Journal of King Saud University – Computer and Information Sciences (2016) **28**, 199–210

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

A blind video watermarking scheme resistant to rotation and collusion attacks

Amlan Karmakar [a], Amit Phadikar [a,*], Baisakhi Sur Phadikar [b], Goutam Kr. Maity [c]

[a] *Department of Information Technology, MCKV Institute of Engineering, Liluah, Howrah 711204, India*
[b] *Department of Computer Science and Engineering, MCKV Institute of Engineering, Liluah, Howrah 711204, India*
[c] *Dept. of Electronics and Communication Engineering, MCKV Institute of Engineering, Liluah, Howrah 711204, India*

Received 6 November 2013; revised 7 May 2014; accepted 4 June 2014
Available online 2 November 2015

KEYWORDS

Video watermarking;
DCT;
Complex Zernike moments;
Rotation attack;
Collusion attacks;
Rayleigh fading

Abstract In this paper, Discrete Cosine Transform (DCT) based blind video watermarking algorithm is proposed, which is perceptually invisible and robust against rotation and collusion attacks. To make the scheme resistant against rotation, watermark is embedded within the square blocks, placed on the middle position of every luminance channel. Then Zernike moments of those square blocks are calculated. The rotation invariance property of the Complex Zernike moments is exploited to predict the rotation angle of the video at the time of extraction of watermark bits. To make the scheme robust against collusion, design of the scheme is done in such a way that the embedding blocks will vary for the successive frames of the video. A Pseudo Random Number (PRN) generator and a permutation vector are used to achieve the goal. The experimental results show that the scheme is robust against conventional video attacks, rotation attack and collusion attacks.

1. Introduction

Digital watermarking is a well-established technique to protect the Intellectual property and digital copyright of the multimedia information such as image, audio or video. It also secures the integrity of digital contents at the time of transferring those through the internet.

In recent times, Moving Picture Experts Group (MPEG) video standard has widespread applications in internet streaming, digital High-definition (HD) handy-cams as well as in mobile phones. So many researchers are working in the field of digital video watermarking (Al-Taweel et al., 2010; Guo-juan and Rang-ding, 2009; Chao et al., 2008). Many video-watermarking algorithms were proposed by many researchers both in the spatial (Al-Taweel et al., 2010) or in the temporal domain (Chao et al., 2008) as well as in the transform or frequency domain (Chao et al., 2008; Al-Taweel et al., 2009; Jing, 2009; Hartung and Girod, 1998) over the last few years.

At the time of developing a video-watermarking algorithm, the researchers should concentrate on the two most important things, i.e., imperceptibility and robustness (Phadikar, 2013). Imperceptibility refers to perceptual transparency, i.e., the

* Corresponding author. Tel.: +91 (33) 2654 9315; fax: +91 (33) 2654 9318.
E-mail addresses: amlan.karma@gmail.com (A. Karmakar), amitphadikar@rediffmail.com (A. Phadikar), baisakhi.sur@gmail.com (B.S. Phadikar), goutam123_2005@yahoo.co.in (G.Kr. Maity).
Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksuci.2014.06.019
1319-1578

algorithm embeds the watermark in video frames in a fashion that the quality of the video frames is not perceptually affected. Robustness refers to the ability of the watermark extraction process to extract the watermark successfully from the video frames, even if the quality of the frames is degraded by various types of intentional or unintentional attacks.

In most of the video-watermarking schemes it is seen that the discussion on robustness is mainly focused on temporal attacks like frame dropping, frame inserting, frame rate changes, etc. (Chao et al., 2008; Al-Taweel et al., 2009; Hartung and Girod, 1998; Liang, 2009). However, very few number of video watermarking schemes are robust against geometrical attacks (Al-Taweel et al., 2010; Guo-juan and Rang-ding, 2009; Jing, 2009; Wei et al., 2010; Bahrushin et al., 2009; Wang et al., 2008; Xu et al., 2008) especially rotation of video frames in any random angle. The rotation attack destroys the synchronization of watermark extraction as it changes the pixel positions of the watermarked video frame. So there is a great challenge for the researches to develop an algorithm which is robust against rotation attack. Zernike moments have a rotation invariance property. One can find out the angle of rotation from the phase information of the Zernike moments. So it is widely used in the field of image as well as video watermarking.

Besides temporal and geometrical attacks, another type of attack for watermarked video is collusion attack (Su et al., 2002). In some situations, it is possible for an attacker to obtain multiple watermarked data. The attacker can often exploit this situation to remove watermarks without knowing the watermarking algorithm. This kind of attack is known as collusion attack. The collusion attack is a different kind of attack and very few numbers of video watermarking schemes can resist collusion attack (Kanócz et al., 2009; Saxena and Gupta, 2007). Moreover, in the literature, very few works are found so far which address both rotation and collusion attacks in a single platform.

In this paper, a blind video watermarking algorithm is proposed in the DCT domain. The scheme embeds watermark before MPEG-4 encoding and extracts the watermark after decoding. So, the scheme can be applied to both uncompressed video or to any compression method (before and after compression). In the present scheme, the watermark information is embedded to every frame. It offers two-fold advantages, namely (1): it increases robustness against frame dropping and frame insertion, (2): one can identify individual frames so that integrity of the video is checked. The Complex Zernike moment is used to make the scheme robust against rotation attack. At the same time, the design of the watermark embedding algorithm is made in such a way that robustness is achieved against collusion attacks, including a Rayleigh fading wireless channel.

The rest of the paper is organized as follows: Section 2 describes the details of Discrete Cosine Transformation (DCT). Section 3 focuses on Complex Zernike moments. Section 4 discusses Type-1 and Type-2 collusion attacks. Section 5 highlights the related works already done in this area. Section 6 explores the algorithm for embedding and extracting watermark. Section 7 presents performance evaluation of the proposed scheme and Section 8 describes conclusions and the scope of future work.

2. Discrete Cosine Transformation (DCT)

Transform coding is a valuable component of contemporary image and video processing applications. Like other transforms, Discrete Cosine Transform (DCT) attempts to decorrelate the image data. It has an excellent energy compaction property for highly correlated images. DCT can be expressed in separable format and exhibits decrease in the entropy of an image. After de-correlation, each transform coefficient can be encoded independently without losing compression efficiency (Khayam, 2003).

The one dimension DCT is defined as:

$$C(u) = \alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{\pi(2x+1)u}{2N}\right] \quad (1)$$

For $u = 0, 1, 2 \ldots N-1$.

The inverse transformation is defined as:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)\cos\left[\frac{\pi(2x+1)u}{2N}\right] \quad (2)$$

For $x = 0, 1, 2 \ldots N-1$.

where,

$$\alpha(u) = \sqrt{\frac{1}{N}},\ \text{For } u = 0 \text{ and } \alpha(u) = \sqrt{\frac{2}{N}} \text{ For } u \neq 0 \quad (3)$$

From Eq. (1) it is understood that for $u = 0$,

$$C(u=0) = \sqrt{1/N}\sum_{x=0}^{N-1} f(x)$$

Thus, the first transform coefficient is the average value of the sample sequence. In literature, this value is referred to as DC coefficient. All other transform coefficients are called AC coefficients. However, in image or video processing applications 2-D DCT is used instead of 1-D DCT and it is the extension of Eq. (1).

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \times \cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (4)$$

For $u, v = 0, 1, 2 \ldots N-1$ and $\alpha(u)$ and $\alpha(v)$ are defined in Eq. (3). The inverse transformation is defined as:

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (5)$$

For $x, y = 0, 1, 2 \ldots N-1$.

In case of video, the temporal redundancy is to be exploited to provide a better compression. Hence, adjacent pixels in consecutive frames show very high correlation. Consequently, this correlations can be exploited to predict the value of a pixel from its respective neighbors. Moreover, most of the image and video data are still available in DCT compressed form. So the scheme is more suitable for real time implementation.

3. Complex Zernike moments

Complex Zernike moments (Guo-juan and Rang-ding, 2009) are constructed using a set of complex polynomials which forms a complete orthogonal set over unit disk of $(x^2 + y^2) \leqslant 1$. The set of such polynomials can be defined as:

$$A_{mn}(x,y) = A_{mn}(r,\theta) = R_{mn}(r)e^{in\theta} \quad (6)$$