Imperceptibility and Robustness Analysis of DWT-based Digital Image Watermarking

Yusnita Yusof[1], Othman O. Khalifa[1]
[1] *Kulliyyah of Engineering, International Islamic University, Gombak, Malaysia*
ynitayusof@yahoo.com

Abstract

Digital watermarking is distinctive depending on the techniques used and its intended applications. This paper concentrates on invisible digital image watermarking using discrete wavelet transform. The work flow involves watermark embedding, attacks and watermark extraction. Two methods are proposed and analyzed to imply the imperceptibility and robustness, among the most important criteria of digital watermarking, using three types of attacks – JPEG compression, blurring and histogram equalization. The results are compared through subjective visual inspections and calculative measurements using PSNR for watermark imperceptibility and SSIM Index for watermark robustness.

I. INTRODUCTION

Digital watermarking has been inspired from security concerns over multimedia contents due to the advances of computer technology. Nowadays, it is easy to obtain, manipulate, distribute and store these contents due to evolution of Internet, excellent multimedia tools and low-cost storage devices. Research community and industry has shown extensive interests in developing and implementing digital watermarking.

Watermarking techniques can be classified into many types, depending on various aspects. For examples, classification may be based on type of content to be watermarked (i.e. image, audio or video), working domain being used (i.e. spatial or transform), information type (i.e. blind, semi-blind or non-blind) and many others which actually determines its intended applications. Several applications are described by Cox et al. in [1] and Katzenbeisser and Petitcolas in [2].

In earlier days, watermarking techniques are commonly implemented in spatial domain. Over the years, more techniques are being implemented in transform domain including DCT, DFT and DWT.

In [3], the authors have made extensive analysis of the watermarking scheme proposed in [4]. Based on their analysis of [4], two different watermarks are embedded in DWT domain by modifying both low and high frequency coefficients. It is observed that the advantages and disadvantages of embedding the watermark in low and middle-to-high frequencies are complement to each other by performing different kind of attacks. As claimed by authors in [3], the scheme has its flaws as it used the same scaling factor for both bands which leads to visible degradation in the image. Thus they generalized the scheme by embedding the same visual watermark in all four bands using first and second level decompositions with different scaling factors. Both [3] and [4] used grayscale cover image and binary visual watermark.

In [5], a scheme is proposed by embedding grayscale watermark DWT coefficients into grayscale host image coefficients by using first level decomposition. The scheme enables using watermark size as much as 25% of host image size and provides simple control parameter which is scaling factor to tailor between data hiding and watermarking purposes, with respect to JPEG compression attack.

In this paper, two methods are generalized based on the three schemes mentioned above. First level DWT coefficients of grayscale watermark are embedded into second level DWT coefficients of cover image in all subbands. The size of watermark is one forth the size of cover image. Embedding gain is used as control variable to compensate between watermark imperceptibility and robustness, by performing three types of attacks – JPEG compression, blurring and histogram equalization. The results are compared and analyzed for three different grayscale images – baby, boat and hill images.

978-1-4244-1692-9/08/$25.00 ©2008 IEEE

II. PROPOSED METHODS

In two-dimensional DWT, each decomposition level yields four bands of data, one low pass band (LL), and three high-pass bands (horizontal HL, vertical LH and diagonal HH).

watermarked image or as an independent measure of its acceptability. Meanwhile, robustness can be evaluated by applying various kinds of signal distortions (attacks) to the watermarked image and measuring detection probability of the watermark after those distortions.

The general workflow is depicted in Figure 1.

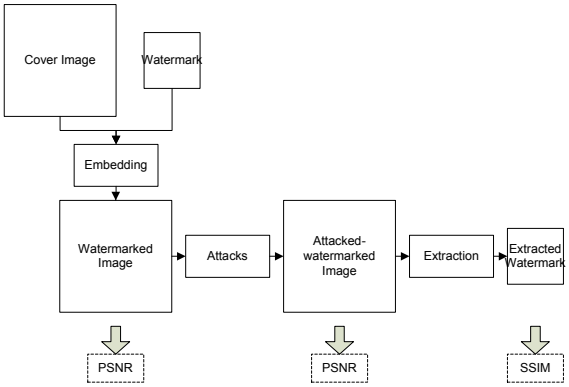


Figure 1. General workflow of image watermarking system

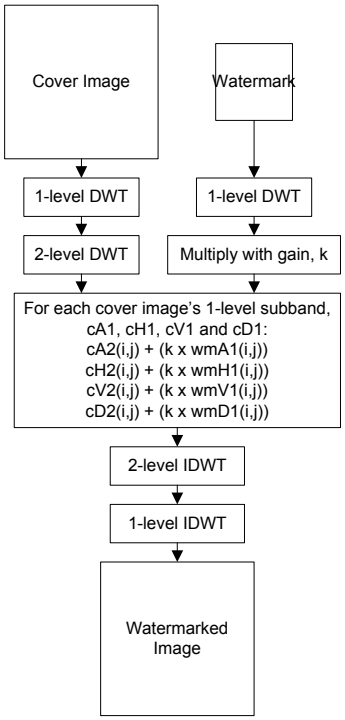


Figure 2. Method A – embedding

The proposed methods are illustrated as Method A and Method B.

Watermark imperceptibility can be expressed either as a measure of similarity between the original and

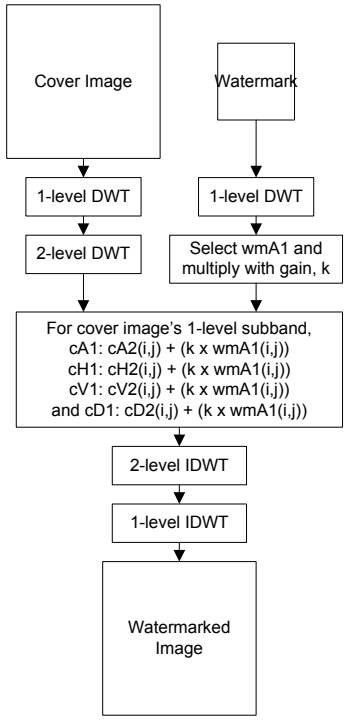
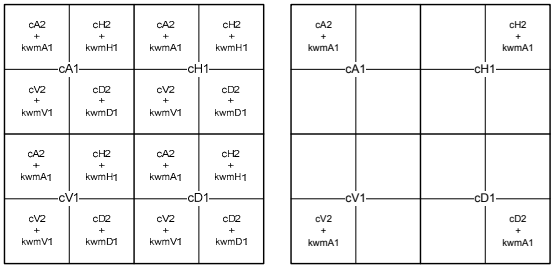


Figure 3. Method B – embedding

During extraction, the process is reversed for both methods. The difference between these two methods is as shown in Figure 4, referring to cover image's subbands used to embed watermark's subbands.



Method A vs. Method B

Figure 4. Embedding subbands

I. EXPERIMENTS AND RESULTS

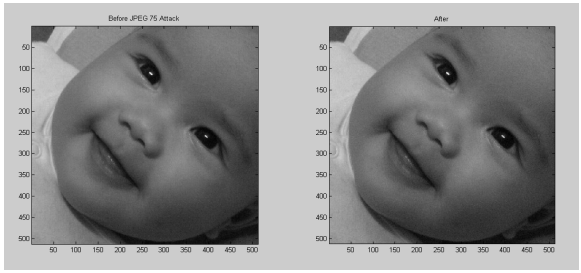
Three different images of size 512x512 are used as cover image with a watermark of size 256x256. The range of embedding gain used is from 2 to 8.

Figure 5. Cover images (baby, boat and hill) and watermark image

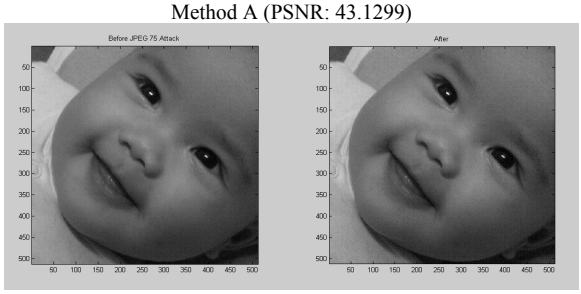
Three types of attacks are performed on the watermarked images with different embedding gain values. It is assumed that the scheme is non-blind where the extraction process requires original cover image and original watermark.

For qualitative visual inspections, the results of both methods are shown using all three images, each image for each attack, respectively with k=2 and k=8.

A. JPEG Quality 75 Compression

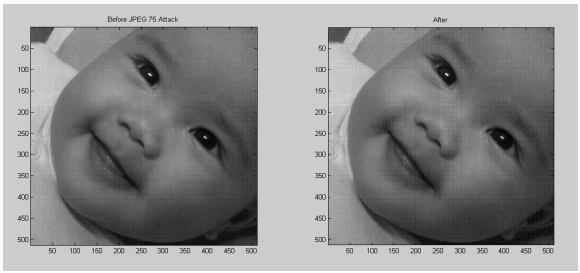


Method A (PSNR: 43.1299)

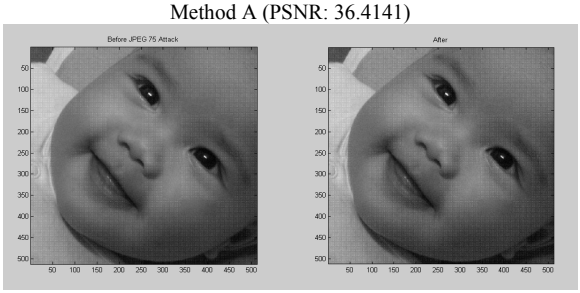


Method B (PSNR: 45.328)

Figure 6. JPEG 75: Watermarked images vs. attacked-watermarked images for k=2

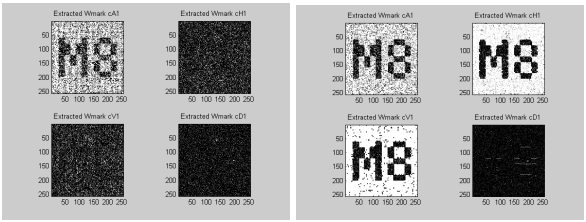


Method A (PSNR: 36.4141)

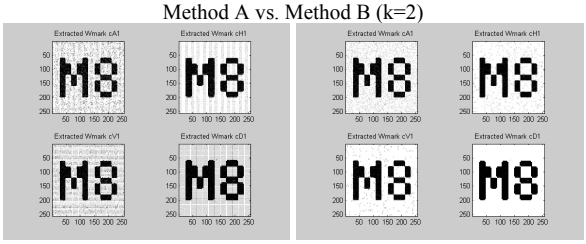


Method B (PSNR: 45.0021)

Figure 7. JPEG 75: Watermarked images vs. attacked-watermarked images for k=8



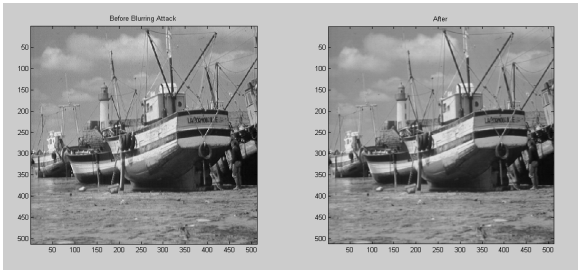
Method A vs. Method B (k=2)



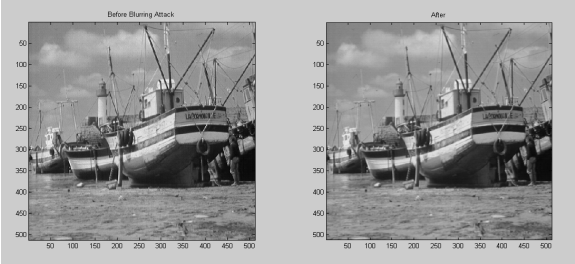
Method A vs. Method B (k=8)

Figure 8. JPEG 75: Extracted watermark in all subbands

B. Blurring

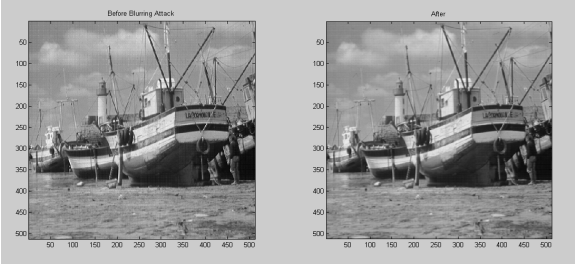


Method A (PSNR: 29.9433)

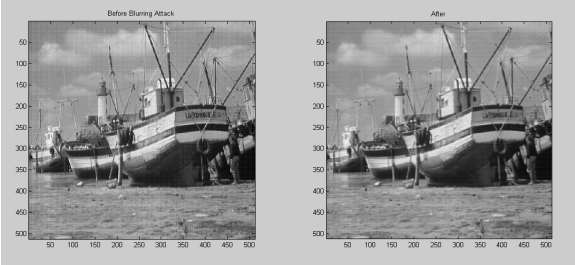


Method B (PSNR: 30.0692)

Figure 9. Blurring: Watermarked vs. attacked-watermarked images for k = 2

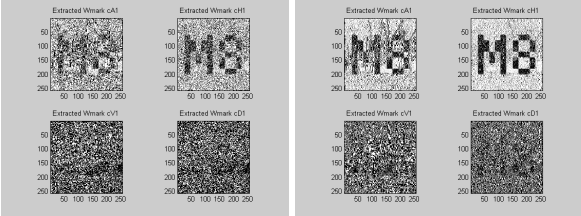


Method A (PSNR: 27.8805)

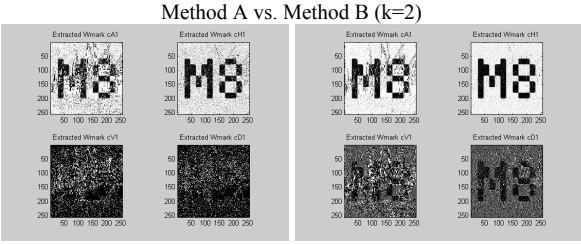


Method B (PSNR: 29.3779)

Figure 10. Blurring: Watermarked vs. attacked-watermarked images for k = 8



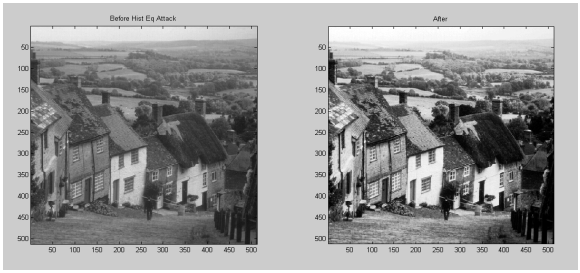
Method A vs. Method B (k=2)



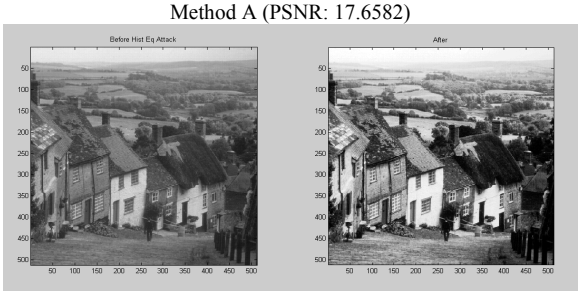
Method A vs. Method B (k=8)

Figure 11. Blurring: Extracted watermark in all subbands

C. Histogram Equalization

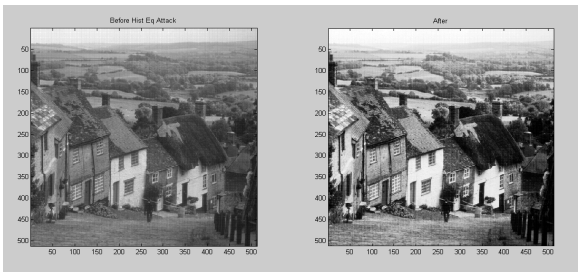


Method A (PSNR: 17.6582)

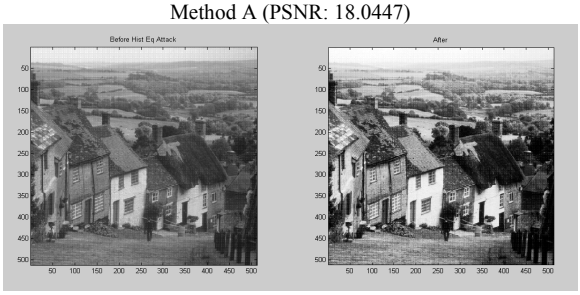


Method B (PSNR: 17.6582)

Figure 12. Histogram equalization: Watermarked vs. attacked-watermarked images for k = 2



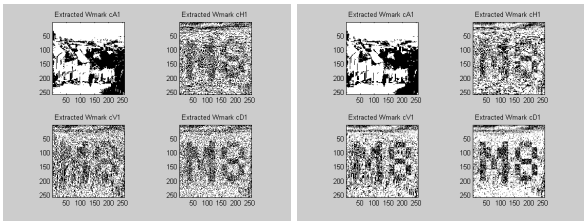
Method A (PSNR: 18.0447)



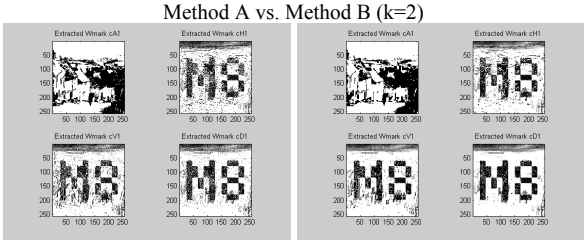
Method B (PSNR: 18.034)

Figure 13. Histogram equalization: Watermarked vs. attacked-watermarked images for k = 8

For quantitative measurements, the cover image perceptibility is determined using PSNR values while the watermark robustness is computed using SSIM Index. Detailed information of SSIM Index is explained in [6]. The results are shown as graphs.



Method A vs. Method B (k=2)



Method A vs. Method B (k=8)

Figure 14. Histogram equalization: Extracted watermark in all subbands

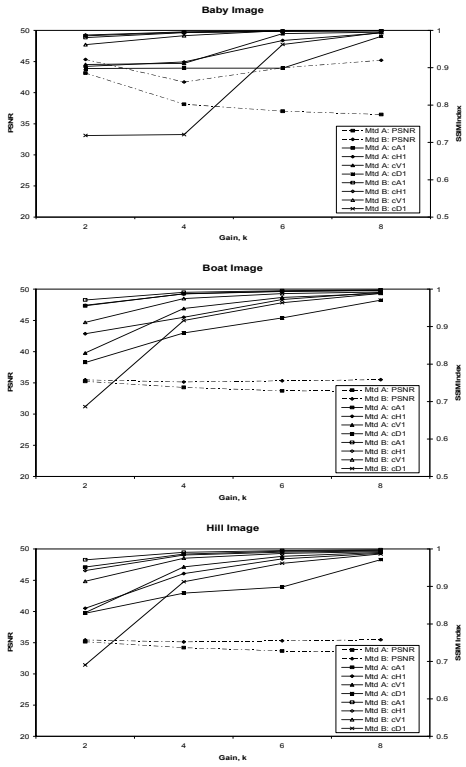


Figure 15. PSNR and SSIM Index for JPEG quality 75 attack

Looking at the graphs in Figure 15, as the gain increases, the PSNR values decrease while most SSIM values increase. In terms of subbands, the low frequency cA1 is the most robust while the high frequency cD1 is the most fragile against these attacks.

This is due to fact that during JPEG compression's quantization process, it discards many of the high-frequency (noise-like) details and preserves the slowly-changing image information.

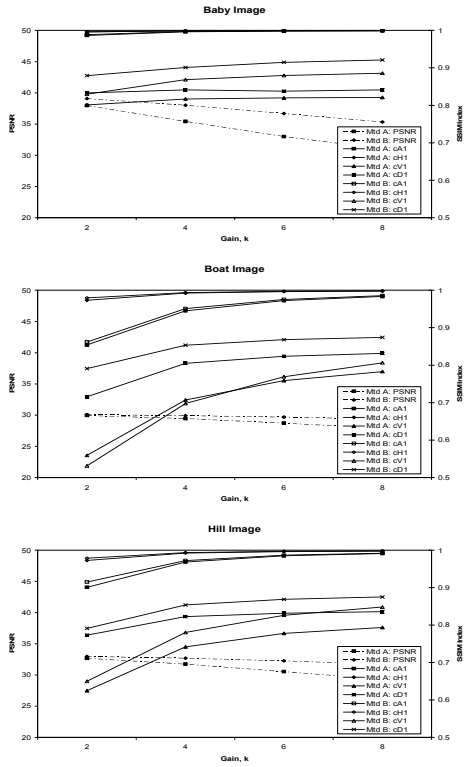


Figure 16. PSNR and SSIM Index for blurring attack

The blurring is performed by using image filtering that convolves the point spread function (PSF) with the input image, in this case, the watermarked image to produce blurred watermarked image.

Evidently showed by the graphs for all images in Figure 16, high frequency cV1 and cD1 subbands are the most affected ones. This conforms with the fact that blurring or also known as smoothing suppresses noise and small fluctuations i.e. in the frequency domain, this process refers to the suppression of high frequencies.

Referring to graphs in Figure 17, SSIM values for cA1 subband are the lowest, almost nearing 0 index for gain = 2 and increased when the gain increased. While for the other three subbands, the SSIM values are almost approaching index 1, which means very good extraction. In terms of PSNR, the values are low, indicating significant difference between watermarked and attacked-watermarked images.

In theory, this attack deals with image's intensity values (brightness level). Remapping or relocating these values throughout the brightness scale could be visually analyzed based from the histograms. Such manipulation definitely affects the wavelet subbands'

coefficients as well, especially on the low frequency cA1 subbands for having the largest magnitudes among all.

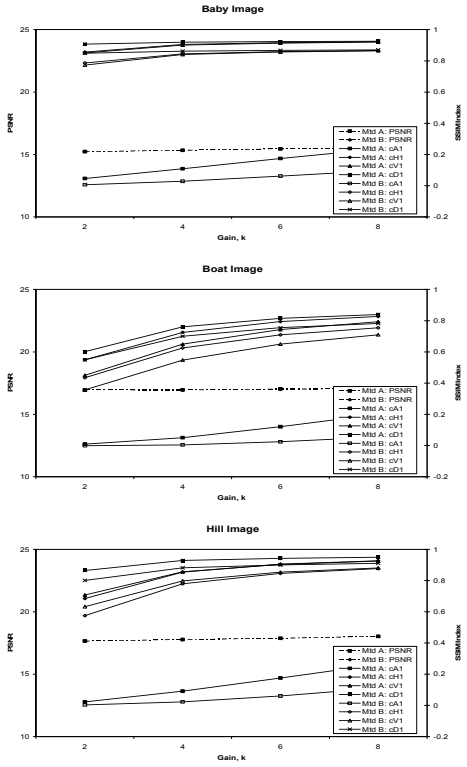


Figure 17. PSNR and SSIM Index for histogram equalization attack

II. DISCUSSIONS AND CONCLUSION

Two methods are proposed and presented. The difference between the two is in the embedding process where method A embeds each watermark 1-level coefficients into each cover image's 2-level DWT coefficients for all four subbands, respectively. While method B embeds watermark 1-level low-pass (LL) coefficients into chosen 2-level DWT coefficients based on 1-level subband for all four subbands, accordingly.

Unlike the previous papers [3]-[5], this analysis used three different images to see the effects of image's characteristics on perceptibility and robustness. With careful inspections, more visible distortions are detected at smooth regions of the cover image (example baby image) compared to regions with more textures (example hill image).

Based on the graphs, it is observed that different embedding gain yields different outcomes of the perceptibility (PSNR) and robustness (SSIM Index). Smaller gain reflects with good cover image's perceptibility but with less robust watermark extraction and vice versa.

Attacks commonly alter either low frequencies or high frequencies, thus embedding watermark in both bands gives advantages in terms of robustness. Low frequencies watermark is robust to attacks with low pass characteristics such as compression and blurring, while high frequencies watermark is robust to modifications such as histogram equalization. Both methods could survive a wide range of attacks as these watermarks might be destroyed in one band, but could still be extracted from the other bands.

Embedding gain acts as control variable to counterbalance between image perceptibility and watermark robustness in finding the best possible results.

Further improvement includes more attacks to be performed on both methods to analyze and summarize its performance in terms of perceptibility and robustness, being the two most important criteria in any watermarking system.

REFERENCES

[1] I. J. Cox, M. L. Miller and J. A. Bloom, *Digital Watermarking*, Morgan Kaufmann Publishers, 2002.

[2] S. Katzenbeisser and F. A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, 2000.

[3] P. Tao and A. M. Eskicioglu, "A Robust Multiple Watermarking Scheme in the Discrete Wavelet Transform Domain", *Optics East 2004 Symposium, Internet Multimedia Management Systems Conference V*, Philadelphia, PA, USA, Oct. 25-28, 2004.

[4] R. Mehul and R. Priti, "Discrete Wavelet Transform Based Multiple Watermarking Scheme," *Proc. of IEEE Region 10 Technical Conference on Convergent Technologies for the Asia Pasific*, Bangalore, India, Oct. 14-17, 2003.

[5] J. J. Chae and B. S. Manjunath, "A Robust Embedded Data from Wavelet Coefficients", *Proc. of the SPIE International Conference on Storage and Retrieval for Image and Video Databases VI*, San Jose, CA, Jan. 28-30, 1998, Vol. 3312, pp. 308-317.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.