Machine Vision and Applications (2011) 22:157–170
DOI 10.1007/s00138-010-0255-2

ORIGINAL PAPER

Automatic multiple view inspection using geometrical tracking and feature analysis in aluminum wheels

Miguel Carrasco · Domingo Mery

Received: 27 August 2008 / Revised: 14 June 2009 / Accepted: 7 February 2010 / Published online: 6 March 2010

Abstract The classic image processing method for flaw detection uses one image of the scene, or multiple images without correspondences between them. To improve this scheme, automated inspection using multiple views has been developed in recent years. This strategy's key idea is to consider as real flaws those regions that can be tracked in a sequence of multiple images because they are located in positions dictated by geometric conditions. In contrast, false alarms (or noise) can be successfully eliminated in this manner, since they do not appear in the predicted places in the following images, and thus cannot be tracked. This paper presents a method to inspect aluminum wheels using images taken from different positions using a method called automatic multiple view inspection. Our method can be applied to uncalibrated image sequences, therefore, it is not necessary to determine optical and geometric parameters normally present in the calibrated systems. In addition, to improve the performance, we designed a false alarm reduction method in two and three views called intermediate classifier block (ICB). The ICB method takes advantage of the classifier ensemble methodology by making use of feature analysis in multiple views. Using this method, real flaws can be detected with high precision while most false alarms can be discriminated.

Keywords Automated inspection · Tracking · Flaw detection · X-ray imaging · Nondestructive testing

M. Carrasco (✉)
Escuela de Ingeniería Informática, Facultad de Ingeniería,
Universidad Diego Portales, Av. Ejército 441, Santiago, Chile
e-mail: miguel.carrasco@mail.udp.cl; mlacarrasco@gmail.com

M. Carrasco · D. Mery
Departamento de Ciencia de la Computación,
Pontificia Universidad Católica de Chile,
Av. Vicuña Mackenna 4860 (143), Santiago, Chile
e-mail: dmery@ing.puc.cl

1 Introduction

Over the last 30 years, the worldwide manufacturing market has faced heavy competition to produce higher quality products while actively reducing prices. This has led to great advances in the technology required for automating production processes but inspection and quality control problems have yet to be fully resolved. Due to these gaps in the industry, several automatic inspection techniques represent an area of high interest and active research. Traditionally, inspection and quality control in manufacturing environments have been carried out by means of an intensive human visual inspection inserted into different phases of the production processes [33]. The economic benefits represent some of the main reasons this kind of inspection is used. The investment cost to install and develop a specialized machine for inspection tasks is very high compared to the cost of training a human operator. Also, human visual inspection has the great advantage of adapting to unforeseen situations and is flexible when faced with any change in the objects' position, orientation or shape. This is because human beings have high cognitive and sensory abilities that allow them to carry out complex reasoning and inferences while inspecting the objects [40].

Various studies have analyzed the performance of human inspection and its main defects (e.g., [10,11,19,32]). According to them, there is a clear consensus that human inspection does not achieve 100% performance in the detection of defect-free products (error-free). Mital et al. [32] determined various factors that affect the performance of manual inspections, such as the rhythm and complexity of the task, the time for inspection, fault density, inspection model, luminosity, inspection strategy, training, age, and gender. Other authors have indicated that human inspection has a maximum of 80% effectiveness [11]. Human inspection does have constraints as well as multiple failures, it is (1) variable, inspection quality is not constant over time because it is dependent on fatigue and monotony caused by the work; (2) irregular, because it depends on the ability, experience and strategy for revision of each inspector; (3) slow, some industries have high production levels and require inspection at a high processing rate, however, human inspection can require more time because handling and observation tasks have limiting factor, such as the speed of human operations; (4) tedious, because the inspection routine can be very repetitive which generates a lower concentration level due to the large number of objects that must be revised in a short period; (5) hazardous, because in some environments, such as under water inspection, the nuclear industry, and the chemical industry, human inspection can be inviable due to the high risk inherent in those systems; (6) complex, the difference between a product with or without defects can be very subtle, and that is not always easily distinguishable by a human operator; (7) inaccessible, in some cases even access to the object to be inspected can be very complex because of the size of the product. All of these factors have lead industry to gradually replace human inspection with automatic visual inspection (AVI) methods which allow contact free inspections to be made of the object.

Since the introduction of AVI methods in the early 1980s [8,20], several systems for quality inspection have been successfully developed using different image processing techniques. The main objective of AVI is to increase productivity ensuring high quality, reliability and consistency standards, i.e., rejecting most of the defective products and accepting all the defect-free products. AVI inspections normally require less time than inspections performed by human operators. Malamas et al. [25] and Kumar [22] have presented extensive reviews of various AVI technologies applied to the manufacturing processes of different products, such as electronic components, textiles, glass, mechanical parts, integrated circuits (IC), etc. Despite their advantages, AVI methods in general also have the following problems: (1) they lack precision in their performance, because of the imbalance between undetected flaws (false negatives) and false alarms (false positives); (2) they are limited by time, the mechanical requirements for placing an object in the desired position can be time consuming; (3) they require high computer cost for determining whether the object is defective or not; (4) they generate high complexity in the configuration and lack of flexibility for analyzing changes in parts design. The issues outlined above show that AVI remains a problem open to the development of new applications.

In many AVI systems, the use of one image to carry out quality inspection is sufficient. However, in other cases where the signal-to-noise ratio is low, the identification of real flaws with little contrast implies the appearance of numerous false alarms. It is precisely in these cases where multiple views can improve the inspection performance in the same way a human inspector uses his sight to see multiple parts of an object to evaluate its quality.

In this paper, we aim to exploit the redundant information from multiple views that contain corresponding parts of the object. The information captured from different viewpoints can reinforce the diagnosis when a single image is insufficient. To discriminate real flaws from false alarms, our system tracks every possible flaw. Only real flaws can be successfully tracked along an image sequence. A real flaw entails a spatio-temporal relation in different views where it appears while a false alarm corresponds to a random event allowing us to distinguish real flaws from other artifacts. Based on this observation, we propose a three-step methodology for detecting real flaws in uncalibrated image sequences of aluminum wheels: segmentation of potential flaws, computation of corresponding points, and tracking of potential flaws with intermediate classifiers. Similar ideas have been presented in [4,5,29,36,41]. The main differences between this contribution and those works lie in the fusion of multiple view geometry and a statistical analysis of each flaw aiming to reduce the number of false alarms while simultaneously improving the true flaws detection in correspondence.

It is important to highlight that our method does not require a calibration process. In general, the calibration process is difficult to carry out in industrial environments due to vibrations and random movements that vary with time. The vibrations of the imaging system induce inaccuracies in the estimated parameters of the multiple view geometric model. Thus, the calibration is not stable and the imaging system must be re-calibrated periodically. In many cases, it might be an extremely complicated procedure for real-time applications and manufacturing systems that cannot be stopped temporarily for calibration purposes [36].

The rest of the paper is organized as follows: Sect. 2 includes a brief discussion of automatic multiple visual inspection; Sect. 3 explains our proposed method for uncalibrated image sequences; Sect. 4 shows the experimental results; and finally, Sect. 5 presents the conclusions and future work.

2 Automatic visual inspection

Currently, one of the most widely used flaw detection systems in industry is the X-ray inspection, extensively used by the automotive and aerospace industry, for detecting flaws like: porosity, cracks, corrosion, inclusions, debris, bubbles, and thickness variations, among others [2,15,34]. It is commonly used because the X-ray attenuation surrounding the flaws is less (or more). The use of X-rays exploits the fact that most material flaws are not visible. However, even in radioscopic images the signal-to-noise ratio (SNR) is low, the flaw signal is slightly greater than the background noise meaning that the identification of real flaws with poor contrast can involve detection of false alarms as well.

Motivated by (human) visual inspections that are able to differentiate between flaws and noise by looking at the objects being tested in motion, a new method of automated inspection was developed using sequences of multiple images [29]. The new inspection methodology called automatic multiple view inspection (AMVI) uses redundant views to perform the inspection task. This novel methodology is opening up new possibilities in the inspection field, mainly by taking into account the useful information in the corresponding different views of potential flaws in the test object. The main idea is to consider as false alarms those potential flaws that cannot be tracked in a sequence of multiple images. Therefore, two or more views of the same object taken from different viewpoints can be used to confirm and improve the diagnosis made by analyzing only one image. AMVI has been developed under two schemes: calibrated and uncalibrated. The calibrated scheme uses a 3D calibration object to estimate corresponding points [29]. Alternatively, the uncalibrated scheme automatically establishes the correspondences from the information contained in the images through a robust correspondence system [4] (see Fig. 2). These steps are equivalent to the work done by an inspector. First, all the possible regions that might contain flaws (or potential flaws) are detected. Second, because of the large number of false alarms that can occur in the identification step, the corresponding positions that each flaw might have in the following views are analyzed, using multiple view tracking (see Fig. 1). Both methods share the following two steps: *identification* and *tracking*.

Identification aims at detecting all the anomalous regions or potential flaws in each image of an object's motion sequence, without a priori knowledge of its structure. There are two general features used to identify them: (1) a flaw is considered as a connected subset in the image, and (2) the differences between the gray levels of the flaw and its neighbors is considerable. Although there are a lot of false alarms detected by this process, the detector has the following advantages: (1) the same detector is applied to all the images; (2) it allows for the identification of potential flaws regardless of the position or the structure of the object under study; in other words, without a priori knowledge of the design of the structure; (3) the detection of real flaws is very high (better than 90%).[1] The process that follows extracts features of each potential flaw after identifying these regions in the previous procedure. This information makes it possible to determine whether a flaw is corresponding in the multiple view analysis, according to the new intermediate classification method.

Tracking aims at "chasing", in subsequent images of a sequence, potential flaws detected in the first step using the positions forced by the geometric restrictions in multiple views [18]. If a potential flaw continues through an image sequence, it is identified as a real flaw and the object is classified as defective. However, if a potential flaw does not have a correspondence in the sequence, it will be considered as a false alarm (details in the segmentation in Fig. 1). A similar idea is also used by radiologists that analyze two different X-ray views of the same breast to detect cancer in its early stages. Thus, the number of cancers flagged erroneously as well as missed cancers may be greatly reduced (see for example Kita et al. [21], where a novel method that automatically finds correspondences in two different views of the breast is presented).

3 Proposed method

In this section, we provide an explanation of the stages in the uncalibrated AMVI process with intermediate classifiers. The proposed scheme has four major steps (A, B, C and D) detailed in Fig. 2. They correspond to the following stages: (A) identification of potential flaws, (B) extraction of control points, (C) tracking, and (D) intermediate classifier block.

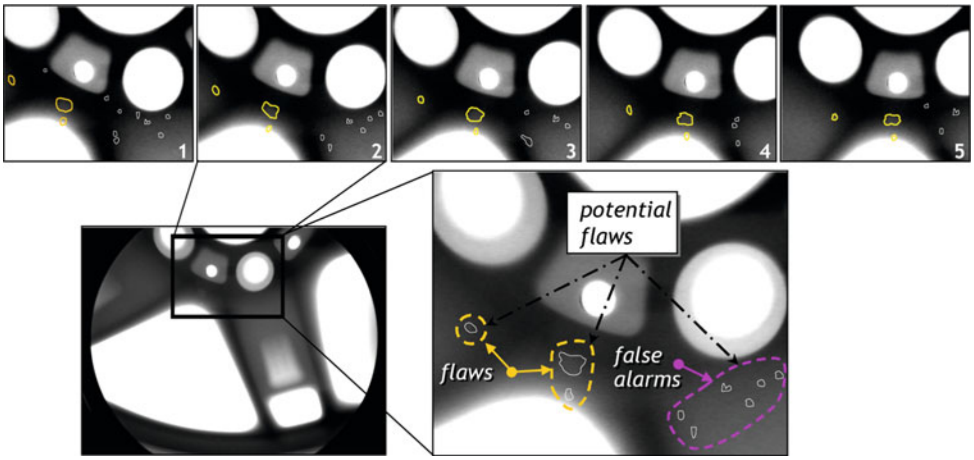
(A) **Segmentation of potential flaws** Numerous investigations have been carried out to segment flaws depending on the product analyzed [3,26,31,35]. Here we used the segmentation and feature extraction method, described in [26], with the aim of identifying multiple regions which may correspond to real flaws. In particular, the segmentation algorithm used is able to detect most real flaws as well as numerous false alarms. The process consists of the following, each potential flaw extracts a set of measurements (described in Table 1) and stores them in a normalised feature vector. For instance, let $m_a^i = [x_a^i, y_a^i, 1]$ be the centre of mass stored in homogenous coordinates of the segmented region i in the ath view, and let $\mathbf{v}_a^i$ be the feature vector of the region i in the ath view. As a result, numerous potential flaws appear as observed in the segmented image (Fig. 1).

(B) **Robust control points** As stated before, our final goal is tracking real flaws in an image sequence. For this purpose, accurate corresponding points between every pair of views are required. In general, the estimation of control points can be solved by various mechanisms that use the intrinsic information of the structures after a process of segmentation, edges extraction, normalization, and smoothing [24,38]. In general, there are two curve alignment categories: methods based on rigid transformations [43] and methods based on

---

[1] See [26] for details on the computation of the segmentation algorithm.

Fig. 1 *Top* Example of a radioscopic sequence of five images after the segmentation step. *Bottom* Example of false alarms and flaws, called potential flaws at this stage



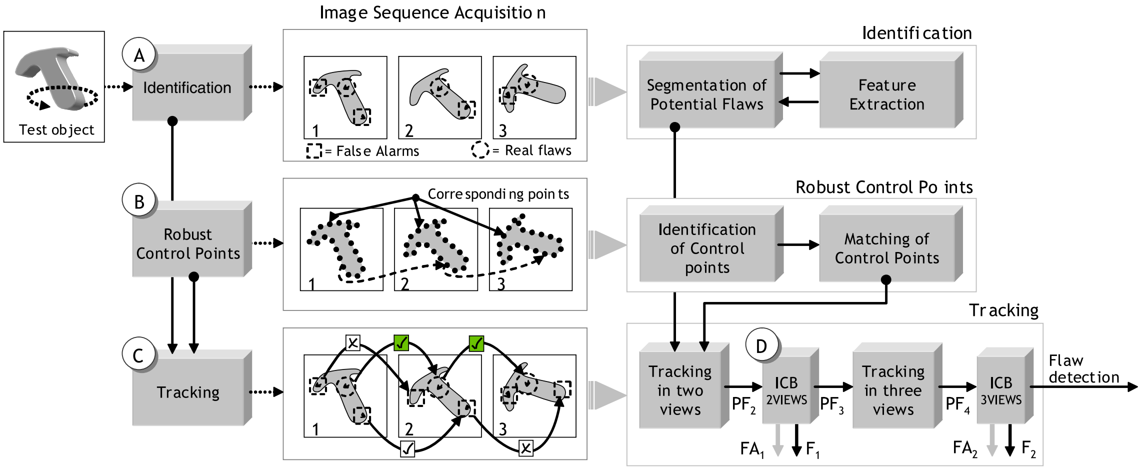


Fig. 2 General block diagram of the uncalibrated automatic multiple view inspection (AMVI) method with the three phases: identification, robust control point and tracking of potential flaws in two and three views with the intermediate classifier block (ICB) method

non-rigid deformation [9]. First, methods based on rigid transformations determine the control points by estimating the rotation, lineal displacement, and scaling parameters [24]. However, due to the rigidity assumption they are sensitive to occlusions, deformations, articulations, perspective projections, and other variations of the edges [38]. Second, methods based on non-rigid deformations try to match one curve over the other. The goal is minimizing a function of elasticity through the transformation of the curve flexion, orientation or stretching. Generally, this transformation is not invariant under rotation and scaling [16], it is very sensitive to noise because it is defined in terms of the curvature, and it requires the evaluation of second-order derivatives [38].

Our investigation proposes a simple and effective curve alignment method by minimizing the Pearson's correlation coefficient using an isometric transformation between two curves. We use this scheme because in the analysis of manufactured products the object being analyzed is usually not deformable. This premise justifies the use of a rigid transformation method with which, given a rotation and a lineal displacement, it is possible to estimate a correspondence between the object's control points. However, due to the object's rotation, some regions can remain occluded, and, therefore, the proposed system must consider that only some regions retain this transformation. The proposed robust system of control points consists of two stages that are detailed below: matching of regions, and matching of control points.

(B.1) **Matching of regions** This consists of establishing correspondences between regions of each view and not between control points. The designed process is composed by four stages: First, segmentation of those regions in which the intensity of the object is distinguishable from the background using Otsu's method [17] (Fig. 3a). Second, extraction of a set of features for each segmented region. This consists of extracting the moments of Flusser-and-Suk [39] of

Table 1 Features extracted from the identification step

Symbol	Feature and description
A	Area: Number of pixels that belong to the region
G	Mean of the grey: Mean of the grey values that belong to the region [30]
D	Mean of the second derivative: Mean of the second derivative values of the pixels that belong to the boundary of the region [30]
F_1	Crossing line profiles: The grey level profiles along straight lines crossing each segmented potential flaw in the middle. The profile that contains the most similar grey levels in the extremes is defined as the best crossing line profile (BCLP). Feature F1 corresponds to the first harmonic of the fast Fourier transformation of BCLP [26]
$K\sigma$	Contrast: Standard deviation of the vertical and horizontal profiles without offset [30]
r	High contrast pixels ratio: Ratio of number of high contrast pixels to area [27]

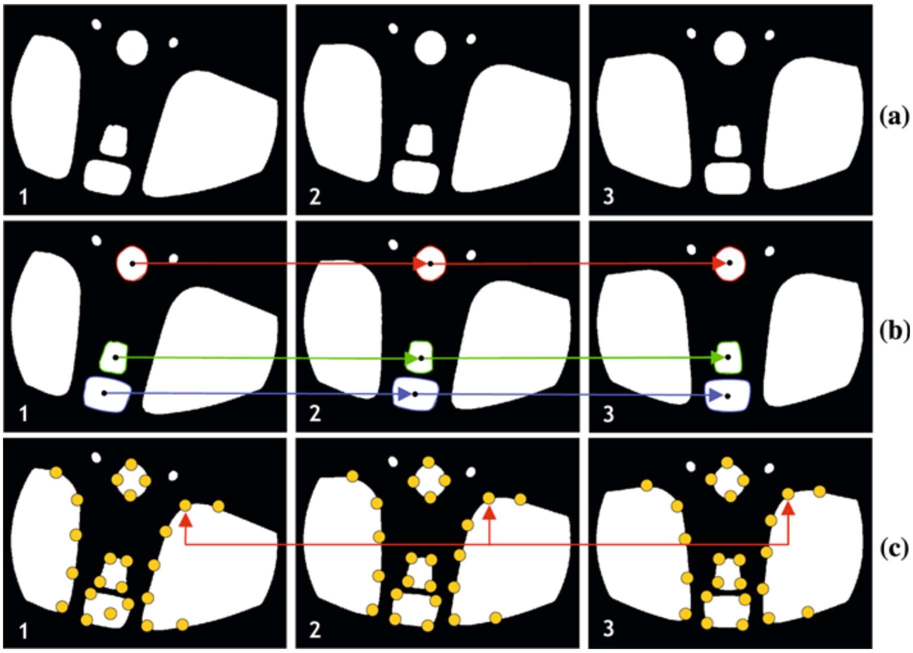


Fig. 3 Matching of regions. **a** Segmentation of regions in a sequence by Otsu's method. **b** Correspondence between regions according to a similarity criterion between the extracted features. **c** Matching of control points on the border (i.e., the curve) of a region

each region in three views. Third, determination of a region-correspondence using the features extracted before by relating those regions with greater similarity. The similarity relation is fulfilled when two or three regions have little variation in their normalized features according to the Euclidean distance metric (Fig. 3b). Fourth, smoothing the edges of each region in correspondence, in order to decrease the noise of each curvature. For that we calculate the perimeter of each segmented region and generate a list in a parametric form as $Z_s = [x_s, y_s]$, where $s = 0, \ldots, L-1$ is the index of the list of pixels ordered in a turning direction, and L is the number of pixels of the region's perimeter. Using this parametric form, we generate the Fourier descriptors, transforming the Z_s coordinates into a complex value $u_s = x_s + j \cdot y_s$. This signal with period L is transformed into the Fourier domain by means of a discrete Fourier transform (DFT):

$$F_n = \sum_{s=1}^{L-1} u_s \cdot e^{-j \cdot \frac{2\pi \cdot s \cdot n}{L}}$$

The modulus of the complex Fourier coefficients describes the energy of each descriptor. Therefore, if we choose the highest energy coefficients (above 98%) and return to real space with the inverse discrete Fourier transform (IDFT) we get a smoother curve with less noise. This transformation produces the same number of points as the original curve. Likewise, the spacing between the original points remains constant. However, when applying the elimination of some

Fourier coefficients, the original curve is transformed into a new curve $C_s = [x'_s, y'_s]$, where, $C_s \neq Z_s$.

(B.2) **M**atching of control points The estimation of control points is a process in which the correspondence of pairs-points on the border of a region is established (Fig. 3c). Using Fourier procedure as described above, we define a curve C_1 corresponding to a region in the first view, and a curve C_2 corresponding to C_1 in the second view. Both curves do not have the same length because they correspond to the perimeter of corresponding regions. However, these regions have an isometric transformation, and in cases of occlusion the curves will have different sizes. For both curves, to keep the same distance and to be aligned, it is necessary to select a section of equal length from each list. Let P, a section of curve C, be such that $P = C(\delta)$, where $\delta = [s_i, \ldots, s_j]$, for $i, j \in [1, \ldots, n]$. In this way, there is a section P_1 in the first view that has the same length as section P_2 in the second view. These sections of the curve do not necessarily have a correspondence, and for that we define a shift operator $\Theta(P, \lambda)$ that displaces the list P by λ positions in a turning direction. Operator Θ uses the function "mod" (modulus after division) to determine the λ relative positions that list C, of length P, must turn.

Using the above definitions, we implemented an alignment function $\mu(\Omega)$ as the maximization of the Pearson's correlation coefficient $\rho(\alpha, \beta)$ [13] between the isometric transformation of a section of P_1, with the shift of section P_2 with a jump λ, composed by four parameters $\Omega = \{\theta, \Delta s_x, \Delta s_y, \lambda\}$

$$\mu(\Omega) = |1 - \rho([R, t][P_1], \Theta(P_2, \lambda))| \rightarrow \min \tag{1}$$

where,

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad t = \begin{bmatrix} \Delta s_x \\ \Delta s_y \end{bmatrix} \tag{2}$$

The minimization of $\mu(\Omega)$ must find Ω parameters to estimate an alignment between sections P_1 and P_2. The main advantage of this function is that it does not require a perfect alignment because the correlation coefficient takes a maximum if the displacement is linear. Another advantage is that curves P_1 and P_2 are open, the alignment determines only sections that are corresponding, allowing control points to be obtained for curves that have partial occlusion in corresponding regions. Also, the use of the parameter λ allows finding a position relation for curve C_2 with P_1, and in this way, while curve P_2 adjusts its shift, curve P_1 adjusts its lineal displacement and rotation angle to become aligned.

(C) **Tracking of potential flaws** In the previous steps, we have segmented all potential flaws along an image sequence and we have established the corresponding points in a sequence. We now turn to the problem of separating real flaws from false alarms. The essential point is that only real flaws can be tracked along an image sequence. A real flaw entails a spatio-temporal relation in different views where it appears, while a false alarm corresponds to a random event.

(C.1) **Two views** If a potential flaw $\mathbf{m}_a^i$ in view a is actually a real i-flaw it must have a corresponding point $\mathbf{m}_b^j$ in another consecutive view b where a potential flaw j was also segmented. According to the *principle of multiple view geometry* [18], points $\mathbf{m}_a^i$ and $\mathbf{m}_b^j$ are in correspondence if matrix $\mathbf{F}_{a,b}$ exists such that

$$\mathbf{m}_b^{j\top} \cdot \mathbf{F}_{a,b} \cdot \mathbf{m}_a^i = \mathbf{m}_b^{j\top} \cdot \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \cdot \mathbf{m}_a^i = 0 \tag{3}$$

where $\mathbf{F}_{a,b}$ is known as the fundamental matrix of the projection of points $\mathbf{m}_a^i$ and $\mathbf{m}_b^j$, in homogenous coordinates as $[x_a^i, y_a^i, 1]^\top$ and $[x_b^j, y_b^j, 1]^\top$, respectively. Once the set of corresponding positions has been generated in both views, we use the method proposed by Chen et al. [7] to make an initial estimation of the fundamental matrix. For the sake of completeness, we briefly describe the Chen's method below.

The method proposed by Chen et al. is based on choosing a subset of candidate points by exploiting the fact that the rank of the fundamental matrix is two and the value of the epipolar restriction is zero ($\mathbf{m}_b^{j\top} \cdot \mathbf{F}_{a,b} \cdot \mathbf{m}_a^i = 0$). Based on this, it is possible to define the fundamental matrix without losing generalization by means of a lineal combination of its values, where it is only necessary to estimate four parameters $\Gamma = \{\alpha, \beta, \alpha', \beta'\}$ according to

$$\begin{aligned} f_3 &= -\alpha' f_1 - \beta' f_2 \\ f_6 &= -\alpha' f_4 - \beta' f_5 \\ f_7 &= -\alpha f_1 - \beta f_4 \\ f_8 &= -\alpha f_2 - \beta f_5 \\ f_9 &= -\alpha' \alpha f_1 - \alpha' \beta f_4 + \beta' \alpha f_2 + \beta' \beta f_5 \end{aligned} \tag{4}$$

Using the above parameters and the point correspondences detected in the control point step, we define a new problem as $\mathbf{A} \cdot \mathbf{f} = 0$, where $\mathbf{f} = [f_1, f_2, f_4, f_5]^\top$. Since the values of $\mathbf{f}$ are not null, only $\det(\mathbf{A}) = 0$ is possible. Likewise, to find a solution of matrix $\mathbf{A}$, the Γ parameters are defined randomly.

Thus, it is necessary to build a matrix **A** as

$$\mathbf{A} = \begin{bmatrix} (x_1-\alpha)(x'_1-\alpha') \dots (x_n-\alpha)(x'_n-\alpha') \\ (x_1-\alpha)(y'_1-\beta') \dots (x_n-\alpha)(y'_n-\beta') \\ (y_1-\beta)(x'_1-\alpha') \dots (y_n-\beta)(x'_n-\alpha') \\ (y_1-\beta)(y'_1-\beta') \dots (y_n-\beta)(y'_n-\beta') \end{bmatrix}^{\top}_{4\times N} \tag{5}$$

Since the matrix **A** is composed by N stereo correspondences, Chen proposed that using only a combination of four restrictions will be enough to seek a fundamental matrix. For this reason, we define a row vector $\mathbf{B}_i$ of the matrix **A** as

$$\mathbf{B}_i(\alpha,\beta,\alpha',\beta') = \begin{bmatrix} (x_i-\alpha)(x'_i-\alpha') \\ (x_i-\alpha)(y'_i-\beta') \\ (y_i-\beta)(x'_i-\alpha') \\ (y_i-\beta)(y'_i-\beta') \end{bmatrix}^{\top} . \tag{6}$$

Then, choosing randomly four restrictions (i, j, k, l) between N rows, the problem that follows must fulfill the following condition

$$\mathbf{B}_{ijkl} = \det\left(\left[\mathbf{B}_i \ \mathbf{B}_j \ \mathbf{B}_k \ \mathbf{B}_l\right]^{\top}\right) = 0 \tag{7}$$

where rows (i, j, k, l) belong to the combinatorial subset of $R = C^N_4$. As we stated before, to estimate the Γ parameters it is necessary to solve the optimization problem by looking for four $\mathbf{B}_i$ restrictions such that (7) is a minimum according to

$$(\alpha,\beta,\alpha',\beta') = \operatorname{argmin}_{\alpha,\beta,\alpha',\beta'} \sum_{1}^{N} |\mathbf{B}_{ijkl}| \tag{8}$$

Next, replacing the Γ parameters in (5) allows us to compute the vector **f** by decomposing into singular values $[\mathbf{S}, \mathbf{V}, \mathbf{D}] = \mathrm{svd}(\mathbf{A})$.

Thus, the solution of (5) corresponds to the last column vector **D**. Finally, the Γ parameters obtained in (8) are replaced in (4) defining an initial solution of $\mathbf{F}_{a,b}$. In our research, we combine this procedure with the RANSAC algorithm, thus the solution with the largest number of inliers is used to compute the fundamental matrix.

The next procedure is to estimate the epipolar line based on previous results. More formally, let $\mathbf{l}^i_a$ be the epipolar line defined as the product between the fundamental matrix $\mathbf{F}_{a,b}$ and the point $\mathbf{m}^i_a$ as

$$\mathbf{l}^i_a = [l^i_{a,x}, l^i_{a,y}, l^i_{a,z}] = \mathbf{F}^{\top}_{a,b} \cdot \mathbf{m}^i_a, \tag{9}$$

where $\mathbf{l}^i_a$ is the epipolar line of flaw i in view b, $\mathbf{m}^i_a$ is the centre of mass of flaw i in view a and $[l^i_{a,x}, l^i_{a,y}, l^i_{a,z}]$ are the coefficients of the epipolar line. Once the epipolar line $\mathbf{l}^i_a$ has been generated, it is necessary to determine the distance between the corresponding flaw in view b. This distance is determined through the *practical bifocal constraint* [18]. Given that the epipolar constraint is applied to points and not to regions, we consider the centre of mass of each potential flaw to be a corresponding point between pairs of images.

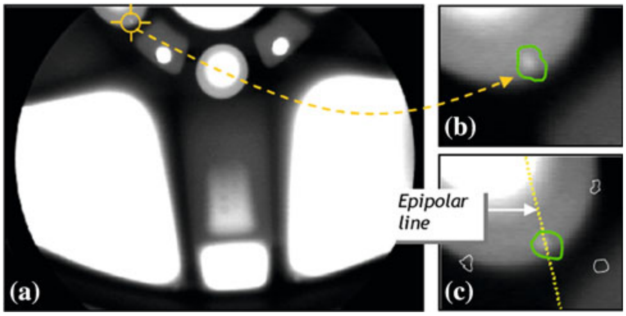


Fig. 4 Epipolar line generated from the fundamental matrix. **a** First view; **b** segmentation of a potential flaw in the first view; **c** intersection of the epipolar line in the second view with one or more corresponding potential flaws

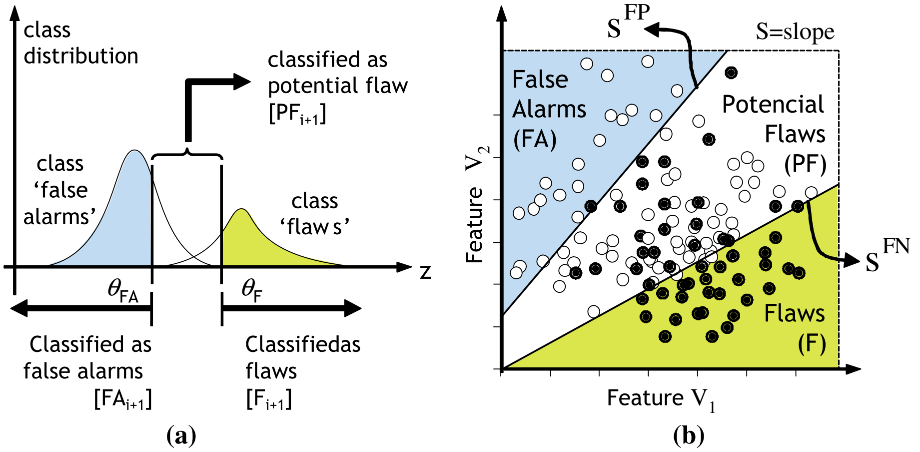
Using the epipolar line $\mathbf{l}^i_a$, we identify the correspondence associated with the potential flaw i as the potential flaw j in view b that satisfies the constraints

$$\frac{|\mathbf{m}^{j\top}_b \, \mathbf{F}_{a,b} \, \mathbf{m}^i_a|}{\sqrt{(l^i_{a,x})^2 + (l^i_{a,y})^2}} < \varepsilon_1 \tag{10}$$

for small $\varepsilon_1 > 0$. If this constraint is fulfilled, a potential flaw is thus found in two views. In this case it could be considered a real flaw with a bifocal correspondence. Otherwise, it is regarded as a false alarm. An example is shown in Fig. 4. The same procedure is applied to every potential flaw in view a that is to be found in view b. It is important to recall that the precision of the fundamental matrix allows the correct determination of potential flaws along the length of the epipolar line. However, if the fundamental matrix is not robust, the epipolar line will be incorrect and the subsequent trifocal tensor process will fail.

(C.2) **Three views** Trifocal analysis allows modeling all the geometric relationships in three views, and is independent of the structure contained in each image [18]. The tensor, a matrix structure similar to the fundamental matrix, only depends on the movement between images and the internal parameters of the cameras. Its main advantage is that it can be calculated from the correspondences of the images without any a priori knowledge of the movement or calibration of the object. This characteristic justifies it because the estimation of the fundamental matrix does not always eliminate all false positives.

Based on previous results, to confirm that a bifocal correspondence indeed represents a real flaw, we try to discover a new correspondence using a third view with the help of trifocal tensors. Let $\mathbf{T} = (T^{rs}_t)$ be a $3 \times 3 \times 3$ matrix representing the trifocal tensor that encodes the relative motion



(a) (b)

Fig. 5 **a** Distribution of classes of potentials flaws between views. **b** Distribution of three classes in two dimensions with a linear separation between F, FA and PF regions

among views a, b, c.[2] Then, we can estimate the hypothetical position of a flaw k in a third view c using the correspondences $\mathbf{m}_a^i$, $\mathbf{m}_b^j$ and the tensor $\mathbf{T}$ as[3]

$$\widehat{\mathbf{m}}_c^k = \frac{1}{\mathbf{m}_a^{i\top}(T_{\cdot}^{13} - x_b^j\, T_{\cdot}^{33})} \begin{bmatrix} \mathbf{m}_a^{i\top}(T_{\cdot}^{11} - x_b^j\, T_{\cdot}^{31}) \\ \mathbf{m}_a^{i\top}(T_{\cdot}^{12} - x_b^j\, T_{\cdot}^{32}) \\ \mathbf{m}_a^{i\top}(T_{\cdot}^{13} - x_b^j\, T_{\cdot}^{33}) \end{bmatrix}. \quad (11)$$

We compare the estimated position with all potential flaws of view c, regarding the potential flaw k as a real flaw if

$$\|\widehat{\mathbf{m}}_c^k - \mathbf{m}_c^k\| < \varepsilon_1, \quad (12)$$

for small $\varepsilon_1 > 0$. If this constraint is fulfilled, we take the potential flaw to be a real flaw, since it complies with the correspondence in three views. Should the potential flaw in the third view not coincide with the projection of the tensor, it is discarded, as it does not fulfill the trifocal condition. In general, given that the trifocal condition is analyzed for the sequences that fulfill the bifocal condition, we reduce the number of false positives generated in two views.

(D) **Intermediate classifier block method** The goal of the ICB method is to eliminate those correspondences between potential flaws that have a low probability of being true positives. The ICB method uses the classifier ensemble methodology [37], in which multiple linear classifiers do the classification and then, through the majority of votes technique, a final decision is made. According to the multiple view hypothesis, the key idea is to consider as false alarms those potential flaws that cannot be tracked in a sequence of multiple images. Nonetheless, there are false alarms that fulfill the above condition and must be eliminated in the multiple views analysis using a partial elimination classification system. The ICB method has as input the distribution of two classes: flaws (F) and false alarms (FA). According to this distribution, the classifier must determine the region of space where there are actually flaws only starting from point θ_F, and false alarms from θ_{FA} (Fig. 5a). Once these regions are extracted only flaws or false alarms (which the classifier cannot verify with high probability the class to which they belong) are assigned to a new class called potential flaw (PF) (Fig. 5b). As a result, the ICB method generates the separation of three classes (F, FA and PF). Accordingly, flaws or false alarms contained in the PF region are used as new potential flaws in the following step of the multiple views analysis. This reduction avoids the analysis of the trajectories of all flaws in correspondence; consequently, improving the performance. The simplest form of the previous classifier is reflected in the linear separation of the F, FA and PF regions, using the V_1 and V_2 features (Fig. 5b). The methodology used by the ICB is composed by a series of stages detailed below (see Fig. 6).

(D.1) **Assessment method of the ICB classifier** Our problem falls within the framework of supervised classification problems, since the class which each potential flaw belongs to is known. Using this information, the classification model is designed by means of the cross-validation method. To compare the results of the various configurations of the classifier we use the ROC curve [14]. The main advantage of the ROC curve is that it allows the comparison to be independent of the sample. For this research, the classes are a set of registers with flaws and false alarms, we determine the sensitivity (S_n) and 1-specificity $(1 - S_p)$ as $S_n = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$ and $S_p = \frac{\mathrm{FP}}{\mathrm{FP+TN}}$, where TP is the number of true positives (classified correctly as

---

[2] See [18] for details on the computation of the trifocal tensors.

[3] The estimated projection in the third view can be improved applying the point-line-point method proposed in [18, pp. 373].

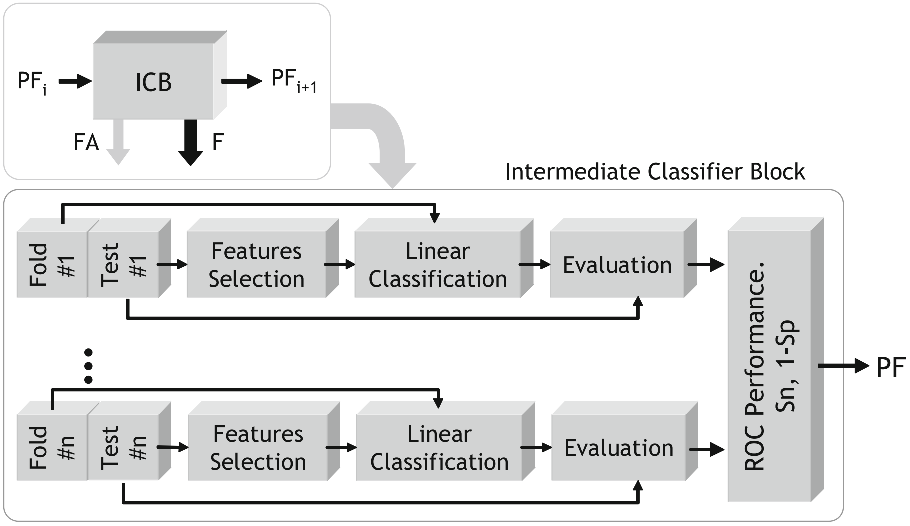


Fig. 6 General model of the internal process of the ICB linear classifier with cross validation selection and automatic feature selection by means of Fisher discriminant

flaws), TN is the number of true negatives (classified correctly as false alarms), FP is the number of false positives (classified incorrectly as flaws), and FN is the number of false negatives (classified incorrectly as false alarms). The objective is for the sensitivity to be maximum (100%) and simultaneously the 1-specificity to be minimum (0%), this way the classifier guarantees an ideal classification for two classes. In practice this is difficult to achieve because it depends on the classifier's internal parameters which can vary with respect to the noise existing in the data.

(**D.2**) **Selection of features** The features selected by the ICB classifier are determined automatically using the information contained in each potential flaw, each of which has an associated feature vector $\mathbf{v}$. Each feature vector is composed of twelve measures extracted previously in the segmentation step (see details in Table 1). Here we used the Take-L-Plus-R feature selection algorithm [12] to determine a combination of features that separate the classification space. The objective of this algorithm is to determine the best features that allow greater separation between the classes space.[4]

In the multiple views analysis above, we combined multiple flaws only by means of a geometric analysis. In this next step we carry out a fusion of its features to seek a separation of its classes. For instance, let $\mathbf{v}_a^i$ be a feature vector of flaw i in view a and let $\mathbf{v}_b^j$ be a feature vector of flaw j in view b. If these two regions are corresponding, in theory the distance of their features should be short, since both regions are the same in two views, otherwise if there are false alarms the distance between them should be longer. With this simple criterion we define a unique vector of features $\mathbf{v}_{ab}^{ij}$ that relates two regions for the bifocal case and $\mathbf{v}_{abc}^{ijk}$ which relates three regions in the trifocal case as

$$\begin{aligned} \mathbf{v}_{ab}^{ij} &= \left[(\mathbf{v}_a^i - \mathbf{v}_b^j)\right] \\ \mathbf{v}_{abc}^{ijk} &= \left[(\mathbf{v}_a^i - \mathbf{v}_b^j)(\mathbf{v}_a^i - \mathbf{v}_c^k)(\mathbf{v}_b^j - \mathbf{v}_c^k)\right] \end{aligned} \tag{13}$$

For the next analysis, and for simplicity, we define matrix $\mathbf{v}$ as $\mathbf{v} \equiv \mathbf{v}_{ab}^{ij} \equiv \mathbf{v}_{abc}^{ijk}$, because the feature vector contains the relations between two and three regions and both elimination processes are independent (see ICB process in Fig. 2).

(**D.3**) **Linear classification** We use a linear discriminative analysis (LDA) classification [12] that allows finding the hyperplanes that best separate the solution space. For that, the classification process must fit the following linear equation

$$\mathbf{w}^{\mathsf{T}} \cdot \mathbf{v} + w_0 > 0\,, \tag{14}$$

where $\mathbf{w} = \Sigma_{\mathrm{W}}^{-1} \cdot (\bar{v}_1 - \bar{v}_2)$ are the hyperplane parameters, Σ_{W} is the interclass covariance matrix, and $\mathbf{v}$ corresponds to the feature vector chosen earlier. Finally, the value w_0 for two classes is determined according to the mean of features $\bar{v}_1$ and $\bar{v}_2$ and the probabilities of each class p_{e1} and p_{e2} according to

$$\begin{aligned} w_0 = &-\frac{1}{2} \cdot (\bar{v}_1 + \bar{v}_2) \cdot \Sigma_{\mathrm{W}}^{-1} \cdot (\bar{v}_1 - \bar{v}_2) \\ &- \log\left(\frac{p_{e1}}{p_{e2}}\right) \end{aligned} \tag{15}$$

Once an initial solution is obtained for the vector $\mathbf{w}$, the optimization problem tries to fit the hyperplanes so that (16)

---

[4] As a criterion function, we used the Fisher discriminant [42].

is the maximum, this ensures that we are obtaining a high performance of (S_n) and $(1 - S_p)$ for each sub-selection of its features.

$$\{\mathbf{w}, w_0\} = \arg\max \{S_n(\mathbf{w}, w_0)\} \ s.t. \ S_p(\mathbf{w}, w_0) = 1 \tag{16}$$

This problem has been solved by the Nelder-Mead Simplex method [23]. Then the information from the selected straight lines and features is used to evaluate the performance of the classifier on the test data.

(D.4) **Joint classification** The linear discrimination analysis model has been used together with the cross-validation technique. This way the optimization process generated by each combination of features generates, as a result, a set of straight lines specific for each combination. Finally, this model is used for the testing data by the classifier, therefore, a set of weak classifiers makes it possible to generate a robust classification. For example, let us assume that we have used a set C consisting of three features in the training phase $C = \{V_1, V_2, V_3\}$. The separation between them generates a three-dimensional volume bounded by the cuts of the two-dimensional separations, containing only potential flaws (PF) (Fig. 7). This three-dimensional volume generated from the combination of the two-dimensional features $[V_1, V_2]$, $[V_1, V_3]$ and $[V_2, V_3]$ contains potential flaws. Conversely, the space outside the three-dimensional volume could be flaws or false alarms, depending on the position in which the hyperplanes are projected.

Our final classification method is based on the use of multiple linear separation models. The objective of the linear separation is to find a dividing line for two classes, but we use the same LDA algorithm with two purposes: First, to find the best separation line that minimizes the FPs subject to $S_n = 1$ defined as $\mathbf{s}^{\mathrm{FP}}$. Second, to find the best separation line that minimizes the FNs subject to a $S_p = 0$, defined as $\mathbf{s}^{\mathrm{FN}}$ (see slopes in Fig. 5b). These two separation lines generate a bi-dimensional separation space, and the total set of combinations of features generates a hyperplane.

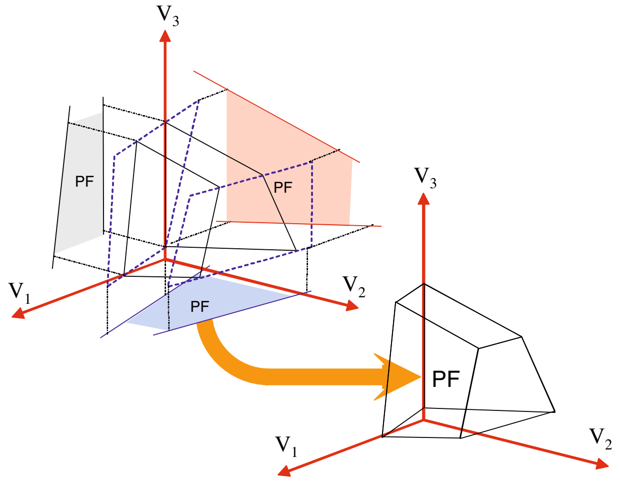


Fig. 7 Three dimensional representation of the ICB classification method

First, we calculate the separation line set in order to evaluate the joint classifier in the testing data for each combination pair. Thus, for the $\mathrm{V_m}$, $\mathrm{V_n}$ combination we generate a linear separation between the straight lines $\mathbf{s}^{\mathrm{FP}}_{m,n}$ and $\mathbf{s}^{\mathrm{FN}}_{n,n}$. More formally, let $\mathrm{PF}_{m,n}$, $\mathrm{FA}_{m,n}$ and $F_{m,n}$ be the space generated by the linear intersection between the features m and n, defined as

$$\begin{aligned}
\mathrm{PF}_{m,n} &= \left[\,\mathbf{v}\ 1\,\right] \cdot \mathbf{s}^{\mathrm{FP}}_{m,n} < 0 \wedge \left[\,\mathbf{v}\ 1\,\right] \cdot \mathbf{s}^{\mathrm{FN}}_{m,n} > 0 = \{0, 1\}\\
\mathrm{FA}_{m,n} &= \left[\,\mathbf{v}\ 1\,\right] \cdot \mathbf{s}^{\mathrm{FP}}_{m,n} < 0 \wedge \left[\,\mathbf{v}\ 1\,\right] \cdot \mathbf{s}^{\mathrm{FN}}_{m,n} < 0 = \{0, 1\}\\
F_{m,n} &= \left[\,\mathbf{v}\ 1\,\right] \cdot \mathbf{s}^{\mathrm{FP}}_{m,n} > 0 \wedge \left[\,\mathbf{v}\ 1\,\right] \cdot \mathbf{s}^{\mathrm{FN}}_{m,n} > 0 = \{0, 1\}
\end{aligned} \tag{17}$$

The next step is to verify the classification for each feature vector $\mathbf{v}$ with unknown values. To that end, let $N = C_2^P$ be the number of possible combinations of a feature vector P. For every vector of length P we generate N combinations of two features, therefore, the classification result requires a set of N results, and then by simple majority vote, the final classification is evaluated. Let us assume that $\mathbf{v_1}$ is the first feature vector and we want to find its classification. Its result for each class $\{\mathrm{PF}, \mathrm{FA}, F\}$ is defined by

$$\mathbf{M}(\mathbf{v_1}) = \begin{bmatrix} \mathrm{PF}_{1,2} & \mathrm{FA}_{1,2} & F_{1,2} \\ \vdots & \vdots & \vdots \\ \mathrm{PF}_{m,n} & \mathrm{FA}_{m,n} & F_{m,n} \\ \vdots & \vdots & \vdots \\ \mathrm{PF}_{N-1,N} & \mathrm{FA}_{N,N-1} & F_{N,N-1} \end{bmatrix}^{\top}_{N\times 3} \tag{18}$$

where the matrix $\mathbf{M}$ is the binary outcome of multiple partial classifications for the feature vector $\mathbf{v_1}$. Finally, the classification of the matrix $\mathbf{M}(\mathbf{v_1})$ is defined as

$$p(class|\mathbf{v}_1) = \max \frac{\sum_1^N \mathbf{M}(\mathbf{v_1})}{\sum_1^N \sum_1^3 \mathbf{M}(\mathbf{v_1})} \tag{19}$$

This process is carried out for each of the testing vectors. In our analysis, we consider the combination of two to seven features. This is because more than seven features turn the performance of the ICB down to zero, and, therefore, it is not possible to filter more false alarms. However this number can vary as a result of the linear classification inserted in each ICB.

4 Experimental results

This section presents the results of experiments carried out on a sequence of 72 radioscopic images of aluminum wheels (see some of them in Fig. 8). The dimensions of the wheel are 470 mm diameter and 200 mm height. The image size is 572 × 768 pixels with a dynamic range of 8 bits. There are 12 known real flaws in this sequence. Three of these flaws were detected by human visual inspection ($\emptyset = 2.0 \sim 7.5$ mm), however, the remaining nine flaws (small holes generated by a drill ($\emptyset = 2.0 \sim 4.0$ mm) in positions making their detection difficult) were not detected. A pattern of 1 mm in the middle of the wheel is projected in the X-ray projection coordinate as a pattern of 2.96 pixels in the image, i.e., the flaws are very small. In addition, since the signal-to-noise ratio in our radioscopic images is low, the flaw signal is slightly greater than the background noise, as illustrated in Fig. 8. In our experiments, the mean gray level of the flaw signal ranged from 2.4 to 28.8 gray values with a standard deviation of 6.1. Analyzing a homogeneous background in different areas of interest, we obtain noise within ±13 gray values with a standard deviation of 2.5. Due to the reasons stated above, the segmentation of real flaws with poor contrast can also involve detection of false alarms.

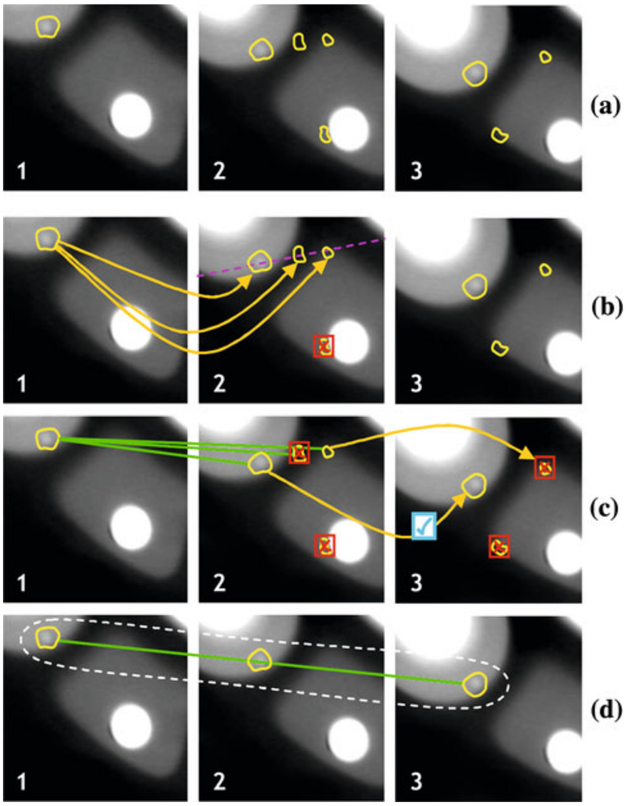


Fig. 8 Generalized flaw tracking process in one sequence of three views. **a** Identification of potential flaws; **b** searching for a match two views with the fundamental matrix; **c** searching for a match in three views with the trifocal tensor; **d** final detection, the false alarms are eliminated without discriminating real flaws

We separated the analysis into three steps. (1. Identification) Potential flaws are automatically identified in each image of the sequence using a single filter without a priori knowledge of the object structure (Fig. 8b). The result of the identification generates a data base that contains 424 registers with twelve features of the total potential flaws detected in the sequence. From them, 214 are real flaws, which correspond to the twelve real flaws mentioned above, and 210 registers are false alarms that must be reduced. (2. Tracking) In this step we separate the analysis into two phases: (a) the detection of pairs of flaws using the estimation of the fundamental matrix in two views, through the epipolar constraint; (b) using the previous results, we re-projected the pairs of potential flaws in the third view using the trifocal tensor estimation. (3. ICB method) Classifiers are inserted into two and three views to filter potential flaws between the views, according to the general model proposed in Fig. 2. All the phases are detailed below.

4.1 Performance with two views

The first phase is to assess the performance of the algorithm in two views using the bifocal method. This consists of determining the corresponding flaws between two images in a sequence through the search for flaws along the epipolar line (Fig. 8c). The results indicate that the model detects 100% of the real flaws that are corresponding in two views (Table 2, 2-Views Track). This validates the assumption of correspondence between the position of real flaws and implies that automatic detection with the fundamental matrix allows the detection of corresponding flaws that are contained on the epipolar line. There is, however, a large number of false alarms in the sequence (198/388 = 51%), which must be reduced using a third view.

4.2 Performance with three views

After completing the matching of possible pairs of flaws in both images, we extend the detection of flaws to the third image in the sequence (Fig. 8d). In this case the performance remains at 100% of real flaws detected in the sequence, however, it has not been possible to eliminate all false alarms (Table 2, Rate of False Alarms). Furthermore, the ICB method in two and three views has allowed the detection of a large part of the real flaws (F) and false alarms (FA) with high probability, allowing them to be separated from the multiple views analysis. Thus, in two views the reduction of false alarms with ICB-2 reaches 24.2%, and with ICB-3 it reaches 4.4%. These results indicate that the proposed method has generated a sustained reduction of potential flaws (PF) in the sequence.

Table 2 Performance of the Uncalibrated Tracking

Step	Flaws in sequence	False alarms in sequence	Rate of real flaws (%)	Rate of false alarms (%)
2-Views Track	190	198	100	51.0
ICB-2	151	94	100	24.2
3-Views Track	137	45	100	11.6
ICB-3	18	17	100	4.4

Table 3 Sensitivity and 1-specificity performance of the ICB classifier, and percentage of flaws reduction of ICB classifier

Views	Features	2	3	4	5	6	7
2 Views	S_n	92.7	95.6	98.4	100	100	100
	$1 - S_p$	2.4	0	0	0	0	0
	ICB reduction	52.3	49.5	43.6	17.3	12.8	10.3
3 Views	S_n	95.8	98	96	99.1	99.1	100
	$1 - S_p$	25	16.7	26.7	22.2	0	6.7
	ICB reduction	81.6	61.17	79.7	89.4	75.3	75.7

4.3 Performance of ICB

The results indicate a clear correlation between the performance of the ICB method for two views and the number of features chosen. By using the five best feature combinations, the performance of the classification is ideal, but there is a clear decrease in the number of false alarms extracted by the ICB method (Table 3, 2 Views). For example, in the case of combining the five best features, only 17.3% of the total potential flaws in correspondence are extracted, leaving 82.7% which is then transferred to the following matching block in three views. In the case of three views, the number of correspondences is drastically reduced because the correspondence of a false alarm in three images has a lower probability of occurrence. The highest performance results occur when the six best features are combined by means of the Take-L-Plus-R selection process included in each ICB (Fig. 6).

With respect to real-time capabilities, we tested our method in two ways, with and without ICB. First, with ICB, the required computation time to process a sequence of three images, was in average 38.3 s, with a remainder of 4.4% of false alarms. Second, without ICB, using the same sequence, the time was in average of 31.1 s,[5] with a remainder of 32.3% of false alarms. These results are very promising because the ICB method can filter the majority of false alarms in sequence, although it did require more time. We want to clarify that the ICB method follows a classification ensemble methodology to filter false alarms; without ICB each potential flaw is tracked in two and three views, thus the number of tests in multiple views is increased.

4.4 Comparison with other methods

Finally, we present a summary of the performance obtained with the calibrated and uncalibrated AMVI method (Table 4). For comparison purposes, we show the performances carried out with the same sequence of X-ray images designed in [29]. The given performances correspond to the 'true positives' and 'false positives'. The true positives are the percentages of flaws correctly detected in a sequence. The false positives (or false alarms) correspond to the percentage of 'non-flaws' that have been classified incorrectly as flaws. Current results indicate that it is possible to obtain 100% of the real flaws in a sequence detected correctly. These results have been generated in spite of the optical and geometric perturbations and the low SNR level that corresponds to X-ray images. However, false alarms remain in the sequence and reducing them has not been possible. Despite the false alarms, our method has achieved better performance than the system proposed by Pizarro et al. [36], mainly because it is not necessary to carry out matching of the potential flaws, only a tracking analysis. According to the results generated in two and three views in [4], the ICB technique has allowed a reduction of 8.7% in the correspondence number in two views, and of 5.5% in the case of three views with a 4.4% remainder, which has been impossible to eliminate so far by geometric analysis.

---

[5] The method was programmed in Matlab 7.0 under Windows XP SP2 on a Pentium Centrino Duo/2 GHz.

Table 4 Comparison between different calibrated and uncalibrated tracking techniques

Method	Images tracked	Year and reference	Analyzed images	True positives (%)	False positives (%)
Calibrated	3	2002 [29]	70	100	25
	4	2002 [29]	70	100	0
	5	2002 [29]	70	83	0
Uncalibrated	2	2005 [28]	24	92.3	10
	2	2006 [4]	70	100	32.9
	3	2006 [4]	70	98.8	9.9
	2	2008 [36]	70	86.7	14
	2	2008 $^{\text{new}}$	70	100	24.2
	3	2008 $^{\text{new}}$	70	100	4.4

5 Conclusions

Automated visual inspection remains an open question. Many research directions have been exploited, some very different principles have been adopted and a wide variety of algorithms have appeared in literature on automated visual inspection. Although there are several approaches in the last 30 years that have been developed, automated visual inspection systems still suffer from (1) detection accuracy, because there is a fundamental trade off between false alarms and missed detections; and (2) strong bottleneck derived from mechanical speed (required to place the test object in the desired positions) and (3) high computational cost (to determine whether the test object is defective or not). In this sense, AMVI offers a robust alternative method that uses redundant views to perform the inspection task. In this paper, we have developed a new flaw detection algorithm using an uncalibrated sequence of images. Using the new uncalibrated AMVI methodology, we have designed a novel system based only on the spatial positions of the structures. The proposed approach uses the projection of the epipolar line, generated by the fundamental matrix and the trifocal tensors in a robust manner, with the purpose of building a motion model without a priori knowledge of the object structure. The key idea of our strategy is to consider as false alarms those potential flaws that cannot be tracked in a sequence of multiple images. In this research, we have introduced the calculation of corresponding points generated artificially through the maximization of the correlation coefficient from two curves and the intermediate classifier block (ICB) method in order to filter false alarms. The method was tested in a sequence of X-ray images of aluminium wheels but the methodology can be applied to other sequences as well by changing the segmentation and control point algorithms as we demonstrated in bottle inspection system with multiple views in [6].

Our results indicate that it is possible to generate an automatic model for a sequence of images which represent the movement between the points and regions they contain. This way we can use as reference points the edges of the structures or areas, without loss of information, using a nonlinear method. The main advantage of our model is the automatic estimation of movement thus avoiding the calibration process. Our future aim is to reduce the number of false alarms by means of a method of final verification of the flaws in correspondence, and an analysis of the ICB classification method with other ensemble classification and probabilistic techniques. Another possibility is to change the Fisher discriminant into another linear dimensionality reduction (LDR) technique inside of each ICB. That way we can maximize the Chernoff distance based in [1]. The essential idea is to increase the distance defining linear class separability. This will allow the separation of more false alarms and flaws in each ICB. Thus, with less potential flaws in a sequence, the process will be faster.

Acknowledgments The author acknowledges the financial support from Escuela de Ingeniería, Pontificia Universidad Católica de Chile and from FONDECYT, Chile (Grant no. 1040210).

References

1. Ali, M., Rueda, L., Herrera, M.: On the performance of chernoff-distance-based linear dimensionality reduction techniques. In: Proceedings of the 19th Canadian Conference on Artificial Intelligence, pp. 469–480, Quebec, Canada (2006)
2. Boerner, H., Strecker, H.: Automated x-ray inspection of aluminum casting. IEEE Trans. Pattern Anal. Mach. Intell. **10**(1), 79–91 (1988)
3. Campbell, J.G., Fraley, C., Murtagh, F., Raftery, A.E.: Linear flaw detection in woven textiles using model-based clustering. Pattern Recognit. Lett. **18**(14), 1539–1548 (1997)
4. Carrasco, M., Mery, D.: Automated visual inspection using trifocal analysis in an uncalibrated sequence of images. Mater. Eval. **64**(9), 900–906 (2006)
5. Carrasco, M., Mery, D.: Automatic multiple visual inspection on non-calibrated image sequence with intermediate classifier block. In: Pacific-Rim Symposium on Image and Video Technology (PSIVT'07), pp. 371–384. Springer, Berlin (2007)

6. Carrasco, M., Pizarro, L., Mery, D.: Image acquisition and automated inspection of wine bottlenecks by tracking in multiple views. In: Proceedings of 8th International Conference on Signal Processing, Computational Geometry and Artificial Vision—ISCGAV'08, pp. 84–89. Rhodes Island, Greece (2008)
7. Chen, Z., Wu, C., Shen, P., Liu, Y., Quan, L.: A robust algorithhm to estimate the fundamental matrix. Pattern Recognit. Lett. **21**, 851–861 (2000)
8. Chin, R.T., Harlow, C.A.: Automated visual inspection: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **4**(6), 557–573 (1982)
9. Cohen, I., Ayache, N., Sulger, P.: Tracking points on deformable objects using curvature information. In: Proceedings of the 2nd European Conference in Computer Vision, pp. 458–466 (1992)
10. Drury, C.: Inspection Performance. pp. 2282–2314. Wiley, New York (1992)
11. Drury, C.G., Saran, M., Schultz, J.: Effect of fatigue/vigilance/environment on inspectors performing fluorescent penetrant and/or magnetic particle inspection. Interim Report 03-G-012, University at Buffalo, Federal Aviation Administration William J. Hughes Technical Center (2004)
12. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd edn. Wiley, New York (2001)
13. Dunn O.J., Clark V.A.: Applied Statistics: Analysis of Variance and Regression. Wiley, New York (1974)
14. Egan, J.: Signal Detection Theory and ROC Analysis. Academic Press, New York (1975)
15. Filbert, D., Klatte, R., Heinrich, W., Purshke, M.: Computer aided inspection of castings. In: IEEE-IAS Annual Meeting, pp. 1087–1095. Atlanta, USA (1987)
16. Gdalyahu, Y., Weinshall, D.: Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. IEEE Trans. Pattern Anal. Mach. Intell. **21**(12), 1312–1328 (1999)
17. Haralick, R.M., Shapiro, L.G.: Computer and Robot Vision. Addison-Wesley, New York (1992)
18. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, 1st edn. Cambridge University Press, Cambridge (2000)
19. Jacob, R., Raina, S., Regunath, S., Subramanian, R., Gramopadhye, A.K.: Improving inspector's performance and reducing errors—general aviation inspection training systems (gaits). In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting Proceedings. Aerospace Systems, Human Factors and Ergonomics Society (2004)
20. Jarvis, J.F.: Visual inspection automation. Computer **13**(5), 32–38 (1980)
21. Kita, Y., Highnam, R., Brady, M.: Correspondence between different view breast X-rays using curved epipolar lines. Comput. Vis. Underst. **83**(1), 38–56 (2001)
22. Kumar, A.: Computer-vision-based fabric defect detection: a survey. IEEE Trans. Ind. Electron. **55**(1), 348–363 (2008)
23. Lagarias, J., Reeds, J.A., Wright, M.H., Wright, P.E.: Convergence properties of the nelder-mead simplex method in low dimensions. SIAM J. Optim. **9**(1), 112–147 (1998)
24. Liu, H., Srinath, M.: Partial shape classification using contour matching in distance transformation. IEEE Trans. Pattern Anal. Mach. Intell. **12**(11), 1072–1079 (1990)
25. Malamas, E., Petrakis, E.G., Zervakis, M.: A survey on industrial vision systems, applications and tools. Image Vis. Comput. **21**(2), 171–188 (2003)
26. Mery, D.: Crossing line profile: a new approach to detecting defects in aluminium castings. Lect. Notes Comput. Sci. **2749**, 725–732 (2003)
27. Mery, D.: High contrast pixels: a new feature for defect detection in x-ray testing. Insight **46**(12), 751–753 (2006)
28. Mery, D., Carrasco, M.: Automated multiple view inspection based on uncalibrated image sequence. Lect. Notes Comput. Sci. **3540**, 1238–1247 (2005)
29. Mery, D., Filbert, D.: Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence. IEEE Trans. Robot. Autom. **18**(6), 890–901 (2002)
30. Mery, D., Filbert, D.: Classification of potential defects in automated inspection of aluminium castings using statistical pattern recognition. In: Proceedings of 8th European Conference on Non-Destructive Testing (ECNDT 2002), Barcelona, Spain (2002)
31. Mery, D., da Silva, R.R., Calôba, L.P., Rebello, J.M.A.: Patter recognition in the automatic inspection of aluminium castings. Insight **45**(7), 475–483 (2003)
32. Mital, A., Govindaraju, M., Subramani, B.: A comparison between manual and hybrid methods in parts inspections, vol. 9, issue 6. Integrated Manufacturing Systems, MCB UP Ltd, Bradford (1998)
33. Newman, T.S., Jain, A.K.: A survey of automated visual inspection. Comput. Vis. Image Underst. **61**(2), 231–262 (1995)
34. Nguyen, V.D., Noble, A., Mundy, J., Janning, J., Ross, J.: Exhaustive detection of manufacturing flaws as abnormalities. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 945–952 (1998)
35. Pedreschi, F., Mery, D., Mendoza, F., Aguilera, J.: Classification of potato chips using pattern recognition. J. Food Sci. **69**(6), E264–E270 (2004)
36. Pizarro, L., Mery, D., Delpiano, R., Carrasco, M.: Robust automated multiple view inspection. Pattern Anal. Appl. **11**(1), 21–32 (2008)
37. Polikar, R.: Ensemble systems in decision making. IEEE Circuits Syst. Mag. **6**(3), 21–45 (2006)
38. Sebastian, T., Klein, P., Kimia, B.: On aligning curves. IEEE Trans. Pattern Anal. Mach. Intell. **25**(1), 116–125 (2003)
39. Sonka, M., Hlavac, V., Boyle, R. (eds.): Image Processing, Analysis and Machine Vision, 2nd edn. PWS Publishing, Pacific Grove (1999)
40. Spencer, F.: Visual inspection research project report on benchmark inspections. Technical Report DOT/FAA/AR-96/65, U.S. Department of Transportation, Federal Aviation Administration, Washington, DC, Office of Aviation Research Washington, D.C. 20591 (1996)
41. Spicer, P., Bohl, K., Abramovich, G., Barhak, J.: Robust calibration of a reconfigurable camera array for machine vision inspection (RAMVI): using rule-based colour recognition. In: Proceedings of the 1st International Conference on Computer Vision Theory and Applications Ultrasonics Symposium (VISAPP), pp. 131–138, Setúbal, Portugal (2006)
42. Stearns, S.: On selecting features for patterns classifiers. In: Proceedings of IAPR International Conference on Pattern Recognition, pp. 71–75 (1976)
43. Umeyama, S.: Parameterized point pattern matching and its application to recognition of object families. IEEE Trans. Pattern Anal. Mach. Intell. **15**(2), 136–144 (1993)