Image Segmentation of Vickers Indentations Using Shape from Focus

Michael Gadermayr and Andreas Uhl

Multimedia Signal Processing and Security Lab
Department of Computer Sciences, University of Salzburg, Austria
{mgadermayr,uhl}@cosy.sbg.ac.at

Abstract. To measure the hardness of a material, an indenter is pressed into the material and the deformation is measured. As we focus on Vickers hardness testing, our exercise is to compute the diagonal lengths of a square indentation. We especially investigate if it is possible to reconstruct the shape of the indentation by the use of the Shape-from-Focus method. We show that the shape information alone does not contain enough information for a robust segmentation. However, we incorporate the depth information into an effective existing approach and achieve significantly better results.

1 Introduction

The Vickers indentation images which should be segmented, approximately fit the following description: The object has a square geometry and is darker colored than the background. The diagonals are aligned horizontally and vertically. Figure 1 shows example images. Whereas the first image is quite perfect, the others suffer from noise and/or low contrast.

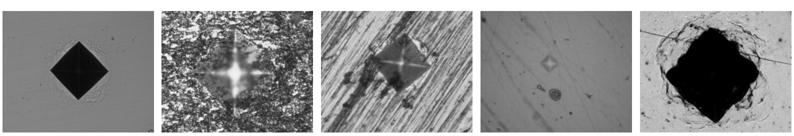


Fig. 1. Example images of Vickers indentations

There are several proposals for automated image segmentation of Vickers indentations. In [1], an exact and precise multi-resolution active contours approach has been introduced. Other methods [2,3] are based on template matching.

The contribution of this work is the proposal of an inclusion of the depth information gathered by the Shape-from-Focus [4] method (applied in [5]) into an existing approach [1]. Whereas the shape information is not robust enough to act as the single feature for a segmentation, in addition with traditional features a gain in performance can be achieved.

This paper is structured in to following way: In Sect. 2 the Shape-from-Focus method is explained and examples are shown. In Sect. 3 a localization method

A. Campilho and M. Kamel (Eds.): ICIAR 2012, Part I, LNCS 7324, pp. 149–157, 2012.

and an exact segmentation method [1] are explained and adjusted to additionally deal with shape information. In Sect. 4, the results are explained and compared with an existing approach. Section 5 concludes this paper.

2 Shape from Focus

Straightforward image processing usually only relies on one image, which has to be segmented. All the information must be gathered from this single two dimensional signal. However, the real world cannot be described by two dimensions, as space has got three. To overcome this inadequacy, a more general approach [4] relies not only on one (focused) image, but on a set of images, with different focus setups. This introduces a kind of depth information, although the signal is sampled with a simple 2D-camera. We investigate this method with reference to Vickers indentation images.

Focus can only be achieved for a specifically defined region. That means, it is not possible to focus an object in the foreground as well as the background simultaneously. Consequently, if pictures with different focus setups of a three dimensional object are taken, information of the third dimension can be obtained, as in the different images, different regions are focused. To estimate the depth of an image point, the image with the highest focus measure (i.e. the point is in focus) at this point has to be evaluated. In the referenced approach, the shape of visibly rough surfaces is determined. As focus can only be measured if differences of pixel values occur, a smooth surface cannot be segmented in this way. To resolve this restriction, an illumination strategy which creates artificial high frequency signals has been proposed [4]. As our hardware should stay unchanged this detail has not been investigated.

2.1 Shape Computation

First of all, to compute the shape of an object, a series of images I_k of an object with different focus levels $k \in L$ must be gathered (L is the set of focus levels). After that, for each point $v = (x, y)$ in each image I_k of a focus series, a focus measure $F(v)$ must be computed. Next, for each point v the focus level $k \in L$ with the highest focus measure $F_k(v)$ is calculated. Each focus level k represents a defined depth level d:

$$d(v) = k \in L : \forall l \in L : F_k(v) \geq F_l(v) \ . \tag{1}$$

Although the depth is not measured absolutely, relative differences are sufficient to determine peaks and valleys of the surface. Whereas this method produces regions of the same discrete level, an approach that generates a smooth shape has been additionally suggested [4]. As our application does not allow a perfect reconstruction of the shape (images contain regions without high frequency) and a smoother approximation is computationally more expensive, we content ourself with the simpler discrete version.

2.2 Focus Measures

As mentioned, a focus measure F is necessary to determine the focus level k with the highest response. In the original publication [4], the authors proposed to use the Sum-modified-Laplacian (SML) operator F_{SML} which is based on the second derivation:

$$F_{SML}(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} ML(x,y) \quad if\ ML(x,y) \geq T\ . \tag{2}$$

$$\begin{aligned} ML(x,y) = & |2 \cdot I(x,y) - I_(x-s,y) - I(x+s,y)| + \\ & |2 \cdot I(x,y) - I(x,y-s) - I(x,y+s)|\ . \end{aligned} \tag{3}$$

s is the step size of the metric, which can be adjusted according to the image properties. The SML operator not only consists of a simple gradient operator. To increase the robustness, a threshold T is introduced, which suppresses very small responses. Moreover some neighboring pixel responses are summed up to achieve a more steady output (adjustable with N).

Alternatively, we investigated a generalization of the Tenengrad focus measure F_T that is based on the first order derivation, which could be used instead of the proposed SML measure.

$$F_T(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} T(x,y) \quad if\ T(x,y) \geq T\ . \tag{4}$$

$T(i,j) = S_x^{*2}(i,j) + S_y^{*2}(i,j)$ and S_x^* and S_y^* are convolutions of the Sobel operators in x and y direction with the image.

Moreover, we investigated the Range metric, which is based on the histogram. As we need a metric for each point separately and not for the whole image, the region r of a defined region surrounding the point is regarded:

$$r(i,j) = \{(x,y)|(|x-i| + |y-j|) \leq T\}\ . \tag{5}$$

T defines the size of the region. The metric is calculated in the following way:

$$F_{range}(i,j) = max(r(i,j)) - min(r(i,j))\ . \tag{6}$$

2.3 Analysis on Indentation Images

In the case of Vickers indentations, the images are containing imprints, which definitely are three dimensional (the middle of the object is farthest away whereas the background is closer to the camera). The constraint of a visibly rough surface often applies (i.e. high frequency, noisy images, low contrast), especially in the case of low quality images which are likely to fail traditional segmentation processes. We decided to focus majorly on low quality images, with high frequencies on the one hand (necessary for this approach) and low contrast (hard to segment for traditional methods) on the other hand.

Figure 2 shows the effect of different focus settings on the resulting image. Whereas in the right image the deepest part of the imprint, in the left image the

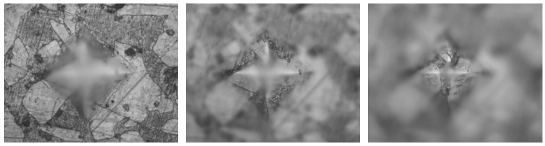


Fig. 2. From left to right the focus starts nearby (background in focus) and ends far away (imprint in focus)

background (shallow part) is focused. Images with the focused region closer to the camera than the background of the image, do not represent useful information, as none of the image points are focused. We decided to select 8 images with different focus levels (focused points farther away) including the focused image. These selected images are processed by the Shape-from-Focus method explained in Sect. 2.1.

Figure 3 shows examples of generated depth information (the left image corresponds with the indentation in Fig. 2). Whereas bright regions (except white regions) have been determined to be nearby the camera, dark ones are expected to be far away. In the case of a weak response (no high frequencies available), the determination is uncertain. Regions, which cannot be reliably determined (threshold T of the focus measure is not reached) because of their missing high frequency information, are marked in white.

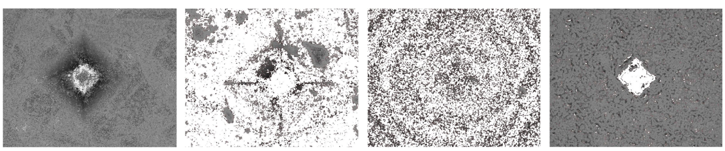


Fig. 3. Reliable (left) and unreliable determinations of depth (others) SML focus metric: N=1, T=7, s=3

3 Active Contours Segmentation Methods

The Shape-from-Focus approach does not segment the images, but extracts additional information. We still have to segment the images, because we need the exact coordinates of the corner points.

As shown in Fig. 3, some indentations can be determined quite exactly by the Shape-from-Focus method, whereas others definitely cannot. Consequently, by regarding the shape information, some images can be segmented, but others cannot be segmented.

One option is, to estimate which images are suitable for the Shape-from-Focus approach and which are not. The unsuitable images can be treated by traditional methods [1,2,3]. The suitable ones' shape information must be extracted from the different focus setups and this information again can be treated by the traditional methods. We did not investigate this option so far.

The second option is, not to discriminate the images, but provide both informations (gray value and depth) to the segmentation method. Unfortunately, most segmentation approaches are based on gray (or edge) information alone. However, in [1] two approaches are discussed, which are based on feature vectors of theoretically arbitrary length. Whereas one of the approaches is used to approximately localize the contour, the other one is used for an exact segmentation. The methods are discussed in the following subsections.

3.1 Approximative Shape Prior Gradient Descent Method

This robust Vickers indentation localization method has been introduced in [1]. 4 parameters (scaling, rotation, shift x, shift y) of a square shape are evolved by a gradient descent algorithm which minimizes an energy criterion. The best fitting parameters are assumed to be these with the lowest energy. Different energy criteria were introduced (edge based, region based, statistical). The method is able to robustly localize the indentations, but it is not able to segment the object exactly because the indentations are no perfect squares.

We focus on the statistical energy criterion, which allows the inclusion of arbitrary feature vectors for each point in the image. The energy is given by E:

$$E = -\int_{\Gamma_{in}} \log(p_{in}(f(v)))dv - \int_{\Gamma_{out}} \log(p_{out}(f(v)))dv \ . \tag{7}$$

Γ_{in} and Γ_{out} are the regions inside and outside of the contour and p_{in} (p_{out}) is the discrete probability density function of the features inside (outside) of the contour. In [1], the authors propose the following feature vector (∇ is the gradient operator, I is the image gray value):

$$f(v) = (I(v), ||\nabla I(v)||) \ . \tag{8}$$

Now we investigate the following feature vectors:

$$f(v) = (I(v), depth_{focus}(v)) \ . \tag{9}$$

$$f(v) = (I(v), ||\nabla I(v)||, depth_{focus}(v)) \ . \tag{10}$$

3.2 The Statistical Level Set Method

For an exact segmentation, a level set method [6] is used. A contour is defined by it's level set and is also evolved by minimizing an energy criterion. Whereas the Shape Prior method only allows to adjust the parameters of a square shape, the level set method allows the evolution of an arbitrary shape. Consequently, a more accurate segmentation can be achieved, but the method requires an approximate localization as initialization. So we used the Shape Prior gradient descent method as an approximative stage and the level set method as a second stage to achieve exact results. For the level set method we also used a statistical energy criterion [7] which allows arbitrary feature vectors:

$$E = -\int_{\Gamma_{in}} \log p_{in}(f(v))dv - \int_{\Gamma_{out}} \log p_{out}(f(v))dv + \alpha|C| \ . \tag{11}$$

$\alpha|C|$ is the contour regularization term which consists of a weighting term α an the length of the contour $|C|$. We investigate this method, as it allows to include the shape information. The region based method used in [1] is not able to deal with the additional information, but only relies on gray values. The feature vectors in equation (9) and (10) are used in the energy criterion (equation (11)).

4 Experiments

We compare the traditional segmentation methods without depth information [1] with the newly introduced versions with depth from focus information.

Our database consists of 25 different indentations. For each indentation, 8 differently focused images are used. Many of the images are hard to segment by traditional methods because of noise and/or low contrast. We investigate if the additional depth information improves the segmentation results.

4.1 Statistical Shape Prior Method

The best results are achieved with the 2 dimensional feature vector in equation (9). The distributions p_{in} and p_{out} are calculated by convolving the empirical distributions with a Gaussian Parcen window ($\sigma = 2$).

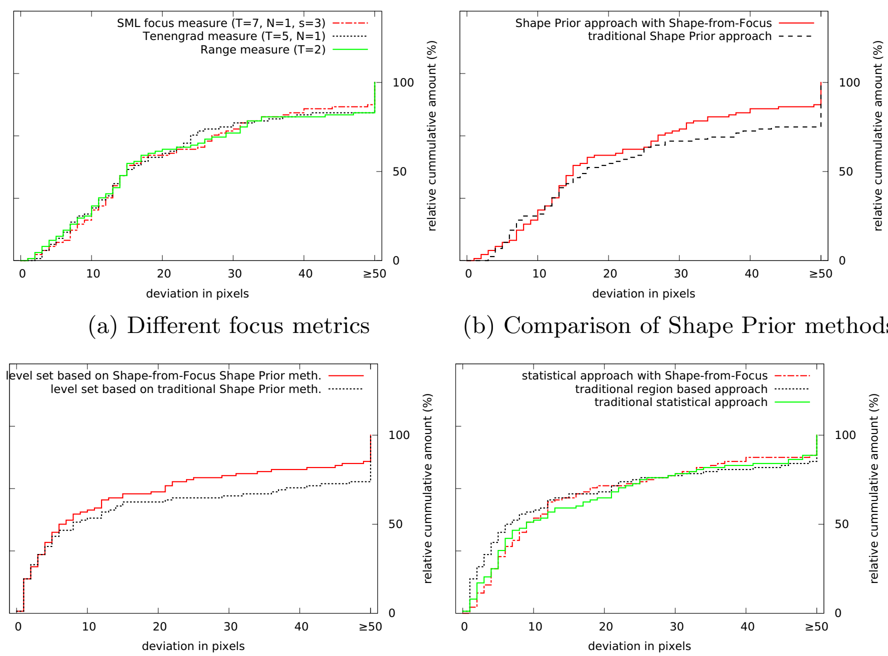
The choice of the focus measure is not very decisive as the results are quite similar. The achieved segmentation performance with the different measures is shown in Fig. 4a. The (cumulative) curves represent the number of indentation corner-points, which are detected within ($\leq$) a given distance. The earlier (lower deviation) the line rises, the more accurate the segmentation is. The Shape-from-Focus method in the following experiments is based on the SML focus measure ($T = 7$, $N = 1, s = 3$) which is slightly more competitive as far as outliers are concerned.

In Fig. 4b the results of the statistical method with depth information is compared with the statistical method without depth information [1]. Whereas the number of edges detected quite exactly (0-25 pixels) is similar, the number of outliers (deviation $\geq$ 50 pixels) can be decreased significantly with the Shape-from-Focus information. As we concentrate on a low outliers ratio (i.e. a high degree of robustness) in the localization stage, the achieved results seem to be more appropriate as initialization for the exact segmentation stage.

4.2 Statistical Level Set Method

First of all, in Fig. 4c the impact of the initialization on the traditional region based level set segmentation [1] is shown. If the method is initialized with the results achieved with the Shape Prior method including the depth information, the results are clearly superior (red line). The depth information used by the Shape Prior approach definitely increases the segmentation performance.

Now we compare the statistical level set approach including shape information with the traditional region based level set method (as before) and the statistical level set approach without shape information. The methods are initialized with the results achieved with the Shape Prior approach including the depth information (Section 3.1). The results are shown in Fig. 4d. In contrast to the approximative Shape Prior approach, the results of the precise level set segmentation approach are more similar. The approach including the depth information (Section 3.2, red line) seems to be slightly more robust than the region based approach (regarding deviations of e.g. ≤ 40 pixels). However, the region based approach tends to be more accurate (regarding deviations of e.g. ≤ 5 pixels). The statistical approach without the depth information (based on gray value and edge information) tends to be in the middle of the other mentioned approaches.



(a) Different focus metrics

(b) Comparison of Shape Prior methods

(c) Level set: Impact of initialization

(d) Comparison of level set methods

Fig. 4. Experimental results

To understand this behavior, we watched the different depth images provided by the depth from focus method (Fig. 3). Images with low noise and high contrast, which can be segmented well without any depth information often have bad depth information (large regions without depth information (marked white)). Consequently this images suffer from the additional information. In opposite the depth information of highly noisy images usually is quite accurate, so the segmentation performance increases. Such an image is shown in Fig. 5. In this image, the addition of depth information leads to a successful segmentation.

Consequently, we only recommend to use the Shape-from-Focus approach if the image is hard to segment by the traditional algorithms (e.g. the multi-resolution approach introduced in [1]). To automatically determine the quality of images, a measure would be required. This has not been investigated so far.

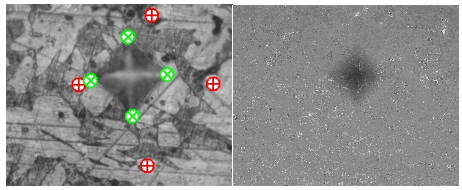


Fig. 5. Achieved corner points with traditional method (⊕) and the proposed approach (⊗) (left) and the corresponding depth information (right)

In Table 1, the execution runtimes of the Shape-from-Focus method are compared with the traditional ones. The algorithms are not optimized for execution speed. As especially the computation of the shape information is computational expensive, only low quality images should be segmented with the Shape from Focus method.

Table 1. Execution runtimes (Intel Core 2 Duo T5500 1.66 GHz)

Process	Traditional [1]	Shape-from-Focus
Shape-from-Focus computation	(not required) 0.0 s	7.3 s
Shape Prior approach	2.2 s	2.2 s
Level set approach	2.1 s	4.2 s
Average total runtime	4.3 s	13.7 s

5 Conclusion

The addition of the Shape-from-Focus information in the existing approximative Shape Prior method definitely is advantageous. The use of the new information in the existing precise level set segmentation methods only is advantageous for images which are hard to segment traditionally. In contrast, the performance on images which can be segmented exactly with traditional methods decreases.

Acknowledgment. This work has been partially supported by the Austrian Federal Ministry for Transport, Innovation and Technology (FFG Bridge 2 project no. 822682).

References

1. Gadermayr, M., Uhl, A.: Dual-resolution active contours segmentation of Vickers indentation images. In: Proceedings of the 5th International Conference on Image and Signal Processing, ICISP 2012 (2012)

2. Gadermayr, M., Maier, A., Uhl, A.: Algorithms for microindentation measurement in automated Vickers hardness testing. In: Tenth International Conference on Quality Control for Artificial Vision (QCAV 2011). Proceedings of SPIE, vol. 8000, pp. 80000M–1–80000M–40. SPIE, St. Etienne (2011)
3. Maier, A., Uhl, A.: Robust automatic indentation localisation and size approximation for vickers microindentation hardness indentations. In: Proceedings of the 7th International Symposium on Image and Signal Processing (ISPA 2011), Dubrovnik, Croatia, pp. 295–300 (September 2011)
4. Nayar, S.K.: Shape from focus system. In: Computer vision and pattern recognition, Proceedings CVPR 1992, pp. 302–308 (1992)
5. Harada, T.: Robust method for position measurement of vertex of polyhedron using shape from focus. Journal of Advanced Mechanical Design, Systems and Manufacturing 4(2), 492–503 (2010)
6. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. Journal of Computational Physics 79(1), 12–49 (1988)
7. Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. International Journal of Computer Vision 72(2), 195–215 (2006)