MAT JOURNALS
Journal of Electronic Design Engineering
Volume 1 Issue 1-2

Text Extraction from Natural Images of Different Languages Using ISEF Edge Detection

Sanjay K. Shah[1], Deven N. Trivedi[2], Keyur Chauhan[1], Ankur Macwan[1], Hitesh Chaukikar[1]

[1] Department of Electronic and Communication Department (EC), Madhuben and Bhanubhai Patel Institute of Engineering for Studies and Research in Computer and Communication Technology, New V.V. Nagar, India

[2]C. U. Shah University, Near Kothariya Village, Wadhwan City, Gujarat, India

E-mail: skshah@mbict.ac.in

Abstract

In this paper, we proposed the algorithm text extraction of different images of languages. In computer vision research area, text is very important in images. Here, we use edge based extraction of text using ISEF (infinite symmetrical edge filter). ISEF is optimal edge detector which gives accurate results for text in images. Text extraction involves detection, localization, tracking and enhancement. Large numbers of technique have been proposed for the text extraction. Our aim is to present robust technique for text extraction of different languages images.

Keywords: *Text detection, ISEF algorithm, edge detector, image processing, extraction*

INTRODUCTION

Current trend of research area of image processing has much interest in content retrieval. It can be derived in the perceptual and semantic content. Perceptual includes colour, intensity, shape and texture and semantic includes objects, events and their relations. Texts are also easily and clearly describe the contents of an image. Since, the text data can be embedded in an image. Up till now it has been extracted by two basic techniques. That are edge and connected component based technique.

A text extraction system receives an input in the form of an image or a sequence of images. The problem can be divided into the following parts.

***Fig. 1:** Architecture of Text Extraction.*

Text extraction divided into the following parts (i) detection (ii) localization (iii) tracking (iv) extraction and enhancement and (v) recognition.

Text detection means to detect the text which is presences in image. Kim's scene-change detection method is not described in detail. He mentioned that very low threshold values are needed for scene-change detection because the portion occupied by a text region relative to the whole image is usually small. Smith and Kanade defined a scene-change based on the difference between two consecutive frames and then used this scene-change information for text detection [1–3]. Text localization methods which divided into two types: region and texture-based. Region-based methods use the properties of the color or gray-scale in a text region or their differences with the corresponding properties of the background.

Text Extraction is done using two basic methods. One is region based while the other is based on texture.

Region Based Methods

Region-based methods use the properties of the color or gray-scale in a text region or their differences with the corresponding properties of the background. These methods also divided into two parts: connected component (cc) based and edge-based.

CC-based methods use approach by grouping small components into successively larger components until all regions are identified in the image. It's a four-stage method: (i) binarization (ii) tentative character component (iii) character recognition (iv) relaxation operation.

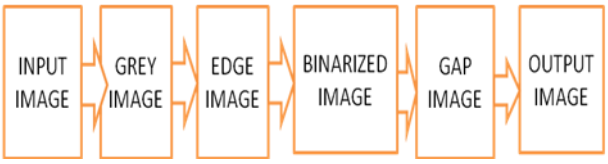

***Fig. 2:** Basic Block Diagram for Connected Component Based.*

Edge-based methods are based on contrast between the text and the background. Edges of the text boundary are identified and merged. Smith and Kanade apply a 3×3 horizontal differential filter to an input image and perform thresholding to find vertical edges and smoothing operation that is used to eliminate small edges [1]. The RGB components of a color input image are combined to give an intensity image *Y* as follows:

Y = 0.299R + 0.587G + 0.114B

Where R, G, and B are the red, green, and blue components.

Fig. 3: *Basic Block Diagram for Edge Based Text Extraction.*

Chen used the canny operator to detect edges in an image [2]. One edge point in a small window is used in the estimation of scale and orientation to reduce the computational complexity. The edges of the text are then enhanced using this scale information. Morphological dilation is performed to connect the edges into cluster.

Texture Based Methods

This methods use the observation that text in images have detect textural properties that distinguish them from the background. It's based on Wavelet, FFT, spatial variance, etc.

which utilize a horizontal window to compute the spatial variance for pixels in a local neighbourhood. Then the horizontal edges in an image are identified using a canny edge detector, and the small edge components are merged into longer lines [4–6].

We are focusing on text extraction by edge based techniques. In Table 1 we summarize the work done in this field up till now.

Table 1: *Review of Edge Detection Technique.*

Author	**Year**	**Approach**	**Features**
Smith and Kanade	1995	3×3 Filter Seeking Vertical Edges	Caption Text, Localization
Xiaoheng Yang, Hiroki Takahashi, Masayuki Nakajima	2004	Investigation of Robust Colour Model for Edge Detection on Text Extraction from Scenery Images	Simple Edge Computation Scheme
Xiaoqing Liu	2005	An Edge-Based Text	Landmarks, Scene

Jagath Samarabandu		Region Extraction Algorithm for Indoor Mobile Robot Navigation	Text, Text Localization and Extraction
Xiaoqing Liu and Jagath Samarabandu	2006	Multiscale Edge-Based Text Extraction from Complex Images	Text Localization and Extraction
Sachin Grover, Kushal Arora, Suman K. Mitra	2009	Text Extraction from Document Images Using Edge Information	Edge Based Feature
Seonghun Lee, Min Su Cho, Kyomin Jungz and Jin Hyung Kim	2010	Scene Text Extraction with Edge Constraint and Text Co linearity	Text Region Candidates

ISEF EDGE DETECTION

In the image the edge can be detected by any of the template based method but the ISEF is the shen-castan infinite symmetrical exponential filter based edge detector [6, 7]. Shen and Castan agree with Canny about the general form of the edge detector: a convolution with a smoothing kernel followed by a search for edge pixels. It is single dimension function [8]. Which can be given real continues function by below equation (1).

$$f(x) = \frac{P}{2} e^{-P/|x|} \qquad (1)$$

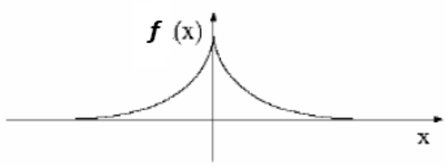

Fig. 4: Exponential Function.

$$f(x,y) = a e^{-p\|x\|+\|y\|} \qquad (2)$$

$$f(x,y) = \frac{(1-b)b^{\|x+y\|}}{(1+b)} \qquad (3)$$

Where, b=thinning factor, 0<b<1. First the whole image will be filtered by the recursive ISEF filter in X and in Y direction, which can be implementing by using equations as given below.

X direction recursion:

$$y1[i,j] = \frac{(1-b)}{(1+b)}\ I[i,j] + by1[i,j-1] \qquad (4)$$

Where j = 1…….N, i=1….M

$$y2[i,j] = \frac{b(1-b)}{(1+b)}\ I[i,j] + by1[i,j-1] \qquad (5)$$

Where j = 1…….N, i=1….M

$$r[i,j] = y1[i,j] + y2[i.j+1]$$

(6)

Y direction recursion:

$$y1[i,j] = \frac{(1-b)}{(1+b)}\ I[i,j] + by1[i,j-1] \qquad (7)$$

Where j = 1…….M, i=1….N

$$y2[i,j] = \frac{b(1-b)}{(1+b)}\ I[i,j] + by1[i,j-1] \qquad (8)$$

Where j = 1…….M, i=1….N

$$y[i,j] = y1[i,j] + y2[i.j+1] \qquad (9)$$

Where b=thinning factor (0<b<1)

After that Laplacian image can be approximated by subtracting the filtered image from the original image. So, at the location of an edge pixel there will be zero crossing in the second derivative of the filtered image. The first derivative of the image function should have an extreme at the position corresponding to the edge in image and so the second derivative should be zero at the same position. So, for thinning purpose we apply non maxima suppression. The gradient at the edge pixel is either a maximum or a minimum. If the second derivative changes sign from positive to negative this is called positive zero crossing and if it changes from negative to positive it is called negative zero crossing. We will allow positive zero crossing to have positive gradient and negative zero crossing to have negative gradient, all other zero crossing we assumed to be false and are not considered to an edge.

Now, gradient applied image has been thinned, and ready for the thresholding. The simple thresholding can have only one cut off but Shen-Castan suggests using Hysteresis thresholding. Spurious response to the single edge caused by noise usually creates a streaking problem that is very common in edge detection. The output of an edge detector is usually threshold, to decide which edges are significant and streaking means the breaking up of the edge contour caused by the operator fluctuating above and below the threshold. Streaking can be eliminated by thresholding with Hysteresis. Individual weak responses usually correspond to noise, but if these points are connected to any of the pixels with strong responses, they are more likely to be actual edge in the image. Such connected pixels are treated as edge pixels if there response is above a low threshold. Finally, thinning is applied to make edge of single pixel.

Below is the block diagram of ISEF algorithm.

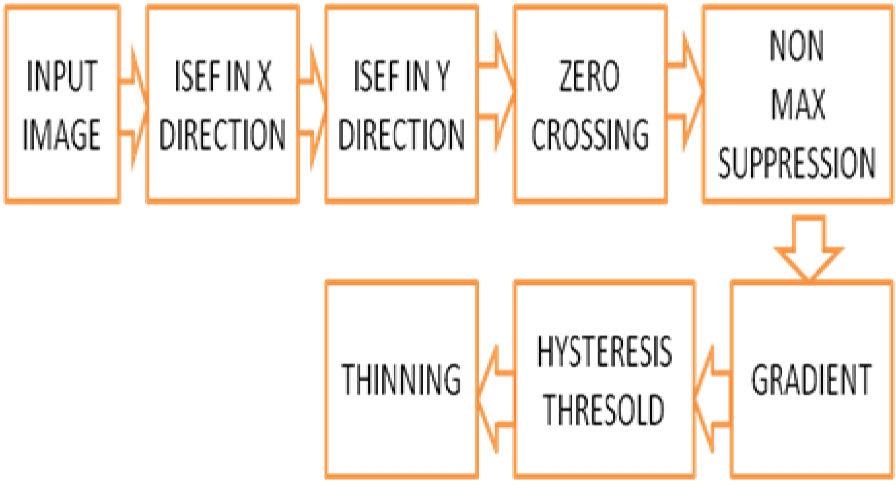

Fig. 5: *Block Diagram of ISEF Algorithm.*

PROPOSED ALGORITHM

In the proposed algorithm an image is taken as an input.

Table 2: *Step of Proposed Algorithm.*

Sr. No.	Steps
1	Capture the image.
2	(For text detection) create a Gaussian pyramid by convolving the input image with a Gaussian kernel.
3	Apply ISEF algorithm to each image of Gaussian pyramid to get edges of images.
4	Resize and combine all Gaussian pyramid images in to one image.
5	(For localization) apply morphological op dilate and closing operations using a sufficiently large structuring element to cluster candidate text regions together.
6	Find weak and strong edges, combine both the edges and apply thinning operation.
7	Eliminate long edges and obtain short edge image.
8	Apply dilation on the short edge image obtained in step 7. Multiply this image with the total image obtained in step 4.

9	Apply dilation on the refined image obtained in step 8. Call this image a reference image.
10	(For character extraction) create final output image with text in white pixels against a plain black background.
11	Display the output image.

RESULTS AND ANALYSIS

We use different types of languages images to check the robustness of ISEF based algorithm. For the experimentation of the proposed algorithm. The performance of proposed ISEF based technique has been evaluated based on its precision and recall rates obtained. Precision and recall rates are calculated as follows:

$$\text{Precision rate} = \frac{\text{correctly detected words}}{\text{correctly detected words} + \text{false positive}} \times 100\%$$

$$\text{Recall rate} = \frac{\text{correctly detected words}}{\text{correctly detected words} + \text{false negative}} \times 100\%$$

Precision rate is defined as above equation in which the false positives, which are the non-text regions in the image and have been detected by the algorithm as text regions. While Recall rate is defined in which the false negatives, which are text words in the image and have not been detected by the algorithm.

Thus, precision and recall rates are useful as measures to determine the accuracy of the proposed ISEF based algorithm in locating correct text regions and eliminating non-text regions.

Original and resultant images are shown in Figures 6–8. Corresponding precision and recall rates are given in Tables 3–5.

Bengali Languages Images' Results

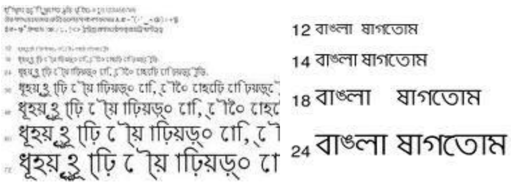

Image 1 Image 2

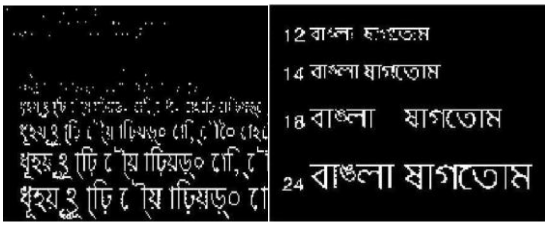

Fig. 6: *Row 1 Bengali Language Images, Row 2 ISEF based Results of Row 1 Images.*

Table 3: *Bengali Languages Images' Results Rates.*

Images	Precision Rate (%)	Recall Rate (%)
Image 1	60.00	70.00
Image 2	225.00	75.00

Kannada Language Images' Results

Image 1 Image 2

Fig. 7: *Row 1 Kannada Language Images, Row 2 ISEF Based Results of Row 1 Images.*

Table 4: *Kannada Language Images' Results Rates.*

Images	Precision Rate (%)	Recall Rate (%)
Image 1	80.00	95.00
Image 2	90.00	96.00

Urdu Langauge Images' Results

Image 1 Image 2

Fig. 8: *Row 1 Urdu Language Images, Row 2 ISEF Based Results of Row 1 Images.*

Table 5: *Urdu Language Images' Results Rates.*

Images	Precision Rate (%)	Recall Rate (%)
Image 1	80.00	92.00
Image 2	82.00	95.00

CONCLUSION

The results obtained by proposed ISEF edge detection based algorithm on a different set of images with respect to precision and recall rates. In every image the recall rates we get are higher than the precision rates. The average precision rates are 70% and the average recall rates are 85%. Thus we can conclude that it is more robust algorithm and novel approach to detect text regions from natural languages images.

REFERENCES

1. M.A. Smith, T. Kanade. *Video skimming for quick browsing based on audio and image characterization*. Technical Report Cmu-Cs-95-186, Carnegie Mellon University. 1995.
2. D. Chen, K. Shearer, H. Bourlard. *Text enhancement with asymmetric filter for video* ocr. Proc. of International Conference on Image Analysis and Processing. 2001; 192p.
3. B. Sin, S. Kim, B. Cho. *Locating characters in scene images using frequency features*. Proc. of

International Conference on Pattern Recognition. 2002; 3: 489–492p.s
4. B.L. Yeo, B. Liu. *Visual content highlighting via automatic extraction of embedded captions on mpeg compressed video*. Proc. of Spie. 1996; 142–149p.
5. Y. K. Lim, S. H. Choi, S.W. Lee. *Text extraction in mpeg compressed video for content-based indexing*. Proc. of International Conference on Pattern Recognition. 2000; 409–412p.
6. J. Ohya, A. Shio, S. Akamatsu. *Recognizing characters in scene images*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1994; 16(2): 214–224p.
7. Kunal J Pithadiya, Chintan K Modi, Jayesh D Chauhan, et al. *Performance evaluation of ISEF and Canny edge detector in Acrylic fiber quality control production.* Proceedings of National Conference on Innovations in Mechatronics Engineering. GCET, V V Nagar, India. 2009; 89–92p.
8. Castan S, Zhao J, Shen J. *New edge detection* methods based on exponential filter. Pattern Recognition, Proceedings 10th International Conference. 1990; 1(16): 709– 711p