Bottle Label Segmentation
Based on Multiple Gradient Information

Yanjuan Chen[1], Sangcheol Park[1], Inseop Na[1], Soohyung Kim[1]
[1]School of Electronics and Computer Engineering
Chonnam National University
Gwangju, Korea

Myungeun Lee[2]
[2]Medical Research Center, Seoul National University
Seoul, Korea

ABSTRACT

In this paper, we propose a method to segment the bottle label in images taken by mobile phones using multi-gradient approaches. In order to segment the label region of interest-object, the saliency map method and Hough Transformation method are first applied to the original images to obtain the candidate region. The saliency map is used to detect the most salient area based on three kinds of features (color, orientation and illumination features). The Hough Transformation is a technique to isolated features of a particular shape within an image. Therefore, we utilize it to find the left and right border of the bottle. Next, we segment the label based on the gradient information obtained from the structure tensor method and edge method. The experimental results have shown that the proposed method is able to accurately segment the labels as the first step of product label recognition system.

Keywords: Hough Transformation, Saliency Map, Local Structure Tensor, Bottle Label Segmentation.

1. INTRODUCTION

Recently, lots of studies pay attention to the smart phone, and a considerable number of researches have been done for the smart phone image processing, such as flower recognition system [1], wine text recognition system [2] and so on. Especially, the text segmentation results affect the accuracy rates of the text recognition systems. However, text segmentation using a whole product image is not a good approach because of complex background. Therefore, we first pay attention to the label region of product bottle which contains the text region. However, it is also challenge task to segment the label in the product bottle images taken in supermarkets because of uneven illumination, strong reflection, hand and other unexpected factors.

There are a large number of methods have been developed for image segmentation. For example, the simplest method, thresholding method, is based on threshold value to turn a gray-scale image into a binary image; the most well-known method is Otsu's method [3]. The clustering methods is also used for segmentation based on distance between a pixel and a cluster center, this algorithm depends on the initial set of clusters and the number of the clusters, such as k-means algorithm [4].

Alternatively, there are many region based techniques for image segmentation, such as region growing methods, split-and-merge methods, and so on. The region growing method is the seeded region growing method. This method takes a set of seeds as input along with the image; the regions are iteratively grown by comparing all unallocated neighboring pixels to the regions. The main limitation of many region based algorithms is the permeation to area where the boundaries between the objects are weak or blurry.

In recent years, curve propagation is a popular technique on image analysis for object segmentation, object tracking, stereo reconstruction. It is to evolve a curve towards the lowest potential of a cost function, where its definition reflects the task to be addressed and imposes certain smoothness constraints. The level set method [5], [6] can be used to efficiently address the problem of curve propagation in an implicit manner. To the similar intensity or concave of the region, it is difficult to segment. And another method has to be mentioned, graph based method. It also can be effectively used for image segmentation based on a cost function. In these methods, the image is modeled as a weighted, undirected graph. Good examples of these methods are normalized cut [7], minimum cut/max flow cut [8] and others. However, the segmentation result depends on the choice of seeds (background and foreground) by users.

These methods described above relay on the assumption that

---
* *Corresponding author. E-mail : chenyanjuancc@hotmail.com*

the regions approximate normal distributions. They often get a relatively worse segmentation results for bottle label captured by mobile phones due to complexity of label region. Therefore, we propose methods to segment the label region of the bottle in the images taken by mobile phones in the supermarket with lot of the light sources based on multiple gradient approaches. First, saliency map is used to select out the most salient region, and the Hough Transformation is used to detect the two straight lines of left and right border to obtain candidate label region. Next, we calculate the local structure tensor of candidate label region, this gradient information and Sobel gradient information are used to detect the label region by 8-directional chain code.

The rest of this note is organized as follows. The pre-processing of system is described in section 2. In section 3, we present the process of candidate label region and object segmentation according to the structure tensor and edge information. The experimental results are shown in section 4 and at last we discuss the conclusion.

2. PREPROCESSING

The saliency map [9], [10] is designed as input to the control mechanism for covert selective attention. A method to location the most salient area is proposed by Itti et al [10] based on computing the position of the maximum in this map by the Winner-Take-All mechanism. Due to the complicated background including uninteresting objects, cabinets and hands holding the product bottle, it is not trivial task to segment an interesting object in our experimental images taken in the supermarkets. To take out complicated background, we first generate a saliency map on the input image Fig. 1(b) and then apply a threshold to the image in order to draw the boundary of the interest region or salient area.

Boundary of the label on the product bottle generally consists of straight lines. Hence, we use the Sobel edge detector to roughly detect straight lines in the interesting region detected by the saliency map such as Fig. 1(d). However, a challenging problem of the object segmentation is the light reflection, because the images are obtained in the supermarkets under the many light sources on the ceiling and most products are made by glass. Therefore, strong light reflection is an unavoidable and undesirable factor preventing from accurately segmenting the interesting objects in the images. For example, many small regions caused by light refection are shown in the original image in Fig. 1(a) and they could affect border detection. In order to reduce these interferences, we label the edge map to sub-regions according to a connect component analysis. Small sub-regions are removed as noise depending on their size and density in Fig. 1(e).

3. LABEL REGION SEGMENTATION

In this section, we divide the contents into two sub-sections, candidate label region generation and label region detection. In the first sub-section, the Hough transform is used to obtain the left and right border to general the candidate label region; and the detection process is described in the second sub-section.
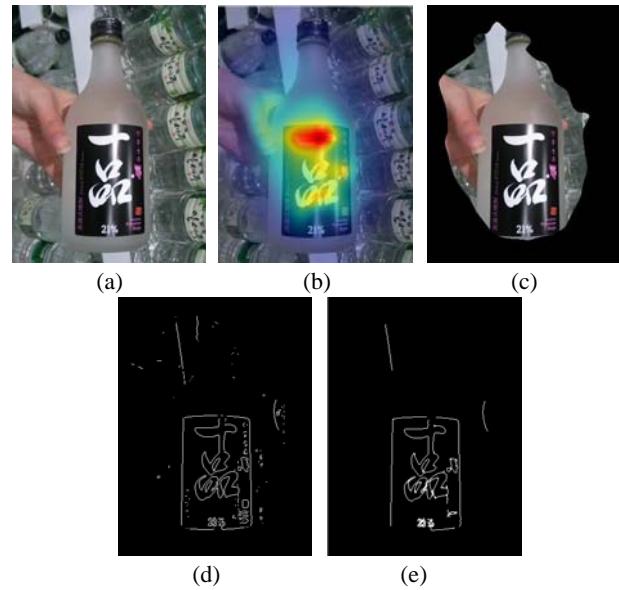

(a)   (b)   (c)

(d)   (e)

Fig. 1. Saliant region including an object of interest and edge map on the region; (a) input image taken by a mobile phone, (b) the saliency map overlaped on the input image, (c) salient region after binarization using a threshold value set to 55% of saliency ma, (d) edge map on the salient area by the sobel edge detector, and (e) results of the noise removal depending on the size of the edge connected component.

3.1 Candidate Label Region Generation

The edge detector can describe the boundary information of an interesting object. However, other factors such as discontinuous edge features or human's hand can produce the unexpected results. Especially, the border of bottle label region forms mostly straight line. Therefore, we use a Hough Transformation detect straight lines in the restricted area defined by a saliency map and edge detector. For example, the transformation is applied to all gradient pixels on the edge to convert to Hough space.

The Hough Transformation [11], [12] is a technique that can be used to isolated feature of a particular shape within an image. It is most commonly used for the detection of regular curves such as lines, circles, ellipses and so on. In this paper, we just consider Hough Transformation method for straight line detection. We can analytically describe a line segment in a number of forms; however this representation fails in case of vertical lines. In [13], R. Duda and P. Hart explored the fact that any line in the *xy* plane, shown in Fig. 2, can be described by a following convenient equation with parametric and normal notion:

$$\rho(\theta) = x \cdot \cos\theta + y \cdot \sin\theta \quad (1)$$

Where $\rho(\theta)$ is the length of a normal from the origin to this line, and θ is the normal orientation of $\rho(\theta)$ with respect to the x-axis; for any point (x, y) on this line, $\rho(\theta)$ and θ are constant. Applying Hough Transformation for each point will produce a

sinusoidal curve in the $\rho-\theta$ space, called parameter space shown in Fig. 3.
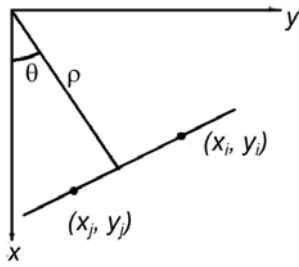

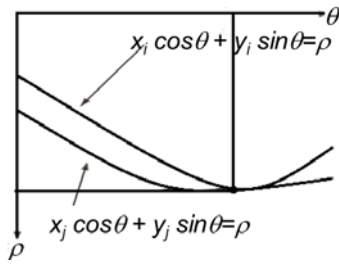Fig. 2. Parametric description of a straight line

Fig. 3. The result of applying Hough Transformation for (x_j, y_j) and (x_i, y_i) points

However, not only two straight lines are detected but also some other lines are detected in the image due to the complex backgound or the light reflection. Hence, we select only two straight lines based on the observation of that the two staight lines on the left and the right boudnary of the label are somehow similar in intensity value as the left and right borders of the label region such as Fig. 4. Therefore, the user's hand could be removed by the selected staight lines. Finally, we are able to crop the interesting region using a horizontal projection anlaysis.
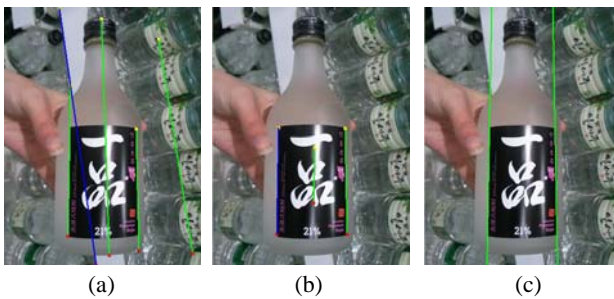

| (a) | (b) | (c) |

Fig. 4. Comparison of the results of the line detection (a) without and (b) with applying the saliency map, respecvtively and (c) the final staight line detection results after selecting two similar lines

3.2 Label Detection

The most difficult task of the label segmentation is the upper and lower boundary curve detection of the product bottle. A simple region-based method can be used to separate the relatively simple label region from the bottle product. However, label region of interest-object is very complicated; it contains not only characters, but also multiple objects with various colors to beautify the label region. These make the label region

segmentation much more difficult. Therefore, in this paper, we apply a local structure tensor to detect the exact shape of the label to accomplish label region segmentation.

The local structure tensor [13] is a different method of representing gradient information. It summarizes the predominant directions of the gradient in a specified neighborhood of a point. It can more accurately describe the local gradient characteristics, and distinguish between the isotropic and uniform cases.

For input image I with two variables values x, y, I_x and I_y are obtained from partial derivatives of x and y as following Eq. (2) and Eq. (3). The local structure tensor matrix is defined by the following equation (4) as per with taking advantage of I_x and I_y values,

$$I_x = \frac{\partial I(x,y)}{\partial(x)} \tag{2}$$

$$I_y = \frac{\partial I(x,y)}{\partial(y)} \tag{3}$$

$$S = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{4}$$

Here, to obtain the gradient information from the local structure tensor matrix, Eigen-decomposition is then applied to the matrix 'S' to general the eigenvalues (λ_1, λ_2) and eigenvectors (e_1, e_2) respectively. e_1 is a unit vector directed normal to the gradient edge while e_2 is tangent direction. And when λ_1 is larger than λ_2, the position has predominated direction that is aligned with gradient by e_1.

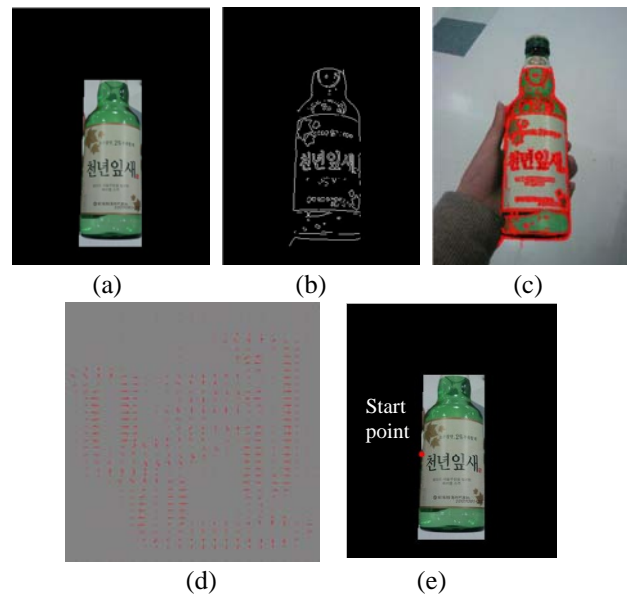
| (a) | (b) | (c) |

| (d) | (e) |

Fig. 5. Example of local structure tensor (a) input candidate label region image (b) sobel edge of candidate label region (c) structure tensor map (d) structure tensor of part region (e) candidate label region with start point by red

The local structure tensor results are shown in Fig. 5. To a candidate label region Fig. 5(a), its Sobel edge map and other gradient information based on the local structure tensor with orientation information by red arrows is shown in Fig. 5(b) and Fig. 5(c) respectively. Fig. 5(d) depicts the expand tensor map of the character "년". Compared with Sobel edge detector, the local structure tensor obviously provides more gradient information on the boundary of the label.
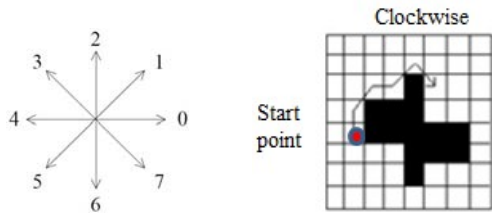

Fig. 6. 8-directional chain code

In order to obtain the exact shape of label of the product bottle, chain code [14] is utilized according to the Sobel edge information and structure tensor information. Depending on the different connectivity definition, there exist 4-directional chain code and 8-directional chain code. Here, we select the 8-directional chain code. First, we set the start point (initial position) of retrieval to the center of left border of candidate label region which obtained from previous sub-section as shown in Fig. 5(e) denoted by red. And then along the clockwise direction, retrieval begins from the start point illustrated in Fig. 6. In this paper, we retrieve the boundary of label region based on the Sobel edge and structure tensor information by comparing their orientation and magnitude in a mask with 3 by 3.

4. EXPERIMENTAL RESULTS

The high-definition mobile phone was used to take the product bottle images in the supermarket. In this paper, a set of 120 images whose dimension is equally 300×400, are used to test and evaluate the performance of the proposed methods.

Although user's hand was connected to an object of interest in an input image, Hough Transformation was able to detect the straight lines on the left and right border of the product label as shown in Fig. 7(b). Besides, the irregular shape of the label region is the most difficult task. For example, upper boundary of the label region consists of the red roof and irregular step line as shown in the middle image of Fig. 7(a). However, our label detection methods based on the Sobel edge and the structure tensor was able to draw a very accurate boundary curve. And the third case in Fig. 7(a) is normal case. All of them have been detected successfully by the proposed methods.

The segmentation results are shown in Table 1. The results are classified into two groups: Success, Partial Success and Failure. Success indicates that the methods segment correctly the label regions. Partial Success means that the segmentation results contain the main label region, and we classified it into two sub-groups again, Over-Segmentation and Under-Segmentation. Over-Segmentation is the results not only contain

the label region correctly, but also contain some un-label region. Under-Segmentation is the results are the main part of the whole label region, where the most typical example is due to the multiple-label. The Failure means that the results do not contains the label region because of the ambiguity and low contrast around boundaries.

(a)

(b)

(c)

Fig. 7. Example of successful results (a) a set of original Images and (b) the detected shape of obejct by green line (c) the segmentation results

Table 1:   Results of product bottle label segmentation

	Total	Success	Partial Success		Failure
			Over	Under	
Number	120	101	2	7	10
Accuracy (%)	100	84.16	7.50		8.34

The segmentation results are compared with the region based active contour, grow cut and our proposed approach in Fig. 8. Region based active contour method is senstive to the intensity. It has difficuty to segment the label region with un-even illumination. Grow Cut is based on color information. It could obtain the relatively good segmentation results based on the selection of seeds(background and foreground) by user.

(a)

(b)

(c)

(d)

Fig. 8. Comparsion with other methods (a) a set of original images
(b) region based active contours (c) grow cut method (d) our proposed
method

5. CONCLUSION

In this paper, we proposed a method to segment the label region in bottle images captured by a smart phone camera. These images are taken in normal indoor lighting environment, such as supermarket. The saliency map and the Hough Transformation methods are used to detect out the candidate label region and reduce the computational time for the local structure tensor. And then we use the 8-adjacence chain code method to detect the label region according to gradient information which is provided by the local structure tensor and Sobel edge map. Experimental results have shown that, our proposed method has been successfully applied to detection for bottle product images and achieved a high detection accuracy of 91.66% under relatively complex background. Although the results are encouraging, our study has several limitations. The method hardly works on the multiple label regions or the label region similar to the background. In addition, strong reflection often negatively affects the segmentation results. Therefore, we plan to extend our works to more complex bottle images with multiple regions, low contrast and light reflection.

REFERENCES

[1]   S. W. Hong. L. Choi, "Automatic Flowers Recognition Using Segmentation," Korea Computer Congress, Vol. 38, No. 1(A), 2011, pp. 463-465.

[2]   J. S. Lee, S. H. Kim, and J. H Park, G. S Lee, H. J Yang, C. W. Lee, "Recognition of Text in Wine Label Images," *IEEE Pattern Recognition on Chinese Conference*, 2009, pp. 1-5.

[3]   N. Otsu., "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems*, *Man, and* Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.

[4]   J. B. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1967, pp. 281-297

[5]   T. F. Chan and L. A. Vese, "Active Contours Without Edge," *IEEE Transactions on Image Processing*, Vol. 10, No. 2, 2001, pp. 266-277.

[6]   M. Rousson, and R. Seriche, *A Variational Framework for Active and Adaptative Segmentation of Vector Valued Images*, Proceeding of IEEE Workshop on Motion and Video Computing, 2002.

[7]   J. B. Shi and J. D. Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, No. 8, 2000, pp. 888-905.

[8]   Y. Boykov, Vladimir Kolmogorov, "An Experimental Comparison of Min-Cut/Flow Algorithms for Energy Minimization in Vision," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 26, No. 9, 2004, pp. 1124-1137.

[9]   B. C. Ko, and J. Y. Nam, "Object-of-interest image segmentation based on human attention and semantic region," Optical Society of Society of America, Vol. 23. Oct. 2006, pp. 2462-2470.

[10]   L. Itti, C. Koch, and E. Niebur, "A Model of Saliecny-Based Visual Attention for Rapid Scene Analysis", *IEEE Trans. Pattern Anal. Mach*. Intell. 20, 1998, 1254-1259.

[11]   D. Ballard, "Generalizing the Hough Transform to Detection Arbitray Shape," Pattern Recognition Vo. 13, No. 2, 1981. pp. 111-122.

[12]   R. Duda and P. Hart, "Use of the Hough Transformation

to Detect Line and Curves in picutures," Communication of the ACM, Vol. 15, No. 1, Jan. 1972, pp. 11-15.

[13] H. Knutsson, "Representing Local Structure Using Tensor," Proceeding of the 6th Scandinavian Conf. on Image Analysis. 1989. pp. 244-251.

[14] R. C. Gonzalez, E. W. and S. L. Eddins, *Digital Image Processing using MATLAB*, Publishing House of Electronics Industry, 2002

Yanjuan Chen
She received the B.S in computer science (Information Security) from Yangzhou University, China in 2010. Since then, she has been a master process student in the Department of Computer Science, Chonnam National University, Korea. Her main research interests include pattern recognition, product image processing and medical image processing

Sangcheol Park
He received received his B.S. and M.S degree in Computer Science from Chosun University, Korea in 1999 and 2001, respectively and his Ph.D. degree in Computer Science from Chonnam National University, Korea in 2006. From 2006 to 2010, he was a member of research staff in Medical Imaging Center, Department of Radiology, University of Pittsburgh. Since 2010, he has been a research professor in the Department of Computer Science, Chonnam National University, Korea. His research interests are medical image processing, pattern recognition, content-based image retrieval, image matching and 3D image reconstruction.

Inseop Na
He received his B.S., M.S. and Ph.D. degree in Computer Science from Chonnam National University, Korea in 1997, 1999 and 2008, respectively. Since 1999, he has been a CEO of in Infovalley Ltd., Korea. His research interests include image processing, pattern recognition, character recognition and digital library.

Soohyung Kim
He received his B.S. degree in Computer Engineering from Seoul National University in 1986, and his M.S. and Ph.D. degrees in Computer Science from Korea Advanced Institute of Science and Technology in 1988 and 1993, respectively. From 1990 to 1996, he was a senior member of research staff in Multimedia Research Center of Samsung Electronics Co., Korea. Since 1997, he has been a professor in the Department of Computer Science, Chonnam National University, Korea. His research interests are pattern recognition, document image processing, medical image processing and ubiquitous computing.

Myungeun Lee
She received the B.S., M.S. and Ph.D. degrees in Electronics Engineering from Mokpo National University, Korea, in 1998, 2001 and 2007, respectively. From 2007 to 2011, she was a postdoctoral fellowship and a research professor in the Department of Computer Science, Chonnam National University, Korea. Since 2011, she has been a senior member of research staff in Medical Research Center, Seoul National University College of Medicine, Korea. Her research interests are medical image processing, pattern recognition, and 3D image reconstruction.