Active Contours Methods with Respect to Vickers Indentations

M. Gadermayr[a], A. Maier[a], and A. Uhl[a]

[a]University of Salzburg, Department of Computer Sciences, Salzburg, Austria

ABSTRACT

We investigate different Vickers indentation segmentation methods and especially concentrate on active contours approaches. In order to circumvent the initialization problem of active contours, we introduce a localization stage. Furthermore, traditional active contours based segmentation approaches are incorporated with the Shape from Focus approach to improve robustness. In order to decrease the overall runtime, moreover, a gradual enhancement approach based on unfocused images is introduced. With three different databases, we compare the proposed methods and we show that the segmentation performances of these methods are highly competitive compared with other approaches in literature.

1. INTRODUCTION

Hardness, which is an important characteristic of solid materials, can be determined with the Vickers Hardness test. A pyramidal indenter is pressed into the material with a defined force and causes an indentation. An important issue is to measure the size (diagonal length) of the approximately square indentation to determine the Vickers Hardness[1] (depends on the applied force and the diagonal length). As the manual measurement of the indentation images not only is expensive, but also interpretive and subjective, a robust and accurate automatized measurement method is highly beneficial. Segmentation algorithms are used to get the positions of the four vertices of the indentation. Having the four vertices, the diagonal lengths can be calculated easily.

There are several proposals for image segmentation of Vickers indentations. One group of algorithms rely on wavelet analysis.[2,3] These methods assume that object borders are perfectly straight lines, which is not always true. Another approach[4] is based on edge detection followed by Hough transform and least squares approximation of lines. Edge finding techniques are based on the assumption that high differences between neighboring pixels imply that these pixels are part of the border. Especially in case of noisy images, this assumption is not right at all. The method introduced in[5] applies thresholding followed by a Hough transform. If thresholding is applied to images, it is necessary to calculate a threshold value which depends on the current image, because a fixed threshold surely does lead to good results. Often a separation of the object from the background with thresholding is not possible. Other suggested methods also binarise the image using thresholding,[6,7] followed up

by morphological closing. Another approach is based on axis projection and Hough transform.[8] This approach is based on the assumption that the objects are perfectly aligned (diagonals vertical/horizontal). Methods relying on template matching ([9,10]) are quite robust to noise. This is because big templates suppress noise as large regions are summed up. The template matching approach introduced in[10] provides robust as well as precise results, but requires an accurate alignment (diagonals horizontal/vertical). A high degree of accuracy is achieved by applying four corner templates instead of one complete square template. In[9] square templates in different sizes and different rotations are matched with the image. The results of this approach are robust but only serve as approximations. In,[11] a refinement strategy is introduced.

This work deals with different kinds of active contours algorithms. The aim of these approaches is that a contour with a defined initialization converges to the real object boundaries. The contour is forced by different energy functionals, which depend on pixel values of the image on the one hand and homogeneity criteria of the contour on the other hand. As this method suffers from bad initializations, we introduce a Shape Prior approach, which produces robust, but not accurate results. These results can deal as initializations for an active contour. Moreover, we incorporate the active contours with the Shape from Focus approach, which extracts additional information of the image, that could increase the segmentation performance. Furthermore, we propose a gradual enhancement approach based on different unfocused images in order to decrease the overall runtime and also increase robustness.

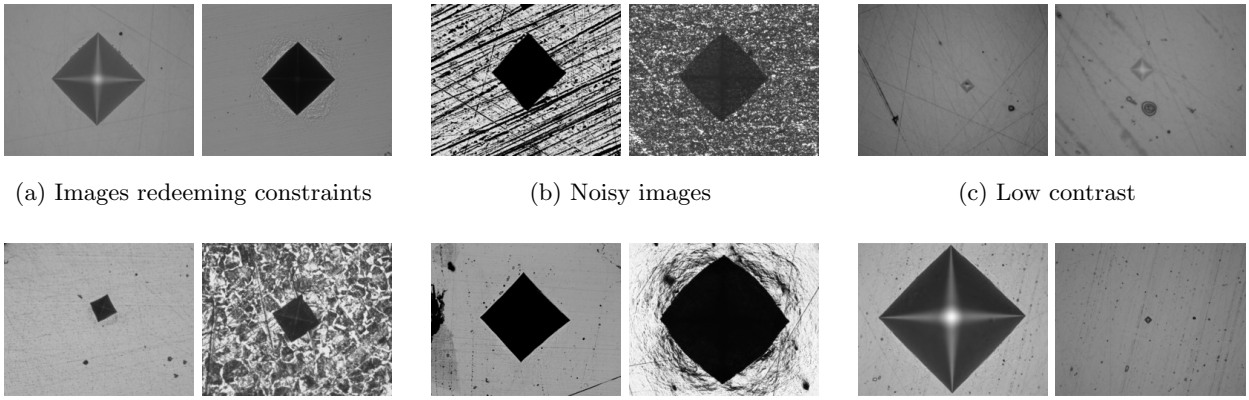
In section 2, different active contours approaches are investigated with reference to Vickers indentation images. In section 3, a new approximative Shape Prior segmentation approach is proposed to achieve good initializations. In section 4, the Shape from Focus approach[12] is incorporated with traditional segmentation methods, in order to improve robustness. In section 5, a new gradual enhancement approach is introduced. Experiments are shown in section 6. Section 7 concludes this paper.

1.1 Vickers images

The images to segment approximately fit the following description:

- square geometry of object to segment
- dark object, bright background
- diagonals are horizontally and vertically aligned
- object is situated close to the center

Figure 1a shows quite perfect images. Images in Fig. 1b show different kinds of noise. Some bright images lack of contrast (Fig. 1c), especially the diagonals' gray scale is the same as the background. Another inconvenience is the fact, that in such images, possible noise is often darker than the imprint. In Fig. 1d, the diagonals are not aligned horizontally and vertically. In Fig. 1e, a concave and a convex curvature is shown. The images in Fig. 1f represent the smallest and largest imprints in our databases.



(a) Images redeeming constraints (b) Noisy images (c) Low contrast

(d) Violation of rotation constraint (e) Shape violation (f) Different sizes

Figure 1: Different categories of Vickers indentation images

2. ACTIVE CONTOURS APPROACHES

2.1 State of the Art Methods

The traditional active contours (snake) model has been introduced in.[13] A snake is a closed curve which iteratively converges at the object's borders, by means of gradient descent of an energy functional. The curve is represented by a sequence of pairs containing x and y coordinates. connected with straight line segments.

Traditional active contours is an attempt to minimize the cost (energy) function E_s which depends on image pixels and the contour's shape by the use of gradient descent. The energy function consists of an additive weighted mixture of internal energy, image energy and external energy. The weights can be adjusted in order to control the behavior of the snake. A snake is represented parametrically by $v(s) = (x(s), y(s))$, where $x(s)$ is the x coordinate and $y(s)$ is the y coordinate for a given element s of the contour.

$$E_s = \int_{Snake} E_{intern}(v(s)) + E_{image}(v(s)) + E_{extern}(v(s))ds \; . \tag{1}$$

The internal energy E_{intern} consists of a linear combination of continuity (first order derivation) and curvature (second order derivation). The continuity term prevents the marker points $v(s)$ from converging to few points

with low image energy (i.e. the contour acts like a membrane). The curvature term prevents the contour from developing sharp corners (i.e. the contour acts like a thin plate).

The image energy E_{image} is computed from image pixels. Usually the image energy is determined by applying an edge operator to the original image. Pixels near edges having a higher gradient represent a lower energy. The idea is that edges are likely to be part of the object's border. λ is a weighting parameter, ∇ is the gradient operator, $||\cdot||$ is the Euclidean norm and I is the image gray value.

$$E_{image} = \lambda \cdot E_{edge} = -\lambda \cdot ||\nabla I(v(s))|| \; . \tag{2}$$

External forces like user interaction might be also regarded (E_{extern}), but this is not further investigated, as hardness image segmentation should run as batch process without any human interaction at all.

The level set method introduced in,[14] is an alternative to the traditional snake model. Apart from other inconveniences, traditional active contours suffer from an explicit parametrization (by frontier points) of the contour. In the level set formulation an explicit parametrization is circumvented by using an intrinsic formulation. The evolving contour is given by its level set Γ.

$$\Gamma = \{(x, y) | \phi(x, y) = 0\} \; . \tag{3}$$

$\phi(x, y)$ is a function which is 1 inside, -1 outside of the region and exactly 0 at the frontier of the evolved shape. The points outside (inside) of the contour are given by Γ_{out} (Γ_{in}):

$$\Gamma_{out} = \{(x, y) | \phi(x, y) < 0\} \; . \tag{4}$$

$$\Gamma_{in} = \{(x, y) | \phi(x, y) > 0\} \; . \tag{5}$$

Evolution of the frontier happens by moving the level set Γ in normal direction $\frac{\nabla \phi}{||\nabla \phi||}$ with a specified speed v. Although the general curve evolution is defined, the level set approach is highly flexible. There exist lots of different ways of calculating the speed function v, which influences the behavior of the evolving level set.

One commonly used edge based level set variation has been introduced in.[15] As with the snake approach, this formulation is based on edge information.

As with the snake approach, the edge based level set approach requires the propagation of edges to increase the capture range (region where the contour converges to the real boundary). To bypass this issues, in[16] a region based approach has been introduced, where the force of the contour is not based on image gradients.

This method is based on the assumption that the object's surface as well as the surface outside of the object are homogeneous as far as its gray value is concerned. The energy term is given in the following equation:

$$E_{CV} = \int_{\Gamma_{in}} (I(v) - \bar{I_{in}})^2 dv + \int_{\Gamma_{out}} (I(v) - \bar{I_{out}})^2 dv + \lambda \int_{\Gamma} ||\nabla \phi(v)|| dv \ . \tag{6}$$

Γ is the image, I is the image gray value, $\bar{I_{in}}$ ($\bar{I_{out}}$) is the average image value inside (outside) of the contour, ∇ is the gradient operator and λ is the curvature weighting term.

The region based approach is based on the assumption that images do not necessarily have strong gradients at their boundaries, but the regions inside and outside of the contour have to be homogeneous as far as the gray value is concerned. This assumption usually is quite appropriate, however it is inappropriate for some kinds of images (e.g. the image in Fig. 2: background consists of regions which are darker than the object).

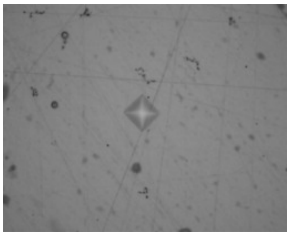


Figure 2: Indentation image with inhomogeneous background

Consequently, a region based model would state that such dark noise pixels are more likely to be part of the object than to be part of background (as background average color is brighter). Obviously, this does not match with reality. To overcome that inconvenience, a statistical approach has been introduced:[17]

$$E_{stat} = -\int_{\Gamma_{in}} \log p_{in}(f(v)) dv - \int_{\Gamma_{out}} \log p_{out}(f(v)) dv + \alpha |C| \ . \tag{7}$$

$\alpha|C|$ is the regularization term. p_{in} and p_{out} are the probabilities of feature vectors f inside and outside of the contour. Intuitively, the energy is low, if both regions are homogeneous with respect to the feature vectors f.

This formulation allows not only to use the gray value as feature, but each feature which could be defined for a specific pixel might be included in an arbitrary feature vector. The following feature vectors are used:

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

$$f(v) = (I(v), \frac{dI(v)}{dx}, \frac{dI(v)}{dy}) \ . \tag{9}$$

2.2 Vulnerability to Initialization

First we discuss the edge based approaches (traditional active contours and edge based level set methods). These approaches are based on the gradients of the image. That means, the contour moves into the direction, where

the gradients are increasing. Close to the object boundaries, this actually is appropriate. However, we do not already know where the objects are located approximately. In Fig. 3 (left), the gradient image and a possible initial contour is shown. Surely, a gradient descent of the contour would never be successful as the capture range is to small. Especially the segmentation of noisy images suffers, because the gradient image shows lots of regions with low image energy (Fig. 3, right), where the contour potentially converges to.

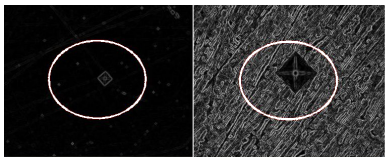


Figure 3: Problems: small capture range (left), noise (right)

To increase the capture range and decrease the effect of noise, large edge operators instead of small ones (Sobel 3x3) are investigated. This causes a propagation of the edge information and thus makes a gradient descent of the contour possible. However, the size of the gradient operator is limited as small objects are blurred too much and moreover, the computational costs become high. The effect is shown in Fig. 4. However, although a big gradient operator is used in the right image, the segmentation fails. In,[18,19] strategies to increase the

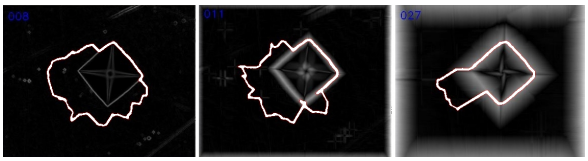


Figure 4: Different gradient operators

capture range are proposed. One big problem of these methods is, that not only edge information is propagated, but also noise. Noisy edge images affect the active contours as shown in Fig. 5.

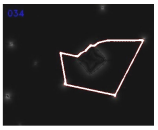


Figure 5: Segmentation failed because of noisy edge image

Edge based algorithms have to be initialized nearby the boundary. Although region based approaches (and the statistical approach) which do not rely on gradients, are known to be less vulnerable to the initial configuration, starting with a general level set is not advantageous. On the one hand, even region based algorithms do not succeed to converge at the desired boundary, if the contour is too far away from the indentation. The contour often does not shrink to converge at the boundaries, but grows instead. This is because the average gray value

of the background might be darker than the average gray value within of the contour. Such an initialization is shown in Fig. 6. On the other hand if such long distances must be managed, computational costs are tremendous (runtime $> 10s$).

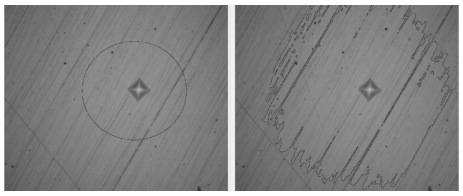


Figure 6: Region based approach fails (left: initialization, right: wrong segmentation)

3. AN APPROXIMATIVE SHAPE-PRIOR METHOD

The methods mentioned so far suffer from converging to local minima if the initialization is inappropriate. In existing Shape Prior level set approaches,[20,21] a weighted shape term is added to the energy function. This increases the robustness, but for a standalone segmentation even these methods are inappropriate. We propose a quite different way for robust segmentation of shapes which are known a priori:[22]

Our approach requires a parametric description of the prior shape. The object that will be segmented will have exactly the prior shape, as not a contour (parametrized by points or level set), but the object description parameters are directly evolved by gradient means of descent. Whereas the traditional active contours as well as level set algorithms allow arbitrary deformation of the initial contour, our approach only allows the evaluation of the following four parameters (the effect on the contour is shown in Fig. 7).

1. r_0: scaling
2. x_0: translation x axis
3. y_0: translation y axis
4. α_0: rotation

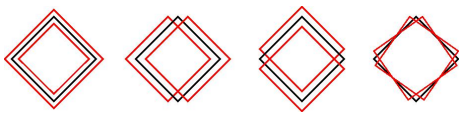


Figure 7: The four degrees of freedom

The contour of the square is given by the points (x, y) with the distance $d = r_0$ to a center (x_0, y_0). d is calculated in the following way, to ensure that the evolved contour has a square shape:

$$\begin{aligned} d = &|(x - x_0) \cdot cos(\alpha) + (y - y_0) \cdot sin(\alpha)| + \\ &|(x - x_0) \cdot sin(\alpha) - (y - y_0) \cdot cos(\alpha)| \ . \end{aligned} \tag{10}$$

Of course, this algorithm will not be able to segment Vickers images perfectly, as Vickers' shape often cannot be described by a perfect square. Though this is not our objective, but we aim at a pretty good approximative segmentation of a very high rate of images. These approximations deal as the initializations for a precise segmentation method.

The regions in- and outside of the square are given by Γ_{in} and Γ_{out}:

$$\Gamma_{in} = \{(x, y) : d < r_0\}. \tag{11}$$

$$\Gamma_{out} = \{(x, y) : d > r_0\}. \tag{12}$$

As with the level-set approach, we define an energy criterion which is minimized by gradient descent. We investigated different energy functions (edge based, region based). Tests showed that the following statistical criterion, which is derived from the statistical level set approach proposed in,[17] is the best choice:

$$E = -\int_{\Gamma_{in}} \log(p_{\Gamma_{in}}(f(v)))dv - \int_{\Gamma_{out}} \log(p_{\Gamma_{out}}(f(v)))dv \ . \tag{13}$$

$f(v)$ is an arbitrary feature of the point v. We investigate the feature vectors in equation (8) and (9). The evolved parameters are collected in the vectors $s_i = (x_0, y_0, r_0, \alpha)$. The vector s_0 is the initialization. s_{n+1} is defined recursively:

$$s_{n+1} = s_n + \lambda(\nabla E) \ . \tag{14}$$

λ which usually is a multiplicative component, is called step size. To allow λ to act as a signum function (one pixel left, stay, one pixel right), which can deal with numerical issues, it is more generally defined as function:

$$\lambda((x_1, ..., x_n)^T) = (sign(x_1), ..., sign(x_n))^T \ . \tag{15}$$

$$\nabla E = (\frac{dE}{dx}, \frac{dE}{dy}, \frac{dE}{dr}, \frac{dE}{d\alpha})^T \ . \tag{16}$$

e.g. the partial derivative of the x dimension is calculated as:

$$\frac{dE}{dx}((x, y, r, \alpha)^T) = E((x + 1, y, r, \alpha)^T) - E((x - 1, y, r, \alpha)^T) \ . \tag{17}$$

Although the introduced approach is already able to deal with local minima caused by noise, we still have not achieved a total invariance to the initialization s_0. Local minima still prevent from a proper localization of the indentation in several cases. The balloon approach[23] introduced for active contours, deals with this problem by adding an energy term, forcing the contour to become smaller or larger. Our approach allows to apply a kind of balloon force in an easy but effective

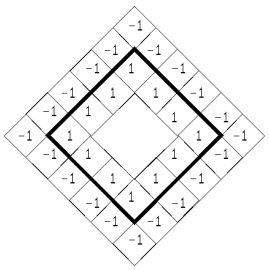


Figure 8: Directed edge template

way: Instead of calculating the radius r_0 by gradient descent, r_0 is simply decreased by one in each iteration of the gradient descent. If the contour starts at the image boundaries (r_0 is large), it necessarily has to cross the object's boundaries, when getting smaller and smaller.

Unlike unforced gradient descent, the proposed balloon-method does not stop before r_0 becomes zero (or a defined minimum). In a second step the history of the gradient descent has to be analyzed, to get the best fitting vector s_{res} from a set of several local minima. In our case the best results are achieved when using the vector s_i with the highest response (achieved by convolution) of the image information to the template (parametrized by s_i) shown in Fig. 8 with a thickness of 3 pixels.

4. INCLUDING SHAPE FROM FOCUS

The approaches mentioned so far only rely 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, the Shape from Focus approach[12] uses a set of images, with different focus setups (as shown in Fig. 10) to generate a depth map.

A focus can only be achieved for a specific 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 (depth) 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 order to measure the focus, high frequency information (i.e. differences between neighboring pixels) is necessary.

4.1 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:

$$d(v) = k \in L : \forall l \in L : F_k(v) \geq F_l(v) \; . \tag{18}$$

Although the depth is not measured absolutely, relative differences are sufficient to determine peaks and valleys of the surface.

4.2 Focus Measures

As mentioned, a focus measure F is necessary to determine the focus level k with the highest response. In[12] the Sum-modified-Laplacian (SML) operator F_{SML} is proposed, which is based on the second order derivation:

$$F_{SML}(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} ML(x,y) \quad if\ ML(x,y) \geq T\ . \tag{19}$$

$$ML(x,y) = |2 \cdot I(x,y) - I(x-s,y) - I(x+s,y)| + |2 \cdot I(x,y) - I(x,y-s) - I(,y+s)|\ . \tag{20}$$

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 investigate 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.

$$F_T(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} T(x,y) \quad if\ T(x,y) \geq T\ . \tag{21}$$

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

Moreover, we investigate the Range metric, which is based on the histogram:

$$F_{range}(i,j) = max(r(i,j)) - min(r(i,j))\ . \tag{22}$$

$$r(i,j) = \{(x,y)|(|x-i| + |y-j|) \leq T\}\ . \tag{23}$$

T defines the size of the regarded region.

4.3 Introduction of Shape from Focus Knowledge into Segmentation

As shown in Fig. 9, the depth estimation based on the Shape from Focus approach often is not able to appropriately separate the indentation from the background. The depth in regions marked white cannot be reliably determined (i.e. region is homogeneous). Otherwise, the darker the region, the farther it is away. Especially in case of high quality images, the shape information often is very unreliable, whereas in case of (noisy) low quality images, the shape information seems to be more reliable than the original image. Consequently, we incorporate the active contours and the Shape Prior method with the Shape from Focus approach.[24] We concentrate on the statistical energy criteria (equation (6), (13)), which allow (unlike the other criteria) the inclusion of arbitrary

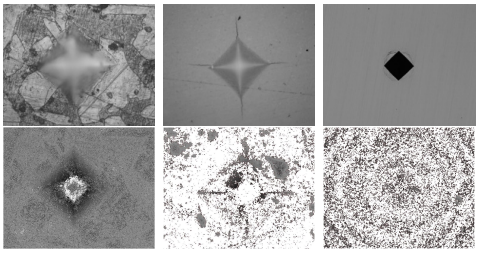


Figure 9: Original images (top) and corresponding depth informations (bottom): Reliable (left) and unreliable (middle, right) depth information

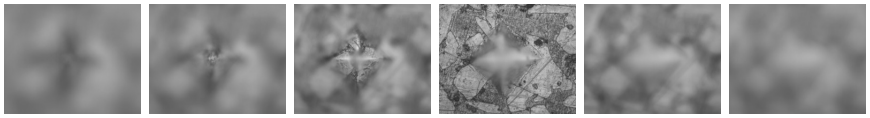


Figure 10: Different focus settings, reaching from $fl << 0$ (left) to $fl >> 0$ (right)

feature vectors. Now we investigate the following feature vectors, instead of the traditional ones in equations 8 and 9:

$$f(v) = (I(v), depth_{focus}(v)) \; . \tag{24}$$

$$f(v) = (I(v), ||\nabla I(v)||, depth_{focus}(v)) \; . \tag{25}$$

5. GRADUAL ENHANCEMENT WITH UNFOCUSED IMAGES

The Shape from Focus method reconstructs the shape (depth information) of surfaces by regarding differently focused images of the same surface.

The following two facts motivate us to propose one more approach dealing with differently focused images:

- Actually, with appropriately unfocused images the segmentation robustness can even be improved (see section 6). This is because the unfocused images are not just similar to low pass filtered images, but in addition the indentation is reinforced if the focus plane is chosen appropriately (see Fig. 10, left images).
- Furthermore, the autofocus system takes a significant amount of time, to provide the focused image. To get the focused image, our system starts with a focus plane farther away than the surface and incrementally gets nearer.

Active contours incrementally converge at the object boundaries. That means, the image might be changed during the segmentation process. Our intention is, to start with unfocused images, which can be segmented robustly. Incrementally, when a "better" image is available, the segmentation approach continues to process the new image in order to refine the results. The focused image is defined to have the focus level zero ($fl = 0$). If

the focus plane is farther away from the camera than the specimen, the focus level is smaller than zero and vice versa. The chosen step size between two focus levels is explained in section 6.

We propose the following 3 steps based on the Shape Prior gradient descent approach:

1. The focus starting setting is chosen that the focus plane is farther away than any part of the specimen ($fl << 0$).
2. Start the proposed first stage gradient descent segmentation algorithm on the unfocused image which is taken with the mentioned focus setting. Approximative results are achieved.
3. Until the end-criterion is reached:
   - Increase the focus level by one step and get the image.
   - Initialize the gradient descent algorithm with the current approximative results and the new image.
   - Increase the initialization variable "radius" by 2 pixels. As a balloon force is used, otherwise the contour could only shrink.
   - Run the algorithm with only 5 iterations to enhance the approximative results.
   - New approximative results are achieved.

The first image to segment is highly unfocused. Consequently, an exact segmentation surely cannot be achieved. However, the blurred image can be segmented robustly. Whereas the first image is segmented as proposed in section 3, the enhanced images are not. These images are initialized according to the current approximative results and only 5 iterations of the gradient descent approach are applied. The proposed policy allows to start the segmentation before the final focused image is available.

5.1 Appropriate Endcriterion

The intention is, that the results could be enhanced until the focused image is reached. Actually, this is not true. The best results are achieved, when stopping with the image of a focus level below zero. In practice, this is not possible, as the focus levels are defined relatively to the focused image. However, when saving the result history, these results can be recovered.

5.2 Speeding up the initial segmentation

Whereas the enhancement steps (3) are fast, the initial step (2) takes quite a long time. As the initial contour starts at the boundary of the image (initial radius is about 50 pixels as the images are downscaled by factor 10),

has to shrink until it collapses and shrinks one pixel per iteration, about 50 iterations are necessary. Whereas a further reduction of the image size affects the segmentation accuracy, increasing the step size of the contour does not, as far as robustness is concerned. Instead of modifying the evolving shape parameters by one per iteration, we propose to increase the step size (i.e. in one iteration, each parameter is adjusted by the positive or negative step size or stays the same). Increasing the step size to 4, we achieved less accurate results after the initial segmentation step, but after the enhancement steps, the results were exactly the same (the results are shown in Sect. 6).

6. EXPERIMENTS

6.1 Database

For testing, two different databases with 150 (DB1) and 216 images (DB2) respectively (resolution 1280x1024 pixels) provided to us were used. These images have been evaluated manually by four people independently. The ground truth was determined by taking the mean of all four measures. Moreover, one more database (DB3) consisting of 25 indentation series, each consisting of 40 images with different focus settings is used to evaluate the Shape from Focus and the Gradual enhancement approach. The quality of these images is considerably lower.

Our aim is to detect the four vertices of the approximately square Vickers indentations. In the following analyses, the distances between detected vertices and the ground truth are measured. In these Figures, for each deviation bin (Euclidean distance in pixels) on the x-axis, the number of vertices detected within the deviation is shown on the y-axis.

6.2 Two Stage Approach without Shape from Focus

The following strategy turned out to be competitive as far as segmentation performance is concerned:

- Stage 1 - Localization: Approximative segmentation with the Shape-Prior algorithm (section 3) on down-scaled images (factor 10).
- Stage 2 - Refinement: The precise region based level set method based on the results (as initializations) of the localization stage.
- For a further improvement of the segmentation accuracy, the local Hough transform[4,8] is applied in an area of 60 pixels around the corners as a postprocessing strategy.

The approximative Shape Prior localization stage already is highly robust, as the number of outliers (e.g. deviation > 50 pixels) is low, as can be seen in Fig. 11a and 11b. The region based refinement stage based on the Shape Prior results but without a Hough transform, is able to improve the accuracy (e.g. deviations ≤ 5 pixels). With the local Hough transform, the accuracy can be improved further more.

In Fig. 11c and 11d, different energy criteria of the refinement stage are compared. The region based criterion is the best choice. The edge based criterion is very vulnerable even with good initializations, as small edge operators are used. In this case, the traditional active contours approach (snake) is used instead of the level set approach. The edge based level set approach is more competitive, but not as competitive as the statistical or the region based level set approach.

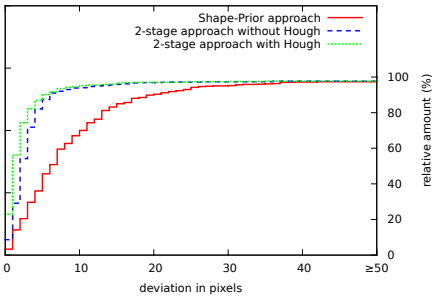
In Fig. 11e and 11f, we compare the proposed 2-stage active contours approach with two existing template matching approaches.[10,11] In the high accuracy range ($0 - 1$ pixels) our 2-stage approach is the best choice. As far as deviations of about 3 to 8 pixels are concerned, the 3-stage template matching approach[11] is slightly more competitive. However, when regarding higher deviations, the corner template matching approach[10] and the proposed method are more reliable. Over the whole range, the proposed 2-stage method seems to be the best alternative. The deduced knowledge seems to be reliable, as both databases show a similar behavior.

6.3 Shape from Focus

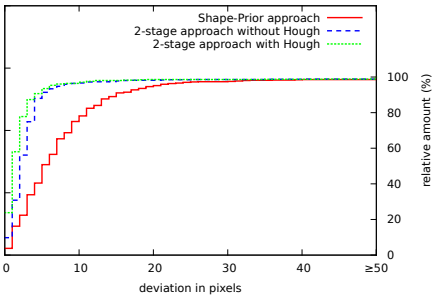
We compare the traditional proposed 2 stage approach without depth information with the versions with depth information. As the Shape from Focus approach requires series of images with different focus levels, for the experiments the image database DB3 is chosen. We decided to utilize only 7 images per indentation with the focus plane farther away than the specimen and the focused image.

First we investigate the impact of the depth information on the first approximative stage. The best results are achieved with the 2 dimensional feature vector (equation (24)). The distributions p_{in} and p_{out} are calculated by convolving the empirical distributions with a Gaussian Parcen window ($\sigma = 2$).

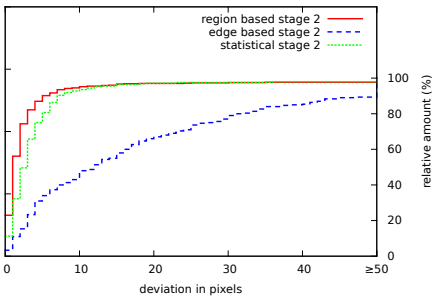
The choice of the focus measure is not very decisive as the results are quite similar. The achieved segmentation performance with the different measures is shown in Fig. 12a. The Shape-from-Focus method in the following experiments is based on the SML focus measure ($T = 7$, $N = 1, s = 3$) which is slightly more competitive as far as outliers are concerned. In Fig. 12b the results of the statistical method with depth information (equation (24)) is compared with the statistical method without depth information (equation (8)). Whereas the number of edges detected quite exactly (0-25 pixels) is similar, the number of outliers (deviation ≥ 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



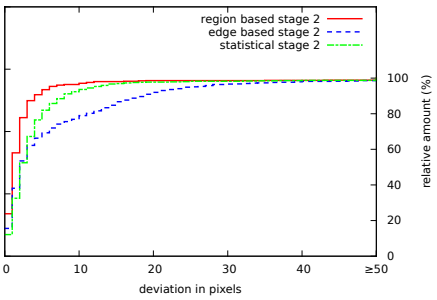
(a) DB1, with and without refinements



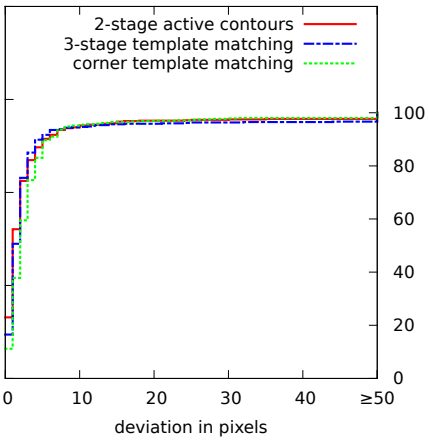
(b) DB2, with and without refinements



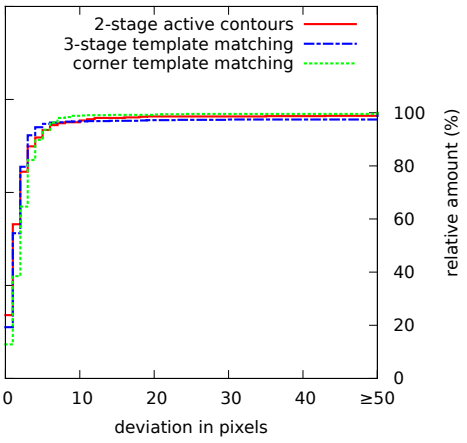
(c) DB1, different energy criteria in stage 2



(d) DB2, different energy criteria in stage 2



(e) DB1, Comparison with approaches in literature



(f) DB2, Comparison with approaches in literature

Figure 11: Comparisons

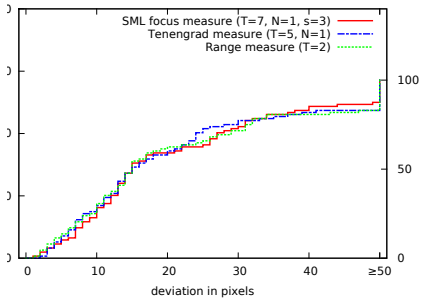
of robustness) in the localization stage, the achieved results seem to be more appropriate as initialization for the exact segmentation stage.

Now we investigate the impact of the shape information on the second stage. First of all, in Fig. 12c the impact of the initialization on the traditional region based level set segmentation is shown. If the method is initialized with the results achieved with the Shape Prior method including the depth information, the results are superior. The depth information used by the Shape Prior approach definitely increases the segmentation performance as far as robustness is concerned.

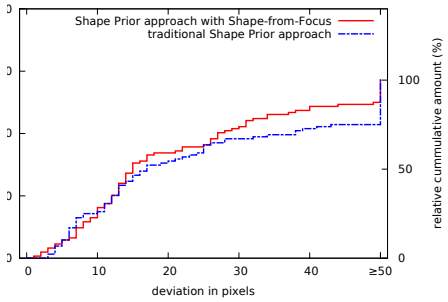
In Fig. 12d, we compare our methods from Fig. 12c with the corner template matching approach[10] and the 3-stage template matching approach.[11] The corner template matching approach definitely suffers from the low quality of the images. The robustness highly declines. The 3-stage template matching approach is slightly more accurate than the 2-stage active contours approach applied to the low quality images. However, the active contours approach additionally based on the shape information delivers even better results than the 3-stage approach, especially as far as outliers are concerned.

Now we compare the statistical level set approach including shape information (best results are achieved with 3 dimensional feature vector in equation (25) and the empirical distributions) with the traditional region based level set method[11] and the statistical level set approach without shape information (best results with equation (9)). The methods are initialized with the results achieved with the Shape Prior approach including the depth information. The results are shown in Fig. 12e. 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 seems to be slightly more robust than the region based approach (regarding deviations of e.g. $\leq$ 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 tends to be in the middle of the other mentioned approaches.

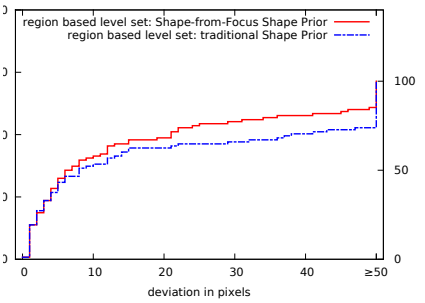
To understand this behavior, we regard the different depth images provided by the depth from focus method (Fig. 9). Images with low noise and high contrast, which can be segmented well without any depth information (middle and right image) often have bad depth information (large regions without reliable depth information (marked white)). Consequently, the segmentation of these images suffer from the additional information. In opposite, the depth information of highly noisy images (left image) usually is quite accurate, so the segmentation performance increases. Such an image is also shown in Fig. 13. In this image, the addition of depth information leads to a successful segmentation. Consequently, we only recommend to use the second stage Shape-from-Focus approach only if the image is very hard to segment by the traditional algorithms as the execution runtime



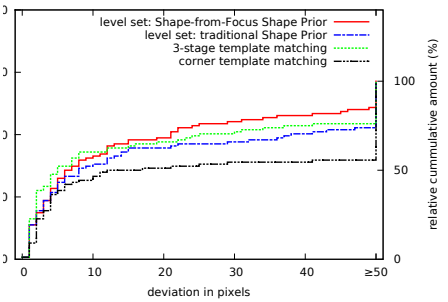
(a) Different focus metrics



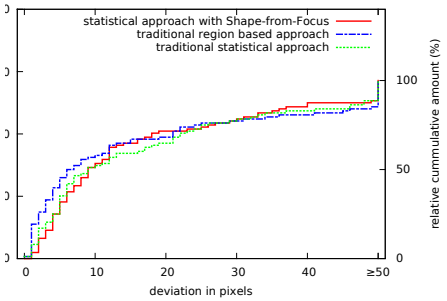
(b) Comparison of stage 1 methods



(c) Stage 2: Impact of initialization



(d) Comparison with methods in literature[10, 11]



(e) Stage 2 with Shape from Focus

Figure 12: Comparisons

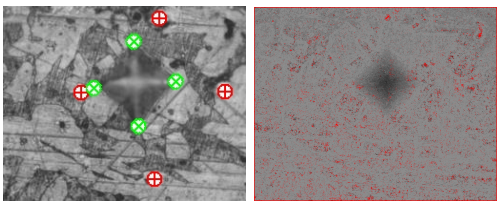


Figure 13: Achieved corner points with traditional method ($\oplus$) and Shape from Focus based approach ($\otimes$) (left) and the corresponding depth information (right)

increases considerably (section 6.5) if the second stage is based on shape information. In the first approximative stage the additional computational costs are low and the performance significantly raises.

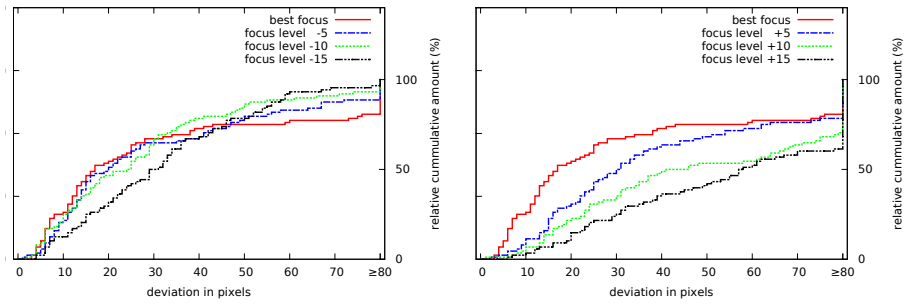
6.4 Gradual Enhancement with Unfocused Images

As this approach also requires differently focused image, DB3 is chosen with the low quality images. The step size between two focus levels is chosen dependently from the optical zoom of the camera (see Table 1). E.g. with a magnification of 10, one focus level less means that the camera with a fixed-focus lens moves 10,000 nm towards the specimen.

Table 1: Focus step size dependent on the optical zoom factor

zoom factor	step size
10 x	10,000 nm
20 x	5,000 nm
40 x	1,000 nm

First of all, we investigate the effect of single unfocused images instead of focused ones on the proposed approximative Shape Prior indentation segmentation approach. Figures 14a and 14b show results of the approximative methods. In Fig. 14a, the focus plane is farther away from the camera compared with the best-focus setting. The robustness (i.e. the number of outliers is low) of the segmentation did not only stay unchanged, but can actually be increased The segmentation accuracy (e.g. the ratio of vertices with a deviation of maximal 20 pixels) slightly decreases. In opposite, if the focus plane is nearer to the camera compared with the best-focus setting (Fig. 14b), the accuracy and the robustness significantly decreases.

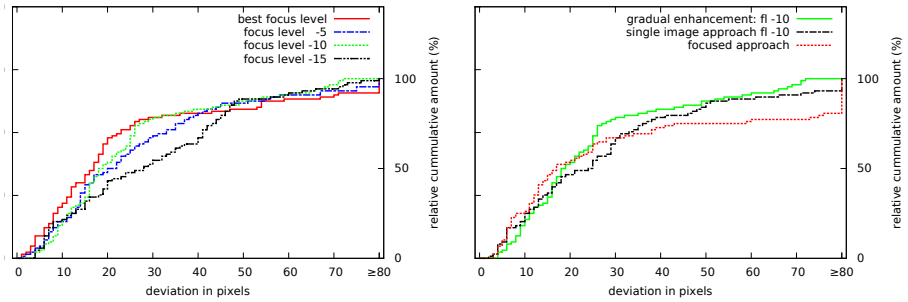


(a) Focal point farther away Shape Prior approach

(b) Focal point nearer[22]

Figure 14: Single unfocused images: Advantageous settings (2 approaches)

Now we regard the proposed gradual enhancement approach: The results with different focus levels as stopping



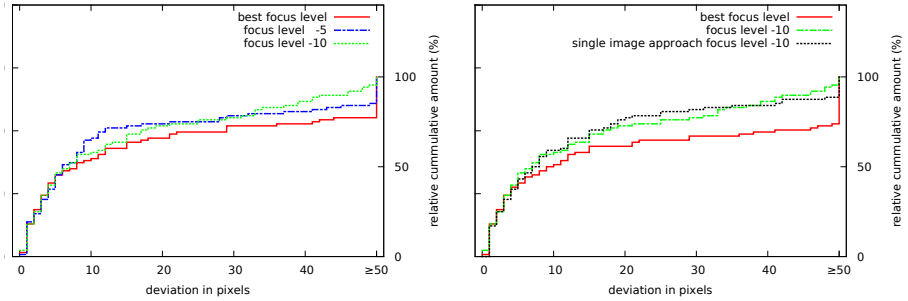
(a) Different stopping levels: first stage (b) Best stopping level vs. focused approach

Figure 15: Gradual enhancement approach

conditions are shown in Fig. 15a. Although the behavior is similar to the behavior with one single unfocused image (best stopping level: $fl = -10$), the effect is smaller. The outliers ratio generally is lower than with the single image approach. In Fig. 15b, the gradual enhancement approach with the best stopping focus level ($fl = -10$) is compared with the best results achieved with one single (unfocused) image and with the results with the focused image. The gradual enhancement approach definitely is more competitive as far as the approximative stage is concerned than the best focus approach and even more robust (less outliers) than the single unfocused image approach.

The results seem to be more similar compared with the single image approach. However, the impact of the different initialization results, on the level set algorithm is considerable, as shown in Fig. 16a. Especially the number of outliers can considerably be decreased when stopping earlier ($fl = -10$). The slightly lower accuracy of the first stage can be compensated by the second stage. Consequently, we define the stop level -10 to be the best choice. In Fig. 16b the achieved results of best configurations (gradual enhancement and single unfocused image) are compared with the traditional approach with focused images. The performance of the methods using unfocused images definitely are higher than the performance of the simple approach dealing with the focused image. The gradual enhancement approach is even slightly more robust (very few outliers) than the single unfocused image approach.

In Fig. 17, the Shape from Focus and the gradual enhancement approach are compared with the traditional method. Only the first stage varies. The second stage is always the same (region based level set approach). As shown, with the gradual enhancement approach in the first stage, the robustness of the overall segmentation can be improved once again.



(a) Effect of init.: level set algorithm (b) Effect of init.: level set algorithm

Figure 16: Gradual enhancement approach: Effect on the precise stage

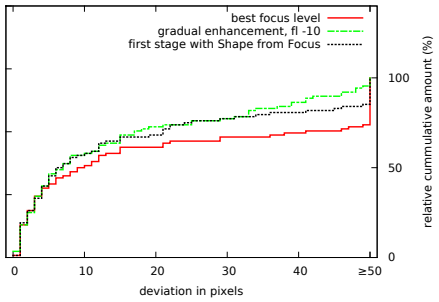


Figure 17: Overall comparison after second stage

6.5 Runtimes

In Table 2, the average runtimes of the different approaches are shown. The traditional 2-stage approach takes about 4.3 seconds per image. If the Shape from Focus approach is included in the first stage, the overall runtime slightly increases, as the shape information must be computed. If the Shape from Focus approach is included in both stages, the runtime significantly raises, as the shape information must be computed from the originally sized images, which is computational expensive. If the gradual enhancement approach based on unfocused images is used, the traditional Shape Prior approach (2.2 seconds per image) is replaced by the initial step which takes

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

Process	traditional 2-stage approach	stage 1 with Shape from Focus	both stages with Shape from Focus
Shape Prior approach	2.2 s	2.2 s	2.2 s
Level set approach	2.1 s	2.1 s	4.2 s
Shape computation	(not required)	0.8 s	7.3 s
Average total runtime	4.3 s	5.1 s	13.7 s

1.0 seconds and a number of enhancement steps (0.14 seconds per step). The runtimes cannot be compared, as with this approach the segmentation could start earlier.

7. CONCLUSION

We showed, that active contours are a precise segmentation approach especially if a good localization of the indentation is already available. The introduced Shape Prior approach provides very robust approximative results, which are used as initialization for the active contours. On the high quality databases, the introduced 2-stage approach produces better results, than existing approaches. When incorporating the mentioned approaches with Shape-from-Focus information, robustness can even more be improved. Especially the localization stage benefits from the additional information. On the low quality database, the traditional 2-stage active contours approach is slightly less competitive than an existing high performing template matching based method. However, with the shape information, the 2-stage active contours approach is even more competitive than this existing approach. Similar results are achieved with the gradual enhancement approach. Nevertheless, the additional advantage of this method is that the overall execution runtime potentially can be reduced.

REFERENCES

[1] ASTM Standard E384, (2010e2), "Standard test method for Knoop and Vickers hardness of materials," in [*ASTM Standards*], ASTM International, West Conshohocken, PA (2010). DOI: 10.1520/E0384-10E02, www.astm.org.

[2] Liming, W., Qu, Z., Yaohua, D., and Miaoxian, Z., "Automatically analyzing the impress image of vickers hardness test using wavelet," *China Mechanics Engineering* **15** (March 2006).

[3] Qu, Z., Guozheng, Y., and Yi, Z., "A new method for quickly and automatically analysis of the image of vickers hardness using wavelet theory," *Acta Metrologica Sinica* **26**, 245–248 (July 2005).

[4] Ji, Y. and Xu, A., "A new method for automatically measurement of vickers hardness using thick line hough transform and least square method," in [*Proceedings of the 2nd International Congress on Image and Signal Processing (CISP'09)*], 1–4 (2009).

[5] Macedo, M., Mendes, V., Conci, A., and Leta, F., "Using hough transform as an auxiliary technique for vickers hardness measurement," in [*Proceedings of the 13th International Conference on Systems, Signals and Image Processing (IWSSIP'06)*], 287–290 (2006).

[6] Mendes, V. and Leta, F., "Automatic measurement of Brinell and Vickers hardness using computer vision techniques," in [*Proceedings of the XVII IMEKO World Congress*], 992–995 (June 2003).

[7] Sugimoto, T. and Kawaguchi, T., "Development of an automatic Vickers hardness testing system using image processing technology," *IEEE Transactions on Industrial Electronics* **44**, 696–702 (Oct. 1997).

[8] Yao, L. and Fang, C.-H., "A hardness measuring method based on hough fuzzy vertex detection algorithm," *IEEE Trans. on Industrial Electronics* **53**(3), 963–973 (2006).

[9] Maier, A. and 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)*], 295–300 (Sept. 2011).

[10] Gadermayr, M., Maier, A., and Uhl, A., "Algorithms for microindentation measurement in automated Vickers hardness testing," in [*Tenth International Conference on Quality Control for Artificial Vision (QCAV'11)*], Pinoli, J.-C., Debayle, J., Gavet, Y., Cruy, F., and Lambert, C., eds., *Proceedings of SPIE*, 80000M–1 – 80000M–10, SPIE, St. Etienne, France (June 2011).

[11] Gadermayr, M., Maier, A., and Uhl, A., "A robust algorithm for automated microindentation measurement in vickers hardness testing," *Journal of Electronic Imaging* (2012). accepted.

[12] Nayar, S. K., "Shape from focus system," in [*Computer vision and pattern recognition proceedings CVPR '92*], 302–308 (1992).

[13] Kass, M., Witkin, A., and Terzopoulos, D., "Snakes: Active contour models," *International Journal of Computer Vision* **1**(4), 321–331 (1988).

[14] Osher, S. and Sethian, J. A., "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations," *Journal of Computational Physics* **79**, 12–49 (November 1988).

[15] Caselles, V., Kimmel, R., and Sapiro, G., "Geodesic active contours," *International Journal of Computer Vision* **22**(1), 61–79 (1997).

[16] Chan, T. and Vese, A., "Active contours without edges," *IEEE Transactions on Image Processing* **10**, 266–277 (February 2001).

[17] Cremers, D., Rousson, M., and 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).

[18] Xu, C. and Prince, J. L., "Gradient vector flow: A new external force for snakes," in [*IEEE Proc. Conf. on Computer Vision and Pattern Recognition*], 66–71 (1997).

[19] Kiser, C., Musial, C., and Sen, P., "Accelerating active contour algorithms with the gradient diffusion field," *Pattern Recognition ICPR 2008 19th International Conference on* , 1–4 (2008).

[20] Chan, T. and Zhu, W., "Level set based shape prior segmentation," *Computer Vision and Pattern Recognition CVPR IEEE Computer Society Conference on* , 1164–1170 (2005).

[21] Zhu, G. P., Zeng, Q. S., and Wang, C. H., "Robust shape based active contour for circle detection," *The Imaging Science Journal* **56**(3), 175–178 (2008).

[22] Gadermayr, M. and Uhl, A., "Dual-resolution active contours segmentation of vickers indentation images," Tech. Rep. 2012-01, Department of Computer Sciences, University of Salzburg, Austria, `http://www.cosy.sbg.ac.at/research/tr.html` (2012).

[23] Cohen, L., "On active contour models and balloons," *CVGIP: Graphical models and image processing* **53**, 211–218 (March 1991).

[24] Gadermayr, M., Maier, A., and Uhl, A., "Image segmentation of vickers indentations using shape from focus," *International Conference on Image Analysis and Recognition* (2012).