Study of Various Methods Used for Text Confining and Extraction from image

Harpreet Singh[1], Deepinder Singh[2]

[1, 2]Department of E.C.E, Punjab Technical University, Punjab, India

Abstract: *Text occur in any of the image play a significant role because it contains very useful information. But extraction of text in any image is not a simple due to difference in shape, size, alignment and complex background. There are various methods used for the text extraction process. In this paper we are reviewed or discussed the various techniques used for the text extraction and finding a technique with maximum efficiency or accuracy.*

Keywords: Text extraction, text localization, text confining, OCR

1. Introduction

In the text extraction process we extract all of the relevant text data from the image. Basically text present in any image contains very useful information which gives us overall idea about that image. Text extraction from image have various useful applications like detection of vehicle license plate, study of articles with maps, tables, diagrams and charts etc., name plate, object identification, image search based on keyword search, street signs, video indexing based on text, page segmentation, address block location etc. Images also can be categorized into various images like document image, scene text image and caption text image. A document image normally contains text with some graphic components as shown in the fig 1. Document image are obtained by scanning printed documents, journals, handwritten historical documents, book cover and degraded document image etc. Text extraction in document image with colored background is difficult because of complexity of background and mix up of fore-ground text color with background color.

Caption text is also known as cutline text or overlay text. In this type of image texts are artificially superimposed on the image at the time of editing. It is usually identify the subject of the image content. Caption text may be as moving text, rotating text, growing text and text of arbitrary size. Caption text image is as shown in the fig 2.

Scene text as shown in the fig 3 occurs in the scene which is captured by the recording device. Text occurs at a time when we shot a video is known as scene text. Basically scene text occur as a part of the scene and contains useful information like advertisement, name of streets, shops, road signs, board signs, food containers, bill boards, text on vehicles etc. generally text extraction from image involves text detection, text localization, text tracking, text binarization, extraction, image enhancement, and recognition of text from any image. The extraction of text from image plays an important role in various applications. But this is not a simple task. So it is a challenging issue to extract the text from any given image because of variation of character font, size, shape, style, text direction and complicated background.
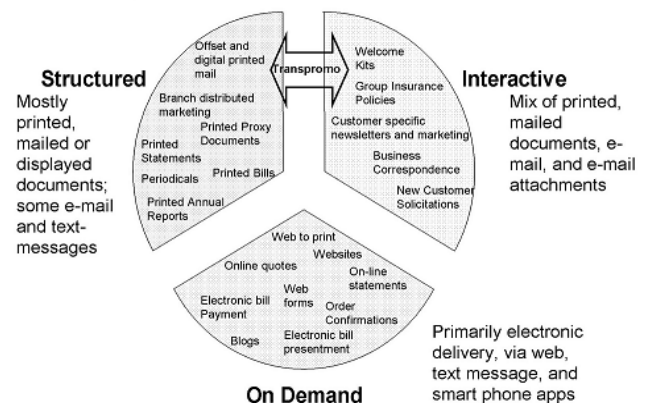

Figure 1: Document image

Figure 2: Caption text image

Figure 3: Scene text image

Before going to text extraction techniques we first study about the properties of the text like size, motion, color, edge etc.

1) **Size:** differences in text size can also make problem for text extraction, but it can be reduced by assuming specific data during text region detection process.
2) **Color:** intensity of color also affects the quality of text extraction. If all characters are of same color then text extraction processes become more effective and easy.
3) **Motion:** this property of text usually applied to the videos containing text, and this refers to the movement of text in vertical or horizontal direction.
4) **Edge:** edges are important and useful feature of text as compared with the other features like color layout or motion.

Basically there are various methods are used for the text extraction from an image. Various methods were based on the mathematical morphology, edge detection method, connected components method, histogram method, skeletonization operation etc. so in this paper we study all of the methods used for the text extraction from image.

2. Methods Used For Text Extraction

Various techniques are used to extract the text from any type of image. Generally we study about the text extraction from scene image only which is more challenging or difficult as compared to other type of images. So review of various techniques used for text extraction is given below:

Vaishnavi Ganesh and Dr.L.G. Malik proposed a method in which big data analysis has been done by using Google Apis. Then filters have been applied to the images for obtaining noiseless images. Then two methods have been applied in the image processing stage. The first method is Color based Partition. Here canny edge detector and k-means clustering have been performed. After this the trained classifiers are applied on the partitions to determine whether each cluster contains text or not. Then the text grouping method uses Hough transform to detect the text. By applying trained classifiers either time or efficiency will be improved.

Prof. Amit Choksi, Nihar Desai et. al proposed a method for text extraction from natural scene images using Prewitt edge detection method. The Edge based method detects text regions with highest Precision (100 %) & Recall (100 %) rate. Thus it is more efficient compared to that of the performance obtained with Gaussian edge detection method. We observed that Gaussian edge detection based methods is not good enough to detect the text regions. In order to add with that Prewitt edge-detector algorithm is more robust since it gives higher recall rate as compared Gaussian edge-detector algorithm. The edge based algorithm is also able to give better results in case of lighting variance compare to Gaussian edge detection method and also on illuminated images.

Prachi Mukherji and Priti Rege had used a new approach having shape features and fuzzy logic to recognize offline character recognition .They segmented the thinned character into strokes using structural features like endpoint, junction points, cross point and thinning. They

used tree and fuzzy classifiers and they gives an average 86.4 accuracy.

Yuming Wang,Naoki Tanaka in uses mathematical morphology to extract text effectively and edge border ratio is utilized to differentiate text region from noise region, using the edge contrast feature of text region in real scene .This proposed method also describes the method which can connect characters in to text strings and distribute text strings to different sub images according to their width of strokes. The algorithm is designed for scene image like signs, indicators as well as magazine covers and it is a robust method but this method does not give satisfied results of RRC value and correctly detected characters.

G.Rama Mohan Babu, P.Sarimaiyee, A.Srikrishna had proposed an algorithm which is insensitive to text orientation, noise and skew .It is free from artifacts that are usually introduced by thresholding using morphological operators. They implement a new text extraction algorithm from a text /graphics mixed images. In this method we use the OCR system to recognize the contained information in any image. With this method value of recall rate and f-score of 84.3 and 84.01 is obtained respectively. so this method does not give satisfied result.

Ming zhao, Shutao, James Kwak had implemented text detection in images using sparse representation with discriminative dictionaries. This method gives the value of F-score of 78.

P.Anish, Prof.S.Mary Joans in developed a system based on digital images and can be easily applied to commercial car parking systems for the use of documenting access of parking services. The character recognition was accomplished with the aid of optical characters by the process of template matching. In this method we use four d i f f e r e n t types of edge detectors named sobel, prewitt, canny, L O G . A comparative analysis on the success rate of proposed system showed overall better success rate by using canny edge detector. The proposed system does not give satisfied results for scene images. But this method is work only for number plate of cars etc.

Partha Pratim Roy, Josep Llad' os and Umapada Pal 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 separate them. But some of the characters can't be separated through connected component analysis.

Antani and R.Kasturi proposed an algorithm which works well for text string separation from mixed text and graphics image but it makes an impractical imagination that the character components in string are aligned straight and does not touch and overlap with graphics.

Coleman and Andrews d e v e l o p e d a method, which operates in an unsupervised mode with connected component method. But some part of text regions are vanished using connected component method. For improving accuracy they used modified m o r p h o l o g i c a l

filter and also proposed a clustering method based on different text sizes. Moreover, their method reduces noise in the resultant image. Experimental result confirms the superiority of their approach. With this method RRC accuracy of 92.8 % is achieved.

Yi-Feng Pan, Cheng-Lin Liu, Xinwen Hou in proposed a system for text extraction from images by learning based filtering and verification. Value of F score of 68 is achieved by this method. Nirmala Shivananda and P.Nagabhushan in proposed a method which is used to separate foreground text from multicolored complex background . in this method we use a connected component analysis with an unsupervised thresholding for the text separation from complex background. With the help of this method we can separate any color , font and size of foreground text in any image. In this method we use an OCR method to extract the foreground text. But noise level cannot be reduced to desired level due to absence of filter.

Vikas Dungre et .al proposed a reviews about the methods of the character recognition. They have introduced image preprocessing techniques for thresholding, skew correction which are used in character recognition process. They have also reviewed the feature extraction using global transformation and series expansion like Fourier transform, wavelets, moments, Gabor transform, statistical features like zoning, projections, crossings and distances ,and some geometrical and topological features commonly practiced .They also reviewed the classification using template matching ,statistical techniques ,neural network, SVM and combination of classifiers for better accuracy is practiced for recognition.

Giorgos Vamvakas et.al have described the structural and statistical features they have used in their approach of character recognition .The statistical features they have used are zoning, projections and profiling, and crossings and distances. With help of zoning they derived local features. Directional histograms of contour and skeleton images are used in the direction features .In addition to normal profile features they described in and out profiles of contour of images .The structural features they have depicted are end point ,loop, horizontal and vertical projection histograms ,radial histograms, radial-out-in and in –out histogram.

Wang Jin et .al proposed a series of recognition systems by using the virtual reconfigurable architecture based on evolvable hardware. A statistical pattern recognition-inspired methodology was introduced to improve the recognition accuracy of the proposed systems.

R.Chandrasekaran,RM.Chandrasekaran, P.Natarajan implemented an approach. In this approach, they have presented a robust approach for text extraction and recognition in images. First, the input image is filtered by the Median filter to remove any noises. Then edges are detected using ZCED algorithm. Then the morphological dilation operation is applied for object localization. All the connected Components are then extracted and all non-text character components are discarded by a two step process,

first by applying simple statistical matrices and then using SVM. For testing with SVM, features are extracted from the test components. These features form the feature vector for SVM. These features are then tested with SVM for recognizing individual characters. In this work, main emphasis has been given in eliminating false positives which is the major drawback in using cc based approach. The proposed method is tested with various types of images, both images with caption text as well as scene text. All related methods given in references are analyzed and the drawbacks are reduced and thereby getting an improved version of the previous works. A much higher recall rate can be achieved by using any other. With this approach F- score of 93 is achieved.

Mohammad Shorif Uddin, Madeena Sultana, Tanzila Rahman, and Umme Saym 2012 proposed a method with their main objective 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. Many researchers have been working on the development of techniques to extract texts from a scene images.

Satish R.Damade and Ranjana S. Zinjore presents a methodology for text extraction from document image. This is a region based approach. In this method firstly color image is converted into grey scale and then binary image. Then canny edge detector is applied on the binary image to detect the edges. Then morphological operations (dilation) are applied with square structuring element. Then dilated image is labeled with help of the property of the bounding box and area and then connected component is extracted. So we get only the text regions. Finally output image is subtracted from input image to visualize it properly. They also compare the result with edge based and connected component with projection profile approach and produce good results than others.

Prof. Amit Choksi et.al proposed a method for text extraction from natural scene image using Prewitt edge detection method. They compare the performance of the Prewitt edge detector method and Gaussian Edge Detector method performance. They have proved that edge detection using Prewitt edge detector gives almost same result compared with connected component algorithm. The Edge based method detects text regions with highest Precision & Recall rate. Thus it is more efficient compared to that of the performance obtained with Gaussian edge detection method. We observed that Gaussian edge detection based methods is not good enough to detect the text regions.

3. Performance Evolution

The performance of this method is calculated using precision, recall and f-score metrics and it is also compared with other methods.

3.1 False Positive

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.

Paper ID: SEP14292

3.2 False Negative

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

3.3 Precision Rate

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.

3.4 Recall Rate

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].

3.5 F-Score

F-score is the harmonic mean of the recall and precision rates. So on comparing all of these results with other methods, it is observed that this method gives better results then compared to all other techniques.

4. Conclusion

There are various applications of the text extraction from image like image indexing and retrieval, document analysis, vehicle license detection and recognition, page segmentation ,technical paper analysis, street signs, name plates, document. Coding, object identification, text based video indexing, video content analysis etc. various methods have been proposed for the text extraction from image for different type of the text size, shape, style, orientation, alignment etc. these properties are used to classify text from background in the image. In this paper we study about various techniques used for the text extraction from the image. All of the methods have some advantages and disadvantages for the text extraction. Based on the literature survey it is cleared that text extraction based on mathematical morphology produce good results as compared to all other techniques. It can extract 100% result for medium and large text present in any image.

References

[1] Mohammad Shrif Uddin,Madeena Sultana,Tanzila Rahman,andUmme 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] R.Chandrasekaran,R.M.Chandrasekaran,P.Natarajan, "Text localization and extraction in images using mathematical morphology and SVM". IEEE,International Conference (ICAESM-2012)March 30,312012.

[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 http.html,(Access ed on: November 22, 2010).

[6] G.Rama Mohan Babu, P. Srimaiyee, A.Srikrishna, " Text extraction from Hetrogeous Images using Mathematical Morphology",Journal of Theoretical and Applied Information Technology', 2010

[7] Yi-Feng Pan, Cheng-Lin Liu, Xinwen Hou, " Fast sceene text localization by learning based filtering and verification", IEEE 17th International Conference on Image Processing, 2010 September 26-29,2010.

[8] Ming zhao, Shutao, James Kwak, "Text detection in images using sparce representation with discriminative dictionaries",Image and Vision computing, 2010.

[9] 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).

[10] 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).

[11] Prachi Mukherji,Priti Rege, " Shape Feature and Fuzzy logic based offline Devanagari Handwritten Optical Character Recognition",Journal of Pattern Recognition, Research 4,pp.52-68,2009.

[12] Wang Jin ,Tang Bin bin,Lei Gai-hui, "Statistical method –based evolvable character recognition system" ,IEEE International Symposium on industrial Electronics(ISIE),pp.804-808,July2009.

[13] G.Vamvakas,B.Gatos,S.Petridis,N.Stamatopoulouset. al.,''Optical character Recognition for hand written characters"ppt,[online].

[14] Ming Zeng, Jianxun Li , ZhangPeng, "The design of top-hat morphological filter and application to infrared target detection", Infrared Physics & Technology, 48(1), pp. 67–76, 2009.

[15] J. Serra, "Image analysis and mathematical morphology", Publisher: Academic Press, 2nd Edition, vol.2, (2008).

[16] 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).

[17] Chitrakala Gopalan and D. Manjula, "Contourlet Based Approach for Text Identification and Extraction from Heterogeneous Textual Images", International Journal of Computer Science and Engineering 2(4) (2008) pp.202-211.

[18] 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).

[19] Edward R. Dougherty, "An Introduction to Morphological Image Processing", Publisher: SPIE Press, vol: TT59, (2003).

[20] 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, 2001, pp. 865-869.

Paper ID: SEP14292
1000

[21] Henk J . A . M . Heijmans a n d Jos B.T.M. Roerdink (Eds), "Mathematical Morphology and its Applications to Image and Signal Processing", Publisher: Springer, 1stEdition, pp. 367-374, (2000).

[22] K. Jain and B. Yu, "Document representation and its application to page decomposition,"IEEE Trans. Pattern Anal. Machine Intell., vol. 20, pp. 294–308, Mar. 1998.

[23] 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) .

[24] John C. Russ, "The Image Processing Handbook", publisher : CRC Press. 4th edition (2002).

[25] Bryan S. Morse, " Lecture 4: Thresholding", Brigham Young university, (1998-2000).

[26] Rafael Gonzalez, Richard E. Woods and Steven L. Eddins, "Digital Image Processing Using Matlab", Publisher: Prentice Hall, 2nd Edition.

[27] K. Jain, and Y. Zhong, "Page Segmentation using Texture Analysis", Pattern Recognition, 29 (5) 743-790, 1996.

[28] T. S. HUANG AND K. AIZAWA, "Image processing: Some challenging problems", Publisher: National Academy of Science, (1992)