International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 5 Issue: 5 26 – 28

A Review on Intelligent Scene Text Recognition of Natural Images

Pooja P. Kohapare, Prof. N. Kottawar, Prof A. Manusmare
Dept. Electronics and Communication Engg,
Gondwana University,
Ballarpur, India

Abstract: This paper provides an algorithm for detection and reading of a particular text given in natural images. Scene text recognition has inspired a good interest for computer vision community in recent years. In this paper we proposed text recognition method integrating structure-guided character detection of natural images present in surroundings. From the dataset, we manually label and extract the text region. Then next we perform statistical analysis of the text region to determine which image features are reliable indicators of text and have low entropy.We use part-based tree structure to model each category of characters so as to detect and recognize characters simultaneously.

Keywords- *Robust reading, character recognition, feature learning, cropped word recognition, part-based tree-structured models (TSMs), posterior probability, word spotting.*

*****

I. INTRODUCTION

With rapid growth of camera-based applications readily available on smart phones and portable devices understanding the pictures or a text taken by the devices semantically has gained increasing attention from the computer vision community in recent years. Among all the information contained in the image or text which carries semantic information, could provide valuable cues about the content of the image and thus is very important for human as well as computer to understand the scenes. Methods for scene text localization and recognition aim to find all areas in an image that would be considered as text by human mark boundaries of the areas (usually by rectangular bounding boxes) and output for real-world images and video processing (i.e. processing of images / videos taken by standard camera or mobile phone) and reading content of each detected area into a digital text format that can be further processed by a computer.

This paper presents an algorithm for detecting and reading text in city scenes. Text includes stereotypical forms such as street sign, hospital signs, bus numbers, shop signs, house numbers, and billboards. Database of city images were taken in San Francisco partly by normally slighted viewers and partly by blind volunteers who were accompanied by slighted guides using automatic camera setting and little practical knowledge [7], were text was located in image. Any natural scenes located at environment can be used for character recognition.

The image content is classified into two categories: perceptual content and semantic content.

Perceptual content includes colors, shapes, textures, intensities and their temporal changes. While semantic content includes objects, events and their relations [5]. Text contains high level of semantic information as compared to visual information. The importance of digital libraries for information retrieval cannot be denied. The ancient historical books contain invaluable knowledge but it is very time consuming to search the required information in this paper books. Different methods have been devised to facilitate the information search. This includes word spotting, optical character recognition etc. [9]. Although a lot of work is already has done in this field, it still remains an inviting and challenging field of research.

II. LITERATURE SURVEY

This section describes previously proposed studies that text extracted from any natural scenes around us which depends on their font sizes and thickness of text denoted on any scene, as in [5] paper revealed Y. Song, A. Liu, L. Pang, S. Lin, has mentioned about text images which contain important contents for information indexing and retrieval, automatic annotation and structuring of images.

As in [8] L. Neumann and J. Matas, describes regarding end-to-end and real-time scene text localization and recognition method is presented. The real-time performance is achieved by posing the character detection problem as an efficient sequential selection from the set of extremely regions (ERs). The ER detector is robust to blur, illumination, color and texture variation and handles low-contrast text.

As paper published [11]. P. Shivakumara, T. Phan, S. Bhowmick, C. Tan, and U. Pal, introduce a novel ring radius transform (RRT) and the concept of medical pixels on character with broken contours in the edge domain for reconstruction. For each pixel, the RRT assign a value which is the distance to the nearest edge pixel. The medical pixels are those which have maximum radius values in their neighborhood.

III. PROPOSED METHODOLOGY

Most of the work on scene text recognition could be roughly classified into three categories:

- Traditional OCR-based method.
- Object recognition method.
- Structure based model for object detection.

Methods for scene text localization and recognition aim to and all areas in an image (or a video) that would be considered as text by human, mark boundaries of the areas (usually by rectangular bound-In boxes) and output a sequence of (Unicode) characters associated with its content. They allow for real-world images and video process- ing (i.e. processing of images/videos taken by a standard camera or a mobile phone) and \reading" content of each detected area into a digital text format that can be further processed by a computer.We have proposed an end-to-end real-time scene text localization and recognition method which achieves state-of-the-art results on standard datasets (we consider a text recognition method real-time if the processing time is comparable with the time it would takea human to read the text).

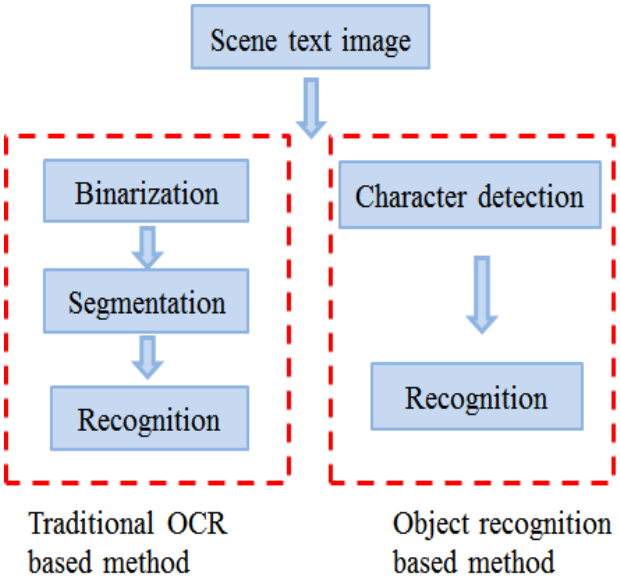

Fig. 1 Illustration of the traditional OCR-based method and object recognition-based method

For traditional OCR-based methods as shown in Fig.1, they focus on the binarization process, which segments text from background, and then the binary image could be segmented into individual characters, which could be recognized by the OCR engine. The binarization results are very disappointing, making it almost impossible for the following segmentation and recognition.

On the other hand, object recognition-based methods assume that scene character recognition is quite similar to object recognition with a high degree of intra-class variation. Proposed a local intensity normalizationmethod to handle lighting variations then used a Gabor transform to obtain local features and finally adopted a lineardiscriminant analysis for feature selection.
Structure-based model, which captures the local appearance properties and the deformable configuration of an object, has inspired great interest for object detection. As mentioned above, a good scene character-recognition
Method should make use of both the local appearance and global structure information. Motivated by the recent progress in object detection, using part-based TSM, we find that we could adopt these models to use both the global structure and local appearance information of characters. We propose to recognize characters by detecting part based tree-structures, which seamlessly combines detection and recognition together.
Then, we convert the candidate detection scores to posterior probabilities via confidence transformation. For word recognition, we combine the detection scores and language model into the posterior probability of the character sequence from the Bayesian decision view.Bigram, trigram, and even higher order language model could be incorporated. The final word-recognition result is obtained by finding the most probable character sequence via Viterbi algorithm.

IV. CONCLUSION

This article studies an effective scene text-recognition method incorporating structure-guided detection and linguistic knowledge into the posterior probability of character sequence from Bayesian decision view. We propose to learn a part based TSM for each category of characters to detect and recognize the characters simultaneously. Since the character specific TSM makes use of both the global structure informationand the local appearance information, the detection results are more reliable. Finally, we get the word-recognitionresult by maximizing the posterior probability of the character sequence via Viterbi algorithm.

V. ACKNOWLEDGEMENT

This research was supported by publisher of this paper. We thank our colleagues from Ballarpur institute of

technology who provided expertise that greatly assisted the research. Also, for sharing their pearls of wisdom with us during the course of this review paper.

REFERENCES

[1] K. Wang and S. Belongie, "Word spotting in the wild," in Proc. 11th ECCV, Sep. 2010, pp. 591–604.

[2] A. Mishra, K. Alahari, and C. V. Jawahar, "Scene text recognition usinghigher order language priors," in Proc.23rd BMVC, 2012, pp. 1–11.

[3] J. Gllavata, R. Ewerth, T. Stefi, and B. Freisleben, "Unsupervised textsegmentation using color and wavelet features," in Image and VideoRetrieval. New York, NY, USA: Springer-Verlag, 2004..

[4] Y. Song, A. Liu, L. Pang, S. Lin, Y. Zhang, and S. Tang, "A novelimage text extraction method based on K-means clustering," in Proc.7th IEEE/ACIS ICIS, May 2008, pp. 185–190.

[5] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in naturalscenes with stroke width transform," in Proc.IEEE CVPR, Jun. 2010,pp. 2963–2970.

[6] X. Chen and A. Yuille, "Detecting and reading text in natural scenes,"in Proc. IEEE CVPR, vol. 2. Jul. 2004, pp. 366–373.

[7] L. Neumann and J. Matas, "Real-time scene text localization andrecognition," in Proc. IEEE CVPR, Dec. 2012, pp. 3538–3545.

[8] W. Niblack, An Introduction to Digital Image Processing. Mundelein,IL, USA: Strandberg, 1985

[9] A. Newell and L. Griffin, "Multistage histogram of oriented gradientdescriptors for robust character recognition," in Proc. IEEE ICDAR,Sep. 2011, pp. 1085–1089.

[10] P. Shivakumara, T. Phan, S. Bhowmick, C. Tan, and U. Pal, "A novel ringradius transform for video character reconstruction," Pattern Recognition.,vol. 46, no. 1, pp. 131–140, 2012.