

Text Extraction and Character Recognition from Image using Mathematical Morphology and OCR Technique

Sukhvinder Singh¹, Surender Kumar Grewal²

¹M. Tech (Student), Department of ECE, DCRUST, Murthal (Sonepat), Haryana, India

²Associate Professor, Department of ECE, DCRUST, Murthal (Sonepat), Haryana, India

Abstract: Images contain various types of useful information that should be extracted whenever required and this information may be in the form of text present in image. Extraction of this information involves detection, localization, extraction, enhancement and recognition of the text from the given image. Mathematical morphology is the foundation of morphological image processing, which consists of a set of operators that transform images according to the characterization. We use some fundamental morphological operations such as dilation, erosion, opening and closing for text extraction and OCR technique for character recognition from image.

Keywords: Dilation, Erosion, Structuring element, F-Score and OCR

1. Introduction

Text Extraction and recognition in Images has become a potential application in many fields like Image indexing, Robotics, Intelligent transport systems etc. For example capturing license plate information through a video camera and extracting license number in traffic signals. This whole process can be broadly divided into three sub parts (a) Text localization, (b) Text extraction, (c) Text recognition. In this paper, we have presented an approach for text extraction and recognition present in images. First of all the input image is filtered by using the Median filter to remove any noises if noise present in image. Then edges of the text region are detected using Canny edge detector. Then the morphological operations dilation and erosion are applied for object localization process, then all the connected Components are then extracted which shows the text part present in image and all non-text character components are then discarded. Features are then extracted from the connected components. Then OCR technique is used to recognize these extracted components. OCR is a powerful technique for the recognition process.

1.1 Related work

The characters on the scene images are difficult to be extracted, because they are intricate with complicated background, and show various shape, size, and direction. Although the importance of this field have led a number of researchers to pay attentions to it and there have been several studies in the character extraction, they are proposed to be optimal ideas for extraction characters, but are limited to simpler background conditions. . B. Marcotegui , J. Fabrizio and M. Cordhas suggested a region based approach [2]. In that first of all text strings are decomposed to letters, and then that letters are grouped to restore words after recognition using SVM. Region based approach is found to be more efficient than other approaches; this approach was ranked first in ImageEval campaign. But the authors accepted that this method also produces lot of FP. CC based or Edge based approaches also have been used for text localization

process. But the main disadvantage in using these approaches are like increase of FP, increase in processing time etc. but they are much easy to implement. Mohammad Shorif Uddin, Madeena Sultana, Tanzila Rahman, and Umme Sayma Busra[1] examined in 2012 their main objective is to extract text from scene images. In this approach they discussed an effective approach for detecting and extracting text from scene images based on morphological features. Nirmala Shivananda and Nagabhushan [6] proposed a hybrid method for separating text from color document images. But this method can't extract text from complex graphics. The algorithm of Antani and R.Kasturi [10] works well for text string separation from mixed text/graphics image, but it makes an impractical assumption that character components in a string are aligned straight and does not touch or overlap with graphics. Ming zhao, Shutao, James Kwak[9] in 2010 had implemented a text detection in images using sparse representation with discriminative dictionaries as a result F-score of 78 is obtained. Yi-Feng Pan, Cheng-Lin Liu, Xinwen Hou[3] in 2010 proposed a system for text extraction from images by learning based filtering and verification .With this approach F-score of 68 is achieved. In our method, we have used an edge based technique combined with mathematical morphological operations for text localization and OCR technique for text recognition.

2. Methodology

Our methodology is divided into four subparts:

- 1) Image Preprocessing
- 2) Text localization and Extraction.
- 3) Text/Non text classification.
- 4) Character Recognition

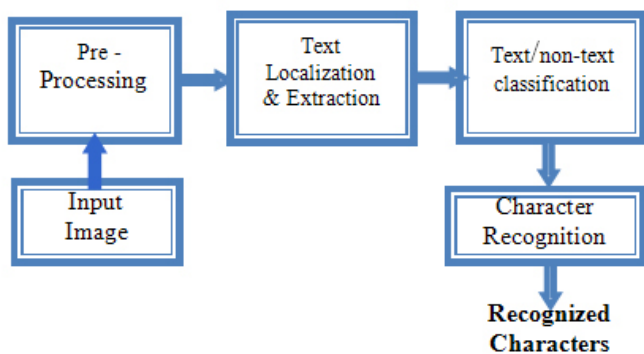


Figure 1: Methodology block diagram

2.1 Image Pre Processing

In this subpart we firstly read image from which text has to be extracted. Then Convert the colored image to gray image to avoid processing overload, after the conversion into gray scale we remove background of image to decrease complexity of image [6]. At last we remove the noise from image by using morphological filtering.



Figure 2: Image preprocessing results

2.2 Text localization and Extraction

In this subpart first of all Conversion of gray image into binary image by defining the threshold of the image is done. After that edges of the components are detected, these components may have text part or non text part, which later on determined on basis of a three way step. After the edge detection morphological operations are applied to the image. Firstly dilation and then erosion of image is done. We choose the structuring element in accordance to the cluster of image text[1]. Generally we preferred rectangular structuring element of dimensions (2*3). After that connected components are extracted and bounding boxes are placed around them.

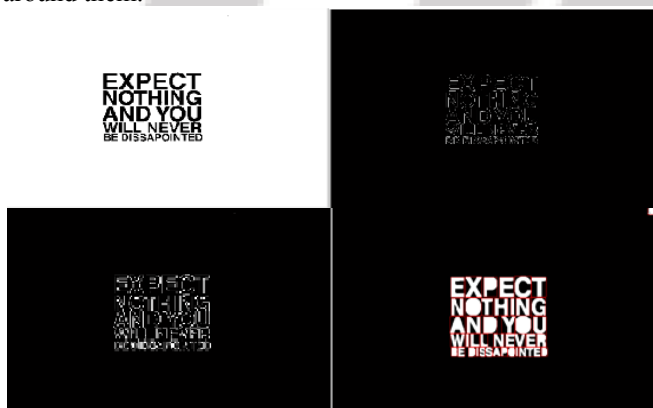


Figure 3: Results of text localization and extraction.

2.3 Text/non-text classification

All these extracted components may contain both text components and Non-text components. They are differentiated by a two way process [5]. First, the initial Bounding Boxes are drawn for all objects, then we used connected components matrices to eliminated the non text part. Thus finally remains the bounding boxes only around the text part only.

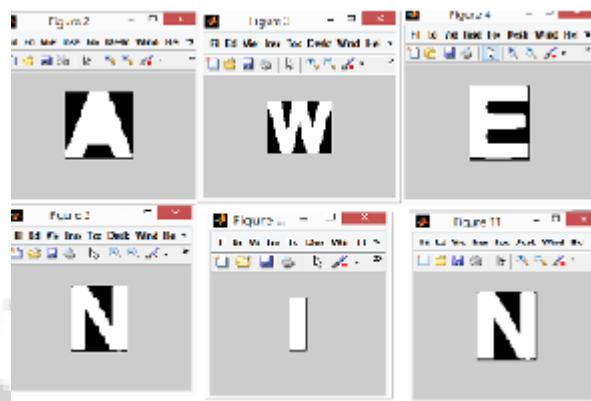


Figure 4: Extracted characters for image

2.4. Character recognition

2.4.1. Optical Character Recognition

The goal of Optical Character Recognition (OCR) is to classify optical patterns (often contained in a digital image) corresponding to alphanumeric or other characters[1]. OCR is computerized processing to recognize individual character is required to convert scanned document into machine encoded form. Computer system equipped with such an OCR system can improve the speed of input operation and decrease some possible human errors. Recognition of printed characters is itself a challenging problem since there is a variation of the same character due to change of fonts or introduction of different types of noises. Difference in font and sizes makes recognition task difficult if pre processing, feature extraction and recognition are not robust. There may be noise pixels that are introduced due to scanning of the image. Besides, same font and size may also have bold face character as well as normal one. OCR technique enables us to successfully recognize the text extracted from an image and convert it into an editable text document.

2.4.2. Clustering of input images

A single threshold value is not sufficient for our experiment. So, we divide test images into three clusters depending on text style, shape, and size[1].

- Cluster1: Containing images of small size text
- Cluster2: Containing images of medium size text
- Cluster3: Containing images of large size text.



Figure 5: Cluster 1 type image

JUDAS
PRIEST
775758
HOLA
DIEGO
1 2 3 1 2
3 6 7 9 4 5

Figure 6: Cluster 2 type image

PLUS
SIZE
PRINTS

Figure 7: Cluster 3 type image

3. Performance Evolution and Results

3.1 Performance evolution

It can be done on the basis of following parameters.

- 1) False Positives (FP) / False alarms are those regions in the image which are actually not characters of a text, but have been detected by the algorithm as text.
- 2) False Negatives (FN) / Misses are those regions in the image which are actually text characters, but have not been detected by the algorithm.
- 3) Precision rate (p) is defined as the ratio of correctly detected characters to the sum of correctly detected characters plus false positives as represented in equation below

$$p = \frac{\text{correctly detected characters}}{\text{correctly detected characters} + \text{FP}}$$
- 4) RRC(Recall rate) = (No. of extracted characters/ No. of characters in image) x 100.
- 5) F-score is the harmonic mean of the recall and precision rates [9].

3.2 Results

Proposed methodology is applied over 100 images on different type of clusters and following results are obtained.

Table 1: Precision rate (p)

Type	Cluster1	Cluster2	Cluster3	Cluster4
Total Character	348	247	198	793
Correctly Detected	262	196	171	629
P(%)	75.28	79.35	86.36	79.31

Table 2: Recall Rate(RRC)

Type	Cluster1	Cluster2	Cluster3	Cluster4
Total Character	348	247	198	793
Correctly Detected	271	217	186	694
P(%)	77.87	87.85	93.40	84.99

Table 3: F-Score

Type	Cluster1	Cluster2	Cluster3	Cluster4
RRC	75.28	79.35	86.36	79.31
p	77.87	87.85	93.40	84.99
F-score	76.57	83.60	89.80	82.15

4. Conclusions

In this paper, we have proposed an improved and robust approach for text localization and extraction technique from images. The proposed method is tested with various clusters of images, both images with the caption text and the scene text. All related methods which are given in references are analyzed and the drawbacks are tried to reduced and thereby getting an improved version of the previous works. In this work, we obtain reduced noise levels and comparatively high F-score, precision rate and RRC value. Main stress has been given in reducing false positives.

References

- [1] Mohammad Shrif Uddin, Madeena Sultana, Tanzila Rahman and Umme Sayma Busra, "Extraction of texts from a scene image using morphology based approach", IEEE Transactions on image processing, 978-1-4673-1154/12(2012)
- [2] B. Marcotegui, J. Fabrizio, M. Cordhas "T-HOG: An effective gradient-based descriptor for single line text regions", R. Minetto et al. / Pattern Recognition 1078-1090, 46 (2013)
- [3] Yi-Feng Pan, Cheng-Lin Liu, Xinwen Hou, "Fast scene text localization by learning based filtering and verification", IEEE 17th International Conference on Image Processing, 2010 September 26-29, 2010
- [4] F. Mollah, S. Basu, M. Nasipuri, D. K. Basu, "Text/Graphic separation for business card images for mobile devices", IAPR International Workshop on Graphics Recognition, pp. 263-270, (2009).
- [5] Partha Pratim Roy, Josep Lladós and Umapada Pal, "Text/Graphics Separation in Color Maps", IEEE Transactions on Image Processing, vol. 7, pp.545-551, (2007).
- [6] Nirmala Shivananda and P. Nagabhushan, "Separation Foreground Text from Complex Background in Color Images", IEEE Transactions Processing vol.10, 306-09
- [7] Henk J.A.M. Heijmans and Jos B.T.M. Roerdink (Eds), "Mathematical Morphology and its Applications to

Image and Signal Processing”, Publisher: Springer, 1st Edition, pp. 3374,2009.

- [10] Lukas Neumann, Jiri Matas, "A method for text local localization and recognition in real-world images", ACCV, Nov. 8-12(2010)
- [11] Ming zhao, Shutao, James Kwak, “Text detection in images using sparce representation with discriminative dictionaries”,Image and Vision computing, 2010.
- [12]D. Crandall, S. Antani, and R. Kasturi, “Robust Detection of Stylized Text Events in Digital Video”, Proceedings of International Conference on Document Analysis and Recognition, pp. 865-869(2001)

Authors Profile

Sukhvinder Singh is pursuing M.Tech form Department of ECE, DCRUST, Murthal (Sonapat), Haryana, India

Surender Kumar Grewal is working as Associate Professor, Department of ECE, DCRUST, Murthal (Sonapat), Haryana, India

