Model Selection Within a Bayesian Approach to Extraction of Walker Motion

Ziheng Zhou, Robert I. Damper and Adam Prügel-Bennett
Information: Signals, Images, Systems (ISIS) Research Group
School of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, UK
{zz02r|rid|apb}@ecs.soton.ac.uk

Abstract

Extracting articulated motion of walking people in image sequences remains a challenge, particularly when we take into account the changes caused by carried objects or the severe motion occlusions by clothing (e.g., a long skirt, a trench coat, etc). In this paper, we propose a Bayesian framework capable of handling such uncertainties by exploiting our strong prior knowledge of how humans walk. In this work, the strong prior is built from a simple articulated model, which can be easily modified to cater for situations such as walkers wearing clothing that obscures the limbs. A model selection process is built into the framework to determine the body configuration of the walker in the given sequence automatically. The statistics of the parameters describing a basic walker are learned from data and the Bayesian framework then allows us to 'bootstrap' to accurate motion extraction on the images of walkers with extra body configurations. We demonstrate our approach on the data of walkers with rucksacks, skirts and trench coats. Results are quantified in terms of average pixel error between automatically extracted body points and corresponding points labelled by hand.

1. Introduction

Extracting full-body articulated motion of walking people from monocular video sequences poses a severe challenge for computer vision. Very often, the motion will be highly articulated and the object will have complex changing boundaries [1]. This problem is even more difficult if we take into account the large variations of body appearances of walkers. Figure 1 gives some examples of the image data used in this work. Figure 1(a) shows a walker with 'normal' appearance. Fig. 1(b) shows a walker carrying a rucksack. It can be seen that the apparent upper-body shape of the walker is significantly changed by the rucksack. In Figs. 1(c) and 1(d), we see the severe occlusions of the

(a) normal

(b) rucksack
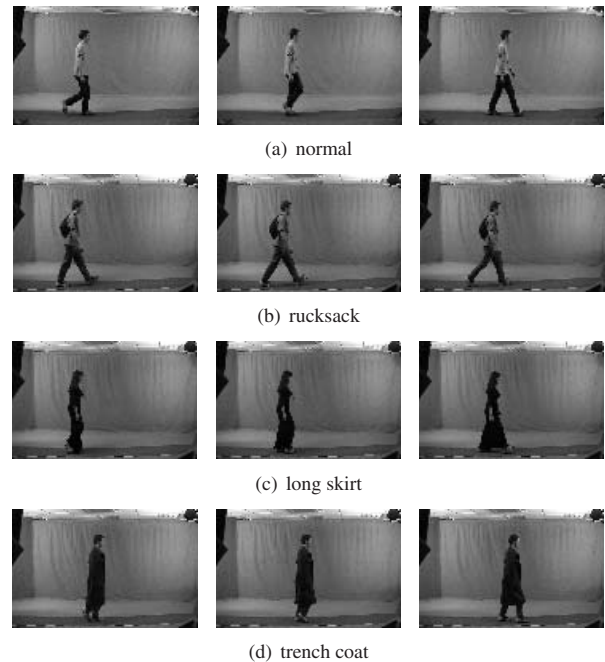
(c) long skirt

(d) trench coat

Figure 1. Examples of the images from which we attempt to extract the full-body articulated motion of the walker. The data include (a) walkers with normal body appearance, (b) carrying a rucksack, and (c) wearing a long skirt or (d) a trench coat.

motion which can be caused by clothing (a long skirt and a trench coat in the two situations). Such occlusions would probably defeat any motion-extraction system which does not build in prior knowledge of how human beings walk.

In this paper, we present a system which is capable of dealing with the uncertainties inherent in this application so as to extract the articulated motion of walking people consistently from the image sequences. To achieve this, we have adopted a Bayesian framework exploiting strong prior information of how humans walk. We know that the human body is highly articulated and, hence, we introduce a

Figure 2. Motion-extraction procedure in the framework. Model selection: determines the body configuration of the walker. Phase location: detects which part of the gait cycle an image comes from. Posterior maximization: extracts the articulated motion.

basic two-dimensional articulated model for a walker with a normal appearance. This can be easily extended by adding more parameters controlling the appearance of extra-body objects such as a carried rucksack or long clothing. At this stage, we consider only walkers moving perpendicular to the camera as this is typical of the current state of the art, and because we have a large database collected under these conditions. For each kind of body configuration, a hidden Markov model (HMM) has been built to capture the dynamic movement of the body. Based on the HMMs, a model selection process is carried out to determine the body configuration of the walker in the given sequence.

The Bayesian framework also allows us to build strong priors using the statistics of the parameters of the basic model and those controlling the extra-body appearance. Such a strategy makes the motion extraction consistent and robust by allowing us to build in our strong prior knowledge of walking (e.g., the temporal variations of the leg movement) no matter what body appearance the walker has. Because we aim to extract the fullest information about articulated motion that the image quality allows (rather than to track walkers in real time), it is highly advantageous to process the whole sequence of images. In this way, and combined with the use of prior information, we can obtain a global solution that, for instance, copes with extreme occlusions.

2. Previous Work

Capturing motion of walking people is an important task in the domain of computer vision, and many systems have been developed to solve this problem [2, 3, 4]. In 1983, Hogg [5] carried out pioneering work using an articulated model to track walking individuals in video sequences. Landmarks in the subsequent development include the publications of Baumberg and Hogg [6], Rohr [7], Ju, Black and Yacoob [8] and Toyama and Blake [9]. Recently, many efforts have been made to extract the contours of walking people accurately for higher levels of human gait analysis (e.g., gait recognition) [10, 11]. The contours are found either by an active shape model [10] or sequential Monte Carlo sampling [11]. Although some good results have been reported, none of the systems yet provides an effective way to extract walking motion consistently for walkers with large variations of their body appearance.

In this paper, we propose a Bayesian framework for extracting the articulated motion of walking people, where

the walkers could have various body appearances as shown in Fig. 1. The framework is capable of dealing with the uncertainties (significant changes of body appearance and limb occlusions) and capturing motion consistently.

3. Proposed Framework

The philosophy of this work is that we impose a consistent framework based on a Bayesian methodology. This has many advantages. It allows us to introduce strong prior knowledge about people and the way they walk (in the form of an articulated model), which eases the problem of learning about walking humans from data. Using Bayesian updating, it also allows us to 'bootstrap' from learning on the training data to coping with the uncertainties of the body appearance. In this section, after outlining the data used, we detail the component parts of the proposed solution. All of them alone are simple, but fitting them together in a Bayesian framework provides us an overall solution for this difficult problem which has not been tackled by any system so far. Figure 2 shows how the articulated motion is extracted from image sequences within the proposed framework.

3.1. Data Description

This research has used a subset of the Southampton Human Identification at a Distance (HiD) database [12]. The subjects walked at a normal speed and were viewed from the side. The image sequences were filmed at 25 frames per second. We used the indoor and supplemental data of the database in this work. Both were filmed under laboratory conditions. Silhouettes were extracted from these data using chroma-key techniques and used in our system as the image input from which articulated motion was extracted. The indoor dataset contains image sequences of just over 100 walkers with normal body configurations. The supplemental dataset contains image sequences of walkers carrying bags, rucksacks, or wearing clothing such as long skirts or trenchcoats which obscure the legs.

3.2. Articulated Model

Because we are viewing walkers from the side, we can build a very simple two-dimensional articulated model for a basic walker. Figure 3 shows this model. There are 16 parameters in total which divide into two groups: those determining the sizes of the body parts which remain con-
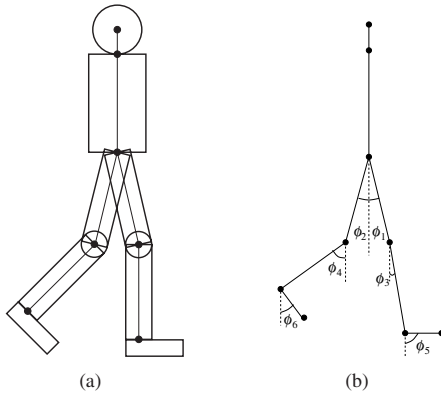
Figure 3. (a) the articulated model and (b) definition of the angles of the model.
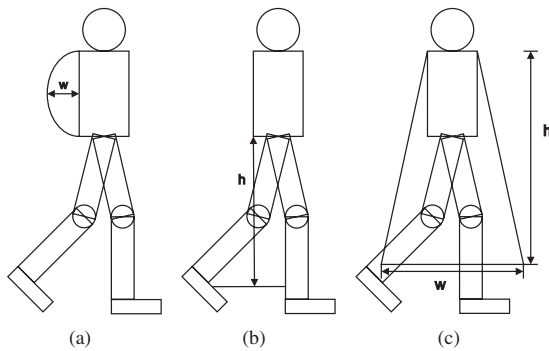

Figure 4. Extended articulated models for (a) a walker carrying a rucksack or (b) wearing a long skirt or (c) a trench coat.

stant for all images in the sequence; and the angles between the body parts, which vary from frame to frame. We refer to these as *static* and *dynamic* variables respectively.

Clearly, the model is only a crude approximation to a real walker. No account is taken of perspective, parts of the body such as the neck, hands and arms are missing, and there is no distinction made between the left and right sides of the body. We rather distinguish between the front and back legs and arms.

The basic model can be simply modified to cater for walkers with different body configurations. Figures 4(a), 4(b) and 4(c) show the extended models for walkers with a rucksack, a long skirt and a trench coat, respectively. A rucksack is represented by a half ellipse, a long skirt by filling the gap between the legs, and a trapezoid is added to stand for a trench coat. A maximum of two more static parameters is added to the basic model to cater for the altered body appearance.

The motion information is extracted by finding the best set of model parameters to fit any given silhouette. In the Bayesian framework, this means determining the likelihood of the image given the model. To achieve this, we generate

from the model a silhouette of the appropriate size. This 'model silhouette' is then matched against the observed data silhouette. In the following, we denote the set of parameters of the articulated model as θ and the model silhouette as $\mathcal{I}(\theta)$.
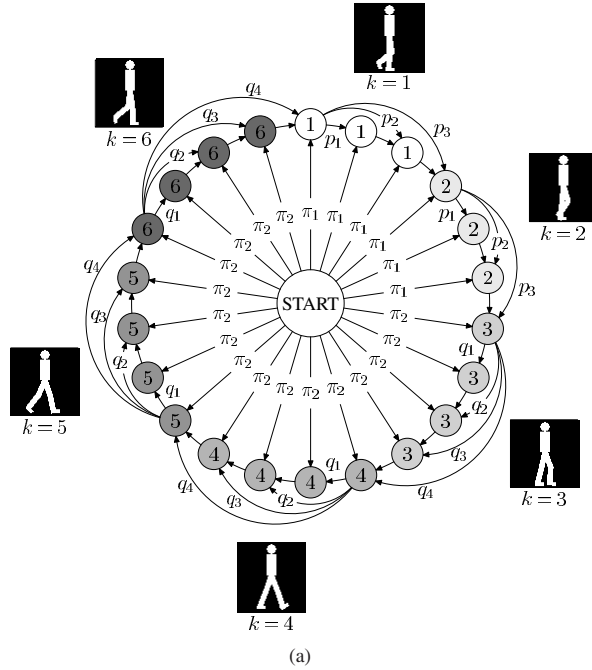
3.3. HMM Decoding

Our prior knowledge of walking tells us that the dynamic parameters of the articulated model (i.e., joint angles) are a strong function of the phase within the gait cycle. We divide a walking cycle into $K = 6$ sections and use a hidden Markov model (HMM) to capture the dynamic information. Since they provide a natural framework for processing sequential stochastic data, HMMs have been very popular in previous work on human gait (e.g., [13, 14]).

Because the walking cycle is a rather simple case of cyclical motion, we took the approach of defining a hand-crafted HMM whose parameters are chosen to reflect this prior knowledge rather than being automatically learned from data. This prior information has been extracted from hand-labelled data: 3 randomly chosen sequences for each of 7 walkers. The definition of the HMM went through some iterations of improvement; its final form (along with its associated parameters) is shown in Figure 5. Here, we used tied states rather than single states with self-transitions to model the state occupancy better.

Estimating the observation probability of an image belonging to a particular state is non-trivial because of the high dimensionality of the image space. We use the pdf projection theorem [15, 16] to learn the probability densities since it can project densities in a high-dimensional raw data space from a low-dimensional feature space avoiding the curse of dimensionality. Here the feature is defined as the distance between the image and a prototype silhouette for the corresponding section of the walk (see the silhouettes in Figure 5(a) for example). The prototypes embody the average articulated-model parameters extracted from that section of the walk. The parameters for the prototype are denoted by $\{\theta_k\}_{k=1}^K$. Starting from some initial values, the average parameters were iteratively refined by the bootstrapping method described below (Section 3.6). Note that Figure 5 only illustrates the prototypes for walkers with a normal appearance.

To measure the distance between the images and the prototypes we use the chamfer distance, which has been popular in object detection and tracking [17, 9]. We denote the chamfer distance between an image x and the model prototype as $\rho(x, \mathcal{I}(\theta_k))$ where function ρ computes the chamfer distance. According to the pdf projection theorem, we compute the probability of an image x being emitted from state k as:

(a)

probabilities	values
π_1	0.033
π_2	0.050
p_1	0.475
p_2	0.475
p_3	0.050
q_1	0.050
q_2	0.530
q_3	0.370
q_4	0.050

(b)

Figure 5. Hand-crafted hidden Markov model used to locate images within the gait cycle: (a) shows the architecture of the HMM with sections labelled by an image of the corresponding prototype; (b) lists the values of transition probabilities.

$$p(\boldsymbol{x}|k) = g(\boldsymbol{x}) \frac{p\Big(\rho(\boldsymbol{x}, \mathcal{I}(\boldsymbol{\theta}_k))\big|k\Big)}{\sum_j p\Big(\rho(\boldsymbol{x}, \mathcal{I}(\boldsymbol{\theta}_k))\big|j\Big) p(j)} \quad (1)$$

where we model the probability of a chamfer distance $\rho(\boldsymbol{x}, \mathcal{I}(\boldsymbol{\theta}_k))$ being equal to r given that $\boldsymbol{x}$ comes from the jth section of the walk by a gamma distribution:

$$p(r|j) = \frac{b^a r^{a-1} \mathrm{e}^{-br}}{\Gamma(a)}. \quad (2)$$

The function $g(\boldsymbol{x})$ is the same for all states k so does not need to be determined. Thus, to use the pdf projection theorem we need to know the parameters a and b for each

of the gamma distributions describing the spread of chamfer matches between the images in section j of the walk and the prototype for section k—giving a total of K^2 gamma distributions. These parameters will be different for walkers with different body appearances.

3.3.1 Model Selection

Given an image sequence, we have to determine the body configuration of the walker before the motion extraction process. This can be simply done once we have constructed the HMMs. Assuming that we know all N types of possible body configurations of walkers, we can build an HMM for each of them. We denote the image sequence as $\boldsymbol{X}$ and the HMMs as $\{\lambda_n\}_{n=1}^N$. These HMMs share the same structure, but with different observation probability densities (different prototype silhouettes), which can be well learned by the pdf projection theorem. For HMM λ_n, we calculate the probability $P(\boldsymbol{X}|\lambda_n)$ using the forward probability calculation algorithm [18]. We choose the body configuration for the walker corresponding to the largest probability.

3.3.2 Locating Phases in the Walking Cycle

Having determined the body configuration, we use the HMM corresponding to that configuration to locate the phase of the images in the walking cycle. This is done by searching the optimal state sequence given an image sequence and the HMM. We implemented the Viterbi algorithm [18] to find the optimal state sequence, indicating which section in the walking cycle an image comes from. This information is used to ease the search for the best-fit models.

3.4. Probabilistic Fitting

To fit the articulated model to silhouette data, the parameters have to be chosen so that the model overlaps with the silhouettes in some sensible way. We use a Bayesian framework to solve this problem, by taking the fit between the model and the silhouettes as a likelihood measure of the data given the model. We can then use a prior on the parameters to ensure that we have a feasible model.

Given silhouette $\boldsymbol{x}$ from section k of the walking phase, a model $\mathcal{M}_k$, and model parameters $\boldsymbol{\theta}$, we define the likelihood probability as:

$$p(\boldsymbol{x}|\boldsymbol{\theta}, \mathcal{M}_k) \propto \mathrm{e}^{-b\,\rho(\boldsymbol{x}, \mathcal{I}(\boldsymbol{\theta}))} \quad (3)$$

where b is a Lagrange multiplier to be determined empirically. We make the additional assumption that the number of images with chamfer distance $\rho(\boldsymbol{x}, \mathcal{I}(\boldsymbol{\theta}))$ equal to r grows as a polynomial r^{a-1}. That is, the distribution of chamfer distances is given by:

$$p\left(\rho(\boldsymbol{x}, \mathcal{I}(\boldsymbol{\theta})) = r \Big| \boldsymbol{\theta}\right) \propto r^{a-1} \mathrm{e}^{-br} \qquad (4)$$

i.e., a gamma distribution. The parameters a and b can be found by fitting empirical data using maximum likelihood.

We assume a Gaussian prior for the parameters:

$$p(\boldsymbol{\theta}|\mathcal{M}_k) \propto \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_k)^{\mathrm{T}} \boldsymbol{C}_k^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta}_k)\right) \quad (5)$$

where $\boldsymbol{\theta}_k$ and $\boldsymbol{C}_k$ are the averages and covariances for the parameters in section k. The average parameters are the same as those of our prototypes described in the previous subsection. The covariance matrix is learned using Bayesian updating as described in Section 3.6. In the next subsection, we discuss the practical details of finding the parameters which maximise the posterior probability.
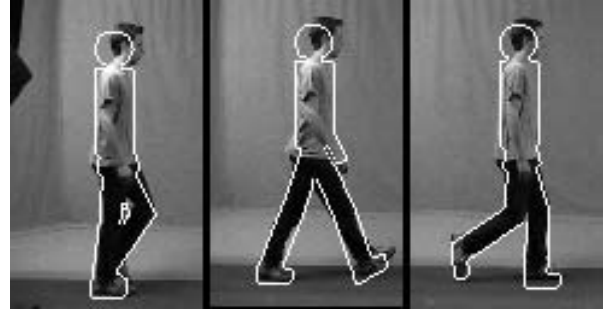
3.5. Optimising Parameters

As previously stated, we separate the parameters into static and dynamic parameters. To fit the static parameters, the images are grouped in a sequence according to which part of the walking cycle they come from. We then fit a model which maximises the posterior for each group where the static parameters are the same for each group while the dynamic parameters vary. Having fitted the static parameters, the dynamic parameters are then fitted frame-by-frame. Two different initial parameter settings are used for the optimisation. These are the average dynamic parameters that were calculated when fitting the static parameters and a linear prediction of the parameters from the two previous frames. We choose the fit that has larger posterior probability.
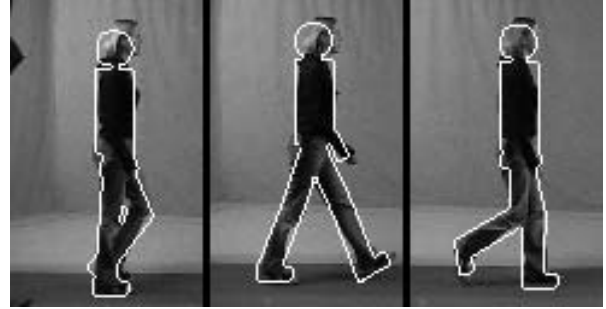
3.6. Building Priors

We first learn the parameter means $\{\boldsymbol{\mu}_k^{\mathrm{b}}\}_{k=1}^K$ and covariances $\{\mathbf{C}_k^{\mathrm{b}}\}_{k=1}^K$ for the basic articulated model for walkers with normal appearances. This is done in three steps:

1. For walkers with normal appearance, 7 subjects were chosen from the database and 3 sequences used for each of them. Every sequence was sorted manually into $K = 6$ groups. The parameters of the model were evaluated as described previously except that we used likelihoods rather than posteriors. Hard constraints were used to obtain a feasible model. For instance, we limited the head radius in the interval from 3 to 5 pixels. The raw estimation of the means and covariances were calculated based on the extracted parameters.

2. The parameters of the HMM were learned from manually-labelled sequences.

(a)

(b)

Figure 6. Typical fitting results for normal walkers: (a) male and (b) female. The outlines of the extracted models are superimposed on the relevant frames in the sequence.

3. The system was given 100 sequences from 50 walkers to refine our prior knowledge.

Figure 6 shows typical fitting results for normal walkers using the refined priors. As can be seen, the extracted models fit the walkers in the images very well.

For the walker with a different body appearance (e.g., wearing a long skirt), we extend the basic model by adding 1 or 2 static parameters. We denote the mean and covariance of the added parameters as $\boldsymbol{\mu}^{\mathrm{a}}$ and $\mathbf{C}^{\mathrm{a}}$. The means $\{\boldsymbol{\mu}_k\}_{k=1}^K$ and covariances $\{\mathbf{C}_k\}_{k=1}^K$ of the parameters of the new model can be simply computed as:

$$\boldsymbol{\theta}_k = \begin{pmatrix} \boldsymbol{\mu}^{\mathrm{a}} \\ \boldsymbol{\mu}_k^{\mathrm{b}} \end{pmatrix} \qquad (6)$$

$$\boldsymbol{C}_k = \begin{pmatrix} \boldsymbol{\mu}^{\mathrm{a}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C}_k^{\mathrm{b}} \end{pmatrix}. \qquad (7)$$

We can see that the refined statistics of the parameters of the basic model are always built in the priors. Such a way of building priors makes the motion extraction consistent and robust since the behavior of the model is constrained all the time.

For a new body configuration, we have to:

1. extend the basic articulated model,

2. hand-label some selected training sequences, as we did for the normal walkers,

3. measure μ^{a} and $\mathbf{C}^{\mathrm{a}}$,

4. estimate the observation probability densities using the new means, and

5. calculate the multiplier b in the likelihood definition.

4. Results

The system has been tested on image data from the Southampton HiD gait database containing four kinds of body configurations of the walkers: normal appearance, carrying a rucksack, wearing a long skirt, and wearing a trench coat. In the rest of this section, we will refer to these data as normal, rucksack, skirt and trench coat, respectively. Background subtraction was performed on the initial color images [12] and the obtained silhouettes were normalized to be of size (120×120) pixels with the subjects centered.

4.1. Model Selection

As mentioned in Section 3.3.1, this framework has the ability to determine automatically the body configuration of the walker in the given image sequence. We built the HMMs $\{\lambda_n\}_{n=1}^{4}$ for the data and learned the observation probability densities for the states. For the normal, rucksack and trench coat data, 10 subjects were chosen from the database and 10 sequences were used for each of them. Since there was only one subject in the skirt data section of the database, we chose 20 sequences for the test. For each sequence $\boldsymbol{X}_i$, we computed the probabilities $P(\boldsymbol{X}_i|\lambda_n)$ to determine the body configuration of the walker in the images. The system selected the correct body configurations for all the testing sequences.

4.2. Fitting Results

After the model selection, we extracted the articulated motion of walkers using the methods described. Figure 7 shows some typical fitting results for the rucksack, skirt and trench coat data. The outlines of the extracted models are superimposed on the relevant frames of the sequences. It can be seen that the extended models fit the walkers with reasonable accuracy, which supports the idea of exploiting strong prior knowledge to compensate for the uncertainties (body-shape changes and limb occlusions) raised by various body configurations of the walkers.

These results, although typical, are illustrative (i.e., qualitative) only. To quantify the results, we manually marked the positions of five fiducial points, namely the hip, front/back knee and front/back ankle, in the original data. Although the manual labelling can never be perfect, we feel this is a reasonable, practical approximation to 'ground

truth'. The labelling was done for 4 sequences of 20 frames each from one walker for each condition (added rucksack, long skirt or trench coat). Distances were calculated from the hip, front/back knee and front/back ankle positions obtained from the extracted models to the fiducial points in the corresponding images. Table 1 shows the means and standard deviations of the distances computed for each of the five points plus the overall mean and standard deviation. As expected, for walkers wearing a long skirt or a trench coat, the errors at the knees are larger than at other points. In spite of this, the overall results (as judged visually) remain good. We believe the level of disruption of body shape which occurs in these sequences would defeat most current approaches to motion extraction for walking people.

5. Conclusions

We have described a Bayesian framework for extracting the articulated motion of walking people in video sequences. The framework has been demonstrated to be capable of dealing with various body configurations of the walkers. This is achieved by exploiting strong prior knowledge of human walking within a Bayesian formalism. We built a simple two-dimensional articulated model to represent a basic walker.

Such a model can be easily extended to cater for the changes of body appearances caused by carried objects or voluminous clothing. HMMs were used to capture the dynamic information of the walk in the walking cycle. Using the HMMs, we determined automatically the body configuration of the walker in a given sequence by a model-selection process. The framework allowed us to build strong priors based on the statistics of the parameters depicting basic walkers and those controlling the extra body appearances (e.g., a carried rucksack, a long skirt or a trench coat). Such a mechanism made the motion extraction consistent and robust.

Although encouraging results have been achieved, there are some limitations of the current system. First, the image sequences were recorded in a well-controlled environment, resulting in high-quality silhouettes. However, the motion occlusions caused by clothing are still realistic and challenging for any motion-extraction system. Testing the framework on outdoor image data with real-world noise is a priority for future work. Second, we have assumed that all subjects walked with their normal speed. For significantly slower or faster walking, it will undoubtedly be necessary to relearn the HMM parameters and priors because of the consequent change of walking style. Finally, we have attempted to prove the model-selection concept on some representative situations, where silhouettes are changed by carrying a rucksack or wearing a long skirt or a trench coat, but this does not cover the whole range of possibilities (e.g., walker may be carrying a parcel or briefcase). Further work
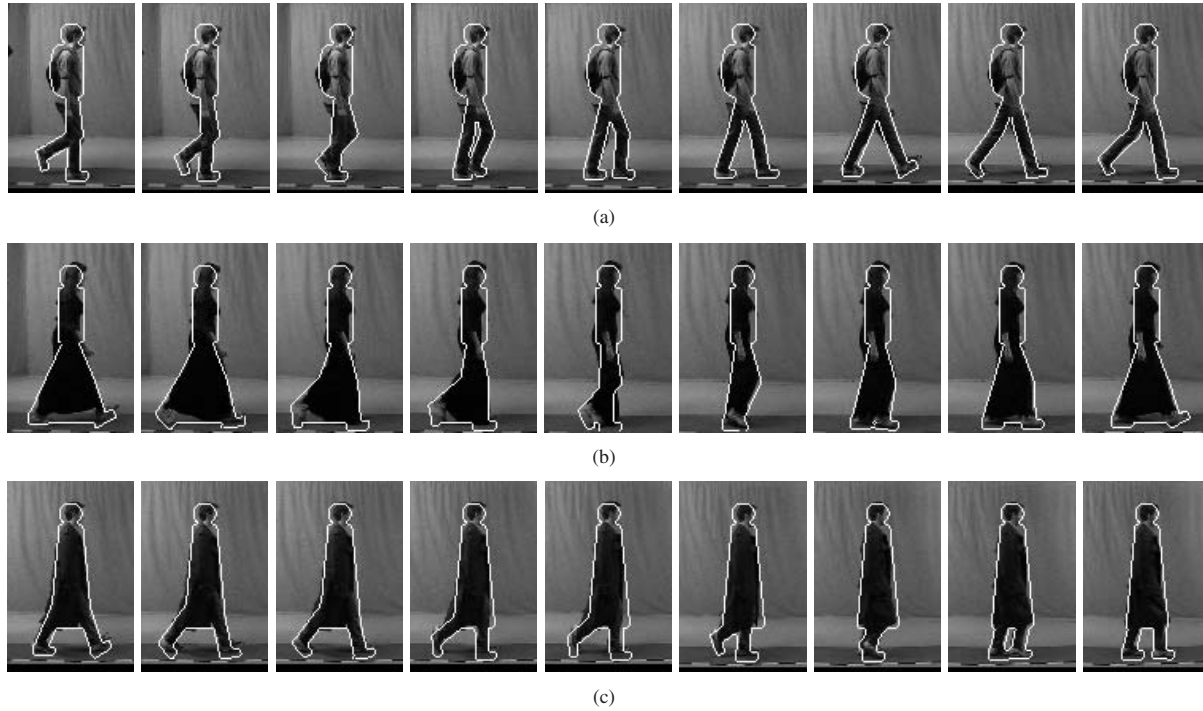
(a)

(b)

(c)

Figure 7. Typical fitting results: (a) carrying a rucksack, (b) wearing a long skirt, and (c) wearing a trench coat.

(pixel)		Hip	Front Knee	Front Ankle	Back Knee	Back Ankle	Overall
Rucksack	Mean	2.24	3.53	3.79	2.47	2.71	2.95
	Standard Deviation	1.57	1.80	1.39	1.58	1.40	1.66
Long	Mean	3.17	6.26	3.72	6.33	4.68	4.83
Skirt	Standard Deviation	2.19	2.70	2.48	2.64	2.41	2.80
Trench	Mean	2.01	6.19	3.63	5.23	3.01	4.02
Coat	Standard Deviation	1.04	1.63	2.24	1.71	3.13	2.55

Table 1. Means and standard deviations of the distances between the five points marked manually on the sequences with rucksack, long skirt or trench coat and those obtained by fitting the extended model.

is needed to extend the approach to deal with such cases.

The methods employed in this paper can be used to good effect in a variety of practical scenarios in which the walkers of interest can not reasonably be fitted with markers. The most obvious application is person identification by gait. Another might be tracking of an individual or individuals in a surveillance setting. With some modifications, the framework could be adapted to applications such as gesture recognition for sign language or cued speech communication for hearing-impaired people.

References

[1] C. Cédras and M. Shah, "Motion-based recognition: A survey," *Image and Vision Computing*, vol. 13, no. 2, pp. 129–155, 1995.

[2] J. K. Aggarwal and Q. Cai, "Human motion analysis: A review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428–440, 1999.

[3] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, 2001.

[4] L. Wang, W. Hu, and T. Tan, "Recent developments in human motion analysis," *Pattern Recognition*, vol. 36, no. 3, pp. 585–601, 2003.

[5] D. C. Hogg, "Model-based vision: A program to see a walking person," *Image and Vision Computing*, vol. 1, no. 1, pp. 5–20, 1983.

IEEE
COMPUTER
SOCIETY

[6] A. M. Baumberg and D. C. Hogg, "An efficient method for contour tracking using active shape models," in *Proceedings of IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, Austin, TX, 1994, pp. 194–199.

[7] K. Rohr, "Towards model-based recognition of human movement in image sequences," *CVGIP, Image Understanding*, vol. 59, no. 1, pp. 94–115, 1994.

[8] S. X. Ju, M. J. Black, and Y. Yacoob, "Cardbord people: a parameterized model of articulated motion," in *Proceedings of 2nd International Conference on Automatic Face and Gesture Recognition*, Killington, VT, 1996, pp. 38–44.

[9] K. Toyama and A. Blake, "Probabilistic tracking with exemplars in a metric space," *International Journal of Computer Vision*, vol. 48, no. 1, pp. 9–19, 2002.

[10] D. K. Wagg and M. S. Nixon, "Automated markerless extraction of walking people using deformable contour models," *Computer Animation and Virtual Worlds*, vol. 15, no. 3–4, pp. 399–406, 2004.

[11] J. Zhang, R. Collins, and Y. Liu, "Representation and matching of articulated shapes," in *Proceedings of the 2004 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'04)*, vol. II, Washington, DC, 2004, pp. 342–349.

[12] J. Shutler, M. Grant, M. S. Nixon, and J. N. Carter, "On a large sequence-based human gait database," in *Proceedings of 4th International Conference on Recent Advances in Soft Computing*, Nottingham, UK, 2002, pp. 66–72.

[13] D. Meyer, J. Pösl, and H. Niemann, "Gait classification with HMMs for trajectories of body parts extracted by mixture densities," in *Proceedings of British Machine Vision Conference (BMVC'98)*, Southampton, UK, 1998, pp. 459–468.

[14] X. Lan and D. P. Huttenlocher, "A unified spatio-temporal articulated model for tracking," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, Washington, DC, 2004, pp. 722–729.

[15] P. M. Baggenstoss, "Class-specific feature sets in classification," *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3428–3432, 1999.

[16] A. Thayananthan, R. Navaratnam, P. H. S. Torr, and R. Cipolla, "Likelihood models for template matching using the PDF projection theorem," in *Proceedings of British Machine Vision Conference (BMVC 2004)*, Kingston, UK, 2004, pagination unknown.

[17] G. Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 6, pp. 849–865, 1988.

[18] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–285, 1989.