ISSN No. 0976-5697

Volume 8, No. 5, May – June 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Text Recognition using Image Processing

Chowdhury Md Mizan, Tridib Chakraborty* and Suparna Karmakar
Department of Information Technology
Gurunanak Institute of Technology
Kolkata, India

Abstract: The goal of Text Recognition is to recognize the text from printed hardcopy document to desired format (like .docx). The process of Text Recognition involves several steps including preprocessing, segmentation, feature extraction, classification, post processing. Preprocessing is for done the basic operation on input image like binarization which convert gray Scale image into Binary Image, noise reduction which remove the noisy signal from image. Segmentation stage for segment the given image into line by line and segment each character from segmented line. Future extraction calculates the characteristics of character. A classification contains the database and does the comparison. Nowadays it plays an important role in office, colleges etc.

Keywords: Text detection, text segmentation, character recognition, scene image.

1. INTRODUCTION

Nowadays all over digitization technology is used. Text Recognition usually abbreviated to OCR[2][5][14], involves a computer system designed to translate images of typewritten text (usually captured by a scanner) into machine editable text or to translate pictures of characters into a standard encoding scheme representing them. OCR began as a field of research in artificial intelligence[24] and computational vision[26]. Text Recognition used in official task in which the large data have to type like post offices, banks, colleges etc., in real life applications where we want to collect some information from text written image. People wish to scan in a document and have the text of that document available in a .txt or .docx format.

2. PRIOR WORK

Preprocessing is the first step in the processing of scanned image[1][9]. The scanned image is checked for noise, skew, slant etc. There are possibilities of image getting skewed with either left or right orientation or with noise such as Gaussian. Here the image is first convert into grayscale and then into binary. Hence we get image which is suitable for further processing.

After pre-processing, the noise free image is passed to the segmentation phase, where the image is decomposed into individual characters.The binarized image is checked for inter line spaces. If inter line spaces are detected then the image is segmented into sets of paragraphs across the interline gap. The lines in the paragraphs are scanned for horizontal space intersection with respect to the background. Histogram[13] of the image is used to detect the width of the horizontal lines.Then the lines are scanned vertically for vertical space intersection. Here histograms[13] are used to detect the width of the words. Then the words are decomposed into characters using character width computation.

Feature extraction follows the segmentation phase of OCR[2][5][14] where the individual image glyph is considered and extracted for features. First a character glyph is defined by the following attributes like height of the character, width of the character.

Classification is done using the features extracted in the previous step, which corresponds to each character glyph. These features are analyzed using the set of rules and labeled as belonging to different classes. This classification is generalized such that it works for single font type. The height of the character and the width of the character, various distance metrics are chosen as the candidate for classification when conflict occurs. Similarly the classification rules are written for other characters. This method is a generic one since it extracts the shape of the characters and need not be trained. When a new glyph is given to this classifier block[10] it extracts the features and compares the features as per the rules and then recognizes the character and labels it.

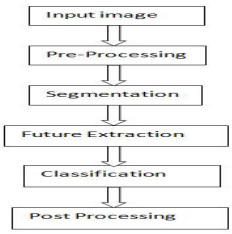

Fig 1.Flowchart of Text extraction process

3. ALGORITHMS

1. Start

2. Scan the textual image.

3. Convert color image into gray image and then binary image.

4. Do preprocessing like noise removal, skew correction etc.

5. Load the DATABASE.
6. Do segmentation by separating lines from textual image.

4. RELATED WORK

Development and progress of various approaches to the extraction of text information fro m the image and video have been proposed for specific application, including page

segmentation, text color extraction[2],video frame [3] text detection and content-based image or video indexing. However, extensive research, it is not easy to design series general-purpose systems. This is because there many possible sources of variation when extracting text . Shaded from the textured background or, from the low-contrast or complex images, or images with variations in font size, style, color, orientation, and alignment. This variation makes the problem very difficult to draw automatically. Generally text-detection methods can be classified into three categories. The first one consists of connected component-based methods, which assume that the text regions have uniform colors and satisfy certain size, shape, and spatial alignment constraints. However, these methods are not effective when the text have similar colors with background. The second one consists of the texture based methods, which assume that the text regions have special texture. Though these methods are comparatively less sensitive to background colors, they may not differentiate the texts from the text-like backgrounds. The third one consists of the edge-based methods. The text regions are detected under the assumption that the edge of the background and the object regions are sparser than those of the text regions. However, this kind of approaches is not very effective to detect texts with large font size. compared the Support Vector Machines (SVM) [1] based method with the multilayer perceptrons (MLP)[1] based one for text verification over four independent features, namely, the distance map feature, the grayscale spatial derivative feature, the constant gradient variance feature and the DCT coefficients feature. They found that better detection results are obtained by SVM rather than by MLP. Mu lti-resolution-based text detection methods are often adopted to detect texts in different scales. Texts with different scales will have different

MPEG VIDEO -> FRAME EXTRACTION ->IMAGE SEGMENTATION -> IMAGE CLASSIFICATION -> TEXT EXTRACTION -.> OCR

Text Extraction

The aim of Optical Character Recognition (OCR) [2][5][7] is to classify optical patterns (often contained in a dig ital image) corresponding to alphanumeric or other characters. The process of OCR[2][5][7] involves several steps including segmentation, feature extraction, and classification. In principle, any standard OCR[2][7] software can now be used to recognize the text in the segmented frames. However, a hard look at the properties of the candidate character regions in the segmented[13] frames or image reveals that most OCR software packages will have significant difficulty to recognize the text.Document images are different fro m natural images because they contain mainly text with a few graphics and images.Due to the very low-resolution of images of those captured using handheld devices, it is hard to extract the complete layout structure (logical or physical) of the documents and even worse to apply standard OCR systems. For this reason, a shallow representation of the low-resolution captured document images is proposed. In case of original electronic documents in the repository, the extraction of the same signature is straightforward; the PDF or PowerPoint form o f the original electronic documents is converted into a relatively high-resolution image (TIFF, JPEG, etc.)[16] on which the signature is computed.

Finally, the captured document's signature is compared to with all the original electronic documents' signatures in order to find a match.

I. Architecture of Text Extraction Process

Text extraction and recognition process comprises of five steps namely text detection, text localization, text tracking, segmentation or binarization[6], and character recognition. Architecture of text extraction process can be visualized in Fig. 2

Text Detection: This phase takes image or video frameas input and decides it contains text or not. It also identifies the text regions in image.

*Text Localization:*Text localization merges the textregions to formulate the text objects and define the tight bounds around the text objects.

*Text Tracking:*This phase is applied to video data only.For the readability purpose, text embedded in the video appears in more than thirty consecutive frames. Text tracking phase exploits this temporal occurrences of the same text object in multiple consecutive frames. It can be used to rectify the results of text detection and localization stage. It is also used to speed up the text extraction process by not applying the binarization[6] and recognition step to every detected object.

*Text Binarization:*This step is used to segment the textobject from the background in the bounded text objects. The output of text binarization is the binary image, where text pixels and background pixels appear in two different binary levels.

Character Recognition: The last module of textextraction process is the character recognition. This module converts the binary text object into the ASCII text.

Text detection, localization and tracking modules are closely related to each other and constitute the most challenging and difficult part of extraction process.

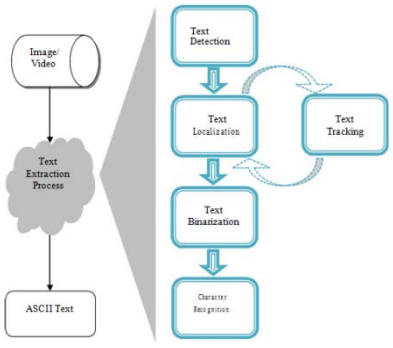

Fig 2. architecture of text extraction process

II. Applications of Text Extraction

Text extraction from images has ample of applications. With the rapid increase of multimedia data, need of understanding its content is also amplifying. Some of the applications of the text extraction are mentioned below.

A. Video and Image Retrieval

Content based image and video retrieval[] is the focus of many researchers for the last many years. Text appearing in the images gives the essence of the actual content of the image and displays the human perception about the content. This makes it a vital tool for indexing and retrieval of multimedia contents [3].This tool can give much better results than the other shape, texture or color based retrieval techniques [8]. Embedded text in the videos and images communicate human discernment about the content, hence it is most suitable for indexing and retrieval of multimedia data.

B. Multimedia Summarization

With the vast increase in the multimedia data, huge amount of information is available. Because of this overwhelming information, problem of overloaded information arise. Text summarization can provide the solution for the problem. Superimposed text in video sequences offer helpful information concerning their contents. Text data appear in video hold valuable knowledge for automatic annotation and generation of content summary. A variety of methods have been presented to deal with this issue. Sports video summarization and News digest are the well known applications of summarization of visual information

C. Indexing and Retrieval of Web Pages

Text Extraction method from web images can truly improve the indexing and retrieval of web pages. Main indexing terms are embedded in the title image or banners. Instead of text, most of the sites use image to present the title of the web page. So to precisely index and retrieve web pages, text within images must be understood. This would result into enhanced indexing and more proficient and accurate searching [10].

Text extraction from web images can also help in filtering of images with offensive language. It is also helpful in conversion of web page to voice.

Above listed applications are not the only examples of text extraction methods. There are plenty of other applications such as voice coding for blinds, intelligent transport system, Image tagging, robot vision and scene analysis etc.

5. CONCLUSION

In this paper we proposed algorithm for solving the problem of offline character recognition. We had given the input in the form of images. The algorithm was trained on the training data that was initially present in the database. We have done preprocessing and segmentation and detect the line.

The paper presents a brief survey of the applications in various fields along with experimentation into few selected fields. The proposed method is extremely efficient to extract all kinds of bimodal images including blur and illumination. The paper will act as a good literature survey for researchers starting to work in the field of optical character recognition.

6. REFERENCES

[1] Ramanathan. R. et al., "A Novel Technique for English Font Recognition Using Support Vector Machines", in Advances in Recent Technologies in Communication and Computing, Kottayam, Kerala, 2009, pp. 766-769.

[2] Line Eikvil, "Optical Character Recognition", NorskRegnesentral, Oslo, Norway, Rep. 876, 1993.

[3] M Usman Raza, et al., "Text Extraction Using Artificial Neural Networks", in Networked Computing and Advanced Information Management (NCM) 7th International Conference, Gyeongju, North Gyeongsang, 2011, pp. 134-137.

[4] Bertolami, Roman; Zimmermann, Matthias andBunke, Horst, 'Rejection strategies for offline handwritten text line recognition', ACM Portal, Vol. 27, Issue. 16, December 2006

[5] C.P. Sumathi, T. Santhanam, G.Gayathri Devi, "A Survey On Various Approaches Of text Extraction InImages", International Journal of Computer Science &Engineering Survey (IJCSES). Vol.3, August 2012, Page no. 27-42.

[6] Datong Chen, Juergen Luettin, Kim Shearer, "A Survey of Text Detection and Recognition in Images andVideos". Institute Dalle Molle d'Intelligence ArtificiellePerceptive Research Report, August 2000, Page no. 00-38.

[7] Xu-Cheng Yin, Xuwang Yin, Kaizhu Huang, and Hong-Wei Hao, "Robust Text Detection in Natural Scene Images", IEEE transaction on Pattern Analysis AndMachine Intelligence, 2013, Vol. 36, Page no. 970 – 983.

[8] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition". International conference on computer vision ICCV 2011, vol. 10, Page no.1457 – 1464.

[9] Xiaobing Wang, Yonghang Song, Yuanlin Zhang,"Natural scene text detection in multi-channel connected component segmentation", 12th International conf. onDocument Analysis and Recognition, pp. 1375-1379, 2013.

[10] Shehzad Muhammad Hanif, Lionel Prevost, "TextDetection and Localization in Complex Scene Images using Constrained AdaBoost Algorithm", 10thInternational Conference on Document Analysis and Recognition, pp.1-9, 2009.

[11]Teofilo E. de Campos, Bodla Rakesh Babu, Manik Varma, "Character Recognition In Natural Images",International conf. on Intelligence Science and Big data Engg., pp. 193-200, 2011.

[12]Lukas Neumann, Jırı Matas, "Real-Time Scene Text Localization and Recognition", IEEE Conf. on Computer Vision and Pattern Recognition, 2012, pp. 3538–3545.

[13]Teofilo E. de Campos, Bodla Rakesh Babu, Manik Varma, "Character Recognition In Natural Images",International conf. on Intelligence Science and Big data Engg., pp. 193-200, 2011.

[14] Honggang Zhang, KailiZhao, Yi-ZheSong, JunGuo, "Text extraction from natural scene image: A survey",Elsevier journal on Neurocomputing ,pp.310-323, 2013.

[15] Cunzhao Shi, Chunheng Wang, Baihua Xiao, Yang Zhang, Song Gao, "Scene text detection using graph model built upon maximally stable extremal region". vol 34, issue 2, 2013, page no. 107-116

[16] K. Bachuwar, A. Singh, G. Bansa, S. Tiwari, "An Experimental Evaluation of Preprocessing Parameters for GA Based OCR Segmentation" in *3rd International Conference on ComputationalIntelligence and Industrial Applications (PACIIA 2010),* 2010,proceedings, Vol. 2, pp. 417 -420.

[17] M.D. Ganis, C.L. Wilson, J.L. Blue, "Neural network-based systems for handprint OCR applications" *in IEEE Transactions on ImageProcessing*, 1998, Vol: 7 , Issue: 8, p.p. 1097 - 1112

[18] R. Gossweiler, M. Kamvar, S. Baluja, "What's Up CAPTCHA? A CAPTCHA Based On Image Orientation", in *WWW*, 2009.

[19] B. Joanna," Building an institutional repository at Loughborough University: some experiences, program: *Electronic library andinformation systems*, 2009.

[20] A. Singh, K. Bacchuwar, A. Choubey, S. Karanam, D. Kumar, "An OMR Based Automatic Music Player", in *3rdInternational Conferenceon Computer Research and Development (ICCRD 2011) in, (IEEE Xplore)*, 2011, Vol. 1, pp. 174-178.

[21] S.L. Chang, T. Taiwan , L.S. Chen, Y.C. Chung, S.W. Chen, " Automatic license plate recognition" in *IEEE Transactions onIntelligent Transportation Systems*, 2004, Vol: 5 , Issue: 1, p.p. 42 – 53

[22] R Plamondon, S. N. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey" *IEEE transaction on patternAnalysis and machine Intelligence*, 2000, 22(1), 63-84

[23] K. S. Fu, J.K. Mui, "A Survey on Image segmentation". *PatternRecognition*, 1981, 13, 3–16

[24] M.S. Kalas, "An Artificial Neural Network for Detection of Biological Early Brain Cancer", *International Journal of Computer Applications*, 2010, 1(6), 17–23.

[25] R. Storn, K. Price, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces", *Journal ofGlobal Optimization*, 1997, 341–359

[26] IP, H.H.-S., S.-L. Chan, "Hypertext-Assisted Video Indexing and Content- based Retrieval", *ACM* 0-89791-866-5, 1997, PP. 232–23.3

[27] N.R. Pal, S.K. Pal, "A Review on Image Segmentation Techniques", *Pattern Recognition*, 1993, Vol 26(9), 1277–1294.