International Journal on Recent and Innovation Trends in Computing and Communication ISSN 2321 – 8169
Volume: 1 Issue: 8 652 – 655

A New Approach for Text String Detection from Natural Scenes By Grouping & Partition

Regula SriLakshmi[1], S.Swetha[2], Thillai Nayagi[3]

Asst.Professor[1,2,3], Avanthi Institute of Engineering & Technology

Abstract: **In this paper we have reviewed and analyzed different methods to find strings of characters from natural scene images. We have reviewed different techniques like extraction of character string regions from scenery images based on contours and thickness of characters, efficient binarization and enhancement technique followed by a suitable connected component analysis procedure, text string detection from natural scenes by structure- based partition and grouping, and a robust algorithm for text detection in images.**

It is assumed that characters have closed contours, and a character string consists of characters which lie on a straight line in most cases. Therefore, by extracting closed contours and searching neighbors of them, character string regions can be extracted;

Image binarization successfully processes natural scene images having shadows, non-uniform illumination, low contrast and large signal-dependent noise. Connected component analysis is used to define the final binary images that mainly consist of text regions.

One technique chooses the candidate text characters from connected components by gradient feature and color feature. The text line grouping method performs Hough transform to fit text line among the centroids of text candidates. Each fitted text line describes the orientation of a potential text string. The detected text string is presented by a rectangle region covering all characters whose centroids are cascaded in its text line. To improve efficiency and accuracy, our algorithms are carried out in multi-scales. The proposed methods outperform the state-of-the-art results on the public Robust Reading Dataset, which contains text only in horizontal orientation. Furthermore, the effectiveness of our methods to detect text strings with arbitrary orientations is evaluated on the Oriented Scene Text Dataset collected by ourselves containing text strings in no horizontal orientations.

Index Terms—**Adjacent character grouping, character property, image partition, text line grouping, text string detection, text string structure.**

____*____

II. INTRODUCTION

A lot of objects on which characters are written exist in our living environment. We humans get much information from these texts. It is expected that robots act in our living environment and support us in the future. If robots can read text on objects such as Packages and signs, robots can get information from them, and they can use it in their activation and support for us. Owing to the progress of OCR, computers have been able to read text in images. However, images have many non-character textures, and they make it difficult to read text by OCR. To cope with that problem, we need to extract character string regions from images. Indexing images or videos requires information about their content. This content is often strongly related to the textual information appearing in them, which can be divided into two groups: 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. Text produced separately from the image is in general a very good key to understand the image which is called artificial text. In contrast to scene text, artificial text is not only an important source of information but also a significant entity for indexing and retrieval purposes. Natural scene images contain text information which is often required to be automatically recognized and processed. Localization of text and simplification of the background in images is the main objective of automatic text detection approaches. However, text localization in complex images is an intricate process due to the often bad quality of images, different backgrounds or different fonts, colors, sizes of texts appearing in them. In order to be successfully recognizable by an OCR system, an image having text must fulfill certain requirements, like a monochrome text and background where the background-to- text contrast should be high. This paper strives toward methodologies that aid automatic detection, segmentation and recognition of visual text entities in complex natural scene images. The algorithms of text extraction from images can be broadly classified under three types. They are gradient feature based, color segmentation based, and texture analysis based. The gradient feature based algorithm is based on the idea that pixels which have high gradient are the candidates of characters since edges exist between a character and background. One technique makes use of closed contours to extract string regions.

Different from document images, in which text characters are normalized into elegant poses and proper resolutions, natural scene images embed text in arbitrary shapes, sizes, and orientations into complex background, as shown in

International Journal on Recent and Innovation Trends in Computing and Communication ISSN 2321 – 8169
Volume: 1 Issue: 8 652 – 655

Fig. 1. It is impossible to recognize text in natural scene images directly because the off-the-shelf OCR software cannot handle complex background interferences and nonorienting text lines. Thus, we need to detect image regions containing text strings and their corresponding orientations. This is compatible with the detection and localization procedure described in the survey of text extraction algorithms . With knowledge of text string orientations, we can normalize them to horizontal. Some algorithms of scene text normalization are introduced . However, the algorithms described in this paper will focus on text detection.

Fig 1.

STRUCTURE-BASED PARTITION AND GROUPING BASED METHOD

The proposed framework consists of two main steps, given here.

Step 1) Image partition to find text character candidates based on gradient feature and color uniformity. In this step, Chucai Yi and YingLi Tian propose two methods to partition scene images into binary maps of non overlapped connected components: gradient-based method and color-based method. A post processing is then performed to remove the connected components which are not text characters by size, aspect ratio, and the number of inner holes.

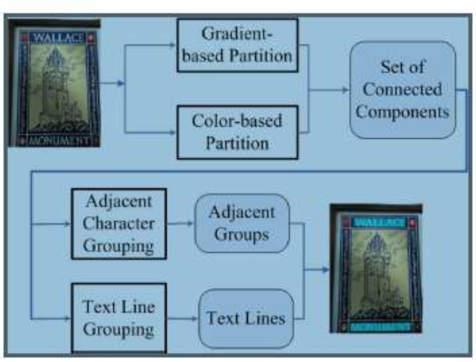


Fig. 2. Flowchart of the proposed framework of text string detection.

Step 2) Character candidate grouping to detect text strings based on joint structural features of text characters in each text string such as character sizes, distances between two neighboring characters, and character alignment. In this step, Chucai Yi and YingLi Tian propose two methods of structural analysis of text strings: adjacent character grouping method and text line grouping method. Adjacent Character grouping method:Text strings in natural scene images usually appear in alignment, each text character in a text string must possess character siblings at adjacent positions. The structure features among sibling characters can be used to determine whether the connected components belong to text characters or unexpected noises. Here, five constraints are defined to decide whether two connected components are siblings of each other.

1) Considering the capital and lowercase characters, the height ratio falls between 1/T1 and T1.

2) The distance between two connected components should not be greater than T2 times the width of the wider one.

3) For text strings aligned approximately horizontally, the difference between Y-coordinates of the connected component centroids should not be greater than T3 times the height of the higher one.

4) Two adjacent characters usually appear in the same font size, thus their area ratio should be greater than 1/T4 and less than T4.

5) If the connected components are obtained from gradient based partition, the color difference between them should be lower than a predefined threshold Ts.In their system Chucai Yi and YingLi Tian set T1=T4=2, T2=3, T3=0.5, T5=40 For two connected components C and C' they can be grouped together as sibling components if the above five constraints are satisfied. When C and C' are grouped together, their sibling sets will be updated according to their relative locations. That is, when C is located on the left of C',

C' will be added into the right-sibling set of C, which is simultaneously added into the left -sibling set of C' The reverse operation will be applied when C is located on the right of C'.

III. IMAGE PARTITION

To extract text information from a complex background, image partition is first performed to group together pixels that belong to the same text character, obtaining a binary map of candidate character components. Based on local gradient features and uniform colors of text characters, we design a gradient-based partition algorithm and a color-based partition algorithm, respectively

A. Gradient-Based Partition by Connecting Paths of Pixel Couples Although text characters and strings vary in font, size, color, and orientation, they are composed of strokes which are rec- Fig. 3. (a) Examples of pixel couples. (b) Connecting paths of all pixel couples are marked as white foreground while other pixels are marked as black background. tangle connected components with closed-width boundaries and uniform torso intensities each pixel is mapped to the width of the stroke in which it is located, and then the consistency of the stroke width is used to extract a candidate character component. In our method, each pixel is mapped to the connecting path of a pixel couple, defined by two edge pixels and on an edge map with approximately equal gradient magnitudes and opposite directions, as shown in Fig.. Each pixel couple is connected by a path. Then the distribution of gradient magnitudes at pixels of the connecting path is computed to extract candidate character component. depicts that a character boundary consists of a number of pixel couples. We model the character by distribution of gradient magnitudes and stroke size including width, height, and aspect ratio. The partitioned components are calculated from connecting path of pixel couple across the pixels with small gradient magnitudes.

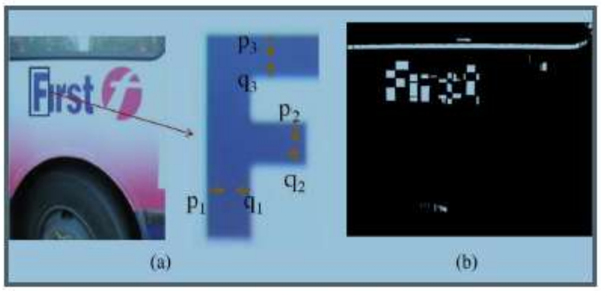


Fig. 3. (a) Examples of pixel couples. (b) Connecting paths of all pixel couples are marked as white foreground while other pixels are marked as black background.

IV. CONNECTED COMPONENTS GROUPING

The image partition creates a set of connected components from an input image, including both text characters and unwanted noises. Observing that text information appears as one or more text strings in most natural scene images, we perform heuristic grouping and structural analysis of text strings to distinguish connected components representing text characters from those representing noises. Assuming that a text string has at least three characters in alignment, we develop two methods to locate regions containing text strings: adjacent character grouping and text line grouping, respectively. In both algorithms, a connected component is described by four metrics: , , , and . In addition, we use to represent the distance between the centroids of two neighboring characters.

A. Adjacent Character Grouping Text strings in natural scene images usually appear in alignment, namely, each text character in a text string must possess character siblings at adjacent positions. The structure features among sibling characters can be used to determine whether the connected components belong to text characters or unexpected noises. Here, five constraints are defined to decide whether two connected components are siblings of each other.

1) Considering the capital and lowercase characters, the height ratio falls between $1/ T_1$and T_1

2) Two adjacent characters should not be too far from each other despite the variations of width, so the distance between two connected components should not be greater than T_2times the width of the wider one.

3) For text strings aligned approximately horizontally, the difference between -coordinates of the connected component centroids should not be greater than T_3 times the height of the higher one.

4) Two adjacent characters usually appear in the same font size, thus their area ratio should be greater than $1/ T_4$ and less than T_4 .

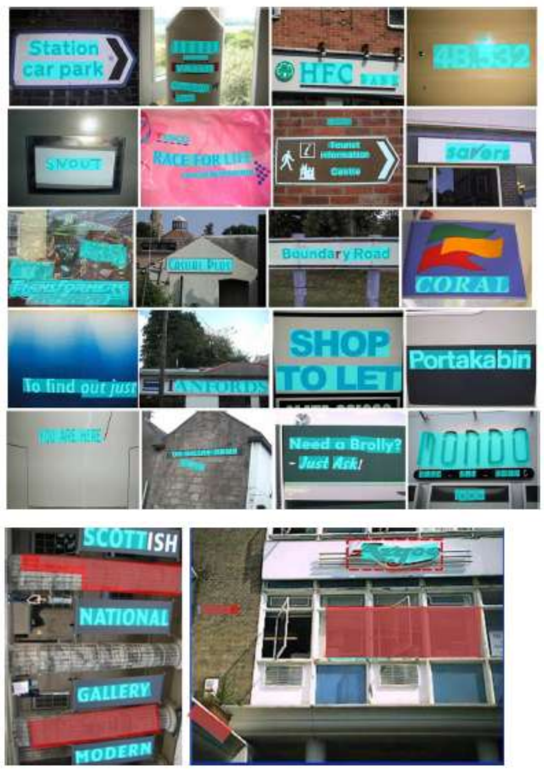
5) If the connected components are obtained from gradient based partition as described in Section III-A, the color difference between them should be lower than a predefined threshold because the characters in the same string have similar colors.

V. EXPERIMENTAL RESULTS

Two datasets are employed to evaluate the proposed algorithms. The first is the Robust Reading Dataset1 from ICDAR 2003. In this dataset, there are 509 images in total, in which 258 images are prepared for training and 251 images for testing. The image regions containing text strings are labeled in a XML file. Each image contains about four text regions on average. All of the text strings in this dataset are in horizontal. In our testing, we selected 420 images which are compatible with the assumption that a text string contains at least three characters with relatively uniform color. Furthermore, to verify that text line grouping can detect text strings with arbitrary orientations, we collect 89 scene images with non horizontal text strings to construct the OSTD. The resolutions of most images are from 600 *450 to 1280 *960. The average number of text strings is two on each image. Text string regions are also manually labeled in the .xml file. This OSTD dataset contains colorful logos, indoor scenes, and street views.

Results and Discussions

The experimental results on the Robust Reading dataset are illustrated, where blue bars denote results of GA, cyan bars denote results of CA, yellow bars denote results of GT, and red bars denote results of CT. The average time of text string detection is presented in the upper boxes. The combination of color-based partition and adjacent character grouping (CA) achieves the highest precision and recall. In most of the cases, color uniformity acts as a stronger indicator to distinguish the connected components of text characters from surrounding background. However color-based partition takes more computing time than gradient-based partition. Also, color-based partition makes adjacent character grouping be performed in each of the color layers. Color-based partition still performs better when adjacent character grouping is replaced by the text line grouping illustrates that text line grouping gives lower efficiency and precision than the adjacent character grouping for either partition. Adjacent character grouping is supported by the information of text orientations while text line grouping is performed for arbitrary text orientations, so its calculation cost is more expensive. Meanwhile, the indetermination of text orientation produces more false positive fitted lines.



VI. CONLCUSION

In this paper we have reviewed and analyzed different methods to find strings of characters from natural scene images. We have reviewed different techniques like extraction of character string regions from scenery images based on contours and thickness of characters, efficient binarization and enhancement technique followed by a suitable connected component analysis procedure, text string detection from natural scenes by structure-based partition and grouping, and a robust algorithm for text detection in images.Also, we have presented an approach to detect, localize, and extract texts appearing in grayscale or color Images as well as locate text strings with arbitrary orientations.

Our future work will focus on developing 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 wayfinding of visually impaired people.

REFERENCES

[1] Extraction of Character String Regions from Scenery Images Based on Contours and Thickness of Characters in MVA2011 IAPR Conference on Machine Vision Applications, June 13 – 1 5, 2011, Nara, APAN
[2] Towards Text Recognition in Natural Scene Images by Basilios Gatos, Ioannis Pratikakis and Stavros Perantonis.
[3] Text String Detection from Natural Scenes by Structure-based Partition and Grouping by Chucai Yi nd YingLi Tian, Senior Member, IEEE
[4] A Robust Algorithm for Text Detection in Images by Julinda Gllavata, Ralph Ewerth and Bernd Freisleben
[5] J. Banerjee, A. M. Namboodiri, and C. V. Jawahar, "Contextual restoration of severely degraded document images," in Proc. IEEEConf. Comput. Vis. Pattern Recognition., 2009,
[6] 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.

About the Authors

Regula Srilakshmi Working as Asst.Professor CSE Department of Avanthi Institute of Engineering & Technology, Hyderabad. She has vast experience in research work.

S.Swetha Working as Asst.Professor CSE Department of Avanthi Institute of Engineering & Technology,Hyderabad. She has vast experience in research work.

Thillai Nayagi Working as Asst.Professor CSE Department of Avanthi Institute of Engineering & Technology,Hyderabad. She has vast experience in research work.