Security of DM quantization watermarking scheme: a practical study for digital images

Patrick Bas, Jarmo Hurri

SECURITY OF DM QUANTIZATION WATERMARKING SCHEMES: A PRACTICAL STUDY FOR DIGITAL IMAGES

Patrick Bas[1], Jarmo Hurri[2]

[1] Laboratoire des Images et des Signaux de Grenoble
961 rue de la Houille Blanche Domaine universitaire
B.P. 46 38402 Saint Martin d'Hères cedex FRANCE
Laboratory of Computer and Information Science
Helsinki University of Technology
P.O. Box 5400 FI-02015 HUT FINLAND
[2] Helsinki Institute for Information Technology
Basic Research Unit, P.O. Box 68
FIN-00014 University of Helsinki FINLAND

Abstract. In this paper, the security of Dither Modulation Quantization Index Modulation schemes for digital images is analyzed. Both pixel and DCT coefficient quantization schemes are investigated. The related works that deal with the security of spread spectrum and quantization schemes are presented and their limits are outlined. The use of independent component analysis (ICA) for natural image is introduced. We show that ICA can be an efficient tool to estimate the quantization noise which is by definition independent of the host signal. We present both a method for estimating the carrier, and an attack that relies on the ICA decomposition of patches of images; our attack scheme is also compared with another classical attack. The results reported in this paper demonstrate how changes in natural image statistics can be used to detect watermarks and devise attacks. Such natural image statistics-based attacks may pose a serious threat against watermarking schemes which are based on quantization techniques.

1 Introduction

After ten years of active development by the watermarking scientific community, many of the proposed watermarking techniques are considered to be mature because they are robust while preserving the quality of the host data. However, if robustness and fidelity are mandatory requirements for an usable watermarking scheme, security is also a very important issue that is not very often addressed. Robustness commonly denotes an ability to decode the watermark after various operations such as compression, filtering, noise addition or geometric transforms. A scheme is considered secure if it is not possible to extract, remove or change the watermark[1][2]. Many watermarking schemes claim to be secure because they use a secret key during the embedding and detection process. However, this

hypothesis is often too weak in real application scenarios and several security attacks have been already proposed based on for example a full access to the detection process [3], the use of a symmetric detection scheme [4], or information leakage when a database of hosts is watermarked using the same secret key[2].
This paper focuses on the security of an important class of watermarking schemes based on quantization called Dither-Modulation (DM) Quantization. The contribution of this paper is to show how changes in natural image statistics can be used to detect watermarks and devise attacks against watermarking schemes. In general, watermarking may change the statistical properties of images, and this can be used to devise detection and attack schemes. Here we show how a method called Independent Component Analysis (ICA) can be used to separate the watermark component when DM Quantization has been used to watermark images. As demonstrated by the results of our paper, such natural image statistics-based attack schemes may pose a serious threat against watermarking schemes.
The rest of the paper is divided in five sections. First, principles of DM Quantization watermarking schemes are presented in section 2. Two different scenarios are presented: secrecy is either established by a secret location selection or by the use of a dither vector. In section 3, the solution proposed by Cayre *et. al* for the Dither-Modulation techniques based on spread transforms is presented, and its limitations with natural images are outlined. We also motivate the use of ICA as a tool to separate the watermark component from the features of the image. Decomposition of natural images into independent basis vectors is presented in section 4. Section 5 presents the two main ideas of the paper: the estimation and identification of the secret carrier and the attack that is used to remove the watermark. These two methods both use decomposition of image blocks into a basis of independent vectors. The performance of the presented scheme against other attacks in term of introduced distortion is also compared. Finally section 6 draws conclusions and gives directions of upcoming works.

2 DM Quantization schemes

2.1 Principles of DM quantization

The class of Quantization Index Modulation (QIM) watermarking schemes was first presented by Chen and Wornell [5]. The principle of the basic QIM technique is to embed a binary message $b(m)$ in an host sample x by applying a quantizer on a modified component to obtain the watermarked component x_w:

$$x_w(x; b) = q(x + d(b)) - d(b)$$

where $q(x)$ is a quantization function with a quantization step equal to Δ, and $d(b)$ is called the dither vector that is function of the transmitted bit b:

$$d(b=1) = \begin{cases} d(b=0) + \Delta/2 & \text{if} \quad d(b=0) < 0 \\ d(b=0) - \Delta/2 & \text{if} \quad d(b=0) \geq 0 \end{cases}$$

One basic solution is to choose $d(b=0) = 0$ which is equivalent to have a set of two disjoint quantizers where each quantizer has a quantization step equal

to Δ and each quantization cell is distant from $\Delta/2$ with the closest one. Such embedding quantization grid is illustrated in Fig.1. The main problem of this

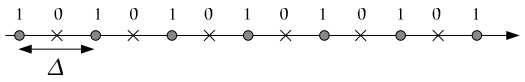


Fig. 1. Quantization grid obtained after applying QIM and a null dither component.

basic and very simple embedding scheme is that it is public, consequently everybody can access to the watermarked components and decode the watermark once the quantization step Δ is known. Computing the pdf of the watermarked components, or in the practical case, a histogram allows one to estimate the parameter Δ. Two solutions have been used to achieve secrecy, either by taking a pseudo-random dither vector, or by selecting the watermarked coefficients at secret locations.

2.2 Secrecy by the use of a dither vector

In [5] the authors propose to use a pseudo-random dither component $d(b = 0)$ that depends on a secret key (the seed of the random number generator for example). In this case, the scheme is the called QIM with Dither Modulation (DM). The dither sequence may for example represent a uniform distribution between $[-\Delta/2; \Delta/2]$. Under the hypothesis that the host signal is locally uniform around the quantization cell, the watermarked signal will stay uniform after the embedding. If we consider samples such as image pixels, the quasi-invariance of the pdf after DM-QIM embedding can also be presumed for smooth distributions and small quantization steps as illustrated on Fig.2. When considering the information available in the pdf, secrecy has then been achieved because the quantization cells are no more disclosed and it is not possible to access the watermarked components without knowing the dither sequence.
Howerver in this case, due to the quantization process, the embedding can be modelled as the addition of uniform noise between $[-\Delta/2; \Delta/2]$ (this property is true even if the dither component is null) and this noise is independent of the host signal. This last property, independence between the watermark signal and the host image, is the main idea of this paper.

2.3 Secrecy by location selection

Another simple way that is used to achieve secrecy in a number of watermarking schemes, including the QIM scheme, is to watermark only a set of selected coefficients of the host image. The selection can be done using a secret key that is used to generate the position of the embedded coefficients. Without this knowledge, the attacker will have to process all of the coefficients to perform a successful attack. This security measure can be of course combined with dither modulation.

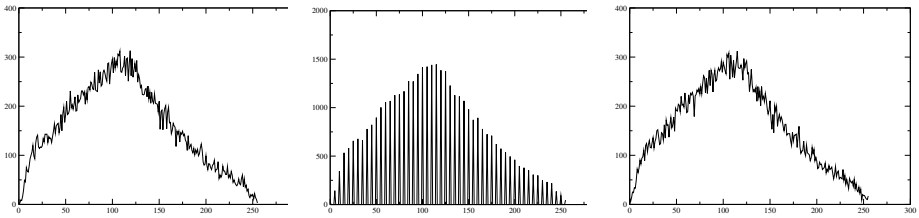


Fig. 2. Histograms of 40000 pixels randomly choosen from different images. After DM, the embedding cannot be observed.

2.4 Studied practical scenarios

This paper addresses both of the previous security measures but with different practical hypotheses. It is important to point out that we assume that the watermark is embedded in a block, and only one "component" of the block is watermarked using the DM-QIM algorithm. By the term "component" we mean either a pixel of the block (which is often the case when the location is secret) or a DCT coefficient of the DCT transform of the block (which can also be used for robust DM-QIM watermarking).

The "database and location" scenario: if secrecy is based on the position of the watermarked coefficients, we assume that we have a large collection of watermarked images that have been watermarked using the same secret key, which is equivalent to assume that for each host vector, the samples have been watermarked at the same locations. We also assume that the watermarked coefficients are randomly distributed in each host signal and that in one block of size $N \times N$ there is only zero or one watermarked coefficient. We call this the database and location scenario. The goal here is to find the position of the watermarked pixels (see Fig.3) by processing one block per image, each block being taken at the same location. When the location is found, we also aim to remove the watermark. Such a scenario can result from an entity using its private key to watermark all the images in its database.

The "one coefficient" per block scenario: in this case we assume that we

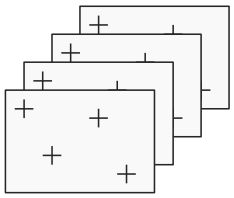


Fig. 3. Database and location scenario: our objective is to find the position of watermarked blocks considering a database of images.

have only one watermarked host image that is partitioned into blocks of size $N \times N$ and in each vector only one component is modified using DM-QIM. This scenario is for example used by watermarking schemes working in the JPEG compressed domain where only one coefficient is modified.
It is important to notice that these two scenarios provide similar outputs: in both cases we have to process a set of blocks of size $N \times N$, and the goal is to estimate the position of the watermarked coefficient and subsequently to remove the watermark.

3 Motivation

3.1 Related research on watermarking security

A consequent work on watermarking security has been done by Cayre *et. al.*[2][6]. The authors have analyzed both the theoretical and practical securities of blind spread spectrum (BSS) watermarking schemes and quantization schemes. In this context the watermark is embedded using a set of pseudo random and orthogonal carriers (one carrier for each transmitted bit) that is modulated and added to the host signal. One part of the mentioned work is devoted to estimate the carriers using a collection of images that have been watermarked using the same carriers but different messages. The authors have expressed such a situation by the following equation:

$$\mathbf{Y} = \mathbf{X} + \alpha \mathbf{U}\mathbf{A}$$

where α is a scale proportional to the power of the watermark and the meaning of the other matrices $\mathbf{Y}$, $\mathbf{X}$, $\mathbf{U}$ and $\mathbf{A}$, that compose this equation is explained in Fig.4. Each image is considered as a realization of independent and identically distributed (iid) gaussian process. The authors have proposed to estimate the

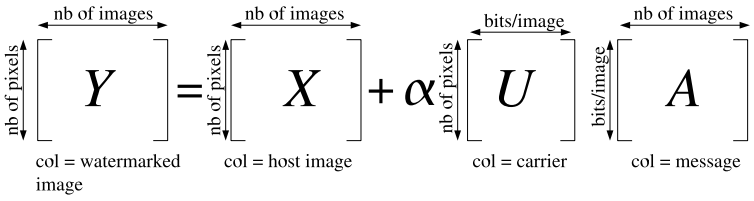
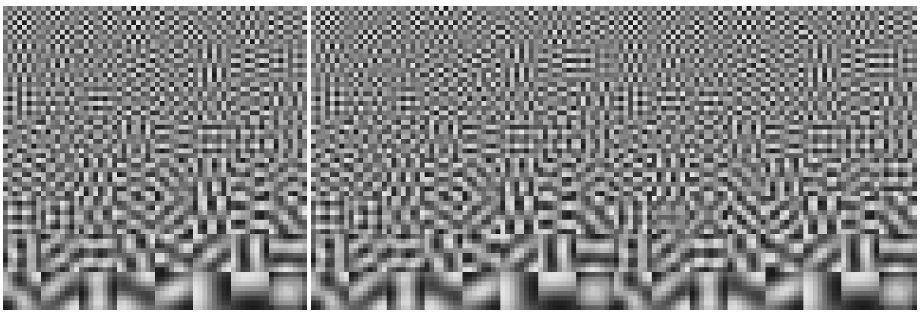


Fig. 4. Details of the model equation proposed by Cayre *et. al.* to model BSS watermarking for a database of images.

matrices $\mathbf{U}$ (the carriers) and $\mathbf{A}$ (the embedded messages) using only the matrix $\mathbf{Y}$ (the set of watermarked images) by using Independent Component Analysis [7]. In this case, the independent sources are given by the embedded messages that are supposed to be different in each watermarked image and therefore independent. Thus ICA is the appropriate tool to estimate both $\mathbf{U}$ and $\mathbf{A}$. The only

limitation of this technique is the fact that both the decoded message and the carriers are estimated up to sign (or bit flipping), this is due to the ICA technique itself. If it is not possible to decode the watermark, it is however possible to remove it and the authors have successfully used this technique to remove watermarks embedded in images.
In [6] the authors have also proposed to deal with the case of one unique carrier and to apply their methodology on quantization schemes. Because independence can not be exploited anymore in this particular case, the authors propose to use classical principal component analysis (PCA) to estimate the carrier. PCA is a method that estimates *principal components*, i.e., uncorrelated components along which the data has variance optima (including the components with the largest and smallest variances)[7]. Considering the host as an iid process, the principal component of the host should be negligible in comparizon with the component given by the carrier.
However in practice the assumption of an iid host is not realistic in the case of natural images when the carrier has the size of a small block (8x8, 16x16, 32x32). For small patches of natural images, the host can not be modelled by and iid process and the PCA give high energies components looking like a DCT basis (cf Fig.5). Consequently the carrier can not be estimated using PCA on small image blocks.



Principal components of original blocks. | Principal components of watermarked blocks, spatial embedding. | Principal components of watermarked blocks, DCT embedding.

Fig. 5. PCA basis vectors for 20000 original and watermarked image blocks. For spatial embedding, the embedding is done on one pixel of each 8×8 block and the resulting PSNR is equal to 51.1dB. For DCT embedding, the embedding is done on the coefficient (5,5) and the resulting PSNR is equal to 43.8dB. No difference is noticeable between the three sets, only several principal components have been displayed with an opposite signe.

3.2 Carrier estimation and blind source separation

Estimating and removing a carrier can be seen has a blind source separation problem where one might have a decomposition of the watermarked image into two subspaces: one representing the watermark and an other one representing the features of the original images. This paper proposes to estimate the secret carrier by exploiting the fact that the carrier is statistically independent of the features of natural images. This property enables firstly to use the decomposition in independent vectors to estimate the secret carrier and secondly to design an attack by processing the independent signal related to the watermark. The next section introduces the model that is used to generate a basis of independent vectors from blocks of images.

4 Independent basis vectors of natural and watermarked images

4.1 Estimation of independent basis vectors

Independent component analysis has been proposed to be used as a generative model of image data [7]. In this model each image block $\mathbf{x}$ can be expressed as a linear combination of independent components referred as the source signals $\{s_i\}$:

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \sum_i \mathbf{a}_i s_i$$

where $\mathbf{A}$ is a constant matrix called the mixing matrix and the vector $\mathbf{a}_i$ denotes the i^{th} column of $\mathbf{A}$. These vectors are called the features or basis vectors. The decomposition of images patches (blocks) into independent basis vectors has been widely used for both image processing and computational modeling of the visual system. The independent basis vectors are similar to edge and lines detectors; neurons performing similar feature detection have been found in the mammalian visual system[8]. Moreover, because in natural image data the distribution of basis vectors is often sparse, the decomposition of images into independent components has also been used for image denoising [9] and image coding [10]. ICA techniques can be used to estimate the mixing matrix $\mathbf{A}$ from a set of N image blocks $\{\mathbf{x}_1, ..., \mathbf{x}_N\}$ (which are unfolded into vectors) using the matrix formulation:

$$\mathbf{X} = \mathbf{A}\mathbf{S}$$

For purpose of illustration each term of this equation is explained in Fig.6. The estimation of the matrices $\mathbf{X}$ and $\mathbf{A}$ can be performed using different ICA algorithms; we have decided to use FastICA because this algorithm achieves good performance both in computational cost and reliability of the extracted basis vectors [7]. An example of a set of ICA basis vectors obtained from 8×8 blocks of natural images is given in Fig.7. Verbally, these basis vectors are typically described as localized, oriented and band-pass filters.

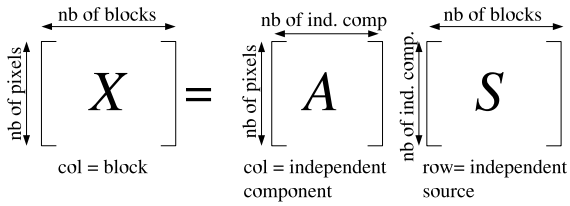


Fig. 6. Details of the equation used for computing independent basis vectors based on images patches.

4.2 Independent basis vectors for watermarked images

The main idea of this paper is to use the decomposition of blocks of images into independent basis vectors to estimate the secret carrier that has been used during DM-QIM watermarking. Note that if the image has been watermarked in the pixel domain, the carrier corresponds to the position of the quantized pixel, and if the image has been watermarked in the DCT domain, the carrier then corresponds to a representation of a DCT basis in the pixel domain. Because these carriers are related to the quantization noise, and we can assume that the quantization noise in independent of the other independent basis vectors, we can estimate the carriers using independent component analysis.
Fig.8 illustrates the capability of ICA to estimate the secret carrier for both the pixel and DCT DM-QIM. In the pixel domain embedding, we have watermarked the central pixel of 20 000 blocks taken from natural images; the resulting PSNR after the embedding was 51.1dB. The estimation of the basis vectors by FastICA using the tanh nonlinearity[11] shows that the secret carrier has been detected in the basis vector that is located on the bottom right of the set of patches. The same test has been completed for DCT embedding. For illustration purposes, we have chosen to watermark the (4,4) coefficient. The resulting PSNR is equal to 43.8 dB. As illustrated in Fig. 8, the spatial representation of this DCT component is also clearly identified on the bottom right corner of the set of basis vectors.

5 Estimating and removing the watermark

The goal of this section is to present a simple scheme to automatically identify the secret carrier and to remove the associated watermark.The estimation of the carrier relies on the decomposition in a basis of independent vectors, as described in the previous section. The removal of the watermark is performed by processing the independent component related to the watermark and applying the mixing operation to generate the attacked image.

5.1 Carrier estimation

Estimation of a secret carrier is a necessary step to perform a successful attack. The carrier will be estimated as a vector that is included in the set of basis vectors

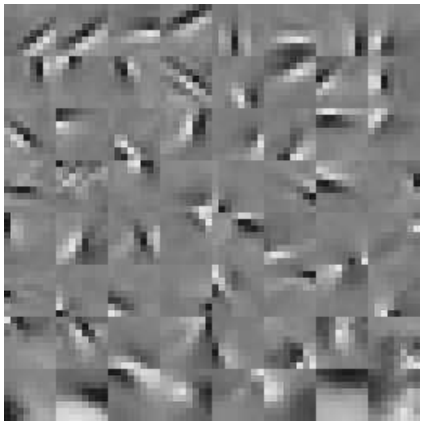
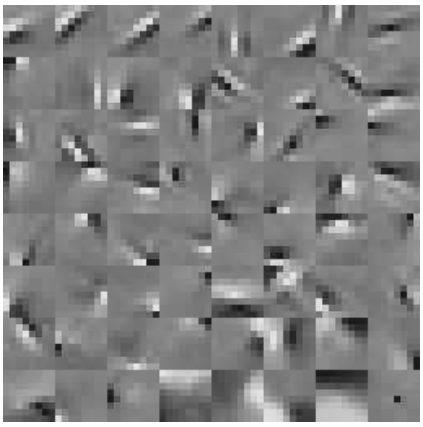
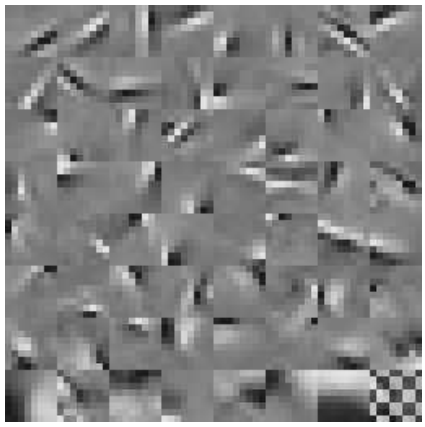


Fig. 7. A complete set of 64 ICA basis vectors, computed from 20000 8×8 pixel samples taken from natural images.



Independent components of watermarked blocks in the pixel domain.



Independent components of watermarked blocks in the DCT domain.

Fig. 8. ICA outputs for 8×8 watermarked blocks. ICA is able to estimate the secret carrier, located at the bottom right block in each set of vectors.

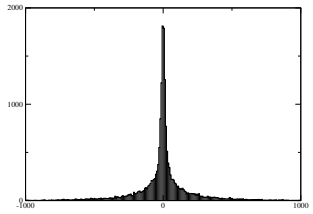
obtained using an ICA algorithm. Note that the proposed method estimates the carrier without any a-priori information about the nature of the transformed coefficient.

Detection criterium: We firt need to select the carrier from the set of basis vectors obtained using the ICA algorithm, e.g. the columns of the mixing matrix **A**. Using the deflationary algorithm which works by estimating the independent vectors one by one, the carrier is extracted ias one of the last estimated vectors. This property is due to the fact that the watermark is the 'least nongaussian' of the independent components, as measured by the objective function maximized by the ICA algorithm. Consequently we have decided to choose the basis vector related to the carrier by taking into account the negentropy of its related component. Negentropy of a source $\mathbf{x}$ is the difference between the entropy of a gaussian source and the entropy of the considered source. It is given by $J(\mathbf{x}) = H(\mathbf{x}_{gauss}) - H(\mathbf{x})$. The negentropy of subgaussian component is lower than the negentropy of the components of natural images which are sparse and have an large negentropy. Consequently, we select the vector with the lowest negentropy as the estimate of the carrier.
Using this criterion, we have noticed that, in the case of the DCT embedding, this estimation of the carrier gives a low estimation error for high frequency components, but gives an high estimation error for low frequency components (see Table 9). This is due to the fact that, for low pass components, the energy of the watermark is too small in comparison with the energy of the host component (see Fig.1). This problem is only present when the embedding is done in a low-frequency DCT coefficient; the selection criterion is successful for pixel embedding and high-frequency DCT embedding.

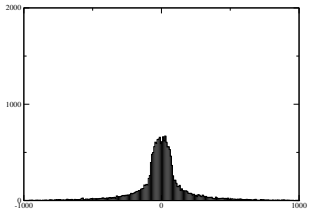
Coeff	(1,1)	(2,2)	(3,3)	(4,4)
MSE	0.0209	0.0216	0.0047	0.0037

Table 1. Mean Square Error between the estimated and true carrier as a function of the watermarked DCT coefficient ($\Delta = 200$).

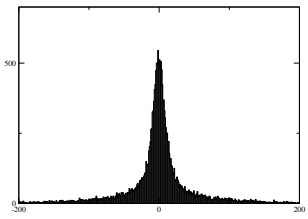
Improvements using pre-filtering of the DCT components: To improve the estimation of the carrier when the embedding is done using a carrier which has low-frequency component, we have reduced the impact of DCT coefficients having high values by only considering coefficients that are under a given threshold (for example Δ). This operation consists of computing the DCT coefficient of each blocks, zeroing the DCT coefficients that are above the threshold, and finally computing the inverse DCT transform of the block before applying the ICA algorithm using the whole set of blocks. Consequently ICA is performed considering only the centre of the distributions of the DCT coefficients, where



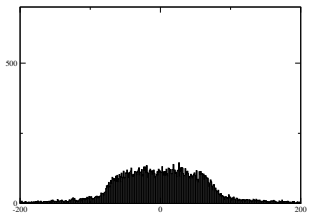
Histogram of the (1;1)
DCT component, original image.



Histogram of the (1;1)
DCT component, watermarked image.



Histogram of the (4,4)
DCT component, original image.



Histogram of the (4,4)
DCT component, watermarked image.

Fig. 9. Histograms of original DCT coefficients and watermarked coefficients ($\Delta = 200$). The subgaussianity is more pronounced for the high frequency component (4,4) than for the low frequency one (1,1).

difference between the original and watermarked component is the most important. Such a pre-filtering technique enables us to achieve low estimations errors (cf. Table 2).

Coeff	(1,1)	(2,2)	(3,3)	(4,4)
MSE	0.0009	0.0015	0.0015	0.0014

Table 2. Mean Square Error between the estimated and true carrier when the pre-filtering of DCT component is used. The DM algorithm was applied to different DCT components (as denoted by column heading in the table).

5.2 Removing the watermark

Once the carrier $\mathbf{a}_k$ that has been used to convey the watermark has been estimated, we have access to the communication channel. It is then possible to design a specific attack to reduce as much as possible the quantization noise produced in DM-QIM and to destroy the watermark. If we denote with $\mathbf{s}_k$ the k^{th} row of matrix $\mathbf{S}$, $\mathbf{s}_k$ contains all the information that is related to the watermark vector. One straightforward way to remove the watermark is to simply

reset $\mathbf{s}_k$. However, if we consider that the component $\mathbf{s}_k$ may contain image information especially for heavy textured patches, we may apply a softer function by keeping high values. This leads to the application of a shrinkage function to each sample of the vector $\mathbf{s}_k$. The used function is depicted in Fig.10; and similar functions have been used for image denoising applications [12] or blind watermarking removal [13]. Our watermark identification and removal procedure can be summarized as follows:

- Build the $\mathbf{X}$ matrix from the set of watermarked image blocks (each block in a column of $\mathbf{X}$).
- Using FastICA, compute $\mathbf{A}$ and $\mathbf{S}$ such as $\mathbf{X} = \mathbf{AS}$.
- Estimate the watermark carrier $\mathbf{a}_k$ and the related source $\mathbf{s}_k$.
- Modify $\mathbf{s}_k$ using a shrinkage function to obtain $\hat{\mathbf{s}}_k$.
- Substitute $\mathbf{s}_k$ by $\hat{\mathbf{s}}_k$ to obtain $\hat{\mathbf{S}}$.
- Compute the matrix $\mathbf{X}_a = \mathbf{A}\hat{\mathbf{S}}$ that represents the attacked set of blocks.

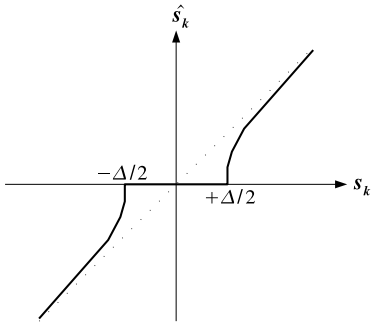


Fig. 10. Shape of the used shrinkage function.

5.3 Attack performance

The goal of this section is to evaluate the performance of the presented denoising attack as a function of the number of observed watermarked blocks. To test the denoising quality we have computed the PSNR between the original and watermarked image and the PSNR between the original and attacked image. We have also computed the resulting Bit Error Rate relative to the attack. For purposes of comparison we have also calculated the PSNR after a straightforward attack that consists of adding $\pm(\Delta/4+\epsilon)$ on several watermarked coefficients to achieve the same BER. The principle of this attack is to move quantized coefficient just to border of the quantization cell that code the opposed bit as illustrated in Fig.11.
The results are depicted in Table 3. Both pixel domain and DCT domain DM-QIM embedding have been tested. For the DCT embedding, the coefficient (4,4)

has been watermarked ((0,0) is the DC coefficient). Due to the fact that the modification of a DCT coefficient affects the whole block, the distortion introduced by the DCT embedding is larger (PSNR=42.8dB) than for embedding in the pixel domain (PSNR=50.9dB). In both cases, the denoising attack is able to increase the PSNR between the attacked and original image in comparison with the PSNR between the watermarked and original image. This means that the power of the watermark has been decreased, and corresponds to a reliable estimation of the watermark component. These results highlight the fact that the estimation error of the secret carrier $\mathbf{a}_k$ and also the independent source $\mathbf{s}_k$ (which depends on the estimation of all the basis vectors) depends on the number of observations which is here represented by the number of blocks that is used to build the matrix $\mathbf{X}$. A gain of 2dB for the PSNR can be achieved using 19000 blocks instead of 1200. Comparison of the presented denoising attack against the moving attack confirms the superiority of independent source estimation in comparison with this additive blind attack which always yields to a PSNR that is above the PSNR between the original and watermarked image.

	Pixel domain				DCT domain			
Initial PSNR	50.9 dB				42.8 dB			
	Denoising attack		Moving attack		Denoising attack		Moving attack	
Nb of blocks	BER	PSNR	BER	PSNR	BER	PSNR	BER	PSNR
19000	37%	54.4dB	37%	49.6dB	41%	46.1dB	41%	40.3dB
4800	42%	53.9dB	42%	49.3dB	43%	45.7dB	43%	40.2dB
1200	40%	52.3dB	40%	49.2dB	37%	44.3dB	37%	40.8dB

Table 3. Denoising performance as a function of different number of blocks.

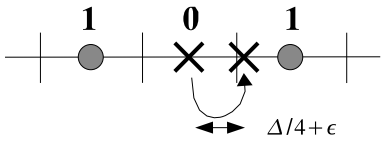


Fig. 11. Principle of the moving attack.

5.4 Remarks on the security of DCT quantization schemes

In this section, the ability to erase the watermark is studied as a function of the position of the watermarked DCT coefficient. We have applied the watermark removal procedure has described previously, for different positions of the DCT coefficient. The resulting distortion between the attacked and original image is

illustrated in Table 4. As a general rule, the efficiency of the proposed attack is lower for the low-frequency coefficient than for the high-frequency ones. For example, if the coefficient (1,0) is watermarked, the PSNR is only 34.1 dB vs. 46.5dB for an embedding in the coefficient (3,3). Such a result is due to the fact that low frequency coefficients convey more information in natural images than high-frequency ones. Consequently it is more difficult to separate the information relative to the image and to the watermark in the first case. This remarks confirms the motivation given in [14] which suggests to watermark low-frequency components for security purposes. Nevertheless it is also important to point out that a compromise between the security and the perceptibility has to be considered by the embedder: the perturbation of low-frequency DCT coefficients produce a more significant impact of the visual system than high frequency ones.

Coeff	(0,1)	(1,0)	(0,2)	(1,1)	(2,0)	(1,2)	(2,1)	(0,3)	(2,2)	(1,3)	(3,0)	(3,1)	(2,3)	(3,2)	(3,3)
PSNR	34.1	35.9	38.9	39.1	40.3	41.5	42.6	42.7	43.7	43.9	44.3	44.8	45.2	45.6	46.5

Table 4. Attack performance for different coefficients (4800 blocks, PSNR=45.29 dB, $\Delta = 150$).

6 Future work and concluding remarks

This paper addresses the security of quantization-based schemes which are supposed to be secure because of the use of the dithering vector and/or secret location of quantized coefficients. We have show that, under several hypothesis (natural images where one coefficient per block has been watermarked, small blocks) it is possible to estimate the secret carrier. Because Principal Component Analysis is not a suitable tool in this case, we have proposed to use Independent Component Analysis (ICA) to estimate the secret carrier and then to perform a removing attack. This is due to the fact that the DM-QIM watermarking process can be seen as the addition of an independent uniform noise that can be extracted using ICA. We have also outlined that the ability to separate the watermark from the original image increases as a function of the position of the DCT coefficient and as a function of the number of processed blocks.
In the future, we would like to see if such an approach can be used for other popular substitutive schemes working in the DCT domain such as the scheme proposed by [15]. In general, we expect that attacks utilizing the statistics of natural images will play an important role in the security of image watermarking schemes.

7 Aknowlegments

The work described in this paper has been supported (in part) by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT, the National French project Fabriano and by the Academy of Finland (project # 205742).

References

1. Kalker, T.: Considerations on watermarking security. In: Proc. of MMSP, Cannes, France (2001) 201–206
2. Cayre, F., Fontaine, C., Furron, T.: Watermarking security part I: Theory. In: Proceedings of SPIE, Security, Steganography and Watermarking of Multimedia Contents VII. Volume 5681., San Jose, USA (2005)
3. Kalker, T.: A security risk for publicly available watermark detectors. In: Proc. Benelux Inform. Theory Symp., Veldhoven, The Netherlands (1998)
4. Furon, T., Duhamel, P.: An asymmetric watermarking method. IEEE Trans. on Signal Processing **51** (2003) 981–995 Special Issue on Signal Processing for Data Hiding in Digital Media and Secure Content Delivery.
5. Chen, B., Wornell, G.W.: Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Transactions on Information Theory **47** (2001) 1423–1443
6. Cayre, F., Fontaine, C., Furron, T.: Watermarking security part II: Practice. In: Proceedings of SPIE, Security, Steganography and Watermarking of Multimedia Contents VII. Volume 5681., San Jose, USA (2005)
7. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. John Wiley & Sons (2001)
8. van Hateren, J.H., Ruderman, D.L.: Independent component analysis of natural image sequences yields spatiotemporal filters similar to simple cells in primary visual cortex. Proc. Royal Society **B** (1998) 2315–2320
9. Hyvärinen, A.: Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation. Neural Computation **11** (1999) 1739–1768
10. J.Ferreira, A., Figueiredo, M.: Class-adapted image compresion using independent component analysis. In: Proc. ICIP, Barcelona (2003)
11. Hyvärinen, A.: Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks **10** (1999) 626–634
12. Hyvärinen, A., Hoyer, P.O., Oja, E.: Sparse code shrinkage: Denoising by nonlinear maximum likelihood estimation. In: Proc. of Advances in Neural Information Processing Systems 11 (NIPS*98). (1999) 473–479
13. Voloshynovskiy, S., Pereira, S., Herrigel, A., Baumgärtner, N., Pun, T.: Generalized watermark attack based on watermark estimation and perceptual remodulation. In Wah Wong, P., Delp, E.J., eds.: EI'2000: Security and Watermarking of Multimedia Content II. Volume 3971 of SPIE Proceedings., San Jose, California USA (2000) 358–370
14. Cox, I., Killian, J., Leighton, T., Shamoon, T.: Secure spread sprectrum watermarking for images, audio and video. In: Int. Conf. on Image Processing (ICIP). Volume 3., IEEE (1996) 243–246
15. Zhao, J., Koch, E.: Embedding robust labels intro images for copyright protection. In: Int. Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies, Vienne, Autriche (1995)