Int. J. Electron. Commun. (AEÜ) 69 (2015) 1868–1879

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

A blind spatio-temporal data hiding for video ownership verification in frequency domain

Majid Masoumi [a,*], Mahsa Rezaei [b], A. Ben Hamza [a]

[a] Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada
[b] Electrical Engineering Department, Azad University South Tehran Branch, Tehran, Iran

ARTICLE INFO

Article history:
Received 31 May 2015
Accepted 24 September 2015

Keywords:
Video processing
Visual quality assessment
Code division multiple access
Lossy compression
Pseudo random number
Human visual system

ABSTRACT

By widespread usage of video-based applications, digital watermarking has been recommended as a solution to prevent illegal and malicious copying and distribution of digital media. In this paper, a blind watermarking system for video ownership verification is proposed. Inspired by the used code division multiple access (CDMA) techniques in image watermarking with 2D form, and considering that video is inherently three dimensional, we expand CDMA techniques in 3D form based on discrete wavelet transform for watermarking of video data. First, a 3D-DWT is performed on the video frames. Then, by using CDMA techniques and two different pseudo random number (PRN) sequences, a binary image (as a watermark) is spread into mid and high frequency sub-bands of wavelet. In extraction step, the original video is not needed, namely, blind detection. Eventually, by defining a heuristic algorithm, the watermark is clearly detectable. Unlike other schemes, which are resistant to a limited series of attacks, the proposed method has admirable robustness against a large range of attacks such as frame averaging, frame dropping, frame transposing, median and low-pass filtering, Gaussian noise, salt and pepper noise, speckle noise, resizing and a selected group of lossy compressions like MJPEG, MPEG and H.264/AVC. In addition, evaluation of well-known visual quality assessment (VQA) metrics like PSNR, MSE, SSIM, MS-SSIM and MOVIE shows the high transparency of the proposed system for human visual system (HVS).

1. Introduction

Nowadays, by propagating the usage of the Internet and multimedia systems, it is easier for digital data owners to transfer multimedia documents across the Internet. Copying or altering these contents by unauthorized people leads to serious copyright violation problems [1]. Therefore, there is an increase in concern over copyright protection of digital contents [2–5]. Traditionally, encryption and control access techniques were employed to protect the ownership of media. These techniques, however, do not protect against unauthorized copying after the media have been successfully transmitted and decrypted. Recently, the watermark techniques are utilized to maintain the copyright [5–8].

Digital watermarking is an effective way to protect copyright of multimedia data even after its transmission. Watermarking is a concept of embedding a special pattern, namely watermark into a multimedia document so that a given piece of copyright information is constantly tied to the data. This information can later prove the ownership, identify an unauthorized person and simply inform the users about the rights-holder or the permitted use of the data [7].

The watermarking techniques can be classified into several groups such as text-based watermarking [9,10], image watermarking [10], video watermarking [11], audio watermarking [12] and 3D watermarking [13]. Since almost 90% of the contents are being transmitted in image and video data [14,15], a number of techniques have been developed for these two groups. Applications of video watermarking contain fingerprinting, copyright protection, broadcast monitoring and video authentication.

The watermarking techniques can be divided into two main categories according to the embedding domain of the cover image: the spatial domain methods and the transform domain methods. The spatial domain methods are the earliest and simplest watermarking techniques but they have low capacity for hiding information, and the watermark can be easily erased by lossy compression. On the other hand, the transform domain approaches insert the watermark into the transform coefficients of the original media, yielding

* Corresponding author. Tel.: +1 5148036486.
E-mail addresses: ma_masou@encs.concordia.ca (M. Masoumi), mahsa.rezaei@yahoo.com (M. Rezaei), hamza@ciise.concordia.ca (A.B. Hamza).

http://dx.doi.org/10.1016/j.aeue.2015.09.015
1434-8411/

more information embedding and more robustness against watermarking attacks. Recent popular transforms contain the discrete cosine transform (DCT) [5], the discrete wavelet transform (DWT) [6], and the discrete Fourier transform (DFT).

Despite the similarity between the image and video watermarking techniques, the video watermarking schemes deals with more challenges like the large volume of inherently redundant data between video frames, unbalance between motion and motionless regions [16], and the real-time requirements in video broadcasting that make the video signals highly susceptible to pirate attacks including frame swapping, frame averaging, frame dropping, and statistical analysis [17].

Basic requirements of digital video watermarking are: invisibility, robustness, and capacity.

Invisibility: a degree that an embedded watermark remains unnoticeable when a user views the watermarked digital media. The information should embed the watermark in the regions of the video frame in which imperceptibility is least affected [18].

Robustness: the resilience of an embedded watermark against being removed by incidental and intended attacks. Major attacks on the videos are filtering, adding noise, compression, scaling, frame dropping, frames averaging and frame swapping [18]. In some of the watermarking techniques, robustness is more a property and not a requirement [19].

Capacity: the amount of information that can reliably be hidden when the scheme provides the ability to change digital data [20–23].

These three requirements of watermarking make a triangle. Improvement in any one of them, affects the other two negatively. We have to find the correct balance between these conflicting requirements of watermarking [18].

1.1. Related works

The early usage of spread spectrum techniques in watermarking goes back to the watermarking of images. The primary examples of applying spread spectrum in image watermarking belongs to Van Schynold et al. [24] who perturbed the LSB of each pixel by random amounts produced by an *m*-sequence generator. Cox et al. [25] definition of spread spectrum was rather different. They applied perturbation to the first 1000 largest DCT coefficients of the entire image. Ruanaidh and Pun [26] presented a watermarking algorithm that spread spectrum the watermark message by an m-sequence based on the direct sequence code division multiple access (DS-CDMA), after that the CDMA encoded message are embedded into images in the DCT domain. Silvestre and Dowling [27] performed the DFT to the original images and formed many independent bands by choosing selected DFT coefficients. Then the watermark message was spread according to the CDMA coding by two orthogonal sequences and embedded in the independent bands. Kohda et al. [28] converted images from RGB to the YIQ domain, and inserted the message in the independent CDMA channels that are formed by the first 15 DCT coefficients of Y, the first 6 DCT coefficients of I and the first 3 DCT coefficients of Q. A RW/CDMA technique was proposed in [29] which maintained the copyright for medical images.

In recent years, watermarking methods are operating in the 3D domains. A watermarking algorithm based on 3D discrete wavelet transform and video scene segmentation was proposed by Zhuang et al. [30], where the watermark is a binary logo which is disordered, and then is inserted into the 3D wavelet coefficients of a selected video scenes. In the extraction step, however, this scheme is not completely blind because it requires the disordered watermark generated during the embedding process to detect the embedded watermark. The DFT of 3D chunks of a video scene was used in [31] for video watermarking, where the embedding and the extraction algorithms are applied to uncompressed video data. A DWT-based watermarking scheme which embeds the watermark in successive frames was proposed in [32]; these frames are selected based on scene change detection. However, it will not be useful if the video scene changes occur rapidly or video sequence contains too many different short scenes. In addition, if the video contains long motionless scenes, the algorithm will face some difficulties in embedding process; because if the frame numbers are not a factor of 2^n, our coefficients will have zero values in executing the inverse DWT. Furthermore, the system will be damaged when faces to temporal attacks like frame dropping, frame transposing and frame averaging. In fact, motionless regions may be statistically compared or averaged to remove watermark. In addition, video watermarking schemes must take into account inter-frame information. Li [33] proposed a scheme based on 3D-DWT and artificial neural network (ANN) in which the watermark is adaptively embedded in the wavelet sub-bands using the relationship among their neighbors and statistical features and. The ANN constructs the relationship in the embedding phase and is saved as the ANN connection weights. The statistical features of coefficients and their relationship are used to extract the watermark in detection step. Therefore, this scheme is not completely blind; on the contrary, a large amount of data is required to detect watermark sequence. Da-Wen [34] proposed a blind watermarking method which embeds the CDMA encoded watermark into the selected sub-block of wavelet coefficients. Each sub-block is $64 \times 64 \times 64$ pixels; anyway restricting the insertion of watermark for some sub-blocks reduces the payload for embedding the watermark information.

Among the delivered techniques in recent years, DWT-based techniques achieved more popularity because of their nice multi-resolution characteristics, frequency spread and also spatial localization. We chose DWT due to it is more computationally efficient than the other transforms. Also, its speed is faster than DCT and DFT as only sum or difference of the pixel have to be calculated.

In this study, a video watermarking scheme which shows the robustness against video attacks and yet enables blind retrieval of the watermark is proposed. Inspired by used CDMA techniques in image watermarking in 2D form, a new CDMA technique for video watermarking in 3D form is proposed. The novelty of our watermarking process contains expanding the CDMA embedding procedure in 3D form based on wavelet domain for video data, and also defining a heuristic algorithm for blind retrieval of the watermark in extraction step. Furthermore, the proposed scheme is robust against format conversions because the watermark is inserted before compression. Otherwise the authentication information will be lost if the video file is converted to a different compression standard.

The rest of the paper is organized as follows: Section 2 gives information about 3D-DWT. Section 3 briefly describes CDMA spread spectrum techniques and its relation with watermarking. In Section 4, the proposed algorithm is explained which includes the embedding and detection process of the watermark. Results and analysis are discussed in Section 5. Finally, Section 6 concludes the delivered video watermarking scheme.

2. Discrete wavelet transform

The wavelet transform is a valuable tool for multi-resolution analysis that has been widely used in image processing applications. The wavelet transform has a number of advantages over other transforms as it provides a multi-resolution description, it allows superior modeling of the human visual system (HVS), the high resolution sub-bands allow easy detection of features such as edges or textured areas in transform domain.