Text Extraction from Road Display Boards using Discrete Wavelet Transformation

Ningappa T. Pujar, Vijaykumar S. M., Manjunath Kammar, and Ramesh M.Badiger

Tontadarya College of Engineering, Gadag

Corresponding Author: ningappap@gmail.com

Abstract — Now days digital cameras have more popularity due to its various applications. Text extraction from natural scene images like road display boards, government office boards and advertisement display boards has become a problem that is to change our everyday lives. The problem is challenging in nature due to variations in the color and font size, illumination change, reflections and text alignment. The method for extraction of text from road display board using kernel learning method with DWT wavelet. The proposed algorithm extraction of text in road display board is one of the important aspects in automation of vehicle.. Properties of text in road display images: motion of the vehicle, distance of image captured height of the road board convert into standard format and achieves desire result. Properties of text in natural scene images: Size, Alignment, Color, Edge and Distortion. The proposed method is insensitive and robust to noise, blur, variation in font size and style, colour, uneven thickness. The text extraction accuracy of 94.80% is achieved.

Index Terms — road display boards, natural scene images, preprocessing, wavelet method, non text objects, kernel learning, morphological operations, DWT.

I. INTRODUCTION

With the rapid development of digital devices, such as digital cameras, mobile phones and PDA's, images are now popular media in our daily lives. Texts, which are often embedded in images, contain lots of semantic information useful for our day today life. Get the text information in the image automatically can help people better understand the image and for further storage, compression, retrieval and other handing.In recent years, the automatic detection and extraction of texts from Images has gained increasing attention. Scene texts such as bottle labels, street signs, road display boards and license plates are text regions, which are present as integral parts of pictures. The text is usually linked to the semantic context of the image, and it constitutes a relevant descriptor for content-based image indexing. Direct recognition of scene text from the images captured by mobile devices could facilitate the development of a variety of new applications, such as translation, navigation, and tour guide services, road display board etc. The purpose of the extraction process is to separate text regions from the road display board image. The purpose of the recognition process is to determine the extracted text regions. Robust extraction of text from road display board's images is an essential step for successful text recognition. A very efficient text extraction method would enable the direct use of commercial OCR engines, which are normally optimized for binaries document images. However, errors due to a poor extraction method could be propagated in the recognition process.

This project presents a new system for detecting and extracting of text from road display boards. Extracting text from unconstrained images of road display board is difficult owing to the lack of any prior knowledge about the text regions, such as the color, font, size, orientation, or even the location of the text. In addition, images usually have uneven illumination, reflections on objects, and inter-reflection between objects owing to uncontrolled lighting conditions and the presence of shadows. These conditions make colors vary drastically, so the text regions may be fragmented, or the boundaries of the text region may be faint. It is also common for outdoor images to have complex layouts in which the content and background are mixed. Shapes in the background can be similar to characters; such complications make extracting text from road display board images a persistent challenge.

Text in images can exhibit many variations with respect to the following properties:

1. Size: Although the text size can vary a lot, assumptions can be made depending on the application.

2. Alignment: The caption texts appear in clusters and usually lie horizontally, although sometimes they can appear as non-planar texts as a result of special effects.

3. Inter-character distance: characters in a text line have a uniform distance between them.

4. Color: The characters tend to have the same or similar colors. This property makes it possible to use a connected component-based approach for text detection.

5. Motion: The same characters usually exist in consecutive frames in a video with or without movement. Road display board text can have arbitrary motion due to camera or object movement.

6. Edge: Most caption and text are designed to be easily read, thereby resulting in strong edges at the boundaries of text and background.

7. Compression: Many digital images are recorded, transferred, and processed in a compressed format.

The research field of text recognition receives a growing attention due to the proliferation of digital cameras and the great variety of potential applications. Such applications include number plate recognition robotic vision, image retrieval, intelligent navigation systems and applications to provide assistance to visual impaired persons. Generally text detection methods can be classified as both edge-based, connected-component based and region -based methods, Region-based methods attempt to detect and localize text regions by texture analysis. Text extraction is the stage where text components are segmented from the Background **[2][9]** .

Examples of natural scene images like road display boards with textual information can be found in Fig. 1.1

Fig 1.1 Natural scene images like road display board images

II. RELATED WORK

[1] Honggang Zhang, Kaili Zhao, Yi-Zhe Song, presented method that extracts text regions in natural scene images using low-level image features and that verifies the extracted regions through a high-level text stroke feature. [2] Yen-Lin Chen, Zeng-Wei Hong, Cheng-Hung Chuang, method for a new knowledge-based system for extracting and identifying text-lines from various real-life mixed text/graphics compound document images. The system first decomposes the document image into distinct object planes to separate homogeneous objects, including textual regions of interest, non-text objects such as graphics and pictures, and background textures. [3] Keechul Junga, Kwang In Kimb, Anil K. Jain presented a method to extract characters from natural scene images.Algorithm works well with the medium sized characters. [4] Nobuo Ezaki and Marius Bulacu, Lambert Schomaker presented a text extraction method for blind persons. [5] Xiaoqing Liu et al. proposed “Multiscale edge based text extraction from complex images”, method which automatically detects and extracts text present in the complex images using the multi scale edge information. This method is robust with respect to the font size, color, orientation and alignment and has good performance of character extraction. [6] Honggang Zhang a, KailiZhao a, Yi-ZheSong b 2013, proposed the areas of information retrieval (IR) and information extraction (IE). The proposed approach, focus on the extraction of the surface information, i.e. information that not requires complex linguistic processing to be categorized. The goal is to detect and extract passages or sequences of words containing relevant information from the prophetic narrations texts.

III. PROPOSED METHODOLOGY

The proposed method uses DWT wavelet for texture feature extraction of text object. Kernel learning method works in eight directions of Gaussian filter the result of all total direction is used for convolution process. After the convolution Otsu used to detect edges and using a function region group eliminating long edges then features are extracted with DWT. Later extract text from road display board image. The method comprises various phases such as image acquisition, preprocessing, and elimination of non-text, texture feature extraction. The block diagram of the proposed model is given in Fig. 3.1. The detailed description of each phase is presented in the following subsections..

3.1 Image acquisition

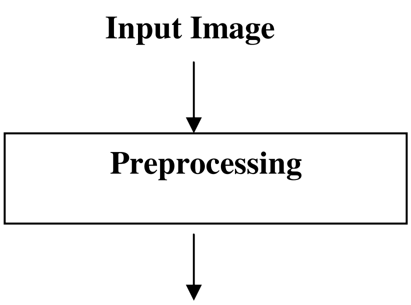
Performing image acquisition in image processing is always the first step because without the image no further processing can be carried out. In our work, road display board images are taken by Nikon Coolpix S2600 camera, these images are collected in highways, bus stand area, market etc, they are in jpg format and database is created. The images includes different color, font, orientation, alignment, complex background and low image contrast.

3.2 Preprocessing: Colored image is given as input in this phase and this input image is resized to the standard scale of image fixed. Then converted into gray scale image using an operation rgb2gray, this operation converts the colored image to gray scale image and then by applying some of the standard techniques to remove noise, distortion etc.

3.3 Detection: In the gray scale image of previous phase, applying kernel learning filter method by eight directions 0^0, 45^0 and etc , with the gaussian kernel filter used to remove the noise and distortion. Then convolving the images in each level with each direction filter. With the help of Otsu threshold edges are detected in given image.

3.4 Text Localization: By using some of the localization techniques to locate the text in images. Dilating the image for better quality of edges. Applying region groups functions to locate exact location of the text is determined and subsequently generating some of the bounding boxes around the text. Morphological operations are used for detection and localization.

3.5. Text Extraction: In this phase, after detection and localization of image, DWT method is applied to extract the text region Eliminating long edges and mapping the features, calculating the area of text in image, finally get the result text extraction from road display board.



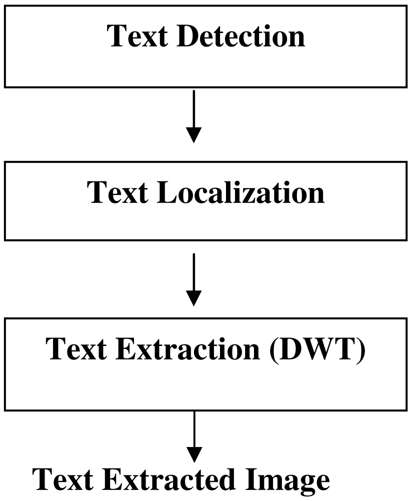


Fig 3.1 : Proposed Model Text Extraction Paradigm

IV. EXPERIMENTS AND RESULT

The proposed method has been evaluated for some samples. The proposed method is robust and insensitive to noise, uneven thickness, blur, variation in font size and style. An accuracy of 94.80% is achieved. The processing of several images dealing with various issues and the overall performance of the system are reported. The following figures show overall operations of proposed model. First converts colored input road display board image into gray scale image, morphological operations feature extraction using DWT wavelet then extract text. As a result the algorithm displays the text detected image of road display boards.

Operations of proposed model:

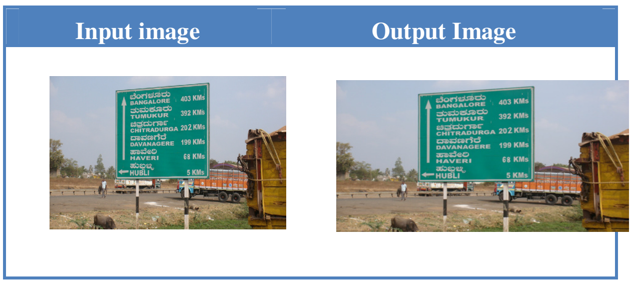


Figure 4.1 Original image is converted into resized image

Input image	Output Image

Figure 4.2 Colored input image is converted to gray scale image

Input image Output Image

Figure 4.3 Edge detected image of gray scale image.

Input image Output Image

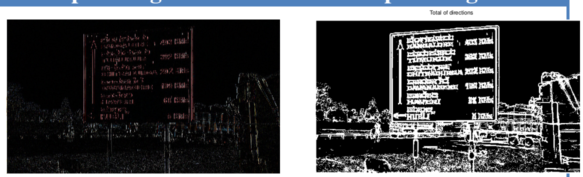


Figure 4.4 Total of eight direction Gaussian filtered image

Input image Output Image

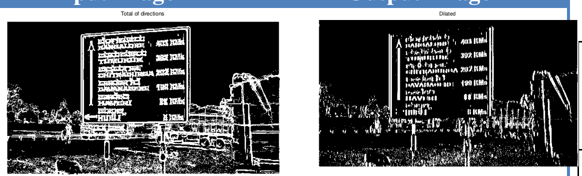


Figure 4.5 Dilated image

Output Image

Input image

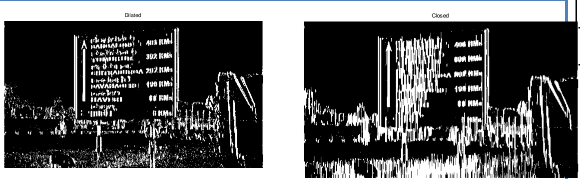


Figure 4.6 Closed image, thinning image, weak edge image

Input image Output Image

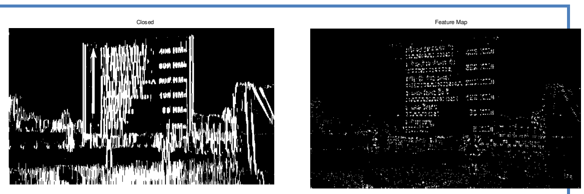


Figure 4.7 Closed image is converted to feature extraction image

Output Image

Input image

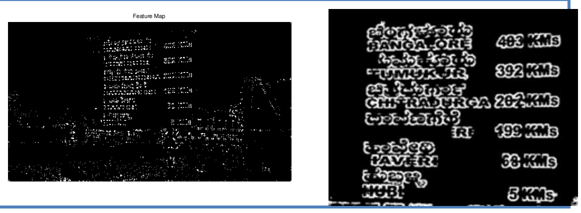


Figure 4.8 Text extracted image

The proposed method uses two performance metrics, the recall rate, and the precision rate, for the quantitative evaluation of text detection and extraction. Based on the non text objects removal, segmentation and extraction of text on several images, evaluated the overall system performance. The table 4.1 and 4.2 shows the experiments of test results and the comparisons of proposed method with other existing methods respectively.

Table 4.1: Experimental Test Results

No of Images tested	False Positive (FP)	False Negative (FN)	True Positive (TP)	True Negative (TN)	Precision Rate (PR)	Recall Rate (RR)
10 images	3	4	98	2	98	97.00
20 images	6	7	96	4	95.04	93.11
30 images	5.5	6	91	5	93.26	95.00
40 images	7	6	80	9	93.02	94.11
Total	21.5	23.00	365	20	379.32	379.22
Average	5.357	5.75	91.25	5.00	94.83	**94.80**

FP-False Positive FN-False Negative TP-True Positive
TN-True Negative PR-Precision Rate RR-Recall Rate

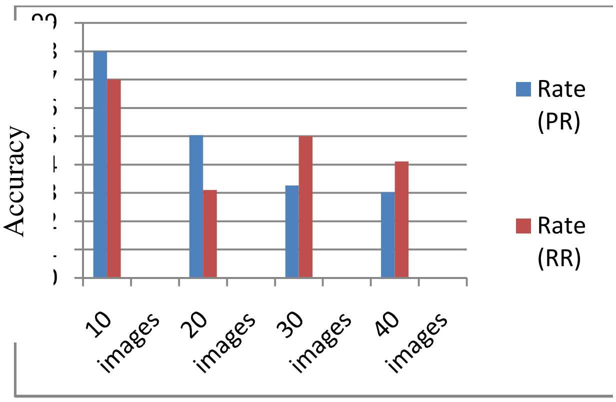
Accuracy of test result is 94.80.

Table 4.2: Comparisons with other methods

Method	Precision Rate (%)	Recall Rate (%)
Proposed method	**94.83**	**94.80**
Prashanth et al [5]	93.8	95.5
Honggang et al [4]	86	93.3

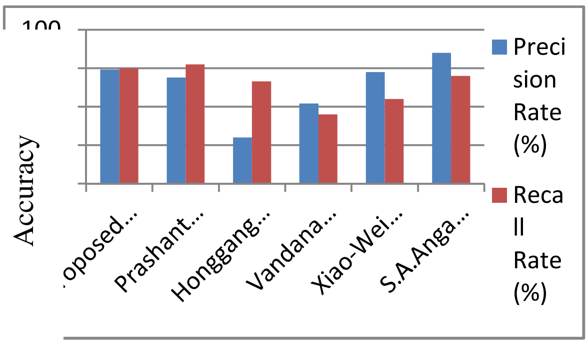
Vandana et al [3]	90.4	89
Xiao-Wei Zhang [20]	94.5	91
S.A.Angadi [2]	97	94

After Comparing with the other methods, proposed method is better than the existing one with higher accuracy. Following graph shows precision rate vs. recall rate of proposed method. Here we tested around 4 set (each set 10 images) images, got good true positive, negative values and also have observed that less false positive and negative values. This gives good precision rate and recall rate in extraction of text from road display boards.



Number of images tested

Similarly the following graph shows comparisons of proposed method with existing methods. Comparing to existing method, the proposed method misses less text information and text extracted are more accurate. It gives better precision rate, recall rate and gives good accuracy rate as compared to existing methods.



Number of images tested

V. CONCLUSION AND FUTURE SCOPE

The experimental results showed that the proposed method is computationally efficient and it gives good extraction rate and low false alarm rate. Comparing to existing method, the proposed method misses less text information extracted texts are more accurate. The proposed method uses kernel learning method and DWT wavelet for feature extraction. It has been observed that the method is robust and insensitive to noise, variation in font size and style orientation, complex background and low contrast image. The extraction of text accuracy of 94.80% is achieved. In the future efficiency of the method can be enhanced by considering new set of texture feature and increase the speed for extracting the text from images.

REFERENCES

[1] Jerod J Weinman, Erik Learned-Miller, Allen R Hanson,"Scene text recognition using similarity and a lexicon with sparse belief propagation 2007/08 Journal namePattern Analysis and Machine Intelligence, IEEE Transactions.

[2] Keechul Junga, Kwang In Kimb, Anil K. Jain, "Text information extraction in images andvideo: a survey",Pattern Recognition 37977 – 997, 2013.

[3] Jianping Zeng, Jiangjiao Duan, Wei Wang, Chengrong Wu: Semantic multi-grain mixture topic model for text analysis. Expert Syst. Appl. 38(4): 3574-3579 (2013).

[4] Yen-Lin Chen, Zeng-Wei Hong, Cheng-Hung Chuang "A knowledge-based system for extracting text-lines from mixed and overlapping text/graphics compound document images." Expert Syst. Appl., 2012.

[5] Seonghun Lee, Jin Hyung Kim: Integrating multiple character proposals for robust scene text extraction.Image Vision Comput. 31(11): 823-840 (2012).

[6] Raid Saabni, Abedelkadir Asi, Jihad El-Sana: Text line extraction for historical document images. Pattern Recognition Letters 35: 23-33 (2014).

[7] S A Angadi, M M Kotabagi, "Text region extraction from low resolution display board images using wavelet" IAEME Volume 4, Issue 1, January – April (2013).

[8] Vandana Gupta ,Kanchan Singh A "Novel Approach for Detection and Extraction of Textual Information from Scanned Document

Images and Scene Images"ijrcsse Volume 3, Issue 10, October 2013.

[9] Honggang Zhang a,n, KailiZhao a, Yi-ZheSong "Text extraction from natural scene image: A survey"Elsevier 2013.

[10] Prashant Madavanavar, Vinaykumar Patil, "Text Region Detection And Extraction From Road DirectionSign Boards". International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 7, July – 2013.

[11] K.Ganapathi Babu, A.Komali, V.Mythry, A.S.K.Ratnam, "Text Detection and Recognition in Images using Edge Based Detection and Bilinear Interpolation", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-2, issue-3, July 2012.

[12] C. Strouthopoulos , N. Papamarkos, A. E. Atsalakis,"Text extraction in complex color documents"(2012).

[13] Ronald R. Yager, Frederick E. Petry: Intuitive decision-making using hyper similarity matching. IFSA/NAFIPS 2012: 386-389.

[14] Mallikarjun Hangarge,B.V.Dhandra,"Offline Handwritten Script Identification in DocumentImages", International Journal of Computer applications (0975 8887)Volume 4 No.6, July 2010.

[15] Sangame S.K.1, Ramteke R.J.1, Rajkumar Benne2,"Recognition of isolatedhandwritten Kannada vowels",Advances in Computational Research, ISSN: 0975–3273, Volume 1, Issue 2, 2009, pp-52-55.

[16] Chuntao Jiang, Frans Coenen, Michele Zito: A survey of frequent subgraph mining algorithms. Knowledge Eng. Review 28(1): 75-105 (2009).

[17] Teofi lo E. De Campos , Bodla Rakesh Babu , Manik Varma ," Character recognition in natural images", 2009.

[18] Xiao-Wei Zhang, Xiong-Bo Zeng, Zhi-Juan Weng "text extraction algorithm under background image using wavelet transform "Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition, Hong Kong, 30-31 Aug. 2008.

[19] Narasimha Murthy K N , Dr. Y S Kumaraswamy " A Novel Method for Efficient Text Extraction from Real Time Images with Diversified Background using Haar Discrete Wavelet Transform and K-Means Clustering "IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 3, September 2011.

[20] Sanjeev Satheesh, "Text Detection and Character Recognition in Scene Images with UnsupervisedFeature Learning" 2011.