Abstract

A digital watermark is an invisible mark embedded in a digital image which may be used for Copyright Protection. This paper proposes that Fourier-Mellin transform-based invariants can be used for digital image watermarking. The embedded marks may be 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. Unfortunately however, digital networks and multimedia also afford virtually unprecedented 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) [6, 2] Discrete Fourier Transform magnitude and phase [5], Wavelets [6], Linear Predictive Coding and Fractals. 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 [6, 2]. The term "perceptually significant" is somewhat subjective but it suggests that a 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 [6, 2] to statistical [7] and psychovisual [3] criteria.

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 [10]. We can recognize objects and patterns independently of changes in image contrast, shifts in the object or changes in orientation and scale. It seems clear that an embedded watermark should have the same invariance properties as the image it is intended to protect.

Digital watermarking is also fundamentally a problem in digital communications [6, 9, 2]. 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. Tirkel and Osborne [11] 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 [11, 2], and is capable of achieving error free transmission of the watermark near or at the limits set by Shannon's noisy channel coding theorem [6, 9]. Note that the shorter is the core information or "payload" contained in a watermark then the greater is the chances of the watermark being communica tion reliably. Spread spectrum is also an example of a symmetric key [8] cryptosystem where system security is based on proprietary knowledge of the keys (or the
seeds for pseudorandom generators) required to embed, extract or remove an image watermark.

Synchronization of the watermark signal is of utmost importance during watermark extraction. If watermark extraction is carried out in the presence of the original 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 original image. If the watermarked image 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. A digital watermark that is invariant to these transformations requires no such search. The tradeoff here is between using a fully invariant representation which may be numerically unstable and expensive to compute with the expense of carrying out a search.

2 Integral Transform Invariants

There are many different kinds of image invariant such as moment, algebraic and projective invariants. In this section we will briefly outline the development of several integral transform based invariants [1].

The invariants described below depend on the properties of the Fourier transform. There are a number of advantages in using a transform based representation. First, using integral transform-based invariants is 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.

2.1 The Fourier Transform

Let the image be a real valued continuous function \( f(x_1, x_2) \) defined on an integer-valued Cartesian grid \( 0 \leq x_1 < N_1, 0 \leq x_2 < N_2 \). Let the two dimensional Discrete Fourier Transform (DFT) \( F(k_1, k_2) \) where \( 0 \leq k_1 < N_1, 0 \leq k_2 < N_2 \) be defined in the usual way [4].

2.1.1 The Translation Property

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

\[
F(k_1, k_2) \exp \left[ -j (a k_1 + b k_2) \right] \Leftrightarrow f(x_1 + a, x_2 + b) \tag{1}
\]

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.

2.1.2 Reciprocal Scaling

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

\[
\frac{1}{\rho} F(\frac{k_1}{\rho}, \frac{k_2}{\rho}) \Leftrightarrow f(\rho x_1, \rho x_2) \tag{2}
\]

2.1.3 The Rotation Property

Rotating the image through an angle \( \theta \) in the spatial domain causes the Fourier representation to be rotated through the same angle:

\[
F(k_1 \cos \theta - k_2 \sin \theta, k_1 \sin \theta + k_2 \cos \theta) \Leftrightarrow f(x_1 \cos \theta - x_2 \sin \theta, x_1 \sin \theta + x_2 \cos \theta) \tag{3}
\]

2.2 Translation Invariance

From the translation property 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.

2.3 Rotation and Scale Invariance

The basic translation invariants described in section 2.2 may be converted to rotation and scale invariants by means of a log-polar mapping. Consider a point \((x, y) \in \mathbb{R}^2\) and define:

\[
x = e^\mu \cos \theta
\]

\[
y = e^\mu \sin \theta
\]

where \( \mu \in \mathbb{R} \) and \( 0 \leq \theta < 2\pi \). One can readily see that for every point \((x, y)\) there is a point \((\mu, \theta)\) that uniquely corresponds to it. Note that in the new coordinate system scaling and rotation are converted to a translation of the \( \mu \) and \( \theta \) coordinates respectively. 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 (LPM) is equivalent to computing the Fourier-Mellin transform [1].

2.4 Rotation, Scale and Translation Invariance

Consider two invariant operators: \( F \) which extracts the modulus of the Fourier transform and \( F_M \) 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:

$$I_1 = [\mathcal{F}_M \circ \mathcal{F}] f(x, y) \quad (5)$$

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

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

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

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

$$= I_1 \quad (6)$$

Hence $I_1 = I_2$ and the representation is rotation, scale and translation invariant. The rotation, scale and translation (RST) invariant just described is sufficient to deal with any combination of rotation, scale and translation transformations in any order [1].

3 Watermarking Implementation

Figure 1 illustrates the process of obtaining the RST transformation invariant from a digital image. Figure 1 is for illustrative purposes only since the process used in practice is more complicated; the main difficulty being that the time and frequency domain are both discretely sampled spaces. The watermark takes the form of a two dimensional spread spectrum signal in the RST transformation invariant domain. Note that the size of the 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 Examples

Figure 2 is a standard image which contains a 104 bit rotational and scale invariant watermark. The watermark is encoded as a spread spectrum signal which was embedded in the $R.S$ invariant domain. Figure 2 was rotated by $143^\circ$ and scaled by a factor of 75% along each axis. The embedded mark which read “The watermark” in ASCII code was recovered from this watermarked image. It was also found that the watermark survived lossy image compression using JPEG at normal settings (75% quality factor). Other methods exist that tolerate JPEG compression down to 5% quality factor [2, 6]; 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.

5 Conclusion

This paper has outlined the theory of integral transform invariants and proposed that this can be used to produce watermarks that are resistant to translation, rotation and scaling. The importance of invertibility of the invariant representation was emphasised. One of the significant points is the application of the Fourier-Mellin transform to digital image watermarking.

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.

---

Figure 1: A diagram of a prototype RST invariant watermarking scheme.
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.

Acknowledgments

This work is supported by the Swiss National Science Foundation (grant no. 5003-45334). 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. Thanks also to Geoff Rhoads for answering our queries about PictureMarc.

References


