Text Detection and Recognition in Images using Edge Based Detection and Bilinear Interpolation

K.Ganapathi Babu, A.Komali, V.Mythry, A.S.K.Ratnam

Abstract*: Text can be anything which conveys meaning through the use of signs. This revelation means that the realm of text can be extended from the form of written discourse to that of any object, regardless of whether it is aesthetic, written or spoken. 'Text' refers to anything which is capable of being read or interpreted. For example 'Text' may represent natural language like English or it may be useful for describing the contents of an image. Identifying text areas and recognizing the text is a complex task in image processing. Because the image may be grayscale or color or it represents a natural image, world map or space images. And the text appears in the image may also too small to recognize. This paper presents the combination of edge based detection and bilinear interpolation methods to identify and recognizing text in different images.

***Index Terms*: Bilinear interpolation, Edge based detection, Image, Text.**

I. INTRODUCTION

Content-based multimedia database indexing and retrieval tasks require automatic extraction of descriptive features that are relevant to the subject materials (images, video, etc.). The typical low-level features that are extracted in images include measures of color [1], texture [2], or shape [3]. Although these features can easily be obtained, they do not give a precise idea of the image content. Extracting more descriptive features and higher level entities, such as text [4] and human faces [5], has recently attracted signi3cant research interest. Text embedded in images and video, especially captions, provide brief and important content information, such as the name of players or speakers, the title, location, date of an event, etc. This text can be a keyword resource as powerful as the information provided by speech recognizers.

Identifying text areas and recognizing the text is a complex task in image processing. Because the image may be grayscale or color or it represents a natural image, world map or space images. In this paper, to identify text areas, local threshold and Boolean function based edge detection method [6] is used. This edge detection method is derived from canny edge detection method [7]. It takes advantage of both local and global thresholding to find edges. Unlike other edge detectors, it converts a window of pixels into a binary pattern based on a local threshold, and then applies masks to determine if an edge exists at a certain point or not. By calculating the threshold on a per pixel basis, the edge detector should be less sensitive to variations in lighting throughout the picture. It does not rely on blurring to reduce noise in the image. It instead looks at the variance on a local level. After identifying the text areas, these are given as input to the OCR system [8] to recognize the characters in the image.

An important problem while recognizing text is that the text appears to small to read by the OCR system. So there is a need to enhance the text area before it is given as input to the OCR system. For this we use bilinear interpolation method. Bilinear method is a resampling method that uses the distance- weighted average of the four nearest pixel values to estimate a new pixel value. These new pixels are calculated when we need to enlarge an image. In this paper we are using this method to enhance the identified text area before it is given as input to the OCR system.

This paper is organized as follows: Section 2 presents the local threshold and Boolean function based edge detection method. Section 3 presents the bilinear interpolation method. Section 4 presents the proposed work. And finally section 5 gives the conclusion.

II. LOCAL THRESHOLD AND BOOLEAN FUNCTION BASED EDGE DETECTION

This edge detector is fundamentally different than many of the modern edge detectors derived from Canny's edge detection method. It takes advantage of both local and global thresholding[9] to find edges. The algorithm is as follows:

1. Apply a local threshold to a 3x3 window of the image. Because this is a local threshold, it is recalculated each time the window is moved. The threshold value is calculated as the mean of the 9 intensity values of the pixels in the window minus some small tolerance value. If a pixel has an intensity value greater than this threshold, it is set to a 1. If a pixel has an intensity value less than this threshold, it is set to a 0. This gives a binary pattern of the 3x3 window.
2. Compare the binary pattern to the edge masks. There are sixteen possible edge-like patterns that can arise in a 3x3 window, as shown below:

Manuscript received on July, 2012.

K.Ganapathi Babu Pursuing M.Tech in Computer Science Engineering at Vignan's LARA Institute Of Technology and Sceince, Vadlamudi, Guntur Dist., A.P., India.

A.Komali Pursuing M.Tech in Computer Science Engineering at Vignan's LARA Institute Of Technology and Sceince, Vadlamudi, Guntur Dist., A.P., India.

V.Mythry Pursuing M.Tech in Computer Science Engineering at Vignan's LARA Institute Of Technology and Sceince, Vadlamudi, Guntur Dist., A.P., India.

A.S.K.Ratnam Head, Department of CSE, Vignan's LARA Institute Of Technology & Sceince,Vadlamudi Guntur Dist., A.P., India.

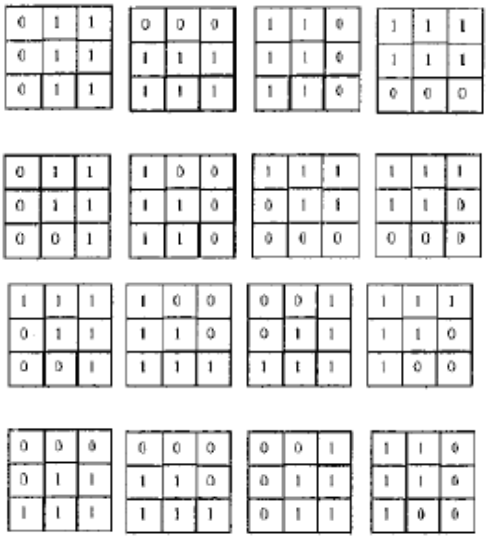

Figure-1

If the binary window obtained in step 1 matches any of these sixteen masks, the center pixel of the window is set to be an edge pixel.

3. Repeat steps 1 and 2 for each pixel in the image as the center pixel of the window. This will give all edges, but it will also give some false edges as a result of noise.
4. Use a global threshold to remove false edges. The variance for each 3x3 window is calculated, which will have a maximum at an edge. This value is then compared with a global threshold based on the level of noise in the image. If the value is greater than the threshold, it is kept as an edge. If it is not greater than the threshold, it is removed.

III. BILINEAR INTERPOLATION

Bilinear interpolation is a smoother interpolation method. Given four neighboring image coordinates f(n_{10} , n_{20}), f(n_{11} , n_{21}), f(n_{12} , n_{22}), and f(n_{13} , n_{23}) , then the geometrically transformed image g(n_1, n_2) is computed as:

$$g(n_1, n_2) = A_0 + A_1 n_1 + A_2 n_2 + A_3 n_1 n_2 \quad \text{------ (1)}$$

which is a bilinear function in the coordinates (n_1, n_2). The bilinear weights A_0, A_1, A_2, A_3 are found by solving

$$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} 1 & n_{10} & n_{20} & n_{10}n_{20} \\ 1 & n_{11} & n_{21} & n_{11}n_{21} \\ 1 & n_{12} & n_{22} & n_{12}n_{22} \\ 1 & n_{13} & n_{23} & n_{13}n_{23} \end{bmatrix}^{-1} \begin{bmatrix} f(n_{10}, n_{20}) \\ f(n_{11}, n_{21}) \\ f(n_{12}, n_{22}) \\ f(n_{13}, n_{23}) \end{bmatrix} \quad \text{----(2)}$$

Thus, g(n_1, n_2) is defined to be a linear combination of the gray levels of its four nearest neighbors. The linear combination defined by (2) is in fact the value assigned to g(n_1, n_2) when the best planar fit is made to these four neighbors. This process of optimal averaging produces a visually smoother result.

IV. PROPOSED WORK

The proposed work uses the combination of local threshold and Boolean function based edge detection method and the bilinear interpolation method. The proposed method is shown in the following figure:

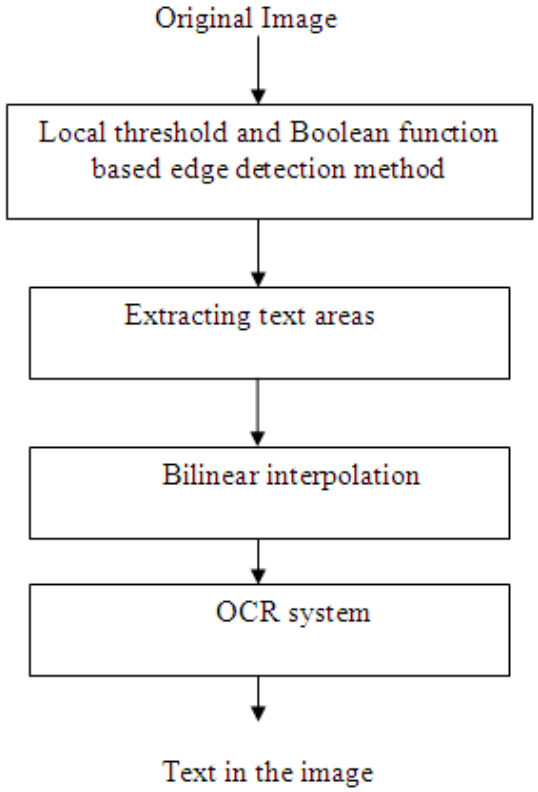

Figure-2 : Proposed method

The steps of the proposed work is as follows:

1. Apply the edge detection method to the original image. The edge detection method used in this paper is local threshold and Boolean function based edge detection method.
2. After applying the edge detection method, we extract the text areas in the given image.
3. The identified text areas might be too small to recognize. So to enhance these text areas, bilinear interpolation method will be used.
4. After enhancing the text areas, these are given as input to the OCR system, which results the 'Text' appeared in the given image.

The proposed method can be illustrated by the following figures:Let the original be,

Figure 3: Original image

After applying the local threshold and Boolean function based edge detection method, we get the result as

Figure 4: Result of edge detection method

After extracting the text areas, we get

Figure 5: Extracted text areas

After applying the bilinear interpolation, the extracted text areas are enhanced and the results are

Figure 6: Enhanced text areas

These enhanced text areas are given as input to the OCR systems which results the text appeared in the given image. The results are,

- EUGENE
- NORTH
- INTERSTATE
- 5

V. CONCLUSION

With the widely use of digital image, text location/ detection and recognition is considered as a key part of the image text information extraction system. In this paper we use two important methods of image processing for detecting and recognizing text in different images like natural images or satellite images. They are edge detection method and bilinear interpolation. The local threshold and Boolean function based edge detection method is derived from the canny's edge detection method and it uses local and global thresholding to find edges in the image. The image may also contain text areas, too small to recognized by the OCR system. So to enhance such type of text areas, bilinear interpolation is used. Thus using the combination of edge detection and bilinear interpolation provides better results.

REFERENCES

1. M. Swain, H. Ballard, Color indexing, Int. J. Comput. Vision 7 (1991) 11–32.
2. B. Manjunath, W. Ma, Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. Mach. Intell. 18 (8) (1996) 837–842.
3. F. Mokhtarian, S. Abbasi, J. Kittler, Robust and eAcient shape indexing through curvature scale space, in: British Machine Vision Conference, 1996, pp. 9–12.
4. D. Chen, H. Bourlard, J.-P. Thiran, Text identification in complex background using SVM, in: International Conference on Computer Vision and Pattern Recognition, 2001, pp. 621–626.
5. R.K. Srihari, Z. Zhang, A. Rao, Intelligent indexing and semantic retrieval of multimodal documents, Inf. Retri. 2 (2/3) (2000) 245–275.
6. M.B. Ahmad and T.S. Choi , Local Threshold and Boolean Function Based Edge Detection, IEEE Transactions on Consumer Electronics, Vol. 45, No 3. August 1999.
7. An approach of Canny edge detection with virtual hexagonal image structure, by Xiangjian He, pages 879-882, Control, Automation, robotics and vision, ICARCV 2008.
8. Historical review of OCR Research and Development, by S. Moric, Y. Suen and k. Yamamoto, P. 1029- 1057, IEEE, 1992.
9. A study of local and global thresholding techniques in Text categorization, Nayer M. Wanas, Dina A. Said, Naia H. Hegazy and Nevin M. Darwish, Fifth Australasian Data Mining conference 2006.

AUTHORS PROFILE

K.Ganapathi Babu pursuing his (M.Tech in Computer Science Engineering) at Vignan's LARA Institute Of Technology and Sceince, Vadlamudi, Guntur Dist., A.P., India. His research interest includes Data Mining and Image Processing.

A.Komali pursuing her (M.Tech In Computer Science Engineering) at Vignan's LARA Institute Of Technology and Sceince, Vadlamudi, Guntur Dist., A.P., India. Her research interest includes Data Mining and Image Processing.

V.Mythry pursuing her (M.Tech in Computer Science Engineering) at Vignan's LARA Institute Of Technology and Sceince, Vadlamudi, Guntur Dist., A.P., India. Her research interest includes Data Mining and Image Processing..

A.S.K.Ratnam, Head, Department of CSE, Vignan's LARA Institute Of Technology & Sceince,Vadlamudi Guntur Dist., A.P., India. His research interest includes Data Mining and Image Processing.