An Asymmetric Public Detection Watermarking Technique

Teddy Furon, Pierre Duhamel

An Asymmetric Public Detection Watermarking Technique

Teddy FURON[1] and Pierre DUHAMEL[2]

[1] THOMSON multimedia R/D France, User Interface Interactivity and Security Lab, 1, av. Belle Fontaine, 35510 Cesson Sévigné, France
furont@thmulti.com

[2] Ecole Nationale Supérieure des Télécommunications de Paris, Laboratoire Traitement Signaux et Images, 46 rue Barrault, 75013 Paris, France
duhamel@sig.enst.fr

Abstract. The new watermarking technique[1] presented in this paper is an example of an asymmetric public detection scheme. The detection process does not need the original picture nor the secret key used in the embedding process. It is the translation, in the watermarking domain, of a public key pair cryptosystem. The key idea is to filter the pseudo-noise sequence before embedding it in the cover-content. Contrary to classical techniques, the heart of the detection algorithm is not a correlation measure but a consistent statistical test hypothesis in spectral analysis. Some signal based considerations show that knowing the public key used in the detection process is no use for pirate who wants to discover the secret key. An example of a copyright protection system for digital content using this technique is briefly described.

1 Introduction

Watermarking is the art of embedding information in a cover-content in a robust and non-perceptible way. Therefore, the quality of a watermarking technique can be expressed in terms of capacity, non-perceptibility and robustness. Another distinction is whether the technique supports private watermark, where the original cover-content is needed to extract hidden information, or public watermark, where one can detect the embedded message without the original content. The terminology of public watermarking was set by B. Pfitzmann [1] during the first international Workshop on Information Hiding and is depicted in Fig. 1. It clearly appears that the embedding and detection processes have a common parameter called the key. Thus, comparing to a cryptography system, current watermarking techniques are symmetric schemes [12]. The key parameter is usually called the secret key in reference to Kerckhoffs cryptographic principle [14]: A cryptographic algorithm can not remain secret, hence the security of a cryptosystem must only rely on the key kept in a safe place. A really important constraint is intrinsic to the symmetry of these schemes. Every entity able to

[1] French patent application number 99-07139 filled on the first of June 1999

detect a watermark shares the same secret key as the watermarker and thus can erase it or change the embedded message. The watermarker has to give his secret key in a secure way only to trusted entities. This constraint restricts drastically the use of watermarking technique in many domains. It is well known that no secret can be stored in consumer electronic devices or software. Smart-cards, which are considered to be the only secure equipment in consumer electronic domain, are not powerful enough or too expensive to support a complete watermark detection process. Indeed, it seems that the only way to use watermark in secure way is the case where a content owner proves its ownership detecting his watermark in the presence of a lawyer in order not to reveal his secret key in public audiences.

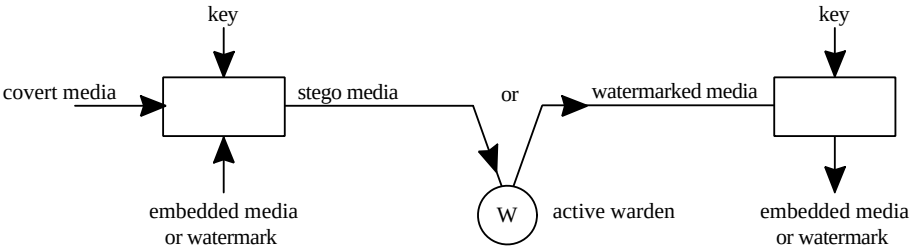

Fig. 1. Usual watermarking scheme and terminology

This major drawback has been solved in cryptography thanks to asymmetric schemes: encryption and decryption processes do not use the same key. The global system is based on a public keys pair where the private key, for example in case of certificate signature, is used in the encryption process and the public key in the decryption process. Obviously, the public key depends on the private key although it is impossible to deduce the private key from it. Transposing this idea in the watermarking domain would mean that the embedding and detection processes do not share a common secret. Moreover, knowing the detection algorithm and all its parameters is not enough to deduce the secret parameter used in the embedding process and besides it does not bring any clue in order to remove the watermark from the stego-content. F. Hartung has already presented a kind of public key watermarking technique based on classical spread spectrum watermaking scheme [19]. But it does not achieve all the desired specifications. Especially, a pirate can remove a part of the watermark. Whereas there is enough watermark signal left to allow the owner to retrieve it thanks to his private key, a detector with the public key can no more detect it.

This paper introduces a new watermark technique that is asymmetric and indeed completely different from classical spread spectrum technique. Following the reasoning of the authors, this paper firstly focuses on the cryptographic RSA scheme in order to derive some useful comparisons with the signal processing domain. To be more precise, this leads to the well known issue of blind

equalization in the digital communication domain. Thanks to these comparisons, an asymmetric watermarking technique is then built up and tested with classical simulations. Finally, the use of these asymmetric watermarking technique is discussed.

2 From Cryptography to Signal Processing

Cryptography and watermarking are obviously linked, at least because watermarking, contrary to steganography, has to be robust and secure. But it is extremely hard to mix in a fruitful way these two scientific domains: Cryptography is based on number theory whereas watermarking tackles real signals. Usually, cryptography and watermarking are used separately, one after the other. The message to be hidden is encrypted before being embedded in the cover-content, or the stego-content is encrypted after the watermark embedding. But these combinations of cryptographic and watermarking techniques do not obviously improve the security or the functionalities of the global system. For example, imagine that one encrypts the message with an asymmetric cryptographic cipher like RSA, before embedding it in the cover-content with a classical spread spectrum technique, in order to build an asymmetric watermarking scheme. It is true that the only person able to embed a message is the RSA and watermark private keys holder. But, a pirate, knowing the detection process because he hacked a software or did the reverse engineering of a consumer electronic device or transformed the stego-content by an efficient attack like Stirmark [13], can erase the watermark whatever the security level of the cryptosystem used. This example is clearly not a good design of asymmetric watermarking technique. In most cases, the weakest link in a security point of view and the most constraining function is the watermark detection. That is the reason why it is necessary to invent a completely different watermarking technique truly asymmetric, getting inspired by the cryptography domain.

2.1 Example of RSA Cryptosystem

The RSA cryptosystem, invented by R. Rivest, A. Shamir and L. Adleman, is one of the most famous asymmetric schemes. This section gives a short description of the RSA encryption scheme. Key generation is done as follows: choose two large prime numbers p and q, compute $n = pq$ and $\Psi(n) = (p-1)(q-1)$ where Ψ is the Euler's function, select a random integer e prime with $\Psi(n)$ and compute the only one integer d such that $ed = 1 \mod \Psi(n)$. (n, e) is the public key whereas d is the private key. As watermarking is similar to signature in concept, the encryption with the private key is detailed: represent the message as an integer m, compute $c = m^d \mod n$, c is then the encrypted message. Decryption is easy and based on Fermat's theorem: use the public key (n, e) to compute $m = c^e \mod n$. See [2] or [14] for further details.

The security of RSA relies on the difficulty of computing the private key d knowing the public key (n, e). The problem stems from factoring n in prime

numbers: because the pirate can not find (p, q) from n (no efficient algorithm known up to now) he cannot compute $\Psi(n)$ nor d, which is the private key. Hence, it seems (it has not yet been proven that breaking RSA is equivalent to factoring) that the security of the system is based on the fact that $(p, q) \longrightarrow n = pq$ is a one way operation.

2.2 One Way Operation in Signal Processing

Compared to RSA algorithm, an asymmetric watermarking technique should add a watermark in the content using a one way operation in order to prevent pirate from removing it (operation equivalent to $(p, q) \longrightarrow n = pq$), but with the property that the detector can still retrieve the watermark (property equivalent to Fermat's theorem). It means that the detector should notice the effect of the watermark without having the knowledge of the cause. A solution was found in the digital communication field. In this domain, signals transmitted are modified by the communication channel. The channel is usually considered as a linear time-invariant (LTI) filter. The receiver has to invert the effect of the channel. This is the role of the equalizer. For this purpose, the transmitter begins to send a reference sequence known by the receiver in order to initially adjust the coefficients of the equalizer. Then, when the really informative signal is transmitted, the receiver is able to compensate the channel effect. As transmitting the referenced sequence takes time and power, communication system are desired not to need such training period. This is the problem of blind equalization based on initial adjustment of the coefficients without the benefit of a training sequence. But these methods are less efficient than classical ones especially at low signal to noise power ratio. A one way operation can be found every time blind equalization is not possible. This asymmetric technique is based on the main idea that passing through a filter, whatever white noise will produce a sequence noticeable by the shape of its power spectrum density (psd). But, with some assumptions explained in the next section, it is impossible to retrieve the original sequence. Hence, while the detection only consists of checking the psd of the watermark, a pirate can not estimate and remove the watermark. This referenced psd is the public key, whereas the private key is the set of the white noise sequence and the filter coefficients.

A Signal Processing Theorem. Second-order statistics like auto-correlation and density spectrum function of the output signal of a LTI filter provide information only on the magnitude of the filter characteristics (except for periodic signals). This statement can be done regarding the following theorem:

Consider a discrete LTI filter which impulse response $\mathbf{h} = \{h_n\}$ is real and frequency response is noted $H(f)$. Its input signal is noted $\mathbf{v} = \{v_n\}$, issued from a stationary random process, and its output $\mathbf{w} = \{w_n\}$. The following equations

(1) can be written:

$$\phi_{\mathbf{ww}}[k] = \sum_{m=-\infty}^{\infty} \phi_{\mathbf{vv}}[k-m] \sum_{u=-\infty}^{\infty} h_u h_{u+m} \quad \text{and} \quad \Phi_{\mathbf{ww}}(f) = |H(f)|^2 \cdot \Phi_{\mathbf{vv}}(f) \tag{1}$$

where $\phi_{\mathbf{ww}}[k]$ is the auto-correlation discrete function of $\mathbf{w}$ and $\Phi_{\mathbf{ww}}(f)$ its power spectrum density. Observing $\Phi_{\mathbf{ww}}(f)$, there is, a priori, no way to estimate $H(f)$ (or equivalently $\mathbf{h}$) due to the phase non-determination [3]. A known exception are minimum (or maximum) phase filter, because the phase of $H(f)$ is related to its magnitude.

Blind Deconvolution. Blind equalization techniques manage however to estimate filter coefficients. In the case of SISO (Single Input Single Output) system, there are mainly three classes of blind equalization algorithms based on maximum likelihood criterion, stochastic gradient iteration or high order signal statistics [3]. These techniques do not work if the broadcast signal is a gaussian white noise [18]. All the information concerning this random process are given by second-order statistics, which, according to the signal processing theorem above, bring only information on the magnitude of the filter process. But in the SIMO case (Single Input Multiple Outputs), an algorithm derived from the subspace method for example, is able to estimate the impulse responses of the different filters. These estimations feed a classical equalizer which retrieve the input of the SIMO bank of filters. In the watermarking technique described hereafter, the SISO case is mandatory to prevent the pirate from estimating the filter and retrieving the secret sequence. Thus, only one filter is used in the embedding process.

Hence, in the SISO case, if sequence $\mathbf{v}$ is issued from a gaussian stationary random process and $\mathbf{h}$ is not the impulse response of a minimum or maximum phase filter, $(\mathbf{v}, \mathbf{h}) \longrightarrow \mathbf{h} \otimes \mathbf{v}$ is a signal processing one way operation ($\otimes$ is the convolution product).

3 Asymmetric Watermarking Technique

3.1 Embedding Process

Let the sequence $\mathbf{x} = \{x_n\}$ of length N represent the content (luminance of pixels [4], DCT coefficients [6], Fourier-Mellin transform coefficients [5], wavelet transform coefficients for still picture, location of grid nodes of computer image, facial animation parameters of MPEG-4 head object, sample of sound [15]...). Assume there is an algorithm based on some Human Perception Model considerations which is able to calculate the amount of noise that each coefficient can bear without perceptible quality loss. Its output is the HPM modulation sequence noted $\mathbf{p} = \{p_n\}$ with $p_n \geq 0 \quad \forall n \in [1..N]$.

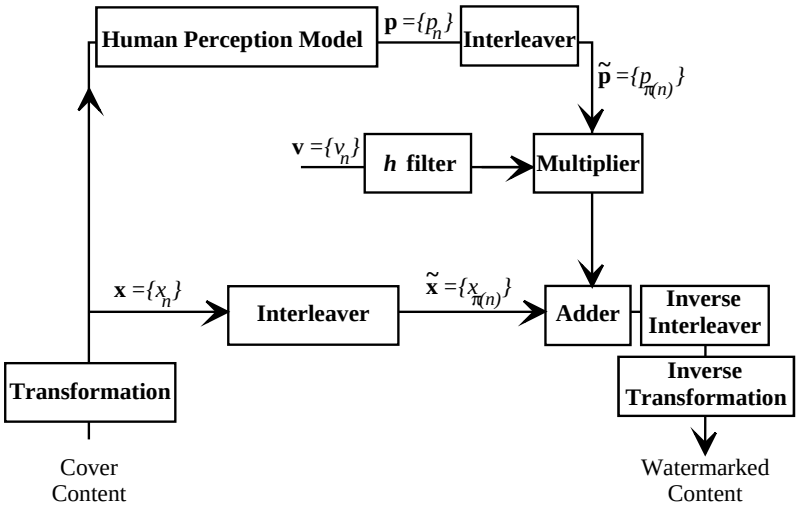

Fig. 2. Design of the embedding process

An interleaver is used in this embedding process. Its role is to mix the sequence $\mathbf{x}$ as a random permutation π. Output sequences are noted with the symbol $\tilde{}$. Hence, $\widetilde{x}_n = x_{\pi(n)} \quad \forall n \in [1..N]$.

According to Fig. 2, embedding process leads to the equation (2).

$$y_n = x_n + p_n.(\mathbf{h} \otimes \mathbf{v})_{\pi^{-1}(n)} \qquad \forall n \in [1..N] \tag{2}$$

with $\mathbf{v}$ white noise distributed as $\mathcal{N}(0, \sigma_{\mathbf{v}})$. Assuming that $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{v}$ are statistically independent sequences, equations (3) hold:

$$\phi_{\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}}[k] = E[\widetilde{y}_n \widetilde{y}_{n-k}] = E[(\widetilde{x}_n + \widetilde{p}_n.(\mathbf{h} \otimes \mathbf{v})_n).(\widetilde{x}_{n-k} + \widetilde{p}_{n-k}.(\mathbf{h} \otimes \mathbf{v})_{n-k})]$$

$$\phi_{\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}}[k] = \phi_{\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}}[k] + E[\widetilde{p}_n \widetilde{p}_{n-k}. \sum_u \sum_m h_u h_m v_{n-u} v_{n-k-m}]$$

$$\phi_{\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}}[k] = \phi_{\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}}[k] + \phi_{\widetilde{\mathbf{p}}\widetilde{\mathbf{p}}}[k].\sigma^2_{\mathbf{v}}.(\mathbf{h} \otimes \mathbf{h})_k \tag{3}$$

where E is the statistical expectation, $\sigma^2_{\mathbf{v}}$ is the variance of the sequence $\mathbf{v}$. Assuming the interleaver is perfect, that is to say its output sequences are white and stationary, then simplifications leads to:

$$\phi_{\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}}[k] = \mu_{\mathbf{x}}^2 + (\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{p}}^2.\sigma_{\mathbf{v}}^2.\sum_u h_u^2).\delta[k] + \mu_{\mathbf{p}}^2.\sigma_{\mathbf{v}}^2.(\mathbf{h} \otimes \mathbf{h})_k \quad (4)$$

and

$$\Phi_{\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}}(f) = \mu_{\mathbf{x}}^2.\delta(f) + (\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{p}}^2.\sigma_{\mathbf{v}}^2.\sum_u h_u^2) + \mu_{\mathbf{p}}^2.\sigma_{\mathbf{v}}^2.\left|H(f)\right|^2 \quad (5)$$

The secret key in the embedding process is the set of sequences $\mathbf{h}$ and $\mathbf{v}$.

3.2 Detection Process

The detection process is based on spectral analysis. Let the sequence $\mathbf{r} = \{r_n\}$ of length N represent the received content. The goal is to test two hypothesis:

- G_0: the received content is not watermarked so the power spectral density of the interleaved received sequence $\widetilde{\mathbf{r}}$ is flat. The estimated psd is expressed as $g_0(f) = \sigma_{\mathbf{r}}^2 + \mu_{\mathbf{r}}^2.\delta(f) \quad \forall f \in]-\frac{1}{2}, \frac{1}{2}]$.
- G_1: the received content is watermarked so the power spectral density of the interleaved received sequence $\widetilde{\mathbf{r}}$ is estimated as
  $g_1(f) = \mu_{\mathbf{x}}^2.\delta(f) + \mu_{\mathbf{p}}^2.\sigma_{\mathbf{v}}^2.\left|H(f)\right|^2 + C \quad \forall f \in]-\frac{1}{2}, \frac{1}{2}]$
  such that $C = \sigma_{\mathbf{r}}^2 - \int_{-\frac{1}{2}}^{\frac{1}{2}} \mu_{\mathbf{p}}^2.\sigma_{\mathbf{v}}^2.\left|H(f)\right|^2.df - \mu_{\mathbf{x}}^2$ according to (5).

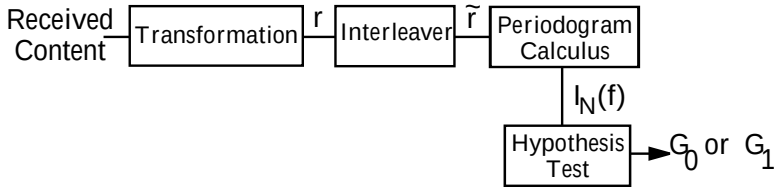

Fig. 3. Design of the detection process

In [9], a test defined by the critical region expressed in (6) is proved to be asymptotically equivalent to the likelihood ratio test.

$$\{\widetilde{\mathbf{r}} \quad | \quad 2.N.[U_{N,0}(\widetilde{\mathbf{r}}) - U_{N,1}(\widetilde{\mathbf{r}})] > d(P_{fa})\} \quad (6)$$

where $d(P_{fa})$ is a threshold depending on a desired false alarm probability and $U_{N,i}$ is the significant quantity related to the hypothesis G_i of the principal part

of the log likelihood. If the estimated power spectral density is strictly positive, then $U_{N,i}$ expression is simply (7).

$$U_{N,i}(\widetilde{\mathbf{r}}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} (\log g_i(f) + \frac{I_N(f)}{g_i(f)})df \qquad i \in \{0,1\} \tag{7}$$

where $I_N(f)=\frac{1}{N}\left|\sum_{k=1}^{N} \widetilde{r}_k.\exp(2\pi jfk)\right|^2$ is the periodogram of the interleaved received sequence $\widetilde{\mathbf{r}}$, following a χ^2 distribution with two degrees of freedom [10].

The public key in the detection process is $|H(f)|$.

3.3 Security Point of View

Classical Spread Spectrum Technique. Keeping the same notation, classical spread spectrum technique can be sum up with the equation (8).

$$y_n = x_n + p_n.v_n \quad \forall n \in [1..N] \tag{8}$$

Detection is done via a correlation: a content is declared watermarked if the correlation with the referenced sequence is superior to a fixed positive threshold.

$$\langle \mathbf{y}|\mathbf{v}\rangle = \langle \mathbf{x}|\mathbf{v}\rangle + \langle \mathbf{p}.\mathbf{v}|\mathbf{v}\rangle > d'(P_{fa}) \tag{9}$$

where $\langle \mathbf{y}|\mathbf{v}\rangle = \sum_n y_n v_n$. The aim of a pirate is to create from a watermarked content a cleared content no more detectable. This can be easily done knowing the sequence $\mathbf{v}$ which is the secret key of this scheme:

$$\mathbf{y}' = \mathbf{y} - \frac{\langle \mathbf{y}|\mathbf{v}\rangle}{\langle \mathbf{v}|\mathbf{v}\rangle}.\mathbf{v} \tag{10}$$

The resulting sequence $\mathbf{y}'$ is not equal to the original one $\mathbf{x}$ due to the HPM modulation sequence $\mathbf{p}$. However, if the quality of the resulting content is correct, the pirate achieved his goal. It clearly highlights the importance of the safety of the secret key. If this scheme is designed to be used widely, such detector can not be implemented in non-secure electronic component. But, 'stealing' the secret key is not the only way to achieve pirate's aim. One can try to estimate the sequence $\mathbf{v}$ via an average process of T different watermarked contents via formula (11).

$$\widehat{\mathbf{v}} = \frac{1}{T}\sum_{k=1}^{T} \mathbf{y}_k = \frac{1}{T}\sum_{k=1\mathbf{x}}^{T} \mathbf{x}_k + (\frac{1}{T}\sum_{k=1}^{T} \mathbf{p}_k).\mathbf{v} \approx \alpha.\mathbf{v} \tag{11}$$

The only solution to avoid this attack is to desynchronise the embedded pseudo-noise sequence $\mathbf{v}$: $y_{n,k} = x_{n,k} + p_{n,k}.v_{n-t_k} \quad \forall n \in [1..N]$. But then, the detector has to get resynchronised. It means that it has to find the delay t_k for a given

content $\mathbf{r}_k$, calculating as many correlation as the number of possible delays. Ton Kalker and *al* [11] give a very efficient and nice implementation of this method. Nethertheless, I.J. Cox and J.P. Linnartz pointed out a more powerful attack. Assuming the pirate has a detector device he can use as many times he likes (which is the case in consumer electronic), he can finally manage to create a cleared content $\mathbf{y}'$ using the detector $O(N)$ times. This security flaw is due to the linearity of the correlation. See [12] for further details.

Asymmetric Technique. Thanks to the design of the asymmetric watermarking technique above, these threats hold no more. Obviously, the sequence $\mathbf{v}$ is stored nowhere in the detection process so nobody can steal it. This sequence is different for each content, so the average attack is useless. And finally, the detection algorithm is not linear which makes the attack of [12] non valid.

Although this watermarking technique has, for the moment, only one bit of capacity, a copy protection system for digital contents can be based on it. The idea is, as usual, to add a header to each copyrighted content. This header contains important data related to the content (identification number, rights granted to the user...) and will be bound to the content via a cryptographic signature. This is usually called a certificate. Pirate can not modify a copyrighted content or its certificate because the digital signature is then no more valid. But, he can remove the certificate, pretending the hacked content is a personal creation or whatever not copyrighted. That is the reason why copyrighted content are watermarked with an asymmetric technique. The role of this watermark is to warn the device that the content it deals with, is copyrighted. This device will read data in the certificate and will check its signature. If the signature is non valid or if no certificate is present whereas the content is watermarked, the device refuses to deal with this content. Notice that the device has no secret key but two public keys: one for the cryptosystem verifying signature, another for the watermark asymmetric detector. The content's owner has two secret keys: one for the cryptosystem making signature, another for the watermark embedding process.

A restriction may appear in the use of this watermarking technique. The comparison with asymmetric cryptosystems is not completely fulfilled. Knowing the public key $|H(f)|$, everybody can build its own private key ($\mathbf{h}'$, $\mathbf{v}'$) provided that $|H'(f)| = |H(f)| \quad \forall f \in]-\frac{1}{2}, \frac{1}{2}]$. As a watermark usually induces a restriction of user's rights like in the copy protection system described above, this fact is not really a dead end. Notice that the owner can still prove its ownership detecting his watermark, via a classical correlation detector, in the presence of a lawyer in order not to reveal his secret key ($\mathbf{h}$, $\mathbf{v}$) in public audiences.

4 Simulation

The goal of these simulations is to prove the validity of this new concept. Details of implementation are first given. A small panel of 512×512 pixels pictures is used (Lena, peppers and mandrill). Only the luminance data coded in 256 grey

levels are watermarked. x_n is the luminance of the pixel located in (i, j) such that $n = i + (j-1).512$.

4.1 Interleaver

Two interleavers are used. The direct interleaver tidies luminance pixels in a row of length 512^2 and mixes it according to a random given permutation vector. The inverse interleaver will apply the inverse permutation and will tidy the resulting row in a 512×512 pixels picture. The random permutation vector is calculated with the Moses and Oakford algorithm [17].

4.2 Human Visual System Modulation Sequence

A basic algorithm is used to calculate the modulation sequence $\mathbf{p}$. The image $\mathbf{X}$ is filtered with a laplacian high-pass filter. The absolute value of this result is tidied in a sequence $\mathbf{p}$.

$$p_{i+(j-1).512} = |\lambda \otimes \mathbf{X}|_{i,j} \qquad \lambda = \begin{pmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{pmatrix} /9 \tag{12}$$

Textured regions or edges lead to high coefficients whereas uniform regions lead to very low values. This follows very roughly the eye's behavior, but no precise theoretical model sustains this choice. This algorithm is very fast and experimental results are satisfactory as noted in [11].

4.3 Role of the Length of Sequences

To embed the watermark, two strategies can be chosen. Add a filtered pseudo-random sequence $\mathbf{w}$ as long as the sequence $\widetilde{\mathbf{x}}$, or use a shorter pseudo-random sequence that one repeats several times before adding it to the sequence $\widetilde{\mathbf{x}}$. This last choice is usually made in classical spread spectrum technique [11]; to detect the watermark, the received sequence $\widetilde{\mathbf{r}}$ is then averaged. Hence, the watermark to cover-content power ratio is increased leading to a better false alarm probability. But, this 'tiling' process is also interesting for a pirate using the average attack. With the asymmetric watermarking scheme described before, the two strategies are illustrated. N_{seq} is the number of time the filtered pseudo-random sequence is repeated. Thus, $N_{seq} * N = 512^2$. Original and watermarked pictures are given to the detector. The result is the quantity $2.N.[U_{N,0}(\widetilde{\mathbf{r}}) - U_{N,1}(\widetilde{\mathbf{r}})]$. The sign of this quantity figures out if the content is watermarked (positive) or not (negative), whereas its absolute value shows how reliable is the decision. It means that the threshold $d(P_{fa})$ is set to zero. Figure 4 plots the average of these quantities calculated for the three different pictures and for 30 different pseudo-random sequences. On the abscissa is the parameter N_{seq}. It appears that high reliability occurs when embedding long sequences rather than short ones tiled several times, thanks to the consistence of the test hypothesis designed. The conclusion of this simulation is to set the parameter N_{seq} to one, which means that the sequence $\mathbf{w}$ is not repeated.

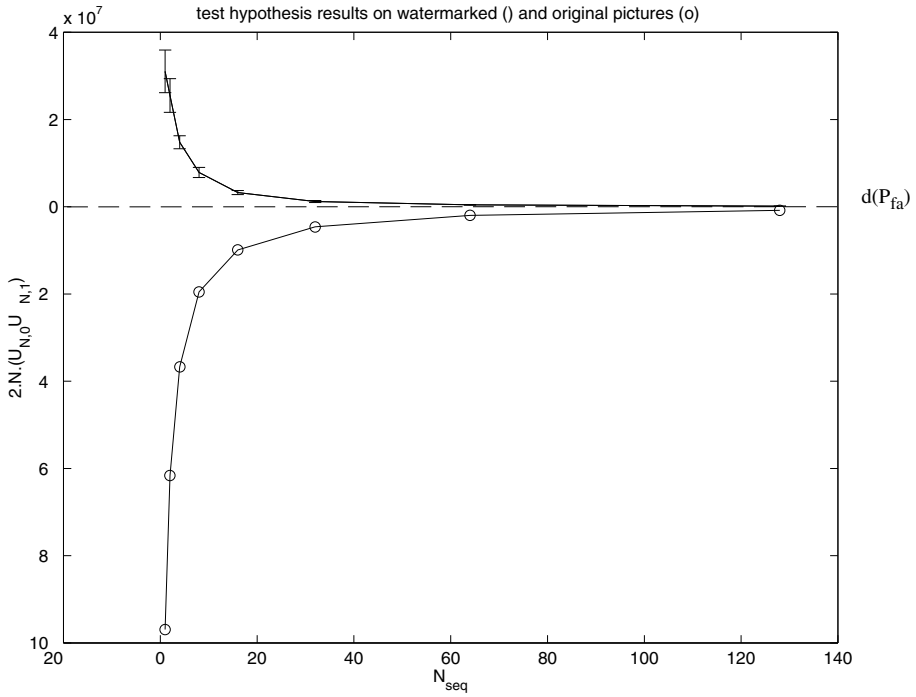

Fig. 4. Role of the length of the sequences

4.4 JPEG Test

The three pictures watermarked with the parameter N_{seq} set to 1, are then compressed with the JPEG algorithm with a quality factor Q. Figure 5 plots the average of the resulting quantities $2.N.[U_{N,0}(\widetilde{\mathbf{r}}) - U_{N,1}(\widetilde{\mathbf{r}})]$. For quality factors Q higher than 10, the results are far bigger than the one computed with the original image. But, as the results are indeed compared to $d(P_{fa})$, the watermark is only robust to $Q = 70$ JPEG compression. This is not a good robustness compared to actual watermarking techniques. Indeed, it can be compared to the first watermarking techniques presented a few years ago. But, authors believe a far better robustness can be reached if DCT or DWT coefficients are watermarked instead of pixels' value. In the same way, Fourrier-Mellin transformation might be useful to derive an asymmetric watermarking technique robust to rotation, scaling and translation processes. The simualtion shows that the basic concept of the watermarking technique is valuable and allows us to foresee a fair robustness using clever transformations like DCT or DWT coefficients.

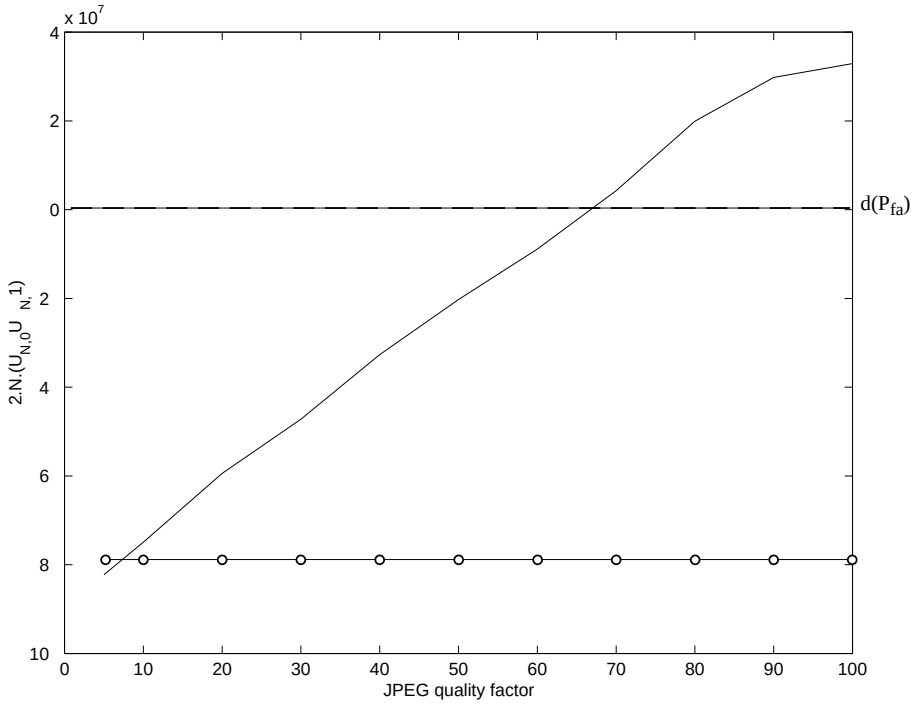

Fig. 5. Robustness to JPEG compression

5 Conclusion

This article is a description of the concept of the first truly asymmetric watermarking technique. The main advantage is that no secret is stored in the detector. Hence, the owner of copyrighted content must not rely on the security of the detection device. Thanks to the consistence of the detector, this technique is absolutely robust against average attack. The simulations show that this concept is valuable and may solve the open issue left in [16] about 'public key steganography against an active warden'. There are several technical problems left which need to be solved: some are related to the concept itself (large size of the public key, synchronization, increase of capacity, domain of application restricted...), others are related to the implementation of it (which transformation has to be used in order to achieve better robustness against scaling, rotation and compression - Can this concept be used for sound samples?).

6 Acknowledgments

The authors would like to acknowledge Eric Moulines (ENST Paris) for very fruitful discussions on test hypothesis in spectral analysis.

References

[1] B. Pfitzmann, "Information Hiding Terminology", in *Proceedings of the First Int. Workshop on Information Hiding*, May 1996.

[2] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, Computer Sciences Applied Mathematics Engineering, CRC press, 1997.

[3] J. G. Proakis, *Digital Communications*, Electrical Engineering Series, McGraw-Hill International Editions, third edition, 1995.

[4] J. R. Smith and B. O. Comiskey, "Modulation and Information Hiding in Images", in *Proceedings of the First Int. Workshop on Information Hiding*, May 1996.

[5] J. J. K. O'Ruanaidh, T. Pun "Rotation, Scale and Translation Invariant Spread Spectrum Digital Image Watermarking", in *Signal Processing, v 66 no 3*, May, 1998.

[6] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Images, Audio and Video", in *Proceedings of IEEE-ICIP'96, v III*, Oct. 1996.

[7] M. Barni, F. Bartolini, V. Cappellini, A. Lippi, and A. Piva, "A DWT-based Technique for Spatio-frequency Masking of Digital Signatures", in *Proceedings of SPIE Security and Watermarking of Multimedia Contents, v 3657*, Jan 1999.

[9] K. Dzhaparidze, *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer Series in Statistics, Springer-Verlag, 1986.

[10] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Computer Science and Scientific Computing, Academic Press, second edition, 1990.

[11] T. Kalker, G. Depovere, J. Haitsma, and M. Maes, "A Video Watermarking System for Broadcast Monitoring", in *Proceedings of SPIE Security and Watermarking of Multimedia Contents, v 3657*, Jan 1999.

[12] I. J. Cox and J. P. Linnartz, "Some General Methods for Tampering with Watermark", in *IEEE Journal on Selected areas in communications, v 16 no 4*, May 1998.

[13] F. A. P. Petitcolas, R. Anderson, and M. Kuhn, "Attacks on copyright marking systems", in *Proceedings of the Second Int. Workshop on Information Hiding*, April 1998.

[14] B. Schneier, *Applied Cryptography*, John Wiley & Sons, Inc., second edition, 1996.

[15] M.D. Swanson, B. Zhu, A.H. Tewfik and L. Boney, "Robust Audio watermarking using perceptual masking", in *Signal Processing, v 66 no 3*, May, 1998.

[16] R. Anderson, "Stretching the limits of steganography", in *Proceedings of the First Int. Workshop on Information Hiding*, May 1996.

[17] D. Knuth, *The art of computer programming*,vol II, Addison-Wesley series in Computer Science and Information Processing, second edition, 1981.

[18] D. Donoho, *On minimum entropy deconvolution*, D. Findley Academic Press in Applied Time Series Analysis, second edition, 1981.

[19] F. Hartung and B. Girod, "Fast Public-Key Watermarking of Compressed Video" , in *Proceedings IEEE International Conference on Image Processing*, October 1997.