Content Protection in Video Data Based on Robust Digital Watermarking Resistant to Intentional and Unintentional Attacks

Majid Masoumi[1] and Shervin Amiri[2]
[1]Department of Electrical Engineering, Islamic Azad University Qazvin Branch, Iran
[2]Scientific Member of electrical engineering Department, Iranian Research Organization for Science and Technology, Iran

Abstract: *Embedding a digital watermark into an electronic document is proving to be a feasible solution for multimedia copyright protection and authentication purposes. In the present paper we propose a new digital video watermarking scheme based on scene change analysis. By detecting the motion scene of video and using CDMA techniques the watermark is embedded into mid-frequency sub-bands of wavelet coefficients. In this experiment in order to enhance the security of our algorithm four keys are considered. Three of them are formed in watermark encryption process and one key is related to CDMA embedding process. Also, with the aim of making a good compatibility between the proposed scheme and Human Visual System (HVS), the blue channel of RGB video is utilized to embed the watermark. Experimental results show the high robustness of the proposed method against both intentional and unintentional attacks during the transfer of video data. The implemented attacks are Gaussian noise, median filtering, frame averaging, frame dropping, geometric attacks and different kinds of lossy compressions including MPEG-2, MPEG-4, MJPEG and H.264/AVC.*

Keywords: *Digital watermarking, scene change analysis, geometric attacks, information security, HVS.*

Received May 13, 2012; accepted December 30, 2012

1. Introduction

In the past decade, the global rife access of internet technologies makes the communication and circulation of digital multimedia contents like images, audio and video very easy. However, this convenience also causes substantial increase in illegal operations such as duplication, modification, forgery, copy-right protection and others in digital media. Therefore, the protection of digital media has become an imperative issue. Recently, digital watermarking has drawn much attention of research community to resolve these pressing problems [3, 26].

The basic procedure for digital watermarking is to embed some hidden information into multimedia data, while the quality of the watermarked data is retained, and the watermark can still be detected under different kinds of intentional and unintentional attacks. An effective watermarking scheme is distinguished by three characteristics, namely, imperceptibility, robustness and capacity. Also, in order to enhance the safety of algorithm another feature, i.e., security, can be added to the main necessities of video watermarking.

- *Imperceptibility:* 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 [13].
- *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 [13]. In some of the watermarking techniques, robustness is more a property and not a requirement [1].
- *Capacity:* The amount of information that can reliably be hidden when the scheme provides the ability to change digital data [18, 10, 11, 29]
  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 [13].
- *Security:* The ability of the watermark to resist against attempts by a sophisticated attacker to remove it or destroy it via cryptanalysis, without modifying the video itself [6].

Applications of video watermarking contain fingerprinting, broadcast monitoring, video authentication and copyright protection. A variety of watermarking algorithms have been proposed in the literature. These algorithms can be broadly classified in two categories according to the embedding domain: spatial and transform domain. Spatial domain approaches [12, 22] are the simplest and the earliest algorithms based on the modification of pixel intensities. These algorithms are less robust against the

attacks. On the other hand, transform domain approaches insert the watermark into transform coefficients, such as Discrete Fourier Transform (DFT) [19], Discrete Cosine Transform (DCT) [2], and Discrete Wavelet Transform (DWT) [23, 25, 30].

Li [15] proposed a scheme based on 3D-DWT and Artificial Neural Network (ANN). First the average frame of the video shot is computed. Then the resulting frame is transformed to the wavelet domain. Afterwards, the LL subband is divided into non overlapping 3×3 blocks. The blocks which satisfy a relationship among the center of block, mean eight neighbors and standard deviation are selected as the training sets. The center of each block is the output while the mean and standard deviation of the neighbor are the inputs. Consequently a trained network for watermark embedding and extraction process is obtained. So in the extraction stage, statistical features of coefficients and their relationship are used which means, this scheme is not completely blind; on the contrary, a large amount of data is required to detect the watermark sequence. On the other hand embedding the watermark information in low frequency coefficient enhances the robustness of system; however it severely destroys the quality of video and its transparency.

A DWT-based watermarking scheme which embeds the watermark in successive frames is proposed in [4]; this scheme defined a general algorithm based on scene change detection for embedding the watermark by altering the magnitude of some wavelet coefficients. In fact they insert the watermark in both motion and motionless parts of video. However, HVS is more sensitive to motionless part of video, so embedding the watermark in this part leads to the perceptibility of the watermark. On the other hand compression algorithms used to destroy the motionless part of video meanwhile maintain the motion part of it. Also, these motionless regions may be statistically compared or averaged to remove the independent watermarks. From the other point of view, dividing the video based on scene-change detection will not be useful if the video changes occur rapidly or contains too many different short scenes, because it demands different algorithms in each scene for embedding and detecting the watermark in the watermarking process.

Furthermore they had no estimation or discussion about the acquiring of frame number based on a factor of 2^n. Since each algorithm which has a 3D shape (e.g., 3D-DWT, 3D-DCT,…) should has a frame number of factor 2^n for embedding and extraction of watermark. Otherwise it faces problems in executing the inverse DWT, because if the number of frames is not a factor of 2^n the coefficients will have zero values which lead to major damage in video quality and destruction of watermark information.

Among the delivered techniques in recent years, the ones which are based on the Discrete Wavelet Transform (DWT) are gaining more popularity due to their excellent spatial localization, frequency spread, and multi-resolution characteristics. The DWT is chosen because it is more computationally efficient than other transform methods. The speed is faster than DCT and DFT as only sum or difference of the pixel have to be calculated.

From the other point of view, watermarking in video domain can be done on compressed or uncompressed video. Generally the watermarking in compressed video has more compatibility for real-time systems [14]. However, the goal of this paper is to embed the watermark in uncompressed video, because format conversions destroy the watermark which is embedded in compressed watermarked video. The huge memory requirement of the algorithms that compute the 3D-DWT is one of the main drawbacks in practical implementations. Originally the algorithms which are based on 3D-DWT and 3D-DCT,… demand high computational complexity and have enormous memory usage. Also, if the 3D algorithms perform on all video streams (containing motion and motionless regions) they require higher computational complexity and are expensive for implementation and eventually cannot be useful for real-time systems. The video watermarking methods which have lower computational complexity (e.g. 2D based algorithms) are suitable for real-time systems.

Many of the existent algorithms meet the imperceptibility necessity quite easily and survey robustness only against a subset of attacks not all kinds of attacks, meanwhile robustness against different attacks is the key challenge in watermarking process.

In this study, a secure video watermarking scheme which is robust against different video attacks and yet is suitable for real-time systems is delivered. The rest of paper is planned as follows: The description of used terminologies i.e. Discrete Wavelet Transform and watermark scrambling are explained in Section 2. Section 3 illustrates the proposed watermarking scheme in three sub-sections. The experimental results are presented in Section 4. Finally, the concluding remarks are given in Section 5.

2. Preliminaries

In this section, we provide the main terminologies which are used in the proposed algorithm to achieve the desired goal. These terminologies are as follows:

2.1. Discrete Wavelet Transform

2.1.1. One Dimensional Discrete Wavelet Transform

A general 1-D discrete wavelet transform can be written as [9]:

$$W(j,k)=\frac{1}{\sqrt{M}}\sum_{x} f(x)2^{\frac{j}{2}}\psi(2^{j}x-k) \quad (1)$$

$$\psi=\begin{cases}1 & 0\le x\le 0.5\\ -1 & 0.5\le x\le 1\\ 0 & \text{otherwise.}\end{cases} \quad (2)$$

Where W represents the wavelet coefficients function, j and k denote the dilation and translation parameters respectively, and M is the length of sequence f.

2.1.2. Two Dimensional Discrete Wavelet Transform

The one-dimensional wavelet transform can be easily extended to two-dimensional functions like images. In two dimensions, a two-dimensional scaling function, $\varphi(x,y)$ and three, two-dimensional wavelets, $\psi^{H}(x,y)$, $\psi^{V}(x,y)$, $\psi^{D}(x,y)$ are required. Each is the product of a one-dimensional scaling function φ and corresponding wavelet ψ. Excluding products that produce one-dimensional results, like $\varphi(x)\psi(x)$, the four remaining products produce the separable scaling function:

$$\varphi(x,y)=\varphi(x)\varphi(y) \quad (3)$$

and separable, "directionally sensitive" wavelets :

$$\psi^{H}(x,y)=\psi(x)\varphi(y) \quad (4)$$

$$\psi^{V}(x,y)=\varphi(x)\psi(y) \quad (5)$$

$$\psi^{D}(x,y)=\psi(x)\psi(y) \quad (6)$$

Given separable two-dimensional scaling and wavelet functions, extension of the one-dimensional DWT to two dimensions is straightforward. The scaled and translated basis functions are:

$$\varphi_{j,m,n}(x,y)=2^{j/2}\varphi(2^{j}x-m,2^{j}y-n) \quad (7)$$

$$\psi^{i}_{j,m,n}(x,y)=2^{j/2}\psi(2^{j}x-m,2^{j}y-n) \quad \text{i=\{H,V,D\}} \quad (8)$$

Where index i identifies the directional wavelets in equations (4) to (6). The discrete wavelet transform of function $f(x,y)$ of size $M\times N$ is then:

$$W_{\varphi}(j_0,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\varphi_{j_0,m,n}(x_0) \quad (9)$$

$$W^{i}_{\psi}(j,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\psi^{i}_{j,m,n}(x_0) \quad \text{i=\{H,V,D\}} \quad (10)$$

As in the one-dimensional case, j_0 is an arbitrary starting scale and the $W_{\varphi}(j_0,m,n)$ coefficients define an approximation of $f(x,y)$ at scale j_0. The $W^{i}_{\psi}(j,m,n)$ coefficients add horizontal, vertical, and diagonal details for scales $j\ge j_0$. Given the W_{φ} and W^{i}_{ψ} of equations. (9) and (10), $f(x,y)$ is obtained via the inverse discrete wavelet transform:

$$f(x,y)=\frac{1}{\sqrt{MN}}\sum_{m}\sum_{n} W_{\varphi}(j_0,m,n)\varphi_{j_0,m,n}(x,y)+\frac{1}{\sqrt{MN}}\sum_{i=H,V,D}\sum_{j=j_0}^{\infty}\sum_{m}\sum_{n} W^{i}_{\psi}(j,m,n)\psi^{i}_{j,m,n}(x,y) \quad (11)$$

2.1.3 Three Dimensional Discrete Wavelet Transform

In three dimension, a three-dimensional scaling function, $\varphi(x,y,z)$ and seven three-dimensional wavelet functions, $\psi^{LLH}(x,y,z)$, $\psi^{LHL}(x,y,z)$, $\psi^{LHH}(x,y,z)$, $\psi^{HLL}(x,y,z)$, $\psi^{HLH}(x,y,z)$, $\psi^{HHL}(x,y,z)$ and $\psi^{HHH}(x,y,z)$ are required. Each is the product of a one-dimensional scaling function φ and corresponding wavelet ψ .Excluding products that produce one-dimensional results, like $\varphi(x,y,z)\,\psi(x,y,z)$, the eight residual products produce the separable scaling function:

$$\varphi(x,y,z)=\varphi(x)\varphi(y)\varphi(z) \quad (12)$$

and separable, "directionally sensitive" wavelets :

$$\begin{aligned}\psi^{LLH}(x,y,z)&=\varphi(x)\varphi(y)\psi(z)\\ \psi^{LHL}(x,y,z)&=\varphi(x)\psi(y)\varphi(z)\\ \psi^{LHH}(x,y,z)&=\varphi(x)\psi(y)\psi(z)\\ \psi^{HLL}(x,y,z)&=\psi(y)\varphi(y)\varphi(z)\\ \psi^{HLH}(x,y,z)&=\psi(x)\varphi(y)\psi(z)\\ \psi^{HHL}(x,y,z)&=\psi(x)\psi(y)\varphi(z)\\ \psi^{HHH}(x,y,z)&=\psi(x)\psi(y)\psi(z)\end{aligned} \quad (13)$$

Given separable three-dimensional scaling and wavelet functions, extension of the one-dimensional DWT to three dimensions is simple. The scaled and translated basis functions are:

$$\varphi_{j,m,n,o}(x,y,z)=2^{j/2}\varphi(2^{j}x-m,2^{j}y-n,2^{j}z-o)$$
$$\psi^{i}_{j,m,n,o}(x,y,z)=2^{j/2}\psi(2^{j}x-m,2^{j}y-n,2^{j}z-o) \quad (14)$$
$$\text{i=\{LLH,LHL,LHH,HLL,HLH,HHL,HHH\}}$$

Where index i identifies the directional wavelets in equation (13). The discrete wavelet transform of function $f(x,y,z)$ of size $M\times N\times O$ is then:

$$W_{\varphi}(j_0,m,n,o)=\frac{1}{\sqrt{MNO}}\sum_{z=0}^{O-1}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y,z)\varphi_{j_0,m,n,o}(x,y,z) \quad (15)$$

$$W_{\psi}^{i}(j,m,n,o)=\frac{1}{\sqrt{MNO}}\sum_{z=0}^{O-1}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y,z)\psi_{j,m,n,o}^{i}(x,y,z) \quad (16)$$

i={LLH,LHL,LHH,HLL,HLH,HHL,HHH}

As in the one-dimensional case, j_0 is an arbitrary starting scale and the $W_{\varphi}(j_0,m,n,o)$ coefficients define an approximation of $f(x,y,z)$ at scale j_0. The $W_{\psi}^{i}(j,m,n,o)$ coefficients add horizontal, vertical, and diagonal details for scales $j \geq j_0$. Given the W_{φ} and W_{ψ}^{i} of equations (15) and (16), $f(x,y,z)$ is obtained via the inverse discrete wavelet transform:

$$f(x,y)=\frac{1}{\sqrt{MNO}}\sum_{o}\sum_{m}\sum_{n}W_{\varphi}(j_0,m,n,o)\varphi_{j_0,m,n,o}(x,y,z)+\frac{1}{\sqrt{MNO}}\sum_{i=LLH}^{HHH}\sum_{j=j_0}^{\infty}\sum_{o}\sum_{m}\sum_{n}W_{\psi}^{i}(j,m,n,o)\psi_{j,m,n,o}^{i}(x,y,z) \quad (17)$$

2.1. Watermark Scrambling

Usually scrambling transform is used in the pretreatment stage of the watermark as a way of encryption. Generally, a meaningful watermark image becomes meaningless and disordered after scrambling. For improving the security and confidentiality, a scrambling method [27] is used to encrypt the binary watermark image. After scrambling, human eyes cannot distinguish the shape of the original watermark. Without the scrambling algorithm and the key, the attacker will not recover the watermark at all even if it has been extracted from the watermarked video. So shuffling the image gives a secondary security for the digital products. This method is defined as follows:

$$\begin{bmatrix}x_{n+1}\\ y_{n+1}\end{bmatrix}=\begin{bmatrix}1 & \alpha\\ \beta & \alpha\beta+1\end{bmatrix}\begin{bmatrix}x_n\\ y_n\end{bmatrix}\bmod(N) \quad (18)$$

Where (x_n, y_n) is the pixels position in an $N \times N$ image; (x_{n+1}, y_{n+1}) is the transformed position after cat map; α and β are the system parameters and must be the positive integers. The determinant value is 1, so cat map is a map which keeping area (no attractor). At the same time, the cat map is one-to-one mapping; each point in matrix can be transformed to another point uniquely. Cat map has two typical factors, which bring chaotic movement: tension (multiply matrix in order to enlarge *x*, *y*) and fold (taking mod in order to bring *x*, *y* in unit matrix). In fact, cat map is a chaotic map. Image position can be scrambled via the iteration of cat map, by considering $(\alpha=7, \beta=3)$, consequently the encrypted image will be achieved. After a period of iteration the watermark image will be attained. So iteration times and system parameters (α, β) can be used as the encryption keys. By using these three keys the security of our scheme will be enhanced.

(a) (b)

Figure 1. Encrypting the watermark; (a) Original watermark (b) Scrambled watermark (Iteration times= 9).

3. The Proposed Watermarking Algorithm

3.1. Preprocessing of Video

Applying independent watermarks to each frame presents a problem, if regions in each video frames remain little or no motion frame after frame. These motionless regions may be statistically compared or averaged to remove the independent watermarks [24]. From the other point of view, dividing the video based on scene-change detection will not be useful if the video changes occur rapidly or contains too many different short scenes, because it demands different algorithms of embedding and detecting the watermark in a watermarking process, especially if they use a 3D-based watermarking algorithm. Also, if the video contains long motionless scenes, the algorithm will face the same difficulties. Therefore, in this experiment, we decided to insert the watermark into motion part of video. In order to detect the motion part of video, the following criterion should be satisfied:

$$D(i,i+1)=\sum_{i=1}^{n}\left|H_i(j)-H_{i+1}(j)\right| \quad (19)$$

For this purpose, only the histograms of green components for all frames are utilized. This is because of the intended motion part of video contains a large texture of green color which distinguish it from the other scenes. In this method *i* is the number of frames, *j* is the green component of frames and *H* is the calculated histogram. So by meeting $D(i,i+1) > threshold(T)$ a scene change will be occurred. With the purpose of obtaining the desired motion part of video *T* is regarded as 5305. In the considered video sequence i.e. wildlife sequence, we have chosen the scene that the birds start to fly, 2 frames from the detected motion part are randomly shown in Figure 2.

Figure 2. Two randomly frames from the detected motion scene.

Two other reasons for choosing the motion scene of video for embedding the watermark can be illustrated as follows; Embedding the watermark in motion scene

leads to less imperceptibility of the watermark, because HVS is less sensitive to motion part of video, on the other hand compression algorithms used to destroy the motionless part of video, at the same time, maintain the motion part of it.

In the Human Visual System (HVS), there are three types of cones that react to the basic three colors: red, green and blue. The number of cones reacted to blue is 30 times smaller than the number of cones reacted to red or green, which means the HVS has lack of sensibility to blue color [21]. Figure 3 shows fraction of light absorbed by each type of cone, here R, G, and B represent red, green and blue colors, respectively. For this reason, the proposed algorithm intends to embed the watermark signal into the blue channel; hence a good compatibility between the watermarking and HVS will be achieved.

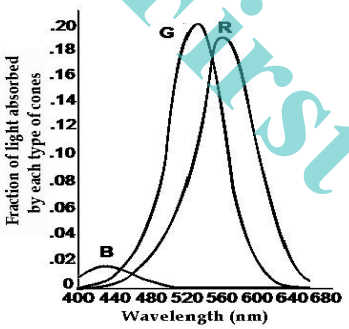


Figure 3. Sensibility of HVS to different wavelength related to three basic colors.

3.2. Watermark Insertion

In this section in order to insert the watermark a Code Division Multiple Access (CDMA) technique is used. In this technique watermarking algorithm can be viewed in terms of a telecommunication system: a message (the watermark) has to be sent through a noise channel (the image/video) to a receiver that has to recover the original message [20]. Generally, in CDMA technique, the watermark information is considered as pseudo random numbers, then by using an algorithm, each bit of watermark information is scattered randomly throughout the video frames, which leads to increase in capacity of embedding and improving the resistance of watermarking system (specially against adding noise). Utilizing CDMA technique makes hard to detect the watermark because it uses wide-band, noise-like signals. Also, by applying pseudo-random numbers which are independent from the data, both the band spread and security of the system will be provided.

The process of embedding the watermark is as follows:

- First of all, a three dimensional wavelet in three levels is applied on the detected motion scene, namely two dimensional DWT in terms of space on each frame and one dimensional DWT within the temporal axis over frames.
- The three dimensional coefficients of HL2, LH2 and HL3, LH3 are chosen for embedding the watermark. Coefficient of LL3 (i.e. the low frequency sub-band) is not watermarked, as video energy is concentrated on lower frequency wavelet coefficient. If it is altered, it will affect on perceptual quality. Also the coefficients of HH (i.e. the high frequency sub-bands) are excluded from embedding the watermark, as data loss usually occurs among the high frequency components due to lossy compression [7]. Coefficients of wavelet decomposition in three levels are shown in Figure 4.

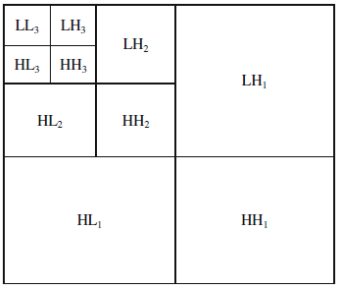


Figure 4. Three levels of discrete wavelet decomposition.

- Embedding process is followed by applying CDMA techniques in a way that pseudo random numbers based on Mersenne–Twister algorithm which is proposed by Nishimura and Matsumoto [16] are created. This method generates numbers with a period of $(2^{19937} - 1)/2$. By using a key, four sets of pseudo random numbers which $W_X(i,j,k) \in \{-1,+1\}$, are achieved based on the adaptive size of wavelet coefficients. So this key is considered as our fourth key in the proposed algorithm which guarantees the security of the proposed watermarking scheme.
  According to magnitudes of the 3D-DWT coefficients, the scrambled watermark is adaptively spread and embedded into these coefficients. Following algorithm shows the embedding process:

$$\begin{cases} C'_X(i,j,k) = C_X(i,j,k) + QW_X(i,j,k) & \text{if } b = 0 \\ C'_X(i,j,k) = C_X(i,j,k) & \text{if } b = 1 \end{cases}$$

  Where $C_X(i,j,k)$ is the 3D wavelet coefficient, $C'_X(i,j,k)$ is the watermarked 3D wavelet coefficient, Q is the modulation index, $W_X(i,j,k)$ is the PRN matrix, b is bit of the scrambled watermark that has to be embedded and $X \in \{HL2, LH2, HL3, LH3\}$.
- Finally by performing inverse 3D wavelet over the motion part of video, the process of inserting the watermark will be accomplished.

3.3. Watermark Detection

During extraction process the original video is not needed, namely, blind detection. The detection is the inverse process of watermark embedding:

- Performing the 3D wavelet on motion part of video, that is, 2D wavelet on each frame and 1D-DWT along temporal axis over frames.
- Selecting the three dimensional coefficients of HL2, LH2, HL3, and LH3 for extracting the watermark.
- Generating four sets of pseudo random numbers based on Mersenne–Twister algorithm by applying the same key which is used in inserting process.
- Calculating the correlation between the extracted coefficients and their relative pseudo-random numbers gives the hidden watermark:

$$\begin{aligned} &if\ \sum\left(\rho(EC_X(i,j,k),\ W_X(i,j,k))\right) > th \\ &watermark\ is\ \text{Availabe} \\ &end\ if \end{aligned}$$

Where $X \in \{\text{HL2, LH2, HL3,LH3}\}$ and $EC_X(i,j,k)$ is the extracted 3D wavelet coefficients. It should be noted that *th* (threshold) is considered by experience. Also the correlation coefficient between the extracted wavelet coefficients and the pseudo random numbers is determined by [31]:

$$\rho(\omega,\omega') = \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(\omega(i,j)-\varpi).(\omega'(i,j)-\varpi')}{\sqrt{(\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(\omega(i,j)-\varpi)^2).(\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(\omega'(i,j)-\varpi')^2)}} \quad (20)$$

Where ϖ and ϖ' are the mean of $\omega(i,j)$ and $\omega'(i,j)$, respectively. The value of ρ will be equal to unity if the watermark is extracted without any error and $\rho = 0$ for an error rate of 50%.

4. Experimental Results

The image which is considered as the watermark has a size of 20×20 which is shown in Figure 1(a). The watermarking process is performed on true colour wildlife video sequence which contains 330 frames and 25 fps with CIF format.

Generally, the accurate measurement of the imperceptibility as perceived by a human observer is a great challenge in image/video processing. The reason is that the amount and visibility of the distortions introduced by the watermarking attacks strongly depend on the actual image/video content [28]. To measure the perceptual quality, the Peak Signal-to-Noise Ratio (PSNR) is calculated which is used to estimate the quality of the watermarked frames in comparison with the original ones. The PSNR [17] is defined as follows:

$$PSNR = 20\log_{10}\left(\frac{\max_i}{\sqrt{MSE}}\right) \quad (21)$$

$$MSE = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}\left\|F_{ij} - \hat{F}_{ij}\right\|^2 \quad (22)$$

where $\max_i = \max\left\{\hat{F}_{ij}, 1 \le i,j \le m\right\}$ and the *MSE* is the mean squared error between the cover frame F and the watermarked frame $\hat{F}$ .This parameter is declared as dB. Based on [8] a PSNR value over 30 dB will be desirable. An example of watermarked frame is shown in Figure 5(a).

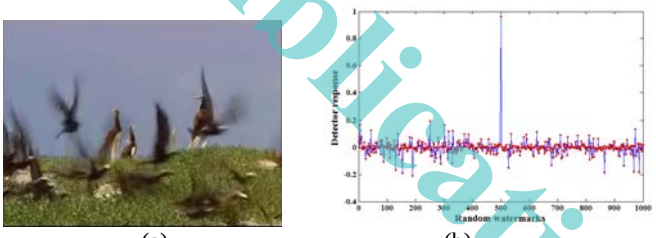
After extracting and refining the watermark, a similarity measurement between the extracted and reference watermark is used for objective judgment of the extraction fidelity which defined as:

$$NC = \frac{\sum_i\sum_j W(i,j)\hat{W}(i,j)}{\sum_i\sum_j [W(i,j)]^2} \quad (23)$$

Which is the cross-correlation normalized by the reference watermark energy to give unity as the peak correlation [5]. In the presented experiment this measurement is used to evaluate the robustness of the proposed scheme.

In order to achieve a high visual quality for watermarked video sequence, the modulation index *Q* should be taken into consideration. In addition, for implementing a good trade-off between the robustness and imperceptibility *Q* plays an important role. So, for this significance, in this experiment *Q* is considered as 3.5. By applying this value there will be an acceptable PSNR of 36 dB. Furthermore the extracted watermark has an admirable NC value of 0.9969 meanwhile *th* is considered as 0.029.

In all robustness evaluations, the watermarked videos are analyzed using 1000 possible watermark signals generated by 1000 different keys; and the embedded watermark generated by the owner corresponds to the key equal to 500.



(a) (b)

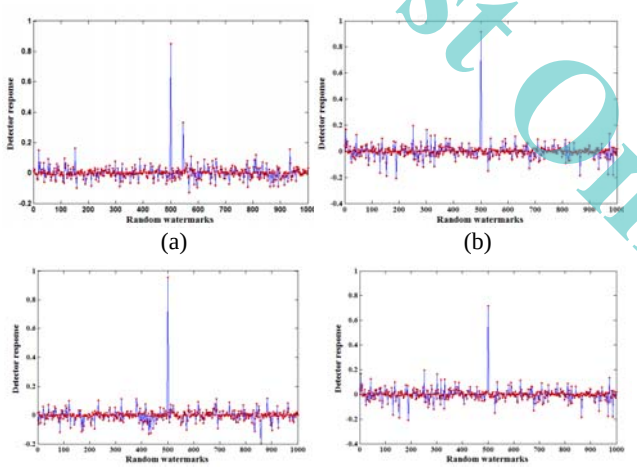
Figure 5. (a) Watermarked frame (b) Detector Response for (a)

In order to show the robustness of the presented scheme against intentional and unintentional attacks different kinds of compression methods like MJEPG, MPEG-2, MPEG-4 and H.264/AVC and also median filtering and Gaussian noise, frame averaging, frame swapping, rotation and rescaling are performed on the proposed watermarking scheme, so the embedded watermark is retrieved using the proposed algorithm and the NC value of the recovered watermark is recorded for all kinds of attacks.

Compression is one of the most basic attacks to

video watermark. In most applications involving storage and transmission of digital video, a lossy coding operation is performed on the video to reduce bit rates and increase efficiency. For this significance, the video watermarking scheme must be robust against lossy compression. Any of aforementioned compressions have specific method for compressing of video frames. The impact of these attacks over the watermarked video is shown in Figure 6.

In order to show the robustness of the proposed algorithm against lossy compression, watermarked video is severely compressed with high Compression Ratio (CR) for all kinds of lossy compression. As Figure 6 shows the proposed algorithm has a good robustness against all of applied compression methods. Furthermore, H.264/AVC lossy compression which recently is used as the most modern compression codec in the worldwide is performed on our scheme. As Figure 6(c) shows our approach has a very good robustness against this widely used kind of compression, too.

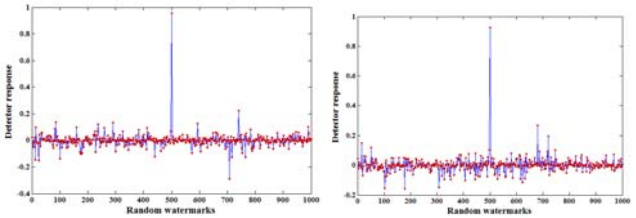


(a) (b) (c) (d)

Figure 6. Watermark robustness against lossy compression; Detector response for (a) MPEG-4 (CR=95:1), (b) MJPEG (CR=95:1), (c) H.264/AVC (CR=93:1), (d) MPEG-2 (CR=95:1), respectively.

Addition of noise is another method to estimate the robustness of the watermark. Generally, addition of noise is responsible for the degradation and distortion of the video. The watermark information is also degraded by noise addition and results in difficulty in watermark extraction. So, in order to test the robustness of the proposed system 1% Gaussian noise with zero mean is added to watermarked video. As Figure 7(a) depicts, the great achieved robustness against noise attack, declares the power of utilized CDMA technique for watermarking process.

On the other hand, one of the most common manipulations in digital videos is filtering. The extracted watermark, after applying 3×3 median filtering is retrieved. By applying this filter, video is significantly degraded and lots of data are lost but the extracted watermark can still be recognized with high NC value. Figure 7 shows the detector response for applied median filtering and Gaussian noise attacks.



(a) (b)

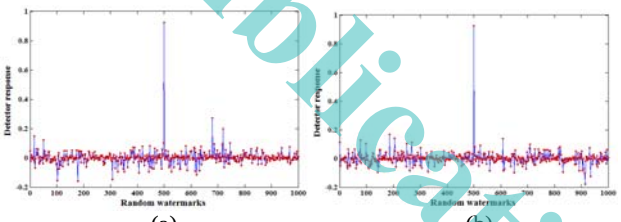
Figure 7.Detector response for (a) Gaussian noise (1%), (b) Median filtering(3×3).

Due to large amounts of data and inherent redundancy between frames, video signals are highly susceptible to pirate attacks, including frame averaging, frame dropping and generally statistical analysis. In frame dropping, selected frames are removed from the watermarked video and replaced by their corresponding original frames. This attack is often used as an effective video watermarking attack, since it leads little or no damage to the video signal. For this purpose 25% of the watermarked video frames are dropped.

Moreover, frame averaging is another significant video watermarking attack. It is clear that the average of multiple frames will remove the dynamic composition of the watermark. Frame averaging is the average of the current frame and its two nearest neighbors to replace the current frame. Averaging is defined by:

$$F_k(i,j)=\frac{[F_{k-1}(i,j)+F_k(i,j)+F_{k+1}(i,j)]}{3} \qquad k=2,3,4,...,n\text{-}1 \qquad (24)$$

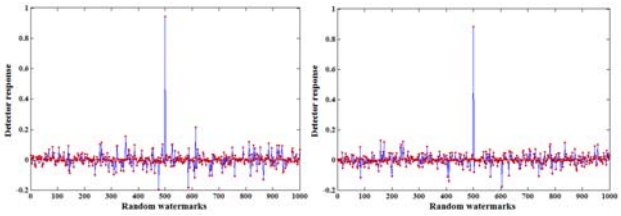
This attack is also performed on 25% of the watermarked video frames. The robustness of the presented algorithm against these two statistical attacks is shown in Figure 8.



(a) (b)

Figure 8. Detector response for (a) Frame averaging (b)Frame dropping.

For more showing the robustness of the proposed scheme another group of attacks in image/video processing domain i.e. geometric attack are tested over the presented algorithm. Rescaling and rotation are two commonly geometric attacks which are widely used in this domain. As Figure 9 shows these attacks have the least effects on the robustness of the proposed scheme.



(a) (b)

Figure 9. Detector response for (a) Rescaling (100-50-100%), (b) Rotation (-0.6º).

5. Conclusions

In this paper a novel watermarking approach for maintaining the copyright in digital videos was investigated. Because of embedding the watermark in blue channel of video, the delivered algorithm has a good compatibility with HVS. Also, the security of our scheme is guaranteed by adding four keys during the embedding process. Furthermore, the blind retrieval of watermark is the strength point of the presented algorithm. So, it can be used for public watermarking applications, where the original video is not available for watermark extraction. As the experimental results showed the proposed method has a good robustness against different kinds of attacks including intentional or unintentional ones.

References

[1] Al-Haj A.M., *Advanced techniques in multimedia watermarking: image, video and audio applications*, Hershey, New York, 2010.

[2] Barni M., Bartolini F., Piva A., "A DCT domain system for robust image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 357–372, 1998.

[3] Bhatnagar G., Jonathan Wu Q. M., Raman B., "A new aspect in robust digital watermarking," *Multimedia Tools Application*, 2011.

[4] Chetan K.R., Raghavendra K., "DWT based blind digital video watermarking scheme for video authentication," *International Journal of Computer Application*, 2010.

[5] Chiou-Tung H., Ja-Ling W., "Digital watermarking for video," *In Proc. 13th International Conf. on Digital Signal Processing (DSP 97)*, pp. 217-220, 1997.

[6] Cox I.J., Miller M.L., Bloom J.A., *Digital Watermarking*, Morgan Kaufmann, San Francisco, 2001.

[7] Cox J., Kilian J., Leighton F.T., Shamoon T., "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Trans. Image Processing*, vol. 6, no. 12, 1997.

[8] Dutoit T., Marqu´es F., Applied signal processing, Springer, New York, 2009.

[9] Gonzalez R. C., Woods R. E., *Digital Image Processing*, third Edition, published by Pearson Education (Singapore) Pte. Ltr., Indian Branch, 482 F.I.F. Patparganj, Delhi 110092, India, 2004.

[10] Hartung F., Girod B., "Watermarking of uncompressed and compressed video," *Signal Processing*, vol. 66, no. 3, pp. 283-301, 1998.

[11] Hsu C.T., Wu J. L., "Hidden digital watermarks in images," *IEEE Trans. Image Processing*, vol. 8, no. 1, pp. 58-68, 1999.

[12] Hwang M. S., Chang C. C., Hwang K. F., "A watermarking technique based on one-way hash functions," *IEEE Trans. Consum Electron* vol. 45, no. 2, pp.286–294, 1999.

[13] Ishtiaq M., Jaffar M. A., Khan M. A., Jan Z., Mirza A. M., "Robust and imperceptible watermarking of video streams for low power devices," *Communication Computer Information Science*, vol. 61, pp. 177-184, 2009.

[14] langelaar G. C., "*Real-time Watermarking Techniques for Compressed Video Data"*, Ph.D thesis, 2000.

[15] Li X., Wang R., "A Video Watermarking Scheme based on 3D-DWT and Neural Network," *Ninth IEEE International Symposium on Multimedia*, Taichung, Taiwan, December, pp. 110-115, 2007.

[16] Matsumoto M., Nishimura T., "Mersenne Twister, A 623-dimensionally equilistributed uniform pseudorandom number generator," *ACM Trans. on Modeling and Computer Simulation*, vol. 8, pp. 3–30, 1998.

[17] Netravali A.N., Haskell B.G., *Digital pictures: representation, compression, and standards*, Plenum, 9ew York, 1995.

[18] Petitcolas F. A. P. , "Watermarking schemes evaluation," *IEEE Signal Processing Magazine*, vol. 17, no. 5, pp. 58-64, 2000.

[19] Pun C. M., "A novel DFT-based digital watermarking system for images," *In Proc. international, conference of signal processing*, Vienna, Austria, vol. 2, pp 1–4, 2006.

[20] Rosa L., "High Capacity Wavelet Watermarking Using CDMA Multilevel Codes," *advancedsourcecode*, 2009

[21] Sayood K., *Introduction to Data Compression*, 2nd Edition, Morgan Kaufmann Publishers, 2000.

[22] Schyndle R. G., Trikel A. Z., Osbrone C. F., "A digital watermark," *In Proc. IEEE international conference on image processing*, Austin, Texas, vol. 2, pp. 86–90, 1994.

[23] Sujatha S., Sathik M., "Blind Wavelet Based Watermarking Technique for Image Authentication", *International Arab Journal of Information Technology (IAJIT)*, vol. 10, no. 3, 2013, to be published.

[24] Swanson M., Zhu B., Tewfik A., "Multi-resolution video watermarking using perceptual models and scene segmentation," *In Proc.*

International Conference on Image Processing, Washington, DC, pp. 558–61, 1997.

[25] Wang Y., Doherty J. F., Van-dyck R. E., "A wavelet based watermarking algorithm for ownership verification of digital images," *IEEE Trans. Image Processing*, vol. 11, no.2 pp.77–88, 2002.

[26] Wang R., Hu L., Xu D., "A Watermarking Algorithm Based on the CABAC Entropy Coding for H.264/AVC," *Journal of Computer Information System*, vol. 7, no. 6 pp. 2132-2141, 2011.

[27] Wei Y., Hao Y., Li Y., "A Multipurpose Digital Watermarking Algorithm of Color Image," *In Proc. IEEE International Conf on Mechatronics and Automation*, Changchun, China, pp. 112-117, 2009.

[28] Winkler S., Gelasca E. D., Ebrahimi T., "Toward perceptual metrics for video watermark evaluation," *In Proc. SPIE, Applications of Digital Image Processing*, pp. 371–378, 2003.

[29] Wolfgang R. B., Podilchuk C. I., Delp E. J., " Perceptual watermarks for digital images and video," *In Proc. of IEEE*, vol. 87, no. 7, pp. 1108-1126, 1999.

[30] Xia X., Boncelet C. G., "A multiresolution watermark for digital images," *In Proc IEEE international conference on image processing*,Washington, DC, USA, vol. 3, pp. 548–551, 1997.

[31] Zhang F., "Image watermarking algorithm based on the code division multiple access technique," *Lecture Notes in Computer Science*, pp. 204–211, 2006.

Majid Masoumi was born in Lahijan, Iran, in 1986. He is a Master student in electrical engineering at Azad University of Qazvin. He received his B.Sc. degree in electrical engineering in 2008. His previously researches include audio watermarking, opto-electronics and photonic crystals. He is currently interested in watermarking, image & video processing, cryptography, optimization, and communication systems. Specially using spread spectrum for making secure the networks against attacks and jammers. He is presently working at Iranian Research Organization for Science and Technology (IROST) as a researcher.

Shervin Amiri was born in Tehran, Iran, in 1966. He received his B.Sc., M.Sc. and Ph.D. from Iran University of Science & Technology (IUST) in communication systems. Now he is a Scientific Member of electrical engineering department in Iranian Research Organization for Science and Technology (IROST). He is supervisor of many PhD and MSc students in the fields of communication system and subsystems.