

Boosted Edge Detection Algorithm for Unstructured Environment in Document Using Optimized Text Region Detection

N.P.Revathy¹, S.Janarthanam² and S.Sukumaran³

^{1&2}Assistant Professor,

^{1&2}Department of Computer Science, Gobi Arts & Science College, Erode, Tamil Nadu, India

³Associate Professor, Erode Arts and Science College, Erode, Tamil Nadu, India

E-Mail: np.revathy@yahoo.in, professorjana@gmail.com, prof_sukumar@yahoo.in

Abstract - Document images are more popular in today's world and being made available over the internet for Information retrieval. The document images becomes a difficult task compared with digital texts and edge detection is an important task in the document image retrieval, edge detection indicates to the process of finding sharp discontinuation of characters in the document images. The single edge detection methods causing the weak gradient and edge missing problems adopts the method of combining global with local edge detection to extract edge. The global edge detection obtains the whole edges and uses to improve adaptive smooth filter algorithm based on canny operator. These combinations increase the detection efficiency and reduce the computational time. In addition, the proposed algorithm has been tested through real-time document retrieval system to detect the edges in unstructured environment and generate 2D maps. These maps contain the starting and destination points in addition to current positions of the objects. This proposed work enhancing the searching ability of the document to move towards the optimal solution and to verify the capability in terms of detection efficiency.

Keywords: Adaptive Smoothing, Character Recognition, Document Layout, Edge Detection, Edge Preserving, Gradient Mapping

I. INTRODUCTION

Now a day's all over digitization technology is used with digitalization of the world, the paper based documents need to be converted into digital to make them handier, searchable and for preserving of the documents. Optical Character recognition is used for this process. OCR can be described as Mechanical or electronic conversion of scanned images where images can be handwritten, typewritten or printed text [2]. One of the main challenges in this field is to recognize the cursive handwritten text. The problem is to align the text, segment individual characters and then recognize the characters. This paper describes the needs, applications, basic terminology and specific research contributions related to handwritten character recognition. The detection techniques line-out the processing steps involved in the detection process and also classify various methods related to each step. Handwriting skill is unique for each person [1]. Each script contains an own set of icons. These icons are known as characters or letters. Each letter has a specified shape.

Text detection is a crucial step in processing textual information in images. Finding the text regions is the first

stage of a standard OCR pipeline in extracting image text[10]. Determining the location of text is also important for high-level image content understanding, the text location indicates the meaning of certain image text element, such as the label of the x- versus y-axis in a graph. Practical applications aside, in this paper, exclusively concerned with optimizing the performance of text detection, which is a fundamental research problem in image text processing.

During the development of the corpus, the clear guidelines are exactly constitutes on an image text region and how to manually mark the image region linked to the string. Then compared the proposed algorithm against three existing state-of-the-art text detection methods. Even though our algorithm applied for processing all images types, it is especially beneficial for images embedded in document publications.

II. RELATED WORK

Briefly look at prior work on image processing algorithms for image text detection, is concerned with separating image text elements from other elements in an image. Ohya *et al.*, [2] presented an algorithm for text detection from document images. In their work, they first detect character components according to gray-level differences and then match the results to standard character patterns captured in a database. The method is very robust to the font, size and intensity variation in the image texts, but is not able to deal with color and orientation changes. Zhong *et al.*, [3] mentioned a connected component method for locating texts in a complex color image. Their method analyzes the color histogram of the RGB space to detect text regions. Jung [4] introduced a neural network approach for identifying text in color images. The text detection problem for texts with different orientations and other distortions, Messelodi and Modena[5] described the use of low level image features such as density and contrast to detect image texts, with the ability to deal with skew in the image text. Hasan and Karam [6] also proposed morphological approach for image text detection, is robust to the presence of noise, text orientation, skew and curvature.

A. Contribution of Our Work

To overcome the limitation, it is necessary to develop an

approach to segment the regions properly. This inspired us to work on the problem of text detection in document images to segment the region along with text region detection. Consumption of time in execution is also important to provide good results in real time [7]. Having this factor in mind has considered the fact that edges are reliable features of text regardless of color or intensity, layout, orientation etc. The edge detection operation is performed using the basic operators of mathematical morphology. Using the edges the algorithm has tried to find out text with connected components. These components have been labeled to identify different components of the text in the document image. Once the components have been identified, the variance is found for each connected component considering the gray levels of those components.

B. Traditional Histogram Analysis Based Text Detection

One of the most popular text region detection methods to analyze the vertical and horizontal projection of histograms in document images. More concretely, given an input image, first detect the edge pixels in the image [8]. Then a vertical and a horizontal projection histogram are derived. It is assumed that text regions generally exhibit higher density of edge pixels than non-text regions. The vertical and horizontal histograms will thus show the highest density of edge pixels in text areas. A density threshold defines the exact dimensions of the text area along the vertical and horizontal histogram. The traditional histogram-based analysis technique does not cope well with distributed and nested text layout. To address this problem to introduce a new iterative pivoting histogram analysis procedure for text region detection.

C. Region Based Text detection

Region-based methods use the property of the gray-scale or color in a text content region or their differences with the corresponding properties of the background of that image. This technique follows bottom-up approach method by grouping the small components into consecutively larger components step by step until all the similar regions are identified in an image [9-11].

Leon *et al.*, [10] presented a method for caption text detection. It included a generic indexing system dealing with semantic concepts are used to automatically detect the text region using common image description. The presented the image description which is a hierarchical region-based image technique and introduced the algorithm for text detection. This region-based algorithm is divided into three stages.

1. Text candidate spotting: It is an attempt to segment the text region from the background.
2. Text characteristics verification: In this stage the text content regions are grouped to discard those non-text regions that were wrongly selected.
3. Consistency analysis for output: Regions representing the text content are enhanced to obtain a more useful representation for an optical character recognition.

D. Edge Based Techniques

Edges are the reliable feature of text in spite of color, intensity, orientations, layout, etc. Edge-based text extraction method is a general-purpose algorithm, which can quickly and efficiently localize and extract the text region from different kinds of images [12]. The edge density is calculated based on the average edge strength of boundaries of the connected component within a window. Taking into consideration of effectiveness and efficiency, four orientation alterations such as 0° , 45° , 90° , 135° were used to calculate the variance of edge orientations, where 0° denotes horizontal edge direction, 90° denotes vertical edge direction, and 45° and 135° are the two diagonal edge directions. Edge detector step is carried out by using a multi-scale technique, where the multi-scale present in an image is produced by Gaussian pyramids after sequentially applying the low-pass filter and down-sample the original image in both vertical and horizontal directions [13].

III. PROPOSED METHOD

Mathematical morphology is a topological and geometrical based method for image analysis [14]. It provides powerful tools for identifying and extracting the geometrical structures and representing shapes in a lot of applications. Morphological based feature extraction methods have been efficiently applied for document analysis and character recognition. It is used to extract the text contrast features from the processed images or video frame. The text feature is invariant against various geometrical image changes like rotation, translation and scaling. Even after the variations in lighting condition the color of the text is also changed but the geometrical features still can be maintained. This technique works robustly under different image orientations and geometrical features.

The canny edge detector is an edge detection operator that uses a multi stage algorithm to detect a wide range of edges in images. The process of canny edge detection algorithms can be listed.

1. Apply Gaussian filters to smooth the image in order to remove the noise.
2. Find the intensity gradients of the image.
3. Apply non-maximum suppression to get rid of spurious to edge detection
4. Apply double threshold to determine potential edges
5. Track edges by hysteresis.

After application of non-maximum suppression, remaining edge pixels provide a more accurate representation of real edges in an image. However, some edges pixels remain that are caused by noise and color variation. It is essential to filter out edge pixels with weak gradient value and preserve edge pixels with a high gradient value. This is accomplished by selecting high and low threshold values. If an edge pixel's gradient value is higher than the high threshold value, it is marked as a strong edge pixel [15]. If an edge pixel's gradient value is smaller than the high threshold value and larger than the low threshold value, it is marked

as a weak edge pixel. The strong edge pixels should certainly be involved in the final edge image, as they are extracted from the true edges in the image. However, there will be some debate on the weak edge pixels, as these pixels can either be extracted from the true edge or the noise /color variations to achieve an accurate result, the weak edges caused by the latter reasons should be removed. Usually a weak edge pixel caused from the true edge will be connected to a strong edge pixel while noise response are unconnected to track the edge connection, blob analysis is applied by looking at a weak edge pixel and its 8-connected neighborhood pixels as long as there is one strong edge pixel that is involved in the blob, that weak edge point can be identified as one that should be preserved.



Fig.1 Region identification by edge detection

For region extraction we have employed canny edge detection, region filtering and finally stroke width technique to extract text regions from MSER [16].Canny edge detection is well known method to detect the wide range of edges in the image. We used canny to detect the edges of the text region in the image; in our input image we have text and some other uselesobjects.We want to detect the only text from that image the MSER identifies the regions and by using canny edge detector we detect the edges of the text regions only so we can eliminate the other things easily. Region filtering is used to identify the properties of different regions presented in the input image using the pixel values, by using those properties we can separate the image into sub images and get the text region image.

IV. EXPERIMENTAL RESULTS

Through visual analysis, it is clear that our method is successful along with segmentation; pixels of region are also identified [17-19]. Our method is simple and fast because of using basic image processing techniques also be helpful in other medical imaging applications, pattern matching, and feature extraction. The main drawback of our method is the manual selection of threshold parameter and size of averaging filter. The Region and boundary descriptors are used for quantitatively characterizing the abnormality from the digital images. Stroke width is useful discriminator for text in images, is the variation in stroke width within eachtextso it is useful to eliminate regions where the stroke width exhibits large variation[20].

Consider an image as shown in Fig. 2 contains some text in foreground and background contains some unwanted things. Every region can be filtered with different colors from 0 to 255. The main use of the canny is to remove on-text regions from the input image the canny edge detection we have to separate the letters from the background and many of the non-text regions have been separate from text. Now the filtering is to remove some of the connected components busing their region properties shown in Fig.3.



Fig. 2 Input Image

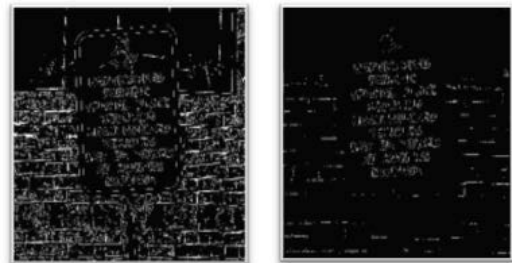


Fig.3 Canny Edges with MSER Regions



Fig.4. Text Before and after Region Filtering



Fig. 5 Visualization of Text with Stroke width

Once all the potential lines are detected, a procedure to apply a threshold is performed to obtain a possible line

separation in the text. Dynamically calculated and it is proportional to the average length of the lines in the text Y histogram values given by Fig.4. This process applied to the histogram aims to remove the regions in the histogram that are not referred to the lines in the text, or the elimination of noises with the text lines. Once the possible text line regions are separated by removing an offset from the histogram determine the average height of these regions to exclude false lines that might be detected text regions with stroke width given in Fig.5. The idea is to find the maximum area that each line might be inscribed, by determining the superior and inferior coordinates in the axis. Fig.3 shows the limits of these regions after the exclusion threshold is applied. The dashed lines are the limits between two adjacent line regions. In this way, the excluded regions are recovered. Note that the limit lines establish the maximum and minimum coordinates for each text line.

V. CONCLUSION

The images have been tested using various retrieval techniques. It has been observed that the threshold depends on a various parameters like the illumination condition reflections and the scan point spread function. This approach has used morphological clearing of images to reduce the number of false positive obtained. The proposed method significantly reduces the number of ambiguous feature matches and greatly improves the image matching performance if text characters by using stroke width along unstructured environment of the document images.

REFERENCES

- [1] F.Alaei, A.Alaei, and M. Blumenstein, *et al.*, "Document Image Retrieval Based On Texture Features and Similarity Fusion, International Conference on Image and Vision Computing", New Zealand, *IEEE*, pp. 1–6, 2017.
- [2] K.L. Bouman., G. Abdollahian, M. Boutin, and E.J. Delp, "A Low Complexity Sign Detection and Text Localization Method for Mobile Applications", *IEEE Transactions on Multimedia*, Vol. 13, No.5, pp. 922-934, 2011.
- [3] J. Ohya, A. Shio, and S. Akamatsu, "Recognizing characters in scene images", *IEEE Trans Pattern Anal Mach Intell*, Vol.16, No. 2, pp:214–220,1994.
- [4] Y. Zhong, K. Karu, AK. Jain, "Locating text in complex color images", *Proc Third International Conference Document Anal Recognition*, Vol.1, pp.146 - 9,1995.
- [5] K. Jung, "Neural network-based text location in color images", *Pattern Recognition Lett*, Vol. 22, No.14, pp.1503–15, 2001.
- [6] S. Messelodi, and C. Modena, "Automatic Identification and Skew Estimation of Text lines in Real Scene Images", *Pattern Recognition* Vol.32, No.5, pp.791–810,1999.
- [7] Y. Hasan, L. Karam, "Morphological Text Extraction from Images", *IEEE Trans Image Process* Vol. 9, No.11, pp. 1978–83,2000.
- [8] S. Janarthanam., *etal.*, "Active Salient Component Classifier System on Local Features for Image Retrieval", *Indian Journal of Science and Technology*, Vol. 10, No.26, 2017.
- [9] YumenYao, and Deyun Chen, "A Multiple Wavelet Capacitance Tomography Image Fusion Method", *Journal of Harbin Institute of Technology University*. Vol. 05, pp. 88-93, 2014.
- [10] FA. Mahdi, MFA.Fauzi and NN. Ahmad, "Image Retrieval Using Most Similar Highest Priority Principle based On Fusion Of Color And Texture Features", *Springer- Verlag*, pp. 765–770, 2012
- [11] X.Yin, X.in, K.Huang, and H.WeiHao, "Robust Text Detection in Natural Scene Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 36, No. 5, 2014.
- [12] Jia.Y. E. Shelhamer., J. Donahue., S. Karayev., J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolution architecture for fast feature embedding", *in Proceedings of the ACM International Conference*, 2014.
- [13] G. Aquino, and L.Lanzarini, "Keyword Identification in Spanish Documents Using Neural networks", *J. Comput. Sci. Technol.* Vol.15, No.2, pp. 55–60, 2015.
- [14] X. Bai, B. Shi, C. Zhang, X. Cai, and Qi L, "Text/Non-Text Image Classification In The Wild With Convolutional neural Networks", *Pattern Recogn*, Vol. 66, pp.437–446, 2017.
- [15] IN. Dutta, N. Chakraborty, AF. Mollah, S. Basu, and R. Sarkar, "Multi-lingual Text Localization from Camera Captured Images Based on Foreground Homogeneity Analysis, in Recent Developments in Machine Learning and Data Analytics", *Springer*, pp.149–158, 2019.
- [16] S. Paul, S. Saha, S. Basu, and M. Nasipuri, "Text Localization in Camera Captured Images Using Adaptive Stroke Filter. In Information Systems Design and Intelligent Applications", *Springer*, pp. 217–225. 2015.
- [17] X. Rong, C. Yi, and Y. Tian, "Unambiguous text localization and retrieval for cluttered scenes, in Computer Vision and Pattern Recognition (CVPR)", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.3279–3287, 2017.
- [18] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition", *IEEE Trans Pattern Anal Mach Intell*, Vol.39, No.11, pp. 2298–2304, 2017.
- [19] Y. Wei, Z. Zhang, W. Shen, D. Zeng, M.Fang, S.Zhou, "Text detection in scene images based on exhaustive segmentation", *Signal Process Image Commun.*, No.50, pp. 1–8, 2017.
- [20] X. Zhou, C. Yao, H. Wen, Y. Wang, S. Zhou, W. He, and J. Liang , "EAST: An Efficient and Accurate Scene Text detector", *In Proc. CVPR*, pp.2642–2651,2017.