

Reliable People Detection Using Range and Intensity Data from Multiple Layers of Laser Range Finders on a Mobile Robot

Alexander Carballo · Akihisa Ohya · Shin'ichi Yuta

Accepted: 14 December 2010 / Published online: 7 January 2011
© Springer Science & Business Media BV 2010

Abstract Reliable people detection is an important task in several areas like security, intelligent environments and human robot interaction. People detection does not depend only upon separation of static environment objects from those showing motion (hopefully humans), a reliable system should be able to detect static people even in cluttered environments.

This work presents a reliable approach for people detection and position estimation using multiple layers of Laser Range Finders (LRF) on a mobile robot. Each layer combines two LRF sensors to scan around the robot's surroundings and are vertically separated to detect distinct parts of the human body. By using AdaBoost we create strong classifiers to detect body parts, candidate segments in each layer are fused for people detection, and we use simple data association to estimate their positions. Additionally, this work introduces laser reflection intensity as a novel property for people detection. First, we present a study of laser intensity and textiles, then introduce new intensity-based features for detection, and propose a method for segment separation using laser intensity. We provide a thorough evaluation of our multi-layered system through several experiments on a mobile robot.

Keywords Multi-layered laser range finder · Laser reflection intensity · People detection · AdaBoost · Sensor fusion

1 Introduction

To successfully interact with humans, a companion mobile robot requires the ability to detect people, to predict how or where we move, what motion patterns we follow, and even to recognize who it is interacting with. Furthermore, human detection must be robust enough so that the robot can operate in our dynamic living and working environments.

Laser Range Finders (LRF) besides being used for obstacle detection are also an important part of people detection and tracking systems. LRF present important advantages over other sensing devices (e.g., vision) like high accuracy, wide view angles, high scanning rates, little sensibility to illumination changes, etc., and are becoming more accessible and safer to use in human environments. Renown examples of such LRF applications are the Tour-Guide robots Rhino and Minerva by Burgard *et al.* [1] and Thrun *et al.* [2], and more recently the Autonomous City Explorer (ACE) by Bauer *et al.* [3], featuring LRF for navigation, people detection and collision avoidance.

Most approaches based on LRF [4–9] place the sensors at the same height (single row) to detect and track some parts of the human body. Due to laser safety regulations, applications using non class-1 lasers are mostly limited to lower positions, mostly about knee height. Thus legs are widely used for human detection and tracking.

One of the existing problems is how to correctly identify humans from laser measurements. Arras *et al.* [10] uses AdaBoost (Schapire *et al.* [11]) to determine which properties

A. Carballo (✉) · A. Ohya · S. Yuta
Intelligent Robot Laboratory,
Graduate School of Systems and Information Engineering,
University of Tsukuba, 1-1-1 Tennoudai, Tsukuba City,
Ibaraki Pref., 305-8573, Japan
e-mail: alexandr@roboken.esys.tsukuba.ac.jp

A. Ohya
e-mail: ohya@roboken.esys.tsukuba.ac.jp

S. Yuta
e-mail: yuta@roboken.esys.tsukuba.ac.jp

to consider to improve detection. Still, detection of multiple people in cluttered environments is difficult especially considering occlusion cases of persons walking side by side or behind objects. One interesting solution was proposed by Hashimoto *et al.* [12], where several LRF are placed around a wheelchair at different heights and orientations for 360° scanning, each sensor with its own processing computer for independent detection and tracking.

From Hashimoto's multi-laser idea one can consider a *multi-layered arrangement of LRF*: rather than a single row (single-layered) system, arrange the sensors in multiple rows each one separated vertically by some height. Multi-layered detection systems are thus tolerant to occlusions, the partial or total occlusion of one layer has little effect when other layers are still available. Multi-layered systems have the additional benefit of complementarity for detection and tracking: a human can be represented by the *union of several sets of scan segments*, one or two small segments at low height (legs), one segment at medium height (waist), a large segment at a high position (chest) and even one small segment at a higher position (head). An immediate drawback is cost for the relative larger number of sensors needed for a multi-layered system (at least two) when a simple 3D LRF can provide tens to hundreds of layers. In spite of that, multi-layered systems can provide multiple scan planes in parallel at the speed of one single scan (in the order of tens to hundreds of Hertz), which is ideal for real-time detection and tracking systems.

Multi-layered detection systems have been developed recently [13, 14]. Gidel *et al.* [13] used a 4 layered sensor for pedestrian detection: the IBEO ALASCA XT scanning in 4 planes simultaneously, spaced over a vertical FOV of 3.2°. Martínez Mozos *et al.* [14] uses 3 layers of 2D laser scanners to simultaneously detect different sets of features for a human: head, chest and legs. They use AdaBoost to train classifiers for every type of feature and use probabilistic voting for detection allowing tolerance to occlusions.

Most of the works on people detection/tracking using LRF require some form of data segmentation (clustering) at some point. Segmentation techniques are categorized in

two main groups: Point-Distance and Predictive [15–17]. A problem using LRF is the capacity to distinguish different objects standing very close, when two physically different objects are very close together (e.g., two humans walking close) their LRF points might get merged into one large segment and become a false-negative in the detection stage. Existing solutions include adding time dimension to LRF segments and measure segment motion between consecutive scans [10], keeping tracks for one or multiple targets [18–21], or else using complementary sensors [22]. Tracking can deal with temporal occlusions due to obstacles, such as the temporal disappearance of legs behind a dust bin. One limitation still present is occlusion of the tracked body feature for an extended time, e.g., the human stops behind the dust bin.

Our main research problem aims to develop a companion robot with the particular objective to study the relationship of an autonomous mobile robot as a member of a group of people, where the robot is to move and behave like another member of the group. For that aim we require a robot and detection system with the following requirements:

- Scan from all around the robot
- Simple yet robust enough for detection and tracking of group members
- Fast response for navigation tasks inside the group

For that goal we introduced a multi-layered people detection system [23]: four LRF sensors are installed in a mobile robot, arranged in pairs on two layers (Figs. 1(a) to 1(b)). Data from each independent layer is processed in parallel: scan data is first divided into segments, for each segment compute a set of features which serve for classification of candidate segments. Candidate segments correspond to complementary body parts from each human around the robot: legs are extracted from the bottom layer and chests from the top layer (Fig. 1(c)). Finally those detected parts are fused to finally detect actual humans and their positions (Fig. 1(d)). Our system is tolerant to occlusions by keeping simple data association.

Our multi-layered system differs from other related works in several ways. The in-sensor layered system in

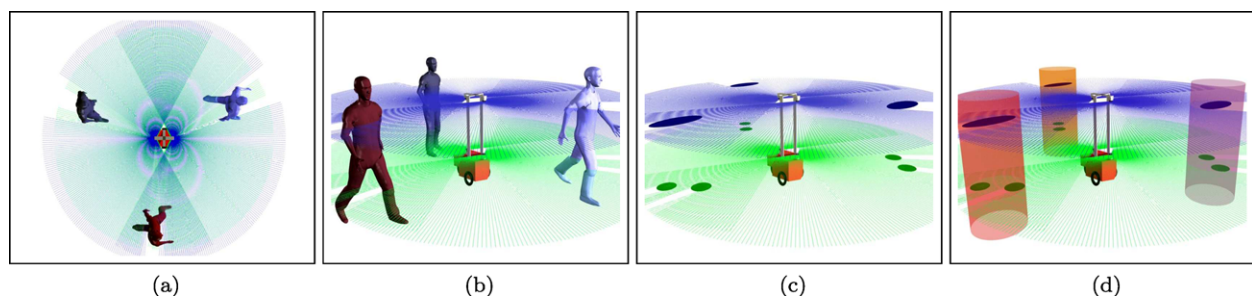


Fig. 1 Scanning from a double layered approach: (a) sensors facing in opposite directions (bird's-eye view), (b) two layers of sensors (frontal view), (c) detection of body parts in each layer (dark elliptical shapes), and (d) people detection by fusion of layers

Gidel *et al.* [13] has a narrow field-of-view so the vertical distance between layers depend on the distance from the sensor, constraining the extraction of different sets of features. On the contrary, our system uses a parallel arrangement of layers (vertically separated), and scan data is read and processed in parallel. Different from Martínez Mozos *et al.* [14], our system is installed on a mobile robot, each layer scans the complete surroundings of the robot, it is designed for multiple people detection, and our approach involves a fusion of layers step using a simple human walk model.

Our approach is simple enough to be implemented in a mobile robot by placing sensor layers in the robot body at different heights from the ground depending on the features to detect, and can be easily extended to support additional layers by installing more sensor pairs. Despite its simplicity, our system handles large number of people with good detection rates, and processing time is well under the sensor scan time.

Different from other related works, we introduce the use of *laser reflection intensity* data for the people detection problem as an extension of our system [24]. First, we present a study of how laser intensity changes with colors and materials of the textiles we commonly wear. We propose a simple method to reduce the effect of target distance over intensity values. We introduce several intensity-based features used for training strong classifiers to detect candidate segments, which allow increasing detection rates. Laser intensity is also introduced here to separate merged laser segments which come from different objects based on their different reflection properties. To our knowledge no other work considers laser reflection intensity for people detection and segment separation.

The contributions of this work can be summarized as follows:

- Multi-layered system for reliable people detection from a mobile robot.
- A study of the effects of textiles and colors on laser reflection intensity for people detection applications.
- Introduction of laser reflection intensity as a new feature for people detection.
- A new method for separation of segments using laser reflection intensity.
- A simple technique for calibration of laser reflection intensity for low power LRF.

The rest of the paper is organized as follows: in Sect. 2 we present an overview of the hardware and software of our system. In Sect. 3 we present our single-layer approach for classification of segments. In Sect. 4 we provide a study of laser reflection intensity as a new tool for people detection, considering how it behaves with common textiles, intensity-based features for detection and its use for merged segment separation. Section 5 presents our multiple-layered

approach to combine individual layer detection results. Section 6 presents experimental results and evaluations of the reliability of our people detection approach and in Sect. 7 we elaborate about results and our current system. Finally, conclusions and future work are left for Sect. 8.

2 System Overview

In this section, we describe our current robot platform and the software architecture of our multi-layered detection system.

2.1 Hardware

The robot used for our research is based on our differential drive *Yamabico* robotic platform depicted in Fig. 2. Two layers of LRF sensors are used (Fig. 2(a)): the bottom layer is about 40 cm from the ground whereas the top layer is about 110 cm. Every layer consists of two LRF sensors: one facing forwards and another facing backwards for a 360° coverage. The robot center-of-mass is in a low position and the upper frame is light-weighted to minimize oscillative motion. The robot dimensions are 40 × 38 × 120 cm (length, width and height).

The four LRF sensors are connected to the main processing system using USB interface (Fig. 2(b)). Scan data is read by four independent driver processes and integrated into a shared memory manager with automatic timestamps. For data processing and robot control we use a notebook computer with an Intel Core 2 Duo processor at 1.83 GHz, 2 GB of RAM, running Linux (kernel 2.6.31) as operating system. Robot wheels are driven by 40 W motors with rotary encoders. A SH2 micro-controller is used for wheel rotation count and velocity control. The actual implemented robot corresponds to Fig. 2(c).

The LRF sensors used in our system are the *URG-04LX* laser range scanners from *Hokuyo Automatic Co., Ltd.* (Fig. 2(d)): small size and weight (5 × 5 × 7 cm and 160 gr), operates at 10 Hz, covers up to a maximum range of 5.6 m with a range resolution of 10 mm, and scans an angular range of 240° with an angular resolution of 0.36°. The laser inside this sensor is a near-infrared (NIR) solid state laser with wavelength 785 nm. Scan data from each sensor consists of 682 points circularly ordered (from right to left) according to scanning step. Kawata *et al.* [25] provides a good analysis of the sensor operation (consider also [26, 27] for a detailed comparison of this sensor with other LRF).

2.2 Software

The processing pipeline of our system is represented in Fig. 3. Our people detection approach involves three ma-

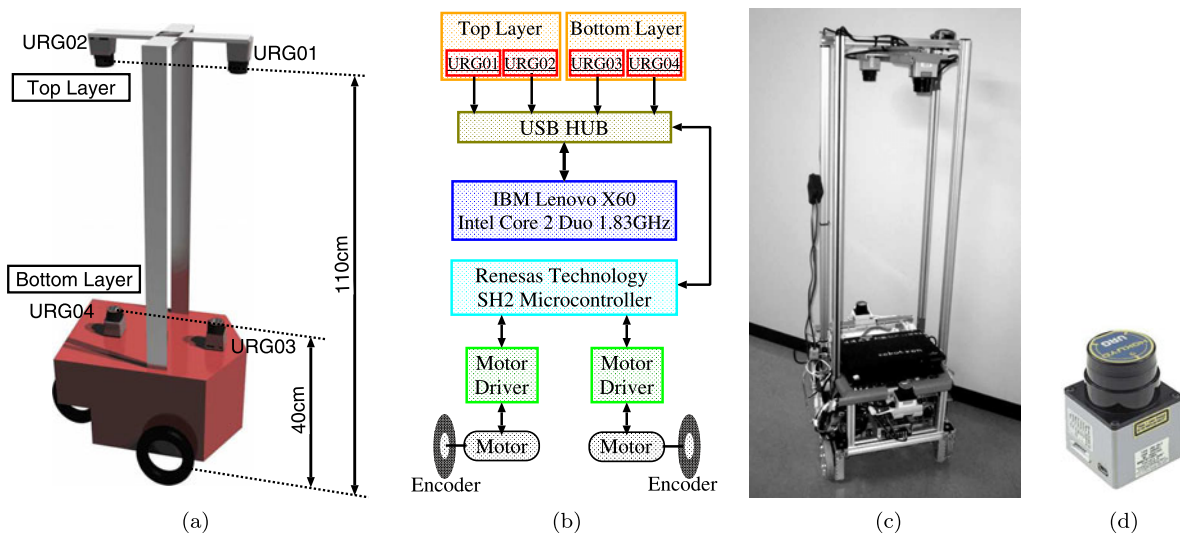


Fig. 2 Our robot system for multiple people detection: (a) sensor layers, (b) hardware architecture, (c) the implemented robot, (d) the *URG-04LX* laser scanner

major steps: single-layer classification, segment separation using laser intensity, and multi-layered people detection. Some outputs for intermediate steps are presented as insets on the figure and correspond to actual scan data. Figure 3(a) is the result of fusion of sensors (the top layer is represented by darker points and the robot is represented as a filled polygon), Fig. 3(b) corresponds to single-layer classification (possible body parts are represented as ellipses). As a novel input in the people detection process, we use the reflection intensity from each of the LRF sensors. Figure 3(c) represents this laser intensity input (brighter points correspond to higher intensity levels). Finally, in Fig. 3(d) the detected human around the robot (represented by a cylinder) by fusion of layers.

In the single-layer classification step, sensors facing opposite directions in the same layer are combined, then fused data is divided into segments, for each segment we compute a set of properties (features), and finally select segments which possibly belong to people (which we call *candidate segments*). Candidate segments are those which most likely correspond to some part of the human body according to some confidence level; we consider here segments from the chest area (top layer) and segments from the legs area (bottom layer).

In an intermediate step, we seek to improve single-layer detection by detecting and solving cases where a human is so close to an object or another human that they form one large LRF segment (their individual LRF segments become merged) and cannot be detected. For that we use the value of the laser intensity reflected from the objects. To further improve detection results, we use laser intensity-based features as part of the segment classification sub-step.

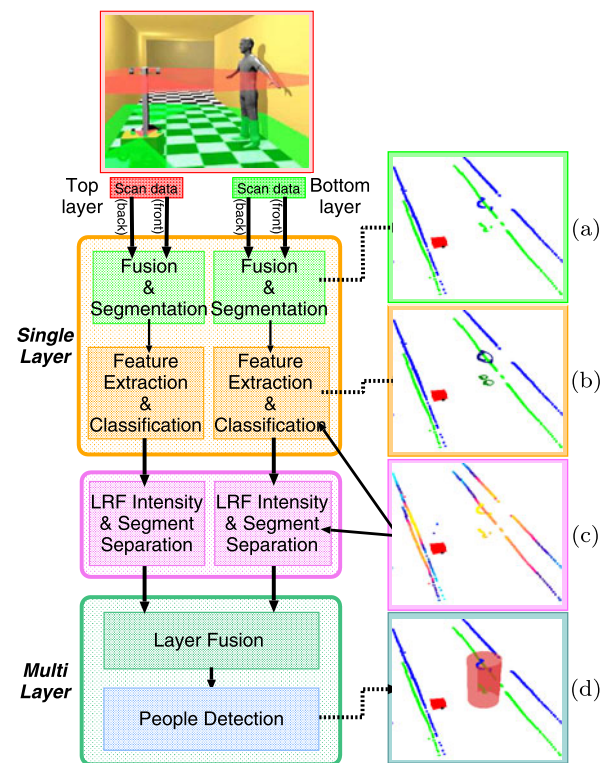


Fig. 3 System overview: (a) fusion of sensors, (b) single layer classification, (c) laser reflection intensity as input, and (d) fusion of layers for people detection

In the final multi-layered step, we fuse the individual detection results from each layer to robustly detect people and their position. Candidate segments from each layer correspond to individual body parts, therefore we try to find for

each chest segment its corresponding leg segments¹ according to a simple human walking model.

The individual steps of single-layered classification and multi-layered people detection are explained in the following sections. Although segment separation using laser intensity is formally part of the single-layered classification step, the use of laser intensity for people detection is a novel contribution on this work, therefore we study in detail some of its applications in a separate section.

3 Single-layer Classification

As described in Sect. 2.2, data from the LRF sensors is read and processed independently in each layer, layer data is divided into segments which are then classified according to their geometrical properties. The steps in Sects. 3.1 and 3.2 are briefly explained here. Please refer to [24] for details.

3.1 Sensor Fusion

The robot scans its surroundings from both sensor layers having two sensors a and b in each layer, one facing forward and one facing backwards. Each sensor a and b provide scan data $\mathbf{C}^a$ and $\mathbf{C}^b$ of the form $\mathbf{C}^a = \{p_1^a, \dots, p_m^a\}$ and $\mathbf{C}^b = \{p_1^b, \dots, p_n^b\}$ for m and n valid scan beams and each scan beam p_i consists of range and direction $\langle r_i, \psi_i \rangle$, where $\psi_i \in [-120^\circ, 120^\circ]$ for our LRF sensors.

We apply a transformation $\mathcal{R}: \mathcal{L} \rightarrow \mathcal{G}$ from each sensor's local coordinate system $\mathcal{L}$ into robot reference system $\mathcal{G}$. With $\mathbf{Q} = \mathcal{R}\{\mathbf{C}\}$ we obtain two new sets $\mathbf{Q}^a$ and $\mathbf{Q}^b$ in $\mathcal{G}$ system, and each scan point $q_k \in \mathbf{Q}$ with transformed range and direction $q_k = \langle \rho_k, \varphi_k \rangle$.

Data from $\mathbf{Q}^a$ and $\mathbf{Q}^b$ is fused considering the overlapping areas in scan readings (dark slices in Fig. 1(a)): for every point q_i^a in $\mathbf{Q}^a$ in one of the overlapping areas we find the nearest point q_n^b in $\mathbf{Q}^b$ in the corresponding overlapping area, then we obtain and store the average point $f_j = \langle \frac{\rho_i^a + \rho_n^b}{2}, \frac{\varphi_i^a + \varphi_n^b}{2} \rangle$ in the fused set $\mathbf{F}^{ab}$; points from $\mathbf{Q}^{a|b}$ outside the overlapping areas are directly copied into $\mathbf{F}^{ab}$.

3.2 Segmentation

Segmentation is understood here as the problem of break-point detection: finding a gap between a set of points that is big enough not to be considered as noise and that clearly divides the set of points in two or more segments. Our segmentation technique falls in the category of Point-Distance (Borges *et al.* [15]). Consider two neighboring points f_j and

$f_k \in \mathbf{F}^{ab}$, separated by a distance $D(f_j, f_k) = |\rho_j - \rho_k|$ and by an angle $\Delta\varphi_{jk} = |\varphi_j - \varphi_k|$. By using a threshold function $\mathcal{T}$ to measure the gap between neighbor points, a new scan point f_n is compared to the last known member f_m of a given segment S_i , if the test $D(f_m, f_n) \leq \mathcal{T}(f_m, f_n)$ fails then they are assigned to different segments S_i and S_j (i.e., f_n marks the beginning of the new segment S_j).

We define an adaptive threshold function $\mathcal{T}$ based on [28], as:

$$\mathcal{T}(f_j, f_k) = \mathcal{E}(f_j) + \min(\rho_j, \rho_k) \sqrt{2(1 - \cos(\Delta\varphi_{jk}))} \quad (1)$$

$\mathcal{E}(\rho_j)$ corresponds to the size of the radial error at range ρ_j , it is defined according to LRF's specifications as:

$$\mathcal{E}(\rho_j) = \begin{cases} 10 \text{ (mm)} & \text{if } \rho_j \leq 1000 \text{ (mm)} \\ 0.01 \times \rho_j \text{ (mm)} & \text{otherwise} \end{cases} \quad (2)$$

An example of the adaptive threshold function $\mathcal{T}$ for segmentation, using actual scan data, is presented in Fig. 4; data was divided into segments using function $\mathcal{T}$ as jump distance and the radius of the circles corresponds to this threshold function. Using function $\mathcal{T}$ we avoid unnecessary break points due to noise with the radial error term $\mathcal{E}$, which additionally allows to keep together neighboring points particularly when they have large range values because the distance between points increases as their range increases. The last term in Eq. 1 computes the distance of the two unitary vectors in the same direction of f_j and f_k .

To locate the closest neighbor f_k for a given point f_j , we keep the scan points in a kdTree data structure and implement a bounded k-NearestNeighbors (k -NN) search, to find the nearest neighbors inside a search radius $\mathcal{T}$ centered at point f_j , based on the radially bounded nearest neighbor (RBNN) by Klasing *et al.* [29].

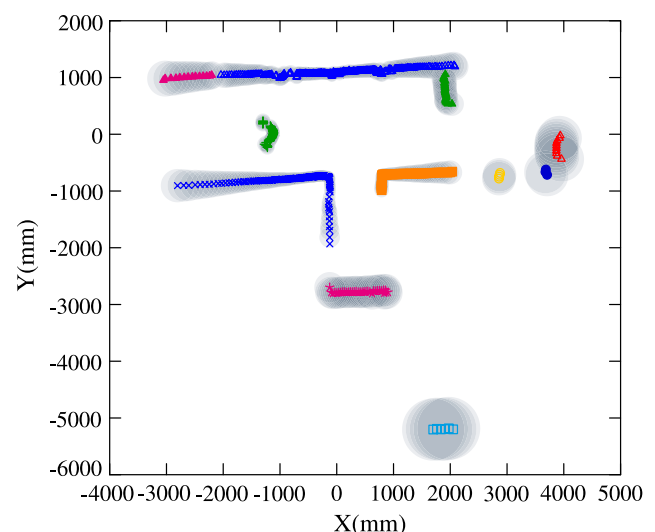


Fig. 4 Adaptive threshold function $\mathcal{T}$ for segmentation (circles have been enlarged to ease visualization)

¹One or two legs, one leg may partially or totally occlude the other or even legs are too close that form a single scan segment, our detection approach allows such cases.

Table 1 Range-based features for people detection (based on [10])

No.	Feature $\mathbf{h}(S_i)$
1.	Number of points (N) of the segment
2.	Width (w), longest side of the bounding-box enclosing the segment $w = \max(W, H)$, with W and H the sides of the box
3.	Size ratio (ℓ), ratio of the sides of the segment's bounding-box $\ell = \max(W, H)/\min(W, H)$
4.	Linearity (L), variance of the residuals to the best fitting line
5.	Circularity (c), variance of the residuals to the best fit circle
6.	Radius (R), from the best fitting circle
7.	Ellipticity (ε), variance of the residuals to the best fit ellipse
8.	Boundary length (bl), average distance between points $bl = \frac{1}{N} \sum_{i=1}^{N-1} D(f_i, f_{i+1})$
9.	Boundary regularity (br), standard deviation of the boundary length
10.	Mean curvature ($\bar{\kappa}$), mean for each of the curvatures κ_i computed from the triangle $\triangle f_{i-1} f_i f_{i+1}$, every three points, as $\kappa_i = 4A/(D(f_{i-1}, f_i)D(f_i, f_{i+1})D(f_{i-1}, f_{i+1}))$ with A the triangle's area
11.	Mean angular difference ($\bar{\phi}$), mean value of the angles ϕ_i formed every three points as $\phi_i = \angle(\overline{f_{i-1} f_i}, \overline{f_i f_{i+1}})$
12.	Normalized number of points ($\hat{N}$), the ratio of the actual number of points and the maximum expected number of points at a given range, $\hat{N} = 2N\rho \tan(\vartheta/2)/w_{\max}$, where ρ is the range to the segment center, ϑ is the angular resolution of the sensor and $w_{\max}$ the maximum segment width

3.3 Feature Extraction and Classification

The idea of feature extraction and classification is to match the sensor readings with geometrical models representing expected behavior of data. As example, if a LRF sensor is scanning a wall it is expected that scan data forms a straight line. Also if the sensor is scanning a human, the expected behavior is a set of points forming an arc. To identify walls we need to classify segments according to their fitness with some line model (linearity), for humans select those segments according to their fitness with some curve model (circularity or ellipticity).

In the least squares sense, fitting scan data to some model (in particular curves) will try to find the set of model parameters that minimizes the sum of squared residuals, using the Cartesian coordinates (x, y) of scan data as factors for the overdetermined system. Fitting a circle to a scan segment which forms a straight line (e.g., a wall), may yield a large circle with small residual error that cannot represent any human being—concerning physical size. As such, fitness to some curve model is not enough for accurate people detection. On the other hand, if we include additional features such as segment width, the circle radius, etc., we can achieve better classification. This means, a rich set of simple features allows accurate guessing of the type of the segment: human or environment object (walls and others).

For every segment S_i we compute a set of n features $\mathbf{h}(S_i) \in \mathbb{R}^n$ to judge whether the segment corresponds to body part (chest or leg) or not. In Table 1 we present the complete list of extracted features from every segment, based on Arras *et al.* [10]. We introduce the *Size ratio* (ℓ) and *Normalized number of points* ($\hat{N}$) features. ℓ corresponds to be ratio of the sides of the bounding-box enclosing the

segment (see Fig. 11) and measures the segment thinness (a segment from a wall will have a larger ℓ than a human segment). $\hat{N}$ is the ratio of the current number of points to the expected maximum number of points at a certain range. Here $w_{\max}$ is the maximum expected width for the body part (currently 800 mm for a chest and 450 mm for legs); since the number of scan points of an object decreases with its range to the sensor,² $\hat{N}$ allows a range-independent measurement of the segment size whereas the number of points feature N establishes the minimum number of points of a segment.

To effectively classify body parts using the list of features $\mathbf{h}(S_i)$, we use the generalized AdaBoost algorithm [11], following the description in Arras *et al.* [10], to train a strong classifier $\mathcal{H}$. For that purpose we first form a set of m labeled training examples $\mathcal{X} = \{(\mathbf{h}(S_1), l_1), \dots, (\mathbf{h}(S_m), l_m)\}$ with $l_i \in \{+1, -1\}$ the segment's label where $l_i = +1$ is for positive examples (a human) and $l_i = -1$ for negative examples (an environment object).

Similar to [14], to create the training vector $\mathcal{X}$ we define obstacle free areas using our LRF data visualization tool: any segment (from any layer) which enters these areas is automatically labeled as positive (a human) and the rest as negative. We give an example in Fig. 5 where a person (labeled as P_0) walks inside a cluttered office: all segments falling inside painted areas in Fig. 5 are labeled as positive samples. This method works both for on-line and offline seg-

²Given the sensor's angular resolution ($\vartheta = 0.36^\circ$), an object of 800 mm wide will have at most 127 scan points at 1000 mm range and only 23 scan points at the maximum range of 5600 mm. To observe 100 points of this object at 1000 mm ($\hat{N} = 0.76$) is reasonable but a 100 points observation at 5600 mm ($\hat{N} = 4.40$) is not.

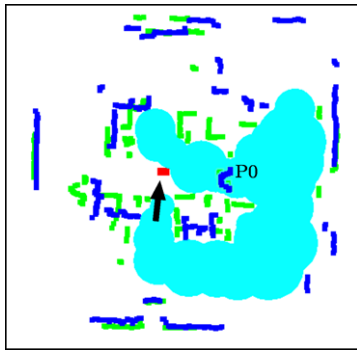


Fig. 5 Automatic labeling of segments for complex scenarios, darker segments correspond to top layer data and the robot is marked by an arrow

ment labeling. We labeled samples for different experimental environments, both from a static and from a mobile robot. Finally, from the large collection of labeled scan data we randomly selected entries to define $\mathcal{X}$ to avoid overfitting.

Based again on [10, 11], the final hypothesis (strong classifier) $\mathcal{H}$ has the form:

$$\mathcal{H}(S) = \text{sign} \left(\sum_{i=1}^T \omega_i g_i(S) \right) \quad (3)$$

where the weak classifier function $g_i(S)$ evaluates the i -th feature $h_i(S)$ as follows:

$$g_i(S) = \begin{cases} +1, & \text{if } s_i h_i(S) < s_i \theta_i \\ -1, & \text{otherwise} \end{cases} \quad (4)$$

After the feature boosting process, the algorithm learns the parameters: ω_i which is a weight applied to the weak classifier g_i , θ_i the threshold for the feature h_i and $s_i \in \{+1, -1\}$ the sign which controls the direction of the inequality in Eq. 4.

Feature classification depends also on the layer Ψ of the segment (top layer or bottom layer). For example the width w of a chest segment is expected to be larger than the same feature for a leg segment and we do not want to misclassify environment objects as humans (e.g., columns). Thus our final strong classifier $\mathcal{H}_\Psi(S)$ is:

$$\mathcal{H}_\Psi(S) = \begin{cases} \mathcal{H}_t(S), & \text{if } S \text{ is in the top layer} \\ \mathcal{H}_b(S), & \text{if } S \text{ is in the bottom layer} \end{cases} \quad (5)$$

The definition of $\mathcal{H}_\Psi$ in Eq. 5 is not limited to two layers, but is rather a physical implementation of the proposed system. Nevertheless, this means that each layer's strong classifier must be learned separately.

One parameter of the AdaBoost algorithm is the number of iterations T in Eq. 3, which also corresponds to the number of weak classifiers that form the final strong classifier $\mathcal{H}_\Psi$. In our implementation we choose $T = 100$ which

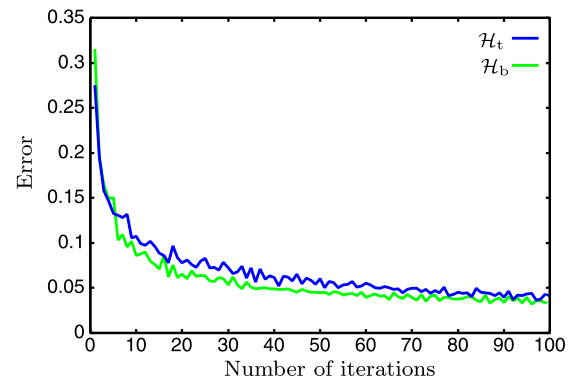


Fig. 6 Classification error vs. the number of iterations T for the strong classifiers $\mathcal{H}_t$ and $\mathcal{H}_b$

showed good performance and small training errors as presented in Fig. 6. This graph was obtained while running training of $\mathcal{H}_\Psi$, for both layers separately.

The final step of this classification process is to label segments as possible body parts or as environment object using the output from $\mathcal{H}_\Psi$ as:

$$\text{label}(S) = \begin{cases} \text{"candidate"}, & \text{if } \mathcal{H}_\Psi(S) > 0 \\ \text{"object"}, & \text{otherwise} \end{cases} \quad (6)$$

4 Laser Intensity for People Detection

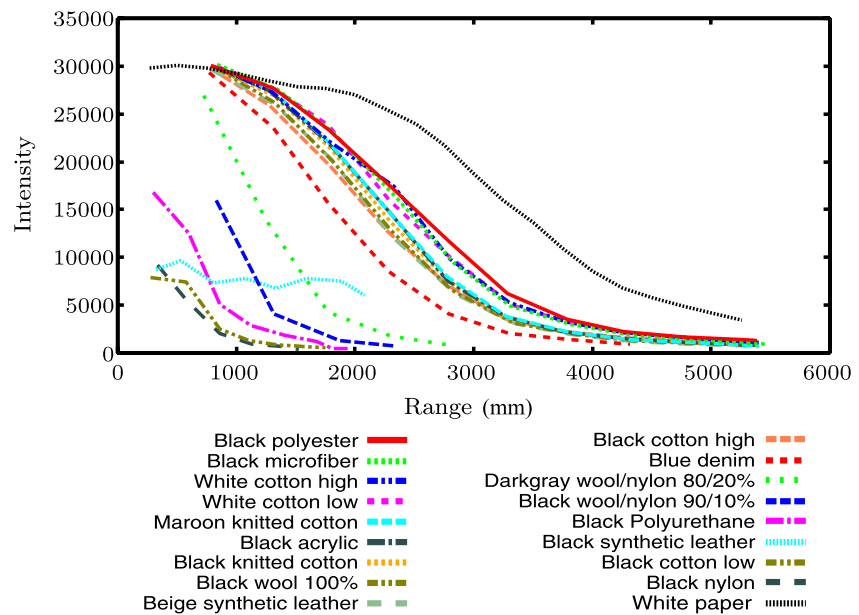
Besides range data, most LRF sensors provide also the reflection intensity (reflectivity) for every laser beam. This property has been rarely considered; one pioneer work using laser intensity from a mobile robot is attributed to Hancock [30] where a model of laser reflection was proposed which characterizes the response of the reflected laser depending on surface albedo, roughness, range to target, etc. Recently laser intensity is gaining interest from the research community. Nüchter *et al.* [31, 32] uses range and intensity data, together with Haar-like features for object classification. In Montemerlo *et al.* [33] laser intensity is used for extracting road lanes from an autonomous vehicle. Ye [34] uses laser range and intensity to detect and remove mixed pixels: spurious range measurements between two different surfaces. In the work by Rusu *et al.* [35], they combine range and intensity information to detect objects in human environments, particularly the handle of a door in a kitchen.

Laser reflection intensity of laser depends upon several properties:

- Color
- Material and Roughness
- Distance to the source (sensor)
- Power of the laser beam, wavelength, etc.

Colors and materials of objects in indoor environments (walls, doors, windows, metallic objects, etc.) are usually

Fig. 7 Different textiles and their raw laser reflection intensity levels for the *URG-04LX*, the selection include diverse materials, colors, textures, and densities



different from clothes. Thus intensity of the reflected laser beam from environment objects is also different from that of humans, and can be considered for separation of people from objects. Some form of identification could also be achieved with a proper laser characterization [30].

The strength of the received intensity value depends on the power of the emitted laser beam. High powered laser scanners achieve longer ranges and can provide detailed range/intensity images (e.g., the picture-like intensity images of street roads in Montemerlo *et al.* [33]). For low-powered eye-safe laser scanners (like the *URG-04LX* range scanner) the intensity value decays quickly with range (as reported by Kawata *et al.* [36]). Furthermore, for such eye-safe sensor the reflected laser arrives at the sensor's photodiode with low power so it must be processed through filters, correction tables and amplifiers in the sensor's circuitry. Particularly the *URG-04LX*'s auto gain controller (AGC) makes sure that intensity is within usable ranges by the sensor and to avoid saturation.

Furthermore, in the *URG-04LX* sensor, if the amplified intensity of any laser beam is over some internal threshold the sensor will report the range data for that beam, if not it is regarded as *out-of-range* in the sensor output. If we read the beam's intensity output from the sensor it will be the value after filtering and amplification, so the physical intensity and the sensor's intensity output are not linearly related. As such, the reflection intensity values returned by this sensor are not absolute. Therefore we attempt to characterize the actual sensor intensity output, in particular for the effects of range and direction of the laser beams in the intensity value.

4.1 Laser Intensity and Textiles

To consider applications of laser intensity for people detection, it is interesting to study how laser intensity changes with range and the type of textiles in clothes, particularly how black textiles affect the intensity for their lower reflectivity.³

In Fig. 7 we present a selection of 17 types of commonly used textiles in the form of shirts, coats and jackets; including materials like cotton, polyester, wool, in various colors and densities (i.e., thickness, low for fresh/thin clothes and high for warm clothes). For each case, a human wearing the target textile stood in front of the robot at a specific range and scan data was captured during 30 s at each position, and the obtained data comes from the chest part. The plotted values correspond to averages for both range and intensity; data was obtained for range intervals of 500 mm (starting at 1000 mm) up to the maximum range of the sensor (5600 mm). Due to poor reflectance, in the cases of black polyurethane, black synthetic leather,⁴ black cotton low-density and black nylon, data was obtained every 250 mm and up to the out-of-range condition. To compare results, the curve for white paper presented in Fig. 8 is also included here as reference.

³This test is not comprehensive and corresponds to measurements of a particular type of LRF sensor, work in the area of near-infrared (NIR) spectroscopy for textiles exist, analyzing the response of textiles to different wavelengths (see for example Haran *et al.* [37]), but for LRF sensors dynamic changes in wavelength, laser type and photo-diodes are not possible or limited.

⁴Black synthetic leather has the highest standard deviation of the selected textiles in all the tested ranges.

White and light colored textiles, and interestingly some black textiles like a polyester suit, acrylic jacket, microfiber shirt and a black cotton T-shirt (high-density cotton) present higher intensity levels up to the maximum scanning range of the sensor. Blue denim (jacket) present a relative lower reflectance and up to a range of about 4500 mm.⁵ Thick wool/nylon coats show even lower levels and ranges. We also compare clothes of the same material but of different colors: no significative differences were found with the exceptions of low density cotton and synthetic leather.

Intensity varies with colors and materials: some portion of the light that is diffusely reflected from objects conveys some information (in the form of spectrum of reflectance [37]) from those objects. Thus, for some materials intensity information can be used for separation between light and dark and also from low to high reflective objects (like white paper and walls differ from textiles).

4.2 Calibration of Laser Intensity

The behavior of intensity depends, among many factors, on the distance (range) to the target. As shown before in Fig. 7, the Hokuyo URG-04LX presents a characteristic *sigmoid* curve for intensity vs range, where decay is slower at the start and end ranges. Inspired on the *Gompertz* growth function [38, 39], we defined the decay function for intensity as a function of range as:

$$\mathcal{F}_\rho(p_i) = a(1 - \exp(b \exp(c\rho_i))) \quad (7)$$

where ρ_i is the range of point p_i , a is the upper asymptote, b and c (both negative numbers) are decay parameters. Non-linear least squares was used to find the parameters (a, b, c) for each sensor separately. Equation 7 does not represent the exact physics of how the actual reflected laser beam (received at the sensor photo-diode) decays with distance because there is no linear relation from the physical laser intensity and the sensor's intensity output; nevertheless, we regard it as the characteristic intensity response from the sensor.

We present this characteristic form in Fig. 8. As standard procedure, data was obtained from a white target (90 g/m², 0.13 mm thick white bond paper) with an angle of incidence close to zero, from a minimum range of 250 mm up to a maximum range of 5250 mm, in steps of 250 mm; the continuous curve corresponds to the function $\mathcal{F}_\rho(p_i)$ in Eq. 7. These curves and those in Fig. 7 were obtained after warming up the sensors for about 60 min to avoid range and intensity drifting due to changes in internal temperature (as also reported by Pascoal *et al.* [26] and Okubo *et al.* [27]). Scan

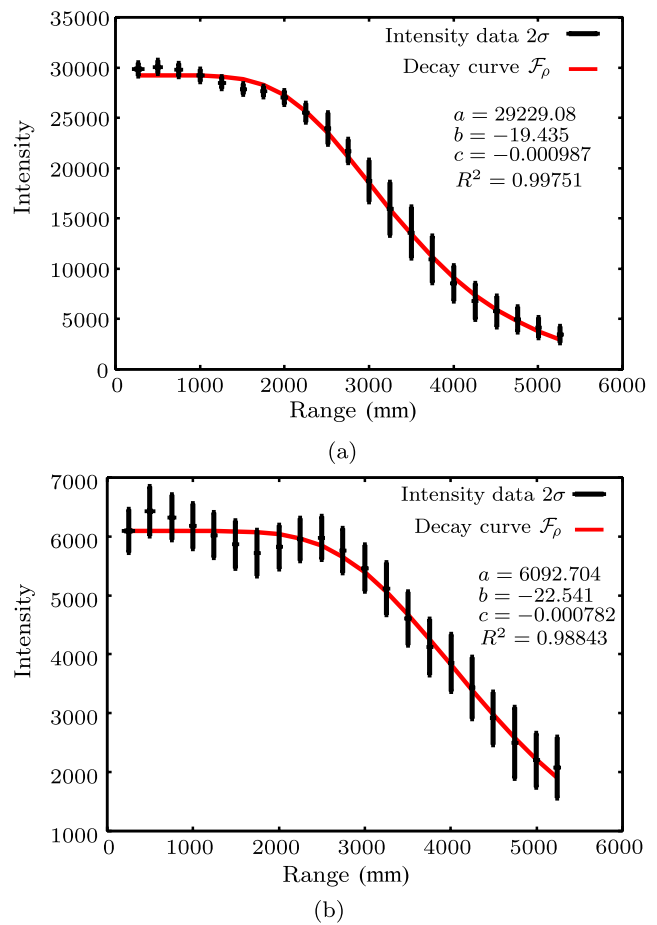


Fig. 8 Characteristic intensity data by varying distance and sigmoid curve $\mathcal{F}_\rho$ for two different URG-04LX sensors, the parameters (a, b, c) of each curve and the correlation coefficient R^2 are also given

data was collected during 60 s for each different range, we also present the intensity error bars for two standard deviations 2σ on the figures.

As depicted in Figs. 8(a) and 8(b) the function $\mathcal{F}_\rho$ changes with the sensor, not only the curves have different parameters but also the scales of intensity values are different (the curve in Fig. 8(a) has a maximum value of 30047 whether that in Fig. 8(b) has a maximum value of 6425⁶). According to Hokuyo personnel, although the model of the sensors and their general operation is the same, they come from two different production lines with different internal calibration parameters in each sensor's circuitry. Thereby, a single calibration curve was not enough to accommodate all the different responses from the sensors. Instead, we obtained the $\mathcal{F}_\rho$ curves for each of the four sensors and calibration of intensity was done for each sensor separately.

Using Eq. 7 as the expected intensity for the range of a given scan point p_i , we normalize (calibrate) the raw inten-

⁵Although trousers did not form part of this study, during tests we observed that blue denim jeans were not detected by the sensor after an average range of 3900 mm.

⁶The maximum intensity values were obtained at an average range of 250 mm from the sensor, we do not consider ranges below this minimum.

sity value reported by the sensor $I(p_i)$ for the effect of range to the target using the decay function $\mathcal{F}_\rho$, for a given scan point as:

$$\mathcal{I}(p_i) = \frac{I(p_i)}{\mathcal{F}_\rho(p_i)} \quad (8)$$

We use the function $\mathcal{F}_\rho$ to normalize the intensity level, that is the maximum expected intensity level for our indoor environments. This normalization level corresponds to that of white paper and, as presented in Fig. 7, reflected intensity levels from textiles are under this maximum. We found that the intensity levels from materials present in common indoor objects (concrete, wood, paper, cardboard, plants, plastics, etc.) are also under this normalization level.

Fusion of sensors in each layer means that if a fused point p_i falls inside an overlapping area (as explained in Sect. 3.1), its intensity value may come from two different sensors a and b , i.e., there are two raw intensity values $I^a(p_i)$ and $I^b(p_i)$. Since each sensor has a different decay function $\mathcal{F}_\rho$, for every fused point p_i we keep their original intensity data before fusion ($I^a(p_i)$ and $I^b(p_i)$), then calibrate the raw intensity value from each sensor to produce $\mathcal{I}^a(p_i)$ and $\mathcal{I}^b(p_i)$, and finally report the average $\mathcal{I}(p_i) = \frac{\mathcal{I}^a(p_i) + \mathcal{I}^b(p_i)}{2}$ as the calibrated intensity for p_i . For points outside overlapping areas calibration of raw intensity is straightforward.

The mean of calibrated intensity $\mu_{\mathcal{I}}$ and the variance of intensity $\sigma_{\mathcal{I}}^2$, after calibration (Eq. 8), serve as measures of an object's intensity. One object (segment) resulting from combining two or more segments with different reflection properties will have a larger variance of intensity values $\sigma_{\mathcal{I}}^2$ resulting from mixing the intensities of the individual objects.

Now, consider how intensity changes between consecutive points p_i^j and p_{i+1}^j in a segment j using the difference of—calibrated—intensity:

$$\mathcal{D}_j(p_i) = \mathcal{I}(p_i^j) - \mathcal{I}(p_{i+1}^j) \quad (9)$$

The variance of intensity differences $\sigma_{\mathcal{D}}^2$ serves as measure the *intensity uniformity*: if the object has uniform reflection (i.e., a smooth surface with a single color) then $\sigma_{\mathcal{D}}^2$ will be rather small, as the differences tend to be small. The *variance of intensity differences* on a segment j is:

$$\sigma_{\mathcal{D}_j}^2 = \frac{1}{N} \sum_{i=1}^N \left(\mathcal{D}_j(i) - \mu_{\mathcal{D}_j} \right)^2 \quad (10)$$

where $\mu_{\mathcal{D}_j}$ is the average of the difference $\mathcal{D}_j$ for all the points in segment j .

4.3 Laser Intensity and People Detection

When two different objects stand very close their individual LRF range segments can become merged into one larger

segment. Additionally some LRF sensors suffer from the *mixed pixels* problem,⁷ that is spurious and noisy scan points between two clearly separated surfaces that are hit by the same laser beam cone or by different pulses (echoes) of the same laser beam at the edge, producing multiple range readings for the same beam (direction); the mixed pixels add continuity to the scan reading, falsely connecting two separated objects and they can get merged during segmentation. Thus the effect of laser reflection intensity for people detection can be better understood in situations where a human cannot be properly distinguished using only range-based features. One example is the case of a person very close to other person or object.

4.3.1 Merged Segments and Laser Intensity

To study the application of laser intensity, we obtained data from three experimental scenarios involving two persons in indoor environments, wearing different colors (in this case dark gray and beige shirts), for the following cases:

- (A) two persons separated
- (B) two persons merged (in contact)
- (C) a person and an object (a wall) merged

Figure 9 presents these three cases: it includes a representation of the humans and the environment for each case, real scan data with actual detection results using our multi-layered system (detected humans are marked here by cylinders) corresponding to each case, and the respective graph of differences of intensity $\mathcal{D}_j$ for the top layer.

Figure 9(a) corresponds to Case A where persons $P0$ and $P1$ are separated and correctly detected. Figure 9(b) corresponds to Case B and here the persons moved closely and their LRF segments became merged (segment 1).⁸ Finally, in Fig. 9(c), corresponding to Case C, they became separated but person $P1$ moved very close to a wall and both $P1$'s and wall's segments became merged (segment 2).

The rightmost column of Fig. 9 shows the graph of differences of calibrated intensity $\mathcal{D}$ for all the segments in the top layer, including their respective variances ($\sigma_{\mathcal{D}}^2$ in Eq. 10), we present other detailed information of these segments in Table 2.

One thing we can observe is that segments corresponding only to walls (segments 0 and 3 in Fig. 9(a), segments 0 and 2 in Fig. 9(b) and segment 0 in Fig. 9(c)) have smaller $\sigma_{\mathcal{D}}^2$ because their reflective property is quite uniform.

⁷Reported by Adams [40], the mixed pixels problem depends on the spot size (beam divergence) of the laser beam. Hancock [30] uses a laser with very small beam divergence, Ye [34] uses thresholding of laser intensity and an EKF for segmentation and breakpoint detection.

⁸For segment separation using intensity we only consider data from top layer, hence we drop the "(t)" and "(b)" terms here.

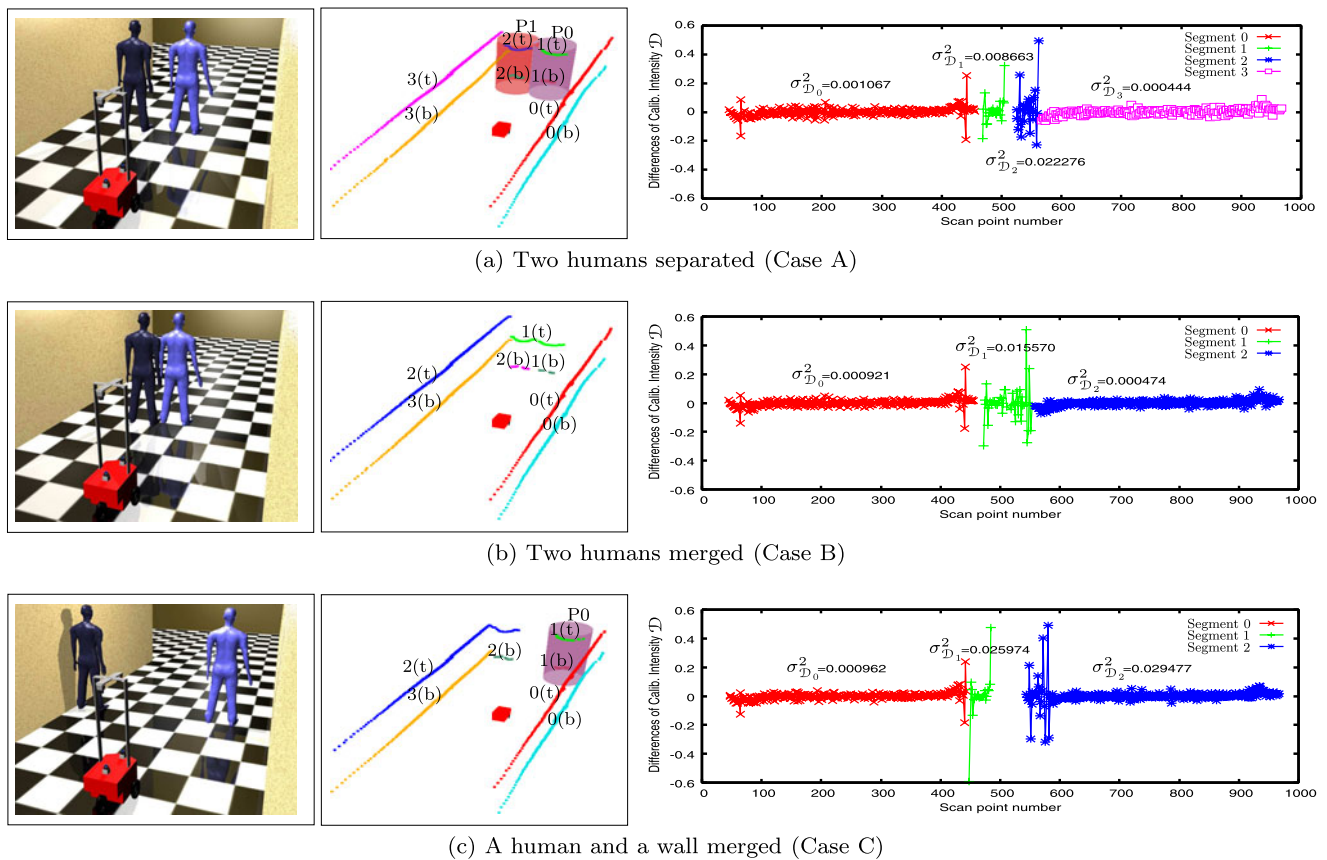


Fig. 9 Three cases for segment separation and laser intensity for two persons, (left) visual representation of the case, (middle) actual scan segments and people detection, (right) difference of intensity curves. (a) segments 1(t) and 2(t) from two persons (P_0 and P_1) are sep-

arated and the persons are detected, (b) two persons very close and forming one segment 1(t), (c) a human very close to a wall forming one large segment 2(t). The terms “(t)” and “(b)” stand for “top” layer and “bottom” layer

Table 2 Segment parameters (top layer)

Case	Seg	μ_D	σ_D^2	ℓ	w
A	0	0.000052	0.001067	121.219	5767.907
	1	0.011512	0.008663	3.789	427.895
	2	0.018666	0.022276	2.667	469.681
	3	0.000736	0.000444	116.857	5513.667
B	0	0.000107	0.000921	111.213	5746.118
	1	−0.005758	0.015570	5.467	922.348
	2	0.000091	0.000474	113.436	5893.558
C	0	0.000085	0.000962	110.988	5073.084
	1	−0.023403	0.025974	2.909	449.623
	2	0.022062	0.029477	8.759	4912.623

In Table 2 we include LRF range and intensity information for each case in Fig. 9 (in this implementation we consider only the top layer). Segments corresponding to people (both separated and merged) are highlighted with bold characters. From range-based features, segments from peo-

ple have a small size ratio ℓ and also small width w value, but for the case of merged segments these range-based features alone are insufficient for people detection. On the other hand, the variance σ_D^2 is large for all the segments from people, even the case of a human merged with a wall (segment

Table 3 Intensity-based features for people detection

No.	Feature $\mathbf{h}(S_i)$
13	Average intensity ($\mu_{\mathcal{I}}$), the average of the calibrated intensity.
14	Intensity variation ($\sigma_{\mathcal{I}}^2$), the variance of calibrated intensity.
15	Average difference of intensity ($\mu_{\mathcal{D}}$), the average of the differences of calibrated intensity.
16	Intensity uniformity ($\sigma_{\mathcal{D}}^2$), the variance of differences of calibrated intensity as in Eq. 10.

2 from Case C). This large variance $\sigma_{\mathcal{D}}^2$ (or intensity uniformity) usually comes from large values of the variance of calibrated intensity $\sigma_{\mathcal{I}}^2$ which provides a measurement the variations of intensity (i.e. diffuse reflection) of an object.

Therefore, intensity uniformity ($\sigma_{\mathcal{D}}^2$) and the intensity variation ($\sigma_{\mathcal{I}}^2$), together with the respective average values $\mu_{\mathcal{D}}$ and $\mu_{\mathcal{I}}$, can be considered as new features for human detection. We introduce these new features and their respective classifiers in Table 3.

These new features cannot be used alone because there can be non-human environment objects with high intensity uniformity values, as it is for cluttered environments, but combined with complementary features, like those in Table 1, they improve detection rates as we will shown in Sect. 6. We include these new features into the training vector $\mathcal{X}$ and apply again AdaBoost to train an augmented strong classifiers $\mathcal{H}_{\Psi}^+(S)$, according to the procedure described in Sect. 3.3, as

$$\mathcal{H}_{\Psi}^+(S) = \begin{cases} \mathcal{H}_t^+(S), & \text{for the top layer} \\ \mathcal{H}_b^+(S), & \text{for the bottom layer} \end{cases} \quad (11)$$

4.3.2 Merged Segments Detection and Separation

We are interested in detecting and separating merged segments where at least one of the segments corresponds to a human. To detect such segments, an intuitive approach is to find wider segments (segments over some $w_{\max}$, a maximum possible width for a human) that are not thin (size ratio ℓ under some $\ell_{\max}$ to reject wide flat wall) with large intensity uniformity ($\sigma_{\mathcal{D}}^2$ over some $\sigma_{\mathcal{D}_{\min}}^2$). $w_{\max}$ and $\ell_{\max}$ were defined as the maximum thresholds θ_w and θ_{ℓ} trained from the strong classifier $\mathcal{H}_t$. $\sigma_{\mathcal{D}_{\min}}^2$ was also obtained from $\mathcal{H}_t$ in a similar way. The remaining problem is how to find the point of separation (breakpoint) that divides the merged segments.

Consider two different objects i and j with differences of intensity $\mathcal{D}_i$ and $\mathcal{D}_j$, both with normal distribution: $\mathcal{D}_i \sim \mathcal{N}(\mu_{\mathcal{D}_i}, \sigma_{\mathcal{D}_i}^2)$ and $\mathcal{D}_j \sim \mathcal{N}(\mu_{\mathcal{D}_j}, \sigma_{\mathcal{D}_j}^2)$, with mean $\mu_{\mathcal{D}}$ and variance $\sigma_{\mathcal{D}}^2$. Consider also that both objects are in contact so that they form one segment of scan data (Cases B and C in Fig. 9, more detail in Fig. 10).

The intensity distribution of the combined objects is the *Gaussian Mixture* $\mathcal{M}$ of both individual intensity distributions:

$$\mathcal{M}(\mathbf{x}) = \omega_i P(\mathbf{x}|\Theta_i) + \omega_j P(\mathbf{x}|\Theta_j) \quad (12)$$

with $\mathbf{x} \in \mathcal{D}_{ij}$ and P the probability density function and the mixture weights ω_i and ω_j . The task of separating merged segments can be considered as the problem of separating the combined intensity distribution $\mathcal{D}_{ij}$ for objects i and j which implies estimating the unknown parameters: $\Theta_i = \langle \mu_{\mathcal{D}_i}, \sigma_{\mathcal{D}_i}^2 \rangle$, $\Theta_j = \langle \mu_{\mathcal{D}_j}, \sigma_{\mathcal{D}_j}^2 \rangle$, with a known parameter $N = 2$ the number of objects that define the mixture.⁹

We use the *Expectation-Maximization* (EM) algorithm to estimate Θ_i and Θ_j and then we separate data into respective distributions $\mathcal{D}_i$ and $\mathcal{D}_j$. After separation into respective distributions, the breakpoint is defined as the average of the last point in $\mathcal{D}_i$ and the first point in $\mathcal{D}_j$.

In Fig. 10 we present how separation of merged segments works for Cases B (left column) and C (right column). To complement the figure and ease visualization, in Figs. 10(a) and 10(d) we include segments' scan data in Cartesian coordinates (the scan point number axis is included here for convenience). The actual break points were computed manually and are indicated in the figures. The curvatures from the two humans (Case B), and also the long wall and to its left the curve of a human (Case C), can be seen in Fig. 10(a) and Fig. 10(d).

In Figs. 10(b) and 10(e) we show a detailed graph of differences of intensity ($\mathcal{D}$) for the merged segments. In each segment, intensity data present different uniformity (vertical distance between points): for Case B (Fig. 10(b)) points on the left half are more tight than those on the right, for Case C (Fig. 10(e)) the first 50 points on the left are more loose than the rest. In Figs. 10(c) and 10(f) we use EM for separation with points divided into respective distributions, the estimated means and covariance ellipses (up to 3σ) and the estimated breakpoints are also shown.

Finally, we formalize the merged segment detection and separation method in Algorithm 1.

5 Multi-Layered People Detection

The purpose of the multi-layered people detection step is to combine candidate segments detected in individual layers increase the confidence on the detection and therefore

⁹We attempt to separate a merged segment in two, applying the separation method on the divided segments is possible.

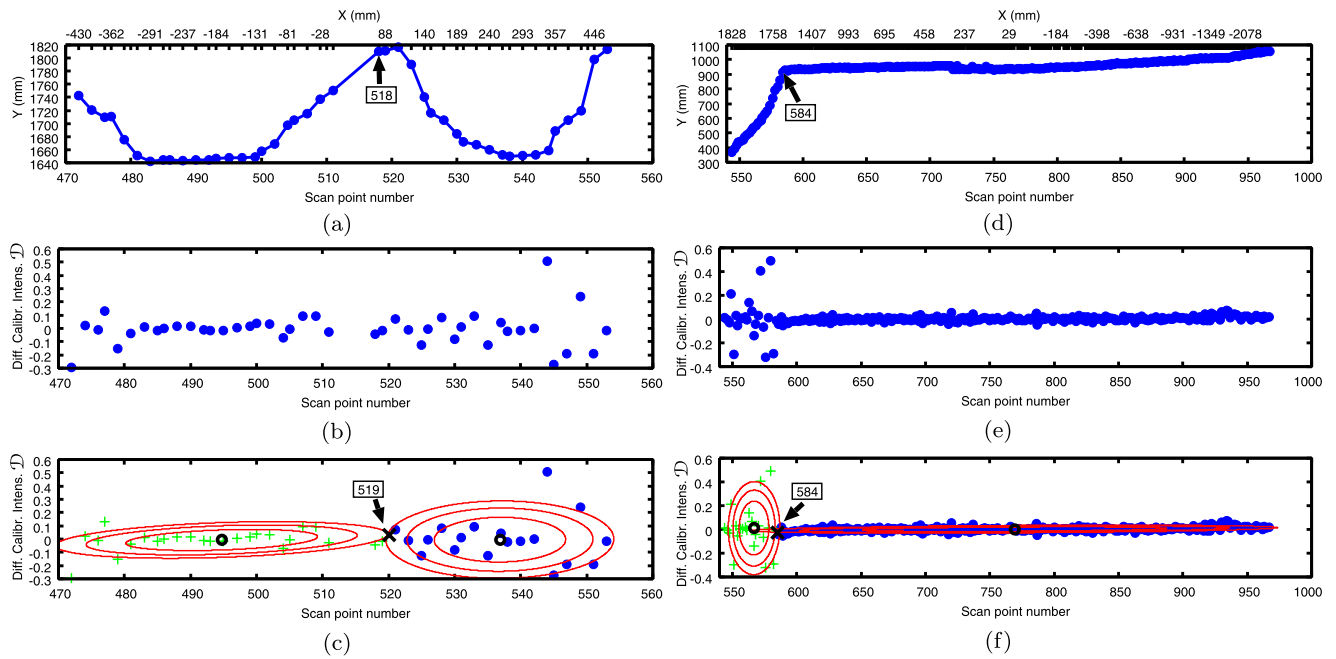


Fig. 10 Segment separation from intensity data for Case B (left column) and Case C (right column): (a) and (d) scan data in Cartesian coordinates with the actual breakpoint indicated, (b) and (e) differences of calibrated intensity D , (c) and (f) Expectation-Maximization

robustness of the system. Additionally, in this step we consider cases where this fusion of layers is not possible (for example when data from one layer is missing), thus our detection system is tolerant to occlusions.

We designed our people detection system under the following assumptions:

1. A human is an average sized adult
2. Data from both layers is necessary to achieve initial detection
3. Humans move at normal walking speeds around the robot

The first two conditions are related, a human is detected if we obtain data from both layers and he/she must be tall enough: be over a height of 110 cm (top layer height). While this limitation is a physical implementation issue which hampers detection of small people, can be eliminated by including additional LRF layers. The second condition limits the response of our system in cases of temporal occlusions: to initially detect a human we require data from both layers, and after that our system is tolerant to occlusions in the bottom layer. The third condition was established to create a simple walking model to fuse individual layer detection results, as will be explained below.

5.1 Fusion of Layers

Fusion of layers consists in finding the corresponding legs for a given chest (see Fig. 11(a) and 11(b)). We project

of D for segment separation with the estimated breakpoint position. “○” marks the estimated mean of each Gaussian and “✕” marks the detected breakpoints, the actual and the detected breakpoints are indicated with an arrow

Algorithm 1 Intensity-based segment separation

```

1: for every non-candidate segment  $S_i$  do
2:   if  $\ell_i < \ell_{\max}$  and  $w_i > w_{\max}$  then /*is not a person*/
3:     if  $\sigma_{D_i}^2 > \sigma_{D_{\min}}^2$  then /*intensity is not uniform*/
4:       Separate  $S_i$  into  $S_{i,1}$  and  $S_{i,2}$  using EM
5:       Feature Extraction for  $S_{i,1}$ ,  $S_{i,2}$ 
6:       Classification for  $S_{i,1}$ ,  $S_{i,2}$ 
7:       Replace  $S_i$  by  $S_{i,1}$  and  $S_{i,2}$  in the list of segments
       /* $S_{i,1}$  and  $S_{i,2}$  are also subject to merged segments check*/
8:   end if
9: end if
10: end for

```

the two layers into one plane and use the distance between bounding-box centers of candidate segments, we set a limit (threshold) to this distance between segments using a searching area which increases the confidence that a chest segment is associated with its corresponding leg(s) segment(s). This fusion method can be applied when more than two layers are available by projecting all candidate segments in one plane and associating them if inside the searching area.

The search area is defined according to the walking step length L , based on Latt *et al.* [41], we selected a normal walking speed of 1.2 ± 0.04 m/s and a corresponding step length of 73.0 ± 3 cm. Consider Fig. 11(a), the average leg

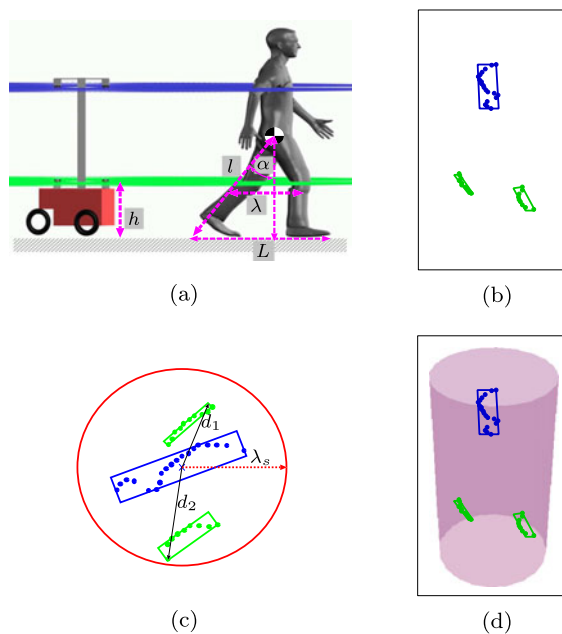


Fig. 11 Walking model and layer fusion: (a) step length L and separation between legs λ at sensor height h , (b) detected features at each layer, (c) search area for association, (d) a detected human represented by a cylinder

length l is about 84 cm, and the height of the bottom layer of sensors h is fixed at 40 cm. L is the step length which depends on the speed, for example 73 cm for an average speed of 1.2 m/s. The separation between legs¹⁰ λ at h is calculated as:

$$\lambda = 2(l \cos(\alpha) - h) \tan(\alpha), \quad \text{where } \alpha = \sin^{-1}\left(\frac{L}{2l}\right) \quad (13)$$

We can define a search radius of $\lambda_s = \frac{\lambda}{2} + \xi$, for some bias ξ , at the center of the chest bounding-box, projected into the bottom layer (bird's-eye view in Fig. 11(c)) to search for the corresponding legs for the chest. Here we compare the distance from the chest center to the furthest corner of the leg bounding-box (d_1 and d_2 in Fig. 11(c)) with the radius λ_s . If one or two leg segments are found connected to the chest then we successfully detected a human (a multi-layered detection), as in Fig. 11(d). By using furthest corners we avoid segments that are only partially inside the search area. In this implementation, we selected $\lambda = 35.42$ cm for normal walking speed and $\xi = 12$ cm.

5.2 Position Estimation and Data Association

For every detected person $\mathcal{P}_i$ we choose the center of the chest segment's bounding-box as $\mathcal{P}_i$'s expected position

¹⁰ Although Fig. 11 shows a normal walk pose, we assume that both left and right legs form symmetric gait angles (i.e., both equal to α).

$\mu_i = [x_i \ y_i]^T$ and the radius λ_s to define the initial covariance of the actual position $\Sigma_i = \lambda_s I$. This position covariance is defined for walking humans, our main target, so it will be rather large for the case of persons standing still or moving slowly. A better estimation will be to keep track of the his/her position through time; this is left as a future work.

After successful detection of a person in current scan data, we associate it with data obtained from the previous scan (history). This is, a *newly detected person* $\hat{\mathcal{P}}_i^t$ in time t is associated with a already seen person from a list $\mathcal{P}^{t-1}$ detected in time $t - 1$. From this association we update the position information of already known people around the robot and compute speed and direction information for every person. To achieve this data association, for every new person $\hat{\mathcal{P}}_i^t$ we find the closest $\mathcal{P}_k^{t-1} \in \mathcal{P}^{t-1}$ using the Mahalanobis distance $d(i, k)$ from $\hat{\mathcal{P}}_i^t$ estimated position and $\mathcal{P}_k^{t-1}$ predicted position for time t from the computed speed and direction.

If the minimum Mahalanobis distance is smaller than some threshold then $\hat{\mathcal{P}}_i^t$ and $\mathcal{P}_k^{t-1}$ are associated in the new list $\mathcal{P}^t$ of detected people around the robot. Otherwise $\hat{\mathcal{P}}_i^t$ is regarded as a new person and integrated in $\mathcal{P}^t$. Successful association means, among other housekeeping tasks, updating the position, speed and direction of the known person and preserving his/her identification k from $\mathcal{P}_k^{t-1}$, so that $\hat{\mathcal{P}}_i^t$ becomes $\mathcal{P}_k^t$ in $\mathcal{P}^t$; failed association means that $\hat{\mathcal{P}}_i^t$ receives some new identification. We also remove entries in list $\mathcal{P}^t$ if not associated during some defined number of scans.

5.3 Occlusion Tolerance

By relaxing this association process we allow *partially detected* persons to become associated to known persons (partial association). A partially detected person consists on just one candidate segment (chest or leg) in one layer for which no corresponding segment(s) could be found in the other layer. This no correspondence can be due to misclassification (false negatives) or due to occlusion by some other object (it falls outside of the search radius λ_s in the fusion process). In the current implementation we allow partial association only for chest segments because occlusions frequently occur in the bottom layer. The same approach can be used for a setup with more than two layers, where a candidate segment in one layer has no corresponding segments in the rest of the layers.

Although this association method is simple and allows occlusion tolerance, we are now working on the implementation of people tracking using Kalman filter and the social forces model (see Luber *et al.* [42]).

6 Experimental Results

We evaluate our detection system under several different scenarios and conditions, we present the experiments and detection results in this section.

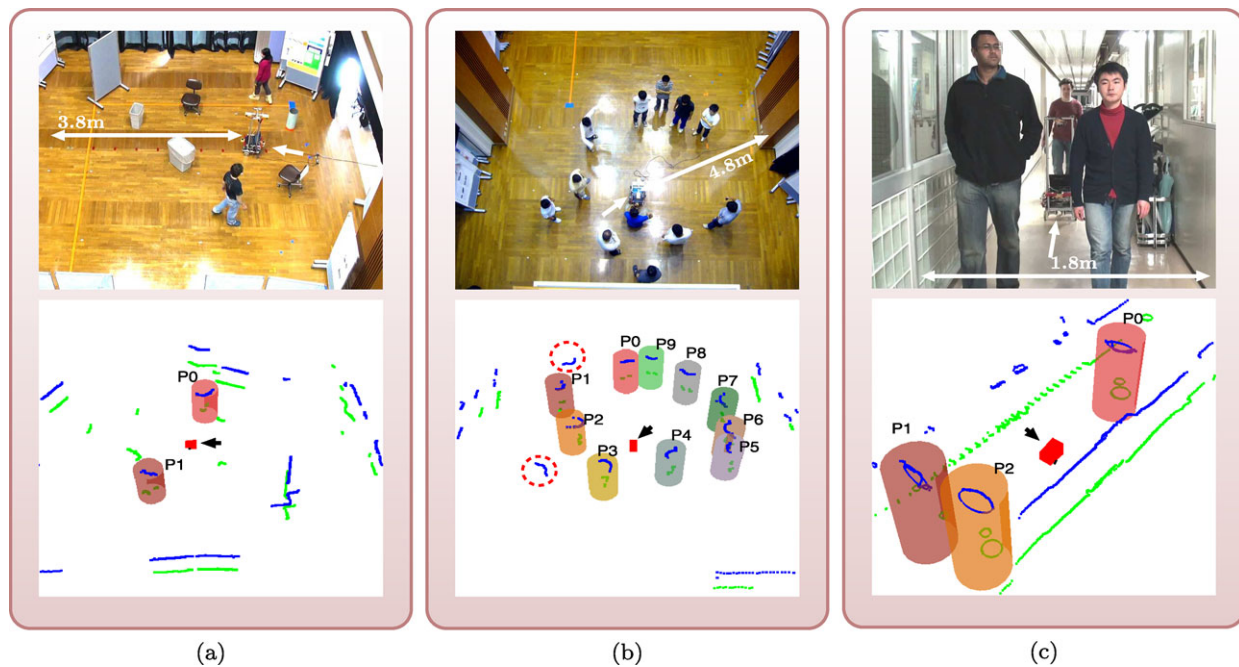


Fig. 12 Scenarios and detection results of the different experiments: “Clutter” test (a) detected persons marked with P_0 and P_1 ; “Crowd” test (b) 10 out of 12 persons detected (P_0 to P_9 , missing persons

marked with circles); “Hallway” test (c) detected people are marked as P_0 , P_1 and P_2

Scan data is read in parallel from each of the four sensors and timestamped using the computer’s clock. Data from each layer is processed in parallel and then integrated in the multi-layer fusion process. The processing time from sensor fusion (single layer processing) to people detection (the last step in multi-layer processing) was in average 52 ms, fast enough given the sensor’s scanning speed of 100 ms.

Normal range data from the *URG-04LX* consists of 682 points circularly ordered from right to left and with an angular resolution of 0.36° . However, to simultaneously obtain range and intensity, the sensor decreases the number of scan points to one half (341 points) and the angular resolution becomes 0.72° . Before obtaining experimental data, we warmed up each of the *URG-04LX*s for about 60 min to avoid range and intensity measurements drifting with temperature.

6.1 Experiment Scenarios

We obtained experimental data from four different environments and conditions:

- Wide area with clutter (“Clutter”)
- Several people around the robot (“Crowd”)
- Long and narrow hallways and the robot moving with people (“Hallway”)
- A cluttered office area (“Office”)

We present pictures of the different environments in the top row of Figs. 12 and 13, in each case some information about

the size of the area for detection is provided and the robot position is marked with an arrow. The “Clutter” environment (Fig. 12(a)) is a wide area with little cluttering (chairs, dust bins, panels) causing occlusions in the bottom layer. Here two humans moved around the robot, moving in front and behind the obstacles. The “Crowd” test (Fig. 12(b)) is a wide environment with no cluttering but includes several people—twelve—of diverse sizes and clothes colors moving around the robot, total occlusions occur frequently and segments merging. Finally the “Hallway” environment (Fig. 12(c)) consists of long and narrow passages with cluttering (shoe-racks, umbrella stands, etc.), where the robot is operated by remote control and moved inside a group of three humans. This case presented also occlusions and frequent merging of segments between people and walls, also when turning a corner people would disappear temporarily. Although our companion robot is intended for open areas with reduced cluttering, as a challenge for our detection system we include the additional “Office” test environment: a research office with desks, book shelves and heavy cluttering at both top and bottom layers (Fig. 13).

In the bottom row of Figs. 12 and 13 we include actual scan data with people detection results. In every figure detected persons are labeled as P_i and marked with a cylinder, where darker segments correspond to top layer data. In all the cases, the robot is indicated with an arrow about the center of the scan data figures.

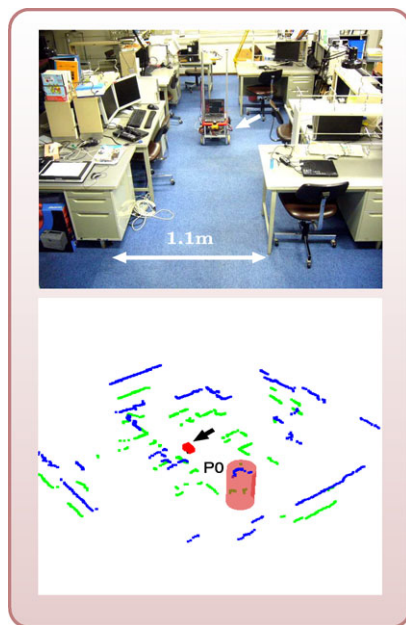


Fig. 13 Scenario and detection result for the “Office” test, the detected person is marked as $P0$

Scan segments were manually labeled into person or object to establish the ground truth, using the method presented in Sect. 3.3. To train the strong classifiers $\mathcal{H}_t$ and $\mathcal{H}_b$, we combined the labeled data from each of the experimental environments and extracted 2183 segments corresponding to positive examples and 9635 segments from negative examples. The same strong classifiers ($\mathcal{H}_t$ and $\mathcal{H}_b$) were used in all the evaluations, we did not train and use a pair $\langle \mathcal{H}_{t_i}, \mathcal{H}_{b_i} \rangle$ of strong classifiers specific for the i -th environment.

Due to our multi-layered configuration, the detection results for the experiments are not based on per-segment classification rates but for each multi-layered scan. A person can be detected even if segments from one layer are missing or misclassified: if the detected person corresponds to the ground truth we count it as a hit for the current scan, if missing in the detection results we count a false negative, and a multi-layered observation of environment objects misclassified as a person is counted as a false positive.

6.2 Range-Only Multi-Layered People Detection

We first evaluate the multi-layered detection rates considering only range data (i.e., without laser intensity features). Therefore no segment separation is performed here. Detection rates for every case are summarized in Table 4.

The “Clutter” test consisted of 473 multi-layer observations (i.e., independent and parallel scans) and 18172 segments, the true detection rate was 89.2%, false positive rate was 7.5% mostly due to misclassification of poster panels and curtains as people, and a false negative rate of 3.3% due

Table 4 Multi-layer detection rates (without laser intensity)

Test	True detection	False positive	False negative
Clutter	89.2%	7.5%	3.3%
Crowd	85.4%	3.0%	11.6%
Hallway	77.4%	9.4%	13.2%
Office	83.3%	12.5%	4.2%

to people walking behind panels and not being detected immediately after reappearing. Of the 473 observations, 166 presented partial occlusion cases (people walking behind clutter), and of those our system detected the humans in 151 observations.

The “Crowd” test consisted of 164 observations (6351 segments) for a true detection rate of 85.4%, complete occlusions (people walking behind people and very close) is largely responsible for the 11.6% false negative rate.

The “Hallway” test had 1238 observations (27662 segments) for a detection rate of 77.4%. Several environment objects, particularly long glass walls with steel frames and glass bricks (see Fig. 12(c) present in about 40% of the route, the small segments in the lower layer are glass bricks), were misclassified as people (9.4% false positive rate), and people walking close to the walls and disappearing behind corners largely accounts for the 13.2% false negative rate.

The “Office” test consisted of 368 multi-layer observations and a total of 23003 segments. The person moved behind some desks and was partially occluded during 36 observations, in 31 observations of those our system successfully detected the person. The true detection rate was 83.3%, false positive rate was 12.5% due to misclassification of books in top layer and chairs in the bottom layer, and a false negative rate of 4.2% due a person walking behind desks and not being detected immediately after reappearing.

6.3 Intensity-Based Multi-Layered People Detection

We repeated the detection evaluation using same scan data from the previous experiment environments, but this time using intensity-based features (the strong classifier $\mathcal{H}_\psi^+$ in Eq. 11) and segment separation. Detection results for every case are summarized in Table 5, all showing an increase in the true detection rates. In the “Clutter” test the number of false negatives did not change much due to the same problem explained above, and the true detection rate increased only 3.4%. In the “Office” test the number of false negatives shows no change but the true detection rate increased 7.8%. The true detection rate for the “Crowd” test increased 5.3% and 10.8% for the “Hallway” test, both with significant decreases in the false negative rates.

Table 5 Multi-layer detection rates (including laser intensity)

Test	True detection	False positive	False negative
Clutter	92.6%	4.2%	3.2%
Crowd	90.7%	2.6%	6.7%
Hallway	88.2%	6.7%	5.1%
Office	91.1%	4.7%	4.2%

Table 6 Single-layer detection rates (with and without laser intensity)

Test	True detection	False positive	False negative
Without			
Clutter	67.3%	21.2%	11.5%
Crowd	76.7%	4.3%	19.0%
Hallway	60.1%	19.7%	20.2%
Office	55.9%	34.3%	9.8%
With			
Clutter	75.5%	18.5%	6.0%
Crowd	86.3%	4.5%	9.2%
Hallway	70.3%	17.0%	12.7%
Office	62.1%	29.6%	8.3%

Although the new intensity-based features contributed to increase the detection rates, segment separation also contributed in reducing the false negative rates (misdetections). In fact we repeated the “Crowd” and the “Hallway” experiments without using segment separation. When compared to the range-only classifier results, the true detection rate increase was minimal (1.5% for the “Crowd” test and 3.2% for the “Hallway” test) since the false negative rates did not change.

6.4 Single-Layered People Detection

To compare results of our multi-layered people detection system (with and without laser intensity), we include the detection results of a single-layered detection system for the four test environments. Using log data from the same experiments, we selected all candidate segments detected by top layer strong classifiers $\mathcal{H}_t$ and $\mathcal{H}_t^+$ as valid persons. In Table 6 we report the single-layered detection results, with and without using laser intensity. As presented, the true detection rates for all the experiments are by far lower than those of the multi-layered system, although laser intensity allows to improve the detection results for this single-layer system.

6.5 Simplified Intensity-Based Classifier

The strong classifier $\mathcal{H}_\Psi^+$ can be considered as primary range-based plus a some intensity-based features. In this

Table 7 Multi-layer detection rates (using $\mathcal{H}_\Psi^-$)

Test	True detection	False positive	False negative
Clutter	90.1%	5.3%	4.6%
Crowd	88.7%	4.2%	7.1%
Hallway	82.3%	11.6%	6.1%
Office	88.7%	6.5%	4.8%

Table 8 Top 5 features per classifier and layer

Range-only $\mathcal{H}_\Psi$	Intensity-based $\mathcal{H}_\Psi^+$	Simplified $\mathcal{H}_\Psi^-$
$\Psi = \text{top}$		
$\widehat{N}$	$\widehat{N}$	w
ℓ	ℓ	ℓ
w	$\sigma_{\mathcal{I}}^2$	$\widehat{N}$
$\bar{\kappa}$	w	$\sigma_{\mathcal{I}}^2$
$\bar{\phi}$	$\bar{\kappa}$	$\mu_{\mathcal{I}}$
$\Psi = \text{bottom}$		
$\widehat{N}$	$\widehat{N}$	$\widehat{N}$
R	R	w
w	w	ℓ
ε	$\sigma_{\mathcal{I}}^2$	$\sigma_{\mathcal{I}}^2$
$\bar{\phi}$	ε	$\mu_{\mathcal{I}}$

experiment we want to verify whether fewer but well-selected range-based features can be used together with our intensity-based features for accurate people detection. Thus we trained a new and simpler strong classifier $\mathcal{H}_\Psi^-$ with only three range-based features (width w , ratio ℓ and normalized number of points $\widehat{N}$ from Table 1) and all the four intensity-based features in Table 3.

Detection results for every case are summarized in Table 7. Compared to the range-only results, the true detection rates show some improvement, in particular the “Hallway” and the “Office” tests which increased of 5.2% and 4.9%. In the “Hallway” test there is a small increase of 2.2% in the false positive rate.

6.6 Best Features per Classifier

Regarding the different strong classifier trained, in Table 8 we present the top 5 features in each of the trained strong classifiers and separated by layer. The selection is based on the feature contribution to classification, measured as the unsigned sum of the feature f weights across all weak classifiers, this is $\sum_{i=1}^T |h_i=f| \omega_i$ (each strong classifier has $T = 100$ weak classifiers so f may appear more than once). Our normalized number of points feature $\widehat{N}$ is present in all the strong classifiers and is dominant in all but one case. In

the case of the simplified classifier $\mathcal{H}_{\psi}^{-}$, the range features $(w, \ell, \hat{N})$ have larger weights and appear in a larger number of weak classifiers. This result is not surprising since these features were selected to define $\mathcal{H}_{\psi}^{-}$ for their relative importance, as shown for the other strong classifiers.

7 Discussion

We confirmed the positive effect of using multiple layers of LRF sensors for human detection. The dimension of the problem is slightly higher than from a single sensor but much smaller when compared with vision and 3D scanners. The arrangement of four sensors in two layers is simple enough to achieve fast processing time (below the sensor's scanning time) and allows high detection rates.

We introduced in this work the use of laser reflection intensity for segment separation and as a new feature detection. Since intensity value decays quickly with range in small low power LRF, a calibration method was required. After calibration of intensity data we hoped to achieve separation by color/material/reflectivity of the segments and use this for identification; however, in its current form is not possible. A better characterization of laser intensity response could be achieved by considering other properties like the angle of incidence of the laser, the physics of the *URG-04LX* photodiodes and the type of materials typical of indoor environments. Then AdaBoost could be applied for learning of object and places according to their laser intensity response.

In this work we used AdaBoost to train strong classifiers $\mathcal{H}_{\psi}$ using range and intensity features, and training data from diverse scenarios. This does not produce an “all-seasons” detector because training a supervised classifier depends on the training sets. Actual environments show varied conditions, there are areas with heavy cluttering, narrow hallways and also wide spaces. Selection of the strong classifier according to the location could be also considered using navigation maps and select the best pair $\langle \mathcal{H}_{t_i}, \mathcal{H}_{b_i} \rangle$ for the environment i according to the current position.

One limitation in the current implementation is that, to get simultaneously range and intensity from the sensors we loose on angular resolution, meaning a reduction of 50% the number of scan points. This means that, regardless of the diffuse reflection properties, small objects (e.g., legs) are not detected by the laser sensor after some range. Some solutions we are considering include *interleaving* range and intensity data, i.e., in the first scan we get range-only data, then intensity-only data in the second scan, and so on. We have already tested this idea on the “Clutter” environment with similar detection results. The drawbacks are the slower scanning time of 200 ms and naturally a slower processing time (about 90 ms in average) because the number of points doubles. Other solutions include of course using different and possibly more costly LRF sensors.

The use of laser intensity and data from low power LRF is limited by the amount of diffuse reflection caused by objects textiles. LRF data from some materials (black nylon, synthetic leather, denim, etc.) cannot be read but inside very short ranges (less than 2.5 m), thus detection system using such low power LRF will have lower accuracy for medium to long ranges.

Although Algorithm 1 was defined for separation of merged segments enclosing a human, we believe that proposed separation using the intensity uniformity criteria and the break point detection (lines 3 and 4) can be applied to other contexts, for example using wall color differences to define localization landmarks, by changing lines 1 and 2 accordingly. In the experiments we did not observe the case of more than two merged segments (i.e., three or more people side by side), this is yet to be tested.

8 Conclusions and Future Works

People detection and position estimation multiple layers of laser range finders was covered in this work. The fusion of candidate segments from each layer increase detection robustness while allowing to estimate each person's position. By relaxing the layers fusion step, we deal with cases of partial occlusion. Although research in the area of detection and tracking exist, the proposed approach is simple, fast, and accurate enough for applications on a mobile robot.

The use of laser reflection intensity data for human detection was introduced in this work. We proposed some intensity-based features, included as part of strong classifiers, to improve segment classification rates. Additionally, laser intensity was proposed to detect and separated merged segments which otherwise will be regarded as false negatives. Since laser intensity is affected by the range to the target, we introduced an intensity calibration method.

As future works, dynamic selection of classifiers based on environment awareness will be considered. The possibility of person identification using laser reflection intensity will also be studied. Regarding our main research problem of a robot as member of a people group, we are working on people and group tracking and plan to develop social forces models for group tracking including group bonding forces. Other future steps of this research include navigation adapting to changes in group formation and assignment of roles inside the group.

References

1. Burgard W, Cremers AB, Fox D, Hähnel D, Lakemeyer G, Schulz D, Steiner W, Thrun S (1998) The interactive museum tourguide robot. In: Fifteenth national conference on artificial intelligence, Madison, WI, July 1998, pp 11–18

2. Thrun S, Bennewitz M, Burgard W, Cremers AB, Dellaert F, Fox D, Hähnle D, Rosenberg CR, Roy N, Schulte J, Schulz D (1999) Minerva: a tour-guide robot that learns. In: 23rd annual German conference on artificial intelligence: advances in artificial intelligence, Pittsburgh, PA, September 1999, pp 14–26
3. Bauer A, Klasing K, Lidoris G, Mühlbauer Q, Rohrmüller F, Sosnowski S, Xu T, Kühnlen K, Wollherr D, Bus M (2009) The autonomous city explorer: towards natural human-robot interaction in urban environments. *Int J Soc Robot* 1(2):127–140
4. Fod A, Howard A, Mataric MJ (2002) Laser-based people tracking. In: IEEE international conference on robotics and automation (ICRA), Washington, DC, May 2002, pp 3024–3029
5. Montemerlo M, Thrun S, Whittaker W (2002) Conditional particle filters for simultaneous mobile robot localization and people-tracking. In: IEEE international conference on robotics and automation (ICRA), Washington, DC, May 2002, pp 695–701
6. Xavier J, Pacheco M, Castro D, Ruano A, Nunes U (2005) 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, April 2005, pp 3941–3946
7. Cui J, Zha H, Zhao H, Shibasaki R (2006) Robust tracking of multiple people in crowds using laser range scanners. In: IEEE 18th international conference on pattern recognition (ICPR), Hong Kong, China, August 2006, pp 857–860
8. Lee JH, Tsubouchi T, Yamamoto K, Egawa S (2006) People tracking using a robot in motion with laser range finder. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Beijing, China, October 2006, pp 2936–2942
9. Zhao H, Chen Y, Shao X, Katabira K, Shibasaki R (2007) 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, April 2007, pp 4739–4745
10. Arras KO, Martínez Mozos Ó, Burgard W (2007) Using boosted features for the detection of people in 2d range data. In: IEEE international conference on robotics and automation (ICRA), Roma, Italy, April 2007, pp 3402–3407
11. Schapire RE, Singer Y (1999) Improved boosting algorithms using confidence-rated predictions. *Mach Learn* 37(3):297–336
12. Hashimoto M, Takahashi K, Matsui Y (2007) Moving-object tracking with multi-laser range sensors for mobile robot navigation. In: IEEE international conference on robotics and biomimetics (ROBIO), Sanya, China, December 2007, pp 399–404
13. Gidel S, Checchin P, Blanc C, Chateau T, Trassoudaine L (2008) Pedestrian detection method using a multilayer laserscanner: Application in urban environment. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Nice, France, September 2008, pp 173–178
14. Martínez Mozos Ó, Kurazume R, Hasegawa T (2010) Multi-layer people detection using 2d range data. *Int J Soc Robot* 2(1):31–40
15. Borges GA, Aldon M-J (2004) Line extraction in 2d range images for mobile robotics. *J Intell Robot Syst* 40:267–297
16. Premebida C, Nunes U (2005) Segmentation and geometric primitives extraction from 2d laser range data for mobile robot applications. In: 5th national festival of robotics, scientific meeting (ROBOTICA), Coimbra, Portugal, April 2005, pp 17–25
17. Nguyen V, Gächter S, Martinelli A, Tomatis N, Siegwart R (2007) A comparison of line extraction algorithms using 2d range data for indoor mobile robotics. *Auton Robots* 23:97–111
18. Mucientes M, Burgard W (2006) Multiple hypothesis tracking of clusters of people. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Beijing, China, October 2006, pp 692–697
19. Arras KO, Grzonka S, Luber M, Burgard W (2008) 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, May 2008, pp 1710–1715
20. Kondaxakis P, Kasderidis S, Trahanias PE (2008) 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, September 2008, pp 3370–3377
21. Luber M, Tipaldi GD, Arras KO (2009) Spatially grounded multi-hypothesis tracking of people. In: Arras KO, Martínez Mozos Ó (eds) IEEE international conference on robotics and automation (ICRA) workshop on people detection and tracking, Kobe, Japan, May 2009, pp 73–80
22. Zivkovic Z, Kröse B (2007) Part based people detection using 2d range data and images. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), San Diego, CA, October 2007, pp 214–219
23. Carballo A, Ohya A, Yuta S (2008) 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, August 2008, pp 677–682
24. Carballo A, Ohya A, Yuta S (2009) Multiple people detection from a mobile robot using double layered laser range finders. In: Arras KO, Martínez Mozos Ó (eds) IEEE international conference on robotics and automation (ICRA) workshop on people detection and Tracking, Kobe, Japan, May 2009, pp 94–100
25. Kawata H, Ohya A, Yuta S, Santosh W, Mori T (2005) Development of ultra-small lightweight optical range sensor system. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Edmonton, Canada, August 2005, pp 1078–1083
26. Pascoal J, Marques L, de Almeida AT (2008) Assessment of laser range finders in risky environments. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Nice, France, September 2008, pp 3533–3538
27. Okubo Y, Ye C, Borenstein J (2009) Characterization of the Hokuyo urg-04lx laser rangefinder for mobile robot obstacle negotiation. In: SPIE conference on defense, security + sensing, unmanned systems technology XI: unmanned robotic and layered systems, vol 7332, Florida, USA, April 2009
28. Dietmayer KC, Sparbert J, Streller D (2001) Model based classification and object tracking in traffic scenes from range-images. In: IEEE intelligent vehicles symposium (IVS), Tokyo, Japan, May 2001
29. Klasing K, Wollherr D, Buss M (2009) Realtime segmentation of range data using continuous nearest neighbors. In: IEEE international conference on robotics and automation (ICRA), Kobe, Japan, May 2009, pp 4287–4292
30. Hancock J (1999) Laser intensity-based obstacle detection and tracking. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, January 1999
31. Surmann H, Nüchter A, Hertzberg J (2003) An autonomous mobile robot with a 3d laser range finder for 3d exploration and digitalization of indoor environments. *Robot Auton Syst* 45(3-4):181–198
32. Nüchter A, Surmann H, Hertzberg J (2005) Automatic classification of objects in 3d laser range scans. In: 8th conference on intelligent autonomous systems (IAS), Amsterdam, Netherlands, March 2005, pp 963–970
33. Montemerlo M, Becker J, Bhat S, Dahlkamp H, Dolgov D, Ettinger S, Hähnle D, Hilden T, Hoffmann G, Huhnke B, Johnston D, Klumpp S, Langer D, Levandowski A, Levinson J, Marcil J, Orenstein D, Paefgen J, Penny I, Petrovskaya A, Pflueger M, Stanek G, Stavens D, Vogt A, Thrun S (2008) Junior: the Stanford entry in the urban challenge. *J Field Robot* 25(9):569–597
34. Ye C (2008) Mixed pixels removal of a laser rangefinder for mobile robot 3-d terrain mapping. In: IEEE international conference on information and automation (ICIA), Hunan, China, June 2008, pp. 1153–1158

35. Rusu RB, Meeussen W, Chitta S, Beetz M (2009) Laser-based perception for door and handle identification. In: IEEE international conference on advanced robotics (ICAR), Munich, Germany, June 2009, pp 1–8
36. Kawata H, Miyachi K, Hara Y, Ohya A, Yuta S (2008) A method for estimation of lightness of objects with intensity data from Sokuiki sensor. In: IEEE international conference on multisensor fusion and integration for intelligent systems (MFI), Seoul, Korea, August 2008, pp 661–664
37. Haran T (2008) Short-wave infrared diffuse reflectance of textile materials. Master's thesis, Physics and Astronomy, Georgia State University, Atlanta, GA, August 2008
38. Franses PH (1994) Fitting a Gompertz curve. *J Oper Res Soc* 45(1):109–113
39. Jukić D, Kralik G, Scitovski R (2004) Least-squares fitting Gompertz curve. *J Comput Appl Math* 169:359–375
40. Adams MD (1993) Amplitude modulated optical range data analysis in mobile robotics. In: IEEE international conference on robotics and automation (ICRA), Georgia, USA, May 1993, vol 2, pp 8–13
41. Latt MD, Menz HB, Fung VS, Lord SR (2008) Walking speed, cadence and step length are selected to optimize the stability of head and pelvis acceleration. *Exp Brain Res* 184:201–209
42. Luber M, Stork JA, Tipaldi GD, Arras KO, (2010) People tracking with human motion predictions from social forces. In: IEEE international conference on robotics and automation (ICRA), Alaska, USA, May 2010, pp 464–469

Alexander Carballo received his Bachelor degree in Computer Engineering from Costa Rica Institute of Technology in 1996, where he worked as lecturer for the undergraduate program of the Department of

Computer Science since 1996. In April 2006 he joined the Intelligent Robot Laboratory at University of Tsukuba as research student, and he is currently a doctoral candidate in Computer Science. His research interests include machine learning, autonomous robot navigation, robotic perception and sensor fusion.

Akihisa Ohya received his Doctoral degree in Engineering from Keio University in 1992. He is an Associate Professor of the Graduate School of Systems and Information Engineering at the University of Tsukuba, where he has been a faculty member since 1992. His research interests includes autonomous navigation of mobile robots and its applications to human daily life environments.

Shin'ichi Yuta received his Bachelor, Master and Doctoral degrees in Electrical Engineering from Keio University, in 1970, 1972, and 1975. From 1975 to 1978, he was a Research Associate at the Tokyo University of Agriculture and Technology in department of Electronics, and from 1978 he joined the University of Tsukuba, and was promoted to full professor in 1992. From 2000 to 2002 he was dean of the College of Engineering Systems, from 2002 to 2004 he was chairman of the Institute of Information Science and Electronics, and from 2004 to 2006 Professor Yuta was appointed vice-president of University of Tsukuba. From April 2006 he is a professor of the Graduate School of Systems and Information Engineering, and was also appointed to be a director of the Industrial Liaison and Cooperative Research Center. He is an expert in robotics, and heading the Intelligent Robot Laboratory at the University of Tsukuba. He has been conducting autonomous mobile robot research projects since 1980. He has published more than 500 technical papers in this field. He is a fellow member of the IEEE (Institute of Electrical and Electronic Engineers), RSJ (Robotics Society of Japan) and other academic societies, where he has received several awards.