Learning small gallery size for prediction of recognition performance on large populations

Rong Wang, Bir Bhanu, Ninad S. Thakoor *

Center for Research in Intelligent Systems, University of California, Riverside, CA 92521, USA

ABSTRACT

This paper addresses the estimation of a small gallery size that can generate the optimal error estimate and its confidence on a large population (relative to the size of the gallery) which is one of the fundamental problems encountered in performance prediction for object recognition. It uses a generalized two-dimensional prediction model that combines a hypergeometric probability distribution model with a binomial model and also considers the data distortion problem in large populations. Learning is incorporated in the prediction process in order to find the optimal small gallery size and to improve the prediction. The Chernoff and Chebychev inequalities are used as a guide to obtain the small gallery size. During the prediction, the expectation–maximization (EM) algorithm is used to learn the match score and the non-match score distributions that are represented as a mixture of Gaussians. The optimal size of the small gallery is learned by comparing it with the sizes obtained by the statistical approaches and at the same time the upper and lower bounds for the prediction on large populations are obtained. Results for the prediction are presented for the NIST-4 fingerprint database.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Recognition systems can classify images, signals, or other types of measurements into a number of classes. In this paper, we mainly focus on biometrics recognition systems. Biometrics can be a fingerprint, a palmprint, a face image, gait, signature, speech, etc. Depending on the application there are two kinds of biometric recognition systems: verification systems and identification systems. Verification (also called authentication) is a one-to-one matching problem [1]. A verification system stores users' biometrics in a database. Then, it compares a person's biometrics signatures with the stored representation to verify if this person is indeed who she/he claims to be. The system can accept or reject the claim according to the verification result. An identification system is more complex than a verification system. In an identification system, for a given query, the system searches the entire database to find out if there are any biometrics signatures that match the query. It conducts a one-to-many matching. There are two kinds of identification systems: the closed-set identification systems and the open-set identification systems [2]. The closed-set identification is the identification for which all potential users are enrolled in the system. Alternatively, the open-set identification is the identification for which some potential users are not enrolled in the system. The verification and the closed-set identification can be considered to be special cases of the open-set identification.

In a practical recognition system, some important parameters for characterizing the system are generally unknown [3]. We need to predict these parameters from a set of available data. In this paper, we provide a prediction model for performance of a closed-set identification system. Since the recognition performance of an algorithm is usually estimated based on limited data, it is difficult to predict its performance for additional data: the limited test data may, after all, not accurately represent a larger population. Before we can evaluate and predict the performance of a recognition algorithm on large populations, we need to answer some fundamental questions. When we use a small gallery to estimate the algorithm performance on large populations, how can we find the optimal size of the small gallery and how accurate is the estimation? Since the prediction is based on the same recognition algorithm, we can give the confidence interval for the performance estimation on a large population [4]. The confidence interval [5] can describe the uncertainty associated with the estimation. This gives an interval within which the true performance of the algorithm for a large population is expected to fall, along with the probability that it is expected to fall there [6]. Bolle et al. [7] presented a bootstrap based approach to compute the confidence interval to evaluate the biometrics system performance.

In this paper, we address the problems associated with the prediction of performance on large populations and the optimal

* Corresponding author. Tel.: +1 951 827 3954.
 E-mail addresses: bhanu@cris.ucr.edu (B. Bhanu),
ninadt@ucr.edu (N.S. Thakoor).

small gallery size. The term small gallery is used to emphasize the size of the gallery used during the design of a biometric system which is small compared to the population. We use a generalized prediction model for a closed-set identification system that combines a hypergeometric probability distribution model with a binomial model. Hypergeometric distribution is a discrete probability distribution which captures the probability of picking a certain number of good samples from a mix of good and bad samples without replacement.

The prediction model takes into account distortions that may occur in large populations. When a physical phenomenon is observed and a quantity corresponding to its properties is measured, the measurement differs from the true underlaying value. This discrepancy is called the distortion. The model also provides performance measurements as a function of the rank, the large population size, the number of distorted images, and match and non-match score distributions.

We model the match score and the non-match score distributions as mixture of Gaussians and use the expectation–maximization (EM) algorithm to estimate its parameters. Given limited data, we can use parametric or nonparametric estimation methods to estimate the data distribution. The expectation–maximization (EM) algorithm [8], one of the parameter estimation methods, assumes that the underlying distribution is known. It is an iterative method to estimate the mixture parameters by maximum likelihood techniques. We introduce learning by feeding back the similarity scores (match scores and non-match scores) to increase the small gallery size. In this way, we can find the optimal size of the small gallery to predict the large population performance.

We also provide the upper and the lower bounds for the prediction performance of a large population. We use two different statistical methods—Chernoff's inequality and Chebychev's inequality—to obtain the relationship between the small gallery size and the confidence interval for a given margin of error. In probability theory, inequalities such as Chernoff's and Chebyshev's are routinely used to provide bounds on the distribution values when minimal information (e.g. mean and standard deviation for Chebyshev's) regarding the distributions is available.

The specific contributions of the paper are:

(1) We use a generalized prediction model that combines a hypergeometric probability distribution model with a binomial model which takes into account distortions that may occur in large populations. Our distortion model includes feature uncertainty, feature occlusion, and feature clutter. In the prediction model, we model the match score and non-match score distributions as a mixture of Gaussians, use the EM algorithm to estimate its parameters and find the number of components of the distributions automatically.
(2) We find the optimal size of a small gallery by an iterative learning process. We use the Chernoff inequality and the Chebychev inequality to determine the small gallery size in theory which is related to the margin of error and the confidence interval. We find the upper bound and a good lower bound on recognition performance on a large population.
(3) Systematic experimental results are shown on a challenging large data set of fingerprint images (NIST-4) with realistic distortion models.

The paper is organized as follows. Related work is presented in Section 2. The details of the technical approach are given in Section 3. It includes the distortion model, the prediction model, and the statistical methods to find the relationship between the optimal small gallery size and the confidence interval. Experimental results are provide in Section 4. The combined model with

learning is tested on the *NIST Special Database 4* (NIST-4) which is the rolled fingerprint database. Conclusions are presented in Section 5.

2. Related work

Many researchers have used statistical approaches to estimate the performance of recognition systems. Usually, these approaches use prediction models based on the feature space or similarity scores. Wayman [9] and Daugman [10] developed a binomial model that used the non-match score distribution. This model underestimates recognition performance for large populations [11]. Phillips et al. [12] developed a moment model, which used both the match score and non-match score distributions.

Pankanti et al. [13] presented a fingerprint individuality model which was based on the feature space and derived an expression to estimate the probability of false matching between two fingerprints based on minutiae. The model measured the amount of information needed to establish correspondence between two fingerprints. Tan and Bhanu [14] presented an improvement over [13] by providing a two-point model and a three-point model to estimate the error rate for the minutiae based fingerprint recognition. Their approach measured minutiae's position and orientation and the relations between different minutiae to find the probability of correspondence between fingerprints. They allowed overlap of the uncertainty area of any two minutiae.

Johnson et al. [15] improved the moment model by using a multiple non-match score set. They averaged match scores of the entire gallery. For each match score, they counted the number of non-match scores larger than the match score leading to an error. They assumed that the match scores are distributed uniformly. Grother and Phillips [11] introduced a joint density function of the match score and the non-match score to estimate both the open-set and the closed-set identification performance. Since the joint density is generally impractical to estimate, they assumed that the match score and non-match scores are independent and their distributions are the same for large populations. They used the Monte Carlo sampling method to linearly interpolate the match score and the non-match score look-up tables. Tabassi et al. [16] and Wein and Baveja [17] used the fingerprint image quality to predict the performance. They defined the quality as an indication of the degree of separation between the match score and non-match score distributions. The farther these two distributions are, the better the system performs.

Ju and Bhanu [18] predicted the gait recognition performance by probabilistic simulation of different within-class feature variance. They provided the upper bound for the recognition performance with regard to different human silhouette resolutions. Li et al. [19] developed an analytical performance characteristic to predict the misclassification statistics of the resulting boosted classifier. The analytic error characterization establishes the relationship between the misclassification statistics and the size of training set and the true distribution parameters.

Wang et al. [20] trained a support vector machine from features based on match and non-match scores to predict success and failure of the face recognition. Scheirer et al. [21] analyzed similarity surfaces to predict algorithmic failures in face recognition for various face recognition algorithms. Aggarwal et al. [22] learned mapping from the image characterization space to the score space to predict performance of face recognition algorithms on unseen data.

Usually, for a biometrics recognition system, the performance margin of error is prespecified. Consequently, providing the upper and lower bounds for the performance is another important topic in the recognition performance prediction. Lindenbaum [23]

proposed a probabilistic method to derive bounds on the number of features required to achieve successful recognition with a certain degree of confidence. This method considered object similarity, bounded uncertainty and occlusion. A similar approach presented in [24] can be used to analyze object recognition with uncertainty, similarity, and clutter. Guyon et al. [4] proposed guaranteed estimators to determine the test set size which gives statistically significant results. The analysis is done for two cases: when the recognition errors are independently identically distributed and when the errors are correlated, along with the assumption of the underlying probability distribution. Boshra and Bhanu [25] presented a method to predict upper and lower bounds on the recognition performance. They predicted performance by considering feature uncertainty, occlusion, clutter, and similarity simultaneously. In their method, performance is predicted in two steps: first, compute the similarity between each pair of models and then use the similarity information along with the statistical model to determine upper and lower bounds for the object recognition performance. Dass et al. [26] proposed a technique which is based on parametric copula model to estimate the test sample size for a predetermined ROC width.

In [27], we presented a binomial model to predict the large fingerprint database recognition performance based on a small gallery. In that work, the probability of false alarm is related with the match score which is more reliable for the prediction of recognition performance. In [28], we presented our early work on a generalized two-dimensional model which integrated a hypergeometric probability distribution explicitly with a binomial distribution. It considered the distortion caused by sensor noise, feature uncertainty, feature occlusion and feature clutter. Our paper [29] introduced the learning of

gallery size within the prediction framework as well as theoretical bounds on the performance to validate the estimated bounds. Additionally, underlying score distributions were modeled as a mixture of Gaussians making them more flexible.

This paper is related to the shorter version of the work in [29] which lacked in experiments. Distortion parameters used in [29] were set arbitrarily and the validation for optimal gallery size was conducted by resubstitution. In this paper, the entire set of experiments is redone where unlike [29], distortion parameters are learned from the data and the prediction is carried out with unseen data. As a result the new results now reflect a practical scenario in the real world.

We list the above approaches and our specific learning approach proposed in this paper in Table 1.

3. Technical approach

While designing a biometric recognition system, two sets of biometrics data are provided: a gallery and probes. The gallery is a set of biometrics templates saved in the database. The probes are a set of queries for the database. A large population is the unknown data set for which the recognition performance of the biometric system needs to be estimated. Based on the given gallery and probes, we would like to estimate the recognition performance of the biometric system on the large population.

Fig. 1 provides the conceptual diagram for the performance prediction system. Fig. 1(a) shows the process of learning the optimal small gallery size. First, we randomly select a small (compared to population) gallery and its corresponding probes from the entire

Table 1
Performance prediction approaches.

| Authors	Application	Approach	Comments
Lindenbaum [23]	Object recognition	Statistical model	Bounds on the number of features were derived considering the object similarity, uncertainty, and occlusion
Lindenbaum [24]	Object recognition	Statistical model	Approach similar to [23], considered the object similarity, uncertainty, and clutter
Guyon et al. [4]	Character recognition	Statistical model	Size of test set which gives statistically significant results was determined using statistical arguments
Wayman [9]	Fingerprint	Binomial model	Non-match score distribution was used to derive error rate equations
Boshra and Bhanu [25]	Object recognition	Statistical model	Similarity information was used along with the statistical model combining uncertainty, occlusion, and clutter
Pankanti et al. [13]	Fingerprint	Feature space	Probability of false match was estimated based on minutiae position and orientation
Daugman [10]	Iris	Binomial model	Non-match score distribution was used; underestimates the performance for large population size
Phillips et al. [12]	Face	Binomial model	Assumed match score and the non-match score are sampled independently
Johnson et al. [15]	Gait	Binomial model	Multiple non-match score sets were used under the assumption that the match score has a uniform distribution
Tan and Bhanu [14]	Fingerprint	Feature space	Probability of error rate was estimated based on minutiae's position, orientation, and relation
Grother and Phillips [11]	Face	Binomial model	The probability of error was related with the match score
Tabassi et al. [16], Wein and Baveja [17]	Fingerprint	Image quality	Fingerprint image quality was used as an indication for the performance prediction
Ju and Bhanu [18]	Gait	Statistical model	Prediction was based on a simulation approach with different within-class feature variance and body parts data
Li et al. [19]	Face	Statistical model	Error characteristics were obtained analytically which related the error with the size of the training set
Dass et al. [26]	Fingerprint	Copula model	Estimated the test sample size for a predetermined ROC width
Wang et al. [20]	Face	SVM	Features based on scores were used to train SVM to detect recognition failure
Scheirer et al. [21]	Face	SVM	Features derived from similarity scores were used in analysis of similarity surfaces
Aggarwal et al. [22]	Face	Multi-dimensional scaling	Mapping from image characterization space to score space was used
Wang et al. [28], Wang and Bhanu [29,27]	Fingerprint [27–29] and ear [28]	Binomial [27–29] and hypergeometric [28,29] model	[27] presented model to predict the large population recognition performance based on a small gallery. Probability of false alarm was related with the match score. In [28], distortions including uncertainty, occlusion, and clutter were modeled for the performance prediction on large populations. The optimal small gallery size and bounds on performance were first estimated in [29]. However, the experiments were conducted with arbitrary distortion parameters and validated with re-substitution
This paper	Fingerprint	Binomial & hypergeometric model	Distortion in large populations is considered and optimal small gallery size is estimated. Performance bounds and confidence in prediction are given

a

b
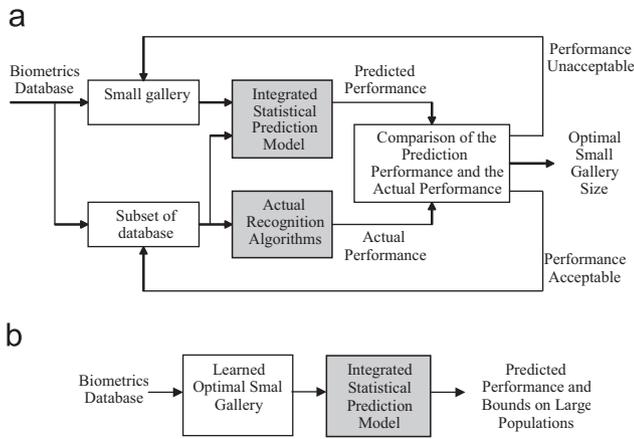
Fig. 1. Conceptual performance prediction system: (a) learning the optimal small gallery size; (b) performance prediction on large populations.

Fig. 2. Our proposed approach for optimal small gallery size selection.

Fig. 3. Our integrated statistical prediction model combined with the distortion model.

database. Then we select a subset of remaining dataset which will be treated as large population. We apply the integrated statistical prediction model proposed in this paper to predict the recognition performance for the selected large population with the selected small sample set (i.e., the small gallery and corresponding probes).

Meanwhile, we use the recognition algorithm to obtain the actual performance for the selected population. We compare the predicted performance and the actual performance. If the predicted performance is acceptable, then we increase the size of the population and repeat the above process, else we increase the size of the small gallery and repeat the above process. When the size of the population cannot be increased, we stop learning and get the optimal small gallery size.

We randomly choose optimal size small galleries to predict the performance by the integrated statistical prediction model. We obtain the performance bounds based on the performance predicted by these small galleries. This process is depicted in Fig. 1(b).

3.1. Methodology for determining the optimal small gallery size

Fig. 2 provides the detailed diagram for the implementation of our approach to predict the optimal small gallery size. For a given biometrics recognition system with data Q of N enrollees, we randomly pick n enrollees from Q to form a small sample set q. By authentication, we can get a set of match scores and non-match scores for this small gallery. Then, we use the expectation–maximization (EM) algorithm to estimate the parameters of the mixture of Gaussians distributions of the match score and non-match scores. Based on these distributions, we use our prediction model, which combines a hypergeometric probability distribution model with a binomial model, to estimate the recognition system performance for a large population Q_1 with N_1 enrollees which is a subset of $Q \backslash q$. We assume that the predicted performance on Q_1 is $\hat{p}$. From the recognition system, we can obtain the match scores and the non-match scores for Q_1 and compute the actual recognition performance p for Q_1. $\tilde{e}$ is the error between the predicted performance and the actual performance, i.e., $\tilde{e} = |\hat{p} - p|$. The margin of error e is the maximum specified error acceptable by the recognition system. If $\tilde{e}$ is larger than the margin of error e then we increase the small gallery size n and feed back match scores and non-match scores to the EM algorithm to estimate the parameters of the similarity score distributions again. Otherwise, we increase the size of the large population Q_1, and repeat this process until Q_1 includes the entire $Q \backslash q$. We use the Chernoff and Chebychev inequalities to find the relationship between the small gallery size and the prediction confidence interval for a given margin of error.
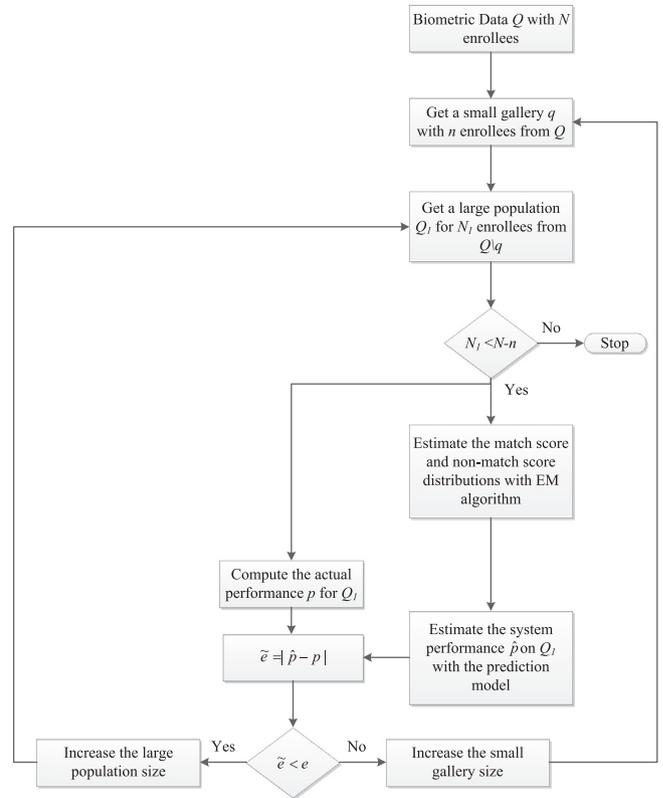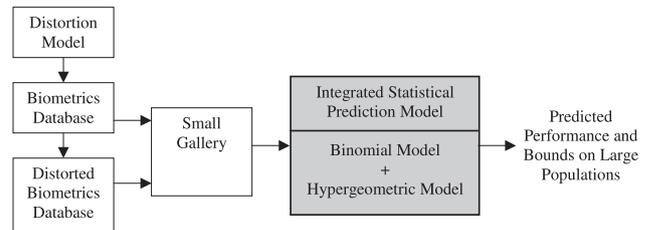
The small gallery size which we got from the inequalities is used to validate the learned optimal small gallery size. We will explain each part of the diagram in detail in rest of this section.

3.2. Integrated statistical prediction model

Fig. 3 shows our integrated statistical prediction model combined with the distortion model to predict the large population performance. Our model integrates three different distributions: mixture of Gaussians to model the match and non-match scores distributions, hypergeometric distribution to model a mix of distorted and undistorted images, and binomial distribution to model the ranking process in the retrieval. The probability of selecting "good" samples (undistorted images) from a mix of "good" (undistorted images) and "bad" (distorted images) samples is given by a hypergeometric distribution. The combinatorial nature of ranking process leads to a binomial distribution.

Our two-dimensional prediction model considers the distortions that occur in large populations. Assume we have two kinds of biometrics images, group #1 and group #2 which differ in quality.

Group #1 is a set of biometrics images without any distortion. Group #2 is a set of biometrics images with distortions. Let the size of these two groups be n_1 and n_2 enrollees respectively. We randomly pick n enrollees from group #1 and group #2 to be our small sample set q. Then, the number of enrollees with distorted images y which are chosen from group #2 follows a hypergeometric distribution

$$f(y) = \frac{C_{n-y}^{n_1} C_y^{n_2}}{C_n^{n_1 + n_2}} \qquad (1)$$

where $n_1 + n_2$ is the total number of enrollees in these two groups (which is same as the large population) and $n-y$ is the number of enrollees chosen from group #1.

For each of the n enrollees in the small sample set q, a pair of images is available, one belonging to the gallery and another a probe. We combine them into the small gallery and corresponding probe set. For each image in the probe set, we compute the similarity scores with the images in the gallery. We have one match score and $n-1$ non-match scores for this image.

With all the similarity scores, we can use the EM algorithm to estimate the parameters of match score and non-match score distributions. From the above discussion, we know that the match score and non-match score distributions depend not only on the scores but also on the number of images with distortion. Let $ms(x|y)$ and $ns(x|y)$ represent the distributions of match scores and non-match scores given the number of distorted images. If the similarity score is higher, then the biometrics signatures are more similar. The error occurs when a given match score is smaller than one or more non-match scores corresponding to the same image. 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

$$NS(x) = \int_x^\infty \sum_{y=0}^n ns(t|y) f(y) \, dt \qquad (2)$$

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

If the size of the large population is N, then for a probe image we can have one match score and $N-1$ non-match scores. We rank the match score and non-match scores in the descending order. For a given number of images with distortion, the probability that the match score x is at rank r is given by the binomial probability distribution

$$\binom{N-1}{r-1} (1-NS(x))^{N-r} (NS(x))^{r-1} \qquad (3)$$

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

$$\int_{-\infty}^\infty \binom{N-1}{r-1} (1-NS(x))^{N-r} (NS(x))^{r-1} ms(x|y) \, dx \qquad (4)$$

Since the large population has one match score and $N-1$ non-match scores, we use $\binom{N-1}{r-1}$ instead of $\binom{N}{r-1}$. By summing over the images chosen from group #2, the probability that the match score is at rank r can be written as

$$\int_{-\infty}^\infty \binom{N-1}{r-1} (1-NS(x))^{N-r} (NS(x))^{r-1} \sum_{y=0}^n ms(x|y) f(y) \, dx \qquad (5)$$

In theory, a match score can be any value within $(-\infty, \infty)$. The probability that the match score is within rank r is

$$P(N,r) = \sum_{i=1}^r \int_{-\infty}^\infty \binom{N-1}{r-1} (1-NS(x))^{N-i} (NS(x))^{i-1} \sum_{y=0}^n ms(x|y) f(y) \, dx \qquad (6)$$

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

$$P(N,r,t) = \sum_{i=1}^r \int_t^\infty \binom{N-1}{r-1} (1-NS(x))^{N-i} (NS(x))^{i-1} \sum_{y=0}^n ms(x|y) f(y) \, dx \qquad (7)$$

When rank $r=1$ the prediction model with threshold t becomes

$$P(N,1,t) = \int_t^\infty (1-NS(x))^{N-1} \sum_{y=0}^n ms(x|y) f(y) \, dx \qquad (8)$$

In this model, we make two assumptions: match scores and non-match scores are independent and large populations have distortions which can be modeled with feature uncertainty, occlusion, and clutter. We use a small gallery to estimate distributions of $ms(x|y)$ and $ns(x|y)$. These distributions are assumed to be mixtures of Gaussians and the parameters for the mixture are estimated using the EM algorithm.

3.2.1. Distortion model

Usually a biometrics recognition system consists of three stages: image acquisition, feature extraction, and matching. Distortion often occurs at these stages and may be caused by sensor noise, feature uncertainty, feature occlusion, and feature clutter. The effects of sensor and image noise are reflected in the feature uncertainty. Performance estimates derived from a small gallery might fail to capture range of distortions that are present in the large population. Thus, we choose to model the distortion explicitly. Our distortion model includes feature uncertainty, occlusion, and clutter. Assume $F = \{f_1, f_2, \ldots, f_k\}$ is the feature set of the biometrics image under consideration, where $f_i = (col, row, o)$, col and row represent the feature's location, o represents the feature's other attributes excluding the location, $i = 1, 2, \ldots, k$. The distortions are modeled as the following:

(a) *Uncertainty*: The uncertainty arises due to perturbation of the true positions of features during acquisition, digitization, preprocessing, etc., of the image. Assume that the uncertainty is uniformly distributed. It represents how likely each feature is to be perturbed. We replace each feature $f_i = (col, row, o)$ with $f'_i = (col', row', o')$ such that

$col - \Delta x \leq col' \leq col + \Delta x$,
$row - \Delta y \leq row' \leq row + \Delta y$,
$(1-\varepsilon)o \leq o' \leq (1+\varepsilon)o$

where ε is a measure of uncertainty in a feature's other attributes, $0 \leq \varepsilon \leq 1$. Δx and Δy are the uncertainty area which are measured in pixels. All uncertainty parameters ε, Δx, and Δy are assumed to be distributed uniformly.

(b) *Occlusion*: The occlusion causes elimination of some features. Assume that the number of features occluded is OC. By assuming each feature is independently and equally likely to be occluded, we choose OC features out of the k features and remove them.

(c) *Clutter*: The clutter in features is caused by spurious features which are incorrectly detected. We add CL additional features, where each feature is generated by picking a feature according to the clutter distribution from the *clutter region* (CR) (the region where clutter features are added). The clutter

PDF determines the distribution of the clutter over the clutter region. The clutter region depends upon the given model to be distorted. We use a bounding box to define the clutter region

$$CR = \{(col, row, o), col_{min} \leq col \leq col_{max},$$
$$row_{min} \leq row \leq row_{max}, o_{min} \leq o \leq o_{max}\}$$

where col_{min} and col_{max} represent the minimum and maximum value of col. Similarly row_{min}, row_{max}, o_{min}, and o_{max} represent the minimum and maximum value of row and o respectively. Features values are picked uniformly from the clutter region.

The distortion model is applied to small gallery data before it is used in the prediction model. Thus, estimates given by the prediction model are affected by the distortion model.

3.3. Estimation of the small gallery size based on statistical inequalities

In this section, we discuss the relationship between the prediction confidence interval and the size of the small gallery which could be used to validate the optimal small gallery size that we obtain through the learning process. We use limited data to estimate a large population recognition performance. Therefore, the prediction value may or may not be accurate enough. We assume that the risk of being wrong is $1-\alpha$. The risk should be equal or greater than the probability that the error between the predicted performance and the actual performance is greater than or equal to the margin of error e of this system

$$\Pr\{|(p-\hat{p})| \geq e\} \leq (1-\alpha) \tag{9}$$

where $\hat{p}$ is the predicted performance for the recognition system which can be obtained from our prediction model, p is the actual performance of the recognition system, and α is the confidence. Since $\Pr\{|(p-\hat{p})| \geq e\} = \Pr\{p \geq \hat{p} + e\} + \Pr\{p \leq \hat{p}-e\}$, $\Pr\{p \geq \hat{p} + e\} \geq 0$, and $\Pr\{p \leq \hat{p}-e\} \geq 0$, inequality (9) can be written as

$$\Pr\{p \geq \hat{p} + e\} \leq (1-\alpha) \tag{10}$$

or

$$\Pr\{p \leq \hat{p}-e\} \leq (1-\alpha) \tag{11}$$

Here, we will solve inequality (10). Inequality (11) can be solved similarly.

We assume that the system recognizes (authenticates) individuals with the probability $\Pr\{X_i = 1\} = t$ and $\Pr\{X_i = 0\} = 1-t$, where $X_i = 1$ means an individual with a given biometrics X_i is recognized correctly, $X_i = 0$ means the opposite, $0 \leq t \leq 1$. According to the Chernoff inequality [30], let $X_1, X_2, \ldots, X_n$ be independent random variables. We define a random variable

$$X = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{12}$$

For any $\epsilon \geq 0$, we have

$$\Pr\left\{X \geq E(X) + \frac{\epsilon}{n}\right\} \leq e^{-2\epsilon^2/n} \tag{13}$$

where $E(X)$ is the mean of X, $E(X) = \hat{p}$. Comparing with inequality (10), we can get

$$1-\alpha = e^{-2\epsilon^2/n} \tag{14}$$

So

$$\epsilon = \sqrt{-\frac{n \ln(1-\alpha)}{2}} \tag{15}$$

Thus, Eq. (13) becomes

$$\Pr\left\{X \geq E(X) + \sqrt{-\frac{\ln(1-\alpha)}{2n}}\right\} \leq 1-\alpha \tag{16}$$

From inequality (10), we know that

$$e = \sqrt{-\frac{\ln(1-\alpha)}{2n}} \tag{17}$$

Thus, we get

$$n = \frac{-\ln(1-\alpha)}{2e^2} \tag{18}$$

Eq. (18) is the relationship between the small gallery size and the confidence interval under the given margin of error for the system with the underlying distribution.

Above, we assume that the system can recognize biometrics with a certain distribution. If we do not know the underlying distribution of the recognition system, then we can use the Chebychev inequality [30] which is distribution independent. Assume $X_1, X_2, \ldots, X_n$ are independent random variables. We define X as

$$X = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{19}$$

For any $\epsilon \geq 0$, we have

$$\Pr\{X - E(X) \geq \epsilon\} \leq \frac{\sigma^2}{2n\epsilon^2} \tag{20}$$

where σ^2 is the variance of X. Comparing with Eq. (10), we have

$$1-\alpha = \frac{\sigma^2}{2n\epsilon^2} \tag{21}$$

From the above equation, we obtain

$$\epsilon = \frac{\sigma}{\sqrt{2n(1-\alpha)}} \tag{22}$$

From Eqs. (20), (21) and (22), we have

$$\Pr\left\{X \geq E(X) + \frac{\sigma}{\sqrt{2n(1-\alpha)}}\right\} \leq (1-\alpha) \tag{23}$$

Then

$$e = \frac{\sigma}{\sqrt{2n(1-\alpha)}} \tag{24}$$

So we have

$$n = \frac{\sigma^2}{2(1-\alpha)e^2} \tag{25}$$

From Eq. (25), we obtain the relationship between the small gallery size and the confidence interval under the given margin of error for the system without the assumption of the underlying distribution. It is known that the Chernoff inequality is much tighter than the Chebychev inequality and the Chebychev inequality is distribution independent [31].

In this subsection, we provide a statistical estimation of the small gallery size. Meanwhile, we learn the similarity score distribution to find the optimal size of the small gallery. The small gallery size which we get from the statistics can be used as a guide for learning. Under the assumptions that the randomly chosen small galleries can represent the distributions of similarity scores for other galleries of the same size, we use different small galleries with the learned optimal size to predict large population performance. We randomly choose several small galleries of the optimal size to predict the large population performance. Then, we obtain the maximum and minimum prediction performance on the large population. In this way, we can provide an upper bound and
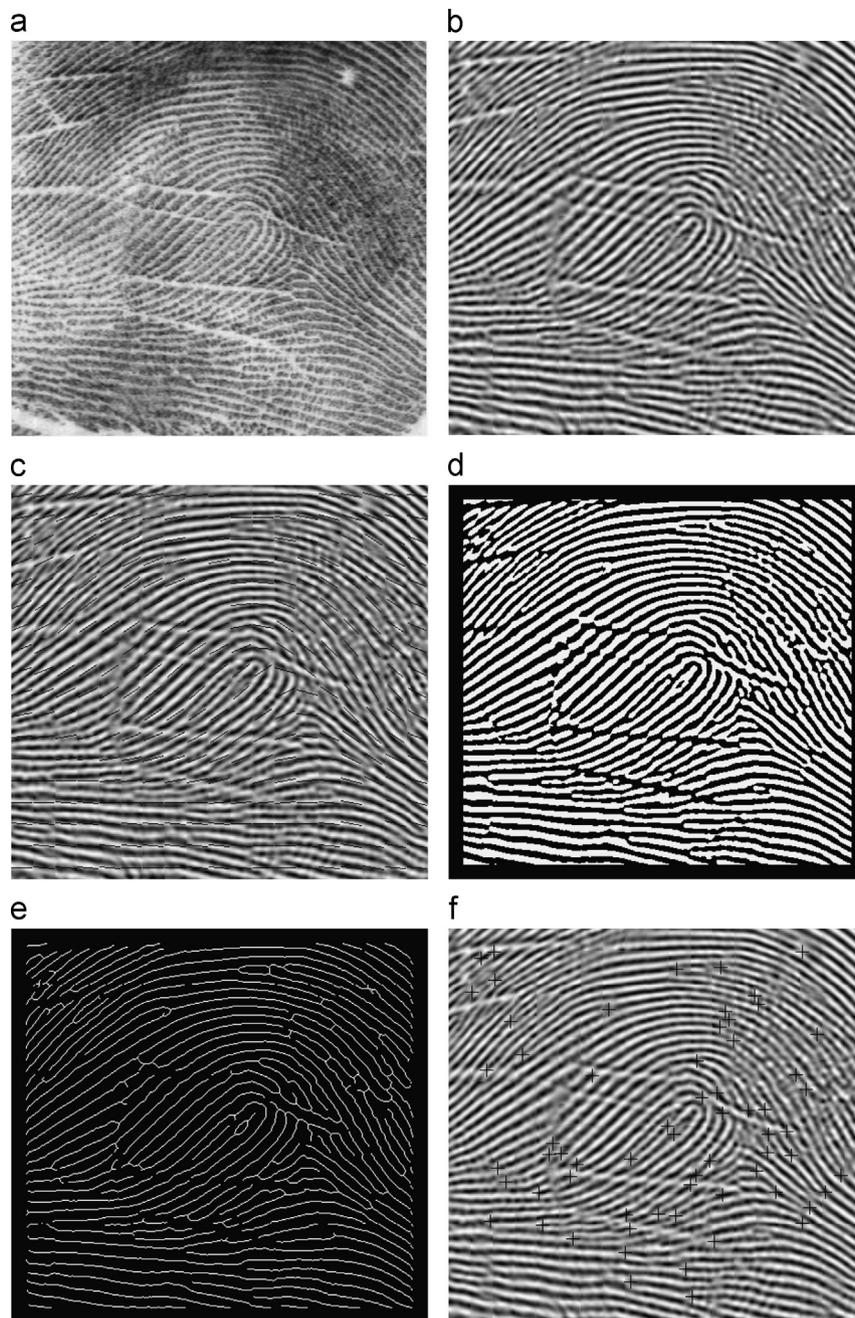
Fig. 4. An example of feature extraction process: (a) original image; (b) image after background removal; (c) smoothed image with local orientation; (d) binarized image; (e) thinned image; (f) feature image, features (endpoints and bifurcations) are marked with '+'s.

a good lower bound for performance prediction on large populations.

3.4. Actual recognition algorithms

3.4.1. Feature extraction

In fingerprint recognition systems minutiae are commonly used as features which are endpoints and bifurcations in the fingerprint ridges [32]. We use a template based approach for minutiae extraction [33], which is based on learning templates by the Lagrange method. Templates for endpoints and bifurcations are learned in the off-line step. During the run time, the learned templates are used adaptively to extract minutiae from fingerprint images. It consists of the following key steps: remove background, compute local orientation, smooth image adaptively, binarize and thin image adaptively, and extract minutiae.

3.4.2. Fingerprint matching

The fingerprint matching algorithm we used is based on the representation of fingerprint minutiae by triangles [34]. For every fingerprint, we first extract minutiae. Then randomly choose any three noncolinear minutiae to form a triangle. Thus one fingerprint can yield thousands of triangles. The features we used to find the corresponding triangles in two fingerprints are: minimum angle α_{min}, median angle α_{med}, triangle handedness ϕ, triangle direction η, maximum side λ, minutiae density χ and ridge counts ξ. The detail of these features can be found in [35,36]. We estimate the
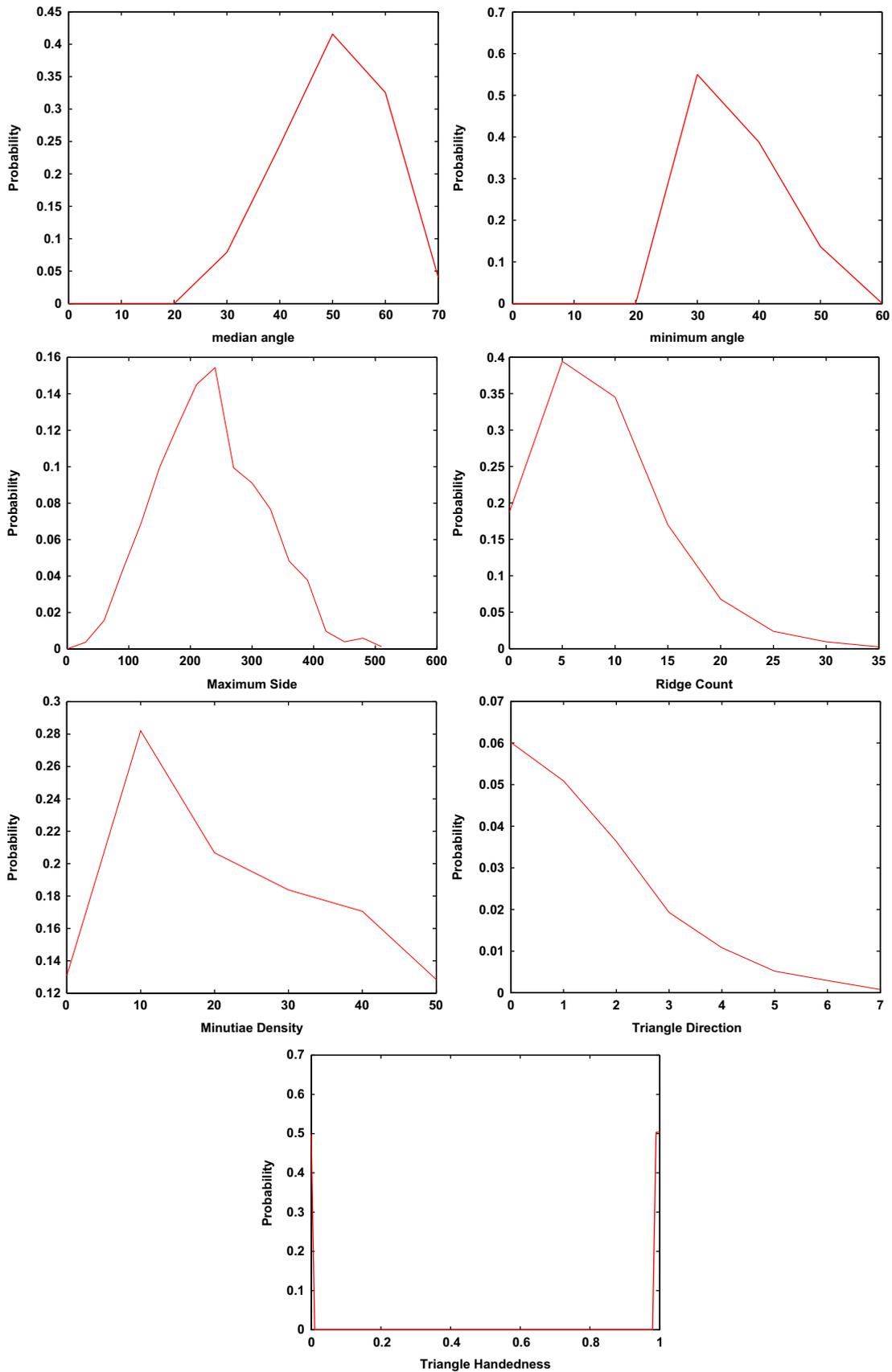
Fig. 5. Fingerprint features distributions obtained from the NIST-4 image: median angle α_{med} (in degree), minimum angle α_{min} (in degree), maximum side λ (in pixel), ridge count ξ (a number), minutiae density χ (a number), triangle direction η (a number), and triangle handedness ϕ (a number, clockwise=0, counterclockwise=1).
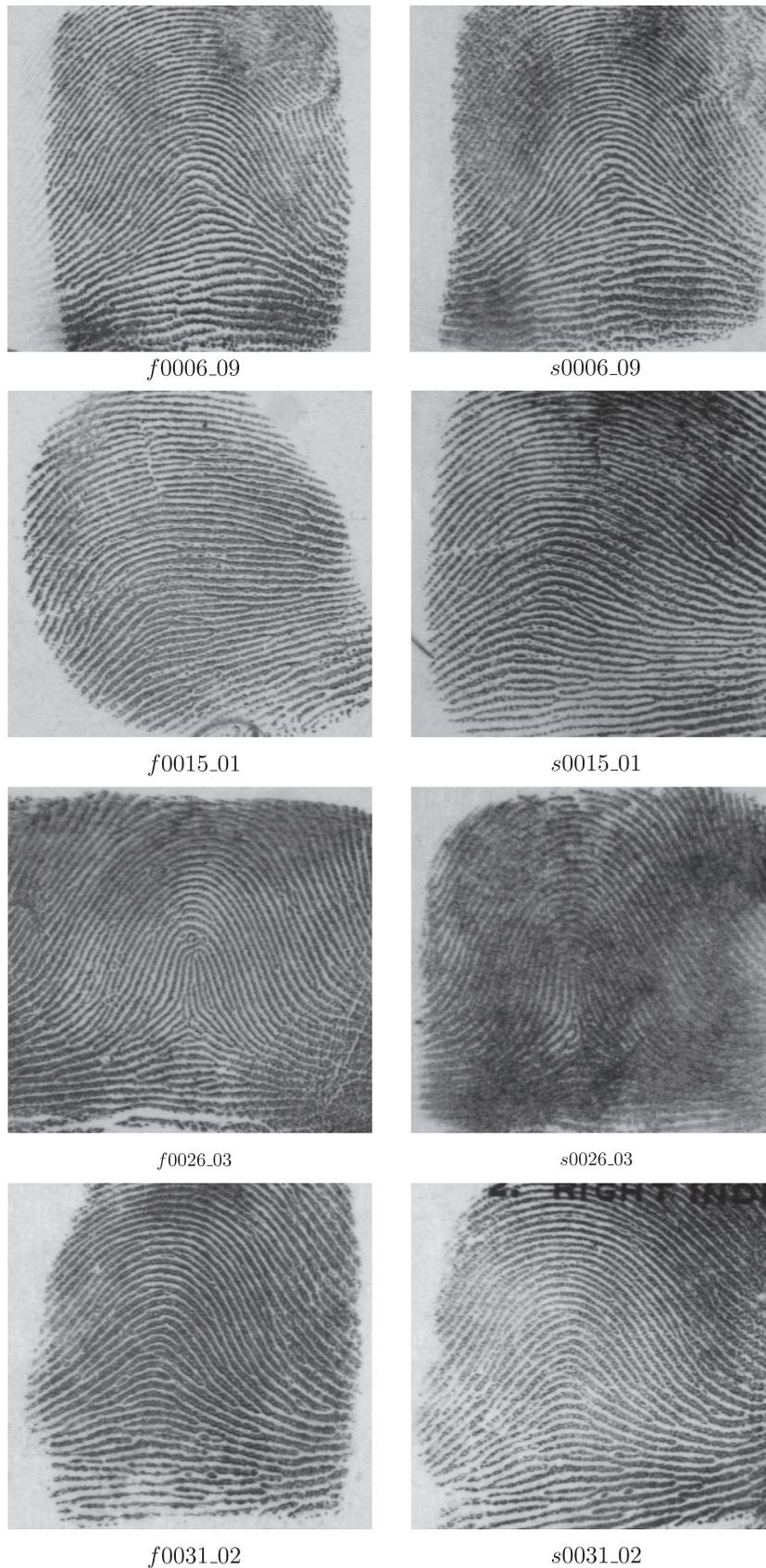
Fig. 6. Sample images from NIST-4 ('f' and 's' represents different impressions of the same fingerprint).

transformation parameters by minimizing the sum of the squared distances between the transformed query fingerprint points and their corresponding template points. Finally, we compute the distance between the query points and the transformed models points to find the corresponding triangles between the model and query fingerprints.

R. Wang et al. / Pattern Recognition 46 (2013) 3533–3547

4. Experimental results

The experiments were conducted with the *NIST Special Database 4* (NIST-4) [34] consisting of 2000 pairs of fingerprints. The experiments were carried out using programs written in C on a Sun UltraSPARC II computer. Compared to [29], all the experiments are redone by learning the *distortion parameters* from the data and by carrying out prediction on *unseen* data to reflect a practical scenario in the real world.

4.1. Actual recognition algorithms

4.1.1. Feature extraction

We use the template based approach which is described in Section 3 to extract minutiae. An example of the feature extraction process is shown in Fig. 4.

4.1.2. Fingerprint matching

For a typical fingerprint from the NIST-4, there are 78 minutiae features and 8862 triangles for which the minimum angles (α_{min}) are greater than $5°$ [34]. The distributions of these features (α_{min}, α_{med}, ϕ, η, λ, χ, ξ) are shown in Fig. 5. Fig. 6 shows four pairs of sample fingerprints from the NIST-4. Their matching results are shown in Table 2. In Table 2, the values on the diagonal are match scores, off diagonal values are non-match scores. The similarity score is the number of triangles that match between the two images. For the correct recognition, the match score should be greater than the non-match score. For the fingerprint $s0026\_03$, the match score is 0, while the non-match score between $s0026\_03$ and $f0006\_09$ is 3, obviously this is not correct. Note from Fig. 6 the quality of $s0026\_03$ is not good. It could not find any corresponding triangle with $f0026\_03$, while it found three corresponding triangles with $f0006\_09$.

Table 2
An example of similarity scores for sample test images matched with database images.

Database	$f0006\_09$	$f0015\_01$	$f0026\_03$	$f0031\_02$
Test				
$s0006\_09$	719	0	0	4
$s0015\_01$	0	106	0	0
$s0026\_03$	3	0	0	0
$s0031\_02$	0	0	0	810

4.2. Integrated statistical prediction model

4.2.1. Distorted data

Since distortions present in large populations might not be reflected in the small gallery, we simulate the distortions in our prediction model to estimate the recognition performance based on small galleries. We randomly choose 300 pairs of fingerprints to learn the probability of minutiae to be distorted. The transformation parameters which include scale (s), rotation (θ), translation (t_x, t_y) are learned in [37]. The range of these parameters are: $0.9 \le s \le 1.1$, $-30° \le \theta \le 30°$, $-128 \le t_x \le 128$, $-128 \le t_y \le 128$. We apply the transformation parameters to the images denoted with 'f'. We denote the transformed images as '$\overline{f}$', the number of minutiae as M_f. We compare the transformed image with the corresponding image which is denoted with 's' to get the number of occluded minutiae OC. We also get the number of cluttered minutiae CL. For the 300 pair of images, we get the mean and variance of occlusion and clutter probabilities. We repeat the above process five times and the histograms of the probabilities of occlusion and clutter are shown in Fig. 7.

The minutiae features used for the fingerprint recognition can be expressed as $f = (col, row, class)$, where col and row are the locations of a minutiae, $class$ is the class of the minutiae which represents whether the minutiae is endpoint (0) or bifurcation (1). We add the distortion to the 2000 pairs of fingerprints. For each fingerprint, we randomly generate the probabilities of occlusion, and clutter according to their mean and variance which are shown in Fig. 7. Using these probabilities, we randomly choose OC minutiae to occlude and CL minutiae to add. Finally, we randomly distort minutiae locations and add the uncertainty uniformly to these. The uncertainty region is chosen as

$$\{(col', row', class'), col-6 \le col' \le col + 6, row-6 \\ \le row' \le row + 6, class' \in \{class, 1-class\}\}$$

Fig. 8 shows examples of the simulated distortion along with the original data.

4.2.2. Prediction results

We randomly choose 50 pairs of fingerprints from two kinds of fingerprint pairs (with and without distortion) as our small gallery following a hypergeometric distribution. For this small gallery, we get 50 match scores and 2450 non-match scores. After we obtain these similarity scores, we use the EM algorithm [38] to estimate the parameters of the match score distribution and the non-match score distribution. The EM algorithm used [38] can find the number of components automatically and for each component the EM algorithm finds its mean, covariance, and weight. In this paper, the similarity scores are the number of matched triangles between two fingerprints, the match scores are positive integers
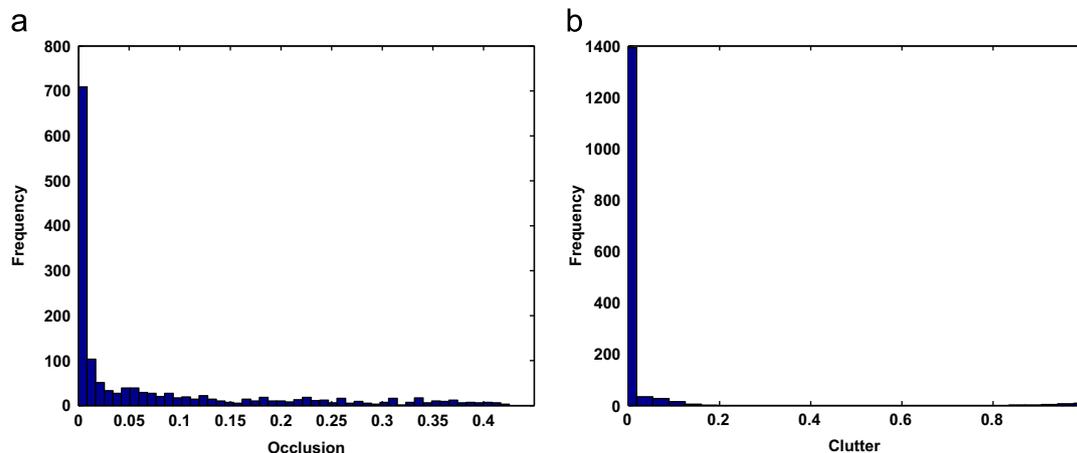
Fig. 7. Histogram of the probabilities of occlusion and clutter. (a) $\mu = 0.073$, $\sigma^2 = 0.012$. (b) $\mu = 0.020$, $\sigma^2 = 0.015$.

and the non-match scores are nearly 0. Table 3 shows the estimation of the match score distribution when the small gallery size is 50. The distributions are represented by the Gaussian mixture model. For each component, we have its mean, covariance matrix, and weight. Fig. 9 shows the match score distribution curve with the small gallery size $n=50$.

By applying the prediction model, we can estimate the fingerprints recognition performance on 2000 pairs of fingerprints based on these 50 pairs of fingerprints. We repeat the experiment seven times by randomly choosing the 50 sample images. Then, we average the results to obtain the prediction performance which is shown in Fig. 10. Here, we choose the subset size $N_1 = 100$ and
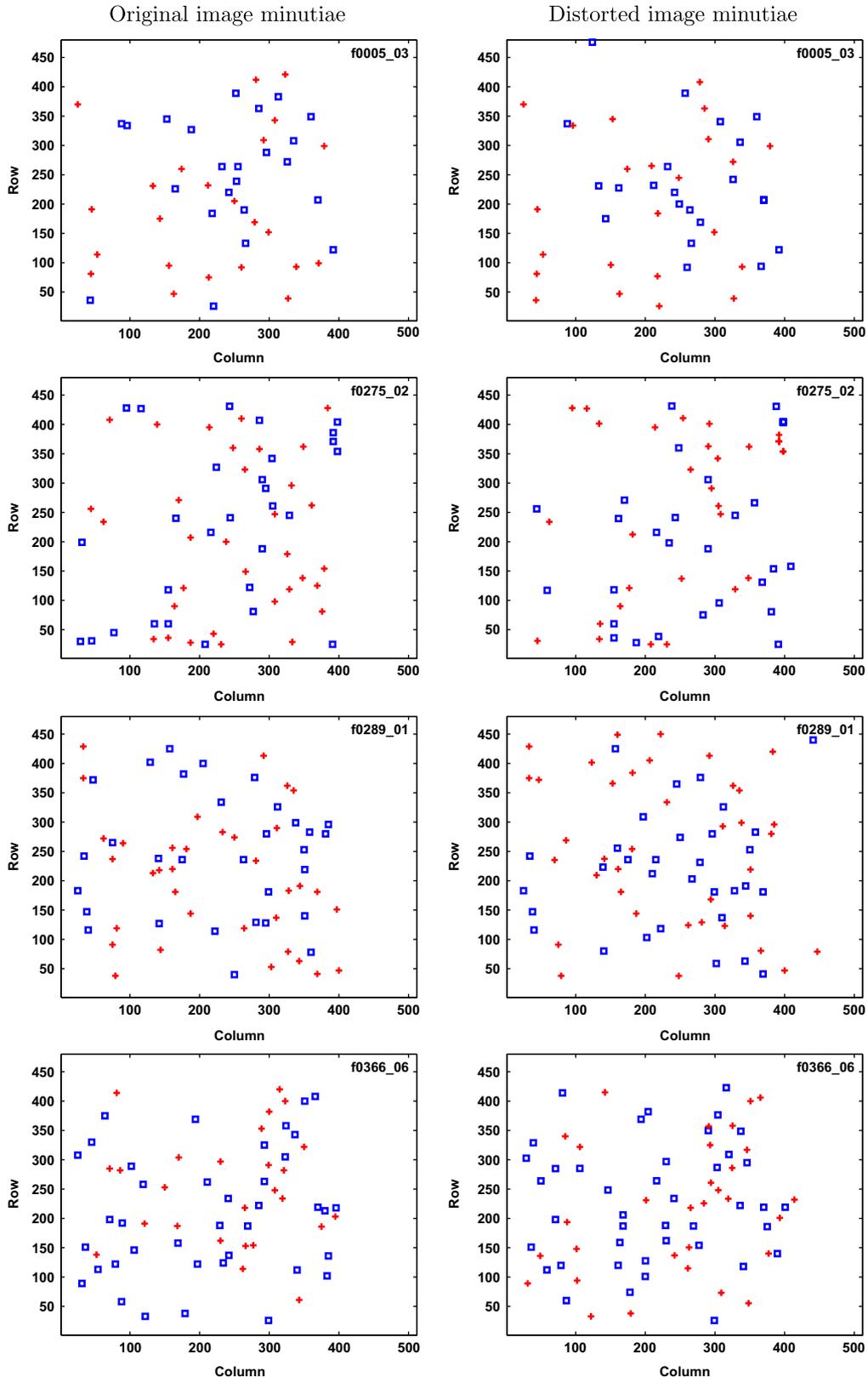

Fig. 8. Examples of the minutiae in the original images and the distorted images (□: endpoint, +: bifurcation).

the margin of error $e=0.06$. From this curve, we can see that for the large population size 100, the error between the prediction performance and the actual performance is 0.159 which is larger than the margin of error.

Now, we apply learning to the prediction process. We increase the small gallery size to $n=100$. We feed back the match scores and the non-match scores from the randomly selected 100 pairs of fingerprints and repeat this process seven times. When the large population size is 100, the absolute error between the prediction performance and the actual performance is 0.139 which is greater than the margin of error 0.06. So, we increase the small gallery size to $n=200$ and repeat the same process. The absolute error is 0.078 when the large population size is 100. Then, we increase the small gallery size to $n=300$ and repeat the same process. The absolute error is 0.022 when the large population size is 100. We increase the large population size in steps of 100 until the large population size $N=2000$. For these three small galleries, most of the non-match scores are 0. Table 4 shows the estimation of the match score distributions with different small gallery sizes. The distributions are represented by the Gaussian mixture model. For each component we have its mean, covariance, and weight. Fig. 11 shows the match score distribution curves on different small gallery sizes. For each small gallery size we provide histogram and distribution estimated by the EM algorithm. We can see that the distributions conform with the histograms very well. Fig. 12 shows the absolute error between the prediction and the actual performance decreases when the gallery size increases. When the small gallery size $n=300$, the absolute error for the large population is smaller than the margin of error 0.06. At this point, we can stop learning the small gallery size.

Since the prediction model is used to predict the recognition system performance on the unseen data, we use 500, 1000, and 1500 pairs of fingerprints as our small gallery to estimate the performance for the other 1500, 1000, and 500 fingerprints. We repeat these experiments for seven times and get the average absolute error between the actual performance and the prediction performance which is shown in Fig. 13. From Fig. 13, we can see

that when the large population sizes are 500, 1000, and 1500, the absolute errors between the prediction performance and the actual performance are 0.020, 0.023, and 0.028. The reason that the absolute errors increase with the population size enlarged is that the small gallery size decreases as the large population size increases due to limited experimental data. This conforms with the conclusion we obtained from Fig. 12. From Fig. 12, we can see that when the small gallery size is 300, the absolute errors between the

Table 4
Match score distribution parameters estimated by the EM algorithm for three different small gallery size.

Small gallery size	Component #	Mean	Variance	Weight
100	4	12.716639	162.708165	0.589401
		86.767241	605.757816	0.195997
		282.570467	8906.413460	0.168444
		808.569072	27 579.158267	0.046110
200	3	12.295530	165.625889	0.580879
		85.899347	711.041188	0.158269
		367.882901	58 819.934339	0.260768
300	6	18.705816	125.831947	0.260355
		78.448414	627.877959	0.173790
		198.259075	2180.769105	0.118167
		1.184599	2.518164	0.255349
		351.612984	3496.592673	0.081996
		595.801606	54 014.280937	0.110343

Fig. 10. Absolute error between the prediction and the actual performance when the small gallery size $n=50$, $N_1=100$, $N=2000$.

Table 3
Match score distribution parameters estimated by the EM algorithm for small gallery size $n=50$.

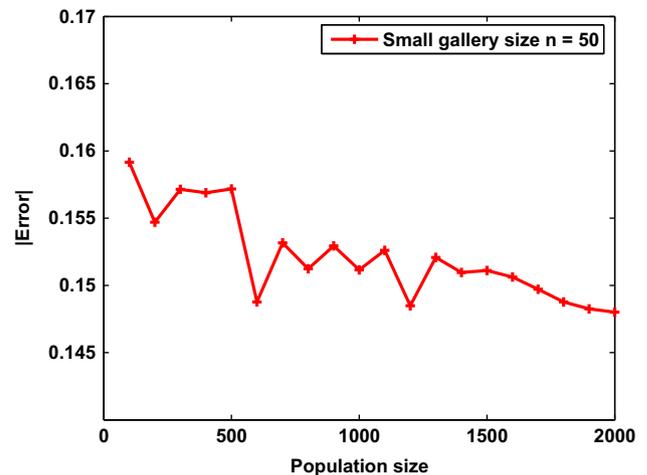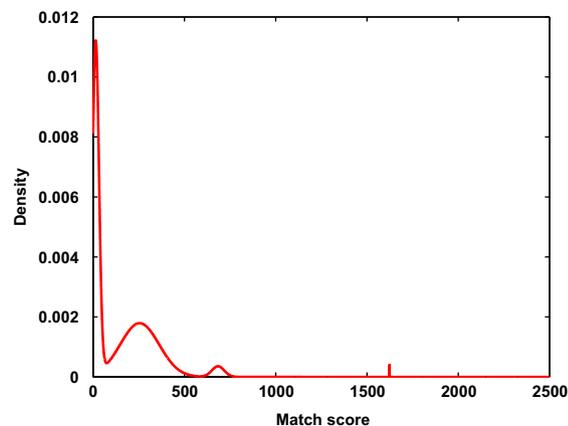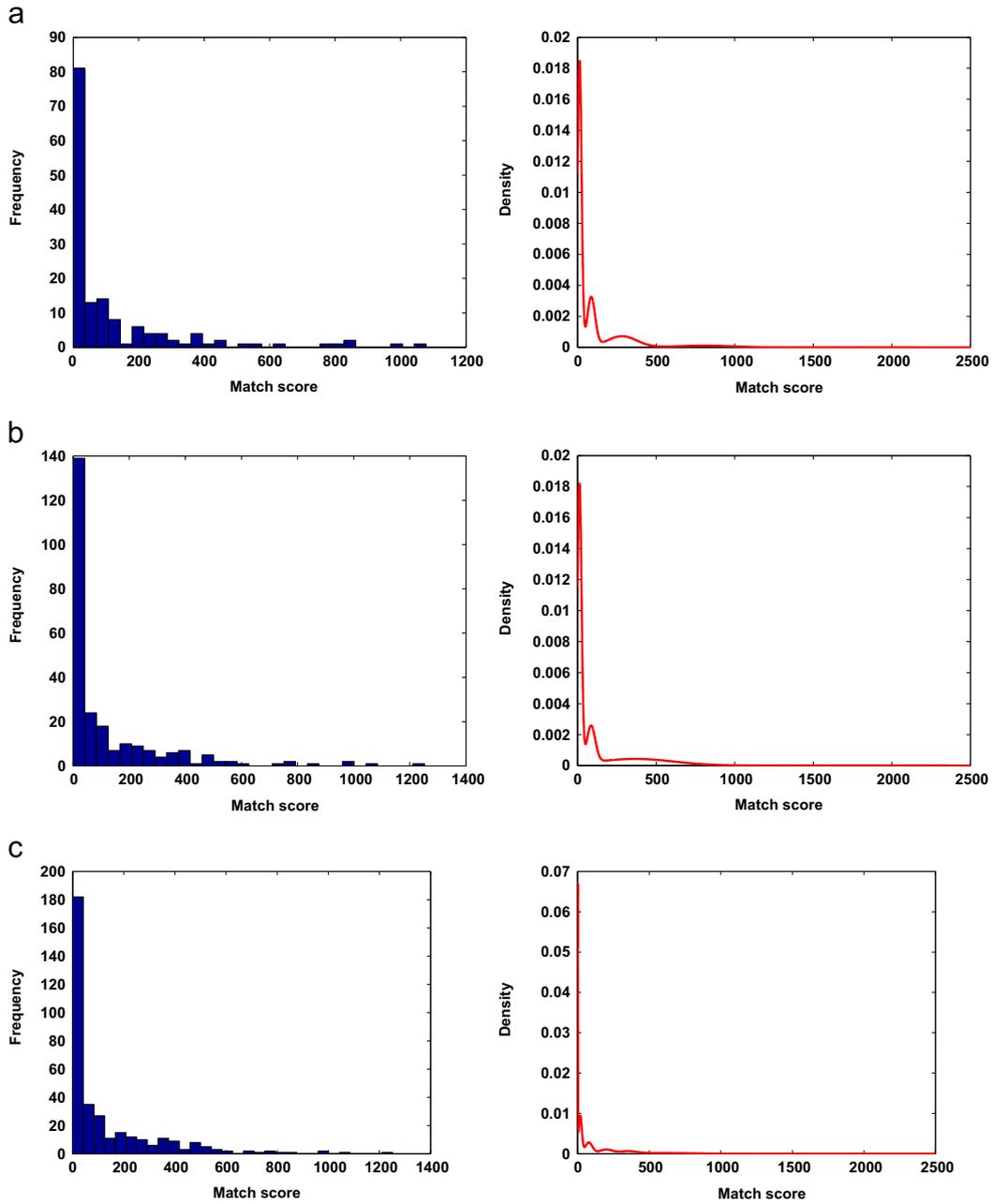| Small gallery size	Component #	Mean	Variance	Weight
50	4	684.458232	1135.136189	0.030087
		14.287706	314.973985	0.492921
		254.830754	11 292.774499	0.476992
		1624.000001	0.000001	0.000001

Fig. 9. Match score histogram and distribution with the small gallery size $n=50$.

Fig. 11. Match score histograms and distributions for different small gallery sizes: (a) four components, Small gallery size $n=100$.; (b) three components, Small gallery size $n=200$.; (c) six components,Small gallery size $n=300$.

prediction performance and the actual performance are 0.024, 0.017, and 0.016 when the large population sizes are 500, 1000, 1500 respectively. The absolute errors obtained from the different prediction evaluation approaches are not significant.

Now we use different small galleries with the learned optimal size to predict large population performance. Then, we select the maximum and the minimum prediction performance as our upper bound and lower bound for the performance prediction on the large population. Fig. 14 gives the upper bound and lower bound on the prediction of large population performance when the small gallery size $n=300$. Since we have 2000 pairs of fingerprints, the actual recognition performance for the distorted images is shown in Fig. 14. Beyond this population size, we can give the bounds for the prediction. From Fig. 14, it can be seen that the actual performance is within the upper bound and lower bound. Our

experiments show that when the small gallery size $n=300$ the prediction error is less than 0.05.

4.3. Estimation of the small gallery size based on statistical inequalities

Table 5 shows different small gallery sizes given different confidence intervals and margins of error for Chernoff inequality and Chebychev inequality ($\sigma^2 = 1$). From the table, we ascertain that the Chernoff inequality is much tighter than the Chebychev inequality. We compare our learning small gallery size with the Chernoff inequality. When the confidence interval $\alpha = 95\%$ and margin of error $e = 0.06$ then the small gallery size $n = 417$. From our experiment for the same margin of error the small gallery size is 300 and the confidence interval is $\alpha = 95\%$. Note that statistical
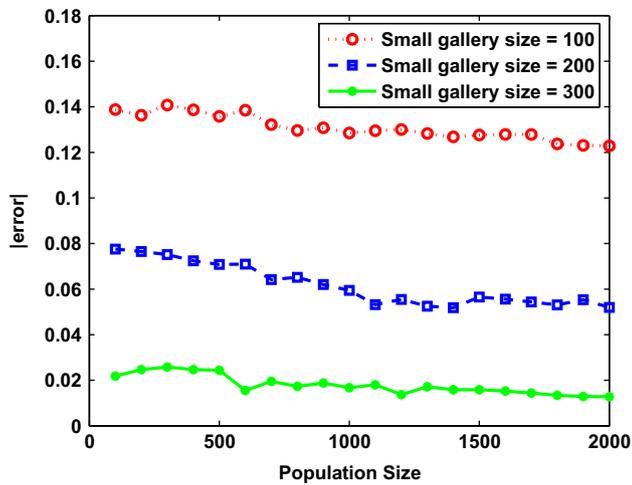
Fig. 12. Absolute error between the prediction and actual performance for different small gallery sizes.
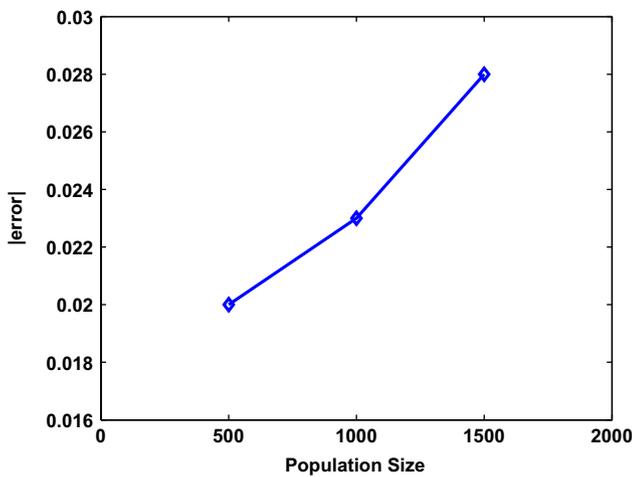

Fig. 13. Absolute error between the prediction and actual performance for the unseen data.
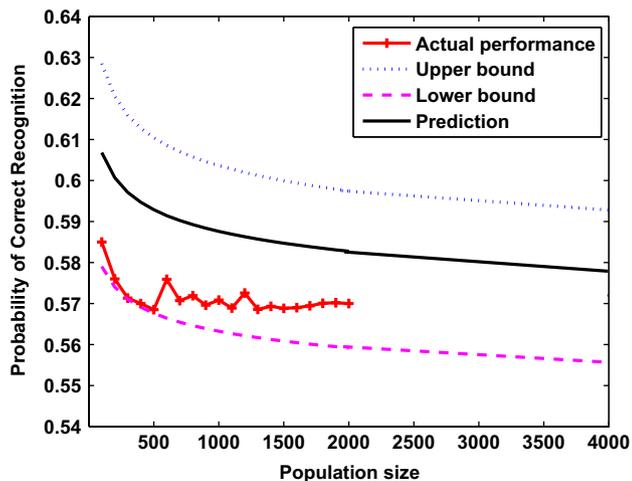

Fig. 14. The upper bound and good lower bound on the large population when the small gallery size is 300. Note that the difference between the upper bound and the lower bound are within 5%.

Table 5
Values of the confidence interval, the margin of error, and the small gallery size for Chernoff inequality and Chebychev inequality ($\sigma^2 = 1$).

$1-\alpha$	0.05	0.05	0.1	0.1	0.15	0.15
e	0.06	0.04	0.06	0.04	0.06	0.04
n (Chernoff)	417	937	320	720	264	593
n (Chebychev)	2778	6250	1389	3125	926	2083

methods give us a loose estimation of the small gallery size. Based on our own recognition system, we can find a more accurate small gallery size by learning.

5. Conclusions

We focus on the fundamental problem of performance prediction for object recognition: what is the optimal size of the small gallery that can give good error estimates and what is the confidence in the estimation? We use a generalized prediction model that combines a hypergeometric probability distribution model with a binomial model, taking into account distortion in large populations. We incorporate learning in the prediction process to find the optimal small gallery size and provide the upper and lower bounds for the performance prediction on large populations. The Chernoff inequality and the Chebychev inequality are used as a guide to obtain the small gallery size and the confidence interval given a margin of error. Experimental results show that the small gallery size obtained from the statistical methods are loose compared to the size provided by the proposed learning method. Using a sufficiently small gallery size we can provide prediction performance on large populations. Since our prediction model is score-based, it can be applied to the prediction problems that the match score and the non-match score are available. We believe that the methodology and results of this research will be useful not only for biometric applications but also to a wide range of applications of signal processing, image processing, computer vision, and pattern recognition.

Conflict of interest

None declared.

References

[1] D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar, Handbook of Fingerprint Recognition, Springer, New York, 2003.
[2] ISO/IEC19795-1, Information Technology—Biometric Performance Testing and Reporting—Part 1: Principles and Framework, 2006.
[3] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press Inc., 1990.
[4] I. Guyon, J. Makhoul, R. Schwartz, V. Vapnik, What size test set gives good error rate estimates? IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (1) (1998) 52–64.
[5] H. Cramér, Mathematical Methods of Statistics, Princeton University Press, 1951.
[6] T.M. Mitchell, Machine Learning, McGraw-Hill Book Company, 1997.

[7] R.M. Bolle, N.K. Ratha, S. Pankanti, Error analysis of pattern recognition systems—the subset bootstrap, Computer Vision and Image Understanding 93 (2004) 1–33.

[8] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classifications, Wiley-Interscience Publication, 2000.

[9] J.L. Wayman, Error-rate equations for the general biometric system, IEEE Robotics & Automation Magazine 6 (1999) 35–48.

[10] J. Daugman, The importance of being random: statistical principles of iris recognition, Pattern Recognition 36 (2) (2003) 279–291.

[11] P. Grother, P.J. Phillips, Models of large population recognition performance, in: IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, vol. 2, 2004, pp. 68–75.

[12] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, M. Bone, Face recognition vendor test 2002, in: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2003, p. 44.

[13] S. Pankanti, S. Prabhakar, A.K. Jain, On the individuality of fingerprints, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (8) (2002) 1010–1025.

[14] X. Tan, B. Bhanu, On the fundamental performance for fingerprint matching, in: IEEE Conference on Computer Vision and Pattern Recognition, Madison, WI, vol. 2, 2003, pp. 18–20.

[15] A.Y. Johnson, J.S.A.F. Boick, Using similarity scores from a small gallery to estimate recognition performance for large galleries, in: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, Nice, France, 2003, pp. 100–103.

[16] E. Tabassi, C.L. Wilson, C.I. Watson, Fingerprint Image Quality, National Institute of Standards and Technology International Report 7151.

[17] L.M. Wein, M. Baveja, Using fingerprint image quality to improve the identification performance of the U.S. visitor and immigrant status indicator technology program, The National Academy of Sciences 102 (21) (2005) 7772–7775.

[18] J. Han, B. Bhanu, Performance prediction for individual recognition by gait, Pattern Recognition Letters 20 (2005) 615–624.

[19] W. Li, X. Gao, Y. Zhu, V. Ramesh, T.E. Boult, On the small sample performance of boosted classifiers, in: IEEE Conference on Computer Vision and Pattern Recognition, San Diego, CA, vol. 2, 2005, pp. 574–581.

[20] P. Wang, Q. Ji, J. Wayman, Modeling and predicting face recognition system performance based on analysis of similarity scores, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (4) (2007) 665–670, http://dx.doi.org/10.1109/TPAMI.2007.1015.

[21] W. Scheirer, A. Bendale, T. Boult, Predicting biometric facial recognition failure with similarity surfaces and support vector machines, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1–8. http://dx.doi.org/10.1109/CVPRW.2008.4563124.

[22] G. Aggarwal, S. Biswas, P. Flynn, K. Bowyer, Predicting performance of face recognition systems: an image characterization approach, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2011, pp. 52–59. http://dx.doi.org/10.1109/CVPRW.2011.5981784.

[23] M. Lindenbaum, Bounds on shape recognition performance, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (7) (1995) 666–680.

[24] M. Lindenbaum, An integrated model for evaluating the amount of data required for reliable recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (11) (1997) 1251–1264.

[25] M. Boshra, B. Bhanu, Predicting performance of object recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (9) (2000) 956–969.

[26] S.C. Dass, Y. Zhu, A.K. Jain, Validating a biometric authentication system: sample size requirements, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (12) (2006) 1902–1913.

[27] R. Wang, B. Bhanu, Predicting fingerprint biometric performance from a small gallery, Pattern Recognition Letters 28 (1) (2006) 40–48.

[28] R. Wang, B. Bhanu, H. Chen, An integrated prediction model for biometrics, in: Audio- and Video-Based Biometric Person Authentication, Springer, New York, 2005, pp. 355–364.

[29] R. Wang, B. Bhanu, Learning models for predicting recognition performance, in: IEEE International Conference on Computer Vision, Beijing, China, vol. 2, 2005, pp. 1613–1618.

[30] A.M. Mood, F.A. Graybill, D.C. Boes, Introduction to the Theory of Statistics, McGraw-Hill Book Company, 1974.

[31] S. Ehrenfeld, S.B. Littauer, Introduction to Statistical Method, McGraw-Hill Book Company, 1964.

[32] A.K. Jain, L. Hong, R. Bolle, On-line fingerprint verification, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (4) (1997) 302–314.

[33] B. Bhanu, X. Tan, Learned templates for feature extraction in fingerprint images, in: IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, vol. 2, 2001, pp. 591–596.

[34] X. Tan, B. Bhanu, Robust fingerprint identification, in: IEEE International Conference on Image Processing, New York, 2002, pp. 277–280.

[35] B. Bhanu, X. Tan, Fingerprint indexing based on novel features of minutiae triplets, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (5) (2003) 616–622.

[36] B. Bhanu, X. Tan, Computational Algorithms for Fingerprint Recognition, Kluwer Academic Publishers, 2003.

[37] X. Tan, B. Bhanu, Fingerprint matching by genetic algorithms, Pattern Recognition 39 (3) (2006) 465–477.

[38] M.A.T. Figueiredo, A.K. Jain, Supervised learning of finite mixture models, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (3) (2002) 381–396.

Rong Wang received her Master's degree from Beijing Institute of Technology, China, in 2002, and PhD degree in Electrical Engineering from the University of California, Riverside, in 2007. Her research interests are computer vision, pattern recognition and machine learning. Currently she works at Pandigital Inc. as a Product Manager.

Bir Bhanu is the Distinguished Professor of Electrical Engineering and a Cooperative Professor of Computer Science and Engineering, Bioengineering, and Mechanical Engineering, and Director of the Center for Research in Intelligent Systems (CRIS), and the Visualization and Intelligent Systems Laboratory (VISLab) at the University of California, Riverside (UCR). Bhanu received the S.M. and E.E. degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology; the PhD degree in Electrical Engineering from the Image Processing Institute, University of Southern California and the MBA degree from the University of California, Irvine. Bhanu's current research interests are computer vision, pattern recognition and data mining, machine learning, artificial intelligence, image, image and video database, graphics and visualization, robotics, human–computer interactions, biological, medical, military and intelligence applications.

Ninad Thakoor received the BE degree in Electronics and Telecommunication Engineering from the University of Mumbai, India, in 2001, and the MS and PhD degrees in Electrical Engineering from the University of Texas at Arlington, in 2004 and 2009, respectively. His research interests include object recognition, stereo disparity segmentation, and structure-and-motion segmentation. Currently he is with the Center for Research in Intelligent Systems at the University of California at Riverside.