ADAPTIVE BIOMETRIC AUTHENTICATION USING NONLINEAR MAPPINGS ON QUALITY MEASURES AND VERIFICATION SCORES

Jinyu Zuo, Francesco Nicolo, Natalia A. Schmid
West Virginia University
Department of CSEE, Morgantown, WV

Harry Wechsler
George Mason University
Department of CS, Fairfax, VA

ABSTRACT
Three methods to improve the performance of biometric matchers based on vectors of quality measures associated with biometric samples are described. The first two methods select samples and matching scores based on predicted values of Quality of Sample (QS) index (defined here as d-prime) and Confidence in matching Scores (CS), respectively. The third method treats quality measures as weak but useful features for discrimination between genuine and imposter matching scores. The unifying theme for the three methods consists of a nonlinear mapping between quality measures and the predicted values of QS, CS, and combined quality measures and matching scores, respectively. The proposed methodology is generic and is suitable for any biometric modality. The experimental results reported show significant performance improvements for all the three methods when applied to iris biometrics.

Index Terms— Quality factors, pattern recognition

1. INTRODUCTION
The most common definition of biometric sample quality is at the image or signal level. A quality checking block is introduced into every biometric system to ensure that enrolled image/signals have sufficient quality to be further processed.

Many recent biometric systems extract a vector of quality measures. The components of a vector of quality measures, however, rarely carry equal weight in terms of their relationship to the performance of the matcher. In practical applications (such as US Visit program), it is required to keep a single biometric quality measure in order to decide if biometric samples are suitable for further processing and matching. Research questions should thus be concerned with (i) what quality measures to use; (ii) how to combine multiple quality measures into a single quality index without losing the information that the vector of quality measures contains; and (iii) how to use this vector to improve performance of biometric systems? This paper addresses all those questions.

Most of the quality based matchers described in the literature involve biometric sample quality at the matching stage by concatenating matching scores due to the original matcher and quality measures. These matchers are known as Q-stack classifiers [1, 2]. In spite of the fundamental theory presented in these works in support of Q-stack classifiers, the improvement of performance is marginal, if at all (see [3], [1]). More noticeable improvements are reported for Q-stack classifiers operating on multiple algorithms or multiple matchers [3, 2].

This paper suggests several methods on the use of biometric sample quality to improve the performance of a single matcher. It targets two main applications for quality measures: 1) to improve performance of a matcher by predicting its QS index or CS score and using them to decide if the underlying biometric sample should be retained or discarded 2) to design a nonlinear matcher that treats a vector of quality measures as a set of weak features.

The remainder of the paper is organized as follows. Sec. 2 describes the three proposed methods. Sec. 3 describes the data sets used and presents experimental results. Sec. 4 summarizes the contributions.

2. METHODOLOGY
This paper advocates the use of predictive tools for the design of quality enhanced matchers. The tools proposed predict a set of quality measures and scores. In each of the three methods described below, the functional relationship between vectors of quality measures and the predicted (estimated) measures is not known and has to be modeled. The modeling problem is stated as a multivariate regression problem:

\[ Y = f(X_1, \ldots, X_K), \quad (1) \]

where \( f \) is a multivariate adaptive mapping, variable \( Y \) is the estimated output variable, i.e. that characterizes the overall quality, confidence in matching score, or quality enhanced decision, and \( X_1, \ldots, X_K \) is a vector of \( K \) input (predictive) variables, such as a vector of quality measures for a biometric sample or a concatenated vector of quality measures and matching scores. Since the true relationship between the input and output variables is not known, it is estimated using a set of labeled training data. The multivariate adaptive mapping \( f(\cdot) \) can be implemented using a variety of multivariate functions and systems. The results reported are obtained using a feed forward neural network (FFNN) (see Sec. 3).
2.1. Quality of Sample (QS)

Here we suggest a single quality index characterizing the overall quality of a biometric sample. Consider a set of biometric samples. Associated with these data are vectors of quality measures. For example, for iris biometrics a matcher may be designed to implement Hamming Distance (HD), while the quality vector may be composed of ten quality measures: 1) iris segmentation score, 2) pupil segmentation score, 3) blur, 4) illumination, 5) dilation, 6) interlacing, 7) lighting, 8) occlusion, 9) off-angle and 10) pixel count (see [4] for details). These data can be used to design a single quality index by fusing the entries of the vector. The main design requirement is that the values of the single quality index must be related to the performance of the original matcher.

We propose to use d-prime as the combined quality index. This index is related to the relative entropy, which is asymptotically related to the performance of the matcher. Since the true distributions, means and variances of matching scores are not known, they are estimated using available labeled data. All labeled data are subdivided into two non-overlapping sets: training set and testing set. The confidence in genuine and imposter scores (CS) is defined for an iris recognition system:

\[
CS_G = \begin{cases} 
0, & \text{if } HD_G < Q(HD_G)_x, \\
-\frac{HD_G - Q(HD_G)_x}{Q(HD_G)_y - Q(HD_G)_x}, & \text{otherwise},
\end{cases}
\]

\[
CS_I = \begin{cases} 
0, & \text{if } HD_I < Q(HD_I)_{1-x}, \\
-\frac{Q(HD_I)_{1-y} - HD_I}{Q(HD_I)_{1-y} - Q(HD_I)_{1-x}}, & \text{otherwise},
\end{cases}
\]

where \(Q(HD_G)_x\) and \(Q(HD_I)_y\) are the quantile points at the quantile \(x\) and \(y\) for genuine and imposter scores, respectively. The levels of the quantiles were optimized empirically. The values resulting in significantly improved verification performance are \(x = 0.7\) and \(y = 0.9\). Fig. 1 illustrates genuine and imposter distributions typical for HDs.

The CS of biometric samples is predicted using a nonlinear adaptive mapping and vectors of quality measures (input parameters). The nonlinear adaptive mapping is trained using a set of labeled data in the form of vectors of quality measures and the corresponding CS values obtained using equations (3) and (4). At the testing stage the CS value is predicted based on vectors of quality measures only. This information is used to keep or discard the corresponding matching score in order to improve the performance of the original matcher. The procedure of predicting the CS of matching scores between two biometric samples \(A\) and \(B\) is:

\[
CS_{AB} = f_{CS}(Q_A, Q_B),
\]

where \(f_{CS}(\cdot)\) is a multivariate adaptive mapping (FFNN in our case) estimated using training data.

2.2. Confidence in Scores (CS)

The second method evaluates the confidence level assigned to matching scores associated with a pair of biometric samples. The confidence in genuine and imposter scores (CS) is

\[
QS_A = \frac{|m(\text{Imp. Scores}_A) - m(\text{Gen. Scores}_A)|}{\sqrt{\text{var(}\text{Imp. Scores}_A\text{)}} + \text{var(}\text{Gen. Scores}_A\text{)}}.
\]
cision made by the classifier. The decision is a binary valued variable corresponding to \{Genuine, Imposter\}. The high dimensional classifier is implemented using a nonlinear adaptive mapping. Denote \(\eta_{AB}\) as the output variable predicted using vectors of the quality measures and the matching score of the biometric samples \(A\) and \(B\). Then the prediction procedure is described as \(\eta_{AB} = \hat{f}_{QST}(Q_A^T, Q_B^T, MS_{AB})\), where \(\hat{f}_{QST}\) is an estimated version of \(f_{QST}\).

3. EXPERIMENTAL RESULTS

The results illustrate the feasibility and utility of our enhanced biometric matchers for iris biometric. We briefly describe the data and experiments below.

All experiments were performed using ICE 2005 dataset [5]. The matching procedures followed a modified log-Gabor implementation by Masek [6] refined by the authors. 26,867 genuine matching scores and 4,331,761 imposter matching scores from 2,953 samples were generated.

3.1. Neural Network

The nonlinear mapping is implemented using Feed Forward Neural Network (FFNN). Training data are assigned labels according to the functional use of the mapping and a set of input and output parameters. The final design is achieved by trading off the complexity and the performance of the network with two hidden layers. For the iris experiments, the first hidden layer of the FFNN is composed of 16 neurons while the second layer is composed of 2 neurons. The training data are divided randomly in two subsets: a learning subset composed of 60\% of training data and a validation subset made of remaining 40\% of data. The training process stops when the mean square error drops below \(10^{-4}\). The experimental results described below are obtained using codes from the Neural Network Toolbox in MATLAB™.

3.2. QS evaluation

The performance of the QS method is evaluated by randomly selecting 1,500 iris images from the ICE 2005 dataset to form the training set. The remaining 1,453 iris images are used to form the testing set. The QS of unlabeled images from the testing set is predicted based on the quality vector (quality factors 1 through 10) and by using a FFNN trained on labeled samples. During performance evaluation, unlabeled images with the value of predicted QS above a preset quantile are retained. Fig. 3 displays three Receiver Operating Characteristic (ROC) curves parameterized by zero, 10\% and 40\% quantile levels. The ROC curve marked as “original” is parameterized by zero quantile level, which means that no poor quality biometric samples were discarded. Note that by discarding only 10\% of iris images with the low predicted QS index, a considerable performance improvement can be achieved. Fig. 4 displays a box plot of the Equal Error Rate (EER) as a function of the quantile used to select iris samples with high QS value. It is a summary of ten independent trials, where training and testing data are sampled at random. It can be observed that regardless of the composition of training and testing data, removing iris images characterized by low predicted QS improves matching performance of the original matcher. The higher the value of QS, the better the performance is.

3.3. CS Evaluation

To assess the performance of the CS method, 20,000 genuine matching scores, 200,000 matching imposter scores and the vectors of quality measures associated with iris images were used to train a FFNN. The remaining data were used for testing. The matcher was designed to be symmetric with respect to quality vectors, that is, if \(Q_A\) and \(Q_B\) are two vectors of quality measures associated with iris image \(A\) and \(B\), training included both the pair \((Q_A, Q_B)\) and the pair \((Q_B, Q_A)\) and the associated matching score. The testing experiment is similar to the experiment of the previous subsection with the difference that pairs of quality vectors are used to predict the CS values.

Fig. 5 shows three ROC curves: the original curve, the curve formed from iris data with the predicted CS values ex-
ceeding 20% quantile and the curve formed from iris data with the predicted CS values exceeding 50% quantile. Performance improves when low confidence matching scores are discarded. Fig. 6 summarizes the results of ten trials. Again, training set is formed by randomly sampling iris images from a larger set. The trends and results are consistent.

### 3.4. QST Evaluation

The matcher is now a FFNN trained and tested as follows. During training the label “1” is assigned to all genuine vectors on the input and the label “-1” is assigned to all imposter vectors on the input to the neural network. During testing the output label is predicted based on the input vector of quality metrics and the original matching score. The output label in this case is a real number. The high dimensional classifier makes decision in favor of Genuine class if the output label is close to 1. It decides in favor of Imposter, if the output label is closer to −1. When the decision threshold varies, the performance of the high dimensional classifier is characterized by the ROC curve.

To assess the performance of the QST method 20,000 genuine scores and 200,000 imposter scores and associated quality vectors were involved in training. The remaining vector-triplets were used for testing. The success of the QST method depends on the data selected for training and testing. A single trial out of set of 20 trials resulted in a perfect separation of genuine and imposter matching scores. In the other cases the % improvement was between 20% and 35%. The results of five first trials are shown in Table 1. In columns 2 and 3 it displays the values of EER without/with quality factors. In columns 3 and 4 the table displays the values of False Reject Rate (FRR) evaluated at 0.001 False Accept Rate (FAR).

### 4. CONCLUSIONS

Three new methods for matching iris biometrics using quality metrics are proposed. The methods are adaptive and use nonlinear mappings for making predictions on quality measures and corresponding verification scores. The experimental results reported illustrate the importance of predictive and selective integration of quality measures for biometric authentication and show significant advantages compared to existing methods. Our future research will focus on (i) designing and analyzing quality based multi-modal fusion schemes; and (ii) developing the concept of quality at different processing levels in biometric systems.

### 5. REFERENCES


