

Phase signature-based image authentication watermark robust to compression and coding

Farid Ahmed^{*a}, Ira S. Moskowitz^b

^aThe Catholic University of America, Dept. of EECS, Washington, DC, USA 20064

^bNaval Research Laboratory, CHACS-Code 5540, Washington, DC USA 20375

ABSTRACT

We present the further development of a watermarking technique that embeds an authentication signal in an image. In this paper, we concentrate on the JPEG 2000 image format. The detection/extraction of this signal can then be used to decide whether the image has gone through any intentional malicious tampering. Therefore, the watermark needs to be fragile to such tampering attacks. On the other hand, we need to make sure that the authentication is robust to change resulting from the watermarking process itself, or from necessary changes such as image compression.

We address the robustness against watermarking process issue in two ways. First, we decompose the image into phase and magnitude values. A signature is then generated from the phase values. In particular, binary phase-only filters and their variants will be utilized for this. This signature is subsequently hidden into the magnitude part by a bit-plane embedding technique. The disjoint operations of signature generation and signature embedding minimize the embedding artifacts of the authentication signal. Secondly, we use wavelet decomposition, whereby, the signature can be generated from one subband, and then it can be embedded in other subband(s), or the same subband.

Keywords: Authentication, Wavelet, JPEG 2000, Watermarking, Correlation, Phasemark™

1. INTRODUCTION

Digital watermarking is the process of hiding or embedding a digital message into digital media. Often the embedded message is a variant from some sort of natural ‘signature’ of the original media, so that it can be used as an authentication watermark, thus authenticating the validity of the content. There have been a number of approaches for the design of authentication watermarks [1], with varying degrees of success. In general, authentication watermarks are variants of semi-fragile watermarks, where the watermark is robust to some desirable image degradation but yet fragile to some undesirable distortions. An example of desirable distortion is image compression.

The core processing in any compression algorithm aims at quantizing and truncating small-valued coefficients in the transform domain. This very processing is opposed to the notion of watermarking, which modifies the coefficients by a small amount [2]. Therefore there has been research [3] that combines both compression and watermarking together to minimize the effects of compression on the effectiveness of watermarking.

The current work is based on the Phasemark™ [4] algorithm, which embeds phase-signature based authentication information into the Fourier magnitudes of an image. We propose to enhance the Phasemark™ performance against JPEG 2000 [5] compression by using a wavelet-domain bit-plane embedding technique. See also [6].

* ahmed@cua.edu

The main motivation of this work is due to the effect that JPEG 2000 compression is more harsh on the high-resolution subbands or frequencies, than compared to the low-resolution ones [2]. Therefore, while embedding watermark signals, one should avoid the high-frequencies/bands as much as possible. JPEG 2000 compression involves a number of processing steps --- codestream syntax, data ordering, quantization, arithmetic coding, discrete wavelet transform (DWT), DC level shifting, etc. In our work, we focus only on how to take advantage of the region of interest that will make the watermarked image more robust against JPEG 2000 compression. The mother wavelet filters used in JPEG 2000, part 1, are integer Biorthogonal 5/3 (for lossless) and the non-reversible 9/7 filter (for lossy). Later parts of the standard allow more flexibility in choosing filters.

The wavelet transform has been used as a popular tool for decomposing a signal or an image at different resolutions [7]. In our current work, the DWT is used for the multiscale/multiresolutional decomposition. An image can be decomposed into a low-resolution smooth image and a detailed image. Let us assume that we are performing one level of decomposition. We denote the smooth image as LL. The detailed image has three components depending upon the directional vector used in the decomposition. These are horizontal (HL), vertical (LH), and diagonal (HH). In [3] the authors mention a number of parameters that affect the compression performance.

2. PROPOSED WAVELET DOMAIN WATERMARKING

A binary phase-only filter (BPOF) will be hidden, after cryptographic manipulation, as an authentication watermark in the original image. Thus, the BPOF forms a signature of the image, which is embedded as the watermark. (All images in this paper are greyscale, but our method will work for color images also [4].) In the detection stage, the extracted and decrypted hidden signature (after the obvious conversion from $\{0,1\}$ values to $\{-1,1\}$ values, respectively) will be correlated with extracted phase-only information (which takes values on the unit complex circle). In addition, the phase quantization in the BPOF has built in tolerance to minor changes to the image, which is a desirable feature for a semi-fragile image authentication watermarks. The above describes our method PhasemarkTM as given in [4]. In this paper, we modify PhasemarkTM, calling it PhasemarkTM 2K by modifying and optimizing our method for JPEG 2000 images. The major modification in PhasemarkTM 2K is that instead of extracting the signature from the entire image and then embedding it back into the entire image, PhasemarkTM 2K extracts the signature from a particular wavelet subband and embeds it into the same or different subband. However, the phase-only information that the extracted signature is correlated against is only from the LL subband.

We use the DWT with the Haar mother wavelet [7]. We use only one level of filtering/decomposition. Future versions of PhasemarkTM will explore using other mother wavelets and higher levels of wavelet filtering. The subbands of interest to us are LL, LH, and HL only, see Figure 1. We extract the watermarking signature only from LL, but may embed in any of the three. No experiments are done embedding the signature into the HH subband because this subband is the least robust to wavelet compression, regardless of the mother wavelet. See Figure 2.

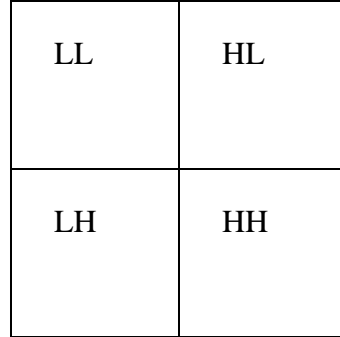


Fig. 1. One level subband decomposition.

(*Embedding*) Figure 3 shows the flowchart of the watermark embedding process. Let SS represent any of the three subbands LL, LH, or HL. The SS subband is transformed to the Fourier frequency domain via the discrete Fourier transform (DFT). The Fourier transform of the SS subband can be expressed as

$$H(u, v) = X(u, v) \exp(j\phi(u, v))$$

Where, $X(u, v)$ is the *magnitude* part of the frequency component given by $|H(u, v)|$, and $\phi(u, v)$ is the *phase* part of frequency $H(u, v)$ given by

$$\phi(u, v) = \arctan\left(\frac{\text{Re}(H(u, v))}{\text{Im}(H(u, v))}\right)$$

The BPOF is 1, if $\cos(\phi)$ is ≥ 0 , and the BPOF is 0 otherwise. Please keep in mind that the original BPOF, which comprises the signature that we embed, is *only extracted* from the LL subband, but we may embed it in LL, LH, or HL (and then look for it there, resp.). The (Fourier frequency) magnitudes span a range of values from 0 to MAX, and the phases are in the interval $(-\pi, \pi]$. In our watermarking embedding process, the phase is kept unchanged and the magnitude part $X(u, v)$ is modulated. The real valued $X(u, v)$ is first transformed to integer valued as follows.

$$R(u, v) = \text{round}[X(u, v)]$$

where the *round(.)* function rounds the operand to the nearest integer value, that makes it representable by a fixed number of q bit planes (depending on what MAX is). Suppressing (u, v) , $R(u, v)$ can be given in bit slice format as $R=R_{q-1}, R_{q-2}, \dots, R_1, R_0$, where R_j is the j^{th} bit-plane of the rounded magnitude, and q is determined by MAX. We now embed an encrypted [4] version of the BPOF (only extracted from the LL subband) into any of the possible SS subbands. We denote the subband that has the signature embedded in it as EB. We do this by determining an i value and replacing R_i with the encrypted BPOF. This gives modified Fourier frequency values for the particular SS subband (EB) we chose for the embedding. We next apply the inverse discrete Fourier transform to these modified Fourier frequency values and thus arriving at a modified EB subband. We now apply the inverse

DWT to all four subbands (of course one of them, EB, has been modified) to arrive at our watermarked image. This image is saved in the JPEG 2000 format at various quality levels.

(Extraction) We start with a test JPEG 2000 image. We test the assumption that the BPOF signature from LL has been embedded into EB. We apply the DWT to the spatial representation of this test image, and arrive at the four subbands LL, LH, HL, and HH. We extract the BPOF from EB (the choice of which subband is EB and the i value is known by both the transmitter and receiver of the watermarked image), by transforming EB into the Fourier frequency domain as above. The discrete Fourier transform of EB is

$$T(u, v) = |T(u, v)| \exp(j\phi_T(u, v))$$

After a rounding operation on Fourier magnitudes, we extract the desired bit plane i , undo the cryptographic processing done during the embedding phase and extract the signature. The extracted (see above description of the BPOF) signature is denoted as $S^i(u, v)$. As an example, if the cryptographic processing is permutation, then the detector generates the inverse permutation matrix and applies that on the extracted bit plane to recover the hidden signature.

Now, from the phase information of LL, the phase-only filter (POF) is computed as

$$H_{POF}^T(u, v) = \exp(-j\phi_T(u, v))$$

LL is used because this is where the original BPOF signature was calculated.

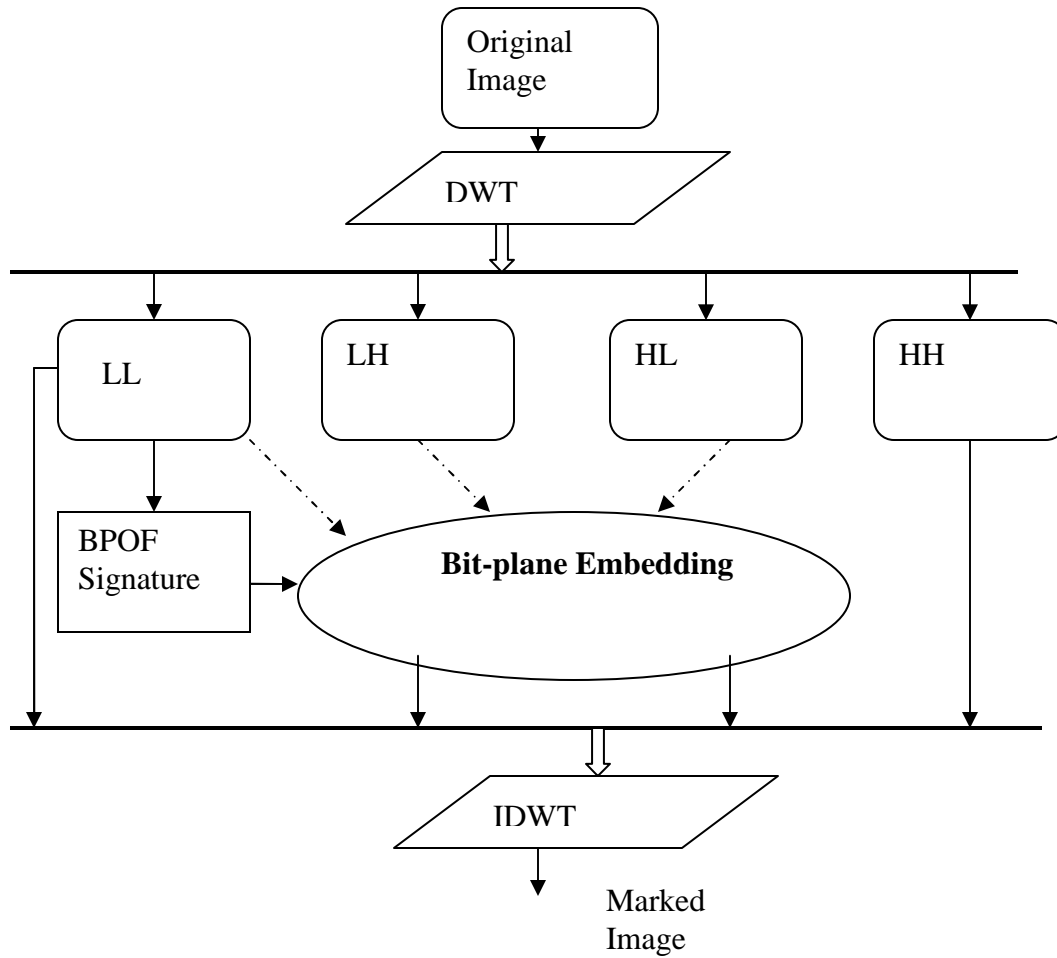


Fig 2. DWT Domain Signature-based Authentication.

The next step is the correlation of the POF with the extracted signal. It is well known that the correlation of two spatial images is given by the IDFT of the product of the DFT of one image with the conjugate DFT of the other image. Therefore, based upon this and successes in Fourier optics (see [4] for historical details) we propose the following as our test for “autocorrelation.”

$$Corr(k,l) = IDFT(H_{POF}^T(u,v) * S'(u,v))$$

This is the test that the “correlator” performs.

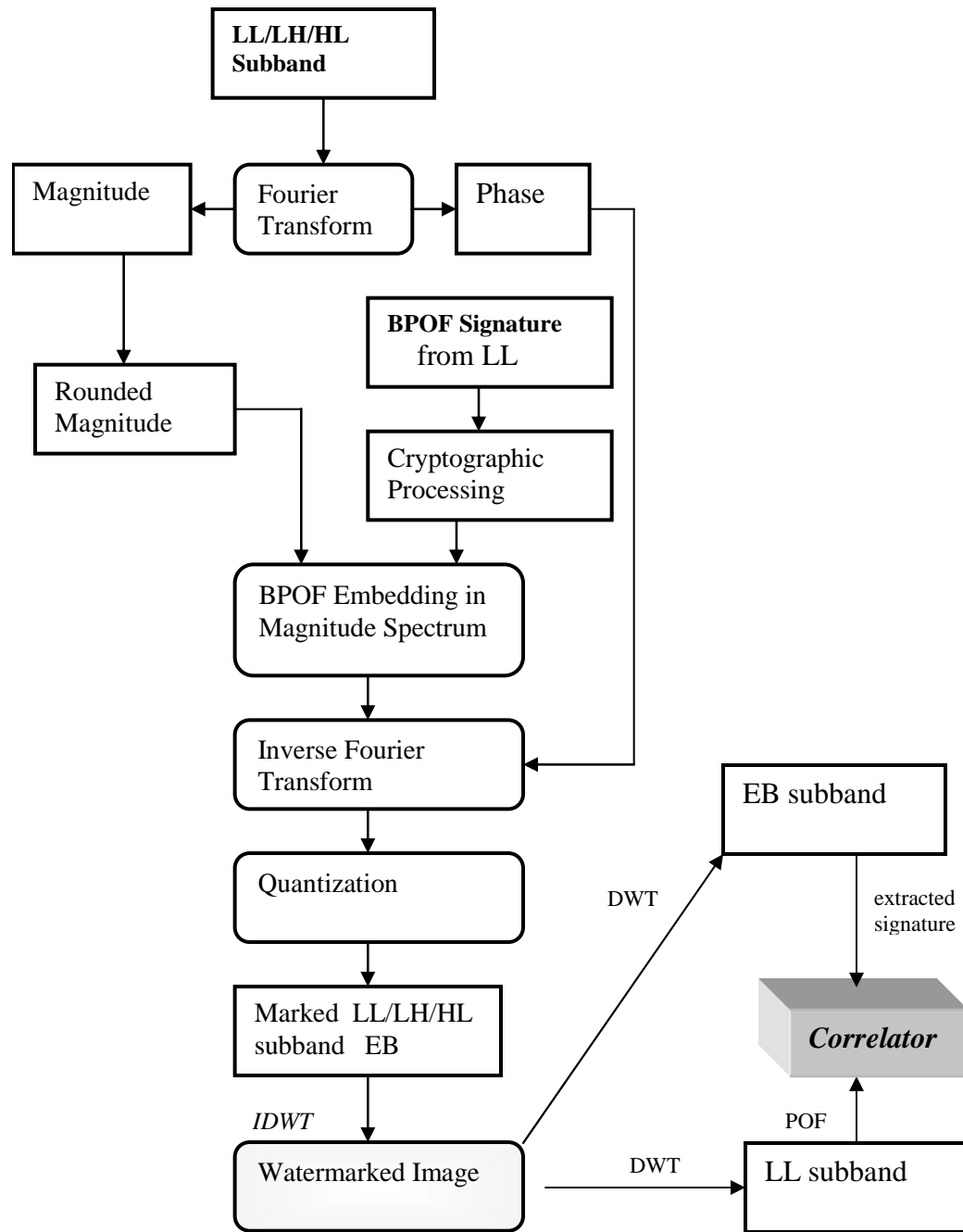


Fig. 3. Bit-plane embedding and simplified detector for Phasemark™ 2K.

3. SIMULATION AND RESULTS

Simulation of the above mentioned algorithm is performed on one set of ten 512x512 greyscale images obtained from the USC SIPI Database [8]. While performing the DWT, we use periodic symmetric extension to take care of the convolution processing of the boundary values to make the transformed image the same size as the original image. The watermarked image is saved as an uncompressed TIFF,

therefore initially (Figs. 4-7) all distortion is solely from the watermarking algorithm (later in Figs. 8&9 we analyze JPEG 2000 compression affects and discuss the worth of our wavelet method).

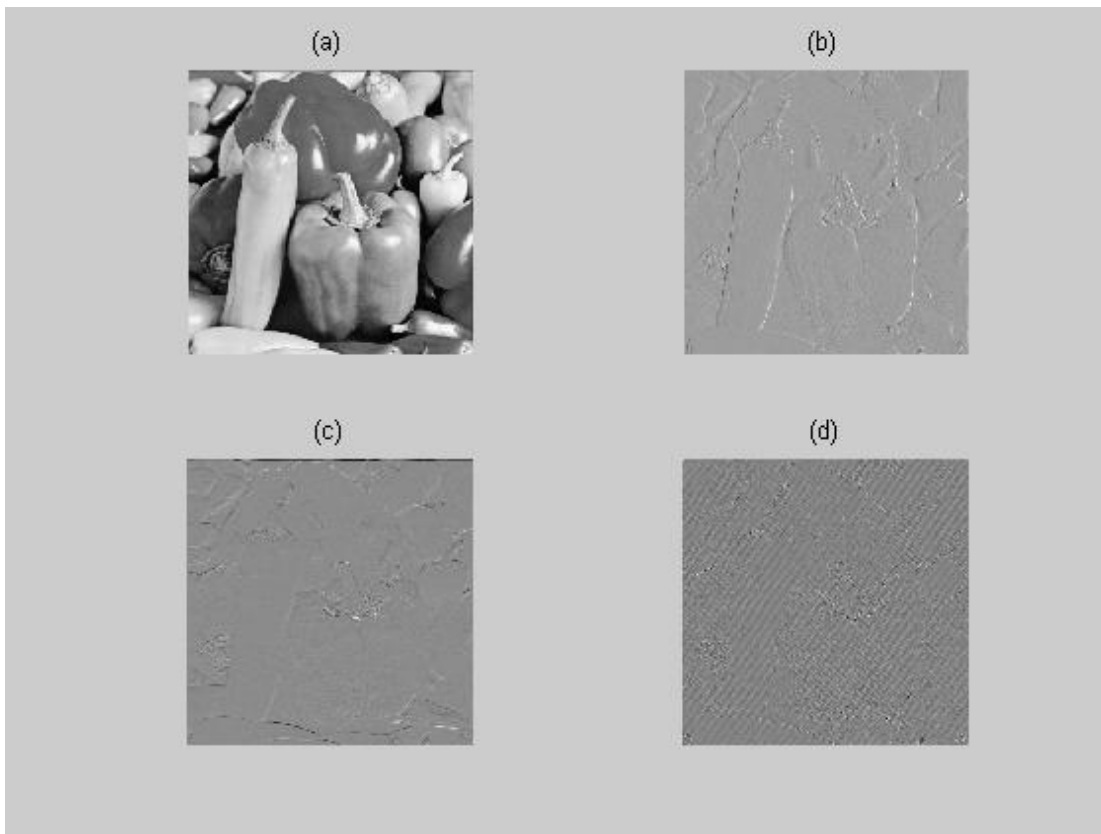


Fig 4: Level 1 Decomposition of the Peppers image, (Normalized for vision) (a) LL, (b) LH, (c) HL, (d) HH.

Before we look at the compression performance, let us define the receiver operating characteristics (ROC) of Phasemark™ 2K. We use two detection metrics PACE and PSR as defined in [4]. We first embed the signature (from LL) in the LH subband. Fig 5 shows how the detection values change with different embedding strengths (chosen bit plane i). It also shows how the quality of the watermarked image (represented by PSNR) changes with different embedding strengths. The embedding strength is determined by the i value. The combined picture then gives us the notion of the ROC. This determines the level of noise already introduced by the watermarking process itself and thus acts as a baseline case for the compression problem in hand. The image used in this experiment was the 'peppers' image.

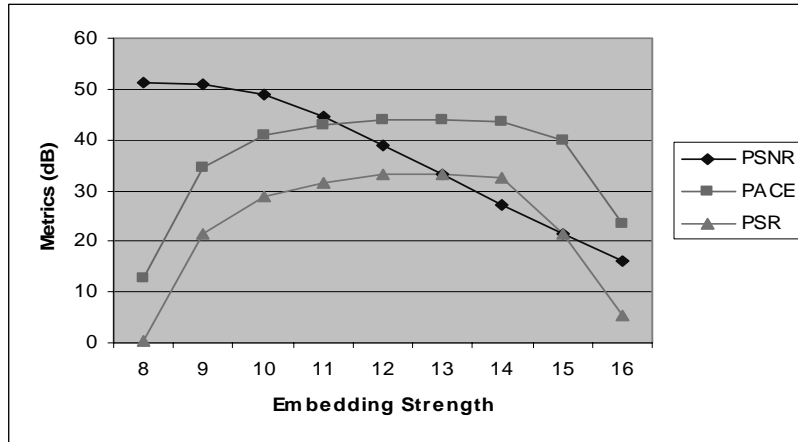
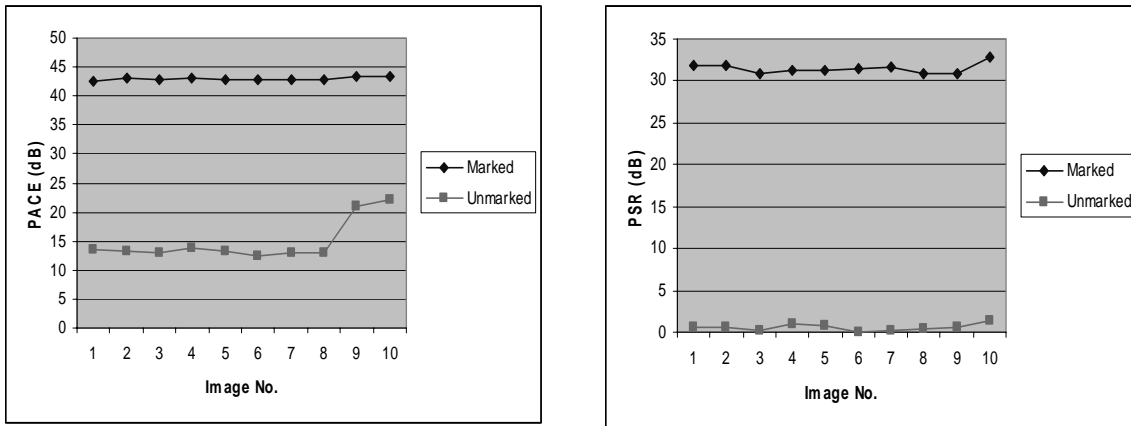


Fig 5. ROC of metrics values vs embedding strength for the peppers image.

Fig. 6 shows the detection metrics for the marked and unmarked images we used. Again, this result is when we embed the BPOF signature in the LH subband only. The embedding plane chosen is $i=12$, resulting in a PSNR value of approximately 39 dB.



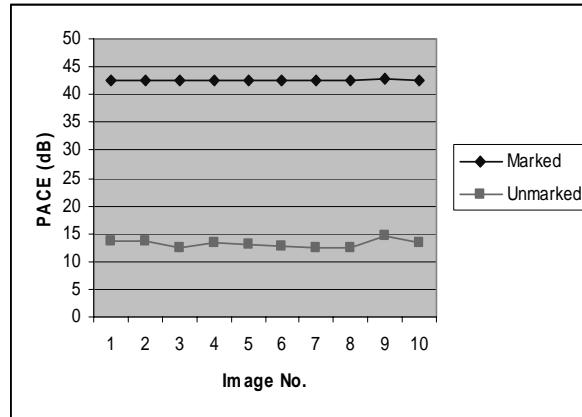
(a)

(b)

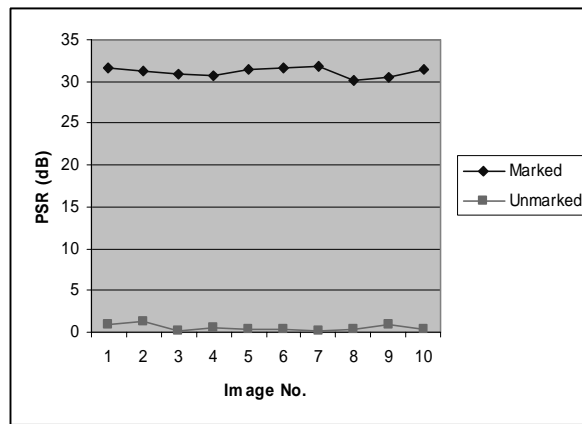
Fig 6. Authentication Performance for LH subband bit-plane embedding (a) PACE, (b) PSR.

Similar results were obtained when we embedded the BPOF signature in the HL subband.

In order to see which subband is friendlier to the watermarking process itself, we next embed the signature in the LL subband (again $i = 12$). The performance metrics are as shown in Fig. 7. It is to be noted that there is an overall performance enhancement (consider the false positives).



(a)



(b)

Fig 7. Authentication Performance for LL subband bit-plane embedding.

This can be explained by the distortion coming out from the watermarking process itself. It is well known that with wavelet decomposition the coefficients in LH or HL subbands are smaller than those in the LL subband. Therefore, when we embed the same signature in these bands, the relative changes in LH or HL subbands are larger than in the LL subband. Consequently, the detection performance is generally better in the LL subband embedding. This is somewhat offset by the increased embedding distortion coming out from the watermark process. *For all of our subsequent results, we use LL subband embedding.*

Next, we look at the compression robustness of the proposed Phasemark™ 2K method. We use the XNVIEW [9] public domain program that has a JPEG 2000 compression utility. XNVIEW uses the Jasper [10] SDK in their implementation. First, we compress the image with a desired quality factor to the .jp2 format. Then we re-convert it to the original format and then feed it to the detector. Fig. 8 shows the results. It shows that the stronger the watermark is, the more robustness it has against JPEG 2000 attack. Comparing it with Fig. 6, it can be easily seen that compression quality up to 40% can still be tolerated to discriminate a marked image from an unmarked one. (Again, strength is the i value).

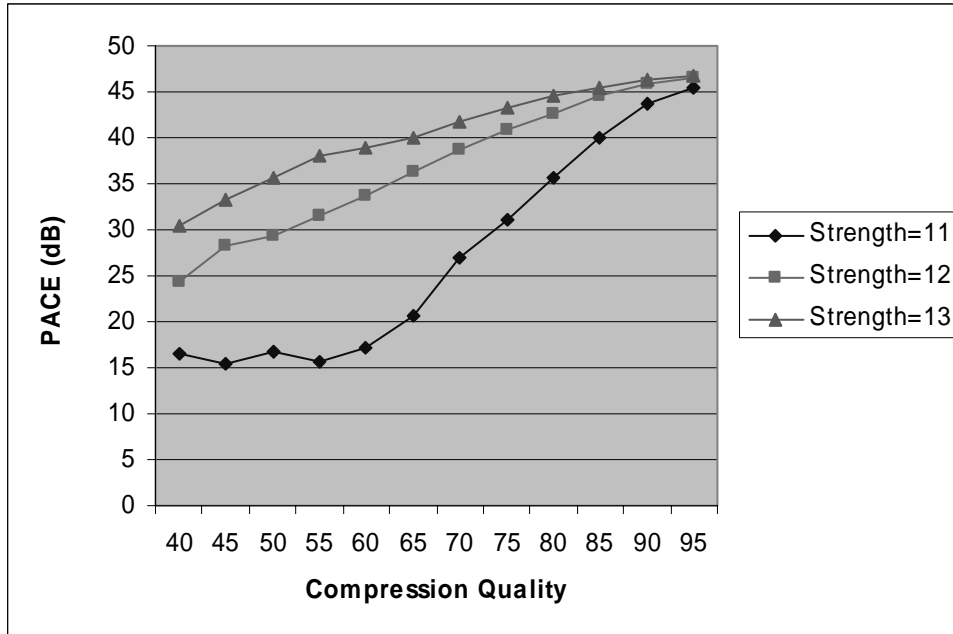
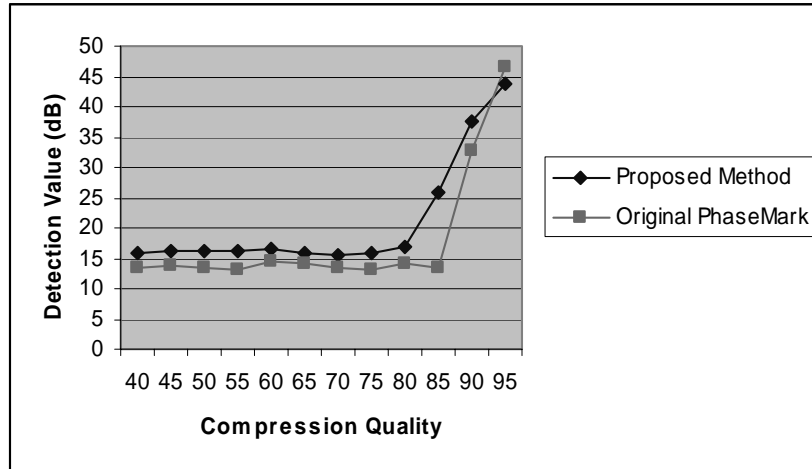


Fig 8. Authentication Performance against JPEG 2000 compression for the proposed method.

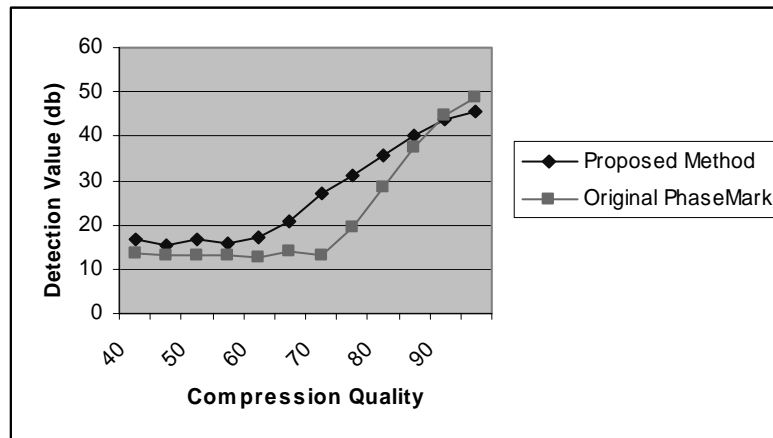
The results are obtained from the ‘peppers’ image, where both the signature and embedding subband is LL.

The three embedding strengths are motivated by the ROC analysis done in Fig. 5.

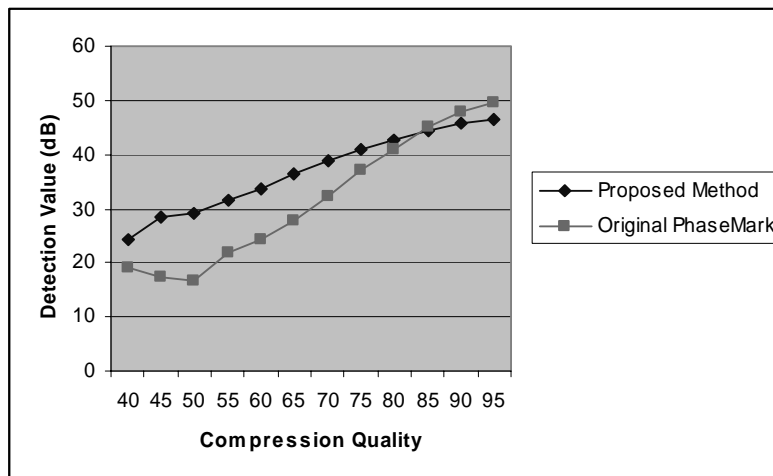
Fig 9 shows the performance enhancement of the proposed Phasemark™ 2K to the original Phasemark™ method. Results are furnished for three different quality of the watermarked image, corresponding to the embedding strengths of $i=10, 11, 12$ with PSNR values of 50.4dB, 45.1dB, 38.6dB, respectively.



(a)



(b)



(c)

Fig 9. Watermarked image (a) PSNR=50.4 dB, (b) 45.1 dB, (c) 38.6 dB.

Two things are clear from these results. First, the proposed method Phasemark™ 2K is uniformly better than the original Phasemark™ for low compression quality JPEG 2000 images. Second, depending upon the quality of the watermarked image, the proposed method yields slightly smaller detection value for very high-quality compression (above 90% for example). This is explained from the correlation operation used in the detection phase. With the first level wavelet decomposition, both the signature image and the embedding image are one fourth of the size of the original signal and embedding space. As a result, the correlation values are smaller now.

4. CONCLUSION

We provided some experimental results to show the enhanced compression tolerance against JPEG 2000 compression attacks using our wavelet-based bit-plane embedding technique Phasemark™ 2K. Future work will look into the performance for different wavelet filters and consider a fusion of detection values from different subband embeddings.

REFERENCES

1. Tong Liu, and Zheng-ding Qiu, "The survey of digital watermarking-based image authentication techniques," 6th International Conference on Signal Processing, 2002, Volume: 2 , 26-30 Aug. 2002, pp. 1556 - 1559 vol.2.
2. V. Fotopoulos and A.N. Skodras, "JPEG2000 parameters against watermarking," DSP 2002, 14th International Conference on Digital Signal Processing, 2002, Volume: 2 , 1-3 July 2002, pp. 713 - 716 vol.2
3. Chuhong Fei, D. Kundur, and R.H. Kwong, "Analysis and design of watermarking algorithms for improved resistance to compression," IEEE Transactions on Image Processing, Volume: 13, Issue: 2 , Feb. 2004, pp. 126 – 144.
4. Farid Ahmed and Ira S. Moskowitz, "Correlation-Based Watermarking Method for Image Authentication Applications," to appear: Opt. Eng., Aug 2004.
5. JPEG 2000 <http://www.jpeg.org/jpeg2000/>.
6. Peter Meerwald and Andreas Uhl, "A Survey of Wavelet domain Watermarking Algorithms," SPIE Symposium, Electronic Imaging, Conference on Security and Watermarking of Multimedia Contents, San Jose, CA, USA, January 20 - 26, 2001.
7. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, NY, 1998.
8. USC-SIPI Image Database, <http://sipi.usc.edu/services/database/Database.html>
9. XNView Software <http://www.xnview.com/>
10. M. D. Adams, "The Jasper Project Homepage," <http://www.ece.uvic.ca/~mdadams/jasper/>, 2002.