IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 4, 2019, accepted August 13, 2019, date of publication August 19, 2019, date of current version August 31, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2936134

A Hybrid Transforms-Based Robust Video Zero-Watermarking Algorithm for Resisting High Efficiency Video Coding Compression

XIAOYAN YU, CHENGYOU WANG, (Member, IEEE), AND XIAO ZHOU
School of Mechanical, Electrical, and Information Engineering, Shandong University, Weihai 264209, China

Corresponding author: Chengyou Wang (wangchengyou@sdu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61702303, and in part by the Shandong Provincial Natural Science Foundation, China, under Grant ZR2017MF020.

ABSTRACT With the rampancy of pirated videos, video watermarking for copyright protection has become a widely researched topic. In this paper, zero-watermarking is applied to videos for the first time to resist high efficiency video coding compression, which can improve the robustness of the watermarking algorithm and ensure the videos' quality. A robust video zero-watermarking algorithm based on the discrete wavelet transform, the all phase biorthogonal transform, and singular value decomposition is proposed. Utilizing the properties of hybrid transforms, robust features can be extracted from videos, and robust zero-watermarks can be constructed. Experimental results demonstrate that the proposed algorithm has strong robustness to high efficiency video coding compression attacks with different quantization parameters. In addition, the algorithm can also resist common image processing attacks, geometric attacks, frame-based attacks, and hybrid attacks. Compared with existing video watermarking algorithms, the proposed algorithm can more accurately and completely reconstruct watermark images.

INDEX TERMS Video signal processing, watermarking, video watermarking, zero-watermarking, copyright protection, robustness, high efficiency video coding, discrete wavelet transform, all phase biorthogonal transform, singular value decomposition.

I. INTRODUCTION

With the rapid popularization of Internet technology and the rise of various short video applications, live broadcast, and video-on-demand (VOD) platforms, the number of online videos has increased dramatically. Videos have penetrated all aspects of people's lives. They enrich people's lives and improve their living standards, but they also pose serious threats to information security, especially the security of video data. In recent years, various pirated videos that seriously infringe upon the legitimate rights and interests of the video owners have flooded the Internet. Copyright protection for digital videos is an issue that must be resolved urgently.

Techniques for copyright protection include cryptography [1], steganography [2], digital fingerprinting [3], and digital watermarking [4]–[27]. Among them, digital watermarking is the most commonly used copyright protection technology for digital products due to its good concealment, robustness, and stability. Depending on the differences of the embedding domains, digital watermarking can be divided into two categories: watermarking in the spatial domain [4] and watermarking in the transform domain [5]–[9]. Watermarking algorithms in the spatial domain are easy to implement but have poor robustness. Watermarking algorithms in the transform domain are robust and have wide application ranges. However, both kinds of algorithms achieve watermark insertion by modifying the original carrier data, which will affect the visual quality of the carrier. As more watermark bits are embedded, the robustness will increase, but the invisibility will be diminished. Therefore, the balance between the robustness and invisibility must be considered when embedding watermarks. To address this tension, zero-watermarking

The associate editor coordinating the review of this article and approving it for publication was Gerard-Andre Capolino.

that does not modify the original carriers was developed. Since each carrier generates its own unique zero-watermark, an intellectual property library must be established to store the zero-watermarks. To date, zero-watermarking technology has been applied to copyright protection in many fields, such as text [10], audio [11], image [12]–[19], video [20]–[24], relational databases [25], and sharing models in cooperative drive engineering [26].

In this paper, zero-watermarking is adopted to resist high efficiency video coding (HEVC) compression for the first time, and a hybrid transforms-based video zero-watermarking algorithm is proposed. The selection of the hybrid transforms is inspired by [27], which is a conventional image watermarking algorithm and proposes a hybrid transforms-based watermarking algorithm with good robustness. The features are extracted according to the parity of the highest bit of the largest singular value in the singular value matrix of the video sequence after the hybrid transforms. The zero-watermark is constructed via XOR operations between the encrypted watermark and the extracted features. The main innovations of this paper are as follows. (1) Zero-watermarking is introduced into videos to resist HEVC compression for the first time. (2) The multiresolution characteristic of the discrete wavelet transform (DWT), the better low-frequency energy gathering characteristic of the all phase biorthogonal transform (APBT), and the stability of singular value decomposition (SVD) are combined to extract robust features from video sequences. (3) Through the watermark postprocessing process, watermark images can be more accurately and completely reconstructed, and the robustness of the algorithm is further improved.

The rest of this paper is organized as follows. Related work is provided in Section II. Section III introduces the three transforms: DWT, APBT, and SVD. The key-frame detection algorithm is also described. Section IV details the proposed algorithm, including the generation and the detection of the zero-watermark. The experimental results and discussion are presented in Section V. Finally, the conclusions and future work are discussed in Section VI.

II. RELATED WORK

Due to the poor robustness of watermarking algorithms in the spatial domain, video watermarking algorithms for copyright protection are mostly concentrated in the transform domain [6]–[9]. Singh [6] embedded watermarks into high-frequency DWT coefficients and used the scale invariant feature transform (SIFT) to improve the method's robustness against rotation attacks. Sathya and Ramakrishnan [7] used the Fibonacci-Lucas transform to scramble watermark images and embedded their singular values into the singular values of video frames' middle-frequency DWT sub-bands. The common disadvantage of these two algorithms is that their robustness to compression attacks is poor. Himeur and Boukabou [8] used the gradient magnitude similarity deviation (GMSD) algorithm to extract key frames and embedded the watermark into the singular values of the low-frequency DWT sub-bands of the key frames. This algorithm is less robust to frame-based attacks, such as frame deletion and frame switching. Gupta *et al.* [9] employed the group search optimization (GSO) method to embed watermarks into low- and middle-frequency DWT coefficients. The main problem with their algorithm is that the embedded watermark has a relatively large impact on the video quality, resulting in low imperceptibility.

With the development of high-definition televisions and movies, the latest video coding standard, H.265/HEVC [28], has been gradually popularized. In the transmission and storage processes, videos will inevitably undergo HEVC compression. Therefore, video watermarking must have high resistance to HEVC compression. There have been some conventional video watermarking algorithms that can resist HEVC compression to a certain extent [29]–[31]. However, when the intensity of HEVC compression is high, the watermark images reconstructed by them will have more obvious false detection bits. In addition, these algorithms embed watermarks by modifying videos, which will decrease the videos' quality. With the development of three-dimensional videos (3DVs), some scholars have begun to study copyright protection for 3DVs [32], [33]. El-Shafai *et al.* [32] proposed two hybrid watermarking schemes to secure the copyright of 3DVs. One is the homomorphic transform-based SVD in the DWT domain, and the other is the three-level discrete stationary wavelet transform in the discrete cosine transform (DCT) domain. On this basis, the authors combined a wavelet-based fusion technique to improve the capacity and robustness without affecting the perceptual quality of the original 3D-HEVC frames [33]. At present, most video watermarking algorithms are still based on 2D videos.

The appearance of zero-watermarking addresses the tension between imperceptibility and robustness. At present, the studies on zero-watermarking are mostly concentrated in the image field. Ghadi *et al.* [12] used the Jacobian matrix model and Jumana *et al.* [13] used the genetic algorithm (GA) to construct zero-watermarks. To improve the security, many algorithms introduced visual cryptography (VC) to zero-watermarking technology [14]–[16]. To improve the robustness to geometric attacks, Xia *et al.* [17] and Wang *et al.* [18] adopted the polar harmonic transform (PHT) and Tsai *et al.* [19] adopted log-polar mapping (LPM) to construct zero-watermarks. Many image zero-watermarking algorithms can be extended into video zero-watermarking algorithms after being modified.

Compared with the image field, the research on zero-watermarking for video copyright protection is still in its infancy, and the number of existing studies is relatively small. Its overall framework is shown in Fig. 1. Liu [20] proposed a non-negative matrix factorization with sparseness constraints on parts (NMFSCP) method to generate zero-watermarks. This algorithm has strong resistance to geometric attacks, such as rotation and cropping, but it has weak resistance to common image processing attacks, such as noise and filtering. To improve the robustness to

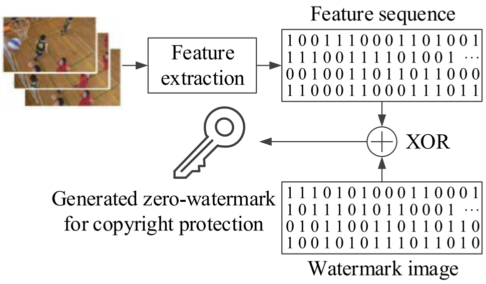


FIGURE 1. Schematic diagram of zero-watermarking algorithms for video copyright protection.

geometric attacks, Li *et al.* [21] adopted the 2D DWT and 3D DCT to obtain stable low- frequency coefficients, and then they conducted the log-polar transform (LPT) to construct zero-watermarks. This algorithm has weak resistance to cropping attacks. Based on the contour wave transform, the pseudo-3D DCT, and SVD, Li *et al.* [22] proposed a zero-watermarking algorithm for copyright protection of animation videos. This algorithm is less robust to compression attacks and cut attacks. Liu *et al.* [23] applied fingerprinting technology to a zero-watermarking algorithm and proposed a hybrid scheme. Its robustness to frame-based attacks is weak. Liu *et al.* [24] applied zero-watermarking for digital rights management (DRM) of 3DVs. The resistance to rotation and cropping attacks of this algorithm is relatively weak. Although the existing video zero-watermarking algorithms can resist some common image processing attacks and a certain degree of geometric attacks, their robustness still needs to be improved. In addition, all of these algorithms do not consider the resistance to HEVC compression and combinations of multiple attacks.

To improve the robustness against HEVC compression attacks and various common attacks and avoid impacts on the videos' quality, a video zero-watermarking algorithm based on hybrid transforms is proposed in this paper. The DWT is first used to concentrate the energy of video frames into the LL sub-band, which can resist attacks such as cropping and sticking to some extent. Then, using the good low-frequency energy accumulation characteristics of the APBT, the direct current (DC) coefficient matrix with more concentrated energy can be obtained. Combined with the stability of the singular values obtained via SVD, robust features can be extracted and zero-watermarks can be generated. The combination of these three transforms ensures the robustness of the proposed algorithm. HEVC compression attacks with different quantization parameters (QPs), common image processing attacks, geometric attacks, frame-based attacks, and hybrid attacks are all taken into account. The experimental results show that compared with existing video watermarking algorithms, the proposed algorithm can more accurately and completely reconstruct watermark images.

III. PRELIMINARY CONCEPTS

To improve the robustness of the algorithm, the advantages of the DWT, APBT, and SVD are combined to extract stable features from video sequences and construct zero-watermarks. To resist various frame-based attacks, the key-frame detection algorithm is used to extract the key frames from each shot of the video sequence. This section mainly introduces the concepts and characteristics of the three transforms and the key-frame detection algorithm.

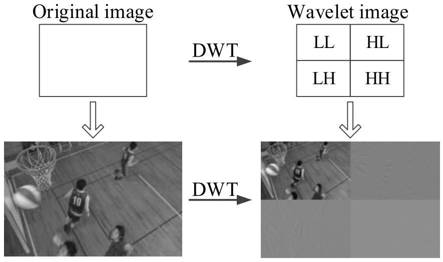


FIGURE 2. Wavelet decomposition diagram of the first frame of the BasketballDrill sequence.

A. DWT

The wavelet analysis method comes from Fourier analysis. Since it has a high-frequency resolution and low time resolution at low frequencies, it is called an image microscope. Compared with the Fourier transform, the advantage of wavelet analysis is that it has good localization in both the time and frequency domains. In the image processing field, the DWT is obtained by discretizing the scale factor and displacement factor of the continuous wavelet transform. Taking the Y component of the first frame of BasketballDrill as an example, a wavelet image can be obtained after performing the DWT, as shown in Fig. 2.

Obviously, most of the energy of the original image is concentrated in the LL sub-band, which is called the approximate sub-band. The LH, HL, and HH sub-bands contain horizontal edge details, vertical edge details, and diagonal edge details of the image, respectively, and are called detail sub-bands. The DWT has good multiscale and multiresolution characteristics. To improve the robustness of the watermarking algorithms, the LL sub-band is usually selected to embed the watermark.

B. APBT

The APBT is developed on the basis of the DCT. Fig. 3 shows the normalized amplitude-frequency response of each filter in the DCT and APBT matrices. It can be observed that compared with the DCT, the APBT has better low-frequency energy accumulation and high-frequency energy attenuation characteristics. In addition, the APBT can eliminate the blocking artifacts that are caused by the block DCT in JPEG compression at a low bit rate. Currently, the APBT has been applied to many fields, such as image compression and digital watermarking.

The APBT can be divided into different types based on the orthogonal transform matrix. In this paper, the all phase discrete cosine biorthogonal transform (APDCBT) [29] is

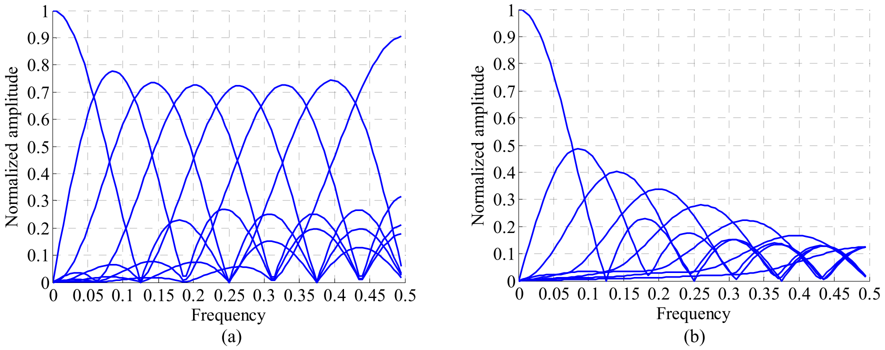


FIGURE 3. The normalized amplitude-frequency response of each filter in different matrices: (a) the DCT matrix and (b) the APBT matrix.

used, and its transform matrix can be expressed as follows:

$$B(i,j)=\begin{cases}\dfrac{N-i}{N^2}, & i=0,1,\cdots,N-1,\ j=0,\\ \dfrac{1}{N^2}[(N-i)\cos\dfrac{ij\pi}{N}-\csc\dfrac{j\pi}{N}\sin\dfrac{ij\pi}{N}], & \\ \qquad i=0,1,\cdots,N-1,\ j=1,2,\cdots,N-1, & \end{cases} \tag{1}$$

where N is the size (height and width) of the image block $\boldsymbol{X}$. Applying the APBT to $\boldsymbol{X}$ can be expressed as follows:

$$\boldsymbol{Y}=\boldsymbol{BXB}^{\mathrm{T}}, \tag{2}$$

where $\boldsymbol{Y}$ is the transform coefficient matrix, and $\boldsymbol{B}$ is the APBT matrix with the same size as $\boldsymbol{X}$.

C. SVD

SVD is a type of matrix decomposition that can be applied to any matrix. Through SVD, a complex matrix can be expressed by multiplying several smaller and simpler submatrices. These obtained submatrices can describe the important characteristics of the matrix [34]. For any matrix $\boldsymbol{A}$, its SVD can be expressed as follows:

$$\boldsymbol{A}=\boldsymbol{USV}^{\mathrm{T}}, \tag{3}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices and $\boldsymbol{S}$ is a diagonal matrix. Since matrix $\boldsymbol{S}$ contains the singular values of matrix $\boldsymbol{A}$, it is called a singular value matrix; it contains the main information of the image and has high stability. Based on this characteristic, SVD is widely used in the field of digital watermarking.

D. KEY-FRAME DETECTION ALGORITHM

A video is composed of a group of image frames and contains multiple shots. The frames that covered in each shot are highly correlated. For a video watermarking algorithm, if the watermark is embedded into the same positions of each frame, it will increase the execution time and lead to poor real-time performance. Furthermore, the watermark may be removed by statistical comparisons and averaging, thereby making the watermarking lose its copyright protection ability. To solve this problem, this paper uses a key-frame detection algorithm based on the correlation between adjacent frames. By comparing the correlation coefficients with the predefined threshold T, the key frames that can represent the main content of the video can be detected. The method for calculating the correlation coefficients is shown in (4), as shown at the bottom of this page where F_k and F_{k+1} are the k-th and $(k+1)$-th frames of the video of size $M \times N$, respectively, and μ_{F_k} and $\mu_{F_{k+1}}$ are the mean values of F_k and F_{k+1}, respectively. The pseudocode of the key-frame detection process is shown in Algorithm 1.

IV. PROPOSED VIDEO ZERO-WATERMARKING ALGORITHM

In this section, the proposed video zero-watermarking algorithm based on hybrid transforms is presented in detail. The algorithm consists of two parts: zero-watermark generation and zero-watermark detection. The specific steps are explained in the following subsections.

A. ZERO-WATERMARK GENERATION

Fig. 4 shows the generation process of a zero-watermark, and the specific steps are as follows:

$$\rho_{F_kF_{k+1}}=\frac{\sum\limits_{m=0}^{M-1}\sum\limits_{n=0}^{N-1}[F_k(m,n)-\mu_{F_k}][F_{k+1}(m,n)-\mu_{F_{k+1}}]}{\sqrt{\left(\sum\limits_{m=0}^{M-1}\sum\limits_{n=0}^{N-1}[F_k(m,n)-\mu_{F_k}]^2\right)\left(\sum\limits_{m=0}^{M-1}\sum\limits_{n=0}^{N-1}[F_{k+1}(m,n)-\mu_{F_{k+1}}]^2\right)}}, \tag{4}$$

Algorithm 1 Key-Frame Detection Algorithm

```
Variable Declaration:
    BasketballDrill: original video sequence
    $F$: read the video sequence
    $Y$: luminance components
    Num: number of frames in the video
    Count: number of extracted key frames
    $k$: location of a frame in the original video
    $L$: locations of extracted key frames
    $p$: correlation coefficients
    $T$: predefined threshold
Key-Frame Detection Procedure:
 1: Read the video sequence and obtain $Y$ components
      $F \leftarrow$ BasketballDrill.yuv (with size of 832×480)
      $Y \leftarrow$ loadyuv($F$, 832, 480, Num)
 2: Parameter initialization
      Count $\leftarrow$ 1
      $L$(Count) $\leftarrow$ 1 // Take the 1st frame as first key frame
 3: Calculate correlation coefficients according to (4)
    and record locations of key frames
      for $k = 2$:Num do
        $p \leftarrow$ Corr2($F_1$, $F_k$)
        if $p < T$ then   // $T$ is adjusted according to the
        number of selected key frames from each video.
          Count = Count + 1
          $L$(Count) $\leftarrow k$
          // Determine the $k$-th frame as a key frame
          $F_1 \leftarrow F_k$  // Replace $F_1$ with $F_k$, and continue.
        end if
    end for
End Procedure
```

Step 1: Apply Algorithm 1 to the video sequence to extract the key frames and record their locations.

Step 2: Process the watermark. Use a pseudorandom number generator to generate a pseudorandom sequence. The XOR operation between the sequence and original watermark $\boldsymbol{W}$ is performed to generate the encrypted watermark.

Step 3: Divide the extracted key frames into n groups according to the size of the watermark image.

Step 4: Obtain Y components of all extracted key frames, and apply the DWT to them to obtain four sub-bands: LL, LH, HL, and HH. In this paper, only the LL sub-bands are selected for the subsequent processing.

Step 5: Apply the block-based APBT to the LL sub-bands to obtain the DC coefficient of each block, and then the DC coefficient matrix $\boldsymbol{D}$ can be obtained.

Step 6: Divide $\boldsymbol{D}$ into nonoverlapping 4 × 4 blocks, and apply SVD to each block to obtain singular value matrices.

Step 7: Extract the highest bit of the maximum singular value of each singular value matrix. If the highest bit is odd, then the feature is defined as 1; otherwise, the feature is defined as 0. After traversing all singular value matrices, the final feature sequence can be obtained.

Step 8: Generate the zero-watermark. Apply XOR between the encrypted watermark and the extracted feature sequence. Note that the encrypted watermark is circularly expanded during XOR operations. A zero-watermark sequence can be generated.

Step 9: Register the zero-watermark with an authority organization.

To more comprehensively illustrate the zero-watermark generation process, its pseudocode is shown in Algorithm 2.

B. ZERO-WATERMARK DETECTION

Fig. 5 illustrates the detection process of the zero-watermark, and the specific steps are as follows.

Step 1: Use the recorded location information to extract the key frames from the suspicious video sequence, and then divide them into n groups.

Step 2: Obtain Y components of all extracted key frames, and apply the DWT to them to obtain the LL sub-bands.

Step 3: Apply the block-based APBT to the LL sub-bands to obtain the DC coefficient of each block. The DC coefficient matrix $\boldsymbol{D}^*$ can be obtained.

Step 4: Divide $\boldsymbol{D}^*$ into 4 × 4 blocks, and apply SVD to each block to obtain the singular value matrices.

Step 5: Extract the highest bit of the maximum singular value of each singular value matrix. If the highest bit is odd, the feature is 1; otherwise, the feature is 0. After traversing all singular value matrices, the final feature sequence can be obtained.

Step 6: Apply XOR between the registered zero-watermark and the extracted feature sequence from the suspicious video. Multiple encrypted watermarks can be obtained.

Step 7: Use the pseudorandom sequence to decrypt these encrypted watermarks, and then n reconstructed watermarks can be generated.

Step 8: Postprocess the watermark. Average these n reconstructed watermarks to obtain a watermark $\boldsymbol{w}_1$. The final reconstructed watermark $\boldsymbol{W}^*$ can be obtained via (5):

$$\boldsymbol{W}^*(i,j) = \begin{cases} 1, & \boldsymbol{w}_1(i,j) > T_1, \\ 0, & \boldsymbol{w}_1(i,j) \le T_1, \end{cases} \tag{5}$$

where T_1 is a predefined threshold in the range of [0.1, 0.5].

Step 9: Compare the correlation between $\boldsymbol{W}^*$ and $\boldsymbol{W}$ and evaluate the robustness of the algorithm.

To more comprehensively illustrate the zero-watermark detection process, its pseudocode is shown in Algorithm 3.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, to evaluate the performance of the proposed algorithm, several experiments are performed using MATLAB R2014a on an Intel Core i5-4590 3.30 GHz CPU. Twenty video files are prepared as carrier videos. In this paper, we select four representative video sequences of size 832 × 480 to illustrate experimental results, which include BasketballDrill, BQMall, PartyScene, and RaceHorses [35]. In addition, the performance of the proposed algorithm on other videos with different resolutions, including

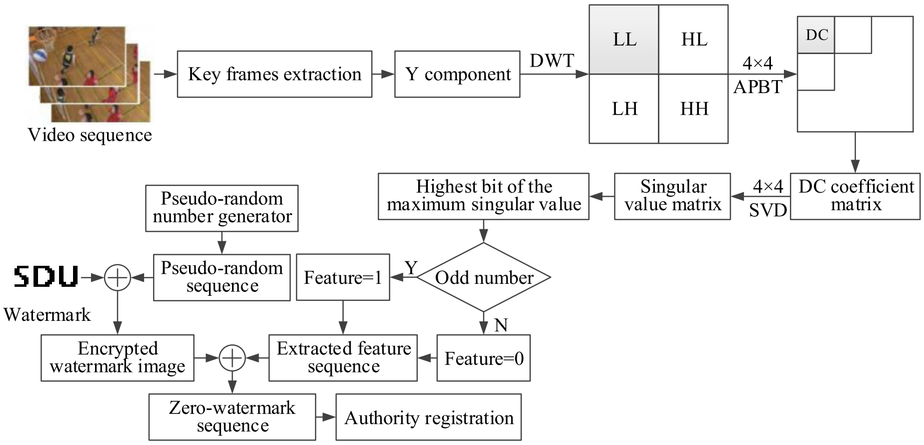


FIGURE 4. Framework of the proposed zero-watermark generation process.

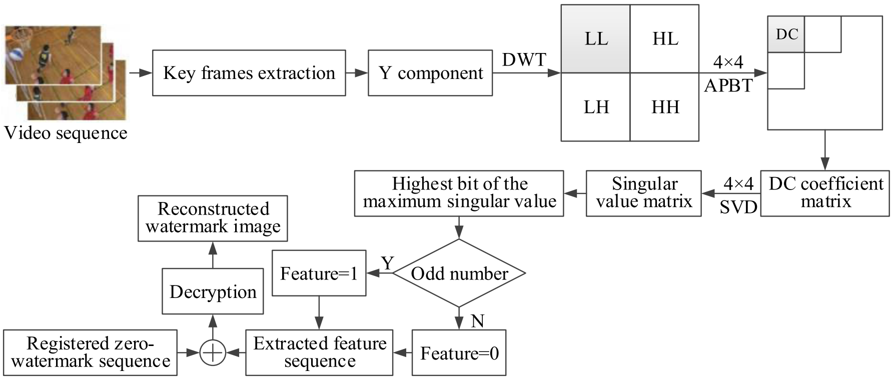


FIGURE 5. Framework of the proposed zero-watermark detection process.

FourPeople (1280 × 720), BQTerrace (1920 × 1080), and Traffic (2560 × 1600), is shown in subsection I. The objects in BasketballDrill and RaceHorses move intensely. The objects in Traffic move moderately, and the objects in FourPeople move slowly. BQMall and BQTerrace are inhomogeneous, and PartyScene has complex textures. The first 100 frames of each video are selected for the experiments. A binary image marked "SDU" with a size of 30 × 26 is used as the watermark image.

Since the zero-watermarking algorithm does not modify the original video, it will not have any impact on the videos' quality. As a result, zero-watermarking algorithms have good imperceptibility. The peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index are two commonly used indexes to evaluate the imperceptibility, and their definitions are shown as (6) and (7), respectively.

$$\mathrm{PSNR} = 10\log_{10}\frac{255^2}{\mathrm{MSE}}(\mathrm{dB}), \tag{6}$$

$$\mathrm{SSIM}(f,f_w) = \frac{(2\mu_f\mu_{f_w}+C_1)(2\sigma_{ff_w}+C_2)}{(\mu_f^2+\mu_{f_w}^2+C_1)(\sigma_f^2+\sigma_{f_w}^2+C_2)}, \tag{7}$$

where the mean square error (MSE) can be defined as:

$$\mathrm{MSE} = \frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j)-f_w(i,j)]^2, \tag{8}$$

where f and f_w are the original and watermarked video frames of size $M \times N$, respectively. In addition, μ_f and μ_{f_w} represent the mean values of f and f_w, respectively. σ_f and σ_{f_w} are the variances of f and f_w, respectively. σ_{ff_w} denotes the covariance of f and f_w, and C_1 and C_2 are two constants to maintain the stability. The higher the PSNR and SSIM are, the better the imperceptibility is. The normalized correlation coefficient (NCC) and bit error rate (BER) are selected as robustness evaluation indexes. The higher the NCC and the lower the BER are, the better the robustness of the algorithm is. The definition of the NCC is shown as (9), and the BER is defined in (10):

$$\mathrm{NCC} = \frac{\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N}\boldsymbol{W}(i,j)\times\boldsymbol{W}^*(i,j)}{\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N}\boldsymbol{W}(i,j)^2}, \tag{9}$$

$$\mathrm{BER} = \frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}|\boldsymbol{W}(i,j)-\boldsymbol{W}^*(i,j)|\times 100\%, \tag{10}$$

where $\boldsymbol{W}$ and $\boldsymbol{W}^*$ denote the original watermark and extracted watermark of size $M \times N$, respectively. The NCC is adopted to estimate the similarity between $\boldsymbol{W}^*$ and $\boldsymbol{W}$, and the BER is adopted to estimate the error rate between $\boldsymbol{W}^*$ and $\boldsymbol{W}$.

Algorithm 2 Zero-Watermark Generation

```
Variable Declaration:
    $F$: read the original video sequence
    $F_k$: key-frames sequence
    $L$: locations of extracted key frames
    Count: number of extracted key frames
    SDU: watermark image
    $W$: read the watermark image
    $m$: pseudorandom sequence
    $W_E$: encrypted watermark
    $n$: the number of groups of key frames
    $Y$: luminance components
    $D$: DC coefficient matrix
    $S$: singular value matrix
    $S_m$: maximum singular value of each singular value
        matrix in string form
    $S_{mh}$: highest bit of each $S_m$
    $F_s$: extracted feature of each block
    $W_z$: generated zero-watermark
Zero-Watermark Generation Procedure:
1:  Read the video sequence and extract key frames
        $F \leftarrow$ BasketballDrill.yuv (video frames with size
        of $832 \times 480$)
        $F_k$, $L$, Count $\leftarrow$ Key-frame Detection Algorithm
2:  Watermark preprocessing
        $W \leftarrow$ SDU (with size of $30 \times 26$)
        $m \leftarrow$randint(1, $30 \times 26$)
        $W_E \leftarrow$ XOR($W$, $m$)
3:  Feature extraction
        $n \leftarrow$ [Count - mod(Count, 2)]/2  // Divide key frames
        into $n$ groups
        for $n$ groups of key frames do
            [LL, LH, HL, HH] $\leftarrow$ DWT($Y$)
            $D \leftarrow$ APBT(LL)
            [$U$ $S$ $V$] $\leftarrow$ SVD($D$)
            $S_m \leftarrow$ num2str(max($S$))
            $S_{mh} \leftarrow S_m(1)$
            if mod($S_{mh}$, 2) $= 1$ then
                $F_s = 1$
            else $F_s = 0$
            end if
        end for
        // The feature sequence $F_{se}$ can be obtained by
        composing all $F_s$
4:  Zero-watermark generation
      $W_z \leftarrow$ XOR($W_E$, $F_{se}$)
End Procedure
```

To estimate the robustness of the proposed algorithm to HEVC compression attacks, HEVC compression attacks with QPs from 16 to 48 with a step of 8 are applied to video sequences. In addition, the robustness to various common image processing attacks (such as noise, filtering, blurring, and sharpening), rotation, scaling, frame-based attacks, and some hybrid attacks is also tested.

Algorithm 3 Zero-Watermark Detection

```
Variable Declaration:
    $F^*$: read the suspicious video sequence
    $F_k^*$: extracted key-frame sequence according to the
        recorded locations $L$
    $n$: number of key-frame groups
    $Y$: luminance components
    $D^*$: DC coefficient matrix
    $S^*$: singular value matrix
    $S_m^*$: maximum singular value of each singular value
        matrix in string form
    $S_{mh}^*$: highest bit of each $S_m^*$
    $F_s^*$: extracted feature of each block
    $W_z$: registered zero-watermark
    $W_E^*$: encrypted watermarks
    $W_D^*$: decrypted watermarks, which contain $W_{D_1}$, $W_{D_2}$,
        $\cdots$, $W_{D_n}$
    $m$: pseudorandom sequence
    $T_1$: predefined threshold
    $W^*$: final reconstructed watermark
Zero-Watermark Detection Procedure:
1:  Read the suspicious video sequence and extract key
        frames
        $F^* \leftarrow$ suspicious BasketballDrill.yuv
        $F_k^* \leftarrow$ extracted key-frame sequence from $F^*$
2:  Feature extraction
        $Y \leftarrow$ loadyuv($F_k^*$, 832, 480, $2n$)
        for $n$ groups of key frames do
            [LL*, LH*, HL*, HH*] $\leftarrow$ DWT($Y$)
            $D^* \leftarrow$ APBT(LL*)
            [$U^*$ $S^*$ $V^*$] $\leftarrow$ SVD($D^*$)
            $S_m^* \leftarrow$ num2str(max($S^*$))
            $S_{mh}^* \leftarrow S_m^*(1)$
            if mod($S_{mh}^*$, 2) $= 1$ then
                $F_s^* = 1$
            else $F_s^* = 0$
            end if
        end for
        // The feature sequence $F_{se}^*$ can be obtained by
        composing all $F_s^*$
3:  Obtain encrypted watermarks and decrypt them
        $W_E^* \leftarrow$ XOR($W_z$, $F_{se}^*$)
        $W_D^* \leftarrow$ XOR($W_E^*$, $m$)  // $W_D^* = \{W_{D_1}, W_{D_2}, \cdots,$
        $W_{D_n}\}$
4:  Watermark postprocessing and watermark reconstruction
        $w_1 \leftarrow \frac{W_{D_1}+W_{D_2}+\cdots+W_{D_n}}{n}$
        for $i = 1:30$ and $j = 1:26$
            if $w_1(i,j) > T_1$ then
            $W^*(i,j) \leftarrow 1$
            else $W^*(i,j) \leftarrow 0$
            end if
        end for
End Procedure
```

TABLE 1. Comparison of the imperceptibility and robustness without attacks among different algorithms.

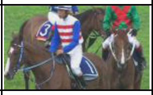
Algorithms	[29]				[21] and Proposed			
Video sequences	BasketballDrill	BQMall	PartyScene	RaceHorses	BasketballDrill	BQMall	PartyScene	RaceHorses
Watermarked frames								
PSNR (dB)	47.3540	49.9318	49.0648	49.8968	Inf	Inf	Inf	Inf
SSIM	0.9978	0.9991	0.9996	0.9983	1.0000	1.0000	1.0000	1.0000
Extracted watermarks								
NCC	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BER	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 2. The robustness of the proposed algorithm to HEVC compression attacks before watermark postprocessing.

Quantization parameter (QP)	Original video sequences							
	BasketballDrill		BQMall		PartyScene		RaceHorses	
	NCC	BER	NCC	BER	NCC	BER	NCC	BER
16	0.9997	0.0002	0.9982	0.0019	0.9986	0.0013	0.9981	0.0018
24	0.9990	0.0009	0.9973	0.0026	0.9973	0.0031	0.9951	0.0048
32	0.9981	0.0019	0.9937	0.0062	0.9939	0.0063	0.9885	0.0113
40	0.9955	0.0046	0.9867	0.0131	0.9858	0.0149	0.9746	0.0250
48	0.9898	0.0097	0.9644	0.0340	0.9687	0.0318	0.9489	0.0506

TABLE 3. The robustness of the proposed algorithm to HEVC compression attacks after watermark postprocessing.

Quantization parameter (QP)	Original video sequences							
	BasketballDrill		BQMall		PartyScene		RaceHorses	
	NCC	BER	NCC	BER	NCC	BER	NCC	BER
16	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
24	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
32	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
40	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
48	1.0000	0.0000	1.0000	0.0000	0.9925	0.0026	1.0000	0.0000

A. WITHOUT ATTACKS

To evaluate the imperceptibility and effectiveness of the algorithm, we use four representative video sequences as examples. The PSNRs and SSIMs are tested between the watermarked videos and original videos, and the watermarks are extracted from the four watermarked video sequences without attacks. The imperceptibility and robustness of the algorithms presented in [29] and [21] and the proposed algorithm without attacks are reported in Table 1.

In Table 1, [29] is a conventional video watermarking algorithm against HEVC compression attacks, and [21] and the proposed algorithm are two zero-watermarking algorithms. We can observe that the three algorithms can completely extract the watermark without attacks. However, the PSNRs of the two zero-watermarking algorithms are infinite and the SSIMs equal 1, both of which are better than those of [29]. Since zero-watermarking algorithms do not modify the original videos, they have good imperceptibility and video quality.

B. HEVC COMPRESSION ATTACKS

In the storage and transmission processes, it is inevitable for videos to undergo HEVC compression. Therefore, the watermarking algorithm must have the ability to resist HEVC compression attacks. Here, we use HEVC compression attacks with QPs from 16 to 48 with a step of 8 to test the robustness of the proposed algorithm. Taking four representative video sequences as examples, the robustness of the proposed algorithm to HEVC compression attacks before watermark postprocessing is reported in Table 2, and the robustness to HEVC compression attacks after watermark postprocessing is reported in Table 3. In addition, the robustness of the proposed algorithm is compared with [29] and [21] when facing HEVC compression attacks. The NCCs of the algorithms under different QPs are shown in Fig. 6, and the BERs under different QPs are shown in Fig. 7.

The larger the QP is, the greater the compression strength, and the more serious the video distortion. From Table 2 and Table 3, we can conclude that the proposed algorithm has good robustness to HEVC compression attacks for the different QPs before and after watermark postprocessing. Even when the QP is 48 and the video quality has been seriously damaged, the watermark can still be well reconstructed. Through the watermark postprocessing, the robustness of the algorithm is further improved. Fig. 6 and Fig. 7 show that all of the three algorithms can resist HEVC compression attacks, and the resistance of the proposed algorithm is strongest, followed by [21], and the worst is [29].

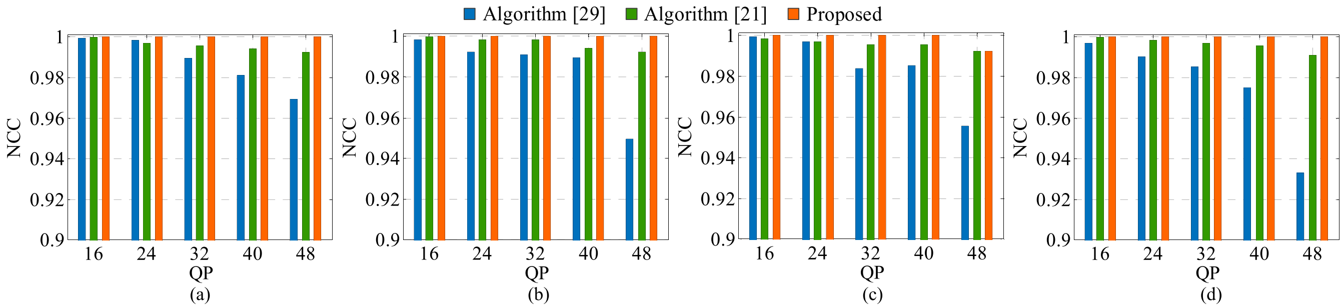


FIGURE 6. Comparison of the NCCs among different algorithms for HEVC compression attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

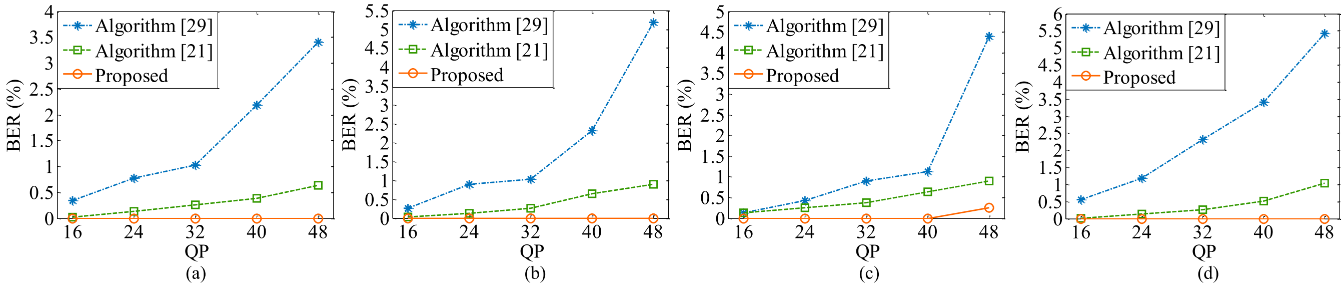


FIGURE 7. Comparison of the BERs among different algorithms for HEVC compression attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

The reason is that zero-watermarking algorithms construct zero-watermarks by extracting robust features, which usually have good robustness, whereas the conventional video watermarking algorithm may lose the embedded watermark information during HEVC compression, thus resulting in poor resistance to HEVC compression with larger QPs. The experimental results effectively prove the robustness of the proposed algorithm to HEVC compression attacks.

C. NOISE ATTACKS

Noise attacks are one of the most common image processing operations. In this paper, taking BasketballDrill as the object, the robustness of the proposed algorithm to different degrees of Gaussian noise (GN) and salt and pepper noise (SPN) is tested and compared with [29] and [21]. The test results are reported in Table 4. In addition, three other video sequences are also tested to estimate the robustness to noise attacks. The comparison results of the NCCs and BERs are shown in Fig. 8 and Fig. 9, respectively.

With increasing noise intensity, the degree of video quality damage becomes greater, and the difficulty of reconstructing watermarks is also increased. As can be observed from Table 4, the proposed algorithm has high robustness to GN and SPN attacks. In addition, its resistance to noise attacks is better than [29] and [21], especially when the intensity of noise attacks is high. From Fig. 8 and Fig. 9, we can obtain the same conclusion.

D. FILTERING ATTACKS

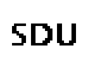
Filtering attacks are also common image processing operations. In image and video processing, filters are usually used to eliminate noises and smooth images. Common filters include the average filter (AF), median filter (MF), and Gaussian filter (GF). Here, taking the BasketballDrill video sequence as an example, the robustness of the proposed algorithm to the three filters is tested and compared with [29] and [21]. The test results are reported in Table 5. To more comprehensively evaluate the resistance of the algorithms to filtering attacks, three other videos are also tested. Comparisons of the NCCs and BERs are shown in Fig. 10. The template size of all filters that are used in the test is 3 × 3.

After suffering filtering attacks, the video will become smooth. Furthermore, different filters have different smoothing effects. Table 5 indicates that the three algorithms have good robustness against filtering attacks, and the proposed algorithm has the robustness that is superior to those of [29] and [21]. Fig. 10 shows that compared with the AF and GF, the robustness of [29] and [21] to MF attacks is relatively weak, while the proposed algorithm has strong robustness to all three filters.

E. BLURRING AND SHARPENING ATTACKS

Blurring and sharpening attacks are two common image processing operations. For blurring attacks, we mainly test disk blurring (DB) and motion blurring (MB) attacks. Sharpening attacks with different pixel radii (0.2 and 1) are also tested. Here, taking the BasketballDrill, BQMall, PartyScene, and RaceHorses sequences as the objects, the robustness of the proposed algorithm to blurring and sharpening attacks is tested and compared with [29] and [21]. The comparison results of the NCCs and the BERs are shown in Fig. 11 and Fig. 12, respectively, where SP represents sharpening attacks.

TABLE 4. Comparison of the robustness among different algorithms to Gaussian noise and salt and pepper noise on BasketballDrill.

Algorithms			[29]	[21]	Proposed	[29]	[21]	Proposed	[29]	[21]	Proposed
Noise attacks			0.01			0.02			0.03		
Gaussian	Attacked frames										
	Extracted watermarks										
	Robustness	NCC	0.9970	0.9970	1.0000	0.9865	0.9940	1.0000	0.9745	0.9895	0.9970
		BER	0.0064	0.0026	0.0000	0.0141	0.0051	0.0000	0.0244	0.0090	0.0026
Salt & pepper noise	Attacked frames										
	Extracted watermarks										
	Robustness	NCC	0.9985	0.9985	1.0000	0.9850	0.9985	1.0000	0.9775	0.9970	1.0000
		BER	0.0038	0.0013	0.0000	0.0154	0.0026	0.0000	0.0218	0.0051	0.0000

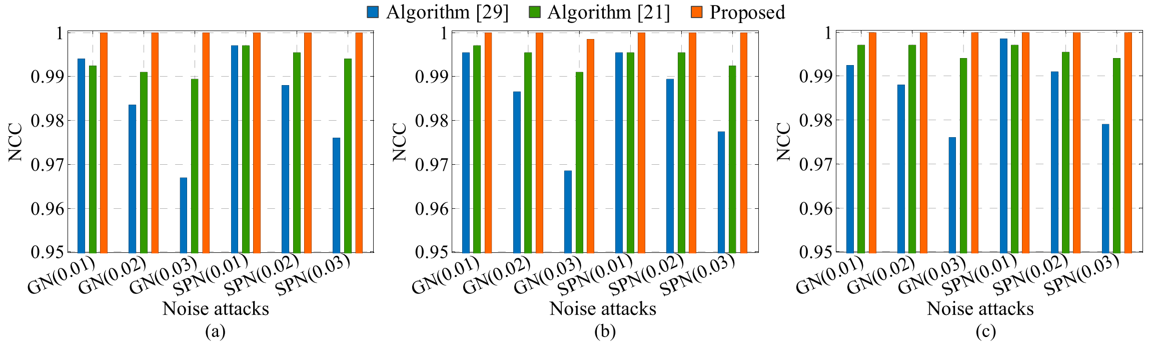


FIGURE 8. Comparison of the NCCs among different algorithms for Gaussian noise and salt and pepper noise on different videos: (a) BQMall, (b) PartyScene, and (c) RaceHorses.

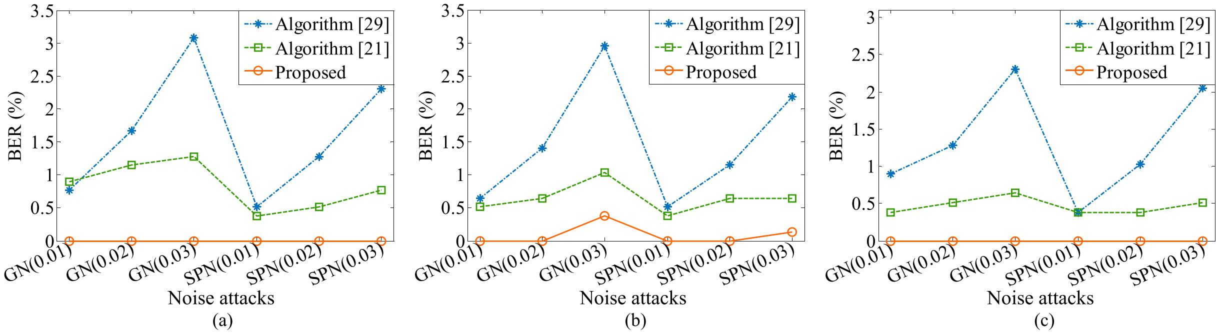


FIGURE 9. Comparison of the BERs among different algorithms for Gaussian noise and salt and pepper noise on different videos: (a) BQMall, (b) PartyScene, and (c) RaceHorses.

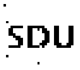
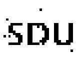
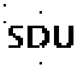
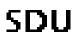
In addition, the parameters in "( )" denote the intensity of the three attacks. For example, DB (2) represents the DB attack of radius 2, and MB (2, 1) represents the MB attack of len 2 and theta 1.

From Fig. 11 and Fig. 12, it can be observed that the three algorithms have good robustness to blurring and sharpening attacks. Among them, the proposed algorithm can completely tolerate blurring attacks, and its robustness to sharpening attacks is also better than the other two algorithms.

F. GEOMETRIC ATTACKS

Rotation, scaling, and cropping are three common geometric attacks. Since a small geometric attack can lead to the problem of watermarking desynchronization, resulting in the

TABLE 5. Comparison of the robustness among different algorithms to average filter, median filter, and Gaussian filter attacks on BasketballDrill.

Algorithms		[29]	[21]	Proposed	[29]	[21]	Proposed	[29]	[21]	Proposed
Filtering attacks		Average filter (3×3)			Median filter (3×3)			Gaussian filter (3×3)		
Attacked frames										
Extracted watermarks										
Robustness	NCC	0.9925	1.0000	1.0000	0.9910	0.9985	1.0000	0.9940	1.0000	1.0000
	BER	0.0090	0.0013	0.0000	0.0103	0.0013	0.0000	0.0077	0.0000	0.0000

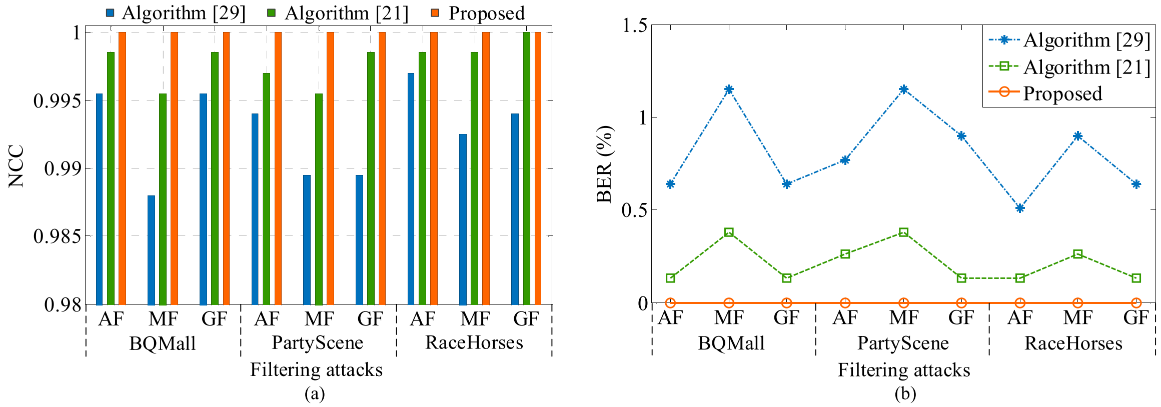


FIGURE 10. Comparison of the robustness among different algorithms to filtering attacks on BQMall, PartyScene, and RaceHorses: (a) NCCs and (b) BERs.

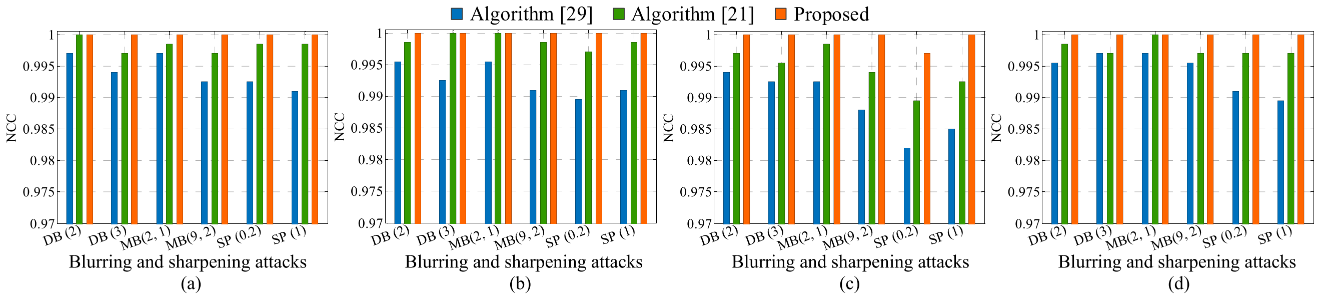


FIGURE 11. Comparison of the NCCs among different algorithms for blurring and sharpening attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

watermark being unable to be correctly extracted, a considerable number of watermarking algorithms have relatively weak resistance to geometric attacks. Here, [29], [21], and the proposed algorithm are applied to the BasketballDrill, BQMall, PartyScene, and RaceHorses sequences to compare their robustness to rotation, scaling, and cropping attacks. In addition, sticking attacks are also tested. The comparison results of the NCCs and the BERs are shown in Fig. 13 and Fig. 14, respectively, where RT stands for rotation attacks, SL stands for scaling attacks, CP represents cropping attacks, and SK represents sticking attacks.

For most video watermarking algorithms, geometric attacks are relatively difficult to resist. As can be observed from Fig. 13 and Fig. 14, all of the three algorithms have high resistance to scaling and sticking attacks. In contrast, the robustness of the three algorithms to rotation and cropping attacks is relatively weak. For the four kinds of geometric attacks, the robustness of the proposed algorithm is slightly better than those of [29] and [21].

G. FRAME-BASED ATTACKS

Frame-based attacks are unique attacks in video watermarking, and they mainly include frame switching (FS), frame dropping (FD), frame replacing (FR), and so on. Here, the robustness to these three frame-based attacks is tested on four videos, and the results are compared with those

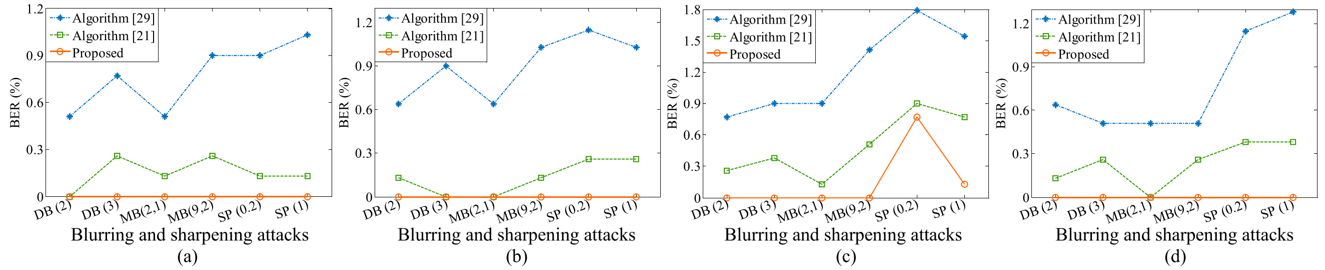


FIGURE 12. Comparison of the BERs among different algorithms for blurring and sharpening attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

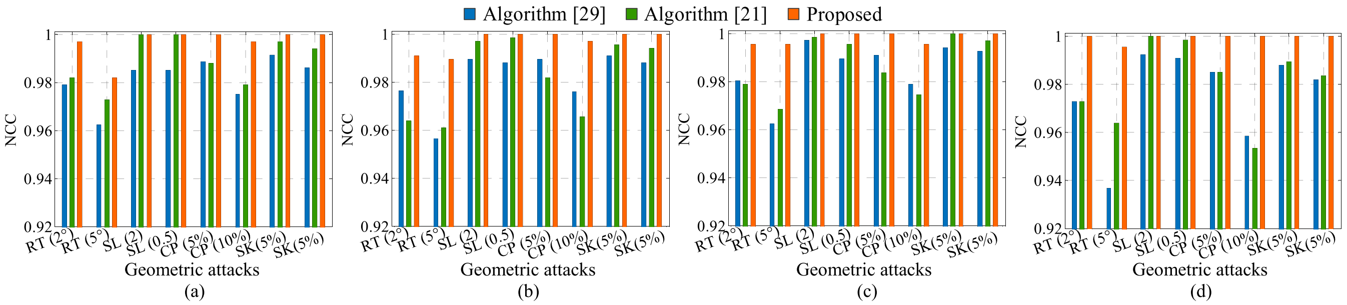


FIGURE 13. Comparison of the NCCs among different algorithms for geometric attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

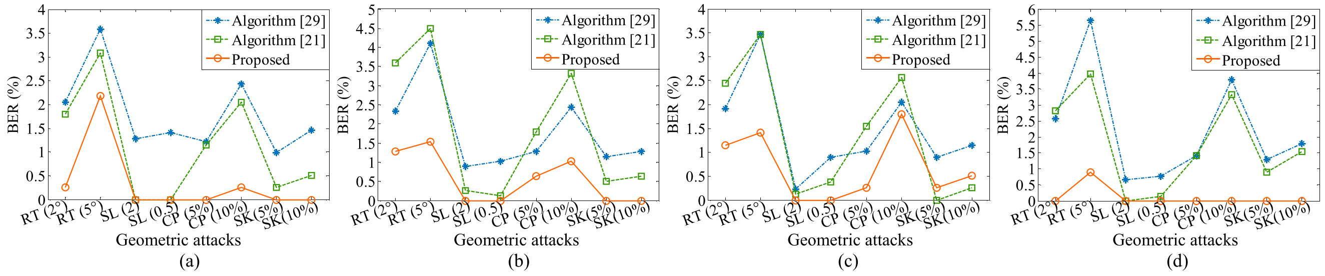


FIGURE 14. Comparison of the BERs among different algorithms for geometric attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

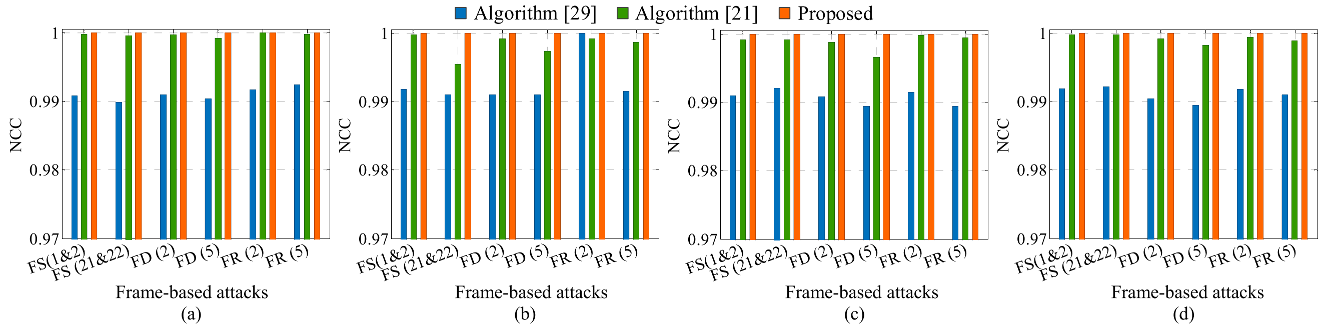


FIGURE 15. Comparison of the NCCs among different algorithms for frame-based attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

of [29] and [21]. The comparison results of the NCCs and BERs are shown in Fig. 15 and Fig. 16, respectively, where the number in "( )" represents the number of attacked frames in each video. In particular, FS (1&2) represents the switch between the first and second frames of the video.

For FS, FD, and FR attacks, both the location and the number of attacked video frames will affect the experimental results. Therefore, two frames in different segments of the video are exchanged, and different number of frames are dropped and replaced in this paper. In each frame-based attack, at least one key frame of the video is attacked. Fig. 15 and Fig. 16 show that all of the three algorithms can resist frame-based attacks well. The resistance of the proposed algorithm and [21] is slightly stronger than [29].

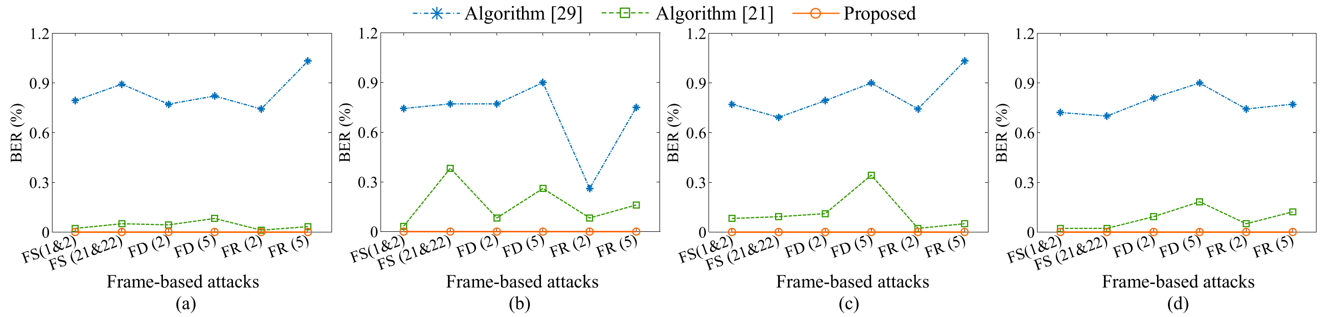


FIGURE 16. Comparison of the BERs among different algorithms for frame-based attacks on different videos: (a) BasketballDrill, (b) BQMall, (c) PartyScene, and (d) RaceHorses.

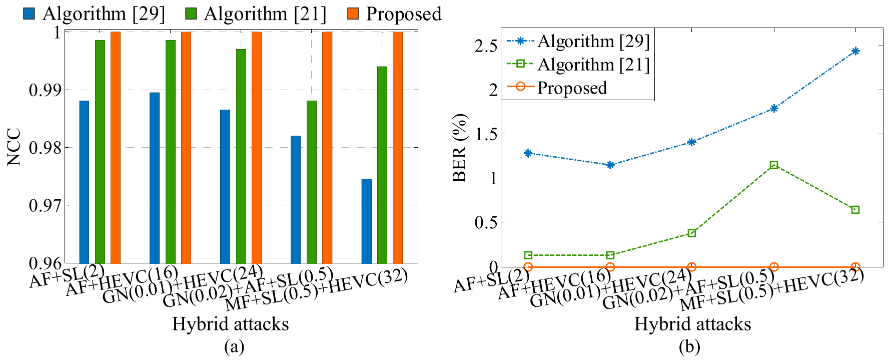


FIGURE 17. Comparison of the robustness among different algorithms to hybrid attacks on BQMall: (a) NCCs and (b) BERs.

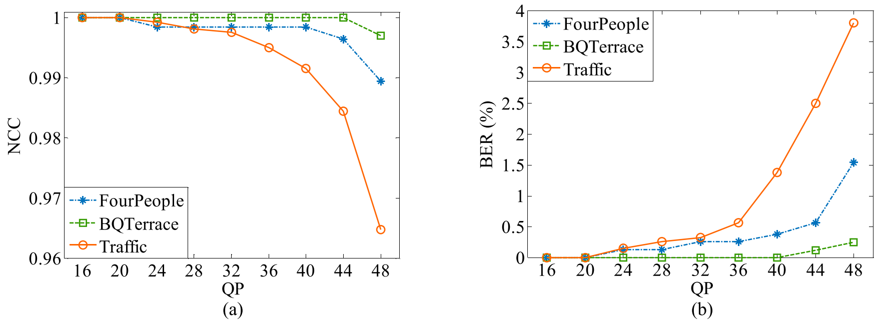


FIGURE 18. Robustness of the proposed algorithm under HEVC compression attacks with different QPs on FourPeople, BQTerrace, and Traffic: (a) NCCs and (b) BERs.

Furthermore, the watermark image that is reconstructed by the proposed algorithm is more accurate than those obtained using the other two algorithms.

H. HYBRID ATTACKS

In the actual video transmission process, videos are usually affected by more than one attack. Most video watermarking algorithms do not consider the robustness when multiple attacks are combined. In this paper, to more comprehensively prove the effectiveness of the proposed algorithm, hybrid attacks, which include two and three kinds of attacks, are considered. Taking the BasketballDrill sequence as an example, after suffering hybrid attacks combining two attacks, the first frames of the video are presented in Table 6, and the comparison of the robustness is reported in Table 7. After suffering hybrid attacks combining three attacks, the first frames of the video are presented in Table 8, and the comparison of the robustness is reported in Table 9. In addition, some hybrid attacks are also performed on the BQMall sequence, and the comparisons of the NCCs and BERs to hybrid attacks are shown in Fig. 17.

From Table 7 and Table 9, it can be observed that the proposed algorithm can completely tolerate hybrid attacks combining several common image processing attacks. In contrast, the watermark images that are reconstructed by [29] and [21] have more obvious false detection bits under hybrid attacks. The same conclusion can be drawn from Fig. 17.

According to the results of the evaluation of the robustness to various attacks, compared with [29] and [21], the proposed algorithm has higher robustness, and watermark images can be more accurately and completely reconstructed.

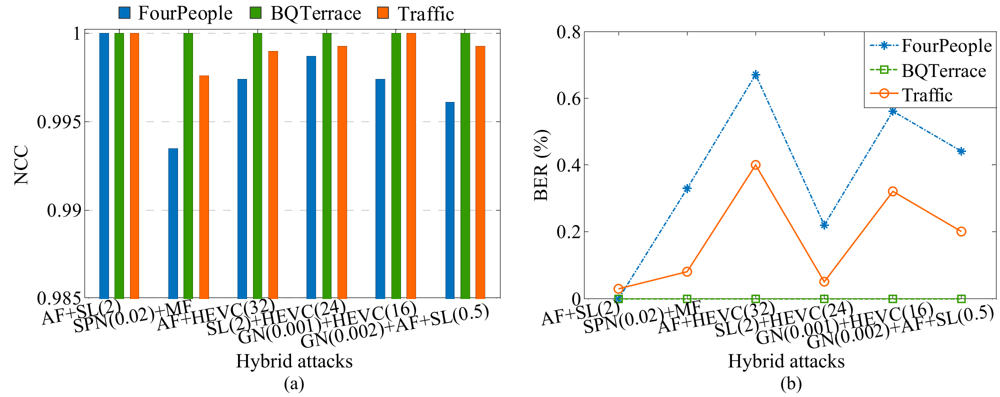


FIGURE 19. Robustness of the proposed algorithm to hybrid attacks on FourPeople, BQTerrace, and Traffic: (a) NCCs and (b) BERs.

TABLE 6. First frames of the attacked BasketballDrill sequence under hybrid attacks combining two attacks.

Attacks	Gaussian noise(0.01) + HEVC(16)	Average filter(3×3) + HEVC(24)	Median filter(3×3) + HEVC(32)	Scaling(0.5) + HEVC(40)
Attacked frames				
Attacks	Gaussian noise(0.02) + Average filter(3×3)	Gaussian noise(0.03) + Scaling(0.5)	Average filter(3×3) + Scaling(2)	Median filter(3×3) + Scaling(2)
Attacked frames				

TABLE 7. Comparison of the robustness among different algorithms to hybrid attacks combining two attacks on BasketballDrill.

Attacks	[29]			[21]			Proposed		
	Extracted watermarks	Robustness		Extracted watermarks	Robustness		Extracted watermarks	Robustness	
		NCC	BER		NCC	BER		NCC	BER
Gaussian noise(0.01) + HEVC(16)		0.9925	0.0090		0.9970	0.0038		1.0000	0.0000
Average filter(3×3) + HEVC(24)		0.9940	0.0077		0.9985	0.0013		1.0000	0.0000
Median filter(3×3) + HEVC(32)		0.9805	0.0192		0.9970	0.0026		1.0000	0.0000
Scaling(0.5) + HEVC(40)		0.9760	0.0231		0.9925	0.0064		1.0000	0.0000
Gaussian noise(0.02) + Average filter(3×3)		0.9895	0.0115		0.9970	0.0026		1.0000	0.0000
Gaussian noise(0.03) + Scaling(0.5)		0.9730	0.0256		0.9955	0.0038		1.0000	0.0000
Average filter(3×3) + Scaling(2)		0.9895	0.0115		1.0000	0.0000		1.0000	0.0000
Median filter(3×3) + Scaling(2)		0.9850	0.0141		0.9985	0.0013		1.0000	0.0000

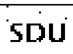
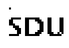
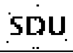
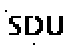
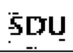
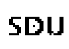
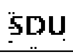
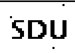
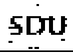
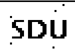
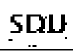
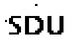
I. DIFFERENT RESOLUTIONS ANALYSIS

An effective video watermarking algorithm should also consider the application in different resolution videos and analyze its effects. Here, we select three representative video sequences with different resolutions to illustrate the experimental results, which include FourPeople (1280 × 720), BQTerrace (1920 × 1080), and Traffic (2560 × 1600). HEVC compression attacks with different QPs are tested on the three videos, and the robustness of the proposed algorithm is shown in Fig. 18. In addition, hybrid attacks are also applied to these videos to test the robustness. The NCCs and BERs of the proposed algorithm to hybrid attacks are shown in Fig. 19.

TABLE 8. First frames of the attacked BasketballDrill sequence under hybrid attacks combining three attacks.

Attacks	Gaussian noise(0.01) + Average filter(3×3) + HEVC(16)	Gaussian noise(0.02) + Median filter(3×3) + HEVC(24)	Gaussian noise(0.03) + Scaling(2) + HEVC(32)
Attacked frames			
Attacks	Average filter(3×3) + Scaling(2) + HEVC(32)	Median filter(3×3) + Scaling(0.5) + HEVC(40)	Gaussian noise(0.02) + Average filter(3×3) + Scaling(0.5)
Attacked frames			

TABLE 9. Comparison of the robustness among different algorithms to hybrid attacks combining three attacks on BasketballDrill.

Attacks	[29]			[21]			Proposed		
	Extracted watermarks	Robustness		Extracted watermarks	Robustness		Extracted watermarks	Robustness	
		NCC	BER		NCC	BER		NCC	BER
Gaussian noise(0.01) + Average filter(3×3) + HEVC(16)		0.9940	0.0077		0.9985	0.0026		1.0000	0.0000
Gaussian noise(0.02) + Median filter(3×3) + HEVC(24)		0.9880	0.0128		0.9970	0.0038		1.0000	0.0000
Gaussian noise(0.03) + Scaling(2) + HEVC(32)		0.9730	0.0256		0.9985	0.0026		1.0000	0.0000
Average filter(3×3) + Scaling(2) + HEVC(32)		0.9835	0.0167		0.9955	0.0038		1.0000	0.0000
Median filter(3×3) + Scaling(0.5) + HEVC(40)		0.9715	0.0269		0.9925	0.0064		1.0000	0.0000
Gaussian noise(0.02) + Average filter(3×3) + Scaling(0.5)		0.9820	0.0179		0.9955	0.0038		1.0000	0.0000

Because each video has its own characteristics, during the complex HEVC compression process, different degrees of distortion may occur due to different block division methods, prediction modes, and so on, resulting in slight differences in the robustness of the proposed algorithm for videos with different resolutions. From Fig. 18, we can observe that the NCC of the proposed algorithm under different QPs is no less than 0.96, and the BER is no greater than 4%, which can further prove the effectiveness of the algorithm against HEVC compression attacks. Fig. 19 shows that the NCC of the proposed algorithm after hybrid attacks is no less than 0.99, while the BER is no greater than 0.8%, which strongly proves that the proposed algorithm is still robust to hybrid attacks when applied to videos with different resolutions.

J. DIFFERENCE ANALYSIS OF EXTRACTED FEATURES

To ensure the uniqueness of zero-watermarks constructed by different videos, the difference between the extracted features should be obvious. Taking the BQMall, FourPeople, BQTerrace, and Traffic as the objects, the percentage of different bits (PDB) between features extracted from them is estimated. The BER is selected as the evaluation index, and the experimental results are shown in Table 10.

TABLE 10. Percentage of different bits between features extracted from BQMall, FourPeople, BQTerrace, and Traffic (%).

Videos	BQMall	FourPeople	BQTerrace	Traffic
BQMall	0	34.15	53.09	32.69
FourPeople	34.15	0	53.85	36.49
BQTerrace	53.09	53.85	0	52.93
Traffic	32.69	36.49	52.93	0
Random	49.62	50.08	51.20	49.42

In Table 10, "Random" is a random binary sequence with the same size as extracted feature sequences, which is used as a reference. As can be observed from the last row of Table 10, the ideal value for the PDB should be around 50%. The PDBs between features extracted from BQTerrace and features extracted from other videos are close to the ideal value. However, the PDBs between BQMall, FourPeople, and Traffic range from 30% to 40%. Table 10 has shown the obvious difference between extracted features from different videos. However, it can also be concluded that the proposed algorithm may have a certain degree of false-alarm problem for some videos in the detection process. The experimental results from part B to part I show that the BERs of the proposed algorithm for all kinds of tested attacks are no

TABLE 11. Comparison of the execution time for watermark embedding and extraction per frame of the different algorithms (second).

Algorithms	Embedding time		Extraction time	
	BasketballDrill	BQTerrace	BasketballDrill	BQTerrace
[29]	0.2122	0.3902	0.0494	0.1404
[21]	0.2461	0.2675	0.1507	0.2457
Proposed	0.1920	0.2197	0.1033	0.1841

TABLE 12. Comparisons among different algorithms in terms of the type, transforms, watermark image, watermark preprocessing, capacity, PSNR, SSIM, the resistance to various attacks, and extraction time.

Algorithms	[29]	[21]	Proposed
Algorithm type	Conventional watermarking	Zero-watermarking	Zero-watermarking
Transforms	APBT, SVD	DWT, 3D-DCT, LPM	DWT, APBT, SVD
Watermark image	Binary	Binary	Binary
Watermark preprocessing	None	Logistic map	Pseudorandom sequence
Capacity/per frame (bits)	780	390	390
PSNR (dB)	49.0618	Infinite	Infinite
SSIM	0.9987	1.0000	1.0000
HEVC compression	Yes (3rd)	Yes (2nd)	Yes (1st)
Noise, filtering, blurring, sharpening attacks	Yes (3rd)	Yes (2nd)	Yes (1st)
Geometric attacks	Yes (2nd)	Yes (2nd)	Yes (1st)
Frame-based attacks	Yes (3rd)	Yes (2nd)	Yes (1st)
Hybrid attacks	Yes (3rd)	Yes (2nd)	Yes (1st)
Time/per frame (second)	0.2616	0.3968	0.2953

greater than 4%. Combined with the Table 10, we can set up a criterion for the proposed algorithm, which is that the algorithm is considered invalid when the BER of detected zero-watermarks exceeds 10%, to eliminate its false-alarm problem to some extent.

K. REAL-TIME ANALYSIS

To compare the real-time performance of the three algorithms, we take BasketballDrill and BQTerrace as examples, and the execution times for watermark embedding and extraction per frame are given in Table 11.

From Table 11, we can see that the embedding time per frame of the proposed algorithm is shorter than those of [29] and [21]. However, its extraction time per frame is longer than [29] and shorter than [21]. The reason is that the proposed algorithm uses watermark postprocessing process to improve the robustness. Correspondingly, it also increases the watermark extraction time. Reference [21] uses the pseudo 3D DCT to extract robust features, which also prolong the embedding and extraction time.

To summarize all compared algorithms, some information and performance of them are listed in Table 12, including the algorithm type, transforms, watermark image, watermark preprocessing, capacity, PSNR, SSIM, the resistance to various attacks, and extraction time.

In Table 12, "Yes" means that the algorithm can resist the attacks. "1st", "2nd", and "3rd" represent the level of the robustness to corresponding attacks. Among them, "1st" means that the algorithm has the strongest robustness, while "3rd" means that the algorithm has the weakest robustness. From Table 12, we can observe that the imperceptibility and robustness of zero-watermarking algorithms are better than those of the conventional watermarking algorithm. Since [21] and the proposed algorithm select coefficients with more concentrated energy to generate zero-watermarks, their capacity is lower than that of [29]. Overall, the performance of the proposed algorithm is better than that of the other two algorithms.

VI. CONCLUSION

In this paper, a hybrid transforms-based robust video zero-watermarking algorithm is proposed. It is the first time that the zero-watermarking has been introduced into videos to resist HEVC compression. Combined with the properties of hybrid transforms, robust features can be extracted from videos and robust zero-watermarks can be constructed, which can ensure the robustness of the proposed algorithm. Additionally, the proposed watermark postprocessing process can further improve the accuracy of the extracted watermarks. Since the proposed algorithm does not modify the video, it will not have any impact on the video's quality. The use of a pseudorandom sequence increases the security of the algorithm. The experimental results show that the algorithm can effectively resist HEVC compression attacks with different QPs. In addition, it has high robustness to some common image processing attacks, rotation, scaling, cropping, sticking, frame-based attacks, and hybrid attacks. In the future, we will consider how to improve the resistance to high-intensity rotation attacks, completely eliminate the false-alarm problem, and apply the algorithm to 3D HEVC videos.

REFERENCES

[1] X. Zhang and X. Wang, "Digital image encryption algorithm based on elliptic curve public cryptosystem," *IEEE Access*, vol. 6, pp. 70025–70034, Nov. 2018.

[2] T. D. Denemark, M. Boroumand, and J. Fridrich, "Steganalysis features for content-adaptive JPEG steganography," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1736–1746, Aug. 2016.

[3] Y. Li and J. Wang, "Robust content fingerprinting algorithm based on invariant and hierarchical generative model," *Digital Signal Process.*, vol. 85, pp. 41–53, Feb. 2019.

[4] Q. Su, Z. Yuan, and D. Liu, "An approximate schur decomposition-based spatial domain color image watermarking method," *IEEE Access*, vol. 7, pp. 4358–4370, 2019.

[5] L. Agilandeeswari and K. Ganesan, "A robust color video watermarking scheme based on hybrid embedding techniques," *Multimedia Tools Appl.*, vol. 75, no. 14, pp. 8745–8780, Jul. 2016.

[6] K. M. Singh, "A robust rotation resilient video watermarking scheme based on the SIFT," *Multimedia Tools Appl.*, vol. 77, no. 13, pp. 16419–16444, Jul. 2018.

[7] S. P. A. Sathya and S. Ramakrishnan, "Fibonacci based key frame selection and scrambling for video watermarking in DWT–SVD domain," *Wireless Pers. Commun.*, vol. 102, no. 2, pp. 2011–2031, Sep. 2018.

[8] Y. Himeur and A. Boukabou, "A robust and secure key-frames based video watermarking system using chaotic encryption," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8603–8627, Apr. 2018.

[9] G. Gupta, V. K. Gupta, and M. Chandra, "An efficient video watermarking based security model," *Microsyst. Technol.*, vol. 24, no. 6, pp. 2539–2548, Jun. 2018.

[10] F. Yang, Y. Zhu, Y. Jiang, and Y. Qing, "A text zero-watermarking method based on keyword dense interval," *Proc. SPIE*, vol. 10420, Jul. 2017, Art. no. 104203K.

[11] Z. Ali, M. S. Hossain, G. Muhammad, and M. Aslam, "New zero-watermarking algorithm using Hurst exponent for protection of privacy in telemedicine," *IEEE Access*, vol. 6, pp. 7930–7940, Jan. 2018.

[12] M. Ghadi, L. Laouamer, L. Nana, and A. Pascu, "A novel zero-watermarking approach of medical images based on Jacobian matrix model," *Secur. Commun. Netw.*, vol. 9, no. 18, pp. 5203–5218, Dec. 2016.

[13] W. Jumana, H. D. Jun, and H. Saad, "A robust optimal zero-watermarking technique for secret watermark sharing," *Int. J. Secur. Appl.*, vol. 8, no. 5, pp. 349–360, Sep. 2014.

[14] A. Rani and B. Raman, "An image copyright protection scheme by encrypting secret data with the host image," *Multimedia Tools Appl.*, vol. 75, no. 2, pp. 1027–1042, Jan. 2016.

[15] A. Rani and B. Raman, "An image copyright protection system using chaotic maps," *Multimedia Tools Appl.*, vol. 76, no. 2, pp. 3121–3138, Jan. 2017.

[16] W. Abdul, Z. Ali, S. Ghouzali, B. Alfawaz, G. Muhammad, and M. S. Hossain, "Biometric security through visual encryption for fog edge computing," *IEEE Access*, vol. 5, pp. 5531–5538, Apr. 2017.

[17] Z. Xia, X. Wang, W. Zhou, R. Li, C. Wang, and C. Zhang, "Color medical image lossless watermarking using chaotic system and accurate quaternion polar harmonic transforms," *Signal Process.*, vol. 157, pp. 108–118, Apr. 2019.

[18] C.-P. Wang, X.-Y. Wang, X.-J. Chen, and C. Zhang, "Robust zero-watermarking algorithm based on polar complex exponential transform and logistic mapping," *Multimedia Tools Appl.*, vol. 76, no. 24, pp. 26355–26376, Dec. 2017.

[19] H.-H. Tsai, Y.-S. Lai, and S.-C. Lo, "A zero-watermark scheme with geometrical invariants using SVM and PSO against geometrical attacks for image protection," *J. Syst. Softw.*, vol. 86, no. 2, pp. 335–348, Feb. 2013.

[20] T. Liu, "A zero-watermarking method to protect intellectual property under strong geometric attacks," in *Proc. 2nd Int. Conf. Multimedia Image Process.*, Wuhan, China, Mar. 2017, pp. 172–176.

[21] D. Li, L. Qiao, and J. Kim, "A video zero-watermarking algorithm based on LPM," *Multimedia Tools Appl*, vol. 75, no. 21, pp. 13093–13106, Nov. 2016.

[22] D. Li, S. Yang, Y. Zuo, Z. Zheng, and L. Cui, "Animation zero watermarking algorithm based on edge feature," in *Proc. Int. Conf. Future Inf. Technol.*, Salerno, Italy, 2018, pp. 565–571.

[23] X. Liu, Y. Zhu, Z. Sun, M. Diao, and L. Zhang, "A novel robust video fingerprinting-watermarking hybrid scheme based on visual secret sharing," *Multimedia Tools Appl.*, vol. 74, no. 21, pp. 9157–9174, Nov. 2015.

[24] X. Liu, R. Zhao, F. Li, S. Liao, Y. Ding, and B. Zou, "Novel robust zero-watermarking scheme for digital rights management of 3D videos," *Signal Process. Image Commun.*, vol. 54, pp. 140–151, May 2017.

[25] A. Khan and S. A. Husain, "A fragile zero watermarking scheme to detect and characterize malicious modifications in database relations," *Sci. World J.*, vol. 2013, Apr. 2013, Art. no. 796726.

[26] S. Martínez, S. Gérard, and J. Cabot, "On watermarking for collaborative model-driven engineering," *IEEE Access*, vol. 6, pp. 29715–29728, 2018.

[27] X. Zhou, H. Zhang, and C. Wang, "A robust image watermarking technique based on DWT, APDCBT, and SVD," *Symmetry*, vol. 10, no. 3, Mar. 2018, Art. no. 77.

[28] J.-L. Lin, Y.-W. Chen, Y.-W. Huang, and S.-M. Lei, "Motion vector coding in the HEVC standard," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 957–968, Dec. 2013.

[29] C. Wang, R. Shan, and X. Zhou, "Anti-HEVC recompression video watermarking algorithm based on the all phase biorthogonal transform and SVD," *IETE Tech. Rev.*, vol. 35, no. s1, pp. 42–58, Jun. 2018.

[30] A. A. Mohammed and N. A. Ali, "Robust video watermarking scheme using high efficiency video coding attack," *Multimedia Tools Appl.*, vol. 77, no. 2, pp. 2791–2806, Jan. 2018.

[31] L. Chen and J. Zhao, "Contourlet-based image and video watermarking robust to geometric attacks and compressions," *Multimedia Tools Appl.*, vol. 77, no. 6, pp. 7187–7204, Mar. 2018.

[32] W. El-Shafai, S. El-Rabaie, M. M. El-Halawany, and F. E. A. El-Samie, "Efficient hybrid watermarking schemes for robust and secure 3D-MVC communication," *Int. J. Commun. Syst.*, vol. 31, no. 4, Mar. 2018, Art. no. e3478.

[33] W. El-Shafai, E.-S. M. El-Rabaie, M. El-Halawany, and F. E. A. El-Samie, "Efficient multi-level security for robust 3D color-plus-depth HEVC," *Multimedia Tools Appl.*, vol. 77, no. 23, pp. 30911–30937, Dec. 2018.

[34] N. M. Makbol, B. E. Khoo, T. H. Rassem, and K. Loukhaoukha, "A new reliable optimized image watermarking scheme based on the integer wavelet transform and singular value decomposition for copyright protection," *Inf. Sci.*, vol. 417, pp. 381–400, Nov. 2017.

[35] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-L1100, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Jan. 2013.

XIAOYAN YU received the B.E. degree in information engineering from Jilin University, China, in 2017. She is currently pursuing the M.E. degree in information and communication engineering with Shandong University, China. Her current research interests include digital image/video watermarking and computer vision.

CHENGYOU WANG (M'16) received the B.E. degree in electronic information science and technology from Yantai University, China, in 2004, and the M.E. and Ph.D. degrees in signal and information processing from Tianjin University, China, in 2007 and 2010, respectively. He is currently an Associate Professor and a Supervisor of master's students with Shandong University, Weihai, China. His current research interests include digital image/video processing and analysis, computer vision, machine learning, and wireless communication technology.

XIAO ZHOU received the B.E. degree in automation from the Nanjing University of Posts and Telecommunications, China, in 2003, the M.E. degree in information and communication engineering from Inha University, South Korea, in 2005, and the Ph.D. degree in information and communication engineering from Tsinghua University, China, in 2013. She is currently an Associate Professor and a Supervisor of master's students with Shandong University, Weihai, China. Her current research interests include wireless communication technology, digital image processing, and computer vision.