

New Video Watermark Scheme Resistant to Super Strong Cropping Attacks

Ming Tong, Tao Chen, Wei Zhang, Linna Dong

School of Electronic Engineerin, Xidian University, Xi'an, China

Email: mtong@xidian.edu.cn

Received January 14, 2012; revised February 28, 2012; accepted March 20, 2012

ABSTRACT

Firstly, the nonnegative matrix factorization with sparseness constraints on parts of the basis matrix (NMFSCPBM) method is proposed in this paper. Secondly, the encrypted watermark is embedded into the big coefficients of the basis matrix that the host video is decomposed into by NMFSCPBM. At the same time, the watermark embedding strength is adaptively adjusted by the video motion characteristic coefficients extracted by NMFSCPBM method. On watermark detection, as long as the residual video contains the numbers of the least remaining sub-blocks, the complete basis matrix can be completely recovered through the decomposition of the nonnegative matrix of the least remaining sub-blocks in residual videos by NMFSCPBM, and then the complete watermark can be extracted. The experimental results show that the average intensity resistant to the various regular cropping of this scheme is up to 95.97% and that the average intensity resistant to the various irregular cropping of this scheme is up to 95.55%. The bit correct rate (BCR) values of the extracted watermark are always 100% under all of the above situations. It is proved that the watermark extraction is not limited by the cropping position and type in this scheme. Compared with other similar methods, the performance of resisting strong cropping is improved greatly.

Keywords: Digital Watermarking; Cropping Attack; Geometric Attacks; Nonnegative Matrix Factorization (NMF); Sparseness Constrain

1. Introduction

The watermarking robustness has been a focus in multimedia research field, and how to resist geometric attacks is the hotspot and difficulty of the study [1]. Geometric attacks destroy the synchronization between the watermark and the videos, which seriously affect the robustness of the watermarking and pose a deadly threat to the watermarking security. At the same time, with the appearance and maturity of a variety of video signal processing tools, the video data can be cropped, copied and tampered in various forms and different degrees much more conveniently, faster and more casually. Especially to strong cropping, the embedded watermark information is cropped directly and enormously and the copyright of digital video products is facing serious challenges. How to extract and recover the complete watermark in residual videos has always troubled researchers [2]. Consequently, the research on video watermark methods resisting to strong cropping attacks can help strengthening the copyright authentication and management, standardizing the market of video products, and solving the technology bottleneck problem of commercializing watermark, which has important academic research value and extensive market application prospect.

A four-level dual-tree complex wavelet transform (DT CWT) is applied to every video frame of the host video in [3], which embeds the watermark in the coefficients of low frequency sub-band. This method takes the advantage of perfect reconstruction, shift invariance, and good directional selectivity of DT CWT, and can resist the cropping attacks with a certain strength. In [4], the watermark is embedded into 8×8 sub-blocks of each I-frame, which is robust to row and column cropping attacks with small strength. Since the geometrical distortions operations for every frame along the time axis in a video sequence are the same [5], the watermark is embedded into the same position in the video frame by modifying the pixels of texture complex or sports intense area. But the information of the watermarks embedded by this method is less, and the robustness is greatly influenced by host carrier characteristics. So this method cannot resist cropping attacks. In [6], the watermark is embedded into the singular value of the coefficient matrix of nonnegative matrix factorization (NMF). This method is robust to noise, filtering and other general attacks, but cannot resist strong cropping attacks. Theoretical analysis and experimental results show that most existing video watermarking methods [2-7] resisting to

geometric attack can resist cropping attacks with smaller strength, but are sensitive to strong cropping attacks.

Nonnegative matrix factorization (NMF) is a matrix factorization method under the condition that all the elements of the matrix are nonnegative, which can greatly reduce the dimensions of the data. The decomposition characteristics meet the experience of human visual perception, and the decomposition results have interpretability and clear physical meaning. Since it is proposed, NMF has been paid to great attentions, and it succeeds in the application to pattern recognition, computer vision and image engineering etc [8-11].

This paper proposes a new video watermarking scheme based on nonnegative matrix factorization with sparseness constraints on parts of the basis matrix (NMFSCPBM), in which the watermark is embedded into the basis matrix of NMFSCPBM. It uses video motion coefficients to adaptively control the watermark embedding strength. This scheme can resist various high-intensity cropping attack, and experimental results show the effectiveness of this scheme.

2. Complete Basis Matrix Recovery and Video Motion Components Extraction Based on NMFSCPBM

The NMF with sparseness constraints (NMFSC) method is proposed in [12], which uses a nonlinear projection operator to achieve the precise control of the sparseness by adding sparseness constraints in all basis vectors. But the data is described insufficiently under higher sparseness constraints. A NMFSCPBM method is proposed in this paper, in which controllable sparseness constraints are added in the part of basis vectors. Not only can the NMFSCPBM method extract video motion components quickly and correctly, and filter out the static background interference completely, but also solve the problems of the NMFSC method under higher sparseness constraints. Meanwhile it reduces the decomposition error and speeds up the convergence rate.

2.1. NMFSCPBM Model Construction

Sparse matrix is the matrix that most of the elements are zero or approach to zero and only a few are nonzero. The sparse degree of a vector is defined as Equation (1).

$$\text{sparseness}(y) = \frac{\sqrt{n} - (\sum |y_i|) / \sqrt{\sum y_i^2}}{\sqrt{n} - 1} \quad (1)$$

where n is the dimension of y [12].

So the NMFSCPBM can be transformed into the following constrained optimization problem that given a nonnegative matrix B of size $m \times n$, solve the basis matrix W of size $m \times r$ and coefficient matrix H of size $r \times n$, where r is the decomposition dimension of B .

If squared Euclidean distance D of primitive matrix and reconstructed matrix is defined as the target evaluation function, shown as Equation (2), which is used to describe the error between primitive matrix B and reconstructed matrix WH , W and H should satisfy the condition of Equation (3).

$$D(B, WH) = \sum_{ij} [B_{ij} - (WH)_{ij}]^2 \quad (2)$$

$$\begin{cases} \min D(B, WH), W \geq 0, H \geq 0 \\ \text{sparseness}(w_i) = s_i, i = 1, 2, \dots, z (z < r) \end{cases} \quad (3)$$

where w_i is the i^{th} column of the basis matrix W , and s_i is the expected sparseness of w_i .

The iteration rules of NMFSCPBM are described as follows:

1) Basis matrix W :

$$\begin{cases} W_{ij} = W_{ij} - \lambda \sum_k (WH - B)_{ik} H_{jk} \\ W_{ij} = \begin{cases} W_{ij}, \text{if } W_{ij} \geq 0 \\ 0, \text{else} \end{cases} \\ w_i = L(w_i) \end{cases} \quad (4)$$

where $L(w_i)$ is performing nonlinear projection operation to w_i [12].

2) Coefficient matrix H :

$$H_{kj} \leftarrow H_{kj} \frac{(W^T B)_{kj}}{(W^T WH)_{kj}} \quad (5)$$

Firstly, the definitions of some parameters used in iterative process are given. L is iterations. s is sparseness degree. W_{new} is the sparse constraint matrix to be added. w_k is the k^{th} column of W_{new} . $\text{sparseness}(w_k, s)$ is the adding sparseness constraints with the sparseness degree s for w_k . w_k is the column vector after the sparse constraint. W_{newj} is sparseness constraints matrix. dw is intermediate variable, λ is penalty factor. β is penalty factor threshold. τ' is the convergence error threshold of objective function. T is transpose operation. Therefore, the iterative steps of NMFSCPBM can be described below. The algorithm input is B , r , and L . The output is W and H .

Step 1. Initialize β , set the loop variable $I = 1$, and initialize W and H to random positive matrices.

Step 2. $H = H \otimes \frac{(W^T B)}{(W^T WH)}$, $D(B, WH) = \|B - WH\|^2$.

Step 3. $dw = (WH - B)H^T$, $\lambda = 1/2$, $j = 1$.

Step 4. Add partly sparseness constraints to W and start the iteration.

(a) $W_{new} = W - \lambda \times dw$.

(b) $w'_k = \text{sparseness}(w_k, s), k = 1, \dots, z. (z < r).$

(c) $D(B, W_{newj}H) = \|B - W_{newj}H\|^2.$

(d) If $D(B, W_{newj}H) < D(B, WH)$ or $\lambda < \beta$, then $W = W_{newj}$, turn to Step 5. Otherwise, return to Step 4(a), and set $\lambda = \lambda/2, j = j+1.$

Step 5. If $\|B - WH\|_F^2 < \tau'$ or $i = L$, then exit. Otherwise, return to Step 2, and set $i = i + 1.$

2.2. Complete Basis Matrix Recovery Based on the Least Remaining Sub-Blocks

First of all, the definition of least remaining sub-blocks is given, followed by the **Theorem 1** and its proof, which is the theoretical basis of resisting strong cropping attack of this scheme. Here, the residual video is defined as the remaining of the watermarked video after it is suffered from cropping attacks.

The number of least remaining sub-blocks is defined as the number of the least complete video sub-blocks needed to recover the complete basis matrix from the residual video.

Set the basic NMF model $B_{m \times n} = W_{m \times r} \times H_{r \times n}$. b_i is the i^{th} column of B . h_i is the i^{th} column of H . $B_{m \times n} = [b_1, b_2, \dots, b_n]$ and $H_{r \times n} = [h_1, h_2, \dots, h_n]$ are substituted in the basic model of NMF, then:

$$[b_1, b_2, \dots, b_n] = W_{m \times r} \times [h_1, h_2, \dots, h_n] \tag{6}$$

Outspread Equation (6), then Equation (7) is obtained.

$$b_i = W_{m \times r} \times h_i \quad i \in [1, n] \tag{7}$$

Firstly, perform blocking preprocessing to video, shown as **Figure 1**. The rules of blocking are as follows. Divide original video V into sub-blocks with size of $a \times a \times K$ along the temporal axis, and outspread each sub-block as one column of nonnegative matrix B . Let C_j be the j^{th} sub-block of V , then:

See the Equation (8).

where $\lceil \cdot \rceil$ denotes rounding operation toward 0, $\text{mod}()$ denotes modulus operation, K is the number of video frames, $i = 1, 2, \dots, Ka^2$, and $k = 1, 2, \dots, K$.

Actually, cropping attacks may occur in anywhere of the video. If video sub-blocks at its position suffer from cropping attacks, the corresponding data in the data matrix

directly set to 0, shown as **Figure 2**. Let c denote the number of remaining complete sub-blocks after strong cropping attacks. For convenience, the remaining complete sub-blocks are rearranged from 1, 2, ..., c according to blocking order. Then the data matrix B' , the coefficient matrix H' and the base matrix W corresponding to the remaining sub-blocks are as follows:

$$B' = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1c} \\ b_{21} & b_{22} & \dots & b_{2c} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mc} \end{bmatrix}, \quad H' = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1c} \\ h_{21} & h_{22} & \dots & h_{2c} \\ \vdots & \vdots & & \vdots \\ h_{r1} & h_{r2} & \dots & h_{rc} \end{bmatrix},$$

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1r} \\ w_{21} & w_{22} & \dots & w_{2r} \\ \vdots & \vdots & & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mr} \end{bmatrix}, \text{ where } r \text{ is the dimension}$$

of nonnegative matrix factorization. According to Equation (7), Equation (9) is obtained.

$$B' = W \cdot H' \tag{9}$$

Equation (9) shows that after the video suffers strong cropping attacks, B' is obtained by Equation (8) and H' and complete matrix W can be obtained by the iteration rules of NMFSCPB. Therefore, the following **Theorem 1** is established.

Theorem 1: When the watermarked video suffers from strong cropping attacks, as long as the residual video contains the number of least remaining sub-blocks and meets $c \geq r$, the complete basis matrix W can be recovered from residual video uniquely and correctly.

Theorem 1 is proved as follows. Transpose Equation (9) to obtain Equation (10).

$$H'^T \cdot W^T = B'^T \tag{10}$$

Take B', H' and W into Equation (10), then Equation (11) is obtained.

See the Equation (11).

Let $w_1 = [w_{11} \ w_{12} \ \dots \ w_{1r}]^T, w_2 = [w_{21} \ w_{22} \ \dots \ w_{2r}]^T, \dots, w_m = [w_{m1} \ w_{m2} \ \dots \ w_{mr}]^T, b_1 = [b_{11} \ b_{12} \ \dots \ b_{1c}]^T, b_2 = [b_{21} \ b_{22} \ \dots \ b_{2c}]^T, \dots, b_m = [b_{m1} \ b_{m2} \ \dots \ b_{mc}]^T,$ then Equation (11) is converted into:

$$B(i, j) = \begin{cases} C_j(\text{mod}(i, a), \lceil i/a \rceil - a \times (k-1), k) & \text{if } \text{mod}(i, a) \neq 0 \\ C_j(a, \lceil i/a \rceil - a \times (k-1), k) & \text{if } \text{mod}(i, a) = 0 \end{cases} \tag{8}$$

$$\begin{bmatrix} h_{11} & h_{21} & \dots & h_{r1} \\ h_{12} & h_{22} & \dots & h_{r2} \\ \vdots & \vdots & & \vdots \\ h_{1c} & h_{2c} & \dots & h_{rc} \end{bmatrix} \cdot \begin{bmatrix} w_{11} & w_{21} & \dots & w_{m1} \\ w_{12} & w_{22} & \dots & w_{m2} \\ \vdots & \vdots & & \vdots \\ w_{1r} & w_{2r} & \dots & w_{mr} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{21} & \dots & b_{m1} \\ b_{12} & b_{22} & \dots & b_{m2} \\ \vdots & \vdots & & \vdots \\ b_{1c} & b_{2c} & \dots & b_{mc} \end{bmatrix} \tag{11}$$

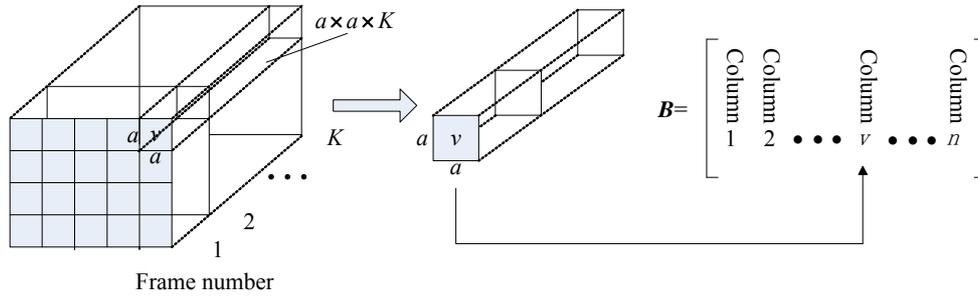


Figure 1. Video blocking of this paper.

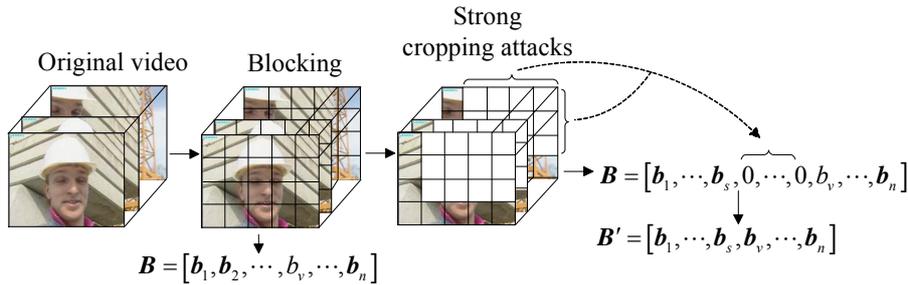


Figure 2. Vide suffer from strong cropping.

$$H^{T^t} \cdot [w_1 \ w_2 \ \dots \ w_m] = [b_1 \ b_2 \ \dots \ b_m] \quad (12)$$

Let $A = H^{T^t}$, then Equation (12) can be converted to the following linear equations, shown as Equation (13), where A is the coefficient matrix of each equations and $W^T = [w_1 \ w_2 \ \dots \ w_m]$ is the unsolved unknown vectors.

$$\begin{cases} Aw_1 = b_1 \\ Aw_2 = b_2 \\ \vdots \\ Aw_m = b_m \end{cases} \quad (13)$$

What follows is: taking the linear equation $Aw_1 = b_1$ as example, the existence and uniqueness of its solution is discussed. By the uniqueness of nonnegative matrix decomposition results, we know that, $Aw_1 = b_1$ is solvable and has a unique solution. In fact, the necessary and sufficient conditions of r -dimensional homogeneous linear equations are $R(A) = R(\bar{A}) = r$, where $\bar{A} = [A \ b_1] = [H^{T^t} \ b_1]$ is the augmented matrix of A , $R(A)$ and $R(\bar{A})$ are the rank of A and \bar{A} respectively, $R(A) \leq \min(r, c)$, and $R(\bar{A}) \leq \min(r+1, c)$. It can be divided into three cases. 1) If $c < r$ and $R(A) = R(\bar{A}) \leq c < r$, $Aw_1 = b_1$ has infinite solutions; 2) If $c \geq r$ and $R(A) = R(\bar{A}) < r$, $Aw_1 = b_1$ has infinite solutions. The above two cases violate to uniqueness of the nonnegative matrix factorizing results, the solutions are rejected. 3) If $c \geq r$ and $R(A) = R(\bar{A}) = r$, $Aw_1 = b_1$ has the unique solution. Similarly, other equations have the same conclusions.

We can see from the above solution process, if $c \geq r$

and $R(A) = R(\bar{A}) = r$, w_1, w_2, \dots, w_m separately has unique solution, or the base matrix W has a unique solution. In other words, when the watermarked video suffers from strong cropping attacks, as long as the residual video contains complete video sub-blocks and the number meets $c \geq r$, the complete basis matrix W can be recovered from residual video uniquely and correctly. Namely, **Theorem 1** is established. Experimental results of this paper verify the correctness of theoretical analysis.

2.3. Video Motion Components Extraction Based on NMFSCPBM

The basis matrix and coefficient matrix can be obtained by nonnegative matrix factorizing. Basis matrix represents the major features of video. Coefficient matrix is the linear projection of nonnegative matrix to basis matrix and represents the local feature weights of video. Video frames can be seen as linear weighted sum of the static components and motion components. In general, the static components are nonsparse, while the motion components are sparse. So motion component can be separated from the static background by controlling sparseness constrains of the basis matrix, and then extract the motion components. The motion component extraction processes based on NMFSCPBM includes:

1) Video pre-processing. Take the target frame of the movement components to be extracted as the center for the original video $V(m_x \times m_y \times K)$ and choose the forward and the next l frames separately. Let the total $2l+1$ frames compose of a video frame group V' . Outspread V' one-dimensionally as one column of non-

negative matrix B , shown as Equation (14), where the size of video frame is $m_x \times m_y$, $i=1, 2, \dots, m_x m_y$, $j=1, 2, \dots, 2l+1$. The parameter l needs a reasonable choice. If l is chosen too large, calculation is risen significantly. If l is chosen too small, there have no obvious motion information between video frames.

$$B(i, j) = V'(\lceil i/m_x \rceil, \text{mod}(i, m_x), j) \quad (14)$$

2) NMFSCPBM factorization. Perform NMFSCPBM to matrix B , and set r as the factorization dimension. Add sparseness constraints to the $(r-1)$ basis vectors, so the basis vectors constrained sparseness $w_i (i=1, 2, \dots, r-1)$ represents the motion components of video [12].

3) Solving Video motion components. Motion components M of target frame can be obtained by weighting and summing to the $(r-1)$ motion components, shown as Equation (15).

$$M = \sum_{i=1}^{r-1} w_i H_{i,l+1} \quad (15)$$

where, $H_{i,l+1}$ is the weighting coefficient of basis vectors w_i corresponding to target frame. For each pixels of the target frame, there is one element in M which is corresponding to it, M has the same size with target frame, the bigger of the element in M , the more intensity of the corresponding pixels moving in target frame.

Motion components extracted by the NMFSCPBM

and other similar methods are shown in **Figure 3**. It can be seen that, the motion components in **Figure 3(b)** extracted by the proposed method filter out the static background interference completely, and show the trajectory clearly compared with **Figure 3(a)**. **Figure 3(c)** cannot distinguish the moving target from static background and cannot show the trajectory. **Figure 3(d)** has improved a little compared with **Figure 3(c)**, but still cannot separate the moving target from static background completely. To quantitatively assessment the effectiveness of the motion components extraction of this paper, the matching rate [13] τ is defined to evaluate as follows:

$$\tau = \frac{\sum_{x,y \in R} I(x,y) \times \text{SMD}(x,y)}{\sqrt{\sum_{x,y \in R} I(x,y) \times I(x,y)} \cdot \sqrt{\sum_{x,y \in R} \text{SMD}(x,y) \times \text{SMD}(x,y)}} \quad (16)$$

where $\text{SMD}(x,y)$ is the motion components extracted, $I(x,y)$ is the manually specified target motion area, and R is the target frame. The more the τ is close to 1, the more matching of the extracted motion features compared with the features specified motion area. When the extracted motion features are exactly the same with the features specified motion area, $\tau = 1$.

The experiment chooses “hall”, “stefan” and “tennis” as test videos. (<http://trace.eas.asu.edu/yuv/index.html>).

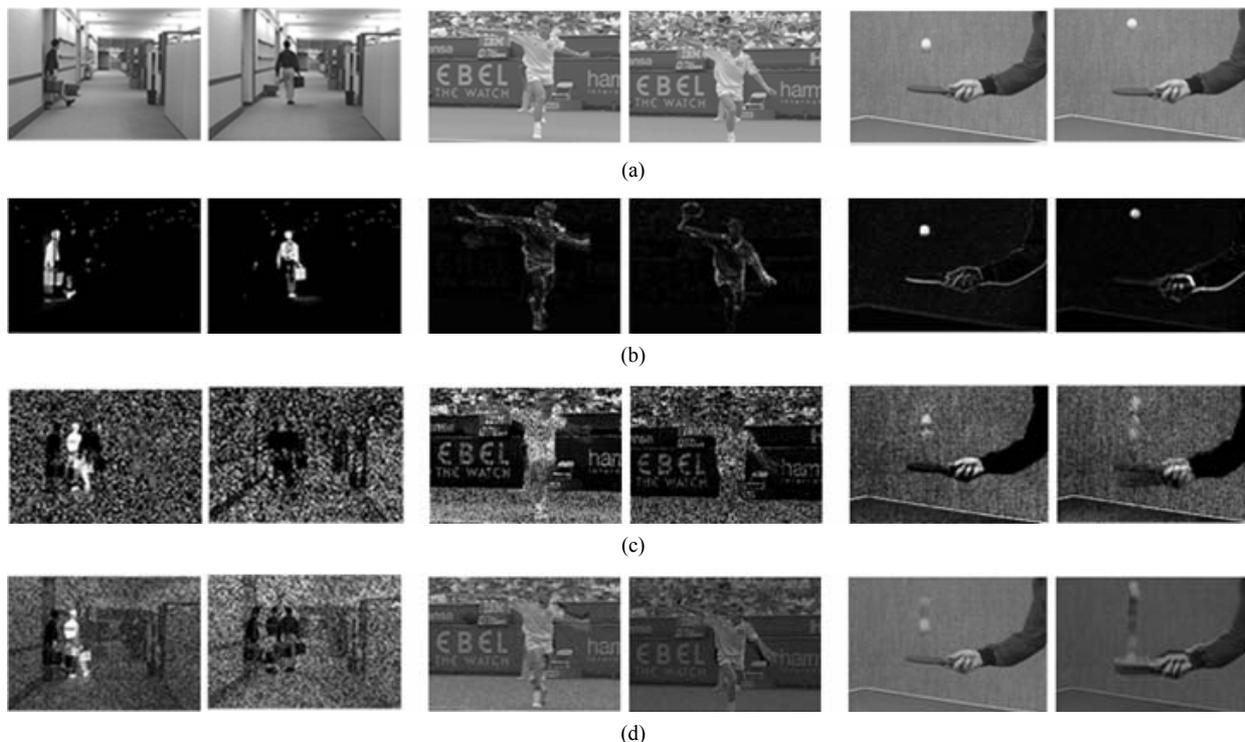


Figure 3. Motion components extraction results obtained by NMFSCPBM method and other similar methods. (a) Original video; (b) Motion components extraction results obtained by NMFSCPBM; (c) Motion components extraction results obtained by NMFSC; (d) Motion components extraction results obtained by NMF.

Compute the matching rate of the 25th fame and the 55th frame in “hall”, the 26th frame and the 30th frame in “stefan”, and the 16th frame and the 18th frame in “tennis”, respectively. The experiment sets $l = 5$, and experimental results are shown in **Figure 4**. It can be seen that the matching rates of this paper are more close to 1 compared with other methods. The average matching rate of this paper is 0.8907, NMF is 0.5371, and NMFSC is 0.4070.

3. Watermark Embedding and Extraction

Watermark embedding. The general watermark methods based on the NMF attempt to change the coefficient matrix by various signal processing operations to accomplish watermark embedding [6], but the robustness is subjected to comparative limitation. As a linear expression method, the significant advantage of NMF lies in that the basis matrix is changeable [10] and that is robust to cropping attack [9]. Therefore, this scheme embeds the watermark into the basis matrix that is decomposed into by NMFSCPBm. When the watermarked video suffers from strong cropping attacks, known by Chapter 2.2 of this paper, as long as the residual video contains the numbers of least remaining sub-blocks, the complete basis matrix W' can be completely recovered and then

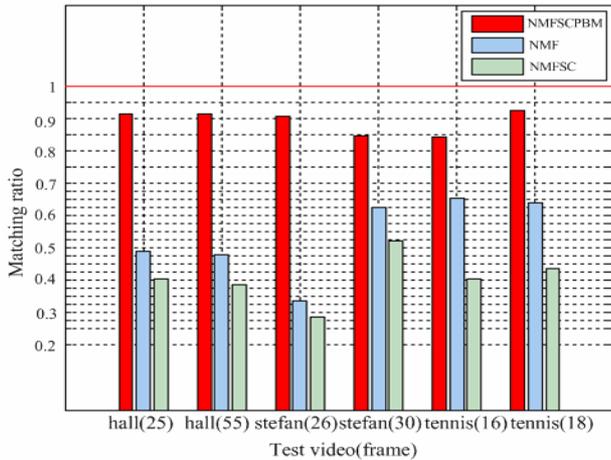


Figure 4. Comparative analysis of matching rate of the motion components obtained by NMFSCPBm method and other similar methods.

the complete watermark can be extracted. The watermark embedding schematic diagram of the proposed scheme is shown in **Figure 5**.

1) Video pre-processing. Perform blocking preprocessing to the original video $V(m_x \times m_y \times K)$. According to the blocking rules Chapter 2.2 of this paper, obtain the nonnegative matrix B .

2) Performing the NMFSCPBm decomposition to B . Calculate the decomposition error according to Equation (17), and retain W for watermark extraction.

$$E = B - WH \tag{17}$$

3) Watermark encryption. Select pseudo-random sequence U as the secret key, perform encryption to watermark signal S , and gain the secret information P , shown as Equation (18).

$$P(k) = \sum_{t=1}^l s_t(k) u_t(k), k = 1, 2, \dots, N \tag{18}$$

4) Watermark embedding. The proposed scheme chooses the N biggest elements $w_1, w_2 \dots w_N$ of the basis matrix W as the positions for watermark embedding, embeds watermark by Equation (19) multiplicative rules, and gains the watermarked basis matrix W' .

$$w_n' = w_n \cdot (1 + I \cdot p_n) \tag{19}$$

5) Watermark embedding strength adaptively adjustment. Perform the blocking processing to video according to Chapter 2.2 of this paper. Extract the video motion components M according to Chapter 2.3 of this paper. Calculate the motion feature coefficient $F(w_i)$ of the row where the i^{th} coefficients locate according to Equation (20). Therefore, the watermark embedding strength of the i^{th} coefficient is $I(w_i) = \alpha \cdot F(w_i) i = 1, 2, \dots, N$, where α is the motion masking weighting factor and $\alpha = 0.013$ is selected by the experiments.

$$F(w_i) = \frac{a^2}{m_x m_y} \sum_j M(i, j) \tag{20}$$

6) Performing nonnegative matrix synthesis by NMFSCPBm. Perform nonnegative matrix synthesis to W', H, E as Equation (21).

$$B' = W'H + E \tag{21}$$

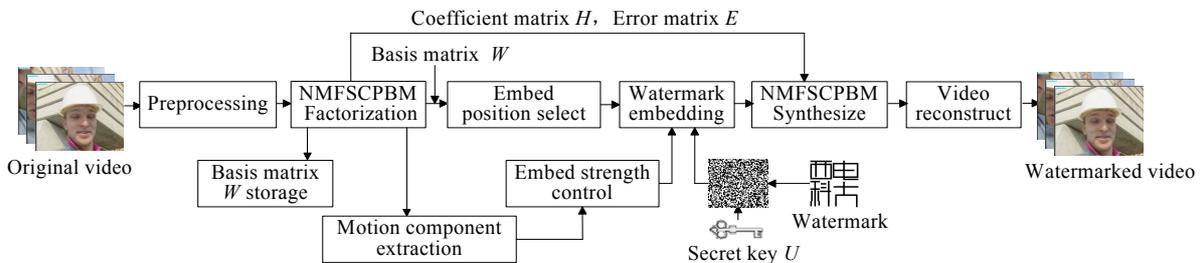


Figure 5. Watermark embedding scheme of this paper.

7) Video reconstruction. Reconstruct B' according to blocking rules, and then output the watermarked video finally.

Watermark extraction. The watermark extraction schematic diagram of the proposed scheme is shown in **Figure 6**. Firstly, perform the blocking preprocessing for the watermarked video that has been attacked according to Chapter 2.2 of this paper, obtaining the nonnegative matrix B'' constituted by the remaining complete sub-blocks in residual video. Secondly, perform the NMFSCPBM decomposition to B'' , getting the watermarked complete basis matrix W' in residual video. Finally, by comparing W with W' , the encrypted watermark is extracted as Equation (22), and then it is decrypted by the secret key U . From the process of watermark extraction, we know that the watermark extraction of this scheme does not need original video. So it is a blind watermark scheme.

$$p = \begin{cases} 0 & \text{if } w \geq w' \\ 1 & \text{if } w < w' \end{cases} \quad (22)$$

4. Experiment and Analysis

In order to verify the validity of this paper, the videos “football”, “mother-daughter”, “tempeste”, “mobile”, “aki-yo”, “hall”, “foreman” and “soccer” with the format of CIF and the length of 300, 260, 260, 300, 300, 300, 300 and 300 frames are selected as the host video for experiment respectively. A 64×64 binary image is selected as the watermark, the size of sub-block is 8×8 , the decomposition dimension r of NMFSCPBM is 32, and the experimental software environment is matlab 7.2. The experiment conducts the test and analysis about the transparency, bit rate constancy, robustness, real-time property, and the algorithm efficiency etc. for this scheme respectively. Due to the length limitation of this paper, only a partial list of results are listed such as the transparency, bit rate constancy, and strong cropping attacks etc.

4.1. Transparency Tests

Figures 7(a) and **(b)** are the I-frame video captures of the original frame and watermarked frame from football sequence. It can be seen that there is no significant visual perceptible distortion before and after the watermark is embedded. From the quantization assessment data shown in **Table 1**, we also know that the difference value Δ PSNR

of video coding before and after the watermark is embedded decreases only by average 0.025dB (PSNR, Peak Signal to Noise Ratio of video coding) and that there is no influence on visual perception. The reason is that the proposed scheme sufficiently considers the visual masking features for watermark embedding position selection and strength control, embeds the watermark into the big coefficients of the basis matrix, and meanwhile embeds comparatively strong strength watermark into the regions where the motion is intense [14], eliminating the flicker influence and guaranteeing the transparency of the method.

4.2. Bit Rate Constancy Tests

Table 2 shows the test results of the bit rate constancy of this paper. The experiment is assessed by the bit increased rate (BIR) before and after the watermark is embedded, as Equation (23).

$$BIR = \frac{\text{watermarked_BR} - \text{original_BR}}{\text{original_BR}} \times 100\% \quad (23)$$

where *original_BR* is the bit rate before the watermark is embedded, and *watermarked_BR* is the bit rate after the watermark is embedded. It can be seen that the increase in video BIR after the watermark is embedded in this paper is under 0.15%, with a good bit rate constancy.

4.3. Strong Cropping Attacks Experiments

Strong cropping attacks include various regular and irregular cropping, such as row and column cropping, edge cropping, top right corner cropping and center cropping etc. The same type and intensity of cropping attacks are synchronously carried out to paper [3], and the comparative analysis is given. The experimental results are assessed by the bit correct rate (BCR) of the extracted watermark, as Equation (24).

$$BCR = \frac{e}{m} \times 100\% \quad (24)$$

where e is the number of correct bits of the extracted watermark, and m is the number of total bits of the extracted watermark. The more the BCR is close to 100%, the more higher of the correct rate. Let the threshold $T = 70.00\%$ that is determined by experiments. If $BCR > T$, the watermark is detected. Some figures and

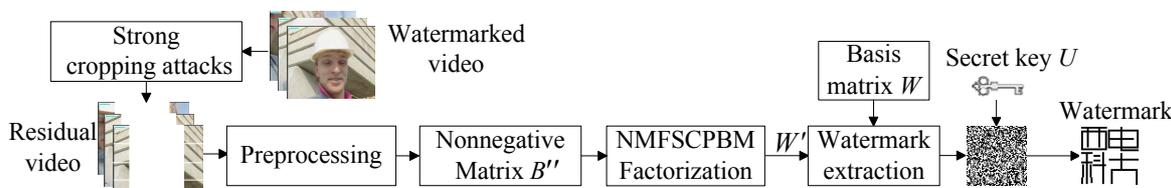


Figure 6. Watermark extraction scheme of this paper.



Figure 7. Strong cropping attacks test results of this paper.

Table 1. Transparency test results of this paper.

Test video	PSNR (dB)		Δ PSNR (dB)
	Without watermark	With watermark	
football	33.25	33.24	0.01
mother-daughter	37.90	37.87	0.03
tempete	32.92	32.88	0.04
mobile	32.22	32.17	0.05
akiyo	38.45	38.44	0.01
hall	36.31	36.28	0.03
foreman	35.15	35.14	0.01
soccer	34.31	34.29	0.02

Table 2. Bit rate constancy tests results of this paper.

Test video	Bit rate (kbps)		BIR
	watermarked_BR	original_BR	
football	1667.98	1668.89	0.0546%
mother-daughter	222.22	222.38	0.0720%
tempete	1101.06	1101.61	0.0500%
mobile	1651.80	1651.83	0.0018%
akiyo	226.49	226.52	0.0130%
hall	415.98	416.05	0.0168%
foreman	489.58	490.33	0.1500%
soccer	654.55	655.44	0.1400%

quantification assessment results of the robustness experiment are shown in **Figure 7** and **Table 3**.

It can be seen from **Figure 7** and **Table 3** that, 1) the BCRs of the proposed scheme are all 100% for various regular and irregular strong cropping attacks. The scheme can recover the complete watermark without any damage, and has a very strong ability to resist strong cropping attacks. Analysis of the main reasons is that, based on the strong robustness of basis matrix to cropping attacks, this scheme embeds the encrypted watermark into the big coefficients of the basis matrix that the host video is decomposed into by NMFSCPBM. When the watermarked video suffers from strong cropping attacks, the complete basis matrix can be completely recovered through decomposition of the residual video by NMFSCPBM, and then the complete watermark can be extracted. The selection of the big coefficients of the basis matrix and the adaptive control of the watermark embedding strength based on the video motion feature coefficients, further increase the watermark embedding strength and improve the robustness. The spread spectrum encryption processing for watermark also increases the robustness while increases the watermark invisibility. 2) About various regular and irregular cropping attacks listed in the experiment, with the same attack intensity compared with this paper, the BCRs of the paper [3] are completely not up to 70% of the threshold requirements. Take the example of the “football” test video, when the

row regular cropping is 97.22%, the BCR of the paper [3] is 51.25%. When the edge cropping is 97.73%, the BCR of the paper [3] is 48.82%. For various irregular cropping attack, as **Figures 7(j)-(o)** shown, the BCRs of the paper [3] are between 47.07% - 56.35%, which can not meet the threshold 70%. The same conclusion is obtained from the mother-daughter, tempete, mobile, akiyo, hall, foreman and soccer test videos.

Figure 8 shows the relation curve of BCRs of the proposed scheme with the cropping attack intensity. It can be seen that the inflexion where the BCR is less than 100% appears on the point with the cropping intensity 97.98%, while the cropping intensity is maximum and the number of the remaining complete sub-blocks in the residual video is $c = 32$, neither more nor less than equaling to the dimension of NMF $r = 32$. Further analysis shows that, with the further increase of the cropping intensity, the BCR decreases rapidly. It is mainly because the number of the remaining complete sub-blocks in the residual video c is less than 32, which indeed does not meet the least number of the remaining complete sub-blocks required in the **Theorem 1** of this paper. This fact makes e that is the number of correct bits of the extracted watermark decrease rapidly and causes the BCR decrease rapidly. The above analysis and results are fully consistent with the results of the theoretical analysis in Chapter 2.2 of this paper.

Table 3. Robustness comparison of this scheme and paper [3] for strong cropping attacks.

Attacks	BCR (%)															
	football		mother-daughter		tempete		mobile		akiyo		hall		foreman		soccer	
	Our	Paper [3]	Our	Paper [3]	Our	Paper [3]	Our	Paper [3]	Our	Paper [3]	Our	Paper [3]	Our	Paper [3]	Our	Paper [3]
No attacks	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Row cropping 94.00%	100	56.30	100	57.38	100	59.07	100	55.48	100	52.86	100	56.21	100	54.40	100	55.03
Row cropping 97.22%	100	51.25	100	51.36	100	52.82	100	51.21	100	45.65	100	51.29	100	53.15	100	54.25
Column cropping 95.45%	100	59.76	100	57.23	100	57.15	100	56.32	100	52.80	100	51.25	100	50.53	100	53.29
Column cropping 97.73%	100	52.05	100	55.15	100	47.41	100	51.26	100	48.99	100	47.26	100	53.34	100	51.77
Row and column cropping 97.98%	100	52.55	100	52.58	100	49.09	100	52.10	100	51.67	100	46.54	100	48.55	100	49.85
Edge cropping 97.73%	100	48.82	100	50.06	100	50.15	100	50.23	100	52.27	100	49.36	100	47.73	100	47.76
Top corner cropping 95.01%	100	50.44	100	49.56	100	52.71	100	51.85	100	53.59	100	50.45	100	52.01	100	50.31
Bottom corner cropping 95.01%	100	47.85	100	49.98	100	51.64	100	50.36	100	52.52	100	43.56	100	53.19	100	52.24
Centre cropping 93.56%	100	53.55	100	50.25	100	50.39	100	50.85	100	49.86	100	51.19	100	54.05	100	49.12
Irregular cropping 94.51%	100	52.80	100	52.01	100	51.26	100	51.36	100	50.71	100	48.57	100	52.16	100	48.58
Irregular cropping 95.04%	100	51.89	100	52.54	100	54.55	100	49.95	100	51.45	100	51.26	100	53.15	100	51.36
Irregular cropping 95.50%	100	55.39	100	54.60	100	56.25	100	50.26	100	55.32	100	51.35	100	51.14	100	52.25
Irregular cropping 95.53%	100	50.85	100	50.76	100	52.35	100	51.25	100	48.07	100	49.68	100	51.65	100	54.33
Irregular cropping 96.17%	100	52.92	100	56.35	100	54.35	100	55.35	100	54.17	100	48.69	100	47.65	100	47.07
Irregular cropping 96.56%	100	51.31	100	52.16	100	53.63	100	52.86	100	50.95	100	50.25	100	51.23	100	48.25

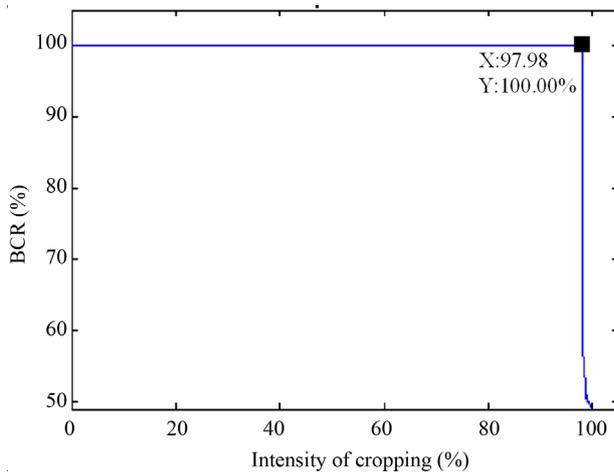


Figure 8. Relation curve of BCRs with the cropping attack intensity in the proposed scheme.

4.4. Computational Efficiency and Decomposing Errors Experiments of NMFSCPBM Methods

Figure 9 shows the convergence curves of decomposition error with the number of iterations of NMFSCPBM, NMF and NMFSC for different data sets. The sparseness constraint of NMFSC and NMFSCPBM is 0.6 respectively and the decomposition dimension is $r = 4$ respectively. Due to the length limitation of an article, **Figure 9** only shows the fore 1000 iterative results of the “football” test video. It can be seen that, 1) the convergence error of this paper is 0.468, 19.379 for the NMFSC and 1.809 for the NMF; 2) the convergence error comes up to the minimum after 77 iterations for NMFSCPBM, 913 for NMFSC, and 410 for NMF. Therefore, the decomposing error of this paper is the lowest in the similar methods, and the convergence rate is also obviously better than that of the similar methods. Other experimental data of the test videos give the same conclusion.

5. Conclusion

A novel video watermarking scheme robust to strong cropping is proposed in this paper. It is characterized as follows. 1) The improved NMFSCPBM method is proposed, which can extract the video motion features accurately, filter out the static background interference completely, and is simple and effective. Solve the problems of the similar methods that cannot describe the data sufficiently under higher sparseness constraints. Meanwhile, reduce the decomposition error and speed up convergence rate. 2) Based on the robustness of basis matrix for shearing attacks, this framework innovatively embeds the encrypted watermark into the big coefficients of the basis matrix that the host video is decomposed into by NMFSCPBM. To achieve the greatest strength of the

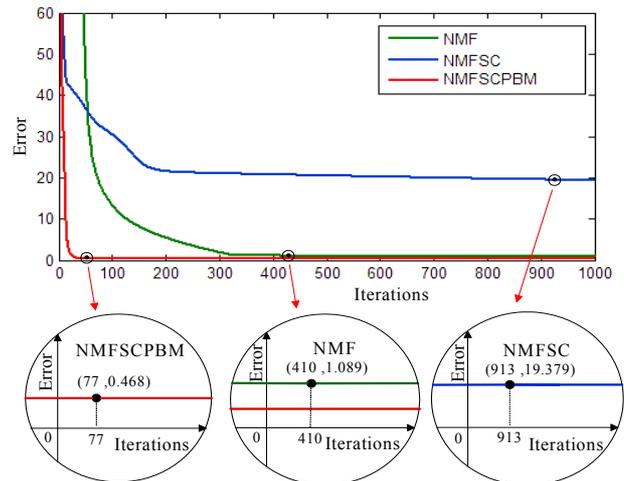


Figure 9. Decomposing errors and computational efficiencies of NMFSCPBM and other similar methods.

watermark embedding with no visual perception, this scheme adaptively adjusts the watermark embedding strength by the video motion feature coefficients extracted by NMFSCPBM method, improving the robustness further. On watermark detection, as long as the residual video contains the numbers of least remaining sub-blocks, the complete basis matrix can be completely recovered through decomposition of the nonnegative matrix of the least remaining sub-blocks in residual videos by NMFSCPBM, and then the complete watermark can be extracted. The experimental results show that the performance of resisting strong cropping attacks of this scheme is improved greatly compared to existing methods, and that the scheme has good transparency, bit rate constancy and real-time property. It is a blind watermark scheme. How to extract and recover the complete watermark from the incomplete sub-blocks of the residual video will be as the further research content for authors.

REFERENCES

- [1] H. Zhang, *et al.*, “Affine Legendre Moment Invariants for Image Watermarking Robust to Geometric Distortions,” *IEEE Transactions on Image Processing*, Vol. 20, No. 8, 2010, pp. 2189-2199. [doi:10.1109/TIP.2011.2118216](https://doi.org/10.1109/TIP.2011.2118216)
- [2] C. V. Serdean, M. A. Ambroze, M. Tomlinso and J. G. Wade, “DWT-Based High-Capacity Blind Video Watermarking, Invariant to Geometrical Attacks,” *IEEE Proceedings of the Vision, Image and Signal Processing*, Vol. 150, No. 1, 2003, pp. 51-58. [doi:10.1049/ip-vis:20030159](https://doi.org/10.1049/ip-vis:20030159)
- [3] L. E. Coria, M. R. Pickering, P. Nasiopoulos and R. K. Ward, “A Video Watermarking Scheme Based on the Dual-Tree Complex Wavelet Transform,” *IEEE Transactions on Information Forensics and Security*, Vol. 3, No. 3, 2008, pp. 466-474. [doi:10.1109/TIFS.2008.927421](https://doi.org/10.1109/TIFS.2008.927421)
- [4] Y. L. Wang and A. Pearmain, “Blind MPEG-2 Video

- Watermarking Robust against Geometric Attacks: A Set of Approaches in DCT Domain,” *IEEE Transactions on Image Processing*, Vol. 15, No. 6, 2006, pp. 1536-1543. [doi:10.1109/TIP.2006.873476](https://doi.org/10.1109/TIP.2006.873476)
- [5] X. M. Niu, *et al.*, “A Video Watermarking against Geometrical Distortions,” *Chinese Journal of Electronics*, Vol. 12, No. 4, 2003, pp. 548-552.
- [6] M. Silja and K. Soman, “A Watermarking Algorithm based on Contourlet Transform and Nonnegative Matrix Factorization,” *Proceedings of the International Conference on Advances in Recent Technologies in Communication and Computing*, Kottayam, 27-28 October 2009, pp. 279-281. [doi:10.1109/ARTCom.2009.198](https://doi.org/10.1109/ARTCom.2009.198)
- [7] A. D’Angelo, Z. P. Li and M. Barni, “A Full-Reference Quality Metric for Geometrically Distorted Images,” *IEEE Transactions on Image Processing*, Vol. 19, No. 4, 2010, pp. 867-881. [doi:10.1109/TIP.2009.2035869](https://doi.org/10.1109/TIP.2009.2035869)
- [8] Z. Y. Yang, *et al.*, “Blind Spectral Unmixing Based on Sparse Nonnegative Matrix Factorization,” *IEEE Transactions on Image Processing*, Vol. 20, No. 4, 2011, pp. 1112-1125. [doi:10.1109/TIP.2010.2081678](https://doi.org/10.1109/TIP.2010.2081678)
- [9] M. Tong, T. Yan and H.-B. Ji, “Strong Anti-Robust WaterMarking Algorithm,” *Journal of Xidian University*, Vol. 36, No. 1, 2009, pp. 22-27.
- [10] D. Gai, X. F. He, J. W. Han and T. S. Huang, “Graph Regularized Nonnegative Matrix Factorization for Data Representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 1, 2011, pp. 1-13. [doi:10.1109/TPAMI.2010.231](https://doi.org/10.1109/TPAMI.2010.231)
- [11] T. Gao and M.-Y. He, “Using Improved Nonnegative Matrix Factorization with Projected Gradient for Single-Trial Feature Extraction,” *Journal of Electronics & Information Technology*, Vol. 32, No. 5, 2010, pp. 1121-1125.
- [12] P. O. Hoyer, “Nonnegative Matrix Factorization with Sparseness Constraints,” *Journal of Machine Learning Research*, Vol. 5, 2004, pp. 1457-1469.
- [13] T. K. Kim, *et al.*, “Video Object Segmentation and Its Salient Motion Detection Using Adaptive Background Generation,” *Electronics Letters*, Vol. 45, No. 11, 2009, p. 542.
- [14] S.-W. Kim, K. R. Rao, S. Suthaharan and H.-K. Lee, “Perceptually Tuned Robust Watermarking Scheme for Digital Video Using Motion Entropy Masking,” *Proceedings of the International Conference on Consumer Electronics of the IEEE ICCE*, Los Angeles, 22-24 June 1999, pp. 104-105. [doi:10.1109/ICCE.1999.785187](https://doi.org/10.1109/ICCE.1999.785187)