

Modified parametric-based AlexNet structure to classify ECG signals for cardiovascular diseases

S.T. Aarthy^{*}, J.L. Mazher Iqbal

Dept. of Electronics and Communication Engineering, Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India

ARTICLE INFO

Keywords:

Cardiovascular disease
Feature extraction
Multiple ECG signals
Knowledge transfer
Parameter-modified deep learning network

ABSTRACT

Patients with cardiovascular disease typically need constant monitoring, and this is made possible by analyzing their electrocardiogram (ECG) signals to determine the specific pattern of variation associated with each disease. The various malfunctions of the internal parts of the organ generates multiple ECG signals with different commonalities, which can then be used to classify the sub-categories of each disease. In this study, a deep learning network for classifying ECG signals were presented and features were extracted to improve disease classification. Totally 5,655 ECG recordings was used. Initially all 30-s signals were converted to RGB images using Continuous Wavelet Transform (CWT), and then we those images were fed to AlexNet that was trained using slightly modified parameters. According to the results, the proposed method not only outperforms the state-of-the-art methods (98.82% accuracy) but also has higher recall (98.9%) and precision (97.9%) in the aggregate.

1. Introduction

In cardiovascular diseases, there are sometimes observable variations in the ECG signal for five distinct reasons such as atrial fibrillation, atrial flutter, Premature ventricular contraction, diastolic heart failure, and systolic heart failure all cause structural discordance within the body [1–3]. Manually observing the variation in similar ECG signals is difficult. Even if it takes time to find and analyze anomalies. Therefore, it is appropriate to deal with this ECG signal variation by first identifying it and then, using the acquired knowledge to transfer information to a deep learning network. Additionally, diagnosing unexpected variations of ECG signal in a quick and inexpensive manner is a top priority [4].

Several recent studies have investigated the possibility of automatic detection of ECG signal variation through the extraction and analysis of convolutional neural network features (CNN) [5–7]. After sufficient knowledge is acquired from the extracted features, the network employs several machine learning algorithms to measure the anomaly part with greater precision. However, this is only possible after a series of pre-processing steps have been carried out. This helps in overcoming the peculiarities of statistical measurements of various types of ECG signals. To produce more precise classification results, a support vector machine (SVM) must be properly trained with relevant training samples [8].

Wavelet transform (WT) is a common format for detecting specific changes when placed in a higher dimension plan. Where hidden features

are explored extensively, suitable approximations can be taken to improve the specific classification tasks [9–12]. Using a large number of layer weights that can contribute the maximum amount of data suitable for sample training and support knowledge transfer, the deep learning framework is able to make highly accurate predictions. Squeeze Net, for instance, plays a central role in extracting useful features from pre-determined goal tasks. Large amounts of data are required for the classification function, but this information is easily gleaned during the training phase and could be used by all parties involved in processing anomalous data [13–17].

The deep learning framework is very adaptable to train the machine learning model from the labeled data to improve detection and classification tasks [18,19]. The goal of this study is to utilize a deep learning network where parametric values are adjusted based on the time frequency representation or scalogram of the ECG signals to classify them into sub-categories corresponding to specific cardiovascular disease subtypes. It allows for the differentiation of ECG signals via textural-shape related features and ensures that a standalone labeled dataset is adequate for deep-learning model training. There has been a transformation in perspective that is both reasonable and effective. When sufficient data is not available for training a deep architecture, then time frequency representations can be used to transfer knowledge to other ECG recognition tasks [20,21].

Section II describes the deep learning strategy used to classify five

* Corresponding author.

E-mail addresses: aarthyvt604@gmail.com (S.T. Aarthy), mazheriq@gmail.com (J.L. Mazher Iqbal).

distinct ECG signals from cardiovascular diseases, as well as the dataset used in this study. Section III discusses the experimental results, which are compared to state-of-the-art methods. Finally, Section IV concludes with a discussion of the future scope.

2. Materials and methods

This section describes the proposed deep learning network, which is a modified version of the Alexnet structure with updated parametric values. Furthermore, the dataset used in this study to perform performance validation of the proposed method is discussed.

2.1. Dataset

The MIT-BIH Arrhythmia Dataset and the PTB Diagnostic ECG Database are two well-known sources of heartbeat signals that were used to compile this dataset, apart from that Realtime database has been collected from the SRM Medical College Hospital and Research Centre, Chennai. These datasets have enough data to use in the training of a deep neural network. **Table 1** summarizes dataset specification and annotation mappings by category.

2.2. Overview

Fig. 1 shows the classification of cardiovascular diseases using a deep learning network trained on five Category extracted from a test ECG signal. This study focuses on classifying the various ECG signals into subcategories of cardiovascular diseases based on the distinctive pattern observed from the trained feature sample using a deep learning network. Parametric values are adjusted based on the scalogram, which is a representation of time and frequency. It makes the separate labeled dataset sufficient to train the deep learning model with this dataset and provides the ability to differentiate between the ECG signal through textural shape related features. It has been accomplished that a transformation of reasonable knowledge has taken place. Time frequency representations can be used to transfer knowledge to other ECG recognition tasks when there is not enough data available for training a deep architecture. First, the raw dataset's textural quality is improved by removing noisy frequency components using FIR filters, this is the first of several steps required to perform classification function.

The CWT (Continuous Wavelet Transform) can be applied to ECG (Electrocardiogram) signals to analyze their time-frequency characteristics. ECG signals are used to monitor the electrical activity of the heart and can provide information about heart function and health. The CWT of an ECG signal can be used to identify features such as the QRS complex, P-wave, and T-wave, which are important for diagnosing different heart conditions. The QRS complex represents the depolarization of the ventricles, the P-wave represents the depolarization of the atria, and the T-wave represents the repolarization of the ventricles. The CWT scalogram of an ECG signal can provide a detailed view of these features in the time-frequency domain. The scalogram can reveal the frequency

Table 1
Dataset specification.

S. No	Items	Comments
1.	Number of Samples	109,446
2.	Number of Categories	5
3.	Sampling Frequency	125Hz
4.	Data Source	MIT-BIH Arrhythmia Dataset
5.	Classes	[‘C1’: 0, ‘C2’: 1, ‘C3’: 2, ‘C4’: 3, ‘C5’: 4]
6.	Type of ECG signal	Atrial fibrillation, Atrial flutter, Premature ventricular contraction, Diastolic heart failure, and Systolic heart failure

content and temporal location of the different ECG features. Additionally, the CWT can be used to denoise the ECG signal, as it can remove unwanted frequency components from the signal.

The Continuous Wavelet Transform (CWT) of an ECG signal $x(t)$ can be expressed mathematically as follows:

$$\text{CWT}(a, b) = \int x(t) * \psi^*((t-b)/a) dt$$

where ψ^* is the complex conjugate of the mother wavelet function $\psi(t)$, a is the scaling parameter (also known as the wavelet scale), and b is the translation parameter (also known as the wavelet shift). The integral is taken over all t , and represents the inner product of the ECG signal with the wavelet function at each time scale and shift.

There are many different types of mother wavelets that can be used for the CWT. Some examples include the Morlet wavelet, the Mexican hat wavelet, the Haar wavelet, and the Daubechies wavelet. Each wavelet has its own strengths and weaknesses, and the choice of wavelet should be based on the characteristics of the signal being analyzed and the goals of the analysis.

The Morlet wavelet, for example, is a popular choice for analyzing signals with oscillatory components. It is defined as:

$$\psi(t) = \pi^{-1/4} * \exp(iw_0 t) * \exp(-t^2/2)$$

where w_0 is the central frequency of the wavelet. The Morlet wavelet has good time-frequency localization and is able to capture both narrowband and broadband features in a signal.

2.2.1. Pre-processing

First, an ECG signal is segmented into 15-s windows, and then one of those windows is chosen. Then, transforming the amplitude values to a scale between zero and one. After that compiling a list of all zero-crossings of the first derivative that correspond to local maxima. Applying a threshold of 0.6 to the normalised value of the local maxima to identify a set of potential candidates for the ECG R-peak. Using the middle value of R-R intervals to determine the window's standard heart rate (T). The 1.5T-long signal segment has to be calculated which corresponds to each R-peak. For giving input to subsequent processing parts, all extracted beats must have the same length.

2.2.2. Signal normalization

A signal is said to be "normalised" when its levels have been equalised. Z-score normalization was utilised on all ECG signals to ensure consistency. Using the Z-score, statisticians can see how far a given value deviates from the mean of the sample. The Z-score is calculated based on the number of standard deviations from the mean. As a result, the issue of dimension scaling can be addressed by simply ignoring offset effects. The texture information contained in the samples of an ECG signal can be used to analyze the subtle variations in the signal's pattern and amplitude at any given time. For a large enough collection of train samples, this holds true without fail, as shown by

$$ECG_{\text{sys}} = ECG_i + (ECG_{i+1} - ECG_i) \times \rho \quad (1)$$

Where ECG_{i+1} are the samples of the random train that, when adjusted for small differences, come the closest to the nearest neighbouring classes. Therefore, clear evaluation of the most distinct train samples required for the actual training phase of a deep learning network can be achieved using equation (1).

2.2.3. Modified parametric based AlexNet architecture

Fig. 2 depicts the proposed architecture of a modified parametric-based AlexNet for classifying five different types of ECG signals associated with cardiovascular diseases. There are fifteen levels to the proposed Modified parametric based AlexNet structure. There are five convolutional layers, three max-pooling layers, four BLSTM layers, and three fully connected layers, as shown in **Fig. 3**. Convolutional processing allows a deep learning network to automatically identify and

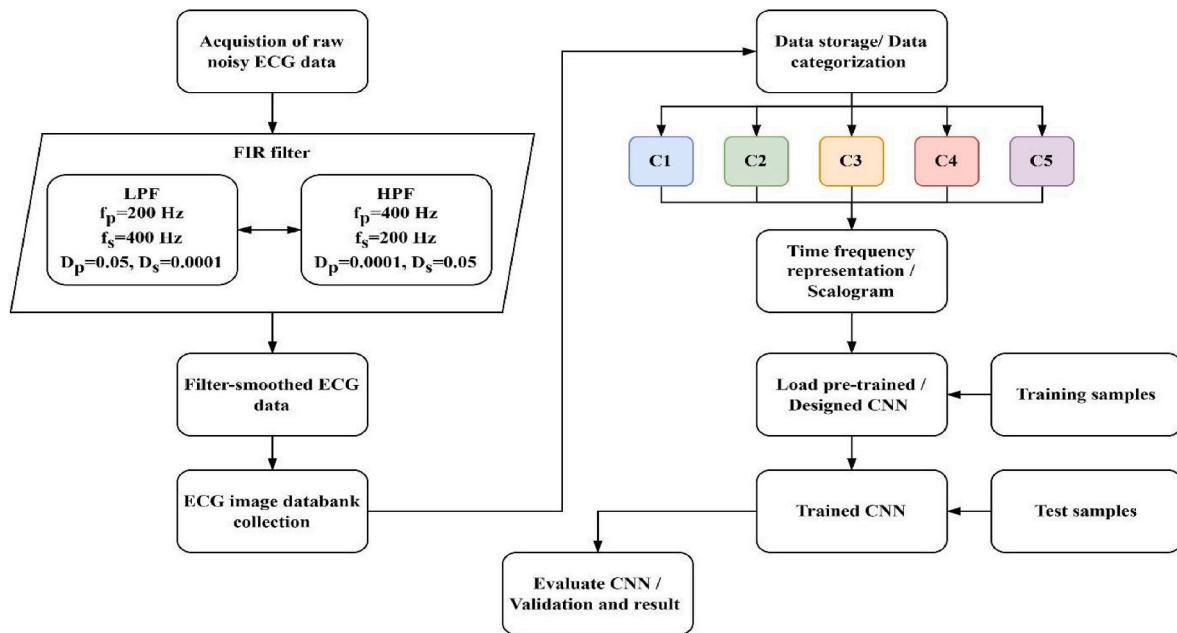


Fig. 1. Flow diagram of the proposed deep learning network for cardiovascular disease classification using five unique ECG pattern from the test ECG signal.



Fig. 2. Screenshot of the FIR filtering process. (a) Noisy ECG signal, and (b) Smoothed ECG signal.

extract the most relevant and useful aspects of incoming data. There is a deeper consideration of context, and the various parts of the ECG scalogram are integrated into the whole. When all of the neuron weights in a deep learning network are the same for the same feature maps, parallel network learning can be achieved, which dramatically shortens the learning time [15]. The filter sizes are input into equation (2) for each of the five convolution layers (Layer 1, Layer 3, Layer 5, Layer 7, and Layer 9).

$$L_n = \sum_{i=1}^{N-1} ECG_i F_{n-i} \quad (2)$$

Where, L_n be the nth layer, ECG_i is the signal value, and F_{n-i} be the filtering function over the subjected instant ECG signal.

The size of the feature map can be reduced by adding a max-pooling layer after each convolution layer. The filter (kernel) size attained in layers 1, 3, and 5 is determined via a brute-force approach in this study. Rectifier Linear Unit (ReLU) Activation Function [12]. The Modified parametric based AlexNet structure layer is then used to extract the temporal information from the feature maps. A temporal analysis is preferred by first segmenting and then componentizing the features obtained with the convolution and pooling method. These were

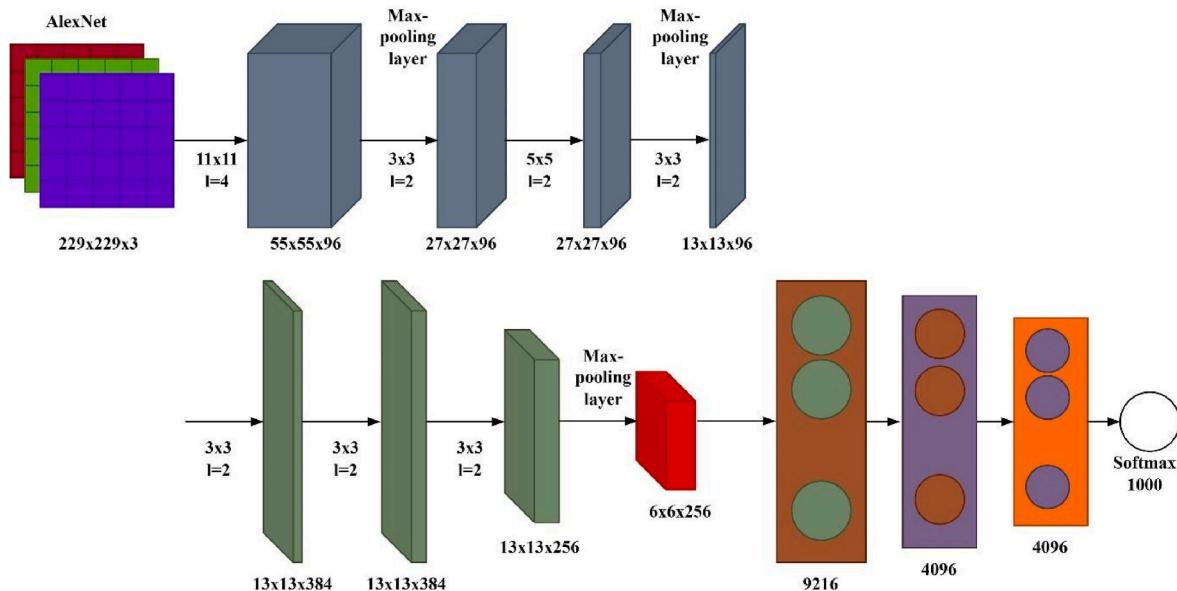


Fig. 3. Modified parametric-based AlexNet structure for classification of five ECG signals for cardiovascular diseases.

disseminated throughout the various AlexNet structure layer nodes so that they might each follow up on the findings [16]. In order to predict cardiovascular diseases, it is necessary to take into account a number of factors. The activation function () used by the AlexNet's structure layer can be seen in equation (3).

$$F_t = (W^t ECG_i, H_{t-1}, S_{t-1}^j) + B_f \quad (3)$$

Where, B_f is the biasing coefficient, W is the weighting factor to adjust the required scale down to unique pattern observation, S_{t-1} is the memory requirement in order to process the acquired time-instant data. There are 175 first layer coupled neurons, 50 s-layer, and 5 third-layer neurons. Each output class is determined by the network's ultimate soft-max function. This model was trained using backpropagation and the Adam optimizer to create the proposed Modified parametric based AlexNet structure. The regularization value in training is 0.01, and the learning rate is 1.103. The model of the deep learning network can use a wide range of batch sizes—34 in total. In order to determine the loss parameter, the definite cross-entropy function is employed. Deep learning network has been through multiple training and testing cycles. There's a cap of 100 reps per sentence. Iterative methods often include a validation phase. First, it creates ten subsets, each of which contains roughly 10% of the original dataset. Using an oversampled subset of the remaining data, it is also possible to train a deep learning network model. The second stage is also very similar to the first, except it occurs before the deep learning network model is trained.

2.3. Max-pooling layer

Nonlinear down sampling, known as "max-pooling," allows convolutional layers to achieve spatial invariance [10] by reducing the feature map's size through convolutional layers. As a result, convergence occurs more quickly, and generalization is improved.

As an illustration, a pooled feature map is created when a feature map is sent to a layer for max-pooling and that layer's operation is focused on the feature map γ . equation (4) states that the max operation selects the largest object:

$$D_j = \max_{i \in m_j} \gamma_j \quad (4)$$

Where, m_j depicts in a feature map the combined regions of j and pooled features are denoted by D_j .

2.4. Softmax

It is common practise to employ softmax regression when dealing with multiclassification issues. equation (5) depicts the hypothesis function:

$$H_\psi(a) = \frac{1}{1 + e^{(-\psi K_a)}} \quad (5)$$

The cost function (ψ) is minimized by training ψ the model parameters. Furthermore, the cost function is derived from the poison model [16].

$$J(\psi) = -\frac{1}{p} \left[\sum_{i=1}^p \sum_{j=1}^Q \{b^{(i)} = j\} \log p(b^{(i)} = j | a^{(i)}; \psi) \right] \quad (6)$$

The classification of an into category j is possible in SoftMax regression is given by

$$p(b^{(i)} = j | a^{(i)}; \psi) = \frac{e^{\psi_j^K a^{(i)}}}{\sum_{j=1}^Q e^j} \quad (7)$$

3. Performance evaluation

This section describes the performance evaluation of the proposed modified parametric based deep learning network for classifying five different ECG signals of cardiovascular diseases. As shown in Fig. 4, cardiovascular diseases cause organ discomfort, which manifests as a condensed form of five separate ECG signals. Time frequency representation provides the ECG signal scalogram which is added to this separate dataset on the existing ECG signal database. Then, all the images in each category are split in half again, this time into training samples (80% of the total data) and test samples (20% of the test data). Four different machine learning classifiers (Support Vector Machine, Support Vector Regression, Squeeze Network, and Google Network) are evaluated alongside the DWT to improve the process of obtaining a distinctive ECG signal pattern for the varying ECG signal caused by internal organ discomfort. The effectiveness of image classification based on textural characteristics is evaluated by comparing the proposed method to these existing ones, specifically about the textural boundaries

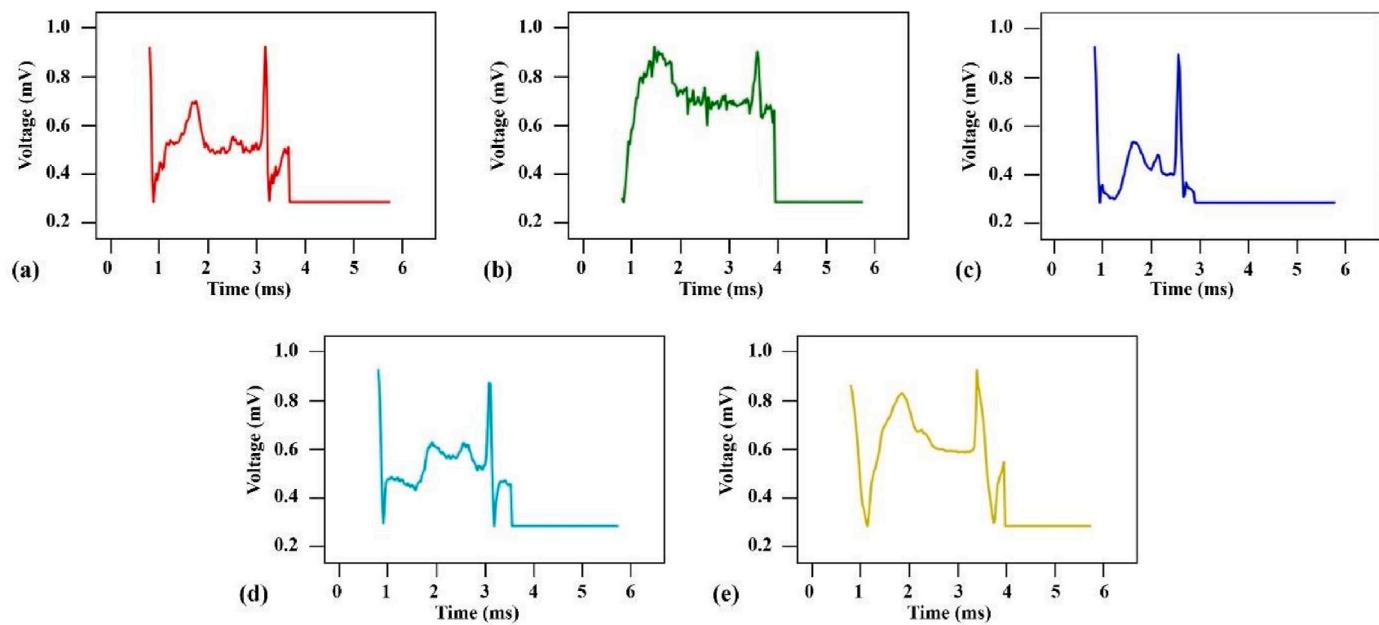


Fig. 4. A compressed form of five unique ECG signals is observed from the cardiovascular diseases due to organ discomfort. (a) Atrial fibrillation, (b) Atrial flutter, (c) Premature ventricular contraction, (d) Diastolic heart failure, and (e) Systolic heart failure.

of the ECG scalogram image. When compared to the DWT-Random Forest [9], DWT-SVM [12], Squeeze Net [15], and Google Net [16], the proposed deep learning network proves to be more effective at classifying data from standard datasets and real time dataset collected from the SRM Hospital. Because of the improvements made to the textural-based feature selection method within the deep-CNN structure and the reduction of noisy occupancy, the built-in AlexNet classifier did well in recognising the unique pattern class. Additionally, the feature extracted from images of ECG scalograms in a variety of formats (rotatory and fully scaled down) is distinguished and compared to trained feature sets. In order to classify the various sets of images, a modified parametric based deep learning network is trained with additional textural features sufficient for tracking field distribution and estimating similar class gaps.

3.1. Evaluation metrics

Calculating recall (R), prediction (P), and accuracy (η) are key performance indicators. The mathematical expression for the three features necessary for this assessment—true positive (TP), false negative (FN), and true negative (TN)—is as follows.

$$\% \text{ of } \eta = \left(\frac{TP_1 + TP_2 + \dots + TP_n}{\sum TP + FP + TN + FN} \right) \times 100 \quad (8)$$

$$\% \text{ of } R = \left(\frac{TP}{TP + FN} \right) \times 100 \quad (9)$$

$$\% \text{ of } P = \left(\frac{TP}{TP + FP} \right) \times 100 \quad (10)$$

A deep learning network is used in the proposed structure for AlexNet to complete parametric updating. This network is also used for classifying ECG signals and extracting features to improve disease classification. Totally 5,655 recordings of a single lead electrocardiogram were used. Initially, Continuous Wavelet Transform (CWT) was used to convert all 30-s signals to RGB images. After that, those images were fed to a transferred AlexNet that was trained using slightly modified parameters. Using the various deep learning networks, the confusion metric for predicting all five classes over the available datasets is graphically displayed in Fig. 5.

Maximum prediction and detection percentages are achieved even with noisy ECG signals using the proposed modified parametric based Alexnet structure. Thus, with the aid of training samples, it achieves nearly perfect class detection, allowing for the clear detection of the inaudible textural boundaries.

3.2. Result discussion

Because of the discomfort felt by a variety of internal organs, the five distinct ECG signals can be identified from the portion of the heart that is located inside the lungs. When determining which training samples to use, these signals clearly understand the extract features that are most significantly distinct from one another. It is contingent on the number of hidden features that are involved in each deep learning network to describe the actual classification outcome. According to this point, several different deep learning networks exhibit variation in the classification task. These networks include DWT-random Forest ($R = 95.97$, and $P = 96.52$), DWT-SVM ($R = 95.01$, and $P = 95.71$), SqueezeNet ($R = 95.18$, and $P = 95.57$), GoogleNet ($R = 97.95$, and $P = 98.29$), and AlexNet [Proposed] ($R = 98.08$, and $P = 97.19$), respectively.

In Table 2 and Table 3, we compare the performance of all existing deep learning networks using the Recall and Precision metric. According to the findings, it can be deduced that the proposed modified parametric based AlexNet structure achieved better classification results when applied to all five different ECG signals. The newly proposed deep learning network outperforms other already existing networks in terms of performance metrics like recall and precision, according to the findings of a comparative analysis of the two types of systems. Based on this, it can be seen that there are significant differences between the existing deep learning network and the proposed network. That implies, DWT-Random Forest ($R = 2.11$, and $P = 0.68$), DWT-SVM ($R = 3.07$, and $P = 1.48$), Squeeze Net ($R = 2.89$, and $P = 1.63$), and Google Net ($R = 0.13$, and $P = -1.10$), respectively.

In order to quantify the specific ECG signal, normalised voltage values are calculated. The proposed deep learning network, however, yields a highly accurate classifier based on a modified parametric AlexNet structure. The proposed AlexNet makes use of a potent compact discriminant feature, the curvature of the scalogram images, which is derived from the time frequency representation of images. Table 4 depicts accuracy evaluations of all known deep learning networks and

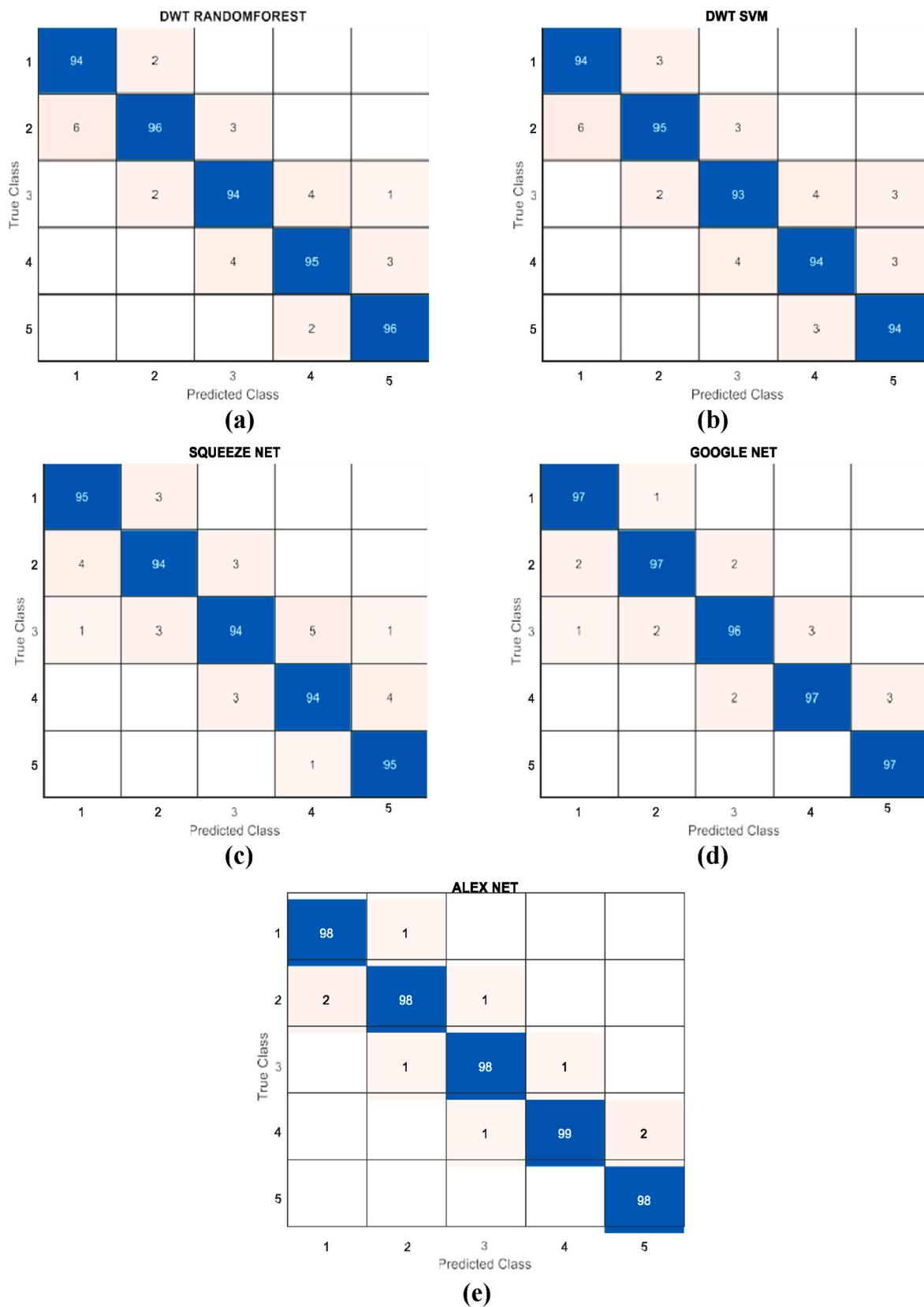


Fig. 5. Evaluation metrics. (a) DWT-Random Forest, (b) DWT-SVM, (c) Squeeze Net, (d) Google Net, and (e) AlexNet (Proposed).

Table 2

Evaluation of every existing deep learning network with respect to the Recall metric.

Type of ECG signal	DWT Random Forest [9]	DWT SVM [12]	Squeeze Net [15]	Google Net [16]	Alex Net [Proposed]
C1	0.95546	0.96152	0.96556	0.97667	0.98182
C2	0.97667	0.97263	0.94334	0.97869	0.98889
C3	0.94738	0.93526	0.93728	0.96253	0.98192
C4	0.95647	0.93324	0.94536	0.97768	0.98596
C5	0.96253	0.94738	0.96758	0.98192	0.99495
Avg.	0.95972	0.95006	0.95184	0.97948	0.98078

Table 3

Evaluation of every existing deep learning network with respect to the Precision metric.

Type of ECG signal	DWT Random Forest [9]	DWT SVM [12]	Squeeze Net [15]	Google Net [16]	Alex Net [Proposed]
C1	0.96586	0.97687	0.98586	0.98495	0.97697
C2	0.92213	0.92516	0.92718	0.95546	0.98192
C3	0.96556	0.93021	0.91001	0.94132	0.97788
C4	0.94334	0.94536	0.94334	0.98596	0.97586
C5	0.99889	0.98778	0.96192	0.99697	0.97697
Avg.	0.965156	0.957076	0.955662	0.982932	0.97192

Table 4

Evaluation of every existing deep learning network with respect to the % of accuracy.

Type of ECG signal	DWT Random Forest [9]	DWT SVM [12]	Squeeze Net [15]	Google Net [16]	Alex Net [Proposed]
C1	97.78	98.18	98.98	99.38	98.52
C2	96.96	97.77	97.16	98.97	97.71
C3	94.93	95.34	94.13	96.54	98.57
C4	95.22	94.74	93.42	95.94	98.29
C5	95.23	96.55	94.43	96.75	99.19
Avg.	95.42	96.52	95.62	97.72	98.85

proposed AlexNet. The percentage of classification accuracy of several existing deep learning networks are given by DWT-random Forest ($\eta = 95.42$), DWT-SVM ($\eta = 96.52$), SqueezeNet ($\eta = 95.62$), GoogleNet ($\eta = 97.72$), and AlexNet [Proposed] ($\eta = 98.85$), respectively. However, there exists significant differences between the existing deep learning network and the proposed network such as DWT- Random Forest (Diff = 3.43), DWT-SVM (Diff = 2.33), Squeeze Net (Diff = 3.23), and Google Net (Diff = 1.13), respectively. Overall accuracy performance of the proposed AlexNet in comparison to other deep learning networks is shown in Fig. 6.

4. Conclusion

This study aims to replace human feature identification with a deep learning-based strategy. It used a deep learning network to classify ECG signals into cardiovascular disease subcategories based on the unique pattern observed from the trained feature sample. Parametric values are adjusted from the time frequency representation or scalogram. It uses textural shape features to distinguish the ECG signal and train the deep learning model with the labeled dataset. These time frequency representations can transfer knowledge to other ECG recognition tasks when data is insufficient for deep architecture training. Around 5,655 single-lead ECG recordings were used to train a deep learning network. Using Continuous Wavelet Transform (CWT), All 30-s signals were transformed into RGB images, these were then fed into a transferred

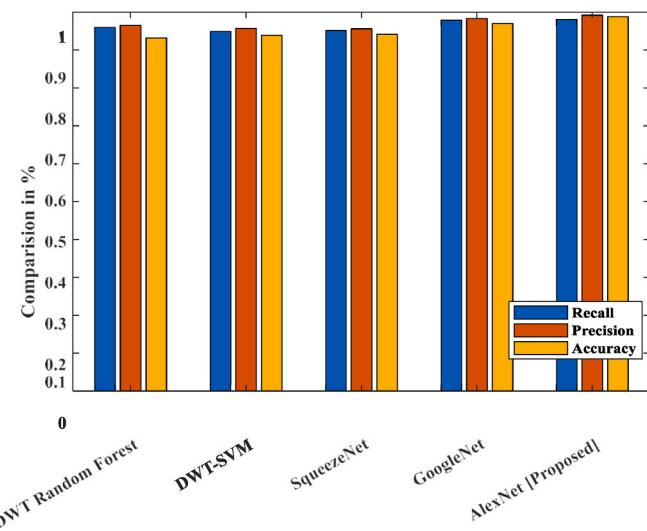


Fig. 6. Proposed AlexNet performance compared to other deep learning networks.

AlexNet that had been trained with slightly adjusted parameters. The simulation results conveyed that the modified parametric-based Alexnet structure classified five ECG signals better around 98.82% accuracy. The proposed deep learning network outperforms other networks in recall (98.9%) and precision (97.9%). Results show that compared to state-of-the-art methods, proposed AlexNet is more accurate in detecting individual ECG signal class and maintain maximum percentage of classification accuracy. Hence, it shows DWT-Random Forest ($R = 2.11$, $P = 0.68$, and $Diff = 3.43$), DWT-SVM ($R = 3.07$, $P = 1.48$, and $Diff = 2.33$), Squeeze Net ($R = 2.89$, $P = 1.63$, and $Diff = 3.23$), and Google Net ($R = 0.13$, $P = -1.10$, and $Diff = 1.13$), respectively.

Credit author statement

Ms. T. Aarthy: Plan and Methodology, Conceptualization, Formal Survey, and Writing Dr.J.L.Mazher Iqbal: Review, Editing, Validation, and Final Drafting.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- B. Salami, A Deep Learning (Neural Network) Approach for ECG Heartbeat Classification: Arrhythmia Detection, 2020.
- D. Zhang, Y. Chen, Y. Chen, S. Ye, W. Cai, M. Chen, An ECG heartbeat classification method based on deep convolutional neural network, Journal of Healthcare Engineering (2021), 2021.
- M. Degirmenci, M. Ozdemir, E. Izci, A. Akan, ECG Heartbeat Classification Based on Signal- To-Image Transformation Using Convolutional Neural Networks, 2020.
- J. Huo, X. Li, J. Liu, J. Tang, J. Rao, H. Deng, A novel arrhythmia classification of electrocardiogram signal based on modified HRNet and ECA, Meas. Sci. Technol. 33 (2022).
- H.A. Deepak, T. Vijayakumar, ECG signal classification with hybrid features using bayesian optimized KNearest neighbors classifier, International Journal of Intelligent Engineering and Systems (2021).
- Z. Liu, G. Yao, Q. Zhang, J. Zhang, X. Zeng, Wavelet scattering transform for ECG beat classification, Comput. Math. Methods Med. (2020), 2020.
- Y.N. Fuadah, K.M. Lim, Optimal classification of atrial fibrillation and congestive heart failure using machine learning, Front. Physiol. 12 (2021).

- [8] W. Zeng, J. Yuan, C. Yuan, Q. Wang, F. Liu, Y. Wang, A novel technique for the detection of myocardial dysfunction using ECG signals based on hybrid signal processing and neural networks, *Soft Comput.* 25 (2021) 4571–4595.
- [9] M. Wu, Y. Lu, W. Yang, S.Y. Wong, A study on arrhythmia via ECG signal classification using the convolutional neural network, *Front. Comput. Neurosci.* 14 (2021).
- [10] P.S. Choudhary, S. Dandapat, An evaluation of machine learning classifiers for detection of myocardial infarction using wavelet entropy and eigenspace features, in: 2020 IEEE Applied Signal Processing Conference (ASPICON), 2020, pp. 222–226.
- [11] B. Fatimah, P. Singh, A. Singhal, D. Pramanick, P. Shivashankaran, R.B. Pachori, Efficient detection of myocardial infarction from single lead ECG signal, *Biomed. Signal Process Control* 68 (2021), 102678.
- [12] A.B. Slama, H. Sahli, R. Maalmi, H. Trabelsi, ConvNet: 1D-convolutional neural networks for cardiac arrhythmia recognition using ECG signals, *Trait. Du. Signal* 38 (2021) 1737–1745.
- [13] N. Omar, M. Dey, M.A. Ullah, Detection of myocardial infarction from ECG signal through combining CNN and Bi-LSTM, in: 2020 11th International Conference on Electrical and Computer Engineering (ICECE), 2020, pp. 395–398.
- [14] A.J. Khalaf, S.J. Mohammed, Verification and comparison of MIT-BIH Arrhythmia database based on number of beats, *Int. J. Electr. Comput. Eng.* 11 (2021) 4950–4961.
- [15] P. Plawiak, M. Abdar, Novel methodology for cardiac arrhythmias classification based on long-duration ECG signal fragments analysis, *Series in BioEngineering* (2019).
- [16] H.A. Deepak, T. Vijaykumar, ECG beat classification using CNN, in: 2022 IEEE International Conference on Data Science and Information System (ICDSIS), 2022, pp. 1–5.
- [17] L. Fang, Construction of Physical Education Quality Evaluation Index and Analysis with Wearable Device, *Computational Intelligence and Neuroscience*, 2022, 2022.
- [18] M.H. Mazidi, M. Eshghi, M.R. Raoufy, Detection of premature ventricular contraction (PVC) using linear and nonlinear techniques: an experimental study, *Cluster Comput.* 23 (2019) 759–774.
- [19] J. Yu, X. Wang, X. Chen, J. Guo, Searching for premature ventricular contraction from electrocardiogram by using one-dimensional convolutional neural network, *Electronics* (2020).
- [20] Q. Mastoi, T.Y. Wah, M.A. Mohammed, U. Iqbal, S. Kadry, A. Majumdar, O. Thinnukool, Novel DERMA fusion technique for ECG heartbeat classification, *Life* 12 (2022).
- [21] Z. Cai, T. Wang, Y. Shen, Y. Xing, R. Yan, J. Li, C. Liu, Robust PVC identification by fusing expert system and deep learning, *Biosensors* 12 (2022).

Biometrics

SBE 4022, Fall 2024

Professor

Ahmed M. Badawi

ambadawi@eng1.cu.edu.eg

<http://scholar.google.com.eg/citations?user=r9pLu6EAAAAJ&hl=en>

TA:

Mohamed Mutair

Mohamed_mutair@eng.cu.edu.eg



قسم الهندسة الحيوية
الطبية والمنظومات



جامعة القاهرة
كلية الهندسة

Existing biometric technologies

- Voice
- Fingerprints
- Face
- Iris
- Ear
- Gait
- Keystroke Dynamics
- DNA
- Signature and Acoustic Emissions
- Odor
- Retinal Scan
- Hand and Finger Geometry

Biometrics

- **Biometrics** is the science of identifying or verifying a person's identity by physical or behavioral characteristics.

- **Verification (1:1 Matching)**

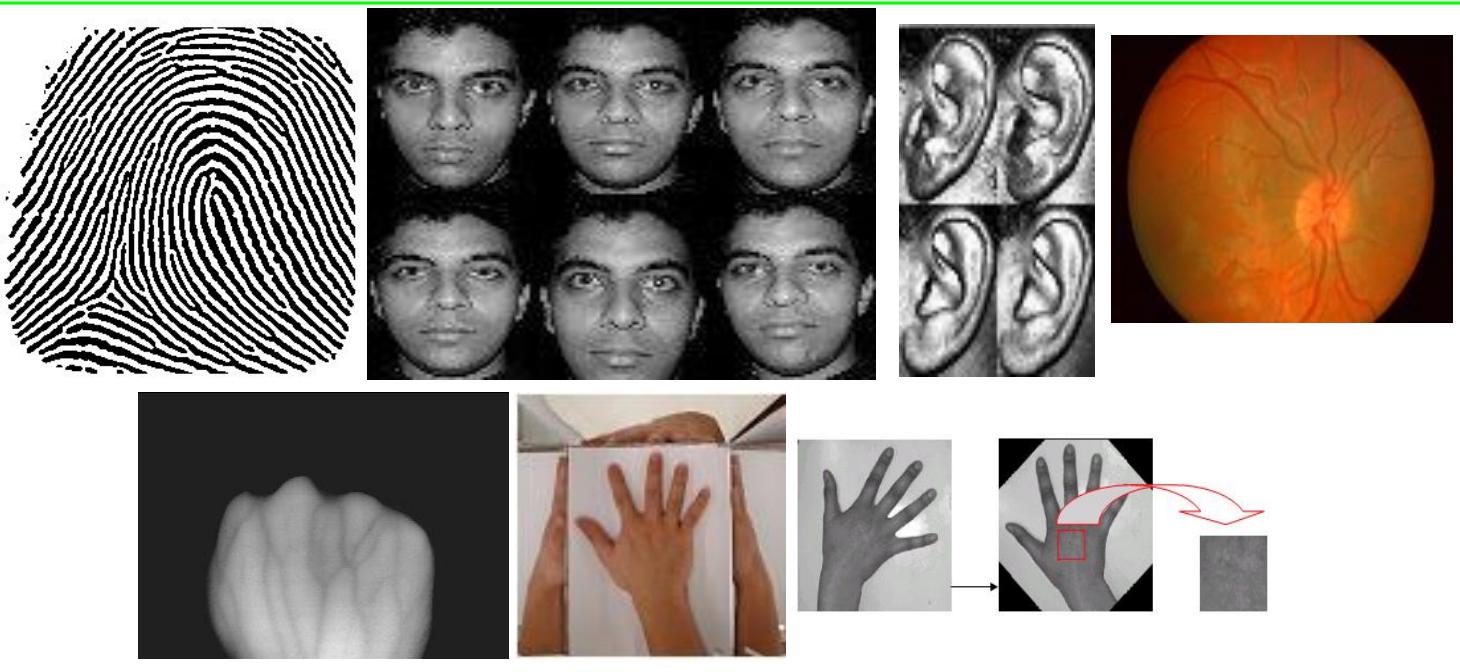
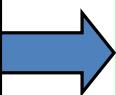
Does this biometric data belong to *this person*?

- **Identification (1:N Matching)**

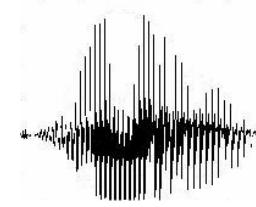
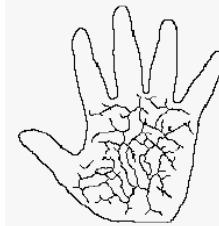
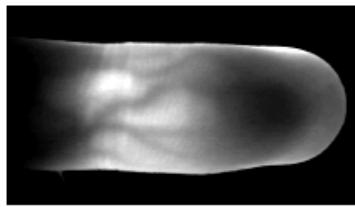
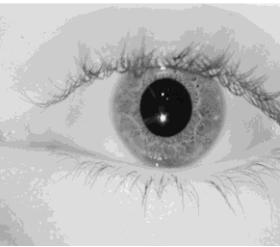
Whose biometric data is this?

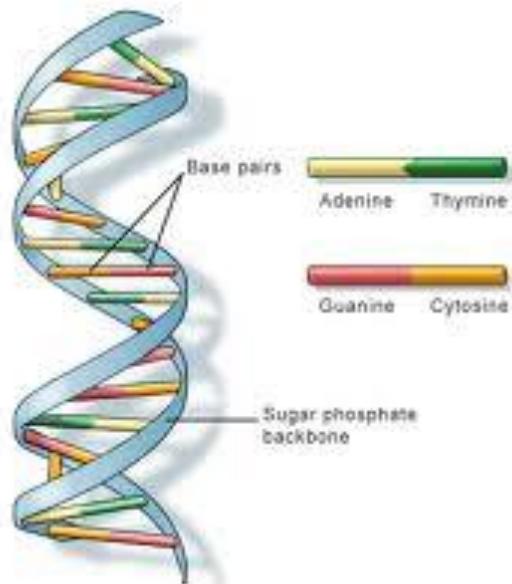
Common Biometrics Modalities (SBE/CU)

Some of our
analyzed
modalities at
SBE



Other
Modalities





U.S. National Library of Medicine



What Biometrics is all about

It is a trade off between
Individuality (FAR) vs Operability (FRR)

Which Biometric system is the best

- Universality (all users possess this biometrics)
- Uniqueness (varies across users)
- Permanence (does not change over time)
- Collectability (can be measured quantitatively)
- Performance (low error rates and processing time)
- Acceptability (is it acceptable to the users)
- Circumvention (can it be easily spoofed)

No Single Biometric modality is **optimal**

Limitations of Biometric Systems

- Noise or distortion in sensed data
- Intra-class variations
- Distinctiveness
- Non-universality
- Send-up attacks

Why Multimodal Biometrics ?

- Noise or distortion in sensed data: (A fingerprint image with a scar – Accumulation of dirt on a fingerprint sensor)
- Intra-class variations:
(Incorrect facial pose - Optical versus solid-state fingerprint sensors).
- Inter-class similarities:
(Number of distinguishable patterns for hand geometry and face are only of the order of 10^5 and 10^3 , respectively)
- Non-universality: (The biometric system may not be able to acquire meaningful biometric data from a subset of users).
- Spoof attacks: (Example: physical traits such as fingerprints and hand geometry are susceptible to spoof attacks).
- Accuracy: (Fusion of more signals increases the performance, **lower EER**)

Faking Commercial Single-Modality Systems

Using photographic or light to capture superficial prints that are seen by naked eye, easy to access, all can be faked



Fake Hands plastic or paper Models



H. Chen, H. Valizadegan, C. Jackson, S. Soltysiak and A.K. Jain, "Fake Hands: Spoofing Hand Geometry Systems", Biometric Consortium 2005, Washington DC, September 2005.

Fake Fingerprints & Palmpoints models by graphite powder on latex

<http://www.securityfocus.com/news/6717>, Accessed in June, 2006.



Merits of Multimodal Biometrics

- Overcome some of the limitations imposed by **Unimodal Biometric Systems** and expected to be more efficient.
- **Hard to fake.**
- **Multimodal Biometric Systems** integrate the evidences presented by multiple sources of information (**Simulates human eyes**).
- Example: **Face and Fingerprints used in e-Passports.**

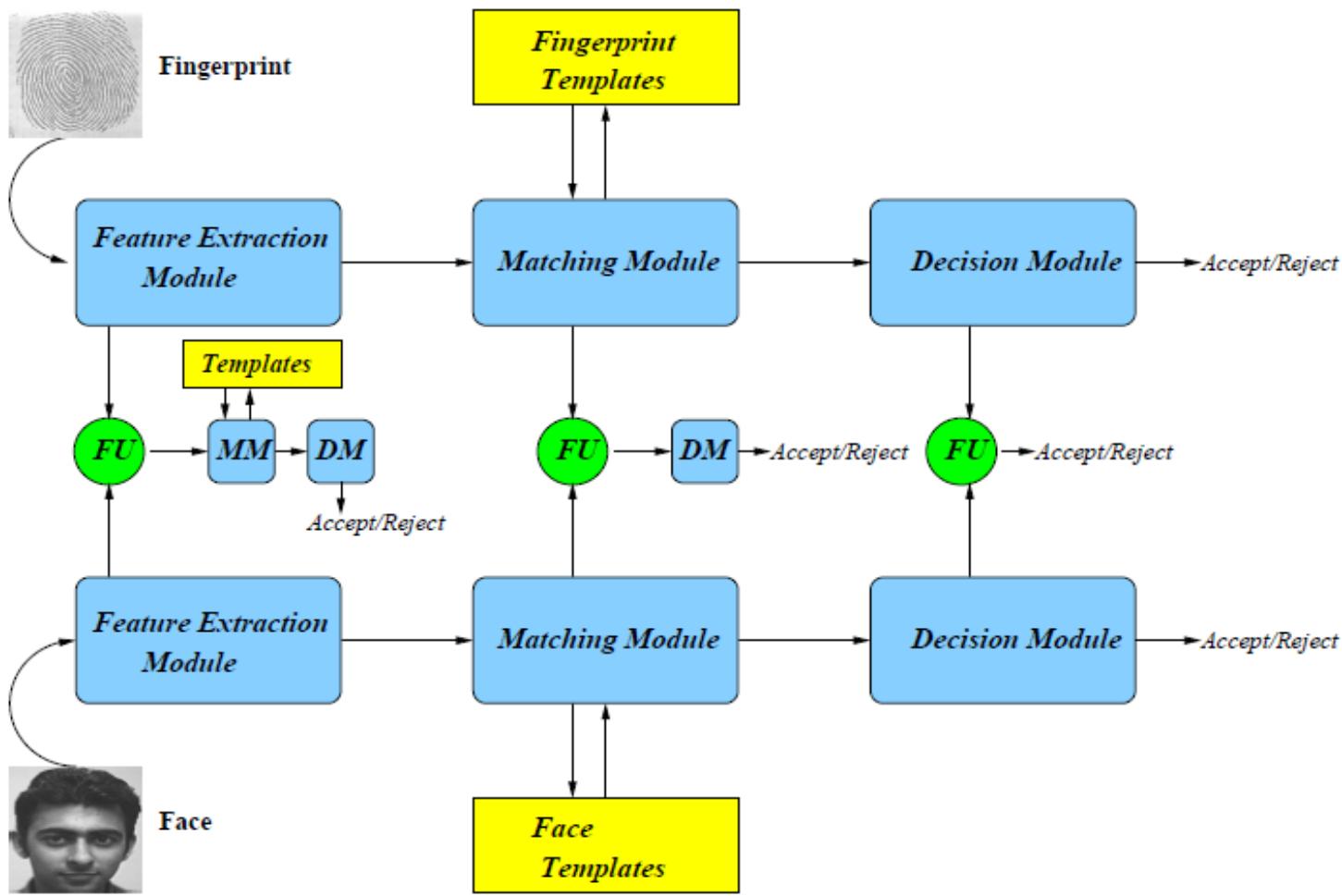


Figure 1: Levels of fusion in a bimodal biometric system;
 FU: Fusion Module, MM: Matching Module, DM: Decision Module.

Arun Ross and Anil K. Jain, "MULTIMODAL BIOMETRICS: AN OVERVIEW", Appeared in Proc. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), pp. 1221-1224, September 2004.

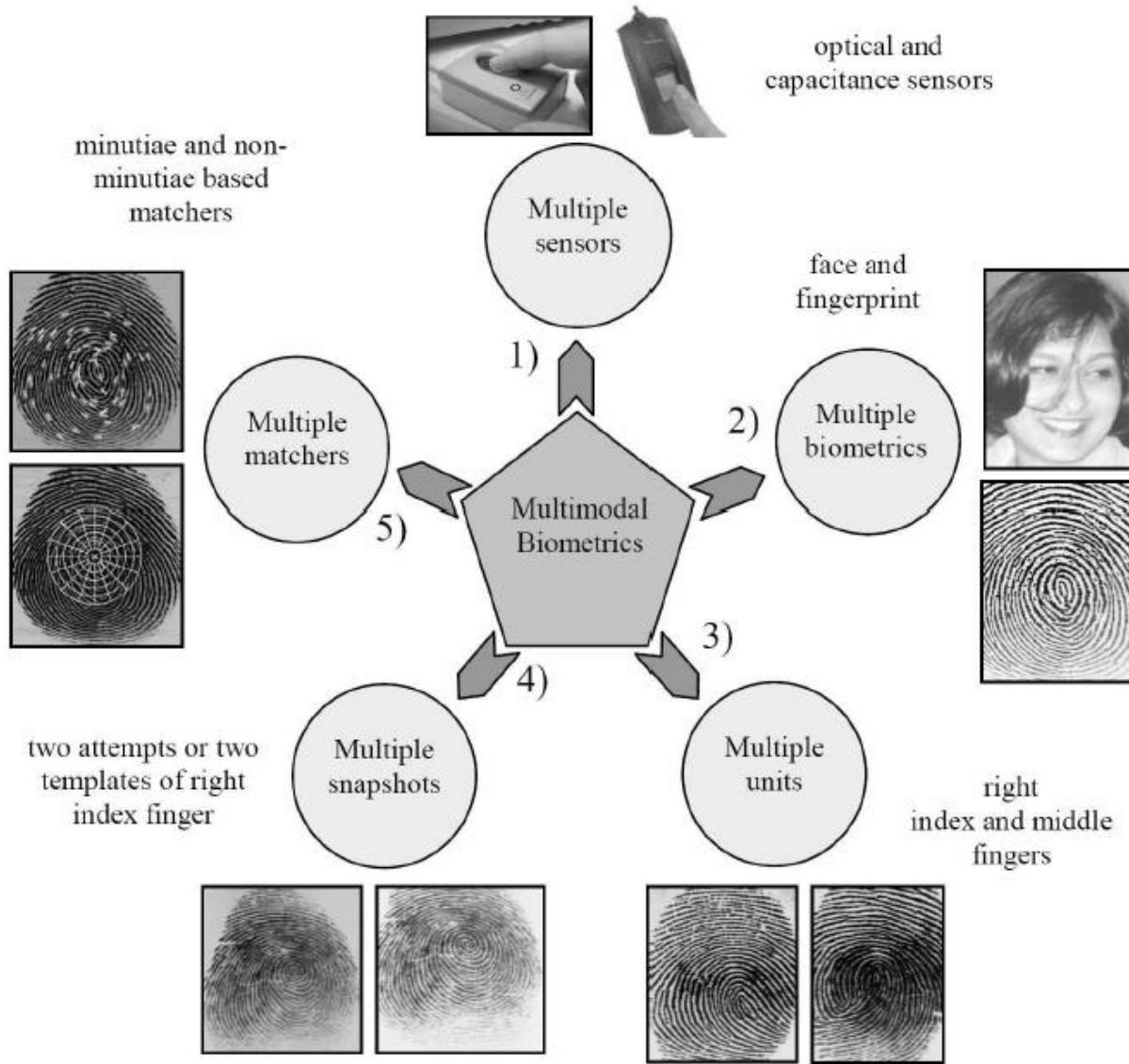


Figure 2: Scenarios in a multimodal biometric system [6].

Arun Ross and Anil K. Jain, "MULTIMODAL BIOMETRICS: AN OVERVIEW", Appeared in Proc. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), pp. 1221-1224, September 2004.

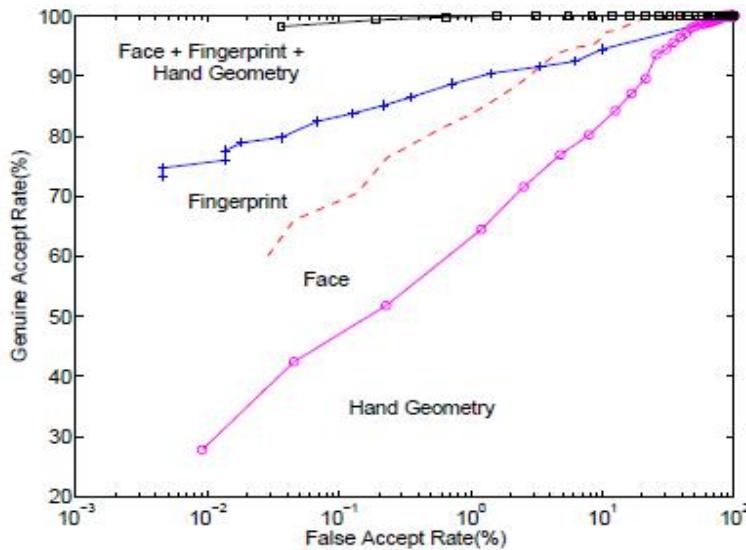


Figure 3: Performance gain using the sum rule [3].

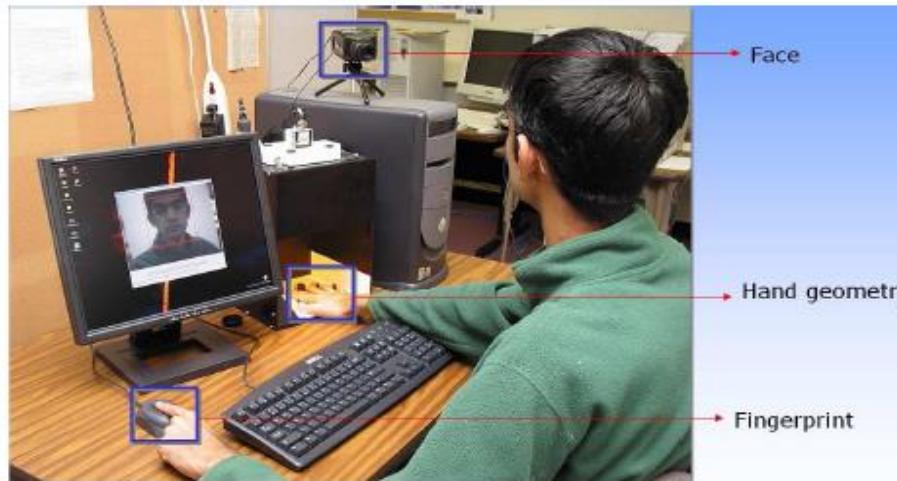
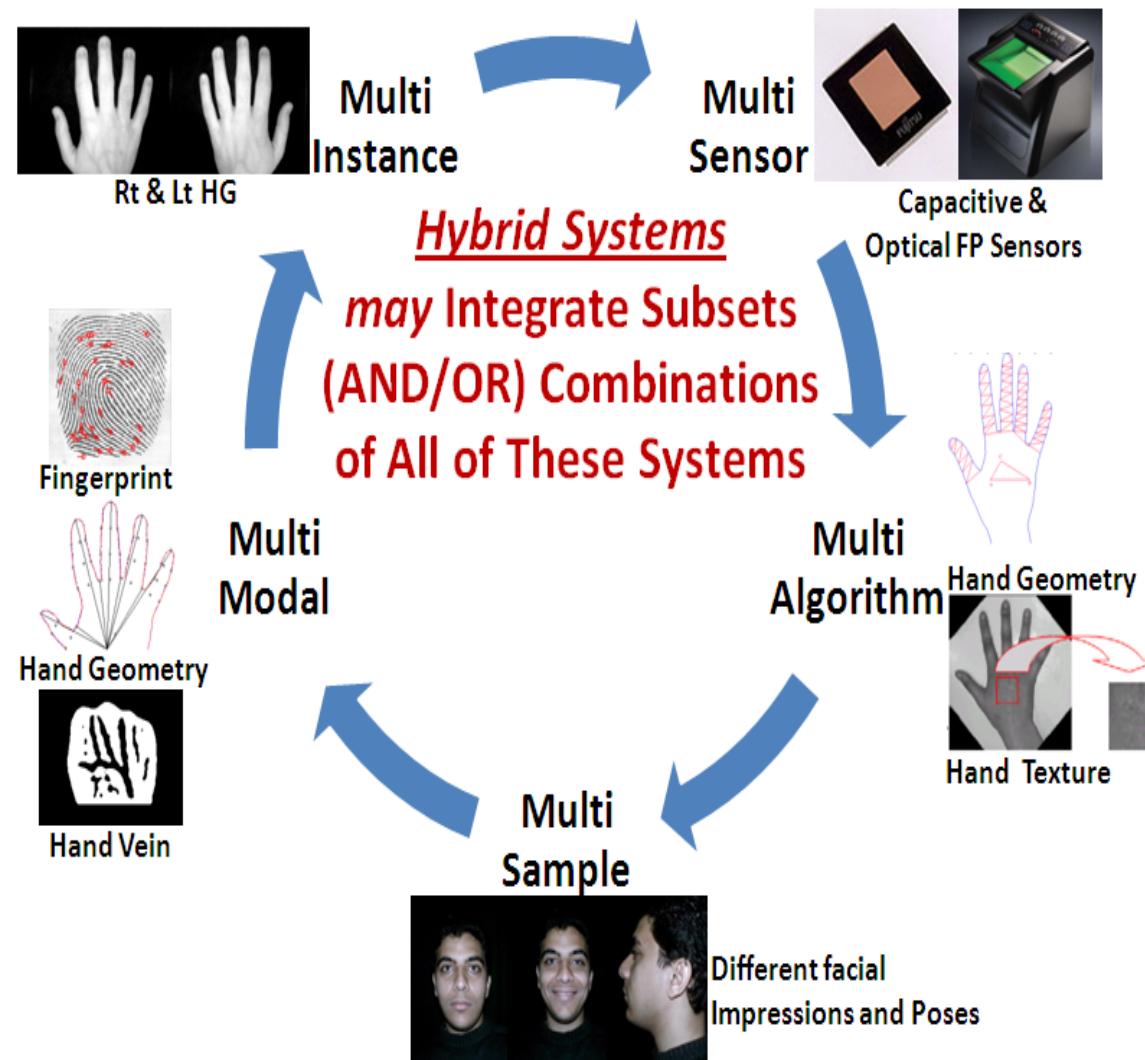


Figure 4: A prototype multimodal biometric login system.

Arun Ross and Anil K. Jain, "MULTIMODAL BIOMETRICS: AN OVERVIEW", Appeared in Proc. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), pp. 1221-1224, September 2004.

MMBS Different Structural Forms



Why Multibiometrics !

Ho, 2002 states that there has been a paradigm shift in the approach to solving pattern recognition problems:

Instead of looking for the best set of features and the best classifier, now we look for the best set of classifiers and then the best combination method.

T. K. Ho, “Multiple Classifier Combination: Lessons and Next Steps”, In H. Bunke and A. Kandel, editors, *Hybrid Methods in Pattern Recognition, vol. 47 of Machine Perception and Artificial Intelligence, World Scientific, p.p. 171-198, 2002.*

Biometric Systems Genuine/Imposter Events

Genuine Events:

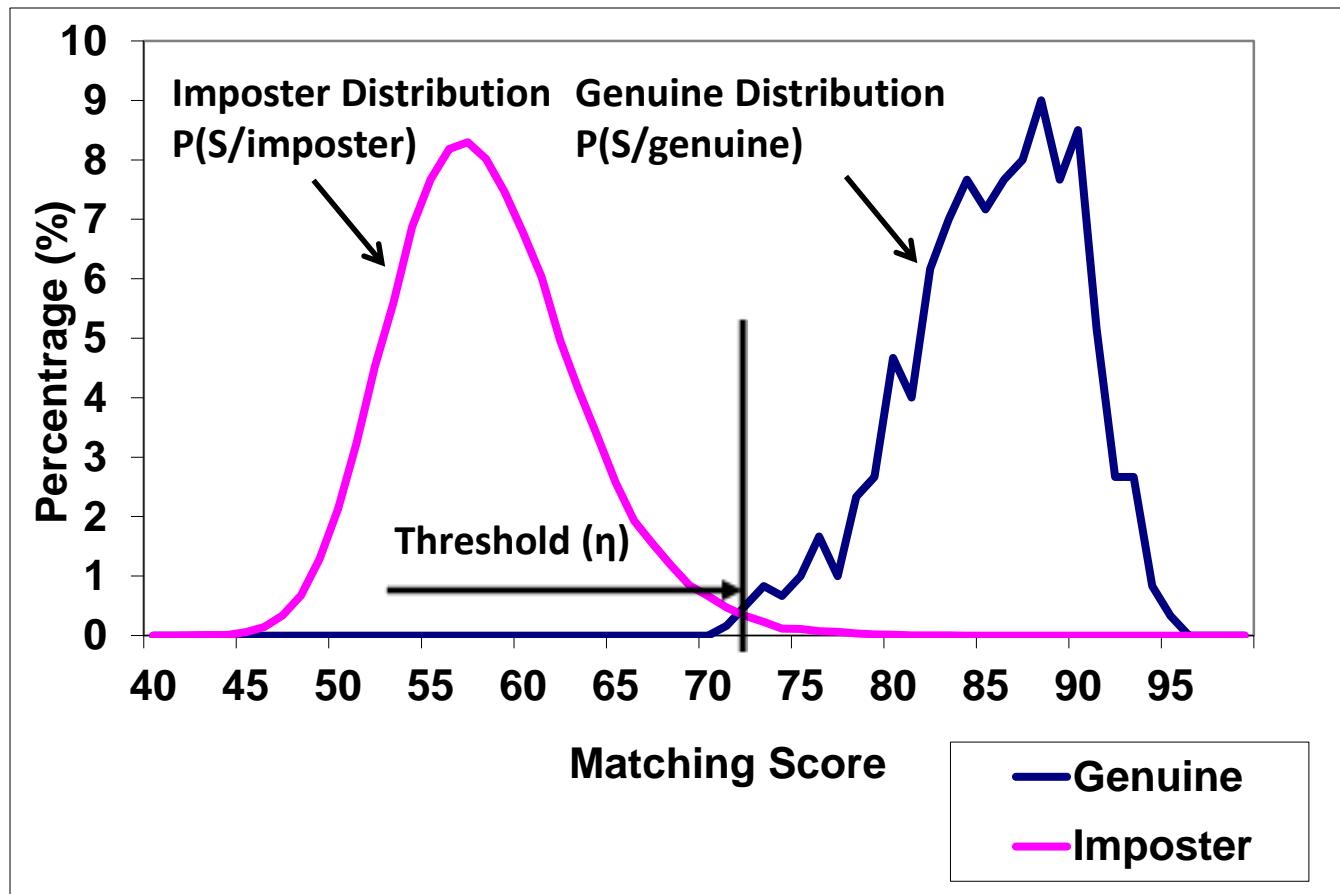
GA: Genuine Accept

GR: Genuine Reject

Imposter Events

FR: False Reject

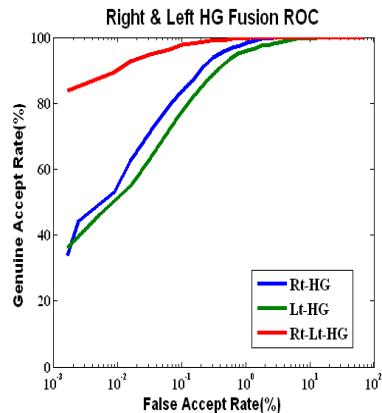
FA: False Accept



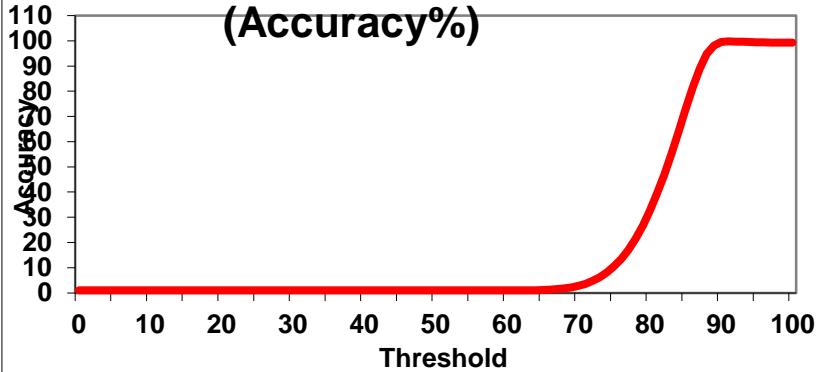
Comparing Biometric Systems Performance

Qualitative Methods

ROC Curves



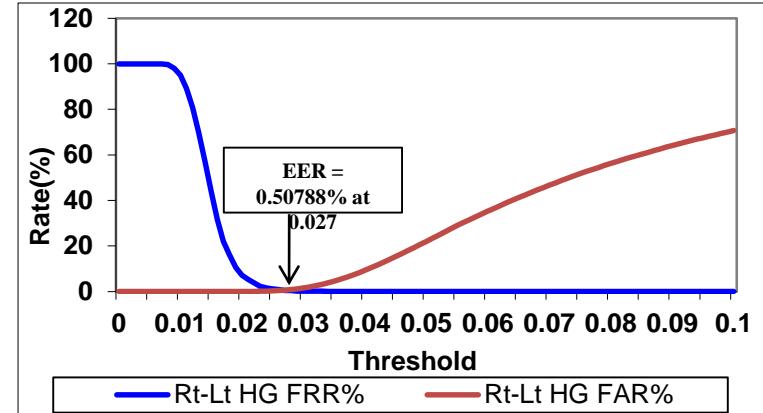
System Performance (Accuracy%)



$$\text{Accuracy (\%)} = \frac{\text{GA} + \text{GR}}{\text{GA} + \text{FR} + \text{GR} + \text{FA}} \times 100 (\%)$$

Quantitative Methods

EER: Equal Error Rate



TER: Total Error Rate

TER = minimum (FRR+FAR)

Multibiometrics

The automated recognition of individuals based on their biological or behavioral characteristics and involving the use of biometric **fusion**

Approaches to Multibiometrics

- Multimodal
- Multialgorithmic
- Multi-instance
- Multisensorial
- Hybrid (combinations of the above)

“Levels” of Fusion

- Decision
- Score
- Feature
- Sample

MMBS Examples

Example 1- Multimodal, decision level fusion with sequential sampling

Application – Attended Physical Access Control

Description – Fingerprint and Iris, “OR” logic, fingerprint first

Advantages

- High user enrollment rate (either modality policy)
- High throughput potential (finger first)
- Low false reject rate

Disadvantages

- Additional hardware costs (two sensor systems)
- Additional enrollment time
- Increased potential for false accepts

Example 2- Multi-instance, score level fusion with sequential sampling

Application – Unattended Logical or Physical Access Control

Description – Multiple Fingerprints, single digit sensor, verification

(1-to-1) with **Sum Rule Combination Fusion**

Advantages

- Low Cost (primary driver)
- Low false reject rate with high security
- Option for higher spoof resistance (using query - response – different fingers)

Disadvantages

- Longer usage time (multiple sequential samples)
- Some users unable to enroll

Example 3- Multisensorial, multialgorithmic, hybrid fusion with simultaneous sampling

Application – Token-less Identification for Privileged Access

Description – Dual sensor Face Recognition (2-D and 3-D), each with multiple matchers, individualized weighted sum rule and voting scheme

Advantages

- User satisfaction (no token to forget, no contact)
- Low failure to enroll rate
- Potentially very high accuracy identification rates

Disadvantages

- Lengthy enrollment time
- Very complex logic requiring tuning
- Costly hardware

Example of Multi-instance fingerprint-verification

Multi-instance (**L and R index**)

- Score-level fusion using simple sum rule (no normalization required)

A real Egyptian style Multi-instance Biometrics system

أبصملك بالعشريه



Emerging MM concepts

Iris – Retina

Single presentation (simultaneous capture)

Significantly cooperative user

Palm print – Hand Geometry

Single presentation (simultaneous capture)

2-D optical scanner (low cost)

Face – Iris

Single image – high resolution camera

Perception of public acceptability

Whole-Hands* (Fingerprint, Hand geometry, hand veins)

Two FP sensors, one NIR camera **Badawi et al**

Score Normalization in Multimodal Biometric Systems

Karthik Nandakumar and Anil K. Jain
Michigan State University, East Lansing, MI

Arun A. Ross

West Virginia University, Morgantown, WV

<http://biometrics.cse.mse.edu>

Multimodal Biometric Systems

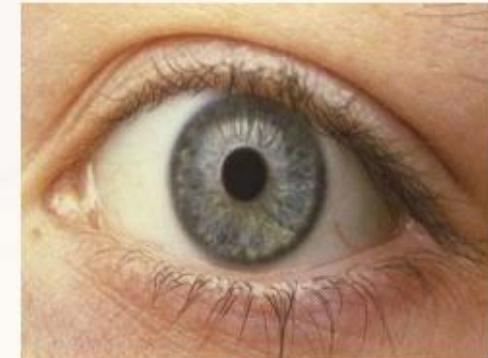
- Multiple sources of biometric information are integrated to enhance matching performance
- Increases population coverage by reducing failure to enroll rate
- Anti-spoofing; difficult to spoof multiple traits simultaneously



Fingerprint

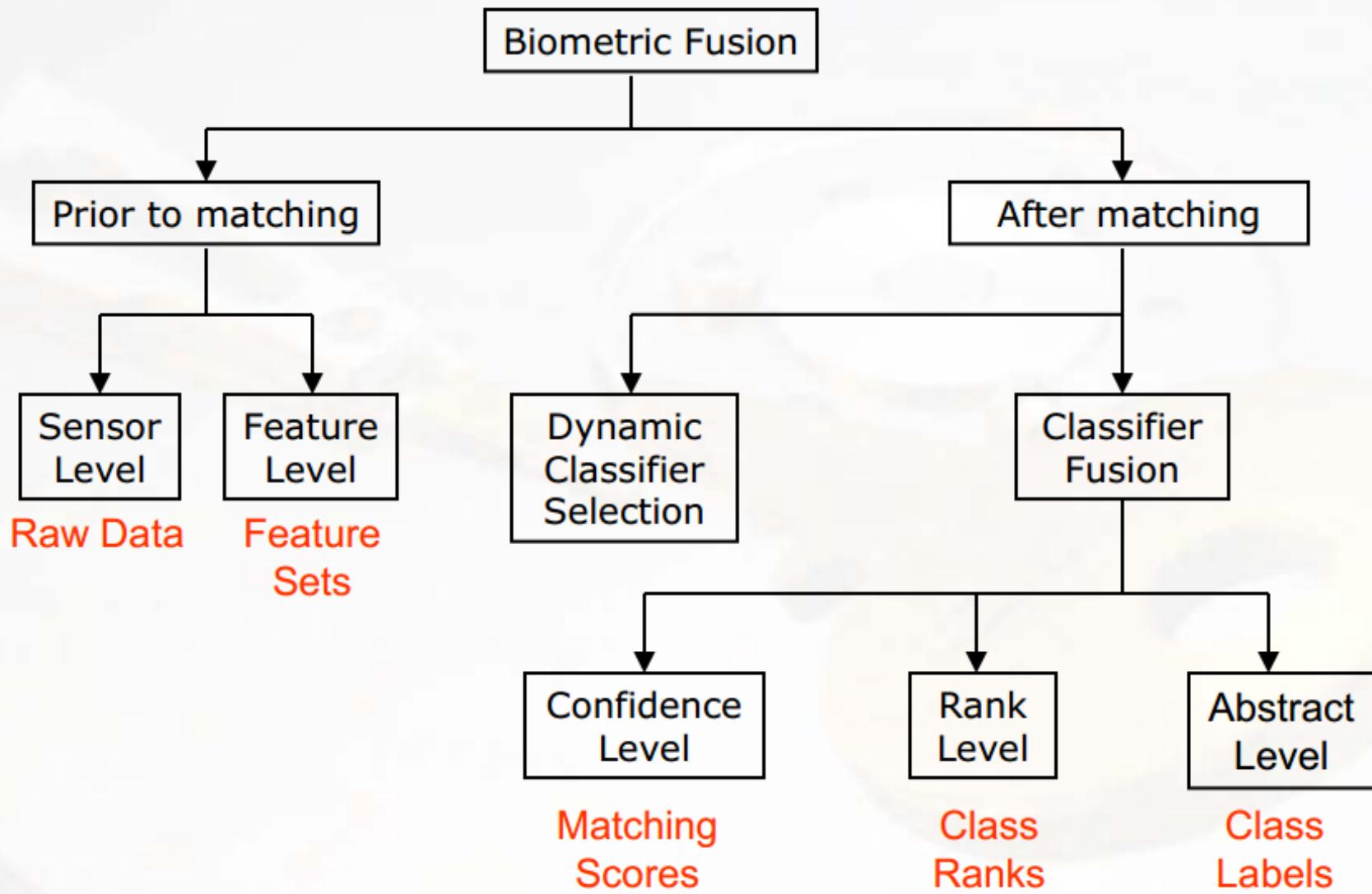


Face



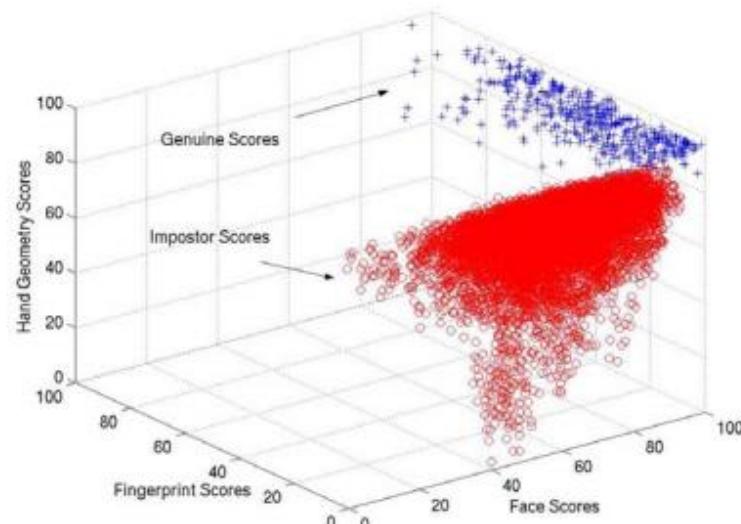
Iris

Fusion in Multimodal Biometrics

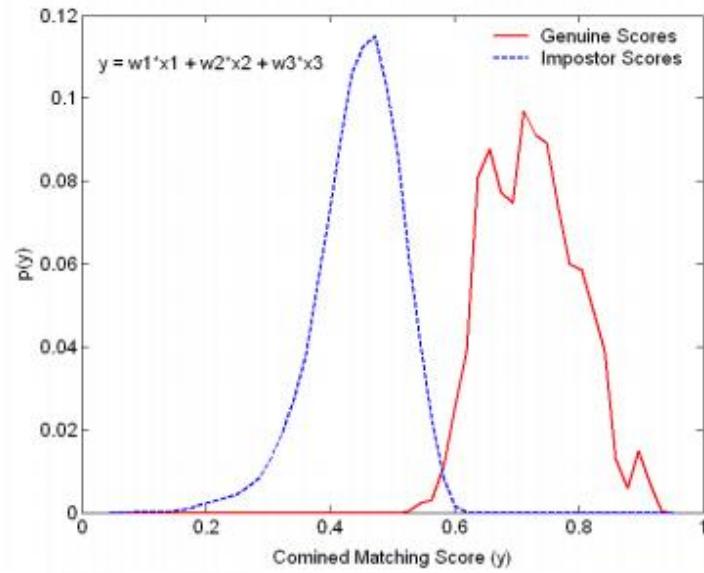


Fusion at the Matching Score Level

- Fusion at the matching score level offers the best tradeoff in terms of information content and ease in fusion



Classification Approach



Combination Approach

- Experiments indicate that the combination approach performs better than the classification approach*

*A. Ross, A.K. Jain, "Information Fusion in Biometrics", Pattern Recognition Letters, Sep. 2003

Fusion Rules*

- Problem: Classify input pattern Z into one of m possible classes (c_1, \dots, c_m) based on evidence provided by R classifiers
- Let \mathbf{x}_i be the feature vector for the i^{th} classifier derived from Z ; \mathbf{x}_i 's are independent
- Assign $Z \rightarrow c_j$, if $g(c_j) \geq g(c_k)$, $1 \leq k \leq m$, $k \neq j$
 - **Product rule:** $g(c_r) = \prod_{i=1}^R P(c_r | x_i)$
 - **Sum rule:** $g(c_r) = \sum_{i=1}^R P(c_r | x_i)$
 - **Max rule:** $g(c_r) = \max_i P(c_r | x_i)$
 - **Min rule:** $g(c_r) = \min_i P(c_r | x_i)$

*Kittler et al., "On Combining Classifiers", IEEE PAMI, March 1998

- **Neural Networks** : Train a MMBPN to weight classes/scores (Khairy/Badawi et al 2010)

Computing the Posteriori Probability

- To make use of the fusion rules, we need to compute $P(c_j | \mathbf{x}_i)$
- At the matching score level, we have only the score s_{ij} (not \mathbf{x}_i)
- Verlinde et al. proposed that

$$s_{ij} = f(P(c_j | \mathbf{x}_i)) + \eta(\mathbf{x})$$

- f is a monotonic function
- η is the error introduced by the biometric system due to problems in acquisition and feature extraction processes
- To estimate η , the biometric system should output a matching score along with a confidence measure on that score indicating the quality of the input feature vector \mathbf{x}_i
- If η is known, we can estimate $P(c_j | \mathbf{x}_i)$ from s_{ij} using non-parametric density estimation methods (e.g., Parzen window)

Score Normalization

- If η is unknown, the errors in the estimation of $P(c_j | \mathbf{x}_i)$ will be very large; hence, it is better to combine the scores directly
- Combination of scores has the following problems:
 - Non-homogeneous scores: distance or similarity
 - Ranges may be different; e.g., [0,100] or [0,1000]
 - Distributions may be different
 - Modify the location and scale parameters of score distributions of individual matchers to transform the scores into a common domain
- Robustness: Should not be affected by the outliers
- Efficiency: Estimated parameters of the score distribution should be close to the true values

Score Normalization Techniques

- **Min-max normalization:** Given matching scores $\{s_k\}$, $k=1,2,\dots,n$ the normalized scores are given by:

$$s' = \frac{s - \min\{s_k\}}{\max\{s_k\} - \min\{s_k\}}$$

- **Decimal scaling:** Used when scores of different matchers differ by a logarithmic factor; e.g., one matcher has scores in the range [0,1] and the other matcher has scores in the range [0, 1000]

$$s' = \frac{s}{10^n},$$
$$n = \log_{10} \max\{s_k\}$$

Score Normalization Techniques

- Z-score:

$$s' = \frac{s - \mu}{\sigma}$$

- Median and Median Absolute Deviation (MAD):

$$s' = \frac{(s - \text{median})}{MAD}$$

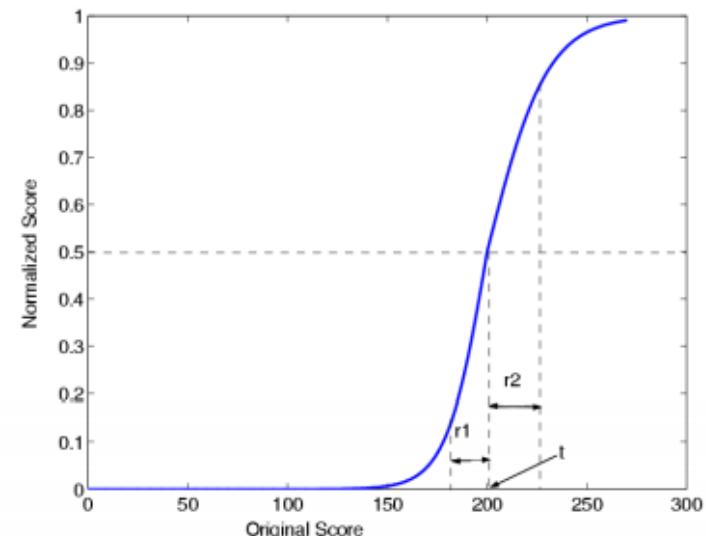
$$MAD = \text{median}(|\{s_k\} - \text{median}|)$$

- Double Sigmoid function:

$$s' = \frac{1}{1 + \exp\left(-2\left(\frac{s-t}{r}\right)\right)}$$

$$r = r_1, \text{ if } s < t$$

$$r = r_2, \text{ otherwise}$$

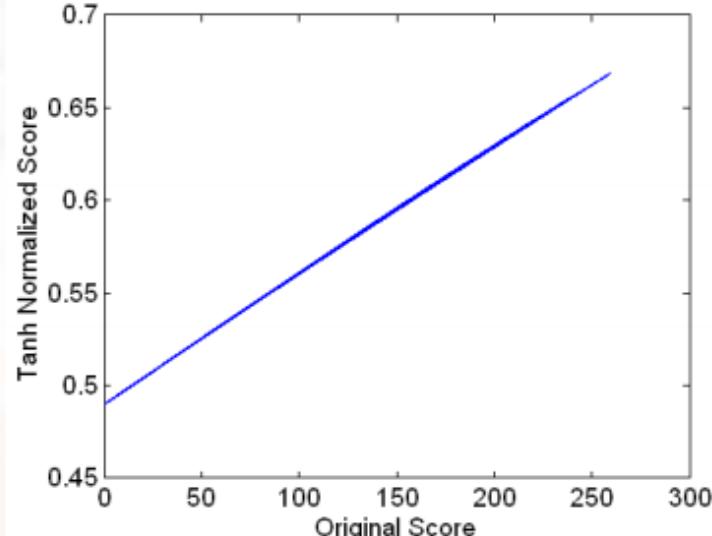


Score Normalization Techniques

- Tanh estimators:

$$s' = 0.5 \left[\tanh \left(0.01 \frac{(s - \mu_{GH})}{\sigma_{GH}} \right) + 1 \right],$$

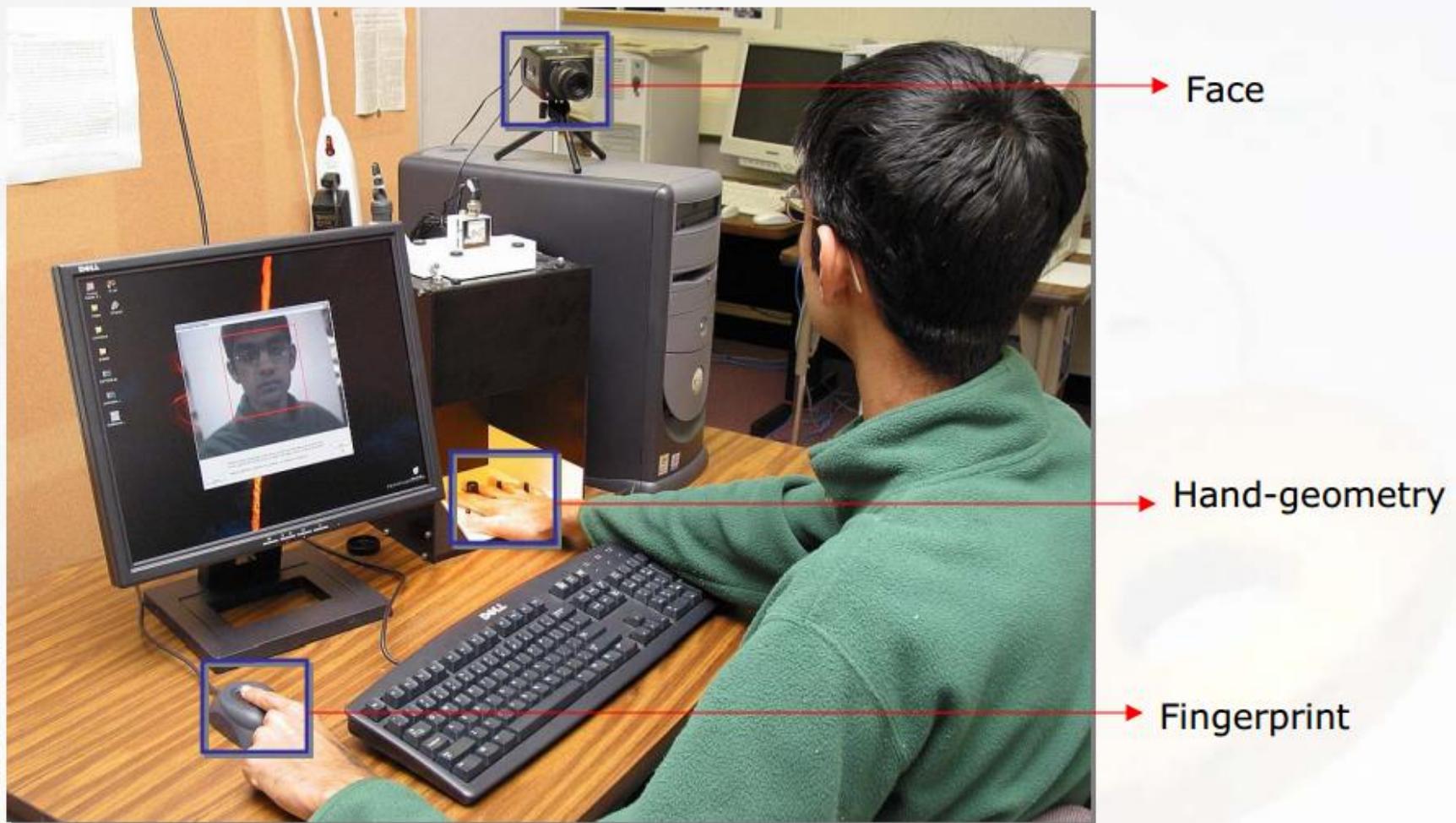
where μ_{GH} and σ_{GH} are the mean and standard deviation estimates of the genuine score distribution as given by Hampel estimators*



- Min-max, Z-score, and Tanh normalization schemes are efficient
- Median, Double Sigmoid, and Tanh methods are robust

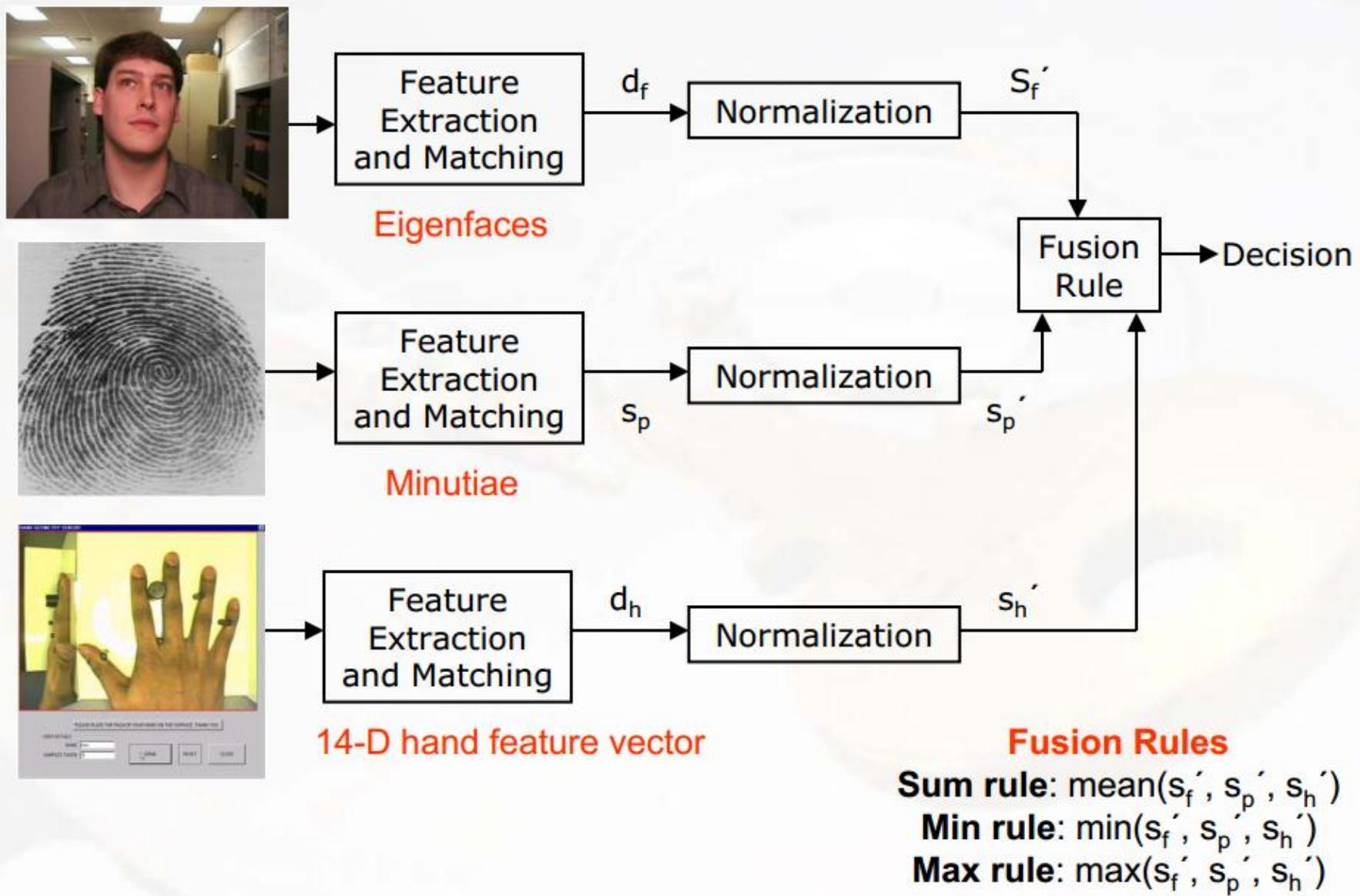
*Hampel et al., *Robust Statistics: The Approach Based on Influence Functions*, 1986

Experimental Setup

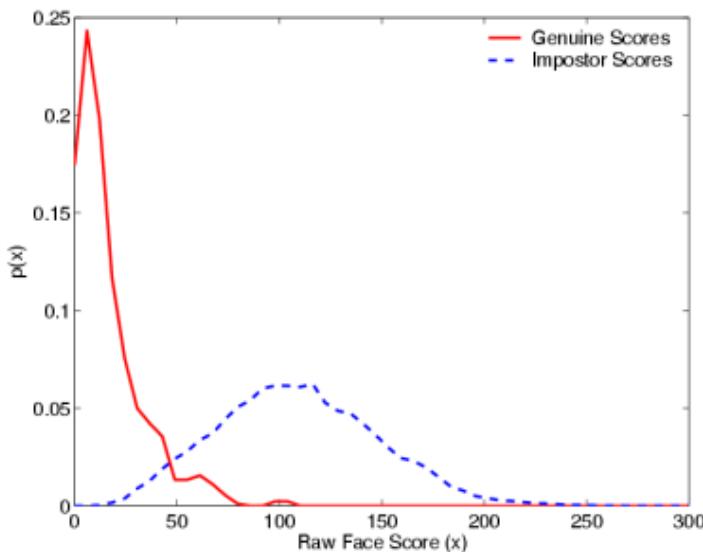


Database of 100 users with three modalities (5 samples/user)

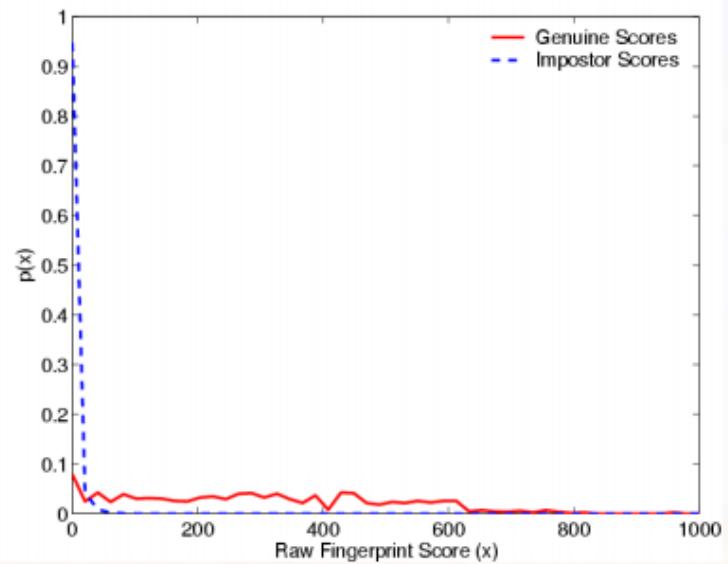
Experimental Setup



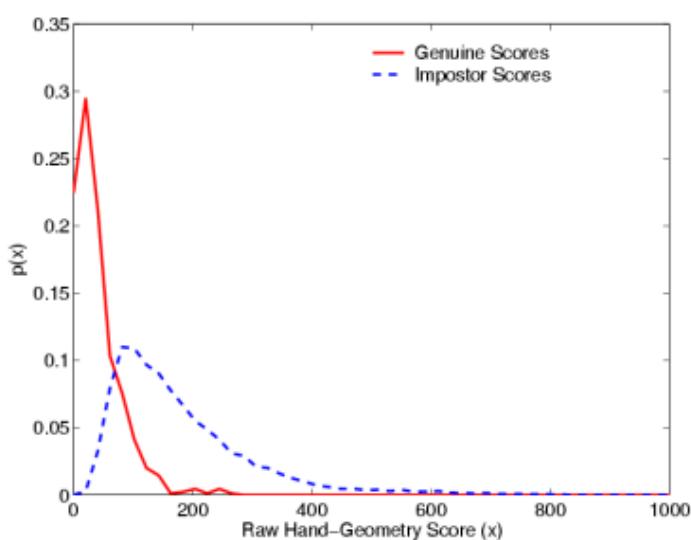
Distribution of Matching Scores



Face

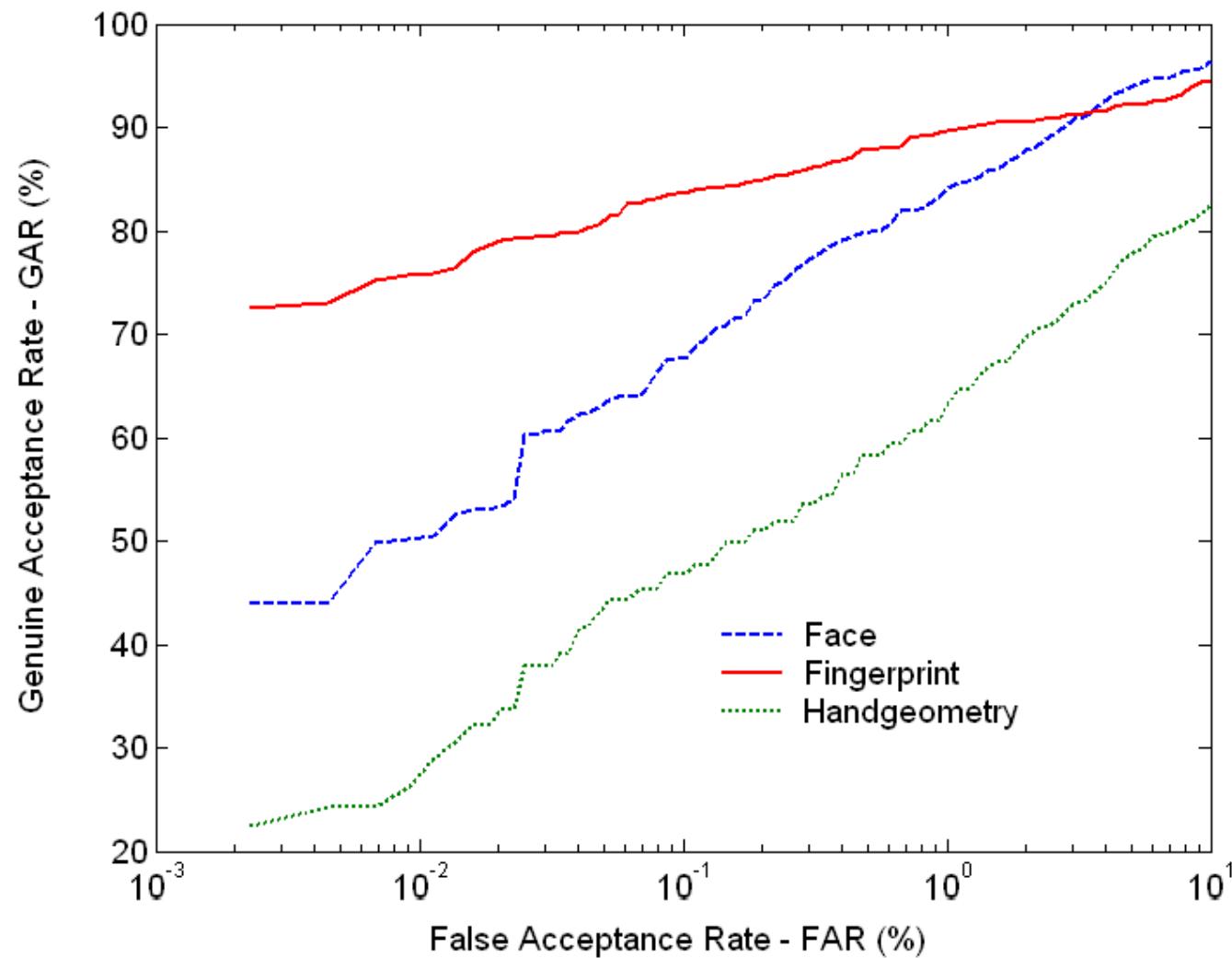


Fingerprint

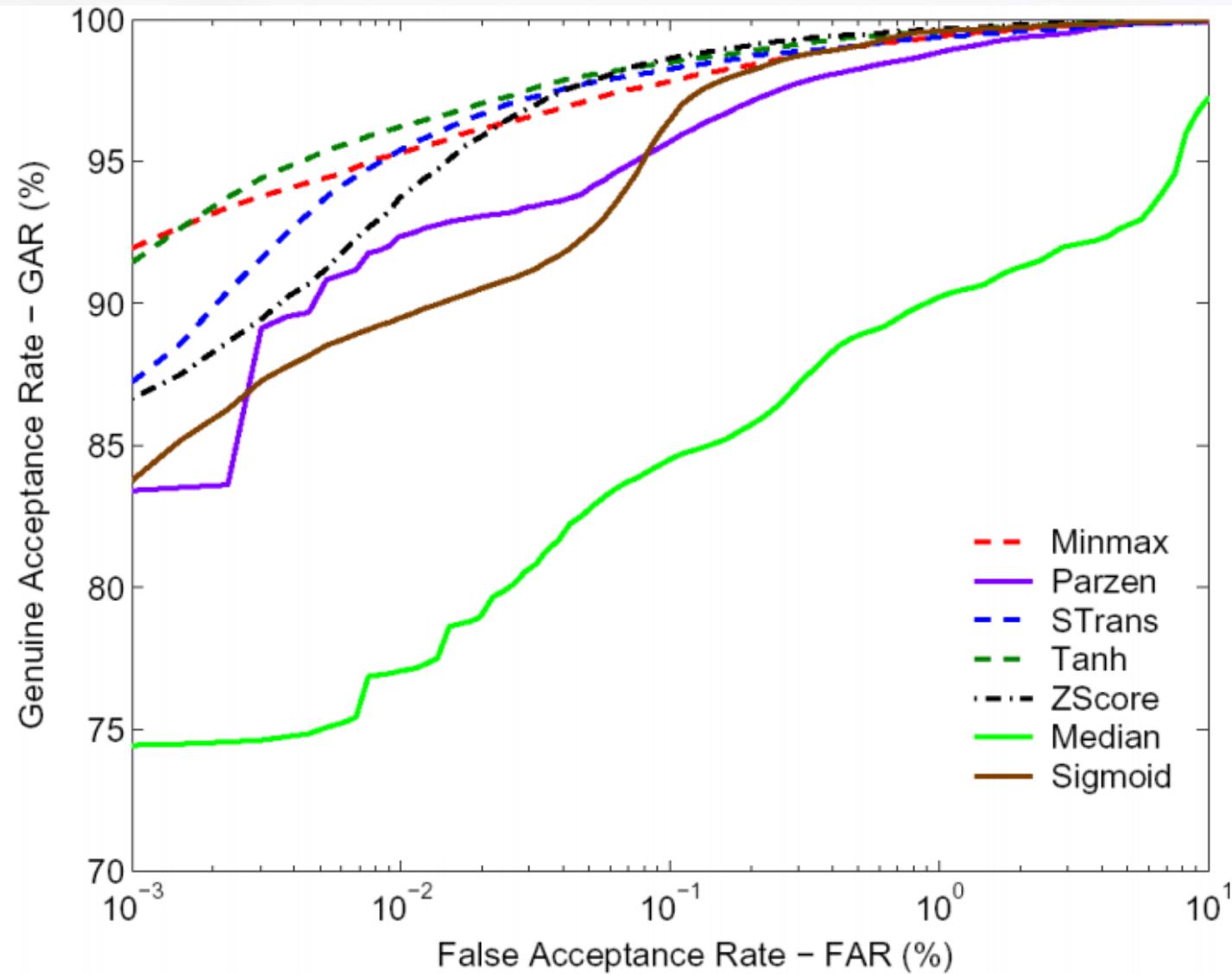


Hand-geometry

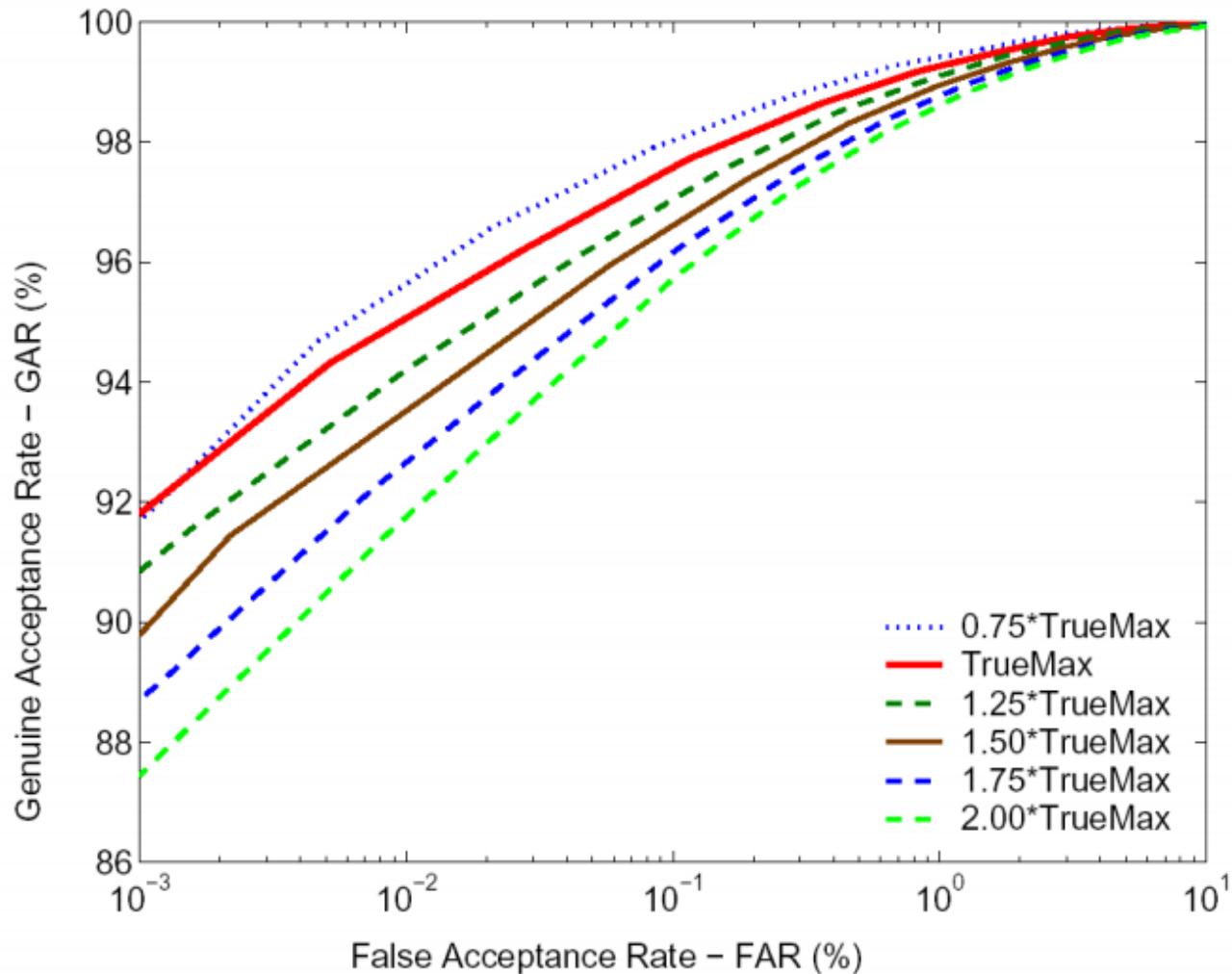
Performance of Individual Modalities



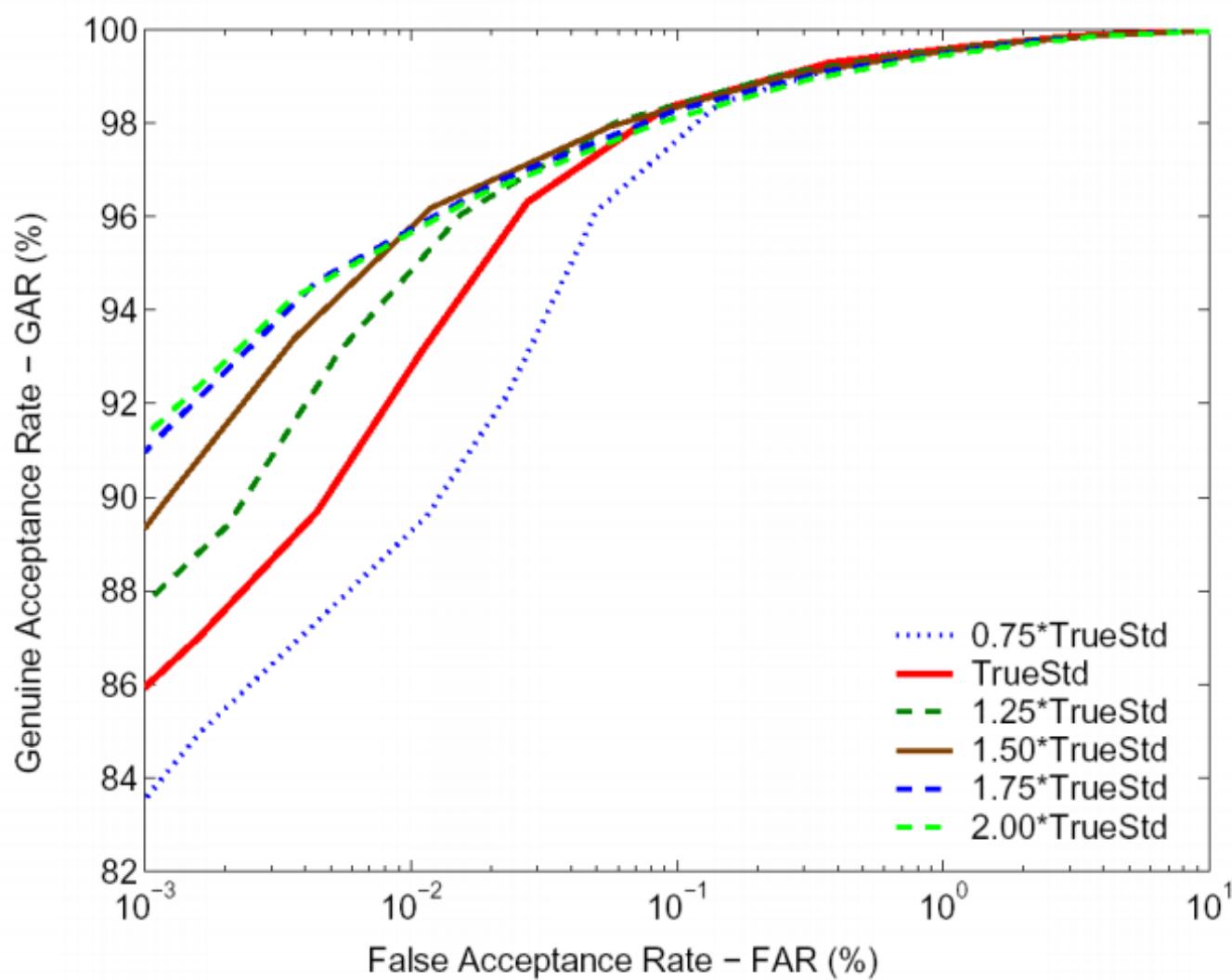
Normalization followed by Sum Rule



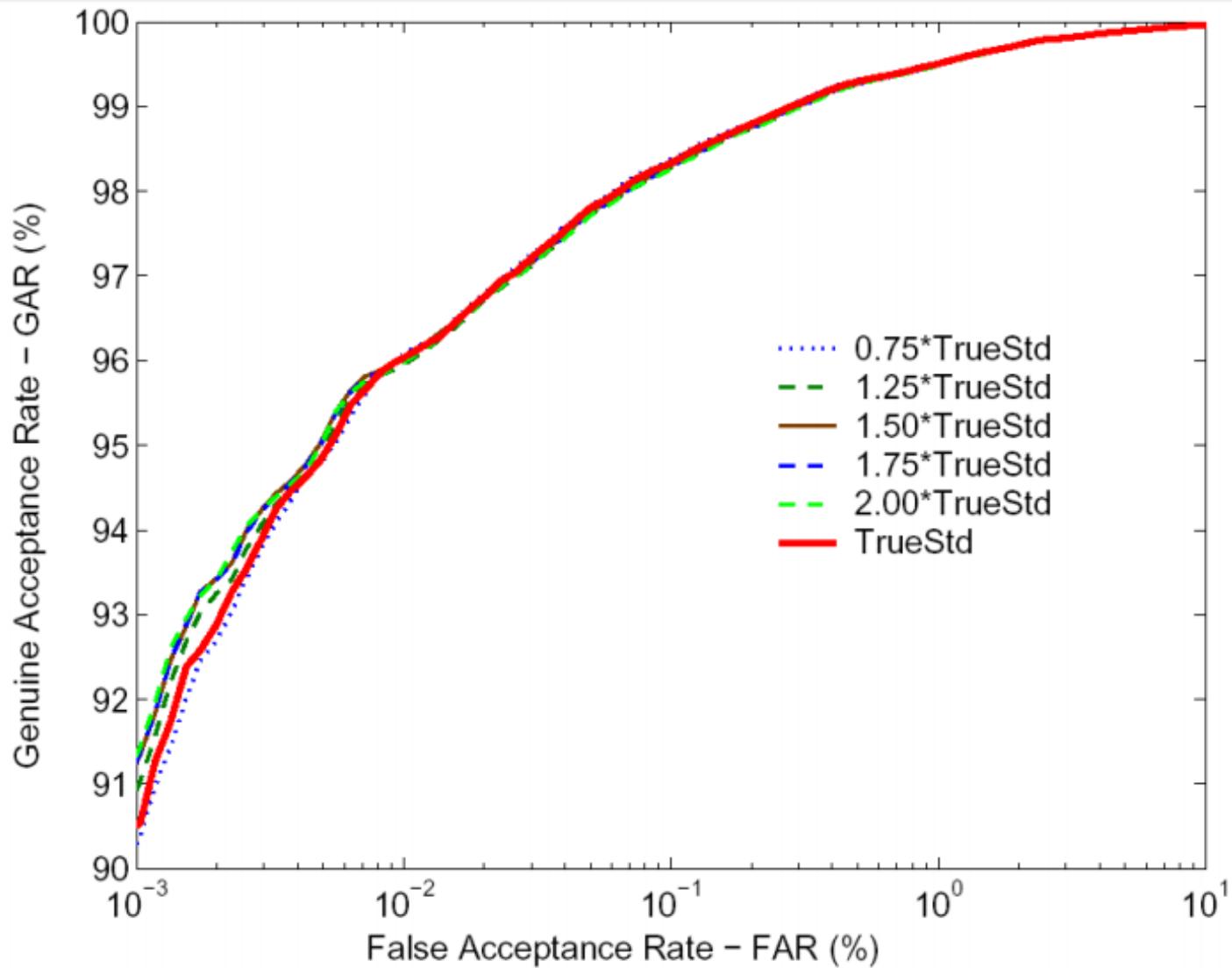
Sensitivity of Min-max to Outliers



Sensitivity of Z-score to Outliers



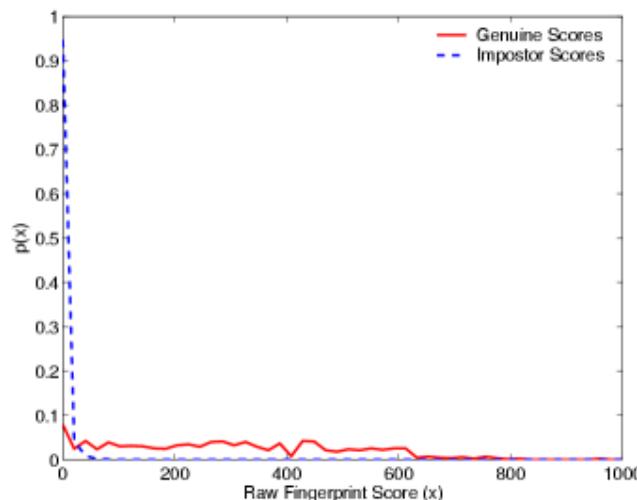
Sensitivity of Tanh to Outliers



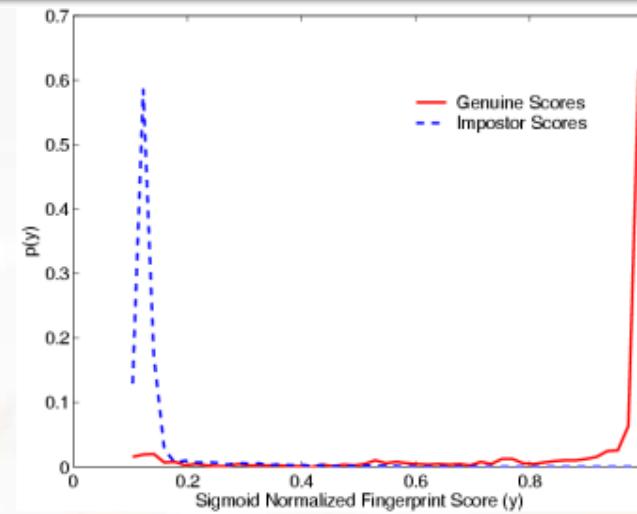
Summary

- The effects of different score normalization techniques in a multimodal biometric system have been studied
- Min-max, Z-score, and Tanh normalization techniques followed by sum rule resulted in the best recognition performance
- Min-max and Z-score methods are efficient but sensitive to outliers in the training data; Tanh is both efficient and robust
- If the parameters of the matching score distribution are known, simple methods like Min-max and Z-score would suffice
- If the parameters are to be estimated from noisy training data, one should choose robust methods like Tanh normalization
- If the system can provide both score and some confidence measure on that score, non-parametric density estimation methods can be used to estimate the posteriori probability

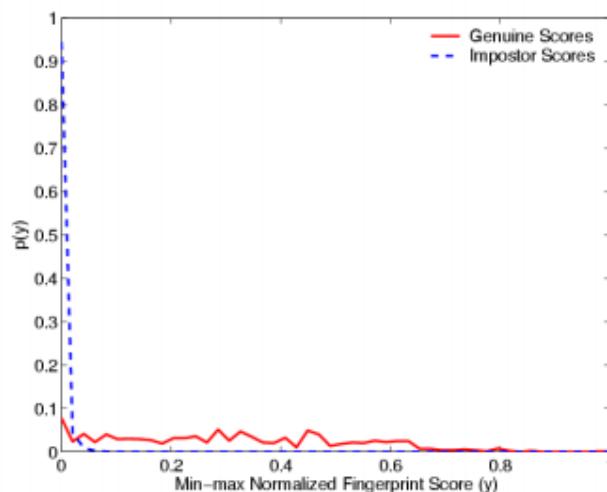
Efficiency of a Normalization Scheme



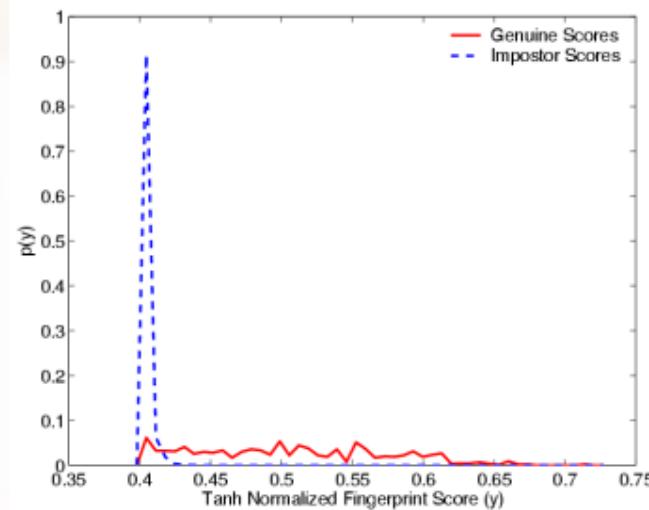
Before Normalization



After Sigmoid Normalization



After Min-max Normalization



After Tanh Normalization

One of our Proposed Multimodal Biometric System (Whole hands) FYI



FP
HG
HV

**NIR Hand
Veins**
(For Rt/Lt Hand)

Fingerprint
Index & Ring
(For Rt/LT Hand)

**NIR Hand
Geometry**
(For Rt/Lt Hand)

And/Or

And/Or

And/Or

And/Or



One of my supervised
PhD students , M.K.Shahin 2010

Fusion & Decision

Measures of performance:

- Sensitivity or GAR = $GA / (GA+FR) \times 100$
- Specificity or GRR = $GR / (GR+FA) \times 100$
- FAR= $FA / (GR+FA) \times 100 = 100 - GRR$
- FRR= $FR / (GA+FR) \times 100 = 100 - GAR$
- Efficiency = $(GA+GR) / (GA+GR+FA+FR)$
- Receiver Operating Characteristic (ROC)
- Equal error rate (EER)
- Total error rate (TER=min (FAR+FRR))
- Failure to enroll (FTE)
- Failure to capture (FTC) or failure to acquire
- Receiver operating characteristic (ROC) showing trade off between (GAR vs FAR)

Evaluated Biometric Systems (1st-to-6th)

3 Rt-Index-FP



4 Rt-Ring-FP



5 Lt-Ring-FP



6 Lt-Index-FP



1 Rt-HG



2 Lt-HG



Feature Fusion

Evaluated Biometric Systems (7th-to-10th)

Rt-Index-FP Rt-Ring-FP



8
+



10
+

Lt-Ring-FP Lt-Index-FP



9
+



Rt-HG



7
+

Lt-HG



Score Fusion

Evaluated Biometric Systems (11th-to-13th)

Rt-Index-FP Rt-Ring-FP

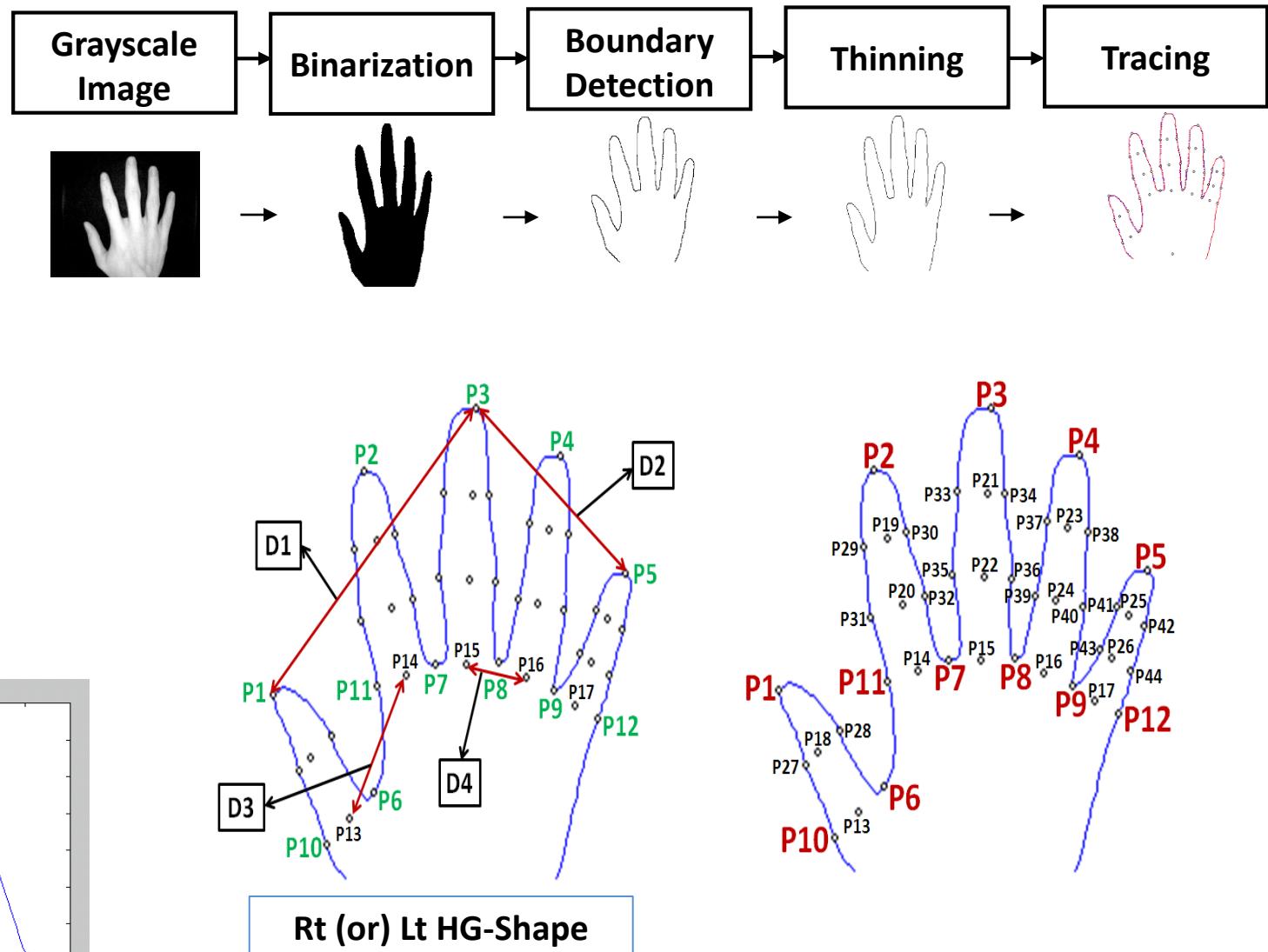


13
+

Lt-Ring-FP Lt-Index-FP



HG Feature Extraction



HG Feature Vector & Pattern Matching

1. HG Feature Vector is 1D array of 29 elements, all of lengths, widths, circumferences, and areas of different fingers
2. Square roots of Finger areas were taken to keep the homogeneity of all HG features in 1D Form

1 29
35.5 1	26.9 2
72.4 0	32.6 9
27.3 1
..... ...	265. 2
..... ...	183.7 3

Matching of HG Feature Vectors via
Euclidian Distance Measure



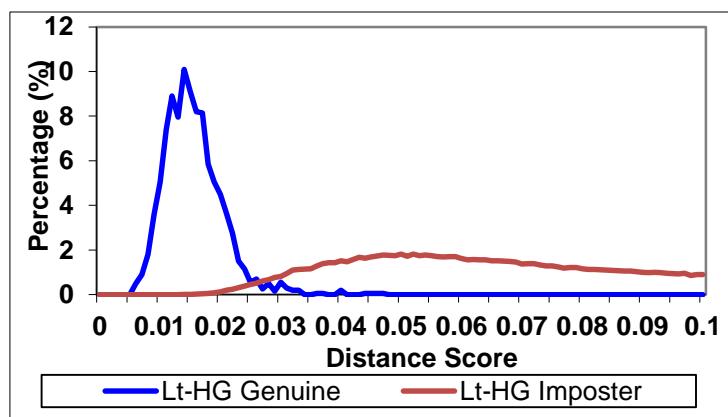
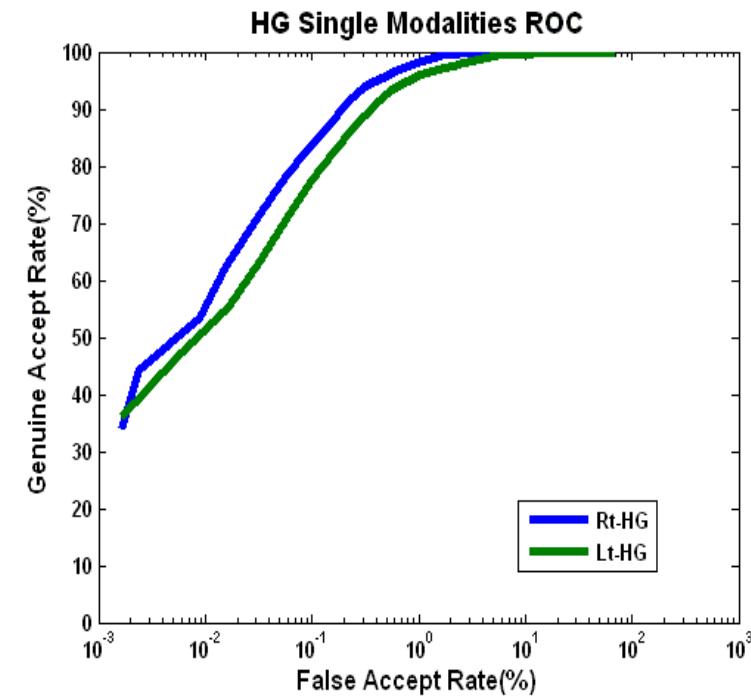
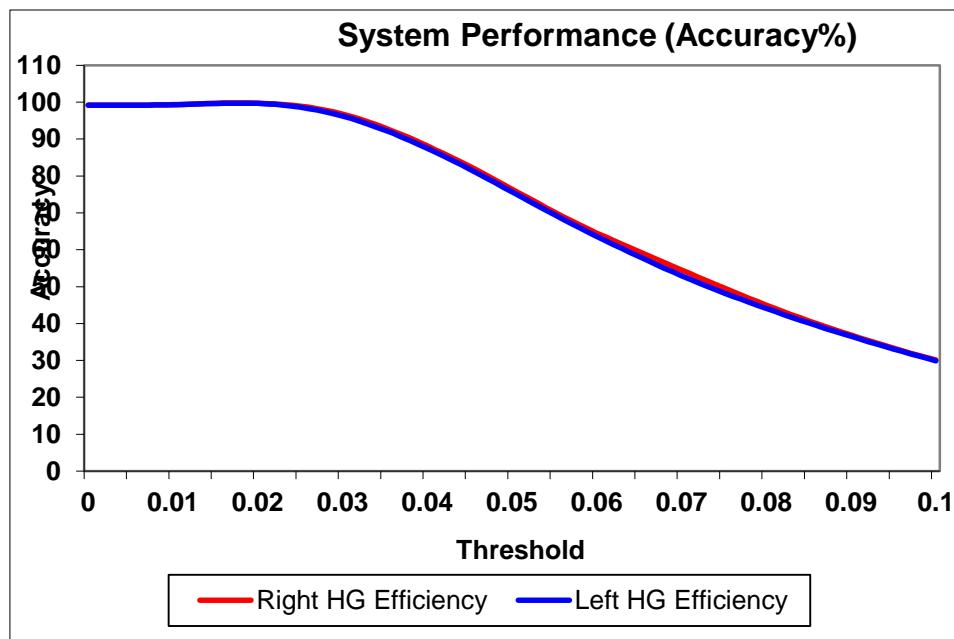
$$\sqrt{\sum_{i=1}^d (y_i - f_i)^2} < T$$

Similarity Between Rt & Lt HG for the Same Person

Results for the probability for the true match between Rt-HG & Lt-HG patterns for the same identity with Mean = 0.27648 and Std. Dev. = 0.008805243.

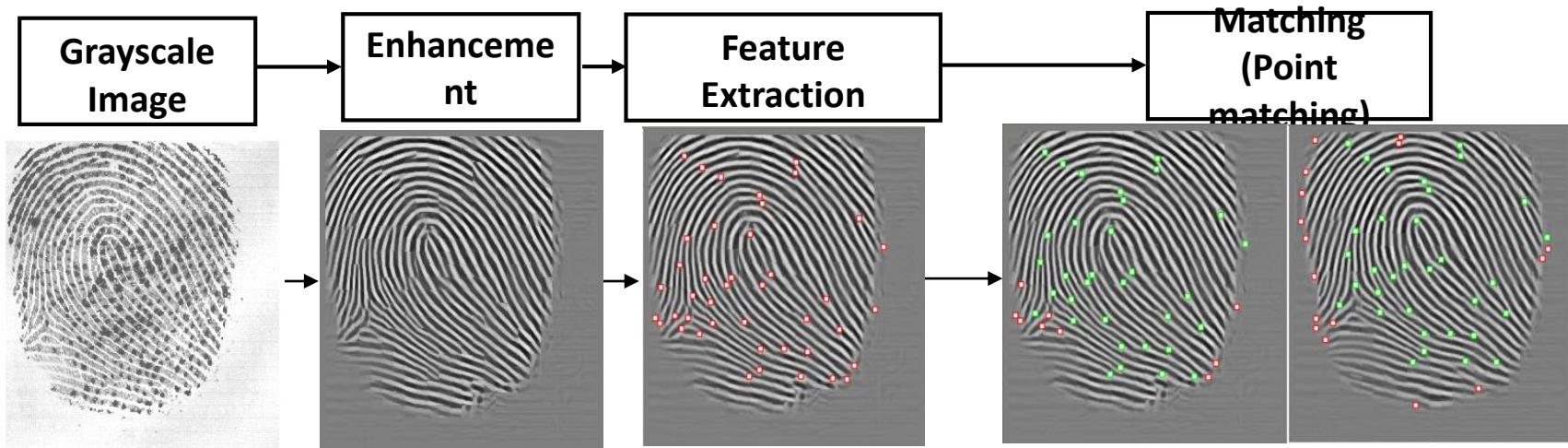
Threshold	Z-score	P(Z)	0.5 - P(Z)	P(Distance <= Threshold) %
0.01	-2.00428	0.4772	0.0228	2.28%
0.011	-1.89071	0.4706	0.0294	2.94%
0.012	-1.77714	0.4625	0.0375	3.75%
0.013	-1.66358	0.4525	0.0475	4.75%
0.014	-1.55001	0.4394	0.0606	6.06%
0.015	-1.43644	0.4251	0.0749	7.49%
0.016	-1.32287	0.4066	0.0934	9.34%
0.017	-1.209	0.3869	0.1131	11.31%
0.018	-1.09573	0.3643	0.1357	13.57%
0.019	-0.98216	0.3365	0.1635	16.35%
0.02	-0.86859	0.3078	0.1922	19.22%

HG Single-Modality Biometrics Results

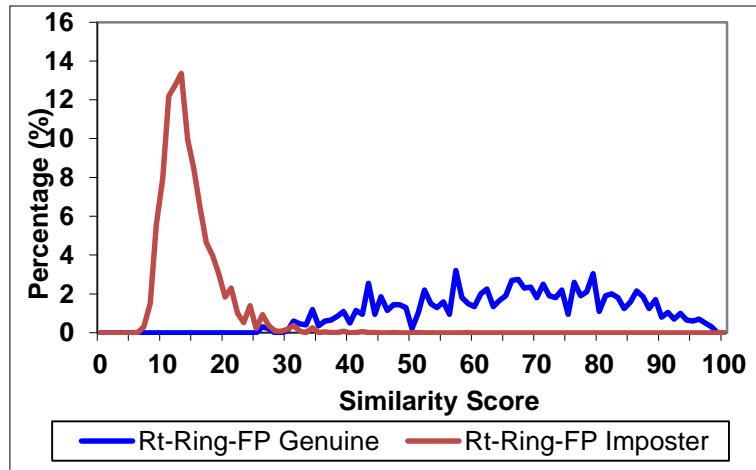
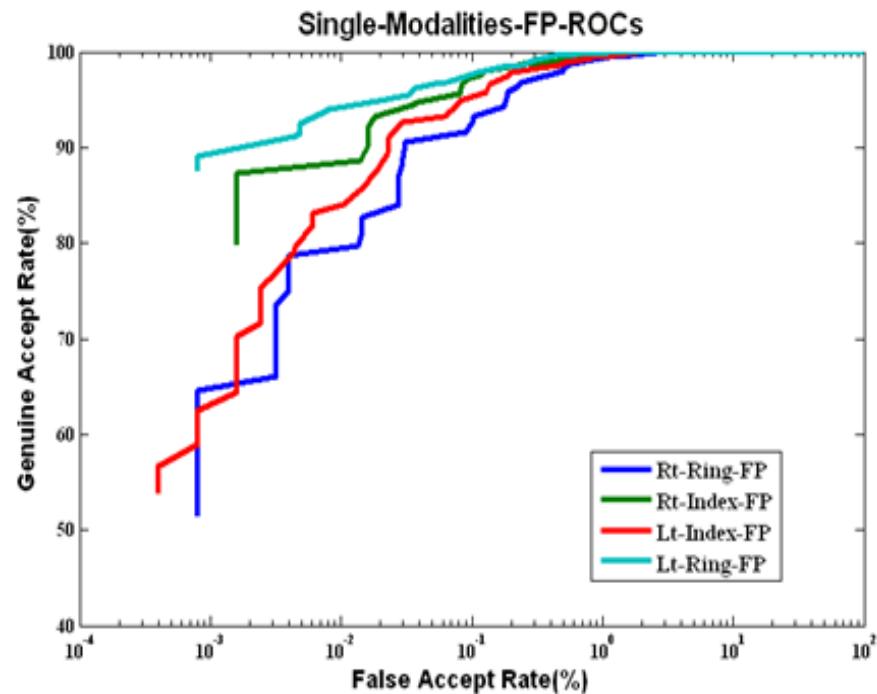
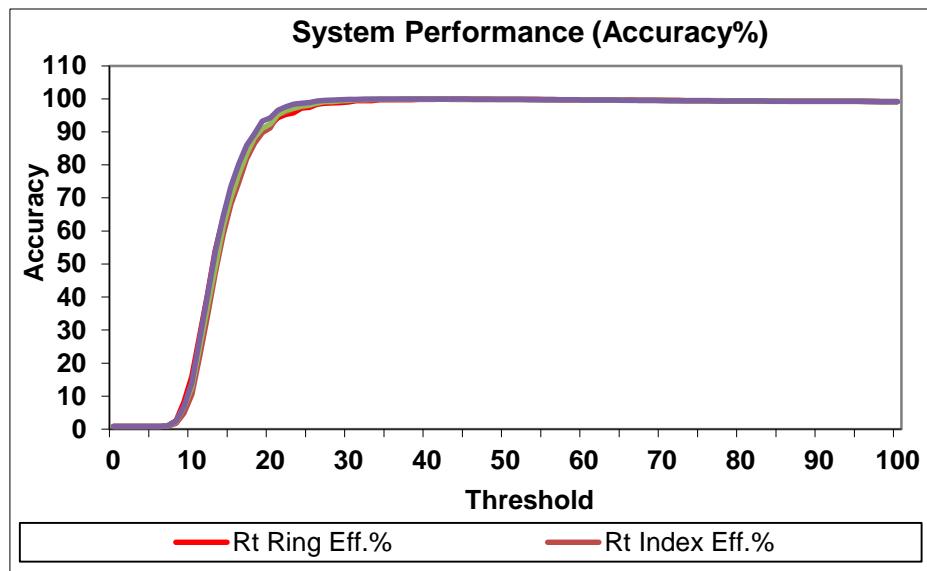


Modality	Accuracy %	TER%	EER%
Rt-HG	99.772345	2.169293	1.192525
Lt-HG	99.719439	4.384950	2.285051

Fingerprint Analysis Diagram



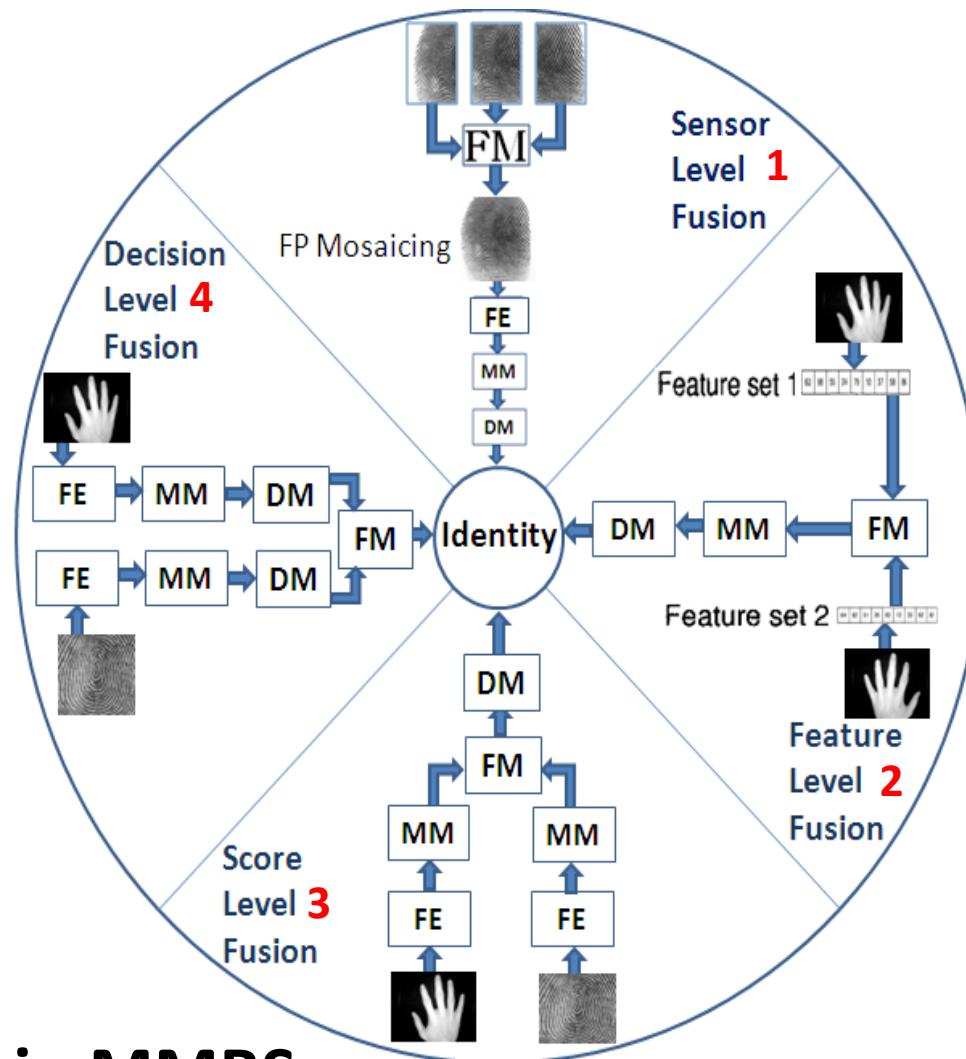
FP Single-Modality Biometrics Results



Modality	Accuracy%	TER%	EER%
Rt Ring FP	99.894188	1.557172	0.778586
Rt Index FP	99.928257	0.928889	0.658636
Lt Index FP	99.911824	1.417071	0.752020
Lt Ring FP	99.943888	0.788384	0.414141

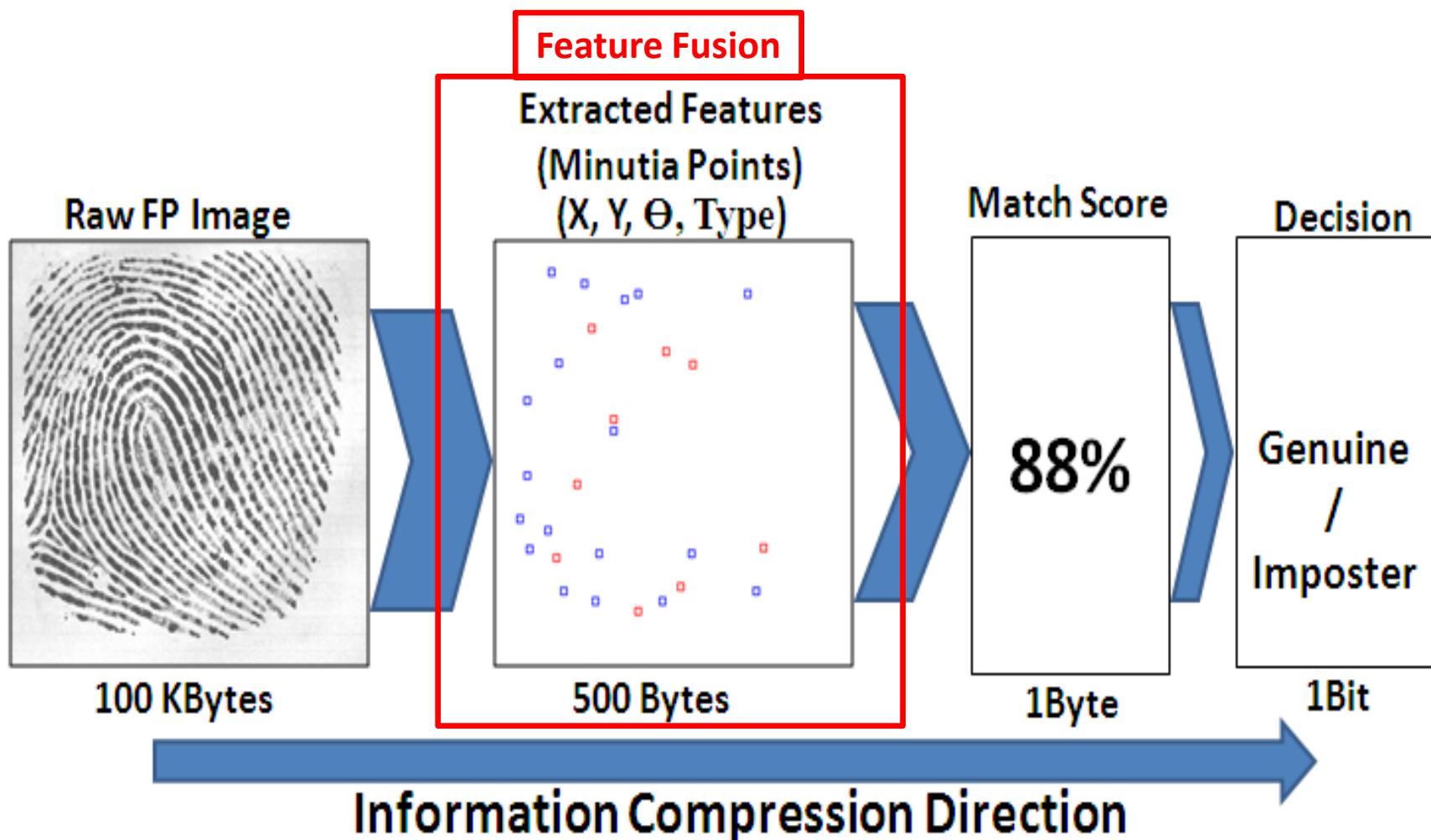
MMBS Fusion Scenarios

FE: Feature Extraction
MM: Matching Module
DM: Decision Module
FM: Fusion Module



Scenarios of Fusion in MMBS

Data is Getting Compressed from Sensor to Decision



Why Feature Fusion !

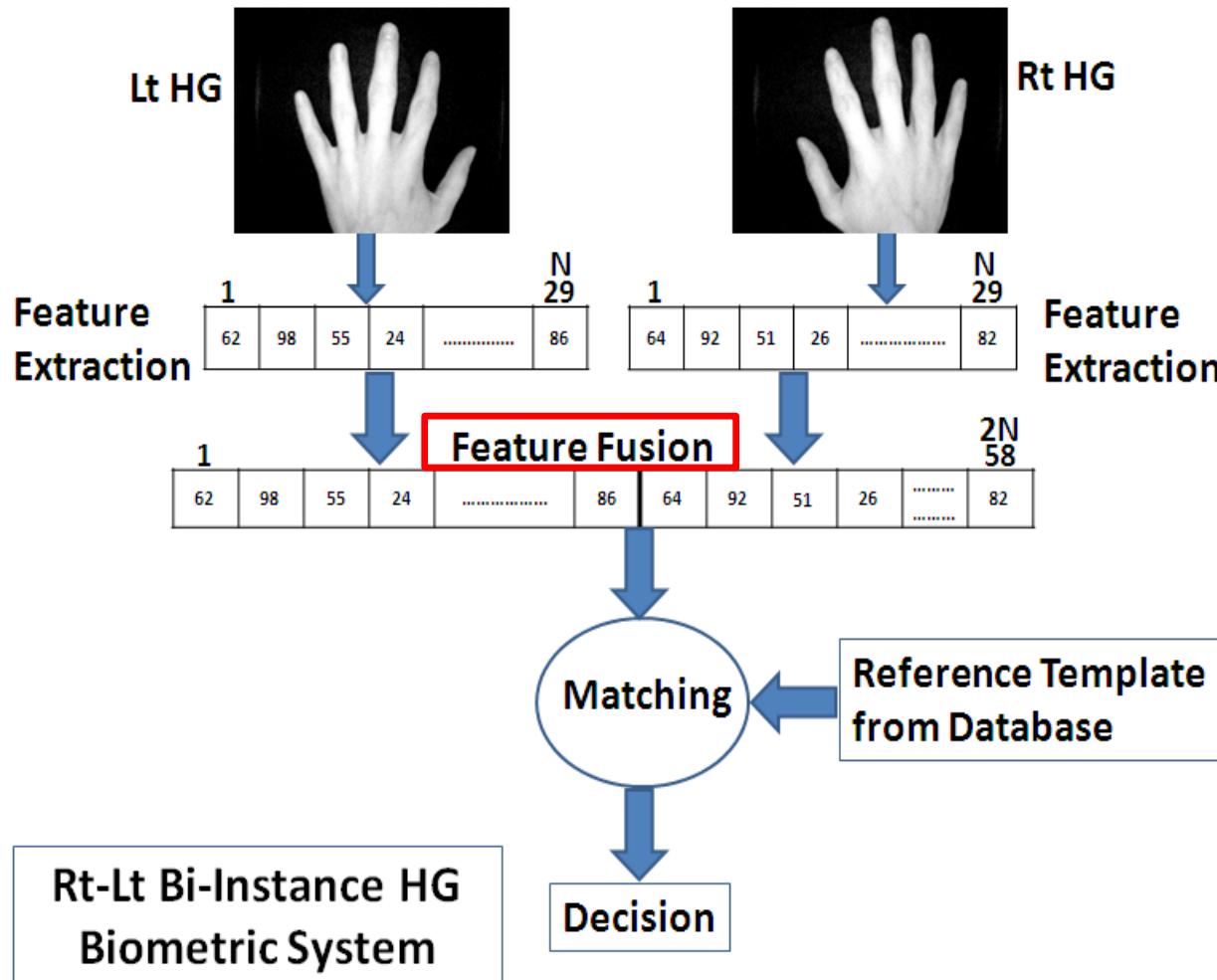
Biometric systems that integrate information at an early stage of processing are believed to be more effective than those which perform integration at a later stage, and this was our choice.

V. M. Mane and D. V. Jadhav, "Review of multimodal biometrics: Applications, challenges and research areas," *Int. Journal of Biometrics and Bioinformatics (IJBB)*, vol. 3, 2009.

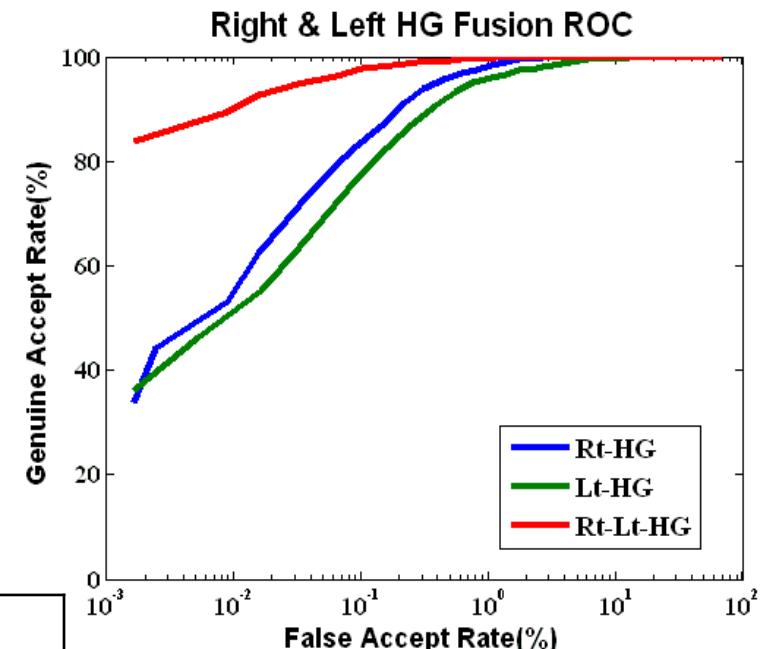
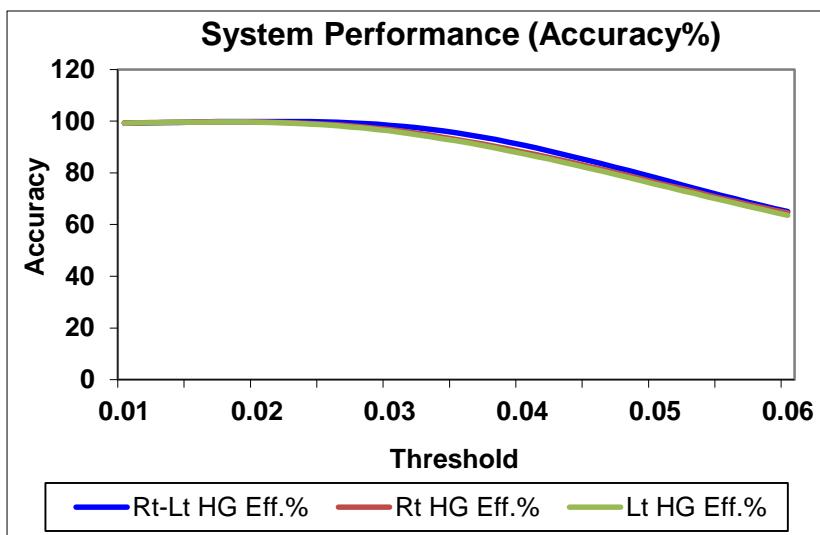
From Single-Modalities —To—> Multi-Instances Using Feature Fusion Methodologies

- Rt-Lt Bi-Instance HG Biometric System: (From Fusing Rt-HG and Lt-HG using Feature-Level Fusion Technique)
- Right Bi-Instance FP Biometric System:
(From Fusing Rt-Ring-FP and Rt-Index-FP using Feature-Level Fusion Technique)
- Left Bi-Instance FP Biometric System:
(From Fusing Lt-Index-FP and Lt-Ring-FP using Feature-Level Fusion Technique)
- Rt-Lt Bi-Instance FP Biometric System:
(From Fusing Rt-Ring-FP, Rt-Index-FP, Lt-Index-FP and Lt-Ring-FP using Feature-Level Fusion Technique).

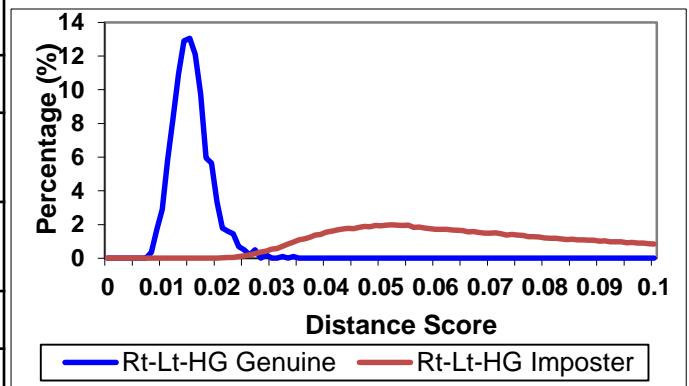
Rt-Lt Bi-Instance HG Biometric System



Rt-Lt Bi-Instance HG System: Results



Modality	Accuracy%	TER%	EER%
HG Single-Modalities			
Rt-HG	99.772345	2.169293	1.192525
Lt-HG	99.719439	4.384950	2.285051
After HG Feature Fusion			
Rt-Lt HG	99.926653	1.015758	0.507879



Fusion for similar traits always increases accuracy (lower EER)

FP Features (Minutiae) Fusion

FP Similarity (Single Modality)

Fusion of Right Ring & Right Index Fingerprint Feature Points

X	Y	Θ	Type	FingerCode
13	120	146	0	1
18	115	152	0	1
83	253	78	1	1

Right Ring
Minutia
Features

X	Y	Θ	Type	FingerCode
29	138	146	0	2
58	272	67	1	2
112	195	56	1	2

Right Index
Minutia
Features

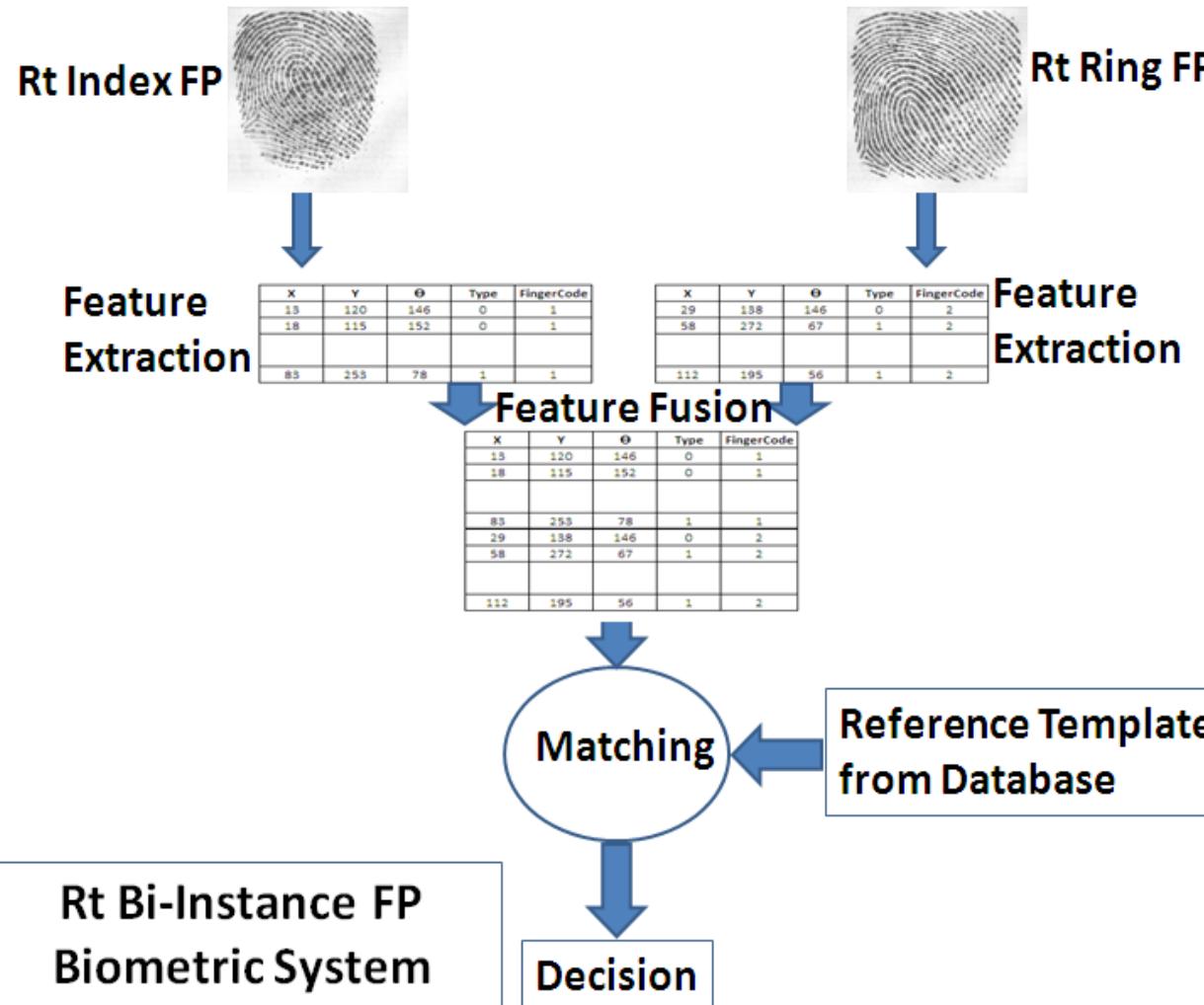
Concatenated (Fused) Minutia Features				
X	Y	Θ	Type	FingerCode
13	120	146	0	1
18	115	152	0	1
83	253	78	1	1
29	138	146	0	2
58	272	67	1	2
112	195	56	1	2

$$S(R, T) = \frac{m}{\min(M_R, M_T)}$$

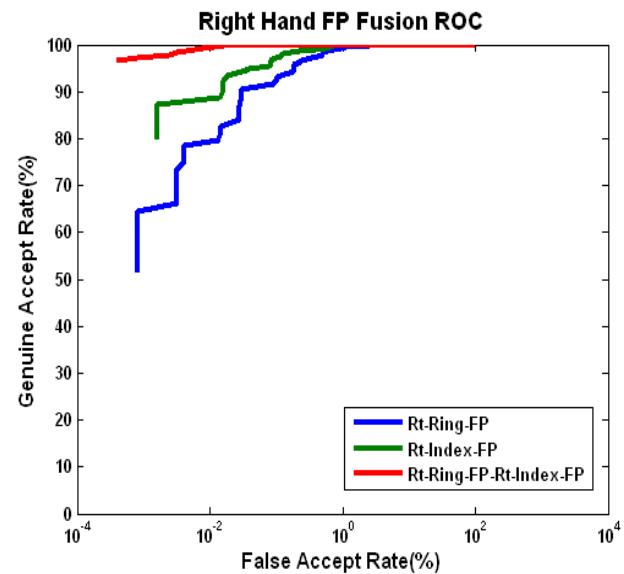
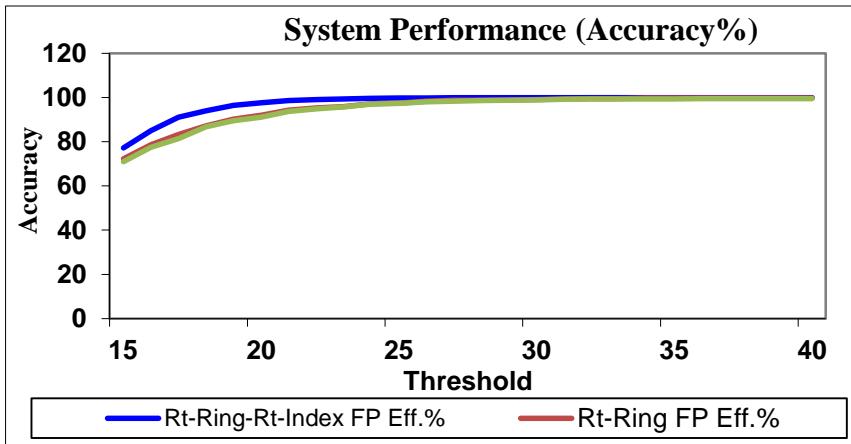
FP Similarity (Multiple Traits)

$$S(R_1, R_2, \dots, R_n, T_1, T_2, \dots, T_n) = \frac{\sum_{i=1}^n m_i}{\sum_{i=1}^n \min(M_{R_i}, M_{T_i})}$$

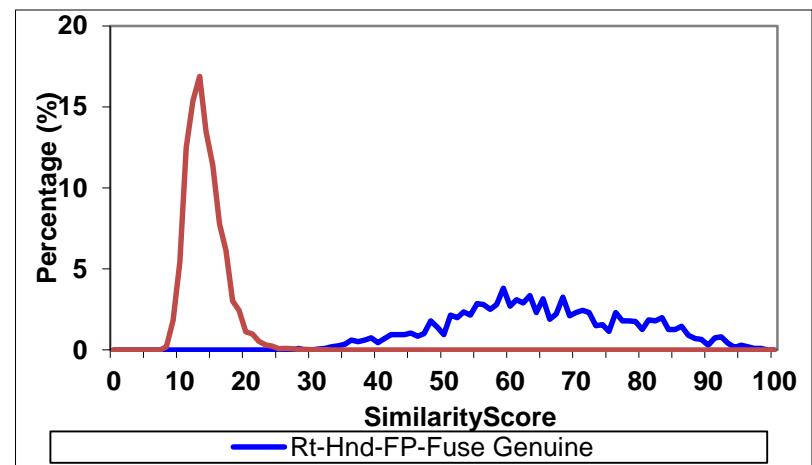
Rt Bi-Instance FP Biometric System



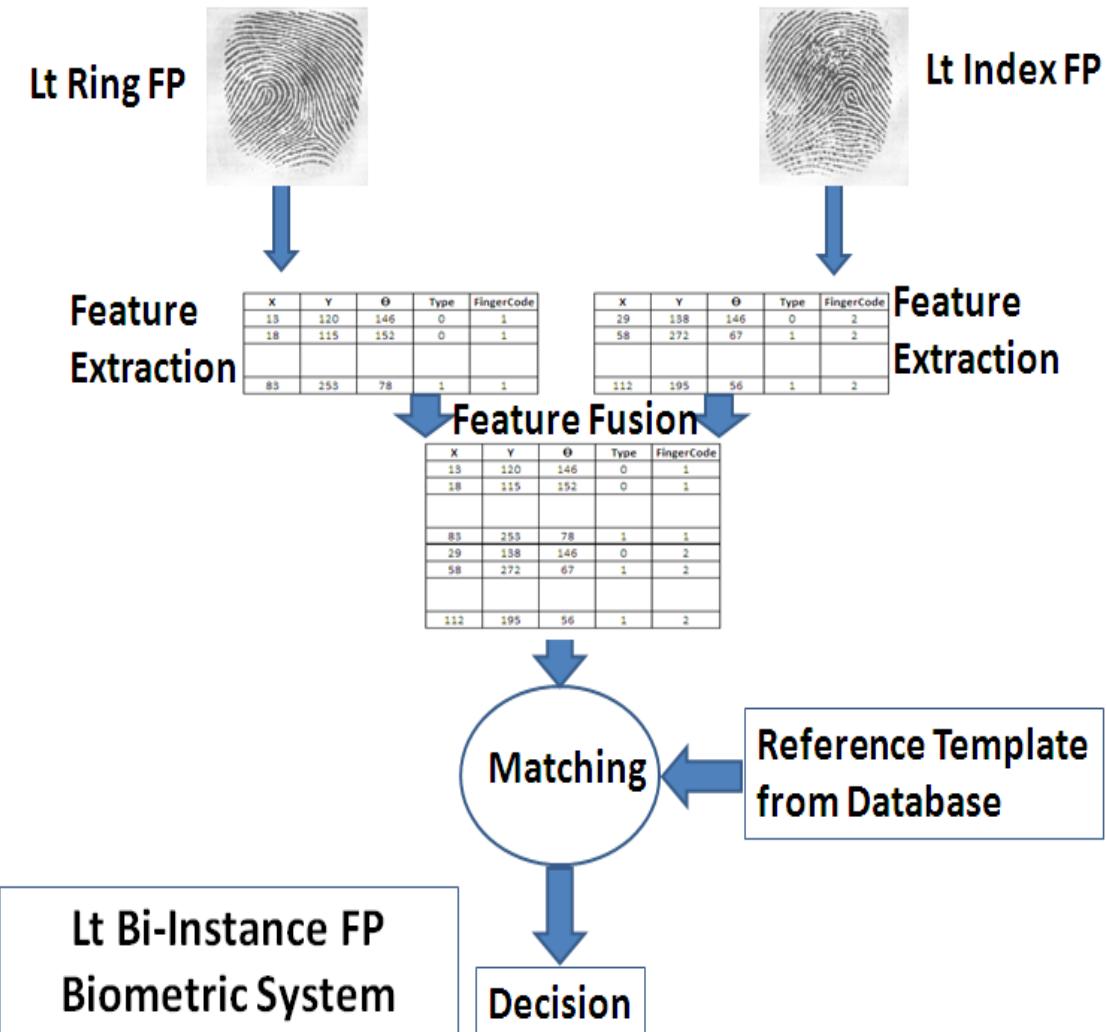
Rt Bi-Instance FP System: Results



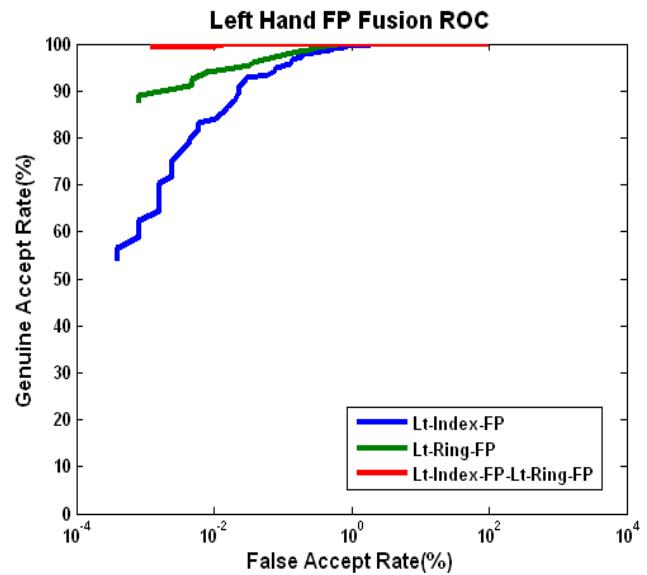
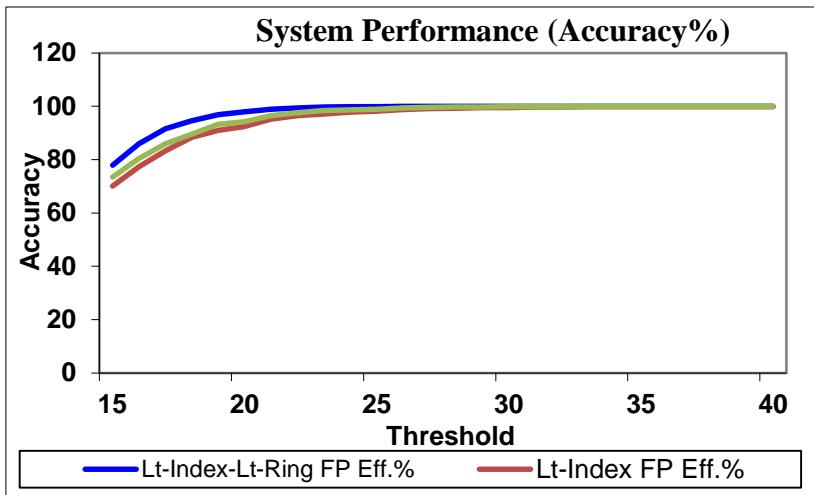
Modality	Accuracy %	TER%	EER%
FP Single-Modalities			
Rt-Ring FP	99.894188	1.55717 2	0.77858 6
Rt-Index FP	99.928257	0.92888 9	0.65863 6
After FP Biometric Feature Fusion			
Rt-Ring-Rt-Index FP	99.986774	0.11232 3	0.08979 8



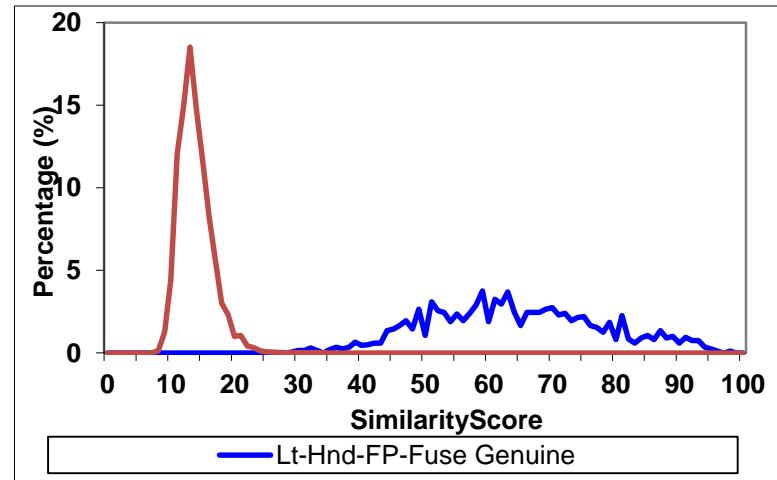
Lt Bi-Instance FP Biometric System



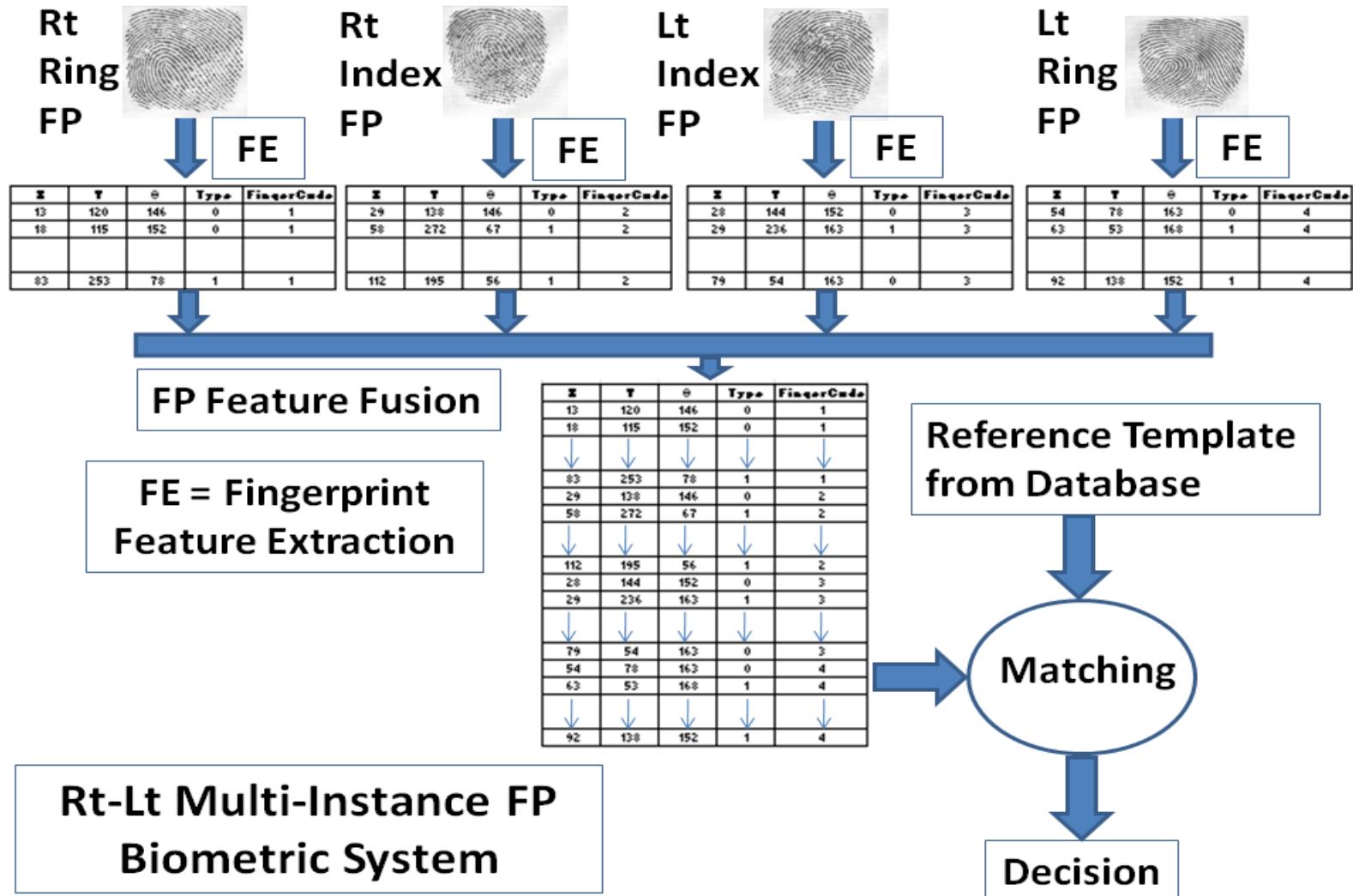
Lt Bi-Instance FP System: Results



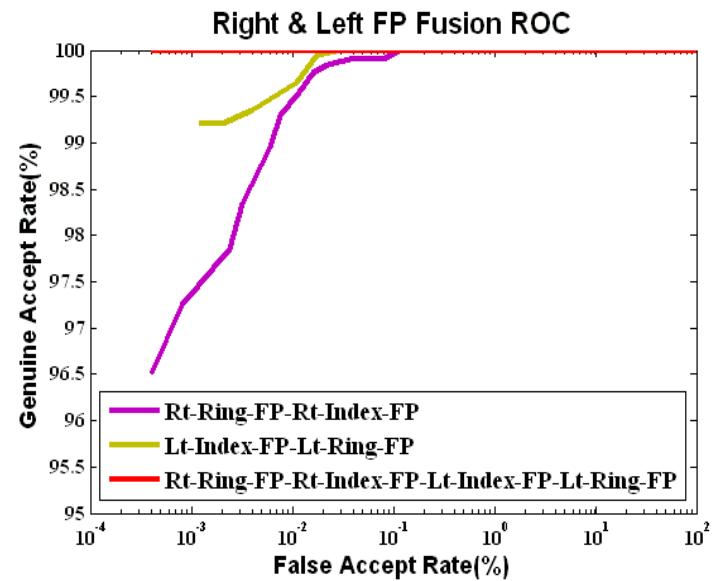
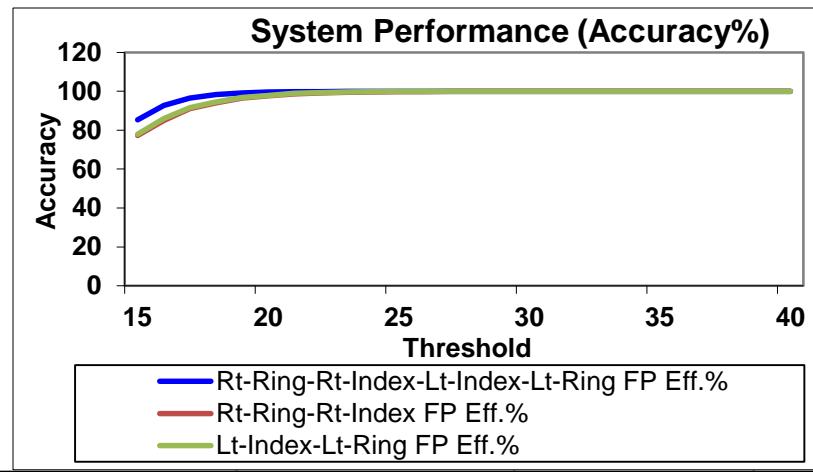
Modality	Accuracy %	TER%	EER%
FP Single-Modalities			
Lt-Index FP	99.911824	1.417071	0.752020
Lt-Ring FP	99.943888	0.788384	0.414141
After FP Biometric Feature Fusion			
Lt-Index-Lt-Ring FP	99.992385	0.026667	0.013333



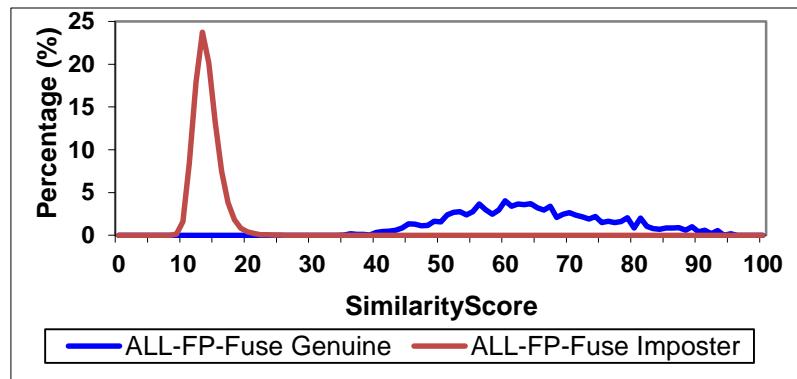
Rt-Lt Multi-Instance FP Biometric System



Rt-Lt Multi-Instance FP System: Results



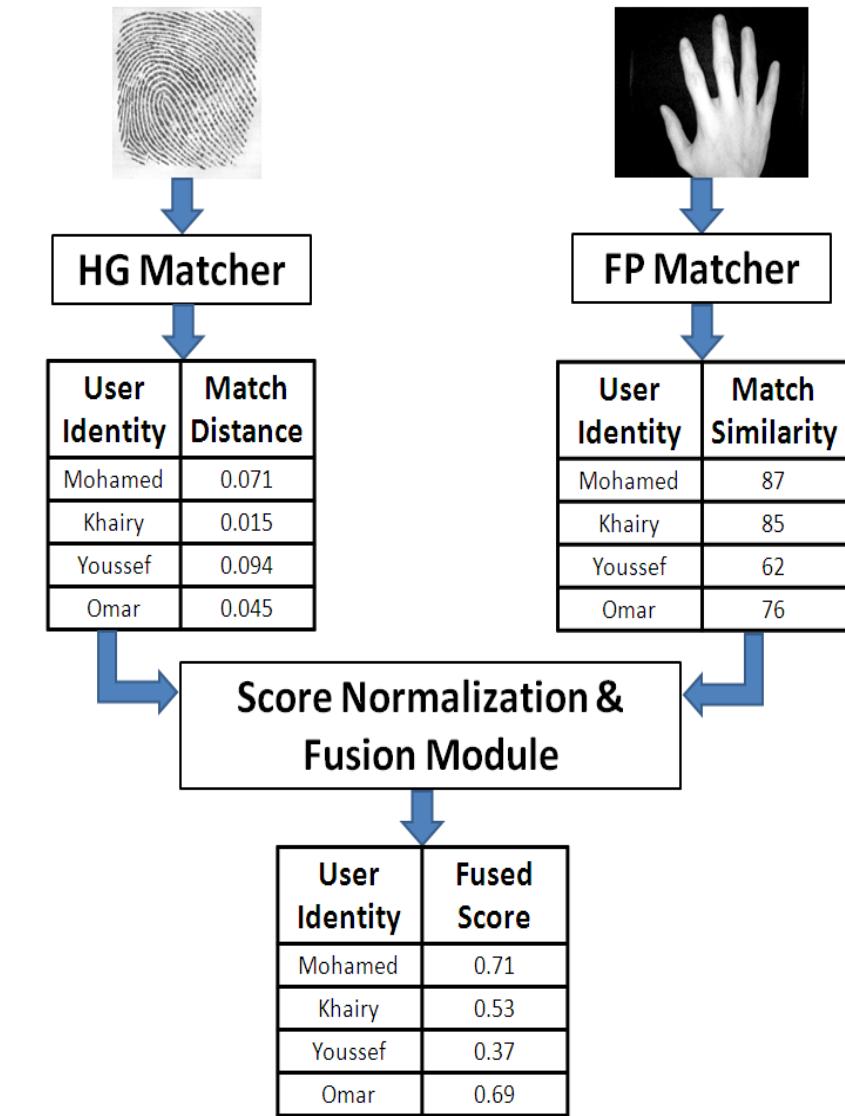
Modality	Accuracy %	TER%	EER%
Rt-Ring- Rt-Index FP	99.987	0.112323	0.089798
Lt-Index- Lt-Ring FP	99.992	0.026667	0.013333
Rt-Ring- Rt-Index- Lt-Index- Lt-Ring FP	99.999	0.002020 2	0.001010 1



Fusion of Different Biometric Traits

Challenges

- The Features are NOT Homogeneous
- Different Matchers Output are NOT Homogeneous
- HG Distance: small when matching is high
- FP Similarity: large when matching is high



Different Matchers O/P Score Normalization

- Score normalization refers to: changing the location and scale parameters of the match score distributions at the outputs of the individual matchers
- Many Techniques can be used for Score Normalization like: Min-Max and Z-Score

Min-Max Scores Normalization

- Min-Max: transform the minimum and maximum scores to 0 and 1, respectively
- Distance scores can be transformed into similarity scores by subtracting the normalized score from 1

$$n_{-}s_j^t = \frac{s_j^t - \min_{i=1}^N s_j^i}{\max_{i=1}^N s_j^i - \min_{i=1}^N s_j^i}$$

Fusing the Normalized Scores

- Several methods can be implemented for merging and fusing the normalized scores like max-score, min-score, sum-of-scores

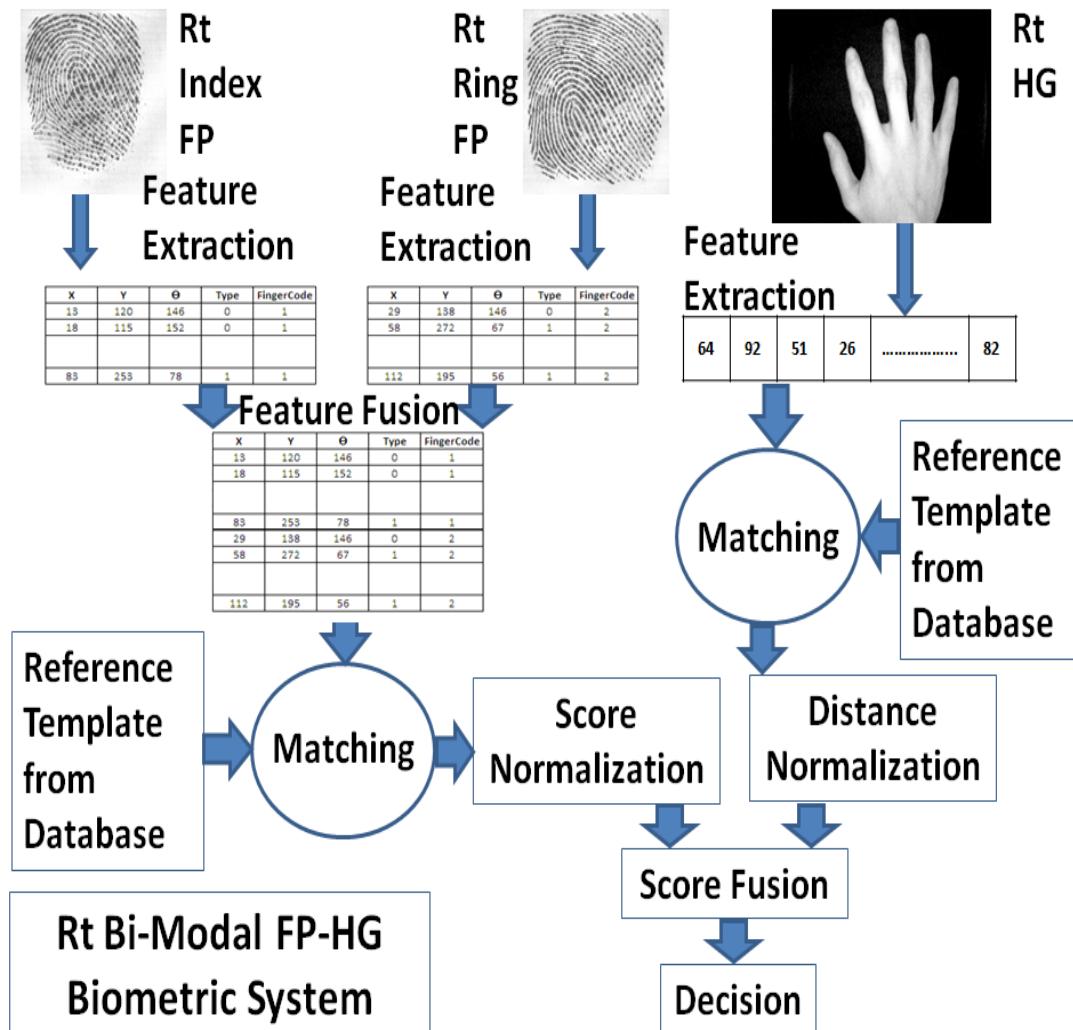
In this work, we used the sum-of-scores Score Fusion Technique. for N normalized scores: n_{S_1} , n_{S_2} , ..., n_{S_N} is given by:

$$n_s = \frac{\sum_{i=1}^N n_{S_i}}{N}$$

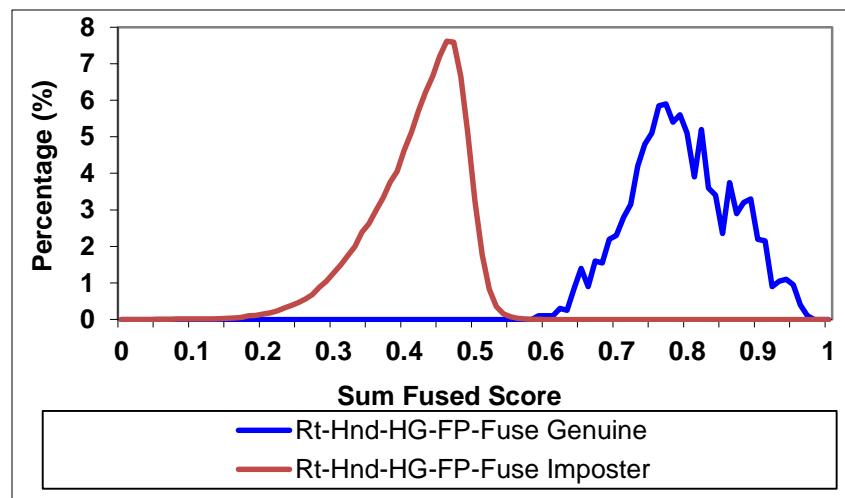
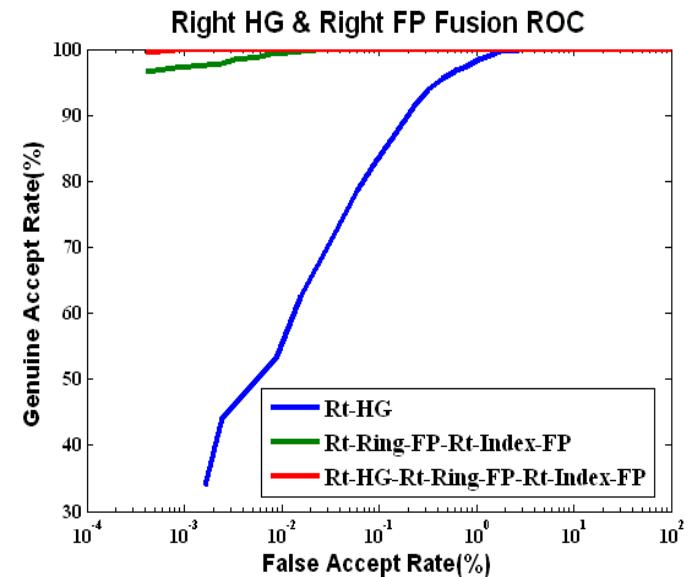
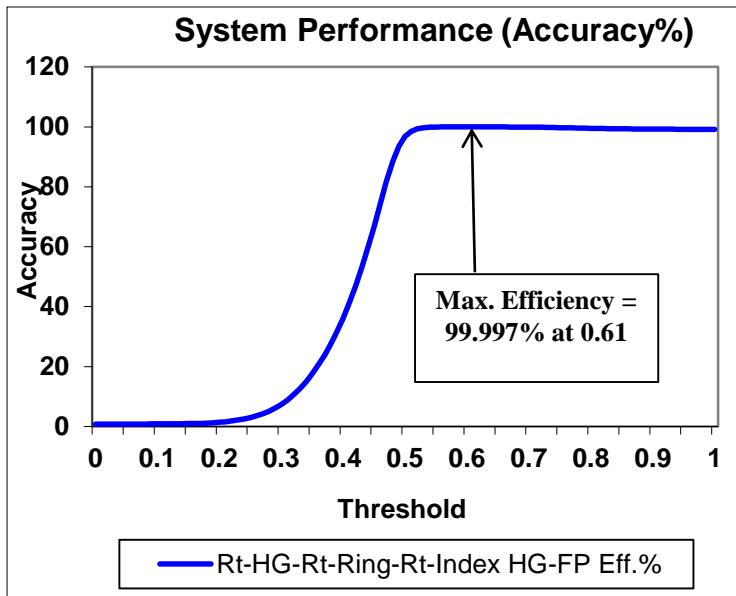
3 MMBSS Using Score Fusion Techniques

- Rt-HG-FP Biometric System:
 - (From Fusing Rt-HG, Rt-Ring-FP and Rt-Index-FP using Min-Max Score Normalization and sum-of-scores Score-Level Fusion Technique)
- Lt-HG-FP Biometric System:
 - (From Fusing Lt-HG, Lt-Index-FP and Lt-Ring-FP using Min-Max Score Normalization and sum-of-scores Score-Level Fusion Technique)
- Rt-Lt-HG-FP Biometric System:
 - (From Fusing Rt-HG, Lt-HG, Rt-Ring-FP, Rt-Index-FP, Lt-Index-FP, and Lt-Ring-FP using Min-Max Score Normalization and sum-of-scores Score-Level Fusion Technique)

Rt-HG-FP Biometric System



Rt-HG-FP Biometric System: Results



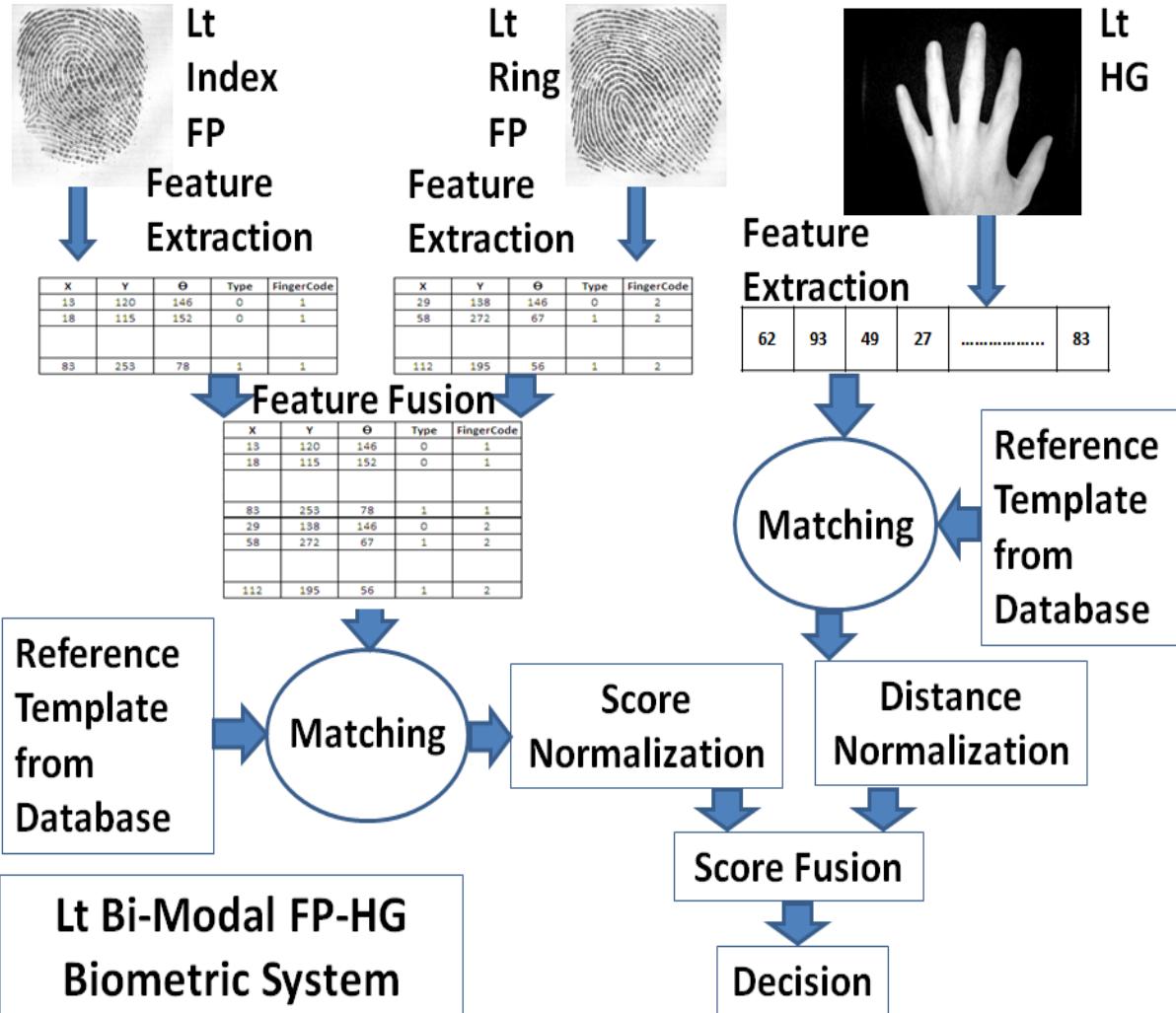
Qualitative Analysis

Rt-HG-FP Biometric System: Results

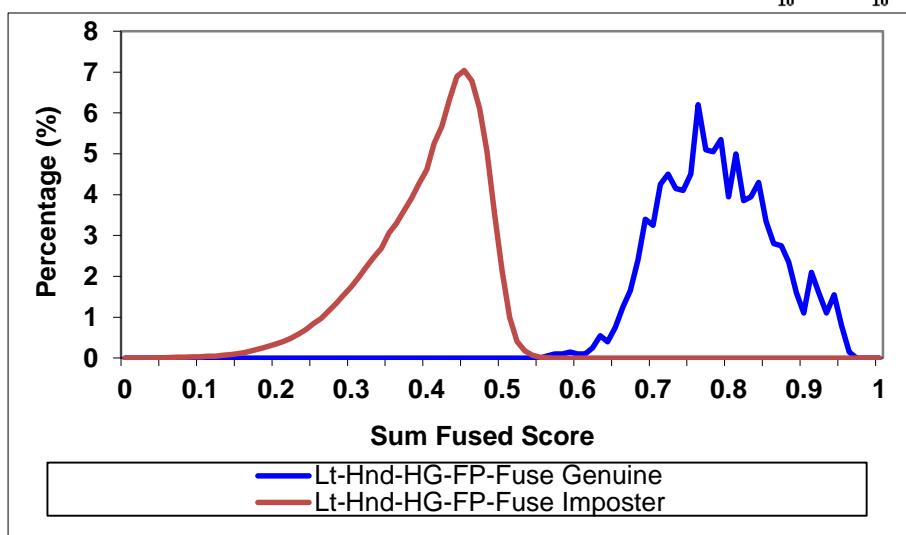
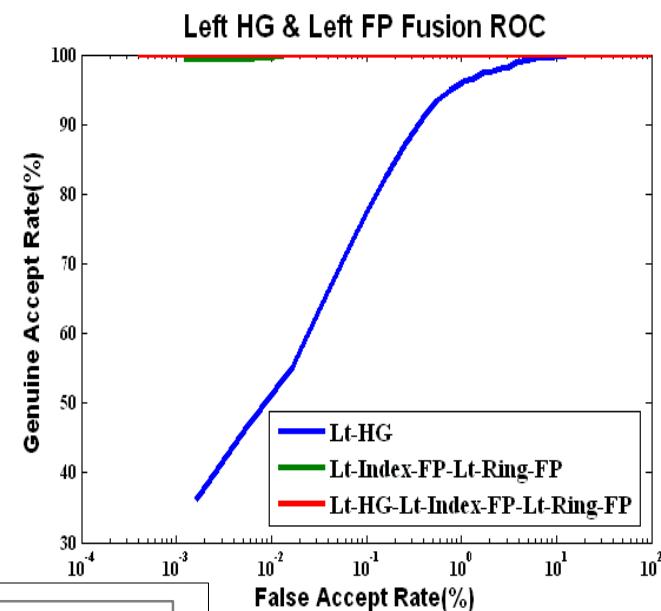
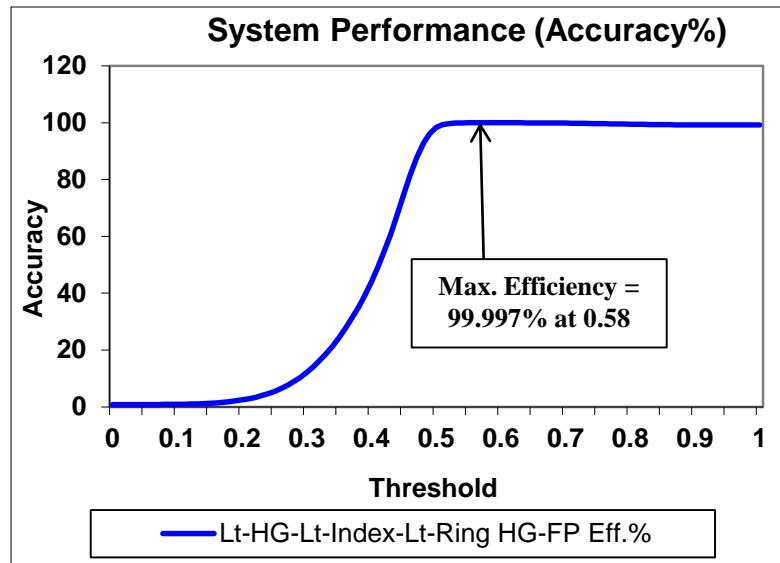
Modality	GAR %	GRR%	FAR%	FRR %	Accuracy %	TER%	EER%
Before Rt-HG & Rt-FP Score Fusion							
Rt-HG	82.90	99.9086 8	0.09131 3	17.10	99.772345 [0.018]	2.16929 3 [0.027]	1.192525 [0.026]
Rt-Ring-Rt- Index FP	99.30	99.9923 2	0.00767 7	0.70	99.986774 [34%]	0.112323 [27%]	0.089798 [28%]
After Rt-HG-FP Score Fusion							
Rt-HG-Rt-Ring- FP-Rt-Index-FP HG-FP	99.85	100.000 0	0.00080 8	0.30	99.996794 [0.61]	0.006868 7 [0.58]	0.003434 3 [0.58]

Quantitative
Analysis

Lt-HG-FP Biometric System



Lt-HG-FP Biometric System: Results



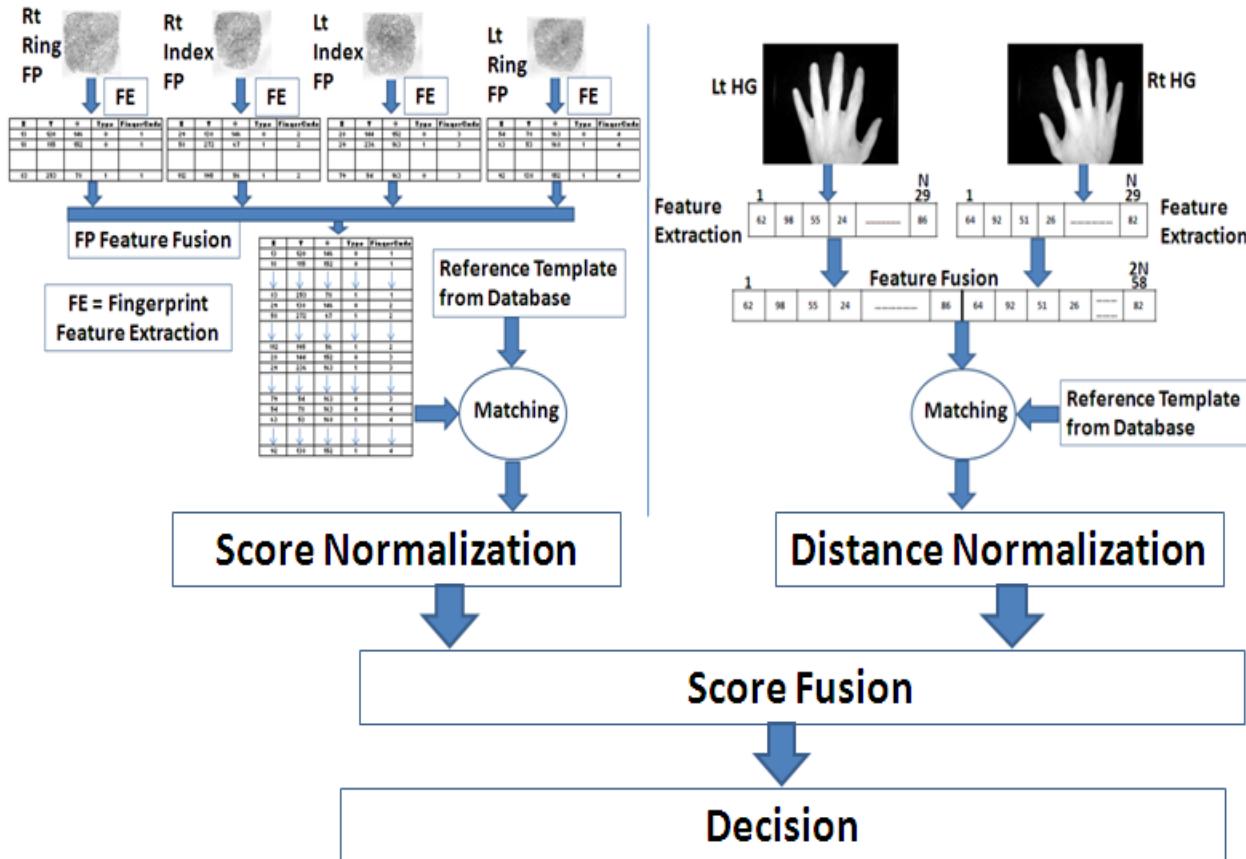
Qualitative Analysis

Lt-HG-FP Biometric System: Results

Modality	GAR %	GRR%	FAR%	FRR%	Accuracy %	TER%	EER%
Before Lt-HG & Lt-FP Score Fusion							
Lt-HG	77.50	99.8989 9	0.101010	22.50 0	99.719439 [0.018]	4.38495 0 [0.026]	2.285051 [0.027]
Lt-Index-Lt-Ring FP	99.20	100.000 0	0.001211 2	0.800	99.992385 [34%]	0.02666 7 [28%]	0.013333 [28%]
After Lt-HG-FP Score Fusion							
Lt-HG-Lt-Index-FP-Lt-Ring-FP HG-FP	100.0 0	100.000 0	0.000808	0.250	99.99719 4 [0.58]	0.0068687 [0.55]	0.003434 3 [0.55]

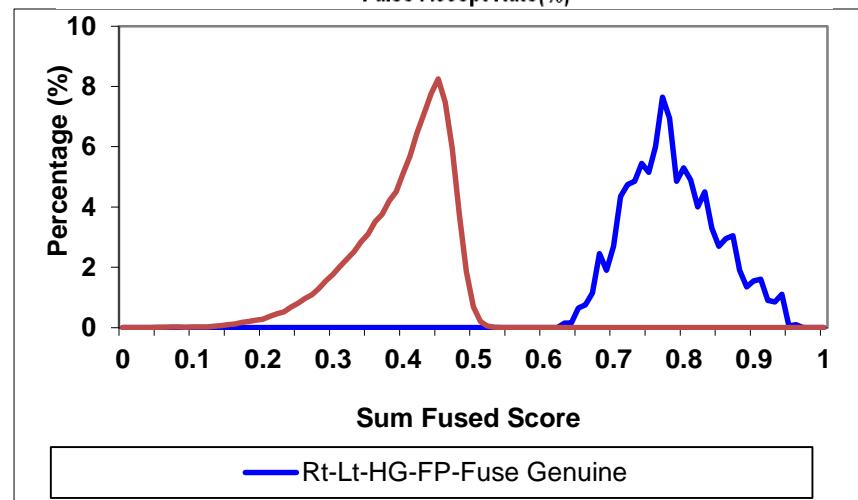
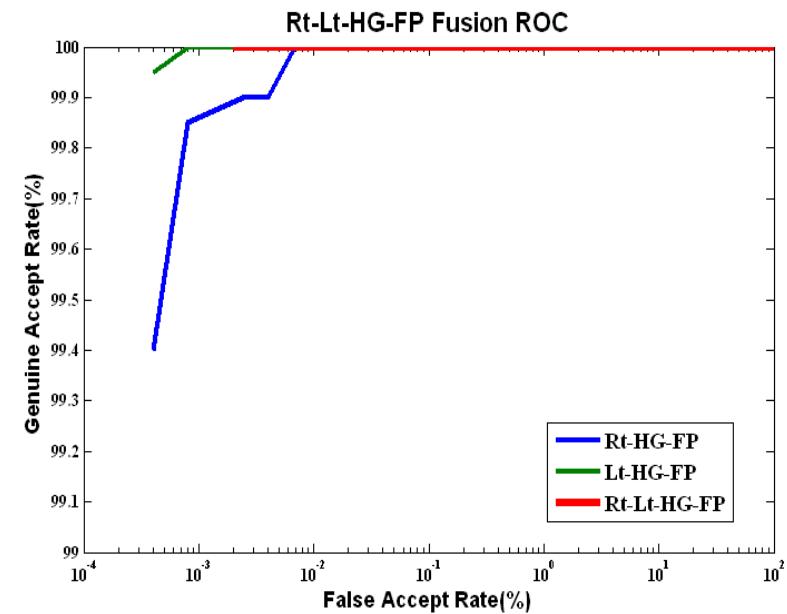
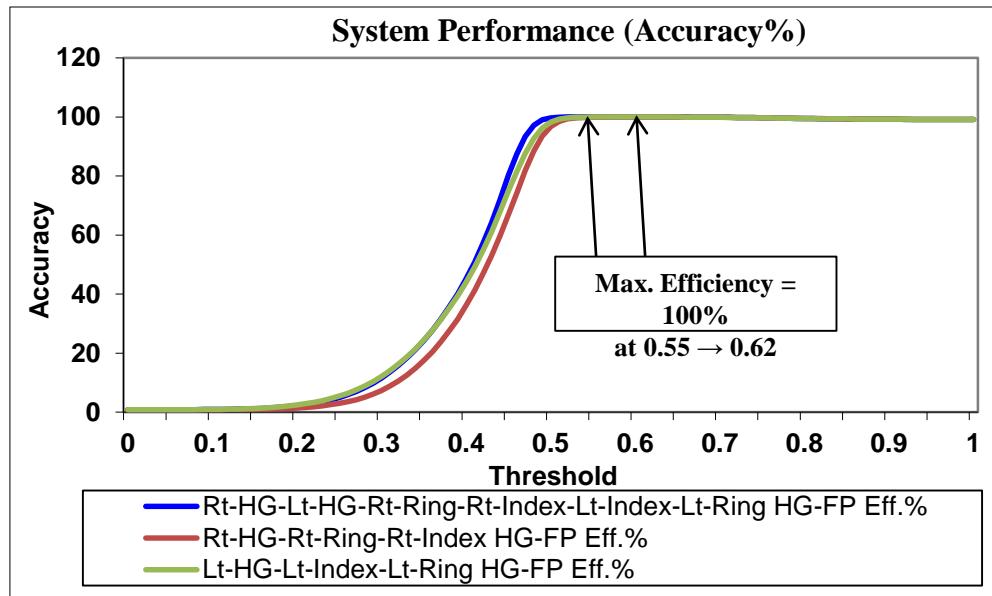
Quantitative
Analysis

Rt-Lt-HG-FP Biometric System



Rt-Lt Bi-Modal FP-HG
Biometric System

Rt-Lt-HG-FP Biometric System: Results



Qualitative
Analysis

Rt-Lt-HG-FP System: Results (Cont.)

Modality	GAR %	GRR %	FAR%	FRR%	Accuracy%	TER%	EER%
After Lt-FP Feature Fusion & Before HG-FP Score Fusion							
Rt-HG-FP	99.85	100	0.00080 8	0.30000 0	99.996794 [0.61]	0.006868 7 [0.58]	0.003434 3 [0.58]
Lt-HG-FP	100.0 0	100	0.00080 8	0.25000 0	99.997194 [0.58]	0.006868 7 [0.55]	0.003434 3 [0.55]
After Lt HG-FP Score Fusion							
Rt-Lt-HG-FP	100.0 0	100	0.00000 0	0.00000 0	100.000000 [0.55-0.62]	0.000000 [0.55- 0.62]	0.000000 [0.55- 0.62]

Quantitative
Analysis

Publications and Contributions

- **Conference:** M. K. Shahin, A. M. Badawi, and M. E. Rasmy, "A multimodal hand vein, hand geometry, and fingerprint prototype design for high security biometrics," *Proc. of 4th IEEE Cairo Int. Biomedical Engineering Conf. (CIBEC2008), Egypt*, Dec. 2008.
- **Journal:** M. K. Shahin, A. M. Badawi, and M. E. Rasmy, "Multimodal Biometric System Based on Near-Infra-Red Dorsal Hand Geometry and Fingerprints for Single and Whole Hands," *International Journal of computer and Information Engineering*, 2011.

Conclusions

- 1. MMBS based on Hand Geometry & Fingerprints was proposed**
- 2. Novel Hand-Made MMBS Prototype was designed**
- 3. For testing and validating the proposed MMBS dataset of 100 persons were experimentally acquired.**
- 4. 3000 images were in our database: (5 Rt-HG, 5 Lt-HG, 5 Rt-Ring-FP, 5 Rt-Index-FP, 5 Lt-Index-FP, and 5 Lt-Ring-FP) for each person.**
- 5. Feature Extraction for HG and FP**
- 6. Feature Fusion for Similar Biometric Traits**
- 7. Score Fusion for Different Biometric Signals**

Modality	FAR%	FRR%	Accuracy%	TER%	EER%
HG Single-Modalities					
Rt-HG	0.091313	17.10	99.772345 [0.018]	2.169293 [0.027]	1.192525 [0.026]
Lt-HG	0.101010	22.50	99.719439 [0.018]	4.384950 [0.026]	2.285051 [0.027]
FP Single-Modalities					
Rt-Ring-FP	0.030707	9.40	99.894188 [42%]	1.557172 [30%]	0.778586 [30%]
Rt-Index-FP	0.018182	6.70	99.928257 [41%]	0.928889 [28%]	0.658636 [30%]
Lt-Index-FP	0.029899	7.30	99.911824 [39%]	1.417071 [27%]	0.752020 [28%]
Lt-Ring-FP	0.008081	6.00	99.943888 [37%]	0.788384 [26%]	0.414141 [25%]
HG Feature Fusion					
Rt-Lt-HG	0.016162	7.15	99.926653 [0.02]	1.015758 [0.027]	0.507879 [0.027]
FP Biometric Feature Fusion					
Rt-FP	0.007677	0.70	99.986774 [34%]	0.112323 [27%]	0.089798 [28%]
Lt-FP	0.001211	0.80	99.992385 [34%]	0.026667 [28%]	0.013333 [28%]
Rt-Lt-FP	0.000404	0.05	99.999198 [35%]	0.0020202 [34%]	0.001010 [34%]
Rt-HG-FP Score Fusion					
Rt-HG-FP	0.000808	0.30	99.996794 [0.61]	0.006868 [0.58]	0.003434 [0.58]
Lt-HG-FP Score Fusion					
Lt-HG-FP	0.000808	0.25	99.997194 [0.58]	0.006868 [0.55]	0.003434 [0.55]
Rt-Lt-HG-FP Score Fusion					
Rt-Lt-HG-FP	0.000000	0.00	100.00 [0.55-0.62]	0.00 [0.55-0.62]	0.00 [0.55-0.62]