International Journal of Advanced Technology in Engineering and Science
Vol. No.5, Issue No. 04, April 2017
www.ijates.com
ijates
ISSN 2348 - 7550

AUTOMATED TEXT DETECTION AND RECOGNITION FROM COMPLEX IMAGES

Devendra Haldankar[1], Pratibha Gupta[2], Priyanka Gawade[3], Sandeep Dhobi[4], Priyanka Bhilare[5]

[1,2,3,4]Student, [5]Professor, Computer, Rajiv Gandhi Institute of Technology, Maharashtra (India)

ABSTRACT

The proposed system detects and extracts the text from different complex images such as scene images and document images. As text detection from images and videos is very essential step to achieve multimedia contents for different application purpose. The proposed system focuses on discrete wavelet transform for localization, so it requires less processing time which is necessary for real time applications. This method will provide better efficiency and the best part of this proposed system is that the extracted text is in the .txt format.

Keywords: DWT(Discrete wavelet transform), Text detection Text extraction, Text recognition,.

I. INTRODUCTION

Text, as one of the most influential inventions of humanity, has played an important role in human life, so far from ancient times. The rapid advancement in the technology and multimedia has digitalized the world. The availability of cameras and other systems contributes large number of images to the world. Ranging from cameras embedded in mobile phones to professional ones, Surveillance cameras to broadcast videos, every day images to satellite images, all these contributes to increase in multi-media data. Most of the images may contain text as part of it, which gives some information about that image. Therefore identification of these texts has relevance in many applications. This shows the importance of the text extraction system in lot of applications. It was stated that in recent years there was a drastic in-crease in multimedia libraries and the amount of data is growing exponentially with time. Generally, the images can be categorized into three based on its type: document images, scene images and digital images.

Document images are the image in the form of the document which includes pdf, doc etc.Scene images are the images that contain the text, such as the advertising boards, banners, which is captured naturally.Images generated by computer softwareare known as digital or born-digital images. The text in these images is called as caption text or overlay text.Compared with document images and born-digital images, the scene images, have more complex foreground/ background, low resolution, compression loss, and severe edge softness. This makes the detection of text from scene text more difficult. There-fore automatic extraction of texts from images or video is a challenging task and research under this field is still under progress.

1.1 Challenges in Scene Text Detection

The major challenges in scene text detection and recognition can be roughly categorized into three types:

• **Diversity of scene text**: In contrast to characters in document images, which are usually with regular font, single colour, consistent size and uniform arrangement, texts in natural scenes may bear entirely different fonts, colours, scales and orientations. These many diversities can add complexity in scene images.

• **Background complexity**: The backgrounds in natural scene images and videos can be very complex. Elements like signs, fences and bricks are virtually undistinguishable from true text, and thus are easily to cause confusions and errors in a system.

• **Interference factors**: Interference factors, for example, noise, blur, distortion, low resolution, non-uniformillumination etc. may give rise to failures in scene text detection and recognition.

1.2 Recent Advances in Scene Text Detection

In recent years, text detection and recognition in natural images have become active research topics in the communities of computer vision, pattern recognition and even document analysis. There are basically two types of methods: texture based methods and component based methods.

Texture based method: Texture based method treat texts as a special type of texture and make use of their textural properties, such as local intensities, filter responses and wavelet coefficients, to distinguish between text and non-text areas in the images. These methods are computationally expensive as all locations and scales should be scanned. Also, these methods handle horizontal texts and they are sensitive to rotation change and scale change.

Component based methods: Component based method first extract candidate components through a variety of ways i.e. color clustering or extreme region extraction and then filter out non-text components using manually designed rules or automatically trained classifiers. these methods are much more efficient, because the number of components to be processed is relatively small. But these methods aren't sensitive to rotation, scale change and font variations.

II. PROPOSED SYSTEM

The proposed Intelligent Text Extraction system automatically identifies and extracts the text present in different form of images i.e. natural or document images. Input is taken in the form of text image and then applying preprocessing methods on it whichremoves 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 preprocessing if some part text data will loss them by thinning and scaling is performed by connectivity algorithm. Then we get connected text character from image. The main advantage of the system is that the extracted text is shown in the .txt file. This method shows better efficiency, precision and recall compared to the existing techniques. This shows the possibility of using this technique in more new and advanced applications.

2.1 System Implementation and Working

As we know that Natural scene images usually suffer from low resolution and low quality, perspective distortion and complex background. To overcome such type of problems we introduced a method which is character descriptor. Following fig shows the flow chart of the proposed system.

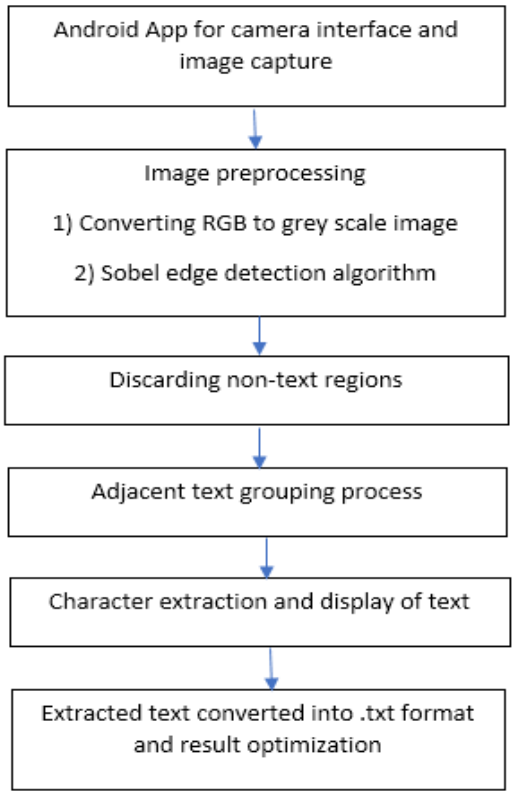

Fig-1: Flowchart of the System

a)Android app for camera interface and image capture

The first module of the system is to develop an android application for camera interfacing and image capture. The android app will capture the input scene image, on which text extraction will be performed.

b) Image preprocessing

The second module of the system is image preprocessing. The preprocessing will be done in order to get noise free and filtered image. The preprocessing will be done in two steps.

1. Converting RGB image into Grey scale image

As we know that, the image captured by camera is generally is color image i.e RGB image. The RGB color model is made up of three colors i.e red, green and blue. The model name comes from the initials of the three additive primary colors:red, green, and blue.The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Each color have8 pixels. The RGB image can be converted into grey scale image with the help of following formula.

;Grey scale image= 0.3r + 0.59g + 0.11b

2. Applying Sobel Edge Detection algorithm

In the second step of image preprocessing, the sobel edge detection have to apply on grey scale image. It is very efficient than the canny edge detection algorithm since it does not consume time as like edge detection algorithm.

Sobel has two main advantages compared to other edge operator: Since the introduction of the average factor, it has some smoothing effect to the random noise of the image.

-1	-2	-1
0	0	0
1	2	1

-1	0	-1
-2	0	2
-1	0	1

Fig-2: Sobel edge operator

The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation (call these Gx and Gy). These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient. The gradient magnitude is given by

;$|x| = \sqrt{Gx^2 + Gy^2}$

3. Adjacent text grouping process

The image got from the previous stage is considered as input in this phase. Text recognition is done by using the tesseract OCR. The text information which appears in text strings is made up of several character members in similar sizes rather than single character, and text strings are normally in approximately horizontal alignment.Adjacent text grouping method as shown in fig-2.

Fig-2:Adjacent text grouping method

c) Text Extraction

To detect and extracting the text from camera captured natural scene images. Text finding is used to get text containing image area then text identification is used to transform image-based information into readable text. This step refers to classify the characters as they are in original image. This is performed by multiplying resultant figure with binary converted novel image. Final result is the plane text dependent on the novel image.

III. CONCLUSIONS

Most of methods fail when the characters are not aligned well or when the characters are too small. Those methods also result in some of missing characters when the characters have very poor contrast with respect to the background. But the pro-posed method is not sensitive to color or intensity of image, and also the uneven illumination and reflection effects.The main advantage of the system is that the extracted text is shown in the .txt file.

This can be used in large variety of application fields such as vehicle license plate detection, object identification, identification of various parts in industrial automation, mobile robot navigation which helps to detect text based land marks.

IV. ACKNOWLEDGEMENT

We are thankful to our project Guide Miss. PriyankaBhilare and all staff members of computer department who have provided us various facilities and have guided us whenever required.

REFERENCES

[1] A. K. Jain, B. Yu, "Automatic text location in images and video frames", *Pattern Recognit., vol. 31, no. 12, pp. 2055-2076, 1998*.

[2] Xiaoqing Liu and Jagath Samarabandu, Multiscale edge-based Text extraction from Complex images, *IEEE, 2006*.

[3] Keechul Jung, Kwang In Kim and Anil K. Jain, Text information extraction in images and video: a survey, *The journal of the Pattern Recognition society, 2004*.

[4] S.Audithan, RM. Chandrasekaran, Document Text Extraction from Document Images Using HaarDiscrete Wavelet Transform,*European Journal of Scientific Research, 2009*