NOVEL BIOMETRIC SPEECH WATERMARKING SCHEME IN IMAGES USING WAVELET TRANSFORM

Vikrant B. Jundale, S.B.Patil
[1]Department of E&TC (P.G), D. Y. Patil college of Engg. & Tech. Kolhapur, India.

ABSTRACT

The rapid development in digital technologies has necessitated the owners to pay great attention in protecting their digital contents. Watermarking is used for content authentication and ownership protection. Biometric data provide unique means for individual identification. Biometric watermarking refers to the incorporation of biometric data in watermarking technology. By embedding biometrics in the host, it formulates a reliable individual identification as biometrics possesses exclusive characteristics that can be hardly counterfeited. In this paper we proposed a method of embedding owner's speech signal. Speech being a biometric data, the watermark signal in this method is expected to be more meaningful and has closer correlation with copyright holder.

KEYWORDS: *Biometric Speech, Watermarking, Common & Geometric attacks, Copyright protection, Authentication.*

1. INTRODUCTION

The rapid growth of digital media and communication network has highlighted the need for Intellectual Property Rights (IRP) protection technology for digital multimedia. Watermarking of multimedia data has become a hotspot for research in recent years. Watermarking can be used to identify the owners, license information, or other information related to the digital object carrying the watermark. Watermarks can provide the mechanism for determining if a particular work has been tampered with or copied illegally.

In this paper, we present a novel algorithm for robust audio watermarking in image using wavelet transform based on image entropy. The motivation of choosing image as a cover is driven by the fact that human visual system is less sensitive than human auditory system thus an image provides better masking effect. The algorithm is based on decomposition of images using Haar wavelet basis. The hidden data can be recovered reliably under certain attacks such as cropping, compression, noise effect, geometrical attacks and contrast enhancement. As a necessary background, a literature survey of the watermarking techniques is presented. The last part of the paper, analyzes the watermarking results of wavelet-based watermarking technique on different images and audio samples, using various quality assessment metrics.

2. RELATED WORK

Recently biometrics is adaptively merged into watermarking technology to enhance the credibility of the conventional watermarking methods [1]. Haiqing Wange et.al [1] has embedded speech watermark in relational data bases. To minimize the errors into relations, which are introduced by watermarking, watermark should small. Haiqing Wange has used wavelet transformation to compress the speech. Vasta, Singh et al. [2] has embedded voice template into the face image for multimodal verification. Mel Frequency Cepstral Coefficients extracted from voice data are embedded into the face image. The MFCC watermarked face image is stored in the database for recognition. For verification, the MFCC coefficients are extracted from the watermarked face image. The extracted

MFC coefficients are matched with the MFC coefficients of query voice data. However from MFC coefficients voice is not reconstructed for speech verification.
Motwani [3] proposed a method to embed voice biometric watermark in 3D models. Voice samples are directly embedded in 3D graphics. However due to high payload of .wav file, 3D mesh files cannot accommodate the entire watermark. Bartow [4] proposed a framework that encodes voice feature descriptors in raw iris images thereby offering an example of a secure biometric system. The proposed watermarking method embeds a representation of a biometric trait, such as speech which corresponds to the identity of owner, in the digital images. As a biometric of human being, speech is inherent and does not change along with time, it is universal and easily quantifiable. Using speech as a watermark the limitations mentioned above can be overcome. The major challenge in embedding the voice watermark is the exorbitant size of watermark itself.
This paper presents the extended work of in which we are increasing the embedding capacity without affecting the perceptual quality of image.

3. PROPOSED WATERMARKING TECHNIQUE

After selecting voice as a biometric watermark, the very important stage is the feature extraction process of biometric trait. In [5] Linear predictive coefficients [LPC] of speech watermark are embedded in horizontal details of the cover object. As each band in wavelet transform is half the size of host, the embedding capacity is limited. In this paper we are presenting a scheme in which embedding capacity of voice watermark is increased for the same size of cover object as compared to [5] without degrading the perceptual quality of image. This objective is met by following two methods.
1) More compression of Speech signal is achieved by applying wavelet compression on speech watermark.
2) Stationary wavelet transform is used as it is shift invariant and as each band is of same size as that of original image, it provides more capacity for data embedding. Watermarking scheme consists of following steps:
A) Watermark preparation.
B) Watermark Embedding.
C) Watermarking Extraction.

3.1 Watermark Preparation

The audio signal is first encoded using PCM encoding technique at 8 KHz sampling rate. As the Speech file consists of large number of samples we have applied wavelet compression technique to reduce the size of audio watermark. Wavelet concentrate speech information (energy and perception) into a few neighboring coefficients. The wavelet prototype function used for analysis is called the mother wavelet. This function is dilated and translated to achieve the basis function at different scales. The discrete normalized scaling and wavelet basis functions are defined as

$$\Phi_{i,K}(l)=2^{i/2}\, h_i\,(2^i\, l - K) \qquad (1)$$

$$\Psi_{i,K}(l)=2^{i/2} g_i\,(2^i\, l - K) \qquad (2)$$

Where are i and k are the dilation and translation parameters and hi and gi are respectively the sequence of low pass and high pass filter. The choice of the mother wavelet function in designing the high quality speech coders is of prime importance. Choosing a wavelet that has compact support in both time and frequency in addition to a significant number of vanishing moments is essential for an optimum wavelet speech compression [6]. Optimum wavelet can be selected based on the energy conservation properties in the approximation part of the wavelet coefficients. Wavelets work by decomposing a signal into different resolution or frequency bands. Choosing a decomposition level for the DWT usually depends on the type of signal being analyzed or some another criteria such as entropy. For processing the speech signal decomposition up to scale five is adequate, with no further advantage gained in processing beyond scale five[7]. For the truncation of small-valued coefficients, global thresholding and by level thresholding is used. With db6 as a mother wavelet and number of decomposition levels four, eighty five percent of compression is achieved, which suffice the requirement of data embedding capacity and perceptual quality of extracted watermark.

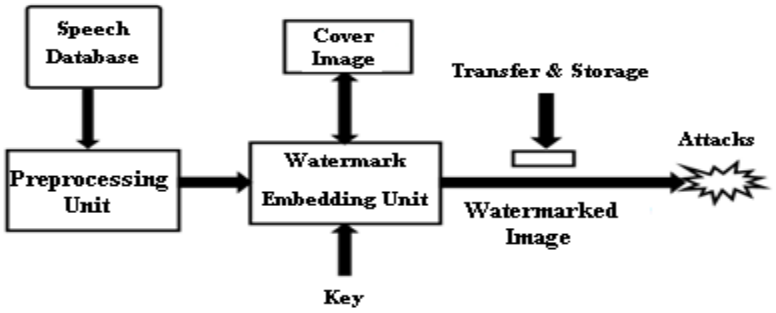

Figure-1 Generic Biometric Watermarking Scheme(Embedding Unit)

3.2 Watermark Embedding

Watermarking scheme is based on stationary wavelet transform domain. Single level decomposition using db2 filter is applied to cover image. The 2D dimensional stationary discrete wavelet transform decomposes an image in to four sub bands namely LL, LH, HL, and HH. As stationary wavelet transform is used, each band is of same size. The decomposed sub bands correspond to the coarse approximation, horizontal, vertical and diagonal details of the image signal respectively. The challenges are how to select the coefficients and which frequency band. Embedding watermark in low frequency band is most resistant to JPEG compression, blurring, adding noise, rescaling, and sharpening while embedding in high frequency is most resistant to histogram equalization, intensity adjustment and gamma correction [8]. Although embedding in low frequency band survives most of the attack, because of its high energy, any modification to its coefficients can be detected easily. We have selected mid frequency band either LH or HL for embedding. While selecting the coefficients in mid frequency band, coefficients which have large perceptual capacity should be selected because they allow stronger watermarks to be embedded and result in least perceptual distortion [9]. Large coefficients in this band are selected for watermarking. The embedding process consists of the following steps:

1) Apply the stationary wavelet transform on image to decompose it into four bands.

2) Select the LH sub band of decomposed image and generate the perceptual mask that identifies the significant perceptual components (watermark indices) of the wavelet coefficients. The method employs the largest 'N' wavelet coefficients, where 'N' is chosen to be equal to length of watermark signal.

3) Insert the watermark into selected wavelet coefficients using additive multiplicative eq.(3):

$$\mathbf{V'_i = V_i + \alpha X_i} \qquad \mathbf{(3)}$$

Where X_i= The i^{th} watermark value

V_i= The original wavelet coefficient. α= the strength of watermark.

As the strength of watermark is increased the robustness of watermark increases and its recovery is good but the image suffers visible degradation. So, the value of the watermark strength is chosen as a tradeoff between perceptibility and robustness. Generate the watermarked image by applying the inverse stationary wavelet transform.

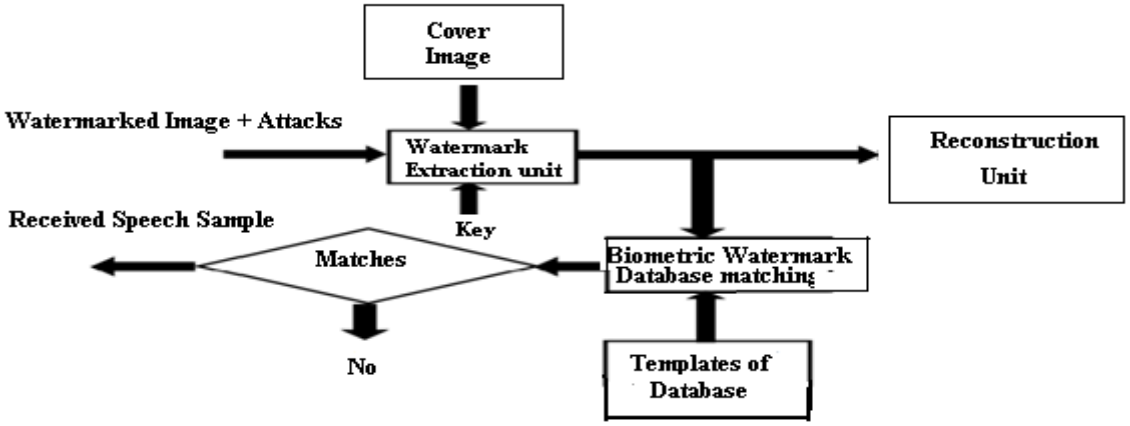

Figure-2 Biometric Watermarking Extracting Unit.

3.3 Watermark Extraction

Watermark extraction process consists of following steps.
1) Apply the stationary wavelet transform on original image and watermarked image to get the four bands
2) Apply the perceptual mask on LH subband to extract the wavelet coefficients that contain the watermark. The perceptual mask used in the watermark insertion process identifies the locations of wavelet coefficients in the recovered image to look for watermark information. Extract the watermark by using the inverse of the insertion equation. The watermark is extracted by using equation (4)

$$X^*_i = V^*_i - V_i / \alpha \quad (4)$$

Where X^*_i= the i_{th} recovered watermark value
3) The extracted watermark is in randomized form. Generate the ordered watermark using the same secret key.
4) Extracted watermark are the wavelet coefficients of compressed speech which are passed to decoder to generate the speech.
5) Quantitative measure used to evaluate the fidelity of extracted watermark with original one is the Similarity Factor (S.F) which is defined by eq. (5).

$$Sim(X, X^*) = (X \cdot X^*) / sqrt(X \cdot X^*) \quad (5)$$

4. RESULTS & DISCUSSIONS

The proposed technique has been experimented on a number of gray scale images and biometric audio clips that were WAV files. To study the results of the proposed technique comprehensively, the images were resized. The results in this chapter are discussed images namely 'Lena' and while as four audio samples of different durations namely 'sample1 (18s)', 'sample2 (15s)'are taken into consideration. The performance of the proposed scheme is gauged by different parameters. The visual imperceptibility of the audio watermark is measured by the Peak Signal to Noise Ratio (PSNR) and entropy of the host image and the watermarked image. The correlation between the original audio clip and the extracted audio clip is monitored by the respective Root mean Square (RMS) values of each. Further the subplots of original and extracted audio samples are shown for different combinations of cover images and audio watermarks. The robustness of the scheme is tested by considering the different types of image processing attacks like cropping, rotation, contrast enhancement, Gaussian noise, speckle noise, poisson noise and salt and pepper noise effect. The experimental result shows that the embedding watermark into subband coefficients is robust against different types of attacks.

5. RESULTS AND ANALYSIS

Images used as a cover image for embedding speech watermark of duration 7sec to 18 sec. are lenna, Baboon, cameraman etc. Figure (a) shows the original image and compressed Biometric watermark of 18sec. Figure (b) shows watermarked image and recovered watermark without any attack. The

watermarked image is subjected to common attacks such as JPEG Compression, salt &pepper noise, histogram equalization, unsharpening, and Gaussian noise which are simulated and checked for extraction. **Fig.(c) to (f)** shows that watermark survives under different common attacks.

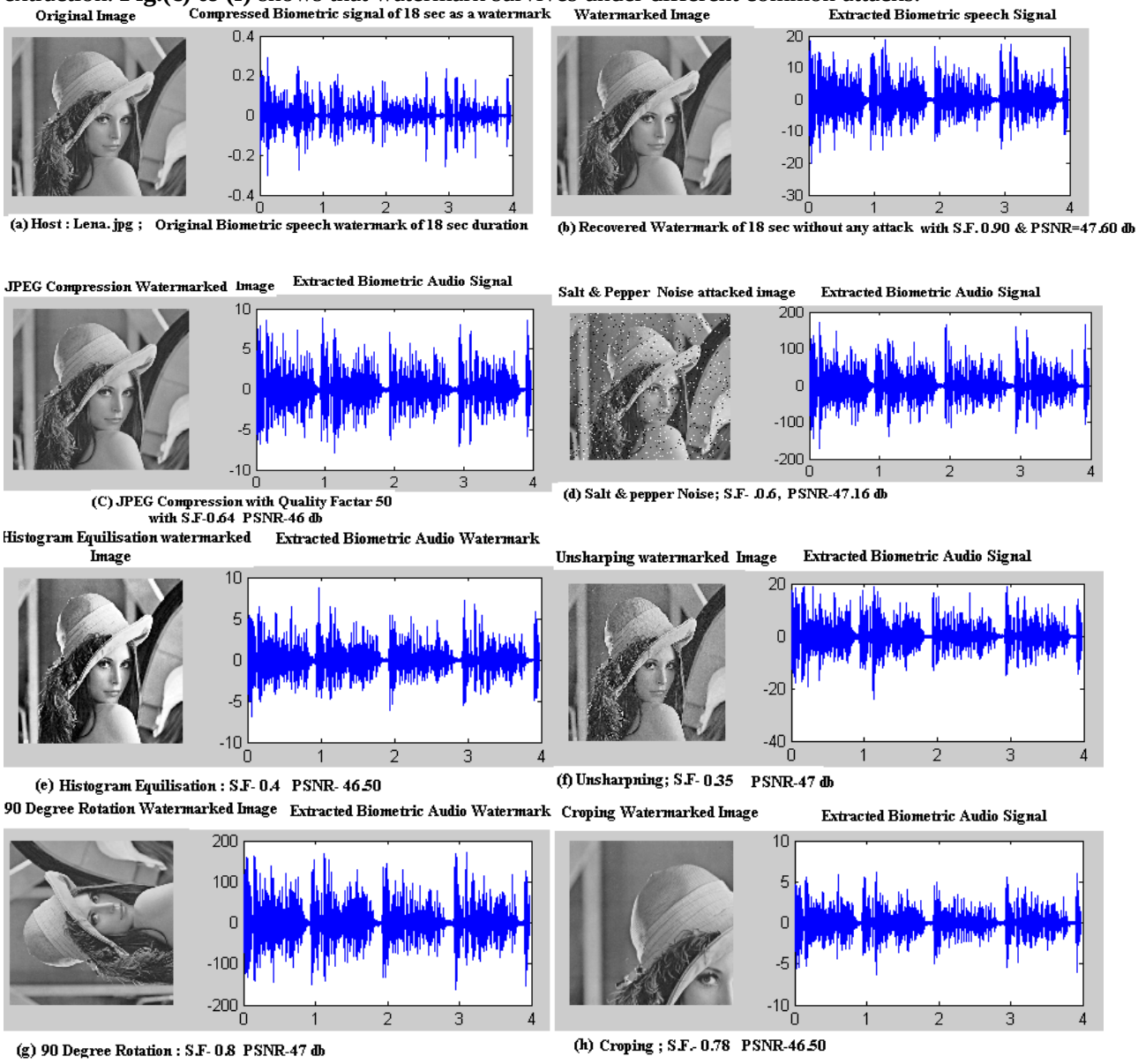

Fig.(g) & (h) shows that watermark survives under Geometric attacks. Strength of watermark plays an important role in maintaining the tradeoff between robustness and imperceptibility. It is varied between one to three.

Table I shows readings which are formulated on empirical results. The perceptual quality of recovered watermark is fair, and as original above Similarity Factor of 0.6. The recovered watermark is noisy but still intelligible for similarity Factor between 0.4 to 0.6. However it does not make any sense if it below 0.4. As the strength of watermark is increased the robustness of watermark increases and its recovery is good but the image suffers visible degradation.

TABLE –I Similarity factor based on empirical result

Speech watermark recovery	
O ≤ S. F ≤ 0.4	**Poor.**
0.4 ≤ S.F ≤ 0.5	**Presence of watermark Detection threshold T=04.**
0.5 ≤ S.F ≤ 0.6	**Average.**
0.6 ≤ S.F ≤ 1	**Good**

TABLE –II

Parameter	Audio sample	No attack	Jpeg compression (Q=50)	Salt & pepper Noise	Histogram Equalized Image	Unsharped Image	Gaussian Noise	90º Rotation	Cropping
PSNR	**A**	**47.60**	**46**	**47.16**	**46.50**	**47.60**	**47.00**	**47.00**	**46.50**
	B	**30.69**	**30.15**	**30.40**	**30.65**	**30.69**	**30.60**	**30.60**	**30.65**
MSE	**A**	**4.71**	**4.47**	**4.37**	**4.57**	**4.60**	**4.71**	**4.71**	**4.52**
	B	**2.15**	**2.15**	**2.12**	**2.10**	**2.14**	**2.15**	**2.15**	**2.18**
RMS	**A**	**2.17**	**2.11**	**2.15**	**2.09**	**2.15**	**2.17**	**2.17**	**2.10**
	B	**1.07**	**1.05**	**1.05**	**1.06**	**1.06**	**1.07**	**1.07**	**1.05**
Entropy (Original)	**A**	**7.45**	**7.44**	**7.45**	**7.45**	**7.45**	**7.45**	**7.45**	**7.45**
	B	**6.20**	**6.17**	**6.20**	**6.17**	**6.20**	**6.20**	**6.20**	**6.17**
Entropy (Recover)	**A**	**7.44**	**7.43**	**7.44**	**7.44**	**7.44**	**7.44**	**7.44**	**7.44**
	B	**6.18**	**6.15**	**6.18**	**6.15**	**6.18**	**6.18**	**6.18**	**6.15**

From the above table it is clear that the proposed algorithm works well and is resistant to different types of attacks. The PSNR, MSE, ENTROPY and RMS values are tabulated in Table-II for one of the three (Lena.jpg) cover images.

6. CONCLUSION

The main objective of this experimentation is to access the viability of using biometric speech as a watermark for copy protection and authentication. Our approach is more intuitive and robust. Using stationary wavelet transform, wavelet coefficients of speech watermark are casted in digital images. When stego images are decoded, the speech data is completely recoverable and intelligible. In addition, the system's ability to cope with added noise and compression of the stego image has been exhibited. The experimental results show that the proposed watermarking scheme suitable for Different signal processing attacks. Compared to previous work [5], with stationary wavelet transform. As each subband is of same size, it provides more capacity for data payload .The proposed watermarking scheme is robust against geometric attacks.

7. FUTURE WORK

Future work should be aimed towards increasing embedding capacity of speech signal. As watermark is embedded into one of the mid frequency band, only half of the image size coefficients are available for watermark casting. Obviously the capacity is upper bounded by the resolution of image. Higher is the resolution of image embedding capacity is increases. Security of the algorithm can be increased by using cryptography techniques with watermarking. By doing that the algorithm becomes more secure because of the existence of two keys, the cryptographic key and the stego key. The scheme can be further elaborated on the color image which can add certain parameters which will act as additional security measure, increasing the payload capacity further. The work can be extended to video watermarking.

REFERENCES

[1] Haiqing Wang, Xinchu Cui and Zaihui cao, "A Speech based Algorithm For Watermarking Relational Databases," 2008 International symposiums on Information Processing

[2] Mayank Vasta, Richa Sigh, Afzel noore," Feature based RDWT watermarking for multi modal biometric system," Science Direct, Image and vision Computing 27(2009)293-304
[3] R. Motwani, S.Dascalu, and F.Harris Jr, " A voice biometric for 3D models," In proceeding of IEEE International Conference on Computer Engineering and Technology, April 2010.
[4] Nick Bartlow, Nathan Kalka, Bojan Cukic, Arun Ross, "Protecting Iris Images through Asymmetric Digital watermarking," IEEE workshop on Automatic identification Advanced Technologies, 2007.
[5] Vandana Inamdar, Priti Rege, Aarti Bang, "Speech based watermarking of digital images," IEEE International Conference, TENCON 2009, Singapore.
[6] R. Kastantin, D.Stefanoiu, G.Feng, M.Martin and M. Mrayati, " Optimal Wavelets for High Quality Speech Coding," Proc. Of EUSIPCO-94, Europian Association for signal Processing, Edinburgh Scotland, 1994, pp.399-402
[7] I.J.Agbinya," Discrete Wavelet Transform techniques in Speech signal Processing," IEEE TENCON Digital signal Processing Application Proceedings, IEEE, New York 1996, pp 514-519
[8] Chirawat Temi, Somsak Choomchuay and Attasit Lasakul," Robust Image Watermarking Using Multiresolution Analysis of Wavelet," in IEEE proceedings ofISCIT 2005.
[9] Wenwu Zhu, Zixiang Xiong ,and, Ya-Qin zhang," Multiresolution watermarking of images and video," IEEE transactions on circuits and systems for video technology, vol. 1 9, No4, June 1999.
[10] Sung Shik koh Chung Hwa Kim ," A cropping, Rotation and scaling Invariant Audio Signal Embedding Watermarking Using LBX Interleaving," 13th international conference on AI and Simulation and Planning in high Autonomy System, AI 2004, Jeju Island Korea October 4-6,2004
[11] Min Wu, Bede Liu ," Multimedia data Hiding" Springer.
[12] G. Langelaar, I. Setyawan, R.L. Lagendijk, "Watermarking Digital Image and Video Data," in IEEE Signal Processing Magazine, Vol. 17, pp 20-43, September 2000
[13] Cox, 1. Kilian, F. Leighton, and T. Shamoon, "Secure spread spectrum Watermarking for multimedia," IEEE Transaction on Image Processing, Vol. 6, pp. 1673-1687, Dec. 1997.
[14] c.Y. low, Andrew B.J. Teoh, "A preliminary study on Biometric Watermarking for offline Handwritten Signature," Proceeding of the 2007 IEEE International Conference on communication,pp.691-695,2007
[15] F, Kutter M, "Multimedia watermarking Technique," IEEE proceeding on Signal Processing," Vol. 87, No.7, pp.I079-I 107, July 1999.
[16] Anil K Jain, Umut Uluday, "Hiding Biometric Data," IEEE transactions on Pattern Analysis and Machine intelligence, Vol. 28, No. I I, November 2003.