Speech Watermarking of Fingerprints Using Perceptual Hashing for Authentication

P Nagaraju[#1], Arjun Jain [*2], Shruthi Sharath [*3], Shwetha S Choudhary [*4], Vinutha G H [*5]

#Associate Professor
Department of Telecommunication Engineering, R.V.C.E, Bangalore, Karnataka, India

**B.E. Student*
Telecommunication Engineering, R.V.C.E, Bangalore, Karnataka, India

***Abstract*— An authentication algorithm is proposed in this paper using speech watermarking of fingerprints. First, the fingerprint image of a person is divided into non-overlapping blocks using fixed dimensions for each block. The gravity centres for each block are determined and used to generate perceptual hash value. In order to identify the speaker, the perceptual fingerprint hashing that represents the speaker must be closely related to the speech signal. Therefore, the perceptual hashing value of the fingerprint image is embedded as a watermark in the speech signal. At the authentication stage, the fingerprint hashing value is extracted from watermark signal and then it is matched with another perceptual fingerprint hashing value of the same user, already stored. The authentication is determined by the result of the matching .This method also provides an additional layer of security, apart from providing authentication.**

***Keywords*— Speech watermarking, Perceptual hashing, Centre of gravity (CoG), Authentication**

I. INTRODUCTION

E-Commerce is a set of services provided by a group of networked bank branches via the internet. Bank customers may access their funds and perform other simple transactions from any of the member branch offices. The current internet banking authentication system is established by utilizing the public key cryptography system and the digital signature technology for authenticating, which are susceptible to attacks. As the carrier of the huge fund flow, internet banking is very easy to be intruded illegally and viciously attacked. Presently, the internet banking systems are exposed to a few information security risks: clients' ID authentication. Authentication is used to guarantee the authenticity of the data prevent from positive attacks such as altering and imitating. Due to unavoidable hacking of the databases on the Internet, it is always quite difficult to trust the information on the Internet. The password is easily forgotten, vulnerable, leaked, while the smart card and USB certificate token are easily stolen, lost, or spurious. Thus the major issue in core banking is the authenticity of the customer. By combining biometrics and cryptography, Biometric encryption can be extensively exploited for information security. One method is by considering the speaker verification. The basic principle of traditional speaker verification technology creates a personality description model for every speaker. The speaker authentication is implemented by comparing the suspicious speech with the saved speaker model. If the similarity exists, the speaker's identity will be confirmed. Otherwise, the authentication doesn't pass. However, speech can be easily imitated due to the development of high capability digital recorder that can achieve the sound features of the speaker at a very low cost. Thus, using only the voice feature reduces the identity reliability. Therefore, other aided features are required in addition to voice features. The biometric feature like fingerprint is considered for its high recognition accuracy, small space for storage, ease of procurement, low equipment cost and hence it can be used for authentication purposes along with voice features.

However, the same individual's fingerprint images may be different due to the changed orientation, strength of pressing, wetness/dryness of the fingers. Though perceptual hashing function produces different numerical hash representations for similar images, it enables same or similar hash values for the same individual's fingerprint having differences in input angle, orientation etc. Perceptual hash values are obtained from gravity centres. Gravity centres are geometric invariant to rotation and translation and thus can tolerate small distortion. Since gravity centres are calculated based on the image content, the perceptual hash values generated due to the gravity centres of different images are different. Thus, even if two fingerprint images have the same perceptual elements, the hash values developed will be different.

II. PROPOSED ALGORITHM

Speech watermarking of fingerprint image using perceptual hashing for authentication consists of two parts, encoding at the transmitter side and decoding at the receiver side. First the fingerprint image is read from the computer, and the gravity centre of the fingerprint image is obtained by using perceptual hashing model. Then the speech signal is read from the computer. The generated perceptual hashing sequence is embedded into the speech signal to form a watermarked signal. At the receiver side the perceptual hash values are extracted from the watermarked signal and compared with the hash values stored at the receiver end. The authentication is determined by the amount of matching. After comparison, if the hash values are same the authentication is granted else it is denied.

A. *Encoding*

The encoding method used at the transmitter side is data hiding. The speech signal is the cover signal used for hiding the

CoG values of the fingerprint image. The speech signal is also termed as the watermarked signal. The encoding process to achieve the desired results comprises of 3 steps and is shown in Fig.1.

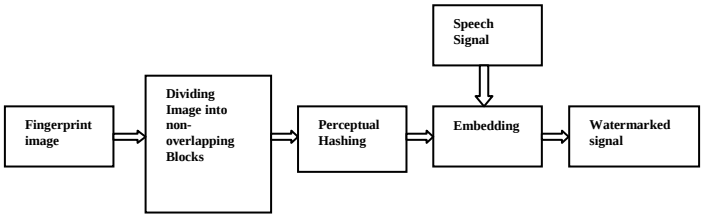

Fig. 1 Block Diagram of the Encoding Process

1) Accepting and Dividing the Image: First the fingerprint is accepted from the fingerprint scanner as a gray scale image as shown in Fig. 2. The accepted image is divided into a number of non-overlapping blocks. Using non-overlapping blocks ensures complete coverage of the image and that no data is repeated. The dimensions should be optimum to ensure coverage of finer details in the fingerprint as well as to achieve a good trade-off between the numbers of CoG values for comparison and to also ensure that the speech signal isn't under perceptible changes.

Fig. 2 Original Image

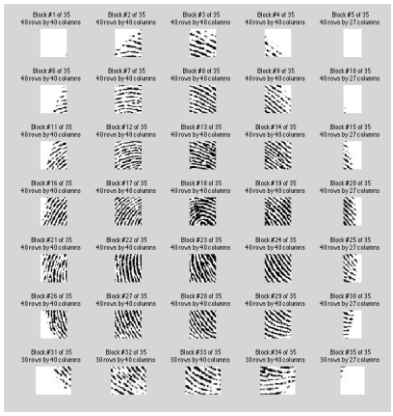

Fig. 3 Image Divided into non-overlapping blocks

2) Finding the Perceptual Hash: The original grey fingerprint image is denoted by f (i, j), $1 \le i \le M$, $1 \le j \le N$ where M=number of rows in the image, N=number of columns in the image. The coordinate of the gravity centre is (mx, my). The intensity value at the specified co-ordinates G (mx, my) is the centre of gravity for that block. The formula shown in equation (1) is applied to each and every block that the original fingerprint image is divided into blocks. The resulting CoG values are converted to 8-bit binary. The binary values so obtained are concatenated as CoG1||CoG2||…CoGn. The binary sequence so obtained represents the perceptual hash. The obtained hash values are compressed using lossless compression technique like Huffman encoding so that it is difficult to detect these in the watermarked signal.

$$\left.\begin{aligned} m_x &= \sum_{i=1}^{M}\sum_{j=1}^{N} i.\, f(i,j) \Big/ \sum_{i=1}^{M}\sum_{j=1}^{N} f(i,j) \\ m_y &= \sum_{i=1}^{M}\sum_{j=1}^{N} j.\, f(i,j) \Big/ \sum_{i=1}^{M}\sum_{j=1}^{N} f(i,j) \end{aligned}\right\} \quad (1)$$

3) Accepting and Embedding Speech: Next part of the algorithm deals with accepting speech sample, processing it and transmitting it. The active speech sample considered is read from the computer. The sample values of the speech signal are considered for embedding. The sample range that falls beyond the centre of the plot is considered and the corresponding sample values are subtracted from the concatenated binary values. This doesn't lead to any changes in the perception of the sound. But, mathematically, the transmitted watermarked speech signal and the original speech signal vary in their characteristics. The overlapping plot of the original speech signal and the watermarked speech signal can be seen in Fig. 4. Except the region where the data is embedded, the spectra are just the same except for high amplitude in a certain region (caused due to embedded CoG).

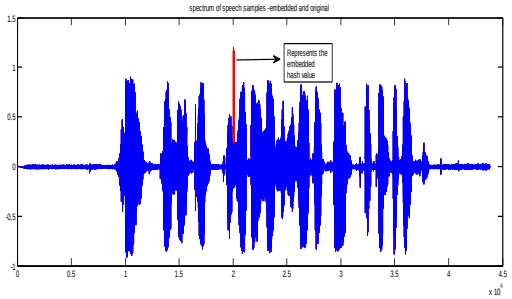

Fig. 4 The overlapping plot of the original speech signal and the watermarked speech signal.

B. *Decoding*

At the receiver, first the obtained watermarked signal undergoes Huffman decoding. The decoding method employed is largely based on comparison of the embedded data with the original data already stored at the receiver. The block diagram of the decoding process is shown in Fig. 5.

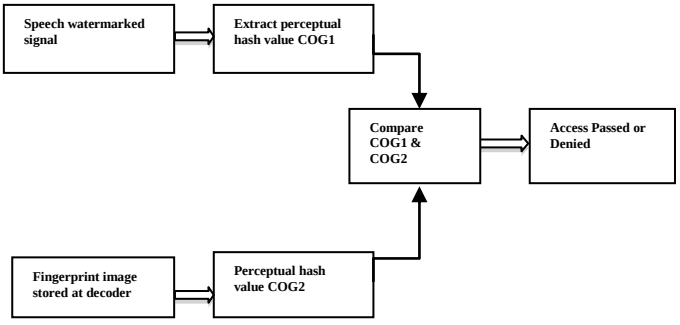

Fig. 5 Block Diagram of the Decoding Process

III. RESULTS AND INFERENCE

For the experiment, about 100 different images were considered randomly. The image set also had images with slightly different features like missing ridges and valleys due to over-pressing, slant orientations etc. The dimensions of the images are 280×187, stored in the format '.jpeg'. The block size dimensions are taken as 40×40 for all images. Consequently, 35 blocks are obtained for each image. The various parameters tested are:

1) Security: The sequence of hash values are arbitrary and are largely dependent on the fingerprint pattern. Hence they are non-periodic and unique for a finger of an individual. Also, the perceptual hash value is sensitive to the orientation of the fingerprint during scanning. Thus, even if the watermark extraction algorithm is known, the correct perceptual hashing value generated by the fingerprint cannot be leaked as it depends on initial orientation as well, which can't be known.

2) Collision: Collision implies two different fingerprint images generating approximate similar perceptual hashing values. We calculate the perceptual hash values for the 100 fingerprint images and the Hamming distance between two perceptual hashing values. Finally, we obtained 3856 matching values which imply a very small percentage of collision and thus a small conflict probability.

3) Authentication Capability: The existing traditional speaker authentication methods are based on extensive speech analysis, feature extraction, modelling. The amount of data to be processed is extremely large, and features extraction is time consuming. This paper introduces the calculation of gravity center of fingerprint image as well as perceptual Hashing technology, effectively reducing the consumption of time. If the size of fingerprint image is 280×187 then the image has 52360 pixels with 8 bits. If the image is stored in binary, the storage space needs 418880 bits. But if we store the perceptual hashing sequence of the fingerprint image based on the gravity centers, the storage space requirement reduces greatly. Since the fingerprint image is randomly partitioned into 35 rectangular blocks, 35 gravity centers are generated and quantified. If we convert the decimal gravity centers coordinate to the binary, the storage space requires only 280 bits. This means, the algorithm can greatly reduce the storage cost, at the same time, greatly increase the running efficiency.

While considering the speaker's voice sample as watermark signal, we need to also look into the environmental factors, such as the speaker's health, age, emotion, environment noise and so on. The proposed method introduces the speaker's fingerprint biological characteristic, which is embedded into the speech to be protected. In fact, this method adds a stable phase to the speech, in the form of fingerprints, so it can avoid the environmental impact. Moreover, the accuracy of speaker identification depends on the accuracy of speaker's fingerprint authentication. If the fingerprint feature can be nicely extracted, the accuracy of speech identification can reach 100% in some cases according to the performance of fingerprint perceptual Hashing technique.

IV. CONCLUSION

The concept of perceptual hashing makes the proposed algorithm more realistic as it takes into account the minute changes in the fingerprint image due to dryness, cuts or bruises. The plot of the original speech signal and the watermarked speech signal look alike because of the compression obtained by Huffman coding. The proposed algorithm provides authentication as it denies access if the speech or fingerprint are not found in the database at the receiver side. The algorithm assures the user of secure communication between the user and the bank and encourages the user to make transactions without fear.

V. REFERENCES

[1] Chetana Hegde , Manu S, P Deepa Shenoy , Venugopal K R , L M Patnaik, "Secure Authentication using Image Processing and Visual Cryptography for Banking Applications", IEEE, International Conference on Advanced Computing and Communications- ADCOM, pp. 65-72, 14-17th December 2008, Chennai.

[2]]S.Saraswathi, "Speech Authentication based on Audio Watermarking", International Journal of Information Technology, pp. 12-17, Vol. 16 No. 1, 2010.

[3] Shervin Shokri, Mahamod Ismail and Nasharuddin Zainal, "Voice Quality in Speech Watermarking Using Spread Plot Technique", International Conference on Computer and Communication Engineering (ICCCE), pp. 399-405, 2008, Kuala Lumpur.

[4] Rupa Patel, Urmila Shrawankar, Dr. V.M Thakare , "Secure Transmission of Password Using Speech Watermarking", IJCST (International Journal of Computer Science and Technology), pp. 221-227, Vol. 2, Issue 3, September 2011.

[5] Zhu Yanqiong, Xu Hui, Gao Zhan, "Security Authentication Scheme Based on Certificateless Signature and Fingerprint Recognition", IEEE, Seventh International Conference on Computational Intelligence and Security (CIS), pp. 380-387 3rd-4th December 2011, Hainan.

[6] C. Soutar, D. Roberge, S. A. Stojanov, R. Gilroy, and B. V. K. Vijaya Kumar, "Biometric encryption using image processing," SPIE(Society of knowledge in photonics and image engineering), Optical Security and Counterfeit Deterrence Techniques II, vol. 3314, 1998, pp. 178–188.

[7] Zhongwei Wang , Ruzhen Dou, Yu Leng, Jianming Wang, "A New Framework of Biometric Encryption with Filter-bank based Fingerprint Feature", 2nd International Conference on Signal Processing Systems (ICSPS), pp. 416-421, Volume 3, 5th-7th July 2010,