Available online at www.sciencedirect.com

SCIENCE DIRECT®

Pattern Recognition Letters 26 (2005) 615–624

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Performance prediction for individual recognition by gait

Ju Han, Bir Bhanu *

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

Received 8 January 2004; received in revised form 23 July 2004

Abstract

Existing gait recognition approaches do not give their theoretical or experimental performance predictions. Therefore, the discriminating power of gait as a feature for human recognition cannot be evaluated. In this paper, we first propose a kinematic-based approach to recognize human by gait. The proposed approach estimates 3D human walking parameters by performing a least squares fit of the 3D kinematic model to the 2D silhouette extracted from a monocular image sequence. Next, a Bayesian-based statistical analysis is performed to evaluate the discriminating power of extracted stationary gait features. Through probabilistic simulation, we not only predict the probability of correct recognition (PCR) with regard to different within-class feature variance, but also obtain the upper bound on PCR with regard to different human silhouette resolution. In addition, the maximum number of people in a database is obtained given the allowable error rate. This is extremely important for gait recognition in large databases.

Keywords: Performance prediction; Individual recognition by gait; Probability of correct recognition (PCR); Bayesian classifier

1. Introduction

Model-based object recognition is concerned with searching for a match to associate components of the given data with corresponding parameters of the object model (Grimson, 1990). The approaches can be classified as global matching or local feature matching. Global matching (e.g., silhouette image matching) approaches consider finding a transformation from a model to an image while feature matching approaches involve establishing a correspondence between local features extracted from the given data and corresponding local features of the object model.

Boshra and Bhanu (2000) present a method for predicting fundamental performance of object recognition. They assume that both scene data and model objects are represented by 2D point features and a data/model match is evaluated using a vote-based criterion. Their method considers data distortion factors such as uncertainty, occlusion,

* Corresponding author. Tel.: +1 9097873954; fax: +1 9097873188.

E-mail addresses: jhan@cris.ucr.edu (J. Han), bhanu@cris.ucr.edu (B. Bhanu).

0167-8655/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2004.09.011

and clutter, in addition to model similarity. This is unlike previous approaches, which consider only a subset of these factors. However, their assumptions make their method only applicable to local feature matching and not to global matching.

In our proposed approach of human recognition by kinematic-based gait analysis, we use global matching because we only have the global human silhouette information before matching. The detailed information for different body parts is obtained after matching. Next, we carry out Bayesian based statistical analysis to evaluate the discriminating power of various features. We address the prediction problem in the context of an object recognition task as follows: (1) scene data are represented by 2D regions where the region pixels are discretized at some resolution, and model objects are represented by 3D volumes; (2) an instance of a model object in the scene data is assumed to be obtained by applying a 3D to 2D transformation to the object; (3) the matching criterion is based on Bayesian theory.

2. Motivation and contributions

Current human recognition methods, such as fingerprints, face or iris biometrics, generally require a cooperative subject, views from certain aspects and physical contact or close proximity. These methods cannot reliably recognize noncooperating individuals at a distance in real-world changing environmental conditions. Moreover, in many applications of personnel identification, many established biometrics can be obscured. Gait, which concerns recognizing individuals by the way they walk, can be used as a biometric without the above-mentioned disadvantages.

As a new biometrics, gait also has some limitations. Gait can be affected by clothing, shoes, or environmental conditions. In addition, special physical conditions such as injury can also change people's walking style. Unlike fingerprint and iris, gait cannot be regarded as a unique characteristic for each person. Although the large gait variation of the same person under different conditions reduces the discriminating power of gait as a biometric, the inherent property of gait still makes it irreplaceable in visual surveillance applications.

In recent years, some approaches have already been employed in automatic gait recognition (i.e., human recognition by gait). Niyogi and Adelson (1994) make an initial attempt in a spatiotemporal (XYT) volume. They first find the bounding contours of the walker, and then fit a simplified stick model on them. A characteristic gait pattern in XYT is generated from the model parameters for recognition. Little and Boyd (1998) propose a model-free approach making no attempt to recover a structural model of human motion. Instead they describe the shape of the motion with a set of features derived from moments of a dense flow distribution. Similarly, He and Debrunner's (2000) approach detects a sequence of feature vectors based on Hu's moments of motion segmentation in each frame, and the individual is recognized from the feature vector sequence using hidden Markov models. To avoid a feature extraction process which may reduce reliability, Murase and Sakai (1996) propose a template matching method to calculate the spatio-temporal correlation in a parametric eigenspace representation for gait recognition. Huang et al. (1999, 2001) extend this approach by combining canonical space transformation (CST) with eigenspace transformation (EST) for feature selection.

However, most existing gait recognition approaches only consider human walking frontoparallel to the image plane. Moreover, none of the existing gait recognition approaches give their theoretical or experiential performance prediction. Therefore, we cannot evaluate the discriminating power of gait as a feature for human recognition. In this paper, we propose a kinematic-based approach to recognize human by gait, and carry out Bayesian based statistical analysis to predict recognition performance. The proposed approach estimates 3D human walking parameters by performing a least squares fit of the 3D kinematic model to the 2D silhouette extracted from a monocular image sequence. The gait features are then generated from the estimated model parameters for human recognition. Our approach eliminates the assumption of human walking frontoparallel to the image plane, which is desirable in many gait recognition applications.

3. Technical approach

3.1. Human kinematic model

A human body is considered as an articulated object, consisting of a number of body parts. The body model adopted here is shown in Fig. 1, where a circle represents a joint and a rectangle represent a body part (N: neck, S: shoulder, E: elbow, W: waist, H: hip, K: knee, and A: ankle). Most joints and body part ends can be represented as spheres, and most body parts can be represented as cones. The whole human kinematic model is represented as a set of cones connected by spheres (Lin, 1999). Fig. 2 shows that most of the body parts can be approximated well in this manner. However, the head is approximated only crudely by a sphere and the torso is approximated by a cylinder with two spheroid ends.

3.2. Matching 3D model with 2D silhouette

The matching procedure determines a parameter vector $\boldsymbol{x}$ so that the proposed 3D model fits the given 2D silhouette as well as possible. For that purpose, two chained transformations transform human body local coordinates (x, y, z) into image coordinates (x', y') (Wachter and Nagel, 1997). The first transformation transforms local coordinates into camera coordinates; while the second transformation projects camera coordinates into image coordinates.

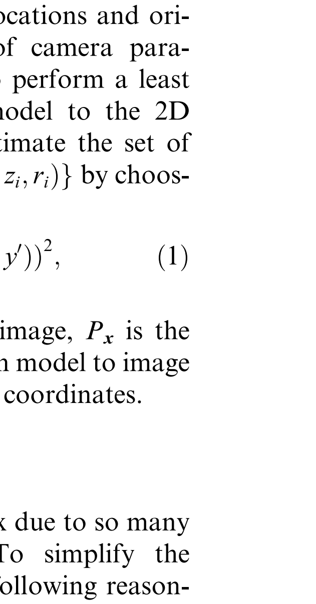

Fig. 1. 3D human kinematic model.

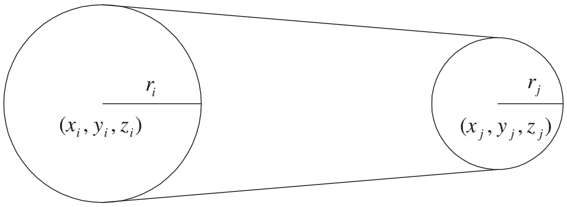

Fig. 2. Body part geometric representation.

Each 3D human body part is modeled by a cone with two spheres $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ at its ends, as shown in Fig. 2 (Lin, 1999). Each sphere $\boldsymbol{s}_i$ is fully defined by four scalar values, (x_i, y_i, z_i, r_i), which define its location and size. Given these values for two spheroid ends (x_i, y_i, z_i, r_i) and (x_j, y_j, z_j, r_j) of a 3D human body part model, its projection $P_{(ij)}$ onto the image plane is the convex hull of the two circles defined by (x'_i, y'_i, r'_i) and (x'_j, y'_j, r'_j).

If the 2D human silhouette is known, we may find the relative 3D body parts locations and orientations with the knowledge of camera parameters. We propose a method to perform a least squares fit of the 3D human model to the 2D human silhouette. That is, to estimate the set of sphere parameters $\boldsymbol{x} = \{\boldsymbol{x}_i : (x_i, y_i, z_i, r_i)\}$ by choosing $\boldsymbol{x}$ to minimize

$$\text{error}(\boldsymbol{x}; I) = \sum_{x', y' \in I} (P_{\boldsymbol{x}}(x', y') - I(x', y'))^2, \tag{1}$$

where I is the silhouette binary image, $P_{\boldsymbol{x}}$ is the binary projection of the 3D human model to image plane, and x', y' are image plane coordinates.

3.3. Model parameter selection

Human motion is very complex due to so many degrees of freedom (DOFs). To simplify the matching procedure, we use the following reasonable assumptions:

- the camera is stationary;
- people are walking before the camera at a distance;
- people are moving in a constant direction;
- the swing direction of arms and legs parallels to the moving direction.

According to these assumptions, we do not need to consider the waist joint, and only need to consider one DOF for each other joint. Therefore, the elements of the parameter vector of the 3D human kinematic model are defined as follows:

- Stationary parameters: radius $r_i(11)$: torso(3), shoulder, elbow, hand, hip, knee, ankle, toe, and head; length $l_i(9)$: torso, inter-shoulder, inter-hip, upper arm, lower arm, thigh, calf, foot, and neck;
- Kinematic parameters: location $(x,y)(2)$; angle $\theta_i(11)$: neck, left upper arm, left lower arm, right upper arm, right lower arm, left thigh, left calf, left foot, right thigh, right calf, and right foot.

With 33 stationary and kinematic parameters, the projection of the human model can be completely determined.

3.4. Silhouette extraction

Assuming that people are the only moving objects in the scene, human silhouette can be extracted by a simple background subtraction method. Notice that an area cast into shadow often results in a significant change in intensity without much change in chromaticity (Nadimi and Bhanu, 2002). Given an image sequence containing moving people and the corresponding background image, for each frame $\boldsymbol{I}_i$ in the sequence, the color value difference $\Delta\boldsymbol{p}_i(x,y) = \|\boldsymbol{p}_i(x,y) - \boldsymbol{p}_b(x,y)\|$ is computed for each pixel, where $\boldsymbol{p}_i(x,y)$ and $\boldsymbol{p}_b(x,y)$ are RGB color values of the pixel at (x,y) in the ith frame and background image, respectively. The chromaticity is computed as $r_c(x,y) = r(x,y)/(r(x,y) + g(x,y) + b(x,y))$ and $g_c(x,y) = g(x,y)/(r(x,y) + g(x,y) + b(x,y))$. We have $\Delta r_{ci}(x,y) = |r_{ci}(x,y) - r_{cb}(x,y)|$ and $\Delta g_{ci}(x,y) = |g_{ci}(x,y) - g_{cb}(x,y)|$. Given thresholds t_1 and t_2, if $(\Delta\boldsymbol{p}_i(x,y) > t_1) \wedge ((\Delta r_{ci}\ (x,y) > t_2) \vee (\Delta g_{ci}(x,y) > t_2))$, the pixel at (x,y) is determined to be part of the moving objects; otherwise, it is part of the background.

After the silhouette has been cleaned by a preprocessing procedure, its height, width and centroid can be easily extracted for motion analysis. In addition, the moving direction of the walking person is determined as follows:

$$\theta = \begin{cases} \tan^{-1}\dfrac{f(h_1 - h_N)}{h_1 y_N - h_N y_1}, & \text{if } y_1 > y_N; \\ \tan^{-1}\dfrac{f(h_1 - h_N)}{h_1 y_N - h_N y_1} + \pi, & \text{otherwise.} \end{cases} \tag{2}$$

where f is the camera focal length, y_1 and y_N are the horizontal centroid of the silhouette in the first and Nth frame, and h_1 and h_N are the height of the silhouette in the first and Nth frame.

3.5. Stationary parameter estimation

The stationary parameters include body part length and joint radius. Notice that human walking is a cyclic motion, so an image sequence can be divided into motion cycles and studied separately. In each walking cycle, the silhouette with minimum width means that the person stands straight and that means the maximum occlusion; the silhouette with maximum width means the least occlusion and, therefore, it is more reliable.

To estimate the stationary parameters, we first select several key frames (4 frames in our experiments) which contain more reliable silhouettes, and then perform matching procedure on the key frames as a whole. The corresponding feature vector thus includes 20 common stationary parameters and $13 * 4$ individual kinematic parameters. Next, we initialize these parameters according to the human statistical information. Then, the set of parameters is estimated from these initial parameters by choosing a parameter vector $\boldsymbol{x}$ to minimize the least square error in (1) with respect to the kinematic constraints.

After the matching algorithm is converged, the estimated stationary parameters so obtained are used for kinematic parameter estimation of other frames. At the same time, the estimated kinematic parameters of key frames are used for prediction. Because even the same person might walk at different speed, we normalize the estimated kinematic parameters of each walking cycle to a fixed-length walking cycle, and the kinematic gait features are generated from the normalized walking cycle.

4. Recognition performance prediction

In this paper, we only use features from stationary parameters for gait recognition. In the above-mentioned stationary parameters, radius parameters will be different if the same person is in different clothes, and are thus not reliable for recognition. Similarly, inter-shoulder and inter-hip length parameters are not reliable when people walk within a small angle along the direction perpendicular to the camera axis. The head region depends on the hair style, which will change if the view changes, and the head representation in our model (sphere) is not precise in some cases, so the estimated neck length is also not reliable. Therefore, the feature vector selected for human recognition in our approach includes six elements: torso length, upper arm length, lower arm length, thigh length, calf length, and foot length, which are not sensitive for recognizing human with different clothes and different walking directions. In this paper, we consider uncertainties for feature vectors in two categories: uncertainties from all factors which are algorithm dependent; uncertainties only from different silhouette resolutions that are algorithm independent.

4.1. Body part length distribution

To predict the performance of recognizing human from body part lengths, we have to know the prior length distributions of body parts over human population. The data are called static anthropometric data shown in Fig. 3. Although the data are surveyed in the British population, the predicted performance on it is applicable in other scenarios. In general, the mean of body part lengths will change but the standard deviation will not change a lot in different populations. Assuming that men and women have the same population, the overall distributions for each of the body part lengths are obtained. In this paper, we only consider that the body part lengths are independently distributed due to the absence of statistical knowledge of their correlation.

4.2. Algorithm dependent performance prediction

Uncertainties of stationary gait features come from various sources: image quantization error, camera calibration error, silhouette segmentation error, matching error, and body part occlusion. To completely model the uncertainties of 3D body part lengths, we have to model all the above-mentioned factors. This is a challenging task because it is difficult to mathematically find the distribution functions of uncertainties for all these factors. A reasonable approach is to estimate the uncertainties from training data. Assuming that feature vectors obtained from a feature extraction algorithm for a person are normally distributed in the given feature space, we can easily obtain the within-class variance from the experimental results on the training data. Then, the obtained within-class variance can be used to predict the recognition performance of this algorithm.

According to the Bayes decision theory, an unknown feature vector $\boldsymbol{x}$ is assigned to class ω_i if $P(\omega_i|x) > P(\omega_j|x) \forall j \neq i$ (Theodoridis and Koutroumbas, 1998). Let $g_i(x) = \ln(p(x|\omega_i)P(\omega_i))$, this decision test becomes classifying $\boldsymbol{x}$ to ω_i if $g_i(x) > g_j(x) \forall j \neq i$.

Assuming the feature vector $\boldsymbol{x}$ for a person ω_i is normally distributed in l-dimensional feature space, the probability distributions for ω_i with respect to $\boldsymbol{x}$ follows $p(\boldsymbol{x}|\omega_i) = \frac{1}{(2\pi)^{\frac{l}{2}}|\sum_i|^{\frac{1}{2}}} \times \exp(-\frac{1}{2}(\boldsymbol{x} - \mu_i)^T \sum_i^{-1} (\boldsymbol{x} - \mu_i))$ for $i = 1, \ldots, M$, where $\mu_i = E[\boldsymbol{x}]$ is the mean value of the ω_i class and $\sum_i$ is the $l \times l$ covariance matrix defined as $\sum_i = E[(\boldsymbol{x} - \mu_i)(\boldsymbol{x} - \mu_i)^T]$. Assume that $\sum_i = \sum$ for all i, maximum $g_i(x)$ implies minimum Mahalanobis distance: $d_M = (\boldsymbol{x} - \mu_i)^T \sum^{-1} (\boldsymbol{x} - \mu_i)$. Thus, feature vectors are assigned to classes according to their Mahalanobis distances from the respective mean vectors.

With the body part length distributions shown in Fig. 3 and the within-class covariance matrix $\sum$ of the features obtained from a feature extraction

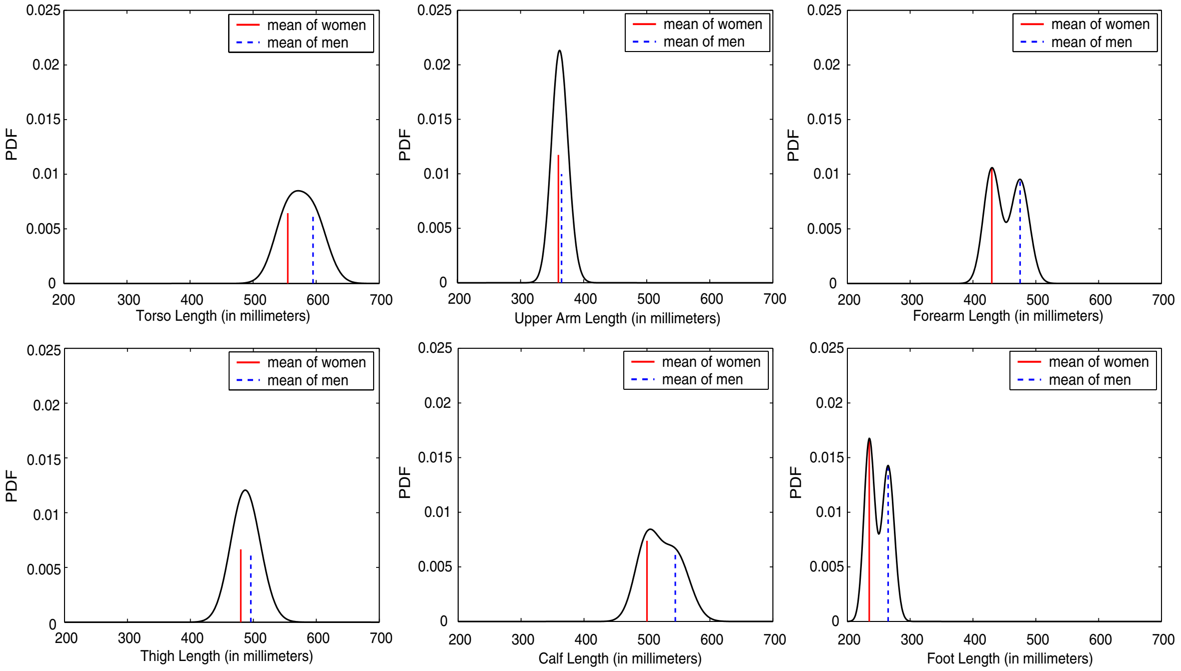

Fig. 3. Anthropometric estimates for British adults aged 19–65 years (Pheasant, 1986).

approach, we can predict its *probability of correct recognition* (PCR) with regard to the number of classes (people) in the database through a simulation approach. In order to provide a direct understanding about how the within-class covariance matrix affects the recognition performance, we assume that $\sum_i = \sigma^2 I$ for all i. This is a reasonable assumption since all the features are lengths, although different features may have slightly different variances and may be slightly correlated. Therefore, maximum $g_i(\boldsymbol{x})$ implies minimum Euclidean distance: $d_E = \|\boldsymbol{x} - \mu_i\|$. Thus, feature vectors are assigned to classes according to their Euclidean distances from the respective mean vectors. In this way, we can obtain a plot directly indicating the relationship between the predicted recognition performance and the within-class standard deviation σ.

4.3. Upper bound on PCR

We have considered the uncertainties that are dependent on feature extraction algorithms. The predicted performance indicates the discriminant power of features extracted by different algorithms, and these algorithms can therefore be compared. However, we still do not know the upper bound on PCR which can be achieved independent of different algorithms. In the ideal case, image quantization errors, i.e., the human silhouette resolution, is the only source of uncertainties. By analyzing the uncertainties given a fixed silhouette resolution, we can obtain the upper bound on PCR with regard to the number of classes (people) in the database.

Given the silhouette resolution r, we can compute the corresponding uncertainty from the body part length L, view angles a and b, and the walking direction c through two steps as shown in Fig. 4. The first step is projecting the 3D length L to length l' in the 2D continuous plane. We obtain the projection of L on the plane at depth h which is perpendicular to the camera axis as follows:

$$L' = L(\cos c + \sin c \tan(a+b)). \tag{3}$$

Fig. 4(a) only shows the case where $a > 0$, $b > 0$ and $c > 0$. We can easily derive the same equation

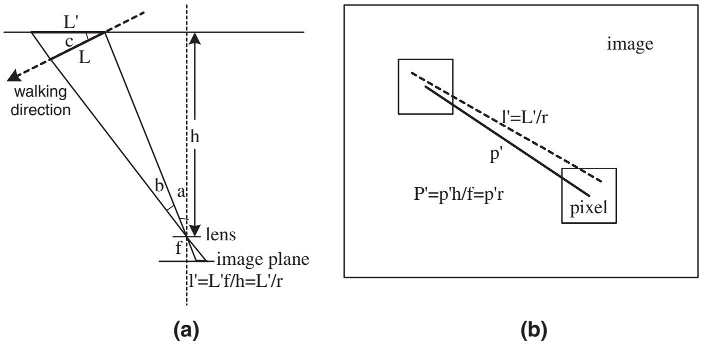

Fig. 4. Uncertainty computation for the given silhouette resolution *r* (in millimeters/pixel), the body part length *L*, view angles *a* and *b*, and the walking direction *c*.

in other cases. Then the corresponding length of L in the continuous image plane can be computed from the following equation:

$$l' = L'f/h = L'r, \tag{4}$$

where f is the camera focal length.

The second step is the image quantization step as shown in Fig. 4(b). For every 2D point falling into a box in the continuous image plane, it's location is represented by the center location of the box in the discrete image plane. Therefore, the corresponding length of L in the discrete image plane is the discrete value p'. From (3) and (4), we can obtain the following results

$$P' = p'h/f = p'r, \tag{5}$$

$$P = \frac{P'}{\cos c + \sin c \tan(a+b)} = \frac{p'r}{\cos c + \sin c \tan(a+b)}, \tag{6}$$

where P is the corresponding length of p' in 3D space and P' is the projection of P on the plane at depth h which is perpendicular to the camera axis. Therefore, the overall error in Fig. 4 is $P - L$. Considering $h \gg L$ in our applications, we have $b \approx 0$, and (6) becomes

$$P = p'r/(\cos c + \sin c \tan a). \tag{7}$$

Assuming the elements in the feature vector are independent and identically distributed, the minimum Euclidean distance classification criteria is still effective.

Assuming that the quantization error is uniformly distributed in the $r \times r$ area, view angle a is uniformly distributed from $-45°$ to $45°$ of arc, and walking direction c is uniformly distributed from $-30°$ to $30°$ of arc, we can predict the recognition performance with regard to the number of classes (people) in the database through a simulation approach. The prediction results are upper bounds on PCR with regards to different human silhouette resolution values.

5. Experimental results

5.1. Performance prediction results

In our experiments, the performance prediction results are obtained through simulation approaches. First we randomly generate the body part lengths of M classes (people) according to the distribution of different body part lengths. Next, for each of the M classes, we randomly generate N instances for this class according to the uncertainties in Section 4.2 or Section 4.3. Then, the recognition rate is obtained by the minimum Euclidean distance classification on the $M \times N$ instances. After this experiment has been repeated for K times, we can obtain the average recognition rate. If $N * K$ is large enough ($N = 100$ and $K = 100$ in our experiments), this average recognition rate can be viewed as the predicted PCR of the given algorithm (Fig. 5), and upper bounds on PCR (Fig. 6). From these prediction results, we

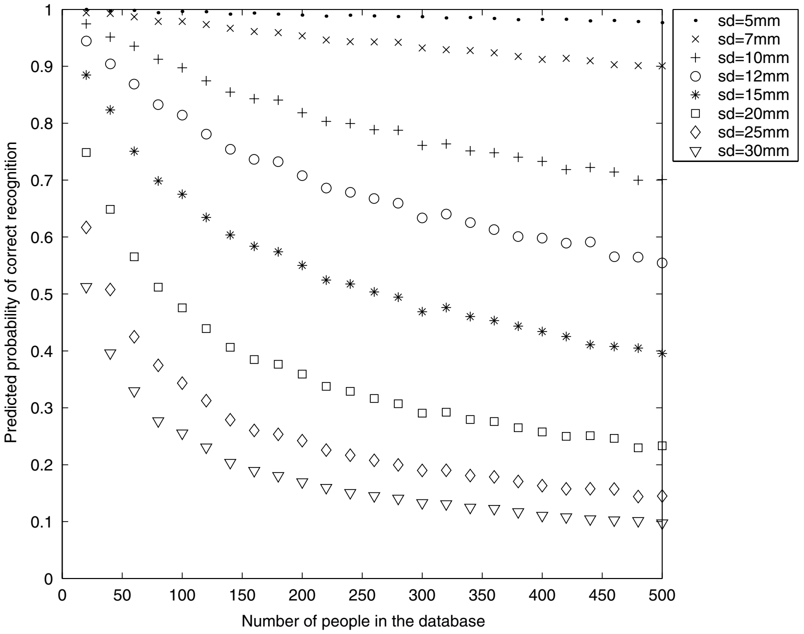

Fig. 5. Predicted PCR with regard to different database sizes and different within-class standard deviation values of the features extracted from different algorithms.

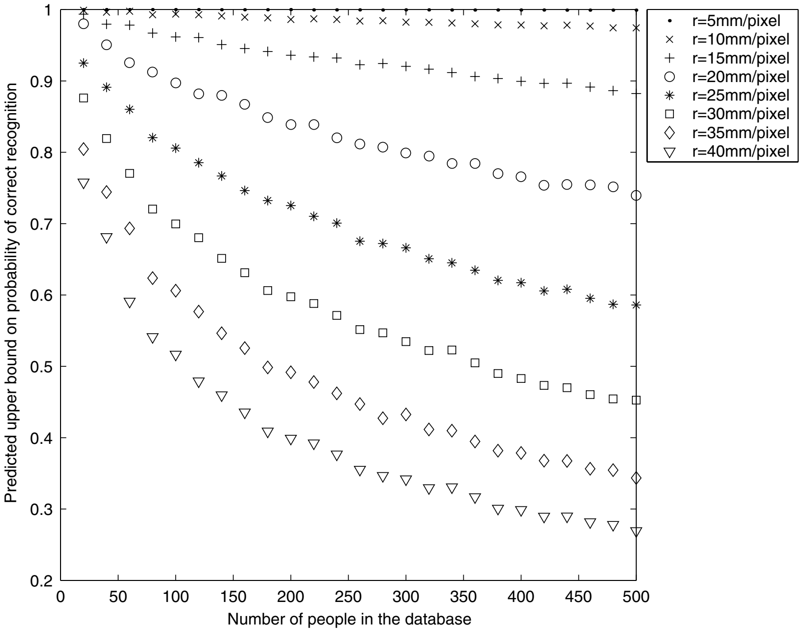

Fig. 6. Predicted upper bound on PCR with regard to different database sizes and different human silhouette resolution values.

can find the corresponding maximum number of people in a database given the allowable error rate. Table 1 shows the corresponding resolution (mm/pixel) for a 1675mm (population average height) person occupying different vertical portions of the frame with different video formats. It is shown that most of these resolutions are good enough for human recognition in databases of less than 500 people.

Our prediction results are based on the assumption that the selected length features are independently distributed with an identical Gaussian distribution. This assumption may not accurately reflect different types of perturbations. In the future, we will investigate the real feature distribution under different types of perturbations.

5.2. Recognition results on real data

The video data used in our experiments are real human walking data recorded in an outdoor environment, and there is only one walking person at the same time. Eight different people walk along different directions (within [−45°, 45°] along the image plane). The size of image frames is 180 × 240. In our experiments, we first manually divide video data into single-cycle sequences, and then select 15 sequences from each person: 10 sequences for training and 5 sequences for testing. Fig. 7 shows sequences in our gait database.

The least square matching algorithm is implemented using a genetic algorithm. The fitness

Table 1
Resolution (mm/pixel) for a 1675mm (population average height) person occupying different vertical portions of the frame with different video formats

Human Silhouette occupancy	VHS (240 lines)	Digital video (480 lines)	High definition (1080 lines)
100% of frame	6.98	3.49	1.55
75% of frame	9.31	4.65	2.07
50% of frame	13.96	6.98	3.10
25% of frame	27.92	13.96	6.20

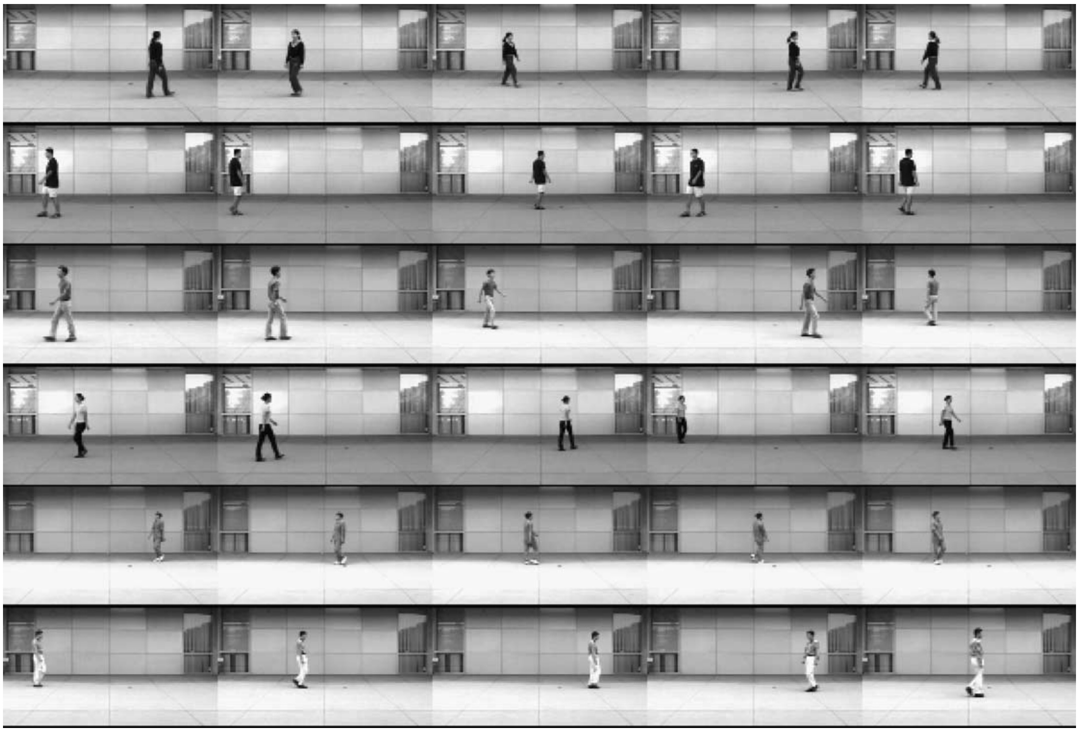

Fig. 7. Sample sequences in our database.

function is computed from the matching error in (1). In the experiments, our approach achieves 60% recognition rate on the training dataset using the Leave-One-Out method. The performance on the testing data is 42% recognition rate. We also compute the average standard deviation for each person in the database which is 20mm, and the corresponding predicted PCR is 87%. The correct recognition rate in our approach is much lower than this PCR because the PCR is computed on the data distributed according to Fig. 3 while the data in our database are not well distributed due to the small data size, i.e., they have more similarity. The human silhouette resolution in our database varies from 20 to 30mm/pixel, and the corresponding predicted upper bound on PCR in the ideal case is from 94.67% to 98.80%. The predicted PCR (87%) is lower than the upper bound because the feature extraction procedure introduces several additional uncertainties such as camera calibration error, silhouette segmentation error, matching error, and body part occlusion.

Note that the use of binary silhouette to fit 3D model suffers from ambiguity as a result of body parts self-occlusion, and the use of least squares makes it sensitive to noise in the silhouette. This problem can be solved by considering the correlation between adjacent frames.

6. Conclusions

In this paper, we proposed a Bayesian based statistical analysis to evaluate the discriminating power of extracted features. Through probabilistic simulation, we not only obtain the probability of correct recognition for our approach, but also obtain the upper bound on the probability of correct recognition with regard to different human silhouette resolution in ideal cases. We obtain the plots characterizing maximum number of people in the database that can be recognized given the allowable error rate. This will guide future research for gait recognition in large databases. The discrepancy between actual and predicted results will be reduced by developing better gait recognition algorithms.

References

Grimson, W., 1990. Object recognition by computer: The role of geometric constraints. The MIT Press.

Boshra, M., Bhanu, B., 2000. Predicting performance of object recognition. IEEE Trans. Pattern Anal. Machine Intell. 22 (9), 956–969.

Niyogi, S., Adelson, E., 1994. Analyzing and recognizing walking figures in xyt, in: Proc. IEEE Conference on CVPR, pp. 469–474.

Little, J., Boyd, J., 1998. Recognizing people by their gait: The shape of motion. Videre: J. Comput. Vis. Res. 1 (2), 1–32.

He, Q., Debrunner, C., 2000. Individual recognition from periodic activity using hidden markov models, in: Proc. IEEE Workshop on Human Motion, pp. 47–52.

Murase, H., Sakai, R., 1996. Moving object recognition in eigenspace representation: Gait analysis and lip reading. Pattern Recognition Lett. 17 (2), 155–162.

Huang, P., Harris, C., Nixon, M., 1999. Recognizing humans by gait via parameteric canonical space. Artificial Intell. Eng. 13, 359–366.

Huang, P., 2001. Automatic gait recognition via statistical approaches for extended template features. IEEE Trans. Systems Man Cybernet., Part B 31 (5), 818–824.

Lin, M., 1999.Tracking articulated objects in real-time range image sequences, in: Proc. International Conference on Computer Vision, pp. 648–653.

Wachter, S., Nagel, H.-H., 1997. Tracking of persons in monocular image sequences, in: Proc. IEEE Workshop on Nonrigid and Articulated Motion, pp. 2–9.

Nadimi, S., Bhanu, B., 2002. Moving shadow detection using a physics-based approach, in: Proc. International Conference on Pattern Recognition, vol. 2, pp. 701–704.

Pheasant, S., 1986. Bodyspace: Anthropometry, Ergonomics and Design. Taylor & Francis.

Theodoridis, S., Koutroumbas, K., 1998. Pattern Recognition. Academic Press.