Eleyon
SOUTH ASIAN JOURNAL OF ENGINEERING AND TECHNOLOGY

Full Length Article

SCENE TEXT RECOGNITION IN IMAGES BY CHARACTER STRUCTURE CONFIGURATION AND DESCRIPTOR

[a] K.DINESHKUMAR ,[a] K.KUMARESAN, [a] V.SENTHILKUMAR, [b] S.KUBENDRAN, [b] G.LINGESH, [b] M.PASUPATHI, [b] K.PRAVEEN

[a] *Assistant Professor,Department of Computer Science and Engineering, Sri K .S .R College of Engineering, Namakkal, Tamil Nadu-637 215, India*

[b] *Student,Department of Computer Science and Engineering, Sri K .S .R College of Engineering, Namakkal, Tamil Nadu-637 215, India*

***Corresponding Author**

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 paper 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 paper 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 system is developed to show the effectiveness of our proposed method on scene text information extraction from nearby objects.
.

Keywords: Character Descriptor, Text Detection, Scene Images.

1 Introduction

Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to- day life for various applications. Various techniques have been developed in Image Processing during the last four to five

decades. Most of the techniques are developed for enhancing images obtained from unmanned spacecrafts, space probes and military reconnaissance flights. Image Processing systems are becoming popular due to easy availability of powerful personnel computers, large size memory devices, graphics softwares etc.

Image analysis is concerned with making quantitative measurements from an image to produce a description. In the simplest form, this task could be reading a label on a grocery item, sorting different parts on an assembly line, or measuring the size and orientation of blood cells in a medical image. More advanced image analysis systems measure quantitative information and use it to make a sophisticated decision, such as controlling the arm of a robot

to move an object after identifying it or navigating an aircraft with the aid of images acquired along its trajectory.

Image analysis techniques require extraction of certain features that aid in the identification of the object. Segmentation techniques are used to isolate the desired object from the scene so that measurements can be made on it subsequently. Quantitative measurements of object features allow classification and description of the image.

Image segmentation is the process that subdivides an image into its constituent parts or objects. The level to which this subdivision is carried out depends on the problem being solved, i.e., the segmentation should stop when the objects of interest in an application have been isolated e.g., in autonomous air-to-ground target acquisition, suppose our interest lies in identifying vehicles on a road, the first step is to segment the road from the image and then to segment the contents of the road down to potential vehicles. Image thresholding techniques are used for image segmentation.

Classification is the labeling of a pixel or a group of pixels based on its grey value. Classification is one of the most often used methods of information extraction. In Classification, usually multiple features are used for a set of pixels i.e., many images of a particular object are needed. In Remote Sensing area, this procedure assumes that the imagery of a specific geographic area is collected in multiple regions of the electromagnetic spectrum and that the images are in good registration. Most of the information extraction techniques rely on analysis of the spectral reflectance properties of such imagery and employ special algorithms designed to perform various types of 'spectral analysis'. The process of multispectral classification can be performed using either of the two methods: Supervised or Unsupervised.

In Supervised classification, the identity and location of some of the land cover types such as urban, wetland, forest etc., are known as priori through a combination of field

works and toposheets. The analyst attempts to locate specific sites in the remotely sensed data that represents homogeneous examples of these land cover types. These areas are commonly referred as TRAINING SITES because

the spectral characteristics of these known areas are used to 'train' the classification algorithm for eventual land cover mapping of reminder of the image. Multivariate statistical parameters are calculated for each training site. Every pixel both within and outside these training sites is then evaluated and assigned to a class of which it has the highest likelihood of being a member.

In an Unsupervised classification, the identities of land cover types has to be specified as classes within a scene are not generally known as priori because ground truth is lacking or surface features within the scene are not well defined. The computer is required to group pixel data into different spectral classes according to some statistically determined criteria. The comparison in medical area is the labeling of cells based on their shape, size, color and texture, which act as features. This method is also useful for MRI images.

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

II. RELATED WORKS

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 are 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.

The main objective of this paper is,

- To extract representative and discriminative features from character patches.
- It combines several feature detectors (Harris-Corner, Maximal Stable Extremal Regions (MSER), and dense
- sampling) and Histogram of Oriented Gradients (HOG) descriptors.
- To generate a binary classifier for each character class in text retrieval.
- To novel stroke configuration from character boundary and skeleton to model character structure.

Rakesh Agrawal Jerry Kiernan et al., 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.

Claudio Lucchese et al., 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.

Vıctor R. Doncel, Nikos et al., 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).

Lukas Neumann and Jiri Matas., In this paper, we present an end-to-end real-time1 text localization and recognition method which achieves stateof-the-art results on standard datasets. The real-time performance is achieved by posing the character detection problem as an efficient sequential selection from the set of Extremal Regions (ERs). The ER detector is robust against blur, low contrast and illumination, color and texture variation2. Its complexity is O(2pN), where p denotes number of channels (projections) used. In the first stage of the classification, the probability of each ER being a character is estimated using novel features calculated with O(1) complexity and only ERs with locally maximal probability are selected for the second stage, where the classification is improved using more computationally expensive features. A highly efficient exhaustive search with feedback loops (adapted from [15]) is then applied to group ERs into words and select the most probable character segmentation.

Rakesh Agrawal and Jerry Kiernan et al., 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 control of a private key known only to the owner of the data.

N. F. Johnson, Z. Duric 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. METHODOLOGY

In the proposed effective algorithms is text recognition from detected text regions in scene image. In scene text detection process, we apply the methods presented in our previous work. Pixel-based layout analysis is adopted to extract text regions and segment text characters in images,

based on color uniformity and horizontal alignment of text characters.

In text recognition process, we design two schemes of scene text recognition. The first one is training a character recognizer to predict the category of a character in an image patch. The second one is training a binary character classifier for each character class to predict the existence of this category in an image patch. The two schemes are compatible with two promising applications related to scene text, which are text understanding and text retrieval.

Text understanding is to acquire text information from natural scene to understand surrounding environment and objects. Text retrieval is to verify whether a piece of text information exists in natural scene. These two applications can be widely used in smart mobile device. The proposed method combines scene text detection and scene text recognition algorithms recognition the character recognizer, text understanding is able to provide surrounding text information for mobile applications and by the character classifier of each character class, text retrieval is able to help search for expect objects from environment.

Similar to other methods, our proposed feature representation is based on the different text format feature descriptors schemes. Different from other methods, our method combines the low-level feature

In addition, characters can be of any colors in the give image. Using Character Skeleton and Character Boundary extraction, the character is recognized

descriptors with stroke configuration to model text character structure.

A. TRAINING SET - ALPHABET IMAGE ADDITION The add alphabet module is used to add the alphabet

character information in the application details such as character, image file path and image and store these information in the “images” file.

B. TRAINING SET - NUMBER IMAGE ADDITION The add number form is used to add the number

information in the application details such as character, image file path and image and store these information in the “images” file.

C. TRAINING SET - SPECIAL CHARACTER IMAGE ADDITION

The add special character form is used to add the special

character information in the application details such as special character, image file path and image and store these information in the “images” file.

D. SELECT IMAGE

In this module, the image file is selected. The image

type may be any format (e.g., bmp, jpg, gif, tiff, etc). Both gray scale and color image can be selected.

E. SCENE TEXT RECOGNITION

In this module, the selected image file is taken for

processing. Then threshold value is set so that different size of character other than in training set can also be extracted from the given image. Spaces between characters are also measured so that words can be extracted out.

IV. CONCLUSION

In this paper 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.

REFERENCES

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

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

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

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

pp. 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. 9,1627–1645, Sep. 2010.

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

[8] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Johsi, "Text extraction and document image segmentation using matched wavelets andMRF model," IEEE Trans. Image Process., vol. 16, no. 8, pp. 2117–2128, Aug. 2007.

[9] 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.

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

[11] . 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.

[12] S. M. Lucas, A. Panaretos, L. Sosa, A. Tang, S.Wong, and R. Young, "ICDAR 2003 robust reading competitions," in Proc. Int. Conf.Document Anal. Recognit., Aug. 2003, pp. 682– 687.

[15] A. Mishra, K. Alahari, and C. V. Jawahar, "Top-down and bottom-up cues for scene text recognition," in Proc. IEEE Conf. Comput. Vis.

Pattern Recognit., Jun. 2012, pp. 1063–6919.

[16] N. Nikolaou and N. Papamarkos, "Color reduction for complex document images," Int. J. Imag. Syst.Technol., vol. 19, no. 1, pp. 14–26, 2009.