Multimed Tools Appl (2018) 77:8603–8627
DOI 10.1007/s11042-017-4754-2

A robust and secure key-frames based video watermarking system using chaotic encryption

Yassine Himeur[1,2] · **Abdelkrim Boukabou**[2]

Received: 17 June 2016 / Revised: 7 March 2017 / Accepted: 25 April 2017 /
Published online: 6 May 2017

Abstract Currently we are facing a wide interest in multimedia security and copyright protection due to the explosion of data exchange in the Internet and the extensive use of digital media. In this paper, we propose a video watermarking method in which watermark information are encrypted using a new chaotic encryption, and then, embedded in the key-frames extracted from the video stream. Under this framework, a simple and fast key-frames extraction algorithm based on gradient magnitude similarity deviation (GMSD) is used. This algorithm can significantly decrease the complexity of video watermarking systems. In order to insert the watermark in a blind manner, new insertion and extraction functions are designed by means of a quantization process. A double transformation domain based on discrete wavelet transform (DWT) and singular value decomposition (SVD) is adopted to robustly embed the watermark with low visual distortion. Evaluation study is conducted to verify the performance through a series of experiments. The proposed system outperforms several recent algorithms found in the literature in terms of the robustness and imperceptibility under potential attacks. Furthermore, the security requirement of the proposed algorithm is achieved with the proposed chaotic encryption procedure.

Keywords Video watermarking · Shot detection · Key-frames · Chaotic encryption · Blind · DWT-SVD

✉ Yassine Himeur
yhimeur@cdta.dz

[1] Telecom Division, Center for Development of Advanced Technologies (CDTA), Baba Hassen, Algiers 16303, Algeria

[2] Department of Electronics, University of MSB Jijel, B.P 98, Ouled Aissa, Jijel 18000, Algeria

1 Introduction

Over the last two decades, with the rapid development of smartphones, wireless (e.g., Wifi, Bluetooth) and wired (e.g., ADSL, optical fiber, power-line) networks, there has been an impressive expansion in the volume of multimedia traffic through the Internet [6]. Depending on a Cisco study, in 2015, 64% of consumer Internet traffic was used for video delivery and it will be up to 80% in 2019 [5]. Further, based on the recent statistics published by YouTube, hundreds of millions hours of video are watched each day, more than 150 hours of video are uploaded each minute, which can reflect a raise of 40% in comparison to the last year [50]. This fast increase in the online video content and its sharing opened new challenges in terms of copyright violation detection, and video search and retrieval. Depending on a study conducted in [37], for a set of 24 queries searched in YouTube, Google Video and Yahoo Video, more than 27% of the restored pertinent videos are considered as copies. A copy or a duplicate of a video corresponds entirely to a subgroup of the frames in the original video, individual frames may be further modified and their temporal order varied.

An effective solution to address these challenges is to embed various kinds of copyright information, also known as watermark or logo, in the host video sequence that we want to protect its content. Digital watermarking has found many applications in image, video and audio [8, 21, 33]. To prove the right property, the owner extracts the inserted watermark and compared it with the original one. A watermarking function is robust if it is perceptually robust, in the sense that the inserted watermark should be kept imperceptible, recognizable and robust to potential attacks such as strong compression, scaling, blurring, adding noise, cropping, adding caption, etc. Generally, watermarking gives value-added protection on the top of data encryption and content based copy detection for effective digital rights management [2, 19, 20].

Still, there is a need for a robust video watermarking system with a higher degree of security, which can be achieved by embedding encrypted images in video sequences. There are several encryption algorithms that can be used for the encryption of digital images such as the diffusion based algorithms and the permutation based algorithms. Diffusion based algorithms are suitable for noise free environments, and are very sensitive to rounding errors [38, 48, 51, 52]. So, the appropriate choice to encrypt images to be embedded as watermarks in video sequences is to use one of the permutation based encryption algorithms, which are less sensitive to attacks [1]. Chaotic encryption can be a practical candidate solution to this problem. In fact, chaotic signals are aperiodic, random, broadband, and can be generated using simple circuits in a frequency band at any power level. Chaotic encryption cannot only secure the watermark system, but also can be used to combat the interference and noise generated by the different cited-above attacks.

A novel video watermarking scheme is proposed based on key-frames extraction and chaotic encryption of the watermark to be inserted. At stage 1, the shots boundaries of the request video sequence are firstly detected and representative key-frames are extracted. The copyright information is then encrypted using the new chaotic encryption. The idea is to cipher the watermark based on a chaotic 2-D logistic function with low correlation between its samples before its insertion. At stage 2, a double transformation based on DWT-SVD is applied to each key-frame. Then, the obtained result is sliced into different blocks, and one bit is inserted in each block using a new insertion/extraction function that allows a blind watermarking process.

The work presented in this paper is motivated by the need to develop a practical video watermarking system in order to protect copyright property efficiently, with a high level of security and a low complexity. The proposed watermarking presents several advantages.

The double transformation DWT-SVD is principally used to robustly insert the watermark information. This idea is based on the fact that a combination of two transforms can improve the results in terms of imperceptibility and robustness since each transform can compensate the drawbacks of the other. Further, a new insertion function based on a quantization procedure is proposed which allows a blind extraction process, i.e., there is no need to original video sequence in the extraction procedure. The watermark coefficients are inserted only in the key-frames which can efficiently reduce the time complexity and make our system a good candidate for real time applications. From another side, by inserting the watermarks only in the key-frames, we can keep a good visual quality of the watermarked video sequences. In order to make a high security level, a chaotic encryption procedure is introduced. Finally, the proposed watermarking system gives a good trade-off between robustness to different attacks, capacity of insertion and imperceptibility.

The rest of the paper is organized as follows. Section 2 briefly summarizes some recent related works found in the literature. Section 3 describes the proposed video watermarking system. Consequently, the key-frames extraction process, chaotic encryption and insertion/extraction procedures are explained in details. Simulation results are given in Section 4 to demonstrate the efficiency of the proposed scheme. Finally, concluding remarks are drawn in Section 5.

2 Related works

Digital watermarking for copyright protection is an active area of research. Commonly, watermarks can be inserted either by directly modifying the samples of the original carrier or by modulating the coefficients in the transform domain, and in most cases, watermarks should be invisible and resistant to potential attacks. In this section, we analyze a set of well-known watermarking schemes. Those methods are classified based on the insertion domain into uncompressed and compressed based techniques.

Uncompressed based techniques make use of spatial, temporal, and transform based domains. Many methods utilize several projections and transforms of video frames to both robustly and invisibly insert watermark information. In [11], e.g., El'Arbi et al. presented a watermarking scheme that inserts watermark coefficients into several scenes of a video in the wavelet domain. By using motion activity analysis, many regions of the original video are separated into perceptually distinct categories according to the motion information and region complexity. To make the watermark imperceptible and less sensitive to automated removal, the watermark coefficients are adaptively inserted with respect to the human visual system (HVS) properties. However, this approach is only suited for videos with significant motion activities. Furthermore, it shows a weak resistance against geometrical attacks and a high time complexity. In [35], binary watermarks are inserted in the detail wavelet coefficients of the middle wavelet sub-bands to watermark the video streams. This scheme combine both the spread spectrum and the quantization-based information hiding techniques, in the sense that every bit of the logo is spread over a number of wavelet coefficients with the use of a secret key using a quantization process. Then, the chosen wavelet detail coefficients picked from different sub-bands are quantized by means of an optimal quantization model based on the HVS properties. However, this scheme has not enough immunity against different external attacks since it provided unacceptable results in terms of imperceptibility with a peak signal-to-noise ratio (PSNR) below 40 dB. In [12, 18, 25, 30], different watermarking schemes have been proposed for image and video watermarking by mean of SVD using the DWT. Basically, images and/or videos are

transformed with the DWT using different resolution levels, then the resulting high frequency band (HH) and middle frequency bands (LH, HL) are watermarked and transformed using the SVD transform. This watermarking concept shows good robustness against several transformations, however, the major drawback is that it presents less security level. In fact, the watermarking system should be highly secured using an encryption algorithm to cope with intruder attacks. In [41], both visual cryptography and scene change detection are used to insert watermark information in DWT domain. Chen and Zhu proposed in [4] a robust watermarking algorithm for video copyright protection that can be conducted in the following four steps. First, the SVD transform is applied to insert watermark indexes. Then, a slope-based insertion process is introduced to embed a 1-bit watermark into several successive blocks in the temporal direction in order to improve the robustness against different attacks. After that, a block insertion algorithm is employed to accord priority for blocks with small variations that can enhance the visual quality of the watermarked video. Finally, temporal synchronization attacks are suppressed using a specific temporal synchronization approach. In spite of having a good imperceptibility and strong robustness against some attacks such as frame dropping and frame insertion attacks, the primary issue of this watermarking system is its sever degradation under some video attacks, e.g., video rotation and cropping. Youssef et al. [49] developed an adaptive video sequence watermarking system by means of fuzzy logic and wavelet transform. This model incorporate the HVS properties of video motion sub-regions in the frequency multi-resolution wavelet domain using a multi-dimensional fuzzy inference perceptual model. However, the use of adaptive fuzzy inference perceptual model makes the watermarking system very complex in time. In [27], Li et al. employed the entropy model for local motion characterization to embed watermark bits. The algorithm firstly associates HVS with the block-matching procedures to extract the motion-related information. Then it uses the entropy model to statistically analyze above motion-related information in order to obtain the motion entropy. This system splits each frame into local partitions, and then, local motion entropy is extracted based on the motion-related information in a local region. Afterward, the motion properties visual masking is estimated depending on the local motion entropy with the motion entropy of frame. Finally, the maximal strength of embedding is determined using both the motion characteristics visual masking and the contents of video frames. As mentioned before, using the motion properties to insert the watermark can cause some problems with video sequences especially where there is no significant motion activity. In [45], the watermark information is inserted in local polar harmonic transform domain to combat geometric attacks. Further, the stable and uniform frame feature points are generated using the improved speeded-up robust feature descriptor, where the probability density gradient is used. Then, the affine invariant local feature regions are adaptively designed based on the variation of local probability density. Finally, a 2D transform, named polar harmonic transform, is introduced to embed watermark in each frame. Even if this watermarking has good performance against geometrical attacks, it should be noted that it fails to keep the watermark under other transformations such as H.264/AVC compression, median filter and frame averaging.

On the other part, compressed based techniques use different coding standards such as MPEG-2, MPEG-4, and H.264/AVC to insert the watermark indexes. In [17], the watermark are embedded in the homogeneous moving object inside a shot of video sequence to confront geometric attacks, e.g., flipping, rotation,..., etc. Intuitively, object based watermarking derives a low insertion capacity and has the least impact on imperceptibility since the object area is generally small and highly textured. This concept can be described in two tasks, firstly, an existing compressed domain motion coherent block detection algorithm [9]

is extended to detect the moving objects within a video shot, and secondly, a watermarking scheme is addressed by embedding within the moving objects to resist rotation, scaling and translation attacks. This approach used an interesting idea for watermarking, however, the performances in terms of robustness are far away from the state of the art performances. This method has serious problems with several attacks such as: salt and pepper noise, Gaussian filter and scaling.

In [39], Tardos probabilistic fingerprinting code is inserted in H.264/AVC compressed video sequences using spread spectrum watermarking technique in both luma and chroma such that the Tardos code is embedded in intra as well as inter frames. The insertion is completed in the nonzero quantized transformed coefficients to prevent the uncontrollable rise in bitrate of video bitstream. This approach has a good performances in terms of payload and imperceptibility, however, it presents a weak resistance to different attacks. In [47], a video watermarking approach based on H.264/AVC codec and selecting host frame is proposed. The watermark information is cropped into small watermarks based on the number of shots in the host video, and small watermarks are respectively inserted into the shots. Furthermore, a scheme is introduced to select host coefficients by means of block classification in the discrete cosine transformation (DCT) domain. The insertion areas of watermark indexes are selected adaptively based on the video content and according to texture properties of host blocks. The major issue with this watermarking system is that it is not blind since the original video sequence is needed in the extraction process.

Moreover, shot boundary detection (SBD) and key-frames extraction play important roles in many video applications. They can be very helpful for video watermarking systems. In [36], a system based on SBD operating directly in the compressed domain is proposed. After extracting local indicators from MPEG macroblocks, AdaBoost is employed for both feature extraction and fusion. Then, shot boundaries are defined using the selected features via pre-filtering and rule-based decision making. After that, similarity between boundary frames of cut candidates is measured and analyzed using phase correlation of dc images in order to create a video summarization. This method seems interesting since it has good results in terms of precision and recall. However, the major issue with this scheme is due to the fact that both the selection and the fusion modules considerably increase the time complexity. In [29], a macroblock classification method is proposed that can be very useful for key-frames based video watermarking application. The idea is based on the classification of macroblocks corresponding to each video frame into different classes, and then, use this class information to describe the frame content. This scheme has good precision to detect and extract key-frames with a low-computation complexity. However, this approach cannot be efficient for videos with low motion activity since it is based on the analysis of the motion vector.

3 Proposed watermarking system

In this section we describe our video watermarking system that involves extracting key-frames using an interesting approach based on GMSD [20], and that encrypts the watermark image by means of a new chaotic encryption before its embedding in the key-frames sequence using a novel embedding strategy, allowing its extraction in a blind manner. To be robust against potential attacks, the encrypted watermark is inserted in DWT-based SVD domain.

3.1 Shot detection and key-frames extraction

In this part we present a simple yet efficient scheme for video summarization that works in the uncompressed domain based on measuring the GMSD between consecutive frames. In fact, the GMSD is very sensitive to any change between consecutive video frames [46] and hence, it has the ability to detect shot boundaries in a video stream. Moreover, it has a good robustness against different attacks such as transcoding (H.264/AVC), edits, noise, cropping, logo insertion, etc. [20]. A flowchart of the key-frame based GMSD extraction algorithm is shown in Fig. 1. Accordingly, visual features are extracted from the video stream to describe its visual content for each frame of an input sequence. After that, a simple and fast algorithm is used to detect groups of video frames with a similar content and select a representative frame per each group. Finally, the selected frames are filtered in order to avoid possible redundant or meaningless frames in the video summary.

The proposed key-frames extraction is based on measuring the distortion between consecutive frames of the whole video sequence in order to detect key-frames with significant change of the visual content. After calculating the GMSD difference between all video frames, a vector is obtained, and each value of the vector is compared to a threshold. Only frames with a distortion $dist$ exceeding the threshold value are considered as key-frames.

In fact, the GMSD is an improved version of the GMS (Gradient Magnitude Similarity) map with pooling strategy. If we consider a reference image (denoted by r) and a distorted image (denoted by d), then the gradient magnitudes of r and d, denoted by $m_r(i)$ and $m_d(i)$, respectively, are computed at a location i as follows [46]:

$$m_r(i) = \sqrt{(r \otimes h_x)^2(i) + (r \otimes h_y)^2(i)} \tag{1}$$

$$m_d(i) = \sqrt{(d \otimes h_x)^2(i) + (d \otimes h_y)^2(i)} \tag{2}$$

where $\otimes$ denotes the convolution operation, h_x and h_y are the Prewitt filters along horizontal (x) and vertical (y) directions, respectively, defined by

$$h_x = \begin{bmatrix} 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \end{bmatrix}, \quad h_y = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 \\ -1/3 & -1/3 & -1/3 \end{bmatrix}. \tag{3}$$

Then the GMS map is calculated using the gradient magnitude images m_r and m_d as follows:

$$GMS(i) = \frac{2m_r(i) + m_d(i) + c}{m_r^2(i) + m_d^2(i) + c} \tag{4}$$

where c represents a positive constant to support numerical stability.

Finally, the GMSD is just the standard deviation of the GMS map, considered as the image quality assessment (IQA) index, such that

$$GMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(GMS(i) - GMSM)^2} \tag{5}$$

where N is the total number of pixels in the image. The threshold used in the key-frames extraction process is computed using the following equation:

$$Thr = \frac{\alpha}{2}(\max_{GMSD} + \min_{GMSD}) \tag{6}$$

where $0 < \alpha < 1$, max and min are the maximum and minimum values obtained when computing the GMSD difference between two consecutive video frames, respectively.

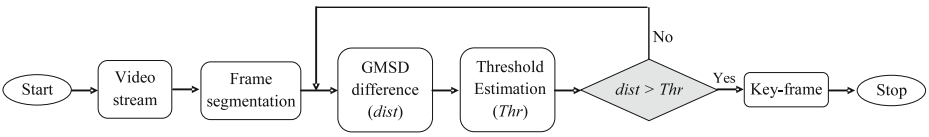


Fig. 1 Flowchart for key-frames extraction

It is worth noting that the estimation of the key-frames is a delicate task; it must be carefully selected to minimize the video size and to be robust against different attacks. If the key-frames extraction procedure is not robust against attacks, then most of the key-frames will be changed after applying an attack, resulting in a poor extraction process [20].

3.2 Watermark embedding and extraction

Both robustness and practicality of the watermarking system are mainly dependent on the insertion and extraction procedures. For this reason, we have developed a new blind insertion/extraction function that permits a good way to hide the watermark logo, allowing its extraction without the need for the original video.

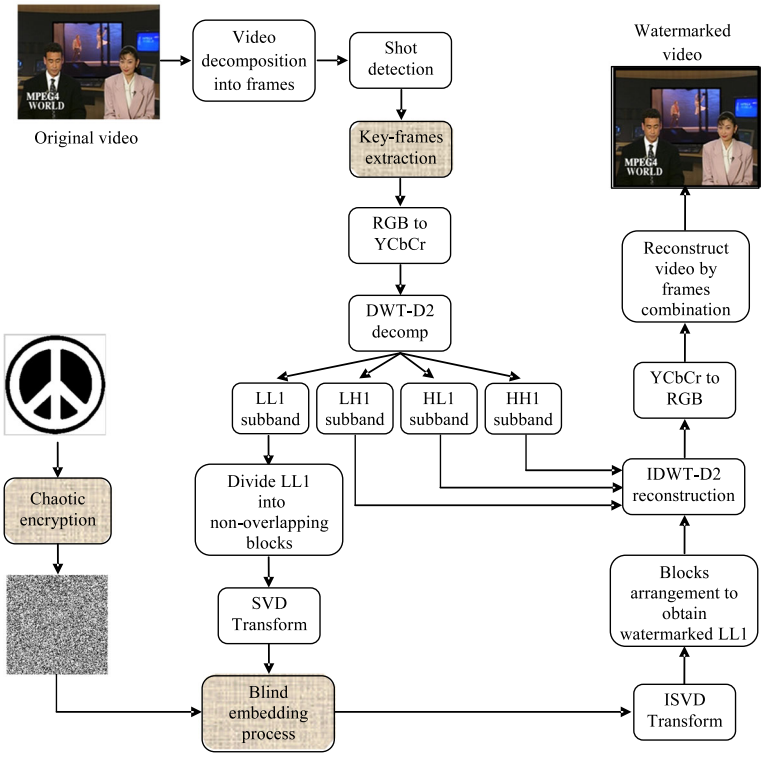


Fig. 2 Block diagram of the proposed video watermark system

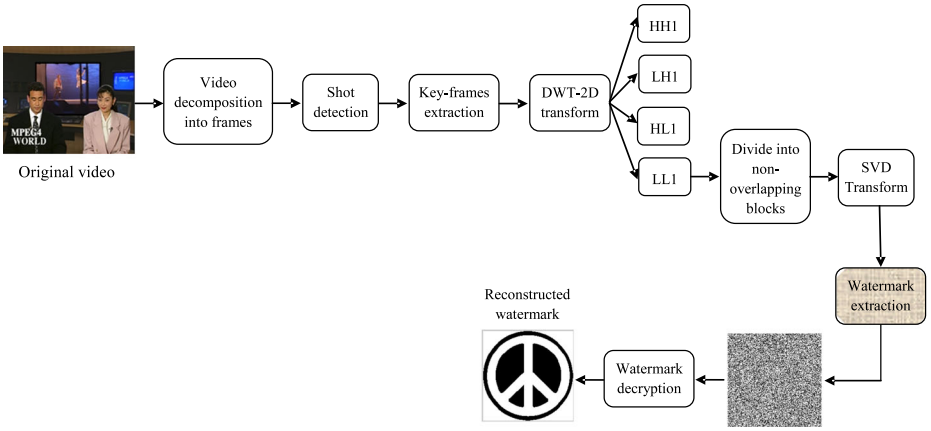


Fig. 3 Watermark extraction process

3.2.1 Watermark embedding

Figure 2 shows the block diagram of the proposed video watermark insertion procedure. The whole video sequence frames are analyzed using GMSD in order to detect shot boundaries, in which a set of key-frames are selected. Each shot boundary is represented by only a key-frame. Next, all obtained key-frames are converted from the RGB color space to the YCbCr color space and then transformed using DWT. After that, the wavelet approximation coefficients of the luminance Y are decomposed using SVD into blocks of size $M \times M$. This procedure is performed on the original blocks Z_l in order to obtain two orthogonal matrices U_l and V_l, and a diagonal matrix S_l such that

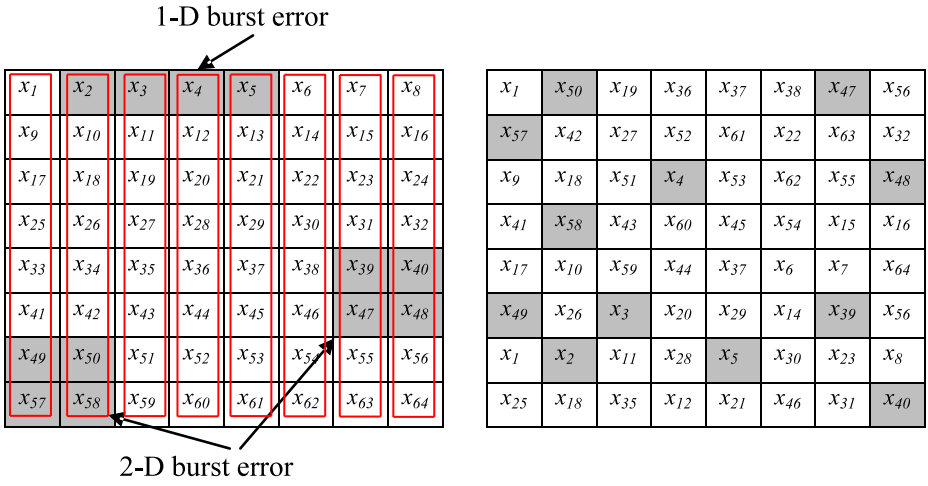
$$Z_l = U_l S_l V_l^T \tag{7}$$

where l is the index of the block. Finally, the watermark bits are inserted in the singular values (SVs), i.e., the S matrix [12]. Unlike the work in [12] where the watermark is inserted in a non-blind manner, the proposed scheme uses a quantization process to embed the watermark in a blind manner. Hence, the original video is not needed in the extraction process.

The wavelet decomposition level is selected to be $L = 1$. The selection of the first decomposition level is a trade-off between the imperceptibility of the watermark, the resistance to attacks and payload. In fact, if the watermark bits are inserted in the LL2 sub-bands or higher levels, then the perceptual quality of the video will be significantly altered, and therefore, the capacity of watermarking will be considerably decreased. For these reasons, the best choice for watermark embedding is the first wavelet partition level.

The embedding process is performed into the S matrix of each block according to an average value AV of the given block as follows:

$$AV = \frac{1}{M_d} \sum_{m=1}^{M_d} \sum_{m=1}^{M_d} |S(m, m)| \tag{8}$$



(a) Without chaotic encryption (b) With chaotic encryption

Fig. 4 Effect of chaotic encryption against burst errors generated by different attacks

where M_d is the number of non-zero coefficients of the textured block. The watermark bits a_l are inserted into the diagonal coefficients $S_l(m, m)$ of each selected block using the following formula

$$S_l^W(m,m) = \begin{cases} -|AV - \Delta| & \text{if } a_l = 1 \text{ and } S_l(m,m) > 0 \\ -|AV + \Delta| & \text{if } a_l = 1 \text{ and } S_l(m,m) < 0 \\ |AV + \Delta| & \text{if } a_l = 0 \text{ and } S_l(m,m) > 0 \\ |AV - \Delta| & \text{if } a_l = 0 \text{ and } S_l(m,m) < 0 \end{cases} \tag{9}$$

such that

$$\Delta = \alpha \, ||S_l(m,m)| - AV| \tag{10}$$

The parameter α represents the strength factor, determined empirically between 0 and 10; the parameter l represents the index of the binary bit of the watermark signal; and parameters m and Δ can be modified to adjust the imperceptibility and robustness. Generally, the index m is selected to match the middle frequency coefficients [26].

3.2.2 Watermark extraction

The extraction process of the digital watermark from the watermarked video is similar to the watermark insertion process. Figure 3 shows such extraction process.

Accordingly, for each watermarked video, we firstly apply the two-level DWT based Daubechies on the key-frames extracted after conversion from RGB to YCbCr color space. Next, the SVD is performed for each $8 \times 8/16 \times 16$ blocks within the low frequency sub-band LL1. The resulting non zero diagonal coefficients S_l are then scanned, and the extracted bits b_l are obtained as follows:

$$b_i = \begin{cases} 1 \text{ if } S_l(m,m) < 0 \\ 0 \text{ if } S_l(m,m) > 0 \end{cases} \tag{11}$$

The recovered one dimensional sequence b_l is transformed into a matrix representation, which is finally decrypted using the proposed chaotic encryption procedure to reconstruct the inserted binary watermark.

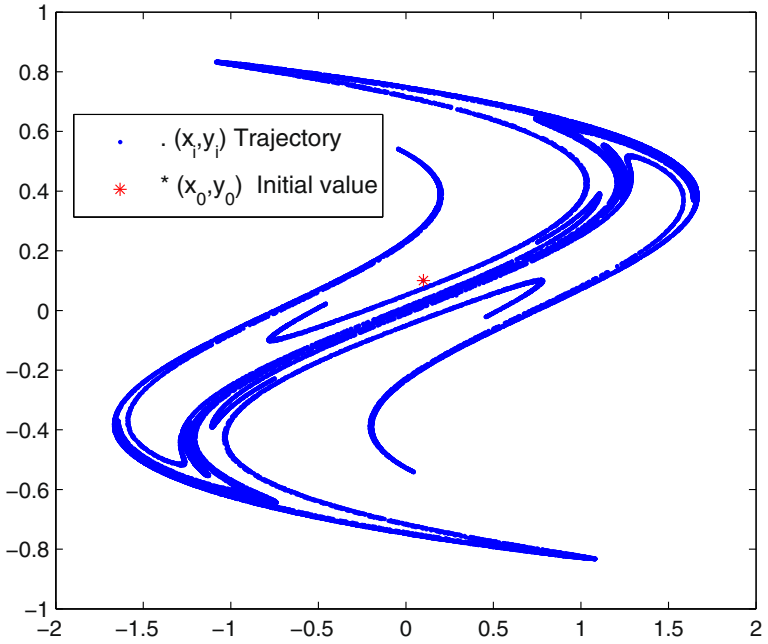


Fig. 5 A trajectory of the 2-D logistic function

3.3 Chaotic encryption

In this framework, a binary image is used as a watermark information. The watermark bits are arranged on a 2-D format as indicated in Fig. 4a. The chaotic 2-D logistic function is applied on each column of the matrix. This function is used to randomize the bits using an iterative mode [42]. Furthermore, the initial conditions, control parameters and number of iterations are used as a secret key. In doing so, the degree of encryption to the watermarking system is highly improved. In fact, even if an intruder can extract the watermark information; he should decrypt it in order to prove the ownership of the video sequence.

As illustrated in Fig. 4b, the chaotic encryption procedure distributes burst errors generated by the different attacks (such as additive Gaussian noise, salt & pepper, cropping, etc.) through different positions, which considerably enhances the robustness of the video watermarking system. The chaotic interleaving is performed on a 2-D vector as follows:

Step 1. Generate a chaotic sequence using the following chaotic 2-D logistic function

$$\begin{cases} x_1(k+1) = a x_1(k) - x_1(k)^3 + x_2(k), \\ x_2(k+1) = b x_1(k), \end{cases} \tag{12}$$

where a and b are the system parameters to exhibit a chaotic behavior, selected as $a = 1.9$ and $b = 0.5$.

Step 2. Construct a chaotic binary sequence $\mathbf{C} = [c_0, c_1, \ldots, c_{N-1}]$ as follows:

$$c_k = \begin{cases} 1, & \text{if } x_1(k) > T, \\ 0, & \text{otherwise}, \end{cases} \tag{13}$$

Fig. 6 Binary watermark used in the simulation

where N is the length of the chaotic binary sequence $\mathbf{C}$ and T is a specific threshold used to generate this sequence from the 2-D logistic function. This step is performed to randomize the data bits inserted in the video key-frames when using the 2-D logistic function. A vector $\mathbf{V}_{in}$ of linear indexes corresponding to $\mathbf{C}$ will be assigned.

Step 3. Divide the bits in the input vector $\mathbf{V}_{in} = [v_1, v_2, ..., v_N]$ into two groups $\mathbf{B}_1 = [b_0^1, b_1^1, \ldots, b_{\frac{N}{2}-1}^1]$ and $\mathbf{B}_2 = [b_0^2, b_1^2, \ldots, b_{\frac{N}{2}-1}^2]$ depending on the bits in the chaotic binary sequence $\mathbf{C}$. For each bit c_k in $\mathbf{C}$, we check the corresponding bit in $\mathbf{V}_{in}$. If the bit is 1, we put the corresponding bit in $\mathbf{V}_{in}$ into $\mathbf{B}_2$. Otherwise we put this bit into $\mathbf{B}_1$.

Step 4. Put $\mathbf{V}_{out} = \mathbf{C}$, then Steps 1 to 4 are repeated M_t times in order to have a flat output distribution. Finally, we get the result value in $\mathbf{V}_{out}$ by concatenating $\mathbf{B}_1$ and $\mathbf{B}_2$.

To recover the input sequence $\mathbf{C}$ from the output sequence $\mathbf{V}_{out}$, we operate in the inverse order, and we must know the value of the secret key K_s composed of initial conditions, control parameters and number of iterations, expressed as $K_s = [x_1(0), x_2(0), a, b, M_t]$.

These parameters are used to construct the secret key K_s, and therefore, a very slight variation in one of these parameters makes the decryption process impossible as will be discussed in the experimental results.

Figure 5 shows the scatter plot of 30,000 points from the trajectory of the 2-D logistic function and the initial value $(x_0, y_0) = (0.111, 0.111)$.

4 Experimental results

The proposed watermarking scheme has been implemented using MATLAB 8.1 platform on a system running on Pentium-core i3 processor with 3.3 GHz and 12GB RAM. Under this framework, a total of 12 CIF (288×352) video sequences in YUV format with a frame rate of 25 fps are employed. Each video sequence consists of 300 frames. Binary images of sizes 54×54 and 27×27 are used as watermarks. Figure 6 shows the binary watermark with a size of 54×54.

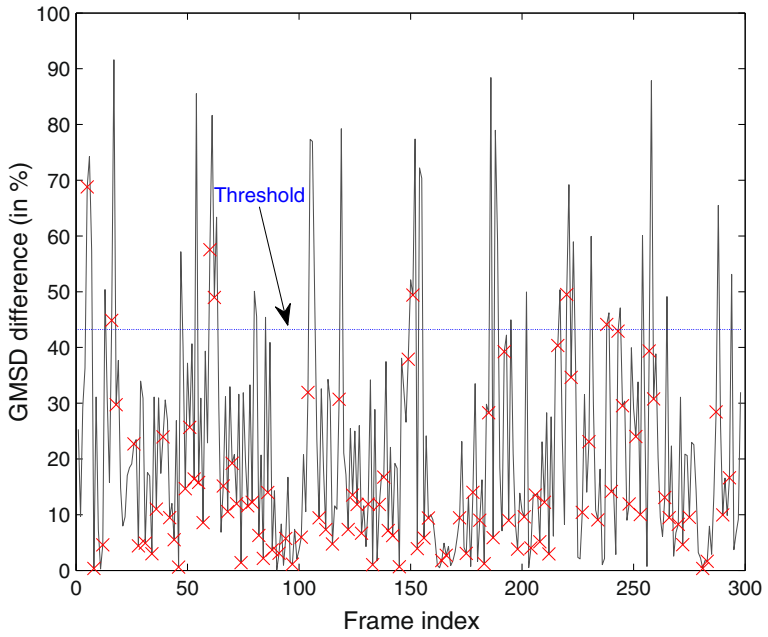


Fig. 7 Example of local max values extracted for 'Salesman' video

Generally, four aspects are often used as performance metrics in watermarking systems, namely, imperceptibility, robustness, payload insertion and computational complexity. We used the PSNR as the imperceptibility metric [32]:

$$PSNR = 10\log_{10}\frac{255^2}{\frac{1}{w_v \times h_v}\sum_{x=0}^{w_v-1}\sum_{y=0}^{h_v-1}(I_{x,y} - I'_{x,y})^2} \tag{14}$$

where I and I' denote the frame in the original video and the corresponding frame in the watermarked video, respectively; (x, y) is the position of a pixel in I or I'; and $w_v \times h_v$

Fig. 8 Example of key-frames extracted from 'Salesman' video

Table 1 Recall, precision and F-measure achieved by different techniques

	DT [34]	STIMO [16]	VSUMM [3]	FASAM [10]	RPCA-KFE [7]	GMSD
Recall	0.66	0.68	0.74	0.86	0.96	0.95
Precision	0.62	0.66	0.73	0.82	0.95	0.93
F-measure	0.65	0.64	0.72	0.84	0.954	0.94

is the resolution of each frame in the original video. For a watermarked video, the average PSNR for all frames is used as the PSNR of the video. A higher PSNR indicates a better performance with respect to the imperceptibility.

The robustness is a measure of the immunity of the watermark against attempts to remove the watermark using different types of digital signal processing attacks. We measured the similarity between original and extracted watermark from the attacked watermarked video using the correlation factor, which is computed using the following formula [31]:

$$NC = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} w(i,j)w'(i,j)}{\sum_{i=1}^{N}\sum_{j=1}^{M} w^2(i,j)w'^2(i,j)} \tag{15}$$

where w and w' are the original and extracted watermark images, respectively. N and M are the number of rows and columns of the watermark images, respectively.

The watermark resistance to different attacks can also be measured by the bit-error rate (BER) [22]. We used the BER defined by (16) to evaluate the performance of our scheme such that

$$BER(w,w') = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} w(i,j) \otimes w'(i,j)}{N \times M} \times 100\% \tag{16}$$

where $\otimes$ is the exclusive or (XOR) operator.

4.1 Key-frames extraction evaluation

4.1.1 Compression ratio

The key-frames compression aims to reduce the amount of video data while preserving the overall contents of the original video. For this reason, we have also evaluated the proposed scheme according to the compactness of the summary (compression ratio). This latter is computed by dividing the number of key-frames in the summary by the length of the video sequence. For a given video sequence, the compression ratio is defined as:

$$CR_{(compr\ ratio)} = 1 - \frac{\gamma_{NKF}}{\gamma_{NF}} \tag{17}$$

where γ_{NKF} is the number of key-frames in the summary, and γ_{NF} is the total number of frames in the video sequence. Ideally, a good summary produced by a key-frame extraction algorithm will present both high quality measure and high compression ratio (i.e., small number of key-frames).

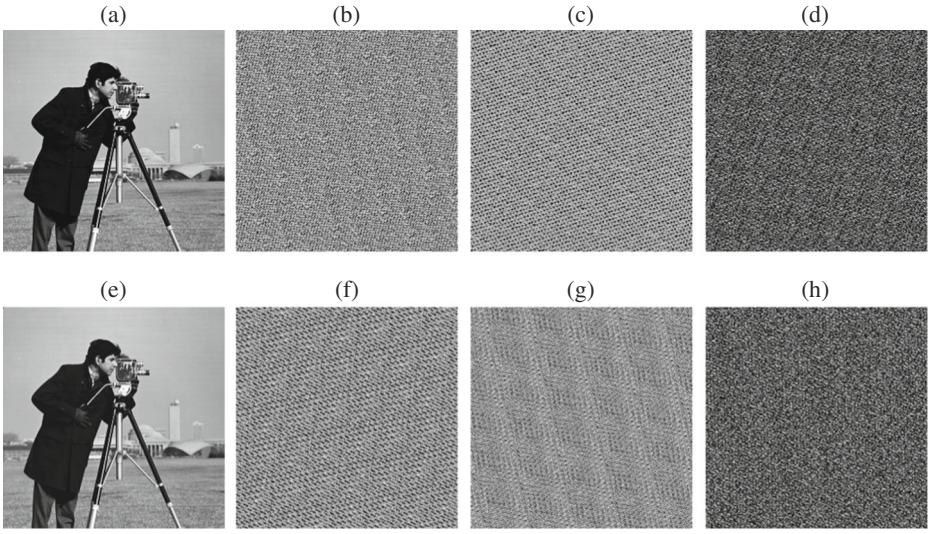


Fig. 9 Key sensitivity results. **a** Plaintext image P; **b** ciphertext image $C^1 = \text{Enc}(P, K^1)$; **c** ciphertext image $C^2 = \text{Enc}(P, K^2)$; **d** ciphertext image difference $\left|C^1 - C^2\right|$; **e** deciphertext image $D^1 = \text{Dec}(C^1, K^1)$; **f** deciphertext image $D^2 = \text{Dec}(C^1, K^2)$; **g** deciphertext image $D^3 = \text{Dec}(C^1, K^3)$; **h** deciphertext image difference $\left|D^3 - D^2\right|$ (K^1 and K^2 are different only for one bit, K^2 and K^3 are also different only for one bit, and $K^1 \neq K^3$)

Using the developed key-frames extraction algorithm, we obtained an average compression ratio of CR = 97.66%. Consequently, the representative key-frames can be extracted accurately and semantically from long video sequences or videos with more transitions, reflecting the video content objectively. An example of key-frames detection by means of estimating the GMSD difference is shown in Fig. 7 which illustrates the positions of the key-frames in the video sequence and how they were selected using the proposed key-frames extraction approach. As it is shown, the extraction process is based on abrupt change of the similarity between the video frames and on the threshold that is fixed experimentally to consider a frame as a key-frame. Moreover, Fig. 8 shows the results of key-frames extraction using the proposed scheme for the case of 'Salesman' video. Only seven key-frames are extracted from the 300 frames that formed the video sequence, while the video content can be clearly acknowledged.

4.1.2 Comparison with non-visual attention based techniques

This section discusses the comparison results of the proposed scheme with other well-known key-frames extraction schemes in the literature. The compared techniques are: Delaunay clustering (DT) scheme [34], STIll and MOving video (STIMO) scheme [16], video summarization based on the VSUMM [3] and FASAM [10] methodologies, and key-frame extraction for video using robust principal component analysis (RPCA-KFE) [7]. Note that the key-frames generated by each method are compared to the corresponding ground truth key-frames. Under this framework, ground truth key-frames (user summaries) drawn generated by [3] and available at http://www.npdi.dcc.ufmg.br/VSUMM. Five human experts were used to generate this ground truth manually after watching the videos. The number of similar and different key-frames between summary and ground truth of each scheme is then calculated. When we apply an approach to extract the key-frames, a frame

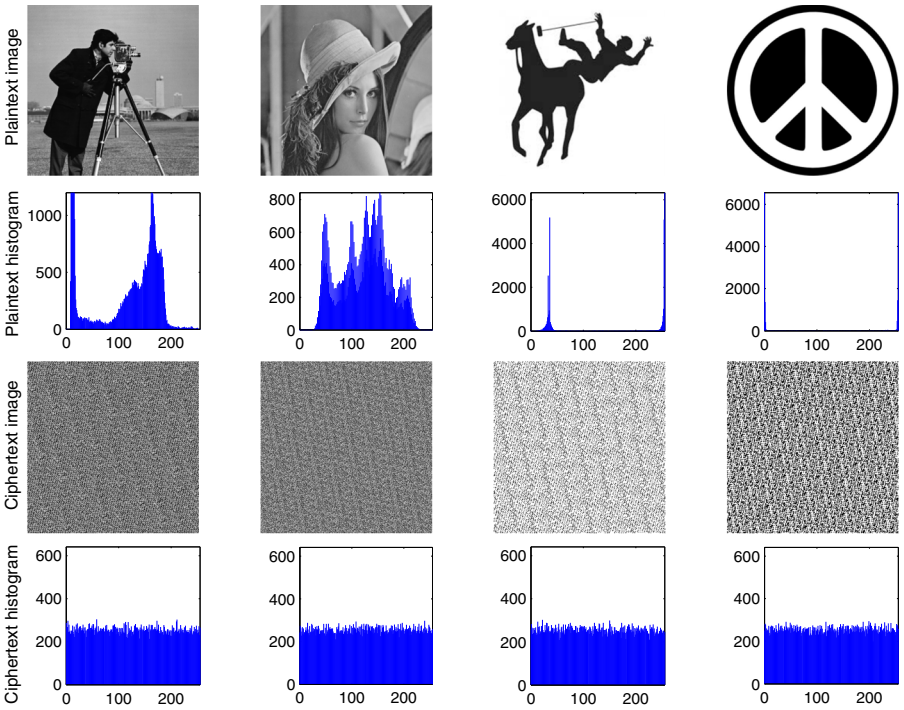


Fig. 10 Histogram analysis of encrypted images

is considered as true positive if it is selected as key frame by both human and the considered approach. A false positive takes place when a frame is considered as key-frame by the approach but not by the user, while a frame selected as key-frame by the user but not by the approach is considered as false negative. The Recall and Precision metrics are then drawn using the whole number of true positive n_{TP}, false positive n_{FP} and false negative n_{FN} as follows

$$Recall = \frac{n_{TP}}{n_{TP} + n_{FN}} \tag{18}$$

and

$$Precision = \frac{n_{TP}}{n_{TP} + n_{FP}} \tag{19}$$

Furthermore, F-measure is employed to consider the averages of Recall and Precision, which is determined as

$$F = 2\frac{Recall \times Precision}{Recall + Precision} \tag{20}$$

Table 1 shows the average values for Recall, Precision and F-measure for all the schemes under consideration. Consequently, the GMSD based scheme achieves good results for the three metrics. It highly outperforms the performances of some well-known methods such as: DT [34], STIMO [16], VSUMM [3] and FASAM [10]. Moreover, the performances of the key-frame based GMSD scheme are comparable to those obtained by RPCA-KFE [7], with the advantage that the GMSD based approach is less complex and very fast. The result illustrates that the GMSD based scheme is valid to segment the shot and extract the key-frames while satisfying a strong robustness.

Table 2 Robustness in term of NPCR and UACI scores

File name	Liao's [28]		Hua's [24]		LAS-IES [23]		Proposed	
	NPCR	UACI	NPCR	UACI	NPCR	UACI	NPCR	UACI
5.1.09	49.8093	16.6687	99.6658	33.5908	99.6064	33.4456	99.5124	33.5214
5.1.10	99.6140	33.5374	99.6475	33.5366	99.6154	33.4946	99.6121	33.4215
5.1.11	49.8138	16.7015	99.6674	33.4398	99.6244	33.5541	99.5943	33.4014
5.1.12	49.8280	17.0621	99.5941	33.4228	99.5703	33.4302	99.5811	33.4158
5.1.13	99.5972	33.6419	99.6445	33.4205	99.6109	33.4438	99.5963	33.4236
5.1.14	99.6368	34.2965	99.5975	33.4696	99.6364	33.4655	99.5945	33.3951
5.2.08	99.6208	33.4267	99.6281	33.4720	99.5870	33.4008	99.5878	33.3978
5.2.09	99.6174	33.4553	99.6197	33.4921	99.6260	33.4804	99.5812	33.4182
5.2.10	99.6292	33.4993	99.6288	33.4914	99.6124	33.4563	99.6100	33.4263
7.1.01	49.8005	16.8228	99.6273	33.5212	99.5992	33.5037	99.6028	33.4474
7.1.02	49.8039	16.8126	99.5892	33.4846	99.6075	33.4237	99.6078	33.4326
7.1.03	49.8096	16.7308	99.6201	33.4647	99.6079	33.4291	99.5811	33.4836
7.1.04	99.6094	33.4778	99.5894	33.5202	99.5988	33.4739	99.5946	33.4782
7.1.05	99.6063	33.4581	99.6185	33.5400	99.6170	33.4362	99.5937	33.4716
7.1.06	99.6048	33.4489	99.6117	33.5254	99.6272	33.3954	99.5912	33.4365
7.1.07	99.6323	33.5216	99.6223	33.5205	99.5931	33.4073	99.6014	33.4313
7.1.08	99.6101	33.4496	99.6151	33.5678	99.6094	33.4332	99.6013	33.4460
7.1.09	49.8100	16.7680	99.6044	33.5223	99.6162	33.4117	99.6148	33.3856
7.1.10	49.8199	16.8557	99.6101	33.4325	99.6045	33.4344	99.6097	33.3941
boat.512	99.6037	33.6291	99.6006	33.5097	99.6154	33.4654	99.6101	33.3973
alaine.512	99.6292	33.4419	99.6128	33.5477	99.6196	33.4225	99.6185	33.4104
gray21.512	99.6254	33.4770	99.6082	33.3930	99.6022	33.4608	99.6034	33.4089
numbers.512	99.6120	33.4503	99.6059	33.3993	99.6141	33.4240	99.5941	33.4561
ruler.512	99.6304	34.0635	99.6265	33.5129	99.6120	33.4262	99.5945	33.4635
5.3.01	498086	49.8086	99.6098	33.4532	99.5931	33.4585	99.6032	33.4392
5.3.02	996163	99.6163	99.6119	33.4853	99.6128	33.4605	99.6108	33.4547
7.2.01	49.8199	33.4685	99.6156	33.4965	996156	33.4556	99.6036	33.4301
testpat.1k	99.6108	33.4786	99.6124	33.4455	99.6072	33.4347	99.5971	33.4146
Mean	81.8195	28.4887	99.6180	33.4887	99.6093	33.4476	99.5965	33.4322
Std	24.3022	0.06116	0.01957	7.97291	0.01332	0.03371	0.01932	0.03173

4.2 Chaotic encryption performances

4.2.1 Space key

As discussed in section, the encryption key of the proposed encryption approach is constituted of five parts, namely, $x_1(0)$, $x_2(0)$, a, b and M_t. The first four parts are considered as a fraction part having double-precision float number with 52-bit length, conforms to the IEEE 754 standard. The last term M_t stores the iteration number with 8-bit length, the total length of the encryption key is therefore of $52 \times 4 + 8 = 216$-bit. Thus, the cipher key

Table 3 PSNR results of watermarked key-frames using the proposed system

Video	16 × 16 blocks	8 × 8 blocks
V1:Foreman	48.311	44.920
V2:Salesman	49.951	46.854
V3:Hall monitor	48.714	44.123
V4:News	49.015	46.444
V5:Mother-daughter	48.441	45.683
V6:Claire	49.222	46.897
V7:Mobile	47.921	44.777
V8:Coastguard	48.111	45.658
V9:Container	49.851	45.562
V10:Akiyo	48.762	46.907
V11:Silent	49.981	45.938
V12:Carphone	48.745	45.670

space has a high robustness to brute-force attacks [43] since it is similar or better than some state-of-the-art encryption methods and standards [13, 14].

4.2.2 Key sensitivity analysis

To have a high security level, each encryption system must be sensitive to the encryption key. Such sensitivity is generally analyzed with reference to two aspects:

Encryption: here we evaluate the difference between two ciphertext images C^1, C^2 with reference to the same plaintext image using two encryption keys K^1, K^2 different only in one bit.

Decryption: here we measure the difference between two decrypted images D^1, D^2 with reference to the same ciphertext image using two encryption keys K^1, K^2 different only in one bit.

Figure 9 depicts the key sensitivity of the proposed chaotic encryption scheme regarding the encryption and the decryption processes, where K^2, K^3 are slightly different from K^1 with only one bit. These results obviously illustrate that the chaotic encryption based 2D

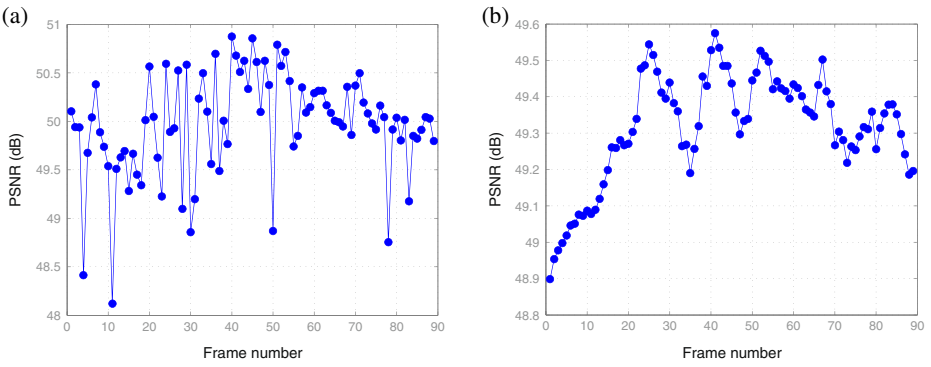


Fig. 11 PSNR curves of all the frames for two watermarked videos: **a** Foreman; **b** Salesman

Fig. 12 Original (*first column*) and watermarked (*second column*) frames

logistic function is highly sensitive to the encryption key for both encryption and decryption processes, meaning that the proposed chaotic encryption has good confusion properties [40].

4.2.3 Histogram analysis

In order to check the image encryption quality, it is very important to evaluate the histogram of encrypted images. A uniformly distributed histogram for ciphertext image is highly needed, since a secure image encryption approach aims to randomize a plaintext image effectively. Figure 10 illustrates different ciphertext histograms extracted from the encrypted images. It is observable that these images cover the format from binary and 8-bit gray. From these results, it is clearly shown that the ciphertext image histograms become very flat after encryption, despite the fact that some plaintext images have highly tilted histograms.

4.2.4 Robustness to differential attack

The robustness to differential attack is another issue that should be addressed in any encryption system. The goal here is to study the effect of difference between inputs and corresponding images outputs [23, 44]. To measure the robustness of an image encryption method against differential attacks, the number of pixel changing rate (NPCR) and the unified average changed intensity (UACI) are evaluated.

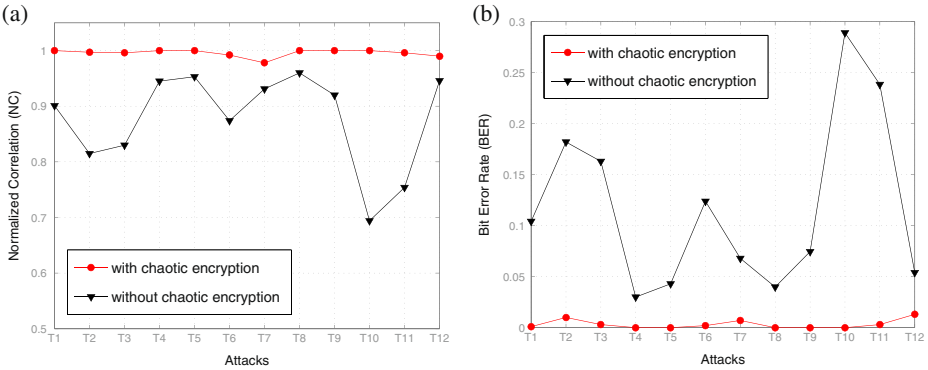


Fig. 13 Effect of using chaotic encryption on watermark robustness against different attacks in term of; **a** NC and, down) BER

Let us consider the plaintext image P, and P^2 as another plaintext image obtained from P by changing one bit of a pixel. Let C^1 and C^2 represent two ciphertext images encrypted from P and P^2, respectively. Therefore, the NPCR and UACI are defined by:

$$NPRC(C^1, C^2) = \sum_{i,j} \frac{A(i,j)}{G} \times 100\% \tag{21}$$

and

$$UACI(C^1, C^2) = \sum_{i,j} \frac{\left|C^1(i,j) - C^2(i,j)\right|}{(L-1) \times G} \times 100\% \tag{22}$$

respectively, where G denotes the total number of pixels, L is the grayscale level, and

$$A(i,j) = \begin{cases} 0, & \text{if } C^1(i,j) = C^2(i,j), \\ 1, & \text{if } C^1(i,j) \neq C^2(i,j), \end{cases} \tag{23}$$

Table 4 NC comparison of different watermarking algorithms against several attacks for 'Foreman' video

Attacks	Scheme in [11]	Scheme in [41]	Scheme in [49]	Scheme in [12]	Scheme in [27]	Scheme in [45]	Proposed
T1:MJPEG	1	0.973	0.95	1	0.54	0.909	1
T2:H.264/AVC (QP = 20)	0.960	0.909	0.871	0.921	1	0.454	0.997
T3:Cropping (15%)	0.663	0.960	0.893	1	0.98	0.909	0.996
T4:Gaussian noise (var = 0.01)	0.982	0.982	0.946	1	0.918	0.979	1
T5:Salt & pepper (var = 0.01)	0.975	0.991	0.951	0.980	0.9	0.979	1
T6:Scaling (100%)	0.670	0.948	0.920	0.952	0.870	0.636	0.992
T7:Blurring	0.941	0.965	0.967	0.974	0.953	0.945	0.978
T8:Sharpening	0.976	0.991	0.9	0.981	0.961	0.909	1
T9:Histogram equalization	1	1	1	1	0.982	1	1
T10:Median filter (3 × 3)	0.918	0.989	0.931	0.991	0.906	0.633	1
T11: Circular filter (radius = 5)	0.933	0.959	0.904	0.939	0.883	0.266	0.966
T12:frame averaging	0.890	0.983	0.96	0.952	1	0.818	0.990

Table 5 BER comparison of different watermarking algorithms against several attacks for 'Foreman' video (in %)

Attacks	Scheme in [11]	Scheme in [41]	Scheme in [49]	Scheme in [12]	Scheme in [27]	Scheme in [45]	Proposed
T1:MJPEG	0	2.69	5.14	0	46.01	9.21	0.1
T2:AVC (QP = 20)	4.13	9.15	12.90	7.91	0	54.6	1
T3:Cropping (15%)	33.69	4.08	10.73	0	2.03	9.1	0.3
T4:Gaussian noise (var = 0.01)	1.84	1.89	5.44	0	8.17	2.09	0
T5:Salt & pepper (var = 0.01)	2.51	0.93	4.92	2.3	9.98	2.1	0
T6:Scaling (100%)	33.12	5.26	8.06	4.84	12.95	36.4	0.2
T7:Blurring	5.95	3.50	3.37	2.62	13.02	5.5	0.7
T8:Sharpening	2.46	0.99	10.07	1.98	3.91	9.1	0
T9:Histogram equalization	0	0	0	0	1.86	0	0
T10:Median filter (3 × 3)	8.22	1.18	6.92	0.97	9.43	36.7	0
T11: Circular filter (radius = 5)	6.77	4.23	9.65	6.09	11.7	73.4	0.3
T12:frame averaging	11.18	1.77	4.03	4.88	0	18.2	1.3

We used the 8-bit grayscale images selected from the USC-SIPI 'Miscellaneous' image dataset to test the robustness of the proposed chaotic encryption approach against some recent and well-known techniques [23]. Table 2 presents the comparison results in term of NPCR and UACI scores. The obtained results show that the characteristics of our approach have excellent performance (NPRC = 99.5965%, UACI = 33.4322). They highly outperform those obtained by Liao's algorithm [28] and they are comparable to those obtained by Hua's [24] and LAS-IES schemes [23]. The proposed approach achieves close average scores of NPCR and UACI to the expected ones provided in [15]. Therefore, we can clearly prove that it has good robustness against differential attack. Furthermore, it should also be noted that our approach presents a low time complexity which will be very helpful for real-time applications.

4.3 Performance with respect to imperceptibility

Imperceptibility tests are essential to perceptual quality assessment since the ultimate judgment is made by human visual perception. Informal visual perception reveals excellent imperceptibility of the embedded watermark using the proposed algorithm. In order to evaluate the quality of a watermarked video, the referenced-based PSNR is evaluated.

The PSNR between query videos and watermarked videos is measured using a block partition of 16×16 and 8×8 for all QCIF video sequences used in the simulation, as shown in Table 3, and using a strength parameter $\alpha = 1$. In fact, the PSNR is measured for only the key-frames of each video sequence to show how the payload affects the imperceptibility. It is worth noting that better PSNR results are obtained using 16×16 blocks for watermark insertion, but gives less payload capacity. An average PSNR value of 48.91 is achieved in this case. On the other hand, the imperceptibility reached by 8×8 blocks remains acceptable as the average PSNR value equals 45.78. Moreover, we can achieve a payload capacity of 99 bps with SVD 16×16 blocks while a capacity of 396 bps can be obtained with SVD 8×8 blocks for CIF video sequence with a resolution of 352×288 like PAL source input format.

To better analyze the visual impact of the watermark insertion process, PSNRs of all the frames for two different watermarked videos (Foreman and Salesman) are plotted in

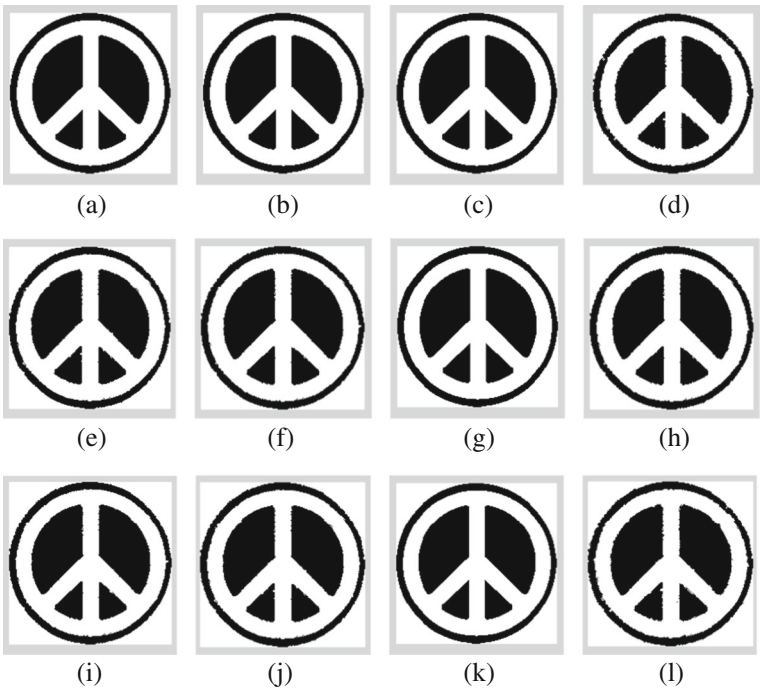


Fig. 14 Extracted watermarks using the proposed algorithm , **a** original watermark, **b** H.264/AVC coding, **c** Salt and pepper noise, **d** MJPEG coding, **e** Cropping (15%), **f** Gaussian noise, **g** Resize, **h** Blurring, **i** Histogram equalization, **j** Median filter, **k** Sharpening, and **l** Circular filter

Fig. 11. The average PSNRs of the two watermarked videos are 48.311 and 49.951 dB, respectively, when only the watermarked key-frames are considered. It is clearly shown that after embedding the watermark information, the PSNR of all the video frames even the watermarked ones are above 48 dB, which ensures a good imperceptibility using the proposed watermarking system. Consequently, we can conclude that combining DWT and SVD with 16×16 block size gives better imperceptibility than using only DWT or SVD as it is proved by the PSNR results.

For subjective observation, the first frames of the two original videos and their corresponding watermarked frames are illustrated in Fig. 12. The results in Figs. 11 and 12 show that the PSNRs of all frames in the two watermarked videos are higher than 48 dB; furthermore, no visible artifacts are observed in the watermarked frames. Therefore, our approach realizes a good imperceptibility.

4.4 Performance with respect to the robustness

4.4.1 Effect of chaotic encryption

In this section we analyze the effect of chaotic encryption on the robustness of the watermarking system. The chaotic encryption is not only used for security issue, but also for adding robustness against the different transformations. In fact, it exploits its randomness capability to distribute burst errors generated by the different attacks in many directions, and

Table 6 Capacity imperceptibility and complexity comparison for 'Foreman' video sequence

Evaluation	Scheme in [11]	Scheme in [41]	Scheme in [49]	Scheme in [12]	Scheme in [27]	Scheme in [45]	Proposed (16 × 16)	Proposed (8 × 8)
Capacity (bits/frame)	99	48	64	32	2	64	99	396
PSNR (dB)	53.37	48.18	50.05	44.85	46.86	48.11	49.82	45.31
Time (sec)	938.11	43.48	584.6	48.37	76.31	136.39	29.16	41.16

then reduce their effects, as it is explained in Section 4.2.1. Figure 13 depicts the effect of using chaotic encryption on the robustness of the proposed watermarking system in terms of NC and BER. Obtained results show that the number of errors after using the chaotic encryption is much lower than that without the chaotic encryption, and the NC of the extracted watermark is very close to 1 when the proposed chaotic encryption is employed for almost the attacks introduced in the simulation. Therefore, we can deduce that the robustness of the proposed approach is also due to the chaotic encryption which is used as a post-processing step.

4.4.2 Comparison study

Tables 4 and 5 present the NC and BER results for different attacks. From theses tables, the proposed scheme can achieve a high robustness, it can outperforms the other recent schemes for almost the attacks used in the evaluation.

The extracted watermarks after different attacks are illustrated in Fig. 14. Obviously, the quality of these watermarks varies from one attack to another one. However, extracted watermarks are remained almost intact after the different attacks which can prove the robustness of the proposed approach.

4.5 Performance with respect to capacity of insertion and complexity

Table 6 illustrates the performances of the proposed watermarking scheme in comparison to other recent schemes in terms of insertion capacity, PSNR and time complexity. From this table, the proposed scheme can achieve a highly payload capacity which can reach 396 bits/frame in contrast to the other schemes. For example, a 32 bits/frame payload can be reached using Faragallah's method [12] and only 2 bits/frame is achieved using Li's approach [27]. We also demonstrate the reduced time complexity of the proposed scheme by comparing the execution time. The computed time is also a matter of only a few seconds. Therefore, the present watermarking system can be a successful candidate for real-time applications.

5 Conclusion

In this paper, we have proposed a watermarking system for uncompressed video sequences for which a watermark information is inserted in the key-frames of the video sequence in a blind manner. For this reason, we have developed a new blind insertion/extraction function that permits a good way to hide the watermark information, and allows its extraction

without using the original video. The proposed system exploits two robust mathematical transforms; DWT and SVD by means of additive method. In addition, a powerful chaotic encryption algorithm is employed to encrypt the watermark information before its insertion in the DWT-SVD domain. Moreover, the new insertion/extraction function allows a blind extraction of the inserted watermark signature.

Encouraging results have been achieved in term of the imperceptibility, robustness, and with respect to the capacity of insertion and complexity. The experimental results demonstrated that the proposed watermarking system is robust to potential attacks such as resolution scaling, blurring, cropping, filtering, H.264 compression, etc. Furthermore, the chaotic encryption algorithm adds a security level to the watermarking scheme. This set of capabilities makes it possible to use the proposed scheme in digital video watermarking applications. The main limitation of our proposed watermarking system is that if a video segment which contains a key-frame is removed, then the watermark cannot be exactly recovered. This issue will be addressed in our future work. Furthermore, more attention will be given to geometrical attacks such as rotation and flipping since almost the watermarking algorithms in the literature fail to resist this kind of attacks. Finally, we will address the possibility to insert the watermark directly in the compressed video stream.

References

1. Abdallah HA, Ghazy RA, Kasban H, Faragallah OS, Shaalan AA, Hadhoud MM, Dessouky MI, El-Fishawy NA, Alshebeili SA, Abd El-samie FE (2014) Homomorphic image watermarking with a singular value decomposition algorithm. Inf Process Manag 50(6):909–923
2. Agilandeeswari L, Ganesan K (2016) A robust color video watermarking scheme based on hybrid embedding techniques. Multimed Tools Appl 75(14):8745–8780
3. Avila SEF, Lopes APB (2011) VSUMM: a mechanism designed to produce static video summaries and a novel evaluation method. Pattern Recogn Lett 32(1):56–68
4. Chen H, Zhu Y (2014) A robust video watermarking algorithm based on singular value decomposition and slope-based embedding technique. Multimed Tools Appl 71(3):991–1012
5. Cisco visual networking index (2012) Forecast and methodology, 2011–2016, Cisco Systems Inc.
6. Cox IJ, Miller ML, Bloom JA, Kalker T, Fridrich J (2008) Digital watermarking and steganography. Morgan Kaufmann
7. Dang C, Radha H (2015) RPCA-KFE: key frame extraction for video using robust principal component analysis. IEEE Trans Image Proc 24(11):3742–53
8. Dawen X, Rangding W, Shi YQ (2014) Data hiding in encrypted h.264/AVC video streams by codeword substitution. IEEE Trans Inf Forensics Security 9(4):596–606
9. Dutta T, Sur A, Nandi S (2013) Mcrd: motion coherent region detection in h.264 compressed video. In: Proceedings international conference on multimedia and expo (ICME). San Jose, CA, USA, pp 1–6
10. Ejaz N, Mehmood I, Baik SW (2014) Feature aggregation based visual attention model for video summarization. Comput Electr Eng 40:993–1005
11. El'Arbi M, Koubaa M, Charfeddine M, Ben Amar C (2011) A dynamic video watermarking algorithm in fast motion areas in the wavelet domain. Multimed Tools Appl 55(3):579–600
12. Faragallah OS (2013) Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain. Int J Electron Commun (AEÜ) 67(3):189–196
13. FIPS PUB 197 (2001) Advanced encryption standard. New York
14. FIPS PUB 46 (1977) Data encryption standard
15. Fu C, Chen JJ, Zou H, Meng WH, Zhan YF, Yu YW (2012) A chaos-based digital image encryption scheme with an improved diffusion strategy. Opt Express 20(3):2363–2378
16. Furini M, Geraci F, Montangero M, Pellegrini M (2010) STIMO: STILl and MOving video storyboard for the web scenario. Multimed Tools Appl 46(1):47–69
17. Gaj S, Patel AS, Sur A (2016) Object based watermarking for h.264/AVC video resistant to rst attacks. Multimed Tools Appl 75(6):3053–3080

18. Ganic E, Eskicioglu AM (2004) Robust DWT-SVD domain image watermarking: embedding data in all frequencies. In: Proceedings the workshop on multimedia and security (MM&Sec). Magdeburg, Germany, pp 166–174
19. Himeur Y, Ait-Sadi K (2015) Joint color and texture descriptor using ring decomposition for robust video copy detection in large databases. In: Proceedings the 15th IEEE international symposium on signal processing and information technology (ISSPIT). Abu Dhabi, UAE, pp 495–500
20. Himeur Y, Ait-Sadi K, Ouamane A (2014) A fast and robust key-frames based video copy detection using BSIF-RMI. In: Proceedings the 11th international conference on signal proceedings and multimedia application (SIGMAP). Vienna, Austria, pp 40–47
21. Himeur Y, Boudraa B (2012) Secure and robust audio watermarking system for copyright protection. In: Proceedings the 24th international conference on microelectronics (ICM). Algiers, Algeria, pp 1–4
22. Hu HT, Hsu LY (2015) Exploring DWT-SVD-DCT feature parameters for robust multiple watermarking against JPEG and JPEG2000 compression. Comput Electr Eng 41(1):52–63
23. Hua Z, Zhou Y (2016) Image encryption using 2D logistic-adjustedSine map. Inf Sci 339:237–53
24. Hua Z, Zhou Y, Pun CM, Chen CLP (2015) 2D sine logistic modulation map for image encryption. Inf Sci 297:80–94
25. Lai CC, Tsai CC (2010) Digital image watermarking using discrete wavelet transform and singular value decomposition. IEEE Trans Instrum Meas 59(11):3060–3063
26. Langelaar G, Setyawan I, Lagendijk RL (2000) Watermarking digital image and video data. A state-of-the-art overview. IEEE Signal Proc Mag 17(5):20–43
27. Li Z, Chen XW, Ma J (2015) Adaptively imperceptible video watermarking based on the local motion entropy. Multimed Tools Appl 74(8):2781–2802
28. Liao X, Lai S, Zhou Q (2010) A novel image encryption algorithm based on self-adaptive wave transmission. Signal Process 90:2714–2722
29. Lin W, Sun MT, Li H, Chen Z, Li W, Zhou B (2012) Macroblock classification method for video applications involving motions. IEEE Trans Broadcast 58(1):34–46
30. Mishra A, Agarwal C, Sharma A, Bedi P (2014) Optimized gray-scale image watermarking using DWT–SVD and firefly algorithm. Expert Syst Appl 41(17):7858–7867
31. Muhammad K, Sajjad M, Baik SW (2016) Dual-level security based Cyclic18 steganographic method and its application for secure transmission of keyframes during wireless capsule endoscopy. J Med Syst 40:114
32. Muhammad K, Sajjad M, Mehmood I, Rho S, Baik SW (2016) A novel magic LSB substitution method (m-LSB-SM) using multi-level encryption and achromatic component of an image. Multimed Tools Appl 75(22):14867–893
33. Muhammad N, Bibi N (2015) Digital image watermarking using partial pivoting lower and upper triangular decomposition into the wavelet domain. IET Image Proc 9(9):795–803
34. Mundur P, Rao Y, Yesha Y (2006) Keyframe-based video summarization using Delaunay clustering. Int J Digit Libr 6(2):219–232
35. Preda RO, Vizireanu DN (2011) Robust wavelet-based video watermarking scheme for copyright protection using the human visual system. J Electron Imaging 20(1):13–22
36. Ren J, Jiang J, Chen J (2009) Shot boundary detection in MPEG videos using local and global indicators. IEEE Trans Circuits Syst Video Techn 19(8):1234–1238
37. Sarkar A, Singh V, Ghosh P, Manjunath B, Singh A (2010) Efficient and robust detection of duplicate videos in a large database. IEEE Trans Circuits Syst Video Techn 20(6):870–885
38. Seyedzadeh SM, Mirzakuchaki S (2012) A fast color image encryption algorithm based on coupled two-dimensional piecewise chaotic map. Signal Process 92(5):1202–1215
39. Shahid Z, Chaumont M, Puech W (2013) H.264/AVC video watermarking for active fingerprinting based on Tardos code. Signal Image Video Proc 7(4):679–694
40. Shannon CE (1949) Communication theory of secrecy systems. Bell Syst Tech J 28(4):656–715
41. Singh RT, Manglem Singh K, Roy S (2013) Video watermarking scheme based on visual cryptography and scene change detection. Int J Electron Commun (AEÜ) 67(8):645–651
42. Sprott JC (2003) Chaos and time-series analysis. Oxford University Press, London
43. Stinson DR (2006) Cryptography: theory and practice. Chapman and Hall CRC, Boca Raton
44. Wu Y, Zhou Y, Noonan JP, Agaianc S (2014) Design of image cipher using latin squares. Inf Sci 264:317–339
45. Xiang-yang W, Yu-nan L, Shuo L, Hong-ying Y, Pan-pan N, Yan Z (2015) A new robust digital watermarking using local polar harmonic transform. Comput Electr Eng 46:403–418
46. Xue W, Zhang L, Mou X, Bovik A (2014) Gradient magnitude similarity deviation: a highly efficient perceptual image quality index. IEEE Trans Image Proc 23(2):684–695

47. Xuemei J, Quan L, Qiaoyan W (2013) A new video watermarking algorithm based on shot segmentation and block classification. Multimed Tools Appl 62(3):545–560
48. Yavuz E, Yazıcı R, Kasapbaşı MC, Yamaç E (2015) A chaos-based image encryption algorithm with simple logical functions. Comput Electr Eng 54:471–483
49. Youssef SM, Abou ElFarag A, Ghatwary NM (2014) Adaptive video watermarking integrating a fuzzy wavelet-based human visual system perceptual model. Multimed Tools Appl 73(3):1545–1573
50. Youtube: Statistics, [online]. Available: http://www.youtube.com/yt/press/statistics.html
51. Zhang M, Tong X (2014) A new chaotic map based image encryption schemes for several image formats Original. J Syst Softw 98:140–154
52. Zhou G, Zhang D, Liu Y, Yuan Y, Liu Q (2015) A novel image encryption algorithm based on chaos and line map. Neurocomputing 169:150–157

Dr. Yassine Himeur received the Master degree in electronic from The University of Science and Technology–Houari Boumediene (USTHB), Algiers, Algeria in 2011 and the PhD degree from Mohammed Seddik Benyahia University of Jijel in 2015. Currently, he is a Senior researcher at Centre de Développement des Technologies Avancées (CDTA), in Algiers, Algeria. His current research interests are: Multimedia security, Multimedia retrieval and Powerline Communication. He has authored several papers in refereed journals and international conference proceedings. He is the recipient of the International Conference on Signal Processing and Multimedia Applications (SIGMAP 2014) Best Paper Award.

Prof. Abdelkrim Boukabou obtained both his Dr. Eng. and habilitation of electronics at Constantine University in 2006 and 2008 respectively. In 2013 he became full professor at Mohammed Seddik Benyahia University of Jijel. Since 2009 Boukabou has been Director of Post Graduate Research in electronics department. His research interests include nonlinear control theory, robotics and automation, power-line communications, and smart grids. He has authored several papers in refereed journals and international conference proceedings. He is also the recipient of the International Conference on Signal Processing and Multimedia Applications (SIGMAP 2014) Best Paper Award.