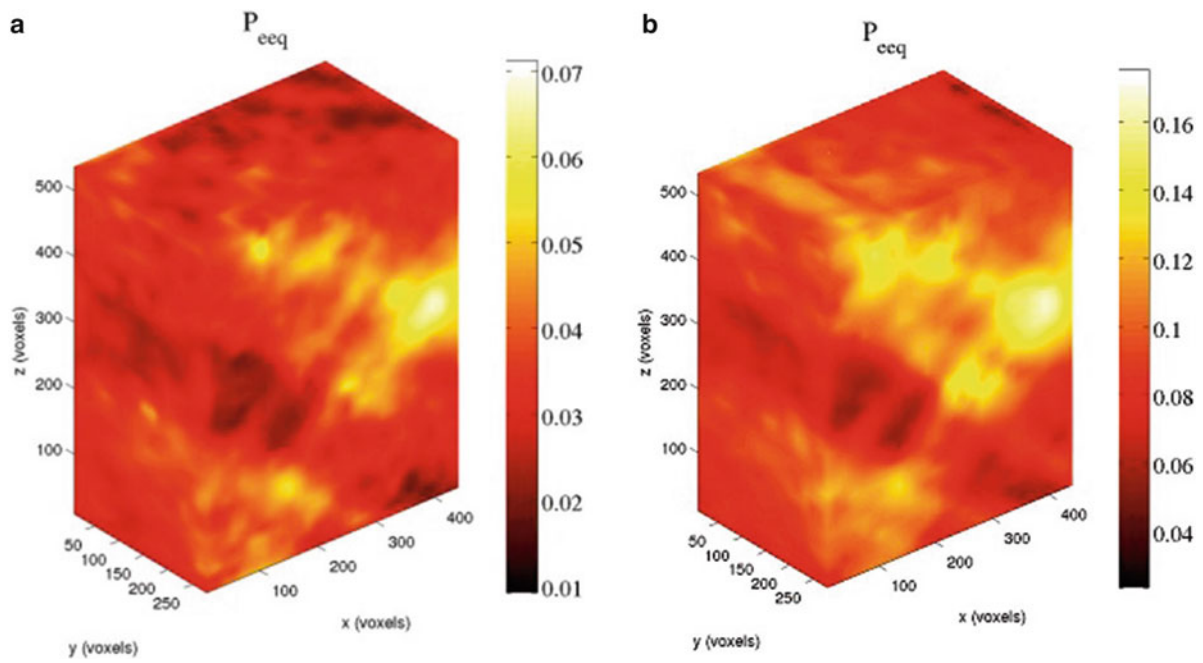
Fig. 17.6 Von Mises' equivalent strain fields measured at early (**a**) and late (**b**) stages of loading of the same volume
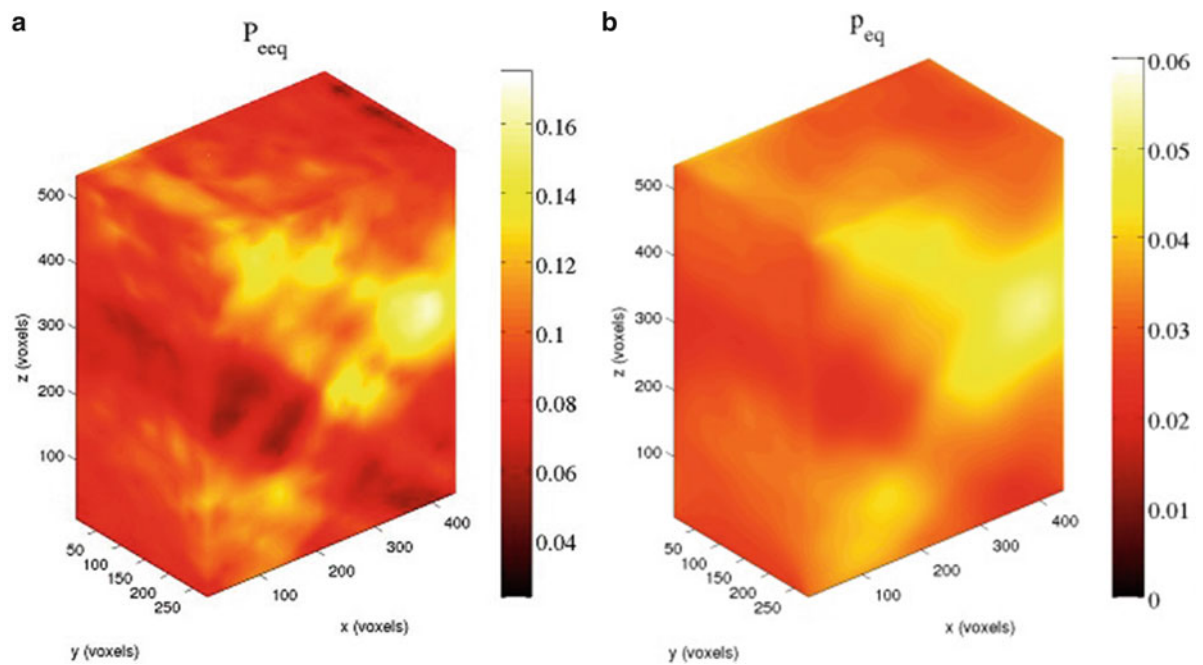
Fig. 17.7 Von Mises' equivalent strain measured at the early stage of loading with a regularization length of 5 voxel (**a**) and 30 voxel (**b**)
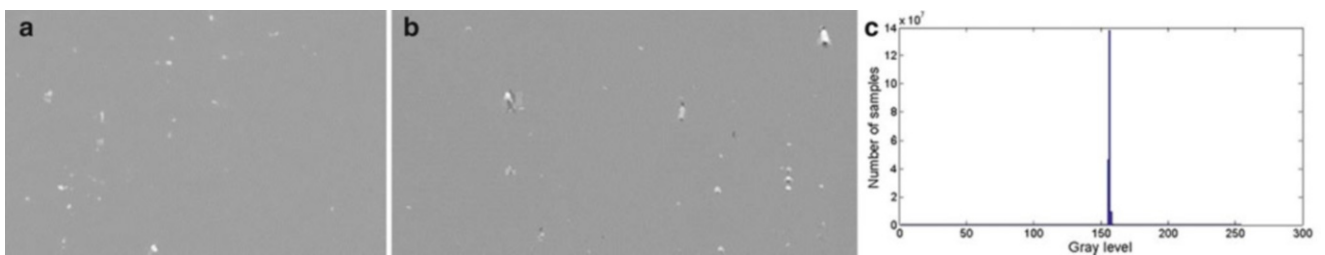
Fig. 17.8 Texture of the filtered scan with the gray level distribution (see Fig. 17.4 for comparison purposes)
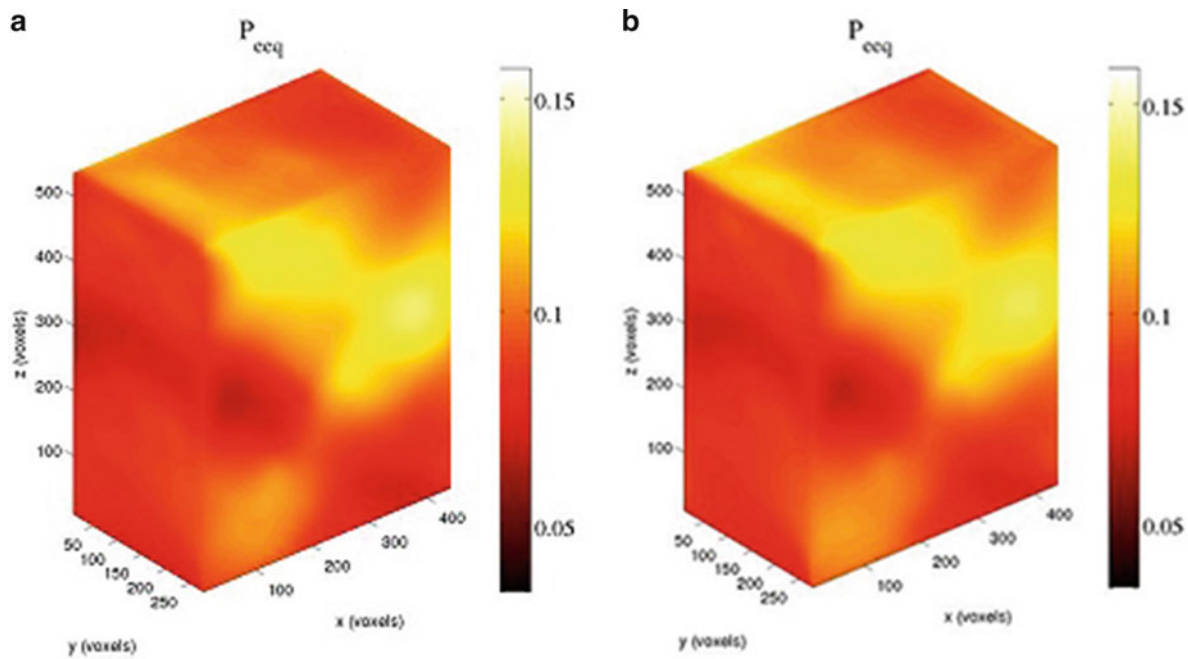
Fig. 17.9 Von Mises' equivalent strain obtained at the late stage with the raw (**a**) and filtered (**b**) volumes (and a mechanical regularization length of 40 voxels)

If regularized DVC performed on the raw scans is to be trusted, the displacement and strain fields measured should be close to those obtained with filtered scans. Figure 17.9 illustrates this point. Von Mises' equivalent strain fields measured at the late stage are almost identical for the raw (a) and filtered (b) volumes.

17.4   Conclusions

Thanks to computed laminography in situ experiments can be performed on plate-like samples with a high definition. The poor contrast of the aluminum alloy can then be successfully registered by global digital volume correlation to measure displacement and strain fields. Results are similar with and without numerical filter to erase reconstruction artifacts, thereby validating that the analyses on the raw images yielded trustworthy kinematic fields. Regularized DVC is even more robust since it can deal with cases where DVC does not work (*i.e.* filtered volumes) and achieves lower uncertainty levels while keeping the same spatial resolution of the measured displacement fields.

References

1. Roux S, Hild F, Viot P, Bernard D (2008) Three dimensional image correlation from X-Ray computed tomography of solid foam. Compos Part A Appl Sci Manuf 39:1253–1265
2. Morgeneyer TF, Helfen L, Mubarak H, Hild F (2013) 3D digital volume correlation of synchrotron radiation laminography images of ductile crack initiation: an initial feasibility study. Exp Mech 53(4):543–556
3. Tomičević Z, Roux S, Hild F (2013) Mechanics-aided digital image correlation. J Strain Anal 48:330–343
4. Morgeneyer TF, Taillandier-Thomas T, Helfen L, Baumbach T, Sinclair I, Roux S, Hild F (2014) In situ 3D observation of early strain localisation during failure of thin Al alloy (2198) sheet. Acta Mater 69:78–91
5. Taillandier-Thomas T, Roux S, Morgeneyer TF, Hild F (2014) Localized strain field measurement on laminography data with mechanical regularization. Nucl Instrum Methods Phys Res B 324:70–79. doi:10.1016/j.nimb.2013.09.033,2014

Chapter 18
Determination of Surface Bi-Axial Stresses Using Raman Spectroscopy

M. Shafiq and G. Subhash

Abstract The forthcoming discussion presents a new method to estimate Residual stresses using Raman spectroscopy. The method has been validated independently through an existing reliable surface strain measurement technique, known as Digital Image Correlation (DIC). Utilizing phonon deformation potentials for diamond and zinc blende crystal structures, a relationship was first introduced that provides a quantifiable estimate between the Raman peak-shift and residual stresses in confined Silicon Carbide particles. The proposed relationship was further validated independently using classical formulation for thick walled cylinders and DIC techniques. The results obtained show excellent agreement between the measured stresses and the derived relationship between Raman peak-shift and the applied stress.

Keywords ZrB_2-SiC • Residual stresses • Raman spectroscopy • Digital image correlation (DIC) • Thick walled cylinders

18.1 Introduction

Raman spectroscopy has been evolved as a widely accepted technique to determine local thermal residual stresses in semiconductor structures [1]. The mismatch in thermal coefficients of expansion between a thin film and substrate material give rise to local strains which manifest themselves as change in atomic force constants [2, 3]. The perturbation of these lattice parameters can be quantified with respect to the virgin state of the material, provided the deformation potentials are known under stress free condition [3]. Thus the sensitivity of Raman wave numbers to applied strain and its high spatial resolution (1 μm) can be effectively utilized to obtain a reliable measurement of local residual stresses in Raman active materials [3].

In this study we have utilized zirconium diboride-silicon carbide (ZrB_2-SiC) composite to develop an analytical framework for determining residual stresses as a result of external confinement. ZrB_2-SiC is a very brittle material and used in high temperature advance aerospace application due to its extremely high melting point and excellent oxidation resistance in highly corrosive environment [3]. However, use of ZrB_2-SiC in such applications requires an understanding of its deformation behavior under a range of loading conditions. Brittle materials are known to perform better when confined by a lateral stress [4]. Therefore confinement studies on the bulk material can provide valuable data for characterizing ZrB_2-SiC.

ZrB_2-SiC was chosen for this study because it offers two advantages. First, the unusual high electrical conductivity of the matrix material (ZrB_2) allowed the machining of cylindrical pallets, which would have not been otherwise possible for the second phase SiC (purely Raman active) material [5, 6]. Furthermore, the matrix material provides confinement on second phase particles i.e. SiC. Second, SiC which is a very brittle material and cannot be economically machined into cylindrical pellets provides an opportunity to probe Raman peaks due to its Raman peaks sensitivity to applied stress [5]. In addition to that, SiC has been studied extensively and its optical phonon deformation potentials can be estimated which are required in the calculation of residual stresses through Raman peak-shift.

In this manuscript a systematic study is conducted by first developing a relationship between the Raman peak-shift to the magnitude of confinement and then estimating the residual stresses through Raman peak-shift on individual SiC particles using two steps. In the first step, the specimen was prepared using thermal shrink-fit metallic sleeves to create external

M. Shafiq • G. Subhash (✉)

Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville 32611, USA

e-mail: subhash@ufl.edu

bi-axial pressure on the lateral surface of ZrB$_2$-SiC composite. The Raman peaks were probed and estimated stresses were compared to the one's obtained from classical formulation for thick walled cylinders. In the second step, the specimen was confined in a shaft-collar and progressively increasing load was applied on the lateral surface of the specimen. The two planar faces were used for Raman peak-shift probing and DIC consecutively. In the final step, the stresses obtained from these two techniques were compared to each other. Hence, the entire procedure was validated using two independent experimental measurement techniques.

18.2 Experimental

18.2.1 Raman Peak-Shift Relation to Residual Stresses

Ganesan et al. [7] showed that the frequencies of three optical phonons in a deformed material can be determined by evaluating the Eigenvalues, $\lambda_j = 1 - 3$, of the following secular equation

$$\begin{vmatrix} p\varepsilon_r^R + q(\varepsilon_\theta^R + \varepsilon_z^R) - \lambda & r\gamma_{r\theta}^R & r\gamma_{rz}^R \\ r\gamma_{r\theta}^R & p\varepsilon_\theta^R + q(\varepsilon_r^R + \varepsilon_z^R) - \lambda & r\gamma_{\theta z}^R \\ r\gamma_{rz}^R & r\gamma_{\theta z}^R & p\varepsilon_z^R + q(\varepsilon_r^R + \varepsilon_\theta^R) - \lambda \end{vmatrix} = 0 \tag{18.1}$$

Where p, q and r are phonon deformation potentials, ε^R and γ^R are residual normal and shear strain tensor components. Under the assumption of small strain the frequency w_j of the each mode of the deformed crystal is related to the unreformed state through the following relationship [1, 5]

$$\Delta w_j = w_j - w_{j_o} \approx \frac{\lambda_j}{2w_{j_o}} \tag{18.2}$$

The residual strain tensor ε in Eq. (18.1) can be related to the residual stress tensor for diamond and zinc blende crystal structure of 3C-SiC using Hooke's law as [5]:

$$\begin{Bmatrix} \varepsilon_r^R \\ \varepsilon_\theta^R \\ \varepsilon_z^R \\ \gamma_{r\theta}^R \\ \gamma_{rz}^R \\ \gamma_{\theta z}^R \end{Bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & 0 & 0 & 0 \\ S_{21} & S_{22} & S_{23} & 0 & 0 & 0 \\ S_{31} & S_{32} & S_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & S_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & S_{66} \end{bmatrix} \begin{Bmatrix} \sigma_r^R \\ \sigma_\theta^R \\ \sigma_z^R \\ \tau_{r\theta}^R \\ \tau_{rz}^R \\ \tau_{\theta z}^R \end{Bmatrix} \tag{18.3}$$

Where, S_{ij} are compliance terms, and σ^R and τ^R are the residual normal and shear stress components, respectively. Considering only the largest principal stresses and cubic crystal structure of 3C-SiC, the following simplifications can be made [5]

$$\begin{aligned} & S_{11} = S_{22} = S_{33} \\ & S_{12} = S_{13} = S_{21} = S_{23} = S_{31} = S_{32} \\ & S_{44} = S_{55} = S_{66} \\ & \sigma_z^R \approx 0 \\ & (\tau_{r\theta}^R, \tau_{rz}^R, \tau_{\theta z}^R, \gamma_{r\theta}^R, \gamma_{rz}^R, \gamma_{\theta z}^R) \approx 0 \end{aligned} \tag{18.4}$$

Using Eq. (18.4), Eq. (18.3) can be reduced to

$$\begin{aligned} \varepsilon_r^R &= (S_{11}\sigma_r^R + S_{12}\sigma_\theta^R) \\ \varepsilon_\theta^R &= (S_{11}\sigma_\theta^R + S_{12}\sigma_r^R) \\ \varepsilon_z^R &= S_{12}(\sigma_r^R + \sigma_\theta^R) \end{aligned} \tag{18.5}$$

Substituting Eq. (18.5) in Eq. (18.1), the Eigenvalues can be obtained as

$$\lambda_1 = \{pS_{12} + (S_{11} + S_{12})q\}\left(\sigma_r^R + \sigma_\theta^R\right)$$
$$\lambda_2 = \{pS_{12} + (S_{11} + S_{12})q\}\sigma_r^R + \{pS_{11} + 2qS_{12}\}\sigma_\theta^R \qquad (18.6)$$
$$\lambda_3 = \{S_{11}p + 2S_{12}q\}\sigma_r^R + \{S_{11}p + (S_{11} + S_{12})q\}\sigma_\theta^R$$

Substituting the above Eigenvalues in Eq. (18.2), the relationship between the Raman peak-shit and the residual stresses are derived as

$$\Delta w_1 = \frac{1}{2w_{01}}\{pS_{12} + (S_{11} + S_{12})q\}\left(\sigma_r^R + \sigma_\theta^R\right) \qquad (18.7)$$

$$\Delta w_2 = \frac{1}{2w_{02}}\left[\{pS_{12} + (S_{11} + S_{12})q\}\sigma_r^R + \{pS_{11} + 2qS_{12}\}\sigma_\theta^R\right] \qquad (18.8)$$

$$\Delta w_3 = \frac{1}{2w_{03}}\left[\{S_{11}p + 2S_{12}q\}\sigma_r^R + \{S_{11}p + (S_{11} + S_{12})q\}\sigma_\theta^R\right] \qquad (18.9)$$

The above peak-shift relations can be used to calculate local residual stresses within individual SiC particles based on Raman peak numbers from undeformed and deformed particles. In a strained zinc blende crystal structure a maximum o two transverse optical (TO) and one longitudinal optical (LO) Raman modes are possible [5]. Equations (18.7) and (18.8) represents TO peaks, while Eq. (18.9) correspond to the LO peak. Because of the symmetricity and clarity of TO peak, we will use Eq. (18.7) for the calculation of residual stresses. However, Eq. (18.7) requires values of phonon deformation potentials p, q and r which are not directly available in the literature. Therefore, mode Grüneisen parameters for hydrostatic stress (γ_o) and uniaxial (γ_s) are used to yield indirect values for p and q, where (γ_o) and (γ_s) are given by [5]:

$$\gamma_o = -\frac{(p + 2q)}{6w_o^2}$$
$$\gamma_s = \frac{(p - q)}{2w_o^2} \qquad (18.10)$$

Where w_o is the Raman peak for the undeformed crystal structure. Assuming $\gamma_o^{TO} = \gamma_s^{TO} = 1.10$ and taking $w_o = 796.98\,cm^{-1}$, the deformation potentials are calculated as [5]

$$p^{TO} = -0.4658 \times 10^6\,cm^{-2}$$
$$q^{TO} = -1.8632 \times 10^6\,cm^{-2} \qquad (18.11)$$

Compliance constants are taken as reported by Tolpygo [8] and are given as

$$S_{11} = 3.7 \times 10^{-13}\ cm^2/dyne$$
$$S_{12} = -1.05 \times 10^{-13}\ cm^2/dyne \qquad (18.12)$$

Substituting Eqs. (18.11) and (18.12) into (18.7), the relationship between the Raman peak-shift and residual stress is given as:

$$\sigma_r^R + \sigma_\theta^R = -358.32\,(\Delta w_1)\,MPa \qquad (18.13)$$

Where Δw_1 is the Raman peak-shift for TO peak position. Eq. (18.13) can be further simplified for the ideal case, when $\sigma_r^R = \sigma_\theta^R$ and the final expression for residual stress calculation is given as

$$\sigma_r^R = -179.16\,(\Delta w_1)\,MPa \qquad (18.14)$$

The above equation provides a relationship between the Raman peak-shift Δw and the magnitude of confinement stress σ_r. It is evident from Eq. (18.14) that compressive and tensile residual stresses induce a positive and negative Raman peak-shift, respectively.

18.3 Specimen Preparation

18.3.1 Shrink Fit Assembly

The microstructure of ZrB_2-SiC composite consists of uniformly distributed SiC particles embedded in ZrB_2 matrix material as shown in Fig. 18.1. The average size of the matrix and particulate phase is 1 μm and 5 μm, respectively. Taking advantage of the metal like high electrical conductivity of the matrix material (ZrB_2) cylindrical pellets were cut using electrical discharge machining (EDM) measuring 4 mm in diameter and 7 mm in length, see Fig. 18.2a.

External pressure was applied on the lateral surface by two approaches (1) installing a thermal shrink fit metal sleeve, see Fig. 18.2b and (2) using a shaft-collar ring, as shown in Fig. 18.3a, b thermal shrink fitting metallic sleeves. The sleeves were machined to have an inner diameter slightly smaller than the specimen diameter. The ceramic-sleeve assembly was achieved by heating the metallic sleeve to a pre-determined temperature in a furnace such that the inner diameter of the sleeve is equal or slightly larger than the specimen diameter and then forcing it onto the ceramic specimen by using an axial force. The pressure due to confinement was calculated by treating the sleeve as thick walled pressure vessel and solving an axisymmetric boundary value problem on the cross-section of a circular elastic ceramic specimen inside an elastic perfectly

Fig. 18.1 SEM image of polished ZrB_2-SiC surface revealing SiC embedded in ZrB_2 matrix [5]
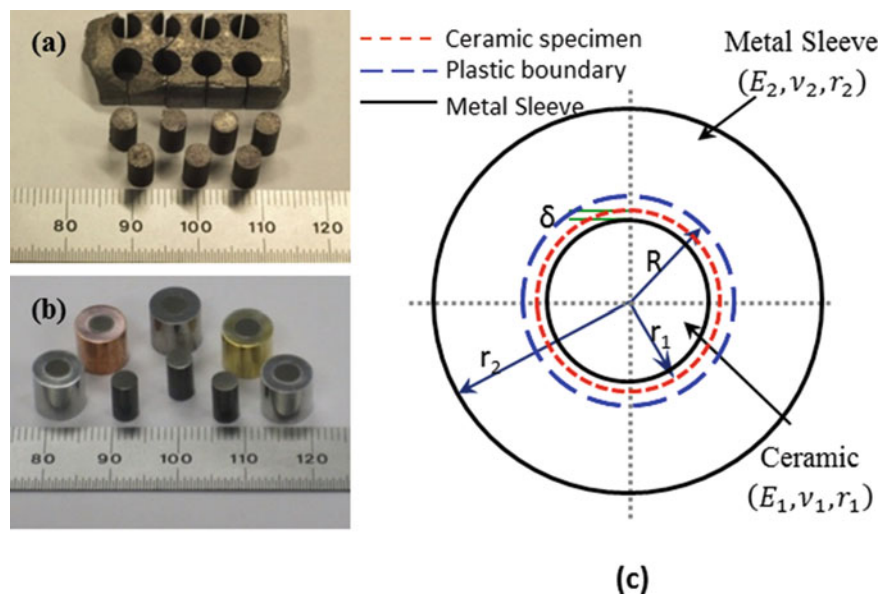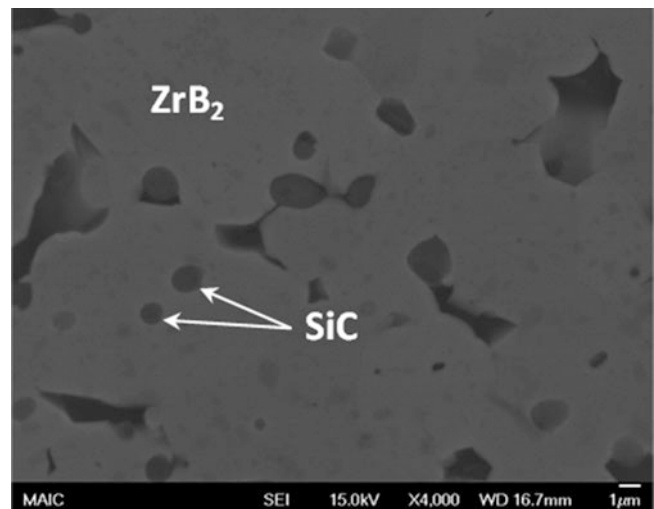
Fig. 18.2 (**a**) Cylindrical specimens of ZrB_2-SiC cut using electric discharge machining (EDM), (**b**) Shrink-fit specimen assembly with various metal sleeves and, (**c**) schematic of an axisymmetric boundary value problem [5]
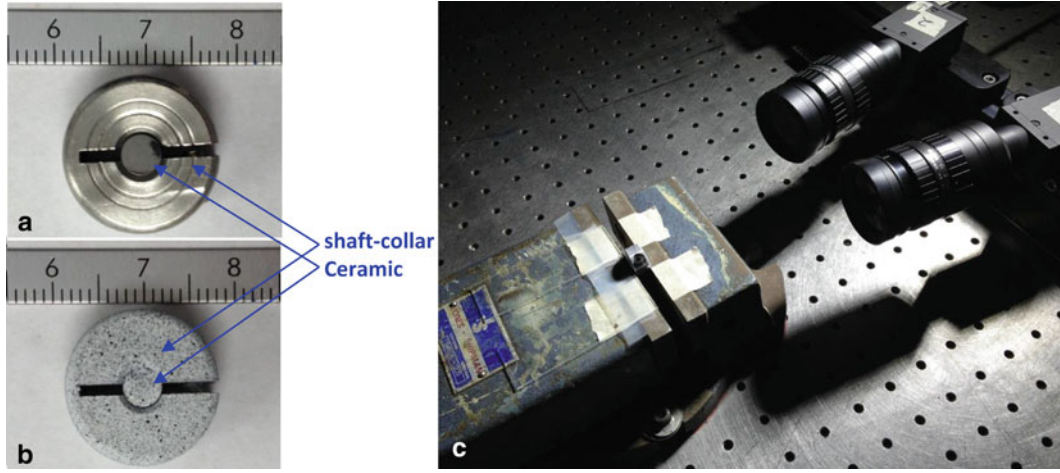
Fig. 18.3 (**a**) Uncoated surface (for Raman measurements) and (**b**) coated surface (for DIC) of shaft-collar ring and ZrB$_2$-SiC assembly, and (**c**) Digital Image Correlation (DIC) setup comprised of cameras, extension tubes, lenses and shaft-collar ZrB$_2$-SiC specimen assembly [5]

Table 18.1 Sleeve materials and estimated confinement pressure using Eqs. (18.15) and (18.16) [5]

Sleeve material	δ (mm)	r_1 (mm)	r_2 (mm)	R (mm)	E (GPa)	v	Pressure (MPa)
Aluminum-2024	0.02	1.99	3.92	2.64	73	0.33	180
Brass-385	0.011	1.99	3.99	2.98	98	0.32	195
Stainless Steel-303	0.014	1.91	3.99	2.52	196	0.25	284

plastic sleeve material with a misfit δ at the interface as shown in Fig. 18.2c and elastic constants (E_1, v_1) and (E_2, v_2), respectively. The resulting expressions are as given below:

$$\frac{\delta}{2r_1} = (1 + v_2)(1 - 2v_2)\left(\frac{\sigma_{Y_2}}{E_2}\right)\left[ln\left(\frac{r_1}{R}\right) + \frac{1}{2}\left(\frac{R}{r_2}\right)^2 - \frac{1}{2}\right] + (1 - v_2^2)\left(\frac{\sigma_{Y_2}}{E_2}\right)\left(\frac{R}{r_1}\right)^2$$
$$+ (1 - v_1)\left(\frac{\sigma_{Y_2}}{E_1}\right)\left[ln\left(\frac{R}{r_1}\right) + \frac{1}{2}\left(1 - \left(\frac{R}{r_2}\right)^2\right)\right] \tag{18.15}$$

$$P = \sigma_{Y_2}\left[ln\left(\frac{R}{r_1}\right) + \frac{1}{2}\left(1 - \left(\frac{R}{r_2}\right)^2\right)\right] \tag{18.16}$$

Where E and v are Young's modulus and Poisson's ratio and subscript 1 and 2 corresponds to the ceramic and metal sleeve, respectively. P is the induced confining pressure, σ_{Y_2} is the yield strength of the metallic sleeve, r_1 and r_2 are the specimen radius and outer radius of the sleeve material, respectively. δ is the misfit between the specimen diameter and the sleeve inner diameter and R is the radius of elastic-plastic boundary in the sleeve cross-section. Different confinement pressures were applied using three different sleeve materials. Due to the difference in elastic constants each sleeve material provided different confinement pressure as listed in Table 18.1.

18.4 DIC on Shaft-Collar Confined Specimens

To validate the Raman peak-shift relation for residual stress determination an unknown level of confinement stress was applied on the lateral surface of cylindrical ZrB$_2$-SiC pellets using shaft-collar rings, as shown in Fig. 18.3a, b. Among the exposed planar surfaces, one is used to measure surface strain by conducting DIC while the other is used for Raman peak-shift measurement. A total of 16 well identified SiC particles were used to obtain Raman peak-shift data.

18.5　Results and Discussion

Figure 18.4 presents the estimated peak-shift data for several SiC particles and the calculated confinement stress for the same using Eq. (18.14). The measurement was conducted on individual SiC particles at various distances from the center and along the radius of the specimen. The data is plotted for three different confinement pressures obtained using three different sleeve materials. The same data is replotted in Fig. 18.5 where the confinement stress calculated from Eqs. (18.15) and (18.16) is also shown. These two values from two different approaches match reasonably well and support the relationship derived in Eq. (18.14).

The Raman peak-shift relation was further validated using DIC. First, an unknown level of confinement was applied using the shaft-collar ring, see Fig. 18.3a, b and the average strain on one planar surface of the specimen was calculated using surface displacement. Corresponding stresses were obtained by multiplying the average strain to the Young's modulus of the bulk material. On the other planar surface of the specimen, a total of 16 cylindrical shaped SiC particles were identified and Raman peak-shift data was obtained from the virgin and deformed state. Both DIC and Raman scans were repeated for three load increments. Table 18.2 shows a comparison of average confinement pressure obtained from DIC measurement on one planar surface and the Raman spectroscopy on individual SiC particles on the other planar surface. Figure 18.6a shows the uniformity of DIC-measured strain on one surface (the red spot is due to stress concentration caused by a rather coarser dot of

Fig. 18.4 The measured Raman peak-shift and the estimated lateral pressure [Eq. (18.14)] along the radius of cylindrical ZrB$_2$-SiC specimens [5]
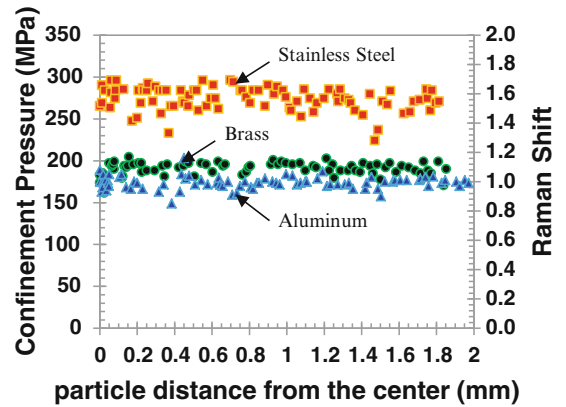
Fig. 18.5 Estimated confinement pressure obtained using Eqs. (18.16) and (18.14) along the radius and as a function of Raman peak-shift [5]
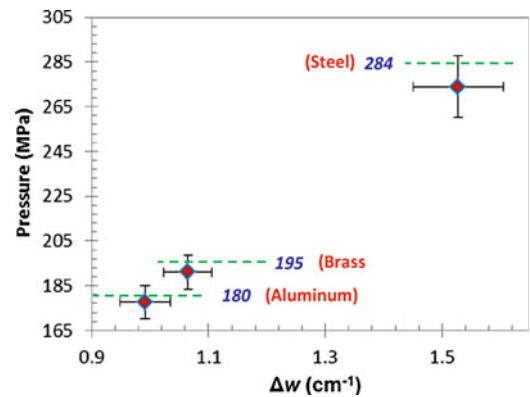
Table 18.2 Shaft-collar ring confinement pressures measured using DIC and [Eq. (18.14)]

Load step	$\Delta w \ (cm^{-1})$	Confinement pressure	
		Eq. (18.14)	DIC (MPa)
0th	0	0	0
1st	0.93	168	189
2nd	2.01	361	379
3rd	2.75	493	511
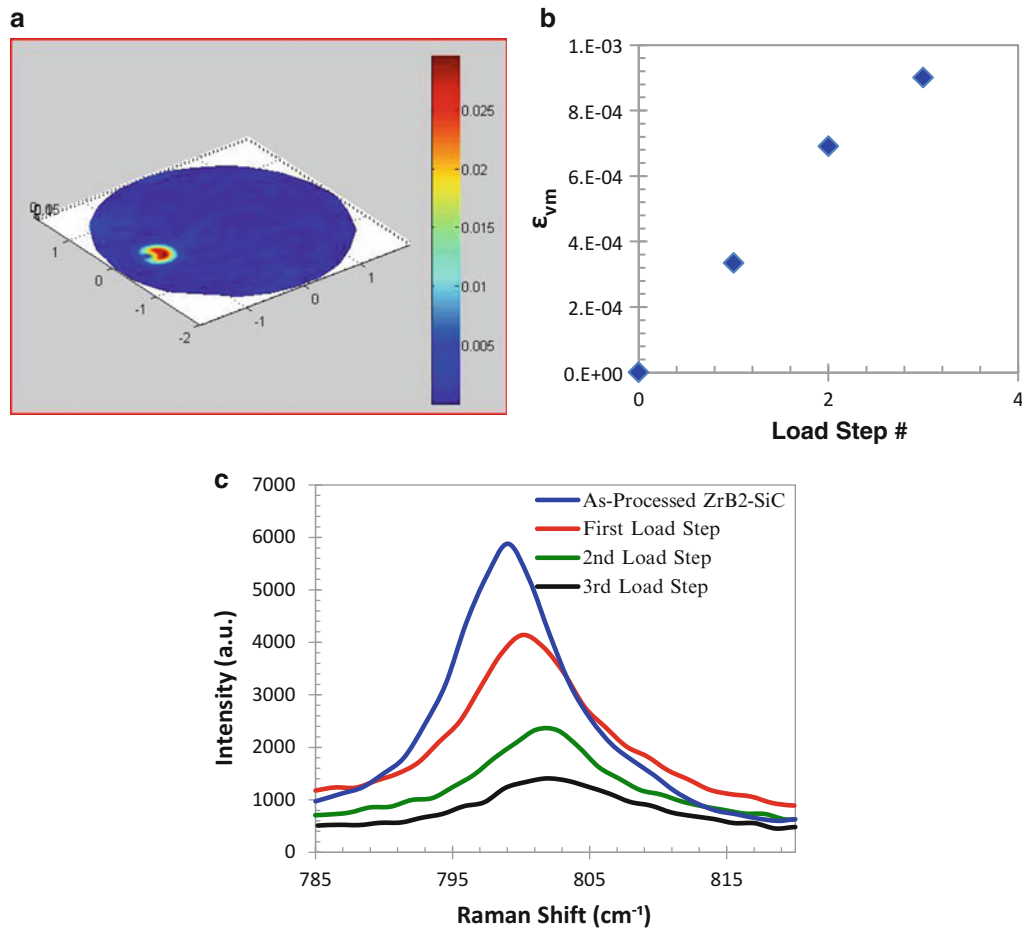
Fig. 18.6 (**a**) Plot of Von-mises strain on one planar face of cylindrical ZrB_2-SiC specimen measured using DIC, (**b**) Average Von-mises strain on the surface of ZrB_2-SiC at different load steps estimated using DIC, and (**c**) Raman peak positions in SiC particles for ZrB_2-SiC specimen in the as-sintered composite as well as three different confinement stresses applied through shaft-collar ring [5]

Valspar paint) and Fig. 18.6b shows the measured average Von-mises strain (ε_{vm}) at three levels of confinement using the shaft-collar ring. Figure 18.6c shows the corresponding TO Raman peaks for these three stress levels, where the Raman peak-shifts to the right with increasing confinement stress.

The Raman peak-shift to higher wave numbers, see Fig. 18.6c, was a direct consequence of macroscopic strain induced within SiC particles due to progressively increasing confinement pressure. The applied external pressure on bulk ZrB_2-SiC specimen governs the magnitude of macroscopic strain induced within SiC particles and hence the degree of distortion of crystal lattice. The Raman peak width or peak asymmetry with confinement pressure is a result of photon damping by lattice imperfections and grain boundaries [5]. The sharp Raman peaks in the as-sintered composite indicates perfect crystallinity as compared to blunt peaks in the deformed composite [5]. Based on the observation of Fig. 18.6c, it can be inferred that the application of biaxial compressive stresses on SiC particles may also cause an increase in the density of stacking faults thus increasing the structural disorder within SiC. Consequently, Raman peak positions are shifted to higher wave numbers [5].

Figure 18.7 summarizes all the results obtained from the DIC on one planar surface of a ZrB_2-SiC specimen and the Raman peak-shift measured stress on selected SiC particles on the other planar surface of the shaft-collar ring specimen assembly. The residual stress values estimated from shrink-fit sleeve assembly were also superposed on the same plot. Interestingly, all three methods predict the same linear relationship between confinement stress and peak-shift. The confinement tress due to processing induced thermal stress was also shown which follows the same relationship.

Fig. 18.7 Summary
of experimental results from
two different and independent
techniques confirming a linear
relationship between Raman
peak-shift and confinement
pressure [5]
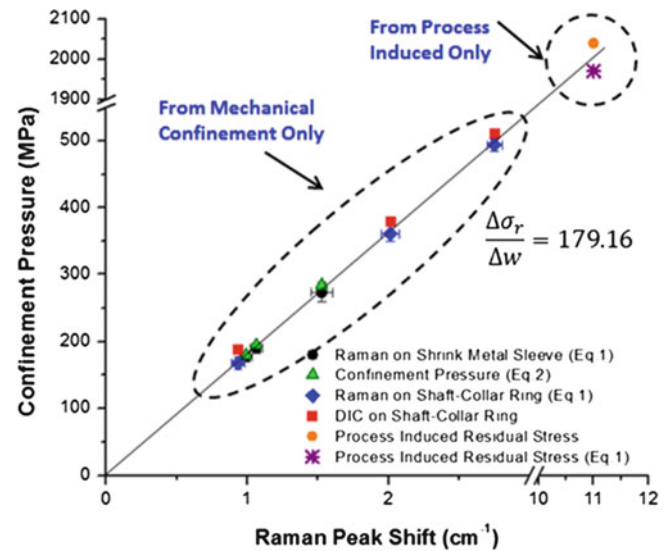
18.6 Conclusion

A new method for determination of residual stresses using Raman spectroscopy in structural ceramics with zinc blende crystal structure was presented and validated. An equation relating the residual stresses and Raman peak-shift was first derived and then verified by comparing the residual stress values determined from shrink-fit metal sleeve assembly. The method is further validated by applying a progressively increasing unknown confinement stress (by shaft-collar ring) which was consecutively and independently determined using DIC. The results from all the methods predict a linear relationship between the applied confinement stress and the Raman peak-shift. The method is simple to apply and can be used widely on any Raman active material to determine the unknown stress.

References

1. Anastassakis E, Pinczuk A, Burstein E, Pollak FH, Cardona M (1970) Effect of static uniaxial stress on the Raman spectrum of silicon. Solid State Commun 8:133–138
2. Kang Y, Qiu Y, Lei Z, Hu M (2005) An application of Raman spectroscopy on the measurement of residual stress in porous silicon. Opt Lasers Eng 43:847–855
3. Ghosh D, Subhash G, Orlovskaya N (2008) Measurement of scratch-induced residual stress within SiC grains in ZrB2–SiC composite using micro-Raman spectroscopy. Acta Mater 56:5345–5354
4. Ravichandran G, Subhash G (1995) A micromechanical model for high strain rate behavior of ceramics. Int J Solids Struct 32:2627–2646
5. Shafiq M, Subhash G (2014) A novel technique for the determination of surface biaxial stress under external confinement using Raman spectroscopy. Exp Mech 54:763–774. doi:10.1007/s11340-014-9851-9
6. Guo S, Kagawa Y (2012) High-strength zirconium diboride-based ceramic composites consolidated by low-temperature hot pressing. Sci Technol Adv Mater 13:045007
7. Ganesan S, Maradudin AA, Oitmaa J (1970) A lattice theory of morphic effects in crystals of the diamond structure. Ann Phys 56:556–594
8. Tolpygo KB (1961) Optical, elastic and piezoelectric properties of ionic and valence crystals with ZnS-type lattice. Sov Phys Solid State 2:2367–2376

Chapter 19
Visualization and Quantification of Quasi-Static and Dynamic Surface Slopes Using a Reflection-Mode Digital Gradient Sensor

Amith Jain, Chandru Periasamy, and Hareesh Tippur

Abstract A full-field, reflection-mode Digital Gradient Sensing (DGS) technique capable of measuring small angular deflections of light rays reflected off specularly reflective planar surfaces is developed. The method is aided by 2D digital image correlation principle to quantify angular deflections of light rays. In this paper, the principle of the method is described and the governing equations relating light ray deflections to surface slopes are presented. The method is demonstrated by simultaneously mapping orthogonal surface slopes of a circumferentially clamped silicon wafer subjected to a central deflection. The curvature fields, proportional to stresses in thin plates, as well as surface topography are evaluated from the measured slopes. Subsequently, the method is demonstrated for studying a stress-wave dominant problem of a free, thin, compliant plate subjected to impact by a rigid spherical ball.

Keywords Optical metrology • Digital image correlation • Thin structures • Slope • Curvature • Topography • Impact loading

19.1 Introduction

Thin reflective engineering structures such as space mirrors, microelectronic devices, photovoltaic devices, to name a few often fail when subjected to mechanical stresses. Hence quantification of deflections, slopes and curvatures under static and dynamic loading conditions is critical. Non-contacting methods to perform such measurements at relatively high sensitivities are important when small deflections are to be detected and quantified in stiff structures and materials. It should also be noted that mechanical stresses in thin structures are proportional to curvatures which in turn are the first and second derivatives of slopes and deflections. As numerical differentiation of measured data is often prone to substantial errors, direct measurement of surface slopes is preferred for evaluating curvatures.

The Coherent Gradient Sensing (CGS) method in reflection mode presented by Tippur [1] was capable of simultaneously measuring slopes and curvature fringes in a thin silicon wafer subjected to mechanical loading. Recently, Periasamy and Tippur [2] have shown that DGS method in reflection-mode is effective to measure surface slopes of thin structures using popular digital image correlation principles.

In the following, experimental details and governing equations of DGS in reflection-mode are described. The method is demonstrated by measuring orthogonal surface slopes for a circular silicon wafer clamped at the edge and subjected to a central deflection. The measured surface slopes are compared with the theoretical counterparts. Subsequently, the measured slopes are differentiated to obtain curvatures and integrated to obtain surface topography. The processed results are compared with the theoretical counterparts as well. The DGS method in reflection-mode is then extended to visualize the evolution of surface slopes due to impact loading of a free-free thin sheet by a steel ball impactor.

A. Jain • C. Periasamy • H. Tippur (✉)
Department of Mechanical Engineering, Auburn University, Auburn, AL 36849, USA
e-mail: htippur@eng.auburn.edu

19.2   Experimental Details

The schematic of the experimental setup for implementing reflection-mode DGS method to measure surface slopes of a planar reflective object is shown in the Fig. 19.1.

The setup consists of a digital imaging device, a 50/50 beam splitter, a planar thin reflective specimen and a planar speckle target. The target plate is coated with alternate mists of black and white paints to create random speckles. The target plate is illuminated with the help of two LED light sources. The digital camera fitted with a long focal length lens and an extension tube is focused on the target plane via the reflective surface of the specimen and through the beam splitter. The beam splitter placed in a stage is oriented at an angle of 45° to the optical axis (or, the z-axis) of the camera.

Deflection of light rays from a reflective surface due to non-uniform deformations is caused by the out-of-plane displacements [2]. Details of the optical path between the specimen and the target plate are shown in Fig. 19.2. *For the sake of simplicity of analysis the target plate and specimen positions are changed suitably.*

Using a ray optics analysis, orthogonal surface slopes (along the *x*- and *y*-axes) due to out-of-plane deformations ($w(x, y)$) are related to angular deflections of light rays [2] as,

$$\frac{\partial w}{\partial x; y} = \frac{1}{2} \tan\left(\phi_{x;y}\right) = \frac{1}{2} \tan\left(\frac{\delta_{x;y}}{\Delta}\right) \approx \frac{1}{2}\left(\frac{\delta_{x;y}}{\Delta}\right). \tag{19.1}$$
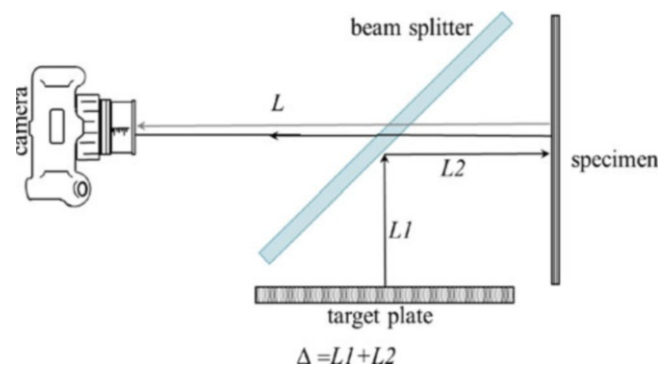
Fig. 19.1 Schematic of DGS experimental set up in reflection mode
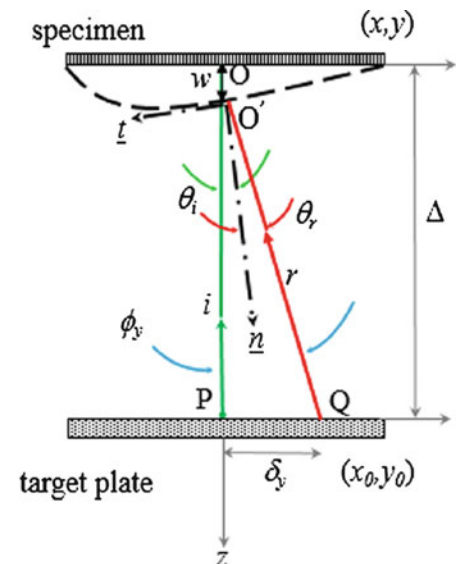
Fig. 19.2 Details of the optical path between the specimen and target in undeformed and deformed specimen configurations. $\underline{n}$ and $\underline{t}$ represent unit normal vectors to the specimen plane. $\theta_{i;r}$ denote incident (i) and reflection (r) angles of light ray due to a small deflection $w(x, y)$ of the specimen

For small angular deflections (or, $\delta_{x;y} \ll \Delta$ in Fig. 19.2) this can be written as,

$$2\frac{\partial w}{\partial x; y} \simeq \phi_{x;y} = \frac{\delta_{x;y}}{\Delta}. \tag{19.2}$$

19.3 Measurement of Surface Slopes

The feasibility of reflection-mode DGS method to measure surface slopes of planar reflective surfaces has been demonstrated by Periasamy and Tippur [2], A schematic of the experimental setup is shown in Fig. 19.3

A 280 μm thick, single face polished silicon wafer of diameter 50.8 mm was used for measurements. The unpolished face of the silicon wafer is glued to the steel washer using an epoxy adhesive. The inner and outer diameters of the washer are 25.7 mm and 76.2 mm, respectively. A central displacement was applied to the silicon wafer using a micrometer mounted on a steel plate with circular aperture placed behind the silicon wafer inside a cylindrical holder. The entire assembly is placed in front of a Nikon D100 camera such that the reflective face of the silicon wafer was facing the camera. The camera was fitted with an extension tube and 70–300 mm lens. Then a beam splitter was positioned between the wafer and the camera at an angle of 45° to the optical axis of the camera.

The speckles on the target plate were recorded with the help of the beam splitter. The target plane was at a distance of 65 mm from the reflective surface. The camera was then focused on the target plate via the reflective face of the silicon wafer. Under no-load conditions an image was recorded which served as the reference image. Then using the micrometer, central out-of-plane displacements, $w = 10$ μm, 20 μm, 30 μm was imposed on the wafer. Figure 19.4 shows the distortion of the speckle relative to the undeformed counterpart at the central region of the silicon wafer due to an applied 30 μm displacement. Deformed images captured were then correlated with the reference image using ARAMIS, to extract in-plane displacement fields δ_x and δ_y. Figure 19.5 shows the surface slopes (solid lines) calculated over the entire circular region for an applied 30 μm displacement. The contour plots are shown in steps of 5×10^{-4} rad.
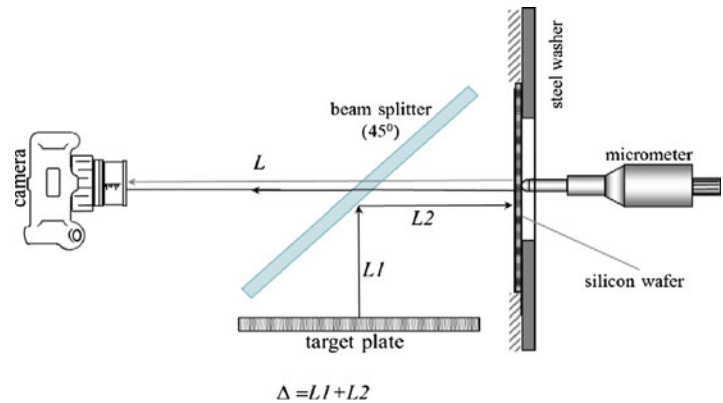
Fig. 19.3 Experimental set-up to measure surface slopes on silicon wafer subjected to central displacement using reflection mode DGS
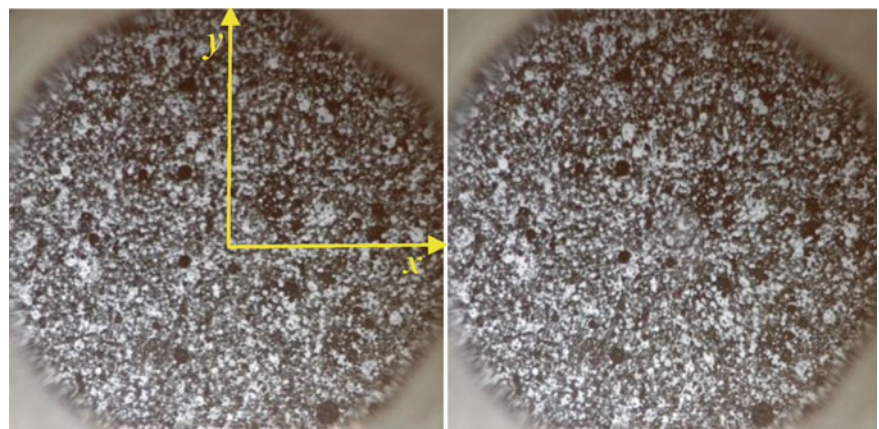
Fig. 19.4 Undeformed (*left*) and deformed (*right*) images captured off the reflective face of silicon wafer subjected to 30 μm central displacement
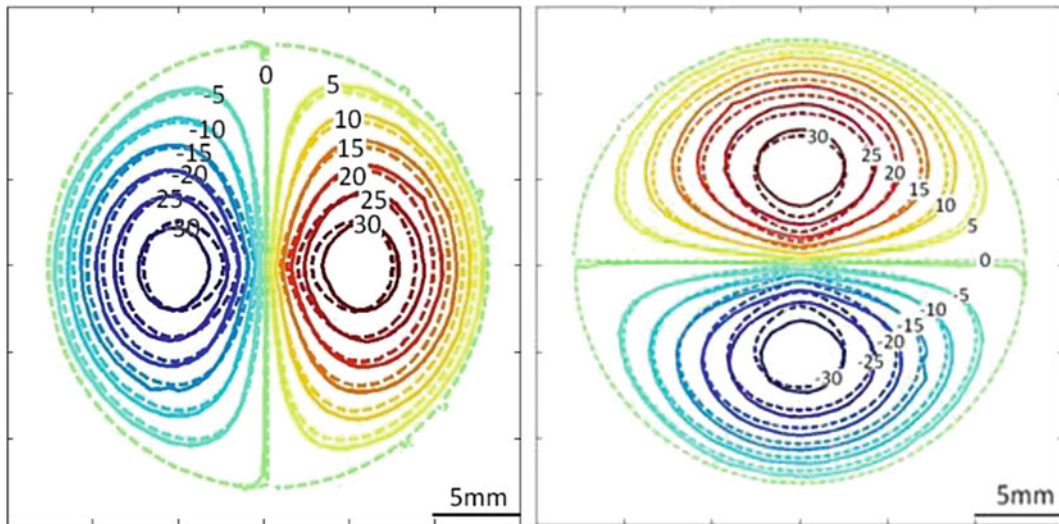
Fig. 19.5 Experimental (*solid lines*) and analytical (*broken lines*) contours of $\frac{\partial w}{\partial x}$ (*left*) and $\frac{\partial w}{\partial y}$ (*right*) comparison for a clamped silicon wafer subjected to 30 μm central displacement. Contour levels are in 1×10^{-4} rad

The surface slope measurements from reflection-mode DGS were compared with those for a clamped elastic plate subjected to a central deflection by a point force given by [1],

$$\frac{\partial w}{\partial x; y} = \frac{4(x;y)w}{R^2} \log\left(\frac{\sqrt{x^2 + y^2}}{R}\right),$$
(19.3)

where w is the out-of-plane displacement and R is the clamped radius of the wafer.

The analytical contours (broken lines) are superimposed on experimentally measured surface slopes in Fig. 19.5. Good qualitative and quantitative agreement between the two results is self-evident.

19.4   Estimation of Curvatures

The δ_x and δ_y measurements from reflection-mode DGS were differentiated once with respect to in-plane spatial coordinates x and y to estimate curvatures. Surface slopes $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ being proportional to δ_x and δ_y, the respective curvatures [2] were calculated as,

$$\frac{\partial^2 w}{\partial x^2; y^2} = \frac{1}{2} \frac{\partial}{\partial x; y}\left(\frac{\delta_{x;y}}{\Delta}\right).$$
(19.4)

The contour plots of curvature measurements (solid lines) are shown in Fig. 19.6. For comparison, the analytical curvature fields given by [1],

$$\frac{\partial^2 w}{\partial x^2; y^2} = \frac{4w}{R^2} \log\left(\frac{\sqrt{x^2 + y^2}}{R}\right) + \frac{4(x^2; y^2)w}{R^2(x^2 + y^2)}.$$
(19.5)

are superimposed on the experimental counterparts. The contour plots are shown in steps of 2×10^{-4} mm^{-1}. Again, good qualitative and quantitative agreement is evident between experimental and analytical results.
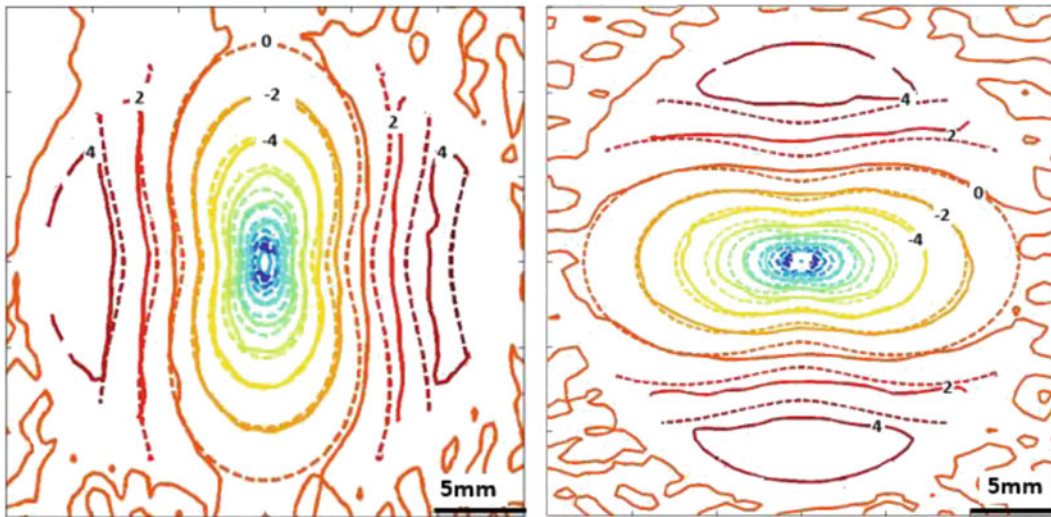
Fig. 19.6 Experimental (*solid lines*) and nalytical (*broken lines*) contours of $\frac{\partial^2 w}{\partial x^2}$ (*left*) and $\frac{\partial^2 w}{\partial y^2}$ (*right*) compared for a clamped silicon wafer subjected to 30 μm central displacement. The contour levels are in 1×10^{-4} mm^{-1}

19.5 Surface Topography

In reflection-mode DGS, δ_x and δ_y measurements being proportional to surface slopes, $\frac{\partial w}{\partial x}$. and $\frac{\partial w}{\partial y}$ can be integrated with respect to in-plane spatial coordinates to estimate topography of the deformed plate. Trapezoidal rule of integration was used to obtain the out-of-plane deformations. Integration was carried along both the in-plane directions. The resulting surface plots of topography are shown in Fig. 19.7. The results from integration show a good agreement between the maximum value at the center with the applied displacement, within an error of approximately 10 %.

Figure 19.8 shows plots of radial variation of out-of-plane displacement w along (x, y = 0) and (x = 0, y), respectively. Again from the two plots, it can be seen that the integration results agree with the applied displacements very well and the deformed shape of the plate for this axisymmetric problem is as expected. The integrated values reveal that the peak value of integrated data is slightly lower than the applied displacement in each case. This is attributed to the micrometer used to apply the central displacement might have some backlash.

19.6 Dynamic Impact of Thin Plates

Periasamy and Tippur [2] have shown that the reflection-mode DGS can be applied to monitor time-resolved surface slopes for slow speed events. Accordingly, the feasibility of the method to study deformation of plate-like structures subjected to stress- wave dominant loading conditions was attempted next.

A PMMA (poly-methyl methacrylate) disk of 3 mm thickness and 76.2 mm diameter was used for impact loading experiment. One face of the disk was made reflective by sputter coating a thin layer of Aluminum. A steel ball of diameter 25.4 mm travelling at a speed of approximately 2 m s^{-1} was used to impact the PMMA disk center. A Cordin (model 550) multi-sensor high-speed camera recording at a speed of ~80,000 frames per second was used to capture deformations due to propagating stress waves in PMMA following impact. The high-speed camera equipped with 32 CCD sensors was capable of capturing images sequentially upon an initial trigger produced by a photo-gate device. Two sets of images, one captured before the impact and the other during the impact event, were correlated using 2D DIC software ARAMIS.

The resulting time-resolved contour plots of angular deflection of light rays in the *x-z* and *y-z* planes are shown in Figs. 19.9 and 19.10, respectively. The qualitative resemblance between these contours with the quasi-static counterparts shown in Fig. 19.5 is clearly evident. As expected, the vectorial sum of the two measurements, shown in Fig. 19.10, depicts the axisymmetric nature of the deformations relative to the impact point. Increase in contour density between two successive images indicates the time- dependent deformations during impact. Analysis of these experimental results including integration of surface slopes and validation of displacements from a finite element simulation is currently being pursued (Fig. 19.11).
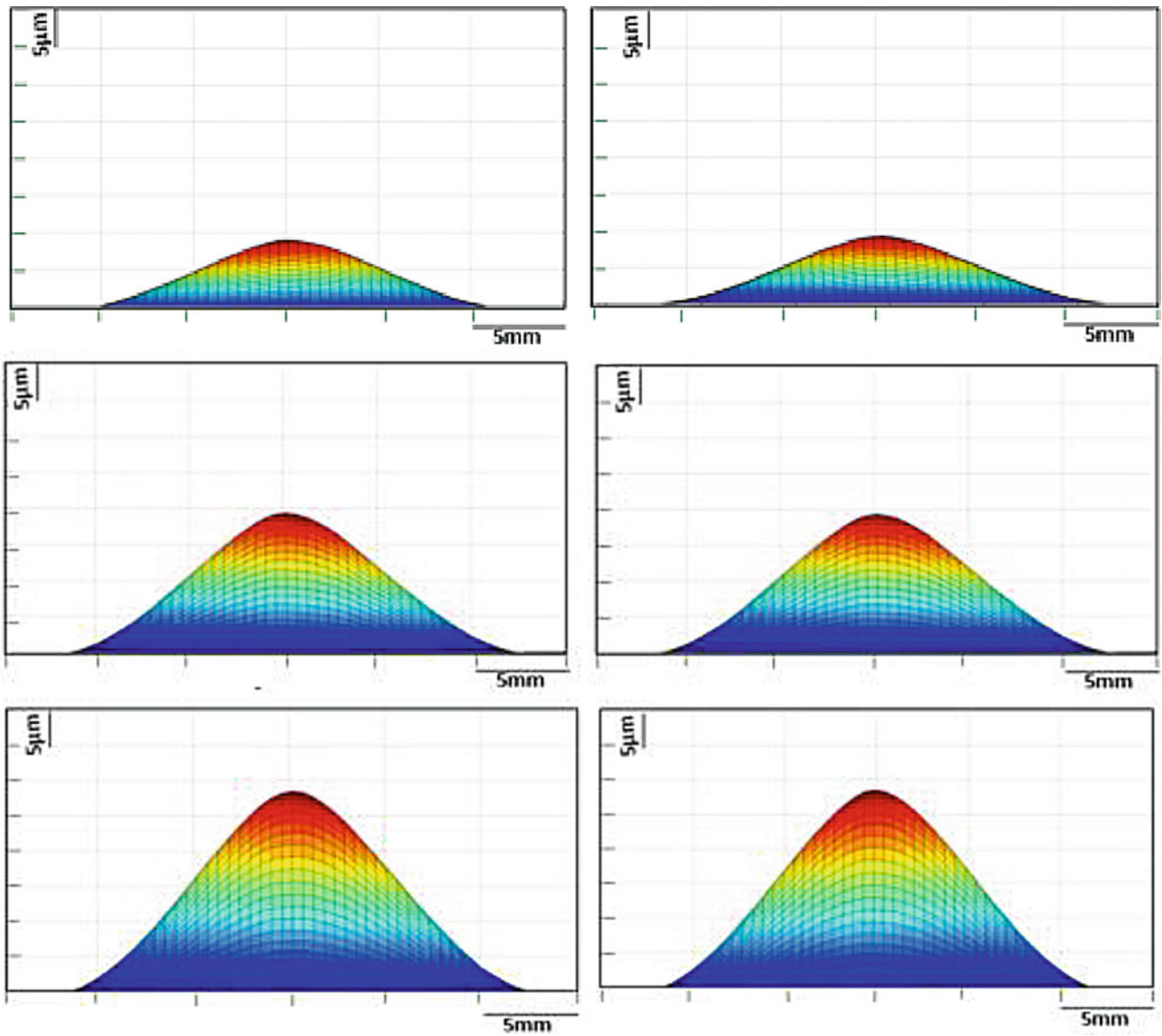
Fig. 19.7 Surface plots of integrated values of $\frac{\partial w}{\partial x}$ (*left*) and $\frac{\partial w}{\partial y}$ (*right*) for 10 μm, 20 μm, 30 μm respectively corresponding to clamped silicon wafer subjected to central out-of-plane displacement
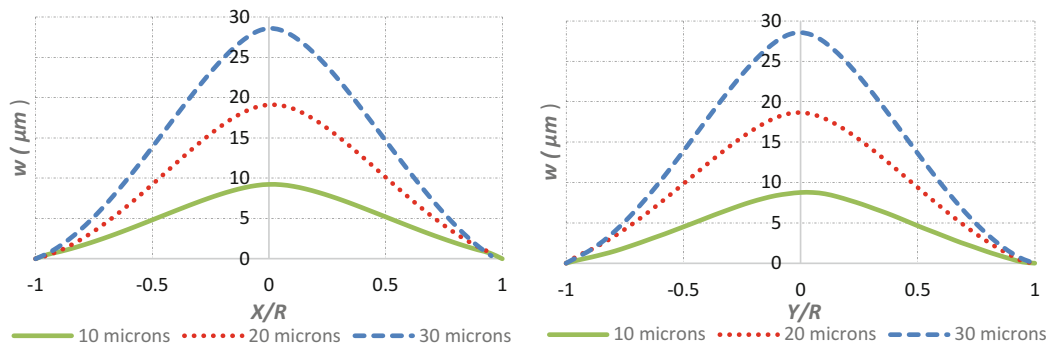
Fig. 19.8 Variation of w (out of plane displacement) along x, $y = 0$ (*left*) and $x = 0$, y (*right*) obtained by integration of $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ respectively
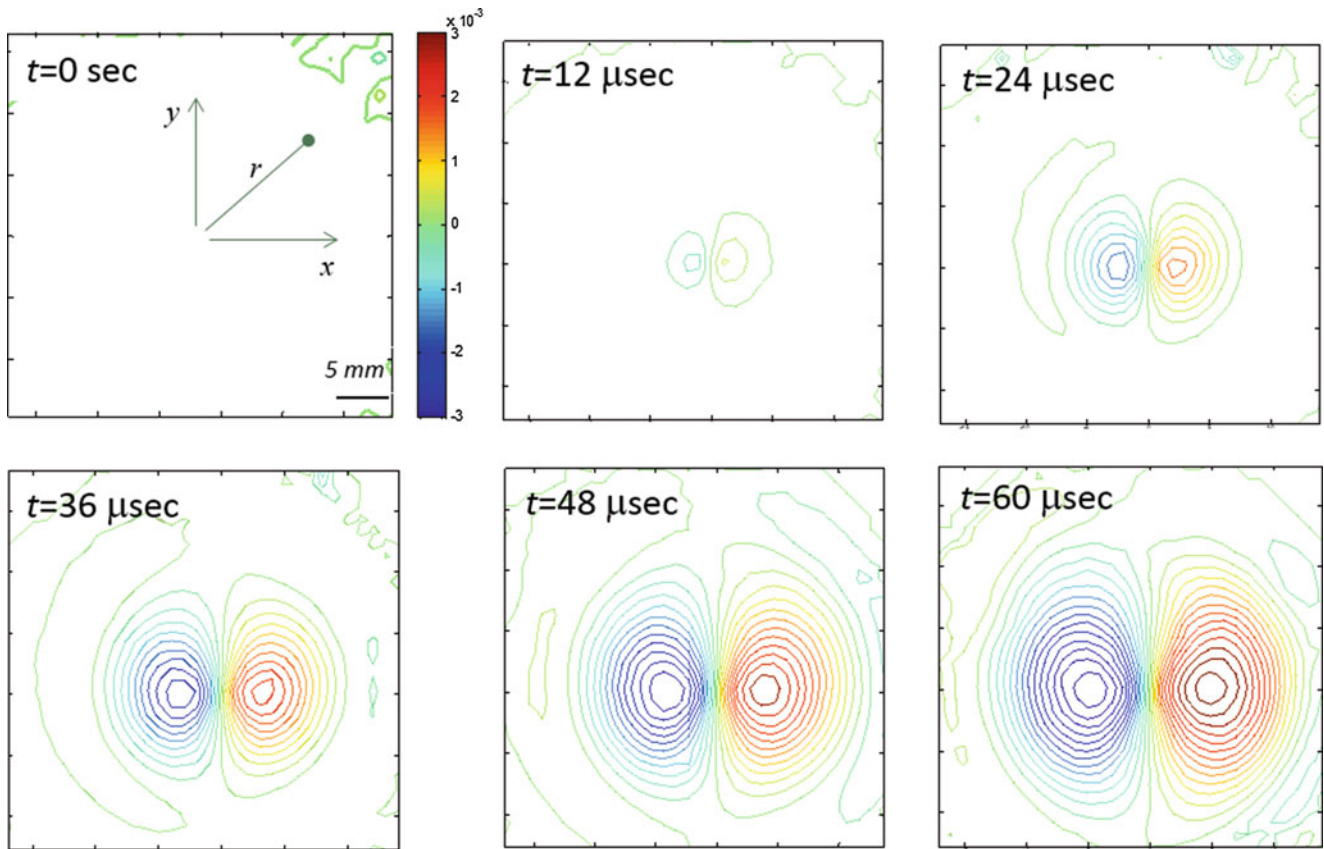
Fig. 19.9 Evolution of ϕ_x contours due to impact of a reflective PMMA disk by a steel ball at ~2 m s^{-1}. The contour levels are in 1×10^{-3} rad
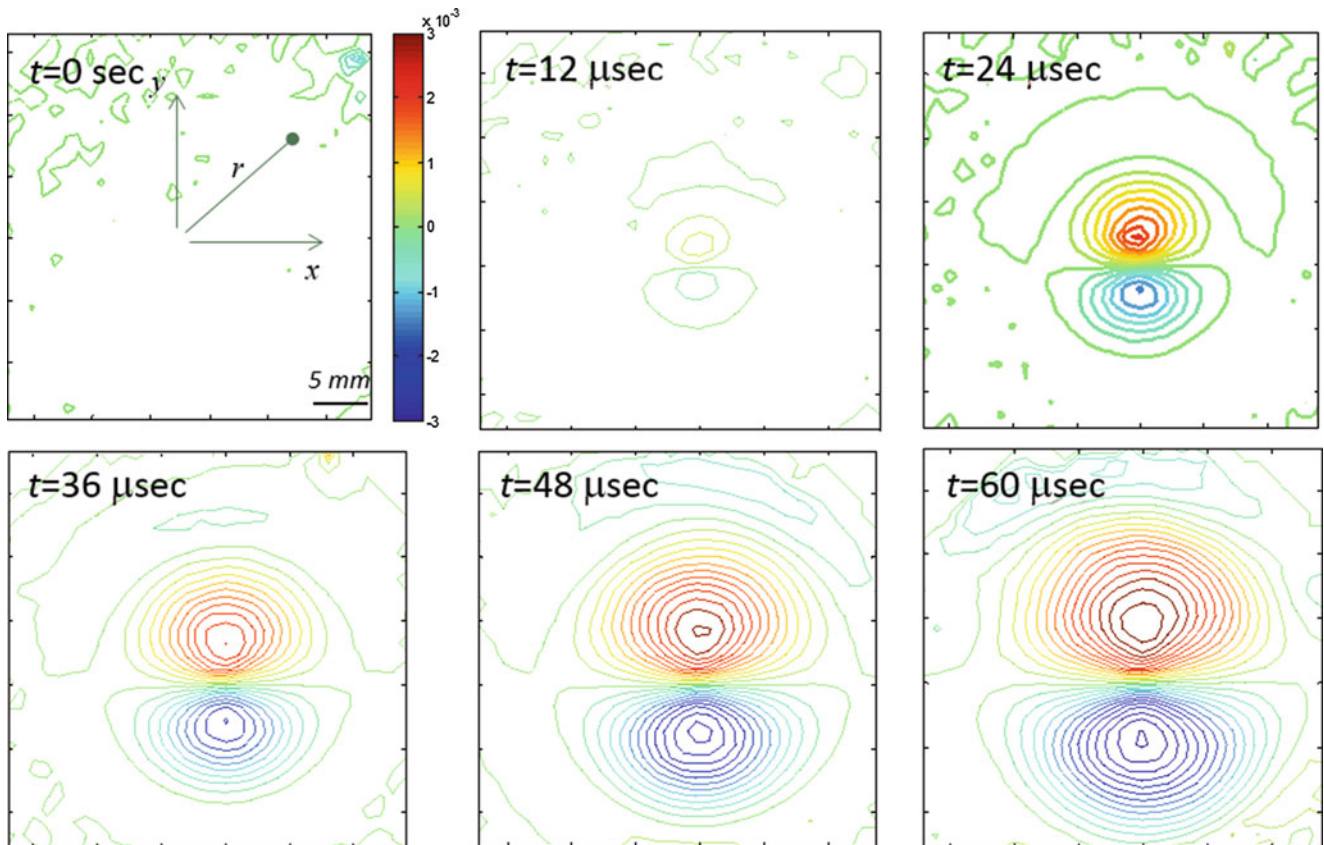
Fig. 19.10 Evolution of ϕ_y contours due to impact of a reflective PMMA disk by a steel ball at ~2 m s^{-1}. The contour levels are in 1×10^{-3} rad
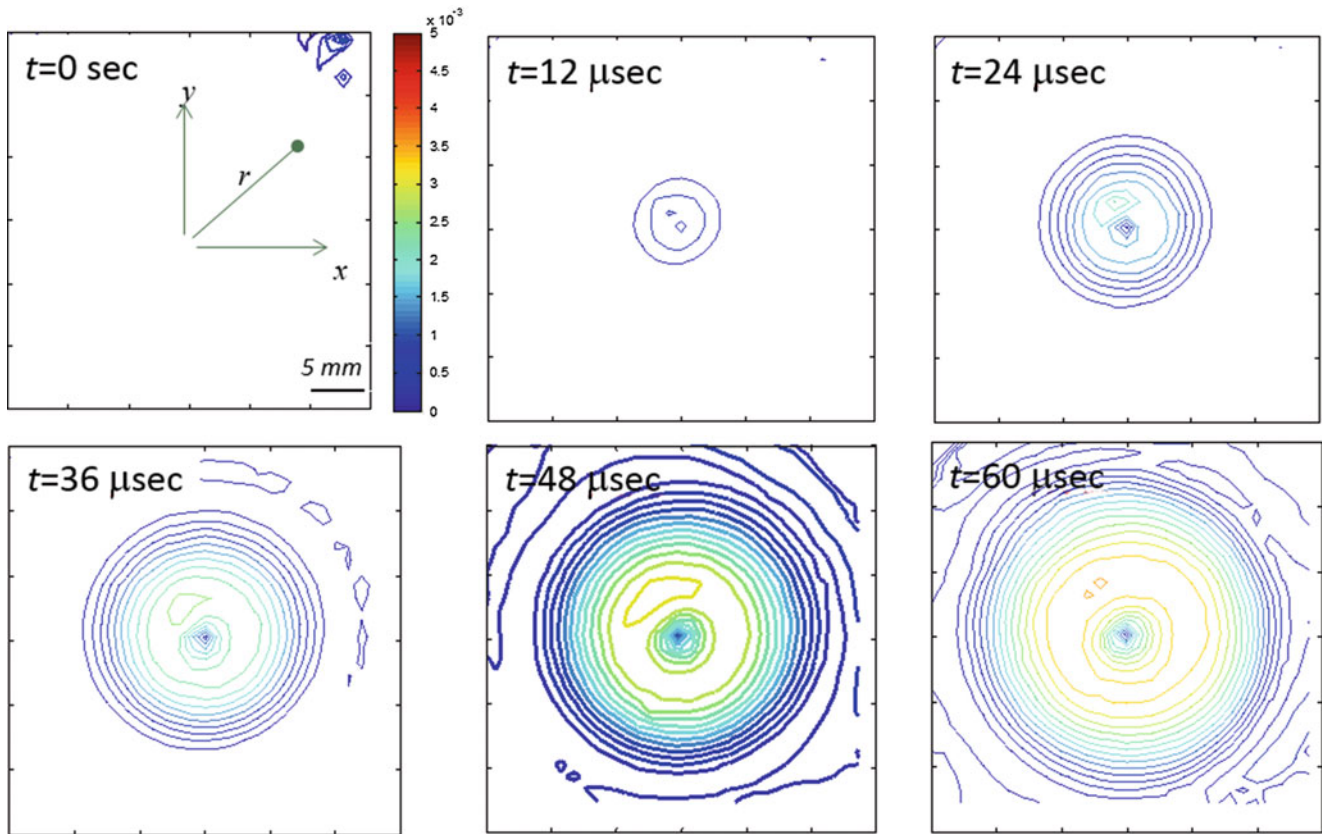
Fig. 19.11 Evolution of $\phi_r\left(\sqrt{\phi_x^2 + \phi_y^2}\right)$ contours due to impact of a reflective PMMA disk by a steel ball at ~2 m s^{-1}. The contour levels are in 1×10^{-3} rad

19.7 Summary

Reflection-mode Digital Gradient Sensing method has been extended to measure surface slopes of thin plates under static and dynamic loading conditions. The curvatures for a clamped silicon wafer subjected to static central deflection were obtained by differentiating the surface slopes and comparing with the closed form solutions successfully. Topography of the deformed wafer was also obtained by integrating the measured surface slopes. Again, a good agreement was realized between the applied displacement and the extracted values by integrating measured slopes. The ability of the method to perform time-resolved dynamic measurement of surface slopes in two orthogonal directions was demonstrated using high-speed photography used in conjunction with reflection DGS for an impact loaded reflective polymer sheet with a steel ball. The measured surface slopes show qualitative resemblance with the quasi-static counterpart and axisymmetric characteristics. Further analyses of the results are currently underway.

References

1. Tippur HV (2004) Simultaneous and real-time measurement of slope and curvature fringes in thin structures using shearing interferometry. Opt Eng 43:3014–3020
2. Periasamy C, Tippur HV (2013) A full-field reflection-mode digital gradient sensing method for measuring orthogonal slopes and curvatures of thin structures. Meas Sci Technol 24(025202):9

Chapter 20
Analysis of Linear Anisotropic Parameters by Using Hybrid Model in Mueller Optical Coherence Tomography

Chia-Chi Liao and Yu-Lung Lo

Abstract This study proposes a Mueller optical coherence tomography (OCT) to extract linear birefringence (LB), linear dichroism (LD) parameters in the optically anisotropic material. The full 4×4 Mueller matrix of the sample is able to be measured by the Mueller OCT. Hence, a hybrid model based on the Mueller matrix derived by the differential Mueller matrix formalism and both considering the forward and backward measured beams for an anisotropic sample containing LB and LD is developed to obtain corresponding anisotropic parameters in OCT system. In contrast to the conventional polarization-sensitive OCT only for LB measurement and Mueller OCT only for depolarization measurement, the proposed Mueller OCT using the hybrid model provides full range measurements of LB and LD parameters in the optical sample. Furthermore, the proposed model is insensitive to the multiplication order of the constituent basis matrices and provides stable measurements in characterizing composite anisotropic properties. The validity of the method is proved in the quarter wave-plates and testing baked polarizer. As such, inclusive of the ability to get in-depth cross-sectional images of the sample by OCT, the proposed Mueller OCT provides an ideal solution for biological and industrial application in which the precise optical properties of an anisotropic material is required.

Keywords Optical coherence tomography • Polarization-sensitive devices • Mueller matrix • Birefringence • Dichroism

20.1 Introduction

Optical coherence tomography (OCT) is a powerful technique for performing in-depth cross-sectional imaging in scattering-type media [1]. Recent enhancements to the traditional OCT structure have enabled the incorporation of polarized control into the system to evaluate the polarization effect or anisotropic properties of certain optical materials. Several researches have been proposed and are related to the use of polarization-sensitive (PS) OCT structures to establish the depth-resolved optical birefringence properties of biological tissues [2–4]. The proposed PS-OCT [2–4] combined the characteristic depth resolution of traditional OCT structures with the polarization sensitivity of scanning polarimetry based on Jones calculus. The feasibility of the proposed structure was demonstrated by measuring the phase retardation of an optically birefringent sample by detecting two orthogonal polarization components at each backscattering or reflection point. However, in general, the polarization of biological tissue is complicated and can not be assumed only by 2×2 Jones matrix. As is known, Stokes vectors and Mueller matrices provide complete representation of anisotropic properties in optical sample. As a result, A Mueller OCT system for measuring the full 4×4 Mueller matrix of biological tissue was proposed [5, 6]. By measuring the Stokes vectors of the light that is backscattered from biological tissues and calculating the Mueller matrix, a complete characterization of the optically anisotropic properties of the sample can be obtained. With the combination of Mueller matrix measurements and OCT, the Mueller matrix of a sample can be measured with in OCT resolution. However, in Ref. [5, 6], the Mueller OCT system only can obtain the depolarization of the sample and the Mueller expression of the sample but not solve the exactly parameters related to anisotropic properties. Hence, an analysis on the Mueller matrix to extract anisotropic parameters is needed.

C.-C. Liao • Y.-L. Lo (✉)
Department of Mechanical Engineering, National Cheng Kung University, No.1, Ta-Hsueh Road, Tainan 701, Taiwan ROC
e-mail: loyl@mail.ncku.edu.tw

About the study in Mueller matrix to represent the polarization effect in sample, Vitkin and his group [7–10] proposed a method using a Mueller matrix decomposition methodology for extracted the linear birefringence (LB), circular birefrinegence (CB), linear dichroism (LD) and depolarization coefficient of a complex turbid media. In their approach, Monte Carlo simulation was used to generate Mueller matrix and decomposed basis matrices and compared with the experimental results. Ossikovski [11] extended the differential matrix formalism including depolarizing media. It is noted that proposed studies [7–10] did not show enough effective parameters of scattering media. Hence, Pham and Lo [12] proposed a decoupled analytical technique for extracting nine effective parameters of anisotropic optical materials in LB, CB, LD, circular dichroism (CD), linear depolarization (L-Dep), and circular depolarization (C-Dep) properties. However, studies [7–10, 12] based on the Mueller matrix decomposition method have a restriction that LB, CB, LD, CD, L-Dep, and C-Dep should be in a restricted sequential order in analytical model. As a result, Ortega-Quijano and Arce-Diego [13, 14] proposed the differential Mueller matrices for general depolarizing media for measurements in transmission and backward direction. However, differential matrix formalism including depolarizing media has problems in limitations in the range of measurements.

In previous studies, the current group developed a PS-OCT structure for use a thermal light source to achieve an ultra-high axial resolution so as to obtain precise and simultaneous measurements of multiple optical parameters of a single-layer birefringent sample [15]. The objective of the current study is to extend this PS-OCT system to measure the Mueller matrix of an anisotropic sample and extract its optical parameters by a hybrid model. According to the Mueller matrix of the sample measured by Mueller OCT, a hybrid model derived by the differential Mueller matrix formalism and both considering the forward and backward measured beams for a composite sample containing LB and LD is developed and used to obtain optically anisotropic parameters. The proposed method can yield full range measurement of the LB and LD parameters in comparison with the differential Mueller matrix formalism. The validity of the method is proved in testing the quarter wave-plates with only LB property and the baked polarizer with LB and LD properties. The new algorithm introduced here has several advantages not only in solving anisotropic parameters in one hybrid sample with extended range measurements but also in obtaining cross-sectional measurement with in OCT resolution.

20.2 Principle of OCT

Figure 20.1 presents a schematic illustration of the proposed Mueller OCT structure. As shown, the structure is similar to that of a traditional OCT system other than the use of a thermal light source to enhance the axial resolution and taking account of polarization effects to enable the measurement of the multiple optically anisotropic parameters of the sample.

As shown in Fig. 20.1, the beam from the thermal light source is split by a non-polarizing beam splitter (NPBS) into two beams, namely a reference beam and a measurement beam. The reference beam passes through a variable wave plate and is reflected by a mirror positioned on a scanning stage to carry out path-length scanning, while the measurement beam passes through a glass plate designed to compensate for the dispersion effect and then is incident upon the sample. During the measurement process, the locations of the low-coherence interferometric signals are acquired at Detector 1 as the scanning stage is driven at a constant velocity. Detector 1 observes the interferometric signal induced by the beams from the sample arm and the reference arm as:

$$I = \frac{I_0}{2} \sqrt{R_s(d)} \exp\left[-\left(\frac{\Delta L \times 2\sqrt{\ln 2}}{l_c}\right)^2\right] \cos(2k\Delta L) \tag{20.1}$$

where d is the penetration depth; $R_s(d)$ is the reflectivity at depth d; $\Delta L = L_r - L_s$ is the path-length difference of the two arms, in which L_s is the path length in the sample arm and includes the optical path in the sample, i.e. $L_s = L_t + nd$ (Optical path in the sample); and l_c is the coherence length of the thermal light source. A fixed mirror is inserted to produce a reference signal, and by using the signal, the thickness and mean refractive index of the sample can be extracted by a simple method proposed by the current group [15]. It is shown that the interferometric signal in OCT can reveal the location of the reflective interface of the sample to produce a cross-sectional image related to the sample.

In measuring the full 4×4 Mueller matrix of the sample, it is assumed that the polarization state of the light incident upon the sample and reflected from the reference arm should be controlled. In order to calculate the Mueller matrix of the sample, four different incident polarization states, namely H (horizontal linear polarization), V (vertical linear polarization), P (45° linear polarization), and R (right-circular polarization) are achieved by rotating the quarter-wave plate and the polarizer.
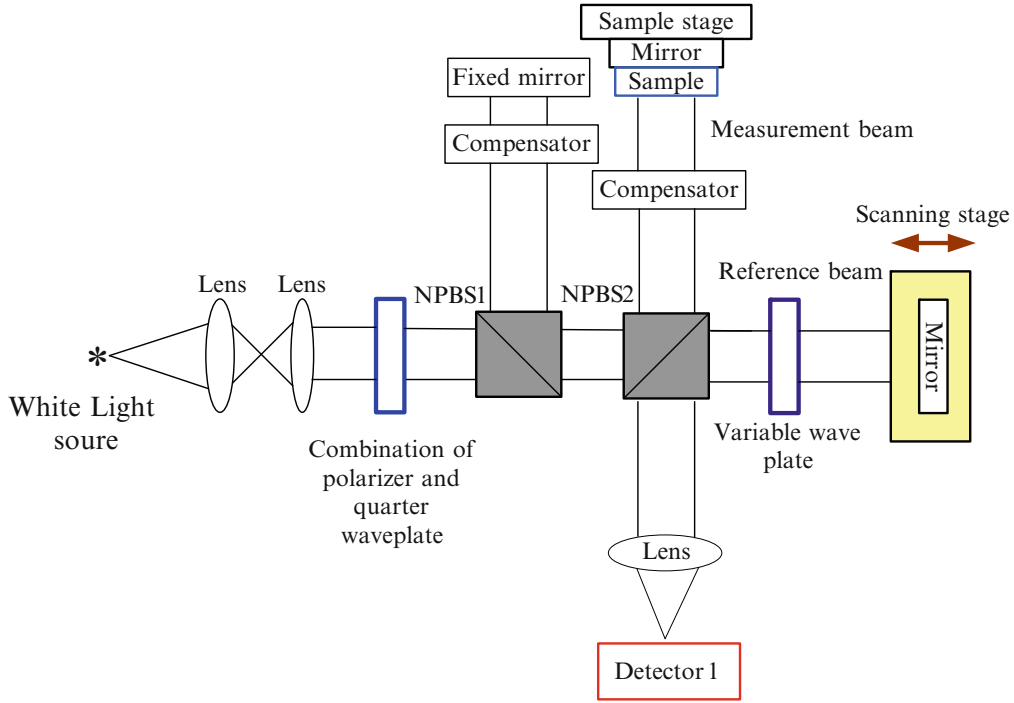
Fig. 20.1 Schematic illustration of proposed Mueller OCT system

Also, the variable wave plate in the reference arm is adjusted to change the reference beam with polarization states, H, V, P, and R sequentially for each four incident polarization. The light intensities of the interferometric signals are measured in the 16 combination of the polarization states of the incident beam and reference beam. When the incident beam is reflected from the sample and brings the optical state of the sample, the reference beam is used to convert the signal into Mueller parameters. Totally, 16 polarization signals are processed to calculate the 16 elements in Mueller matrix by following formula: [5]

$$
\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix}
$$

$$
= \begin{bmatrix} HH + HV + VH + VV & HH + HV - VH - VV & 2PH + 2PV - M_{11} & 2RH + 2RV - M_{11} \\ HH - HV + VH - VV & HH - HV - VH + VV & 2PH - 2PV - M_{21} & 2RH - 2RV - M_{21} \\ 2HP + 2VP - M_{11} & 2HP - 2VP - M_{12} & 4PP - 2PH - 2PV - M_{31} & 4RP - 2RH - 2RV - M_{31} \\ 2HR + 2VR - M_{11} & 2HR - 2VR - M_{12} & 4PR - 2PH - 2PV - M_{41} & 4RR - 2RH - 2RV - M_{41} \end{bmatrix}
$$

$$(20.2)$$

where M_{ij} represents the Mueller matrix element of the ith row and the jth column. The symbols consisting of double polarization states represent an OCT measurement, with the incident polarization state denoted by the left letter and the reference polarization state denoted by the right letter. For example, HV refers to an interferometric signal acquired by the horizontal linear polarized incident beam and vertical linear polarized reference beam.

In developing the optically equivalent model of the anisotropic material containing the linear birefringence and dichroism in OCT, influences of the forward measured beam, the mirror and the backward measured beam as shown in Fig. 20.2 are considered.

The macroscopic Mueller matrix of the sample containing LB and LD can be calculated from differential Mueller matrices by inverse differential calculation [16, 17]. In the OCT system, however, the measured beam will pass through the sample for the forward and backward directions. It means both the forward and backward passing through the sample should be considered for the Mueller matrix of the sample measured by OCT. Hence, a hybrid model for the OCT is proposed.

Fig. 20.2 Schematic diagram of the model of the measurement on the sample containing LB/LD in OCT system
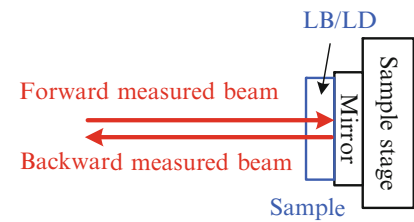
The macroscopic Mueller matrix of the LB/LD sample for the forward measured beam, $\mathbf{M_L}(\alpha, \beta, \theta_d, D)$ is derived from the differential matrix by inversing differential calculation, where α, β, θ_d, and D represent the orientation angle of LB, phase retardation of LB, the orientation angle of LD, and linear dichroism of LD, respectively. In contrast, the macroscopic Mueller matrix for the backward measured beam should be obtained in the assumption that $\mathbf{M_L}(-\alpha, \beta, -\theta_d, D)$. Thus, the equivalent Mueller matrix of the sample in OCT can be expressed as

$$\mathbf{M_{L,\,OCT}} = \mathbf{M_{L,\,Backward}}\mathbf{M_{Mirror}}\mathbf{M_{L,\,Forward}} \tag{20.3}$$

where $\mathbf{M_{L,Forward}}$ represent the macroscopic matrix and derived from the differential matrix containing the LB and LD properties by inverse differential calculation. $\mathbf{M_{L,Backward}}$ is the LB/LD macroscopic matrix for the backward measured beam. $\mathbf{M_{Mirror}}$ means the Mueller matrix of the mirror. By using Eq. (20.3), the Mueller matrix of the sample in OCT system can be constituted and it is noted that the macroscopic Mueller matrix $\mathbf{M_{L,OCT}}$ is highly complicated. For the complicated formula, a Genetic Algorithm [18–20] can be employed to obtain values of LB and LD parameters. By using the Mueller matrix measured from the OCT in Eq. (20.2) as target function, unknown parameters in $\mathbf{M_{L,OCT}}$ given in Eq. (20.3) are extracted in a GA. For the hybrid model in OCT, parameters related to the linear birefringence (α, β) and linear dichroism (θ_d, D) can be obtained in the full range as: $0 < \alpha < 180°, 0 < \beta < 360°, 0 < \theta_d < 180°, 0 < D < 1$.

20.3 Simulation Results

In this section, the ability of the proposed Mueller OCT analyzed by the hybrid model to extract corresponding anisotropic optical parameters of interest over the measurement ranges is confirmed. Thereafter, simulations are performed to evaluate the accuracy of the results obtained from the hybrid model for samples with various parameters. In performing the analytical simulations, the measured Mueller matrix is calculated via a process of a hybrid model with corresponding parameters for a hypothetical sample. The theoretical input values of corresponding parameters are inserted into measured Mueller matrix, and then the hybrid model in GAs is used to extract the corresponding parameters. Finally, extracted values of the corresponding optical parameters are compared with the theoretical values in order to quantify the accuracy of the corresponding extraction method. (Note that for each of the extracted parameters, the theoretical input value is varied over the full range, i.e., α: 0 ~ 180°, β: 0 ~ 360°, θ_d: 0 ~ 180°, and D: 0 ~ 1).

The ability of the proposed hybrid model to determine the optical parameters of LB/LD composite samples was evaluated by extracting the values of α, β, θ_d, and D and then comparing the extracted values with the known input values. The input parameters are specified as follows: $\alpha = 30°$, $\beta = 60°$, $\theta_d = 35°$, $D = 0.5$. Figure 20.3a presents the extracted values of α for input values of α over the full range (i.e., α: 0 ~ 180°) given parameter settings of $\beta = 60°$, $\theta_d = 35°$, $D = 0.4$. Figure 20.3b–d present the equivalent results for parameters β, θ_d and D, respectively. In Fig. 20.3, extracted α, β, θ_d, and D of the LB/LD sample are compared to the theoretical input values. In general, the results presented in Fig. 20.3 show that the proposed hybrid model enables all four parameter of the LB/LD composite sample to be measured over the full range.

20.4 Experimental Results

The whole system according to our presented structure in Fig. 20.1 is established. The light source for the Mueller OCT structure was provided by a halogen lamp (Model R150-BM1, TECHNIQUIP Co., USA) with a color temperature of 3,250 K. A halogen illumination source was chosen quite specifically here since it has a broad spectral bandwidth and therefore provides an enhanced axial resolution. The light from the halogen lamp was filtered through a light guide such that only the visible wavelengths were passed. Hence, the light entering the Mueller OCT structure had a spectral bandwidth of
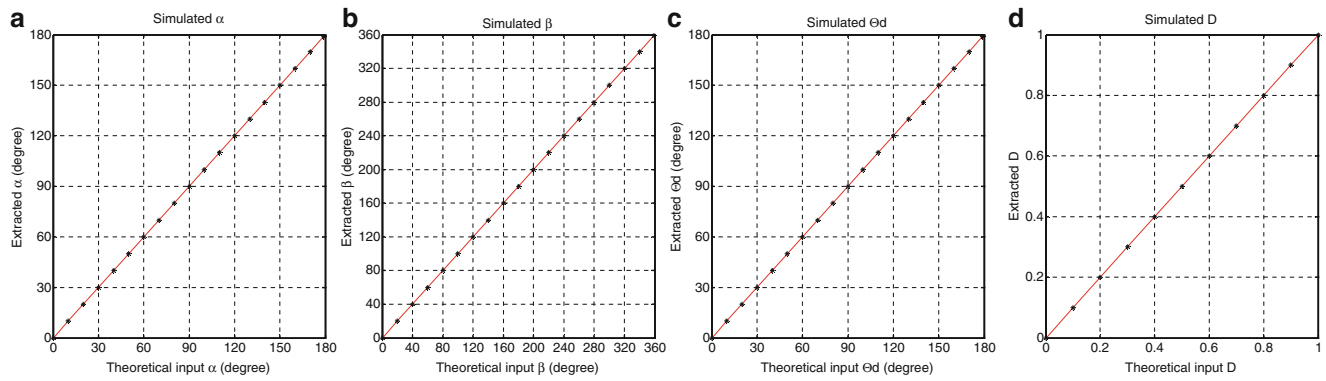
Fig. 20.3 Simulated results for α, β, θ_d, and D of a LB/LD sample solved by hybrid model in GAs. Set theoretical input: (**a**) α: 0 ~ 180°, $\beta = 60°$, $\theta_d = 35°$, D = 0.5; (**b**) β : 0 ~ 360°, $\alpha = 30°$, $\theta_d = 35°$, D = 0.5; (**c**) θ_d: 0 ~ 180°, $\alpha = 30°$, $\beta = 60°$, D = 0.5; (**d**) D: 0 ~ 1, $\alpha = 30°$, $\beta = 60°$, $\theta_d = 35°$

$\Delta\lambda = 210$ *nm* and a central wavelength of $\lambda_0 = 603$ *nm*. The photo-detector used in the Mueller OCT structure were acquired from New Focus Corporation, USA, (Model 2001) and had an operating wavelength range of 300–1,050 nm, a maximum conversion gain of 9.4×106 V/W. The scanning stage was acquired from Sigma Koki Co, Ltd., Japan (Model SGSP 20–85). During the scanning experiments, the stage was moved with a constant velocity of 1,000 μm/s to induce Doppler shift frequency for modulation. In this research, the velocity of the scanning stage is significant to determine the location from the oscilloscope, not the stage controller, and hence the velocity has to be defined carefully. The scanning stage was controlled using a Mark-204MS stage controller (Sigma Koki Co., Ltd; Japan) via the GPIP interface of a PC. However, this study will concern about the anisotropic properties of the sample, the thickness and the mean refractive index of the sample are not discussed in experiments. Additionally, a variable wave plates (New Focus Corporation, USA, Model 5540) is used to change the polarization of the beam in reference arm.

When using a broadband light source in a Mueller OCT structure, the reference arm and the sample arm must be perfectly symmetrical in order to prevent the resolution from being degraded by dispersion mismatches. The various polarization of the beam in reference arm should be applied for the Mueller matrix measurement. Therefore, the variable-wave plate at the reference arm was adjusted to achieve the desired polarization states in the reference beam. After setting the variable wave plate, a polarizer and quarter wave plate can be put in front of the detector to inspect the polarization state is right. Finally, the interferometric signals observed by the photo-detector were recorded by an oscilloscope (LeCroy Corporation; USA; Model: Wave Runner 6050) with a maximum sampling rate of 5 GS/s and a bandwidth of 500 MHz. The validity of the proposed measurement method is evaluated using two optical samples, namely (20.1) a quarter wave-plate; and (20.2) a baked polarizer.

The performance of the Mueller OCT structure was investigated using a zero-order quarter wave plate (Thorlabs, AQWP05-630) served as a linear birefringent sample. The effect of dispersion has a more significant impact when measuring the phase retardation of a birefringent sample than when measuring its thickness and mean refractive index since dispersion readily affects the amplitude of the interferometric signals and therefore results in a greater error when extracting the apparent phase retardation. Therefore, the dispersion effect must be suitably compensated. In this measurement, it is confirmed by the experimental inspection that the D263T glass plate can be used for compensation purposes.

The main experimental steps are accomplished according to the description in Sect. 20.2. The intensity of 16 combinations of the polarization states at different degree were measured to reproduce the Mueller matrix of the measured sample by using Eq. (20.2). However, it should be noticed that because the calibration in experiments is difficult in successive measurements, experimental results just are shown in the orientation angle for the range: 0 ~ 20°. Finally, the Mueller matrix of the sample can be obtained and used in GAs to extract the principal angle of LB, α, the phase retardation, β, the principal angle of LD, θ_d, and the linear dichroism, *D*. Figure 20.4 displays the extracted results in α, β, θ_d, and *D*. The measured optical axis orientation angle and phase retardation of LB have good agreements with the quarter wave plate, while the measured LD approximates zero and orientation angle of LD are random since the quarter wave plate theoretically does not contain the linear dichroism property.

The study achieves a Mueller OCT system based on Mueller matrix method to measure the linear birefringence material and has got initial results. Errors in experiments can be explained in two reasons. Firstly, the measured zero order quarter wave plate is made up by two half wave plates to produce the achromatism in 400–700 nm. Because the composite structure
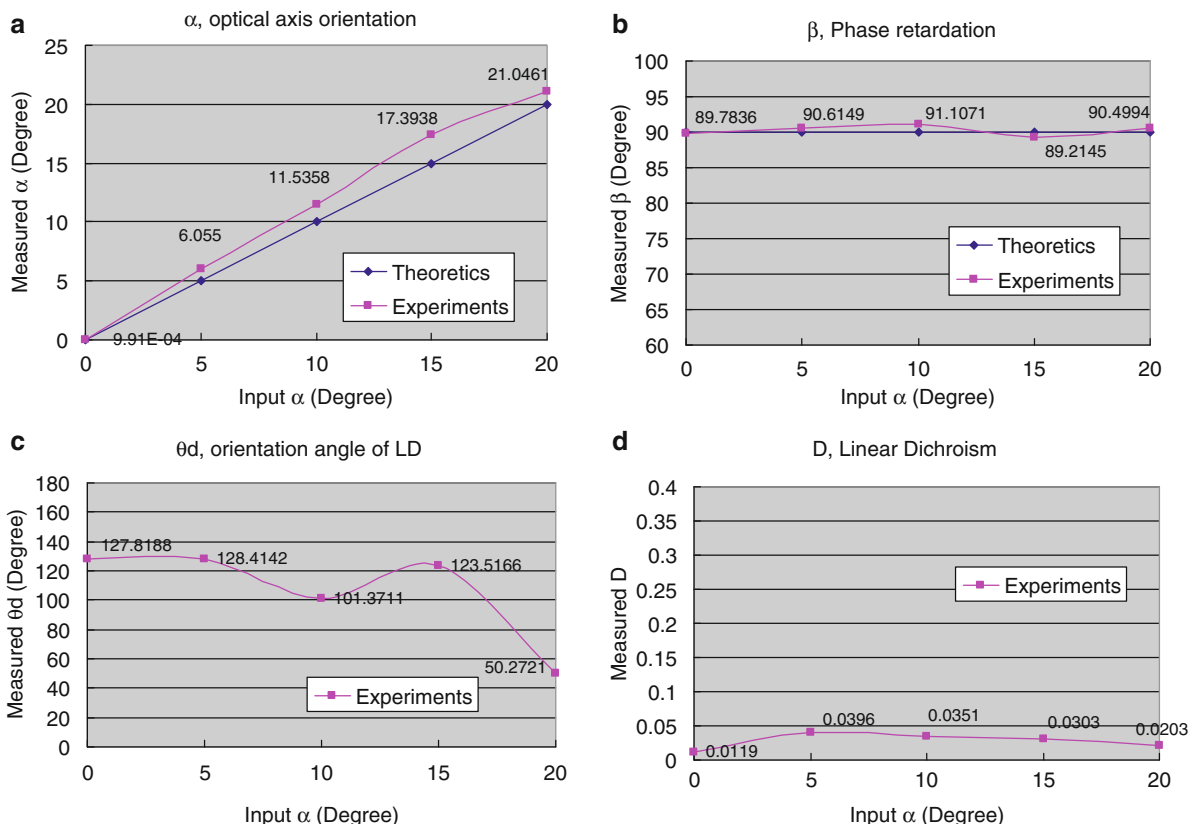
Fig. 20.4 Comparison of theoretical and experimental results for (**a**) optical axis orientation angle of LB; (**b**) phase retardation of LB; (**c**) orientation angle of LD; (**d**) linear dichroism

in the sample, the spatial coherence of the interferometric signal may be destroyed when the sample is rotated at specific angles. In fact, the intensity of light passing through the zero-order quarter wave plate is various while in different principal angles despite the incident light is unpolarized. It also can explain why only the range in $0 \sim 20°$ of the orientation angle of LB is measured. The composite structure of the sample will induce some difficulties in the calibration in successive measurements. Secondly, the dispersion effect has an influence on the measured intensity and the width of interferometric signals, and the use of this light source requires an effective dispersion compensation scheme. So, we may not achieve enough dispersion compensation in this study.

Additionally, a polymer polarizer (LLC2-82-18S, OPTIMAX Co.) baked in an oven at $150 °C$ for 100 min are served as a testing sample for the LB and LD properties. Due to the prolonged exposure of the polarizer to a high-temperature environment, the input light leaks through one of the LD axes. Thus, as expected, the measured value of linear dichroism has a value of 0.9864 closed to 1 and the measured value of the phase retardance is found to be $19.2262°$. In comparison with the previous study [12], the result is slight different because the central wavelength of the light in the Mueller OCT system is different from that in Stokes polarimeter. However, it also can confirm that the proposed analytical model enables the parameters of hybrid samples with both LB and LD properties to be accurately determined.

20.5 Conclusions and Discussions

This study has developed a Mueller OCT to measure the Mueller matrix expression of an anisotropic sample with a high depth resolution. According to the measured Mueller matrix, an analytical model based on use of the differential calculation method and both considering the forward and backward measured beams to extract effective parameters of a composite sample containing LB and LD properties measured by the Mueller OCT is proposed. In the calculation by using a genetic algorithm, the model is suitable for the LB/LD composite sample and able to provide the measurement over the full range.

The validity of this model to extract LB and LD parameters in the full range is confirmed by the simulation. The experiments on the quarter wave plate for only LB property and the baked polarizer for both LB and LD properties have demonstrated the proposed method. As a result, the proposed approach has significant potential for applications in the cross-sectional image of the anisotropic materials. For example, it can be used in collagen and muscle structure characterization (based on LB measurement only) without using compensation technique or pretreatment. This is necessary to provide the data to identify the biological or medical materials. Hence, the new algorithm introduced here provides not only the measurement in anisotropic parameters of a composite sample containing LB and LD properties simultaneously with full range measurement but also the cross-sectional measurement with in OCT resolution.

References

1. Huang HD, Swanson EA, Lin CP, Schuman JS, Sinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA, Fujimoto JG (1991) Optical coherence tomography. Science 254:1178–1181
2. Hee MR, Huang D, Swanson EA, Fujimoto JG (1992) Polarization-sensitive low-coherence reflectometer for birefringence characterization and ranging. J Opt Soc Am B 9:903–908
3. de Boer JF, Milner TE, van Gemert MJC, Nelson JS (1997) Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence tomography. Opt Lett 22:934–936
4. Hitzenberger CK, Goetzinger E, Sticker M, Pircher M, Fercher AF (2001) Measurement and imaging of birefringence and optic axis orientation by phase resolved polarization sensitive optical coherence tomography. Opt Express 9:780–790
5. Yao G, Wang LV (1999) Two-dimensional depth-resolved Mueller matrix characterization of biological tissue by optical coherence tomography. Opt Lett 24:537–539
6. Jiao S, Yao G, Wang LV (2000) Depth-resolved two-dimensional stokes vectors of backscattered light and Mueller matrices of biological tissue measured with optical coherence tomography. Appl Opt 39:6318–6324
7. Ghosh N, Wood MFG, Li SH, Weisel RD, Wilson BC, Li RK, Vitkin IA (2009) Mueller matrix decomposition for polarized light assessment of biological tissues. J Biophoton 2:145–156
8. Ghosh N, Wood MFG, Vitkin IA (2009) Polarimetry in turbid, birefringent, optically active media: a Monte Carlo study of Mueller matrix decomposition in the backscattering geometry. J Appl Phys 105:102023-1–102023-8
9. Guo X, Wood MFG, Ghosh N, Vitkin IA (2010) Depolarization of light in turbid media: a scattering event resolved Monte Carlo study. Appl Opt 49:153–162
10. Ghosh N, Vitkin IA (2011) Tissue polarimetry: concepts, challenges, applications, and outlook. J Biomed Opt 16:110801-1–110801–29
11. Ossikovski R (2011) Differential matrix formalism for depolarizing anisotropic media. Opt Lett 36:2330–2332
12. Pham TTH, Lo YL (2012) Extraction of effective parameters of anisotropic optical materials using decoupled analytical method. J Biomed Opt 17:025006-1–025006-17
13. Ortega-Quijano N, Arce-Diego JL (2011) Depolarizing differential Mueller matrices. Opt Lett 36:2429–2431
14. Ortega-Quijano N, Arce-Diego JL (2011) Mueller matrix differential decomposition for direction reversal: application to samples measured in reflection and backscattering. Opt Express 19:14348–14353
15. Liao CC, Lo YL, Yeh CY (2009) Measurements in multiple optical parameters of the birefrigent sample using polarization-sensitive optical coherence tomography. IEEE J Lightwave Technol 27:483–493
16. Azzam RMA (1978) Propagation of partially polarized light through anisotropic media with or without depolarization: a differential 4×4 matrix calculus. J Opt Soc Am 68:1756–1767
17. Liao CC, Lo YL (2013) Extraction of anisotropic parameters of turbid media using hybrid model comprising differential- and decomposition-based Mueller matrices. Opt Express 21:16831–16853
18. Michalewicz Z (1994) Genetic algorithm + data structure = evolution programs. Springer, New York
19. Cheng HC, Lo YL (2005) The synthesis of multiple parameters of arbitrary FBGs via a genetic algorithm, and two thermally modulated intensity. J Lightwave Technol 23:2158–2168
20. Yu TC, Lo YL (2007) A novel heterodyne polarimeter for the multiple-parameter measurements of twisted nematic liquid crystal cell using a genetic algorithm approach. IEEE J Lightwave Technol 25:946–951

Chapter 21
Characterization of Time-Dependent Mechanical Behaviors of Dental Composites by DIC

T.Y. Chen, C.L. Hsu, and S.F. Chuang

Abstract Dental resin composites are widely applied in dental restoration. These resin composites have some problems due to shrinkage occurred. Since shrinkage stress could be presented by resin matrix polymerization as tensile stress and causes structure failure. In this study, four types of dental composites (Z350 flow, universal composite Z250, packable composite P60, and low-shrinkage composite LS) were chosen to investigate their tensile creep and recovery behavior. The specimens were divided into two groups: (1) after curing about 5 min, and (2) stored in artificial saliva for 30 days. With force control, the specimen image under uni-axial tension was recorded and analyzed by using digital image correlation (DIC) method. Four-parameter fluid model is found applicable to fit the experimental obtained strain. The results showed that the longitudinal strain of 5-min case has higher creep and residual strain than that of 30 days. The creep of LS was the smallest in four resin composites. The Poisson's ratio of 5-min case was increased about 0.5 with time except for LS. However, the Poisson's ratio of 30-day cases varied up and down insignificantly with time.

Keywords Dental resin composites • Shrinkage • Creep and recovery behavior • Digital image correlation • Poisson's ratio

21.1 Introduction

Today, dental resin composites are widely applied in dental restoration such as bonding the porcelain and metal crown [1], bonding the core and post [2], as pit and fissure sealants [3], and direct cavities restoration for aesthetic condition. But, these resin composites had many worrying problems such as post-operative sensitivity, secondary caries, interface debonding, and remaining tooth fracture [4]. These problems are mainly attributed to shrinkage occurred by resin matrix polymerization, and researchers found many methods, higher molecular weight, higher filler content, ring-opening polymerization to reduce shrinkage for wrestling with those problems [5]. However, these problems still exist in clinic case. On the other hand, some studies focus on the time-dependent material properties and long-term stability such as: creep [6], wear resistance [7], and mechanical modulus [8]. These tests were used to estimate the potential risk in different type resin composites.

Since shrinkage stress could be presented as tensile stress to cause structure failure. The study of tensile creep behaviors of dental composites is needed. In this study, a simple rectangular beam model subjected to tensile force was used to study the time dependent mechanical behavior of dental composite resins. Digital image correlation (DIC) method [9, 10] was employed to analyze the deformation. Finally the mechanical behavior of dental composite resins was fit by a generally physical model, and characterized.

T.Y. Chen (✉) • C.L. Hsu
Department of Mechanical Engineering, National Cheng-Kung University, Tainan 70101, Taiwan ROC
e-mail: ctyf@mail.ncku.edu.tw

S.F. Chuang
Institute of Oral Medicine, National Cheng-Kung University, Tainan 70101, Taiwan ROC

21.2 Materials and Methods

Four types of dental composites as listed in Table 21.1, flowable composite, Z350 flow; universal composite, Z250; packable composite, P60; and low-shrinkage composite LS were chosen to study. Rectangular beam specimens were prepared using a plastics mode with internal dimensions 30 mm × 5 mm × 1.5 mm. Z350 flow, P60, Z250 and LS specimens were cured at room temperature by using a light curing unit with a tip of 8 mm diameter (Optilux 501, Kerr, Orange Co., USA). The specimens were irradiated 10 s (20-s for LS) by eight times along the length of the upper and lower surface. The intensity of irradiance is 650 mW/cm^2.

Those specimens divided into two groups: (1) after curing about 5 min ($S_{5\,m}$), this group is used to simulate shrinkage stress effect of post curing, (2) after the specimens stored in artificial saliva at 37 °C in incubator for 30 days ($S_{30\text{-day}}$), this group is used to simulate long-term stability of the mechanical behavior of dental composites. Composition of the artificial saliva is listed in Table 21.2

A photograph of the experimental system is shown in Fig. 21.1. The specimen is fixed on a jig by an adhesive (mxbon® gel 909, Cartell Chemical Co., Taiwan). Creep and recovery test were implemented by using a universal testing machine (AG-I 5kN, Shimadzu, Kyoto, Japan) under the force control. In order to enhance the correlation, surface of specimens were sprayed paint to have broader gray level distribution. The deformed specimen images were recorded by an image acquisition system assembled with a Charge-coupled device (CCD) camera (MTV-12V1E, Mintron Co., Taiwan), an optical microscope (Zoom microscope ML-Z07545D, Moritex Inc., Japan) a SCSI interfaced image acquisition card and a personal computer. The image capture rate was 1 frame/5 s for the first 10 min and 1 frame/30s for the rest of time. The stored mages were processed by using the commercial DIC software VIC-2D 2009 (Correlated solutions, West, Columbia, SC, USA) to obtain longitudinal strains and Poisson's ratios of a 5 × 5 mm^2 area near the center of specimen. The resolution is 8 μm per pixel.

Table 21.1 Four dental composites (3 M ESPE) tested

ID	Material	Matrix monomer	Composite type	Filler content (Vol%)
Z350 flow	Filtek™ Z350	BisGMA, TEGDMA	Nanofill	60
Z250	Filtek™ Z250	UDMA, BisEMA, BisGMA, TEGDMA	Universal	60
P60	Filtek™ P60	UDMA, TEGDMA BisEMA, BisGMA,	Packable	61
LS	Filtek™ LS	Silorane	Low shrinkage	55

Table 21.2 Composition of the artificial saliva

Composition of 1,000 ml	
Potassium chloride	1.20 g
Sodium chloride	0.84 g
Potassium phosphate	0.26 g
Calcium-chloride-dihydrate	0.14 g
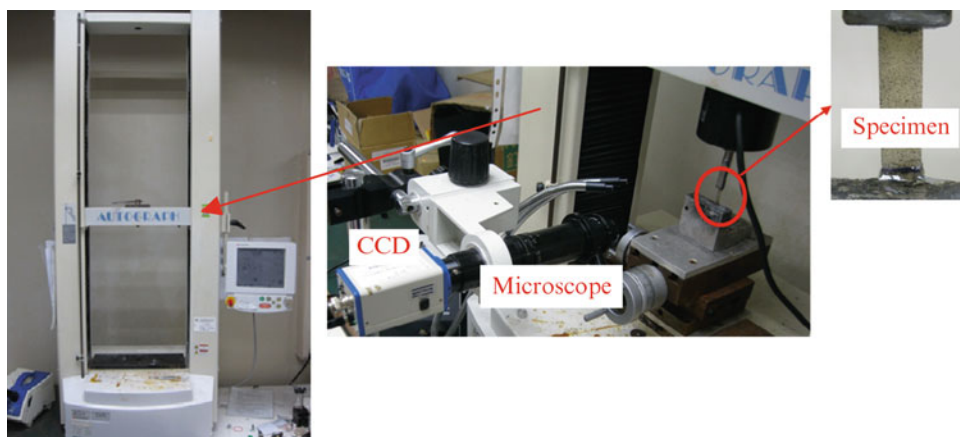
Fig. 21.1 Creep and recovery test system

Fig. 21.2 Loading profile for creep and recovery test
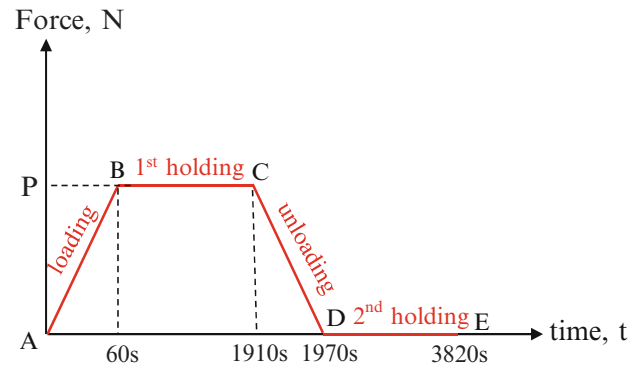
Fig. 21.3 Four-parameter model
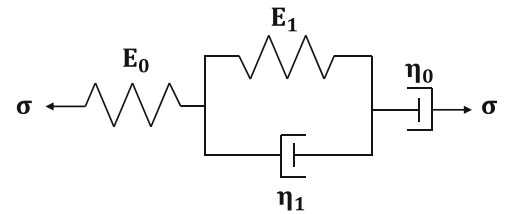
In order to judge whether the viscoelastic behavior is linear or not, isochronal test is done for all dental composite resins. During the test, specimens were subjected to three different uni-axial tensile forces P (45 N, 75 N and 105 N) with loading rate P N/min. The loading profile is shown in Fig. 21.2. Each creep and recovery experiment cycle consists of four steps: loading (A to B), first holding (B to C) with force P for 1,850 s, unloading (C to D) and secondary holding (D to E) with zero force for 1,850 s. Finally, those creep behavior curves were fitted by using the four-parameter fluid model and evaluated.

A linear rheological model (four-parameter fluid) is shown in Fig. 21.3. This model represents the most general behavior possible for a viscoelastic material that includes instantaneous elasticity, delayed elasticity, and flow.

The creep response of this model can be derived as the following equation.

$$\frac{\varepsilon(t)}{\sigma} = \left[\frac{1}{E_0} + \frac{1}{E_1}\left(1 - e^{-t/\tau}\right) + \frac{t}{\eta_0} \right] = J(t) \tag{21.1}$$

where σ is the constant stress applied to specimens, E_0 represents the instantaneous elastic modulus, E_1 and η_1 are the part to describe primary creep, τ (η_1/E_1) is relaxation time, $1/\eta_0$ is the slope of steady-state creep (strain rate) and $J(t)$ is creep compliance. Equation (21.1) was employed to fit the creep behavior by least squares method during the first holding part to obtain $J(t)$.

If stress varies with time, Eq. (21.1) will become a convolution integral:

$$\varepsilon(t) = \int_0^t J(t - \tau) \frac{d\sigma}{d\tau} d\tau \tag{21.2}$$

The creep strain can be obtain by integrating the fitted $J(t)$ and σ using Eq. (21.2).

21.3 Results and Discussion

Figure 21.4 shows the typical longitudinal strains measured for the samples under the creep and recovery test. It can be observed that the order of creep strain of 5 min-group was: Z350 flow > P60 ≈ Z250 > LS, and the order of the creep strain of 30 day-group was: Z350 flow > LS > P60 ≈ Z250. The behavior of 5 min-group exhibited higher creep, strain rates and residual strains than that of 30 day-group.

The creep in 5 min-group are small than that in 30 day-group. This phenomenon may be attributed to the fact that specimens measured after 5 min had a lower degree of conversion than the stored specimens. Because dental composite

Fig. 21.4 Longitudinal strain (75 N) vs. time
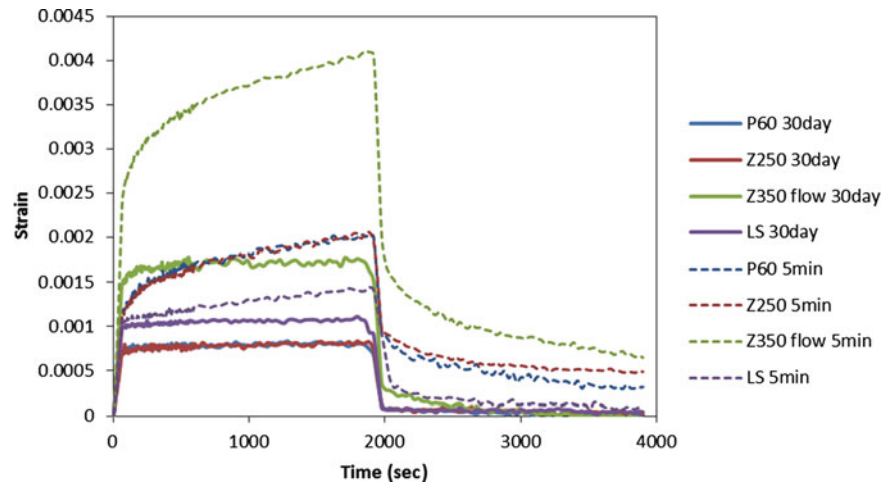
Fig. 21.5 Isochronous stress–strain diagrams and its straight light fit of LS in 5 min-group at 4 chosen time
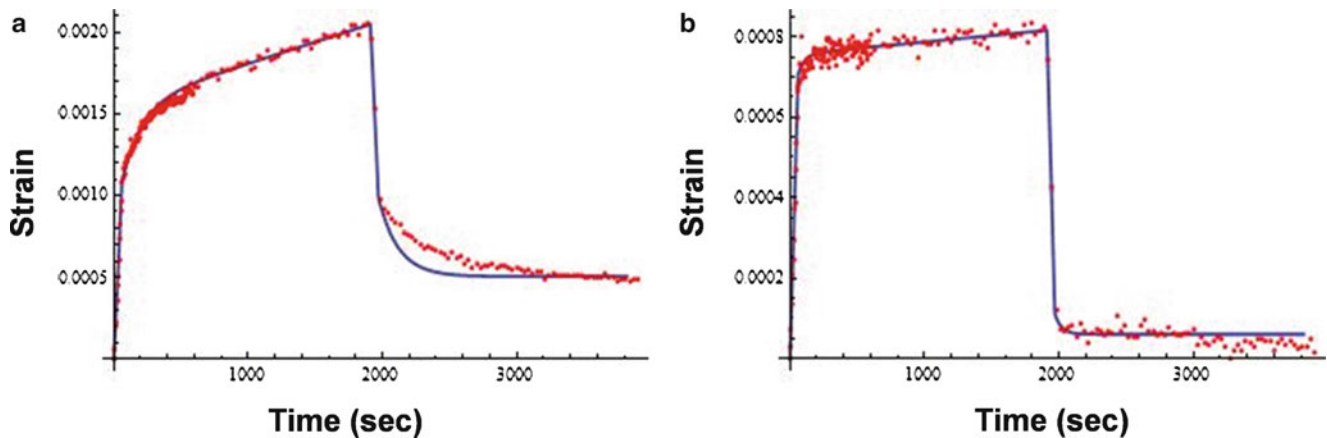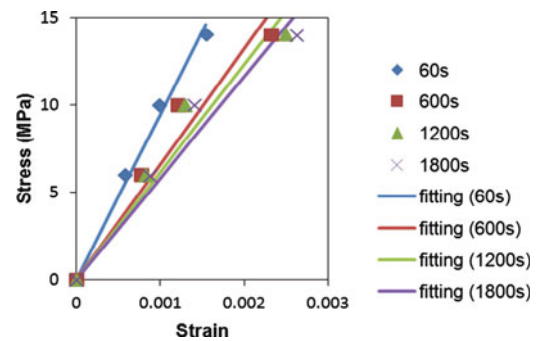
Fig. 21.6 Time dependent strain of Z250 fitted by four-parameter fluid: (**a**) 5 min, and (**b**) 30 days

resins of post irradiation storage at higher temperatures may lead to an enhanced mobility of dimethacrylate monomers and a greater flexibility of the chains in the network, leading to further cross-linking and thereby a more rigid network [6]. However, the creep change of LS is the smallest among them. This is attributed to the doubled curing time.

The measured strains at different time by DIC can be plotted verse stress to form an isochronous stress–strain diagram. A typical isochronous stress–strain diagram of LS in 5 min-group at four different times (60, 600, 1,200, and 1,800 s) is shown in Fig. 21.5. A straight line is well fitted to the data of all composite resin with coefficient of correlation greater than 0.97 (almost 0.99). Therefore, the time-dependent mechanical behavior of dental composite resin could be considered as linear.

The measured time dependent strain of Z250 fitted by four-parameter fluid model is shown in Fig. 21.6. Relaxation time (τ) of 150 s can be obtained from the fitted curve, which is the same for all composite resin studied. The strain rate of

Table 21.3 Strain rate of steady-state creep ($1/\eta_0$)

	Z350 flow	Z250	P60	LS
5 min	4.047E-05	2.650E-05	1.719E-05	0.669E-05
30 days	0.340E-05	0.318E-05	0.136E-05	0.165E-05

Unit: 1/GPa-s
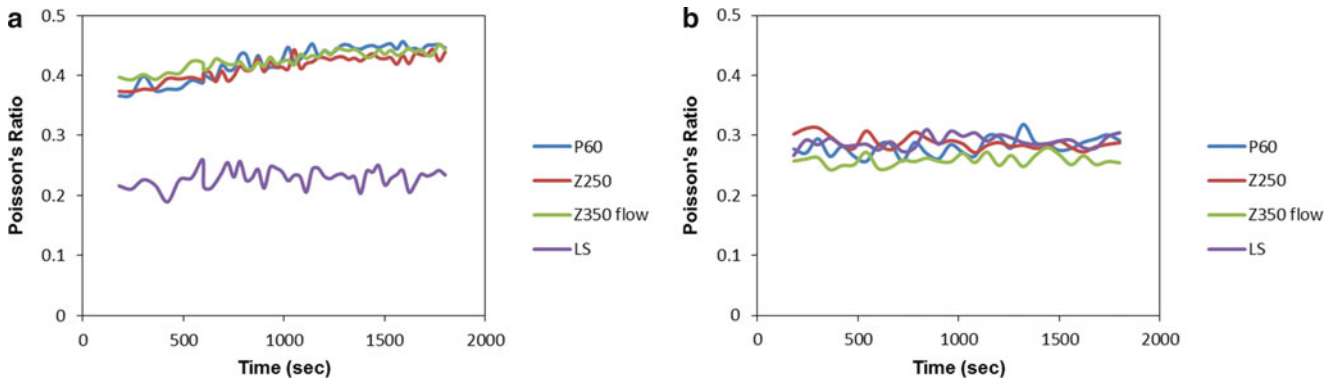
Fig. 21.7 Poisson's ratio vs. time: (**a**) 5 min-group and (**b**) 30 day-group

steady-state creep of all composite resin obtained is given in Table 21.3. The order of the strain rate of 5 min-group are: Z350 flow > Z250 > P60 > LS, however, the order of the strain rate of 30 day-group are: Z350 flow > Z250 > LS > P60.

Figure 21.7 shows the variation of Poisson's ratios of P60, Z250, Z350 flow and LS in 5 min-group and 30 day-group. It can be observed in Fig. 21.7a that Poisson's ratios of 5 min-group increase from 0.35 to 0.45 with increasing time for all composite resins except for LS, which is near constant and is the lowest. The reason is that the inner structure of dental composite resins is unstable. However, Poisson's ratio of all composites resins are similar in 30 day-group (Fig. 21.7b) and does not varied much with increasing time.

21.4   Summary

In this study, the tensile creep and recovery behavior of four dental composites (Z350 flow, universal composite Z250, packable composite P60, and low-shrinkage composite LS) were investigated. With force control, the specimen image under unit-axial tension was recorded and analyzed by using digital image correlation (DIC) method. Four-parameter fluid model is found applicable to fit the experimental obtained strain. The results show that the longitudinal strain of 5-min case has higher creep and residual strain than that of 30 days. The creep of LS was the smallest in four resin composites. The Poisson's ratio of 5-min case was increased about 0.5 with time except for LS. However, the Poisson's ratio of 30-day cases varied up and down insignificantly with time.

References

1. Sun R, Suansuwan N, Kilpatrick N, Swain M (2000) Characterisation of tribochemically assisted bonding of composite resin to porcelain and metal. Am J Dent 28(6):441–445
2. Vano M, Goracci C, Monticelli F, Tognini F, Gabriele M, Tay FR, Ferrari M (2006) The adhesion between fibre posts and composite resin cores: the evaluation of microtensile bond strength following various surface chemical treatments to posts. Int Endod J 39(1):31–39
3. Beun S, Bailly C (2012) Physical, mechanical and rheological characterization of resin-based pit and fissure sealants compared to flowable resin composites. Dent Mater 28:349–359
4. Braga RR, Ferracane JL (2004) Alternatives in polymerization contraction stress management. Crit Rev Oral Biol Med 15:176–184

5. Tantbirojn D, Pfeifer CS, Braga RR, Versluis A (2011) Do low-shrink composites reduce polymerization shrinkage effects? J Dent Res 90 (5):596–601
6. Baroudi K, Silikas N, Watts DC (2007) Time-dependent visco-elastic creep and recovery of flowable composites. Eur J Oral Sci 115:517–521
7. Bhamra GS, Fleming GJP (2012) Influence of halogen irradiance on short- and long-term wear resistance of resin-based composite materials. Dent Mater 25(2):214–220
8. Schmidt C, Ilie N (2012) The mechanical stability of nano-hybrid composites with new methacrylate monomers for matrix compositions. Dent Mater 28:152–159
9. Grantham SG, Siviour CR, Proud WG, Field JE (2004) High-strain rate Brazilian testing of an explosive simulant using speckle metrology. Meas Sci Technol 15:1867–1870
10. Chuang S-F, Chang C-H, Chen TY-F (1963) Contraction behaviors of dental composite restorations—finite element investigation with DIC validation

Chapter 22
Deformation Distribution Measurement from Oblique Direction Using Sampling Moire Method

Daiki Tomita, Motoharu Fujigaki, and Yorinobu Murata

Abstract A sampling Moire method is one of a deformation measurement method of a large structure such as a building and a bridge. The accuracy of phase difference analysis is from 1/100 to 1/1,000 of the grating pitches. This method can analyze two-dimensional phases from a two-dimensional grating image attached on the surface of a specimen. The phase difference distribution before and after deformation gives displacement distributions for the *x*- and the *y*-directions. Grating phase also gives position for the *x*- and the *y*-directions. Therefore, the relationship between a point on an image and the corresponding point on the object can be obtained without any calibrations. A sampling Moire camera we proposed has an FPGA with an algorithm of the sampling Moire method inside the camera. It can output two-dimensional phases and phase differences from a single shot two-dimensional grating image in real-time. In this paper, a measurement accuracy of the deformation distribution measurement using the sampling Moire camera placed in the oblique direction is estimated. Experimental result of the time-series deformation measurement of a moving cantilever is shown.

Keywords Digital image processing • Sampling Moire method • Sampling Moire camera • Real-time • Displacement distribution measurement

22.1 Introduction

It is required to develop a deformation measurement method without contact for large structures in real-time. The deformation measurement can be applied to detect a prediction of a landslide and the health monitoring of a large structures such as a bridge, a steel tower, and tall buildings.

Phase analysis method of a grating is one of the useful methods to measure deformation accurately. It is better to measure deformation from one-shot image in the case of a moving object or a vibrating object. A spatial fringe analysis method using sampling technique to analyze phase distribution of Moire images produced from one-shot grating image in a computer were proposed by Arai et al. [1]. This method is applied to phase analysis of two-dimensional grating image. We called this method a sampling Moire method. This method was applied to deformation and strain measurement [2–4]. The accuracy of phase difference analysis is in the range between 1/100 and 1/1,000 of the grating pitch. The accuracy of this method is very high. This is due to the phase of the grating pattern which is analyzed using a phase-shifting method. The phase-shifting method can be performed by generating phase-shifted Moire images in a computer via a one-shot grating image. The analysis speed is faster than a Fourier transform method [5] or a digital image correlation method [6].

The sampling Moire method has an advantage. The camera calibration is not necessary because the grating pattern attached on the object has a known pitch of grating. In the case when a large structure is measured, it is difficult to install a camera in the front of the object. This advantage can be applied to the measurement of large structures.

D. Tomita (✉)
Graduate School of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama 640-8510, Japan
e-mail: s152029@sys.wakayama-u.ac.jp

M. Fujigaki • Y. Murata
Faculty of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama 640-8510, Japan
e-mail: fujigaki@sys.wakayama-u.ac.jp

In this paper, the accuracy of dynamic displacement measurement by using a sampling Moire camera placed in the oblique direction is confirmed. Experimental results of the time-series deformation measurement of a moving cantilever are shown.

22.2 Principle of Sampling Moire Method

22.2.1 Spatial Fringe Analysis Method Using Sampling Technique

The principle of the phase analysis method for a one-dimensional grating in a sampling Moire method is essentially the same as the spatial fringe analysis method using a sampling technique [1]. In this method, a specimen grating pattern placed on an object is recorded by a digital camera. Though the digitized image shows only the grating, a Moire fringe pattern appears by thinning out the pixels; that is, by sampling the image with a constant pixel pitch. Figure 22.1 illustrates the appearance of Moire fringe patterns by this method. In this figure, only three horizontal lines are shown. Figure 22.1a shows the deformed grating pattern attached to the specimen. The pitch of the grating in this figure is 1.125 times larger than that of the sampling points. The recorded image is shown in Fig. 22.1b, in which no Moire fringe pattern can be discerned. Figure 22.1c shows the Moire fringe patterns when the recorded image is sampled at every Nth pixel (in the figure, $N = 4$). The four images in Fig. 22.1c are obtained by using the first, second, third and fourth pixels of Fig. 22.1b as the sampling start point, respectively. This process corresponds to the phase-shifting of the fringe pattern. These sampled images shown in Fig. 22.1c are interpolated using neighboring data. Then, the four phase-shifted Moire images shown in Fig. 22.1d are obtained from a single picture in Fig. 22.1b.

The k-th phase-shifted images can be expressed approximately as follows:

$$I_k(i,j) = a(i,j) \cos \left[\theta(i,j) + k\frac{2\pi}{N} \right] + b(i,j) \qquad (k = 0, 1, \ldots, N-1) \tag{22.1}$$

Here, $a(i, j)$ represents the amplitude of the grating intensity, $b(i, j)$ represents the background intensity and $\theta(i, j)$ is the initial phase value. The phase distribution of the Moire pattern can be obtained by the discrete Fourier transform (DFT) algorithm using Eq. (22.2) as follows:

$$\tan \theta = -\frac{\sum_{k=0}^{N-1} I_k \sin \left(k\frac{2\pi}{N} \right)}{\sum_{k=0}^{N-1} I_k \cos \left(k\frac{2\pi}{N} \right)} \tag{22.2}$$

Fig. 22.1 Principle of the phase-shifting Moire method for 1-D grating pattern. (**a**) Deformed grating pattern on specimen, (**b**) Recorded image by digital camera, (**c**) Moire fringe patterns obtained when every N- (=4) pixel from the first, second, third and fourth sampling point is picked up from (**b**), respectively, (**d**) Linear interpolated images from (**c**), respectively, (**e**) Phase θ_m distribution analyzed from (**d**)
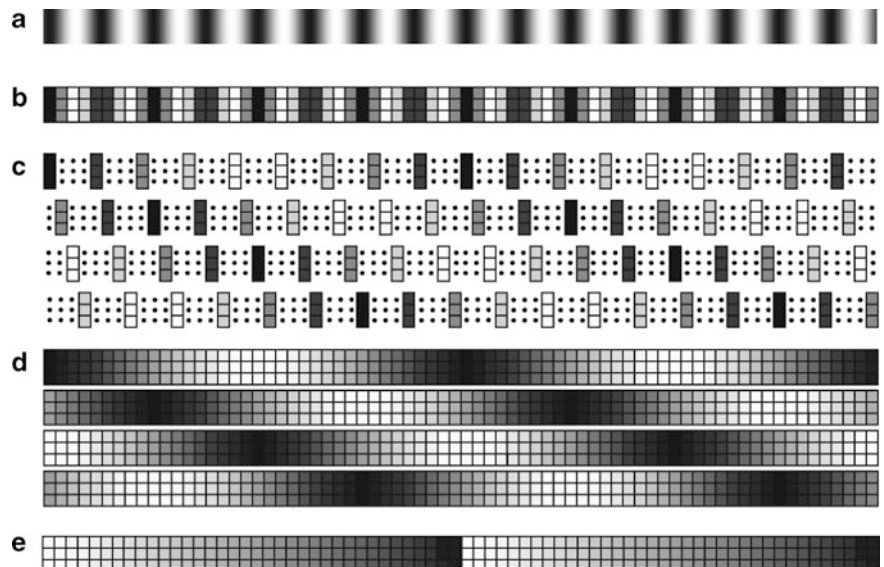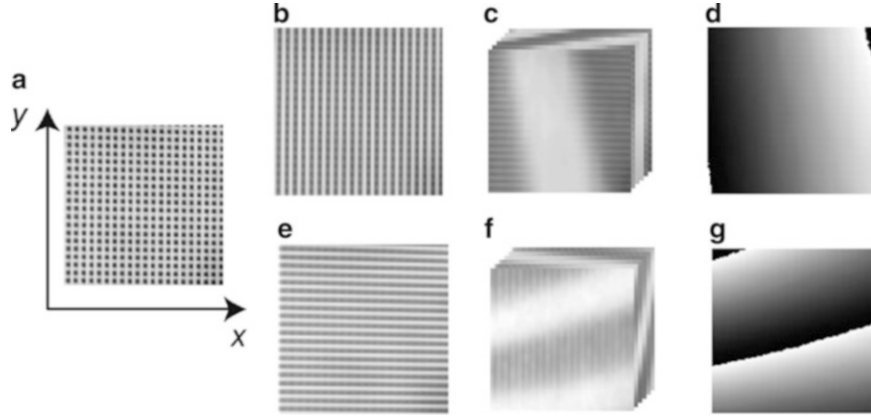
Fig. 22.2 Process of phase analysis by using smoothing process and the phase-shifting Moire method for 2-D grating pattern. (**a**) 2D grating image, (**b**) *Vertical* stripe image, (**c**) Phase shifted sampling Moire image for x-direction, (**d**) Moire phase distribution image for x-direction, (**e**) *Horizontal* stripe image, (**f**) Phase shifted sampling Moire image for y-direction, (**g**) Moire phase distribution image for y-direction

The phase θ_m of the Moire pattern is defined as the difference between the grating phase θ_g and the phase θ_r of the reference grating (i.e., the sampling phase) as follows:

$$\theta_m = \text{Wrap}\left(\theta_g - \theta_r\right) \tag{22.3}$$

Here, Wrap is the phase wrapping function defined as Eq. (22.4).

$$\text{Wrap}(\theta) = \begin{cases} \tan^{-1}(\sin\theta/\cos\theta) & (\cos\theta \geq 0) \\ \tan^{-1}(\sin\theta/\cos\theta) + \pi & (\cos\theta < 0,\ \sin\theta \geq 0) \\ \tan^{-1}(\sin\theta/\cos\theta) - \pi & (\cos\theta < 0,\ \sin\theta < 0) \end{cases} \tag{22.4}$$

The grating phase θ_g can be obtained as follows:

$$\theta_g = \text{Wrap}(\theta_m + \theta_r) \tag{22.5}$$

22.2.2 Phase Analysis for 2-D Grating Pattern

Figure 22.2 shows examples of two dimensional phase analysis by the sampling Moire method. Figure 22.2a shows the 2-D grating image captured by a camera. Figure 22.2b shows a grating image obtained after smoothing process in the *y*-direction. By this process, *x*-dimensional phase distribution can be analyzed because the 2-D grating is converted into the pattern image as one-dimensional grating pattern image. Figure 22.2c shows phase-shifted sampling Moire images produced from Fig. 22.2b. Figure 22.2d shows the phase distribution for *x*-direction produced from Fig. 22.2c by phase-shifting method. Figure 22.2e shows a grating image obtained after the *y*-dimensional smoothing process.

22.2.3 Sampling Moire Camera

Figure 22.3a, b show a photograph and a block diagram of a developed sampling Moire camera, respectively [7]. The camera is composed of a CMOS sensor, an FPGA, memory modules and a USB interface. A two-dimensional grating image on the object is taken by the CMOS sensor in synchronization with a camera trigger. The algorithm mentioned previous was written into the FPGA. The two-dimensional grating images analyzed from one-shot image by the FPGA and outputted in real-time.
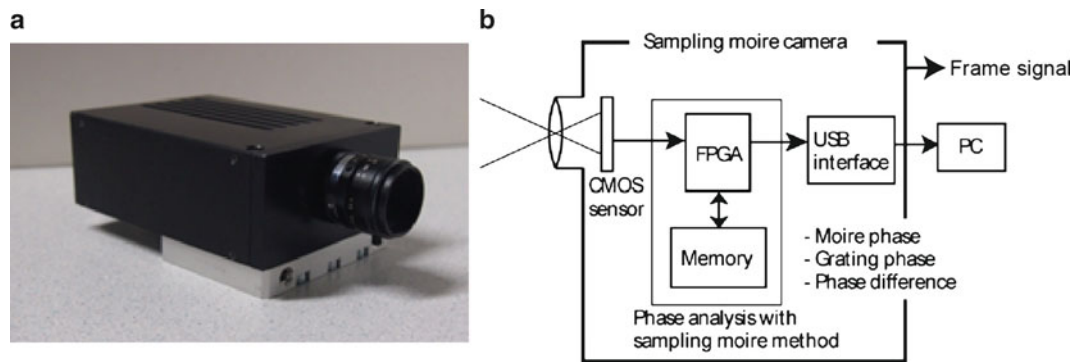
Fig. 22.3 Composition of sampling Moire camera which can measure dynamic displacement measurement in real-time. (**a**) Photograph (**b**) Block diagram
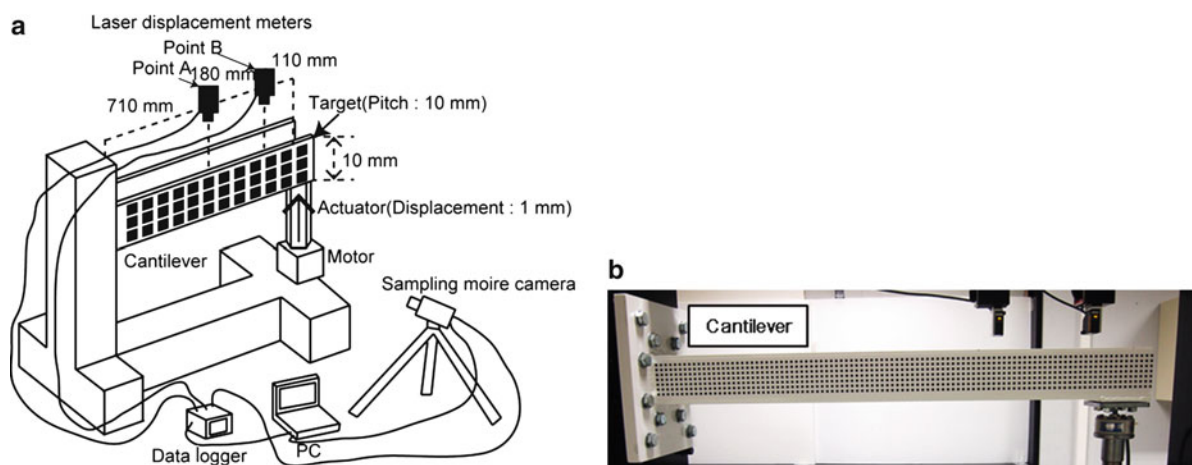
Fig. 22.4 Experimental composition of dynamic displacement measurement by using sampling Moire camera. (**a**) Illustration (**b**) Cantilever (The length is 1 m)

The size of this camera is 86 mm × 44 mm × 128 mm. The weight is 0.66 kg. The effective pixel size is 2,048 × 1,024 pixels. The frame rate is 19.8 fps with a high-resolution mode (2,048 × 1,024 pixels), 117 fps with a normal mode (512 × 512 pixels), and 288 fps with a high-speed mode (128 × 128 pixels).

22.3 Experiment

22.3.1 Displacement Measurement Using a Cantilever

An application on dynamic deformation measurement of a cantilever is demonstrated. Figure 22.4 shows an experimental composition. The length of the cantilever is 1 m. Two-dimensional grating patterns are pasted on the cantilever. The pitch of the grating is 10 mm. The quantity of displacement to give the right-side end of the cantilever is 1 mm. The vibration to give a cantilever is 1 Hz. Two laser displacement meters are installed above the cantilever as Point A and Point B. The frame signal of the sampling Moire and the output of the laser displacement meters are recorded by a data logger with a sampling period of 1 ms.

Camera angles are 0°, horizontal direction 60°, and vertical direction 30 and horizontal direction 30° such as Figs. 22.5 and 22.6. The frame rate of the camera is 50 fps and the exposure time is 5 ms.

Figure 22.7 shows captured images by the sampling Moire camera. Figure 22.8 shows images of displacement distribution measured by the sampling Moire camera. Figure 22.9 shows results of displacement distribution measured by the
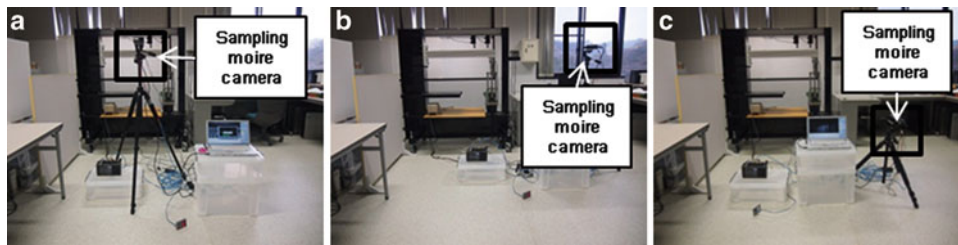
Fig. 22.5 Camera angles of sampling Moire camera. (**a**) 0° (**b**) *Horizontal* direction 60° (**c**) *Vertical* direction 30 and *horizontal* direction 30°
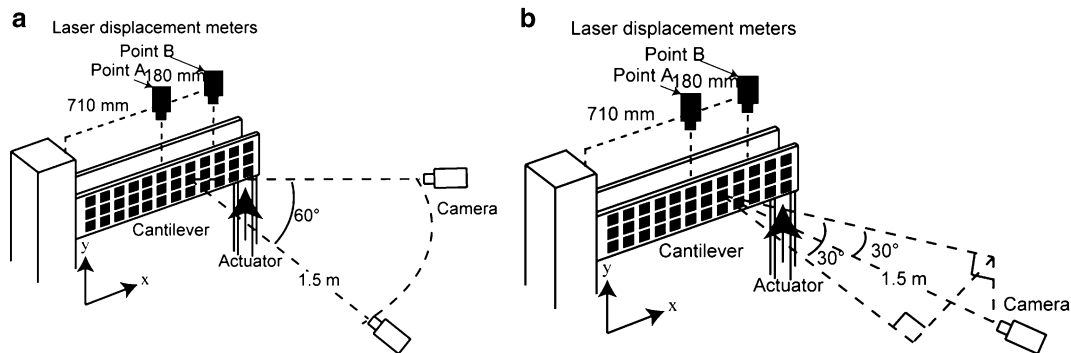
Fig. 22.6 Illustration of camera angles of sampling Moire camera. (**a**) 0° and *horizontal* direction 60° (**b**) *Vertical* direction 30 and *horizontal* direction 30°

Fig. 22.7 Images captured images by sampling Moire camera. (**a**) 0°, (**b**) *Horizontal* direction 60°, (**c**) *Vertical* direction 30 and *horizontal* direction 30°
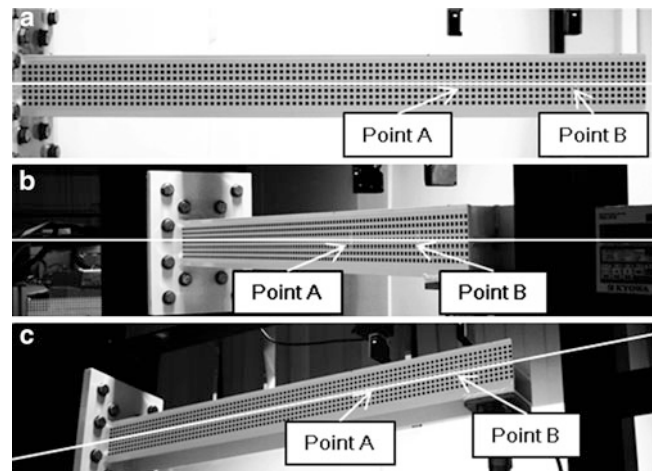
sampling Moire camera and the laser displacement meters when the camera angle is 0°, horizontal direction 60°, and vertical direction 30 and horizontal direction 30°, respectively. Figure 22.10 shows results of real-time displacement measured by the sampling Moire camera and the laser displacement meters when the camera angle is 0°, horizontal direction 60°, and vertical direction 30 and horizontal direction 30°, respectively. Figure 22.11 shows results of error of the difference between the sampling Moire camera and the laser displacement meters when the camera angle is 0°, horizontal direction 60°, and vertical direction 30 and horizontal direction 30°, respectively.

The difference of displacement measured by the sampling Moire camera and the difference of displacement measured by the laser displacement meters is 0.05 mm or less when the camera angle is 0°, horizontal direction 60°. The difference of displacement measured by the sampling Moire camera and the difference of displacement measured by the laser

Fig. 22.8 Images of
displacement distribution.
(**a**) 0°, (**b**) *Horizontal*
direction 60°, (**c**) *Vertical*
direction 30 and *horizontal*
direction 30°
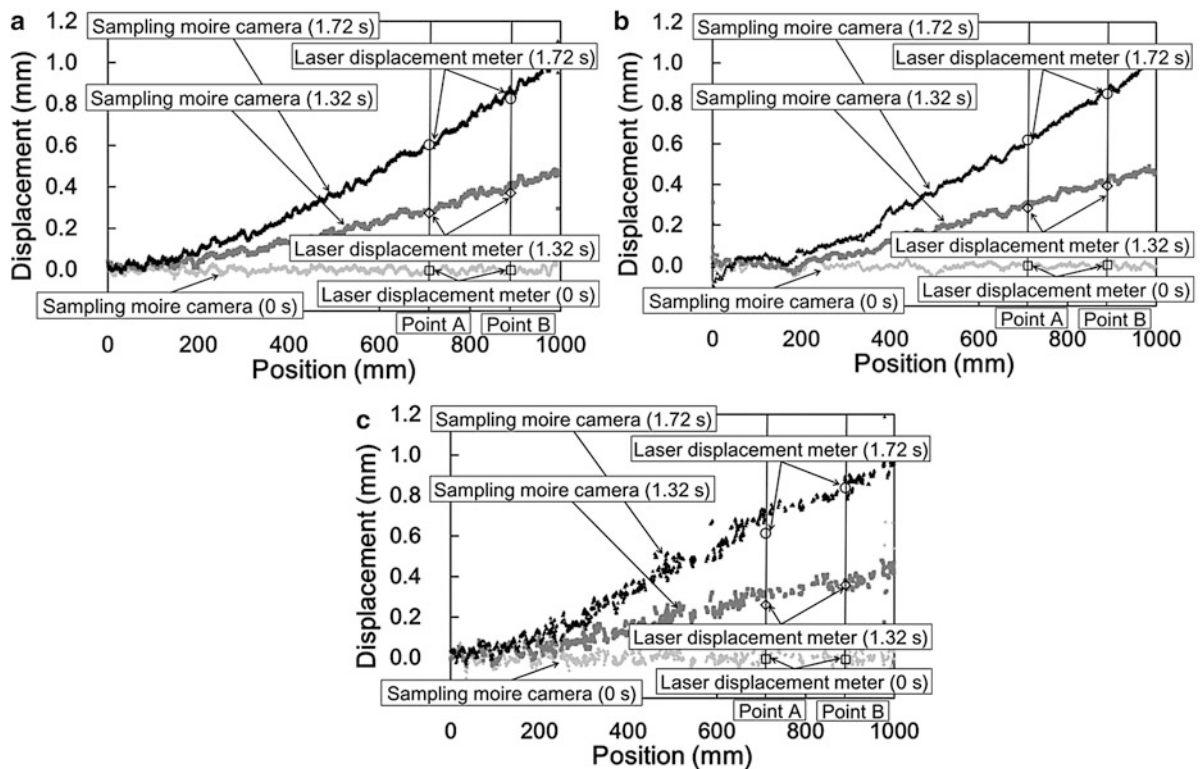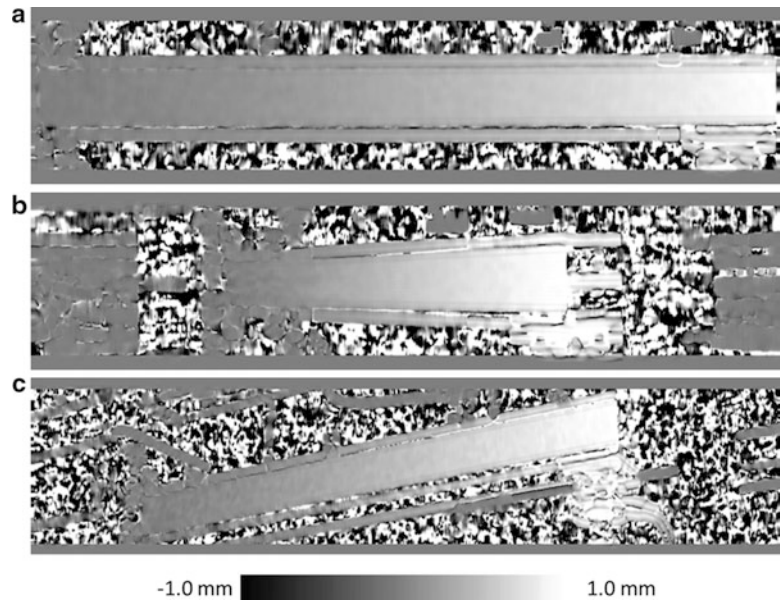
Fig. 22.9 Displacement distribution results (**a**) 0°, (**b**) *Horizontal* direction 60°, (**c**) *Vertical* direction 30 and *horizontal* direction 30°

displacement meters is less than 0.1 mm when the camera angle is vertical direction 30 and horizontal direction 30°. They are 1/200 and 1/100 of the pitch of the grating, respectively. It is confirmed that the effectiveness of a deformation distribution measurement from oblique direction using sampling Moire method. The error occurred periodically. It has an ability to reduce the error with finding out the cause of the periodic error.
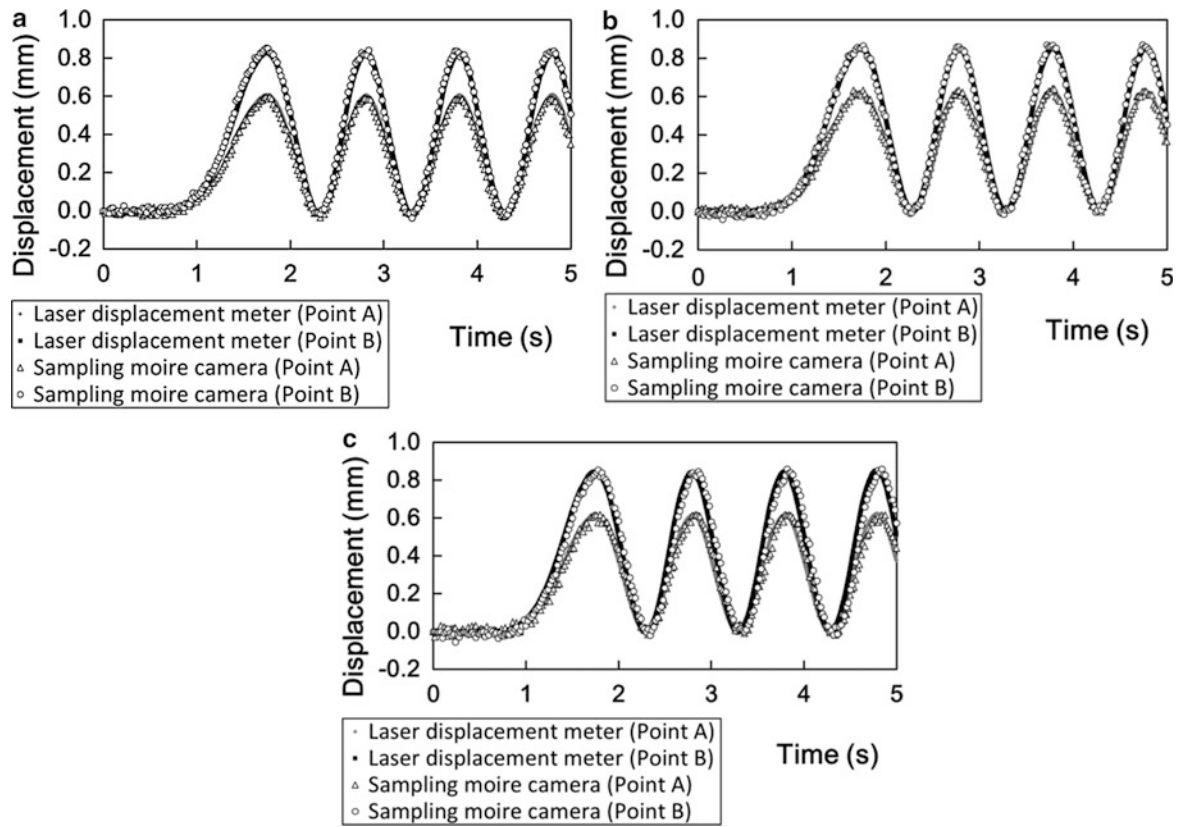
Fig. 22.10 Real-time displacement measured by a sampling Moire camera and laser displacement meters (**a**) 0°, (**b**) *Horizontal* direction 60°, (**c**) *Vertical* direction 30 and *horizontal* direction 30°
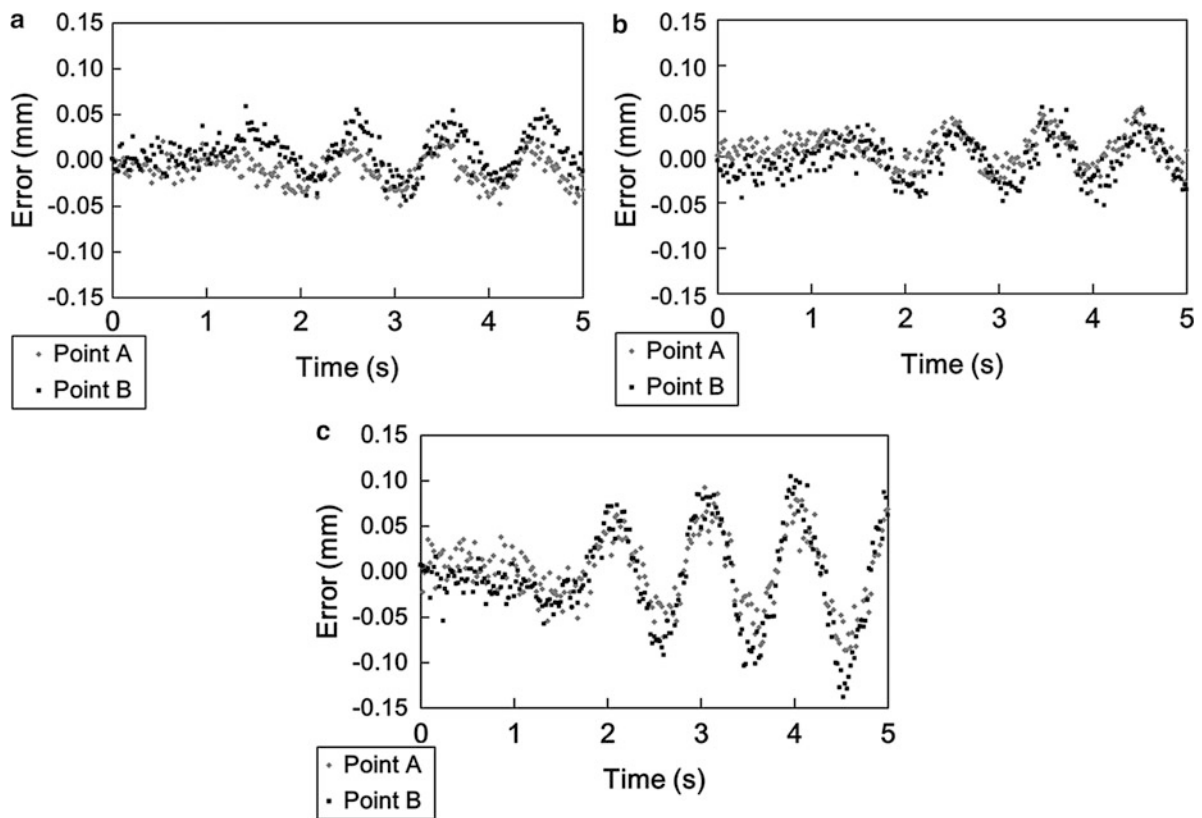
Fig. 22.11 Error of difference between a sampling Moire camera and laser displacement meters. (**a**) 0° (**b**) *Horizontal* direction 60°, (**c**) *Vertical* direction 30 and *horizontal* direction 30°

22.4 Conclusions

In this paper, a real-time displacement measurement of a steel cantilever using a sampling Moire camera placed in the oblique direction was performed. The results showed that the error was 1/200 and 1/100 of the pitch of the grating. The effectiveness of a deformation distribution measurement from oblique direction using sampling Moire method was confirmed. It is expected that the periodically error can be reduced with improvement the sampling Moire method.

References

1. Arai Y, Yokozeki S, Shiraki K, Yamada T (1997) High precision two-dimensional spatial fringe analysis method. J Mod Opt 44(4):739–751
2. Ri S, Fujigaki M, Morimoto Y (2010) Sampling Moire method for accurate small deformation distribution measurement. Exp Mech 50(4):501–508
3. Shimo K, Fujigaki M, Masaya A, Morimoto Y (2009) Development of dynamic shape and strain measurement system by sampling Moire method. In: Proceedings of SPIE, 7522, ICEM2009
4. Hanada R, Miyazawa M, Fujigaki M, Simo K, Morimoto Y (2013) A sampling Moire method to measure the dynamic shape and strain of rotating tires. Tire Sci Technol 41(3):214–225
5. Takeda M, Ina H, Kobayashi S (1982) Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. Opt Soc Am 72(1):156–160
6. Yoneyama S, Kitagawa A, Kitamura K, Kikuta H (2006) In-plane displacement measurement using digital image correlation with lens distortion correction. JSME Int J Ser A 49(3):458–467
7. Fujigaki M, Sasatani Y, Masaya A, Kondo H, Nakabo M, Hara T, Morimoto Y, Asai D, Miyagi T, Kurokawa N (2011) Development of sampling Moire camera for real-time phase analysis. Appl Mech Mater 83:48–53

Chapter 23
Automatic Stress Measurement by Integrating Photoelasticity and Spectrometry

Po-Chi Sung, Yu-An Chiang, Wei-Chung Wang, and Te-Heng Hung

Abstract By integrating the white light photoelasticity (WLP) and spectrometry, a novel stress measurement method called transmissivity extremities theory of photoelasticity (TEToP) has been recently developed. The key of TEToP is to establish the systematic three-dimensional relationship of transmissivity with stress and wavelength (3DRoTwSW) and further derive the stress quantifying formula (SQF). Based on the high resolution and sensitivity of the spectrometer and the high measurement accuracy of the WLP, in contrast to the traditional photoelasticity, TEToP can measure the low level stress even in low birefringence materials. However, until now, TEToP has not been used in higher level stress measurement. The primary purpose of this paper is to extend the application of TEToP to higher level stress measurement. By analyzing the 3DRoTwSW, the correlation between different transmissivity extremities linear equations (TELEs) can be proved definitely and the feasibility of the development of TEToP on automatic stress measurement can be further confirmed. Furthermore, the measurement results of the specimens with different thicknesses were investigated. By exploring the connection between the parameters in TEToP and thicknesses of specimens, the SQFs of different thicknesses can be further derived. In this paper, both glass and PSM-1 were used to verify the generality of TEToP.

Keywords Transmissivity extremities theory of photoelasticity • Transmissivity extremities linear equation • Automatic stress measurement • White light photoelasticity • Spectrometer

23.1 Introduction

In recent years, the development of the plane display industry has been booming, especially the thin film transistor-liquid crystal display (TFT-LCD) [1]. Asia has become the number one war zone of the plane display industry. With those extremely intense competitions, how to increase the resolution and defect-free rate of products effectively rather than massive trials and errors has become a grand challenge.

In the field of experimental mechanics, photomechanic methods have full-field, non-destructive, non-contact and remote-sensing and high sensitivity characteristics [2]. They can be used to detect defects and improve design and manufacturing process of components. Digital photoelasticity [3] is the only photomechanic method which is able to measure the state of stress directly. In the past, many efforts were devoted to automate the digital photoelasticity [4–12]. In 2012, based on the digital photoelasticity, Wang et al. [13, 14] proposed a stress measurement method named transmissivity extremities theory of photoelasticity (TEToP) to make the measurement of low level stress possible.

The principle of TEToP is to analyze the spectroscopy in the white light photoelasticity (WLP) [7–12] and further to construct the systematic three-dimensional relationship of transmissivity with stress and wavelength (3DRoTwSW). By using the systematic 3DRoTwSW, the stress quantifying formula (SQF) can be derived and the stresses can be determined by substituting the measured transmissivity spectrum (TS) into the stress quantifying formula (SQF).

With high resolution of the spectrometer and high sensitivity of the WLP to the variety of light intensity, TEToP can measure the stresses more precisely than traditional photoelasticity.

P.-C. Sung • Y.-A. Chiang • W.-C. Wang (✉) • T.-H. Hung
Department of Power Mechanical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan ROC
e-mail: wcwang@pme.nthu.edu.tw

The 3DRoTwSW and the correlation between TELEs were verified in this paper. By investigating the relationship between parameters of TEToP and thicknesses of specimens, the relationship of stress and thickness can be derived. Finally, two transparent materials, PSM-1 and glass, were employed to prove the generality of TEToP.

23.2 Theory of the Automatic Stress Measurement

23.2.1 TEToP [13, 14]

Figure 23.1 shows the transmissivity spectrogram of the experimentally obtained 3DRoTwSW of the PSM-1 material. In Fig. 23.1, it can be found that the transmissivity has a systematic relationship between stress and wavelength. Furthermore, as indicated by gray dotted lines in Fig. 23.1, the transmissivity extremities (TEs) of each wavelength display linear distributions. The TEToP is obtained by combining the TELEs of those linear distributions with the TS equation of the isochromatic fringe order to reconstruct the 3DRoTwSW. According to the stress-optic law, the TS equation of the isochromatic fringe order can be expressed as [13]

$$T_\lambda = T_{b\lambda} + T_{a\lambda} \sin^2\left(\frac{\pi t C}{\lambda} S\right) \tag{23.1}$$

where $T_{a\lambda}$ and $T_{b\lambda}$ are the amplitude and background transmissivity corresponding to the wavelength λ; S is the in-plane principal stress difference (PSD); C is the stress-optic coefficient (SOC) and t is the thickness of the specimen. To make T_λ be free of the influence of $T_{a\lambda}$ and $T_{b\lambda}$, normalization was implemented as

$$T_{N\lambda} = \frac{T_\lambda - T_{b\lambda}}{T_{a\lambda}} = \sin^2\left(\frac{\pi t C}{\lambda} S\right) \tag{23.2}$$

where $T_{N\lambda}$ denotes the normalized TS. In Fig. 23.1, the first TELE can be expressed by

$$S = A\lambda + B \tag{23.3}$$

where A is the slope and B is the intercept.
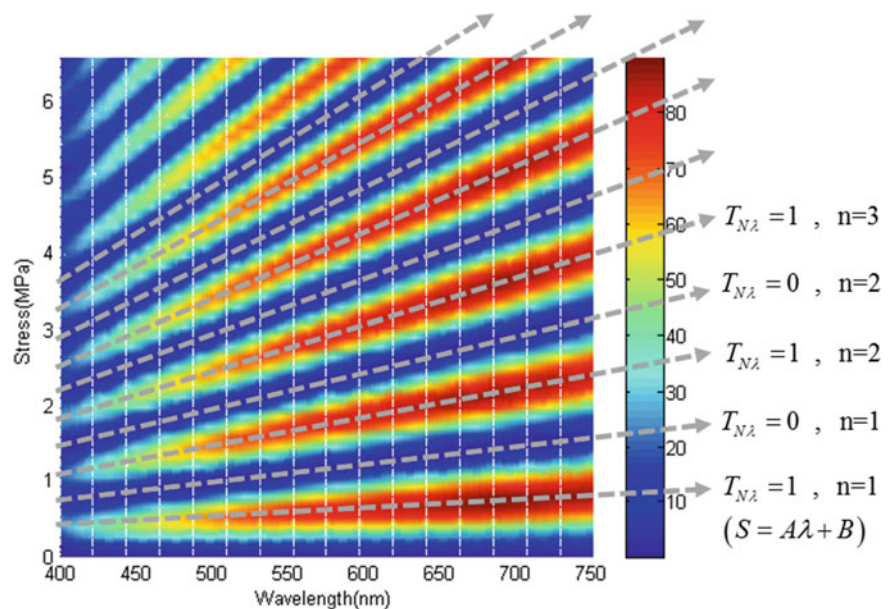
Fig. 23.1 The transmissivity spectrogram of 3DRoTwSW of the PSM-1 material [14]

Because $T_{N\lambda}$ is equal to 1 at the first TELE, by substituting Eq. 23.3 into Eq. 23.2, Eq. 23.2 becomes

$$1 = \sin^2\left[\frac{\pi t C}{\lambda}(A\lambda + B)\right] \tag{23.4}$$

By combining Eqs. 23.2 and 23.4, it can be shown that [13, 14]

$$S = \frac{2}{\pi}(A\lambda + B)\sin^{-1}\sqrt{T_{N\lambda}} \tag{23.5}$$

Equation 23.5 represents the reconstructed 3DRoTwSW. Utilizing this systematic relationship, S can be determined by substituting the measured TS into the following equation [13, 14]

$$S_m = \frac{2}{\pi}(A\lambda + B)\sin^{-1}\sqrt{\frac{T_{m\lambda} - 2T_{b\lambda}}{T_{a\lambda}}} \tag{23.6}$$

where S_m and $T_{m\lambda}$ are the measured S and TS, respectively. Equation 23.6 is called the SQF of TEToP. By applying the SQF to determine S, it is obvious that the stress optical coefficient (SOC) is no longer needed. Moreover, SQF can be used with all kinds of light sources either monochromatic or multi-wavelength. S can be determined by substituting $T_{m\lambda}$ and wavelength of the light source into Eq. 23.6 when a monochromatic light source is used. On the other hand, S can be determined by taking average of S values of every wavelength when a multi-wavelength light source is adopted.

23.2.2   Determination of Parameters of Tetop Based on the Regression Analysis

Before measurement, parameters in Eq. 23.6, i.e. $A\lambda + B$, $T_{a\lambda}$, and $T_{b\lambda}$ should be known beforehand. Therefore, a calibration specimen made of the same material as the test specimen should be prepared. Sequentially apply the load to the calibration specimen, the database can be established by recording Ss and the corresponding TSs for each load. By further analyzing the database by the normalization and linear fitting, $A\lambda + B$, $T_{a\lambda}$, and $T_{b\lambda}$ can be obtained.

The total stress in the calibration specimen can be expressed as

$$S = S_a + S_r \tag{23.7}$$

where S_r is the residual stress and S_a is the applied stress.

Substituting Eq. 23.7 into Eq. 23.1, the measured TS can be obtained as

$$T_\lambda = T_{b\lambda} + T_{a\lambda}\sin^2\frac{\pi t C(S_a + S_r)}{\lambda} \tag{23.8}$$

In practice, Eq. 23.8 can be rewritten in the form of the constant term and first-order harmonic terms of the Fourier series function. It should be mentioned here that the constant term and first-order harmonic terms of the Fourier series function is called Fourier 1 function in the Matlab [15].

$$T_\lambda = (c_{0\lambda} - c_{1\lambda}) + 2c_{1\lambda}\sin^2\left[\frac{S_a + (c_{3\lambda}/2 - c_{2\lambda})}{2c_{3\lambda}}\pi\right] \tag{23.9}$$

where $c_{0\lambda}$, $c_{1\lambda}$, $c_{2\lambda}$, and $c_{3\lambda}$ are the fitting results obtained by the regression analysis of the Fourier 1 function. Therefore, by comparing Eq. 23.8 with Eq. 23.9, the undetermined parameters $T_{b\lambda}$, $T_{a\lambda}$, S_r, and C in Eq. 23.8 can be expressed by $c_{0\lambda}$, $c_{1\lambda}$, $c_{2\lambda}$, and $c_{3\lambda}$, i.e.

$$\begin{cases} 2T_{b\lambda} = c_{0\lambda} - c_{1\lambda} \\ T_{b\lambda} = 2c_{1\lambda} \\ S_r = \dfrac{c_{3\lambda}}{2} - c_{2\lambda} \\ C = \dfrac{\lambda}{2tc_{3\lambda}} \end{cases} \tag{23.10}$$

In Eq. 23.9, when T_λ is an extremity, the argument of $\sin^2$ function is $\pi/2$, i.e.

$$\frac{S_a + (c_{3\lambda}/2 - c_{2\lambda})}{2c_{3\lambda}} \pi = \frac{\pi}{2} \tag{23.11}$$

Substituting S_r of Eqs. 23.10 and 23.7 into Eq. 23.11, Eq. 23.11 can be expressed as

$$S = c_{3\lambda} \tag{23.12}$$

In a transmissivity spectrogram such as Fig. 23.1, Eq. 23.12 represents the position of the first TE at each wavelength. Thus, the first TELE can be directly obtained by linear fitting the Eq. 23.12. Based on the above analysis, the first TELE and $T_{b\lambda}$, $T_{a\lambda}$, S_r, and C can be determined even with nil or low level stress in the calibration specimen.

23.2.3 Correlation Between Teles

From Eq. 23.5, different orders of TELEs can be derived from Eq. 23.5 as [13, 14]

$$T_{N\lambda} = \begin{cases} 0, \text{when } S = 2n(A\lambda + B) \\ 1, \text{when } S = (2n - 1)(A\lambda + B) \end{cases} \text{where } n = 1, 2, \ldots \tag{23.13}$$

where n is the order of TELE. In Eq. 23.13, it can be found that $T_{N\lambda}$ is maximum when S is the odd times of $A\lambda + B$. On the other hand, $T_{N\lambda}$ is minimum when S is zero or the even times of $A\lambda + B$.

23.2.4 Relationship Between SQF and Thickness of Specimen

Equations 23.5 and 23.6 were derived when the thickness of test specimen is constant. With further development of the relationship between SQF and thickness of the test specimen, S can be measured even when the specimen thickness is not constant.

The SOC is a material constant which is independent of the thickness of the test specimen, it can be expressed as follows [13, 14]

$$C = \frac{\lambda}{2t(A\lambda + B)} \tag{23.14}$$

Equation 23.14 shows that $A\lambda + B$ (i.e. the first TELE defined in Eq. 23.3) is inversely proportional to the thickness of the specimen. Thus, the TELE of another thickness t_1 can be expressed by the TELE of thickness t as

$$A_{t1}\lambda + B_{t1} = \frac{t(A\lambda + B)}{t_1} \tag{23.15}$$

Substituting Eq. 23.15 into Eq. 23.6, the new expression of S_m becomes

$$S_m = \frac{2}{\pi}(A_{t_1}\lambda + B)\sin^{-1}\sqrt{\frac{(T_{Sm\lambda} - 2T_{b\lambda})}{T_{a\lambda}}} \quad = \frac{t}{t_1}\frac{2}{\pi}(A\lambda + B)\sin^{1}\sqrt{\frac{(T_{Sm\lambda} - 2T_{b\lambda})}{T_{a\lambda}}} \tag{23.16}$$

Consequently, if the TELE of thickness t is known, the stresses can be measured in any test specimen whose thickness is known.

23.3 Experimental Arrangements

23.3.1 Experimental Setup

As shown in Fig. 23.2, the experimental setup includes a halogen light source, a polarizer, an analyzer, a CCD camera, a mask and a spectrometer. The test specimen was clamped on the loading frame of a tensile test machine. There is a small hole drilled at the center of the mask. In addition, the whole surface of the mask was covered with a piece of dark black paper. Both the small hole and the dark black paper applied on the mask were used to eliminate the stray light from the environment and make sure that the spectrometer only receives the light intensity from the measurement point.
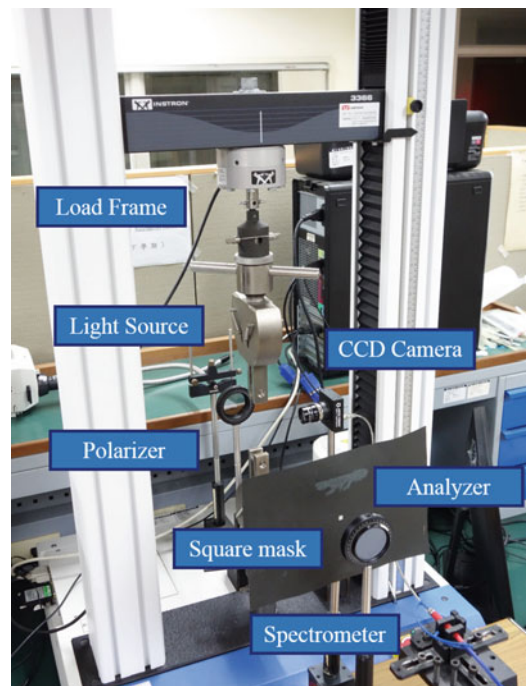
Fig. 23.2 The experimental setup of TEToP

23.3.2 Experimental Procedure

In applying TEToP, TS is essential for determining S when SQF is employed. Thus, before the measurement, it must be ensured that the halogen light source, mask, and the spectrometer are aligned on the same horizontal line. After that, TSs emitted from the dark-field plane polariscope were recorded for the establishment of the database.

The calibration specimen was pinned-fixed on the loading frame. To make sure the optical path is normal to the surface of the specimen during the loading process, a CCD camera was used to monitor the specimen from the lateral side of the specimen.

At the beginning of the test, an initial axial load was applied to the test specimen. For the case of PSM-1, for example, the initial axial load was 400 N. And then the load applied on the test specimen was reduced sequentially. The reduction rate of load was about 20 N per step when the load was larger than 15 N. To maintain the sufficient measurement accuracy, the reduction rate became 0.5–1 N when the load was lower than 15 N. After recording the corresponding Ss of the applied loads and the corresponding TSs of S s, the database was established.

By the regression analysis of Fourier 1 function, the parameters $A\lambda + B$, $T_{a\lambda}$, and $T_{b\lambda}$ can be obtained from the database. Additionally, in order to investigate the relationship between the TELE and thickness, the ratios of thickness, A and B of each thickness were also obtained.

23.4 Experimental Results and Discussions

23.4.1 Results of Teles of PSM-1 Material

Figure 23.3 is the transmissivity spectrogram corresponding to the Ss of PSM-1 specimen. The thickness of the specimen is 5.80 mm and the maximum load applied is 400 N.

The fitting results of Fourier 1 function of TSs at wavelengths 450, 550, 650, and 750 nm are shown in Fig. 23.4. Most of the fitting results match well with the original data. It shows the feasibility of the regression analysis of the Fourier 1 function. The TELEs of the 5.80 mm PSM-1 specimen are shown in Fig. 23.5. It is obvious that TELEs almost coincide to the linear fit results.

To investigate the proportional relations in various thicknesses, the 6.00 mm thickness PSM-1 specimen was selected as the reference specimen. The comparison of $\frac{t_1}{6.00}$, $\frac{A_{6.00}}{A_{t_1}}$, and $\frac{B_{6.00}}{B_{t_1}}$ of the first TELE was made and results are summarized in Table 23.1. As indicated in Table 23.1, for each t_1, the values of $\frac{t_1}{6.00}$, $\frac{A_{6.00}}{A_{t_1}}$, and $\frac{B_{6.00}}{B_{t_1}}$ are very close. The largest difference is 2.29 %. Consequently, the relationship between SQF and thickness is thus verified.
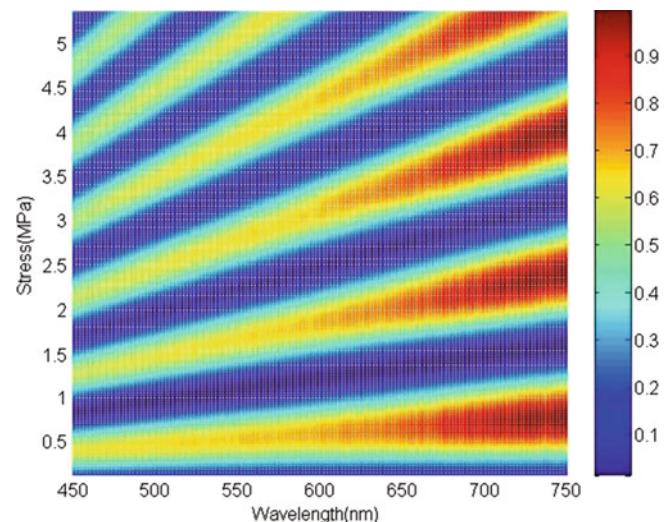
Fig. 23.3 The transmissivity spectrogram corresponding to Ss of 5.80 mm thickness PSM-1 specimen
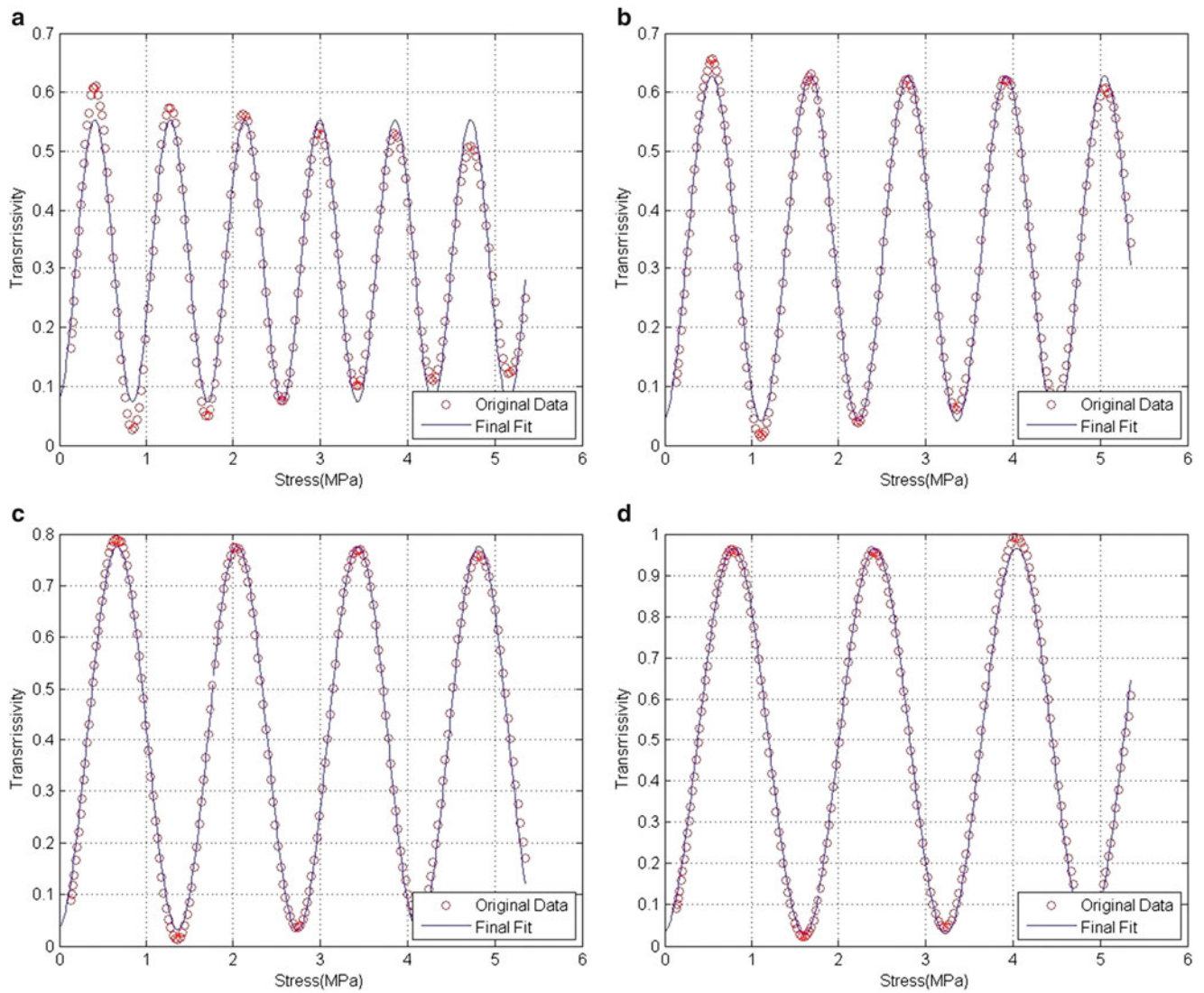
Fig. 23.4 The fitting results of Fourier 1 function of TSs of different wavelengths for 5.80 mm PSM-1 specimen. (**a**) 450 nm, (**b**) 550 nm, (**c**) 650 nm, (**d**) 750 nm

Fig. 23.5 The TELEs of 5.80 mm PSM-1 specimen
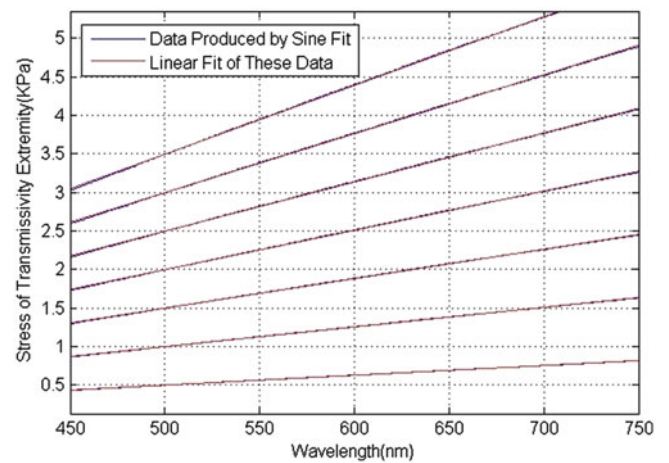
Table 23.1 Comparison of $\frac{t_1}{6.00}$, $\frac{A_{6.00}}{A_{t_1}}$, and $\frac{B_{6.00}}{B_{t_1}}$ of the first TELE of different thicknesses for PSM-1 specimens

t (mm)	$\dfrac{t}{t_{6.00}}$	$\dfrac{A_{6.00}}{A_{t_1}}$	Diff. (%)	$\dfrac{B_{6.00}}{B_{t_1}}$	Diff. (%)
3.00	0.500	0.505660	1.13	0.509063	1.81
5.80	0.967	0.945882	−2.18	0.947864	−1.98
9.75	1.625	1.652055	1.66	1.662208	2.29

Fig. 23.6 The transmissivity spectrogram corresponding to Ss of 4.85 mm glass specimen
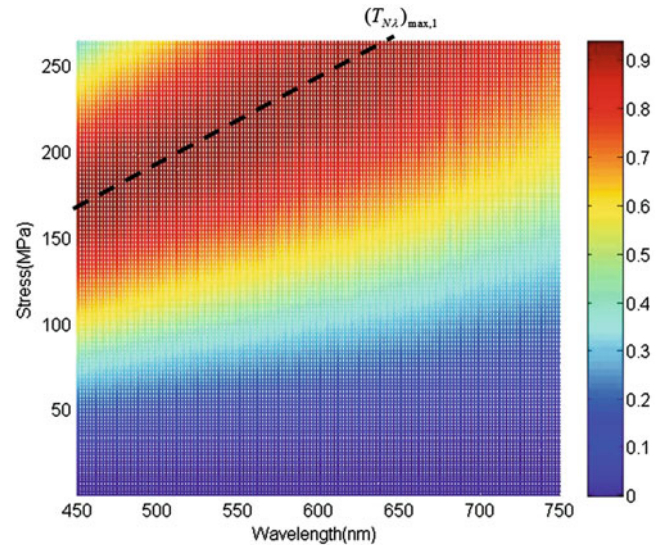
23.4.2 Results of Teles of Glass Material

Figure 23.6 is the transmissivity spectrogram corresponding to Ss of glass specimen. The specimen thickness is 4.85 mm. The maximum load applied is 2,500 N which is much larger than the load applied on the PSM-1 material, however, because of the low birefringence of the glass material, only the data of low fringe order were obtained.

The fitting results of Fourier 1 function of TSs at 450, 550, 650, and 750 nm wavelengths are shown in Fig. 23.7. All results match well with the original data. The TELE of 4.85 mm thickness glass specimen is shown in Fig. 23.8. It is obvious that the difference between the position of TE of regression analysis of Fourier 1 function and the linear fitting results of TELE becomes larger when the wavelength increases.

As depicted in Fig. 23.7, different peak values of transmissivity appear at 450, 550, and 650 nm, respectively. However, no peak value appears at 750 nm. Thus this no appearance of peak value at 750 nm produces difference between the fitting results of Fourier 1 function and the experimental results. In the analysis of the proportional relations in various thicknesses, the 4.85 mm thickness specimen was chosen as the reference specimen. The comparison of $\frac{t_1}{4.85}$, $\frac{A_{4.85}}{A_{t_1}}$, and $\frac{B_{4.85}}{B_{t_1}}$ of the first TELE was analyzed and the results are summarized in Table 23.2. As indicated in Table 23.2, for each t_1, $\frac{t_1}{4.85}$, and $\frac{A_{4.85}}{A_{t_1}}$ are quite close; however, −6.94 and −4.64 % difference between $\frac{t_1}{4.85}$ and $\frac{B_{4.85}}{B_{t_1}}$ were found for t_1 of 1.85 and 7.75 mm, respectively. Again, the difference is caused by the no appearance of a peak value at 750 nm so that the first TELE cannot be obtained. For all engineering purposes, those two relatively larger differences are still acceptable. Essentially speaking, the relationship between SQF and thickness for the glass material is again verified. It is believed that with the improvement of the experimental setup, the relationship between SQF and thickness proposed in this paper can be again substantiated for the glass material.
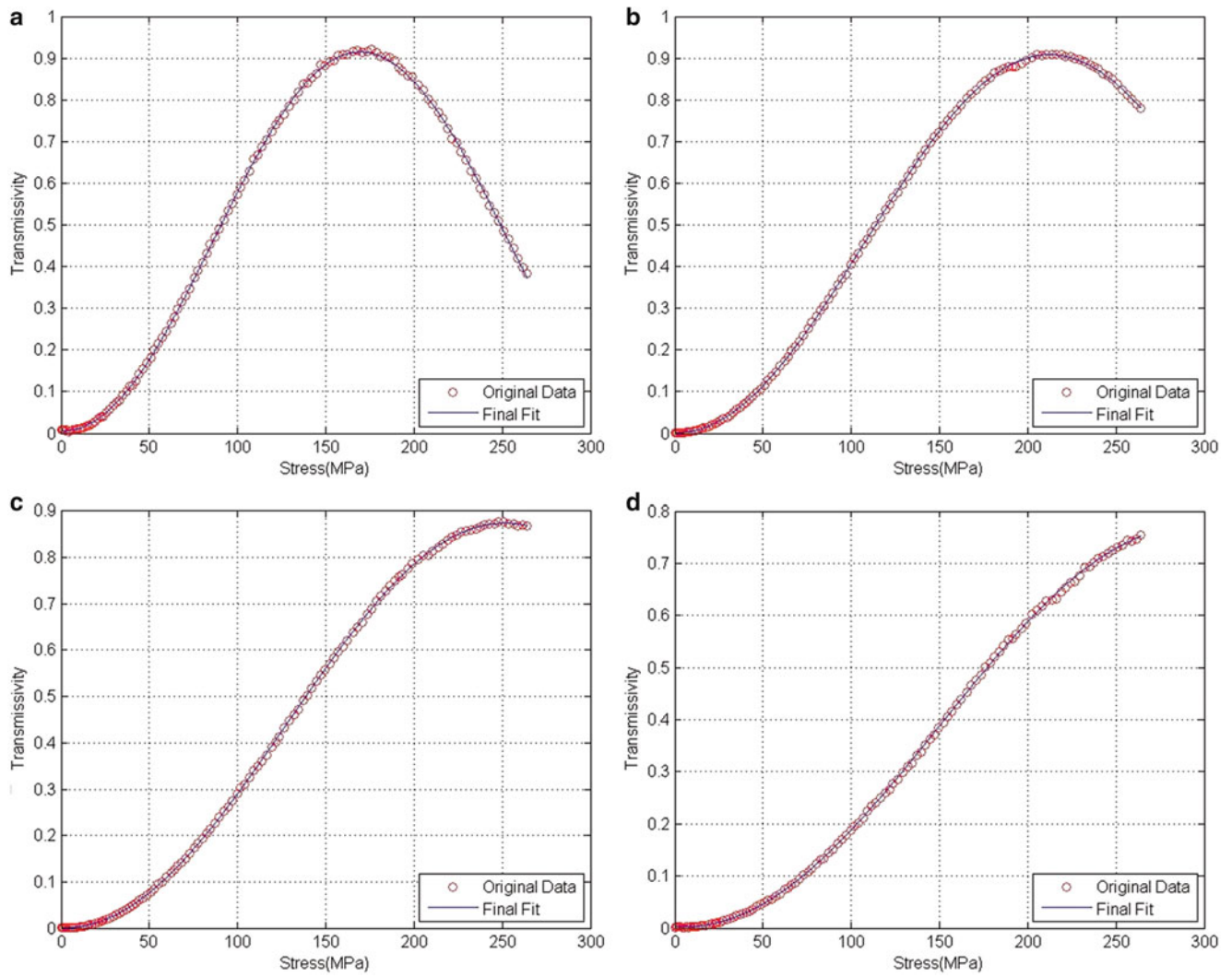
Fig. 23.7 The fitting results of Fourier 1 function of TSs of different wavelengths for 4.85 mm glass specimen. (**a**) 450 nm, (**b**) 550 nm, (**c**) 650 nm, (**d**) 750 nm

Fig. 23.8 The TELEs of 4.85 mm glass specimen
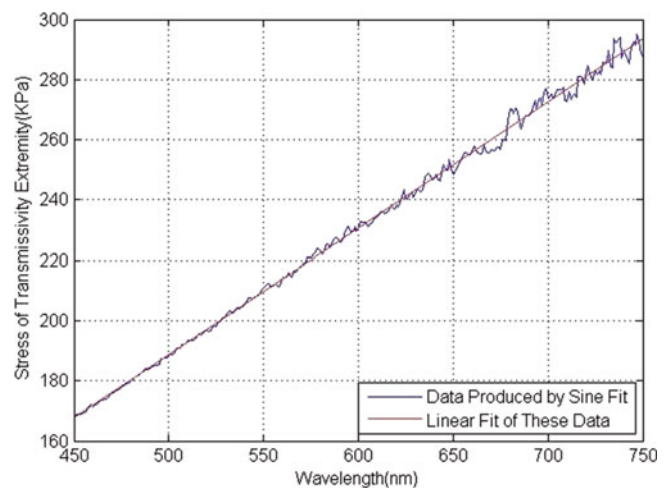
Table 23.2 Comparison of $\frac{t_1}{4.85}$, $\frac{A_{4.85}}{A_{t_1}}$, and $\frac{B_{4.85}}{B_{t_1}}$ of the first TELE for glass specimens of different thicknesses

t (mm)	$\dfrac{t}{t_{4.85}}$	$\dfrac{A_{4.85}}{A_{t_1}}$	Diff. (%)	$\dfrac{B_{4.85}}{B_{t_1}}$	Diff. (%)
1.85	0.381443	0.376767	−1.23	0.354953	−6.94
3.85	0.793814	0.797132	0.42	0.796476	0.34
7.75	1.597938	1.585356	0.51	1.485296	−4.64

23.5 Conclusions

In this paper, systematic 3DRoTwSW was successfully reconstructed by using TEToP. Based on the regression analysis of the Fourier 1 function, parameters $A\lambda + B$, $T_{a\lambda}$, and $T_{b\lambda}$ can be obtained from the database. Moreover, the TELEs can be derived by linear fitting TEs of different orders with the relationship between S and $T_{N\lambda}$.

The proportional relationship of TELEs and the relation between ratios of thickness and TELE of different thicknesses of specimens were verified by using PSM-1 material and glass material. Therefore, the generality of TEToP was proved.

References

1. Bureau of Foreign Trade (2013) Information and communication technology industry development program. Bureau of Foreign Trade, Ministry of Economic Affairs, Taiwan
2. Frocht MM (1984) Photoelasticity, vol 2. Wiley, New York, Chapter 4
3. Ramesh K (2000) Digital photoelasticity: advanced techniques and applications. Springer, Berlin
4. Ajovalasit A, Barone S, Petrucci G (1998) A review of automated methods for the collection and analysis of photoelastic data. J Strain Anal Eng Des 33(2):75–91
5. Petrucci G (1997) Full field automatic evaluation of an isoclinic parameter in white light. Exp Mech 37(4):420–426
6. Liu T, Asundi A, Boay CG (2001) Full-field automated photoelasticity using two-load-step method. Opt Eng 40(8):1629–1635
7. D'Acquisto L, Petrucci G, Zuccarello B (2002) Full field automated evaluation of the quarter wave plate retardation by phase stepping technique. Opt Lasers Eng 37:389–400
8. Wakayama T, Otani Y, Umeda N (2005) Real-time measurement for birefringence dispersion using double retarder. Proc SPIE 5888:588807/1–588807/6
9. Ajovalasit A, Barone S, Petrucci G (2007) Phase shifting photoelasticity in white light. Opt Lasers Eng 45:596–611
10. Wakayama T, Otani Y, Umeda N (2008) One-shot birefringence dispersion measurement based on channeled spectrum technique. Opt Commun 281:3668–3672
11. Otani Y, Wakayama T, Oka K, Umeda N (2008) Spectroscopic Muller matrix polarimeter using four-channeled spectra. Opt Commun 281:5725–5730
12. Otani Y, Wakayama T (2008) Two-dimensional measurement of birefringence dispersion. In: Proceedings of the international symposium to commemorate the 60th anniversary of the invention of holography, Springfield, MA, 27–29 Oct, pp 178–182
13. Wang WC, Hwang CH, Sung PC, Chen WR, Lai GT (2012) An apparatus for quantifying unknown stress and residual stress of a material and a method thereof. USA invention patent, patent application number: 1006085
14. Sung PC, Wang WC, Hwang CH, Lai GT (2012) A novel stress measurement method by integrating white light photoelasticity and spectrometry. In: 2012 IUTAM symposium on advances of optical methods in experimental mechanics, paper no. G-10, Taipei, Taiwan, 3–6 Nov 2012, pp 1–11
15. Website: http://www.mathworks.com/

Chapter 24
Observation of Fiber-Matrix Interfacial Stresses Using Phase-Stepping Photoelasticity

Takenobu Sakai, Yasunori Iihara, and Satoru Yoneyama

Abstract In this study, the single fiber composite was used to evaluate the stress transfer between a single fiber and a matrix. Single steel fiber was inserted to the epoxy resin, and it was applied the tensile load. Applied load was from 0 N to 177 N, and then the photoelastic images were taken by a digital CCD camera. In the photoelastic analysis, the phase-stepping and phase-unwrapping technique were used to measure the stress difference distribution more accurately. Obtained data were included the 3D information, therefore, we converted the stress difference data, including the 3D information to 2D photoelasticity using the geometrical information. The maximum shear stress was placed on the side surface of the steel fiber near the end of fiber, and the maximum normal stress difference was measured near the surface of the end of the fiber. The results of stress distribution on the interface between the fiber and the resin indicated that the stress concentration was observed in the middle part of the end of the fiber and the side surface near the end of the fiber.

Keywords Photoelasticity • Single fiber composites • Stress measurement • Optical method • Stress distribution • Phase-stepping • Phase-unwrapping

24.1 Introduction

Composite materials are used in various products, especially for automobile, airplane and aerospace industry. For product design, a lot of engineers use the FEM techniques in service. However, it is still necessary to obtain experimentally more information about the Micromechanics near an interface under load and to describe the stress transfer at the interface more precisely.

CFRPs are consisted with the epoxy resin for the matrix and the carbon fiber, and they have the fiber/matrix interface. CFRPs have the anisotropic properties, therefore, they are strong for tensile load, but are weak for the bending and compression load. They are sometimes fractured by the impact load as the falling object, the footprint of engineers, and so on. These CFRPs have the fiber edges in the composites because of their fibers being a failure at the inside of composites. In such situation, stress concentration will occur around the fiber ends. Of course, it is important not only to know the strength of the composites, but also to know about the stress concentration state around the fiber ends is required for the product reliability. For the research about the fiber/matrix interface [1, 2], there are many papers considering the approximation theory, therefore, it is necessary to analyze in more detail.

Cox gave the first solution of fiber stress distribution along the fiber/matrix interface with both the matrix and fiber undergoing elastic deformation [3]. After that, many researchers have used the photoelasticity for measuring experimentally the stress field in a matrix and the stress distribution at the interface for single-fiber epoxy model composites. However, it is necessary for the more detail analyze because of their approximated data [2, 4–7].

In this study, the photoelastic measurements, that can measure the stress distribution in full fields and the stress concentration, were used to reveal the stress information around the fiber/matrix interface of the single fiber composite.

T. Sakai (✉)
Department of Mechanical Engineering, Tokyo Metropolitan University, Hachioji, Tokyo, Japan
e-mail: t-sakai@tmu.ac.jp

Y. Iihara • S. Yoneyama
Department of Mechanical Engineering, Aoyama Gakuin University, Sagamihara, Kanagawa, Japan

And the phase shift method and the phase connect method were used to measure them with high accuracy. The final goal of this study is to know the stress concentration of epoxy resin around the carbon fiber experimentally in the microscopic view. This paper shows the fundamental study for the final goal as using the steel fiber instead of the carbon fiber.

24.2 Photoelastic Measurement

24.2.1 Phase-Stepping Method

An arrangement of optical elements in a polariscope is shown in Fig. 24.1. This setup consists of a monochromatic light source, a frosted glass, a linear polarizer P_1 whose optical axis is vertical, 2 quarterwave plates Q_1 and Q_2 whose fast axes make angles β and γ with the ox axis (horizontal axis) respectively, a linear polarizer (analyzer) whose optical axis makes an angle θ with the ox axis, 3 interference filters (500, 550 and 600 nm) and a monochromatic CCD camera. Between 2 quarterwave plates, a birefringent material as a single fiber embedded epoxy resin with retardation δ whose fast axis subtends as angle φ with the ox axis.

The angle φ of the principal axis of the specimen is interpreted as the principal stress direction, i.e., the isoclinic parameter. Similarly, the retardation δ of the specimen, that is, the isochromatic parameter relates the principal stress difference as

$$\delta = 2N\pi = 2\pi \frac{C_\sigma d}{\lambda}(\sigma_1 - \sigma_2) \tag{24.1}$$

where N is the isochromatic fringe order, C_σ is the stress-optic coefficient, d the thickness of the specimen, λ is the wavelength of the monochromatic incident light, and σ_1 and σ_2 are the principal stresses.

For the phase-stepping method, seven combinations of the angular positions β, γ and θ of the retarders and the analyzer are used. The arrangements of β, γ and θ, and the corresponding light intensities $I_1 \sim I_7$ used are shown in Table 24.1. The amplitude and the background bias are omitted in this table. Using the seven light intensity values $I_1 \sim I_7$ are also shown in Table 24.1. The isoclinic parameter (principal direction) φ, the retardation Δ of the retarder, and the isochromatic parameter (retardation) δ can be obtained as [8]

Since the retardation Δ must be the positive value, the sign of the function $\sin\Delta$ in Eq. (24.4) can be determined as positive. Using Eq. (24.2) to (24.4), the phase values δ and φ of the isochromatic and isoclinic parameters are determined
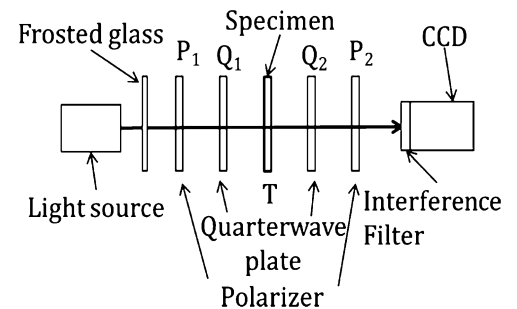
Fig. 24.1 Arrangement of optical elements

Table 24.1 Optical arrangements and light intensity equations

	β,rad	γ,rad	θ,rad	Light Intensity, I
I_1	$\pi/4$	0	0	$I_1 = \frac{1}{2}\{1 + \sin\delta\sin\Delta\sin2\varphi - \cos\Delta(\cos2\varphi + \cos\delta\sin^2 2\varphi)\}$
I_2	$-\pi/4$	0	0	$I_2 = \frac{1}{2}\{1 - \sin\delta\sin\Delta\sin2\varphi - \cos\Delta(\cos2\varphi + \cos\delta\sin^2 2\varphi)\}$
I_3	$\pi/4$	$-\pi/4$	$-\pi/4$	$I_3 = \frac{1}{2}\left(1 + \sin\delta\sin\Delta\cos2\varphi + \cos\Delta\sin^2\frac{\delta}{2}\sin4\varphi\right)$
I_4	$-\pi/4$	$-\pi/4$	$-\pi/4$	$I_4 = \frac{1}{2}\left(1 - \sin\delta\sin\Delta\cos2\varphi + \cos\Delta\sin^2\frac{\delta}{2}\sin4\varphi\right)$
I_5	$-\pi/4$	$-\pi/4$	$\pi/2$	$I_5 = \frac{1}{2}\{1 - \cos\delta\sin^2\Delta + \sin\delta\sin2\Delta\sin2\varphi + \cos^2\Delta(\cos^2 2\varphi + \cos\delta\sin^2 2\varphi)\}$
I_6	$-\pi/4$	$\pi/4$	$\pi/2$	$I_6 = \frac{1}{2}\{1 + \cos\delta\sin^2\Delta + \cos^2\Delta(\cos^2 2\varphi + \cos\delta\sin^2 2\varphi)\}$
I_7	$\pi/4$	$\pi/4$	$\pi/2$	$I_7 = \frac{1}{2}\{1 - \cos\delta\sin^2\Delta - \sin\delta\sin2\Delta\sin2\varphi + \cos^2\Delta(\cos^2 2\varphi + \cos\delta\sin^2 2\varphi)\}$

even if the retardation Δ of the retarder is unkown. This method is not necessarily a quarterwave plate specified for the wavelength of the light used.

$$\tan 2\varphi = \frac{I_1 - I_2}{I_3 - I_4} = \frac{\sin \delta \sin \Delta \sin 2\varphi}{\sin \delta \sin \Delta \cos 2\varphi} \tag{24.2}$$

for, $\sin\delta \neq 0$

$$\cos \Delta = \frac{I_5 - I_7}{2(I_1 - I_2)} = \frac{\frac{1}{2} \sin \delta \sin \Delta \sin 2\varphi}{\sin \delta \sin \Delta \sin 2\varphi} \tag{24.3}$$

for, $\sin\delta \neq 0$ and $\sin 2\varphi \neq 0$

$$\tan \delta = \frac{2(I_3 - I_4) \sin \Delta}{(-I_5 + 2I_6 - I_7)} = \frac{\sin \delta \sin^2 \Delta \cos 2\varphi}{\cos \delta \sin^2 \Delta \cos 2\varphi} \tag{24.4}$$

for, $\cos 2\phi \neq 0$

24.2.2 Phase-Unwrapping

Corrected isochromatic parameters for each wavelength are represented as δ_{c1}, δ_{c2} and δ_{c3}. The unwrapped phase δ'_i ($i = 1,2,3$) can be expressed as [9]

$$\delta'_i = \delta_{ci} + 2N_i\pi \tag{24.5}$$

where Ni is the fringe order. Then using Eq. (24.1), the unwrapped phase values at all points satisfy the equation as;

$$\frac{\lambda_i}{C_i} \delta'_i = const. \tag{24.6}$$

where, Ci is the stress-optic coefficient for each wavelength. Substituting Eq. (24.5) into Eq. (24.6), a function fi can be defined as;

$$f_i = \frac{\lambda_i}{C_i} (\delta_{ci} + 2N_i\pi) = const. \tag{24.7}$$

Phase unwrapping can be performed by searching fringe order N_i, which is consistent with Eq. (24.7). The error function E is defined as

$$E = (f_1 - f_2)^2 + (f_2 - f_3)^2 + (f_3 - f_1)^2 \tag{24.8}$$

By finding the set of fringe orders N_i, which make the value of E minimum, the unwrapped phase δ'_i can be obtained using Eq. (24.5). In searching the set of N_i, taking the inequalities

$$N_1 > N_2 > N_3 \tag{24.9}$$

into account, the set of N_i that does not satisfy the inequalities is excluded.

Fig. 24.2 Schematic
drawing of the Optical path
and the cross section of the
single-fiber composite
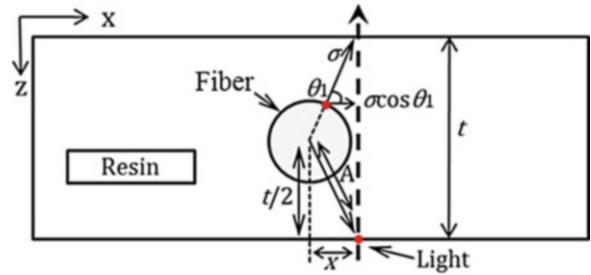
24.2.3 Conversion from 3D to 2D Photoelastic Information

In the present study, single-fiber composite was used to evaluate the stress distribution around the fiber and on the fiber/matrix interface. The specimens have the thickness as 2.3 mm, and the used steel fiber's diameter is 0.7 mm, therefore, it must be considered the thickness of the composite on the optical path direction.

θ_1 shows the minimum angle to the point on the optical path that have the distance x from the center of the steel fiber as shown in Fig. 24.2. If there is the normal stress σ on the arbitrary point on the fiber/matrix interface, the stress on the optical path through the specimen can be described as $\sigma\cos\theta_1$.

It is assumed that the normal stresses are generated at the every point on the fiber surface. The stress obtained by 2D-photoelasticity on the optical path shows the average value of the integral of the each stress at each point. It is, therefore, necessary to convert from the stress obtained in 2D-photoelasticity to the stress considered the thickness direction for 3D-information.

On the optical path that penetrates a certain point on the specimen, the stress obtained from the isochromatic parameters considering the thickness t can be described as follows;

$$\sigma' = \frac{1}{t}\int_0^t \sigma\cos\theta_1 dt \tag{24.10}$$

where, σ' is the stress in 2D, σ is the stress in 3D and t is the thickness of the specimen. At first, θ_1 is shown as the function of t,

$$\sin\theta_1 = \frac{\frac{t}{2}}{A} = \frac{t}{2A} \tag{24.11}$$

$$\therefore \ \theta_1 = \sin^{-1}\frac{t}{2A} \tag{24.12}$$

Using the Pythagorean theorem, A can be determined with the distance x and thickness t. Because A is constance, by differentiation with θ, Eq. (24.11) will be;

$$\cos\theta_1 = \frac{dt}{d\theta} \tag{24.13}$$

Equations (24.12), (24.13) are substituted for Eq. (24.10), then

$$\sigma' = \frac{1}{t}\int_{-\sin^{-1}\left(\frac{t}{2A}\right)}^{\sin^{-1}\left(\frac{t}{2A}\right)} \sigma\cos^2\theta_1 d\theta \tag{24.14}$$

$$\therefore \int \cos^2\theta d\theta = \frac{\theta}{2} + \frac{\sin 2\theta}{4} + C' \tag{24.15}$$

where, C' is the constant of integration. Equation (24.15) is the definite integral, therefore, it can be ignored the definite integral. Then σ' can be described by σ. In the present paper, σ' was used to evaluate the stresses on the side surface of the fiber/matrix interface and the around of the fiber.

Fig. 24.3 Tensile testing system and observation area in composite
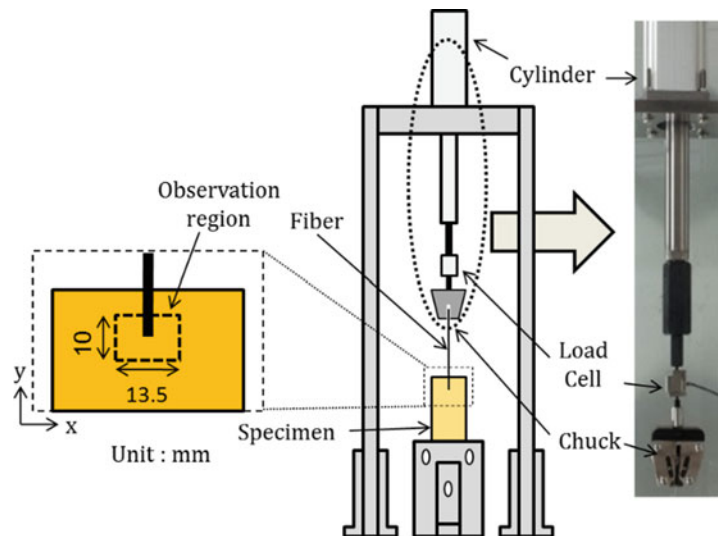
Table 24.2 Fringe order and photoelastic constant in each wavelength

Wavelength [nm]	500	550	600
Fringe order	2.80	2.47	2.21
Photoelastic constant [mm^2/N]	1.220×0^{-5}	1.199×10^{-5}	1.158×10^{-5}

24.2.4 Materials Preparation and Experimental Procedure

Used composites were consisted with the epoxy matrix (Epikote 828, 871 and the cure agent HN-2200 as the ratio of 5:5:8) and the steel fibers (φ 0.7 mm). The single fiber was embedded to epoxy resin by 10 mm as shown in Fig. 24.3.

The single-fiber composite specimens were subjected to uniaxial tension to investigate the stress distribution around the interface of matrix and fiber. The used tensile testing system is shown in Fig. 24.3. Observation area in the composite is 10×13.5 mm around the fiber end. The loading was interrupted temporarily at various levels of applied matrix stress to capture the birefringence patterns at the fiber-end.

The photoelastic fringe patterns are collected by a monochromatic CCD camera with a resolution of 640×480 pixels and 256 gray levels. Three monochromatic lights of wavelengths 500 nm, 550 nm and 600 nm emitted from a halogen lamp with interference filters are sued as the light source in order to perform the tricolor method [10]. The average retardation value of Δ of the retarder obtained by Eq. (24.3) are evaluated as 1.66 rad for the wavelength 500 nm, 1.54 rad for 550 nm and 1.48 rad for 600 nm. The fringe order and the photoelastic constant in each wavelength were obtained by the 4-point bending tests, and their results are shown in table 24.2.

24.2.5 Matrix Stress Field from Photoelastic Analysis

For the tensile tests of the single-fiber composites, applied loads were 0 N and 177 N. To consider the residual stress that was occurring when the composite was fabricated, the photoelasticity in the 0 N condition was investigated. In 177 N condition, the crosshead was stopped, and then the photographs were taken immediately. It took a few minutes to take them, therefore, the viscoelastic behavior was almost not occurred during taking photos.

To make the isochromatic and isoclinic images, 7 photographs based on Table 24.1 for each wavelength were taken and phase-stepped. Figure 24.4 shows (a) and (d) isochromatic, (b) and (e) isoclinic and (c) and (f) corrected retardation with ambiguity images for 0 N ((a), (b) and (c)) and 177 N ((d), (e) and (f)) of wavelength as 550 nm. And three-wavelength method [10] was used to correct the retardation where ambiguity exists.
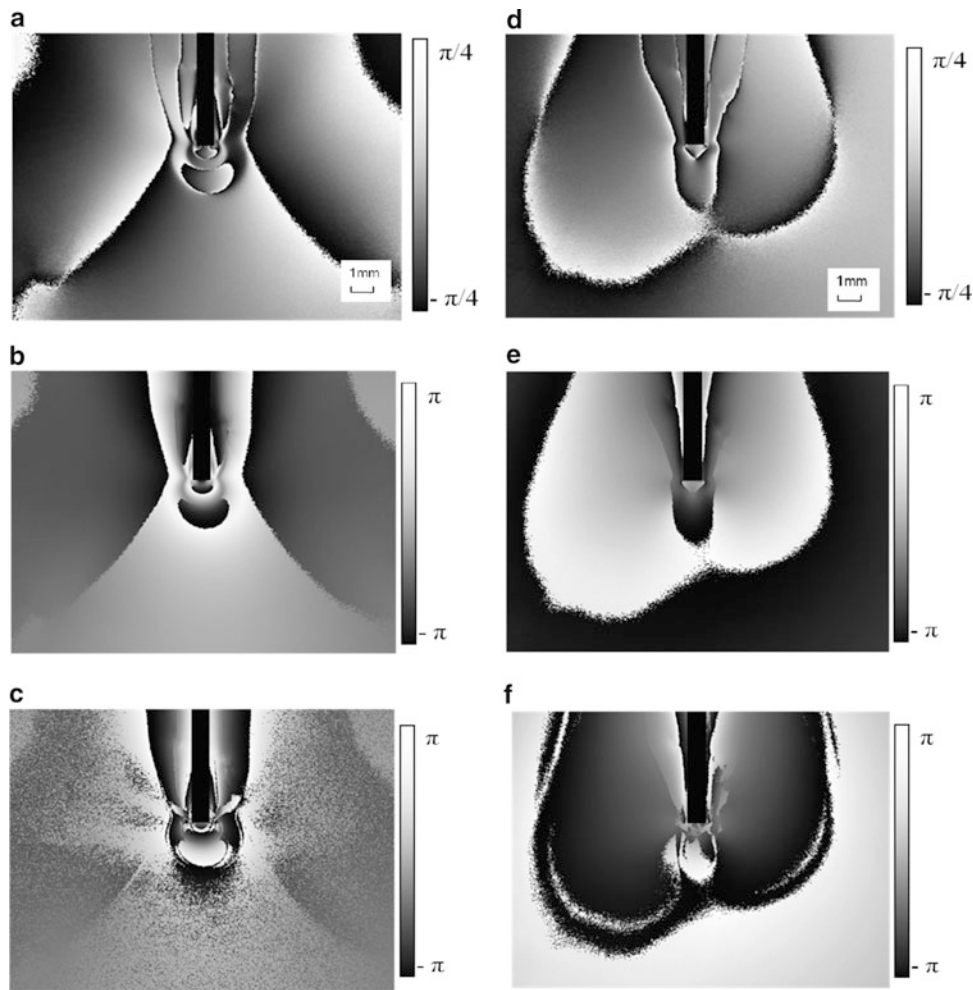
Fig. 24.4 (**a**), (**d**) Wrapped principal direction, (**b**), (**e**) Wrapped retardation, (**c**), (**f**) Corrected retardation with ambiguity. (**a**), (**b**) and (**c**) are under 0 N, and (**d**), (**e**) and (**f**) are under 177 N. Every images are under 550 nm wavelength

Isoclinic images under 0 and 177 N, which were obtained by phase-stepped and phase-unwrapped, are shown in Fig. 24.5a, b, respectively. These isoclinic images show the principal direction of stress. The change in direction became more complex near the fiber ends, therefore, the principal stress in matrix changed suddenly around the fiber.

The stress difference distribution around the fibers under 0 N and 177 N are shown in Fig. 24.5c, d, respectively. As shown in this figure, the principal stress difference concentration was occurring around fiber on each load condition. That is why the residual stress was occurring when the single fiber composite was fabricated. The stress difference gradually increased along the side surface of the fiber from the bottom to the end of the fiber, and the difference under 177 N (36 ~ 39 MPa) was higher than under 0 N (28 ~ 38 MPa). The maximum stress difference around fiber is 38 MPa for 0 N and 39 MPa for 177 N. It was observed that the stress difference were increased around the fiber end, and the part of the stress concentration moved from the end of the fiber to the side when the load was applied.

24.3 Results and Discussion

24.3.1 *Shear Stress and Normal Stress Difference*

To remove the effect of the residual stress on the stress distribution for discussing the stress concentration around the fiber caused by the applied stress, the stress value of before loading must be subtracted from that of after loading. Therefore, the normal stress difference and the shear stress are calculated using the principal stress difference and principal direction. The normal stress difference and the shear stress are
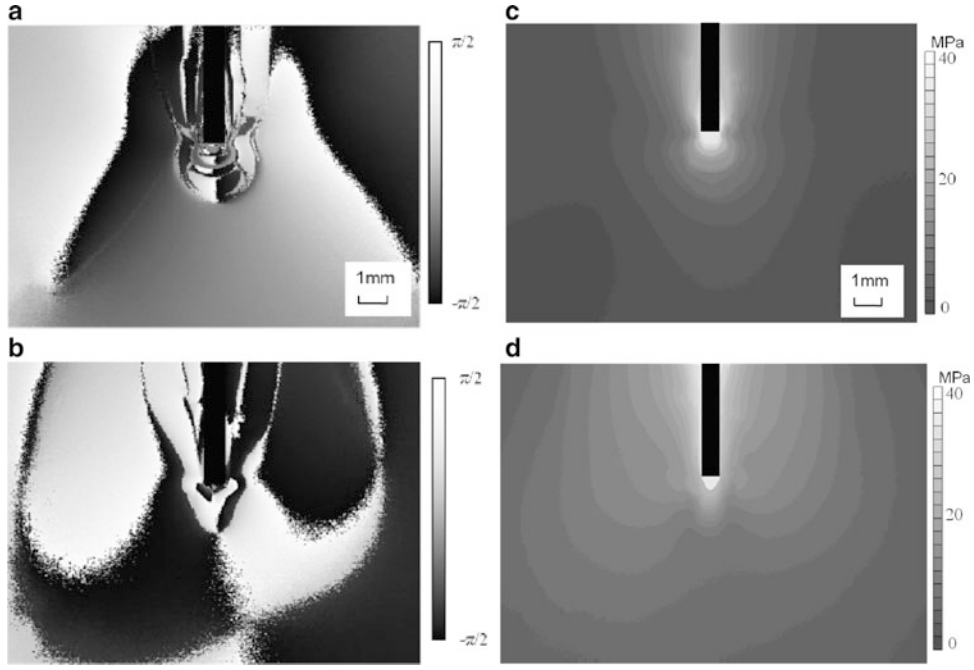
Fig. 24.5 (**a**), (**b**) Isoclinic images under 0 N and 177 N, (**c**), (**d**) principal stress difference distribution under 0 N and 177 N
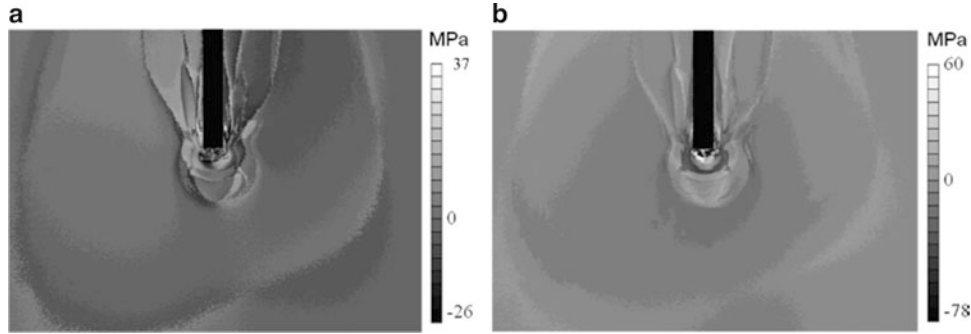
Fig. 24.6 Stress distribution map considered the residual stress (**a**) τ_{xy} and (**b**) $(\sigma_y - \sigma_x)$

$$\tau_{xy} = \frac{1}{2}(\sigma_1 - \sigma_2)\sin 2\phi \tag{24.16}$$

$$(\sigma_y - \sigma_x) = (\sigma_1 - \sigma_2)\cos 2\phi \tag{24.17}$$

where, shear stress is τ_{xy}, the normal stress difference is $(\sigma_y - \sigma_x)$, the principal stress difference is $(\sigma_1 - \sigma_2)$ and principal stress direction is ϕ. Figure 24.6 shows (a) shear stress and (b) normal stress difference. As shown in these figures, the strong shear stresses are distributed around the end and the small area near the side and the end of fiber, and the maximum shear stress are at a distance of 5 mm from the end of the fiber.

The maximum normal stress difference is placed at a distance of ~0.4 mm from the edge of the fiber, and the high level normal stress difference is distributed along the side wall of the fiber. Therefore, in the big picture, it is revealed that the shear stress is dominant at the side of the fiber, and the normal stress difference is mainly placed at the end of the fiber.
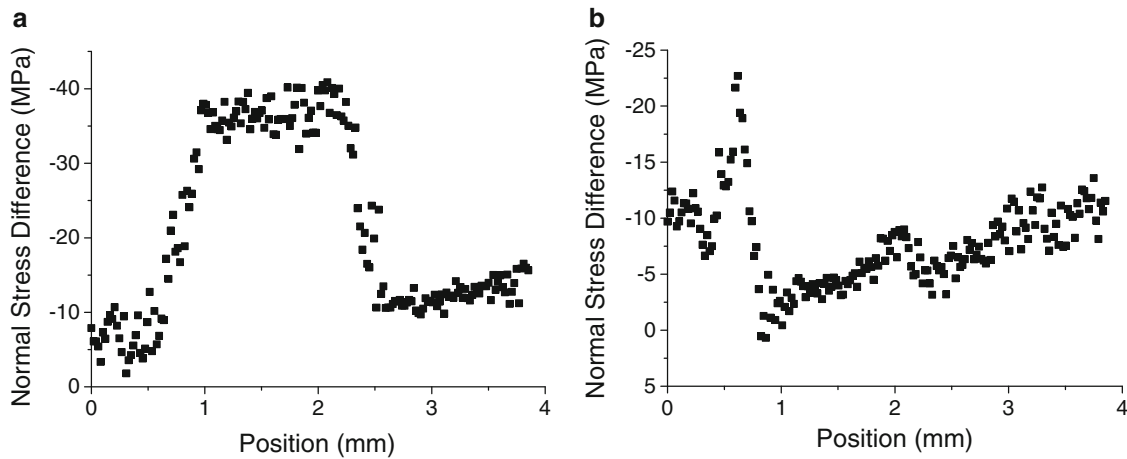
Fig. 24.7 Normal stress difference ($\sigma_y - \sigma_x$) distribution along the fiber at (**a**) *left side*, (**b**) *right side*
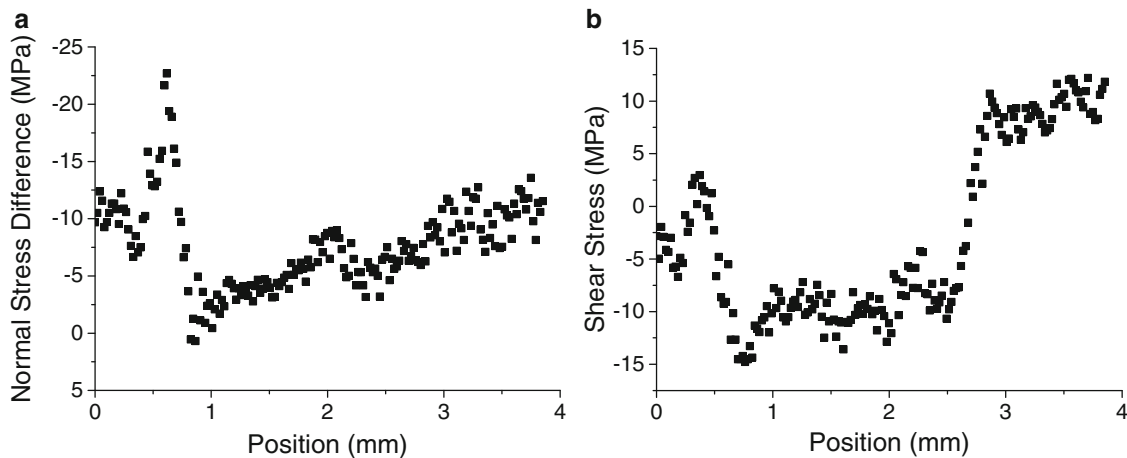
Fig. 24.8 Shear stress difference τ_{xy} distribution along the fiber at (**a**) *left side*, (**b**) *right side*

24.3.2 Stress Distribution at the Fiber/Matrix Interface

In the last section, the stress value was calculated from the isochromatic and isoclinic images, therefore, it could be considered the direction of the stresses. In this section, the stresses were converted from 2D to 3D value using Eq. (24.15), and the stress distribution of the full field was not obtained, but the stress distributions on the fiber/matrix interface were calculated. To consider the effect of loading on the stress distribution around the fiber, the subtracted value from the data of 177 N to 0 N was discussed.

Figure 24.7 shows the normal stress distribution on the side wall of the fiber. As shown in these figures, there are big difference between the left and the right side surface. The difference might be occurred by the manufacturing process. When the making of single fiber composites, the emitted fiber is slightly bent, the distributions had the difference between the both sides. On the left side, there is the stress concentration ranged from 0.5 to 2.3 mm from the end of the fiber, and on the right side, the stress concentration is placed only at 0.5 mm from the end. It is assumed that the adhesion between fiber and composite is not satisfied with the right side. However the shear stress distribution (in Fig. 24.8) showed the almost same distribution on the both sides. Therefore, the stress will concentrate on the distance of the short length from the end of the fiber.

Figure 24.9 shows the stress distribution at the end surface of the fiber. On the corner of the end surface, there are compressive stresses, however the each stress was increased toward the center of the fiber. As shown in this results, stress concentration area is different from the results obtained by Zhao et., al. [6]. This is because that this area considered no 3D
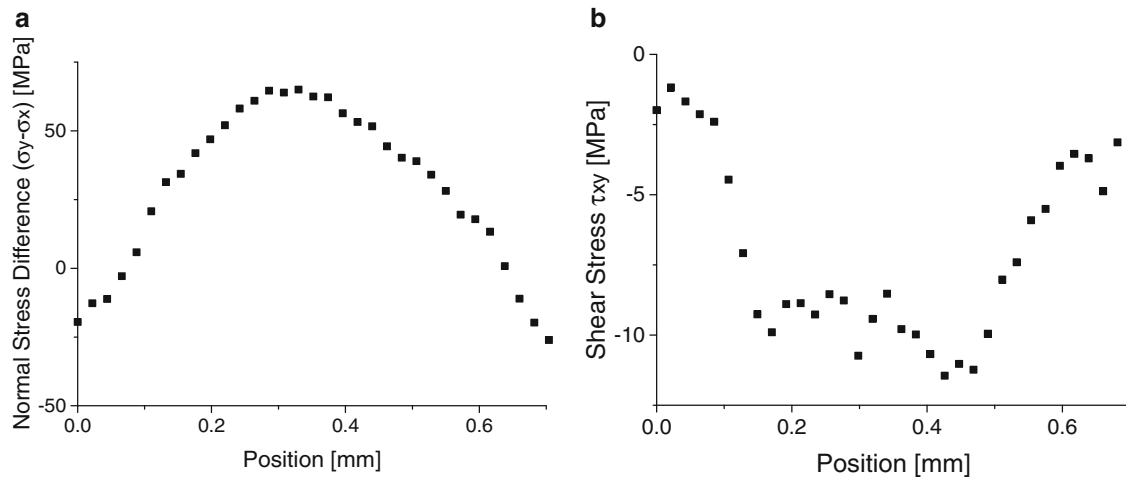
Fig. 24.9 Stress difference along the end surface of the fiber (**a**) ($\sigma_y - \sigma_x$) (**b**) τ_{xy}

information. For our future work, we have to consider the 3D information in fiber ends and applicability for the Abel inverse transform.

Consequently, using this photoelastic technique, the stress distribution around the single fiber in the composites was clarified. In the future works, several fibers will be embedded in the epoxy resin, and the photo-viscoelastic properties will be considered.

24.4 Conclusions

In this study, the stress distribution of the single fiber composite was obtained by the photoelasticity with the phase-stepping and phase-unwrapped technique. The stresses were calculated with the photoelastic and geometrical information. The stress distribution and stress direction, distribution, they showed not only the effect of the loading but also that of the manufacturing process. The normal stress difference and the shear stress without the effect of the residual stress were calculated and evaluated. The stress concentrated in the middle of the end of the fiber and the side surface at the distance slightly longer.

For our future work, we have to consider the 3D information in fiber ends and applicability for the Abel inverse transform.

References

1. Koyanagi J, Kawai J, Ogihara S, Watanabe K (2010) Carbon fiber/matrix interfacial shear strength evaluated by single fiber pull-out test considering effects of resin meniscus. Exp Mech 10(4):407–412
2. Zhao FM, Hayes SA, Young RJ, Jones FR (2006) Photoelastic study of the stress transfer in single fiber composites. Comp Interfaces 13-8-9:757–772
3. Cox HL (1952) The elasticity and strength of paper and other fibrous materials. Br J Appl Phys 3:72–79
4. Piggott MR (1991) Failure processes in the fibre-polymer interphase. Compos Sci Technol 42:57–76
5. Kelly A, Tyson WR (1965) Tensile properties of fibre-reinforced metal: copper/tungsten and copper/molybdenum. J Mech Phys Solids 13:329–350
6. Zhao FM, Hayes SA, Patterson EA, Young RJ, Jones FR (2003) Measurement of micro stress fields in epoxy matrix around a fibre using phase-stepping automated photoelasticity. Compos Sci Technol 63:1783–1787
7. Flores-Johnson EA, Vazquez-Rodriguez JM, Herrera-Franco PJ, Gonzalez-Chi PI (2011) Photoelastic evaluation of fiber surface-treatments on the interfacial performance of a polyester fiber/epoxy model composite. Compos A: Appl Sci Manuf 42:1017–1024
8. Yoneyama S, Kikuta H (2006) Phase-stepping photoelastisity by use of retarders with arbitrary retardation. Exp Mech 46:289–296
9. Yoneyama S, Nakamura K, Kikuta H (2009) Absolute phase analysis of isochromatics and isoclinic using arbitrary retarded retarders with tricolorimage. Opt Eng 48:123603
10. Tong L, Boay CG (2001) A new three-wavelength algorithm for extracting information of photoelastic parameters. Asian J Phys 10(1):87–94

Chapter 25

Stabilizing Heteroscedastic Noise with the Generalized Anscombe Transform: Application to Accurate Prediction of the Resolution in Displacement and Strain Maps Obtained with the Grid Method

M. Grédiac and F. Sur

Abstract The objective of this paper is to show that it is possible to predict the noise level in displacement and strain maps obtained with the grid method, but that actual noise of camera sensors being heteroscedastic, it is necessary to stabilize this noise in grid images prior to employing the predicting formulas. The procedure used for this purpose relies on the Generalized Anscombe Transform. This transform is first described. It is then shown that experimental and theoretical resolutions in strain maps obtained with the grid method are in good agreement when this transform is employed.

Keywords Displacement • Generalized Anscombe transform • Grid method • Noise • Strain measurement • Metrological performance

25.1 Introduction

Assessing the metrological performance of full-field measurement techniques is a key-issue in experimental mechanics. In this context, this paper deals with the a priori determination of the resolution in displacement and strain maps obtained with the grid method, the spatial resolution being fixed by the grid image processing procedure. Two points are addressed more specifically:

1. closed-form formulas which provide the link between resolution, spatial resolution and standard deviation of sensor noise in grid images taken by a camera are given and discussed. A specific predictive formula is obtained for both the displacement and for strain maps;
2. experiments are carried out on real grid images to check whether these formulas are verified.

As in most of the studies discussed in the literature dealing with noise prediction in displacement and strain maps obtained with various full-field measurement techniques, the closed-form formulas mentioned in point 1 above are obtained here assuming that noise in the grid images is white. This means that its characteristics (especially its variance) are constant throughout the grid images. In statistics, such a noise is referred to as *homoscedastic*.

The actual response of camera sensors is however not homoscedastic since its standard deviation significantly depends on brightness at any pixel [1], and the latter obviously changes from one pixel to another in any contrasted image. Thus actual noise in images grabbed by cameras is not homoscedastic but *heteroscedastic*. This leads any predictive formulae for the noise in final displacement or strain maps obtained under the assumption of homoscedastic noise not to be satisfied in practice, or to be at least more or less severely impaired by too rough this assumption.

Stabilizing the changing variance of a heteroscedastic noise, thus transforming a heteroscedastic noise into a homoscedastic one, is theoretically possible when this noise is e.g. Poisson-Gaussian. This specific transform is the so-called Generalized Anscombe Transform [2]. This transform involves some noise parameters of the camera which need therefore to be identified.

M. Grédiac (✉)
UMR CNRS 6602, Clermont Université, Université Blaise Pascal, Institut Pascal, BP 10448, 63000 Clermont-Ferrand, France
e-mail: michel.grediac@univ-bpclermont.fr

F. Sur
Laboratoire Lorrain de Recherche en Informatique et ses Applications, UMR CNRS 7503 Université de Lorraine,
CNRS, INRIA Projet Magrit, Campus Scientifique, BP 239 54506, Vandoeuvre-lès-Nancy, Cedex, France

The grid method, which is employed here to retrieve displacement and strain maps, is first briefly presented. The main steps of the GAT are then described. It is finally applied to a set of real images grabbed by a CCD camera. Other phenomena may corrupt images such as microvibrations or light flicker. They are eliminated using another transform which is detailed in another paper [3]. It is finally shown through experiments that applying these transforms to actual grid images leads the abovementioned formulas to be very satisfactorily verified, thus illustrating that noise in final displacement and strain maps can be accurately and reliably predicted.

25.2 Obtaining Displacement and Strain Maps the with Grid Method

The grid method consists first in depositing a crossed grid on the surface under investigation in order to track the change in the geometry of the grid while loading increases, and to deduce the 2D strain fields from these images. The typical value of the pitch (denoted p) is 0.2 mm. The procedure for depositing the grids is given in [4].

Processing grid images consists first in extracting the phases along directions x and y both in the reference and in the current images. Phase extraction is carried out with the Windowed Fourier Transform (WFT) [5], in which the envelope considered in the present study is Gaussian, as in [6]. The displacement is obtained from the phase changes between current and reference grid images using the following equation:

$$u_i = -\frac{p}{2\pi}\Delta\Phi_i, i = x, y \tag{25.1}$$

Thus

$$\varepsilon_{ij} = -\frac{p}{4\pi}\Delta\left(\frac{\partial\Phi_i}{\partial x_j} + \frac{\partial\Phi_j}{\partial x_i}\right), i, j = x, y \tag{25.2}$$

25.3 Predictive Formulas

After pioneering work in [7], [8] gives a closed-form expression for the variance of the noise in the phase or the phase derivative images from the variance of the noise in the grid image. These expressions are obtained assuming that:

- the grid images are corrupted by a homoscedastic Gaussian white noise of constant variance;
- the grid used as a spatial carrier does not feature any defect such as pitch variation, so the noise in the maps is only due to sensor noise;
- the envelope of the kernel used to extract the phase and phase derivative maps is a Gaussian function. Its standard deviation is denoted σ.

In [8], this result was establish to feed the Wiener filter used to deconvolve strain images, as in [9]. These predictive equations for the variance can easily be rewritten as links between resolution, spatial resolution, standard deviation of noise in the grid images, and various parameters characterizing the grid profile.

Resolution can be defined by the "change in quantity being measured that causes a change in the corresponding indication greater than one standard deviation of the measurement noise" [10]. Note that this quantity is not defined by any standard, so any multiple of one standard deviation could be considered: this would only change the equations given below. *Spatial resolution* is defined here by the shortest distance between two spatially independent measurements. Since a Gaussian window is used in the approach employed here, it is reasonable to define this quantity by a multiple of the standard deviation σ of the Gaussian window, say 6σ according to the three-sigma rule.

The links between all these quantities are as follows [11]

$$\begin{cases} \sigma_\Phi \times d_\Phi = \dfrac{3}{\sqrt{2\pi}} \times \dfrac{a}{K} \times \sigma_{im} & \text{(25.3a)} \\[4mm] \sigma_{\Phi'} \times d_{\Phi'}^2 = \dfrac{9}{\sqrt{\pi}} \times \dfrac{a}{K} \times \sigma_{im} & \text{(25.3b)} \end{cases}$$

where:

- σ_Φ is the resolution in the phase map as defined above;
- d_Φ is the spatial resolution in the phase map as defined above;
- a is the area of one pixel. Dimensions being given in pixels, a = 1 in the following;
- K is the modulus of the WFT used to extract the phase;
- σ_{im} is the standard deviation of the sensor noise assumed here to be homoscedastic.

Similar equations are deduced for both the displacement and the strain:

$$
\begin{cases}
\sigma_u \times d_u = p\dfrac{3}{2\pi^{3/2}} \times \dfrac{s}{K} \times \sigma_{im} & (25.4a) \\[3mm]
\sigma_\varepsilon \times d_\varepsilon^2 = p\dfrac{9\sqrt{2}}{2\pi^{3/2}} \times \dfrac{s}{K} \times \sigma_{im} & (25.4b)
\end{cases}
$$

where:

- σ_u is the resolution in the displacement map as defined above;
- d_u is the spatial resolution in the displacement map as defined above;
- p is the pitch of the grid deposited on the specimen.

Several remarks can be drawn from these equations:

- Equations (25.3a) and (25.4a) show that the metrological performance of the phase and the displacement determination are characterized by a relationship similar to that deduced from the Heisenberg uncertainty principle [12] in the sense that the resolution and spatial resolution cannot be simultaneously arbitrarily low;
- differentiation involved in the calculation of the phase derivative and strain components causes the spatial resolution to be squared in the left-hand side of Eqs. (25.3b) and (25.4b). Consequently, improving the spatial resolution by a factor 3 impairs the resolution by a factor 9. A similar property has already been discussed for the global performance of DIC in [13] for instance;
- both products are proportional to the standard deviation of the noise σ_{im} in the grid image, which seems quite logical. They are also inversely proportional to the modulus K of the WFT of the grid image. This modulus can easily be obtained at any pixel of the grid image since the WFT is calculated for the determination of the phase anyway. K is actually proportional to three quantities [11]: the brightness, the contrast of the grid and the first coefficient of the development of the grid profile in Fourier series, denoted d_1 here. Finding that noise is all the more magnified as brightness and contrast are low is logical. d_1 is maximum for a sine profile, so any printed grid should have a profile tending to a sine to minimize noise sensor propagation. If rectangular profile are employed, the black and white lines should at least have the same width [11];
- it is finally worth mentioning that in theory, the phase distribution retrieved by the WFT does not depend on the nature of the line profile of the grid (ignoring the) whereas the noise level in the displacement map depends on this profile. The same remark holds for the displacement.

25.4 Noise Model and Noise Normalization with the Generalized Anscombe Transform

A classic model for the camera response function is to consider that it is linear. It means that the grey level measured at a pixel is proportional to the quantity of light photons arriving at a CCD sensor cell, plus a dark current [1]:

$$
s(x, y) = g\eta_{p(x,y)} + \delta(x, y) \tag{25.5}
$$

where s is the observed signal and g is the gain of the electronic system. $\eta_{p(x,y)}$ is the number of collected light photons at a pixel (x,y). This is a random variable which follows a Poisson distribution of parameter p(x, y). This parameter is assumed to be spatially independently distributed. $\delta(x, y)$ models the dark current and readout noise. This is a Gaussian white noise of

mean μ_d and variance σ_d^2. $\eta_{p(x,y)}$ and $\delta(x, y)$ being assumed to be independent, we can deduce the following linear relation between the expectation $E(s(x, y))$ and the variance $Var(s(x, y))$ of $s(x, y)$:

$$Var(s(x, y)) = gE(s(x, y)) + \sigma_d^2 - g\mu_d \tag{25.6}$$

The noisy signal is therefore a heteroscedastic random variable whose variance linearly depends on its expectation, thus illustrating that noise is heteroscedastic. This heteroscedastic noise can be changed into a homoscedastic one by using the GAT introduced above [1]. This transform stabilizes to 1 the variance of any Poisson-Gaussian variable following the model given in Eq. (25.6). The GAT writes as follows:

$$GAT(s(x, y)) = \frac{2}{g} \sqrt{gs(x, y) + \frac{3}{8}g^2 + \sigma_d^2 - g\mu_d} \tag{25.7}$$

In practice, the parameters involved in the GAT are those driving the linear equation between variance and expectation given in Eq. (25.6). In theory, they can be found either from one single image, as in [1] for instance, or by considering a stack of images of the same static scene (here a grid), calculating the mean and variance of noise at any pixel, plotting the corresponding point cloud in the variance-mean plane and finally fitting the data by a straight line whose slope and intercept are g and $\sigma_d^2 - g\mu_d$, respectively. Note that g, σ_d and μ_d are not always given in datasheets of cameras, it therefore necessary to identify them using a dedicated procedure as in [3] for instance.

25.5 Experimental Results

The objective here is to check whether the above theoretical predictions are satisfied in practice and to observe the benefit of the GAT on grid images grabbed by a CCD sensor. The brightness being not constant throughout any grid image and the spatial resolution being fixed, it is worth mentioning, by considering Eqs. (3–4) above, that the resolution is not a constant quantity: it varies throughout the maps since it is inversely proportional to the brightness which is in general heterogeneous. Equations (25.3–25.4) must therefore be verified at any pixel of a stack of grid images.

Finding a state of reference for a strain distribution is an issue. A mere translation test was therefore performed to check the abovementioned equations, as in [14] for instance, the strain distribution being theoretically null in this case. In this study, the specimen was an aluminum rectangular plate on which a grid was transferred using the technique described in [4]. The pitch of the grid was 0.2 mm and the distance between camera and specimen was such that 5 pixels were employed to digitize a grid pitch. The grid was lit by three flexible and movable light guides fed by a cold light source. 200 grid images of this grid image were grabbed before and after translating the specimen by 0.1 mm. A SENSICAM QE camera featuring a 12-bit 1,040 × 1,376-pixel CCD sensor was used for this purpose. The mean value and standard deviation for the temporal distribution of gray level in grid images can be easily obtained pixelwise from this stack of images. Phase, phase derivative, displacement and strain in the corresponding maps can also be easily calculated at any pixel of the maps extracted from these images, and then compared to their theoretical predictions.

Heteroscedastic noise in the grid images is illustrated in Fig. 25.1a which represents a close-up of the standard deviation of the gray level distribution. In this figure, the size of each pixel is 40 × 40 μm.

The grid pattern is easily recognizable by comparing this map with a close-up of the same zone in the grid (see Fig. 25.1b). More precisely, the standard deviation of the gray level distribution is all the greater as the gray level is high.

The GAT has been applied to normalize noise throughout the grid images. Parameters involved in Eq. (25.7) were first identified, so the GAT could then be applied to normalize noise. It can then be checked that the standard deviation of the gray level distribution at any pixel is randomly distributed around one, as illustrated in Fig. 25.1c. It must be pointed out that unavoidable microvibrations that occurred during grid image grabbing (in a laboratory context) significantly corrupted this noise characterization despite their tiny amplitude of some micrometers. This effect can be detected in Fig. 25.1a: the bright center of all the cells of the grid pattern is slightly upwardly stretched since microvibrations occur along the vertical direction. A specific procedure described in [3] was therefore employed to get rid this effect before applying the GAT.

Equations 25.4 are then employed to predict the resolution in displacement and strain maps. Since the GAT has been applied, σ_{im} is equal to one in these equations. In the same way, K is the modulus of the normalized images, not that of the initial raw images. The comparison between observed and predicted resolutions along the vertical direction is shown in Fig. 25.2 as an example. It can also be checked that the mean value of the strain at any pixel is very close to zero [11].
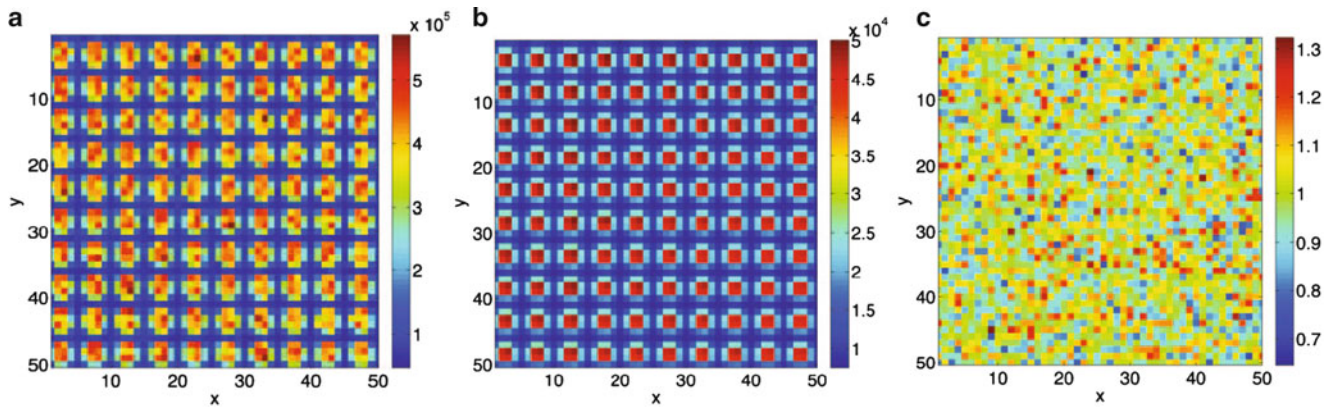
Fig. 25.1 Benefit of the GAT. Close-up of a mean grid image and corresponding spatial variance distributions. (**a**) Variance of the noise (**b**) gray level distribution (**c**) normalized variance obtained after eliminating microvibrations and applying the GAT
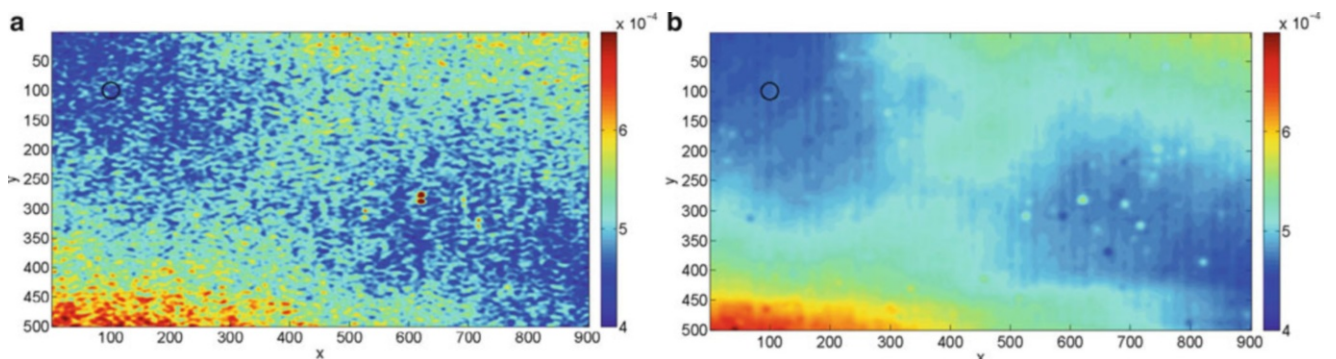
Fig. 25.2 Resolution in the strain map. (**a**) Observed with a stack of 200 strain maps (**b**) Predicted with Eq. (25.4a)

Several remarks can be drawn from this figure:

– the global aspect of the experimental and theoretical distributions of the resolution are very similar. The small blobs visible in the experimental maps are due to noise correlation that is not accounted for in Eq. (25.4a). The next step will be to check whether it corresponds to noise correlation predicted in [8];

– a slight vertical pattern as well as some isolated outliers are visible in the predicted map. They are due to grid printing defects;

– the resolution is not constant throughout the map. This is due to the non-uniform lightning. By comparing this distribution directly with the gray level of the grid image or the distribution of the modulus of the WFT, it can be easily checked that the noise level is all the lower as the brightness is high. It is worth mentioning that obtaining a uniform lighting is tricky in practice: even though the grid image observed to the naked eye on the screen of the companion PC of the camera could appear to be uniform (as in the current example);

– observed and predicted distributions for the displacement are also in very good agreement [11] (not shown here);

– a black circle is superposed at the top left of each map. The diameter is equal to the spatial resolution, considered here to be equal to $6 \times \sigma = 30$ pixels. The reader can see at a glance the metrological performance of the technique for retrieving the strain: the circle provides the spatial resolution which is constant throughout the map while the resolution changes from one point to each other because the brightness changes throughout the grid images in this example. Impairing the spatial resolution in the strain map (thus enlarging the circle) would cause the noise level to decrease and the map to be more blurred, the observed distribution being equal to the actual one convolved by the envelope of the kernel of the WFT [8].

25.6 Conclusion

Closed-form expressions are given to predict how noise propagates between camera sensor and displacement and strain maps obtained with the grid method. Noise of camera sensor being heteroscedastic, it must first be stabilized with the GAT to obtain experimental results in very good agreement theoretical expectations. reliably estimating the noise level distribution in strain maps is crucial, not only to assess the metrological performance of the grid method as a tool for measuring strain maps, but also to perform efficient deconvolution of the strain maps to retrieve the actual strain distributions [6].

References

1. Healey GE, Kondepudy R (1994) Radiometric CCD camera calibration and noise estimation. IEEE Trans Pattern Anal Mach Intell 16(3):267–276
2. Murthagh F, Starck JL, Bijaoui A (1995) Image restoration with noise suppression using a multiresolution support. Astron Astrophys 112:179–189
3. Sur F, Grédiac M (2014) Sensor noise modeling by stacking pseudo-periodic grid images affected by vibrations. IEEE Signal Process Lett 21(4):432–436
4. Piro JL, Grédiac M (2004) Producing and transferring low-spatial-frequency grids for measuring displacement fields with moiré and grid methods. Exp Tech 28(4):23–26
5. Surrel Y (2000) Fringe analysis. In: Rastogi PK (ed) Photomechanics, vol 77, Topics in applied physics. Springer, Berlin, pp 55–102
6. Badulescu C, Grédiac M, Mathias J-D (2009) Investigation of the grid method for accurate in-plane strain measurement. Meas Sci Technol 20(9):095102
7. Surrel Y (1997) Additive noise effect in digital phase detection. Appl Opt 36(1):271–276
8. Sur F, Grédiac M (2014) Towards deconvolution to enhance the grid method for in-plane strain measurement. Inverse Probl Imaging 8(1):259–291
9. Grédiac M, Sur F, Badulescu C, Mathias J-D (2013) Using deconvolution to improve the metrological performance of the grid method. Opt Lasers Eng 51:716–734
10. Chrysochoos A, Surrel Y (2012) Basics of metrology and introduction to techniques. In: Grédiac M, Hild F (eds) Full-field measurements and identification in solid mechanics. Wiley, Weinheim, pp 1–29
11. Grédiac M, Sur F (2014) Effect of sensor noise on the resolution and spatial resolution of the displacement and strain maps obtained with the grid method. Strain 50(1):1–27
12. Mallat S (2008) A wavelet tour of signal processing: the sparse way, vol 3. Academic, San Diego
13. Hild F, Roux S (2012) Comparison of local and global approaches to digital image correlation. Exp Mech 52(9):1503–1519
14. Haddadi H, Belhabib S (2008) Use of a rigid-body motion for the investigation and estimation of the measurement errors related to digital image correlation technique. Opt Lasers Eng 46:185–196

Chapter 26
Experimental Evaluation of the Warping Deformation in Thin-Walled Open Section Profiles

Sandro Cammarano, Giuseppe Lacidogna, Bartolomeo Montrucchio, and Alberto Carpinteri

Abstract The analysis of thin-walled open section beams can be performed by means of Vlasov's theory of the sectorial areas. It is well-known that this type of profiles, when subjected to torsional actions, are characterised by the warping deformation and, consequently, by a further stress state whose intensity is comparable with that produced by mere flexural deformations. In the literature many papers are focused on the structural behaviour of these elements, but, to the Author's best knowledge, none proposed an experimental technique to evaluate first-hand this particular behaviour. In order to verify the classical theory of the sectorial areas, in the present paper an experiment regarding a thin-walled open section profile subjected to flexural and torsional loads is performed. With the help of a specific optical device, suitable for precision measurements, the warping displacements of a U-shaped section are easily acquired. These are compared to those derived, first, from an analytical formulation, originally devised to deal with vertical thin-walled bracings belonging to the structural core of a tall building, and, secondly, from a FE program, in which the steel profile is modelled by means of thin-shell elements. The numerical comparison confirms the reliability of the analytical formulation.

Keywords Thin-walled section • Warping deformation • Vlasov's theory • Experimental evaluation • Optical device

26.1 Introduction

Thin-walled open section profiles are usually employed in high-rise buildings to provide the adequate horizontal resistance against the effects derived from earthquakes and wind. The limits imposed by legislation in terms of lateral displacements of these constructions have caused the appearance of several resistant systems. Among all, open section bracings prove to be particularly convenient since, in addition to contributing to the total stiffness, they can house elevator shafts and stairwells, which are necessary for the functional needs required by the building occupants and useful for a rapid usability of the floor spaces.

When subjected to torsional actions, the transversal section of such elements undergoes an out-of-plane distortion, known in literature as warping deformation. Associated with this phenomenon, the bimoment action produces a further stress state, similar to that generated by flexural actions, which cannot be disregarded especially when their structural contribution is decisive for the lateral stability of the building.

Timoshenko [1] and Vlasov [2] were the pioneers in the analysis of thin-walled open section profiles. In particular, a comprehensive theory was proposed by Vlasov who demonstrated that the well-known Saint Venant's theory was actually a very particular case of a more general formulation. Many authors tried to extend his original approach, which, however, represents a milestone in the study of structural mechanics.

To the authors' best knowledge, papers regarding experimental tests for the assessment of the warping deformation of thin-walled open section beams are almost absent in literature. Some of them, based on Vlasov's theory, are related to the

S. Cammarano (✉) • G. Lacidogna • A. Carpinteri
Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
e-mail: sandro.cammarano@polito.it; giuseppe.lacidogna@polito.it; alberto.carpinteri@polito.it

B. Montrucchio
Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
e-mail: bartolomeo.montrucchio@polito.it

experimental evaluation of the free vibrations of these elements, which in dynamics show a coupled behaviour between the flexural modes and between torsional and flexural modes. The papers by Ambrosini [3, 4] and Klausbruckner and Pryputniewicz [5] are glaring examples of these analyses.

In statics, experimental tests for the identification of the warping deformations are very rare and, however, not actually focused on this purpose. In this direction two works can be underlined: the first is the paper by Rendek and Baláž [6] who focused their attention on the evaluation of the efficiency of the analytical approach known as Generalized Beam Theory (GBT). To this aim, a strain-measuring test on a steel cantilever beam of complex open section, subjected to various loading cases, was performed. Exploiting the presence of strain gauges located along the beam, the results obtained analytically and experimentally were compared in terms of longitudinal stresses. The type of test permitted to understand the influence of the out-of-plane distortion of the section on the structural capacity of the element, but no information regarding the warping deformation itself were stressed by the authors.

Focusing the attention on a particular type of distortion which occurs in composite elements, the paper by Dufort et al. [7] dealt with the cross-section warping of beams by examining the displacement field through the use of an optical technique in conjunction with the employment of microcomputer and methods of image analysis. Once the beam was deformed, the warping of the section could be acquired through the difference between the measured displacements and a linear trend of displacements obtained joining the longitudinal displacements related to the top and the bottom part of the section. In this way, an indirect evaluation of the distortion of the profiles could be performed.

Therefore, from the foregoing it is evident that a direct evaluation of the warping deformation of thin-walled open section profiles is still missing in literature. To this aim, in the present paper a specific experimental test is performed by means of a sophisticated optical device applied to the section of an horizontal steel cantilever subjected to transversal concentrated actions. The results obtained experimentally and through a Finite Element (FE) model allow to validate a computer program based on the analytical theory of the sectorial areas and devised for the evaluation of the structural behaviour of high-rise buildings stiffened by any kind of horizontal bracings. This code implements the formulation proposed by Carpinteri et al. [8–10] which computes the transversal displacements of tall buildings and evaluates the external load distribution between the bracings constituting the resistant core of the structure. The effectiveness of the method is finally highlighted through the execution of a numerical example in which the benefits provided by the analytical formulation compared to FE program are evident.

26.2 Evaluation of the Warping Deformation

The theory of the sectorial areas proposed by Vlasov is rarely used to analyse the structural behaviour of thin-walled open section beams. In effect, to the authors' best knowledge, papers focused on experimental techniques able to evaluate first-hand their warping deformation are quite absent in literature. Therefore, with the aim of verifying the efficiency of the theory of the sectorial areas, an experiment which takes into account a thin-walled open section profile subjected to flexural and torsional loads is here proposed. The sample is a slender steel beam having an U profile, on which an optical device, suitable for precision measurements, is applied in order to evaluate the out-of-plane distortion of the section. The comparison of the results obtained experimentally with those coming from a FE model and computed analytically allows to appreciate the degree of accuracy of Vlasov's theory.

In the following sections, a summary of the analytical approach which, starting from the theory of the sectorial areas, allows to identify the stiffness matrix of thin-walled open section beams subjected to transversal loads is proposed as well as the main phases of the experimental test are described in depth.

26.2.1 Computation of the Stiffness Matrix for Thin-Walled Open Section Beams Subjected to Transversal Actions

A linear-elastic, isotropic and homogeneous beam characterised by a thin-walled open section in a right-handed reference system OXYZ is considered (Fig. 26.1). The variables ξ, η and ζ represent the translations of the origin O along the directions X, Y and Z respectively, whereas the variable ϑ describes the transversal rotation of the section. In a general loading case, the differential equilibrium equations of the beam can be written, according to Vlasov's theory, as:
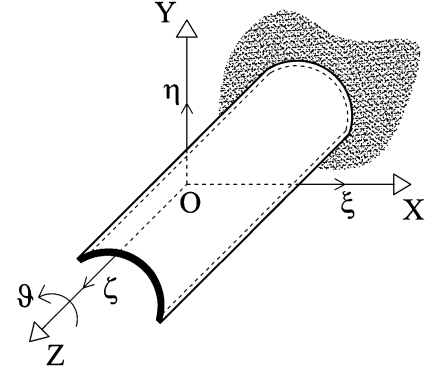
$$E\left(\zeta'' A - \xi''' S_y - \eta''' S_x - \vartheta''' S_\omega\right) = q_z$$

$$-E\left(\zeta''' S_y - \xi^{IV} I_y - \eta^{IV} I_{xy} - \vartheta^{IV} I_{\omega y}\right) = q_x$$

$$-E\left(\zeta''' S_x - \xi^{IV} I_{xy} - \eta^{IV} I_x - \vartheta^{IV} I_{\omega x}\right) = q_y$$

$$-E\left(\zeta''' S_\omega - \xi^{IV} I_{\omega y} - \eta^{IV} I_{\omega x} - \vartheta^{IV} I_\omega\right) - GJ_t \vartheta'' = m$$

(26.1)

Nevertheless, this system can be simplified considering some hypotheses related to the reference systems: if the origin O coincides with the centroid of the section, the static moments S_x and S_y become equal to zero. Similarly, if the axes X and Y are principal, the product of inertia I_{xy} is also null. Finally, if we consider the shear centre as the sectorial pole and the sectorial origin as the principal one, the sectorial static moment S_ω and the sectorial products of inertia $I_{\omega x}$ and $I_{\omega y}$ vanish. Consequently, the system (26.1), initially constituted by coupled equations, is now composed by four independent expressions:

$$EA\zeta'' = q_z$$

$$EI_y \xi^{IV} = q_x$$

$$EI_x \eta^{IV} = q_y$$

$$EI_\omega \vartheta^{IV} - GJ_t \vartheta'' = m$$

(26.2)

For the sake of simplicity, in the fourth equation the term proportional to the torsional stiffness GJ_t is disregarded. Its contribution is added later, in the final expression of the stiffness matrix of the element. If the external forces are only represented by transversal actions, Eq. (26.2a) can be neglected so that the remaining equations can be organised in a matrix form exploiting the vectors of the displacements δ and the actions q:

$$\delta = \left\{ \begin{matrix} \xi \\ \eta \\ \vartheta \end{matrix} \right\} \quad q = \left\{ \begin{matrix} q_x \\ q_y \\ m \end{matrix} \right\}$$

(26.3)

$$E\mathbf{I}\delta^{IV} = q$$

(26.4)

The term $\mathbf{I}$ is a diagonal 3×3 matrix whose diagonal coefficients are given by I_y, I_x and I_ω, whereas all the other elements are equal to zero. Since $\mathbf{I}$ is symmetrical and positive definite until the sectorial moment of inertia I_ω is different from zero, the inversion of the expression (26.4) highlights the vector of displacements δ.

$$\delta^{IV} = \mathbf{I}^{-1} q / E$$

(26.5)

If the actions are represented by concentrated loads in correspondence to N sections (Fig. 26.2a), Eq. (26.5) is valid in each segment included between two consecutive sections and turns into:

Fig. 26.2 Thin-walled open section beam subjected to concentrated transversal actions (**a**); continuity conditions between different sections (**b**)
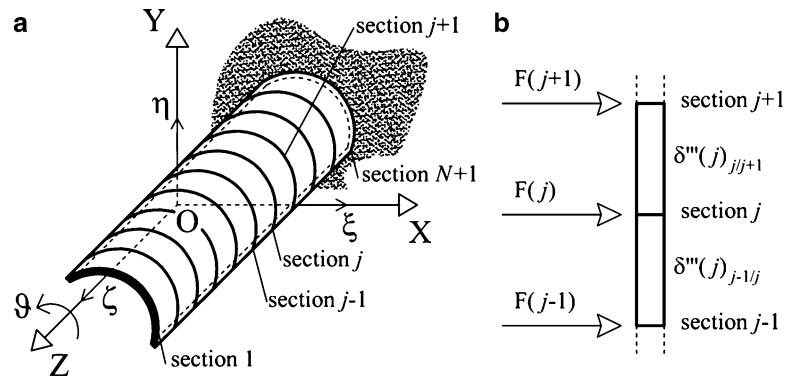
$$\delta^{IV} = \mathbf{I}^{-1}0/E \tag{26.6}$$

or, recalling the concentrated actions,

$$\delta''' = -\mathbf{I}^{-1}F/E \tag{26.7}$$

being the vector F connected to the vector q:

$$q = -\frac{\partial}{\partial z}F = -\frac{\partial}{\partial z}\begin{Bmatrix} F_x \\ F_y \\ M_Z \end{Bmatrix} \tag{26.8}$$

The variables ξ, η and ϑ are acquired integrating Eq. (26.6) in each single segment and exploiting the adequate boundary conditions. Such conditions are continuity and global conditions. The continuity conditions for the j-th section ($j = 2,\ldots, N$) are four, two kinematic and two assuring the equilibrium:

$$\begin{aligned} \delta(j)_{j-1/j} &= \delta(j)_{j/j+1} \\ \delta'(j)_{j-1/j} &= \delta'(j)_{j/j+1} \\ \delta''(j)_{j-1/j} &= \delta''(j)_{j/j+1} \\ \mathbf{EI}\delta'''(j)_{j-1/j} &= \mathbf{EI}\delta'''(j)_{j/j+1} + F(j) \end{aligned} \tag{26.9}$$

If the beam is constrained according to a cantilever scheme, the global conditions for the free edge ($j = 1$) are given by equilibrium conditions:

$$\begin{aligned} \delta''(1) &= 0 \\ \delta'''(1) &= -\mathbf{I}^{-1}F(1)/E \end{aligned} \tag{26.10}$$

whereas, for the opposite edge ($j = N + 1$), we consider kinematic conditions:

$$\begin{aligned} \delta(N + 1) &= 0 \\ \delta'(N + 1) &= 0 \end{aligned} \tag{26.11}$$

For the sake of simplicity, the displacements ξ, η and ϑ of all the sections are gathered in a single 3Nvector δ, composed by the translations along the X and Y axes and finally the rotations. Similarly a global 3 N vector $\underline{F}$ containing the shear forces F_x, F_y and torque moments M_Z is defined. Performing the procedure of integration previously mentioned, a

Fig. 26.3 Scheme of the thin-walled U-shaped section steel beam

U-shaped profile [t 3 mm]   L-shaped profile [t 2 mm]
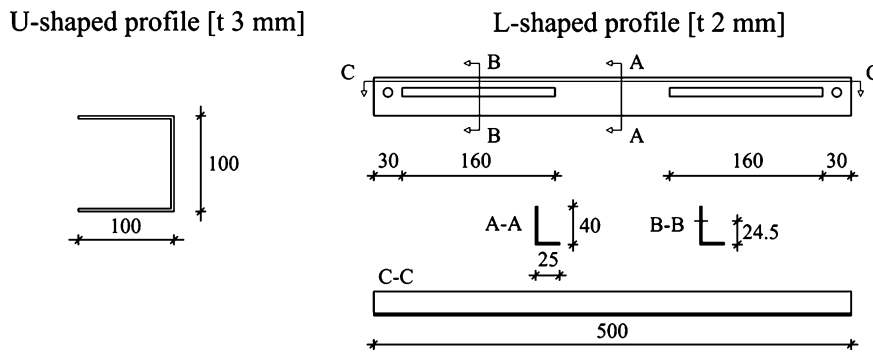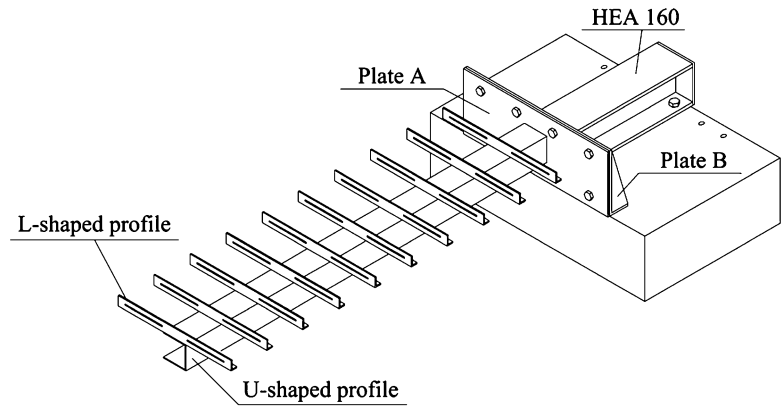
Fig. 26.4 Geometrical properties of the components of the experimental test

relationship between δ and $\underline{F}$ through the compliance matrix $\mathbf{C}$ or its inverse, the stiffness matrix $\underline{\mathbf{k}}$, is obtained. Finally the contribution proportional to the torsional inertia J_t is added to the torsional component of the calculated stiffness matrix.

$$\underline{\delta} = \mathbf{C}\underline{F} = \underline{\mathbf{k}}^{-1}\underline{F} \tag{26.12}$$

Thanks to Eq. (26.12) the displacements of the sections in terms of ξ, η and ϑ can be computed. Then, differentiating the latter with respect to the coordinate z and recalling Eq. (26.13) from Vlasov's theory, the axial displacements which identify the warping of the section can be simply acquired.

$$\mathrm{w}(z, s) = -\vartheta^{'}(z)\omega \tag{26.13}$$

26.2.2   Experimental Evaluation of the Warping Deformation in Thin-Walled Open Section Beams

Experimental investigations focused on the evaluation of the effective out-of-plane distortion of thin-walled open sections subjected to torsional actions are almost absent in the literature of the past 50 years. Therefore it is possible to acknowledge that the contribution described below proves to be particularly innovative for the analysis of these thin-walled profiles.

The assessment of the warping deformation is performed by means of an experimental test on a 1.6 m long steel beam, constrained according to a cantilever scheme having as cross-section a thin-walled U profile. The rigid constraint is realised at one end by means of welding on a steel plate, which is bolted to a fixed system (Fig. 26.3, plate A). On the upper flange of the section, at distance of 0.16 m each other, some L-shaped section beams, conveniently altered for the application of the external actions, are welded. The geometrical dimensions of the components are shown in Fig. 26.4. The applied loads are represented by masses, each of 3.08 kg, positioned at 0.23 m far from the shear centre of the U-shaped section. In this way shear forces and anticlockwise torsional moments are simultaneously produced (Fig. 26.5).

Fig. 26.5 Application of the
external concentrated loads
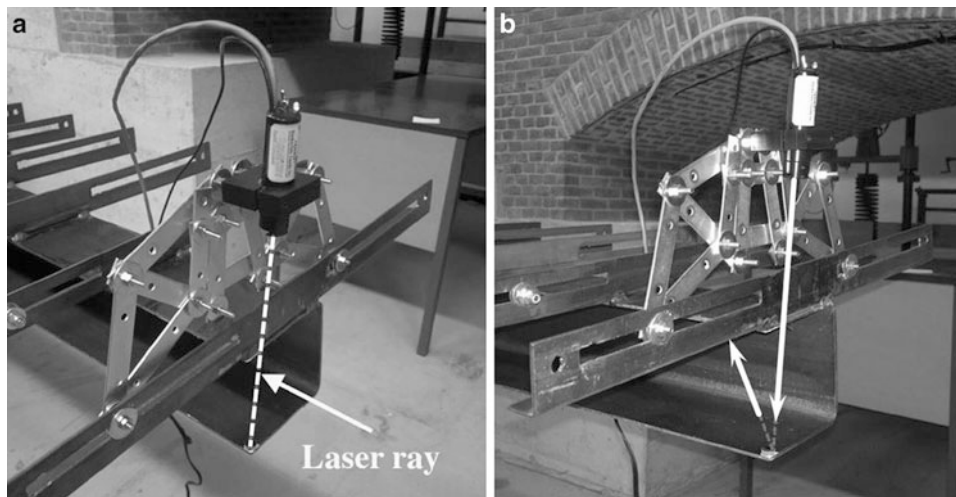
Fig. 26.6 Picture of the little steel frame, applied to the extreme L-shaped profile, to which the laser is fixed (**a**); scheme of the reflective effect of the mirror (**b**)

Fig. 26.7 Positions of the
spherical mirror and the ruler
which allows the
measurement of the warping
of the section
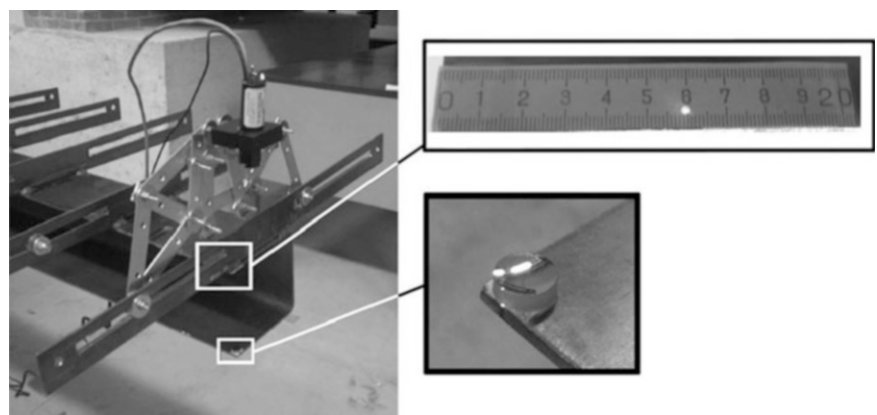
The detection of the axial displacements which define the warping of the section is realised by means of three main instruments. The first is a steel frame connected through bolts to the L-shaped section beam fixed at the free end of the U-shaped section beam; the second is a laser, attached to the steel frame and directed to the free end of the bottom flange of the U-shaped profile (Fig. 26.6a); the last is a spherical mirror, having specific geometrical properties for precision

Fig. 26.8 Relative axial displacement s between the laser (A → A′) and the mirror (B → B′) caused by the warping of the section
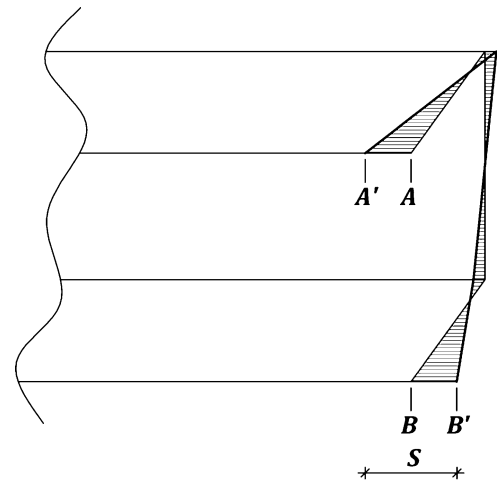
Fig. 26.9 Amplification of s due to the geometrical characteristics of the spherical mirror

ET [mm] = 3.63
CT [mm] = 3
d [mm] = 10
R [mm] = 20
h [mm] = 90.37
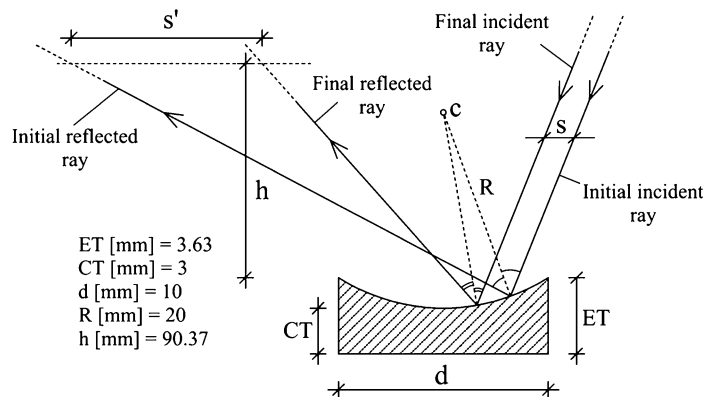
measurements (instrumental error 1/1000 mm), fixed in the point of incidence between the ray and the bottom flange of the main beam. Due to the presence of the mirror, the laser ray is reflected on the inner side of the upper flange of the U-shaped section beam, where a ruler is pre-emptively pasted (Figs. 26.6b and 26.7).

Because of the concentrated torques, the section twists around its shear centre and undergoes, at the same time, a distortion out of its own plane. This deformation is given by the system of axial displacements of the points the section, once the pure flexural deformations are subtracted. Experimentally, it is obtained considering the relative axial displacement of two symmetrical points of the section: in particular the points corresponding to the laser and to the mirror (Fig. 26.8).

As a consequence of the displacement s of the incident ray caused by the warping deformation, on the upper flange of the U-shaped section the displacement s' related to the corresponding reflected ray can be detected. Thanks to the curvature of the mirror, the term s' represents an amplification of s and its measurement becomes much easier (Fig. 26.9).In this case, the geometrical characteristics of the mirror are such that the amplification factor, in other words the ratio between s' and s, is about 12.

The experimental analysis is conducted in presence of different loading conditions: in the first phase the L-shaped section beams are loaded progressively from z = 0.48 m to z = 1.44 m, being z = 0 m the constrained end of the main beam; then, the unloading process is performed following the same order (Fig. 26.10). In Fig. 26.11 the experimental results in terms of the amplified displacement s' during the loading phase are shown. It is evident that the use of the optical device facilitates the evaluation of displacements which are, otherwise, invisible to the naked eye.

The sample is also modelled in a Finite Element program in order to verify the efficiency of the analytical approach in the evaluation of the structural behaviour of thin-walled open section profiles. In this case, the U-shaped section beam is constructed by means of thin shell elements. In Table 26.1 the numerical comparison in terms of the relative displacement s between the analytical and the FE method is reported. The greatest error is about 4 %, thus confirming the good convergence between the approaches. A further demonstration of the reliability of the analytical formulation can be acquired
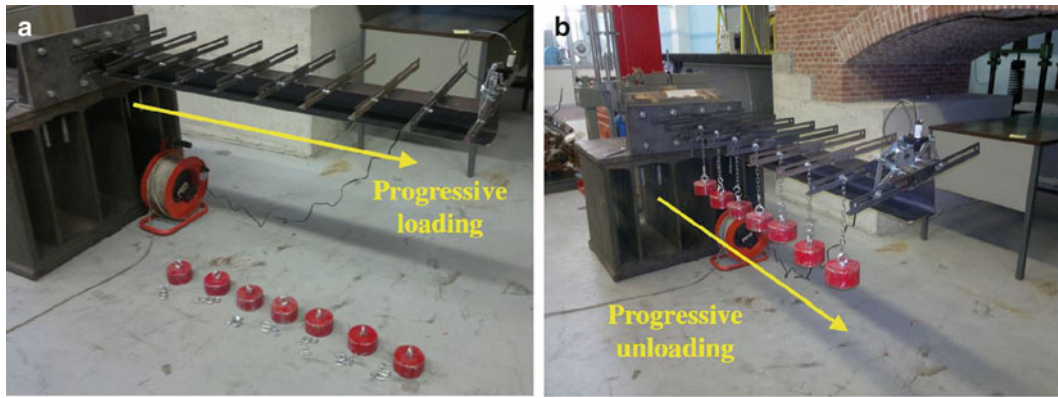
Fig. 26.10 Loading (**a**) and unloading (**b**) phase
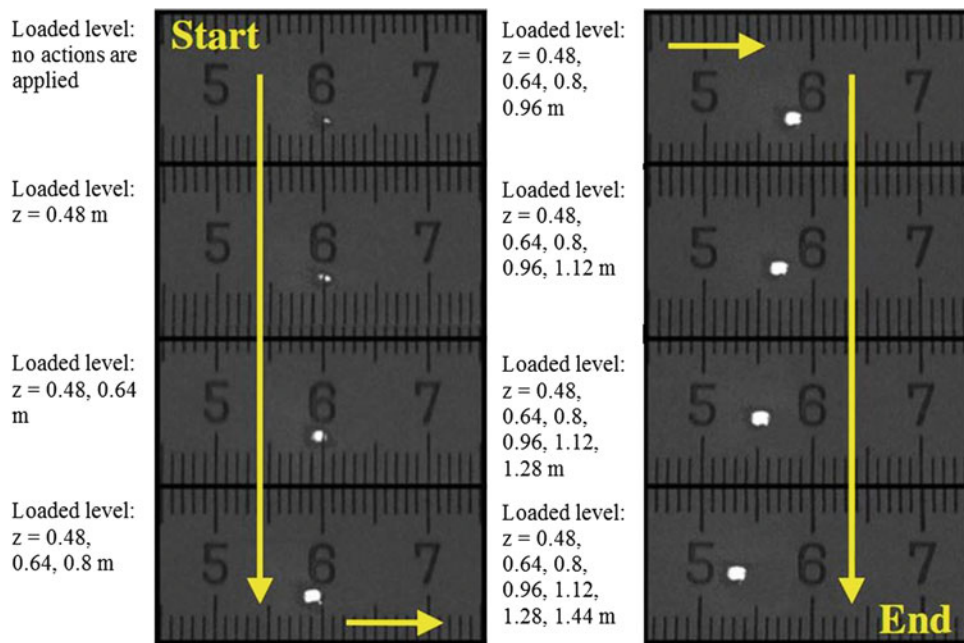
Fig. 26.11 Measurement of the displacement s′ during the loading phase: the white dots represent the consecutive positions of the reflected ray on the ruler

Table 26.1 Comparison between analytical and FE method

Type of process		Loading				Unloading			
			s [mm]				s [mm]		
Level N.	z [m]	Loaded level	Analytic.	FEM	Err. [%]	Loaded level	Analytic.	FEM	Err. [%]
1	0.48	1	7.61E-03	7.56E-03	0.61	1÷7	2.52E-01	2.46 E-01	2.58
2	0.64	1, 2	2.13E-02	2.10E-02	1.23	2÷7	2.44 E-01	2.38 E-01	2.64
3	0.80	1÷3	4.29 E-02	4.22 E-02	1.56	3÷7	2.31 E-01	2.25 E-01	2.71
4	0.96	1÷4	7.44 E-02	7.32 E-02	1.74	4÷7	2.09 E-01	2.03 E-01	2.79
5	1.12	1÷5	1.18 E-01	1.16 E-01	1.86	5÷7	1.77 E-01	1.72 E-01	2.94
6	1.28	1÷6	1.76 E-01	1.73 E-01	2.00	6, 7	1.34 E-01	1.30 E-01	3.23
7	1.44	1÷7	2.52 E-01	2.46 E-01	2.58	7	7.54 E-02	7.26 E-02	3.95

Error [%] = (Analytical–FEM)/FEM × 100
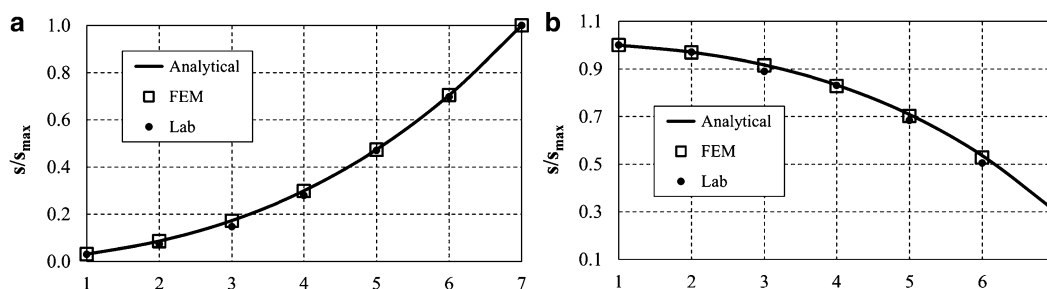
Fig. 26.12 Comparison of the results in terms of normalised values of the axial displacement s, during the loading (**a**) and unloading (**b**) process

comparing the previous numerical results with the experimental ones: Fig. 26.12 shows the normalised values of the axial displacement s during both the loading conditions.

Since the analytical approach is characterised by enough accuracy and the material is considered in the linear elastic domain, the proposed formulation can be extended to examine greater elements, which are often employed to stiffen horizontally tall buildings, being part of the vertical resistant core of the structural skeleton.

26.3   Analysis of Thin-Walled Open Section Bracings in High-Rise Buildings

The analytical formulation introduced in the previous section can be directly included in the three-dimensional formulation proposed by Carpinteri et al. [8, 9] for the evaluation of the structural behaviour of high-rise buildings stiffened by various vertical bracings, randomly arranged in the building plan.

FE methods are usually employed in the design of tall complex structures; nevertheless the implementation of the analytical approach can provide some benefits which are crucial especially in the preliminary phases of the design process. First of all, since the modelling times prove to be reduced, several analyses aimed to identify the best structural solution can be carried out; furthermore, the clearness of the method limits the possibility of unexpected errors and allows to elicit information related to the parameters which govern the structural behaviour of the constructions. Secondly, the external load distribution between the vertical bracings constituting the structural skeleton of the building can be distinctly described. Finally, aimed to perform a preliminary design of each structural element, the internal actions, bimoment included in the case of thin-walled open section members, can be directly evaluated. In order to verify the aforementioned benefits, a numerical example concerning a tall building loaded by transversal static actions is executed. The resistant core of the structure is composed by three thin-walled open section members, having different cross-sections.

The building is a 40-storey construction whose storey height is constant and equal to 5 m. The bracings are made of concrete whose mechanical properties are given by Young's modulus and Poisson's ratio equal to 3×10^4 MPa and 0.18 respectively. The dimensions and the geometrical properties of the vertical elements are shown respectively in Fig. 26.13 and in Table 26.2. The actions are represented by concentrated resultant forces applied to the floors in correspondence to the origin of the global coordinate system OXYZ (Fig. 26.13). In particular, along the X direction a resultant force of 99 kN for each floor is considered, whereas in Y direction a force equal to104 kN per storey is applied; finally a clockwise torque of 183.5 kNm is considered at each level.

This structural skeleton is analysed implementing the present formulation and employing a FE program. The results are expressed in terms of the transversal displacements, in particular translations and rotations of the floors according to the global coordinate system XYZ. The corresponding curves are shown in Fig. 26.14 from which it is self-evident the remarkable convergence of the approaches, confirming the benefit derived from the analytical formulation in the early phases of the design process. Further information regarding the internal actions can be obtained from the proposed method: in Fig. 26.15 the distributions of the shears T_x, T_y and the torque M_z between the bracings are pointed out.

This approach proves to be an useful support in the design process since it simplifies the evaluation of the most stressed element as well as allows an accurate comprehension of the force flow within the elements constituting the structural skeleton of the construction. In addition, unlike FE simulations, in the case of thin-walled open section profiles, the

Fig. 26.13 Internal core system of a tall building constituted by thin-walled open section shear walls (measures in meters)

H = 200 m
h = 5 m
E = 30000 MPa
v = 0.18
F_x = 99 kN/floor
F_y = 104 kN/floor
M_z = -183.5 kNm/floor
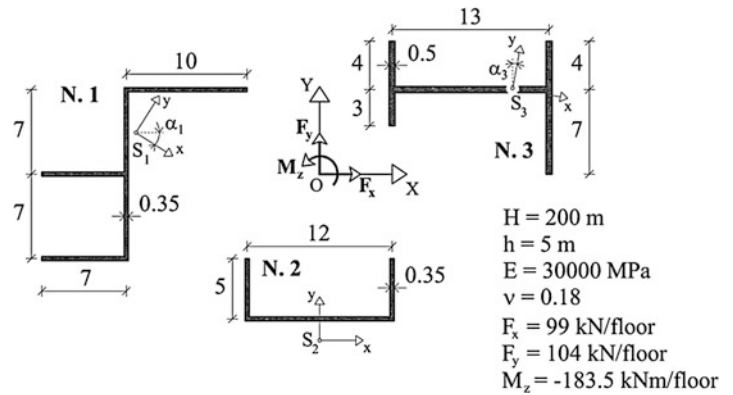
Table 26.2 Geometrical properties of the shear walls

Element N.	1	2	3
I_x [m^4]	483.45	19.22	71.18
I_y [m^4]	80.76	176.40	469.99
I_ω [m^6]	3611.67	487.50	1997.06
J_t [m^4]	0.54	0.31	1.29
x_s [m]	−15.21	0.00	16.45
y_s [m]	3.43	−13.79	6.97
α [°]	−32.45	0.00	−8.69
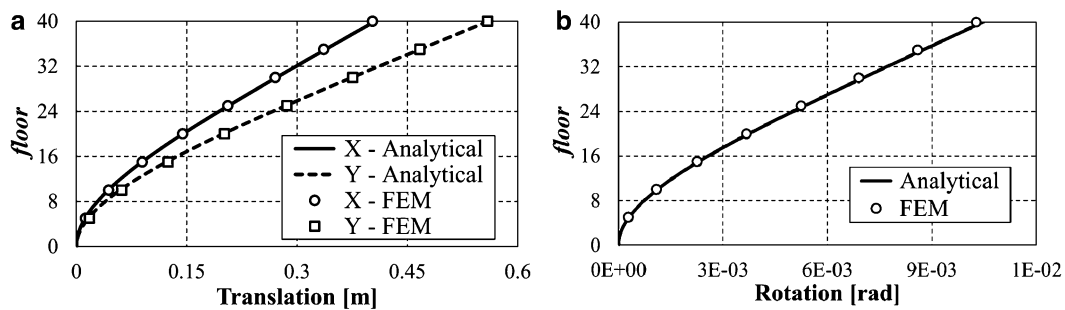
Fig. 26.14 Comparison between the analytical and FE method in terms of displacements of the building

bimoment action, which is associated with the warping deformation of open sections, can be computed and, consequently, the additional state of stresses defined by normal and tangential components can be derived. Since their intensity can be comparable to the one derived from pure flexural deformations and, therefore, can affect the structural behaviour of the entire building, at least an approximate evaluation of the bimoment action has to be carried out. In this direction, the results acquired by the analytical formulation are reported in Figs. 26.16 and 26.17: in the first figure the trends of the bimoment acting in each shear wall are shown, whereas in the second the additional stress state in terms of normal components is highlighted. In the last figure, a comparison with the stress state caused by the pure bending behaviour is underlined from which it is possible to recognise an increase in tension up to 20–30 %.
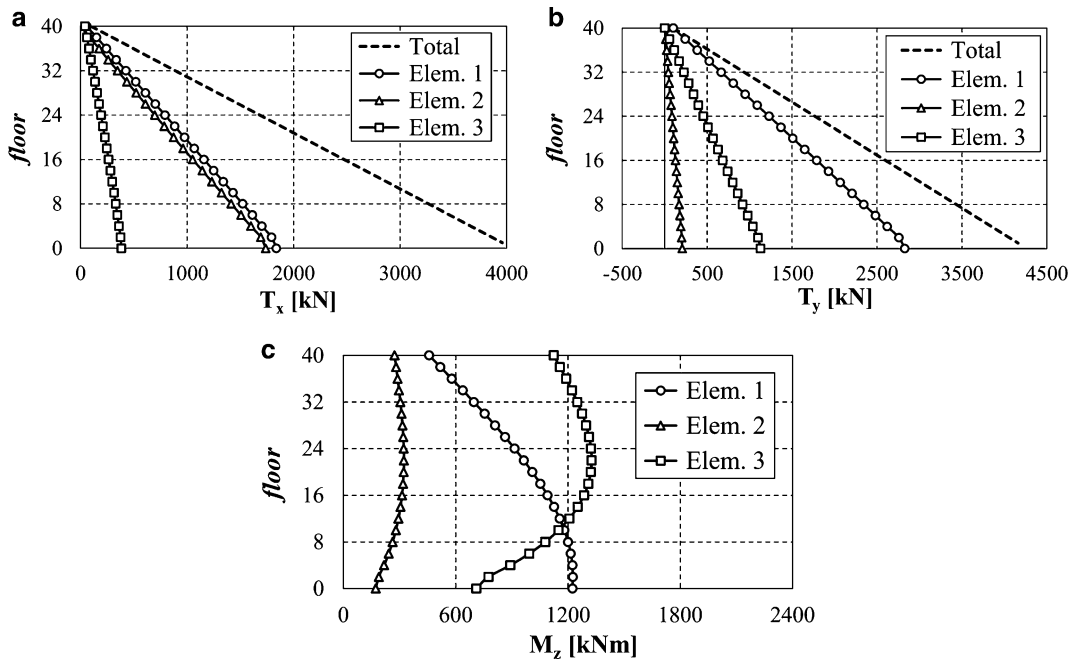
Fig. 26.15 External load distribution between the components of the core system in terms of shears in X (**a**) and Y (**b**) direction and torque (**c**)

Fig. 26.16 Bimoment action
in the shear walls constituting
the horizontal resistant system
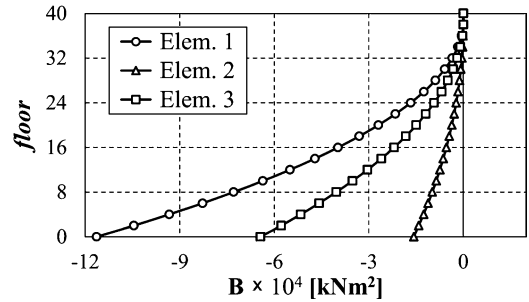
Element N.	1	2	3
M_x [kNm]	-350313.09	-20827.37	-116807.38
M_y [kNm]	2776.34	171628.43	26628.93
B [kNm²]	-116436.02	-15716.44	-64382.83
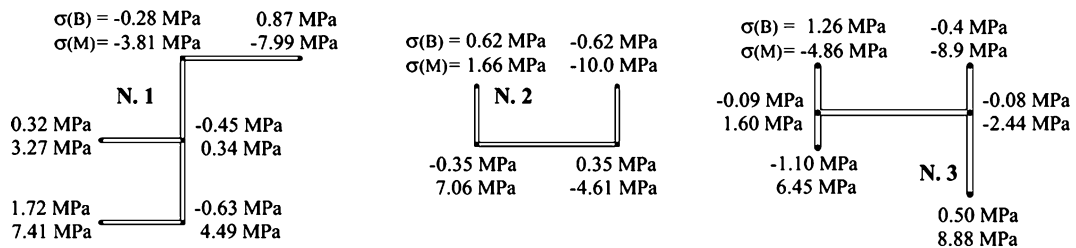
Fig. 26.17 Comparison of the stress state caused by bending and warping of the walls

26.4 Conclusions

The paper is focused on the analysis of the warping deformation of thin-walled open section beams, subjected to torsional actions. To this aim, an experimental test regarding a thin-walled U-shaped section profile loaded by transversal forces is performed. The use of an optical device allows to identify the out-of-plane distortion of the section. The results obtained experimentally are compared to those calculated exploiting a FE program and implementing the analytical formulation described by Vlasov's theory. Since a good convergence is observed, the proposed approach is extended to analyse the structural behaviour of high-rise structures stiffened by thin-walled open section bracings.

References

1. Timoshenko S (1945) Theory of bending, torsion and buckling of thin-walled members of open cross section. J Frankl Inst 239:201–361
2. Vlasov V (1961) Thin-walled elastic beams, 2nd edn. US Science Foundation, Washington, DC
3. Ambrosini D (2009) On free vibration of non-symmetrical thin-walled beams. Thin-Walled Struct 47:629–636
4. Ambrosini D (2010) Experimental validation of free vibrations from non-symmetrical thin walled beams. Eng Struct 32:1324–1332
5. Klausbruckner MJ, Pryputniewicz RJ (1995) Theoretical and experimental study of coupled vibrations of channel beams. J Sound Vib 183:239–252
6. Rendek S, Baláž I (2004) Distortion of thin-walled beams. Thin-Walled Struct 42:255–277
7. Dufort L, Grédiac M, Surrel Y (2001) Experimental evidence of the cross-section warping in short composite beams under three point bending. Compos Struct 51:37–47
8. Carpinteri Al, Carpinteri An (1985) Lateral loading distribution between the elements of a three-dimensional civil structure. Comput Struct 21:563–580
9. Carpinteri A, Lacidogna G, Puzzi S (2010) A global approach for three dimensional analysis of tall buildings. Struct Des Tall Spec Build 19:518–536
10. Carpinteri A, Corrado M, Lacidogna G, Cammarano S (2012) Lateral load effects on tall shear wall structures of different height. Struct Eng Mech 41:313–337

Chapter 27
In Situ Study of Plastic Flow at Sliding Metal Surfaces

A. Mahato, Y. Guo, N. Sundaram, and S. Chandrasekar

Abstract An in situ study of deformation and flow at sliding metal interfaces is described. The model system used—a hard wedge sliding against a metal surface—is representative of wear, and surface conditioning processes typical of manufacturing. By combining high speed imaging with image analysis, important characteristics of unconstrained plastic flow intrinsic to these processes are highlighted. These characteristics include unusual fluid-like flow, surface folding, wear particle formation by folding, and large plastic strains on the surfaces. Implications for sliding wear and quality of surfaces generated by manufacturing processes are discussed.

Keywords Metals • Wear • Surface folding • Image correlation • High speed photography

27.1 Introduction

Unconstrained plastic flow is intrinsic to wear of sliding metal surfaces by abrasion, adhesion and delamination [1], and surface conditioning processes such as burnishing, machining and peening [2]. In all of these instances, the surfaces show extensive deformation, as do the chips and particles generated by the processes. This plastic flow is a consequence of high contact pressures, on the order of the material hardness, applied by an indenter (asperity, tool) sliding on the surface. A basic understanding of the characteristics of surface plastic flow will be of value in analyzing/controlling wear at sliding metal interfaces and improving the quality of surfaces created by the conditioning processes [3].

The typical experimental approach used to date in studying surface flow, to paraphrase Tabor's comments [4], is to "investigate the resultants of the wear system at various stages and to attempt to reconstruct a picture of what is going on". The problem with this "post-mortem" examination is that evidence of the critical (unit) material removal events is masked or obliterated by subsequent events. Nevertheless, this type of examination has enabled important insights to be obtained into flow/deformation phenomena occurring in wear and surface generation [5, 6]. These insights have enabled analytical models of wear and conditioning processes to be formulated using continuum approaches—slip line field (SLF), shake-down and finite element analyses (FEA) [7, 8].

A complementary, perhaps more attractive, experimental approach to analyzing surface flow phenomena is using in situ studies of the process zone. Improvements in imaging methods and image analysis techniques offer much scope for such in situ studies, provided the regions of interest can be accessed in a model system representative of the unit events in the process. The sliding contact between a sharp wedge and a metal surface under conditions of plane strain (2D) deformation provides a natural model system for conducting this type of study.

A. Mahato (✉) • S. Chandrasekar
Center for Materials Processing and Tribology, Purdue University, West Lafayette, IN 47907, USA
e-mail: mahato@purdue.edu

Y. Guo
M4 Sciences LLC, West Lafayette, IN 47906, USA

N. Sundaram
Department of Civil Engineering, Indian Institute of Science, Bangalore 560012, India

The present work is an in situ study of material flow in a sliding system representative of sliding wear and surface conditioning. The observations describe material flow in the process zone, strain and strain rate fields, and wear particle formation. Unusual fluid-like flow phenomena, surface folding, a mechanism of wear particle formation, and large plastic strains are the characteristics of unconstrained plastic flow revealed by this in situ observational approach.

27.2   Experimental

The model system (Fig. 27.1) consists of a metal workpiece (WP) in the form of a plate sliding against a hard steel wedge indenter at velocities <50 mm/s [9, 10] under conditions of plane-strain deformation. Various indenter rake angles (α) are used to explore loading conditions typical of sliding, and to study the flow and deformation phenomena. The angle (α) is adjusted by changing the orientation of the indenter and the depth of tool-work interaction (h_o) is varied by using a micro-stage (Fig. 27.1). The WP system is commercially pure Al (Al 1100, grain size ~ 200 μm) in an initial annealed condition. The tool-WP region is observed using a high-speed imaging system. The flow patterns and quantitative details of the deformation are obtained by application of Particle Image Velocimetry (PIV) to the image sequences [9, 11]. A sharp indenter edge (<10 μm edge radius) is used to minimize rubbing-induced deformation at the indenter-WP contact.

Deformation parameters are measured using Particle Image Velocimetry (PIV), an image analysis technique. In its basic form, PIV involves use of tracers or particles dispersed in the medium and tracking the motion of particle ensembles by digitizing high-speed images of the 'flow' [12, 13]. In the present adaptation of the method, the role of the particles is played by 'asperities'—roughness features deliberately introduced onto the WP surface being imaged by abrading with 1,200 grit SiC paper. This side is imaged using a high-speed CMOS camera (PCO dimax) coupled to an optical microscope assembly (Nikon Optiphot). An optically transparent glass plate lightly constrains the imaged side so as to minimize side (out-of-plane) flow of material during the sliding. The CMOS image sensor has a full resolution of 2,016 × 2,016 pixel and physical size of 22.18 × 22.18 mm. Images can be recorded at up to 1,279 frames per second (fps) using the full sensor area. Higher imaging speeds are achieved by reducing the employed sensor area; for example, the camera can operate at framing rates of up to ~5,000 fps when the image size is 1,296 × 720 pixels. The images are analyzed using the PIV to estimate displacement/velocity fields in the deformation zone, and adjoining particle/workpiece [9, 11]. Strain rate and strain, including tensor components, principal strains, and effective scalar values are estimated from the displacement fields. Equally, importantly stream-, streak- and path- lines of flow (analogous to fluid flows) are obtained from the displacement and velocity fields [10]. These lines are particularly useful for visualizing characteristic features of the flow.
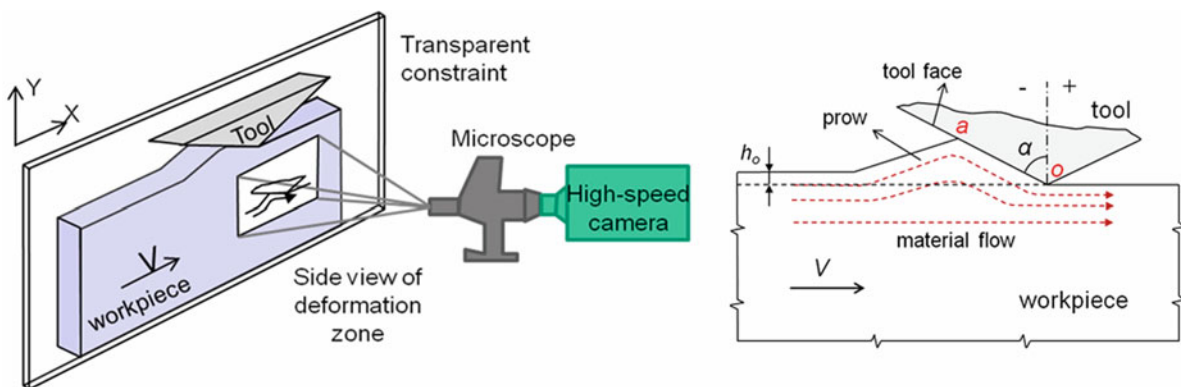
Fig. 27.1 Schematic of the experimental 2-D sliding setup (*left*) and nomenclature (*right*)

27.3 Results

The in situ imaging and the PIV analysis have revealed interesting, including hitherto unknown, aspects of plastic flow in sliding contact.

Flow patterns Fig. 27.2 shows one frame selected from a high speed image sequence of lubricated sliding of Al with $\alpha = -75°$. A prow of material builds up ahead of the wedge, and a smooth laminar flow is revealed by the superimposed streakline pattern derived from the high speed images. This flow is similar to that usually assumed in SLF analysis of sliding.

A dramatic change in the flow occurs when α is increased. Figure 27.3a–c shows the developing flow pattern in three frames from a high speed image sequence taken over duration of 0.14 s for $\alpha = -65°$. The interesting features of this flow pattern are the formation of small "bumps" ahead of the wedge, and the occurrence of self-contacts between successive bumps that result in fold-like features (folds) on the surface. With sliding, these bumps grow in surface height to ~ 50–100 μm while coming closer together and interacting to make a self-contact (fold); one such fold may be seen at white arrow in Fig. 27.3a. The propagation and evolution of this fold is shown in Fig. 27.3b, c. Fold formation precedes contact with the wedge face. Additional folds, quite well developed, occur near the crest of the prow in Fig. 27.3b. The development of these folds is also well revealed by observing the changes in the curvature of 3 streak lines immediately below the "surface". As a fold traverses the wedge tip and leaves the contact, it changes its orientation by stretching and rotation leaving behind crack-like features on the surface in the wedge wake, see for example at white arrow in Fig. 27.3c.

Fig. 27.2 Prow formation and laminar flow in sliding of aluminum. *Streak lines* highlight the smooth flow. $\alpha = -75°$, $V = 5$ mm/s, 1,000 fps, fluid: Mobil 1. SD—sliding direction of WP
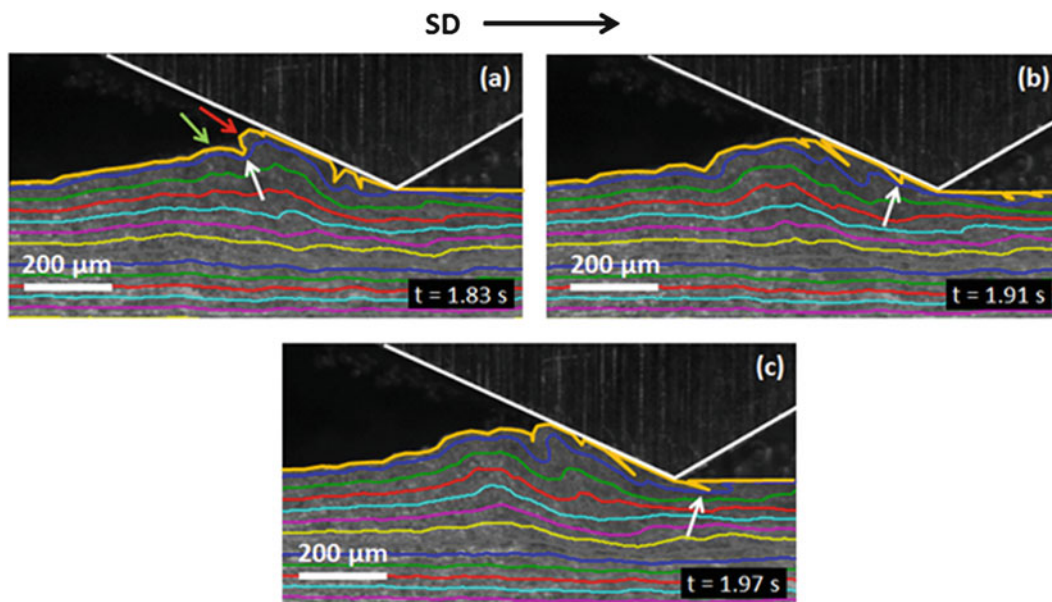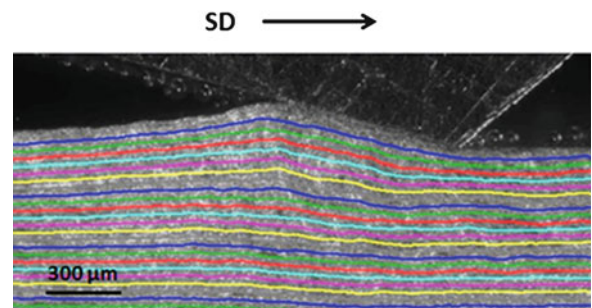
Fig. 27.3 Three frames with superimposed streak-lines from an image sequence showing development of bumps and folds in sliding of commercially pure aluminum. $\alpha = -65°$, $V = 5$ mm/s, fluid: Mobil 1, 1,000 fps. *Green* and *red arrows* show two consecutive bumps making self-contact (fold). The *white arrow* shows the evolution of a specific fold into a crack-like feature on the residual surface. The *bold yellow lines* demarcate the surface and various surface features. SD—sliding direction of WP
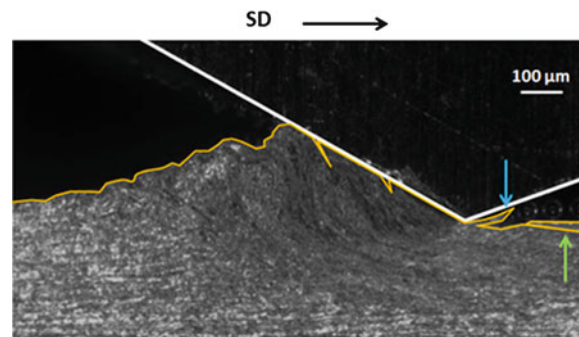
Fig. 27.4 A frame from a high-speed image sequence showing the formation of tear- and crack- like features on the Al surface. The *blue* and *green arrows* point to a tear and a complementary crack-like remnant formed, respectively, by the splitting of a fold as it exits the contact. $\alpha = -60°$, $V = 5$ mm/s, 1,000 fps, fluid: Mobil 1. SD—sliding direction of WP
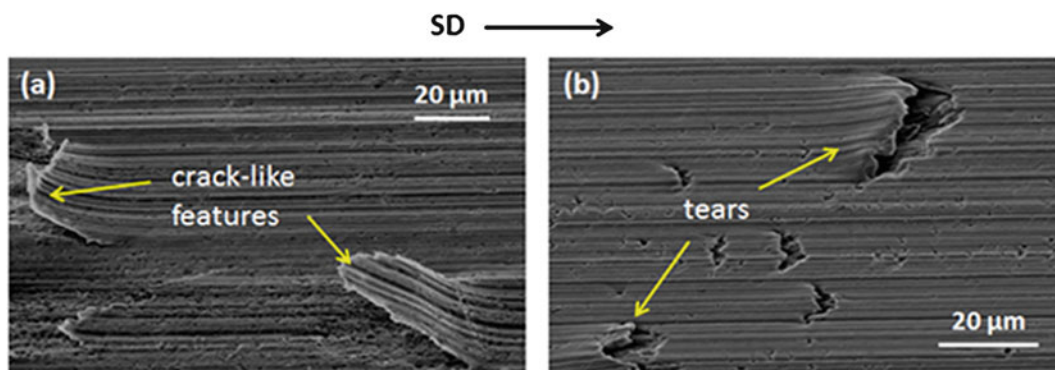
Fig. 27.5 Scanning electron micrographs showing (**a**) crack-like features and (**b**) tears on the aluminum surface in the wake of the wedge. SD—sliding direction of the workpiece

The streak lines at this sliding condition are quite sinuous indicating that the flow is highly non-laminar and vortex-like, in contrast to Fig. 27.2 [10]. This flow is also quite different from that commonly assumed in triboplasticity [7, 8].

Wear particle formation Bumps and folds become more frequent at smaller negative rake angles. Furthermore, a fold at these rake angles, as it emerges from the wedge region, is "torn off" from the surface leaving behind a tear (Fig. 27.4, blue arrow) and a complementary crack-like feature (Fig. 27.4, green arrow). The tear is usually seen to project upwards and in a direction opposite to the direction in which the crack-like feature is oriented. Often, these crack-like features are embedded in the subsurface and may not be visible in a surface examination. Metallographic examination of the subsurface, however, is adequate to reveal these embedded cracks. The occurrence of the tear and the complementary crack-like feature is likely due to the intense straining that the original fold undergoes in the sliding, see the discussion in the ensuing with regard to Fig. 27.6b. During a subsequent sliding pass of the wedge over this surface, the tear- and crack- like features are both "dislodged" from the surface resulting in wear particles.

Figure 27.5a, b are SEM images of the crack-like features and tears, respectively, on the WP surface in the wake of the wedge. These folds and tears have remarkable resemblances to features commonly observed on surfaces produced by sliding wear [5, 14], delamination [15] and even abrasion [16]. However, the past studies have attributed mechanisms unrelated to folding as the origins of these features. For example, delamination wear particles have been attributed to repeated straining of the surface over many sliding passes analogous to subsurface fatigue. Folding and the associated formation of tears and crack-like features particles have been observed also in other material systems such as copper and brass [10], and in both dry and lubricated sliding.

The observations show that particles with morphologies akin to those of wear particles can develop in sliding by a new type of folding mechanism and that this can happen in as little as 1–2 sliding passes. It suggests a mechanism for sliding wear that does not require chip formation by a cutting type process occurring due to asperities interacting with a surface. Chip formation
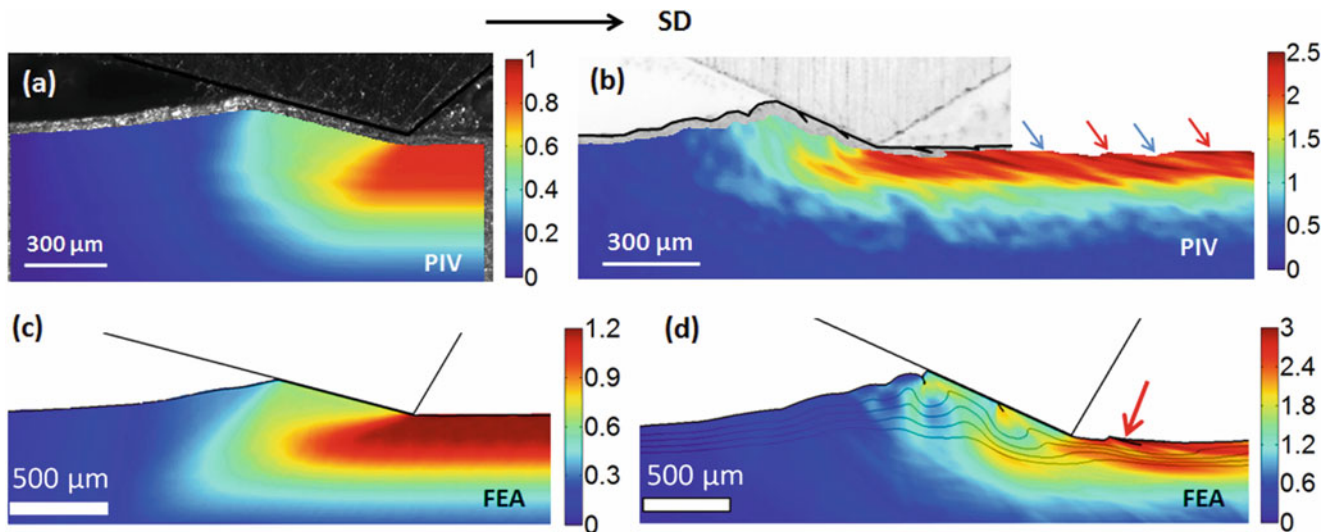
Fig. 27.6 Subsurface strain distribution for Al1100-O from experiment (PIV) and FEA. (**a**) and (**c**) show laminar type flow with steady, homogeneous strain field upon time averaging strains at spatial locations. (**b**) and (**d**) show non-laminar type flow (Fig. 27.3) with a highly heterogeneous strain field; in this case time averaging is not useful. The *red arrows* show strain localization on the folded surface while the *blue arrows* point to less strained regions adjoining the folds (unfolded surface). SD—sliding direction of workpiece

has often been used as the basis for explaining sliding wear in the so-called moderate-to-severe wear regime [7, 8]. But this type of wear particle formation requires asperity incidence angles much greater than those typical of sliding asperity contacts.

Strain field The strain distributions for the two cases of sliding corresponding to Figs. 27.2 (laminar) and 27.3 (non-laminar) are quite different, as seen from Fig. 27.6a (laminar, $\alpha = -75°$) and b (non-laminar, $\alpha = -65°$). The laminar strain field is quite homogeneous (Fig. 27.6a) with an average surface strain value of ~ 1. In contrast, the non-laminar strain field is characterized by significant strain heterogeneity (Fig. 27.6b) with the strain pattern showing a "lamellar" structure arising from the recurring folds. Each lamella corresponds to the residue from a fold. The peak strain value associated with a fold on the residual WP surface is $\sim$2.5, substantially higher than that noted for the laminar flow strain field.

Analysis of the folding Heterogeneous plastic flow arising from the grain structure in polycrystalline metals has been identified as the cause for the bump-formation (via differential deformation) and ensuing folding observed under the non-laminar flow conditions [10]. This was tested by FEA of the sliding. The tool was modeled as an analytical rigid surface, while the WP mesh consisted of about 55,000 4-noded, plane-strain (CPE4R) elements using reduced integration. The response of the Al 1100 WP was assessed for a heterogeneous workpiece "microstructure" and $\alpha = -75°$ and $\alpha = -65°$. Flow stress data of Al1100-O under quasi-static conditions (strain rate of 5×10^{-3}/s) was obtained from Armstrong et al. [17]. The strain-rate-sensitivity 'n' was estimated to be 0.0194 based on data from Lindholm et al. [18] and used to calculate yield ratios at higher rates. A 'microstructure' was created by subdividing the workpiece into cells using Voronoi subdivision. Two microstructural phases were introduced into the sample such that 45 % of the material (cells) is soft and 55 % is hard. The initial yield stress ratio between the soft and hard phases was set at 0.88, but with identical elastic properties. The finite-sliding contact pair algorithm was used to model contact between the wedge and workpiece. A capped Coulomb friction model, with $\mu = 0.2$ and maximum interfacial shear stress of 127.5 MPa, was used. Further details about this type of simulation may be found elsewhere [10].

Figure 27.6c, d shows the flow/strain fields in the laminar ($\alpha = -75°$) and fold-forming (non-laminar, $\alpha = -65°$) cases as computed by the FE simulations. In the former, the flow pattern is steady, and no bumps or folds are observed. The strain field shows good qualitative agreement with the measured strain field in Fig. 27.6a. However, at $\alpha = -65°$, the heterogeneous workpiece simulation reproduced the bump formation and folding as shown in Fig. 27.6d. The flow is highly unsteady (non-laminar type) in this case with much higher surface strains. It should be noted that a homogeneous specimen does not fold even when $\alpha = -65°$. The simulations, besides confirming the hypothesis about microstructure heterogeneity being central to the triggering of folds, reproduce key characteristics—flow pattern, strain fields—of the deformation noted in the experiments.

27.4 Conclusions

The high-resolution, in situ study of deformation produced by a sliding wedge has revealed unique aspects of unconstrained plastic flow at metal surfaces. An unusual, highly non-laminar type of flow is seen when the wedge has large negative rake angles typical of surface asperities in sliding or tools used in surface conditioning processes. This vortex-like flow triggers the formation of folds on the surface by self-contacts occurring between surface bumps initiated by the flow; and tears and crack-like features on the workpiece surface in the wake of the wedge. The tears and crack-like features can cause wear particle formation by delamination in as few as 1–2 sliding passes. The observations emphasize the importance of considering material heterogeneity in sliding, surface conditioning and machining processes. They also point to inadequacies in the common continuum approaches used to model sliding wear and surface deformation processes, which implicitly assume a "laminar" type of surface plastic flow. Another key feature of the unconstrained flow is occurrence of large residual plastic strains (>1) on the surfaces. The results demonstrate the power of the observational approach for characterizing large plastic flow phenomena in engineering systems.

References

1. Hutchings IM (2003) Tribology: friction and wear of engineering materials. Butterworth Heinemann, Oxford
2. Oxley PLB (1989) The mechanics of machining: an analytical approach to assessing machinability. Halsted, New York
3. Whitehouse DJ (1978) Surface topography and quality and its relevance to wear. In: Saka N, Suh NP (eds) Fundamentals of tribology: proceedings of the international conference on the fundamentals of tribology. MIT Press, Cambridge, pp 17–52
4. Tabor D (1977) Wear-A critical synoptic view. In: Glaeser WA, The American Society of Mechanical Engineers (eds) Wear of materials. The American Society of Mechanical Engineers, New York, pp 1–11
5. Samuels LE, Doyle ED, Turley DM (1981) Sliding wear mechanisms. In: Rigney DA (ed) Fundamentals of friction and wear of materials. American Society for Metals, Materials Park, OH, pp 13–41
6. Scott D, Seifert WW, Westcott VC (1974) The particles of wear. Sci Am 230:88–97
7. Johnson KL (1995) Contact mechanics and the wear of metals. Wear 190(2):162–170
8. Kopalinsky EM, Oxley PLB (1995) Explaining the mechanics of metallic sliding friction and wear in terms of slipline field models of asperity deformation. Wear 190(2):145–154
9. Guo Y, Saldana C, Compton WD, Chandrasekar S (2011) Controlling deformation and microstructure on machined surfaces. Acta Mater 59(11):4538–4547
10. Sundaram NK, Guo Y, Chandrasekar S (2012) Mesoscale folding, instability, and disruption of laminar flow in metal surfaces. Phys Rev Lett 109(10):106001
11. Gnanamanickam EP, Lee S, Sullivan JP, Chandrasekar S (2009) Direct measurement of large-strain deformation fields by particle tracking'. Meas Sci Technol 20(9):095710
12. Adrian RJ (1991) Particle-imaging techniques for experimental fluid mechanics. Annu Rev Fluid Mech 23(1):261–304
13. Raffel M, Willert CE, Wereley ST, Kompenhans J (2007) Particle image velocimetry: a practical guide, 2nd edn. Springer Verlag, Berlin
14. Rigney DA (1988) Sliding wear of metals. Annu Rev Mater Sci 18(1):141–163
15. Jahanmir S, Suh NP (1977) Mechanics of subsurface void nucleation in delamination wear. Wear 44(1):17–38
16. Malkin S (1989) Grinding technology: theory and applications of machining with abrasives. Halsted, New York
17. Armstrong PE, Hockett JE, Sherby OD (1982) Large strain multidirectional deformation of 1100 aluminum at 300 K. J Mech Phys Solids 30(1):37–58
18. Lindholm US (1964) Some experiments with the split Hopkinson pressure bar. J Mech Phys Solids 12(5):317–335

Chapter 28
Stiffness Investigation of Synthetic Flapping Wings for Hovering Flight

Kelvin Chang, Anirban Chaudhuri, Jayson Tang, Jordan R. Van Hall, Peter Ifju, Raphael Haftka, Christopher Tyler, and Tony Schmitz

Abstract Tests on a single active degree of freedom flapping platform are used to investigate the relationship between span-wise/chord-wise stiffness and hovering performance. The intended application is to establish constraints in a multi-objective optimization (thrust-power) that avoid selection of wings that perform poorly. It can also have utility as an alternative engine for identifying favorable performance. The procedure used to make the stiffness measurements is detailed along with the post-processing approach. Twelve wing designs, adapted from a previous study, were tested in both directions to extract a figure of merit that combines both stiffness values into a non-dimensional parameter (SC_{ratio}). The wings were also tested for thrust performance and current consumption across three different flapping frequencies (20, 25, and 30 Hz). A comparison is provided that identifies the added benefit of considering power consumption when selecting a wing for favorable performance. The data for 20 and 25 Hz flapping frequencies suggest a decrease in efficiency with increased SC_{ratio}, while the 30 Hz flapping frequency data was unimodal. This suggests the presence of a point or region on the spectrum of SC_{ratio} that provides optimum efficiency.

Keywords Optimization • MAV • Flapping • Stiffness • Thrust

28.1 Introduction

Flapping wing Micro Air Vehicles (MAVs) have captured the interest of many due to their high maneuverability and small package size, making them effective for surveillance and even deployment of small (<20 g) payloads. Their capability expands beyond their fixed wing siblings in that they can fly forward like most birds while also possessing the capability to stop on a dime and hover; they can even handle large gusts of wind. The hummingbird is a prime example of these capabilities. The AeroVironment Nano Hummingbird is an example of a feature rich controllable flapping MAV that has similar flight capability to a real hummingbird [1]. Improvements in technology and understanding of hovering flight over the next years will advance these types of flapping vehicles to smaller sizes and even more capability. One area of interest is the wing structure and, specifically how its reinforcement plays a role in hovering flight. A primary research task is to identify what type of reinforcement or stiffness is needed for efficient hovering.

Natural examples such as insects [2, 3] and hummingbirds [4] exhibit high levels of complexity. High fidelity models for hovering flight have yet to mature due to the complexity of how thrust is produced. Low Reynolds number unsteady flow modeled for different types of wings can be challenging/costly [5–9]. This provides an opportunity for experimental approaches to assist in developing a better understanding. Earlier work identified correlations between experimental wing deformation and thrust production [10]. In our previous work, improvements of the manufacturing and testing of the synthetic wings [11] was achieved by systematically decreasing the manufacturing and testing uncertainties. This led to a successful experimental optimization through surrogate-based optimization using adaptive sampling, where the output (thrust) from experiments acted as the objective of the optimization [12]. Since all MAVs have limited on-board power source, we also took

K. Chang (✉) • A. Chaudhuri • J. Tang • J.R. Van Hall • P. Ifju • R. Haftka
Mechanical and Aerospace Engineering Department, University of Florida, Gainesville, FL 32611, USA
e-mail: kc3635@ufl.edu

C. Tyler • T. Schmitz
Mechanical Engineering and Engineering Science Department, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

power into consideration and explored various multi-objective formulations [13]. Our future pursuit is a more expansive multi-objective optimization, utilizing both thrust and power consumption and a more generalized design space.

A study by Combes and Daniel identified scaling laws between geometric dimensions on the wing, such as span, and the flexural stiffness for a variety of insect forewings [14]. This suggests that stiffness is important for both efficiency and thrust production. This work can be effective in identifying whether scaling laws exist in a synthetic platform where there is only one active degree of freedom, and the materials differ from those used in nature.

This paper explores the effect of synthetic wing stiffness on the thrust production and power consumption. The objective is to use the stiffness ratio between span-wise and chord-wise measurements as a simple indicator of efficiency. We study the effect of the stiffness ratio on the thrust, power and an efficiency indicator (thrust/power). The final step is to use a surrogate (or response surface or metamodels) for approximating the stiffness ratios of a wing and then use the surrogate in a multi-objective optimization as an extra constraint, deterring sampling in the region that is inefficient. This can save time and prevent waste of resources. In this paper, we explore this idea of identifying efficient wing designs through stiffness ratio measurements. The simplicity of these measurements make it easy to model, and can provide alternative optimization opportunities.

28.2 Wing Construction

The components of the wing include a plastic frame, a unidirectional carbon fiber rod, and a nylon membrane. The plastic used for the frame was Acetal resin (DuPont™ Delrin®) and was chosen for is machinability and stiffness. The 250 μm thick plastic sheets were CNC machined using a four-fluted square endmill (1 mm diameter), using a 3-axis Haas TM-1 milling machine. A channel was milled into the leading edge along the whole span that assisted in locating a 0.5 mm diameter carbon fiber rod (Avia Sport Graphlite®). A cyanoacrylate adhesive was then utilized to bond the carbon rod to the frame. The assembly is completed by applying a Capran® nylon film (Honeywell) to the bottom of the skeleton, using an adhesive transfer film (Scotch® 3 M 9471LE) Fig. 28.1.

28.3 Thrust and Current Testing Procedure

A custom flapping mechanism is used to test the wing pairs. This single active degree of freedom device is fixed to a six axis force-torque sensor (ATI Nano 17), allowing for thrust and lift forces to be measured. A Maxon EC16 brushless motor drives a slider crank mechanism that oscillates the wing clamps Fig. 28.2. They allow for a stroke angle of 96° and are fixed in the other two rotational axis; therefore, the passive rotation of the wing is due to twisting of the wing structure. The wings are tested at 20, 25, and 30 Hz flapping frequencies, and data is collected for 10 s for each frequency. To account for any offset in the load measurement, a tare is taken for 6 s before the flapping begins. The average force in the 10 s window is used to represent the thrust at that frequency. A second order Butterworth low pass filter is used to reduce high frequency noise in the measurement. A 16-bit National Instruments USB-6251 data acquisition device was utilized, while a LabVIEW virtual
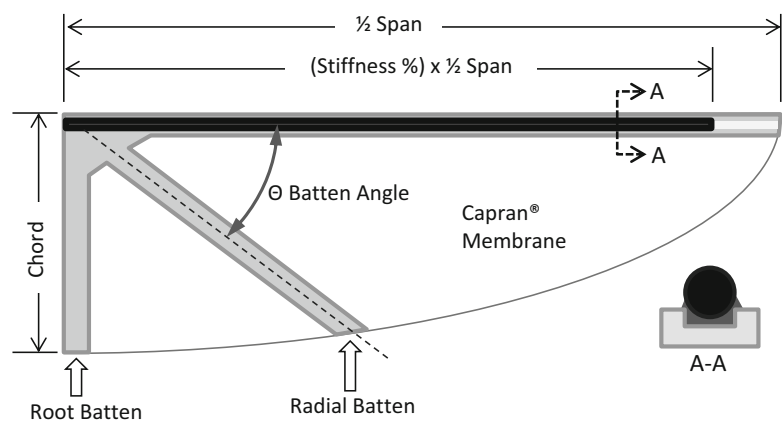
Fig. 28.1 The components and terminology used in this optimization are presented in the illustration. The 0.5 mm diameter commercial off the shelf carbon fiber rod is placed in a channel and extends a portion of the span described by the stiffness percentage design variable
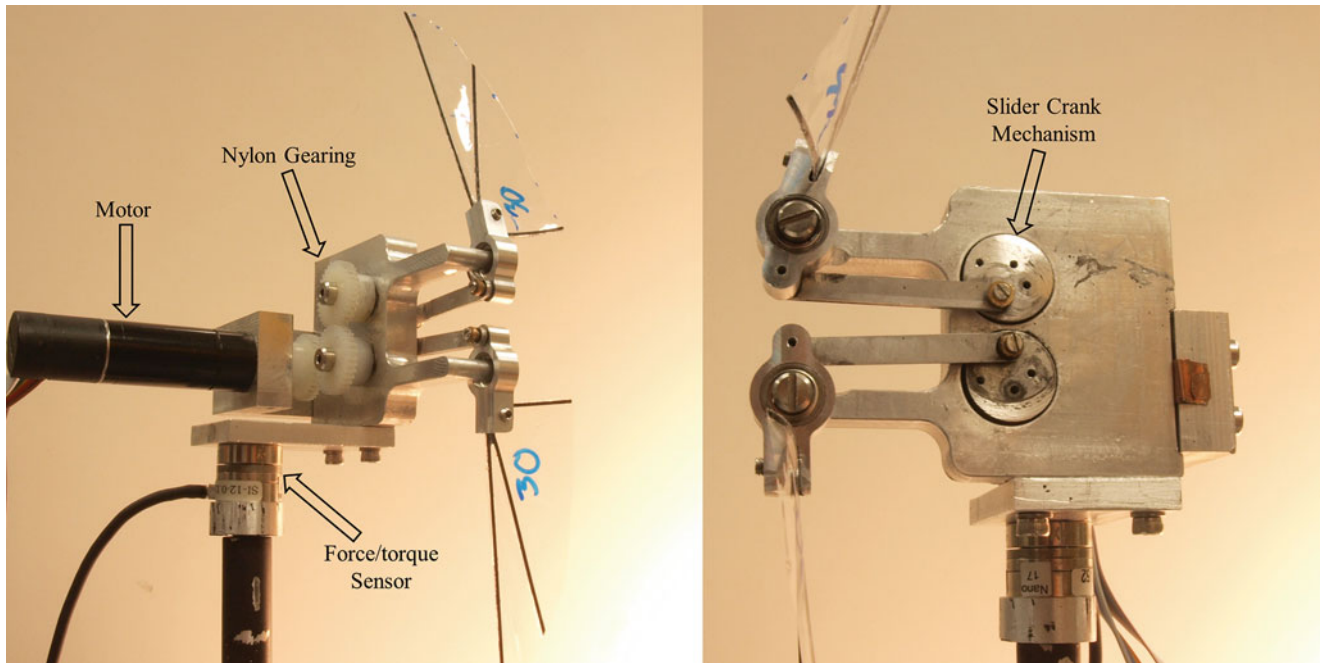
Fig. 28.2 Two views of the flapping mechanism

instrument was used to control and monitor the testing. Ten tests were conducted for each wing design, and the average of the ten trials was taken to be the wing's thrust performance. The current was acquired from measuring the voltage across a shunt resistor that was in line with the DC power supply. The average current measurement was taken in a similar manner to the thrust. The 6 s tare period accounted for any bias due to steady on electronic components such as the microcontroller. Averages for the current were taken for each frequency (20, 25, 30 Hz). The thrust efficiency indicator was taken to be the average thrust in grams over the average current in amps Eq. 28.1.

$$Efficiency\ Indicator = \frac{Avg.Thrust}{Avg.Current} \qquad (28.1)$$

28.4   Stiffness Testing Procedure

The stiffness measurements pertain to the slope of the load displacement data collected and were performed using a line load along the span and chord-wise directions at 50 % of their respective lengths as shown in Fig. 28.3. The slope of the load–displacement output was extracted in each direction and combined to make a non-dimensional parameter that describes the ratio of the stiffness measurements; it is termed the SC ratio Eq. 28.2.

$$SC_{ratio} = \frac{Spanwise\ stiffness}{chordwise\ stiffness} \qquad (28.2)$$

The experimental setup consists of a clamp to hold the wing, a line load fixture, and a load cell (ATI nano17) Fig. 28.4. The assembly is driven by a universal testing machine (Test Resources 311 Electromechanical frame), allowing for the displacement to be monitored continuously. The test was performed in displacement control at $10\ \frac{mm}{min}$, and the sampling rate was $10\ \frac{samples}{second}$.

The stiffness has dependence on the effective length and reinforcement, so each wing was tested until the load–displacement data became visually non-linear, leading to various maximum loads. The main focus was to extract enough data in the early linear regime, while avoiding the non-linear regime at large deflections. When the loading begins, the noise floor is significant. Fitting this data would create a false impression of the stiffness. The early stages of loading can also have incorrect stiffness because of incomplete engagement with both the leading edge and radial batten. Therefore, the

Fig. 28.3 The crosshatching
illustrates the boundary
conditions used in the stiffness
tests in both the span and
chord-wise directions. The
dotted lines show where
the line load was applied
for every wing
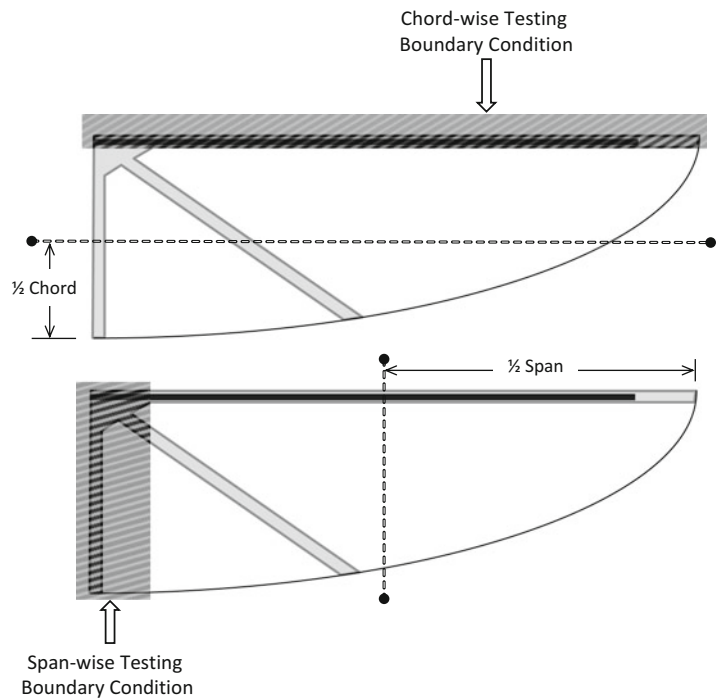
Fig. 28.4 The experimental
setup is shown for the stiffness
tests. The wing is lowered
onto the line load fixture,
while the load cell takes
measurements of force,
and the testing machine
monitors displacement
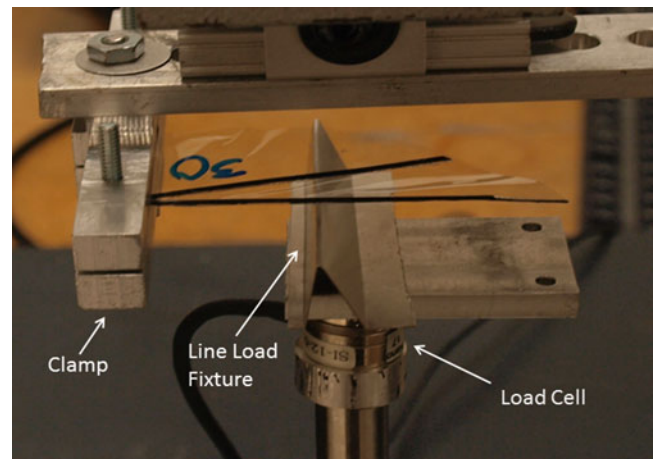
load-displacement data requires different quantities of deflection to arrive at a linear relationship. This is wing specific and is based on the flatness of the wing which is dictated by manufacturing errors. This can be seen in Fig. 28.5a.

To account for these issues, the data was filtered by discarding data leading to outlier residuals in the linear regression that pertained to percentiles beyond 99 %. With each 1 % of the data that was discarded, the remaining points were re-fit. The p-values for the coefficients of the linear fit were monitored to assure significant data was not being discarded. This was performed three times for each direction, accounting for 3 % of the data, and then the fit was deemed to be appropriate to use. This approach was preferred over the case specific removal of data, because it provided a systematic means of correcting for outlier data.

28.5 Results and Discussion

The designs used in this stiffness study stem from a preliminary single objective optimization [12] and subsequent thrust power Pareto front study [13]; they all follow the same design conventions. Both use three variables to fully describe the wing (Aspect Ratio #, Stiffness %, and Batten Angle) (Refer to Fig. 28.1) and were fabricated in the same manner.
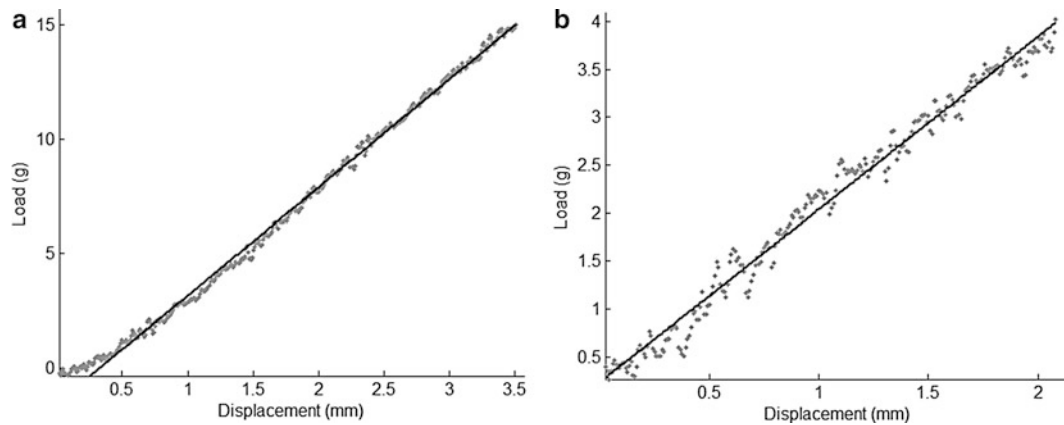
Fig. 28.5 (**a**) The span-wise load displacement data is fit with a linear regression model. (**b**) The chord-wise load displacement data in *grey* is fit as well. All the data points are provided, though 3 % were removed using a systematic filtering approach

Table 28.1 The wings that were tested along with their descriptions

Wing	Aspect ratio #	Stiffness %	batten angle	SC_{ratio}
1	3.6	100	80	1.15
3	6	100	30	0.177
13	5.2	100	80	0.275
1	2	95.6	58	7.64
8	1	90.1	21	35.4
9	4	99.8	73	2.01
20	1	83.4	42	23.1
23	4	91.3	22	5.63
30	3	81.4	16	2.61
32	4	85.2	13	1.39
34	2	88.9	48	15.3
36	2	97.5	36	12.3

Table 28.2 The average thrust and current are presented for each of the frequencies tested

	20 Hz flapping frequency		25 Hz flapping frequency		30 Hz flapping frequency	
Wing	Thrust (g)	Current (A)	Thrust (g)	Current (A)	Thrust (g)	Current (A)
1	2.362	0.056	4.427	0.107	6.083	0.165
3	3.521	0.072	7.166	0.160	10.698	0.326
13	3.095	0.061	5.648	0.122	7.628	0.218
1	1.004	0.033	2.302	0.065	3.528	0.095
8	0.653	0.031	1.581	0.059	2.336	0.086
9	2.437	0.050	5.038	0.101	7.205	0.159
20	0.864	0.032	1.887	0.060	3.193	0.087
23	2.219	0.049	4.572	0.103	6.076	0.164
30	1.761	0.045	3.772	0.090	5.507	0.137
32	2.166	0.049	4.752	0.105	7.282	0.173
34	0.659	0.035	1.690	0.067	2.834	0.098
36	0.886	0.034	2.133	0.067	3.310	0.100

Twelve wings were chosen to test based on a high stiffness percentage (>80), as the lower stiffness percentage wings were generally poor performers. The first three wing designs were identified as Pareto optimal wings in a study by Chaudhuri et al. [13]. The other nine were borrowed from the same study. For continuity, the wing numbers correspond to those used in the cited studies and are presented in Table 28.1. Table 28.2 provides the average thrust and current measured data for all the wings tested (10 trials for each wing design).
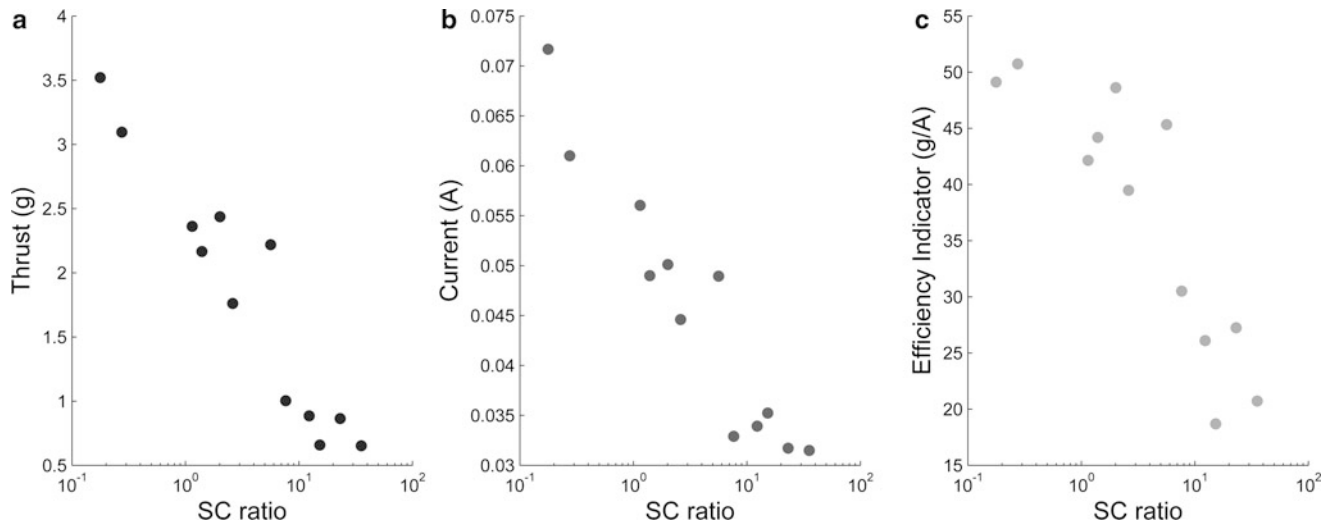
Fig. 28.6 (a) Thrust, (b) current, and (c) efficiency indicator plotted against SC_{ratio} for 20 Hz flapping frequency
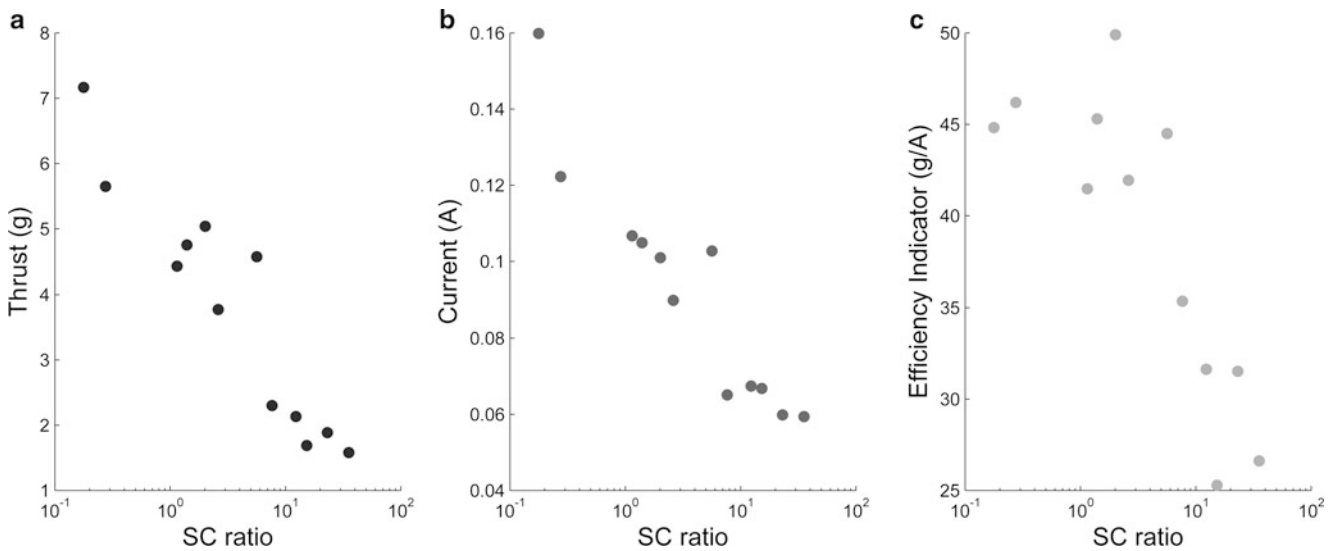
Fig. 28.7 (a) Thrust, (b) current, and (c) efficiency indicator plotted against SC_{ratio} for 25 Hz flapping frequency

The thrust performance for all of the flapping frequencies shows a trend of reduced thrust with higher SC_{ratio}. Based on prior experience, this is expected since limited chord-wise stiffness has shown to be detrimental to thrust production. The thrust output reduces dramatically, dropping by more than 80 % over two decades for all of the frequencies.

The current plots follow a similar trend in that the current consumption decreases with the increase of SC_{ratio}. This can be attributed to the excessive chord-wise compliance, causing a lower angle of attack over the stroke, and lending to reduced drag force and power required to drive the wing.

The 30 Hz efficiency indicator plot seems to follow a different trend; it seems unimodal. The wings that possessed high thrust output were found to consume considerable energy, making them less desirable. With the change of the distribution, there is evidence that using thrust or efficiency indicator as an objective will alter the results of the optimization. Using thrust as an objective would suggest wings with low SC_{ratio}, while using efficiency indicator as the objective would drive the optimization toward wings with a SC_{ratio} close to 1 Figs. 28.6, 28.7, and 28.8. Taking efficiency indicator as the optimization's objective opens consideration for energy consumption and paints a more complete picture of the wing's performance. This can be important in the application to MAVs which likely possess limited energy stores.

Observing the other frequencies, the results indicate a decreasing trend in the efficiency of the wing as the SC_{ratio} increases for both 20 and 25 Hz flapping frequencies. The data for 30 Hz flapping frequency is unimodal, suggesting that
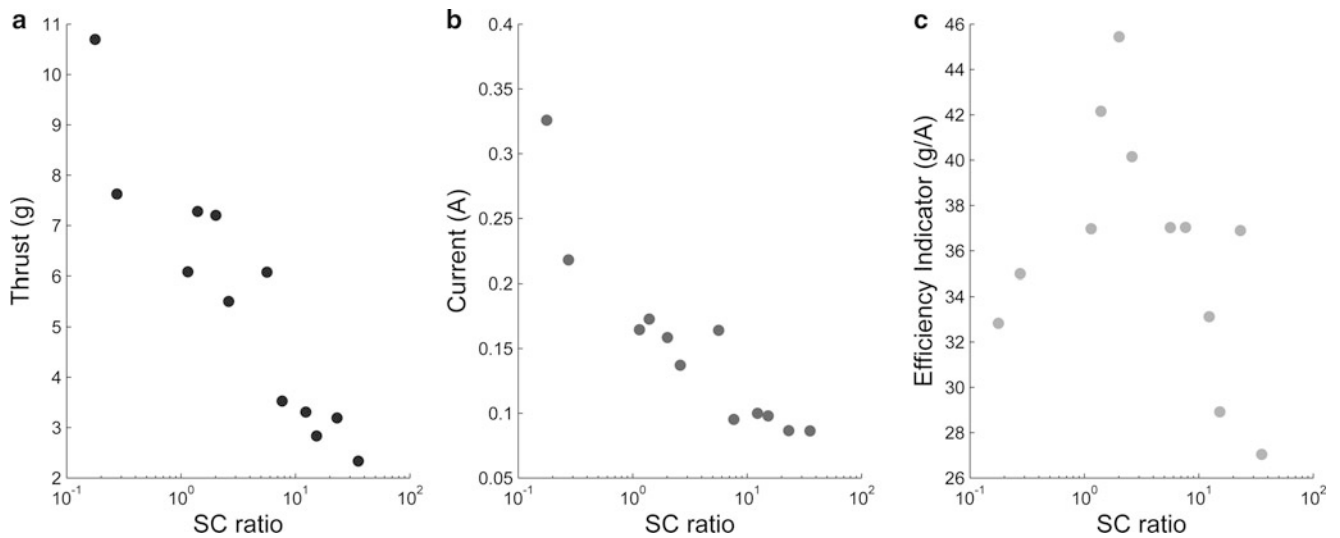
Fig. 28.8 (**a**) Thrust, (**b**) current, and (**c**) efficiency indicator plotted against SC_{ratio} for 30 Hz flapping frequency

Table 28.3 This illustrates a selection table that has application to an optimization as a constraint or to a user who is looking to select wing designs that possess performance beyond a specified threshold (e.g. >3 g thrust)

	SC_{ratio} range		
	20 Hz	25 Hz	30 Hz
Efficiency indicator >35 g/A	0.177–5.63	0.177–7.64	1.15–23.1

Table 28.4 The SC_{ratio} range is displayed for different example thresholds

Criterion	SC_{ratio} range at 30 Hz
Thrust >6 g	0.177–5.63
Current <0.15A	2.61–35.4
Efficiency indicator >40 g/A	1.39–2.01

The thresholds can be changed to accommodate different types of flapping wings (e.g. Different construction)

there is a "sweet spot," where the ratio of the span-wise and chord-wise stiffness makes for optimum efficiency. Based on prior experiences, this is justifiable since wings with excessive bias between their span-wise and chord-wise stiffnesses are expected to perform poorly.

From the perspective of all the efficiencies across the three frequencies, the 30 Hz flapping frequency data at lower SC_{ratio} values shows less efficiency compared to the other frequencies, then matches or surpasses the efficiency of the other two frequencies as SC_{ratio} increases. This indicates a tradeoff characteristic for the wing designs in that wings that are fit for low frequency flapping might not be as efficient at higher frequencies and vice-versa.

Table 28.3 provides an example of how a threshold can be utilized by an optimization or MAV designer. For instance, requiring the efficiency indicator to be greater than 3 g can be a constraint; this translates to a requirement that SC_{ratio} measurements stay within 1.15 and 23.1 for 30 Hz flapping frequency, otherwise it is unlikely that the wing will surpass the designated threshold. The simplicity of the measurement procedure lends to the possibility of creating a model that can predict SC_{ratio}, possibly eliminating the need to physically test wings to realize whether the wing meets the constraint. Table 28.4 provides more sample thresholds such as the requirement for minimum thrust. The threshold will depend on the wing and user preference, because this measurement procedure can be applied to wings of different construction and size. It can also be seen that using efficiency indicator as the criterion for constraint holds an advantage as it is easy to decide a common threshold across different frequencies, while thrust and current are frequency specific thresholds.

28.6 Conclusion

The results of this study have suggested the feasibility of using a simple static test to identify regions of a design domain with poor performance, taking a role as an extra constraint on the optimization. Future work will look to identify the uncertainty in the measurements with extended testing and attempt alternative test procedures. A model will be developed and calibrated such that it can extrapolate, predicting stiffnesses so that it can be used in conjunction with an optimization.

References

1. Keenmon M, Klingebiel K, Won H, Andriukov A (2012) Development of the nano hummingbird: a tailless flapping wing micro air vehicle. In: 50th AIAA aerospace sciences meeting, Nashville
2. Ellington CP (1999) The novel aerodynamics of insect flight: applications to micro-air vehicles. J Exp Biol 202:3439–3448
3. Young J, Walker S, Bomphrey RJ, Taylor GK, Thomas ALR (2009) Details of insect wing design and deformation enhance aerodynamic function and flight efficiency. Science 325(5947):1548–1552
4. Altshuler DL, Dudley R, Ellington CP (2004) Aerodynamic forces of revolving hummingbird wings and wing models. J Zool 264(4):327–332
5. Platzer MF, Jones KD, Young J, Lai JCS (2008) Flapping-wing aerodynamics: progress and challenges. AIAA J 46(9):2136–2149
6. Ansari SA, Zbikowski R, Knowles K (2006) Aerodynamic modeling of insect-like fapping flight for micro air vehicles. Prog Aerosp Sci 42:129–172
7. Shyy W, Aono H, Chimakurthi S, Trizila P, Kang C, Cesnik C, Liu H (2010) Recent progress in flapping wing aerodynamics and aeroelasticity. Prog Aerosp Sci 46(7):284–327
8. Shkarayev S, Silin D (2010) Applications of actuator disk theory to membrane flapping wings. AIAA 48:2227–2234
9. Sun M, Tang J (2002) Unsteady aerodynamic force generation by a model fruit fly wing in flapping motion. J Exp Biol 205:55–70
10. Wu P, Ifju P, Stanford B (2009) Flapping wing structural deformation and thrust correlation study with flexible membrane wings. AIAA 48(9):2111–2122
11. Chang K, Rue J, Ifju P, Haftka R, Schmitz T, Tyler C, Chaudhuri A, Ganguly V (2014) Analysis of thrust production in small synthetic flapping wings. In: Conference proceedings of the society for experimental mechanics, Chicago
12. Chaudhuri A, Haftka RT, Ifju P, Villanueva D, Chang K, Rue J, Tyler C, Schmitz T (2013) Experimental optimization and uncertainty quantification of flapping wing of a micro air vehicle. In: 10th world congress on structural and multidisciplinary, Orlando
13. Chaudhuri A, Haftka RT, Chang K, Van Hall JK, Ifju P (2014) Thrust-power pareto fronts based on experiments of a small flapping wing. In: 10th AIAA multidisciplinary design optimization conference
14. Combes SA, Daniel TL (2003) Flexural stiffness in insect wings I. Scaling and the influence of wing venation. J Exp Biol 206:2979–2987

Chapter 29
A Generic, Time-Resolved, Integrated Digital Image Correlation, Identification Approach

J.P.M. Hoefnagels, J. Neggers, Benoît Blaysat, François Hild, and M.G.D. Geers

Abstract A generic one-step Integrated Digital Image Correlation (I-DIC) inverse parameter identification approach is introduced that enables direct identification of constitutive model parameters by intimately integrating a Finite Elements Method (FEM) with Digital Image Correlation (DIC), directly connecting the complete time sequence of experimental images to the sought model parameters. The problem is cast into a transparent single-minimization formulation with explicit expression of the unknowns, being the material properties and, optionally, experimental uncertainties such as misalignments. The tight integration between FEM and DIC creates an information dialogue that yields accurate material parameters while providing necessary regularization to the DIC problem, making the method robust and noise insensitive. Through this method the versatility of the FEM method is translated to the experimental realm, simplifying the existing experiments and creating new experimental possibilities.

A convincing demonstration of the method has been achieved by successful identification of three challenging models from three very different (virtual) experiments, all thoroughly analyzed on accuracy and noise sensitivity: identification of a mixed-mode interface model from a mixed-mode delamination test, a 10-parameter history- and rate-dependent glassy polymer model from a simple tensile test, and simultaneous identification of two material models from a single bulge test of a structured membrane.

Keywords Digital image correlation (DIC) • Integrated DIC • Inverse parameter identification • Finite elements method (FEM) • Numerical validation

29.1 Introduction

This paper discusses a method which provides the direct identification of constitutive model parameters by intimately integrating an (analytical or FEM) mechanical model with Digital Image Correlation (DIC), connecting the experimentally obtained images for all time increments directly to the unknown material parameters. The problem is formulated as a single minimization problem which incorporates all the experimental data. It allows for precise specification of the unknowns, which can be, but is not limited to, the unknown material properties. The tight integration between the mechanical model and DIC creates enables direct identification of the unknown parameters while providing necessary regularization of the DIC problem, making the method robust and noise insensitive. Through this approach, the versatility of the FEM method can be translated to the experimental realm, enhancing the analyses of existing experiments and creating new experimental possibilities.

J.P.M. Hoefnagels (✉) • J. Neggers • B. Blaysat • M.G.D. Geers
Department of Mechanical Engineering, Eindhoven University of Technology, Den Dolech 2, 5612AZ, Eindhoven, The Netherlands
e-mail: j.p.m.hoefnagels@tue.nl

F. Hild
LMT Cachan, avenue du Préesident Wilson 61, 94235 Cachan Cedex, France

29.2 Principle

The basis of the introduced method is the global digital image correlation (G-DIC) (cf. [1, 2]) can be explained as follow. Considering two images, (f) and (g). (f) leads to the reference state, and (g) the response of the structure under a prescribed loading. One can introduce a cost function, $I\left(\vec{u}\right)$, a function of the displacement $\vec{u}$. This function gives a measure over a zone of interest (ZoI) between both pictures (f) and (g), when (g) has been corrected with the displacement $\vec{u}$:

$$I\left(\vec{u}\right) = \int_{ZoI} \left\{ f\left(\vec{x}_0\right) - g\left(\vec{x}_0 + \vec{u}\left(\vec{x}_0\right)\right) \right\}^2 dS \tag{29.1}$$

This approach is defined as a global one since the displacement has a definition over the entire zone of interest. The objective within the image correlation is to reduce this cost function, in other words, the objective is to find the optimal displacement, $\vec{u}^{opt}$, leading to the smallest value of $I\left(\vec{u}^{opt}\right)$. This step can only be conducted by defining $\mathcal{U}$, a kinematically admissible space for $\vec{u}$. One can then write:

$$\vec{u}^{opt} = \underset{\vec{u} \in \mathcal{U}}{\mathrm{Argmin}} \left\{ I\left(\vec{u}\right) \right\} \tag{29.2}$$

The novelty of the presented method lies in the definition that defines this admissible kinematic space with respect to the delamination test and the delamination mechanisms that occur. One denotes this Global-DIC as an integrated one aproach, i.e. I-GDIC, because information about the material behavior has been included within the Global-DIC procedure, see Fig. 29.1.

The method is demonstrated from three challenging models identified from three very different (virtual) experiments, all thoroughly analyzed on accuracy and noise sensitivity: (a) identification of a mixed-mode interface model from a mixed-mode delamination test, (b) a 10-parameter history- and rate-dependent glassy polymer model from a simple tensile test, and (c) simultaneous identification of two material models from a single bulge test of a structured membrane.
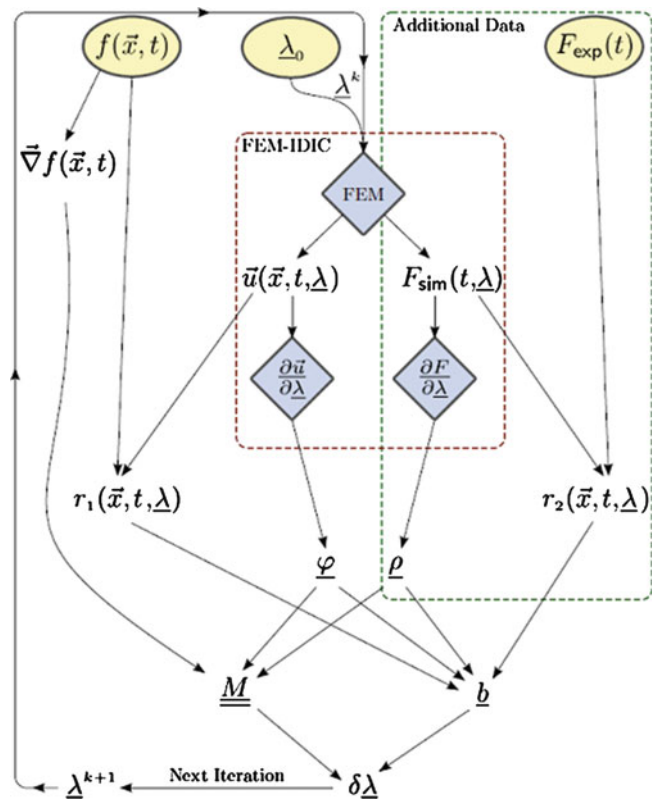
Fig. 29.1 Block diagram of the proposed IDIC framework

29.3 Identification of a Mixed-Mode Interface model from a Mixed-Mode Delamination Test

In this case, we are after the interface behavior, therefore, a parameterized description of the interface traction profile in the process zone [using 3 degrees of freedom (D.O.F.'s)] is used to define the kinematically admissible space for $\vec{u}$. This is schematically shown in Fig. 29.2a, b. The I-GDIC algorithm optimizes the displacement field to the measured displacement field, thereby directly yielding the correct crack tip opening displacement (CTOD) profile, Fig. 29.2d, but the routine also directly yield the optimum parameter values of the D.O.F.'s for describing the interface traction profile. Together, the optimized CTOD and interface traction profile can be combined to reconstruct the interface material behavior in the form of its traction-separation curve (Fig. 29.2d).

The optimized CTOD is extracted from the converged displacement solution, which is combined with the optimized interface traction profile (resulting from the optimized D.O.F.'s) to reconstruct the interface material behavior in the form of its traction-separation curve, which is shown in Fig. 29.3. For each of the 5 different material interface behaviors
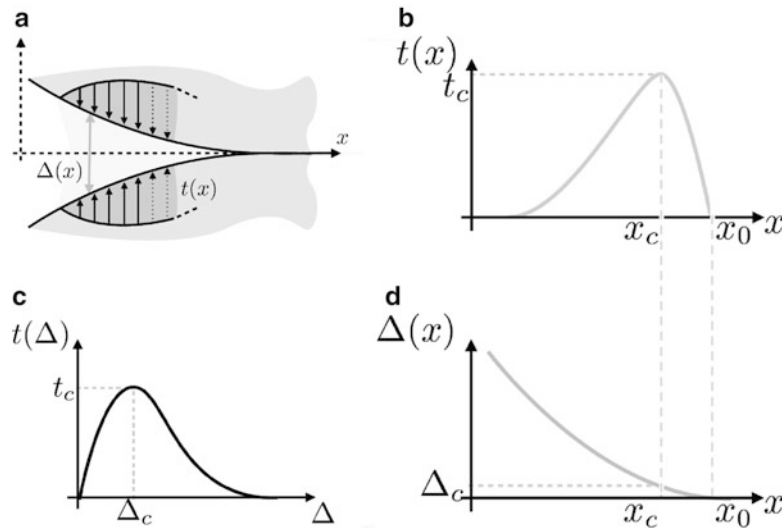
Fig. 29.2 (**a**) Schematic of the parameterized description of the interface traction profile in the process zone, where the *length of the arrows* indicate the magnitude of the traction. (**b, d**) the interface traction profile (**b**) and its corresponding crack tip opening displacement (CTOD) profile (**d**) depicted as function of its three parameters, i.e. t_c, x_c, and x_0, where x_0 is the position at the interface of first deviation from elastic interface behavior, while t_c is the maximum traction, which occurs at interface position x_c and which corresponds to the critical opening delta. The CTOD profile (**d**) results directly from the optimized displacement field reconstructed by the I-GDIC algorithm. (**c**) The reconstructed interface behavior in the form of its traction-separation curve (**c**) is computed by combining the interface traction profile with optimized parameters (**b**) with the corresponding optimized CTOD profile (**d**)
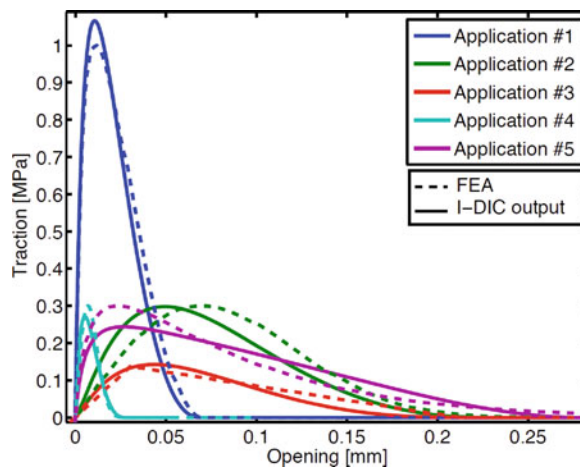
Fig. 29.3 Comparison of the input FE traction separation laws (*dashed lines*) with the traction separation law recovered with the I-GDIC routine (*solid lines*), for five different material interface behaviors used in the FE simulations. Note that these results are preliminary and may be subject to future change

tested, ranging in fracture toughness and ductility, two traction-separation curves are plotted: the input one (dashes lines) used for the finite element analysis and the output curve (red lines), which are recovered from the images through the introduced I-GDIC procedure. The figure shows that for all these different material interface behaviors the input traction-separation curve can be recovered fairly accurate, which indicates the accuracy and robustness of the developed I-GDIC algorithm is extracting the material interface behavior from measured images of the interfacial crack propagation of a delamination test.

29.4 10-Parameter History- and Rate-Dependent Glassy Polymer Model from a Simple Tensile Test (Preliminary Results)

As a second proof-of-principle example the constitutive properties of PolyCarbonate (PC) are identified. PC is a glassy polymer which is rate dependent (visco-elastic) and history dependent. The polymer chains tend to diffuse to a low energy state which compacts the material, thereby increasing the yield stress. Material flow induces rejuvenation which has a reverse effect on the yield stress, causing flow-induced softening. It is this softening behavior combined with the stress- and temperature-dependent viscosity which makes this material a challenging test case for the proposed IDIC method.

As a proof-of-principle experiment the identification procedure was initiated with an initial guess for each parameter equal to 110 % of their reference value. After 30 iterations convergence is reached, and values for all 10 EGP parameters have been identified. To evaluate the quality of the identification, the residual field (stack) has been analyzed. The residual field is reduced to almost zero, the only visible features are due to the sub-pixel interpolations.

A typical identification procedure for the EGP model requires a plethora of experiments, each using separate samples to identify each parameter individually. This is not only cumbersome, but identifying the history dependent aging parameter is troublesome since it tends to vary from sample to sample, or even within one sample. The proposed IDIC method showed the ability to identify all 10 EGP parameters under the condition that a reasonable initial guess is provided. The convergence radius can be improved by differentiating between parameters with high sensitivity and other which have less sensitivity (Fig. 29.4).

29.5 Simultaneous Identification of Two Material Models from a Single Bulge Test of a Structured Membrane (Preliminairy Results)

Instead of identifying multiple constitutive parameters of a single material, it is also possible to identify the constitutive parameters of heterogeneous samples. For instance the material response of micro-scale specimens where the geometrical length scale interacts with the intrinsic microstructural length scales, tends to deviate from the bulk material response and is influenced by neighboring materials or phases. For those cases, it is interesting to identify the materials closely to the situation where they are applied in the eventual device. Examples of this are the structured metal interconnects adhered to elastomer substrates as encountered in stretchable electronic applications.

The presented test case in this example consists of a 2 μm thick elastomer membrane (1×6 mm^2) with a line of high purity aluminum (Al) deposited on top of the elastomer membrane. The Al film has a thickness of 200 nm, and is 100 μm wide covering the full 6 mm length of the sample. The sample is loaded in a bulge test setup, where a pressure difference is applied causing the membrane to bulge outwards [3], see Fig. 29.5a. The bulge profile (including the surface roughness) is measured using optical confocal microscopy.

As a first proof-of-principle for this example, the proposed IDIC method is applied to identify the five constitutive parameters, see Fig. 29.5b, where the initial guess is taken equal to 130 % of to the reference values for the material parameters. Similarly, as with the previous (EGP) example, the residual field (stack) is analyzed for present patterns in the residue (not shown). The residual almost vanishes completely, which is expected since no additional acquisition noise was added. The only remaining visible features in the residual are the scars left behind by the sub-pixel interpolation.

For this virtual experiment, the reference displacement field is again known, enabling a direct comparison between the obtained and reference displacement field, see Fig. 29.5c. Compared to conventional DIC algorithms, where typically a displacement error of 1 % is considered as good, the achieved adequate accuracy of the proposed method is noteworthy, see Fig. 29.5d.
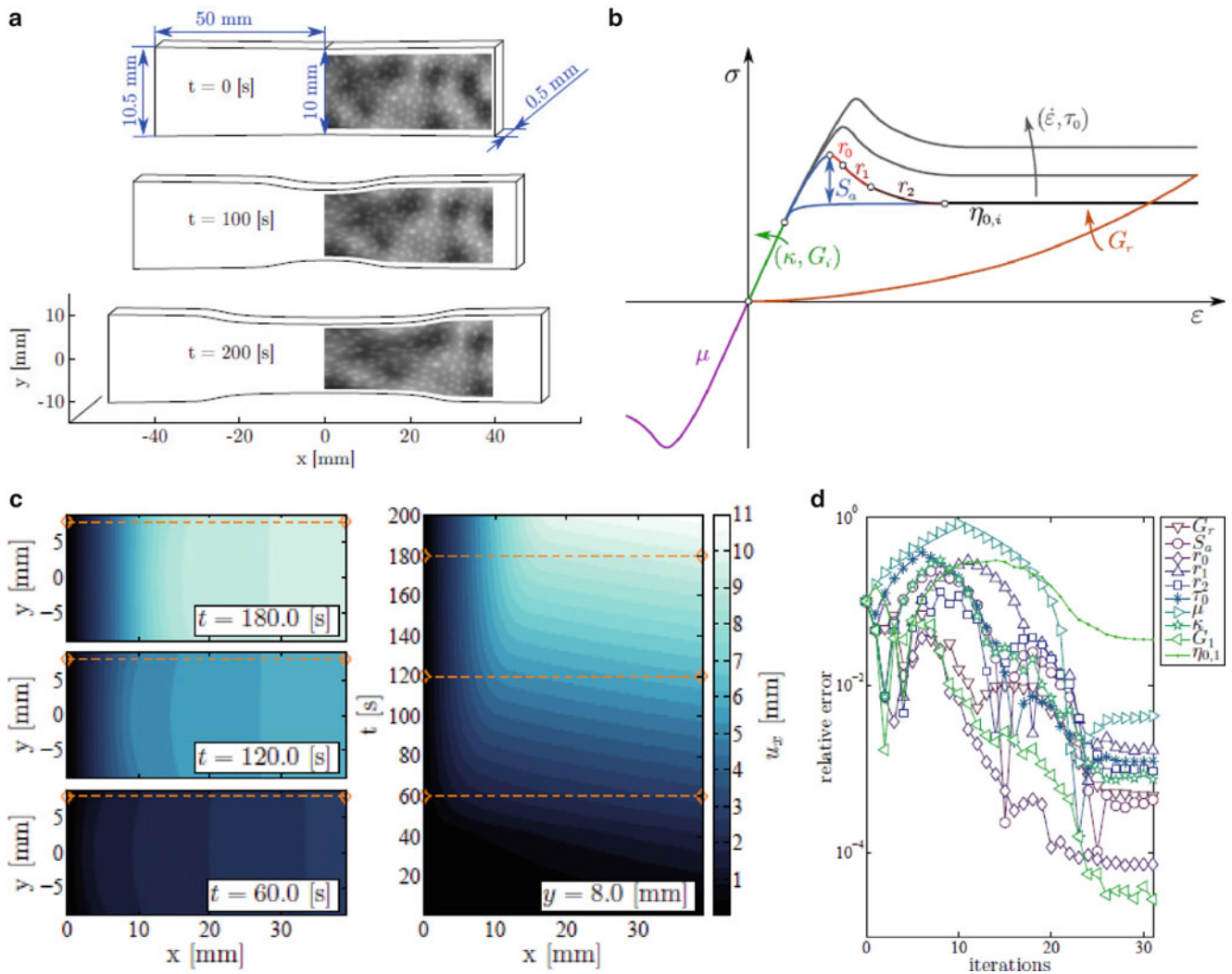
Fig. 29.4 (**a**) The contour of the mesh at the undeformed, intermediate, and deformed state. The adopted DIC pattern is depicted in the mesh. (**b**) A sketch of the influence of the 10 EGP parameters on the stress–strain response. (**c**) The x-component of the displacement field for three time increments and for a single y-plane, for all time increments. (**d**) The convergence behavior of the IDIC routine, identifying all 10 EGP material parameters

The bulge test example was chosen to prove that the proposed IDIC method can be applied to identify the constitutive parameters of multiple materials in one sub-structured monolithic sample. This is important, especially for micro-electronics applications, where different materials shaped in thin and small structures deviate in properties from their bulk counterparts. The identification of these parameters is a challenge at that scale. The presented example shows that the adopted identification procedure is robust (i.e. relatively insensitive to the initial guess and acquisition noise). Additionally, this example showed the relative ease with which the proposed IDIC method was extended to incorporate quasi-3D topographical profilometer data.

29.6   Conclusions

A generic one-step Integrated Digital Image Correlation (I-DIC) inverse parameter identification approach is introduced that enables direct identification of constitutive model parameters by intimately integrating a Finite Element Method (FEM) with Digital Image Correlation (DIC), directly connecting the complete time sequence of experimental images to the sought
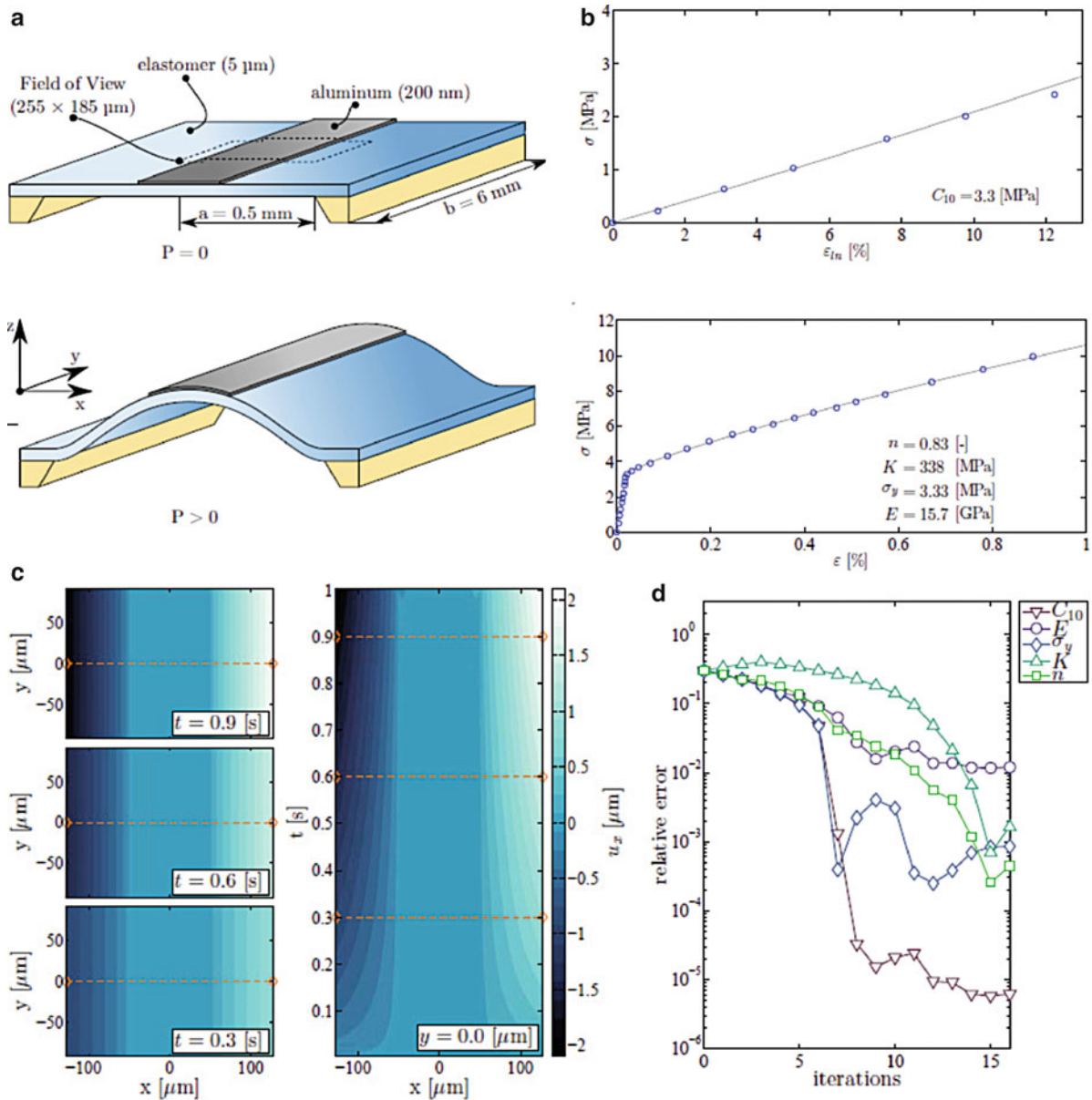
Fig. 29.5 (a) The elastomer membrane sample dimensions with aluminum structure on *top*, deformed to a bulged shape by a pressure difference. (b) The experimental reference material response and the fitted models which will be used as a reference material in the virtual experiment for the TPU elastomer substrate (*top*), fitted with a neo-Hookean model, and the high purity aluminum (*bottom*), fitted with an elasto-plastic model with isotropic hardening. (c) The x-component of the displacement field for three time increments and for a single y-plane, for all time increments. (d) Convergence of the IDIC method in terms of relative error for this bulge test example

model parameters. A convincing demonstration of the method has been achieved by successful identification of three challenging models from three very different (virtual) experiments, all thoroughly analyzed on accuracy and noise sensitivity: identification of a mixed-mode interface model from a mixed-mode delamination test, a 10-parameter history- and rate-dependent glassy polymer model from a simple tensile test, and simultaneous identification of two material models from a single bulge test of a structured membrane.

References

1. van Beeck J, Neggers J, Schreurs PJG, Hoefnagels JPM, Geers MGD (2013) Quantification of roughening in metal sheet stretching using global digital image correlation. Exp Mech. doi: 10.1007/s11340-013-9799-1
2. Neggers J, Hoefnagels JPM, Hild F, Roux S, Geers MGD (2014) Direct stress–strain measurements from bulged membranes using topography image correlation. Exp Mech. doi: 10.1007/s11340-013-9832-4
3. Neggers J, Hoefnagels JPM, Geers MGD (2012) On the validity regime of the bulge equations. J Mater Res 27(09):1245–1250

Chapter 30
Multiscale FE-Based DIC for Enhanced Measurements and Constitutive Parameter Identification

Laurent Robert, Jean-Charles Passieux, Florian Bugarin, Christoph David, and Jean-Noël Périé

Abstract Since they provide a large amount of information, full-field measurement techniques like Digital Image Correlation (DIC) allow the identification of several material parameters from a single non-homogeneous test. However the level of uncertainty associated with the identified parameters depends on the displacement measurement uncertainties, which are related to the spatial resolution of the measurement. To overcome the well-known compromise between spatial resolution and uncertainty, a multiscale approach to Finite Element DIC (FE-DIC) is proposed by considering additional nearfield images to improve locally the resolution of the measurement. An accurate estimation of the nearfield/farfield transformation is obtained by a dedicated global DIC method to bridge precisely the measurements at both scales. This multiscale FE-DIC measurement is then associated to a multiscale Finite Element Model Updating (FEMU) identification technique. After being validated on synthetic test cases, the method is applied to a tensile test carried out on an open-hole specimen made of glass/epoxy laminate. The four in-plane orthotropic elastic parameters are identified. Results show that the multiscale approach greatly improves the uncertainties of both the measured displacements and the identified material parameters.

Keywords Digital image correlation • Multiscale • Finite element method updated • Mechanical identification • Uncertainty assessment

30.1 Introduction

Over the three last decades, full-field measurement techniques have become increasingly popular in the community of experimental mechanics [1]. Among them, Digital Image Correlation (DIC) is probably the most used thanks to its (apparent) simplicity [2], although it suffers sometimes from a lack of information about its metrological performances [3]. DIC is also used for inverse constitutive parameters identification, e.g. from heterogeneous mechanical tests. In fact, with conventional identification methods for composite materials [4] more than one test is usually required for the identification of the constitutive parameters. Since they provide a sufficiently large amount of information, full-field measurement techniques allow the identification of several parameters from a single non-homogeneous test [5, 6]. For that purpose, several identification strategies based on full-field measurements have been recently developed [1, 7]. However a drawback of this approach is that obviously the level of uncertainty associated with the identified parameters depends on the displacement measurement uncertainties, which are related to the spatial resolution of the measurement [3, 8]. The most restricting DIC measurement uncertainties are the random errors, which are linked to the subset size (in the local DIC approaches [2, 3, 9]) or to the number of pixels per element (in the global DIC approaches) [8], thus defining the spatial resolution (expressed in pixel). More precisely, the higher the number of pixels per element (or per subset) is, the smaller uncertainties are, this compromise is classical to DIC. It is worth noting that a kinematic model (i.e. mesh) sufficiently rich to catch strain gradients, may lead to high spatial resolution, to the detrimental of larger measurement errors. One way to overcome the well-known spatial resolution/uncertainty compromise is to perform a multiscale DIC measurement by

L. Robert (✉) • J.-C. Passieux • F. Bugarin • C. David • J.-N. Périé
Université de Toulouse; ICA (Institut Clément Ader), INSA/UPS/Mines Albi/ISAE,
Campus Jarlard, F-81013 Albi, France
e-mail: laurent.robert@mines-albi.fr

adjusting, on a single specimen, the scales of the images to be taken in order to decrease the measurement uncertainty in specific interesting regions. In this paper, we address the problem of providing an accurate constitutive parameter estimation from a single non-homogeneous tensile test using a new nearfield/farfield multiscale approach that utilizes a Finite Element-based DIC measurement method and a Finite Element Model Updating (FEMU) identification procedure [5, 10, 11]. In a first step, we propose to use two cameras that acquire images with two different image resolutions (in pixel/mm) to measure the displacement fields by FE-DIC on the surface of the specimen. A series of images capture the full specimen (farfield images: at the scale of the structure), while a second series of images zoom on a structural detail (nearfield images: for example, in a local region where the displacement field is particularly sensitive to the parameters to be identified). The FE simulation mesh is used for the FE-DIC measurement at both scales. A fundamental point, an image registration process, based on a global DIC approach, have been developed to automatically and accurately reposition the nearfield image into the farfield image and to place precisely the mesh on the nearfield image. At this stage, measurement uncertainties are evaluated through synthetic unbiased images (shifted or strained as a virtual open-hole tensile test). In a second step, the proposed inverse multiscale identification method is presented. Based on the Finite Element Method Updated (FEMU), it takes advantage of the multiscale FE-DIC by using the same FE mesh without any displacement field interpolation. On the one hand the farfield FE-DIC measurement provides representative Dirichlet boundary conditions for the numerical simulation [6, 12]. On the second hand, the nearfield FE-DIC measurement provides a high spatial resolution kinematics field for the test/calculation displacement comparison, in a region where the model parameters are particularly sensitive. After being validated on a previous synthetic test case, the inverse method is applied to an open-hole tensile test of a laminated glass/epoxy coupon. Results show that the multiscale approach greatly reduces the uncertainty associated to the identified parameters.

30.2 Methods

30.2.1 Multiscale DIC Measurement and Image Registration

Digital Image Correlation [2, 13, 14] consists in seeking the displacement field **u** that register an image $\mathcal{I}_1$ into another image $\mathcal{I}_0$ of a specimen in two different loading conditions. Following [8], a weak form of the optical flow conservation is written globally over the whole region of interest (ROI):

$$u^{\star} = \underset{u(x)}{\arg \min} \int_{ROI} (\mathcal{I}_0(x) - \mathcal{I}_1(x + u(x)))^2 dx \qquad (30.1)$$

In practice, the unknown displacement field $u : \mathcal{I}_0 \to \mathcal{I}_1$ is sought in an approximation subspace $\mathcal{U}^N$, spanned by a finite dimension interpolation basis $\phi_i(\mathbf{x})$ as follows:

$$u = \sum_{i=1}^{N} \phi_i(x) q_i \qquad (30.2)$$

where **q** is the corresponding vector of degrees of freedom (dof) q_i. A large choice of interpolations can be used in this framework (see [15] and references within). In this work, a DIC method based on Finite Element kinematics (FE-DIC) has been developed following [8, 16]. The stationarity conditions associated to the minimization of the linearized problem (30.1) yields a set of linear systems:

$$Mq^k = b^k \qquad (30.3)$$

where **M** is a $N \times N$ matrix called the correlation operator and **b** the corresponding right-hand-side:

$$M_{ij} = \int_{ROI} \left(\nabla \mathcal{I}_0^{\top} \phi_i\right)\left(\nabla \mathcal{I}_0^{\top} \phi_j\right) dx$$

$$b_i^k = \int_{ROI} \left(\nabla \mathcal{I}_0^{\top} \phi_i\right)\left(\mathcal{I}_0(x) - \mathcal{I}_1(x + u^{k-1})\right) dx$$

where $u^{k-1} = \sum_{j=0}^{k-1} \sum_{i=1}^{N} \phi_i \, q_i^j$ is the approximation of the displacement at iteration $k - 1$. Since $x + u^{k-1}$ may be non-integer, a classical spline gray level interpolation is used to evaluate the right-hand-side. The definition of the approximation subspace $\mathcal{U}^N$ has a direct impact on the accuracy of the estimation. First N should be far lower than the number of pixels in the ROI because of the ill-posedness nature of the correlation problem. Namely, the larger N, the larger are measurement uncertainties. However, $\mathcal{U}^N$ should be rich enough to accurately represent the a priori unknown displacement. Namely, measurement uncertainties result from a compromise between the accuracy of the displacement interpolation, that manages what was called the "mismatch error" in [3], and the so-called "ultimate error" when the displacement interpolation is sufficiently accurate according to the true deformation of the image [17]. In the case of FE-DIC, a mesh that would be optimal for simulation purposes, is not necessarily optimal for the DIC measurement. Thus the choice of a common mesh is not, in general, an easy task in the context of identification. Most often, digital images are taken at one single resolution. In this case, a first way to use a simulation mesh including small elements is to search only for solutions that have some mechanical [18] regularity. The aim of this article is to explore another route, which consists in using images of the same speckle at more than one resolution.

The multiscale (or multi-resolution) measurement technique proposed herein, is designed to adapt the image resolution to the mesh and not the reverse. So the case of image pairs taken from two cameras at two different scales is considered, one being denoted *farfield* and the second *nearfield*. It is important to note that in this work both the optical axes of the cameras are set to be perpendicular to the flat specimen. Typically, the nearfield camera takes pictures which cover approximately the whole gauge region, while the farfield camera concentrates over a smaller area with a ratio between the two optical resolutions of around 5. So for practical considerations, the nearfield camera is mounted on a translation stage and is retracted to take a picture with the farfield camera. A key point of the proposed multiscale DIC method thus consists in registering accurately the nearfield reference image $\mathcal{I}_0^n$ in the farfield reference image $\mathcal{I}_0^f$. In this study, this multiscale transformation $t(x)$ is estimated by considering only the speckle of both scales. Thus a dedicated DIC algorithm is devised for that purpose. Indeed, $t(x)$ is sought to minimize the gray level conservation equation:

$$t^{\star} = \arg \min_{t(x)} \int_{ROI} \left(\mathcal{I}_0^n(x) - \mathcal{I}_0^f(x + t(x)) \right)^2 dx \tag{30.4}$$

where the ROI corresponds, here, to the entire nearfield image $\mathcal{I}_0^n$. In the context of the near/farfield registration, and since the studied specimen are assumed to be planar, the proposed contribution consists in seeking the transformation $\mathbf{t}$ as an arbitrary homography H relating $\mathcal{I}_0^n$ to $\mathcal{I}_0^f$. An homography is an invertible mapping of points (and lines) on the projective plane $\mathbb{P}^2$, represented by a non-singular 3×3 matrix H. In other words, for a given point $x \in \mathcal{I}_0^n$ and its corresponding point $x_i' \in \mathcal{I}_0^f$ we have the constraint: $x' = Hx$. Note that H can be multiplied by an arbitrary non-zero constant without modifying the projective transformation. Thus H is an homogeneous matrix with only 8 degrees of freedom even though it contains 9 parameters. So, the dimension of the approximation subspace is reduced to 8 for the whole nearfield image. Finally, the solution is computed by a Levenberg-Marquardt algorithm applied to the following problem:

$$H^{\star} = \arg \min_{H \in \mathcal{M}_{3,3}(\mathbb{R})} \int_{ROI} \left(\mathcal{I}_0^n(x) - \mathcal{I}_0^f(Hx) \right)^2 dx \tag{30.5}$$

where the numerical integration over the ROI $\mathcal{I}_0^n$ is performed by a mid-pixel rectangle method following [15]. Since the scales are very different between near and farfield images, this algorithm has to be initialized with a coarse approximation of the homography, estimated between two images by finding a set of r matched points (m_i, m_i'). Several algorithms have been compared for extracting and matching interest points, in most of the examples that have been processed, the SIFT algorithm was the most efficient because it provides a large sets of matched points for most of our configurations. Details of the implementation can be found in [19].

The method has been first applied on synthetic multiscale images. The set of synthetic speckle-pattern images was obtained using the TexGen software, already used in the French DIC community for assessing errors and uncertainties of DIC packages, see [3, 17, 20] for details of the software. We just recall here that the speckle-pattern image is generated by a photometric mapping and an 8-bit digitization of a continuous texture function (that can be deformed as wanted) computed for each integer pixel of the image. In this work, in order to perform a virtual mechanical test, the continuous displacement field is calculated from the analytical solution $\mathbf{u}^L$ of an infinite orthotropic open hole plate in vertical remote tension. Theoretical solution of this problem has been already used for identification purposes in [5]. Furthermore it has been also shown in [5] that the region near the hole is very sensitive to the constitutive parameters. A first pair (farfield images, with a

Fig. 30.1 Reference 1, 000 ×
1, 000 pixels definition image
generated with TexGen for the
farfield image $\mathcal{I}_0^f$ (*left*) and the
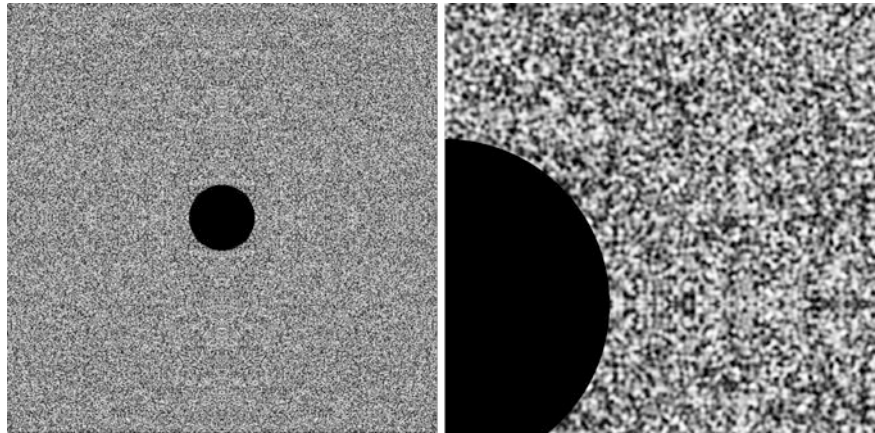nearfield image $\mathcal{I}_0^n$ (*right*)

Fig. 30.2 The raw discrepancy
map $\mathcal{I}_0^n(x_i) - \mathcal{I}_0^f(Hx_i)$ (in gray
levels) with the initial
homography estimated by SIFT
(*left*) and the one measured by the
global DIC technique (*right*)
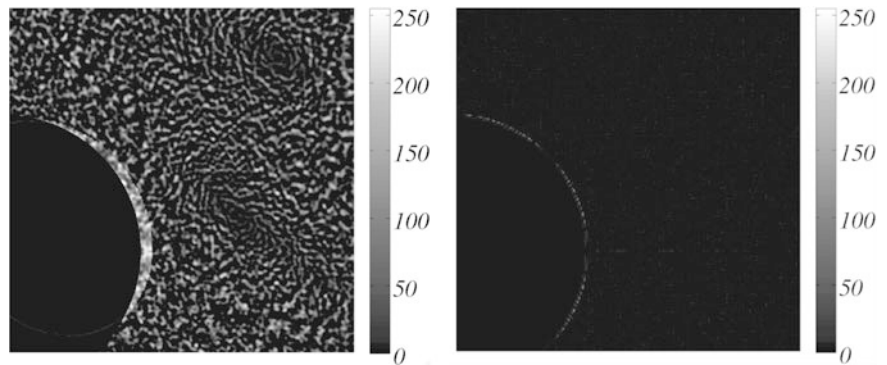
FOV of 26 mm wide) of images $(\mathcal{I}_0^f, \mathcal{I}_1^f)$ is generated with a resolution of 26 μm per pixel. A second pair (nearfield images, with a FOV of 5.2 mm wide) of images $(\mathcal{I}_0^n, \mathcal{I}_1^n)$ is generated, but with a resolution of 5.2 μm per pixel (5 times smaller). The mean size of the speckle is 34 μm, which corresponds to 1.3 pixels in the farfield images and 6.5 pixels in the nearfield images (Fig. 30.1).

Results in term of raw discrepancy maps (in gray levels for a 256 gray levels range) $\mathcal{I}_0^n(x_i) - \mathcal{I}_0^f(Hx_i)$ are plotted in Fig. 30.2. Theoretically, with exact matched points the result of the initial value of H estimated by SIFT should be accurate. But, in practice, the couples (m_i, m_i') are not properly matched. The gray level conservation is also poorly verified since the standard deviation of the residual map is equal to 23.5 % of the dynamic range of the image, see Fig. 30.2(left). However it provides a sufficiently good approximation to initialize the DIC algorithm of equation (30.5). The latter provides a very good estimation of the transformation H since the gray level conservation seems accurately verified: standard deviation of the discrepancy map is equal to 1.14 % of the dynamic range, as shown in Fig. 30.2(right).

30.2.2 A Priori Error Analysis of Multiscale Images

As mentioned above, in DIC, the total measurement uncertainties is classically viewed as a competition of the so-called ultimate and model errors [3, 21]. Knowing the exact analytical displacement field $\mathbf{u}_{ref}$ in the synthetic images (see Sect. 30.2.1), it is possible to compute the *random error* of a displacement field $\mathbf{u}$ measured by FE-DIC as the standard deviation $\sigma(\cdot)$ of the discrepancy $\Delta u = u - u_{ref}$ over the pixels overlapped by finite elements within the ROI. The discrepancy term $\Delta\mathbf{u}$ can be expressed in several chosen uncertainty:

- **ultimate error**. A simple rigid body translation along *x*-axis is imposed to the reference image, here by a shift in the Fourier space [8]. Such a displacement field does belong to the finite element approximation subspace $\mathcal{U}^h$ so the ultimate error only considers the errors inherent to the DIC technique in situation where the adopted kinematic model of the DIC algorithm perfectly fits the actual displacement field in the image [17]. In this work, the magnitude of the *x*-component of the displacement is set to $u_{ref}^{shift} = 0.5$ pixel since it maximises the standard uncertainty in the case of noiseless images [17].

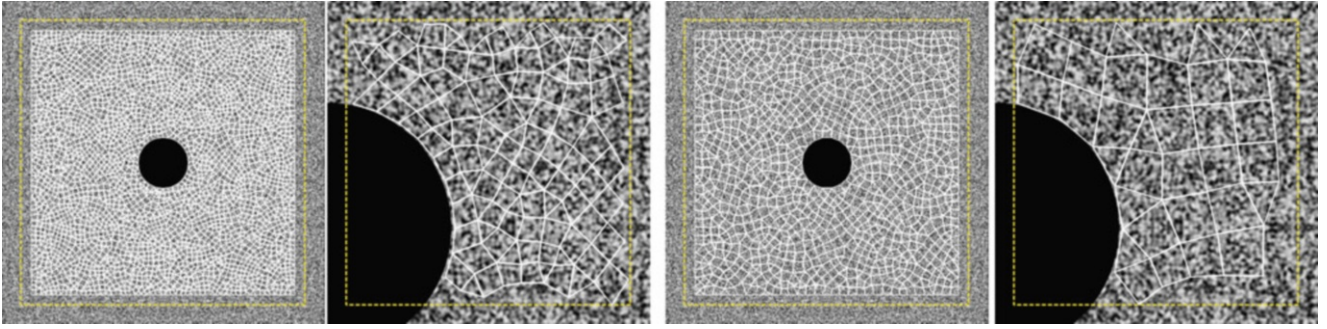**Fig. 30.3** Example of unstructured meshes used for the a priori uncertainty analysis: from left to right and for the farfield images, the average element size is respectively 411 and 719 μm (which corresponds to 16 and 28 pixels per element width), and 79 and 138 pixels per element width in the associated nearfield image

Fig. 30.4 Evolution of the ultimate (*wedge*), model (*circles*) and total (*solid line*) random errors as a function of the element size in millimeter for both nearfield (*in red*) and farfield images (*in black*) (Color online)
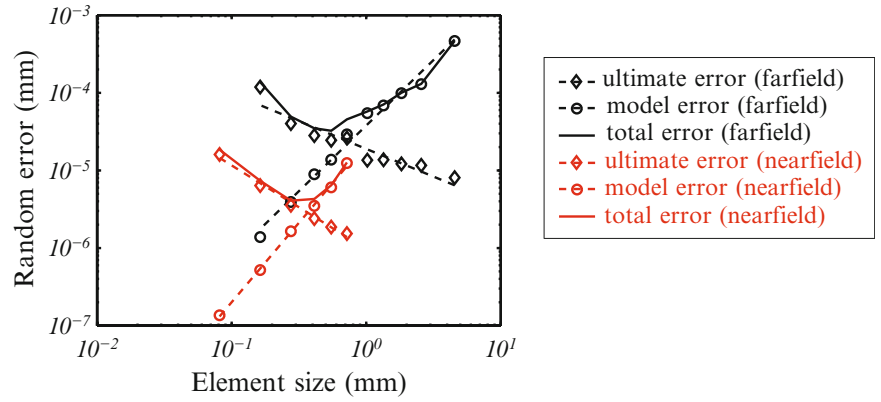
The FE-DIC measurement yields an inexact displacement map $\mathbf{u}_m$ which is used to estimate the ultimate random error $\sigma^{ult} = \sigma(u_m - u_{ref}^{shift})$

- **model error**. It corresponds to the so-called kinematical *interpolation error* in the computational mechanics jargon [21]. It only consists in the evaluation of the distance between a non-constant analytical displacement field and its projection on the finite element approximation subspace $\mathcal{U}^h$. No DIC is performed at this stage. In this paper, the mechanical analytical displacement field described in the previous section serves as the reference u_{ref}^L. Its projection $\mathbf{u}_{proj}$ on the FE approximation subspace is computed in the least square sense:

$$u_{proj} = \arg \min_{u \in \mathcal{U}^h} \sum_{i=1}^{m} \left(u_{ref}^L - u \right)^2 \tag{30.6}$$

which only requires the resolution of a linear system whose operator is the finite element mass matrix. The model error is thus estimated by $\sigma^{mod} = \sigma(u_{proj} - u_{ref}^L)$

- **total error**. The mechanical analytical field u_{ref}^L is prescribed to the reference image as described in Sect. 30.2.1. The measured displacement $\mathbf{u}_m$ between these synthetic images is computed by performing a FE-DIC measurement. The total error is then computed as $\sigma^{tot} = \sigma(u_m - u_{ref}^L)$. This quantity, which measures the exact error between the measured and reference displacement maps, takes into account both sources of uncertainties.

This a priori performance analysis is performed with both fields of view as a function of the characteristic mesh size (which corresponds to the *spatial resolution* i.e. the subset size for subset-based DIC approaches). Therefore, a set of eleven unstructured finite element meshes are generated. Their elements size are rather tightly clustered around the mean value that ranges from 78 μm to 5 mm, as shown in Fig. 30.3. The mesh is adjusted on the farfield image and transferred to the nearfield image thanks to the inverse of the optimized homography $(H^\star)^{-1}$, ensuring that the same meshes are used for both near and farfield images analyses. Figure 30.4 presents the evolution of ultimate, model and total random errors in millimeter as a function of the element size in millimeter, for both nearfield (in red) and farfield (in black) images. When nearfield

and farfield analyses are considered independently, it can be observed that the larger the elements (or equivalent, the more pixels per element), the lower is the ultimate error. Conversely, the larger the elements, the higher is the model error. The overmentioned compromise can be seen graphically on this figure, since the total error results from the competition of these two antithetical behaviors. When nearfield and farfield curves are compared, it appears that the model errors seems to broadly follow the same trend. Theoretically, they should be aligned, since this error simply depends on the physical mesh size (mm). In practice, it is not exactly the case, since the number of elements considered for computing this error is not the same in nearfield and farfield analyses as shown in Fig. 30.3. Conversely, the ultimate error associated to the nearfield image is much lower than that of the farfield, for a given element size. This gain can be explained, almost in its entirety, by the resolution ratio. As a result, the total error is logically shifted by the same ratio, along the direction of the model error. As a conclusion, for a given physical element size, the total error is thus significantly reduced thanks to such a multiscale approach. In other words, for a given target error, the multiscale measurement makes it possible to use much smaller elements. Consequently, it makes the use of a simulation mesh for measurement purposes more flexible.

30.2.3 Multiscale FEMU

The Finite Element Model Updating (FEMU) method is a popular, intuitive and versatile identification technique [7, 10]. It consists in updating a set of p constitutive parameters p in a FE analysis in order to reduce, in the least squares sense, the distance between the measured and the simulated quantities. Different optimization techniques, norms and cost functions may be used to exploit measured displacement fields [5, 7, 10, 11]. In our case, a hybrid cost function is built as follows:

$$p^\star = \arg \min_{p \in \mathbb{R}^p} \sqrt{\parallel R_u(p)\parallel^2 + \parallel R_f(p)\parallel^2} \tag{30.7}$$

and the displacement and force residuals read respectively:

$$R_u(p) = \frac{u_s(p) - u_m}{\parallel u_m \parallel} \text{ and } R_f(p) = \frac{F_s(p) - F_m}{F_m} \tag{30.8}$$

where $\mathbf{u}$ denotes the displacement dof vector and F the resultant force while $._m$ and $._s$ stand respectively for the measured and the simulated quantities. A Levenberg-Marquardt algorithm is used to solve the minimization problem (30.7). Instead of using this L^2-norm, one could alternatively solve a weighted least squares problem. In particular, the FE-DIC correlation matrix $\mathbf{M}$ is related to the inverse of the covariance matrix for the degrees of freedom [8]. The use of $\mathbf{M}$ allows thus for a convenient weighting of the degrees of freedom [22]:

$$p^\star = \arg \min_{p \in \mathbb{R}^p} \sqrt{\parallel R_u(p) \parallel_M^2 + \parallel R_f(p) \parallel_2^2} \tag{30.9}$$

where the $\mathbf{M}$-norm is defined as $\parallel R_u \parallel_M^2 = R_u^T M R_u$. It should be noted that the matrix $\mathbf{M}$ is symmetric and positive by construction. When invertible, it can be decomposed using the Cholesky decomposition as follows $M = LL^T$. Thus the M-norm of $\mathbf{R}_u$ can be written using the L^2 norm as $\parallel R_u \parallel_M = \parallel L^T R_u \parallel_2$ which makes it compatible with a standard implementation of the Levenberg-Marquardt algorithm.

As mentioned before, the use of a locally refined mesh is relevant for the simulation but it has to be coarse enough to measure accurately FE-DIC displacements. To avoid the delicate issue that arises because large uncertainties in the displacement field will obviously yield large uncertainties on the identified parameters, a multiscale FEMU approach is thus developed. Farfield measurements are used to define Dirichlet boundary conditions in the FE simulation. They are indeed mandatory to compute the reaction force $F_s(p)$. However, the FEMU takes advantage of nearfield FE-DIC measurements. Once the simulation mesh adjusted on the reference nearfield image following the multiscale registration technique, one measures better resolved nodal displacements $\mathbf{u}_m|_{near}$ in a restricted region where it is particularly sensitive to the sought parameters. In this multiscale approach of FEMU, only those nearfield measurements are compared to the simulated displacements. The residual vector turns to $R_u(p) = (u_s(p) - u_m|_{near})/ \parallel u_m|_{near} \parallel$.
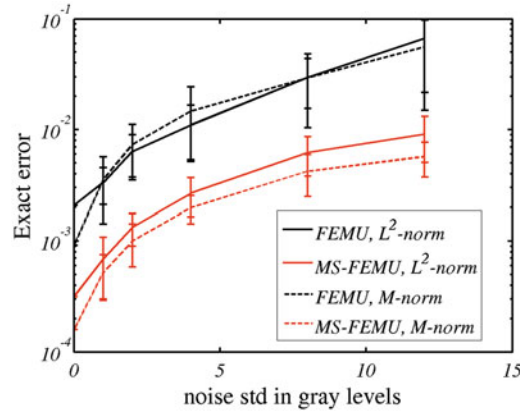
Fig. 30.5 Relative error $e(p)$ of the identified parameters as a function of the image noise for the standard single-camera identification technique (FEMU) and the proposed multiscale identification method (MS-FEMU) with the cost function expressed in the L^2-norm and M-norm (note that the errorbars are also in logscale)

30.2.4   A Priori Analysis of the FEMU Identification Robustness

For that purpose, at each scale, an image is generated by warping the real reference images (Fig. 30.7) with a displacement corresponding to the solution of the FE model with a known set of parameters denoted p_{ref}. Both identification procedures are then applied to these two image pairs. To compare the approaches, the relative error between the reference p_{ref} and the identified parameters $p^\star$ is quantified by $e(p^\star) = \|p_{ref} - p^\star\|/\|p_{ref}\|$. The standard FEMU (i.e. that uses in the cost function only the displacement field measured from the farfield images) and multiscale FEMU approaches give respectively $e(p^\star) = 2.08 \times 10^{-3}$ and $e(p^\star) = 3.15 \times 10^{-4}$. Thus the identification accuracy is improved by almost one order of magnitude on $e(p)$ with only a ratio of 5 between the image resolutions, in the proposed multiscale method. In addition, in order to further compare the robustness of the methods, the impact of the image noise is evaluated. A Gaussian noise with zero mean and a standard deviation ranging from 1 to 12 gray levels, is added to the farfield and the nearfield images. For each level of noise, both the identification methods are run for 20 random samples. The evolution of the standard and multiscale identification errors are plotted in Fig. 30.5 as a function of the image noise. It can be seen that the proposed multiscale procedure reduces also the sensitivity of the identification with respect to image noise roughly by one order of magnitude. Finally, the use of the M-norm helps to further improve the identification accuracy ($e(p^\star) = 1.58 \times 10^{-4}$ compared to 3. 15×10^{-4} previously) and the noise robustness of the identified process.

30.3   Application

30.3.1   Experimental Set-Up

The proposed multiscale methodology is now applied to a real experiment. A quasi-static tensile test is performed on an open hole glass/epoxy composite coupon. The base plate was manufactured by stacking four pre-preg plies made of 8-harness satin balanced woven fabric (8-HS) in the same draping direction. It was formed using a vacuum bag technique and cured in a polymerization oven. The 1.26 mm thick laminated plate was cut with a diamond wheel into a straight-sided 30 $\times$ 250 mm $[0]_4$ coupon (i.e. orientated at $0°$ with respect to the warp direction). Once the tabs glued, the gauge section is 150 mm long. Finally, a 10 mm hole is drilled in the centre of the specimen. The macroscopic behavior of the studied thin laminate is assumed to be orthotropic without in-plane bending-twisting coupling. The major material axis is aligned with the tensile direction. In the following, a 2D stress state is assumed. The test was carried out on an electromechanical tensile machine (Instron 5800). The loading was periodically interrupted after a load increment of approximately 0.5 kN up to 5 kN. In between the steps, the loading rate was around 0.25 mm/mn. At each load step, once the load stabilized, both the farfield and nearfield images were recorded. Two CCD cameras (AVT Dolphin F-145B) have been used to capture these images (definition: 1, 392 $\times$ 1, 040 pixels, 8-bit). Assuming that the specimen undergoes in-plane deformations, both the optical

Fig. 30.6 Experimental
setup. The two optical axes
of the CCD cameras are
perpendicular to the laminate
plane. A translation stage
is used to retract the nearfield
camera in order to capture
the farfield image

Fig. 30.7 Multiscale digital
images with corresponding
finite element mesh.
The *yellow lines* correspond
to the ROIs (Color online)
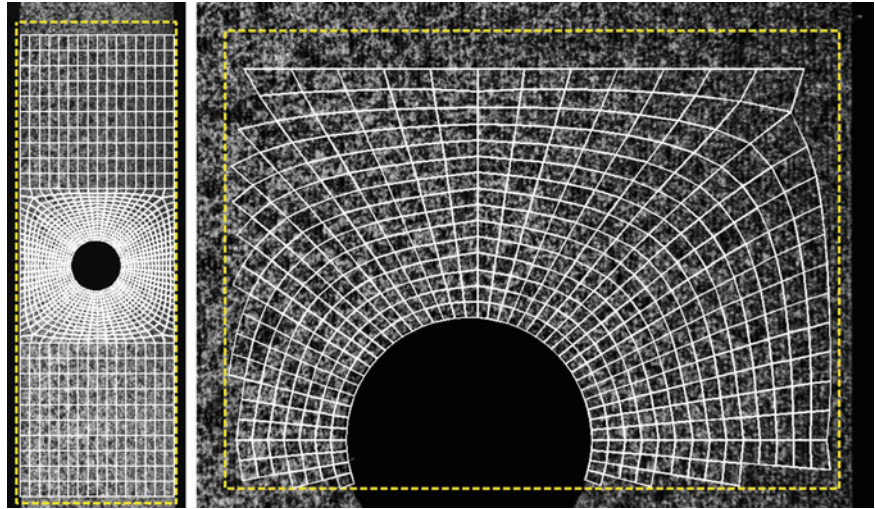
axes are set to be perpendicular to the laminate plane. The first camera takes pictures which cover approximately the whole gauge region, while the second concentrates over a smaller area located around the hole, provided by the sensitivity analysis. Here the ratio between the two optical resolutions is around 5. The nearfield camera, mounted on a translation stage, is then retracted to take a picture of the farfield region, see Fig. 30.6.

A black and white speckle is sprayed on top of the surface in order to provide a random texture suited for the DIC. As expected in such a situation, the global load/displacement response is linear elastic. A typical simulation mesh is used in the FE-DIC to measure the displacement field at both scales. The multiscale images and corresponding mesh positions are plotted in Fig. 30.7.

30.3.2 DIC Error Analysis

A performance analysis is performed on the real images, in order to assess the efficiency of the multiscale approach. A typical FE mesh is built for the simulation, see e.g. Fig. 30.7. It is irregular but structured and made of 4-noded bilinear elements whose size ranges gradually from 2.8 to 0.64 mm near the hole. Therefore the spatial resolution for the DIC displacement measurement varies from 33 to 145 (respectively 7–30) pixels in the nearfield (respectively farfield) image. In this case, the FE mesh is assumed to be optimized for the simulation. As a consequence, only the ultimate error is considered in this section. From the real reference images $\mathcal{I}_0^f$ and $\mathcal{I}_0^n$, two series of synthetic deformed images are generated by a subpixel shift in the Fourier space whose magnitude ranges between 0 and 1 pixel [8, 9, 17]. A FE-DIC measurement is then performed at both scales and the ultimate random error (standard uncertainty) and systematic error (bias) are computed from the discrepancy between measured and prescribed displacement fields, as described in [17] for each value of the shift. Results are plotted in Fig. 30.8. Note that for the farfield images, only measurement values inside the ROI corresponding to the nearfield images are considered in the error calculation.

First, with the FE-DIC approach, typical *bell*-shaped and *S*-shaped curves are also obtained for random and systematic errors respectively, in good agreement with the literature on subset-based DIC methods [17]. By using such a multiscale
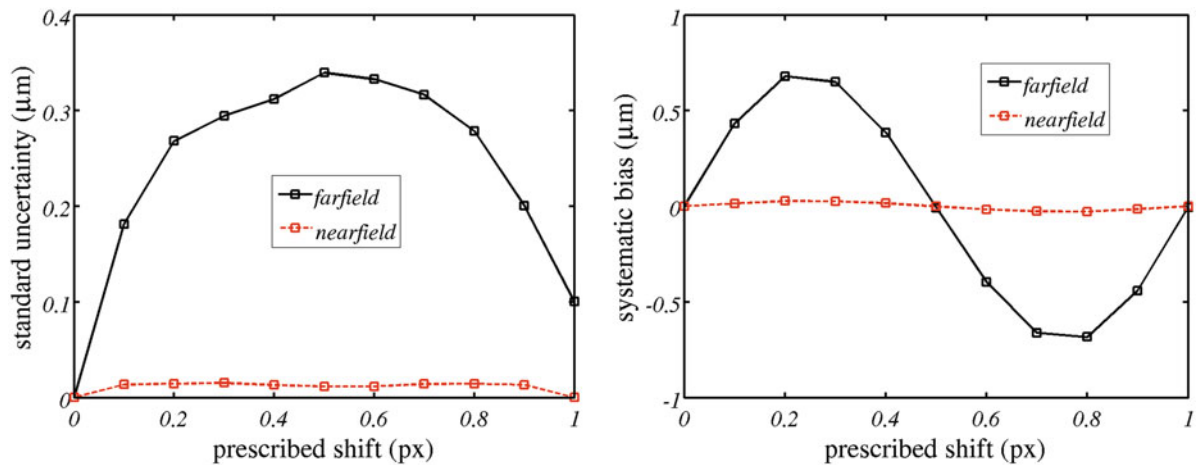
Fig. 30.8 Random (*left*) and systematic (*right*) errors (in μm) as a function of the prescribed shift in pixel with the unstructured mesh

approach, the gain is hence double, since it takes advantage of both (i) the improvement of the image resolution (number of pixel/mm) and (ii) the improvement of the DIC spatial resolution in pixel in the image (number of pixel/element width), for a given mesh. This corresponds respectively to the vertical (i) and the horizontal (ii) translation between nearfield (red) and farfield (black) curves in Fig. 30.4. Thanks to that, the ratio between nearfield and farfield uncertainties is more than one order of magnitude for an image resolution ratio of only 5.

30.3.3   Identification of Elastic Properties

The ten pairs of images recorded during the tensile test are used to measure both farfield and nearfield displacement fields. The FEMU procedure is initialized using the elastic parameters identified classically [4] (reference values): $E_l = 21.53$ GPa, $E_l = 20.59$ GPa, $v_{lt} = 0.15$ and $G_{lt} = 3.54$ GPa. Figure 30.9 shows the evolution of the parameters identified with both standard and multiscale FEMU methods, at the nine last load steps.

The identification results demonstrate that the longitudinal modulus can correctly be estimated from both FEMU analyses. At the first loading steps, the multiscale approach produces more realistic values for all the parameters. Moreover, except for the transverse modulus E_t, the values of the parameters are close to the reference ones. It is worth remembering that the sensitivity of the displacement field with respect to this parameter is lower than for the other parameters, and that the highest sensitivities are restricted to a very small area. It is therefore logical that both approaches fail to provide relevant results for this parameter. On the other hand, the evolution of the identified parameters as a function of the loading step is much more regular with MS-FEMU than with standard FEMU, in particular at the first steps, where the signal to noise ratio is bad. This is in accordance with the uncertainty analysis of Fig. 30.5. Paradoxically, the identification of the elastic properties of a composite laminate is quite a tricky problem. In order to stay in the elastic domain (including in the hole vicinity), the loading must be sufficiently low, which leads to small strain levels. The resulting signal to noise ratio makes standard FEMU fail to identify accurately elastic parameters. Conversely, high levels of loading, inevitably lead to material degradations (at least locally) which invalidate the elastic assumption. In this regard, the proposed multiscale identification technique is a good alternative since is proved to reduce significantly the noise sensitivity. A complete analysis of the results reveals that the parameters seem to evolve significantly at a very early stage of the loading. The shear modulus G_{lt} continuously decreases all along the tensile experiment. This could be related to the development of damage in the plate. The model considered herein (orthotropic linear elastic) is unable to model this phenomenon which is in addition most prominent in the nearfield region. This may explain the reason why the parameters identified from the mono and the multiscale approaches can not be compared for the higher loading steps. Indeed a precise analysis of the difference between measured displacements $\mathbf{u}_m$ and the displacements $u_s(p^\star)$ simulated with the optimized parameters $p^\star$ has shown that some discrepancies are present, hardly visible at 2 kN, but significant at 5 kN. This simply confirms that the chosen model is (obviously) not able to describe the observed behavior throughout the tensile test, particularly in the vicinity of the hole where damage is known to localize.
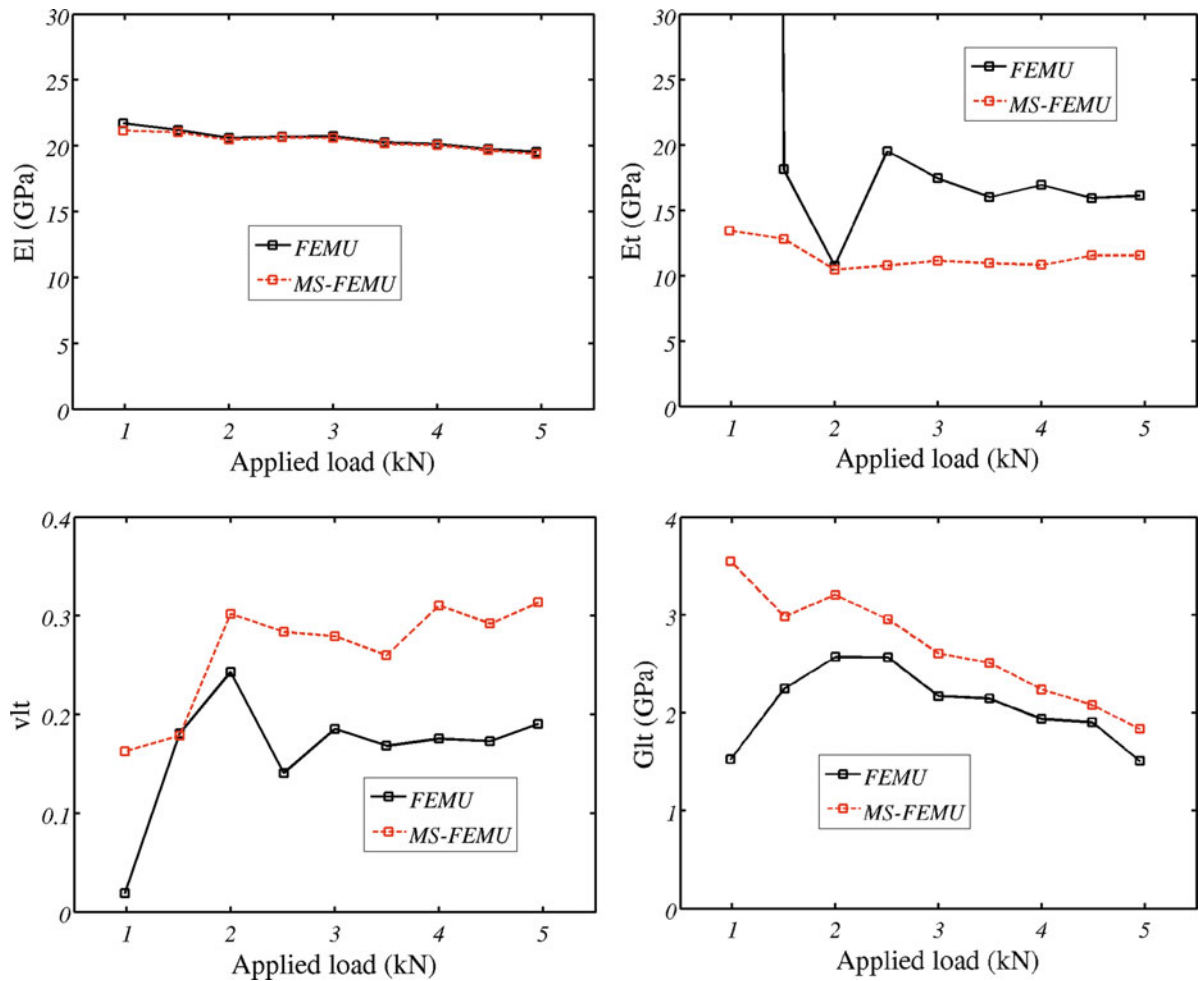
Fig. 30.9 Identified parameters from different loading magnitude. From top left to bottom right: E_t; E_l; v_{lt}; G_{lt}. Standard FEMU (*solid line*) and MS-FEMU (*dashed line*)

30.4   Conclusions

Connection between simulation and full-field measurement was originally a critical task. With the advent of FE-DIC methods [8], it is now possible to bridge efficiently both of them with a common language: a FE mesh [22]. However, choosing an appropriate mesh and/or spatial resolution may be quite tricky because of the spatial resolution/uncertainty compromise [3, 17]. In addition, this choice is also constrained by hardware limitations. Moreover, in the context of identification, the spatial resolution is often limited because the field of view generally needs to include the boundaries of the specimen [12]. At the same time, a high spatial resolution is required in the regions (sometimes very small) where the displacement is sensitive to the parameter to identify. All these remarks led us to devise both a new DIC methodology and an associated FEMU technique able to take the best of images taken at two different resolutions. Thus, in this paper, (a) a dedicated DIC method was proposed for the automatic and accurate registration of the farfield image in the nearfield image and (b) an hybrid multiscale cost function was used in the FEMU technique. The effectiveness of this multiscale approach was assessed with synthetic multi-resolution speckle pattern images. The results show that the proposed multiscale method significantly improves both measured displacements and identified parameters. It is shown that even with a ratio of 5 between the image resolutions, the measurement and identification uncertainties can be reduced by one order of magnitude which is one of the most interesting output of the study. Not only the uncertainties are reduced, but it is shown that the proposed method is also more robust with respect to image noise by approximately one order of magnitude. Finally, it is also shown that the selected behavior model may be, at least locally, not adapted.

References

1. Grédiac M, Hild F (eds) (2012) Full-field measurements and identification in solid mechanics. Wiley, New York. ISBN: 9781848212947
2. Sutton MA, Orteu J-J, Schreier HW (2009) Image correlation for shape, motion and deformation measurements – basic concepts, theory and applications. Springer, New York. ISBN: 978-0-387-78746-6
3. Bornert M, Brémand F, Doumalin P, Dupré JC, Fazzini M, Grédiac M, Hild F, Mistou S, Molimard J, Orteu JJ, Robert L, Surrel Y, Vacher P, Wattrisse B (2009) Assessment of digital image correlation measurement errors: methodology and results. Exp Mech 49(3):353–370
4. ASTM D3039-00 (2000) Standard test method for tensile properties of polymer matrix composite materials, vol. 15
5. Molimard J, Riche R, Vautrin A, Lee J (2005) Identification of the four orthotropic plate stiffnesses using a single open hole tensile test. Exp Mech 45(5):404–411
6. Sztefek P, Olsson R (2008) Tensile stiffness distribution in impacted composite laminates determined by an inverse method. Compos Part A Appl Sci Manuf 39(8):1282–1293
7. Avril S, Bonnet M, Bretelle AS, Grédiac M, Hild F, Ienny P, Latourte F, Lemosse D, Pagano S, Pagnacco E, Pierron F (2008) Overview of identification methods of mechanical parameters based on full-field measurements. Exp Mech 48(4):381–402
8. Besnard G, Hild F, Roux S (2006) "Finite-element" displacement fields analysis from digital images: Application to Portevin-Le Châtelier bands. Exp Mech 46(6):789–803
9. Wang YQ, Sutton MA, Bruck HA, Schreier HW (2009) Quantitative error assessment in pattern matching: effects of intensity pattern noise, interpolation, strain and image contrast on motion measurements. Strain 45(2):160–178
10. Kavanagh KT, Clough RW (1971) Finite element applications in the characterization of elastic solids. Int J Solids Struct 7:11–23
11. Lecompte D, Smits A, Sol H, Vantomme J, Hemelrijck DV (2007) Mixed numerical-experimental technique for orthotropic parameter identification using biaxial tensile tests on cruciform specimens. Int J Solids Struct 44(5):1643–1656
12. Réthoré J, Muhibullah, Elguedj T, Coret M, Chaudet P, Combescure A, (2013) Robust identification of elasto-plastic constitutive law parameters from digital images using 3D kinematics. Int J Solids Struct 50:73–85
13. Lucas B, Kanade T (1981) An iterative image registration technique with an application to stereo vision. In: Proceedings 1981 DARPA imaging understanding workshop, 121–130
14. Sutton M, Wolters W, Peters W, Ranson W, McNeill S (1983) Determination of displacements using an improved digital correlation method. Image Vis Comput 1(3):133–139
15. Hild F, Roux S (2012) Comparison of local and global approaches to digital image correlation. Exp Mech 52(9):1503–1519
16. Sun Y, Pang J, Wong CK, Su F (2005) Finite element formulation for a digital image correlation method. Appl Optic 44(34):7357–7363
17. Amiot F, Bornert M, Doumalin P, Dupré JC, Fazzini M, Orteu JJ, Poilâne C, Robert L, Rotinat R, Toussaint E, Wattrisse B, Wienin JS (2013) Assessment of Digital Image Correlation measurement accuracy in the ultimate error regime: main results of a collaborative benchmark. Strain 49(6):483–496
18. Leclerc H, Périé JN, Roux S, Hild F (2011) Voxel-scale digital volume correlation. Exp Mech 51(4):479–490
19. Passieux J-C, Bugarin F, Périé J-N, Robert L, David C (2014) Multiscale displacement field measurement using digital image correlation: application to the identification of elastic properties. Exp Mech. doi: 10.1007/s11340-014-9872-4
20. Orteu JJ, Garcia D, Robert L, Bugarin F (2006) A speckle-texture image generator. In: Speckle'06 International Conference, SPIE Proceedings, vol. 6341
21. Réthoré J, Elguedj T, Simon P, Coret M (2010) On the use of NURBS functions for displacement derivatives measurement by digital image correlation. Exp Mech 50(7):1099–1116
22. Leclerc H, Périé JN, Roux S, Hild F (2009) Integrated digital image correlation for the identification of mechanical properties. Lect Note Comput Sci 5496:161–171

Chapter 31
Uncertainties of Digital Image Correlation Near Strain Localizations

Mark A. Iadicola and Adam A. Creuziger

Abstract Estimates of the strain errors resulting from digital image correlation (DIC) measurements are desired for many uses. One application is the measurement of the strain localization near failure in forming limit testing of sheet metals. This work measures and statistically characterizes the displacement measurement uncertainties for a typical DIC system. These uncertainties are found to have nearly Normal probability distributions and were used as inputs into a Monte Carlo analysis to determine the resulting strain uncertainty. A limited parameter study was made using the Monte Carlo analysis. The results demonstrate that the strain measurement uncertainty is quantifiable, and reducing the virtual gauge length (over which the strain is determined) tends to increase the strain measurement uncertainty. Based on the results, curves relating the strain uncertainty to the DIC analysis parameters can be developed. These curves suggest a balance must be chosen between the optimum processing parameters to minimize the strain error or the parameters to minimize the virtual gauge length. For some engineering applications (e.g. forming limit testing), a smaller virtual gauge length might be preferred even if it results in a higher strain uncertainty, as long as that uncertainty is quantifiable.

Keywords Digital image correlation • DIC • Measurement uncertainty • Monte Carlo method • Strain localization

31.1 Introduction

There has been an effort to apply digital image correlation (DIC) to measure the evolving strain fields during standard forming limit testing [1]. These limit strains result in an inhomogeneous strain field just prior to failure of the sheet metal specimen. Various analysis methods of the DIC strain field data have been considered including: spatial [1], temporal [2], or combined [3]. All of the methods use the variation at or near the strain localization prior to failure. Use of these measurements in documentary standards requires determination of the uncertainty of the measurements. ASTM E2208-02 "Standard Guide for Evaluating Non-Contacting Optical Strain Measurement Systems" [4] describes some requirements for assessment of the uncertainties in optical strain measurement systems by comparing the results on the same length scales as other established calibrated test measurement methods (e.g. calibrated extensometers), but the use of these methods at smaller length scales requires either more advanced calibration or a numerical assessment. The work presented here builds on prior work by the author [5] and attempts to assess the expected measurement uncertainty of strain at these smaller length scales of interest. In addition to the specific assumptions listed in the body of this paper, certain general simplifications are made to facilitate the initial analysis including: the pattern is assumed to not breakdown as strain increases; that there is no loss of correlation; that the uncertainty of the measurement is not affected by the location within the image due to higher order lens effects [6]; and the pattern roughness is beyond the DIC measurement capability (partly verified in [5]).

M.A. Iadicola (✉) • A.A. Creuziger
National Institute of Standards and Technology, 100 Bureau Drive, STOP 8553, Gaithersburg, MD 20899, USA
e-mail: mark.iadicola@nist.gov

31.2   DIC Setup

A three-dimensional DIC (3D-DIC) system was used for all of the measurements presented here. Two 8-bit grey-scale 2MP-CCD cameras with compact 35 mm lenses were mounted at a stereo angle ≈33° with a stand-off distance of ≈380 mm. The resulting average magnification was 23.3 pixels/mm with a field of view approximately 73 mm by 52 mm (Note: although many DIC analysis parameters and results are naturally in units of pixels, it is often desirable to describe the results in physical units of the specimen. In this work when the physical values are presented they are calculated from the pixel values using the average magnification given here). To reduce motion between the specimen and cameras, the specimen was bolted to a translation stage (X-direction motion), which was bolted to an optical bread-board attached by extrusions to the cameras. Unfortunately, this "C" shaped structure still allows a torsional vibration that can result in motion of the cameras in the Y-direction, but was similar hardware setups used for forming limit testing. Unless stated specifically, for all the results described here, a fan was used to reduce the effects of air currents (similar to [6]). Although the specimen was nominally aligned to the camera pixel arrays, some rotation of the resulting measurements was necessary to align the X-axis of the DIC measurements to the X-direction of the DIC patterns (see next section) and translation stage. Variable intensity fiber optic lighting was used to achieve good brightness in the resulting digital images taken with 15–20 ms exposures at an f-stop = 8. Commercial DIC analysis software was used for position measurement, but not for the strain determination. DIC performs correlation on small square subsets of each image pair. In this work, the size of the subset was one of the parameters varied for study. The calibration of the system used a reference dot grid and 161 image pairs, to reduce calibration artifacts (similar to [6]). An optimized 6-tap correlation interpolation function [7, pp. 107–109] was used for all of the results presented here. Although the choice of correlation function has been shown to affect sub-pixel accuracy [7, 8], the uncertainties are typically <0.03 pixels compared to the uncertainties of interest here, which are between 0.012 pixels to 0.256 pixels. The confidence margin for the correlation (correlation tolerance) was constant for all except the smallest correlation subset window size. For the smallest subset, the confidence margin was twice as large as the others to achieve correlation (which resulted in increased displacement uncertainties for this subset size). The specimen was translated through four rigid-body motions in the X-direction, nominally X = (0, 1, 2, 3) mm. At each position the displacement was held, and 50 image pairs were taken for subsequent DIC analysis.

31.3   Specimen and Patterning

The specimen used in the measurements presented in this work was 40 mm by 120 mm with six grooves of various depths and widths embossed at approximately 10 mm spacing centered along the 120 mm length. The axis of each groove is parallel to the 40 mm length. The long axis of the specimen was oriented in the X-direction, and the Y-direction was oriented along the surface of the specimen in the short direction. The grooves after patterning application have depths between 50 and 150 μm and widths between 1.5 mm to 4 mm. The size and depth of the grooves were selected to simulate those measured by a profilometer on forming limit specimens after localized plastic deformation just prior to failure (see [5]). The forming limit test methods [1, 9] are examples where the application of 3D-DIC is very desirable.

Five black and white patterns labeled A through E (in Fig. 31.1a) were applied to the specimen using a sheet of photographic quality paper and spray adhesive. Figure 31.1b shows close-up images of each pattern. The close-up images are 41 pixel square and match the size of the largest subset used. Patterns A, B, and C are printed patterns designed to have minimal directional bias and nearly equal amounts of black and white. The three (A, B, and C) use the same vector graphic (at 100, 50, and 25 % scales, respectively). Pattern C had the smallest features, approximately 3–4 pixels in extent (which is the smallest feature size typically used in DIC [10]). Unfortunately, the finer pattern results in a less bimodal intensity distributions and more intermediate gray values (see Fig. 31.1c). Pattern D was created using black spray paint after the paper was adhered to the specimen, and resulted in a less uniform, but a more easily created pattern. Pattern E was applied to the paper using an ultra-fine tip permanent marker.

The overall shape of the grooves after application of the patterns was measured using a profilometer and compared with DIC measurements of the same. These results showed that the DIC subsets tended to average the shape of the groove. For shallow grooves, the DIC measured depth matched the profilometer data, however deeper (sharper) grooves resulted in DIC measured depths less than the profilometer data [5].
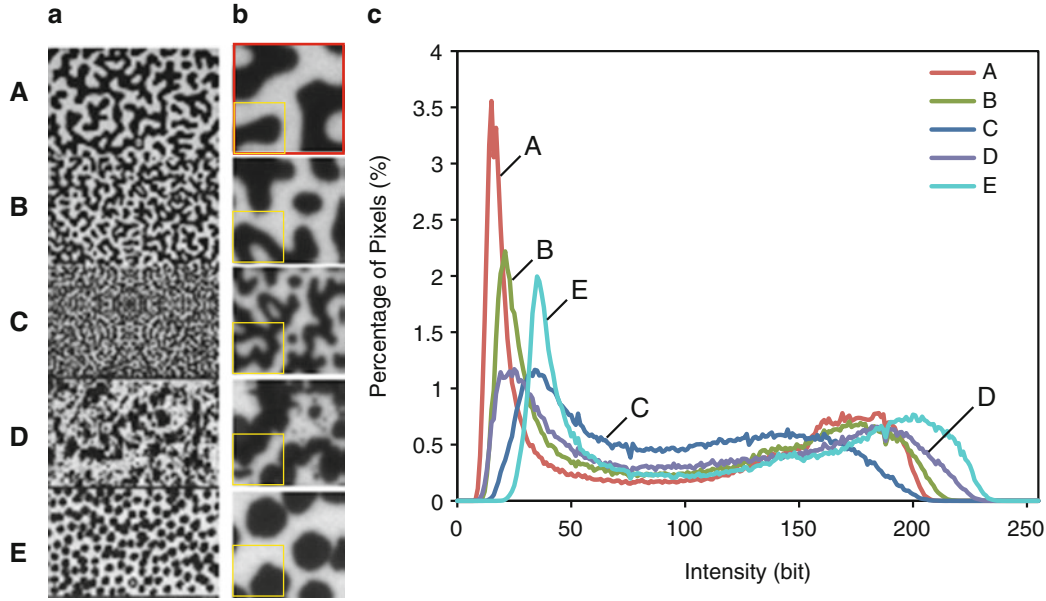
Fig. 31.1 A section of a DIC image ($X \approx$ *horizontal* and $Y \approx$ *vertical* in the image), (**a**) patterns labeled A through E, with 41 pixel *square* close-up views (**b**) smaller (*yellow square* is 19 pixels on each side). (**c**) Intensity histogram of the patterns for a typical image [5]

31.4 Displacement Uncertainties and Distributions

The values and statistical distributions of the measured 3D displacements (u, v, and w in the X, Y, and Z directions, respectively) were determined from image pairs taken after each rigid body displacement of the specimen. These correlations were performed with a two-pixel spacing of the subset center points. This constant value was intended to facilitate comparison of the displacement data only, and was NOT used for strain determination. The mean values ($\overline{u}$) of these displacements (u_i) for all of the correlated points ($i = 1$ to N) in each image pair should not affect the uncertainty of the calculated strains since their cumulative effect will be zero. The total number of points (N) for one pattern in each image pair is typically >62,000. A critical factor affecting the strain uncertainty is the standard deviation (σ_u) about the mean ($\overline{u}$) of the displacements defined for the X-direction as,

$$\overline{u} = \frac{\sum\limits_{i=1}^{N} u_i}{N} \tag{31.1}$$

and

$$\sigma_u = \sqrt{\frac{\sum\limits_{i=1}^{N} (u_i - \overline{u})}{N - 1}} \tag{31.2}$$

and similarly for the other directions (Y and Z displacements u and v, respectively). The distributions of the measured displacements are further analyzed to verify their degree of symmetry (skewness, γ_u) and sharpness of their peak (excess kurtosis, β_u) defined as,

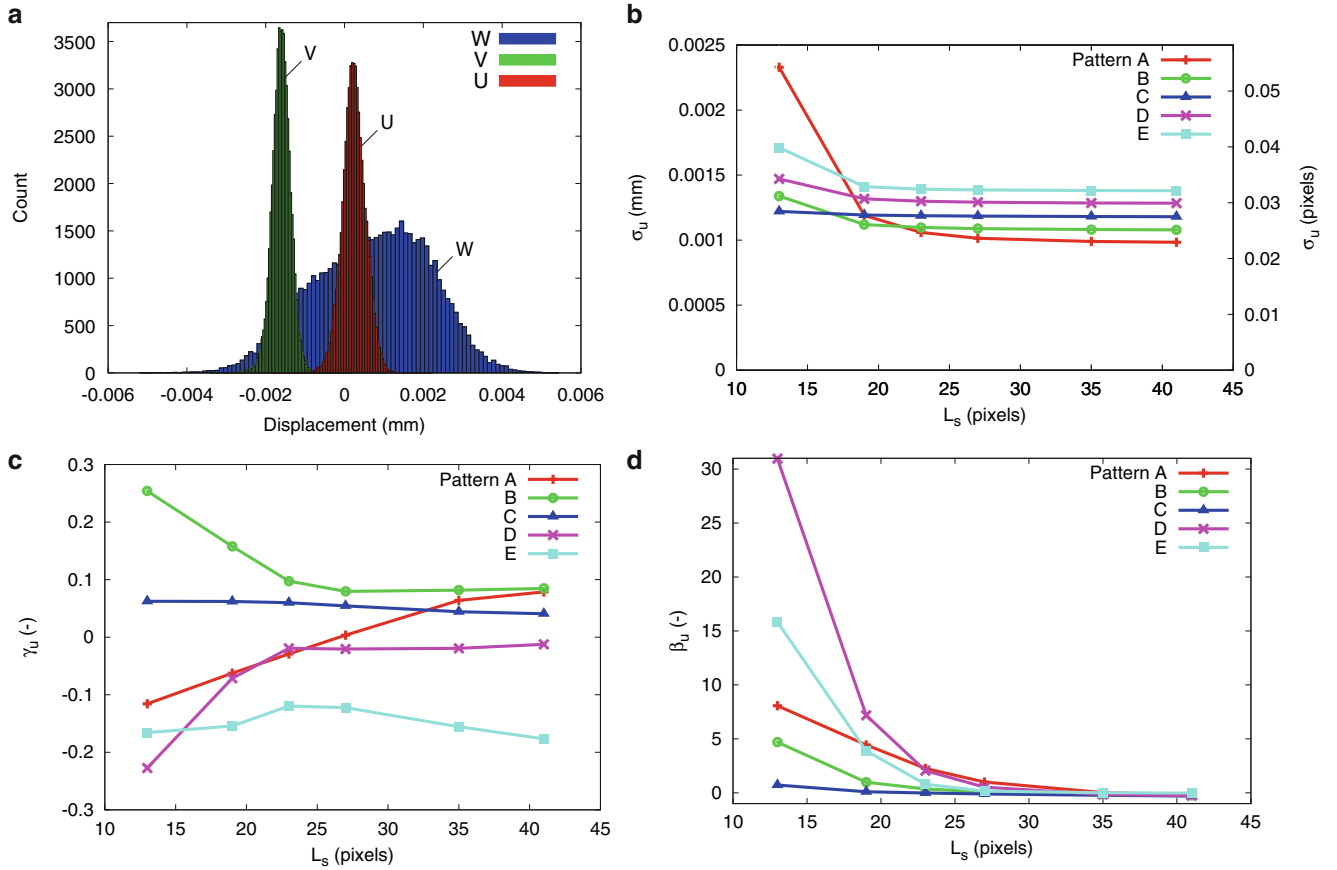$$\gamma_u = \frac{1}{N\sigma_u^3} \sum\limits_{i=1}^{N} (u_i - \overline{u})^3 \tag{31.3}$$

Fig. 31.2 Example distributions of displacements u, v, and w (**a**) for a case with no applied displacement and correlated using a subset of 23 pixels square on pattern E. Standard deviations (**b**), skewness (**c**), and excess kurtosis (**d**) distributions of u displacement (at zero applied displacement with a fan) for patterns A through E at various subset lengths (L_s)

and

$$\beta_u = \frac{1}{N\sigma_u^4} \sum_{i=1}^{N} (u_i - \overline{u})^4 - 3 \tag{31.4}$$

for the displacement in the X-direction, and similarly for the other directions (Y and Z displacements v and w, respectively). Figure 31.2a shows the histogram of the displacements (u, v, and w) for one typical image pair, and Fig. 31.2b–d show the results for the u displacements for standard deviation, skewness, and excess kurtosis, respectively. Here and in the remainder of this work the average value over all 50 image pairs is reported for each σ, γ, and β. An alternative method would be to calculate these for all of the data points in all of the image pairs simultaneously, but the results do not differ significantly. Figure 31.2b–d also show the variation due to pattern (A through E) and the selected correlation subset sizes (L_s in pixels). The values of the distribution shapes approach the values for a Normal probability distribution ($\gamma = 0$ and $\beta = 0$) as the subsets become larger or the patterns become finer. Similar results were seen for all of the displacements applied to the sample, therefore a Normal probability distribution will be assumed for the determination of strain uncertainty (discussed in the next section).

Since a critical parameter that will affect the strain uncertainty is the standard deviation of the displacements, we now look a bit more closely at the variation with the displacement direction. Figure 31.3a–c plot the variation of the standard deviation of the displacements in the X, Y, and Z directions, respectively, for image pairs with no applied displacement and no fan moving the air currents. The results are plotted against the subset size for the five patterns used. The trends are quite similar to Fig. 31.2b showing strong dependence on the pattern and subset, but the overall scale of the uncertainty varies between u, v, and w, typically with $\sigma_v > \sigma_u$ and σ_w values being quite a bit larger than either σ_u or σ_v (similar trends were seen in [6]). The trends and amplitude of standard deviations for the translated cases were all similar to Fig. 31.3.

One unexpected variation was between the sets of images at zero translation with and without the fan. Although the case without the fan showed larger mean amplitude (not shown), the standard deviations were smaller and more clustered
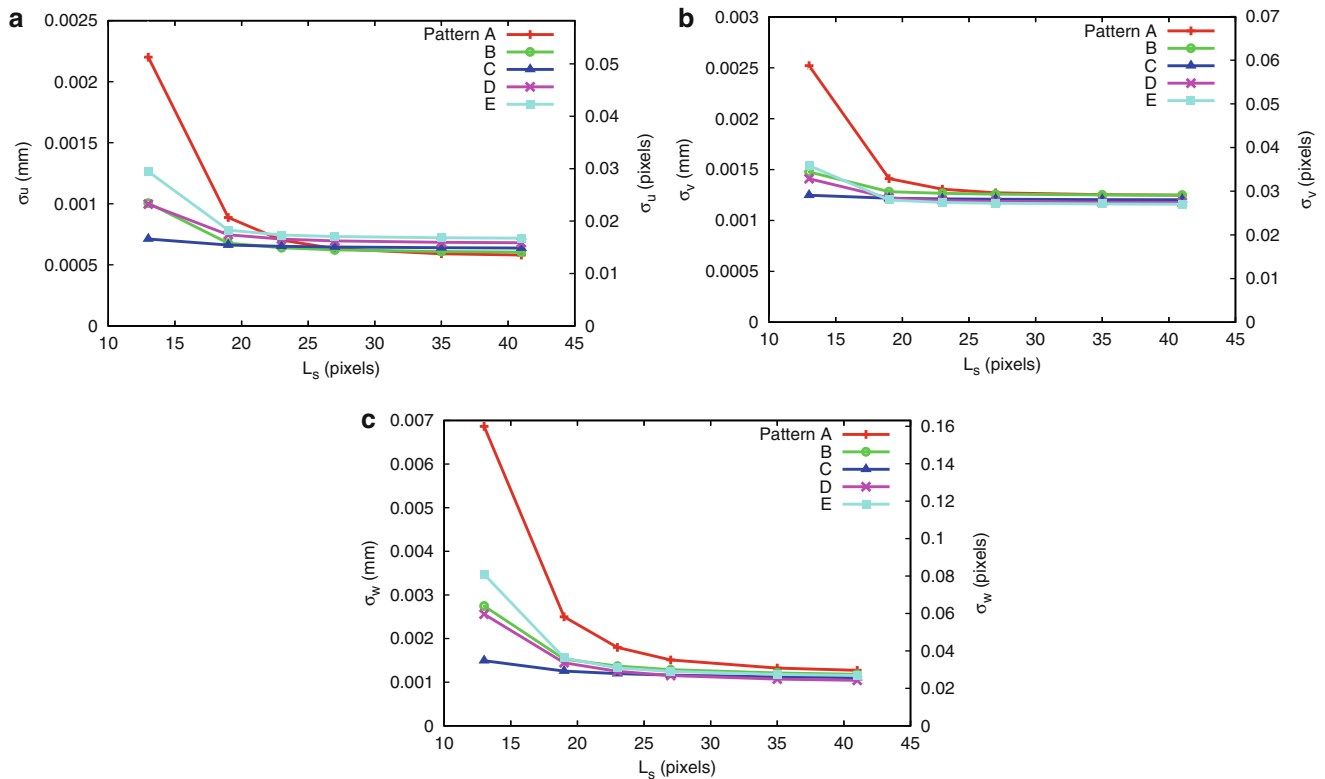
Fig. 31.3 Standard deviations of (**a**) u, (**b**) v, and (**c**) w displacements (at zero displacement without a fan) for patterns A through E at various subset lengths (L_s in pixels)

(σ_u in Fig. 31.3a) than the case with the fan (σ_u in Fig. 31.2b). Looking at the results for the 50 image pairs for each of these cases showed that five images in the case with the fan were outliers by $> 3\sigma_u$. The reason for this result is not clear, but could be due to the torsional vibration that is not controlled by the camera mounting structure. Further investigation is needed. Typically one would expect lower standard deviation values with the fan than without (similar to [6]), but after removing the outlying images the resulting standard deviations vary with subset size and pattern approximately the same magnitude with and without the fan. Therefore, pattern and subset size seem to be stronger influences on the uncertainties measured than other factors.

Similar analyses of the displacement distributions and standard deviations were done splitting the data sets into measurements inside and outside of the specimen grooves to see if the 3D features affected the results. No major difference was observed, and the variation was not systematic, i.e. the standard deviation values inside the grooves were as often lower as they were higher than outside of the grooves. Displacement standard deviations both inside the grooves and outside the grooves had the same dominant trends with pattern variation and subset size variation as seen above.

31.5 Strain Uncertainty Determination

A parametric study of the strain uncertainty was performed using the Monte Carlo method. The parameters that were varied include: (1) the level of applied displacements (nominally uniaxial strain equivalents in the X-direction); (2) the spacing between displacement measurements (referred to here as node spacing L_n) for strain calculation; and (3) the values of displacement uncertainties (based on the previous section results).

Fig. 31.4 Displacement
uncertainties σ_u, σ_v, and
σ_w used for the Monte Carlo
analysis of strain uncertainty
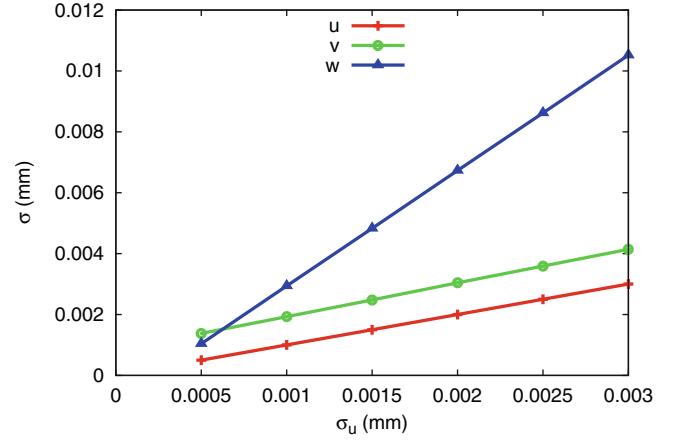
The strain measure used was the in-plane Lagrangian strain,

$$\varepsilon_{xx} = \frac{\partial u}{\partial x} + \frac{1}{2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2\right] \tag{31.5}$$

$$\varepsilon_{yy} = \frac{\partial v}{\partial y} + \frac{1}{2}\left[\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2\right] \tag{31.6}$$

$$\varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) + \left[\left(\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial x}\frac{\partial w}{\partial y}\right)^2\right] \tag{31.7}$$

for one triangle between three displacement nodes (which would be the smallest area for a strain value determination). The initial (reference) positions for the nodes for the Monte Carlo analysis were oriented to the X and Y directions, with the Z values constant for the three points. These nodes were similar to a position on the surface of a flat uniaxial specimen oriented in the X-direction at a location slightly away from the longitudinal centerline. Therefore, no additional rotations were needed in the formulation. Uniaxial strains,

$$\varepsilon_{xx} = \text{constant}, \varepsilon_{yy} = -\frac{1}{2}\varepsilon_{xx}, \text{and } \varepsilon_{xy} = 0, \tag{31.8}$$

were applied as numerically displacements (u, v, and w) to the nodes. This was done at three strain levels (0.002, 0.10, and 0.20) to see if the magnitude of applied strain affects the calculated strain uncertainty. In DIC analysis, the displacement node spacing (L_n) is typically set to $L_n \geq 0.5L_s$ [6]. Unfortunately, larger L_n values increase the physical extent of the measurement (i.e. the effective gauge length), and result in more spatial smoothing. In cases where the spatial variation is of interest, this spatial smoothing is not desirable. Therefore, for the parametric study the reference spacing (L_n) of the node points was varied from 1 pixel to 25 pixels, which frequently resulted in $L_n < 0.5L_s$.

The displacement uncertainty ranges used in the Monte Carlo analysis were based on the results of the previous section. The trends between the amplitudes of σ_u, σ_v, and σ_w seen in Fig. 31.3 were similar for all of the rigid body displacements. These trends were abstracted into simple linear relations (Fig. 31.4) between σ_u and the other two displacement uncertainties (σ_v and σ_w) and used as the input into the Monte Carlo parameter study. For example in one run of the Monte Carlo analysis, if an input value of $\sigma_u = 0.0020$ mm was used then $\sigma_v = 0.0030$ mm and $\sigma_w = 0.0067$ mm.

The Monte Carlo analysis was performed in the computer language R using the National Institute of Standards and Technology Uncertainty Machine [11]. Based on initial numerical testing, 300 000 realizations for each parameter combination studied resulted in stable results for determined standard deviations (S) in the calculated strains. Note the measured displacement standard deviations from the section above are denoted by σ, where as the calculated deviations in strain from the Monte Carlo analysis are denoted as S, for clarity. The resulting stable strain value distributions appeared to be approximately Normal probability distributions, with the exception of two cases with the smallest node spacing and largest displacement uncertainties. Although the mean values of the resulting strain distributions were not equal to the

numerically applied strains they were always within less than one standard deviation of the resulting strain; therefore the first standard deviations of strain are shown here and are referred to as the calculated strain uncertainties (S). For brevity, only the resulting strain uncertainties for the ε_{xx} strains are reviewed here and are denoted as ($S_{\varepsilon xx}$).

31.6   Results

Figure 31.5a shows the predicted uncertainty of ε_{xx} with a fixed node spacing of 10 pixels (approximately 0.429 mm) at three uniaxial strain levels (0.002, 0.10, and 0.20) for a range of displacement uncertainties. For each σ_u value a fixed value of σ_v and σ_w were used (see Fig. 31.4) in the calculations. The resulting strain uncertainty values ($S_{\varepsilon xx}$) are not as sensitive to the applied uniaxial strains as they are to the applied displacement uncertainty values. This is seen in the small spread in strain uncertainty between the three strain levels shown in Fig. 31.5a. The relationship between the displacement uncertainties and strain uncertainties are almost linear with zero intercept values (as shown by the trend lines in Fig. 31.5a).

Figure 31.5b shows the variation in strain uncertainty with node spacing (1 pixel to 25 pixels, or approximately 0.043–1.073 mm based on the average magnification) at a single applied uniaxial strain of 0.10 in the X-direction. Each set of data points was for a selected value of σ_u, and its associated σ_v and σ_w (see Fig. 31.4). The general variation of strain uncertainty with node spacing is intuitive. Smaller spacing between the node points results in higher strain uncertainties where as larger spacing results in lower uncertainties. This is understandable since a reduction in the gauge length of the strain measurement results in a smaller denominator in the calculation of strain, and thus accentuates the displacement uncertainty effect. The relationship seems to follow a power law trend as shown by the black lines in Fig. 31.5b.

Figure 31.5 shows the strain uncertainty variation with displacement uncertainty, but does not directly relate the strain uncertainty to the subset size or DIC pattern used. This suggested that the results are decoupled and would apply for any displacements with Normally distributed standard deviations in this range of values, and could be applied to other DIC setups if characterized similar to the Displacement Uncertainties and Distributions section. We now demonstrate how the results can be combined for more practical use. Here the strain uncertainty for one specific pattern and rigid body displacement is considered to demonstrate how the previous results can define the relationship directly between the strain uncertainty and the DIC parameters (i.e. selection of pattern, subset size, and node spacing for strain calculation). The measured displacement uncertainty values used are for the case of no displacement and no fan (Fig. 31.3a) for the painted pattern D. This pattern was selected since it is a very common patterning method. For this demonstration, the results of Fig. 31.5b were used with varied node spacing (1 pixel to 25 pixels in 2 pixel increments) and a fixed applied uniaxial strain of 0.10 (which was shown in Fig. 31.5a to have only a small influence on the resulting strain uncertainty). The results of the demonstration calculation are shown in Fig. 31.6, which plots the resulting strain uncertainty variation versus the virtual gauge length (L_v). Here the virtual gauge length is the sum of one subset length and one node spacing (i.e. $L_v = L_s + L_n$), which is the total length used to determine a single strain value. The natural unit for L_v is pixels, but to better convey the
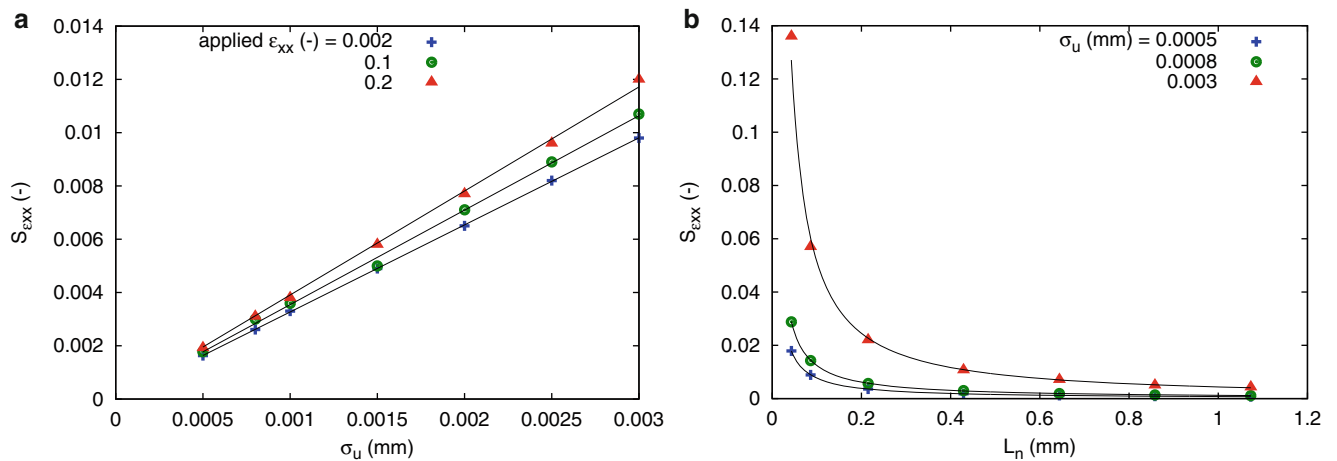

Fig. 31.5 Calculated strain uncertainty ($S_{\varepsilon xx}$) variation: (**a**) with u displacement uncertainty (σ_u in mm) for three applied uniaxial strain levels (see legend) and $L_n = 10$ pixels ≈ 0.429 mm, and (**b**) with varying node spacing (L_n in mm) at an applied 0.10 uniaxial strain for three σ_u values (see legend)

Fig. 31.6 Calculated strain uncertainty ($S_{\varepsilon xx}$) variation with virtual gauge length (L_v in mm) for six subset sizes (L_s in pixels, legend) using the displacement uncertainty results of pattern D (Fig. 31.3) and an applied 0.10 strain (Fig. 31.5b)
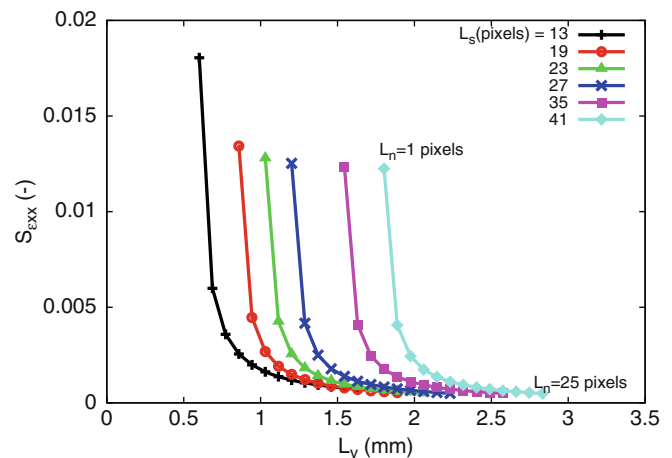
length that is being averaged, L_v is shown in Fig. 31.6 in physical units of mm based on the average magnification of the system (recall 23.3 pixels/mm). Each curve shown resulted from varying the node spacing for one subset size (shown in the legend). The lowest point for a given virtual gauge length describes the best combination of subset size and node spacing to reduce the strain uncertainty for that virtual gauge length. Note that the virtual gauge length will spatially average the results similar to the shape averaging seen in the measurement of the grooves [5] due to subset size, but over an even larger gauge length. If more neighboring triangles were used to determine the strain values the error would be further reduced by one over the square-root of the number of triangles, based on a typical calculation of uncertainty propagation [12, pp. 49–50]. However this will result in even larger virtual gauge lengths and more spatial averaging.

The results in Fig. 31.6 show that a balance must be chosen between the optimum processing parameters to minimize the strain uncertainty or the parameters to minimize the virtual gauge length. For some engineering applications (e.g. forming limit testing), a smaller virtual gauge length might be preferred even if it results in a higher strain uncertainty, as long as that uncertainty is quantifiable. One might imagine the ± 0.006 strain uncertainty (measured over a short length, e.g. 0.64 mm in Fig. 31.6) on a strain of 0.10 (near a strain gradient) as acceptable for the forming limit example mentioned in the Introduction. The alternative situation would be a case were spatial strain variation might be more gradual. In that case, a larger virtual gauge length ≈ 2 mm could be acceptable, and based on the demonstration case (Fig. 31.6) the minimum strain uncertainty ≈ 0.0008 could be achieved using $L_s = 27$ pixels and $L_n = 19$ pixels, resulting in $L_v = 46$ pixels or 1.97 mm.

31.7 Conclusions

An estimation of the strain uncertainty resulting from digital image correlation measurements of local strains has been presented. Measured displacement uncertainties were characterized as Normal probability distributions and used as input into a Monte Carlo analysis. The results of the analysis are estimates of the strain uncertainty for a wide range of DIC analysis parameters. The strain uncertainty results are limited to the effects of displacement uncertainty and displacement node spacing. The effects of spatial averaging on 3D shape measurement due to correlation subset size were discussed briefly in previous work [5] and were not considered here. Also not considered here is the reduction of correlation and increased noise due to pattern breakdown with large strains. Although the result is therefore only a partial solution, the relationships between the strain uncertainty and the input displacement uncertainties and node spacing were determined. In addition, this result was used to demonstrate how a relationship can be determined between the uncertainty in local strain measurement and the analysis parameters for the given: hardware, equipment setup, five test patterns, and resulting virtual gauge lengths. These results suggest that for a given (large) acceptable strain uncertainty a minimum virtual gauge length may be determined, thus minimizing any smoothing of the spatial variation in strain that is being measured. Conversely, for a given acceptable amount of spatial averaging (i.e. virtual gauge length) the parameters that result in the smallest strain uncertainty may be determined.

References

1. ISO (2008) Metallic materials-sheet and strip-determination of forming-limit curves-part 2: determination of forming-limit curves in the laboratory. ISO12004-2:2008(E)
2. Sriram S, Huang G, Yan B, Geoffroy J-L (2009) Comparison of forming limit curves for advanced high strength steels using different techniques. In: Proceedings of SAE international, 2009-01-1173 SAE International
3. Wang K et al (2014) Measuring forming limit strains with digital image correlation analysis. J Mater Process Technol 214:1120–1130
4. ASTM (2011) Standard guide for evaluating non-contacting optical strain measurement systems. E2208-02, ASTM International
5. Iadicola MA (2014) Augmented use of standard mechanical testing measurements for sheet metal forming: digital image correlation for localized necking. In: Yoon JW, Stoughton TB, Rolfe B, Beynon JH, Hodgson P (eds) NUMISHEET 2014 part b general papers, vol 1567, AIP proceedings. American Institute of Physics, Melville, NY, pp 614–619
6. Ke XD et al (2011) Error assessment in stereo-based deformation measurements part II: experimental validation of uncertainty and bias estimates. Exp Mech 51:423–441. doi:10.1007/s11340-010-9450-3
7. Sutton MA, Orteu J-J, Schreier HW (2009) Image correlation for shape, motion and deformation measurements. Springer, New York. doi:10.1007/978-0-387-78747-3
8. Reu PL (2011) Experimental and numerical methods for exact subpixel shifting. Exp Mech 51:443–452. doi:10.1007/s11340-010-9417-4
9. ASTM (2008) Standard test method for determining forming limit curves. E2218-02, ASTM International
10. Reu PL (2013) A method for overlapping two DIC views by using a two-tone speckle pattern. In: Proceedings of the SEM annual meeting 2013
11. http://www.nist.gov/itl/sed/gsg/uncertainty.cfm
12. Holman JP (1994) Experimental methods for engineers, 6th edn. McGraw-Hill, New York

Chapter 32
Pre-qualifying DIC Performance Based on Image MTF Correlation Coefficient

Chi-Hung Hwang, Wei-Chung Wang, Yung-Hsiang Chen, Jia-He Chen, Yan-Ting Wu, Jheng-Yong Lyu, and Ya Hsi-Chiao

Abstract Attempts on developing the method for qualifying digital image correlation (DIC) method results, especial the displacement field, is presented in this paper. By taking a series images from center and margin of imaging field, DIC method is implemented to calculate displacement fields with respect to center of imaging field as origin of movement. Concurrently, residual displacements are obtained by subtracting mean value from the obtained displacement fields. The residual displacements are considered to be errors and mainly introduced by the imaging system of DIC. In this paper, modulation transfer function (MTF) is adopted to relate the residual displacement, DIC quality, with imaging system performance. By investing the MTFs cross-correlation values of all positions with respect to the reference image, a potential qualitative method for inspecting the DIC results based on sub-image MTF cross-correlation is proposed in this paper.

Keywords Digital image correlation • Modulation transfer function • Sub-image set • Cross-correlation • Field of view

32.1 Introduction

Digital image correlation method (DIC) is one of the image based optical strain measurement methods, the resolution can easily reach sub-pixel level thank to the improvement of personal computer computing power. DIC is first proposed in early 1980s [1, 2]; and now based on optical setups, DIC can be categorized as 2D DIC, 3D DIC and V-DIC methods [3–7]. In principle, DIC can be applied to determine displacement and strain filed from two digitized data matrixes (the images); therefore, the applications are extended to various imaging systems such as optical instruments, scanning microscopes, electronic beam microscopic systems and many others which make the DIC become a powerful method for different scales [8–11]. DIC method has been widely used for variety deformation measurements, such as surface strain of amorphous glassy polymer, soft materials, and even the deformation of thin plates subject to explosive blast.

While DIC is applied for displacement and strain field determination, imaging system must be calibrated. By calibration, the object-image coordinate relation can be obtained, and the affections of lens distortion and lens-sensor misalignment can also be minimized. Calibration procedures consist of taking a series of images of a target at giving position with various rotated angles; the calibration procedure can ensure DIC accuracy as the object is moved within the calibration region. However, in some DIC application scenarios, for example long-term displacement monitoring, the cross-correlations might be ill posed because large displacement might move objects out-of-depth-of-field (DOF) or to the margin of field-of-view (FOV) and the reference image and the second image set can be obtained from different viewing area, unpredictable DIC results are concluded.

Considering Modulation transfer function (MTF) is always used to determine optical and imaging system performances and can be evaluated by random patterns [12–15]. Hwang et al. [16] improved the displacement standard derivation from μm to sub-μm level by fitting MTFs-coefficient curve at Nyquest frequency of the imaging system. It is important to develop a method for per-evaluating the reference image and second-image-set are mechanical related or not. Considering images

C.-H. Hwang (✉) • Y.-H. Chen
ITRC, NARL, Hsinchu 30073, Taiwan ROC
e-mail: chhwang@narlabs.org.tw

W.-C. Wang • J.-H. Chen • Y.-T. Wu • J.-Y. Lyu • Y. Hsi-Chiao
Department of Power Mechanical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan ROC

captured out-of-focus or at margin of imaging filed, image contract becomes poor. Low contract is equal to poor visibility. MTF is a parameter based on image visibility which is generally used to verify imaging system performance. Therefore, in this paper, a qualitative method based on sub-image is proposed for DIC performance verification.

32.2 Principles

32.2.1 Digital Image Correlation Method

The DIC method tracks the surface pattern of an objective target among states; the displacement between two different states can be determined by searching the most-likely sub-image from the corresponding sub-images. Algorithms, such as the predefined cross-correlation and the sum of squared differences criterion, can be used to calculate the similarity between the sub-image in reference image and possible mapping sub-image in the deformed image [17] to determine the mean displacement of sub-images. By repeating the process, the whole field displacement can be determined.

32.2.2 Modulation Transfer Function

In general, an optical system performance can be determined by power of decomposing targets in object space and always evaluated by modulation transfer function (MTF). MTF is a spatial frequency, ξ, determined parameter and can be defined as

$$MTF(\xi) = \frac{M_{image}(\xi)}{M_{object}(\xi)}; \qquad (32.1)$$

where

$$M(\xi) = \frac{E_{\max} - E_{\min}}{E_{\max} + E_{\min}};$$

E_{max} and E_{min} are the maximum and minimum irradiances respectively. MTF indicates the image visibilities of an optical system at different spatial frequencies; that is, the ability to identify two near points on the object space. According to the physical definition of MTF, the MTF value would be limited to diffraction limit of an optical system. To determine the MTF at specified spatial frequency, a sinusoid pattern target of giving spatial frequency is always used. By Eq. (32.1), MTF of assigned spatial frequency, ξ, along a giving axis can be determined with a sinusoid pattern rotated at associated angle. MTF can also be determined with spatial-frequency domain method by the following relations;

$$\left|PSD_{image}(\xi)\right| = |MTF(\xi)|^2 \left|PSD_{object}(\xi)\right|; \qquad (32.2)$$

where PSD_{image} and PSD_{object} are the power spectrum densities (PSD) of image and object. Typically, there are three different patterns are generally used for MTF determination; the slit pattern which is a one-dimensional plus function and can determine one-dimensional MTF over wide spatial frequency range according to slit width; a point pattern which can be treated as a two-dimensional exciting function for two-dimensional MTF determination. The random pattern target is used for 2 dimensional MTF measurements. For a typical imaging lens, MTF value reaches maximum value at the center of the lens and then degrades to minimum value around the edge of the lens.

32.3 Experimental Setup

In this paper, a 2D-DIC measurement system with single camera is implemented as the test platform, as shown in Fig. 32.1. The main parameters of the system are tabulated in Table 32.1, including CCD Camera, lens, linear stage, and DIC software. In this test platform, the CCD pixel size is 4.4 μm ×4.4 μm and the corresponding Nyquist frequency is about 110 lp/mm.

Fig. 32.1 Experimental layout

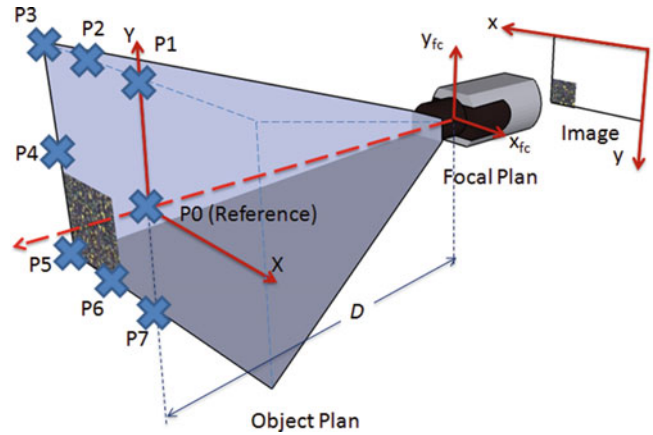**Table 32.1** System specifications

Specifications	Descriptions
CCD camera	Marlin F-201B CCD camera.
	Resolution: 1,628 × 1,236 pixels. Cell size: 4.4 μm. Max frame rate at full resolution: 12 fps. A/D: 12 bit.
Lens	Schneider Kreuznach Xenoplan 1.9/35, 2/3″ 35 mm lens, F#1.9.
Linear stage	T-LSM50A—Zaber Technologies Inc.
	Travel Range: 50.8 mm. Accuracy: ±8 μm. Repeatability: <1 μm.
DIC software	Vic-2D 2007, Correlated Solutions, Inc.
Specimen	Alumnus plate, dimension is 51 mm × 51mm.

Fig. 32.2 Default positions and DIC determined locations
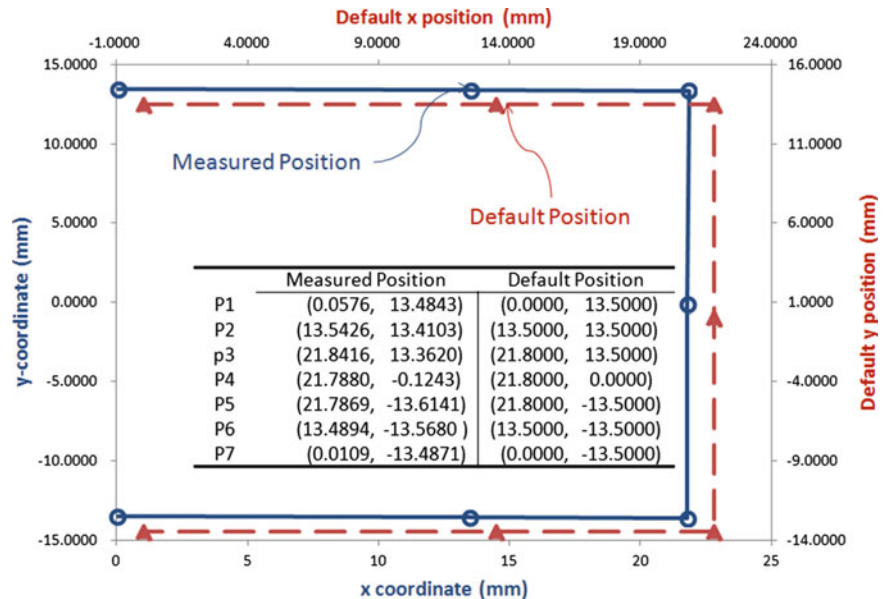
As for the measurement specimen (a random target) is made of aluminum and the surface was prepared with standard surface preparing process [18]. The specimen is clamped along the bottom edge and then mounted on a lab jack with fine screw makes the measurement specimen can be moved vertically along Y-axis. Then the lab jack is fixed on a computer controlled step-motor stage that enable the specimen perform in-plan displacements along X-axis. In this study, the specimen are first placed at position P1, and the camera is fine adjusted to ensure the camera's line of view is aligned perpendicular to the out-normal of the specimen surface to mitigate possible pointing error. After taking a reference image, the measurement specimen began to move to seven different positions, P1 to P7, as shown in Fig. 32.2; where the red-triangles indicate the default positions of specimen should be placed and the blue-circles are the measurement positions where the specimen

is placed. To define the margin of the camera image plan, the measurement specimen is moved up- down- and right- till the edge of the specimen reach the boundaries of the image plane. In Fig. 32.2, the red-triangles identify the defaulted measurement points, and the blue-circles mark the positions where the specimen are placed, and the positions are defined by the mean displacement of specimen calculated by DIC.

32.4 Results and Discussions

32.4.1 Displacement Measurement Results

The raw random pattern images at different position and the calculated U-, V-displacements are shown in Fig. 32.3a, the total displacement for each movement is shown in Fig. 32.3b. Inspecting Fig. 32.3a, b, all the surface plots indicated the calculated
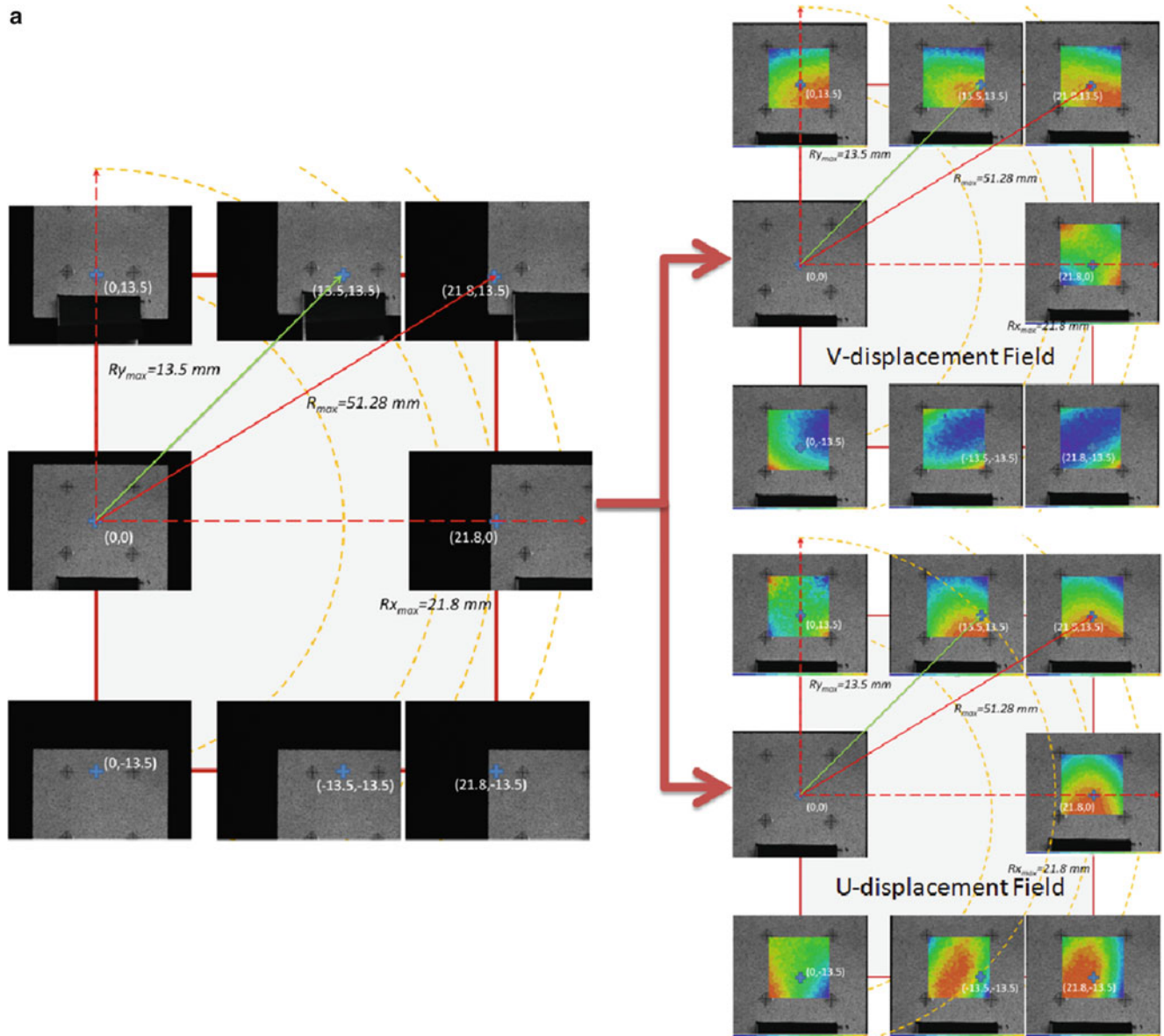

Fig. 32.3 Raw images and the corresponding displacement fields computed by DIC (unit: mm); (**a**) U-, V- displacement field computed by DIC method, (**b**) total displacement
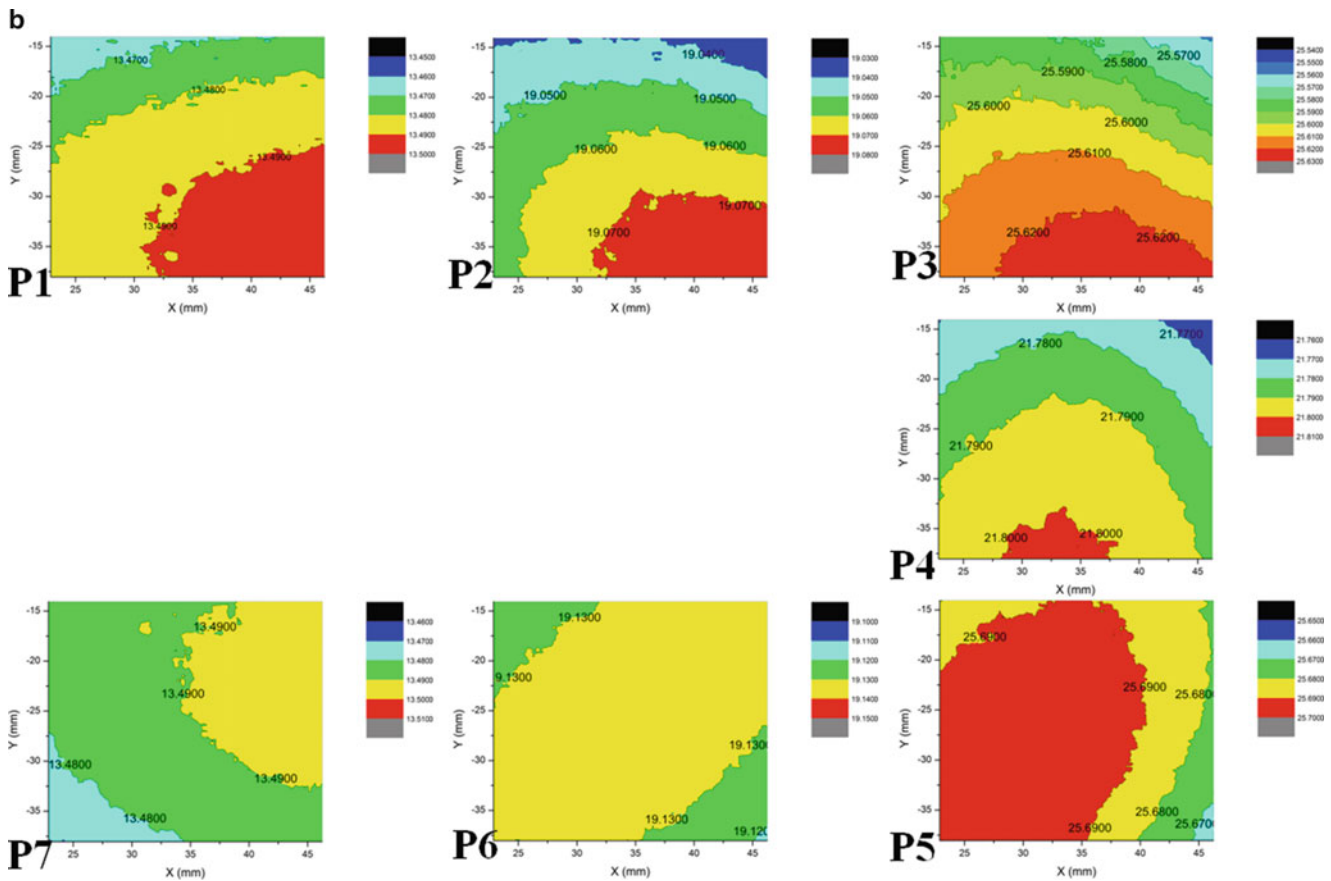
Fig. 32.3 (continued)

Table 32.2 DIC computed mean displacement and associated displacements standard derivations

Position	Location (mm)	Displacement (mm)					
		U- field		V- field		Total displacement field	
		Mean	σ_U	Mean	σ_V	Mean	σ
P1	(0, +13.5)	0.0576	0.0017	13.4843	0.0079	13.4844	0.0079
P2	(+13.5, +13.5)	13.5426	0.0090	13.4103	0.0066	19.0588	0.0108
P3	(+21.8, +13.5)	21.8416	0.0133	13.3620	0.0072	25.6047	0.0149
P4	(+21.8, 0)	21.7880	0.0084	−0.1243	0.0033	21.7884	0.0084
P5	(+21.8, −13.5)	21.7869	0.0067	−13.6141	0.0036	25.6908	0.0070
P6	(−13.5, −13.5)	13.4894	0.0026	−13.5680	0.0026	19.1326	0.0032
P7	(0, −13.5)	0.0109	0.0033	−13.4871	0.0051	13.4871	0.0051

displacement fields are not uniform but with displacement gradients. According to Fig. 32.3a, for up-position (Y is 13.5 mm), said position P1 to P3, the V- displacement gradients are available from bottom-right corner toward upper edge of the specimen, and the maximum gradients rotate clock-wisely as the specimen moved from upper-center to the upper-right position because. While specimen moved down to the X-axis, the V- displacement surface became a saddle surface. As the specimen is moved along low-position (Y is −13.5 mm), V- displacement fields are of concave shape from up-left to bottom-right, while the specimen moved from right to left, the concave shape changed and there is the maximum displacement gradient from bottom-left to up-right at position P7. As for the U- field, at up-position, at position P1 and P7 are saddle shape. Other than P1, for the up-positions P2 and P3, the gradients of U- displacement fields, in general, are from the bottom- to the up- edge, and U- displacement field at P4 has the geometrical characteristic. For low-position, the U- displacement fields are of convex shape and the maximum U-displacement region shrinks to center of specimen as the specimen position changed from P6 to P5. The corresponding mean displacements and standard derivations are tabulated in Table 32.2; in this study, the

standard derivations, σ_U, of U-displacement filed are smaller at positions P1, P7 and P6, and σ_U are increased as specimen is moved towards right-edge. Meanwhile, σ_U of up-positions P2 and P3 are related higher than low-positions P6 and P5. As for V-displacement field, the standard derivations, σ_V, at up-positions are higher than low-positions at the same nominal X-positions. Subtracting the mean U- and V- displacements, as shown in Fig. 32.4, the residual displacement fields δU- and δV- are plot with 0.1 pixel increments. Figure 32.5 is plot based on the data extracted from the δU- and δV- fields along X- and Y-axis, the δU- of P1 and P7 along X-axis δV- of position P4 along Y-axis are linear functions of X- and Y- respectively. A shown in Fig. 32.5b, c, δUs of the other five positions can be well fitted by the second degree polynomials; meanwhile, δUs have maximum around the center and become related smaller at the two ends. Regarding δV, as shown in Fig. 32.5d, e, the δVs can also be fitted by a second degree polynomials. δVs of up-positions, as shown in Fig. 32.5d, are negative at upper-edge of the spacemen and become positive at bottom-edge of the specimen. Different from δVs of up-positions, δVs of low-positions have negative and also minimum values at the center and related higher values at two ends. Considering the specimen is moved by a step-motor along X-direction and moved along Y-axis by driving fine screw of lab jack manually, the specimen is assumed to be deformation free during movement and derivations of displacement fields are not caused by external loads. Although the magnitudes of residual displacement, from 0.2 to 1.5 pixels, are small, but the residual displacement δU- and δV- indicate there might be errors introduced by the imaging system.

32.4.2 Experiments of Modulation Transfer Function

Considering MTF is always used to evaluate the performance of the optical and/or imaging system, if the optical/imaging system has problem such as out-of-defcous, misalignment, or poor SNR, the MTF would degard. Therefore, in this study, MTF is considered to be the indicator of DIC performance. The random pattern images captured for DIC analysis are also used for MTF calculation. The typical MTF curves along X-direction computed from 256×256 pixels of images captured at different positions are shown in Fig. 32.6. From subplot of Fig. 32.6, the MTF around Nyquist frequency (110 lp/mm in this study) P5 is close to P0 and P6 is also higher than other MTFs of up-positions. All the MTFs are fitted with four degree polynomials, and then cross-correlating all MTFs curves of different positions with MTF at P0, the maximum values of each cross-correlation results are listed in Table 32.3. All the cross-correlation values of fitted MTF curves are over 0.999 at all positions. Considering the imaging system of the DIC system utilizes a low distortion lens and the random target moved in-plan, therefore the optical distortions might be small in this study, and the changes of MTFs obtained from full frame images are insensitive to MTF cross-correlation values. The cross-correlation of MTFs of two images can be similar to calculate the averaged distribution density of random dots in various distances or sizes, therefore, sub-images from P0 and other positions are used to calculate MTF and corresponding MTF cross-correlation. As shown in Fig. 32.6, Sub-1, Sub-2 and Sub-3 are three sub-images with 128×128 pixels in size cropped from up-left corner, center and bottom-right corner of the original image of P3, and also three sub-images are taken out-off image of P0. Again, the MTFs of corresponding sub-images are calculated, then cross-correlating the MTFs data of corresponding sub-images both from P0 and P3. The cross-correlation results are summarized in Table 32.4, and the full image MTF cross-correlation values are calculated from images at P0 and P3 with 592×592 pixels in size. In general, smaller sub-image size has higher MTF cross-correlation value. Therefore, if the cross-correlation values of sub-images are smaller than the values from full images, the averaged random pattern characteristic are different in the assigned region, then the computed displacement need to be reinvestigated (Fig. 32.7).

32.5 Conclusions

In this paper, DIC method calculates the displacement field with series images which are taken at center and margin field-of-view of the imaging system. The specimen is moved in-plane without external load; however, the calculated displacement fields are not uniform and the displacement gradients exist. Then residual displacement fields are obtained by remove the mean-displacement; the magnitudes of residual displacement are small, 0.2–1.5 pixels in this study, but the residual displacement δU and δV indicate extra-errors might be introduced by the imaging system. MTF is considered to be the indicator of DIC performance because MTF is always used to evaluate the performance of the optical and/or imaging system. However, in this paper, the imaging system utilizes low distortion lens and the random target moved in-plan, therefore the changes of MTFs obtained from full frames are insensitive to the cross-correlation values. In this paper, a qualitative method based on sub-image is proposed to indicate the DIC performance. The process begins from evaluating full
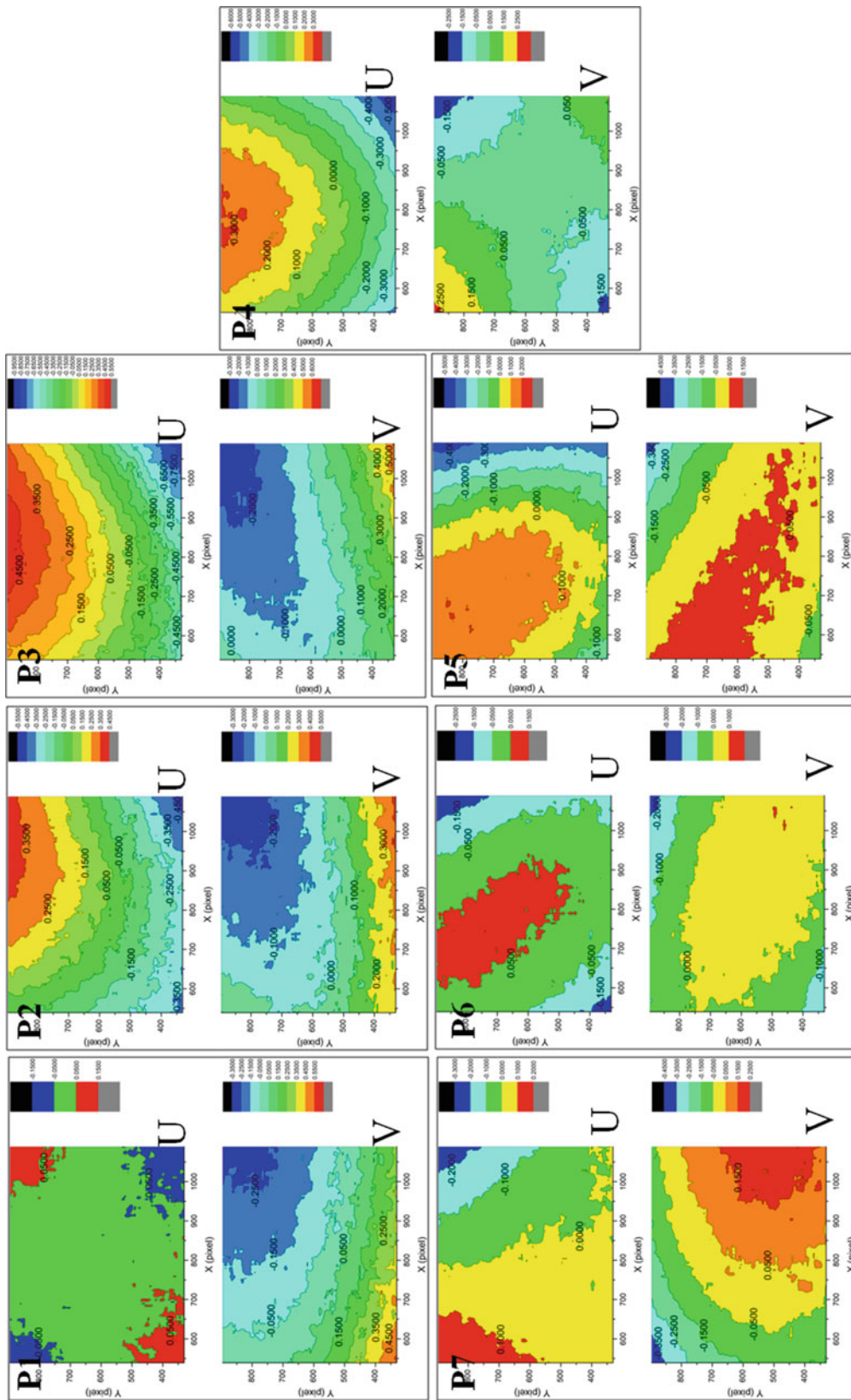
Fig. 32.4 Residual U-, V- displacement fields by removing mean displacement (unit: pixel)
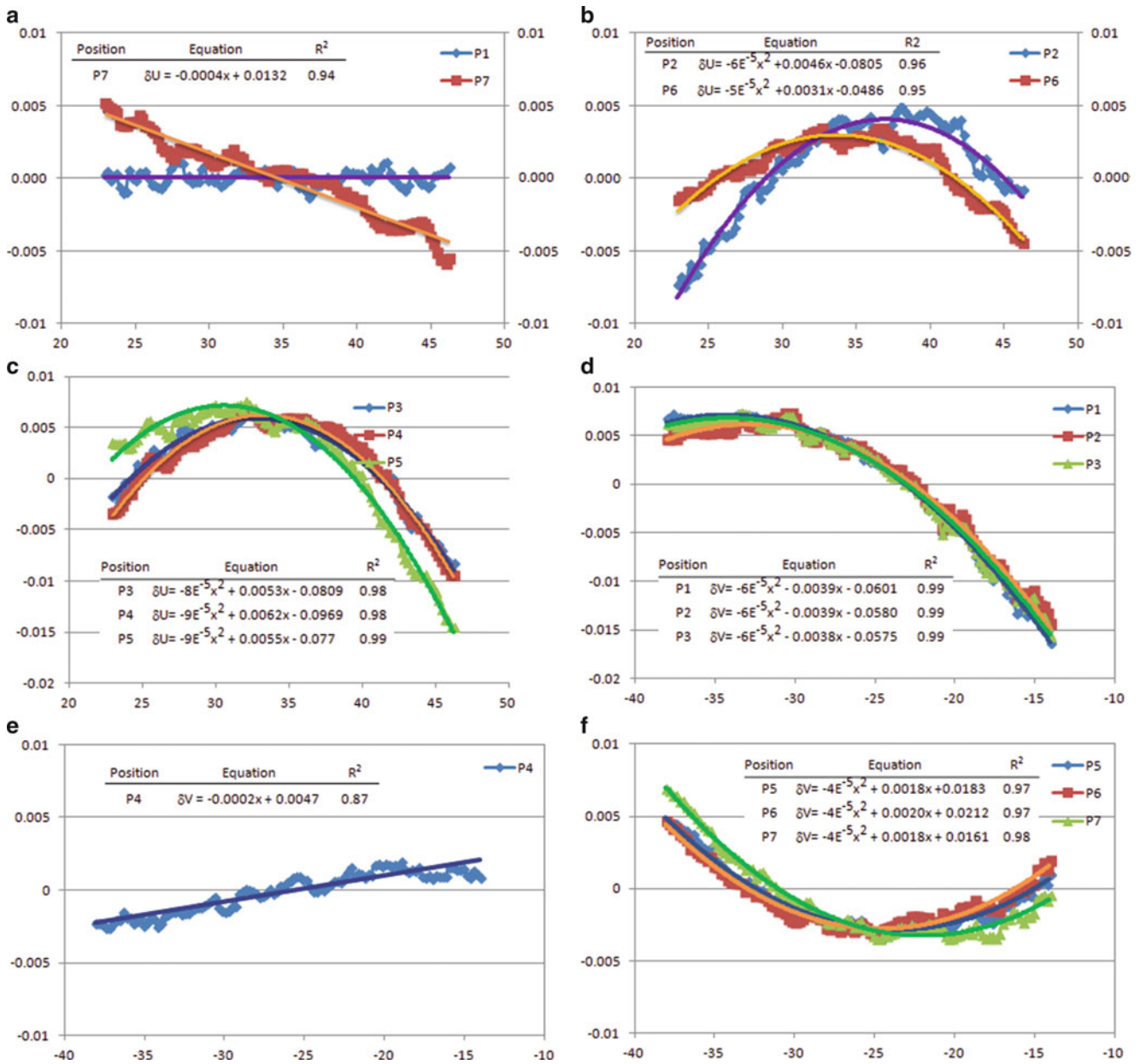
Fig. 32.5 Fitting curves of residue displacement δ U and δ V at varied positions (unit: pixel) (**a**) δ U of P1 and P7 along X-axis (**b**) δ U of P2 and P6 along X-axis (**c**) δ U of P3, P4 and P5 along X-axis (**d**) δ V of P1, P2 and P3 along Y-axis (**e**) δ V of P4 along Y-axis (**f**) δ V of P5, P6 and P7 along Y-axis

image frame MTF cross-correlation value with respect to reference image. The sub-image MTFs are then calculated with respect to corresponding sub-image of reference one. Comparing the cross-correlation values obtained from full images and sub-images, and suggesting the DIC results have to be checked if the cross-correlation values of sub-images are far below the value obtained from full images.
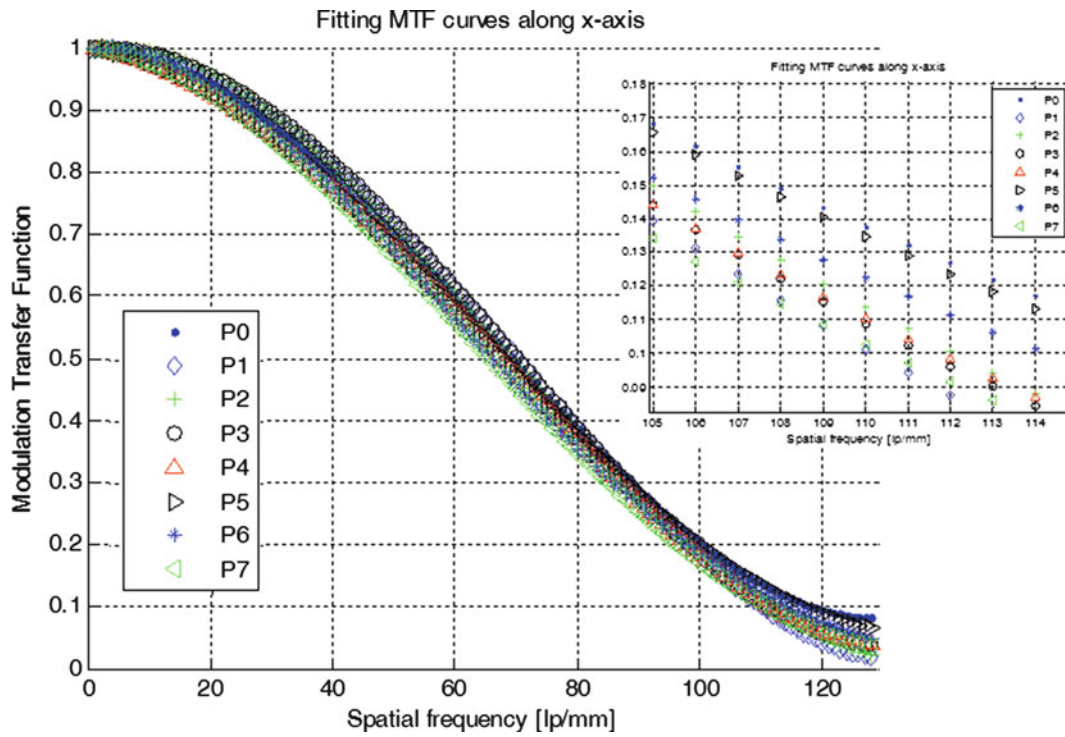
Fig. 32.6 MTF curves from images of random target at different positions

Table 32.3 MTFs cross-correlation with image P0 and the sampling images have 256 × 256 pixels in size

Position		P1	P2	P3	P4	P5	P6	P7
Cross-correlation value	X-direction	0.9994	0.9996	0.9996	0.9997	0.9999	0.9995	0.9992
	Y-direction	0.9999	0.9998	1.0000	0.9999	0.9995	0.9997	0.9997

Table 32.4 P0-P3 MTF Correlation factors of three Defined Sub-images

| | | Sub-images (128 × 128) | | |
	Full image	Up-left (Sub-1)	Center (Sub-2)	Low-right (sub-3)
X-direction	0.9988	0.9947 (−)	0.9992 (+)	0.9978 (−)
Y-direction	0.9983	0.9927 (−)	0.9975 (−)	0.9961 (−)
45-direction	0.9682	0.8229 (−)	0.9805 (+)	0.9869 (+)

(−): the cross-correlation value of sub-images is smaller than obtained from full frame of image
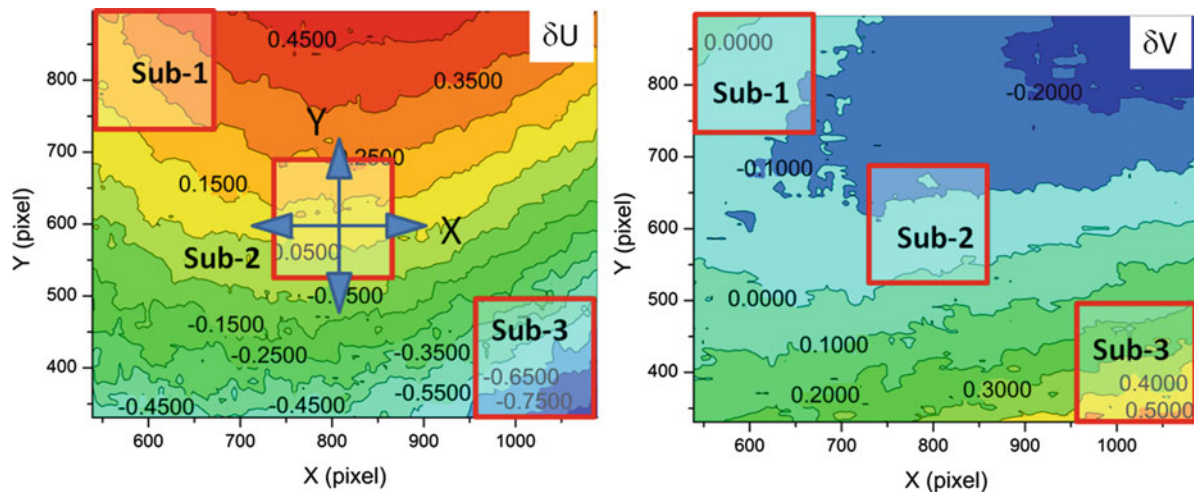(+): the cross-correlation value of sub-images is higher than obtained from full frame of image

Fig. 32.7 Calculating MTF with sub-images with 256 × 256 pixels in size

References

1. Peters W, Ranson W (1982) Digital imaging techniques in experimental stress analysis. Opt Eng 21:427–431
2. Sutton M, Wolters W, Peters W, Ranson W, McNeill S (1983) Determination of displacements using an improved digital correlation method. Image Vis Comput 1(3):133–139
3. Vendroux G, Knauss WG (1998) Submicron deformation field measurements: Part 1. Developing a digital scanning tunneling microscope. Exp Mech 38(1):18–23
4. Luo PF, Chao YJ, Sutton MA, Peters WH (1993) Accurate measurement of three-dimensional deformations in deformable and rigid bodies using computer vision. Exp Mech 33(2):123–132
5. Tiwari V, Sutton MA, McNeill SR (2007) Assessment of high speed imaging systems for 2D and 3D deformation measurements: methodology development and validation. Exp Mech 47(4):561–579
6. Sutton MA, Orteu J-J, Schreier HW (2009) Image correlation for shape, motion and deformation measurements–basic concepts, theory and applications. Springer, New York
7. Bay BK, Smith TS, Fyhrie DP, Saad M (1999) Digital volume correlation: three-dimensional strain mapping using X-ray tomography. Exp Mech 39(3):217–226
8. Schreier H, Garcia D, Sutton M (2004) Advances in light microscope stereo vision. Exp Mech 44(3):278–288
9. Doumalin P, Bornert M, Caldemaison D (1999) Microextensometry by image correlation applied to micromechanical studies using the scanning electron microscopy. In: Proceedings of international conference on advanced technology in experimental mechanics, Ube, Japan, pp 81–86
10. Gianola DS, Sedlmayr A, Monig R, Volkert CA, Major RC, Cyrankowski E, Asif SAS, Warren OL, Kraft O (2011) In situ nanomechanical testing in focused ion beam and scanning electron microscopes. Rev Sci Instrum 82(6), 063901
11. Zhu T, Sutton MA, Li N, Orteu JJ, Cornille N, Li X, Reynolds AP (2011) Quantitative stereovision in a scanning electron microscope. Exp Mech 51(1):97–109
12. Daniels A, Boreman GD, Ducharme AD, Sapir E (1993) Random targets for modulation transfer function testing. SPIE 1969:184
13. Hong QH, Lettington AH, Macdonald J (1996) Measuring the MTF for focal plane arrays using random noise targets. Meas Sci Technol 7(7):1087
14. Levy E, Peles D, Opher-Lipson M, Lipson SG (1999) Modulation transfer function of a lens measured with a random target method. Appl Opt 38(4):679–683
15. Fernandez-Oliveras A, Pozo AM, Rubino M (2010) Comparison of spectacle-lens optical quality by modulation transfer function measurements based on random-dot patterns. Opt Eng 49(8):083603–083606
16. Hwang C-H, Wang W-C, Hung T-H, Chen J-H (2014) Indicating DIC potential correlation errors with optical modulation transfer function. Adv Opt Methods Exp Mech 3:137–144
17. Pan B, Xie HM, Wang ZY (2010) Equivalence of digital image correlation criteria for pattern matching. Appl Opt 49(28):5501–5509
18. Vic-2D Testing Guide Correlated Solutions (2007)

Chapter 33
Analysis of E-Beam Microlithography and SEM Imaging Distortions

A. Guery, F. Latourte, F. Hild, and S. Roux

Abstract Surface patterning by e-beam lithography and SEM imaging distortions are studied via digital image correlation. The surface of a stainless steel specimen is marked with a numerically-generated random pattern by microlithography. The global distortions from the reference pattern are first quantified by digital image correlation between the virtual reference pattern and the actual SEM image both in secondary and backscattered electron imaging modes. A second order polynomial basis reveals sufficient to capture most of the distortions. They result from both patterning and imaging techniques. To separate the two contributions without resorting to an external caliper, it is proposed to analyze a series of images of the same patterned surface acquired after rotations of the specimen by different angles. The apparent displacement fields are expressed as a static field, corresponding to the imaging distortion, and another one that rotates together with the specimen. Because large rotations are considered, the problem is nonlinear in the entire set of parameters characterizing each displacement field, but can be solved with an iterative scheme. The obtained patterning distortions appear smaller than those due to imaging on wide field images.

Keywords SEM • Microlithography patterning • Digital image correlation • Patterning distortions • Imaging distortions

33.1 Introduction

Kinematic measurements are performed to study the mechanical behavior of materials. This can be done at the microstructure scale by using a Scanning Electron Microscope (SEM). Such spatially dense experimental information offers a wide scope of possibilities, such as for example the comparison with simulations to identify material parameters [1] associated with crystal plasticity constitutive equations [2]. Yet, the imaging distortions induced by the SEM electromagnetic environment should be taken into account to ensure the accuracy of the experimental information [3].

To perform kinematic measurements using Digital Image Correlation (DIC), a gray level pattern is required on the observed surface of the sample. When the natural texture of the material is not suitable, an artificial one may be deposited onto the surface. In the case of SEM imaging, a common practice is to resort to microgrid marking using e-beam microlithography [4]. In this study, the same marking technique is employed, but has been tailored to create a numerically-generated random speckle, which presents the advantage (for DIC) to not induce spatially periodic patterns.

A. Guery (✉)
EDF R&D, Site des Renardières, Avenue des Renardières - Ecuelles, 77818 Moret-sur-Loing, France

LMT Cachan, ENS Cachan, CNRS UMR 8535, PRES UniverSud Paris, 61 Avenue du Président Wilson,
94235 Cachan Cedex, France
e-mail: adrien.guery@lmt.ens-cachan.fr

F. Latourte
EDF R&D, Site des Renardières, Avenue des Renardières - Ecuelles, 77818 Moret-sur-Loing, France

F. Hild • S. Roux
LMT Cachan, ENS Cachan, CNRS UMR 8535, PRES UniverSud Paris, 61 Avenue du Président Wilson,
94235 Cachan Cedex, France

Global distortions between the theoretical (or "reference") pattern and its SEM image, resulting from patterning and imaging, are measured by DIC. A partition method of these two contributions is performed using several SEM images of the pattern after a set of rotations of the sample in the SEM.

33.2 Global Distortions Measurement by Digital Image Correlation

A gold speckle is deposited onto the surface of an in-situ specimen by a process of microlithography patterning. The material is 316LN austenitic stainless steel. The surface is prepared by mechanical polishing with colloidal silica finish. The process is divided into five main steps (see Fig. 33.1). The principle consists of exposing a resist film, deposited onto the surface, with the electron beam of the SEM only at specific positions following a pattern design. For that purpose, a Raith Elphy Quantum commercial system is used to control the FEI Quanta FEG 600 SEM. After the dissolution of the exposed resist, a very thin layer of gold is deposited over the entire surface, and a final dissolution step of the resist film leaves gold only in the region that were first swept by the e-beam. In the present case, the speckle pattern is chosen to be a known distribution of disk-shaped spots of random radius and position.

The speckle covers a 400 μm side square surface. Figure 33.2a shows an SEM image of the imaged gold pattern using backscattered electron (BSE) detector at a working distance of 14 mm, with an acceleration voltage of 10 kV. Images are acquired with a physical size of 240 nm per pixel and stored in 16 bit format. Thanks to the histogram of gray levels of the SEM image, a reference image of the known distribution of spots is built. Its background is set as a homogeneous gray

Fig. 33.1 Surface gold patterning using e-beam microlithography: resin film deposition (**a**), electron beam exposure (**b**), exposed resist film development (**c**), gold deposition (**d**), residual resist film dissolution (**e**)
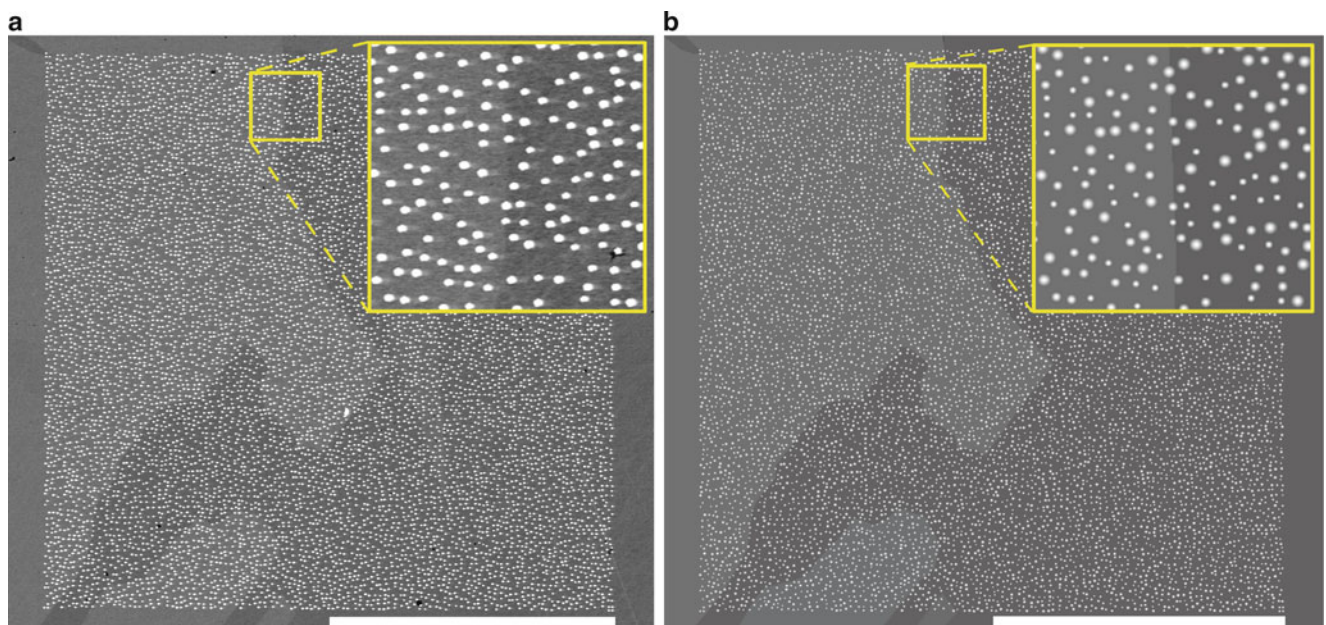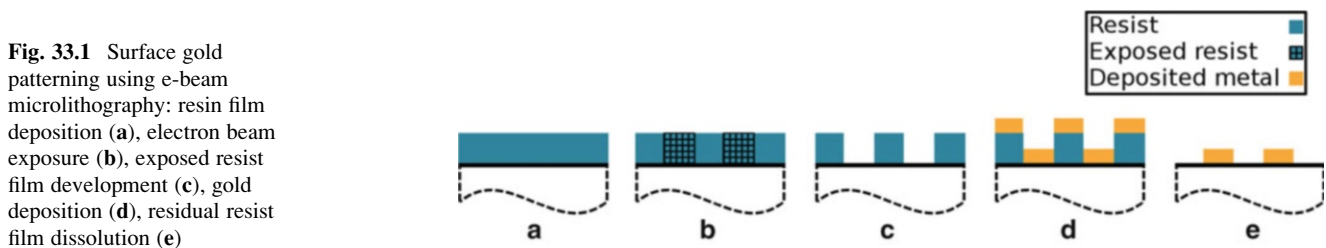

Fig. 33.2 SEM image of the gold pattern using BSE detector (**a**), and corresponding reference image generated numerically (**b**). The scale bar is 200 μm

level read at the peak of the histogram corresponding to the uncovered alloy. Spots are set as a Gaussian-shaped distribution of gray levels whose mean value is read at the peak of the histogram related to the gold disks. Because the underlying microstructure of the material involves different gray levels (see Fig. 33.2a), this process is performed regionally after a simple first segmentation of the main regions. Figure 33.2b shows the reference image obtained from the SEM image in Fig. 33.2a.

The displacement fields between the SEM image, denoted f, and the computer-generated image, g, of the pattern are measured using finite element based DIC [5], which is based on the conservation of the texture

$$f(\mathbf{z}) = g(\mathbf{z} + \mathbf{u}(\mathbf{z})) \tag{33.1}$$

where $\mathbf{z}$ is the complex-valued coordinates of pixels in f, and $\mathbf{u}$ the sought displacement field. The minimization of the gray level residuals leads to the iterative solution of linear systems [6]

$$[\mathbf{M}]\{\delta\mathbf{u}\} = \{\mathbf{s}\} \tag{33.2}$$

with

$$M_{ij} = \sum_{ROI} (\boldsymbol{\psi}_i \cdot \nabla f)(\mathbf{z})\left(\boldsymbol{\psi}_j \cdot \nabla f\right)(\mathbf{z}) \tag{33.3}$$

and

$$s_i = \sum_{ROI} (f - \widetilde{g})(\mathbf{z})(\boldsymbol{\psi}_i \cdot \nabla f)(\mathbf{z}) \tag{33.4}$$

where, at each iteration, $\{\delta\mathbf{u}\}$ is a vector of the displacement corrections, and $\widetilde{g}$ the deformed image corrected by the current estimation of the displacement field. $\boldsymbol{\psi}_i$ are chosen basis functions for the displacement discretization, in the present case an unstructured mesh of 3-noded triangular elements (about 40 pixel/9.6 μm sides). With this DIC technique, a Root Mean Square (RMS) value of the displacement uncertainty of 0.02 pixel (or 5.0 nm) is measured (between 2 SEM images with no motion).

In addition to rigid body motions, the obtained field includes distortions resulting from both patterning and imaging errors. The separation of these two contributions is not reachable so far.

33.3   Partition of Distortions

A partition method of global distortions between patterning and imaging distortions is now presented. A set of ten rotations of an angle θ^k ranging from 0° to 90° of the observed pattern is applied inside the SEM chamber (via the stage rotation) prior to imaging. Displacement fields, denoted $\mathbf{u}^k$, are then measured using DIC between the acquired images and the reference image of the pattern. Because large rotations are prescribed, the previous DIC formulation has been adapted considering finite transformation kinematics [7]. The partition method consists of considering each displacement field $\mathbf{u}^k$ as a composition of three transformations, namely, a patterning distortions field $\mathbf{v}_g$, identical for each $\mathbf{u}^k$, a rotation of an angle θ^k, and an imaging distortions field $\mathbf{v}_d$, assumed to be time-independent. This procedure is described in Fig. 33.3.
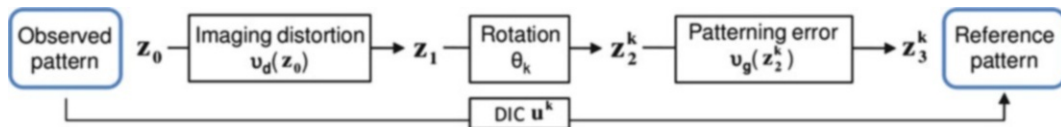

Fig. 33.3 Schematic of the partition of each of the k measurements of global distortions

The two unknown distortion fields are chosen to be described as a linear combination of polynomials expressed in the complex plane as

$$\mathbf{d}_{kl}(\mathbf{z}) = \mathbf{z}^k \overline{\mathbf{z}}^l \quad \text{with} \quad 0 \leq k, \ 0 \leq l, \ k+l \leq p \tag{33.5}$$

where $\overline{\mathbf{z}}$ denotes the conjugate of $\mathbf{z}$, and p the order of the series. Analytical expressions generally proposed in the literature [8] are described via this chosen basis too. A second order expansion of this polynomial basis appeared sufficient to describe both patterning and imaging distortions fields. Thus, $(a_1^d, \ldots, a_6^d)$ and $(a_1^g, \ldots, a_6^g)$ denote the unknown complex-valued coefficients of the polynomial series describing $\mathbf{v}_d$, respectively $\mathbf{v}_g$.

For each of the k displacement measurements, the deformed coordinate system in the reference image is given by

$$\mathbf{z}_3^k = \mathbf{z}_0 + \mathbf{u}^k(\mathbf{z}_0) \tag{33.6}$$

With the proposed partition method, an approximation of $\mathbf{z}_3^k$, denoted $\hat{\mathbf{z}}_3^k$, is proposed as

$$\hat{\mathbf{z}}_3^k = \left(\mathbf{z}_0 + v_d\left(\mathbf{z}_0, a_1^d, \ldots, a_6^d\right) \right) \ e^{i\theta^k} + v_g\left(\hat{\mathbf{z}}_2^k, a_1^g, \ldots, a_6^g\right) \tag{33.7}$$

with

$$\hat{\mathbf{z}}_2^k = \left(\mathbf{z}_0 + v_d\left(\mathbf{z}_0, a_1^d, \ldots, a_6^d\right) \right) \ e^{i\theta^k} \tag{33.8}$$

The error fields are then expressed as

$$\mathbf{r}^k = \mathbf{z}_3^k - \hat{\mathbf{z}}_3^k \tag{33.9}$$

These fields are minimized in a *weighted* least squares criterion with respect to the sought coefficients $(a_1^d, \ldots, a_6^d)$ and $(a_1^g, \ldots, a_6^g)$

$$\left(a_1^d, \ldots, a_6^d, a_1^g, \ldots, a_6^g\right) = \operatorname{argmin}\left(\sum_k \{\mathbf{r}^k\} \left[\mathbf{M}^k\right] \{\mathbf{r}^k\} \right) \tag{33.10}$$

where $[\mathbf{M}^k]$ is the DIC matrix associated to the measured degrees of freedom of $\mathbf{u}^k$, which is shown to be proportional to the inverse of the covariance matrix of measured degrees of freedom [6], and therefore appears adequate to weight the error fields in the least squares criterion.

The problem is nonlinear because of large rotations. However, it becomes linear if considering one of the functions $\mathbf{v}_d$ or $\mathbf{v}_g$ as known, and the other one is sought, and vice-versa. A few iterations are needed for converging to a minimum of the residual function. Figure 33.4 shows the change of the RMS error fields $\mathbf{r}^k$ as a function of the measured rotation angle of the pattern between the first and the last iteration.

Obtained imaging distortions and patterning errors at convergence are shown in Fig. 33.5. Note that, in this figure, the element of the polynomial series describing magnification changes (i.e., $\mathbf{d}_{10}(\mathbf{z}) = \mathbf{z}$, with a real amplitude) is not considered. Its amplitude is due to small magnification changes between the SEM image and the reference picture. Electron beam microlithography appears to involve smaller errors (in the 400 nm range) than SEM imaging does (in the 600 nm range). The imaging distortion magnitude is less than that observed by Sutton et al. [3] (of the order of 2.5 µm) but at a magnification about two times lower, and with a (very) different SEM. Moreover, the maximum levels of patterning errors are reached at the edges of the area of interest, which is consistent with the second step of the patterning process. The exposure of each elementary pattern (circular disks) is performed by electron beam deflection from the center of the square surface to mark, and thus it is expected to be less accurate far from the center. Last, the same experiment has been conducted with twice the previous magnification with a 100 µm side square speckle and very similar results have been obtained.

Fig. 33.4 RMS value of the error fields expressed in pixels as a function of the measured rotation angle of the pattern in degrees. The physical size of one pixel is 240 nm
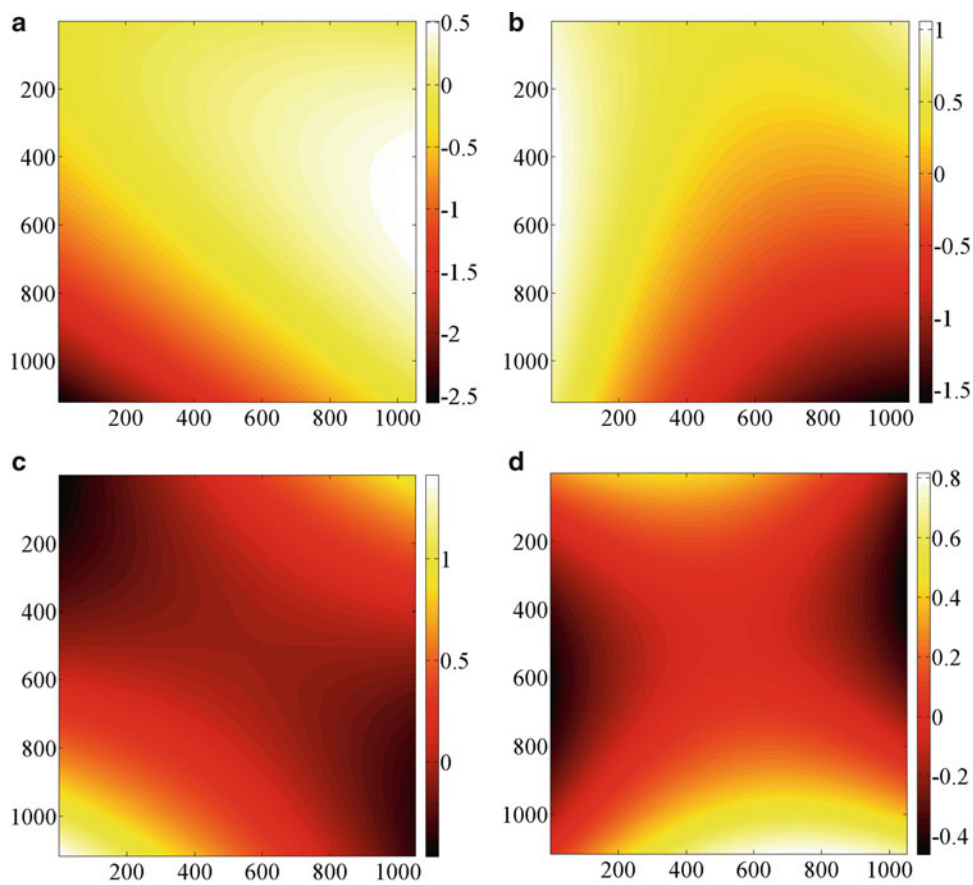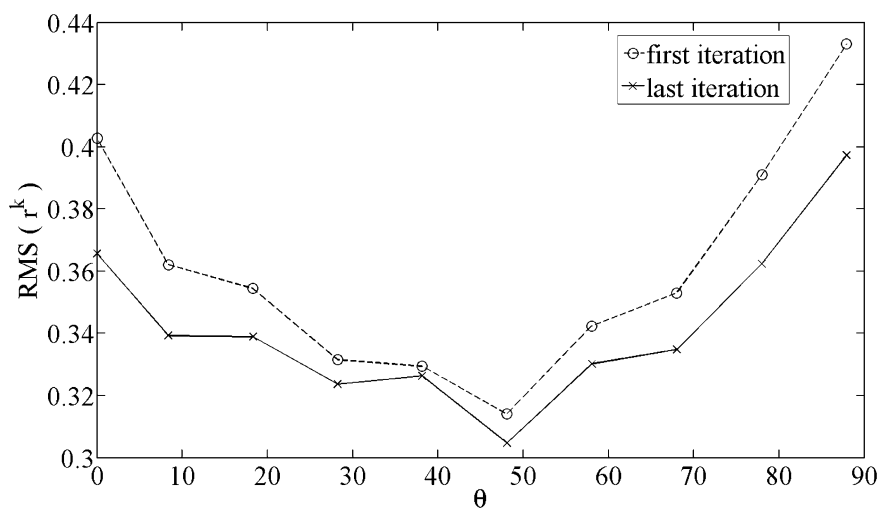

Fig. 33.5 *Horizontal* (**a**) and *vertical* (**b**) components of imaging distortions. *Horizontal* (**c**) and *vertical* (**d**) components of patterning errors. The fields are expressed in pixels, whose physical size is 240 nm

33.4 Conclusions

A gold patterning process has been introduced to mark stainless steel surfaces with a numerically-generated random pattern using microlithography. Global distortions have been quantified using digital image correlation between an SEM image of the pattern and its corresponding reference image, including both patterning and SEM imaging biases. It has then been shown that it is possible to separate these two contributions by collecting several evaluations of this global distortion with various angles of rotation of the observed pattern. The patterning has been shown to involve lower errors than SEM imaging distortions. From these analyses, it would be relevant to check for any additional drifts due to SEM environment too. The subsequent use of SEM images for kinematic measurements requires a good characterization and quantification of the measurement errors.

References

1. Héripré E, Dexet M, Crépin J, Gélébart L, Roos A, Bornert M, Caldemaison D (2007) Coupling between experimental measurements and polycrystal finite element calculations for micromechanical study of metallic materials. Int J Plast 23(9):1512–1539
2. Roters F, Eisenlohr P, Hantcherli L, Tjahjanto DD, Bieler TR, Raabe D (2010) Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: theory, experiments, applications. Acta Mater 58:1152–1211
3. Sutton MA, Li N, Garcia D, Cornille N, Orteu JJ, McNeill SR, Schreier HW, Li X (2006) Metrology in a scanning electron microscope: theoretical developments and experimental validation. Meas Sci Technol 17:2613–2622
4. Allais A, Bornert M, Bretheau T, Caldemaison D (1994) Experimental characterization of the local strain field in an heterogeneous elastoplastic material. Acta Metall Mater 42(11):3865–3880
5. Leclerc H, Périé J-N, Roux S, Hild F (2009) Integrated digital image correlation for the identification of mechanical properties, MIRAGE 2009. LNCS 5496:161–171
6. Besnard G, Hild F, Roux S (2006) "Finite-Element" displacement fields analysis from digital images: application to portevin-le chatelier bands. Exp Mech 46(6):789–803
7. Guery A, Latourte F, Hild F, Roux S (2014) Characterization of SEM speckle pattern marking and imaging distortion by digital image correlation. Meas Sci Technol 25:12
8. Sutton MA, Orteu JJ, Schreier HW (2009) Image correlation for shape, motion and deformation measurements: basic concepts, theory and application, Springer handbook of experimental solid mechanics. Springer, Berlin

Chapter 34
Displacement Measurements Using CAD-Based Stereo-DIC

J.-E. Dufour, B. Beaubier, F. Hild, and S. Roux

Abstract The aim of this study is to measure displacement fields during an experiment on a structure by using a priori information about its surface representation in a stereoDIC framework. The main goal of the present method is to measure the displacement and strain fields in formalisms completely consistent with the surface description via global stereoDIC. Although the majority of stereoDIC methods provide clouds of 3D displacement vectors associated with clouds of 3D points, which are subsequently interpolated to get continuous fields, the proposed CAD-based stereoDIC allows continuous 3D displacement fields to be measured directly. Therefore, there is a direct link between the measured fields and computational results since common shape and kinematic bases are considered. No additional interpolation or data manipulation is needed with such approaches. One practical example will illustrate the use of this novel technique.

Keywords Stereo-correlation • Global approach • Displacement measurement • DIC • CAD representation

34.1 Introduction

StereoDIC is a powerful method to measure 3D shapes, 3D displacement and 2D strain fields. This technique has been utilized at different scales, from the smallest ones using SEM images to large structures. It is mostly used as a medium scale measurement method in place of standard 2D-DIC method to avoid artifacts associated with out-of-plane motions or to measure 3D displacements of a 3D geometry [1–3].

The present work is dedicated to the measurement of continuous 3D displacement (and 2D strain) fields. It follows up a previous work focusing on calibration and shape measurement [4]. Unlike other stereovision techniques, the present method provides a dense description of the measured fields (i.e. 3D shape, 3D displacements and 2D strains) based upon freeform surface using NURBS descriptions of the observed structure [5]. Moreover, the number of degrees of freedom is reduced and this description provides a direct link with CAD software. It is worth noting that the description used herein (i.e. with NURBS) can be replaced by any parametric formulation of fields such as facets or mechanics-related fields (e.g. finite element or boundary element methods).

J.-E. Dufour (✉)
LMT Cachan, ENS Cachan, CNRS UMR 8535, PRES UniverSud Paris, 61 Avenue du Président Wilson,
94235 Cachan Cedex, France

Snecma Villaroche, Rond point René Ravaud-Réau, 77550 Moissy-Cramayel, France
e-mail: dufour@lmt.ens-cachan.fr

B. Beaubier
LMT Cachan, ENS Cachan, CNRS UMR 8535, PRES UniverSud Paris, 61 Avenue du Président Wilson,
94235 Cachan Cedex, France

PSA Peugeot Citroën, Centre Technique de Vélizy B, Zone aéronautique Louis Bréguet,
78140 Vélizy Villacoublay, France

F. Hild • S. Roux
LMT Cachan, ENS Cachan, CNRS UMR 8535, PRES UniverSud Paris, 61 Avenue du Président Wilson,
94235 Cachan Cedex, France

The outline of the paper is as follows. First, the formulation of the displacement measurement using a continuous description of the observed surface is presented. Then, the particular case of NURBS representation is detailed. Finally an industrial application is presented (i.e. welding of a roof top).

34.2 CAD-Based Stereo-DIC Formulation for Displacement Measurements

Let us denote by f^l and f^r the reference pictures (respectively acquired with the left and right cameras) and g^l and g^r the pictures in the deformed configuration (respectively shot with the left and right cameras). A global approach [6] to stereoDIC consists of minimizing the functional η^2

$$\eta^2 = \iint\limits_{ROI} \left(\underbrace{g^l(\mathbf{x}^l + \mathbf{u}^l(\mathbf{U})) - f^l(\mathbf{x}^l)}_{\rho^l(v,w)} \right)^2 dvdw \; + \; \iint\limits_{ROI} \left(\underbrace{g^r(\mathbf{x}^r + \mathbf{u}^r(\mathbf{U})) - f^r(\mathbf{x}^r)}_{\rho^r(v,w)} \right)^2 dvdw \tag{34.1}$$

where v and w are the parameters defining the observed surface, $\mathbf{x}^l$ and $\mathbf{x}^r$ the coordinates in the left and right cameras, and $\mathbf{U}$ the sought displacement field expressed in the 3D space of coordinates $\mathbf{X}$. To assess the quality of the registration, local correlation residuals ρ^l and ρ^r are constructed for both cameras. Let us decompose the displacement field $\mathbf{U}$ on an arbitrary basis of fields $\mathbf{\Phi}_i$

$$\mathbf{U}(\mathbf{X}) = \sum_i a_i \mathbf{\Phi}_i(\mathbf{X}) \tag{34.2}$$

where the amplitudes a_i are to be determined. A Taylor expansion of this functional leads to

$$\eta^2_{lin} = \iint\limits_{ROI} \left(g^l(\mathbf{x}^l) + \nabla f^l(\mathbf{x}^l) \cdot \frac{\partial \mathbf{u}^l}{\partial a_i}(\mathbf{U}) da_i - f^l(\mathbf{x}^l) \right)^2 dvdw \; + \; \iint\limits_{ROI} \left(g^r(\mathbf{x}^r) + \nabla f^r(\mathbf{x}^r) \cdot \frac{\partial \mathbf{u}^r}{\partial a_i}(\mathbf{U}) da_i - f^r(\mathbf{x}^r) \right)^2 dvdw \tag{34.3}$$

Minimizing η^2_{lin} leads to a series of linear problems to determine the corrections to the amplitudes a_i gathered in a column vector $\{d\mathbf{a}\}$

$$\left([\mathbf{C}^l] + [\mathbf{C}^r] \right) \{d\mathbf{a}\} = \left([\mathbf{b}^l] + [\mathbf{b}^r] \right) \tag{34.4}$$

where $[\mathbf{C}^l]$ and $[\mathbf{C}^r]$ are the stereoDIC matrices, and $\{\mathbf{b}^l\}$ and $\{\mathbf{b}^r\}$ the stereoDIC vectors.

34.3 Nurbs Description

In this study, a new algorithm to measure displacement fields is proposed. It is based upon the *a priori* knowledge about the geometry of the observed surface. The description of the latter is a freeform surface using Non-Uniform Rational Basis Splines (or NURBS Patches [5]). The linearized functional to minimize is rewritten as

$$\eta^2_{lin} = \iint\limits_{ROI} \left(g^l(\mathbf{x}^l) + \nabla f^l(\mathbf{x}^l) \cdot \frac{\partial \mathbf{u}^l}{\partial \mathbf{P}_{ij}}(\mathbf{U}) d\mathbf{P}_{ij} - f^l(\mathbf{x}^l) \right)^2 dvdw + \iint\limits_{ROI} \left(g^r(\mathbf{x}^r) + \nabla f^r(\mathbf{x}^r) \cdot \frac{\partial \mathbf{u}^r}{\partial \mathbf{P}_{ij}}(\mathbf{U}) d\mathbf{P}_{ij} - f^r(\mathbf{x}^r) \right)^2 dvdw \tag{34.5}$$

The solution to the stereoDIC problem is thus obtained as the motion $d\mathbf{P}_{ij}$ of the control points $\mathbf{P}_{ij}$ parameterizing the surface. The deformation of the observed surface represents the 3D displacement field between the reference and deformed configurations. The freeform surface is composed of the same parametric space as the original part. For any point belonging to the surface, the 3D displacement is known (Fig. 34.1).
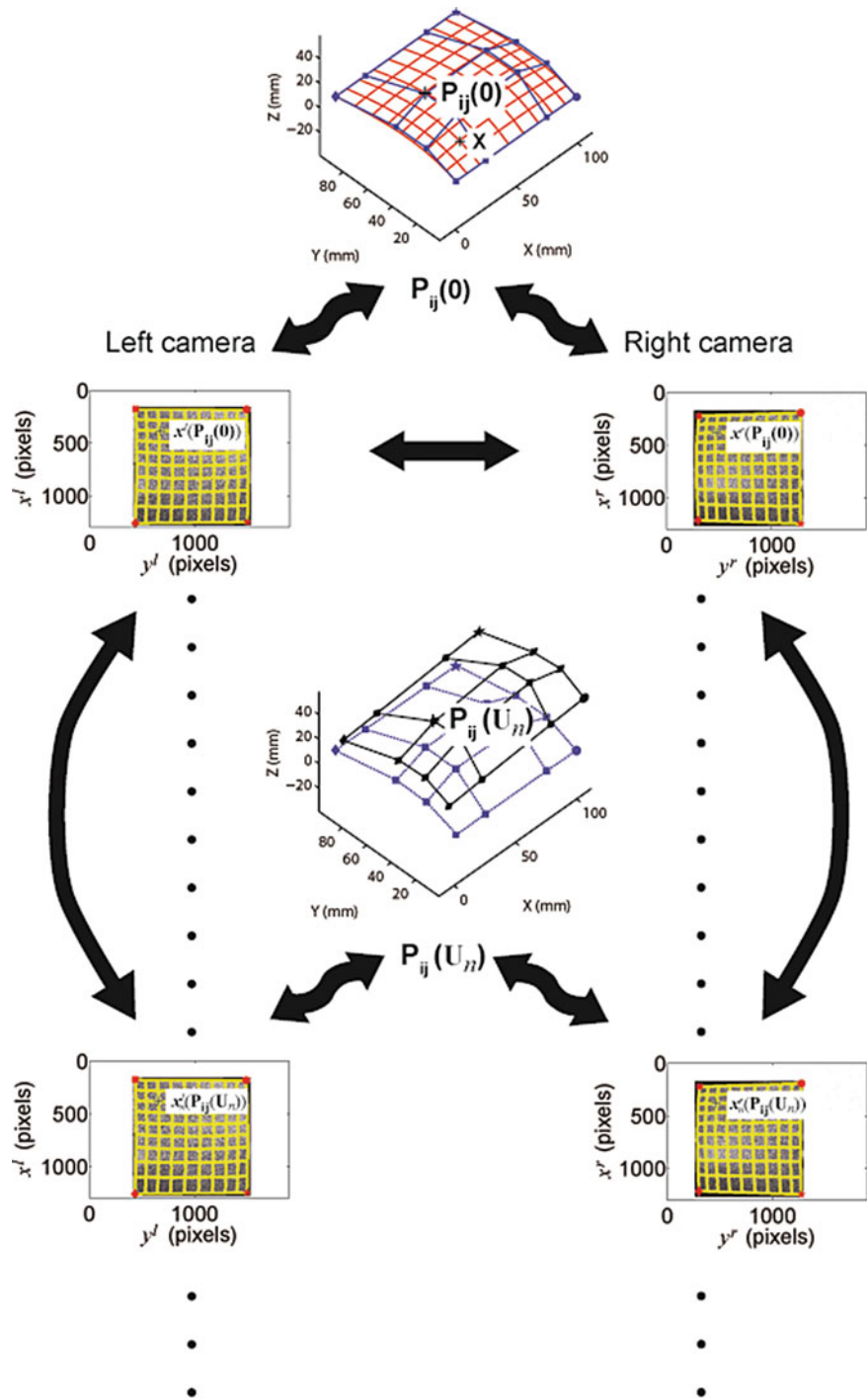
34.4 Example on an Industrial Part

The geometry considered in this section is a 2 m^2 automotive roof (Fig. 34.2). This work focuses on measuring the displacement of the surface induced by the welding brazing process of this part onto the car body.

The calibration of the stereoscopic setup is performed using a self-calibration method [4]. Figure 34.3 shows the theoretical shape used as a first guess during the calibration step.

Each pair of images is processed using the algorithm described earlier. In the present case, 800 image pairs are analyzed to get the NURBS modeling of displacement fields or shape changes. The measured displacement is about 6 mm along the vertical axis (Fig. 34.4), which is consistent with the experimental setup (i.e. the roof is pushed against the car body during
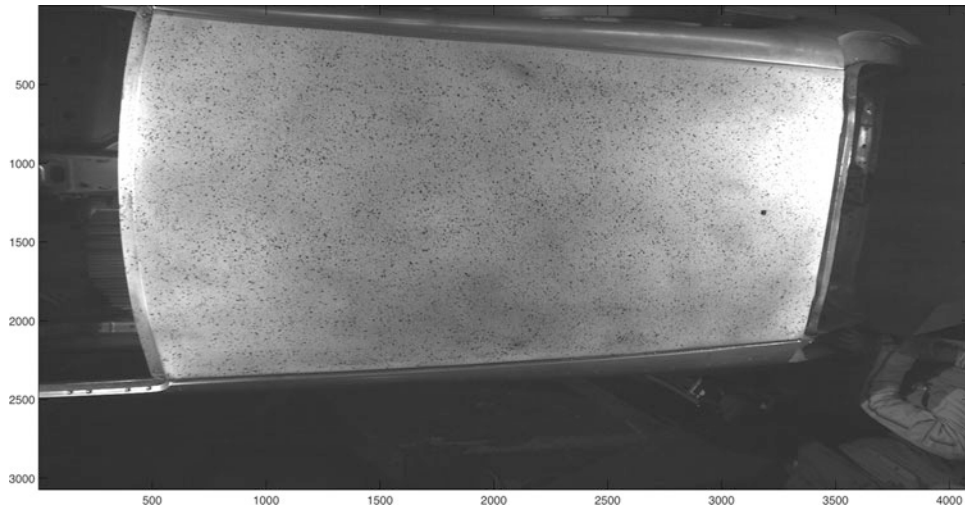
Fig. 34.2 Industrial part used in this study (from the *left point of view*: $f^l(x^l, y^l)$)

Fig. 34.3 Theoretical shape used as first guess (dimensions in mm). 12×7 control points are used to define the considered surface
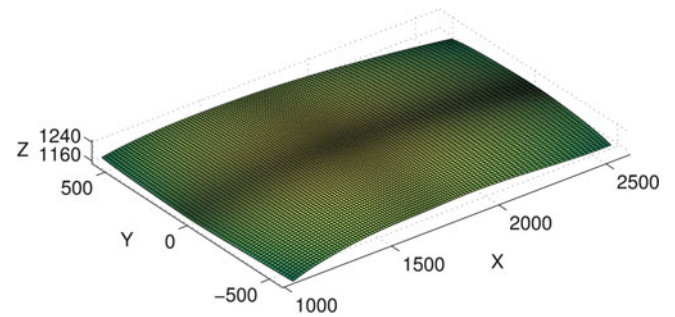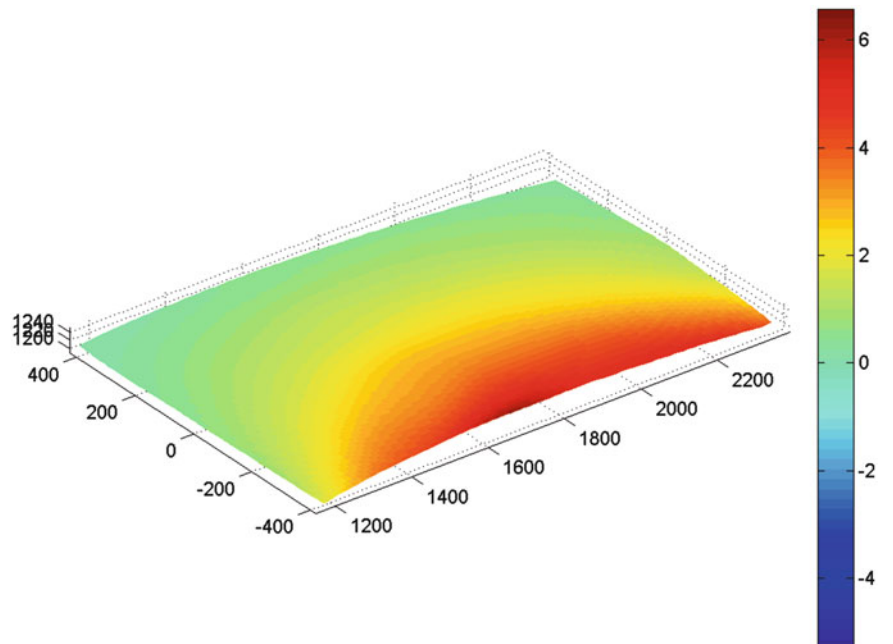

Fig. 34.4 Displacement field U_z measured along the *vertical* axis between the deformed and reference configuration during the welding process
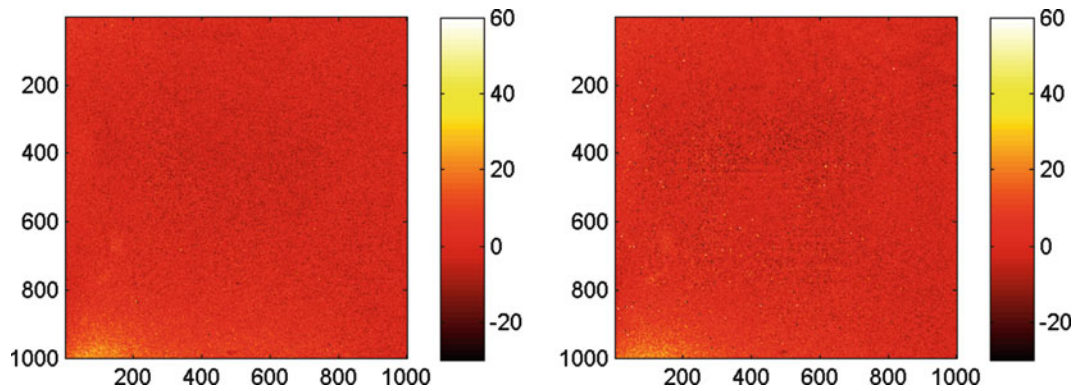
Fig. 34.5 Gray level correlation residuals for the *left* (ρ^l) and *right* (ρ^r) registrations. The RMS level is equal to 2.1 and 2.3 %, respectively

the welding operation). The number of iterations needed to achieve convergence of the algorithm is less than 5 in most cases. Note that the displacement field obtained at a given image pair is used as the initialization for the next pair.

The correlation residuals (Fig. 34.5) show that the left and right registrations are accurate over the whole region of interest. The brighter areas correspond to the position of the welding torch. The root mean square level is of the order of 2 % of the dynamic range of the pictures in the reference configuration. Given the very severe experimental conditions, these values indicate that the results are trustworthy.

34.5   Conclusions

By using an a priori regularized approach to stereoDIC 3D continuous displacement fields are measured and expressed in a NURBS formalism, which is totally consistent with the geometrical model used. This framework can be very useful as it allows direct comparisons to be performed between measurements and computations with a single 'mesh' (e.g. for identification purposes or metrological characterization of the 3D shape).

References

1. Sutton MA, Orteu JJ, Schreier HW (2009) Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications. Springer, New York
2. Garcia D, Orteu JJ, Penazzi L (2002) A combined temporal tracking and stereo-correlation technique for accurate measurement of 3D displacements: application to sheet metal forming. J Mater Process Technol 125–126:736–742
3. Besnard G, Lagrange JM, Hild F, Roux S, Voltz C (2010) Characterization of necking phenomena in high speed experiments by using a single camera. EURASIP J Image Video Process 2010(21595):15
4. Beaubier B, Dufour J-E, Hild F, Roux S, Lavernhe S, Lavernhe-Taillard K (2014) CAD-based calibration and shape measurement with stereoDIC: principle and application on test and industrial parts. Exp Mech. doi:10.1007/s11340-013-9794-6
5. Piegl LA, Tiller W (1995) The NURBS book. Springer, New York
6. Hild F, Roux S (2012) Digital image correlation. In: Rastogi P, Hack E (eds) Optical methods for solid mechanics. Wiley, Weinheim, pp 183–228, Chap. 5

Chapter 35
Single-Camera-Based 3D DIC for Fast-Speed Measurement

Hien Kieu, Zhaoyang Wang, Hieu Nguyen, and Minh Le

Abstract The 3D digital image correlation (DIC) is a very popular method for 3D shape and deformation measurements in the optics and mechanics fields. In general, the 3D DIC technique requires two calibrated cameras to simultaneously capture images of object from two different views. In practice, this can be challenging for fast- or high-speed measurements because of the cost and synchronization issues. In this paper, we demonstrate that the 3D DIC can be implemented by using a single camera with a mirror-based approach. Experiments are presented to verify the practicability of the proposed scheme.

Keywords Digital image correlation • 3D reconstruction • High-speed measurement • High-speed camera • Calibration

35.1 Introduction

The 3D reconstruction of objects is among the most interesting topics in a few fields and has a wide range of applications. In recent years, the 3D digital image correlation (3D DIC) technique has been widely used in experimental mechanics to measure 3D shape, deformation and motion with high accuracy from two images [1]. The 3D DIC measurement typically requires the following key components: speckle or texture patterns on object surface, two synchronized cameras, accurate camera intrinsic and extrinsic parameters, and pixel-by-pixel full-field image matching [2]. As technologies evolve, it is usually easy to obtain these components. However, for high-speed measurements, it can be economically challenging to have two high-speed cameras ready, which has prevented the technique from being used by small businesses and individuals.

In this paper, we present a technique for high speed 3D shape measurement by 3D DIC using only one high speed camera. Four mirrors are used to create the required two views in a single image from the high-speed camera. The captured image contains two different views of the interest object, allowing stereo capturing with a frame rate of the high-speed camera. Experiments have been carried out to prove the validity of the technique in terms of speed and accuracy.

35.2 Principle

Traditionally, two high speed cameras are calibrated and used to capture the target object from two different views and the images are then analyzed by using 3D DIC calculation to produce 3D shape results. To significantly reduce the measurement's cost, we propose a system that uses only a single high-speed camera for the fast-speed 3D shape and deformation measurements. The system setup shown in Fig. 35.1 helps illustrate the principle. Two mirrors 3 and 4 reflect

H. Kieu (✉) • H. Nguyen • M. Le
Department of Electrical Engineering, The Catholic University of America, 620 Michigan Ave. N.E.,
Washington, DC 20064, USA
e-mail: 81kieu@cardinalmail.cua.edu

Z. Wang
Department of Mechanical Engineering, The Catholic University of America, 620 Michigan Ave. N.E.,
Washington, DC 20064, USA

Fig. 35.1 The system setup for high-speed 3D measurement using a single camera. 1, 2, 3, 4 are mirrors
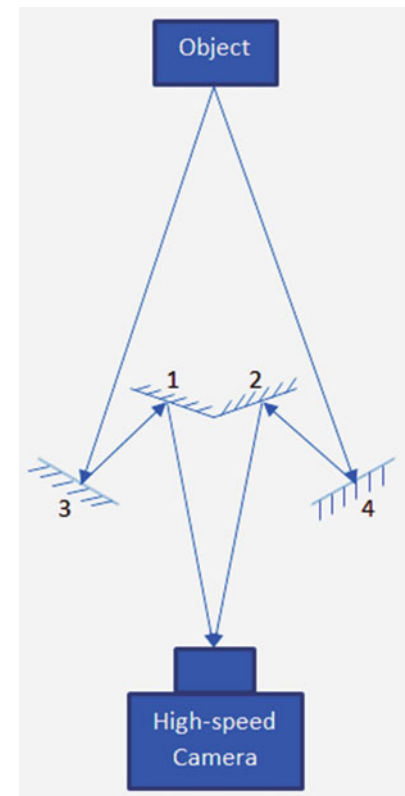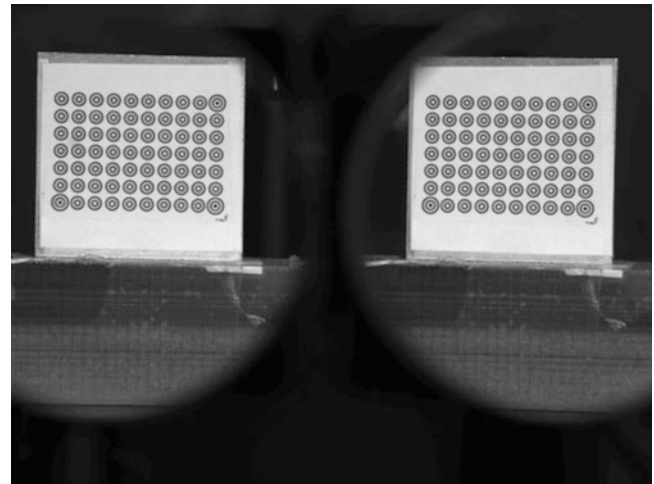

Fig. 35.2 One image from the set of images for calibration

light from object to mirrors 1 and 2 which then reflect light to the camera. 3D DIC analysis will be carried out on the two views in the captured image as if they are captured from two separate high-speed cameras.

Images captured from the high-speed camera always have two different views of the object which can be thought to be captured from two different virtual cameras. Like the general 3D DIC in which the stereo cameras must be calibrated before 3D measurements can be conducted from stereo matches, our virtual cameras also need to be calibrated. Figure 35.2 shows an example of our calibration image where the captured image contains two views of the calibration board obtained from two mirrors. The calibration images are treated like pairs of images captured from two virtual high-speed cameras, and two sets of the camera intrinsic and extrinsic parameters are then obtained.

After calibrating the two virtual cameras by using a set of images of the calibration board at different poses [3], the system is then ready for 3D shape measurement. Technically, the 3D DIC analysis involves 2D DIC processing and triangular geometry using the aforementioned cameras information. 2D DIC is a technique that finds the displacement of every pixel in the interested region of the object at different stages or different views [4]. In our analysis, the left view from the image is chosen to be the reference view and the right view is chosen to be the target view. 2D DIC can track the pixel of interest in the reference view by choosing a square region $(2M + 1) \times (2M + 1)$ called subset around that pixel and finding a corresponding region which consists of corresponding pixels of those in the reference subset that minimizes the correlation coefficient expressed as:

$$C = \sum_{i=-M}^{M} \sum_{j=-M}^{M} \left[af(x_i, y_i) + b - f\left(x_i', y_i'\right) \right]^2 \tag{35.1}$$

where a and b are the scale factor and the offset of the intensity, respectively, $f(x_i, y_i)$ and $f(x_i', y_i')$ indicate the intensity values at a representative pixel in the reference view and the potential matching pixel in the target view, respectively. There are other correlation criteria that can be used depending on the quality of the images and how the object is captured [5]. The coordinates of the pixel (x_i', y_i') in the target view is expressed as a transformation function of the coordinates of its corresponding pixel in the reference view:

$$\begin{cases} x_i' = x_i + \xi + \xi_x \Delta_x + \xi + \xi_y \Delta_y + \xi_{xx} \Delta_x^2 + \xi_{yy} \Delta_y^2 + \xi_{xy} \Delta_x \Delta_y \\ y_i' = y_i + \eta + \eta_x \Delta_x + \eta + \eta_y \Delta_y + \eta_{xx} \Delta_x^2 + \eta_{yy} \Delta_y^2 + \eta_{xy} \Delta_x \Delta_y \end{cases} \tag{35.2}$$

where i and j range from $-M$ to M, $\Delta_x = x_i - x_0$, $\Delta_y = y_i - y_0$, and $\xi, \xi_x, \xi_y, \xi_{xx}, \xi_{yy}, \xi_{xy}, \eta, \eta_x, \eta_y, \eta_{xx}, \eta_{yy}$, and η_{xy} are the transformation parameters. The solutions of $\xi, \xi_x, \xi_y, \xi_{xx}, \xi_{yy}, \xi_{xy}, \eta, \eta_x, \eta_y, \eta_{xx}, \eta_{yy}, \eta_{xy}, a$, and b, for the minimum value of correlation coefficient C can be found by an iterative method such as the Levenberg-Madquardt algorithm [6]. The solution will provide the matches of most or all the pixels in the region of interest between the reference and target views, which can help reconstruct the 3D shape of the object. Most of the times, DIC analysis will not interrogate every interested pixel. Instead, pixels on a grid with a constant step size are selected. Then an interpolation scheme is carried out to find matching for all the remaining pixels [7]. This can save much of the computational cost.

Since the system can capture the object in high-speed for 3D measurement, it is desired that the DIC analysis can operate in a fast enough manner for real-time measurement, which can be accomplished by properly decreasing subset size, increasing step size and calculation step size inside the subset [8].

35.3 Experiment

A few experiments have been conducted to demonstrate the validity of the technique. Here, one of the experiments, the 3D surface shape measurement of a cubic object, is shown. Speckle patterns have been prepared on the object surface for matching purpose. Figure 35.3 shows the image captured from the system and it gives two different views of the front surface of the object. These two views are analyzed by 2D-DIC algorithm to find the matches in the speckle region. As mentioned in Sect. 35.2, the DIC analysis for the speckle-pattern region is accomplished based on a good initial guess, which is in this experiment, is automatically found by Scale-Invariant Feature Transform (SIFT) [9]. This helps the initial guess selection become less tedious [10]. The matching information is then integrated with the camera intrinsic and extrinsic information to yield the 3D shape of the surface, which is shown in Fig. 35.4.

Another experiment is carried out on a ping pong ball which is shown in Fig. 35.5. The ball is captured by the camera and its 3D image is reconstructed as shown in Fig. 35.6 thanks to the two different views provided by the mirrors. Noted that no speckle pattern was projected on the ball and the DIC analysis uses natural pattern of the ball for matching [11].

It should be noted that the accuracy verification test and more high-speed experimental results will be presented by the authors during the conference.

Fig. 35.3 A cubic object with speckle patterns on its front surface has its two different views captured in a single image
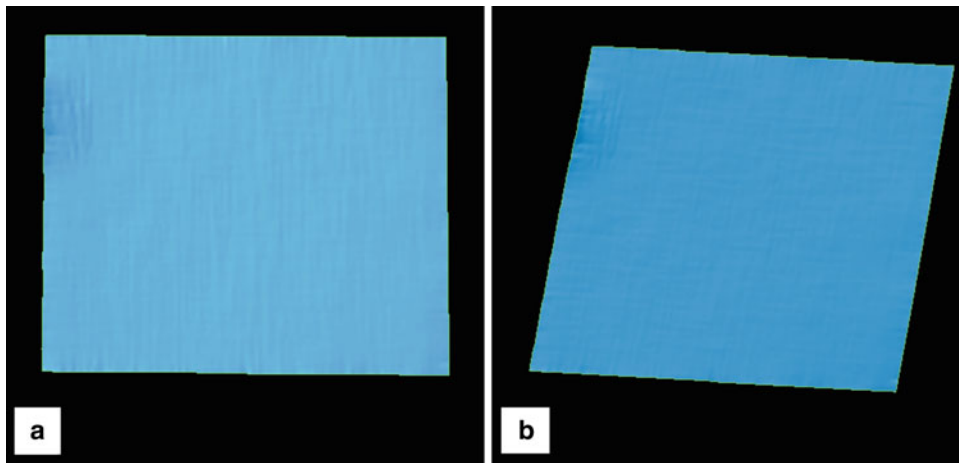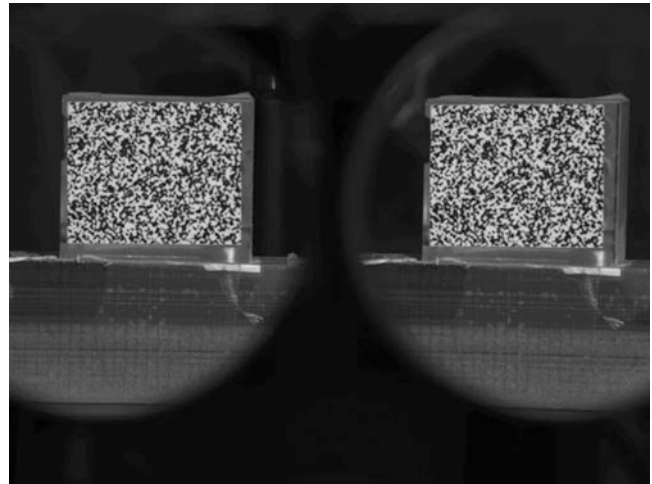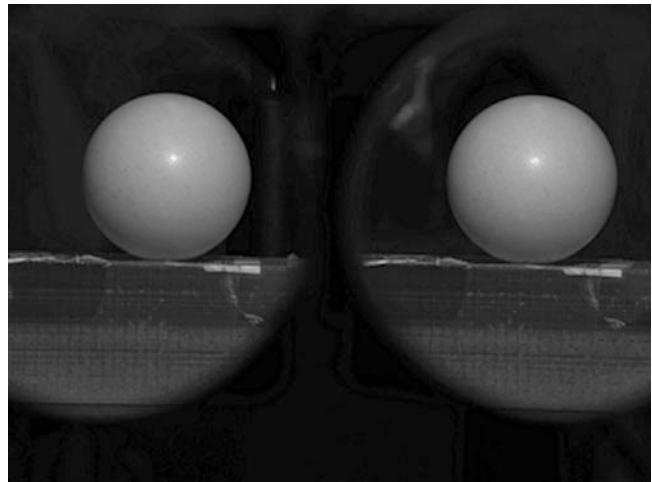

Fig. 35.4 The 3D shape reconstruction of the object front surface; (**a**) and (**b**) are two representative plots

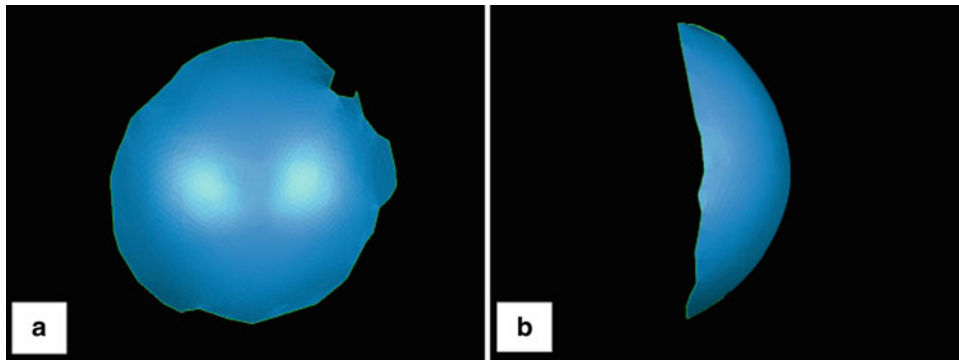**Fig. 35.5** A ping pong ball being captured by the system with one high-speed camera

Fig. 35.6 3D surface of the ping pong ball reconstructed from the image in Fig. 35.5; (**a**) and (**b**) are the front view and the side view of the ball, respectively

35.4   Conclusion

As technology evolves to enhance our life as well as to lower the resources required, it is preferable to use less materials to accomplish any specific task without compromising the performance and speed of the current techniques.

This paper presents a method that uses only a single high-speed camera to carry out high-speed 3D shape and deformation measurements. This has considerably reduced the expensive cost of building a high-speed 3D DIC system. The proposed method is easy and simple to implement, and the accuracy and practicability have been verified by experiments. It is expected that the proposed method will help broaden the applications of high-speed 3D DIC technique to more scientific and engineering applications.

References

1. Hu Z, Xie H, Lu J, Wang H, Zhu J (2011) Error evaluation technique for three-dimensional digital image correlation. Appl Opt 50:6239–6247
2. Pan B, Xie H, Yang L, Wang Z (2009) Accurate measurement of satellite antenna surface using 3D digital image correlation technique. Strain 45(2):194–200
3. Vo M, Wang Z, Luu L, Ma J (2011) Advanced geometric camera calibration for machine vision. Opt Eng 50:110503
4. Pan B, Qian K, Xie H, Asundi A (2009) Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. Meas Sci Technol 20:062001
5. Pan B, Xie H, Wang Z (2010) Equivalence of digital image correlation criteria for pattern matching. Appl Opt 49:5501–5509
6. Lourakis MIA, Argyros AA (2005) Is Levenberg-Marquardt the most efficient optimization algorithm for implementing bundle adjustment? In: Tenth IEEE international conference on computer vision (ICCV), vol 2, Beijing
7. Luu L, Wang Z, Vo M, Hoang T, Ma J (2010) Accuracy enhancement of digital image correlation with B-spline interpolation. Opt Lett 36(16):3070–3072
8. Wang Z, Hoang T, Nguyen D, Urcinas A, Magro J (2010) High-speed digital image correlation method: comment. Opt Lett 35:17–2891
9. Lowe D (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110
10. Wang Z, Vo M, Kieu H, Pan T (2014) Automated fast initial guess in digital image correlation. Strain 50(1):28–36
11. Kieu H, Pan T, Wang Z, Le M, Nguyen H, Vo M (2014) Accurate 3D shape measurement of multiple separate objects with stereo vision. Measur Sci Technol 25(3):035401

Chapter 36
Three-Dimensional Digital Image Correlation Using a Single Color-Camera

Wade Gubbels and Gary S. Schajer

Abstract Three-dimensional Digital Image Correlation is typically done using two cameras that view the measured object from differing oblique directions. The measured images are independent and must be spatially connected using a detailed calibration procedure. This places a large demand on the practitioner, the optical equipment and the computational method. A novel approach is presented here where a single color-camera is used in place of multiple monochrome cameras. The color-camera measures three independent Red-Green-Blue (RGB) color-coded images. This feature greatly reduces the scale of the required system calibrations and spatial computations because the color images are physically aligned on the camera sensor. Out-of-plane information is obtained by triangulation using oblique illumination from separate colored light sources. The 3-D Digital Image Correlation is completed by combining the 2-D correlations for each RGB color. An example apparatus is described and some typical results are presented.

Keywords 3-D DIC • Color DIC • Diffraction grating • Telecentric lens

36.1 Introduction

Digital Image Correlation (DIC) is a full-field non-contacting optical method used to measure two-dimensional (2-D) and three-dimensional (3-D) surface displacements and strains. Traditionally in DIC, a high contrast random pattern or speckle is applied to the object surface. Image matching and tracking techniques are used to compare the grayscale images of the speckled surface taken before and after deformation. 2-D DIC is realized using a single fixed camera, commonly oriented with the viewing axis normal to the object surface. Because of this orientation, the method has minimal sensitivity to out-of-plane surface displacements. The different correlation algorithms used to measure the in-plane displacements are conceptually straightforward; Pan [1] gives a comprehensive review of commonly used 2-D correlation techniques.

Currently, a dual-camera or stereoscopic arrangement is employed to measure 3-D displacements simultaneously. The two cameras are arranged so that they view the test specimen surface from different oblique angles. Before accurate measurements can be made a detailed calibration procedure is required to cross-reference the viewing parameters of each camera and their position and orientation relative to one another and the test specimen [2]. The required accuracy of this calibration places a much greater demand on the precision and resolution of the cameras used in the stereoscopic arrangement than with a single-camera system. Once the proper calibration is completed the image sets are then analyzed using sophisticated 3-D correlation algorithms [3, 4].

The 2-D single-camera method requires relatively modest calibration procedures, so it would certainly be a great advantage if it were possible to measure 3-D displacements using a single camera. Unfortunately, the conventional arrangement of using a single monochrome camera does not provide sufficient image information. Some work has been done recently to adapt the traditional single-camera method to obtain more information from grayscale images. Xia [5] developed a novel diffraction assisted technique, but the demonstrated measurement sensitivity for out-of-plane displacements was relatively modest. Tang [6] incorporated a photogrammetry method by offsetting the camera axis and Sun [7] extracted the out-of-plane information from the change in magnification due to perspective errors from the camera

W. Gubbels • G.S. Schajer (✉)

Department of Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada
e-mail: schajer@mech.ubc.ca

lens. While these last two methods show accurate out-of-plane results, each requires a detailed calibration procedure in which multiple camera locations or known displacements are required. These recently developed techniques all aim to extract extra information from a grayscale image taken from a single camera. A different approach can be taken where additional information can be obtained by changing from a monochrome to a color-camera. This single color-camera approach has recently been used in hybrid fringe projection and DIC methods [8, 9]. In these cases the multiple color signals available in a color-camera provide separate fringe pattern measurements, used to indicate out-of-plane displacements [10], from a surface speckle, used to identify in-plane DIC measurements. These hybrid methods are capable of measuring 3-D displacements, but calibration is required for accurate fringe projection measurements. Here, the available information from a color-camera is used to separate multiple speckle patterns which are used to obtain 3-D displacements with modest calibration.

A typical color-camera sensor is constructed with an interlaced pattern of Red, Green and Blue (RGB) pixels. The most commonly used colored pixel arrangement was introduced by Bayer [11]. The three color signals are recorded individually on the sensor and, in conventional use, are combined to create full-color images. However, the three color signals are independent and just one color, say green, is sufficient to do 2-D DIC measurements. This leaves the remaining two color signals, say red and blue, available to measure the out-of-plane displacements. Such measurements can be done using triangulation by projecting a random speckle pattern of either red or blue on to the surface of the object. Then a further 2-D DIC evaluation combined with triangulation can identify the out-of-plane surface displacements. This procedure has the advantage that all of the various colored pixels have precisely known locations relative to each other, so it is not necessary to do extensive calibrations to cross-reference the pixels in two separate cameras, as in the conventional 3-D technique. This paper describes an optical system using a single color-camera with a structured color illumination arrangement that was successfully developed to achieve 3-D displacement measurements of a planar specimen.

36.2 Triangulation

In traditional 2-D DIC a fixed camera views a planar object whose surface is perpendicular to the viewing (z) direction. If a speckle pattern is painted on the object then 2-D DIC can be used to identify the local in-plane (x and y) displacements from the measured images. There is no sensitivity to out-of-plane (z) displacements, except that due to change in viewing perspective. For out-of-plane displacements, the triangulation arrangement shown in Fig. 36.1 can be used instead. In this arrangement, the object surface is left unpainted. Instead, a speckle pattern is projected onto the surface using a projector system inclined at angle θ to the z direction in the x-z plane. This arrangement has opposite properties compared with the traditional one. In-plane (x and y) displacements of the object do not change the position of the projected speckle observed by the camera. However, a z-displacement, dz, causes an x-displacement, dx, of the projected speckle that can be identified by 2-D DIC from the image measured by the camera. Triangulation gives

$$dz = dx/\tan\theta \tag{36.1}$$

The proposed single-camera 3-D DIC system combines these two measurements and does them simultaneously using the separate color signals from a color camera. A painted speckle pattern on the object surface is illuminated in one color, say green. This allows the x and y displacements to be identified by doing DIC using the green signal from the camera. Simultaneously, a second speckle pattern is projected at angle θ on the object surface in a second color, say blue. This second arrangement allows the z displacement to be identified by doing DIC using the blue signal from the camera and Eq. 36.1. In this basic arrangement, the third camera color, red, is unused.
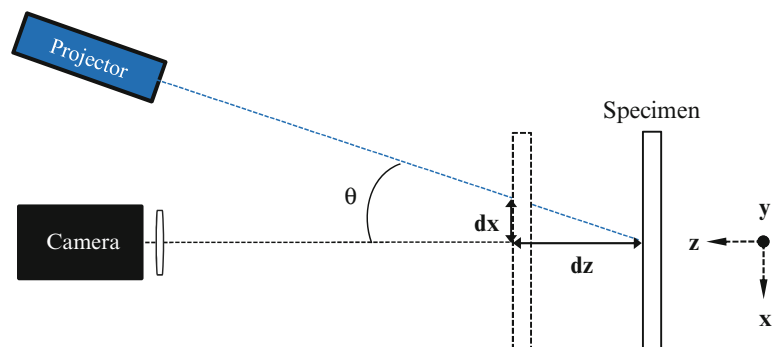

Fig. 36.1 Plan view of triangulation setup for measuring out-of-plane displacements

The basic arrangement illustrated in Fig. 36.1 can provide accurate displacement measurements, but several practical design issues need to be addressed. These include:

1. The diverging image configuration typically used by projector systems causes change in the in-plane size of the projected speckle pattern. This causes perspective changes that cause local displacements of the projected image in addition to those caused by triangulation, thus introducing geometrical error.
2. A similar perspective error occurs with the camera imaging system. The typically diverging view of a camera causes apparent x and y displacements to occur when the object distance z is significantly changed.
3. The object surface is oriented obliquely to the speckle pattern projector, so differing areas on the object can be at significantly different distances from the projection lens. This can make it difficult to focus the projected speckle pattern accurately over all the illuminated area.
4. The color signals measured by the camera must measure each illumination color independently; else cross-talk will occur.
5. The painted and projected speckle patterns must be independent and uncorrelated; else cross-talk will also occur.
6. For Eq. 36.1 to apply, the object surface must everywhere be perpendicular to the viewing (z) direction.

The first five of these issues will be addressed here in regards to a planar specimen. The sixth requires the use of additional color signals and will be addressed in a future report.

36.3   Speckle Projector Design

A custom telecentric projector was created to overcome design issue 1, the varying in-plane image size created by a conventional projector. The telecentric arrangement creates an essentially parallel projection [12]. Figure 36.2 shows a ray diagram for the telecentric projector. A laser source is directed onto a diffusion plate that expands the light into a conical beam. This beam is collimated by the first lens, which directs the light though a transparent film on which a random speckle pattern had been applied. This transparency forms the object of the double telecentric lens formed by the following two lenses and intermediate aperture. The near parallel rays emerging from the final imaging lens project an image that remains constant in size over a significant image distance range known as the telecentric depth. The designed projector creates a telecentric depth (=depth of field) of approximately 5 cm allowing for significant z displacements to be measured using triangulation.

The telecentric projector enables the projected speckle pattern to maintain the same size during out-of-plane motion of the measured surface. However, perspective errors will be apparent in the images captured by a camera with a conventional lens (design issue 2). This can be handled in two possible ways. One possibility is to use a telecentric lens on the camera as well. This arrangement requires that the aperture of the telecentric lens be at least as large as the measured object, thus constraining the measured object to be of fairly modest size. An alternative approach is to use a conventional camera lens and correct the apparent in-plane displacements using knowledge of the out-of-plane displacements found from triangulation
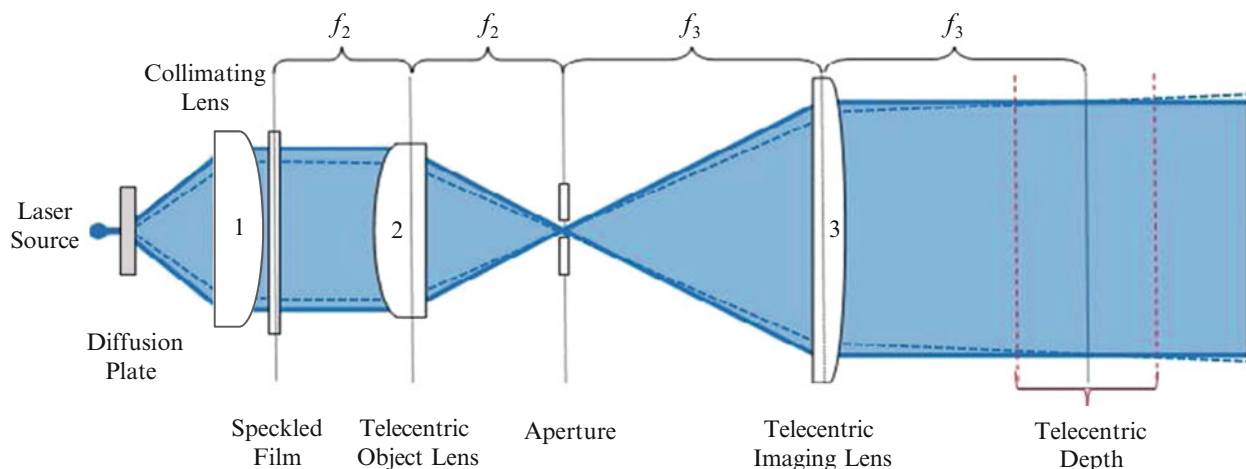

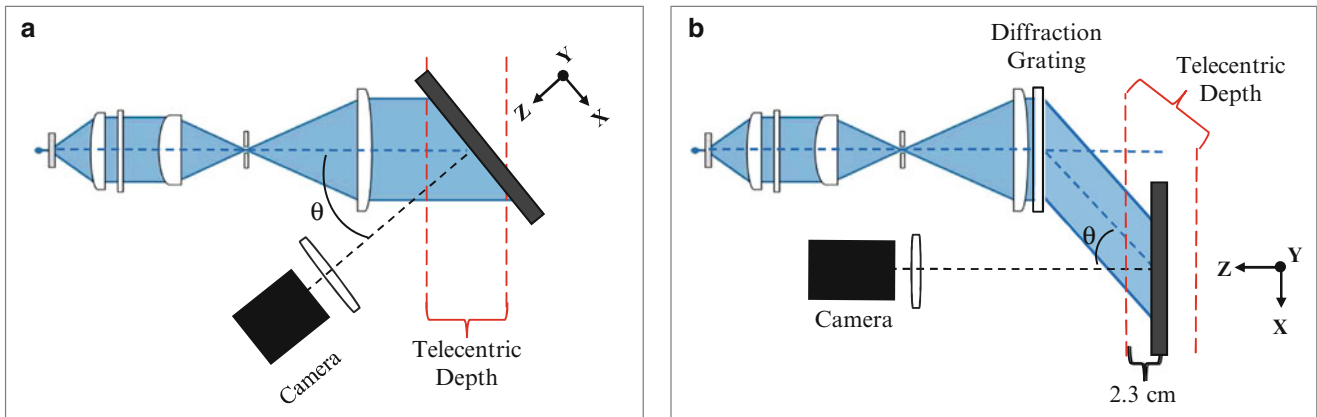**Fig. 36.2** Telecentric projector ray diagram

Fig. 36.3 Oblique projection. (**a**) direct projection, (**b**) projection through a diffraction grating

of the projected speckle pattern. The perspective errors due to the camera lens should be equal in x and y, for both the painted-on and projected speckle patterns. Triangulation of the projected speckle pattern creates only x-displacements, so, any apparent stretching in the y direction in the projected image is due to perspective errors and can be subtracted from each image to restore the original image scale.

For triangulation to be used, the projection system must illuminate the surface at an angle relative to the camera viewing (z) axis. One possibility is to orient the projector as shown in Fig. 36.3a, but this arrangement creates a greatly varying distance between imaging lens and object. This makes focusing the entire projected image difficult. The available telecentric depth-of-field may or may not be sufficient for the task (design issue 3). Even if the telecentric depth-of-field is sufficient, likely not much field depth will remain to accommodate object motion.

Variable distance between imaging lens and specimen can be eliminated by using a linear diffraction grating, as shown in Fig. 36.3b. Light passing through the diffraction grating diffracts through an angle

$$\theta = \sin^{-1}\left(\frac{\lambda}{d}\right) \tag{36.2}$$

where λ is the wavelength of the light used and d is the line spacing of the diffraction grating [13]. The outgoing diffracted light creates a parallelogram projection. Therefore, aligning the camera and projection arrangement parallel to one another insures that the optical path lengths of the projection remain equal across the specimen surface. This arrangement fixes the incidence angle θ to be the diffraction angle defined in Eq. 36.2.

A $\lambda = 405$ nm blue laser and a diffraction grating with a line spacing d = 1,000 nm are used in the projection arrangement. Using these values in Eq. 36.2, the projection angle is 23.9°. This particular arrangement therefore has the capability of measuring z displacements within a range ±2.3 cm around the center of the telecentric depth.

To summarize, design issues 1–3 have been addressed by developing a custom projector using a telecentric lens with a linear diffraction grating. The customized system creates projected blue patterns tailored for accurate and economic calculation of the out-of-plane displacements but, before in-plane displacements, obtained from the green signal, can be made simultaneously, steps must be taken to ensure the two color signals can be independently separated from a full-color image.

36.3.1   Color Separation

For successful multi-color measurements it is important that each color is measured independently, with minimal cross-talk from other colors. The color-camera used here (Prosilica GC 2450, Allied Vision Technologies, Burnaby, Canada) uses a Bayer style mask to filter the light reaching the pixels on the camera sensor [14]. The filters are arranged in a repeating pattern of Red-Green-Blue (RGB) pixels. To allow realistic rendition of color scenes where a wide range of wavelengths are present, the filters are designed to have moderately wide bandwidths that overlap each other. However, these overlaps are undesirable here because they cause crosstalk between the separately colored signals and consequently impair measurement accuracy.

The effect of the color overlaps can be removed by mathematical calibration. For a beam of color 1 (red) only, with intensity I_1 the measured average responses R_1, G_1, B_1 are:

$$R_1 = A_{11}I_1 \quad G_1 = A_{21}I_1 \quad B_1 = A_{31}I_1 \tag{36.3}$$

where A_{ij} are calibration coefficients relating the response of pixel color i and illumination color j. In this measurement, the numerical value of the light intensity I_1 is not explicitly known. However, it is sufficient to work in terms of scaled quantities and to normalize the calibration coefficients A_{ij}:

$$a_{11} = \frac{A_{11}}{\max(A_{11} + A_{21} + A_{31})} \quad a_{21} = \frac{A_{21}}{\max(A_{11} + A_{21} + A_{31})} \quad a_{31} = \frac{A_{31}}{\max(A_{11} + A_{21} + A_{31})} \tag{36.4}$$

Similarly, with beams of colors 2 (green) or 3 (blue) only, subscript "2" or "3" replaces subscript "1" above.

When all three beams are present simultaneously, the measured responses m_i are:

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} R_1 + G_1 + B_1 \\ R_2 + G_2 + B_2 \\ R_3 + G_3 + B_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} \tag{36.5}$$

which in matrix form, becomes:

$$\mathbf{a\,I} = \mathbf{m} \tag{36.6}$$

The individual beam intensities $\mathbf{I}$ can be recovered in scaled form from the measured responses $\mathbf{m}$ by matrix solution. The normalization in Eq. 36.4 ensures that the scaled results have a similar range as the original measurements. This calibration procedure need only be performed once for given illumination sources. In tests using single-color illumination, overlaps between different color pixels were typically 5–15 %. After applying Eq. 36.6, the overlaps reduced to around 1 %. Therefore, the technique effectively resolves design issue 4.

36.4 Laser Speckle Cross-Talk

A laser is a convenient light source for the telecentric projector because it behaves almost as a point source and can therefore be accurately collimated. In addition, the light is monochromatic, so will pass through the diffraction grating at a specific angle θ and so preserve the focus of the projected image. However, a consequence of using highly monochromatic laser light is the occurrence of laser speckle [15]. While this speckle can be useful in experimental methods such as interferometry and holography [16], it is undesirable here because it adds to and mixes with the projected speckle pattern from the transparent film. The projected speckle pattern is fixed in space and the measured specimen moves through that projected speckle pattern to form the observed triangulation that indicates the out-of-plane displacements. However, the laser speckle appears as though it is fixed to the specimen surface. Thus, the image measured in the projected color, blue here, contains both in-plane and out-of-plane components, creating cross-talk and consequent error in the associated DIC results. Fortunately, a simple modification of the projection system can remove the laser speckle. This can be done by rotating the diffusion plate that expands the light from the laser source. The rotating plate randomly moves the speckle pattern created by the laser light, while leaving the projected speckle pattern unaffected. If the motion is greater than the laser speckle size and occurs within the exposure time of the image acquisition, the details of the laser speckles are smeared out and a projected pattern image, free of laser speckle is acquired. For an exposure time of 50 mS, a diffuser plate rotation speed of 0.5 Hz is sufficient.

36.5 Painted Speckle Crosstalk

A further source of cross-talk exists between the fixed (painted-on) and projected speckle patterns. Of necessity, the speckles in the painted-on pattern must have a distinctly different light intensity compared with that of the background surface so as to create sufficient contrast for effective functioning of the DIC. However, this large contrast difference imprints the painted
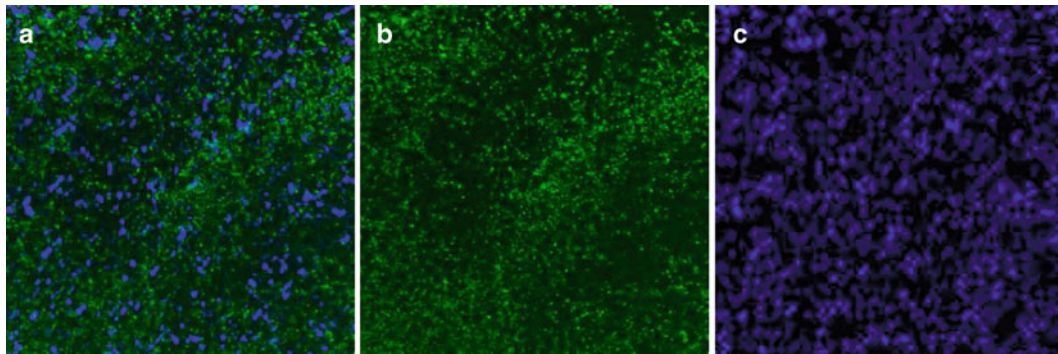
Fig. 36.4 Color image separation. (**a**) *full-color image*, (**b**) separated *green image*, (**c**) separated *blue image* (This figure is shown in color in the electronic version of this paper)

speckle pattern onto the projected speckle pattern image, and significantly impedes the functioning of the DIC on the projected speckle pattern. Mathematical techniques were attempted to try to remove the common portions of the painted-on (green) and projected (blue) images from the measured blue image but the procedure was not adequately effective. Therefore, an alternative approach was pursued.

The chosen approach uses the very small spherical glass beads that are commonly used in retro-reflective paint. These glass beads have the property of reflecting incident light back towards its source. If the green light source that illuminates the surface is placed nearly in-line with the camera viewing axis, the light that strikes the beads reflects back to the camera sensor and appears as a green speckle. However, the blue light, which is projected onto the surface at an oblique angle to the viewing axis, reflects back in the blue illumination direction and therefore is invisible to the camera.

It turns out that application of the retro-reflective beads can be done very conveniently. The glass beads, which range in diameter from 50 to 100 μm, are simply sprinkled onto the surface while a newly applied thin layer of white paint is still wet. The sprinkled beads naturally form a random pattern that is very suitable for DIC work. The high reflectivity of the beads in the illumination direction causes a large fraction of the incident light to return to the camera, thus allowing the use of a light source of only modest power. Since the off-axis reflectivity of the beads is similar to the diffuse reflection from the matte white painted surface, the bead pattern does not significantly appear within the projected speckle pattern observed by the camera.

Figure 36.4 shows the separation of a single full-color image, taken using the retro-reflective bead technique, into independent blue and green images using the color separation of Eq. 36.6. (This figure is shown in color in the electronic version of this paper). The separated green image in Fig. 36.4b shows that the applied surface pattern from the retro-reflective beads is clear and has minimal cross-talk from the projected speckle image shown in Fig. 36.4c. Similarly, the separated blue image in Fig. 36.4c shows that the projected speckle pattern is clear and has minimal cross-talk with the applied retro-reflective surface pattern in Fig. 36.4b. These images show that by creating the surface pattern with retro-reflective beads the cross-talk between the painted speckle and projected speckle images is removed, thereby resolving design issue 5. Thus all design issues 1–5 have been resolved within the custom projection system shown in Fig. 36.3b. The completed speckle projector system was then tested to evaluate the effectiveness of the proposed 2-D color measurement system for identifying 3-D displacements.

36.6 Experimental Setup, Procedure and Results

Figure 36.5 shows the optical arrangement used for testing and evaluation. The test specimen is a flat steel plate with white painted surface and attached retro-reflective beads. The specimen is mounted to a 3-axis translation stage and positioned with the painted surface normal to the camera viewing (z) axis. The camera is fixed 107 cm from the specimen and a 75 mm zoom lens is used to image the surface. A 1 W, white LED is positioned slightly above the camera viewing axis and the light is passed through a green gel filter to illuminate the surface. The telecentric projector consisting of: the laser source, rotating diffusion plate, collimating lens, speckled transparent film, telecentric lenses with intermediate

Fig. 36.5 Experimental setup
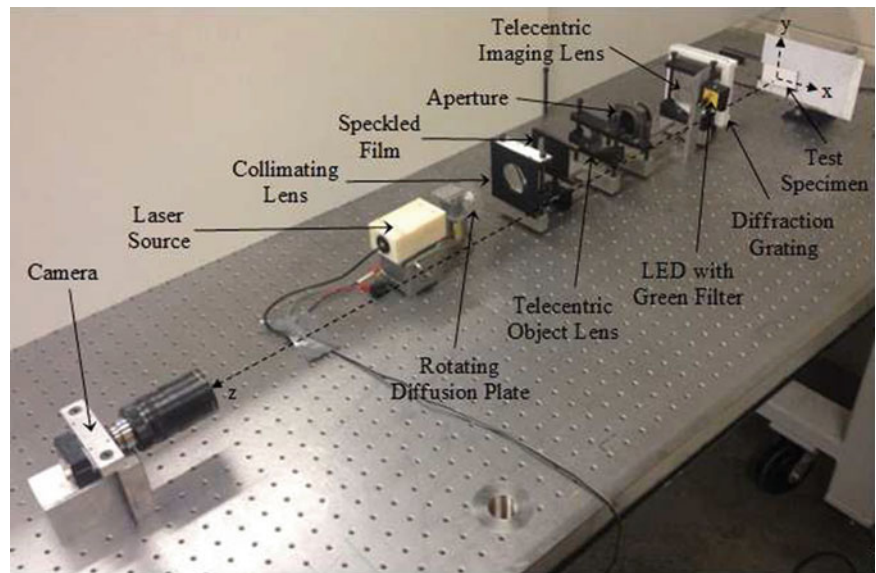
aperture, and linear diffraction grating, is fixed in the same plane and parallel to the camera viewing direction. The projection passes through the linear diffraction grating and is refracted on to the specimen surface to be viewed by the camera.

Two experiments were conducted to demonstrate the capability to measure in-plane and out-of-plane displacements simultaneously. Each experiment includes ten incremental displacements of 100 μm. The specimen displacements were made using a precision linear actuator (CMA-25CCL Newport). A single reference image was taken in each experiment before any displacements were made. A "displaced" image was taken after each 100 μm displacement. The images were acquired, separated by color signal, and analyzed using custom software.

The measured images contained 2,448 × 2,050 pixels, from which a central region of interest of 500 × 500 pixels was extracted. Each reference image was divided into a 5 × 5 array of subsets centered at the midpoint of the region of interest. Each subset size was 50 × 50 pixels and spaced such that there was no overlap between subsets. The magnification of the camera was set so that the region of interest was 25 mm^2. At this magnification, each image pixel represents an area on the specimen surface of approximately 50 μm^2. 2-D DIC analysis was performed between each "displaced" image and the initial reference image, therefore measuring total displacement rather than incremental.

The first experiment consists of ten incremental displacements of 100 μm in the positive x direction. The results in Fig. 36.6a and b shows the mean displacement and standard deviations of the displacements at each increment determined by DIC for the 25 measured subsets. The mean measured x-displacements have a maximum absolute error of 4 μm (0.08 pixels) for the first 8 increments, while the final two increments show a somewhat higher absolute error of 10 μm (≈0.20 pixels). The standard deviations among the 25 measured areas are less than 3 μm (≈0.06 pixels) in all ten increments in the x direction. As expected, the measured y and z displacements all cluster around zero. The standard deviations of the y-displacements are also less than 3 μm in all cases, while the z-displacements show a slightly higher maximum standard deviation of 7 μm (≈0.14 pixels).

In the second experiment, the 10 displacement increments are made in the positive z direction. Figure 36.6c and d shows the results for the mean and standard deviations of the DIC measured displacements for each increment. The x and y displacements are now expected to be zero and the measured values closely agree. The standard deviations of the x and y results are less than 4 μm (≈0.08 pixels) in all cases and show an approximately linear increase. The out-of-plane (z) displacements, obtained from the blue signal, also show high accuracy and precision in the second experiment. The absolute error follows a linear trend throughout the measurement set, reaching a low of less than 5 μm (≈0.10 pixels) at the fourth increment and a high of 30 μm (≈0.60 pixels) at the final increment. The standard deviations of the mean z-displacements are relatively constant however and are less than 8 μm (≈0.16 pixels) for all increments.

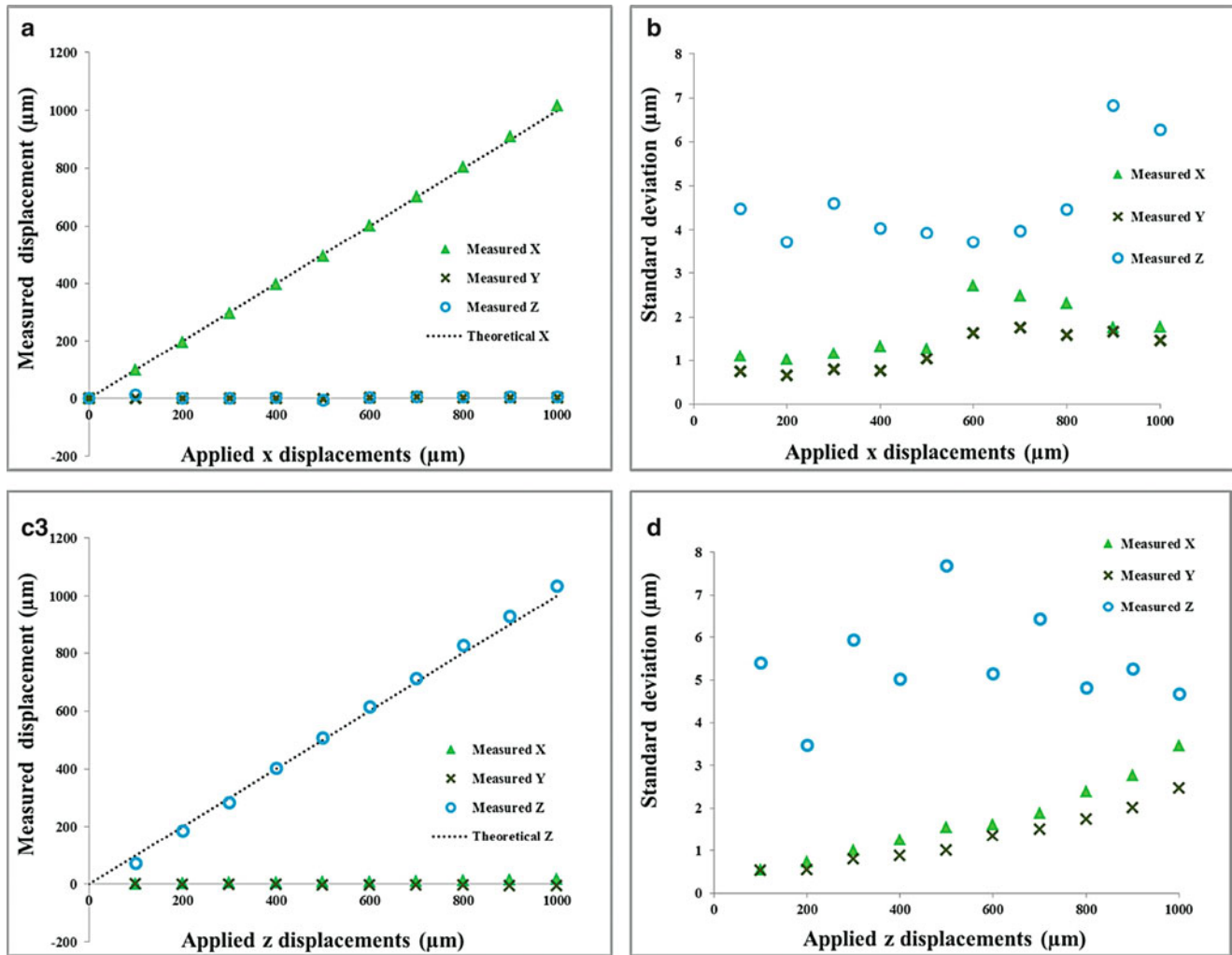**Fig. 36.6** Displacement measurement accuracy tests. (**a**) mean measurements for applied x-displacements, (**b**) standard deviations for applied x-displacements, (**c**) mean measurements for applied z-displacements, (**d**) standard deviations for applied z-displacements

36.7 Discussion

The presented results confirm that the proposed single-camera system can accurately measure 3-D displacements. While still relatively low, the standard deviations for out-of-plane (blue) measurements are consistently twice those for in-plane (green) measurement, for all increments in both experiments. One possible explanation for this occurrence is the pixel arrangement of a Bayer color sensor, which contains 50 % green pixels and 25 % each of red and blue pixels. Thus, the lower density of blue pixels may be expected to result in reduced DIC accuracy. This issue applies to all colors because a single color image extracted from a Bayer sensor has a lower spatial resolution compared with a monochrome sensor. However, the high pixel count of modern camera sensors significantly reduces the consequences. If needed, the issue could be avoided entirely by using a color sensor such as the Foveon X3, which has RGB color sensitivity at every pixel [17]. In the apparatus used here, the camera had been placed rather far from the measured object to accommodate a different application. Thus, simply moving the camera closer to the object would here be an effective way of increasing the pixel resolution of the camera.

A second reason for the differing DIC accuracy of the in-plane and out-of-plane responses is the greatly different scales of the projected and painted speckle patterns. The retro-reflective beads create a "painted-on" speckle pattern that has very fine and sharply defined features. In comparison, the projected speckle pattern has a relatively large and more smoothly varying speckle pattern. This is a characteristic of the film transparency used in the speckle projector, which was made by printing on a plastic film using a computer printer. Clearly, more sophisticated methods can be developed for preparing film

transparencies and these will be explored in the future. The objective is to achieve similar spatial scale for the painted-on and projected speckle patterns. With this in place, it is expected that more similar in-plane and out-of-plane measurement accuracies will be achieved.

36.8   Conclusion

A single color-camera system with multiple color light sources has been demonstrated that is able to identify 3-D displacements of a flat surface. In an initial prototype apparatus, in-plane displacements could be identified within 0.08 pixels and out-of-plane displacements within 0.16 pixels. The single color-camera system needs minimal calibration to allow it to work effectively, certainly much less than that required from a comparable 3-D two-camera system. In addition, the demands on the experimentalist and on the optical quality of the equipment are much less. The prototype system shown here was assembled mostly from remnant laboratory items, typically not of too high quality, yet the resulting system worked quite well. Certainly, the achievable results can be expected to improve significantly with more sophisticated optical construction.

The use of retro-reflective glass beads for the "painted-on" speckle pattern gave a sharply defined, high-contrast surface image, both helpful features for good DIC accuracy. In the present application, the retro-reflective character of the beads was also helpful in concentrating the speckle pattern into the desired color image and excluding it from the other color images. The experience here with the retro-reflective glass beads suggests that they may also be useful for conventional DIC work using a monochrome camera. If the measured specimen is illuminated using a light source adjacent to the camera, the retro-reflective action of the glass beads will cause a large fraction of the light to return in the direction of the camera, much larger than would occur by ordinary diffuse reflection. Thus, retro-reflective glass beads would be a good surface decoration choice in low-light or high-speed imaging applications.

The single color-camera system, while highly capable, does have limitations. The amount of out-of-plane displacement that may be measured is limited by the telecentric depth of the projection arrangement. This depth can be increased by using lenses with longer focal lengths, but very large displacements are difficult to accommodate. In addition, for accurate triangulation using just two measurement colors, the specimen must be planar and remain nearly normal to the camera viewing axis during deformation. This reduces the range of measurements to small displacement applications. Work is currently being done to develop a multi-color projection arrangement capable of measuring 3-D displacements of non-planar specimens. This work will be presented in a future report.

References

 1. Pan B, Qian K, Xie H, Asundi A (2009) Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. Meas Sci Technol 20(6):3–9
 2. Orteu J-J (2009) 3-D computer vision in experimental mechanics. Opt Lasers Eng 47(3–4):282–291
 3. Sutton MA, Orteu J-J, Schreier H (2009) Image correlation for shape, motion and deformation measurements. Springer, New York, pp 225–228
 4. Sutton MA, Wolters WJ, Peters WH, Ranson WF, McNeill SR (1983) Determination of displacements using an improved digital correlation method. Image Vis Comput 1(3):133–139
 5. Xia S, Gdoutou A, Ravichandran G (2012) Diffraction assisted image correlation: a novel method for measuring three-dimensional deformation using two-dimensional digital image correlation. Exp Mech 53(5):755–765
 6. Tang Z, Liang J (2012) Photogrammetry-based two-dimensional digital image correlation with nonperpendicular camera alignment. Opt Eng 51(2), 023602
 7. Sun W, Dong E, He X (2007) 3D displacement measurement with a single camera based on digital image correlation technique. Proc SPIE 6723(1), 67230E
 8. Mares C, Barrientos B, Blanco A (2011) Measurement of transient deformation by color encoding. Opt Exp 19(25):25712–25722
 9. Felip-Sese L, Siegmann P, Diaz F, Patterson E (2014) Simultaneous in-and-out-of-plane displacement measurements using fringe projection and digital image correlation. Opt Lasers Eng 52:66–74
10. Gorthi SS, Rastogi P (2010) Fringe projection techniques: whither we are? Opt Lasers Eng 48:133–140
11. Parulski K, Spaulding K (2002) Color image processing for digital cameras. In: Sharma G (ed) Digital color imaging handbook. CRC Press, Boca Raton

12. Luster BS (2013) Using telecentric lenses in inspection systems http://www.vision-systems.com. Accessed 10 Sep 2013
13. Huntley MC (1982) Diffraction gratings. Academic Press, New York, pp 22–24
14. AVT (2013) Prosilica GC technical manual, Ver. 2.0.8, Allied Vision Technologies, Stadtroda, Germany
15. Goodman J (1976) Some fundamental properties of speckle. J Opt Soc Am 66(11):1145–1150
16. Cloud G (1998) Optical methods of engineering analysis. Cambridge University Press, Cambridge
17. Foveon Inc. (2014) Direct image sensors. http://www.foveon.com. Accessed 5 Jan 2014

Chapter 37
Multi-Camera DIC Offers New Dimensions in Material Testing

Thorsten Siebert and Vinh Tran

Abstract Optical full field techniques, like digital image correlation (DIC) become more and more standard tools for the determination of material parameters. Especially for anisotropic materials the multidimensional information is important. For example, up to now this is limited to 2D strain information. The use of the innovative cluster approach for DIC and a multi camera setup in a standard tensile test allows the access to the third dimension and measure not just the strain on one surface but on both surfaces and in thickness direction.

With the cluster approach the points to be evaluated are not defined by an image of one camera, like in conventional DIC systems, but on the object. In this case every object point seen by two or more cameras can be measured, independently of the arrangement of the cameras. In a front back side arrangement e.g. during a tensile test, both sides of the sample are measured simultaneously. As all cameras are calibrated in a common coordinate system the thickness of sample and any changes are measured directly. In this way the strain in thickness direction is measureable directly.

As an example we present the measurement on a simple aluminium sample up to failure. The difference in the strain in direction of thickness and on the surface indicates anisotropic material properties caused by the process of manufacturing.

Keywords Digital image correlation • Multi camera DIC • Strain measurement • Material testing • R value

37.1 Introduction

Now a day, Digital Image Correlation (DIC) technique has been successfully proven to be an accurate analysis tool for displacement, vibration and strain analysis. Most of the systems and applications are divided into 2D or 3D measurements and therefore referred to measurements using one (2D) or two cameras (3D) [1]. Taking the development of the DIC technique into account it started with one camera systems which are able capturing a 2D projection of the sample and so limited to plane objects with in-plane displacement only. This limitation was caused by the technical state of the art of digital image processing and camera technology. The results of these measurements gave a big advantage in the understanding of e.g. inhomogeneous materials. The next step increased the number of cameras from 1 to 2. This additional camera opens the application from plane objects into 3D objects, from in-plane deformations to 3D deformations and therefore overcomes the limitations of DIC system using single cameras. This more on information and flexibility must be bought by additional complexity in the determination of the projection parameter, geometrical setup and the DIC algorithm [2]. The effort paid for itself. If possible people like to get the 3D information.

Why stopping using two cameras? What advantage might additional cameras, like 3, 4 or more, give?

T. Siebert (✉)
Dantec Dynamics GmbH, Kässbohrerstr. 18, 89077 Ulm, Germany
e-mail: Thorsten.Siebert@Dantecdynamics.com

V. Tran
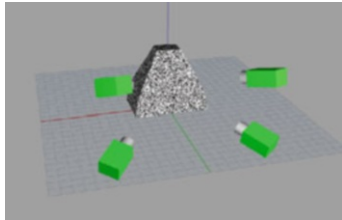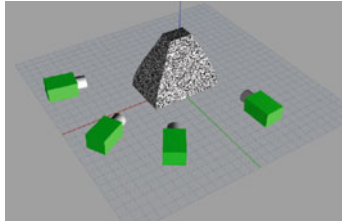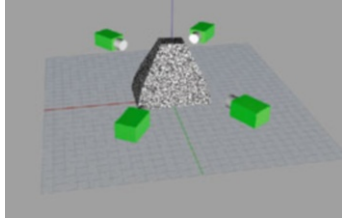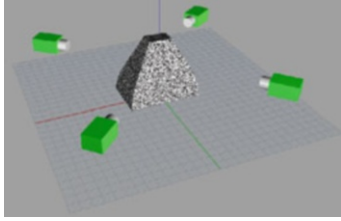Dantec Dynamics Inc, 750 Blue Point Rd, Holtsville, NY 11742, USA

37.2 DIC Using the Cluster Approach

As 3D information can only be calculated from objects points viewed from two cameras, the field of measurement is limited by the overlap of the field of view of the cameras. These restrictions apply to non plane objects mainly. By adding a 3rd camera the measureable area, especially on curved objects or corners, can be enlarged. In addition the quality of the measurement is increased in the area where all three cameras view the object [3]. This addition of a 3rd camera doesn't cost any further change in the principle of how the DIC algorithm is used. But still the area, which can be evaluated, is restricted by the view of one, so called reference camera. As in the image of this camera the points to be measured are defined. This principle is still based on the first one camera systems. An additional 4th camera may increase the resolution of the DIC system but does not enlarge the field of measurement. Unless the way of defining the measurement points on the object is changed.

By using the cluster approach for DIC systems this limitation can be overcome. Here the points to be evaluated are not defined in a single camera view but on the object itself [4]. Every object point seen by at least of any two cameras is used. This innovative approach opens for DIC technique new dimensions in application and information.

By using e.g. a four camera setup and the cluster approach the following different configurations are possible (Table 37.1):

Table 37.1 Different configurations of a four camera multi camera setup

One side arrangement All cameras see the object from one side		• Increased accuracy • Extension of field of measurement
Corner arrangement Cameras look at two sides and the corner		• Measurement around a corner • Increased accuracy on the sides and at the corner
Double side arrangement Cameras look at front and back side		• Simultaneous measurement of front and back side • Data in common coordinate system • Direct measurement of thinning
All side arrangement Cameras look around the object		

37.3 Multi Camera DIC in Material Testing

In the following we demonstrate the advantages using the cluster approach in a front-backside setup for the material testing application. Assuming a plane sample in a tensile testing configuration, in the first approach one might think it's a simple 2D problem and a single camera setup might be sufficient. But experiences have shown that there are deviations from an ideal plane situation. Therefore a 3D measurement is recommended. Using a classical two camera configuration, the displacement and strain on the surface viewed by the cameras. Here e.g. the Poisson ratio can be calculated from the ratio of the strain in and perpendicular to the loading direction. But in order to understand the complete behavior some assumption are made; for example regarding material isotropy or a bending of the sample during the loading. For a complete characterization of the material the behavior in the direction of the thickness is required. Especially for rolled materials from this the Lankford coefficient, also called R value, can be determined [5].

In a front-back side configuration all cameras are mounted on a common frame, two cameras view the front and two cameras view the back side of the sample (Fig. 37.1). For the determination of the projection parameter of each camera and the geometrical position of the camera the calibration is performed using a double sided calibration target. The front and back side of this target having calibration pattern with a known position relative to each other. In this way all cameras are calibrated in the same coordinate system. For the evaluation the two cameras reconstruct the front side and the other cameras the backside of the sample. As both sides are in the identical coordinate system the distance between them is the thickness directly. During the loading process not just the deformation and strain in both sides independently, but also the thinning is measured. No assumptions about the sample are required and so the strain in the direction of the thickness can be measured directly as well as it might be important for composite materials, the displacement and strain on both surfaces independently.

37.4 Measurement of Three Strain Components During Tensile Test

As an example the results of a tensile test of a 1 mm thick rolled aluminum t-bone sample are presented. In this experiment the X direction is the direction of loading and the Y direction is perpendicular to the load in the surface and the Z direction is in the direction of thickness of the sample.

The test started without preload and last until failure. The tensile test machine was operating in constant traverse seed mode. During the test two cameras imaging the front side of the sample and two cameras the back side. All cameras are operating in portrait format and acquire the images synchronized. Due to the shape of the sample, the region of interest was reduced in order to increase the acquisition speed.

Fig. 37.1 Multi camera setup for tensile test measurement

Fig. 37.2 Strain in X direction over time measured at front and back side
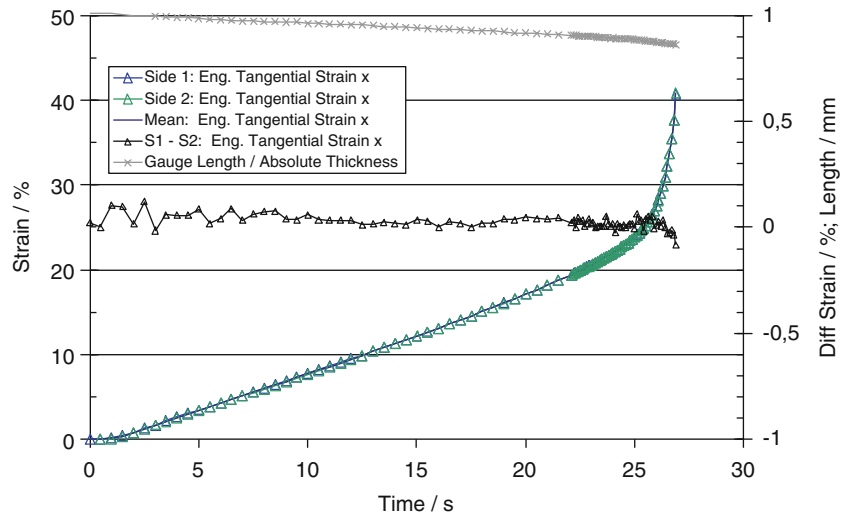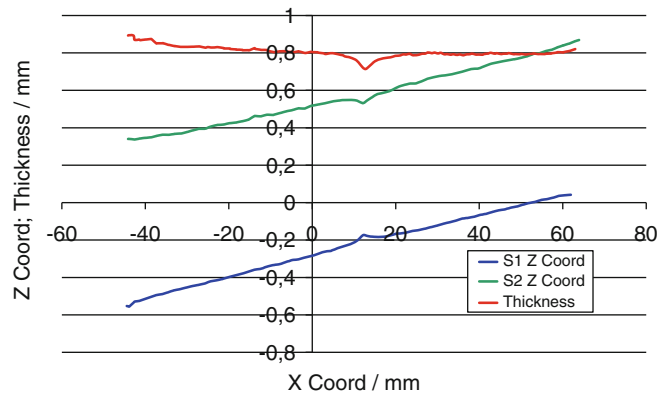

Fig. 37.3 Z coordinate on front and back side and thickness along a line before failure

37.4.1 Comparison of Strains on Front and Back Side

At the first step the strain values at the front and backside are compared. A mismatch of the data would indicate a bending of the material before or during the test or might be caused by an uneven fixation of the sample. In the following Fig. 37.2 the strain in X direction and their differences at the point of failure are plotted over the time. The strain data show up to just before the failure no significant deviation before failure, just in the last 1 s a higher strain on one side is indicated. In addition to the strain the thickness of the sample over time is shown in the upper part. It can be seen, that the thickness is reduced down to about 0.75 mm before failure.

From these results one can state that the boundary conditions of the test were good. No bending or unsymmetrical clamping of the sample was found. In the next step, the strain perpendicular to the loading direction is investigated. The strain in Y direction is the value which is also accessible by a conventional DIC system. For these standard setups the missing information about changes in the thickness can only be estimated by making assumptions regarding the overall behavior and the measured distance of the object points in regard to the DIC system.

37.4.2 Measurement of Changes of Thickness

At the first step the z coordinate along two lines in the center on both surfaces of the object are compared. Therefore the coordinate system is defined at the unloaded step on the surface of the object. At a measurement step just before failure the Z coordinates and the thickness of the sample along a vertical line are shown in Fig. 37.3. It can be seen that the object was tilted about 0.5 mm during the load. Also the thinning of the sample in the base material and at the point of failure are clearly visible.

Fig. 37.4 Strain
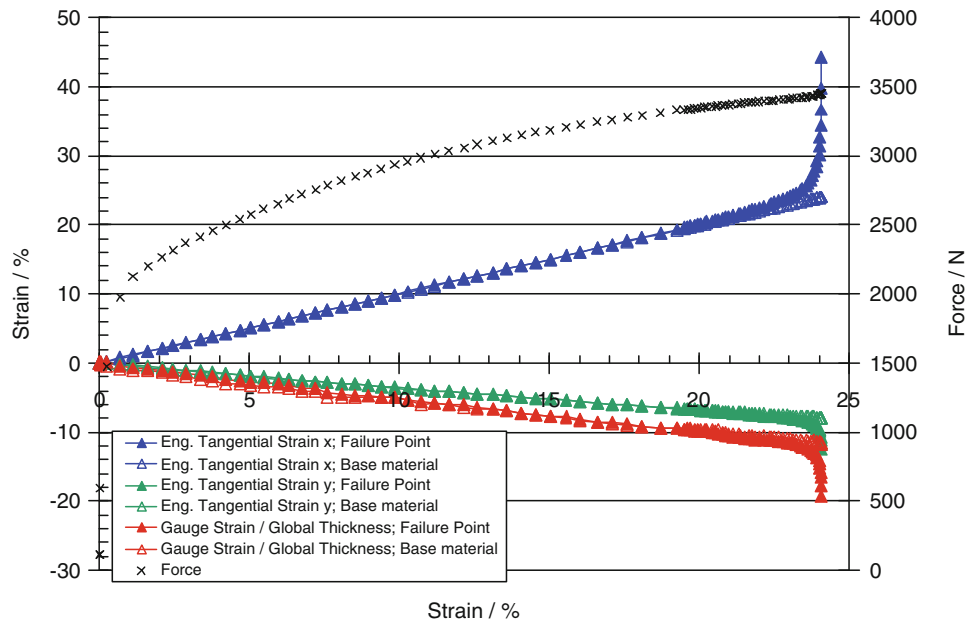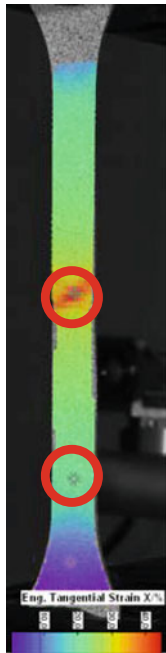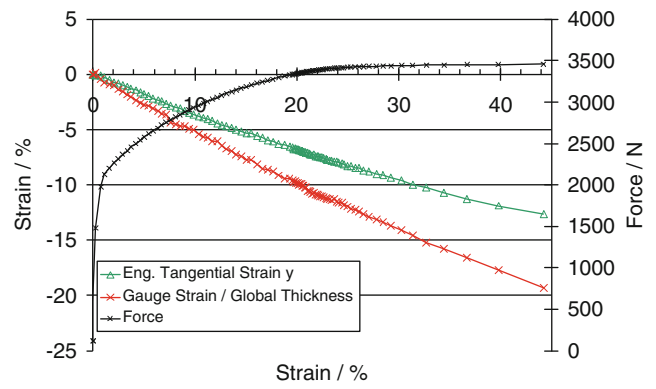perpendicular over strain
in loading direction

Fig. 37.5 Comparison of all strain components at point of failure and base material

37.4.3 Measurement of Third Strain Component

Using the Multi Camera DIC approach the changes in thickness can be measured directly and so the strain in that third direction is measurable. The strain determined from this is shown in the Fig. 37.4 as well. The strains in Y direction and in thickness are plotted over the strain in X direction and show significant deviations from the beginning. All values are calculated at the point of failure.

The strain in the thickness direction is higher than in Y direction. This might be caused by an-isotropic behavior of the material like it is caused by the rolling direction of the material, which is not known for this sample. In addition to the strain the force, recorded from the tensile testing machine is plotted in the graph.

Figure 37.5 compares all three strain components at the point of failure with a point away from the hot spot in the base material. The strains at the point of failure and a point at the base material are plotted over the strain in loading (X) direction at the base material. The location of the points is indicated at the image of the sample at the left side. In addition to the strains the force is plotted in the graph.

37.5 Conclusion

We have discussed and shown the advantages of using more than two cameras for performing DIC measurements. The use of the cluster approach for a multi camera DIC measurements of tensile test sample allows the direct measure of all three strain direction during the compete test for each object point. This is one dimension more than conventional DIC systems are able to deliver. The extension of standard DIC systems to Multi Camera Systems requires a new way of defining the points to be correlated but this innovative approach increases the flexibility of the arrangement of the multi cameras and so the applications dramatically. In addition new information, like the third strain component in tensile testing, is accessible directly.

References

1. Sutton MA, McNeil SR, Helm JD, Chao YJ (2000) Advances in 2-D and 3-D computer vision for shape and deformation measurements. In: Rastogi PK (ed) Photomechanics, vol 77, Topics in applied physics. Springer, New York, pp 323–372
2. Sutton MA, Orteu J-J, Schreier HW (2009) Image correlation for shape, motion and deformation measurements. Springer, New York. ISBN 978-0-387-78746-6
3. Siebert T, Crompton MJ (2012) An approach to strain measurement uncertainty for DIC using the SPOTS calibration procedure. In: Proceedings of the SEM international conference and exposition on experimental and applied mechanics, paper 107, 11–12 June 2012
4. Orteu J-J, Bugarin F, Harvent J, Robert L, Velay V (2011) Multiple-camera instrumentation of a single point incremental forming process pilot for shape and 3D displacement measurements: methodology and results. Exp Mech 51:625–639
5. http://en.wikipedia.org/wiki/Lankford_coefficient. Accessed Feb 2014

Chapter 38
Using Sampling Moiré to Extract Displacement Information from X-Ray Images of Molten Salt Batteries

Phillip L. Reu, Enrico Quintana, and Kevin Long

Abstract Full-field axial deformation within molten-salt batteries was measured using x-ray imaging with a sampling moiré technique. This method worked for in situ testing of the batteries because of the inherent grid pattern of the battery layers when imaged with x-rays. High-speed x-ray imaging acquired movies of the layer deformation during battery activation. Numerical validation of the technique, as implemented in this paper, was done using synthetic and numerically shifted images. Typical results of a battery are shown for one test. Ongoing work on validation and more test results are in progress.

Keywords Moiré • Full-field measurements • X-ray • Battery

38.1 Introduction

Sampling moiré is a full-field measurement technique that uses the grid of a digital imaging detector with a grid pattern on the measurement object to create moiré patterns [1]. The shifting moiré patterns will reflect the underlying deformation or motion of the sample. An advantage over other moiré techniques is that only one image is required to create phase-shifted moiré patterns and allows for the use of high-speed imaging with phase shifting. Another advantage is that a reference grating is not required, simplifying the experimental setup, by using the CCD rows and columns on the detector as the reference grating. The phase distribution of the object deformation can be calculated from the moiré fringes using the discrete Fourier transform (DFT). Relative deformation can be calculated by subtracting the phase of the current load step from that of a reference image in an unloaded configuration.

Sampling moiré was used in this experiment to measure the deflection of the layers in a molten-salt battery. This battery type naturally lends itself to this technique because of the repeating layers of uniform thickness that make up the battery. The differing density of the layers creates a natural grid pattern when imaged with x-rays. This approach also allows the battery to be imaged in situ, with the case, insulation and all of the ancillary components, rather than as an unconstrained stack and will yield results that reflect the layer motion in the actual battery configuration.

38.2 Experimental Setup

To image the molten-salt battery, a collimated x-ray source (Philips 225 kV, 2.5 kW) was aligned with the battery such that a cross-sectional view of the battery layers was seen (Fig. 38.1). An intensified x-ray scintillator screen (Model AI-9448-F) created a visible image of the attenuated x-rays. A Phantom V9 camera with a K2 lens using a CF4 high-magnification adapter imaged the intensified x-ray image. The camera resolution was $1,632 \times 1,200$ pixels at 500 frames per second (fps). The camera resolution greatly exceeded that of the scintillator screen, resulting in empty magnification. Therefore, a 4×4-pixel binning reduced the noise of the image by averaging a pixel region without losing effective resolution. The binned images were of higher quality; however, the noise was still larger than desired. There were also a number of

P.L. Reu (✉) • E. Quintana • K. Long
Sandia National Laboratories, PO Box 5800, Albuquerque, NM 87185, USA
e-mail: plreu@sandia.gov

Fig. 38.1 Sampling moiré X-ray experimental setup. (Need to replace battery image)
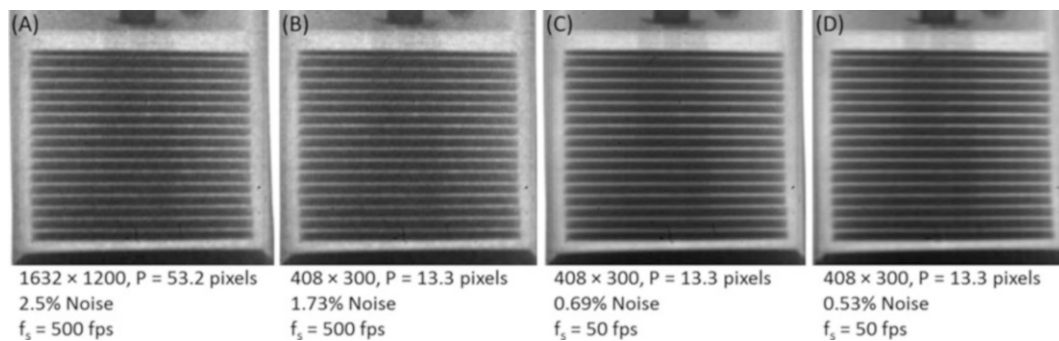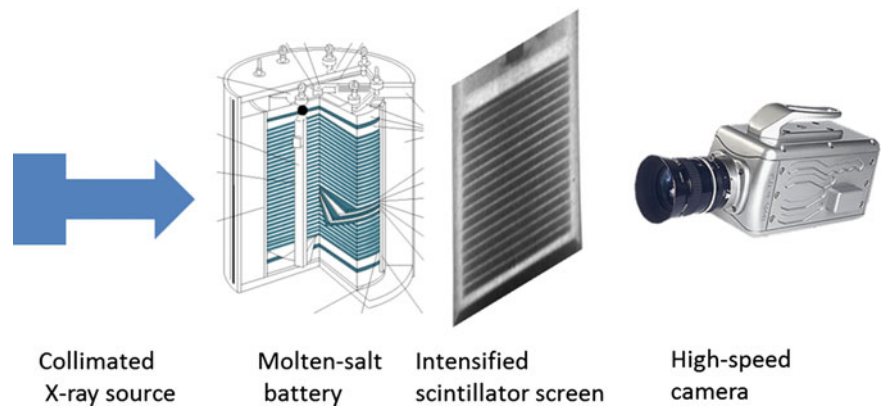

Collimated X-ray source | Molten-salt battery | Intensified scintillator screen | High-speed camera

| (A) 1632 × 1200, P = 53.2 pixels 2.5% Noise f_s = 500 fps | (B) 408 × 300, P = 13.3 pixels 1.73% Noise f_s = 500 fps | (C) 408 × 300, P = 13.3 pixels 0.69% Noise f_s = 50 fps | (D) 408 × 300, P = 13.3 pixels 0.53% Noise f_s = 50 fps |

Fig. 38.2 Pre-processing steps before analysis with the sampling moiré algorithm: (**a**) Raw image, (**b**) Binned image, (**c**) 10 frame average image, and (**d**) FFT filtered image

imperfections in the image due to dust. These were removed from the image with a 2D FFT filter, biased to low-pass filter the image in the horizontal direction. After a preliminary analysis showed that the motion of the batteries was slow relative to the frame rate, the images were averaged in groups of 10. This reduced the effective frame rate from 500 fps to 50 fps. The processing steps and the resulting images are shown in Fig. 38.2.

It is important to consider what the image being recorded represents. The image is the average x-ray attenuation along a (nearly) straight path. A collimated source is assumed for this because the point source distance was large, approximately 2 m from the sample, relative to the sample size, making the x-rays approximately parallel at the sample location. Because of this, the measured deflection is an average of the deformation of each layer throughout the thickness of the battery. When the layers tilt, this has the effect of decreasing the contrast between neighboring layers. This effect was not quantified in the current analysis.

38.3 Moiré Code Validation

The sampling moiré code was validated both numerically and experimentally. The first section briefly presents the sampling moiré theory and mathematics. The next two sections describe the validation approach and present the results.

38.3.1 Sampling Moiré Theory

Sampling moiré is an efficient method of creating moiré fringe patterns from a single image. Moiré patterns are created when two overlapping grids are superimposed on one another. The resulting moiré pattern can subsequently be used to measure the motion or deformation of one of the two patterns. The undeformed pattern is the "reference" against which the motion of the other grid can be measured. In sampling moiré the CCD detector functions as the reference grid, with the added

Fig. 38.3 Sample grid pattern from battery on a row of pixels showing the pitch (P) and thinning (T)
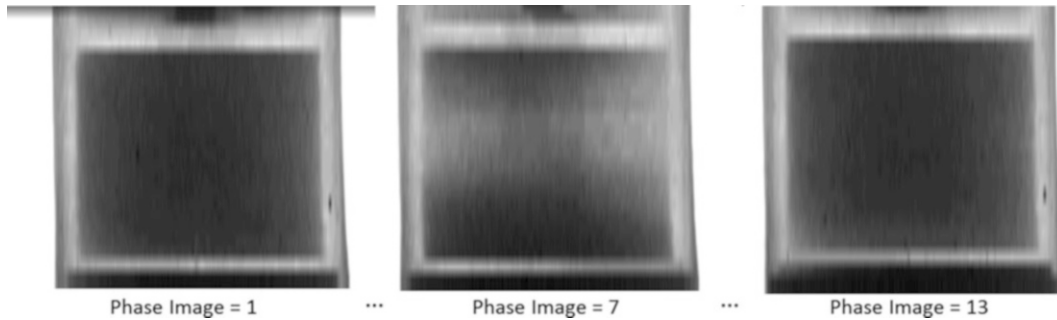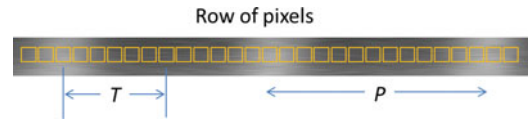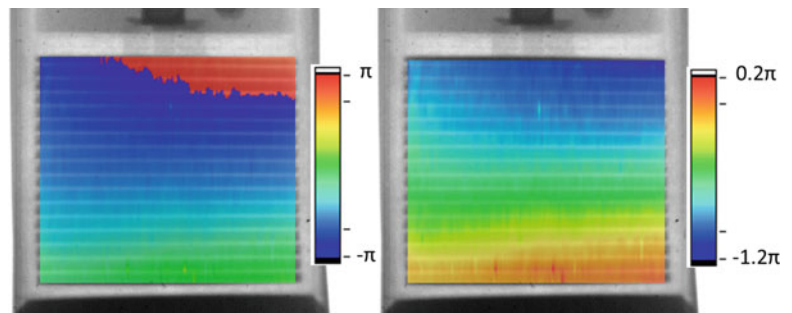

Fig. 38.4 Three representative images (from a total of 13) showing the moiré fringe pattern phase shift in the grid region of the battery

**Fig. 38.5** (*Left*) Wrapped moiré phase and (*right*) unwrapped moiré phase

advantage of being able to dynamically adjust the reference period by selecting, T, the distance between pixels. The sample must of course have a grid pattern, P, which may either be applied artificially or be inherent in the sample as in this experiment. The recorded image of the sample may then be "thinned" by a factor T, by taking every T th pixel and resulting in T decimated images. The decimated images created by the thinning process are then interpolated to return them to their original resolution. The interpolation used for this paper is either a linear or spline interpolation. The process is illustrated in Fig. 38.3. It is important to match T as closely as possible to P, to avoid errors [2]. To do this, the frequency of the fringe period was measured by taking the fundamental frequency as calculated using a 1D FFT. The fringe period for the experimental setup was measured to be 13.3 pixels. Consequently the images were thinned with $T = 13$ pixels. Three representative thinned and interpolated images are shown in Fig. 38.4. Because the fringe and thinning period are nearly identical, the resulting moiré fringes are accordingly very broad.

The resulting moiré fringes are phase shifted relative to each other by using a different start pixel for each of the thinned images. The shifted phase value can be calculated from the phase shifted moiré fringes by a number of methods, a convenient one being the discrete Fourier transform (DFT) [3]. Figure 38.5 shows a wrapped phase map ($\pm\pi$) created using the DFT and three of the 13 shifted moiré fringe images created from Fig. 38.2c. The wrapped phase must then be unwrapped by any number of different phase unwrapping methods. Relative motion can be measured by taking a reference image and calculating the phase as shown in Fig. 38.5 (right) and then subtracting this phase from any subsequent images that have either deformation or translation.

38.3.2   Sampling Moiré Validation

The moiré code was validated using both synthetically created images and experimentally shifted images.

Fig. 38.6 Translation validation
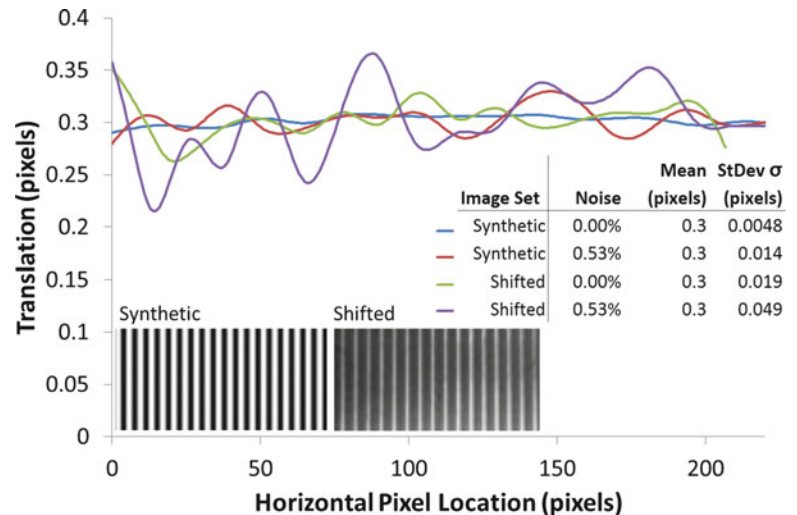
38.3.2.1 Sampling Moiré Numerical Validation

Both translated and deformed images were simulated. The shifted images were created using the Fourier shift theorem and implemented as outlined in [4]. The approximate noise level of the images was first quantified by taking 5 images pre-test and calculating the mean and standard deviation on a pixel-by-pixel basis. This procedure highlighted noise problems that are inherent in x-ray imaging as shown in Fig. 38.2. The noise level used was for the 0.53 % noise case. Two types of images were created. The first image type labeled "synthetic" was a sinusoidal pattern with a reference image with 0° phase shift and a translated image with a phase shift equivalent to 0.3 pixels. The second image was the actual undeformed x-ray pattern shown previously and numerically shifted. Noise for both cases was added as a Gaussian distribution on a pixel-by-pixel basis. For the actual deformed image it should be noted that it is not possible to have a truly "noise free" case, as even with averaging and filtering some noise will remain in the images. This is one reason why these images return noisier results. (Other reasons could be: non-uniformity of the grid features, high-frequency components aliasing during the shifting, non-uniform grey values throughout the image.) The images were analyzed with the sampling moiré code and a translation was measured. The results for a line cut near the middle of the image are shown in Fig. 38.6. The inset table also indicates the mean and standard deviation for the entire image area.

To validate the ability to measure varying displacement fields across the sample, a varying sine wave pattern was created. The strain was applied with a Fourier expansion of the image by a factor of 100, and then using a linear interpolation of the intensity at the shifted pixel location to simulate the pattern. A small deformation of 0.3 pixels was chosen to have a representative displacement while remaining well above the noise floor of 0.02 pixels for the real image without noise. The smallest detectable displacement will then between 0.02 and 0.05 pixels depending on the noise level of the images (Fig. 38.7).

38.4 Experimental Results

The tested batteries were analyzed using the sampling moiré approach as outlined here. This analysis process yields a dense set of data that can be analyzed in a number of different ways. The grid direction yields a displacement in only the *y*-direction. Figure 38.8 shows a time-lapse overlay of the *y*-deformation for six frames. The deformation is down at the middle of the battery and the bottom moves up 0.01 mm. Figure 38.9 shows one of the frames with cross-sectional cuts of the deformation gradients overlaid on the 2D deformation field.

Fig. 38.7 Varying sine wave. Ideal sine pattern (*top*), X-ray image numerically deformed (*bottom*). Both with noise of zero and 0.53 % noise
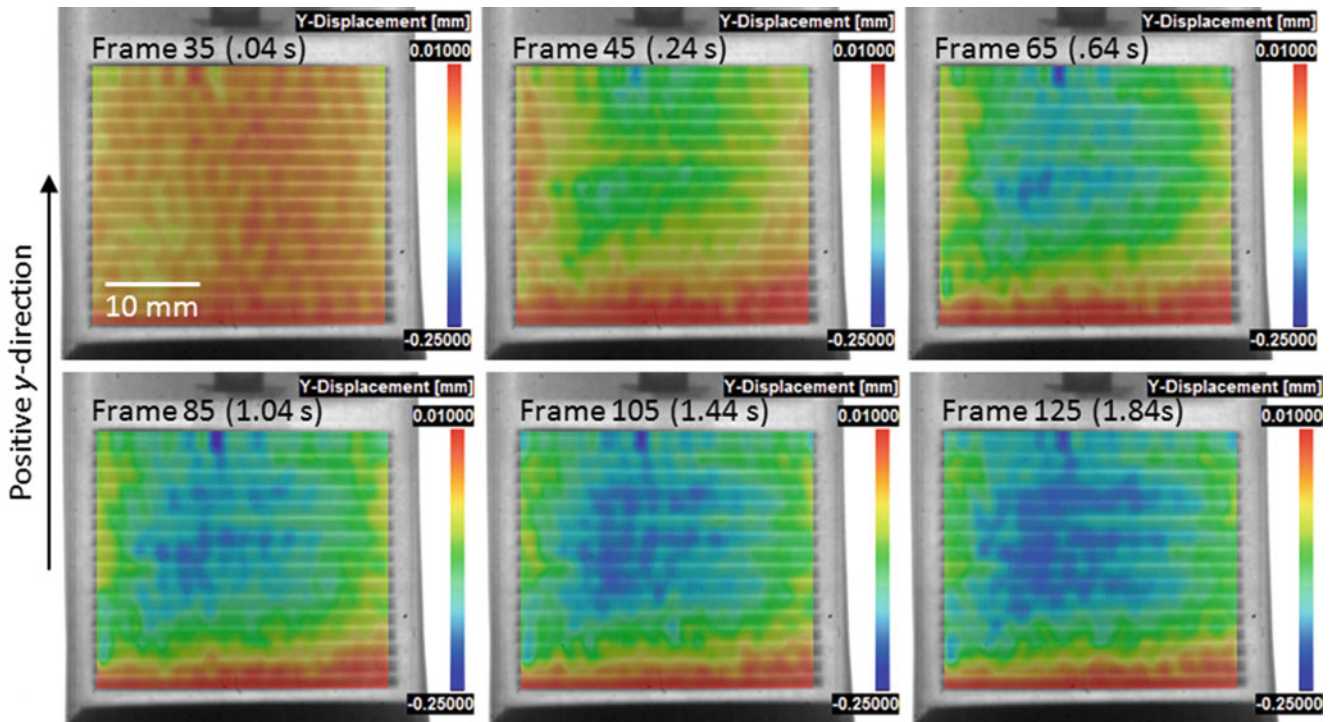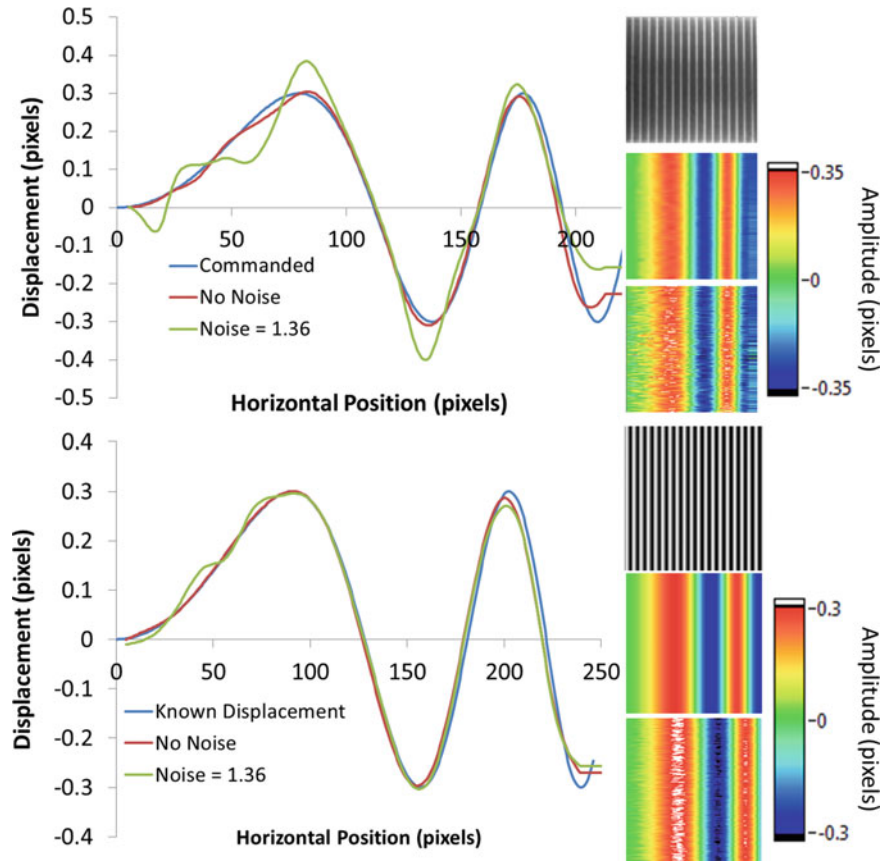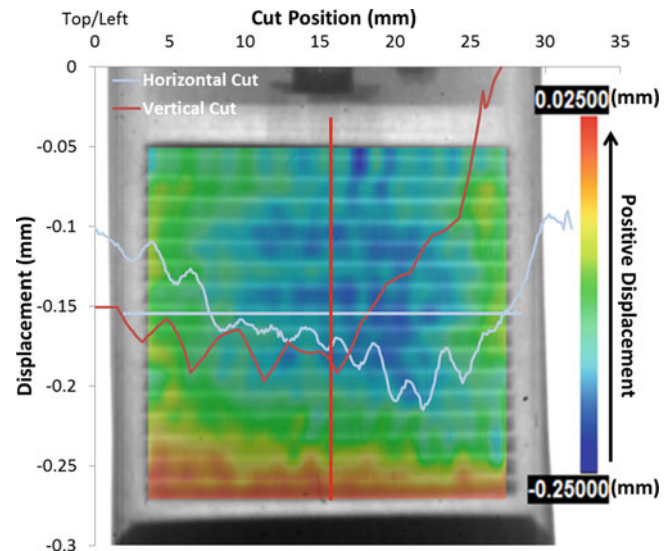
Fig. 38.8 Time lapse image (identical scales) of the molten-salt battery

Fig. 38.9 A typical result
(maybe). We need some
more testing

38.5   Conclusions

X-ray moiré has proven to be a valuable measurement technique to probe the deformation of the layers of a molten-salt battery. Outstanding problems for analysis are to determine the effect of rotation of the layers on the measurement. Preliminary testing indicates that a loss of contrast in the fringes occurs in the central region. Experimental verification of the method are also ongoing using printed fringe patterns to look at the causes of error as it relates to the molten-salt battery experiments.

References

1. Ri S, Fujigaki M, Morimoto Y (2010) Sampling Moiré method for accurate small deformation distribution measurement. Exp Mech 50(4):501–508
2. Ri S et al (2012) Accuracy of the sampling Moiré method and its application to deflection measurements of large-scale structures. Exp Mech 52(4):331–340
3. Morimoto Y, Fujisawa M (1994) Fringe pattern analysis by a phase-shifting method using Fourier transform. Opt Eng 33(11):3709–3714
4. Reu P (2011) Experimental and numerical methods for exact subpixel shifting. Exp Mech 51(4):443–452

Chapter 39
High-Speed Digital Holography for Transient Response of the Human Tympanic Membrane

I. Dobrev, C. Furlong, J.J. Rosowski, and J.T. Cheng

Abstract The tympanic membrane (TM; eardrum) transforms sound energy from the ear canal into mechanical vibrations of the ossicles (middle–ear bones), whence energy enters the inner ear and stimulates auditory hair cells for hearing. Our past efforts toward understanding of the energy transformation within the middle-ear has been focused on the TM's response to single tones. In this paper, we report results of the development of high-speed holographic capabilities to measure the transient characteristics of the TM.

We have developed a high-speed holographic system (HHS) that incorporates a transient (i.e., >40 kHz) acquisition method to quantify the displacements on the TM surface excited by sharp (i.e., <50 μs) acoustic clicks. The performance of the HHS is compared with Laser-Doppler Vibrometery (LDV) measurements on both artificial membrane and human cadaveric TM. The instantaneous response of >100 k data points on the TM surface measured by the HHS can be used to quantify spatially-dependent motion parameters such as modal frequencies, mode shapes, time constants, and acoustic delays. In addition, the local material properties across the surface of the TM can be inferred.

Keywords Acoustic-solid interaction • Digital holography • High-speed cameras • Local phase correlation • Otology

39.1 Introduction

Ongoing hearing research efforts to understand the human hearing process are mainly focused on tonal acoustic excitation response of the human ear [1–5]. However, studying the transient response of the human ear, and of the human tympanic membrane (TM) in particular, could expand our understanding of the processes by which acoustical energy transformation and transmission from the TM to the ossicles chain occur [1]. Current state-of-the-art clinical methods to measure the

I. Dobrev (✉)
Center for Holographic Studies and Laser Micro-mechaTronics - CHSLT, 100 Institute Road, Worcester, MA 01609, USA

Department of Mechanical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA
e-mail: ivo_d@wpi.edu

C. Furlong
Center for Holographic Studies and Laser Micro-mechaTronics - CHSLT, 100 Institute Road, Worcester, MA 01609, USA

Department of Mechanical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA

Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02114, USA

J.J. Rosowski • J.T. Cheng
Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02114, USA

transient acousto-mechanical response of the TM rely on averaged acoustic or point-wise displacement information [6–8]. These methods may be insufficient to fully describe the complex patterns that unfold across the TM surface.

In this paper, we report our progress of the development and implementation of a full-field, high-speed, holographic system (HHS) that utilizes a novel 2 + N frame acquisition method based on hybrid spatio-temporal local correlation (LC) phase sampling approach [9, 10]. The HHS utilizes the temporal resolution of a high-speed camera without imposing constrains on its spatial resolution and without the need of specialized optical setups [11, 12]. Automatic execution and synchronization of high-speed measurements is achieved by a custom control system. The high temporal (i.e., >40 kHz) and spatial (i.e., >100 k data points) measurement resolutions of the HHS enable the investigation of the complex spatio-temporal behavior of the TM at a sufficient level of detail to further expand our knowledge of the hearing processes.

39.2 Methods

We have developed and implemented a novel high-speed 2 + N frame acquisition method based on hybrid spatio-temporal local correlation (LC) phase sampling approach [9], that allows quantification of the TM transient deformation by utilizing two reference frames, I_{ref}, $I_{ref + \pi/2}$, and N consecutive deformed frames, $(I_{def})_{i, i \in 1, 2.. N}$, recorded before and throughout the evolution of an event, as shown in Fig. 39.1.

According to Fig. 39.1, the 2 + N frame acquisition method starts by recording two reference frames, I_{ref} and $I_{ref + \pi/2}$, temporally separated to bypass the piezo-transducer (PZT) phase shifter settling time after the introduction of a $\pi/2$ phase shift [13]. While keeping the PZT at its final phase shift position, a set of deformed frames, $(I_{def})_{i, i \in 1, 2.. N}$, are recorded at rapid sampling rate (e.g., >40 kHz). The corresponding spatio-temporal variation of the double-exposure optical phase change, $\phi(m, n, t)$, is quantified by correlating the intensities of each deformed and reference frames with

$$\phi(m, n, t) = tan^{-1}\left[\frac{\rho\left(I_{ref}(m, n), \quad I_{def}(m, n, t)\right)}{\rho\left(I_{ref+\pi/2}(m, n), I_{def}(m, n, t)\right)}\right], \tag{39.1}$$

where ρ is the Pearson's correlation coefficient for finite discrete sets having the recoded intensities as arguments and computed based on a spatial kernel (i.e., with dimensions of 3×3 or 5×5 pixels) centered around each measurement point, (m, n), of a pair of reference and deformed frames [9].

39.3 Validation

The measuring capabilities of the HHS was characterized relative to LDV measurements of the transient responses of latex and human TM samples and at several predefined locations, as shown in Fig. 39.2. The full-field transient responses of both samples were recorded with the HHS set at 42 kHz sampling rate and with a 6.62 μs exposure time (corresponding to a shutter speed of 151 kHz). The point-wise LDV measurements were sampled at 84 kHz. Both samples were acoustically excited by a 50 μs click. Representative results of the comparisons between HHS and LDV based on the time waveforms of the maximum and minimum responses of the surface of each sample are shown in Fig. 39.3. The differences between HHS and LDV results are <5 % relative to the maximum displacement amplitude of the samples. The temporal locations of the local maxima and minima in the time waveforms measured by the HHS and LDV differ by <10 μs.

Fig. 39.1 A timing diagram demonstrating the principle of operation of the high-speed 2 + N frame acquisition method based on a hybrid spatio-temporal phase sampling approach
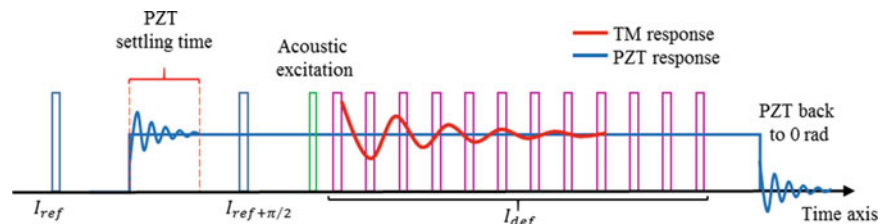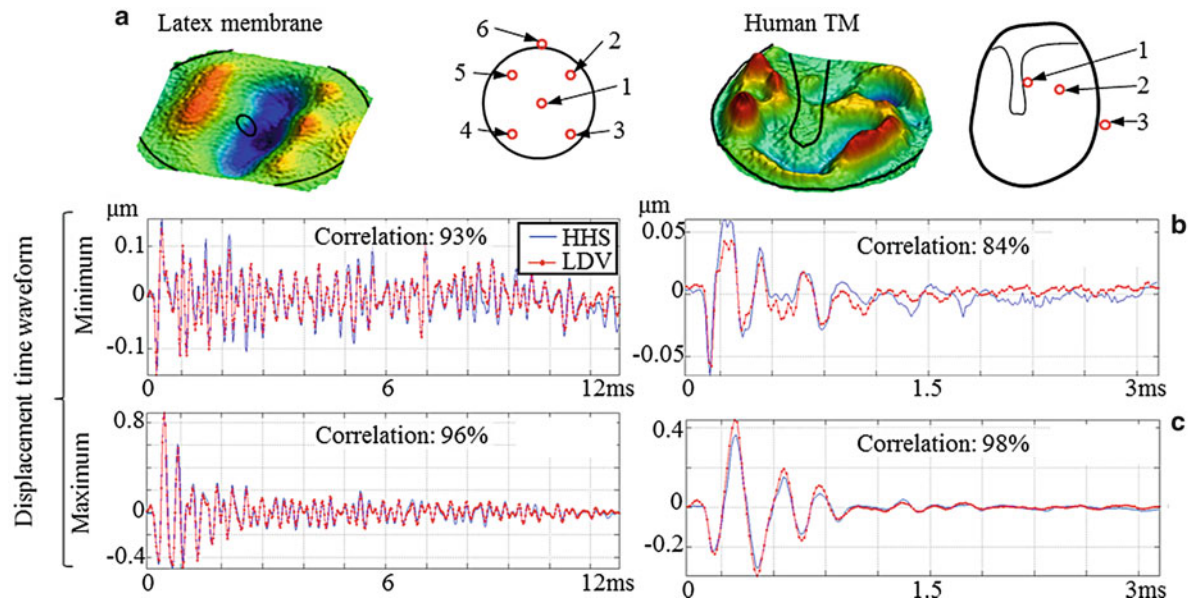
Fig. 39.2 Representative full-field-of-view HHS displacement measurement of the surface of latex and human TM samples. In order to characterize the measuring capabilities of the HHS relative to LDV, holographic displacements are sampled at specific discrete locations on the samples: (**a**) full-field-of-view holographic measurement showing the location of the discrete points that are used to identify the minimum and maximum amplitude of deformations; (**b**) and (**c**) show HHS and LDV time waveforms, for each sample, corresponding to locations having the minimum and maximum deformations. Correlation coefficient between HHS and LDV is on average >95 %. The time axis is relative to the beginning of the acoustic excitation, which was a 50 μs click with 115 dB maximum SPL
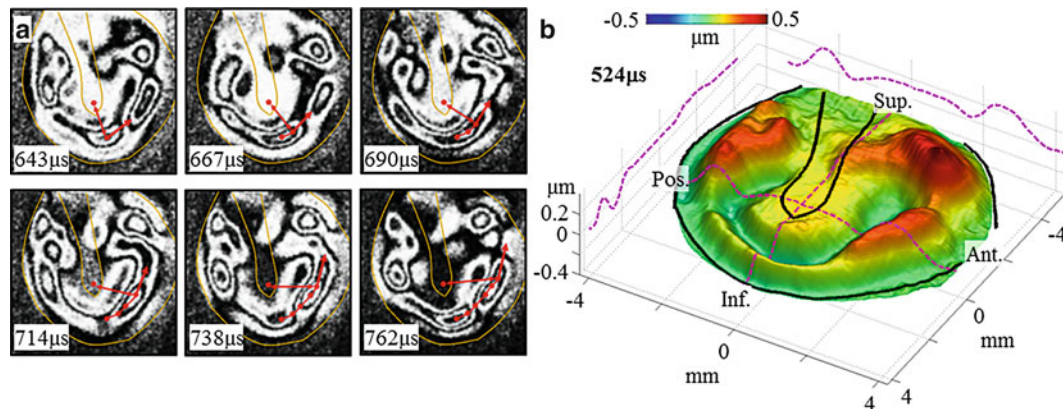

Fig. 39.3 Representative measurements of acoustically induced transient response of a human TM sample obtained with our high-speed digital holography methods: (**a**) modulation maps recorded between 640 and 770 μs after the beginning of the acoustic stimulus; (**b**) 3D displacement map at 524 μs showing 0.63 μm p-p displacement. Outlines of the outer boundary of the membrane and the manubrium are indicated with *solid lines*. Analysis of the measurements indicate circumferentially traveling waves propagating at an estimated speed of 24 m/s. The acoustic excitation was a 50 μs click with 115 dB maximum SPL

39.4 Applications

39.4.1 Full-Field Transient Response of the Human TM

The HHS allows for high acquisition speed (>42 kHz) and high resolution (>100 k data points) enabling measurements of the time waveform of displacement, velocity, and acceleration simultaneously across the full surface of the TM. Representative full-field-of-view displacement measurements at specific temporal instances of the acoustically induced transient response of a human TM sample are shown in Fig. 39.3.
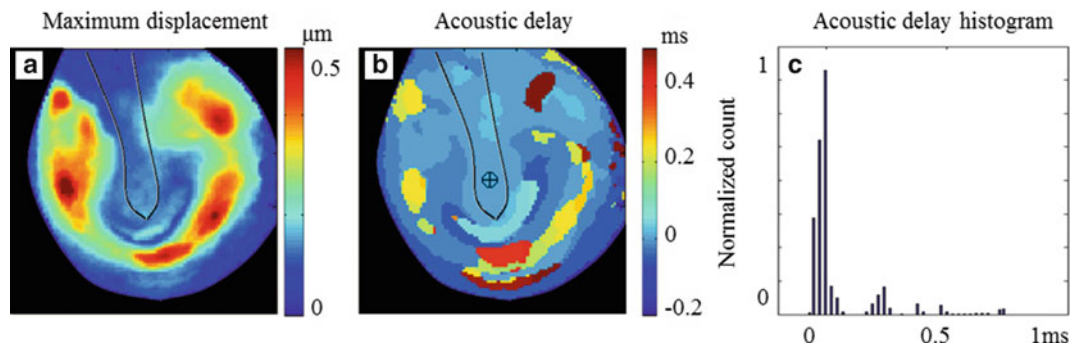
Fig. 39.4 Characteristic maps showing the spatial dependence of the maximum displacement and the acoustic delay across the surface of a human TM sample: (**a**) maximum displacement; (**b**) acoustic delay relative to the umbo (marked with ⊕); and (**c**) histogram of (**b**). Outline of the manubrium is indicated with a *solid line*
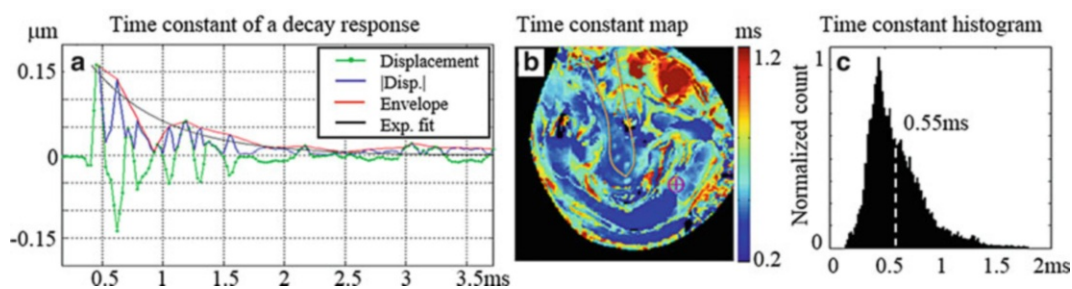

Fig. 39.5 Spatial dependence of the time constant of a human TM sample: (**a**) representative plot showing the quantification of the time constant based on an automatic decay envelope estimation and exponential fitting of the time waveform at a specific location marked with ⊕; (**b**) spatial dependence of the time constant across the surface of the TM; and (**c**) histogram of (**b**) indicating a time constant range of 0.3–1.5 ms with mean of 0.55 ms

39.4.2 Spatial Dependence of Acousto-Mechanical Parameters of the TM

39.4.2.1 Maximum Response and Acoustic Delay Map

By analyzing all time waveforms at all spatial points, we can automatically extract global temporal maximum magnitude of displacements and their temporal locations (peak times). The acoustic delay of the entire surface of the TM can be quantified by referencing the peak time of all points relative to the umbo, as shown in Fig. 39.4.

The acoustic delay data shown in Fig. 39.4 suggest that >50 % of the points on the TM surface between the inferior half of the rim of the TM and the manubrium are within −50 to −20 µs acoustical delay relative to the umbo with smaller acoustical delays (i.e., −50 µs) at the inferior boundary of the TM. This indicates that the acousto-mechanical response of the TM reaches its maximum at the periphery of the surface of the membrane before the umbo. This is in agreement with previous reported theories that suggest that the acousto-mechanical energy is transferred radially from the periphery to the center of the TM [7].

39.4.2.2 Spatial Distribution of the Time Constant of Decay Response

By automatically fitting a decay curve to all time waveforms at all spatial points, we can quantify the spatial dependence of the time constant of the decay response, as shown in Fig. 39.5. The time constant of the response at every point of the TM is related to the acoustic phase delay, mechanical damping, and local material properties.
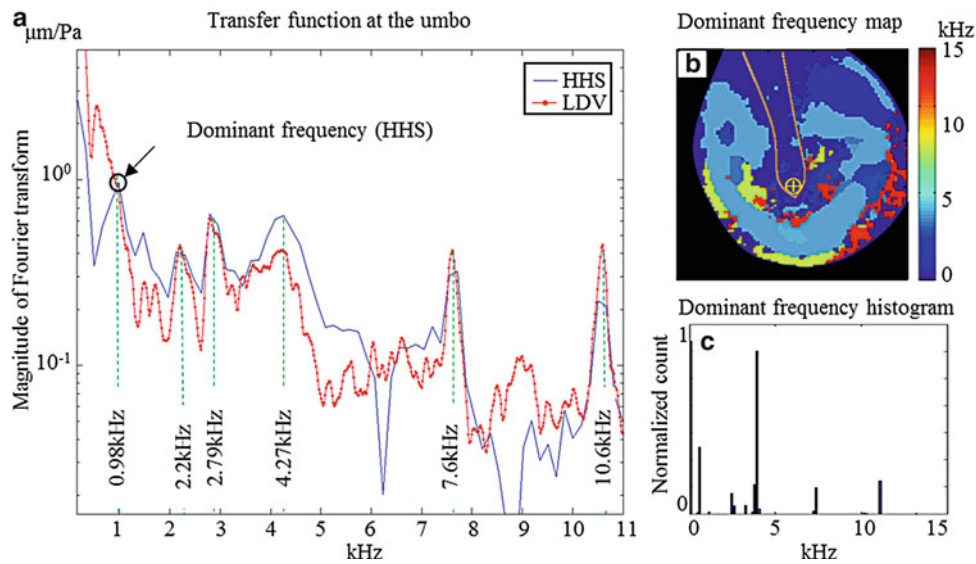
Fig. 39.6 Spatial dependence of the dominant modal frequency of a human TM sample: (**a**) power spectrum of the transfer function of the tip of the umbo (marked with ⊕) measured with HHS (*solid line*) and LDV (*dotted line*) with indicated dominant frequency of 0.98 kHz; (**b**) map of the dominant frequency at every point across the surface of the TM; (**c**) histogram of the dominant frequency map. *Dashed lines* in (**a**) indicate six automatically determined modal frequencies based on the local maxima of the transfer function. Detected modal frequencies at the umbo based on the HHS and LDV differ by <5 %. Outline of the manubrium in (**b**) is indicated with a *solid line*

39.4.2.3 Spatial Dependence of Dominant Frequencies

Based on the magnitude of the FFT of the time-waveform of each point and the pressure time-waveform of the sound wave front (based on calibrated microphone measurements), we can determine the transfer function (TF) of each point across the surface of the TM [5]. Based on automatic identification of the local maxima of the TF, we can determine the modal frequency at each point as shown in Fig. 39.6. By automatically performing this operation across the full surface of the TM and extracting the global maxima at each point, we can determine the spatial distribution of the local dominant frequencies, as shown in Fig. 39.6b. The detected modal frequencies at the umbo based on the HHS and LDV differ by <5 %. The measured dominant frequency range is within previously reported data [1–5].

39.5 Conclusions and Future Work

Our developed high-speed holographic system (HHS) provides high temporal (i.e., <24 µs) and spatial (i.e., >100 k data points) resolutions, which enable measurements of the nanometer (i.e., <15 nm) scale displacements of the full surface of human TM excited by transient acoustic stimuli. Our methods allow quantification of spatially-dependent motion parameters such as modal frequencies, time constants, and acoustic delays. Such capabilities can be used to infer local material properties across the surface of the TM.

Future work should be focused on the analysis and interpretation of the measured transient displacement time waveforms to extract medically meaningful information on the pathologies of the TM. Fundamental research work should be focused on the explanation of the initial transient dynamics of the TM, its relationship to the energy transfer into the middle-ear, as well as its connection to previous steady-state dynamics research [1–5]. Further improvements of the HHS should include optimization of the optical design, spatial resolution, as well as its packaging for in-vivo applications and for medical research.

References

1. Rosowski JJ, Cheng JT, Ravicz ME, Hulli N, Hernandez-Montes M, Harrington E, Furlong C (2009) Computer-assisted time-averaged holograms of the motion of the surface of the mammalian tympanic membrane with sound stimuli of 0.4–25 kHz. Hear Res 253(1):83–96
2. Rosowski JJ, Dobrev I, Khaleghi M, Lu W, Cheng JT, Harrington E, Furlong C (2012) Measurements of three-dimensional shape and sound-induced motion of the chinchilla tympanic membrane. Hear Res 301:44–52. doi:10.1016/j.heares.2012.11.022
3. Khaleghi M, Lu W, Dobrev I, Cheng JT, Furlong C, Rosowski JJ (2013) Digital holographic measurements of shape and 3D sound-induced displacements of tympanic membrane. Opt Eng 52(10), 101916. doi:10.1117/1.OE.52.10.101916
4. Puria S, Allen JB (1998) Measurements and model of the cat middle ear: evidence of tympanic membrane acoustic delay. J Acoust Soc Am 104:3463
5. Cheng JT, Hamade M, Merchant SN, Rosowski JJ, Harrington E, Furlong C (2013) Wave motion on the surface of the human tympanic membrane: holographic measurement and modeling analysis. J Acoust Soc Am 133:918
6. Kemp DT (1978) Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 64:1386
7. Puria S (2003) Measurements of human middle ear forward and reverse acoustics: implications for otoacoustic emissions. J Acoust Soc Am 113:2773
8. Decraemer WF, Khanna SM, Funnell WRJ (1999) Vibrations at a fine grid of points on the cat tympanic membrane measured with a heterodyne interferometer. In: EOS/SPIE international symposia on industrial lasers and inspection, conference on biomedical laser and metrology and applications, Munchen
9. Dobrev I, Furlong C, Rosowski JJ, Cheng JT, Harrington EJ (2014) Implementation and evaluation of single frame recording techniques for holographic measurements of the tympanic membrane in-vivo. Proc SEM 2013(3):85–95
10. Georgas PJ, Schajer GS (2012) Modulo-2pi phase determination from individual ESPI images. Opt Lasers Eng 50(8):1030–1035
11. Pedrini G, Osten W, Gusev ME (2006) High-speed digital holographic interferometry for vibration measurement. Appl Opt 45(15):3456–3462
12. Novak M, Millerd J, Brock N, North-Morris M, Hayes J, Wyant J (2005) Analysis of a micropolarizer array-based simultaneous phase-shifting interferometer. Appl Opt 44(32):6861–6868
13. Physik Instrumente (PI) (2014) Application notes. http://www.pi-usa.us/pdf/Piezo-Actuators_Ceramics-www.pdf. Accessed 15 Jan 2014

Chapter 40
Displacement and Strain Measurement with Multiple Imaging Head Using PSDHI

Motoharu Fujigaki, Hiroki Minamino, and Yorinobu Murata

Abstract The development of compact and conventional strain distribution measurement equipment for practical use is required for the evaluation of the health monitoring and the life-lengthening characteristics of infrastructures, such as steel bridges. Phase-shifting digital holography is a convenient method that can be used to measure displacement and strain distributions on the surface of an object. To simplify the optical setup is required for producing more compact equipment. An imaging head with four imaging devices was developed to producing more compact equipment by authors. The optical axis of each imaging sensor was directed to the object. Each imaging device can detects the phase difference distribution on the surface of the specimen. The four phase differences gives three directional displacement distributions and two directional strain distributions. In this paper, a prototype of a displacement and strain distribution measurement system with the multiple imaging devices is developed. The evaluation of the displacement and strain measurement with the prototype are shown.

Keywords Displacement and strain measurement • Multiple imaging devices • Phase-shifting digital holography • Sensitivity matrix

40.1 Introduction

The development of compact and conventional strain distribution measurement equipment for practical use is required for the evaluation of the health monitoring and the life-lengthening characteristics of infrastructures, such as steel bridges. Phase-shifting digital holography is a convenient method that can be used to measure displacement and strain distributions on the surface of an object [1]. Many researchers are studying this method, and several compact types of equipment are being developed [2–8]. Authors have also developed compact types of equipment for strain distribution measurements [9, 10]. However, several object waves are necessary to measure the in-plane displacement and strain distributions in conventional equipments. Simplifying the optical setup is required to produce more compact equipment. To simplify the optical setup is required for producing more compact equipment.

We proposed a method to simplify the optical setup using multiple imaging sensors with off-axis optical setup [11, 12]. The imaging sensors were placed in parallel. However, the resolution of the phase analysis was not high because this method uses the second-order reconstructed images. Recently, an imaging head with four imaging devices was developed to producing more compact equipment by authors. The optical axis of each imaging sensor was directed to the object. Each imaging device can detects the phase difference distribution on the surface of the specimen. The four phase differences gives three directional displacement distributions and two directional strain distributions.

_____

M. Fujigaki (✉) • Y. Murata
Faculty of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama 640-8510, Japan
e-mail: fujigaki@sys.wakayama-u.ac.jp

H. Minamino
Graduate School of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama 640-8510, Japan

In this paper, a prototype of a displacement and strain distribution measurement system with the multiple imaging devices is developed. The evaluation of the displacement and strain measurement with the prototype are shown.

40.2 Principle of Displacement and Strain Measurement with Multiple Imaging Sensors

Figure 40.1 shows a positional relationship among a light source, imaging devices and sensitivity vectors. The angle between an incident wave into an object and the scattered waves to the direction for image sensor i is shown as θ_i in this figure. The direction of the sensitivity vector e_i is the direction of the bisected angle of θ_i. The vector d is the displacement vector at the point P as shown in Fig. 40.1, when a point P on an object is displaced into the point P'. The phase difference $\Delta\phi_i$ obtained with a digital holographic interferometry is expressed as follows.

$$\Delta\phi_i = e_i \cdot d \tag{40.1}$$

The displacement vector and the sensitivity vector have each component of the x, y, and z directions. Thus, they are expressed as follows:

$$\Delta\phi_i = \begin{pmatrix} e_{ix} & e_{iy} & e_{iz} \end{pmatrix} \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = e_{ix}d_x + e_{iy}d_y + e_{iz}d_z, \tag{40.2}$$

where $e_i = (e_{ix}, e_{iy}, e_{iz})$ and $d = (d_x, d_y, d_z)$.

When image sensors 1, 2, ... N are places in the independent positions, N sensitivity vectors are obtained. In this case, Eq. (40.2) can be rewritten as Eq. (40.3),

$$\begin{pmatrix} \Delta\phi_1 \\ \Delta\phi_2 \\ \vdots \\ \Delta\phi_N \end{pmatrix} = S \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix}, \tag{40.3}$$

where subscripts 1, 2, ... N mean each imaging device. The matrix S called a sensitivity matrix is defined as Eq. (40.4).

$$S = \begin{pmatrix} e_{1x} & e_{1y} & e_{1z} \\ e_{2x} & e_{2y} & e_{2z} \\ \vdots & \vdots & \vdots \\ e_{Nx} & e_{Ny} & e_{Nz} \end{pmatrix} \tag{40.4}$$

Each element of the matrix S is constant number provided by the positional relationship between a light source and the corresponding imaging device. When the pseudo inverse matrix S^+ of the sensitivity matrix S is defined as Eq. (40.5), Eq. (40.3) can be rewritten as Eq. (40.6).
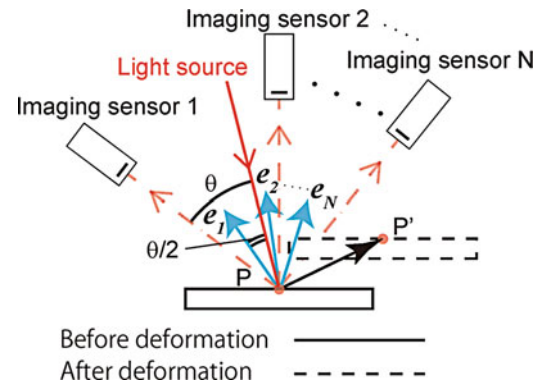

Fig. 40.1 Positional relationship among a light source, imaging devices and sensitivity vectors

$$S^+ = \begin{pmatrix} f_{1x} & f_{2x} & \cdots & f_{3x} \\ f_{1y} & f_{2y} & \cdots & f_{3y} \\ f_{1z} & f_{2z} & \cdots & f_{3z} \end{pmatrix} \tag{40.5}$$

$$\begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = S^+ \begin{pmatrix} \Delta\phi_1 \\ \Delta\phi_2 \\ \vdots \\ \Delta\phi_N \end{pmatrix} \tag{40.6}$$

This equation means that the displacement components d_x, d_y and d_z can be obtained from phase difference $\Delta\phi_1, \Delta\phi_2, \ldots \Delta\phi_N$ obtained by each image sensors. Displacement components d_x, d_y and d_z for the directions x, y and z, respectively, are expressed as follows;

$$d_x = \sum_{k=1}^{N} f_{kx}\Delta\phi_k, \quad d_y = \sum_{k=1}^{N} f_{ky}\Delta\phi_k, \quad d_z = \sum_{k=1}^{N} f_{kz}\Delta\phi_k. \tag{40.7}$$

Strains e_x and e_y are obtained by performing partial differentiation for them as follows;

$$\varepsilon_x = \frac{\partial d_x}{\partial x} = \sum_{k=1}^{N} f_{kx}\frac{\partial \Delta\phi_k}{\partial x}, \quad \varepsilon_y = \frac{\partial d_y}{\partial y} = \sum_{k=1}^{N} f_{ky}\frac{\partial \Delta\phi_k}{\partial y}. \tag{40.8}$$

That is, strain distributions can be derived from the phase difference distributions obtained with imaging sensor $1, 2, \ldots N$.

40.3 Displacement and Strain Measurement Device with Multiple Imaging Head

Figure 40.2 shows a prototype of strain measurement system with multiple imaging head. The imaging head has 4 CMOS imaging sensors (Omnivision, OV6930). The pixel size of this sensor is 400×400 pixels.

A glass plate is placed between an imaging head and an object. A part of an incident wave is reflected on the surface of the glass plate and the reflected wave arrives respective imaging sensors as reference waves. This glass plate is attached on PZT devices. The phase of the reference wave is shifted with changing voltage applied to the PZT devices. The voltage is also applied to the other 4 PZT devices connected in series. A mirror attached on the PZT devices is placed in an interferometer as shown in Fig. 40.2b. This interferometer makes a signal of an imaging trigger. While a period of phase-shift of a reference wave, four times of imaging trigger signals appear.
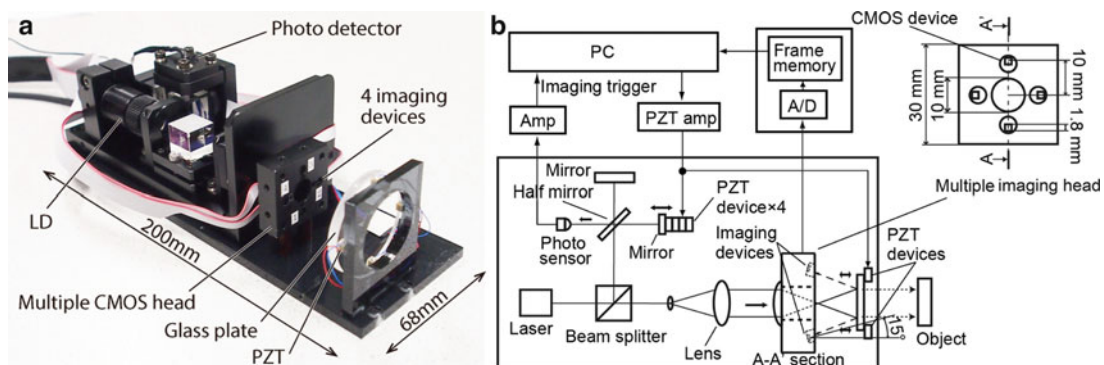

Fig. 40.2 Prototype of strain measurement system with multiple imaging head (**a**) Photograph (**b**) Block diagram

40.4 Experiment

40.4.1 Displacement Measurement

An experiment of displacement measurement using the optical setup of the prototype shown in Fig. 40.2. A He-Ne laser with 5 mW is used as a light source. An aluminum plate placed on a 3-axes PZT stage is used as a specimen as shown in Fig. 40.3a. Figure 40.3b shows a photograph of the specimen. The size of the analysis area is 5 mm × 5 mm. The outside area of the analysis area is covered with a piece of black flocked paper. The movable area can be shifted into the any direction with the 3-axes PZT stage.

In this experiment, the amount of moving is 200, 400 and 100 nm for the *x*, *y* and *z*-direction, respectively. Figure 40.4 shows phase difference distributions before and after moving the specimen obtained with a windowed PSDHI [13] with 64 windows. Figure 40.5 shows displacement distributions obtained from the phase difference distributions with Eq. (40.7). The sensitivity matrix is obtained using a movable plate attached on the 3-axes PZT stage in advance.

Table 40.1 shows the result of displacement measurement. The measured displacements are given as the difference between the fixed part and the movable part. The errors for the *x*, *y* and *z* directions are -10 nm, 25 nm and 6 nm, respectively.

Fig. 40.3 Experimental setup (**a**) Prototype of strain measurement system and specimen (**b**) Photograph of specimen
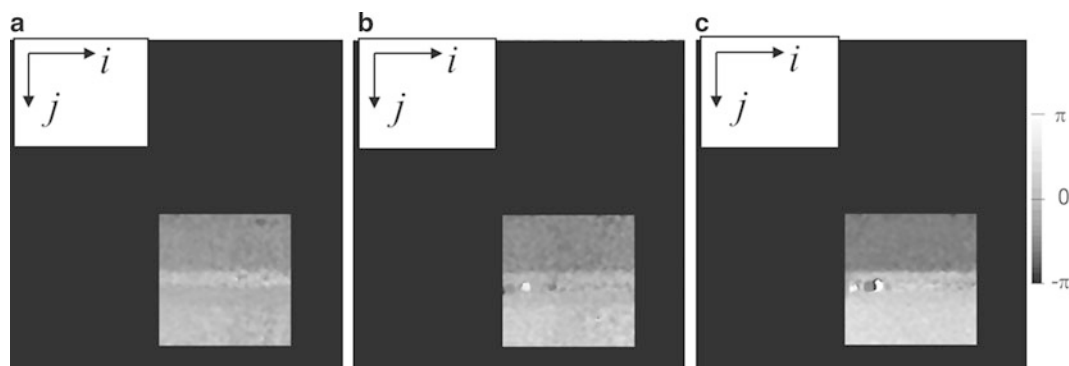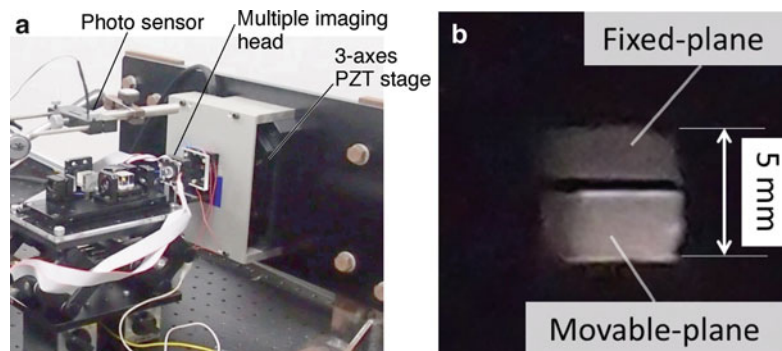

Fig. 40.4 Phase difference distributions (**a**) Sensor 1 (**b**) Sensor 2 (**c**) Sensor 3
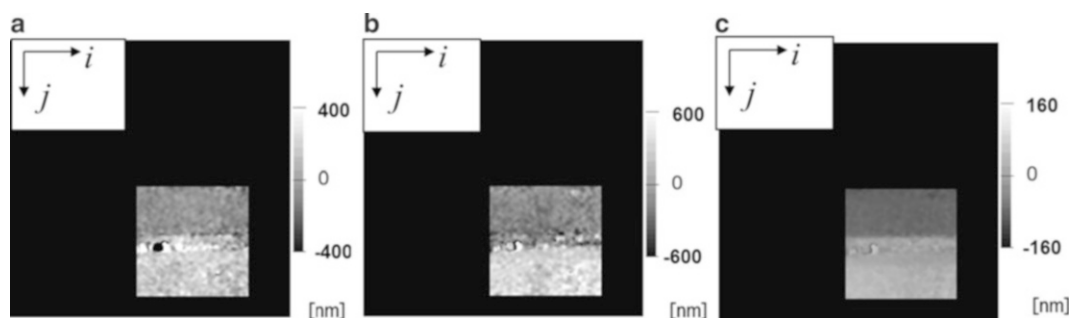
Fig. 40.5 Displacement distributions (**a**) *x*-direction (**b**) *y*-direction (**c**) *z*-direction

Table 40.1 Results of displacement measurement (Unit: nm)

	x	y	z
Given displacement	200	400	100
Measured displacement	190	425	106
Error	−10	25	6
Standard deviation	29	46	4

Fig. 40.6 Specimen (Cantilever) (**a**) Photograph (**b**) Sketch
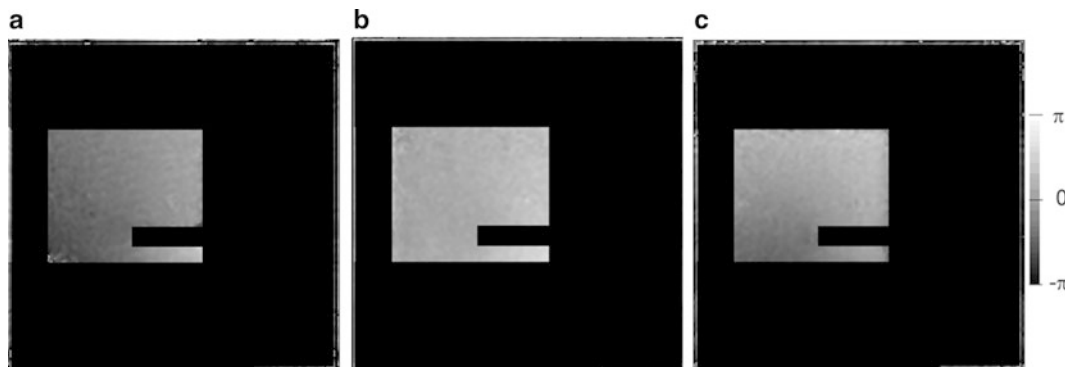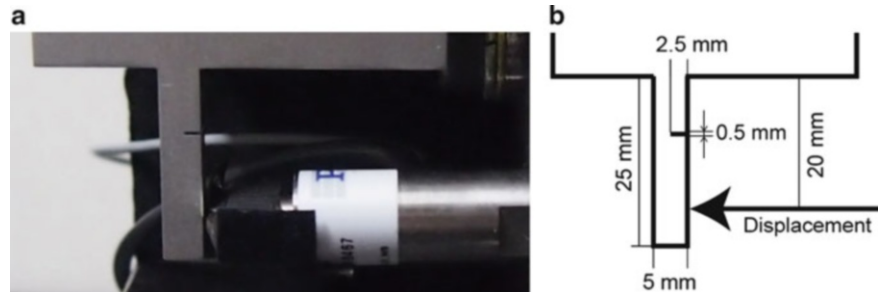
Fig. 40.7 Phase difference distributions (**a**) Sensor 1 (**b**) Sensor 2 (**c**) Sensor 3

Fig. 40.8 Strain distributions (**a**) *x*-direction (**b**) *y*-direction
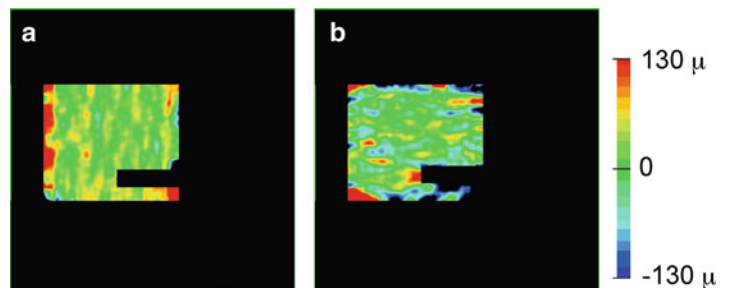
40.4.2 Strain Measurement

A cantilever shown in Fig. 40.6 is used as a specimen for an experiment of strain measurement. This cantilever has a cutout with 0.5 mm width and 2.5 mm length as shown in Fig. 40.6b. Displacement of 10 μm is applied at the point of 20 mm apart from fixed end by a PZT actuator.

Figure 40.7 shows phase difference distributions before and after deformation. Figure 40.8 shows strain distributions obtained from the phase difference distributions with Eq. (40.8). An area near the tip of the cutout in the *y*-directional strain distribution shown in Fig. 40.8b has large strain.

40.5 Conclusions

A prototype of a displacement and strain distribution measurement system with the multiple imaging devices was developed. Experimental result of the displacement measurement with a flat plate moved by a 3-axes PZT stage was shown. It was confirmed the effectiveness of the prototype system and the displacement measurement method using multiple imaging device because the error of displacement measurement was small. This prototype was applied to a strain measurement of a deformed cantilever with a cutout.

References

1. Yamaguchi I, Zhang T (1997) Phase-shifting digital holography. Opt Lett 22–16:1268–1270
2. Kolenovic E, Osten W, Klattenhoff R, Lai S, Von Kopylow C, Jüptner W (2003) Miniaturized digital holography sensor for distal three-dimensional endoscopy. Appl Opt 42–25:5167–5172
3. Michalkiewicz A, Kujawinska M (2004) Development of a miniature digital holographic interferometry system. Proc SPIE 5484:467–474
4. Pedrini G, Alexeenko I (2004) Miniaturized optical system based on digital holography. Proc SPIE 5503(5503):493–498
5. Kebbel V, Kolenovic E, Klattenhoff R, Jüptner W (2005) Miniaturized optical sensor for mechanical testing in industrial environment. In: Proceeding of the European conference on spacecraft structures, materials and mechanical testing 2005 (ESA SP-581), 1561–1568
6. Michalkiewicz A, Kujawinska M, Marc P, Jaroszewicz LR (2006) Miniaturized low-cost digital holographic interferometer. Proc SPIE 6188:618806
7. Saucedo AT, Santoyo FM, De la Torre-Ibarra M, Pedrini G, Osten W (2006) Endoscopic pulsed digital holography for 3D measurements. Opt Express 14–4:1468–1475
8. Michalkiewicz A, Kujawinska M, Stasiewicz K (2008) Digital holographic cameras and data processing for remote monitoring and measurements of mechanical parts. Opto Electron Rev 16–1:68–75
9. Fujigaki M, Matui T, Morimoto Y, Kita T, Nakatani M, Kitagawa A (2005) Development of real-time displacement measurement system using phase-shifting digital holography. In: Proceedings of the 5th international conference on mechanics of time dependent materials (MTDM05), pp 160–163
10. Fujigaki M, Kido R, Shiotani K, Morimoto Y (2008) High-speed and compact strain measurement system by phase-shifting digital holography. In: Proceedings of the international symposium to commemorate the 60th anniversary of the invention of holography, pp 316–323
11. Fujigaki M, Shiotani K, Nishitani R, Masaya A, Morimoto Y (2009) Off-axis reconstruction method for displacement and strain distribution measurement with phase-shifting digital holography. In: Osten W, Kujawinska M (eds) Fringe09, 6th international workshop on advanced optical metrology. Springer, New York, pp 764–769
12. Fujigaki M, Nishitani R, Morimoto Y (2010) Strain measurement using phase-shifting digital holography with two cameras. In: Proceedings of ICEM 14, EPJ web of conferences, 6, 30001(CD-ROM)
13. Morimoto Y, Matsui T, Fujigaki M, Kawagishi N (2007) Subnanometer displacement measurement by averaging of phase-difference in windowed digital holographic interferometry. Opt Eng 46-2, 025603

Chapter 41
Simultaneous ESPI Measurements Using Multiple Wavelengths and a Color Camera

Guillaume Richoz and Gary S. Schajer

Abstract ESPI is typically done using a single monochromatic light source with a monochrome camera. This arrangement enables full-field measurements of a single deformation quantity according to the particular arrangement of the optical system. If a further deformation quantity is to be measured, then the associated optical arrangement must be used sequentially. Here, an alternative approach is described where multiple interferometric measurements are simultaneously made using a color camera imaging monochromatic light sources of different wavelengths. The Red-Green-Blue (RGB) sensors of a conventional Bayer type camera can be read separately, thereby providing three independent color signals and independent ESPI phase maps. An example case is demonstrated here where two Michelson interferometers are combined to form a shearography camera with surface slope sensitivity in two perpendicular directions. By the use of two colors, both surface slopes can be measured simultaneously. The availability of the two surface slopes gives the opportunity for the data to be integrated numerically to give the surface displacement shape. This application is of significant practical interest because the surface displacement measurement can be made under field conditions by taking advantage of the well-known stability of shearography measurements.

Keywords ESPI • Shearography • Color camera • RGB • Surface slopes • Surface displacement shape

41.1 Introduction

Electronic Speckle Pattern Interferometry (ESPI) is an imaging technique that uses coherent laser light to visualize and quantify small surface displacements of an object [1, 2]. The technique has been applied to the solution of a wide range of problems in strain/stress and vibrational analysis, Non-Destructive Testing (NDT), quality control, design validation and optimization, and in fluid flow visualization [3–5]. With modern data processing systems, a large amount of experimental data can be processed and displayed effectively in real time.

ESPI is a very versatile method for measuring surface deformations and many different deformation quantities can be measured depending on the optical configuration used. Typical measured quantities are in-plane displacements, out-of-plane displacements, out-of-plane slopes and in-plane strains. In all these cases a separate optical configuration is needed for each measurement; if further measurements are required then further optical systems are needed.

Multi-measurement ESPI experimental setups have already been studied [3, 6]. These involve using multiple optical configurations that are activated sequentially to measure each desired quantity separately. This approach works well for measurement of static deformations, but is not well suited to dynamic measurements. In addition, the measurements are typically made using several separate cameras, and these must be carefully aligned to ensure they record consistent measurement areas.

The use of a color camera allows simultaneous and spatially aligned measurements of multiple independent signals, one for each measured color. This is possible because the sensor of a color camera contains interlaced sets of color sensitive pixels, Red, Green, and Blue (RGB) in the conventional Bayer arrangement [7]. Each of the colored pixels can be read separately, thus allowing independent and simultaneous measurement of three different colored speckle images.

G. Richoz • G.S. Schajer (✉)
Department Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada
e-mail: schajer@mech.ubc.ca

Such simultaneous measurement is particularly valuable for surface vibration measurements. Techniques exist for making ESPI measurements from a single optical image [8], thereby allowing high-speed sequential acquisition of ESPI data. Now, three independent measurements can be made simultaneously, so automatically making them coordinated in time. In addition, the three speckle images are mutually aligned on the Bayer sensor, so also automatically making them coordinated in space. All of these features extend the range of use of ESPI measurements and simplify optical design.

A further use of multiple color measurements is to identify absolute phase. Interferometry measurements using a single color can identify optical phase only within an angular range 0–2π. Phase angles outside that range "wrap" at intervals of 2π, thus causing an ambiguity where a phase angle of π cannot be distinguished from 3π, 5π, etc. This ambiguity could be resolved using simultaneous measurements with light sources of different wavelengths. The 2π wrapping occurs at different spatial intervals for the two phase maps, from which it will be possible to resolve the 2π ambiguity and determine absolute phase. Although straightforward in concept, such multi-color measurements are awkward to achieve with conventional ESPI systems, requiring sequential measurements and possibly a second camera. However, with the proposed method, the required measurements could be made simultaneously using a single color-camera.

This paper illustrates the use of multiple separate-color measurements using shearography [3] as an example. This technique measures out-of-plane rotations within the measured area. It is a popular technique because its parallel path optical configuration and its low sensitivity to rigid-body motions makes it relatively stable compared with other ESPI measurement types [9, 10] and thus more practical for field use. A multi-color shearography setup for full-field quantitative measurements of slope in two orthogonal directions is presented here. The setup uses two of the three color measurements available within a Bayer type camera. The third available color measurement will be used to determine absolute phase. The construction of the proposed shearography system is described and its capabilities are demonstrated through some example measurements.

41.2 Two-Axis Shearograpy Example

ESPI is a very flexible technique and it can be used to measure many different deformation quantities, depending on the optical geometry used. Figure 41.1 shows an example arrangement using a modified Michelson interferometer. This is a common setup for doing shearography [1–3]. In this arrangement, a camera views an object illuminated by coherent monochromatic light source through the interferometer. The configuration of the beam splitter and mirrors provides two alternative optical paths. If the components were exactly aligned, the two paths would recombine to image the same points on the object. However, a deliberately induced tilt in one of the mirrors causes a lateral shift in one path so that the paths do not exactly recombine; instead the camera views two slightly displaced images. In this way, each pixel on the camera receives light from two adjacent points on the illuminated object. The light from the pairs of points interfere on the camera pixels to produce the measured speckle pattern. For this optical arrangement the measured phase map corresponds to the out-of-plane surface slope in the direction of the image shift.

Many techniques can be used to retrieve the phase map from ESPI speckle patterns [11]. The phase-stepping method is straightforward to implement and is chosen here. In phase-stepping, a set of N phase-stepped speckle patterns is recorded. Typically the required phase stepping is done using a piezo actuator that moves a mirror within one of the optical paths, thereby varying by controlled amounts the path length and hence the received phase. The shearography arrangement in Fig. 41.1 uses this method. Commonly, N = 4 steps of $\lambda/4$ are used.

The shearography technique in Fig. 41.1 measures the surface slope in the image shearing direction determined by the direction of the mirror tilt used. If measurement of a surface slope in a different direction is required, either the mirror tilt
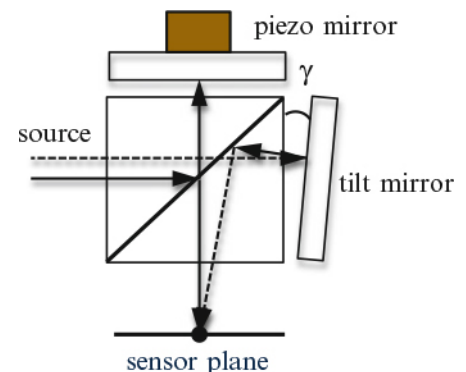

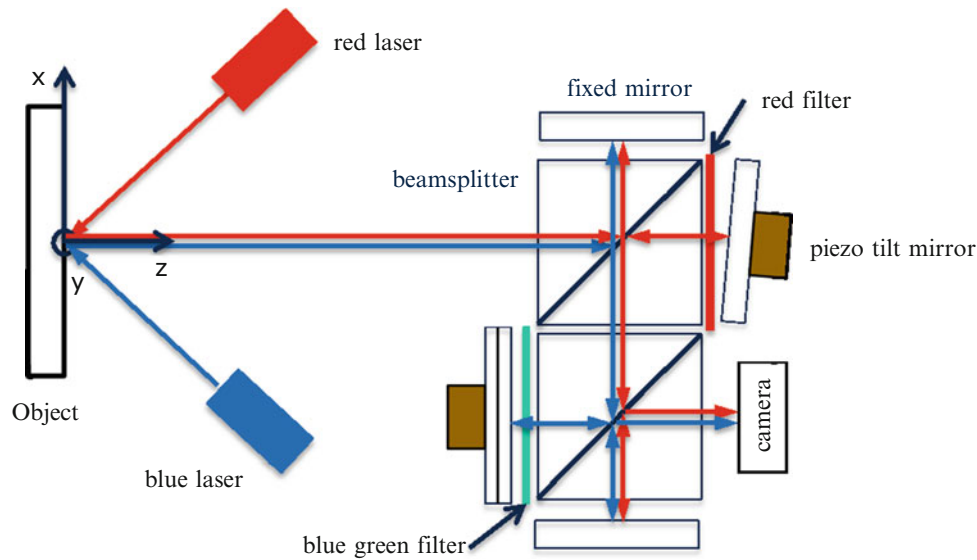**Fig. 41.1** Example ESPI arrangement for doing shearography

Fig. 41.2 Double shearography interferometer using two-color illumination

would have to be changed and a further measurement made, or else a second interferometer system containing the desired mirror tilt would have to be used in parallel. Both options are awkward.

The use of a double shearing interferometer using multiple coherent light sources of different wavelengths enables simultaneous interferometry in two image shearing directions. Figure 41.2 illustrates the proposed arrangement. Two modified Michelson interferometers are arranged in series, one with a mirror tilted in the x-direction on the specimen, and the second with a mirror tilted in the y-direction. Under ordinary circumstances this would result in the camera imaging four sheared images. However, the proposed arrangement includes an additional red filter to pass only red light through the stepped mirror of the first interferometer, and a blue-green filter to pass only blue and green light through the stepped mirror of the second. Thus, illumination by red light will display only the shearing in the x-direction, while illumination by green or blue light will display only shearing in the y-direction. This concept is developed here, where the target object is simultaneously illuminated by three coherent laser sources of different wavelengths, corresponding to red, green and blue. The light passes from the object through the double interferometer to a color camera that is able to record each color separately. This color separation allows separate measurement of the two shearing directions and hence separate identification of the surface slopes. Green and blue could be used redundantly to measure the same surface slope in order to identify the absolute phase via the "Two-Wavelength" method.

41.3 Color Measurements

The sensor of a typical color camera consists of a grid of pixels on which color filters have been fixed in a mosaic pattern. The Bayer pattern [7] is the most common design, with a sequence of 50 % green pixels and 25 % each of red and blue pixels. Figure 41.3 shows the transmission curves for the pixel color filters for a typical color camera. These filters are designed to be wide and have significant overlaps so that the camera is able to represent realistically the full range of colors present in a general image. However, for the present purpose a much tighter color response is needed. For good results, each color measurement should be entirely independent; overlap ("cross-talk") between color response is to be avoided. This behavior can be achieved by calibration. For example, consider illumination by a beam of color 1 (red) only, with average intensity I_1. The measured responses of the Red, Green and Blue pixels are $R_1, \; G_1, \; B_1$ where:

$$R_1 = A_{11} I_1 \quad G_1 = A_{21} I_1 \quad B_1 = A_{31} I_1 \tag{41.1}$$

Fig. 41.3 Absolute quantum efficiency graph of the Prosilica GE680C color camera—Allied Vision Technologies
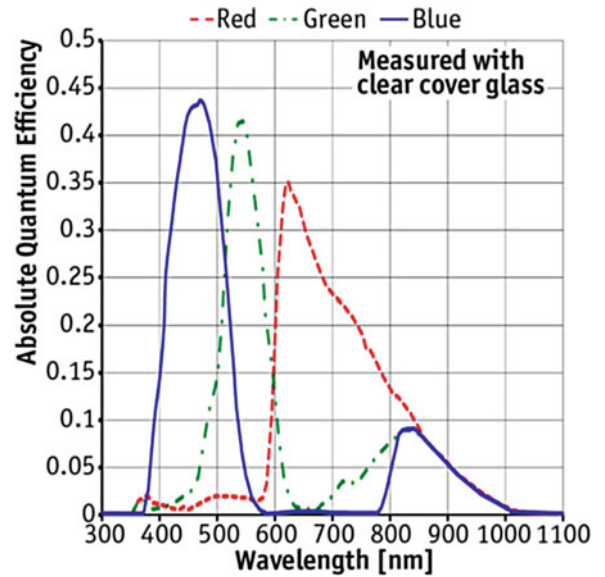
and A_{ij} is the response of pixel color i to illumination color j. In general, the numerical value of I_1 is not known. However, it is sufficient for the present purpose to work in terms of scaled quantities and to normalize the calibration coefficients A_{ij}:

$$a_{11} = \frac{A_{11}}{max(A_{11} + A_{21} + A_{31})} \quad a_{21} = \frac{A_{21}}{max(A_{11} + A_{21} + A_{31})} \quad a_{31} = \frac{A_{31}}{max(A_{11} + A_{21} + A_{31})} \quad (41.2)$$

Thus, the largest of a_{11}, a_{21}, a_{31} equals 1. Similarly, with beams of colors 2 or 3 only, subscript "2" or "3" replaces subscript "1" above. When all three beams are present simultaneously, the measured responses are:

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} R_1 + G_1 + B_1 \\ R_2 + G_2 + B_2 \\ R_3 + G_3 + B_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} \quad (41.3)$$

which becomes in matrix form:

$$aI = m \quad (41.4)$$

The individual beam intensities I can be recovered from the measured responses m by matrix solution. The scaling in Eq. (41.2) causes the computed I values to be in the same range as the original pixel readings.

41.4 Phase Stepping

A common five steps scheme [11] is used here with $\psi_n = [0°, 90°, 180°, 270°, 360°]$ as step angles. Using these values, the phase difference between the illumination and reference beams is:

$$\tan \theta = \frac{7(I_2 - I_4)}{4\,I_1 - I_2 - 6\,I_3 - I_4 + 4\,I_5} \quad (41.5)$$

The phase difference is extracted according to Eq. (41.5) while the object under analysis is in its reference state. Then, the same sequence of stepped measurements is made while the object is in the deformed state. The phase change between the initial and deformed states can be determined by subtraction of the phase values at each pixel and a phase change map can be formed. For practical use the phase map can be further processed by filtering [12] and unwrapping [13].

ESPI measurements implemented with phase-stepping for phase determination traditionally need only one coherent light source, an interferometer, and one monochrome camera. The use of multiple colors in the ESPI technique requires simultaneous determination of the phases of the corresponding colors. The red, green, and blue colors used have greatly different frequencies and so do not interfere with each other. In other words, working these colors allows simultaneous measurement of three independent phase differences. The ESPI setup used in this research uses a red laser of 660 nm, a green laser of 532 nm, and blue laser of 473 nm. A 90° phase step on the green color will not be equivalent of a 90° phase step on a red or blue color. Here, the green signal is taken as a reference and is stepped at a phase of 90°, thus stepping the red signal at a lower angle reduced by the ratio of the green and red wavelengths $\psi_n = [0°, 90°, 180°, 270°, 360°] \times 532/660$ Similarly, the stepping of the blue signal is increased by the ratio of the green and blue wavelengths $\psi_n = [0°, 90°, 180°, 270°, 360°] \times 532/473$. For these non-orthogonal steps, the phase can be determined by solving:

$$
\begin{bmatrix}
n & \sum \sin \psi_i & \sum \cos \psi_i \\
\sum \sin \psi_i & \sum \sin^2 \psi_i & \sum \sin \psi_i \cos \psi_i \\
\sum \cos \psi_i & \sum \sin \psi_i \cos \psi_i & \sum \cos^2 \psi_i
\end{bmatrix}
\begin{bmatrix}
C_1 \\
C_2 \\
C_3
\end{bmatrix}
=
\begin{bmatrix}
\sum I_i \\
\sum \sin \psi_i \; I_i \\
\sum \cos \psi_i \; I_i
\end{bmatrix}
\tag{41.6}
$$

where

$$
\theta = \arctan \left(\frac{C_2}{C_3} \right) \tag{41.7}
$$

Equations (41.6) and (41.7) reduce to Eq. (41.5) for the case where $\psi_n = [0°, 90°, 180°, 270°, 360°]$. For the general case of a non-orthogonal angle ψ, the format of Eq. (41.5) is preserved, but with non-integer coefficients. Thus, the use of non-orthogonal step angles involves no extra mathematical burden. A five-step scheme was chosen here rather than the conventional four-step method so that a substantial fraction of a wavelength is covered even for steps less than 90°. This helps maintain phase evaluation accuracy or non-orthogonal steps.

41.5 Results

An initial test of the color separation method of Eq. (41.4) was made. The three lasers red $\lambda = 660$ nm, a green $\lambda = 532$ nm, and blue $\lambda = 473$ nm were sequentially used to illuminate a white test specimen. Measurements were made using a Prosilica GE680C color camera (Allied Vision Techologies, Inc, Burnaby, Canada). The raw individual color images from the camera showed cross-talk between color measurements in the 5–15 % range. After application of Eq. (41.4), the cross-talk was reduced to around 1 %. Thus, the matrix inversion method for calibrating the color separation was found to be effective.

After completion of the color calibration, a test was done to assess the ability of the color camera to do simultaneous ESPI measurements. This was done by using a single-axis shearography interferometer as shown Fig. 41.1 to measure the surface slope of the overhanging beam shown in Fig. 41.4. Two lasers, red and blue simultaneously illuminated the same beam area between x_1 and x_2 in Fig. 41.4b. A micrometer screw applied a sequence of 8×1 mm displacement increments at the right side of the beam in Fig. 41.4a. This displacement-controlled loading was chosen to avoid the need to know the beam material Young's modulus precisely when determining the "theoretical" surface slopes to compare with the ESPI measured values. In addition, evaluation of the slope difference between two points x_1 and x_2 avoids the need to determine the absolute slope.

Figure 41.5a shows the wrapped phase map extracted from the blue image; it contains 15.4 vertical fringes. Figure 41.5b shows the corresponding wrapped phase map extracted from the red image; it comprises 11.6 fringes. The two phase maps measure the same deformation and therefore the fringe numbers should be in the inverse ratio of the wavelengths. Here, $N_{blue}/N_{red} = 1.33$ and $\lambda_{red}/\lambda_{blue} = 1.39$. In Fig. 41.5c, the Mean Absolute Percentage Error (MAPE) is around 2 % for the red and blue signal. The max Absolute Percentage Error (APE) is below 5 % for both colors.

A further measurement was done to illustrate two-direction shearography using the arrangement shown in Fig. 41.2. The same overhanging beam was used as in single-axis interferometer measurement. In this two-direction case, the measured fringes differ not only in their spacing but also in their orientation. The fringes in Fig. 41.6a indicate the surface slope in the longitudinal direction of the beam, with approximately 11 vertical fringes. The fringes in Fig. 41.6b indicate the surface slope in the transverse direction of the beam, with approximately three horizontal fringes. These fringes represent the surface anticlastic curvature of the beam, a phenomenon controlled by the action of Poisson's ratio [14]. The quality of these

Fig. 41.4 Overhanging beam. (**a**) plan view, (**b**) side view

$L = 100 \, [mm]$ $x_1 = 35.0 \, [mm]$
$a = 488.5 \, [mm]$ $x_2 = 68.8 \, [mm]$
$b = 25.4 \, [mm]$ $\delta x = 2.72 \, [mm]$
$t = 2.4 \, [mm]$ δx: shear distance
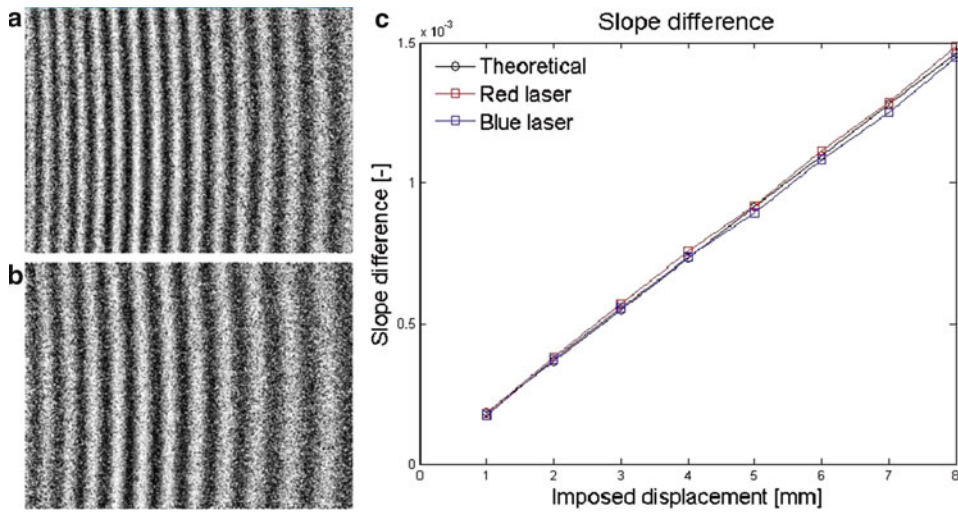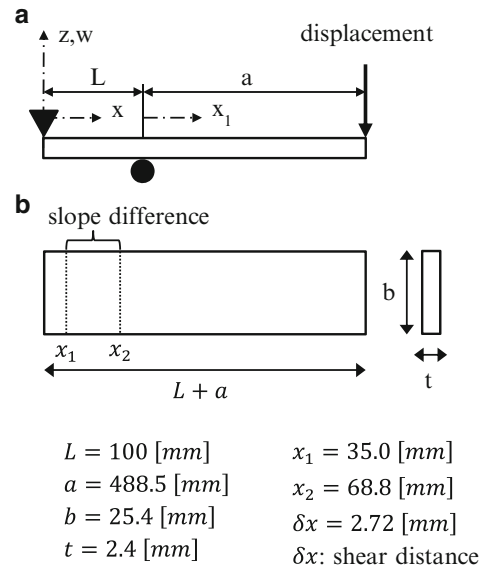
Fig. 41.5 Slope difference measurements in x direction using a single-axis interferometer (**a**) wrapped phase map *blue laser*, (**b**) wrapped phase map *red laser*, (**c**) slope difference results

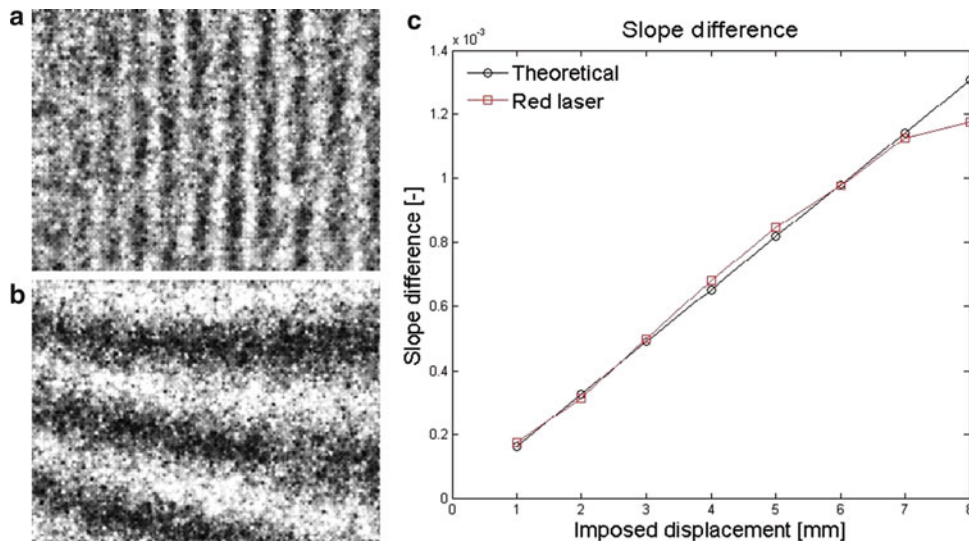**Fig. 41.6** Slope difference measurements with a two-axis interferometer (**a**) wrapped phase map with filtering with *red laser* $= \partial w/\partial x$, (**b**) wrapped phase map with filtering with *blue laser* $= \partial w/\partial y$, (**c**) slope difference results

initial measurements is rather poor, for reasons not yet clarified. Work is presently underway to understand and rectify this issue. In Fig. 41.6c, the mean absolute percentage error is around 4 % for the red results over the eight measurements. The max absolute percentage error remains below 10 % for the blue results.

41.6 Discussion

Simultaneous ESPI measurements using multiple wavelengths and a color camera is shown to be both possible and effective. The fringe patterns in Fig. 41.5 from the single-axis shearography arrangement in Fig. 41.1 show good contrast. However, the fringe patterns from the double-axis shearography arrangement in Fig. 41.2 are quite noisy. This is not an artifact of the use of multiple colors because the single-axis shearography results are clean. Although not yet fully investigated, part of the reason may be that only half of the light proceeds through each beamsplitter while half returns towards the light source. Thus, there is a significant loss in overall received light intensity. In addition, the retro-reflected light from the second beamsplitter interferes with the interferometry taking place in the first beamsplitter. These issues are presently being pursued.

The use of two colors enables simultaneous, full-field and non-contact measurements of two surface slopes. The resulting surface slopes $\partial w/\partial x$ and $\partial w/\partial y$ could be integrated to determine the surface displacement shape. Given the well-known stability of shearography measurements, this could provide a robust displacement measurement technique that is much less delicate than typical ESPI measurements.

The example use of multiple color measurements described here uses two of the three available colors, leaving the third color (green here) available for further use. A practical application for the third color is the determination of the absolute phase on one of the shear measurements, a feature that is typically not available when unwrapping ESPI phase maps because of the associated modulo 2π ambiguity. The two-wavelength method [15, 16] can be applied for absolute phase identification, for example using blue and green light in this case. For this to be done effectively, the wavelength difference between the two colors must be small, rather smaller than with the presently available RGB light sources. Additional light sources are being acquired.

41.7 Conclusions

The use of a color camera for doing simultaneous ESPI measurements has been demonstrated. This capability is possible by using the three separate color measurements available with a Bayer style camera. The technique is valuable because it allows the use of much simpler and more compact optical equipment and because it provides multiple independent measurements that are automatically coordinated in both time and space.

Some cross-talk exists between the raw measurements when using a typical Bayer style camera. Fortunately, it can the largely eliminated using a simple mathematical adjustment, thus producing three independent measurements. Example multiple simultaneous ESPI measurements have been demonstrated here using a standard single-axis shearography apparatus and independent ESPI fringe patterns were obtained in two different colors. Further example measurements were made using a double shearography apparatus designed to measure surface gradients in two directions. They were also effective, but the resulting fringe patterns have lower visibility because of internal reflections. All measurements were quantitatively in good agreement with the theoretically expected values. The results confirm that the multi-color technique is an effective tool for quantitative measurements. Use of multiple wavelengths and a color camera is not limited to shearography but opens the doors to simultaneous multi-measurement for most optical configurations of the ESPI family.

Work is presently underway to extend the use of a third color in order to identify the absolute phase. The numerical integration of both orthogonal surface slopes is ongoing. This will be described in a future report.

References

1. Leendertz JA (1970) Interferometric displacement measurement on scattering surfaces utilizing speckle effect. J Phys E Sci Instrum 3(3):214–218
2. Butters JN, Leendertz JA (1971) Speckle pattern and holographic techniques in engineering metrology. Opt Laser Technol 3(1):26–30
3. Steinchen W, Yang L (2003) Digital shearography: theory and application of digital speckle pattern shearing interferometry. SPIE Press, Bellingham, WA, p 310
4. Hack E, Riner M (2001) 3D ESPI and 3D shearography measurements applied to NDT and FEM analysis validation for industrial quality control. Proc SPIE 4398:155–167
5. Burnett M, Bryanston-Cross PJ (1996) Measurements of transonic shock structures using shearography. Proc SPIE 2861:124–135
6. Hack E, Kaestle RP (1999) Multi-wavelengths shearography for optical whole-field strain measurements. In: Colloque Nouveaux Moyens Optiques Pour L'Industrie II, Mittelwihr/Colmar (F), Proceedings, 17–19 November 1999, pp 31–36
7. Bayer BE (1976) Color imaging array. US patent 3,971,065, US Patent and Trademarks Office, Washington DC, 1976
8. Georgas P, Schajer GS (2012) Simultaneous measurement of full-field vibration modes using ESPI. Exp Mech 53(8):1461–1466
9. Rastogi P, Hack E (2012) Optical methods for solid mechanics: a full-field approach. Wiley, Chichester
10. Yang LX, Steinchen W, Schuth M, Kupfer G (1995) Precision measurement and nondestructive testing by means of digital phase shifting speckle pattern and speckle pattern shearing interferometry. J Int Meas Confed 16(3):149–160
11. Sirohi RS, Chau FS (1999) Optical methods of measurement: wholefield techniques. Dekker, New York
12. An Y, Schajer GS (2011) Pixel quality evaluation and correction procedures in ESPI. Exp Tech 35(5):23–29
13. Giglia DC, Pritt MD (1998) Two-dimensional phase unwrapping: theory, algorithms, and software. Wiley-Interscience, New York
14. Pomeroy RJ (1970) The effect of anticlastic bending on the curvature of beams. Int J Solids Struct 6(2):277–285
15. Schnars U, Jüptner W (1992) Messung von Absolutverformungen mit Methoden der holografischen Interferometrie und der Elektronischen Specklemuster-Interferometrie. Laser und Optoelektronik 24(1):59–63
16. Marwitz H (1990) Holografische Konturgenerierung, Chapter 18 in Praxis der Holografie—Grundlagen, Standard- und Spezialverfahren, Marwitz, H (ed.) Expert Verlag, Stuttgart

Chapter 42
Some Practical Considerations in High-Speed 3D Shape and Deformation Measurement Using Single-Shot Fringe Projection Technique

Minh Le, Zhaoyang Wang, and Hieu T. Nguyen

Abstract Fringe projection technique is one of the most widely used techniques for three-dimensional (3D) shape and deformation measurements by Wang Z, Nguyen DA, Barnes JC (2010) Some practical considerations in fringe projection profilometry. Opt Lasers Eng 48:218–225. Generally, it is a difficult task to conduct high-speed measurement with this technique because the speed of projecting fringes is restricted by the projector. In addition, the technique usually employs multiple-shot method such as the phase-shifting scheme, which can considerably cut down the measurement speed. To make the technique suitable for high-speed measurement, single-shot approach is normally adopted; however, a number of problems can arise, such as those associated with phase calculation, phase unwrapping, shadow and occlusion, discontinuity, multiple objects, accuracy, etc. In this paper, we propose some practical considerations to cope with the limitations in existing single-shot fringe projection technique for high-speed 3D shape and deformation measurements. Experimental supports are presented.

Keywords Fringe projection • Single-shot • High speed • Shadow and occlusion • Phase unwrapping

42.1 Introduction

Fringe projection techniques have been studied and used extensively in academia and industry due to their numerous advantages such as high accuracy and automated analysis [1]. The techniques normally require determination of the full-field fringe phases. At present, there are many algorithms available for calculating the phase distributions of the projected fringes on the objects, including but not limited to phase-shifting, Fourier transform (FT), Wavelet transform (WT). While multiple-shot techniques like the phase-shifting scheme can greatly suppress noise by using multiple images [2], the techniques are not suitable for dynamic measurements where objects quickly move or change their shapes. On the contrary, single-shot methods such as FT and WT algorithms require only a single image for each 3D reconstruction, so they are insensitive to vibration, motion and fast deformation. With single-shot algorithms, however, the noise and discontinuity issues encountered in real applications can bring substantial errors to the 3D shape measurement results. In this paper, we present a few considerations to cope with the limitations in existing single-shot fringe-projection-based 3D shape measurement techniques. We mainly focus on high-speed measurement, so a high-speed camera, Phantom v310, is used in our work. We also use FT algorithm as an example, but it can be easily extended to other algorithms such as WT one. Experimental support will be presented to verify the validity of the proposed approaches.

M. Le (✉) • H.T. Nguyen
Department of Electrical Engineering, The Catholic University of America, Washington,
620 Michigan Ave. N.E, Washington, DC 20064, USA
e-mail: 26letran@cardinalmail.cua.edu

Z. Wang
Department of Mechanical Engineering, The Catholic University of America, Washington,
620 Michigan Ave. N.E, Washington, DC 20064, USA

42.2 Principles

42.2.1 Image Acquisition

In this paper, we aim to measure the 3D shape of a dynamic object at high speed, so a Phantom v310 high speed camera with the speed up 10,000 fps at 640 × 480 pixels resolution is used to capture images. Also, the computer generates a sinusoid fringe pattern which is then projected on the reference plane and the object for measurement through a DLP projector. The captured images will afterwards be sent back to the computer to determine the phase map as well as the out-of-plane height distribution of the object. Noticeably, the DLP projector uses a RGB color wheel in order to display the fringe pattern. As a result, the high speed camera will capture a series of Red-Green-Blue images as well as transition images between Red, Green, and Blue channels. Since the three-channel images are basically identical, they will not bring a problem to the measurement. After that, the FT fringe analysis is carried out to find the natural phase of fringe projected on the object. The final result is finally displayed in 3D on the computer monitor. A schematic of the system is illustrated in Fig. 42.1.

42.2.2 Fourier Transform method

In 1983, Takeda et al. presented the Fourier transform method for fringe analysis [3]. In this method the intensity distribution of the fringe patterns $I(x, y)$ captured by a camera can be expressed as:

$$I(x,y) = Im(x,y) + Ia(x,y) \cos [\phi(x,y)] \qquad (42.1)$$

where $I(x, y)$ is the intensity distribution of the fringe patterns, $I_m(x, y)$ is the mean intensity, I_a is the intensity modulation amplitude, ϕ is the angular phase information of the interferogram, and (x, y) represents all the points in the x-y plane of the object and the interferogram.
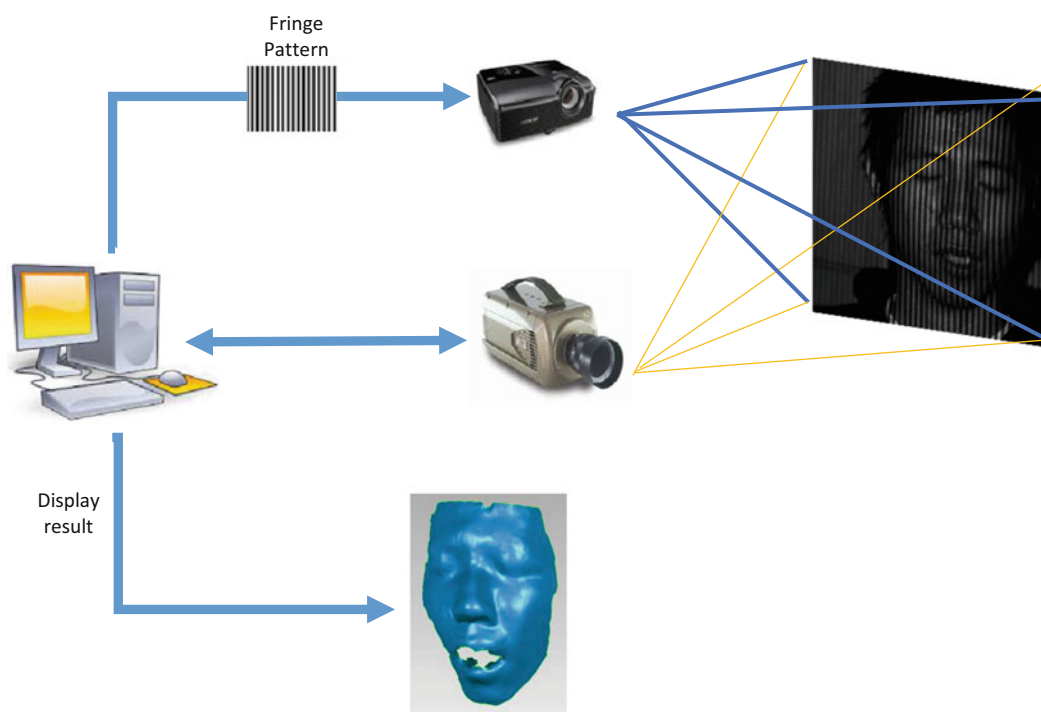

Fig. 42.1 Schematic of the system

The principle of FT is to extract the fundamental frequency component of fringe pattern in the frequency domain. Therefore, the captured intensity in Eq. (42.1) can also be expressed as:

$$I(x,y) = Im(x,y) + c(x,y) + c^*(x,y)$$ (42.2)

with:

$$c(x,y) = \frac{1}{2}I_a(x,y)e^{i\phi(x,y)}$$ (42.3)

where * denotes the complex conjugate.

After a two-dimensional Fourier transform (2D-FTP), the spatial frequency domain representation of the pattern becomes:

$$F(\zeta,\eta) = A(\zeta,\eta) + C(\zeta,\eta) + C^*(\zeta,\eta)$$ (42.4)

where $A(\zeta,\eta)$ is the transform of $Im(x,y)$ and $C(\zeta,\eta)$ and $C^*(\zeta,\eta)$ are the positive and negative frequency spectra of the modulated carrier fringes, ζ and η are the spatial frequencies that represent intensity changes with respect to spatial distances.

At the frequency domain, $C(\zeta,\eta)$ can be isolated, then perform an inverse Fourier transform for $C(\zeta,\eta)$. Finally, $c(x,y)$ can be obtained at the spatial domain and the phase information ϕ can be calculated from:

$$\phi = \arctan\frac{Im[c(x,y)]}{Re[c(x,y)]}$$ (42.5)

where $Im[]$, $Re[]$ represent the imaginary and real part of $c(x,y)$, respectively.

The wrapped phase achieved from Eq. (42.5) varies from $-\pi$ to π. The wrapped phase then needs to be converted to a continuous natural distribution by a phase unwrapping process in order to obtain the height distribution of the object.

## 42.3 Practical Considerations in Single Shot 3D Imaging

42.3.1 Shadows and Discontinuities

Generally, phase unwrapping refers to the process of recovering the unknown integral multiple of 2π to be added to the measured wrapped phase at each pixel. For an impeccable unwrapped phase process, it must be a trivial procedure and path independent. However, in practice it is not possible because of inevitable factors such as local shadows or low fringe modulation, irregular surface brightness, geometric discontinuities which are quite common features in the observed fringe patterns. As a result, the phase unwrapping process is path dependent and thus becomes a challenging problem, especially for complex shape measurement.

42.3.2 Reliability-Guided Map Phase Unwrapping Process

A challenge problem associated with the fringe projection systems is the unwrapping of phase maps resulted from complex object surface shapes [4]. In 2000, Su et al. proposed a phase unwrapping algorithm to cope with local shadows and discontinuities [5]. This algorithm is based on using a reliability guided parameter map which is obtained by applying digital weighted filtering in the frequency domain. Specifically, the modulation function is employed for phase unwrapping in phase measuring profilometry. Because the value of modulation function in the areas of local shadow and abrupt discontinuities is lower than the value in the other parts. So modulation can be a helpful guide to determine an optimized phase unwrapping path.

Fig. 42.2 Example of phase unwrapping process
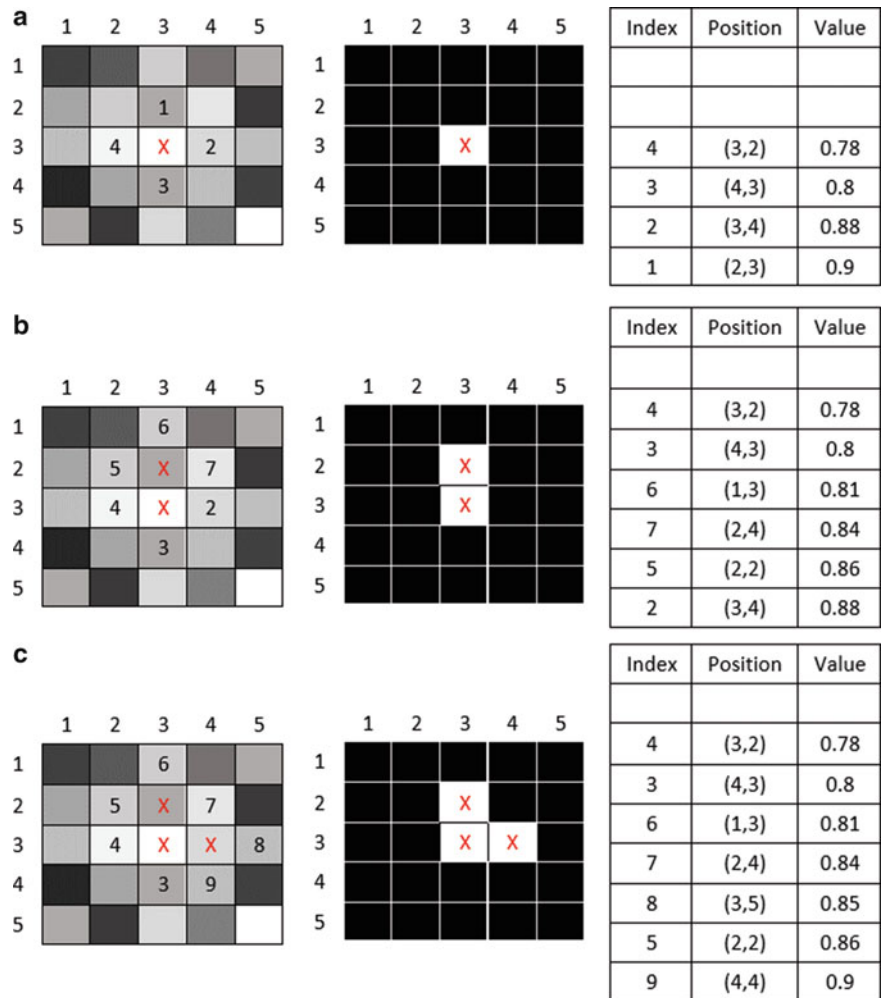
The modulation function can be calculated as:

$$M(x, y) = I_a(x, y) = \sqrt{(Im[c(x, y)])^2 + (Re[c(x, y)])^2} \tag{42.6}$$

In this algorithm, a binary mask is built to identify the wrapped and unwrapped points. The pixels to be unwrapped are placed in the queue in order. All initial value in the binary mask is set to be 0, representing that the corresponding phase is not be unwrapped. Later any point having been wrapped is set to be 1. The details of this algorithm is as follows [6]:

1. Figure 42.2a is an example of 5 × 5 pixels area. First, determine the maximum value of reliability parameter map, start the phase unwrapping process from this point. At the same time, mark the corresponding point in the binary mask as 1, then put its four neighboring points into a queue, and order them according the parameter value from higher to lower one, the pixel with maximum parameter value is put at the bottom of the queue.
2. Remove the pixel with maximum parameter value from the bottom of queue and begin its phase unwrapping on the basis of the starting point. If the difference between this point and the starting point is more than π; the natural phase of this point equals to its wrapped phase subtracting 2π: If the difference is less than π; the natural phase equals to its wrapped phase adding 2π. Finish phase unwrapping of this point according to this rule, then mark the corresponding point in the binary mask as 1, and put the points which have not been unwrapped among its four neighboring points into a queue, and order them from higher to lower one as in step (1) as shown in Fig. 42.2b.
3. Each time, remove the point with maximum value in the queue, and proceed its phase unwrapping on the basis of the very former unwrapped point. Then mark its corresponding point in the binary mask as 1, and put the points which have not been unwrapped among its 4 neighboring points into the queue again, order them from higher to lower one. Repeat step (1) and step (2), as shown in Fig. 42.2c.

When the queue is empty, which means that all the points have been unwrapped phase unwrapping is finished.

This approach makes the path of phase unwrapping be always along the direction from the pixel with higher parameter value to the pixel with low parameter value. Therefore, in the worst case the error is limited, if there is any, to local minimum areas.

42.4   Experimental Results

To verify the proposed idea in this paper, an experiment is conducted by capturing one of the authors' face along with the motion of the face. Specifically, we tried to build the 3D of the human face while the person is talking. Normally, it is hard to capture the 3D details of the motion of a talking face, but with the assistance from high speed camera and single-shot technique, the 3D motion of the face can be acquired clearly. In this experiment, the Phantom camera's speed was set at 800 fps with 512×512 pixels resolution. The captured fringe patterns are shown Fig. 42.3

After that, the proposed algorithm is applied to calculate the natural phase from captured images, and the 3D shapes are then reconstructed, as shown in Fig. 42.4.

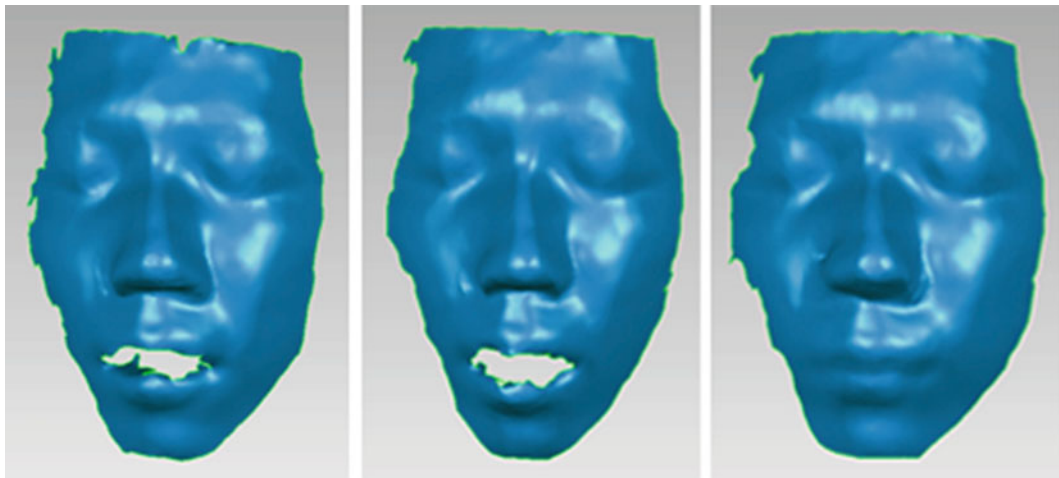**Fig. 42.3** The deformed fringe on the talking face

Fig. 42.4 3D shape measurement result of a talking face

42.5 Conclusion

This paper presents a few practical considerations in single-shot high speed 3D shape measurements, along with a simple yet robust solutions with high accuracy, fast speed and easy implementation capabilities. A quality map is generated to aid performing a reliable phase unwrapping process. The experimental result has shown the efficiency of this technique for measuring the 3D shape of a dynamic object.

References

1. Zhang Z (2012) Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques. Opt Lasers Eng 50:1097–1106
2. Burton DR, Lalor MJ (1994) Multichanel Fourier fringe analysis as an aid to automatic phase unwrapping. Appl Opt 33(14):2939–2948
3. Takeda M, Motoh K (1983) Fourier transform profilometry for the automatic measurement of 3-D object shapes. Appl Opt 22(24):3977–3982
4. Chen K, Xi J, Yu Y (2013) Quality-guided spatial phase unwrapping algorithm for fast three-dimensional measurement. Opt Commun 294:139–147
5. Xianyu S, Xue L (2001) Phase unwrapping algorithm based on fringe frequency analysis in Fourier-transform profilometry. Opt Eng 40(4):637–643
6. Xianyu S, Chen W (2004) Reliability-guided phase unwrapping algorithm: a review. Opt Lasers Eng 42:245–261

Chapter 43
Fast-Speed, High-Accuracy and Real-Time 3D Imaging with Fringe Projection Technique

Hieu Nguyen, Zhaoyang Wang, Hien Kieu, and Minh Le

Abstract In recent years, improving and extending the capabilities of the three-dimensional (3D) imaging to fast-speed and high-accuracy performance has been a very active research topic. Although the latest advances may help provide real-time and high-resolution measurements, to simultaneously achieve high-accuracy performance is still a big challenge. In this paper, we describe a novel 3D shape measurement and 3D imaging system based on using the fringe projection technique. It employs a generalized system setup, an ultrafast multiple-frequency phase-unwrapping algorithm, a flexible system calibration method, a robust gamma correction technique, and a graphical-programming-unit-based (GPU-based) data processing scheme to achieve fast-speed, high-accuracy and fast-speed 3D shape measurement and 3D imaging performance. The real-time system can operate at a speed of 45 3D-frames/second with relative accuracy of 1/2,500 or at a reduced speed of 10 3D-frames/second with enhanced accuracy of 1/10,000.

Keywords Fringe analysis • 3D imaging system • Phase shifting • 3D shape measurement • Real-time

43.1 Introduction

Optical three-dimensional (3D) shape measurement and geometry reconstruction has emerged as an important tool for obtaining 3D geometry information of the physical objects and the surrounding scenes. The relevant techniques have been employed in numerous practical applications of 3D shape measurement such as reserve engineering, industrial inspection and quality control, biomedical applications, entertainment and education applications [1–3]. The image-based full-field 3D imaging techniques mainly fall into two categories: stereo vision and structured-light-based techniques [4–7]. Although stereo vision is probably the most widely used technique and have the same schematic system of using two devices for 3D shape measurement, the time-consuming nature of the stereo-based method due to the matching of stereo images is still a big challenge for performing real-time 3D shape reconstruction. By replacing one of cameras with a projector and using simpler processing algorithms, the structured light system offers a few advantages of full-field performance, easy implementation and real-time features. The 3D imaging system based on fringe projection profilometry (FPP), is one of structured-light techniques and is well-known for achieving real-time performance in the long run time.

Once the FPP-based approach can reach the high speed, another notable issue is the system accuracy when it copes with complex shape or multiple objects following the ever-raising measurement standard for 3D imaging systems. The goal of this paper is introduce a novel accurate 3D shape measurement and 3D imaging system with the capabilities of high resolution and high speed performance. Specifically, the high speed and high accuracy performance can be archieved by employing hardware enhancements such as modified DLP projector, the syncroniztion between projector-camera pairs and

H. Nguyen (✉) • H. Kieu • M. Le
Department of Electrical Engineering and Computer Science, The Catholic University of America,
620 Michigan Ave. N.E, Washington, DC 20064, USA
e-mail: 29nguyen@cardinalmail.cua.edu

Z. Wang
Department of Mechanical Engineering, The Catholic University of America,
620 Michigan Ave. N.E,
Washington, DC 20064, USA

numerous advanced techniques like a generalized system set up, an effective gramma correction, an accurate camera calibration method, and amultiple-frequency fringes prjection technique. Subsequently, owing to the robust use of a graphics processing unit (GPU), the system can achieve real-time 3D imaging while keeping very high accuracy. Various experiments have been conducted to verify the validity and practicability of the approach, which give the great promise of satifying the 3D imaging demands in many scientific and engineering applications.

43.2 Hardware Enhancements

A basic FPP-based 3D imaging and shape measurement system usually contains a digital projector, a digital camera, and a computer. The basic concept of the FPP-based 3D shape measurement involves projecting a set of structured light patterns onto the objects of interest. Then, the height/depth information stored in the distorted fringe patterns is captured by high speed camera for post processing. For the purpose of easy analysis of the fringe patterns, straight sinusoidal patterns in a vertical or horizontal direction are usually employed to achieve full-field 3D shape measurements. In our work, a modified DLP projector and an external circuit board are introduced to the system to substantially boost the imaging speed. Figure 43.1 shows a geometrical illustration of the proposed system.

43.2.1 Customized DLP Projector for Pattern Illumination

A single-chip digital light processing (DLP) projector has an interesting mechanism that can be utilized to significantly speed up the 3D imaging performance. Inside the DLP projector, there is a color wheel between the lamp and the DLP chip. The color wheel is divided into four sectors, including three in primary colors of red (R), green (G), and blue (B), and a transparent colorless sector (clear sector) for enhancing the brightness of the projected image or video. The DLP chip is synchronized by the rotating motion of the color wheel, which allows the RGB channels of an image or a video frame to be projected sequentially at a sufficiently high rate so that human eyes can see a composite colorful image or video. The timeline for a projection running at a frequency of 90 Hz, i.e., each projection cycle lasts for 11.11 ms, is shown in Fig. 43.2.

To take advantage of the above projection mechanism, the color wheel is removed from the DLP projector; this leads to gray scale image projection for any color image. Meanwhile, a high speed CMOS camera with an exposure time of 3.66 ms is employed to capture each RGB channel of the image. Consequently, three different images can be captured in one cycle if they are encoded respectively in red, green, and blue channels and combined into a single source image. Figure 43.3 is a computer generated source image whose RGB channels are three uniformly-phase-shifted sinusoidal patterns. As a result, the projecting speed virtually becomes three times faster.
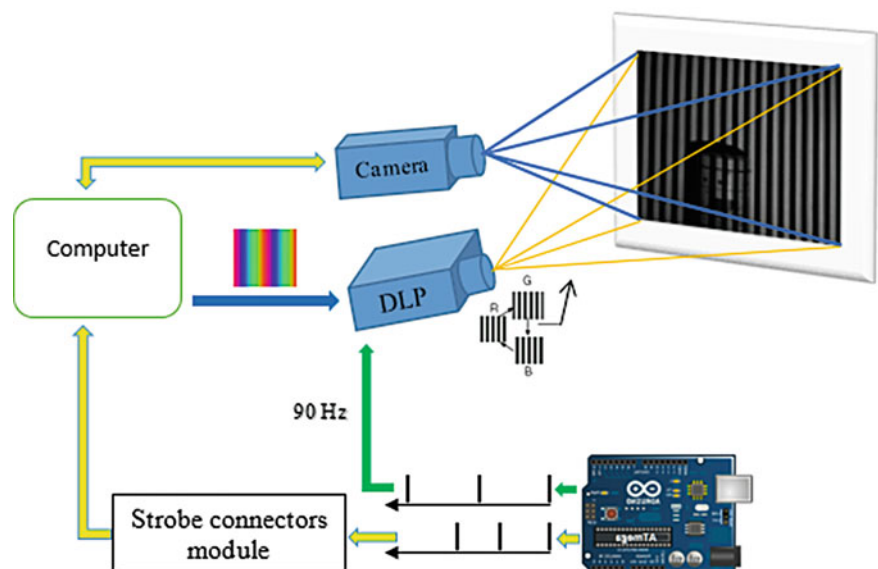

Fig. 43.1 Schematic of the proposed system
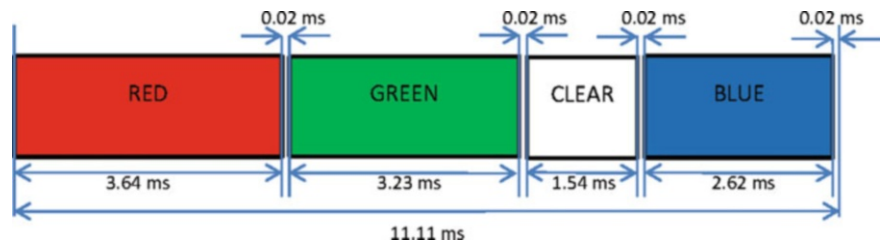
Fig. 43.2 Timeline for a projection rate of 90 Hz

Fig. 43.3 Composite image of three phase-shifted sinusoidal fringe patterns encoded in RGB channels
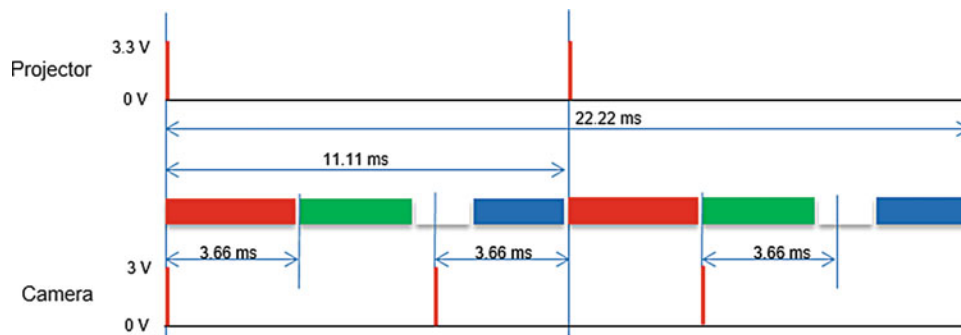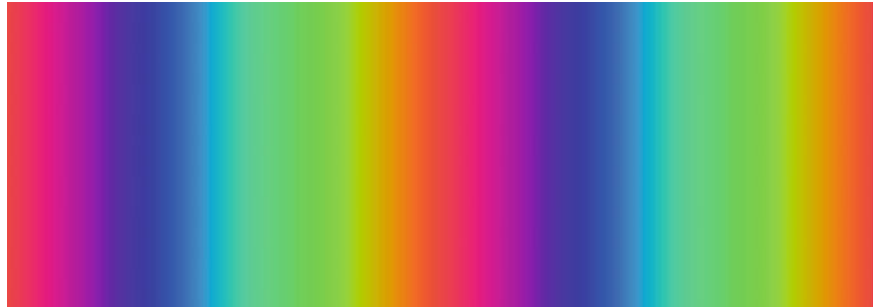

Fig. 43.4 Synchronization of the projecting and capturing processes

43.2.2 Projecting and Capturing Synchronization with Externally Triggering Circuit

An external triggering circuit based on a microcontroller on Arduino board controls and synchronizes the capturing process of the high speed camera with the projecting process of the DLP projector. In order to trigger the DLP projector externally, the internal triggering circuit, which is normally attached to the panel of the color wheel, is removed. Two wires, ground and input, are wired out of the projector after that. The ground wire is then linked with the ground of the Arduino board whereas the input wire is connected to one output pin of the Arduino board to get the desired control signal. To trigger the camera with the same circuit, a LVTTL (Low Voltage Transistor-Transistor Logic) (3.3 V) trigger and strobe connector module is needed. This module will transmit the trigger signal to the frame grabber card to control the camera. With this microcontroller, there is no longer delay for the synchronization between camera capturing and projector projecting. As mentioned previously, in the ideal case, three images can be obtained in each projection cycle. In reality, the camera employed in the system has a minimum period of 5 ms between two shots. For this reason, two projection cycles for each capture of three images are adopted in the system. Figure 43.4 demonstrates the timeline for capturing three images in two projection cycles.

43.3 Advanced Approach in FPP

43.3.1 Gamma Correction Scheme

In the FPP 3D imaging technique, the full-field shape information of the object can be achieved from captured fringe images. The initial fringe patterns are typically generated with the following sinusoidal function:

$$I_{0_i}(x,y) = I_c \left[1 + \cos \left(\frac{2\pi f x}{W} + \delta_i \right) \right]$$

(43.1)

where I_0 is the intensity of the pattern at pixel coordinate (x, y), i indicates the ith image, W is the width of the pattern, f is the fringe frequency, i.e., number of fringes in the image, δ is the phase-shifting amount, and I_c is a constant denoting the intensity modulation amplitude. Depending on the intensity produced by the lamp, all the commercial DLP projectors are employed a gamma modification to enhance the visual effect of the projected images. This leads to non-sinusoidal error to the desired fringe patterns which affect the accuracy of the whole system. In our system, a robust scheme [8] is employed and can be considered as an effective solution where the initial fringe patterns created by above equation are pre-modified to compensate non-sinusoidal error. With this scheme, the equation for generating the source fringe pattern can be modified as:

$$I = 2I_0 . \left[\frac{1}{2} + \frac{1}{2} \cos \left(\frac{2\pi f x}{W} + \delta \right) \right]^{\frac{1}{\gamma}}$$

(43.2)

With the new subscript γ indicates the gamma constant which DLP projectors adopt. The advantage of this approach is that the correction is made with the patterns before projecting onto the objects of interest. By using the appropriate solution for gamma encoding, this approach does not introduce any additional calculation in the 3D shape determination part while helps the system achieve higher accuracy in the phase detection and the eventual 3D imaging.

43.3.2 Multiple-Frequency Fringes Approach for Phase Unwrapping

FPP-based techniques generally employ phase shifting approach to obtain the full-field wrapped phase distributions of the projection fringes. The wrapped phase must be unwrapped to obtain the real phase distributions, which are required for the FPP 3D imaging. Although there are numerous phase unwrapping methods, the problem with fringe discontinuities due to the complex shape of objects or multiple objects still presents. In this paper, this issue can be well addressed by using multi-frequency fringes projection approach [9]. The new approach uses various fringe frequencies to automatically determine the unwrapped full-field phase distributions from the wrapped ones in a fast manner. The basic principle of the algorithm is as follows. The lowest-frequency fringe pattern with one single-fringe on the entire field is utilized to get full-field phase distributions and the required integer order offsets for higher fringe frequencies. The algorithm for the unwrapped phase distributions of the higher frequencies can be described as:

$$\emptyset_i(u, v) = \emptyset_i^w(u, v) + 2\pi . INT \left[\frac{\emptyset_{i-1}^{uw} \cdot (f_i / f_{i-1}) - \emptyset_i^w}{2\pi} \right] \ with \ (i = 2, 3, \ldots, n)$$

(43.3)

where i indicates the ith projection fringe pattern, and uw and w denote unwrapped phase and wrapped phase, respectively. In the equation, n is the number of various fringe frequencies with $n \geq 2$; f is the relative fringe frequency or the number of fringes in the projection pattern with $f_n \geq f_{n-1} \geq \ldots \geq f_1 = 1$; INT represents an operator to take the rounding integer of a decimal number. It is noted that the ratio between two adjacent fringe frequencies f_i/f_{i-1} should not be large (e.g. a ratio less than 5) to avoid the noise effect and ensure the accuracy reliability of the measurement system.

43.3.3 3D Shape Reconstruction Algorithm and System Calibration

To obtain the 3D shape of an object, the primary task in the FPP-based 3D imaging technique is to determine the out-of-plane height and depth information. The mathematical governing equation of the out-of-plane shape determination can be written as [10, 11]:

$$Z = \frac{1 + c_1\emptyset + (c_2 + c_3\emptyset)i + (c_4 + c_5\emptyset)j}{d_0 + d_1\emptyset + (d_2 + d_3\emptyset)i + (d_4 + d_5\emptyset)j} \tag{43.4}$$

where Z is the out-of-reference-plane dimensions information at corresponding pixel coordinate (i,j), is the unwrapped phase determined by Eq. (43.3), and the coefficient c_1–c_5 and d_0–d_5 are constants determined by geometrical and other relevant parameters. It should be noted that the above equation is rigorous for the ideal case; for further enhancement of the 3D imaging accuracy, the Eq. (43.4) can be rewritten with extra terms of second order i and j:

$$Z = \frac{1 + c_1\emptyset + (c_2 + c_3\emptyset)i + (c_4 + c_5\emptyset)j + (c_6 + c_7\emptyset)i^2 + (c_8 + c_9\emptyset)j^2 + (c_{10} + c_{11}\emptyset)ij}{d_0 + d_1\emptyset + (d_2 + d_3\emptyset)i + (d_4 + d_5\emptyset)j + (d_6 + d_7\emptyset)i^2 + (d_8 + d_9\emptyset)j^2 + (d_{10} + d_{11}\emptyset)ij} \tag{43.5}$$

The existing calibration of the FPP-based 3D shape measurement has a considerable limitation in practice because it requires a number of precisely made gage blocks of various sizes. To deal with this problem, a flexible calibration [12] has been introduced with its advantages of easy implementation, full-field and high accuracy measurement. The flexible calibration method requires capturing a set of images with different fringe frequencies at various positions on camera calibration board (at least three positions). The high-resolution low-noise calibration board can be obtained at each position by averaging the captured fringe images. It is also noted that the first position of the calibration board is very important (closest and parallel to reference plane) due to its capability to generate a virtual reference plane. The height of each corresponding point $(X_{c,ij}, Y_{c,ij}, Z_{c,ij})$ in the camera coordinate associated to virtual reference plane can be calculated as:

$$Z_{ij} = \frac{AX_{c,ij} + BY_{c,ij} + CZ_{c,ij}}{\sqrt{A^2 + B^2 + C^2}} \tag{43.6}$$

where A, B, C are the planar coefficients of the reference plane and Z_{ij} is the depth information of jth control point on calibration board obtained at the pth position.

43.3.4 Using CUDA and OpenGL Interoperation to Reconstruct Ultrahigh Accuracy 3D Images with Multiple Frequencies

To increase the measurement accuracy while maintaining the real-time feature of the 3D imaging and 3D reconstruction, multiple-frequency fringes are used. For multiple-frequency fringe patterns, the patterns for projection need to be changed periodically. However, at high rate of capturing, the pattern switch can disrupt the captured images, which leads to 3D results with large errors. The problem comes from the instability in time of the process of transferring data from the memory of the computer (RAM) to the memory of the graphic cards (GPU). Since the patterns are traditionally generated by the CPU and stored on the memory of the computer, every time a new pattern needs to be displayed, the CPU will transfer the pattern data from host memory to global memory of graphic device. The time required for this data transaction is not always a constant. In fact, this period can vary from 30 to 200 ms. This instability prevents the system from performing 3D shape measurements accurately and stably.

This problem can be well resolved by the implementation of CUDA programming. CUDA (an acronym for Compute Unified Device Architecture) is a parallel computing architecture developed by NVIDIA. CUDA allows programmers to access the memory of the graphical device as well as take advantage of the GPU to process data. So by using CUDA, the whole process of projection, including generating data and displaying it on the screen, can be done entirely within the GPU. No data transaction or communication is needed between the host and the graphical device. In addition, due to its natural function, the graphical device is able to handle the switching pattern in very stable manner. Furthermore, in order to display

the patterns quickly and reliably, OpenGL is employed. OpenGL is a software interface to graphics device and is used to draw pixel or vertices on the screen. By combining OpenGL with CUDA, the image data generated by CUDA part can be transferred and displayed at a very fast speed because they share data through the common memory in the frame buffer.

43.4 Experiments

To demonstrate the validity and practicability of the proposed system, two experiments have been conducted and presented here. In the experiments, the camera and projector are arranged arbitrarily according to the generalized set up; the distance between camera and the reference plane is about 1.2 m, and the reference plane has a physical width of around 411 mm.

In the first experiment, three different fringe frequencies (e.g. 1, 3, and 9) are combined with the widely used three-step phase-shifting technique to significantly increase the accuracy performance of the measurement system. A gage board with 9 different gage blocks on it is used to test the measurement accuracy of the technique. The largest measurement error came from the largest block, which is a cube of 101.60 mm width/height/depth. The measurement result is 101.64 ± 0.33 mm with a standard deviation of 0.07 mm. Through the experiment, the accuracy can be calculated as the ratio of out-of-plane measurement error to the in-plane dimension, which gives a relative accuracy of (101.64–101.6 mm)/411 mm ≈ 1/10,000. On the contrary, if only one fringe frequency is used to reconstruct the 3D images, the relative accuracy of the measurement system is attenuated to 1/2,500. Figure 43.5 have shown the final detected 2D and 3D shape maps of the measured gage blocks.

In the second experiment, by using only two projection cycles to capture three images from a single fringe frequency, we successfully reconstruct the 3D images at a speed of 45 fps. Moreover, with the aforementioned GPU-based scheme, three-frequency fringe patterns are used together with a synchronization between the camera and projector. This gives a speed of 10 fps with an accuracy-enhanced 3D imaging. The fast speed system have been employed to reconstruct the 360° 3D models of object with different shapes and colors. Figure 43.6 shows some representative examples.
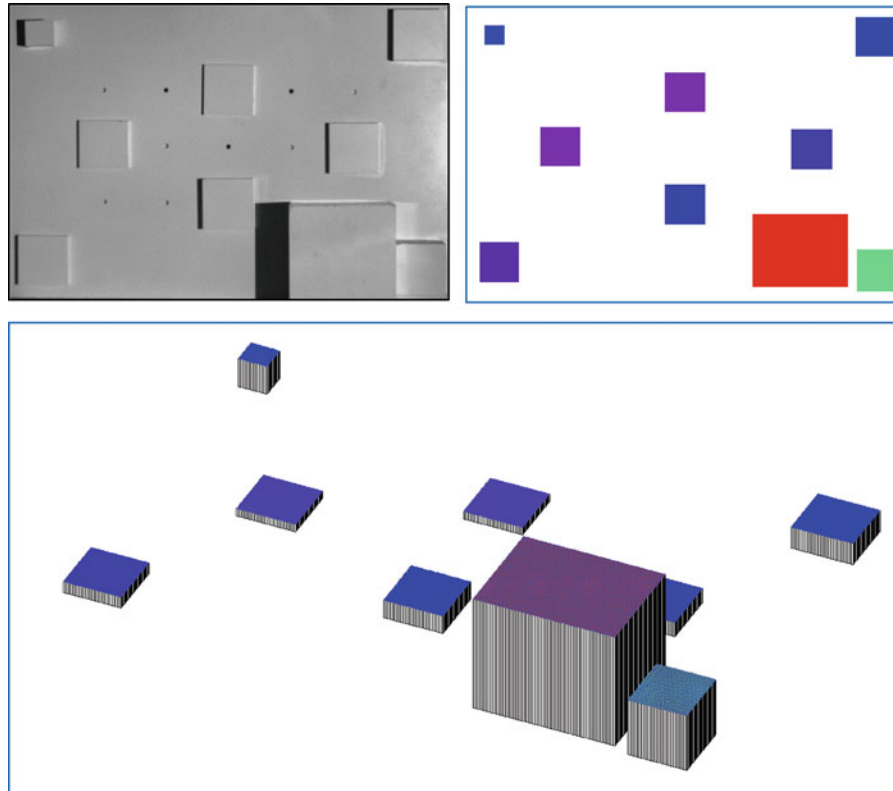

Fig. 43.5 Accuracy test with a gage block target and the results plots in 2D and 3D
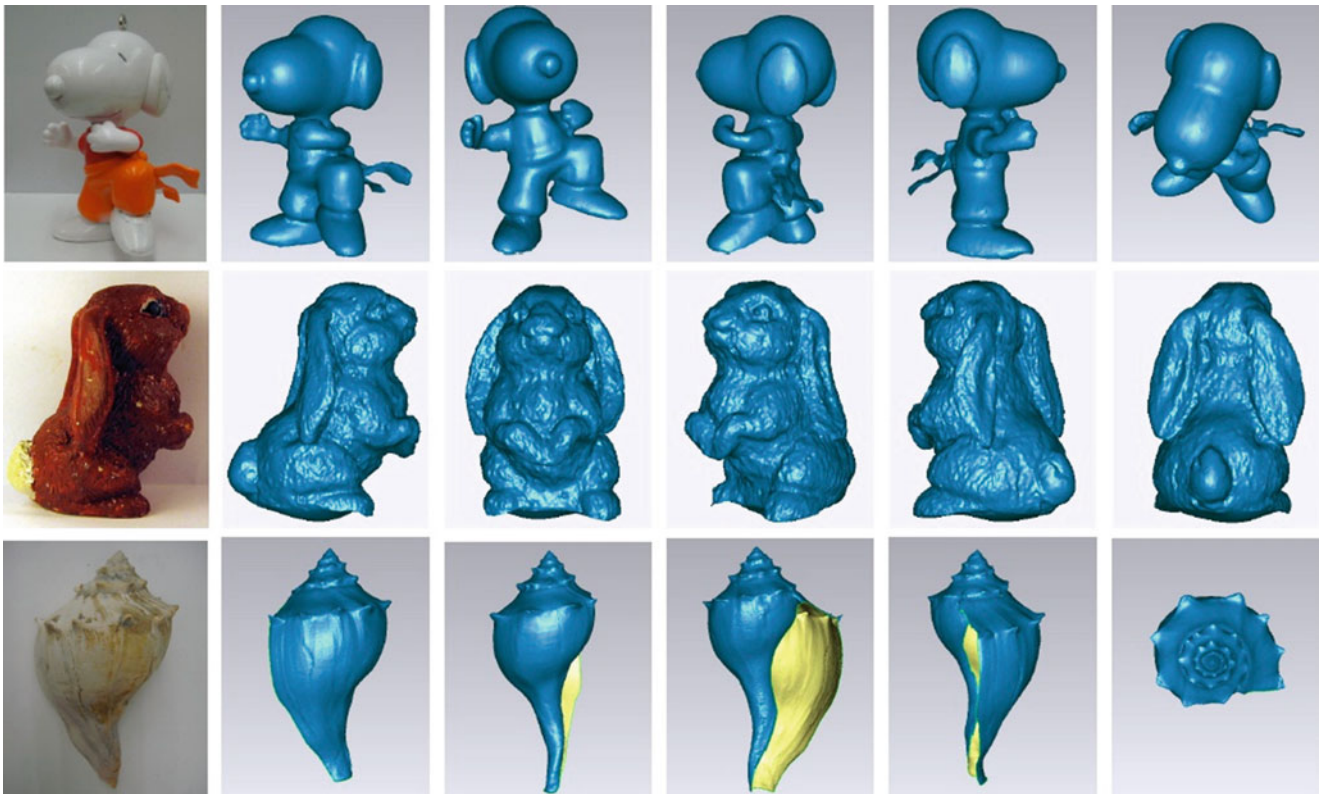
Fig. 43.6 Full 3D reconstructed models of a snoopy, a rabbit and a sea shell

43.5   Conclusion

This paper presents a high-accuracy FPP-based 3D imaging technique with a capability of measuring multiple objects in a real-time manner. The technique uses a series of advanced technical approaches to cope with the limitations of the existing techniques. The validity and practicability of the technique have been verified by experiments.

As technologies evolve, there are high demands for the 3D imaging techniques to possess a number of advanced technical features. We believe that the technique presented in this paper is capable of satisfying the ever-increasing application requirements in enormous scientific and engineering practice.

References

 1. Burke J, Bothe T, Osten W, Hess C (2002) Reverse engineering by fringe projection. Proc SPIE 4778:312–324
 2. Park SC, Chang M (2009) Reverse engineering with a structured light system. Comput Ind Eng 57:1377–1384
 3. Geng J (2011) Structured-light 3D surface imaging: a tutorial. Adv Opt Photon 3:128–160
 4. Sutton A, Orteu J, Schreier H (2009) Image correlation for shape, motion, and deformation measurements. Springer, New York
 5. Kieu H, Pan T, Wang Z, Le M, Nguyen H, Vo M (2014) Accurate 3D shape measurement of multiple separate objects with stereo vision. Meas Sci Technol 25:035401
 6. Chen F, Brown G, Song M (2000) Overview of 3-D shape measurement using optical methods. Opt Eng 39:10–22
 7. Gorthi SS, Rastogi P (2010) Fringe projection techniques: whither we are? Opt Lasers Eng 48:133–140
 8. Hoang T, Pan B, Nguyen D, Wang Z (2010) Generic gamma correction for accuracy enhancement in fringe-projection profilometry. Opt Lett 35:1992–1994
 9. Wang Z, Nguyen D, Barnes J (2010) Some practical considerations in fringe projection profilometry. Opt Lasers Eng 48:218–225
10. Du H, Wang Z (2007) Three-dimensional shape measurement with an arbitrarily arranged fringe projection profilometry system. Opt Lett 32:2438–2440
11. Wang Z, Du H, Han B (2006) Out-of-plane determination in generalized fringe projection profilometry. Opt Express 14:121–122
12. Vo M, Wang Z, Pan B, Pan T (2012) Hyper-accurate flexible calibration technique for fringe-projection-based three-dimensional imaging. Opt Express 20:16926–16941

Chapter 44
DIC Strain Analysis of Pipeline Test Specimens Containing Metal Loss

Leonardo D. Rodrigues, José L.F. Freire, and Ronaldo D. Vieira

Abstract The elastic and plastic strain data of two tubular specimens, with longitudinal and circumferential metal loss at areas on their external surface, tested under internal pressure are presented and analyzed. The specimens were cut from two longitudinally welded tubes made of low carbon steel with nominal outside diameters of 76.7 and 102 mm and nominal wall thicknesses of 2.04 and 3.2 mm respectively. Each of the tested specimens had milled metal loss areas simulating external longitudinal and circumferential corrosion defects. Digital Image Correlation was employed to determine elastic and plastic strains at the metal loss defects and their results were verified and compared with data collected with electrical resistance strain gages and results obtained with a finite element numerical solution.

Keywords Pipeline • Corrosion • Metal loss defects • Circumferential defects • Longitudinal defects • Strain distributions • DIC • Finite elements • Strain gages

44.1 Introduction

This paper had four objectives. One was to apply the Digital Image Correlation (DIC) method to measure surface elastic and plastic strains in actual thin pipes loaded by internal pressure which had machined longitudinal and circumferential metal losses. The second objective was to verify, by means of confronting laboratory tests and numerical finite element tests, its accuracy and applicability to determine strain distributions in possible field measurements. The third objective was to determine longitudinal and circumferential strain distributions on defects that are very wide in the circumferential direction but have short lengths in the longitudinal directions. The fourth and final objective was to show that strain distributions could be directly achieved on the surface of a composite repair constituted by a carbon-fiber-reinforced-epoxy (CFRE) composite repair. Measurement of strains on the composite repair free surface can help to assess strains occurring on the defect surface of the test specimen that are located below the reinforcement repair during the tests and possibly under field operation.

Two cylindrical thin wall pipe specimens were used in the tests. Figure 44.1 shows the sketch of the specimens called S1 and S2. DIC measurements were made on areas A1 (long-narrow and short-wide defects), A2 (short-narrow defect) and B2 (composite surface of a longitudinal long-narrow reinforced defect). Rosette strain gage measurements were made at a nominal area (B1), and on areas B2, C1 and C2 to be compared with the DIC results.

44.2 Experimental Methods

The raw material used in the tests was composed of longitudinally welded pipes made of low carbon steel with S_y yield strength and S_u ultimate strength equal to 290 and 350 MPa, respectively.

The lengths of the two original pipes were approximately 6 m. The nominal outside diameter and the nominal wall thickness of the pipes were, respectively, 76.7 and 2.04 mm for specimen named S1 and 102 and 3.2 mm, respectively, for

L.D. Rodrigues (✉) • J.L.F. Freire • R.D. Vieira
Department of Mechanical Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil
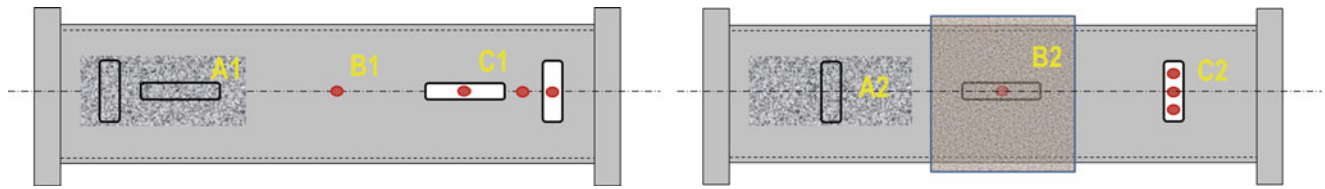e-mail: leo45dr@gmail.com; jlfreire@puc-rio.br; rdvieira@puc-rio.br

Fig. 44.1 Sketch of specimens S1 (*left*) and S2 (*right*). A1 and A2 are areas of DIC strain measurements; B1, C1 and C2 are areas of rosette strain gage measurements; B2 is the area with a longitudinal defect that had a strain gage rosette bonded to the metal surface and it was repaired with layers of an epoxy reinforced carbon fiber system CFRE

Table 44.1 Data on specimens S1 and S2

Specimen	Defect (metal loss) area	Defect	t^*	L	w	$\frac{d}{t}$	$\frac{L^2}{D_e\,t}$	
S1	C1	Long-SG	0.88	80	15	0.57	41	$D_e = 76.7$ mm
		Circ-SG	0.92	15	82	0.55	1.4	$t = 2.04$ mm
	A1	Long-DIC	1.19	80	15	0.42	41	a1 $= 70$ mm
		Circ-DIC	1.39	15	82	0.32	1.4	a2 $= 30$ mm
								See Fig. 44.2
S2	B2	Long-SG	1.2	200	20	0.58	138	$D_e = 112$ mm
	A2	Circ-DIC	1.2*	20	120	0.58	1.2	$t = 3.2$ mm
	C2	Circ SG	1.2**	20	120	0.58	1,2	a1 $= 150$ mm
		–	C1	C2	C3	C4	C5	See Fig. 44.3
		Circ-DIC—t*	1.36	1.42	1.32	1.12	1.11	
		Circ-SG—t**	1.22	1.34	1.28	1.13	1.05	

*, ** nominal thickness: variation of the thickness for five measured points 20 mm apart along the circumferential mid-line of the defects
t* and t** (remaining thickness) (mm); L (defect length) (mm); w (defect width) (mm);
d/t (Ratio of defect depth to nominal wall thickness); D_e (external diameter) (mm)
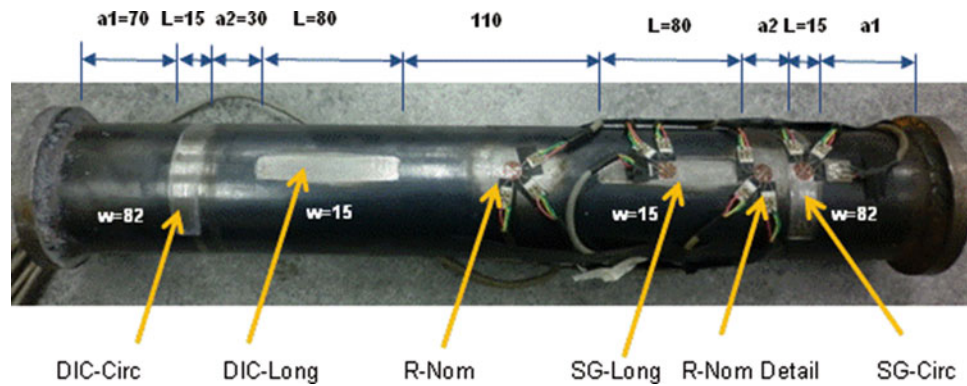

Fig. 44.2 Cylindrical pipe pressure vessel specimen S1 showing location of defects and strain gage rosettes

specimen named S2. The specimens had circumferential defects and longitudinal defects milled to simulate external and uniform depth corrosion (or erosion) defects. These defects were machined opposite to the tube seam weld.

The tubular specimens were closed with plane heads that were welded on before the defects were machined. Sketches of the specimens S1 and S2 are respectively presented in Fig. 44.1. The so called longitudinal and circumferential defects machined in both specimens S1 and S2 had longitudinal lengths that are considered, respectively, equivalent to long length and short length corrosion defects.

Table 44.1 gives data of specimens S1 and S2. Figure 44.2 shows specimen S1 before being tested and Fig. 44.3 shows specimen S2 details.

The thin walled specimens were loaded with internal pressure only. The water pressurizing system used an air driven pump with a maximum capacity of 70 MPa. Strain gage and manometer data readings were made with a Lynx ADS 2000 system. Strain gages rosettes were manufactured by Kyowa (KFC-5-120). A cyanoacrylate based adhesive was used. The strain gages were connected to the Lynx system using the traditional 3-wire arrangement. As expected, the most-strained
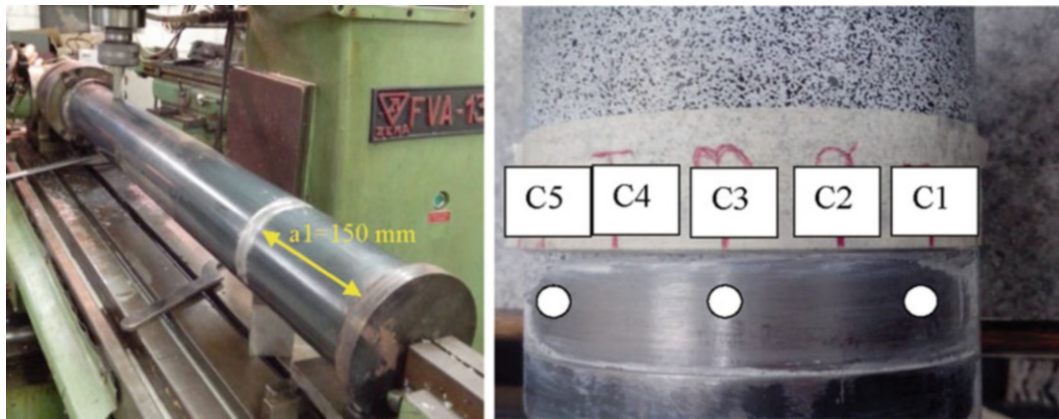
Fig. 44.3 Cylindrical pipe pressure vessel specimen S2 showing machining of the circumferential defect (*left*) and location of remaining thickness measurement points (*right*). Three 5 mm length strain gage rosettes were bonded on the circumferential defect (points marked with *circles*)

gages stopped measuring after a total deformation of about 4 % was reached. This occurred during the (burst) tests. The DIC strain measurements were performed with a 3-D stereo Correlated Solutions system. The system consisted of two CCD cameras, two lenses, a tripod, calibration targets (4, 5 and 9 mm depending on the inspection area were used in the tests), and two software—one for image acquisition (VIC-Snap 2009) and another for data analysis (VIC-3D 2009).

44.3 Results and Discussion

44.3.1 Specimens S1

Four pressure tests were performed. In the first three tests, pressure was increased from 0 to 5 MPa in five steps. Strain gage readings and DIC images were taken at each step. In one of these DIC tests both longitudinal and circumferential defects were visualized at the same time (global test). The second and third tests were performed in such a way that the defects were magnified in the pictures (DIC local tests). In other words, the second test was performed with full vision of the longitudinal defect and the third test was performed with full vision of the circumferential defect.

Elastic strain and pressure data collected during the first three tests are presented in Fig. 44.4. Elastic strains measured by the strain gage and DIC techniques along with pressure variation are depicted in Fig. 44.4a for the longitudinal defect areas. Only the circumferential strains are presented in Fig. 44.4a due to the fact that they were much larger than the longitudinal strains in all tests. All circumferential strain gage and DIC determined data are plotted in Fig. 44.4b for comparison purposes. At this point it is important to note that the DIC data presented in Fig. 44.4 were corrected by a factor equal to the calculated ratio between the thicknesses of the similar points where strain gage and DIC measurements were taken. It is possible to see that strains independently measured by both techniques coincided satisfactorily, and this can be demonstrated by the proximity of the data points with a 45° straight line. The handling of these data revealed an average deviation of -8 µε and a root mean square deviation of 79 µε.

A Finite Element solution was carried out for specimen S1 using the software ANSYS. Results for this solution are plotted in Fig. 44.5 for the circumferential strains that occur along the central line of the longitudinal defect when the internal pressure is 5 MPa. This solution is compared with the DIC solution for the same defect central line and both strain distributions agree satisfactorily.

The forth test was performed in two steps. First, under zero pressure, initial images were taken from the global field that covered both defects. Next, the specimen was taken off from the optical DIC bench and positioned inside the test bunker to be pressurized up to bursting at a pressure equal to 14.2 MPa.

Rupture, as expected by the dimensions that resulted from the defect machining process, occurred in the area of the longitudinal defect with the smallest thickness (C1 of Fig. 44.1, the one instrumented by the strain gage rosette). It should be noted that strain gage readings were taken during the bursting test up to the point the circumferential gage located in the center of the longitudinal defect stopped measuring, due to the very large strain reached.
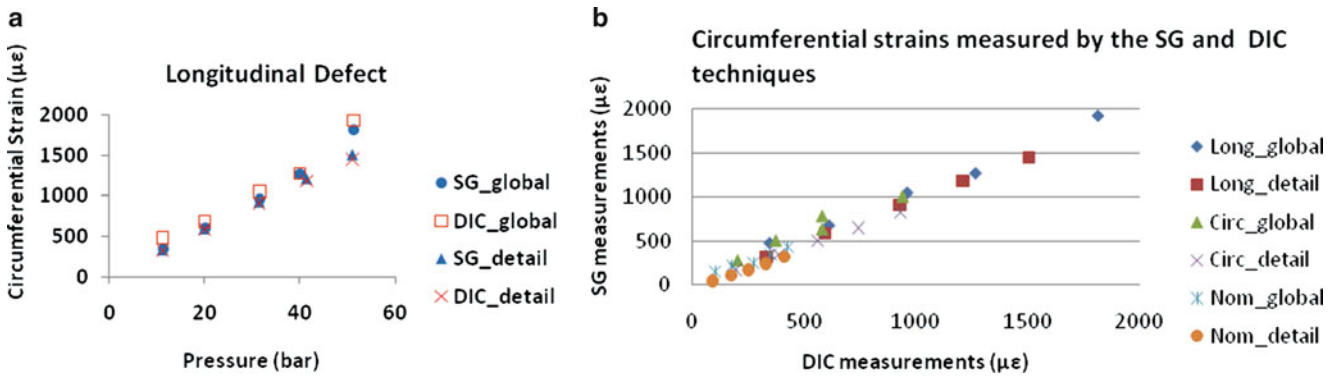
Fig. 44.4 DIC and SG elastic strain measurements collected in the pressure vessel test; (**a**) circumferential strains at the central point of the longitudinal defect; (**b**) comparison of DIC and SG measurements at similar points of the specimen
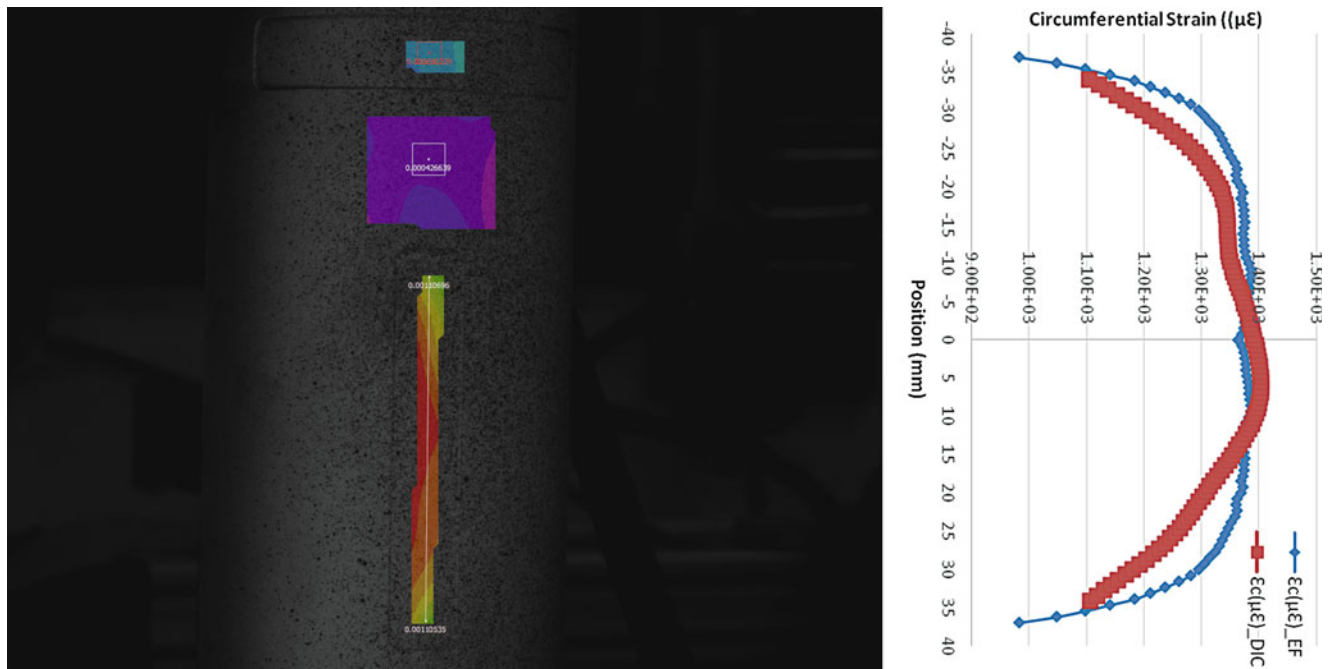

Figure 44.5 Finite Element and DIC results for circumferential strains along the central line of the specimen S1 while under 5 MPa internal pressure

After bursting, the specimen was again positioned on the optical bench as close as possible to its previous (initial) position in order to be viewed and photographed again by the cameras. In this way it was possible to collect data to determine the remaining plastic strains that occurred in the defects' area covered by the DIC analysis (A1 of Fig. 44.1). The image of the area viewed by one of the cameras used in the DIC analysis is presented in Fig. 44.6.

Figure 44.6 shows the magnified area of the longitudinal defect that was analyzed by the DIC method. This area is similar to the rupture area but it was kept from fracturing, as expected, due to its much larger remaining thickness. The full field DIC strain analysis of this region is presented in Fig. 44.6 showing the circumferential strain distribution ($\varepsilon_c = \varepsilon_{xx}$). The central point of the longitudinal defect presents a rather large circumferential strain.

The circumferential distribution shown in Fig. 44.6 is compatible with the plot of strains varying with pressure shown in Fig. 44.7. This Fig. 44.7 shows the variation of circumferential strains measured by the strain gages under pressure. These values were measured during the fourth (rupture) test. It can be seen that the center of the longitudinal defect (where fracture actually happened) was much more deformed than the other instrumented points. This observation is compatible with the strain plot of Fig. 44.6.

Fig. 44.6 Circumferential strain distribution measured by the DIC technique for points located along a *vertical line* that crosses the *center line* of the circumferential and longitudinal defects; (**a**) image of the specimen viewed by one of the CCD cameras; (**b**) image of the strain response given by the VIC-3D 2009 software; (**c**) strain distribution
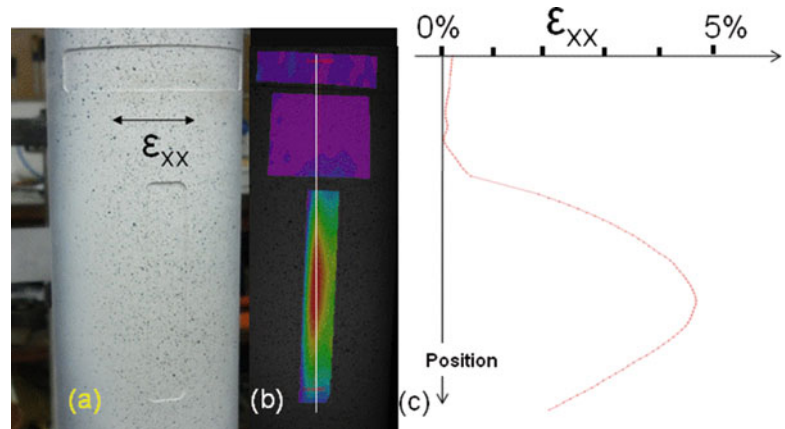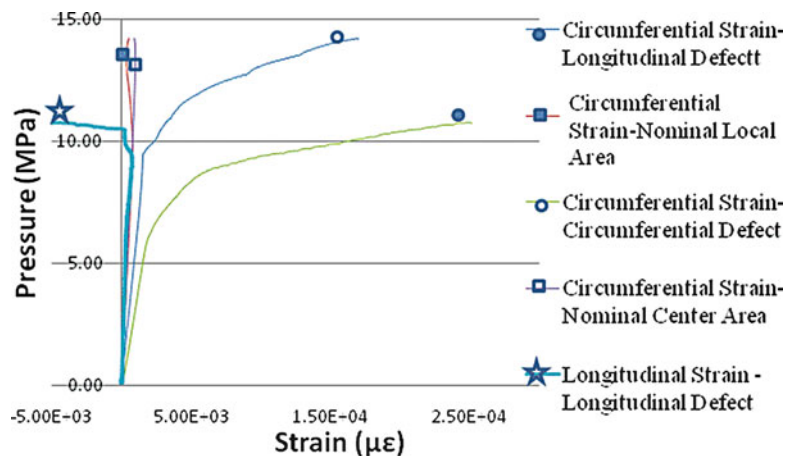

Fig. 44.7 Circumferential and longitudinal strains measured by the strain gage rosette located in the *center* of the longitudinal defect and circumferential strains measured by the other rosettes located inside the circumferential defect and in the so called nominal-global (*center* of the specimen) and nominal-detail (between the circumferential and the longitudinal defects) areas

Although not shown in this paper, it has to be stated that the measured longitudinal strains at the heavily plastic deformed points were very small when compared to the circumferential strains measured for the same points. This can be seen in the plot of strains shown in Fig. 44.7. The plot presents the circumferential and longitudinal strains measured by the strain gage rosette located in the center of the longitudinal defect. The strains were measured during test four before rupture occurred and show the elastic-plastic behavior of the material point. One can see that the longitudinal strain is very small when compared to the circumferential strain. One can also see that it changes its positive increasing trend after plastic behavior starts. This behavior has been observed before in [1] and it is explained by the restriction offered by the thicker and near walls of the defect, and by a decrease in thickness of this region to guarantee plastic strain behavior at volumetric strain equal to zero.

44.3.2 Specimen S2

Distributions of circumferential and longitudinal total strains along the central line of the circumferential defect (positions A2 and C2 of the S2 specimen depicted in Fig. 44.1) are plotted in Fig. 44.8 for an internal pressure equal to 8 MPa. These total strains were determined by using the DIC, strain gage and Finite Element techniques and were linearly corrected to the same normalized remaining thickness of 1.11 mm. This correction was made to consider the variation of the remaining thicknesses along the defect lengths that occur inside the circumferential defects of positions A2 and C2 as showed in Table 44.1. Considering that the experimental solutions (DIC and Strain Gages) were applied to different locations in the actual test specimen (A2 and C2) and also that the numerical solution used a constant remaining thickness equal to 1.11 mm, it is possible to see that the DIC, Strain Gage and Finite Element solutions agreed fairly well (in terms of shape of the distributions and deviation around 150 $\mu\epsilon$).

Fig. 44.8 Distributions
of circumferential and
longitudinal total strains
determined by the DIC, Strain
Gage and the Finite Element
techniques along the central
line of the circumferential
defect (positions A2 and C2
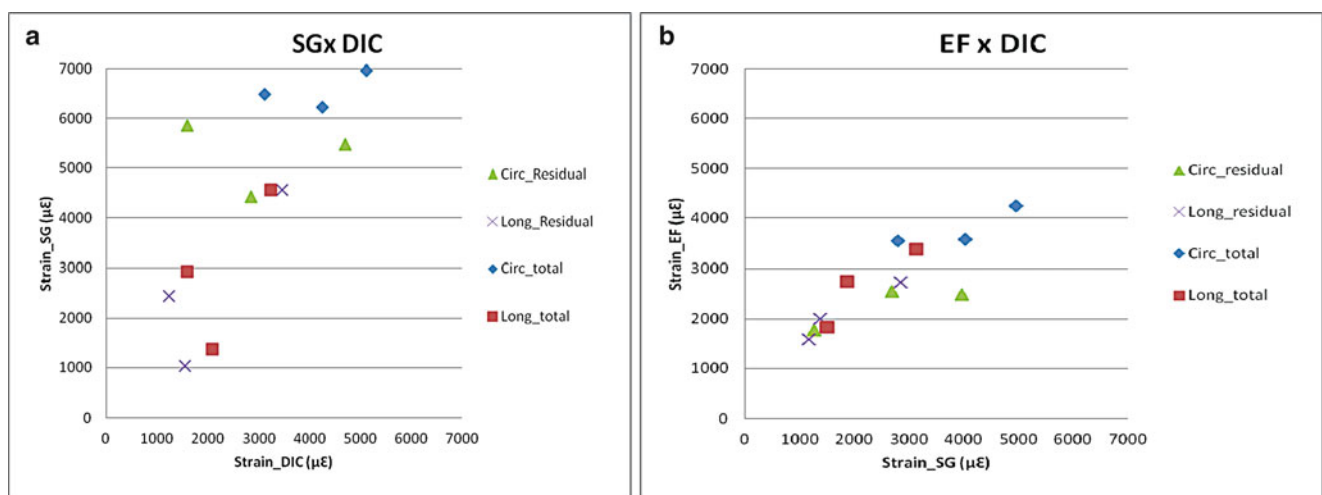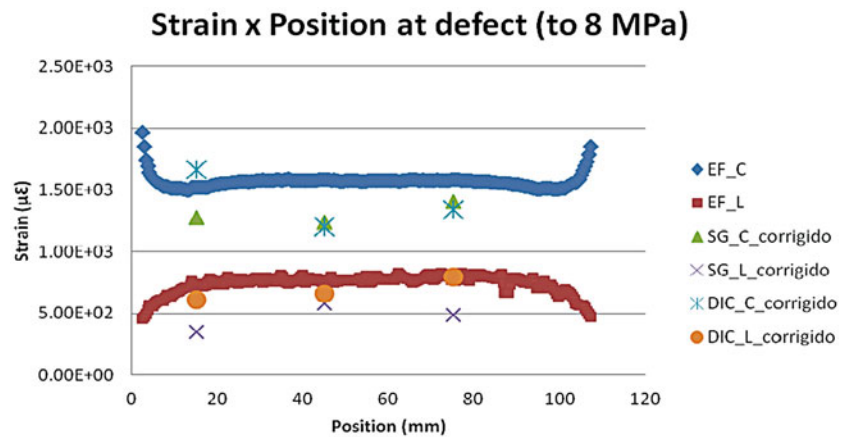of the S2 specimen depicted
in Fig. 44.1)

Fig. 44.9 Total circumferential and longitudinal strain (plastic plus elastic) and total residual circumferential and longitudinal strain (after unloading) at the central points of the circumferential defects (positions A2 and C2 of the S2 specimen depicted in Fig. 44.1); (**a**) SG and DIC techniques; (**b**) Finite element and DIC techniques

Total circumferential and longitudinal strain (plastic plus elastic) and total residual circumferential and longitudinal strains (after unloading) at the central point of the circumferential defects (positions A2 and C2 of the S2 specimen depicted in Fig. 44.1) are plotted in Fig. 44.9a, b. These Fig. 44.9a, b show comparisons between the experimental DIC and Strain Gage techniques and the DIC and the Finite Element techniques. In these cases three pressure loading steps were considered in sequence: zero-10 MPa-zero, zero-12 MPa-zero and again zeo-12 MPa-zero. The second loading step, increase from zero to 11 MPa, caused full plasticity inside the defects, and then it was decreased to zero. The first and third loading steps caused small plastic strains. Again, these total remaining strains were determined by using the DIC, Strain Gage and the Finite Element techniques and were linearly corrected to the same normalized remaining thickness of 1.11 mm. The closeness of the data points to 45° lines shows that the results agreed satisfactorily (average deviation around 150 με).

The strain distribution along the thickness in the area that contains the longitudinal defect and the carbon fiber epoxy resin repair is known to vary, starting with higher values at the metal surface and decreasing to lower values at the CFRE surface. Figure 44.10a shows the strain results obtained with a rosette (circumferential and longitudinal gages) bonded on the metal loss external surface of specimen S2 while the pressure increases from zero to 11 MPa. It can be seen that the circumferential strain is elastic (linear with pressure) and below 1,800 με. This value of strain is lower than the circumferential strain that would occur if there was not the fiber reinforcement. Figure 44.10b shows the circumferential and longitudinal strain variations with increasing pressure measured on the surface of the CFRE repair using the DIC technique. It can be seen that the strain behavior is linear with pressure and that the surface strains in the CFRE are lower than the strains occurring in the metal loss surface as mentioned in [1]. Figure 44.10c shows ratios between the strain measured in the metal loss and CFRE
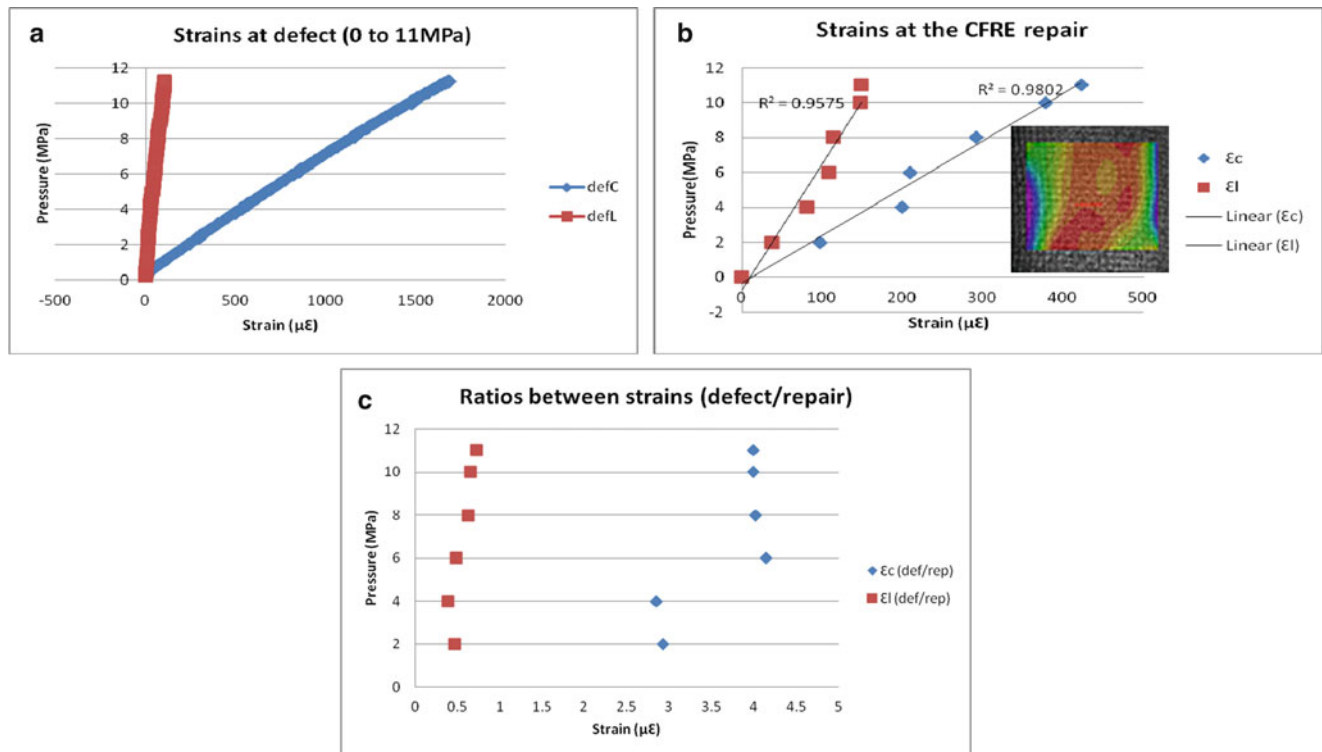
Fig. 44.10 (**a**) Strain results obtained with a rosette (circumferential and longitudinal gages) bonded on the metal loss external surface of specimen S2; (**b**) Strain results obtained with DIC (circumferential and longitudinal gages) on the composite surface; (**c**) Ratio of strains measured on the metal surface to strains measured on the composite surface

surfaces. This is an interesting result because it gives an insight on the strain distribution occurring in the metal-loss defect surface by means of measuring strains in the reinforcement repair surface using a non-destructive and full field technique such as the DIC technique.

44.4 Conclusions

This investigation applied the digital image correlation technique—DIC—to thin walled cylindrical vessels containing longitudinal and circumferential metal loss defects. The DIC was applied to the pipe metal surface and also to the surface of a carbon fiber reinforced epoxy repair. The DIC has proven to be a good choice for this type of experiment when full field surface strains have to be measured and elastic and plastic strains are expected. The comparison of DIC and Strain Gage measured strains and Finite Element calculated strains at similar points of the specimens tested agreed satisfactorily.

Reference

1. Freire JLF, Vieira RD, Diniz JC (2007) Effectiveness of composite repairs applied to damaged pipeline. Exp Tech 31(5):59–66

Chapter 45
Experimental Inference of Inter-Particle Forces in Granular Systems Using Digital Image Correlation

Nikhil Karanjgaokar and Guruswami Ravichandran

Abstract The observed features of granular materials can be successfully predicted using models based on relationship of inter-particle forces and macroscopic properties. In the current work, a drop-tower experimental setup was developed for the impact testing of 2D assembly of cylinders and 3D assembly of spheres with impactor velocity of around 6 m/s. This drop tower setup was used to load 2D granular assembly of polyurethane and polycarbonate cylinders of $1.25''$ length with 3 different diameters of $1/4''$, $3/8''$ and $1/2''$. A high speed camera was used for recording the images of approx. 800×800 resolution at speeds between 10,000 and 15,000 fps to monitor the deformation of the cylinders. The recorded images were used to obtain the granular fabric and kinematics for each grain and average strains were obtained from images using digital image correlation. The experimental data was subsequently used to infer the interparticle forces between individual grains in the assembly using a Granular Element Method based optimization process.

Keywords Digital image correlation • Granular materials • Granular element method • Impact testing • Force chains

45.1 Introduction

The macroscopic behavior of granular materials has not been successfully described using continuum mechanics and phenomenological description of macroscopic response has long been viewed as a viable alternative [1, 2]. Inter-particle forces between individual grains in a granular assembly have been linked to the macroscopic behavior of granular media like constitutive behavior [3, 4], wave propagation [5] and friction [6]. Modeling techniques like discrete element method embracing the link between inter-particle forces and macroscopic properties have also emerged but these models need to be validated through experiments at the particle scale.

In the past, photoelasticity based techniques have been used to observe inter-particle forces in granular media [7, 8]. Although photoelasticity is a useful tool for visualization of inter-particle forces, these studies have been limited to birefringent grains and require the knowledge of the boundary conditions. Most recent techniques for force-inference in granular media require some inherent assumptions about grain shape or law governing the contact mechanics [9, 10].

The recently proposed Granular Element Method (GEM) [11–13] overcomes all of these challenges and is the most versatile and quantitative force-inference technique to-date. GEM is only restricted by the availability of experimental techniques to extract full-field grain displacement fields and particle edges; however, GEM and all previous methods have only been developed for studying inter-particle forces in materials in static equilibrium. The method presented in [13] provides a new tool for quantitatively measuring dynamic inter-particle force transmission in granular media. The method is an extension of GEM, providing a new mathematical foundation and inverse problem formulation which accounts for dynamic particle behavior.

The current paper employs a drop-tower based experimental approach to load an assembly of polyurethane and polycarbonate cylinders and uses high speed imaging to record to the deformation process. Digital image correlation is used to obtain the particle kinematics and strain fields in the granular assembly. The experimental measurements like

N. Karanjgaokar (✉) • G. Ravichandran
Graduate Aerospace Laboratories of the California Institute of Technology, 1200 E California Blvd.
MC 105-50, Pasadena, CA 91106, USA
e-mail: nikhiljk@caltech.edu

granular kinematics, contact points and strain fields are then used to infer the inter-particle forces using the GEM based approach described in [13].

The Sect. 45.2 of this paper discusses the experimental approach designed for the drop-tower testing of the granular assemblies while some preliminary results obtained from the experiments are presented along with some discussion in Sect. 45.3. The final section explores future applications of this method and presents some of its limitations.

45.2 Experimental Setup and Procedure

The drop-tower experimental setup used for testing granular assemblies in the current work is shown in Fig. 45.1a. As seen in the figure, hollow aluminum tube (9′ long) is mounted on a rigid load frame and a rigid impactor can free-fall under gravity through the hollow tube. The impactor velocity when it reaches fixture holding the granular assembly is around 6 m/s. The fixture holding the cylindrical grains has rigid walls on the sides and the bottom face while the slotted top face allows a flat slider to move vertically on impact.

The drop tower setup was used to conduct impact experiments on two different materials: Polyurethane (Durometer 80A) and Impact Resistant Polycarbonate. For both materials, cylindrical grains with a length of 1.25″ and three different diameters (1/4″, 3/8″ and 1/2″) were randomly arranged as granular assemblies. The elastic modulus for Polyurethane rods was 9.35 MPa with a Poisson's ratio u = 0.49 while the elastic modulus for Polycarbonate rods was 2.6 GPa with a Poisson's ratio u = 0.37. The coefficient of friction for polyurethane and polycarbonate was 0.5 and 0.3 respectively.

In order to facilitate the use of digital image correlation (DIC), the flat faces of the cylindrical grains were patterned with spray paint with appropriate speckle density for different diameters (see Fig. 45.1b). The images during the impact experiments were recorded with Phantom v710 high speed camera at an approx. resolution of 800×800 pixels between 10,000 and 15,000 fps.

The DIC software VIC-2D was used to calculate the displacement and velocity fields for each grain separately. The displacement fields were then used to obtain the εxx, εxy and εyy strain fields in each grain and calculate the average strain in each grain. The velocity fields from DIC in each grain were then used to calculate the acceleration fields and average accelerations in the grains by numerical differentiation.
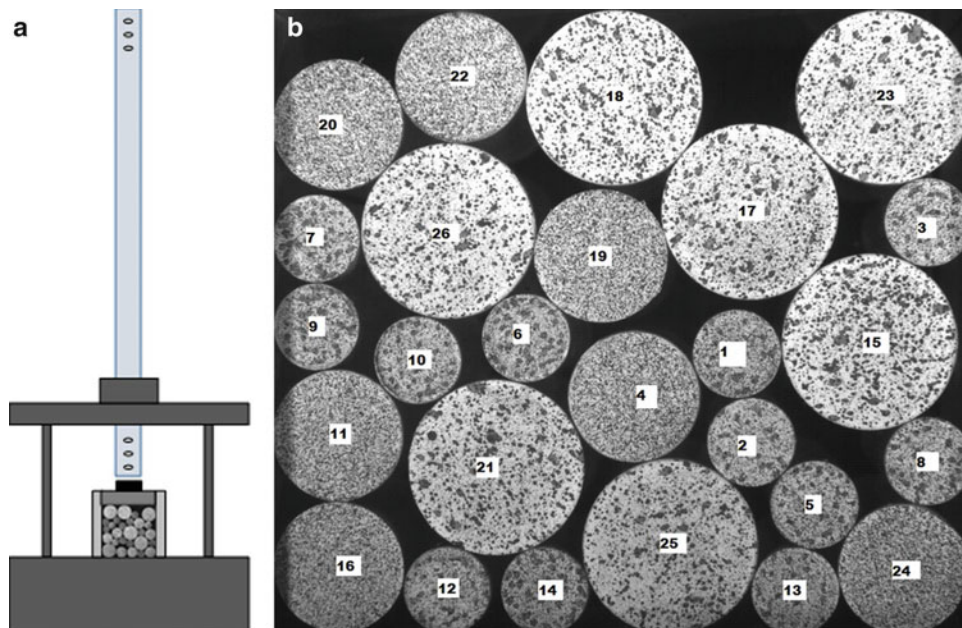

Fig. 45.1 (**a**) A schematic representation of the drop tower setup for loading the granular assembly of cylindrical grains, (**b**) a representative image recorded by the high speed camera for a granular assembly of polyurethane cylindrical grains of three different diameters with speckle-patterns

Fig. 45.2 The particle edges (*red circles*) and contact points (*blue circles*) for a representative image as detected by the circular Hough transform based approach described above
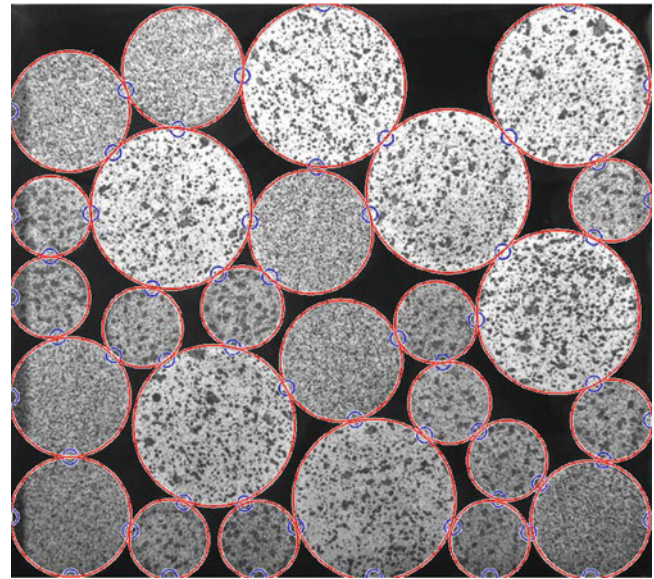
The recorded images were also used to obtain the position of the centroid and the locations of the particles contacts at each time instant (see Fig. 45.2). This was achieved by converting the recorded grayscale image to a binary image and the particle centroids and radius were then detected using a circular Hough transform with Matlab's image processing toolbox. The resulting centroids and radii were then used to find contacting particles: if the distance between two centroids was less than two times the radii of the corresponding particles, the particles were taken to be contacting with appropriate normal and tangent vectors. Similarly, contact points between particles and the boundaries of the fixture were also detected.

45.3 Results and Discussion

The Fig. 45.3 shows the representative εxx, εxy and εyy strain fields for polyurethane grains at some time instant during the impact experiment. In order to compute the average stress values from average strain values for each grain, linear elastic model is used for both materials. While the assumption is definitely valid for grains with small strains, it may no longer be valid for polyurethane grains with large strains. But hyperelasticity, plasticity and rate dependency based models for polyurethane are not considered for scope of this paper.

The analysis of the velocity fields obtained from the DIC results reveals that although the wave velocity for the polyurethane is around 93 m/s, the wavefront moves at approximately 5 m/s through the granular media.

The numerical optimization presented in [13] links particle positions, accelerations, contact points, and stresses to inter-particle forces. Three particle-scale equations accomplish this connection: momentum balance, stress-force relations, and constraint equations. These equations are combined in a multiobjective optimization problem that can be solved to obtain inter-particle forces. In [13], the authors have ignored the effect of local accelerations and have thus ignored any effect of only the mean particle acceleration in their analysis. In the current work, we have studied the effect of local acceleration gradients within the particles and have included integral terms with moments of local accelerations about the particle centroid. In [13], the plane stress assumption has been considered in constitutive response but as seen in Fig. 45.4, the comparison of plane stress and plane strain conditions for same experiment at a particular instant reveals good agreement between the two cases in terms of inter-particle forces.

The good agreement between the plane stress and plane strain conditions indicate that the contribution of elastic energy stored in the grains is limited. High inter-particle forces are observed due to collisions between grains indicating the
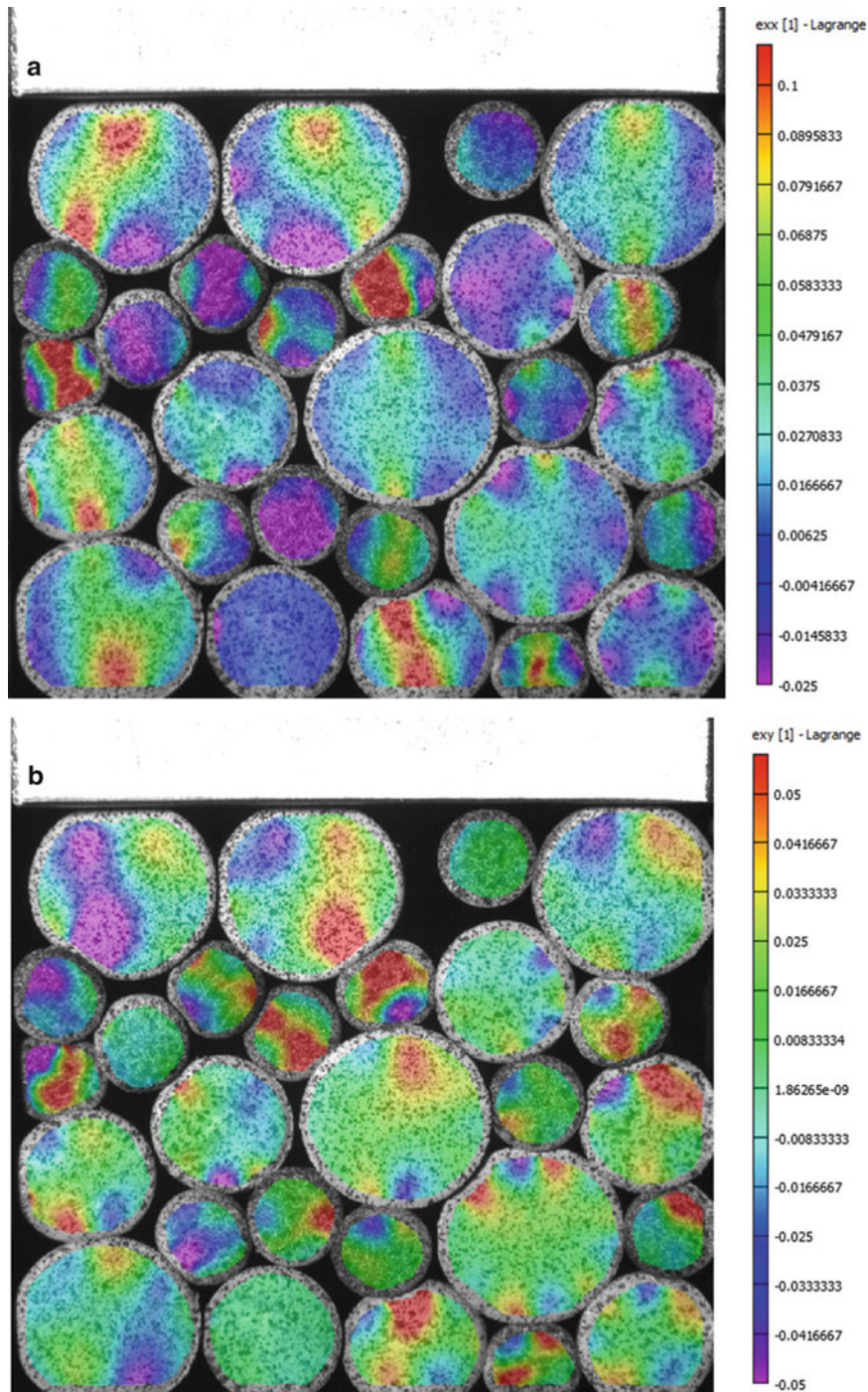
Fig. 45.3 The full-field strain for impact experiment for polyurethane grains for (**a**) εxx, (**b**) εxy and (**c**) εyy at some instant
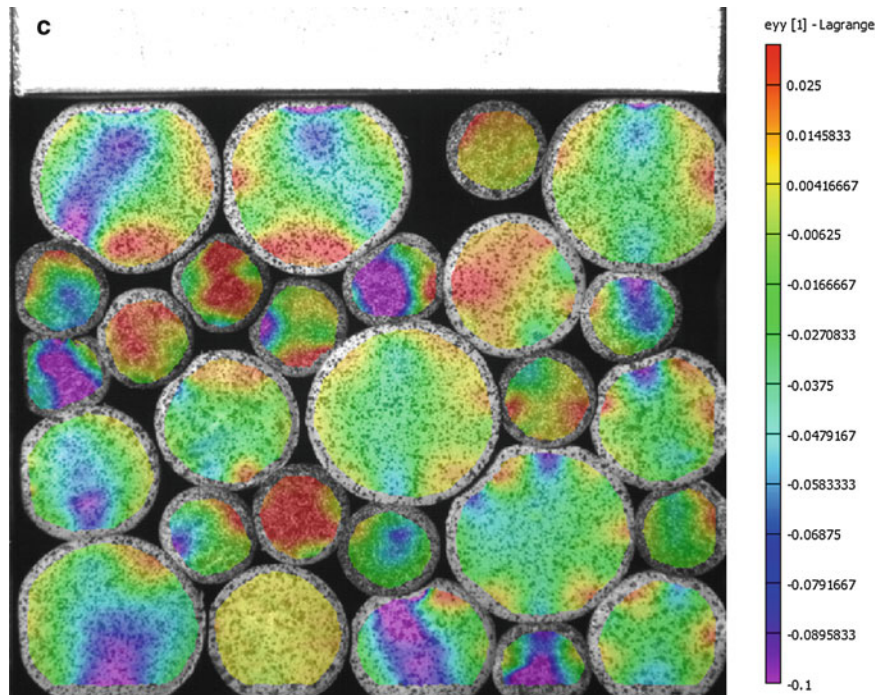
Fig. 45.3 (continued)

importance of collisions in energy dissipation in the granular media. A comparison between Fig. 45.4a and b and Fig. 45.4c and d reveals that although ignoring influence of local acceleration fields does results in errors up to 10 N in the inter-particle force values in the current granular assembly.

A similar analysis was also conducted for polycarbonate grains but due to the comparably higher wave speeds in polycarbonate, the impact event lasted for a shorter time and hence fewer images could recorded for the experiments. Since the polycarbonate grains undergo smaller deformations, the images were recorded at a higher magnification and inter-particle forces were calculated for plane strain and plane stress assumptions. As seen in Fig. 45.5, there is no significant difference between plane stress and plane strain cases for polycarbonate. The collisions once again resulted in large inter-particle forces in polycarbonate granular system, thus reiterating the importance of collisions in the energy dissipation in granular media.

45.4   Conclusions

A method for quantitatively measuring dynamic inter-particle force transmission in granular materials is presented in the current work. The proposed method has been demonstrated to have the ability to test two different materials with modulus orders of magnitude apart. These results illustrate a new opportunity to experimentally validate theories linking microscopic inter-particle forces with macroscopic behavior in a variety of applications. The extension of the current approach to testing of stiffer materials is currently restricted by high-speed cameras that can provide high resolution images.

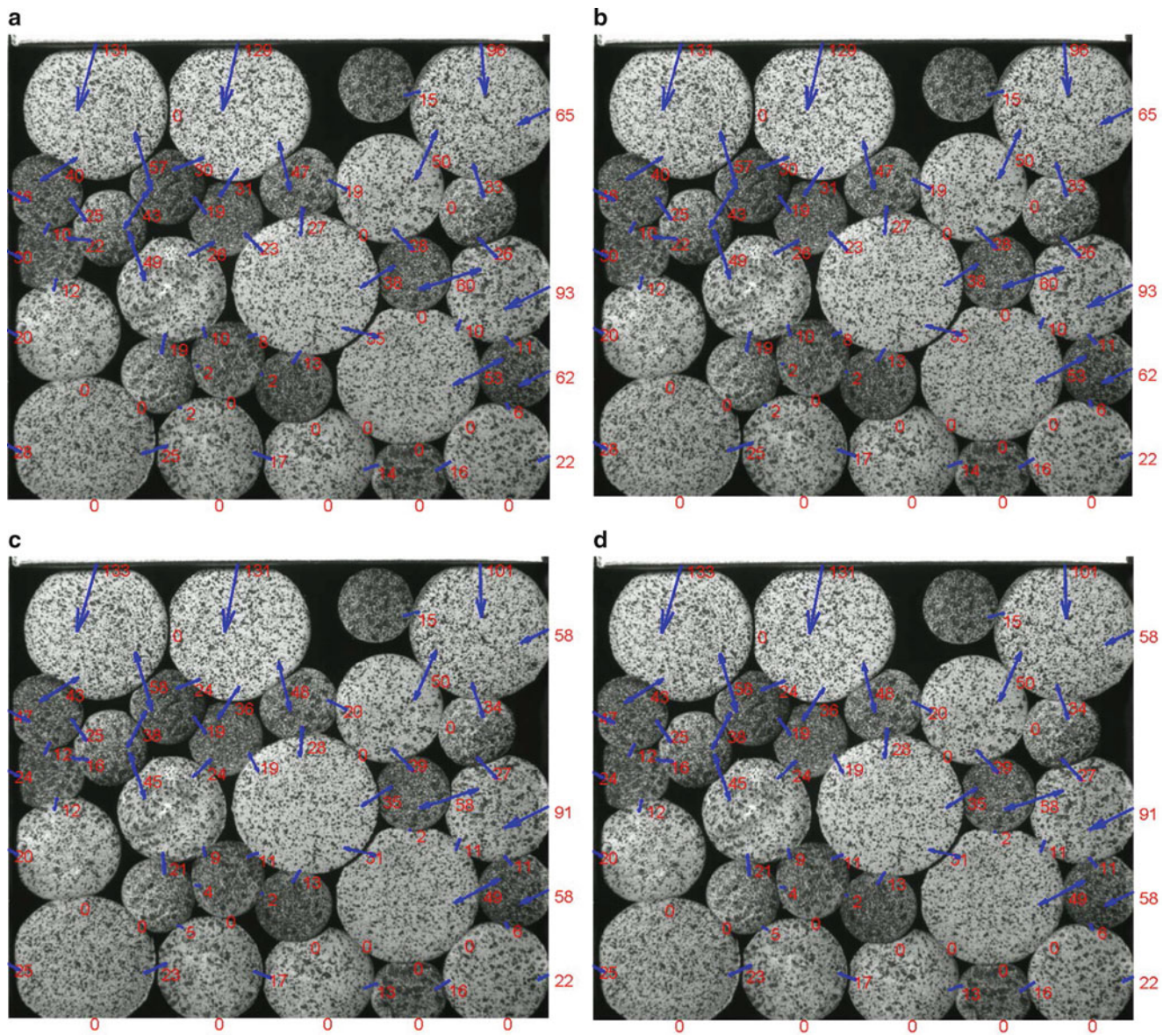**Fig. 45.4** The inter-particle forces (N) obtained in polyurethane grains using the GEM based approach described in [13] along with terms accounting for gradients local acceleration fields in each grain for (**a**) plane stress and (**b**) plane strain. Inter-particle forces (**N**) in polyurethane grains ignoring influence of local acceleration fields for (**c**) plane stress and (**d**) plane strain
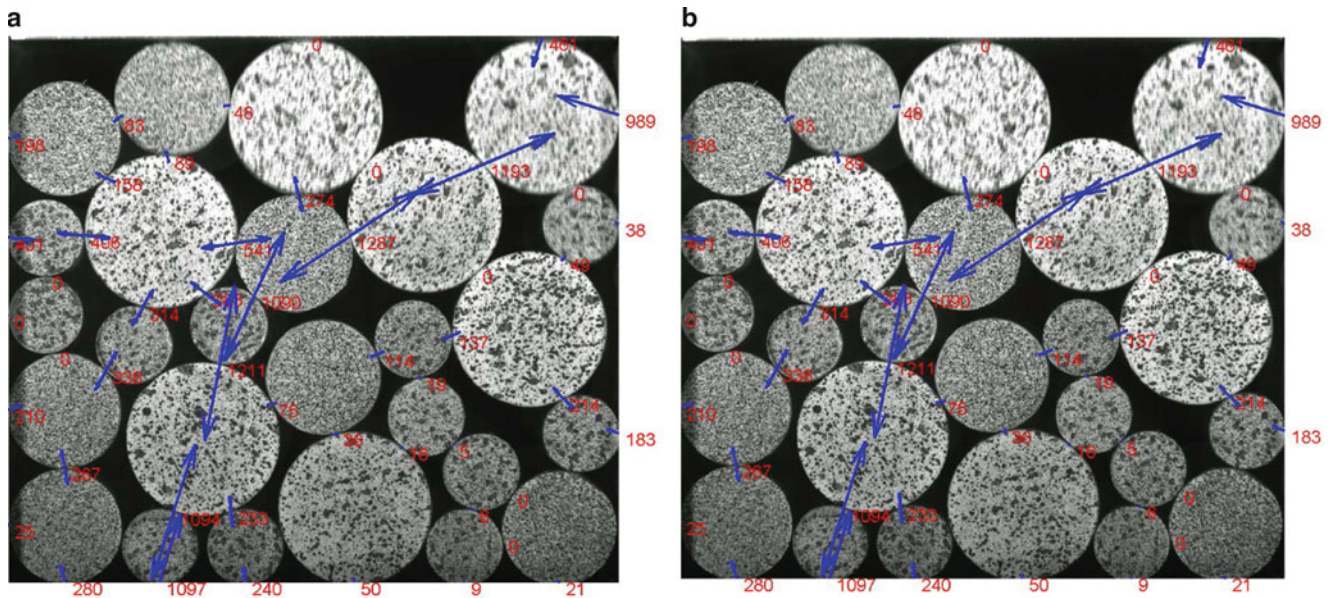
Fig. 45.5 The inter-particle forces (**N**) obtained in polycarbonate grains using the GEM based approach described in [13] along with terms accounting for gradients local acceleration fields in each grain for (**a**) plane stress and (**b**) plane strain

References

1. Andrade JE, Borja RI (2006) Capturing strain localization in dense sands with random density. Int J Numer Methods Eng 67:1531–1564
2. Dafalias YF, Popov EP (1975) A model of nonlinearly hardening materials for complex loadings. Acta Mech 21:173–192
3. Christoffersen J, Mehrabadi MM, Nemat-Nasser S (1981) A micromechanical description of granular material behavior. J Appl Mech 48 (2):339–344
4. Bathurst RJ, Rothenburg L (1990) Observations on stress-force-fabric relationships in idealized granular materials. Mech Mater 9(1):65–80
5. Somfai E, Roux J-N, Snoeijer JH, Van Hecke M, Van Saarloos W (2005) Elastic wave propagation in confined granular systems. Phys Rev E 72:021301
6. Clark AH, Kondic L, Behringer RP (2012) Particle scale dynamics in granular impact. Phys Rev Lett 109:238302
7. Dantu P (1957) Contribution a l'etiude mecanique et geometrique des milieux pulverulents. In: Proceedings of the 4th international conference on soil mechanics and foundation engineering, vol 1. Butterworths, London, pp 144–148
8. Wakabayashi T (1957) Photoelastic method for determination of stress in powdered mass. In: Proceeding of the 7th national congress applied mechanics. Japan National Committee for Theoretical and Applied Mechanics, Science Council of Japan, Tokyo, Japan, pp 153–158
9. Saadatfar M, Sheppard AP, Senden TJ, Kabla AJ (2012) Mapping forces in a 3D elastic assembly of grains. J Mech Phys Solids 60(1):55–66
10. Zhou J, Long S, Wang Q, Dinsmore AD (2006) Measurement of forces inside a three-dimensional pile of frictionless droplets. Science 312 (5780):1631–1633
11. Andrade J, Avila CF (2012) Granular element method (GEM): linking inter-particle forces with macroscopic loading. Granul Matter 14:51–61
12. Hurley R, Marteau E, Ravichandran G, Andrade JE (2014) Extracting inter-particle forces in opaque granular materials: beyond photoelasticity. J Mech Phys Solids 63:154–166
13. Hurley R, Lim K-W, Andrade JE (2014) Measuring dynamic force transmission in granular materials (in review)

Chapter 46
High Pressure Burst Testing of SiC$_f$-SiC$_m$ Composite Nuclear Fuel Cladding

Luis H. Alva, Xinyu Huang, George M. Jacobsen, and Christina A. Back

Abstract Silicon carbide fiber reinforced silicon carbide matrix (SiC$_f$-SiC$_m$) composite tube is being developed as an accident tolerant fuel cladding material for light water reactors. The mechanical robustness of the material is of critical importance to ensure that the nuclear fuel and fission products are contained during both normal operation and accident conditions, such as the loss of coolant accident (LOCA) encountered in Fukushima Daiichi Plant. We report the development and application of a high pressure burst testing method to evaluate the mechanical robustness of SiC$_f$-SiC$_m$ composite cladding. The internal high pressure is generated using a rubber tubing placed within the SiC$_f$-SiC$_m$ sample tube. A unique test rig was designed to seal the ends of the rubber tubing while pressurizing it up to 2,000 bar by hydraulic oil. The expanding rubber tubing confined by the SiC$_f$-SiC$_m$ sample tube thus exerts a controlled uniform internal pressure to the SiC$_f$-SiC$_m$ sample. The full-field strain distribution of the outside surface of the sample was captured by 3D digital image correlation (DIC) method. The acoustic emission (AE) technique was used to detect damage events during the high pressure burst testing.

Keywords Nuclear fuel cladding • Digital image correlation • Internal high pressure testing • Strain measurement

Nomenclature

LOCA	Loss of coolant accident
DIC	Digital image correlation
SiC$_f$-SiC$_m$	Silicon carbide fiber–silicon carbide matrix
CVI	Carbon vapor infiltration
O.D	Outside diameter
I.D	Inside diameter
MPa	Mega Pascals, 10^{+6} Pascals
AJ	Atta Joules, energy unit, 10^{-18} J
Psi	Pounds per square inch
PLS	Proportional limit stress
μ-strains	Microstrains

L.H. Alva • X. Huang (✉)
University of South Carolina, 300 Main St., Columbia, SC 29208, USA
e-mail: huangxin@mailbox.sc.edu

G.M. Jacobsen • C.A. Back
General Atomics, San Diego, CA 92186, USA

46.1 Introduction

Zircaloy tube has been successfully used as fuel cladding in light water reactors for many decades. The 2011 accident of the nuclear power plant in Japan (Fukushima Daiichi) highlighted a known weakness of the Zircaloy cladding, i.e., low melting temperature and undesirable reactivity of Zircaloy with water at high temperature. The metal-water reaction produces hydrogen, an explosion hazard. Alternative accident tolerant fuel claddings are being actively developed. Among them engineered silicon carbide (SiC) structures consisting of a layer of monolithic SiC or other material for hermeticity and an silicon carbide fiber reinforced silicon carbide matrix ceramic composite (SiC_f-SiC_m) [1, 2] layer for strength and toughness are a promising candidate.

Nuclear fuel cladding tube is designed to hold the fuel pellets (typically UO_2) in the designed geometry and distribution within the reactor core, allowing efficient transfer of heat from the nuclear fuel pellet to the coolant, and prevent the release of radioactive fission products into the coolant loop and outside environment [3]. The mechanical robustness of the cladding tube is of critical importance to ensure that the radioactive fission products are retained during both normal operation and accident conditions, such as the loss of coolant accident (LOCA) encountered in Fukushima Daiichi plant. As a ceramic material, the mechanical robustness of SiC based ceramic composite cladding needs to be carefully studied to understand its short-term and long-term failure modes, and to establish the probability of mechanical failure under relevant conditions. Mechanical testing is necessary to acquire material properties and to validate prediction of analytical models.

It is important to consider the stress state of the materials during service when selecting mechanical testing technique to determine the relevant mechanical properties. In the case of the fuel cladding materials, the failure typically occurs by fracture due to excess hoop stress [4] generated by internal pressure and temperature gradient. The mechanical characterization of cylindrical ceramic components has been made by tensile tests, flexural beams bending test, and C-ring and O-ring tension or compression test. For these tests, high stress occurs only in a small portion of the sample, and large stress gradient exists. These are not representative of the working conditions of the nuclear fuel cladding. With the exception of local PCMI (Pellet Cladding Mechanical Interaction) and clad-grid interaction, the fuel cladding tube will experience differential pressure exerted on the outside and inside surfaces and temperature gradient induced thermal stress and irradiation swelling induced stress [5]. The internal pressure test better represents the load condition of ceramic composite cladding tubes during normal operating and severe accident conditions. The pressurized cylindrical tube test, or burst test, is very reliable, relatively easy to implement, and economical, and can be applied to small and large samples [6].

For lab testing, the controlled internal high pressure can be generated by a fluid media, a rubber plug [7], or an expanding solid plug (with split segments) [8]; in case of ceramics and SiC composites the preferred test is the expanding internal plug method because is considered safer than using a gas or liquid media because of the high internal pressure needed [9]. We report a novel high pressure burst testing method for evaluating the mechanical robustness of the SiC_f-SiC_m with internal pressure being generated by hydrostatic pressure contained within a rubber tubing (as bladder) placed inside the SiC_f-SiC_m sample tube. The use of rubber bladder not only improves the pressure uniformity but also allows the testing of samples which have liquid permeable porous wall. This approach also reduces the axial friction problems encountered in the rubber plug test and removes the uncertainty of the applied pressure in the rubber plug test. Acoustic Emission (AE) is a passive sensing technique to detect damage initiation and propagation during the test. Digital Image Correlation (DIC) method is a non-contact full field strain and displacement measurement method used to calculate the strains on the outer surface of the sample during the test. Both can be applied during the pressure test and provide rich information for better understanding mechanical responses of the materials.

46.2 Experimental Details

The SiC_f-SiC_m sample tubes were fabricated at General Atomics (GA) using nuclear grade SiC fibers, pyrolytic carbon interface coating, and CVI SiC matrix [10]. This sample does not have a monolithic SiC layer, does not represent a fully engineered multilayer SiC cladding design, however, the SiC_f-SiC_m composite layer is the primary load-bearing component. The sample tube has an O.D. 10.85 mm, wall thickness of 1.45 mm and length 39 mm. The speckle pattern was first applied to the sample by painting small spots of around 5 pixels by 5 pixels (as viewed by CCD camera) on the surface of the sample using regular white paint, as shown in Fig. 46.1. Two sets of strain rosette gages (Micro Measurement, 350 Ω, active length 3.18 mm) were installed in the sample to monitor the axial and hoop strains on the tube surface at diametral locations, one of the two rosette gage can be seen in the picture in Fig. 46.1. The strain data from the gages are to be compared with the strains calculated with the DIC method. Due to the inherent surface roughness of the composite tube, it is challenging to obtain

Fig. 46.1 Speckle pattern (1)
and strain gages installed
in sample; Sample without
speckle pattern (2)

Fig. 46.2 Position of cameras
respect to sample
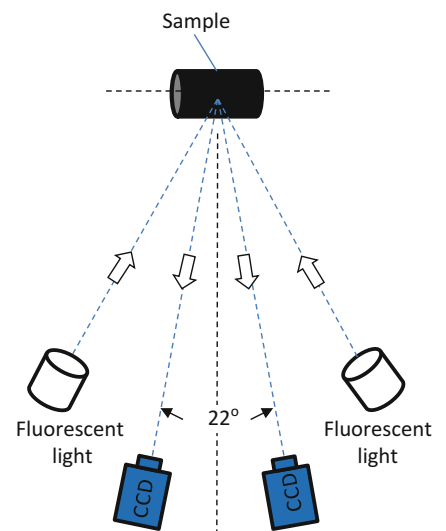
uniform adhesion of rosette gages to the surface. Some spots under the foil gage may not be bonded as strong as other spots. At large strain, the foil gage tends to debond locally and results in errors in measured strain. However, at low strain amplitudes, it is believed that the strain gages reading will closely represent the average strain of the tube surface under internal pressure.

The internal pressure in this study is generated using a piston-type hydraulic pressure generator (Model 37-6-30 from High Pressure Equipment Co. Erie PA) that feeds hydraulic oil into flexible rubber tubing placed within the SiC_f-SiC_m sample tube. Loading-unloading tests were performed using a SiC_f-SiC_m sample tubes. During the test the internal pressure was increased by displacing the high pressure cylinder shaft of the hydraulic pressure generator, from zero to a predetermined internal pressure and then decreasing to zero, completing a cycle. The loading-unloading continues for the next predetermined internal pressure. Each loading-unloading cycle is achieved in less than 60 s to reduce the effect of slow crack growth. The internal pressure was varied in the range of 6.9 MPa (1,000 psi) to 96.5 MPa (14,000 psi), the pressure was monitored with an analog pressure gage. Simultaneously images of the SiC_f-SiC_m sample tube were acquired with CCD cameras. These images were later used to calculate the strains on the outside surface by DIC method. Data from the strain gages attached to the aluminum adapters was recorded during the experiment using a NI-cDAQ-9172 chassis with two NI-9237 modules and LabView software. To detect the evolution of damage within the SiC_f-SiC_m sample tube, the signal from AE sensor was recorded for posterior analysis. The AE equipment is a Micro-II Digital AE System (Physical Acoustics Corporation) equipped with a NANO-30 AE sensor and 60 DB pre-amplifier. The data was analyzed using the AEWin software. The AE sensor (NANO-30) was coupled to the sample with play dough and held in place by rubber band.

The digital image correlation (DIC) method is applied to measure the surface strains. DIC is a non-contact full-field strain measurement technique with a broad range of applications [11]. Due to the curved nature of the sample, three-dimensional DIC method enabled by two cameras is used. Two CCD cameras from Point Grey (GRAS-30S5M) are used for image acquisition. Figure 46.2 shows the spatial arrangement of the cameras with respect to the sample. The camera-to-sample

distance is 220 mm, and the angle between the two cameras is 22°. The progressive scan CCD chip in the camera is Sony ICX625AL, with 2/3″ or 11.016 mm diagonal length, 5.05 million (2,448 by 2,048) square pixels. Two 25 mm lenses (NT63-780 from Edmund Optics) are attached to the cameras. The sample is illuminated by two compact fluorescent light lamps. The Vic-3D software from Correlated Solutions, Inc. was used to analyze the collected images and to calculate the strains on the surface of the SiC_f-SiC_m sample tube. The correlation algorithm used is based on the normalized square differences criterion which has better tolerance to changes in lighting. The analysis was done using a subset of 29 pixels, and a step of 7 pixels. The image acquisition is made using the Vic-Snap 8 software using a frame rate of 2 per second which is the maximum allowed with the current hardware.

46.3   Results and Discussions

A typical hoop and axial strain mapping of the outer surface of the sample obtained by correlating the images of the SiC_f-SiC_m sample tube is shown in Fig. 46.3, for an internal pressure of 96.5 MPa (14,000 psi). The DIC correlation software reproduces the irregularity of the surface of the sample with great detail. The values of the Lagrange strains on the surface of the sample are shown in different colors, where each color corresponds to a strain value according to the scale bar on the right of Fig. 46.3. The hoop-strain plot shows that the deformation of the sample is not uniform along the surface; the calculated strains have different values depending on the location on the surface of the sample. The non-uniformity is likely due to the texture of the filament wound composite tube which can affect outer surface strain distribution and pressure distribution at inner surface. The foil gage bonded nearby may have also affected the DIC reading. It is important to mention that the typical error for strains when calculated using DIC is about 200 μ-strains [11, 12], this must be taken into account when interpreting the values of strains in the results.

The axial strains from DIC (shown in Fig. 46.3.) shows that the sample is under very low compression loading, as expected from the rubber bladder approach. The average axial strain is −88 μ-strains in a range of −400 to 25 μ-strains and strain resolution of the DIC method is ~200 μ-strains. The slight negative axial strains at the internal pressure of 96.5 MPa (14,000 psi) is likely due to the Poisson effect of the sample in combination with a small tensile force induced by the aluminum adapters (which tend to shrink) due to its smaller ID. It is believed that the aluminum tube exerted a small tensile force on the SiC sample at the stepped lap joint transmitted through friction.

The hoop stress in the SiC_f-SiC_m sample tube is calculated at the outer diameter using thick-walled cylinder theory for the case of internal pressure only [13] using the equation,

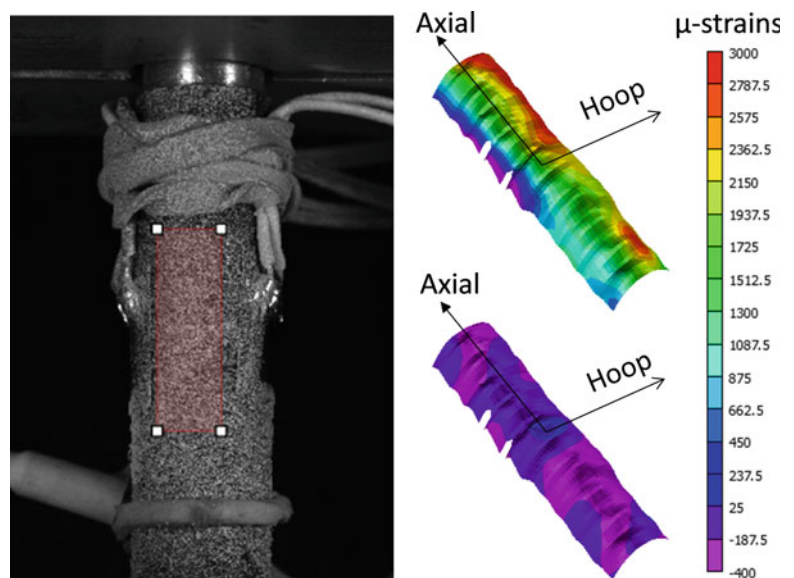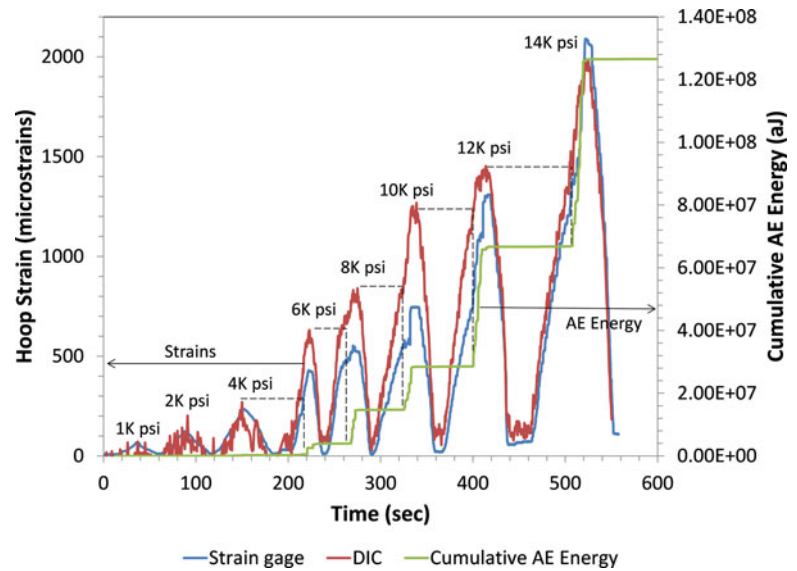$$\sigma_h = P \frac{2r_i^2}{r_o^2 - r_i^2} \qquad (46.1)$$

Fig. 46.3 DIC plotting of hoop and axial Lagrange strains at an internal pressure of 96.5 MPa (14,000 psi)

Fig. 46.4 Cumulative AE
Energy and Calculated Hoop
Stress vs. Time

where:

σ_h = hoop stress on the outer surface
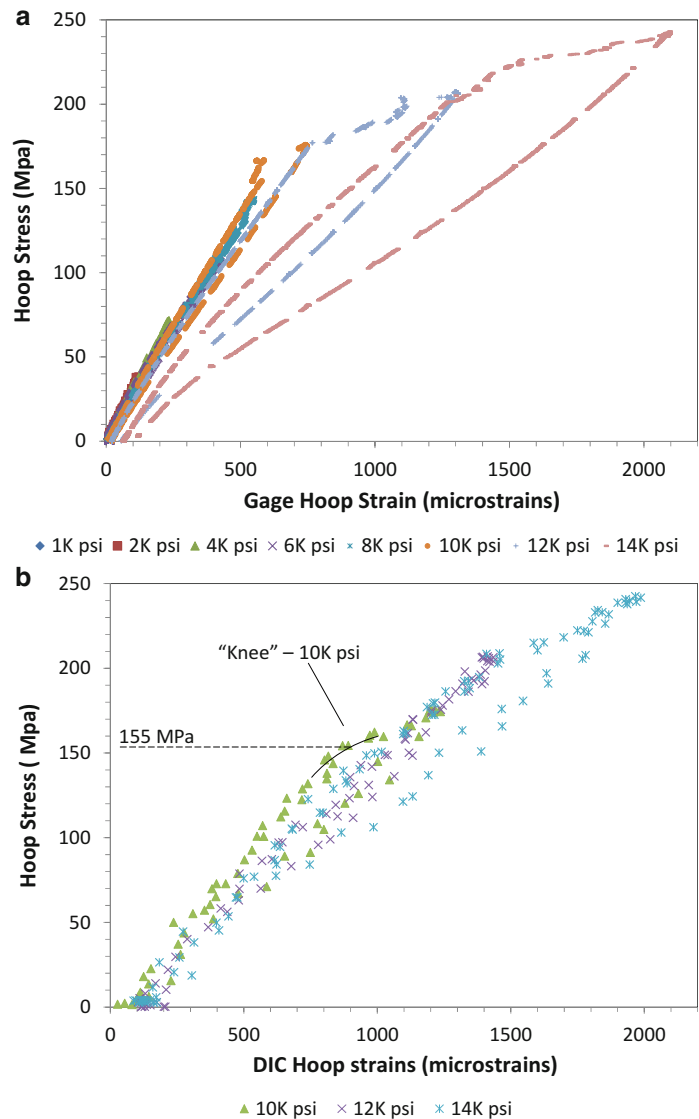P = internal pressure
r_i = inner radius of the sample
r_o = outer radius of the sample

The internal pressure (P) in Eq. (46.1) is back calculated using the hoop and axial strains measured on the aluminum adapter, based on thick-wall cylinder theory and generalized Hooke's law. The calculated internal pressure P matches closely to that indicated on the analog pressure gage of the pressure generator.

During the test the internal pressure was varied between 6.9 MPa (1,000 psi) and 96.5 MPa (14,000 psi), in loading-unloading cycles. Results for the cumulative AE energy and the average hoop strains plotted vs. the time are shown in Fig. 46.4. Surface strain data measured by the DIC method and the strain gage are both shown in Fig. 46.4. For internal pressures of less than 27.6 MPa (4,000 psi) the strain calculated using the DIC method and strain gage readings track each other fairly well, except that the DIC data have more noise than the strain gage readings. Above 27.6 MPa (4,000 psi) the obtained DIC strain values are higher than that of the strain gage readings, except at 96.5 MPa (14,000 psi). The SiC$_f$-SiC$_m$ sample has an uneven and textured surface which contains rather large wavy fiber tows, as shown in Fig. 46.1. The active element of strain gage is about 2 ~ 3 × the width of the fiber tow. Before mounting the strain gage, the SiC$_f$-SiC$_m$ sample surface was not abraded due to concerns of abrasive affecting the strength properties of the sample; hence the adhesion quality may not be as good and as uniform as what can be achieved on abraded flat metal surfaces. Under large hoop stress, the fiber tows will tend to rotation (scissoring) and align with the principal stress direction. This might have induced some local debonding of the strain gage, which reduces the strain gage reading. However, the edge of the foil gage may still be bonded, so when the sample stretch to a certain extent, the active element gage gets pulled and bend over the protrusion of the fiber tow, the gage bending strain may have caused the strain gage readings exceeding that of the DIC value in the last loading-unloading cycle. In contrast, the DIC, as a non-contact method, can tolerate surface irregularities much better than the strain gages up to relatively high strain amplitude before the speckle paint layer starts to spall off.

The cumulative AE energy is an indicator of the damage accumulation process in the SiC$_f$-SiC$_m$ sample during the test. The hoop stress on the outer surface of the sample can be calculated by Eq. (46.1); when this hoop stress is below 41.4 MPa (6,000 psi internal pressure) the cumulative AE energy is very low. The AE energy sharply increases whenever the strain surpasses the highest peak strain of the past loading history, known as Kaiser effect [14]. The AE energy curve thus forms a distinct stair shape. The rise of AE energy is believed to be associated with internal material damage events, such as matrix cracking and fiber sliding. For this test, the SiC$_f$-SiC$_m$ sample tube did not fail at the end of this test (pressurized up to 14,000 psi) and no apparent cracks or permanent deformation could be seen by visual examination however some internal damage has occurred as indicated by the emission of acoustic waves [15]. In another test, failure of the sample occurred at an internal pressure of 131 MPa (19,000 psi).

Fig. 46.5 Hoop stress and strain on the outside surface of the sample. Strain gage (top) and DIC Method (bottom)

The hoop stress–strain curves on the outer surface of the sample are constructed using the measured hoop strain and the calculated hoop stress (from internal pressure). The results are shown in Fig. 46.5. The strain in the top figure is measured with strain gage, the strain in the bottom figure is measured with DIC method. The hysteresis of the loading-unloading cycles are clearly seen for each loading and unloading cycle in the top stress–strain curves based on the strain gage data. The hysteresis is not as clear in the stress–strain curve based on the DIC hoop strains (lower figure), the last three hysteresis loops can be vaguely seen. The reason is the lower strain resolution and the low density of data points due to sustainable frame rate limit. The bandwidth of the data bus (from CCD camera to computer) and the write speed of the hard drive limits the usable frame rate to 2 frame per second. A "knee" (bend) is clearly indicated in the stress–strain curve from strain gage data. The knee corresponds to an internal pressure of 69 MPa (10,000 psi). For ceramic composite, the "knee" is typically observed in the uniaxial tension test where the knee indicates the PLS (Proportional Limit Strain); for the case of the S_iC_f-S_iC_m sample it is believed to be the stress level when cracking of the SiC matrix occurs [10]. At the knee point, the hoop stress at the outside diameter is around 160 MPa, the hoop strain is 850 μ-strains based on DIC data and 550 μ-strains based on the strain gage measurement.; this difference in the strain value is also found in Fig. 46.4.

46.4 Conclusions

A test rig for high pressure burst test of nuclear fuel cladding tubes was designed, constructed and used to test S$_i$C$_f$-S$_i$C$_m$ nuclear fuel cladding samples. The rubber tubing can effectively transmit the internal pressure to the sample and the end sealing features resisted a pressure of 131 MPa (19,000 psi) without oil leaking. This scheme allowed us to perform high internal pressure test of tubular samples of various length (from 0.5 in. to 10s of inches) and samples which may be permeable to liquids. End effects were minimized by a metal adapter that partially relieves the bladder pressure at the end of the tube. With the rubber tubing separating the cladding tube and the pressurizing media, the sample will not be directly exposed to the pressurizing media, and the risk of sample contamination minimized. This makes it possible to apply it as a proof test method for quality control. The as-fabricated cladding tubes can be internally pressurized and the AE technique can be used to detect internal damage event and serve as pass (or fail) criterion. The burst test method can also be applied to other load bearing tubular ceramics components, such as hot gas filters [7] in coal fired power plants, and advanced gun barrels [9], where internal pressure is an important service load.

Our preliminary work showed that the surface deformation of the sample can be captured by the DIC method, however using DIC data alone, the "knee" behavior of the stress–strain curve is blurred due to relatively higher noise level of the DIC strain data. The strain capture by the traditional foil gage on the sample clearly indicates the "knee" behavior, which starts around 600 micro strain. The advantage of the DIC, the non-contact and full field measurement capability, provides strain distribution on the outer surface of the sample. The highly non-homogeneous strain field due to the intrinsic texture of the composites is clearly revealed by DIC. The strain gage rosette reading represents an average strain of a finite area on the surface. The thick-walled cylinder theory was used to calculate hoop stress from internal load. The AE sensor captured the acoustic emission produced inside the sample as shown in the results. The cumulative AE energy suggested the gradual progression of damage in the sample represented by the Kaiser effect. More experiments are needed to differentiate AE signal from material and from the friction between the adapter and the sample. A more thorough study of the material using the high pressure test rig is currently in progress.

References

1. Yueh K, Carpenter D, Feinroth H (2010) Clad in clay. Nucl Eng Int:14–16
2. Herderick ED, Cooper K, Ames N (2012) New approach to join SiC for accident-tolerant nuclear fuel cladding. Adv Mater Process 170(1):24–27
3. Azevedo CRF (2011) Selection of fuel cladding material for nuclear fission reactors. Eng Fail Anal 18:1943–1962
4. Desquines J, Koss DA, Motta AT, Cazalis B, Petit M (2011) The issue of stress state during mechanical tests to assess cladding performance during a reactivity-initiated accident (RIA). J Nucl Mater 412:250–267
5. Ben-Belgacem M, Richet V, Terrani KA, Katoh Y, Snead LL (2014) Thermo-mechanical analysis of LWR SiC/SiC composite cladding. J Nucl Mater 447:125–142
6. Chuck L, Graves GA (1997) Hoop tensile strength and fracture behavior of continuous fiber ceramic composite (CFCC) tubes from ambient to elevated temperatures. J Compos Technol Res 19(3):184–190
7. Huang X (2001) Mechanics and durability of fiber reinforced porous ceramic composites. Dissertation, Department of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University
8. Nilsson KF, Martin O, Chenel-Ramos C, Mendes J (2011) The segmented expanding cone-mandrel test revisited as material characterization and component test for fuel claddings. Nucl Eng Des 241(2):445–458
9. Carter R (2006) Compressed Elastomer method for internal pressure testing. ARL-TR-3921, Army Research Laboratory
10. Back CA, Khalifa HE, Deck CP, Samuli B, Hilsabeck T (2012) Fabrication of SiC-SiC composites for fuel cladding in advanced reactor designs nuclear. Prog Eng 57:38–45
11. Sutton MA, Orteu JJ, Schreider HW (2009) Image correlation for shape, motion and deformation measurements. Springer, New York
12. Lyons JS, Liu J, Sutton MA (1995) High-temperature deformation measurements using digital-image correlation. Exp Mech 36:64–70
13. Seely FB, Smith JO (1952) Advanced mechanics of materials. Wiley, New York
14. Physical Acoustics Corporation (2007) PCI-2 Based AE System Manual
15. Nozawa T, Ozawa K, Tanigawa H (2013) Re-defining failure envelopes for silicon carbide composites based on damage process analysis by acoustic emission. Fusion Eng Des 88:2543–2546

Chapter 47
Low Cost Digital Image Correlation (DIC) for Monitoring Components Undergoing Fatigue Loading

R.K. Fruehmann and J.M. Dulieu-Barton

Abstract An alternative approach to dealing with noise in digital image correlated data. The target application is for non-destructive inspection of components outside of laboratory conditions, where oscillatory loading (either as part of the normal service or especially excited) is used to produce the strain changes required to determine the strain field. The approach is to use a lock-in amplifier, obtaining the reference signal from an accelerometer on the structure under scrutiny. The challenge of low sampling rates relative to the loading excitation is considered, and it is demonstrated that by careful selection of sampling and loading frequency combinations, standard low speed, low cost digital cameras can be used to obtain high fidelity strain measurements with a fine spatial resolution.

Keywords Digital image correlation • Non-destructive testing • Lock-in amplifier • Cyclic testing • Brazilian disc

47.1 Introduction

The work described in the paper is part of a larger project focusing on the development of full-field optical non-destructive inspection (NDI) techniques for use in a service environment. The use of a strain based approach is considered advantageous because it enables the service strains to be evaluated, thereby giving a quantitative assessment of the severity of damage or defects with respect to material failure criteria. NDI techniques that can be applied without the use of a test machine have been considered, and the approach deemed most versatile was to use cyclic loading, for example natural frequency excitations or vibrations resulting from normal service loads such as pumps or wind induced vibrations, typically providing rapidly changing, transient strain states. An additional challenge comes from operators favouring inspections to be conducted at low strain levels, to avoid damage resulting from the inspection process.

Digital image correlation (DIC) is a well-established, non-contacting technique for obtaining full-field surface strain data [1]. The technique is versatile and simple to implement, and well suited to large deformations. Part of the versatility stems from the ability to use a wide range of low cost digital cameras, and standard off the shelf components for lighting and camera fixation. However, accurate results at low strain levels is more challenging, and with the addition of transient strain changes, a modified approach is required to achieve the fidelity required for NDI. The use of a cyclic load provides an opportunity for filtering, known to the authors from the field of thermoelastic stress analysis (TSA), a competing non-contact stress analysis technique that uses thermal imaging to measure the small temperature changes associated with the thermoelastic effect [2]. In TSA, a lock-in (LI) amplifier is applied, using a reference signal to reject noise in the temperature measurement. The approach presented in this work uses the idea of the LI amplifier to filter strains derived from DIC to improve the strain resolution. However, when employing standard digital cameras for DIC, the additional problem of low sampling rates and the consequent under sampling of the signal of interest, poses an additional challenge. In the paper, an experiment that utilises a 'Brazilian disc' (i.e. a disc in two-point diametral compression) demonstrates the improvement in strain resolution achieved by using the LI amplifier to filter DIC strain data. Through careful selection of loading and sampling frequencies strain changes occurring faster than the available sampling rate can none-the-less be captured accurately.

R.K. Fruehmann • J.M. Dulieu-Barton (✉)
Faculty of Engineering and the Environment, University of Southampton, Highfield, Southampton SO17 1XR, UK
e-mail: janice@soton.ac.uk

47.2 Methodology

In its basic form, the LI amplifier multiplies a measurement signal with a reference signal and then integrates the result over a number of periods [3, 4]. The output is a value proportional to the amplitude of the measurement signal. The effect is like a band pass filter; the larger the number of periods in the analysed signal, the narrower the band pass and the better the noise rejection. For this, it is important that the reference signal has a high signal to noise ratio and that it has exactly the same frequency as the component of the measurement signal of interest.

The implementation of the LI amplifier for DIC in this work can be divided into three parts. First, a series of images of the specimen are recorded while undergoing a cyclic load. A reference signal value is also recorded with each image. In the current work a servo-hydraulic test machine (Instron 8802 with a 100 kN actuator and load cell) was used so the reference signal was taken from the load cell. However, the reference signal could come from any source directly related to the specimen loading (e.g. strain gauge, accelerometer...) provided the signal to noise ratio is large. The second step is to obtain a corresponding series of strain fields, each calculated relative to the first image in the series. These strain values form the set of measurement signals: one signal for each correlation cell in the DIC. Finally, the reference signal is converted into two reference signals, 90° out of phase with each other, and with their amplitude normalised to 1. The LI algorithm was then implemented using Matlab 2010b. For this, the strain images were imported into 2D arrays and the amplitude of the strain was obtained for each element, giving an image of strain amplitudes.

The use of standard digital cameras for DIC leads to a restriction on the maximum sampling rate, typically in the range of 1–10 Hz. According to Shannon's sampling theorem [5], the sampling frequency must be twice the maximum signal frequency in a band-limited signal, known as the Nyquist condition. This would limit the measurement signal frequency to 0.5–5 Hz, severely restricting the range of practical applications. However, the problem of aliasing that arises when confronted with non-band limited signals provides an opportunity where it is possible to analyse a signal with a frequency greater than the sampling frequency for specific cases, sometimes referred to as sub-Nyquist sampling [6]. In this work, where the sampling frequency and measurement frequency are known a priori and can be adjusted to suit, it is possible to setup the sampling and loading conditions to enable sub-Nyquist sampling. Hence it becomes possible to apply DIC using standard low frame rate digital cameras, even for relatively high excitation frequencies.

47.3 Experimental Arrangements and Results

The Brazilian disc set-up was used because when loaded the disc has a strong biaxial strain field with regions of relatively uniform, low strains as well as a very large strain gradient in the y-direction towards the loading points. This forms a challenging problem as both high strain resolution and high spatial resolution are required, two criteria that generally contradict each other, i.e. high strain resolution requires large interrogation cell sizes which in turn leads to low spatial resolution, and vice versa. Furthermore, an analytical solution exists describing the strain field; this enables the accuracy of the experimental results to be assessed. The disc used in this study was manufactured from a 6 mm thick 6061-T6 aluminium plate, with a diameter of 80 mm, loaded between two flat mild steel plates. A load range of −9 kN was applied; for the static case the specimen held at −0.5 and −9.5 kN, for the dynamic case the specimen was loaded at an amplitude of 4.5 kN about a mean of −5.0 kN.

A Manta G504-B from Allied Vision with a 2,452 × 2,056 sensor with 3.45 μm pitch, and an 8 bit grey scale was used with a 50 mm Nikon lens. The frame rate was set at 2 Hz. For the static case, images were recorded at each load, to provide a strain field. The strains obtained using the static test represent current standard practice. For the dynamic load cases, 251 images were recorded. For the LI analysis, the number of strain fields used was then reduced, depending on the loading and recording frequency, to cover an exact integer number of load cycles to meet the truncation requirements [7]. To demonstrate the ability to obtain accurate strains from a sub-Nyquist sampling, dynamic tests were run at 0.75 (Nyquist condition met), 7.1 and 21.1 Hz (sub-Nyquist sampling). The strains were calculated using DaVis version 8. A cell size of 21 × 21 pixels with an overlap of 11 pixels was used for all the data presented here. All data were processed using the same circular mask and seeding point (placed towards the centre of the masked area).

A visual comparison of the standard and the LI approaches is presented in Fig. 47.1, where the y-direction strains are shown. The image is divided into quarters; the top left shows the results from a static test using one image obtained at −0.5 kN and a second at −9.5 kN to obtain the strain field, the top right shows the result obtained by applying the LI amplifier to obtain the strain amplitude of the disc loaded at 0.75 Hz, the bottom left shows the result loaded at 7.1 Hz

Fig. 47.1   strains (ε_{yy}—*vertical axis*)
calculated by the standard approach
(*top left*), using the LI amplifier applied
to a 0.75 Hz loading frequency (*top right*),
using the LI amplifier applied to a 7.1 Hz
loading frequency (*bottom left*) and the
theoretical value (*bottom right*)
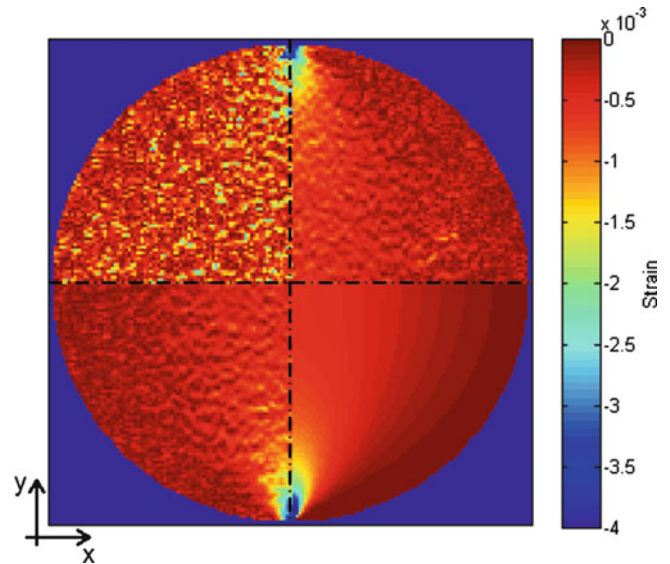
(sub-Nyquist sampling) and the bottom right shows the theoretical strains, calculated using an analytical solution presented in [8]. The difference between the static case and the two dynamic load cases is significant. Especially where the strains are small towards the centre of the disc, these lie significantly below the noise threshold of the DIC calculation. A larger cell size would lower the noise threshold at the expense of spatial resolution. To enable the strain concentrations at the loading points to be captured, a larger cell size is not desirable. The solution to use the LI amplifier to obtain an amplitude during a cyclic loading clearly improves the noise threshold. No difference can be discerned between the results from the loading at 0.75 and 7.1 Hz, demonstrating that sub-Nyquist sampling is able to provide accurate results. A discrepancy between experiment and theory can be seen close to the loading points. The theory assumes a perfect point load which has infinite strain; in the experiment, a small plastic deformation leads to spreading of the contact point to a contact surface. This results in a finite strain at the load introduction points.

A quantitative comparison between the theoretical and experimentally obtained strains shows that the strains obtained using the LI amplifier lie within 5 % of the theoretical predictions at over more than 65 % of the disc area. Large departures from the theory exist near the edges of the disc where the strains are close to zero, and hence a small difference in the strain value gives a large percentage error, and near the loading points where the theoretical strain tends towards infinity. By contrast, the standard approach lies within 5 % of the theory over less that 50 % of the area. This is due to the much higher noise threshold associated with a single correlation.

47.4   Conclusions

It is demonstrated that by using knowledge of the temporal variation of the strain field, it is possible to use fairly simple signal processing ideas to improve the resolution of DIC strain calculations. The result is a twofold improvement of the data; the base strain resolution is improved, thereby enabling smaller strains to be resolved with greater accuracy. And this is achieved without needing to compromise on spatial resolution, thereby enabling finer resolution of regions with high strain concentrations. The example of the Brazilian disc has been used in this work to demonstrate the potential of the approach. However, these improvements in both strain and spatial resolution also have benefits for a wide range of practical applications, when trying to identify small localised defects, for example: in the field of non-destructive inspection, or the initiation and development of damage, for example: crack formation and growth in a fatigue test. Furthermore, the approach is easily implemented using standard digital cameras, thereby rendering it cost efficient and readily accessible to a wide spectrum of potential users.

References

1. Sutton MA, Orteu J-J, Schreier H (2009) Image correlation for shape, motion and deformation measurements. Springer, New York
2. Dulieu-Barton JM, Stanley P (1998) Development and applications of thermoelastic stress analysis. J Strain Anal 33:93–104
3. Temple PA (1975) An introduction to phase-sensitive amplifiers: an inexpensive student instrument. Am J Phys 43:801–807
4. Scofield JH (1994) Frequency-domain description of a lock-in amplifier. Am J Phys 63:129–133
5. Shannon CE (1949) Communication in the presence of noise. Proc IRE 37:10–21
6. Mishali M, Eldar YC (2011) Sub-Nyquist sampling: bridging theory and practice. IEEE Signal Process Mag (November):98–124
7. Unser M (2000) Sampling – 50 years after Shannon. Proc IEEE 88:569–587
8. Muskhelishvili NI (1977) Some basic problems of the mathematical theory of elasticity, 2nd edn. Noordhoff International, Groningen

Chapter 48
Tensile Response and the Associated Post: Yield Heating of Polycarbonate

C. Allan Gunnarsson, Bryan Love, Paul Moy, and Tusit Weerasooriya

Abstract It is well known that amorphous polymers, such as polycarbonate (PC), will exhibit adiabatic heating due to the large plastic work that occurs when undergoing significant plastic deformation. However the extent of adiabatic heating has not been investigated with respect to strain rate, with full-field temperature measurements performed on the specimen during deformation. In this study, ASTM tensile dog bone PC specimens were used to investigate the rate-dependent mechanical response from quasi-static to intermediate (~ 10 s^{-1}) strain rates using a traditional servo-hydraulic load frame. To determine the variations in yield and post-yield response at different locations of the gage area of the specimen, digital image correlation was used to measure the full-field surface strains. In addition, an InSb thermal camera was used concurrently to measure the full-field temperature distribution in the gage area during the deformation. The material experienced non-uniform heating as high as 50–70 °C, and showed significant rate sensitive mechanical response. In this paper, the experimental techniques and results are discussed.

Keywords DIC • Thermal response • Polycarbonate • Strain rate effect • Post yield behavior

48.1 Introduction

Polycarbonate (PC) is a thermoplastic polymer with a density of approximately 1.2 g/cm^3 that can be easily molded and thermoformed. This is due in part to the low glass transition temperature (T_G) of 150 °C and melting point of about 267 °C [1–3]. The mechanical properties of polymers are dependent upon two key factors: the rate of deformation and the material temperature. Polymers typically exhibit an increase in the yield strength and modulus, and a decrease in strain to failure as the strain rate increases from quasi-static ($\sim 10^{-3}$/s) to dynamic (10^3/s) [4, 5]. In this study, the tensile behavior of PC is studied as a function of strain rate between quasi-static (0.001/s) and intermediate ($\sim$5/s) strain rates. Digital image correlation (DIC) was used to measure localized plastic strain concentration in the necked region. A thermal (infrared) camera was used to measure the temperature change in the specimen at both strain rates. As the specimens yielded and plastically deformed, temperature in the specimen increases due to plastic work, with the largest temperature increase concentrated in the necked region of the specimen.

48.2 Experiments

ASTM D638.10 standard polymer tensile dog bone specimens were water-jet cut from PC sheet of 5.75 mm thickness and polished using wetted 600 grit sandpaper [6]. An Instron servo-hydraulic test frame was used to load the specimens in tension. Optical cameras were used to make DIC full-field strain measurements, while a thermal camera was used to make detailed temperature measurements of the specimen gauge area during deformation. The experiments were conducted at controlled displacement rates. These displacement rates caused the specimens to deform at a constant strain rate (when obtained as an average over the entire gage area).

C.A. Gunnarsson (✉) • B. Love • P. Moy • T. Weerasooriya
Army Research Laboratory, 2800 Powder Mill Rd, Adelphi, MD 20783, USA
e-mail: carey.a.gunnarsson.civ@mail.mil

48.3 Results

48.3.1 Strain Rate Effect on Thermal Behavior

Heat generation during plastic strain caused temperature increases to occur in the specimens for both strain rates. The thermal measurements show that significant heat generation begins when the localized strain (neck) begins to form for both rates. The maximum temperature change in the gage area as a function of time is shown in Fig. 48.1, along with the strain histories (discussed in the next section), for the (a) quasi-static and (b) intermediate strain rates. At the quasi-static rate, the temperature in the neck increases quickly, and then levels off, having risen 8 °C. The rate of temperature increase during the period of neck formation (from ~90 to 110 s in Fig. 48.1a) was approximately 0.25 °C/s. At the intermediate rate, the measured maximum change in temperature was approximately 28 °C, more than triple the temperature change for the quasi-static rate. The temperature increase never reached a constant value, continuing to increase even as the specimen reached failure; this is contrasted by the behavior observed at the quasi-static rate. The rate of temperature increase for the intermediate rate is not very linear; at the time of initial necking (0.37–0.38 s), the temperature is increasing at a rate of ~820 °C/s, while from 0.38 s to failure (0.405 s), the temperature increase rate decreases to ~375 °C/s.

48.3.2 Strain Rate Effect on Mechanical Behavior

The deformation strain rate has a significant effect on the mechanical response of PC. The mechanical (stress–strain) response for PC is shown in Fig. 48.2 for both quasi-static and intermediate strain rates. The yield stress increased from 66 MPa at the quasi-static strain rate to 82 MPa at the intermediate strain rate. The material response at the intermediate strain rate continued to be about 15–20 MPa higher for the post-yield behavior, as the material strain hardened and failed. The modulus remained the same at 1.9 GPa for both strain rates.

   The strain data shown in Fig. 48.2 were obtained using two different methods. The first method used DIC to extract the strain history during the tensile experiments averaged over the entire gage area; this method would be analogous to the deformation measurement technique (extensometer) specified in the ASTM standard. The second method of strain measurement used DIC to extract the maximum strain value regardless of location; for most of the experiments, this maximum strain was at the neck. Using the maximum strain value takes into account the strain concentration that occurs at the neck and gives a better perspective of the post-yield behavior of the material. The major difference between the two methods in the material response is what occurs during the softening region just after yield. The stress–strain response using the maximum strain method shows a reduced amount of softening when compared to the material response using the averaged gage area strain.
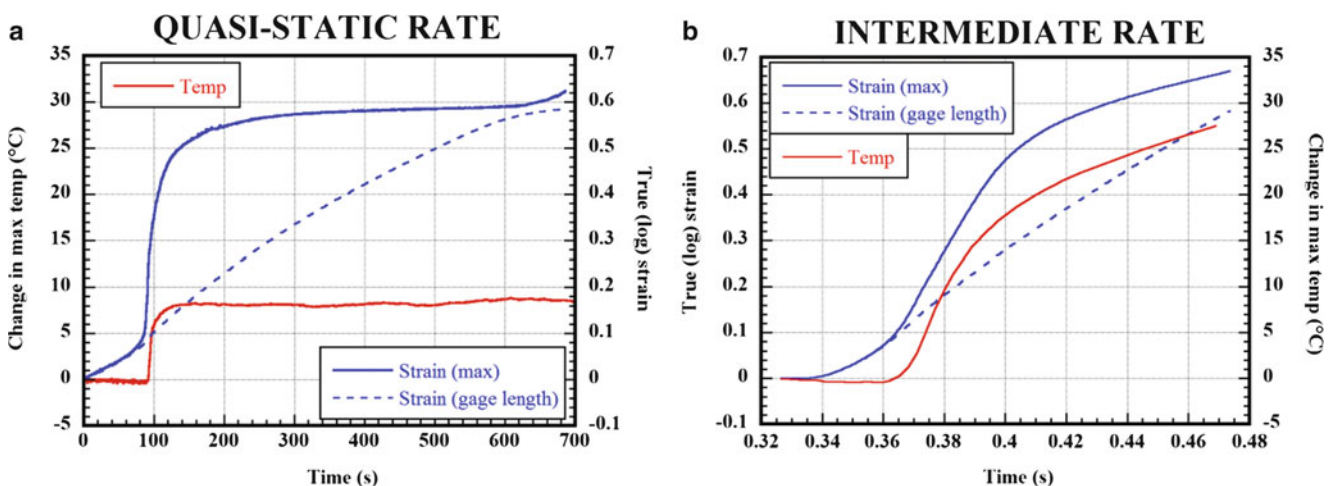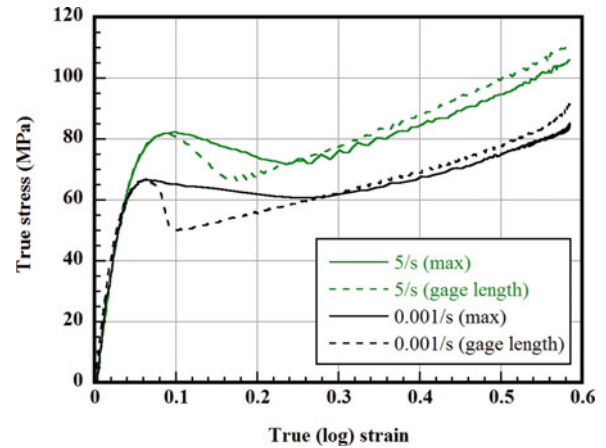

Fig. 48.1 Max temperature change and strain histories for PC at (**a**) quasi-static and (**b**) intermediate rates

Fig. 48.2 Comparison of mechanical response at quasi-static and intermediate strain rate for PC

The difference in the amount of softening between the two methods decreased as the strain rate increased from quasi-static to intermediate rate. The softening behavior is attributed to increased material temperature caused by heat generated during plastic work; the intermediate rate experienced greater temperature increase, and therefore more softening. Using the maximum strain data, the maximum softening from yield at the quasi-static strain rate was 5.5 MPa or 8.3 %, while at the intermediate strain rate it was 10.7 MPa, or 13.4 %.

The strain rate measured for these experiments was determined using the averaged gage length, as they were relatively constant and linear (Fig. 48.1). At the quasi-static strain rate, the averaged strain rate was 0.001/s, while the maximum strain rate was much higher ($\sim 10^{-2}$/s) initially during neck formation (60–120 s), and then dropped to about 10^{-4}/s (>120 s).

At the intermediate strain rate, the averaged strain rate was 5/s. The maximum strain rate was less varied; it reached 11/s during neck formation (0.36–0.40 s) and then fell to 2.4/s after the neck was formed (>0.40s). Using the averaged strain rate allows these results (at these strain rates) to be compared with other research performed according to the standard, although it is important to consider the non-uniformity of the strain distribution and strain rate.

48.4 Conclusions

A technique was developed to enable measurement of the mechanical and thermal response of polymers, and specifically PC, at quasi-static to intermediate strain rates using standard polymer test specimen geometry. PC is highly rate sensitive, with significant changes in response to increasing in strain rate. The yield stress and tensile strength both increased as strain rate increased, while the modulus remained constant. At both strain rates, there was a significant amount of heat generated during plastic deformation causing the material temperature to increase; this heating contributes to the softening of the material's response after yield. Increased strain rate also increased the amount of heat generated. The temperature increase in the specimen at the neck increased by over a factor of three as the strain rate increased from quasi-static (8 °C) to intermediate (28 °C).

References

1. Myers FS, Brittain JO (1973) Mechanical relaxation in polycarbonate-polysulfone blends. J Appl Polym Sci 17:2715–2724
2. Petersen RJ, Corneliussen RD, Rozelle LT (1969) Polymer reprint 10:385
3. IUPAC (1997) The IUPAC compendium of chemical terminology, 66th edn, p 583
4. Lo YC, Halldin GW (1984) The effect of strain rate and degree of crystallinity on the solid-phase flow behavior of thermoplastic. In: ANTEC'84, pp 488–491
5. Kaufman HS (1977) Introduction to polymer science and technology. Wiley, New York
6. Walter T, Love B, Moy P, Weerasooriya T (2014) High resolution digital image correlation of necking phenomena in polycarbonate at intermediate strain rates. In: Proceedings of the society for experimental mechanics annual conference and expo. Greenville, SC, 2–5 June, 2014

Chapter 49
Passive 3D Face Reconstruction with 3D Digital Image Correlation

Hien Kieu, Zhaoyang Wang, Minh Le, and Hieu Nguyen

Abstract Digital Image Correlation (DIC) is a widely used technique in the optics and mechanics fields for shape and deformation measurements. However, it is normally not used to measure the 3D shapes of human faces because of the requirement of speckle-pattern fabrication. In this paper, we demonstrate that the 3D DIC can be employed to reconstruct 3D images of human faces without using active speckle patterns either projected or fabricated on the faces. The technique involves a combination of the 3D DIC algorithm and the scale-invariant feature transform (SIFT) algorithm.

Keywords Digital image correlation • Passive face reconstruction • Subset size • Calculation step size • Step size

49.1   Introduction

3D human face reconstruction is a very interesting topic in computer vision and has been attracting many researchers because of its wide range of applications such as face recognition, animation for movies and video games, face surgery in medicine, and so on. There have been a number of techniques for reconstructing human face but most of them either require complicated and special setup or yield low accuracy result [1–3]. In order to obtain a high accuracy 3D image, it requires the subject to be painted or projected with patterns or markers which make obtaining dense and accurate point matches possible [4–8]. In addition, lots of high resolution cameras have been used to obtain high-density correspondences as well as to avoid occlusions [2, 4]. Intensive and uniform light is very important and has been used to obtain good result in other previous techniques [2]. These issues would not only increase the cost of the system but also make the subject feel uncomfortable. Exposure and running time are also issues from previous techniques that need to be addressed. In our presented techniques, there is no need for template facial geometry, pattern paint or projection, or intensive and uniform light. The system contains only two cameras in normal room-light condition and the capture is one-time exposure. Our system is simple, easy to setup, low-cost, and comfortable for the captured subject and our algorithm is fast, accurate, and able to produce dense 3D points without any post processing.

3D Digital Image Correlation (3D DIC) has been known for its high accuracy in finding the displacement of dense 3D points on speckle-painted objects [9]. In this paper, 3D DIC is exploited to give a fine point matches on human face's surface in which pores and wrinkles are used as passive speckle patterns. The reason that the ability to act as pattern of pores has not gone popularly noticed is that they are difficult to see with medium resolution and are also easily affected by noise. In order to use pores for the image correlation, high resolution camera must be used and step size in the subset must be small. We figured out that a 2,048 $\times$ 1,536 high definition camera can capture the pores if it is tuned to best focus and barely capture the whole face. Because small subset size will be easily affected by noise, large subset size needs to be used. However, for a large subset size, the computation cost can be very high if the calculation step size inside the subset is small whereas the detail of the face reconstruction will be missing if the calculation step size in the subset is large. Thus, we have introduced a multi-subset concept in which we are not using one single subset size but multiple subset sizes which can help find the

H. Kieu (✉) • M. Le • H. Nguyen
Department of Electrical Engineering, The Catholic University of America, 620 Michigan Ave. N.E, Washington, DC 20064, USA
e-mail: 81kieu@cardinalmail.cua.edu

Z. Wang
Department of Mechanical Engineering, The Catholic University of America, 620 Michigan Ave. N.E, Washington, DC 20064, USA

disparity map of the face with details and in short computational time. Together with our already existed high-accuracy camera calibration algorithm [10], the mentioned point matches obtained from 3D DIC give us the reconstructed face with high accuracy in a matter of minutes. We also employ Scale-Invariant Feature Transform (SIFT) for an automatic finding of the initial guess for the DIC algorithm [11].

49.2 Algorithm

Scale-Invariant Feature Transform is first applied to the pair of captured stereo images. The reference image is continuously convolved with Gaussian window and down sampled by 2 to give a set of blurred images at different scales. These images will constitute the Gaussian pyramid [12]. By subtracting each image in the Gaussian pyramid by its adjacent image, the Difference of Gaussian (DoG) pyramid is constructed. Any pixel in the DoG pyramid which is larger or smaller than its 26 neighbor pixels will be considered local extremum. These pixels are feature points for matching. Based on the pixel intensities around each feature point, a descriptor is then constructed in a form of a 128-dimensional vector. SIFT will also be applied to extract feature points with their unique descriptors in the target image. Then the FLANN algorithm is used to detect the matching of feature points between two images according to the Euclidian distance between SIFT descriptors [13]. SIFT is very robust and can cope well with the scaling, translation, rotation, distortion (affine transformation), occlusion and illumination-variation issues encountered in real images. An example of SIFT matching is shown in Fig. 49.1 where SIFT detects and matches feature points on the reference (left) and the target (right) images.

The pixel that is closest to the centroid of any three SIFT points can be the seed point for the DIC analysis to begin. DIC analysis for a pixel in the target image is calculated based on the intensity information in a region $(2M + 1) \times (2M + 1)$ around it. DIC algorithm estimates the displacement of points in that region by a transformation equation that transforms 2D coordinate of every point in that region of the reference image to the corresponding coordinate of the same physical point in the target image [14]:

$$\begin{cases} x_i' = x_i + \xi + \xi_x\Delta_x + \xi + \xi_y\Delta_y + \xi_{xx}\Delta_x^2 + \xi_{yy}\Delta_y^2 + \xi_{xy}\Delta_x\Delta_y \\ y_i' = y_i + \eta + \eta_x\Delta_x + \eta + \eta_y\Delta_y + \eta_{xx}\Delta_x^2 + \eta_{yy}\Delta_y^2 + \eta_{xy}\Delta_x\Delta_y \end{cases} \tag{49.1}$$

where i and j range from $-M$ to M, $\Delta_x = x_i - x_0$, $\Delta_y = y_i - y_0$, and $\xi, \xi_x, \xi_y, \xi_{xx}, \xi_{yy}, \xi_{xy}, \eta, \eta_x, \eta_y, \eta_{xx}, \eta_{yy}$, and η_{xy} are the transformation parameters. The determination of these parameters can be achieved by minimizing the least-squares-based correlation coefficient defined as:

$$C = \sum_{i=-M}^{M} \sum_{j=-M}^{M} \left[af(x_i, y_i) + b - g\left(x_i', y_i'\right) \right]^2 \tag{49.2}$$

where a is a scale factor, and b is the offset of the intensity, and $f(x_i, y_i)$, $g(x_i', y_i')$ indicate the intensity values at a representative pixel in the reference subset and the potential matching pixel in the target subset, respectively. There are other
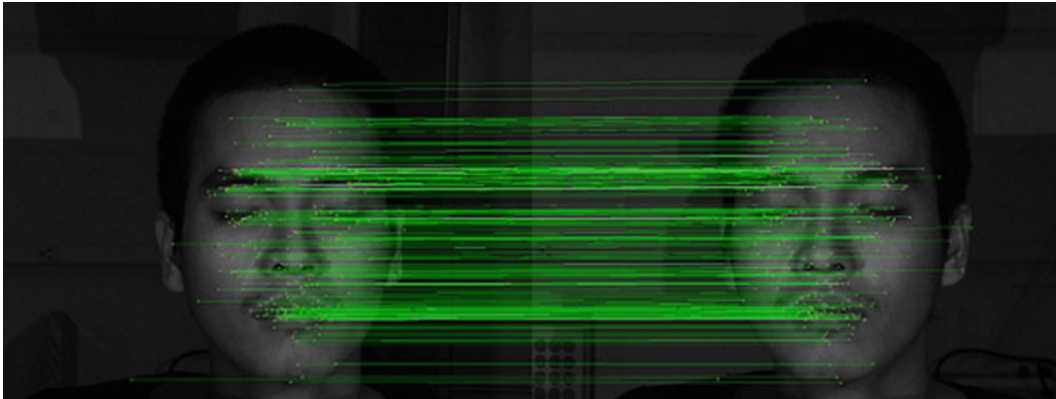

Fig. 49.1 SIFT matching result on an experiment with human face

correlation criteria but in this specific technique, we only use the correlation criterion mentioned above [15]. The solution of ξ, ξ_x, ξ_y, ξ_{xx}, ξ_{yy}, ξ_{xy}, η, η_x, η_y, η_{xx}, η_{yy}, a, and b, for the minimum value of correlation coefficient C is found by an iterative method called Levenberg-Madquardt [16]. Most of the times, DIC analysis will not interrogate every interested pixel. Instead, pixels on a grid with a constant step size are selected. Then an interpolation scheme is carried out to find matching for all the remaining pixels [17]. This can save much of the computational cost.

From the seed point, the parameter solution of the four neighbor pixels at left, right, top, bottom with a step size of N are obtained using the seed point's parameter as their initial guess. The smaller N is, the finer the result will be, and the computation time will be also longer. Normally, we set N to 3 or 5. After finding the parameter solution by the Levenberg-Madquardt algorithm, their correlation coefficients are calculated. The information of those four pixels is stored in an array. After that, the correlation coefficients and parameter solutions of the four neighbor pixels of the one in the array that has the smallest correlation coefficient are calculated. If any pixel among four mentioned pixels was calculated, it will not be calculated again. These additional pixels are again put into the array. The process continues by propagating in such way until all the points are calculated. A processed point will not be propagated if its analysis fails the convergence or disparity limitation criteria, or its correlation coefficient is above the pre-defined threshold, or all its neighbors have already been processed. By doing this, the analysis will stop as it reaches any bad region and prevent the analysis from producing incorrect results [18].

DIC analysis will first operate will M = 61 and with a calculation step size of m = 4. This means that not all the pixels inside the subset $(2M + 1) \times (2M + 1)$ will be used to calculate the correlation coefficient C. Instead, only pixels on a square grid with a distance between two eyes of the grid m = 4 are used to calculate the correlation coefficient of the pixel around which the subset is drawn. After the DIC analysis with M = 61, and m = 4 is finished, the algorithm is repeated with smaller M and m. In our options, we set M = 61, 41, 21 and m = 4, 3, 2. When the same pixel is analyzed again in later analysis with different subset sizes, the new parameter solution will be updated if the correlation coefficient obtained is smaller than the correlation coefficient generated by the current parameter solution. By doing this way, the final result is always the best result among all results obtained from different subsets.

At the beginning of the process, large subset and calculation step size are used to get the rough result of the reconstruction face because large subset is more invariant to noise and thus gives higher fidelity in matching. We have to set calculation step size large to save a large amount of computation time although by doing that, we lose information of the pixels that are not calculated and thus the result rendered are not detailed. However, the result can be refined by repeating DIC analysis with smaller subset sizes and calculation step sizes [19]. Despite the fact that small subsets are easily affected by noise and are susceptible to errors that are propagated during the analysis, the results are not misleading like when a single small subset size is chosen because the small subset analysis is based on roughly correct results obtained from the previous analysis with larger subset. This explains how our multi-subset concept works. The number of subset layers, subset sizes, calculation step size for each subset can be adjusted by the users according to the image quality. Note that in this paper, we use three layers of subset sizes of 21, 41, and 61 with different calculation step sizes of 2, 3, and 4, respectively. By sufficient experiments, our proposed algorithm has proved to be robust, fast, and accurate [20].

49.3 Results

Figure 49.2 shows some of our various experiments to render 3D reconstruction from stereo 2D images without selecting initial guess and boundary for DIC. Note that the multi-subset process has helped the algorithm to successfully reconstruct the human face without paints, markers or any pattern projection. The process was carried out in a fully automatic manner with no need for selecting initial guess for the DIC analysis. The DIC analysis for this experiment is operated with subset sizes of 21, 41, 61, with the calculation step size of 2, 3, 4, respectively. The grid step size is 5 and interpolation is carried out afterwards to get full-field result. With our existing flexible camera calibration technique, 3D result is obtained from full-field matching with high accuracy [21].

49.4 Conclusion

This paper presents a simple system for 3D face reconstruction. Only two cameras are used for the stereo matching and with camera information, 3D image of the face is reconstructed. Thanks to the employment Digital Image Correlation with multi-subset sizes, the proposed algorithm can find matches based on the natural patterns on human's face which are very small
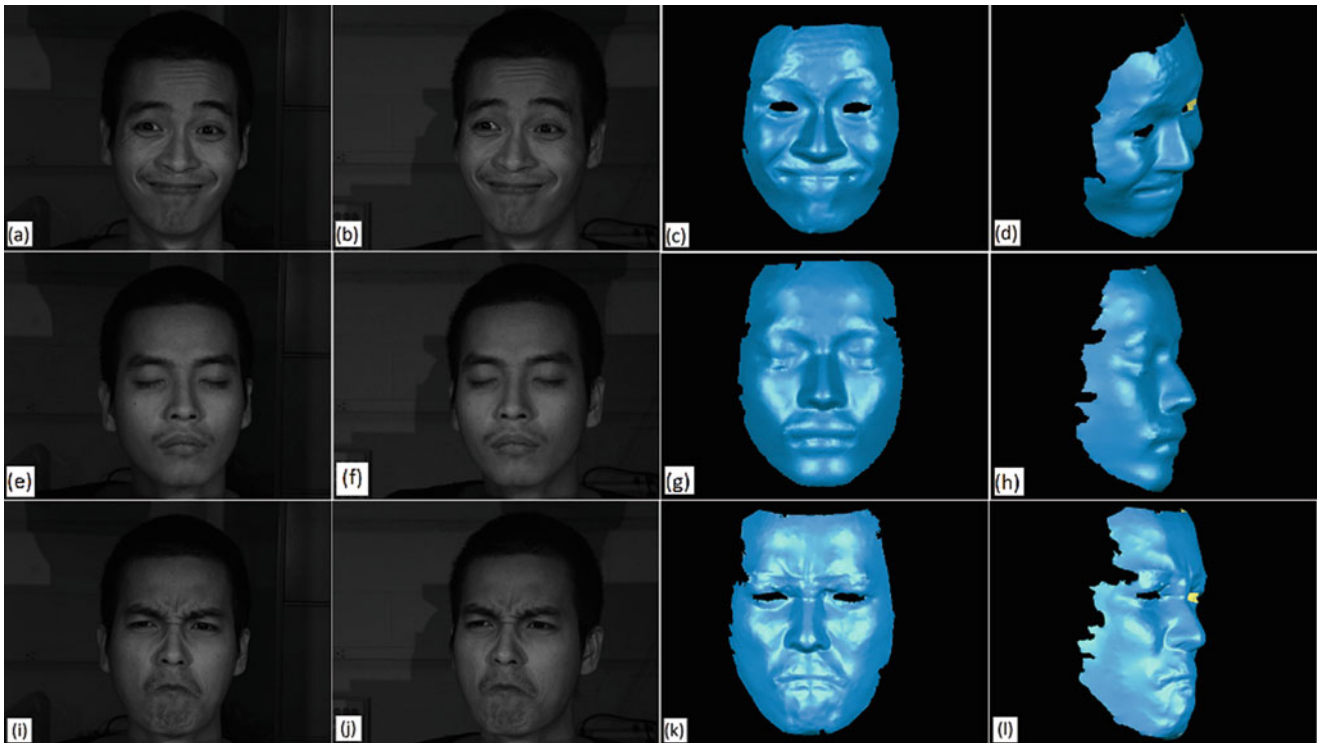
Fig. 49.2 Demonstration of 3D view result for multiple separate objects and human face, the first two columns show images captured from two cameras and the latter two columns show two views of the 3D reconstruction

and easily affected by noise. SIFT is also applied to automatically find the initial guess for DIC, which helps skip the manual guess selection part which is sometimes tedious and can be incorrect. In addition, failure in DIC analysis to propagate matching points does not stop the analysis. Instead, the proposed scheme continues the DIC analysis from other SIFT points which can act as starting points. This has relatively enhanced the reliability of the DIC analysis on complex regions which do not have clear speckle patterns. In summary, thanks to all the modification in the traditional DIC algorithm, we are able render a 3D image of human's face using a simple system with only two stereo cameras and without any additional special setup.

As technology evolves to make humans' lives more comfortable, any process that serves humans must be simple, fast, and should not be irritating. Therefore, a passive 3D face reconstruction system in a short time, with high accuracy is demanded. Because of the system's low cost and simple setup, we hope that it can be further developed to become a commercial 3D face reconstruction system.

References

1. Garcia E, Dugelay J-L (2001) Low cost 3D face acquisition and modeling. In: International conference on information technology: coding and computing, Las Vegas, NV, pp 657–661
2. Bradley D, Heidrich W, Popa T, Sheffer A (2010) High resolution passive facial performance capture. In: ACM transaction on graphics (Proceedings of SIGGRAPH). ACM, New York, p 41
3. Hossain MS, Akbar M, Starkey JD (2013) Inexpensive construction of a 3D face model from stereo images. In: The third international conference on communications and information technology (ICCIT)
4. Zhang L, Snavely N, Curless B, Seitz SM (2004) Spacetimefaces: high resolution capture for modeling and animation. In: ACM transaction on graphics (Proceedings of SIGGRAPH), pp 548–558
5. Wang Y, Huang X, Lee C-S, Zhang S, Li Z, Samaras D, Metaxas D, Elgammal A, Huang P (2004) High resolution acquisition, learning and transfer of dynamic 3-D facial expressions. In: Computer graphics forum, pp 677–686
6. Furukawa Y, Ponce J (2009) Dense 3D motion capture for human faces. In: 2009 IEEE conference on computer vision and pattern recognition
7. Bickel B, Botsch M, Angst R, Matusik W, Otaduy M, Pfister H, Gross M (2007) Multi-scale capture of facial geometry and motion. In: ACM transaction on graphics (Proceedings of SIGGRAPH), New York, p 33

8. Ma W-C, Jones A, Chiang J-Y, Hawkins T, Frederiksen S, Peers P, Vukovic M, Ouhyoung M, Debevec B (2008) Facial performance synthesis using deformation-driven polynomial displacement maps. ACM Trans Graph (Proc SIGGRAPH Asia) 27(5):121

9. Pan B, Xie H, Yang L, Wang Z (2009) Accurate measurement of satellite antenna surface using 3D digital image correlation technique. Strain 45(2):194–200

10. Vo M, Wang Z, Luu L, Ma J (2011) Advanced geometric camera calibration for machine vision. Opt Eng 50:110503

11. Wang Z, Vo M, Kieu H, Pan T (2014) Automated fast initial guess in digital image correlation. Strain 50(1):28–36

12. Lowe D (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110

13. Muja M, Lowe DG (2009) Fast approximate nearest neighbors with automatic algorithm configuration. In: VISAPP international conference on computer vision theory and applications

14. Pan B, Qian K, Xie H, Asundi A (2009) Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. Meas Sci Technol 20:062001

15. Pan B, Xie H, Wang Z (2010) Equivalence of digital image correlation criteria for pattern matching. Appl Opt 49:5501–5509

16. Lourakis MIA, Argyros AA (2005) Is Levenberg-Marquardt the most efficient optimization algorithm for implementing bundle adjustment? In: tenth IEEE international conference on computer vision (ICCV), p 2

17. Luu L, Wang Z, Vo M, Hoang T, Ma J (2010) Accuracy enhancement of digital image correlation with B-spline interpolation. Opt Lett 36(16):3070–3072

18. Kieu H, Pan T, Wang Z, Le M, Nguyen H, Vo M (2014) Accurate 3D shape measurement of multiple separate objects with stereo vision. Meas Sci Technol 25(3), 035401

19. Wang Z, Hoang T, Nguyen D, Urcinas A, Magro J (2010) High-speed digital image correlation method: comment. Opt Lett 35(17):2891

20. Hu Z, Xie H, Lu J, Wang H, Zhu J (2011) Error evaluation technique for three-dimensional digital image correlation. Appl Opt 50:6239–6247

21. Vo M, Wang Z, Pan B, Pan T (2012) Hyper-accurate flexible calibration technique for fringe-projection-based three-dimensional imaging. Opt Express 20(15):16926–16941

Chapter 50
On the Meso-Macro Scale Deformation of Low Carbon Steel

Suraj Ravindran, Behrad Koohbor, and Addis Kidane

Abstract The multiscale deformation response of low carbon steel is investigated. The meso and macro scale displacement and strain fields for specimen subjected to pure tension are measured using in-situ multiscale digital image correlation technique. The specimen is specked with different scale pattern ranging from 5 to 500 μm size. The smallest scale, 5 μm, speckles are used for local meso-scale deformation measurement. In this case an optical microscope is used to record the local information within 1 mm square field of view. On the other hand, the larger size, 500 μm, speckles are used to measure the continuum level deformation. In this case two digital cameras with 5 megapixel resolution are used in 3D arrangement by considering the entire width of the specimen inside the field of view. Both the optical microscope and the digital camera systems are triggered simultaneously to acquire the deformation at the same time scale. The displacement and strain fields are extracted using digital image correlation. The effect of local deformation on the overall displacement and strain of low carbon steel is presented by comparing with the macro scale deformation and strain fields. Furthermore, microstructure images are obtained by optical microscope and used for the analysis of local strain field coupled with the strain field from digital image correlation.

Keywords Multiscale experiments • Digital image correlation • RVE • Subset • Low carbon steel

50.1 Introduction

The deformation behavior of polycrystalline materials are highly influenced by the microstructure of the material. Many researchers have experimentally measured inhomogeneous polycrystalline deformation at length scales ranging from 1 mm to 10 μm. Some of the experimental investigation performed in late 1990s in conjunction with the electron back scattered diffraction revealed the strong influence of microstructural properties on deformation behavior and development of plastic heterogeneities in crystalline materials [1, 2]. Therefore, it is important to study the influence of microstructural properties (size, orientation etc.) in deformation behavior of the polycrystalline materials in order to achieve appropriate micro scale model of the materials. There are different computational homogenization schemes available to model the deformation behavior in the polycrystal at microscale [3]. However, it is difficult to obtain the size of representative volume element (RVE) to incorporate the deformation behavior of material at plastic strain. In this paper, we have attempted to obtain the size of the RVE through meso scale experiments.

Two important methods used for obtaining the displacement field at smaller scales are micro grid method and digital image correlation. Though micro grid method can produce accurate results for large deformation, it is not accurate enough to capture small strains. Digital image correlation is an optical based experimental technique developed late 1980s [4]. Steps involved in this method are decorating the specimen with black speckles in white background. Then obtain image before and after loading, images are cross correlated using correlation functions to obtain the displacement field. There are numerous studies has been carried out using digital image correlation to obtain the strain field at micro scale to nano scale [5, 6]. Majority of the work was performed using scanning electron microscope or optical microscope in conjunction with the digital image correlation technique. Most of the work that has been done is in one scale in other words considering specific

S. Ravindran (✉) • B. Koohbor • A. Kidane
Department of Mechanical Engineering, University of South Carolina, 300 Main Street, Columbia, SC 29208, USA
e-mail: muthiram@email.sc.edu

number of grains in an experiment, which will provide the deformation behavior at one scale in connection with the microstructure in polycrystal. There is limited number of experimental studies available for obtaining the RVE size which can represent the macro scale behavior of material incorporating the microscale deformation field. In this study we have used different element sizes to find the RVE size of the low carbon steel.

50.2 Experimental Procedure

50.2.1 Specimen Preparation and Experimental Procedure

The material used in this study is low carbon steel. Material was obtained in sheet form with nominal thickness of 1 mm. Dog bone specimens were cut with dimensions $16 \times 4 \times 1$ mm as shown in Fig. 50.1. Ratio of the gage length to width of the specimen was maintained to be 4 to obtain the actual material constitutive behavior under tension [7]. First the specimens are full annealed in an oven to remove the residual stress in the specimens from cutting. After that, each specimen is mechanically grinded using grit size ranging from 240 to 1,200. Then it is polished using 5 μm, and 3 μm alumina powder for 20 min and colloidal silica slurry of 0.05 μm for 1 h. Etching is done using Nital (90 ml Methanol + 8 ml HNO_3) in order to take the images of grain structure before and after deformation. The average grain size of the specimen is 20 μm.

An inexpensive method of speckling was developed to make the speckle pattern for the micro scale deformation measurement. Specimen was initially painted with plain white paint, before white paint get dried black toner powder is blown on to the specimen at high pressure. After drying it is seen that the black toner powder particles adhered to the white paint firmly. Speckling pattern at 6× magnification is shown in Fig. 50.2. The grey scale intensity of the pattern follows a bell shaped distribution Fig. 50.2, which is suitable for the digital image correlation. Size of the black dots in the speckle pattern varies from 3 to 7 μm. For aligning the strain field with the grain structures, the area of consideration is required to be marked before the experiment. Four 20 μm micro indentation marks are used to capture rectangular area of interest.

Fig. 50.1 Dimensions of the dog bone tensile specimen used in this work (Dimensions in mm)
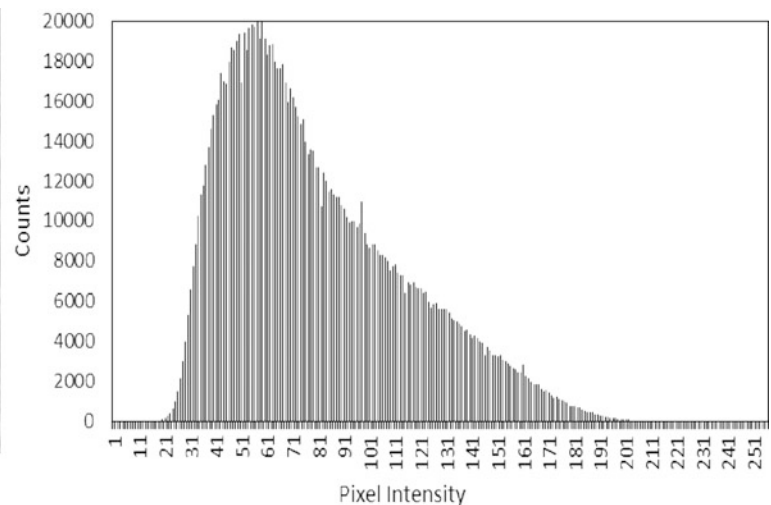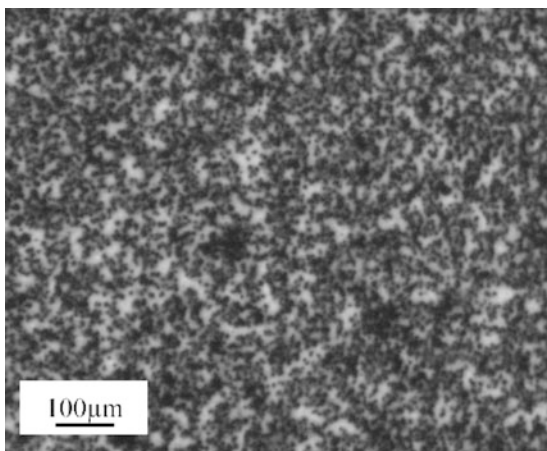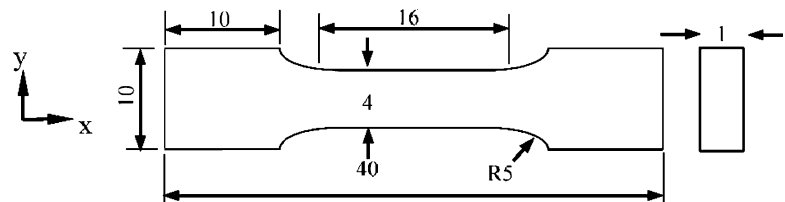

Fig. 50.2 Speckling pattern and the grey scale distribution of the speckles

Fig. 50.3 Small tensile machine used for testing and stress strain curve obtained
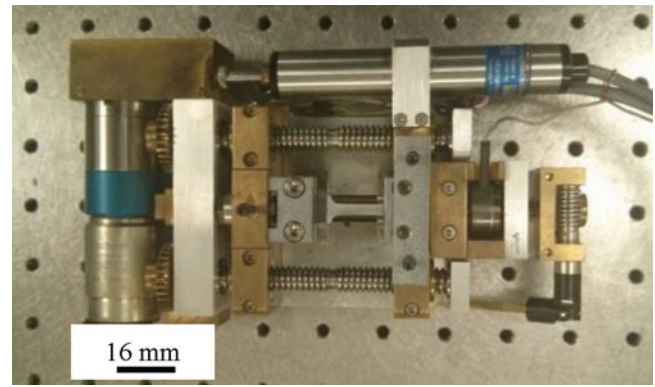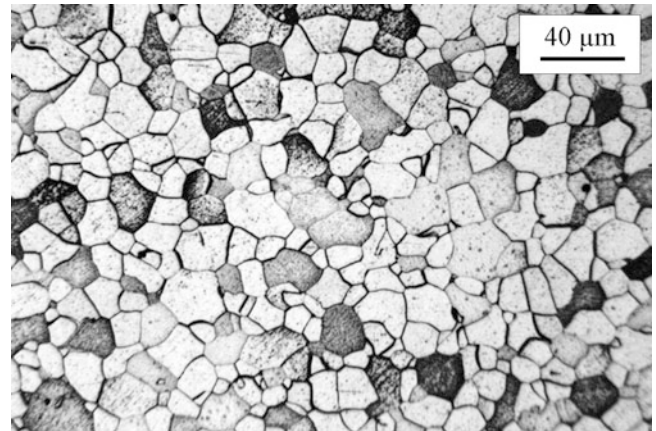

Fig. 50.4 Optical micrograph of the low carbon steel examined in this work (Average size of grain 20 μm)

The specimen is then uni-axially deformed at a loading rate of 1.7 mmS^{-1} in X-direction using micro-tensile testing machine shown in Fig. 50.3. This small tensile stage testing machine has a load cell of 2,250 N. Microscope was arranged in such a way that, the specimen is facing the microscope and the four marked position are inside the field of view. Experiment was conducted on vibration isolated table to extract images without the effect of floor vibration. Images are taken during loading using a camera from QI digital imaging with a resolution of 1,280 × 1,024 pixels. During loading images are taken at the rate of 1 images/s. Both, camera and tensile frame are triggered at the same time so that both the load data and the image acquired during deformation of the specimen will match.

50.3 Result and Discussion

The micrograph showing the microstructure of the examined low carbon steel specimen is shown in Fig. 50.4. A typical global stress–strain curve form the macro size experiment is shown in Fig. 50.5. The microscopic images taken at global strain of 1.4 and 3.4 % are analyzed using commercial digital image correlation software Vic-2D, by Correlated Solutions, Inc. It is seen that the subset size have effect on capturing strain field at smaller length scale. For a subset size of 12 and 16 μm, the strain field shows inhomogeous pattern. As shown in Fig. 50.6a and b, at a subset size of 12 and 16 μm the local strain varies from about 1.2 to 1.8 % for a corresponding global strain of 1.4 %. As the subset size increases, the strain starts to become homogeneous staring from a subset size of 31 μm, as shown in Fig. 50.6d–f. The entire field is homogenized to a strain between 1.4 and 1.5 %. The axial strain is converged to a constant value at the subset size of 41 μm as shown in Fig. 50.7. From this observation, the smallest length scale which can represent (RVE) the deformation behavior of the material at a global strain of 1.4 % is 41 μm (2 grains).
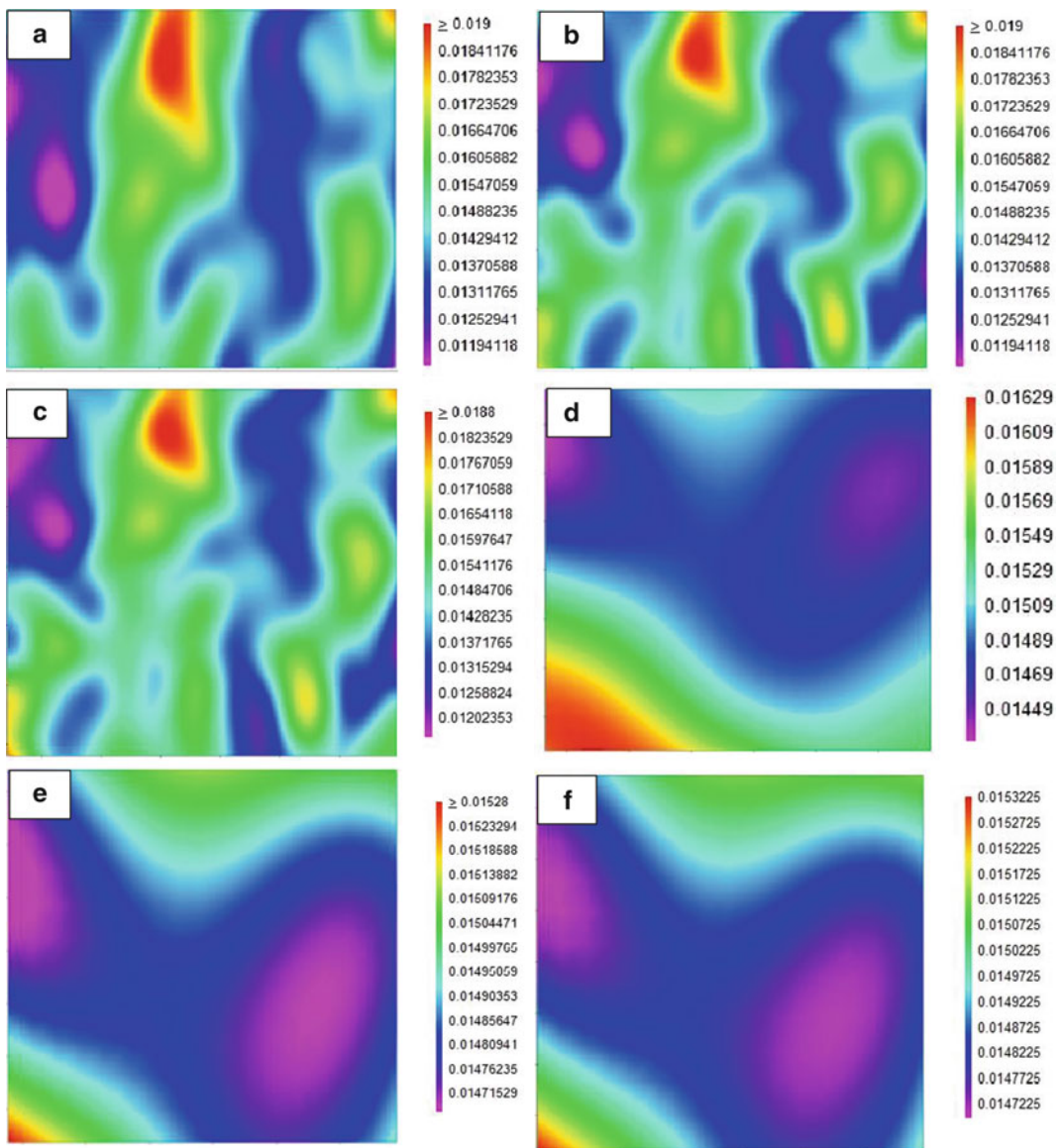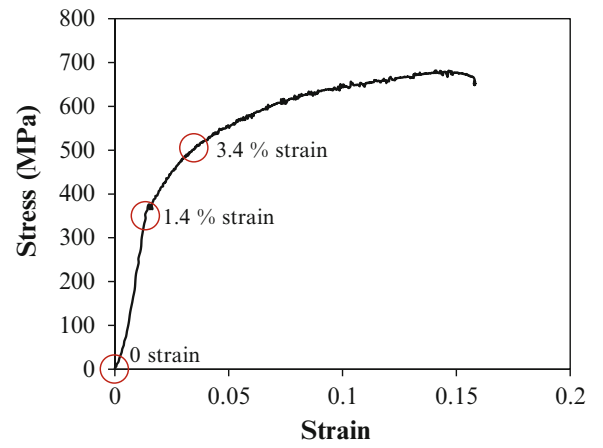
Fig. 50.5 Typical global
stress–strain curve

Fig. 50.6 Axial strain field for subset size of (**a**) 12 μm, (**b**) 16 μm, (**c**) 21 μm, (**d**) 31 μm, (**e**) 41 μm and (**f**) 51 μm, for global axial strain of 1.4 %

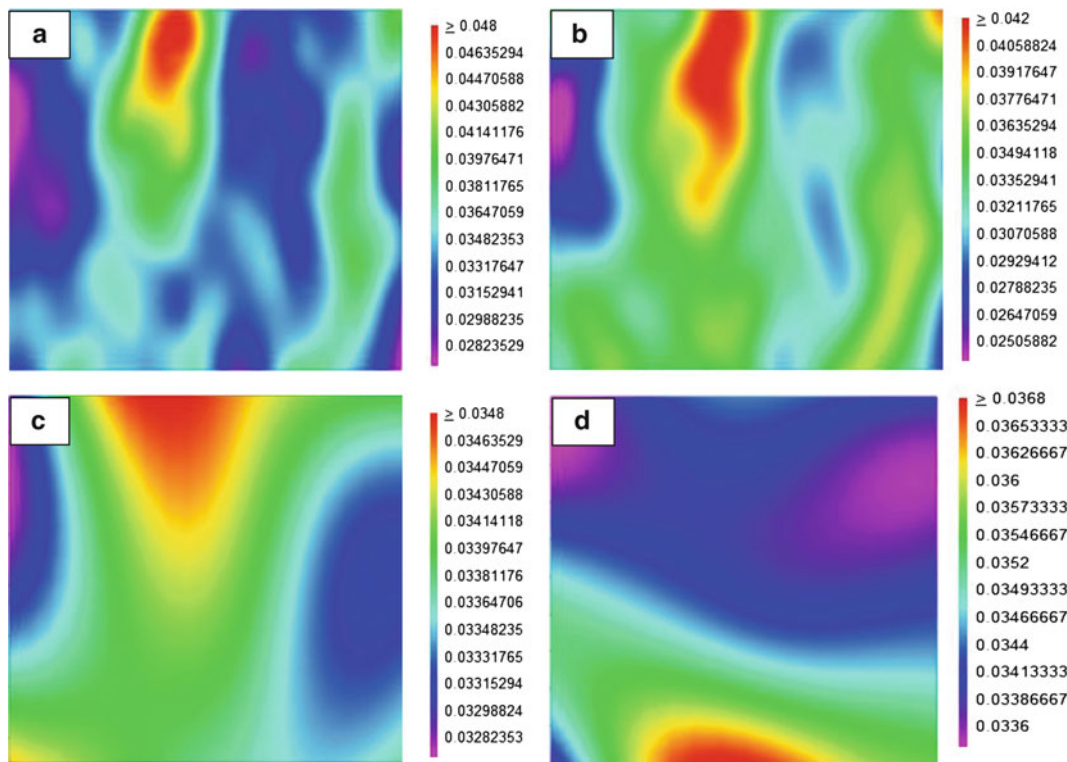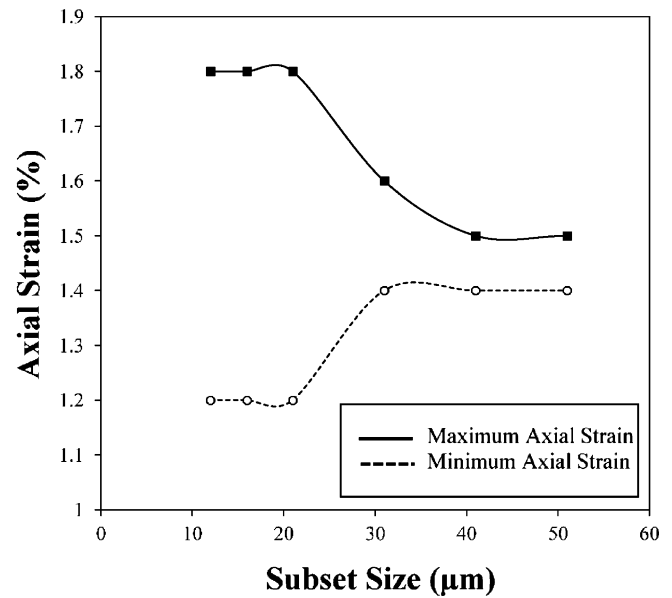**Fig. 50.7** Local strain as a function of subset size for the global strain of 1.4 %

Fig. 50.8 Axial strain field for subset size of (**a**) 12 μm, (**b**) 21 μm, (**c**) 31 μm and (**d**) 51 μm, for global axial

For a global strain of 3.4 % obtained from macroscopic experiment, that the local strain field is inhomogeneous at a subset size of 12–21 μm, as it was observed in the case of 1.4 % global strain. As shown in Fig. 50.8a and b, the strain varies from 2.8 to 4.6 % for a corresponding global strain of 3.4 %. Again the strain field starts to homogenize at subset size of 31 μm and above. Looking at Fig. 50.9, one can easily notice that, the strain is fairly homogeneous at a subset size above 41 μm. The representative volume at a global strain of 3.4 % is around 41 μm. One can see a bit effect of strain on the RVE size,

Fig. 50.9 Local strain as
a function of subset size for
the global strain of 3.4 %
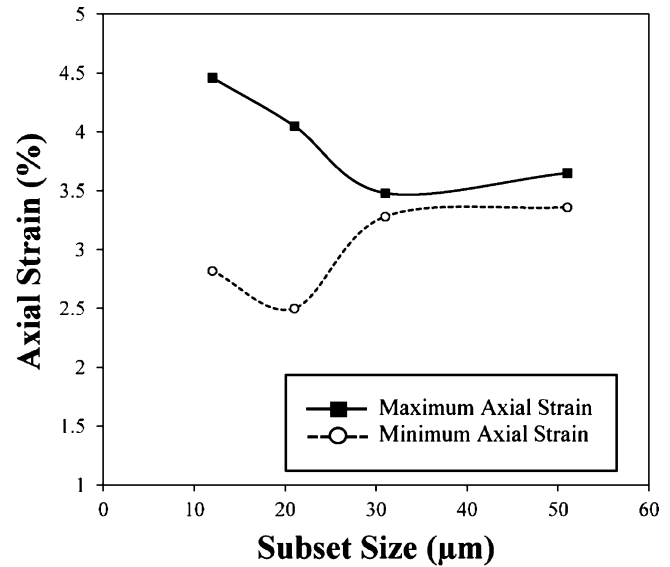
as the global strain get larger the RVE size become smaller and smaller. Further investigations are required to confirm the observation.

50.4 Conclusion

The representative elemental volume of low carbon steel is investigated using in-situ multiscale digital image correlation technique. A novel image speckling technique is developed. The effect of local deformation on the overall displacement and strain of low carbon steel is presented by comparing with the macro scale deformation and strain fields. It was found that the RVE size of the polycrystalline material fairly depends on the subset size and it is found to be around at least twice the size of the small grain.

References

1. Schtleber M, Zhao Z, Raabe D (2002) Experimental investigation of plastic grain interaction. Mater Sci Eng A 336:81–87
2. Hoc T, Crepin J, Gelebart L, Zaoui A (2003) A procedure for identifying the plastic behavior of single crystals from the local response of polycrystals. Acta Mater 51(18):5477–5488
3. Haudaigui FE, Forest S, Gourgues AF, Jeulin D (2006) Representative volume element sizes for copper bulk polycrystals and thin layers. In: Symposium 3M mat'eriaux, mechanical microcircuits, on the interfaces: from atom to polycrystal, CEA Saclay, INSTN, pp 141–153
4. Chu TC, Ranson WF, Sutton MA (1985) Applications of digital image correlation techniques to experimental mechanics. Exp Mech 25(3):232–244
5. Vendroux G, Schmidt N, Knauss WG (1996) Submicron deformation field measurements - part III. Demonstration of deformation determinations. Exp Mech 38(3):154–160
6. Berfield TA, Patel JK, Shimmin RG, Braun PV, Lambros J, Sottos NR (2007) Micro- and nanoscale deformation measurement of surface and internal planes via digital image correlation. Exp Mech 47(1):51–62
7. ASTM Standard E8 (2011) Standard test methods for tension testing of metallic materials. In: ASTM international, West Conshohocken, PA

Chapter 51
Feasibility of Non-Contacting Measurement of Wind-Induced Full-Field Displacements on Asphalt Shingles

Rahim Ghorbani, Xing Zhao, Fabio Matta, Michael A. Sutton, Addis Kidane, Zhuzhao Liu, Anne Cope, and Timothy Reinhold

Abstract Understanding of deformations and the progressive failure mechanisms of asphalt shingles under wind loads is key to develop wind-resistant roofing systems, as well as standard test methods to characterize strength under representative wind loads. In fact, failure of shingles rated as resistant to winds up to 150 mph have been reported at speeds below 115 mph. Damage associated with failure of roof shingles continues to be a major source of insurance claims. Though pressure measurements can be taken at discrete points using pressure taps, no technology has been successfully deployed to measure full-field deformations on roof shingles subjected to wind loads. This paper reports on a feasibility study of three-dimensional digital image correlation (3D-DIC) as a non-contacting technique to measure full-field displacements of roof shingles under high wind loads. Feasibility is assessed based on evidence from load testing of three-tab shingles mounted on a full-scale roof panel specimen that was subjected to straight winds with speed up to 155 mph. Uplift displacements were measured on a target shingle tab. The natural color variations on the shingle exposed surface were used to provide a suitable speckle pattern for 3D-DIC measurements. It is shown that consistent 3D-DIC uplift displacement maps can be obtained up to failure. The evidence gained also highlights the importance of understanding the influence of time-dependent shingle material deformations, together with the progressive physical damage along the sealant strip.

Keywords Asphalt shingle • Speckle pattern • Three-dimensional digital image correlation • Uplift displacement • Wind load

51.1 Introduction

Roof asphalt shingles are the most common type of sloped-roof cover for residential construction in the US. Modern shingles typically consist of light-weight fiber-reinforced laminates including a reinforcing mat, asphalt coating and surface granules [1]. Asphalt shingles provide a water-shedding barrier for the roof sheathing into which the shingles are nailed. A self-sealing strip bonding the top shingle near its exposed edge to the one below minimizes the penetration of water and provides resistance against wind-induced uplift. Self-sealing strips, which were introduced in the 1950s, consist of heat-sensitive adhesive strips that attach the leading edge of a given shingle to the shingle underneath, providing resistance to wind-induced uplift and allowing the use of lighter and more economical shingle mats [2].

Asphalt shingles are rated based on their resistance to wind loads according to their performance when tested following applicable standards, i.e., ASTM D3161 [3] and ASTM D7158 [4]. However, there is evidence from field observations that asphalt shingles rated for resistance against 150 mph winds (i.e., H-rating per ASTM D7158 [4]) performed poorly when

R. Ghorbani (✉) • F. Matta
Department of Civil and Environmental Engineering, University of South Carolina, 300 Main St., Columbia, SC, USA
e-mail: ghorbanr@email.sc.edu; fmatta@sc.edu

X. Zhao • M.A. Sutton • A. Kidane
Department of Mechanical Engineering, University of South Carolina, 300 Main St, Columbia, SC, USA
e-mail: zhao29@email.sc.edu; sutton@sc.edu; kidanea@cec.sc.edu

Z. Liu • A. Cope • T. Reinhold
Insurance Institute for Business and Home Safety Research Center 5335, Richburg Rd, Richburg, SC, USA
e-mail: zliu@ibhs.org; acope@ibhs.org; treinhold@ibhs.org

subjected to wind speeds with 115 mph or less 3-s gusts during Hurricane Ike [5]. Similar evidence was collected after Hurricane Andrew [6] and Hurricane Hugo [7], suggesting that the effects of real-world wind loads are not sufficiently captured through existing standard load tests. The discrepancy between rated and real-world performance is due in part to the fact that standard test methods [3, 4] represent idealized shingle simulations [8], and in part to the fact that the shingles being tested do not experience the effects of the supply and storage chain, for example by being weighted down under stacks of other shingles (resulting in a thinner sealant strip which may provide a less effective bond). The understanding of the correlation between uplift forces and wind loads, and the wind-induced failure mechanisms, is complicated by the difficulty to measure and map uplift deformations and pressures with good spatial resolution. In fact, the use of contacting distributed point-wise sensors such as strain gauges and pressure taps may affect the mass and stiffness of light-weight shingles, thereby altering the deformations under high winds. While these practices are time-consuming and labor-intensive, no alternatives have been implemented to date.

This paper reports on a feasibility study of using three-dimensional digital image correlation (3D-DIC) [9–12] as a non-contacting technique to measure uplift displacements on asphalt shingles subjected to wind loads. Feasibility is assessed through wind load testing of a full-scale roof panel where a target shingle was subjected to straight winds up to failure. The experiment was performed at the Insurance Institute for Business & Home Safety (IBHS) Research Center (Richburg, SC) using the newly-built COMET (Component Material Evaluation Testing) facility, which can deliver straight winds having speed up to 180 mph.

51.2 Experimental Program

51.2.1 Specimen and Target Shingle

Typical three-tab asphalt shingles used in residential construction were attached to a full-scale roof test panel with a surface area of 1,270 by 1,680 mm (Fig. 51.1). One shingle was replaced with a similar shingle having a lighter color, herein referred to as the "target shingle" (Fig. 51.2), to provide a better natural speckle pattern compared to that of the darker shingles, where a limited variation in the gray-level distribution was observed. It is noted that a natural speckle pattern, if effective, is preferable to painted patterns, which require additional test setup work and may affect the flexural stiffness of the shingles.

In fact, to ensure effective subset matching in DIC analysis, it is necessary that the variation in the gray level of a given speckle pattern be sufficiently large to distinguish each subset from other subsets [11]. For example, suitable speckle patterns can be obtained using the natural color variations on the specimen surface [13] or they can be painted on the measurement surface [11, 14, 15]. For the target shingle tab shown in Fig. 51.3a, Fig. 51.3b presents a histogram of the gray levels in the natural speckle pattern.

The bell-shaped distribution of the gray level in Fig. 51.3b indicates that the natural speckle pattern of the target shingle is suitable for DIC analysis.
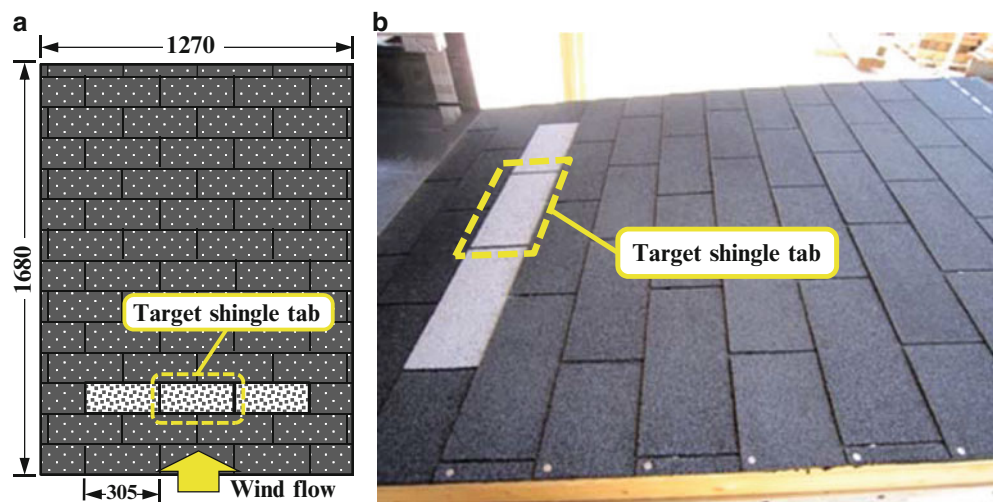

Fig. 51.1 Roof panel specimen with target shingle: (**a**) schematic of plan view; and (**b**) photograph. Dimensions in mm
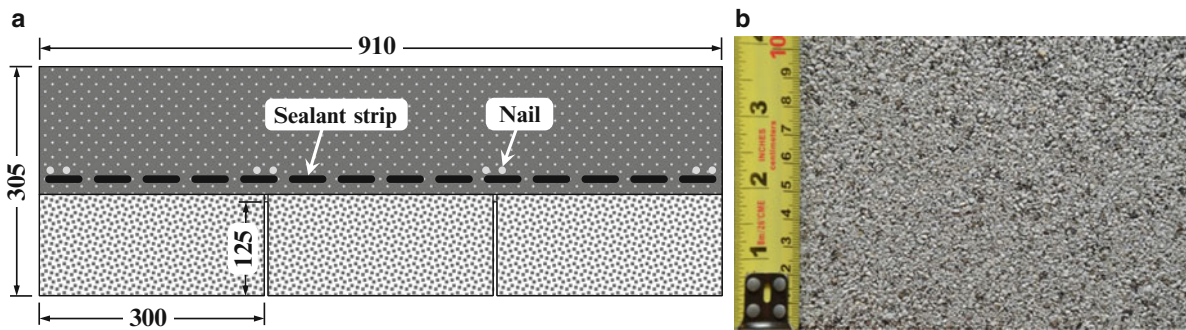
Fig. 51.2 Target shingle: (**a**) schematic; and (**b**) close-up view of natural speckle pattern. Dimensions in mm
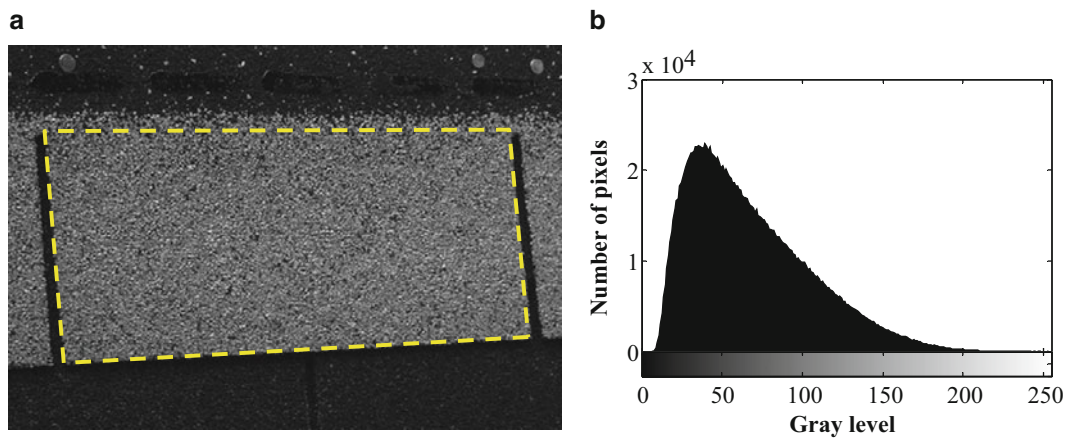

Fig. 51.3 Natural speckle pattern: (**a**) DIC image of target shingle tab; and (**b**) histogram of gray level
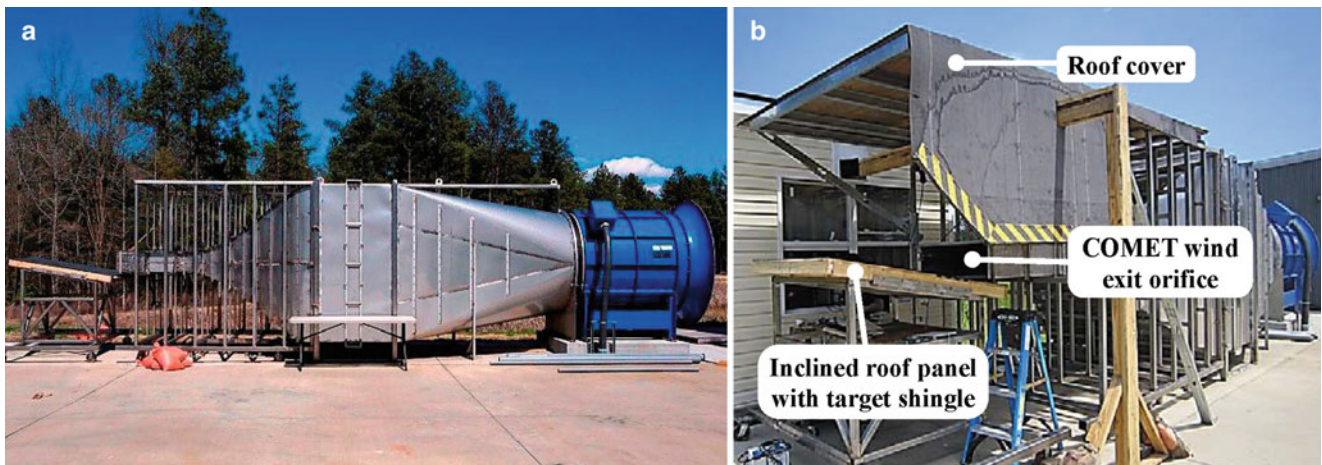

Fig. 51.4 Wind load test setup

51.2.2 Test Setup and Wind Load Protocol

The wind load test was conducted using the IBHS COMET facility shown in Fig. 51.4a. The machine includes a 1.7-m diameter vane-axial fan equipped with 350-hp medium-voltage electric motor, and is capable of delivering horizontal winds with a velocity up to 180 mph through a 915-mm wide by 305-mm high rectangular exit orifice. In addition, a variable-speed drive allows the wind speed to be varied during a test by 70 mph or more in less than 3 s.
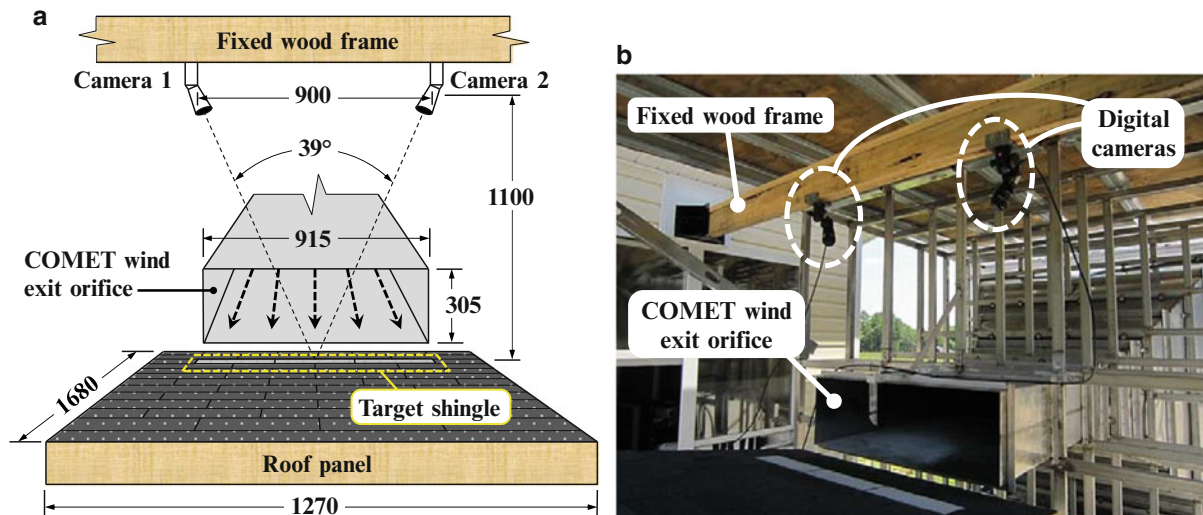
Fig. 51.5 3D-DIC setup: (**a**) schematic of front view (towards wind exit orifice); and (**b**) photograph. Dimensions in mm

The setups for the wind load test and DIC measurements are illustrated in Figs. 51.4b, 51.5, respectively. The roof panel specimen with the target shingle was rigidly attached to a fixed steel support frame with a 9.5° angle of inclination with respect to the horizontal plane. The specimen was then positioned in front of the COMET wind exit orifice. A roof cover was built above the specimen to minimize changes in ambient light during the load test. The stereo-vision setup used for the 3D-DIC measurements consisted of a pair of 5.0 MP (2,448 × 2,048 pixels) CCD digital cameras (Point Grey Grasshopper GRAS-50S5M-C) equipped with 28-mm lenses (AF Nikkor 28 mm f/2.8D). The cameras were rigidly attached to a fixed wood frame to minimize wind-induced vibrations and spaced 900 mm apart, resulting in a stereo angle of 39°. A desktop computer was used to record the images and perform DIC analysis using a dedicated software (Vic-3D, Correlated Solutions [16]). The stereo-vision system was calibrated by taking 130 images of a calibration grid in front of the target shingle with different orientations.

The wind load was imparted by means of straight winds with increasing wind speed of 19, 56, 84, 98, 113 127, 141 and 155 mph. Prior to failure of the sealant strip at a speed of 155 mph, each speed level was maintained for approximately 1 min. Images for DIC analysis were acquired at a frequency of 5 Hz. This relatively high frequency was used to reduce the likelihood of acquiring blurred images due to wind-induced vibrations.

51.3 Results and Discussion

51.3.1 Selection of Subset Size for DIC Analysis

A suitable subset size for DIC analysis must be large enough to ensure that there is a sufficiently distinctive pattern contained in the area used for subset matching. The subset size is typically determined based on the size of speckles. An effective rule of thumb is that each subset of size $N \times N$ should contain at least 3×3 speckles to ensure a reasonable DIC measurement accuracy [11]. Larger subsets containing more speckles may result in better accuracy but with the tradeoff of a reduced spatial resolution [17].

To investigate the effect of subset size on the accuracy of shingle uplift displacements, 31 images of the target shingle tab shown in Fig. 51.3a were analyzed at zero wind speed. One image was chosen as the reference. For all other images, the mean and standard deviation of the measured displacement components perpendicular to the roof panel surface for all data points was calculated to evaluate the bias error and the standard deviation error. A negligible change in bias was noted. Figure 51.6a shows the standard deviation error ('STD') for the uplift displacement measurements for subset sizes ranging from 31×31 to 91×91 pixels.

The results show that the STD decreases nearly linearly as the subset size increases. While using smaller subsets that provide higher spatial resolution is desirable, it was noted that for subset sizes equal or smaller than 51×51 pixels may
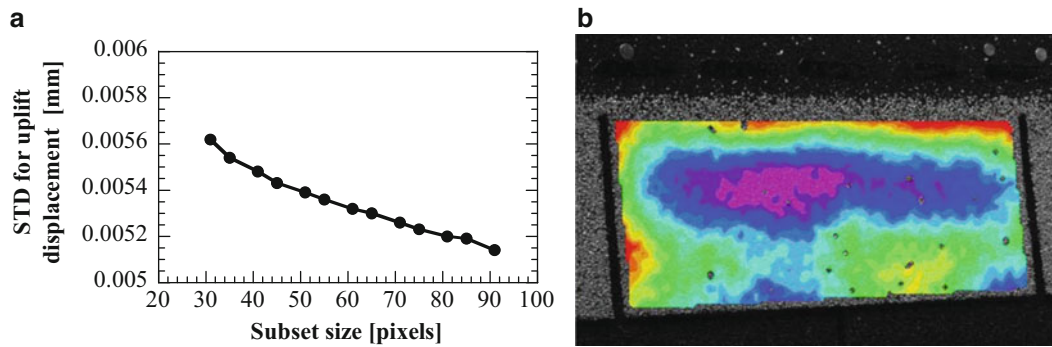
Fig. 51.6 Selection of subset size for DIC analysis: (**a**) standard deviation error of uplift displacement as function of subset size at zero wind speed; and (**b**) uplift displacement contour using 51 × 51 pixel subsets showing numerous null data points
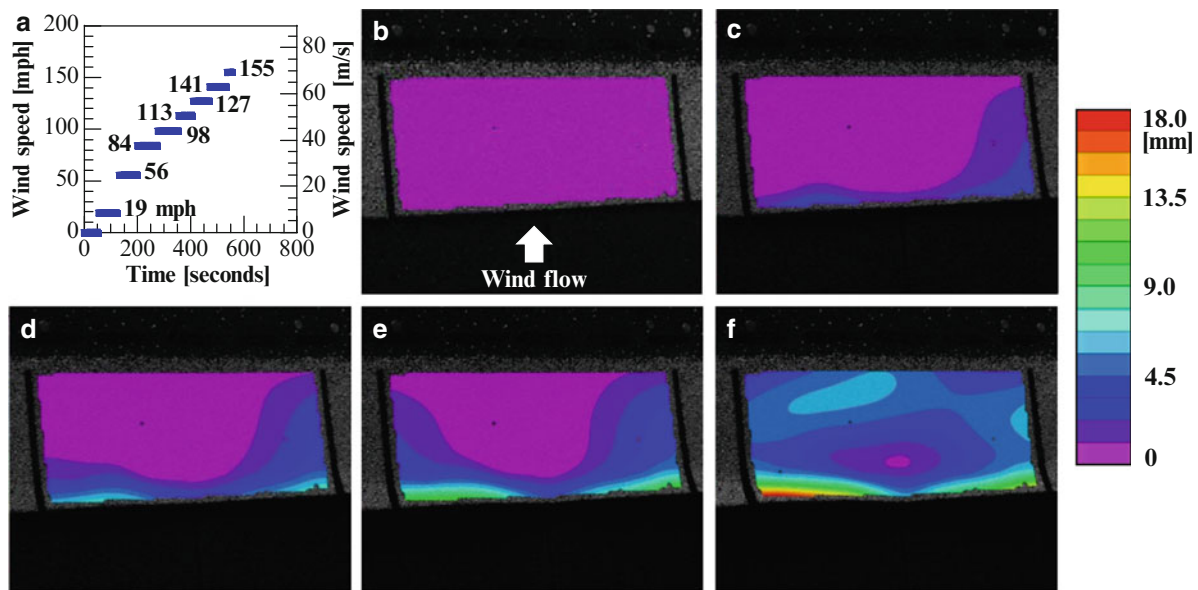

Fig. 51.7 3D-DIC measurements: (**a**) time history of wind speed; and (**b–f**) representative uplift displacement maps at wind speed of 0, 113, 127, 141 and 155 mph

yield DIC displacement maps with a relatively large number of null data points, as illustrated in Fig. 51.6b. Therefore, 61 × 61 pixel (corresponding to 8.8 × 8.8 mm) subsets were used. For this subset size, a standard deviation error of 0.0053 mm was estimated at zero wind speed as shown in Fig. 51.6a.

51.3.2 3D-DIC Measurements of Wind-Induced Uplift Displacements

Figure 51.7 presents the time history of wind speeds along with representative 3D-DIC uplift displacement maps for the target shingle tab at wind speeds ranging from 0 to 155 mph, when the sealant strip detached from the surface of the shingle located below the target shingle. The displacement contours clearly illustrate the deformed shape of the tab at increasing wind speeds, highlighting the uplifting of the free leading edge as well as the criticality of the corner locations where the sealant strip is reasonably subjected to more severe conditions, which may include combined shear and tension stresses.

The potential to use accurate full-field 3D-DIC uplift displacement maps to better understand the progressive and complex failure mechanisms of asphalt shingles under wind loads is demonstrated in Fig. 51.8.

Figure 51.8a shows a representative uplift displacement map at a wind speed of 155 mph, using the marker 'P1' to indicate a specific point near the failure area. Figure 51.8b presents the time-history of the uplift displacement, where the
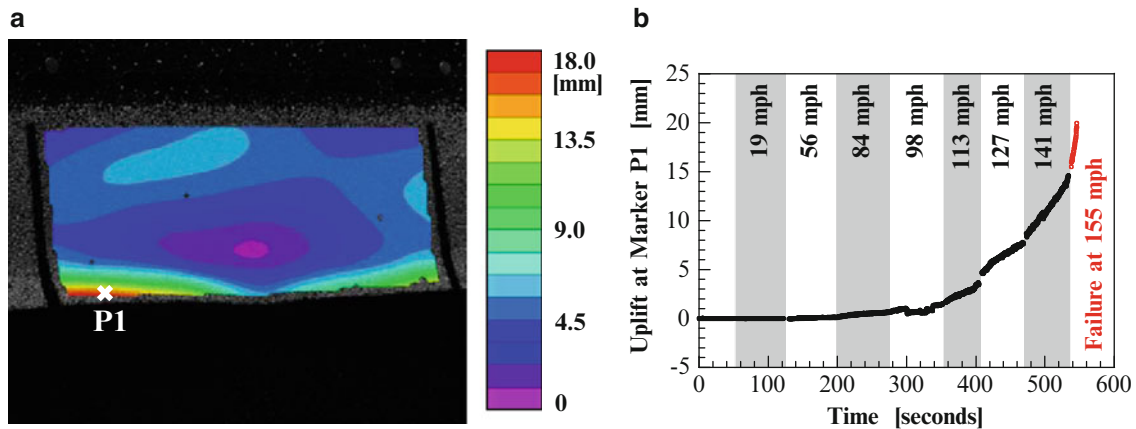
Fig. 51.8 3D-DIC uplift displacement at discrete point at increasing wind speed: (**a**) displacement map showing target point indicated by Marker P1; and (**b**) time history of measured uplift displacement at Marker P1 point up to failure

blank and gray bands indicate the time-periods for each constant wind speed reported in Fig. 51.7a. This evidence shows that consistent measurements were made at any given time irrespective of the wind speed. In fact, the displacement variability was always smaller than 0.1 mm (i.e., less than 0.6 % of the maximum uplift displacement measured on the target shingle tab). In addition, the progressively increasing displacements measured at higher wind speeds suggest that shingle and sealant visco-elastic deformations may significantly contribute to the increasing uplift displacement under high-speed but steady winds, which is the typical loading condition experience in an ASTM D3161 [3] test. It is noted that while partial debonding of the sealant strip is reasonably a contributing mechanism, especially near failure, no significant debonding was observed visually during the experiment at wind speeds below 155 mph, and the deformation eventually recovered after wind loading ceased. Additional research is necessary to better characterize these mechanisms, for example including wind load tests on roof panels where dedicated side-view cameras are used.

51.4 Conclusions

This paper presents preliminary results from a pilot study supporting the feasibility of non-contacting 3D-DIC surface deformation measurements on roof asphalt shingles subjected to high wind loads up to failure. The following conclusions are drawn.

1. For a given commercially available asphalt shingle, the natural color and surface texture can be selected to provide a gray-level distribution that is suitable for 3D-DIC subset matching.
2. For the setup used in the experiment presented in this paper, a subset size of 61×61 pixels was found to offer a good compromise between measurement accuracy and spatial resolution. The use of smaller subsets resulted in displacement maps having a relatively large number of null data points.
3. Consistent 3D-DIC uplift displacement measurements were obtained at any given time and wind speed up to failure. The inevitable vibration of the leading edge of the target shingle tab under steady winds was sufficiently contained as the displacement variability never exceeded 0.1 mm, which is less than 0.6 % of the maximum uplift displacement measured).
4. 3D-DIC displacement maps can be used to describe the progressive change of the deformed shape of an asphalt shingle at increasing wind speeds. Full-field 3D-DIC measurements offer a powerful tool to better understand the progressive and complex failure mechanisms under wind loads. For example, for the shingle tested in the experiment reported in this paper, the progressively increasing displacements measured at higher wind speeds highlight the importance of better understanding the influence of time-dependent shingle material deformations on the wind-induced uplift loads and failure mechanism, together with the progressive physical damage along the sealant strip.

References

1. Cullen W (1992) The evolution of asphalt shingles: survival of the fittest? In: Professional roofing, National Roofing Contractors Association, R4-R8, June 1992
2. Cash C (1995) Asphalt roofing shingles. In: Proceedings of the 11th conference on roofing technology, Gaithersburg, MD
3. ASTM D3161 (2009) Standard test method for wind-resistance of asphalt shingles (fan-induced method). ASTM International, West Conshohocken, PA
4. ASTM D7158 (2011) Standard test method for wind resistance of asphalt shingles (uplift force/uplift resistance method). ASTM International, West Conshohocken, PA
5. Liu Z, Pogorzelski H, Masters F, Tezak S, Reinhold T (2010) Surviving nature's fury: performance of asphalt shingle roofs in the real world. Institute for Business & Home Safety
6. Smith TL (1995) Improving wind performance of asphalt shingles: lessons from Hurricane Andrew. In: 11th conference on roofing technology, Gaithersburg, MD
7. Smith TL, McDonald JR (1990) Roof wind damage mitigation: lessons from Hugo. In: Professional roofing, National Roofing Contractors Association, 30–33, Nov 1990
8. Peterka JA, Cermak JE, Cochran L, Cochran B, Hosoya N, Derickson R, Harper C, Jones J, Metz B (1997) Wind uplift model for asphalt shingles. J Archit Eng 3(4):147–155
9. Helm JD, McNeill SR, Sutton MA (1996) Improved 3-D image correlation for surface displacement measurement. Opt Eng 35(7):1911–1920
10. Luo PF, Chao YJ, Sutton MA, Peters WH III (1993) Accurate measurement of three dimensional deformations in deformable and rigid bodies using computer vision. Exp Mech 33(2):123–133
11. Sutton MA, Orteu JJ, Schreier HW (2009) Image correlation for shape, motion and deformation measurements. Springer, New York, p 364
12. Sutton MA (2013) Computer vision-based, noncontacting deformation measurements in mechanics: a generational transformation. Appl Mech Rev Trans ASME 65:23
13. Küntz M, Jolin M, Bastien J, Perez F, Hild F (2006) Digital image correlation analysis of crack behavior in a reinforced concrete beam during a load test. Can J Civ Eng 33:1418–1425
14. Ghorbani R, Matta F, Sutton MA (2014) Full-field displacement measurement and crack mapping on masonry walls using digital image correlation. Adv Opt Methods Exp Mech 3:187–196
15. Ghorbani R, Matta F, Sutton MA (2014) Full-field deformation measurement and crack mapping on masonry walls using digital image correlation. Exp Mech (in review) (special issue: DIC Methods and Applications)
16. Vic-3D (2007) Correlated Solutions Inc., Columbia, SC
17. Bornert M, Brémand F, Doumalin P, Dupré JC, Fazzini M, Grédiac M, Hild F, Mistou S, Molimard J, Orteu JJ, Robert L, Surrel Y, Vacher P, Wattrisse B (2009) Assessment of digital image correlation measurement errors: methodology and results. Exp Mech 49(3):353–370