PROCEEDINGS OF SPIE

An edge-from-focus approach to 3D inspection and metrology

Fuqin Deng, Jia Chen, Jianyang Liu, Zhijun Zhang, Jiangwen Deng, et al.

SPIE.

An edge-from-focus approach to 3D inspection and metrology

Fuqin Deng[1], Jia Chen[2], Jianyang Liu[3], Zhijun Zhang[1], Jiangwen Deng[1], Kenneth S.M. Fung[1], and Edmund Y. Lam[4]

[1] ASM Pacific Technology Limited, Kwai Chung, Hong Kong;
[2] Shenzhen Graduate School of Harbin Institute of Technology, Shenzhen, China;
[3] Southwest Jiaotong University, Chengdu, China;
[4] Imaging Systems Laboratory, Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong.

ABSTRACT

We propose an edge-based depth-from-focus technique for high-precision non-contact industrial inspection and metrology applications. In our system, an objective lens with a large numerical aperture is chosen to resolve the edge details of the measured object. By motorizing this imaging system, we capture the high-resolution edges within every narrow depth of field. We can therefore extend the measured range and keep a high resolution at the same time. Yet, on the surfaces with a large depth variation, a significant amount of data around each measured point are out of focus within the captured images. Then, it is difficult to extract the valuable information from these out-of-focus data due to the depth-variant blur. Moreover, these data impede the extraction of continuous contours for the measurement objects in high-level machine vision applications. The proposed approach however makes use of the out-of-focus data to synthesize a depth-invariant smoothed image, and then robustly locates the positions of high contrast edges based on non-maximum suppression and hysteresis thresholding. Furthermore, by focus analysis of both the in-focus and the out-of-focus data, we reconstruct the high-precision 3D edges for metrology applications.

Keywords: three-dimensional image acquisition, edge detection, industrial inspection, 3D metrology

1. INTRODUCTION

In recent years, there has been increased attention regarding the depth information from images.[1,2] For example, in the semiconductor industry, high-resolution non-contact 3D reconstruction is often required for checking the quality of the wire bonding process in automatic optical inspection (AOI) systems.[3] The IC chips contain small feature sizes, requiring high resolution and high precision. For instance, a $20\times$ magnified objective lens with a 0.4 numerical aperture (NA) may be used to resolve the sharp edges such as the outer contours of the centering ball of an unfinished IC chip, as shown in Figure 1(a). Yet, the high resolvability of the large NA objective lens results in a small field of view (FOV) and a shallow depth of the field (DOF). Figure 1(b) is another image at a different focal level, in which only the boundaries of the metal pads on the IC chip are sharply focused. Besides high resolution, we often also need to design a machine vision system with a large FOV and DOF so that we can achieve an automatic visual inspection of large IC samples. Furthermore, 3D depth of the object boundary is needed for some high-precision metrology applications.

In this paper, we propose an imaging system and computational method for 3D edge reconstruction in high-precision metrology applications, such as inspecting the quality of the IC chips and improving the wire bonding process for IC chip fabrication in the semiconductor industry. We first review the non-contact optical 3D reconstruction methods in Section 2. Then, we present the details of the proposed method in Section 3. After that, in Section 4, we show some experimental results for the proposed method and illustrate some industrial metrology applications in semiconductor fabrication. Finally, some concluding remarks are given in Section 5.

Further author information: (Send correspondence to Edmund Lam)
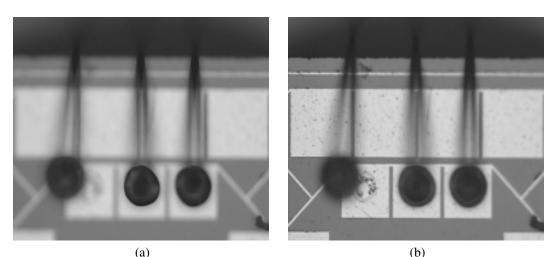E-mail: elam@eee.hku.hk

Figure 1. Images of a part of an unfinished sample from the IC chip production line. (a) The boundary of centering ball is well focused; (b) The boundaries of the pads are well focused.

2. OVERVIEW OF NON-CONTACT OPTICAL 3D RECONSTRUCTION AND METROLOGY

Over the years, various non-contact optical 3D reconstruction techniques have been developed, with their advantages and disadvantages. Lee and Kweon designed a compact system with single camera, which can obtain a stereo pair from a single shot using a biprism.[4] However, this design reduces the resolution of the image and limits the FOV for each reconstruction. Interference-based methods such as phase-shifting interferometry have been proposed for high-resolution surface metrology, but most of these systems can only achieve a small measurement range and they are sensitive to vibration.[5] By adding pinholes in front of the illumination source and the sensor, confocal imaging can also ensure high-resolution imaging by eliminating the out-of-focus light. However, it involves time-consuming scanning process for surface profilometry of stationary samples. Some laser patterns such as dots or lines have been used in laser triangulation system for surface profilometry. Usually, speckle is an inevitable factor and it degrades the performance.[6] Other structured-light methods have been developed by projecting pre-defined structured patterns and identifying the unique pattern for depth recovery.[7] These illumination-coded approaches can be implemented by an add-on projection grating, an LCD or a DLP projector, and they provide an efficient full-field surface profilometry method for metrology applications.[8] However, like other triangulation-based 3D metrology techniques, these structured-light methods fail at the occluded regions.[9]

To eliminate the occlusion problem by triangulation-based 3D reconstruction techniques, shape-from-focus (SFF) provides a powerful means of recovering the shapes of rough surfaces,[10] but for general surfaces with sudden height variations, the accuracy of depth around the edge is seriously degraded by the out-of-focus light from the neighboring points. Moreover, the erroneously reconstructed depth at the low contrast region brings difficulties in extracting valuable information in later metrology applications.

3. THE EDGE-FROM-FOCUS TECHNIQUE

3.1 The principle of depth reconstruction

Figure 2(a) shows the coaxial setup of our focusing system, equipped with high-magnification objective lenses and a large aperture. There are some advantages for such a configuration. First, based on coaxial illumination and imaging, the focusing system can reduce the occlusion effect in many triangulation-based 3D reconstruction systems such as the fringe pattern projection and the laser triangulation systems. Second, it provides a fine lateral resolution for metrology applications. Equipped with high-magnification objective lenses, it can resolve the micrometer-level features as many microscopes do in biomedical applications. Third, with a large NA, it can image the edge features with a high SNR. This large aperture also contributes to a shallow DOF, which ensures a fine axial resolution for depth estimation. Moreover, there is a motorized imaging system along the optical axis to capture the narrow-depth images of the inspected object at every height levels, with an X-Y table to translate the object for extending the FOV. Accordingly, not only can we obtain
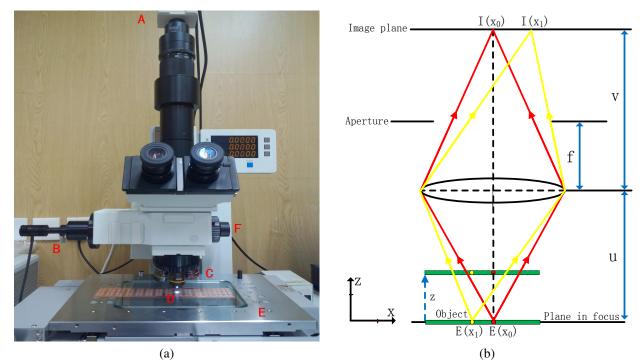
Figure 2. (a) The focusing system for 3D metrology. A: The CCD sensor; B: The light source for coaxial illumination; C: The objective lens; D: The inspected IC samples; E: The X-Y table for translating the inspected object on the X-Y plane; F: The motion system for translating the optical system along Z axis; (b) The geometry model for image formation.

images with high-resolution in both axial and lateral directions, but we can also extend both the DOF and the FOV freely for various metrology applications.

Figure 2(b) uses a single thin lens model to illustrate the image formation process and the principle of the depth reconstruction based on the focused and the defocused images. Let the focal length and the image distance be f and v, respectively. Based on the Gaussian lens law, the plane at the object distance u satisfying the lens equation

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \tag{1}$$

will be sharply focused. Denoting this in-focus plane in the object space as the reference plane (i.e., $z = 0$), we consider the image $I(x)$ of an object $E(x)$, where x is the coordinate on the image plane (generally, x is a two-element vector).

For a specific image point at x_0, we focus on the cone of light collected by the lens for imaging this point, as illustrated by the red rays in Figure 2(b). In the first scenario, when the object plane coincides with the reference plane, the cone of rays from the feature point $E(x_0)$ on the object are collected for imaging the point $I(x_0)$. Since the nearby feature point $E(x_1)$ is not within this cone, the rays from $E(x_1)$ do not contribute to the image point $I(x_0)$. Instead, they are collected by the lens for imaging another point $I(x_1)$. Since both features are sharply focused, we can easily distinguish them for high-precision 2D metrology applications.

However, such a sharp image may not always be obtained in practice due to the limited DOF. In another more common scenario, we consider the image $I_1(x_0)$ when the object is off the focused plane by z_1 along the optical axis. Then, the rays from a small region around the point $E(x_0)$, including the rays from the point $E(x_1)$ contribute to the image point $I_1(x_0)$. Therefore, a defocused image is formed on the image plane. In a linear and spatially invariant imaging system, the defocused image $I_1(x)$ can be modeled as the convolution of the object $E(x)$ with a blurring function $B(z_1)$, whose width is proportional to the offset z_1. Assuming $N_1(x)$ to be the additive noise in the image $I_1(x)$, we have

$$I_1(x) = E(x) * B(z_1) + N_1(x), \tag{2}$$

where $*$ denotes the convolution operation.

In an analogous manner, we can obtain K images $I_k(x)$ at different focus levels z_k, where $k = 1, \ldots, K$. Since the image becomes more defocused when the inspected object deviates farther from the in-focus plane, we can infer the depth of the object based on both the in-focus and out-of-focus data. Let such an operation be described abstractly as $\mathcal{H}\{\cdot\}$; we compute the focus measure for each point focal level z_k by

$$F_k(x) = \mathcal{H}\{I_k(x)\}. \tag{3}$$

Various focus operators, some gradient-based and some statistics-based, have been proposed to measure the degree of the focus.[11] For example, we can choose the Sobel operators S_1 and S_2 to define the focus detection operation such that

$$F_k(x) = \mathcal{H}\{I_k(x)\} = \sqrt{\left(I_k(x) * S_1\right)^2 + \left(I_k(x) * S_2\right)^2}, \tag{4}$$

where

$$S_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad S_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{5}$$

This focus measure computes an approximation of the gradient of the image intensity, and gives an efficient and effective model for the edge contrast of feature points. At different focus levels, we can obtain the corresponding focus measure in an analogous manner and then formulate the focus curve $F_k(x)$ along different z levels. Since the resulting focus measure achieves the maximum when the measured edge point is sharply focused and decreases gradually when this edge point is out of focus, the z level with the maximum focus measure indicates the relative offset of this measured point to the reference plane, i.e., the depth. After reconstructing the heights for the other points in an analogous manner, we have the depth map for the inspected object. Different from conventional SFF technique, our proposed EFF includes the following two major procedures: 2D edge localization and 3D edge reconstruction.

3.2 2D edge localization

There have been various edge detection methods such as template matching approaches, statistical approaches, and partial differential equation approaches.[12] Among them, Canny's edge detection remains a mainstream technique.[13] Most of them assume that the image is captured by a linear system with uniform blurring on the whole FOV, and therefore few of them are suitable for handling the depth-variant blurring. Moreover, due to the shallow DOF, different parts of the object are blurred differently, which introduces difficulties in choosing appropriate parameters for extracting the valid edges. It is therefore difficult to extract all the valid edges of the inspected object effectively from the captured raw images. As different parts of the object are in focus and distributed within different images, one may need to perform edge detection within each image, and stitch all these trivial edges from different images to synthesize a full edge map of the object.

For both robustness and efficiency considerations, our edge detection method consists of the following four steps, where $I_k(x)$ $(k = 1, \ldots, K)$ are the captured image sequence of the inspected object at different focal levels z_k around the focal plane.

1. Synthesize the image with depth-invariant smoothing by

$$I_s(x) = \frac{1}{K} \sum_{k=1}^{K} I_k(x). \tag{6}$$

2. Calculate the magnitude $M(x)$ and the gradient direction $\theta(x)$ for the edge candidates based on $I_x(x) = I_s(x) * S_1$ and $I_y(x) = I_s(x) * S_2$. Then,

$$M(x) = \sqrt{I_x^2(x) + I_y^2(x)}, \quad \text{and} \quad \theta(x) = \tan^{-1} \frac{I_y(x)}{I_x(x)}. \tag{7}$$

3. Apply non-maximum suppression along the above gradient direction for each edge candidate.

4. Apply hysteresis thresholding and link the valid edges.

There are several advantages for the above edge detection strategy. First, by summing the images with different blurring effect at different focal levels, we synthesize a uniformly smoothed image so that we can use existing techniques to locate all the valid edges from this single synthesized image efficiently, instead of detecting the distributed trivial edges at different parts of the object from different raw images combined with a subsequent merging process. Second, the averaging operation for image synthesis helps to remove most of the noise, so that simple convolution operations can be used to calculate the magnitude and the gradient direction of the edges. Third, since the magnitudes of true edges are local maxima along the gradient direction, we can mask out the false edges by suppressing the non-maximum magnitudes. This suppression technique also helps to reduce the number of edge candidates for depth estimation and eliminate the possible candidates with false depth values around the high-contrast true edges. Fourth, true edges are usually connected to form contours of the object and false edges are usually randomly distributed in most real images. Therefore, a double-threshold strategy is adopted to link the valid edges and remove the isolated false edges. In this strategy, we use the high threshold to locate the strong edges and then take them as edge seeds for contour growing. A low threshold is used to identify all the edges connecting these edge seeds. Note that a similar edge detection strategy has also been used by other researchers,[13] where usually a uniformly blurring is assumed within the image, while our method can handle the image sequence with depth-variant blurring due to the high-resolution focusing system.

Since setting absolute threshold in conventional edge detection methods may change the number of resulting edges quite differently after degradation of the illumination source, we choose the threshold based on the relative magnitude of the edges to ensure detection of a given percentage of the edge points within the FOV. For some industrial applications in measuring objects of the same type, the percentage of the edge points over the total pixels are almost the same and they can be determined by the shapes of the objects for a given FOV in advance. Based on this thresholding strategy, the number of detected edges are almost the same, and more repeatable edge detection results can be achieved by the proposed method even when the intensity of the illumination changes.
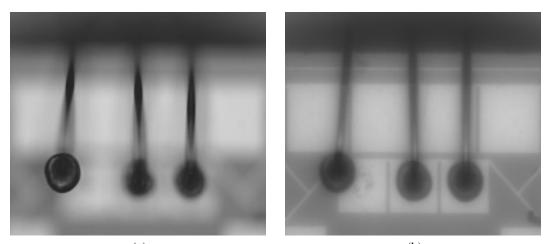
3.3 3D edge reconstruction

After the non-maximum suppression and hysteresis thresholding, most of the false edges and low contrast points are eliminated. These operations are helpful for further high-level 2D metrology applications. However, for some high precision applications such as measuring the height of some important features points, 2D information is not enough. Therefore, we consider the 3D reconstruction of edge points for the inspected object. Let x_j $(j = 1, \ldots, M)$ be the resulting edge points of the inspected object detected by the previous 2D edge identification strategy. For these edge points, we can assume that they achieve the unique maximum focus measure when they are on the in-focus plane of the system. Then, we formulate the depth estimation as an optimization problem at these edge points. Using the symbols defined above, we calculate the focus measure for each measured edge point x_j, and reconstruct its depth $Z(x_j)$ from the K raw images at different focal levels z_k around the focused plane by

$$\begin{aligned} \underset{z_k}{\text{maximize}} \quad & F_k(x_j) \\ \text{subject to} \quad & F_k(x_j) = \mathcal{H}\{I_k(x_j)\} \\ & k = 1, \ldots, K. \end{aligned}$$

After finding the index k of the best focused image for each measured point, we interpolate a quadratic function with the discrete focus measure $F_{k-1}(x_j)$, $F_k(x_j)$ and $F_{k+1}(x_j)$. Then, we refine its depth $Z(x_j)$ by locating the position with the maximal focus measure on this interpolated quadratic focus curve.

As mentioned in Section 3.1, many focus operators can be used to measure the degree of the focus. The values of the focus measures may be different by using different focus operators, but these focus measures all achieve their peak values when the object is on the in-focus plane. For simplicity, similar to the edge detection procedure, the Sobel operators S_1 and S_2 are chosen as the focus operators for estimating the focus measure during depth reconstruction. Since we only need to calculate the focus measures for true edge points in our formulation and the number of edge point M usually is much smaller than the total pixels within the image, the proposed method involves much less computation, compared with the full-field depth reconstruction by the SFF. Furthermore, since most of the false edges have been eliminated by our 2D edge identification procedure, we do not need any complicated and robust 3D feature extraction technique to extract the useful 3D data from the full-field depth image by the SFF.

Figure 3. (a) One of the raw images from the focusing system; (b) The synthesized image by summing all the raw images.
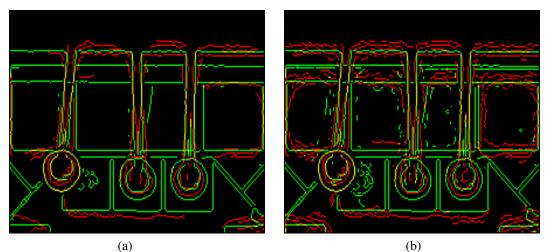

Figure 4. (a) Comparison of edge detection results with larger threshold; (b) Comparison of edge detection results with smaller threshold. Red and green edges are the detection results of the raw image and our synthesized image, respectively.

4. EXPERIMENTS

In the above section, we have presented a high-resolution focusing system and the framework of EFF for reconstructing the 3D edges of inspected objects by focus analysis. In experiments, we apply them in 3D metrology and inspection applications for IC chips after the wire bonding process. To resolve the details of the edge features, a $20\times$ objective lens with NA 0.4 is used. The final resolution is 0.83μm per pixel and the sizes of these images are 768×576 pixels. Since the height range of the inspected samples is usually within 200μm, to cover this measurement range, we move our motorized focusing system along the optical axis (i.e., the z-axis) at every 6μm and then capture 35 of the partially focused images, including the wires, the balls and the pads of the IC chip.

Now, we illustrate some results based on the proposed EFF. Figure 3(a) shows one of the captured raw images for the unfinished IC chip from our high-resolution focusing system. We see that only a small part of the wire loops and the outer boundary of the left ball are well focused. This imaging property limits us to perform image processing only in a small region of interest at a time and it is inconvenient or inefficient to fuse these distributed data from the raw images at different focal levels for metrology applications. However, with our proposed method, we can synthesize a uniformly smoothed image (b) for later image processing and applications.
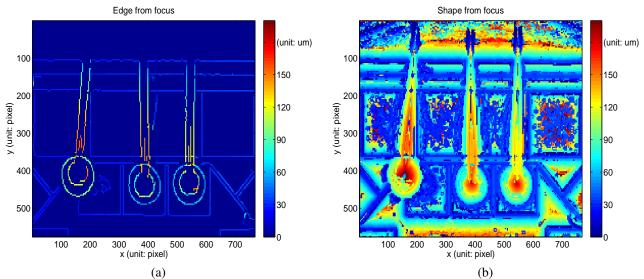
Figure 5. (a) The reconstructed 3D edge map by EFF; (b) The reconstructed depth by SFF.

Since edges are the most important data for most machine vision applications, we consider the edge detection performance in our system. Figure 4(a) shows the edge detection results on the raw image and our synthesized image, respectively. For a fair comparison, we use the same non-maximum suppression strategy. The upper threshold is 0.8 and the lower threshold is 0.4, for both images in the hysteresis thresholding stage. These thresholds select the points with edge magnitude larger than 80% of the data, as the strong edges and marks 40% of the data with small magnitude as the non-edge positions. At the other edge candidate positions, we grow from all the strong edges and link all the possible edges from these strong seeds. As illustrated in Section 3.1, the light rays from the high contrast boundaries of the object bring the variation of the image intensities to their neighboring points when they become out of focus. Therefore, many false edges near the boundaries of the wires and balls are detected from the raw image. Some boundaries of the pads are also missing due to the serious blurring caused by out of focus. In comparison, the proposed method can highlight most of the true edges on the boundaries of wires, balls and pads from our synthesized image.

We also test the sensitivity of the edge detection results to the chosen thresholds. Figure 4(b) shows the detection results at another set of thresholds, with upper one equal to 0.6 and lower one equal to 0.3. We see that the number of detected edges are quite different in the raw image, and more false edges are extracted when we decrease the threshold, as shown in the red edge map. Note we can still obtain the clean and accurate green edge map based on the proposed edge strategy even with different thresholds. Due to this insensitivity to the thresholds, it is much easier for us to choose the parameters for detecting the edges from our synthesized image, compared with the raw image.

From the above experiments, we see that in order to locate the edges positions of the inspected object, we only need to perform edge detection from a single synthesized image instead of all the raw images. Moreover, we can obtain a cleaner and more accurate edge map than that from the raw images.

Based on the 2D edge map, on each edge point, by locating the image and the z value from the motor corresponding to the maximum focus measure, we can reconstruct the depth of the edge. Figure 5(a) shows the 3D edges by our EFF. Comparing with the depth map in (b), reconstructed by the SFF, we have a much cleaner 3D edge map which highlights the depth of the boundaries of wires, balls and pads. While it is difficult to extract useful 3D boundaries of object from (b) for further metrology and inspection applications, based on our reconstructed 3D edges, we can effectively recognize the left ball is higher than the other as the depth image shows. However, it is difficult and inconvenient to achieve such result based on the reconstructed shape from the SFF, since there are so many false height values within the reconstructed depth. Similarly, we can inspect the shape of the wire loop along the boundaries of the wires and measure the height between the top part of each wire to the corresponding pad for checking the quality of the wire bonding process. Furthermore, instead of storing the whole image sequences with a huge amount of data, we only need to store this sparse 3D edge map for further machine vision applications.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we present a focusing system and a framework to reconstruct high-resolution 3D edges of the measured object. The proposed EFF approach can effectively suppress the possible false heights caused by the out-of-focus light from nearby high contrast boundaries and at non-texture regions with low contrast. In order to reconstruct the depth at the low contrast regions precisely, in the future, we will consider to project high contrast patterns onto the surface of the object, and then reconstruct the dense 3D surface of the object by analyzing the focus and defocus of these projected patterns.

Acknowledgment

REFERENCES

[1] Deng, F., Liu, C., Sze, W., Deng, J., Fung, K. S. M., Leung, W., and Lam, E. Y., "Illumination-invariant phase-shifting algorithm for three-dimensional profilometry of a moving object," *Optical Engineering* **51**, 097001 (September 2012).

[2] Deng, F., Liu, C., Sze, W., Deng, J., Fung, K. S., and Lam, E. Y., "An INSPECT measurement system for moving objects," *IEEE Transactions on Instrumentation and Measurement* (Accepted 2014).

[3] Wang, C. and Sun, R., "The quality test of wire bonding," *Modern Applied Science* **3**, 50–56 (December 2009).

[4] Lee, D. H. and Kweon, I. S., "A novel stereo camera system by a biprism," *IEEE Transactions on Robotics and Automation* **16**, 528–541 (October 2000).

[5] de Groot, P. J., "Vibration in phase-shifting interferometry," *Journal of the Optical Society of America A* **12**, 354–365 (February 1996).

[6] Dorsch, R. G., Häusler, G., and Herrmann, J. M., "Laser triangulation: Fundamental uncertainty in distance measurement," *Applied Optics* **33**, 1306–1314 (March 1994).

[7] Salvi, J., Pagès, J., and Batlle, J., "Pattern codification strategies in structured light systems," *Pattern Recognition* **37**, 827–849 (April 2004).

[8] Deng, F., Sze, W. F., Deng, J., Fung, K. S., Leung, W., and Lam, E. Y., "Regularized multiframe phase-shifting algorithm for three-dimensional profilometry," *Applied Optics* **51**, 33–42 (January 2012).

[9] Scharstein, D. and Szeliski, R., "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision* **47**, 7–42 (May 2002).

[10] Nayar, S. K., "Shape from focus," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**, 824–831 (August 1994).

[11] Pertuza, S., Puiga, D., and Garciab, M. A., "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition* **46**, 1415–1432 (May 2013).

[12] Deng, F., Fung, K. S., Deng, J., and Lam, E. Y., "An edge detection algorithm based on rectangular gaussian kernels for machine vision applications," in [*Image Processing: Machine Vision Applications*], *Proceedings of the SPIE* **7251**, 72510N (January 2009).

[13] Canny, J., "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8**, 679–698 (November 1986).