

BRAIN. Broad Research in Artificial Intelligence and Neuroscience

e-ISSN: 2067-3957 | p-ISSN: 2068-0473

Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science

Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)

2025, Volume 16, Issue 4, pages: 185-211

Submitted: April 24th, 2025 | Accepted for publication: August 23rd, 2025

Workbench Review for Biometrics Face Verification Systems Analysis Studies

Mohamed Soltane

Electrical Engineering Department, Faculty of Technology Yahia Fares University of Medea, Algeria; Ministry of Higher Education and Scientific Research, Algeria (MHESR), Algeria.

soltane.mohamed@univ-medea.dz &

soltane.mohamed.3099@gmail.com

<https://orcid.org/0000-0002-5474-5740>

Mourad Oussalah

CMVS, Faculty of ITEE, University of Oulu, Finland.

<https://orcid.org/0000-0002-4422-8723>

Aissa Belmeguenai

Electrical Engineering Department, Faculty of Engineering Sciences, 20 August 1955 University of Skikda, 21000 Skikda, Algeria.

Abstract: *Computer vision is a promoted field of artificial intelligence (AI) that uses machine learning algorithms to teach computers and systems to derive meaningful information from digital images. Pattern recognition has long been a goal of computer vision, but recently reliable automated face recognition has become a realistic target of biometrics development research. In this paper pattern recognition versus verification systems has been diagnosed, the contribution of classifier analysis to Biometrics Face Verification performance is examined. The discussion covers classification paradigms, the use of multiple observations, and their judicious fusion at the data; it should improve the correct decision performance at the different decision fusions levels. The fusion tasks reported in this development work were carried through fusion of two well-known face recognisers, ICA I and ICA II. Novel strategy based on Likelihood Ratio Fusion within scores is employed, in which face recognition accuracy is optimised and limitations of single recogniser have been reduced. The performance of the analysis studies were tested based on three different face databases and the Dynamic Still Face Databases eNTERFACE2005 (Stylianou et al., 2005), with simulation results demonstrating significant performance achievements.*

Keywords: *classifier; fusion; biometrics; face verification; PCA; LDA; ICA and Likelihood; parameter estimate.*

How to cite: Soltane, M., Oussalah, M., & Belmeguena, A. (2025). Workbench review for biometrics face verification systems analysis studies. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 16(4), 185-211. <https://doi.org/10.70594/brain/16.4/11>

1. Introduction

BIO-METRIC: A stemming composite word synthesis of “bio” and “metric”, a body measurements and calculations related to human characteristics and features. The science of biometrics is concerned with the accurate measurement of unique biological characteristics of an individual in order to securely identify them to a computer or other electronic system. Biological characteristics measured typically include fingerprints, voice patterns, retinal and iris scans, facial patterns, and even the chemical composition of an individual's DNA.

Biometric Authentication (BA) is generally acknowledged as providing a high-level of protection because it is difficult to steal or transfer biological material or features from one user to another. Biometric authentication (*Am I whom I claim I am?*) involves confirming or denying a person's *claimed identity* based on his/her physiological or behavioural characteristics. BA is becoming an important alternative to traditional authentication methods such as keys (“something one has”, i.e., by possession) or PIN numbers (“something one knows”, i.e., by knowledge) because it is essentially “who one is”, i.e., by biometric information. Therefore, it is not susceptible to misplacement or forgetfulness (Sofia & Panagiotis, 2004; Girija & Michael, 2006; Norman & Samy, 2004).

These biometric systems for personal authentication and identification are based upon physiological or behavioural features which are typically distinctive, although time varying, such as fingerprints, hand geometry, face, voice, lip movement, gait, and iris patterns. Multi-biometric systems, which consolidate information from multiple biometric sources, are gaining popularity because they are able to overcome limitations such as non-universality, noisy sensor data, significant intra-user variations, and susceptibility to spoof attacks that are commonly encountered in mono-biometric systems.

The ultimate goal of designing pattern recognition systems is to achieve the best possible classification performance for the task at hand. Traditionally, this objective has led to the development of different classification schemes for any given pattern recognition problem to be solved. The results of an experimental assessment of the different designs would then be the basis for choosing one of the classifiers as a final solution to the problem. It has been observed in such design studies that, although one of the designs would yield the best performance, the sets of patterns misclassified by different classifiers do not necessarily overlap. This suggested that different classifier designs potentially offered complementary information about the patterns to be classified, which could be harnessed to improve the performance of the selected classifier.

On the other hand, appearance-based approaches, such as PCA, LDA, and ICA based methods, have significantly advanced face recognition techniques. These approaches generally operate directly on image-based representations, features into a subspace derived from training images. In addition these linear methods can be extended using nonlinear kernel techniques to address nonlinearity in face recognition. Although kernel methods may achieve good performance on training data, this may not be the case for unseen data due to their higher flexibility compared to linear methods and the possibility of overfitting.

Several studies on multi-classifiers biometric verification systems have been reported in literature. Reddy et al., (2024) presents a novel approach to biometric recognition using a multi-modal system that combines face and iris data. By employing a Gradient Neural Network, the method significantly enhances the accuracy, sensitivity, and specificity of biometric identification systems. Incorporating advanced feature extraction techniques via ResNet-101 and WaveNet, the system effectively manages biometric inputs to provide robust security against unauthorised access. The results demonstrate the method's superiority over traditional biometric classifiers, affirming its potential for practical applications in sensitive domains. Somashekhar & Nijagunarya (2018) present a multimodal biometric system incorporating both face and fingerprint data, using advanced feature extraction methods such as Gabor and SIFT, combined at both the feature and decision levels. Their results demonstrate improved accuracy and robustness over traditional unimodal systems, suggesting such an approach is promising for large-scale identity authentication. Sun &

Redei (2022) investigates the implementation of CNN-based models for facial recognition, inspired by the DeepFace architecture but optimised for different resources and constraints. The research achieved notable accuracy and demonstrates the potential for substantial results with appropriate model architecture and data handling. Alharbi & Alshanbari (2023) introduced a face-voice multimodal biometric authentication system enhancing security through the integration of advanced models: FaceNet for face recognition and Gaussian mixture models for voice recognition. Their work explores methodologies of noise removal and feature extraction, emphasising the reduced equal error rate (EER) when compared to previous systems to improve overall authentication accuracy. Batagelj & Solina (2006) present a comparative study for Face recognition applications based on the most popular appearance-based face recognition projection methods (PCA, LDA and ICA), tested under equal working conditions regarding preprocessing and algorithm implementation on the FERET dataset with its standard tests. They report that the L1 metric gives the best results in combination with PCA and ICA1, while COS is superior to any other metric when used with LDA and ICA2. Thamaraimanalan et al., (2021) categorise data variability based on biometric sample capture and analysis, implementing the image quality parameters to enhance biometric authentication systems and analyse both hardware and software-based methods to improve fake property detection, advocating the use of specialised sensors and software algorithms. Becker & Ortiz (2009) present an evaluation of multiple face recognition algorithms specifically applied to Facebook, highlighting discrepancies in effectiveness when transitioning from controlled datasets to real-world data. Through extensive testing across different datasets, the study focuses on algorithm efficiency alongside accuracy, proposing a balance between practicality and performance for advancements in face recognition technology. Delac et al., (2012) conduct a detailed comparative study of PCA, ICA, and LDA in face recognition tasks, highlighting the importance of selecting the appropriate combination of algorithm and metric for specific tasks. They outline the methodology, experimental design, and evidence, drawing conclusions about the effectiveness of different techniques for handling temporal changes, illumination, and expression variations in face recognition.

This paper proposes a system that analyses face biometrics verification systems using ICA I and ICA II recognisers, evaluated on three different face databases: the Olivetti Research Laboratories (ORL, 1994) face database (ORL, 1994; Samaria & Harter, 1994), the Indian face database (Jain & Mukherjee, 2002) and the eNTERFACE2005 Dynamic Face Database (Stylianou et al., 2005). These two recognisers are capable of considering the problem of dimensionality by eliminating redundant features and reducing the feature space. The system can handle variations (illumination, pose, etc.) up to a significant level. The integrated system can overcome the drawbacks of individual recognisers.

2. Biometrics Face Recognition

Face recognition, authentication, and identification are often confused. Face recognition is a broad topic that includes both face identification and face authentication (also called verification). On the one hand, face authentication is concerned with validating a claimed identity based on the image of a face, and either accepting or rejecting the identity claim (one-to-one matching). On the other hand, the goal of face identification is to identify a person based on the image of a face. This face image must be compared with all the registered persons (one-to-many matching). Thus, the key issue in face recognition is to extract the meaningful features that characterise a human face. Accordingly, there are two major tasks: Face detection and face verification.

2.1. Face Detection

Face detection is concerned with finding whether or not there are any faces in a given image (usually in gray scale) and, if present, return the image location and content of each face. This is the first step of any fully automatic system that analyses the information contained in faces (e.g., identity, gender, expression, age, race and pose). While earlier work dealt mainly with upright

frontal faces, several systems have since been developed that are able to detect faces fairly accurately with in-plane or out-of-plane rotations in real time. For biometric systems that use faces as non-intrusive input modules, it is imperative to locate faces in a scene before any recognition algorithm can be applied. An intelligent vision-based user interface should be able to determine the user's focus of attention (i.e., where the user is looking at) in order to respond accordingly. To detect facial features accurately for applications such as digital cosmetics, faces need to be located and registered first to facilitate further processing. It is evident that face detection plays a vital and critical role in the success of any face processing systems.

For the results presented in this paper, only size normalisation of the extracted faces was applied. All face images were resized to 130 x 150 pixels, using a bi-cubic interpolation. After this stage, a position correction algorithm was developed, based on detecting the eyes into the face, applying rotation, and resizing to align the eyes of all pictures to the same coordinates. The face detection and segmentation tasks presented in this paper were performed based on the "Face analysis in Polar Frequency Domain" method proposed by Zana et al., (2005). First, the Fourier-Bessel (FB) coefficients were extracted from the images. Next, the Cartesian distance between all of the Fourier-Bessel transformation (FBT) representations was computed, and each object was redefined by its distance to all others. Images were transformed by a FBT up to the 30th Bessel order and 6th root, with an angular resolution of 3°, thus obtaining 372 coefficients. These coefficients correspond to a frequency range of up to 30 and 3 cycles per image of angular and radial frequency, respectively.

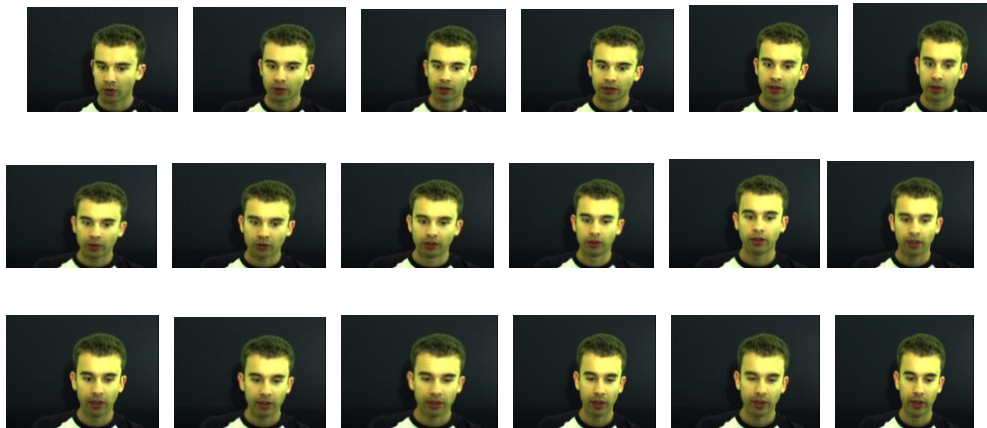


Figure 1. 18 still faces extracted from video for user 1 of eINTERFACE2005 dynamics face database (Stylianou et al., 2005)



Figure 2. Face & Eyes Detections for different users from TWO different DATABASES, First row from eINTERFACE 2005 (Stylianou et al., 2005) and Second row from ORL 94 (ORL, 1994; Samaria & Harter, 1994)

Figure 1 shows 18 still faces extracted from video for user 1 of the eNTERFACE2005 dynamics face database (Stylianou et al., 2005). Figure 2 shows the face and eyes detections for different users from two different databases, First row from eNTERFACE 2005 (Stylianou et al., 2005) and Second row from ORL 94 (ORL, 1994; Samaria & Harter, 1994). And Figure 3 shows the face normalization for the same users.



Figure 3. Face Normalization for the above users.

Polar Frequency Analysis:

The FB series is useful for describing the radial and angular components in images (Zana et al., 2005). FBT analysis starts by converting the coordinates of a region of interest from Cartesian (x, y) to polar (r, θ) . The $f(r, \theta)$ function is represented by the two-dimensional FB series, defined as:

$$f(r, \theta) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} A_{n,i} J_n(\alpha_{n,i} r) \cos \cos(n\theta) + \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} B_{n,i} J_n(\alpha_{n,i} r) \sin \sin(n\theta) \quad (1)$$

Where J_n is the Bessel function of order n , $f(R, \theta) = 0$ and $0 \leq r \leq R$. $\alpha_{n,i}$ is the i^{th} root of the J_n function, i.e. the zero crossing value satisfying $J_n(\alpha_{n,i}) = 0$ is the radial distance to the edge of the image. The orthogonal coefficients $A_{n,i}$ and $B_{n,i}$ are given by:

$$A_{0,i} = \frac{1}{\pi R^2 J_1^2(\alpha_{n,i})} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=R} f(r, \theta) r J_n\left(\frac{\alpha_{n,i}}{R} r\right) dr d\theta \quad (2)$$

if $B_{0,i} = 0$ and $n = 0$;

$$[A_{n,i} \ B_{n,i}] = \frac{2}{\pi R^2 J_{n+1}^2(\alpha_{n,i})} \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=R} f(r, \theta) r J_n\left(\frac{\alpha_{n,i}}{R} r\right) [\cos \cos(n\theta) \ \sin \sin(n\theta)] dr d\theta \quad (3)$$

if $n > 0$.

An alternative method to polar frequency analysis is to represent images using polar Fourier transform descriptors. The polar Fourier transform is a well-established mathematical operation in which, after converting the image coordinates from Cartesian to polar (as described above), a conventional Fourier transformation is applied. These descriptors are directly related to radial and angular components, but are not identical to the coefficients extracted by the FBT.

2.2. Face Verification

Feature Extraction:

A. Principal Component Analysis (PCA)

Principal component analysis (PCA) (Turk & Pentland, 1991) is based on the second-order statistics of the input image, which tries to attain an optimal representation that minimises the

reconstruction error in a least-squares sense. The eigenvectors of the covariance matrix of the face images constitute the eigenfaces. The dimensionality of the face feature space is reduced by selecting only the eigenvectors possessing largest eigenvalues. Once the new face space is constructed, when a test image arrives, it is projected onto this face space to yield the feature vector—the representation coefficients in the constructed face space. The classifier determines the identity of the individual, according to a similarity score between the test image's feature vector and the PCA feature vectors of the individuals in the database.

PCA is closely related to the Karhunen-Loève Transform (KLT) (Shakhnarovich & Moghaddam, 2004; Loève, 1955), which was derived in the signal processing context as the orthogonal transform with the basis $\Phi = [\phi_1, \dots, \phi_N]^T$ that for any $k \leq N$ minimise the average L_2 reconstruction error for data points x

$$\epsilon(x) = \|x - \sum_{i=1}^k (\phi_i^T x) \phi_i\|. \quad (4)$$

It can be shown that, under the assumption that the data is zero-mean; the formulations of PCA and KLT are identical. Without loss of generality we shall hereafter assume that the data is indeed zero-mean; that is, the mean face $\bar{x}$ is always subtracted from the data.

The basis vectors in KLT can be calculated in the following way. Let X be the $N \times M$ data matrix whose columns $x_1, \dots, x_M$ are *observations* of a signal embedded in R^N ; in the context of face recognition, M is the number of available face images and $N = mn$ is the number of pixels in an image. The KLT basis Φ is obtained by solving the eigenvalue problem $\Lambda = \Phi^T \Sigma \Phi$, where Σ is the covariance matrix of the data

$$\Sigma = \frac{1}{M} \sum_{i=1}^M x_i x_i^T, \quad (5)$$

$\Phi = [\phi_1, \dots, \phi_m]^T$ is the eigenvector matrix of Σ , and Λ is the diagonal matrix with eigenvalues $\lambda_1 \geq \dots \geq \lambda_N$ of Σ on its main diagonal, so that ϕ_j is the eigenvector corresponding to the j -th largest eigenvalue. Then it can be shown that the eigenvalue λ_i is the variance of the data projected on ϕ_i .

Thus, to perform PCA and extract k principal components of the data, one must project the data onto Φ_k - the first k columns of the KLT basis Φ , which correspond to the k highest eigenvalues of Σ . This can be seen as a linear projection $R^N \rightarrow R^k$ that retains the maximum energy (i.e. variance) of the signal. Another important property of PCA is that it *de-correlates* the data: the covariance matrix of $\Phi_k^T X$ is always diagonal.

The main properties of PCA are summarised by the following:

$$x \approx \Phi_k y, \quad \Phi_k^T \Phi_k = I, \quad E\{y_i y_j\}_{i \neq j} = 0 \quad (6)$$

Namely, approximate reconstruction, ortho-normality of the basis Φ_k and de-correlated principal components $y_i = \phi_i^T x$, respectively. Where PCA is successful in finding the principal manifold where it is less successful, due to clear nonlinearity of the principal manifold.

PCA may be implemented via Singular Value Decomposition (SVD): The SVD of a $M \times N$ matrix X ($M \geq N$) is given by

$$X = U D V^T, \quad (7)$$

Where the $M \times N$ matrix U and the $N \times N$ matrix V have ortho-normal columns, and the $M \times N$ matrix D has the singular values of X on its main diagonal and zero elsewhere. It can be shown that $U = \Phi$, so that SVD allows efficient and robust computation of PCA without the need to estimate the data covariance matrix Σ in Eq. (5). When the number of examples M is much smaller than the dimension N .

B. Linear Discriminate Analysis (LDA)

Linear Discriminate Analysis (LDA) also called Fisher Linear Discriminant (FLD) (Belhumeur et al., 1997; Bn-Yacoub et al., 1999) is an example of a class specific subspace method that finds the optimal linear projection for classification. Rather than finding a projection that maximises the projected variance as in principal component analysis, FLD determines a projection, $y = \Phi_F^T x$, that maximises the ratio between the between class scatter and the within-class scatter. Consequently, classification is simplified in the projected space.

Consider a c -class problem, with the between-class scatter matrix given by:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (8)$$

And the within-class scatter matrix by:

$$S_W = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad (9)$$

where μ is the mean of all samples, μ_i is the mean of class i , and N_i is the number of samples in class i . The optimal projection Φ_F is the projection matrix which maximises the ratio of the determinant of the between-class scatter to the determinant of the within-class scatter of the projections,

$$\Phi_F = \arg \frac{|\Phi^T S_B \Phi|}{|\Phi^T S_W \Phi|} = [w_1 \ w_2 \ \dots \ w_m] \quad (10)$$

where $\{i = 1, 2, \dots, m\}$ is the set of generalised eigenvectors of S_B and S_W , corresponding to the m largest generalised eigenvalues $\{\lambda_i \mid i = 1, 2, \dots, m\}$.

However, the rank of S_B is $c-1$ or less since it is the sum of c matrices of rank one or less. Thus, the upper bound on m is $c-1$.

To avoid the singularity, one can apply PCA first to reduce the dimension of the feature space to $N-c$, and then use FLD to reduce the dimension to $c-1$. This two-step procedure is used in computing "Fisher Faces".

C. Independent Component Analysis (ICA)

Independent component analysis (ICA) (Hyvärinen & Oja, 2000; Lu & Rajapakse, 2001; Vaseghi & Jetelová, 2009; Bartlett et al., 1998) is a statistical method for linear transforming an observed multidimensional random vector X into a random vector Y whose components are stochastically as independent from each other as possible. Several procedures to find such transformations have been recently developed in the signal processing literature, relying either on Comon's (Comon, 1994) information-theoretic approach or Hyvärinen's maximum negentropy approach (Hyvärinen, 1997). The fundamental goal of ICA is to find a representation $Y = MX$ (M not necessarily a square matrix) in which the transformed components Y_i are the least statistically dependent. ICA leads to meaningful results whenever the probability distribution of X is far from Gaussian and this is the case that we are interested in this paper. For the face recognition task were proposed two different architectures: Architecture I, which has statistically independent basis images (ICA I), and Architecture II, which assumes that the sources are independent coefficients (ICA II). These coefficients give the factorial code representation. Architecture I provides a more localised representation for faces, while ICA Architecture II, similar to PCA, provides a more holistic representation. ICA I produces spatially localised features influenced only by small parts of an image, thus isolating particular parts of faces. For this reason, ICA I is optimal for recognising facial actions and suboptimal for recognising temporal changes in faces or images taken under different conditions. Preprocessing steps of the methods ICA involves a PCA process by vertically centering (for ICA I), and whitened PCA process by horizontally centering (for ICA II). ICA Architecture I include a PCA by vertically centering (PCA I):

$$P_v = X_v V^T \Lambda^{-1/2} \quad (11)$$

where X_v is the vertically-centered training image column data matrix. Symbols Λ and V correspond to largest eigenvalues and eigenvectors of S_T matrix respectively:

$$S_T = \sum_{i=1}^M (x_i - \mu_v) \cdot (x_i - \mu_v)^T, \quad \mu_v = \frac{1}{N} \sum_{j=1}^N x_i \quad (12)$$

In contrast to standard PCA, PCA I removes the mean of each image while standard PCA removes the mean image of all training samples. ICA Architecture II includes a whitened PCA by horizontally centering (PCA II):

$$P_w = P_h \cdot \left(\frac{1}{M} \Lambda \right)^{-1/2} = \sqrt{M} X_h \Lambda^{-1/2}, \quad (13)$$

where P_h is the projection matrix of standard PCA method:

$$P_h = X_h \Lambda^{-1/2}. \quad (14)$$

Matrix X_h contains in rows horizontally-centered training images. PCA II is actually the whiten version of standard PCA.

3. Multimodal Biometric Fusion Decision

The process of biometric user authentication can be outlined as follows (Veeramachaneni et al., 2005): (a) acquisition of raw data; (b) extraction of features from these raw data; (c) computing a score for the similarity or dissimilarity between these features and a previously given

set of reference features; and (d) classification with respect to the score, using a threshold. The outcomes of the decision processing steps are either *true* or *false* (or *accept/reject*) for verification purposes, or the user's identity for identification scenarios.

The fusion of different signals can be performed (1) at the raw data or the feature level, (2) at the score level, or (3) at the decision level. Each of these approaches has its own advantages and disadvantages. For *raw data* or *feature-level* fusion, the base data must be compatible for all modalities and a common matching algorithm (processing step c) must be used. If these conditions are met, the separate feature vectors of the modalities easily could be concatenated into a single new vector. This level of fusion has the advantage that only one algorithm for further processing steps is necessary instead of one for each modality. Another advantage of fusing at this early stage of processing is that no information is lost due to prior processing steps. The main disadvantage is the demand for compatibility among the different raw data or features. The fusion at *score level* is performed by computing a similarity or dissimilarity (distance) score for each single modality. For joining these different scores, normalization should be done. The straightforward and most rigid approach for fusion is at the decision level. Here, each biometric modality results in its own decision; in the case of a verification scenario this is a set of *true's* and *falsies*. From this set a kind of voting (majority decision) or a logical *AND* or *OR* decision can be computed. This level of fusion is the least powerful due to the absence of detailed information. However, the advantage of this fusion strategy is the ease and the guaranteed availability of all single-modality decision results. In practice, score-level fusion is the best-researched approach, which appears to result in better improvements of recognition accuracy compared to the other strategies.

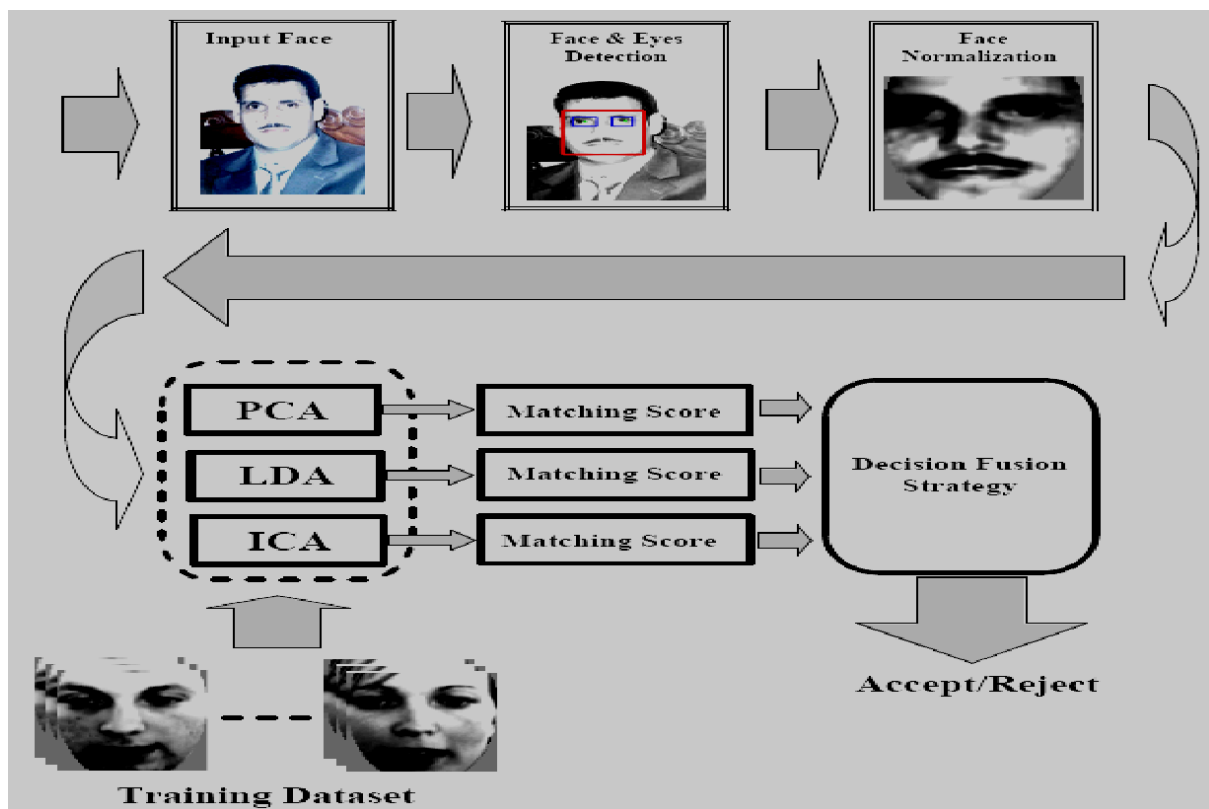


Figure 4. Multi-Classifier Face Verification System Framework.

Theoretical Analysis for Decision Level Fusion:

The fusion scheme using these two modalities is denoted by S . Verification system based only on modality one is denoted by S_1 , while on modality two by S_2 (Dongliang et al., 2005). If Γ is an algorithm, then the task is to find which acts on independent sources so that the output is maximised. It can be written as:

$$\hat{\Gamma} = \Gamma(S_1, S_2). \quad (15)$$

The performance indices in biometrics authentication system are false acceptance rate denoted FAR which means wrongly identifying an impostor to be an enrollee, and false rejection rate denoted by FRR which means wrongly identifying an enrollee as an impostor.

$$FAR(t) = P(w_0) = \int_{R_1} p(w_0) dX = 1 - \int_{R_0} p(w_0) dX, \quad (16)$$

$$FRR(t) = P(w_1) = \int_{R_0} p(w_1) dX \quad (17)$$

where w_1 denotes the genuine user while w_0 denotes the imposter one. R_0 and R_1 are two exclusive sets in the real axis. Both FAR and FRR are desirable to be as low as possible in the authentication system. For any biometrics authentication system, whatever classifier takes, there exists a great risk of error. From the viewpoint of Bayesian decision theory, this is represented by the following equations for a two class problem,

$$E(t) = C_r \times FRR(t) + C_a \times FAR(t), \quad (18)$$

$$E_i(t_i) = C_r^i \times FRR^i(t_i) + C_a^i \times FAR^i(t_i), \text{ for } i = 1, \dots, N \quad (19)$$

where, N is the total modalities number, C_r denotes the loss function pertinent to the false rejection, and C_a denotes the loss function for the false acceptance. For simplicity, it assume that $C_a = C_a^i$ and $C_r = C_r^i$.

Decision level fusion:

The integrated system is denoted by Ψ . The outputs by individual systems Ψ_1 and Ψ_2 , are called scores, which stand for the probability of claimant to be a genuine or an impostor. For any fusion strategies, an error is expressed as (18) and (19). If it assumes that $E_1(t_1) \leq E_2(t_2) \leq \dots \leq E_N(t_N)$, then it is easily known it is sufficient to prove that $E(t) \leq E_1(t_1)$. For a two-modality and Bayesian rule Fusion:

$$\text{Decide } w_0, \text{ if } (X_1, X_2) \in R \quad (20)$$

$$\text{Decide } w_1, \text{ otherwise}$$

Where

$$R = \{ (X_1, X_2) | C_r p(w_0) \geq C_a p(w_1) \}.$$

Since Ψ_1 and Ψ_2 are independent, it have:

$$p(w_0) = p_1(w_0) p_2(w_0), \quad (21)$$

&

$$p(w_1) = p_1(w_1)p_2(w_1), \quad (22)$$

Then:

$$\begin{aligned} FAR(t) &= 1 - \int_{R_0} p(w_0) dX_1 dX_2 \\ &= 1 - \int_{R_0} p_1(w_0) dX_1 \int_{R_0} p_2(w_0) dX_2 \quad (23) \\ &= 1 - \left(1 - FAR^1(t_1)\right) \left(1 - FAR^2(t_2)\right). \end{aligned}$$

&

$$\begin{aligned} FRR(t) &= 1 - \int_{R_0} p(w_1) dX_1 dX_2 \\ &= 1 - \int_{R_0} p_1(w_1) dX_1 \int_{R_0} p_2(w_1) dX_2 \quad (24) \\ &= FRR^1(t_1) FRR^2(t_2). \end{aligned}$$

From the Equations (23) & (24) it can be obviously seen that:

$FAR(t) = FAR^1(t_1)$ And $FAR(t) = FAR^2(t_2)$. Thus the two combined modalities cannot improve the false acceptance rate by the Bayesian decision rule. Otherwise $FRR(t) = FRR^1(t_1)$ and $FRR(t) = FRR^2(t_2)$. Hence the false rejection rate of the combined system is reduced compared to individual sub-classifiers.

Maximum Likelihood Parameter Estimation:

Given a set of observation data in a matrix X and a set of observation parameters θ the ML parameter estimation aims at maximising the likelihood $L(\theta)$ or log likelihood of the observation data $X = \{X_1, \dots, X_n\}$

$$\hat{\theta} = L(\theta). \quad (25)$$

Assuming that it has independent, identically distributed data, it can write the above equations as:

$$L(\theta) = p(\theta) = p(\theta) = \prod_{i=1}^n p(\theta). \quad (26)$$

The maximum for this function can be found by taking the derivative and setting it equal to zero, assuming an analytical function.

$$\frac{\partial}{\partial \theta} L(\theta) = 0. \quad (27)$$

The incomplete-data log-likelihood of the data for the mixture model is given by:

$$L(\theta) = \log(\theta) = \sum_{i=1}^N \log(\theta) \quad (28)$$

which is difficult to optimise because it contains the log of the sum. If it considers X as incomplete, however, and posits the existence of unobserved data items $Y = \{y_i\}_{i=1}^N$ whose values inform us which component density generated each data item, the likelihood expression is significantly simplified. That is, it assumes that $y_i \in \{1 \dots K\}$ for each i , and $y_i = k$ if the i -th sample was generated by the k -th mixture component. If it knows the values of Y , it obtains the complete-data log-likelihood, given by:

$$L(\theta, Y) = \log \log p(\theta) \quad (29)$$

$$= \sum_{i=1}^N \log \log p(\theta) \quad (30)$$

$$= \sum_{i=1}^N \log \log (p(\theta)p(y_i, \theta)) \quad (31)$$

$$= \sum_{i=1}^N \left(p_{y_i} + \log \log g(\mu_{y_i}, \Sigma_{y_i}) \right) \quad (32)$$

which, given a particular form of the component densities, can be optimised using a variety of techniques (Pham, 2008).

4. Experiments and Results Discussion

The experiments were performed using the Olivetti Research Laboratories (ORL, 1994) face database (ORL, 1994; Samaria & Harter, 1994), the Indian Face database (Jain & Mukherjee, 2002), and the eNTERFACE2005 Dynamic Face Database (Stylianou et al., 2005):

The ORL 94 (ORL, 1994; Samaria & Harter, 1994) Olivetti Research Laboratories (ORL, 1994) database of faces provides 10 sample images of each of 40 subjects (in PGM files format) in which four subjects are females and thirty six are males. The primary age range of the individuals captured seems to range from the late 20's to the mid 30's. Clearly there is an under representation of young and old people. The different images for each subject provide variation in views of the individual such as lighting, facial features (such as glasses), and slight changes in head orientation. It chose to use this face database because it seemed to be a standard set of test images used in much of the literature we encountered dealing with face recognition.

INDIAN FACES DATABASE 2002 (IFD2002; Jain & Mukherjee, 2002): This database contains human face images of 60 subjects (640 x 480 JPEG file format) with eleven different poses for each individual in which thirty nine are males and twenty two are females. The images were captured in February 2002 on the campus of the Indian Institute of Technology Kanpur. All images have a bright, homogeneous background, and the subjects are in an upright, frontal position. For each individual, it has included the following pose for the face: looking front, looking left, looking right, looking up, and looking up to the left, and looking up to the right, and looking down. In

addition to the variation in pose, images with four emotions—neutral, smile, laughter, and sad/disgust—are also included for every individual.

eNTERFACE 2005 Still Faces Databases (Stylianou et al., 2005): This database was extracted from video encoded in raw UYVY, AVI 640 x 480, 15.00 fps, with uncompressed 16bit PCM audio; mono, 32000 Hz, little endian. Uncompressed PNG files were extracted from the video files for feeding the face detection algorithms. The capturing devices for recording the video and audio data were: Allied Vision Technologies AVT Marlin MF-046C 10 bit ADC, 1/2" (8mm) progressive scan SONY IT CCD, and Shure SM58 microphone (frequency response 50 Hz to 15000 Hz. Unidirectional (Cardiod) dynamic vocal microphones). Thirty subjects participated in the experiments, of whom twenty-six were males and four were females. For each subject, the database obtained from eNTERFACE 2005 (Stylianou et al., 2005) was utilised. For each user of the still face experts, seventy-two still face images per subject were randomly selected for training, and the remaining forty-eight samples were used for subsequent validation and testing. Three sessions of the still face database were used separately: session one was used for training the still face experts, while sessions two and three were used to obtain expert opinions of known impostors and true claims.

For both Independent Component Analysis Architectures I and II, two novel processing design were employed, generating scores using three different similarity measures algorithms: Cosine Similarity, City Block, and Square Euclidean Distances Similarity (McCune & Grace, 2002; TIBCO, 2013). The performance of the analysis design for the likelihood ratio was evaluated based on EER (Equal Error Rate) of the ROC (Receptive Operation Characteristics) plotting evaluation.

For algorithmic processing design one (One Image Processing): at the preprocessing phase, each representation of the databases training and testing is normalised (divided by the square root of its power); at the fundamental processing phase, it computing the scores (Similarity Distances Measures) of the testing normalised representation with normalised training representation. For ORL, 1994 Face Databases, 60% of the data size are used for training and 40% of the data size are used for testing, resulting in $4 \times 6 = 24$ Similarity Distances for each user; thus $40 \times (4 \times 6) = 960$ target scores. For IFD2002 $60 \times (4 \times 6) = 1440$ target scores. For eNTERFACE 2005 still faces databases (Stylianou et al., 2005), 60% of the data size are used for training and 40% of the data size are used for validation and testing, resulting in $48 \times 72 = 3456$ Similarity Distances for each user; thus $30 \times (48 \times 72) = 103680$ target scores. These target scores, noted (*tscores*) are built using Normal Law of probability $N(\text{tscores}, t_m, t_s^2)$ of the mean t_m and standard deviation t_s . Scores for Distances Similarity Measures between the normalised representations of test users with the normalised representation of the training users are then computed; this results in $4 \times 6 \times 39 = 936$ scores for each user, and a total of $40 \times (4 \times 6 \times 39) = 37440$ non-target scores for ORL, 1994; and $60 \times (4 \times 6 \times 59) = 84960$ for IFD, 2002. For eNTERFACE 2005, this yields $48 \times 72 \times 29 = 100224$ scores for each user; with a total of $30 \times (48 \times 72 \times 29) = 3006720$ non-target scores. These non-target scores, noted (*nscores*), are built using Normal Law of probability $N(\text{nscores}, n_m, n_s^2)$, with mean n_m and standard deviation n_s .

Computing the logarithmic ratio of the likelihood for *tscores* and *nscores* with:

$$LR_{\text{tscores}} = \log \log \left(N(\text{tscores}, t_m, t_s^2) \right) - \log \log \left(N(\text{tscores}, n_m, n_s^2) \right) \quad (33)$$

&

$$LR_{\text{nscores}} = \log \log \left(N(\text{nscores}, t_m, t_s^2) \right) - \log \log \left(N(\text{nscores}, n_m, n_s^2) \right) \quad (34)$$

Then; Evaluated the ROC and the EER.

For algorithmic processing designs two (Multi-Image Processing), during the pre-processing phase each representation of the database for training and testing is normalised (divided on the square root of its power).

In the main processing phase, the scores (Similarity Distances Measures) of the testing normalised representation with normalised training representation are computed. For eINTERFACE 2005 still faces databases (Stylianou et al., 2005), 60% of the data size are used for training and 40% of the data size are used for validation and testing.

This results in $48 \times 72 = 3456$ Similarity Distances for each user. For each of them, the mean of these 3456 Distances is computed, providing 30 target scores for the 30 users. These target scores, noted (mtscores), are built using Normal Law of probability $N(mtscores, mtm, mts^2)$ of the mean mtm and standard deviation mts. Subsequently, the scores of Distances Similarity Measures between the normalised representations of the test users and the normalised representation of the training users are computed, resulting in $48 \times 72 \times 29 = 100224$ scores for each user. The mean of these 100224 scores, and the non-target scores are computed. These non-target scores, noted (mnscores), are built using Normal Law of probability $N(mnscores, mnm, mns^2)$ of the mean mnm and standard deviation mns.

Computing the Logarithmic Ration of the Likelihood for tscores and nscores with:

$$LR_{tscores} = \log \log \left(N(mtscores, , tm, mts^2) \right) - \log \log \left(N(mtscores, mnm, mns^2) \right) \quad (35)$$

&

$$LR_{nscores} = \log \log \left(N(mnscores, mtm, mts^2) \right) - \log \log \left(N(mnscores, mnm, mns^2) \right) \quad (36)$$

Then; Evaluated the ROC and the EER.

Figure 5 presents the Pattern Recognition Analysis Performance studies.

The Systems Used for the simulation studies are: i5 Intel Core Processor (2.66Ghz, 3MB L3 cache); 8 GB DDR3 Memory, Windows 8 Operating Systems, and MATLAB 2014a Programming Tools.

Figure 6 shows the Verification Analysis Studies based on Olivetti Research Laboratories (ORL, 1994) Face databases for Independent Components Analysis Classifier Fusion of Likelihood Ratio.

Figure 7 presents the Verification Analysis Studies based on Indian Face databases for Independent Components Analysis Classifier Fusion of Likelihood Ratio.

Figure 8 displays the Verification Analysis Studies based on eINTERFACE2005 Dynamic Face databases (Stylianou et al., 2005) for Independent Components Analysis Classifier Fusion of Likelihood Ratio.

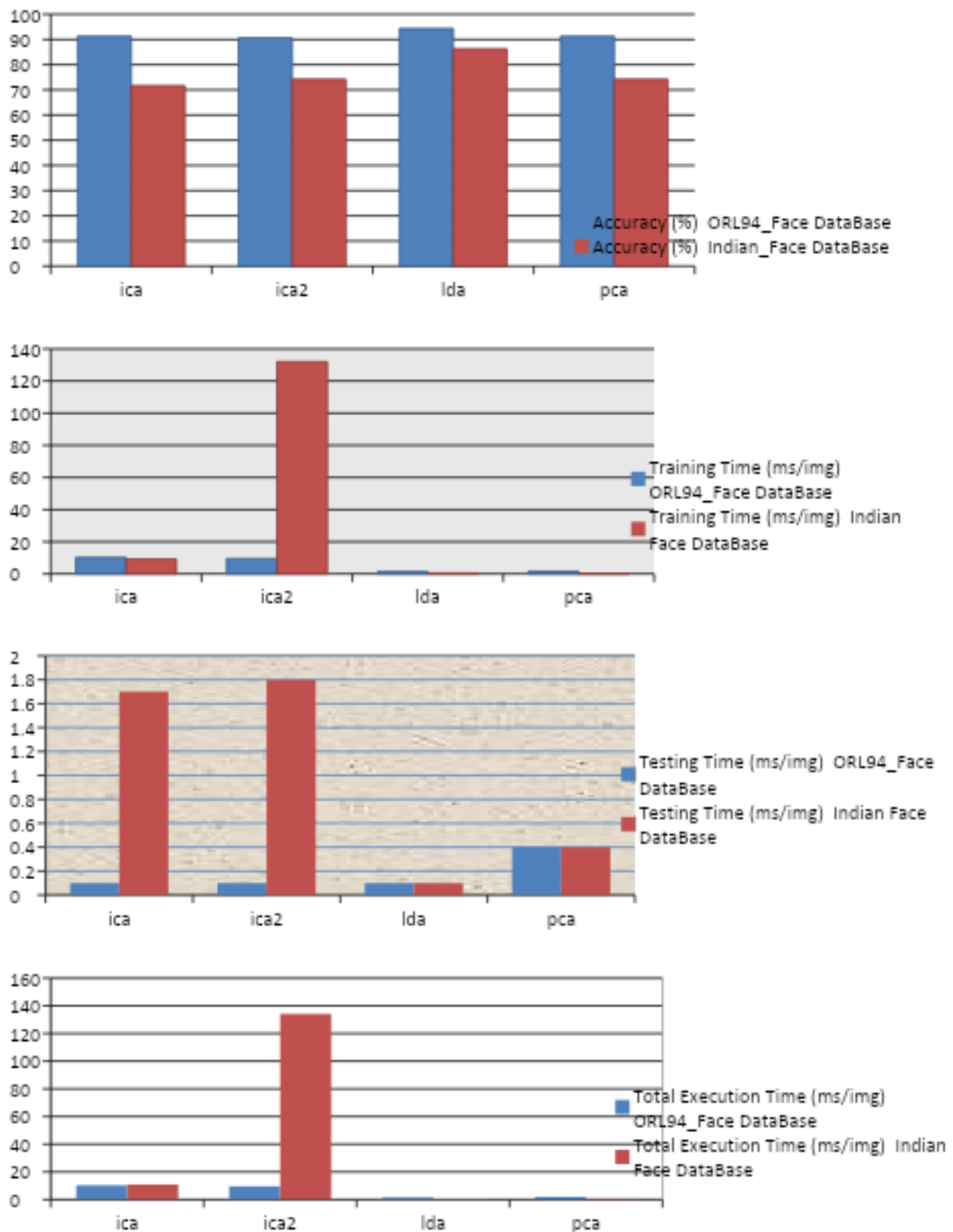
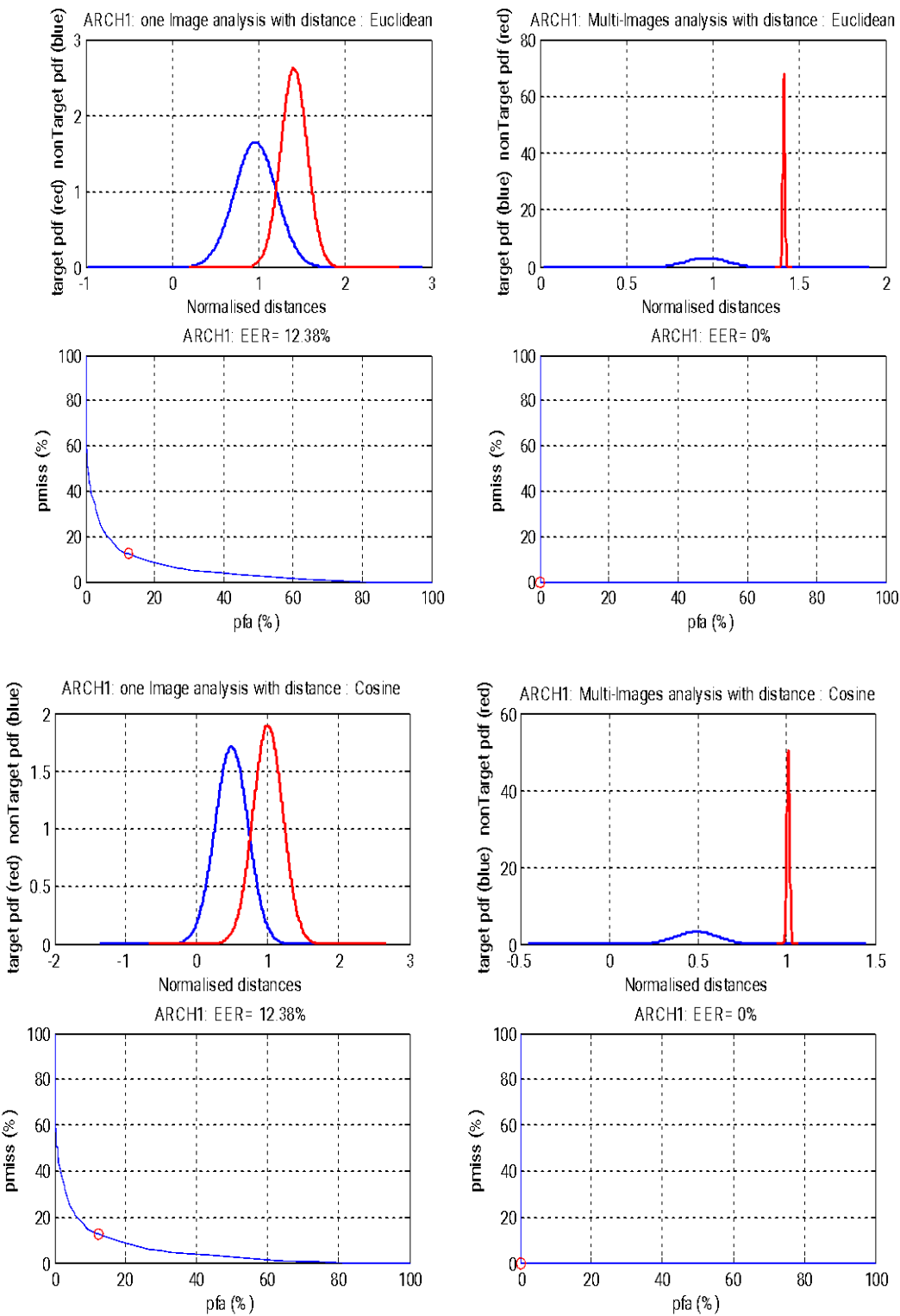
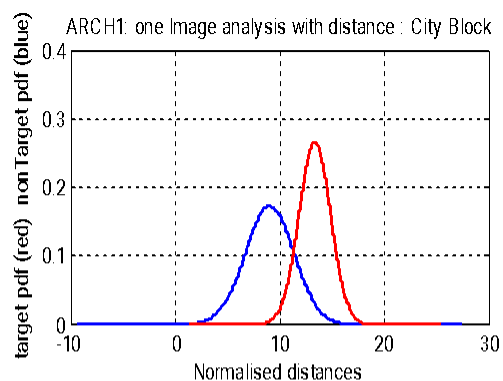
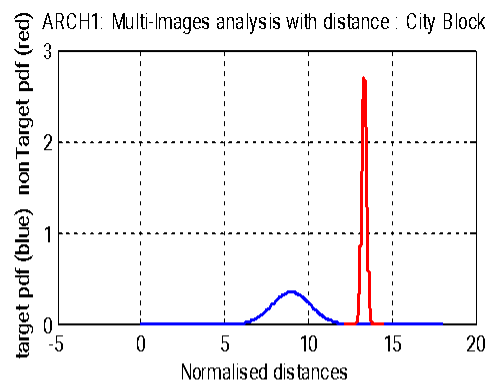
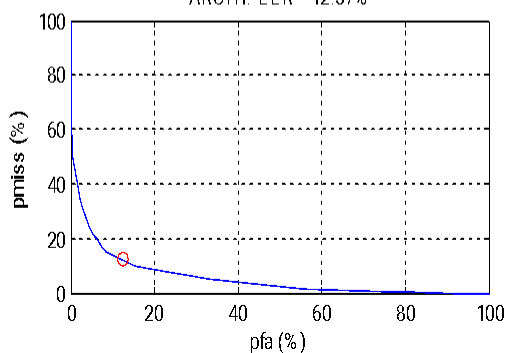


Figure 5. Pattern Recognition Analysis Performance studies

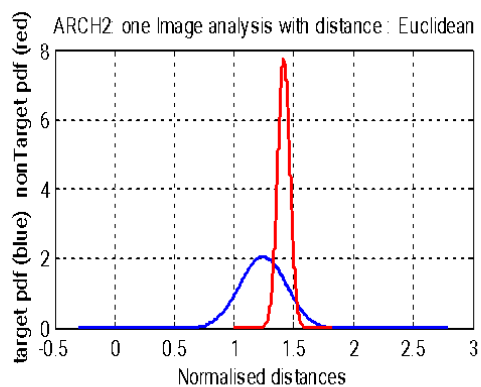
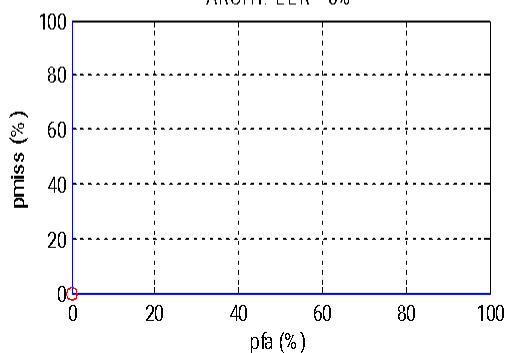




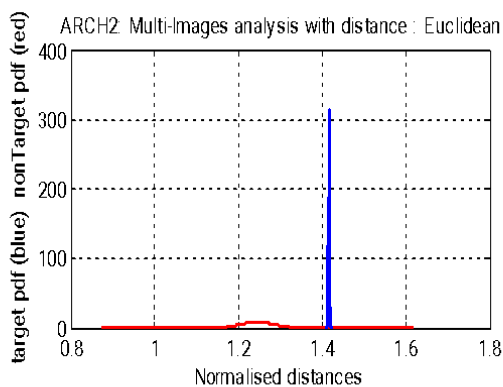
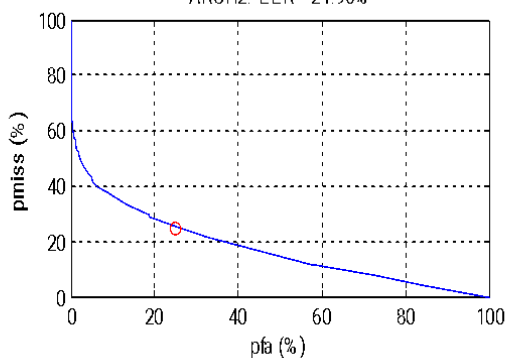
ARCH1: EER = 12.57%



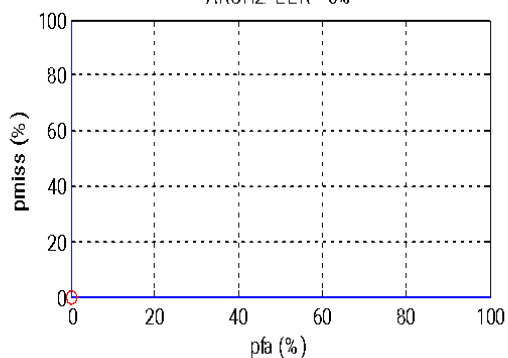
ARCH1: EER = 0%



ARCH2: EER = 24.96%



ARCH2: EER = 0%



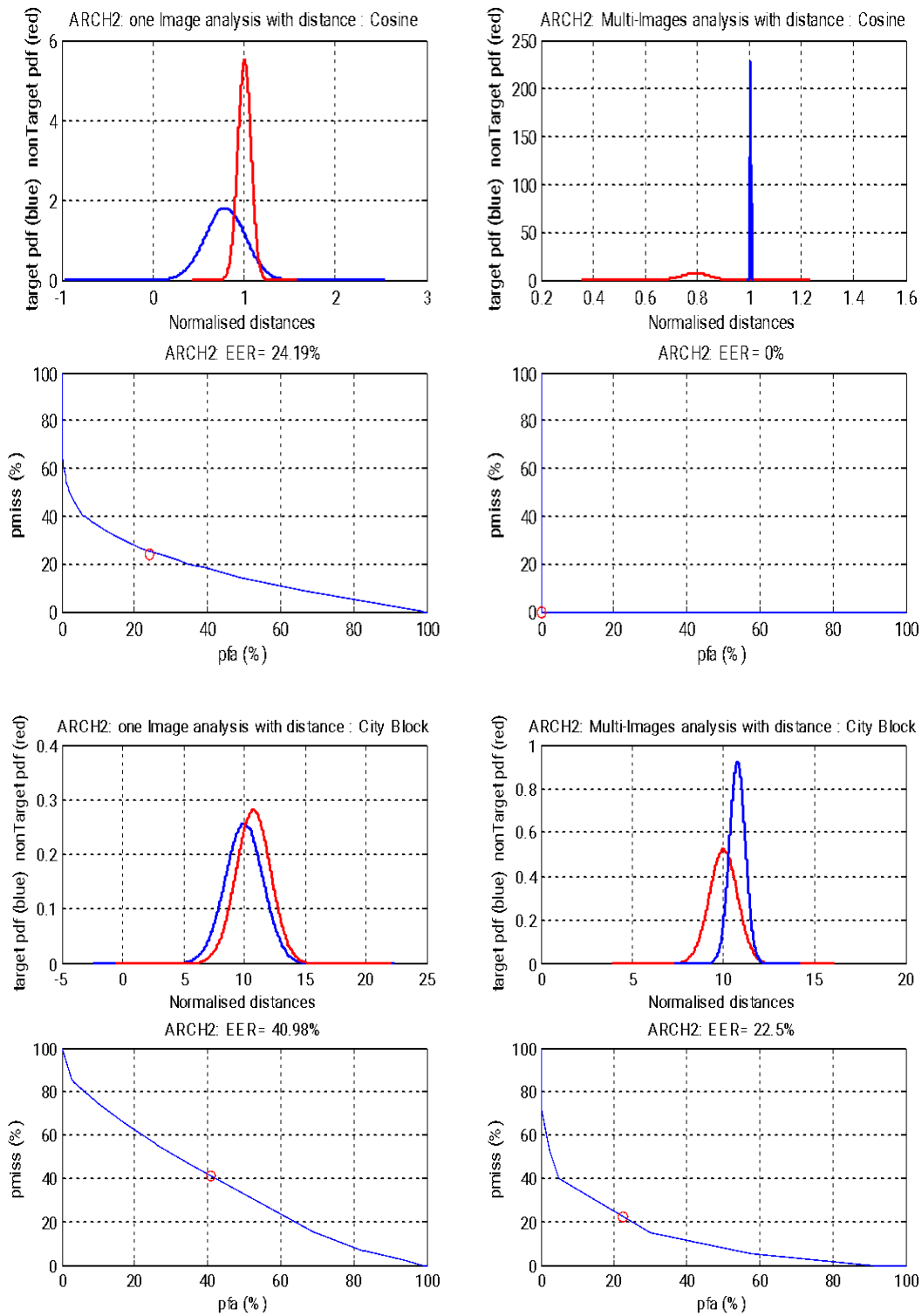
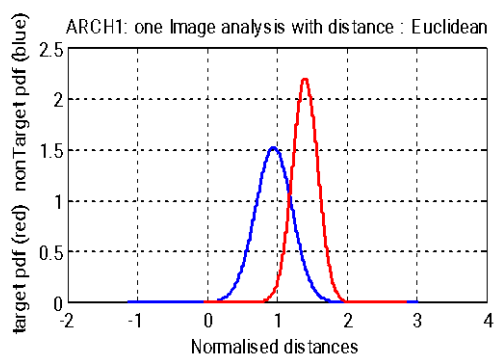
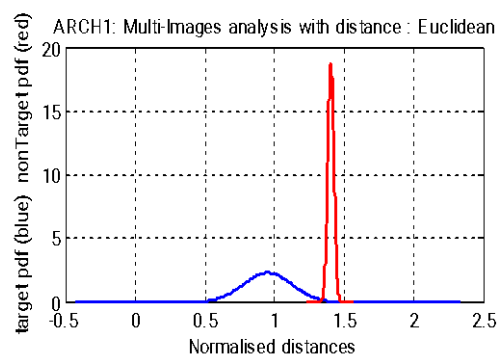
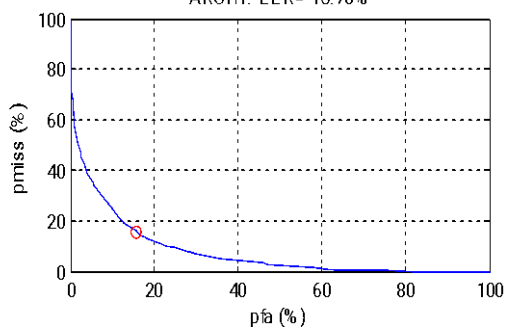


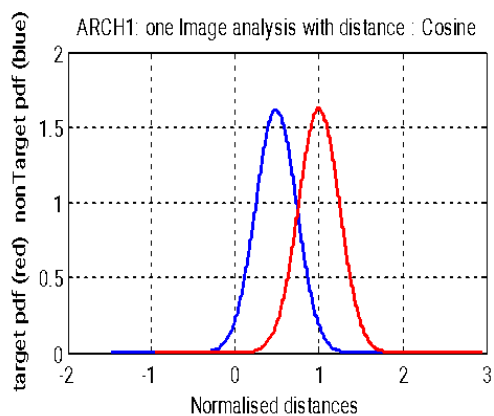
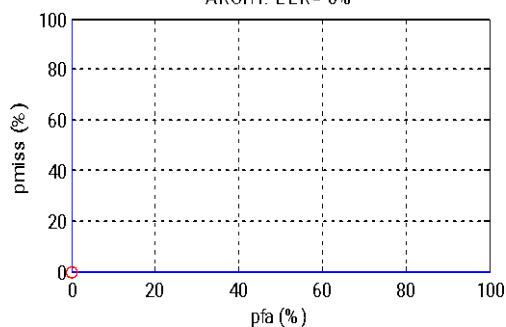
Figure 6. Verification Analysis Studies Based Olivetti Research Laboratories (ORL, 1994) Face databases for Independent Components Analysis Classifier Fusion of Likelihood Ratio



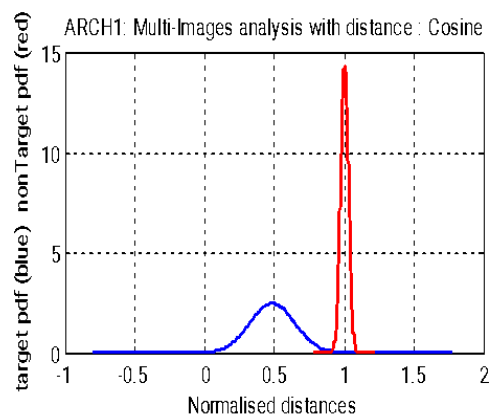
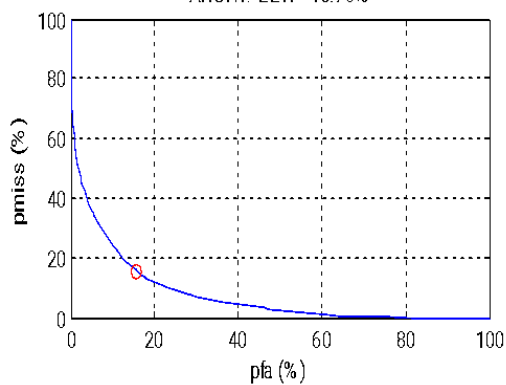
ARCH1: EER= 15.78%



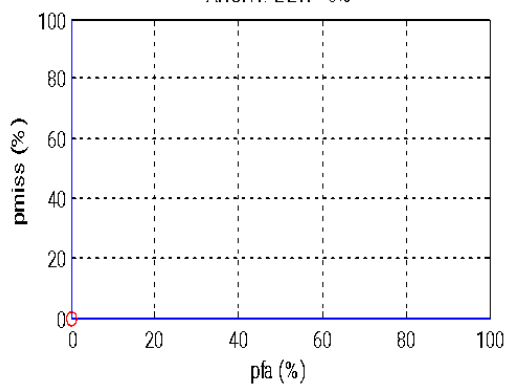
ARCH1: EER= 0%

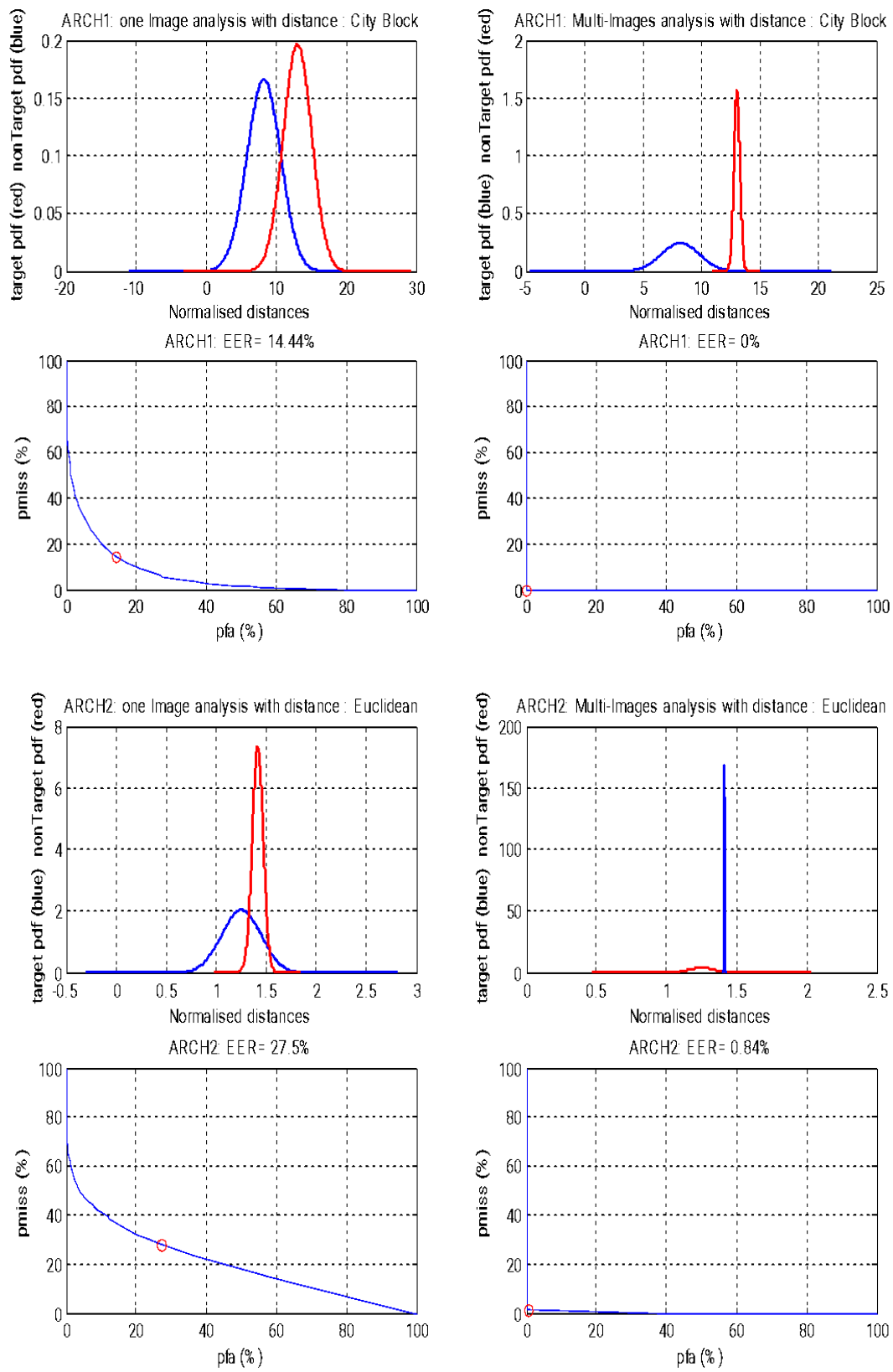


ARCH1: EER= 15.78%



ARCH1: EER= 0%





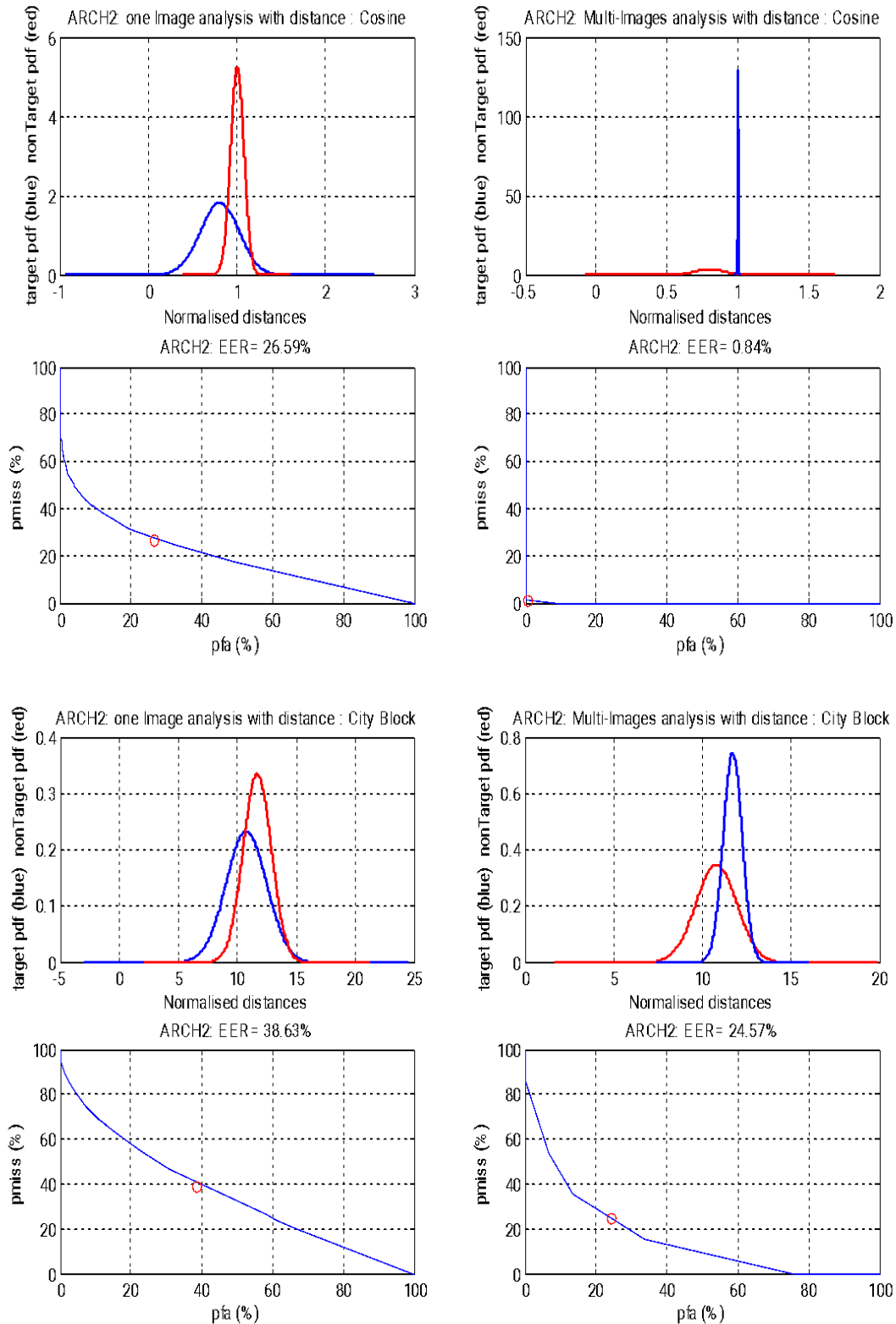
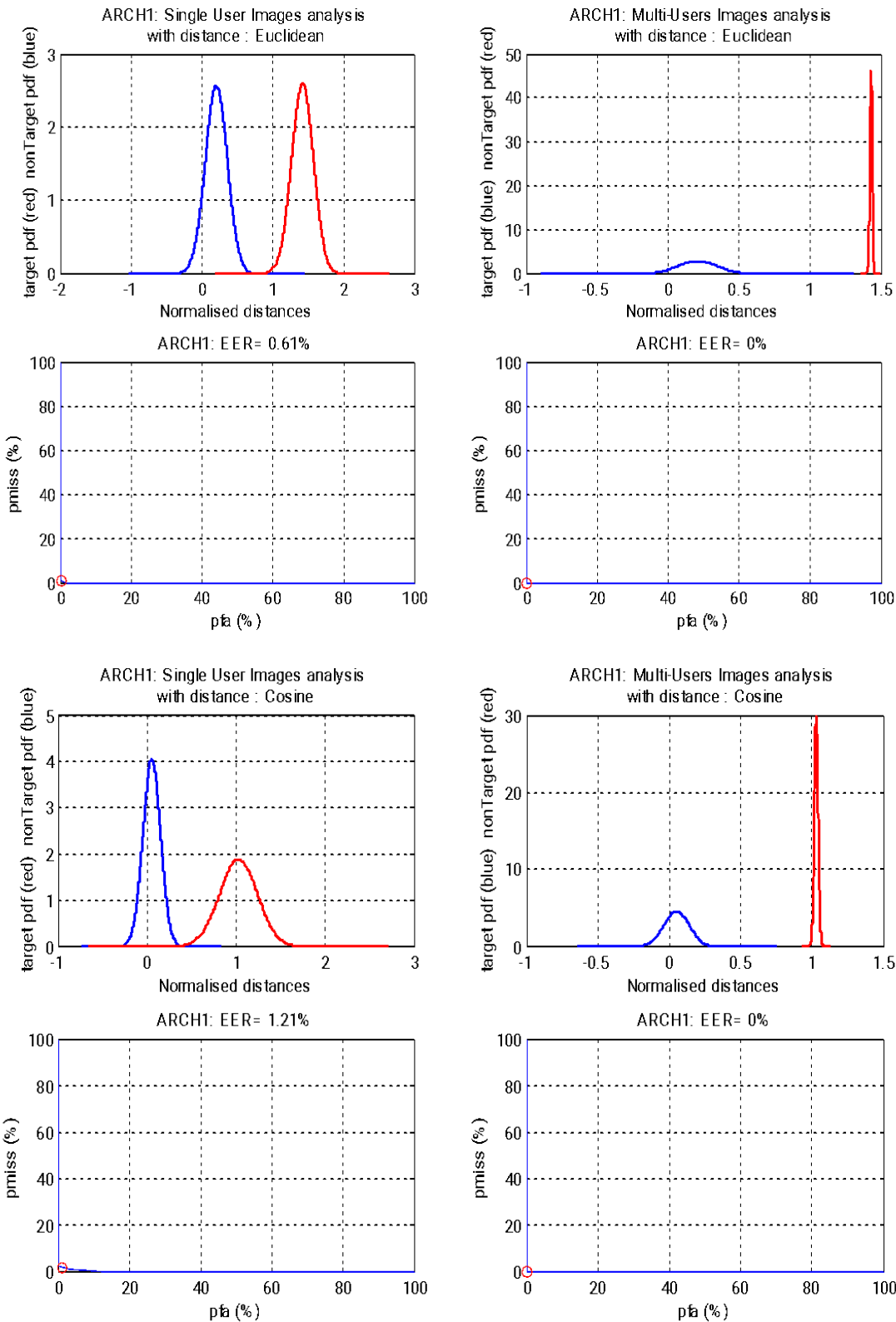
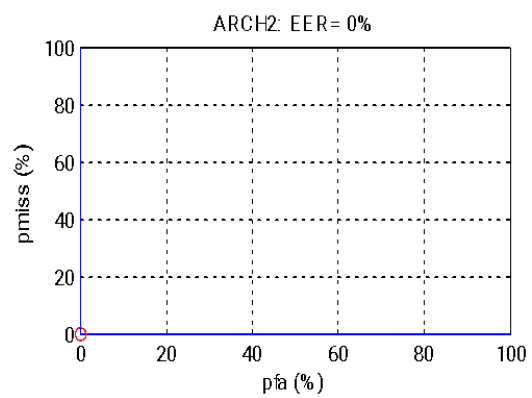
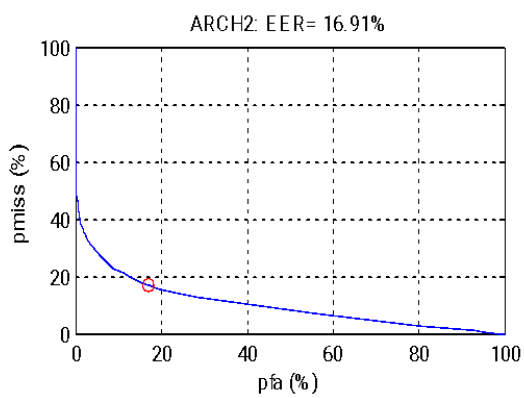
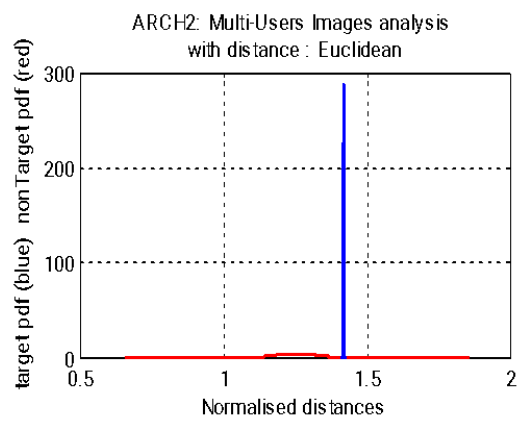
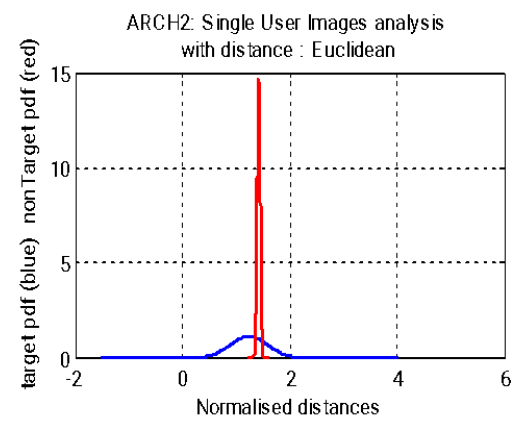
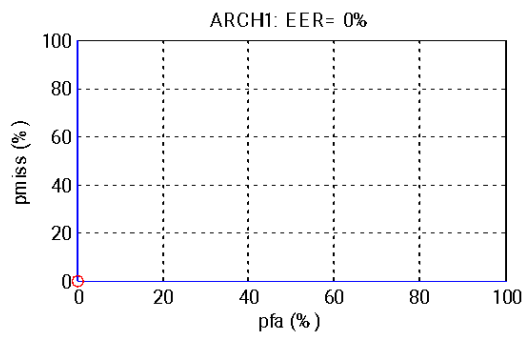
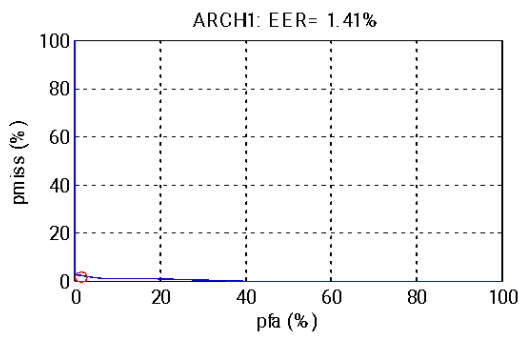
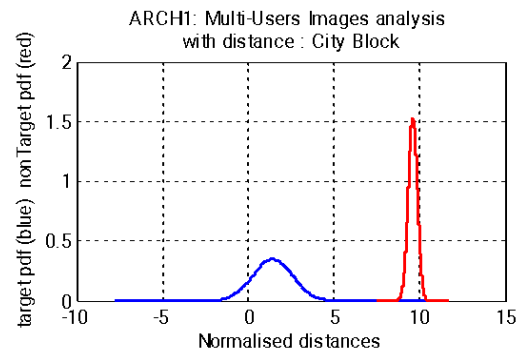
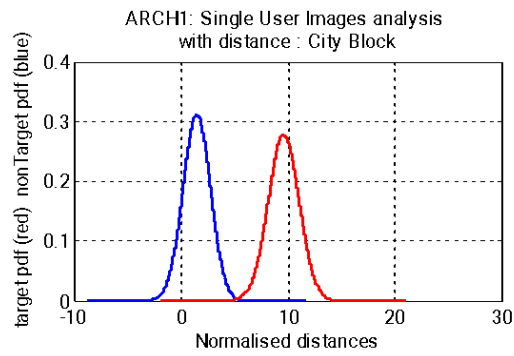


Figure 7. Verification Analysis Studies Based Indian Face databases for Independent Components Analysis Classifier Fusion of Likelihood Ratio





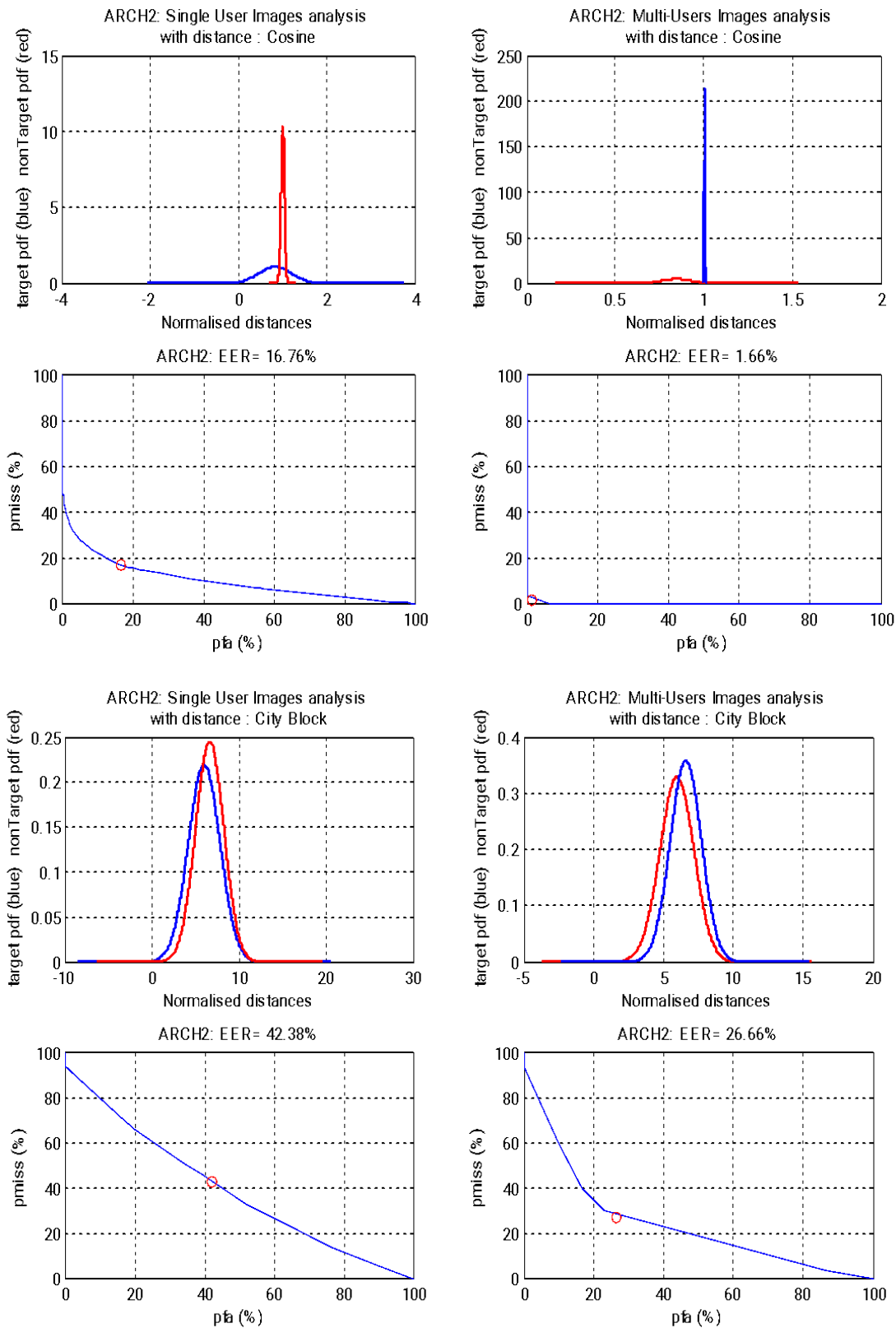


Figure 8. Verification Analysis Studies Based eINTERFACE2005 Dynamics Face databases for Independent Components Analysis Classifier Fusion of Likelihood Ratio

5. Conclusions

The question that always arises is how it can obtain the “best” (in terms of accuracy); the solution is fusion. It is well known that the “best” fusion solution is one that satisfies the Neyman-Pearson (NP) criterion (Soltane & Bakhti, 2012). The NP criterion characterises the fusion

solution that maximises the True Accept Rate (TAR) for a fixed value of False Accept Rate (FAR) (Van Trees & Bell, 1968), (Korves et al., 2005), (Soltane & Bakhti, 2012).

In the case of a verification system, two error rates are evaluated which vary in opposite directions: the false rejection rate FRR (rejection of a legitimate user called “the client”) and the false acceptance rate FAR (acceptance of an impostor) (Soltane & Bakhti, 2012). The decision to accept or reject an individual is thus taken by comparing the system’s output to a threshold (called the decision threshold). The values of FAR and FRR are thus dependent on this threshold, which can be selected to minimise the overall error of the system (Dorizzi et al., 2005; Soltane & Bakhti, 2012). The decision threshold must be adjusted according to the desired application characteristics. High security applications require a low FAR, which increases the FRR, while low-security applications are less demanding in terms of FAR. EER denotes Equal Error Rate and it’s the point where FAR equals FRR.

This threshold must be calculated afresh for each application, to adapt it to the specific population concerned (Dorizzi et al., 2005; Soltane & Bakhti, 2012). This is done in general using a small database recorded for this purpose.

Performance capabilities have been traditionally shown in the form of ROC (receiver- or relative-operating characteristic) plots (Soltane & Bakhti, 2012), in which the probability of a false-acceptance is plotted versus the probability of a false-rejection for varying decision thresholds. Unfortunately, with ROC plots, curves corresponding to well-performing systems tend to bunch together near the lower left corner, impeding a clear visualization of competitive systems (Soltane & Bakhti, 2012).

The paper has presented a human authentication method of Biometric face based Independent Component Analysis Architectures I and II. Single biometric authentication has the fundamental problems of high FAR and FRR. A framework for a novel fusion strategy at the score level has been proposed. The likelihood ratio-based fusion rule within the scores of the classifier achieves a significant recognition rate. The results demonstrated EER values of 0.61% and 0.00%, for Algorithmic Processing design One and Two, of algorithmic design based on ICA ARCH I fusion of Likelihood Ratio, within which scores are achieved based on Euclidean Measures Similarity. Thus, based on the experimental results, it has shown that EER can be reduced down significantly.

References

- Alharbi, B., & Alshanbari, H. S. (2023). Face-voice based multimodal biometric authentication system via FaceNet and GMM. *PeerJ Computer Science*, 9, e1468. <https://doi.org/10.7717/peerj-cs.1468>
- Bartlett, M. S., Lades, H. M., & Sejnowski, T. J. (1998). Independent component representations for face recognition. In *Proceedings of Human Vision and Electronic Imaging III*. SPIE.
- Batagelj, B., & Solina, F. (2006). Face recognition in different subspaces: A comparative study. In *Proceedings of the 6th International Workshop on Pattern Recognition in Information Systems (PRIS), ICEIS 2006* (Vol. 1, pp. 71–80). SCITEPRESS. <https://doi.org/10.5220/0002500200710080>
- Becker, B. C., & Ortiz, E. G. (2008). Evaluation of face recognition techniques for application to Facebook. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition* (pp. 1–6). IEEE. <https://doi.org/10.1109/AFGR.2008.4813471>
- Belhumeur, V. D., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition using class-specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 711–720. <https://doi.org/10.1109/34.598228>
- Bn-Yacoub, S., Abdljaoued, Y., & Mayoraz, E. (1999). *Fusion of face and speech data for person identity verification* (IDIAP Research Report RR 99-2003). IDIAP Research Institute.
- Comon, P. (1994). Independent component analysis: A new concept. *Signal Processing*, 36, 287–314. [https://doi.org/10.1016/0165-1684\(94\)90029-9](https://doi.org/10.1016/0165-1684(94)90029-9)

- Delac, K., Grgic, M., & Grgic, S. (2012). *A comparative study of PCA, ICA and LDA*. Croatian Telecom & University of Zagreb, Faculty of Electrical Engineering and Computing.
- Sofia, G., & Panagiotis, P. (2004). DNA smart card for financial transactions. *The ACM Student Magazine*.
- Dongliang, H., Henry, L., & Li, W. (2005). *Fusion of dependent and independent biometric information sources* (Technical report). Department of Electrical & Computer Engineering, University of Calgary.
- Dorizzi, B. (2005). *Biometrics at the frontiers: Assessing the impact on society — Technical impact of biometrics* (Background paper). Institute of Prospective Technological Studies, DG JRC – Sevilla, European Commission.
- Girija, C., & Michael, W. (2006). Audio-visual multimodal fusion for biometric person authentication and liveness verification. In *Proceedings of the NICTA-HCSNet Multimodal User Interaction Workshop (MMUI 2005)*. Australian Computer Society.
- Hyvärinen, A. (1997). Independent component analysis by minimization of mutual information. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)* (pp. 3917–3920).
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4–5), 411–430. [https://doi.org/10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5)
- Jain, V., & Mukherjee, A. (2002). *The Indian Face Database* [Dataset].
- Korves, H. J., Nadel, L. D., Ulery, B. T., & Bevilacqua Masi, D. M. (2005). *Multi-biometric fusion: From research to operations*. MTS MitreTek Systems, Sigma Summer, 39–48.
- Loève, M. (1955). *Probability theory*. Van Nostrand.
- Lu, W., & Rajapakse, J. C. (2001). *Constrained independent component analysis* [Technical report]. School of Computer Engineering, Nanyang Technological University.
- McCune, B., & Grace, J. B. (2002). *Distance measures* (Chapter 6). University of Massachusetts. <http://www.umass.edu/landeco/teaching/multivariate/readings/McCune.and.Grace.2002.chapter6.pdf>
- Norman, P., & Samy, B. (2004). *Database, protocol and tools for evaluating score-level fusion algorithms in biometric authentication* (IDIAP Research Report RR 04-44). IDIAP Research Institute.
- Olivetti Research Laboratory. (1994). *ORL face database* [Dataset]. <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- Pham, M. T. (2008). *On estimating the parameters of Gaussian mixtures using EM* (Doctoral dissertation). In *Principled asymmetric boosting approaches to rapid training and classification in face detection*. School of Computer Engineering, Nanyang Technological University, Singapore. http://www.sharkdolphin.com/pmtri/EM/phamminhtri_EM.pdf
- Reddy, S. S. S., Bommala, H., Sakthidharan, G. R., & Vatin, N. I. (2024). Multi-modal biometric recognition for face and iris using gradient neural network (Gen-NN). In *MATEC Web of Conferences: ICMED 2024* (Vol. 392, Article 01078). <https://doi.org/10.1051/mateconf/202439201078>
- Samaria, F. S., & Harter, A. C. (1994, December). Parameterisation of a stochastic model for human face identification. In *Proceedings of the 1994 IEEE Workshop on Applications of Computer Vision* (pp. 138–142). IEEE. <https://doi.org/10.1109/ACV.1994.341300>
- Shakhnarovich, G., & Moghaddam, B. (2004). *Face recognition in subspaces* (Technical Report No. TR2004-041). Mitsubishi Electric Research Laboratories.
- Soltane, M., & Bakhti, M. (2012). Multi-modal biometric authentications: Concept issues and applications strategies. *International Journal of Advanced Science and Technology*, 48.
- Somashekhar, B. M., & Nijagunarya, Y. S. (2018). Face and fingerprint fusion system for identity authentication using fusion classifiers. *International Journal of Computer Science & Engineering Survey*, 9(1), 2.

- Stylianou, Y., Pantazis, Y., Calderero, F., Larroy, P., Severin, F., Schimke, S., ... Valsamakis, A. (2005). GMM-based multimodal biometric verification. In *eNTERFACE 2005: The Summer Workshop on Multimodal Interfaces* (July 18–August 12). Faculté Polytechnique de Mons, Belgium.
- Sun, Q., & Redei, A. (2022). Knock knock, who's there: Facial recognition using CNN-based classifiers. *International Journal of Advanced Computer Science and Applications*, 13(1), 9–16.
- Thamaraimanalan, T., Pranavakumar, S., Lingeswaran, R. A., Kabileshar, R. M., & Sarankarthick, M. (2021). *Multi biometric authentication using SVM and ANN classifiers*. Irish Interdisciplinary Journal of Science & Research. SSRN. <https://ssrn.com/abstract=3814661>
- TIBCO Software Inc. (2013). *Distance measures overview*. https://docs.tibco.com/pub/spotfire/5.5.0-march-2013/UsersGuide/hc/hc_distance_measures_overview.htm
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86. <https://doi.org/10.1162/jocn.1991.3.1.71>
- Van Trees, H. L., & Bell, H. L. (1968). *Detection, estimation, and modulation theory* (Part I). John Wiley & Sons.
- Vaseghi, S., & Jetelová, H. (2009). *Principal and independent component analysis in image processing* [Technical report]. Brunel University, Department of Electronics and Computer Engineering; Institute of Chemical Technology, Department of Computing and Control Engineering.
- Veeramachaneni, K., Osadciw, L. A., & Varshney, P. K. (2005). An adaptive multimodal biometric management algorithm. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, 35(3), 344–356. <https://doi.org/10.1109/TSMCC.2005.848191>
- Zana, Y., César-Júnior, R. M., Feris, R. S., & Turk, M. (2005). Face verification in polar frequency domain: A biologically motivated approach. In G. Bebis, R. Boyle, D. Koracin, & B. Parvin (Eds.), *Lecture Notes in Computer Science: Vol. 3804. Advances in Visual Computing (ISVC 2005)* (pp. 183–190). Springer. https://doi.org/10.1007/11595755_23