UNIVERSITY
OF TRENTO

DEPARTMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY

38050 Povo – Trento (Italy), Via Sommarive 14
http://www.dit.unitn.it

AN IMPROVED ASYMMETRIC WATERMARKING SCHEME SUITABLE FOR COPY PROTECTION

G. Boato, F.G.B. De Natale, and C. Fontanari

July 2005

Technical Report DIT-05-056

An Improved Asymmetric Watermarking Scheme Suitable for Copy Protection

G. Boato[1]*, F. G. B. De Natale[1] and C. Fontanari[2]

[1]Dept. of Information and Communication Technology, University of Trento,
Via Sommarive 14, I-38050, Trento, Italy, tel +39 0461 883917, fax +39 0461 882093

[2]Dept. of Mathematics, Fac. of Information Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129, Torino, Italy, tel +39 011 5647567 fax +39 011 5647599

boato@dit.unitn.it; denatale@ing.unitn.it; claudio.fontanari@polito.it

Abstract

A major limitation of some recently proposed asymmetric watermarking techniques based on linear algebra lies in the strong dependence of the watermark on the original image. The present letter suggests an alternative scheme which is not only secure against projection attack but allows also the insertion of arbitrary watermarking sequences.

I. Introduction

Watermarking security is arising a great deal of interest in both accademia and industry (see for instance [1] and [2]). The analogy with public key cryptography suggests to consider asymmetric structures, involving a private key for embedding and a public key for detection. However, this is by no means sufficient in order to make a watermarking scheme secure: as remarked in [2], § 5, almost all available asymmetric watermarking schemes can be defeated by a standard closest point or projection attack (see Section III below for details).

In order to protect the boundary of the detection region from unauthorized access, it has been proposed to exploit sophisticated mathematical tools such as the theory of fractals (see for instance [3]). From this point of view, the approach of the recent paper [4] sounds particularly appealing, since it introduces an asymmetric watermarking scheme based just on elementary linear algebra, which is proven to be secure under projection attack. Unfortunately, in order to achieve such a property, the watermark cannot be chosen arbitrarily, but it turns out to be heavily dependent on the host image (see in particular statement c) of the Theorem on p. 787, which shows that the watermark is forced to be a suitable multiple of a sequence deterministically extracted from the original image). As a consequence, the proposed method is appropriate just for copyright protection, where only one key is assigned to each image, but definitely not for copy protection, where every recipient is identified by its own key.

In this letter, we propose a substantial improvement of this approach, which makes it suitable also for copy protection, allowing the insertion into any image of an arbitrary watermarking sequence. The robustness of the method against standard image degradation operators relies on the choice of the watermark space proposed in [5] for the symmetric case: therefore, we can just refer to the simulation results in [5], as the authors of [4] do in Section V.B. 2) on p. 789. On the other hand, the security of the method can be discussed in a purely mathematical setting exactly as in [4] (see Theorem 1 below). This makes of no value a specific experimental analysis at this stage.

II. Watermarking procedure

We are going to describe a subspace asymmetric watermarking procedure. In asymmetric watermarking the encoding and decoding algorithms as well as the detection key are known, while the embedding key is kept secret. As in [4], let us fix an integer $n \geq 1$ and a feature space $\mathcal{X}$ (for instance, the space corresponding to the entries in the top left corner of the DCT) and, after a singular value decomposition analysis, decompose it into two orthogonal subspaces $\mathcal{W}$ of dimension $2n$ (the perceptually robust component) and $\mathcal{V}$ (the component more susceptible to standard modifications of the image). Next, we split $\mathcal{W}$ into two orthogonal subspaces $\mathcal{G}$ and $\mathcal{H}$ of dimension n and we choose matrices G and H whose columns form an orthonormal basis of $\mathcal{G}$ and $\mathcal{H}$, respectively. Finally, we pick an arbitrary watermarking sequence $w \in \mathbb{R}^n$.

A. *Embedding and detection*

Let $\phi_o \in \mathcal{X}$ be the feature vector associated to the original image. We write

$$\phi_o = \psi_o + \sigma_o \tag{1}$$

where $\psi_o \in \mathcal{W}$ and $\sigma_o \in \mathcal{V}$, and

$$\psi_o = Gs + Ht \tag{2}$$

The watermark embedding is defined by

$$\phi_w = \phi_o + Gw \tag{3}$$

where G is the embedding key. Thanks to the properties of the subspace $\mathcal{W}$, the watermark turns out to be robust against standard image degradations.

Next we choose a symmetric matrix A (i.e., $A^T = A$) satisfying

$$A(s + w) = s + w \tag{4}$$

and an orthogonal matrix B (i.e., $B^T = B^{-1}$) satisfying

$$Bt = \mu(s + w) \tag{5}$$

with $\mu := \|t\|/\|s+w\|$ and we define

$$D = AG^T + \mu BH^T \tag{6}$$

which is released to the public and is the crucial ingredient in the detection phase. The matrix A, whose existence is ensured by the trivial choice A=I (the identity matrix), introduces a further degree of freedom which can be exploited to minimize false-positive probability. As far as B is concerned, we point out the following easy fact.

Lemma 1: If $s+w \neq 0$, then we can construct an orthogonal matrix B satisfying (5).

Proof: If $t=0$, just take B equal to the identity matrix. For $t \neq 0$, let $a_1 := \frac{t}{\|t\|}$ and $b_1 := \frac{s+w}{\|s+w\|}$ and complete them to orthonormal bases $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ of $\mathbb{R}^n$ (for instance, complete them to arbitrary bases and then apply the standard Gram-Schmidt orthonormalization process). If M (resp., N) is the matrix with a_i^T (resp., b_i^T) as the i-th column ($i = 1, \ldots, n$), then $M(1,0,\ldots,0)^T = a_1^T$ and $N(1,0,\ldots,0)^T = b_1^T$. The matrix $B := NM^T$ is orthogonal since it is product of orthogonal matrices and $Ba_1^T = NM^Ta_1^T = N(1,0,\ldots,0)^T = b_1^T$, so (5) holds. ■

Notice that the assumption $s+w \neq 0$ is trivially satisfied since imperceptibility of the watermark implies $\|w\| << \|s\|$.

Let now ϕ_e be an extracted feature. The watermark detection is accomplished by the decision function

$$\delta(\phi_e) = \begin{cases} 1 & \textit{if } |\mathrm{sim}(s+w, D\phi_e)| \geq \varepsilon \\ 0 & \textit{otherwise} \end{cases} \tag{7}$$

where $0 \leq \varepsilon << 1$ is a suitable threshold and

$$\mathrm{sim}(s+w, D\phi_e) = \frac{(s+w)^T D\phi_e}{\|s+w\|\|D\phi_e\|} \tag{8}$$

Definitions (7) and (8) for the detector are motivated by the following result, which shows that the watermark is perfectly detected in the feature vector associated to the watermarked image.

Proposition 1: We have $\mathrm{sim}(s+w, D\phi_w) = 1$.

Proof: From definitions (6), (3), (1), (2) it follows that $D\phi_w = (AG^T + \mu BH^T)(Gs + Ht + \sigma_o + Gw) = A(s+w) + \mu Bt = (1+\mu^2)(s+w)$ by conditions (4) and (5) (notice that $G^TG = H^TH = I$, $G^TH = H^TG = 0$ since the columns of G and H are orthonormal bases of mutually orthogonal spaces). Hence from (8) we deduce

$$\mathrm{sim}(s+w, D\phi_w) = \frac{(1+\mu^2)(s+w)^T(s+w)}{(1+\mu^2)\|s+w\|^2} = 1$$

■

Notice that, in order to work, the detector needs only the matrix D and the vector $s+w$. Therefore, if we take (G, H, A, B) as a secret key and $(D, s+w)$ as a public key, we obtain an asymmetric watermarking scheme as in [4].

III. Security analysis

Security of the watermark refers to the inability by not authorized users to decode the embedded sequence. As discussed in [2], § 5, the crucial problem for asymmetric watermarking security is represented by the projection attack. As explained in [4], III.B., p. 786, a projection attack replaces the feature vector ϕ_w associated to the watermarked image with a feature vector $\tilde{\phi}$ satisfying

$$\|\tilde{\phi} - \phi_w\| = \min \|\phi - \phi_w\|^2 \tag{9}$$

under the constraint

$$\delta(\phi) = \mathrm{sim}(s + w, D\phi) = 0 \tag{10}$$

Hence, $\tilde{\phi}$ is the non-watermarked feature vector closest to ϕ_w. By definition (8), condition (10) says that $(s + w)^T D\phi = 0$, i.e., ϕ has to lie on the hyperplane through the origin of the feature space having normal vector $a = D^T(s + w)$. As a consequence, the feature vector $\tilde{\phi}$ satisfying condition (9) is the projection of ϕ_w onto this hyperplane, which is given by

$$\tilde{\phi} = \phi_w - \frac{a^T \phi_w}{\|a\|^2} a \tag{11}$$

Our main result is the following.

Theorem 1: For every choice of the watermark w, our scheme is secure under projection attack.

Proof: By (6) we have $a = D^T(s + w) = (GA^T + \mu HB^T)(s + w) = G(s + w) + Ht$ since $A^T(s + w) = A(s + w) = s + w$ by (4) and $\mu B^T(s + w) = t$ by (5). On the other hand, if we let $\psi_w = \phi_w - \sigma_o$, from (3), (1), (2) it follows that $\psi_w = \phi_o + Gw - \sigma_o = Gs + Ht + \sigma_o + Gw - \sigma_o = G(s + w) + Ht$. Hence we see that $a = \psi_w$ and from (11) we deduce

$$\tilde{\phi} = \phi_w - \frac{\psi_w^T \phi_w}{\|\psi_w\|^2} \psi_w = \phi_w - \psi_w = \sigma_o$$

by definition of ψ_w. Since $\sigma_o \in \mathcal{V}$ is the fragile part of the original feature vector, we conclude as in [4], III.B., p. 786, that the image reconstructed from $\tilde{\phi}$ has a high probability of being perceptually distorted. ■

We stress that the corresponding result in [4] implies that w is a multiple of s (see statement c) of the Theorem on p. 787). On the contrary, the security of our scheme does not depend on a specific watermark, thus making it suitable also for copy protection.

IV. Conclusion

We have presented an asymmetric watermarking scheme which improves a recent proposal [4] by allowing the insertion into the host image of an arbitrary, image independent, sequence of data. In particular, it has been checked that the method is robust against the most dangerous attack for asymmetric schemes, namely, the projection attack.

Future work will deal with related implementation issues and a full experimental assessment in a simulation environment. Further investigation will concern the optimal choice of matrix A with respect to detection performances.

References

[1] T. Kalker, ”Considerations on watermarking security,” *IEEE Fourth Workshop on Multimedia Signal Processing*, 2001, pp. 201–206.

[2] M. Barni, F. Bartolini, T. Furon, ”A general framework for robust watermarking security,” *Signal Processing*, vol. 83, pp. 2069–2084, 2003.

[3] M. F. Mansour, A. H. Tewfik, ”Secure detection of public watermarks with fractal decision boundaries,” *European Signal Processing Conference – EUSIPCO 2002*, vol. I, pp. 295-298, 2002.

[4] J. Tzeng, W.-L. Hwang, I-L. Chern, ”An asymmetric subspace watermarking method for copyright protection,” *IEEE Trans. Signal Processing*, vol. 53, pp. 784–792, Feb. 2005.

[5] J. Tzeng, W.-L. Hwang, I-L. Chern, ”Enhancing image watermarking methods with/without reference images by optimization on second order statistics,” *IEEE Trans. Image Processing*, vol. 11, pp. 771–782, Jul. 2002.