International Journal of Engineering Research
Volume No.3 Issue No: Special 2, pp: 22-25

ISSN:2319-6890)(online),2347-5013(print)
22 March 2014

A Review of Biometric Identification In Signal Processing

B.RamaSubbaiah, Asst.Prof, Dept.of IT, RGMCET, Nandyal, e.mail: subhash45229@gmail.com
M.Suleman Basha, Asst.Prof, Dept.of IT, RGMCET, Nandyal, e.mail: sulemanrgmit@gmail.com
B.Ravi Teja, Asst.Prof, Dept.of IT, RGMCET, Nandyal, e.mail: raviteja550@gmail.com

Abstract : This paper describes person identity by finger-print, Face recognition, and voice information using Bio-metrics tool. The person is modeled by their features using Gaussian Mixture Model (GMM). Bio metrics is frequently used in signal processing applications. Thus, we concentrated on the methodology of biometrics for person identification which is useful in industrial and military security systems. The statistical values are measured by GMM in pattern recognition, Face recognition, and voice Recognition. These statistical values will helps for modeling of person using Bio-metric technique. The voice features are mapped into a Mel-Frequency-Cepstral-Coefficients (MFCC) form. The process of indentifying a person using MFCC data is described in this paper.

Keywords: GMM, MFCC, Speaker Recognition, Fingerprint recognition, Face recognition, Healing

1. INTRODUCTION

The use of computers to identify people exist digital computation –filter banks having been used in speech and speaker recognition. Speaker [1] and fingerprint [2] pattern recognition are among the first applications in signal processing. A “wide, diverse market” for computer-based human recognition has identified along with potential applications in “industrial and military security systems”. Facial recognition followed [4]. By the mid- 1970’s, the first operational fingerprint and hand geometry systems were fielded and biometric system testing had begun [5]. Iris recognition systems became available in the mid- 1990’s. Today there are competitive, government-sponsored algorithm test programs in speaker (National Institute of Standards and Technology Speaker Verification Competition), fingerprint (Fingerprint Verification Competition) and facial (Facial Recognition Vendor Test) recognition. In this paper, we will review the biometric identification as per investigation to be made as the idea of speaker identification, fingerprint and facial recognition.

2. BIOMETRIC SYSTEMS

The following figure depicts a general biometric system consisting of data set, transmission, signal processing, storage space and decision sub-systems. It accounts for both enrollment and application systems. Our concern here is exclusively with the signal processing sub-system. As biometric data can be speech, fingerprint or handwriting dynamics, generally we are not commerce with “images”. To simplify our language, we refer to raw signals simply as “samples”.

The biometrics signal processing sub-system is collected of four modules: segmentation, feature extraction, quality control, and pattern matching. The first (segmentation) module must agree on if a biometric signal exists in the received data stream (signal detection) and, if yes, extract the signal from the adjoining data.

The second module must squeeze the signal in some way to protect or improve the between-individual difference while minimizing the within-individual deviation. The yield of this module is a set of mathematical features, which may or may not comprise a vector.

The third module must do a numerical “understanding check” on the extracted features. If the understanding check is not fruitfully passed, the system may be able to alert the user to resubmit the biometric pattern. If the biometric system is at last not capable to bring into being a satisfactory feature set from a user, a “failure-to-enroll” or a “failure-to-acquire” will be said to have occurred. “Failure-to-enroll/acquire” may be due to crash of the segmentation algorithm, in this case no feature set will be bent. The quality control module might even impact the resolution process, directing the decision subsystem to take up higher requirements for matching a poor quality input sample.

The fourth module match up to sample feature sets with enrolled ”templates” from the database and produces a numerical “score”. When both template and features are vectors, the contrast may be as simple as a relating to distance. The pattern matching module determines the reliability of the practical features with the stored generating model. Some pattern matching modules may even direct the characterized recomputation of features.

The conclusion sub-system is considered separately from the pattern matching module. One module might construct a simple “match” or “no match” decision based on the output score from the pattern matcher. An additional module might finally “accept” or “reject” a user’s claim to character (or non-

International Journal of Engineering Research ISSN:2319-6890)(online),2347-5013(print)
Volume No.3 Issue No: Special 2, pp: 22-25 22 March 2014

character) based on multiple measures or measure-dependent verdict criteria.

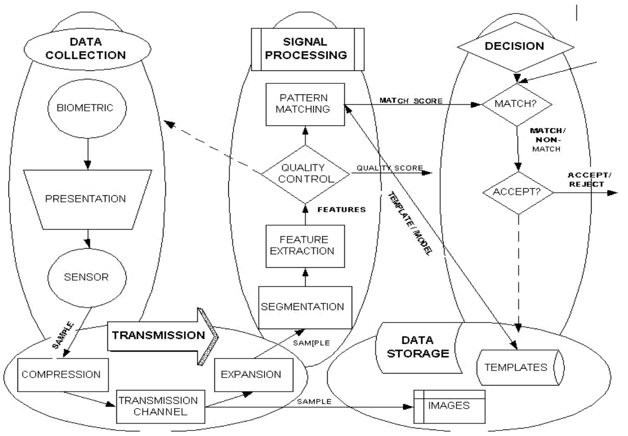

Figure : Example Biometric System

3. FEATURE EXTRACTION

Although the segmentation (signal detection and extraction) problem is interesting, we will restrict our review to the signal processing techniques used in the feature extraction module.

3.1 Speaker Recognition

Speaker recognition has refined of all biometric technologies. The challenging issue of the speech recognition is to identify the speaker voice by the converted “mel-scale” cepstral (Mel-cep) coefficient. We model the speaker by “Gaussian Mixture Model” (GMM).Similar technique of speech recognition is applied for biometric identification.

The approach is described as follows: speech samples are placed in 50% overlapping frames, the length of sample frames are determined by the correlation of given speech. The window function is applied to go to zero at the ends using a function for overlapping and non-overlapping frames with pleasing spectral properties. A Discrete Fourier Transform (DFT) is performed on the windowed frame to extract the local spectral properties.

The logs of the spectral energies are computed for a “Mel-scale” frequency in dynamic time warping. The “Mel-scale” models human voice or any related measuring metrics, its keeping uniform DFT spacing up to some Hz, and interpolating log energies for frequencies increasing by a factor of 1.1.

We apply GMM on each speaker has a number, K. The GMM mean super vectors are calculated for each speaker in a high-dimensional space. The high-dimensional space is mapped into low-dimensional space by PCA. Each speaker is represented as a point by two-dimensional principle axes. So samples from each state are assumed to come from a distribution with some shape in N space. Clearly, if the variance of each of the N Mel-cep coefficients is equal, the distributions have a “round” shape in the N space. Non-round states can be “whitened” using the Eigen system of the state’s covariance matrix. The Eigen system will helps to recognize the variance between any two speakers. The distance from the sample speaker to the centric of the generating state will be assumed normally distributed after the whitening transformation of the distance. Due to a most amazing mathematical relationship, the whitened transformation requires only inversion of the covariance matrix rather than actual extraction of the eigensystem.

Speech clustering needs ‘K-means’ algorithm for grouping of similar speakers based on Mel-cep coefficients of the frames. The mean vector and covariance matrices for each of the K clusters is computed and stored as the enrollment “model”. When a sample Mel-cep vector is received from some unknown speaker, then we estimate probability. We calculate the probability distance of the received sample to each and every centroids for each of the enrolled person’s states, selecting the greatest probability over the K centroids. Among these multiple frames, the maximum likelihood estimator is calculated enrolled speaker from the product of greatest probabilities for each of the sample frames. The “model” allows us to compensate for intra-class speech variation.

For text-constrained speech, the transitions between states should also follow probabilistic model. With the addition of state-transition probabilities, the method known as a “Hidden Markov Model” (HMM).The GMM is a single state of HMM. This highly successful “model” based approach is not general in biometrics.

Speaker recognition methods incompatibly appear to be based on a perception-related measure, the “Mel-cep” coefficient, and not on any physiological feature of speech production. Recently, it has been shown [6] that application of linear discriminate analysis (LDA) to the spectral energy vectors of various phonemes yields most favorable classification projections that look like the properties of human hearing. This suggests that the **perceptual** “Mel-scale” is finest for classification of the products of speech **production.** If speech production is evolutionarily modified to the perception mechanism, then possibly automatic speaker recognition is, after all, based on physiologically-related features.

Recently [8] the field of text-independent speaker recognition is moving past only dependence on the acoustics of the speech and into “ideolectal” schemes, recognizing the

International Journal of Engineering Research ISSN:2319-6890)(online),2347-5013(print)
Volume No.3 Issue No: Special 2, pp: 22-25 22 March 2014

speaker by characteristic sound utterances. Perhaps the added use of distinctive behaviors can be extended to other biometrics.

3.2 Fingerprint recognition

The prevail approach to fingerprint recognition is based on physiological measures – the “precise” points. Precise points are the details in the fingerprint ridges and are used by fingerprint examiners to find out fingerprint matches. Most, but not all, commercially-available automatic fingerprint matching systems are based on a compact definition of “minutiae”, recognizing two types: ridge endings and division (splits). Two fingerprints are said to match to the extent that some number of finer points match in type, location, and ridge hill prior to the minutia point.

Images of fingerprints are contact directly with an imaging device, so normalization in size is not necessary. Unhappily, inconsistencies in contact location and rotation, and the synthetic deformation of the skin must be accounted for by the matching algorithm. The following steps are taken to find the minutiae:

Flow field estimation: A Gabor filter, adjust to the regular periodicity of finger ridges and of scope determined by average ridge directional permanence and rationality, can be placed on the pattern and rotated to disclose ridge direction.

Ridge finding: Knowing the direction of the ridge field, we can look for rapid changes in gray-scale pixel values to indicate ridge edges.

Skeletonizing and binarizing: Having found the track and boundaries of the ridge, we can trim down its width to one pixel and give all ridge pixels a value of 1. Gorge pixels are valued at 0. This is almost equivalent to the energy normalization process in speaker confirmation.

Healing: Breaks in the ridges caused by skin furious and damage must be “healed” by extending not working ridges. If the extensions in the flow field pick up another ridge ending the extension is kept.

Minutiae extraction: Using the following filter, output of 1 indicates a ridge ending, 2 a ridge, 3 a junction.

1	1	1
1	0	1
1	1	1

Check: The original gray-scale fingerprint image can be upturned in polarization (black->white) and the process frequent. Endings and bifurcations also invert under this process, allowing a check of the extracted minutiae.

Matching of sample and enrolled prints must allow for make of the minutiae “constellations” to reimburse for artificial skin deformation during collection.

While minutiae removal seeks to recognize physiological structures, there are different approaches based on association in either the image or convert domains, which do not link directly to ridge physiology.

Can fingerprints be represented as a generating model in the same way as speech? This opposite problem of creating a model from a fingerprint has not been attempted, but the advance problem of generating a fingerprint from a model has been rather flourishing. In the future, we might see fingerprint appreciation based on evaluation of a sample print to a model created from enrollment images. Would the presence or absence of chapping and furious be the fingerprint analogue of “ideolectics”?

3.3 Face recognition

In a commercial manner obtained face recognition systems looks for similarities at the image-level and do not go to model the basic facial structure. Confusion exists because of the use of the expression of “facial features” to describe the mathematical decompositions of the images.

The entire commercially-available systems expect a full-frontal facial image. This image might be taken at a variable range, so the first task after segmentation is to normalize the size and lighting of the face. This is the same of the energy normalization process used in speaker verification. While color data can be used for segmentation, by the normalization stage of the process, images are gray-scale Size normalization is done by finding the eyes, then image conservatory or decimation to normalize the “interocular” distance. So the systems do really be familiar with these physiological structures. Lighting modification is able by normalizing the pixel gray-scale histogram, with the fascinating effect of removing any sense of skin color from the method.

At this situation, signal processing methods are separated. The two largest facial recognition companies are accept that to use image rotting closely related to Principal Component Analysis (PCA). The image is unwrapped into a vector, then estimated into a reduced set of basis vectors. These basis vectors can be the global eigenvectors related with the largest Eigen values of a covariance matrix of different training image.

Sample and enrolled faces are compared based on the projections of each into the selected basis space, may be using a Euclidean distance or a neural network trained on the weightings of enrolled faces.

Possibly systems could be optimized to recognize members of a particular national/gender group by using only that national group in the formation of the basis vectors. This has never been done due to the potential socio-political ramifications.

Additional commercial methods of facial recognition use global 2-dimensional Fourier transform coefficients as input into a neural network, with the exemption of eye location for size normalization, no physiological features are used in commercial facial recognition systems.

Model-based recognition has become popular, because of its potential to overcome sensitivities of image-based techniques to lighting and present angle. A general 3-D facial model is also available on 2-D data by estimating of lighting direction and present angle, then adjusting the physiological features of the model to produce the observed 2-D image. If enrollment images can be collected under convenient lighting conditions, lighting direction estimation will be unnecessary and a better model will result. Pattern matching can be done by estimating present and lighting direction from a sample 2-D image. The resulting 2-D image of the model can be compared to the sample image using image-based techniques.

We may speculation that facial expression or lip movement might be used as the "ideolectics" of facial recognition in the future.

4. SUMMARY

In unkindness variation in signal types, there is some unity in the signal processing approaches in biometrics. Spectral breakdown over locally associated regions frequently follows an energy normalization method. A clear physiological understanding of extracted "features" appears to be unnecessary. Model-based approaches, which might compensate best for intra-class variation, are not universal, recent work in speaker recognition leads to assumption that addition of familiar behavioral patterns with the observed physiology might lead to more correct systems.

5. REFERENCES

i. *A. Rosenberg, "Automatic speaker verification: A review", and B.S. Atal, "Automatic recognition of speakers from their voices", Proc. IEEE, 64(4), 1976, pp. 460-487*

ii. *M. Trauring, "On the automatic comparison of finger ridge patterns", Nature, Vol. 197,pp. 938-940, 1963*

iii. *M. Trauring, "Automatic comparison of finger ridge patterns", Report No. 190, Hughes Research Laboratories,March, 1961, Rev. April 1963*

iv. *T. Kanade, Computer Recognition of Human Faces. Birkhauser, Stuttgart, 1977*

v. *A.J. Mansfield and J.L. Wayman, "Best Practices for Testing and Reporting Biometric Device Performance", issue 2.0, U.K. Biometric Working Group, www.cesg.gov.uk/technology/biometrics*

vi. *H. Hermansky, "Human Speech Perception: Some Lessons from Automatic Speech Recognition", "Text, Speech and Dialogue", V. Matousek, et al (Eds.), Lecture Notes in Artificial Intelligence, LNAI 2166, pp. 187-196, Springer-Verlag, 2001*

vii. *G. Doddington, "Speech Recognition Based on Ideolectal Differences Among Speakers", and W. Andrews, et al, "Phonetic Speaker Recognition", Proc. Eurospeech 2001, Aalborg, Denmark, Sept. 3-7, 2001*