Circuits, Systems, and Signal Processing (2024) 43:264–301
https://doi.org/10.1007/s00034-023-02462-8

Blind Semi-fragile Hybrid Domain-Based Dual Watermarking System for Video Authentication and Tampering Localization

Amal Hammami[1] · Amal Ben Hamida[1] · Chokri Ben Amar[1] · Henri Nicolas[2]

Received: 6 July 2022 / Revised: 17 July 2023 / Accepted: 18 July 2023 /
Published online: 4 August 2023

Abstract

In this paper, a blind semi-fragile dual watermarking system for video content authentication and tampering localization is proposed. In this method, two watermarks are tailored for each host video frame. Indeed, the frame index is firstly binary modulated and then, serves as a primary watermark. Subsequently, a second content-based authentication watermark is built up using frame blocks texture features. To improve the security aspect, Torus automorphism mapping is applied under the second watermark before being embedded. The suggested scheme operates in the hybrid domain. In fact, the primary watermark is infused in the host frame following a spatial domain-based embedding technique. However, a frequency domain-based watermarking method, which combines Lifting Wavelet Transform (LWT) and Singular Value Decomposition (SVD), is involved to hide the second watermark. Herein, the most textured blocks are chosen as second watermark holders to enhance the watermarked video perceptual quality. During the detection process, the dissimilarity between the second extracted watermark and its reconstructed version as well as the mismatch between the observed index and the extracted one enable to reveal spatial and temporal tampering, respectively. Experimental results show that the proposed scheme achieves a good visual quality with a large embedding capacity. Furthermore, it can efficiently withstand

✉ Amal Hammami
amal.hammami@enis.tn

Amal Ben Hamida
amal.benhamida@enis.tn

Chokri Ben Amar
chokri.benamar@ieee.org

Henri Nicolas
henri.nicolas@u-bordeaux.fr

[1] REsearch Groups in Intelligent Machines, University of Sfax, National Engineering School of Sfax, Sfax 3038, Tunisia

[2] Bordeaux Computer Science Research Laboratory, University of Bordeaux 1, 33405 Talence, France

non-malicious processing while being fragile to frames manipulations and content modification attacks. Moreover, it ensures an accurate tampered areas localization.

Keywords Blind semi-fragile video watermarking · Content authentication · Tampering detection and localization · Lifting wavelet transform · Singular value decomposition · Torus automorphism mapping

1 Introduction

In recent years, the deployment of digital videos is proliferating rapidly. In fact, this multimedia medium is applied in different applications such as intelligent surveillance, event of interest detection and future behaviors inspection. Besides, videos often serve as criminal evidences and insurance claims especially in surveillance context. Hence, stored videos credibility and authenticity must be guaranteed. Tackling this issue is a challenging concern notably with the multimedia technologies tremendous development. Indeed, the emergence of advanced editing software makes it increasingly easy for unauthorized users to forge video content without a noticeable perception quality degradation. Therefore, reliable content authentication procedure becomes an inevitable need in the digitally advanced community. Digital encryption, which converts the original video content into unreadable version, has been widely used as a solution to protect videos authenticity and trustworthiness [13, 19, 48]. However, this traditional authenticity protection technique exhibits many shortcomings. As first, preserving the actual video content is one of the most important requirements in practical applications due to the fast transmission [47]. Moreover, encryption methods do not enable to accurately locate tampered areas [49].

To address these defects, digital watermarking has been introduced and has become an important research discipline in the security field [16, 21, 39]. A watermarking system usually consists of two processes: an embedding process and an extraction one. The first procedure encapsulates a secret information referred as watermark into the host carrier content. The to-be-hidden watermark may be an image, a random binary sequence or a message. The extraction process allows recovering the inserted information from the watermarked content [21]. Watermarking techniques are widely exploited in many applications such as copyright protection, broadcast monitoring and data authentication [4]. Generally, a watermarking scheme is acknowledged in terms of three main properties namely (i) robustness: the hidden watermark should be resilient to several attacks; (ii) imperceptibility: the watermark embedding should not yield any conspicuous difference between the original content visual quality and the watermarked one; (iii) capacity: it determines the data amount that can be embedded into the host multimedia content [11, 24]. Obviously, the aforementioned characteristics are antagonist. In fact, increasing the capacity degrades significantly the watermarked content quality as well as diminishes the watermark robustness against attacks. Thus, an effective watermarking system must assure, depending on the dedicated application, a reasonable compromise between these properties [22].

Based on the robustness criterion, watermarking algorithms are split into three classes: robust, fragile and semi-fragile watermarking schemes [4]. Robust schemes have the ability to recover the watermark after undergoing modifications by any attack. Conversely, the fragile watermarking systems are designed to be susceptible to any small distortion. Semi-fragile watermarking is known as a robust and fragile algorithms modified version. For this watermarking type, the discrimination between malicious and non-malicious attacks is an important requirement. Indeed, semi-fragile schemes are introduced to allow legitimate manipulations including noises addition and compression while rejecting malicious distortions such as frames content changing by object deletion or insertion [22]. The main advantage of semi-fragile watermarking adoption in video content authentication purpose is that it jointly provides a greater potential for tamper modifications characterization and a good resilience level to certain manipulations especially the compression process, which represents a demanding feature in practical applications.

In addition to the robustness criterion, blindness is also commonly used to classify the watermarking systems as non-blind, semi-blind and blind ones [23]. A non-blind scheme necessitates the knowledge of the original video as well as the original watermark in order to extract the concealed information from the watermarked content. Actually, non-blind watermarking algorithms are not appropriate for practical applications where the non-watermarked cover video version is not available at the decoder. As far as semi-blind systems are concerned, only the original watermark is required during the extraction process. Besides, it is stated that these two categories, i.e., the non-blind and semi-blind schemes, lead for several information storage and bandwidth consumption issues as the original video version and/or the secret signature need to be saved and transmitted along with the watermarked sequence. Lastly, neither the host video nor the original watermark are required to successfully extract the hidden information by a blind detector. This characteristic makes the blind scheme gaining more popularity compared to the other two watermarking classes as it avoids the storage and bandwidth overload problems.

In this work, a blind semi-fragile dual watermarking scheme for video content authentication with a high robustness level to common processing, an efficient intentional attacks localization ability and a good watermarked video quality is proposed. The developed technique involves simultaneously the Least Significant Bit (LSB) technique, Torus automorphism mapping, the Lifting Wavelet Transform (LWT) and the Singular Value Decomposition (SVD). The remainder of this paper is framed as follows. A review on video watermarking techniques is provided in Sect. 2. LWT, SVD, Torus automorphism mapping and texture features are briefly introduced in Sect. 3. Section 4 explains in detail the proposed scheme and the associated processes. Performance evaluation results and comparative analysis are investigated in Sect. 5. Finally, Sect. 6 deals with this work conclusions and futures perspectives.

2 Related Works

As reported earlier, video watermarking has become an interesting research focus in the information security field in last decades. Thereby, various video watermarking techniques have been proposed in the literature. These techniques often differ in their design strategies, although they rely on some similar features. In fact, according to the working domain, acknowledged also as the embedding domain, watermarking methods are set into the spatial, the frequency, or the hybrid techniques [30].

Watermarking techniques that perform in the spatial domain conceal the watermark bits by directly adjusting the host video frames pixels values. In [3], a spatial domain-based watermarking scheme for content authentication and tamper detection is presented. In order to encapsulate the watermark, the authors use LSB technique, which consists in substituting the least significant bits of the original video frame pixels with the watermark bit. Simulations results show that this technique gives a good imperceptibility and an efficient forged areas detection. Mohammed A.A et al. introduce another video watermarking scheme performing in the spatial domain [35]. In this algorithm, the watermark is inserted via an additive method in selected lowest pixels values from the luminance component of the host frame. According to test results, this scheme has a high compression resilience with a good preserved watermarked video quality. In [36], a watermarking algorithm is proposed to verify the digital video content integrity and authenticity. In order to fulfill security requirement, the watermark is ciphered by XOR-ing it with an arbitrary image generated using Arnold Cat Map [1]. Then, the encrypted watermark is embedded following a LSB pixel-wise algorithm. Simulation experiments demonstrate the ability of this scheme to identify the modified regions in maliciously manipulated frames. The video watermarking proposed by Guangxi et al. in [8] is executed in the spatial domain by fine-tuning the luminance component pixels values. This method includes an embedding area selection strategy based on luminance adaptive and edge mask. The proposed approach can withstand a variety of attacks such as compression and scaling. Besides, it achieves a good imperceptibility. Watermark embedding in the spatial domain is recognized with its simplicity and inexpensive computational complexity. Furthermore, it does not lead to a significant distortion to the cover video quality. As a result, this technique is more convenient for applications in low computing power environment.

In contrast to the former, frequency watermarking schemes initially convert pixels values into the transform domain using a specific transformation. Then, the watermark is embedded into certain frequency bands coefficients. A wide range of transformations, including but not limited to Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), LWT and SVD, are extensively applied in the watermarking field. A watermarking scheme involving the DWT is proposed in [41]. For transparency improvement, the watermark is infused only on a single frame per scene. In fact, this watermarking methodology applies a scene-change detector to identify the scene-changed frames that are used as watermark information carriers. Then, these latter are decomposed using the 3 levels DWT and the to-be-hidden watermark is inserted in the LL sub-band coefficients. This approach can successfully sustain different manipulations. Furthermore, it yields watermarked videos with a good quality. Another DWT-based semi-fragile watermarking approach for video authentication is proposed in [46]. The authentication-based watermark used in this method is a noise-like image generated by arranging a random binary sub-sequence bits in the spectral domain. The built up watermark is embedded in the DWT frequency domain using a frame-by-frame watermarking algorithm. Based on the experiments results, the suggested design is convenient for solving authentication task as well as localizing the tampered frames.

Ponni and Ramakrishnan introduce a video watermarking combining the DWT and the SVD [40]. They use a scene change detection technique to identify key scenes that exhibit frequent changes. Later, key frames are chosen based on Fibonacci sequence and used as watermark information holders. The insertion is fulfilled by embedding the singular value of the watermark, which is encrypted using Fibonacci–Lucas transform [51], in the selected frames singular values. This methodology is robust against various attacks while achieving a good imperceptibility level. A modified version of the above-described scheme is proposed in [38]. Likewise, this technique performs in the transform domain using the DWT and the SVD. This scheme differs from the method in [40] by embedding the principal components of the watermark instead of the singular values in the specific selected frames. This technique satisfies the robustness and imperceptibility aspects. Relying on the same general concept, Abdulla and Navas advance a semi-blind watermarking scheme in multi-frequency domain [2]. In this methodology, only the lower entropy frames are used as watermark holders in order to reduce the yielded visual distortion. Under this framework, two mathematical transforms namely 2-level DWT and SVD are jointly exploited through an additive method to conduct the insertion function. Based on the simulations outcomes, this system successfully preserves the perceptual quality of the watermarked videos. Furthermore, it performs robustly under numerous manipulations.

Bhardwaj et al. present a video watermarking system in the LWT domain [6]. In this scheme, the authors exploit the mathematical relationship among the frames number in the cover video, the coefficient block size and the capacity to select a few frames for the insertion process. LWT is performed to each selected frame luminance, which is decomposed into different frequency sub-bands. Then, the watermark is injected into the lower frequency band via a quantization method. This scheme achieves a noteworthy robustness and imperceptibility enhancement. Another frequency video watermarking approach using SVD and Multiresolution Singular Value Decomposition (MSVD) is employed in [37]. Only significant frames, which are identified based on motion estimation, are chosen to host the to-be-hidden information. This latter is firstly scrambled and then, blindly inserted using a Quantization Index Modulation-based algorithm [7]. According to the experimental findings, this method gives a very good imperceptibility as well as a high robustness to different geometric and image processing-based attacks.

A semi-fragile watermarking system for video content authentication is provided in [15]. The used authentication signature in this algorithm is composed of the timing information and invariant features, which are extracted by a spatial analysis of macroblocks. The watermark embedding is executed by flipping the sign of nonzero QDCT (Quantized DCT) coefficients in a set of random selected Group of Pictures (GOP). The prime advantage of this methodology is that it is insensitive to mild processing while being sensitive to intentional distortions. Likewise, another semi-fragile watermarking approach using DCT is suggested in [10]. The frame number and the relationship among the DCT nonzero coefficients are used to generate the watermark, which is concealed in the medium frequency sub-bands coefficients. This algorithm has a negligible effect on the watermarked video quality. Besides, it shows a good immunity against non-malicious manipulations. Similarly, a chromatic residual DCT-based watermarking scheme is presented in [45]. In this system, an authentication data are generated through the intra-prediction mode categorization. Actually, the chroma blocks with the higher number of nonzero (NNZ) items are selected as watermark holders. In fact, the watermark is hidden within the mid-frequency quantized AC coefficients of chromatic 4 × 4 sub blocks. The simulations results demonstrate that this method is resilient against compression while exhibiting a high sensitivity toward malicious tampering at both the temporal and spatial level. Another frequency domain-based watermarking solution is introduced in [31]. Herein, the watermark holders are chosen tacking into account the human visual system specifications. Then, a three-dimensional discrete cosine transform-based embedding algorithm is involved to hide the watermark information in the selected host blocks pixels. The obtained results show that this approach offers a reasonable sustainability level as well as good watermarked videos quality. Although frequency domain-based techniques are high in computation, they exhibit an efficient robustness to a wide variety of attacks.

The third category of watermarking is hybrid domain-based techniques. The watermark embedding is fulfilled combining the two areas viz. the spatial domain and the frequency one. In [28], authors use both of LSB and DCT to develop a new video watermarking scheme. A first watermark is concealed in the blue channel of the host frame using LSB as spatial domain-based embedding technique. Then, two levels DCT is applied to the initially watermarked frame and next, the second watermark is inserted into each transformed coefficients. This method shows a high invisibility level and a good resilience to several manipulations. Another hybrid video watermarking algorithm, which combines three transformations together, i.e., LSB, DWT and DCT, is presented in [14]. A first mark is doubly inserted in the host frame using DCT and LSB-based embedding techniques. The final watermarked frame is obtained by hiding a second watermark in the initially marked one following an embedding algorithm in the DWT domain. Evaluation results ascertain the robustness and the invisibility criteria of this scheme. Kerbiche et al. develop a similar video watermarking scheme, which jointly uses DWT, SVD and LSB [30]. In this methodology, crowdsourced regions and moving objects, both generated from the mosaic host video frames, are adopted as watermarks carriers. Since two signatures are considered, a DWT-SVD-based insertion algorithm as well a LSB based one are carried out during the embedding process. This approach has a high robustness and a good imperceptibility.

Algorithms based on hybrid domain are more suitable for watermarking schemes that involve multi-watermarks embedding because they guarantee high imperceptibility and robustness levels as well as a less computational complexity. For this reason, the dual watermarking approach proposed in this paper operates in the hybrid domain. Indeed, LSB-based embedding method is used as a combination with LWT-SVD based one to carry out the insertion of two distinct watermarks in each host video frame.

3 Preliminaries

LWT, SVD, texture features and Torus automorphism mapping are simultaneously utilized as fundamental components to build up the proposed watermarking scheme. A brief description of each of these components is given in the following subsections.

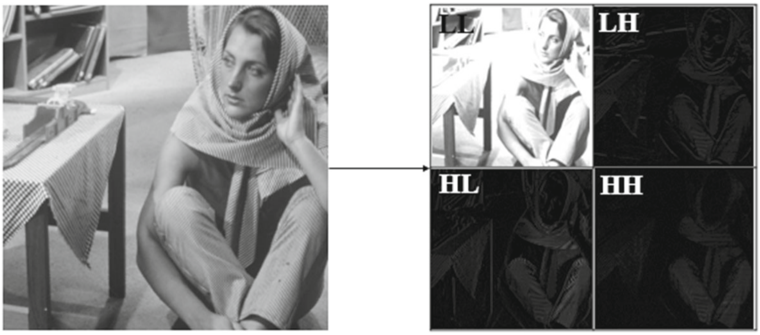


Fig. 1 Lifting wavelet transform application to Barbara image

3.1 Lifting Wavelet Transform (LWT)

The lifting wavelet transform was initially introduced by Sweldens as an alternative concept for traditional DWT. This domain transform technique produces reversible integer wavelet and scaling coefficients rather than floating-point ones. Hence, it avoids aliasing effects and information loss. Allowing a wavelet transform fully in-place calculation, lifting scheme necessitates the half computations number that is involved in conventional wavelet transform.

The signal decomposition into frequency sub-bands by LWT is achieved using iteration operations of three fundamental steps that are termed as splitting, prediction and update. Readers can refer to [12, 44] for the detailed lifting scheme mathematical description. For visual perception, Fig. 1 depicts Barbara image decomposition using lifting scheme. LWT is known for its attractive properties including a good robustness to noise addition, a less computation complexity, a reduced distortion and a better frequency localization [34]. These features make LWT suitable for several applications such as denoising [9], features extraction [25] and watermarking [6].

3.2 Singular Value Decomposition (SVD)

SVD is a well-known algebraic numerical tool used for complex matrix decomposition. First, SVD was proposed by Beltrami and Jordan only for square matrices, and then, Eckart and Yong have generalized this technique to handle rectangular matrices [33]. For any matrix M of an arbitrary size m*n, the singular value decomposition can be expressed using the following formula:

$$M = U \times S \times V^{t} \tag{1}$$

where S is a diagonal matrix that contains only positive values called the singular values and arranged in descending order. U and V are orthogonal matrices termed as the left-hand side singular vectors and the right-hand side ones, respectively. In image analysis, singular values matrix defines the image luminance, while the corresponding singular vectors pair describes the image geometry. SVD becomes an interesting algebra transform for signal processing area owing to its prominent properties such as its maximum energy packing, multivariate analysis and adaptation to distortions ability [23, 27, 29]. It is worth mentioning that after applying SVD the obtained U matrix holds the main and greatest coefficients values and it exhibits two interesting characteristics [5]. Indeed, all elements in its first column have the same sign. Moreover, these elements are strongly correlated.

3.3 Texture Features

Texture is recognized as an important image descriptor that is broadly applied in several pattern recognition and computer vision applications. Up to now, no unique texture definition has been proposed by researchers. However, it can be defined as a set of complex visual patterns. These latter are composed of spatially arranged entities having color, brightness, size or shape as well as distinguished by heaviness, uniformity and smoothness [42]. Several texture features are proposed in the literature. Energy, entropy and kurtosis are among the commonly texture features used to classify image as a single entity or its blocks as sub-entities. Energy, calculated using (2), is among the statistical measures that provide most meaningful texture information [17].

$$\text{Energy} = \sum_{i=1}^{n} \sum_{j=1}^{m} I^2(i, j) \tag{2}$$

where I is an image of $m*n$ size. The energy amount describes the uniformity of the distribution. In a blocks cluster, textured blocks are identified by higher energy values compared to their neighbors, while non-textured ones are recognized by lower energy values.

Entropy is a statistical metric, which measures the randomness of image content dispersion. It can be expressed using the following formula:

$$\text{Entropy} = -\sum_{i=1}^{255} p(x_i) \log_2 p(x_i) \tag{3}$$

where p(x_i) is the symbol x_i occurrence of probability. Entropy is an efficient indicator for image information magnitude [32]. High entropy value reveals a randomly dispersed intensities distribution; thus, it indicates a textured block. In contrast, low entropy value is associated with uniform intensities distribution through untextured block.

Kurtosis represents the normalized fourth-order moment [20]. It measures the flatness of a distribution. Textured blocks, which have a significant deal of information, are identified by low kurtosis values. Contrariwise, high kurtosis value designates a non-textured block. The kurtosis mathematical expression is given in (4).

$$\text{Kurtosis} = \sigma^{-4} \sum_{i=1}^{255} (x_i - \mu)^4 p(x_i) - 3 \tag{4}$$

where p(x_i) is the symbol x_i occurrence of probability, $\mu = \sum_{i=1}^{n} x_i\, p(x_i)$ is the pixels mean value and $\sigma = \sqrt{(x_i - \mu)^2 \times p(x_i)}$ is the variance square root.

3.4 Torus Automorphism Mapping

Torus automorphism mapping represents a chaotic map version that allows spatially dispersing pixels locations [18]. In this technique, for each point only one unique mapping point is assigned. Torus automorphism mapping expression corresponding to one dimension sequence is defined as:

$$X_{\text{mapped}} = (K \times X) \mod (L) + 1 \tag{5}$$

X and X_{mapped} represent the bit positions before and after the Torus automorphism mapping. L is the sequence length, and K denotes the mapping secret key. Torus automorphism is commonly adopted as content scrambling method especially in watermarking application to enhance the embedded watermark security [22, 32].

4 Proposed Approach

The proposed approach is a semi-fragile dual video watermarking system devoted for content authentication and tampering detection and localization. As shown in Fig. 2, this scheme consists of three processes: the watermarks construction, the watermarks embedding and the detection.

In particular, the main contributions in this work are outlined as below:

1. A content-based watermarks generation procedure is proposed. Indeed, in the advanced dual watermarking technique two distinct authentication information are designed using frame content characteristics and then, individually encapsulated in each host video frame. The built up content-based watermarks permit to fulfill the tasks of (i) discrimination between intentional manipulations and non-intentional ones for a faithful authenticity verification, (ii) the detection of inter-frames as well as intra-frame forgeries and (iii) the localization of maliciously tampered areas.
2. To improve the overall proposed scheme imperceptibility attribute, a texture analysis-based block classification method is introduced. This technique permits to optimally choose the watermark holders in concordance with the given video frame characteristics and thereby to lessen the visual artifacts introduced by the embedding process.
3. To proficiently establish the trade-off between the different semi-fragile watermarking requirements, a hybrid domain-based watermarks embedding procedure is proposed. In fact, a LSB-based insertion method is employed to successfully hide the first watermark in the spatial domain. However, the second watermark is embedded using a multi frequential algorithm utilizing both of the LWT frequency bands specifications and the SVD coefficients characteristics to strengthen the scheme performance. The established embedding algorithms not only permit to successfully infuse the watermarks data in the host video frames but also to guar-

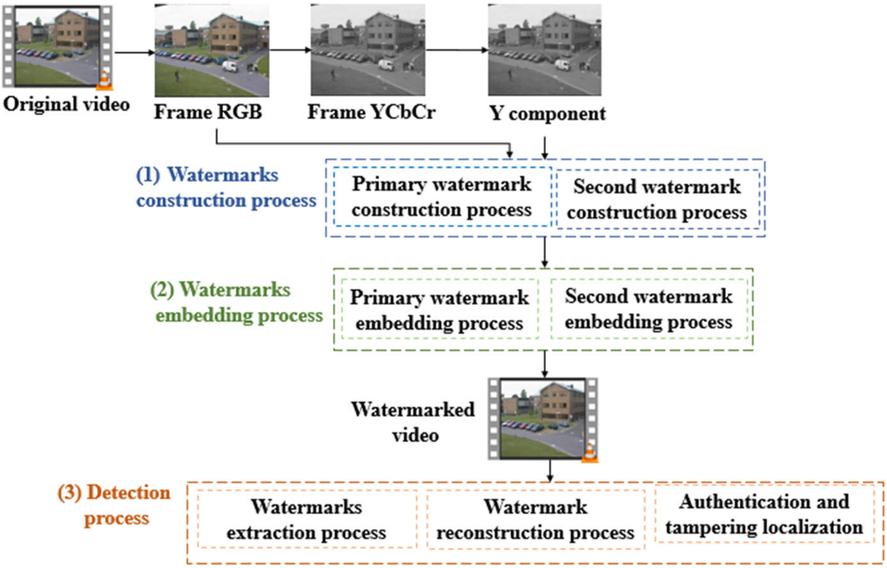

Original video
Frame RGB
Frame YCbCr
Y component
(1) Watermarks construction process
Primary watermark construction process
Second watermark construction process
(2) Watermarks embedding process
Primary watermark embedding process
Second watermark embedding process
Watermarked video
(3) Detection process
Watermarks extraction process
Watermark reconstruction process
Authentication and tampering localization

Fig. 2 Overall flow diagram of the proposed video watermarking approach

antee a blind detection at the receiving side. Indeed, the proposed detector is fully capable of blindly extracting the two-hosted watermarks. This allows avoiding information storage and bandwidth consumption issues as neither the unwatermarked video version nor the secret signature need to be saved and transmitted along with the watermarked sequence.

4. To boost the watermark ability to precisely locate spatial distortions, a glide window-based detection strategy is proposed. It allows investigating the inconsistency between the extracted and the reconstructed second watermark versions for a rigorous content authentication with an accurate tamper localization.

The watermarks construction, embedding and detection processes are separately explored in the incoming subsections.

4.1 Watermarks Construction Process

The proposed scheme is a dual watermarking scheme, in which two watermarks are tailored for each video frame, as illustrated in Fig. 3, to detect and locate each tampering type.

In fact, video forgery techniques are often classified to two sets namely intra-frame and inter-frames forgery [43]. The former, also termed as spatial tampering, refers to the host frame content modification such objects manipulation through object addition or deletion. Inter-frames forgeries, also called temporal tampering, designate a manipulation conducted at time level such as frames dropping, replacing or swapping.

The watermark construction process starts by dividing the host video into RGB frames. Subsequently, the primary watermark is built up using the current frame index

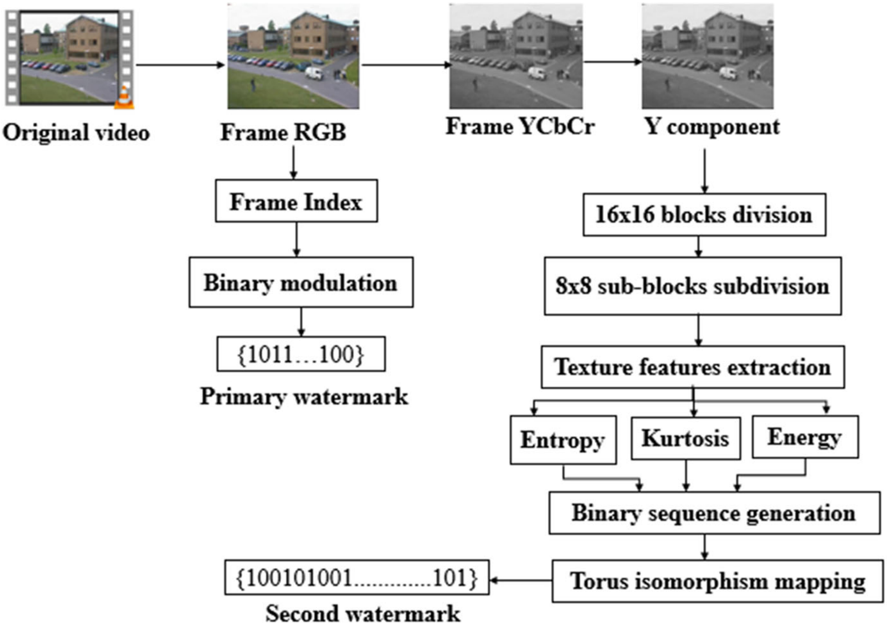


Fig. 3 Overall flow diagram of the proposed watermarks construction process

as timing information. Indeed, the frame index is modulated to a binary sequence for convenient insertion. Then, it is served as watermark to detect and locate temporal tampering.

Next, a block-based technique is adopted to generate the second watermark, which is used to reveal spatial tampering. First, each RGB frame is converted to YCbCr space color. Y component, which allows ensuring the best compromise between the robustness and the imperceptibility, is segmented into non-overlapping blocks of 16*16 size. Following, a texture analysis is applied to each block. In fact, each 16*16 block is further decomposed into four sub-block of 8*8 size. Then, three texture features including the energy, entropy and kurtosis are investigated to evaluate the texture aspect of the segmented sub-blocks as follows:

step1: For each sub-block among the four, ones associated with the same block, calculate the energy, entropy and kurtosis values denoted, respectively, by V_{energy}, V_{entropy} and V_{kurtosis}.

step2: Calculate $Mean_{\text{energy}}$, $Mean_{\text{entropy}}$ and $Mean_{\text{kurtosis}}$ the average of each feature values.

step3: Associate a Boolean variable B to each feature and compute its value based on the following rules:

If $V_{\text{energy}} > Mean_{\text{energy}}$: $B_{\text{energy}} = 1$; Otherwise, $B_{\text{energy}} = 0$
If $V_{\text{entropy}} > Mean_{\text{entropy}}$: $B_{\text{entropy}} = 1$; Otherwise, $B_{\text{entropy}} = 0$
If $V_{\text{kurtosis}} < Mean_{\text{kurtosis}}$: $B_{\text{kurtosis}} = 1$; Otherwise, $B_{\text{kurtosis}} = 0$

step4: Identify the sub-blocks texture status based on the rules illustrated in Table 1.

Table 1 Used rules for sub-blocks texture identification

B_{energy}	B_{entropy}	B_{kurtosis}	Texture status
0	0	0	Non-Textured
0	0	1	Non-Textured
0	1	0	Non-Textured
0	1	1	Textured
1	0	0	Non-Textured
1	0	1	Textured
1	1	0	Textured
1	1	1	Textured

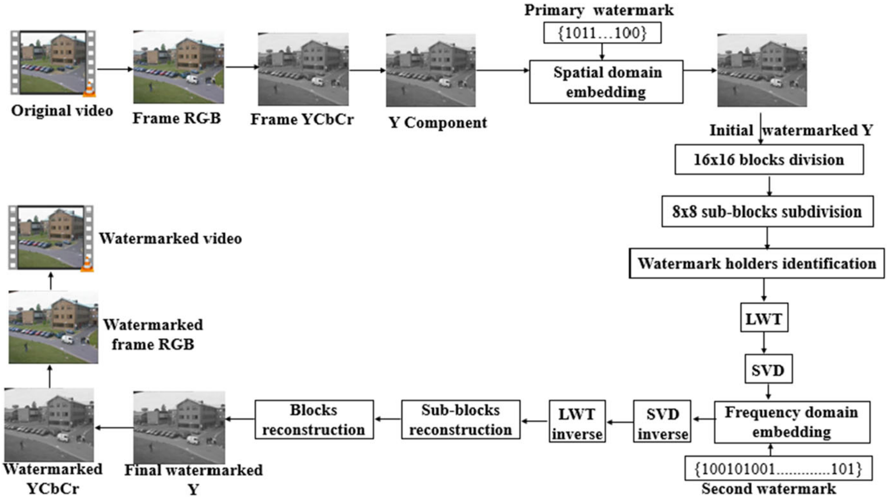


Fig. 4 Overall flow diagram of the proposed watermarks embedding process

After classifying all the four sub-blocks based on texture aspect, the corresponding block is regarded as textured block if it holds more than one textured sub-block and thereby, the current generated watermark bit is 1. Otherwise, it is labeled as non-textured block and the corresponding watermark bit is 0. All the generated bits are combined in one-dimensional binary sequence, which is encrypted using Torus automorphism mapping to further improve the scheme security aspect. Finally, the scrambled sequence is used as a second watermark.

4.2 Watermarks Embedding Process

As mentioned previously, two different watermarks are embedded in each video frame. Since the proposed watermarking scheme is a hybrid domain-based approach, the watermarks embedding is conducted in both spatial and frequency domains. The embedding process flow diagram is given in Fig. 4.

To begin the process, the video is decomposed into consecutive RGB frames. To further satisfy the trade-off between the imperceptibility and the robustness requirements, the RGB frame is converted to the YCbCr color model as its components are less correlated than the RGB ones. To improve the imperceptibility aspect, among the constituent channels, only the luminance component Y is implied in the watermarking by virtue of the fact that it provides a high relevancy to the human visualization capacity. Explicitly, the human visual system is less sensitive toward distortions at luminance level compared to modifications in chrominance components (Cb and Cr). In addition, the involvement of the Y band in the watermarking further strengthens the scheme sustainability owing the fact that the luminance element is lenient against different attacks particularly compression operations. Hence, the primary watermark is hidden in Y component pixels using the LSB-based insertion method where the least significant bit of pixel is substituted by the current watermark bit.

Thereafter, the second watermark is embedded inside of the initially watermarked Y component. Indeed, this latter is split into 16*16 blocks as illustrated in Fig 4. Then, every bit in the second watermarked is expanded into three copies which are inserted in the same block. Indeed, each block is divided in four equal-sized sub blocks, which are classified into textured and non-textured sub-blocks based on texture analysis as already discussed in Sect. 4.1. The three most textured sub-blocks are selected as the best embedding locations for the three expanded similar bits. The explanation behind watermarking the strongly textured sub-blocks is that the human visual system (HVS) is less sensitive to changes in textured areas that hold many details rather than in non-textured ones.

Taking into account that increasing the decomposition level can induce serious perceivable distortions to the final watermarked video quality as well as and significantly reduce the embedding payload capacity, a single level LWT is applied to transform the selected candidate sub-blocks to the frequency domain. This operation generates four different frequency sub-bands termed as low-frequency sub-band (LL), high-frequency sub-band (HH) and mid-frequency sub-bands (LH) and (HL). Inserting a mark in LL sub-band, which contains the greatest frame energy amount, guarantees a good robustness, whereas it yields an apparent distortion in the video watermarked quality. In contrast, involving the high-frequency sub-band provides high perceptual quality while making the watermark fragile to attacks. Among the two mid-frequency bands, HL is more susceptible to human visual system (HVS). Thereby, LH is recognized as the most appropriate sub-band for the embedding since it successfully ensures a favorable balance between the mutually inversely related transparency and sustainability aspects.

For enhancing the proposed scheme performance, the LWT is associated with SVD. Thus, the selected sub-band (LH) is factorized to the product of three matrices namely U, S and V using SVD. As already highlighted in Sect. 3.2, the U matrix first column coefficients exhibit a stable relationship by having the same sign and being strongly correlated values. Besides, it holds the main and greatest coefficients values. In the proposed scheme, these attractive properties are exploited to establish an embedding algorithm that permits to increase the robustness and provide further coherence with the human visual system as well as guarantee a blind detection at the receiving side. In order to find the most correlated two elements among the U matrix first column coefficients, the average of difference between each couple of coefficients values, denoted by D_{avg}, is calculated for different videos and summarized in Table 2. As can be noticed from this table, U(2,1) and U(3,1) are the closet coefficients since they exhibit the less difference of values. Hence, the watermark bits are encapsulated within the U matrix by adjusting these two values via the formulas below:

If $W_{\text{embedding}} = 0$

$$\begin{cases} U_{\text{watermarked}}(2,1) = sign(U(2,1)) \times |U_{\text{avg}} + T| \\ U_{\text{watermarked}}(3,1) = sign(U(3,1)) \times |U_{\text{avg}} - T| \end{cases} \tag{6}$$

Else

$$\begin{cases} U_{\text{watermarked}}(2,1) = sign(U(2,1)) \times |U_{\text{avg}} - T| \\ U_{\text{watermarked}}(3,1) = sign(U(3,1)) \times |U_{\text{avg}} + T| \end{cases} \tag{7}$$

Where U(2,1) and U(3,1) are the original values of the second and the third coefficients in U matrix. $U(2,1)_{\text{watermarked}}$ and $U(3,1)_{\text{watermarked}}$ are the watermarked values corresponding to the second and the third coefficients in U matrix. U_{avg} is the mean between U(2,1) and U(3,1) values computed using (8), and T is the embedding strength factor.

$$U_{\text{avg}} = \frac{|U(2,1) + U(3,1)|}{2} \tag{8}$$

The next step is to apply the SVD inverse and the LWT inverse to obtain the watermarked sub-block. The three watermarked sub-blocks are merged with the residual non-watermarked one to get the corresponding watermarked 16*16 block. Similarly, watermarked blocks are combined to create the finally watermarked Y component. To reconstruct the RGB watermarked frame, this latter is merged with the non-watermarked chrominance channels, i.e., Cr and Cb and then, converted to RGB space color. Lastly, the above-explained steps are repeated for all the host frames to obtain the watermarked video.

4.3 Detection Process

Figure 5 presents the flow diagram of the detection process. It includes three stages: the reconstruction, the extraction and the authentication and tampering localization. The reconstruction phase allows regenerating the watermark, which is used to detect spatial tamper, by applying the second watermark construction steps as described in the Sect. 4.1. $W_{\text{reconstructed}}$ is the label utilized to denote the obtained watermark. Execution concrete steps of the other two stages, i.e., the extraction and the authentication and tampering localization are introduced in the subsequent subsections.

4.3.1 Watermarks Extraction Process

During this process, two watermarks are blindly extracted. This allows avoiding information storage and bandwidth consumption issues as neither the unwatermarked video version nor the secret signature need to be saved and transmitted along with the watermarked sequence.

Table 2 The average difference between each couple of coefficients in first column of U component for different videos

	Test	Camera1	Video1	Coastguard	News	Foreman	Paris
$D_{\mathrm{avg}}(U_{11}, U_{21})$	0.02735	0.03160	0.04454	0.06177	0.06196	0.03146	0.05139
$D_{\mathrm{avg}}(U_{11}, U_{31})$	0.04102	0.04535	0.05831	0.07667	0.08636	0.05203	0.07140
$D_{\mathrm{avg}}(U_{11}, U_{41})$	0.05735	0.066871	0.06785	0.09143	0.10950	0.06936	0.09168
$D_{\mathrm{avg}}(U_{21}, U_{31})$	0.02733	0.03159	0.04434	0.06160	0.06088	0.03134	0.05135
$D_{\mathrm{avg}}(U_{21}, U_{41})$	0.05738	0.06689	0.06813	0.09166	0.10976	0.06955	0.09175
$D_{\mathrm{avg}}(U_{31}, U_{41})$	0.03500	0.04292	0.04966	0.06589	0.06179	0.03360	0.05251

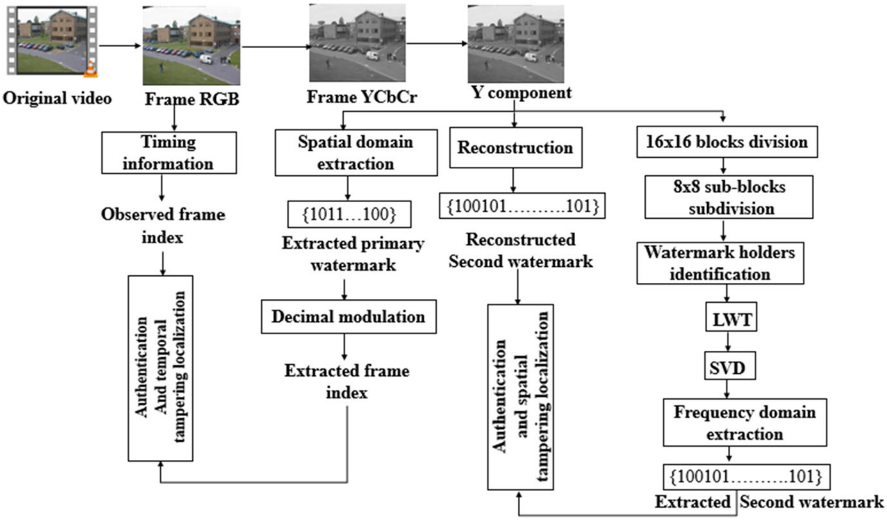

Original video
Frame RGB
Frame YCbCr
Y component
Timing information
Observed frame index
Authentication And temporal tampering localization
Spatial domain extraction
{1011...100}
Extracted primary watermark
Decimal modulation
Extracted frame index
Reconstruction
{100101..........101}
Reconstructed Second watermark
Authentication and spatial tampering localization
16x16 blocks division
8x8 sub-blocks subdivision
Watermark holders identification
LWT
SVD
Frequency domain extraction
{100101..........101}
Extracted Second watermark

Fig. 5 Overall flow diagram of the proposed detection process

The extraction process begins with the same watermarks insertion steps as shown in Fig. 5. After partitioning the watermarked video into RGB frames and converting them to the YCbCr color model, Y channel is retained for further processing. To get the primary extracted watermark, the least significant bits are extracted from the Y pixels values and aggregated in one binary sequence, which is later converted to decimal number representing the extracted frame index.

In order to extract the second watermark, the same luminance component Y is decomposed into 16*16 blocks. From each block, which is subdivided into 8*8 sub-blocks, three copies (C_1,C_2,C_3) for the same watermark bit $W(i)_{\text{extracted}}$ are extracted. In fact, the watermark holders among the resulting sub-blocks are identified with the same procedure used during the watermark embedding process. Recognized watermark carriers are subjected to LWT following by SVD. Each copy of the watermark bit C_j is extracted from the considered $U_{\text{extracted}}$ matrix through the following rules:

$$\begin{cases} C_j = 0 \quad If \quad U_{\text{extracted}}(2,1) > U_{\text{extracted}}(3,1) \\ C_j = 1 \quad \text{Otherwise} \end{cases} \tag{9}$$

Where $U_{\text{extracted}}(2, 1)$ and $U_{\text{extracted}}(3, 1)$ denote the extracted values corresponding, respectively, to the second and the third coefficients in the $U_{\text{extracted}}$ matrix first column. C_j is the watermark bit copy relative to the sub-block number j, with j ∈ {1, 2, 3}.

The extracted watermark bit final value $W(i)_{\text{extracted}}$ assigned to a considered block is evaluated by applying a majority voting method on the three versions (C_1,C_2,C_3) issued from its corresponding sub-blocks as explained in Table 3. All the extracted watermark bits are concatenated together to obtain the final extracted second watermark $W_{\text{extracted}}$.

Table 3 Majority voting rules for watermark bit value investigation

C_1	C_2	C_3	$W(i)_{\text{extracted}}$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

4.3.2 Authentication and Tampering Localization

Establish an efficient strategy to verify the video content authenticity and to detect and locate the tampered area is the proposed approach main focus. In temporal authentication stage, the observed frame index and the extracted index from the current frame are compared. A mismatch between these two indexes permits to detect and locate a performed temporal tampering.

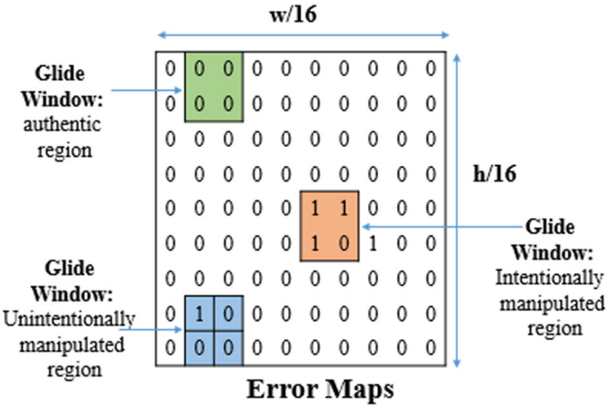
On the other hand, the spatial authentication is conducted for each watermarked video frame by inspecting the similarity between the extracted second watermark $W_{\text{extracted}}$ and its reconstructed version $W_{\text{reconstructed}}$ using two measures namely Bit Error Rate (BER) and Normalized Correlation (NC) as well as two thresholds termed as Tr_{BER} and Tr_{NC}. Definitions and mathematical descriptions of these metrics are provided in Sect. 5.3. The authentication decision is made after comparing the BER and the NC measures to their corresponding thresholds. The considered frame is judged as authentic only if BER is below its threshold Tr_{BER} and NC is above Tr_{NC}. Otherwise, the given frame of $h{*}w$ size is regarded as non-authentic.

Altered content regions are located in non-authentic watermarked video frame based on a tampering localization strategy including the following steps. First, an error map E_{map}, which is a $\frac{h}{16} * \frac{w}{16}$ size matrix, is elaborated by XOR-ing the extracted second watermark with the reconstructed one as shown in (10).

$$E_{\text{map}} = W_{\text{extracted}} \oplus W_{\text{reconstructed}} \tag{10}$$

To classify the 16 × 16 blocks, issued from the Y component segmentation, as authentic, intentionally manipulated or non-intentionally manipulated, the constructed error map E_{map} coefficients are scanned using a 2x2 sized glide window. Obviously, the authentication sensitivity depends on the glide window dimensions. In fact, the 2x2 size enables to increasing the localization ability by accurately finding the tampered blocks. Thus, the glide window is opened from each location in the E_{map} and then, the values of the four held in elements are examined. Based on these latter, maliciously tampered regions are identified.

As E_{map} is constructed by XOR-ing the extracted and the reconstructed second watermark versions (application of the Exclusive OR between the two watermarks),


w/16
h/16
Glide Window: authentic region
Glide Window: Intentionally manipulated region
Glide Window: Unintentionally manipulated region
Error Maps

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	1	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 6 Sample of tampering localization with Error map and 2*2 glide window size

a nonzero element in E_{map} coefficients corresponds to a mismatch between the two investigated watermarks. Therefore, if at least two nonzero values are clustered among the glide window elements, the corresponding 16x16 block localized in the considered frame is flagged as an intentionally manipulated block content. If there are only zero elements inside the glide window elements, the corresponding block is considered as authentic. Otherwise, the associated block is deemed as unintentionally manipulated.

Figure 6 depicts an illustration of tampering localization using the error map E_{map} with the glide window.

5 Simulations Findings and Comparisons

To testify the proposed system performance and prove the effectiveness of the used techniques and procedures to build up its overall framework, several simulations are carried out under diverse videos including test, camera1, video1, coastguard, news, foreman and paris. The test videos can be classified into two sets. The first one includes well-known standard videos, which are commonly applied in the watermarking field and often utilized to evaluate state-of-the-art schemes. The second set involves different surveillance sequences.

Actually, the developed watermarking scheme is evaluated in terms of the three requirements. In Sects. 5.1 and 5.2, the capacity and the imperceptibility are scrutinized to demonstrate the proposed scheme ability to hide a large amount of watermark data without raising a rigorous visual quality degradation and thereby highlight the proficiency of (i) the involved texture features-based watermarking positions selection strategy and (ii) the proposed watermarks insertion algorithm.

As previously stated, the proposed semi fragile video watermarking system is devoted for authenticity verification as well as tampering detection and localization goals. Hence, its semi-fragility performance is evaluated in Sect. 5.3 to substantiate its reliability in terms of discrimination between unintentional attacks and intentional ones along with maliciously tampered content localization. More precisely, Sect. 5.3 is divided into two subsections namely Sects. 5.3.1 and 5.3.2. Section 5.3.1 is dedi-

Table 4 Capacity values for used videos for simulation experiments

Videos	Capacity per frame (bits)
Test	1188
Camera1	1188
Video1	900
Coastguard	1188
News	297
Foreman	297
Paris	1188

cated to analyze the robustness against common processing, while Sect. 5.3.2 deals with the fragility to temporal and spatial intentional attacks investigation in addition to tampering localization.

The conducted simulations in the above-mentioned two subsections permit to attest the appropriateness of (i) the advanced content-based watermarks generation procedure to construct two distinct authentication information that allow to successfully fulfill the tasks of attacks types characterization, inter-frames as well as intra-frame forgeries detection and maliciously tampered areas localization (ii) the various beneficial techniques and procedures used to build up the hybrid domain-based watermarks embedding process particularly the LSB, the LWT and the SVD techniques and (iii) the proposed glide window-based spatial distortions detection strategy for an accurate tamper localization.

In the sequel, simulations results belonging to each performance level, i.e., capacity, imperceptibility and semi-fragility are reported in detail, thoroughly discussed and compared with previous works. It is worth mentioning that the embedding strength factor is empirically set to T = 0.05 in the performed experiments.

5.1 Capacity Results

As explained before, every watermark bit is inserted in three copies in one 16*16 size block during the embedding process. Hence, the information amount that can be concealed into a given frame with h*w dimension, denoted by Cap, can be expressed by (11).

$$Cap = 3 \times \frac{h \times w}{16 \times 16} \tag{11}$$

The different test videos capacity results are illustrated in Table 4. From this table, it can be inferred that the proposed scheme offers a large payload capacity. The main reason behind the achieved results is the use of block-based embedding method with the bit expansion mode.

5.2 Imperceptibility Results

In order to evaluate the proposed watermarking scheme imperceptibility, the Peak Signal to Noise Ratio (PSNR) is used as measure. This metric allows assessing the

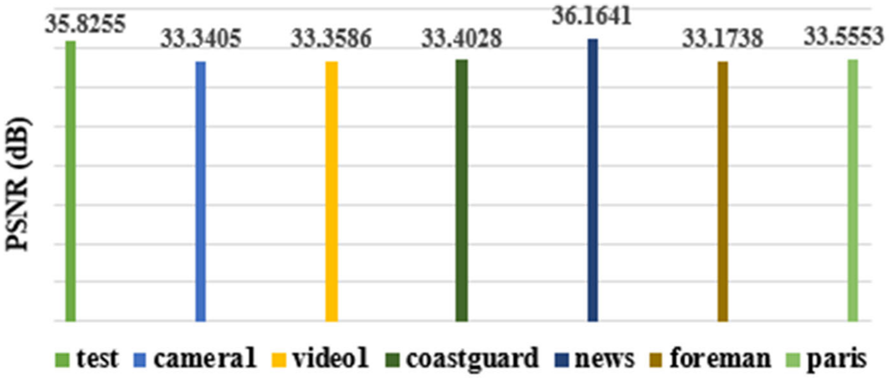


Fig. 7 Imperceptibility results: The obtained PSNR values for all watermarked videos

quality difference between the original host video and its watermarked version with respect to human visual system [21]. The PSNR value, which is proportional to the Mean Square Error (MSE) value, is computed using (12).

$$PSNR = 10 \times \log \frac{255^2}{MSE} \tag{12}$$

PSNRs of all watermarked videos are shown in Fig. 7. From this figure, it is observed that the obtained PSNR values are ranged between 33.1738 dB and 36.1641 dB. As the marked frames are estimated to have high perceptual quality when the PSNR is higher than 30 dB [26, 50], the obtained results correspondingly disclose the good imperceptibility level, achieved using the proposed dual watermarking scheme.

For subjective observation and better visual comparison, examples of unwatermarked original frames and their corresponding watermarked frames are plotted in Fig. 8. According to Fig. 8, no perceptible distortion is visually detectable in the dual watermarked frames. Thus, our watermarking approach preserves the video perpetual quality after the embedment of a watermark data with a significant size. Involving an efficient texture features-based method to choose the best watermark holders positions in coherence with the human visual system, selecting the mid-frequency sub-band for the watermarking and combining the LWT and the SVD during the embedding process enable to successfully fulfill the imperceptibility requirement.

5.3 Semi-fragility Results

To assess the semi-fragility performance, two standard measurement parameters are considered namely the Normalized Correlation (NC) and the Bit Error Rate (BER) which are computed through (13) and (14), respectively [23]. NC is used to inspect the difference between the extracted watermark $W_{\text{extracted}}$ and the reconstructed one $W_{\text{reconstructed}}$, while BER is defined as the watermark erroneous bits rate overall the correct ones.

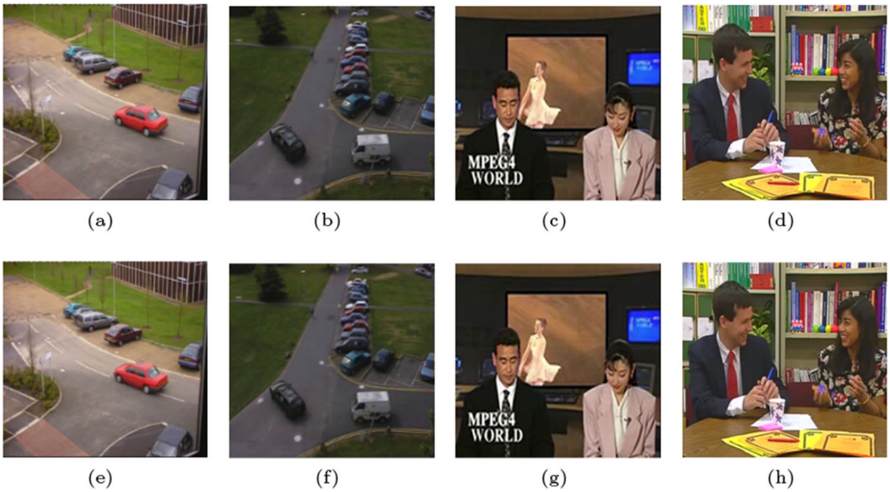


Fig. 8 Up: non-watermarked host frames: **a** test.avi **b** camera1.avi **c** news.avi **d** paris.avi, Down: watermarked frames: **e** test.avi **f** camera1.avi **g** news.avi **h** paris.avi

$$NC = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n} W(i,j)_{\text{extracted}} \times W(i,j)_{\text{reconstructed}}}{\sqrt{\sum_{i=0}^{m}\sum_{j=0}^{n} W(i,j)^2_{\text{extracted}}}\sqrt{\sum_{i=0}^{m}\sum_{j=0}^{n} W(i,j)^2_{\text{reconstructed}}}} \tag{13}$$

$$BER = \sum_{i=0}^{m}\sum_{j=0}^{n}(\frac{W(i,j)_{\text{extracted}} \oplus W(i,j)_{\text{reconstructed}}}{m \times n} \tag{14}$$

Where, the watermarks dimensions are represented by m and n and $\oplus$ is the XOR operator.

Attacks characterization is a paramount property that must be satisfied by an authentication-based watermarking scheme. Indeed, a semi-fragile scheme for authentication goal should survive admissible processing which preserve semantic content. On the other hand, it should detect malicious attacks, which attempt to alter the data general semantic, and be able to locate the edited regions in the tampered watermarked video frames. To these ends, two thresholds Tr_{NC} and Tr_{BER} are, respectively, associated with NC and BER. As prescribed in Sect. 4.3.2, the content is deemed as authentic only if BER is inferior to Tr_{BER} and NC is superior to Tr_{NC}. Otherwise, the content is regarded as non-authentic. Tr_{NC} and Tr_{BER} are empirically set to 0.9 and 0.1, respectively. Several experiments are directed to test the robustness of the proposed watermarking scheme against non-malicious modifications and its fragility to malicious ones. The following subsections detail and discuss the obtained results.

5.3.1 Robustness to Non-malicious Attacks Results

The proposed scheme robustness is investigated against common processing operations including Gaussian noise, salt and pepper noise, speckle noise, brightening, sharping, histogram equalization, H.264 compression and MJPEG compression. NC and BER values obtained after performing the already listed non-malicious attacks as well as in the case of absence of attacks are recorded in Table 5 and Table 6. If no attack is carried out, the complete hidden information can be successfully extracted since the obtained NC and BER are, respectively, 1 and 0 as shown in Table 5 and Table 6. This property is fundamental for an authentication scheme where a faithful authenticity decision and an accurate tamper localization are required.

After adding Gaussian noise, the procured NC values vary between 0.92973 and 0.98894; thereby, they are above the fixed threshold $Tr_{\mathrm{NC}} = 0.9$. Besides, the obtained BER values are ranged between 0.02197 and 0.08135; hence, they do not exceed their relative threshold $Tr_{\mathrm{BER}} = 0.1$. Based on these results, it can be concluded that the proposed scheme can resist this manipulation type. Similarly, the detector is able to efficiently extract the encapsulated watermark from all watermarked videos processed by the two next considered non-malicious attacks viz. salt and pepper and speckle noises. In fact, according to the experiments results tabulated in Table 5 and Table 6, the achieved NC values under salt and pepper attack reach 0.99751 and the obtained BER is ranged between 0.00494 and 0.06144. Likewise, the resulting NC and BER after performing speckle noise attack are, respectively, superior to 0.95018 and inferior to 0.09661. The resilience to Gaussian, salt and pepper and speckle noises is issued from involving the LWT, which is broadly recognized by its high immunity to noises addition, in the watermarking algorithm.

Afterward, three different enhancement-based attacks namely brightening, sharpening and histogram equalization are applied to the watermarked videos in order to test the proposed scheme robustness to them. As shown in Table 5 and Table 6, the minimum NC and the maximum BER associated with brightening attack are 0.98731 and 0.02514, respectively. As far as the sharpening manipulation is concerned, the achieved NC vary between 0.99489 and 0.99999, while the BER values lie between 0 and 0.01017. For the last considered enhancement attack, i.e., the histogram equalization manipulation, the obtained NC values are beyond 0.98704 and the BER values do not exceed 0.02569. Hence, all the acquired NC measurements are above Tr_{NC}, while BER is below Tr_{BER}, which confirm that the embedded signature detection can be successfully fulfilled after applying the enhancement-based attacks. The selection of the most convenient blocks for the watermark insertion along with the simultaneous use of the intrinsic characteristics relative to two powerful transforms, i.e., LWT and SVD to build up the proposed watermarking scheme processes are the fundamental arguments for the explanation of this strong resilience to enhancement-based attacks.

As seen from Table 5 and Table 6, the last considered non-malicious attacks are the H.264 and MJPEG compression. The averages of NC and BER values obtained under H.264 compression are equal to 0.99169 and 0.01592, respectively. However, 0.00314 and 0.99841 are, respectively, the achieved BER and NC averages after compressing the watermarked video based on MJPEG format. The above outcomes demonstrate the proposed scheme immunity to these attacks. The main reason for this result is the

Table 5 Robustness results: The NC values after the non-malicious attacks application

	Test	Camera1	Video1	Coastguard	News	Foreman	Paris
No attacks	1	1	1	1	1	1	1
Gaussian Noise(0.01)	0.97935	0.96165	0.98894	0.97264	0.92973	0.97093	0.96082
Salt and Pepper(0.01)	0.99577	0.99618	0.99619	0.99596	0.99476	0.99751	0.99623
Salt and Pepper(0.02)	0.98589	0.98710	0.98634	0.98723	0.98285	0.99023	0.98712
Salt and Pepper(0.03)	0.97337	0.97555	0.97513	0.97553	0.96870	0.97928	0.97543
Speckle Noise(0.01)	0.99997	0.97539	0.99940	0.99928	0.99872	0.98912	0.99713
Speckle Noise(0.02)	0.99878	0.96513	0.99647	0.99711	0.99453	0.97172	0.99073
Speckle Noise(0.03)	0.99464	0.95722	0.99049	0.99248	0.98872	0.95018	0.98117
Brightening	0.99974	1	0.98923	0.99606	0.99326	0.98731	0.99672
Sharpening	0.99999	0.99982	0.99987	0.99793	0.99489	0.99903	0.99797
Histogram equalization	0.99800	0.99963	0.98940	0.99473	0.99092	0.98704	0.99641
H.264 compression	0.99996	0.99998	1	0.97041	0.99813	0.97414	0.99926
MJPEG compression	0.99979	1	1	0.99888	0.99026	1	0.99995

Table 6 Robustness results: The BER values after the non-malicious attacks application

	Test	Camera1	Video1	Coastguard	News	Foreman	Paris
No attacks	0	0	0	0	0	0	0
Gaussian Noise(0.01)	0.04083	0.07517	0.02197	0.05389	0.08135	0.05717	0.07676
Salt and Pepper(0.01)	0.00842	0.00762	0.00759	0.00805	0.01041	0.00494	0.00750
Salt and Pepper(0.02)	0.02798	0.02560	0.02710	0.02534	0.03390	0.01937	0.02557
Salt and Pepper(0.03)	0.05251	0.04826	0.04907	0.04829	0.06144	0.04088	0.04850
Speckle Noise(0.01)	0.00004	0.04858	0.00118	0.00143	0.00255	0.02157	0.00572
Speckle Noise(0.02)	0.00243	0.06848	0.00704	0.00575	0.01087	0.05552	0.01843
Speckle Noise(0.03)	0.01066	0.08367	0.01890	0.01495	0.02236	0.09661	0.03727
Brightening	0.00051	0	0.02140	0.00784	0.01342	0.02514	0.00653
Sharpening	0	0.00034	0.00025	0.00412	0.01017	0.00192	0.00404
Histogram equalization	0.00397	0.00073	0.02107	0.01049	0.01803	0.02569	0.00715
H.264 compression	0.00007	0.00002	0	0.05744	0.00360	0.04810	0.00145
MJPEG compression	0.00040	0	0	0.00222	0.01932	0	0.00009

selection of the middle frequency sub-band as appropriate watermark carrier, which improves the overall proposed scheme robustness attribute and particularly against compression processing.

All the above reported evaluations outcomes ascertain that our proposed scheme has an outstanding capability to survive a large set of incidental distortions including noises addition attacks, enhancement-based operations and compression processing. These important findings are obtained owing to the exploitation of the attractive and complementary characteristics of both LWT and SVD transforms throughout the different watermarking stages. Moreover, apart from improving the transparency aspect, the texture feature-based method involved during the insertion process permits to select the proper watermarking positions with an important stability level, which further improves the advanced algorithm performance in terms of robustness.

5.3.2 Fragility to Malicious Attacks Results and Tampering Localization

The watermarked videos are mischievously tampered in order to scrutinize the proposed approach effectiveness against malicious attacks and its ability to detect tampering locations. Malicious attacks are often classified into temporal and spatial attacks. As explained in Sect. 4.3.2, a mismatch between the observed frame index and the extracted one from the current watermarked frame permits to detect and locate a performed temporal distortion. However, spatial tampering is detected and tampered segments are located by assessing the difference between the extracted second watermark and its reconstructed version using an error map and a glide window.

Firstly, watermarked videos are applied to different temporal attacks, including (i) frame dropping; (ii) frame replacing and (iii) frame swapping.

Without conducting any temporal manipulation, the observed frame index as well as the extracted one are shown in Fig. 9a. As expected, the graph shows that the two indexes are identical which indicates that no temporal tampering is occurred.

Frame dropping attack is simulated by deleting the frames 40-60. The corresponding authentication result is given in Fig. 9b. It can be noticed that the watermarked video contains 80 frames, but the last extracted frame index is 100. In addition, it can be seen that the observed and extracted indexes keep equal until frame 39, then a jump between frames 40 and 60 is occurred. Based on these findings, it can be claimed that the frames 40–60 are dropped.

Figure. 9c displays the authentication result after applying frame replacing attack by changing the frames ranged from 40 to 60 with the twenty first frames. According to Fig. 9c, the observed index and the extracted one are consistent for frames 1–39 and 61–100, but a total mismatch is inferred from frame 40 to frame 60. These drawn observations demonstrate that frames 40–60 are tampered in temporal domain.

Figure 9d illustrates the result of frame swapping manipulation where the frames 25-49 are sequentially replaced by the frames 50 to 75 and vice versa. The observed and the extracted indexes are unaffected until frame number 24 and after frame number 75 as depicted by the plot in Fig. 9d. However, the indexes of the two ranges namely 25–49 and 50–75 are exchanged. Therefore, this temporal manipulation is successfully detected.

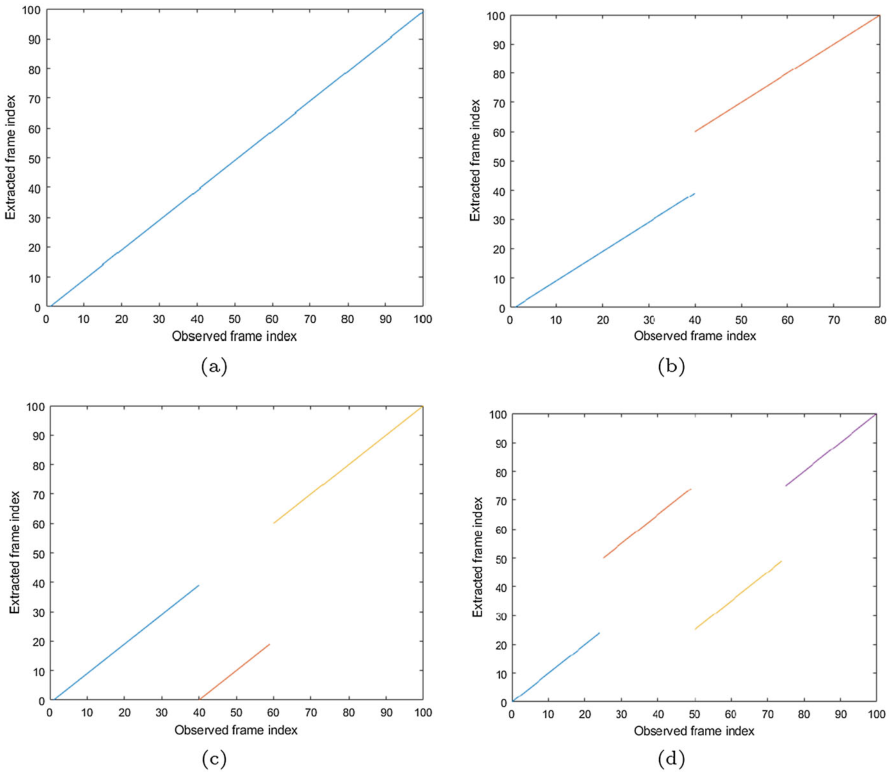
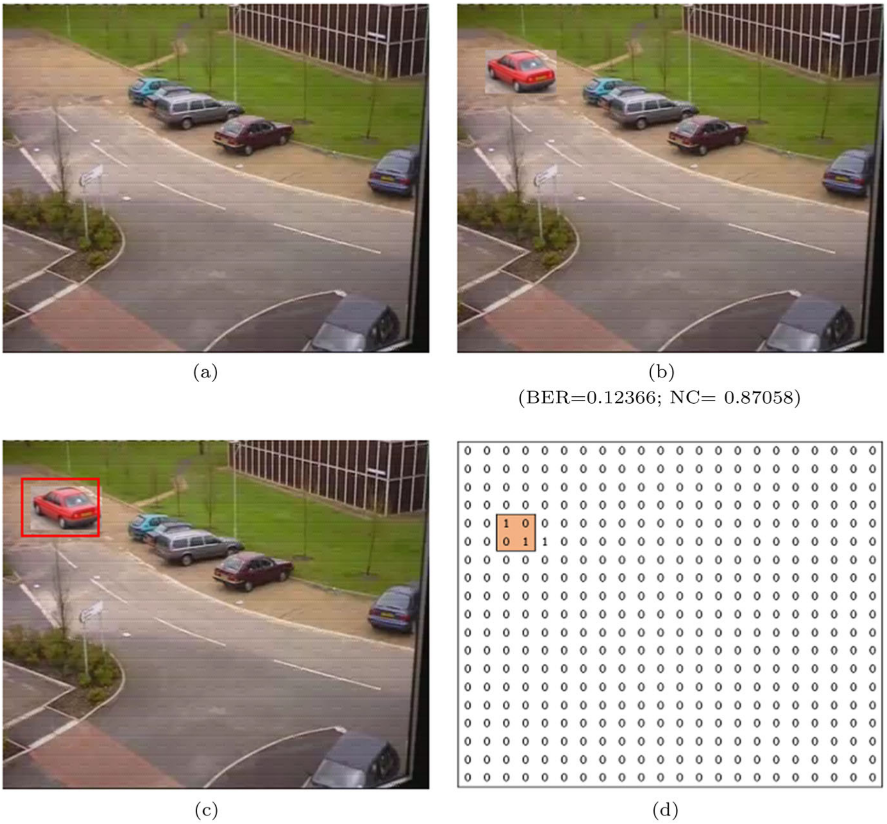


Fig. 9 Temporal tampering detection and localization results: **a** without tampering **b** frame dropping attack **c** frame replacing attack **d** frame swapping attack

In summary, the above-discussed experimental findings confirm the proposed semi-fragile watermarking approach reliability in detecting and precisely locating all the simulated inter-frames tampering. This high performance is resulted from the suitable choice of the current frame index as timing information for the construction of the first authentication information that is used to mark the considered cover video frame.

On the other hand, watermarked videos are subjected to another set of attacks in order to test the proposed scheme performance in detecting and locating spatial distortions namely (i) object removal and (ii) object insertion attacks. Thus, the visual content of different watermarked videos frames is deliberately manipulated by deleting or adding a specific object. In the presence of these two tampering manipulation types, the obtained authentication metrics, i.e., NC and BER are, respectively, above and below their preset thresholds Tr_{NC} and Tr_{BER} as indicated in Fig. 10 and Fig. 11. These results represent an explicit proof about the intentional modifications that the frames had undergone. The glide window-based tampering localization strategy described in Sect. 4.3.2 is then performed to determine the altered regions yielded by object insertion or object removal attacks as illustrated in Fig. 10 and Fig. 11, respectively.

For better illustration of the tampering localization ability, several watermarked frames, their maliciously tampered versions by one of the two considered spatial forg-



(a)

(b)
(BER=0.12366; NC= 0.87058)

(c)

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

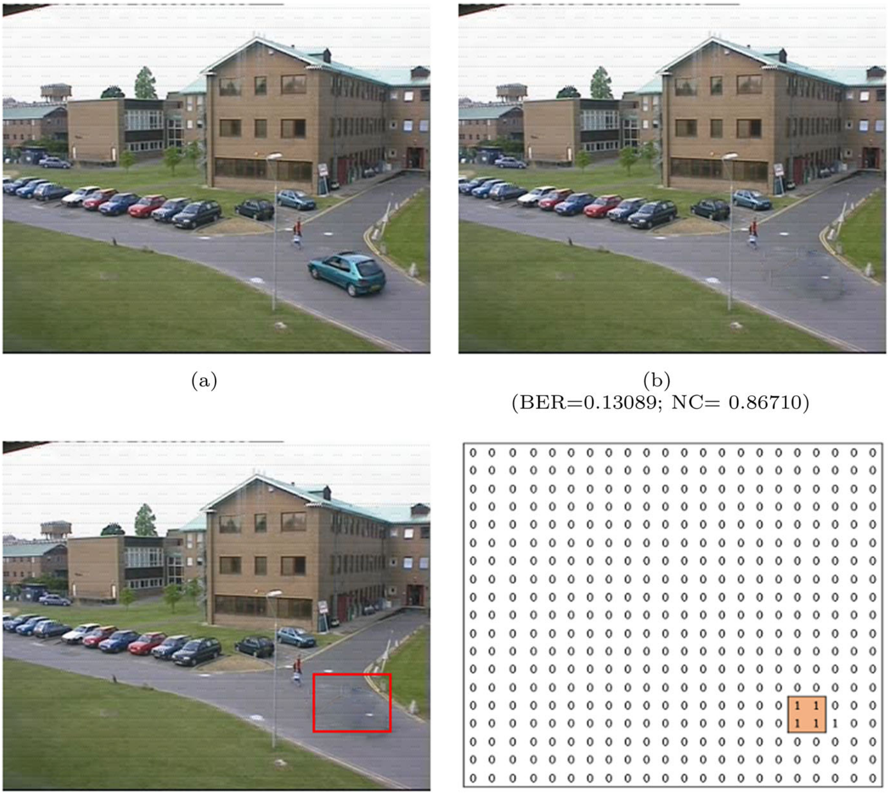
(d)

Fig. 10 Spatial tampering detection and localization results after object insertion attack: **a** watermarked frame **b** tampered frame with BER=0.12366 and NC= 0.87058 **c** corresponding error map **d** localized tampering

eries and their corresponding versions with the localized tampering are depicted in Fig. 12 and Fig. 13. As obvious from these illustrations, the proposed scheme successes to detect and concisely locate these kinds of tampering manipulations in the spatial domain. This achievement is due to the involvement of an effective watermark generation strategy, which benefits from 3 independent texture features to elaborate to be embedded authentication data that faithfully describe the current frame content characteristics. Indeed, being subjected to a malicious attack contributes to an inconsistence between the watermarked and the intentionally tampered frame versions at texture information level. Taking advantage of this property as well as the glide window concept allows ensuring reliable spatial tamper detection and localization.

5.4 Comparative Study

To further validate its performance, the proposed technique is compared with seven previous existing watermarking techniques, which are discussed in [2, 6, 31, 37, 38, 40,



(a)

(b)
(BER=0.13089; NC= 0.86710)

(c)

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(d)

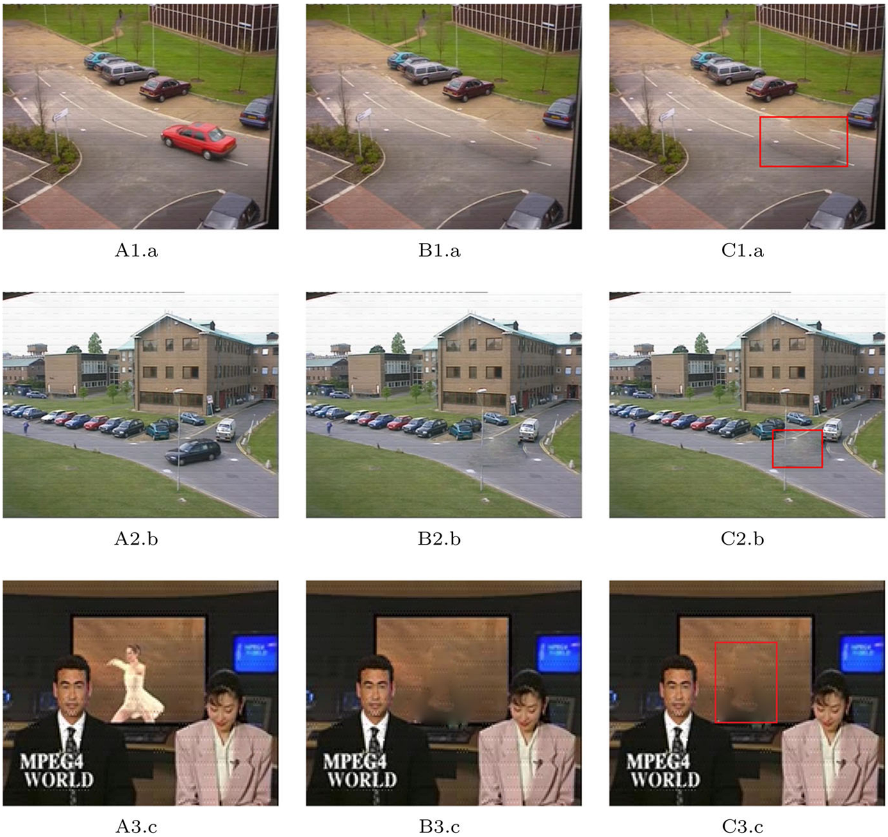
Fig. 11 Spatial tampering detection and localization results after object removal attack: **a** watermarked frame **b** tampered frame with BER=0.13089 and NC= 0.86710 **c** corresponding error map **d** localized tampering

41], with respect to robustness, capacity and imperceptibility properties. Moreover, a comparison in terms of tampering localization ability is conducted with state-of-the-art systems cited in [10, 15, 45, 46]. For a faithful comparison seeking, the experiments comparing the robustness, which are depicted in Tables 7 and 8, are redone using watermarks of length 90 bits.

The comparison in terms of robustness between our technique and that in [6] is provided in Table 7. Referring to Table 7, it is quite evident that the proposed technique carries superior robustness against Gaussian noise, salt and pepper and speckle noise as well as sharpening, histogram equalization and MJPEG compression. In addition, the two watermarking methods provide a comparable robustness performance against brightening attack. In brief, it is noted that the most findings of the proposed technique surpass those reached by the considered comparative algorithm. This superiority is resulted from the use of LWT as a combination with the SVD to carry out the watermarking in the multi-frequential domain instead of the mono-frequential one, which contributes to a considerable improvement in terms of the resilience performance.

Table 7 Robustness comparison (NC) between the proposed approach and the watermarking scheme [6]

	Foreman		News		Average	
Attacks	[6]	Proposed scheme	[6]	Proposed scheme	[6]	Proposed scheme
Gaussian noise (0.01)	0.8105	0.9689	0.8418	0.9285	0.8261	0.9487
Salt and pepper (0.01)	0.9395	0.9972	0.9395	0.9953	0.9395	0.9963
Speckle noise (0.01)	0.9883	0.9881	0.9766	0.9982	0.9824	0.9932
Sharpening	1	0.9989	0.9922	0.9900	0.9961	0.9946
Histogram equalization	0.8867	0.9860	0.8027	0.9899	0.8447	0.9890
Brightening	–	–	1	0.9927	1	0.9927
MJPEG compression	0.9980	1	0.9805	0.9893	0.9892	0.9947


A1.a B1.a C1.a

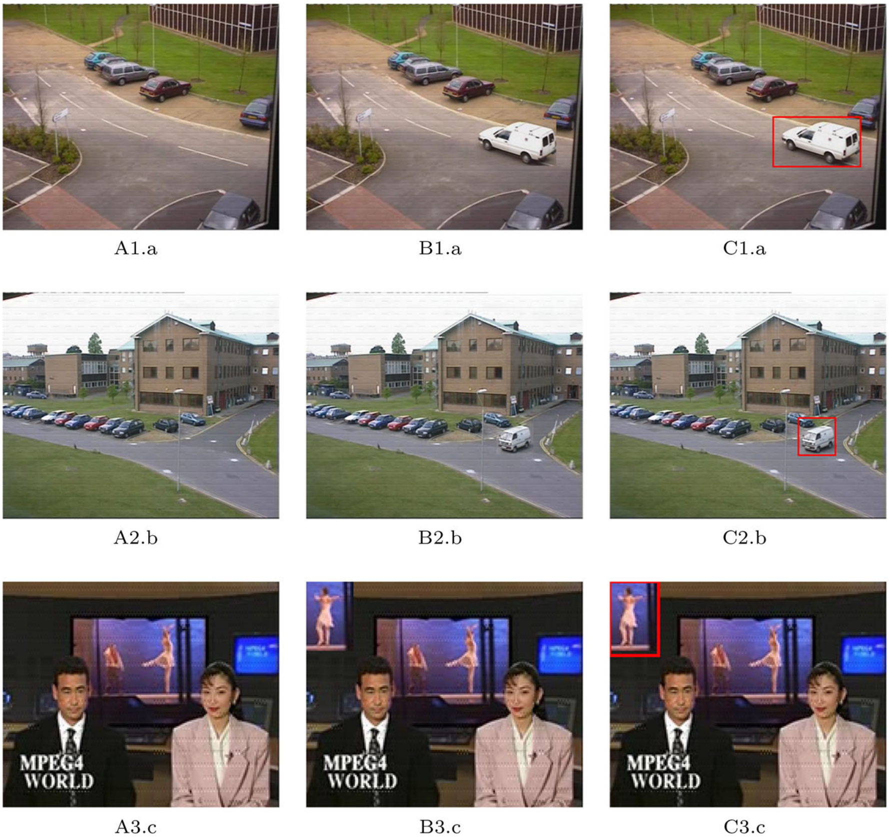
A2.b B2.b C2.b

A3.c B3.c C3.c

Fig. 12 Samples of **A** Watermarked frames, **B** their tampered versions by object removal attack and their corresponding located tampered area from videos. (**a**) test, (**b**) camera1, (**c**) news

On the other hand, Table 8 depicts the comparison of the proposed scheme robustness with those presented in [2, 31, 37, 38, 40, 41]. Indeed, the NC values tabulated in Table 8 demonstrate that our watermarking technique outperforms all the techniques cited in this table in terms of the resistance to noises addition attacks except salt and pepper to which the method introduced in [37] exhibits a slightly better robustness level. The involvement of LWT transform, which is broadly recognized by its noteworthy immunity to noises addition, in the watermarking processes is the prime reason for our scheme good sustainability against this type of manipulations.

Moreover, based on the comparison results summarized in Table 8, it is revealed that our proposed system is the most robust to sharpening and brightening as it attains the high performance value when compared to the existing watermarking approaches. Furthermore, it is observed that a competitive survival level against histogram equalization is ensured by the proposed system as well as the previous work in [2], which has a slightly enhanced NC value. Selecting the embedding positions in coherence with the human visual system capacity along with the watermarking execution in the multi-frequency domain allow reaching this robustness degree in the presence of these


A1.a B1.a C1.a

A2.b B2.b C2.b

A3.c B3.c C3.c

Fig. 13 Samples of (**A**) Watermarked frames, (**B**) their tampered versions by object insertion attack and their corresponding located tampered area from videos. (**a**) test, (**b**) camera1, (**c**) news

three enhancement-based attacks. In addition, while analyzing Table 8, it is inferred that our approach outperforms the techniques [37, 38, 40, 41] with respect to the sustainability criterion to the MJPEG compression. For H.264 compression, both of our algorithm and the one in [37] have the same strong resilience degree to this manipulation. The proposed system robustness to compression processing is the result of the middle frequency sub-band use as watermark information holder.

PSNR and capacity measurements values corresponding to our algorithm and to the previous works in [2, 6, 31, 37, 38, 40, 41] can be found in Table 9. According to this latter, the related schemes in [2, 6, 31, 37, 38, 40, 41] achieve better imperceptibility results than the proposed one because the embedding payload in our advanced system is noticeably larger as compared to those methods. Actually, our watermarking scheme capacity is about 3.5 times greater than the comparative works as shown in Table 9. Moreover, in the proposed system, dual watermarks are embedded in every frame in the host video. However, for the approaches [6, 37, 38, 40, 41] and [31] only a few number of frames or blocks are selected for the watermarking, which denotes a less visual content modification. In the same vein, it has to be pointed out that the reasons

Table 8 Robustness comparison (NC) between the proposed approach and the watermarking schemes [2, 37, 38, 40, 41] and [31] on Foreman video

Attacks	[40]	[38]	[37]	[41]	[2]	[31]	Proposed scheme
Gaussian noise (0.001)	0.9792	0.8377	0.9008	0.4193	0.9800	0.6470	1
Salt and pepper (0.01)	0.9636	0.7379	1	0.9207	0.9800	0.6870	0.9972
Speckle noise (0.001)	0.9890	0.8734	0.9479	0.9541	1	0.7100	1
Sharpening	0.9909	0.7301	–	0.9060	–	–	0.9989
Brightening	–	–	–	–	0.9500	–	0.9860
Histogram equalization	–	–	–	–	0.9970	0.7350	0.9860
MJPEG compression	0.9857	0.9317	0.8485	0.9524	–	–	1
H.264 compression	–	–	1	–	–	–	1

Table 9 Imperceptibility and capacity comparison between the proposed approach and the watermarking schemes [2, 6, 37, 38, 40, 41] and [31] on Foreman video

	[40]	[38]	[6]	[37]	[41]	[2]	[31]	Proposed scheme
PSNR	35.03	39.78	40.2	41.83	67.84	59.00	40.60	33.17
Capacity	–	–	90	90	–	–	–	297

for the high imperceptibility of the scheme in [2] are its involvement of a holders selection strategy that allows to choose a limited number of frames to convey the watermark information as well as its semi-blindness nature. Actually, the method is a semi-blind SVD-DWT watermarking technique that uses the unwatermarked video format during the extraction. Therefore, it adopts a very small embedding factor to fuse the watermark information within the cover video. This implies that the perceptual quality is barely degraded. In meantime, our proposed approach is a fully blind scheme. Thus, our provided PSNR value remains reasonable and satisfactory especially as it is superior to 30 db [26, 50].

To compare the proposed watermarking scheme performance regarding the spatial distortions localization ability with the scheme presented in [15], we deliberately tampered the original watermarked frame from news.avi video by deleting the blue TV from the upper right as shown in Fig. 14b. Figures 14c and 14d illustrate the tampering localization findings belonging to the work [15] and our proposed one, respectively. From these figures, it can be seen that the two approaches fulfill the localization requirement. However, it is clearly noticed that our proposed technique ensures a more concise maliciously edited areas identification. In fact, unlike the watermarking scheme in [15], which involves a 3x3 sized gliding window, our methodology relies on a 2x2 glide window-based tampered regions recognition strategy. Actually, the used 2x2 size enables to more improve the tampering detection reliability as well as further increase the tamper localization accuracy.

As far as the temporal tampering localization ability is concerned, the comparative study, recapitulated in Table 10, confirms that all the techniques cited in this table succeed in localizing the frame dropping-based attack. Moreover, except the work in [46] the other schemes can efficaciously recognize frame replacing-based tampering. Nevertheless, only the proposed approach and the system suggested in [46] are able to successfully identify the swapping-based inter-frames manipulation. Therefore, only our approach can offer the capability to jointly find out and properly locate the considered three temporal tampering attacks.

The comparison with other methods testifies the proposed algorithm effectiveness in terms of capacity, imperceptibility, robustness and tampering localization. It efficiently strikes a favorable compromise between the different requirements. This performance is derived from the joint use of various beneficial techniques and procedures to build up its overall framework.

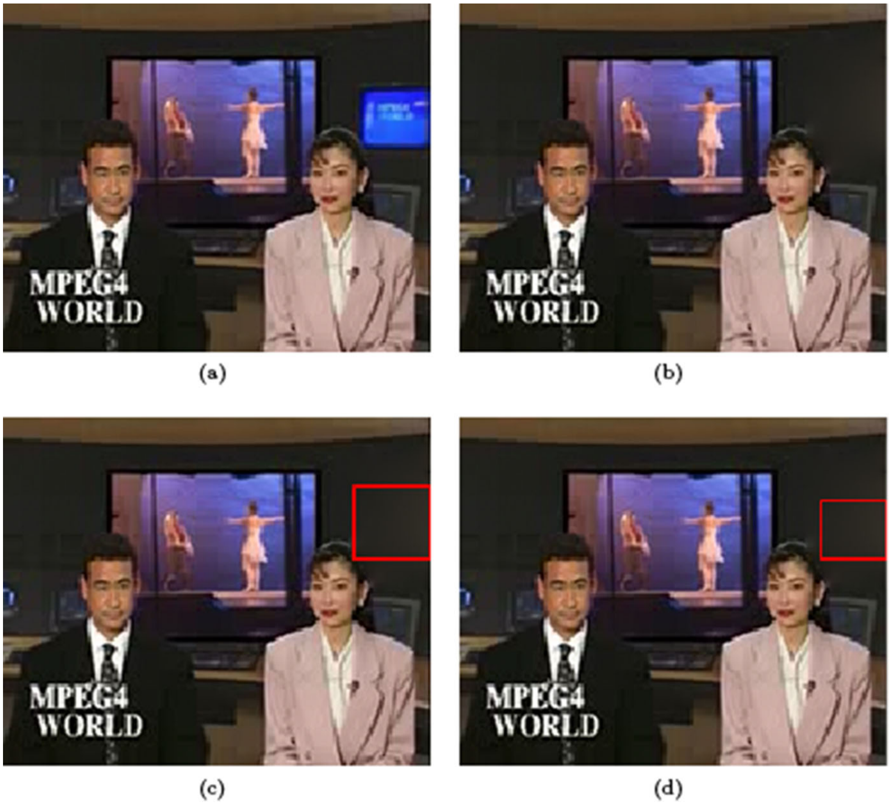


Fig. 14 Spatial tampering localization ability comparison between the proposed approach and the watermarking scheme [15]: (**a**) watermarked frame (**b**) tampered frame (**c**) localized tampering by the scheme [15] (**d**) localized tampering by our scheme

Table 10 Temporal tampering localization ability comparison between the proposed approach and the watermarking schemes [10, 15, 46] and [45]

Temporal tampering	[10]	[46]	[15]	[45]	Proposed approach
Frames dropping	✓	✓	✓	✓	✓
Frames replacing	✓	×	✓	✓	✓
Frames swapping	×	✓	×	×	✓

6 Conclusion and Future Works

In this paper, a semi-fragile dual video watermarking scheme was proposed for content authentication and tampering detection and localization. Timing information and reliable texture features are used to design two authentication watermarks, which are used to detect and locate intra-frame and inter-frames manipulations. The embedding process is done in the hybrid domain. Indeed, the first watermark is embedded within the host frames pixels using LSB method. To meet the security requirement, the second watermark is encrypted by Torus isomorphism mapping. Afterward, it is encapsulated into the most textured sub-blocks in bit expansion mode following a LWT-SVD-based embedding method. In the detection side, the two hidden watermarks are blindly extracted. The mismatch between the observed frame index and the extracted one from the current frame indicates a temporal tampering occurrence. Whereas, the extracted second watermark and its reconstructed version are processed as inputs to a glide window-based tampering localization procedure in order to accurately determine the spatial forgeries. Several experiments were performed to assess the proposed algorithm performance. Simulations results demonstrate the effectiveness of the proposed semi-fragile scheme method in attacks types characterization. In fact, it efficiently survives incidental processing such as compression and noises addition while being sensitive to intentional distortions such as content modification and frames manipulations with a concise location ability. Moreover, it carries a good watermarked video perceptual quality with a large payload capacity. In the follow-up work, the proposed approach will be extended to ensure the tampered regions recovery.

Acknowledgements The research leading to these results received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the Grant Agreement Number LR11ES48.

Availability of Data and Materials The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. N.A. Abbas, Image encryption based on independent component analysis and Arnold's Cat Map. Egypt Inform. J. **17**(1), 139–146 (2015). https://doi.org/10.1016/j.eij.2015.10.001
2. N.T.B. Abdulla, K.A. Navas, Robust video watermarking resilient to inadvertent attacks, in *International Conference on Power Electronics and Renewable Energy Applications (PEREA)*, pp. 1–5 (2020). https://doi.org/10.1109/PEREA51218.2020.9339797
3. F. Arab, M. Zamani, VW16E: a robust video watermarking technique using simulated blocks. Intell. Syst. Ref. Libr. **115**, 193–221 (2017). https://doi.org/10.1007/978-3-319-44270-9_9
4. M. Asikuzzaman, M.R. Pickering, An overview of digital video watermarking. IEEE Trans. Circuits Syst. Video Technol. **28**(9), 2131–2153 (2018). https://doi.org/10.1109/TCSVT.2017.2712162
5. S. Belilita, N. Amardjia, T. Bekkouche, I. Nouioua, Combining SVD-DCT image watermarking scheme based on Perona-Malik diffusion. Elektron. Elektrotechn. **25**(4), 68–74 (2019). https://doi.org/10.5755/j01.eie.25.4.23973
6. A. Bhardwaj, V.S. Verma, R.K. Jha, Robust video watermarking using significant frame selection based on coefficient difference of lifting wavelet transform. Multimed. Tools Appl. **77**(15), 19659–19678 (2018). https://doi.org/10.1007/s11042-017-5340-3
7. B. Chen, G.W. Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding. IEEE Trans. Inf. Theory **47**(4), 1423–1443 (2001). https://doi.org/10.1109/18.923725
8. G. Chen, C. Kang, D. Wang, X Zhao, Y Huang, A robust video watermarking algorithm based on spatial domain, in *International Conference on Energy and Environmental Protection (ICEEP)*, pp. 412–419 (2018). https://doi.org/10.2991/iceep-18.2018.71

9. H. Chen, W. Xu, N. Broderick, J. Han, An adaptive denoising method for Raman spectroscopy based on lifting wavelet transform. J. Raman Spectrosc. **49**(4), 1–11 (2018). https://doi.org/10.1002/jrs.5399
10. L. Chen, Y. Yi, L. Kai, T. Lihua, A semi-fragile video watermarking algorithm based on H.264/AVC. Wirel. Commun. Mob. Comput. **2020**, 8848553–8848563 (2020). https://doi.org/10.1155/2020/8848553
11. S.C. Chu, H.C. Huang, Y. Shi et al., Genetic watermarking for zerotree-based applications. Circuits Syst. Signal Process. **27**, 171–182 (2008). https://doi.org/10.1007/s00034-008-9025-z
12. I. Daubeches, W. Sweldens, Factoring wavelet transform into lifting steps. J. Fourier Anal. Appl. **4**(3), 247–269 (1998). https://doi.org/10.1007/BF02476026
13. B. Dhanalaxmi, S. Tadisetty, Multimedia cryptography—a review, in *International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, pp. 764–766 (2017). https://doi.org/10.1109/ICPCSI.2017.8391817
14. D. Dhaou, S. Ben Jabra, E. Zagrouba, An efficient anaglyph 3d video watermarking approach based on hybrid insertion, in *International Conference on Computer Analysis of Images and Patterns*, pp. 96–107 (2019). https://doi.org/10.1007/978-3-030-29891-3_9
15. M.E. Farfoura, S.J. Horng, J.M. Guo et al., Low complexity semi-fragile watermarking scheme for H.264/AVC authentication. Multimed. Tools Appl. **75**(13), 7465 (2016). https://doi.org/10.1007/s11042-015-2672-8
16. E. Farri, P. Ayubi, A blind and robust video watermarking based on IWT and new 3D generalized chaotic sine map. Nonlinear Dyn. **93**(4), 1875–1897 (2018). https://doi.org/10.1007/s11071-018-4295-x
17. S. Fekri-Ershad, Texture classification approach based on energy variation. Int. J. Multimed. Technol. **2**(2), 52–55 (2012)
18. B. Feng, X. Li, Y. Jie, C. Guo, H. Fu, A novel semi-fragile digital watermarking scheme for scrambled image authentication and restoration. Mobile Netw. Appl. **25**, 82–94 (2020). https://doi.org/10.1007/s11036-018-1186-9
19. V.R. Folifack Signing, T. Fozin Fonzin, M. Kountchou et al., Chaotic Jerk system with Hump structure for text and image encryption using DNA coding. Circuits Syst. Signal Process. **40**, 4370–4406 (2021). https://doi.org/10.1007/s00034-021-01665-1
20. M. Ghadi, L. Laouamer, L. Nana, A. Pascu, A blind spatial domain-based image watermarking using texture analysis and association rules mining. Multimed. Tools Appl. **78**, 15705–15750 (2019). https://doi.org/10.1007/s11042-018-6851-2
21. A. Hammami, A. Ben Hamida, C. Ben Amar, A robust blind video watermarking scheme based on discrete wavelet transform and singular value decomposition, in *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Application*, pp. 597–604 (2019). https://doi.org/10.5220/0007685305970604
22. A. Hammami, A. Ben Hamida, C. Ben Amar, H. Nicolas, Regions based semi-fragile watermarking scheme for video authentication, in *International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, pp. 96–104 (2020). https://doi.org/10.24132/JWSCG.2020.28.12
23. A. Hammami, A. Ben Hamida, C. Ben Amar, Blind semi-fragile watermarking scheme for video authentication in video surveillance context. Multimed. Tools Appl. **80**, 7479–7513 (2021). https://doi.org/10.1007/s11042-020-09982-4
24. H.T. Hu, T.T. Lee, Robust complementary dual image watermarking in subbands derived from the Laplacian pyramid, discrete wavelet transform, and directional filter bank. Circuits Syst. Signal Process. **41**, 4090–4116 (2022). https://doi.org/10.1007/s00034-022-01975-y
25. Y. Huang, C. Liu, X. Zha, Y. Li, An enhanced feature extraction model using lifting based wavelet packet transform scheme and sampling-importance-resampling-analysis. Mech. Syst. Signal Process. **23**(8), 2470–2487 (2009). https://doi.org/10.1016/j.ymssp.2009.06.003
26. Q. Huynh-Thu, M. Ghanbari, Scope of validity of PSNR in image/video quality assessment. Electron. Lett. **44**(13), 800–801 (2008). https://doi.org/10.1049/el:20080522
27. M.S. Islam, N. Naqvi, A.T. Abbasi et al., Robust dual domain twofold encrypted image-in-audio watermarking based on SVD. Circuits Syst. Signal Process. **40**, 4651–4685 (2021). https://doi.org/10.1007/s00034-021-01690-0
28. K. Jain, U.S.N. Raju, A digital video watermarking algorithm based on LSB and DCT. J. Inf. Secur. Res. **6**(3), 92–97 (2015)
29. A. Kanhe, G. Aghila, A DCT-SVD-based speech steganography in voiced frames. Circuits Syst. Signal Process. **37**, 5049–5068 (2018). https://doi.org/10.1007/s00034-018-0805-9

30. A. Kerbiche, S. Ben Jabra, E. Zagrouba, V. Charvillat, Robust video watermarking approach based on crowdsourcing and hybrid insertion, in *International Conference on Digital Image Computing: Techniques and Applications*, pp. 1–8 (2017). https://doi.org/10.1109/DICTA.2017.8227489
31. M.R. Keyvanpour, N. Khanbani, M. Boreiry, A secure method in digital video watermarking with transform domain algorithms. Multimed. Tools Appl. **80**, 20449–20476 (2021). https://doi.org/10.1007/s11042-021-10730-5
32. M. Khan, F. Masood, A novel chaotic image encryption technique based on multiple discrete dynamical maps. Multimed. Tools Appl. **78**, 26203–26222 (2019). https://doi.org/10.1007/s11042-019-07818-4
33. V. Klema, A. Laub, The singular value decomposition: its computation and some applications. IEEE Trans. Autom. Control **25**(2), 164–176 (1980). https://doi.org/10.1109/TAC.1980.1102314
34. R. Mehta, N. Rajpal, V.P. Vishwakarma, A robust and efficient image watermarking scheme based on Lagrangian SVR and lifting wavelet transform. Int. J. Mach. Learn. Cyber. **8**, 379–395 (2017). https://doi.org/10.1007/s13042-015-0331-z
35. A.A. Mohammed, N.A. Ali, Robust video watermarking scheme using high efficiency video coding attack. Multimed. Tools Appl. **77**, 2791–2806 (2018). https://doi.org/10.1007/s11042-017-4427-1
36. R. Munir, H. Harlili, A secure fragile video watermarking algorithm for content authentication based on Arnold Cat Map, in *International Conference on Information Technology (InCIT)*, pp. 32–37 (2019). https://doi.org/10.1109/INCIT.2019.8912074
37. I. Nouioua, N. Amardjia, S. Belilita, A novel blind and robust video watermarking technique in fast motion frames based on SVD and MR-SVD. Secur. Commun. Netw. **2018**(10), 1–17 (2018). https://doi.org/10.1155/2018/6712065
38. H. Prasetyo, C.H. Hsia, C.H. Liu, Vulnerability attacks of SVD-based video watermarking scheme in an IoT environment. IEEE Access **8**, 69919–69936 (2020). https://doi.org/10.1109/ACCESS.2020.2984180
39. C. Priya, C. Ramya, Robust and secure video watermarking based on cellular automata and singular value decomposition for copyright protection. Circuits Syst. Signal Process. **40**, 2464–2493 (2021). https://doi.org/10.1007/s00034-020-01585-6
40. S.P.A. Sathya, S. Ramakrishnan, Fibonacci based key frame selection and scrambling for video watermarking in DWT-SVD domain. Wirel. Pers. Commun. **102**, 2011–2031 (2018). https://doi.org/10.1007/s11277-018-5252-1
41. D. Shukla, M. Sharma, Robust scene-based digital video watermarking scheme using level-3 DWT: approach, evaluation, and experimentation. Radioelectron. Commun. Syst. **61**(01), 1–12 (2018). https://doi.org/10.3103/S0735272718010016
42. P. Simon, V. Uma, Review of texture descriptors for texture classification, in *Data Engineering and Intelligent Computing*, pp. 159–176 (2018). https://doi.org/10.1007/978-981-10-3223-3_15
43. K. Staffy, A. Naveen, D.S. Raahat, Video inter-frame forgery detection approach for surveillance and mobile recorded videos. Int. J. Electr. Comput. Eng. **7**(2), 831–841 (2017). https://doi.org/10.11591/ijece.v7i2.pp831-841
44. W. Sweldens, The lifting scheme: a custom design construction of biorthogonal wavelets. Appl. Comput. Harmonic Anal. **3**(2), 186–200 (1996). https://doi.org/10.1006/acha.1996.0015
45. L. Tian, H. Dai, C. Li, A semi-fragile video watermarking algorithm based on chromatic residual DCT. Multimed. Tools Appl. **79**, 1759–1779 (2020). https://doi.org/10.1007/s11042-019-08256-y
46. Y. Vybornova, A New watermarking method for video authentication with tamper localization, in *International Conference on Computer Vision and Graphics (ICCVG)*, pp. 201–213 (2020). https://doi.org/10.1007/978-3-030-59006-2_18
47. U.A. Waqas, M. Khan, S.I. Batool, A new watermarking scheme based on Daubechies wavelet and chaotic map for quick response code images. Multimed. Tools Appl. **79**, 6891–6914 (2020). https://doi.org/10.1007/s11042-019-08570-5
48. D. Xu, Commutative encryption and data hiding in HEVC video compression. IEEE Access **7**, 66028–66041 (2019). https://doi.org/10.1109/ACCESS.2019.2916484
49. M.A. Yongqiang et al., Research on color image watermarking algorithm based on Quaternion Fourier transform, in *Conference Series: Materials Science and Engineering*, p. 012050 (2020). https://doi.org/10.1088/1757-899X/799/1/012050
50. S. Yuxin, T. Chen, X. Min, C. Mingming, L. Zhenkun, A DWT-SVD based adaptive color multi-watermarking scheme for copyright protection using AMEF and PSO-GWO. Expert Syst. Appl. **168**, 114414 (2021). https://doi.org/10.1016/j.eswa.2020.114414

51. J. Zou, RK. Ward, D. Qi, The generalized Fibonacci transformations and application to image scrambling, in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. iii-385 (2004). https://doi.org/10.1109/ICASSP.2004.1326562

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.