ORIGINAL ARTICLE

Automated Extraction of Text from Images using Morphology Based Approach

Md. Shorif Uddin[1], Tanzila Rahman[2], Umme Sayma Busra[3], and Madeena Sultana[4]*

ISSN: 2186-0114
http://www.IJEI.org

ARTICLE HISTORY

Received: 09 July, 2012
Revised: 12 August , 2012
Accepted: 13 August, 2012

Published online:

14th August, 2012

Vol.1, No.1, 2012

Abstract

Existing text strings play an important role in understanding a scene image. Scene images differ from document images, which are composed of text characters of various size, shape, direction, and situation along with complicated backgrounds, such as map, picture or painting, etc. Hence, the extraction of texts on scene images is a difficult as well as challenging task. Mathematical morphology based algorithm finds applications to extract texts from scene images. In this paper a modified morphological filter along with an automatic clustering technique is described. Experimental result confirms the superiority of this approach compared to existing morphological method.

Keywords: *Text extraction, Thresholding, OCR, scene image.*

I. INTRODUCTION

In real life situations, text data embedded in images contain useful information for automatic annotation, indexing and structuring of images, document analysis, vehicle license plate extraction, technical paper analysis, and object oriented data compression. Furthermore, text printed on the cover of magazine, signs, indicators, billboards etc. always mixes with photos and designs [1]. This kind of texts in scene images may contain much information and thus need to separate text strings from scene image.

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 [2]. It is useful for visually impaired persons and foreign travellers to recognize signs and indicators. In addition, recognition of magazine cover can help in inserting information into the database quickly for library to improve the efficiency of classification [1]. OCR software 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 image. In this paper, we discuss an effective approach for detecting and extracting text from scene image based on morphological features.

A variety of approaches have been proposed for extracting text from images. Nirmala Shivananda and Nagabhushan [1] proposed a hybrid method for separating text from color document images. But this method can't extract text from complex graphics. Partha Pratim Roy, Josep Llad´os and Umapada Pal [3] proposed a method for separating text from color map based on connected component analysis and grouping of characters in a string. This approach can detect the characters connected to graphics and can separate them. But some of the characters can't be separated through connected component analysis. The algorithm of Fletcher and Kasturi [4] 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

AUTHORS INFO

[1]e-mail: shorifuddin@gmail.com
Jahangirnagar University

[2*]e-mail: tanzila.rahman@yahoo.com
Jahangirnagar University

[3]e-mail: saymabushra@gmail.com
Jahangirnagar University

[4]e-mail: deena.sultana@gmail.com
Jahangirnagar University

*Corresponding author
e-mail: tanzila.rahman@yahoo.com
Tel: 08801711263933

or overlap with graphics. Coleman and Andrews [5] developed a method, which operates in an unsupervised mode with connected component analysis. But some text regions are vanished using connected component analysis.

Recently Lixu Gu [6] proposed a mathematical morphology based approach for separation texts from scene. However, it has some limitations. Hence, we focus on the limitations of Gu's approach and then propose a modified morphological filter. As text strings are of different shapes and sizes, we use an automatic clustering technique to cope with this issue. The 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 3. Section 4 provides the limitations of Gu's algorithm. Clustering of input images are described in section 5. In section 6, Our modified morphological filters are described and a noise removal approach is used in section 7. In section 8 experimental results with scene images are discussed. Finally the conclusion are drawn in section 9.

II. Principle of Text Extraction

Comparing to the documentary texts, extraction of the texts in scene images is much more challenging. A lot of researchers succeeded in extracting single text string from images, but cannot deal with images including many text strings [7] due to several problems listed below [8]:

- Different types of objects such as structural bars, company logos and smears may be mixed with characters.
- Difference between background and text color.
- The font, style, and size of the characters may vary.
- Uneven lighting conditions in scene images.

Here we describe briefly Gu's algorithm [6] that uses mathematical morphology to extract text effectively. Mathematical morphology provides us the theory and tools to capture geodesic information. Hence structure segmentation and shape representation have been performed by morphological approaches. Top-hats transformation (TT) is an excellent tool for extracting bright or dark objects from uneven background. But due to complicated segmentation problems, the TT alone cannot provide satisfactory solutions, when a series of structure elements with various sizes have been employed. In our study, the gray scale particle segmentation for some complex cases is realized successfully in an iterative manner. Here we try to implement a method which was derived based on TT idea[9]. A typical block diagram of the text extraction method using mathematical morphology is given in Fig. 1.

III. Gu's Text Extraction Technique

The whole text extraction process is divided into two distinct parts [6]:

i. Primary process
ii. Text extraction process.

A. Primary Process

The decomposition procedure is implemented by following morphological algorithm, which is applied to produce binary images.

$$X_i = \left|\left(X_0 - X_0 \circ r_i B_{Disk}\right) - \left(X_0 - X_0 \circ r_i B_{Disk}\right)\right|_B - X'_{i-1} \quad (1)$$

$$X'_j = \bigcup_{0<j\le i} X_i,\ X'_1 = \varnothing$$

$$X_i = \left|\left(X_0 \bullet r_i B_{Disk} - X_0\right) - \left(X_0 \bullet r_{i-1} B_{Disk} - X_0\right)\right|_B - X'_{i-1} \quad (2)$$

$$X'_j = \bigcup_{0<j\le i} X_i,\ X'_1 = \varnothing$$

Where, X_0 is the original gray scale image, X'_i is the decomposed binary image, and B denotes threshold operation in a defined value. Equation (1) is used to decompose the images where characters are brighter than the background; if the characters are darker than background, (2) will be applied. The procedure of this processing is start with r_iB_{Disk}, a series of sub-images X'_i are produced in a recursive manner. The processing will be stopped when the image $X_0 \circ r_iB_{Disk}$ or $X_0 \bullet r_iB_{Disk}$ have no characters remained.

B. Text Extraction Process

This is the most important part of Gu's method. Extraction processing is divided into 3 distinct steps:

1. Feature emphasis
2. Character extraction
3. Noise reduction / refinement

1) Feature Emphasis: In this stage, the decomposed sub images are processed by a morphological filter to reduce noises and emphasize character region.

$$E_i = \begin{bmatrix} (((X'_i \circ r_{i-1}B_{Disk}) \oplus r_{i+1}B_{Disk}) \circ r_{2i}B_{Disc}) \times X_0 & i \le 10 \\ (X'_i - X'_{10}) \times X_0 & i > 10 \end{bmatrix} \quad (3)$$

$$E_i = \begin{bmatrix} ((X'_{ki} \circ r_{i-1}B_{Disk}) \oplus r_{i+1}B_{Disk}) \circ r_{2i}B_{Disc} & i \le 10 \\ (X'_{ki} - X'_{k10}) & i > 10 \end{bmatrix} \quad (4)$$

2) Character Extraction: Since character regions are the main component in E_i, those hold the peak values in the histogram. The peak values which are bigger than the average of all peak values are searched and is the selected peak is used as threshold for E_i to extract characters from it. The extracted characters are in Hi.

3) Noise Reduction/Refinement: The extracted characters are broken in Hi and there remain diverse noises. A morphological filter derived from conditional dilation is implemented to refine the extracted characters.

$$R_{i0} = H_i \circ r_{i-1}B_{square}$$

$$R_{in} = (R_{i(n-i)} \oplus r_5 B_{Disc}) \bigcap |X_0|_B \quad (5)$$

if $R_{ik} = R_{i(k-i)}$ then stop.

Finally, sub-images R_{ik} are united to obtain the entire resultant image X_r denoted by

$$X_r = X_{r-1} \bigcup R_{ik} \quad (6)$$

IV. Limitation of Gu's Approach

In (3) and (4), Gu proposed a morphological filter based on morphological opening. 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.

Again in (6), all sub-images R_{ik} are united to obtain the entire resultant image X_r. But the resultant image contains same region more than once as a result of union and much noise remains.

V. Clustering of Input Images

A single threshold value is not sufficient for primary and text extraction process. So, we divide test images into three clusters depending on text style, shape, and size.

- Cluster 1: Consists of images with small texts such as color map.
- Cluster 2: Consists of images with medium texts such as business card.
- Cluster 3: Consists of images with large texts such as color cover image.

In our experiment, after converting the input images into gray images we used clustering approach. For this, we calculate total number of connected component using 8-point connectivity. Because image with large text has less number of text region than image with small text. Thus, Images of cluster 1 contain large number of component than cluster 2. Again, Cluster 3 contains less component than cluster 2. Table I shows the threshold scores, which are used in our experiment. These threshold scores are taken based on experimentation.

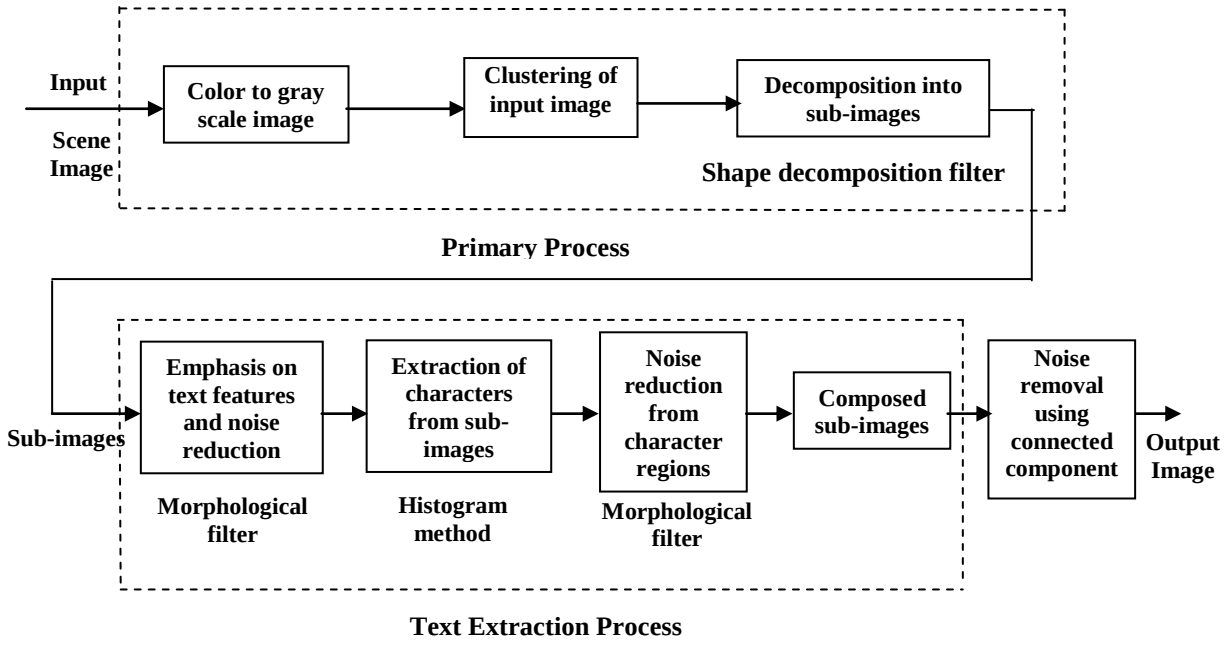


Fig. 1. Block diagram of text extraction method.

VI. Modified Morphological Filter

Instead of (3) and (4), we proposed a new morphological filter in (7), based on morphological closing for the sub-images where i≥2. This filter will be filled small holes which are smaller than the structuring elements.

$$E_i = \begin{bmatrix} (((X_i' \circ r_{i-1}B_{Disk}) \oplus r_{i+1}B_{Disk}) \circ r_{2i}B_{Disc}) & i=1 \\ (X_i' \bullet r_{i-1}B_{Disk}) & 2 \le i \end{bmatrix} \quad (7)$$

We also modified (6) as follows

$$X_r = X_{r-1} \bigcap R_{ik} \quad (8)$$

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 gain an output image which contains only text regions with some noise.

VII. Noise Removal

To reduce noise we use a filter based on label matrix obtained using 8 connectivity. Hence the area which are smaller than the threshold value will be removed from the output image.

The threshold value will be different for different cluster. As cluster 1 contains small region it's threshold value will be smaller than cluster 2. The threshold value for cluster 3 will be larger than cluster 2. In our experiment, we use threshold value 1 for cluster 1 and 2 for cluster 2 and 5 for cluster 3. Thus noise which are smaller than text region will be removed.

VIII. Experimental Results

To evaluate the performance of our method for extracting texts from scene images, we used total 60 images where each cluster consists of 20 images. All the images were taken from web. Gray scale and color cover images of books/journals/magazines, business card, color map, billboard are used in our experiment to demonstrate the efficiency of our method. We focused on this type of image because they have variations in font size, style and background color.

In our experiment, primary processing is conducted using (1) and (2), the stopping condition is preferred to j≤1.

TABLE I. The Values which are used in our experiment for Clustering the input images

	Cluster 1	**Cluster 2**	**Cluster 3**
Number of component	Larger than 120	71-119	Smaller than 70

(a) Original image with 495 x 310 pixels.

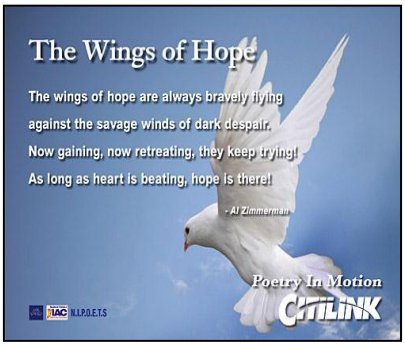
(b) Image after primary processing.

(c) Image after extraction process.

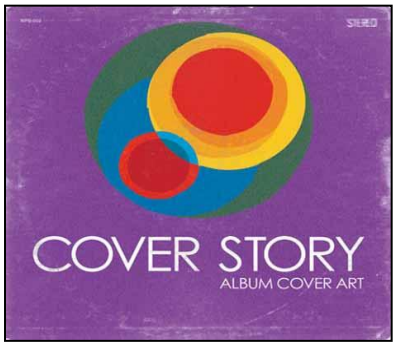
(d) Image after noise removal.

Fig. 2. Image of different steps for text extraction.

(a) Sample image from of cluster 1.

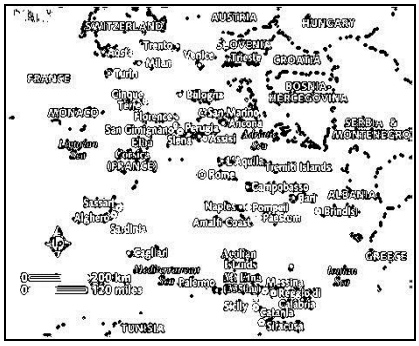


(b) Sample image from of cluster 2.

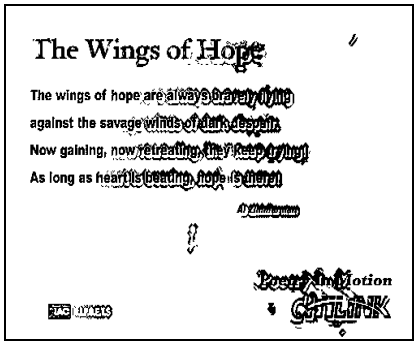


(c) Sample image from of cluster 3.

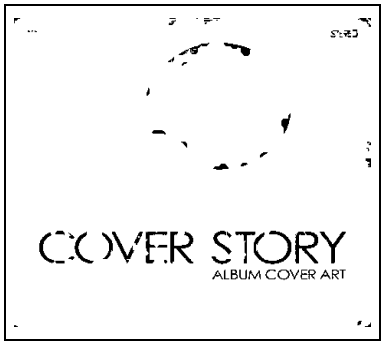
Fig. 3. Sample images of different clusters.



(d) Resultant image of cluster 1.

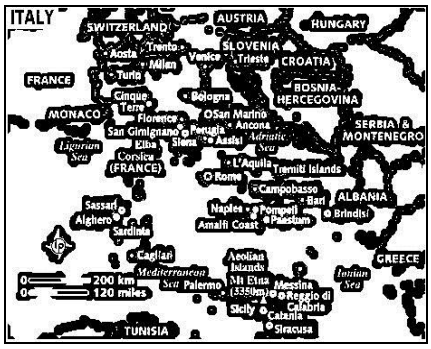


(e) Resultant image of cluster 2.



(f) Resultant image of cluster 3.

Fig. 4. Output images of different clusters shown in Fig. 3 using our method.



(g) Resultant image of cluster 1.

(h) Resultant image of cluster 2.

(i) Resultant image of cluster 3.

Fig. 5. Output images of different clusters shown in Fig. 3 using Gu's method.

Fig.2 contains images of different steps for extracting text.

Fig. 3 shows some sample images from three different clusters. The corresponding resultant images after digital negative are given in Fig.4 and Fig.5, respectively, using our proposed and Gu's method.

Above figures show that using our method the resultant images contain less noise than Gu's method. For quantification of accuracy, recall rate of characters are calculated to evaluate performance. Recall rate of characters (RRC) is defined as follow:

RRC = (No. of extracted characters/ No. of characters in image) × 100.

Table II shows the result of extraction performance using our method and Gu's method, which confirms the superiority of our technique.

TABLE II. Extraction Performance for Characters

	Our Method				**Gu's method**
	Total	**Cluster 3**	**Cluster 2**	**Cluster 1**	**Total**
Total Characters	420	23	63	334	420
Extracted Characters	396	23	63	310	375
RRC (Accuracy)	94.29%	100%	100%	92.81%	89.28%

IX. CONCLUSION

A modified method based on morphological approach for extraction of texts from scene images is discussed. For improving accuracy we proposed a modified morphological filter and also proposed a clustering method based on different text sizes. Moreover, our method reduces noise in the resultant image. Experimental result confirms the superiority of our approach compared to existing morphological approach. We hope that our method is useful for extracting texts from scene images. Our future concentration will be on the recognition of text characters using vertical and horizontal projections and converting text images into editable form.

ACKNOWLEDGMENT

Authors were supported by the Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh.

REFERENCES

[1] 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.

[2] A. 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.

[3] 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.

[4] L.A. Flecher and R. Kasturi, "A robust algorithm for text string separation from mixed text/graphics images," *IEEE Transactions on PAMI*, vol. 10, no. 6, pp. 910-918, 1988.

[5] G.B. Coleman and H.C. Andrews, "Image separation by clustering", in *Proc. of IEEE*, vol. 67, pp. 773-785, 1979.

[6] 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.

[7] Yuming Wang, Naoki Tanaka, "Text string extraction from scene image based on edge feature and morphology," *IEEE Transactions on Image Processing*, vol. 10, pp. 323-328, 2008.

[8] Kongqiao Wang, Jari A.Kangas, "Character location in scene images from digital camera," *Pattern recognition,* vol. 36, pp. 2287-2299, 2003.

[9] 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 Ed., pp. 367-374, 2000.

[10] Md. Shorif Uddin, Madeena Sultana, Tanzila Rahman and Umme Sayma Busra, "Extraction of Text from a scene Images using Morphology Based Approach", in *Proc. of IEEE Intl. Conf. on Informatics, Electronics & Vision (ICIEV)*, 2012, Dhaka, Bangladesh.