International Journal of Engineering & Technology, 8 (1.9) (2019) 157-160

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Implementation of Maximally Stable Extremal Region for Text Segmentation on Food Package

Marlindia Ike Sari[1*], Rini Handayani[2], and Sidiq Maulana[3]

1,2,3 Computer Engineering, School of Applied Science, Telkom University, Bandung 40257, Indonesia
**Corresponding author E-mail: ike@tass.telkomuniversity.ac.id*

Abstract

Many technologies help people with vision disabilities. It helps these people to walk, read, and other activities. However, when these people have to shop and there is no one to help, they can't determine what product in their hand is. In this research, we applied image processing to recognize the name of the product based on text in the food package. We applied Maximally Stable Extremal Region (MSER) for text segmentation in the food package. The model is applied to predict the result of implementation MSER for text segmentation in the food package. We implement MSER with input from camera board in Raspberry Pi. Light intensity and type of food package give a different result. The result of this research has average 85%. It shows that MSER works for text segmentation in the food package.

Keywords: *Text Segmentation, MSER, image processing, food package*

1. Introduction

Food packages have many types of background, simple and complex ones. The font that used in food package has many types and size. This is consistent with the characteristic of natural scene image with text information (Chucai Yi, 2012). As well as natural scene image, text information in food package generally has similar or same color in many cases. Thus, text character could be separated from the background.



(a) (b)

Fig. 1: Example of food package (a) complex background (b) simple background

Fig 1 is the example of food package with complex and simple background. Food package with complex background has many color and texture on its background. Food package with simple background has solid color as background. This system builds to help people with low/no vision abilities to determine what product in their hand. This research only discusses about segmentation text area. The full system needs more research to make this system ready and available used by people with low/no vision abilities.

MSER is one of method for segmentation that conforms to the character of the natural scene image. MSER can detect similar pixel by select salient region among extremal region on an image ((Petra Bosilj). The MSER was introduced for the first time in (J. Matas, 2002). In (Matas, 2006) MSER was used for object detection. In (Pedro Martins a, 2016) was discussed about features in MSER. In (Kethineni Venkateswarlu, 2015) MSER used for text detection, but the input image was from a digital image.

In (J. Matas, 2002) the algorithm of MSER was shown in the quotation below:

In this research, we applied MSER for segmentation text are on image from food package that generally contains text information, and the image was captured by camera board on Raspberry Pi. The result of this system was displayed on computer.

In this research, we applied:

• MSER for text segmentation on food package.

Copyright © 2019 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

• Input image captured from camera board in Raspberry Pi.

• The system was processed by general computer.

The rest of the paper is organized as follows: Section Material and Methods discusses design and implementation of the system. Section Results and Discussion discusses experimental results. Summary of this paper and future research is discussed in latest section.

2. Material and Methods

This system was build using Computer and Single Board Computer (SBC), Raspberry Pi. Food package captured by camera board that installed on Raspberry Pi. Captured image sent to the computer using communication network via wired communication (using Ethernet cable). The computer was used for system process and display output. The block diagram of the system is shown in Fig. 2.

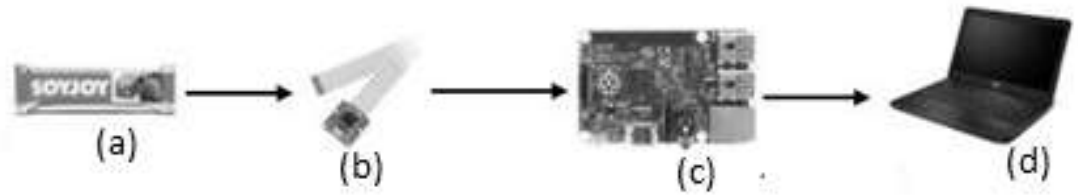


Fig. 2: Hardware component that be used to implement system (a) food package (b) camera board (c) Raspberry Pi (d) computer

The specification hardware in this research is:

1. Raspberry Pi:

    a. Type: Raspberry Pi 2 Model B

    b. CPU: 700 MHz ARM1176JZF0-S core(ARM11 family)

    c. SoC Broadcom BCM2835 (CPU,GPU,DSP and SDRAM)

    d. GPU : Broadcom Video IV, OpenGL ES2.0, 1080p30 h.264/MPEG-4 AVC high-profile decoder

    e. Memory : 512Megabytes (MiB)

    f. SD Card : SanDisk Ultra Micro SD Class 10 32GB

    g. Operating System: Raspbian

2. Camera Board:

    a. Small board size : 25mm x 20mm x 9mm
    b. A 5MP (2592X1944 pixels) Omnivision 5647 sensor in a fixed focus module
    c. Support 1080p30, 720p60 and, 640x480p60/90 video record

3. Computer:

    d. Processor : Intel Core i3
    e. Memory : 4 GB
    f. Harddisk : 500GB
    g. Operating System : Windows 7 Ultimate 64-bit
    h. GPU : Intel HD Graphics Family, 1GB

The input image was processed to get the segmented area. The processed was shown in Fig. 3.

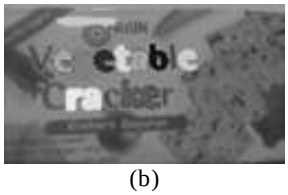


Fig. 3: Image processing

The original image was a color image with RGB format color. The size of the image was 2592X1944pixel. The resolution of the image depends on camera setting. In this research, we use default setting of camera board. The quality of the image depends on light intensity and distance object (food package) to camera board. In this research distance between object and camera board was 10-12 cm, and did on around 1 PM, in outdoor with the sun as light for the image.



(a)



(b)



(c)



(d)

Fig. 4: Pre-processing image (a) Original image in Grayscale format (b) Detect, Extract and Visualize the MSER Regions in grayscale image (c) Convert MSER pixel lists to a binary mask (d) Erode a binary image with a disk structuring element

The next process was converting the RGB image to gray scale image (see in Fig 4(a)) and then do MSER process. The output was a binary image. If the MSER process succeeded, only main text area that gets value '1' in the binary image. The highlight of output MSER was shown in Fig 4(b) and the output binary image in Fig 4 (c). The process that was showed in Fig 4 (b) and (c) is implementing the process of MSER methods. Erode process was used to get a clear image from the useless pixel. The output of erode is the segmented text area that ready to be input for another process i.e text recognition.

3. Result and Discussion

The result of implementation this system shows in Table 1 and 2. Percentage of the segmented area means that comparison of segmented text area to supposed segmented text area. The illustration of calculation percentage of segmented area is shown in Fig. 5



(a)



(b)



(c)

Fig. 5: Illustration for calculation result percentage (a) original image (b) supposed segmented text area (b) tested segmented text area

Fig 5 (a) is the original image, (b) is supposed segmented area, contain all text 'Vegetable Crackers', 17 fonts. But in (c) only 14 fonts that into the segmented area. The calculation of segmented area is (14/17)*100%=82.35%

The system was tested to 10 food packages, with each package tested 24 times. These 10 examples are chosen because they represent different type of backgrounds and fonts. Each sample was tested 24 times with different light and angle.

Table 2: is example of calculation percentage segmented text area.

Segmented Area=(∑(Segmented text/supposed segmented text)*100%)/testing times

Table 2: Example of calculation percentage in 'Vegetable Cracker"

Name of product	Number of test	Segmented Area (number of font)		Percentage
		Supposed	Testing	
Vegetable Cracker	1	17	9	53%
	2	17	14	71%
	3	17	12	82%
	4	17	15	88%
			Average	74%

Table 3: Result of System Testing

Name of Product	Image of Food Package	Segmented Area
Vegetable Cracker		74%
Broniz		92%
ABC		67%

Name of Product	Image of Food Package	Segmented Area
Buavita Guava		96%
Mie ABC		96%
Soyjoy		94%
TARO		87%
Ultra Milk		81%
Roma Malkist		63%
Potatoz Krekez		100%

From Table 3, shows that the best result is 'Potatoz Krekez's image with 100% segmented text area. "Potatoz Krekez" image have a simple background. The image with minimum result is 'Roma Malkist's Image with 63% segmented text area. 'Roma Malkist's image have complex background. The result of this testing give an average of result is 85%.

The result needs to increase. The weakness of camera board that used in this research is no auto focus, and the captured image has to be in ideal condition i.e good light intensity, distance, and type of food package.

3. Conclusion

MSER can implement for text segmentation of food package. The next research is increasing the percentage of segmented text area and give more feature such as text recognition (may use OCR methods) and text to speech. Portable version or mobile version needed to build for easy accessibility for needed people.

References

[1] Chucai Yi, a. Y. (2012). Localizing Text in Scene Images by Boundary Clustering, Stroke Segmentation, and String Fragment Classification. IEEE TRANSACTIONS ON IMAGE PROCESSING, 4256-4268.
[2] J. Matas, O. C. (2002). Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. Proc. Of British Machine Vision Conference (BMVC), (pp. 384–396).
[3] Kethineni Venkateswarlu, S. M. (2015). Text Detection on Scene Image Using MSER. International Journal of Research in Computer and Communication Technology, 452-456.
[4] Matas, ˇ. S. (2006). Object Recognition using Local Affine Frames on Maximally Stable Extremal Regions. In Toward Category-Level Object Recognition, (pp. 83-104).
[5] Pedro Martins a, P. C. (2016). On the completeness of feature-driven maximally stable extremal regions. Pattern Recognition Letter, 9-16.
[6] Petra Bosilj, E. K. (n.d.). Beyond MSER: Maximally Stable Regions using Tree of Shapes.