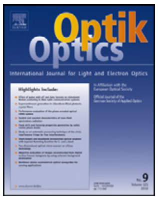
Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo



Blind image watermarking based on discrete fractional random transform and subsampling

Hao Luo, Fa-Xin Yu*, Zheng-Liang Huang, Zhe-Ming Lu

School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310029, PR China

ARTICLE INFO

Article history:
Received 11 June 2009
Accepted 27 December 2009

Keywords:
Blind image watermarking
Discrete fractional random transform
Image subsampling
Tamper detection
Self-embedding

ABSTRACT

This paper proposes a blind watermarking scheme based on discrete fractional random transform. The watermark information can be a binary sequence, a gray level image or a set of decimal fractions sampled from a given source signal. The host image is subsampled into four subimages, and the high correlations among their discrete fractional random transform coefficients are exploited for watermark embedding. Based on this self-reference strategy, the watermark can be extracted without the aid of the host image. As a fragile watermarking technique, our scheme can be used in tamper detection. Besides, it can be used in self-embedding for a large payload is provided. Meanwhile, security of the watermark is preserved due to the randomness of the discrete fractional random transform. Experimental results demonstrate the effectiveness of our scheme.

1. Introduction

Digital watermarking [1] plays an important role in multimedia information security. It is a process of embedding some secret data in the host media such as images, videos, audios, 3D meshes, etc. So far various digital image watermarking techniques have been presented in literatures. These schemes can be categorized according to different aspects as follows.

According to the watermark perceptibility, they can be divided into visible and invisible watermarking methods. Visible watermarking [2,3] can be used in applications such as copyright announcement and advertisement, while invisible watermarking [4–15] is used for covert communication, traitor tracing, etc.

According to the watermark robustness against attacks, they can be divided into robust, semi-fragile and fragile approaches. Robust watermarking [4,5,8,9,16], usually designed for copyright protection, can resist most intentional or unintentional attacks, e.g., rotation, scaling and translation. In contrast, fragile watermarks [11] are sensitive to any alterations on the watermarked image including common image operations (e.g., JPEG compression, low-pass filtering) and malicious attacks (e.g., cropping). They can be used for content authentication, tamper detection and localization, etc. Semi-fragile watermarking [15] can resist common image operations, while not robust to malicious attacks.

According to the watermark embedding domain, they can be divided into spatial domain, transform domain and compression domain based techniques. Generally speaking, the spatial domain based schemes [7,10,11] have lower computational complexity than the other two kinds of methods. For instance, the watermarking methods based on the least significant bitplane (LSB) modification [11,17] are such a kind of classical spatial domain based techniques. Watermarks in the transform domain based schemes [4–6,16] usually exhibit good robustness. The compression domain based techniques insert the watermark during compression or on compressed images. The associated compression techniques consist of JPEG [18], JPEG2000 [19], vector quantization (VQ) [12–14], block truncation coding (BTC) [20], etc.

According to whether the host image is required or not in watermark extraction, the available methods can be divided into blind, semi-blind and non-blind techniques. In the blind ones [6,7,9–11,21], the host image is not required. On the contrary, it must be provided in the non-blind methods [4,5]. In the semi-blind schemes [16], the host image is not required while some prior knowledge is usually supplied as auxiliary information for watermark extraction.

According to whether the host image can be perfectly recovered or not after watermark extracted, they can be divided into reversible and irreversible schemes. The host image can be perfectly recovered in the reversible methods [10], while cannot in those irreversible [4–7,9,11–16].

Generally a watermarking scheme may have two or more properties as mentioned above. For instance, the methods in [4,5] are

* Corresponding author at: MinZhuGuan, HuaJiaChi Campus, Zhejiang University, No. 268 KaiXuan Road, Hangzhou 310029, PR China. Tel.: +86 571 86971612; fax: +86 571 86971612.

E-mail addresses: luohao@zju.edu.cn (H. Luo), aeeizju@gmail.com (F.-X. Yu), hyland@zju.edu.cn (Z.-L. Huang), zheminglu@zju.edu.cn (Z.-M. Lu).

0030-4026/$ – see front matter © 2010 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ijleo.2009.12.018

not only robust but also transform domain based, while that in [22] is a reversible method in the VQ-compressed domain.

Nowadays, most existing transform domain watermarking schemes are based on discrete cosine transform (DCT) [6,23], discrete wavelet transform (DWT) [8], discrete Fourier transform (DFT) [16], etc. Recently, Guo et al. [4] propose a novel transform domain watermarking method based on discrete fractional random transform (DFRNT). The DFRNT [24] is derived from the discrete fractional Fourier transform (DFrFT). In Guo et al.'s scheme, the phase shift keying (PSK) is adopted to adjust the phase of complex coefficients in the DFRNT domain with reference to the watermark bit to be embedded. This scheme is robust against some familiar attacks including cropping, noising and low-pass filtering. In 2008, another DFRNT based watermarking scheme is proposed in [5]. It further demonstrates the method maintains higher robustness than those based on DCT, DFT and DFrFT.

These DFRNT based watermarking methods have two common characteristics, i.e., both of them belong to robust and non-blind techniques. However, in many scenarios, the host image cannot be released after watermark embedded. In those cases, the non-blind DFRNT based methods are no longer appropriate. In addition, as aforementioned, fragile watermarking can be widely used in tamper detection and self-embedding [7]. Motivated by this, we propose a blind and fragile watermarking scheme based on DFRNT. In it, the watermark can be accurately extracted if the watermarked image is intact, without the host image aided. Moreover, its utilization in tamper detection, localization and recovery is also demonstrated.

Our method starts from subsampling the host image into four equal-sized subimages. Then each subimage is transformed into DFRNT domain coefficients. It is easy to understand that high correlation among the spatial subimages' pixels still exists more or less in their corresponding DFRNT coefficients. Thus, watermark can be embedded and extracted with respect to the correlation of two subimages' coefficients. A high visual quality watermarked image can be acquired even if much watermark data is hidden. As long as the watermarked image suffers alterations, the coefficient correlation is destroyed. Therefore, our scheme can be used in tamper detection and localization. In particular, it can be easily extended for self-embedding. That is, if the host image content is used as a watermark, the tampered area can be recovered to some extent.

The rest part of this paper is organized as follows. Section 2 briefly reviews the DFRNT transform and two non-blind watermarking methods based on it. Section 3 extensively describes the watermark embedding and extraction procedures of our scheme. Experimental results are shown in Sections 4 and 5 concludes the paper.

2. Related work

2.1. DFRNT transform

In [24], Liu et al. introduced the transform named DFRNT for the first time. It can be used for one or two-dimensional discrete signal analysis. Specially, it is effective for image encryption [24], secret sharing [25] and watermarking [4,5]. The DFRNT of a two-dimensional image I can be represented as

$$C = R^{\alpha} I (R^{\alpha})^{\mathrm{T}} \tag{1}$$

where R^{α} and α denote the kernel transform matrix and the fractional order of DFRNT, respectively. The superscript T means matrix transposition. The output matrix C is the transform coefficients in DFRNT domain. To construct a transform matrix R^{α}, we must generate a random matrix P in advance as

$$P = \frac{Q + Q^{\mathrm{T}}}{2} \tag{2}$$

where Q is a real nonsingular matrix with elements randomly generated. Suppose $(V = v_1, v_2, \ldots, v_n)$ denotes the normalized eigenvector matrix of P. Obviously, any two eigenvectors (two columns of V) are orthonormal because P is a real symmetric matrix. Then we can obtain R^{α} as

$$R^{\alpha} = V D^{\alpha} V^{\mathrm{T}} \tag{3}$$

where D^{α} is a diagonal matrix defined as

$$D^{\alpha} = diag\left[1,\ \exp\left(-i\frac{2\pi\alpha}{t}\right),\ \ldots,\ \exp\left(-i\frac{2(N-1)\pi\alpha}{t}\right)\right] \tag{4}$$

where t denotes the periodicity of DFRNT.

Clearly, DFRNT is a random transform due to the randomness of R^{α}. In other words, different P or Q corresponds to different R^{α}, and further different DFRNT. Specifically, when $\alpha = rt/2$ (r is a constant integer), the output of DFRNT is real for a real signal [24].

2.2. Non-blind watermarking based on DFRNT

The two watermarking schemes based on DFRNT are briefly reviewed. In [4], the host image is firstly transformed into DFRNT coefficients. Then this coefficient matrix is partitioned into a set of non-overlapping blocks. Next, in each block several coefficients with the largest amplitudes are selected for watermark embedding. The watermark is hidden by PSK, i.e., slightly changing the phase of selected coefficient according to the watermark bit. At last, the watermarked coefficients are transformed into the watermarked image via inverse DFRNT. The watermark extraction is the inverse process of watermark embedding by comparing the DFRNT coefficients' phases of the host image with those of the watermarked version.

Different from the first one, the second scheme [5] exploits the fact that when α is set as half periodicity, the output of DFRNT remains real for a real signal [24]. It starts from transforming the host image and the watermark image into coefficients in DFRNT domain, respectively. Then similar as the scheme in [4], some coefficients of the host image with the largest amplitudes are selected. The watermark embedding is achieved by slightly changing the amplitudes of these selected coefficients according to the watermark image's coefficients. At last, the watermarked coefficients are inverse DFRNT transformed into the watermarked image. The watermark extraction is also a comparison mechanism similar as that in the first scheme [4], with the host image needed.

In these two DFRNT based schemes, the selected coefficients of the host image for watermark embedding are with largest amplitudes. This is because small variations on larger coefficients are more likely to result in smaller distortions of the host image introduced by watermark embedding. In this way, the imperceptibility of the watermark is better preserved. In addition, as the host image must be provided during watermark extraction, they belong to non-blind watermarking techniques.

3. Proposed scheme

The main difference between our scheme and the reported two methods lies in the host image is not required in watermark extraction. That is, some substitute reference information must be provided for watermark extraction. Actually, much information redundancy exists in natural image pixels, especially in small local blocks. Thus, we can exploit this high correlation among them as reference information, namely the principle of self-reference. This