A New Wavelet-Based Digital Watermarking for Video

Xiamu Niu, Shenghe Sun

P.O.Box 339, Department of Automatic Test and Control,

Harbin Institute of Technology (150001),

Harbin, P.R China

niuxiamu@public.hr.hl.cn

***Abstract*-- This paper proposes a method of embedding a digital watermark image in video. In the watermarking, the decomposed watermark image with different resolution is embedded in the corresponding resolution of the decomposed video by means of multiresolution signal decomposing. The experimental results show that the proposed method is robust against the attack of frame dropping, averaging and lossy compression.**

***Index terms*-- Video watermarking, Discrete Wavelet Transform, Multimedia security**

I. INTRODUCTION

Recently there has been great interest in watermarking techniques for protecting the copyright ownership of multimedia data distributing in Internet. Considering large amounts of data and inherent redundancy between frames, video watermarking poses many problems. On one hand, using the identical watermark for each frame, the attacker could collude with frames from completely different scenes [1]. Using independent watermark for each frame, on the other hand, the attacker could take advantage of motionless regions in successive video frames to remove the watermark by comparing and averaging the frames statistically [1][2]. By multiresolution temporal representation of video sequence [2], two types of watermark (identical and independent watermark) are used for embedding them in motionless and motion regions of the video respectively. Since the multiresolution temporal representation is along the temporal axis of the video, the robustness of watermarking for each frame (especially for I-frame) should be considered in addition for the sake of surviving Moving Picture Expert Group (MPEG) video lossy compression. Also, for large amounts of data in video, the information of watermark should be correlative with each other as well as abundant in quantity. This paper proposes a new method of embedding a digital watermark image in the video. By taking advantage of Two-Dimensional (2-D) multiresolution signal decomposing and One-Dimensional (1-D) multiresolution temporal representation of video sequence, both the watermark image and the original video are decomposed into multiresolution representations with different structure. The decomposed watermark image of different resolution is embedded in the corresponding resolution of the decomposed original video. As the 1-D multiresolution temporal representation of the video is only for the temporal axis of the video, each frame along spatial axis is decomposed into 2-D multiresolution representations for watermarking the spatial detail of the frame as well as the motion and motionless regions of the video. By still image decomposition technique, the gray-level digital watermark is decomposed into a series of binary digital images. The decomposed binary digital images, which are correlative with each other and also abundant in quantity, are as series of independent watermarks. Experimental results show that the proposed techniques are robust enough against frame dropping, averaging and MPEG lossy compression.

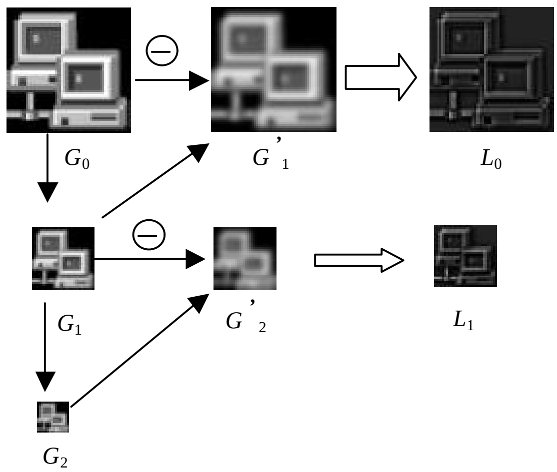


Fig.1 The first three levels of the Pyramid Structure of an image

II. DECOMPOSITION OF 2-D VISUALLY RECOGNIZABLE WATERMARK

2.1 Multiresolution hierarchical structure of watermark image

Considering a watermark, in origin, should be a visually recognizable pattern, we adopt 2-D gray-level digital image(64×64) as a watermark. In order to embed the watermark invisibly, the watermark information should adapt itself to the detail of original video. Hence, we decompose the watermark into a multiresolution hierarchical structure of images L_0, L_1 and G_2 by the resolution-reduction method described in literature[3] as shown in Fig.1, where G_0 (64×64), the watermark image, is level 0 of the Gaussian pyramids, G_1 (32×32) is level 1 of the Gaussian pyramids, G_2 (16×16) is level 2 of the Gaussian pyramids, G'_1 (64×64) and G'_2 (32×32) are the Gaussian Pyramid interpolation results of G_1 (32×32) and G_2 (16×16) respectively, L_0(64×64) is level 0 of the Lapacian pyramids, and L_1(32×32) is level 1 of the Lapacian pyramids.

2.2 Decomposing the gray-level watermark image into a series of bitplanes

L_0, L_1 and G_2 are still gray-level images. We decompose them into binary bitplanes for the sake of watermarking. A 256-gray (8-bit-gray scale) image of size $M \times N$ can be decomposed into 8 bitplanes which can be viewed as ordinary two-dimensional binary watermark images for watermarking[4].

III. DECOMPOSITION OF VIDEO

3.1 procedure of multiresolution temporal representation of video

For watermarking the motion content of video, we decompose the video sequence into multiresolution temporal representation with a 2-band or 3-band perfect reconstruction filter bank by 1-D Discrete Wavelet Transform (DWT) along the temporal axis of the video. To enhance the robustness against the attack on the identical watermark for each frame, the video sequence is broken into scenes[2][6] and the length of the 1-D DWT depends on the length of each scene. Let N be the length of a video scene, F^k be the k-th frame in a video scene and WF^k be the k-th wavelet coefficient frame. The wavelet frames are ordered from lowest frequency to highest frequency[2], i.e., WF^0 is a DC frame. The procedure of multiresolution temporal representation is shown in Fig.2.

3.2 2-D multiresolution representation of the wavelet coefficient frames

The multiresolution temporal representation mentioned above is only along the temporal axis of the video. The robustness of spatial watermarking for each frame (especially for I-frame) should be considered in addition for the sake of surviving MPEG video lossy compression. Hence, the wavelet coefficient frame WF^k is decomposed into multiresolution representation by the 2-D DWT[5][7]. Fig.3 shows the three-scale discrete wavelet transform with 3 levels.

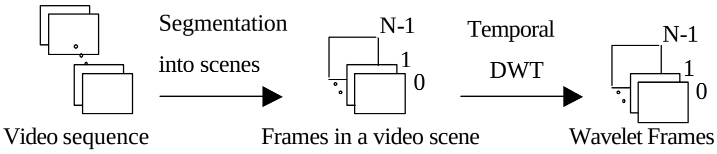


Fig.2 Procedure of multiresolution temporal representation

IV. EMBEDDING AND EXTRACTING PROCEDURE

Fig.4 shows the watermarking procedure. Assume that original video is a series of gray-level images of size 352 × 240, and the watermark image is a 8-bit-gray scale image of size 64 × 64.

4.1 Generating the pyramid structure of the watermark image and decomposing them into bitplanes

The watermark image is decomposed into hierarchical structure of images L_0, L_1 and G_2 by the resolution-reduction method and each of the hierarchical structure of images is decomposed into 8 bitplanes for watermarking conveniently as described in section II.

4.2 Pseudo-random permutation of the bitplane

For robustness to the common picture-cropping processing, a fast two-dimensional pseudo-random number traversing method[4] is used to permute each bitplane of the watermark image to disperse its spatial location for the sake of spreading the watermarking information.

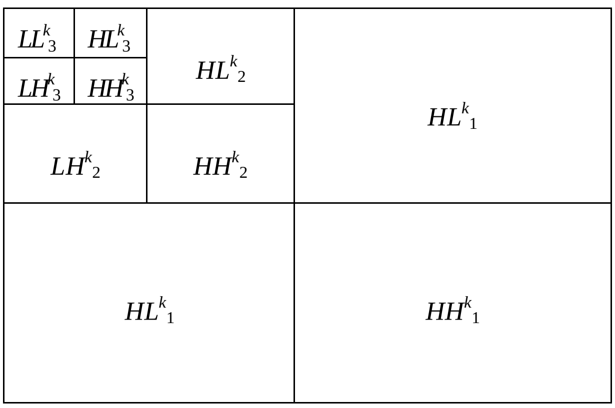


Fig.3 A three-scale wavelet decomposition with three levels of the *k*-th wavelet coefficient frame in a video

4.3 Modification of 2-D DWT Coefficients

In our multiresolution embedding algorithm, the differential resolution layers of the watermark are embedded in the detail components of the original wavelet coefficient frames at the same resolution as shown in Fig.4. The lowest and highest frequency levels of the original wavelet coefficient frames are left unmodified. Since there are 8 bitplanes in each resolution layers of the watermark, each bitplane is embedded in each original wavelet coefficient frame. Hence, there are 8 original wavelet coefficient frames are watermarked. Considering the video energy is concentrated on lower wavelet coefficient frames and the watermark should survive MPEG lossy compression, we embed the 8 bitplanes in middle wavelet coefficient frames that are corresponding to middle frequency subbands along the temporal axis of the video. Let the middle wavelet coefficient frames be from *m*-th wavelet coefficient frame to (m+7)-th wavelet coefficient frame in a video scene. The bitplanes (16×16×8) of G_2 with the low-resolution information of watermark are embedded in low frequency subbands HL^k_3 (44×30×k, where k=m,.., m+7) and the redundance of the bitplanes is embedded in LH^k_3 (44×30×k, where k= m,.., m+7). The bitplanes (32×32×8) of L_1 and their redundance are embedded in subbands HL^k_2 (88×60×k, where k= m,.., m+7) and LH^k_2 (88×60×k, where k= m,.., m+7) respectively. Also, the bitplanes (64×64×8) of L_0 plus their redundance are embedded in subbands HL^k_1 (176×120×k, where k= m,.., m+7) and LH^k_1 (176×120×k, where k= m,.., m+7). If the length of the selected scene is very long, i.e, N>>8, the watermark bitplanes can modify the lower and higher frequency subband frames as well as middle frequency subband frames in the scene at the same time. Otherwise, if the length of the selected scene is very short, the frames within the scene are transformed into 2-subband by DWT. We have to modify the lower or higher frequency subband frames, especially modify the both. By embedding the watermark information in the lower and higher frequency frames, it is robust enough against the attacking on the watermark in motionless or motion regions of video sequence.

4.4 Inverse of the 2-D and 1-D DWT

By inversing the watermarked 2-D and 1-D DWT wavelet coefficient frames, we obtain the watermarked video.

4.5 Watermark extracting procedure

The watermark extraction requires both the original and the watermarked video. The extraction steps are as follows:

1) Both the original and the watermarked video are broken into scenes and decomposed each scene into multiresolution temporal representation (a series of wavelet coefficient frames) by 1-D DWT along the temporal axis of the video.
2) Each wavelet coefficient frame of both the original and the watermarked video is transformed by 2-D DWT into multiresolution representation of three levels.
3) For each level, the DWT Coefficients of the original and watermarked wavelet coefficient frames are compared to extract the 8 bitplanes of images L_0, L_1 and G_2.
4) The pseudo-random permutation is reversed according to the predefined pseudo-random order for these bitplanes.
5) By composing these bitplanes into the gray-level images G_2, L_1, and L_0, the extracted watermark is reconstructed.

V. EXPERIMENTAL RESULTS

Simulations are performed to evaluate the performance of the algorithms mentioned above. We use the first 440 frames of size 352×240 from American Football video for our tests. There are several scenes in the video sequence under test. We choose two scenes as shown in Fig.5 for instance where scene Ⅰ has 60 frames and scene Ⅱ have 16 frames. We transforms the frames scene Ⅱ into 2-subband frames and embeds the watermark bitplanes in low frequency subband. Since there are 60 frames in scene Ⅰ, we decompose the frames of scene Ⅰ into 3-subband frames and modify the lower and higher frequency subband frames as well as middle frequency subband frames of the scene Ⅰ at same time. The watermarked frames and the extracted watermarks are shown in Fig.6. To test frame dropping and interpolation, the even index frames of the scene Ⅰ are dropped[2]. The missing frames are replaced by the average of the two neighboring frames of scene Ⅰ. Fig.7(a) is the extracted watermark image of dropped and averaged version of watermarked frames of the scene Ⅰ. After 20 watermarked frames are dropped and replaced by corresponding original frames in watermarked frames of scene Ⅰ, the extracted watermark image is given in Fig.7(b). Fig.7(c) demonstrates the extracted watermark from scene Ⅰ after the Football video is compressed by MPEG (ratio=18.8). Fig.7(d) the extracted watermark from scene Ⅱ after the Football video is compressed by MPEG (ratio=18.8). Fig.7(e) shows the watermark extracted from watermarked frames of scene Ⅰ which are captured non-uniformly by MPEG player.

VI. CONCLUSION

This paper proposes a multiresolution watermarking technique for video. The watermark is a gray-level image whose decomposed products are correlative with each other and also abundant in quantity. The decomposed watermark image with different resolution is embedded in the corresponding resolution of the decomposed video. As the multiresolution representation of the video is for spatial axis of each frame as well as the temporal axis of the video, it is convenient for watermark to be adapted to motion and motionless regions of the video and the spatial detail of each frame. Experimental results show that the technique proposed here has more visually recognizable information of copyright ownership compared to the traditional watermark and is robust enough against the attack of frame dropping, averaging and lossy compression.

REFERENCES

[1]R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, ¡°Perceptual Watermarks for Digital Images and Video,¡± Proceedings of the IEEE, Vol.87, No.7, July 1999, pp.1108-1126

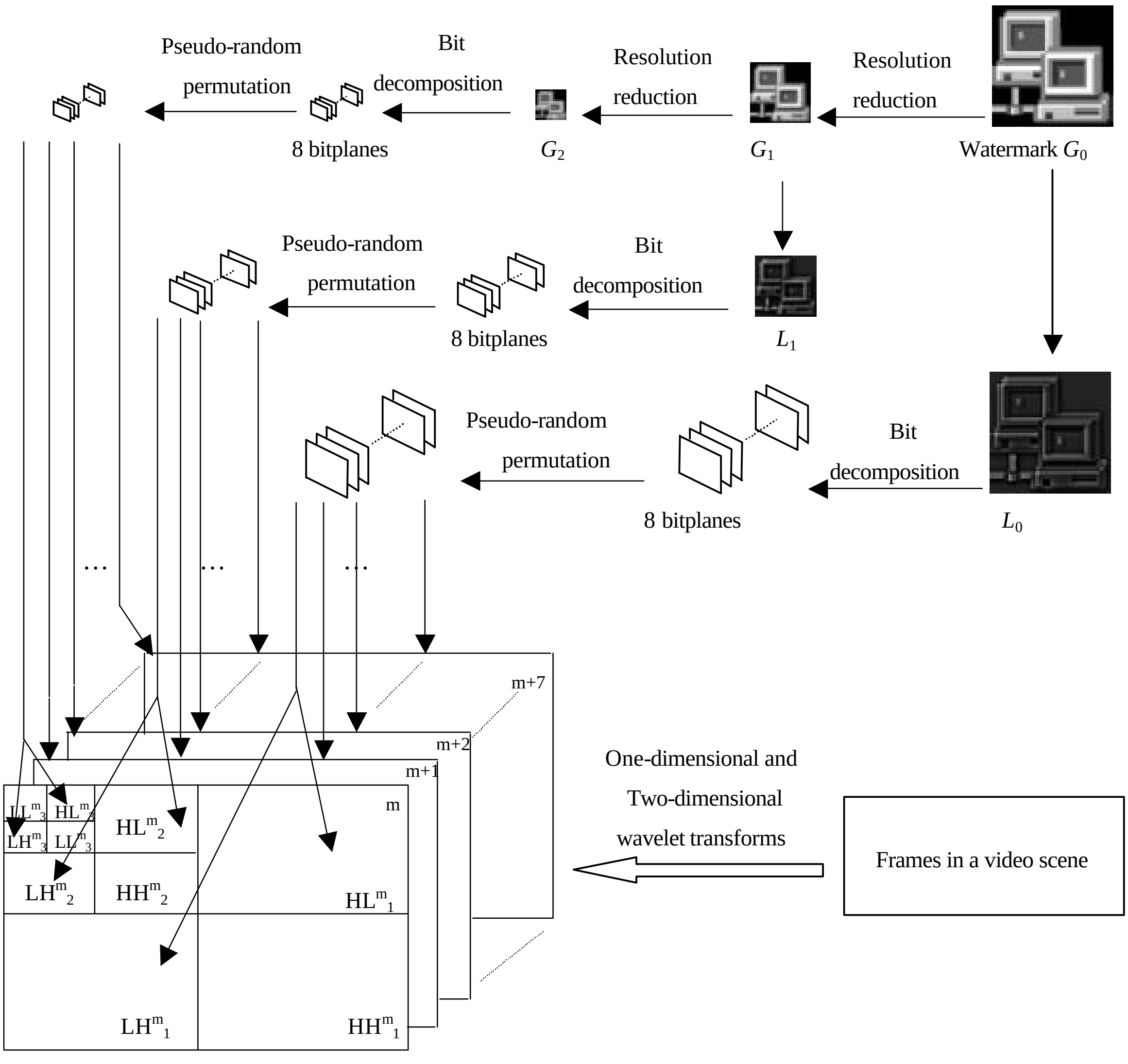
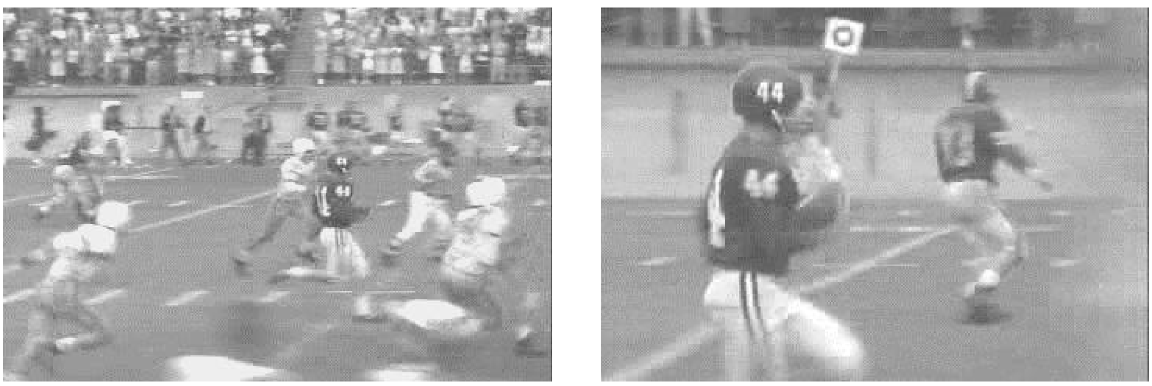


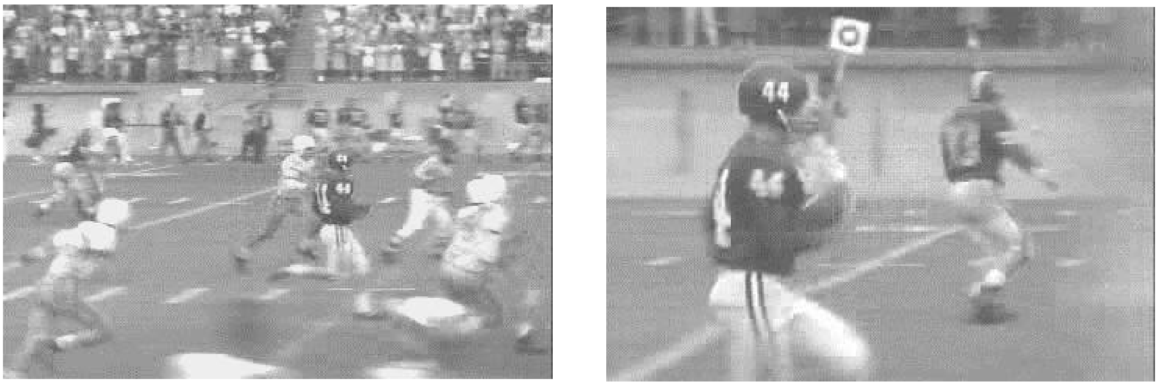
Fig.4 Digital watermark embedding scheme diagram



(a) (b)

Fig.5 The original frame(352×240) from American Football video

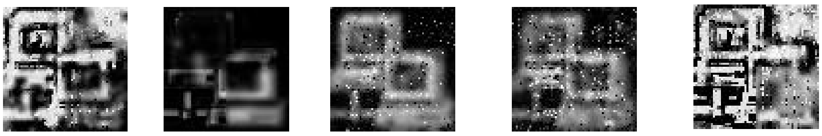
(a) scene I (b) scene II



(a) (b)

(c) (d)

Fig.6 (a) One of the watermarked frame of scene I (PSNR=36.76)
(b) One of the watermarked frame of scene II (PSNR=33.34)
(c) The extracted watermark (NC=1.0) from watermarked scene I
(d) The extracted watermark (NC=1.0) from watermarked scene II



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

Fig.7 (a) The extracted watermark (NC=0.69) of dropped and averaged version of the watermarked frames of scene I
(b) The extracted watermark (NC=0.58) after dropped 20 watermarked frames of scene I
(c) The extracted watermark (NC=0.82) from scene I after the Football video is compressed by MPEG (ratio=18.8)
(d) The extracted watermark (NC=0.72) from scene II after the Football video is compressed by MPEG (ratio=18.8)
(e) The watermark (NC=0.63) extracted from watermarked frames of scene I which are captured non-uniformly by MPEG player

[2]M. D.Swanson, B. Zhu and A. H. Tewfik, ¡§Multiresolution Scene-Based Video Watermarking Using Perceptual Models,¡¨ IEEE Journal on Selected Areas in Communications, Vol.16, No.4, May 1998, pp.540-550

[3]P. J. Burt and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code, " IEEE Trans. on Communications, Vol.31, No.4, April 1983, pp.532-540

[4] Xiamu Niu, Zheming Lu and Shenghe Sun, ¡§Digital Watermarking of Still Images with Gray-Level Digital Watermarks,¡¨ IEEE Trans. on Consumer Electronics, Vol.46, No.1, Feb. 2000, pp137-145

[5]J. M. Shapiro, "Embedded Image Coding Using Zerotree of Wavelet Coefficients," IEEE Trans. on SP, Vol.41, No.12, Dec. 1993, pp.3445-3462

[6] J. Nam and A. H. Tewfik, ¡§Combined audio and visual streams analysis for video sequence segmentation,¡¨ in Proc. 1997 Int. Conf. Acoustics, Speech and signal Processing, 1997, pp.2665-2668

[7]Seung-Jong Choi and John W. Woods, ¡§Motion-Compensated 3-D Subband Coding of Video,¡¨ IEEE Trans. on Image Processing, Vol.8, No.2, Feb. 1999, pp.155-167