An Approach for Text Detection in Images

[1]Sandeep Sharma, [2]Jai Prakash
[1,2]Dept. of CSE, Madan Mohan Malaviya Engineering College, Gorakhpur, UP, India

Abstract
We present a novel text detection technique for the application of image in databases of consumer pictures and videos. Images are annotated using information only from the picture itself. It needs to be fast, efficient and robust in order to feed an OCR classifier with the correct input. In other words, segmented regions must correspond to the actual text. A lot of work has been done for detecting text in images and a lot has to be done. This paper gives detailed survey on image to text detection mechanism. This paper gives the description of work has been done for automatically detection of text from images, localize and extract horizontally aligned text in images (and digital videos) with complex backgrounds is presented.

Keywords
Text Detection, Text Segmentation, Color Uniformity, Preprocessing

I. Introduction

A text in any form or place has high purpose and contains more information related to the place and helps us to understand the objective more easily. The rapid growth in digital technologies and gadgets outfitted with megapixel cameras and invention of latest touch screen method in digital devices like mobile, PDA, etc., increase the demand for information retrieval and it leads to many new research challenges. Text detection and segmentation from natural scene images are useful in many applications. the primary property of scene text such as, high contrast against background, uniform colors are difficult to preserve in real application. When the system scans whole image for texts, text pixels with low contrast and non-uniform lighting could be confused as background due to similar colors.

An overview on text detection algorithms is given in [1]. Ezaki et al. [2] exploit the color uniformity property and process non-overlapping windows in the image. Fisher's Discriminant Rate (FDR) is used to classify each window into text and background. Clark and Mirmehdi [3] calculate statistical measures of orientation and gradient information and classify those using a neural network. Zhu et al. [4] extract Connected Components (CCs) from a Niblack binarization. Each CC is described using geometric, edge, shape, stroke and spatial features and is classified into text/background using an AdaBoost classifier. Hanif et al. [5] also use AdaBoost classification but use a more extended feature set. We want to detect typewritten text of any size, font and color in gray-level images. For this reason, our algorithm is based on a set of text features, reused from the previously mentioned papers. It is hard to compare the performance of the available literature, because of the use of different performance measures and different datasets.

1. Text appearing accidentally in an image that usually does not represent anything important related to the content of the image. Such texts are referred to as scene text .
2. Text produced separately from the image is in general a very good key to understand the image it is called artificial text.

Recently, face detection and recognition is being applied in cameras and at large popular image sharing websites. Several text detection methods have been proposed based on edge detection, binarization, spatial-frequency image analysis and mathematical morphology Generally text detection methods can be classified as either edge-based, connected-component based and texture-based methods . According to the best results were achieved using edge based text detection. It obtained top overall performance among 4 methods including mathematical morphology and color-based character extraction. Edge-based text detection has also been used in combination with edge profiles. Park et al. use them for automatic detection and recognition of Korean text in outdoor signboard images. However, they assume that a single text sign is located around the center line of the image. Edge profiles have also been used for detecting text in video data. Shivakumara et al. use edge profiles in combination with additional edge features to eliminate false positives selection.

The remaining part of our paper is organized as follows: In section II, we will discuss the text detection algorithm from image to text detection and in section III, we will discuss the result and analysis of this method and finally in section IV, we will conclude the paper and give the future scope of this paper.

II. Text Detection ALGO

Our approach processes each image in several steps: from pixels to candidate regions and from candidate regions to word regions. First, we select text candidates by using color uniformity. Next, more advanced textual features are extracted for the candidates for classifying the window into text orbackground. Remaining candidates are grouped into words. The block diagram is shown as

A. Preprocessing
B. Feature extraction
C. Classify
D. Grouping

A. Preprocessing

First, the image is converted to gray-level, because we want to be able to detect text, even in low-light conditions where color information is absent or noisy. Next, the input image is down sampled several times to create a scale-space pyramid, enabling the detection of text at different size levels. Each scaled version of the original image is processed independently. In the sequel, we explain the successive processing steps for a single scaled version of the image. The image is densely scanned with a sliding window. Each position of the window defines a text candidate, which is assumed to contain either a single character, or background.

We use overlapping windows of size 32×32 pixels. Candidate windows that do not satisfy the constraints of text are discarded. We assume that characters have a uniform color with a high contrast to a uniform background. Therefore, the color distribution of a candidate with text will ideally show two peaks. Each candidate window is locally binarized according to its color distribution and CCs are extracted.

B. Text Candidate Features Classification

From the CCs remaining after preprocessing stage, 5 features will be extracted and used for the final classification into text and background. These features present the intrinsic properties of text, namely, color, shape and orientation information.

First, the color features are represented by the Fisher's Discriminant

Rate (FDR) [2]. Second, assuming that characters contains many strong edges, the shape features are the number of edge pixels after a binary edge detection stage and a measure of the compactness of the CC [4]. Furthermore, we calculate the M1, M2 and M4 measures from [3].

C. Grouping

In this step, character candidates that are not likely to form a word with the neighboring candidates are removed. Since candidates represent complete characters, we assume that candidates are spatially close and have similar a size and color and are possibly part of a word. For each individual property, we define an independent distance function and we multiply the distance values to obtain the final similarity score based on the mutual correlation on the occurrence of neigboring character features. For each candidate, the final score is thresholded to provide the candidates to be merged into words. We group the remaining candidates that are aligned and are spatially close into a word candidate. Since the word candidate can sometimes contain multiple words, a splitting operation is applied to divide the line of text into separate words, based on the histogram of inter-character distances.

III. Result and Analysis

We evaluated our algorithm using the very challenging public ICDAR2003 dataset [2]. We trained the classifier on the training set (258 images) and tested on the test set (251 images). .We obtain a maximum detection rate (recall) of 70% using the measurements of [2] and an equal-error-rate of 58%. . Note that the applied measure is very strict. For several images we obtained a recall of only 80%, although we detected the text perfectly. The average processing time of the unoptimized software implementation is 1.6 seconds per image on the ICDAR dataset. VGA-resolution images are processed in 670 msec. This is much faster than the algorithm with best performance [7]. Our proposed text detection module provides a suitable extension to the available toolset for the automatic annotation of images to allow the archiving and retrieval in a consumer-oriented database.

IV. Conclusion

In This paper, we have studied the define what is image to text detection paradigm the we have done the literature review in image to text detection and we have also analyze the various performance issues and research challenges as well as we have examines the various image to text detection algorithms and we have done the comparative analysis of the algorithms on the basis of various parameters and we have done the analysis on the basis of advantages and disadvantages. So we can say that image to text detection is a growing these days and become the important area of research and a number of work has been done in this field and various work has to be done in this field and can propose various fine algorithms in the field of image to text detection which can perform well.

References

[1] Jian Liang, David Doermann, Huiping Li,"Camera-based Analysis of Text and Documents: a Survey", International Journal on Document Analysis and Recognition, pp. 84–104, 2005.

[2] Simon M. Lucas et al.,"ICDAR 2003 Robust Reading Competitions: Entries, Results, and Future Directions", International Journal on Document Analysis and Recognition, Vol. 7, No. 2, pp. 105–122, 2005.

[3] P. Clark, M. Mirmehdi,"Combining Statistical Measures to find Image Text Regions", Proc. IEEE International Conference on Pattern Recognition (ICPR), September 2000, Vol. 1, pp. 450–453.

[4] Kaihua Zhu et al.,"Using AdaBoost to Detect and Segment Characters from Natural Scenes", First International Workshop on Camera-Based Document Analysis and Recognition (CBDAR), August 2005.

[5] Shehzad Muhammad Hanif, Lionel Prevost, Pablo Augusto Negri,"A Cascade Detector for Text Detection in Natural Scene Images", Proc. International Conference on Document Analysis and Recognition (ICDAR), July 2009.

[6] Rob G. J. Wijnhoven, Peter H. N. de With,"Fast Training of Object Detection using Stochastic Gradient Descent", Proc. IEEE International Conference on Pattern Recognition (ICPR), August 2010, pp. 424–427.

[7] S.M. Lucas,"ICDAR 2005 Text Locating Competition Results", Proceedings 8th International Conference on Document Analysis and Recognition (ICDAR), August 2005, Vol. 1, pp. 80–84.

Sandeep Sharma has received his B.Tech degree from Pranveer Singh Institute Of Technology ,Kanpur in Computer Science and Engg(2005-09). He is doing M.Tech from MMMEC, Gorakhpur in Computer Science and Engg. He is doing his thesis under the area of image processing. He has published one paper in IJARCS journal which covers the topic "A Survey on Image To Text Detection Methodology".