to appear in *Workshop on Information Hiding, LNCS, Springer Verlag*

Secure Copyright Protection Techniques for Digital Images

Alexander Herrigel[a] · Joseph Ó Ruanaidh[b]
Holger Petersen[a] · Shelby Pereira[b] · Thierry Pun[b]

[a]r[3] security engineering ag
{herrigel,petersen}@r3.ch
P.O. Box
CH-8301 Glattzentrum

[b]University of Geneva
CUI - Vision Group
24, rue du Général-Dufour
CH-1211 Geneva

Abstract. This paper[2] presents a new approach for the secure and robust copyright protection of digital images. A system for generating digital watermarks and for trading watermarked images is described. The system is based on a new watermarking technique, which is *robust* against image transformation techniques such as *compression, rotation, translation, scaling* and *cropping*. It uses modulation of the magnitude components in Fourier space to embed a watermark and an accompanying template and, during watermark extraction, reads a template in the log polar transform of the frequency domain. The template is used for analyzing scaling and rotation suffered by the watermarked stego-image. The detection of the watermarks is also possible without any need for the original cover-image. In addition, the system applies asymmetric cryptographic protocols for different purposes, namely embedding/detecting the watermark and transferring watermarked data. The public key technique is applied for the construction of a one-way watermark embedding and the verification function to identify and prove the uniqueness of the watermark. Legal dispute resolution is supported for the multiple watermarking of a digital image without revealing the confidential keying information.

1 Introduction

The current rapid development and deployment of new IT technologies for the fast provision of commercial multimedia services has resulted in a strong demand for reliable and secure *copyright protection* techniques for multimedia data. Copyright protection of digital images is defined as the process of proving the intellectual property rights to a court of law against the unauthorized reproduction, processing, transformation or broadcasting of a digital image. Depending on the law in various countries, this process may be based on a prior registration

[1] All methods, procedures and schemes presented in this paper are based on the European patent application No. 97 810 708.4

[2] This work has been founded by the Swiss National Science Foundation under the SPP program (Grant. 5003-45334) and by the EC (ESPRIT Project No. 25530: Jedi-Fire).

of the copyright with a trusted third party. After successful registration, the copyright ownership is legally bound by a copyright notice, which is required to notify and prove copyright ownership.

Digital watermarking is a method for marking data sets, such as images, sound or video. A stego data set consists of the original data, the cover data set and a digital watermark that does not affect the data set's usability but that can be detected using dedicated analysis software or systems. Watermarking can, for example, be used for marking authorship or ownership of a data set.

Quite a number of different approaches [CKLS95,Caro95,TRSM93,TiSO95] [MaTa94,SmCo96,Schn95,CMYY97,ORDB96,ORDW96,DaSc96,ZhKo94,Zhao96] [DBQM96] to digital watermarking have been proposed but only some of them implemented in commercial products. Due to the very short time and minimal effort needed for copying and distributing digital multimedia data, protection against copyright infringements is an important issue for the copyright owner and should form an integral part of the exploitation process for Internet based trading services. Today, the Internet community has not identified or accepted adequate copyright protection techniques. This is in direct contrast to the provision of secure transaction protocols, such as SSL [FrKK96].

2 State-of-the-art

Digital watermarking can be seen as being fundamentally a problem in digital communications [CKLS95]. Early methods of encoding watermarks consisted of no more than incrementing an image component to encode a binary '1' and decrementing to encode a '0' [Caro95]. Tirkel et al. [TRSM93] and van Schyndel et al. [TiSO95] have applied the properties of $m-$sequences to produce oblivious watermarks resistant to filtering, cropping and reasonably robust to cryptographic attack. Matsui and Tanaka [MaTa94] have applied linear predictive coding for watermarking. Their approach to hide a watermark is to make the watermark resemble quantization noise. Tirkel and Osborne [TRSM93] were the first to note the applicability of spread spectrum techniques to digital image watermarking. Spread spectrum has several advantageous features. It offers cryptographic security (see [TRSM93]) and is capable of achieving error free transmission of the watermark at the limits given by the maximum channel capacity [SmCo96]. Fundamental information theoretic limits to reliable communication have been discussed by some authors (see [SmCo96]). The shorter is the payload of a watermark, the better are the chances of it being communicated reliably.

Spread spectrum is an example of a symmetric key cryptosystem [Schn95]. System security is based on proprietary knowledge of the keys (or pseudo random seeds) which are required to embed, extract or remove an image watermark. One provision in the use of a spread spectrum system is that it is important that the watermarking be non-invertible because only in this way can true ownership of the copyright material be resolved [CMYY97]. Ó Ruanaidh et al. [ORDB96] and Cox et al. [CKLS95] have developed perceptually adaptive transform domain methods for watermarking. In contrast to previous approaches the empha-

sis was on embedding the watermark in the most significant components of an image. The general approach used in these papers is to divide the image into blocks. Each block is mapped into the transform domain using either the Discrete Cosine Transform (DCT) [PeMi93], the Hadamard Transform [Cham85] or the Daubechies Wavelet Transform [PTVF92]. Information has been embedded using the DCT [ORDW96] or FFT magnitude and phase, wavelets (see refs. of [ORDW96]), Linear Predictive Coding [MaTa94] and fractals [DaSc96]. J.-F. Delaigle at al. [DBQM96] have applied signature labelling techniques for the copyright protection of digital images.

The industrial importance of digital copyright protection has resulted in a number of products, either based on specific watermark techniques or additional registration services. They include the PictureMarc system by Digimarc, SureSign (former FBI's Fingerprint) by HighWater Signum, IP2 system by Intellectual Protocols, the Argent system by Digital Information Commodities Exchange and the Tigermark system from NEC. Further some prototypes have been developed among which are the PixelTag system by the MIT Media Lab and the SysCop system from Zhao and Koch of the Frauenhofer-Institut für Graphische Datenverarbeitung [ZhKo94,Zhao96]

3 Overview

We envision the watermark system operating an open environment like the Internet with different interconnected computers. Users can be located anywhere and can sell or buy images. If legal dispute resolution for multiple watermarks is needed the Copyright Holder (H) sends copyright information and authentic image information to the Copyright Certificate Center (C). After having received a copyright certificate from C, the copyright holder can sell his digital images, for example, via an image shopping mall, to an image buyer (B). The Public Key Infrastructure (PKI) supports the distribution of authentic public keys between all parties which are needed for mutual authentication, non-repudiation and confidentiality. The communication channels between the parties are shown in figure 1.

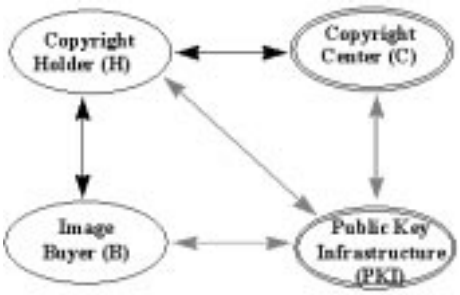

Fig. 1. Communication channels between identified parties

Our approach enables the secure generation and transmission of watermarked data using an asymmetric key pair like applied in public-key cryptography. The cover data set is watermarked, while the watermark is encoded using one or both of these keys. The resulting stego data set is then transmitted to a second party, while the same keys are used for establishing a secure transmission between the parties.

During the trading process, the involved parties use asymmetric key pairs and a key agreement protocol for establishing a symmetric key between them. The party creating the watermark can embed a *private*, a *detection* and a *public watermark* in the data set. The public watermark can be detected by third parties while the private and detection watermark can only be detected by the copyright holder.

After embedding the digital watermark into the image, the information describing it, such as the image identifier, a textual description and related information is transmitted authentically to a registration party that permanently stores a record relating to this stego data set and issues a copyright certificate which it stores and transmits to the copyright holder.

A template pattern is added to the Fourier transform of an image to be watermarked. For checking the watermark, the Fourier transform of the stego-image is calculated. From this Fourier transform, the log polar mapping transform is generated, which is then searched for the modulation pattern. Using the log polar transform of the Fourier transform has the advantage that scaling and rotation of the stego-image are expressed as translations. This allows an easy search for rotation and scaling using cross-correlation techniques. The magnitude components of the Fourier transform of each image block is modulated using the same pattern in each block. This method provides robustness against cropping of the stego-image because the magnitude spectrum is invariant to circular translations (cyclic shifts) and cropping leads to a circular translation of the watermark in each block.

4 Copyright Protection and Image Owner Authentication

Depending on the proof-level to be provided for copyright protection, our approach provides three increasing levels of reliability, namely: individual copyright protection, copyright protection with registered cryptographic keys and copyright protection with C on the basis of registered cryptographic keys. The present method is based on an image owner authentication technique, described below, which embeds and detects the Image Authentication Data (IAD) as the payload of a watermark. The applied image owner authentication technique is based on a perceptually adaptive spread spectrum technique. This technique provides a reliable means of embedding robust watermarks. Such a technique will be discussed in section 5. In addition, spread spectrum is a form of symmetric cryptosystem. In order to embed or extract a watermark, it is necessary to know the exact values of the seed used to produce pseudo random sequences used to encode a watermark. The seeds are considered to be cryptographic keys

for watermark generation and verification. System security is therefore based on proprietary knowledge of private keys, which provide in addition the necessary security parameters needed for a secure communication (mutual authentication, integrity, confidentiality, non-repudiation) in the trading process of digital images. Because spread spectrum signals are statistically independent (and therefore virtually orthogonal), the present method and apparatus encodes more than one watermark in an image at the same time, namely a private, a detection and a public watermark.

The *detection watermark* is embedded under a fixed random seed which allows H to efficiently search for his images on the Internet. The *public watermark* indicates that the image is copyright material and provide information on true ownership. At the same time there is a secure *private watermark* whose secrecy depends on the private key of H. Since the public key of H is registered, H can prove that he is the only person in the possession of the adequate private key and therefore the generator of the private watermark. The system also provides the secure registration (mutual authentication, integrity, non-repudiation) of watermark encoded images (data sets) at C. Derived data of the stego-image is registered at C. A signed digital copyright certificate is generated by C and transmitted to H. If an unauthorized third party has also encoded watermarks in the same image, conflicting claims in copyright disputes can be resolved, as only one of the two parties has a copyright certificate for the image containing only its watermark. The other party, who redistributed the original watermarked image, has only a certificate on the image where both watermarks are embedded and thus can be identified as the cheating party.

Watermark protection with registered cryptographic keys and C based copyright protection are based on a PKI. The PKI issues on request public key certificates such as X.509 certificates, containing the public key of the party, its distinguished name and a time stamp. Every certificate is signed with the PKI's private key and trust is built on the validity of the authentic copy of the PKI's public key (we assume that the public key of the PKI is accessible, authentically distributed and verifiable by every party).

4.1 Cryptographic mechanisms

The following cryptographic mechanisms are used in the description [MeOV96]:

- A probabilistic digital signature scheme $(\mathcal{G}, \mathcal{S}, \mathcal{V})$, with key generation algorithm $\mathcal{G}$, signature algorithm $\mathcal{S}$ and verification algorithm $\mathcal{V}$. The key generation algorithms $\mathcal{G}$ returns a key pair $(x_{\mathcal{Z}}, y_{\mathcal{Z}})$ for entity $\mathcal{Z}$. Signature generation for a message m is described as $\sigma := \mathcal{S}(x_{\mathcal{Z}}, m)$ and the verification of this signature by $\mathcal{V}(y_{\mathcal{Z}}, m, \sigma) \in \{true, false\}$.
- $AKAP(A, B, K_{AB}, x_A, y_A, x_B, y_B)$*:* Asymmetric key agreement protocol (e.g. [ISO 95]) with entity A's keypair (x_A, y_A), entity B's keypair (x_B, y_B), between the entities A and B. After the protocol, the two entities have agreed on a symmetric key K_{AB}.

- *OIAE(X, Y, CI, SI)*: Oblivious image owner authentication embedding algorithm with seed X, payload Y, cover-image CI and resulting stego-image SI.
- *OIAV(X, SI, Y)*: Oblivious image owner authentication detection algorithm with seed X, stego-image SI and the resulting payload Y.
- h_1, h_2: collision resistant hash functions with a hash value of appropriate length.

4.2 Individual Copyright Protection

During individual copyright protection only the copyright holder H with distinguished $name_H$ and asymmetric key pair (x_H, y_H) is involved. The following steps are applied:

1. H retrieves the cover-image CI, generates a unique image identifier $ID_I := h_1(name_H)||SN_I$, where SN_I is an image serial number and stores ID_I.
2. *Embedding of private watermark:*
   (a) H generates the stego-image SI_0 applying the transformation $OIAE(h_2(\sigma), ID_I, CI, SI_0)$, where CI denotes the cover-image, SI_0 denotes the resulting stego-image and $\sigma := \mathcal{S}(x_H, ID_I)$.
   (b) H stores σ together with CI in a protected database.
3. *Embedding of detection watermark:*
   H generates $OIAE(h_2(x_H), SN_I||SN_I, SI_0, SI^*)$. The key x_H could also be replaced by a random secret key k_H, which is used for every embedding and securely stored in a database.
4. *Embedding of public watermark:*
   (a) H generates a public $IAD_I :=$ "Copyright by" $||CDSig||$ "All Rights Reserved" applying $CDSig := \mathcal{S}^3(x_H, Initials||year)$.
   (b) H partitions IAD_I into blocks $BL_i, 1 \leq i \leq P$ of length 128 bits[4].
   (c) H generates the stego-image SI applying for every $i, 1 \leq i \leq P$, the transformation $OIAE(h_2(y_H||i||y_H), BL_i, CI_i, SI_i)$, where $CI_i := SI_{i-1}$ is the cover-image (stego-image from the previous iteration), $CI_1 = SI^*$ and SI_i is resulting stego-image after iteration i.

   The resulting stego-image is $SI := SI_P$.
5. H stores SI and might generate a signed copyright certificate $SigSI := \mathcal{S}(x_H, SI||TS)$ where TS is a time stamp.

Cryptographic properties: Besides the robustness of the watermarks against various image transformations, which is discussed in section 5.4, the embedding of the private watermark offers useful cryptographic properties:

[3] A signature scheme with message recovery is used here.
[4] This is the maximum length of a payload that can be robustly embedded by the spread spectrum technique.

- The seed for embedding the private watermark is *probabilistic*, as it depends on the output of a probabilistic signature scheme. Thus even if the private watermark in one image is detected this doesn't allow an attacker to find the private watermarks in other protected images.
- The seed for embedding is *image-dependent*. Thus an attacker who knows a valid image ID can't embed this using a different seed, as this wouldn't fit with the recovered ID from the signature with message used to generate the seed. The same is true for an attacker who knows a valid signed seed and wants to use this to embed his own image ID.
- The private watermark offers *non-repudiation*, as it can only be generated by the copyright holder, who knows the corresponding private key. This allows the proof of ownership to a judge.
- The signature σ remains *secret* until H has to prove his ownership to a court. Even in this case, he doesn't have to reveal a secret to the court, which would enable it to generate valid private watermarks instead of H afterwards.
- The payload is *very short*. One could think of embedding the signed image ID using a random seed instead of the described method. This leads to a longer message, which is not as robust as the chosen one, if we assume that the image ID consists of, for example, 12 bytes. The shortest known signature schemes with message recovery already produce a signature 20 bytes long [NyRu94,HoMP94], i.e. their output is 80% longer than our payload.

4.3 Trading of digital images

The copyright holder H and the image buyer B with distinguished $name_B$ are involved in the trading of digital images. Suppose (x_H, y_H) is the asymmetric key pair of H and (x_B, y_B) is the asymmetric key pair of B. Suppose H has an authentic copy of y_B and B has an authentic copy of y_H before they start any communication. The following steps are applied during the trading of digital images:

1. H and B execute $AKAP(H, B, K_{HB}, x_H, y_H, x_B, y_B)$ for the generation of a shared symmetric session key K_{HB}.
2. B generates the Trading Request Envelope $TRE :=< TD||SigTD >$, with transmission data $TD := \{ID_I||ExpTime||name_B||name_H\}$ and $SigTD := \mathcal{S}(x_B, h_2(TD))$. $ExpTime$ is the expiry time of the TRE, which avoids later replay of the same envelope. B transmits TRE to H.
3. H receives TRE and verifies TD, applying $\mathcal{V}(y_B, h_2(TD), SigTD) \stackrel{?}{=} true$ where $TD := ID_I||ExpTime||name_B||name_H$. If TD has been successfully verified the next step is executed. In any other case, the processing and communication between H and B is stopped.
4. If the verification was successful, H retrieves with ID_I the corresponding stego-image SI and generates the Trading reSponse Envelope $TSE :=< TD||SigTD >$, with $TD := K_{HB}[SI]||ExpTime||name_H||name_B$ and $SigTD := \mathcal{S}(x_H, h_2(TD))$. H transmits TSE to B.

5. B receives TSE and verifies TD by applying $\mathcal{V}(y_H, h_2(TD), SigTD) \stackrel{?}{=} true$, where $TD := K_{HB}[SI]||ExpTime||name_H||name_B$. If the verification is true, then TD has been successfully verified.
   B then deciphers $K_{HB}[SI]$ and checks IAD applying for every $i, 1 \leq i \leq P$, the following transformation: $OIAV(h_2(y_H||h_2(i)||y_H), SI, PL_i)$, where SI_i denotes the stego-image and PL_i the detected payload of the i-th public watermark. (If P is not known, the procedure is iterated until no more public watermarks can be detected).
   IAD_I is then generated by concatenating PL_i, i.e. $IAD_I := PL_1||\ldots||PL_N, 1 \leq i \leq P$. IAD_I should be of the format "Copyright by" $||CDSig||$ "All Rights Reserved". The message $Initials||year$ is recovered from $CDSig$ and the signature is verified applying $\mathcal{V}(y_H, Initials||year, CDSig) \stackrel{?}{=} true$. If the verification is correct, B has verified H as the copyright holder.

Remark
In the case of a legal copyright dispute, H can retrieve IAD_I and construct the corresponding unique image ID_I. Since the generation of the same asymmetric key pair by two distinguished entities is very unlikely, the construction of the unique image ID_I provides a high level of proof against copyright infringement. In the case of watermark protection with registered keys, the generation of the same asymmetric key pair by two distinguished entities can be prevented.

4.4 Copyright Protection with Registered Keys

Copyright protection with registered cryptographic keys needs three parties, namely H with $name_H$, B with $name_B$ and the PKI with $name_I$. Suppose (x_H, y_H), (x_B, y_B), (x_I, y_I) are the unique key pairs of H, B and I respectively. Suppose, H has an authentic and actual copy of $Cert_B$ which signature was verified with the authentic copy of y_I and B has an authentic and actual copy of $Cert_H$ which signature was verified with the authentic copy of y_I. Then the same steps as for the individual watermark protection are applied.
Remark:
Since the generated asymmetric key pairs are unique, H can be uniquely identified if no additional watermarks by unauthorized persons have been encoded into a given SI. The C based watermark protection described below provides the necessary countermeasures to prevent this threat.

4.5 Copyright Protection with a Copyright Center

The system for copyright protection with Copyright Center has four participants, namely H with $name_H$, B with $name_B$, the PKI with $name_I$ and C with $name_C$. Suppose (x_H, y_H), (x_B, y_B), (x_I, y_I) and (x_C, y_C) are the unique key pairs of H, B, I and C, respectively. H has an authentic copy of $Cert_B$ and $Cert_C$ whose signatures were verified with the authentic copy of y_I, B has an authentic copy

of $Cert_H$ and $Cert_C$ whose signatures were verified with the authentic copy of y_I and C has an authentic copy of $Cert_H$ and $Cert_B$ whose signatures were verified with the authentic copy of y_I. The following steps are applied:

1. Steps 1-4 of the individual watermark protection protocol are applied, where the unique image ID_I is computed as $ID_I := h_1(name_H||name_C)||SN_I$.
2. H generates a thumbnail $Thumb$ from SI and stores it together with the stego-image SI.
3. H and C execute the following steps for the secure registration and the generation of copyright certificates:
   (a) H generates the Copyright Request Data $CRD := \{h_2(SI)||h_2(Thumb)||ID_I\}$ and the Copyright Request Envelope $CRE :=< TD||SigTD >$, with $TD := \{CRD||ExpTime||name_H||name_C\}$ and $SigTD := \mathcal{S}(x_H, h_2(TD))$. H transmits CRE to C.
   (b) C receives CRE and verifies it. For this, C requests the certificate $Cert_H$ that belongs to $name_H$ in CRE and obtains the authentic public key y_H. C checks the signature on TD by applying $\mathcal{V}(y_H, h_2(TD), SigTD) \stackrel{?}{=} true$, where $TD = \{CRD||ExpTime||name_H||name_C\}$. Then C checks the semantical correctness of TD. If all verifications are passed, then CRE has been successfully verified and the next step is executed.
   (c) C generates the digital *Copyright Certificate* by executing $SigCCD := \mathcal{S}(x_C, CCD)$, with $CCD := \{name_C||name_H||SN||ID_I||h_2(SI)||h_2(Thumb)||UCCN\}$ and $UCCN :=$ "Copyright" $||year||name_H||$ "All rights reserved". C stores the copyright certificate $CC := CCD||SigCCD$ together with $h_2(Thumb)$ in its database, generates the Copyright Certificate Envelope $CCE :=< TD||SigTD >$, with $TD := \{CC||ExpTime||name_C||name_H\}$ and $SigTD := \mathcal{S}(x_C, TD)$ and transmits it to H.
   (d) H receives CCE and verifies it by requesting the certificate $Cert_C$ that belongs to $name_C$ in CCE. H obtains the authentic public key y_C which he uses to check $SigTD$ on TD by applying $\mathcal{V}(y_C, h_2(TD), SigTD) \stackrel{?}{=} true$, where $TD := \{CC||ExpTime||name_C||name_H\}$. H checks the semantic correctness of TD and if all verifications are correct then H stores CC in its database.
4. H and B might execute the image trading protocol described in section 4.3

Remark
B may check the copyright certificate requesting C to transfer an authentic copy of the copyright certificate for a given image identifier ID_I. Except the data transfer, the applied protocol is similar to the one described above. If B would like to transfer a specific copyright of a CI to another legal party, he may initiate a copyright revocation request with C. The different phases of this request are analogous to the copyright request.

5 Oblivious Image Owner Authentication

The watermarking technique comprises the following components:

1. An error-control coding technique for the payload to be transmitted in the watermark.
2. An encoding technique to encode the resulting message.
3. A reliable method for embedding the encoded message in the image without introducing visible artefacts.

Components 1 and 2 apply to embed watermarks in any type of data while component 3 is specific to the embedding of watermarks in images.

5.1 Error control coding

Error control coding is applied to the message prior to encoding. Symbol based Reed Solomon (RS) codes are applied for this purpose. The advantages are the following:

- RS codes correct symbol errors rather than bit errors and
- RS codes can correct erasures as well as errors.

Erasures can be factored out of the key equation, which means that "erased" symbols can be ignored. They do not play any role in the error control mechanism. In a sense, an erasure is useless redundancy. Being able to discard erased symbols has two advantages:

- If the posterior probability of a received symbol is low, it may be ignored.
- RS codes only come in standard sizes. For example a 255 x 8 bit code is common. Most commonly used RS error control codes appear to be too large to be used in watermarking. However, it is possible to make almost any RS code fit a watermarking application by judiciously selecting symbols as being erased (because they were never embedded in the image in the first place).

5.2 Encoding the message

During encoding, the message to be transmitted in the watermark is transformed into a form suitable for use in the modulation of image components. At the same time, it is encrypted using a suitable key.

One can easily combine spread spectrum based watermarking with the cryptographic key distribution techniques described earlier. The encoding procedure has access to the cryptographic keys x_H and y_H (or their hash values), which are used to generate the seeds for the pseudo-random sequences as described below. Knowledge of the corresponding key is required for recovering the message from the watermark.

A watermark can be embedded or extracted by the key owner using the same key. From the point of view of embedding watermarks in images given the cryptographic keys the sequences themselves can be generated. A good spread spectrum sequence is one which combines desirable statistical properties such as uniformly low cross correlation with cryptographic security.

Suppose we are given a message M (for example, that was provided with error coding). The message has the binary form $b_1 b_2 \ldots b_L$, where b_i are its bits. This can be written in the form of a set of symbols $s_1 s_2 \ldots s_M$ – most generally by a change in a number base from 2 to B. The next stage is to encode each symbol s_i in the form of a pseudo random vector of length N, wherein each element of this vector either takes the value 0 or 1. N is, for example, in the order of 1000 to 10000 (typically in the order of 10% of the total number of image coefficients (Fourier components) that can, theoretically, be modulated).

To encode the first symbol a pseudo random sequence v of length $N + B - 1$ is generated. To encode a symbol of values where $0 < s < B$ the elements $v_s, v_{s+1} \ldots v_{s+N-1}$ are extracted as a vector r_1 of length N. For the next symbol another independent pseudo random sequence is generated and the symbol encoded as a random vector r_2. Each successive symbol is encoded in the same way. Alternatively, one can use any the N cyclic shifts of the pseudo random sequence v. Note that even if the same symbol occurs in different positions in the sequence, no collision is possible because the random sequences used to encode them are different — in fact they are statistically independent. Finally the entire sequence of symbols is encoded as the summation: $m = \sum_{i=1..M} r_i$

The pseudo-random vector m has N elements, each varying between 0 and M. In a next step, the elements of m are offset to make their mean zero. When decoding the watermark, a vector m' (read-out message) is derived from the stego-image. In oblivious watermarking, m' corresponds to the modulated Fourier coefficients. Due to distortions suffered by the stego-image due to image processing, in general m' will not be equal to but will be "statistically similar" to m. To decode s from m', the elements of m' are first offset to make their mean zero. Then, starting from the (known) seed, the first random sequence v of length $N + B - 1$ is generated and the correlation of v with m' is calculated. The peak of the correlation indicates the offset s_1 in the random sequence that was used for generating r_1. Then, the next random sequence v is generated and cross-correlated with m' to retrieve s_2, etc. Reliable communications of the system are best accommodated by using $m-$sequences that possess minimum cross correlation with each other. This is the same as maximizing the Euclidean distance between vectors $v_1, v_2, v_3 \ldots$. If M is sufficiently large, the statistical distribution of the message m should approach a Gaussian distribution (according to the Central Limit Theorem). A Gaussian distributed watermark has the advantage that it is somewhat more difficult to detect. The variance increases with order $M^{1/2}$; in other words, the expected peak excursion of the sequence is only order $M^{1/2}$.

5.3 Oblivious Image Authentication Algorithm

In this section, we describe how to embed the encoded message m in the image in the form of a watermark. The method is designed for robustness to operations generally applied to images such as translation, cropping, rotating and scaling. (The method is not designed for other types of data such as sound or text.) In order to achieve robustness against circular translation, each image block is first

subjected to a Fourier transform and the same watermark is embedded in each block so the watermark tiles the entire image. Then, message m modulates the Fourier components. In addition to this, a template is embedded in the image, which can be used for detecting a rotation and scaling of the image when reading the watermark. Given a cover image the steps for embedding a watermark in the image are as follows:

1. If the image is a colour image, then compute the luminance component (for example, by simply replacing each pixel by $g/2 + r/3 + b/6$, where g, r and b are its green, red and blue components) and use these values for the following calculations.
2. Divide the image into adjacent blocks of size 128 $\times$128 pixels.
3. Map the image luminance levels (or grey levels for a black and white image) to a perceptually "flat" domain by replacing them with their logarithm. The logarithm is a good choice because it corresponds to the Weber-Fechner law which describes the response of the human visual system to changes of luminance. This step ensures that the intensity of the watermark is diminished in the darker regions of the image where it would otherwise be visible.
4. Compute the FFT (Fast Fourier Transform) of each block. From the real and imaginary components obtained in this way, calculate the corresponding magnitude and phase components. The magnitude components are translation invariant and will therefore be used in the following modulation steps. (However, it is possible to derive translation invariants from the phase spectrum as well, which could also be modulated).
5. Select the magnitude components to be modulated. To encode a message m of length N, a total number of N components are modulated. In non-oblivious watermarking, any components can be modulated. For oblivious watermarking, because of the interference of the cover-image with the watermark, the largest magnitude components are avoided and only a band of frequencies are used. These mid band frequencies are chosen because they generally give a good compromise between robustness and visibility of the watermark. There are two methods for selecting the components to be modulated:
   - The selection of the components to be modulated does not depend on the given image. Rather, the same components are selected for every image. The author as well as the reader of the watermark know the positions of the components to be selected in advance.
   - The largest components (inside the allowable frequency range) are used for modulation using a perceptually adaptive approach.
6. Add a template by a second modulation of the magnitude components. This is described in more detail below.
7. Compute the inverse FFT using the phase components and the modulated magnitude components.
8. Compute the inverse of the perceptual mapping function of step 3. For Weber-Fechner law mapping, the inverse function is an exponential.
9. Replace each watermarked block in the image to obtain the stego-image.

10. If the image is a colour image, then re-scale the red, green and blue components by the relative change in luminance introduced by embedding a watermark. Typically, the red, green and blue pixels occupy a byte each in program memory. If overflow or underflow occurs then the pixel is set to the upper bound 255 or lower bound 0 respectively.

When selecting the components to be modulated, care must be taken to preserve the symmetry imposed on the Fourier components $F(k_1, k_2)$ by the fact that the image block is real valued: $F(k_1, k_2) = F^*(N_1 - k_1, N_2 - k_2)$, where N_1, N_2 designate the size of the image block. Once the magnitude components $(M_1, ...M_N)$ to be modulated are chosen, the corresponding value m_i of message m is added to or subtracted from the corresponding selected magnitude component M_i. Addition is applied, if the difference between the corresponding phase component P_i and a "reference phase" is between 0 and π. Subtraction is applied if the difference is between π and 2π. Note that the reference phase for each watermark component should be the same for each block. This provides robustness against translation and cropping (see below). Before adding/subtracting the values m_i to/from M_i, the vector m can be scaled to adjust the magnitude of its elements to those of the components M_i. Generally, the elements m_i should be in the same order of magnitude as the components M_i. The depth of modulation or amplitude of the embedded signal should depend on the objective measure of the perceptual significance. The lower the perceptual significance, the higher should be the amplitude of the watermark. However, for simplicity, the amplitude for all components is usually kept constant.

Template: As mentioned above, a template is added to the image in step 6. The following steps have to be executed:

1. Apply a log-polar map to the magnitude components, i.e. transform them into a polar coordinate system $(\Theta, log - r)$ with an angular and a logarithmic radius axis respectively. In this representation, a scaling of the image leads to an offset of the components along the $log - r$ axis. A rotation of the image leads to an offset along the Θ axis. Preferably, low pass filtering is used for interpolating the frequency space components during this mapping. The magnitude components belonging to very low or high frequencies are not mapped. The following modulation is only applied to components in midband frequency range.
2. Select the magnitude components in the log-polar coordinate system to be modulated. Typically, as few as 10 components can be modulated. The pattern T formed by the selected components in log polar space should be such that its auto-correlation under translation is weak. There should be as little ambiguity as possible in matching the shifted stego-image with the template. For this purpose, the indices of the selected components should be co-prime or at least be derived from a two-dimensional random sequence.
3. Map the modulated points by a change of coordinates back into frequency space. The pattern T formed by the selected components in log polar space

is predefined and known to the reader of the watermark. It must be noted that the calculation of the log-polar transform is not required for embedding the template. The pattern T of the components to be modulated in log-polar space can be mapped back to the corresponding components in frequency space.

As will be explained below, the template is not required for non-oblivious watermarking (since the original image can be used as the "template, in other words, the stego-image can be registered with the original image). Also if the image is a colour image then the luminance image is used during the following operations:

1. Divide the image into adjacent blocks of size 128×128.
2. Map the image luminance levels (or gray levels) to the perceptually "flat" domain by replacing them with their logarithm.
3. For each block compute the FFT.
4. Determine the rotation and scaling that the image suffered by finding the template in log-polar space and compensate for the rotation and scale.
5. Read the modulated components to generate message m'.
6. Once that the message m' is recovered, it is demodulated and error corrected using the methods described earlier.

Finding the template
The steps for finding the template are as follows:

1. Apply a log-polar mapping to the magnitude components of the Fourier transform. The magnitude components in the very low or high frequency range are not mapped. The log-polar mapping is only applied to components in midband frequency range.
2. For oblivious watermarking, calculate the normalized cross correlation of the components in log-polar space with the template pattern T that was used for generating the template in step 6 and find the point of best correlation. If the image has neither been rotated or scaled, this point is at origin. Scaling leads to a corresponding offset along the $log - r$ axis, rotation to a corresponding offset along the Θ axis.
3. For non-oblivious watermarking, the log polar transform of the Fourier components of the cover-image can be used instead of template pattern T for retrieving scaling and rotation. The cross correlation can be calculated efficiently using conventional Fourier techniques.

5.4 Properties of the watermark

In the following, some of the properties of the watermark generated using the steps described above are discussed.

Robustness to cropping: One feature of translation invariants developed using the Fourier transform is that they are invariant to circular translations (or cyclic shifts). This is used to construct watermarks that are invariant to cropping. As mentioned above, the image is split into blocks and the watermark is applied to each block. In other words, the same modulation pattern is applied to the Fourier components of each block, wherein the modulation pattern is given by the corresponding encoded messages m. Suppose that the watermark in a standard size block will be of the form: $T = [AB; CD]$ where the submatrices A, B, C and D are of arbitrary size. A circular translation of such a watermark is of the form: $S = [DC; BA]$. The original stego-image is tiled with watermarks in the pattern $[TTTT; TTTT; TTTT]$. A little thought demonstrates that a cropped section of the matrix will carry a watermark in the form $[SSSS; SSSS; SSSS]$. When reading the watermark of the cropped image, each block carries the watermark S. Since S is a circular transform of T, it can be read directly in the Fourier domain using the steps outlined above. Note, however, that the cover-image is not tiled, only the watermark is. Therefore, while cropping merely induces a circular translation of the watermark in each block, the change of image in each block is not a circular translation.

The optimum size of block depends on a number of different factors. A size that is a power of two is useful because the FFT can be used. The block size also must be small enough to withstand cropping but large enough to comfortably contain a watermark. Heuristically, the best compromise for block size is 128.

Robustness to scaling and rotation: As mentioned earlier, reading the template in log-polar space allows to detect and measure any scaling and/or rotation that was applied to the image. This information can then be used for reading the watermark. Since the reader knows the pattern that was used for modulating the magnitude components, he can identify the modulated components in the scaled and rotated image and extract the message m'. Note that the apparatus does not explicitly use a rotation and scale invariant watermark but instead searches the parameter space of rotations and scales. Since searching the space of rotation and scales in the frequency or space domain is quite complicated, the log-polar map is used to map these parameters to Cartesian coordinates where searching can be carried out using efficient correlation techniques.

Robustness to translations: By translation, we mean zero padding of the image such as would occur if an image were placed on a scanner and scanned. In this case the effect on the watermark blocks may be understood in terms of simple signal and image processing theory. Effectively, zero padding is a multiplication by a rectangular window function in the spatial domain. In the frequency domain this approximates to a convolution with a cardinal sine function. Generally, this blurring of frequency space is not severe. If more than about one third of an watermark block is present the watermark can still be decoded.

Robustness to compression: The watermark is also resistant to compression. It is well known that transform based image compression techniques favour low frequencies in the sense that low frequency content in the original image is better preserved in the compressed image. For this reason, it would seem that a low frequency watermark would be better. However, as mentioned earlier in oblivious watermarking it is necessary to avoid low frequencies for embedding information because the image interferes with the watermark at those frequencies. Fortunately, a compromise is possible: judicious selection of a band of frequencies leads to a watermark that is both oblivious and is sufficiently resistant to lossy image compression. One helpful factor is that there are relatively few low frequency components in which to embed a spread spectrum signal. Using midband frequencies actually improves the robustness of the mark because of the increased redundancy of the encoding of the payload.

Figure 2 shows an image that was quite strongly watermarked using the techniques described in this paper. Figure 3 shows this image after JPEG compression was applied at 15% quality factor where it is found that that the storage required is less than 1% of that needed for the original image. Not surprisingly, the quality is very low and the image is of little commercial use. A 104 bit watermark "The watermark" (in ASCII code) can be recovered from the JPEG compressed image.

Fig. 2. A watermarked image of Lena

Fig. 3. A watermarked image of Lena compressed at 15% quality factor. The compression ratio is 100:1

Robustness to other attacks: Specialised watermark removal algorithms have been proposed in the literature. StirMark [Kuhn97] uses both contrast based attacks and geometric attacks. The geometric attacks are not directly addressed using the method proposed in this paper. However, we can give one good example of the robustness of the spread spectrum watermark described in this paper to a contrast based attack. The results are as follows:

- Operation: stirmark -i0 -o0 -d128 :
  - decoded watermark: "The watermark",
  - agreement with original = 100.00 percent.

  Operation: stirmark -i0 -o0 -d256 :
  - decoded watermark: "The wAtereak",
  - agreement with original = 96.15 percent.

The stirmark distorted image is shown in figure 4. It has obviously been very severely distorted. It was even surprising to the authors that this distortion only destroyed 4% of the bits of the watermark. Other attacks such as UnZign [UnZi97] have no effect on the mark.

Robustness of the template pattern: The template pattern is essential to determine the rotation and scaling of the image. Thus its robustness is highly

Fig. 4. The watermarked image of Lena after being attacked using "StirMark"

important. There is a good case for arguing that the template is somewhat more robust than the watermark. The reason for this is, that the template actually has to carry less information than the watermark. For example, suppose it is only possible to recover the watermark if the rotation angle and scaling factor recovered using the template are both 99.9% accurate. To give 99.9% accuracy one needs $\log_2(1000) \approx 9$ bit. Specifying both the angle and the scaling factor therefore needs around 18 bit which is considerably less than the amount of information contained in a typical watermark (about 128 bit).

Redundancy: The watermark is embedded in blocks of a fixed size with exactly the same watermark embedded in each block. This means that the watermark can be recovered from a single block only. However the chance of extracting the watermark correctly increases if one has more than one block. Therefore, the reliability of watermark extraction increases with the size of image, as one might expect.

6 Conclusions and future work

We have presented a new approach for the copyright protection of digital images. This approach is based on asymmetric cryptographic protocols and techniques. In contrast to any other scheme, the combination of a spread spectrum based

technique in conjunction with an asymmetric cryptographic technique allows the construction of a one-way function, since only H is able to verify the private watermark. In addition, he may prove that he has the adequate key by verifying the signature of the seed for the payload data. Even if the different phases of the approach are known in the public, the security of our approach is not compromised. Compared to other approaches, the following new properties have been identified:

1. Different security services for the communication, such as *mutual authentication, integrity, confidentiality* and *non-repudiation* are supported along with the protection against copyright infringement by the system with one asymmetric cryptographic key pair.
2. The present technique enables a strong binding relation between the image ID, the image and H if H registers his copyright at C. If an image is watermarked later by an unauthorized person, the time stamp in the copyright certificates resolves the copyright ownership.
3. H does not have to reveal his private cryptographic key if ownership verification has to be applied by a different legal party.
4. The present technique supports the *transferral of copyrights*. If a copyright is transferred to another legal party, corresponding copyright revocation certificates may be generated.
5. Digital signatures are used for the security of the communication between different parties and for the authenticity of the data embedded in a public watermark of an image. No signature labelling techniques of the complete image are applied by the system.
6. Circular translation invariants are used as a means of constructing digital watermarks that are invariant to cropping.
7. In contrast to some known techniques, the system does not require a database of watermarks since only the keys are required to embed or extract a watermark.
8. Information is retrieved from the log polar domain of the Fourier transform. Frequency components are modulated which are oblivious to the cover-image but which also have the property that they form an unambiguous non-repeated pattern in log-polar space. They are used for determining the degree of rotation and scaling suffered by a stego-image in the absence of the cover-image. Co-prime frequencies are useful for generating such a pattern or template. Uniform random sampling of log polar space is another method that can be applied.

To demonstrate the feasibility of the approach, a Java/C++ based copyright protection and authentication environment for digital images has been implemented. An example of this copyright protection environment in action is shown in figure 5. The PKI, H, C and B application processes all implement a Graphical User Interface and a server, supporting both console users and other requests through a socket interface.

The approach presented for the copyright protection of digital images can also be extended to other data such as video, audio and binary data. We are

Fig. 5. The watermarked image of Lena being registered for copyright protection using a Java console

actually investigating new spread spectrum techniques for the watermarking of audio and binary data.

References

[CKLS95] C. Cox, J. Killian, T. Leighton and T. Shamoon, "Secure spread spectrum communication for multimedia", Technical report, N.E.C. Research Institute, 1995.

[Caro95] G. Caronni "Assuring Ownership Rights for Digital Images" in H. H. Brueggemann and W. Gerhardt-Haeckl, editors, Reliable IT Systems VIS '95, Vieweg, Germany, 1995, pp. 251-264.

[Cham85] W. G. Chambers, "Basics of Communications and Coding", Oxford Science Publications. Clarendon Press Oxford, 1985.

[CMYY97] S. Craver, N. Memon, B. Yeo and M. Yeung, "Can invisible marks resolve rightful ownerships ?", IS&T/SPIE Electronic Imaging '97 : "Storage and Retrieval of Image and Video Databases", 1997.

[DaSc96] P. Davern and M. Scott, "Fractal based image steganography", Proc. International Workshop in Information Hiding, LNCS, Springer, 1996, pp. 279 – 294.

[DBQM96] J.-F. Delaigle, J.-M. Boucqueau, J.-J. Quisquater and B. Macq, "Digital Images protection techniques in a broadcast framework: An overview", Laboratoire de Télécommunications et de Télédéction, Université Catholique de Louvain, 1996.

[FrKK96] A. Freier, P. Karlton and P. Kocher, "SSL Version 3.0", Netscape Communications, Version 3.0, November 1996.

[HoMP94] P. Horster, M. Michels, H. Petersen, "'Meta signature schemes giving message recovery based on the discrete logarithm problem"', Proc. 2nd Int. Workshop on IT-Security, September, 1994, pp. 82 – 92.

[ISO 95] ISO/ IEC 11770-3, "Information technology-Security techniques-Key management, Part 3: Mechanisms using asymmetric techniques", 1995.

[Kuhn97] M.G. Kuhn, "StirMark", http://www.cl.cam.ac.uk/~fapp2/watermarking/image_watermarking/stirmark/, November 1997.

[MaTa94] K. Matsui and K. Tanaka, "Video-Steganography : How to secretly embed a signature in a picture", IMA Intellectual Property Project Proceedings, January 1994, pp. 187 – 206.

[MeOV96] A.J. Menezes, P.C. Van Oorschot and S.A. Vanstone, "Handbook of Applied Cryptography", CRC Press, 1996.

[NyRu94] K.Nyberg, R.Rueppel, "Message recovery for signature schemes based on the discrete logarithm problem", LNCS 950, Advances in Cryptology: Proc. Eurocrypt '94, Springer, (1994), pp. 182 – 193.

[ORDB96] J. K. Ó Ruanaidh, W. J. Dowling and F. M. Boland, "Phase watermarking of images", IEEE International Conference on Image Processing, September 1996.

[ORDW96] J. J. K. Ó Ruanaidh, W. J. Dowling and F. M. Boland, "Watermarking digital images for copyright protection", IEEE Proceedings on Vision, Image and Signal Processing, Vol. 143, No. 4, August 1996, pp. 250 – 256.

[PeMi93] W. B. Pennebaker and J. L. Mitchell, "JPEG Still Image Compression Standard", Van Nostrand Reinhold, New York, 1993.

[PTVF92] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, "Numerical Recipes in C", Cambridge University Press, second edition, 1992.

[SmCo96] J. Smith and B. Comiskey, "Modulation and information hiding in images", Proc. Workshop in Information Hiding, LNCS 1173, Springer, 1996, pp. 207 – 226.

[Schn95] B. Schneier, "Applied Cryptography", Wiley, 2nd edition, 1995.

[TRSM93] A. Z. Tirkel, G. A. Rankin, R. G. van Schyndel, W. J. Ho, N. R. A. Mee and C. F. Osborne, "Electronic watermark", Dicta-93, December 1993, pp. 666 – 672.

[TiSO95] A. Z. Tirkel, R. G. van Schyndel and C. F. Osborne, "a two-dimensional digital watermark", Proc. ACCV'95, December 1995, pp. 378 - 383.

[UnZi97] "UnZign", http://www.altern.org/watermark/, 1997.

[ZhKo94] J. Zhao and E. Koch, "Embedding robust labels into images for copyright protection", Technical report, Fraunhofer Institute for Computer Graphics, Darmstadt, Germany, 1994.

[Zhao96] J. Zhao, "A WWW Service To Embed And Prove Digital Copyright Watermarks", Proc. Of the European Conference on Multimedia Application, Services and Techniques, May 1996.

A Glossary

Some key terms used in the description of the digital copyright protection scheme are explained here.

Image An image in either digital or physical form. It may constitute a still image or a video frame. It can also refer to other types of data, such as video and audio data.

Signal A signal in either digital or physical form. It may refer to one dimensional or multidimensional signals such as image and video signals.

Image Copyright Holder (H) A party (or a process acting on behalf of it) "owning" a digital image. This is the party that generates the watermarks.

Image Buyer (B) A party (or a process acting on behalf it) which obtains (e.g. by purchase) via electronic means a specific image from H.

Image Authentication Data (IAD) The authentication data used in the image authentication process.

Stego-image Implies that an image or data is marked (i.e. it has an IAD embedded in it). The stego-image is also referred to as the stego data set.

Cover-image Implies that an image or data is unmarked (i.e. it has no IAD embedded in it). The cover-image is also referred to as the cover data set.

Watermark The form the IAD takes when it is suitable for embedding in a signal.

Image Copyright Certificate Center (C) An organization (or a process which acts on behalf it) which registers ownership for a specific image. Successful registration is based on a verification procedure such as checking the name and postal address of H, information how ownership was acquired, the title of the image, a description of the type of image (artistic, literary, musical, dramatic) and date and place of first publication. After registration a digital copyright certificate is generated.

Digital copyright certificate Digital copyright data which comprise the copyright certificate data and a digital signature.

Public watermark A watermark that can be detected using a publicly available key.

Detection watermark An image independent watermark that can be detected using a secret key.

Private watermark An image dependent watermark that can only be detected using a private key. It is not possible for an unauthorized third party to overwrite or delete the private watermark without the cryptographic secret keying information.

Payload The core of the hidden Image Authentication Data in bit form without error control coding applied.

Image ID The unique image ID is represented by an 8 byte hash determining H and C, followed by 4 bytes assigned by H for unique identification of each of his images.

Oblivious A watermarking technique which does not require the cover-image for extracting the mark. In other words, only the stego image is required to extract the mark when using an oblivious marking scheme.

Template A hidden message encoded in the image. By detecting the template, the scaling (zooming) and rotation suffered by a stego-image can be determined.