



# Extraction of Hidden Text from Images using DWT

<sup>1</sup>V. Beslin Geo, <sup>2</sup>K. Sakthidasan @ Sankaran, <sup>3</sup>P. Archana, <sup>4</sup>S.M. Umarani

<sup>1,3,4</sup>Assistant Professor, <sup>2</sup>Associate Professor,

<sup>1,2,3,4</sup>Department of Electronics and Communication Engineering,

Hindustan Institute of Technology and Science, Chennai – 603103, India.

\*Corresponding author E-mail: [sakthidasan.sankaran@gmail.com](mailto:sakthidasan.sankaran@gmail.com)

## Abstract

Extraction of hidden text in web images, computer screen images, news, games and e-learning is a very important task in image processing. Compression of digital images leads to poor visual quality of background and text images. Digital images are significantly considered and segmented using DWT into text and background blocks. Huffman coding is used to perform the lossless compression process in the compressed text pixels and the SPIHT algorithm is employed to compress the background pixels. The result of DWT segmentation shows fringes in the segmented text image. The proposed method uses connected region and edge detection approach which provides a segmented text from digital video stills. The segmented text is converted to binary image using luminance thresholding which leads to fine quality of extracted text.

**Keywords**—Compression of Digital images; Filters; Text segmentation; Thresholding.

## 1. Introduction

The combination of text, graphics and natural pictures is called as compound images. Text in the image provides more information in web pages and e-learning. Compression of images leads to blurriness. Compression algorithm is classified as lossless and lossy techniques. The main motive is to reduce the visual quality loss during compression. Henceforth, lossless compression technique is used for compressing text in an image. The following three steps are involved for the compression of the compound images,

1. The images are segmented into regions of homogeneous data types and grouped together.
2. Selecting the compression algorithm for each data types.
3. Segmentation technique not only provides low bit rate but also lowers the distortion levels of the compressed image.

Segmentation algorithms are primarily based on three approaches namely layer-based segmentation, object-based segmentation and block-based segmentation. The common method used for block-based segmentation is DWT based compression as shown in Fig 2. The images are segmented into text and blocks of pixels using the geometric structures of each DWT transformed image block. Irrespective of font style, background complexity and style, the DWT method minimizes the misclassification error. The proposed method uses object-based segmentation where uses split and merge algorithm and edge detection approach for video still pictures. This proposed method results in reduced noise in segmented text from the video stills.

## 2. Related Works

Xiao chuan Chen et al. [11] applied statistical study of empirical matrix (EM) in an image to identify the hidden data.

Using EM projection histogram, PH features like moments of PH and 53 characteristic functions are extracted. To boost the performance, PE [prediction-error] image features are pulled out and incorporated. SVM is utilized as a classifier. Yuan Liu et al. [12] recommended 3 methods to derive the feature vector such as pixel domain based Robert gradient energy; DCT domain based variance of Laplacian parameter and extracted wavelet coefficients based higher-order statistics. BPA neural network is used as a classifier.

Xiang yang Luo et al. [13] applied WPT to decompose picture into 3 scales and to get the sub band coefficients. From the sub bands Multi-order absolute characteristic function of histogram features are pulled out. Lastly vector for each picture, all the features are standardized also merged to a 255-D feature. Back propagation neural set-up is utilized as a classifier.

Yuan-luTu et al. [14] designed a technique to extract the features using image luminance and chrominance components. The extracted features are found to be in DCT and DWT domains. Wavelet high-order statistics is used instead of wavelet characteristic function. Non linear SVM is employed as a classifier. Jing-Qu Lin et al. [15] intended Binary Similarity Method (BSM) for JPEG images to capture the 7th and 8th bit planes of non-zero DCT coefficients and to calculate the 14 features of every image. SVM is exploited as a classifier.

Zhi-Min He et al. [16] utilized RBFNN for step analysis. DCT and Markov features are found to be as inputs of neural networks. Sheikhan et al. [17] designed a technique to extract the features from sub band images and Contour let coefficients. 54 features are reduced using Analysis of Variance (ANOVA) method. Non linear SVM is found to be a classifier. Lie et al [18] recommended the gradient energy and statistical variance as two features to detect the presence of hidden content in spatial or DCT domain. Shi et al. [19] suggested a technique that uses statistical characteristic function of the prediction-error image, the test image, and its wavelet sub bands as special features.

### 3. Dwt Using Block-Based Segmentation

#### A. System

The block-based segmentation using DWT consist is processed in two stages: coding and segmentation. The process is as follows,

1. The image pixels are segmented into blocks of text and background parts.
2. Compress the texts blocks using lossless Huffman coding.
3. Then compress the background blocks with lossy SPIHT algorithm.

#### B. Text segmentation using Two-dimensional DWT

The Two Dimensional DWT investigates an image to spate horizontal, vertical and diagonal information. There are two types of stages and they are,

1. High pass and low pass filters are used for filtering image of rows  $N \times N$
2. One dimensional convolution filters is applied to the filtered image.

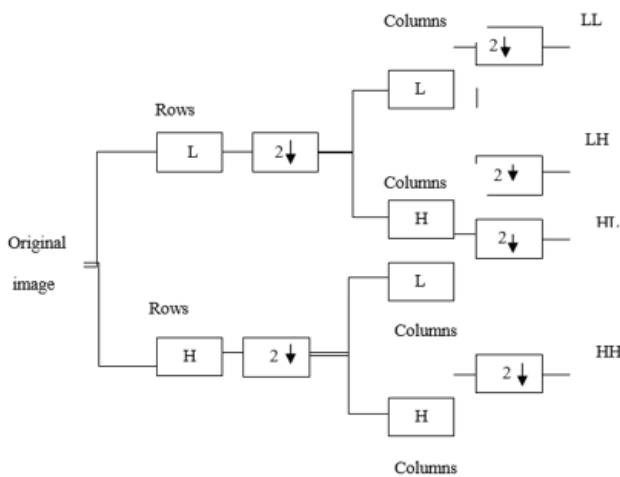


Fig. 1. DWT decomposition using Single-level two-dimensional.

An  $(N/2) \times (N/2)$  sub-image is produced by the Two Dimensional DWT. These sub images has four subbands namely low-pass (LL), Diagonal(HH), Vertical (LH and Horizontal (HL). This is illustrated in Fig 1

To reconstruct the photo from sub-images (LL, LH, HL and HH) of two dimensional DWT, details are recollected by means of up-sampling and convolution with the corresponding fusion filters and low-pass estimate. Up-sampling illustrates the insertion of a zero row after every available zero column or row.

#### C. Process of text extraction

The multi-resolution attributes of the wavelet transform can decay the signal to sub-bands at numerous scales. The edges are enhanced in all aspects saying vertical, horizontal and diagonal using sub-bands (LH, HL, HH). The text that are present in vertical, horizontal or diagonal directions are characterised highly by the high contrast edges. Based on this a high clarification text block is obtained by the high-frequency wavelet coefficients.

#### D. Algorithm of DWT block-based text extraction

1. The image from the database is equally divided into blocks.
2. Set threshold for sub-bands as  $T_1, T_2, T_3$  respectively.
3. Repeat steps 4 through 5 for all input image blocks.
4. In order to decompose into sub-bands DWT is applied to each block
5. The high-frequency sub-bands mean and standard deviation are calculated for each  $8 \times 8$  block using the below mentioned formulas.

$$\mu_{\theta} = \sqrt{\frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H [I_{\theta}(i, j)]^2} \quad (1)$$

$$\sigma_{\theta} = \sqrt{\frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H [I_{\theta}(i, j) - \mu_{\theta}]^2} \quad (2)$$

If the standard deviation ( $\sigma$ ) of high frequency sub-band value is greater than the threshold value, then the corresponding image block is considered as text block. Or else, the corresponding image block is considered as background block.

$$type(B_i) = \begin{cases} 0, & \text{text blocks if } \sigma_{LH} \geq T_1 \vee \sigma_{HL} \geq T_2 \vee \sigma_{HH} \geq T_3 \\ 1, & \text{background blocks otherwise} \end{cases} \quad (3)$$

Where,

$\theta \in \{LH, HL, HH\}$  and  $I_{\theta}(i, j)$  are belong to the wavelet coefficients of high-frequency sub-band  $\theta$ ,  
 $W$  is high-frequency sub-band width,  
 $H$  is high frequency sub-band height,

$B_i$  is  $i^{th}$  block in the image and  $\vee$  is logical OR operator,

The sub bands (HL, LH and HH) which counterparts to slant, horizontal and vertical whacks of English literacies have an important statistic standard deviation boundary as  $T_1, T_2$  and  $T_3$ . From the text blocks of different images, the standard deviation is plotted and the thresholds  $T_1, T_2$  and  $T_3$  are calculated.

Fig. 2. Flowchart of text segmentation.

#### E. SPIHT algorithm of background pixels

The background pixels, pictures and text are found in the compound images. To keep the greater subsets with unrelated measurements together, it is necessary to partition the tree which is done by the SPHIT Procedure. The results obtained from these partitioning appear in binary format. Providing a substantial plot encrypting these binaries are transmitted to the decoder. The SPHIT Algorithm sends the binary representation of the numeral significance of the wave coefficient.

Instigating from the bottom resolution band, these wavelet coefficients are divided once more. The divided coefficients are then combined into an array of  $2 \times 2$ , which produces the band 1 coefficients. The bottom determination band coefficients are also divided into an array of  $2 \times 2$

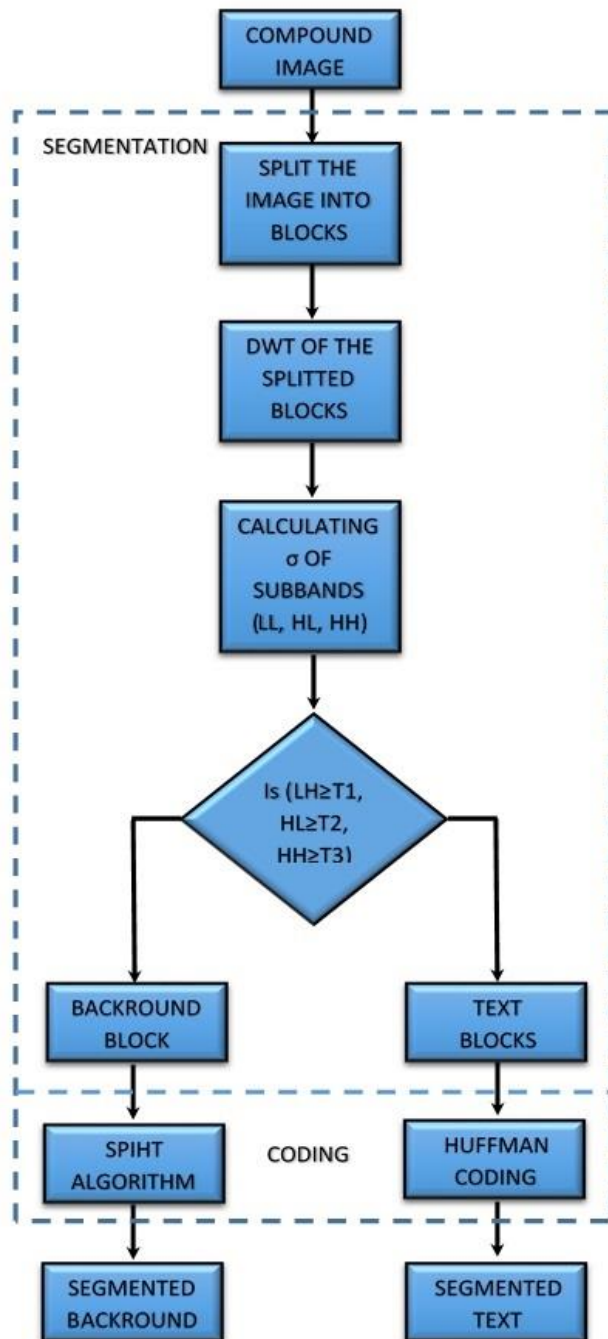
#### F. Lossless coding of text pixels

From the early encoded image in the horizontal direction, an optimal prediction of text blocks is discovered by the prediction method. There are two modes in the method

1. Find out the finest counterpart among the current block and its instantaneous left most blocks in the same row. This process has to be carried for each txt block that is considered
2. For an optimal prediction of orientation block, a lowest block dissimilarity in chosen

Before sending information of each block to the decoder, encoding is done in the prediction residual by Huffman encoder

### 4. Proposed Method



Digital video has become one of the most important in education, news and games. To extract information from a huge amount of video data, we need to extract text from video and recognize the extracted text. Text is important part in video. Fig. 4 shows the flow of the proposed system.

#### A. Text extraction

Text extraction in video consists in three steps,

1. Find text region in original images.
2. The text is separated from background.
3. Binarize the segmented image using threshold.

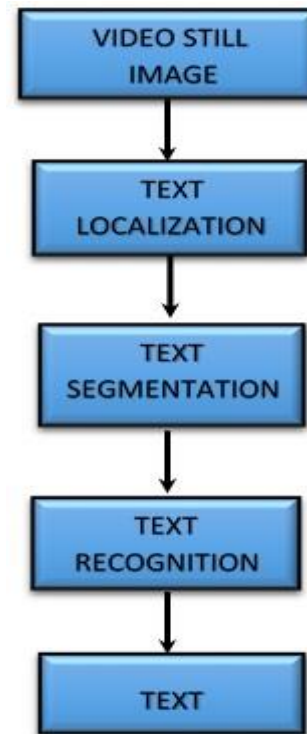


Fig. 3. Process of text extraction from video still.

#### B. Connected region approach

The letters are considered as a homogeneous region. Then divide the video still into homogeneous regions. Split and Merge algorithm is useful in doing such division. Process is as follows;

1. When there is a non-homogeneous region, then split it into four regions.
2. If two adjacent regions are homogeneous, then they can be merged.
3. Using size characterizations of the text the insignificant regions will be deleted.

#### Token Verification & Validation

Text contrasts with background. Edge detection used to find the edges of the text. The number of edges in x and y direction will be calculated. If the edges are higher than the threshold then it will be considered as a text area. Then each text area will be binarized using the luminance threshold.

## 5. Results and Discussions

This section give a detailed discussion of the process of extracting the text in the images using the proposed DWT technique. Fig. 4a shows the compound image. Fig. 4b shows the result of DWT block-based text segmented part. Fig. 4c shows the result of DWT block-based segmented picture part.



Fig. 4a. Sample compound image.



Fig. 4b. Segmented text part.



Fig. 4c. Segmented picture part.

The final result, is the text which will be in white and the background in black. Fig. 5a shows the video still and Fig. 5b shows the segmented text information from the video still.



Fig. 5a. Video still.



Fig. 5b. Segmented text from video still.

## 6. Conclusions

In this paper, the proposed method can be improved with colour segmentation and deleting false regions. Transform domain techniques are implemented in image, audio, and video steganography to enhance the amount of secret information hidden in the cover object, maintaining good quality of stego object and extracted secret information.

## References

- [1] Efficient block prediction-based coding of computer screen images with precise block classification, Ebenezer Juliet, S., JemiFlorinabel, D., IT Dept., Dr.SivanthiAditanar Eng. Coll., Tiruchendur, India. Institution of Engineering and Technology 2011.
- [2] Li, X., Lei, S.: 'Block-based segmentation and adaptive coding for visually lossless compression of scanned documents'. Proc. Int. Conf. on Image Processing, 2001.
- [3] Keslassy, I., Kalman, M., Wang, D., Girod, B.: 'Classification of compound images based on transform coefficient likelihood'. Proc. Int. Conf. on Image Processing, October 2001.
- [4] Lin, T., Hao, P.: 'Compound image compression for real-time computer screen image transmission', IEEE Trans. Image Process., 2005.
- [5] Optimizing Block-Thresholding Segmentation for Multilayer Compression of Compound Images Ricardo L. de Queiroz, Senior Member, IEEE, Zhigang Fan, Member, IEEE, and Trac D. Tran, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 9, NO. 9, SEPTEMBER 2000.
- [6] Gonzalez, R.C., Woods, R.E., Eddins, S.L.: 'Digital image processing using MATLAB' (Prentice-Hall, Upper Saddle River, NJ, 2004).
- [7] Li, D. Doermann and O. Kia. Automatic Text Detection and Tracking in Digital Video. *IEEE Transactions on Image Processing*. Vol. 9, No. 1, pp. 147-156, Jan. 2000.

- [8] Rainer Lienhart and Frank Stuber. Automatic Text Recognition in Digital Videos. *Proc. SPIE 2666: Image and Video Processing IV*, pp. 180-188, 1996.
- [9] Rainer Lienhart and Axel Wernicke. Localizing and Segmenting Text in Images, Videos and Web Pages. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.12, No. 4, pp. 256-268, April 2002.
- [10] V. Wu, R. Manmatha, E.M. Riseman. Textfinder: An Automatic System to Detect and Recognize Text in Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, Issue 11, pp. 1224-1229, Nov. 1999.
- [11] Xiaochuan Chen, Yunhong Wang, Tieniu Tan and Lei Guo, "Blind ImageSteganalysis Based on Statistical Analysis of Empirical Matrix Pattern Recognition", 18th International Conference on Pattern Recognition, Vol.3, pp. 1107 – 1110, 2006
- [12] Yuan Liu, Li Huang, Ping Wang, and Guodong Wang, "A blind imagesteganalysis based on features from three domains", Proceedings of Control and Decision Conference (CCDC), pp. 2933-2936, 2008.
- [13] Xiangyang Luo, Fenlin Liu, Jianming Chen and Yining Zhang, "Imageuniversal steganalysis based on wavelet packet transform, Multimedia Signal Processing", IEEE 10th Workshop on Digital, pp. 780 – 784, 2008.
- [14] Yuanlu Tu, and Shengrong Gong, "Universal steganalysis Using Color Correlation and Feature Fusion", International Symposium on Information Science and Engineering, Vol.1, pp. 107 – 111, 2008.
- [15] Jing Qu Lin and Shang Ping Zhong, "JPEG Image Steganalysis Method Based on Binary Similarity Measures", Proceedings of Eighth International Conference on Machine Learning and Cybernetics, Baoding, pp. 2238-2243, 2009.
- [16] Zhi Min He, W. Wing Y NG, P. K. Patrick Chan and S. Daniel Yeung, "Blind Steganalysis with High Generalization Capability for different Image Databases", Proceedings of the International Conference on Machine Learning and Cybernetics, Guilin, pp. 1690-1695, 2011.
- [17] Mansour Sheikhan, M. Shahram Moin, and Mansoureh Pezhmanpour, "Blind Image Steganalysis via Joint Cooccurrence Matrix and Statistical Moments of Contourlet Transform", 10th International Conference on Intelligent Systems Design and Applications, IEEE, 2010.
- [18] W. N. Lie, G. S. Lin and C. L. Wu, "Robust image watermarking on the DCT domain", In Proceedings of IEEE International Symposium on Circuits and Systems, Geneva, Switzerland, vol.1, pp. 228-231, 2000.
- [19] Y. Q. Shi, G. R. Xuan and D. K. Zou, "Image Steganalysis Based on Moment of Characteristic Function Using Wavelet Decomposition, Prediction Error Image, and Neural Network", IEEE International Conference on Multimedia and Expo, IEEE press, Amsterdam, Netherlands, pp. 269-272, 2005