IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

Automatic Detection and Segmentation of Text in Signages

[1] Mrs.S.Sathiya, [2] Dr.M.Balasubramanian, [3] N.Ananthi,
[1 & 2] Assistant professor, Dept of Computer Science & Engg, Annamalai University, Chidambaram.
[3] ME-CSE, Dept of computer Science & Engg, Annamalai University, Chidambaram.

Abstract

The objective of this paper is to design an efficient system for automatic detection and segmentation of text from signages especially for Indian traffic conditions. In this work three major steps are proposed. First step deals with detecting the board from the loaded image using fast hybrid grayscale reconstruction. The second step uses image segmentation technique to get the rectangular board with text using toggle mapping morphological segmentation. The third step segments the individual characters using row and column wise segmentation.

Keywords- Fast Hybrid Grayscale Reconstruction, Toggle Mapping Morphological Segmentation.

I. INTRODUCTION

With the rapid development of digital devices, images and videos are now popular in our daily lives. Texts which are often embedded in images and videos contain lots of semantic information useful for video comprehension. In recent years, the automatic detection of texts from images and videos has gained increased attention. However, the large variations in text fonts, color, style, size as well as the low contrast between text and the often complicated background, often makes text detection extremely challenging. Text information in natural scene images serves as important clues for many image-based applications such as scene understanding, content-based image retrieval, assistive navigation, and automatic geo-coding. However, locating text from complex background with multiple colors is a challenging task. The text detection is a method, which is used to detect the text in the rectangular location board. In general, according to Indian Transport System the location board is present at the left side of the road. While driving on the road, the rider sometimes fails to notice the location board besides due to the following reasons: Diverted by some factors viz., personal tension, handling mobiles or chatting with the co-passenger, Not clearly visible because of the snowfall, Hided by shadows, Light illumination affecting the board.

Fig. 1. Signage

1.1 Usage Of Text Detection

Incase if the rider fails to note the place and crossed the board means, this process assist him to know the location where he is at present. The Initial step is Image acquisition, the image containing the location board is loaded to the system. The next step detects the rectangular board from the complex background. The detected object is now contains only the text.

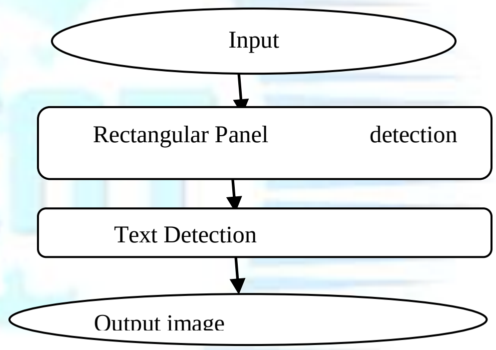

Fig. 2. Diagram block of proposed system.

II. LITERATURE REVIEW

In [1]., K. Jain and B. Yu suggested an Automatic text location without character recognition capabilities to extract image regions that contain only text. These regions can then be either fed to an optical character recognition module or highlighted for a user. In [2]., R. Lienhart and A. Wernicke proposed a method in which text lines are identified by using a complex-valued multilayer feed-forward network trained to detect text at a fixed scale and position. The network's output at all scales and positions is integrated into a single text-saliency map, serving as a starting point for candidate text lines.In [3]., Q Ye, W Gao, W Wang and W Zeng proposed an algorithm for detecting text in images and video frames. The algorithm contains two steps: initial detection and verification. In the first step, edge feature and morphology operation are employed to locate edge-dense image blocks. Empirically rules are applied on these blocks to get candidate text. In the second step, wavelet-based features are employed to represent the texture property of text.

III. METHODOLOGY OF THE WORK

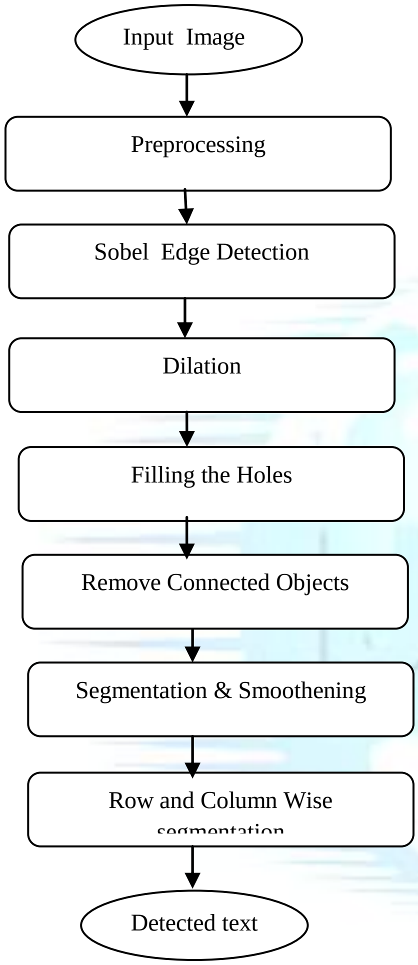

Fig.3.Object Detection

3.1 Preprocessing

This proposed system involves the following steps.

1. Get the color image as input.
2. Reading the RGB color image and converting into gray image.
3. Now calculate the monochrome luminance by combining the RGB values according to the NTSC standard, which applies coefficients related to the eye's sensitivity to RGB colors.

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

$$I = 0.2989*r+0.5870*g+0.1140*b;$$

Where,
I is an intensity image with integer values ranging from a minimum of zero,
r, g and b are the red, green and blue components respectively.

The above image preprocessing functions transform inputs into a better form for the use. The color image is considered as input, it contains the signage anywhere. The color image is then converted into gray level image. This paper is proposed to detect text in the signages. Hence the object to be detected from the image is the rectangular board.

3.2 Edge detection and morphological operations

Initially, the gray level image is obtained from the image preprocessing module, here the edges are now detected using the sobel edge detector. The sobel edge detector uses the Sobel operator. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function.

This proposed system uses the two morphological operations viz., dilation and erosion. The dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries.

Fig.4. Dilated image

3.3 Filling the holes and remove connected objects

A hole is defined as an area of dark pixels surrounded by lighter pixels To fill the holes in the above dilated image a separate algorithm is used called as fast hybrid grayscale reconstruction. The idea is to start with the two first scanning of the sequential algorithm. During the second one (anti-raster), every pixel p such that its current value could still be propagated during the next raster scanning, i.e. such that

$$\exists_q \in N_G^-(p) J(q) < J(p) \text{ and } J(q) < I(q);$$

is put into the queue. The last step of the algorithm is then exactly the same as the breadth-first propagation step of the FIFO algorithm. However, the number of pixels to be considered during this step is considerably smaller. This algorithm is described below in pseudo-code:

Algorithm: fast hybrid grayscale reconstruction

- I : mask image (grayscale)
- J : marker image, defined on domain $D_I, J \leq I$.
  Reconstruction is determined directly in J
- Scan DI in raster order :
  Let p be the current pixel;
  J(p) $\leftarrow (max\{J(q), q \in N_G^+(p) \cup \{p\}\}) \wedge I(p)$
- Scan DI in anti-raster order :
  Let p be the current pixel;
  J(p) $\leftarrow (max\{J(q), q \in N_G^-(p) \cup \{p\}\}) \wedge I(p)$
  If there exist $q \in N_G^-(p)$ such that J(q) < J(p) and J(q) < I(q)
  fifo_add(p)
- Propagation step: While fifo_ empty() = false
  p← fifo_first()
  For every pixel $q \in N_G(p)$:
  If J(q) < J(p) and I(q) ≠ J(q)
  $J(q) \leftarrow \min\{J(p), I(q)\}$
  Fifo_add(q)

This algorithm seems to offer the best compromise for computing grayscale reconstructions.

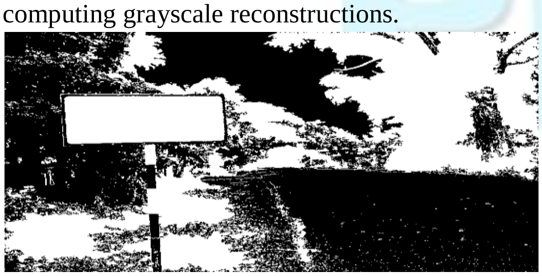

Fig.5. Image with filled holes

The rectangular board has been successfully segmented, but it is not the only object that has been found. Any objects that are connected to the border of the image can be removed using geodesic dilation. It suppresses structures that are lighter than their surroundings and that are connected to the image border. The connectivity in the imclearborder function was set to 4 to remove diagonal connections. When further dilations do not change the marker image any more, the processing is finished. The final dilation creates the reconstructed image.

Fig.6. after removing connected objects

3.4 Image Segmentation and Smoothening

Finally, in order to make the segmented object look natural, the object is smoothen by eroding the image twice with a rectangle structuring element. The rectangle structuring element is created using the strel function. Segmentation step is done by Toggle mapping which is a generic operator. Bigger homogeneous regions are mostly left unchanged, only a small dilation of these boundaries is performed. This method is called Toggle Mapping Morphological Segmentation (TMMS).

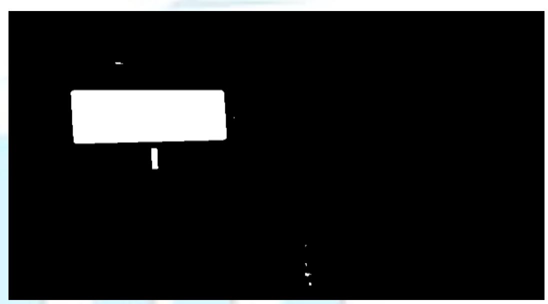

Fig.7. Segmentation

In order to make the segmented object look natural, the object is smoothen by eroding the image twice with a rectangle structuring element. The rectangle structuring element is created using the `strel` function. The smoothened object is multiplied with the gray image obtained from the color image produces image that contains only a text region is obtained.

IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

Fig.8. Smoothening

3.5 Row and Column Wise Segmentation

This step involves segmentation of image in row and column wise. The first step segments the image in horizontal manner. It counts the white pixels from left to right, if the white pixel is found moves it to the array. The above step is repeated through the last pixel of the image. Simultaneously, this segmented image is mapped with the original gray image. The above set of steps is followed for the column wise segmentation. Except that it counts the white pixels from top to bottom and moves it to the array and finally produces the image contains only the location board.

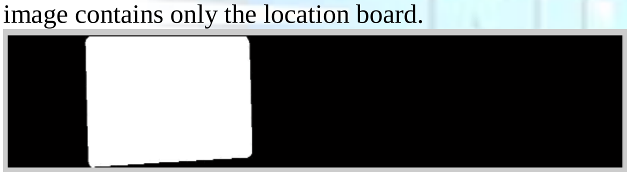

Fig. 9.Row-Wise Segmentation

Fig.10. Detected Board

IV. EXPERIMENTAL RESULTS

It has been analyzed a total of 173 kilometers of two different highways of the Tamil Nadu road network. In this stretch, there are 84 panels located at the left side of the road. Traffic panels contain not only words, but also numbers and different symbols such as direction arrows or petrol station indications. In addition, there are several samples for every panel, because each one usually appears in different frames at different distances. Since the detection and recognition have been carried out for every frame independently. There are three ranges of distance are defined: short distance when the panel is less than 30 meters far, medium distance when it is in the range 30-50 meters and long distance when the panel is further than 50 meters. The nearer the panel is, the better the performance of the algorithm is. However, the recognition rate for symbols remains above 70% no matter the distance of the panel is, due to the fact that symbols are typically bigger than characters and numbers, being easier to segment.

Some examples that this project cannot handle to locate the text information because of very small size, overexposure, characters with non-uniform colors or fade, and occlusions caused by other objects such as wire mesh.

Fig.11.Multiple Boards

Fig.12.Occlusion by Leaf

V. CONCLUSION AND FUTURE WORK

This method detects traffic panels which are supposed to be high-density text areas in the image in first place and text inside the panel regions secondly, a different approach has been implemented. It has been showed that the performance of the proposed method is quite good, as it detects most of the words in almost all the panels at short distance. Numbers and symbols detection needs to be improved, as well as text detection at long distance.

In the future it is proposed to make use a priori of the regulation on traffic panels, which describes how they must be designed, in order to make a more effective recognition algorithm. For instance, in the above side of the panels, it is supposed to appear just numbers and symbols. The main problem is that there is image distortion and the panel edges are not straight in the image when the panels are near. It is also necessary to improve the estimation of the distance of the panels in the image in order to geo-localize the traffic panels more and also attempt to improve the efficiency and transplant the algorithms into a navigation system prepared for the way finding of visually impaired people.

VI. REFERENCES

[1] K. Jain and B. Yu, “Automatic text location in images and video frames,” Pattern Recognition. vol. 31, No.12, pp. 2055-2076, 1998.
[2] R. Lienhart and A. Wernicke, “Localizing and segmenting text in images and videos,” IEEE

trans.on Circuits and Systems for Video Technology,Vol.12, No.4, April, 2002.

[3] Q Ye, W Gao, W Wang, W Zeng "A robust text detection algorithm in images and video frames" ICICS PCM 2003.

[4] K. Junga, K. In Kimb, Anil K. Jainc," Text information extraction in images and video: a survey", Pattern Recognition.vol.37, pp. 977 – 997, 2004.

[5] Q.Yea, Q.Huangb, Wen Gaoa,b, D. Zhao, "Fast and robust text detection in images and video frames", Image and Vision Computing.vol. 23, pp.565–576, 2005.

[6] J. Gllavata1, Ralph Ewerth1 and Bernd Freisleben, "A Robust Algorithm for Text Detection in Images", SFB/FK 615.

[7] V.N. Manjunath Aradhyaa, M.S. Pavithrab and C. Naveenac,"A Robust Multilingual Text Detection Approach Based on Transforms and Wavelet Entropy", Procedia Technology.vol.4, pp. 232 – 237, 2012.

[8] Ming Zhao, Shutao Li , James Kwok ,"Text detection in images using sparse representation with discriminative dictionaries", Image and Vision Computing.vol. 28, pp.1590–1599, 2010.

[9] ´A. Gonz´alez, L.M. Bergasa and et al., "Text Recognition on Traffic Panels from Street-level Imagery", 2012 Intelligent Vehicles Symposium, June 3-7, 2012.

[10] Christian Wolf, Jean-Michel Jolion," Model based text detection in images and videos: a learning approach", Technical Report LIRIS RR-2004-13