Video Watermarking Scheme against Geometrical Distortions

Hu Shanshan
The Dept. of Aeronautics Ammunition Engineering
The First Aeronautics College of Air Force
Xinyang, Henan, China
e-mail: fromdanger123@126.com

Liu Fang
The Dept. of Aeronautics Ammunition Engineering
The First Aeronautics College of Air Force
Xinyang, Henan, China
e-mail: lyxlf2733@sina.com

***Abstract*—A new video watermarking scheme against geometric distortions is proposed, which is most suitable for DCT-encoded compressed video, like MPEG-2, MPEG-4, H.264, etc. To lower the computational cost, 2-D DFT coefficients are constructed directly from the block DCTs by exploiting a fast inter-transformation. Then a watermark is embedded in a RST (rotation, scaling, and translation) invariant domain generated by performing the Fourier-Mellin transform (FMT) on the 2-D DFT magnitudes. For resisting different video format conversion, the watermark detection is performed in the spatial domain. The experimental results show that the proposed scheme is transparent and robust to typical signal processing attacks, lots of commonly geometric distortions including scaling, rotation and some combination of several attacks, and frame dropping.**

KeyWord—video watermarking; fourier-mellin transform; geometric distortions

I. INTRODUCTION

Although research on digital watermarking has made a great progress, the geometric attack is still a major threat to the watermarking applications. The geometric distortions desynchronize the watermark information while preserving the visual quality, and pose problems for many existing watermarking detectors. Moreover, watermark synchronization is more difficult to handle in the applications of blind watermarking detection.

In general, existing image watermarking methods resisting against geometrical distortions can be categorized into three classes: (1) synchronization recovery techniques; (2) geometrically invariant methods; (3) feature-based local watermarking. The first class can be divided into image registration based scheme [1], exhaustive search based scheme [2], template based scheme [3] and inserting a periodic watermark pattern [4]. The first one is a non-blind watermarking method. The computation of exhaustive search is large and its applicability is limited while the template and periodic watermark pattern are easily removed by collusion attacks.

In the second class, O'Ruanaidh [5] has proposed a RST watermarking method in which the FMT is performed on the DFT magnitudes of the original image. Consequently, the original image is transformed into a truly RST invariant domain. A watermark embedded in this domain could be successfully extracted under RST attacks.

The feature-based local watermarking method [6] can resist local geometric distortions including cropping attacks. However, its computational complexity is very high because of the operation of feature point extraction, area partition and normalization. It is impractical for video watermarking.

Most compressed video data are stored as block DCT coefficients and motion vectors. Any geometric processing leads to repartition of these blocks, which in turn totally changes the coefficients and motion vectors. This implies that conventional block-based watermarking techniques are vulnerable to any geometrical attacks. Therefore, how to construct a geometric invariant domain directly from the compressed domain becomes an essential process to solve the problem of both resisting geometric attacks and reducing computing time.

In this paper, a robust video watermarking scheme using FMT is proposed for DCT-encoded compressed videos. In Sec. 2, introduce the fast inter-transformation between block DCTs and 2-D DFT and properties of 2D-DFT and FMT. A watermarking scheme is proposed in Sec. 3. The experimental results are shown in Sec. 4. Sec. 5 concludes.

II. BACKGROUND

Shift invariance of the DFT, rotation and scaling invariance of the FMT are employed to produce an invariant domain. A fast inter-transformation between block DCT coefficients and 2D-DFT is described in Section A, and the introduction to FMT is given in Section C.

A. Fast Inter-Transformation between Block DCTs and 2-D DFT

Our proposed watermarking scheme starts in DFT domain. Every video frame is an orderly set of block DCT coefficients. In [7], Brynmor *et al.* fully revealed a concise and linear relationship between all linear, invertible transforms and any separable sub-block geometry. Hence, the 2-D DFT coefficients can be directly obtained from the block DCTs without involving the IDCT, which results in low computational cost.

A frame I with a size of $LS \times MS$ can be divided into $L \times M$ blocks with a size of $S \times S$, one of which is denoted as b_{ij}. The DCT coefficients C_{ij} of blocks can be expressed as

$$C_{ij} = B_1 \times b_{ij} \times B_1^T \tag{1}$$

where B_1 is the transform matrix of block DCT. B_1 is an orthogonal matrix. So the whole image can be expressed as

$$I=\begin{bmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & B_1 \end{bmatrix}^{-1}_{LS\times LS} \times \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1M} \\ C_{21} & C_{22} & \cdots & C_{2M} \\ \vdots & \vdots & \cdots & \vdots \\ C_{L1} & C_{L2} & \cdots & C_{LM} \end{bmatrix} \times \begin{bmatrix} B_1^T & 0 & \cdots & 0 \\ 0 & B_1^T & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & B_1^T \end{bmatrix}^{-1}_{MS\times MS} \quad (2)$$

The matrices on the right of (2) are denoted as B_4, C_{part} and B_5 respectively.

On the other hand, the 2-D DFT coefficients of image I can directly be computed using (3).

$$F = B_2 \times I \times B_3^T \quad (3)$$

where B_2 and B_3 are defined as

$$B_2(h,l)=\sqrt{\frac{1}{L}}\exp\left(\frac{-j2\pi hl}{L}\right), B_3(h,l)=\sqrt{\frac{1}{M}}\exp\left(\frac{-j2\pi hl}{M}\right) \quad (4)$$

Substituting I into (3) with (2), F can be obtained using this expression:

$$F = A_1 \times C_{part} \times A_2 \quad (5)$$

where $A_1 = B_2 \times B_4$, $A_2 = B_5 \times B_3^T$. Actually, for frames in the same video, they are constants and can be calculated in advance. Significant savings on computing cost can be expected.

B. Properties of 2D-DFT

Assuming that $F(\lambda,\tau)$ is the 2-D DFT of the image $f(x,y)$, Suppose that the RST parameters are ϕ, c and (x_0,y_0) respectively. $F^{'}(\lambda,\tau)$ is the DFT of the transformed image:

$$\left|F^{'}(\lambda,\tau)\right| = |c|^{-2}\left|F(c^{-1}(\lambda\cos\phi+\tau\sin\phi), c^{-1}(-\lambda\sin\phi+\tau\cos\phi))\right| \quad (6)$$

This equation is independent of the translational parameters, which is the translation property of the Fourier transform.

C. Fourier-Mellin Transform

The FMT is a log-polar mapping (LPM) followed by a DFT. If the sampling rates, N_ρ and N_θ, of LPM on the radial and angular direction are constant, the properties of LPM include scaling linearity and rotation linearity [8].

If applying the LPM to the Fourier magnitude of an image on the log-polar plane (ρ,θ), (6) will be rewritten as:

$$\left|F^{'}(\rho,\theta)\right| = |c|^{-2}\left|F(\rho-\ln c, \theta-\phi)\right| \quad (7)$$

It demonstrates that that image scaling results is a translational shift of $\ln c$ along the log-radius ρ axis, that image rotation results in a cyclical shift of θ along the angle θ axis, and that image translation has no effects in the LPM domain. According to the translation property of DFT, after applying DFT to both sides of the Eq. (6), the Fourier magnitude of the two LPM is related by:

$$\left|M^{'}(\zeta_\rho,\zeta_\theta)\right| = \left|M(\zeta_\rho,\zeta_\theta)\right| \quad (8)$$

This equation demonstrates that the amplitude of FMT is invariant to scaling, rotation and transform.

III. THE PROPOSED SCHEME

A. Watermark Embedding

First of all, the video should be partially decoded to obtain the 2-D block DCT coefficients of the luminance of frames. Our watermarking embedding process is shown in Fig. 1. Firstly, calculate 2-D DFT from the block DCTs by using (5). Next, compute RST invariant coefficients $\left|M(\zeta_\rho,\zeta_\theta)\right|$ via FMT.

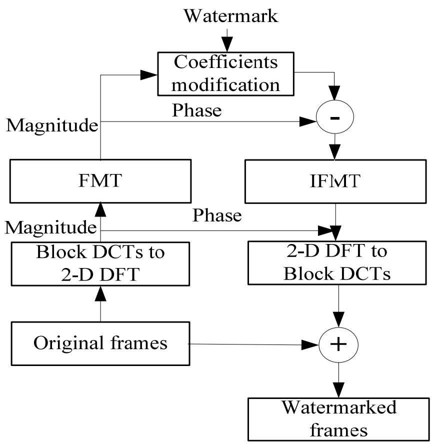

Figure 1. Framework for watermark embedding

The generated watermark W is embedded in N points of the upper half part of $\left|M(\zeta_\rho,\zeta_\theta)\right|$ which satisfy the frequency bands $f_1 < \sqrt{\zeta_\rho^2+\zeta_\theta^2} < f_2$. These N points are shuffled randomly and divided into l groups $G_j (j=1,2,\cdots,l)$. Each group has $n=N/l$ FMT magnitudes and it is further divided into two equal-size sub-groups denoted as G_j^A and G_j^B, and one watermark bit can be correspondingly embed into the group G_j. Let P_j^A and P_j^B denote the sums of the coefficients in group G_j^A and G_j^B respectively:

$$P_j^A = \sum_{(\zeta_{\rho k},\zeta_{\theta k})\in G_j^A}\left|M(\zeta_{\rho k},\zeta_{\theta k})\right|, P_j^B = \sum_{(\zeta_{\rho k},\zeta_{\theta k})\in G_j^B}\left|M(\zeta_{\rho k},\zeta_{\theta k})\right|, \quad (9)$$

Define the ratio between G_{ij}^A and G_{ij}^B as

$$D_j = \frac{P_j^A - P_j^B}{P_j^A + P_j^B}, \quad (10)$$

Modulate D_j by using (12) to meet the relation in (11) for embedding watermark bits:

$$D_j \times w_j \ge T, \quad (11)$$

where T is a threshold for robustness. If the relation (11) does not hold, P_j^A and P_j^B must be adjusted by modifying the coefficients $\left|M(\zeta_\rho,\zeta_\theta)\right|$ in the group:

$$\left|M_j^{w'}(\zeta_{\rho k},\zeta_{\theta k})\right| = \begin{cases} (1+{}^{\operatorname{sgn}(w_j)}\!/\!_{P_j^A})\times\left|M_j(\zeta_{\rho k},\zeta_{\theta k})\right|, & (\zeta_{\rho k},\zeta_{\theta k})\in G_j^A \\ (1-{}^{\operatorname{sgn}(w_j)}\!/\!_{P_j^B})\times\left|M_j(\zeta_{\rho k},\zeta_{\theta k})\right|, & (\zeta_{\rho k},\zeta_{\theta k})\in G_j^B \end{cases}, \quad (12)$$

where α_j is given by (13):

$$\alpha_j = \frac{T\times(P_j^A+P_j^B)-\operatorname{sgn}(w_j)(P_j^A-P_j^B)}{2}, \quad (13)$$

Subsequently, shuffle $\left|M^{w'}(\zeta_\rho,\zeta_\theta)\right|$ back to their original locations. Then modify the low half part of the middle frequency band to remain the symmetry to maintain the symmetry to 2-D DFT DC coefficient.

Furthermore, the differences between the modulated and the original FMT magnitudes are inversely transformed to the differences of 2-D DFT magnitudes, and then to the differences of block DCT coefficients, which are added to the original block DCT coefficients in one frame afterwards. The aim of this process is to lessen the visual degradation.

B. Watermark Detection

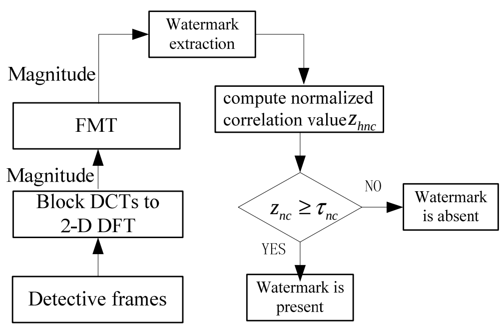

Figure 2. Framework for watermark detection

In our proposed scheme, watermark detection is performed in the spatial domain along with the decoded video playing, which is shown in Fig. 2. Therefore, watermark extraction is independent of the formats of video coding.

The watermark detection is similar to the embedding process. First, the 2-D DFT is calculated in the spatial domain. Afterwards, the RST invariant domain is generated by FMT. The magnitude of the upper half at the same middle frequency with watermark-embedding are selected, shuffled randomly and grouped. Then all the sums $P_j^A{}'$ and $P_j^B{}'$ are computed. The detected watermark bits can be blindly determined by Eq. (14).

$$W_j^e = \operatorname{sgn}(P_j^A{}'-P_j^B{}') = \begin{cases} +1, & if\ P_j^A{}'-P_j^B{}' \ge 0 \\ -1, & if\ P_j^A{}'-P_j^B{}' < 0 \end{cases}, \quad (14)$$

Finally, compute the normalized correlation value z_{nc} between W and W^e:

$$z_{nc} = \frac{W \bullet W^e}{|W|\times|W^e|} \quad (15)$$

where $|W|$ is the length of W. If z_{nc} is smaller than the threshold τ_{nc}, there is no watermark hidden in the region.

IV. Experiments And Discuss

In the experiments, the "Flower-Garden" and "Carriage" sequences are selected as test sequences. The sampling rates, N_ρ and N_θ for LPM are 300 and 300 respectively. And the watermark embedding bands f_1 and f_2 are 60 and 80, respectively. The threshold T is set at 0.4 and length N is set at 640. The number n of neighboring coefficients in one group is set at 8. The detection threshold τ_{nc} is chosen as 0.5.

A. Assessment of Perceptual Quality

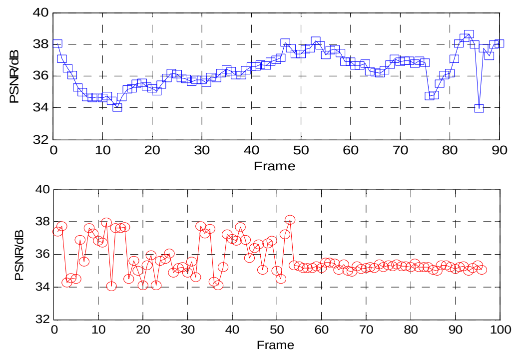

Figure 3. The PSNR curve of watermarked test sequences

The visual quality can be assessed objectively by measuring the Peak Signal-to-Noise Ratio (PSNR) of watermarked frames compared to original frames. The results are illustrated in Fig. 3. The average PSNR values of watermarked video frames are higher than 36dB. Perceptually, the original video and the watermarked video are visually indistinguishable. This implied that the watermarking scheme is able to achieve video transparence.

B. Estimation of Robustness

In practice, watermarked content will be subjected to a variety of distortions before reaching the detector. Watermarking schemes are designed to survive legitimate and everyday usage of content and malicious attacks, e.g. signal processing attacks and geometric distortions. The average normalized correlation value z_{mean} of test video sequences can be used to estimate the robustness of the watermark. The high z_{mean} is, the more robust the watermark is. Therefore, should consider the watermarking scheme to be robust if z_{mean} is higher than the threshold τ_{nc}.

The experimental results of signal processing attacks are shown TABLE Ⅰ. It is clear from these results that the algorithm is very robust against these attacks.

TABLE I. ROBUSTNESS VERSUS SIGNAL PROCESSING ATTACKS

Signal processing attacks	z_{mean}	
	Flower-garden	*Carriage*
Median filter 2×2	1.0	1.0
Median filter 3×3	1.0	1.0
Gaussian filter 3×3	1.0	1.0
Gaussian noise (δ =4)	1.0	1.0
Gaussian noise (δ =10)	1.0	0.916
Gaussian noise (δ =20)	0.875	0.833
Video bit rate reduction(5 Mb/s)	1.0	0.916
Video bit rate reduction(3 Mb/s)	0.916	0.875
Video bit rate reduction(1 Mb/s)	0.833	0.792

The test results for some general geometric distortions are shown in TABLE Ⅱ It shows that our algorithm can successfully resist some local geometric attacks, such as random removal of some rows and columns, which is referred to as jitter attack. It is also shown that the proposed algorithm successfully resists scaling, linear geometric transforms and shearing. The values of z_{mean} are higher than 0.79 roughly. Finally, in terms of rotation + cropping and rotation + scaling + cropping, the strength of cropping and scaling increase in proportion to the degree of rotation and hence the z_{mean} decrease. Our algorithm remains robust when the rotational angle is smaller than 20 degree. However, it is sensitive to cropping and the respect ratio changes. Since the watermark is embedded in frames in each GOP repeatedly, it can be detected successfully even if only several frames are left. That means the frame dropping attack is not a real threat to this proposed method.

TABLE II. ROBUST VERSUS GEOMETRICAL ATTACKS

Geometric distortion attack	z_{mean}	
	Flower-garden	*Carriage*
Row 1 and column 5 removal	1.0	1.0
Row 5 and column 17 removal	0.833	0.916
Random bending	0.667	0.68
Shearing-x-0%-y-5%	0.916	0.958
Shearing-x-5%-y-5%	0.791	0.833
Rotation 5°+ cropping	0.816	0.858
Rotation 10°+ cropping	0.733	0.746
Rotation 20°+ cropping	0.624	0.633
Rotation 2°+ scaling + cropping	0.833	0.875
Rotation 5°+ scaling + cropping	0.754	0.791
Rotation 10°+ scaling + cropping	0.64	0.667
Rotation 20°+ scaling + cropping	0.52	0.6
Scaling 0.5×	1.0	1.0
Scaling 1.5×	1.0	1.0
Scaling 2.0×	1.0	1.0
Linear[1.007,0.010;0.010,1.012]	1.0	1.0
Linear[1.013,0.008;0.011,1.008]	1.0	1.0

V. CONCLUSION

Targeting the problem of resisting geometric attacks, we proposed a new video watermarking scheme with high robustness. In the watermark embedding, with fast inter-transformation between block DCT coefficients and 2-D DFT employed, the computational cost is reduced significantly. And then the 2-D DFT magnitude is transformed into RST invariant domain by using FMT. A watermark is embedded in this domain. The watermark detection is performed in spatial domain along with video playing, which makes it insensitive to video format conversions. The experimental results show that this scheme is transparent and robust to typical signal processing attacks, lots of commonly geometric distortions and frame dropping.

REFERENCES

[1] Lai Chao-Hung, Wu Jiunn-Lin, “Robust image watermarking against local geometric attacks using multiscale block matching method,” Journal of Visual Communication and Image Representation, vol. 20, no. 6, 2009, pp. 377-388.

[2] S. Baudry, P. Nguyen, H. Maitre, “Estimation of geometric distortions in digital watermarking,” Proc. IEEE International Conference on Image Processing, vol. 2, 2002, pp. 885-888.

[3] Motallebi, F., Aghaeinia, H., “RST invariant wavelet-based image watermarking using template matching techniques,” Proc. 1st International Congress on Image and Signal Processing, vol. 5, 2008, pp. 720-724.

[4] D. Delannay, B. Macq, “Generalized 2-D cyclic patterns for secret watermark generation,” Proc. International Conference on Image Processing, vol. 2, 2000, pp. 77-79.

[5] Jjko O'ruanaidh, T. Pun. “Rotation, scale and translation invariant spread spectrum digital image watermarking,” Signal Processing, vol. 66, no.3, 1998, pp. 303-317.

[6] Hae-Yeoun Lee, Hyungshin Kim, Heung-Kyu Lee, “Robust image watermarking using local invariant features,” Optical Engineering, vol. 45, no. 3, 2006, pp. 037002(1-11).

[7] Brynmor J. Davis, S. Hamid Nawab, “The Relationship of Transform Coefficients for Differing Transforms and/or Differing Subblock Sizes,” Proc. IEEE Transaction on Signal Processing, vol.52, no.5, 2004, pp. 1458-1461.