www.ijiarec.com

ISSN:2348-2079

Volume-5 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Scene text recognition in images by character descriptor and structure configuration

1.Mr.S.Jagadeesan.,M.Sc.,M.C.A.,M.Phil.,M.E., Assistant Professor
2. Mr. R.Sethupathi Final M.C.A.
Department of MCA, Nandha Engineering College,(Autonomous), Erode-52.
jagadeesan12398@gmail.com, sethupathi15987@gmail.com

Abstract—Camera-Based text information serves as effective tags or clues for many mobile applications associated with media analysis, content retrieval, scene understanding, and assistant navigation. In natural scene images, text characters and strings usually appear in nearby sign boards and provide significant knowledge of surrounding environment and objects. Text characters and strings in natural scene can provide valuable information for many applications. Extracting text directly from natural scene images is a challenging task because of diverse text patterns and variant background interferences. This project proposes a method of scene text recognition from detected text regions. In text detection, the previously proposed algorithms are applied to obtain text regions from scene image. The project designs a discriminative character descriptor by combining several state-of-the-art feature detectors and descriptors. Second, it models character structure at each character class by designing stroke configuration maps. The design is compatible with the application of scene text extraction in images. The project is developed to show the effectiveness of our proposed method on scene text information extraction from nearby objects. The demo system also provides us some insight into algorithm design and performance improvement of scene text extraction.

Index Terms—Scene text detection, scene text recognition, mobile application, character descriptor, stroke configuration, text understanding, text retrieval, mobile application.

I. INTRODUCTION

The first step in the software development life cycle is the identification of the problem. As the success of the system depends largely on how accurately a problem is identified. The natural scene images, text characters and strings usually appear in nearby sign boards and hand-held objects and provide significant knowledge of surrounding environment and objects.

Text-based tags are much more applicable than barcode or quick response code because the latter techniques contain limited information and require pre-installed marks. To extract text information by mobile devices from natural scene, automatic and efficient scene text detection and recognition algorithms are essential. However, extracting scene text is a challenging task due to two main factors is cluttered backgrounds with noise and non-text outliers, and diverse text patterns such as character types, fonts, and sizes. The frequency of occurrence of text in natural scene is very low, and a limited number of text characters are embedded into complex non-text background outliers. Background textures, such as grid, window, and brick, even resemble text characters and strings. Although these challenging factors exist in face and car, many state-of-the-art algorithms have demonstrated effectiveness on those applications, because face and car, have relatively stable features.

In this proposed system to solve these challenging problems, scene text extraction is divided into two processes text detection and text recognition. Text detection is to localize image regions containing text characters and strings. It aims to remove most non-text background outliers. Text recognition is to transform pixel-based text into readable code. It aims to accurately distinguish different text characters and properly compose text word. So, this project identifiesthat,ithelpsfor effective communication.

II. RELATED WORKS

In this section, we present a general review of previous work on scene text recognition respectively. While text detection aims to localize text regions in images by filtering out nontext outliers from cluttered background [3], [7], [10], text recognition is to transform image-based text information in the detected regions into readable text codes. Scene text recognition is still an open topic to be addressed. In the Robust Reading Competition of International Conference on Document Analysis and Recognition (ICDAR) 2011 the best word recognition rate for scene images was only about 41.2%. In general, scene text characters are composed of cross-cutting stroke components in uniform colors and multiple orientations, but they are usually influenced by some font distortions and background outliers.

We observe that text characters from different categories are distinguished by boundary shape and skeleton structure, which plays an important role in designing character recognition algorithm. Current optical character recognition (OCR) systems [2], can achieve almost perfect recognition rate on printed text in scanned documents, but cannot accurately recognize text information directly from camera-captured scene images and videos, and are usually sensitive to font scale changes and background interference which widely exists in scene text. Although some OCR systems have started to support scene character recognition, the recognition performance is still much lower than the recognition for scanned documents.

Many algorithms were proposed to improve scene-image-based text character recognition. Weinman et al. combined the Gabor-based appearance model, a language model related to simultaneity frequency and letter case, similarity model, and lexicon model to perform scene character recognition. Neumann et al. proposed a real time scene text localization and recognition method based on extremal regions Smith et al. built a similarity model of scene text characters based on SIFT, and maximized posterior probability of similarity constraints by integer programming. Mishra et al. adopted conditional random field to combine bottom-up character recognition and top-down word-level recognition.

Luet al. [13] modeled the inner character structure by defining a dictionary of basic shape codes to perform character and word retrieval without OCR on scanned documents. Coates et al. [4] extracted local features of character patches from an unsupervised learning method associates with a variant of K-means clustering, and pooled them by cascading sub-patch features. In complete performance evaluation of scene text character recognition was carried out to design a discriminative feature representation of scene text character structure. In a part-based tree structure model was designed to detect text characters under Latent-SVM [8], and recognize text words from text regions under conditional random field. In Scale Invariant Feature Transform (SIFT) feature matching was adopted to recognize text characters in different languages, and a voting and geometric verification algorithm was presented to filter out false positive matches.

In generic object recognition method was imported to extract scene text information. A dictionary of words to be spot is built to improve the accuracy of detection and recognition. Character structure was modeled by HOG features and cross correlation analysis of character similarity for text recognition and detection. In Random Ferns algorithm was used to perform character detection and constructed a system for query-based word detection in scene images.

Differences include:

A multimedia object consists of a large number of bits, with considerable redundancy. Thus, the watermark has a large cover in which to hide. A database relation consists of tuples, each of which represents a separate object. The watermark needs to be spread over these separate objects.The relative spatial/temporal positioning of various pieces of a multimedia object typically does not change. Tuples of a relation on the other hand constitute a set and there is no implied ordering between them.Portions of a multimedia object cannot be dropped or replaced arbitrarily without causing perceptual changes in the object. However, the pirate of a relation can simply drop some tuples or substitute them with tuples from other relations.

Claudio Lucchese et al [2001] [2] describe the sharing is an important aspect of scientific or business collaboration. However, data owners are also concern with the protection of their rights on the datasets, which is many cases have been obtained after expensive and laborious procedures. The ease of data exchange through the Interne has compounded the need to assemble technological mechanisms for effectively protecting one’s intellectual or pragmatic property. Trajectories abound in applications such as GPS tracking experiments, video and motion capture data, and even image shapes can be considered as 2-dimensional trajectories. We provide ownership assurances on such datasets using watermarking principles. While there is a rich literature on watermarking for multimedia datasets, previous work is primarily concerned with watermarking a single object and not a collection of objects. Here, we consider the watermarking problem from a new perspective, by focusing on the additional maintenance of the inter-relationship between objects.

The technique embeds a secret key in each of the dataset objects, distorting them imperceptibly, while taking special consideration in retaining the original neighboring object. We call this operation Neighbor Preserving (NP) watermarking. Guaranteeing preservation of the nearest objects is very important for an array of search and mining operations, such as similarity search or Nearest-Neighbor(NN)-classification. Contrary to privacy-preserving approaches for data-mining that first add noise and then reconstruct the original data distributions based on the known noise model, our approach learns/calculates the largest amount of noise that can be added, so that nearest neighbors are not distorted. While the naive algorithms for determining the said watermarking power are costly, we show efficient ways of speeding up the process making it more than 2 orders of magnitude faster, thus allowing the technique to be applicable to large datasets.

Vıctor R. Doncel, Nikos et al [2000] [3] describe a watermark is a hidden information within a digital signal, used primarily for copyright protection of multimedia data. Its main features are the imperceptibility of the imposed medications and its persistence against processing (attacks) that may result in its removal, either intentionally or unintentionally. A general framework for digital water marking has been presented, whereas provides an excellent overview of the watermarking principles and techniques. Digital watermarking has been mainly applied to still image, audio and video data. However, little work has been done in watermarking vector graphics data, that are typically used in Geographic Information Systems (GIS) or in Computer Aided Design (CAD).

Digital watermarking is a polygonal line, which are a key graphics primitive in vector graphics data and thus can be used for the copyright protection of such data. Furthermore, the method can be used for the watermarking of MPEG-4 natural video data by watermarking the outline of the Video Objects in MPEG-4 stream. In that case, the method should be accompanied by a way of extrapolating existing textures in case the watermarked boundary defines a bigger area than the original one.

This paper extends the work presented. The same embedding method is adopted here, and efforts focus the design of a new, enhanced performance detector. Theoretical and experimental analysis show that a substantial improvement in detection performance can be achieved if the statistics of the watermarked polygon are considered.

Rakesh Agrawal and Jerry Kiernan et al [2006] describe watermarking database relations to deter their piracy, identify the unique characteristics of relational data which pose new challenges for watermarking, and provide desirable properties of a watermarking system for relational data. A watermark can be applied to any database relation having attributes which are such that changes in a few of their values do not affect the applications. An effective watermarking technique geared for relational data. This technique ensures that some bit positions of some of the attributes of some of the tuples contain specific values. The tuples, attributes within a tuples, bit positions in an attribute, and specific bit values are all algorithmically determined under the controlof a private key known only to the owner of the data.

N. F. Johnson, Z. Duric [2000] describe the piracy of digital assets such as software, images, video, audio and text has long been a concern for owners of these assets. Protection of these assets is usually based upon the insertion of digital watermarks into the data. The watermarking software introduces small errors into the object being watermarked. These intentional errors are called marks and all the marks together constitute the watermark. The marks must not have a significant impact on the usefulness of the data and they should be placed in such a way that a malicious user cannot destroy them without making the data less useful. Thus, watermarking does not prevent copying, but it deters illegal copying by providing a means for establishing the original ownership of a redistributed copy. The increasing use of databases in applications beyond “behind-the-firewalls data processing” is creating a similar need for watermarking databases.

III. EXISTING SYSTEM

Firstly, given a synthesized character patch from the training set, the existing system obtains **character boundary** and **character skeleton** by applying **discrete contour evolution** (DCE) and **skeleton pruning** on the basis of DCE. The DCE simplifies characters into polygons of visual parts with a small number of vertices, and we define the polygon as a character boundary.

Then skeleton pruning is performed to obtain a refined character skeleton. Since DCE and skeleton pruning are invariant to deformation and scaling, they provide stable results of boundary and skeleton, which are used as a universal description of the structure of characters in the same category but with different fonts and sizes.

However, DCE-based skeleton pruning can only obtain a coarse character skeleton by boundary analysis, but not predict its stroke configurations.Thus it further locates all the strokes that compose the character skeleton in the next two steps. Secondly, it estimates the stroke width and orientation on sample points of character boundary. Next, stroke width and orientation at each boundary sample point b is estimated.

Drawbacks

Only character recognition is considered.

All the text will be extracted if color is same.

Image type is not discussed.

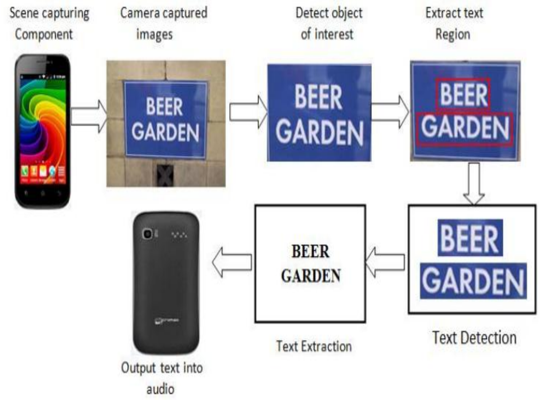
Spaces between characters are not measured.

IV. PROPOSED WORK

Text detection and recognition used to detect text in complex background images. It takes text image as input and then applying preprocessing methods on it to remove noise from image by converting color image to gray ,binarization which helps to efficient and accurate text identification from image which is input to OCR, within pre-processing if some part text data will loss them by thinning andscaling is performed by connectivity algorithm. Then we get connected text character from image. Then text recognition is done.

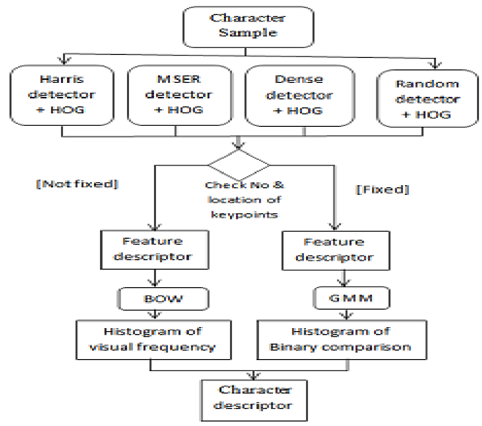
The proposed framework is divided into three stages. Here applied text detection and text recognition to the image and recognize. The text detection uses to quickly extract text region in images with a very less false positive rate. To provide the recognition for accurate result proposed system to test the text image is segmented, assuming a different number of classes in the image each time.

It is a demanding problem to automatically localize objects and text Region of Interest from captured images with cluttered backgrounds, because text in natural scene images is most likely surrounded by various surroundings outlier noise and text characters usually appear in different fonts and colors. For the content orientations, this thesis assumes that text strings in scene images keep in the section of vertical locations.



Several algorithms have been developed for localization of text area in scene images. We can divide them into categories component Based and operations Based. To recognize and extract text from difficult backgrounds with multiple and variable text patterns, here propose a text localization algorithm that combines area-based layout analysis and horizontal-based text classifier preparation, which define characteristic maps based on stroke orientations and boundary distributions. To generate delegate and discriminative text features to decide text characters from environment outliers.

FLOW CHART



Algorithm

1. Initially set SWT = ∞
2. Find edge by canny edge detector.
3. Follow the ray $r=p+n \cdot dp$, $n>0$ until another edge is found.
4. If $dq= -dp \pm \Pi/6$ then $SWT= |p - q|$ and dp++ else discard the ray.
5. If SWT ratio <= 3 then group neighbouring pixels.
6. If two letters are having similar stroke width, they can be grouped.

7. The output is a set of rectangles designating bounding boxes for detected words.

8. Search the text on web or in database.

9. Match the word, and retrieve the related information.

10. Display retrieved information on mobile screen

V. CONCLUSION

In this project is detects text regions from natural scene image, and recognizes text information from the detected text regions. In scene text detection, layout analysis of color decomposition and horizontal alignment is performed to search for image regions of text strings. In scene text recognition, two schemes, text understanding and text retrieval, are respectively proposed to extract text information from surrounding environment. The proposed system character descriptor is effective to extract representative and discriminative text features for both recognition schemes. The model text character structure for text retrieval scheme designed a novel feature representation, stroke configuration map, based on boundary and skeleton. Quantitative experimental results demonstrate that our proposed method of scene text recognition outperforms most existing methods.

In addition, varying size of character images are saved in training set. Threshold value is set so that different size of character other than in training set can also be extracted from the given image. Any image type can be given as source image. Spaces between characters are also measured so that words can be extracted out. In addition, characters can be of any colors in the give image.

REFERENCE

[1] X. Bai, L. J. Latecki, and W.-Y. Liu, "Skeleton pruning by contour partitioning with discrete curve evolution," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 3, pp. 449–462, Mar. 2007.

[2] R. Beaufort and C. Mancas-Thillou, "A weighted finite-state framework for correcting errors in natural scene OCR," in Proc. 9th Int. Conf. Document Anal. Recognit., Sep. 2007, pp. 889–893.

[3] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image Process., vol. 13, no. 1, pp. 87–99, Jan. 2004.

[4]A. Coates et al., "Text detection and character recognition in scene images with unsupervised feature learning," in Proc. ICDAR, Sep. 2011, pp. 440–445.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2005, pp. 886–893.

[6] T. de Campos, B. Babu, and M. Varma, "Character recognition in natural images," in Proc. VISAPP, 2009.

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

[8] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[9] T. Jiang, F. Jurie, and C. Schmid, "Learning shape prior models for object matching," in Proc. CVPR, Jun. 2009, pp. 848–855.

[10] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Johsi, "Text extraction".

[11] L. J. Latecki and R. Lakamper, "Convexity rule for shape decomposition based on discrete contour evolution," Comput. Vis. Image Understand., vol. 73, no. 3, pp. 441–454, 1999.

[12] Y. Liu, J. Yang, and M. Liu, "Recognition of QR code with mobile phones," in Proc. CCDC, Jul. 2008, pp. 203–206.

[13] S. Lu, L. Li, and C. L. Tan, "Document image retrieval through word shape coding," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 11, pp. 1913–1918, Nov. 2008.