International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Vision based Traffic Panel Text Information and Sign Retrieval

Mounika Mantha†* and J Krishna Chaithanya†

†Department of ECE, Vardhaman Engineering College, Hyderabad, Telangana State, India

Accepted 15 July 2015, Available online 16 July 2015, **Vol.5, No.4 (Aug 2015)**

Abstract

We propose a novel system for the automatic detection and recognition of text in traffic sign boards. Maximally stable extremal regions (MSERs) and hue, saturation, and value color thresholding are used to locate a large number of candidates, which are then reduced by applying constraints based on temporal and structural information. A recognition stage interprets the text contained within detected candidate regions. Individual text characters are detected as MSERs and are grouped into lines, before being interpreted using optical character recognition (OCR).

Keywords: *Maximally stable extremal regions (MSERs), hue, saturation, value color thresholding, Optical Character Recognition (OCR).*

1. Introduction

This paper presents the implementation of detecting traffic panels text in street-level images and to recognize the information, as an real application to intelligent transportation systems (ITS). Intelligent transportation systems (ITS) are advanced applications which, without embodying intelligence as such, aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated, and 'smarter' use of transport networks. Information and communication technologies are applied in the field of road transport, including infrastructure, vehicles and users, and in traffic management and mobility management, as well as for interfaces with other modes of transport. Traffic Panel text and sign recognition is important systems, automatic vehicles, and inventory purposes. Fig 1shows the traffic panel images and traffic signs are to detect the text in road panels. In order to compute the distance from the vehicle to the panels, a function that the text height in pixels has been obtained. After computing the text vector of the vehicle, word coordinates are computed for each panel.

This paper is aimed at to obtain a traffic signs on a given road and gather information about their state. However, the system focused on the position of the possible text and signs and not on their recognition. In this paper, to apply some modifications and new functionalities are used to read the information contained in traffic panel street images. This work mainly deals on the task of automatically detecting text on road signs from traffic panel street image and using that information in a driver assistance system.

Figure 1: Traffic Panel Image

2. System Description

The system is developed using an ARM Cortex A7 processor based Raspberry pi board, the Raspberry Pi is based on the Broadcom BCM2836 system on a chip (SoC), which includes an ARM1176JZFS700 MHz processor, Video Core IV GPU, and was originally shipped with 512 megabytes of RAM, later upgraded (models B and B+) to 1 GB. The system has Secure Digital (SD) (models A and B) or Micro SD (models A+ and B+) sockets for boot media and persistent storage. Raspbian is a free operating system based on Debian optimized for the Raspberry Pi hardware. An operating system is the set of basic programs and utilities that make your Raspberry Pi run. However, Raspbian provides more than a pure OS: it comes with over 35,000 packages, pre-compiled software bundled in a nice format for easy installation on your Raspberry Pi. We have developed 3 types of text recognition i.e.

*Corresponding author: **Mounika Mantha; J Krishna Chaithanya** is working as Associate Professor

- Text identification over the Traffic Panel
- Sign Detection
- Optical Character Recognition over live image with TTS

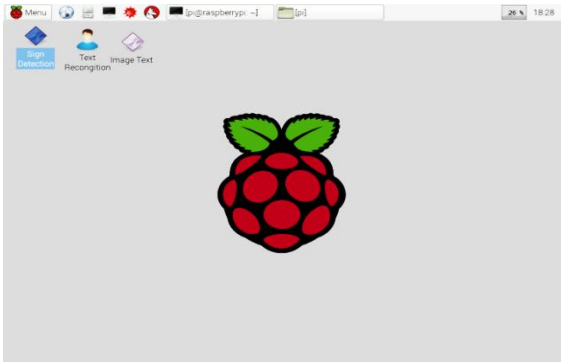

Figure 2: Application Icons on Raspberry pi

In the fig 2, the developed programs can be viewed over the raspberry pi board Operating system.

3. Panel Detection and Text Identification

In this module we identify the panel based on the prediction of the SVM. Which cluster is predicted as a panel that index is used to detecting the area in the traffic panel image. The detected partition portion is merged together to predict the panel in the image. Finally the identified panel is displayed in the images. The thresholding operation identifies the image pixels that are having the pixels similar to the input threshold value.

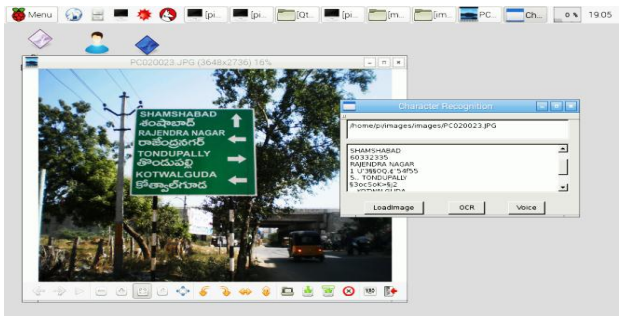

Figure 3: Text Recognition over Traffic Panel

Binary images may contain numerous imperfections. In particular, the binary regions produced by simple thresholding are distorted by noise and texture. Morphological image processing pursues the goals of removing these imperfections by accounting for the form and structure of the image. Morphological image processing is a collection of nonlinear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. The morphological operation eliminates the unwanted pixels and identifies the exact text portions.

Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. The morphological operation eliminates the unwanted pixels and identifies the exact text portions.

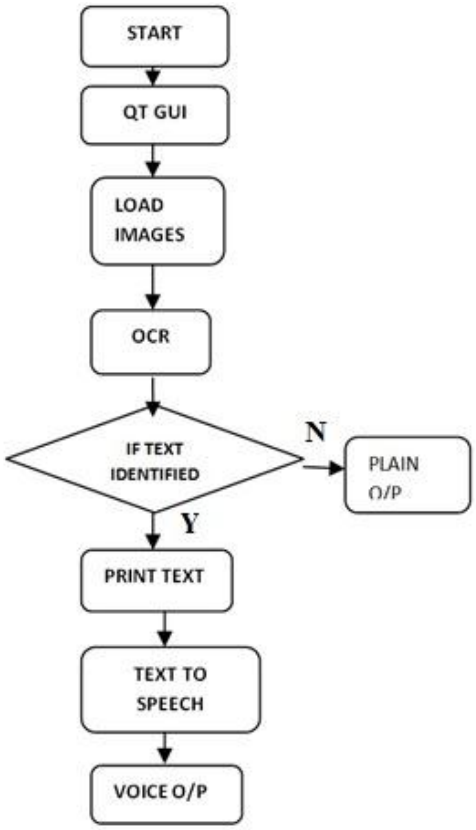

Figure 4: Flow of Text Recognition

The above figure 4 resembles the flow chart of the steps for recognition of the steps on the traffic panels.

4. Shape-based Sign Detection

Several approaches for shape-based detection of traffic signs are recurrent in literature. Probably the most common approach is using some form of Hough transform. Approaches based on corner detection followed by reasoning or approaches based on simple template matching are also popular. Generalized Hough transform is a technique for finding arbitrary shapes in an image. The basic idea is that, using an edge image, each pixel of the edge image votes for where the object center would be if that pixel were at the object boundary. The technique originated early in the history of computer vision. It was extended and modified numerous times and there are many variants. Here we will present work by Loy and Barnes, as it was intended specifically for traffic sign detection and was used independently in several detection systems. Loy and Barnes propose a general regular polygon detector and use it to detect traffic signs. The detector is based on their fast radial symmetry transform, and the

overall approach is similar to Hough transform. First the gradient magnitude image is built from the original image. The gradient magnitude image is then threshold so that the points with low magnitudes, which are unlikely to correspond to edges, are eliminated. Each remaining pixel then votes for the possible positions of the center of a regular polygon. One pixel casts its vote at multiple locations distributed along a line which is perpendicular to the gradient of the pixel and whose distance to the pixel is equal to the expected radius of the regular polygon (see figure 5).

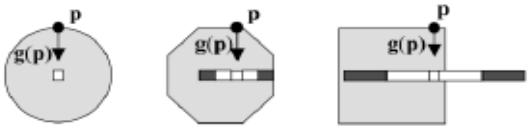

Figure 5: Different Polygon Shapes

Notice that there are actually two lines which satisfy these requirements, one in the direction of the gradient and the other in the opposite direction. Both can be used if we don't know in advance whether signs will be lighter or darker than the background.

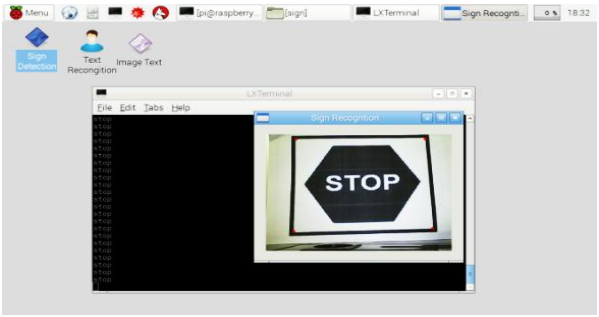

Figure 6: Symbol of STOP bound using polygon

The length of the voting the votes towards the end of the line have negative weights, to minimize the influence of straight lines in an image which are too long to be polygon edges. The resulting vote image is labeled Or. In addition to the vote image, another image called equiangular image is built.

The proposed procedure favors equiangular polygons by utilizing the following property: if the gradient angles of edge pixels of an n-sided regular polygon are multiplied by n, the resulting angles will be equal gradient angle of the pixel multiplied by the number of sides of the sought regular polygon. The pixel then again casts its vote on locations determined by the voting line, except that this each point in this image fig 6 represents a vector which is the sum of all contributing votes. The votes coming from edges of equiangular polygons will have the same slope, so the magnitudes of vote vectors in equiangular polygon centroids should be the largest. Finally, the vote image and the norm of the equiangular image are combined to produce the overall response.

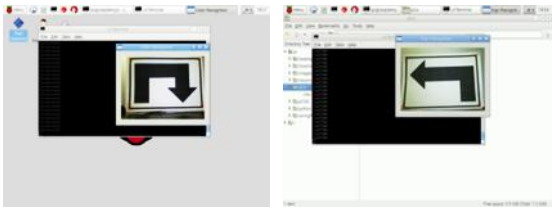

Figure 7: Results of different Signs Detection

Computational complexity of this method is O(Nkl), where l is the maximum length of the voting line, N is the number of pixels in an image and k is the number of radii being considered. The main weakness of the approach is that the radius of the sought polygon should be known in advance, which is not always easy to accomplish. The results are shown in figure 7.

5. OCR Over Live Image with TTS

Optical Character Recognition [T. M. Breuel, 2001] has been an active subject of research since a decade. The rapid growth of digital libraries worldwide poses new challenges for document image analysis research and development. Digital libraries offer access to larger document collection, and at a faster speed. A number of OCR software [M. Kramer, 2007] – [U.-V. Marti, 2002] available in market claims 99% recognition accuracy, but in practice these accuracy rates are rarely achieved. Most systems breakdown when input document images are highly degraded. OCR is a process which associates a symbolic meaning with objects (letters, symbols an number) with the image of a character. It is defined as the process o converting scanned images of machine printed into a computer process able format. A methodology is implemented to recognition sequence of characters and the line of reading. As part of the software development [S. Yanadume,2004] the Open CV (Open source Computer Vision) libraries is utilized to do image capture of English text, to do the character recognition. Optical character recognition (OCR) is the translation of captured images of printed English text into machine-encoded text. It is widely used to convert books and documents into electronic files for use in storage and document analysis. OCR makes it possible to apply techniques such as machine translation, text-to-speech and text mining to the capture / scanned page.

The Hardware system is composed by following parts: an image capturing camera [2006], Raspberry Pi board to run image recognition programs on it and a Headphone to deliver the output speech. The system block diagram is shown in Fig 8.

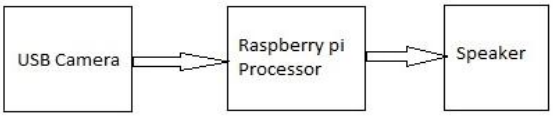

Figure 8: OCR System Hardware Diagram

A. Image Capturing

The first step in which the device is moved over the printed page and the USB camera captures the images of the text. The quality of the image captured will be high so as to have fast and clear recognition due to the high resolution camera.

B. Pre-processing

Preprocessing stage consists of three steps: Skew Correction, Binarization and Noise removal. The captured image is checked for skewing. There are possibilities of image getting skewed with either left or right orientation. Here the image is first brightened and binarized.

The function for skew detection checks for an angle of orientation between ±15 degrees and if detected then a simple image rotation is carried out till the lines match with the true horizontal axis, which produces a skew corrected image. The noise introduced during capturing or due to poor quality of the page has to be cleared before further processing.

C. Segmentation

After pre-processing, the noise free image is passed to the segmentation phase. It is an operation that seeks to decompose an image of sequence o characters into sub-image of individual symbol (characters). The binarized image is checked for inter line spaces. If inter line spaces are detected then the image is segmented into sets of paragraphs across the interline gap. The lines in the paragraphs are scanned for horizontal space intersection with respect to the background. Histogram of the image is used to detect the width of the horizontal lines. Then the lines are scanned vertically for vertical space intersection. Here histograms are used to detect the width of the words. Then the words are decomposed into characters using character width computation

D. Feature Extraction

Feature extraction [X.Wang,2010] [F. Einsele,2007] is the individual image glyph is considered and extracted for features.

First a character glyph is defined by the following attributes: (Álvaro González,2014) Height of the character; (X.Wang,2010) Width of the character; (F. Einsele,2007) Numbers of horizontal lines present—short and long; Numbers of vertical lines present—short and long; (V. I. Levenshtein,1966) Numbers of circles present;(S. Wachenfeld,2006) Numbers of horizontally oriented arcs; (T. M. Breuel,2001) Numbers of vertically oriented arcs; (S. Wachenfeld,2007) Centroid of the image; (F. Einsele,2007) Position of the various features; (ACM symposium, 2008) Pixels in the various regions.

E. Image to Text Converter

The ASCII values of the recognized English characters are processed by Raspberry Pi board.

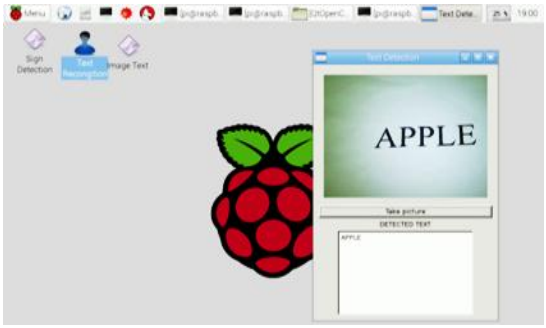

Figure 9: Image to Text conversion result

Here each of the characters is matched with its corresponding template and saved as normalized text transcription. This transcription is further delivered to audio output.

F. Text to Speech

The scope of this module is initiated with the conclusion of the receding module of Character Recognition. The module performs the task of conversion of the transformed English text to audible form. The Raspberry Pi has an on-board audio jack, the on-board audio is generated by a PWM output and is minimally filtered. Using Open source text to speech algorithm “FLITE”, the audio output of the recognized text is provided through the speakers.

Conclusion

In this work we evaluate the Text recognition over Traffic Panel using techniques and existing state-of-the-art OCR engines for recognition of live image characters and text-lines. Due to very low resolution and small size, OCR of screen rendered text requires specialized approaches. Segmentation of screen rendered text lines is also challenging due to touching of characters with each other in some instances. This touching appears as a side effect due to anti-aliasing in rendering process. Recognition results show that approach based, open vocabulary, segmentation free techniques performs quite well in recognition of low resolution text.

References

Álvaro González, Luis M. Bergasa and J. Javier Yebes (February 2014), Text Detection and Recognition on Traffic Panels From Street-Level Imagery using Visual Appearance IEEE Transactions On Intelligent Transportation Systems, Vol. 15, No. 1, pp. 228-238.

http://www.babylon.com/

X.Wang, L. Huang, and C. Liu (Sep.2010), A video ext location method based on background classification, Int. Jour. on

DocumentAnalysis and Recognition, vol. 13, no. 3, pp. 187–207.

http://en.wikipedia.org/wiki/Phishing.

V. I. Levenshtein (1966), Binary codes capable of correcting deletions,insertions and reversals,Soviet Physics Doklady,vol. 10, no. 8, pp. 707–710

S. Wachenfeld, H.-U. Klein, and X. Jiang (2006), Recognition of screen-rendered text, Proceedings of the 18th International Conference on Pattern Recognition, vol. 02, pp.1086–1089.

T. M. Breuel (2001), Segmentation of hand printed letter strings using a dynamic programming algorithm, International Conference on Document Analysis and Recognition, vol. 02.

S. Wachenfeld, H.-U. Klein, and X. Jiang (2007), Annotated datab ases for the recognition of screen-rendered text, Proceeding sof the Ninth International Conference on Document Analysis and Recognition, vol. 02, pp. 272–276.

F. Einsele, R. Ingold, and J. Hennebert (2007), A HMM-based approach to recognize ultra low resolution anti-aliased words, Proceedings of the 2nd international conference on Pattern recognition and machine intelligence, pp. 511–518.

A language-independent (2008), open-vocabulary system based on HMMs for recognition of ultralow resolution words, proceedings of the 2008 ACM symposium on Applied computing, pp. 429–433.

C. Jacobs, P. Y. Simard, P. Viola, and J. Rinker (2005), Text recognition of low-resolution document images, Proceedings of the Eighth International Conference on Document Analysis and Recognition, pp. 695–699.

S. Yanadume, Y. Mekada, I. Ide, and H. Murase (2004), Recognition of very low-resolution characters from motion images captured by a portable digital camera, PCM (1), pp.247–254.

M. Kramer (2007), Optical character recognition using Hidden Markov Models, Master's thesis, Technical University of Kaiserslauter.

F. Jelinek (1997), Statistical methods for speech recognition, Cambridge, MA, USA: MIT Press.

I. Bazzi, R. Schwartz, and J. Makhoul (June 1999), An omni font open vocabulary OCR system for English and Arabic, IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, pp. 495–504.

U.-V. Marti and H. Bunke (2002), Using a statistical language model to improve the performance of an HMM-based cursive hand writing recognition systems, River Edge, NJ, USA: World Scientific Publishing Co., Inc , pp. 65–90.

Ms. Mounika Mantha is working towards a Master of Technology in Embedded Systems from Electronics and Communication Engineering (ECE) in prestigious Vardhaman College of Engineering, Hyderabad. She has completed B.Tech from Scient institute of Technology.

Mr. J. Krishna Chaithanyais presently working as an Associate Professor of Electronics Comm. Engineering in prestigious Vardhaman College of EngineeringHyderabad, India. B.Tech (Electronics Instrumentation & Control Engineering)2006, from N.B.K.R Institute of Science & Technology, Affiliated to S V University, Tirupati. M.Tech (Electronics and Communication Engineering), 2008 from SNIST, Affiliated to JNTU, Hyderabad. Ph.D. in Electronics and Communication Engineering (Pursuing), S V University, Tirupati.