Proceedings of the 2007 IEEE/RSJ International
Conference on Intelligent Robots and Systems
San Diego, CA, USA, Oct 29 - Nov 2, 2007

Laser Tracking of Human Body Motion Using Adaptive Shape Modeling

Dylan F. Glas, Takahiro Miyashita, Hiroshi Ishiguro, and Norihiro Hagita
ATR Intelligent Robotics and Communication Laboratories
2-2-2 Hikaridai, Seika-cho, Souraku-gun, Kyoto, 619-0288, Japan
{dylan,miyasita,ishiguro,hagita}@atr.jp

Abstract— **In this paper we present a method for determining body orientation and pose information from laser scanner data using particle filtering with an adaptive modeling algorithm. A parametric human shape model is recursively updated to fit observed data after each resampling step of the particle filter. This updated model is then used in the likelihood estimation step for the following iteration. This method has been implemented and tested by using a network of laser range finders to observe human subjects in a variety of interactions. We present results illustrating that our method can closely track torso and arm movements even with noisy and incomplete sensor data, and we show examples of body language primitives that can be observed from this orientation and positioning information.**

I. INTRODUCTION

Advances in sensing technology and computing power have pushed the frontier of robotics to a point where it is not only conceivable but expected that robots will soon be working among humans in a social context. In such applications, accurate position estimation and tracking of people in the vicinity of the robot are essential, both for safety and for smooth, context-aware interaction with humans.

Indeed, for robots to operate within society, it will be necessary for them to detect the subtle cues of gesture and body positioning that we naturally notice and respond to subconsciously. Although many existing robotic applications track the locations of humans, the wealth of information encoded in posture and movement is often discarded in the process.

We are developing methods of identifying a variety of human behaviors and understanding social context using data from ubiquitous sensor networks. In addition to data such as location, speed, and relative positioning within a group, pose-related information such as a person's body orientation can provide valuable insight into the dynamics of a social interaction.

Scanning laser range finders are a popular tool for human-tracking and navigation applications due to their precision, effective sensing distance, and ease of use. Rough human positions can easily be determined from laser scanner data, and techniques such as particle filters enable more robust position estimates to be achieved by reducing errors due to noise and occlusion.

In this paper, we propose an extension to this general position estimation technique, in which the body orientation and pose are tracked in addition to position and velocity. By incorporating a shape model into the tracking algorithm, highly accurate estimates of torso and arm position can be extracted, and we will show that these position estimates can provide information valuable to the understanding of social situations.

II. RELATED WORK

Human tracking is a well-developed field, and many aspects of the problem have already been explored extensively. Our work involves the use of particle filters to estimate human positions based on the nonlinear and noisy data of laser scan measurements. Particle filters are a common tool in the robotics community and have often been used in conjunction with laser scan data for the purposes of robot localization and mapping [1], [2]. A general overview of robotic applications of particle filters can be found in [3].

Much of the human-tracking research to date has been based on leg tracking [4], [5], [6]. Some reasons for this include simplicity of shape (legs are roughly circular and look the same from any angle) and visibility (legs are narrow and tend not to completely occlude objects behind them). Another consideration is that many robots use laser sensors for obstacle avoidance, and for that reason already have laser sensors mounted near the ground.

In our work, the laser sensors were part of an environmental sensor network used exclusively for human tracking, rather than robot-mounted sensors used for obstacle avoidance. For this reason, we were able to mount the sensors higher, above obstacles like furniture and ground clutter, to a height where the positions of the arms and torso could be clearly observed.

Although less common than leg-tracking, torso-level tracking is not without precedent in research. For example, a real-time torso-level human tracking system utilizing particle filters to process laser scan data was developed by Almeida et al. in [7]. Their system, however, was focused specifically on position tracking, whereas our work is concerned with observing body orientation and pose in addition to position.

III. HUMAN DETECTION AND TRACKING WITH ADAPTIVE SHAPE MODELING

A. Algorithm Overview

Our tracking technique is based on particle filtering, the basics of which will be very briefly explained here. For a more in-depth explanation, [8] provides a thorough treatment of particle filters and many other state-estimation techniques.

Particle-filtering is a method of estimating the state of a system by using a cloud of "particles", each of which

represents a hypothesis about that state. The following four-step procedure is performed at each iteration of a traditional particle filter.

1) **Update** The state of each particle is updated by applying an internal *motion model*, reflecting the dynamics of the system, to the previous state estimate. The motion model used in our work is described in Section III-D.
2) **Assign Weights** Particles are then assigned weights according to a *likelihood model*, which evaluates the likelihood of the state represented by each particle based on the latest sensor data. Our likelihood model is described in Section III-E.
3) **Estimate State** At this point, the state estimate is calculated, generally as a weighted average of the states of the particles.
4) **Resample** Finally, the particles are removed or propagated based on their weights to produce a new set of particles which more accurately reflects the true state of the system.

In this way, the cloud of particles converges on the most likely state and follows it over time.

One often-criticized shortcoming of particle filters is their poor performance in high-dimensional spaces. Typical particle filters used for people-tracking tend to use a four-dimensional state space to model the position (x, y) and velocity(v_x, v_y) of each person. Our human shape model, described in Section III-B, has six variables. Tracking them all in addition to position and velocity would more than double the dimensionality of the state space, severely hindering the efficiency of the particle filter.

To address this problem, we assume that the variations in the body shape model parameters from one iteration of the particle filter to the next are small enough that their effect on the position and velocity estimates are negligible. We consider this assumption to be generally justifiable (with one exception, explained below) since the subjects move slowly with respect to the update rate of the sensors (approximately 38 Hz in our case). This assumption allows us to use the shape model from the previous iteration when calculating likelihood weights for the particles, rather than trying to estimate body shape simultaneously with position.

Our extended particle filter algorithm is illustrated in Figure 1. A particle filter is used for tracking position and velocity, while the parameters of the shape model are updated after each state estimation step and incorporated into the likelihood calculations of the particle filter during the next iteration. The algorithm we use for updating the model parameters is described in Section III-F.

Note that there is one model parameter for which the above assumption does not hold. Due to the nonlinear nature of fitting a shape to incomplete and noisy data, we have observed that small changes in body orientation angle (θ in our model) can substantially affect the particle filter's position estimates. Since this violates our earlier assumption, θ cannot be treated as an independent variable. To address this problem, the likelihood calculations only consider the

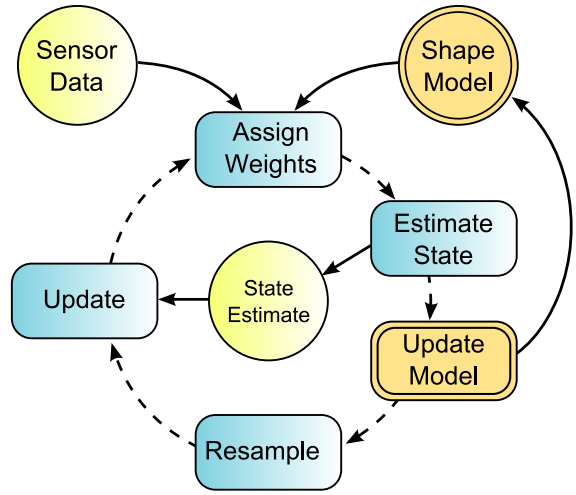

Fig. 1. Diagram of our particle filter implementation using an adaptive shape model. The shape model is updated after every state estimation step, and that updated model is used for shape-matching to determine particle weights during the next iteration.

optimal body orientation angle, which is treated as a pure function of position. During likelihood calculations, a fast rotational shape-matching algorithm is used to determine the body orientation angle θ_{opt} which maximizes the likelihood for each particle based on its (x,y) location. In this way, isolation of the body shape model from position and velocity is preserved, and the state space searched by the particle filter remains limited to four dimensions.

B. Human Shape Model

In our experience with laser scanning data, we have observed large variations in cross-sectional contour shape between subjects. This is due in part to individual differences in body shape, and also to differences in height. The scan plane of our sensors for this experiment was at a height of 90 cm, with the result that for taller subjects, the scan was taken roughly at wrist-height, whereas for shorter subjects, the scan plane was above elbow-height. Consequently, measured arm movement was more pronounced for the taller subjects.

Another factor in contour shape is the type of clothing being worn. For example, a loose shirt or a heavy coat can make a person's torso appear unusually large or asymmetrical, as can a backpack or purse.

Taking these factors into consideration, the amount of variation between subjects makes it difficult to develop a precise, yet generalizable, model. Thus a simple three-circle model was used for determining body orientation.

Our model is illustrated in Figure 2. A central, large circle represents the person's torso, and two smaller circles represent the arms. This model has six parameters which can be varied to best match a subject's cross-sectional body contour.

The parameters describing the state of this model are summarized in Table I. The two parameters of primary interest to us are θ and φ.

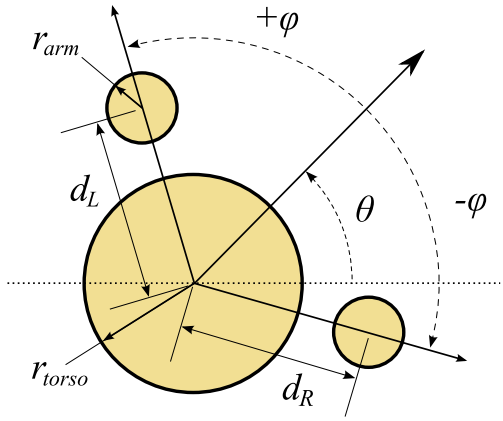

Fig. 2. Our three-circle model, with the six variable parameters indicated.

We have designated θ to represent the angle midway between the two arms. When a subject is standing still, this coincides with the direction of torso orientation. While the subject is walking, the swinging of the arms and torso cause θ to oscillate around the direction of motion.

The parameter φ represents the angle of separation between each arm and the center angle designated by θ. This tends not to vary far from 90°, as the arms swing in alternate directions during walking.

TABLE I
MODEL PARAMETERS

Parameter	Description
θ	Average direction of body orientation
φ	Arm separation angle
d_L	Distance of left arm from body
d_R	Distance of right arm from body
r_{arm}	Arm radius
r_{torso}	Torso radius

C. Data Association

The problem of associating detections with features can become quite complex, particularly with large numbers of objects and many occlusions. A variety of approaches have been developed for this problem, some of which are described in [5], [9], [10].

In our case, the humans move slowly with respect to the sensor scan rate, and there tend to be few ambiguities, so complex approaches such as these are not necessary. A simple nearest-neighbor matching is used for associating data points with humans. We will consider three aspects of data association: continuous tracking, detection of entry events, and detection of exit events.

1) Continuous Tracking: In our implementation, an individual particle filter is associated with each human upon detection. This particle filter is updated during each step and accordingly moves with the human it is tracking until it has been determined that the filter is no longer tracking the person.

To avoid the case of multiple particle filters tracking the same human, a repulsive component is included in the likelihood model. A value described by a Gaussian distribution centered around each human being tracked is subtracted from the likelihood assigned to each particle, effectively creating a repulsive force which prevents the convergence of two particle filters on a single human.

2) Entry Events: The mechanism for detecting untracked humans is based on a first-pass analysis of the raw sensor data. The range readings from each sensor are thresholded using a pre-calculated background model, and a set of candidate locations is created by identifying segments of contiguous foreground data points. The rules for this segmentation model approximate humans as ellipses with a minor axis of 15 cm and a major axis of 55 cm, taking into consideration the possibility of partial occlusions and using heuristic rules for outlier rejection. Finally, the center of each ellipse is identified as a human candidate.

Particularly for distant and partially-occluded readings, these detections can be unreliable due to noise and the simplicity of the estimation rules. To improve the reliability of detections, a 2-D windowing filter is applied to each candidate position. At each step, human candidates within a threshold distance of a windowing filter are absorbed into that filter, and filters with no associated human candidates are removed. Any filter which survives for a specified minimum number of time steps is considered to be reliable human position estimate, and only then is a particle filter assigned to track the human at that position.

By filtering the sensor data in this way, the incidence of false detections in noisy environments is greatly reduced.

3) Exit Events: The remaining task is the detection of the disappearance of a human. When a person exits the experimental area, no more data is available to reflect that person's presence, and the particle filter naturally disperses. When minimum probability or maximum statistical dispersion threshold values are crossed, the human is assumed to have left the area and the particle filter is removed.

D. Motion Model

The purpose of the motion model is to approximate the probability of a state $\mathbf{x}_t$ based on the previous state $\mathbf{x}_{t-1}$. The state vector we used to model the motion of humans includes four variables: x-position, y-position, walking speed, and direction of motion.

As has been observed in [11], the modeling of human motion presents difficulty because it is neither Brownian in nature, nor can it be modeled as a smooth linear function, since people may stop or change direction abruptly. Thus, as a compromise between the two, we modeled the (x, y) motion as the sum of a linear velocity vector and a Gaussian noise component.

E. Likelihood Model

The purpose of the likelihood model is to approximate the value of $p(z_t|\mathbf{x}_t^{[m]})$ for particle $m, (m = 1..M)$ and

measurement vector z_t taken at time step t of the particle filter.

Laser scan data provides two qualitatively distinct types of information useful for estimating human positions: *occupancy information*, indicating whether a certain point is occupied or empty, and *edge information*, indicating a contour which may correspond with the edge of a detected object. The likelihood model used here is expressed as the product of two terms: $p_{occ}(z_t|\mathbf{x}_t)$, based on occupancy information, and $p_{fit}(z_t|\mathbf{x}_t)$, based on a fit between a proposed human position and the observed edge data, as shown in Eq. 1:

$$p(z_t|\mathbf{x}_t^{[m]}) = p_{occ}(z_t|\mathbf{x}_t^{[m]}) \cdot p_{fit}(z_t|\mathbf{x}_t^{[m]}) \tag{1}$$

The occupancy information component is computed using a coverage grid, in which coverage information from all six sensors is aggregated. The edge information component is computed by estimating correctness-of-fit using a radial accumulator array. Both of these components are explained in detail below.

1) Coverage Grid: The coverage grid is a 5cm resolution grid representing the coverage of the experimental space by the laser scanners. In a real-world sensor network, noise due to calibration errors, measurement errors, and timing errors often produces paradoxical situations, such as the front and back edges of a person overlapping, or one sensor indicating that a region is empty where another sensor observes a person. Thus, a simple grid of binary values would not adequately represent the information contained in the sensor data.

Instead, each cell $C_{x,y}$ in the grid holds a value expressing the number of sensors which currently confirm that square to be empty. Thus, a cell with a value of zero is in an unobservable region and so may possibly be occupied by a human. A cell with a value of three has been positively observed by three sensors to be unoccupied and thus is most likely actually empty.

To account for the possibility of conflicting data (*i.e.* one or more sensors erroneously observing an occupied space to be empty), the occupancy likelihood $p_{occ}(z_t|\mathbf{x}_t^{[m]})$, is modeled as falling off exponentially as a function of the value $C_{x_t^{[m]},y_t^{[m]}}$ stored in the coverage grid for that location, as shown in Eq. 2. This method of calculation serves to drive the particles into shadow regions in a smooth way, reducing sensitivity to sensor noise. K_{occ} is a tuning parameter which controls the noise tolerance of the function.

$$p_{occ}(z_t|\mathbf{x}_t^{[m]}) = \exp\left(-K_{occ} \cdot C_{x_t^{[m]},y_t^{[m]}}\right) \tag{2}$$

To speed up calculations, the entire grid is not populated at every time-step. Instead, a just-in-time grid update policy is adopted, whereby the value of a cell $C_{x,y}$ is calculated only the first time it is requested. That cell of the grid is then cached for further accesses, which is particularly efficient for dense clouds of particles covering a small number of grid cells.

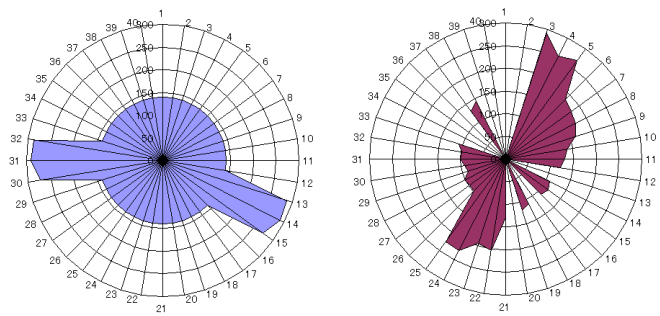

Fig. 3. Examples of populated radial accumulator arrays. Left: Populated with human shape model data (R_{ref}). Right: Populated with sensor data (R_m).

2) Radial Accumulator: The second component is the radial accumulator, an angular array which keeps track of the distance of detected data points from a proposed center point. By grouping the data points into discrete angular bins, the radial accumulator reduces the complexity of distance calculations while retaining the overall shape characteristics of the cluster of data points.

The metric chosen for evaluating the degree of fit between the model and the data points was a normalized root sum square calculation of radial distance. For each particle m at position (x_m, y_m), the surrounding cluster of points P_m are mapped into a radial accumulator array R_m with N equiangular divisions. For every set of points $P_m^n, (n = 1..N)$ falling in a particular bin, the distance of each point from position (x_m, y_m) is computed and the maximum value of that set is stored in R_m^n, until all points have been assigned. Let R'_m indicate the set of bins which have at least one point assigned to them, and let N' be the number of occupied bins. A second radial accumulator array R_{ref} is populated with expected distance values for each bin based on the human body model using parameters from the previous time step. Fig. 3 shows examples of the two radial accumulator arrays.

Next, the normalized root sum square error ϵ_j is calculated across all angular subdivisions, omitting bins which contain no data points.

$$\epsilon_j = \frac{1}{N'}\sqrt{\sum_{n \in R'_m} (R_m^n - R_{ref}^{n+j})^2} \tag{3}$$

To find the optimal value of θ, this procedure is repeated for j rotations of the model, and the value j_{opt}, which minimizes the error ϵ_j, is calculated. θ_{opt} is defined as the angle corresponding to bin index j_{opt}, and the θ parameter of the human shape model associated with this particle is then set to θ_{opt}.

Finally, the edge-fit probability p_{fit} is calculated as a zero-mean Gaussian function of the minimum error $\epsilon_{j_{opt}}$, with an arbitrary standard deviation of σ_{fit} treated as a tunable sensitivity parameter.

$$p_{fit}(\mathbf{x}_t^{[m]}) = \frac{1}{\sigma_{fit}\sqrt{2\pi}} \exp\left(-\frac{\epsilon_{j_{opt}}^2}{2\sigma_{fit}^2}\right) \tag{4}$$

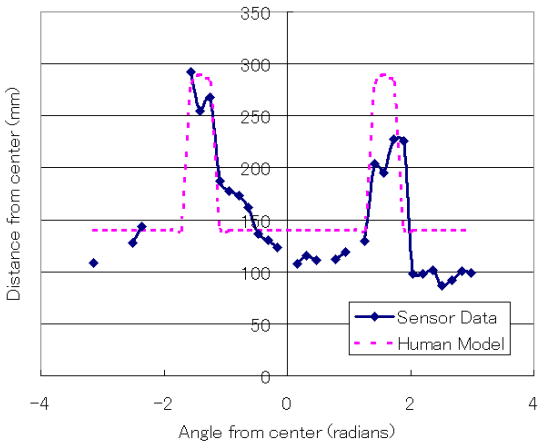

Fig. 4. Comparison of sensor data (R_t) and human shape model (R_{ref}) after optimizing rotation angle.

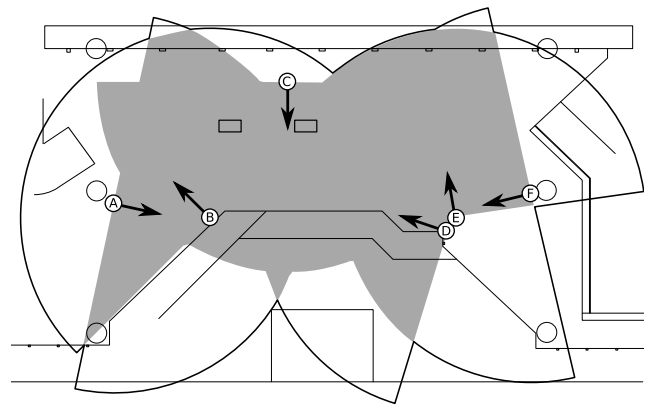

Fig. 5. Arrangement of laser scanners in experiment room. The solid black line shows the limit of the sensor range. The gray area indicates area covered by two or more sensors.

Fig. 4 shows the contents of the two radial accumulators after θ-optimization.

F. Model Adaptation

After the evaluation of each particle filter is complete, the parameters of the corresponding human shape model are adjusted to fit the current sensor observations. In our case, we have chosen to leave the parameters r_{arm} and r_{torso} fixed. We next adjust the remaining parameters as follows:

a) Optimizing φ: Since φ has been observed not to vary over a wide range, we limit it to between $\frac{\pi}{2} \pm \frac{\pi}{10}$ radians. Next, we discretize this angular range into 10 steps and construct a human shape model for each step, performing the θ-optimization procedure described in Section III-E.2 for each model. By minimizing ϵ_j across all models, we obtain optimal values for θ and φ.

Due to symmetry, arbitrary 180° flipping can be a problem. To remedy this, the range of possible θ values is constrained to be within ±90° of the previous θ value, and if necessary, the model is rotated 180° to be within 90° of the time-averaged motion direction vector.

b) Optimizing d_L and d_R: Next, the arm distances are adjusted. To produce smoother transitions between steps, the arm distances are not immediately set to optimal values at every time step. Rather, models are generated for single-step increases and decreases of arm distance, and each arm is set to the length that minimizes ϵ_j.

IV. EXPERIMENT

We verified the performance of our system using data taken in the lobby of the NICT building, near our laboratory at ATR, shown in Figure 5. Experimental subjects were instructed to walk through the lobby several times under a number of different conditions, *e.g.* individually, in groups, wandering aimlessly, walking purposefully, making U-turns, and stopping to ask for directions.

Raw data from a network of laser range finders monitoring this area was recorded for each trial, which we processed offline to determine human positions. In this experiment, no real-time tracking was performed.

A. Setup and Procedure

The area of interest in our experimental environment was a space within the lobby roughly 19 meters long and 8 meters wide. We used six SICK LMS-200 laser scanners, set to scan an angular area of 180° at a resolution of 0.5°, covering a radial distance of 8 meters with a nominal system error of ± 20 mm, providing readings of 361 data points every 26 ms. These were placed around the periphery of the experimental area, as illustrated in Figure 5. This arrangement was chosen such that every point within the area of interest would be covered by at least two sensors, which is important for reducing occlusions.

The sensors were mounted at a uniform height of 90cm, slightly above waist-level for most subjects. Tables, benches, and a small mobile robot were also placed within the walking area, but all of these were below 90cm and thus not visible to the laser scanners.

Twelve adults participated as subjects in this experiment, although at any given time only a subset of the group was walking within the sensor area. Six trials were conducted, and a total of 172 minutes of raw sensor data was collected.

B. Results

Two aspects of the results of this experiment will be considered here. The first is the accuracy of our method in tracking the subjects' motions, and the second is the ability to interpret this data in terms of actual body language and behavior.

1) Tracking Individuals: Quantifying the accuracy of this tracking technique is challenging due to the lack of more precise measurement techniques to establish a ground truth for evaluation. A side-by-side visual comparison of the raw data with the model-based estimate is perhaps the most effective indicator of the tracking accuracy.

Figure 6 shows raw data from five frames taken during the course of a single stride, and compares them with the

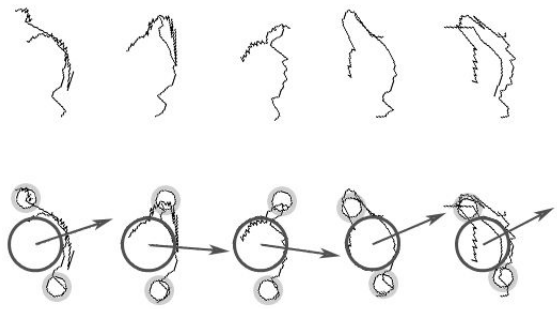

Fig. 6. Example of arm and torso movement during a single stride. Top: Five frames of raw data from laser scanners taken at 320ms intervals. Bottom: Corresponding human shape model positions for each frame.

model-based estimates for those time frames. Note that the swinging of the arm is clearly visible from the data, and that the model follows this movement closely.

Another indicator of the tracking accuracy of our technique is the resolution of movement that is visible over time. Figure 7 shows a sample path walked by one of the subjects during our experiment. The variations in θ due to the swinging of the arms and torso with each stride are quite clearly visible, with little noise present. The more subtle change in angle as the subject walks around a curving path is also quite clearly visible from the data.

These tracking results were then visually compared with video recorded during the experiment. The subjects' arm-swinging motions were observed to match with the data. The subject's torso rotations were not as exaggerated as the variations of θ in our model, which suggests the possibility that modeling the motion of the arms during walking may offer a better estimate of torso orientation.

Interestingly, our tracking results for this trial indicated an asymmetry of motion, with one arm moving much more than the other. Inspection of the video revealed that this was not a tracking error at all, but an idiosyncrasy of the subject's walking style, an observation which raises the question of whether observation of patterns in walking movement tracked by this model may reveal other information as well, such as aspects of personality or mood.

2) Observing Interactions: In addition to the model's tracking accuracy, it is important to consider what information can be observed regarding groups of people in social situations.

Figure 8 shows three scenes from our experiment. In the top scene, two subjects are seen walking together. The model correctly shows that they are walking side-by-side, facing slightly towards each other.

In the center scene, one of the two subjects is asking a third subject for directions. The model clearly shows the social situation, in which Subjects A and B are focusing their attention on subject C.

The bottom scene illustrates the tracking of a group of subjects. Again, the group dynamic is apparent, in that all of the subjects are listening to instructions from Subject A.

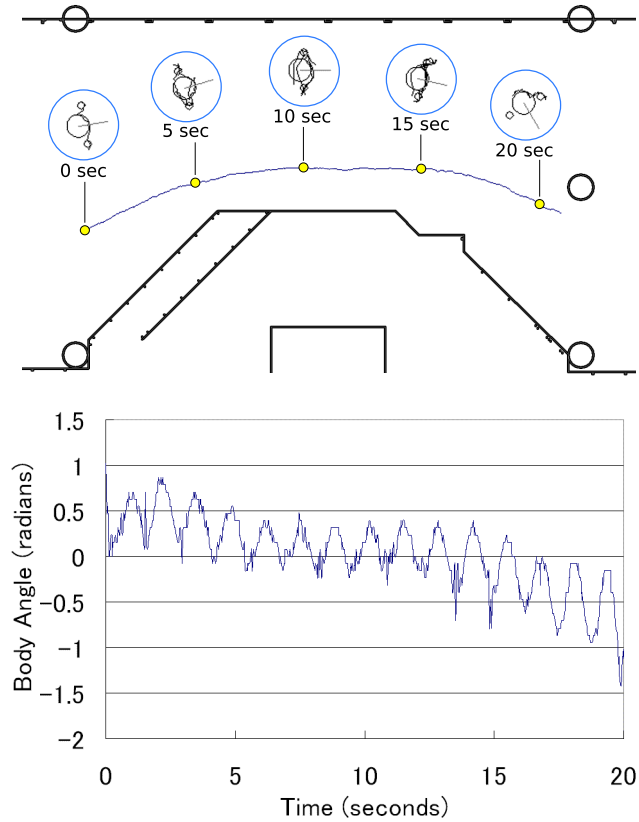

Fig. 7. Body angle tracking during 20 seconds of walking motion. Top: Body position captures along the walking path. Bottom: Observed body angle variations (in room-centric coordinates). The periodic oscillations are due to natural arm-swinging motion during each stride.

(Note that the model is unable to correctly determine the direction of Subject A because he is sitting and holding his arms in an unusual position.)

All three of these examples illustrate information that could not have been determined from purely location-based tracking data.

V. CONCLUSIONS AND FUTURE WORK

A. Conclusions

In this paper we have presented a novel human-tracking technique using adaptive shape modeling to determine the position and pose of a person based on a set of laser range scanner observations.

The proposed method uses particle filters and a parametric human shape model. The likelihood model used for the particle filter is composed of two components: one based on an evaluation of the possibility of occupancy based on sensor coverage, and one based on minimizing the error between observed body contours and the shape of the human model. The shape model is then updated after each time step to fit the observed data as closely as possible.

Our technique provides not only position, but also pose and body orientation information which can be useful for inferring meaning in a social context, such as determining

Fig. 8. Scenes from the experiment.

the direction of a person's attention or the social relationship between two people walking together.

We have demonstrated with real experimental data that the proposed technique is able to very closely track the movements of several subjects simultaneously, and we have shown that the information this method provides can help in evaluating social situations.

B. Future Work

The idea of extracting social contextual information from movement and positioning data is a powerful one. In this paper, the effectiveness of our proposed tracking method has been demonstrated in offline processing of experimental data. An important next step is the implementation of this algorithm within a real-time tracking system. To this end, a variety of techniques are available for improving the performance of particle filters in general ([12], [13], [14], [15], [16]), and several optimizations are being implemented for our algorithm itself.

Other refinements of this approach, such as using an elliptical torso model or varying the torso and arm radii may produce better tracking accuracy. Perhaps the motion model or parameter adaptation rules can be improved based on patterns in walking style. In addition, the parameter θ is currently a rather abstract quantity. An analysis of the relationship between arm swinging, walking speed, and body orientation would enable a more direct estimation of facing direction to be derived from the variation in θ over time.

This technology opens up many interesting possibilities for human-robot interaction research. An investigation into the subtleties of meaning encoded in leg stance, torso orientation, and head direction would provide useful insights into how to determine direction of attention. Perhaps differences in arm movement during walking or torso movements while standing can yield information about emotional state or intention. A closer examination of torso orientation during a variety of social interactions may also yield insights into understanding nuances of group dynamics, another field which is important to the development of interactive social robots.

VI. ACKNOWLEDGMENTS

This study was performed through Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

REFERENCES

[1] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte carlo localization for mobile robots," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 1999)*. ICRA, May 1999.

[2] M. Montemerlo, W. Whittaker, and S. Thrun, "Conditional particle filters for simultaneous mobile robot localization and people-tracking," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2002)*. Washington, DC: ICRA, 2002.

[3] S. Thrun, "Particle filters in robotics," in *Proceedings of the 17th Annual Conference on Uncertainty in AI (UAI)*, 2002.

[4] A. Brooks and S. Williams, "Tracking people with networks of heterogeneous sensors," in *Proc. Australian Conf. on Robotics and Automation (ACRA 2003)*, Brisbane QLD, Australia, Dec.1–3 2003.

[5] D. Schulz, W. Burgard, D. Fox, and A. B. Cremens, "People tracking with mobile robots using sample-based joint probabilistic data association filters," *International Journal of Robotics Research (IJRR)*, vol. 22, no. 2, pp. 99–116, 2003.

[6] A. Morate, "People detecting and tracking using laser and vision," Master's thesis, Royal Institute of Technology, Stockholm, Sweden, 2005.

[7] A. Almeida, J. Almeida, and R. Araujo, "Real-time tracking of moving objects using particle filters," in *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE 2005)*, Dubrovnik, Croatia, Jun. 2005.

[8] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.

[9] M. Gelgon, P. Bouthemy, and J.-P. L. Cadre, "Recovery of the trajectories of multiple moving objects in an image sequence with a pmht approach," *Image and Vision Computing*, vol. 23, pp. 19–31, 2005.

[10] S. Coradeschi and A. Saffiotti, "An introduction to the anchoring problem," *Robotics and Autonomous Systems*, vol. 43(2-3), 2003.

[11] A. Bruce and G. Gordon, "Better motion prediction for people-tracking," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2004)*, New Orleans, LA, USA, 2004.

[12] D. Fox, "Kld-sampling: Adaptive particle filters," in *Advances in Neural Information Processing Systems 14*. MIT Press, 2001.

[13] A. Doucet, N. de Freitas, K. Murphy, and S. Russell, "Rao-blackwellised particle filtering for dynamic bayesian networks," in *The 16th Annual Conference on Uncertainty in Artificial Intelligence, San Francisco, CA*. Morgan Kaufmann Publishers, 2000, pp. 176–183.

[14] A. Soto, "Self adaptive particle filter." in *IJCAI*, L. P. Kaelbling and A. Saffiotti, Eds. Professional Book Center, 2005, pp. 1398–1406.

[15] K. Choo and D. Fleet, "People tracking using hybrid monte carlo filtering," in *Proc. of the IEEE International Conference on Computer Vision (ICCV 2001)*, 2001, pp. 321–328.

[16] J. F. G. de Freitas, M. Niranjan, A. Gee, and A. Doucet, "Sequential monte carlo methods for optimisation of neural network models," Cambridge University Engineering Department, Cambridge, England, Technical Report TR-328, Nov. 1998.