



A video watermark algorithm based on tensor feature map

Shanqing Zhang¹ · Xiaoyun Guo¹ · Xianghua Xu¹ · Li Li¹

Received: 30 March 2020 / Revised: 25 October 2020 / Accepted: 3 December 2022/
Published online: 5 January 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Video has become one of the main ways of information transmission with the development of the Internet. Video copyright protection becomes an urgent task. Video watermark technology embeds copyright into the redundant information of the carrier, and video copyright protection is achieved. However, most video watermark algorithms do not use the correlation and redundancy among adjacent frames of a video and are weak to resist frame attacks. In order to make up this shortage and improve robustness, a video watermark algorithm based on a tensor feature map is proposed. A grayscale video segment with the same scene is selected and represented as a 3-order tensor, a high-order singular value decomposition is performed on the video tensor to obtain a stable core tensor and three factor matrices. A feature tensor is obtained by the mode-3 product of the video tensor with the transpose of the factor matrix that contains a time axis. It is called a tensor feature map. Since the tensor feature map contains the main information of each frame of a video, the watermark is distributed in each frame of a video by embedding the watermark into the tensor feature map. The first-order discrete wavelet transform and discrete cosine transform are used to embed the watermark into the tensor feature map. The experimental results show that the proposed watermark algorithm based on the tensor feature map has better transparency and is robust to common video attacks.

Keywords Video watermark · High-order singular value decomposition · Tensor feature map · Feature tensor

1 Introduction

Video usage and transmission are becoming more and more popular with the development of information technology. Due to the nature of digital information, pirated videos are easily

✉ Xianghua Xu
xhxu@hdu.edu.cn

¹ School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China

made by modifying the contents based on a non-destructive copy of the original video, video copyrights are infringed on the Internet [11]. The need for video copyright protection, proof of ownership, and tampering verification become urgent. Digital watermark technology is the most popular method among the solutions to these problems. Video watermark technology uses a digital watermark in the video carrier. It can embed copyright into the redundant information of the carrier. It is invisible and resistant to malicious attacks. In this way video copyright protection and content certification are achieved.

Video watermark technology is divided into spatio-temporal domain and transform domain watermark technologies, depending on the domain of watermark embedding [5]. The basic idea of time-space video watermarking algorithm is to embed watermarks into the original pixel domain. The algorithm based on plane [16] embedded the watermark information into the lowest four bit planes of the video brightness and used a pseudo-random sequence to determine which plane should be replaced. This algorithm has the advantages of low computational complexity and easy implementation, but its robustness against compression and geometric attacks still needs to be improved. Transform domain watermark technology embeds the watermark in a transform domain of a video signal, for example, the cosine domain or wavelet domain. Although transform domain watermark algorithms are more robust to common attacks, the computational cost is often high. In order to overcome these shortcomings, watermark algorithms based on Singular Value Decomposition (SVD) have attracted researchers. Literature [4] introduced singular value decomposition into digital watermark in 2001 and embedded the watermark image into the singular value of a carrier image. The singular value representing the image's intrinsic characteristics, is stable, can improve algorithm's robustness (especially, to geometric attacks). Literature [21] proposed a real-time video watermarking scheme applicable to MPEG-1 and MPEG-2 videos. In their proposed scheme, they embed watermarks into histograms calculated from low-frequency subbands in the DWT domain. They claim that their scheme provides high robustness and transparency for geometric distortions, including rotation with cropping, zooming, frame drop, aspect ratio variation and swapping. Most singular value decomposition watermark algorithms embed the watermark by modifying singular value [10, 15]. The maximum singular value in a singular value matrix is usually modified to embed the watermark. The watermark image is well concealed. High-Order Singular Value Decomposition (HOSVD) is a generalization of matrix singular value decomposition in multi-dimensional space [19]. It has been widely used in computer vision analysis such as face recognition [6] and gesture recognition [9]. More and more researchers have used tensor in digital watermark, they have proposed many watermark algorithms based on high-order singular value decompositions recently. Because tensor can intuitively represent the structural characteristics of a video, compared to the traditional watermark algorithms, video watermark algorithms based on tensor can make better use of the correlation and redundancy among adjacent frames of a video and further improve algorithms' robustness.

However, most current tensor-based watermark algorithms use the stability of the core tensor and embed the watermark into the core tensor directly to improve the robustness of the watermark algorithm. Among them, Zhang et al. [23] represented a grayscale video as a 3-order tensor, they obtained core tensor using tensor decomposition (i.e. High-Order Singular Value Decomposition) on the video tensor, embedded the watermark by quantizing the maximum of the core tensor. The correlation and redundancy among adjacent frames of a video were used to encapsulate video into the tensor, thereby robustness was improved. However, the transparency of algorithm was at the cost of the capacity of the algorithm. Since

only one bit of watermark was embedded into one tensor, thousands of video frames are required to embed a complete watermark. In order to make full use of the correlation among RGB channels, Xu et al. [22] represented a color image as a tensor, embedded watermark by the parity quantization of core tensor, their algorithm is robust. Zhang et al. [24] represented a color image as a 3-order tensor, extracted the feature map that contains RGB channels of color information from the color image by the mode-3 product of the color image tensor and the transpose of the factor matrix that contains the color information. Embedding the watermark into the feature map guarantees robustness and improves transparency. Because a video sequence consists of a series of still images, the correlation and redundancy among adjacent frames of a video are considered, we propose a video watermark algorithm based on the tensor feature map. The watermark embedding process is shown in Fig. 1. Firstly, a grayscale video segment with the same scene is represented as 3-order tensor, and high-order singular value decomposition is performed on the video tensor to obtain a stable core tensor and three factor matrices. The three factor matrices represent the information of row, column and time axis of video frames, respectively. Secondly, a feature tensor of the same size as video tensor is obtained by the mode-3 product of the video tensor and the transpose of the factor matrix of the time axis. The first front slice of the feature tensor contains the main information and is called tensor the feature map. Finally, Discrete Wavelet Transformation (DWT) is performed on the tensor feature map to obtain four sub bands: LL, LH, HL, and HH. Sub band LL is the coefficients after low-pass filtering and concentrates the most energy of an image. Sub band LL is divided into blocks of size 8×8 , Discrete Cosine Transform (DCT) is performed on each block and the watermark is embedded into the coefficient matrix of transform domain. The robustness and transparency of the algorithm is ensured.

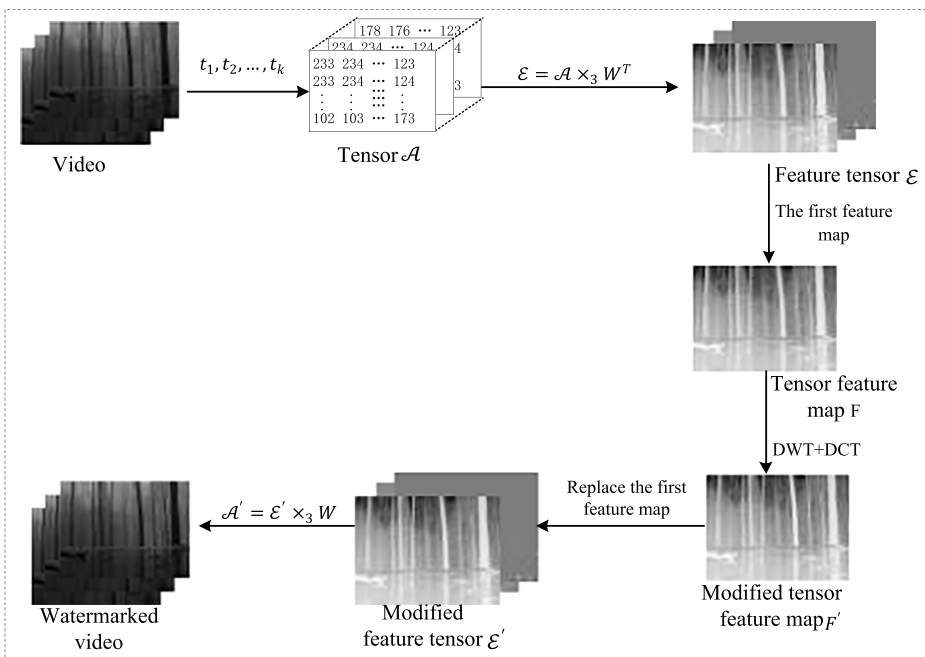


Fig. 1 Embedding process of watermark (extraction is the inverse process of embedding)

The innovations of the study are as follows:

- (1) The tensor is used to video watermark, the redundancy and correlation among adjacent frames of a video are used, the robustness to common video attacks is improved.
- (2) The tensor feature map that contains the information of video frames is proposed, the watermark is distributed in different video frames by embedding the watermark into the tensor feature map, then the robustness and transparency are ensured.

The next sections of this paper are organized as follows. The basic knowledge related to the proposed algorithm is introduced in Section 2, the watermark embedding and extracting algorithms are described in details in Section 3, the experimental results are in Section 4 to verify the performance of the proposed watermark algorithm. The final section is about conclusions.

2 Tensor

The following notations are used in this paper:

Scalar: (a, b, etc); Vector: ($\mathbf{a}, \mathbf{b}, etc$); Matrix: (A, B, etc); High order tensor: ($\mathcal{A}, \mathcal{B}, etc.$).

A tensor is the generalization of a vector or a matrix, a vector or a matrix may be regarded as the 1-order or 2-order tensor, respectively [14]. A N -order tensor is defined as $\mathcal{A} \in R^{I_1 \times \dots \times I_N}$. A grayscale video is represented as a 3-order tensor, its three dimensions are the width, height, and duration of the video.

2.1 Slices of tensor

A slice is a two-dimensional part of tensor, that is, a matrix extracted from a tensor [13]. The two indexes are fixed and the rest index is used. The horizontal, vertical, and front slices of a 3-order tensor \mathcal{A} are shown in Fig. 2. They are represented by $\mathcal{A}_{i::}$, $\mathcal{A}_{:j}$, $\mathcal{A}_{::k}$ respectively.

2.2 Tensor expansion

Tensor has many advantages in representing multi-dimension data. For example, the feature information of a video is retained to the maximum extent after encapsulating the video as a

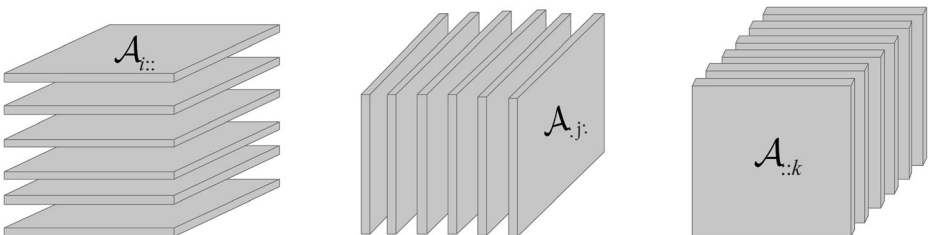


Fig. 2 Three slices of a 3-order tensor

tensor, the correlation and redundancy among adjacent frames are fully used. However, high-dimensional characteristics need heavy calculation. A tensor is usually expanded into matrices according to different orders to facilitate calculation. Tensor expansion is the process of rearranging elements in a tensor into matrices in order and is defined as follows [20]:

Given a N -order tensor $\mathcal{A} \in R^{I_1 \times \dots \times I_N}$, \mathcal{A} is unfolded into a series of matrices $A_1, \dots, A_N, A_k \in R^{I_k \times (I_1 \times \dots \times I_{k-1} \times I_{k+1} \times \dots \times I_N)}$. The expansion of a 3-order tensor is shown in Fig. 3. The third order tensor \mathcal{A} is sliced according to three axes of the space coordinate system and then the tensor slices are rearranged into the second order matrix A_i .

2.3 The mode- n product of tensor and matrix

The mode- n product of a N -order tensor $\mathcal{A} \in R^{I_1 \times \dots \times I_N}$ and a matrix $B \in R^{J \times I_n}$ is denoted as $\mathcal{A} \times_n B$, where $\mathcal{A} \times_n B \in R^{I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N}$. The mode-1 product of a 3-order tensor and a matrix is shown in Fig. 4 [8]. The entries are given by:

$$(\mathcal{A} \times_n B)_{i_1 \times \dots \times i_{n-1} \times j \times i_{n+1} \times \dots \times i_N} = \sum_{i_n=1}^{I_n} a_{i_1 \dots i_N} b_{j i_n} \quad (1)$$

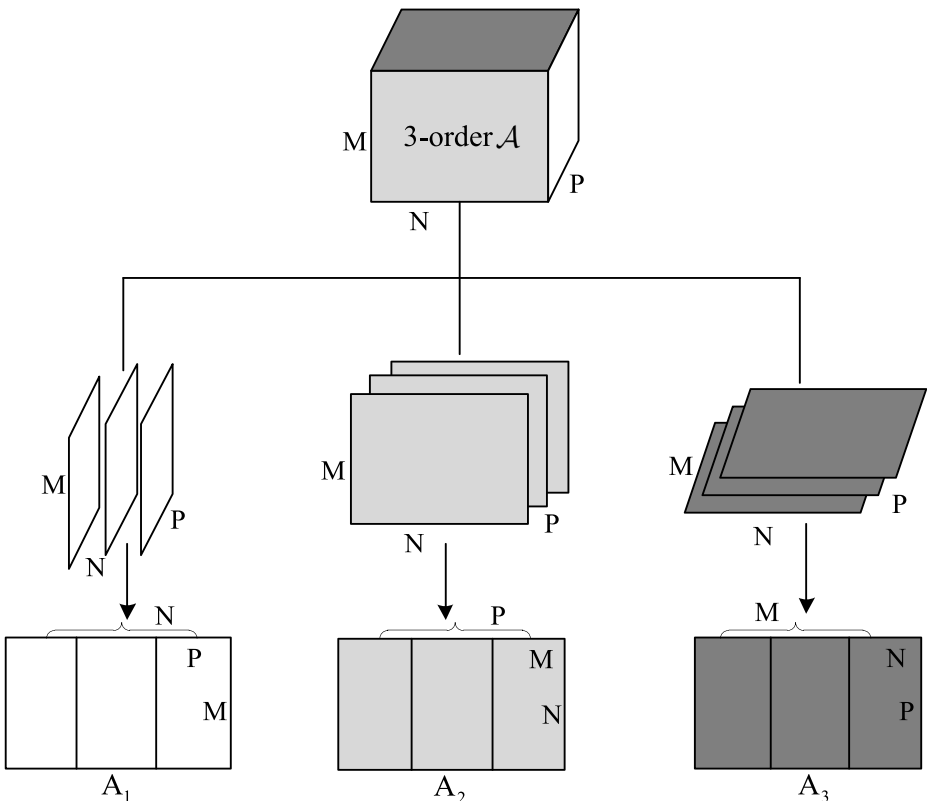


Fig. 3 Matrix expansion of a 3-order tensor

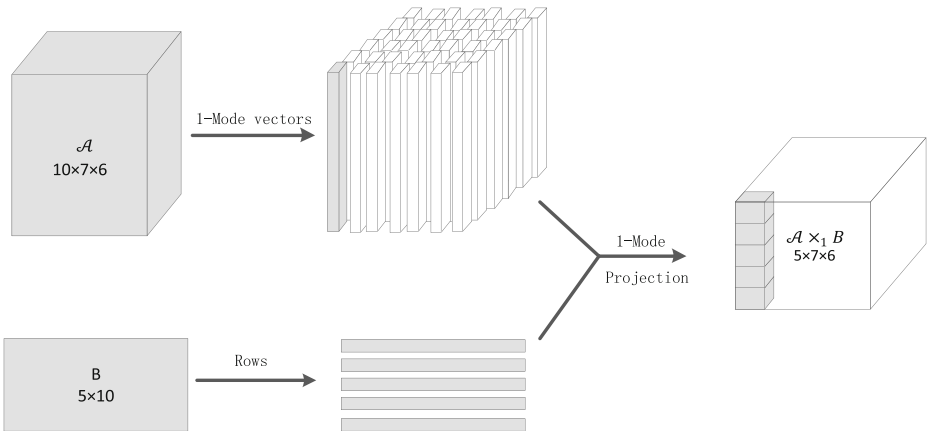


Fig. 4 The mode-1 product of a 3-order tensor and a matrix

2.4 High-order singular value decomposition

For a two-dimensional matrix, the singular value decomposition of image I of size $I_1 \times I_2$ is represented by the product of three matrices, namely:

$$I_{I_1 \times I_2} = U_{I_1 \times I_1} \cdot S_{I_1 \times I_2} \cdot V_{I_2 \times I_2}^T \tag{2}$$

where U and V are the left and right singular matrix of matrix I , respectively. S is a diagonal matrix that contains the singular values of matrix I , and its elements are arranged in decreasing order [17].

High-order singular value decomposition(HOSVD) is the generalization of matrix's singular value decomposition in high-dimension. Given a 3-order tensor \mathcal{A} of size $M \times N \times K$, the high-order singular value decomposition of \mathcal{A} is represented as follows [7]:

$$\mathcal{A} = \mathcal{S} \times_1 U \times_2 V \times_3 W \# \tag{3}$$

where factor matrices U , V , and W satisfy:

$$A_1 = U \cdot D_1 \cdot V_1^T, \tag{4}$$

$$A_2 = V \cdot D_2 \cdot V_2^T, \tag{5}$$

$$A_3 = W \cdot D_3 \cdot V_3^T. \tag{6}$$

where A_1 , A_2 , and A_3 are the expanded matrices of the tensor \mathcal{A} along three directions, respectively. The core tensor \mathcal{S} is represented as:

$$\mathcal{S} = \mathcal{A} \times_1 U^T \times_2 V^T \times_3 W^T \# \tag{7}$$

An example of HOSVD decomposition of a 3-order tensor is shown in Fig. 5, where $1 \leq R_1 \leq M$, $1 \leq R_2 \leq N$, $1 \leq R_3 \leq K$.

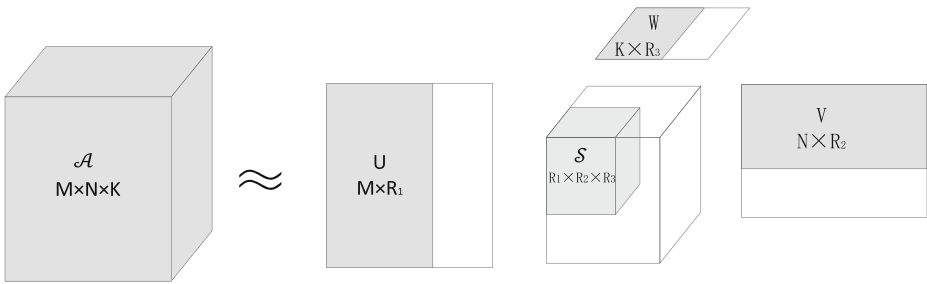


Fig. 5 HOSVD of a 3-order tensor

In order to make full use of the correlation and redundancy among adjacent frames of a video, the feature tensor of representing the time axis is extracted from the video tensor, that is, $\mathcal{E} = \mathcal{S} \times_1 U \times_2 V$, and its size is the same as that of the video tensor. Expression (3) can be written as follows:

$$\mathcal{A} = \mathcal{E} \times_3 W \# \tag{8}$$

Since the rows in W represent the correlation among different frames of a video, according to (8), the linear relationship between W and video tensor \mathcal{A} is represented as follows:

$$\mathcal{A}_t = W_{t,1} \mathcal{E}_1 + W_{t,2} \mathcal{E}_2 + \dots + W_{t,K} \mathcal{E}_K \# \tag{9}$$

where \mathcal{A}_t represents the t -th frame of the video tensor \mathcal{A} , and \mathcal{E}_i represents the i -th front slice of feature tensor \mathcal{E} . $W_{t,i}$ represents the i -th value of the t -th row in the factor matrix W .

According to Expression (8), the feature tensor of representing the time axis is obtained using the modulo-3 product of the video tensor \mathcal{A} and the transpose of factor matrix W , that is:

$$\mathcal{E} = \mathcal{A} \times_3 W^T \# \tag{10}$$

and

$$\|\mathcal{E}_1\|^2 \geq \|\mathcal{E}_2\|^2 \geq \dots \geq \|\mathcal{E}_K\|^2 \tag{11}$$

while \mathcal{E}_i represents the i -th front slice of \mathcal{E} , and $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_K$ are regarded as the first, second, ..., and K -th feature map of the video, respectively. $\|\mathcal{E}_i\|$ represents the Frobenius norm of the i -th feature map \mathcal{E}_i . The extracted feature map after HOSVD decomposition of a video of size 480*640*4 is shown in Fig. 6. The first feature map obviously contains the main information of the video, and is therefore called the tensor feature map.

3 The proposed watermark algorithm

In order to make full use of the correlation and redundancy among adjacent frames of a video, we represent a grayscale video segment with the same scene as a 3-order tensor, and extract the feature tensor of representing the time axis from the video tensor through high-order singular

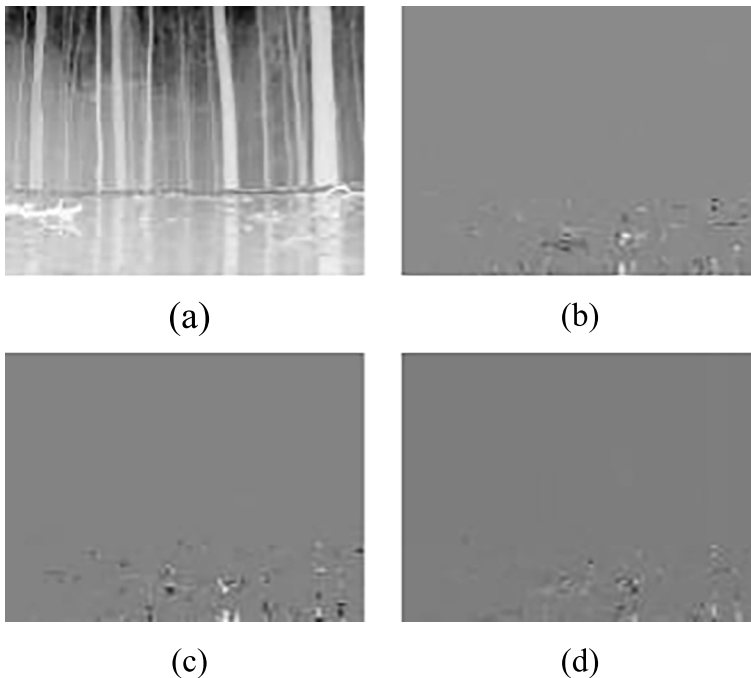


Fig. 6 Examples of feature map

value decomposition. The first front slice of a feature tensor contains the main information of each frame of the video and is called a tensor feature map.

3.1 The process of watermark embedding

Given a video segment V of resolution $M \times N$, the size of the watermark B is assumed to be $m \times n$. The correlation and redundancy among adjacent frames of a video considered, a grayscale video of K frames is represented as a 3-order tensor, the size of tensor \mathcal{A} is $M \times N \times K$. The process of watermark embedding is as follows:

- (1) Watermark image preprocessing. In order to eliminate the spacious correlation among binary watermark pixels, the watermark image is scrambled using Logistic chaotic sequence. The scrambled watermark is denoted as B' .
- (2) Obtaining the tensor feature map. The feature tensor \mathcal{E} is obtained by the mode-3 product of the video tensor \mathcal{A} and the transpose of factor matrix W of containing time axis, and the first front slice of \mathcal{E} is a tensor feature map F .

$$\mathcal{E} = \mathcal{A} \times_3 W^T \# \quad (12)$$

W is obtained using $A_3 = W \cdot D_3 \cdot V_3^T$.

- (3) Obtaining sub band LL. Sub band LL is obtained by the level-1 DWT transformation on tensor feature map.

- (4) Obtaining the transform coefficient matrix. Sub band LL is divided into blocks of size 8×8 , the DCT transform is performed on each block to obtain the transform coefficient matrix.
- (5) Embedding the watermark by modifying the mid-frequency coefficient of the transform coefficient matrix.
 - a. A mid-frequency coefficient $y_{i,j}(u, v)$ in the transform coefficient matrix and its six adjacent coefficients in Zig-Zag sorting are selected. We use $u = 4$ and $v = 4$.
 - b. Calculating the average of the six adjacent coefficients.

$$avg_{i,j} = \frac{[y_{i,j}(u+3, v-3) + y_{i,j}(u+2, v-2) + y_{i,j}(u+1, v-1) + y_{i,j}(u-1, v+1) + y_{i,j}(u-2, v+2) + y_{i,j}(u-3, v+3)]}{6} \tag{13}$$

- c. Modifying $y_{i,j}(u, v)$ according to the scrambled binary watermark B' .

$$y'_{i,j}(u, v) = \begin{cases} avg_{i,j} + T & \text{if } B'(i, j) = 1 \\ avg_{i,j} - T & \text{if } B'(i, j) = 0 \end{cases} \tag{14}$$

Where $1 \leq i \leq m, 1 \leq j \leq n$, T is the quantization intensity and its value is discussed below.

- (6) Reconstructing the tensor feature map after embedding the watermark.
 - a. IDCT transformation is performed on each modified transform coefficient matrix, and every block is merged into the watermarked sub band LL in order.
 - b. IDWT transform is performed on the watermarked sub band LL, the watermarked tensor feature map $F'T'$ is obtained.
- (7) Reconstructing the watermarked video
 - a. The watermarked tensor feature map is merged into the first front slice of the feature tensor, the watermarked feature tensor \mathcal{E}' is obtained.
 - b. Watermarked video \mathcal{A}' is reconstructed from the watermarked feature tensor \mathcal{E}' .

3.2 The process of extracting watermark

Watermark extraction, the reversed process of watermark embedding, is as follows:

- (1) Obtaining the watermarked tensor feature map. The feature tensor \mathcal{E}' is the mode-3 product of the watermarked video \mathcal{A}' and the transpose of factor matrix W' of representing time axis. The first front slice of \mathcal{E}' is the watermarked tensor feature map F' .

$$W' \text{ is obtained from } A'_3 = W' \cdot D'_3 \cdot V'^T \mathcal{E}' = \mathcal{A}' \times_3 W'^T \tag{15}$$

- (2) Obtaining the watermarked sub band LL. The level-1 DWT transform is performed on the watermarked tensor feature map, sub band LL is obtained.
- (3) Obtaining a transform coefficient matrix. The watermarked sub band LL is divided into blocks of size 8×8 . DCT transform is performed on each block to obtain a transform coefficient matrix.
- (4) Watermark extraction is determined according to the relationship between element $y'_{i,j}$ (u, v) and the average $avg'_{i,j}$ of its six adjacent coefficients in the transform coefficient matrix.

$$B'(i, j) = \begin{cases} 1, & \text{if } y'_{i,j}(u, v) > avg'_{i,j} \\ 0, & \text{if } y'_{i,j}(u, v) \leq avg'_{i,j} \end{cases} \quad (16)$$

- (5) Watermark B' is restored by scrambling reversely using Logistic chaotic sequence.

4 Experimental results and performance analysis

In our experiment, a set of 480×640 videos and a watermark B with size of 30×40 are used. In one scene of the video, 8 frames in the video are represented as a third-order tensor, then the video tensor is decomposed with HOSVD to a gain core tensor and three factor matrix. Secondly, the feature tensor is obtained by multiplying the video tensor and the factor matrix transposes with the time axis information. Finally, the Discrete Wavelet Transformation (DWT) is applied on the first slice of the feature tensor to obtain four subbands (LL, LH, HL, HH). Further, LL subband is divided into 8×8 subblocks, and Discrete Cosine Transform (DCT) is performed on each subblock, thus embedding watermark into the Transform domain coefficient matrix, which ensured the robustness and transparency of the algorithm.

4.1 Metrics

In order to verify the performance of the proposed video watermark algorithm, the following metrics are used to measure the robustness and transparency of the algorithm. Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE) are used to evaluate the transparency of the algorithm objectively. MSE is calculated as follows:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |I(i, j) - I'(i, j)|^2 \quad (17)$$

where M and N are the height and width of a single frame image, respectively. I and I' are the original and the watermarked video frame, respectively. The smaller the MSE value is, the smaller the difference between a single watermarked image and the original image will be. PSNR is easily calculated using MSE as follows:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (18)$$

The smaller the PSNR of the watermarked video frame from the original video frame, the greater the distortion of a single watermarked frame image will be [1]. In addition, the

robustness of the algorithm is evaluated using normalized correlation coefficient (NC). The expressions are as follows:

$$NC = \frac{\sum_{i=1}^m \sum_{j=1}^n (B'(i, j) \cdot B(i, j))}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (B(i, j))^2} \sqrt{\sum_{i=1}^m \sum_{j=1}^n (B'(i, j))^2}} \tag{19}$$

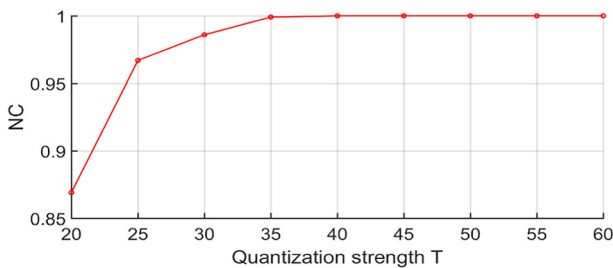
where m and n are the width and height of the watermark, and B and B' are the original and extracted watermark, respectively. The larger NC is, the more robust the algorithm will be.

4.2 Experimental parameters

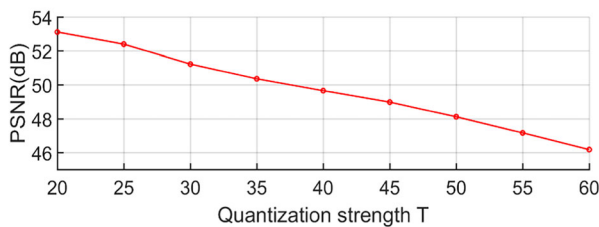
The size of the test video is 480×640 , and the size of the watermark is 30×40 . A grayscale video of K frames is regarded as a whole unit, that is, the size of the video tensor is $480 \times 640 \times K$, and $K = 8$.

The relationship among quantization intensity T , NC of the watermark and the average PSNR of the video is shown in Fig. 7 when there is no attack. NC of watermark increases as quantization intensity T increases, and the average PSNR of the video decreases as quantization intensity T increases. When $T \geq 40$, NC of watermark is 1, and the average PSNR of the video is above 45db. Both the robustness and transparency of algorithm are considered, we set $T = 50$.

The original watermark image and the scrambled watermark image using Logistic chaotic sequence are shown in Fig. 8. Scrambled watermark image eliminates spatial correlation between binary watermark pixels.



(a)



(b)

Fig. 7 a relationship between quantization strength T and watermark NC b relationship between quantization strength T and video average PSNR

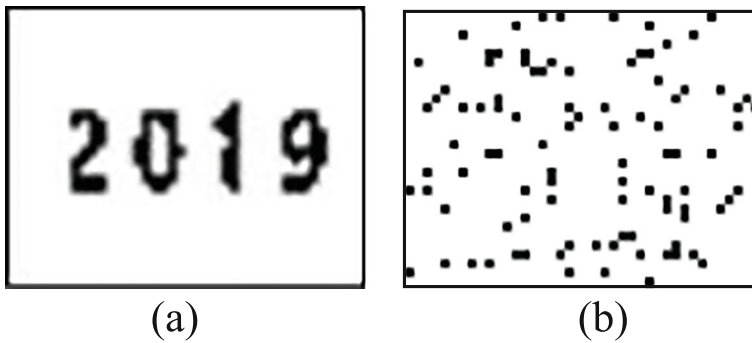








Fig. 8 a original watermark image b scrambled watermark image

4.3 Transparency test

In our method, embedding watermark in the tensor feature map can improve the transparency of the watermarking algorithm. The experimental result of the proposed watermark algorithm is shown in Table 1. The difference between the original and the watermarked video is not obvious to human eyes. In addition, in Table 1, we list the PSNR of our algorithm using DWT, our algorithm without DWT and the algorithm in [23]. Our algorithm has higher transparency than the algorithm in [23], and our algorithm using DWT has higher transparency.

Table 1 Watermarked video examples

Original video	Watermarked video	PSNR with DWT	PSNR without DWT	PSNR [23]
		47.8913	47.9850	44.8913
		48.1160	47.4322	45.1160
		48.1417	44.6049	45.4417

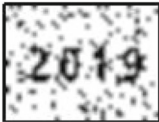

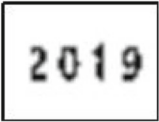
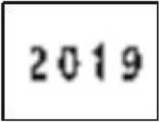
4.4 Robustness test

Some experiments are used to test the robustness of the proposed watermark algorithm to video attacks (e.g. Frame deletion, scaling, rotation, cropping, and noise adding). We used a set of grayscale videos as test data, the experimental results show that the watermark algorithm based on the tensor feature map is robust to these attacks. NC of the extracted watermark for a tensor of size $480 \times 640 \times 8$ with the deletion from 1 to 5 frames are shown in Table 2. NC of the extracted watermark is high enough even if 50% of the total video frames are deleted. The proposed algorithm is more robust than the algorithm in [23] in the case of frame deletion. Because the watermark is evenly distributed in every frame of a video, every frame contains a part of watermark. The proposed algorithm can extract the

Table 2 Results of frame deletion attack

Number of frames deleted	NC (Proposed)	NC [23]
1	1	1
2	0.9992	0.9933
3	0.9958	0.9925
4	0.9817	0.9533
5	0.9475	0.8825

Table 3 Results of scaling attack

Scaling ratio	Extracted watermark	NC(ours)
0.7		0.9367
0.9		0.9825
1.2		1
1.5		1

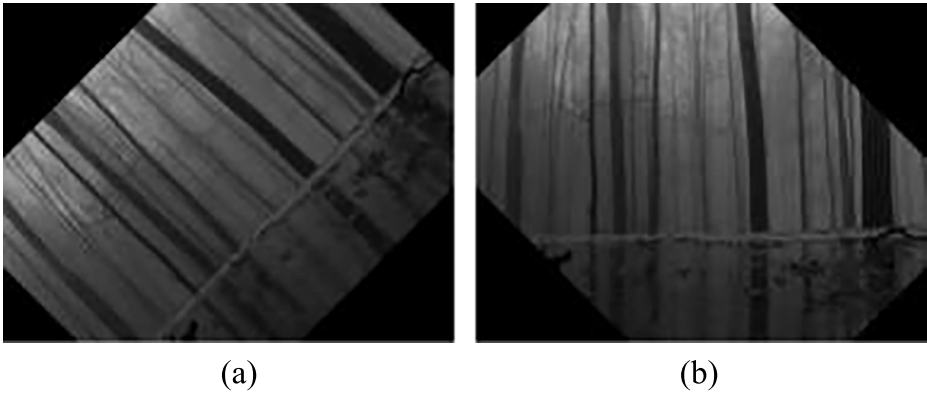


Fig. 9 Example of rotation correction

embedded watermark based on the embedded information in only a small portion of total video frames.

Geometric attacks usually cause information to be out of sync, therefore most watermark extraction algorithms are ineffective in such cases. Scaling and rotation are common geometric attacks. Attacks with the scaling ratios of 0.7, 0.9, 1.2, and 1.5 are used in our experiment. NC of the extracted watermark at different ratios are shown in Table 3. The results show that no matter whether the video is reduced or enlarged, the proposed algorithm can extract the watermark correctly.

In the experiment with rotation attack, clockwise rotation angles from 5 to 90 degrees are used. The rotation correction algorithm [12, 18] is used before the watermark extracting. An example of video rotation correction is shown in Fig. 9. The extracted watermark and NC are shown in Table 4. Both our algorithm and the algorithm in literature [23] have the performance of resisting rotation attack.

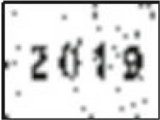
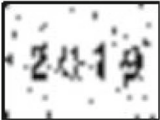
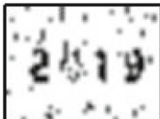
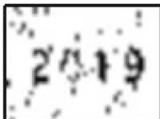
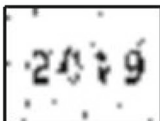
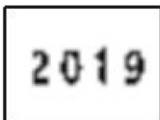
The extracted watermark and NC after video cropping attack are shown in Table 5. Our algorithm has a slightly better robustness against cropping attacks than literature [23].

Compression is the most vulnerable video attack. Video watermark algorithms should be robust to compression attack. H.264 compression is used in watermarked video. The watermark and NC extracted using different compression ratios are shown in Table 6. The experimental results show that the proposed algorithm is robust to compression attack, when the compression ratio is above 60%, our algorithm can fully extract the watermark information, and the watermark information extracted when the compression ratio is 20% has only slight error.

In addition, the video is susceptible to noise pollution both in storage and transmission. The performance of the algorithm is tested with noise attack. Common noises are salt and pepper noise, Poisson noise, Gaussian noise and so on. Noise-enhanced videos are shown in Fig. 10, and NCs of the extracted watermark are shown in Table 7. All NCs in Table 6 are above 0.9, therefore the algorithm is robust to common noise attacks.

In order to further prove the effectiveness of our algorithm, we compare it with the watermarking schemes proposed in literature [2, 3]. In [3], all the videos are of size 352×288 with number of frames 300, the watermark of size 32×32 is presented. In [2], all the videos are of size 352×288 with number of frames 300, the watermark of size 23×46 is

Table 4 Results for rotation attack

Rotation angle	Extracted watermark	NC(ours)	NC[23]
15		0.9800	0.9817
30		0.9750	0.9775
45		0.9667	0.9725
60		0.9717	0.9700
75		0.9825	0.9650
90		1	0.9917

presented. In Table 8, we use NC values to evaluate the robustness of each algorithm, and it can be seen that our algorithm is more robust for equalization, frame deletion, and video compression.

5 Conclusions

In order to make full use of the correlation and redundancy among adjacent frames of a video, a video is encapsulated as a tensor and high-order singular value decomposition is used to obtain the feature tensor of representing its main energy in this study. The first front slice of the feature tensor is the tensor feature map that contains the main information of each frame of a video. According to the stability of the tensor feature map, a video watermark algorithm based on the tensor feature map is proposed. Different from the general video watermark algorithms,

Table 5 Results for cropping attack


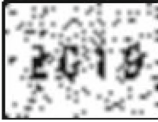
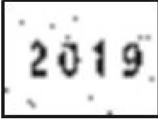
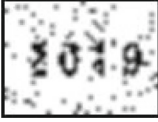
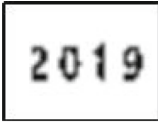
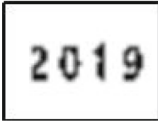
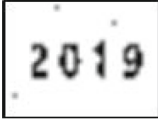
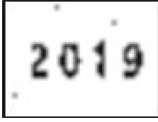
Cropping attack	Extracted watermark	NC	NC[23]
Crop left 1 column		0.9800	0.9317
Crop left 2 column		0.9275	0.9121
Crop top 1 line		0.9933	0.9758
Crop top 2 line		0.9600	0.9583

Table 6 Results for compression attack

Compression ratio	Extracted watermark	NC(ours)
80%		1
60%		1
40%		0.9983
20%		0.9983

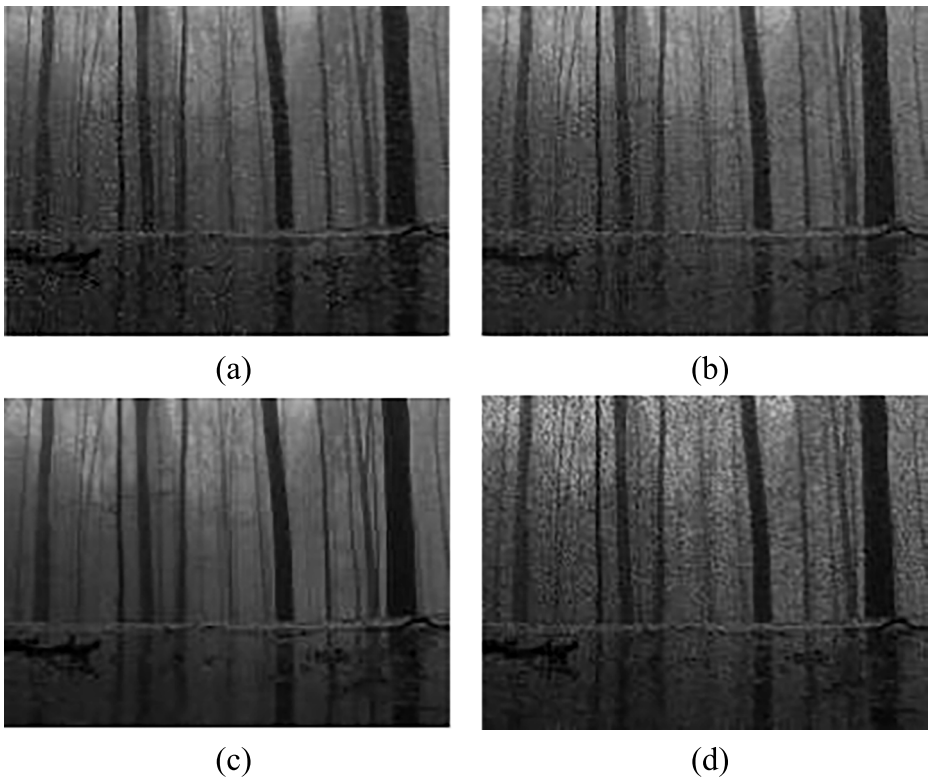


Fig. 10 Example of video noise adding

Table 7 Results for noise attack

Noise type	(mean, variance)	NC(ours)
Gaussian	(0, 0.01)	0.9258
	(0, 0.02)	0.9333
	(0, 0.04)	0.9158
Salt & Pepper	(0, 0.01)	0.9800
	(0, 0.02)	0.9550
	(0, 0.04)	0.9017
Speckle	(0, 0.01)	0.9983
	(0, 0.05)	0.9750
	(0, 0.1)	0.9583
Poisson	\	0.9983

Table 8 Comparison with [2, 3] for different attacks

Algorithm	Equalization	Frame Deletion(50%)	Video Compression(50%)
[3]	0.7793	\	0.9941
[2]	1	0.9263	0.9007
Our	1	0.9813	0.9983

tensor is used in video watermark, the correlation and redundancy among adjacent frames of a video are used to improve the robustness of the algorithm to common video attacks. The watermark information is distributed in each frame of a video by embedding the watermark into the tensor feature map. The robustness and transparency of the algorithm are ensured. However, because a cropped video causes energy loss to its feature tensor, the robustness of the algorithm to cropping attack should be further improved in the future.

Acknowledgments This work was supported by the Key Research and Development Project of Zhejiang Province, under Grant 2020C01067, Public Welfare Technology and Industry Project of Zhejiang Provincial Science Technology Department under Grant LGG18F020013, LGG19F020016.

Declarations

Conflict of interest The author declare no conflicts of interest.

References

1. Akhter R, Sazzad ZMP, Horita Y (2010) No-reference stereoscopic image quality assessment. *Proc SPIE Int Soc Opt Eng* 7524(2):75240T–75240T-12
2. Bahrami Z, Akhlaghian Tab F (2016) A new robust video watermarking algorithm based on SURF features and block classification. *Multimed Tools Appl* 7(1):327–345
3. Bhardwaj A, Verma VS, Jha RK (2017) Robust video watermarking using significant frame selection based on coefficient difference of lifting wavelet transform. *Multimed Tools Appl* 77(15):19659–19678
4. Chandra DVS (2002) Digital image watermarking using singular value decomposition
5. Cox IJ, Miller ML, Bloom JA (2005) Digital watermarking. Springer, Berlin
6. Fu Y, Ruan Q, Luo Z (2019) FERLrTc: 2D+3D facial expression recognition via low-rank tensor completion. *Signal Process* 161:74–88
7. Gao S, Guo N, Zhang M (2019) Image denoising based on HOSVD with iterative-based adaptive hard threshold coefficient shrinkage. *IEEE Access* 7:13781–13790
8. Hemalatha P (2018) Gregarious star factorization of the tensor product of graphs. *Discrete Math Algorithms Appl* 10:1850055
9. Hsieh CY, Lin WY (2017) Video-based human action and hand gesture recognition by fusing factored matrices of dual tensors. *Multimed Tools Appl* 76(6):7575–7594
10. Jia S-I (2014) A novel blind color images watermarking based on SVD. *Optik* 125(12):2868–2874
11. Juergen S (2005) Digital watermarking for digital media. Information Science Publishing, pp 2–5
12. Koivisto PK (2005) Elimination of rotation axes offset error in antenna radiation pattern measurements using spherical wave expansion. *AEU – Int J Electron Commun* 59(7):379–383
13. Kong WW (2014) Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform. *Infrared Phys Technol* 63:110–118
14. Liu Y, Shang F (2013) An efficient matrix factorization method for tensor completion. *IEEE Signal Process Lett* 20(4):307–310
15. Makbol NM, Khoo BE (2013) Robust blind image watermarking scheme based on redundant discrete wavelet transform and singular value decomposition. *AEU Int J Electron Commun* 67(2):102–112
16. Mobasseri BG (2000) A spatial digital video watermark that survives MPEG. *Proceedings international conference on information technology*, pp 68–73
17. Nouioua I, Amardjia N, Belilita S (2018) A novel blind and robust video watermarking technique in fast motion frames based on SVD and MR-SVD. *Secur Commun Netw* 10:1–17
18. Rudant L, Belmakkadem K, Vuong TP (2015) Analysis of open-slot antenna radiation pattern using spherical wave expansion. *Iet Microw Antennas Propag* 9(13):1407–1411
19. Shi Q, Cheung YM, Zhao Q (2018) Feature extraction for incomplete data via low-rank tensor decomposition with feature regularization. *IEEE Trans Neural Netw Learn Syst* 1–15
20. Sun W, Chen Y, Huang L (2018) Tensor completion via generalized tensor tubal rank minimization using general unfolding. *IEEE Signal Process Lett* 1-1

21. Wang L, Ling H, Zou F, Lu Z (2012) Real-time compressed-domain video watermarking resistance to geometric distortions. *IEEE Multimedia*
22. Xu H, Jiang G, Yu M (2018) Color image watermarking based on tensor analysis. *IEEE Access* 99:1–1
23. Zhang SQ, Guo XY, Xu XH (2019) A video watermark algorithm based on tensor decomposition. *Math Biosci Eng* 16(5):3435–3449
24. Zhang F, Luo T, Jiang G (2019) A novel robust color image watermarking method using RGB correlations. *Multimed Tools Appl* 78(14):20133–20155

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.