International Journal of Scientific Engineering and Research (IJSER)
www.ijser.in
ISSN (Online): 2347-3878
Volume 1 Issue 1, September 2013

Text Localization and Extraction in Images Using Mathematical Morphology and OCR Techniques

Rishpa Sachdeva[1], Pooja Nagpal[2]

[1]M. Tech (Student), Department of ECE, JCDM College of Engineering (Sirsa), Haryana, India

[2]Associate Professor, Department of ECE, JCDM College of Engineering (Sirsa), Haryana, India

Abstract: ***In the digitalization of the world, it becomes more and more important to extract text from image. Because text data present in images contain useful information for automatic annotation, indexing and structuring of images. Furthermore, text printed on the cover of magazine, signs, indicators, billboards etc always mixes with photos and designs. This kind of texts in scene images may take much information and thus need to separate text string from scene image. Hence, the extraction of texts in scene images is a difficult as well as challenging task. Recently, mathematical morphology based algorithm finds applications to extract texts from scene images. In this paper we proposed a technique for text extraction from an image. The process uses morphology and OCR techniques. The first, feature extraction stage analyzes the set of isolated characters and selects a set of features that can be used to uniquely identify characters. The performance depends upon the calculation of F-score. RRC and reduced noise level. Due to insufficiency of a single threshold value, we have divided input images into different clusters depending on the size of texts.***

Keywords: scene image, mathematical morphology, thresholding, closing and opening operation, text extraction, OCR

1. Introduction

In this paper, we have presented a robust approach for text extraction and recognition in image. Texts in scene images may contain much information and thus separation of text strings is an important Issue. It is observed that graphic backgrounds are commonly found in most business card images. In order to recognize the text information from the card, the text and background contents must be separated. Recognition of magazine cover can help in inserting information into the database quickly for library to improve the efficiency of classification. OCR [1] technique enables us to successfully extract the text from an image and convert it into an editable text document. Extraction of texts is important in form processing, map interpretation, bank cheque processing, postal address sorting, and engineering drawing interpretation. Hence, our main objective is to extract text from scene images. In this paper, we discuss an effective approach for detecting and extracting text from scene images based on morphological features [2].

For achieving better result we use Lixu Gu's [6] approach, which is based on mathematical morphology. We also focus on the limitations of Gu's approach and compare our method with this approach to achieve better results. The performance of this approach depends on the calculated values of various parameters like F-score, RRC [4] and noise level. Remainder of this paper is organized as follows: section 2 provides a description of the principle of the proposed technique. Description of Gu's method is discussed in section 2. Section 3 provides the limitations of Gu's algorithm.

Our modified method using OCR is described in section 5 and clustering of input images are described in section 6. In section 7, experimental results with scene images are discussed. Finally, the conclusions are drawn in section 8.

2. Principle of Text Extraction

Text extraction is more challenging due to many problems listed below intermingled texts with other objects, such as:

(a) Structural bars, company logos, and smears
(b) Slight difference between background and colour texts.
(c) Varying font, style, and size of the characters in texts.
(d) Uneven lighting conditions in scene images.

Here we describe and implement Gu's algorithm that uses mathematical morphology [3] to extract text effectively. This approach includes two distinct stages – Primary processing and Extraction processing. In primary processing, we give input image and use a new shape decomposition filter based on morphological recursive opening and closing. The input image can be colour or gray scale image. The proposed filter decomposes input image into several sub images based on the size of the characters. Extraction processing includes three steps – feature emphasis, character extraction and noise reduction. In feature emphasis [5] step, a new morphological filter used to emphasis characters features in sub images and removes most of the noise. By histogram method characters are extracted from sub images in character extraction step. Lastly, using a morphological filter based on closing is implemented for noise reduction.

3. Limitation of Gu's Approach

In This, Gu proposed a morphological filter based on morphological opening [7]. But we know that morphological opening removes the small regions, which are smaller than the structuring elements. Thus if we use opening for all sub-images some of the expected regions will be vanished with the increasing disk size. As a result we don't get all characters. To overcome this limitation in our experiment, an image is divided into a series of sub images according to the size of characters. Thus different sub images contain different graphical parts as noise but all have common text regions. If we intersect these sub-images we obtain an output image which contains only text regions.

4. Methodology

The proposed work is divided into four modules, and is shown in figure1

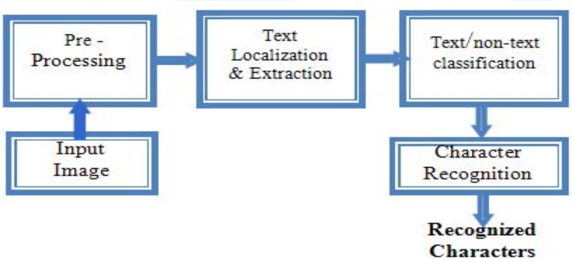

Figure 1: Methodology

Characters in scene images may be characterized by word, "they float on the background or sink under it with a flat surface". This characteristic underlines the basic idea of our new method. The aim of our new method is to separate the characters from the background. The whole process is divided into two distinct stages:

1. Primary processing
2. Extraction processing.

- Feature emphasis
- Character extraction
- Text/non text classification
- Character recognition
- Noise reduction

4.1 Pre-processing

In the pre-processing step, first the input RGB image is converted to gray-scale image. This conversion is done in order to reduce the processing overload. Median filtering is then applied to the Gray-scale image to remove any noises present in that.

4.2 Text Extraction

In this phase, the image obtained from the previous step is binarised. Then the Morphological dilation operation is performed on this image.

4.3 Text/non-text classification

All these extracted components may contain both text components and Non-text components. They are separated and eliminated by a two way process.

4.4 Character recognition

Characters are recognized by using OCR [1] and the samples of templates used are given in fig 2.

Figure 2: Samples of templates

4.5 Noise Reduction

In the end noise reduction is done by using a morphological filter.

5. 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.

Cluster 1: Containing images of small size text
Cluster 2: Containing images of medium size text
Cluster 3: Containing images of large size text

Figure 3: Cluster 1 type image

Figure 4: Cluster 2 type image

International Journal of Scientific Engineering and Research (IJSER)
www.ijser.in
ISSN (Online): 2347-3878
Volume 1 Issue 1, September 2013

Figure 5: Cluster 2 type image

Figure 6: Containing both Small and medium size text

(a) (b)

Figure 6: (a) and (b) Cluster 3 type images

6. Experimental Results

To evaluate the performance of our method for extracting texts from scene images, we used 50 images in our experiment to demonstrate the efficiency of our method Fig. 2 to 6 shows some sample images from three different clusters. The resultant images after digital negative are given in Fig.7 and Fig. 8, using our proposed and Gu's method. Resultant figures show that our method produces less noise than Gu's method. Some results of images of these three clusters are given below:

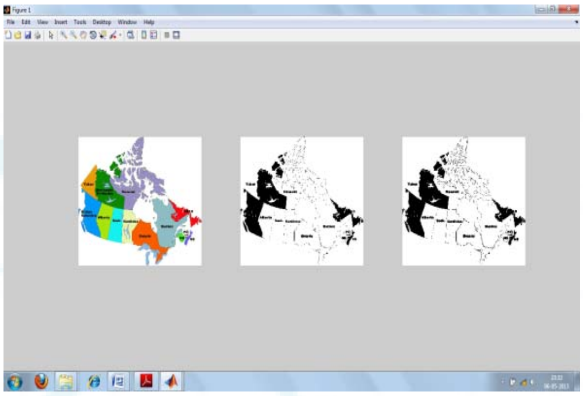

Figure 7: Cluster 1 [our Method]

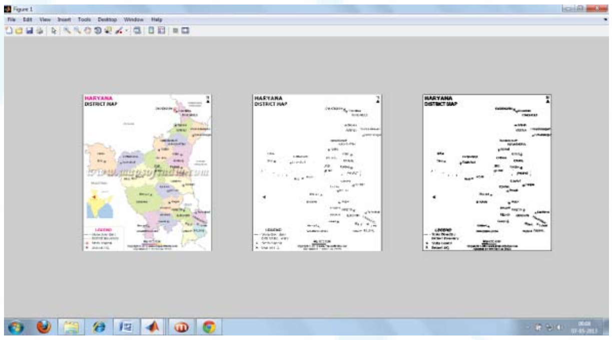

Figure 8: Cluster 1 [GU's Method]

Figure 9: Cluster 2[our method]

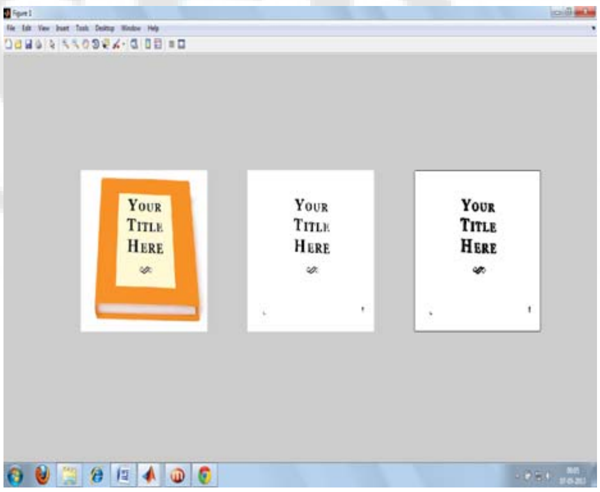

Figure 10: Cluster 3[our method]

Figure 11: Cluster 3[our method]

Table 1: Extraction Performance for Character strings

	Total	*Cluster 3*	*Cluster 2*	*Cluster 1*
Total Character string	152	10	57	85
Extracted Character string	141	9	53	78
RRC (Accuracy)	97.23%	98.70%	92.57%	87.89%
F-score	95.2	94.3	93.2	91.8
Noise level(threshold value)	0.45	0.65	0.57	0.64

Table 2: Shows the Comparison of our method with others

Parameters	*P*	*R*	*F-score*
Gu's method	87.8	89	91
Method[1]	95	92	93
Our method	96	95	94.3

7. Performance Evaluation

The performance of the proposed method is evaluated using precision, recall and f-score metrics and it is compared with the existing methods [9].

7.1Definition 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.

7.2Definition 2

False Negatives (FN)/ Misses are those regions in the image which are actually text characters, but have not been detected by the algorithm.

7.3Definition 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 = correctly detected characters / [correctly detected characters + FP]

7.4 Definition 4

Recall rate (r) is defined as the ratio of the correctly detected characters to sum of correctly detected characters plus false negatives as represented in equation below.

r = correctly detected characters / [correctly detected characters + FN]

7.5 Definition 5

F-score is the harmonic mean of the recall and precision rates. The proposed method is compared with the results of methods given in the references. From this it is observed that our method gives better results when compared to other techniques.

Table I shows the performance analysis extraction using our method and of. Increased accuracy figure confirms nice performance of our clustering method

8. Conclusion

In this paper, we have proposed an improved text localization and extraction technique from images. The proposed method is tested with various types of images, both images with caption text and scene text. All related methods given in references are analyzed and the drawbacks are reduced [10] and thereby getting an improved version of the previous works. In this work, we obtain reduced noise levels and comparatively high F-score value, main emphasis has been given in eliminating false positives [8], making it efficient for large size text and noise elimination which is the major drawback in Gu's based approach.

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] http://en.wikipedia.org/wiki/Digital_image_processing, (Accessed on: November 20, 2010).

[3] http://www.ciesin.org/docs/005-477/005-477.html, (Accessed on: November 22, 2010).

[4] http://pippin.gimp.org/image_processing/chap_dir.html, (Accessed on: November 22, 2010).

[5] http://java.sun.com/products/javamedia/jai/forDevelopers/jai1_0_1guide/unc/ Properties.doc.html (Accessed on: November 22, 2010).

[6] Nirmala Shivananda and P. Nagabhushan, "Separation of Foreground Text from Complex Background in Color Document Images", IEEE Transactions on Image Processing, vol.10, pp.306-309, (2009)

[7] 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).

[8] Partha Pratim Roy, Josep Llad´os and Umapada Pal, "Text/Graphics Separation in Color Maps", IEEE Transactions on Image Processing, vol .7, pp.545-551, (2007).

[9] Lixu Gu, Naoki Tanaka , R. M. Haralick and Toyohisa kaneko, "The Extraction of Characters from Scene Image Using Mathematical morphology", IAPR

Workshop on Machine Vision Applications, vol.2, pp.12-14, (1996) .
[10] Rafael Gonzalez, Richard E. Woods and Steven L. Eddins, "Digital Image Processing Using Matlab", Publisher: Prentice Hall, 2nd Edition.

Authors

Rishpa Sachdeva is currently pursuing master's degree in Electronics and Communication from JCDM College of Engineering Sirsa, India

Pooja Nagpal is Associate professor at JCDM College of Engineering, Sirsa, India