A Novel Video Hiding Method for Biometric Identification

Minghai Yao[1,a], Qi Miao[2,b]

[1]Department of College Foundation Education, Bohai University, Jinzhou, 121013, China

[2]School of Computer Science and Information Technology, Northeast Normal University, Changchun, 130024, China

[a]yao_ming_hai@163.com, [b]qim801@nenu.edu.cn

Abstract: In this study, a novel video hiding method is proposed for protecting the transmitted biometric information. Different from the most of existing information hiding methods, the hiding content is an image set rather than a single image or feature vector. For the sake of making the hiding result more invisible and robust, human visual model is adopted by considering both inter-frame and intra-frame motion information. Moreover, the biometric image set is hidden into the video sequence discretely by using discrete wavelet transform method. In particular, in order to detect the integrity of stego-video and guarantee the exactly extraction, the frame number is embedded into the corresponding frame as one watermark. Finally, the extracted image set is used for biometric identification. The identification accuracy is used to verify the performance of hiding method. Experiments show that the proposed hiding method can reach good invisibility and robustness, and ensure the effectiveness of biometrics simultaneously.

1. Introduction

In computer vision, biometrics has been regarded as an efficient and reliable identification technique, such as face [1], palmprint [2]. However, the digital biometrical information is unsecure and easily intercepted or attacked in the remote recognition, since it is extremely convenient for hackers to copy or process multimedia data. For the sake of solving the security problem, watermarking techniques have been studied to protect the digital information and some biometric watermarking methods have been described [3-4]. In this work, a novel and effective image set hiding method based video is presented to secure biometrics. Different from most of the existing biometric information hiding methods, the hiding content is an image set rather than a single image or some features, aiming to make the method more robust. For improving the invisibility and robustness, the biometric data is discretely hidden into the video sequence by considering the motion information of objects adequately. It is worth noting that the frame number as a watermark is embedded into the corresponding frame to ensure the extraction accurately.

2. The Proposed Hiding Method

Palmprint images have been proved to be one reliable feature for personal verification and studied widely [5]. In this study, the palmprint image set is employed as the secret information for remote identification. Several public video sequences are used as the cover database. Given a palmprint image set, a video is selected randomly from the cover database as the transmission carrier. First, as a watermark, the frame number is embedded into the corresponding frame. Then, two motion analysis methods are used to locate the embedding frames and positions. Finally, the image set is embedded into the fast motion frames with fast motion regions, aiming to increasing the invisibility.

In general, the secret information is hidden into the regions with fast moving. So without loss of generality, the palmprint image set is embedded into the moving fast regions in this study. The temporal difference is a simple and effective method for motion detection, which takes consecutive

video frames to determine the absolute difference. Suppose I_n represents the nth frame, the difference image D_n of I_n can be calculated:

$$D_n = I_n - I_{n-1} \tag{1}$$

where I_{n-1} is the former frame of I_n.

The greater D_n is, motion variations are more significant and the motion activity is higher. The method based on adjacent frames difference is available. The blocks which include moving objects in these frames are located for biometric data embedding. In this paper, we adopt the robust real-time multiresolution algorithm which is available to detect moving objects in the presence of camera movement. The 2D affine model is used here to describe the camera motion. This model can deal with many kinds of motion such as translation, rotation, scaling and deformation.

$$\begin{cases} u(X_i) = a_1 + a_2 x_i + a_3 y_i \\ v(X_i) = a_4 + a_5 x_i + a_6 y_i \end{cases} \tag{2}$$

where $X_i = (x_i, y_i)$ is the position of a pixel point.

This model can also be expressed in the matrix form as follows:

$$V_A(X_i) = \begin{bmatrix} u(X_i) \\ v(X_i) \end{bmatrix} = B(X_i)A, \quad B(X_i) = \begin{bmatrix} 1 & x_i & y_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_i & y_i \end{bmatrix} \tag{3}$$

where $A^t = (a_1, a_2, a_3, a_4, a_5, a_6)$ denotes motion parameters which need to be estimated, $V_A(X_i)$ is the motion vector at point X_i.

To estimate motion parameter vector $\Theta = (A^t, \xi)$, the minimization is done between two consecutive frames I_{t+1} and I_t [6].

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{X_i} DFD_{\Theta}(X_i) \quad DFD_{\Theta}(X_i) = I(X_i + B(X_i)A, t+1) - I(X_i, t) + \xi \tag{4}$$

where ξ is an intensity variation parameter which is used when global illumination changes occur. The minimization is solved through iteratively reweighted least squares. If the weight of X_i is more close to 1, the possibility that the pixel belongs to the background is bigger. Otherwise, it is almost considered as a foreground pixel.

$$M_t(X_i) = \begin{cases} 1 & if \quad W_t(X_i) < \tau \\ 0 & otherwise \end{cases} \tag{5}$$

where W_t is the weight map obtained after the last iteration of IRLS and τ is the binary threshold.

Fig.1 shows the examples of motion analysis for the video 'News'. Fig. 1(a1) and (b1) are two adjacent frames (352×288), (c1) gives the absolute difference results of (a1) and (b1), and (d1) is the block based motion probability image magnified by a factor 255, the size of each block is 8×8. (e1) is the corresponding block based motion probability image.

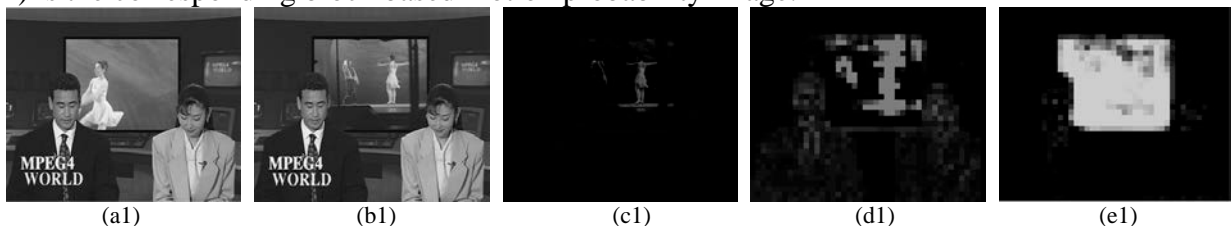

| (a1) | (b1) | (c1) | (d1) | (e1) |

Fig. 1. The motion analysis of the video 'News'

Taken a video sequence as a carrier, the frame number is embedded into related frame to ensure

accurate extraction of biometric image set. In this study, LSB method is employed for embedding the watermark. DWT is adopted to embed the palmprint image set. The embedding processes can be depicted as follows:

Step 1: Converting images. First, each pixel value of palmprint image is converted eight bits. Then, the binary stream is got by linking all pixel values sequentially. It is that the N transmitted images with the size of $m \times n$ can be converted to a binary stream $B = B(s) \in \{0,1\}$ $s = 1,2,\ldots,N \times m \times n \times 8$.

Step 2: Locating embedding positions. DWT with one level is carried on each watermarked frame. On the basis of the length of binary sequence B, the video frames with larger motion activity of approximations sub-band are chosen as embedded frames based on temporal difference. Then, some blocks with higher motion activity are chosen for embedding palmprint image set based on real-time multiresolution algorithm.

Step 3: Embedding binary sequence B. Inspiring by literature [7], we embed the the binary sequence B into the approximation coefficients of DWT in the regions with larger motion activity. The embedding rule can be written as

$$V_{k,ij} = V_{k,ij} + \alpha \times B(s), \qquad 1 \le k \le 300 \tag{6}$$

where $V_{k,ij}$ is the approximation coefficient of kth frame, α is the embedding parameter.

For one received stego-video, we first extract watermarks from the video and determine whether the stego-video is changed during the transfer. Then, the binary sequence *B'* is extracted in terms of the motion activity by combing the watermarks

$$B'(s) = (V_{k,ij}^{'} - V_{k,ij})/\alpha, i, j = 1,2,\cdots,m/2 \tag{7}$$

where $V_{i,j}^{'}$ and V are the corresponding low-frequency coefficients of the stego-image and cover image with one-level wavelet transform respectively. If $B'(s) > T$, then $B'(s) = 1$, otherwise $B'(s) = 0$, where T is the decision threshold.

3. Experimental Results and Analysis

For the sake of evaluating the performance of our proposed method, four public video sequences are utilized as carriers. Fig. 2 exhibits the first frames of the four videos.

| News | Foreman | HandballC | Highway |

Fig. 2. The first frames of the four video sequences

The palmprint database from Hong Kong Polytechnic University is used as the transmitted biometric data. In experiments, Five images of each individual are used for training and the others for transmitted testing set. So, N is equal to 5. The size of each image is 128×128. It is worth noting that the length of binary sequence B is $5 \times 128 \times 128 \times 8 = 655360$ bits. In the method of block-based motion analysis, the binary threshold τ is set 0.3, the size of each block is 16×8 and the embedding capacity is 128 bits. Besides, the first 32 blocks with the larger activity is selected to embed the binary sequence. Therefore, the embedding capacity is $4 \times 4 \times 8 \times 32 = 4094$ bits per frame. That is, the whole binary sequence needs 160 frames for embedding. The embedding parameter is set $\alpha = 8$ and the extracted threshold is set T=0.2.

Imperceptibility Evaluation: Table1 gives the averaged PSNR values of the four video. Generally, the image quality is acceptable if the PSNR value is larger than 35 dB. Therefore, these results

indicate that our method can reach good stego-video quality.

Table 1. PSNR values of four stego-videos

Video sequences	News	Foreman	HandballC	Highway
PSNR (dB)	46.60	46.62	46.56	44.38

Robustness Evaluation: Table2 shows the identification accuracy in terms of various attacks with different attack factors.

Table 2. Identification results under various attacks

Spatial filtering	Standard deviation	0.35	0.40	0.45	0.50	0.55
	Identification accuracy	97%	84%	56%	44%	31%
Noise addition	Noise density	0.01	0.02	0.03	0.04	0.05
	Identification accuracy	99%	98%	90%	64%	48%
Scaling	Scaling factor	1.1	1.2	1.3	1.4	1.5
	Identification accuracy	91%	97%	93%	83%	99%
Frame dropping	Number of dropped frames	4	6	8	10	12
	Identification accuracy	99%	98%	98%	99%	99%
Frame averaging	Number of averaged frames	4	6	8	10	12
	Identification accuracy	98%	99%	98%	98%	98%

Seen from type of spatial attack, the method is robust against attacks with small factors. Even the number of dropped frames is equal to 12, the identification rate still can keep 99%. The reason of these good results can be analyzed that since we attack some frames from the video randomly and the embedded frames might not be chosen for attacking. For the attack of frame swapping, we can extract the palmprint set completely since the extracted watermarks can judge and recover the swapped frames. So, the identification rate will not be affected by this kind of attack.

4. Conclusion and Future Work

A novel and effective video hiding method has been presented for protecting the biometric data in this paper. This method differs with most of existing biometric data hiding methods, an image set rather than a single image or feature is embedded into the unrelated video to increase the validity of biometric identification. For the sake of obtaining better invisibility and robustness, the motion analysis is adequately considered and the biometric data is discretely hidden into the frequency coefficients of transform domain. Specifically, the watermarks are embedded into the video for detecting the integrity of stego-video. A large number of experiments indicate that our proposed method can protect the biometric data efficiently and further ensure the effectiveness of biometrics.

References

[1] Y. Yan, H. Wang, D. Suter, "Multi-sub region based correlation filter bank for robust face recognition," Pattern Recognition, vol. 47, no. 11, pp. 3487-3501, 2014.

[2] X. Q. Wu, Q.S. Zhao, W. Bu, "A SIFT-based contactless palmprint verification approach using iterative RANSAC and local palmprint descriptors," Pattern Recognition, vol. 47, no.10, pp. 3314-3326, 2014.

[3] J. X. Wang, J. Q. Ni, Y. J. Hu, "An efficient reversible data hiding scheme using prediction and optimal side information selection," Journal of Visual Communication and Image Representation, vol. 25, no. 6, pp. 1425-1431, 2014.

[4] X. W. Li, D. Q. Yang, B.H.Chen and Y. Q. Zhang,"Two-factor authenticated key agreement

protocol based on biometric feature and password," Journal on Communications,vol38, no.3,pp. 89-95, 2017.

[5] H.Wang, W.H. Kang, X.P. Chen, "Palm print and palm vein joint recognition system based video," ACAT OPTICA SIMICA, vol. 38, no.2 pp. 1-8, 2018.

[6] A. Bugeau and P. Pérez, "Detection and segmentation of moving objects in complex scenes," Computer Vision and Image Understanding, vol. 113, no. 4, pp.459-476, 2009.

[7] P.N. Tao and A.M. Eskicioglu, "A robust multiple watermarking scheme in the Discrete Wavelet Transform domain," Proc. SPIE, vol. 5601, pp. 133-144, 2004.