SECRET WAVELET PACKET DECOMPOSITIONS FOR JPEG 2000 LIGHTWEIGHT ENCRYPTION

Dominik Engel and Andreas Uhl

Department of Scientific Computing
University of Salzburg, Austria
Email: {dengel,uhl}@cosy.sbg.ac.at

ABSTRACT

A lightweight encryption scheme for JPEG 2000 based on the wavelet packet transform is proposed. This scheme significantly reduces the amount of data to be encrypted compared to full encryption and other partial or selective encryption schemes, at the cost of increased computational complexity in the compression pipeline. We investigate the applicability of this approach in two scenarios: for providing full confidentiality and for its utility as a transparent encryption scheme. We evaluate the presented scheme in the context of each scenario with respect to its impact on compression performance, its complexity, the level of security it provides, and its applicability.

1. INTRODUCTION

Secure transmission and access control in multimedia applications have moved to the center of attention in a significant number of recent research programs. There are several reasons for this development, like the need for privacy and confidentiality in multimedia applications that are becoming increasingly popular and widespread, such as mobile video conferencing. A reason for efficient access control is the commercial interest of owners of multimedia content to secure revenue streams by preventing unauthorized access. In this paper, we discuss secret wavelet packet decompositions for (a) providing confidentiality and (b) providing transparent encryption in the context of JPEG 2000. An important focus for our approach is to significantly lower the amount of data to be encrypted, while still maintaining a level of security that is sufficient for typical multimedia applications.

a) Encryption for Confidentiality: If strict confidentiality is required, full encryption of a multimedia bitstream with a traditional encryption cipher, such as AES, is not always the best option to provide security and access control. In many multimedia applications other requirements, such as retaining bitstream compliance and scalability, low demands in terms of computational complexity, and increased functionality, outweigh traditional security requirements [1]. In the context of wavelet coded image data, different methods have been proposed to achieve some of these advantages. Selective encryption of vital parts of the JPEG 2000 packet data is proposed by [2]. Scrambling coefficient signs in code-blocks is proposed by [3]. [4] investigate scrambling in the context of motion JPEG 2000 coding, integrated into their scalable streaming concept. [5] and [6] discuss the problem of marker emulation in the context of JPEG 2000 encryption and propose solutions that allow to retain standard bistream compliance.

b) Transparent Encryption: The term “transparent encryption” is introduced by [7] to refer to encryption schemes in which portions of the original data are accessible in degraded quality even without key. The full quality version is restricted to legitimate users. In a scheme for transparent JPEG 2000 encryption, [8] propose to encrypt about 85% of the packet data in resolution progressive mode. Parameterized wavelet filters are investigated as a tool for transparent encryption by [9].

We present a lightweight encryption scheme for JPEG 2000 that is based on randomly generated wavelet packet decompositions. The wavelet packet decomposition [10] is a generalization of the pyramidal (or Mallat) wavelet decomposition, where recursive decomposition may be applied to any subband and is not restricted to the approximation subband. This results in a larger space of possible decomposition structures, of which the pyramidal decomposition structure is only one element. The proposed scheme keeps the randomized wavelet packet decomposition trees secret.

Recent work by Pommer and Uhl [11] proposes the use of wavelet packets for providing confidentiality in a zerotree-based wavelet framework. While the work presented here transfers the idea and the central algorithm to JPEG 2000, the entirely different nature of JPEG 2000 as compared to the codec used in [11] leads to a novel situation, especially for potential attacks. The present work has its main focus on the comprehensive evaluation of secret wavelet packets in the different environment. Therefore, the algorithm for producing randomized wavelet packets will only be discussed briefly, more details can be found in [11]. In the following, the compression performance of randomized wavelet packets in JPEG 2000 are assessed, and the parameters used for the generation of randomized wavelet packets are adapted accordingly. Furthermore, with a slight modification to the original algorithm, transparent encryption can be supported naturally. In contrast to the codec used in [11], JPEG 2000 is not based on zerotrees. The impact this has on possible attacks marks a key difference to the previous work. As the superimposition of decomposition structures on a set of wavelet packet coefficients becomes impossible, different modes of attack have to be considered. We evaluate the presented methods regarding their security and usability in terms of (a) providing confidentiality and (b) providing transparent encryption.

2. SELECTIVE ENCRYPTION WITH RANDOMIZED WAVELET PACKETS

Part two of the JPEG 2000 standard allows wavelet decompositions that differ from the pyramidal decomposition [12]. In order to maximize keyspace size for the proposed encryption scheme, we implemented full support for arbitrary isotropic wavelet decomposition structures in JPEG 2000, based on the JJ2000 reference implementation (`http://jj2000.epfl.ch/`).

Wavelet packet decomposition structures can be represented as a sequence of binary decomposition decisions. If during generation of randomized wavelet packet decompositions, the probability for decomposition is the same at each decision, i.e. for each subband, then

This work was partly funded by the Austrian Science Fund (FWF), project number P15170. Dominik Engel was supported by the Austrian Academy of Sciences.

(a) Lena (b) Barbara (c) D105

Fig. 1. Test Images

shallow wavelet decompositions are more probable than deep ones. This bias potentially restricts the keyspace in the context of the proposed encryption scheme. This restriction is resolved by introducing two parameters, *base value (bv)* and *change factor (cf)*, that can be used to influence the probability of decomposition at a single decision point, based on a base probability and a factor that grows or shrinks with the current decomposition depth. Other parameters that determine the mechanism of random wavelet packet decompositions are the maximum global decomposition depth for all subbands *(g)*, the maximum *(m)* and minimum *(n)* decomposition depth for the approximation subband, and the seed for the pseudo-random number generator (PRNG).

Keeping the decomposition structure secret can be achieved in two ways: either the decomposition tree itself is used as the key and not included in the bitstream, or header information containing the used wavelet packet decomposition structure is encrypted with a traditional cipher. If a PRNG is used for generating the wavelet packet structure, the relevant information consists only of the seed and the parameters which are discussed below, otherwise the information consists of the complete wavelet packet structure. In either way, the amount of information that has to be encrypted is very small.

To achieve *transparent encryption*, we introduce an additional parameter p that can be used to optionally specify the number of higher pyramidal resolution levels. If p is set to a value greater than zero, the pyramidal wavelet decomposition is used for resolution levels R_0 through R_p and wavelet packets are used for the higher resolution levels, starting from R_{p+1}. With resolution-layer progressions in the final bitstream, standard JPEG 2000 codecs can be used to obtain resolutions R_0 to R_p. Note that if higher pyramidal resolution levels are used, n should be set to a sufficiently large value, ideally to the same as m, in order to avoid wavelet packet decompositions which are very similar to the pyramidal decomposition.

3. COMPRESSION PERFORMANCE AND COMPLEXITY

For the proposed encryption scheme to be feasible in a general application scenario, the compression performance with wavelet packets in JPEG 2000 has to be comparable to the results obtained with the pyramidal decomposition. The three test images we used are shown in Figure 1. D105 from the collection of Brodatz textures is used to have a specimen of oscillatory patterns.

In the comparison of compression quality of the wavelet packet decomposition and the pyramidal wavelet decomposition it is noteworthy that there are wavelet packet decomposition structures that produce better results than the pyramidal structure. As can be seen in Figures 2(a) and 2(b), this effect is stronger for the image Barbara, because it contains more oscillatory patterns that favor energy compaction with wavelet packets. On average, the wavelet packet decompositions are outperformed by the pyramidal decomposition. At the end of this section, we will present parameter settings that minimize the difference in compression performance.

In the following, we assess test runs which leave one parameter fixed and vary the other parameters through their respective ranges (leading to a total of 68796 test runs per image for g up to 7 and 28665 test runs for g up to 5). If not noted otherwise, the compression rate is 0.25 bpp. The average, minimum and maximum peak-signal-to-noise-ratio (PSNR) for each value of the fixed parameter are plotted. Of the five parameters that influence compression performance, three also affect the number of possible decomposition trees, namely g, m, and n. We will discuss the impact of these parameters in more detail.

The other two parameters, *bv* and *cf*, if not set to extreme values, only affect the probability distribution over the possible decompositions, but not the number of potential decompositions. Our results suggest that furthermore they also have only marginal influence on compression performance. [11] suggest to set the *bv* to 1 and *cf* to 0. While this is sensible from an image compression point of view, it also means a restriction in keyspace, as deeper decompositions are less probable than shallow decompositions. Because the two parameters do not have much impact on compression quality, we propose to set them to values based on the maximum decomposition depth that make every tree decomposition equally likely, regardless of its individual decomposition depth. In this manner, we can avoid giving an attacker the advantage of knowing at which point to start a search in the space of decomposition structures.

The parameter specifying the maximum global decomposition level, g, has a major effect on both compression performance and the number of potential decomposition trees. As can be seen in Figure 2(a), for high overall levels of decomposition, compression performance quickly degenerates. For the size of images used, this is mainly due to the fact that there is an increase in header information for complex wavelet packet decompositions, while no further gain in energy compaction is achieved. We therefore limit our observations to a maximum decomposition depth of 5, which is a reasonable value considering the size of our test images.

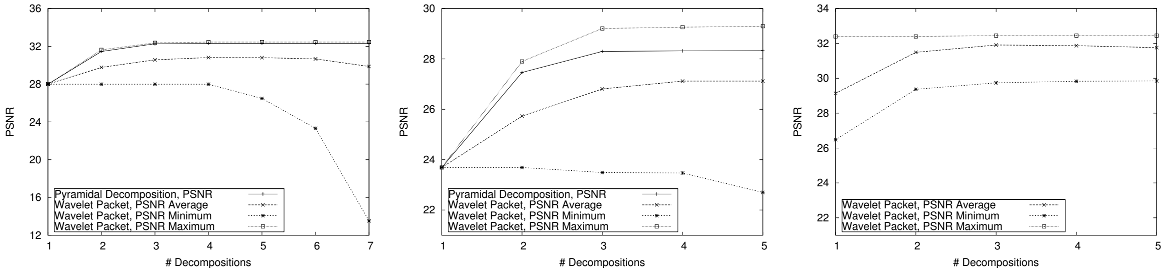
Figure 2(c) shows that the minimum decomposition depth of the approximation subband, n, if set too low, has a significant impact on compression performance. Settings of one or two levels produce compression results that are not competitive and should be avoided. Discarding these smaller settings only marginally affects the number of possible wavelet packet decompositions. With regard to maximum decomposition depth of the approximation subband, m, our results show that there is no reason to restrict this parameter. Considering keyspace size as well as compression performance, we propose to set m to the same value as g.

Table 1. Compression results for suggested parameter settings

	Lena	Barbara	D105
Pyramidal, PSNR	32.31	28.33	17.65
WP, Max. PSNR	32.45	29.19	21.54
WP, Min. PSNR	29.91	26.65	16.94
WP, Avg. PSNR	31.74	28.21	18.81

Taking the above observations and the size of the used test-images into account, we can suggest settings that remove non-competitive compression results: $g = 5$, $n = 3$, and $m = 5$. Table 1 shows the PSNR performance for each of the three test-images for the pyramidal decomposition, and the maximum, minimum and average PSNR for the test runs with the suggested settings (819 test runs for each image). It can be seen that the average compression performance for typical natural images is comparable to the performance of the pyramidal decomposition. As expected, for D105 the maximum PSNR is significantly above the pyramidal decomposition – however, overall compression quality for D105 is generally very low.

Wavelet packets bring an increase in complexity as compared to the pyramidal wavelet decomposition: The order of complexity for



(a) Maximum level of overall decomposition depth (Lena)

(b) Maximum level of overall decomposition depth (Barbara)

(c) Minimum level of approximation subband decomposition depth (Lena)

Fig. 2. Compression performance of randomized wavelet packet decompositions

a level l full wavelet packet decomposition of an image of size N^2 is $\mathcal{O}(\sum_{i=1}^{l} 2^{2(i-1)} \frac{N^2}{2^{2(i-1)}})$ compared to $\mathcal{O}(\sum_{i=1}^{l} \frac{N^2}{2^{2(i-1)}})$ for the pyramidal decomposition, with the randomized wavelet packet decompositions ranging in-between. This has to be taken into account for potential application scenarios. The effort for encryption is dramatically reduced compared to full encryption and other partial or selective encryption schemes, but the wavelet packet transform introduces computational demands in the compression pipeline.

4. SECURITY

In this section we evaluate the security of the proposed scheme with regard to providing (a) confidentiality and (b) transparency. If AES or a cipher of similar security is used, the option of breaking the cipher with which the wavelet packet structure is encrypted, becomes infeasible. Likewise, an exhaustive search for the wavelet decomposition structure is not feasible for sufficiently deep wavelet packet decompositions. The number of possible wavelet decompositions reaching up to level $n+1$ can be determined by

$$f(n) = \sum_{i=0}^{4} \left(\begin{array}{c} 4 \\ i \end{array} \right) \cdot (f(n-1))^i \qquad (1)$$

where $f(0) = 1$. For a maximum decomposition depth g, there are $f(g-1)$ possible decompositions. For $g = 5$, 2^{261} possible decompositions exist, and even with the parameter n set (which removes some decompositions), the complexity of an exhaustive search is higher than a brute-force attack against a 256-bit-key AES cipher. For $g = 6$, the number of possible wavelet decompositions is 2^{1046}. The attacker is therefore left with the option of trying to (partially) reconstruct the image from the unencrypted data.

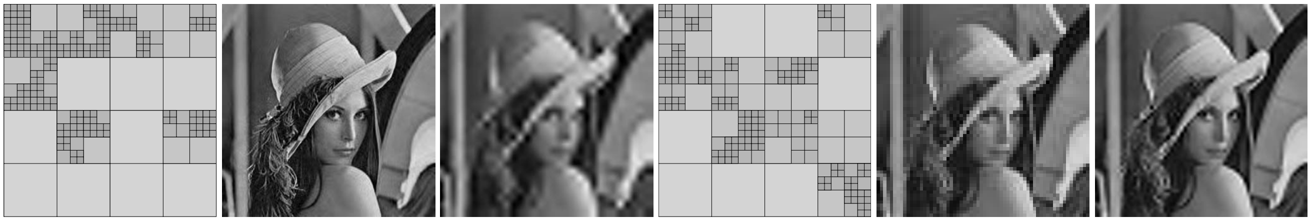
a) Confidentiality: In the case of the codec used in [11], data and the spatial organization of the coefficients can be accessed directly (when a uniform scalar quantizer is used), and the challenge for the attacker lies mainly in superimposing the correct subband structure to make sense of the coefficient data. In the case of JPEG 2000, the coding of the data is in a much lower degree spatially determined. The fact that the subband structure is crucial in decoding the JPEG 2000 packet stream, for associating the coefficient data contained in the packets with code-blocks, poses a major problem for the attacker.

The difficulty of an attack depends on the level of information the attacker has about the bitstream, i.e. how much information except the wavelet packet decomposition structure is not accessible in the header. Encryption of additional header information may help to increase security. If, for example, the number of resolutions and quality layers is known or obtained by the attacker, the progression order used for encryption does not make any difference for the attack: knowing the resolution and layer of each packet allows the attacker to reorganize the packets. This makes one major flaw of the proposed scheme apparent when it comes to providing full confidentiality. The packets of resolution R_0 of any wavelet packet decomposition are the same as the packets produced by a pyramidal decomposition of the same image. Without additional precautions, the first resolution is therefore accessible. This flaw can be weakened by additionally encrypting header information. If the number of resolutions is not known, then the attacker also lacks knowledge of the size of the approximation subband. However, due to the strong limitations on the minimum and maximum number of resolutions, this measure does not sufficiently increase search complexity. In a similar way, additionally hiding the number of quality layers makes the interpretation of the sequence of packets more difficult for the attacker, but fails to provide the level of security needed for full confidentiality. Because most of the energy is contained in the approximation subband (i.e. R_0), an attacker will be able to get a good idea of the visual data by forcing the data from the packets with the highest payload into a pyramidal decomposition.

Effectively, the presented scheme is not able to restrict access to the lower resolution levels in a manner adequate for providing confidentiality. The solution of encrypting the lower resolutions as such introduces additional computational overhead. Nevertheless, future work should investigate adding the security provided by secret wavelet packets in the higher resolution levels to approaches that selectively encrypt the lower resolutions, e.g. [2].

b) Transparency: In contrast to encryption for full confidentiality, in a transparent encryption scheme the accessibility of R_0 is desired, security is only required for the full quality version. Because hiding the number of resolution levels and quality layers falls more into the category of providing security through obscurity, but does not affect vulnerability of the proposed scheme in principle, in the following we assume that apart from the subband structure every detail about the packet stream is known to the attacker. In particular this means that we assume the attacker to know the size of the image and the number of resolution levels contained in the bitstream. Therefore the attacker is assumed to also know the size of the approximation subband. We also assume that no precinct partitioning is used.

Let p be the value of the parameter controlling the number of higher pyramidal resolutions. Then the packets of resolutions R_0 through R_p are the same as for the corresponding pyramidal decomposition and can be decoded by any standard JPEG 2000 decoder without quality loss. For resolution levels higher than R_p, the data does not fit the pyramidal decomposition anymore, and decoding will eventually fail. In order to obtain an image of higher quality than R_p, an attacker could try to read a fraction of the coefficient data of R_{p+1} into the pyramidal structure and then attempt a full resolution reconstruction. However, typically the intersection of the randomly generated decomposition structures and the pyramidal structure is too small to obtain data that allows reconstruction at a substantial quality gain (compared to R_p), as is shown in Figure 3(c) for $p = 1$ and in Figure 3(f) for $p = 2$.



(a) Decomposition Structure, $p = 1$
(b) Reconstruction at full resolution
(c) Result of attack
(d) Decomposition Structure, $p = 2$
(e) Reconstruction at resolution 2
(f) Result of attack

Fig. 3. Reconstruction Examples

When trying to reconstruct the full quality image, the attacker's problem is how to assign packet data to code-blocks, i.e. spatial location. JPEG 2000 employs tag trees to efficiently encode inclusion information for each code-block [12]. Context synchronization between encoder and decoder is used to signal redundant inclusion information. Conceptually, one tag tree is maintained for each subband. The packet headers contain the inclusion information produced by the tag trees for the code-blocks in each subband. The information of location of code-blocks is thus dispersed throughout the code-stream. A promising point of attack is the header of the first packet of a resolution, because it has to contain output from the tag tree of *each* subband contained in the resolution. By analyzing a number of such packets, an attacker could possibly infer the number of subbands. However, the attacker would still lack information regarding the size of each subband and therefore could not uniquely determine the spatial layout of the code-blocks.

With the right overall number of code-blocks (if the maximum code-block size is set to be no larger than the smallest subband, this number is actually known), it would seem that all the data could be retrieved correctly. This data could then be filled into subband structures that have the number of subbands derived from the number of tag trees. The result could be reconstructed and matched against a measure that is correlated to image quality, like the variance, to find the best result. However, even if the right overall number of blocks is known or obtained, the wrong spatial sequencing of the code-blocks remains a major problem for higher resolution levels, because it restricts access to other information necessary to determine the value of individual coefficients: The number of skipped most significant bitplanes in each code-block is contained in another set of tag trees. A wrong spatial placement for an individual block implies a wrong answer from the tag tree for the number of skipped most significant bitplanes, and thus a misinterpretation of this block's codewords. Apart from trying different spatial configurations of code-blocks, the attacker would therefore also have to make an assumption regarding the skipped most significant bitplanes.

5. CONCLUSION

We have shown that wavelet packets can help to reduce computational demands for lightweight encryption. The reduction of complexity for encryption comes at the cost of a rise in complexity in the compression pipeline. The actual applicability of the presented approach depends on the scenario in which it is to be used. We have found the level of security to be too low for applications that demand full confidentiality. In terms of transparent encryption the scheme is successful, as it can naturally be used to grant access to lower levels of resolution while keeping the higher resolution levels secure. The comparatively high computational complexity that is introduced during compression and decompression remains a drawback compared to other lightweight methods. However, the amount of data that needs to be encrypted is extremely small. This presents a significant advantage of the presented approach compared to other suggestions. Encryption effort is minimal and due to the small amount of data to be encrypted this approach can be used directly in public key schemes and benefit from their superior key management.

In future work we will turn to how random wavelet packet decompositions can be combined with other methods of selective encryption to achieve support for full confidentiality. Moreover, it will be investigated how other lightweight encryption methods can benefit in their level of security through a combination with the wavelet packet transform.

6. REFERENCES

[1] A. Uhl and A. Pommer, *Image and Video Encryption. From Digital Rights Management to Secured Personal Communication*, vol. 15 of *Advances in Information Security*, Springer-Verlag, 2005.

[2] Roland Norcen and Andreas Uhl, "Selective encryption of the JPEG2000 bitstream," in *Communications and Multimedia Security. Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security, CMS '03*, A. Lioy and D. Mazzocchi, Eds., Turin, Italy, Oct. 2003, vol. 2828 of *Lecture Notes on Computer Science*, pp. 194 – 204, Springer-Verlag.

[3] Raphaël Grosbois, Pierre Gerbelot, and Touradj Ebrahimi, "Authentication and access control in the JPEG 2000 compressed domain," in *Applications of Digital Image Processing XXIV*, A.G. Tescher, Ed., San Diego, CA, USA, July 2001, vol. 4472 of *Proceedings of SPIE*, pp. 95–104.

[4] S.J. Wee and J.G. Apostolopoulos, "Secure scalable streaming and secure transcoding with JPEG2000," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'03)*, Barcelona, Spain, Sept. 2003, vol. I, pp. 547–551.

[5] H. Kiya, D. Imaizumi, and O. Watanabe, "Partial-scrambling of image encoded using JPEG2000 without generating marker codes," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'03)*, Barcelona, Spain, Sept. 2003, vol. III, pp. 205–208.

[6] Yongdong Wu and Robert H. Deng, "Compliant encryption of JPEG2000 codestreams," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'04)*, Singapure, Oct. 2004, IEEE Signal Processing Society.

[7] Benoit M. Macq and Jean-Jacques Quisquater, "Cryptology for digital TV broadcasting," *Proceedings of the IEEE*, vol. 83, no. 6, pp. 944–957, June 1995.

[8] A. Uhl and Ch. Obermair, "Transparent encryption of JPEG2000 bitstreams," in *Proceedings EC-SIP-M 2005 (5th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services)*, P. Podhradsky et al., Eds., Smolenice, Slovak Republic, 2005, pp. 322–327.

[9] Dominik Engel and Andreas Uhl, "Parameterized biorthogonal wavelet lifting for lightweight JPEG 2000 transparent encryption," in *Proceedings of ACM Multimedia and Security Workshop, MM-SEC '05*, New York, NY, USA, Aug. 2005, pp. 63–70.

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

[11] 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.

[12] D. Taubman and M.W. Marcellin, *JPEG2000 — Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, 2002.