Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2008, Article ID 261317, 13 pages
doi:10.1155/2008/261317

Research Article

A Full-Body Layered Deformable Model for Automatic Model-Based Gait Recognition

Haiping Lu,[1] Konstantinos N. Plataniotis,[1] and Anastasios N. Venetsanopoulos[2]

[1] *The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Canada M5S 3G4*
[2] *Department of Electrical and Computer Engineering, Ryerson University, Toronto, Ontario, Canada M5B 2K3*

Correspondence should be addressed to Haiping Lu, hplu@ieee.org

Received 17 May 2007; Accepted 15 September 2007

Recommended by Arun Ross

This paper proposes a full-body layered deformable model (LDM) inspired by manually labeled silhouettes for automatic model-based gait recognition from part-level gait dynamics in monocular video sequences. The LDM is defined for the fronto-parallel gait with 22 parameters describing the human body part shapes (widths and lengths) and dynamics (positions and orientations). There are four layers in the LDM and the limbs are deformable. Algorithms for LDM-based human body pose recovery are then developed to estimate the LDM parameters from both manually labeled and automatically extracted silhouettes, where the automatic silhouette extraction is through a coarse-to-fine localization and extraction procedure. The estimated LDM parameters are used for model-based gait recognition by employing the dynamic time warping for matching and adopting the combination scheme in AdaBoost.M2. While the existing model-based gait recognition approaches focus primarily on the lower limbs, the estimated LDM parameters enable us to study full-body model-based gait recognition by utilizing the dynamics of the upper limbs, the shoulders and the head as well. In the experiments, the LDM-based gait recognition is tested on gait sequences with differences in shoe-type, surface, carrying condition and time. The results demonstrate that the recognition performance benefits from not only the lower limb dynamics, but also the dynamics of the upper limbs, the shoulders and the head. In addition, the LDM can serve as an analysis tool for studying factors affecting the gait under various conditions.

1. INTRODUCTION

Automatic person identification is an important task in visual surveillance, and monitoring applications in security-sensitive environments such as airports, banks, malls, parking lots, and large civic structures, and biometrics such as iris, face, and fingerprint have been researched extensively for this purpose. Gait, the style of walking of an individual, is an emerging behavioral biometric that offers the potential for vision-based recognition at a distance where the resolution is not high enough for the other biometrics to work [1–4]. In 1975 [5], Johansson used point light displays to demonstrate the ability of humans to rapidly distinguish human locomotion from other motion patterns. Similar experiments later showed the capability of identifying friends or the gender of a person [6, 7], and Stevenage et al. show that humans can identify individuals based on their gait signature in the presence of lighting variations and under brief exposures [8]. Recently, there has been increased research activities in gait recognition from video sequences. Vision-based gait capture is unobtrusive, requiring no cooperation or attention of the observed subject and gait is difficult to hide. These advantages of gait as a biometric make it particularly attractive in human identification at a distance. In a typical vision-based gait recognition application, a monocular video sequence is used as the input.

Gait recognition approaches can be broadly categorized into the model-based approach, where human body structure is explicitly modeled, and the model-free approach, where gait is treated as a sequence of holistic binary patterns (silhouettes). Although the state-of-the-art gait recognition algorithms are taking the model-free approach [3, 4, 9–12], from the literature of the anthropometry and the biomechanics of human gait [13, 14], human body is structured with well-defined body segments and human gait is essentially the way locomotion is achieved through the movement of human limbs. Therefore, for detailed analysis and in-depth understanding of what contributes to the observed

gait (and gait-related applications), it is natural to study the movement of individual human body segments, rather than treating human body as one whole holistic pattern. For example, contrary to common beliefs that cleaner silhouettes are desired for successful recognition, a recent study [15] shows that automatically extracted (noisy) silhouette sequences achieve better recognition results than very-clean (more accurate) manually segmented silhouettes [16], and the explanation in [16] is that there are correlated errors (noise) contributing to the recognition in the noisy silhouette sequences. On the other hand, the model-based approach [17–21] extracts gait dynamics (various human body poses) for recognition and appears to be more sound, but it is not well studied and less successful due to the difficulties in accurate gait dynamics extraction [1, 3]. For these existing model-based gait recognition algorithms, only the dynamics of the lower body (the legs) are used for recognition, except in [20], where the head x-displacement is also used. However, in the visual perception of a human gait, the dynamics of the upper-body, including the arms, the shoulders, and even the head, contributes significantly to the identification of a familiar person as well. Therefore, it is worthwhile to investigate whether it is feasible to extract the upper-body dynamics from monocular video sequences and whether the gait recognition performance can benefit from it.

Motivated by the discussions above, the earlier version of this paper proposed a new full-body articulated human body model for realistic modeling of human movement, named as the layered deformable model (LDM) [22]. It is inspired by the manually labeled body-part-level silhouettes [15] from the "gait challenge" data sets, which were created for studying gait recognition from sequences free from noise and background interference, and it is designed to closely match them in order to study gait recognition from detailed part-level gait dynamics. In this paper, more detailed descriptions and in-depth discussions on the LDM and the pose recovery algorithms proposed in [22] are provided; and furthermore, the LDM is applied to the automatic model-based gait recognition problem. An overview of the LDM-based gait recognition is shown in Figure 1. A coarse-to-fine silhouette extraction algorithm is employed to obtain silhouettes automatically from a monocular video sequence and human body pose recovery algorithms are then developed to estimate the LDM parameters from the silhouettes. The pose recovery algorithms developed here do not rely on any tracking algorithm. Hence, it is fully automatic and does not suffer tracking failures as in [17], where manual parameter estimation is needed when the tracking algorithm fails due to the problems of body part self-occlusion, shadows, occlusion by other objects, and illumination variation in the challenging outdoor environment. Next, the dynamic time warping (DTW) algorithm is utilized for matching body part dynamics and the combination scheme in AdaBoost.M2 is adopted to integrate the various part-level gait dynamics. The gait recognition experiments are carried out on a subset of the gait challenge data sets [9, 15] and several interesting observations are made.

The rest of this paper is organized as follows: Section 2 describes the LDM. In Section 3, human body pose recovery algorithms are presented in more details for manual silhouettes and automatically extracted silhouettes, followed by a brief discussion on the computational complexity. The LDM-based gait recognition module is then proposed in Section 4. Finally, the experimental results are reported in Section 5 and conclusions are drawn in Section 6.

2. THE LAYERED DEFORMABLE MODEL

As discussed in [22], in model-based gait recognition, the desirable human body model should be of moderate complexity for fast processing while at the same time it should provide enough features for discriminant learning. In other words, a tradeoff between the body model complexity (concerning the efficiency) and the model descriptiveness (concerning the accuracy) is sought. It is not to be as detailed as a fully deformable model used for realistic modeling (e.g., of animated characters in movies) in computer graphics and animations, while it must model limbs individually to enable model-based recognition. The existing model-based gait recognition algorithms [17–21] regard the lower-body (the legs) dynamics as the discriminative features and almost completely ignore the upper-body dynamics. Such ignorance is partly due to the difficulty in accurate extraction of the upper-body dynamics and their assumption that the leg dynamics are most important for recognition. However, in our opinion, the upper-body dynamics (the arms, shoulders, and head) provide us with valuable information for identification of a person as well. Therefore, gait recognition algorithms based on a full-body model are expected to achieve better results than those relying on only the lower-body dynamics.

Although there are works making use of the full-body information, such as the seven-ellipse representation in [23] and the combination of the left/right projection vectors and the width vectors in [24], these representations are rather heuristic. Since the biomechanics of human gait is a well-studied subject, it is helpful to develop a human body model by incorporating knowledge from this area. At the same time, as a vision-based approach, the information available for model estimation is limited to what can be extracted from a camera at a distance different from the marker-based studies in biomechanics of human gait [14].

The human full-body model named as the layered deformable model (LDM) was first proposed in [22] for the most commonly used fronto-parallel gait (side-view), although it can be designed for gait from various viewing angles. Without loss of generality, it is assumed that the walking direction is from the right to the left. This model is inspired by the manually labeled silhouettes provided by the University of South Florida (USF) [15], where the silhouette in each frame was specified manually for five key sets: the gallery set, probes B, D, H, and K. (In typical pattern recognition problems, such as human identification using fingerprints, face, or gait signals, there are two types of data sets: the gallery and the probe [9]. The gallery set contains the set of data samples with known identities and it is used for training. The probe set is the testing set where data samples of unknown identity are to be identified and classified via matching with corresponding entries in the gallery set.) In

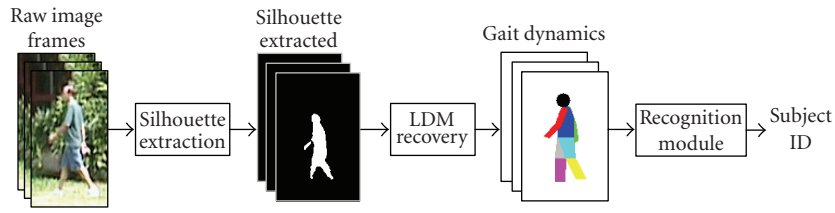


Figure 1: Overview of the proposed automatic LDM-based gait recognition.

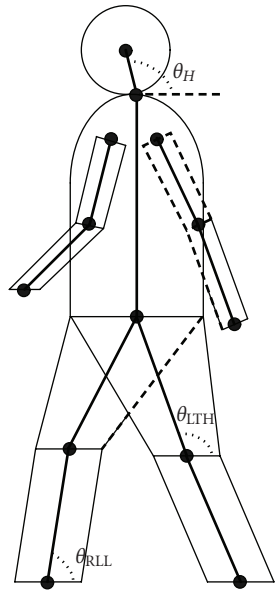


Figure 2: The layered deformable model.

addition, more detailed specifications in terms of body parts were provided. These manual silhouettes are considered to be the ideal "clean" silhouettes that can be obtained from the raw video sequences.

Following [22], the LDM consists of ten segments modeling the ten body parts: the head (a circle), the torso (a semiellipse on top of a rectangle), the left/right upper arms (rectangles), the left/right lower arms (quadrangles), the left/right upper/lower legs (quadrangles). The feet and the hands are not modeled explicitly since they are relatively small in size and difficult to detect consistently due to occlusion with the "background" (e.g., covered by grass). Figure 2 is an illustration of the LDM, which matches closely to the manual silhouettes in [15]. The model is defined based on a skeleton model, which is shown as thick lines and black dots in the figure.

The LDM is specified in [22] using the following 22 parameters that define the lengths, widths, positions, and orientations of body parts, with the number of parameters for each category in brackets:

(i) lengths (6): the lengths of various body parts L_H (the radius of the head), L_T (the torso), L_{UA} (the upper arm), L_{LA} (the lower arm, including the hand), L_{TH} (the thigh), and L_{LL} (the lower leg, including the feet);

(ii) widths (3): the widths (thickness) of body parts W_T (the torso, which is equal to the width of the top of the thigh), W_K (the knee), and W_A (the arm, assuming the same width for the upper and lower parts);

(iii) positions (4): the global position (x_G, y_G), which is also the position of the hip joint, and the shoulder displacement $(dx_{\text{Sh}}, dy_{\text{Sh}})$;

(iv) body part orientations (9): θ_{LTH} (the left thigh), θ_{RTH} (the right thigh), θ_{LLL} (the left lower leg), θ_{RLL} (the right lower leg), θ_{LUA} (the left upper arm), θ_{RUA} (the right upper arm), θ_{LLA} (the left lower arm), θ_{RLA} (the right lower arm), and θ_H (the head, the neck joint angle). The body part orientation is measured in the angle between the major axis of the body part and the horizontal axis, following the biomechanics conventions in [13]. In Figure 2, θ_{RLL}, θ_{LTH}, and θ_H are labeled for illustration.

In addition to the 22 parameters for the LDM, the height of the human full-body is denoted as H_F.

Furthermore, in order to model the human body self-occlusion (e.g., between legs, arms, and torso), the following four layers are introduced in [22], inspired by the layered representation in [25]:

(i) layer one: the right arm;
(ii) layer two: the right leg;
(iii) layer three: the head, the torso, and the left leg;
(iv) layer four: the left arm

where the first layer is furthest from the camera (frequently occluded) and the fourth layer is the closest to the camera (seldom occluded). Figure 3 shows each layer as well as the resulted overlaid image. As seen from the figure, self-occlusion is explained well with this model. Let $\mathbf{L}_j$ denote the image of layer j, where $j = 1, 2, 3, 4$. The gait stance image $\mathbf{I}_g$ obtained by overlying all layers in order can be written as

$$\mathbf{I}_g = \mathbf{L}_4 + B(\mathbf{L}_4) * (\mathbf{L}_3 + B(\mathbf{L}_3) * (\mathbf{L}_2 + B(\mathbf{L}_2) * \mathbf{L}_1)), \quad (1)$$

where "$*$" denotes the elementwise multiplication and $B(\mathbf{L}_j)$ is the mask obtained by setting all the foreground pixels (the body segments) in $\mathbf{L}_j$ to zero and all the background pixels to one. The difference of this layered representation from that in [25] is that here the foreground boundary is determined uniquely by the layer image $\mathbf{L}_j$ and there is no need to introduce an extra mask.

As described in [22], the LDM allows for limb deformation and Figure 4 shows an example for the right leg deformation. This is different from the traditional 2D

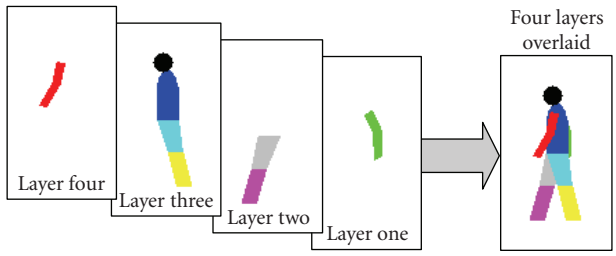


Figure 3: The four-layer representation of the LDM.

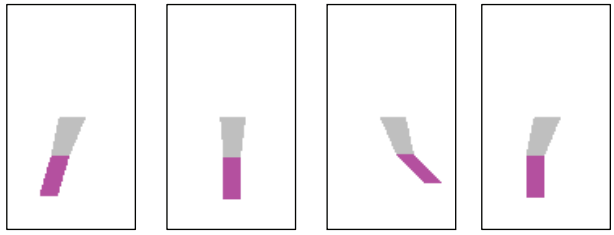


Figure 4: Illustration of the right leg deformation.

(rectangular) models and visual comparison with the manual silhouettes [15] shows that the LDM matches well with human's subjective perception of human body (in 2D).

On the whole, the LDM is able to model human gait realistically with moderate complexity. It has a compact representation comparable to the simple rectangle (cylinder) model [17] and its layered structure models self-occlusion between body parts. At the same time, it models simple limb deformation while it is not as complicated as the fully deformable model [26]. In addition, the shoulder displacement parameters model shoulder swing observed in the manual silhouette sequences, which is shown to be useful for automatic gait recognition in the experiments (Section 5.2), and they also relate to viewing angles.

3. LDM-BASED HUMAN BODY POSE RECOVERY

With the LDM, the pose (LDM parameter) estimation problem is solved in two phases. The estimation of the LDM parameters from the manually labeled silhouettes is tackled first, serving as the ground truth in pose recovery performance evaluation and facilitating the studies of the ideal-case model-based gait recognition. In addition, statistics from the LDM parameters obtained from the manual silhouettes are used in the following task of direct LDM parameter estimation for the silhouettes extracted automatically from raw gait sequences.

3.1. *Pose estimation from manually labeled silhouettes*

For each gait cycle of the manual part-level labeled silhouettes, the LDM parameters for a silhouette are estimated by processing each individual segment one by one. As suggested in [22], some parameters, such as the limb orientations, are more closely related to the way one walks and hence they are more important to gait recognition than the others, such as the width parameters. Therefore, the limb orientation parameters are estimated first using robust algorithms for high accuracy.

3.1.1. *Estimation of limb orientations*

For reliable estimation of the limb orientations (θ_{LTH}, θ_{RTH}, θ_{LLL}, θ_{RLL}, θ_{LUA}, θ_{RUA}, θ_{LLA}, and θ_{RLA}), it is proposed in [22] to estimate them from reliable edge orientations, that is, they are estimated from either the front or the back edges only, decided by the current stance (pose/phase). For instance, the front (back) edges are more reliable when the limbs are in front (at back) of the torso. The number of reliable edge pixels is denoted by R. This method of estimation through reliable body part information extends the leading edge method in [18] so that noise due to loose cloths are greatly reduced. The mean-shift algorithm [27], a powerful kernel-based algorithm for nonparametric mode-seeking, is applied in the joint spatial-orientation domain, and the different scales in the two domains are taken care of by using different kernel sizes for different domains. This algorithm is applied to the reliable edges of each limb individually with a preprocessing by a standard Gaussian lowpass filter to reduce noise. Let an edge pixel feature vector $\mathbf{p}_i = [\mathbf{p}_i^s; p_i^o]$, where $\mathbf{p}_i^s$ is the spatial coordinate vector of 2×1 and p_i^o is the local orientation value, estimated through the gradient. Denote by $\{\mathbf{p}_i\}_{i=1:R}$ the R reliable edge pixel feature vectors. Their modes $\{\mathbf{q}_{i,c}\}_{i=1:R}$ (defined similarly) are sought by iteratively computing

$$\mathbf{q}_{i,j+1} = \frac{\sum_{i=1}^{R} \mathbf{p}_i \cdot k\left(\left\|\left(\mathbf{q}_{i,j}^s - \mathbf{p}_i^s\right)/h_s\right\|^2\right) \cdot k\left(\left\|\left(q_{i,j}^o - p_i^o\right)/h_o\right\|^2\right)}{\sum_{i=1}^{R} k\left(\left\|\left(\mathbf{q}_{i,j}^s - \mathbf{p}_i^s\right)/h_s\right\|^2\right) \cdot k\left(\left\|\left(q_{i,j}^o - p_i^o\right)/h_o\right\|^2\right)} \tag{2}$$

until convergence, where $k(\cdot)$ is a kernel, h_s and h_o are the kernel bandwidths for the spatial and orientation domains, respectively, with the initialization $\mathbf{q}_{i,1} = \mathbf{p}_i$. The modes (points of convergence) are sorted in descending order based on the number of points converged to it. The dominant modes (modes at the top of the sorted list) represent body part orientations and the insignificant modes (modes at the bottom of the sorted list) are ignored.

This estimation process is illustrated in Figure 5. Figure 5(a) shows the edges of one arm and our algorithm is applied to its front edge since it is in front of the torso. The orientations (in degrees) of the front edge points are shown in Figure 5(b) and the converged orientation values for each point are shown in Figure 5(c). After the mode sorting, two dominant (top) modes (for upper arm and lower arm) are retained and they are shown in Figure 5(d) where the converged point positions are highlighted by setting their orientation values to a larger number (140 degree).

3.1.2. *Estimation of other parameters*

With the limb orientations and positions estimated, the joint (e.g., elbow, shoulder, knee) positions can be determined easily and the lengths (L_{UA}, L_{LA}, L_{TH}, and L_{LL}) and widths (W_K

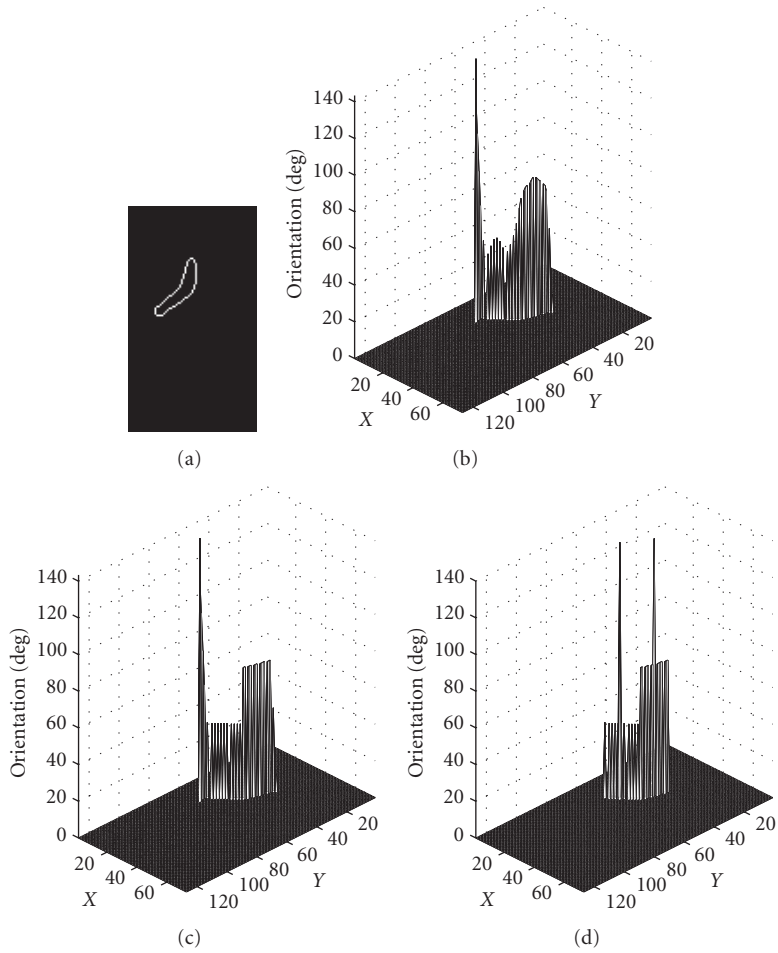


Figure 5: Illustration of limb orientation estimation through mean-shift. (a) The edges of one arm. (b) The orientation versus the spatial coordinate for the front edge in (a). (c) The orientation versus the spatial coordinates after mean-shift. (d) The position of the dominant modes are highlighted by setting their orientation values to 140 degree.

and W_A) of upper and lower limbs are estimated from them using simple geometry, as discussed in [22]. The torso width (W_T), torso length (L_T), and global position (x_G, y_G) are estimated from the bounding box of the torso segment. For the head, the "head top" (the top point of the labeled head) and the "front face" (the left most point of the labeled head) points are estimated through Gaussian filtering and averaging. These two points determine the head size (L_H) and the head center, partly eliminating the effects of hair styles. The neck joint angle (θ_H) can then be estimated from the head center and the neck joint position (estimated from the torso). The shoulder displacement ($dx_{\text{Sh}}, dy_{\text{Sh}}$) is determined from the difference between the neck and the shoulder joint positions.

3.1.3. Postprocessing of the estimations

Due to the imperfection of manual labeling and the pose recovery algorithm in Sections 3.1.1 and 3.1.2, the estimated LDM parameters may not vary smoothly and they need to be smoothed within a gait sequence, since according to biomechanics studies [13], during walking, body segments generally enjoy smooth transition and abrupt (or even unrealistic) change of body segment orientations/positions is not expected. The two-step postprocessing procedure proposed in [22] is modified here. The first step still applies a number of constraints such as the interframe parameter variation limits and the body part orientation limits. The head size (L_H) is fixed to be the median over a cycle and the interdependence between orientations of the same limbs are enforced to realistic values by respecting the following conditions:

$$\begin{aligned} \theta_{\text{LTH}} &\le \theta_{\text{LLL}}, \quad & \theta_{\text{RTH}} &\le \theta_{\text{RLL}}, \\ \theta_{\text{LUA}} &\ge \theta_{\text{LLA}}, \quad & \theta_{\text{RUA}} &\ge \theta_{\text{RLA}}. \end{aligned} \tag{3}$$

In the second step of postprocessing, a moving average filter of window size n is again applied to the parameter sequences, while a parameter sequence is expanded through circular shifting before the filtering and truncated accordingly after the filtering to avoid poor filtering at the two ends (the boundaries).

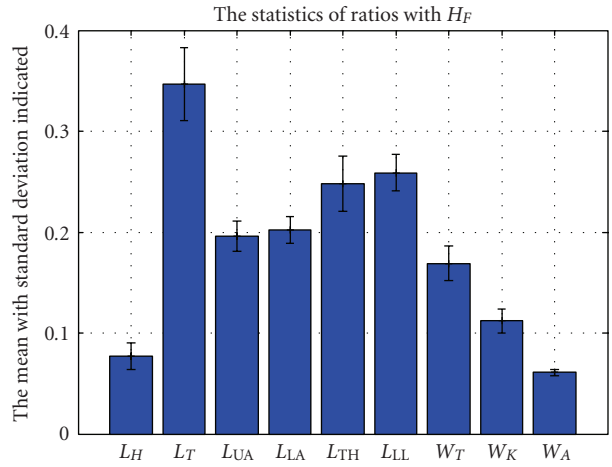


FIGURE 6: The means and standard deviations of the ratios of the length and width parameters over the full-body height for the gallery set of manual silhouettes.

3.2. Pose estimation from automatically extracted silhouettes

In practice, the pose recovery process needs to be automatic and it is infeasible to obtain silhouettes manually. Therefore, an automatic silhouette extraction algorithm is required to produce silhouettes for pose recovery.

3.2.1. Coarse-to-fine automatic silhouette extraction

In [28], we have developed a localized coarse-to-fine algorithm for efficient and accurate pedestrian localization and silhouette extraction for the gait challenge data sets. The coarse detection phase is simple and fast. It locates the target quickly based on temporal differences and some knowledge on the human target such as the shape and the motion of the subject. Based on this coarse detection, the fine detection phase applies a robust background subtraction algorithm based on Markov thresholds [29] to the coarse target regions and the detection obtained is further processed to produce the final results. In the robust background subtraction algorithm [29], the silhouettes of moving objects are extracted from a stationary background using Markov random fields (MRF) of binary segmentation variates so that the spatial and temporal dependencies imposed by moving objects on their images are exploited.

3.2.2. Shape parameter estimation

As pointed out in [22], since the shape (length and width) parameters are largely affected by cloths and the silhouette extraction algorithm used, they are not considered as gait dynamics for practical automatic model-based gait recognition, which is to be shown in the experiments (Section 5). Therefore, coarse estimations can be used for these LDM parameters. The statistics of the ratios of these parameters to the silhouette height H_F are studied for the gallery set of manual silhouettes and the standard deviations in these values are found to be quite low, as shown in Figure 6, where the standard deviations are indicated by the error bars. Therefore, fixed ratios to the height of the silhouette are used in the shape parameter estimations for the automatically extracted silhouettes as in [22], based on the gallery set of manual silhouettes.

3.2.3. Automatic silhouette information extraction

With the help from the ideal proportions of the human eight-head-high figure in drawing [30], the following information is extracted for the LDM parameter (pose) estimation from the automatically extracted silhouettes; (more detailed information regarding body segment proportions from anthropometry is available in [14], where body segments are expressed as a fraction of body height, however, the eight-head figure is simpler and more practical for the application of vision-based gait analysis/recognition at a distance):

(i) the silhouette height H_F, the first row $y_{\min}$, and the last row $y_{\max}$ of the silhouette;

(ii) the center column c_H of the first $H_F/8$ rows (for the head position);

(iii) the center column of the waist c_W is obtained as the average column position of the rows of the torso portion (rows $H_F/8$ to $H_F/2$) with widths within a limited deviation (± 0.3) from the expected width ($0.169 \cdot H_F$) of the torso portion (to avoid distraction by arms); in case that the torso portion is heavily missing, more rows from the below (leg portion) are added until a certain number (5) of rows within the limits are found, these conditions are relaxed further in case of failure;

(iv) the limb spatial-orientation domain modes and the number of points converged to each mode of the front and back edges are obtained through the mean-shift procedure described in Section 3.1.1 for the left and right lower legs (last $H_F/8$ rows) and the left and right lower arms (rows $3 \cdot H_F/8$ to $5 \cdot H_F/8$). For the upper arms (rows $H_F/8$ to $3 \cdot H_F/8$), due to the significant collusion with the torso in silhouettes, similar information is extracted only for the front edge of the left upper arm and the back edge of the right upper arm.

3.2.4. Position and orientation parameter estimation

The silhouette information extracted in the previous section is used for the estimation of the position and orientation parameters. The global position is determined as

$$x_G = c_W, \qquad y_G = y_{\min} + L_T + 2 \cdot L_H. \tag{4}$$

The head orientation θ_H is then calculated through estimating the neck joint $(x_G, y_G - L_T)$ and the head centroid $(c_H, y_{\min} + L_H)$.

Next, the limb orientations are estimated. The left or right limb orientations in this section refer to the orientations estimated for the left or right limb in the silhouettes, respectively. The next section will discuss the correct labeling of the actual left and right limbs for a subject.

Input: The gallery gait dynamics $\mathbf{X}_g \in \mathbb{R}^{T_g \times P}, g = 1, \ldots, G$
Algorithm:
Initialize $\mathbf{D}_1(g,c) = 1/(G(G-1))$, $\mathbf{D}_1(g,g) = 0$,
Do for $p = 1{:}\, P$:
(1) Get hypothesis $\{h_p(\mathbf{x}_{p_g}, c) \in [0,1]\}$ by scaling $\{\mathrm{DTW}(\mathbf{x}_{p_g}, \mathbf{x}_{p_c}), c = 1, \ldots, C\}$.
(2) Calculate ϵ_p, the pseudo-loss of h_p, from (6) and set $\beta_p = \epsilon_p/(1-\epsilon_p)$.
(3) Update $\mathbf{D}_p : \mathbf{D}_{p+1}(g,c) = \mathbf{D}_p(g,c) \cdot \beta_p^{(1/2)(1+h_p(\mathbf{x}_{pg},g)-h_p(\mathbf{x}_{pg},c))}$, and normalize it: $\mathbf{D}_{p+1}(g,c) = (\mathbf{D}_{p+1}(g,c)/(\sum_g \sum_c \mathbf{D}_{p+1}(g,c)))$.
Output: The final hypothesis: $h_{\mathrm{fin}}(\mathbf{X}) = \arg\max_c \sum_{p=1}^{P} (\log(1/\beta_p)) h_p(\mathbf{x}_p, c)$.

Algorithm 1: Combination of the LDM parameters for gait recognition.

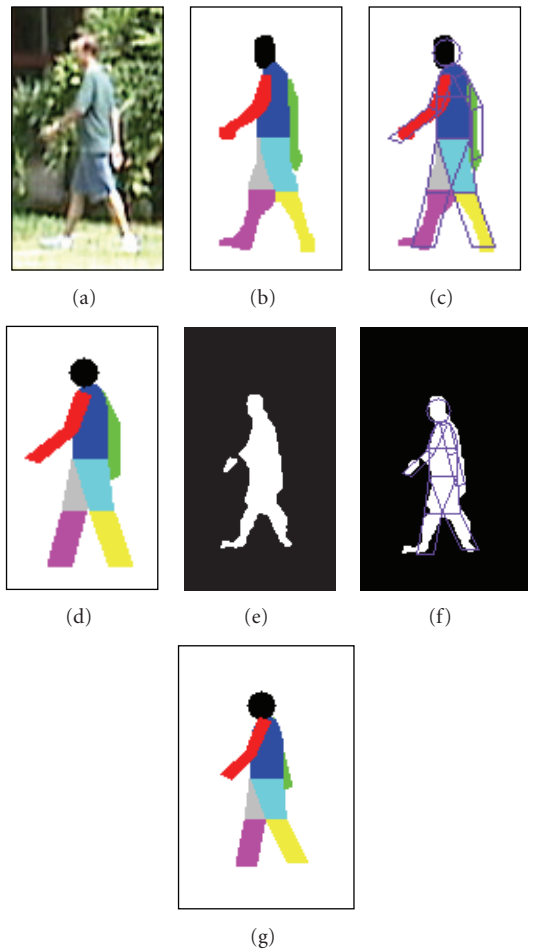


Figure 7: An example of human body pose recovery: (a) the raw image frame, (b) the manual silhouette, (c) the recovered LDM overlaid on the manual silhouette, (d) the reconstructed silhouette for the manual silhouette, (e) the automatically extracted silhouette (auto-silhouette), (e) the recovered LDM overlaid on the auto-silhouette, (f) the reconstructed silhouette for the auto-silhouette.

For the lower leg orientations (θ_{LLL} and θ_{RLL}), if the difference of the front and back edge estimations exceeds a threshold T_{LL} (15) and they have similar number of converged points, the estimations that will result in smaller changes are chosen, compared to the estimations in the last frame. Otherwise, the front and back edge estimations are merged using weighted average if their difference is less than the threshold T_{LL}. If none of these two cases is true, the estimation that has a larger number of points converged to it is taken. A bias of N_{LL} (5) points is applied to the estimation for the reliable edge, that is, N_{LL} is added to the number of converged points of the front edge for the left lower leg and to that of the back edge for the right lower leg. The lower arm orientations (θ_{LLA} and θ_{RLA}) are estimated similarly.

The row number of the left and right knees is set to row $(y_{\max} - H_F/4)$ of the silhouette. Since the lower leg orientations are estimated and the points on the lower legs (the positions) are also available, the knee positions are determined through simple geometry. The thigh orientations (θ_{LTH} and θ_{RTH}) are then calculated from the hip joint position (x_G, y_G) and the knee joint positions. The upper arm orientations (θ_{LUA} and θ_{RUA}) are set to the estimations from Section 3.2.3.

The shoulder displacement $(dx_{\mathrm{Sh}}, dy_{\mathrm{Sh}})$ is estimated from the left arm since the right arm is mostly severely occluded when walking from the right to the left. The points (positions) on the upper and lower left arms together with their estimated orientations determine the elbow position. The shoulder position can then be calculated based on L_{UA}, θ_{LUA} and the elbow position and it is compared with the neck joint position to give dx_{Sh} and dy_{Sh}.

The constraints described in the first step of postprocessing in Section 3.1.3 are enforced in the estimation above. A number of rules are applied to improve the results and they are not described here to save space, for example, when one leg is almost straight (the thigh and the lower leg have the same orientation) and its orientation differs 90 degree by a large amount (15 degree), the other leg should be close to straight too.

3.2.5. *Limb switching detection for correct labeling of left and right*

In previous section, the orientations for limbs are estimated without considering their actual labeling of left or right. This problem needs to be addressed for accurate pose recovery. Without loss of generality, it is assumed that in the first frame, the left and right legs are "switched," that is, the left leg is on the right and the right leg is on the left and we attempt to label the limbs in subsequent frames correctly. The opposite case (the left leg is on the left and the right leg is on

the right) can be tested similarly and the one results in better performance can be selected in practice.

To determine when the thighs and lower legs switch, the variations of respective lower-limb orientations are examined. From our knowledge, in normal gait, the arms have the opposite "switching" mode. The arms switch in opposite direction of the thighs. In addition, we set the minimum time interval between two successive switches to be 0.37 second, which is equivalent to a minimum number of frames of 11 for a 30 frames per second (fps) video.

A number of conditions are examined first to determine when the lower legs switch.

(i) If switched, the sum of the changes in the left and right lower leg orientations (compared with those in the previous frame) is lowered by a certain amount Δ_{LL} (30).
(ii) When the lower leg with thigh at the back (right) is almost vertical (90 ± 5 degree) in the previous frame, its orientation (in degree) is decreasing instead of increasing. This condition is set by observing the movement of the lower legs in crossing.
(iii) When the thighs are just switched, the sum of the changes in the left and right lower leg orientations (compared with those in the previous frame) is less than a certain amount Δ_{LL} if the lower legs are switched.
(iv) None of the above three conditions are satisfied after the thighs have been switched for 0.13 second (4 frames for a 30 fps video).

Similarly, thigh switching is determined by examining the following conditions.

(i) Either thigh orientation is within 90 ± 15 degree or the lower legs are just switched.
(ii) If the thighs are switched, the sum of the changes of the left and right thigh orientations is less than a certain amount Δ_{TS} (28).
(iii) The differences of the left and right thigh orientations are less than a certain amount Δ_D (25) in the previous frame and in this frame.
(iv) The thigh orientation difference is increasing (decreasing) in the previous frames but it is decreasing (increasing) in this frame.
(v) A thigh orientation is within 90 ± 5 degree in the previous frame, and it is increasing (decreasing) in previous frames but it is decreasing (increasing) in this frame.
(vi) If the lower legs are switched, the sum of the changes of the left and right lower leg orientations is less than a certain amount Δ_{LS} (38).
(vii) The column number of the right knee minus that of the left knee is less than -3.

Finally, the estimations are smoothed through the two-step postprocessing described in Section 3.1.3.

3.3. Comments on the computational complexity

It can be seen from the above that with silhouettes as the input, the LDM pose recovery takes a rule-based approach to incorporate human knowledge into the algorithm, rather than the popular tracking-based approach [17]. Most of the calculations are simple geometric, with the only exceptions to be the mean-shift procedure, which is a very efficient algorithm, and the lowpass filtering procedures. Therefore, the proposed algorithm is very efficient compared to the tracking algorithm in [17] based on particle-filtering, which is a sample-based probabilistic tracking algorithm with heavy computational cost. For example, in experiments, the pose estimation from 10005 automatically extracted silhouettes (with average size of 204.12×113.69) took only 94.798 seconds on a 3.2 GHz Pentium 4-based PC (implemented in C++), which is equivalent to a processing speed of more than 105 frames per second, which is much faster than the commonly used 30/25 fps video capturing speed. An additional benefit is that incorrect estimations of the parameters, due to the challenges in outdoor setting, do not lead to tracking failures.

4. LDM-BASED GAIT RECOGNITION THROUGH DYNAMIC TIME WARPING

From a gait cycle, the LDM parameters are estimated using the pose recovery algorithms in previous section for recognition. Let $\mathbf{X} \in \mathbb{R}^{T \times P}$ denote the LDM parameters describing the gait dynamics in a gait cycle, where T is the number of frames (silhouettes) in the gait cycle and P is the number of LDM parameters. The LDM parameters are arranged in the order as shown in Table 1. Thus, $\mathbf{X}(t, p)$ denotes the value of the pth LDM parameter in the tth frame and the sequence for the pth LDM parameter is denoted as $\mathbf{x}_p \in \mathbb{R}^{T \times 1}$. For the automatic LDM-based gait recognition, the maximum P is 12 since the LDM parameters for $p > 12$ (the shape parameters) are proportional to the full-body height ($p = 9$). For gait recognition from the manual silhouettes, the maximum P is 21. Since, in this work, there is only one cycle for each subject, the number of classes C equals to the number of samples G for the gallery set $C = G$.

For the LDM-based gait recognition, the first problem to be solved is the calculation of the distance between two sequences of the same LDM parameter,for example, $\mathbf{x}_{p_1} \in \mathbb{R}^{T_1 \times 1}$ and $\mathbf{x}_{p_2} \in \mathbb{R}^{T_2 \times 1}$. Since there is only one cycle for each subject, a simple direct template matching strategy, the dynamic time warping (DTW), is adopted here. The DTW is an algorithm for measuring the similarity between two sequences that may vary in time or speed based on dynamic programming [31] and it has been applied to gait recognition in [32–34]. To calculate the distance between two sequences, for example, a gallery sequence and a probe sequence, of possibly different lengths (e.g., $T_1 \neq T_2$) through DTW, all distances between the gallery sequence point and the probe sequence point are computed and an optimal "warping" path with the minimum accumulated distance, denoted as DTW $(\mathbf{x}_{p_1}, \mathbf{x}_{p_2})$, is determined. A warping path maps the time axis of a sequence to the time axis of the other sequence. The start and end points of a warping path are fixed and the monotonicity of the time-warping path is enforced. In addition, the warping path will not skip any point. Euclidean distance

Table 1: The arrangement of the LDM parameters.

p	1	2	3	4	5	6	7	8	9	10	11	12
LDM	θ_{LTH}	θ_{LLL}	θ_{RTH}	θ_{RLL}	θ_{LUA}	θ_{LLA}	θ_{RUA}	θ_{RLA}	H_F	θ_H	dx_{Sh}	dy_{Sh}
p	13	14	15	16	17	18	19	20	21	—	—	—
LDM	L_T	L_H	L_{UA}	L_{LA}	L_{TH}	L_{LL}	W_T	W_K	W_A	—	—	—

Table 2: The four key probe sets.

Probe	Difference with the gallery	Number of subjects
B	Shoe	41
D	Surface	70
H	Briefcase	70
K	Time (including shoe, cloths, etc.)	33

is used here for measuring the distance between two points. The details of the DTW algorithm can be found in [31].

A distance is calculated for each parameter and a combination scheme is needed to integrate the gait dynamics (parameters) of each body part for gait recognition. The combination scheme used in AdaBoost.M2 [35] is adopted here to weight the different LDM parameters properly, as shown in Algorithm 1. AdaBoost is an ensemble-based method to combine a set of (weak) base learners, where a base learner produces a hypothesis for the input sample. As seen in Algorithm 1, the DTW distance calculator, with proper scaling, is employed as the base learner in this work. Let $\mathbf{X}_g \in \mathbb{R}^{T_g \times P}, g = 1, \ldots, G$, be the LDM gallery gait dynamics, where G is the number of gallery subject. In the training phase, each parameter sequence $\mathbf{x}_{p_g}$ is matched against all the sequences for the same parameter $\mathbf{x}_{p_c}, c = 1, \ldots, C$, using DTW and the matching scores are scaled to the range of $[0, 1]$, which are the outputs of the hypothesis h_p. Similar to AdaBoost.M2, the pseudoloss ϵ_p is defined with respect to the so-called mislabel distribution $\mathbf{D}_p(g, c)$ [35], where p is the LDM parameter index here. A mislabel is a pair (g, c) where g is the index of a training sample and c is an incorrect label associated with the sample g. Let B be the set of all mislabels:

$$B = \{(g, c) : g = 1, \ldots, G, c \neq g\}. \tag{5}$$

The pseudoloss ϵ_p of the pth hypothesis h_p with respect to $\mathbf{D}_p(g, c)$ is given by [35]

$$\epsilon_p = \frac{1}{2} \sum_{(g,c) \in B} \mathbf{D}_p(g, c)\left(1 - h_p(\mathbf{x}_{p_g}, g) + h_p(\mathbf{x}_{p_g}, c)\right). \tag{6}$$

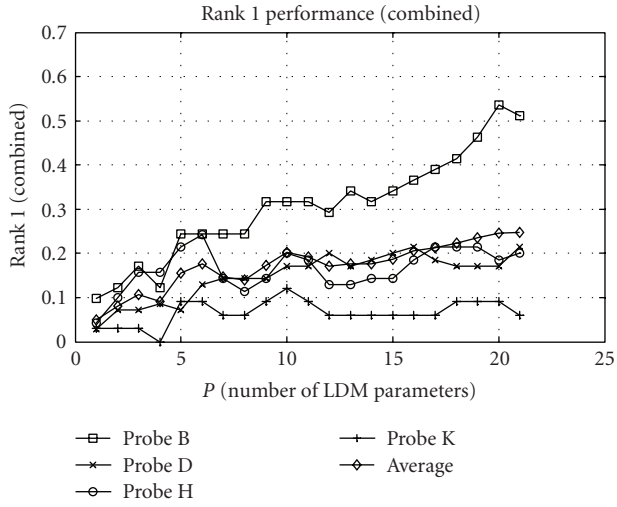
Following the procedures in Algorithm 1, $\log(1/\beta_p)$, the weight of each LDM parameter p, is determined.

5. EXPERIMENTAL RESULTS

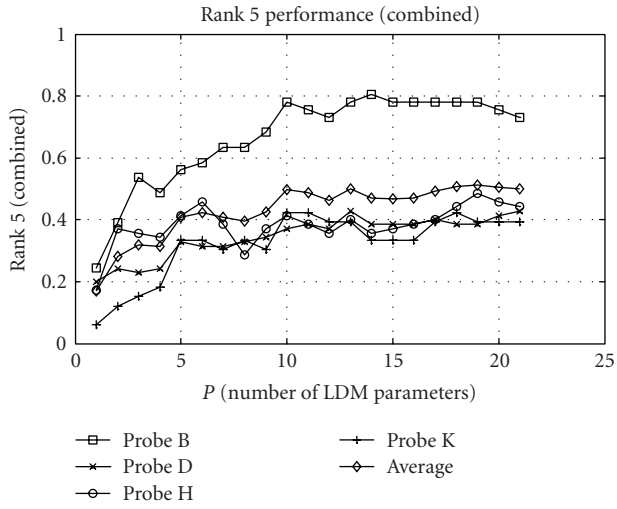
The experiments on LDM-based gait recognition were carried out on the manual silhouettes created in [16] and the corresponding subset in the original "gait challenge" data sets, which contains human gait sequences captured under various outdoor conditions. The five key experiments of this subset are gallery, probes B, D, H, and K. The differences of the probe sets compared to the gallery set are listed in Table 2, together with the number of subjects in each set. The number of subjects in the gallery set is 71. Each sequence for a subject consists of one gait cycle of about 30∼40 frames, and there are 10005 frames in the 285 sequences. For the mean-shift algorithm in the pose recovery procedure, we set the kernel bandwidths $h_s = 15$ and $h_o = 10$ and use the kernel with the Epanechnikov profile [27]. For the running average filter, a window size $n = 7$ is used.

An example of the human body pose recovery for the manual silhouettes and automatically extracted silhouettes are shown in Figure 7, and the qualitative and quantitative evaluations of the human body pose recovery results are reported in [22], where the reconstructed silhouettes from the automatically extracted silhouettes have good resemblance with those from the manual silhouettes. This paper concentrates on the gait recognition results. The rank 1 and rank 5 results are presented, where rank k results report the percentage of probe subjects whose true match in the gallery set was in the top k matches. The results on the manual silhouettes help us to understand the effects of the body part dynamics as well as the shapes when they can be reliably estimated and the results on the automatically extracted silhouettes investigate the performance in practical automatic gait recognition.

Table 3 compares the rank 1 and rank 5 gait recognition performance of the baseline algorithm on the manual silhouettes (denoted as BL-Man) [15], the component-based gait recognition (CBGR) on the manual silhouettes (CBGR-Man) [36], the LDM-based algorithm on the manual silhouettes (LDM-Man), and the LDM-based algorithm on the automatically extracted silhouettes (LDM-Aut). (Note that the baseline results cited here are consistent with those in [15, 16, 36], but different from those in [1, 9] since the experimental data is different. There are two essential differences. The first difference is that in this work, there is only one cycle in each sequence, while in [1, 9], there are multiple cycles. The second difference is that in this work, gait recognition is from the part-level gait dynamics, while in [1, 9], as shown in [15], correlated errors/noise is a contributing factor in recognition performance.) The BL-Man algorithm matches the whole silhouettes directly while the CBGR-Man algorithm uses componentwise matching. Since they both treat gait as holistic patterns, we refer to them as the holistic approach. For the LDM-Man and LDM-Aut algorithms, the indicated recognition rates are obtained with $P = 21$ (all LDM parameters) and $P = 11$, respectively. The shoulder vertical displacement dy_{Sh} ($p = 12$) is excluded for the best performing LDM-Aut algorithm (resulting in $P = 11$) because the estimated dy_{Sh} in this case is not helpful in identification, as



(a) Rank 1 versus the number of LDM parameter



(b) Rank 5 versus the number of LDM parameter

Figure 8: The gait recognition performance for the manual silhouettes.

to be shown in Figure 9 (Section 5.2). The recognition rates reported in brackets for the LDM-Man are obtained with the same set of LDM parameters as in the LDM-Aut, that is, $P = 11$.

5.1. Gait recognition with the manual silhouettes

The detailed gait recognition results using the manual silhouettes are reported in Figure 8, where the averaged recognition rates are shown in thicker lines. There are several interesting observations from the results. First, the inclusion of the arm dynamics ($p = 5, 6, 7, 8$), the dynamic of the full-body height ($p = 9$), and the head dynamic ($p = 10$) improves the average recognition rates, indicating that the leg dynamics ($p = 1, 2, 3, 4$) are not the only information useful for model-based gait recognition. A similar observation is made recently in [36] for the holistic approach, where the arm silhouettes are found to have similar discriminative power as the thigh silhouettes.

Secondly, it is observed that the length and width parameters concerning the shape provide little useful discriminative information when clothing is changed, that is, probe K. Furthermore, for the rank 5 recognition rate (Figure 8(b)), including the shape parameters ($p > 12$) results in little improvement on the performance, indicating that shapes are not reliable features for practical model-based gait recognition, even if the body-part level silhouettes can be obtained ideally, which agrees with intuition since shapes are largely affected by clothing. On the other hand, from Figure 8(a), the rank 1 recognition rate for probe B, which is captured under the conditions with the same clothing and only difference in shoes, benefits the most from the inclusion of the shape parameters.

Another interesting observation is that for probe H, where the subject carries a briefcase with the right arm, the inclusion of the right arm dynamics ($p = 7$ and $p = 8$) results in performance degradation for both rank 1 and rank 5, which can be explained by the fact that the right arms are not moving in the "usual way." This information could be utilized to improve the gait recognition results through, for example, excluding the right arm dynamics if it is known or detected that the subject is carrying some objects (while there is no carrying in the gallery). Moreover, these clues drawn from the LDM gait dynamics could be useful in applications other than gait recognition, such as gait analysis for the detection of carrying objects or other abnormalities.

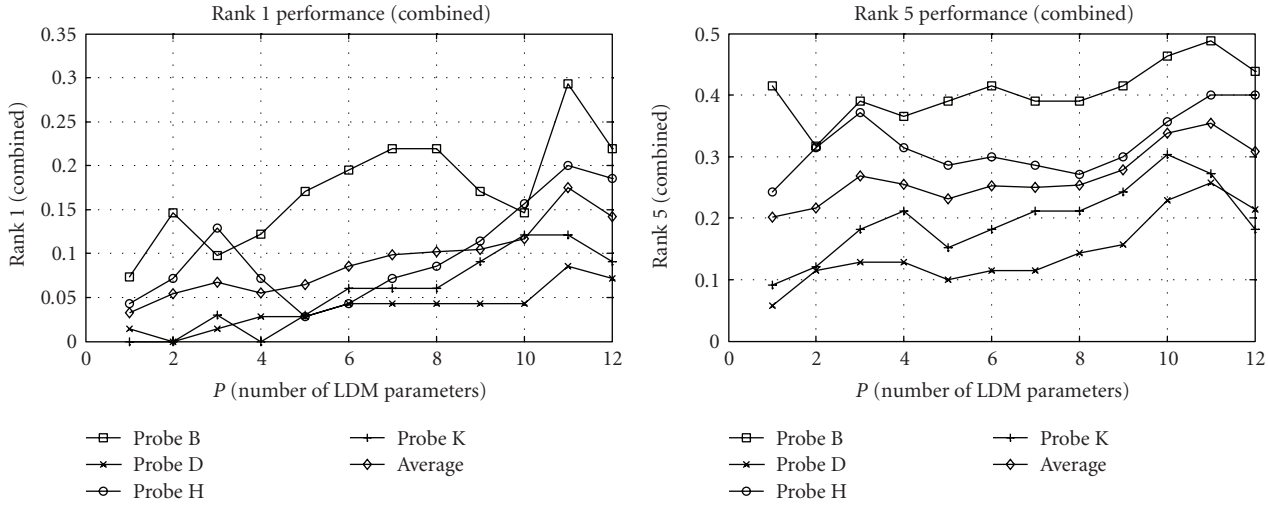
5.2. Gait recognition with automatically extracted silhouettes

In [15], studies on the holistic recognition show that "the low performance under the impact of surface and time variation can not be explained by the silhouette quality," based on the fact that the noisy silhouettes (extracted semi-automatically) outperforms the manual (clean) silhouettes due to correlated errors/noise acting as discriminative information. Different from [15], the LDM-based gait recognition achieves better results on the manual (clean) silhouettes than on the automatically extracted (noisy) silhouettes, especially in the rank 5 performance, as shown in Table 3, suggesting that more accurate silhouette extraction and body pose recovery algorithms could improve the performance of automatic model-based gait recognition, which agrees with our common belief.

It is also observed that the LDM-based results on the automatically extracted silhouettes are the worst on probe D, where the rank 1 and rank 5 recognition rates are only about half of those on the manual silhouettes. This difference is due to the fact that our model-based gait recognition relies purely on the gait dynamics and it seems that a different surface significantly affects the accurate estimation of the LDM parameters. This suggests that by knowing the fact that the surface is different, the silhouette extraction and body pose recovery algorithms should be modified to adapt to (to work better on) the different surface. Another interesting observation is that for probe H (with briefcase), the LDM-based approaches

Table 3: Comparison of the LDM-based and holistic gait recognition algorithms.

	Rank 1 recognition rate (%)				Rank 5 recognition rate (%)			
Probe	BL-Man	CBGR-Man	LDM-Man	LDM-Aut	BL-Man	CBGR-Man	LDM-Man	LDM-Aut
B	46	49	51(32)	29	66	78	73(76)	49
D	23	26	21(17)	9	39	53	43(39)	26
H	9	16	20(19)	20	36	46	44(39)	40
K	12	15	6(9)	12	39	39	39(42)	27
Average	23	27	25 (19)	18	45	54	50 (49)	35



(a) Rank 1 versus the number of LDM parameter.

(b) Rank 5 versus the number of LDM parameter.

Figure 9: The gait recognition performance for the automatically extracted silhouettes.

(both LDM-Man and LDM-Aut) outperform the holistic approach in rank 1, especially the BL-Man, implying that the proposed LDM-based gait recognition approach suits situations with "abnormality" better than the holistic approach.

Figure 9 depicts the detailed gait recognition results for the automatically extracted silhouettes and the averaged recognition rates are shown in thicker lines too. Similar to the results on the manual silhouettes, the inclusion of the dynamics of the arms, the full-body height, the head, and even the shoulder's horizontal dynamic ($p = 11$) improves the average recognition rates, indicating again that there are other gait dynamics other than the leg dynamics that are useful for model-based gait recognition. In addition, it is worthy to note from Table 3 (the results in brackets for LDM-Man and the results for LDM-Aut) that for probe K, which is captured with six months time difference from the gallery set, the inclusion of the shape information degrades both the rank 1 and rank 5 recognition rates from 9 to 6 and from 42 to 39, respectively. While the recognition results for probes B, D, and H, captured with the same clothing, improves with the shape parameters, which confirms again that shape information works only for the same (or similar) clothing.

6. CONCLUSIONS

Recently, gait recognition has attracted much attention for its potential in surveillance and security applications. In order to study the gait recognition performance from the dynamics of various body parts, this paper extends the layered deformable model first introduced in [22] for model-based gait recognition, with 22 parameters defining the body part lengths, widths, positions, and orientations. Algorithms are developed to recover human body poses from the manual silhouettes and the automatically extracted silhouettes. The robust and efficient mean-shift procedure, average filtering, and simple geometric operations are employed, and domain knowledge (including estimation through reliable edges, anthropometry, and biomechanics constraints) is incorporated to achieve accurate recovery. Next, the dynamic time warping is employed for matching parameter sequences and the contributions from each parameter are weighted as in AdaBoost.M2. The experimental results on a subset of the gait challenge data sets show that the LDM-based gait recognition achieves comparable results (and better results in some cases) as the holistic approach. It is demonstrated that the upper-body dynamics, including the arms, the head, and the shoulders, are important for the identification of individuals as well. Furthermore, the LDM serves as a powerful tool for

the analysis of different factors contributing to the gait recognition performance under different conditions and it can be extended for other gait-related applications. In conclusion, the LDM-based approach proposed in this paper advances the technology of automatic model-based gait recognition.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their insightful comments. The authors would also like to thank Professor Sudeep Sarkar from the University of South Florida for kindly providing them with the manual silhouettes as well as the gait challenge data sets. This work is partially supported by the Ontario Centres of Excellence through the Communications and Information Technology Ontario Partnership Program and the Bell University Labs at the University of Toronto. An earlier version of this paper was presented at the Seventh IEEE International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 10–12 April 2006.

REFERENCES

[1] M. S. Nixon and J. N. Carter, "Automatic recognition by gait," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 2013–2024, 2006.

[2] A. Kale, A. Sundaresan, A. N. Rajagopalan, et al., "Identification of humans using gait," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1163–1173, 2004.

[3] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, "Gait recognition: a challenging signal processing technology for biometrics identification," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 78–90, 2005.

[4] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1505–1518, 2003.

[5] G. Johansson, "Visual motion perception," *Scientific American*, vol. 232, no. 6, pp. 76–88, 1975.

[6] J. Cutting and L. Kozlowski, "Recognizing friends by their walk: gait perception without familiarity cues," *Bulletin of the Psychonomic Society*, vol. 9, no. 5, pp. 353–356, 1977.

[7] C. D. Barclay, J. E. Cutting, and L. T. Kozlowski, "Temporal and spatial factors in gait perception that influence gender recognition," *Perception and Psychophysics*, vol. 23, no. 2, pp. 145–152, 1978.

[8] S. V. Stevenage, M. S. Nixon, and K. Vince, "Visual analysis of gait as a cue to identity," *Applied Cognitive Psychology*, vol. 13, no. 6, pp. 513–526, 2000.

[9] S. Sarkar, P. J. Phillips, Z. Liu, I. Robledo, P. Grother, and K. W. Bowyer, "The human ID gait challenge problem: data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.

[10] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "Gait recognition using linear time normalization," *Pattern Recognition*, vol. 39, no. 5, pp. 969–979, 2006.

[11] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, 2006.

[12] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," *IEEE Transactions on Neural Networks*, vol. 19, no. 1, 2008.

[13] D. A. Winter, *The Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological*, University of Waterloo Press, Waterloo, Ontario, Canada, 2nd edition, 1991.

[14] D. A. Winter, *The Biomechanics and Motor Control of Human Movement*, John Wiley & Sons, New York, NY, USA, 2005.

[15] Z. Liu and S. Sarkar, "Effect of silhouette quality on hard problems in gait recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 2, pp. 170–183, 2005.

[16] Z. Liu, L. Malave, A. Osuntogun, P. Sudhakar, and S. Sarkar, "Toward understanding the limits of gait recognition," in *Proceedings of SPIE Processings Defense Security Symposium: Biometric Technology for Human Identification*, pp. 195–205, April 2004.

[17] L. Wang, H. Ning, T. Tan, and W. Hu, "Fusion of static and dynamic body biometrics for gait recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 149–158, 2004.

[18] C. Y. Yam, M. S. Nixon, and J. N. Carter, "Automated person recognition by walking and running via model-based approaches," *Pattern Recognition*, vol. 37, no. 5, pp. 1057–1072, 2004.

[19] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic extraction and description of human gait models for recognition purposes," *Computer Vision and Image Understanding*, vol. 90, no. 1, pp. 1–41, 2003.

[20] D. K. Wagg and M. S. Nixon, "On automated model-based extraction and analysis of gait," in *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 11–16, Seoul, Korea, May 2004.

[21] R. Zhang, C. Vogler, and D. Metaxas, "Human gait recognition," in *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop*, pp. 18–18, Washington, DC, USA, June 2004.

[22] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A layered deformable model for gait analysis," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, pp. 249–256, Southampton, UK, April 2006.

[23] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," in *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 148–155, Washington, DC, USA, May 2002.

[24] N. Cuntoor, A. Kale, and R. Chellappa, "Combining multiple evidences for gait recognition," in *Proceedings of the IEEE International Conference on Multimedia & Expo (ICME '06)*, vol. 3, pp. 113–116, Toronto, Ontario, Canada, July 2006.

[25] N. Jojic and B. J. Frey, "Learning flexible sprites in video layers," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I199–I206, Kauai, Hawaii, USA, 2001.

[26] J. Zhang, R. Collins, and Y. Liu, "Representation and matching of articulated shapes," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. II342–II349, Washington, DC, USA, July 2004.

[27] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.

[28] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Coarse-to-fine pedestrian localization and silhouette extraction for the gait challenge data sets," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '06)*, vol. 2006, pp. 1009–1012, Toronto, Ontario, Canada, 2006.

[29] J. Migdal and W. E. L. Grimson, "Background subtraction us-

ing Markov thresholds," in *Proceedings of the IEEE Workshop on Motion and Video Computing (MOTION '05)*, pp. 58–65, 2005.

[30] A. Zaidenberg, *Drawing the Figure from Top to Toe*, World, 1966.

[31] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, USA, 2000.

[32] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos, "Gait recognition using dynamic time warping," in *Proceedings of the IEEE 6th Workshop on Multimedia Signal Processing (WMSP '04)*, pp. 263–266, Siena, Italy, 2004.

[33] A. Kale, N. Cuntoor, B. Yegnanarayana, A. N. Rajagopalan, and R. Chellappa, "Gait analysis for human identification," in *Proceedings of the International Conference on Audio and Video Based Person Authentication*, Guildford, UK, 2003.

[34] R. Tanawongsuwan and A. Bobick, "Gait recognition from time-normalized joint-angle trajectories in the walking plane," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. II726–II731, Kauai, Hawaii, USA, 2001.

[35] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proceedings of the 13th International Conference on Machine Learning*, pp. 148–156, Desenzano sul Garda, Italy, June 1996.

[36] N. V. Boulgouris and Z. X. Chi, "Human gait recognition based on matching of body components," *Pattern Recognition*, vol. 40, no. 6, pp. 1763–1770, 2007.