An Integrated Prediction Model for Biometrics

Rong Wang, Bir Bhanu, and Hui Chen

Center for Research in Intelligent Systems University of California,
Riverside Riverside, California 92521, USA
{rwang,bhanu,hchen}@vislab.ucr.edu

Abstract. This paper addresses the problem of predicting recognition performance on a large population from a small gallery. Unlike the current approaches based on a binomial model that use match and non-match scores, this paper presents a generalized two-dimensional model that integrates a hypergeometric probability distribution model explicitly with a binomial model. The distortion caused by sensor noise, feature uncertainty, feature occlusion and feature clutter in the gallery data is modeled. The prediction model provides performance measures as a function of rank, population size and the number of distorted images. Results are shown on NIST-4 fingerprint database and 3D ear database for various sizes of gallery and the population.

1 Introduction

The goal of pattern recognition is to classify patterns into a number of classes. Patterns can be images, signals or any other type of measurements that need to be classified [1]. Currently, in order to ensure the high confidence in security, biometrics (e.g. fingerprint, palm, face, gait, signature and speech) are used. Depending on application there are two kinds of biometric recognition systems: verification systems and identification systems [2]. A verification system stores users' biometrics in a database. Then it will compare a person's biometrics with the stored features to verify if this person is who she/he claims to be. This is a one-to-one matching problem. The system can accept or reject this person according to the verification result. An identification system is more complex, where for a query the system searches the entire database to find out if there are any biometric features saved in the database that can match the query. It conducts one-to-many matching [2].

Usually a biometric recognition system consists of three stages: image acquisition, feature extraction and matching. Distortion often occurs in these stages which is caused by the sensor noise, feature uncertainty, feature occlusion and feature clutter. In a biometric recognition system before we can widely use the recognition algorithm we need to evaluate its performance on a large population. Since we have very limited data, we can build a statistical model which is based on a small gallery to estimate its performance on large population. Considering the distortion problem that may occur in large population we present an integrated model which considers the distortion to predict the large population performance from a small gallery. Unlike the previous approaches based on a binomial model that use match and non-match score distributions, we present a generalized two-dimensional model that integrates a hypergeometric model explicitly with a binomial model.

T. Kanade, A. Jain, and N.K. Ratha (Eds.): AVBPA 2005, LNCS 3546, pp. 355–364, 2005.

Our paper is organized as follows. In section 2 we present the related work. In section 3 we describe the distortion model which includes uncertainty, occlusion and clutter. The detail of the integrated model are given here. Results are shown in section 4. The integrated model is tested on NIST-4 fingerprint database and 3D ear database for various sizes of small gallery and the large population. Conclusions are given in section 5.

2 Related Work

Until now the prediction models are mostly based on the feature space and similarity scores [3]. Tan et al. [4] present a two-point model and a three-point model to estimate the error rate for the point based fingerprint recognition. Their approach not only measures minutiae's position and orientation, but also the relations between different minutiae to find the probability of correspondence between fingerprints. They assume that the uncertainty area of any two minutiae may overlap. Hong et al. [5] present a method to predict the upper and lower bound for object recognition performance. They consider the data distortion problem in the prediction. In their method performance is predicted in two steps: compute the similarity between each pair of model; use the similarity information along with the statistical model to determine an upper and lower bound for recognition performance. Johnson et al. [6] build a *cumulative match characteristic* (CMC) model that is based on the feature space to predict the gait identification performance. Mahalanobis distance and L_2 norm are used to compute similarity within the feature space. They make an assumption about the density that the population variation is much greater than the individual variation. When this assumption is invalid this approach cannot be used.

Wayman [7] and Daugman [8] develop a binomial model that uses the non-match score distribution. This model underestimates recognition performance for large galleries. Phillips et al. [9] create a moment model, which uses both the match score and the non-match score distributions. Since all the similarity scores are sampled independently, their results underestimate the identification performance. Johnson et al. [10] improve the moment model by using a multiple non-match scores set. They average match scores on the whole gallery. For each match score they count the number of non-match scores that is larger than this match score, which leads to an error. In reality the distribution of match score is not always uniform. Grother et al. [11] introduce the joint density function of the match and non-match scores to solve the underestimation problem.

In this paper we present a generalized two-dimensional model that integrates a hypergeometric model with a binomial model to predict the large population performance from a small gallery. It considers the data distortion problem in the large population. The number of distorted images follows hypergeometric distribution. Like Hong et al. [5] our distortion model includes feature uncertainty, occlusion and clutter. The distortion model needs users to define some parameters, such as feature uncertainty *probability density function* (PDF), occlusion amount, clutter amount, clutter region, and clutter PDF etc. Then according to the different numbers of distorted images we get the distributions of match score and non-match score. After this we use the CMC curve [6] to rank all these scores. A CMC curve can show different probabilities of recognizing biometrics depending upon how similar the features for this query biometrics are in comparison to the other biometrics in the gallery. Finally we use a binomial distribution to compute the probability that the match score is within rank r. In this paper we consider the performance when the rank $r = 1$.

3 Technical Approach

We are given two sets of data: gallery and probe. Gallery is a set of biometric templates saved in the database. For each individual there is one template saved in the gallery. Probe is a set of query biometrics. Large population is the unknown data set whose recognition performance needed to be estimated based on the given gallery and probe set.

3.1 Distortion Model

Our distortion model includes feature uncertainty, occlusion and clutter. Assume $F = \{f_1, f_2, \cdots, f_k\}$ is feature set of the biometrics, where $f_i = (x, y, t)$, x and y represent feature's location, t represents feature's other attributes except for location. Then the distortion algorithm [5] does the following:

a) Uncertainty: Assume the uncertainty PDF follows uniform distribution. It represents how likely each feature is to be perturbed. We replace each feature $f_i = (x, y, t)$ with $f_i^{'}$ which is chosen uniformly at random from the set

$$\{(x', y', t'), (x', y') \in 4NEIGHBOR(x, y), (1 - \alpha)t \leq t' \leq (1 + \alpha)t\}$$

where α is a coefficient, usually $0 \leq \alpha \leq 1$. $4NEIGHBOR(x, y)$ means 4 points around (x,y), they are $\{(x-1, y), (x+1, y), (x, y-1), (x, y+1)\}$. The unit is pixel.

b) Occlusion: Assume the number of features to be occluded is O. Uniformly choose O features out of the k features, remove these features.

c) Clutter: Add C additional features, where each feature is generated by picking a feature according to the clutter PDF from the *clutter region* (CR). The clutter PDF determines the distribution of clutter over the clutter region. Clutter region is used to determine where clutter features should be added. The clutter region typically depends upon the given model to be distorted. We usually use a bounding box to define the clutter region

$$CR = \{(x, y, t), x_{min} \leq x \leq x_{max}, y_{min} \leq y \leq y_{max}, t_{min} \leq t \leq t_{max}\}$$

where x_{min} and x_{max} represent the minimum and maximum value of x, the same definition for y_{min}, y_{max}, t_{min} and t_{max}.

We define the distortion region of feature f, denoted by $DR(f)$, as the union of all features that could be generated as uncertain version of f. In order to simplify, we use uniform PDF for uncertainty and clutter. In fact other PDFs are also possible and can be implemented.

3.2 Prediction Model

Our two-dimensional prediction model considers the distortion problem which is much more conform with the reality than our previous work [3]. Assume we have two kinds of different quality biometric images, group #1 and group #2. Group #1 is a set of good quality biometric images without distortion. Group #2 is a set of poor quality biometric images with feature uncertainty, occlusion and clutter. In general, the size of these two groups are N_1 and N_2. We randomly pick p images from group #1 and group #2. Then the number of distorted images y which are chosen from group #2 should follow hypergeometric distribution.

$$f(y) = \frac{C_{p-y}^{N_1} C_y^{N_2}}{C_p^{N_1+N_2}} \tag{1}$$

where

$$\begin{aligned} C_{p-y}^{N_1} &= \frac{N_1!}{(p-y)!(N_1-p+y)!} \\ C_y^{N_2} &= \frac{N_2!}{y!(N_2-y)!} \\ C_p^{N_1+N_2} &= \frac{(N_1+N_2)!}{p!(N_1+N_2-p)!} \end{aligned}$$

where $N_1 + N_2$ is the total number of images in these two groups, $p - y$ is the number of images chosen from group #1.

In order to simplify the description we assume sizes of gallery and probe set are all n. For each image in the probe set we compute the similarity scores with every image in the gallery. Then we have one match score and $n - 1$ non-match scores for this image. Here we assume that the match score and the non-match score are independent. Thus for a given number of distorted images we get a set of match scores $M_i = [m_{i,1}, m_{i,2}, \cdots, m_{i,n}]$ and a set of corresponding non-match scores

$$NM_i = \begin{bmatrix} n_{i,1,1} & \cdots & n_{i,n,1} \\ \vdots & \ddots & \vdots \\ n_{i,1,(n-1)} & \cdots & n_{i,n,(n-1)} \end{bmatrix}$$

where i represents the number of images selected from group #2, $i = 1, 2, \cdots, n$. Now for a given number of distorted images i, jth image has a set of similarity scores which include one match score and $n - 1$ non-match scores

$$S_{ij} = \left[m_{i,j} \; n_{i,j,1} \cdots n_{i,j,(n-1)} \right]$$

where $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, n$.

If we have enough match scores and non-match scores then we can estimate their distributions. From above we know that the similarity score distributions depend not only on the similarity scores but also on the number of images with distortion. Here we assume $ms(x|y)$ and $ns(x|y)$ represent the distributions of match scores and non-match scores given the number of distorted images. Assume if the similarity score is higher then the biometrics are more similar. The error occurs when a given match score is smaller than the non-match score. For a given number of distorted images the probability that the non-match score is greater than or equal to the match score x is $NS(x)$ where

$$NS(x) = \int_x^\infty ns(t|y)f(y)dt \tag{2}$$

Then the probability that the non-match score is smaller than the match score is $1 - NS(x)$.

Here we assume that the similarity score distributions are similar for small gallery and large population. If the size of large population is N, then for jth image we can have a set of similarity scores, which include one match score and $N-1$ non-match scores. We rank the similarity scores in decreasing order. Then for a given number of images with distortion the probability that the match score x rank r is given by the binomial probability distribution

$$\mathrm{C}_{r-1}^{N-1}\left(1 - \int_x^\infty ns(t|y)f(y)\right)^{N-r}\left(\int_x^\infty ns(t|y)f(y)\right)^{r-1} \tag{3}$$

Integrating over all the match scores, for a given number of images with distortion the probability that the match scores rank r can be written as

$$\int_{-\infty}^\infty \mathrm{C}_{r-1}^{N-1}\left(1 - \int_x^\infty ns(t|y)f(y)\right)^{N-r}\left(\int_x^\infty ns(t|y)f(y)\right)^{r-1} ms(x|y)f(y)dx \tag{4}$$

We integrate over all the number of images chosen from group #2, the probability that the match scores rank r can be written as

$$\int_{-\infty}^\infty \mathrm{C}_{r-1}^{N-1}\left(1 - \int_x^\infty \sum_{y=0}^{n} ns(t|y)f(y)\right)^{N-r}$$
$$\left(\int_x^\infty \sum_{y=0}^{n} ns(t|y)f(y)\right)^{r-1} \sum_{y=0}^{n} ms(x|y)f(y)dx \tag{5}$$

In theory the match scores can be any values within $(-\infty, \infty)$. We get the probability that the match scores are within rank r is

$$P(N,r) = \sum_{i=1}^{r}\int_{-\infty}^\infty \mathrm{C}_{r-1}^{N-1}\left(1 - \int_x^\infty \sum_{y=0}^{n} ns(t|y)f(y)\right)^{N-r}$$
$$\left(\int_x^\infty \sum_{y=0}^{n} ns(t|y)f(y)\right)^{r-1} \sum_{y=0}^{n} ms(x|y)f(y)dx \tag{6}$$

Considering the correct match take place above a threshold t, the probability that the match score is within rank r becomes

$$P(N,r,t) = \sum_{i=1}^{r}\int_{t}^\infty \mathrm{C}_{r-1}^{N-1}\left(1 - \int_x^\infty \sum_{y=0}^{n} ns(t|y)f(y)\right)^{N-r}$$
$$\left(\int_x^\infty \sum_{y=0}^{n} ns(t|y)f(y)\right)^{r-1} \sum_{y=0}^{n} ms(x|y)f(y)dx \tag{7}$$

For the problem where rank $r = 1$ then the prediction model with the threshold t becomes

$$P(N,1,t) = \int_t^{\infty} \left(1 - \int_x^{\infty} \sum_{y=0}^{n} ns(t|y)f(y)\right)^{N-1} \sum_{y=0}^{n} ms(x|y)f(y)dx \quad (8)$$

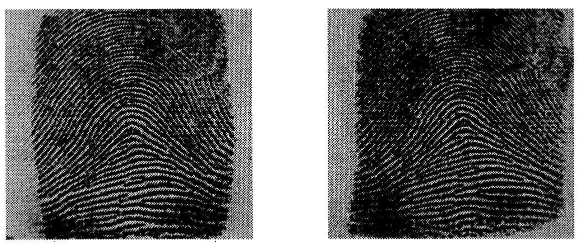
In this model we make two assumptions: match scores and non-match scores are independent and their distributions are similar for large population. In this model N is the size of large population whose performance needs to be estimated. Small size gallery is used to estimate distributions of $ms(x|y)$ and $ns(x|y)$.

4 Experimental Results

In this section we verify our model on NIST-4 fingerprint database and ear database for different sizes of small gallery and large population. Then we compare the performance of our integrated model with our previous binomial model on the NIST-4 fingerprint database.

4.1 Integrated Prediction Model

Fingerprint Database: All the fingerprints we use in the experiments are from *NIST Special Database 4* (NIST-4). It consists of 2000 pairs of fingerprints, each of them is labeled 'f' or 's' that represent different impressions of a fingerprint followed by an ID number. The images are collected by scanning inked fingerprints from paper. The resolution of the fingerprint image is 500 DPI and the size is 480×512 pixels. Figure 1 is a pair of fingerprints from NIST-4 database.



f0006_09 s0006_09

Fig. 1. Sample images from NIST-4

Usually the minutiae features are used for fingerprint recognition which can be expressed as $f = (x, y, c, d)$, where x and y are the locations of the minutiae, c is the class of minutiae, and d is the direction of minutiae. We define the percentage of minutiae with distortion for one fingerprint as g. In our experiments we choose $g = 5\%$, 7%, and 8% respectively. By applying distortion model to these 2000 pairs of fingerprints according to different distortion percentages, we get 6000 pairs of distorted fingerprints. Assume the number of minutiae is num_j, usually one pair of fingerprints have different number of minutiae so $j = 1, 2, \cdots, 4000$. The distortion model is as following:

(a) Uncertainty: Uniformly choose $U = g \times num_j$ minutiae features out of the num_j replace each $f_i = (x, y, c, d)$ with $f_i^{'}$ chosen uniformly at random from the set

$$\{(x', y', c', d'), (x', y') \in 4NEIGHBOR(x, y), c' = c \pm 1, d' = d \pm 3°\}$$

where $i = 1, 2, \cdots, U$.

(b) Occlusion: Uniformly choose $O = g \times num_j$ minutiae features out of the num_j remove these minutiae.

(c) Clutter: Add $C = g \times num_j$ additional minutiae, where each minutiae is generated by picking a feature uniformly at random from the clutter region. Here we choose the clutter region as

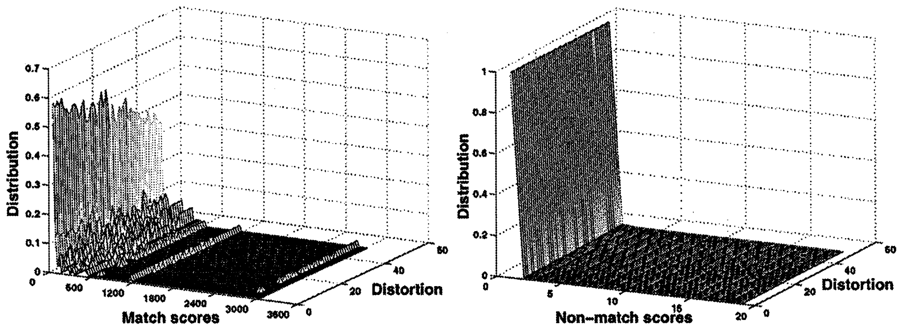
$$CR = \{(x, y, c, d), 50 \le x \le 450, 60 \le y \le 480, c = \{0, 1, 2, 3, 4\}, 10° \le d \le 350°\}$$

In our experiments we use the uniform distribution as the uncertainty PDF and the clutter PDF. The number of features with uncertainty, occlusion and clutter are the same. By adding different percentage of minutiae with distortion g we have four groups of fingerprint images, each group has 2000 pairs of fingerprints. Group #1 is the original fingerprints in NIST-4, group #2 is the fingerprints with $g = 5\%$, group #3 with $g =$ 7%, and group #4 with $g = 8\%$.

Assume our small galley size $n = 50$. We randomly pick up 50 fingerprints pairs from group #1 and group #2. Then the number of fingerprints chosen from group #2 which denoted by y follows hypergeometric distribution,

$$f(y) = \frac{C_y^{50} C_{50-y}^{50}}{C_{50}^{100}} \tag{9}$$

Now we have 50 pairs of images including the original images and the distorted images. The images labeled with 'f' are used as the gallery and the others labeled with 's' are used as the probe set. We use fingerprint verification approach which based on the triplets of minutiae to compute the similarity scores for these fingerprints [12]. Then we get the distributions of the match score and the non-match score. Figure 2 is the distributions of the match score and the non-match score for different number of distorted images. From Figure 2 it's clear that these distributions depend not only on similarity scores also on the number of distorted images. Here we choose the threshold for correct match $t = 12$. For the verification problem we consider the case when rank $r = 1$. This small gallery $n = 50$ applies in the integrated prediction model which can predict the large population performance, here we choose $N = 6000$. Now we get the prediction result for $g = 5\%$. By repeating the above process we get the estimation results for $g = 7\%$ and $g = 8\%$. Average these three prediction values we get the estimation result for large population $N = 6000$. We choose different size of small gallery $n = 70$. By repeating the above process we obtain the estimation results for large population. Figure 3 shows the absolute error between the predicted and experimental verification performance. The absolute error is smaller than 0.08 when the population size is larger than 1000. That means our integrated prediction model can efficiently predict the fingerprint recognition performance for large population.



(a) (b)

Fig. 2. Similarity scores distributions. (a) Match scores distribution. (b) Non-match scores distribution

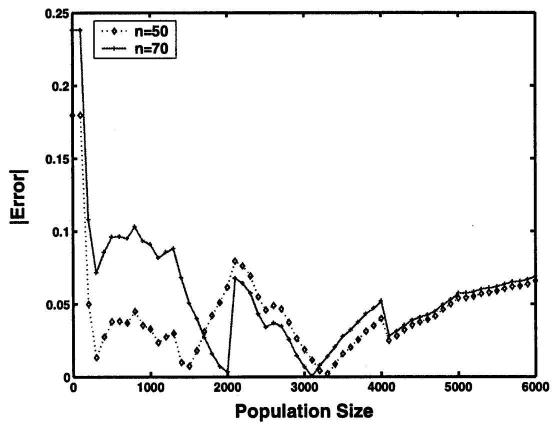
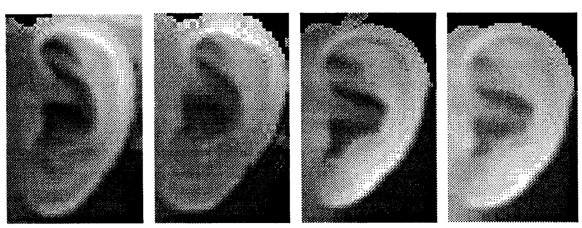


Fig. 3. Absolute error between the integrated prediction model and experimental fingerprint recognition performance



model_1 test_1 model_2 test_2

Fig. 4. Sample images from 3D ear database

Ear database: Ear data we use in this experiment are acquired by using Minolta Vivid 300. The image contains 200×200 grid points which has 3D coordinate (x, y, z). This data set has 52 subjects and every subject has two range images which are taken at different viewpoints. Figure 4 shows two pairs of ear. We add Gaussian noise $N(0, \sigma =$

Table 1. Prediction using the integrated model and experimental ear recognition performance

Gallery Size	Experiment	Prediction
50	88.00%	83.28%

Table 2. Predicted ear recognition performance for different sizes of large population by the gallery of 52 objects

Gallery Size	Prediction Results
100	81.67%
150	81.22%
200	81.07%
250	81.01%
300	80.99%

$0.6mm$) to these images. Then we have two image groups: group #1 has 52 images without noise, group #2 has 52 images with Gaussian noise. We randomly choose 52 images from these two image groups as our small gallery we can predict the recognition performance for different large population sizes. Table 1 shows the comparison of the predicted and actual recognition performance with rank $r = 1$. The error between them is 0.0472 which indicate that our integrated prediction model can predict ear recognition performance for large population. Table 2 shows predicted recognition performance for different sizes of large population by the small gallery of 52 objects.

4.2 Comparison with Previous Work

In our previous work [3], we use binomial model to predict the fingerprint recognition performance when rank $r = 1$. In this model the prediction problem is expressed as:

$$P(N, r, t) = \int_{t}^{\infty} \left(1 - \int_{x}^{\infty} ns(t)dt\right)^{N-r} ms(x)dx \tag{10}$$

Compared with equation (8) binomial model did not consider the distortion problem in large population. Figure 5 is the prediction error between the integrated model and the binomial model under the same small gallery size for fingerprint database. The prediction error made by integrated model is much smaller than that of the binomial model which indicate that the integrated model is suitable for the distortion problem.

5 Conclusions

We have presented an integrated model which can predict large population performance from a small gallery. This model considers the distortion problem which happens in large population. Results are shown on NIST-4 fingerprint database and 3D ear database for various sizes of small gallery and population. From the above results we can see that compared with previous approaches our model improve the prediction results and can be used to predict the large population performance. In this paper we mainly focused on the biometrics recognition system, in fact this prediction model can be used to predict other kind of object recognition system.

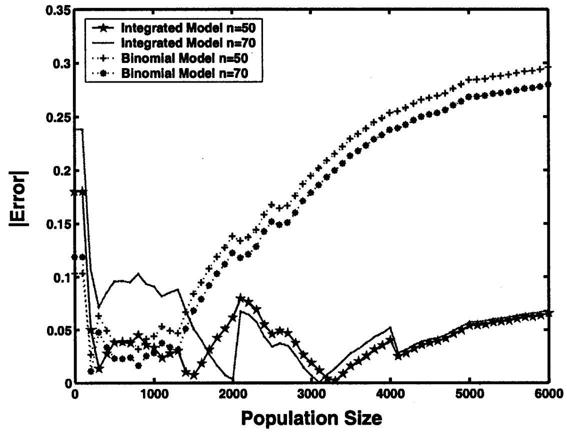


Fig. 5. Prediction error between the integrated model and the binomial model for fingerprint recognition performance

References

1. S. Theodoridis and K. Koutroumbas, "Pattern Recognition," *Academic Press*, Second Edition, 2003.
2. D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, "Handbook of Fingerprint Recognition," *Springer (New York)*, 2003.
3. B. Bhanu, R. Wang, and X. Tan, "Predicting fingerprint recognition performance from a small gallery," *ICPR Workshop on Biometrics: Challenges arising from Theory to Practice*, pp. 47-50, 2004.
4. X. Tan, and B. Bhanu, "On the fundamental performance for fingerprint matching," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.2, pp. 18-20, 2003.
5. E. S. Hong, B. Bhanu, G. Jones and X. B. Qian, "Performance modeling of vote-based object recognition," *Proceedings Radar Sensor Technology IX*, vol. 5077, pp. 157-166, August 2003.
6. A. Y. Johnson, J. Sun, and A. F. Bobick, "Predicting large population data cumulative match characteristic performance from small population data," *The 4th International Conference on AVBPA*, pp. 821-829, June 2003.
7. J. L. Wayman, "Error-rate equations for the general biometric system," *IEEE Robotics & Automation Magazine*, vol. 6, issue 1, pp. 35-48, 1999.
8. J. Daugman, "The importance of being random: statistical principles of iris recognition," *Pattern Recognition*, 36(2):279-291, February 2003.
9. P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone, "Face Recognition Vendor Test 2002," *Evaluation Report*, March 2003.
10. A. Y. Johnson, J. Sun and A. F. Boick, "Using similarity scores from a small gallery to estimate recognition performance for large galleries," *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pp. 100-103, 2003.
11. P. Grother, and P. J. Phillips, "Models of large population recognition performance," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Volume 2, pp.68-75, 2004.
12. X. Tan, and B. Bhanu, "Robust fingerprint identification," *Proc. IEEE International Conference on Image Processing (ICIP)*, vol. 1, pp. 277-280, 2002.