Chapter 3

Fuse: Online Learning Framework for an Adaptive Biometric System

Existing biometrics techniques are unable to provide significant levels of accuracy in uncontrolled noisy environments in several applications such as assisting law enforcement agencies to control crime and fraud. Further, scalability is another challenge due to variations in data distribution with changing conditions. This chapter presents a novel adaptive context switching algorithm coupled with online learning to address both these challenges. The proposed framework, termed as QFuse, uses the quality of input images to dynamically select the best biometric matcher or fusion algorithm to verify the identity of an individual. The proposed algorithm continuously updates the selection process using online learning to address the scalability and accommodate the variations in data distribution. The results on the WVU multimodal database and a large real world multimodal database obtained from a law enforcement agency show the efficacy of the proposed framework.

3.1 Introduction

A biometric system classifies an individual as genuine or impostor based on modalities such as face, fingerprint and iris. A traditional unimodal biometric system extracts features for the given biometric modality and compares it with stored database templates and computes a match score [159]. For verification settings (1 : 1 matching), the match score is classified as genuine or impostor. Unification of different biometric samples or evidence (such as face, fingerprints, and iris) to verify the identity of an individual is referred to as multicentric. Such multimodal systems offer additional benefits over unimodal systems such as resiliency to noise and malfunction, universality, and improved accuracy. However, the performance of a biometric system degrades when probe (query) images are
Figure 3.1: Illustrating examples of multimodal images from the same subject with varying quality (a) fingerprint, (b) iris (the last two rows demonstrate the unwrapped iris images and occlusion mask that indicate iris feature occlusion), and (c) face images.

of lower quality compared with the images that the system has encountered during training. As shown in Figure 3.1, the quality of probe images may degrade because of several reasons, such as improper illumination, improper interaction with the sensor (e.g. pose variations), and different kinds of noise or blur introduced in the probe image when the image is captured in an uncontrolled environment (applicable in many real-world applications). Fingerprints can suffer from dryness, iris images can have occlusion or improper illumination, while face images can be of low resolution or have pose variations.

Biometric systems generally do not facilitate case-based switching for selecting an appropriate classifier or combination of classifiers. Moreover, any biometric system or case-based switching criteria that is learned on limited (in terms of availability and variety) training data performs adequately only if the test data distribution is similar to the training data distribution. Since new users continue to be enrolled into the system, a biometric
Figure 3.2: General concept of the proposed QFuse algorithm that selects a single information source or fusion of multiple sources based on its reliability and the quality of information.

The system needs to be re-trained with new enrolments to accommodate the variations caused due to incremental data and maintain the accuracy levels. However, in many real world applications, re-training biometric classifiers may not be pragmatic as it requires all the training data in batch mode.

It is well understood that in different operating scenarios, information from some sources may be more useful than others. Hence, a mechanism is required to efficiently combine evidences based on situational cues. To incorporate this facility, the thesis proposes QFuse, an online learning algorithm for adaptive biometric fusion that incorporates image quality in the dynamic selection of unimodal classifiers and their fusion. Figure 3.2 shows the generalized concept of the proposed algorithm that selects information source(s) based on the reliability of each source and certain discriminatory cues (quality) of the information. To accommodate the variations in data distributions and sustain the performance with increasing number of users, the thesis also proposes to update the classifiers used in the proposed algorithm, QFuse, in an online manner. Specifically, to address the first challenge, the proposed algorithm uses different unimodal classifiers arranged serially in decreasing order of their reliability (accuracy) to process gallery-probe pairs one modality at a time. The serial arrangement of classifiers is based on our assertion that a unimodal classifier can efficiently match a good quality gallery-probe pair. However, unimodal clas-
sifiers yield conflicting results when the quality of gallery-probe pair degrades. In such cases, complementary information from multiple unimodal classifiers can be efficiently combined to yield correct results. Secondly, the research proposes to update QFuse in an online manner with only the new enrolment data. It provides a huge benefit in terms of computational time as well as improved accuracy (later validated in experimental results). The major contributions of this research are summarized below:

1. A serial context switching algorithm is proposed to select the most appropriate constituent unimodal classifier or fusion algorithm for a given set of gallery-probe pair based on its quality.

2. An online learning algorithm is proposed to update the context switching rule with the new incremental enrolment data to address the variations in data distribution.

The experiments are performed on two databases: WVU multimodal database [50] and a large scale multimodal database obtained from law enforcement agencies. The experimental analysis shows that the proposed framework not only improves the verification accuracy but also significantly reduces the computation time required for updating the framework.

### 3.2 Literature Review

Unification of multiple biometric information can be performed via two approaches: (1) matcher fusion and (2) dynamic matcher selection [205]. In matcher fusion, all the constituent matchers are used and their evidences are combined using fusion rules [100], [158], [172], [129]. On the other hand, dynamic matcher approaches include selecting the most appropriate matcher or a subset of specific matchers [74], [187], [152]. In the biometric literature, matcher fusion approaches have received significant attention [159], [158]; however, dynamic matcher selection has not been extensively explored. Traditionally, in multi-modal fusion approaches, designing a fusion scheme and performance evaluation are considered as two different stages. However, Toh et al. [181] proposed to simultaneously optimize the target performance and design of classifier fusion algorithm based on an approximation of the total error rate. Marcialis et al. [119] proposed serial fusion of face and fingerprint matchers where significant reduction in verification time was achieved. However, their approach did not consider the quality of the gallery-probe pair and was based only on the match score distribution of genuine and impostor scores. Traditional multi-biometric systems work on static fusion rules which may not adapt itself to the
dynamically changing environment and thus degrade the performance as the environment changes. Geng et al. [73] proposed a context aware fusion scheme that takes into account the viewing angle and distance of the subject from the camera to select an appropriate fusion scheme for improved performance. Abaza and Ross [?] proposed including image quality in the fusion scheme to enhance the performance in presence of weak matchers or low quality input images. Alonso-Fernandez et al. [15] proposed a method for efficient combination of match scores from different devices (sensors) depending on the quality of data source. Different modalities may yield heterogeneous scores; therefore, score normalization is required to transform these scores into a common domain. Jain et al. [89] analyzed different normalization and fusion techniques in the context of a multimodal biometric system. Veeramachaneni et al. [191] proposed using particle swarm optimization to switch between different fusion rules for combining decisions received from multiple biometric sensors. Kumar et al. [107] proposed a hybrid particle swarm optimization based approach for adaptive combination of multiple biometric modalities. Raghavendra et al. [157] proposed an efficient fusion scheme to combine complementary information from different biometric modalities at match score level. They also proposed a particle swarm optimization (PSO) procedure for reducing the dimensionality of feature space by identifying a subspace of the large dimension features. For evaluating, comparing, and benchmarking quality-dependent, client-specific, cost-sensitive score-level fusion algorithms, Poh et al. [150] prepared a score and quality database. They also reported the baseline experimental results for evaluating the above three types of fusion scenarios. Faundez-Zamuy [62] analyzed different types of data fusion and stages in a biometric system where fusion can be applied. Poh et al. [153] proposed a user-specific and selective fusion strategy to combine multiple biometric modalities. Their algorithm assigned a different set of fusion parameters to a given enrolled user and selected a subset of modalities for fusion. Their approach achieved better performance at a reduced computational cost based on a criterion called B-ratio that ranked subjects based on their match score statistics. Recently, Huang et al. [87] proposed general multimodal recognition framework which is termed as adaptive bimodal sparse representation-based classification. Utilizing a two-phase sparse coding strategy for precise quality assessment of face and ear images, the framework combined multimodal features for improved performance. Kanhangad et al. [96] propose a dynamic match score combination approach for palm print and 3D hand geometry techniques.

Vatsa et al. [185] proposed a parallel algorithm to select an appropriate constituent unimodal matcher or the fusion algorithm. Their algorithm supported biometric image
quality based and case-based switching for improved recognition performance. However, their approach performed switching in a parallel manner and required all the biometric modalities to be processed upfront; therefore, it was computationally more expensive. Recently, Nair et al. [126] proposed multinomial and geometric models for multibiometric systems in which the framework predicts the matching subjects in a multi-view/multimodal biometric environment via a case based switching approach. In a preliminary version of this manuscript, Bhatt et al. [34] proposed a serial context switching algorithm to address the limitations of a parallel framework and achieve better performance.

Updating a classifier’s knowledge using online learning has been actively studied in the machine learning community [39]; however it is applicability has recently been realized in the biometrics community. In the biometrics literature, incremental learning approaches with principal component analysis [156] and linear discriminant analysis [183] have been shown to be robust against the variations introduced due to incremental data. Singh et al. [171] introduced an online learning approach for updating a face matcher. Later, Kim et al. [99] proposed an online learning algorithm for biometric score fusion. Recently, Bhatt et al. [35] proposed to use labeled as well as unlabeled information for updating biometric matchers using an online co-training approach. While most of the existing approaches in biometrics literature have focused on updating a unimodal matcher, this work, to the best of our knowledge, is the first to incorporate online learning in a quality based modality switching algorithm.

3.3 QFuse: Quality Based Context Switching with Online Learning

The proposed quality based context switching algorithm recognizes individuals captured in uncontrolled environment where the quality of probe images is low. In this research, “context” refers to a biometric modality, i.e. face, fingerprint, and iris or their multi-modal fusion. Therefore, context switching refers to switching from one modality to another depending on certain cues (the quality) obtained from a given gallery-probe image pair. The proposed algorithm efficiently matches individuals using one of the unimodal matchers when the gallery-probe pair is of good quality and dynamically switches to fusion of multiple matchers across different modalities when the gallery-probe pair is of poor quality. The algorithm is further trained in an online manner to adapt the variations introduced due to new enrolments. This section first presents the quality assessment algorithms used in the context switching algorithm followed by the proposed algorithm.
<table>
<thead>
<tr>
<th>Energy spectrum</th>
<th>0.01</th>
<th>0.70</th>
<th>0.08</th>
<th>0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edge spread</td>
<td>0.49</td>
<td>0.31</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>Blockiness</td>
<td>0.41</td>
<td>0.24</td>
<td>0.22</td>
<td>0.58</td>
</tr>
<tr>
<td>Activity</td>
<td>0.36</td>
<td>0.19</td>
<td>0.26</td>
<td>0.46</td>
</tr>
<tr>
<td>ZC-rate</td>
<td>0.19</td>
<td>0.41</td>
<td>0.48</td>
<td>0.43</td>
</tr>
<tr>
<td>Pose</td>
<td>0.79</td>
<td>0.84</td>
<td>0.67</td>
<td>0.69</td>
</tr>
</tbody>
</table>

**Figure 3.3:** Sample images and their corresponding quality assessment scores for (a) face, (b) fingerprint, and (c) iris images. These quality scores are obtained after segmentation and utilized by support vector machines for context switching.
3.3.1 Quality Assessment

It has been shown in literature that quality assessment scores of a biometric sample can be indicative of its recognition performance [16, 32, 33]. The proposed algorithm computes a quality vector for a given image using computationally inexpensive quality assessment techniques. The quality vector comprises of both image quality metrics and modality specific quality metrics. Each of the quality metric is briefly discussed below. Further details regarding the quality metrics used in this research are described in 6.

- **Image quality metrics**: The first set of quality metrics is related to the quality of input gallery and probe images. We have used three algorithms for quality assessment:

  *No-reference quality*: Quality degradation due to compression artifacts can be computed by estimating the blockiness and activity of an image, as proposed by Wang et al. [199]. To effectively utilize the quality metric, three separate estimations of degradation in the image, namely blockiness (B), activity (A) and zero-crossing rate (ZC) are computed and combined in both horizontal and vertical directions.

  *Edge spread*: Motion and off-focus blur are measured using edges and adjacent regions [64]. It is estimated as the difference in image intensity with respect to the local maxima and minima of pixel intensity at every row of the image.

  *Spectral energy*: Block-wise spectral energy is calculated using Fourier transform components [132] which represent sudden changes in illumination and specular reflection. The image is divided into non-overlapping blocks and the spectral energy is computed as the magnitude of Fourier transform components within each block.

- **Modality specific image quality**: The next set of quality parameters is related to biometric information. For each modality, a specific set of parameters (details in Appendix A) is calculated which represents the usability of the image (here we focus only on three modalities, viz. face, fingerprint, and iris).

  *Face quality*: Pose variations in face degrade the performance as some of the facial features may not be visible. Such variations reduce the usability of the face image and a good quality image may not be useful for recognition. In this research, pose is estimated geometrically based on positions of eyes and mouth.

  *Fingerprint quality*: For fingerprint images, Chen et al. [45] proposed to measure the quality of ridge samples by Fourier energy spectral density concentration in particular frequency bands where strong ridges manifest.
Iris quality: Kalka et al. [95] presented quality assessment of iris images based on the evaluation of six separate quality factors (defocus, motion blur, occlusion, specular reflectance, illumination, and pixel count).

Using the above mentioned quality assessment algorithms, a quality vector is computed for a given image. Figure 3.3 shows quality metrics obtained from sample images of each modality. For a given gallery-probe pair, the quality vector of both gallery and probe images are concatenated to form the quality vector $Q$, represented as $Q = [Q_g, Q_p]$, where $Q_g$ and $Q_p$ are the quality vectors of gallery and probe images respectively. This quality vector is then used in the proposed context switching algorithm to dynamically select a biometric matcher/fusion algorithm. These quality metrics are based on using computationally inexpensive cues to determine the work flow of the system as they are known to be indicative of the ‘ability’ of a given sample to be identified by a given biometric modality. However, these may be replaced by other metrics [210] or meta-data depending on the specific use-case scenario. As shown in Figure 3.4, the quality vector for each modality is computed sequentially one at a time when the modality is requested. Unlike existing approaches that process all modalities simultaneously, the proposed approach...
enhances the computational ease by processing each modality serially. The quality vector of the next modality is computed only when the previous modality either fails to efficiently classify the given gallery-probe pair or is not available for processing.

3.3.2 Context Switching Algorithm

As mentioned before, context switching refers to switching from one modality to another depending on the quality of a given gallery-probe image pair. It is observed that when the quality of a given gallery-probe pair is good, uni-modal matchers are sufficient to classify the given pair as genuine or impostor; however, when the quality is poor, fusion of different modalities is required for classification. The proposed context switching algorithm is hierarchical in nature (as shown in Figure 3.4) and the selection of biometric modality/unimodal matcher is posed as a classification problem. In this research, Support Vector Machine (SVM) is used because (i) it has been shown to provide better performance for higher dimensional classification tasks [166], and (ii) the learning is dependent on representative samples rather than the numbers of training samples. At level-0, the quality vectors of the gallery-probe pairs are processed by the first SVM to determine the biometric modality to be used and at level-1, the matcher for the biometric modality is selected using the second SVM. The algorithm allows each modality to be processed one at a time, and the control is passed to other modality only if the first modality is not sufficient to match the given gallery-probe pair. The context switching algorithm is divided into three stages: training the SVMs, online learning during new enrolments, and dynamic matcher selection during probe verification. Each of the three steps is explained below in detail.

3.3.2.1 Training the SVMs

The SVMs for each biometric modality are trained independently using the labeled training data. For each modality, SVM classifier at level-0 and level-1 are referred to as $SVM_{L0}$ and $SVM_{L1}$ respectively. Further details about training the SVM at each hierarchical level are elaborated below:

1. At level-0, a binary SVM is trained for the $j^{th}$ biometric modality using the labeled training data $\{x_{ji}, y_{ji}\}$. Here, input $x_{ji} = [Q_{gi}, Q_{pi}]$ is the quality vector of the $i^{th}$ gallery-probe pair in the training set. $y_{ji} = \{-1, +1\}$ is the label such that $\{-1\}$ is assigned when the gallery-probe pair can be correctly classified using the matchers in the given modality, otherwise, $\{+1\}$ is assigned (i.e. control is switched to other
modality). Using the labeled training data, three $SVM_{L0}$ are trained: one for each modality.

2. Similarly, SVMs at level-1 are also binary classifiers trained for selecting one of the unimodal matchers for a given modality. For the $j^{th}$ biometric modality, input to $SVM_{L1}$ is also the quality vector $[Q_g, Q_p]_{ij}$ of the $i^{th}$ gallery-probe image pair in the training set. As shown in Figure 3.5, the labels are assigned based on the distribution of the genuine-impostor scores and the likelihood ratios [127]. For a matcher, if the score corresponding to a gallery-probe pair is greater than the maximum genuine score (i.e. confidently matched as impostor) or is less than the minimum impostor score (i.e. confidently matched as genuine), then this matcher can efficiently match the pair as genuine or impostor. Label $\{-1\}$ is assigned when matcher1 can confidently match the gallery-probe pair, otherwise, label $\{+1\}$ is assigned to select matcher2. If both the matchers correctly match the given gallery-probe pair then likelihood ratio is used to break the tie. The gallery-probe pair is assigned the label corresponding to the matcher that is more likely to match the gallery-probe pair as genuine (likelihood > 1) or impostor (likelihood < 1). The scores from the matchers are converted to distance scores (wherever required) before assigning the labels. For each modality, the labeled data is then used to train $SVM_{L1}$.

### 3.3.2.2 Online Learning

As mentioned previously, large scale programs such as US Visit (now OBIM: Office of Biometric Identity Management)$^1$ and Aadhaar$^2$ continuously enroll new subjects on a regular basis while performing probe verification. With increase in the number of enrollment, the quality and match-score distributions tend to drift. The classifiers trained with small training samples are unable to generalize well to this concept drift and require re-training to accommodate the variations in data distribution. Re-training the classifiers in batch mode with all the existing data is not pragmatic as it requires large amount of time. Online learning provides an efficient way to sustain the performance by addressing the variations in data (match score and quality score) distribution introduced by the newly enrolled individuals. The idea of context switching algorithm that can evolve with increasing number of new users is novel and to the best of our knowledge, this work presents the first approach for online context switching in biometric literature. Labeled information

---

1. [http://www.dhs.gov/obim](http://www.dhs.gov/obim)
Figure 3.5: Illustrating the process of assigning labels: genuine and impostor match score distributions are used to assign labels to the input gallery-probe quality vector $Q = [Q_g, Q_p]$ during SVM training.

from only “newly enrolled individuals” is used to update the proposed context switching algorithm (i.e. decision boundary of SVM classifiers) in an online learning (incremental+decremental) manner. In this research, we add or remove one sample at a time to update SVM using the incremental+decremental method proposed by Cauwenberghs and Poggio [39]. They proposed a solution for $N \pm 1$ samples that can be obtained using the $N$ old support vectors and the sample to be added or removed. SVMs are first trained using an initial training database and a decision hyperplane is obtained, as illustrated in Section 3.3.2.1. Online learning algorithm for updating the SVM with additional labeled instances from new enrolments and $m$ support vectors learned on an initial labeled training data $D_L$ is described in Algorithm 1.

**Algorithm 1** Online learning with new enrolments

**Input:** A SVM model, $N$ additional labeled instances $\{x_i, z_i\}$.

**Process:** Online training of SVM model

for $i = 1 \text{ to } N$ do
  Predict labels: $SVM(x_i) \rightarrow y_i$.
  if $y_i \neq z_i$ then
    Update $SVM$ decision boundary with labeled instance $\{x_i, z_i\}$ & $m$ support vectors.
  end if.
end for.

**Output:** Updated SVM model.
During new enrolment, a unique identification is assigned to every user. Impostor scores are computed by comparing a new enrollee with the stored gallery. For genuine scores, the enrollee is compared with its own multiple samples captured during enrolment. The labels (ground truth) corresponding to the enrollee are known during enrolment and are compared with the predictions of the SVM classifiers. The decision boundary of SVM classifier is updated in online fashion only for the instances for which the classifier makes incorrect predictions. The process of updating SVM decision boundary using the new available instances and the previous support vectors is elaborated below:

1. Let \( x_i \) be the instance for which SVM needs to be updated and \( z_i \) is the associated label.

2. SVM decision hyperplane is recomputed using the \( m \) trained support vectors and the new training instance \( \{x_i, z_i\} \) using standard batch mode, as explained in Section 3.3.2.1.

3. The number of support vectors may increase on recomputing the hyperplane. To avoid the number of support vectors from growing in an uncontrolled manner, a threshold \( \lambda \) is introduced that controls the number of support vectors. If the number of support vectors is more than \( m \pm \lambda \), then the farthest support vector from the current decision hyperplane is selected.

4. The farthest support vector is then removed from the list of support vectors and is added to a separate list, \( l \). The classifier with remaining \( m + \lambda - 1 \) support vectors is the updated classifier.

5. The support vectors in the list \( l \) are used to test the updated classifier. If there is any misclassification, Step 2 is repeated to minimize the classification error on the removed support vectors.

Online learning is used to update both the classifiers (\( SVM_{L0} \) and \( SVM_{L1} \)) in each modality. It facilitates to update the context switching algorithms with the varying quality and match score distribution. In this work, SVM with radial basis function (RBF) kernel is used where \( \gamma = 6 \).
3.3.2.3 Context Switching during Verification

For verification, the trained SVMs are used to select the modality and the most appropriate matcher for matching an individual using the quality of gallery-probe pairs. Each biometric modality is used one at a time and the second modality is invoked only when the unimodal matchers in the first modality are unable to classify the gallery-probe pair. The proposed online context switching algorithm processes a given gallery-probe pair as explained below:

1. The quality scores of the gallery-probe pair $Q = [Q_g, Q_p]$ corresponding to the first modality are computed and provided as input to the trained SVM$_{L0}$ of the first modality. Based on the quality vector, SVM$_{L0}$ predicts if the matchers in this modality can be used to correctly classify the pair or not.

2. If it predicts that unimodal matchers in the first modality can be used to correctly classify the given gallery-probe pair, then SVM$_{L1}$ in first modality is used to predict the unimodal matcher that should be selected.

3. Otherwise, if SVM$_{L0}$ predicts that the matchers in the first modality cannot efficiently classify the gallery-probe pair, quality vector corresponding to the second modality is computed and provided as input to the corresponding SVM$_{L0}$ of that modality. SVM$_{L0}$ pertaining to the second modality predicts whether the matchers in this modality can be used to correctly classify the given gallery-probe pair or not.

4. It is done serially for all the modalities until an appropriate unimodal matcher is selected.

5. If none of the unimodal matchers can efficiently classify the given gallery-probe pair, normalized score level fusion [158] of unimodal matchers across all modalities is selected to process the gallery-probe pair.

It should be noted that for probe verification, the algorithm does not require computing the match score between the gallery-probe pair to select the most appropriate matcher and SVM prediction is based only on the quality vector ($Q = [Q_g, Q_p]$) of the gallery-probe pair. Further, the proposed context switching algorithm is generic and can be modified to add or remove biometric modalities and matchers within a modality.
3.4 Experimental Results

To evaluate the effectiveness of the proposed context switching algorithm, the performance is evaluated on two multimodal biometric databases captured in significantly different conditions. The performance of the algorithm is also compared with individual unimodal matchers and normalized score level fusion. The effectiveness of the proposed online learning in updating the context switching algorithm is also compared with batch training. Additionally, a comparison is made with the parallel quality based context switching approach by Vatsa et al. [187]. Section 3.4.1 presents the unimodal matchers used in the algorithm, Section 3.4.2 illustrates the database characteristics and experimental protocol and Section 3.4.3 presents the results and key observations.

3.4.1 Unimodal Matchers

Three biometric modalities namely, face, fingerprint, and iris are used in the proposed context switching algorithm. For face, two matchers are used: Uniform Circular Local Binary Pattern (UCLBP) [12] as face matcher1 and Speeded Up Robust Features (SURF) [58] as face matcher2. UCLBP is computed with circular encoding of eight neighboring pixels evenly positioned on a circle of radius two. SURF is a scale and rotation invariant descriptor [58] that computes the descriptor from the spatial distribution of gradient information around the interest points. To match two corresponding UCLBP features or SURF descriptors, $\chi^2$ distance measure is used. Fingerprint matcher1 (NBIS) uses a minutiae based fingerprint matching algorithm. A commercial fingerprint matching software (Neurotechnology Veri-Finger) is used as fingerprint matcher2. Iris matcher1 is an implementation of the algorithm proposed by Vatsa et al. [184] which uses curve evolution based segmentation and 1-D log polar Gabor filters. Commercial iris matching software (Neurotechnology Veri-Eye) is used as iris matcher2. In each case, standard parameters from the corresponding cited work are used. Further, match scores are normalized using min-max normalization and sum rule is used for score level fusion [159].

3.4.2 Database and Experimental Protocol

The performance of the proposed algorithm is evaluated on two databases: WVU multimodal database [50] and the database provided by a Law Enforcement Agency (referred to as the “LEA” database).

1http://www.nist.gov/itl/iad/ig/nbis.cfm
• **WVU Multi-modal database** [50] comprises fingerprint, face, and iris images corresponding to 270 subjects with multiple samples per subject. The database is divided into three parts: 1) initial training, 2) online learning, and 3) testing. The training database comprises images pertaining to 108 subjects (40% of the total database) and the remaining images pertaining to 162 subjects are used for online learning and performance evaluation. The gallery comprises two images per subject and the remaining images are used as probe.

• **LEA database** is a multimodal database captured in unconstrained real world conditions with uncooperative users and comprises fingerprint, face, and iris images. The database is noisy, has both good and poor quality images, and has missing data as well; for instance, information from all the modalities are not available for every individual. The images are divided in two sets, set A and set B, and each set consists of 18,000 samples. LEA database is divided into initial training, online learning, and testing. The initial training is performed on images corresponding to 4,500 individuals, the online learning of the context switching algorithm is performed on the next 4,500 individuals, and the performance is evaluated on images corresponding to 9,000 individuals. In cases where one or more biometric modality is missing, the algorithm uses samples from other biometric modalities to perform context switching.

Face images in both the databases are normalized and the size of each detected image is $196 \times 224$ pixels. The proposed context switching algorithm selects the most appropriate matcher to process the gallery-probe pair based on the quality. QFuse also evolves with new enrolments to accommodate the variations in data distribution in online manner. During training, the parameters of feature extractors are learned and SVMs are trained using the gallery-probe quality vector and match score as explained in Section 3.3.2.1. During online learning, the decision boundary of SVM classifiers is modified to update the context switching rule as described in Section 3.3.2.2.

3.4.3 Results and Analysis

Figures 3.6 and 3.7 show the ROC curves comparing the performance of the proposed context switching algorithm with different unimodal matchers and sum-rule fusion. QFuse is also compared with a parallel quality based context switching framework of Vatsa et al. [187]. The key results and analysis are listed below.
Figure 3.6: ROC curves of the proposed quality based context switching algorithm QFuse and comparison with different unimodal matchers, sum-rule fusion and Vatsa et al. [187] on the WVU database [50].

Figure 3.7: ROC curves of the proposed quality based context switching algorithm and comparison with different unimodal matchers and sum-rule fusion on the LEA database.
Figures 3.6 and 3.7 show that the proposed quality based context switching algorithm outperforms the unimodal matchers by at least 3.8% and 10.6% on the WVU multimodal and LEA databases respectively. It also outperforms sum-rule fusion of unimodal matchers by at least 0.9% and 7.4% on the two databases respectively. The improvement is attributed to the fact that when images in a given modality are of bad quality, the context switching algorithm selects another modality or uses sum-rule [158] fusion of different unimodal matchers to process the gallery probe pair.

At lower FARs, the proposed approach yields improved verification performance compared to the parallel quality based context switching approach of Vatsa et al. [187]. The results show that comparable performance is achieved despite an additional evidence theory based fusion approach employed in the context switching framework of Vatsa et al. [187] along with the sum rule fusion. It is also observed that the proposed algorithm is computationally less expensive (at least 1.5 times faster) due to its serial nature, avoiding computation when strong evidence is obtained from a unimodal matcher. Since the parallel nature of the existing approach requires that all the evidences of the identity (i.e., all captured biometric modalities) should be available before processing and the LEA database contains instances of missing data, Vatsa et al’s algorithm fails to process majority of the samples from the LEA database. Therefore, in this research, the results of the proposed and existing algorithms are compared only on the WVU database.

The results suggest that the performance of unimodal matchers reduces significantly on the real world challenging LEA database due to noisy images. The proposed quality based context switching algorithm sustains its performance across noisy images and even when images from some modalities are not available. Since the size of the LEA database is large, the results also suggest that online learning makes it scalable as it adapts to changes in data distributions.

To evaluate the effectiveness of QFuse, the total number of individuals available for online learning are divided into 10 equal size batches and the average performance is reported by considering each batch sequentially. As shown in Table 3.2 and Figures 3.8a and 3.8b, online learning provides a minor improvement in verification accuracy over batch/offline training. However, Table 3.2 reports the overall training time for online learning and batch/offline training. The combined results (accuracy and time)
validate our assertion that the proposed online context switching algorithm provides significant reduction in training time (at least 1/4 of the batch/offline training time) and sustains the performance with increasing number of users.

- In the proposed algorithm, quality scores of a particular biometric modality are computed only when that modality is requested in the serial framework. Moreover, the algorithm can skip a modality, if the images for that modality are not successfully captured. Computationally, on an Intel i5 processor with 4GB RAM, the proposed algorithm requires an average of 3.1 seconds for quality assessment, feature extraction, context switching, and matching.

- Figures 3.9 and 3.10 illustrate few examples where the proposed context switching algorithm selects different unimodal matchers or their fusion for different quality gallery-probe image pairs. Since the images in the LEA database are captured in extremely harsh unconstrained real world environment with uncooperative users and are of poor quality, as shown in Table 3.3, 36.3% instances are processed using multimodal fusion. On the other hand, in the WVU multimodal database, which is prepared in controlled lab conditions with only selected irregularities; fusion is selected for only 14.7% instances. In both the cases, the results show that our assertion (i.e. for a good quality gallery-probe pair, unimodal matchers are sufficient and fusion should be used only for poor quality images) holds true and the proposed algorithm improves both accuracy and time.

- In this research, different modalities are arranged serially and each gallery-probe pair is processed starting from the strongest biometric modality. However, the hierarchy of the classifiers can also be decided based on the particular application scenario, based on performance, user convenience or other domain specific information.

The proposed quality based serial context switching algorithm can be easily extended to include other biometric modalities, unimodal matchers, and fusion rules. Since the LEA database is collected in-field by law enforcement officials, improved results (in terms of accuracy, computational time, and scalability) suggest that such a algorithm is very useful in real world large scale applications.

To evaluate the effectiveness of the proposed context switching algorithm, the performance is evaluated on two multimodal biometric databases captured in significantly different conditions. The performance of the algorithm is also compared with individual unimodal matchers and normalized score level fusion. The effectiveness of the proposed
Figure 3.8: Improvement with online learning over batch/offline learning on the a) WVU multimodal database [50] and b) LEA database.

Table 3.1: Verification accuracy of individual matchers, fusion algorithms, and QFuse at 0.01% false accept rate (FAR).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy (%)</th>
<th>WVU</th>
<th>LEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face matcher 1</td>
<td>84.6</td>
<td>32.1</td>
<td></td>
</tr>
<tr>
<td>Face matcher 2</td>
<td>80.7</td>
<td>25.6</td>
<td></td>
</tr>
<tr>
<td>Finger matcher 1</td>
<td>87.4</td>
<td>43.7</td>
<td></td>
</tr>
<tr>
<td>Finger matcher 2</td>
<td>91.3</td>
<td>50.3</td>
<td></td>
</tr>
<tr>
<td>Iris matcher 1</td>
<td>89.1</td>
<td>30.1</td>
<td></td>
</tr>
<tr>
<td>Iris matcher 2</td>
<td>82.8</td>
<td>36.7</td>
<td></td>
</tr>
<tr>
<td>Sum-rule Fusion</td>
<td>94.2</td>
<td>53.5</td>
<td></td>
</tr>
<tr>
<td>Vatsa et al. [187]</td>
<td>94.7</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Proposed</td>
<td>95.1</td>
<td>60.9</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Comparing the verification accuracy and computational time of QFuse when the training is performed in online and offline manner.

<table>
<thead>
<tr>
<th>Database</th>
<th>Training</th>
<th>Accuracy (%)</th>
<th>Training time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WVU</td>
<td>Online learning</td>
<td>95.1</td>
<td>13.2</td>
</tr>
<tr>
<td></td>
<td>Batch/offline training</td>
<td>95.3</td>
<td>82.5</td>
</tr>
<tr>
<td>LEA</td>
<td>Online learning</td>
<td>60.9</td>
<td>95.4</td>
</tr>
<tr>
<td></td>
<td>Batch/offline training</td>
<td>59.7</td>
<td>393.7</td>
</tr>
</tbody>
</table>
Figure 3.9: Illustrating sample cases where unimodal a) fingerprint, b) iris, and c) face matchers is selected.
Table 3.3: Illustrating percentage of instances processed by individual components involved in the proposed context switching algorithm.

<table>
<thead>
<tr>
<th>Matcher</th>
<th>% instances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WVU</td>
</tr>
<tr>
<td>Face matcher 1</td>
<td>10.6</td>
</tr>
<tr>
<td>Face matcher 2</td>
<td>8.9</td>
</tr>
<tr>
<td>Finger matcher 1</td>
<td>19.5</td>
</tr>
<tr>
<td>Finger matcher 2</td>
<td>17.3</td>
</tr>
<tr>
<td>Iris matcher 1</td>
<td>15.6</td>
</tr>
<tr>
<td>Iris matcher 2</td>
<td>13.4</td>
</tr>
<tr>
<td>Sum-rule fusion</td>
<td>14.7</td>
</tr>
</tbody>
</table>

Figure 3.10: Illustrating a sample case where multimodal fusion is selected.
online learning in updating the context switching algorithm is also compared with batch training. Additionally, a comparison is made with the parallel quality based context switching approach by Vatsa et al. [187].

3.5 Summary

In biometrics, evidence fusion paradigm has been widely used to establish the identity of an individual with greater confidence. Extensive research has been performed for controlled environment with cooperative users and higher accuracy has been achieved. However, there is a need to enhance the capabilities when operating under uncontrolled environment with noisy data. This thesis presents QFuse, a context switching algorithm coupled with online learning that fills the gap in the current state-of-the-art. The proposed algorithm analyzes the biometric samples that may be from diverse sensors with varying quality. It adaptively makes a decision if a unimodal biometric matcher can reliably verify the individual or a fusion rule is required. This research also updates the context switching algorithm using online learning approach in order to (1) make it scalable and (2) adapt to drift in data distribution due to new enrolments. The experimental results show that the proposed algorithm optimizes the accuracy and computation time for challenging large scale applications. Further, it is our assertion that this algorithm can be easily extended for other multimatcher problems in pattern recognition and machine learning applications.