ROBUSTNESS AGAINST UNAUTHORIZED WATERMARK REMOVAL ATTACKS VIA KEY-DEPENDENT WAVELET PACKET SUBBAND STRUCTURES*

Werner M. Dietl[1] *and Andreas Uhl*[2]

[1]Software Component Technology Group, Department of Computer Science, ETH Zürich
ETH Zentrum, RZ J8, 8092 Zürich, SWITZERLAND
e-mail: dietlw@inf.ethz.ch

[2]Department of Scientific Computing, Salzburg University
Jakob-Haringerstr.2, A-5020 Salzburg, AUSTRIA
e-mail: uhl@cosy.sbg.ac.at

ABSTRACT

We propose the use of random wavelet packet decompositions as a way to increase the security of watermarking systems against unauthorized removal attacks. Experimental attacks based on coefficient quantization show that using a secret key-dependent subband structure to hide the watermarking domain significantly increases the watermark correlation under an attack as compared to a classical pyramidal wavelet watermark.

1. INTRODUCTION

Watermarking has gained high popularity as a method to protect intellectual property rights on the Internet [1, 2]. Watermark robustness denotes the ability of inserted watermarking data to withstand non-specifc cover data manipulations or attacks. This property has been investigated in great detail and several tools have been developed to assess it (e.g. Stirmark and Checkmark).

In contrast to robustness, watermark security aims at withstanding specific attacks which exploit knowledge about the watermark embedding procedure. According to Kerkhoffs principle, security may be achieved by employing secret key data during the watermark embedding stage. In previous work, pseudo-random skipping of coefficients has been proposed to provide watermark security by Wang [3] or Kundur [4], but skipping significant coefficients reduces the capacity of the systems. Fridrich [5] introduced the concept of key-dependent Fourier-like basis functions in order to protect a watermark from hostile attacks. However, the approach suffers from the storage requirements for generating numerous orthogonal patterns of the size of the host image. To provide key-dependent basis functions in the wavelet domain, parametrized wavelet filters were proposed by Dietl et al. [6] and the resistance of the resulting scheme against unauthorized detection and removal attacks has been shown.

A second possibility to employ key-dependent basis functions in the wavelet domain is to embed the watermark sequence using a secret wavelet packet decomposition [7] and to use the subband structure as embedding key. Wavelet packets have not found too much attention in the watermarking community yet. Wang [8] uses a fixed non-standard decomposition to embed a watermark sequence in the middle frequencies of an image to achieve better robustness. The algorithm by Tsai [9] uses wavelet packets, but the selection is not specified and no experimental results are provided. Vehel [10] uses the wavelet packet decomposition structure itself as the watermark sequence. In recent work [11] we have shown a watermarking system based on secret wavelet packet subband structures to be robust against unauthorized detection attacks. In this work, we focus on unauthorized removal attacks. Section 2 reviews the employed system and gives some performance results related to resistance against unauthorized detection. In Section 3 we entirely focus on robustness against unauthorized removal attacks, Section 4 concludes the paper.

2. WATERMARK EMBEDDING IN KEY-DEPENDENT SUBBAND STRUCTURES

The basic system design is shown in Fig. 1. For the forward wavelet transformation we use a secret wavelet packet tree and then embed the watermark in the generated wavelet coefficients. After embedding the watermark we apply the inverse transformation using the same wavelet packet tree to generate the watermarked image.

Our experimental system is based on an algorithm proposed by Wang et al. [3] which embeds the watermark se-

*This work has been partially supported by the Austrian Science Fund, project no. 15170

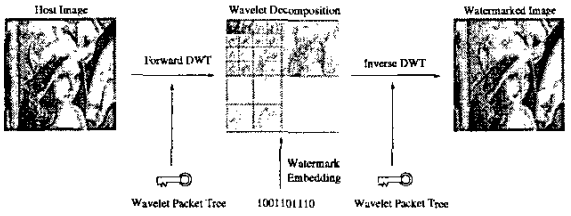


Fig. 1. Basic system design

quence based on Successive Subband Quantization (SSQ). Within a selected subband all unselected coefficients $C_s(x, y)$ that are larger than a threshold T_s are used to embed a watermark element W_k according to

$$C'_{s,k}(x, y) = C_s(x, y) + \alpha_s \beta_s T_s W_k \ . \qquad (1)$$

The two factors α_s and β_s are used to determine the embedding strength of the algorithm.

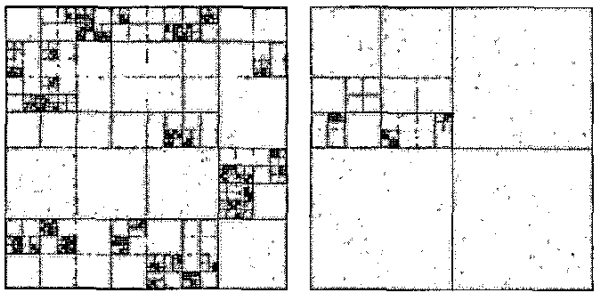
2.1. Generation of Subband Structures

The wavelet packet tree is generated by a random process that depends on a secret seed number. We use two methods to randomly construct a tree.

1. Decomposition 1: a 50% probability is used for each subband to decide whether it should be further decomposed or not. This decomposition strategy gives us the full range of possible decomposition trees, but could also result in trees that are poorly suited for watermark embedding (e.g. when only high frequency subbands are used). See Fig. 2.a for an example.

2. Decomposition 2: the focus is on building a wavelet tree that has a good resolution of the middle and low frequencies, which are best suited for watermark embedding. No decomposition of the three detail subbands on the first level (HL_1, LH_1 and HH_1) is performed, only the first approximation subband is further decomposed. Therefore, more emphasis is put on decomposing in the middle frequencies. See Fig. 2.b for an example.

The size of the keyspace available for embedding corresponds to the vast number of possible wavelet packet trees. According to [12], for a decomposition with n+1 levels there are $f(n) = \sum_{i=0}^{4} \binom{4}{i} \cdot (f(n-1))^i$ possible trees ($f(0) = 1$). Considering decomposition 1, for 4 decomposition levels this results in around 2^{65} trees and for 7 levels around 2^{4185} trees are possible. Using decomposition 2 we basically loose all the trees that are below the three top-level detail subbands. Therefore we have only around 83521 trees for 4 levels, but still around 2^{1046} trees for 7 levels.

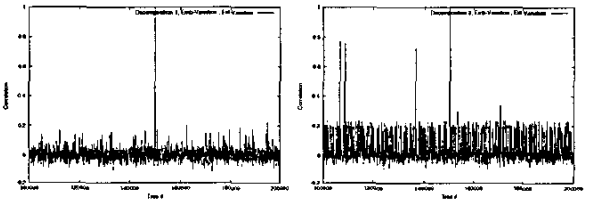


(a) Decomposition 1 (b) Decomposition 2

Fig. 2. Wavelet packet subband structures as used for watermark embedding

2.2. Robustness against Unauthorized Detection

For a security assessment we embed a 1000 element watermark with 40dB PSNR into the Lena image. We use a fixed tree for embedding and then use a set of 20000 randomly chosen trees with which we try to detect the watermark.



(a) Decomposition 1 (b) Decomposition 2

Fig. 3. Security assessment using 7 decomposition levels

Fig. 3 compares the response for decomposition 1 and 2 with 7 decomposition levels. For decomposition 1 there is one clear peak at the correct tree also used for embedding and low correlation for all other trees. However, for decomposition 2 there is one peak at the position of the correct tree but also three other trees with more than 60 percent correlation. There are also many other tree numbers with a correlation of around 0.20. Because for decomposition 2 we do not allow decompositions at the top level and also have a different probability distribution at the lower levels we have more common subtrees than in decomposition 1. As a consequence, embedding variations are used to push the security of decompositon 2 to the same level as shown by decompsition 1 (see [11] for details).

Additionally we find that both wavelet packet decomposition strategies lead on average to higher robustness against

JPEG 2000 compression as the pyramidal schemes. Especially at low bitrates decomposition 2 is superior to decomposition 1 which is due to the emphasis on mid-range frequencies. For detailed results see [11].

3. ROBUSTNESS AGAINST UNAUTHORIZED REMOVAL ATTACKS

In the context of watermark robustness, unauthorized removal is achieved by masking or synchronization attacks which disable the watermark detector to extract the watermark properly [1] but do not harm the watermark content systematically. Knowledge about the embedding process as assumed in the context of watermark security allows to mount an elimination attack where the watermark content is actually destroyed.

For our experiments we embed a watermark of length 1000 into the Lena image by employing a 7 level decomposition that results in 40 dB PSNR. We use a "secret" wavelet packet decomposition 2 tree since it offers higher robustness as compared to a decomposition 1 tree and might therefore be also harder to attack. As reference we again use the standard Wang algorithm with a pyramidal decomposition using the Daubechies 6 and the Biorthogonal 7/9 filters for embedding. For embedding, 250 randomly chosen wavelet packet trees are tested and we calculate the average, minimum and maximum of all results after the attacks.

For a realistic attack the adversary only has access to the watermarked image. By applying the wavelet coefficient selection on the already watermarked image he is likely to select different coefficients from the ones that were used for embedding (which is good for security). The attack therefore will have to modify more coefficients and hope that the correct coefficients are attacked. Of course, the quality of the image must not be damaged too much, otherwise the value of the image is lost to the attacker.

For the attack we use a pyramidal decomposition and attack between 100 and 20000 coefficients. In particular, we apply a fixed quantization step size to the selected coefficients. A step size of 100 turned out to be most effective at removing the watermark information and still preserving the image quality.

Fig. 4 compares the correlation of the different systems, when we use the Biorthogonal 7/9 filter for the attack. The standard Wang system which also uses the Biorthogonal 7/9 filter has a correlation of around 0.40 after 3000 coefficients have been attacked. Interestingly the same system using the Daubechies 6 filter shows correlation higher than 0.9 in the same range which already indicates the importance of the filter choice.

The wavelet packet system still has a correlation of more than 0.75 on average after 20000 coefficients have been attacked. The low minimum for the wavelet packet system

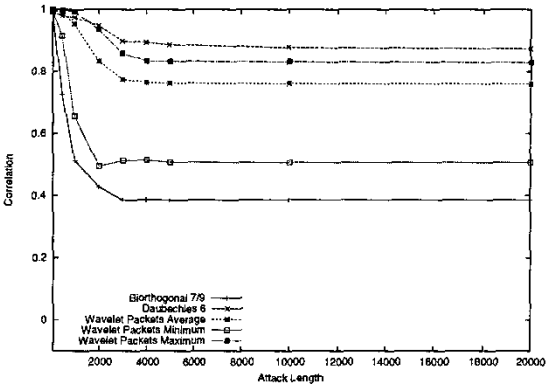


Fig. 4. Watermark correlation under attack using the Biorthogonal 7/9 filter.

can be explained by the possibility of a similarity between the pyramidal decomposition used for the attack and the tree that is used for embedding. Because both decompositions use the Biorthogonal 7/9 filter the minimum can get very close to the behavior of the standard system. In order to avoid this situation, subband structures too similar to the pyramidal scheme could be excluded from the set of admissible candidates.

Fig. 5 shows the resulting image quality after performing the attack. No matter which algorithm has been used for embedding, PSNR is still at 35 dB after having attacked 3000 coefficients, which makes this attack a serious one (significant drop of watermark correlation whereas image quality could be maintained).

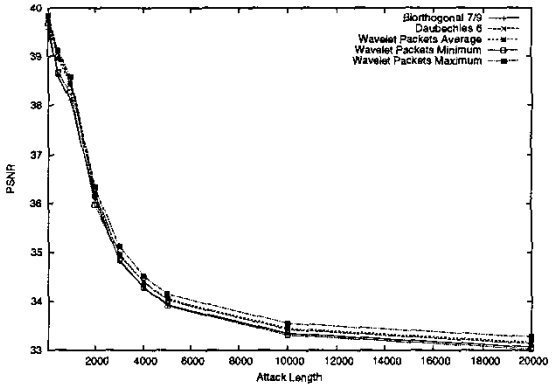


Fig. 5. Quality of the attacked images (attack uses Biorthogonal 7/9 filter).

Figure 6 shows the result when the Daubechies 6 filter is used for the attack decomposition. The performance of the wavelet packet system is better, because now the decompo-

sition filters do not match. The minimum wavelet packet system correlation is now more than 40 percent above the correlation of the standard Daubechies 6 system for 20000 attacked coefficients.

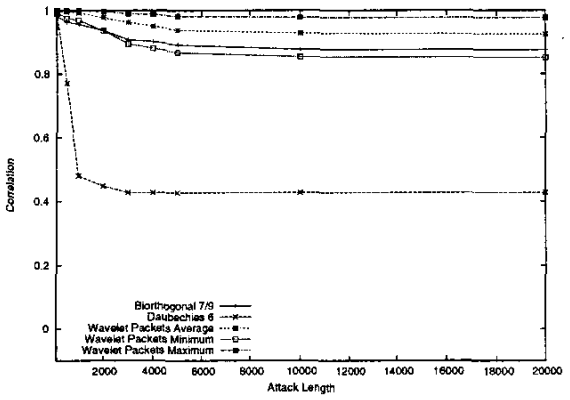


Fig. 6. Watermark correlation under attack using the Daubechies 6 filter.

4. CONCLUSIONS

Consequently, we see a clear advantage of our proposed system in the context of an unauthorized removal attack based on coefficient quantization. With the standard Biorthogonal 7/9 and Daubechies 6 embedding methods the correlation drops to around 40 percent under the attack while the image quality is still reasonable. This means that an unwatermarked image can be obtained with this attack without a severe quality reduction. With the wavelet packet technique we get significantly higher correlation under attack and could therefore still proof ownership of the images. Additionally, we note that the match of filters used for embedding and attack plays an important role as well which explains the good results when using secret filters for watermark embedding as a means to enhance security [6]. Our results suggest a combination of both techniques which results in an enormous keyspace.

5. REFERENCES

[1] Ingemar J. Cox, Matthew L. Miller, and Jeffrey A. Bloom, *Digital Watermarking*, Morgan Kaufmann, 2002.

[2] Stefan Katzenbeisser and Fabien A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, Dec. 1999.

[3] Houng-Jyh Wang and C.-C. Jay Kuo, "Watermark design for embedded wavelet image codec," in *Proceedings of the SPIE's 43rd Annual Meeting, Applications of Digital Image Processing*, San Diego, CA, USA, July 1998, vol. 3460, pp. 388–398.

[4] Deepa Kundur, "Improved digital watermarking through diversity and attack characterization," in *Proceedings of the ACM Workshop on Multimedia Security '99*, Orlando, FL, USA, Oct. 1999, pp. 53–58.

[5] Jiri Fridrich, "Key-dependent random image transforms and their applications in image watermarking," in *Proceedings of the 1999 International Conference on Imaging Science, Systems, and Technology, CISST '99*, Las Vegas, NV, USA, June 1999, pp. 237–243.

[6] Werner Dietl, Peter Meerwald, and Andreas Uhl, "Protection of wavelet-based watermarking systems using filter parametrization," *Signal Processing (Special Issue on Security of Data Hiding Technologies)*, vol. 83, pp. 2095–2116, 2003.

[7] M.V. Wickerhauser, *Adapted wavelet analysis from theory to software*, A.K. Peters, Wellesley, Mass., 1994.

[8] Y. Wang, J. F. Doherty, and R. E. Van Dyck, "A wavelet-based watermarking algorithm for copyright protection of digital images," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 77–88, Feb. 2002.

[9] Min-Jen Tsai, Kuang-Yoo Yu, and Yi-Zhang Chen, "Wavelet packet and adaptive spatial transformation of watemark for digital image authentication," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'00)*, Vancouver, Canada, Sept. 2000.

[10] Jacques Levy-Vehel and Anne Manoury, "Wavelet packet based digital watermarking," in *Proceedings of the 15th International Conference on Pattern Recognition*, Barcelona, Spain, Sept. 2000.

[11] Werner Dietl and Andreas Uhl, "Watermark security via secret wavelet packet subband structures," in *Communications and Multimedia Security. Proceedings of the Seventh IFIP TC-6 TC-11 Conference on Communications and Multimedia Security*, A. Lioy and D. Mazzocchi, Eds., Turin, Italy, Oct. 2003, vol. 2828 of *Lecture Notes on Computer Science*, pp. 214–225, Springer-Verlag.

[12] A. Pommer and A. Uhl, "Selective encryption of wavelet-packet encoded image data — efficiency and security," *ACM Multimedia Systems (Special issue on Multimedia Security)*, vol. 9, no. 3, pp. 279–287, 2003.