Rotation, scale and translation invariant spread spectrum digital image watermarking

O'Ruanaidh, Joséph John; Pun, Thierry

Rotation, Scale and Translation Invariant Digital Image Watermarking

Joseph J.K. Ó Ruanaidh and Thierry Pun

Centre Universitaire d'Informatique, Université de Genève, 24 rue Général Dufour, CH-1211 Genève 4, Switzerland

A digital watermark is an invisible mark embedded in a digital image which may be used for Copyright Protection. This paper describes how Fourier-Mellin transform-based invariants can be used for digital image watermarking. The embedded marks are designed to be unaffected by any combination of rotation, scale and translation transformations. The original image is not required for extracting the embedded mark.

1 Introduction

Computers, printers and high rate digital transmission facilities are becoming less expensive and more widespread. Digital networks provide an efficient cost-effective means of distributing digital media. The popularity of the World Wide Web has clearly demonstrated the commercial potential of the digital multimedia market. Unfortunately however, digital networks and multimedia also afford virtually unprecendented opportunities to pirate copyrighted material. The idea of using a robust digital watermark to detect and trace copyright violations has therefore stimulated significant interest among artists and publishers. As a result, digital image watermarking has recently become a very active area of research. Techniques for hiding watermarks have grown steadily more sophisticated and increasingly robust to lossy image compression and standard image processing operations, as well as to cryptographic attack.

Many of the current techniques for embedding marks in digital images have been inspired by methods of image coding and compression. Information has been embedded using the Discrete Cosine Transform (DCT) [16,34,5,6] Discrete Fourier Transform magnitude and phase [15], Wavelets [16], Linear Predictive Coding [13] and Fractals [9,22]. The key to making watermarks robust has been the recognition that in order for a watermark to be robust it must be embedded in the *perceptually significant* components of the image [16,5,6]. The term "perceptually significant" is somewhat subjective but it suggests that a

⋆ This work is supported by the Swiss National Science Foundation (grant no. 5003-45334)

good watermark is one which takes account of the behaviour of human visual system. Objective criteria for measuring the degree to which an image component is significant in watermarking have gradually evolved from being based purely on energy content [16,5,6] to statistical [20] and psychovisual [27,10] criteria.

Digital watermarking is also fundamentally a problem in digital communications [16,25,5,6]. In parallel with the increasing sophistication in modelling and exploiting the properties of the human visual system, there has been a corresponding development in communication techniques. Early methods of encoding watermarks were primitive and consisted of no more than incrementing an image component to encode a binary '1' and decrementing to encode a '0' [3,16]. Tirkel and Osborne [29] were the first to note the applicability of spread spectrum techniques to digital image watermarking. Since then there has been an increasing use of spread spectrum communications in digital watermarking. It has several advantageous features such as cryptographic security [29,30,6], and is capable of achieving error free transmission of the watermark near or at the limits set by Shannon's noisy channel coding theorem [16,25].

Spread spectrum is an example of a symmetric key [24] cryptosystem. System security is based on proprietary knowledge of the keys (or the seeds for pseudorandom generators) which are required to embed, extract or remove an image watermark. One proviso in the use of a spread spectrum system is that it is important that the watermarking process incorporate some non-invertible step which may depend on a private key or a hash function of the original image. Only in this way can true ownership of the copyright material be resolved [8].

The ability of humans to perceive the salient features of an image regardless of changes in the environment is something which humans take for granted [26,14]. We can recognize objects and patterns independently of changes in image contrast, shifts in the object or changes in orientation and scale. Gibson [12] makes the hypothesis that the human visual system is strongly tied to the ability to recognize invariants. It seems clear that an embedded watermark should have the same invariance properties as the image it is intended to protect. In this paper, we propose that an image watermark should be, so far as possible, encoded to be *invariant* to image transformations. We shall also demonstrate how image invariants can be used to construct watermarks that are unaltered by some of the most basic operations encountered in image processing; namely rotation, translation and changes of scale.

1.1 Nomenclature

This paper will make use of terms agreed during the 1996 Workshop on Information Hiding [18]. The term "cover image" will be used to describe the unmarked original image and "stegoimage" for an image with one or more hidden embedded marks. One significant deviation from the recommended steganographic nomenclature is the frequent use of the term "watermark" to describe the embedded mark. The authors believe this usage is perfectly acceptable because it has become the norm.

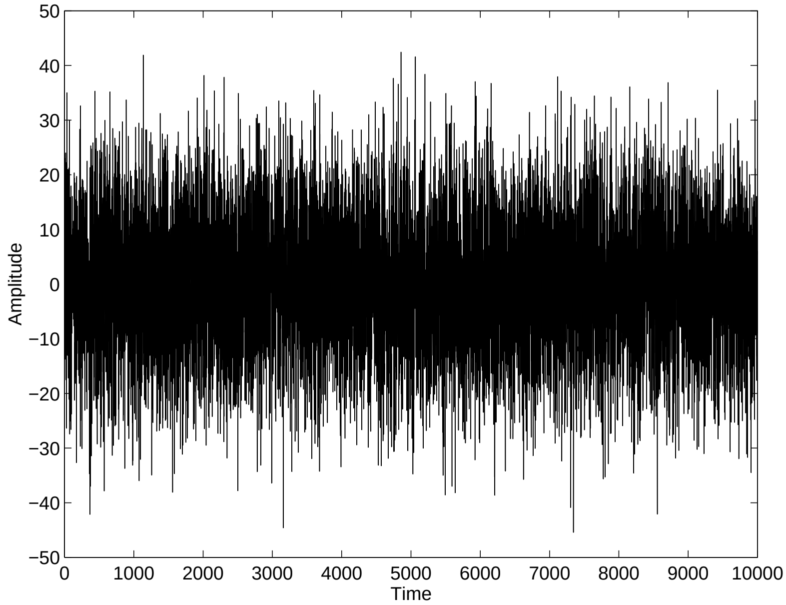

Fig. 1. *An example of a spread spectrum signal used as a digital watermark.*

2 Spread Spectrum

Pickholtz et al. [19] define spread spectrum communications as follows:

> Spread spectrum is a means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information; the band spread is accomplished by a code which is independent of the data, and a synchronized reception with the code at the receiver is used for despreading and subsequent data recovery.

Spread spectrum systems are also capable of approaching the Shannon limit for reliable communication. The fundamental information theoretic limits to reliable communication and its implications to digital watermarking have been discussed by some authors [16,25]. Note that the smaller is the number of bits of core information or “payload” contained in a watermark, the greater the chance of it being communicated without error.

Cox et al [7,6] recover a watermark by explicitly computing the correlation between the (noise corrupted) watermark recovered from the image with the perfect watermarks stored in a database. This is a very robust technique for watermark recovery but it is not very useful in practice because of the need for access to the database of marks and the large amount of computation required. In this paper the approach is similar to other spread spectrum approaches in that the watermark is embedded in the form of a pseudorandom sequence. However the approach is different to that of Cox in that it does not require access to a database of watermarks and is not particularly expensive computationally. In common with other spread spectrum techniques, in order to embed a mark or to extract it, it is important to have access to the key which is simply the seed used to generate pseudo-random sequences. In the case of a public watermarking scheme the key is generally available and may even be contained in publically available software. In a private watermarking scheme the key is proprietary. A mark may be embedded or extracted by the key owner which in our model is the Copyright Holder. In this form spread spectrum is a symmetric key cryptosystem. The infrastructure required to generate, issue and store the keys is not described here.

From the point of view of embedding watermarks in documents given the keys or seeds the sequences themselves can be generated with ease. A good spread spectrum sequence is one which combines desirable statistical propeties such as uniformly low cross correlation with cryptographic security. Examples of sequences used in spread spectrum systems used in digital watermarking include m-sequences, Gold codes, Kasami codes and Legendre sequences.

2.1 CDMA coding of digital watermarks

A method for encoding binary messages which can later be recovered given knowledge of the key used is described here. Suppose we are given a message which, without loss of generality, is in binary form $b_1, b_2 \cdots b_L$ where b_i are the bits. This can be written in the form of a set of symbols $s_1, s_2 \cdots s_M$, most generally by a change in a number base from 2 to B with $L \leq M \log_2 B$. The conversion from base 2 to a base which is a power of two is trivial. The next stage is to encode each symbol s_i in the form of a pseudorandom vector of length N. To encode the first symbol a pseudorandom sequence $\vec{v}$ of length $N + B - 1$ is generated. To encode a symbol of values where $0 \leq s < B$ the elements $v_s, v_{s+1} \cdots v_{s+N}$ are extracted as a vector $\vec{r_1}$ of length N. For the next symbol another independent pseudorandom sequence is generated and the symbol encoded as a random vector $\vec{r_2}$. Each successive symbol is encoded in the same way. Note that even if the same symbol occurs in different positions in the sequence that 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 :

$$\vec{m}(t_i) = \sum_{j=1}^{L} \vec{r_j}(t_i) \tag{1}$$

The pseudo-random vector $\vec{m}$ is decoded by generating all of the random vectors $\vec{r_i}$ in turn and recovering the symbols which the largest value of cross correlation. In this example the pseudo-random generator (PRG) is an m-sequence generator but this is not material to the issue since any "good" generator will do. In addition, one may use two dimensional or higher dimensional arrays in place of the pseudorandom vectors described in the communications system above. One interesting point is that for M sufficiently large the statistical distribution of the message m should approach a Gaussian. This follows from the Central Limit Theorem. A Gaussian distributed watermark has the advantage that it is more difficult to detect. The variance increases with order M - in other words, the expected peak excursion of the sequence is only order

M . One can expect that a message with M = 100 symbols will only have ten times the ampliude of a message with M = 1 symbols. This is very good from the point of view of minimising the visibility of the watermark

Figure 1 shows a spread spectrum signal $s(t)$ composed of a linear combination of L random vectors $r_i(t)$ as given by equation 1. Each random vector is specifically chosen to represent a particular symbol occupying a position in the message. A symbol may be composed of any number of bits. In our case each symbol is eight bits long and the number of random vectors L is nineteen. This is a form of Direct Sequence Code Division Multiple Access (DS-CDMA) spread spectrum communications. The encoded message in Figure 1 reads "This is a watermark".

This form of spread spectrum is resistant to cropping (providing it is resynchronised), non-linear distortions of amplitude and additive noise. Also, if it has good statistical properties it should be mistaken for noise and go undetected by an eavesdropper. The specific choice of method for generating the pseudorandom sequence has direct implications for reliability and cryptographic security of the embedded mark. Pseudorandom number generators described in watermarking literature include Gold Codes, Kasami codes, m-sequences [32,29,33,30] and perfect maps [31].

There are however some drawbacks to using direct sequence spread spectrum. Although a spread spectrum signal as described above is extremely resistant to non-linear distortion of its amplitude and additive noise it is also intolerant of timing errors. Synchronization is of the utmost importance during watermark extraction. If watermark extraction is carried out in the presence of the cover image then synchronization is relatively trivial. The problem of synchronizing the watermark signal is much more difficult to solve in the case where there is no cover image. If the stegoimage is translated, rotated and scaled then synchronization necessitates a search over a four dimensional parameter space (X-offset, Y-offset, angle of rotation and scaling factor). The search space grows even larger if one takes into account the possibility of shear and a change of aspect ratio.

In this paper, the aim is to investigate the possibility of using invariant representations of a digital watermark to help avoid the need to search for synchronization during the watermark extraction process.

2.2 Error control codes

It is desirable to incorporate some form of error control coding into the above scheme. The method is symbol based rather than binary bit based as in normal error codes. Because in this implementation each symbol may be correctly received or not, one finds that errors in the bit stream after despreading will occur in bursts, where each burst is due to an incorrectly decoded symbol. Reed Solomon (RS) codes [4,28,1] are powerful codes which are particularly suited to this application. RS codes can correct both errors (the locations of which are unknown) and erasures (the locations of which are exactly known). The probability of a false detection is extremely low. Reed Solomon codes are particularly suited to this application for the following reasons : RS codes correct symbol errors rather than bit errors. 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 - an erasure is useless redundancy. We recognise that this property of being able to discard erased symbols has two advantages : If the posterior probability of a received symbol is low then it may be ignored. RS codes only come in standard sizes. For example a 255x8 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 document in the first place). For a symbol length of eight bits the corresponding RS code (based on a Galois extension field GF(2^8)) will be 255 symbols long. This is considerably longer than a watermark (typically approximately 100 bits only). However, this is not a problem since the unneeded symbols can be flagged as erasures and they play no part in the decoding process.

3 Integral Transform Invariants

There are many different kinds of image invariant such as moment, algebraic and projective invariants [23,26]. In this section we will briefly outline the development of several integral transform based invariants [26].

The invariants described below depend on the properties of the Fourier transform. There are a number of reasons for this. First, using integral transform-based invariants is a a relatively simple generalization of transform domain watermarking. Second, the number of robust invariant components is relatively large which makes it suitable for spread spectrum techniques. Third, as we shall see, mapping to and from the invariant domain to the spatial domain is well-defined and it is in general not computationally expensive.

3.1 The Fourier Transform

Let the image be a real valued continuous function $f(x_1, x_2)$ defined on an integer-valued Cartesian grid $0 \leq x_1 < N_1, 0 \leq x_2 < N_2$.

The Discrete Fourier Transform (DFT) is defined as follows:

$$F(k_1, k_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(x_1, x_2) e^{-j2\pi x_1 k_1/N_1 - j2\pi x_2 k_2/N_2} \qquad (2)$$

The inverse transform is

$$f(x_1, x_2) = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} F(k_1, k_2) e^{j2\pi k_1 x_1/N_1 + j2\pi k_2 x_2/N_2} \qquad (3)$$

The DFT of a real image is generally complex valued.

This leads to a

magnitude and phase representation for the image:

$$A(k_1, k_2) = [F(k_1, k_2)] \tag{4}$$

$$\Phi(k_1, k_2) = \angle F(k_1, k_2) \tag{5}$$

We now discuss the properties of the Fourier representation that are crucial to the construction of translation, rotation and scaling invariants.

3.1.1 The Translation Property

Shifts in the spatial domain cause a linear shift in the phase component.

$$F(k_1, k_2) \exp\left[-j(ak_1 + bk_2)\right] \leftrightarrow f(x_1 + a, x_2 + b) \tag{6}$$

Note that both $F(k_1, k_2)$ and its dual $f(x_1, x_2)$ are periodic functions so it is implicitly assumed that translations cause the image to be "wrapped around". We shall refer to this as a *circular translation.*

3.1.2 Reciprocal Scaling

Scaling the axes in the spatial domain causes an inverse scaling in the frequency domain.

$$\frac{1}{\rho} F\left(\frac{k_1}{\rho}, \frac{k_2}{\rho}\right) \leftrightarrow f(\rho x_1, \rho x_2) \tag{7}$$

An important example of this property is the Fourier transform of a delta function (which is infinitely narrow) which has a uniformly flat amplitude spectrum (and is infinitely wide).

3.1.3 The Rotation Property

Rotating the image through an angle θ in the spatial domain causes the Fourier representation to be rotated through the same angle.

$$\begin{aligned} &F(k_1 \cos\theta - k_2 \sin\theta, k_1 \sin\theta + k_2 \cos\theta) \\ &\quad \leftrightarrow f(x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta) \end{aligned} \tag{8}$$

Note that the grid is rotated so the new grid points may not be defined. The value of the image at the nearest valid grid point may be estimated by interpolation.

3.2 Translation Invariance

From property 6 of the Fourier transform it is clear that spatial shifts affect only the phase representation of an image. This leads to the well known result that the DFT magnitude is a circular translation invariant. An ordinary translation can be represented as a cropped circular translation.

It is less well known that it is possible to derive invariants based on the phase representation. To do this involves eliminating the translation dependent linear term from the phase representation. Brandt and Lin [2] present two such translation invariants, namely the *Taylor invariant* which removes the linear phase term in the Taylor expansion of the phase and the *Hessian invariant* which removes this linear phase term by double differentiation.

We shall see in section 3.3 that properties 7 and 8 allow one to extend the basic translation invariants to cover changes of rotation and scale.

3.3 Rotation and Scale Invariance

The basic translation invariants described in section 3.2 may be converted to rotation and scale invariants by means of a *log-polar mapping.*

Consider a point $(x, y) \in \Re^2$ and define:

$$\begin{aligned} x &= e^{\mu} \cos \theta \\ y &= e^{\mu} \sin \theta \end{aligned} \tag{9}$$

where $\mu \in \Re$ and $0 \leq \theta < 2\pi$. One can readily see that for every point (x, y) there is a point (μ, θ) that uniquely corresponds to it.

The new coordinate system has the following properties:

Scaling is converted to a translation.

$$(\rho x, \rho y) \leftrightarrow (\mu + \log \rho, \theta) \tag{10}$$

Rotation is converted to a translation.

$$\begin{aligned} &(x \cos (\theta + \delta) - y \sin (\theta + \delta), x \sin (\theta + \delta) + y \cos (\theta + \delta)) \\ &\quad \leftrightarrow (\mu, \theta + \delta) \end{aligned} \tag{11}$$

At this stage one can implement a rotation and scale invariant by applying a translation invariant in the log-polar coordinate system. Taking the Fourier transform of a log-polar map is equivalent to computing the Fourier-Mellin transform:

$$F_M(k_1, k_2) = \int_{-\infty}^{\infty} \int_{0}^{2\pi} f(e^{\mu}\cos\theta, e^{\mu}\sin\theta)\exp\left[i(k_1\mu + k_2\theta)\right] d\mu\, d\theta \qquad (12)$$

The modulus of the Fourier-Mellin transform is rotation and scale invariant.

Many useful invariants are derived by finding an alternative coordinate system in which the effect of the transformation is replaced by a translation and applying a translation invariant operator in the new coordinate syatem. Squire [26] demonstrates how such invariants can be derived formally using the methods of Lie Group algebra.

3.3.1 The Commutative Property

It is interesting to show that the single parameter group of rotation transformations $\mathcal{R}(\theta)$ and the single parameter group of scale tranformations $\mathcal{S}(\rho)$ commute.

$$\begin{aligned} \mathcal{R}(\theta) \circ \mathcal{S}(\rho)\, f(x, y) &= \mathcal{R}(\theta)\, f(\rho x, \rho y) \\ &= f(\rho x \cos\theta - \rho y \sin\theta, \rho x \sin\theta + \rho y \cos\theta) \\ &= \mathcal{S}(\rho)\, f(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \\ &= \mathcal{S}(\rho) \circ \mathcal{R}(\theta)\, f(x, y) \end{aligned} \qquad (13)$$

Similarly one can show [2] that the two parameter group of translation transformations $\mathcal{T}(\alpha, \beta)$ commutes neither with $\mathcal{R}(\theta)$, nor with $\mathcal{S}(\rho)$ nor with the joint transformation $\mathcal{RS}(\theta, \rho)$.

3.4 Rotation, Scale and Translation Invariance

Consider two invariant operators: $\mathcal{F}$ which extracts the modulus of the Fourier transform and $\mathcal{F_M}$ which extracts the modulus of the Fourier-Mellin transform. Applying the hybrid operator $\mathcal{F_M} \circ \mathcal{F}$ to an image $f(x, y)$ we obtain:

$$I_1 = [\mathcal{F_M} \circ \mathcal{F}]\ f(x, y) \qquad (14)$$

Let us also apply this operator to an image that has been translated, rotated and scaled:

$$I_2 = [\mathcal{F_M} \circ \mathcal{F} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\ f(x, y)$$

$$= [\mathcal{F_M} \circ \mathcal{R}(\theta) \circ \mathcal{F} \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\ f(x, y) \tag{15}$$

$$= \left[\mathcal{F_M} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\frac{1}{\rho}) \circ \mathcal{F} \circ \mathcal{T}(\alpha, \beta)\right]\ f(x, y) \tag{16}$$

$$= [\mathcal{F_M} \circ \mathcal{F}]\ f(x, y) \tag{17}$$

$$= I_1 \tag{18}$$

Hence $I_1 = I_2$ and the represention is rotation, scale and translation invariant. Steps 15 and 16 follow from properties 8 and 7 of the Fourier transform respectively. The contraction in equation 17 is due to the invariance properties of $\mathcal{F}$ and $\mathcal{F_M}$.

The rotation, scale and translation ($\mathcal{RST}$) invariant just described is sufficient to deal with any combination or permutation of rotation, scale and translation in any order [2].

To give a concrete example of its application, consider a copy of a stegoimage placed on a scanner from which we wish to extract an embedded mark. The image may be reduced or increased in size and will be, more often than not, at an angle of $\pm\epsilon$, $\pm 90 \pm \epsilon$ or even $180 \pm \epsilon$ degrees where $\pm\epsilon$ is some small random angle. The image is also likely to be translated. Using the invariants derived above it should be possible to to extract an embedded mark regardless of orientation, scale or position.

3.5 Complete and Strong invariants

Brandt and Lin [2] define the important concept of *completeness*. A complete invariant represents "all the information contained in the image modulo the given transformation". In this sense a complete invariant is *almost* invertible. If two images have the same complete translation invariant then, by the definition of completeness, one must be a shifted version of the other. Such an invariant cannot be inverted uniquely because the mapping to the invariant domain is not a bijective function. Brandt and Lin [2] present an example where a complete Hessian invariant is inverted to yield the original image, albeit with the origin shifted and image wrapped around at the edges.

Ferraro and Caelli [11] in an earlier paper defined the related concept of *strong* invariance. "An integral transform is defined to be invariant in the strong sense if ..." the amplitude representation is constant for all states of the transformation and different states are uniquely encoded in the phase component. The phase component may therefore be used to invert the invariant representation.

For convenience, the invariants used in this paper are strongly invariant. In image watermarking it is more convenient to use strong invariants because the last stage of the process of *embedding* a mark involves inverting the invariant representation to obtain the (marked) stegoimage. Invertibility is of no concern whatsoever during the extraction process.

4 Watermark Implementations

In this section we describe two different prototype schemes for embedding watermarks in digital images using $\mathcal{RST}$ invariants. Typically, the watermark is embedded in a gray scale image or the luminance component of a colour image.

4.1 General scheme

Figure 2 illustrates the process of obtaining the $\mathcal{RST}$ transformation invariant from a digital image. The watermark takes the form of a two dimensional spread spectrum signal in the $\mathcal{RST}$ transformation invariant domain. Note that the size of the $\mathcal{RST}$ invariant representation depends on the resolution of the log-polar map which can be kept the same for all images. This is a convenient feature of this approach which helps to standardise the embedding and detection algorithms.

4.1.1 Embedding a watermark

A Fourier transform (FFT) is first applied which is then followed by a Fourier-Mellin transform (A log-polar mapping (LPM) followed by a Fourier transform (FFT)). The invariant coefficients preselected for their robustness to image processing are marked using a spread spectrum signal. The inverse mapping is computed as an inverse FFT (IFFT) followed by an inverse Fourier-Mellin transform (An inverse log-polar mapping (ILPM) followed by an inverse FFT) Note that the inverse transformation from RST invariant domain to the image domain uses the phase computed during the forward transformations from image domain to the RST invariant domain.

4.1.2 Extracting a watermark

A watermark may be extracted without or without a cover image. In the case where there is no cover image the image is transformed to the $\mathcal{RST}$ invariant domain and the watermark is decoded. This is similar in principle to the scheme described by Smith and Comiskey [25] whose approach is to "treat the image as noise" and overcome the interference from the stegoimage using spread spectrum communication. When a cover image is available it should be subtracted from the stegoimage and the difference transformed to the $\mathcal{RST}$ invariant domain (since the operations in Figure 2 are linear with respect to image amplitude). Subtracting the cover image improves the performance of the detector because, as Smith and Comiskey point out, it eliminates the noise interference due to the stegoimage [25]. In many cases, image contrast may be distorted, for example by a scanner, in which case the effects of change of contrast must be compensated for in some way. Cox et al. [5,6] describe a method known as dynamic histogram warping [7] to carry this out.

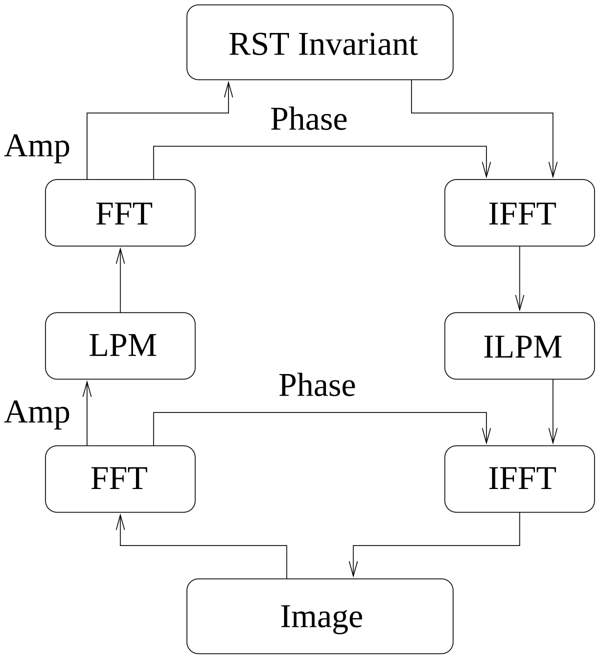

Fig. 2. *A diagram of a prototype RST invariant watermarking scheme.*

4.1.3 Practical considerations

There are a number of complications in implementing the processing steps depicted in Figure 2. The stegoimage must be real which in turn means that its amplitude spectrum ($A(k,l)$ where $0 \leq k < M$ and $0 \leq l < N$) as well as being positive ($A(k,l) \geq 0$) must also be positively symmetric:

$$A(k,l) = A(M-k, N-l) \tag{19}$$

The log-polar map of a positively symmetric matrix consists of two identical halves. This follows from the fact that the positive symmetry condition in equation 19 is written in polar coordinates as:

$$A(r,\theta) = A(r, \pi + \theta) \tag{20}$$

where $(M/2, N/2)$ is the centre of rotation. Since both halves of the log-polar map are identical then only one half need be used in the upper FFT of Figure 2. The spread spectrum signal is determined from the amplitude spectrum of this FFT. Applying the above in reverse gives an embedding algorithm which yields a real valued watermark.

The scheme described in Figure 2 works in principle but has some serious deficiencies in practice. The first difficulty is that both the log-polar mapping and the inverse log-polar mapping can cause a loss of image quality. The change of coordinate system means that some form of interpolation should be used.

Two simple forms of interpolation, nearest neighbour and bilinear interpolation [21], are in common use. Non-stationary low pass filtering can improve the performance by eliminating frequency aliasing. In practice the resolution of the log-polar map must be at least 512×512 for even a quite poor quality image. The second difficulty is numerical. Interpolation only performs well if neighbouring samples are of the same scale. This makes the computation of the Fourier-Mellin transform of the modulus of a Fourier transform somewhat problematic. A typical Fourier transform representation of an image is quite badly behaved in this respect since there are generally a few components of relatively large magnitude. This difficulty is resolved in the next section.

4.2 Cover Image Independent Scheme

The problems in embedding watermarks using the previous implementation described in Figure 2 can be circumvented by using the method illustrated in Figure 3. In this case the mark must be embedded in the $\mathcal{RST}$ invariant domain independently of the original image. The advantage of using this approach is that the distortions caused by the inverse log-polar map are suffered only by the embedded mark itself and do not affect the stegoimage. Figure 4 shows the corresponding detection process which is relatively straightforward.

Note that when embedding the mark there is no phase component available for the first inverse Fourier transform. The first FFT operates on a random phase signal to keep the amplitude distribution of the inverse FFT reasonably flat. This is beneficial to the inverse log-polar map which performs best when the input is a smooth image. The second FFT uses the phase component directly from the cover image. The advantage in doing this is that matching the phase component of the embedded mark to that of the cover image helps to hide it because the embedded mark resembles the cover image. This follows from the research of Oppenheim and Lim [17] which demonstrates that image phase is far more important to image structure than image amplitude.

5 Examples

Figure 5 depicts a standard image of a mandrill. Figure 6 is the log-polar map of Figure 5. This image was computed using 600 grid points along the θ (angle) axis, 600 grid points along the μ (log-radial) axis and bilinear interpolation. Figure 7 is the inverted log-polar map computed using just 100 angular and 100 log-radial grid points and nearest neighbour interpolation. Note that the restoration grows progressively poorer away from the centre.

Figure 5 is in fact a stegoimage which contains a 104 bit rotational and scale invariant watermark. The watermark is encoded as a spread spectrum signal which was embedded in the RS invariant domain. Figure 5 was rotated by 143^{O} and scaled by a factor of 75% along each axis to give the image shown in Figure 8. The embedded mark which read "The watermark" in ASCII code was recovered from this stegoimage. It was also found that the watermark survived

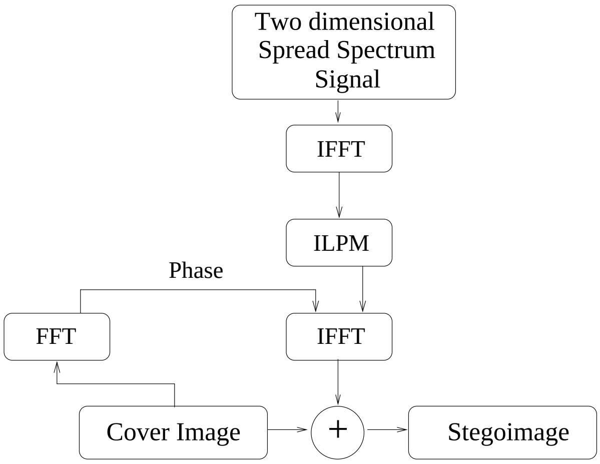

Fig. 3. *A method of embedding a mark in an image which avoids mapping the cover image into the RST invariant domain.*

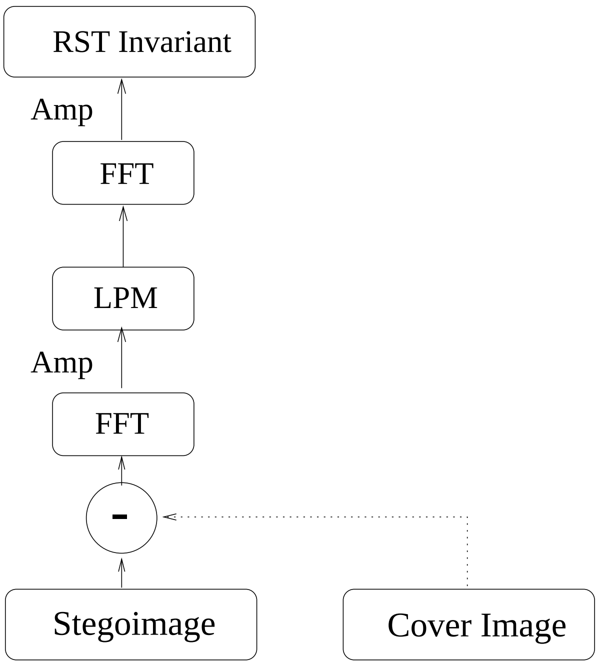

Fig. 4. *A scheme to extract a mark from an image.*

lossy image compression using JPEG at normal settings (75% quality factor). Other methods exist that tolerate JPEG compression down to 5% quality factor [7,6,16,15]; work is underway to combine these with this approach. In addition, the mark is also reasonably resistant to cropping and could be recovered from a segment approximately 50% of the size of the original image.

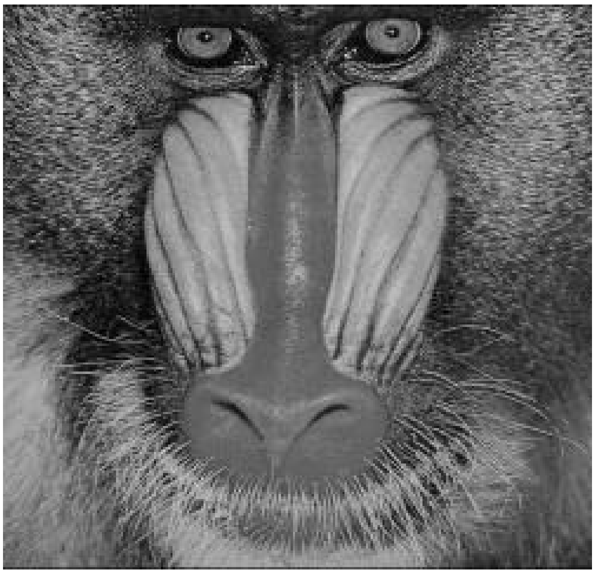

Fig. 5. *A standard* 500×480 *image of a mandrill.*

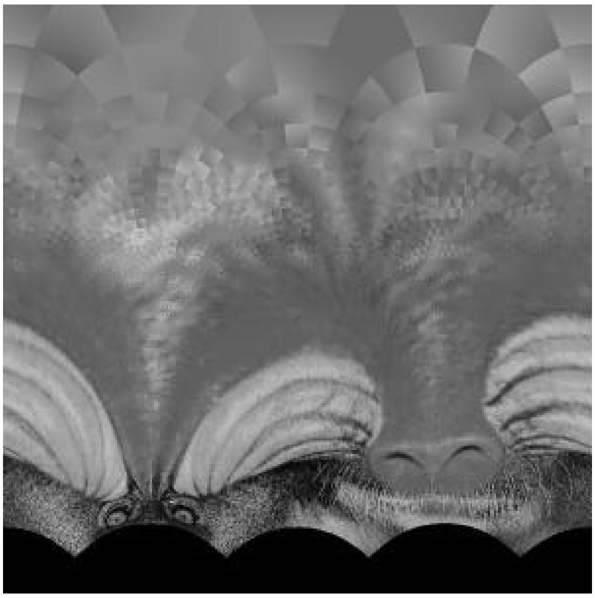

Fig. 6. *A log polar map of the image of a mandrill. The log-polar map employs bilinear interpolation and the log-polar grid is* 600×600 *samples.*

6 Conclusion

This paper has outlined the theory of integral transform invariants and showed that this can be used to produce watermarks that are resistant to translation,

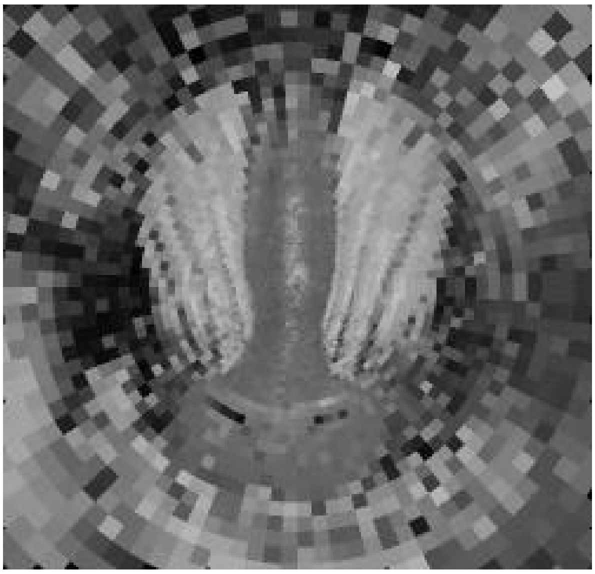

Fig. 7. *The image of a mandrill reconstructed from a log polar map of size 100 × 100 samples. This reconstruction uses nearest neighbour interpolation.*

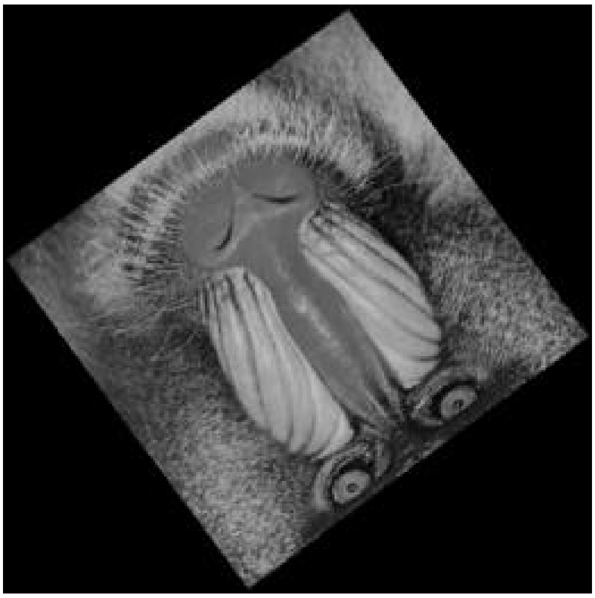

Fig. 8. *A watermarked image of a mandrill that has been rotated by 143 degrees and scaled by 75%. The embedded mark was recovered from this image.*

rotation and scaling. The importance of invertibility of the invariant representation was emphasised. One of the significant points is the novel application of the Fourier-Mellin transform to digital image watermarking.

There are several advantages in using integral transform domain marks. The main advantage is that the transforms can be computed very quickly (although in practice it has been found that the inverse log-polar mapping is a computational bottleneck). In addition, transform space contains a large number of samples which can be used to hide a spread spectrum signal.

An example of a rotation and scale invariant watermark was presented. As one might expect, this proved to be robust to changes in scale and rotation. It was also found to be weakly resistant to lossy image compression and cropping. The robustness of the embedded mark to these attacks will be greatly improved with future work.

On its own, the invariant watermark discussed in this paper cannot resist changes in aspect ratio or shear transformations. There is no obvious means of constructing an integral transform-based operator that is invariant to these transformations. However, work is currently in progress to find a means of searching for the most likely values of aspect ratio and shear factor, and then to apply the necessary corrections during watermark extraction.

In addition to the above, we intend to investigate the possible use of phase-based complete invariants. This would have some advantage over only marking strong invariants, since a complete invariant presents a maximal number of potential communications channels through which watermark information may be transmitted.

Acknowledgement

We wish to thank Dr David McG. Squire, Sergei Starchik and Dr Feng-Lin for their extremely helpful advice on the theory of invariants and Dr A. Z. Tirkel for many stimulating conversations and for exchanging many ideas. We are also grateful to Dr Alexander Herrigel and Adrian Perrig for their useful comments.

References

[1] R. E. Blahut. *The theory and practice of error control codes.* Addison-Wesley, 1983.

[2] R. D. Brandt and F. Lin. Representations that uniquely characterize images modulo translation, rotation and scaling. *Pattern Recognition Letters*, 17:1001–1015, August 1996.

[3] G. Caronni. Assuring Ownership Rights for Digital Images. In H. H. Brueggemann and W. Gerhardt-Haeckl, editors, *Reliable IT Systems VIS '95.* Vieweg Publishing Company, Germany, 1995.

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

[5] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum communication for multimedia. Technical report, N.E.C. Research Institute, 1995. ftp://ftp.nj.nec.com/pub/ingemar/papers/watermark.ps.Z.

[6] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for images, audio and video. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 243–246, Lausanne, Switzerland, September 16-19 1996.

[7] I. Cox, S. Roy, and S. L. Hingorani. Dynamic histogram warping of image pairs for constant image brightness. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-95*, Austin, Texas, 1995.

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

[9] P. Davern and M. Scott. Fractal based image steganography. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 279–294, Cambridge, UK, May/June 1996. Springer Verlag.

[10] J.F. Delaigle, C. De Vleeschouwer, and B Macq. A psychovisual approach for digital picture watermarking. submitted to the Journal of Electronic Imaging, 1996.

[11] Mario Ferraro and Terry M. Caelli. Lie transform groups, integral transforms, and invariant pattern recognition. *Spatial Vision*, 8(1):33–44, 1994.

[12] James Gibson. *The Senses Considered as Perceptual Systems*. Houghton-Mifflin, Boston, Massachusetts, 1966.

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

[14] R. Milanese, S. Gil, and T. Pun. Attentive mechanisms for dynamic and static scene analysis. *Optical Engineering*, 34(8):2428–2434, August 1995.

[15] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Phase watermarking of digital images. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 239–242, Lausanne, Switzerland, September 16-19 1996.

[16] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland. Watermarking digital images for copyright protection. *IEE Proceedings on Vision, Image and Signal Processing*, 143(4):250–256, August 1996. Invited paper, based on the paper of the same title at the IEE Conference on Image Processing and Its Applications, Edinburgh, July 1995.

[17] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):529–541, May 1981.

[18] B Pfitzmann. Information hiding terminology. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 347–350, Cambridge, UK, May/June 1996. Springer Verlag.

[19] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread spectrum communications – a tutorial. *IEEE Transactions on Communications*, COM-30(5):855–884, May 1982.

[20] I Pitas. A method for signature casting on digital images. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 215–218, Lausanne, Switzerland, September 16-19 1996.

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

[22] J. Puate and F. Jordan. Using fractal compression scheme to embed a digital signature into an image. http://ltswww.epfl.ch/~kutter/publications/fvt.html, November 1996.

[23] T. H. Reiss. *Recognizing planar Objects Using Invariant Image Features*. Lecture Notes in Computer Science. Springer-Verlag, 1993.

[24] B. Schneier. *Applied Cryptography*. Wiley, 2nd edition, 1995.

[25] J. Smith and B. Comiskey. Modulation and information hiding in images. In Ross Anderson, editor, *Proceedings of the First International Workshop in Information Hiding*, Lecture Notes in Computer Science, pages 207–226, Cambridge, UK, May/June 1996. Springer Verlag.

[26] D. McG. Squire. *Model-based Neural Networks for Invariant Pattern Recognition*. PhD thesis, Curtin University of Technology, Perth, Western Australia, October 1996.

[27] M. D. Swanson, B. Zhu, and A. Tewfik. Transparent robust image watermarking. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 211–214, Lausanne, Switzerland, September 16-19 1996.

[28] P. Sweeney. *Error Control Coding: An Introduction*. Prentice-Hall, 1991.

[29] A. Z. Tirkel, G. A. Rankin, R. G. van Schyndel, W. J. Ho, N. R. A. Mee, and C. F. Osborne. Electronic watermark. In *Dicta-93*, pages 666–672, Macquarie University, Sydney, December 1993.

[30] A. Z. Tirkel, R. G. van Schyndel, and C. F. Osborne. A two-dimensional digital watermark. In *ACCV'95*, pages 378–383, University of Queensland, Brisbane, December 6-8 1995.

[31] A.Z. Tirkel. Image and watermark registration. Submitted to Signal processing, January 1997.

[32] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne. A digital watermark. In *IEEE Int. Conf. on Image Processing ICIP-95*, pages 86–90, Austin, Texas, 1994.

[33] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne. Towards a robust digital watermark. In *Dicta-95*, pages 504–508, Nanyang Technological University, Singapore, December 5-8 1995.

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