Pose Estimation of Multiple People using Contour Features from Multiple Laser Range Finders

Takashi Matsumoto, Masamichi Shimosaka, Hiroshi Noguchi, Tomomasa Sato and Taketoshi Mori

Abstract— **Laser based tracking systems have been developed for mobile robotics and intelligent surveillance areas. Existing systems estimate only human positions. In this paper, we propose a method for human pose estimation represented by human head and waist position using only laser range finders. Two features of human cross-sectional contours are extracted from laser scanner data scanning on the height of waist. This method estimates human pose by using these features in the Bayesian filtering framework. Moreover, we develop a new particle filter framework with two transition models and two resampling steps. In this framework, position estimation and pose estimation are performed by many hypotheses. Our experimental results demonstrate the effectiveness of the method in pose estimation of multiple people by using only several laser scanners.**

I. INTRODUCTION

Recently, expectations towards robotic systems which enable daily life assistance are rising. Among many, researches on home environment with distributed sensors are active, and believed to offer practical applications. Such an environment is referred as smart environment, smart space and so on. Aware Home [1] and Sensing Room [2] are examples of such systems. Such smart environments observe the space using distributed sensors, extract useful information from the obtained data and provide various services to users. Therefore, a number of researches on human measurement using distributed sensors have been tried.

In those researches, most methods estimate only human position because it is based on the use of passive sensors. However, to provide appropriate service to the human according to the circumstances, not only positions but also human poses are essential information. Therefore, our goal is set to estimate positions and poses of multiple people in the environment by multiple laser range finders.

Among a variety of sensors, cameras and laser range finders are used in many works. Laser range finders are especially popular tools for human tracking and navigation applications due to their precision, effective sensing distance, and ease of use. Moreover, in comparison with cameras, laser range finders are resistant to environmental changes and simple to apply to various environments. Because camera images are influenced by illumination and the installation of cameras into home environment often gives feelings of resistance, we use only laser range finders.

Takashi Matsumoto, Masamichi Shimosaka, Hiroshi Noguchi, Tomomasa Sato and Taketoshi Mori are with Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan {matumoto, simosaka, noguchi, tmori}@ics.t.u-tokyo.ac.jp, tomomasasato@jcom.home.ne.jp

There are a number of researches on laser based human tracking. They tried various scanning techniques such as horizontal scanning at ankle-height [3], [4], [5], [6], horizontal scanning at waist-height [7], [8] and so on. In those researches, the methods based on leg tracking by horizontal scanning at ankle-height are the most popular. 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). For example, Zhao et al. [3] used Kalman filter and a walking model for leg tracking. Cui et al. [4] proposed a method based on the joint particle filter using multi-level observations. These methods succeeded in tracking multiple people in various environments. However, these methods were focused specifically on position tracking. On the other hand, Glas et al. [7] proposed a torso-level tracking method based on the particle filter, which estimates not only positions but also body orientation and arm position. However, it is intended for walking people and not estimates human poses. Therefore, we propose a method for human pose estimation represented by human head and waist position using cross-sectional contours on the height of waist. We focus on cross-sectional contours obtained by horizontal scanning at waist-height and develop a pose estimation system using only laser range finders.

In the past, human pose estimation from camera image sequences has been most proposed. Among those works, discriminative approaches [9], [10], [11] which model and predict state condition directly from observation was proposed, and it is reported to be effective for pose estimation. Therefore, we adopt a discriminative, example based approach, in which human pose candidates are defined beforehand and human poses are estimated by a comparison of likelihoods between input data and human pose candidates. Moreover, because it is not sufficient to estimate robustly from changing human contours, we incorporate the pose estimation by example based approach into the particle filter framework [12]. In our work, for applying two phases; estimation of the center of a contour and human pose estimation, we develop a new particle filter framework with two transition models and two resampling steps.

Our system is intended for multiple people, and we shows the extension to multiple people estimation. In this paper, we will show the effectiveness of the proposed method in pose estimation of multiple people.

This paper is organized as follows. We show the human pose estimation using contour features in section II. In section III, we define a overall framework for pose estimation
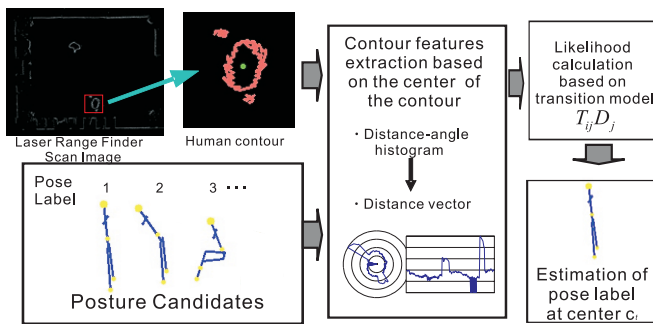
Fig. 1. Flow of human pose estimation using contour features

Fig. 2. Classification of human poses

of multiple people. We mention experimental results of the proposed method in section IV. Finally, conclusion is discussed in section V.

II. POSE ESTIMATION USING MULTIPLE LASER RANGE FINDERS

Human pose is estimated by using only laser range finders. First, we show the human pose estimation phase using contour features from multiple laser range finders in Fig.1.

We adopt an example based approach for human pose estimation. In the example based approach, human pose candidates are defined beforehand. Then, inference is achieved by a comparison of likelihoods between input data and human pose candidates. Moreover, we introduce Bayesian inference in the framework in order to search through the state space efficiently.

A. Human Measurement by Laser Range Finders

For the purpose of estimating not only human position but also human pose, we perceive 2D human cross-sectional contours on the height of waist. In usual approach, the feet data obtained by horizontal scanning at an elevation of about 20 cm above the ground are used. However, the feet data show only human position. In contrast, cross-sectional contours on the height of waist change in parallel with human poses. Therefore, we use cross-sectional contours on the height of waist obtained by horizontal scanning at an elevation of about 90 cm above the ground. 90 cm is the height of targets' waist, and we suppose that targets of our system are about 150-190 cm tall.

The coordinate system is represented with their X and Y axes aligned on the ground plane and the Z-axis representing the vertical direction from the ground plane. In our approach, a human pose is composed of the position of the head $p_h = (x_h, y_h, z_h)$, the position and the rotation around the Z-axis of the waist $p_{w,t} = (x_w, y_w, z_w, \theta)$. These parameters can represent human poses such as standing, sitting or bending in indoor environment. Estimating such poses enables a wide variety of support by robotic systems. z_h, z_w, θ are directly estimated by our approach. Other parameters are determined by these parameters z_h, z_w, θ in human skeleton model. Output values of z_h, z_w are reduced to discrete values $1, ..., 5$. As shown in Fig.2, these discrete values are sufficient to
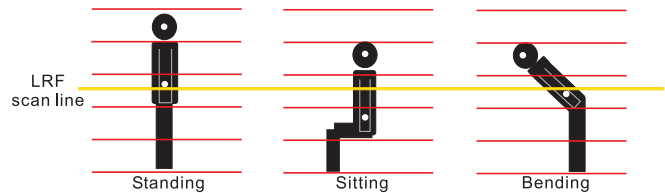
represent human poses such as standing, sitting or bending in indoor environment.

B. Human Contour Features

In example based approach, the choice of the distance for comparing two samples is crucial, as the distance is used for a comparison of likelihoods between input data and human pose candidates. Moreover, features require high discriminative power as well as rapid evaluation. Therefore, two features are extracted from human contours the center of which is defined beforehand.

First, as a feature that requires lower computational cost, we use distance-angle histogram. Detected data points of human contour are plotted polar coordinates using the center of the human contour as the origin. In the coordinates, the area is divided into b_r parts for the radial coordinate and b_ω parts for the angular coordinate and detected data points in each area is accumulated into each histogram bin. To discriminate the front-back orientation, b_ω is set to 2 and b_r is set to 4. The resulting histogram is composed of $b = b_r b_\omega$ bins, where b is 8. This rough histogram is used in the reduction of human pose candidates because the feature requires little computational cost. In the radial coordinates, a distance is limited within 500mm because the radius of human cross-sectional contours is below 500mm even when the target is bending.

Second, we use radial distance vector. The radial distance vector is an angular array which keeps track of the distance of detected data points from a proposed center point. The distance is obtained by the angle 2 degree. Thus the number of elements of the array is 180. Human contours obtained by laser range finders often suffer a loss. This feature can be extracted from such incomplete contours. Fig.3 shows examples of distance-angle histogram and radial distance vector of two poses.

C. Likelihood Evaluation Based on Multiple Features

In our approach, likelihood between input data and pose candidates is calculated by two features and likelihood evaluation employs a two step process.

In the first step, likelihood is calculated by the distance-angle histogram. Likelihood between a histogram $q(p)$ of the pose candidate p predicted by the transition model and a histogram $q(c_t)$ obtained by a proposed center c_t is based on the Bhattacharyya Coefficient. The likelihood is computed by

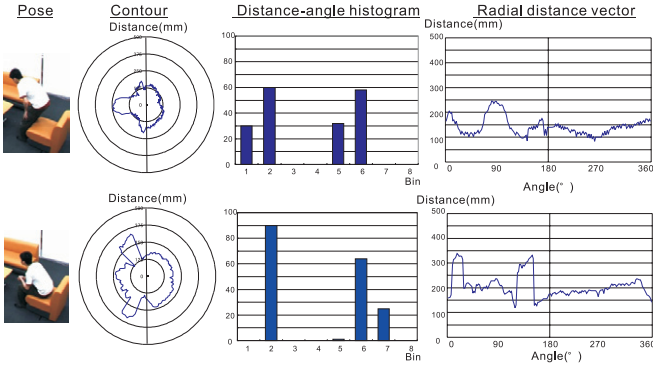$$D_{his}(q(c_t), q(p)) = \sum_{k=1}^{b} \sqrt{q(c_t; k)q(p; k)}, \qquad (1)$$

Fig. 3. Human contour features

where b is the number of histogram bins and b is set to 8. In the first step, pose candidates are eliminated if the likelihood score is below a threshold ζ. This function leads to the reduction of computational cost.

In the second step, likelihood is calculated by the radial distance vector. The metric chosen for evaluating the degree of fit between the pose candidate and the detected data points was a normalized root sum square calculation of radial distance. For each contour that has the center c_t, the surrounding cluster of points are mapped into a radial distance vector array R_{c_t} with N equiangular divisions. A radial distance vector array r_p^n is extracted from the pose candidate p. Likelihood D_{vec} of the normalized root sum square is calculated across all angular subdivisions as below, omitting bins which contain no data points.

$$D_{vec}(c_t, p) = \frac{1}{N}\sqrt{\sum_{n=1}^{N}(r_t^n)^2},\qquad (2)$$

$$r_t^n = \begin{cases} |r_{c_t}^n - r_p^n| & \text{if } |r_{c_t}^n - r_p^n| \le R_{max} , \\ R_{max} & \text{otherwise}, \end{cases}\qquad (3)$$

where each subdivision has the area of 2 degree and the maximum of N is 180. The maximum of distance between the corresponding points $r_{c_t}^n - r_p^n$ is set to R_{max}. This threshold function contributes to improving the robustness against noise such as data noise or non-defined contours of human pose that arms or hands changes the contour beyond the scope of the assumption.

D. Pose Estimation by Bayesian Inference

The example based approach requires high computational cost and often gives the estimation result that lacks continuity. To address these problems, we introduce Bayesian inference into the framework.

Given the observations up to time t, $z_{1:t}$, the aim is to estimate the posterior distribution $p(x_t|z_{1:t})$ of the state x_t. When estimating the posterior distribution $p(x_t|z_{1:t})$, making consideration of both the transition model $p(x_t|x_{t-1})$ and likelihood $p(z_t|x_t)$ enables Bayesian inference.

In our approach, the transition model $T_{i,j}$ from pose i to pose j is defined by the head height z_h, waist height z_w and
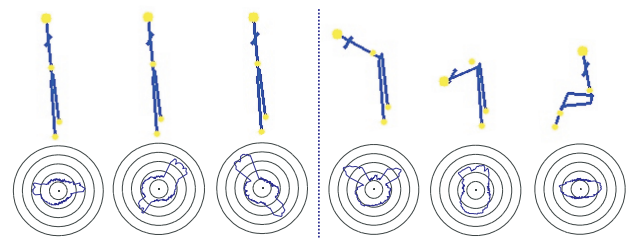

Fig. 4. Pose candidates

orientation θ of these poses as below.

$$T_{i,j} = \begin{cases} 1 & \text{if } (\ (\ |z_{h,j} - z_{h,i}| \le 1\ ) \\ & \quad \&\& (\ |z_{w,j} - z_{w,i}| \le 1\ ) \\ & \quad \&\& (\ |\theta_j - \theta_i| \le \eta_i\ )\ ), \\ 0 & \text{otherwise}, \end{cases}\qquad (4)$$

where $z_{h,i}$ is the head position of the pose i, and $z_{w,i}$ and θ_i are the waist position of the pose i. Values of z_h, z_w are discrete values $1, ..., 5$ as shown in Fig.2, and the transition to the next step is allowed. η_p is the maximum of the transition angle of the pose p in pose data set used in the stage of construction of pose candidates. In order to prevent getting caught into a local solution, $T_{i,j}$ is binarized as 0, 1.

E. Construction of Pose Candidates

Pose candidates are generated by hierarchical clustering.

Using a full body motion capture system and laser range finders, pose data and its contours are collected from subjects that each perform a variety of motions. For efficient motion data use, each pose is rotated and adjusted to the same orientation. First, these contours are divided by human head position $z_h = 1, ..., 5$ and human waist position $z_w = 1, ..., 5$. From each group, we extract several representative contours through an aggregative hierarchical clustering. In the clustering process, we use the radial distance vector as the distance between contours. Finally, pose candidates are constructed by rotating all pose candidates that have representative contours at intervals of a step degree Θ. Each pose candidate contains several representative contours and information of the head height, waist height and orientation. Fig.4 shows examples of poses and contours extracted as pose candidates.

III. POSE ESTIMATION BY MANY HYPOTHESES BASE ON THE PARTICLE FILTER FRAMEWORK

In our work, because two features are based on the center of a contour, estimation of the center is also an important task. If the estimation of the center is not accurate, the same contours have different features. Moreover, by the presence of incomplete contours that suffer a loss by occlusion, an approach where the center is estimated by averaging the laser points is not suited.

In addition, contours extracted from entirely different poses are often very similar. These contours also cause the failure of pose estimation.

Therefore, we develop a new particle filter framework with two transition model and two resampling steps for estimation
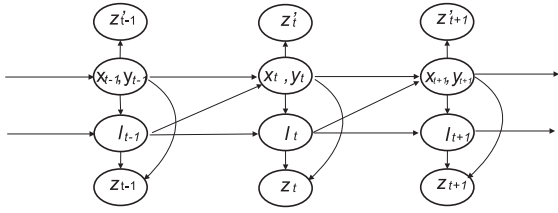
Fig. 5.   Graphical model

Fig. 7.   Evaluation model by background subtraction

of the center and human pose. In this framework, the center is estimated by many hypotheses and human pose is estimated by many hypotheses of the center.

A. Particle Filter

The particle filter is a Bayesian sequential importance sampling technique, which recursively approximates the posterior distribution using a finite set of weighted samples. The posterior distribution at time t is approximated by a set of discrete samples $\{ s_t^{(n)} \}$ $(n = 1...N)$ with importance weights $\{ \pi_t^{(n)} \}$ $(n = 1...N)$. Particle filter simulates this distribution by the following three-step recursion.

1) Selection: Select samples $\{ s_{t-1}'^{(n)} \}$ $(n = 1...N)$ in proportion to weight $\{ \pi_{t-1}^{(n)} \}$ $(n = 1...N)$ corresponding to sample $\{ s_{t-1}^{(n)} \}$ $(n = 1...N)$.

2) Prediction: Propagate samples $\{ s_{t-1}'^{(n)} \}$ $(n = 1...N)$ with state transition probability $p(x_t|x_{t-1} = s_{t-1}')$ and generate new samples $\{ s_t^{(n)} \}$ $(n = 1...N)$ at time t.

3) Update: Update weights $\pi_t^{(n)} = p(y_t|x_t = s_t^{(n)})$ corresponding to sample $s_t^{(n)}$ by evaluating a likelihood through observations. Normalize $\pi_t^{(n)}$ so that the sum of $\{ \pi_t^{(n)} \}$ $(n = 1...N)$ is equal to 1. As a result, estimated state at time t is equal to the expectation of the set of samples $\{ (s_t^{(n)}; \pi_t^{(n)}) \}$ $(n = 1...N)$.

B. Particle Filter with Multiple Transition Models and Multiple Resampling Steps

In our case, the sample state is set to $s_t = (S, l) = (x, y, l)$. The state is compose of a discrete index $l = (z_h, z_w, \theta)$ which labels the human pose and the continuous variable $S = (x, y)$ that denotes the center position of a contour in a 2D space.

The graphical model in Fig.5 describes the dependencies between our variables. The process density on the state sequence is modeled as a first order auto regressive process $p(x_t|x_{t-1})$. According to the independence assumption in the graphical model, the equation of the process density is:

$$p(x_t|x_{t-1}) = p(l_t|S_t, l_{t-1})p(S_t|S_{t-1}, l_{t-1}). \qquad (5)$$

Moreover, we introduce a two stage resampling process for estimation of the center and human pose.

Fig.6 shows the overall process and the concrete procedures of the proposed framework are given below.

1) Generate the new sample set $s_{t-1}'^{(i)}$ in proportion to weight $\pi_{t-1}^{(i)}$ at time $t - 1$.

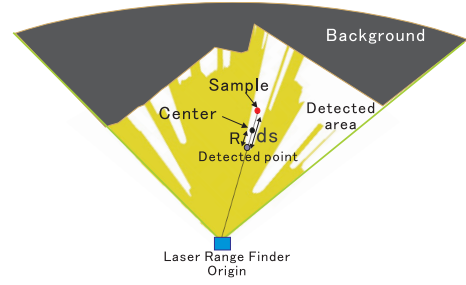2) Propagate the samples with the transition model $p(S_t|S_{t-1}, l_{t-1})$, and generate the sample set $s_t''^{(i)}$.

3) Update weights π_S using the sample state $S_t = (x, y)$.

4) Generate the new sample set $\hat{s}_t^{(i)}$ in proportion to weight $\pi_S^{(i)}$, and add the small diffusion.

5) Propagate the samples with the transition model $p(l_t|S_t, l_{t-1})$, and generate the sample set $s_t^{(i)}$.

6) Update weights π_t using the sample state $s_t = (S_t, l_t)$.

In this framework, the procedures are divided into two stages by two resampling steps. We describe details of these stages below.

C. The First Stage : Contour Center Estimation

Samples s_{t-1}' are propagated with the first transition model $p(S_t|S_{t-1}, l_{t-1})$. We assume uniform straight motion of a target position between two successive frames. Transition model $p(S_t|S_{t-1}, l_{t-1})$ is denoted as below.

$$S_t = S_{t-1} + \tau v_{t-1} + \omega_t, \qquad (6)$$
$$\omega_t = (\Gamma_{l_{t-1}}|v_{t-1}| + \gamma_{l_{t-1}})\nu, \qquad (7)$$

where τ is the time interval between frames, v_{t-1} is the previous velocity of the target, ω is a system noise added to s_{t-1}', and ν is a zero-mean Gaussian noise with a standard deviation 1. $\Gamma_{l_{t-1}}$ and $\gamma_{l_{t-1}}$ are constant determined beforehand by amount of the motion of human pose l, and v_{t-1} is estimated by the estimated target's position in two successive frames. We control the diffusion factor ω adaptively by the velocity v_{t-1} and the pose state l_{t-1} of the sample s_{t-1}'. Such control of the system noise ω contributes to improving the robustness against sudden abrupt motion and the accuracy of the center estimation.

The sample weights π_S of the first stage are evaluated by the likelihood of the center position from the observation z_t' and new samples are selected in proportion to sample weight. The weights π_S are evaluated by background subtraction model of each laser range sensor as below.

$$\pi_S = \sum_{i=1}^{M} \exp \frac{-|d_s - R_l|}{\lambda}, \qquad (8)$$

where d_s is distance between detected point by background subtraction model of sensor i and the sample s_t'', and R_l is the radius of the human contour determined by the pose l of the sample s_t''. Using this model, this function evaluates
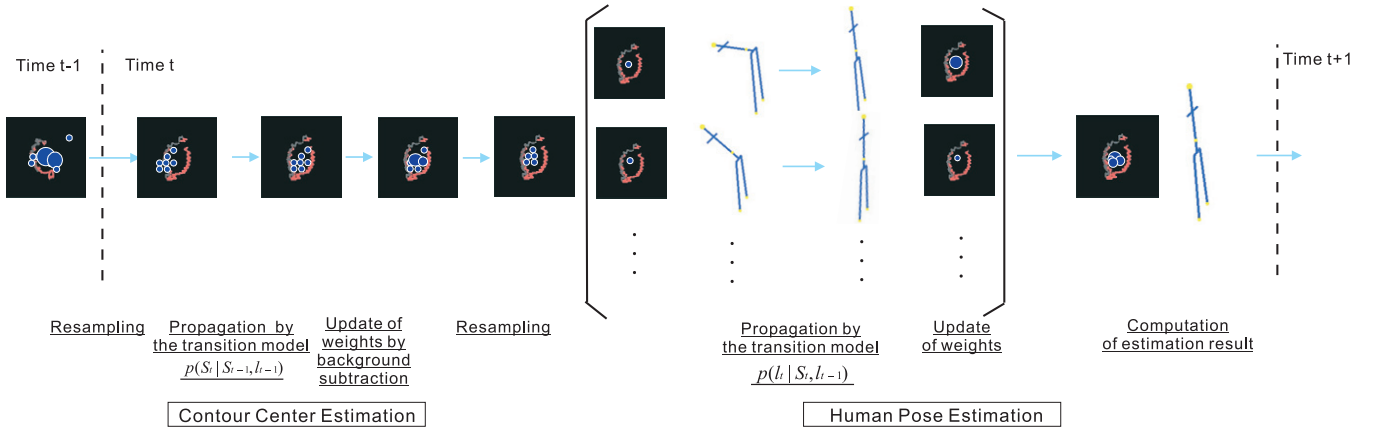
Fig. 6. Flow of the overall process

samples by distance between the sample and the center of the contour that we assumed. We show the background subtraction model in Fig.7

Finally, new samples are selected in proportion to sample weight π_S and are added the small diffusion. These procedures lead to a dense set of samples around the true center position.

D. The Second Stage : Human Pose Estimation

In the second stage, human poses are estimated from each sample state as described in section II. This function is implemented by the second transition model $p(l_t|S_t, l_{t-1})$. Transition model $p(l_t|S_t, l_{t-1})$ is denoted as below.

$$T_{l_{t-1}, l_t} D_{his}(q(S_t), q(l_t)) > \zeta, \tag{9}$$

$$l_t = \underset{l}{\arg\max} \; D_{vec}(S_t, l), \tag{10}$$

where $D_{his}(q(S_t), q(l_t))$, $D_{vec}(S_t, l)$ are likelihood computed in Eq.1, 2. Pose label l_{t-1} transits to pose label l_t that is estimated from the position $S = (x, y)$ of the sample $\hat{s}_t$ by the method described in section II. In this framework, each sample has a pose estimation framework by Bayesian inference and each sample transits independently.

The weights π are evaluated by the likelihood $D_{vec}(S_t, l_t)$ based on the sample state $s_t = (S_t, l_t)$ as below.

$$\pi_t = \exp(-\kappa D_{vec}(S_t, l_t)) \tag{11}$$

Because estimated state at time t is equal to the expectation of the set of samples $\{ \; (s_t^{(n)}; \pi_t^{(n)}) \}$, estimation results by each sample are integrated.

E. Extension to Multiple People Pose Estimation

Traditional particle filters perform poorly at consistently maintaining the multi-modality in the target distribution that often results from multiple targets. Vermaak et al. [13] introduced a mixture particle filter (MPF), where each component is modeled with an individual particle filter that forms part of the mixture. The MPF enables tracking multiple targets simultaneously. In our approach, for the purpose of multiple people estimation, our method is extended to multiple single-target particle filter framework as the MPF. Each filter

tracks each person independently. Human contours obtained by laser range sensors scarcely join each other and not require a segmentation of human existing area. Therefore, the extension can be simplified.

IV. EXPERIMENTAL RESULTS

A. Setup and Procedure

The area of interest in our experimental environment was a space within the room roughly 4 meters long and 4.5meters wide. We used four SICK LMS-200 laser scanners, set to scan an angular area of $100°$ at a resolution of 0.25, covering a radial distance of 8 meters with a nominal system error of ±15 mm, providing readings of 401 data points every 53 ms.

The sensors were mounted at a uniform height of about 90cm, just waist-level for most subjects. Tables and sofas were also placed within the area, but all of these were below 90cm and thus not visible to the laser scanners.

In the construction of pose candidates, we used about 1000 frames of data obtained by a full body motion capture system NatruralPoint OptiTrack. We extract 22 representative contours through clustering process. Θ is set to 10 and $22 \times 36 = 792$ pose candidates are constructed.

In this experiment, we set $R_{max} = 100(mm), \lambda = 220, \kappa = 5.0, \zeta = 0.75(N/180)$ respectively. Moreover, we processed 150 samples per one target.

The overall process is executed by one PC (Intel Core2 Duo E6600 2.40GHz, Memory 2GByte). The tracking runs online at a frame rate of 16.6 Hz when tracking two people, 12.5 Hz when tracking three people. In the below estimation experiments, we used the data obtained at 10 fps.

B. Pose Estimation of a single person

In order to investigate quantitatively the estimation accuracy by the proposed method, we performed several experiments. In the experiments, we assumed a variety of movements of a single person: walking, sitting, bending, etc. The true position is extracted from the motion data collected by a full body motion capture system NatruralPoint OptiTrack at the same time. Fig.8 shows an example of estimation results.
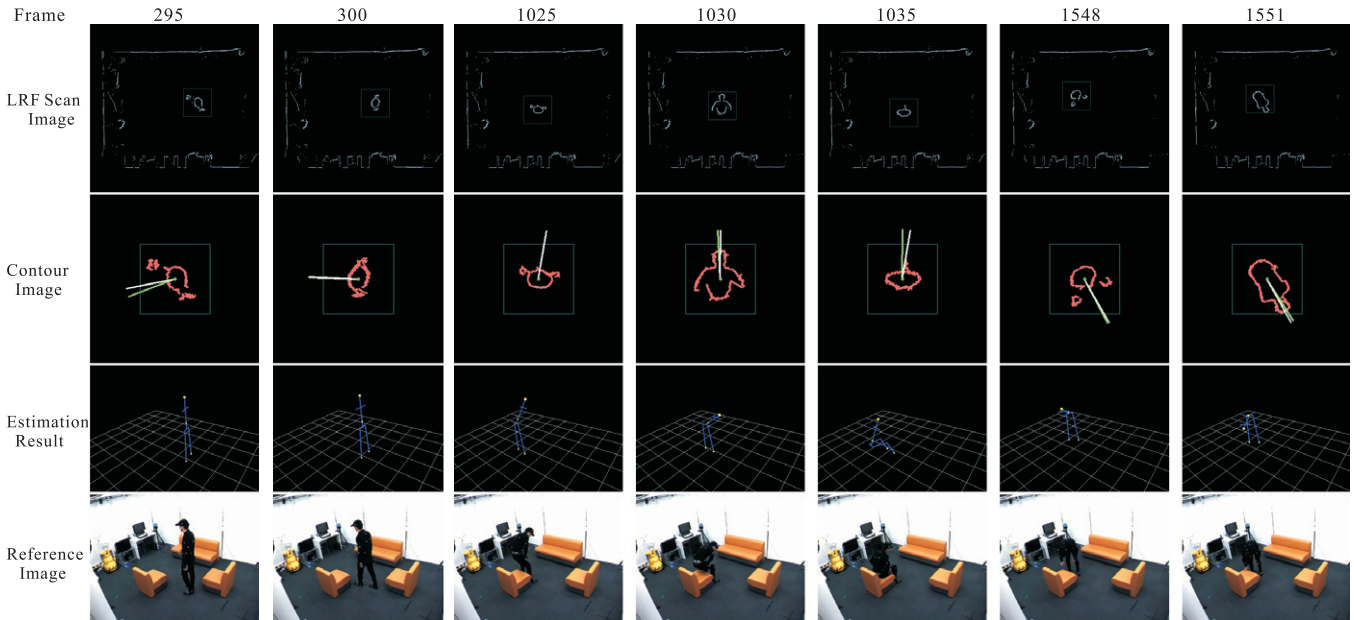
2194

Fig. 8. Experimental results on single human pose estimation

In contour image of Fig.8, white lines denotes estimated waist orientation and green lines denotes the true waist orientation. Reference image is captured for the reference at the same time by a camera. We can see from Fig.8 that the pose is well estimated through a series of motions. Contours extracted from different poses in frame 300 and 1035 are so similar that it may be difficult to distinguish these poses. However, we can see that it is possible to estimate these poses by our method. This is attributed to making consideration of time-series information. Then, due to symmetry, arbitrary 180° flipping can be a problem in the contours as like frame 295 or 300. Against such contours, our method also estimates correctly. We can also confirm the robustness of our approach against the various motions of arms or hands from frame 295 and 300.

Moreover, we computed estimation accuracy. If head position z_h and waist position z_w are correct and error of rotation angle θ is below 30° or 45°, we regard the result as a correct estimation and computed accuracy. We used 5616 frames that contain 3608 walking frames and 1247 sitting frames because the motion of the daily life in home environment is assumed. As a result, estimation accuracy(%) is 82.5 if error of rotation angle θ is below 30°, and 85.6 if error of rotation angle θ is below 45°. Moreover, estimation accuracy is 86.7 when we evaluated only the head and waist position.

We also computed recall rate, precision rate, F-measure at each poses that not contain rotation. Recall rate, precision rate, F-measure are defined as below.

$$\text{recall rate} = \frac{\text{correctly estimated poses}}{\text{total poses}} \quad (12)$$

$$\text{precision rate} = \frac{\text{correctly estimated poses}}{\text{total poses estimated by method}} \quad (13)$$

$$\text{F-measure} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (14)$$

TABLE I
ESTIMATION RESULTS AGAINST STANDING AND SITTING

tolerance	Standing			Sitting		
	Recall	Precision	F-measure	Recall	Precision	F-measure
30°	84.3	88.1	86.2	92.1	94.6	93.3
45°	87.8	91.7	89.7	92.1	94.6	93.3
only position	89.6	93.6	91.5	92.1	94.6	93.3

Table I shows recall rate, precision rate, F-measure at standing pose and sitting pose.

We confirm from estimation accuracy that our method estimates human poses sufficiently. As for the orientation, the accurate estimation can be also confirmed. If we can get the orientation within an error 30°, it is possible to provide appropriate supports by robotic systems.

All F-measure scores at sitting poses are above 92.1 and the high performance against sitting poses is determined. In contrast, the F-measure score at standing poses within 30° angle errors is 86.2 and less than that at sitting poses. This is due to the rapid and various motions of arms that change the contour. However, the rough orientation is well estimated.

C. Pose Estimation of Multiple People

In order to demonstrate the effectiveness and robustness of the method against multiple people, we performed several experiments. In the experiments, we assumed up to four people. Fig.9 shows some of the results. In top image, extracted contours are superimposed upon scan data.

The effectiveness of our method in the movements of multiple people can be confirmed from Fig.9. In frame 710, there is an example of incomplete contours that suffer a loss by occlusion. Even in this case, our method can maintain correct estimation. This is attributed to pose estimation by many hypotheses. We can also see that the contours from

Frame	690	710	890	1040	1080

LRF Contour
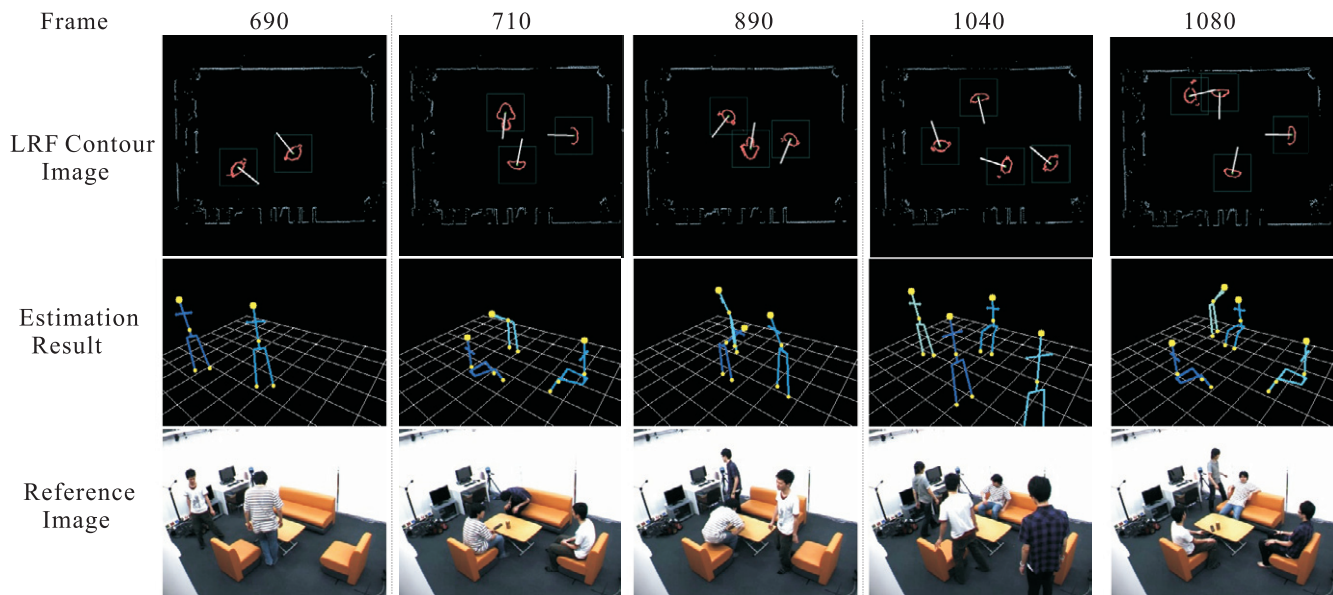Image

Estimation
Result

Reference
Image

Fig. 9.   Experimental results on pose estimation of multiple people

laser range finders can be obtained stably and hardly suffer an extreme loss. It is a superior property of laser range finder. Consequently, human poses of multiple people can be estimated with the almost same accuracy as that of a single person. Then we used the same pose candidates set for all people. We can confirm that differences among individuals hardly affect the estimation accuracy.

V. CONCLUSION

In this paper, we proposed a method for pose estimation of multiple people by using only laser range finders. In order to estimate not only position but also pose, we used human cross sectional contours on the height of waist.

The proposed method uses the example based approach and a new particle filter framework with two transition models and two resampling steps. Position estimation and pose estimation are performed by many hypotheses.

In our experiments, estimation accuracy(%) is 82.5 if error of rotation angle θ is below $30°$, and 85.6 if error of rotation angle θ is below $45°$. Thus we showed the effectiveness of the proposed method in estimation of human poses such as standing, sitting or bending in indoor environment. Successful pose estimation for multiple people was also demonstrated in our experiments. It was shown that pose estimation from human cross sectional contours on the height of waist is effective in pose estimation of multiple people.

Future tasks are experiments in the other wide environment such as airports, train stations, and shopping malls. Because the laser range finder covers a radial distance of 20 meters with a nominal system error of ±4 cm, it is expected that the propose method is effective in those environments.

REFERENCES

[1] Cory D. Kidd, Robert Orr, Gregory D. Abowd, Christopher G. Atkeson, Irfan A. Essa, Blair MacIntyre, Elizabeth Mynatt, Thad E. Starner, and Wendy Newstetter. The Aware Home: A Living Laboratory for Ubiquitous Computing Research. In *Second International Workshop on Cooperative Buildings - CoBuild'99*, pages 191–198, 1999.
[2] Taketoshi Mori, Hiroshi Noguchi, Aritoki Takada, and Tomomasa Sato. Sensing Room: Distributed Sensor Environment for Measurement of Human Daily Behavior. In *International Conference on Networked Sensing Systems*, pages 40–43, 2004.
[3] Huijing Zhao and Ryousuke Shibasaki. A novel system for tracking pedestrians using multiple single-row laser-range scanners. *IEEE Transactions on Systems, Man, and Cybernetics, Part A:Systems and Humans*, 35(2):283–291, 3 2005.
[4] Jinshi Cui, Hongbin Zha, Huijing Zhao, and Ryosuke Shibasaki. Laser-based interacting people tracking using multi-level observations. In *IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS 2006)*, pages 1799–1804, 10 2006.
[5] Xiaowei Shao, Huijing Zhao, Katsuyuki Nakamura, Kyoichiro Katabira, Ryosuke Shibasaki, and Yuri Nakagawa. Detection and tracking of multiple pedestrians by using laser range scanners. In *International Conference on Intelligent Robots and Systems(IROS2007)*, pages 2174–2179, 2007.
[6] Xuan Song, Jinshi Cui, Xulei Wang, Huijing Zhao, and Hongbin Zha. Tracking interacting targets with laser scanner via on-line supervised learning. In *IEEE International Conference on Robotics and Automation(ICRA2008)*, pages 2271–2276, 2008.
[7] Dylan F. Glas, Takahiro Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. Laser tracking of human body motion using adaptive shape modeling. In *IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS 2007)*, pages 603–608, 11 2007.
[8] Ajo Fod, Andrew Howard, and Maja J Mataric. A laser-based people tracker. In *IEEE International Conference on Robotics and Automation*, pages 3024–3029, 2002.
[9] Yuichi Sagawa, Masamichi Shimosaka, Taketoshi Mori, and Tomomasa Sato. Fast online human pose estimation via 3d voxel data. In *International Conference on Intelligent Robots and Systems*, pages 1034–1040, 2007.
[10] R. Okada and B. Stenger. A single camera motion capture system for human-computer interaction. *IEICE Transactions*, E91-D(7):1855–1862, July 2008.
[11] Gregory Shakhnarovich, Paul Viola, and Trevor Darrell. Fast Pose Estimation with Parameter Sensitive Hashing. In *Proc. of ICCV*, volume 2, pages 750–757, 2003.
[12] Michael Isard and Andrew Blake. Condensation-conditional density propagation for visual tracking. *Internanional Journal on Computer Vision*, 29(1):5–28, 1998.
[13] Jaco Vermaak, Arnaud Doucet, and Patrick Perez. Maintaining multi-modality through mixture tracking. In *International Conference on Computer Vision*, pages 1110–1116, 2003.