Performance analysis of various wavelet filters for Gujarati text localization in images

Jagin. M. Patel[1] & Apurva. A. Desai[2]
[1] Asst. professor, [2] Head of department and professor
[1]. M.K. Institute of Computer Studies, Bharuch, Gujarat, India
[2]Department of Computer Science, Veer Narmad South Gujarat University, Surat, Gujarat, India

Received: April 01, 2019 Accepted: May 14, 2019

ABSTRACT: *This paper presents a comparative analysis of Wavelet-based Gujarati text detection from an image using wavelet filters namely Haar, Symlets, Daubechies, Bior and Coiflets. This system uses a discrete wavelet transform (DWT) method to detect Gujarati text from images. A set of rules and geometric properties have been devised on the connected component to localize the actual text regions. Finally, bounding box information is employed to localize the Gujarati text. The experiment on a variety of images shows that Haar wavelet filter can detect text regions effectively.*

Key Words: *Gujarati text localization, wavelet, DWT*

1. Introduction

Texts, which are appeared in images and videos, hold semantic data. So, textual image are good source of data for content-based indexing as well as retrieval. In recent years, research on the automatic text localization from images has increase. However, few criteria such as enormous varieties in textual fonts, sizes, styles; the little contrast between the text and non-text background, create problem in text identification.

In literature, Work is found for text discovery for a range of languages like English, Chinese[1,2], Gurmukhi [3], Japanese[4], Arabic[5,6], Bangla [7,8]. Uyghur[9,10]. Few studies set their concentration on Gujarati OCR [11, 12,13] also.

Towards text extraction for Gujarati language, the first and initial work for Gujarati language is found in [14,15], which has focused on segmentation and recognition of Gujarati printed numeral from the image. In this work, researcher used edge detection, dilation, and connected component analysis to detect Gujarati numeral from images. This methodology successfully worked on a variety of images like a textured, map, magazine cover pages, and noisy images.

Here we investigate the wavelet method DWT and various filters (Haar, Daubechies, Daubechies Least Asymmetric, Coiflet, and Bior) to determine Gujarati text.

1.1 Characteristics of Gujarati script

Gujarati is an Indo-Aryan language, one of the official languages of India. It has twelve vowels (known as Swar), thirty-six consonants (referred to as Vyanjan) called as basic symbols, and ten digits.

In addition to basic symbols, diacritics (dependent vowel modifier) are utilized to connect of vowels with the core consonants. The diacritics can be placed ahead of, after, over, or underneath the basic symbol. Gujarat script can be partitioned into three logical zones: Upper, Middle and Lower [17]. Spotting Gujarati text in complex background images becomes a difficult problem due to the distinctive features of the Gujarati text described in [16].

The rest of the paper is organized as follows. The characteristics of DWT and wavelets are described in section 2. The proposed approach is discussed in section 3. The dataset and analysis protocols are structured in section five. Experimental results and comparison with wavelets are presented in section 5 and the conclusion is provided in section 6.

2. Wavelets

Haar Wavelet

Haar wavelet is the most straightforward among every accessible wavelet. It is discontinuous and its coefficients are either 1 or -1. It utilizes just two scaling and wavelet work coefficients. It is proficient in computation than other wavelet since it calculates pairwise summation and differences of neighboring pixels.

Symlets Wavelet

Symlets is also known as the Daubechies least asymmetric wavelets. It is comparative as Daubechies wavelet with a special case that it has a minimal phase while the Daubechies wavelet has a maximal phase. Symlet is closely symmetric. The scaling filters used are close direct stage filter.

Daubechies Wavelet

Daubechies is conceptually more complex and normally has a higher calculation overhead. It uses overlapping windows, so the results reflect all changes between pixel intensities. It is computed the ensuing low/high pass filtered signal. Daubechies is smoother compare to the Haar wavelet.

Coiflets Wavelet

Coflets have better symmetry over the Daubechies wavelets. It permits a good estimation of polynomial function on different resolutions [18]. The Coif wavelet and the Daubechies wavelet are equally powerful in some feature.

Bior Wavelets

Biorthogonal wavelet is an invertible however not necessarily orthogonal. This wavelet permits more degree of freedom than orthogonal wavelets [19]. This wavelet has two scaling function, which may produce different multi-resolution analysis, and two different wavelet functions.

The Discrete Wavelet Transform (DWT)

Wavelet decomposition gives a very good approximation of images and carries out a multi-resolution analysis of an image [20]. In the field of image processing, the DWT can decompose an image into well-behaved subbands within the frequency domain. With use of low-pass with high-pass 2-D filter, DWT divides image into four subbands: one average subband (LL) and three detail subbands (LH, HL, HH) as shown in figure 1 and delineated as:

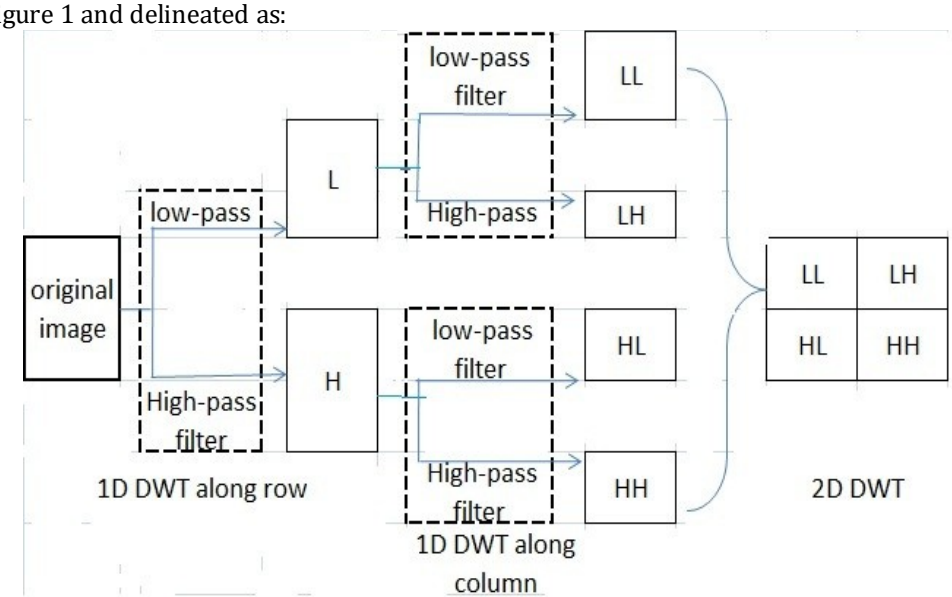

Fig-1: 2-D DWT decomposition scheme

1. LL subband centers on average components.
2. LH subband signifies horizontal edges data.
3. HL subband refers to vertical edges information.
4. HH subband concentrates on the diagonal edges data.

3. IMPLEMENTATION

Figure. 2 shows the procedure for wavelet-based text detection and steps are scribed below:

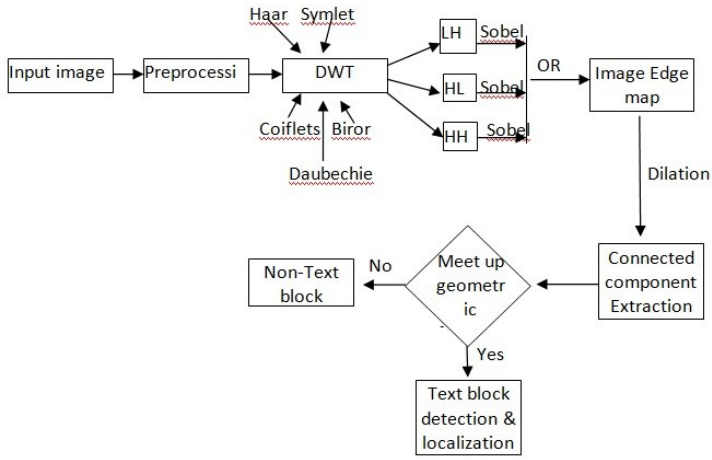

Figure 2: Text detection process

Step-1: Input colour image.
Step-2: Separate luminance component Y from the image using Eq. (1).

$$[Y] = [0.299 \quad 0.587 \quad 0.114] * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

Then median filter is applied on Y to reduce noise.
Step-3: Use various types of wavelet family with DWT. Here Haar, Symlets, Daubechies, Bior and Coiflets have been used.
Step-4: Obtain detailed features viz. LH, LL and HH of an image.
Step-5: Extract Gujarati candidate text edges by applying Sobel operator, Weighted 'OR' and dilation operation on LH, LL and HH.
Step-6: Connected component (CC) labeling is performed using method described in [21].
Step-7. For each CC, derive geometric features such as Height (H), Width (W), Aspect Ratio (AR), Size(S) = H *W, Area (A) = Total no. of foreground (ON) pixels.
Step 8: The false positives are eliminated by using several criteria derived based on geometric features of CC. Criteria are:

(i) S/A < T1
(ii) AR < T2
(iii) H < T3
(iv) S < T4
(v) W < T5

If the discovered text blocks meet any of the above mention heuristics then treat them as false positives and remaining blocks are considered as candidate components. In this study, values of thresholds T1, T2, T3, T4, and T5 are set experimentally which is 2.2, 0.5,25,310 and 40 respectively.
Step-9: The location of every candidate components is projected on the original image to get the corresponding text candidates.

4. DATASET AND EVALUATION PROTOCOL

4.1. Dataset

There are numerous public datasets widely used for performance assessment within the domain of text detection and/or recognition. Few available datasets are ICDAR, Oriented Scene Text Database (OSTD), MSRA-TD500, KAIST, CHAR74K etc. Most of the datasets contain text in English and Chinese language and there is no dataset available for Gujarati text.

As no dataset of images for Gujarati text recognition is available for this work, we have constructed a dataset consist of images which are collected from a various source, like book pages, magazines, newspapers, advertisements, logo etc. The texts may be in various fonts, sizes and colors. The quality of each image may vary. These images mostly center on horizontal text.

4.2 Evaluation protocol

The discovered text blocks in an image are marked by their bounding boxes. Here, performance is assessed at the block level rather than the word or character level. The following parameters are outlined for each discovered block:

1. True Positive (TP): It is a detected block that contains a partially or fully text. It may cover a single character or a whole word or a part of the word or a line.
2. False positive (FP): An identify block that does not have text.

Based on the number of blocks in these two categories, precision is defined as:

Precision = TP /(TP+FP)

5. Results and Discussion

The proposed algorithm is tested on over 800 images. These images contain texture, variations in text font, color, size. Figure. 3 shows the processed images by various wavelets. The detected text blocks in an image are represented by their bounding boxes with red colour.

Figure 3: Output of various wavelets

In this research, researchers have compared five wavelet families: Haar, Symlets (symN, N=1,4), Daubechies (dbN, N = 2,3,4), Biorthogonal (biorNN = 1_1,1_3,1_5,2_2), Coiflets (coifN, N = 1,2). For each wavelet basis, the number of decomposition levels two is considered.

The performance of the proposed technique with various wavelets is summarized in Table 1. From this table, it can be found that results of Haar orthogonal wavelet are superior to other wavelet, where precision is 0.92 for text detection.

Table 1. : Performance (%) of the proposed methods for various wavelet filters

Wavelet type	**Haar**	bior1_1	bior1_3	bior1_5	bior2_2	coif1	coif2	sym1	sym4	db2	db3	db4
p	**0.92**	0.87	0.9	0.89	0.74	0.84	0.76	0.74	0.78	0.77	0.78	0.74

6. Conclusion

In the present study, we proposed an algorithm for Gujarati text localization using 2D DWT wavelet features. Here, researchers have applied wavelet functions namely Haar, Daube, Symlets, Coiflets, and Biorthogonal in order to localize text of Gujarati. For the Gujarati text detection, this algorithm gives maximum success rate with text features which are extracted using Haar wavelet.

References

1. X. Jiang, J. Lian, Z. Xia, X. Feng and A. Hadid. 2016. Fast Chinese character detection from complex scenes. Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), Oulu, pp. 1-4.
2. X. Ren, K. Chen, X. Yang, Y. Zhou, J. He and J. Sun. 2015. A new unsupervised convolutional neural network model for Chinese scene text detection. IEEE China Summit and International Conference on Signal and

Information Processing (ChinaSIP), Chengdu, pp. 428-432..
3. D.V. Sharma and S. Singh. 2015. Gurmukhi text detection and localization in natural scene images. International Journal of Advanced Research in Computer Science and Software Engineering, 5(12): 442–458, 426-428, 2015
4. M. Iwamura, T. Kobayashi, and K. Kise. 2011. Recognition of multiple characters in a scene image using arrangement of local features. In Proc. IEEE International Conference on Document Analysis and Recognition, pp. 1409–1413.
5. H Gaddour, S Kanoun, and N Vincent. 2016. A New Method for Arabic Text Detection in Natural Scene Image Based on the Color Homogeneity. In International Conference on Image and Signal Processing, Springer International Publishing pp. 127-136.
6. O. Zayene, S. M. Touj, J. Hennebert, R. Ingold and N. E. Ben Amara. 2016. Data, protocol and algorithms for performance evaluation of text detection in Arabic news video. In Advanced Technologies for Signal and Image Processing (ATSIP), 2nd International Conference on, pp. 258-263.
7. Wang, Yaqi, Liangrui Peng, and Shengjin Wang, 2016. A Multi-stage Method for Chinese Text Detection in News Videos. Procedia Computer Science 96: 1409-1417.
8. Ghoshal, Ranjit, Anandarup Roy, and Swapan Kr Parui. 2011. Recognition of Bangla text from scene images through perspective correction. In Image Information Processing (ICIIP), International Conference on, pp. 1-6.
9. Y. Song et. al. 2017. Robust and parallel Uyghur text localization in complex background images. Machine Vision and Applications, 28 (7), pp. 755-769.
10. C. Yan, H. Xie, S. Liu, J. Yin, Y. Zhang and Q. Dai. 2018. Effective Uyghur Language Text Detection in Complex Background Images for Traffic Prompt Identification. In IEEE Transactions on Intelligent Transportation Systems, 19(1): 220-229.
11. Desai, A. A. 2015. Support vector machine for identification of handwritten Gujarati alphabets using hybrid feature space. CSI transactions on ICT, 2(4): 235-241
12. Desai, A. A. 2010. , "Gujarati handwritten numeral optical character recognition through neural network. Pattern Recognition, 43(7) : 2582–2589.
13. Desai, A. A. 2010. Handwritten Gujarati numeral optical character recognition using Hybrid feature extraction technique. In Proceeding of international conference on image processing, Computer vision and pattern recognition, (IPCV'10), Vol. 2, pp. 733–739
14. Patel, J. M. and Desai, A. A. 2012. A comparison of four edge detection methods for identifying Gujarati Numerals from images. VNSGU Journal of Science & Technology, 3(2): 113-124, 2012, ISSN: 0975-5446
15. Patel, J. M. and Desai, A. A. 2014. Segmentation and Recognition of printed Gujarati numeral from Image. International Journal of Engineering Research & Technology, 3(2).
16. Patel, J. M. and Desai A. A. 2018. Gujarati Text Localization, Extraction and Binarization from Images. International Journal of Computer Sciences and Engineering, 6(8):714-724.
17. Patel C. N. and Desai A. A. 2011. Zone identification for Gujarati handwritten words. In Proceedings of international conference on emerging applications of information technology (EAIT 2011), IEEEXplore, pp. 19–20.
18. Huang S.J. and Hsieh C.T. 2002. Coiflet wavelet transform applied to inspect power system disturbance generated signals. IEEE Transactions on Aerospace and Electronic Systems, 38, 1, 204–210.
19. Ali, Suhad A. and Ashwaq T. Hashim. 2016. Wavelet transform based technique for text image localization. Karbala International Journal of Modern Science 2, no. 2: 138-144.
20. Khobragade, Kavita Anandrao, and K. V. Kale. 2014. Multi-wavelet based feature extraction algorithm for iris recognition. International Journal of Engineering and Innovative Technology (IJEIT) Volume 3.
21. Haralick, Robert, M., and G. S. Linda. 1992. Computer and Robot Vision. Addison-Wesley Longman: Boston, MA, USA, pp. 28–48.