Geometric attack resistant watermarking in wavelet transform domain

Choong-Hoon Lee and Heung-Kyu Lee

Division of Computer Science, Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Guseong-Dong, Yuseong-Gu, Deajon, Korea

{chlee,hklee}@mmc.kaist.ac.kr

Abstract: In this paper, we propose an autocorrelation function (ACF) based watermarking scheme in the discrete wavelet transform (DWT) domain. Conventional ACF-based watermarking embeds a watermark in the spatial domain due to its detection mechanism. We show that the autocorrelation (AC) peaks, which play an important role in estimating the applied geometric attacks in ACF-based watermarking, can also be extracted by embedding the watermark in the DWT domain. In the proposed scheme, a periodic watermark is embedded in the DWT domain by considering the AC peak strength and noise visibility. The proposed scheme also deals efficiently with the image shift problem in the detection process by using the undecimated DWT. Experimental results show that the proposed scheme yields stronger AC peaks than the spatial domain scheme does and, as a result, shows improved robustness against combined geometric-removal attacks.

OCIS codes: (100.0100) Image Processing, (100.2000) Digital Image Processing.

References and links

1. J. J. K. O' Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," Signal Processing, **66**, 303-317 (1998).
2. M. Kutter, S. K. Bhattacharjee, and T. Ebrahimi, "Towards second generation watermarking schemes," in *Proceedings of IEEE Int. Conference on Image Processing* (Institute of Electrical and Electronics Engineers, New York, 1999), pp. 320-323.
3. S. Pereira and T. Pun, "Fast robust template matching for affine resistant image watermarking," in *International Workshop on Information Hiding*, LNCS **1768** (Springer-Verlag, Berlin, Germany, 1999), pp. 200-210.
4. M. Kutter, "Watermarking resisting to translation, rotation, and scaling." in *Multimedia systems and applications*, Proc. SPIE **3528**, 423-431 (1998).
5. I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. on Image Processing, **6**, 1673-1687 (1997).
6. M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," Signal Processing, **66**, 357-372 (1998).
7. J. S. Lim, *Two-dimensional signal and image processing* (Prentice Hall, New Jersey, 1990).
8. F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn. "Attacks on copyright marking systems," in *International workshop on information hiding*, LNCS **1525** (Springer-Verlag, Berlin, Germany, 1998), pp. 218-238.
9. R. Polikar, "The wavelet tutorial", http://users.rowan.edu/polikar/WAVELETS/WTtutorial.html
10. E. J. Stollnitz, T. D. DeRose, and D. H. Salesin, "Wavelets for computer graphics: A primer," IEEE Computer Graphics and Applications, **15**, 76-84 (1995).
11. A.B. Watson, G.Y.Yang, J.A.Solomon, and J.Villasenor, "Visual thresholds for wavelet quantization error," in *Human Vision and Electronic Imaging*, B. E. Rogowitz and J. P. Allebach, eds., Proc. SPIE **2657**, 382-392 (1996).

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

12. A. Gyaourova, C. Kamath, and I. K. Fodor, "Undecimated wavelet transforms for image de-noising," Technical report, Lawrence Livermore National Laboratory, UCRL-ID-150931 (2002).
13. G. Beylkin, "On the representation of operators in bases of compactly supported wavelets." SIAM J. Numer. Anal., **29**, 1716-1740, (1992).
14. M. Lang, H. Guo, J. E. Odegard, and C. S. Burrus, "Nonlinear processing of a shift invariant DWT for noise reduction," in *Mathematical Imaging: Wavelet Applications for Dual Use*, Proc. SPIE **2491**, 640-651 (1995).
15. M. Kutter and F. A. P. Peticolas, "A fair benchmark for image watermarking systems," in *Security and Watermarking of Multimedia Contents*, P. W. Wong and E. J. Delp, eds., Proc. SPIE **3657**, 226-239 (1999).
16. H. C. Huang, J. S. Pan, and H. M. Hang, "Watermarking based on transform domain," in *Intelligent Watermarking Techniques*, J. S. Pan, H. C. Huang, and L. C. Jain, eds. (World Scientific, Singapore, 2004), pp.147-163.
17. M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise Masking," IEEE Trans. on Image Processing, **10**, 783-791 (2001).
18. V. Darmstaedter, J.-F. Delaigle, J. J. Quisquater, and B. Macq, "Low Cost Spatial Watermarking," Comput. & Graphics, **22**, 417-424 (1998).
19. W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," in *Storage and Retrieval for Image and Video Database III*, Proc. SPIE **2420**, 165-173 (1995).
20. C. I. Podilchuk and W. J. Zheng, "Image-adaptive watermarking using visual models," IEEE Journal on Selected Areas in Communications, **16**, 525-539 (1998).
21. S. Voloshynovskiy, F. Deguillaume, and T. Pun, "Content adaptive watermarking based on a stochastic multiresolution image modeling," in *Tenth European Signal Processing Conference (EUSIPCO'2000)*, Tampere, Finland, Sept. 2000.
22. S. Voloshynovskiy, A. Herrigel, N. Baumgartner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," in *International Workshop on Information Hiding*, LNCS **1768** (Springer-Verlag, Berlin, Germany, 1999), pp. 212-236.
23. I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking* (Morgan Kaufmann Publishers, San Francisco, Calif., 2002).
24. T. Kalker, G. Depovere, J. Haitsma, and M. Maes, "A video watermarking system for broadcast monitoring," in *Security and Watermarking Multimedia Contents*, P. W. Wong and E. J. Delp, eds.,Proc. SPIE **3657**, 103-112 (1999).

1. Introduction

Geometric attacks are thought of as one of the most dangerous attacks in the digital watermarking world. Although several watermarking schemes that handle geometric attacks have been introduced [1, 3, 4, 2], each of them has problems.

ACF-based watermarking is known to have great potential for combating geometric attacks and normal signal processing attacks [21, 23]. It handles geometric attacks by embedding a periodic watermark pattern. Due to the periodicity, periodic peaks are found in the ACF of the watermark. The watermark detector estimates the applied geometric transform by referring to the peak pattern in the ACF of the extracted watermark. The watermark signal is detected after inverting the estimated geometric transform. Because of this two-phase detection mechanism, the correct detection of the AC peaks, as well as the watermark signal, is crucial for detection of the watermark. However, the AC peaks are not sufficiently robust.

Due to the geometric attack estimation mechanism, the watermark embedding and detection of the ACF-based watermarking have been performed in the spatial domain. Although transform domain watermarkings [18, 19] require higher computational complexity than do spatial domain watermarkings [6, 5, 20], it is generally known that they are more robust than spatial domain watermarkings [15, 16]. Thus, if we can implement ACF-based watermarking in the transform domain, we may achieve improved robustness.

To realize ACF-based watermarking in the frequency domain, the watermark embedding in that domain should produce periodic AC peaks in the spatial domain. It is not easy to satisfy this requirement with full frame transform, such as DCT (Discrete Cosine Transform) and DFT (Discrete Fourier Transform), since the transform coefficient change affects the entire image. Unlike full frame transforms, however, the DWT (Discrete Wavelet Transform) has spatial-frequency locality [9, 10]. It implies that signal embedding in the wavelet coefficient affects the

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

image locally. Thus, we hypothesized that the periodicity in the wavelet coefficients can also be extracted in the spatial domain. Experimental findings bore out this hypothesis. Figure 1 shows the ACF of the extracted signal from the spatial domain of the Lena image, in which a periodic signal is embedded in the wavelet subbands. We can see the periodic peaks in the figure.

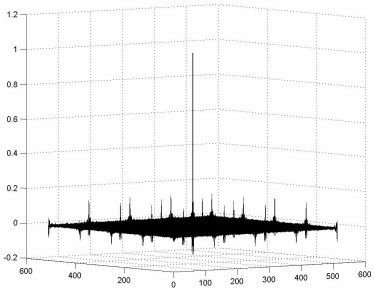


Fig. 1. AC peaks example of the Lena image marked in the DWT domain

In this paper, we present an ACF-based watermarking scheme that operates in the DWT domain. The proposed method embeds a periodic watermark pattern into the DWT domain. The geometric attack estimation is performed as in conventional ACF-based watermarking, by using the ACF of the watermark in the spatial domain. In the detection process, the undecimated wavelet transform is used to compensate for the image shift. In the experimental results, the proposed scheme shows stronger AC peaks and improved detection performance against geometric-removal attacks than does ACF-based watermarking in the spatial domain.

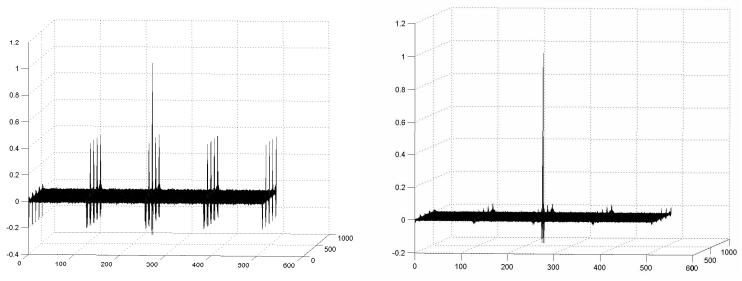
2. Watermarking algorithm

2.1. Watermark embedding in discrete wavelet transform domain

In this section, we explain how a watermark is embedded in the DWT domain. To determine the embedding strength for each sub-band level, we tested the strength and survivability of the AC peaks according to the sub-band level at which the watermark is embedded. Figure 2 shows the peak strength when a periodic watermark is embedded in the first-level sub-bands. As shown in the figure, the initial peak strength is very high. After JPEG compression, however, the peak strength is reduced drastically. By contrast, in the case of second-level embedding, as shown in Fig. 3, the initial peaks are less strong than those generated by the first-level embedding, but the peak strength is not much reduced after JPEG compression. Consequently, we can predict that the AC peaks generated by first-level embedding play an important role in geometric attack estimation when the marked images are not attacked or receive a weak attack. However, when strong attacks are applied, the peaks generated by second-level embedding will play the major role. Thus, to gain the maximum advantage, the watermarks are embedded in both levels.

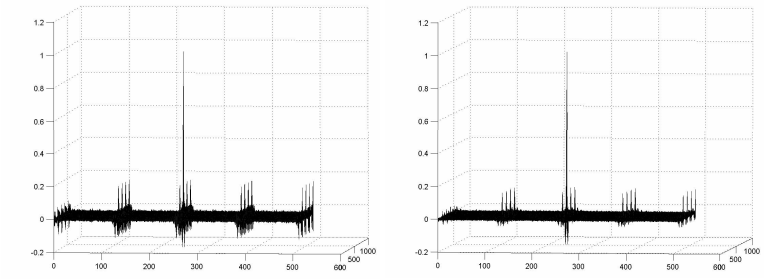
Figure 4 shows the embedding structure of the proposed scheme. The image is first decomposed by DWT up to the second level. In the figure, I_j^θ denotes the sub-band in the jth level in the direction θ ($\theta = 1$: horizontal, 2: diagonal, 3: vertical). To embed the watermark into two sub-band levels, two different periodic watermarks are generated. To obtain a period of $M \times M$ in the spatial domain, a watermark with period $\frac{M}{2^j} \times \frac{M}{2^j}$ is embedded in the jth level sub-bands. For the watermark pattern for the first-level sub-bands, a random number sequence of size $M/2 \times M/2$ that follows a standard normal distribution is generated with a user key. In the same way, a basic block of size $M/4 \times M/4$ is generated for the second-level sub-bands. Each watermark block is repeated up to the corresponding sub-band size.

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005



(a) Before JPEG (b) After JPEG

Fig. 2. Peak strength before and after JPEG compression by embedding a periodic watermark in the first-level sub-band



(a) Before JPEG (b) After JPEG

Fig. 3. Peak strength before and after JPEG compression by embedding a periodic watermark in the second-level sub-band

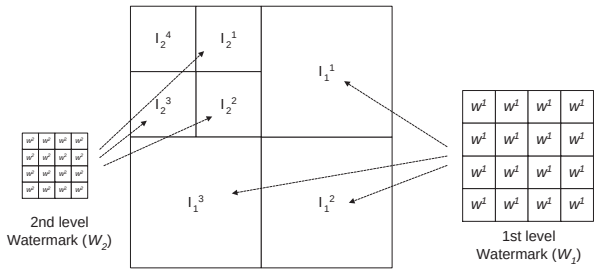


Fig. 4. Periodic watermark embedding in the DWT domain

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

The generated periodic watermark patterns W_1 and W_2 are embedded into the sub-bands I_1^θ and I_2^θ. In the service of image quality, the watermark is not embedded in I_2^4, which contains the DC components of the image. The watermarks are embedded as follows:

$$I_j^{\theta'}(x,y) = I_j^\theta(x,y) + \alpha\lambda_j^\theta(x,y)W_j(x,y) \tag{1}$$

where α and λ are global and local weighting factors, respectively.

A certain amount of research has already been done on visual masking models on the wavelet transform [20, 17, 11]. In the study reported in this paper, we applied the NVF (Noise Visibility Function) model [22] to the wavelet domain for the local weighting factor. The NVF is a function that grades the noise visibility in a defined image area by using the local texture information. The NVF has higher values in those regions in which noise is readily visible. Thus, we can control the strength of the watermark embedding by using the NVF. Since the DWT coefficients contain the local information, the NVF model can be adopted in the DWT domain. The NVF in the wavelet domain is calculated as follows:

$$NVF_j^\theta(x,y) = \frac{1}{1+\frac{D}{\sigma_{jmax}^{\theta 2}}\sigma_j^{\theta 2}(x,y)}, \tag{2}$$

where $\sigma_j^{\theta 2}(x,y)$ and $\sigma_{j_{max}}^{\theta 2}$ are local variances on (x,y) and the maximum of the local variance of the sub-band in direction θ and jth level sub-band. D is a user-defined constant in [50,100]. The higher the value of D, the higher the difference of the NVF values between in the plain region and in the textured region we have.

It is generally known that visual sensitivity to noise differs according to the direction of the sub-band. Noise in the sub-band that runs in the diagonal direction is more difficult to perceive than is noise in the sub-bands that run in the vertical and horizontal directions. We also use this property as a parameter for the weighting factor. The direction-based sensitivity is modeled by

$$\Theta^\theta = \begin{cases} \sqrt{2} & \text{if } \theta = 2 \text{ (diagonal direction)} \\ 1 & \text{otherwise} \end{cases}. \tag{3}$$

In this paper, we determine the weighting factor according to the sub-band level by considering the expected attack strength. If the marked image is not expected to be exposed to strong attacks, the watermark should be more strongly embedded in the first-level sub-bands. By contrast, if strong attacks are expected, then a higher embedding weight should be assigned to the second level. The weighting factor according to the sub-band level is denoted by L_j. For the experiments, we assigned a higher weight to the second-level sub-bands. For the experiment, we set $L_1 = 0.7$ and $L_2 = 1$.

Finally, we use the local weighting factor:

$$\lambda_j^\theta(x,y) = L_j\Theta^\theta[(1-NVF_j^\theta(x,y))\cdot S + NVF_j^\theta(x,y)\cdot S_1]. \tag{4}$$

S and S_1 are the user-defined weighting factors for textured and plain regions, respectively. $NVF_j^\theta(x,y)$, which has a value between 0 and 1, has a high value (near 1) in plain regions, and a low value (near 0) in textured regions. Therefore, in the equation, S_1 affects the embedding strength in plain regions more than S does. By contrast, S affects the strength in textured regions more. Therefore, S should be set to a higher value than S_1. For the experiment, we have set $S = 5$ and $S_1 = 1$.

2.2. Watermark detection using undecimated wavelet transform

The watermark detection follows the two-step detection mechanism of conventional ACF-based watermarking: (1) the geometric attack estimation and (2) watermark signal detection.

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

2.2.1. Geometric attack estimation

Geometric attacks are estimated by using the AC peaks of the estimated watermark signal. For this process, we should extract the watermark periodicity in the spatial domain. Although the watermark is embedded in the transform domain, due to the locality of the DWT, we can extract the watermark periodicity from the spatial domain by using high pass filters or noise removal filters. In our method, the periodic signal is extracted by using the Weiner filter [7]:

$$I^{-}(x,y) = \mu(x,y) + \frac{\sigma^2(x,y) - s^2}{\sigma^2(x,y)}(I(x,y) - \mu(x,y)), \tag{5}$$

where $\mu(x,y)$ and $\sigma^2(x,y)$ are the local mean and local variance of the original image, respectively. s^2 is the noise variance. Since the noise variance is not available, we uses the average of the local variances for s^2. The extracted signal E is given by

$$E = I - I^{-}. \tag{6}$$

Then, the extracted signal E is expected to have periodicity. To find out the periodicity, the ACF of the extracted signal E is calculated. The ACF can be calculated by FFT-based fast correlation calculation method [24] as

$$ACF = \frac{IFFT(FFT(E) \cdot FFT(E)^{*})}{|E|^2}, \tag{7}$$

where '*' denotes the complex conjugation. If the testing image is marked one, we can see a periodic peak pattern as in Fig. 1 in the ACF. The applied geometric attacks are estimated and reversed by using the AC peak pattern.

The AC peaks are detected from the ACF by applying an adaptive threshold as

$$ACF(x,y) > \mu_{acf} + \alpha_{acf}\sigma_{acf}, \tag{8}$$

where μ_{acf} and σ_{acf} denote the average and standard deviation of the autocorrelation function, respectively. α_{acf} is a user defined value. α_{acf} should be defined by considering the false negative and false positive error rate. Supposing that AC values of non-peaks in ACF follow a normal distribution N(μ_{acf}, σ_{acf}), we can calculate the false positive error rate as follows. If we define a random variable X that follows a standard normal distribution N(0,1), the probability that an AC value is higher than $\mu_{acf} + \alpha_{acf}\sigma_{acf}$ equals the probability that X is higher than α_{acf}. Thus, the false positive error rate of the AC peak detection when the threshold is $\mu_{acf} + \alpha_{acf}\sigma_{acf}$ is calculated by

$$P_{fp_{AC}} = P(AC_{\text{non-peak}} > \mu_{acf} + \alpha_{acf}\sigma_{acf}) = P(X > \alpha_{acf}) = \int_{\alpha_{acf}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) dx, \tag{9}$$

where $P(A)$ denotes probability of the event A. $AC_{\text{non-peak}}$ is a random variable that follows the normal distribution N(μ_{acf}, σ_{acf}).

The geometric attack is estimated by finding the base peak pair from the detected AC peaks. Here, we name the pair of the nearest two peaks (in vertical and horizontal direction) from the center of the ACF the *base peak pair*. An example is shown in Fig. 5. Using the offset information of the base peak pair, we can calculate the period of the watermark and rotation angle.

The base peak pair is found as follows. Since the peaks are distributed periodically, if we know the base peak pair, we can find all other peaks in the ACF by using the base peaks offset

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

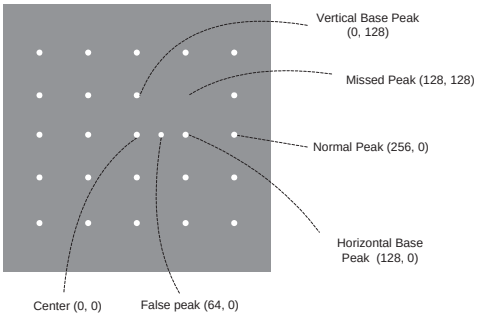


Fig. 5. Peak example for geometric transform estimation algorithm

information. For example, if the peak pair on [(0,128),(128,0)] is the base peak pair, we know that there will be peaks on (128,128), (256, 0), (0, 256), and so on. The base peak pair is found by using this property. For every possible peak pair, we count the number of peaks that can be found by using the peak pair. We will call this number the *peak count* of the peak pair. Then, we can select the peak pair with the highest peak count as the base peak pair. This method works well in normal cases, but there do exist cases in which errors occur. For example, a peak may be falsely detected. In Fig. 5, there exists a false peak on (0, 64). In such a case, the peak pair on [(0,64),(128,0)] is selected as the base peak pair, since every peak that can be found by the peak pair on [(0,128),(128,0)] can be found also by the peak pair on [(0,64),(128,0)]. In order to avoid this problem, we introduce another term, *peak ratio*, which refers to the ratio of the number of actually found peaks to the number of expected peaks with the testing peak pair as

$$\text{Peak Ratio} = \text{Peak Count}/\text{Expected Peak Count}. \tag{10}$$

The *expected peak count* (the number of expected peaks) of a peak pair can be calculated by referring to the image size and the offset of the peak pair. For example, suppose that the testing peak pair is on [(0,128),(128,0)] and the image size is 512×512. Then, if the testing peak pair is the base peak pair, there must be $\frac{512}{128} \times \frac{512}{128} = 16$ peaks in the ACF in the ideal case. Thus, the expected peak count of the testing peak pair is 16.

Using the peak count and peak ratio, we can find the base peak pair by defining another term, *weighted peak count*, as follows.

$$\text{Weighted Peak Count} = \text{Peak Count} \times \text{Peak Ratio}. \tag{11}$$

Then, the peak pair that has the highest weighted peak count is selected as the base peak pair.

In the above example, although the peak count of the peak pair on [(0,128),(128,0)] is lower by 1 than that of the peak pair on [(0,64),(128,0)], the peak ratio is about twice of that of peak pair on [(0,64),(128,0)]. Thus, the peak pair on [(0,128),(128,0)] is selected as the base peak pair.

Finally, geometric attacks, such as rotation, scaling, and aspect ratio change are estimated and reversed by using the offset information of the selected base peak pair.

2.2.2. Watermark signal detection

The watermark signal is detected from the DWT sub-bands of the geometrically restored image. The geometric attack estimation method described in the previous section does not handle image shift. Thus, we should try to detect the watermark from every possible shifted version of the image. In the spatial domain method, this operation can be performed efficiently by using

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

the FFT-based correlation calculation [24]. The problem is that DWT is not shift-invariant. That is, a shift in the spatial domain does not entail a shift in the DWT domain. Therefore, when the marked image is shifted, the image should by transformed by DWT on every possible shift in order to detect the watermark properly. This requires enormous computing time.

Several studies have been done on shift-invariant wavelet transform. One well-known approach is the undecimated wavelet transform [12, 14, 13]. Normally, the shift-variant property of the wavelet transform is caused by the decimation process. After wavelet transform, we have two sub-bands, each of which is half the size of the original signal. Since the decomposed sub-bands have only half the precision of the original, they cannot represent every shift in the spatial domain. If a signal is shifted by an odd offset, the wavelet transform result is completely different from that of the original signal. However, if the shift is even, the wavelet transform result is the shifted version of that of the original signal. By using this property, the undecimated DWT achieves shift invariance. For example, if we have two versions of the wavelet transform results of a signal (by transforming the signal directly and by transforming the signal after shifting by an odd offset), then we can express every possible (even and odd) shift in the spatial domain by shifting the sub-bands of one of the transform versions. The *shift4* algorithm is a 2D expanded undecimated DWT [12]. The *shift4* algorithm produces four wavelet transform results from a non-shifted, a horizontally 1 pixel shifted, a vertically 1 pixel shifted, and a diagonally 1 pixel shifted image. With the four transform results, we can express every possible shift in the spatial domain.

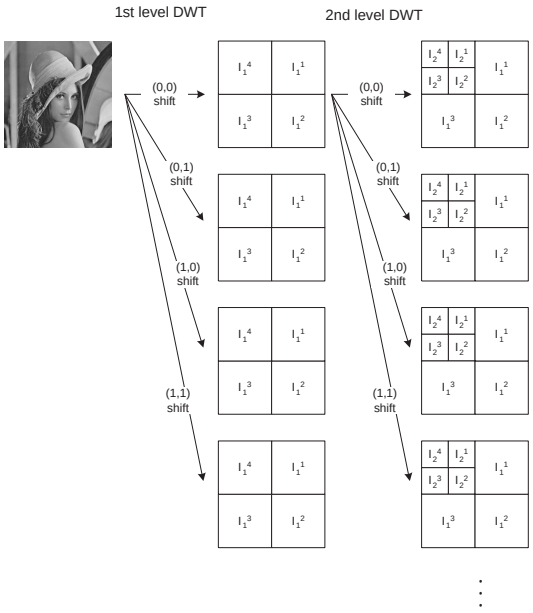


Fig. 6. Image decomposition by *shift4* algorithm

To detect the watermark from the shifted image, the marked image is decomposed first by the *shift4* algorithm up to the second level. After first-level wavelet decomposition, we have four transformed images. The lowpass sub-band in each transform result is transformed again by the *shift4* algorithm. Finally, we have 16 transform results. This process is shown in Fig. 6. By shifting the sub-bands of one of the 16 transform results by appropriate offset, we can express every possible shift in the spatial domain.

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

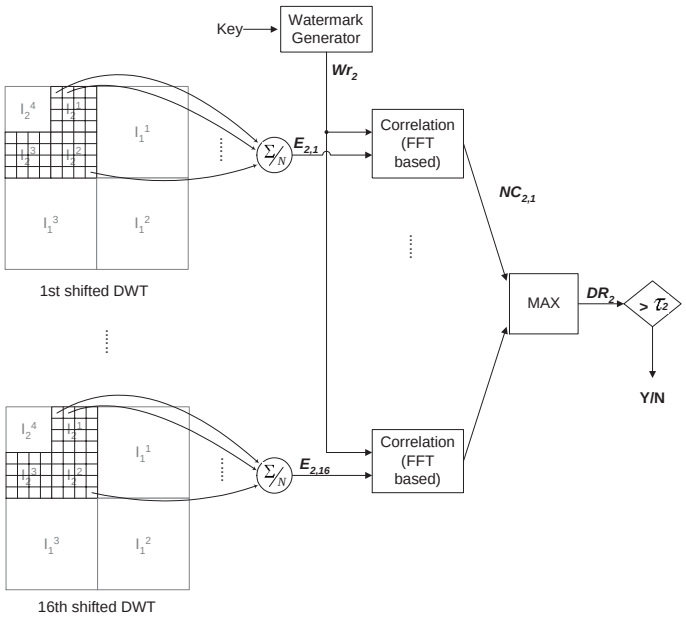


Fig. 7. Correlation based detection process in the second level subband

The embedded watermark is detected from the first- and second-level sub-bands in each transform result. Figure 7 shows a detection process in the second-level sub-bands. At first, the sub-bands that contain a watermark signal are segmented into a basic watermark pattern size ($M/4 \times M/4$ for the second level and $M/2 \times M/2$ for the first-level). The average of all segments in each transform result is calculated. The watermark is detected by calculating the correlation between the segment average $E_{j,k}$ and the reference watermark pattern Wr_j while shifting the segment average by every possible shift. k denotes the transform result index. ($1 \leq k \leq 16$ for the second level and $1 \leq k \leq 4$ for the first-level.) This process can be performed in reduced time with FFT by

$$NC_{j,k} = \frac{IFFT(FFT(E_{j,k}) \cdot FFT(Wr_j)^*)}{|E_{j,k}|\,|Wr_j|}. \tag{12}$$

Among all possible shifts in all transform results, only one shift is valid and, therefore, the correlation on the valid shift has the highest value. Thus, we find the maximum correlation (detector response) value among every possible shift for each sub-band level as

$$DR_j = \max_{x,y,k}\{NC_{j,k}(x,y)\}. \tag{13}$$

Finally, the decision for the watermark detection is made by

$$DR_1 > \tau_1 \ \text{ or } \ DR_2 > \tau_2, \tag{14}$$

where τ_j is the threshold given adaptively by

$$\tau_j = \mu_{nc_j} + \alpha_{nc_j}\sigma_{nc_j}, \tag{15}$$

where μ_{nc_j} and σ_{nc_j} are the average and standard deviation of $NC_{j,k}$, respectively. α_{nc_j} is a user defined value. α_{nc_j} should be set also by considering the false positive and false negative error rate of the watermark detection. Differently from the AC peak detection, the watermark detection uses the maximum value among the correlations. Thus, the calculation

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

method of the false positive error rate is a little different. If we suppose that the correlation values between an unmarked block and the reference pattern follow a normal distribution, the probability that each correlation value is higher than the threshold can be calculated in the same way as in Eq (9). (Let $P_{fp_{NC}}$ denote this probability.) The probability that the maximum value in a group of correlation values is higher than the threshold equals $1-P(\text{all correlation values are less than the threshold})$. Thus the false positive error rate is calculated by

$$P_{fp_{max}} = 1-(1-P_{fp_{NC}})^R, \tag{16}$$

where R is the number of correlation values.

Although undecimated wavelet transform reduces the computing time, the computational cost is still great for the watermark detector since 16 DWTs of full image are required. However, the computing time can be reduced further by reordering the detection process. In the original detection method, the marked image is decomposed by *shift4* algorithm first, and then the result sub-bands are segmented and averaged. Since the DWT is a linear transform, we can reorder this process. That is, (1) the image is segmented and averaged first, and (2) the averaged block is transformed by *shift4* algorithm. This reordering reduces the computation time drastically by reducing the size of input data for the DWT.

In the reordered method, the marked image is segmented into blocks of size $M \times M$ $(b_1, b_2, \ldots, b_N)$. Then, the average of the blocks is calculated by

$$b_{avg}(i,j) = \frac{1}{N}\sum_{n=1}^{N} b_n(i,j), (1 \le \mathrm{i,j} \le M). \tag{17}$$

Then, the average block b_{avg} is transformed by the *shift4* algorithm up to the second-level. Then, we can calculate $E_{j,k}$ by averaging the three directional sub-bands (horizontal, vertical, and diagonal) in each level (j) in each transform result (k). The watermark is detected from with $E_{j,k}$ by Eq. (12) - (14). This reordered detection method yields the same result as the original method described above. The reordered method transforms the block of size $M \times M$ by *shift4* algorithm while the original method transforms the full image. This results in computing time being much reduced.

3. Experimental results

This section presents the experimental results for the proposed watermarking scheme. We tested the AC peak strength and watermark signal detector response, and the watermark detection performance following geometric-removal attacks.

To compare the performance of the proposed method with spatial domain watermarking, the latter method is modeled as

$$I' = I + \alpha_s \lambda_s W_s, \tag{18}$$

where α_s and λ_s denote the global and local weighting factors, respectively. For λ_s, we used the NVF-based weighting factor:

$$\lambda_s = (1-NVF)\cdot S + NVF \cdot S_1. \tag{19}$$

The NVF is calculated in the spatial domain by the same method as Eq. (2). W_s is the periodic watermark pattern of 128×128 period. The basic watermark block is a random number sequence with standard normal distribution. During the watermark detection, the geometric attack is estimated in the same method as the proposed scheme. After the estimation, the extracted signal E in Eq. (6) is restored into the original geometry. The restored signal is segmented into blocks of size 128×128, and the blocks are averaged. Then, the watermark is detected from

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

the average block by using the maximum correlation between the average block and reference watermark pattern as in the proposed scheme. The FFT based correlation calculation is also used here.

For the proposed method, 64×64 and 32×32 period watermark patterns are used for the first and second sub-band levels, respectively, to obtain a 128×128 period in the spatial domain.

3.1. Time complexity analysis

Since both the proposed scheme and spatial scheme have the same geometric attack estimation step, we only compare the computing time of the watermark signal detection step. In the reordered detection method, the proposed scheme has four DWTs of $M \times M$ size blocks (first-level decomposition) and 16 DWTs of $M/2 \times M/2$ size blocks (second-level decomposition). For computing the correlation, three FFTs are required for each $E_{j,k}$. Since the orders of complexities of FFT and DWT of $N \times N$ size block are O($N^2 \log N$) and O(N^2), respectively, the approximate computing time for the watermark signal detection is

$$\begin{aligned} & 4M^2 + 16\left(\tfrac{M}{2}\right)^2 + 3 \times 4 \times \left(\tfrac{M}{2}\right)^2 \log\left(\tfrac{M}{2}\right) + 3 \times 16 \times \left(\tfrac{M}{4}\right)^2 \log\left(\tfrac{M}{4}\right) \\ & = 8M^2 + 3M^2(\log M - \log 2) + 3M^2(\log M - \log 4) \\ & = 6M^2 \log M - M^2. \end{aligned} \tag{20}$$

Since the spatial domain method requires three FFTs of $M \times M$ size blocks to detect the watermark signal, the computing time of the spatial domain method is approximately $3M^2 \log M$. Thus, the computing time of the watermark signal detection of the proposed scheme is less than twice of that of the spatial domain method, and both schemes have the same order of complexity O($M^2 \log M$).

If we consider the geometric attack estimation step, the computing time gap is very small. For calculating an ACF in the geometric attack estimation step, three FFTs of the image of size $X \times Y$ are required. Thus, the computing time of this process is approximately $3XY(\log Y + \log X)$. Since $X, Y >> M$, the watermark signal detection step occupies a minor portion of the overall computing time. Thus, considering the detection procedure as a whole, the computing time gap between two schemes is minor. Moreover, since M is fixed in a watermarking system, the difference is constant.

3.2. Robustness test of the AC peaks and watermark signal

In this section, we test the robustness of the AC peaks and watermark signal. Since geometric attacks are estimated by using AC peaks, we can predict robustness to them by testing the AC peak strength.

We tested the strength of the AC peaks and watermark signal after JPEG compression, which is one of the most popular watermark attacks. For this test, we used 700 photo images (512×512 size) that were collected randomly from the internet. The images are marked by each scheme. The average PSNR (Peak Signal to Noise Ratio) of the marked images was 38dB. The PSNR between the original image I and marked image I' of size $X \times Y$ is calculated by

$$PSNR = 20 \log_{10}\left(\frac{255}{\sqrt{\frac{1}{XY}\sum (I(x,y) - I'(x,y))^2}}\right). \tag{21}$$

Fig. 8 shows the histogram of the AC peak values of both schemes after JPEG quality 50% compression. As seen in the figure, in neither scheme are the AC peak values separated well from the non-peak values. However, the proposed scheme shows better separation and higher AC peak values than the spatial method. The average peak strengths of the proposed scheme

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

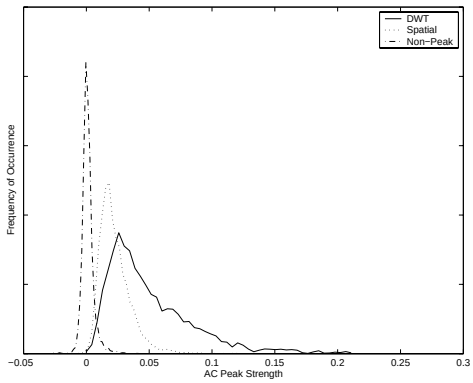


Fig. 8. Distribution of AC peaks after JPEG quality 50% compression. (The histogram of non-peaks is scaled down vertically for better illustration.)

and spatial method are 0.0504 and 0.0228, respectively. That is, the AC peaks can be detected with lower error probability by the proposed scheme than by the spatial method. Consequently, the proposed scheme is expected to show better geometric attack estimation capability than the spatial method.

The histogram of the detector responses is shown in Fig. 9. Differently from the AC peak results, the watermark detector responses show a clear separation between correct mark and no watermark, except for first-level sub-band results for the DWT method. In these test results, the spatial domain method can detect the watermark clearly. The proposed scheme can also detect watermark well from the second-level sub-bands.

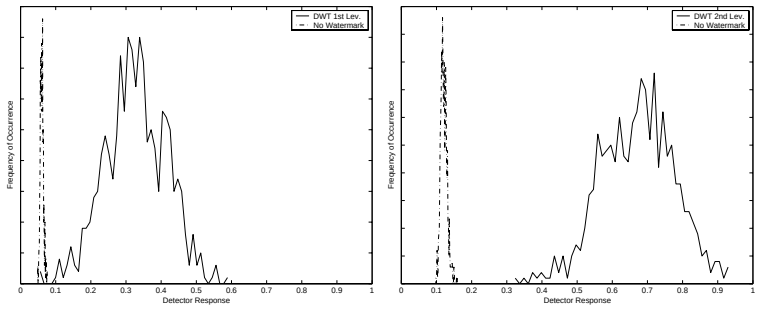
We analyzed the error probability of AC peak and watermark detection after JPEG 50% compression with the ROC (Receiver Operating Characteristic) curve. In order to calculate the ROC curves, we first found an appropriate theoretical distribution model for each histogram. Figure 10 shows the distribution models for the detector response and AC peak strength. In the figure, we can see that the measured histogram of the detector responses is a good fit with the normal distribution model. Differently from the detector responses, the AC peak strengths do not follow a normal distribution. We found that the histogram of the AC peak strength fits well into the gamma distribution model. In the same way, we used the normal distribution model for detector responses from unmarked images and the AC values of the non-peaks. The ROC curves were calculated by using these theoretic distribution models.

The ROC curves for the AC peaks and watermark detection after JPEG 50% compression are presented in Fig. 11. In the figure, we can see that the DWT domain method shows much lower error probability of the AC peaks detection than the spatial domain method does. The EER (Equal Error Rate) of the AC peak detection of the DWT domain method (0.0894) is less than half that of the spatial domain method (0.2268). (The EER refers to the error rate when the false positive and false negative error rates are the same.)

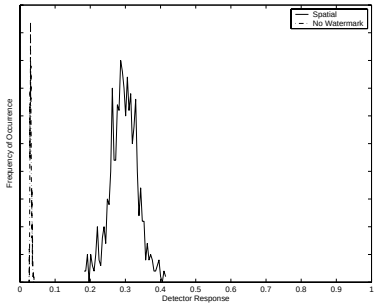
By contrast, the watermark detection error rate of the proposed scheme is a little higher than that of the spatial domain scheme. Although the proposed scheme shows higher detector responses in the second level sub-band than the spatial domain method as in Fig. 9, the DWT method yields poorer ROC curves, because the variance of the detector responses is higher than with the spatial method. Nevertheless, the error rate of the DWT domain method in the second sub-band level is still very low (EER$\approx 1.43 \times 10^{-5}$), despite the JPEG compression.

Overall, the error probability of the AC peak detection is much higher than that of the watermark signal detection. Thus, success in watermark detection depends more on the detection of

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005



(a) DWT 1st level subband (b) DWT 2nd level subband



(c) Spatial domain

Fig. 9. Distribution of the detector responses after JPEG 50% compression. (The histogram of the response from unmarked images is scaled down vertically for better illustration.)

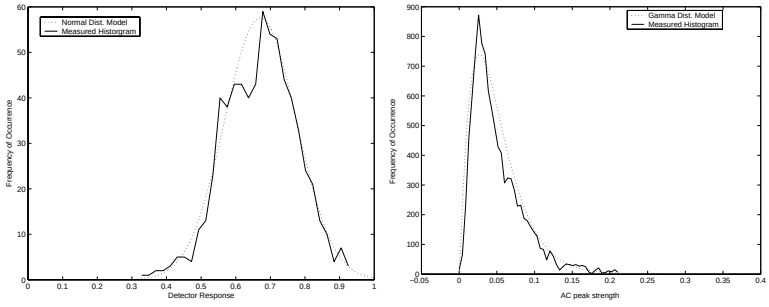
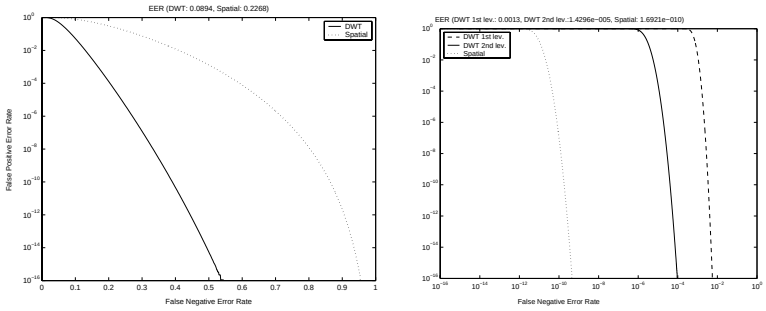


Fig. 10. Theoretical distribution models for detector responses and AC peaks. (a) Distribution model for detector response from DWT second level, (b) Distribution model for AC peak strength of the DWT watermarking.

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005



(a) AC Peak Detection (b) Watermark Detection

Fig. 11. ROC curves of the AC peak and watermark detection after JPEG quality 50% compression.

the AC peak than on that of the watermark signal. As seen in the results, the proposed scheme shows stronger AC peaks than the conventional spatial domain scheme. Consequently, we can expect that the proposed scheme will show better watermark detection performance after geometric attacks.

3.3. Watermark detection test against geometric attacks

In this section, we show the actual watermark detection results after geometric-removal attacks. The Stirmark benchmarking tool [8] was used for the geometric attack tool. The stirmark provides several geometric attacks: Row-column removing (5), cropping (9), flip (1), linear geometric distortion (3), aspect ratio change (8), rotation (16), rotation + scale (16), scale (6) and shearing (6). (The numbers in the parenthesis denote the number of attacks in each class.)

The 15 images in Fig. 12 were marked first by both schemes (PSNR = 38dB). The marked images were each attacked by the stirmark geometric attack and JPEG 50% quality compression. The detection tests were performed on the attacked images.

For the detection threshold, we set $\alpha_{acf} = 3.5$ in Eq. (8) and $\alpha_{nc_j} = 6$ in Eq. (15). With these value, the false positive error rates of the AC peak detection and watermark signal detection are about 2.3×10^{-4} by Eq. (9) and 1.6×10^{-5} by Eq. (16), respectively. (If we do not consider the maximum correlation finding in the watermark signal detection, the probability that each correlation value from unmarked image is higher than the threshold is about 9.9×10^{-10}.) We set the threshold for AC peak detection a little low because the AC peaks are vulnerable to attacks and the process of geometric attack estimation in Section 2.2.1 can work well even with a few false peaks.

Table 1 shows the detection results. For every attack class, the proposed scheme showed better detection results than the spatial method. In all tests, the watermark signals remained in the images following the attacks. All detection failures were caused by AC peak detection failure. Since the proposed scheme yields stronger AC peaks, it showed better detection results after geometric-removal attacks. Totally, among 1050 detection tests, the DWT domain method succeeded in 881 tests while the spatial domain scheme showed 664 successes.

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005

Fig. 12. Test images for the watermark detection experiment

Table 1. Watermark detection results after Stirmark geometric attacks and JPEG 50% compression. The number in the parentesis is the total number of attacks for each attack class. For example, R-C Remove has 75 attacks (15 images $\times$ 6 attacks).

	DWT	Spatial
R-C Remove (75)	65	57
Cropping (135)	135	134
Flip (15)	15	15
Linear transform (45)	35	31
Ratio Change (120)	107	76
Rotation (240)	187	119
Rotation & Scaling (240)	194	118
Scaling (90)	65	52
Shearing (90)	78	62
Total (1050)	881	664

4. Conclusion

In this paper, we presented a new ACF-based watermarking method in the DWT domain. Because of the detection mechanism, conventional ACF-based watermarking has been restricted to spatial domain methods. We found that AC peaks can be also extracted by embedding a periodic watermark pattern into the DWT domain. The AC peak strength and survivability according to the embedding sub-band level are investigated and considered for the watermark embedding. A periodic watermark is embedded in wavelet sub-bands by considering noise visibility, and geometric attacks are estimated as in conventional ACF-based watermarking. By adopting the undecimated wavelet transform, we also solved the shift handling problem in the detection step. The AC peaks were able to survive attacks better under the proposed scheme than under the spatial domain method. The proposed scheme showed better detection performance against geometric, especially combined geometric-removal attacks, than the spatial domain method.

Acknowledgments

This work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Advanced Information Technology Research Center (AITrc).

#6076 - $15.00 US Received 16 December 2004; revised 11 February 2005; accepted 16 February 2005