

Automatic Image Segmentation using Sobel Operator and k -Means Clustering: A Case Study in Volume Measurement System for Food Products

Joko Siswanto^{1,2}, Anton Satria Prabuwono^{1,3}, Azizi Abdullah¹, Bahari Idrus¹

¹Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 UKM, Selangor D. E., Malaysia

²Faculty of Engineering, Universitas Surabaya, Jl. Kali Rungkut, Surabaya, 60293, Indonesia

³Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh 21911, Saudi Arabia

joko_siswanto@ubaya.ac.id, antonsatria@ftsm.ukm.my, azizia@ukm.edu.my, bahari@ukm.edu.my

Abstract—Image segmentation is one of important step in visual inspection of food product using computer vision system. However, segmentation of food product image is not easily performed if the image has low contrast with its background or the background in acquired image is not homogeneous. This paper proposes k -means clustering combined with Sobel operator for automatic food product image segmentation. Sobel operator was used to determine region of interest (ROI) and k -means clustering was then employed to separate object and background in ROI. The area outside ROI was considered as background. The proposed method has been validated using 100 images of food product from ten different types. The validation results show that the proposed segmentation method achieves good segmentation result.

Keywords—food product; segmentation; Sobel operator; k -means clustering

I. INTRODUCTION

Automatic visual inspection (AVI) is a branch of computer vision used to control the quality of manufactured products automatically. [1, 2]. AVI system is very suitable for automatic visual inspection in agriculture and food industry, because it is non-destructive inspection method [3]. AVI for food products is performed by extracting geometric and surface features from the image of object. Generally, it is usually achieved by using a camera connected to a computer and involves five steps. One of these steps is segmentation [4].

In segmentation step, the image of object is decomposed into area of interest and background. One of important step in visual inspection is segmentation. The results of next step depend on segmentation result. Inaccurate segmentation result will lead to inaccurate inspection result [5]. Segmentation techniques commonly used in visual quality inspection of food products are thresholding-based, gradient based, region-based, and classification based segmentation. However, no one of these approaches can produce accurate segmentation result for various food products [4]. Therefore, the combination of several approaches can be considered to obtain accurate segmentation result.

Thresholding-based segmentation is a simplest approach. It is performed on a grayscale image by determining a threshold value T to separate objects from its background and to produce

a binary image. A pixel in the grayscale image with intensity greater than T is set as object pixel, otherwise as background pixel. The threshold value can be determined manually by considering the distribution of pixel intensity or automatically. Otsu [6] has proposed a method to determine the threshold value automatically by maximizing between class variance in grayscale image. However, thresholding-based segmentation is not easily performed if the object has low contrast with its background. Gradient-based segmentation is performed using convolute gradient operators, such as Sobel operator [7]. Sobel operator computes the approximation of gradient of image intensity both in vertical and horizontal directions. This approach only extracts the edge of object. Gradient-based segmentation is usually used in extracting size and shape features [8, 9].

During visual inspection using CVS, the image of inspected object is usually captured using homogenous background color, such as black [9-12] and white [8, 13-17]. The color of background is selected depending on the color of inspected object, such that image segmentation can be easily performed [9, 15]. However, the color of background in acquired image is not always homogenous due to the influence of illuminations or the presence of shadows. In this condition, segmentation in the whole image is not easy. Therefore, segmentation in an area around the object could be considered to minimize inhomogeneity background.

Siswanto, et al. [18] have proposed a CVS for volume measurement of irregularly shaped food product based on Monte Carlo method. The system consisted five cameras, a computer, a light source, and a black background. The images of measured object were captured using five cameras, one from top view and four from surrounding views. The images were then processed to produce binary images. The binary images were used to perform volume measurement based on Monte Carlo method. To reduce measurement error, a heuristic adjustment was applied to the result of Monte Carlo method. In this CSV, a grayscale image from HSV image was firstly constructed by weighted sum of H, S, and V components to produce a binary image. The grayscale image was then segmented using thresholding-based segmentation. The threshold value was determined automatically using an iterative procedure as described by Gonzalez and Woods [19].

Although good segmentation result was achieved for several types of food product, this method is not fully automatic. The values of the weights used to construct grayscale image were chosen manually by user, such that the optimum segmentation result would be obtained.

Image segmentation of food product can be regarded as a clustering problem with the number of cluster equal to two. k -means clustering is a simple method that is often applied in image segmentation [5]. Generally, k -means clustering is used to divide a set of objects into some cluster, such that the distance from each object in each cluster to its cluster center is minimized [20]. With k -means clustering, image segmentation is not only performed on grayscale image but also on multichannel image, such as RGB or HSV image. Several image segmentation methods using k -means clustering have been proposed, such as [21], [22], [23], [24], [25] and [26]. However, almost all of them did not use k -means clustering to separate object from its background.

Coleman and Andrews [21] computed 12 features from an image and then used these features to segment the image into several (2 – 6) regions by employing k -means clustering. Marroquin and Girosi [22] have presented some extensions to k -mean algorithm by using state variable. The method was used to segment a scientific image of heart into two cluster, interior and background. Ng, et al. [23] used k -means clustering combined with improved watershed algorithm to for medical image segmentation. They reported that the proposed method can reduce the number of partitions of up to 92% compared with conventional watershed algorithm. Chen, et al. [24] have proposed fast image segmentation based on k -means clustering with histogram in HSV color space. The proposed method can automatically estimate the initialization of centroids and the number of cluster from image histogram. Ray and Turi [25] have proposed a method to determine the number of cluster in k -means clustering and application in color image segmentation. They used a simple validity measure based on intra-cluster and inter-cluster distance measures to determine the number of cluster. The proposed method was tested to segment synthetic and natural images into several numbers (2-10) of clusters. Chitade and Katiyar [26] have proposed satellite image segmentation method using k -means clustering. The proposed method performed k -means clustering in a^* and b^* channels of La^*b^* color space.

This paper proposes k -means clustering combined with Sobel operator for a fully automatic food product image segmentation method. The region of interest (ROI) is determined using Sobel operator, before applying k -means clustering.

II. MATERIALS

The image samples used to validate the proposed method were 100 images of food product from ten different types. The samples consisted of 10 images of red delicious apple, 10 images of green apple, 10 images of potato, 10 images of orange, 10 images of tomato, 10 images of chukanan mango, 10 images of egg, 10 images of ya pear, 10 images of packham pear, and 10 images of carrot. Five Logitech web cameras HD Webcam 270h were used to acquire images, one from the top view and four from the surrounding views. The images were

captured using a black background in RGB format. Fig. 1 shows the example of images used to validate the proposed method. Although the images were captured using a black background, the black background in acquired image is not homogeneous due to the influence of illumination or the present of shadow, such as in the image of red delicious apple, image of potato, image of orange, image of tomato, image of chukanan mango, image of ya pear, image of packham pear, image of carrot in Fig. 1.

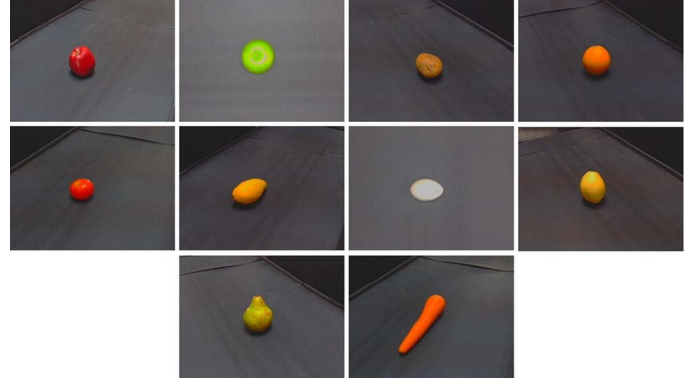


Fig. 1. Images used to validate the proposed method

III. PROPOSED METHOD

The main steps for the proposed method are pre-processing, edge detection, determine ROI, k -means clustering, and morphological operation. Fig. 2 depicts the flowchart of the proposed method.

A. Pre-processing

Firstly, the image was transformed to HSV format. H, S, and V component were obtained from R, G, and B component using a transformation as described in [27]. The object could be easily separated from its background in S and V components. Fig. 3 (a) and (b) show the example of image in S and V components, respectively. To reduce noises, a 50×50 Gaussian filter [28] with $\sigma = 10$ was applied to both S and V components. A grayscale image G was then constructed by calculating the average of S and V component, as shown in Fig. 3 (c).

B. Edge Detection

The edge of object was obtained by convoluting the grayscale image G with Sobel operator. Sobel operator performs edge detection by determining the gradient of the grayscale image G with respect to horizontal and vertical directions. The gradient of a pixel in G is calculated by using operators as in (1). A pixel in G is an edge if $(G_x^2 + G_y^2)^{1/2} \geq T$ in that pixel, for some threshold value T.

$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 2 \end{bmatrix} \quad (1)$$

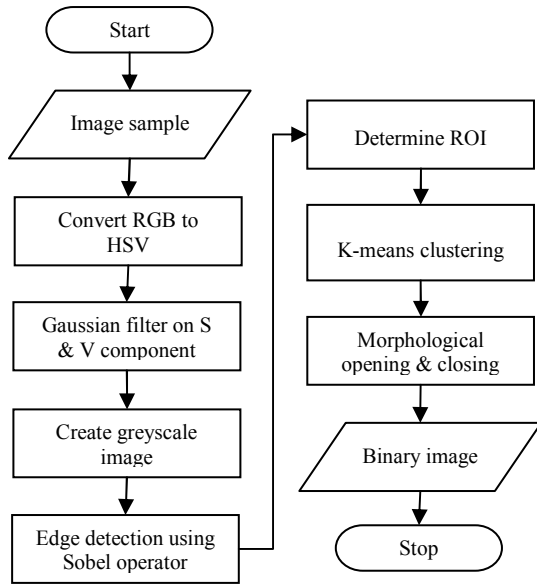


Fig. 2. The steps of the segmentation method

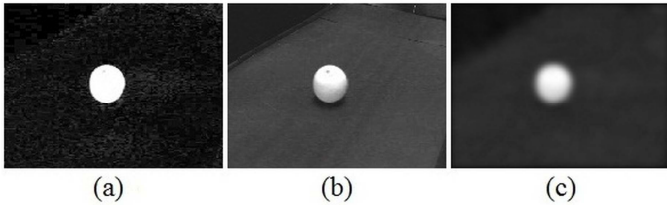


Fig. 3. Captured image: (a) in S and (b) V components (c) grayscale image

C. Region of Interest (ROI)

To determine ROI, the axis-aligned minimum bounding box of the edge of object was first created, as shown in Fig. 4 (a). The bounding box was determined by finding the minimum and maximum coordinates of edge both in x and y directions. ROI was defined by enlarging the bounding box by 50 pixels in four directions. The original image was then cropped according to this ROI. Fig. 4 depicts ROI and cropped image according to ROI.

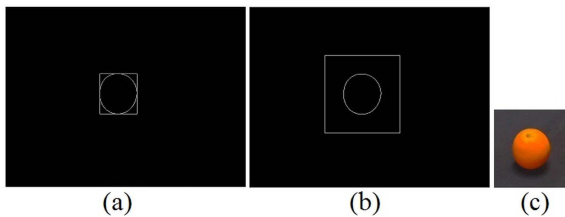


Fig. 4. (a) minimum bounding rectangle (b) ROI (c) cropped image

D. k -means Clustering

For image segmentation, k -means clustering was applied on S and V component of cropped image to preserve the information contained in these components. The pixel values of S and V component in cropped image were arranged as two dimensional vectors $\mathbf{x}_i = (s_i, v_i)$, $i = 1, 2, \dots, n$. Where s_i and v_i are the values of S and V component of pixel i , respectively; and n is the number of pixel in cropped image.

k -means clustering was used to divide $\{\mathbf{x}_i, i = 1, 2, \dots, n\}$ into two groups to produce a binary image B, as in the following steps:

1. Initialize the cluster centers \mathbf{m}_1 and \mathbf{m}_2 .

2. For i from 1 to n do step 3.

3. Set

$$b_i = \begin{cases} 1, & \|\mathbf{x}_i - \mathbf{m}_1\| < \|\mathbf{x}_i - \mathbf{m}_2\| \\ 0, & \text{otherwise} \end{cases}$$

4. Set $\mathbf{m}_1 = \sum_{i=1}^n b_i \mathbf{x}_i / \sum_{i=1}^n b_i$, $\mathbf{m}_2 = \sum_{i=1}^n (1-b_i) \mathbf{x}_i / \sum_{i=1}^n (1-b_i)$.

5. Repeat step 2 to 4 until \mathbf{m}_1 and \mathbf{m}_2 converge.

6. If $\|\mathbf{m}_1\| < \|\mathbf{m}_2\|$ then do step 7

7. For i from 1 to n set $b_i = 1 - b_i$.

8. Construct B from b_i , $i = 1, 2, \dots, n$ according to the position of pixel i in cropped image.

E. Morphological Operation

Sometimes, there were a few object pixels that clustered as background pixels or vice versa. To solve this problem, morphological closing and morphological opening [29] with a disk-shaped structuring element radius 10 were used to remove the black spots in the object and the white spots in the background, respectively. The binary image B was then enlarged to a binary image that has same size with original image according to the position of ROI. All pixels outside ROI were considered as background pixels with binary value 0. Fig. 5 depicts resulted binary image (b) and its original image (a).

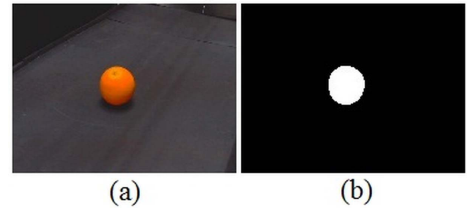


Fig. 5. (a) original image (b) resulted binary image

IV. VALIDATION

All image samples were segmented using the proposed method. Manual segmentation was also performed on all samples for validation. To perform manual segmentation, all image samples were edited manually using photo editor software by replacing the color of object and background with white and black, respectively. Mean square error (MSE) was calculated to obtain the accuracy of the proposed method compared with manual segmentation. MSE was obtained by calculating the average of the squares difference between two binary images, as in (2).

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (B_1(i, j) - B_2(i, j))^2 \quad (2)$$

Where B_1 and B_2 are binary images obtained using an image segmentation method and manual segmentation, respectively. M and N are the dimension of binary images. The value of MSE is in the range $[0,1]$. The lower MSE, the better segmentation accuracy is.

In this study, the proposed methods was also compared with segmentation method proposed by Siswanto, et al. [18] and image segmentation on the whole image using k -means clustering. To perform image segmentation using k -means clustering on the whole image, the pixel values of S and V component in the whole image were arranged as two dimensional vectors $y_i = (s_i, v_i), i = 1, 2, \dots, m$. Where s_i and v_i are the values of S and V component of pixel i , respectively; and m is the number of pixels in the whole image. The set of vectors $\{y_i, i = 1, 2, \dots, m\}$ was then partitioned into two clusters using k -means clustering to produce a binary image by following the steps described in Sec. III-D. The proposed method was implemented in MATLAB 7.10.0 (R2010a) and tested on an Intel® Core™ 2 Duo CPU T6600 @ 2.20 GHz, RAM 2.00 GB, 32-bit Windows 7 OS.

V. RESULT AND DISCUSSION

The results of the proposed method are shown in Fig. 5. The segmentation results using segmentation method proposed by Siswanto, et al. [18] and k -means clustering on the whole image are also shown in Fig. 6. It can be observed from the figure that the proposed method obtains good segmentation result for various images of food product. Although the image samples have inhomogeneous background, the proposed method can still produces good segmentation. As can be seen in Fig. 6 (b) and Fig. 6 (c), the results of the proposed method are similar to the results of segmentation method proposed by Siswanto, et al. [18].

Segmentation using k -means clustering on the whole image fails for some image samples, including red delicious apple, potato, chukanan mango, egg, ya pear, and packham pear, as shown in Fig. 6 (d). This failure can be caused by two factors. The first factor is the image has inhomogeneous background. Inhomogeneous background can caused a background pixel with the values of S and V component close to object pixel is grouped to object pixel. The second factor is k -means clustering reaches local minimum. This factor may occur if the image contains many outlier pixels, such as background pixels with very dark black color due to the influence of illumination or the present of shadow. If an outlier pixel is chosen as one of center cluster then this cluster only contain outlier pixels. These results indicate that the determination of ROI plays an important role in image segmentation using k -means clustering.

Table I summarizes MSE of the proposed method and the other two segmentation methods. It can be seen in Table I that

in average MSE of the proposed method (0.0013) is similar to MSE of segmentation method proposed by Siswanto, et al. [18] (0.0014) and lower than MSE segmentation using k -means clustering on the whole image (0.1399). This result shows that the proposed method has same performance compared with segmentation method proposed by Siswanto, et al. [18]. Moreover, the proposed method is more accurate than segmentation using k -means clustering on the whole image.

The proposed method is slightly slower than segmentation method proposed by Siswanto, et al. [18]. On average, the proposed method takes 0.479 seconds, while segmentation method proposed by Siswanto, et al. [18] takes 0.339 second to segment an image. On the other hand, the proposed method is much faster than segmentation using k -means clustering on the whole image which takes approximately 2.551 seconds to segment an image.

TABLE I. MSE OF PROPOSED METHOD (M_1), SEGMENTATION METHOD PROPOSED BY SISWANTORO, ET AL. [18] (M_2), AND k -MEANS CLUSTERING ON THE WHOLE IMAGE (M_3)

Sample image	Average MSE		
	M_1	M_2	M_3
Red delicious apple	0.0011	0.0017	0.0883
Green apple	0.0014	0.0012	0.0015
Potato	0.0012	0.0009	0.2582
Orange	0.0011	0.0015	0.0011
Tomato	0.0009	0.0007	0.0010
Chukanan mango	0.0013	0.0019	0.0852
Egg	0.0013	0.0008	0.5320
Ya pear	0.0015	0.0015	0.1690
Packham pear	0.0015	0.0014	0.2605
Carrot	0.0021	0.0022	0.0022
All images	0.0013	0.0014	0.1399

VI. CONCLUSION

A method for automatic food product image segmentation has been proposed. The method employed Sobel operator to find ROI which has homogeneous background. k -means clustering was used to partition ROI into two cluster, object and background, in HSV color space. The area outside ROI was considered as background. The validation results have shown that segmentation method based on k -means clustering combined with Sobel operator is similar to segmentation method proposed by Siswanto, et al. [18] and more accurate than segmentation on the whole image using k -means clustering.



Fig. 6. (a) original image (b) proposed method (c) proposed by Siswantoro, et al. [18] (d) *k*-means clustering on the whole image

ACKNOWLEDGMENT

The authors would like to thank Ministry of Education Malaysia and Universiti Kebangsaan Malaysia for financial support to conduct this research under Grants No. ERGS/1/2013/ICT07/UKM/02/2.

REFERENCES

- [1] D. T. Pham and R. J. Alcock, *Smart Inspection Systems Techniques and Applications of Intelligent Vision*: Academic Press, 2003.
- [2] A. S. Prabuwo, R. Sulaiman, A. R. Hamdan, and A. Hasniaty, "Development of Intelligent Visual Inspection System (IVIS) for Bottling Machine," in *TENCON 2006. 2006 IEEE Region 10 Conference*, 2006, pp. 1-4.
- [3] K. Patel, A. Kar, S. Jha, and M. Khan, "Machine vision system: a tool for quality inspection of food and agricultural products," *Journal of Food Science and Technology*, pp. 1-19, 2011.
- [4] C.-J. Du and D.-W. Sun, "Recent developments in the applications of image processing techniques for food quality evaluation," *Trends in Food Science & Technology*, vol. 15, pp. 230-249, 2004.
- [5] C. Zheng and D.-W. Sun, "2 - Image Segmentation Techniques," in *Computer Vision Technology for Food Quality Evaluation*, D.-W. Sun, Ed., ed Amsterdam: Academic Press, 2008, pp. 37-56.
- [6] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, pp. 62-66, 1979.
- [7] I. E. Sobel, "Camera models and machine perception," Ph.D. Dissertation, Stanford University, 1970.
- [8] C.-J. Du and D.-W. Sun, "Multi-classification of pizza using computer vision and support vector machine," *Journal of Food Engineering*, vol. 86, pp. 234-242, 2008.
- [9] T. Y. Wang and S. K. Nguang, "Low cost sensor for volume and surface area computation of axi-symmetric agricultural products," *Journal of Food Engineering*, vol. 79, pp. 870-877, 2007.
- [10] L. Fernández, C. Castellero, and J. M. Aguilera, "An application of image analysis to dehydration of apple discs," *Journal of Food Engineering*, vol. 67, pp. 185-193, 2005.
- [11] K. Kılıç, İ. H. Boyacı, H. Köksel, and İ. Küsmenoğlu, "A classification system for beans using computer vision system and artificial neural networks," *Journal of Food Engineering*, vol. 78, pp. 897-904, 2007.
- [12] M. Soltani, M. Omid, and R. Alimardani, "Egg volume prediction using machine vision technique based on pappus theorem and artificial neural network," *Journal of Food Science and Technology*, pp. 1-7, 2014/04/10 2014.
- [13] C.-J. Du and D.-W. Sun, "Estimating the surface area and volume of ellipsoidal ham using computer vision," *Journal of Food Engineering*, vol. 73, pp. 260-268, 2006.
- [14] C.-J. Du and D.-W. Sun, "Pizza sauce spread classification using colour vision and support vector machines," *Journal of Food Engineering*, vol. 66, pp. 137-145, 2005.
- [15] M. Khojastehnazhand, M. Omid, and A. Tabatabaefar, "Determination of orange volume and surface area using image processing technique," *Int. Agrophys.* 23, pp. 237-242, 2009.
- [16] A. B. Koc, "Determination of watermelon volume using ellipsoid approximation and image processing," *Postharvest Biology and Technology*, vol. 45, pp. 366-371, 2007.
- [17] M. Omid, M. Khojastehnazhand, and A. Tabatabaefar, "Estimating volume and mass of citrus fruits by image processing technique," *Journal of Food Engineering*, vol. 100, pp. 315-321, 2010.
- [18] J. Siswanto, A. S. Prabuwo, A. Abdullah, and B. Idrus, "Monte Carlo Method with Heuristic Adjustment for Irregularly Shaped Food Product Volume Measurement," *The Scientific World Journal*, vol. 2014, p. 10, 2014.
- [19] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed.: Prentice Hall, 2002.
- [20] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques: Concepts and Techniques*, 3rd ed.: Elsevier Science, 2011.
- [21] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering," *Proceedings of the IEEE*, vol. 67, pp. 773-785, 1979.
- [22] J. L. Marroquin and F. Girosi, "Some extensions of the k-means algorithm for image segmentation and pattern classification," *DTIC Document*1993.
- [23] H. Ng, S. Ong, K. Foong, P. Goh, and W. Nowinski, "Medical image segmentation using k-means clustering and improved watershed algorithm," in *Image Analysis and Interpretation, 2006 IEEE Southwest Symposium on*, 2006, pp. 61-65.
- [24] T.-W. Chen, Y.-L. Chen, and S.-Y. Chien, "Fast image segmentation based on K-Means clustering with histograms in HSV color space," in *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, 2008, pp. 322-325.
- [25] S. Ray and R. H. Turi, "Determination of number of clusters in k-means clustering and application in colour image segmentation," in *The 4th international conference on advances in pattern recognition and digital techniques*, 1999, pp. 137-143.
- [26] A. Z. Chitade and S. Katiyar, "Colour based image segmentation using k-means clustering," *International Journal of Engineering Science and Technology*, vol. 2, pp. 5319-5325, 2010.
- [27] C. Wen, Y. Q. Shi, and X. Guorong, "Identifying Computer Graphics using HSV Color Model and Statistical Moments of Characteristic Functions," in *Multimedia and Expo, 2007 IEEE International Conference on*, 2007, pp. 1123-1126.
- [28] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing A Practical Approach with Examples in Matlab*. West Sussex, UK: John Wiley & Sons, Ltd, 2011.
- [29] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*, International Student Edition ed.: Thomson Learning, 2008.