ISSN: 2250–3676

[IJESAT] INTERNATIONAL JOURNAL OF ENGINEERING SCIENCE & ADVANCED TECHNOLOGY Volume-2, Issue-4, 961 – 964

A NOVEL METHOD FOR TEXT EXTRACTION

S.T.Deepa[1], Dr.S.P.Victor[2]

[1] *Department of Computer Science, Shri Shankarlal Sundarbai Shasun Jain College for Women,Tamil Nadu, India,* ***deepatheodore@gmail.com***

[2] *Department of Computer Science dept. St.Xavier's College, TamilNadu, India,* ***victorsp@rediffmail.com***

Abstract

Text detection and recognition in real images taken in unconstrained environments, such as street view images, remain surprisingly challenging in Computer Vision. Extraction of text and caption from images and videos is important and in great demand for video retrieval, annotation, indexing and content analysis. In this paper, we propose a text extraction algorithm using Dual Tree Complex transform. It is demonstrated that the proposed method achieved reasonable accuracy of the text extraction for moderately difficult examples.

Index Terms:*component. Scene text extraction, morphological dilation, binarization*

---------- *** ----------

1.INTRODUCTION

Text extraction from images and video sequences finds many useful applications in document processing and parts identification in industrial automation. Text extraction may employ binarization or directly process the original image. Text block filtering and integration firstly are used to obtain the clean background and clear text with high contrast for the effective recognition [1]. The character of the stroke like structure is adequately considered. Lastly the missed text pixels in former steps are recalled by reliable local region growing.

Text extraction is done by combining the adaptive binarization and perceptual color clustering method [2]. Adaptive binarization method can handle gradual illumination changes on character regions so it can extract whole character regions even though shadows and/or light variations affect the image quality. Perceptual color clustering method complementary can extract text regions which have similar color distances so that it can prevent the problem of the binarization method.

A novel method for overlay text detection and extraction from complex videos is proposed in [3]. The detection method is based on the observation that there exist transient color between inserted text and its adjacent background. The transition map is first generated based on logarithmial change of intensity and modified saturation. Linked mas are generated to make connected components for each candidate region and then each connected component is reshaped to have smooth boundaries.

Wavelet transform was used to detect edges in image [4]. The results are improved output by threshold most important edges of image and erasing some non-text lines. The gray level edge image aimed was converted to binary image with a global threshold. A new projection profile method was employed for estimating local of text.

An algorithm for detecting and tracking text in digital video is presented in [5]. The system implements a scale-space feature extractor that feeds an artificial neural processor to detect text blocks. The text tracking scheme consists of two modules: a sum of squared difference based module and contour based module to refine the position.

A system that reads the text encountered in natural scenes with the aim to provide assistance to the visually impaired persons is present in [6]. Automatic text recognition from natural images receives a growing attention because of potential applications in image retrieval, robotics and intelligent transport system. Camera based document analysis becomes a real possibility with the increasing resolution and availability of digital cameras. However in the case of a blind person, finding the text region is the first important problem that must be addressed because it cannot be assumed that the acquired image contains only characters.

Text localization and recognition in images is important for searching information in digital photo archives. A robust text localization approach is presented in [7] which can automatically detect horizontally aligned text with different sizes, fonts, colors and languages. A wavelet transform is applied to the image and the distribution of high frequency wavelet coefficients is considered to statistically characterize text and non-text areas. K-means algorithm is used to classify

text area in the image. The detected text areas undergo a projection analysis in order to refine. Their localization and a binary segmented text image is generated.

Despite advances in the archiving of digital video, we are still unable to efficiently search and retrieve the portions that interest us. An updated system for detection and extraction of unconstrained variety of text from general purpose video is present in [8]. The text detection results from a variety of methods are fused and each single text instance is segmented to enable if for OCR.

A non-parametric and unsupervised method of automatic threshold selection for picture segmentation is presented in [9]. An optimal threshold is selected by the discriminant criterion namely so as to maximise the separability of the resultant classes in gray level.The procedure is very simple, utilizing only the zeroth and the first order cumulative moments of the gray level histogram.

A detailed analysis of multilingual text characteristics including English and Chinese is performed. Based on the analysis [10] propose a comprehensive, efficient video text detection, localization and extracted method which emphasizes the multilingual capability over the whole processing. The method is also robust to various background complexities and text appearances. The text detection is carried out by edge detection, local thresholding and hysteresis edge recovery. The text extraction consists of adaptive thresholding, dam point labelling and inward filtering.

A robust text segmentation method in complex background is present in [11]. The method utilizes the K-means algorithm to decompose a detected text block into different binary image layer. An effective post processing is followed to eliminate background residues in each layer. A group of robust constraints to characterize general text regions based on color, edge and stroke thickness is proposed.

A novel method for embedded text segmentation is proposed in [12]. The basic idea is used on two properties of embedded text the color of text pixels is subject to Gaussian distribution, the local part and global part of embedded text shares the same color distribution. A two-step text segmentation approach: in the coarse segmentation step, a 1D Gaussian function is adopted to model the color distributions of text pixels.

he method proposed in [13] is novel as it departs from a strict feed forward pipeline and replaces it by a hypotheses verification framework simultaneously processing multiple text line hypothesis, uses synthetic fonts to train the algorithm eliminating the need for time consuming acquisition and labelling of real world training data and exploits maximally stable external regions which provides robustness to geometric and illumination conditions.

We propose the DT CWT as a useful front-end for many multi-dimensional signal analysis and reconstruction tasks

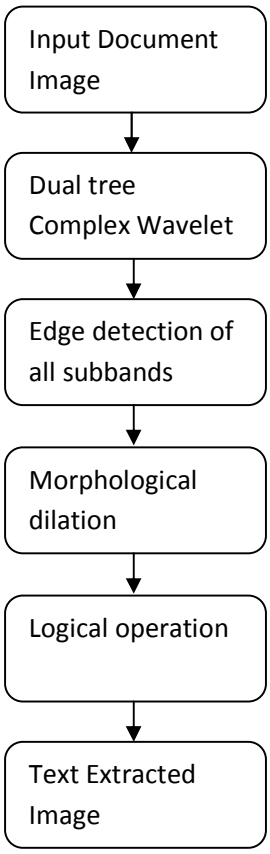

Figure 1. Flow Chart of the approach proposed

2.METHODOLOGY

The methodology proposed describes the text extraction using dual tree complex wavelet transform. Applying the Dual tree complex wavelet in the level of 3 into the input text image, we get the higher and lower subbands, in all the subbands detecting the edges.

A). Text Region Localization:

Morphological Dilation:

Here Detected edges of all subbands are dilated .The basic effect of the dilation operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (i.e. white pixels, typically) by adding pixels to the boundaries of the objects in an image. Thus areas of foreground pixels grow

in size while holes within those regions become smaller. The dilation operator takes two pieces of data as inputs. The first is the image which is to be dilated. The second is a (usually small) set of coordinate points known as a structuring element (also known as a kernel). It is this structuring element that determines the precise effect of the dilation on the input image. Various structuring elements have been experimented to find the suitable one & dilated images are produced using a disk shaped structuring element of 6 pixels radius. Here Dilation is performed to enlarge or group the identified text regions.

Figure 2: Original Image

B). Text Region Extraction:

All subband edges after dilation are combined with addition followed by AND operation which forms the text region as in eqn.1 below. The output of this logical operation is 1 only when all the inputs are 1. Then it is mapped to the original image to get text regions.

$$O_i = (W_B \cup S_p) \cap IP \qquad (1)$$

Text extraction includes many applications. One of the applications is sports video mining. Text can be extracted for content analysis and database retrieval. We can view the frames of our interest, say viewing the cricket from our desired over or wicket. In future we will do video mining using text extraction.

3.EXPERIMENTAL RESULTS

The performance of the proposed system is evaluated in this section. We first extract the lower and higher coefficients from the input text image. Then detect the edges for the all coefficients, in that edge detect image applying the morphological dilation image, this is for text localization. Then extract the text for applying the Logical AND operation and superimpose with the original image. Finally we get the text extracted image. The images are shown below

Figure 3. Image after text localization

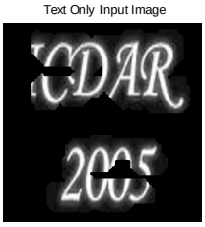

Figure 4. Image with text only using proposed method

Comparison with the existing methods

To compare the proposed text extractor to other systems, we evaluate our performances on the ICDAR 2005 dataset. It is composed of 499 color images, captured with digital cameras and resolutions, of book covers, road signs, household objects, posters etc. Some images are shown below.

The precision rate, recall rate and fmeasure are defined as
Precision rate (PR) = CD / (CD + FP) (2)
Recall rate (RR) = CD/ (CD + MC) (3)
Fmeasure(f) = (2*PR*RR) / (PR + RR) (4)

Where CD is correct character detected, FP is false position and MC is missed character.The performance results is shown in the following table.

System	Precision Rate PR)	Recall rate (RR)	Fmeasure (f)
Proposed system	0.71	0.62	0.66
Hinnerk Becker	0.62	0.67	0.64
Alex Chen	0.60	0.60	0.59

The criteria is also presented as graph

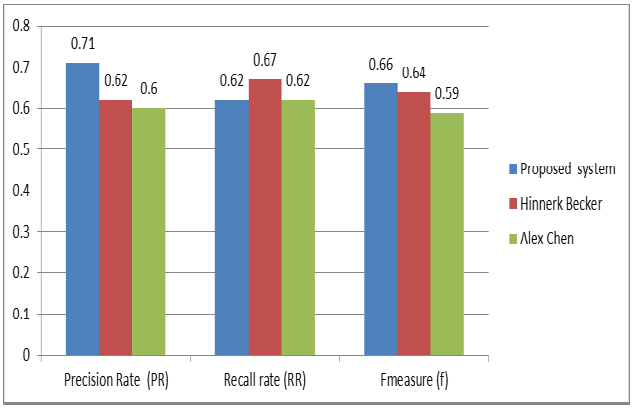

Figure 5: Graph For The Comparison Study

CONCLUSION

This paper presents a novel approach on text extraction. The method uses Dual Tree complex wavelet transform, morphological dilation and logical operation. It is robust against various conditions such as shadows, degradations, non-uniform illuminations, highlights, specular reflections, different font style and size and low contract images. The experiment result showed that the proposed method reasonably extract text regions with eliminating most non-text regions well.

REFERENCES

[1]. Li-Jie Li, Jin Li and Lei Wang, "An Integration Text Extraction Approach In Video Frame", Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010.

[2]. SeongHun Lee, JaeHyun Seok, KyungMin Min and JinHyung Kim, "Scene Text Extraction using Image Intensity and Color Information", IEEE Chinese Conference on Pattern Recognition, 2009.

[3]. Wonjun Kim and Changick Kim, "A New Approach for Overlay Text Detection and Extraction from Complex Video Scene", IEEE Transactions on Image processing, vol. 18, no. 2, February 2009.

[4]. Davod Zaravi, Habibrostami and Alireza Mallahzadeh, "A new algorithm for text localizing", 2010 2nd International conference on Computer Technology and development (ICCTD 2010)

[5]. H.Li, D.Doermann, O.Kia, "Automatic text detection and tracking in digital video", IEEE Trans. Image Processing, vol 9, no.1, pp.147 – 167, 2000.

[6]. N. Ezaki, M.Bulacu, L.Schomaker, "Text detection from natural scene images" towards a system for visually impaired persons", Proc. Of 17th Int. Conf. on Pattern Recognition, vol. II, pp. 683 – 686, 2004.

[7]. J.Gllavata, R.Ewerth, B.Freisleben, "Text Detection in images Based on Unsupervised Classification of High Frequency Wavelet Coefficients", Proc. Of 17th Int.Conf. on Pattern Recognition (ICPR), vol. 1, 1004, pp. 425 - 428.

[8]. S.Antani, D.Crandall and R.Kasturi, "Robust extraction of text in video". Proc.Int.Conf.Pattern Recognition, Sep 2000, Vol 1.pp.831 - 834

[9]. N.Ostu, "A threshold selection method from gray-level histograms", IEEE transactions on systems, MAN and Cybernetics, Vol.smc – 9, No:1, January 1979.

[10]. Michael.R.Lyu, Jiquiang Song and Mincai, "A comprehensive method for multilingual video text detection, localization and extraction", IEEE transactions on circuits and systems for video technology, vol. 15, No.2, February 2005.

[11]. L.Fu, W.Wang and Y.Zhan, "A robust text segmentation approach in complex background based on multiple constraints", Lecture notes in Computer Science, Springer-Verlang, Berlin Heidelberg, Vol.3767, pp.594 – 605, 2005

[12]. X.Wang, L.Huang and C.Liu, "A novel method for embedded text segmentation based on stroke and color, 2011 International conference on Document analysis and recognition.

[13]. L.Neumann and J.Matas, "A method for text localization and recognition in real world images", 10th Asian Conference on Computer vision, Queenstown, New Zealand, Nov 8-12, 2010.

AUTHORS' PROFILE

Mrs.S.T.Deepa earned her MCA degree from University of Madras, Chennai. The Mother Teresa University awarded her M.Phil degree. She is the Head of the Department of Computer Science, Shri Shankarlal Sundarbai Shasun Jain College for Women, Chennai, Email:deepatheodore@gmail.com

Dr. S. P. Victor earned his M.C.A. degree from Bharathidasan University, Tiruchirappalli. The M. S. University, Tirunelveli, awarded him Ph.D. degree in Computer Science for his research in Parallel Algorithms. He is teaching Computer Science since 1988. He is the Head of the Department of Computer Science, and the Director of the Computer Science Research Centre, St. Xavier's college (Autonomous), Palayamkottai, Tirunelveli. The M.S. University, Tirunelveli, Bharathiar University, Coimbatore, and Mother Theresa University, Kodaikanal have recognized him as a research guide. So far two candidates have completed their Ph.D. degree under his guidance, and 12 candidates pursuing research. He has published several research papers in national and international journals. He has organized Conferences and Seminars at state and national level. E-mail: victorsp@rediffmail.com