Mobile Camera Based Text Detection and Translation

Ravindra Bandal, Adesh Jadhav, Vitthal Kale.

B.E computer engineering,Navsahyadri Education Society'sGroup of Institution's , pune.

Abstract -

Text in a natural image directly carry rich high-level semantic information about a scene, which can be used to assist a wide variety of applications, such as image understanding, image indexing and search,geolocation or navigation, and human computer interaction. However, most existing text detection and recognition systems are designed for horizontal or near-horizontal texts. Witt the increasingly popular computing-on-the-go devices, detecting texts of arbitrary orientations from images taken by such devices under less controlled conditions has become an increasingly important and yet challenging task.

In this project, we are using a new algorithm to detect texts of arbitrary orientations in natural images. Our algorithm is based on a two-level classification scheme and utilize two sets of features specially designed for capturing both intrinsic and orientation-invariant characteristics of texts. To better evaluate the proposed method and compare it with other existing algorithms, we generate a more extensive and challenging dataset, which includes various types of texts in diverse real-world scenes.

Keywords:-

Text detection , fuzzy clustering ,edge profile , signboard image , OCR.

1. Introduction

Character which can be used to assist a wide variety of applications, such as image understanding, image indexing and search, geolocation or navigation, and human computer interaction. However, most existing text detection and recognition systems are designed for horizontal or near-horizontal texts. With the increasingly popular computing-on-the-go devices, detecting texts of arbitrary orientations from images taken by such devices under less controlled conditions has become an increasingly important and yet challenging task. In this project, we are using a new algorithm to detect texts of arbitrary orientations in natural images. Our algorithm is based on a two-level classification scheme and utilize two sets of features specially designed for capturing both intrinsic and orientation-invariant characteristics of texts. To better evaluate the proposed method and compare it with other existing algorithms, we generate a more extensive and challenging dataset, which includes various types of texts in diverse real-world scenes.We are also using a new evaluation protocol, which is more suitable for benchmarking algorithms designed for texts of varying orientations. Experiments on conventional benchmarks and the new dataset demonstrate that our system compares favorably with the state-of-the-art algorithms when handling horizontal texts and achieves significantly enhanced performance on texts of arbitrary orientations in complex natural scenes..

2. Existing system

In case of existing system it is difficult to make changes into the existing document but with the help of this application it's easy to make changes into the file.

3. Problem statement

The mobile camera based text detection and translation is apply all the words in the English language, or a more technical for specific field.This technique can be problematic if the document contains words in cursive, like proper nouns. Tesseract uses its dictionary to influence the character segmentation step, for improved accuracy. The output stream may be a plain text stream or file of characters, but more sophisticated OCR systems can preserve the

original layout of the page and produce, for example, an annotated PDF that includes both the original image of the pageand a searchable textual representation..

4. Proposed system

In this system we are providing the effective way to make more editable document . This technique can be problematic if the document contains words in cursive, like proper nouns. Tesseract uses its dictionary to influence the character segmentation step, for improved accuracy. The output stream may be a plain text stream or file of characters, but more sophisticated OCR systems can preserve the original layout of the page and produce.

5.System Requirement

A) Hardware Interfaces:

Processor   :   At   Least   Pentium Processor

Ram   :   512 MB

Hard Disk   :   2 GB

B) Software Interfaces:

Operating System   :   Windows XP/7/8.

Technology   : ADT, Java (JDK 1.6).

Front End   : Android

Database   : MySQL.

Java (JDK 1.6)

Java is a general purpose programming language with a number of features that make the language well suited for use on the World Wide Web. Small Java applications are called Java applets and can be downloaded from a Web server and run on your computer by a Java-compatible Web browser, such as Netscape Navigator or Microsoft Internet Explorer.

6. System Architecture:

The proposed system consists of four stages:

(1) component extraction

(2) component analysis

(3)candidate linking

(4) chain analysis

which can be further categorized into two procedures, bottom-up grouping and top-down pruning, as shown in Fig. In the bottom-up grouping procedure, pixels are first grouped into connected components

and later these connected components are aggregated to form chains; in the top-down pruning procedure non-text components and chains are successively identified and eliminated. These two procedures are applied alternately when detecting text in images.

6.1 Component extraction:

At this stage, edge detection is performed on the original image and the edge map is fed to the SWT module to produce an SWT image. Neighboring pixels in the SWT image are grouped together recursively to form connected components using a simple association rule.

6.2 Component analysis:

Many components extracted at the component extraction stage are not parts of texts. The component analysis stage is aimed to identify and filter out those non-text components. First,the components are filtered using a set of heuristic rules that can distinguish between obvious spurious text regions and true text regions . Next, a component

level classifier is applied to prune the non-text components that are hard for the simple filter.

6.3 Candidate linking:

The remaining components are taken as character candidates. The first step of the candidate linking stage is to link the character candidates into pairs. Two adjacent candidates are grouped into a pair if they have similar geometric properties and colors. At the next step, the candidate pairs are aggregated into chains in a recursive manner.

6.4 Chain analysis:

At the chain analysis stage, the chains determined at the former stage are

verified by a chain level classifier. The chains with low classification scores (probabilities) are discarded. The chains may be in any direction, so a candidate might belong to multiple chains; the interpretation step is aimed to dispel this ambiguity. The chains that pass this stage are the final detected texts. The remainder of this paper is organized as follows. Section II presents the details of the proposed method, including the algorithm pipeline and the two sets of features. Section III introduces the proposed dataset and evaluation protocol. The experimental results and discussions are given in Section IV. Section V concludes the paper and points out potential directions for future research.
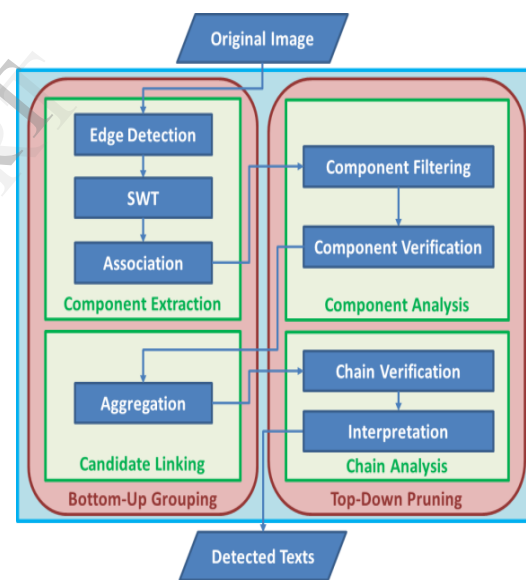
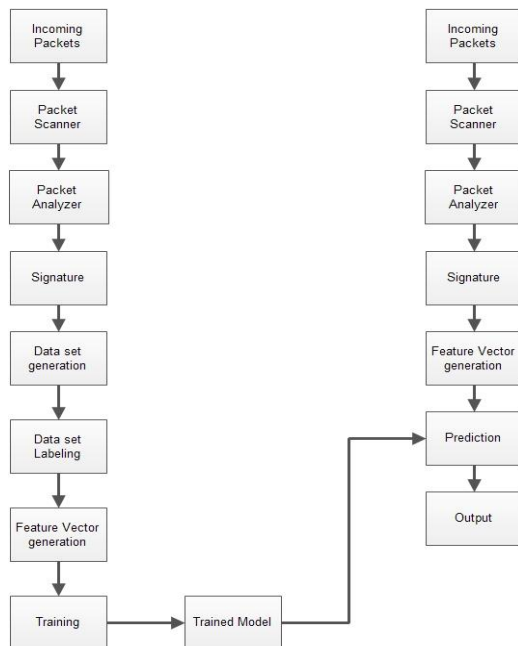Figure6.5 : System pipelined architecture

Workflow of System

Figure: 6.6 Workflow of System

7. Algorithms for converting color to grayscale

How do you convert a color image to grayscale? If each color pixel is described by a triple (R, G, B) of intensities for red, green, and blue, how do you map that to a single number giving a grayscale value? There are following three algorithms.

The **lightness** method averages the most prominent and least prominent colors: (max(R, G, B) + min(R, G, B)) / 2.

The **average** method simply averages the values: (R + G + B) / 3.

The **luminosity** method is a more sophisticated version of the average method. It also averages the values, but it forms a weighted average to account for human perception. We're more sensitive to green than other colors, so green is

weighted most heavily. The formula for luminosity is 0.21 R + 0.71 G + 0.07 B.

The example sunflower images below

Original image

Lightness

Average

Luminosity

The lightness method tends to reduce contrast. The luminosity method works best overall. However, some images look better using one of the other algorithms. And sometimes the three methods produce very similar results.

8.Thresholding

Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images.During the thresholding process, individual pixels in an image are marked as "object" pixels if their value is greater than some threshold value (assuming an object to be brighter than

the background) and as "background" pixels otherwise. This convention is known as threshold above. Variants include threshold below, which is opposite of threshold above; threshold inside, where a pixel is labeled "object" if its value is between two thresholds; and threshold outside, which is the opposite of threshold inside. Typically, an object pixel is given a value of "1" while a background pixel is given a value of "0." Finally, a binary image is created by coloring each pixel white or black, depending on a pixel's labels.

8.1 Threshold selection

The key parameter in the thresholding process is the choice of the threshold value (or values, as mentioned earlier). Several different methods for choosing a threshold exist; users can manually choose a threshold value, or a thresholding algorithm can compute a value automatically, which is known as automatic thresholding . A simple method would be to choose the mean or median value, the rationale being that if the object pixels are brighter than the background, they should also be brighter than the average. In a noiseless image with uniform background and object values, the mean or median will work well as the threshold, however, this will generally not be the case. A more sophisticated approach might be to create a histogram of the image pixel intensities and use the valley point as the threshold. The histogram approach assumes that there is some average value for the background and object pixels, but that the actual pixel values have some variation around these average values. However, this may be computationally expensive, and image histograms may not have clearly defined valley points, often making the selection of an accurate threshold difficult. One method that is

relatively simple, does not require much specific knowledge of the image, and is robust against image noise

9. Template matching

Template matching is a technique in digital image processing for finding small parts of an image which match a template image. It can be used in manufacturing as a part of quality control a way to navigate a mobile robot, or as a way to detect edges in images.

9.1 Approach

Template matching can be subdivided between two approaches: feature-based and template-based matching. The feature-based approach uses the features of the search and template image, such as edges or corners, as the primary match-measuring metrics to find the best matching location of the template in the source image. The template-based, or global, approach, uses the entire template, with generally a sum-comparing metric (using SAD, SSD, cross-correlation, etc.) that determines the best location by testing all or a sample of the viable test locations within the search image that the template image may match up to.

9.2 Feature-based approach

If the template image has strong features, a feature-based approach may be considered; the approach may prove further useful if the match in the search image might be transformed in some fashion. Since this approach does not consider the entirety of the template image, it can be more computationally efficient when working with source images of larger resolution, as the alternative approach, template-based,

may require searching potentially large amounts of points in order to determine the best matching location.

9.3 Template-based approach

For templates without strong features, or for when the bulk of the template image constitutes the matching image, a template-based approach may be effective. As aforementioned, since template-based template matching may potentially require sampling of a large number of points, it is possible to reduce the number of sampling points by reducing the resolution of the search and template images by the same factor and performing the operation on the resultant downsized images (multiresolution, or pyramid, image processing), providing a search window of data points within the search image so that the template does not have to search every viable data point, or a combination of both.

10.Advantages

• Easier and less expensive using a mobile device equipped with real time machine translation.

• Real time mobile translation can help people at travelling.

• The portability of mobile phones makes it convenient.

11. Disadvantages

• No cursive character detection.

• Not suitable for low image quality.

12. Application

• More quickly make textual versions of printed documents.

• Converting handwriting in real time to control a computer

13. Conclusion

We have presented a text detection system that detects texts of arbitrary directions in complex natural scenes. Our system compares favorably with the state-of-the-art algorithms when handling horizontal texts and achieves significantly enhanced performance on texts of arbitrary orientations. Furthermore we have presented an approach for automatic detection and binarization of texts for application to mobile systems. The proposed method can robustly detect and binarize main texts from signboard images. Firstly, we perform detection of main text region using edge histogram method with horizontal and vertical direction in edge map image. After text region verification, detected region is segmented by fuzzy c-mean clustering and each region is distinguished as text region.

14. Reference:

[1].IEEE TRANSACTION ON Mobile based text detection and translation. Jonghyun park, Toan Nguyen and Gueesang lee. IEEE TRANSACTION Nov.2012.

[2]. IEEE TRANSACTION ON Detecting Texts of Arbitrary Orientations in Natural Images.   Cong yao,XinZang,Xiang Bai,Member IEEE ,Wenyu Liu . IEEE TRANSACTION May 2012.

[3]. a-Soft: An English Language OCR Junaid Tariq. Department of Computer science. COMSATS Institute of Information Technology, Islamabad,Pakistan junaid tariq@ comsats.edu.pk Umar Nauman Department of Computer science. COMSATS Institute of Information Technology, Islamabad, Pakistan.

[4]. Graph Matching Method for Character Recognition in Natural Scene Images Jieun Kim*, Ho-sub Yoon** University of Science and Technology / Department of Computer Software and engineering, Daejeon, Republic of Korea.INES 2011 • 15th International Conference on Intelli gent Engineering Systems • June 23–25, 2011, Poprad, Slovakia.

[5]. Application of clustering Technique for Tissue image segmentation and comparison mohit Agarwal, Gaurav Dubey, Ajay Rana 22 International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013 ISSN 2229-5518.

[6]. Digital Image Encryption Algorithm Based on Chaotic Block and Pixel Mapping Table . Nadia Al-Rousan, Hazem Al-Najjar . International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013 ISSN 2229-5518.