A 3-D Video Watermark Embedding Technology in DCT-CS Domain

Longfei Cai, Huimin Zhao, Jun Cai, and Li Zhu

***Abstract*—Video watermark is an important tool using for protecting digital content in information security field. In video watermark achievement, the embedding method is an important technology. For protection fingerprint images by the video watermark, the paper proposes a novel 3-D bedding scheme which can hide digital fingerprint images in DCT-CS domain. The extensive experiments have shown that the proposed scheme can effectively obtain the better trade-off between robustness and statistical transparency.**

***Index Terms*—Digital watermark, discrete cosine transform, compressed sensing, robustness.**

I. Introduction

In video information hiding field, digital watermarking is an effective means to protect multimedia content by imperceptibly embedding secret data into the host media in which digital video is a very promising host signal that can carry a large amount of data (payload) and its potential for secret communications is largely unexplored [1]-[3]. Since a video is formed from a sequence of frames, it presents the data owner with the possibility to embed and send a large amount of watermark data [4], [5].

However, the allowable bit-rate of the video stream is determined via the bandwidth of the transmission medium. Thereby, it is a great challenge to design a robust information hiding system where the hiding parameters scale properly with reduced bitrates and one has to trade-off data embedding capacity to obtain robustness at higher compression rates. For some video compression formats, such as MPEG-2 and MPEG-4, the number of variable parameters are many more (i.e., variation in bit-rate, Group of Pictures size, and so on) [4], [5]. Therefore, design of the video information hiding needs to be synthetically considered, in which video watermarking scheme is robust to variation in these parameters and achieves high compression ratios from both spatial and temporal domains [6], [7].

For satisfying robustness, the embedding method of the watermark signal is an important technology by means of information and communication theory. In [4], Biswas *et al.* proposed an adaptive compressed MPEG-2 watermarking scheme. In the scheme, the spatial spread spectrum watermark is embedded directly into the compressed streams by modifying discrete cosine transform (DCT) coefficients. Although it is substantially more effective and robust against spatial attacks such as scaling, rotation, frame averaging, and filtering, but the effectiveness is not obvious for resisting to temporal attacks like frame dropping and temporal shifting. For MPEG-4 video, Barni *et al.* presented a watermarking method of MPEG-4 Video objects in [5]. The algorithm proposed embeds a watermark in each video object by imposing a particular relationship between some predefined pairs of quantized DCT coefficients in the luminance blocks of pseudo-randomly selected macroblocks (MBs). The system presents some robustness against common manipulations such as re-coding at lower bitrates and frame dropping, however it do not thinks a good candidate for copyright protection and digital rights management (DRM) in information hiding applications. At present, a robust video adaptive watermarking method was proposed for copyright protection in discrete wavelet transform (DWT) domain [6]. In [6], the incorporation of visual model, the scheme proposed has resulted in an efficient watermarking scheme for effective copyright protection of the video. But for authentication of the video content, the method is non-effective. Hence, application of the method is limited. In [7], Chetan and Raghavendra *et al.* proposed a robust blind digital video watermarking scheme with scrambled watermarks data based on scene changes for authentication of digital video, which embeds different parts of a single watermark into different scenes of a video in DWT domain. However, the scheme has a tremendous influence for video bit-rate and invisibility under certain communication condition.

Compressed sensing (CS) theory has been developed recently and has provided a suitable method for identifying the best trade-off relation as described in [8]-[10]. It has been shown that many signal processing algorithms performed in the CS domain have very close performance as performed in the original domain [11]-[13]. Based on the CS theory, Zhao *et al.* proposed an image semi-fragile watermarking algorithm in [10]. In the algorithm, the measurement values of CS are registered as the zero-watermarking, and can recover the tampered image with the watermarking information. Lu *et al.* [14] has proposed a secure image retrieval system through random projection in CS domain. Furthermore, ref. [15] shows that CS transformation can achieve computationally secure encryption. In [16], Zhang *et al.* proposed a novel

Manuscript received September 20, 2014; revised December 22, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61272381, in part by Science and Technology Project of Education Department of Guangdong Province of China under Grant 2013KJCX0118, and in part by Science and Technology Project of Guangzhou City of China under Grant 2014J4100078.

Longfei Cai is with the School of Information Engineering, Guangdong Engineering Polytechnic, Guangzhou 510520, China (tel.: 086-020-37395921; e-mail: cailongfei2@126.com).

HuiMin Zhao, Jun Cai, and Li Zhu are with the School of Electronics and Information, Guangdong Polytechnic Normal University, Guangzhou 510665, China (tel.: 086-020-38256736; e-mail: zhaohuimin@gdin.edu.cn, gzhcaijun@gmail.com, zhuli@gdin.edu.cn).

DOI: 10.7763/IJMLC.2015.V5.508

watermarking scheme which employed a CS technique to retrieve the coefficients by exploiting the sparseness in the DCT domain. In [17], Wang and Zeng *et al*. proposed a scheme of integrated secure watermark detection and privacy preserving storage in the CS domain, in which the multimedia data and secret watermark pattern were presented to the cloud for secure watermark detection in a CS domain to protect the privacy. These works indicate that signal processing or watermarking data-mining in the CS domain is feasible and is computationally secure under certain conditions.

It is obvious from above description, the emerging theory of CS indicates that the watermark models CS -based can also be used to simplify the acquisition of high-dimensional signals that might otherwise be difficult to collect or encode [18]. Rather than collecting an entire ensemble of signal samples, CS requires only a small number of random linear measurements, with the number of measurements proportional to the sparsity level of the signal.

Inspired by the CS theory, we explore a novel video watermark embedding scheme to reduce the computational cost and simultaneously maintain a good robustness and transparency. Additionally, in order to avoid the distortion of the chrominance quality of video data, we mainly focus on the luminance component to perform our embedded scheme.

The rest of this paper is organized as follows. Section II describes the related works for CS. Section III shows the scheme proposed in this paper.

II. Related Techniques of CS Theory

In [19], and [20], the CS asserts that when a signal can be represented by a small number of non-zero coefficients, it can be perfectly recovered after being transformed by a limited number of incoherent, non-adaptive linear measurements. Suppose a signal $f \in R^N$ is a *K*-sparse vector (only *K* out of the *N* elements of *f* are nonzero) and can be transformed to $x \in R^M, M < N$, where $x = \Psi f$ and Ψ is called as the sparse matrix. For images, typical choices of Ψ include the DCT and DWT. If Ψ satisfies RIP (Restricted Isometry Property), [19], [20] shows solving the bellow optimization problem

$$\min \| f \|_1 \qquad s.t. \qquad x = \Psi f \tag{1}$$

This is equivalent to finding the sparsest solutions to $x = \Psi f$, provided $M \geq Ck \log(N/K)$, where *C* is a small constant. The CS theory states that such a signal *x* can be reconstructed by taking only *M* linear projection, non-adaptive measurements as follows

$$y = \Phi x = \Phi \Psi f \tag{2}$$

where y is an $M \times 1$ sampled vector, and Φ is an $M \times N$ measurement matrix that is incoherent with Ψ, i.e., the maximum magnitude of the element in $\Phi\Psi$ is small.

Equation (1) presents an *l*1 minimization problem which can be solved by orthogonal matching pursuit algorithm. It has been shown that it is feasible for many signal processing algorithms to be performed in the CS domain [13], [19], and [20]. For (2), if the entries of matrix Φ are generated from a Gaussian distribution with zero mean and variance $\sigma \in 1/m$, Φ is a RIP matrix with overwhelming probability in [19], [20]. The Gaussian CS matrix suits include the seeds and a random function.

In the practical application, an image with the size $N = N_1 \times N_2$ is divided generally into $B \times B$ blocks, and each block is sampled using an appropriately-sized measurement matrix Φ_B in CS domain. That is, suppose that x_i is a sparse vector representing block *i* of the input image $x \in R^M, M < N$. The corresponding measurement sample y_i is then:

$$y_i = \Phi_B \cdot x_i \tag{3}$$

where length of the y_i signal is *M*, and Φ_B is a $M \times B^2$ measurement matrix such that size of the *m* is $\lfloor M \cdot B^2 / N \rfloor$ and *M* is samples needed by the CS measurement for the whole image. In this way, Φ has a block-diagonal structure as

$$\Phi = \begin{bmatrix} \Phi_B & 0 & \cdots & 0 \\ 0 & \Phi_B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \Phi_B \end{bmatrix} \tag{4}$$

Therefore, the overall technique above was called block CS (BCS) [21].

III. Proposed Scheme

In this paper, we propose an effective video watermarking embedding method based on the DCT-CS domain. In MPEG coding, the video sequence is first divided into groups of pictures or frames (GOP) considered I-frame, B-frame, and P-frame [4]. According to the video sequence characteristics, the B-frame and P-frame are dependent on the I-frame. And the raw video data can also be considered as a sequence of still images. In our approach, the watermark obtained by CS technology is mainly embedded into the luminance component of each I-frame in the uncompressed domain. Fig. 1 shows the flowchart of the proposed scheme. Details of the proposed techniques are described in the following subsections.

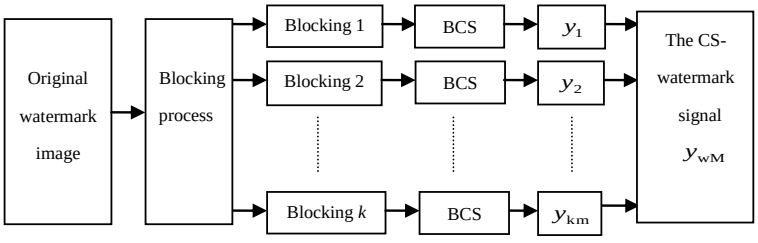


Fig. 1. Sketch of the proposed scheme in DCT-CS domain.

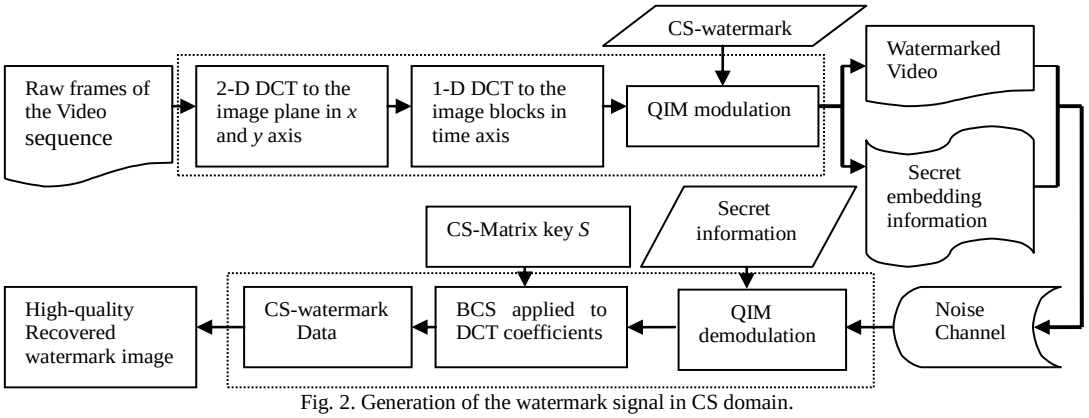


Fig. 2. Generation of the watermark signal in CS domain.

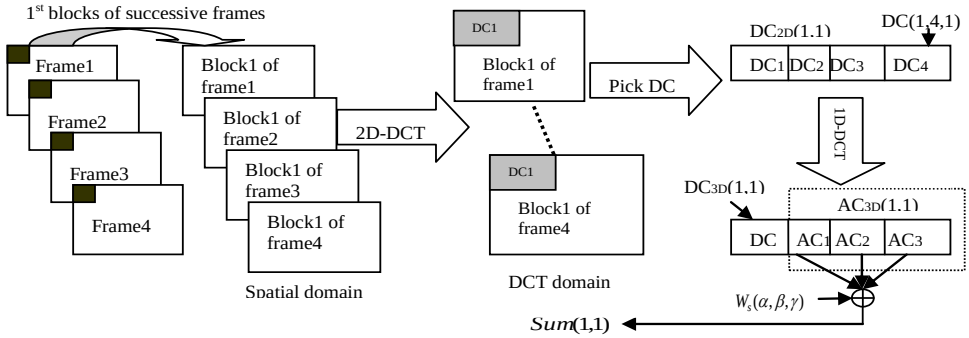


Fig. 3. 3D-DCT process worked on the $Sum(1,1)$.

A. CS-Watermark Signal

In the scheme presented, we divide firstly the original watermark image into a lot of nonoverlapping blocks, in which blocking criteria of the watermark image is jointly decided by size of the robust watermark signal and positioning accuracy of the watermark as extraction and recover purpose. Next, each block of the watermark image is carried out by a sparse basis matrix in which we use DCT as the sparse basis Ψ, and form various DCT coefficients blocks. Simultaneity, measurement matrix Φ_B of the BCS is deployed to sense these DCT coefficients independently within each block. The process is simply random linear projection, and can be achieved by inner product operation of corresponding two elements between Ψ and Φ_B. Here, according to principle of CS theory in [19], and [20], selection of Φ_B is incoherent with Ψ. Since the sparse basis Ψ is a type of DCT matrix, we can solve the constraint by designing an appropriate measurement matrix Φ_B according to CS principle in [21]. Finally, the watermark signal that will be called CS-watermark signal in this paper, is produced by combining the all sampling values of measurement matrix Φ_B. The realization principle of the CS-watermark signal is shown in Fig. 2.

B. CS-Watermark Embedding in DCT-CS Domain

In the paper, for video data of I frame, we take several successive frames as a group. Every frame within a group will be divided into a number of blocks which will be transformed into the DCT domain by the 3-D DCT method according to the principle of Fig. 1.

In Fig.1, the first, we take four consecutive frames as a group, and every frame within a group is divided some blocks. Next, the DC value of each block located in the same position of successive frames for a group is transformed into the DCT domain again. After transforming the second DCT process, we will obtain a new DC value and several AC values. Afterwards, the sum of all absolute AC values with weights is expressed as

$$Sum(\alpha,\beta)=\sum_{l} W_s(\alpha,\beta,\gamma)\,|\,AC(\alpha,\beta,\gamma)\,| \qquad (5)$$

where $Sum(\alpha,\beta)$ is the sum of all AC values, $W_s(\alpha,\beta,\gamma)$ and $AC(\alpha,\beta,\gamma)$ denote the corresponding weight value and the γ th AC value corresponding to β th block of successive frames within the α th group, respectively. Here, the initial weight value can be decided by the owner. Repeating the above steps for all blocks of frames with the same group, a sequence of sums of every block will be acquired. Fig.3 shows a simple process in which each of the four frames of 1st group is separated into the 8×8 size blocks.

After computing $Sum(\alpha,\beta)$ for all blocks, we will calculate threshold $T(\alpha)$ which will affect the robustness and transparency of the embedding CS-watermark signal. $T(\alpha)$ is associated with the characteristic of video and the number of bits to be embedded, and it can be adjusted by the information hiding owner depending on the demand and the trade-off between robustness and transparency

For estimation to $T(\alpha)$, combining our experimental data and result of reference [22], suppose that $sum_G(\alpha,\beta)$ denotes a converting form of a Gaussian function of $Sum(\alpha,\beta)$, we design an empirical formula of $T(\alpha)$ by principle of probability distribution. We sum up the probabilities of all $Sum(\alpha,\beta)$ corresponding to the same $sum_G(\alpha,\beta)$ to gain

the occurrence frequency of every $sum_G(\alpha,\beta)$. So, the distribution of $sum_G(\alpha,\beta)$ can be derived by the probability of every $sum_G(\alpha,\beta)$ value. If $Num_{bit}(\alpha)$ and Num_{Block} denote the number of embedding bits in the ith group and total blocks in a frame, respectively, meanwhile, $\mathrm{Prob}(\eta,\alpha)$ denotes the probability value of $sum_G(\alpha,\beta)=\eta$. Then, based on our experiments, $T(\alpha)$ is expressed as shown in (6)-(8).

$$\mathrm{Prob}_{th}(\alpha)=1-\frac{Num_{bit}(\alpha)}{Num_{block}} \tag{6}$$

$$\mathrm{Prob}_{th}(\alpha)\geq\sum_{t=0}^{R}\mathrm{Prob}(\eta,\alpha),\qquad \eta=Sum_G(\alpha,\beta)\geq 0 \tag{7}$$

$$T(\alpha)=R+\frac{512}{R\times\log(\frac{512}{R})\times\log(Max_{sum}(\alpha))} \tag{8}$$

when $\mathrm{Prob}_{th}(\alpha)$ and $\mathrm{Prob}(\eta,\alpha)$ are given, variable R can be estimated by (20), and we can further derive $T(\alpha)$. The threshold $T(\alpha)$ can be regarded as the tolerance range for the quantizing process in the QIM method, which is obtained depending on the video property.

After determining the $T(\alpha)$ by (8), we can obtain an integer quotient as follows:

$$Q(\alpha,\beta)=\left\lfloor\frac{Sum(\alpha,\beta)}{T(\alpha)}\right\rfloor \tag{9}$$

In order to embed the CS-watermark data, we utilize the QIM to perform the embedding operation in [23]. Based on the QIM, the embedding domain is divided into several regions. The interval of every region is the same, which equals to the $T(\alpha)$, and an index is assigned to every region. So, every region represents a bit (0 or 1) of the CS-watermark data. On the other hand, in order to the robustness of the CS-watermark data, the value of $Sum(\alpha,\beta)$ will be changed to median value in the corresponding section to resist for the distortion embedded. Fig. 4 is an example of embedding procedure.

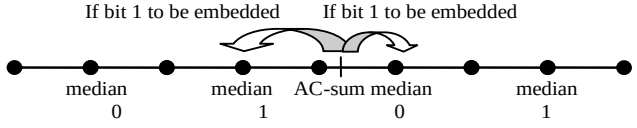


Fig. 4. The watermark signal embedded by QIM.

Since $Sum(\alpha,\beta)$ consists of several AC values, the modification of $Sum(\alpha,\beta)$ caused by embedded watermark is equal to the change of AC values. We known that the low frequency component is more robust and visually sensitive than the high frequency component. That is, if the low frequency component is modulated, it will cause the distortion more seriously, but it has a higher ability to resist attacks than the high frequency components does. Therefore, we apply the weights to modulate the $Sum(\alpha,\beta)$ defined as

$$\begin{aligned}Sum'(\alpha,\beta)=&\sum_{l}W_s(\alpha,\beta,\gamma)\,|\,AC(\alpha,\beta,\gamma)\,|\\&+W_e(\alpha,\beta,\gamma)D(\alpha,\beta)\end{aligned} \tag{10}$$

where $D(\alpha,\beta)=Sum'(\alpha,\beta)-Sum(\alpha,\beta)$, $AC(\alpha,\beta,\gamma)$ and $W_e(\alpha,\beta,\gamma)$ denote γ th AC value corresponding to β th block of successive frames within the α th group, respectively. $W_e(\alpha,\beta,\gamma)$ could be adjusted by the information hiding owner but the sum of $W_e(\alpha,\beta,\gamma)$ must be equal to one.

In this way, according to $Q(\alpha,\beta)$ and the bits to be embedded, we can obtain $Sum(\alpha,\beta)$ modified by QIM as follow

$$\begin{aligned}&Sum(\alpha,\beta)=Sum(\alpha,\beta)\\&\quad if\quad Q(\alpha,\beta)=2p,and\quad w_m=0\\&\quad if\quad Q(\alpha,\beta)=2p+1,and\quad w_m=1\\&Sum(\alpha,\beta)=Sum'(\alpha,\beta)\\&\quad if\quad Q(\alpha,\beta)=2p,and\quad w_m=1\\&\quad if\quad Q(\alpha,\beta)=2p+1,and\quad w_m=0\end{aligned} \tag{11}$$

where $w_m\in[0,1]$ denotes the embedded CS-watermark bit of every block of the watermark image from the measurement samples in CS domain. p is a random nonnegative integer determined the selection of quantizer with step size $\triangle$. Based on (10), it is easy to know whether $Sum(\alpha,\beta)$ needs to change or not when $Q(\alpha,\beta)$ is an even or odd value corresponding to the signal w_m. So, after determining the embedding position by (11), the original value of $Sum(\alpha,\beta)$ in the position can be modulated to the median value of the corresponding section by $Sum'(\alpha,\beta)$. Thus, by repeating the above procedures until all CS-watermark bits are embedded, the embedding process will be finished. Finally, all embedding positions, weights $W_s(\alpha,\beta,\gamma)$, and the threshold $T(\alpha)$ will be recorded as the secret information of embedding key.

Usually, the QIM is independent of video image content, and therefore the method may lead to serious degradation of visual quality. However, in our scheme, we modify the traditional QIM algorithm because a small amount of the CS-watermark signal are embedded into the video stream and alterative $Sum(\alpha,\beta)$ of several AC values of frames can be changed by (5)-(10), and thus maintain good visual quality and avoid distortion.

C. Extraction of the CS-Watermark Signal

The extraction process is the inverse operation of the embedding process as in Fig. 1. Firstly, the raw video sequence is separated into several groups of frames and each frame is divided into blocks. Next, the secret information of

embedding key is applied to acquire the embedding positions. After determining the embedding blocks, we only transform the selected blocks into the DCT domain rather all blocks of the frame in the video stream. By applying 3-D DCT to the selected blocks, we can further obtain an estimated value $\hat{Sum}(\alpha,\beta)$ of $Sum(\alpha,\beta)$, which is the sum of AC values of β th blocks in CS-watermarked frames within the α th group by using (4). Then we compute the quotient $\hat{Q}(\alpha,\beta)$ derived from $\hat{Sum}(\alpha,\beta)$ divided by the threshold $T(\alpha)$ by using (4), which is recorded in the secret embedding information. After computing $\hat{Q}(\alpha,\beta)$, we can exactly decide which bit is embedded by (11) expressed as

$$The\,CS-watermark\,\hat{w}_M=\begin{cases}1, if & \hat{Q}(\alpha,\beta)=2p+1\\ 0, if & \hat{Q}(\alpha,\beta)=2p\end{cases} \quad (12)$$

By repeating the above steps, the embedded CS-watermark bit $\hat{w}_M$ can be exactly ensured one by one until all bits are extracted. Finally, if $\hat{y}_{wM}$ denotes estimated data of the embedded CS-watermark, then $\hat{y}_{wM}=[\hat{w}_1,\hat{w}_2,...,\hat{w}_M]^T$ can be recovered based on the secret key in transmission of the video stream.

D. High- Recovery Quality of the Original Watermark Image in CS Domain

In CS domain, the watermark image is reconstructed from nonadaptive linear projections at the watermark generating side by viewing the decoding step as an inverse problem that is cast as a sparsity-regularized convex optimization process in [8], and [16]. Therefore, for a extracted CS-watermark signal $\hat{y}_{wM}=\Phi\cdot\hat{x}=\Phi\Psi\cdot\hat{f}$, if $\hat{f}$ is supported on a fixed (but arbitrary) set with K none zero entries, BCS with smoothed projected Landweber reconstruction (BCS-SPL) of [21] can recover $\hat{x}$ exactly with high probability when the size M of measurements samples satisfies

$$M\geq C(K\sqrt{N/B}(\log N)^2) \quad (13)$$

for some constant *C*.

According to secure key *S* in encoding side, we can reconstruct measurements matrix Φ. Thus, for data $x_k(i,j)$ of each DCT block *k* ($1\leq k\leq N/64, 1\leq i\leq 8$ and $1\leq j\leq 8$), process of BCS-SPL reconstruction is as follows

$$\begin{aligned}&\hat{V}_1(u,v)=\Phi^T\cdot\hat{y}_{w1}\\ &\hat{V}_k(u,v)=\hat{V}_{k-1}(u,v)+\Phi^T(\hat{y}_{wm}-\Phi\hat{V}_{k-1}(u,v))\\ &\hat{x}_k(i,j)=Wiener(\Psi_{IDCT}^{-1}(\hat{V}_k(u,v)))\end{aligned} \quad (14)$$

where Ψ_{IDCT}^{-1} is IDCT operation of sparse basis Ψ based-DCT, and Wiener filtering takes place in the spatial domain of reconstructing image for eliminating the block effect, and executing condition of iteration in (14) is

$$\begin{aligned}&D_k=\|\hat{x}_k(i,j)-\hat{x}_{k-1}(i,j)\|_2, for \quad k=k+1\\ &until \quad |D_k-D_{k-1}|<10^{-2}\end{aligned} \quad (15)$$

In essence, BCS-SPL reconstruction applies a Landweber step on each DCT block using measurements matrix Φ. Therefore, we can finally recover high-quality image of the original watermark data by (14)-(15) in DCT-CS domain.

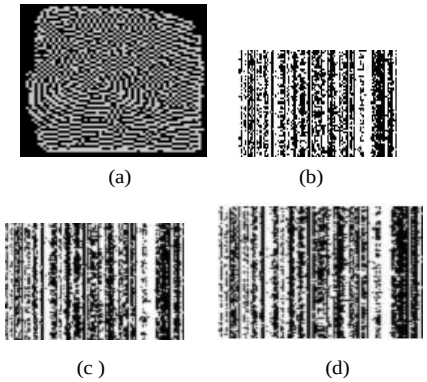
IV. Experiment Process and Results

To test and verify the performance of the CS-watermark signal proposed, the experimental results are compared with H. Huang *et al.* [22] and W. Kong *et al.* [24] to perform various attacks, including MPEG compression, noise contamination, and filtering.

A. Experiment Objects

In our experiment, the CS-watermark image is denoted by a gray level fingerprint with the size 160×160, in which the goal of fingerprint image used for the original watermark signal is to explore an new secure application of e-commerce with the principle of biological recognition in information hiding field. Before experiment, we captured the fingerprint image from sensor FPS110 of Verdicom Inc., and preprocessed the image by binaryzation method [25]. Next, the DCT matrix will be used as a sparse basis Ψ in the paper, and set sparsity, *K*=35. Finally, we obtain the CS-watermark signal base on Fig. 2. In our experiments, size of blocking *B* is 16, but other larger values like 32 and 64 usually yield similar results. The original fingerprint image and the measurement samples of various size of CS are illustrated in Fig. 5. (a), (b), (c) and (d) respectively.

Fig. 5 shows obviously that measurement values of CS have some random properties. In other words, generating procedure of the CS-watermark signal is essentially an encryption process. This is because that the measurement values of CS are decided via measurement matrix while the matrix has pseudo-random entries that can be generated by using a secure key shared between the owner and authorized users.

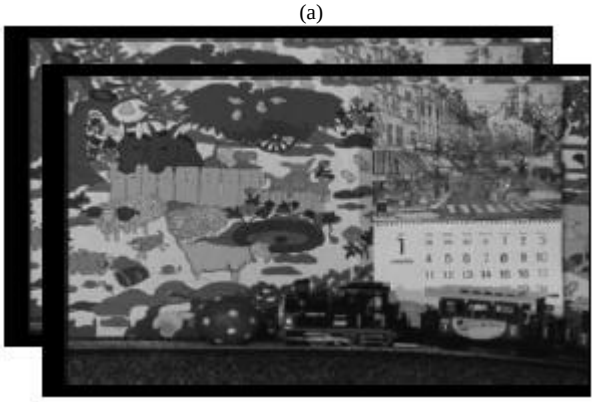


(a) (b) (c) (d)

Fig. 5. (a) Original watermark fingerprint image with the size 160×160 and (b), (c), (d) measurement samples of the image with size of 80×80, 80×100, 100×100, respectively.

For video signal, experiment used two real videos stream for demonstration. The size of each video is 720×480, and every video consists of 80 frames. Fig. 6 is the test videos. We take four successive frames as a group and each frame is divided into numbers of 8×8 blocks. In (9), the weights, $W_s(\alpha,\beta,\gamma)$, γ =1,2,3, are set to 1. The embedding weights, $W_e(\alpha,\beta,\gamma)$, γ =1,2,3, are set to 1/6,1/3, and 1/2, respectively. In addition, we embed 640 bits into a group. Therefore, we need to ten groups to embed one CS-watermark image. Meanwhile, in order to find the obvious comparison result, the watermark data of the fingerprint image are also generated by DCT method of H. Huang *et al*. [33] and SVD method of W. Kong *et al.* [35], respectively. In this case, size of the watermark data of methods of H. Huang *et al*. and W. Kong *et al.* are four times more than one of our proposed method.

(a)



(b)

Fig. 6. Test videos. (a) Basketball. (b) Mobile.

B. Performance Measure

We have known that the transparency and the robustness for information hiding are usually used to measure the system performance. For transparency, it is an important factor in information hiding. We use the peak signal-to-noise ratio (*PSNR*) as a criterion to estimate the invisibility for three methods in our experiment. The *PSNR* can clearly judge the received image quality by comparing the degree of diversity between the received image and the original one. The *PSNR* and mean square error (*MSE*) are expressed as

$$PSNR = 10\log\frac{H_{\max}^2}{MSE}$$
$$MSE = \frac{1}{I_h \times I_w}\sum_{i}^{I_h}\sum_{j}^{I_w}|H_1(i,j) - H_2(i,j)|^2 \quad (16)$$

where $H_{\max}$ is 255 gray value for a gray-level image, and $H_1(i,j)$ and $H_2(i,j)$ denote the received image and the original image corresponding to i and j coordinates in 2D space. I_h and I_w denote the height and width of the image, respectively.

Robustness is also one of important performances in video watermarking. In our experiment, a measure of the normalized correlation (*NC*) used for calculating the difference between the extracted the CS-watermark $\hat{W}(i,j)$ in receiver side and the original CS-watermark $W(i,j)$ in sender side. *NC* is defined as

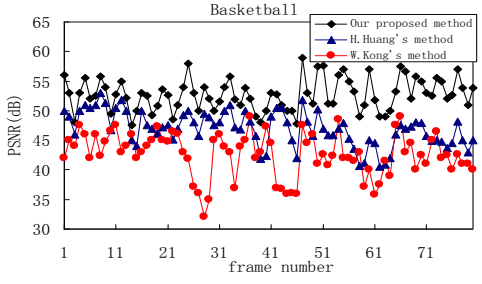
$$NC = \frac{\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} W(i,j)\hat{W}(i,j)}{\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}[W(i,j)]^2} \quad (17)$$

where $N = N_1 \times N_2$ denotes the size of the watermark image, and N_1 and N_2 are the height and width of one.

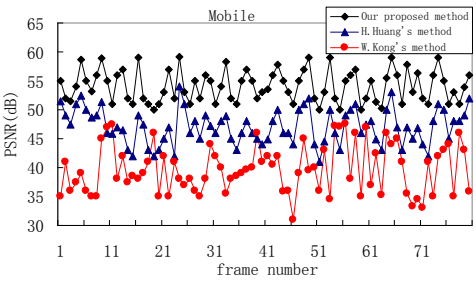
C. Experiment Results

1) The transparency of the system

In our proposed system, we embed the data of one watermark image into every 40 frames. The PSNR values of 80 frames are illustrated in Fig. 7 compared with our method, Huang *et al.*'s method, and Kong *et al.*'s method. It is obvious that the global performance of PSNR values is superior to Huang *et al.*'s method and Kong *et al.*'s method. That is, our proposed hiding data system only causes very slight distortion of video signal and simultaneously provides higher visual quality.



(a)



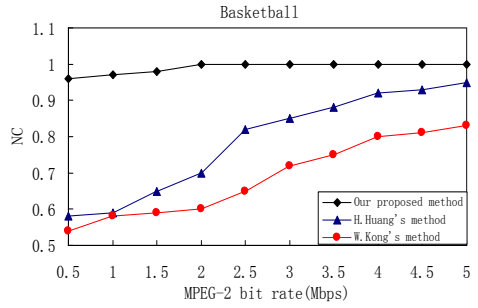
(b)

Fig. 7. PSNR comparison with three methods.

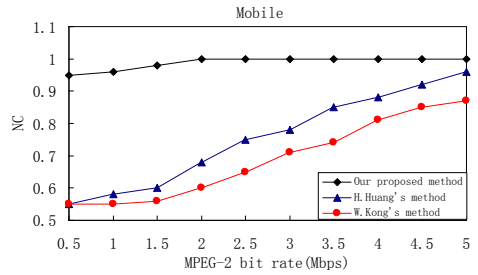
2) The robustness of the system

Compress process: Fig. 8 and Fig. 9 show the results after those of compressions under the different bit rates compared

with our method, Huang *et al.*'s method, and Kong *et al.*'s method. Obviously, the higher the NC value, the better the robustness of the hiding system. On the other hand, after embedded watermark, the hiding data can almost be extracted even though the bit rate is as low as 0.5Mbps. In other words, the results show that our proposed method can efficiently resist compression attacks of MPEG-2 and H.264 for transmission systems of the different bit rates.

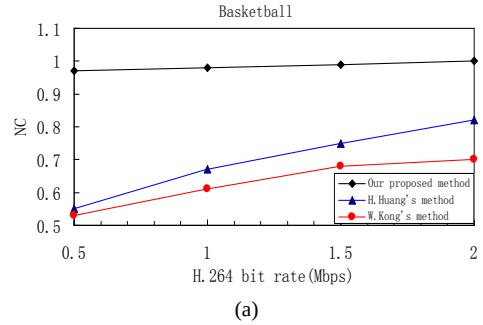


(a)

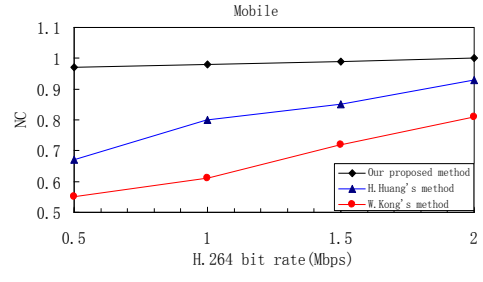


(b)

Fig. 8. Comparison results under the different MPEG-2 bit rates.



(a)



(b)

Fig. 9. Comparison results under the different H.264 bit rates.

Filtering and noise attacks: In order to study deeply the robustness of the information hiding system, we consider also various intentional or unintentional attacks, such as Gaussian noise (mean is 0 and variance is 0.05) and "pepper & salt" (density is 0.1) attacks, wiener filter attacks to demonstrate the performances of the system. Fig. 10-Fig. 13 present the experimental results of the watermark image recovered by three methods after various attacks. From these results, we can understand no matter what the attacks are, the NC values of the fingerprint image recovered from our proposed hiding system can still exceed 0.997, and the image can reconstruct with higher quality than the methods of Huang *et al.* and Kong *et al*. In other words, the information hiding system proposed in this paper have a better ability to resist various attacks. In this case, we take fully advantage of CS in which a signal can be retrieved with high probability by a relatively small number of measurements.

(a) (b) (c)

Fig. 10. Comparison result against wiener filtering attack for basketball video. (a) Our proposed method, NC=1; (b) H.Huang's method, NC=0.782; (c) W.Kong's method, NC=0.623.

(a) (b) (c)

Fig. 11. Comparison result against wiener filtering attack for Mobil video. (a) Our proposed method, NC=0.997; (b) H.Huang's method, NC=0.758; (c) W.Kong's method, NC=0.606.

(a) (b) (c)

Fig. 12. Comparison result against Gaussian noise attack for basketball video. (a) Our proposed method, NC=1; (b) H.Huang's method, NC=0.862; (c) W.Kong's method, NC=0.766.

(a) (b) (c)

Fig. 13. Comparison result against Gaussian noise attack for mobile video. (a) Our proposed method, NC=1; (b) H.Huang's method, NC=0.896; (c) W.Kong's method, NC=0.802.

V. Conclusions

In this paper, we have proposed a robust 3-D video watermark embedding scheme in DCT-CS domain. The watermark data is generated by measurement values of BCS which can express all features of the original watermark image and itself possess an encryption property from random elements of the CS measurement matrix. Since the hiding system proposed in the paper needs only to embed a smaller amount of the CS-watermark data into video signal, it is obvious that the proposed system can effectively resist compressions, noises, and filtering attacks and can maintain a good performance in transparency and robustness. On the

other hand, utilizing a priori knowledge of measurement matrix in CS domain, the system can also reconstruct higher quality the watermark image than ones of DCT-based and SVD-based. In further, we will research deeply information forensics by CS theory.

REFERENCES

[1] B. G. Haskell, A. Puri, and A. N. Netravali, *Digital Video: An Introduction to MPEG-2*, Chapman and Hill, 1997.

[2] J. J. Eggers, J. K. Su, and B. Girod, "A blind watermarking scheme based on structured codebooks," *IEE Colloquium (Digest)*, vol. 39, pp. 27-47, 2000.

[3] J. Chou, S. Pradhan, and K. Ramchandran, "A robust blind watermarking scheme based on distributed source coding principles," in *Proc. 8th ACM International Conference on Multimedia (ACM Multimedia*, New York: ACM Press, 2000, pp. 49−56.

[4] S. Biswas, R. Das, and M. Petriu, "An adaptive compressed MPEG-2 video watermarking scheme," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 5, pp. 1853-1861, Oct. 2005.

[5] M. Barni, F. Bartolini, and N. Checcacci, "Watermarking of mpeg-4 video objects," *IEEE Transactions on Mutimedia*, vol. 7, no. 1, pp. 23-32, Feb. 2005.

[6] A. Abdulfetah, X. Sun, and H. Yang, "Robust adaptive video watermarking scheme using visual models in DWT domain," *Information Technology Journal*, vol. 9, no. 7, pp. 1409-1414, June 2010.

[7] K. Chetan and K. Raghavendra, "DWT based blind digital video watermarking scheme for video authentication," *International Journal of Computer Applications*, vol. 4, no. 10, pp. 19-26, Aug. 2010.

[8] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489-509, 2006.

[9] D. L. Donoho and Y. Tsaig, "Extensions of compressed sensing," *Signal Processing*, 2006, vol. 86, no. 3, pp. 533-548.

[10] C. Zhao and W. Liu, "Block compressive sensing based image semi-fragile zero-watermarking algorithm," *Acta Automatica Sinica*, vol. 38, no. 4, pp. 609-617, April 2012.

[11] R. Calderbank, S. Jafarpour, and R. Schapire, *Compressed Learning: Universal Sparse Dimensionality deduction and Learning in the Measurement Domain*, preprint, 2009.

[12] D. Hsu, S. M. Kakade, J. Langford, and T. Zhang, "Multi-label prediction via compressed sensing," *Neural Information Processing Systems (NIPS)*, 2009.

[13] M. Davenport, P. Boufounos, M. Wakin, and R. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445-460, 2010.

[14] W. Lu, A. L. Varna, and M. Wu, "Security analysis for privacy preserving search for multimedia," in *Proc. IEEE 17th Inter. Conf. on Image Processing*, 2010.

[15] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "On the security and robustness of encryption via compressed sensing," in *Proc. IEEE Military Communications Conference*, 2008, pp.1040-1046.

[16] X. Zhang, Z. Qian, Y. Ren, and G. Feng, "Watermarking with flexible self-recovery quality based on compressive sensing and compositive reconstruction," *IEEE Transaction on Information Forensics and Security*, vol. 6, no. 4, pp. 1223-1232, 2011.

[17] Q. Wang, W. Zeng, and J. Tian, "Integrated secure watermark detection and privacy preserving storage in the compressive sensing domain," in *Proc. IEEE International Workshop on Information Forensics and Security*, November 2013, Guangzhou, China, pp. 67-72.

[18] T. Bianchi and A. Piva, "Secure watermarking for multimedia content protection: A review of its benefits and open issues," *IEEE Signal Processing Magazine*, vol. 30, pp. 87-96, 2013.

[19] D. Donoho, "Compressed sensing," *IEEE Transaction on Information Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

[20] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, issue. 2, pp. 21-30, Mar. 2008.

[21] E. J. Fowler, M. Sungkwang, and E. W. Tramel, "Multiscale block compressed sensing with smoothed projected landweber reconstruction," in *Proc. 19th European Signal Processing Conference (EUSIPCO 2011)*, Barcelona, Aug. 29-Sep. 2, 2011, pp. 564-568.

[22] H. Y. Huang, C. H. Yang, and W. H. Hsu, "A video watermarking technique based on pseudo-3-d dct and quantization index modulation," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 625-627, 2010.

[23] Y. S. Seo, W. G. Kim, Y. H. Huh, W. G. Oh, and C. J. Hwang, "QIM watermarking for image with tow adaptive quantization step-sizes," in *Proc. 9th Int. Conf. Advanced Communication Technology*, 2007, pp. 997-800.

[24] W. Kong, B. Yang, D. Wu, and X. Niu, "SVD based blind video watermarking algorithm," in *Proc. First Int. Conf. Innovative Computing, Information and Control*, 2006, pp. 265-268.

[25] H. Zhao, "Fingerprint system and dsp fast processing technology," *Research & Progress of Solid State Electronics*, vol. 24, no. 3, pp. 337-342, 2004.

Longfei Cai was born in Guangdong, China in 1976. He received the B.S degree from the School of Computer, South China Normal University, and the M.S. degree from School of Software Engineering, Huazhong University of Science and Technology. His research interests include computer process, and information security technology.

Huimin Zhao was born in Shanxi, China, in 1966. He received the B.S. and M.Sc. degrees in signal processing in 1992 and 1997 from Northwestern Polytechnical University, Xian, China, respectively. He received the Ph.D. degree in electrical engineering from the Sun Yat-Sen University in 2001. At present, he is a professor of the Guangdong Polytechnic Normal University. His research interests include image, video and information security technology.

Jun Cai received the B.S degree from Hunan normal universty, Changsha, China, the M.S degree from Jinan University, Guangzhou, China, and the Ph.D. degree from Sun Yat-Sen University, China in 2003, 2006 and 2012, respectively. He is currently an instructor with the School of Electronic and Information, Guangdong Polytechnic Normal University, Guangzhou, China. Complex network, traffic modeling and anomaly detection have been of particular interest over recent years.

Li Zhu was born in Guangxi, China in 1969. She received the B.S. degree in pressure processing from Northwestern Polytechnical University, China, in 1991, and the M.S. degree in engineering computer technology from Guangdong University of Technology, China, 2007. At present, she is a senior engineer of the Guangdong Polytechnic Normal University. Her research interests include multimedia processing, computer vision and information security technology.