Text string detection from natural scenes english language essay

Linguistics, English

Mr. Rampurkar Vyankatesh Vijaykumar., VPCOE Baramati. Mrs. Gyankamal J. Chhajed. Asst. Prof., VPCOE Baramati.

Abstract

In this paper we have presented different methods to find strings of characters from natural scene images. We have presented different techniques like extraction of character string regions from scenery images based on siblings 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. Color-based partition performs better than gradientbased

partition, but it takes more time to detect text strings

on each color layer. The text line grouping is able to extract

text strings with arbitrary orientations. The combination of

color-based partition and adjacent character grouping (CA)

gives the best performance.

Key terms

Connected Components Analysis, Adjacent Character

grouping, Image partition, Text line grouping, Text string

detection, Color reduction technique, Edge detection.

1. Introduction

Text information in natural scene images serves as importantclues for many image-based applications such asscene understanding, content-based image retrieval, assistivenavigation, and automatic geocoding. However, locating text from a complex background with multiplecolors is a challenging task. A lot of objects on whichcharacters are written exist in our living environment. We humans get much information from these texts. Itis expected that robots act in our living environmentand support us in the future. If robots can read texton objects such as packages and signs, robots can getinformation from them, and they can use it in their activationand support for us. Owing to the progress ofOCR, computers have been able to read text in images. However, images have many non-character textures, andthey make it difficult to read text by OCR. To cope withthat problem, we need to extract character string regionsfrom images. Indexing images or videos requires informationabout their content. This content is often stronglyrelated to the textual information appearing in them, which can be divided into two groups: Text appearingaccidentally in an image that usually does not representanything important related to the content of the image. Such texts are referred to as scene text. Text producedseparately from the image is in general a very good keyto understand the image which is called artificial text. In contrast to scene text, artificial text is not only animportant source of information but also a significantentity for indexing and retrieval purposes. Natural sceneimages contain text information which is often requiredto be automatically recognized and processed. Localizationof text and simplification of the background in imagesis the main objective of automatic text detection approaches. However, text localization in complex imagesis an intricate process due to the often bad quality of images, different backgrounds or different fonts, colors, sizesof texts appearing in them. In order to be successfullyrecognizable by an OCR system, an image having textmust fulfill certain requirements, like a monochrome textand background where the background-to-text contrastshould be high. This paper strives toward methodologiesthat aids automatic detection, segmentation and recognitionof visual text entities in complex natural sceneimages. The algorithms of text extraction from images can bebroadly classified under three types. They are gradientfeature based, color segmentation based, and textureanalysis based. The gradient feature based algorithmis based on the idea that pixels which have high gradientare the candidates of characters since edges existbetween a character and background. In this paper, 1Chucai Yi and YingLi Tian propose a new framework toextract text strings with multiple sizes and colors, andarbitrary orientations from scene images with a complexand cluttered background [3]. The proposed frameworkconsists of two main steps: a) image partition to findtext character candidates based on gradient feature andcolor uniformity. In this step, Chucai Yi and YingLiTian propose two methods to partition scene images intobinary maps of non overlapped connected components: gradient-based method and color-based method b) Charactercandidate grouping to detect text strings basedon joint structural features of text characters in eachtext string such as character sizes, distances between twoneighboring characters, and character alignment. In thisstep, Chucai Yi and YingLi Tian propose two methodsof structural analysis of text strings: adjacent charactergrouping method and text line grouping method. Figure 1: Examples of text in natural scene images [3]

2. Related Work

One of the characteristics of common characters in realimages is that most of them are capable of producingclosed contour when edge extraction process is applied. So Tomohiro Nishino[1] takes an approach to detectclosed contours from images. Moreover, it is assumedthat a character string consists of characters which lieon a straight line. From these assumptions, characterstring should be found from regions where closed contoursare arranged with regularly. Assuming that charactersincluded in a character string are aligned horizontally, string regions can be extracted by detecting horizontallyaligned closed contours. Tomohiro Nishino explainshow to detect the horizontally aligned closed contour. First, a circumscribed rectangle of a closed contouris calculated. Next, the rectangle is slid to right by somepixels as much as the width of the rectangle. If the rectangleincludes the center of a circumscribed rectangle ofanother closed contour, these two closed contours are assumedto be aligned horizontally and to be included inthe same character string. Closed contours which are isolatedare assumed not to be characters. Circumscribedrectangles of each character string are assumed to bestring regions. By this process, string regions of horizontallyaligned closed contours are extracted. Both closedand unclosed contours which lie left or right of characterstring regions are extracted, and each thickness iscalculated. These contours are added to the characterstring region if they have the similar thickness to that ofcharacters in that region. Basilios Gatos[2] produces gray level image and invertedgray level image. Then, calculate the two correspondingbinary images using an adaptive binarizationand image enhancement technique. In the sequel, theproposed technique involves a decision function that indicateswhich image between binary images contains textinformation. An efficient algorithm which can automatically detect, localize and extract horizontally aligned text in images(and digital videos) with complex backgrounds is presentedby by Julinda Gllavata, Ralph Ewerth and BerndFreisleben[4]. The proposed approach is based on theapplication of a color reduction technique, a method foredge detection, and the localization of text regions usingprojection profile analyses and geometrical properties. 2Figure 2: The flowchart of the framework [3]

3. Programmer's design

Fig. 2 depicts the flowchart of the framework. Theproposed framework consists of two main steps, givenhere. Step 1) Image partition to find text character candidatesbased on gradient feature and color uniformity. In this step, Chucai Yi and YingLi Tian propose twomethods to partition scene images into binary maps ofnonoverlapped connected components: gradient-basedmethod and color-based method. A post processing isthen performed to remove the connected componentswhich are not text characters by size, aspect ratio, andthe number of inner holes. Step 2) Character candidate grouping to detect textstrings based on joint structural features of text charactersin each text string such as character sizes, distancesbetween two neighboring characters, and character alignment. In this step, Chucai Yi and YingLi Tian proposetwo methods of structural analysis of text strings: adjacentcharacter grouping method and text line groupingmethod.

4. Image Partition

To extract text information from a complex background, image partition is first performed to group togetherpixels that belong to the same text character, obtaininga binary map of candidate character components. Based on local gradient features and uniform colors oftext characters, we design a gradient-based partition algorithmand a color-based partition algorithm, respectively. Figure 3: We compare results of four morphological operatorswith result of our gradient-based partition.[3]Figure 4: Some examples of color-based partition, wherethe left column contains original images and othercolumns contain the corresponding dominant color layers.[3]

5. Connected Components Grouping

The image partition creates a set of connected componentsS from an input image, including both text charactersand unwanted noises. Observing that text infor-3mation appears as one or more text strings in most naturalscene images, we perform heuristic grouping andstructural analysis of text strings to distinguish connectedcomponents representing text characters fromthose representing noises. Assuming that a text stringhas at least three characters in alignment, we developtwo methods to locate regions containing text strings: adjacent character grouping and text line grouping, respectively. In both algorithms, a connected componentC is described by four metrics: height, width , area , andcentroid. In addition, we use D to represent the distancebetween the centroids of two neighboring characters.

6. Adjacent Character Grouping

Text strings in natural scene images usually appear inalignment, each text character in a text string must possesscharacter siblings at adjacent positions. The structurefeatures among sibling characters can be used todetermine whether the connected components belong totext characters or unexpected noises. Here, five constraintsare defined to decide whether two connectedcomponents are siblings of each other. 1) Considering the capital and lowercase characters, theheight ratio falls between 1/T1 and T1. 2) The distance between two connected componentsshould not be greater than T2 times the width of thewider one. 3) For text strings aligned approximately horizontally, the difference between Y-coordinates of the connectedcomponent centroids should not be greater than T3 timesthe height of the higher one. 4) Two adjacent characters usually appear in the samefont size, thus their area ratio should be greater than1/T4 and less than T4. 5) If the connected components are obtained from gradientbased partition, the color difference between themshould be lower than a predefined threshold Ts. In theirsystem Chucai Yi and YingLi Tian set T1= T4= 2, T2= 3, T3= 0. 5, T5= 406. 1. Mathematical ModelFor two connected components C and C' they can begrouped together as sibling components if the above fiveconstraints are satisfied. When C and C' are groupedtogether, their sibling sets will be updated according totheir relative locations. That is, when C is located onthe left of C', C'will be added into the right-sibling setof C, which is simultaneously added into the left-siblingset of C' The reverse operation will be applied when C islocated on the right of C'. To create sibling groups correspondingto complete text strings, Chucai Yi and YingLiTian merge together any two sibling groups SG(C1) andSG(C2) when their intersection contains no less than twoconnected components. At this point, each sibling groupcan be considered as a fragment of a text string. Repeatthe merge process until no sibling groups can bemerged together. Text string in scene images can be describedby corresponding adjacent character groups. Toextract a region containing a text string, Chucai Yi andYingLi Tian calculate rectangle covering all of the connectedcomponents in the corresponding adjacent charactergroup.

Advantages:-

1) The structure features among sibling characters canbe used to determine whether the connected componentsbelong to text characters or unexpected noises. 2) Character grouping is performed to combine the candidatetext characters into text strings which contain atleast three character members in alignment. 4Figure 5: Two detected adjacent character groupsmarked in red and green [3]

7. Text Line Grouping

In order to locate text strings with arbitrary orientations, Chucai Yi and YingLi Tian develop text linegrouping method. To group together the connected componentswhich correspond to text characters in the samestring which is probably non horizontal, Chucai Yi andYingLi Tian use centroid as the descriptor of each connectedcomponent. In order to locate text strings with arbitrary orientations, Chucai Yi and YingLi Tian develop text linegrouping method. To group together the connected componentswhich correspond to text characters in the samestring which is probably non horizontal, Chucai Yi andYingLi Tian use centroid as the descriptor of each connectedcomponent. 7. 1. Mathematical ModelGiven a set of connected component centroids, groups of collinear character centroids are computed, asshown below. M= {m | C 2 S & m = centroid (C)}L = {G | G is subset of M, | G| _ 3, they are charactercentroids & are collinear}Where M denotes the set of centroids of all of the connectedcomponents obtained from image partition, andL denotes the set of text lines which are composed of textcharacter centroids in alignment. Chucai Yi and YingLi Tian design an efficient algorithmto extract regions containing text strings. At first, they remove the centroids from the set M if areas oftheir corresponding connected components are smallerthan the predefined threshold Ts. Then, three pointsmi, mj , mk are randomly selected from the set to formtwo line segments. They calculate the length difference, and incline angle difference between line segments mimjand mjmk as shown_d= D (mi , mj) / D (mj , mk)__=| _ij-_jk | , if | _ij-_jk | _ _/2=| _ij -_ -_jk | , if | _ij-_jk | > _/2The three centroids are approximately collinear if1/T6__d_T6 and ___T7Here T6= 2, T7= _/12 Thus, they compose a preliminaryfitted line lu= mi, mj , mk , u = index of fitted line. Othercollinear centroids along lu can be added into the end positionsto form a complete text string increasingly. Fornow, each text string is described by a fitted line. Thelocation and size of the region containing a text stringis defined by the connected components whose centroidsare cascaded in the corresponding fitted line.

Advantages:-

1) In order to locate text strings with arbitrary orientations, Chucai Yi and YingLi Tian develop text linegrouping method. 2) To group together the connected components whichcorrespond to text characters in the same string whichis probably non horizontal, they use centroid as the descriptorof each connected component.

8. Image binarization and connected component analysis based method

Proposed methodology is presented in Figure 7. Starting from the scene image, Basilios Gatos produces5Figure 6: Resulting fitted lines from centroids cascading. Red line corresponds to text region while cyan lines arefalse positives to be removed [3]gray level image and inverted gray level image. Then, calculate the two corresponding binary images using anadaptive binarization and image enhancement technique. In the sequel, the proposed technique involves a decisionfunction that indicates which image between binary imagescontains text information. In first fig. the originalbinary image is selected while in second fig. the invertedbinary image is selected. Finally, a procedure that detectsconnected components of text areas is applied. 8. 1. Data independence and Data Flow architectureData Flow architecture is represented in figure 8. 8. 2. Turing Machinestate transition diagram is represented in figure 9.

9. Conclusion

Due to the unpredictable text appearances and complexbackgrounds, text detection in natural scene imagesis still an unsolved problem. In this paper we have presentedmethods to find strings of characters from naturalscene images. We have presented techniques likeextraction of character string regions from scenery imagesbased on siblings and thickness of characters, efficientbinarization and enhancement technique followedby a suitable connected component analysis procedure, text string detection from natural scenes by structurebasedpartition and grouping, and a robust algorithmFigure 7: a) Flowchart of the proposed method for textdetection in natural scene images (original binary imageis selected). b)Flowchart of the proposed method for textdetection in natural scene images (inverted binary imageis selected).[2]Figure 8: Data Flow architecturefor text detection in images. Also, we have presented anapproach to detect, localize, and extract texts appearingin grayscale or color Images as well as locate text stringswith arbitrary orientations. 6Figure 9: state transition diagramOur future work will focus on developing learningbased methods for text extraction from complex backgroundsand text normalization for OCR recognition.