Region Adaptive Digital Image Watermarking System using DWT-SVD algorithm

Chunlin Song and Peng Xiao
Faculty of Internet of Things
Jiangnan University
Wuxi, China
songchunlin@jiangnan.edu.cn, zhongpengxiao@gmail.com

Sud Sudirman and Madjid Merabti
School of Computing and Mathematical
Liverpool John Moores University
Liverpool, UK
{s.sudirman, m.merabti}@ljmu.ac.uk

Abstract: **Improving the robustness of watermark in withstanding attacks has been one of the main research objectives in digital image watermarking. In this paper we propose a novel region-adaptive watermarking technique that can provide improvements in both robustness and visual quality of the watermarks when compared to the original, non-region-adaptive, embedding technique. The proposed technique, which is derived from our previously published research finding, shows that the relative difference in spectral distributions between the watermark data and the host image plays an important role in improving the watermark robustness and transparency.**

***Keywords* – Quad-tree algorithm, discrete wavelet transform, singular value decomposition, region-adaptive watermarking system.**

I. INTRODUCTION

As one of the most popular and viable techniques in protecting copyrights in digital media, watermarking technology has received enormous level of attention of researchers and practitioners alike. Unfortunately, due to the same reason, watermarking technology has also attracted the attentions of hackers and criminals alike who are interested in breaking the watermarks in order to crack the copyright protection system. As a result, there is a constant challenge on the researchers to keep improving the robustness of the watermarking technique while at the same time maintaining its transparency as to not intruding any legitimate use of the media.

Progress in this area has been steady as can be seen from a healthy number of publications in the field and the sheer number of institutes around the world that deal with the issue [1]. In the more specific field of digital image watermarking, one of the most notable techniques is region-based image watermarking [2]. The paper described a method for embedding and detecting chaotic watermarks in large images. An adaptive clustering technique is employed in order to derive a robust region representation of the original image. The robust regions are approximated by ellipsoids, whose bounding rectangles are chosen as the embedding area for the watermark. The drawback of this technique is due to limited number of suitable regions for storing the watermark the watermark storing capacity can be low.

In this paper, we present a novel watermarking technique which works by adaptively embedding the watermark data into different region of the host image. The rationale of our approach is based on the research finding we came into in our previous work [3-4]. This finding will be detail in this paper for convenience.

II. RATIONALE OF THE REGION-ADAPTIVE APPROACH

Our previous work [3-4] in this field dealt with determining the effect of different removal attacks on watermark signals. The experiment conducted involved subjecting watermarked images to a number of watermark removal attacks and observe the results in both spatial and frequency domains.

In that work, we analyzed the effect of seven different watermark attacks namely Gaussian noise attack, salt and pepper noise attack, Gaussian smoothing attack, sharpen attack, histogram equalization attack, median filter attack and JPEG compression attack.

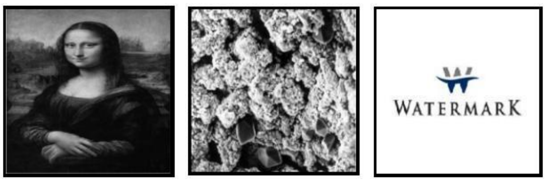
The analysis was carried out using two image analysis tools namely image histogram [10] and Fourier Transforms [11]. In the image processing context, the histogram of an image refers to the distribution of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. Fourier transform converts image data into its frequency spectrum. As with many other signal transforms, Fourier Transform converts time/spatial domain signals into alternative representations that are more amenable for certain types of analysis. The Fourier spectrum of an image is often visualized as a grayscale image whose intensity corresponds to the strength or magnitude of spectrum. The coordinate of these pixels corresponds to the frequency in x-y directions. The low frequency is located at the centre of the image and high frequency is located around the edge. Fourier spectrum has an important characteristic in which its total energy is preserved. This means both low and high frequency components are complementary to each other. In other words, if the high frequency component of an image increases, the low frequency component of that image subsequently decreases. Therefore, any shift in frequency band can be easily detected as the change in spectrum strength in both low and high frequency regions.

The analysis of removal attacks [3] on watermarks is summarized in Table I. Our analysis indicates that the watermark attacks could be divided into two general categories: high frequency (HF) watermark attack and low frequency (LF) watermark attack. The former category attacks the high frequency part of the signal while keeping most of the low frequency component relatively unchanged and the latter category attacks vice versa. Gaussian noise attack, salt and pepper noise attack, sharpen attack and JPEG compression attack belongs to HF watermark attack, whereas Gaussian smoothing attack, median filter attack and histogram equalization attack are more akin to the LF watermark attack.

TABLE I. EFFECTS OF DIFFERENT REMOVAL WATERMARK ATTACKS IN SPATIAL AND FREQUENCY DOMAINS [3]

Watermark attacks	**Effect in spatial domain**	**Effect in frequency domain**
Gaussian smoothing attack	Reduces the variation in image pixel values.	Acts as low pass filter
Gaussian noise attack	Increases the variation in pixel values	Similar effect to a high pass filter
Salt & pepper noise attack	Same as Gaussian noise attack	Same as Gaussian noise attack.
Median filter attack	Similar, albeit much smaller, effect as with Gaussian smoothing attack.	Similar effect to a low pass filter
Histogram equalization attack	Reduces the number of unique gray-scale values and make the histogram more uniformly distributed	Similar, albeit more moderate, effect as Gaussian smoothing attack
Sharpen attack	Reduces the overall image intensity and amplifies differences around edges	Acts as high pass filter
JPEG Compression attack	Reduces the variation in image pixel values and creating blocks, or uniform regions, in the image.	Similar effect to a high pass filter.

This finding suggests that the spectral distribution of the watermarked image plays an important role in the overall robustness of the watermark algorithm to a particular attack. Based upon this finding, we hypothesize that in order to maximize the robustness of the watermarked data, the spectral distribution of the host data and the watermark data should be similar. Furthermore, in order to counter both types of attacks the watermark data should both possess strong high frequency and low frequency part.



(a) (b) (c)

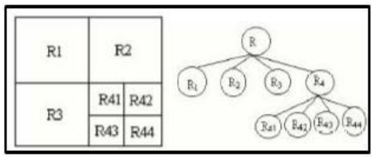
Figure 1. (a) The host image, (b) the HF watermark image and (c) the LF watermark image

To test this hypothesis we develop our proposed region-adaptive watermark technique. This technique employs two unique solutions namely a) it uses two watermark images, each with strong high frequency or low frequency components and b) it embeds different parts of the watermark images into the host image based on the difference of their spectral distributions. Examples of the host image and High Frequency (HF) and Low Frequency (LF) watermark images used in our experiment are shown in Figure 1. The details of our region-adaptive watermarking technique are described in the next section.

III. REGION-ADAPTIVE WATERMARKING

The region-adaptive watermarking technique we propose in this paper works by embedding parts of the rectangular watermark image into selected regions in the host rectangular image. The selection process works by matching the watermark and host image regions with similar spectral distributions. To improve the watermark embedding and extraction speed, we decide not to use regions with arbitrary shape and orientation. Instead, we will use non-overlapping squares of varying sizes as our regions.

To do this we first divide the host image using Quad Tree Partition technique. Assuming that the host image is a square image, Quad Tree Partition works by dividing the image into four equal sized square blocks. A test is carried out on each block to check whether it meets the criterion of homogeneity. If a block meets the criterion, it will not be divided further otherwise it will be subdivided again into four blocks. This process is repeated until every block meets the criterion. The result may have blocks of several different sizes. Figure 2 illustrates a quad tree partition of an image up to 2 levels. The parent node in the tree represents the entire image region and its four descendant's nodes represent the large disjoint sub regions.



(a) (b)

Figure 2. Quad Tree Partition (a) Resulting Blocks, (b) The quad tree structure

The criterion used in deciding the partitioning of a region takes into account the pixel characteristics of that region. In general, the test checks if the characteristics of pixels in that region are too heterogeneous then the test will yield negative and vice versa. To achieve this, an image segmentation algorithm is applied to the host image prior to the partitioning process. Each segment of the host image should represent areas in which pixel characteristics are homogenous. By performing this process, we also avoid computationally expensive task of recalculating pixel characteristics in each region during the partitioning process.

There are a number of image segmentation algorithms that can be used such as Blobworld [6], Gabor Filters [7], and

Markov Random Fields [8] to name a few. In this paper, the Markov Random Field (MRF) technique as described in [8] is used due to its computation speed and statistical categorization of image textures.

A. Discrete Wavelet Transform(DWT)

The embedding of the watermark data into the host image is performed within the partitions produced using the above described process and in the wavelet domain. Wavelet transform has recently become an important tool in image processing and watermarking due to the good energy compaction properties they possess. The transforms are based on small waves, called wavelets, of varying frequency and limited duration. A wavelet series is a representation of a square-integrable function by a certain orthonormal series generated by a wavelet. Furthermore, the properties of wavelet could decompose original signal into wavelet transform coefficients which contains the position information. The original signal can be completely reconstructed by performing Inverse Wavelet Transformation on these coefficients. Wavelet transformation of digital signals can be done by means of the DWT technique.

The DWT separates an image into four parts namely a lower resolution approximation component (LL) as well as horizontal (HL), vertical (LH) and diagonal (HH) detail components. The LL sub-band is the result of low-pass filtering both the rows and columns and contains a rough description of the image. The HH sub-band is high-pass filtered in both directions and contains the high-frequency components along the diagonals. The HL and LH images are the results of low-pass filtering on one direction and high-pass filtering in other direction.

After the image is processed by the wavelet transform, most of the information contained in the original image is concentrated into the LL image. LH contains mostly the vertical detail information which corresponds to horizontal edges. HL represents the horizontal detail information from the vertical edges. The process can be repeated to compute multiple 'scale' wavelet decomposition as shown in Figure 3 and 4.

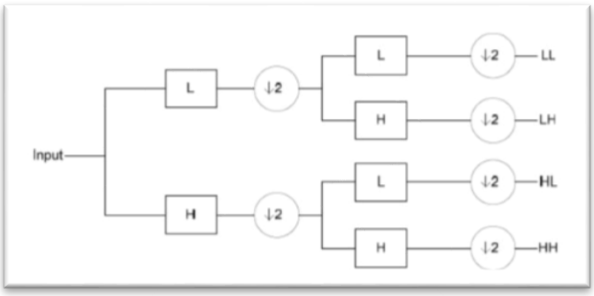


Figure 3.The decomposition step of an image into four sub-bands

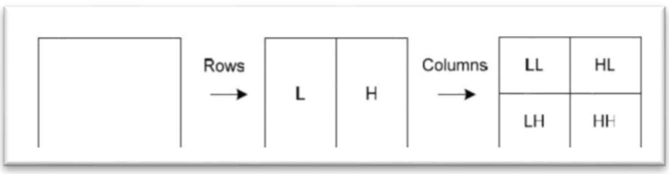


Figure 4.One decomposition step

DWT is very suitable to identify the areas in the host image where a watermark can be embedded effectively. In particular, this property allows the exploitation of the masking effect of the human visual system such that if a DWT coefficient is modified, only the region corresponding to that coefficient will be modified. In general most of the image energy is concentrated at the lower frequency sub-bands LL and therefore embedding watermarks in these sub-bands may degrade the image significantly. Embedding in the low frequency sub-bands, however, could increase robustness significantly. On the other hand, the high frequency sub-bands
HH include the edges and textures of the image and the human eye is not generally sensitive to changes in such sub-bands. This allows the watermark to be embedded without being perceived by the human eye. The compromise adopted by many DWT-based watermarking algorithm, is to embed the watermark in the middle frequency bands LH and HL where
acceptable performance of imperceptibility and robustness could be achieved[19]. Based on this argument, in this proposed technique we will embed the watermark images into the mid frequency, i.e., the LH sub-band of the wavelet coefficients.

B. Singular Value Decomposition (SVD)

Every real matrix A can be decomposed into a product of 3 matrices $A = UDV^T$, where U and V are orthogonal matrices, $U^TU = I$, $V^TV = I$, and D is an m*n rectangular diagonal matrix with nonnegative real numbers on the diagonal. The diagonal entries of D are called the singular values of A, the columns of U are called the left singular vectors of A, and the columns of V are called the right singular vectors of A. This decomposition is known as the *Singular Value Decomposition (SVD)* of A. It is important to note that each singular value specifies the luminance of an image layer while the corresponding pair of singular vectors specifies the geometry of the image layer.

In SVD-based watermarking, several approaches are possible. A common approach is to apply SVD to the whole host image, and modify all the singular values to embed the watermark data. An important property of SVD-based watermarking is that the largest of the modified singular values change very little for most types of attacks. A theoretical analysis of the effects of ordinary geometric distortions on the singular values of an image is provided in a recent paper [9]. In proposed watermarking algorithm and [9], a hybrid scheme based on DWT and SVD is presented, the main connection between them will decompose the host image into four sub-bands first, and then apply SVD to embed watermark data by modifying the singular values. However, there are two notable differences between them:

- In proposed region based algorithm, LL subband is used as opposed to HL and LH subband in [9].
- The watermark coefficients are embedded in specific region of host image as opposed to the entire image.

C. Watermark Insertion

Figure 5 indicates insertion steps of region based watermarking system. The details will be introduced below:

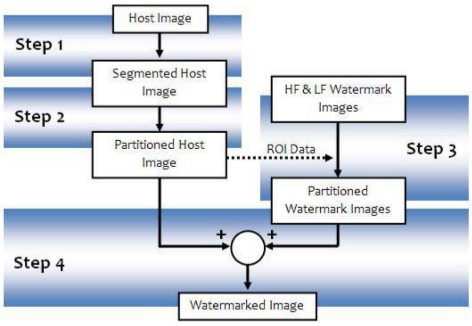


Figure 5. Insertion Step

Step 1: Markov Random Field Segmentation

The goal of MRF segmentation is to simplify the representation of a host image into segmented image, which could be easy to analyse HF and LF region of interest (ROI). HF regions can be seen as having lots of abrupt tonal transitions in a small space. Conversely, LF region are those where the tone remains relatively constant throughout these small areas and the transitions are gradual.

Step 2: Quad Tree Segmentation

The Quad Tree segmentation method splits the MRF segmented image into regions of different sizes depending upon the intensities of MRF segmented image. Large regions represent mainly the LF ROI of host image and small regions indicate HF ROI of host image.

Step 3: HF and LF Watermark Image Partition

At the end of step 2, region of interests are identified in the form of quad tree partitions. This information is then used to partition the HF and LF watermark images. HF watermark image will split into 64 block-based squares and there are 32 squares after LF watermark image partition. The size of HF and LF watermark blocks should be same as HF and LF regions of host image.

Step 4: Region Based Watermark insertion Algorithm

The image to be watermarked is first decomposed through DWT in four levels. The watermark consisting of a pseudorandom binary sequence is inserted by modifying the singular value decomposition. The blocks of HF and LF watermark will insert into LL subband of HF and LF ROI of host image by applying DWT-SVD insertion algorithm directly. Merging the block-based HF and LF ROI of watermarked image is the final step of watermark insertion.

D. Watermark Extraction

Figure 6 below indicates extraction steps of region based watermarking system.

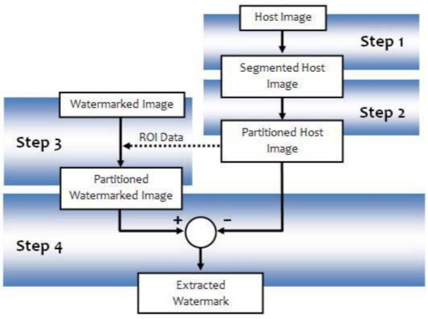


Figure 6. Extraction Step

Step 1 & 2: Image Segmentation and Partition

The first 2 steps in watermark extraction are same as step 1&2 in watermark insertion. The segmented and partitioned ROI of host image will apply to step 4 for extracting HF and LF watermark images. The position of HF and LF ROI are recorded.

Step 3: Partition Watermarked Image

The recorded position of HF and LF ROI is applied to watermarked image. The block-based partitioned of HF and LF ROI of watermarked image is identified.

Step 4: Region Based Watermark Extraction Algorithm

The partitioned watermarked image is first decomposed through DWT in four levels. And then apply SVD to LL subband image. After that, the singular values of watermarked image in LL subband is modified with the singular values of visual host image.The blocks of HF and LF watermark image will extract out of HF and LF ROI of watermarked image by applying extraction algorithm directly. Merging the partitioned and extracted blocks of HF and LF of watermark image is the last step of watermark extraction.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

We applied the region-adaptive watermark technique described in the previous section to the images shown in Figure 1. The host image has dimension of 1024*1024 pixels and both watermark images have dimension of 256x256 pixels. The HF watermark image is divided into 64 blocks of 32x32 pixels and the LF watermark image is divided into 16 blocks of 64x64 pixels.

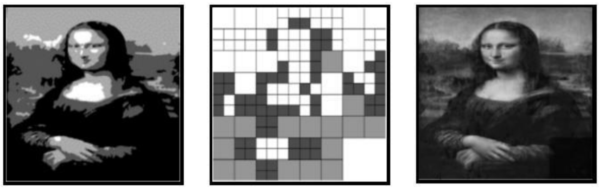
As a quantitative measure of the degradation effect caused by the attacks we use Peak-Signal-to-Noise Ratio (PSNR). The formulation between the original and the attacked watermarked signals can be found described in [3-4]. High PSNR values indicate lower degradation hence indicating that the watermarking technique is more robust to that type of attack.

Two experiments were conducted to test the algorithm. The first experiment is aimed to verify that inserted watermarks images can be extracted with minimal distortion. The second experiment measures the robustness of the proposed region-adaptive technique and compares it to the original DWT-SVD algorithm [9].

A. Watermark insertion and extraction verification.

The first tries to prove that watermark images could be inserted and extracted completely by using proposed region-adaptive technique. Figure 7 shows some in termediate results of the watermark insertion process. The figure shows the segmented host image, the parts of the host image where the different watermarks are inserted and the resulting watermarked image. Figure 8 shows the extracted watermark images.

As can be seen from Figure 7, the extracted watermark images are similar to the original watermark images shown in Figure 1.



(a) (b) (c)

Figure 7. (a) MRF segementated host image, (b) watermark insertion region (dark grey – HF and light grey – LF) and (c) watermarked image

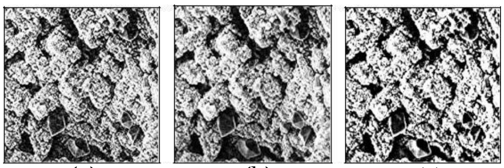
(a) (b)

Figure 8. Extracted (a) HF watermark, and (b) LF watermark images

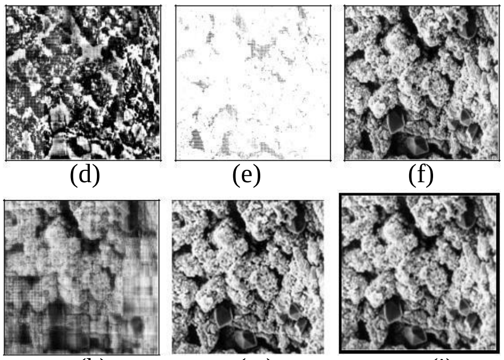
B. Robustness comparison

To test the robustness of the propo sed watermarking scheme, six watermark removal attacks are applied to the watermarked image. They are Gaussian nois e, salt and pepper noise, sharpen, smoothing, histogram equalization and JPEG compression attack. At the same time, rotation, translation and scaling belongs to geometric attacks are also applied. The severity of these attacks can be adjusted by modifying their corresponding parameter values. Definitions of these parameters can be found is given in [3]. Figu res 9 and 10 show the extracted watermark images after differen t attacks.

To compare the robustness of the propos ed technique with the original DWT-SVD technique, we conducted the experiment using 50 different host images. The parameters of the attack algorithms used in the experime nt are the same as those used in Figures 9 and 10. The PSNR v alues between the unmodified watermark image and the atta cked watermarked image are then averaged. These results are summarized in Table II.



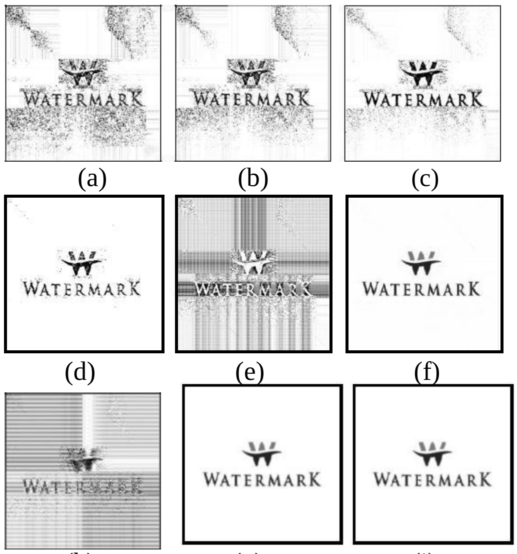
(a) (b) (c)



(d) (e) (f)

(h) (g) (i)

Figure 9. Extracted HF watermark image (a) Gaussian noise with ı = 0.005 (b) Salt and Pepper Noise wit h ı= 0.005 (c) Sharpen with ı = 1.0 (d) Smoothing with ı = 200 (e) Histogram Equalization with Ș = 200 (f) JPEG Compression with ı = 90 (g) Rotation with angle = 30 (h) Translation with movement = 10 (i) Scale



(a) (b) (c)

(d) (e) (f)

(h) (g) (i)

Figure 10. Extracted LF watermark image (a) Gaussian noise with ı = 0.005 (b) Salt and Pepper Noise wit h ı= 0.005 (c) Sharpen with ı = 1.0 (d) Smoothing with ı = 200 (e) Histogram Equalization with Ș= 200 (f) JPEG Compression with ı = 90 (g) Rotation with angle = 30 (h) Translation with movement = 10 (i) Scale

TABLE II. PSNR VALUES OF T HE ATTACKED WATERMARKED IMAGE

Attack Method	**Region Based Algorithm**		**Original DWT-SVD**	
	HF	*LF*	*HF*	*LF*
Gaussian noise	8.1761	9.4111	6.6154	6.4518
Salt & pepper noise	9.6057	10.2655	7.7526	7.4676
Sharpen	10.0342	14.2673	8.2857	6.712
Smoothing	4.3913	11.5172	7.1321	11.3947
Histogram Equalization	4.9438	21.1282	4.9859	19.5592

JPEG Compression	7.0533	8.8899	12.1745	13.0725
Rotation	5.4148	5.8431	4.8386	0.5908
Scale	8.7134	15.8311	5.0674	1.1318

As can be seen from the results shown in Table II, the proposed region-based approach produced higher PSNR values in most types of attacks including Gaussian noise attack, salt and pepper noise attack sharpen attack, smoothing attack and rotation attack. In addition, the robustness of histogram equalization, JPEG compression and translation has similar results between proposed region-adaptive algorithm and original DWT-SVD algorithm.

V. CONCLUSION

The contribution of this paper to the advancement of digital image watermarking technology can be summarized as follows:

We have reviewed an analysis of a number of attack types on image watermarking. The analysis was carried out using two image analysis tools namely image histogram and Fourier spectrum for spatial and frequency domain analysis respectively. Using the results of our experiments, we argue that existing techniques have different sensitivity and robustness levels to different attacks. The results also uncover a number of common similarities between different types of watermark attack

We have derived a hypothesis based upon the analysis result that in order to maximize the robustness of the watermarked data, the spectral distribution of the host data and the watermark data should be similar. In addition, we also suggested that in order to counter most types of attacks the watermark data should possess both strong high frequency and low frequency components except histogram equalization, JPEG compression and translation attack.

We have presented a novel digital image watermarking technique that takes into account the results of our previous analysis and testing of the hypothesis. Our technique utilizes a number of technologies namely dual watermarking, image segmentation and partitioning, and DWT-SVD to fulfill the design criteria set to prove the hypothesis. The experiment results show that our technique is more robust to attacks than the original DWT-SVD technique. This result further provides proof of the hypothesis we had previously derived.

As a future work, we are currently looking at the possibility of using our technique to provide detection and possibly correction to any attacks to watermarked data. This application will be useful especially if we are to use this technique in real world copyright protection system.

ACKNOWLEDGMENT

The work was sponsored by Jiangnan University of Science & Technology Young Scholar Grant No. JUSRP11462, Key Project of Jiangnan University Fund Grant No. JUSRP51414A and National Science Foundation of China Grant No.61304264 .

REFERENCES

[1] Petitcolas, F.A.P., Anderson, R.J., Kuhn, M.G., Information Hiding—A survey, Proceeding of the IEEE, Special Issue on Protection of Multimedia Content, 1062-1078, July 1999.

[2] A. Nikolaidis and I. Pitas, "Region-based image watermarking," *Image Processing, IEEE Transactions on*, vol. 10, no. 11, pp. 1726-1740, 2001

[3] C.Song, S.Sudirman and M.Merabti, "A Spatial and Frequency Domain Analysis of the Effect of Removal Attacks on Digital Image Watermarks", Proc 11th of PostGraduate Network Symposium, 119-124, June, 2010.

[4] C.Song, S. Sudirman, M.Merabti and D.L.Jones, "Analysis of Digital Image Watermark Attacks", 6th IEEE International Workshop on Digital Rights Management, 2010.

[5] P.N.Tao and Eskicioglu, "A robust multiple watermarking scheme in the Discrete Wavelet Transform domain", Internet Multimedia Management System Conference, vol. 5601,133-144, 2004.

[6] Chad, C., Magan, T., Serge, B., Joseph H., Jitendra M.: Blobworld: A System for Region-Based Image Indexing and Retrieval, Visuan Information and Inforamtion Systems, Lecture Notes in Computer Science, Vol: 1614/1999, 1999.

[7] Jain, A.K., Farrokhnia, F.: Unsupervised Texture Segmentation Using Gabor Filters, Systems, Man and Cybernetics, Conference Proceedings., 14-19, Nov, 1990

[8] Erdogan, C., Deliang, W,: Texture Segmentation Using Gaussian-Markov Random Field and Neural Oscillator Networks, IEEE Transactions on Neural Networks, Vol. 12, No. 2, 394-404, March 2001.

[9] Chin-Chin Lai and C.-C. Tsai, "Digital Image Watermarking Using Discrete Wavelet Transform and Singular Value Decomposition," IEEE Transaction on Instrumentation and Measurement, Vol.59, no. 11, 2010

[10] Kotoulas L and Andreadis I, "Colour Histogram Content-based Image Retrieval and Hardware Implementation", IEEE Proceedings on Circuits, Devices and Systems, Vol. 150, Issue 5, 2003.

[11] Vundela U, Kumar, T.S and Suresh, R, "Computation of Fractional Fourier Transform using Filter Bank Approach and its Application". 2013 International Conference on Computer Communication and Informatics (ICCCI), coimbatore, 1-5, 2013.