INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 954-958
RECEIVED : 27/03/2016. PUBLISHED : 05/04/2016
April 5, 2016

TEXT DETECTION IN NATURAL IMAGES AND VIDEO USING MAXIMALLY STABLE EXTREMAL REGIONS AND TEXT CANDIDATE CLASSIFICATION

[1]P.Rumaniya, [2]S.Ramesh,
PG Scholar, Dept of Communication systems, Mahendra Institute of Technology,
Asst prof, Dept of ECE, Mahendra Institute of Technology.

Abstract:
The performance of text detection and recognition in colour images is analyzed and compared to solve the technical challenges. Maximally stable Extremal Regions (MSERs) are extracted as character candidates by a fast and effective pruning algorithm. Text candidates are formed by grouping the character candidates using single link clustering algorithm. Here, clustering threshold and distance weights are known automatically by self training distance metric learning algorithm. Character classifier is used to estimate the probability of text candidates corresponding to non text ones. Text classifier is to identify the texts by eliminating text candidates having high non result reveals the better performance of the proposed work.

Keywords –MSER, Classifer, Algorithm.

1. INTRODUCTION

The area of scene text recognition focuses on the problem of recognizing arbitrary text in images of natural scenes. Examples of scene text include street signs, business signs, grocery item labels, and license plates. Potential applications of scene text recognition include improving navigation for people with low vision, recognizing and translating text into other languages, improving image retrieval and aiding autonomous navigation for cars and robots. This problem is similar to the area of optical character recognition for documents. However, unlike images of documents that usually to learn repeated patterns have standard fonts, structured text on a plain background and are usually captured in a controlled setting, images of natural scenes have many characteristics that make them midcult to analyze. They often contain more extreme lighting variation, may include unusual or highly stylized fonts, often vary widely in color and texture and may be captured from a wide range of viewing angles. In addition, scene text images usually contain only a few words, so it is more midcult of appearance. Content-based image indexing refers to the process of attaching labels to images based on their content. Image content can be divided into two main categories: perceptual content and semantic content . Perceptual content includes attributes such as color, intensity, shape, texture, and their temporal changes, whereas semantic content means objects, events, and their relations. A number of studies on the use of relatively low-level perceptual content for image and video indexing have already been re-ported. Studies on semantic image content in the form of text, face, vehicle, and human action have also attracted some recent interest . Among them, text within an image is of particular interest.

2. METHODOLOGY

Content-based image indexing refers to the process of attaching labels to images based on their content. Image content can be divided into two main categories: semantic content and perceptual content [4]. A number of studies on the use of relatively low- level perceptual content [5] for image and video indexing have already been reported Perceptual content includes attributes such as color, texture, shape, intensity and their temporal changes, whereas semantic content means events, objects, and their relation

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 954-958
RECEIVED : 27/03/2016. PUBLISHED : 05/04/2016
April 5, 2016

between them. Studies on semantic image content in the form of vehicle, text, face, and human action have also attracted some recent interest [6]. It is well known that scene text is more difficult to detect and very

Fig.1.Document images

little work has been done in this area. In contrast to caption text, scene text can have any orientation and may be damaged by the point of view projection. Moreover, it is often affected by variations in scene and camera parameters such as illumination, focus, motion, etc. Page layout analysis usually deals with document images Readers may refer to papers on document segmentation analysis [7]for more examples of document images. Although images acquired by CD covers, scanning book covers or other multi-colored documents have similar characteristics as the document images.

3. RELATED WORK

Automatic detection and recognition of text in images done using different techniques proposed. Generally, Text detection mechanism can be divided into two categories: connected component (CC)-based methods and Edge-based methods. Edge-based methods [8] approved a sliding window scheme, which is basically a brute force scheme which requires a lot of local decisions. Therefore, the region-based methods have attentive on an efficient binary classification (text versus non-text) of a small image area. In other words, they have focused on to determine whether a given patch is a part of a text region or not. Limitations of Edge-based methods are high computational complexity and the difficulty to select the best features for scene text detection. On the other hand CC- based methods are simple and efficient text detection approach. Connected component methods generated separate CC regions. Connected component based methods use a bottom-up approach by grouping small components into sequentially larger components until all regions are identified in the image. A geometrical analysis is needed to join the text components using the spatial arrangement of the components so as to filter out non-text components and mark the boundaries of the text regions in the images. The input image to the proposed system has a complex background with text built in it. The first stage is Image Pre -processing, which serves to remove

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 954-958
RECEIVED : 27/03/2016. PUBLISHED : 05/04/2016
April 5, 2016

the noise from the input image and generates a clear binary image. Text segmentation is the next stage, where we differentiate each character from the entire word by circumscribing them into boxes and saving them each separately. The final stage is text recognition, where the segmented characters [9] are compared

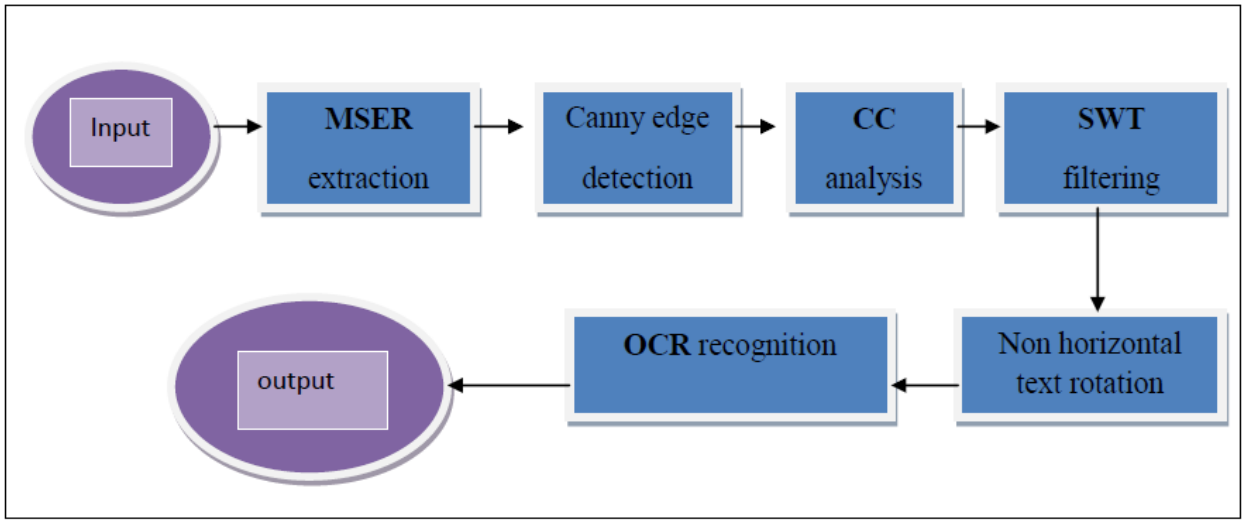

Fig.2. flowchart of our text detection algorithm

to the stored character matrices and as a result, the closest match for each character is displayed as a separate picture.

4. ANALYSIS

The flowchart of our text detection algorithm is shown in Fig.3. At the input of the system, the image intensities are linearly adjusted to enhance the contrast. Subsequently, MSER regions are efficiently extracted from the image [10] and enhanced using Canny edges obtained from the original gray-scale image As a next step, the resulting CCs are filtered using geometric constraints on properties like aspect ratio and number of holes. The stroke width information is robustly computed using a distance transform and objects with high variation in stroke width are rejected. Text candidates are grouped pair wise and form text lines. Finally, words within a text line are separated, giving segmented word patches at the output of our system. In the method propose two level text/ nontext filtering using a novel image operator stroke width transform and cc analysis.

Input image

The proposed method is work on non-horizontal text detection by rotating the text to make it as a horizontal text. Then the text regions are localized using parallelograms. Detected text regions are recognized using OCR recognizer [11]. For the generation of candidates, we extract CCs in images. This step includes Detect MSER region, canny edge detection and CC analysis. Maximally Stable Extremal Regions (MSER)[12] have become one of the commonly used region detector because of their high repeatability and partly because they are somewhat complementary to many other commonly used

Output image after text extraction

detectors . only the MSER algorithm could provide the stable binary results and also help us find most of the text components. And then as shown in above image the next step is to use the canny edge operator as Edge detection is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities. And since written text is typically placed on clear background, it tends to produce high response to edge detection image the Stroke width is defined as the length of a straight line from a text edge pixel to another along its gradient direction. The basic motivation of our stroke width extraction algorithm is that stroke width almost remains the same in a single character; however, there is significant change in stroke width in non-text regions as a result of their irregularity. And last sep is to use the optical character recognition in which the segmentation of text from a muddled scene can greatly improve OCR results. Since the algorithm already produced a well segmented text region, we can use the binary text mask to improve the accuracy of the recognition results. The performance of text detection and recognition in colour images is analyzed and compared to solve the technical challenges. Maximally stable Extremal Regions (MSERs) are extracted as character candidates by a fast and effective pruning algorithm. Text candidates are formed by grouping the character candidates using single link clustering algorithm. Here, clustering threshold and distance weights are known automatically by self training distance metric learning algorithm. Character classifier is used to estimate the probability of text candidates corresponding to non text ones. Text classifier is to identify the texts by eliminating text candidates having high non text probabilities. The simulation result reveals the better performance of the proposed work. Text Extraction from image is concerned with extracting the relevant text data from a collection of images. Rapid development of digital technology has resulted in digitization of all categories of materials. Lot of resources are available in electronic medium. Many existing paper-based collections, historical manuscripts books, journals, scanned document, video images, maps, posters, broadsides, newspapers, micro facsimile, microfilms, university archives, book plates, graphic materials, coins, currency, stamps, business cards, advertisements, web pages are converted to images and these images present many challenging research issues in text extraction and recognition.Text extraction from images have many useful applications in document analysis, detection of vehicle license plate, analysis of article with tables, maps, charts, diagrams, keyword based image search, identification of parts in industrial automation, content based retrieval, object identification, street signs, text based video indexing, page segmentation, document retrieving, address.

CONCLUSION

In this proposed system a new text detection method that effectively detects and recognizes texts of different directions in natural images. In this paper CC-based text detection algorithm is proposed to overcome the difficulties of grouping the characters and remove false positives of text in images. The system is stable and robust. All the system parameters remain the same throughout all the experiments. For future work we will focus on learning based methods for text extraction from complex backgrounds and text normalization for OCR recognition. We also attempt to improve the efficiency and transplant the algorithms into a navigation system prepared for the way finding of visually impaired people.

REFERENCES:

[1] J. Ohya, A. Shio, and S. Akamatsu, Recognizing Characters in Scene Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 16 (2) (1994) 214-224.

[2] K. Jung, "Text information extraction in images and video: A survey," Pattern Recognition, vol. 37, no. 5, pp. 977–997, May 2004.

[3] J. Liang, D. Doermann, and H. P. Li, "Camera-based analysis of text and documents: a survey," IJDAR, vol. 7, no. 2-3, pp. 84–104, 2005.

[4] Antani, S., Kasturi, R., and Jain, R., "A Survey on the Use of Pattern Recognition Methods for Abstraction, Indexing, and Retrieval of Images and Video", Pattern Recognition 35 945-965,2002.

[5] Lyu, M.R., Song, J., Cai, M.: "A comprehensive method for multilingual video text detection, localization, and extraction", IEEE Trans. Circuits Syst. Video Technol. 15(2), 243–255 ,2005.

[6] Leon, M., Vilaplana, V., Gasull, A. and Marques, F., "Caption text extraction for indexing purposes using a hierarchical region-based image model," IEEE ICIP 2009, El Cairo, Egypt, 2009.

[7] Xiaoqing Liu and Jagath Samarabandu, " Multiscale Edge-Based Text Extraction From Complex Images" , Multimedia and Expo, 2006 IEEE International Conference,2006.

[8]Leon, M., Vilaplana, V., Gasull, A. and Marques, F., "Caption text extraction for indexing purposes using a hierarchical region-based image model," IEEE ICIP 2009, El Cairo, Egypt, 2009.

[9] Hyung Il Koo, Member, IEEE, and Duck Hoon Kim, Member, IEEE "Scene Text Detection via Connected Component Clustering and Nontext Filtering" in IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 6, JUNE 2013

[10] M. Basavanna , P. Shivakumara, S. K. Srivatsa, G. Hemantha Kumar,_Multi-Oriented Text Detection In Scene Images,_ in International Journal of Pattern Recognition and Artificial Intelligence , Vol. 26, No. 7 Dec 2012.