Improving Security of JPEG2000-Based Robust Hashing using Key Dependent Wavelet Packet Subband Structures

GEROLD LAIMER
University of Salzburg
Department of Computer Sciences
J.-Haringerstr.2, 5020 Salzburg
AUSTRIA

ANDREAS UHL
University of Salzburg
Department of Computer Sciences
J.-Haringerstr.2, 5020 Salzburg
AUSTRIA

Abstract: We discuss a JPEG2000 packet data based hashing scheme for robust image authentication. Motivated by attacks against the approach, key-dependency is added by means of employing a randomized wavelet packet scheme in the wavelet decomposition stage. Attacks can be prevented effectively employing key dependency but the parameters required for this approach lead to decreased compression robustness which has to be compensated by an decrease of the hash length (which of course reduces security).

Key–Words: robust image hashing, perceptual hashing, image authentication, JPEG2000, wavelet packets

1 Introduction

The widespread availability of digital image and video data has opened a wide range of possibilities to manipulate these data. in particular, different image processing and image manipulation tools offer a variety of possibilities to alter image data without leaving traces which are recognizable by the human visual system.

Classical cryptographic tools to check for data integrity like the cryptographic hash functions MD-5 or SHA are designed to be strongly dependent on every single bit of the input data. While this property is important for a big class of digital data (for instance compressed text, executables, ...), classical hash functions cannot provide any form of robustness and are therefore not suited for typical multimedia data.

In order to ensure the integrity and authenticity of digital visual data, algorithms have to be designed which consider the special properties of such data types and should assess visual appearance or perceptual content instead of digital representation. On the one hand, such an algorithm should be robust against compression and format conversion, since such operations are a very integral part of handling digital data. On the other hand, such an algorithm should be able to recognize a large amount of different intentional manipulations to such data.

The use of robust hash algorithms for media authentication has been extensively researched in recent years. A number of different algorithms [1, 4, 6, 9, 11] has been proposed and discussed in literature.

A robust visual hashing scheme usually relies on a technique for feature extraction as the initial processing stage, often transformations like DCT or wavelet transform are used for this purpose. Subsequently, the features (e.g. a set of carefully selected transform coefficients) are further processed to increase robustness and/or reduce dimensionality (e.g. decoding stages of error-correcting codes are often used for this purpose). Note that the visual features selected are usually publicly known and can therefore be modified. This might threaten security, as the hash value could be adjusted maliciously to match that of another image.

For this reason, security has always been a major design and evaluation criteria [9, 11] for these algorithms. In this work we investigate the security of a JPEG2000 based robust hashing scheme which has been proposed in earlier work [5, 6]. We describe a severe attack against the original scheme and propose key dependent wavelet packet decomposition structures in the wavelet transform stage of JPEG2000 encoding as key-dependency scheme for the JPEG2000 based robust hashing scheme. After reviewing JPEG2000 basics, Section 2 discusses JPEG2000 based hashing and presents an attack against this scheme. In Section 3 the employed wavelet packet decomposition is shortly described. Subsequently, we discuss properties of the key-dependent hashing approach like actual key-dependency of the hash values, robustness, and sensitivity against image alterations. Section 4 concludes this paper.

2 JPEG2000 based (Robust) Hashing

Most robust hashing techniques use a custom and dedicated procedure for hash generation which differs substantially from one technique to the other. Several techniques have been proposed using the wavelet transform as a first stage in feature extraction (e.g. [5, 11]). The employment of a standardized image coding technique like JPEG2000 (based on a wavelet transform as well) for feature extraction offers certain advantages:

- Widespread knowledge on properties of the corresponding bitstream is available.
- A vast hardware (e.g. Analog Devices ADV202 chip) and software (official reference implementations like JJ2000 or Jasper and additional commercial codecs) repository is available.
- In case visual data is already given in JPEG2000 format, the hash value may be extracted with negligible effort (parsing the bitstream and extracting the hash data).

2.1 JPEG2000 Basics

The JPEG2000 [10] image coding standard uses the wavelet transform as energy compaction method. After the transform, the coefficients are quantized and encoded on a codeblock (i.e. independent, non-overlapping blocks of transform coefficients) basis using the EBCOT scheme, which renders distortion scalability possible. Thereby the coefficients are grouped into codeblocks and these are encoded bitplane by bitplane, each with three coding passes (except the first bitplane). The codeblock size can be chosen arbitrarily with certain restrictions.

main header	packet header	packet data	... ...	packet header	packet data

Figure 1: JPEG2000 bitstream structure

The final JPEG2000 bitstream (see Fig. 1) is organized as follows: The main header is followed by packets of data (packet bodies) each of which is preceded by a packet header. A packet body contains CCPs (codeblock contribution to packet) of codeblocks that belong to the same image resolution (wavelet decomposition level) and layer (which roughly stand for successive quality levels). Depending on the arrangement of the packets, different progression orders may be specified (e.g. resolution and layer progression order).

2.2 JPEG2000 Authentication and Hashing

In previous work [5, 6] we have introduced a robust hashing scheme which employs parts of the JPEG2000 packet body data as robust hash – we denote this approach JPEG2000 PBHash (Packet Body Hash). An image given in arbitrary format is converted into raw pixel data and compressed into JPEG2000 format (or it is eventually already given in JPEG2000 format). Due to the embeddedness property of the JPEG2000 bitstream, the perceptually more relevant bitstream parts are positioned at the very beginning of the file. Consequently, the bitstream is scanned from the very beginning to the end, and the data of each data packet – as they appear in the bitstream, excluding any header structures – are collected sequentially and concatenated to be then used as visual feature values. Note that it is not required to actually perform the entire JPEG2000 compression process – as soon as the amount of data required for hash generation has been output by the encoder, compression may be stopped. JPEG2000 PBHash has been demonstrated to exhibit high robustness against JPEG2000 recompression and JPEG compression [5] and provides satisfying sensitivity with respect to intentional local image modifications [6].

The visual information contained in the hash string (i.e. concatenated packet body data) may be visualized by decoding the corresponding part of the bitstream by a JPEG2000 decoder (including the header information for providing the required context information to the decoder). Fig. 2 shows the visual information corresponding to a hash length of 50 byte of the images displayed in Figs. 5 and 6 (in fact, the images shown are severely compressed JPEG2000 images).

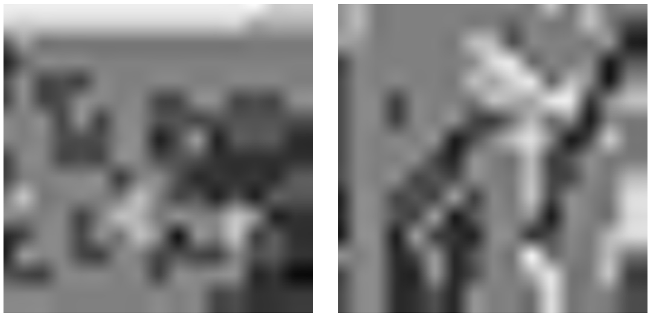

Figure 2: 50-bytes images of the test images Goldhill and Lena.

Unless noted otherwise, we use JPEG2000 with layer progression order and output bitrate set to 1.0 bit per pixel. The length of the hash and the wavelet decomposition depth employed can be used as parameters to control the tradeoff between robustness and sensitivity of the hashing scheme [3] – obviously a shorter hash leads to increased robustness and decreased sensitivity (see [5, 6] for detailed results). A certain minimal decomposition depth (e.g. down to decomposition level 3) is a must and a short hash string requires a higher decomposition depth for sensible employment of the JPEG2000 PBHash.

In Table 1 we show normalized byte-differences among several images with the following parameter settings: hash-length 50 byte with decomposition level 5. As it is desired, the normalized byte difference attains its maximum (or values close to its maximum) for independent images.

Another important aspect of a hashing scheme is the utilization of the available hash space. If we con-

	Graves	Houses	Plane	Lena	Surfside
Goldhill	1.0	1.0	0.98	0.98	1.0
Graves		0.98	1.0	1.0	1.0
Houses			1.0	1.0	1.0
Plane				1.0	0.98
Lena					1.0

Table 1: Normalized byte-differences between various images

sider a hash length of 32 bytes, we have 2^{256} possible hash strings. To get an idea, how this huge number of hash values is used, we conduct another experiment. We calculate the hash strings for 200 natural and artificial images. Now we look at each of the 32 byte positions and plot a histogram of their actual values attained. The assumption is, that if we get a nicely distributed histogram for all positions, chances are good, that the hash space is well utilized. Figure 3 shows the histograms for the first 16 bytes at wavelet decomposition level 6. We notice, that for each byte the possible values from [-1.7,127] are used nearly equally. The only problem seams to be the first byte.

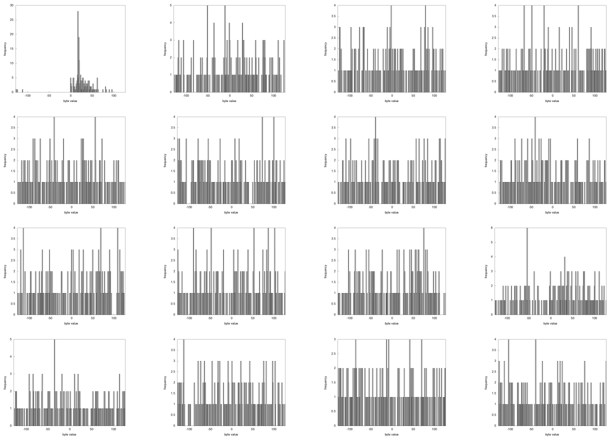

Figure 3: Hash space utilization for bytes 1-16 (wlev 6)

Figure 4 shows the histograms for the first byte at different wavelet decomposition levels. This shows, that with increasing decomposition level the used values drastically decrease. While at level 3 the utilization is quite good, at level 8 only a small amount of possible values is used. According to these results, we can say that the hash space utilization seems to be better at smaller wavelet decomposition levels, although it affects only the first byte of the hash string.

2.3 Attacks against the JPEG2000 PBHash

In order to demonstrate the definite need for key-dependency in the JPEG2000 PBHash procedure, we conduct attacks against the approach using the sightly modified images as displayed in Figs. 5 and 6.

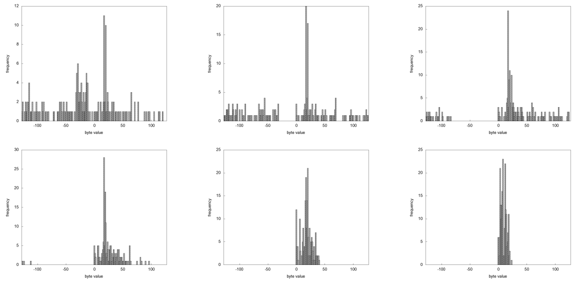

Figure 4: Decreasing hash space utilization for the first byte (wlev 3 – 8)

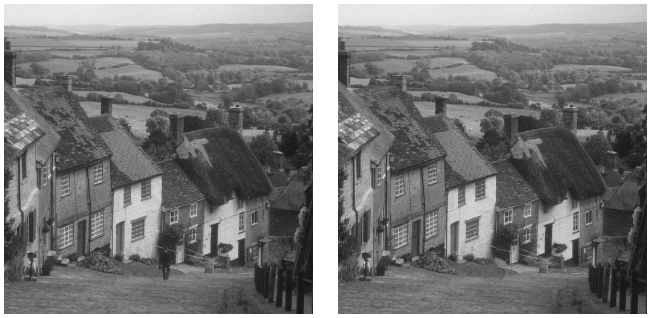

Figure 5: Test image Goldhill (original and with man removed).

A possible attacker aims at maliciously tampering the modified image in a way that the hash string becomes similar or even identical to the hash string of the original image while preserving the visual content (this is the attacked image). In this way, the attacked image would be rated as being authentic by the hashing algorithm.

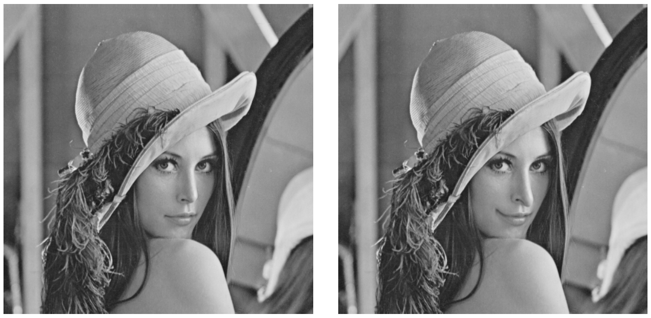

Figure 6: Test image Lena (original and with a grin).

The attack actually conducted works as follows: Both, the original and the modified image are considered in a JPEG2000 representation matching the parameters used for the JPEG2000 PBHash (if they do not match this condition, they are converted to JPEG2000). Now the first part of the bitstream of the original image (corresponding to the packet body data used for hashing) is exchanged with the corresponding part of the bitstream of the modified image resulting in the attacked image. Obviously, if the attacked image remains in JPEG2000 format, its hash exactly matches that of the original. But even if both, the original and the attacked image are converted back to their source format (e.g. PNG) and the JPEG2000 PBHash is applied subsequently it turns out that the hash strings are still identical. Fig. 7 shows the corresponding attacked Goldhill and Lena images. Their hash strings are identical to those of the respective originals.

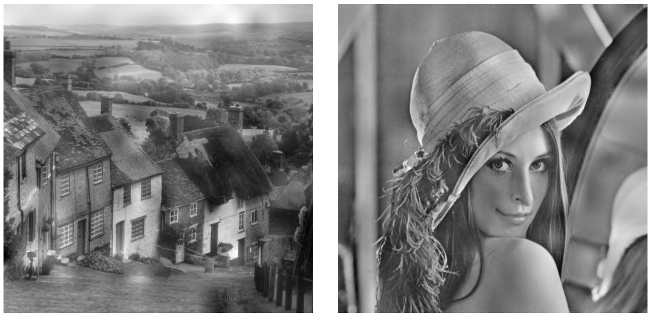

Figure 7: Attacked Goldhill and Lena images.

The demonstrated attack shows that the JPEG2000 PBHash is highly insecure in its original form and requires a significant security improvement to be useful as a reliable authentication hashing scheme.

3 Key-dependent JPEG2000 PBHash

Key-dependency schemes used in the construction of robust hashes include key-dependent transformations [1, 2, 4], pseudo-random permutation of the data [7], randomized statistical features [11, 9], and randomized quantization/clustering [3].

In recent work [2] we have proposed to use Pollens' orthogonal filter parameterization as a generic key-dependency scheme for wavelet-based visual hash functions. Since this parameterization does not easily integrate with lifting based biorthogonal JPEG2000 filters we propose to use a different strategy in this work, compliant to the JPEG2000 Part 2 compression pipeline. JPEG2000 Part 2 allows to extend JPEG2000 in various ways. One possibility is to employ different wavelet packet subband structures as the strictly pyramidal scheme specified in Part 1 of the standard. This is discussed to be used as key-dependency scheme in the following subsection.

Using a key-dependent hashing scheme, the advantage of the JPEG2000 PBHash to generate hash strings from already JPEG2000 encoded visual data by simple parsing and concatenation is lost. An image present as JPEG2000 file needs to be JPEG2000 decoded into raw pixel data and re-encoded into the key-dependent JPEG2000 domain (using the key-dependent wavelet packet decomposition) for generating the corresponding hash string.

3.1 Wavelet Packets

In the classical wavelet transformation only the low-low-sub-band can be further decomposed, resulting in the typical pyramidal structure. Wavelet packet decomposition removes this constraint and allows to further decompose any sub-band (see Fig. 8 for a comparison). The decision which sub-bands are decomposed is either determined by a given structure or based on some measure of optimality.

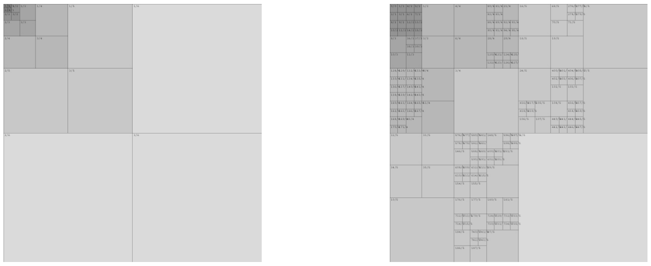

Figure 8: Classical wavelet decomposition vs. arbitrary wavelet packet structure.

By using a pseudo random number generator to decide, if a sub-band should be further decomposed, we can make the decomposition structure key dependent. This approach has been shown to be effective in selective image encryption [8] where also a detailed analysis of the available key space is given (see also [2]).

In order to confirm the results shown in Table 1 also for wavelet packets, we choose an arbitrary fixed key and compute the hash values for 200 independent images. Fig. 9 displays the histograms of the corresponding normalized byte-differences which exhibit the desired property as well (note that this result is just an arbitrary choice as it does not depend on the actual key used).

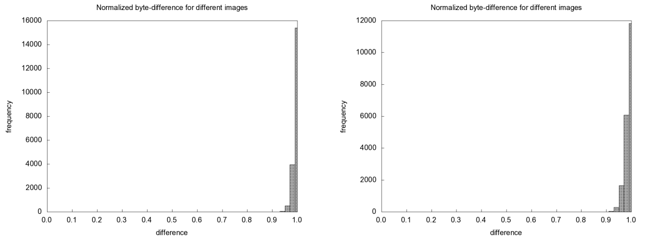

Figure 9: Normalized byte-differences for 200 images using the same key, decomposition depths 7 and 8.

Therefore, also wavelet packet based JPEG2000 PBHash can be considered to be distinctive enough to be used for image authentication purposes.

3.2 Key Dependency

In the following, we investigate the impact of choosing different keys on the resulting hash string, i.e. whether the resulting hash is really sufficiently depending on the key used during JPEG2000 compression. We take an image and generate its hash string with specified settings (i.e. fixed number of bytes extracted from the JPEG2000 bitstream and a certain wavelet decomposition depth) – this procedure is repeated for 100 randomly chosen keys and the normalized byte-difference among all hash strings is computed. Figs. 10 and 11 show the resulting histograms for the images Lena and Goldhill, where the hash string is 50 bytes long and decomposition depths 6, 7, and 8 are compared.

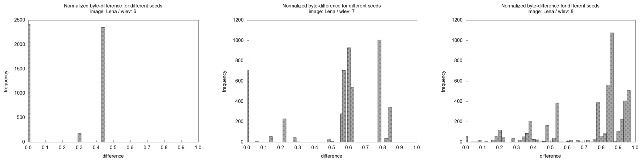

Figure 10: Normalized byte differences for decomposition depths 6, 7, and 8 (Lena image).

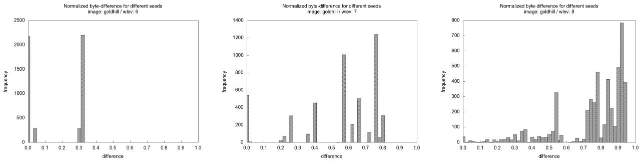

Figure 11: Normalized byte differences for decomposition depths 6, 7, and 8 (Goldhill image).

Whereas we notice large bars even for the zero-difference bin (which means that in these cases the 50 bytes of the hash are in fact identical) for decomposition levels 6 and 7, a much better distribution is found for decomposition level 8 (i.e. a concentration of most bars in the large difference region). The reason is that for lower decomposition depths mostly approximation subband data is included into the hash (since here this subband is larger as for deeper decompositions) – and of course the approximation subband is identical for different wavelet packet subband structures. As a consequence we have to derive that only hashes generated with a wavelet decomposition depth of 8 exhibit actual key dependency. Therefore, we restrict our attention to this settings in the following.

3.3 Properties: Sensitivity and Robustness

While sensitivity against intentional image modifications and robustness with respect to image compression has been discussed in detail for the key-independent JPEG2000 PBHash in previous work [5, 6], the impact of the different filters used in the key-dependency scheme on these properties of the hashing scheme is not clear yet. Therefore, we conduct several experiments on these issues.

The first experiment investigates the sensitivity against the modifications of the images shown in Fig. 5 and 6. We apply the JPEG2000 PBHash to the original and the modified images with the same key and record the number of bytes required to detect the modification (i.e. starting from the begin of the two hash strings, the position / number of the first unequal byte is recorded). This procedure is repeated for 100 different random keys and the results for decomposition depth 8 are shown in Fig. 12. The solid line represents the value obtained with the key-independent JPEG2000 PBHash while the dots represent 100 key-dependent results.

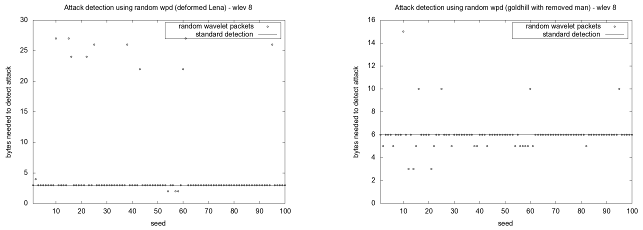

Figure 12: Number of hash byte required to detect Lena's grin and the missing man in Goldhill (hash strings generated with 100 random keys vs. "standard" JPEG2000 PBHash, decomposition depth 8).

First, it is obvious that sensitivity is equal to the key-independent JPEG2000 PBHash for most keys. Second, there is no clear trend with respect to the sensitivity of the "standard" JPEG2000 filter as compared to the wavelet packet versions for the remaining keys. For the Lena image, more keys show decreased sensitivity (more bytes are needed to detect the changes) whereas for Goldhill, more keys exhibit increased sensitivity. Overall, the influence of varying subband structures on sensitivity seems to be of minor importance.

The second property investigated in this subsection is robustness to common image transformations. As a typical example we select JPEG2000 re-compression. We apply the JPEG2000 PBHash to the original and compressed images (bitrate 0.5 bpp) with the same key and record the number of bytes required to detect the modification (i.e. starting from the begin of the two hash strings, the position / number of the first unequal byte is recorded). This procedure is repeated for 100 different random keys and the results for decomposition depth 8 are shown in Fig. 13. The solid line represents the value obtained with the key-independent JPEG2000 PBHash while the dots represent 100 key-dependent results.

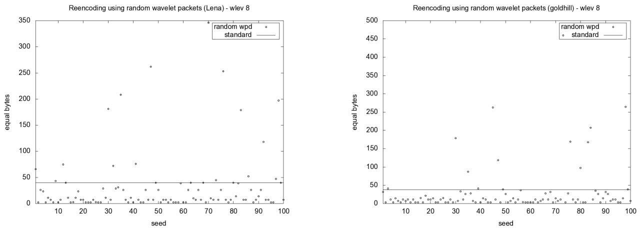

Figure 13: Number of hash bytes required to detect that the Lena and Goldhill images got compressed to 0.5 bpp (hash strings generated with 100 random keys vs. "standard" JPEG2000 PBHash, decomposition depth 8).

It is clear that compression robustness is significantly reduced for most keys considered. A small number of keys also shows increased robustness. The reason for this behaviour is that for most wavelet packet subband structures also more high frequency information is included into the hash string. Since this high frequency information is especially affected by compression, compression is detected in "low-ranked" bytes of the hash. A possible solution is to reduce hash length to 10 bytes: with this settings, most modifications still can be detected (see Fig. 12) and robustness to compression is maintained to a large extent.

4 Conclusion and Future Work

Key-dependency is added to a JPEG2000 packet data based hashing scheme by means of employing a wavelet packet decomposition scheme in the wavelet decomposition stage. However, this type of key-dependency comes at a certain cost: due to reduced robustness of most employed subband structures the hash length has to be decreased to maintain robustness as compared to the scheme without key-dependency (to 10 bytes). This leads to a reduced key-space of course and therefore lower security.

In future work we will investigate if the attacks demonstrated against the scheme without key-dependency can be actually prevented effectively using wavelet-packet based key-dependency.

Acknowledgements: This work has been partially supported by the Austrian Science Fund, project no. 19159 and by the European Commission through the IST Programme under contract IST-2002-507932 ECRYPT. The use of Dominik Engel's wavelet packet implementation is gratefully acknowledged.

References:

[1] Jiri Fridrich. Visual hash for oblivious watermarking. In Ping Wah Wong and Edward J. Delp, editors, *Proceedings of IS&T/SPIE's 12th Annual Symposium, Electronic Imaging 2000: Security and Watermarking of Multimedia Content II*, volume 3971, San Jose, CA, USA, January 2000.

[2] A. Meixner and A. Uhl. Security enhancement of visual hashes through key dependent wavelet transformations. In F. Roli and S. Vitulano, editors, *Image Analysis and Processing - ICIAP 2005*, volume 3617 of *Lecture Notes on Computer Science*, pages 543–550, Cagliari, Italy, September 2005. Springer-Verlag.

[3] V. Monga and B. L. Evens. Perceptual image hashing via feature points: Performance evaluation and trade-offs. *IEEE Transactions on Image Processing*, 15(11):3452–3465, 2006.

[4] V. Monga and M. K. Mihcak. Robust image hashing via non-negative matrix factorizations. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing 2006 (ICASSP'06)*, Toulouse, France, April 2006.

[5] R. Norcen and A. Uhl. Robust authentication of the JPEG2000 bitstream. In *CD-ROM Proceedings of the 6th IEEE Nordic Signal Processing Symposium (NORSIG 2004)*, Espoo, Finland, June 2004. IEEE Norway Section.

[6] R. Norcen and A. Uhl. Robust visual hashing using JPEG2000. In D. Chadwick and B. Preneel, editors, *Eighth IFIP TC6/TC11 Conference on Communications and Multimedia Security (CMS'04)*, pages 223–236, Lake Windermere, GB, September 2004. Springer-Verlag.

[7] H. Özer, B. Sankur, N. Memon, and E. Anarim. Perceptual audio hashing functions. *EURASIP Journal on Applied Signal Processing*, 2005(12):1780–1793, 2005.

[8] A. Pommer and A. Uhl. Selective encryption of wavelet-packet encoded image data — efficiency and security. *ACM Multimedia Systems (Special issue on Multimedia Security)*, 9(3):279–287, 2003.

[9] A. Swaminathan, Y. Mao, and M. Wu. Robust and secure image hashing. *IEEE Trans. on Information Forensics and Security*, 1(2):215–230, June 2006.

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

[11] Ramarathnam Venkatesan, S.-M. Koon, Mariusz H. Jakubowski, and Pierre Moulin. Robust image hashing. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'00)*, Vancouver, Canada, September 2000.