

Extended Feature-Fusion Guidelines to Improve Image-Based Multi-Modal Biometrics

Dane Brown^{*}
Council for Scientific and Industrial Research
Information Security
Pretoria, South Africa
dbrown@csir.co.za

Karen Bradshaw
Rhodes University
Computer Science Dept.
Grahamstown, South Africa
k.bradshaw@ru.ac.za

ABSTRACT

The feature-level, unlike the match score-level, lacks multi-modal fusion guidelines. This work demonstrates a practical approach for improved image-based biometric feature-fusion. The approach extracts and combines the face, fingerprint and palmprint at the feature-level for improved human identification accuracy. Feature-fusion guidelines, proposed in recent work, are extended by adding the palmprint modality and the support vector machine classifier. Guidelines take the form of strengths and weaknesses as observed in the applied feature processing modules during preliminary experiments. The guidelines are used to implement an effective biometric fusion system at the feature-level to reduce the equal error rate on the SDUMLA and IITD datasets, using a novel feature-fusion methodology.

CCS Concepts

•Security and privacy → Biometrics; •Computing methodologies → Image processing; *Supervised learning by classification*;

Keywords

multi-modal biometrics; feature-level fusion; guidelines; face; fingerprint; palmprint

1. INTRODUCTION

Biometric systems often have to contend with degraded quality of the data being modelled as well as inconsistencies in data acquisition. Additionally, their widespread use has increased the risk of identity theft by forgers [14]. In parallel to these developments, fusing multiple sources of biometric information has been shown to improve the recognition accuracy, security and robustness of a biometric system [11].

Initially, development of multi-modal biometrics focused on fusing at the matching score-level. Recently, biometric

fusion at the feature level has been studied and shown to outperform the matching score-level [20]. Feature-level fusion integrates feature sets corresponding to two or more biometric modalities. The widely used matching score-level fusion does not utilize the rich discriminatory information available at the feature-level. Comprehensive studies have been conducted at the matching score-level across most biometric modalities. These studies often compare the best fusion methods, forming fusion guidelines that can be used in future applications of a similar nature [19, 21]. An example is the use of min-max normalization to achieve a low false acceptance rate (FAR) and Bayes-based normalization to achieve a low false rejection rate (FRR), in general. Additionally, the weighted sum rule is preferred for general fusion and the product rule is preferred for very high quality input data fusion. Feature-level fusion, on the other hand, is not only more complex, but also lacks guidelines, as it is a lesser studied problem. The general guideline for feature-level fusion is the use of uncorrelated feature sets among different modalities before fusion [24]. The opposite holds true when fusing multiple samples of the same modality.

It is unclear from the literature whether fusing the face, fingerprint and palmprint at the feature-level before or after feature transformation is an important factor. However, it is often performed after transforming the feature space using linear or non-linear methods [9, 20, 25]. In either case, transforming features to lower the dimensionality is an important factor in biometric feature-fusion. Additional factors include the size, quality and number of training samples used for effective data modelling of the dataset. Hence, determining an appropriate feature selection and transformation scheme is key when combining information at the feature-level [14, 19]. Raghavendra *et al.* also state that applying very similar feature selection and transformation to different modalities can yield a very efficient multi-modal biometric system [19]. Biometric modalities are thus independent yet complementary.

This paper combines the face, fingerprint and palmprint biometrics at the feature-level and investigates the above factors. The resulting combinations of feature processing modules are expected to produce an improved recognition performance compared with the individual modalities. Moreover, the interactions of the resulting combinations are expected to serve as guidelines for face, fingerprint and palmprint feature-fusion and other image-based multi-modal systems. Therefore, these guidelines are first determined based on comparisons with related studies using multiple datasets. There are limited studies that fuse the face, fingerprint and

^{*}Rhodes University student, based at the CSIR.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SAICSIT '16, September 26 - 28, 2016, Johannesburg, South Africa

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-4805-8/16/09...\$15.00

DOI: <http://dx.doi.org/10.1145/2987491.2987512>

palmprint at the feature-level for human identification. The proposed fusion methodology makes use of the guidelines to construct accurate multi-modal biometric human identification systems while correlating the performance to the number of training samples required.

The rest of the paper is organized as follows: Section 2 presents the related studies found in the literature. Section 3 discusses the different modules used at all image processing stages in this paper. Section 4 discusses the implemented face, fingerprint and palmprint fusion methodology. The experimental analysis and results are discussed in Section 5. In Section 6, conclusions are given and ongoing research is outlined.

2. RELATED STUDIES

Rattani *et al.* produced a multi-modal biometric system by fusing face, fingerprint and palmprint images [20]. They showed that good dimension reduction and normalization algorithms allow feature-level fusion to produce a lower equal error rate (EER) than matching score-level fusion. Their feature-level fusion approach emphasizes the extraction of image regions of both modalities.

The study concluded that fusing information from uncorrelated traits such as the face, fingerprint and palmprint at the feature-level increases human identification accuracy. A significantly better FAR and FRR (and ERR) were achieved at the feature-level compared to the matching score-level. Other face, fingerprint and palmprint studies include [13] and [23], which both use a Curvet transform followed by SVM classification. However, these studies were tested using datasets that are not publicly available.

Yao *et al.* proposed an accurate multi-modal biometric system that fused the face and palmprint [25]. Four Eigen-based face and palmprint feature-fusion algorithms were compared. The proposed method filters the EigenFaces and EigenPalms with a Gabor filter followed by the weighted concatenation of the resulting feature vectors.

The AR Face dataset was organized into a face dataset, with a resolution of 60×60 , and limited to the first 20 images per 119 individuals. Images in the palmprint dataset, provided by Hong Kong Polytechnic University, were matched to those used in the face dataset with the same number of samples per individuals. The proposed fusion method obtained its highest accuracy of 95% with six training samples, but reached 91% with only one training sample. Ahmad *et al.* proposed a similar system that was tested on ORL Face and Poly-U palmprint datasets [1]. The resulting accuracy was a 99.5% genuine acceptance rate (GAR), with an unspecified FAR. The near-perfect accuracy is attributed to the use of the relatively small-scale ORL Face dataset, limiting the fused dataset to 40 individuals and 10 samples per individual.

Comparing face, fingerprint and palmprint feature-fusion systems under identical conditions is a non-trivial task for the following reasons:

1. Availability of multi-modal biometric datasets is relatively low.
2. Private uni-modal datasets are used extensively in the literature.
3. Some feature-fusion systems perform human verification only.

4. The lack of EER, FAR and FRR metrics in most related studies limits a direct comparison to only system accuracy.

In Section 4.1, the preliminary experiments, we reimplemented Ahmad *et al.*'s algorithm and compared it with various combinations of feature processing modules. These combinations consist of feature selection, transformation and classification algorithms that constitute our proposed systems.

3. FEATURE PROCESSING MODULES

This section discusses the relevance of various feature processing modules used in the preliminary and proposed multi-modal biometric systems.

3.1 Feature Selection

Feature selection is used to choose or find appropriate features to allow a certain objective function to be optimized. The objective function typically aims to reduce the feature space by removing unwanted features, while keeping features that are highly representative of the underlying image class [4]. The entire image can be considered and contained within the results. However, based on the biometric, specific regions contain less noise and more discriminatory information. Considering the entire image during fusion, rather than core regions, can lead to the phenomenon known as "the curse of dimensionality problem" [19]. These regions are typically centred around the core for alignment during image registration.

In the case of fingerprints the texture pattern is known to contain richer information than singular points and minutiae [16]. The texture patterns consist of ridges and valleys on the surface of the finger and are known as global features. The local features are points on the fingerprint known as minutiae and singular points. Global features can be extracted for use in biometric fusion at the feature-level, while local features can be used to help align the global features. These features can also be extracted from palmprints as they share many characteristics of the fingerprint.

Facial texture patterns consist of global contour and pore features. Local features, known as facial landmarks, consist of the eyes, nose and mouth. These landmarks can be used to align global features in a similar way to fingerprints and palmprints.

Texture patterns can be classified to evaluate the recognition performance of biometrics. Numerous texture classification methods can be used, such as Eigen, Fisher and Local Binary Pattern Histograms (LBPH). Performing a comparison on these texture classification methods can form the basis for finding optimal feature selection and transformation schemes to classify different biometric modalities.

Global and local features often require algorithms to improve their clarity and consistency over multiple samples. This is particularly the case with contours and pores in face images, principal lines in palmprint images, and ridges and valleys in fingerprint images [18]. The biometric recognition process is often initiated by aligning the input image before other feature selection techniques are performed for improved feature discrimination [4, 10, 17, 27].

3.2 Image Registration

Image registration is an important first step in biometric recognition. At the local level, there are unique points within biometric data depending on the considered biometric modality. These points are determined for automated image registration based on the modality.

The fingerprint contains such points, located on ridge curvatures, that are either unique or sharper than those in other areas. The core point, also known as the reference point, is often defined as the sharpest concave ridge curvature based on the orientation field of the fingerprint image [12]. This point is especially useful as it serves as a guide during image registration, which is important for normalizing features. The core point is used to allocate a region of interest (ROI), which minimizes the discrepancy of stretch and alignment differences among multiple fingerprint samples of the same finger. Many previous approaches to reference point determination critically rely on the local features such as Poincaré index or other properties of the orientation field [12]. Poincaré works well in good quality fingerprint images, but fails to correctly localize reference points in poor quality fingerprints such as partials or fingerprints with poor ridge and valley contrast. The solution used in this paper applies an edge preserving non-local means (NL-means) filter [6] before applying Poincaré.

A ROI can also be determined for facial images. The key points are the eyes, nose and mouth, which are used during facial image registration. These points are used to create a border around the face, centred at the nose, which helps to avoid typical changes to the face such as different hair, occluded ears and neck. A large number of Haar-like features are organized to create a classifier cascade. Haar cascading is a popular method for detecting features that are used as key points. Multiple Haar cascades are iterated when the selected one fails to detect a key point.

The ROI for palmprint images is determined by extracting the principal lines and applying the iterative closest point method, which estimates the translation and rotation parameters between the input and test image by minimizing the distance between the two sets of correspondence points [15]. This method ensures that the key points between the index and little fingers, used as a boundary for the ROI, correspond between the two images.

The face, fingerprint and palmprint image registration methods each use a fall-back mechanism in case their respective key point detectors fail. This is determined based on the confidence score of the LBPH texture classifier.

3.3 Feature Normalization and Discrimination

The following three feature selection algorithms are particularly useful at reducing the differences among multiple same-class data or improving the discrimination among different classes of data.

3.3.1 Pixel Normalization

The pixel values of an image or certain image regions are set to a constant mean and variance to compensate for slight inconsistencies in lighting and contrast. This is essential for normalizing image biometric data.

3.3.2 Histogram Equalization

Histogram equalization is an effective way of automatically setting the illumination across a dataset as a consistent

amount [4]. The greyscale range is distributed uniformly by applying a non-linear transformation with a slight side effect on the histogram shape. Histogram equalization is often more effective than pixel normalization, but should be avoided in most histogram-based matching methods. Both pixel normalization and histogram equalization improve the consistency among multiple data samples.

3.3.3 Frequency Filtering

The quality of the input image plays an important role in the performance of the feature extraction and matching algorithm [8]. A bandpass filter can be used to increase the amplitude of the mean component of the image. This has the effect of increasing the dominant spectral components while attenuating the weak components.

The Gabor filter is often used to filter frequencies based on the texture of the face, fingerprint and palmprint. It is constructed using a special short-time Fourier transform by modulating a two-dimensional sine wave at a particular frequency and orientation with a Gaussian envelope. A Gabor filter requires much tuning for specific orientations and frequencies to isolate the undesired noise while preserving the structure of a particular biometric [7].

The Laplacian of Gaussian (LOG) filter can remove unwanted features on the high frequency spectrum before enhancing the remaining features, effectively increasing the DC component. However, a side effect can occur when applying the filter to badly aligned images. This side effect is also prevalent in images with inconsistent lighting. The increased feature discrimination of LOG further highlights the difference among multiple samples, leading to bad training and testing sets and consequently a lower recognition accuracy. The image registration procedure, in Subsection 3.2, as well as lighting normalization is thus imperative to the success of an image-based feature-fusion biometric system.

3.4 Feature Transformation for Reduction and Classification

Image classification algorithms aim to express the most relevant image properties. Feature transformation is used to express a feature vector in an alternate space to improve discrimination. This often allows for a reduction in dimensionality and intra-class variation – multiple samples contain near-identical information. The following image texture classification methods are considered.

3.4.1 Eigen

Statistical classifiers can be used to maximize inter-class variation to discriminate effectively between different individuals.

The Eigen classifier uses Principal Component Analysis (PCA) to represent statistically key features that define a given feature set. An efficient model can be constructed from principal components, retaining key features of samples in one class. The distances among Eigenvalues are compared between the trained model and the model to be tested during matching.

Given N sample images x , the total scatter matrix is defined as [3]:

$$S_t = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (1)$$

where $\mu \in \mathbb{R}^n$ is the mean image obtained from the samples.

3.4.2 Fisher

PCA is a powerful technique to maximize total scatter of data and lower the dimensionality of a dataset, but it does not effectively consider the rest of the classes within a particular class. This can lead to a loss of some discriminative information, on an inter-class level. On the other hand, Linear Discriminant Analysis (LDA) performs extra class-specific dimensionality reduction and is referred to as the Fisher classifier. Fisher classification clusters same class data tightly and maximizes the separation of different classes in a lower-dimensional representation.

Given C classes the between-class scatter matrix is defined as [3]:

$$S_b = \sum_{i=1}^C N_k(\mu_k - \mu)(x_k - \mu)^T \quad (2)$$

and the within-class scatter matrix is defined as:

$$S_w = \sum_{i=1}^C \sum_{x_k \in \mathbb{X}_i} (x_k - \mu)(x_k - \mu)^T \quad (3)$$

The dimensionality is lower than the Eigen method as $C - 1$ is the maximum number of non-zero generalized Eigen-values.

The Fisher classifier learns a highly discriminative class-specific transformation matrix, but inconsistent data within classes, caused by, for example, varying illumination, has a greater negative impact on class separation performance compared with the Eigen method. The Fisher method generally requires more training data than the Eigen method in non-ideal conditions. The extra reduced dimensionality that LDA offers allows for the Fisher method to have a lower training and testing time compared with the Eigen method. This is especially prevalent when using larger datasets and a greater number of samples per class.

3.4.3 Support Vector Machines

The support vector machine (SVM) is a supervised statistical learning model that has been used extensively in pattern recognition problems [5]. Its advantage over many other classifiers is that the training time is unaffected by the high dimensionality of feature vectors from images. SVMs were originally used to solve binary class problems, but are easily extended to support multi-class problems. In a binary class problem, SVMs aim to maximize a mathematical function given a collection of data points that consist of two classes that can be separated by a decision boundary.

Consider that the two classes $S^+ = \{x_i | y_i = 1\}$ and $S^- = \{x_i | y_i = -1\}$ are linearly separable [5]. This results in at least one boundary that can be formed between them. The data points of sets S^+ and S^- that are located on the boundaries of the margin are known as the support vectors. A simple rescale of w for all x_i that are support vectors holds that:

$$w \cdot x_i + b = 1 \quad (4a)$$

$$w \cdot x_i + b = -1 \quad (4b)$$

The distance d between the decision boundary and the margin can be expressed as:

$$d = \frac{2}{\|w\|} \quad (5)$$

In higher-dimensional space, the decision boundary that achieves the maximum margin between sets S^+ and S^- is known as the optimal hyperplane, which allows an SVM to model new data points more accurately.

3.4.4 Spatially Enhanced Histogram Matching

Regular histogram matching is one of the simplest image matching methods. LBPH is another texture feature descriptor for images. A basic local binary pattern operator assigns a label to every pixel of an image by thresholding the 3×3 pixel neighbourhood, consisting of a centre pixel value that is compared to its neighbours, resulting in a binary representation of the neighbourhood. The comparison transforms the neighbour pixel to a 1 if it is larger than the centre pixel and a 0 if it is smaller than the centre pixel. The transformed neighbourhoods are uniform such that the histogram has a separate bin for every pattern neighbourhood. The result is a concatenated histogram consisting of all the neighbourhoods.

In this paper a special kind of LBP operator called extended LBP (ELBP) is used. Instead of being limited to directly adjacent neighbours, the neighbourhood is extended to include interpolated pixels, based on a circular mask, that capture fine texture. This operator uses spatially enhanced histogram matching that enables partial matching and automatic pixel normalization on a pixel level, circular neighbourhood level and image level. This results in the distinct advantage of illumination, scale and rotation invariant texture classification compared with Eigen and Fisher [2].

Training the advanced histogram model is also significantly faster than the former two methods. Furthermore, the training time is independent of the image resolution and it produces the smallest model size. Given m circular neighbourhoods, their corresponding spatially enhanced histograms are determined, with a size of $m \times n$, where n is the length of a single histogram.

4. FEATURE-FUSION GUIDELINES

This section discusses the development and use of guidelines for biometric feature-fusion by utilizing the feature processing modules discussed in the previous section.

4.1 Preliminary Experiments for Developing the Guidelines

Feature-fusion guidelines determined in this subsection are limited to the feature processing modules discussed in Section 3. The datasets that were used in the preliminary experiments are also discussed. The guidelines were determined by analyzing the results of the following preliminary experiments.

The SDUMLA fingerprint dataset consists of fingerprint images that vary in quality, from partials with absent core points to fingerprints with average quality ridges. The ORL Face dataset is of average quality with images captured from various angles. PolyU palmprints consists of very high quality images, captured in a controlled environment with hand restricting user-pegs.

Fingerprint and face pseudo multi-modal datasets were formed by pairing SDUMLA Fingerprint right index fingers [26] with ORL Face [22], limiting the individuals to 40. Palmprint and face and palmprint and fingerprint pseudo multi-modal datasets were similarly formed, by pairing PolyU palmprints with the above face and fingerprint datasets, re-

spectively. The interactions of the feature processing modules were investigated by conducting preliminary experiments on the organized pseudo multi-modal datasets, summarized in Table 1.

Ahmad *et al.*'s best method, as reimplemented in this paper, was compared with preliminary systems that exhausted the combinations of feature processing modules discussed in Section 3. Their system was extended to include a fingerprint version – aligned using the automated approach described in 3.2. Using five training samples, their system achieved an EER of 26.71% when fusing the face and fingerprint, 11.77% when fusing the palmprint and fingerprint, and 7.33% when fusing the palmprint and face, using one training sample. Their system achieved an EER of 6.5% when fusing the face and fingerprint, 1.77% when fusing the palmprint and fingerprint, and 0.77% when fusing the palmprint and face. Their reimplemented system was verified to be a true representation of the original system. Moreover, both the original and the reimplemented systems achieved a 99.5% system accuracy for palmprint and face fusion using the same dataset.

The next set of experiments were performed on our proposed systems. Preliminary test results showed LBPH to be the best performing classifier when no feature discrimination algorithms were used. LBPH achieved an EER of 19.64% when fusing the face and fingerprint, using one training sample. Moreover, it achieved an EER of 0% when fusing the face and fingerprint, using five training samples. This is attributed to its robustness to bad alignment and lighting even with a bad training set. On the other hand, Eigen and Fisher classifiers performed poorly with EERs higher than 31% for one training sample in both cases. Eigen and Fisher achieved a 7.5% EER and 6.67% EER, respectively using five training samples. However, by adding the feature selection algorithms discussed in Section 3.3, which significantly improved discrimination, the accuracy of Eigen and Fisher increased, although there was little improvement for LBPH. Upon individual observation, the LOG filter caused inconsistent improvements across test subjects.

The best performing preliminary system used histogram equalization followed by LOG and the Eigen or Fisher classifier. Furthermore, feature-fusion, by column concatenation of images, after feature selection and before feature transformation was found to be most effective. The resulting EERs, when using a single training sample, were 14.67% and 16.89% for Eigen and Fisher, respectively. Both Eigen and Fisher achieved 0% EERs when using five training samples. Furthermore, Eigen and Fisher always achieved 0% EER when fusing the palmprint with either the face or the fingerprint, regardless of the number of training samples. The results also confirmed Raghavendra *et al.*'s assertion – the same feature transformation should be applied to different modalities.

4.2 Methodology for Applying Guidelines

The following subsections detail the proposed multi-modal biometric solution, based on the guidelines determined in the previous subsection, and illustrated in Figure 1. The figure provides an overview specific to face and fingerprint fusion. A similar process would be followed in the case of other modality combinations. The popular PolyU palmprint dataset resulted in perfect accuracy across all preliminary systems that fused palmprints with the face, fingerprint or

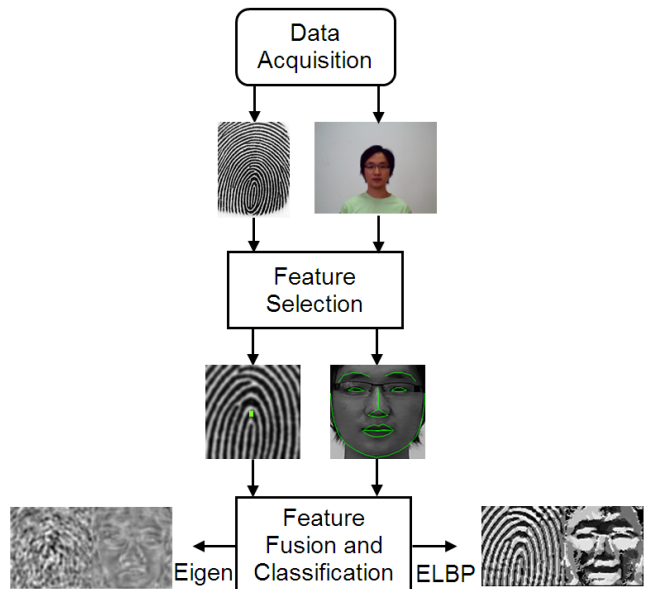


Figure 1: Proposed Methodology

both, and a near-perfect accuracy when using only the palmprint with histogram equalization followed by LOG and the Eigen or Fisher classifier. Therefore, a more challenging palmprint dataset was used in the final experiments.

4.2.1 Data Acquisition

The SDUMLA multi-modal dataset (face and fingerprint) [26] and the IITD palmprint dataset [?] were used in the final experiments discussed in Section 5. Eight samples per individual were selected from each dataset as follows:

1. Fingerprints – Eight samples of the left thumbprint were selected from the fingerprint images consisting of partials with absent core points, poorly-defined ridges and well-defined ridges.
2. Face – Eight samples of frontal faces were selected. The samples consisted of different poses and props, namely, normal, smile, frown, surprise, look down, shut eyes, hat and glasses.
3. Palmprint – Eight samples of the left palmprint was used. Since only six samples exist per individual's hand, two extra samples with added noise were generated by performing contrast reduction and a random four point affine transformation within 15 degrees and 5 pixel units on the last test sample for each individual. The challenging IITD palmprint was captured with a touchless sensor without user-pegs, thus resulting in wrinkle and scale variations.

4.2.2 Feature Selection

The face, fingerprint, and palmprint datasets were automatically cropped to 75×75 using multiple Haar cascades, Poincaré index with NL-means, and the iterative closest point, respectively. The Haar cascades were used to align the face based on the detected landmarks as outlined in Figure 1. LBPH served as both a fall-back mechanism and a final stage of alignment, for face, fingerprint and palmprint datasets.

Table 1: Proposed feature-fusion guidelines.

Stage	Name	Advantage	Disadvantage	Suggested Use
Feature Selection	Pixel Normalization	Reduces inconsistent lighting.	Not very effective for big changes in lighting.	All biometrics affected by lighting. The first step of quality enhancement.
	Histogram Equalization	Reduces inconsistent lighting.	Minor histogram distortion. Introduces some noise.	Biometrics that are affected by lighting. Histogram-shape invariant classifiers such as Eigen and Fisher.
	NL-means Filter	Denoises and preserves edges.	Can remove fine texture.	Use before Poincaré index and other algorithms that do not rely on fine texture.
	LOG Filter	Improved feature discrimination, before transforming to the Eigen space.	Requires consistent lighting.	Remove noise in the upper and lower frequencies before fusion.
	Gabor Filter	Improved feature discrimination, especially for fingerprints.	Requires tuning per application.	Adjust frequencies at a specific orientation and scale. Classifiers such as LBPH benefit significantly when using a well-tuned Gabor filter.
	Modified ELBP Operator	Minimizes inconsistent lighting.	Introduces noise.	Biometrics that are affected by lighting. Classifiers such as Eigen and Fisher.
Feature Transformation and Classification	Eigen Classifier	LOG and modified ELBP supplement this classifier.	Slow training. Relatively high dimensionality.	Small image regions. Useful for fused image-based biometrics.
	Fisher Classifier	LOG and modified ELBP supplement this classifier. Low dimensionality.	Requires a lot of good training data.	When Eigen dimensionality is too high.
	LBPH Classifier	Works well on faces. Very robust to pose angles.	Benefits the least from feature selection and feature-fusion.	General purpose classifier. Face images. Applications with low storage requirements.
	SVM Linear Classifier	LOG and modified ELBP supplement this classifier. Fast training and testing.	–	Single training sample applications.

Histogram equalization was applied to the Eigen and Fisher methods to reduce their shared weakness of dynamic lighting. Pixel normalization was instead applied to LBPH as histogram equalization caused negative effects on the spatial histogram. The ELBP operator was considered due to the inconsistent results achieved in preliminary experiments across test subjects when using the LOG filter. The ELBP operator was found to outperform histogram equalization and pixel normalization in terms of dynamic lighting in the preliminary experiments. However, this resulted in increased noise in the Eigen space, causing a reduction in accuracy compared with histogram equalization and pixel normalization. ELBP especially increased the noise when combined with LOG.

The ELBP operator was modified to improve Eigen and Fisher without increasing noise as follows: The typical parameters of ELBP used for LBPH texture classification – one pixel radius and eight neighbour pixels – were multiplied by four and averaged with the normalized original image to enable its use as a standalone feature selector. This reduced the variation in data of multiple samples across individuals and enabled its use with the LOG filter. This combination of the modified ELBP operator and LOG filter was chosen based on the guidelines, and thus two novel feature selectors are included in the final results in Section 5.

4.2.3 Feature-fusion and Classification

The enhanced face, fingerprint and palmprint feature vectors were combined using serial vector fusion. The fused vector is transformed to the Eigen feature space in the case of Eigen or Fisher. When the LBPH classifier is used, the feature vector is transformed to the spatial-histogram feature space. The classifiers divide the fused dataset into classes as explained in Section 3.4. The Eigen representation is reconstructed using only five principal components in Figure 1, for visualization. The ELBP image is also shown in the figure instead of the spatial histogram, for visualization. Both of these visualizations are produced using the baseline system.

5. RESULTS

The final experiments using the proposed fusion methodology are presented in this section. Figures 2, 3 and 4 illustrate the best performing bi-modal fusion combination (the fingerprint and face in all three cases). The following suffixes identify the applied feature selection and transformation schemes that were combined with the Eigen, Fisher or LBPH classifier: Histogram equalization is the baseline system, referred to as Eh; LOG is referred to as L; Modified ELBP is referred to as LBP; LOG followed by modified ELBP is referred to as LLBP; and PCA reduction is referred to as PCA.

5.1 Experiment 1

Referring to Figure 2, the baseline fusion always outperforms the face, fingerprint and palmprint. LBPL is the best performer for one training sample with an EER of 9.16% while LLBP yields an EER of 0.31% with five training samples.

Referring to Figure 3, the Fisher classifier performs similarly to Eigen, but with a noticeably weaker performance when using two training samples.

Referring to Figure 4, the baseline fusion produces a lower accuracy than the face when using three training samples.

Moreover, LBPH is a good face texture classifier. LBPH does not respond well to any of the feature processing modules described in Section 3. However, reducing the principal components improved the EER by 3% on average.

Table 2 focuses on the best accuracies achieved by the various Eigen, Fisher and LBPH methods, using five training samples. LBPHPCA produces the best face EER, 4.77% lower than EigenLLBP and FisherLLBP face methods, which share an EER of 8.49%. The fingerprint accuracies are similar, but EigenLBPL is slightly better with an EER of 5.35%. The results show that LBPH does not benefit from fusion as greatly as Eigen and Fisher. The fused EigenLLBP method achieves a 1.58% lower EER than FisherLBPL. The LBPL and LLBP methods for Eigen and Fisher produced very similar accuracies in all cases, except for EigenLLBP, which achieved a 0.95% lower EER than EigenLBPL.

5.2 Experiment 2

Referring to Figure 5, various fusion combinations that use EigenLLBP and SVMLLBP methods are compared. The face and fingerprint fused systems outperformed the other combinations, on average, in the case of both Eigen and SVM with a lowest EER of 0.31% and 1.26%, respectively. This is significant when compared with the face and palm fused systems, which achieved the lowest accuracies. The SVM achieves the lowest EER for one training sample regardless of modality, but does not significantly improve when using more than three training samples.

Fusing all modalities resulted in a 0.27% and 0.13% EER for EigenLLBP and SVMLLBP, when using a single training sample, respectively. Using two or more training samples resulted in a perfect accuracy for both methods. None of the other methods achieved a perfect accuracy except for the standard linear SVM – when using three or more samples. The standard linear SVM achieved a similar improvement in accuracy as Eigen did when applying the LLBP algorithm.

5.3 Summary of Results

Both the preliminary and final results confirm that the modified ELBP operator was successfully used together with the LOG filter to significantly improve feature discrimination and verify the hypothesis of requiring guidelines, presented in this paper. The LBPH method did not improve significantly from fusion or the feature processing modules described in Section 3. However, LBPH proved to be a robust classifier and was thus used in the proposed systems for improved image registration and as a fall-back mechanism, preventing system failure. On the other hand, the Eigen classifier showed the most significant improvement particularly when using more than three training samples. This is due to the benefit of increased feature discrimination on a classifier that relies on maximizing total scatter. Therefore, this benefits Fisher to a lesser extent. Eigen achieves the highest recognition accuracy of all the fusion schemes when using five samples, at a near-zero EER for the best bi-modal results and a zero EER for the tri-modal results. SVMLLBP achieved the best EER, on average, when using one to three training samples. Furthermore, the SVM classifier performed the best using one training sample. This is of particular benefit to applications that require low training time and data.

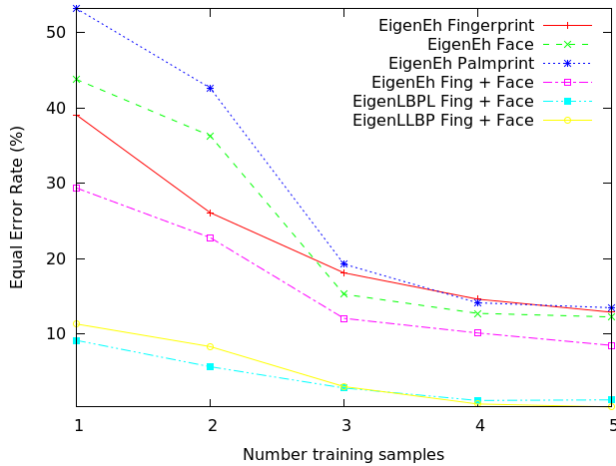


Figure 2: Comparison of the Eigen systems with the lowest EER.

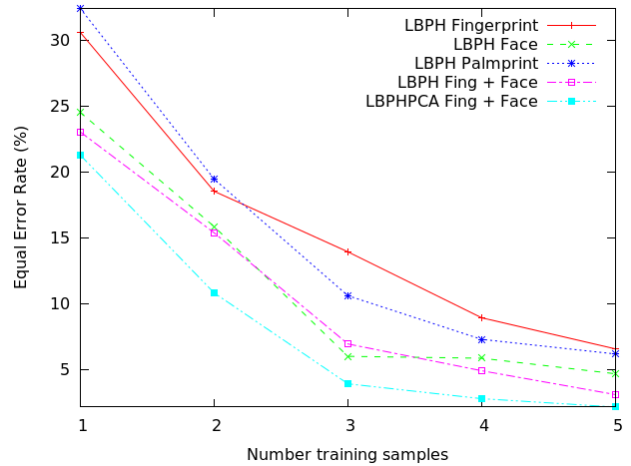


Figure 4: Comparison of the LBPH systems with the lowest EER.

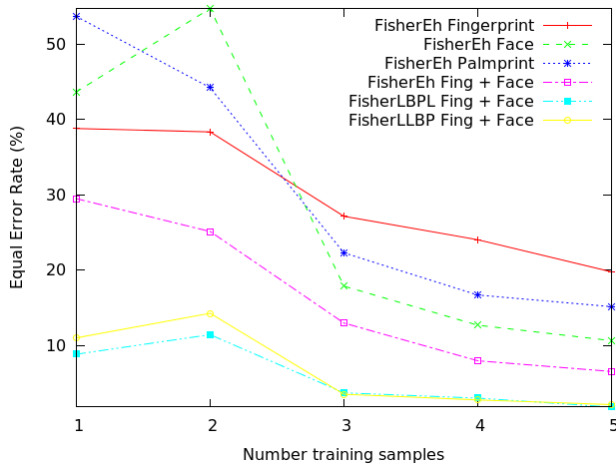


Figure 3: Comparison of the Fisher systems with the lowest EER.

Table 2: Best feature selector for the face, fingerprint, palmprint and fused dataset (in % EER).

	Face (%)	Fingerprint (%)	Palmprint (%)	Fing + Face (%)
Eigen	8.49	5.35	6.49	0.31
Fisher	8.49	6.60	7.60	1.89
LBPH	4.72	6.60	5.71	2.20

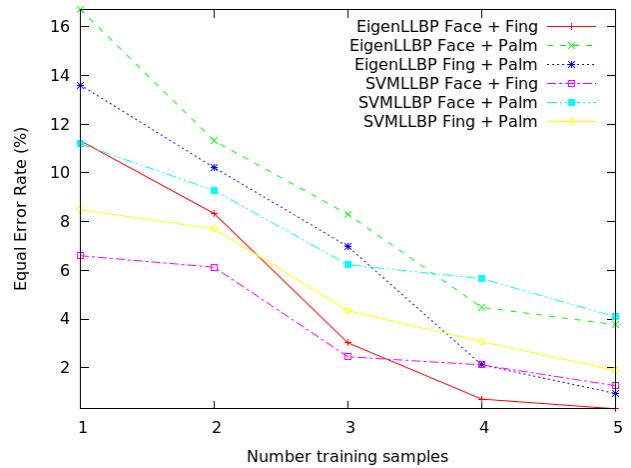


Figure 5: Comparison of Eigen and SVM fused systems when using LLBP.

6. CONCLUSION AND ONGOING RESEARCH

A comparison was performed on the face, fingerprint, palmprint and their fused combinations. The comparison was extended to include two combinations of the modified ELBP operator with the LOG filter, which is applied to the Eigen, Fisher and SVM classifier. This extended comparison is an effort to improve the recognition performance of feature-fusion achieved in our previous work and suggest more guidelines for future systems. The contribution of extended feature-fusion guidelines, in particular, is promising and encourages further improvements to the field.

The LBPH classifier achieved the best accuracy before feature selection and showed promise regarding its robustness to misalignment, dynamic lighting and scaling. Eigen in particular, showed substantial improvement when applying the two novel feature selection combinations before feature transformation. More fusion combinations were compared in the case of the Eigen and SVM classifiers. The face and fingerprint fused systems achieved the best accuracies in terms of modality combinations. At the feature-level, well-known

techniques often form the basis of these systems without reasoning. Moreover, multi-modal biometric studies often propose recognition systems that integrate many feature processing modules, such as quality enhancement and image registration during pre-processing, as well as feature selection, feature transformation and post-processing techniques. Thus, isolating the contributions that measure progress in the state-of-the-art is a time consuming non-trivial problem. However, this research serves as a foundation for selecting appropriate features for image-based multi-modal fusion as it contributes more guidelines to the research area.

Specifying a complete set of formal guidelines for multi-modal biometric fusion at the feature-level requires covering a substantial series of experiments. Therefore, a comprehen-

sive review of the important factors covered in this paper is being further investigated with additional experimentation on more image-based biometric modalities, datasets and feature processing modules. This ongoing research will form the basis for implementing optimal feature-fusion schemes, given a particular application in future.

7. ACKNOWLEDGMENTS

The authors would like to thank the CSIR, Information Security department for their financial support. Thank you to the authors of the publicly available datasets used in this paper.

8. REFERENCES

- [1] M. I. Ahmad, W. L. Woo, and S. S. Dlay. Multimodal biometric fusion at feature level: Face and palmprint. In *2010 7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP)*, pages 801–805. IEEE, 2010.
- [2] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006.
- [3] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.
- [4] A. C. Bovik. *Handbook of image and video processing*. Academic Press, 2010.
- [5] D. Brown. *Faster upper body pose recognition and estimation using Compute Unified Device Architecture*. PhD thesis, University of Western Cape, 2013.
- [6] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.*, volume 2, pages 60–65. IEEE, 2005.
- [7] G. S. Budhi, R. Adipranata, and F. J. Hartono. The use of Gabor filter and back-propagation neural network for the automobile types recognition. In *2nd International Conference SIIT 2010*, 2010.
- [8] S. Chikkerur, A. N. Cartwright, and V. Govindaraju. Fingerprint enhancement using STFT analysis. *Pattern Recognition*, 40(1):198–211, 2007.
- [9] A. Deshmukh, S. Pawar, and M. Joshi. Feature level fusion of face and fingerprint modalities using Gabor filter bank. In *2013 IEEE International Conference on Signal Processing, Computing and Control (ISPCC)*, pages 1–5. IEEE, 2013.
- [10] J. Feng and A. Jain. Fingerprint reconstruction: From minutiae to phase. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(2):209–223, Feb 2011.
- [11] O. N. Iloanusi. Fusion of finger types for fingerprint indexing using minutiae quadruplets. *Pattern Recognition Letters*, 38(0):8 – 14, 2014.
- [12] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *IEEE Transactions on Image Processing*, 9(5):846–859, 2000.
- [13] M. V. Karki and S. S. Selvi. Multimodal biometrics at feature level fusion using texture features. *International Journal of Biometrics and Bioinformatics*, 7(1):58–73, 2013.
- [14] D. Kaur and G. Kaur. Level of fusion in multimodal biometrics: A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(2):242–246, 2013.
- [15] W. Li, B. Zhang, L. Zhang, and J. Yan. Principal line-based alignment refinement for palmprint recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 42(6):1491–1499, 2012.
- [16] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.
- [17] D. Peralta, I. Triguero, R. Sanchez-Reillo, F. Herrera, and J. Benitez. Fast fingerprint identification for large databases. *Pattern Recognition*, 47(2):588–602, Feb 2014.
- [18] P. Porwik and K. Wrobel. The new algorithm of fingerprint reference point location based on identification masks. In *Computer Recognition Systems*, pages 807–814. Springer, 2005.
- [19] R. Raghavendra, B. Dorizzi, A. Rao, and G. H. Kumar. Designing efficient fusion schemes for multimodal biometric systems using face and palmprint. *Pattern Recognition*, 44(5):1076–1088, 2011.
- [20] A. Rattani, D. R. Kisku, M. Bicego, and M. Tistarelli. Feature level fusion of face and fingerprint biometrics. In *Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–5, 2011.
- [21] S. Ribaric and I. Fratric. Experimental evaluation of matching-score normalization techniques on different multimodal biometric systems. In *Electrotechnical Conference, MELECON 2006. IEEE Mediterranean*, pages 498–501. IEEE, 2006.
- [22] F. S. Samaria and A. C. Harter. Parameterisation of a stochastic model for human face identification. In *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, pages 138–142. IEEE, 1994.
- [23] P. Sharma and M. Kaur. Multimodal classification using feature level fusion and SVM. *International Journal of Computer Applications*, 76(4):26–32, 2013.
- [24] Z. Wang, C. Liu, T. Shi, and Q. Ding. Face-palm identification system on feature level fusion based on CCA. *Journal of Information Hiding and Multimedia Signal Processing*, 4(4):272–279, 2013.
- [25] Y.-F. Yao, X.-Y. Jing, and H.-S. Wong. Face and palmprint feature level fusion for single sample biometrics recognition. *Neurocomputing*, 70(7):1582–1586, 2007.
- [26] Y. Yin, L. Liu, and X. Sun. SDUMLA-HMT: a multimodal biometric database. In *Biometric Recognition*, pages 260–268. Springer, 2011.
- [27] J. Zou, J. Feng, X. Zhang, and M. Ding. Local orientation field based nonlocal means method for fingerprint image de-noising. *Journal of Signal and Information Processing*, 4:150, 2013.