ISSN:0975-9646

Remya A. R. et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (4) , 2013, 560 - 563

Authenticating Voice Communication with Barcode as the Watermark

Remya A R, Dr. A Sreekumar

Department of Computer Applications,
Cochin University of Science and Technology,
Kochi-22, Kerala, India

***Abstract*— Authenticating each voice communication by embedding a watermark into its frequency domain guarantees more security to the communication scheme. The suggested system makes use of the mel-frequency cepstral coefficients voice signal features in developing its watermark. Extracted feature vectors are used in the design of an authentic watermark and are embedded in the frequency domain of the communicating voice.**

***Keywords*— Authentication, Watermarking, Feature Vectors, Barcode, Fourier transform etc.**

I. INTRODUCTION

Voice Signal authentication of each communication is of extreme importance in today's information-based society, individual as well as the fields of government, forensics, military, financial, legal, diplomacy, corporation and medicine, the information have to be authenticated to avoid the unauthorized or illegal behaviour and to prevent the misuses. [1]

This project suggests a voice signal authentication scheme for each audio communication.

Pre-processing module starts its framing and windowing procedures upon receiving the recorded voice signal. On completion of the pre-processing module, the feature extraction module extracts the mel-frequency cepstral coefficient values. Extracted coefficients undergo a vector quantization scheme in order to reduce the number of results and will get saved into a database.

The originality of the audio signal can be confirmed by executing the classification module, and provides an assurance if the same person's audio signal has been received earlier or not.

Authentic Barcode watermark is prepared with the help of an online Barcode generator by giving the quantized feature vectors in a specific sequence as input. The generated Barcode will function as the authentic watermark in our authentication system.

Now the authentic Barcode watermark is given to the Watermark embedding module, the embedding is done performing the fast Fourier transformation (fft) and on completion of the embedding process, the inverse transformation (ifft) will be executed before communicating the watermarked voice signal to the desired recipient.

The receiver may suspect a watermarked voice communication and for each voice communication they received an fft transformation is performed to confirm the presence and extraction of the embedded watermark. The extracted Watermark will be scanned using an online Barcode scanner and the obtained feature values can be compared with the original database values in order to finalize the originality of the message.

The overall communication system will thus act as a strong voice authentication scheme to avoid any illegal behaviour.

II. PRELIMINARIES

Authentication of each voice communication with the help of its own feature values gives the communication a different level of security.

[14]**MEL-FREQUENCY CEPSTRAL COEFFICIENTS [MFCC]:** In audio signal processing, the mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency.

For speech/speaker recognition, the most commonly used acoustic features are mel-scale frequency cepstral coefficients.

MFCC takes human perception sensitivity with respect to frequencies into consideration, and therefore are best for speech/speaker recognition.

[2-4] demonstrates audio watermarking as well as audio steganographic techniques in cepstral coefficients.

BARCODE: A barcode is an optical machine-readable representation of data relating to the object to which it is attached. Originally barcodes systematically represented data by varying the widths and spacing of parallel lines, and may be referred to as linear or one-dimensional (1D).

8 21935 11112 4

FAST FOURIER TRANSFORMS: Y = fft(X) returns the discrete Fourier transform (DFT) of vector X, computed with a fast Fourier transform (FFT) algorithm.

If X is a matrix, fft returns the Fourier transform of each column of the matrix. If X is a multidimensional array, fft operates on the first non-singleton dimension.

The execution time for fft depends on the length of the transform. It is fastest for powers of two. It is almost as fast for lengths that have only small prime factors.

It is typically several times slower for lengths that are prime or which have large prime factors.

AUTHENTICATION: Authentication prevents any individual participated in a communication from denying or refrain from their presence of participation. [5-6]

A. *Related Areas*

An audio watermarking scheme in the time domain presented in [7] reveals the watermark without the use of original signal in communication.

An audio watemarking methodology [8] by expoliting the temporal and frequency perceptual masking criteria, constructs its watermark by breaking each audio clip into smaller segments and adding a perceptually shaped pseudo random noise.

Cepstral domain of each audio signal is exploited in the watermark embedding module is presented in [9] which embeds the watermark in the cepstral coefficients of the signal using spread spectrum techniques.

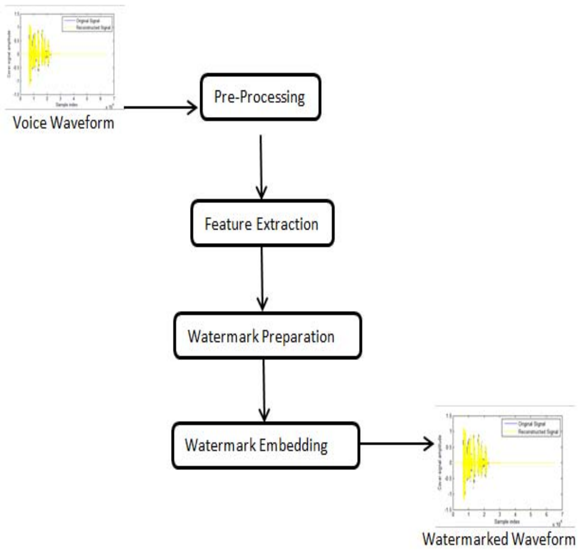
Frequency hopping watermark embedding in the spectral domain is suggested in [10] which use the power spectrum coefficients of the audio signal for embedding the watermark.

Modified patch work algorithm in transform domain is described in [11] is robust to withstand some attacks defined by Secure Digital Music Initiative (SDMI).

Many works have been conducted and being conducted in this area of Research.

III. The Watermarking System

Overall system architecture on the sending window can be summarized as follows:



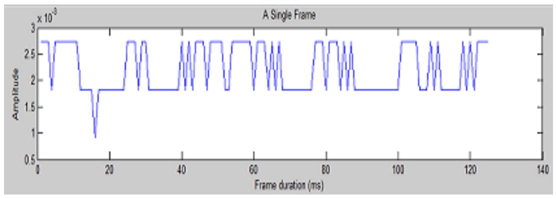
Sending Window

A. *Pre-Processing Step*

Frame-based analysis and windowing function of each audio signal estimates the short-term processing of each signal. Noise removal is not incorporated in our work in order to preserve the originality of the voice and the communicating environment.

1) *Framing:*

Original voice signal are decomposed into a set of overlapping frames. Recorded voice signals are segmented into frames ranging from 2ms to 50ms.



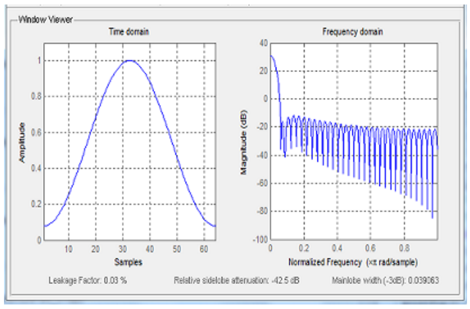
A Sample Frame

2) *Windowing:*

This step involves the estimation of temporal and spectral components of each frame by the analysis of its time and frequency components. The time-frequency mapping is usually matched to the analysis properties of the Human Auditory System [12].

Windowing is usually performed to reduce the signal discontinuities by applying smooth transitions at the beginning and ending of each frame and hence control the spectral leakage effects. That is, it minimizes the spectral distortions or sharp discontinuities at the outer portions. Generally employed windowing techniques are hamming windowing, hanning Windowing, Bartlett etc.

Reducing the spectral leakage in the frequency spectrum of each frame may results in a modest loss of spectral resolution.

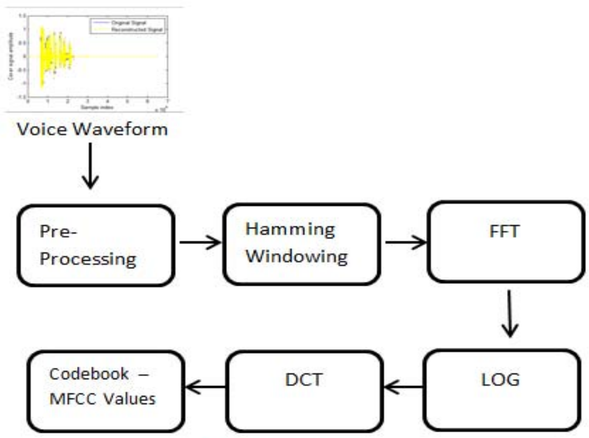


Hamming Window

B. *Feature Extraction*

The Feature extraction module extracts the mel-frequency cepstral coefficients of the recorded voice signal, which are the best feature vector in uniquely identifying the speech/speaker involved in the communication.

Feature extraction is done by performing Fourier transform of the signal, where Fourier analysis decomposes the sampled signal into its fundamental periodic components such as sines and cosines [13].



MFCC Calculation

To find the Mel-frequency cepstral coefficients (MFCC) each chunk of data is windowed with Hamming window, Calculate the magnitude of the FFT, where the FFT is used in shifting these values and the resulting values are converted into a set of filter bank outputs. Taking the base 10 logarithm and its cosine transform will reduce the dimensionality of the obtained results.

The resulting feature vectors such as MFCC values of each signal undergoes vector quantization technique and the results obtained are save into the database.

C. Barcode Generation

Quantized MFCC values are streamed in desired format and submit as input to an online Barcode generator. Obtained Barcode holds all the quantized feature vectors of the recorded voice signal which will be treated as the watermark in the proposed system.

-1.748-1.558-1.418-1.579-1.9

D. Watermark Embedding

The embedding module gets the generated Barcode and embeds it in the frequency domain of the voice signal using fft transform. Once the embedding process finishes, apply the inverse fft of the water-marked signal before communicating the signal to desired recipient.

Let

$$\mathbf{Y = y_1y_2y_3....y_n},$$

represents the original voice signal and $\mathbf{W_f}$ represents the generated FeatureMark.

Then the embedding process can be summarized as follows:

$$\mathbf{Y^{W}{}_{f} = \rho(Y, W_f)},$$

where $\mathbf{Y^{W}{}_{f}}$ corresponds to the FeatureMarked voice signal.

Watermark embedding procedures are usually be performed either in the frequency domain, time domain or in the transform domain of an audio signal. The intention of this task is to exploit the frequency and temporal masking properties of the Human Auditory System(HAS).

This system hides the Barcode watermark by exploiting the fast Fourier transforms of the original voice signal. Perform the inverse fast Fourier transform towards the end of the embedding module will revert the changes occurred as part of the embedding transformation. Now the watermarked signal can be send to the desired recipient.

There might be a mutual understanding between the sender and receiver in the actual embedding and extraction algorithm.

E. Watermark Extraction

At the receiving end, the receiver suspect each communication as the watermarked signal and perform the desired algorithm to detect and extract the embedded watermark. Obtained watermark, that is the Barcode can be scanned to obtain the values embedded in it and compare the values to confirm the originality of the signal.

According to [13] Watermark retrieval process can be grouped in any of the scenarios - when the input audio is employed for decoding and when the input audio is not available for watermark detection:

Case 1: Input audio required for decoding

$$W' = \varsigma\,(Y^{w}, Y)$$

Case 2: Input audio NOT required for decoding

$$W' = \varsigma\,(Y^{w})$$

The proposed voice authentication scheme behaves like the first category and hence not comes under the blind water-marking scheme. The extracted watermark is unambiguous and provides authentic information about voice signal ownership.

-1.748-1.558-1.418-1.579-1.9

For a successful digital audio watermarking scheme: the watermark should meet the imperceptibility criteria, it should have perceptual as well as statistical transparency and should be robust to signal manipulations.

F. Authentication

Voice signal authentication is achieved by comparing feature values embedded in the Barcode with the feature values saved in the data base.

By making use of an online Barcode scanner, the features present in the watermark can be extracted and will be screened for comparing it with the original database values.

Euclidean distance calculation helps in finalizing the closeness between the feature vector values of the original database and the Barcode.

Algorithmic steps can be summarized as follows:

Record Voice Signal

Step 1: Perform the pre-processing
- framing and windowing

Step 2: Extract feature vectors
- mfcc coefficients using fft

Step 3: Feature Vectors will be given as input to a Barcode generator

Step 4: Using fft, embed the watermark in the voice signal

Step 5: Perform the inverse fft and send the watermarked signal to the other end

Step 6: Extract out the watermark by applying the fft on the watermarked signal

Step 7: Obtain the feature vectors from the extracted watermark using an online Barcode scanner

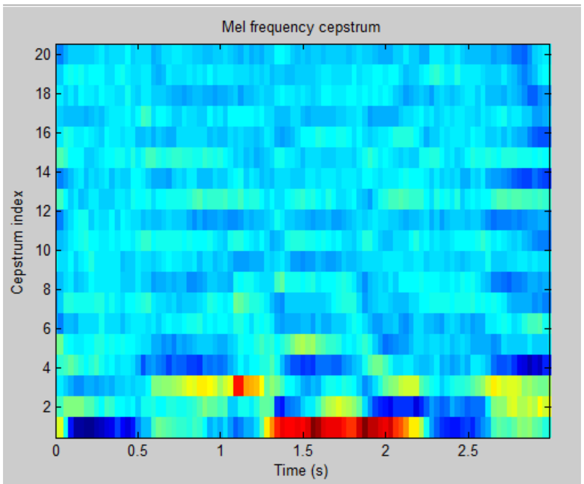
Step 8: Feature comparison to confirm the authenticity of the voice signal

IV. EXPERIMENTAL RESULTS

The proposed system was verified to different data sets of voice signals of different persons. Recorded Signals are pre-processed with framing and windowing terminology.

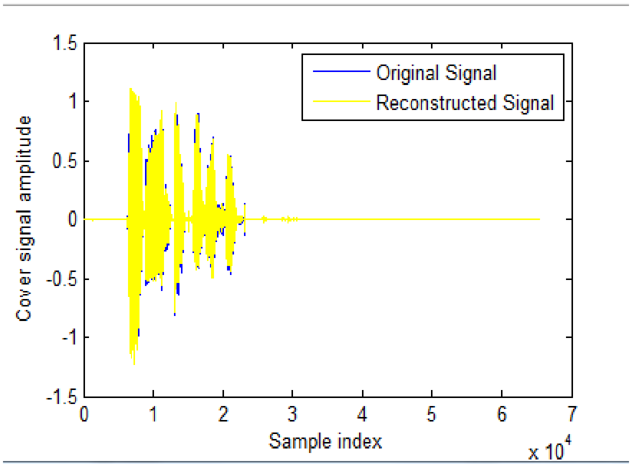
Different data frames obtained for each voice signal are given as input to the feature extraction module.

Feature extraction module accepts different data frames of each voice signal as input and extracts the MFCC coefficients of these individual frames. The MFCC values obtained will be vector quantized and saved into the database. The database values are given to the Barcode generator to generate the Barcode which is the watermark to authenticate the overall communication. Obtained Barcode is given to the watermark embedding module which streams the watermark to embed in the original voice signal by transforming it with the fast Fourier transforms (fft). It performs the inverse transform function after the completion of the embedding process.



The watermarked voice signal is send to the desired recipient, the watermark is obtained by applying the fast Fourier transforms of the Stego signal and by using an Online Barcode scanner, and the embedded features can be extracted out.

Authentication of the voice signal is achieved by comparing the extracted Barcode features with the database values.



However, this system functions as a voice authentication scheme for audio signal communication.

V. CONCLUSIONS

This project with the use of voice signal features and its classification provides a better scheme for authenticating each voice communication.

This system makes use of online Barcode generators and Barcode scanners for creating its watermark. The generated watermark as well as the database feature vectors will be utilized for the comparative analysis and to confirm the originality of the message.

VI. FUTURE WORK

An enhanced version of this scheme which aims to modify the feature values so that even though the presence of watermark is detected it does not reveal the actual feature vectors of an individual communicator.

Thus the new scheme with feature values improves the security of the system where the modification performed with the feature values will serve as the key to the entire system.

ACKNOWLEDGEMENT

This work was funded by the Department of Science and Technology, Government of India under the INSPIRE Fellowship(IF110085)

REFERENCES

[1] Katzenbeisser Stefan and APP Fabien. Information hiding techniques for steganography and digital watermarking. Artech House, London, UK, 2000.

[2] Kaliappan Gopalan. Audio steganography by cepstrum modification. In Acoustics, Speech, and Signal Processing, Proceedings.(ICASSP'05). IEEE International Conference on, volume 5, pages v–481. IEEE, 2005.

[3] Kaliappan Gopalan. Robust watermarking of music signals by cepstrum modification. In Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, pages 4413–4416. IEEE, 2005.

[4] Christian Kraetzer and Jana Dittmann. Mel-cepstrum based steganalysis for voip- steganography. Proceedings of SPIE, Security, Steganography, and Watermarking of Multimedia Contents IX, 6505:650505–1, 2007.

[5] Tom Coffey and Puneet Saidha. Non-repudiation with mandatory proof of receipt. ACM SIGCOMM Computer Communication Review, 26(1):6–17, 1996.

[6] Steve Kremer, Olivier Markowitch, and Jianying Zhou. An intensive survey of fair non-repudiation protocols.Computer communications, 25(17):1606–1621, 2002.

[7] Paraskevi Bassia, Ioannis Pitas, and Nikos Nikolaidis. Robust audio watermarking in the time domain. Multimedia, IEEE Transactions on, 3(2):232–241, 2001.

[8] Mitchell D Swanson, Bin Zhu, Ahmed H Tewfik, and Laurence Boney. Robust audio watermarking using perceptual masking. Signal Processing, 66(3):337–355, 1998.

[9] Sang-Kwang Lee and Yo-Sung Ho.Digital audio watermarking in the cepstrum domain. Consumer Electronics, IEEE Transactions on, 46(3):744–750, 2000.

[10] Nedeljko Cvejic and Tapio Seppänen. Spread spectrum audio watermarking using frequency hopping and attack characterization. Signal processing, 84(1):207–213, 2004.

[11] In-Kwon Yeo and Hyoung Joong Kim. Modified patchwork algorithm: A novel audio watermarking scheme. Speech and Audio Processing, IEEE Transactions on, 11(4):381–386, 2003.

[12] Andreas Spanias, Ted Painter, and Venkatraman Atti. Audio signal processing and coding. Wiley-Interscience, 2006.

[13] Ingo Mierswa and Katharina Morik. Automatic feature extraction for classifying audio data. Machine learning, 58(2-3):127–149, 2005.

[14] Web References: for defining various terms and methodologies used in this paper.