

# Text Localization in the Image of Complex Background using Discrete Wavelet Transform

S. Prabakaran\* and Maghna Luthra

Department of CSE, SRM University, Kattankulathur, Chennai - 603203, Tamil Nadu, India;  
Prabakaran.s@ktr.srmuniv.ac.in, meghna.limat@gmail.com

## Abstract

**Objectives:** To propose suitable method for localizing text in images. **Methods/Statistical Analysis:** Many methods of text localization use techniques based on region and textures using statistical approaches. In this paper, a domain transform based method has been proposed. The proposed approach based on wavelet transform utilizes the properties of wavelet in localizing the text. Particularly, it uses wavelet energy in differentiating text region and non text region in an image.

**Findings:** The proposed approach was applied on many bench marked images to localize the text and found effective in achieving its purpose. Many existing methods of text localization based on region and texture, need special software or complex calculations and methods. Some of these methods require the heuristic functions like threshold and other inputs. Also, some methods are suitable for small text while some of them are suitable for large text. The proposed wavelet based method for text localization works well for small as well large text. It doesn't need any special software. It doesn't have any input constraints like threshold. Also it is quite simple, robust and more efficient in comparison with the existing approaches. **Applications:** Localizing text in images contributes much in the extraction as well as retrieval of information from an image in the web. Also, it is an important step in recognizing the text present in the image.

**Keywords:** Text localization, Wavelet Transform, Wavelet Energy

## 1. Introduction

In modern age, internet is an unavoidable resource for day to day information in life. Mostly, multimedia sources like images contribute more as information resource. One of the high level semantic features in an image is the text. Localizing such text in images contributes much in the extraction as well as retrieval of information from an image in the web. Approaches used for localizing the text in image fall into two types based on the region and texture. First variety based on region utilizes the color based properties in the region of a text. It also utilizes the comparison of the color properties of the text region with that of the background. It also utilizes the building blocks of an image like color contour, edge etc. It identifies these elements and merges them to form boxes bounding text. The second variety of approaches based on textures<sup>1</sup> utilizes the concept that the texts have different texture in

images in comparison with the background of the image. There are many techniques to analyse the properties of texture of a text region. One of them is developing features<sup>2</sup> which result in false positives as region of images is to be detected. Some of them are blocks which are denser at edges with multi-line texts included in them. This approach was tested on a data set of different type of images. Precision and recall were the two metrics used to evaluate the performance of this system on the text box level. In another approach<sup>3</sup>, spatial variance for pixels using a local neighbourhood approach was computed using a window of  $1 \times 21$  size of the image. Then, a Canny edge detector was used to identify the edges. These edge components were merged to form a longer one. The upper and lower boundaries of a text line were identified by joining edges being in opposite directions. Horizontal components only were detected by this approach. In another work<sup>4</sup>, a hybrid approach for localizing the texts

\*Author for correspondence

in scenes using component information a CRF model was constructed for CC analysis. The model parameters were optimized using algorithms of MCE learning and graph cuts inference. The construction of a CRF model and the knowledge of relationship among the neighbouring components limited the use of this approach. There is some other approaches<sup>5-8</sup>. In analyzing the images for localizing the text, they utilize heuristic approach, segmenting regularities etc. All the methods involved complex steps and calculations and sometimes need preprocessing. Also, the need of correct heuristic functions and of choosing a suitable threshold is challenging one. Some of these methods involve complex segmentation approaches too. In this paper, a simple effective and efficient method based on discrete wavelet transform is discussed. It does not need any threshold or heuristic function or complex merits to be calculated.

## 1.1 Discrete Wavelet Transform

Projecting a data from its original domain into another domain sometimes helps in extracting information that is hidden. Wavelet transform is also one such domain transformation technique. It is a linear as well as lossless one<sup>9,10</sup>. It transforms the data into coefficients. Objects like images and audios may be represented mathematically for informativeness using wavelets. A function represents each wavelet. This function can be viewed as a combination of both high pass and low pass filters. While subjecting a signal to the wavelet transform, the signal is decomposed into groups of co-efficients. They are available in two scales: Coarse and fine. The first one projects global details of a signal and the second one captures a signal's local features. A signal is passed simultaneously through two filters: low pass and high pass, for doing discrete wavelet transform for it. During this step, the signal and the impulse response of the filter  $g$  are convoluted at each pass. Every time the signal passes through a filter, its frequency is halved. As per Nyquist's rule half of the samples may be discarded. For this, down sampling or decimation by factor 2 is done. The process of wavelet transform for a signal may be represented by two equations as follows:

$$y_{\text{low}}(n) = \sum_{k=-\infty}^{k=+\infty} x(k) \cdot g(2n-k) \quad (1)$$

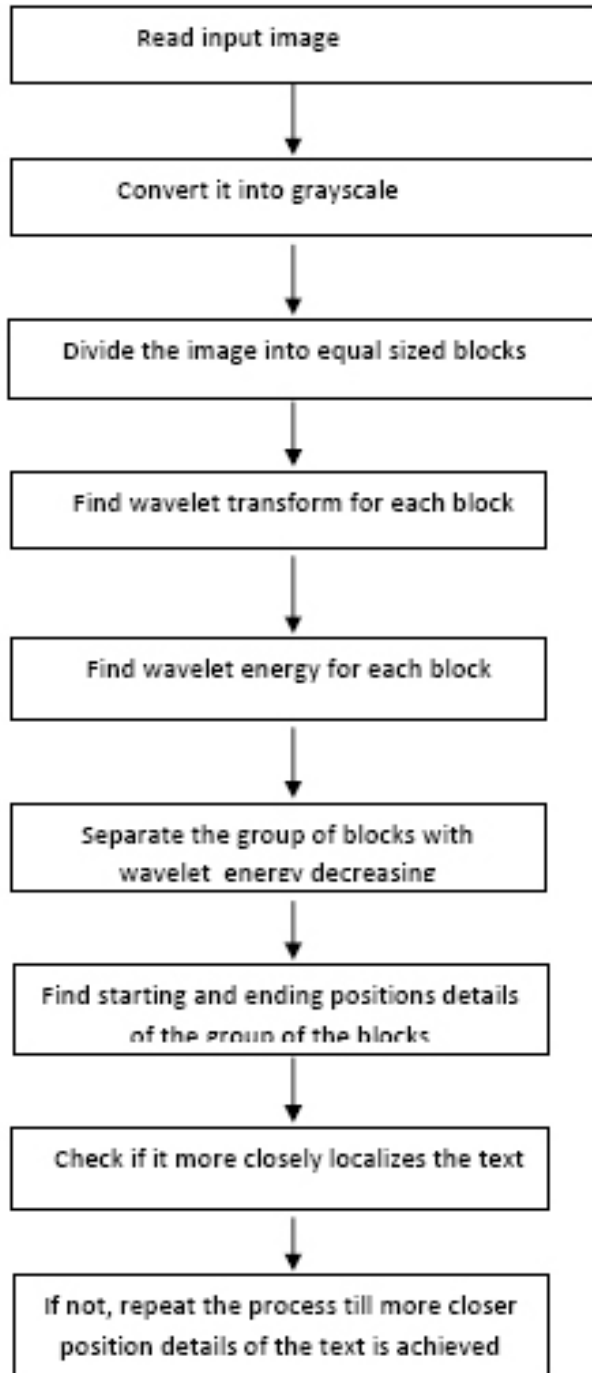
$$y_{\text{high}}(n) = \sum_{k=-\infty}^{k=+\infty} x(k) \cdot h(2n-k) \quad (2)$$

The detailed coefficients,  $y_{\text{high}}(n)$  and  $y_{\text{low}}(n)$ , the approximation co-coefficients occur in order in the resulting wavelet co-efficient vector. The process continues recursively till the number of approximation coefficients, called 'resolution' as 1. The mathematical representation of signals using wavelet is informative as well as economical. Also, many software packages are easily available for accessing wavelets. The wavelets can spatially adapt to the features, such as discontinuities, of a function. They also work well in a situation of varying frequency behavior. The multi resolution analysis of the wavelet facilitates the analysis of local behavior of a particular region of a data. We may refer more about wavelets in<sup>11-14</sup>. Some of the interesting applications of the wavelet in various fields may be referred at in<sup>15-18</sup>. The sum of the squares of the wavelet coefficients is called as wavelet energy. The energy of an image spectrum in the wavelet domain is distributed among various components. A portion of the energy is distributed in detailed coefficients while the remaining portion is distributed in the approximation coefficients.

## 2. Methodology

The method discussed in the present work is based on discrete wavelet transform. In this work, the image in which the text has to be localized is resized into the standard size set to work upon. The image is then converted to gray scale. It is then divided into number of blocks of equal sizes. The wavelet transform is applied on each and every blocks and the corresponding percentage of wavelet energies among various components were calculated. It was observed that the blocks of the image containing the text portions behaved differently from others. The behavioral pattern thus identified was used in localizing the text in test images. The process was repeated by dividing the blocks identified as containing text to get finer position details of the text in the image. It was observed that the group of blocks containing text area contained decreasing wavelet energy. The proposed method is depicted in Figure 1. The proposed method of text localization could also be summarised as follows:

- Step 1: Read the input image and convert it into greyscale image
- Step 2: Divide the image into equal sized blocks
- Step 3: Find the group of blocks with wavelet energy continuously decreasing.



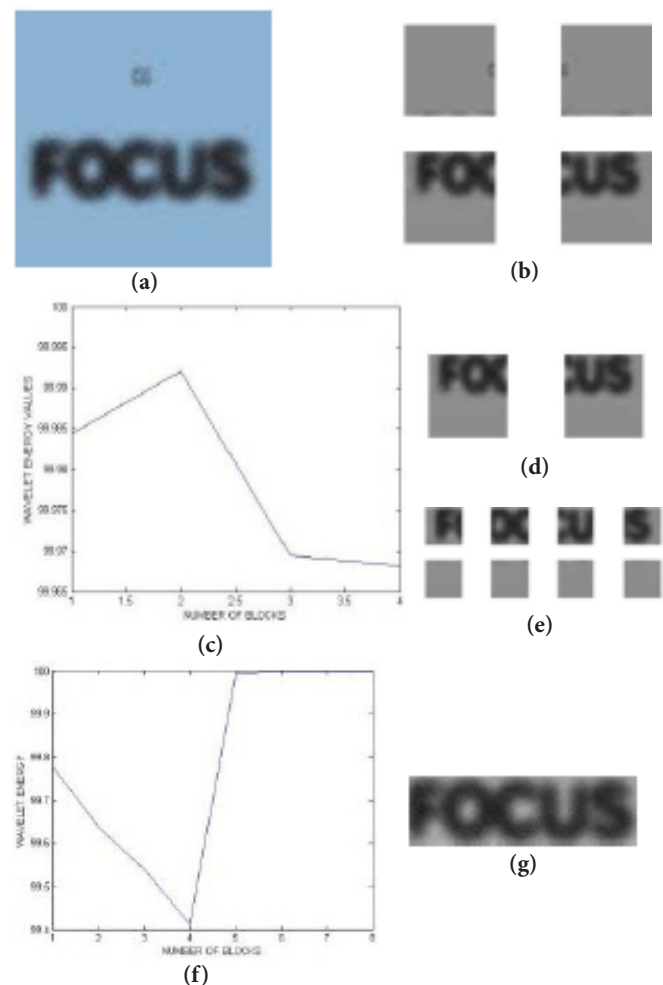
**Figure 1.** The proposed method.

Step 4: Separate the blocks identified in the previous step and identify their position or location details.

Step 5: If the text contained in these blocks are exactly localized in terms of their position, then stop. Else, repeat the steps 1 to 5 by taking the blocks identified in step 4 till a more closure position details about the text present in the image is found

### 3. Result and Analysis

The proposed method was tested with the six images with text embedded in them. Each image was divided into blocks and the blocks, selected as per the proposed algorithm, were depicted in the following Figure. These selected groups of blocks were further subjected to the proposed method and more closure details about the text could be localised. This could be repeated till the text could be localised more closely. The origin for pixel positions were fixed at left bottom side of the image. The results of applying the proposed algorithm on the image 'focus.jpg' is summarised in Figure 2. From Figure 2 (a) to Figure 2(g), it may be observed that the text in the



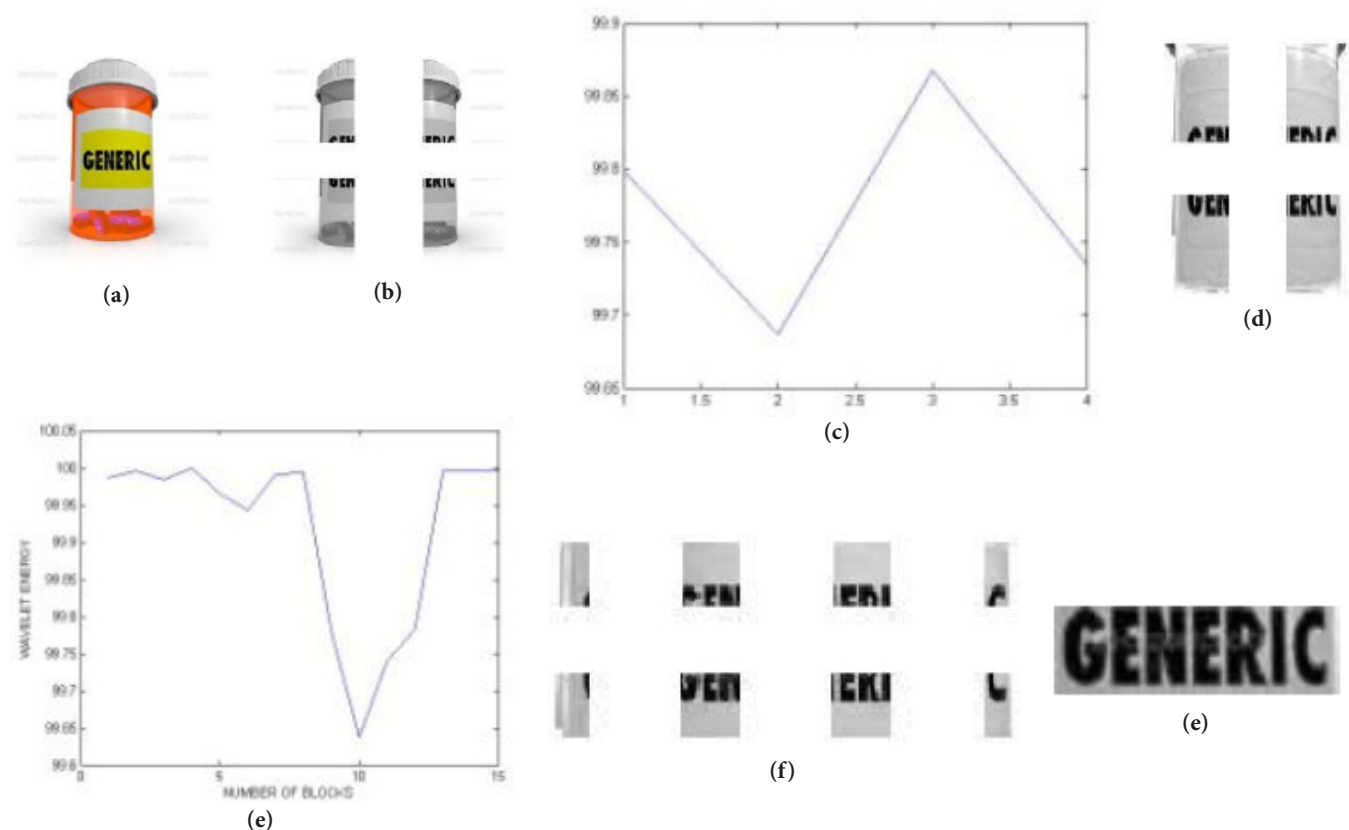
**Figure 2.** Text localisation for Focus.jpg using the proposed algorithm. (a) Original image (b) Image into blocks (c) Wavelet energy of the image (d) Selected blocks according to the proposed (e) Blocks in (d) further approach using the graph in (c) divided (f) Wavelet energy of the blocks in (e) (g) localized text in the image selected using (f)

focus.jpg was effectively localised and the final positions of the text could be seen in Figure 2(g). This result is presented in Table 1. By applying the proposed algorithm in vertical direction of the image, the positions of the text could be further refined. The proposed algorithm was then administered on the image 'medicinebottle.jpg'. The

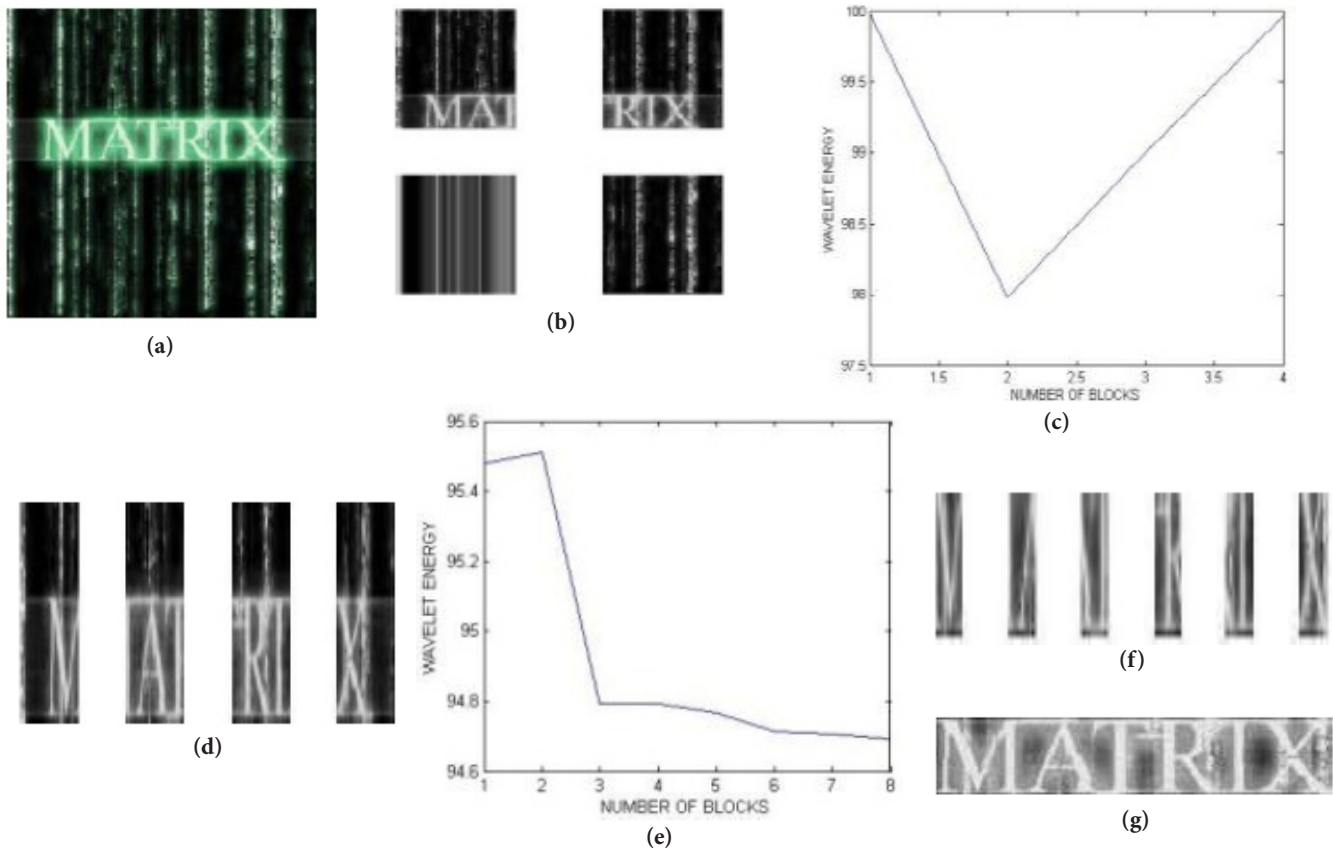
**Table 1.** Identified positions of the text in the image through the proposed algorithm

S.No	Image name	Text localized at the co-ordinates	
		Iteration 1	Iteration2
1	Focus.jpg	(0,0),(0,128), (128,256),(256,0)	(0,64),(0,128), (128,256),(64,256)
2	Medicine bottle.jpg	(0,0),(0,256), (256,256),(256,0)	(64,64),(64,192), (192,64),(192,192)
3	Matrix.jpg	(0,128),(0,256), (256,256),(256,128)	(0,128),(0,192), (256,192),(256,128)
4	ae.jpg	(128,0),(256,0), (128,128),(256,128)	(64,0),(64,128), (256,0),(256,128)
5	id.jpg	(128,0),(256,0), (128,128),(256,128)	(64,0),(64,128), (256,0),(256,128)

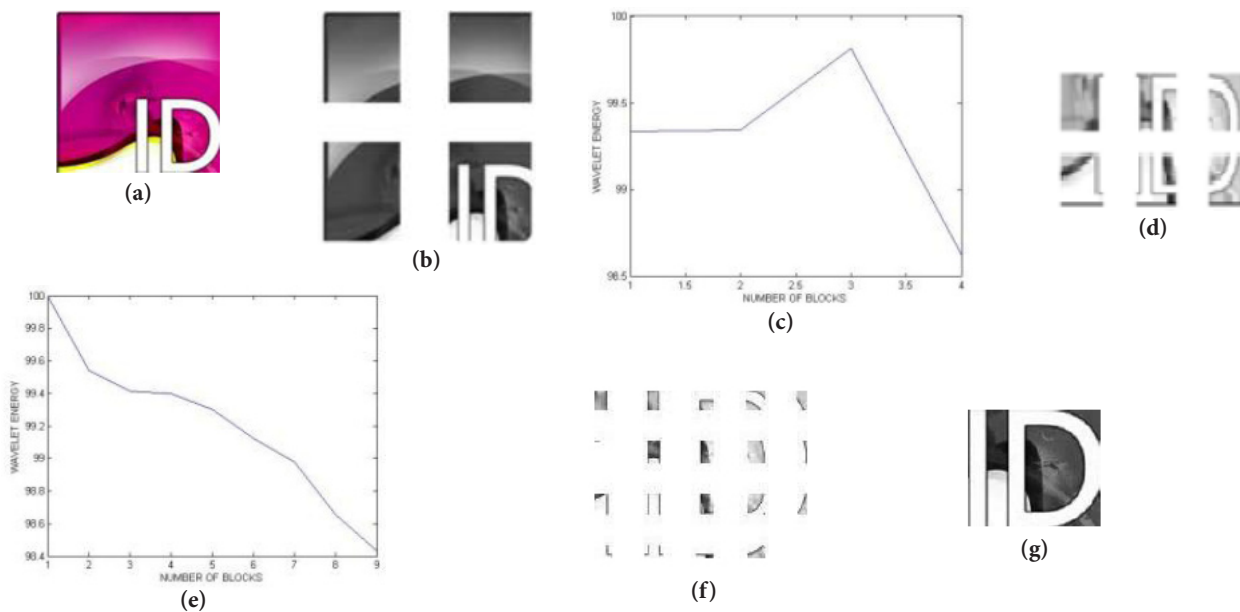
corresponding result is summarized in Figure 3. From Figure 3 (a) to Figure 3(g), it may be observed that the text in the medicinebottle.jpg was effectively localised and the final positions of the text could be seen in Figure 3(g). This result is presented in Table 1. By applying the proposed algorithm in vertical direction of the image, the positions of the text could be further refined. The results of applying the proposed algorithm on the image 'matrix.jpg' is summarized in Figure 4. From Figure 4 (a) to Figure 4(g), it may be observed that the text in the medicinebottle.jpg was effectively localised and the final positions of the text could be seen in Figure 4(g). This result is presented in Table 1. By applying the proposed algorithm in vertical direction of the image, the positions of the text could be further refined. The results of applying the proposed algorithm on the image 'id.jpg' is summarized in Figure 5. From Figure 5 (a) to Figure 5(g), it may be observed that the text in the id.jpg was effectively localised and the final positions of the text could be seen in Figure 5(g). This result is presented in Table 1. By applying the proposed algorithm in vertical direction



**Figure 3.** Text localisation for Medicinebottle.jpg using the proposed algorithm. (a) Original image (b) image into blocks (c) Wavelet energy of the image (d) selected blocks according to (c) divided further (e) wavelet energy of (d) (f) Selected blocks according to (e) are divided further (g) localized text in the image

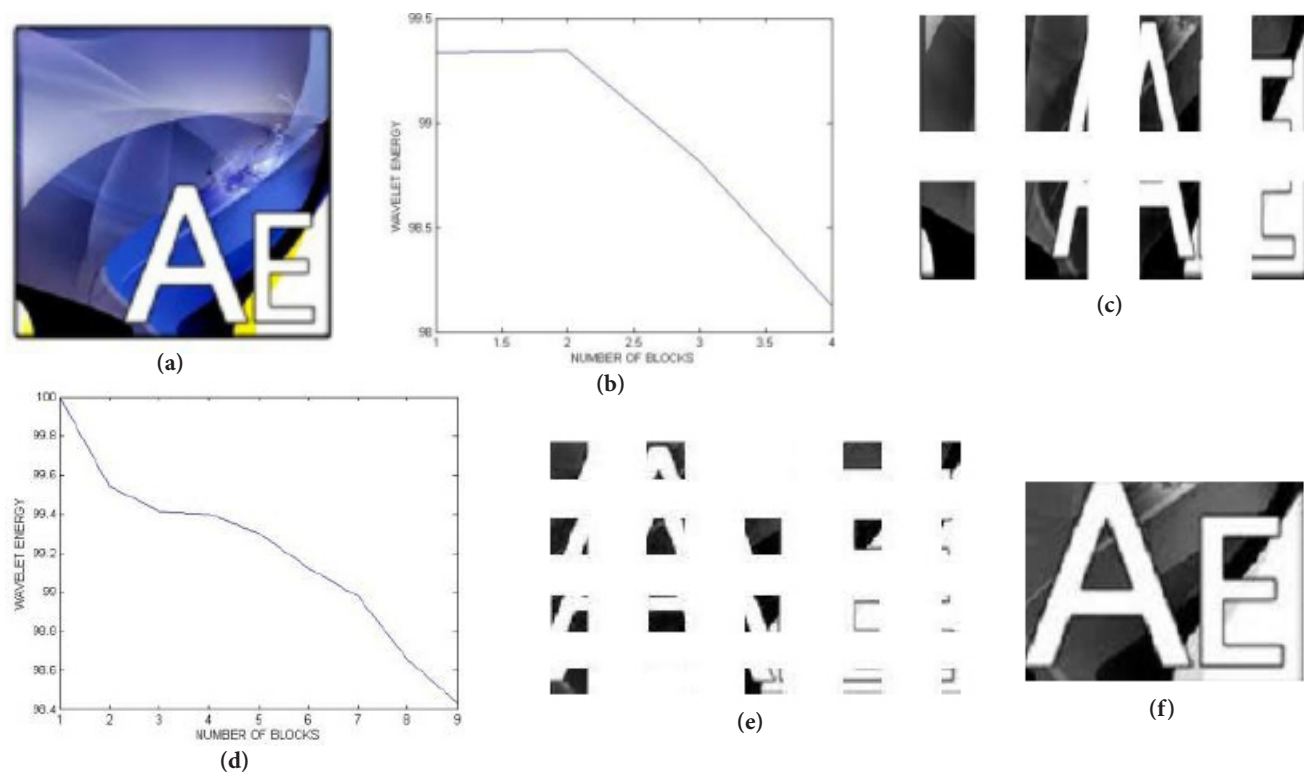


**Figure 4.** Text localisation for Matrix.jpg using the proposed algorithm. (a) Original image (b) Image into blocks (c) Wavelet energy of the image (d) Selected blocks according to (c) are divided further (e) Wavelet energy of (d) (f) Selected blocks from (e) are divided further (g) localized text in the image



**Figure 5.** Text localisation for id.jpg using the proposed algorithm. (a) Original image (b) image into blocks (c) wavelet energy of the blocks in (b) (d) Selected blocks using (c) are divided further (e) Wavelet energy of (d) (f) Selected blocks from (e) are divided further (g) localized text in the image





**Figure 6.** Text localisation for ae.jpg using the proposed algorithm. (a) Original image (b) wavelet energy of the four blocks of the divided image (a) (c) Selected blocks using (b) are divided further (d) Wavelet energy of (c) (e) Selected blocks using (d) are divided further (f) localized text in the image

of the image, the positions of the text could be further refined. The results of applying the proposed algorithm on the image 'ae.jpg' is summarized in Figure 6. From Figure 6 (a) to Figure 6(f), it may be observed that the text in the ae.jpg was effectively localised and the final positions of the text could be seen in Figure 6(g). This result is presented in Table 1. By applying the proposed algorithm in vertical direction of the image, the positions of the text could be further refined.

### 3.1 Advantages of the Proposed Algorithm

Unlike other existing approaches like morphological and texture based approaches, the present method does not involve complex calculations or special software. It also does not need to identify image sub structures like edges or color contours as in region based approaches. Also, the limitations like the need of correct heuristic functions and the need of choosing a suitable threshold in the existing approaches are not present in the present approach. No direct physical quantities need not be analysed in the present approach. Hence, the present approach is more robust since it is free from the need of analysis of existing physical

quantities in the image which may affect the performance of other existing approaches. It uses the wavelet energy which is a hidden property of the image in the wavelet domain. Also, it is more robust since wavelet transform automatically smoothens the data and the wavelet energy depends on the image nature. By administering the proposed method both in vertical and horizontal directions, more closer or exact position of the text could be localized. Hence, in comparison to the earlier methods, the present method is more robust, faster and efficient.

## 4. Conclusion

Text localization in the Images is important in many applications like content-based image analysis and extraction of information and its retrieval in web image. In this paper, a wavelet transform based approach for text localization was proposed. The proposed algorithm was tested on many bench marked image datasets. The results were encouraging despite its simplicity. In comparison to the earlier methods, the proposed method is faster, more robust and efficient mainly due to the use of the

wavelet based hidden properties instead of direct physical parameters. By administering the proposed method both in vertical directions, closer or exact position of the text could be localized.

## 5. References

1. Hanif SM, Lionel Prevost L. Texture based text detection in natural scene images: a help to blind and visually impaired persons. In: Hersh MA editor. Conference and Workshop on Assistive Technologies for People with Vision and Hearing Impairments Assistive Technology for All Ages CVHI 2007. 2007 Jan. p.1–6.
2. Situ L, Liu R, Lim Tan C. Text localization in web images using probabilistic candidate selection model. Proceedings of International Conference on Document Analysis and Recognition. 2011 Sep. p. 1359–63.
3. Yu Z, Karu K, Jain AK. Locating Text in Complex Color Images. Proceedings of the Third International Conference on Document Analysis and Recognition. 1995 Aug. p. 146–9.
4. Pan Y, Hou X, Lin Liu C. Text Localization in Natural Scene Images based on Conditional Random Field. Proceedings of 10th International Conference on Document Analysis and Recognition. 2009 Jul. p. 6–10.
5. Kennedy Babu N. A Novel Approach for Segmentation of Text. International Journal of Imaging Science and Engineering. 2008 Jan; 2(1):139–41.
6. Ma Y, Wang C, Xiao B, Dai R. Usage-Oriented Performance Evaluation for Text Localization Algorithms. Proceedings of 9th International Conference on Document Analysis and Recognition. 2007 Sep. p. 1033–7.
7. Gllavata J, Ewerth R, Freisleban B. A Text Detection, Localization and Segmentation System for OCR in Images. Proceeding of IEEE Sixth International Symposium on Multimedia Software Engineering. 2004 Dec; 310–7.
8. Aghajari G, Shanbehzadeh J. A Text Localization Algorithm in Color Image via New Projection Profile. Proceeding of the International Multi Conference of Engineers and Computer Scientists. 2010 Mar. p.1–4.
9. Aldroubi A, Unser M. Wavelets in Medicine and Biology. CRC Press, Boca Raton. 1996
10. Zhou X, Wang X, Dougherty E R. Nonlinear probit gene selection and wavelet based feature selection. Journal of Biological Systems. 2004 Sep; 12(3):371–86.
11. Strang G. Wavelets and dilation equations: a brief introduction. Society for Industrial and Applied Mathematics Review. 1989 Dec; 31(4):614–27.
12. Daubechies. Ten Lectures on Wavelets. Capital City Press, Montpelier, Vermont. 1992.
13. Mallat S. A Wavelet Tour of Signal Processing. Academic Press, San Diego. 1998.
14. Chui CK. An Introduction to Wavelets. Academic Press, Boston MA. 1992.
15. Mercy Theresa M, Subbiah Bharathi V. CAD for lung nodule detection in chest radiography using complex wavelet transform and shearlet transform features. Indian Journal of Science and Technology. 2016 Jan; 9(1):1–12
16. Roodaki M, Jafari Behbahani Z. Two-dimensional legendre wavelets and their applications to integral equations. Indian Journal of Science and Technology. 2015 Jan; 8(2):157–64.
17. Parthasarathy MB, Srinivasan B. Increased Security in Image Cryptography using Wavelet Transforms. Indian Journal of Science and Technology. 2015 Jun; 8(12):1–8.
18. Somasoundaram T, Subramaniam NP. High Performance Angiogram Sequence Compression using 2D Bi-orthogonal Multiwavelet and Hybrid SPIHT-DEFLATE Algorithm. Indian Journal of Science and Technology. 2015 Jul; 8(14):1–6.