A VIDEO WATERMARKING BASED ON 3-D COMPLEX WAVELET

Jingwei Wang, Xinbo Gao, Juanjuan Zhong

School of Electronic Engineering, Xidian University, Xi'an 710071, China
National Key Laboratory of ISN, Xidian University, Xi'an 710071, China

ABSTRACT

A video watermarking embedding and detection algorithm is proposed in complex wavelet transform (CWT) domain. The host video sequence is firstly segmented into shots. Then each shot is projected onto 3-D CWT domain. To achieve robustness while keeping imperceptibility, a perceptual mask derived from 3-D complex wavelet coefficients is introduced to weight the watermarks and the results is added back to the complex wavelet coefficients. Finally the inverse 3-D CWT is applied to obtain the watermarked video shots. The experimental results illustrate that the proposed algorithm has more advantages in remaining video qualities while keeping the same resistance to attacks over that in discrete wavelet transform (DWT) domain.

Index Terms— CWT, perceptual mask, DWT

1. INTRODUCTION

Digital watermark is a promising technique to protect data from illegal copying. The goal of watermarking is to embed unnoticeable information which represents the ownership of media contents. It has found widely applications in copyright protection, broadcast monitoring, data authentication *etc*. The ideal properties of a digital watermark consist of the imperceptibility and the robustness. That is, the watermarked video should retain the quality of the original video as closely as possible. In addition, the watermark should be robust to attacks, which can protect itself from being removed.

Watermarking techniques can be performed either in the spatial domain, or in the transform domain which is more robust. Consequently, most of research focus on the latter. A watermarking approach operating in 3-D discrete Fourier transform (DFT) domain has been proposed in [1]. A spread spectrum approach in 3-D Wavelet Transform (WT) domain, which first performs a 2-D spatial WT and then a 1-D temporal WT, has been presented in [2]. A video watermarking technique using the 3-D Discrete Cosine Transform (DCT) has been described in [3]. Furthermore, in order to make the embedded watermark imperceptible, a perceptual mask has been presented in [4].

Dual-tree CWT proposed by Kingsbury in [5] adopted two wavelet trees with real coefficients which were operated in parallel to achieve the real and imaginary parts of the complex wavelet coefficients. As a result, the CWT has a redundancy of 2:1 for 1-D signals. As for 3-D signals, the two trees operate on each dimension separately, so the redundancy becomes 8:1. Many researchers clarified that this redundancy must be modified before embedding [6][7]. However, the redundant information includes ampler directional details and large embedding capacity for watermark than that of the DWT, which can help us embed watermark in a more imperceptible and robust way. Therefore, a new watermark embedding method in 3-D CWT domain is proposed in this paper. By adopting an effective embedding mask, one can process the redundancy in a far more imperceptible way rather than modify the complex wavelet coefficients.

2. WATERMARK EMBEDDING USING PERCEPTUAL WEIGHING AND DETECTION

In the proposed watermarking method, the video sequence is firstly segmented into non-overlapping spatio-temporal plot units (shots). Each shot is processed as follows, which is also shown in Fig.1.

Step 1: The shot is projected onto the one level 3-D CWT domain, and then we obtain 8 main subbands and their subsidiary subbands. As each main subband has 7 susidiary subbands, there are 56 subbands in total;

Step 2: Perceptual masks are evaluated from the coefficients of the 56 subsidiary subbands in CWT domain;

Step 3: The watermark to be embedded is got by multiplying the original watermark sequence with the perceptual mask and the weight of the original watermark;

Step 4: The watermark is embedded into the 56 subsidiary subbands;

Step 5: The 3-D ICWT is performed to realize the transformation from CWT domain to spatio-temporal domain, and then the watermarked video shot is obtained.

Finally, all the watermarked video shots are assembled to form the watermarked video sequence.

This work was supported by the Key Project of Chinese Ministry of Education (No.104173), the Program for New Century Excellent Talents in University (NCET-04-0948), and the Program for Changjiang Scholars and Innovative Research Team in University, China.

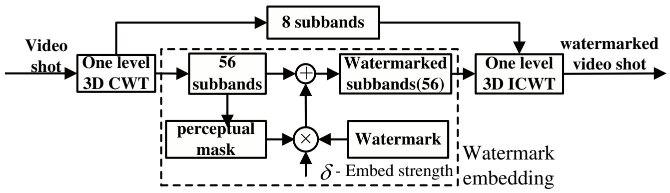


Fig. 1. The proposed watermarking approach using perceptual mask in 3-D CWT domain.

In the following, the proposed watermarking method will be discussed in details.

2.1. Watermark Embedding

By evaluating the variance on the 56 subbands, we can weight with higher value in the area where objects are fast-moving, which is insensitive to eyes. To reduce the computational complexity, the mask is evaluated on the main subbands and some corrective factors are applied to weight the corresponding subsidiary subbands. Specifically, a 3-D sliding window of size $2 \times 2 \times 2$ is applied on the main subbands to evaluate its local variance $\sigma_m(\mathbf{n}), \mathbf{n} = (n_1, n_2, n_3)$, which is then normalized. The masks of its 7 subsidiary subbands are obtained by using some corrective factors to modify the $\sigma_m(\mathbf{n})$. As specified in the following, α is the corresponding order number of its 7 subsidiary subbands.

$$\sigma_s^{(\alpha)}(\mathbf{n}) = \sigma_m(\mathbf{n}) \cdot \begin{cases} \sqrt{2}, \alpha = 3, 5, 6, 7 \\ 1, \alpha = 1, 2, 4 \end{cases} \tag{1}$$

Let $u(\mathbf{n})$ be a shot composed by N_3 frames of dimension $N_1 \times N_2$ and $v(\mathbf{n})$ be its corresponding form after 1-level 3-D CWT. $v_s^{(\alpha)}(\mathbf{n})$ ($\alpha = 1, \cdots, 7$) denotes the 7 subsidiary subbands in CWT domain which belong to one main complex wavelet subband. Watermark is a pseudo-random binary sequence with values in the set {-1,1}. To guarantee the imperceptibility of watermark as well as robustness, perceptual weighing is performed. As a result, the embedding strength of watermark is inflexible and depends on the masking characteristic of human visual system.

In more details, the embedding is performed according to the following rule:

$$\widetilde{v}_s^{(\alpha)}(\mathbf{n}) = v_s^{(\alpha)}(\mathbf{n}) + \delta \cdot m(\mathbf{n}) \cdot \sigma_s^{(\alpha)}(\mathbf{n}) \tag{2}$$

where δ represents the embedding strength, $m(\mathbf{n})$ denotes the watermark sequence rearranged in 3-D, and $\sigma_s^{(\alpha)}(\mathbf{n})$ is the weighing function which is obtained by Eq.(1).

2.2. Watermark Detection

The correlation parameter ρ between the watermarked 3-D complex wavelet coefficients and the watermark sequence to be tested can be expressed as follows:

$$\rho = \frac{1}{C} \sum_{\alpha=1}^{7} \sum_{\mathbf{n} \in \Omega} \widetilde{v}_s^{(\alpha)}(\mathbf{n}) \cdot m(\mathbf{n}) \tag{3}$$

where $\Omega = [0, N_1/2] \times [0, N_2/2] \times [0, N_3/2]$, and C represents the normalization term. When we determine whether the watermark is present or not, the correlation ρ need to be compared with a threshold T. With the Neyman-Person criterion [8], the probability of false detection is

$$P_f = \frac{1}{2} erfc\left(T/\sqrt{2\sigma_\rho^2}\right) \tag{4}$$

where σ_ρ^2 denotes the variance of ρ, which is assumed conforming the Normal distribution. By requiring a probability of false detection $P_f \leq 10^{-8}$, a threshold $T = 3.968\sqrt{2\sigma_\rho^2}$ is obtained. As detailed in [8], the correlation variance σ_ρ^2 is obtained by means of the following estimator

$$\sigma_\rho^2 \approx \frac{1}{C^2} \sum_{\alpha=1}^{7} \sum_{\mathbf{n} \in \Omega} \left(\widetilde{v}_s^{(\alpha)}(\mathbf{n})\right)^2 \tag{5}$$

3. EXPERIMENTAL RESULTS AND ANALYSIS

In order to verify the effectiveness of the proposed watermarking approach in 3-D CWT domain (called CWT algorithm in the following), we conduct three groups of experiments to compare with two methods in DWT domain (called DWT algorithms in the following). The first one is used to test the detection performance of the CWT algorithm, and the other two experiments with different methods are employed to compare the performance of perceptibility and resistance to attacks. The experiments are performed on three uncompressed videos separately: the *football*, *tennis*, and *coast* sequence, which are composed by $N_3 = 360$ frames with the size of $N_1 \times N_2 = 240 \times 180$ pixels. Every video is partitioned into several shots with fixed length of 72 frames. The size of the embedding watermark is $45 \times 60 \times 36$ and the embedding strength δ is 0.6 in the CWT algorithm. The resistance performance to three kinds of attacks, namely frame averaging attack, frame cutting attack and Gaussian noise attack, have been tested after normalizing the correlation ρ of different methods.

3.1. Watermark Detection Results

Fig.2 shows the detector response expressed in Eq.(3) to 200 different marks with the embedded mark at position 100. It is evident that the real watermark is the only one that gives a detection value higher than the threshold, which also means that the CWT algorithm is effective for blind detection.

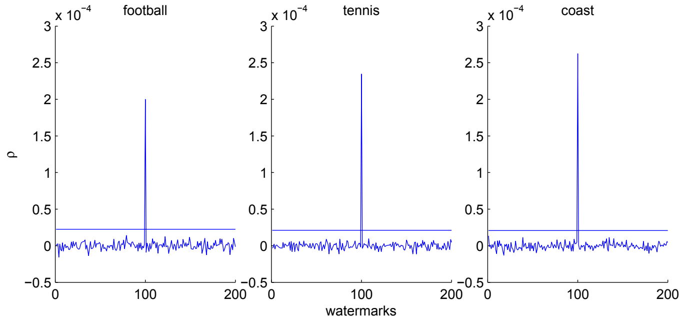


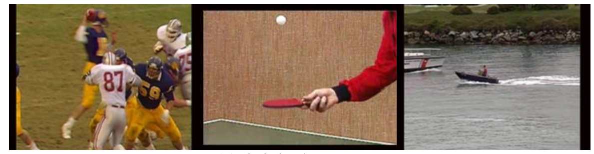
Fig. 2. Detection results of the CWT algorithm for different sequences.

3.2. Comparison under the Same Parameter

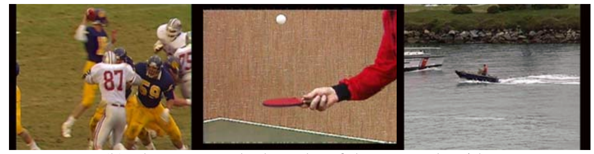
In this experiment, the DWT algorithm has the same embedded areas and the embed strength as the CWT algorithm. After 1-level 3-DWT, 7 subbands (except the low frequency subband-LLL) are watermarked with the perceptual mask described in Eq.(1). However, it must be pointed out that although we embed watermark in the same area in the CWT and DWT algorithm, the number of embeddable subbands in CWT domain is 8 times than that in DWT domain. The performance can be evaluated from two aspects. The quality of the watermark embedding video can be analyzed by the comparison of the PSNR and the detection results with different attacks reflect the resistance to attacks.

(1) *Perceptibility*: Although the video quality appears no different in Fig.3(a)-(c), from the comparison of PSNR in Fig.3(d), it is found that the PSNR of the CWT algorithm is higher than that of the DWT algorithm, which means CWT algorithm can achieve higher perceptual vision qualities.

(2) *Resistance to attacks*: The detection results with different number of frame averaging, frame cutting and different PSNR with Gaussian noise are shown in Fig.4(a)-(c) respectively. The more averaged frames are, the closer to still sequences the video shot is. Due to the fact that movement is one of the main areas we can hide our watermark with the 3-D method, the detection results decline with the increase of the number of averaged frames. It is also evident that the CWT algorithm shows a bit strong at resisting averaging attack, while it gains apparent advantage over the DWT algorithm in the case of cutting attack. The higher the PSNR is, the weaker the Gaussian noise added to every frame in a sequence is. As a result, the detection results are worse with lower PSNR. As CWT has more directional information than DWT, the perceivable distorted area in CWT algorithm is smaller than that in DWT algorithm. Therefore, the CWT algorithm is more sensitive to noise than the DWT algorithm.

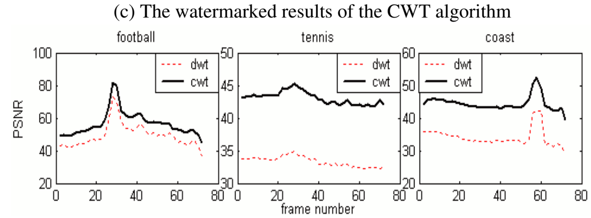


(a) Original sequences



(b) The watermarked results of the DWT algorithm

(c) The watermarked results of the CWT algorithm



(d) Comparison of PSNR between different sequence

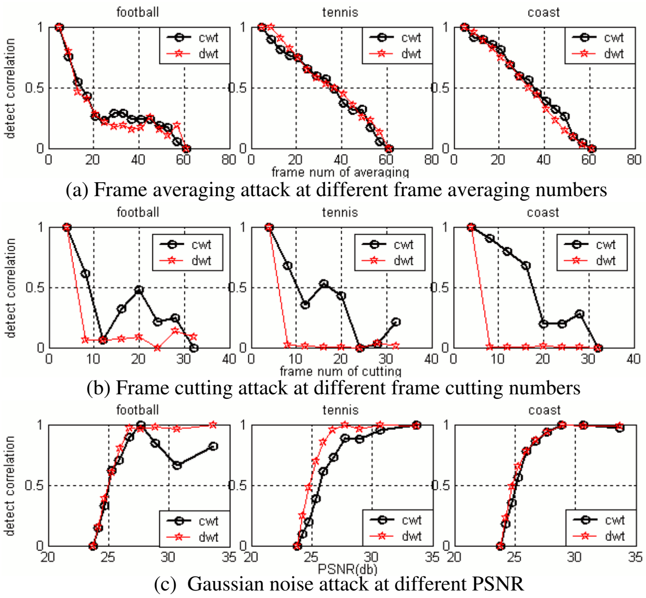
Fig. 3. Watermark embedding sequences and comparison of PSNR between them.

3.3. Comparison under the Same PSNR

For easier comparison, two ways are applied on the DWT algorithm to achieve the same PSNR results as the CWT algorithm. One is adjusting the embedding strength of the DWT algorithm, and the other is resizing the embedding areas of the DWT algorithm. For simplicity, the two methods are denoted by DWT1 method and DWT2 method respectively.

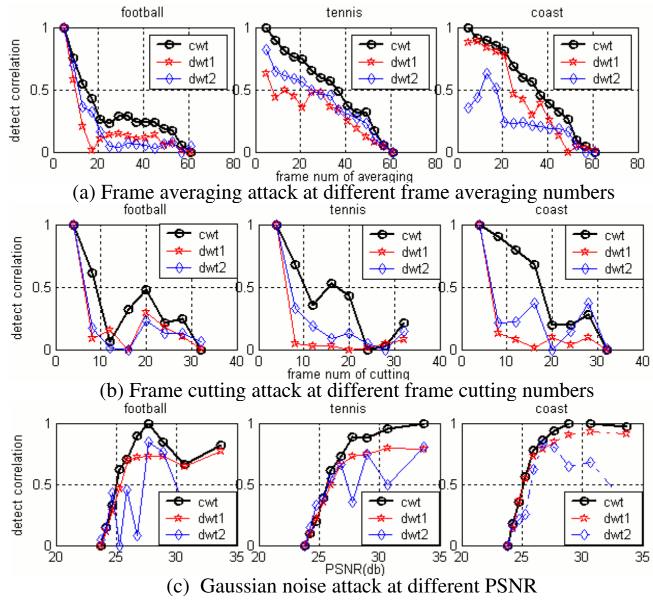
(1) *Perceptibility*: Comparison of the embedding strength and repeat times of watermark between the CWT algorithm and these two DWT methods are showed in Table 1. It is found that the embedding strength of the DWT1 method must be lower than 0.6 and the repeat times of watermark decrease to keep the same PSNR with the CWT algorithm, due to the fact that both the embedding size in one subband and the number of embedded subbands in the DWT2 method must shrink. However, due to the existence of perceptual mask, the degree of decreasing is not constant. If the repeat time of watermark embedded is lower than 1, it means that in DWT domain, one can embed watermark only in one subband which is watermarked partially.

(2) *Resistance to attacks*: Experimental results of frame averaging, frame cutting, and Gaussian noise are shown in Fig.5(a)-(c) respectively. It is clear that the CWT algorithm shows more superiority to attack resistance than DWT1 and DWT2 methods.



(a) Frame averaging attack at different frame averaging numbers
(b) Frame cutting attack at different frame cutting numbers
(c) Gaussian noise attack at different PSNR

Fig. 4. Comparison of detection results with different attack.



(a) Frame averaging attack at different frame averaging numbers
(b) Frame cutting attack at different frame cutting numbers
(c) Gaussian noise attack at different PSNR

Fig. 5. Comparison of detection results with different attack.

Table 1. Comparison of embedding strength and repeat times between the CWT algorithm and the DWT algorithm.

method	football	tennis	coast
CWT(embedding strength)	0.6	0.6	0.6
DWT1	0.3	0.18	0.18
CWT(repeat times)	56	56	56
DWT2	1	0.85	0.5

4. CONCLUSIONS

In this paper, we present a blind detecting video watermarking algorithm in CWT domain. During the embedding, a perceptual mask derived from 3-D complex wavelet coefficients is exploited. With comparison between the CWT and DWT algorithm, it is illustrated that the proposed method has higher PSNR, larger embedding capacity and noticeable stronger resistance to attack than the DWT algorithms. Furthermore, the CWT algorithm has more repeat times of embedded watermark than methods in DWT domain. Consequently, it is more effective when spread spectrum is needed. Although experimental results show that the proposed algorithm has advantages in detection performance, we need to further optimize the video watermark embedding and detection procedures so that it can be performed in real time.

5. REFERENCES

[1] F. Deguillaume, G. Csurka, J. O'Ruanaidh, and T. Pun, "Robust 3D DFT video watermarking," *Proc. SPIE, Security and Watermarking of Multimedia Contents II*, vol.3971, pp.997-1000, 2000.

[2] Ying Li, Xinbo Gao, and Hongbing Ji, "A 3D wavelet based spatial-temporal approach for video watermarking," *Proceedings of International Conf. on Computational Intelligence and Multimedia Applications*, pp.260-265, Xi'an, China, September 27-30, 2003.

[3] J. H. Lim, D.J. Kim, H.T. Kim, and C.S. Won, "Digtal video watermarking using 3D-DCT and Intra-Cubic correlation,"*Proc. SPIE, Security and Watermarking of Multimedia Contents III*, vol.4314, pp.54-72, 2001.

[4] Patrizio Campisi, and Alessandro Neri, "Video watermarking in the 3D-DWT domain using perceptual masking," *Proceedings of IEEE International Conference on Image Processing*, vol.1, pp.997-1000, 2005.

[5] N. G. Kingsbury, "Image processing with complex wavelets," *Phil. Trans. Royal Society London*, vol.A. 357, pp.2543-2560, September 1999.

[6] P. Loo, and N. G. Kingsbury, "Digital watermarking with complex wavelets," *Proc IEE Colloquium on Secure Images and Image Authentication, IEE, London*, vol.10, pp.1-7, 10 April, 2000.

[7] P. Loo, "Digital watermarking using complex wavelets," *PhD. thesis*, University of Cambridge, March 2002.

[8] M. barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. on Im. Proc.*, vol.10(5), pp.783-791, May 2001.