Proceedings of the IEEE ICRA 2009
Workshop on People Detection and Tracking
Kobe, Japan, May 2009

Multiple People Detection from a Mobile Robot using Double Layered Laser Range Finders

Alexander Carballo, Akihisa Ohya and Shin'ichi Yuta

***Abstract*— This work presents our method for people detection on the surroundings of a mobile robot by using two layers of multiple LRFs, allowing to simultaneously detect two set of different features for every person: chest and legs areas. A person model is created according to the association of these features and a volume representation allows to estimate the current person position. We present experimental results of multiple people detection in an indoor environment. The main problem of our research the development of a mobile robot acting as member of a group of people, simple but accurate people detection and tracking is an important requirement.**

I. INTRODUCTION

Companion robots are becoming more part of daily life and are designed to directly interact with people. One necessary subsystem for such robots is detection, recognition and tracking of people as well as obstacles in the environment.

Laser Range Finders (LRF), besides being used for obstacle detection are also an important part of people tracking systems. The Tour-Guide robots Rhino and Minerva by Burgard *et al*[1] and Thrun *et al*[2] featured LRFs for people detection and collision avoidance. LRF present important advantages over other sensing devices like high accuracy, wide view angles, high scanning rates, etc., and are becoming more accessible and safer (meaning class 1 lasers) for usage in human environments.

Most approaches based on LRFs ([3], [4], [5], [6], [7], [8]) place the sensors in the same height (single row or scan plane) to detect and track some feature of the human body. Due to laser safety regulations, applications using non class-1 lasers are mostly limited to a low position, mostly about knee height or below. Thus legs are widely used as features for human detection and tracking.

In Fod *et al* [3] a row of several LRFs on different positions in a room were used for tracking moving objects, future positions are estimated according to a motion model. Montemerlo *et al* [4] also uses LRF from a mobile robot for people tracking and simultaneously robot localization by using conditional particle filters. Xavier *et al* [5] focused on people detection using a fast method for line/arc detection but from a fixed position. Zhao *et al* [6] proposed a walking model to improve position prediction by including information about leg position, velocity and state. The later model was then used by Lee *et al* [7] and by Zhao *et al* [8] but this time from a mobile robot.

Intelligent Robot Laboratory, Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba City Ibaraki Pref., 305-8573, Japan, +81-29-853-6168. {acs,ohya,yuta}*@roboken.esys.tsukuba.ac.jp*

A common problem is how to correctly identify people features from laser measurements. Arras *et al* [9] using range data and Zivkovic *et al* [10] using range data and images, employ a learning method, particularly *AdaBoosting*, to determine which properties and in what amounts to consider to improve detection. However, detection of multiple people in cluttered environments is difficult especially considering occlusion cases of people walking side by side.

Most tracking applications can deal with temporal occlusions due to obstacles, such as the temporal disappearance of the legs behind a dust bin. Multiple target tracking in cluttered environments including crossings tracks is part of most current works [11], [12], [13]. Mucientes *et al* [11] extends the problem of single person tracking by considering clusters of tracks (people) using Multiple Hypothesis Tracking (MHT). Arras *et al* [12] also uses MHT for tracking without a leg swinging-motion model but introducing an occlusion state, low level tracks (legs) are associated to a high level track (people). Kondaxakis *et al*[13] present also a multi-target approach using JPDA with a grid map where occupancy counters of each cell decrease with time to identify background objects.

One limitation still present in those systems is occlusion of the tracked body feature for an extended time, for example if the person stopped behind the dust bin. MHT based systems will delete of the occluded track if it is missing for more than some maximum time. The usage of additional features can overcome this problem, provided that they are separated over some distance (height) where occlusion stops. Instead of a single layer system one can consider a multi-layered arrangement of class-1 LRFs on a mobile platform. Multiple features have the additional benefit of complementarity for detection and tracking: a person can be described by the union of a set of small swinging segments at low height (legs), a bigger segment at medium height (waist) and a larger segment at a high position (chest). This idea was proposed in our previous work [14]. A multi-layered system to extract multiple features is of course possible as long as the person height is over some minimum value.

A multi-layered system has also being considered previously [15], [16]. Gidel *et al*[15] used a 4-layer laser sensor for pedestrian detection, scanning planes are not parallel so that slight tilting of the vehicle do not affect detection. However, the vertical distance between features on the target pedestrian depend on the distance from the sensor. Hashimoto *et al*[16] use 3 LRFs around a wheelchair for 360^o scanning at 3 different heights, each sensor with its own processing computer performing detection and tracking.

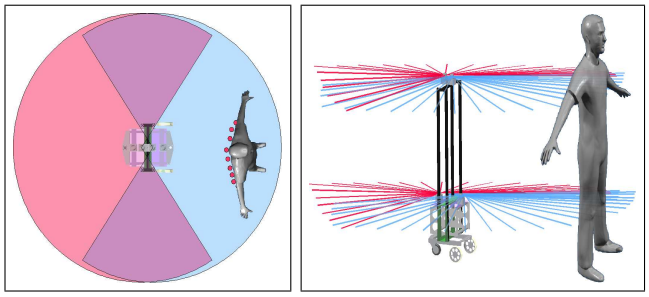
For every sensor, scan data is mapped into a occupancy grid map, then target tracking and tracks association is performed. Tracking in overlapping areas is done by cooperation of respective computers and covariance intersection.

Our approach is then similar to Hashimoto's[16]: we have sensors are arranged in two parallel planes for 360^o scanning, separated at different heights from the ground depending on the features to detect. However, we perform all computing in a single computer, sensors in the same layer are fused to combine their individual readings and then layers are also fused for people detection.

The rest of the paper is organized as follows. In section II present an overview of our current system. Section III presents our approach for fusion of multiple sensor layers, including feature extraction, people detection and position estimation. Section IV presents experimental results for the different fusion steps and for people detection. Finally, conclusions and future work are left for section V.

II. System Overview

Fig. 1 represents our layered approach, every layer has two sensors facing opposite directions for 360^o scanning (Fig. 1(a)), and two layers are used to extract features from upper and lower parts of a person's body (Fig. 1(b)).



(a) (b)

Fig. 1: Scanning from a double layered approach: (a) opposite facing sensors (top view) and (b) two layers of sensors (lateral view).

The processing pipeline of our system is best understood referring to Fig. 2. Our people detection approach (before tracking) involves four steps: fusion of sensors, segmentation, feature extraction and layer fusion. The outputs for some of the steps are depicted as inlets in the figure: Fig. 2(a) is the result of fusion of sensors (the top layer represented in red and the lower in green), Fig. 2(b) corresponds to geometrical feature extraction (features of people is shown), and in Fig. 2(c) the detected people around the robot.

Our method involves two fusion steps: fusion of sensors in a single layer and then fusion of layers. In the first step, sensors facing opposite directions in the same layer are fused to produce a 360^o representation of robot's surroundings. There is overlapping of scan data from both sensors (darker areas in Fig. 1(a)) so this fusion step must deal with data duplication. Then, in the multiple layer fusion step, raw data from every layer is processed to extract features

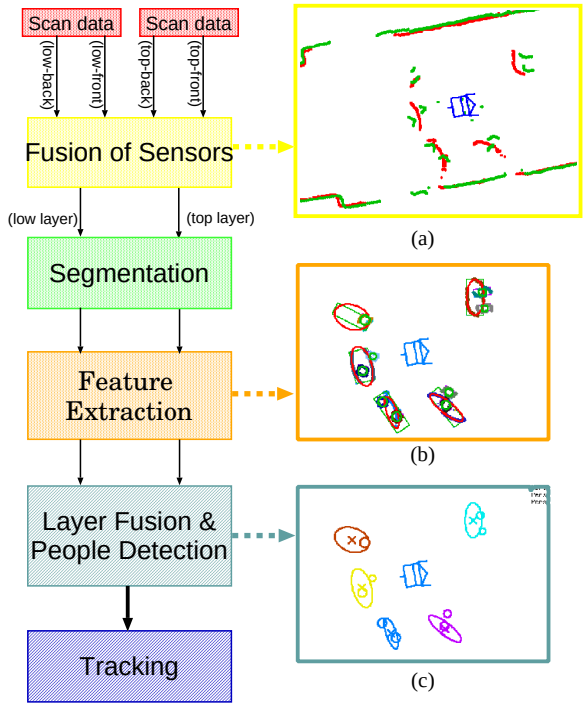


Fig. 2: System overview.

corresponding to people, then a people model is computed and from it allowing people detection and person position and direction estimation.

After fusion of sensors in every layer, geometrical features are extracted: large elliptical shapes corresponding to chest areas and smaller circular shapes for legs. Fusion of extracted features allows creating a cylindrical volume and from it the estimated person position is computed. A simple yet logical assumption here is that an elliptical shape corresponding to a chest is always associated to one or two circular shapes corresponding to legs (if no occlusions due to clutter are considered), and that the large elliptical shape (chest) is *always over* the set of small circles (legs). Fig. 3 illustrates this concept, here we present a sequence of continuous scan images from a person walking (as seen from above), both upper layer (large arc-like shape, chest) and lower layer (small arc-like shapes, legs) are visible.

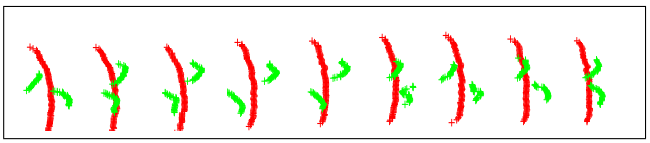
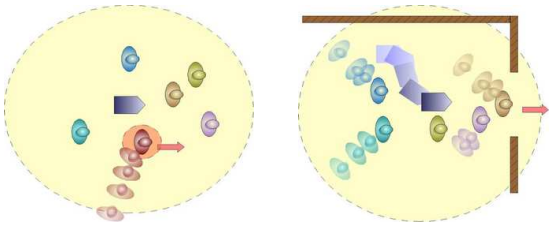


Fig. 3: A sequence of walking steps using actual scan data using sensors from both upper layer (darker points on large curve) and lower layer (smaller curves).

Our main research goal aims to develop a companion robot with the objective to study the relationship of an autonomous mobile robot and a group of multiple people in a complex environment like public areas, where the robot is to move and behave as another member of the group, while

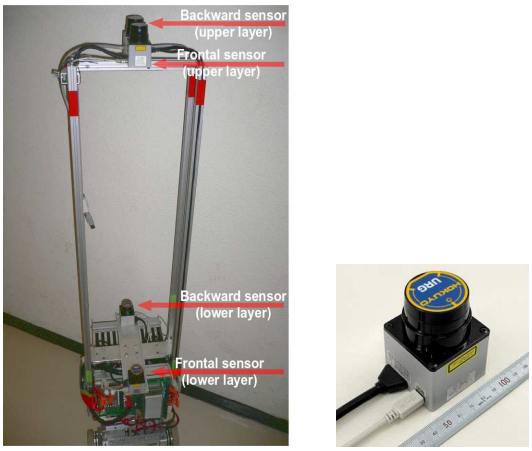
achieving navigation with obstacle avoidance. Some of the basic functions of such companion robot are depicted in Fig. 4, while the robot acts as another group member it has to detect, recognize and track the fellow human members (Fig. 4(a)) and also move in the environment like the rest of the members do (Fig. 4(b)).



(a) (b)

Fig. 4: Companion Robot with a group of people: group members recognition (a) and obstacle avoidance (b).

The robot used for our research is depicted in Fig. 5. The robot (Fig. 5(a)) is based on *Yamabico* robotic platform [17]. Two layers of LRF sensors are used, the lower layer is about 40*cm* from the ground while the upper layer is about 120*cm*. Every layer consists of 2 LRF sensors, one facing forwards and another facing backwards for a 360^o coverage (Fig. 1 and 5). The sensors used in our system are the *URG-04LX* laser range scanners (Fig. 5(b), [18] provides a good description of the sensor's capabilities).



(a) (b)

Fig. 5: Our robot system for multiple people detection and tracking (a), four *URG-04LX* are used (b).

III. Fusion of double layered LRF sensors

Sensors in the same layer are facing opposite directions, individual scan data are combined into a 360^o representation. The next step is fusion of both sensor layers, here data will be divided into clusters with a segmentation function and then clusters will be classified according to their geometrical properties. Finally only those segments that match people features will be selected and joined into a 3D model from where people position is obtained.

A. Segmentation

Data clustering can be considered as the problem of break-point detection and finding breaking points in scan data can be considered as the problem of finding a threshold function $\mathscr{T}$ to measure separation of adjacent points. Every pair of neighboring points p_j and p_k are separated by an angle α which is proportional to the sensor's angular resolution (true for points of two adjacent scan steps) and by a distance $\mathscr{D}(p_j, p_k)$. Points are circularly ordered according to the scanning step of the sensor.

A cluster $\mathscr{C}_i$, where $\mathscr{C}_i = \{p_i, p_{i+1}, p_{i+2}, \cdots, p_m\}$, is defined according to a cluster membership function $\mathscr{M}$

$$\mathscr{M}(p_j, p_k) = (\theta_k - \theta_j) \leq \alpha \wedge \mathscr{D}(p_j, p_k) \leq \mathscr{T}(p_j, p_k) \quad (1)$$

such that for every pair $\langle p_j, p_k \rangle$ of adjacent points, the Euclidean distance $\mathscr{D}(p_j, p_k)$ between them is less than a given threshold function $\mathscr{T}(p_j, p_k)$ for p_j, p_k. A new point p_n is compared to the last known member p_m of a given cluster $\mathscr{C}_i$ as $\mathscr{M}(p_m, p_n)$.

Now, the threshold function $\mathscr{T}$ is defined for a pair of points, as in the work of Dietmayer [19], as:

$$\mathscr{T}(p_i, p_j) = C_0 + C_1 min(r_i, r_j) \quad (2)$$

with $C_1 = \sqrt{2(1 - cos(\alpha)}$. Dietmayer's work includes the constant C_0 to adjust the function to noise and overlapping. In our case C_0 is reemplaced by the radius $\mathscr{R}$ of the accuracy area for p_i as base point plus a fixed threshold value (10*cm* in our case). $\mathscr{R}$ is defined according to the *URG-04LX* sensor specifications [18], [20] as:

$$\mathscr{R}(p_i) = \begin{cases} 10 & \text{if } 20mm \leq r_i \leq 1000mm \\ 0.01 \times r_i & \text{otherwise} \end{cases} \quad (3)$$

The proposed threshold function $\mathscr{T}$ uses this accuracy information $\mathscr{R}$ when checking for break points, if two neighboring points have a large range value, it will be most probable that they form part of the same cluster for their bigger accuracy areas.

There is also a cluster filtering step that will drop segments very small to be considered of significance.

B. Feature Extraction

The idea of *feature extraction* is to match the sensor readings with one or more geometrical models representing expected behaviour of the data. For example if a LRF sensor data scanning a wall, then the *expected behaviour* of *a wall* scan data is *a straight line*. Also if the same sensor is to scan *a person* then the expected behaviour is a set of points forming *an arc*. So in order to identify walls a first requirement is to correctly associate the scan data with some straight line model, for people the same: associate a set of scan points to an arc shape (a circle or an ellipse).

Before applying any fitting method, it is important to have some information about the shape of the cluster that allows selecting the method. The information about clusters is extracted as a set of indicators like number of points, standard deviation, distances from previous and to next clusters, cluster curvature, etc.

One of the indicators is the cluster's *linearity*; our approach here is to classify the clusters into *long-and-thin* and those rather *short-and-thick*. The rationale behind this is that, straight line segments tend to be long and thin, round obstacles, irregular objects, etc., do not have this appearance.

Linearity is achieved by computing the covariance matrix Σ for the cluster $\mathscr{C}_i$ and then its eigenvalues λ_{max} and λ_{min} that define the scale and its eigenvectors $\mathtt{v}_1$ and $\mathtt{v}_2$ orientation (major and minor axes) of the dispersion of $\mathscr{C}$. The ratio $\ell = \lambda_{max}/\lambda_{min}$ defines the degree of longness/thinness of the cluster. We set threshold values for ratio $\mathscr{L}$ and for λ_{max}.

The *ellipticality* factor ε is computed as the standard deviation σ of the residuals of a ellipse fitting processes using the Fitzgibbon method [21]. The distance between a cluster point and an ellipse is computed using *Ramanujan's* approximation.

Only clusters with good ellipticality value are selected and segments passing the linearity criteria (that is lines) can be easily rejected since they do not belong to people.

We assign a weight value $\mathtt{w}$ to every indicator i and compute an scoring function $\mathbb{S}$ for every segment j in in layer Ψ, where $\Psi \in \{top, low\}$, as:

$$\mathbb{S}^j = \sum_i^n \mathtt{w}_i^\Psi \mathscr{H}_i^\Psi(I_i^j) \tag{4}$$

where $\mathscr{H}_i^\Psi : \mathbb{R} \to \{-1, 1\}$ is a binary classifier function for the i-th indicator which compares whether the given indicator is under some threshold value. Table I presents an example of indicators and their classifiers, the actual list of indicators is similar to that presented by Arras *et al* in [9]. Weight values $\mathtt{w}_i$ and thresholds for every indicator i were defined after experimental validation.

TABLE I: Example of indicators and their classifiers

Indicator	Classifier	Meaning
width w	$w \leq W_{max}^\Psi$	a leg or a chest has a width no bigger than the threshold
linearity ℓ	$\ell \leq \ell_{max}^\Psi$	leg and chest features are not linear
curvature k	$k \geq k_{min}^\Psi$	leg and chest features are curved
ellipticality ε	$\varepsilon \leq \varepsilon_{max}^\Psi$	the fitting error of ellipse for chest under the threshold

C. People Model and Position Detection

3D projection of two planes of scan data from the layered sensors can be used to represent the position and direction of a person. The set of geometrical features extracted from the former step are mostly ellipses and circles. If they belong to a person another important criteria should be meet: the large elliptical segment should come from the upper layer and the small circles from the lower layer. No large ellipses are possible for a person in the leg area. The small circles can not be over the large ellipse (the person height is restricted according to the height of the upper layer).

To properly establish the previous requirements, it is necessary to associate segments in the upper layer with those in the lower layer, this is to find the corresponding legs for a given chest. Latt *et al.* [22] present a study about how human motion, step length, walking speed, etc. are selected to optimize stability. Their study present data about different speeds people prefer when walking. If the average values of step length are used then it is possible to define the limits of motion of the legs with respect to the projected chest elliptical area. Figure 6 helps understanding this idea. The average leg height h is about $84cm$, and the height of the lower layer of sensors l is fixed at $40cm$. s is the step length which depends on the speed, for example $73cm$ for an average speed of $1.2m/s$ [22]. d is calculated as:

$$d = 2(H-l)tan(\theta), \text{ where } \theta = sin^{-1}(\frac{s}{2h}). \tag{5}$$

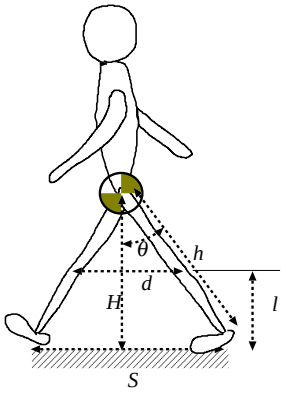


Fig. 6: Simple representation of human step to compute the distance d between leg segments while walking.

According to [22] the step lengths for three different walking speeds are presented in Table II. In this table we include the parameter d from Fig. 6 about the distance between leg segments when walking at the different speeds.

TABLE II: Step length according to speed and distance between leg segments d

Mode	Speed[a]	Step Length[b]	Distance between leg segments d [c]
normal	$1.2 \pm 0.04m/s$	$73.0 \pm 3cm$	$34.40cm$
very slow	$0.5 \pm 0.05m/s$	$47.0 \pm 3cm$	$22.29cm$
very fast	$2.1 \pm 0.1m/s$	$86.0 \pm 6cm$	$39.64cm$

[a,b] Values according to Latt *et al.* [22].
[c] estimated from Eq. 5.

With an estimation of the maximum value for d, the separation of legs at the lower layer height, we can set a search radius of $\frac{d}{2} \pm \xi$ at the center of the chest elliptical area projected into the lower layer to search for the corresponding legs for the chest. We use average walking step length from Latt *et al.* [22], at normal walking speed, to compute the value for d.

IV. Experimental Results

The robot used for our research was presented in Fig. 5, the computer operating the robot is a Intel Pentium Core Duo based notebook running (Linux kernel 2.6.24) as

operating system and robot control board is powered by a Hitachi SH-2 processor. The robot system uses *4 URG-04LX* range scanners from *Hokuyo Automatic Co., Ltd.*[20], small size (50*x*50*x*70mm), covers distances up to 5.6*m*, distance resolution of 10*mm* and angular resolution of 0.36^o, angular range of 240^o operating at 10*Hz*. Scan data from each sensor consists of 682 points circularly ordered according to scanning step.

Data from each sensor is read every 100*ms* by a driver processes and registered in parallel into a shared memory system (*SSM*[23]) based on IPC messaging and multiple ring-buffers with automatic timestamping, one driver process per sensor. SSM also allows to record raw sensor data into log files and to play it back with the same rate as the sensor (10*Hz* in this case).

Client processes read scan data from the ring-buffers according to sensor's pose (those in the top layer and those on the low layer), pairs of LRF sensors are processed in the fusion step, sensor layers are further fused and finally people position is computed. The processing time for the two layers (4 sensors), from single layer fusion to people position detection, was below 40*ms*, fast enough given the sensor's scanning speed.

We performed an experiment for people detection and position estimation from a mobile robot. In the experiment, 5 persons walked around the robot and additional person was taking the experiment video. Log data from each sensor was recorded, people position detection tests were performed off-line by playing back this log data using our SSM system. Fig. 7 corresponds to the group of people surrounding the robot.

Fig. 7: An experiment for multiple people position estimation using the proposed method.

Fig. 8 shows results of LRF data segmentation and feature extraction: raw data from each layer (top layer in Fig. 8(a)) is divided into clusters (Fig. 8(b)) and each cluster's indicators analyzed to extract those segments with human-like features and average sizes (Fig. 8(c)).

In this figure, arrows in Fig. 8(a) represent the location of people in the environment, most of them were successfully detected in the results of feature extraction (8(c)). However one of them has a height below the standard so top-level sensors were actually scanning his neck area, accordingly his chess ellipse is smaller than the allowed values, therefore was rejected. Another interesting case is the segment marked as "column" in Fig. 8(a), although its curvature and linearity indicators classify it as person, the boundary length and segment width were far bigger than the allowed values, reducing its scoring and marking it for rejection.

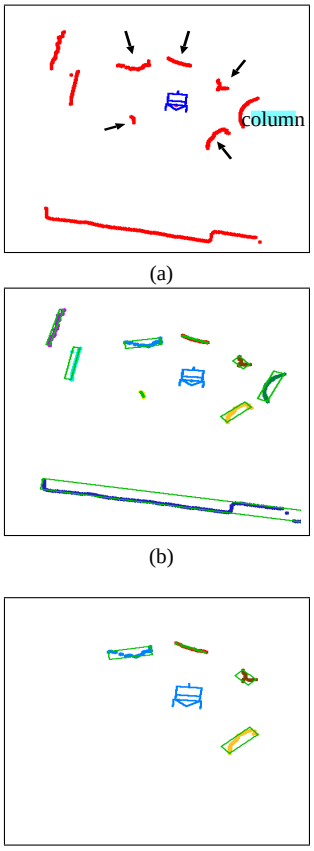


Fig. 8: Results of LRF data segmentation and feature extraction: raw data ((a)) is segmented ((b)) and then classified ((c)).

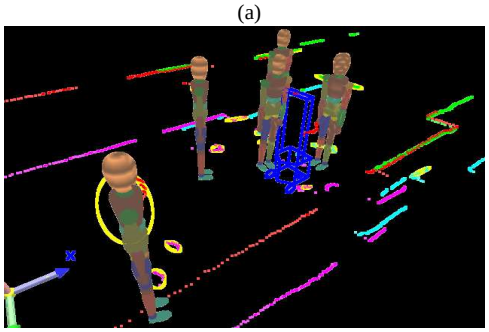
Fig. 9 shows the results of an experiment for people detection and position estimation from a mobile robot. In the experiment, 5 persons walked around the robot and additional person was taking the experiment video (Fig. 9(a)). Log data from each sensor was recorded, people position detection tests were performed off-line by playing back this log data using our SSM system.

A 3D tool was created to visualize inspect how the people detection worked; in Fig. 9(b) chest ellipses and leg circular ellipses are detected then we place a 3D wooden doll, as a representation of a person, in the estimated position the person should have. Results were verified by human operator comparing the experiment video with results.

The members have varied body sizes, from broad and tall to thin and short. Some of the members have a height a little under the average, as result their chest ellipses were not correctly detected in the people detection step. As presented in Fig. 9(b), the person to the right of the robot (represented with blue line segments) is missing although circles from legs are present.

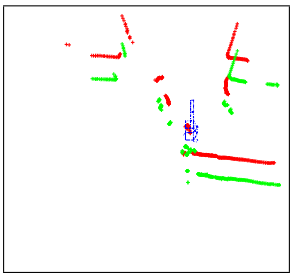
Additional snapshots of experimental results are presented

(a)

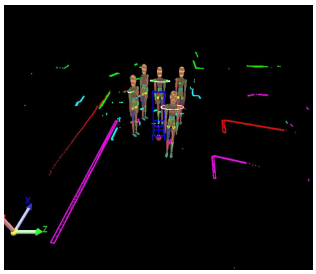


(b)

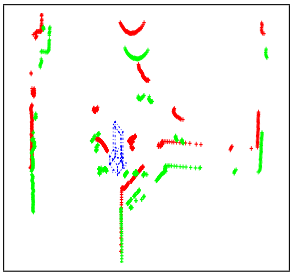
Fig. 9: Results of the people detection, (a) snapshot (b) 3D models in estimated positions of people in the experiment.



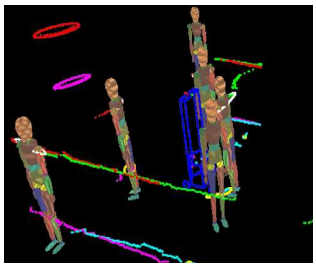
(a)



(b)



(c)



(d)

Fig. 10: Experimental results with raw scan data ((a) and (c)) and the corresponding people detection and position estimation ((b) and (d)).

in Fig. 10, the robot is represented in all cases as blue line segments. Fig. 10(a) and 10(c) shows raw scan data from both layers (red for the upper layer and green for the lower one), and in Fig. 10(b) and 10(d) a 3D representation of the human detection and position estimation. In the cases of 3D representation, the raw scan data is plotted together with wooden dolls enclosed in the estimated people positions represented with elliptical shapes, a large one for the chest area and smaller ones for the extracted leg areas.

In Fig. 10(c) there are two rather large arc-like segments in the raw scan image and two large elliptical shapes in the 3D representation in Fig. 10(d), in both layers. That is the column inside the indoor environment, as already explained in Fig. 8(a). The people detection method discards this elliptical object because its dimensions are larger than the expected for people, those elliptical objects are represented with red color in this figure. Also we do not expect large elliptical objects from the lower layer so discarding this column as a non human object was simple.

V. Conclusions and Future Works

The problem of multiple people position detection by fusion of multiple LRF sensors arranged in a double layer structure was presented in this paper. Instead of using different sensors of complementary capabilities, we used the same type but at different heights (layers), this gives a different perspective which also helps solving simple cases of occlusion where one sensor is occluded and the other is not.

The addition of an extra layer of LRFs to detect chest elliptical areas improve the estimation of people position as the lower part of body (the legs) move faster and wider than the chest area. The combination of both areas creates a 3D volume which helps locating the position of the person more closely related to the center of this 3D volume and as a measure of the possible direction the person is facing. Although research exists in the area of detection and tracking, the proposed approach is simple and fast enough to be used for real time detection of people in robot's surroundings.

As future work, multiple people tracking will be considered. Also the effectiveness of our method in cluttered environments will be studied. Future steps of our research include understanding people group motion and recognition of group members.

References

[1] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "The interactive museum tour-guide robot," in *Fifteenth National Conference on Artificial Intelligence*, (Madison, WI), pp. 11–18, July 1998.

[2] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Häehnel, C. R. Rosenberg, N. Roy, J. Schulte, and D. Schulz, "Minerva: A tour-guide robot that learns," in *23rd Annual German Conference on Artificial Intelligence: Advances in Artificial Intelligence*, (Pittsburgh, PA), pp. 14–26, September 1999.

[3] A. Fod, A. Howard, and M. J. Matarić, "Laser-based people tracking," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Washington D.C.), pp. 3024–3029, May 2002.

[4] M. Montemerlo, S. Thrun, and W. Whittaker, "Conditional particle filters for simultaneous mobile robot localization and people-tracking," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Washington, D.C.), pp. 695–701, May 2002.

[5] J. Xavier, M. Pacheco, D. Castro, A. Ruano, and U. Nunes, "Fast line, arc/circle and leg detection from laser scan data in a player driver," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Barcelona, Spain), pp. 3941–3946, April 2005.

[6] J. Cui, H. Zha, H. Zhao, and R. Shibasaki, "Robust tracking of multiple people in crowds using laser range scanners," in *IEEE 18th International Conference on Pattern Recognition (ICPR)*, (Hong Kong, China), pp. 857–860, August 2006.

[7] J. H. Lee, T. Tsubouchi, K. Yamamoto, and S. Egawa, "People tracking using a robot in motion with laser range finder," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Beijing, China), pp. 2936–2942, October 2006.

[8] H. Zhao, Y. Chen, X. Shao, K. Katabira, and R. Shibasaki, "Monitoring a populated environment using single-row laser range scanners from a mobile platform," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Roma, Italy), pp. 4739–4745, April 2007.

[9] K. O. Arras, Ó. Martínez Mozos, and W. Burgard, "Using boosted features for the detection of people in 2d range data," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Roma, Italy), pp. 3402–3407, April 2007.

[10] Z. Zivkovic and B. Kröse, "Part based people detection using 2d range data and images," in *IEEE International Conference on Intelligent Robots and Systems (IROS)*, (San Diego CA), pp. 214–219, October 2007.

[11] M. Mucientes and W. Burgard, "Multiple hypothesis tracking of clusters of people," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Beijing, China), pp. 692–697, October 2006.

[12] K. O. Arras, S. Grzonka, M. Luber, and W. Burgard, "Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Pasadena CA), pp. 1710–1715, May 2008.

[13] P. Kondaxakis, S. Kasderidis, and P. E. Trahanias, "A multi-target tracking technique for mobile robots using a laser range scanner," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Nice, France), pp. 3370–3377, September 2008.

[14] A. Carballo, A. Ohya, and S. Yuta, "Fusion of double layered multiple laser range finders for people detection from a mobile robot," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, (Seoul, Korea), pp. 677–682, August 2008.

[15] S. Gidel, P. Checchin, C. Blanc, T. Chateau, and L. Trassoudaine, "Pedestrian detection method using a multilayer laserscanner: Application in urban environment," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Nice, France), pp. 173–178, September 2008.

[16] M. Hashimoto, Y. Matsui, and K. Takahashi, "People tracking with in-vehicle multi-laser range sensors," in *2007 Annual Conference of the Society of Instrument and Control Engineers (SICE)*, (Kagawa, Japan), pp. 1851–1855, September 2007.

[17] S. Yuta, S. Suzuki, and S. Iida, "Implementation of a small size experimental self-contained autonomous robot-sensors, vehicle control, and description of sensor based behavior," *Lecture Notes in Control and Information Sciences: Experimental Robotics II, The 2nd International Symposium, Toulouse, France, 1991*, vol. 190, pp. 344–358, 1993.

[18] H. Kawata, S. Kamimura, A. Ohya, J. Iijima, and S. Yuta, "Advanced functions of the scanning laser range sensor for environment recognition in mobile robots," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, (Heidelberg, Germany), pp. 414–419, September 2006.

[19] K. C. Dietmayer, J. Sparbert, and D. Streller, "Model based classification and object tracking in traffic scenes from range-images," in *IEEE Intelligent Vehicles Symposium (IVS)*, (Tokyo, Japan), May 2001.

[20] Hokuyo Automatic Co. Ltd., "Hokuyo Scanning Range Finder (SOKUIKI) Sensor." `http://www.hokuyo-aut.jp`.

[21] M. Pilu, A. W. Fitzgibbon, and R. B. Fisher., "Ellipse-specific direct least-square fitting," in *IEEE International Conference on Image Processing (ICIP)*, vol. 3, (Lausanne, Switzerland), pp. 599–602, September 1996.

[22] M. D. Latt, H. B. Menz, V. S. Fung, and S. R. Lord, "Walking speed, cadence and step length are selected to optimize the stability of head and pelvis acceleration," *Experimental Brain Research*, vol. 184, pp. 201–209, January 2008.

[23] E. Takeuchi, T. Tsubouchi, and S. Yuta, "Integration and synchronization of external sensor data for a mobile robot," in *Society of Instrument and Control Engineers (SICE)*, vol. 1, (Fukui, Japan), pp. 332–337, August 2003.