Problems and Challenges in Multimedia Networking and Content Protection

G.Voyatzis and I.Pitas
Department of Informatics
University of Thessaloniki, Thessaloniki, 54006 GREECE
fax: +3031-996304, email: {voyatzis,pitas}@zeus.csd.auth.gr

Abstract

The digital networked environment necessitates the development of protection techniques for multimedia product access and distribution. This paper refers to protection schemes for copyright and content originality of multimedia products through invisible watermarking. In particular, we present the basic protection schemes and fundamental concepts. The elementary partial algorithms, which can consist an overal watermarking system, are given and their basic characteristics are studied. Efficient protection demands from watermarks to have special features and obey conditions that should be satisfied strictly. Beside watermarking, the importance and the necessity of a proper product registration is stated.

1 Introduction

Transmission, manipulation and storage of images in digital format is rapidly becoming an everyday practice. Desktop publishing, digital libraries, databases and the World Wide Web are only some of the application areas that are strongly related to digital technology. The rapid evolution of digital technology makes the development of reliable and robust schemes for protecting multimedia products from piracy a matter of urgency. Piracy attacks include illegal access to transmitted data in networks, malicious content modification and retransmission of illegitimate copies [1, 2]. The impact of such attacks might be very large both in financial (profit losses by unauthorised access and use of data) and security terms. In the following, we clarify three important notions: key cryptography, authentication and copyright protection.

Public or Secrete key Cryptography
Data transmitted through network communication lines may be protected from unauthorised receivers by applying techniques based on cryptography [3]. Original data are encrypted by the providers using a private key. The users can decrypt the received data using an algorithm implemented either in hardware or in software. Necessary condition for successful decryption is the possession of the provider's private (secret) key or an associated public (or partally public) key.

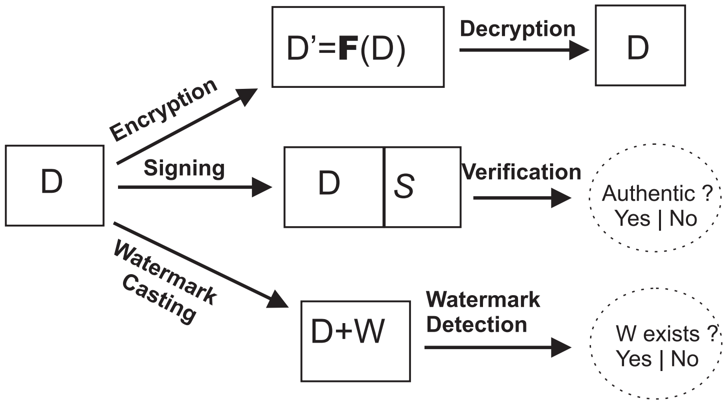

Figure 1: Schematic representation of data encryption, authenticity verification and watermarking.

Fast implementation of encryption-decryption algorithms is highly desirable. Furthermore, increase of data size due to encryption should remain within reasonable limits. The key size should be sufficient for preventing intentional decoding by trial and error procedures.

Authentication
Data content manipulation can be performed for various legal or illegal purposes (compression, noise removal, malicious data modification). The modified product is not authentic with respect to the original one. Therefore, users should be able to check the originality of the content of a digital product. Content verification can be performed by attaching *digital signatures* to the transmitted data. A digital signature is an encoded message that matches the content of a particular authentic digital product [3]. Authenticity verification procedures are based on public algorithms and public keys. Any "worth noting" modification performed in the product or in the signature data should cause verification failure. However, secure and efficient signatures for multimedia products (which are usually of large amount of data) are not feasible. Generally, the size of the signatures is proportional to the size of the signed data.

Copyright Protection
Reproduction of a digital product is easy and inexpensive. In a networked environment (like the Word Wide Web) retransmission of copies all over the world is possible. Copyright ownership can be violated by persons who illegally claim the product exploitation rights. A copyright protection technique, used by television channels, is the insertion of a visible logo in the digital image. However, such a logo can be easily removed or replaced and, subsequently, any evidence about the legal owner is lost. The problem of protecting the intellectual property of digital products has been treated in the last few years with the introduction of the notion of watermarks. Watermarks modify slightly the digital data to embed non-perceptible encoded copyright information. Watermarks may contribute to content verification too.

In this paper we discuss about watermarking schemes for protecting multimedia data. Fundamental concepts and algorithms are presented in order to protect effectively copyright and

Figure 2: The elementary distribution system of digital products. Piracy for copyright violation or malicious modifications and retransmissions is possible.

originality of contents. Finally, the necessity of product registration is remarked for accomplishing a reliable protection system.

2 Watermarking fundamental concepts

The manipulation of digital data and their delivery to the users can be carried out through web services. The user may have access to an interesting product either directly from the site of the original owner or a digital library or through an intermediary editor. A digital product may appear in its original form or as a part of a larger multimedia application. In this paper we consider an elementary digital product distribution system presented in figure 2. The user gets a digital product from an editor who is the copyright owner and, possibly, the original creator. Access and transmission take place in a global network environment. Pirates may harm the copyright owner by reproducing and retransmitted illegally digital products. Furthermore, pirates may proceed to product modification providing the users with non-authentic products. Therefore, both editor and user need protection:

- The copyright owner demands an efficient protection scheme that provides a browser for his/her copyright property in whole accessible domain.
- The users demand to receive authentic products when they follow a legal way for their purchase.

We should mention that users do not harm other users because they do not retransmit or expose (in public domains) products. Thus, pirates are met as unauthorised editors.

Watermarking aims to embed in the multimedia products invisible information carried by a signal called watermark. The watermarks are generated privately and, afterwards, should be detected by using private or public keys according to their use.

2.1 Watermarking for copyright protection

The watermark (called also copyright label or invisible stamp in this case) either carries specific information for the legal owner (e.g. a logo) or it is a random but unique signal for the particular owner. The protection scenario is the following:

- Each copyright owner possesses a unique long number (or a set of numbers) that constitutes the private "watermark key" K_{pr}.

- By using the private key and a *public* or *private* algorithm, the copyright owner alters the digital data being watermarked (format headers and other additional information is excluded).
- The copyright owner, using a detection algorithm, can check or decode the alterations, performed by himself, and uses such information as an indication of his/her legal ownership on the particular product.

A crucial point in the above protection scenario is a potential of the watermark by a pirate. In this case the watermarking protection scheme fails to provide protection.

2.2 Watermarking for authentication

As mentioned in the introduction, the solution to the problem of checking digital product authentication is approached by digital signatures. Authenticity is related to an original reference product and usually deals with the originality of contents, the name of original creator, the time of production, the copyright owner, etc. Watermarks mainly refers to the part of authenticity associated with the originality of contents (the terms data integrity, content verification or tamper proofing are also used). We draw up the following basic scheme:

- The original creator possesses a unique private key K_{pr}.
- The private key and the product are inserted in an algorithm and the data are altered in order to embed the authenticity information. The algorithm should provide also a public key K_{pub}.
- The receiver is in the position to check authenticity. He can use the public key K_{pub} and a public algorithm that provides a binary answer indicating authenticity or not.

Similarly to digital signatures, watermarks for content verification should be fragile. Pirates do not aim at watermark removal but at reproducing the watermark and creating false authenticity proofs.

3 Watermarking elementary procedures

3.1 A basic Watermark definition

Various definitions for digital watermark can be met in literature. Watermarks for digital images are handled as LSB manipulation [4], hidden mark codes [5, 6], invisible textures [7], secret constrain in transform domains [8, 9] etc. We can define as watermark a digital signal W, which is superimposed on digital products through an embedding process [10]. It is convenient to describe W as a signal that consists of binary or, more generally, ternary components:

$$W = \{w(\vec{k}); \mid w(\vec{k}) \in \{-1, 0, 1\}\ ,\ \vec{k} \in \hat{W}^d\} \quad (1)$$

where $\hat{W}^d$ denotes the watermark digital data space (grid) of dimension $d = 1, 2, 3$ for audio, still images and video respectively. Vector $\vec{k}$ stands for pointing the grid positions.

A watermarking protection scheme is composed by considering the following elementary procedures:

3.2 Watermark Generation

Let $\mathbf{W}$ be the set of possible watermark signals. According to the requirement for the existence of an associated key we consider the finite key space $\mathbf{K}$. If $\mathbf{X}$ denotes the set of still digital images, a watermark generation procedure should be defined by the following function :

$$\mathcal{G} \ : \mathbf{X} \times \mathbf{K} \rightarrow \mathbf{W} \quad , \quad W \ = \ \mathcal{G}(X, K) \tag{2}$$

where $K \in \mathbf{K}$ is the watermark key and $X \in \mathbf{X}$ is the product where the watermark will be embedded. For any particular product X and a given watermark signal W the key extraction should be impossible.

3.3 Watermark embedding

By considering the watermark signal $W = \{w(\mathbf{k})\}$, produced by $\mathcal{G}$, the embedding process is defined as a superposition of W onto the original product $X_o = \{x(\mathbf{k})\}$. We denote the embedding procedure by $\mathcal{E}$ and we define it as follows:

$$\mathcal{E} : \mathbf{X} \times \mathbf{W} \times I\!R \rightarrow \mathbf{X} \quad , \quad X_w \ = \ \mathcal{E}(X_o, W; l) \tag{3}$$

The real-valued parameter l is associated to the embedding watermark energy or, equivalently, to the watermark visibility. In practice, instead of a single parameter l, an embedding mask L is required for achieving satisfactory embedding. L is formed by taking under consideration the perceiving characteristics of the human visual/auditory system.

3.4 Web Searching

Illegal copies of digital products are looked for inside accessible and suspicious web domains. Therefore, watermarking should be combined by an automated web-crawling procedure, denoted by $\mathcal{S}$, which provides the watermark detection procedure with the products located in the particular domains:

$$X = \mathcal{S}(NetworkDomain) \ , \quad X \in \mathbf{X} \tag{4}$$

3.5 Watermark Detection

The detection algorithm is denoted by $\mathcal{D}$ and defined as follows:

$$\mathcal{D} \ : \ \mathbf{X} \times \mathbf{K} \rightarrow \{0, 1\} \tag{5}$$

$$\mathcal{D}(X, W) = \begin{cases} 1 & \text{if } W \text{ exists in } X \\ 0 & \text{otherwise} \end{cases}$$

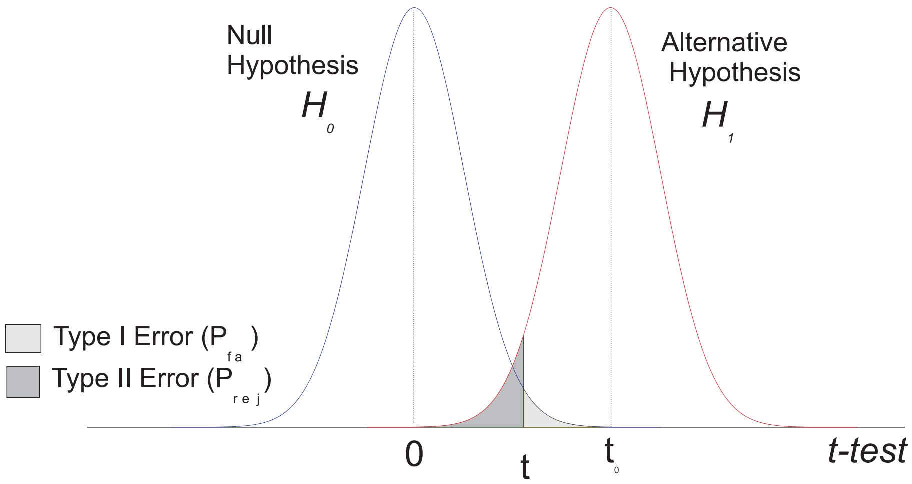

Figure 3: Watermark detection through *t-test* statistic. Alternative and Null hypotheses represents the probabilities for watermark existence and not existence respectively. Errors of type I and II are non-zero for any value of t, derived from the *t-test*.

Watermark detection is performed without resorting to the original products. The use of originals causes serious obstructions for achieving efficient computations and automation. We proceed to the watermark detection, by generating formerly the watermark using $\mathcal{G}$. $\mathcal{G}$ is based exclusively on the product X to be checked and the key K.

In practice, the realisation of $\mathcal{D}$ implies the following errors:

Type I error: Watermark is detected although it does not exist in the data (false positives).

Type II error : Watermark is not detected in the data although it exists (false negative).

The above errors occur with specified probabilities of false alarm (P_{fa}) and rejection (P_{rej}) respectively. Let $c = 1 - P_{fa}$ denotes the certainty of a positive detection, then:

$$c \geq c_{thres} \implies \text{watermark exists} \tag{6}$$

The parameter c_{thres} is the *certainty level* for detection and is chosen by the provider who applies detection. Hypothesis testing can be used for statistical certainty estimation and error manipulation [21]. Generally, when false positives become insignificant ($P_{fa} \to 0$) the probability to reject a watermark increases ($P_{rej} \to 1$) and vice versa. This situation is illustrated in figure 3 where t-test statistic is considered.

3.6 Product Search/matching in Library

It is natural and without significant cost, the providers or the original creators to store their products in a personal library denoted by $\mathcal{L}$. By given a product X, the provider should be able to apply a "matching procedure" $\hat{m}$ in order to check if X is included in his/her library $\mathcal{L}$:

$$\hat{m}(X, \mathcal{L}) = \begin{cases} 1 & \text{if } X \in \mathcal{L} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

A matching procedure, generally, is based on algorithms of great complexity. Therefore, it is not convenient and practically feasible to apply $\hat{m}$ to all products in the interested web domains. Procedure $\hat{m}$ is used for ascertaining ownership of a particular product when watermark detection is performed with low, but not negligible, certainty.

4 Main features of a watermarking scheme

A. Copyright Protection

Watermarks should possess the following features:

Perceptual Invisibility: The modifications caused by watermark embedding, should not degrade the product perceived quality. However, even hardly visible differences may become apparent when the original product is directly compared to the watermarked one. Therefore, such differences remain unnoticed by an observer because the original product is accessible only by the legal owner.

Complexity: Watermark signals should be characterised by great complexity. This is necessary in order to be able to produce an extensive set of sufficiently well distinguishable watermarks. An enormous set of watermarks prevents the recovery of a particular watermark by trial and error procedures. In the majority of cases the complexity of a watermark is directly related to the size of the product where it is applied.

Associated key: Watermarks should be associated with an identification number so called *watermark key*. The key is used to cast, detect and remove a watermark. Subsequently, the key should be *private* and characterise exclusively the legal owner. Any digital signal, extracted from a digital product, is assumed to be a *valid watermark* if and only if it is associated with a key via a well established algorithm. This condition prevents the creation of *counterfeit* watermarks discussed extensively by Craver et al [22].

Automated detection/search: Watermarks should combine easily with a search procedure that scans automatically any publicly accessible domain in the network.

Trustworthy detection: Watermarks should constitute a sufficient and trustworthy proof of ownership on a particular product. Detection false alarms should appear very rarely (hopefully never). A particular watermark is a credible evidence for proving copyright ownership when its demonstration in a digital image is followed with insignificant error probability. However, low certainty detection may be performed in order to reduce the rejection probability during web monitoring.

Statistical invisibility: Watermarks should not be recovered using statistical methods. For example the possession of a great number of digital products, watermarked with the same key, should not dispose the watermark by applying statistical methods. Therefore, watermarks should be product dependent.

Multiple Watermarking: We should be able to embed a sufficient number of different watermarks in the same image. Each watermark should be detectable by using the corresponding unique key. This feature seems necessary because we cannot prevent someone from watermarking an already watermarked product. It is also convenient in cases where the copyright property is transferred from one owner to another (a fingerprinting like process [2]). We mention that the legal image owner is the only one that can dispose a copy containing *only* his/her own watermark.

Robustness: A digital image can undergo a great deal of different modifications that deliberately (piracy attacks) or not (compression, filtering for noise removal, resizing) affect the embedded watermark. Obviously, a watermark that is to be used as a means of copyright protection should be detectable up to the point that the host product quality remains within acceptable limits and for any kind of modification e.g.

1. Lossy Compression up to a certain quality level that do not produces product degradations.
2. Filtering for removing noise, improving product quality etc. Possible filters for intentional watermark removal should be also accounted for.
3. Geometric distortions (e.g. scaling, rotation, cropping, image/frame reflection and line/column/frame extraction or insertion).
4. Changing presentation form (e.g. A/D, D/A conversion, printing and rescanning etc).
5. , Color correction, contrast enhancement, histogram equalization etc.

B. Content verification

Protection for originality of contents demands perceptual invisibility, watermark complexity, key validity, trustworthy detection and statistical invisibility as they analysed previously. However, fragile (not robust) watermarks are required in this case i.e. watermarks which should not be detected when modifications have been performed. However, watermark robustness should be desirable in some cases where modifications do not destroy authenticity, e.g.

1. *High quality compression*
2. *Cropping of non-interested parts*
3. Other *insignificant modifications* to incorporate the product in a multimedia environment.

Roughly speaking, authenticity is guaranteed by the original creator. Therefore, just his/her watermark is legal. Subsequently, multiple watermarking is not necessary.

5 Watermarking and registration for efficient protection

Watermarking techniques, developed up today, show various disadvantages that prevents the formation of a universal trustworthy protection system based exclusively on watermarking. Some important disadvantages are the following:

Instability in respect to key loss/theft. Watermarking is based on a constant private key. When a pirate find a private key then he/she acquires directly the possibility to remove the corresponding watermarks from all products. Subsequently, violation of copyright, for all products of the particular provider, becomes an easy task. Also, products with forged authenticity proofs can be distributed.

Effective robustness/fragility. Watermarks for copyright protection should be robust to all possible modifications. Also, content verification demands watermarks with proper fragility to various kind of processing. The above demands are not easily achieved.

Watermark efficiency against new processing techniques or attacks. Watermark robustness, security, and resistance to attacks is tested for the present processing techniques. Watermarking can not guarantee its efficiency in the future. New compression techniques or filters may be developed in the future that easily remove watermarks from already distributed products.

Insecure Public detection. Watermark detection by using public keys is essential for content verification. However, public detection creates serious difficulties for developing secure watermarking against piracy.

Because of the above mentioned problems, watermarking can not provide a reliable protection scheme just by it self. However, it can constitute a significant part of an overall protection scheme. We propose that product registration, as a complementary action to watermarking, is necessary.

Product registration to a trusted authority is a well established way for protecting intellectual property rights of various products, e.g. books, software packages etc. Registration information can be used to form indisputable proofs for original ownership and legal rights. A protection system based on product registration requires the following actions:

1. The provider is registered in a trusted authority, which provides him/her with a watermarking unique key.

2. Watermark casting is applied by using the registered key.

3. The watermarked product is registered to a trusted authority before distribution.

The registered copy is time-stamped and its originality and legal ownership is marked. Registration after watermarking may contribute for efficient protection in various ways. For example we note the following:

- The provider proceeds to an automated watermark searching through the network sites. When low certainty searching level is used, the reliability of a positive result is reinstated by searching the library $\mathcal{L}$. The demonstration of the registered copy in a court of law is the proof for copyright ownership.

- Content verification can be performed by the registration authority and a dedicated public server for this purpose. The user, who wants to verify the contents of a product, should access the particular server where the product has been registered. The server can verify the integrity of contents by performing private watermarking and, afterwards, replies the result to the user.

6 Conclusion

Protection of digital products, distributed through the global network environment, is a challenging and urgent research topic. Watermarking is a new technique that promises effective particular solutions to problems dealing with the issues of copyright and content verification.

Copyright protection can be succeeded by providing any provider with the capacity to search and locate efficiently in the network his/her products and check their legal use. This capability is given by the products that will be distributed. Afterwards, the detection of the created personal watermark can indicate and alarm on the suspicious products. Piratical attacks, which aim to reduce the capability to detect watermarks or to reduce their capacity to indicate legal ownership, should be encountered. Watermarks should be sufficiently robust to any product modification, which does not degrade the product quality, Also, resistance to intentional watermark removal and forgery should be provided. Watermark detection of high certainty is a strong indication of ownership. We demonstrated a set of basic features that should obey the watermarks in order to satisfy the above mentioned demands. However, there are still unsolved problems and obstructions for using watermarking as it is. Product registration, manipulated by third trust parties, can complete a watermarking system for efficient and reliable copyright protection.

Content verification can be achieved using watermarking too. Watermark existence should indicate originality of contents. Therefore, pirates, who maliciously tamper and retransmit products, aim to preserve the capacity of watermark detection. Watermarks, designed for content verification, should be fragile to any product modification that destroys authenticity. Watermark forgery should not be possible. Content verification, mainly is interested to users, who want to know if the products that they receive through network devices, are originals. Subsequently, the capability for public watermark detection seems necessary. However, public detection is technically difficult and causes serious obstructions for creating secure private watermarks, proper registration and watermark detection by a third trust party has been proposed.

Further research is necessary for developing watermarking schemes and techniques that satisfy, as much as possible, the basic demands. Then, the contribution of the watermarking technique in an overall protection scheme will become valuable and trustworthy.

References

[1] B. M. Macq and J. J. Quisquater. Cryptology for digital TV broadcasting. *Proceeding of the IEEE*, 83, pp. 944–957, 1995.

[2] J. J. Quisquater J. F. Delaigle, J. M. Boucqueau and B. Macq. Digital images protection techniques in a broadcast framework : An overview. In *Proceedings of ECMAST'96*, vol 2, pp. 711–727, Louvain-la-Neuve, Belgium, 1996.

[3] D. R. Stinson. *Cryptography, Theory and Practice.* CRC Press, New York, 1995.

[4] G. Voyatzis and I. Pitas. Chaotic mixing of digital images and applications to wateramrking. In *Proceedings of ECMAST'96*, vol 2, pp. 687–694, Louvain-la-Neuve, Belgium, 1996.

[5] D. L. Hecht. Embedded data clyph technology for hardcopy digital documents. In *Proceedings of SPIE*, vol 2171, 1995.

[6] J. Brassil, S. Low, N.Maxemchuk, and L. Ó Gorman. Electronic marking and identification techniques to discourage document copying. In *Proceedings of Infocom'94*, pp. 1278–1287, 1994.

[7] W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for data hiding. *IBM Systems Journal*, 35, pp. 313–335, 1996.

[8] E. Koch and J. Zhao. Towards robust and hidden image copyright labeling. In *Proceedings of 1995 IEEE Workshop on Nonlinear Signal and Image Processing*, pp. 452–455, N. Marmaras, Greece, 1995.

[9] A. G. Bors and I. Pitas. Image watermarking using dct domain constraints. In *Proceedings of ICIP'96*, volume III, pp. 231–234, Lausanne, Switzerland, 1996.

[10] I. Pitas. A method for signature casting on digital images. In *Proceedings of ICIP'96*, volume III, pp. 215–218, Lausanne, Switzerland, 1996.

[11] N. Nikolaidis and I. Pitas. Robust image watermarking in the spatial domain. *Signal Processing, sp.issue on Copyright Protection and Access control*, to appear in 1998.

[12] G. Voyatzis and I. Pitas. Digital image watermarking using mixing systems. *Computer & Graphics*, 22, no 3, 1998.

[13] A. Piva, M.Barni, and F. Bartolini. Dct-based watermark recovering without resorting to the uncorrupted original image. In *Proceedings of ICIP'97*, vol I, pp. 520–523, Atlanta, USA, 1997.

[14] D. Kundur and D. Hatzinakos. A robust digital image watermarking method using wavelet-based fusion. In *Proceedings of ICIP'97*, vol I, pp. 544–547, Atlanta, USA, 1997.

[15] X. G. Xia, C. G. Boncelet, and G. R. Arce. A multiresolution watermark for digital images. In *Proceedings of ICIP'97*, vol I, pp 548–551, Atlanta, USA, 1997.

[16] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6, no 12, pp. 1673–1687, 1997.

[17] M.D.Swanson, B.Zhu, A.H. Tewfik, and L. Boney. Robust audio watermarking using perceptual masking. *Signal Processing, sp.issue on Copyright Protection and Access control*, to appear in 1998.

[18] J. Ó Ruanaidh and T. Pun. Rotation, scale and translation invariant digital image watermarking. In *Proceedings of ICIP'97*, vol I, pp. 536–539, Atlanta, USA, 1997.

[19] M. Kutter, F. Jordan, and F. Bossen. Digital watermarking of color images using amplitude modulation. *Journal of Electronic Imaging*, 7, no 2, pp. 326–332, 1998.

[20] J. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Phase watermarking of digital images. In *Proceedings of ICIP'96*, vol. III, pp. 239–242, Lausanne, Switzerland, 1996.

[21] A. Papoulis. *Probability & Statistics.* Prentice Hall, 1991.

[22] S. Craver, N. Memon, B-L. Yeo, and M. Yeung. Resolving rightful ownerships with invisible watermarking techniques : Limitations, attacks and implications. *to appear in IEEE Journal of Selected Areas in Communications*, 1998.