Rotation, scale and translation invariant digital image watermarking

O'RUANAIDH, Joséph John, PUN, Thierry

Abstract

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.

Reference

O'RUANAIDH, Joséph John, PUN, Thierry. Rotation, scale and translation invariant digital image watermarking. In: *IEEE International Conference on Image Processing, 1997. Proceedings (ICIP 97)*. 1997. p. 536-539

DOI : 10.1109/ICIP.1997.647968

æ
æ

ROTATION, SCALE AND TRANSLATION INVARIANT DIGITAL IMAGE WATERMARKING

Joseph J.K. Ó Ruanaidh and Thierry Pun

Centre Universitaire d'Informatique,
Université de Genève,
Switzerland

ABSTRACT

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 comb ination 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 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 methods including the Discrete Cosine Transform (DCT) [5, 2] Discrete Fourier Transform magnitude and phase, Wavelets, 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 [5, 2]. 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 [5, 2] to statistical and psychovisual 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 [4]. 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 [5, 7, 2]. In parallel with the increasing sophistication in modelling and exploiting the properties of the human visual system, there has been a corresponding application of advanced digital communication techniques such as spread spectrum [7, 5].

Synchronization of the watermark signal is of the utmost importance during watermark extraction. If watermark extraction is carried out in the presence of the cover image [1] 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. 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 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.

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

[1] This paper will make use of terms agreed during the 1996 Workshop on Information Hiding [6]. The term "cover image" will be used to describe the unmarked original image and the term "stegoimage" for an image with one or more hidden embedded marks.

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 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 [3].

2.1.1. The Translation Property

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

$$F(k_1, k_2) \exp\left[-j(ak_1 + bk_2)\right] \leftrightarrow f(x_1 + a, x_2 + b) \tag{1}$$

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

$$\begin{array}{l} F(k_1 \cos\theta - k_2 \sin\theta, k_1 \sin\theta + k_2 \cos\theta) \\ \quad \leftrightarrow f(x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta) \end{array} \tag{3}$$

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

$$\begin{array}{rcl} x & = & e^{\mu} \cos\theta \\ y & = & e^{\mu} \sin\theta \end{array} \tag{4}$$

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. Note that in the new coordinate system *scaling* and *rotation* are converted to a translation of the μ and θ 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 is equivalent to computing the Fourier-Mellin transform [1].

2.4. 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:

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

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

$$\begin{aligned} I_2 &= [\mathcal{F_M} \circ \mathcal{F} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\ f(x, y) \\ &= [\mathcal{F_M} \circ \mathcal{R}(\theta) \circ \mathcal{F} \circ \mathcal{S}(\rho) \circ \mathcal{T}(\alpha, \beta)]\ f(x, y) \quad (6) \\ &= \left[\mathcal{F_M} \circ \mathcal{R}(\theta) \circ \mathcal{S}(\frac{1}{\rho}) \circ \mathcal{F} \circ \mathcal{T}(\alpha, \beta)\right] f(x, y) \quad (7) \\ &= [\mathcal{F_M} \circ \mathcal{F}]\ f(x, y) \quad (8) \\ &= I_1 \quad (9) \end{aligned}$$

Hence $I_1 = I_2$ and the represention is rotation, scale and translation invariant. Steps 6 and 7 follow from the rotation and reciprocal scaling properties of the DFT [1]. 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 [1].

3. WATERMARKING IMPLEMENTATION

Figure 1 illustrates the process of obtaining the $\mathcal{RST}$ transformation invariant from a digital image. Figure 1 is for illustrative purposes only since the process used in practice is more complicated. 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. 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 *RS* invariant domain. Figure 2 was rotated by 143^O and scaled by a factor of 75% along each axis. The embedded mark which read "The watermark" in ASCII code was recovered from this stegoimage. 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, 5]; 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 showed that this can be used to produce watermarks that are resistant to translation, rotation and

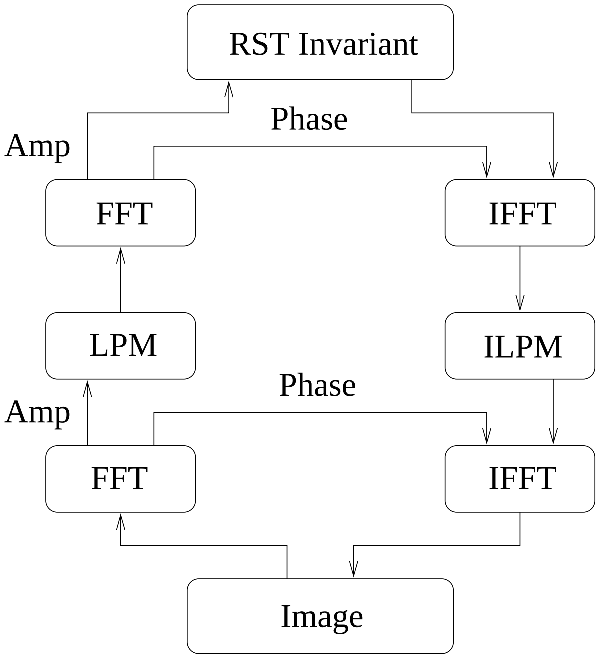

Figure 1: *A diagram of a prototype RST invariant watermarking scheme.*

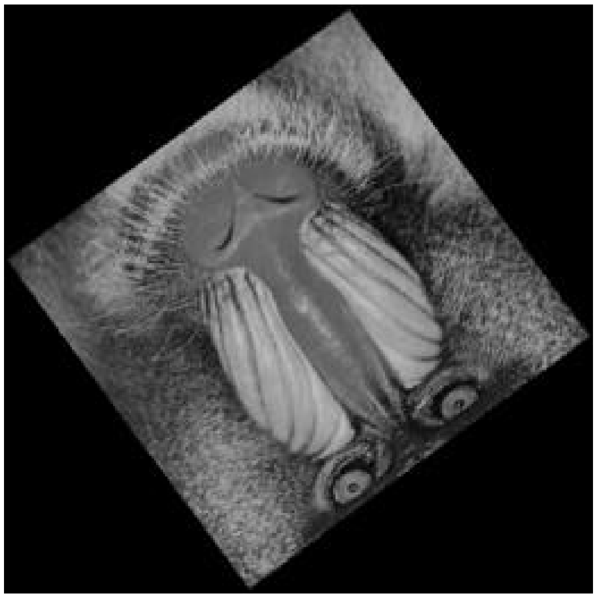

Figure 2: *A watermarked image of a mandrill that has been rotated by 143 degrees and scaled by 75%. The embedded mark was recovered from this image.*

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.

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.

6. REFERENCES

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

[2] 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.

[3] Jae S. Lim. *Two-Dimensional Signal and Image Processing*. Prentice-Hall International, 1990.

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

[5] 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.

[6] 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.

[7] 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.