

# OCR IN ROUND SURFACE BASED ON COMPONENT LABELING: SURVEY PAPER

Sonal M. Thakkar  
*P.G Scholar*  
*CSE Department*  
*Gujarat technological university,*  
*Ahmedabad, India*

Prof. Kalpesh. L. Patel  
*Professor*  
*CSE Department*  
*Gujarat technological university,*  
*Ahmedabad, India*

## ABSTRACT

*Optical Character Recognition (OCR) refers to the process of translating, the handwritten or printed text into a format that is understood by machines for the process of editing, searching, and indexing. The recognition of character in round surface from scanned images of documents has been problem that has received much attention in the field of image processing, pattern recognition and artificial intelligence. In this method, find angle of every character in image and then recognize the character. The proposed method has three flows: Font file Teaching flow, Running Teaching flow and verification flow. In font file teaching flow create the font file using the image processing technique. In Running Teaching flow save reference data and parameters from font file and match character. In verification flow match all segments with reference data. Component labeling method used make segmentation easy to separate the single line in round surface into characters. The proposed approach is implemented using the Opencv2.4.4 and visual studio 2008.*

**Keywords** - *Component labeling, Image processing technique, Optical Character Recognition.*

---

## 1. INTRODUCTION

Optical Character Recognition (OCR) systems have been numerously developed for the recognition of printed script in many languages. The recognition of characters from scanned images of documents has been problem that has received much attention in the Field of image processing, pattern recognition and artificial intelligence. Classical methods in pattern recognition do not suffice for the recognition due to the size, shape, and style [1].

Here, an image can be obtained from the digital camera. OCR was preferable by researchers due to its high accuracy compared to other techniques [2]. OCR was used to extract to characters and alphanumeric from the digital image [2]. In this method, recognition system was simulated using OPENCV and tested on real digital image. There are several steps to make this system works: capture the image using digital camera, extract the character and numbers in an image by applying the algorithm or recognition techniques on selected round area, apply image segmentation technique to get individual character and finally using OCR application to recognize each character. The researches applied image processing techniques such as image acquisition, image pre-processing, character

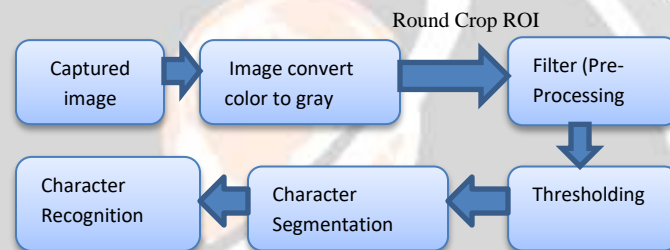
segmentation and character recognition on the captured digital image of cosmetic product in round surface. Segmentation was commonly used in the procedure to extract the character and numbers from the image after filtering all the distortions such as noise, space mark, light variance and etc. After segmentation, template matching in character recognition is applied to recognize each letters in selected area.

The study introduces the application of image processing technique in OPENCV to recognizing the character and numbers in 180 degree circle of images. This is the main objective of this study. The development of GUI for round surface images recognition system becomes the second objective of this study.

Limitations of this study are: (1) it is not possible for 360 degree circle. (2) It is also not possible for multiple lines in round surface images.

## 2. METHODOLOGY

The methodology for this study consists of three major stages. The stages are pre-processing, segmentation and matching process. The pre-processing stages includes the processing of raw image in round surface which is captured by using digital camera until obtaining specific part of an image. Next step is the segmentations process where the process used for detecting characters on the image using the contour tracking method. Finally matching process is applied to recognize the character from the `extracted image with real characters. In fig.1 shows the block diagram of our methodology.



**Fig. 1: Block diagram of our methodology**

### 2.1 Collection of image

In this study, samples of image captured using digital camera will be created and checked for the quality of the samples such as clear image, less disturbance and good view of image. If the image is accepted, the image will be analysed in OPENCV by applying image processing technique. If the image is rejected, a new image will be taken. The process will be continuous until obtaining a good image before proceeding to the next steps. In order to process image, image processing toolbox will be used for pre-processing and filtering image.



**Fig. 2: sample of round surface image**

## 2.2 Pre-processing of image

The image originally in RGB format of color will be converted to black and white format. It will help in identifying the selected region of image and minimize the number of color used in image. After filtering the image, the character will be identified by using segmentation technique. The purpose of this step is to recognize the character of the image. Some features will be extracted to obtain accurate character recognition. In pre-processing digital image, the color (RGB) image was converted to gray code first before converting the image into black and white in order to minimize the number of color used for each image in round surface.

## 2.3 Thresholding

**Thresholding** is the simplest method of [image segmentation](#). From a [grayscale](#) image, thresholding can be used to create [binary images](#). The simplest thresholding methods replace each pixel in an image with a black pixel if the image intensity  $I_{ij}$  is less than some fixed constant  $T$  (that is,  $I_{ij} < T$ ), or a white pixel if the image intensity is greater than that constant. But it may not be good in all the conditions where image has different lighting conditions in different areas.

In this study, we are used adaptive thresholding method. In this, the algorithm calculates the threshold for small regions of the image. So we get different thresholds for different regions of the same image and it gives us better results for images with varying illumination. In Gaussian adaptive threshold method, threshold value is the weighted sum of neighborhood values where weights are a Gaussian window.



Fig. 3: Original image



Fig. 4: result of adaptive Gaussian thresholding

## 2.4 Character segmentation

In character segmentation, we will use connected component labeling using contour tracing for detect any characters in the images. Using contour tracing technique to extract the contour of a binary image and also detect the angle of the character in round image. For example, if we're given a pattern like that of [Fig 5](#) below, the contour traced by the algorithms will be similar to the one shown in [Fig 6](#) (the blue pixels represent the contour). This could be acceptable in some applications but in other applications, like character recognition, we would want to trace the interior of the pattern as well in order to capture any holes which identify a certain character. ([Fig 7](#)

below shows the "complete" contour of the pattern) As a result, a "whole searching" algorithm should be used to first extract the holes in a given pattern and then apply a contour tracing algorithm on each hole in order to extract the complete contour.

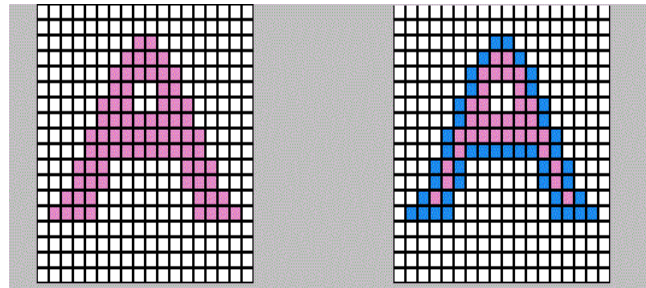


Fig. 5

Fig. 6

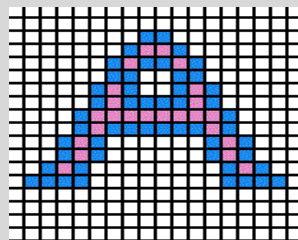


Fig. 7

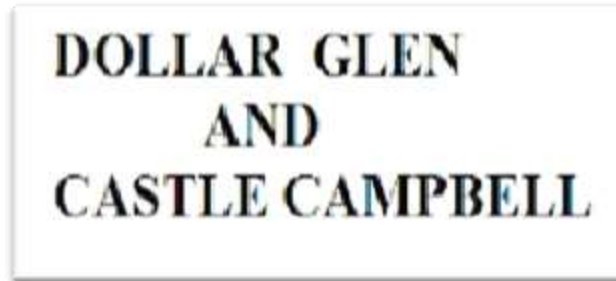
After the contour tracing, segment the all character in the image. In fig.8 shows example of character segmentation.



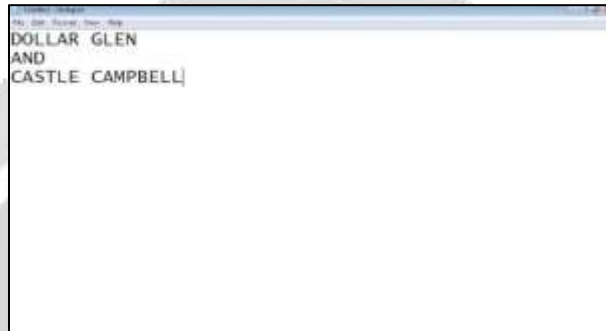
Fig. 8: Example of character segmentation

**2.5 Character Recognition:**

After the segmentation process, we will apply the recognition for recognize the character in the images. Template matching in character recognition is applied to recognize each letters in the figure. The results are displayed in notepad in terms of text as illustrated in fig.9



**Fig. 9: Example of character recognition**



**Fig. 10: Results of applying OCR**

### 3. Application of simple OCR

- Assistive technology for blind and visually impaired users.
- Data entry for business documents
- Automatic reading of postal mails for sorting of bank checks, bills and other documents
- Automatic number plate recognition
- Language processing such as spells checking, indexing, grammar checking etc.

### 4. Application of OCR in round surface

- Indian gas regulator
- Cosmetic product in round surface

### 5. Conclusion & future work

As a conclusion, a captured image of the cosmetic product in round surface was chosen as a sample image to be used in pre-processing technique in order to recognize the characters of image in round surface. Three main techniques were applied in pre-processing technique. There are cropping technique filtering technique and morphology technique. The characters and numbers in image are correctly identified. Besides, the GUI is successfully working and displaying necessities information and results. This study use for only single line and only maximum 180 degree circle.

For the future work, might need to improve the incapability of the system to multiple line and 360 degree circle.

## 6. References

- [1]Mrs. B. Vani, Ms. M. Shyani Beulah and Mrs. R. Deepalakshmi, “High accuracy Optical Character Recognition using learning array of ANN”. 2014 International Conference on circuit, Power and Computing Technologies[ICCPCT].
- [2]Norizam Sulaiman, Sri Nor Hafidah Mohammad Jalani, “Mahfuzah Mustafa, Kamarul Hawari, Development of Automatic Vehicle Plate Detection System” , 2013 IEEE 3<sup>rd</sup> International Conference on System Engineering and Technology 19-20 Aug,2013, Shah Alam, Malasiya.
- [3]Archana S. Sawant , Prof. D.G.Chougule,” Script Independent Text Pre-processing and Segmentation for OCR,International Conference on Electrical” , Electronics, signals, Communication and Optimization(EESCO)-2015 IEEE.
- [4]Rohollah Mazrae Khoshki, Subramaniam Ganesan, “Improved Automatic License Plate Recognition (ALPR) system based on single pass Connected Component Labeling (CCL) and reign property function” , IEEE-2015.
- [5]Sushruth Shastry, Gunasheela G, Thejus Dutt, Vinay D S and Sudhir Rao Rupanagudi, “ ‘ i ‘ – A novel algorithm for Optical Character Recognition(OCR)” , IEEE-2015.
- [6]Mande Shen, Hansheng Lei, “Improving OCR Performance with Background Image Elimination” , 2015 12<sup>th</sup> International Conference on fuzzy Systems and knowledge diuscovery(FSKD) IEEE2015.
- [7]Ayatullah Faruk Mollah, Nabamita Majumder, Subhadip Basu and Mita Nasipuri, “Design of an Optical Character Recognition System for Camera bades Handheld Devices”,International Journal of Computer Science Issues, vol, 8, Issue 4, No 1, July 2011 Issn:1694-0814.
- [8]Takahiro Ota, Toshikazu Wada, “Classification based character segmentation guided by Fast-Hessian-Affine regions” ,IEEE 2011
- [9]W.T Chan, C.P. Tso, T.Y. Lo, K. S. Sim,” A method in Applying OCR in Processing Of Thermal Images”, 2011 IEEE International Conference on signal and image processing Applications(ICSIPA2011).
- [10]Deepayansarkar, “optical Character Recognition using neural networks,” University of wisconsin – Madison.
- [11] “Optical Character Recognition: a guide to the frontier”, George nagy, Thomas A. Nartker, Stephen V. Rice, Procs Document recognition VII, Vol: 3967 58-69.
- [12] Simon Haykin, Neural Networks: Acomprehensive foundation, 2ndEdition, Prentice Hall, 1998
- [13] Optical Character Recognition using Artificial Neural Networks RakeshBhujade, BLB-International Journal of Science & Technology Vol.1, No. 2 (2010), 143-152 (ISSN 0976-3074)
- [14] Online Handwritten Character Recognition Using an Optical Back propagation Neural Network by Walid A. Salameh Princess Summaya University for Science and Technology, Amman, Jordan
- [15]S.R. Dubois and F.H. Glanz (1986), “An Auto regressive model for two dimensional shapeclassification”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume8(Issue 1), pp. 55-66.