Smart Approach for Detection, Recognition and Conversion of Devanagari Text from Road Side Traffic Panels: A Review

Gayatri H. Khobaragade*
M. Tech. Student
Deptt. of Computer Science
G.H.R.A.E.T. College, Nagpur, India

Deepak Kapgate
Assistant Professor
Deptt. of Computer Science
G.H.R.A.E.T. College, Nagpur, India

Abstract

Paper presents an approach of detection, recognition which actually detects the presence or absence of road traffic panels on highway images and to extract the information contained on street traffic panels. Sometimes traffic panel recognition and detection is difficult, due to its number of variation in panel colors, types, shapes and the large variation in the information depicted in road side traffic panels. The proposed system detects street traffic panels using appropriate algorithms that find the exact panel among all the ordinary advertisement panels. Text recognition methods can be dividing into two types: region-based method, and texture-based method. The detection stage detects the presence of text from the given street image and separate all images and symbols. The main goal of this paper is to present the review of different methods presented for text detection, recognition, conversion as well as for extraction on images. Challenging part of street text detection systems as well as Devanagari text conversions is that, sometimes-street images varying with color, shape, size, font, alignment, orientation.

Keywords: Intelligent Transportation Systems, Edge Profile Base Detection, Transform Color Histogram (LCQ), Local Colour Quantization, Artificial Neuron Fizzy Interference System (ANFIS), Rule Based Machine Transliteration (RBMT).

***Author for correspondence** gargihk@gmail.com

1. Introduction

Street panels always located right /left side of the road or above the road to ensure protection of highway man made infrastructure and the safety of the local travelling public through the proper management of a standard Indian organization under Highways Protection and Road Traffic Act [21]. There are two kinds of road traffic panels, i.e., those in green background with white foreground and those in yellow background with black foreground. The main agenda is to detect panels extract text from those accurate panels and finally using some Devanagari lipi to foreign language algorithms convert it into international language (English language). Automatic

classification of road traffic panels can be very useful for inventory purpose and maintenance purposes & it will useful for driver assistance applications and autonomous vehicles. In the form of text, strings and iconic symbol road traffic panels provide essential information. Texts can be in any form, which contain more related data to the location and help us to understand the correct goal easily. The necessity of text detection in images are to understand the content of particular panel for that it directly carry high-level semantic information about a particular scene, which can be used to assist a variety of applications, like image understanding, navigation, and indexing, human computer interaction and image search, detection the text of input images may contains any kind of text, as well as the existence or non-existence of text therefore the text detection stage detects the text from images. Text detection and segmentation from natural images use the primary property of scene text like, high contrast over similar background, colors would be difficult to preserve in real application. The system scans whole image for the existence of text pixel and static texts with low contrast and non-consistent lighting could be confused because of similar colors [22]. Extraction of text from images is an important problem in many applications like document processing, image indexing. Text extraction in images includes five stages, among which text detection and text localization are closely related and more challenging stages, which attract the attention of most researchers. The goal of the above stages is to generate correct bounding boxes among all text objects in images and provide a unique identity to each one.

The offline text recognition process works on images, which are generated by optical scanner. All Indian scripts, are cursive in nature but sometimes hard to recognize by machines. The main features of these vowels is that it can be written as independent character, & diacritical marks method can be written in any direction i.e. below, above, after or before the consonant to which they belongs [4]. The characters made by vowels are called conjuncts & writing the vowels in above pattern called as modifiers. Sometimes two or more consonants can be combined and produce the new shapes. These kinds of new shape clusters can be called as compound characters. Devanagari is used for writing many official languages such as Marathi, Hindi, Sindhi, Sanskrit, Konkani and Nepali, near about 500 million people use Devanagari script for documentation in northern and central parts of India. Hence, the detection of Devanagari lipi conversion algorithms is difficult to implement but in India, there are several regions where some instruction or panels are in Devanagari text so that many Indians and specially foreign peoples can't understand those road side panels may be include some warning, traffic or construction panels that responsible for accidents, for breaking rule and regulation in India. To overcome such problem we can use Devanagari detection as well as conversion algorithms is must.

The paper is organized as: latest techniques for panel detection reviewed in section 2. Section 3 describes performance evaluation as well as some techniques of text detection. Section 4 includes a few methods of text recognition. Section 5 includes effective techniques of text extraction. Section 6 organized some Devanagari conversion methods & finally section 7 presents summary and conclusions.

2. Panel Detection Techniques
A. *Visualise (VISUAL Inspection of Signs and Panels)*

Visualize technique is destined to complement a patented automatic visual analysis system of traffic signs and panels beside the right or left side of road. It is an automatic inspection system, which inbuilt in the vehicle, which accomplishes vertical/horizontal sign posting analysis at conventional driving speeds. It also automatically calculates the retro reflection values at particular distance of road traffic panels Visualize allows for a something better and advance changes in the alertness of road signposting, decision making and state supporting planning on the management and infrastructure operators' side [3, 14]. This particular system depends on light retro reflection principle. It uses an active infrared illuminator with perfectly known features as PLS (pattern light source). Small part of the IR light that meets the traffic panels and signs, symbols is reflected. Stereoscopic system captures the reflected light, this may be made up from high-resolution cameras [23]. Visualize technique is based on SIFT (Scale-Invariant Feature Transform) descriptors to recognize single characters and symbols, and it is based on Hidden Markov Models (HMMs) to recognize single words [21].

B. Shape Recognition Algorithm

Shape Recognition Algorithm proposed by Memane et al. [16]. It is a simple, yet effective method of analyzing the shapes and colors of objects. The algorithm was use to recognize the shapes of all cells and the Extent parameter of normal healthy cells was noted. As cancer cells differ from healthy cells in terms of shape as well as color, their Extent values as well as RGB content do not match those of healthy cells and can therefore be easily identified and quantified. Shape detection & recognition algorithm consists of following steps: Read/capture image, Converting RGB image to Black and White, Recognize boundaries of objects, Finding areas of objects and area filtering, Finding inclination of object, Finding bounding box of the object, Finding ratio of areas for given object, color recognition [14].

C. TCH

It is very efficient panel detection method, which gives 95% higher rate for all the laterals (Green & Yellow) & upper panels and gives the dimensionalities of TCH, are 45. Hue Histogram is computed from the hue and saturation color models, while TCH is computed from the Red, Green and Blue color components after normalizing each channel independently. TCH takes the training time lower than other descriptors like SIFT & Hue Histogram [11].

3. Text Detection Techniques

Roadside traffic panels are used to read and save the information that was depicted in the panels automatically. Detection is used for checking the presence of text within the particular object or image. Various text detection techniques are as follows:

A. Hypotheses Generation Approach

Hypotheses generation full text detection approach was introduced by Fabrizio & Marcotegui [13]. This approach focused on combination of hypothesis generation and hypothesis validation step for the generation of potential boxes & filter false detections. The hypothesis generation approach depends on an efficient segmentation method, which is based on a morphological operator. Using Fourier, Zernike moments, Pseudo and Original polar descriptor regions are filtered & classified with the help of shape descriptors. Classification process based on three SVM classifiers, which are combined in a late fusion scheme. Finally, detected characters are grouped for the generation of text box hypotheses. Using dedicated descriptors by the HOG

approach validation step processed through Global SVM classification. Hypotheses approach detects text from roadside traffic panel's and from those images, which are varying in shape, size & color. Therefore, a text box localization system was developed with few possible hypotheses on text for focusing urban context on street side images [10, 17].

B. Local Color Quantization (LQL)

In LCQ, every color portion assumed as a text without confirmation either it is real or not [22]. To reduce blight, color quantization takes place before processing; and static input image will be converted into a 256-RBG image [22]. Features of text field and text lines will be used to find connected components, candidents that are extracted for each merged colors. Then for each color LCQ being executed [12]. Local Color Quantization gives better result from HUE.

C. Edge Profile based Text Detection

It is introduced by Peer & Ikica [4] which is simple, fast and efficient because of its edge-orientated behaviour which lacks of being enough text-aware. Hence, non-text regions are often extracted. To avoid this limitation, Peer & Ikica proposes some heuristic rules to reduce false positives detection. First horizontal edge profile is computed, and then HP presents a useful clue where the text might appear in the image. Hence particular static text will be assume as horizontally [5] aligned and vertical characteristics expose by characters, after that on vertical edge-map horizontal profile is computed. However, cut out the parts of the characters can be extending over/below the vertical boundary. There are various heuristic rules proposed by Peer & Ikica [4]. Connected components do not overlap with every eliminated Vertical Text Boundaries (VTB). Therefore, the text characters always appear close and sometimes due to same height, completely overlapping regions that extend so far below or above a particular [5] VTB are eliminated. Similarly too small overlapping regions as compared to the VTB height are eliminated. At last, particular heuristic rules eliminate non-text candidate regions, such as closeness rule and the aspect ratio. Hence close candidate text regions are merged into a single area [2, 4].

4. Text Recognition Techniques

Text recognition is a technique, which automatically recognizes letters and words and symbols from an image of printed text, natural scenes, static images and traffic panels. Recognition is a method which is commonly used for digitizing printed texts therefore they can be electronically searched, edited, displayed on-line, stored easily, and also used in machine processes like text-to-speech, machine translation, key data extraction. Following are the text recognition methods:

A. BLSTM

Sankaran & Jawahar [8] introduced BLSTM recognition technique for the Indian script of Devanagari lipi. The efficiency of Devanagari script is not yet comparable to Roman counterparts because of writing, style, and size complexity [24]. This technique use Recurrent Neural Network that is also called as BLSTM (Bidirectional Long Short Term Memory). The most common reasons for high word error rate is that word to character segmentation that is not required in this particular. As compared to the OCR system, the reduction of word error rate and character error rate is about 20% & 9% respectively [6, 7].

B. ANFIS

ANFIS technique is suggested by Sable & Nirve [12]. In this technique, initially scanned image converted from RGB scale to gray-scale. To obtain raw individual character samples, initially image is splitted into individual character blocks with the help of MATLAB script. To obtain a clean dataset, noise removal and pre-processing techniques will be use on raw samples. Scanning process focus some irregularities like speckle noise and salt and pepper noise, which are commonly, appear in the output image. Noise reduction is also known as smoothing or noise filtering. This is one of the most basic processes in image processing. At last, remove the header line. In character recognition, some of the features will be use like Color dominant, Histogram, GLCM, Affine moment invariant, which gives better results compare to others. ANFIS (Artificial Neuron Fizzy Interference System) technique will be use for recognition process only that gives the best result compare to other recognition techniques. Recognition rate of all Devanagari character is near about 95%.

C. Off line Handwritten Character Recognition

It can be implementing by two ways that is: online character recognition, and offline character recognition; and off-line character recognition will be further divided into machine printed, and handwritten character recognition respectively. Off-line handwritten Devanagari character recognition method proposed by Jangid [3] that uses three feature extraction techniques based on recursive sub divisions of the character image, zone density of the pixel and leading distribution of neighboring background to foreground pixels. Offline character recognition consists of image acquisition, preprocessing, region labeling, and recognition. Based on these four steps we can recognize off line text efficiently [12]. The proposed methodology obtained 94.89% recognition accuracy.

5. Text Extraction Techniques

Extraction of text is nothing but the extract only text region from panel. Text present within an image contains useful information regarding automatic indexing, annotation, and structuring of images. Extraction of text information involves localization, detection, extraction, tracking & enhancement as well as recognition of the text from a given image. Following are the various text extraction techniques:

A. Binarization Technique

It can be divided into local & global thresholding [6]. Pre-processing is must for reduction of noise. Hence, to make binarization successful, pre-processing is must and to remove non-text areas and filling the gaps in character strokes, post-processing of binarized image is required. For segmenting the image, the initial step applied will be image thresholding or image binarization. In this technique, input image is forwarded to the pre processing and after that, it will be threshold. It is divided into local thresholding and global thresholding and both are combining to perform binarization. Therefore, post-processing is completed. The input RGB image is converted to a grey scale image. Therefore, the images are captured under natural conditions; these may contain different types of noise. It can be minimized; a pre-processing stage is required for smoothing of background texture, contrast enhancement between background and text areas. In this we applied a low-pass filter to gray scale image [18].

B. *Component Extraction*

It is the process of dividing a street image into multiple fragments, which is called as super pixels. The goal of extraction is to reduce the calculation difficulties within the input image and make it easier to analyze [19]. Image extraction is used to trace objects as well as boundaries within the images. Labeling the pixels of image is nothing but the image extraction/ segmentation, where the pixels with same labels can share common characteristics such as intensity, color, and texture; and edge detection is used in most image processing applications to acquire sharp region boundaries. Tehsin et al. [22] proposed segmentation method consists of two processes that is splitting and merging. Splitting can be performing by the traditional region-based segmentation techniques, but merging is uses novel fuzzy-based method. In splitting process, original image is passing to the edge detection after that dilation & finally, association. Whereas in merging, result of splitting is passing to the feature extraction, fuzzification, fuzzy inference engine, de-fuzzification. Final output will be segmented image [15].

6. Devanagari Conversion Techniques

Text from Devanagari script to Roman script requires a detailed understanding of consonant cluster behaviors in Devanagari scripts. Many Indians as well as foreigners cannot understand regional languages due to lack of knowledge. In India, some panels are written in regional languages, which cannot understand by foreign people, therefore, they need some kind of conversion technique to understand such languages. Following are some conversion techniques available for conversion from Devanagari to Roman scripts:

A. *Machine Transliteration (MT) Approach*

It is proposed by Dixit & Dhore [9]. Proposed phonetic model translate Indian Devanagari name entities into English using hybrid approach and full consonant approach. For schwa deletion, stress analysis approach will be perfect. By using linguistic approach [20], Dhore developed rule-based phonetic model. Uni-encoded Marathi or Hindi caption entity text will be given to the syllabification module [20]. The main function of this module is to separate the input entity into syllabic number of units. After that, each syllabic unit will be converted into English using phonetic map [20]. It is implemented with the help of mapping and transliteration memory is done by writing the manual rules. Total 15, 244 named entities were tested. Hence the highest accuracy of the proposed method is about 74.14% and error rate is near about 26.80%. This particular application can be easily ported on movable devices, as there is no need of multi-lingual and bi-lingual databases to be trained [20].

B. *RBMT*

This system is proposed by Singh & Kaur [23]. The system consists of collection of grammar rules, lexicon rules and to process these rules, software program will be require. First strategy is the rule based approach, which is developed in the machine translation [20]. RBMT consists of lots of things to do with syntactic, morphological and semantic data about the target language and source language [20]. Linguistic rules will be made up from this one. For removing the scriptural barrier and language, transliteration systems will be useful [20]. YamCha toolkit is used for the implementation of SVM. Hindi & Marathi to English transliteration consists of three basic steps like pre-processing, training of bilingual corpus & testing of additional data. Training phase includes bigram, trigram, fourgram & fivegram [1].

7. Conclusion

The paper presented various panel detection techniques as well as methods with the advance functionality of feature extraction. Panel detection will be used for detection of panels among various panels like warning panels, ordinary advertisement panels, and traffic panels. From all the available panel detection methods, TCH gives 95% higher rate for all types of upper & lateral panels. Paper also presents some text detection & recognition methods that check the existence of text. As per review of techniques, hypothesis generation & off-line handwritten character recognition will gives near about 94% high rate for both detection & recognition. Some of the text extraction techniques based on SVM classification approach will gives best result. Text can be easily extracted from road traffic panels with the advance segmentation approach. Binarization gives satisfactory result of extraction. Finally, Devanagari conversation methods have been applied over final extracted text. Transliteration of Hindi to English gives 74.14 % result. Thus, for Devanagari to English conversation transliteration will be the best method as compared with many others.

References

[1] Kumaran, S., & Rajesh, R. (2009). Edge detection techniques for image segmentation: a survey of soft computing approaches. *International Journal of Recent Trends in Engineering. 1*(2), 250-254.

[2] Gonzalez, A., Bergasa, M., & Yebes, J. (2010). Automatic information recognition of traffic panels using SIFT descriptors and HMMs. *In*: Proceedings of International IEEE Annual Conference on "Intelligent Transportation Systems" (pp.1289-1294).

[3] Jangid, M. (2011). Devanagari isolated character recognition by using statistical features. *International Journal of Computer Science and Engineering. 3*(2), 2400–2407.

[4] Peer, P., & Ikica, A. (2011). An improved edge profile based method for text detection in images of natural scenes. *In*: Proceedings of International Conference on "Computer as a Tool (EUROCON)" (pp. 1-4).

[5] Sharma, S., & Prakash, J. (2011). A survey of image to text detection methodology. Department of Computer Science & Engineering. pp. 1-4.

[6] Garcia-Garrido, V. (2011). Automatic traffic signs and panels inspection system using computer vision. *IEEE Transactions on Intelligent Transportation Systems. 12*(2), 485-499.

[7] Sumathi, P., & Santhanam, T. (2012). Techniques and challenges of automatic text extraction in complex image: a survey. *Journal of Theoretical and Applied Information Technology. 35*(2), 225-235.

[8] Sankaran & Jawahar, V. (2012). Recognition of printed Devanagari text using BLSTM neural network pattern recognition (ICPR). *In*: Proceedings of 21st International Conference of IEEE.

[9] Dixit, S., & Dhore, M. R. (2012). Hindi and Marathi to English NE transliteration tool using phonology and stress analysis. *In*: Proceedings of Coling Demonstration Papers (pp.111-118).

[10]   Ashoka, N., & Bera, R. (2012). Feature extraction technique for neural network based pattern recognition. *International Journal on Computer Science and Engineering. 4*(3), 331-339.

[11]   Rathod, H., Dhore, L., & Dhore, M. (2013). Hindi & Marathi to English transliteration using SVM. *International Journal on Natural Language Computing. 2*(4), 1-17.

[12]   Sable, S., & Nirve, A. (2013). Optical character recognition for printed text in Devanagari using ANFIS. *International Journal of Scientific & Engineering Research. 4*(10), 236-241.

[13]   Fabrizio, J., & Marcotegui, B. (2013). Text detection in street level images. Springer Verlag. 519-533.

[14] Gonzalez, A., Bergasa, M., & Yebes, J. (2013). Traffic panels detection using visual appearance. *In*: Intelligent Vehicles Symposium (IV) (pp.1221 – 1226).

[15] Agarwal, S., & Hemarjani, N. (2013). Offline handwritten character recognition with Devanagari script. *IOSR Journal of Computer Engineering. 12*(6), 82-86.

[16]   Memane, R., Phatak, M., & Agarwal, P. (2013). 2D geometric shape and color recognition using digital image processing. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 2*(6), 2479-2487.

[17]   Singh, S. (2013). Optical character recognition techniques: a survey. *International Journal of Advanced Research in Computer Engineering & Technology. 2*(6), 24-32.

[18]   Bawa, K., & Sethi, K. (2014). A binarization technique extraction for Devanagari text from camera based images. *International Journal of Signal & Image Processing. 5*(2), 72-78.

[19]   Garg, A., & Jinda, S. (2014). To extract feature of handwritten Devanagari script. *International Journal of Advanced Research in Computer and Communication Engineering. 3*(7), 7501-7503.

[20]   Sahoo, S., & Tikar, S. (2014). Review methods of scene text detection and its challenges. *International Journal of Electronics and Communication Engineering. 5*(1), 74-81.

[21]   Ingle, M., & Gumaste, P. (2014). Handwritten Devanagari script recognition using phase correlation. *In*: Proceedings of IRF International Conference (pp. 96-99).

[22]   Tehsin, S., Masood, A., & Kausar, V. (2014). Fuzzy based segmentation for variable font sized text extraction from images/videos. *Mathematical Problems in Engineering*. pp. 1-10.

[23]   Singh, P., & Kaur, K. (2014). Review of machine transliteration systems. *International Journal of Advanced Engineering Applications. 7*(3), 72-80.

[24]   Gonzalez, A. (2014). Text detection and recognition on traffic panels from street-level imagery using visual appearance. *IEEE Transactions on Intelligent Transportation Systems*. 1-11.