# Person-Following Algorithm Based on Laser Range Finder and Monocular Camera Data Fusion for a Wheeled Autonomous Mobile Robot

Elvira Chebotareva[1(✉)], Ramil Safin[1], Kuo-Hsien Hsia[2], Alexander Carballo[3], and Evgeni Magid[1]

[1] Laboratory of Intelligent Robotic Systems (LIRS), Intelligent Robotics Department, Higher Institute for Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russian Federation
elvira.chebotareva@kpfu.ru,safin.ramil@it.kfu.ru
https://kpfu.ru/erobotics

[2] Department of Electrical Engineering, National Yunlin University of Science and Technology, Douliu, Taiwan
khhsia@yuntech.edu.tw

[3] Institute of Innovation for Future Society, Nagoya University, Nagoya, Japan
alexander@g.sp.m.is.nagoya-u.ac.jp

**Abstract.** Reliable human following is one of the key capabilities of service and personal assisting robots. This paper presents a novel person tracking and following approach for autonomous mobile robots that are equipped with a 2D laser rangefinder (LRF) and a monocular camera. The proposed method does not impose restrictions on a person's clothes, does not require a head or an upper body to be within a camera field of view and is suitable for low height indoor robots as well. The algorithm is based on a metric that takes into an account parameters obtained directly from LRF and monocular camera data. The algorithm was implemented and tested in the Gazebo simulator. Next, it was integrated into a control system of the TIAGo Base mobile robot and successfully validated in university environment experiments with real people. In addition, this paper proposes a new criterion of algorithm performance estimation, which is a function of false positives number and traveled distances by a person and by a robot. Further this criterion is used to compare performance of the proposed method with the Multiple Instance Learning (MIL) tracker in simulated and in real world environments.

**Keywords:** Mobile robot · Human tracking · Human following algorithm · Laser range finder · Monocular camera · Multisensor tracking · Accuracy score · ROS · Gazebo

## 1 Introduction

Person-following algorithms are used by various types of service robots, including robotic suitcases, cargo robots, autonomous wheelchairs [33], shopping robots

© Springer Nature Switzerland AG 2020
A. Ronzhin et al. (Eds.): ICR 2020, LNAI 12336, pp. 21–33, 2020.
https://doi.org/10.1007/978-3-030-60337-3_3

[14], rescue robots [26,28], and others [31]. A person-following robot simultaneously deals with three problems: a self-localization, followed person coordinates detection (and tracking) and path planning [21]. A significant variety of works on person-following algorithms for autonomous mobile robots, which were proposed in the past decades, could be roughly divided into laser range finder (LRF) based approaches, camera-based approaches and mixed approaches [13].

Mobile robots are often equipped with a 2D LRF. While these on-board LRFs are generally employed for simultaneous localization and mapping (SLAM) tasks [34], they could be also utilized for tracking of a target person's legs position. Major LRF-based human tracking solutions employ a geometric approach assuming that human legs or body have a specific shape and geometric parameters [16,27,35] or apply well-established machine learning techniques [4,12,24]. Yet, the assumption that person clothes allow distinguishing two separate legs in LRF data is not applicable when clothes completely cover legs, e.g., a person wearing a skirt or a long overcoat. Moreover, while LRF-based tracking is mainly successful in large open spaces, they often fail distinguishing human legs from other obstacles in a cluttered environment. Multidimensional LRF data based approaches could significantly improve accuracy of human detection, position estimation and tracking [8]. Alternatively, a 3D LiDAR could be used to track objects in 3D space [20,36], but it is important to emphasize that these sensors are quite expensive.

Additional visual sensors, such as mono cameras, stereo cameras, and 360° cameras, are used to increase person-detection and tracking reliability [32]. In [30] authors applied skeleton tracking by Microsoft Kinect SDK and an auxiliary tracker, which utilized Camshift algorithm. In [15] a mobile robot used an RGB-D camera based simple 2D tracker of a human. In [23] an RGB-D camera based framework for a human following robot combined deep learning and variational Bayesian techniques. A stereo vision based CNN tracker for a person-following robot was proposed in [10]. In [18] a person-following robot used a monocular camera for a person tracking and identification. However, if a robot has a low height, then a human head and upper body are not always contained in a camera's field of view. This makes visual person detection and tracking difficult. To increase a person identification and tracking reliability, it is advisable to jointly use a LRF and a camera. A person-following algorithm in [19] fused data of a LRF and a panoramic camera. In [17] authors suggested a person tracking and following algorithm for a mobile robot, which was equipped by a right-side monocular camera and a frontal LRF.

This paper presents a human following algorithm for an autonomous person-following mobile robot of a low height. The algorithm is based on a LRF and a monocular camera data fusion. We implemented and tested our approach with the PAL Robotics TIAGo Base [29] wheeled differential drive robot. The robot has a cylindrical shape with 54 cm diameter and 30 cm height. Relatively small dimensions allow the robot to traverse environments of various types, including cluttered office rooms [6]. In addition to original Hokuyo URG-04LX-UG1 LRF [1], which is located in the front of the robot at 10 cm height from a floor level, we

equipped the TIAGo Base with a monocular Web camera of $640 \times 480$ resolution. The camera was placed on the top of the robot, in its front, at a height of 40 cm from the floor level. The person-following algorithm was initially validated in the Gazebo simulator [3,22] and then in a real world indoor environment.

## 2 Human Tracking and Following

### 2.1 Human Tracking

To successfully follow a person, a robot needs to track him/her continuously and to be capable of searching and identifying the person in case of loss. Our algorithm simultaneously tracks a person in two 2D images: the first image is constructed from LRF data and the second is a monocular camera image. Using LRF data we constructed a 2D image of $500 \times 500$ pixels, which contains LRF distance readings toward environment obstacles. Closely located pixels that corresponded to occupied locations (a static or a dynamic obstacle, including a human) were connected with straight lines. Figures 1a, b show examples of LRF-based images with a person in front of the TIAGo Base robot: a person in trousers with clearly distinguished legs (Fig. 1a) and a person in a skirt (Fig. 1b).

Our previous work [9] compared five popular trackers of OpenCV library: KCF, TLD, Median Flow, MOSSE and MIL [5]. They were compared in a virtual world of the Gazebo simulator for a person tracking task using LRF data. MIL [5] tracker demonstrated less false positives and a longer path length of successful tracking (without loosing a tracked person) than the other four approaches. However, it should be emphasized that a common disadvantage of all five trackers was a large number of false positives when a person approached any virtual obstacles including walls, furniture or other people. Therefore, to successfully track a person in LRF data, a tracker should be capable to distinguish a person from another object within the LRF data. Combining strong features of LRF-based and camera-based tracking we could obtain better tracking results. Yet, to combine these data, a person should simultaneously appear within a camera and an LRF field of view, which allows (at least partial) matching of LRF-based and camera-based images. This paper proposes a new method of LRF and camera data fusion that elicits a more precise human tracking and following.

Both in virtual experiments within the Gazebo environment and real world experiments with TIAGo Base robot we use the following notations in this paper:

- $LRF_{img}$ is $n \times n$ pixel binary image, which is constructed using LRF data;
- $LRF_p$ is $d \times d$ pixel fragment of $LRF_{img}$ that contains a person (exhaustive target person data within $LRF_{img}$ image, which could reflect the entire person or only a part of his/her body);
- $C_{img}$ is an image from a camera;
- $C_p$ is a fragment of $C_{img}$ that corresponds to $LRF_p$ fragment (i.e., the person).

Our method requires at least $n = 500$ and $d = 75$, which were obtained empirically. In this case, $500 \times 500$ pixel image $LRF_{img}$ corresponds to a square

pattern that reflects virtual or real world data within $5 \times 5$ m$^2$ square. When tracking starts we save the first $C_p$ as a sample image $S$ and keep updating $C_p$ variable with time. If a person is simultaneously within a camera and an LRF fields of view, a new value of $C_p$ is obtained and compared with $S$. In addition to tracking the target person within LRF data we track the person in camera data using an independent tracker. $C_\tau$ denotes an image of the person, which is obtained using the independent tracker. Figure 1c demonstrates examples of images $LRF_p$, $C_p$, $C_\tau$ and $S$.

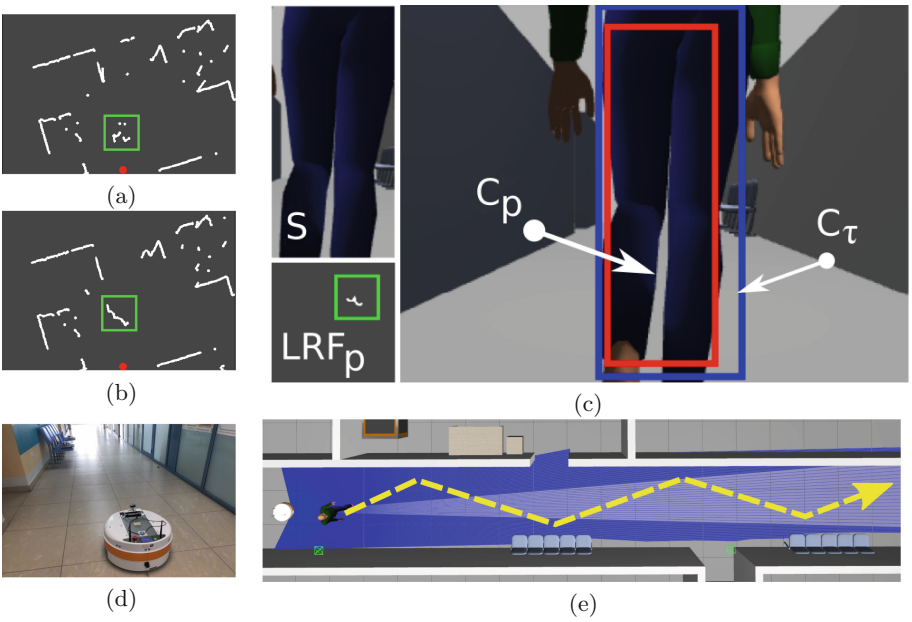


**Fig. 1.** (a) LRF data based 2D image. Black pixels correspond to empty or unknown (due to an occlusion) space, white pixels correspond to visible obstacles. The red dot denotes the TIAGo Base robot position. Legs of a person in trousers form two curves (inside the green square). (b) LRF data based 2D image. A person in a long skirt forms the single curve (inside the green square). (c) Examples of $LRF_p$, $C_p$, $C_\tau$, $S$ images in the Gazebo virtual experiments. (d) Real world experiments with the TIAGo Base robot inside the 2-nd Study Building of KFU, 35 Kremlevskaya street, Kazan. (e) A person trajectory in virtual experiments in Gazebo. (Color figure online)

A person is displayed in $LRF_p$ (LRF data based) image as a set of points and there is a particular $C_p$ (monocular camera data based) image that corresponds to $LRF_p$. The idea of our method is based on the construction of an abstract metric space. The points of this space are ordered pairs of the form $(LRF_p, C_p)$. The tracked person is associated with some point $(LRF_p, C_p)$ of this metric space at each moment. We introduce a feature system for each pair $(LRF_p, C_p)$. At the initialization time images $LRF_p$ and $C_p$ correspond to a target person.

To determine a position of a person at the next moment, we construct a metric, which allows us to compare objects within the LRF and camera field of view with the initial sample.

We denote a correlation coefficient of normalized histograms of $C_p$ and $S$ as $corr(C_p, S)$. $LRF_p(x, y)$ denotes a pixel value at $(x, y)$ cell of an image, where x and y are integer numbers so that $0 \leq x, y \leq d-1$. $LRF_p(x, y)$ takes a value of 0 for an empty cell and 1 for a cell with an obstacle, which could be a person or another object (in Fig. 1a black pixels are assumed to be an empty space, white pixels correspond to visible obstacles). Thus $LRF_p$ image could be viewed as a matrix with zeros and ones. We process $LRF_p$ matrix in a such way that each column of a resulting matrix $LRF_p^N$ has at most one nonzero element. The convention is to keep only the lowest index element if several nonzero elements appear in the column. Next, using $LRF_p^N$ matrix, we construct function $f(x)$ on the set $[0, d-1]$. The value of $f(x)$ is a row number of a nonzero element of column $x$ and is zero if $LRF_p(x, Y) = 0$ for all $Y \in [0, d-1]$.

Consider an abstract space $M$ whose elements are ordered pairs $(LRF_p, C_p)$. Let $P = P(LRF_p, C_p)$ be a point of this space. Using a set of $\{n_i\}, i = [0..7]$ functions for each point of space $M$ we describe a target human appearance using a number of mathematical features of his/her shape as follows:

- $n_0(P)$ denotes function $f(x)$, which is constructed for point $P$ (its $LRF_p$ part) by the method described above;
- $n_1(P)$ denotes a number of local minima of $f(x)$;
- $n_2(P)$ denotes a number of discontinuity points of $f(x)$;
- $n_3(P)$ denotes a cardinality of a set of points at which $f(x) \neq 0$;
- $n_4(P)$ denotes a *height* of a tracked curve in terms of columns of $LRF_p$ image $n_4(P) = |\max f(x) - \min f(x)|$;
- $n_5(P)$ denotes an indicator function that allows detecting whether $LRF_p$ image contains a straight line as follows:

$$n_5(P) = \begin{cases} 0 & if |corr(x, f(x))| < \sigma_1 \\ 1 & if |corr(x, f(x))| \geq \sigma_1 \end{cases}, \tag{1}$$

  where $corr(x, f(x))$ is a linear correlation coefficient of $x$ and $f(x)$, $\sigma_1$ is an empirically established threshold value, which was set to 0.9. $n_5(P)$ allows reducing probability of false positives on walls and other objects with flat vertical surfaces;
- $n_6(P_0, P)$ denotes a result of comparing $C_p$ image and the one that is obtained by an independent tracker

$$n_6(P_0, P) = |corr(C_{p_0}, C_\tau) - corr(C_p, C_\tau)|, \tag{2}$$

  where $P_0(LRF_{p_0}, C_{p_0})$, $P(LRF_p, C_p) \in M$; $C_{p_0}, C_p \subset C_{img}$ and $C_{p_0} \neq \varnothing$, $C_p \neq \varnothing$; $C_\tau$ was previously defined as an image of the person, which is obtained by the independent tracker.

– $n_7(P_0, P)$ denotes a result of comparing $C_p$ image with the initial sample of the human image

$$n_7(P_0, P) = \begin{cases} 0, & if\ corr(C_{p_0}, C_p) \geq \sigma_2 \\ 1, & if\ corr(C_{p_0}, C_p) < \sigma_2 \end{cases}, \tag{3}$$

where $\sigma_2$ is an empirically established threshold value, which was set to 0.8 in our experiments.

For space $M$ we set the following metric for measuring a distance between two distinct points $P_0$ and $P_1$:

$$\begin{aligned} \rho(P_0, P_1) &= \lambda_0 \max |n_0(P_0) - n_0(P_1)| \\ &+ \sum_{i=1}^{5} \lambda_i |n_i(P_0) - n_i(P_1)| + \sum_{i=6}^{7} \lambda_i n_i(P_0, P_1), \end{aligned} \tag{4}$$

where $\lambda_i > 0$ are weighting coefficients that allow to adjust metric in Eq. (4) sensitivity to particular parameters. In our experiments all values $\{\lambda_0, .., \lambda_7\}$ were set to 1 and a research of their correlation and selection is left as a part of our future work.

Using metric defined by Eq. (4) we constructed a tracker that fuses LRF and monocular camera data. At an initial moment of time a target person is the closest object to a robot, which allows setting $P$ as a set of a sample image $S$ of the target person and the corresponding $n_i$, $i = [0..7]$. Next, at each step the algorithm could find point $P_0$, which is the closest one to $P$.

As an independent tracker that provides image $C_\tau$ we used the MIL tracker [5]. To eliminate the problem of loss and false positives by the MIL tracker, we re-initialize it with a newly obtained target person image every time when $\rho(P, P_0) < \sigma_\tau$ holds. Here $\sigma_\tau$ is an empirically established threshold value, which is set to 7 in our experiments.

### 2.2 Person-Following Algorithm

Using the above described tracker a robot can determine a position and coordinates of a person. The next task is to plan a path from current robot position $S_{loc}$ to local target $T_{loc}$, which is located in a close vicinity of a person, at a safe distance from him/her [25]. Points $S_{loc}$ and $T_{loc}$ are selected in a robot centered system of coordinates, and thus $S_{loc}$ is the origin (0,0) of the robot centered system. If a straight line ($S_{loc}$,$T_{loc}$) intersects an obstacle, the algorithm applies a local planner, e.g., the ROS *move_base* package. Upon reaching $T_{loc}$, the algorithm sets $S_{loc}$ to $T_{loc}$, defines new $T_{loc}$ and plans a path to new $T_{loc}$.

The person-following algorithm is presented in Algorithm 1. At a start, a person stands in front of a robot and the robot determines local to LRF coordinates $\widetilde{\Psi}$ of a nearest object using LRF. These coordinates correspond to $P(LRF_p(\widetilde{\Psi}), C_p(\widetilde{\Psi}))$ point of space $M$. We assume that $LRF_p(\widetilde{\Psi})$ and $C_p(\widetilde{\Psi})$ images always contain an image of the person. To determine coordinates of the person at the next moment, we build an array of 100 points $\Psi_i$ (i=1..100), located

no farther than Euclidean distance $\varepsilon$ (empirically selected as $\varepsilon = d$) from a previous position of the person. The selection of 100 points is an empirically defined trade-off between accuracy and execution time. This array corresponds to array of points $P_i(LRF_p(\Psi_i), C_p(\Psi_i))$. First, we fix a minimum value $\rho(P, P_i)$ as $P_0$ point; then we select a point with index $i$ that has a minimal value $\rho(P, P_i)$ within the array. If $\rho(P, P_i) < \theta$ (empirically selected as $\theta = 34$ based on an analysis of a set of virtual pilot experiments), then the robot moves to point $\Psi_i$; otherwise the robot begins searching for a person. During the search, the robot compares all detected objects with the person using Eq. (4).

**Algorithm 1.** Person-following algorithm

1: Get $LRF_{img}$ and $C_{img}$ images
2: Find nearest object coordinates $\widetilde{\Psi}$
3: $P := (LRF_p(\widetilde{\Psi}), C_p(\widetilde{\Psi}))$
4: $\Psi_{prev} := \widetilde{\Psi}$
5: Initialize tracker for $C_\tau$
6: **while not** "Return command" **do**
7: Update $LRF_{img}$, $C_{img}$, $C_\tau$
8: $P_0 := (LRF_p(\Psi_{prev}), C_p(\Psi_{prev}))$
9: $min_\rho := \rho(P, P_0)$
10: **for all** $|\Psi_i - \Psi_{prev}| < \varepsilon$ **do**
11: $P_i := (LRF_p(\Psi_i), C_p(\Psi_i)$
12: **if** $\rho(P, P_i) < min_\rho$ **then**
13: $min_\rho := \rho(P, P_i)$
14: $\Psi_{goal} := \Psi_{min}$
15: **end if**
16: **end for**
17: **if** $min_\rho < \theta$ **then**
18: **if** $min_\rho < \sigma_\tau$ **then**
19: Initialize tracker for $C_\tau$
20: **end if**
21: Move to $T_{loc}(\Psi_{goal})$
22: $\Psi_{prev} := \Psi_{goal}$
23: **else**
24: Run person search procedure
25: **end if**
26: **end while**

## 3 Experiments

This section describes experimental work with TIAGo Base robot in the Gazebo simulator virtual environment and in real world environment.

### 3.1 Simulation in Gazebo

For real world experiments, we used a corridor of the 2-nd study building of Kazan Federal University (Fig. 1d). A part of the corridor and one room of the Laboratory of Intelligent Robotic Systems (LIRS) were modelled in the Gazebo simulator for virtual experiments (Fig. 1e). Algorithm 1 was implemented in the Gazebo simulator using the TIAGo Base robot simulation packages [2]. A 38 m piecewise linear trajectory of a target person (Fig. 1e) was used in 400 virtual experiments in the Gazebo simulation (Fig. 2a). Figure 2b demonstrates $LRF_{img}$, $LRF_p$, $C_{img}$, $C_p$ and $C_\tau$ that correspond to the particular frame of Fig. 2a.

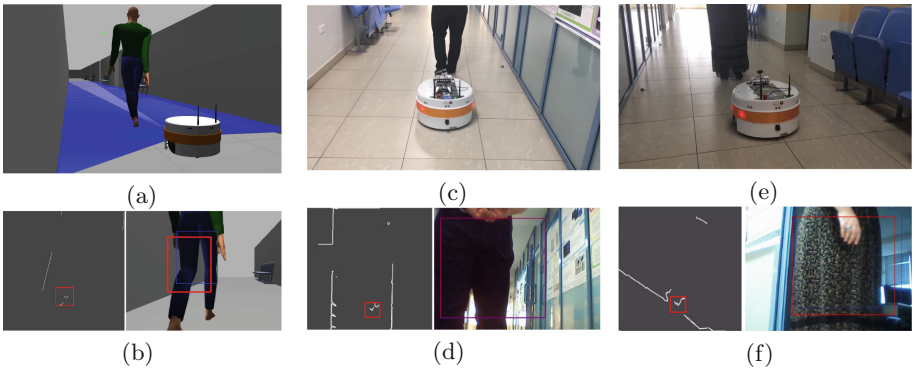


**Fig. 2.** (a) The Gazebo simulation frame of Algorithm 1. (b) Tracking a person using LRF (left) and camera (right) data in the Gazebo. On the left: $LRF_{img}$ image with $LRF_p$ area denoted with the red square. On the right: $C_{img}$ image with $C_p$ area (inside the red rectangle) and $C_\tau$ area (inside the blue rectangle). (c) TIAGo Base follows a person in pants. (d) Tracking the person in the pants using LRF (left) and camera (right) data. In the left image the target person is in the red square" (e) TIAGo Base follows a person in a long dress that covers legs. (f) Tracking the person in the long dress using LRF (left) and camera (right) data. In the left image the target person is in the red square. (Color figure online)

In our previous work [9], we attempted tracking a straight line walking human in LRF data with KCF, TLD, Median Flow, MOSSE and MIL trackers. Their performance comparison demonstrated a clear superiority of the MIL tracker. Therefore in this work the MIL tracker was selected as an independent tracker (within Algorithm 1) as well as the only competitor for Algorithm 1 evaluation. We used standard parameters of the MIL tracker of the OpenCV library: a radius for gathering positive instances during init – 3, a search window size – 25, negative samples to use during init – 65, radius for gathering positive instances during tracking – 4, positive samples to use during tracking – 100000, negative samples to use during tracking – 65, features – 250.

We evaluate the accuracy of person-following algorithms according to an average traveled distance, tracker false positives, and an average person speed. On a false positive occurrence, we stopped the experiment and measured a distance traveled by the robot. For algorithm performance evaluation, the following accuracy score was used:

$$\mu = \frac{1}{2}\left(1 + \frac{\sum_i l_i}{\sum_i L_i} - \frac{N_f}{N}\right), \mu \in [0, 1], \tag{5}$$

where $N_f$ is a number of false positives, $l_i$ is a distance traveled by a person in $i$-th experiment, $L_i$ is the person's path length, $N$ is a total number of experiments. The four aforementioned values correspond to an entire set of experiments for an algorithm with a particular human walking speed. The ratio $\sum_i l_i / \sum_i L_i$ allows evaluating whether a robot succeeded to complete an entire path without losing the person. The ratio $N_f/N$ reflects an average number of false positives per experiment. The parameter $\mu$ reaches its maximum of $\mu = 1$ if the robot succeeded to follow the person along the entire path ($\sum_i l_i = \sum_i L_i$) and there were no false positives ($N_f = 0$). The minimum possible value is $\mu = 0$ if the robot did not start (i.e., immediately lost the human at the start and thus $\sum_i l_i = 0$) and each test was featured with a false positive ($N_f = N$).

A pedestrian walking speed depends on several factors including a person age [7] and environment conditions [11]. Taking into account that the maximum speed of the TIAGo Base robot is only 1 m/s, we tested the algorithms for two human walking speeds of 0.5 and 0.9 m/s. Table 1 presents virtual experimental results in the Gazebo simulation for MIL tracker and for Algorithm 1 (denoted as Alg1). Both algorithms were tested at 0.5 m/s and 0.9 m/s walking speed of a human with 100 runs for each speed.

In total, we conducted 400 virtual experiments. In each experiment, a person walked along a corridor according to the preplanned route (Fig. 1e). The route length in each experiment was 38 m. When a false positive was detected, the experiment stopped and the total distance traveled by the robot was measured at this point. MIL tracker demonstrated 41 false positives within 100 experiments at a speed of 0.5 m/s and 49 false positives within 100 experiments at a speed of 0.9 m/s. Algorithm 1 did not provide any false positives within 100 experiments both at 0.5 m/s and 0.9 m/s speed. The total accuracy score was calculated with Eq. (5). For MIL tracker the score was 0.67 at 0.5 m/s and 0.57 at 0.9 m/s, while Algorithm 1 demonstrated maximal possible accuracy score (of 1) at both speeds due to a monocular camera and LRF data fusion approach.

### 3.2 Experiments in a Real World Environment

A set of real world experiments was conducted in the environment that served as a source for Gazebo environment modelling (Fig. 1d). One male and one female participated in real world experiments. The male wore black trousers (Fig. 2c, 2d) and the female tested one pair of trousers and 2 different long one-piece garments (Fig. 2e, 2f). We asked the experiment participants to stand at a short distance

**Table 1.** Experimental results in the Gazebo simulation and in real environment for MIL tracker and for Algorithm 1 (denoted as Algorithm 1).

| Environment | Simulation | | | | Real world | |
|---|---|---|---|---|---|---|
| Tracker | MIL | MIL | Alg1 | Alg1 | MIL | Alg1 |
| Average human walking speed, m/s | 0.5 | 0.9 | 0.5 | 0.9 | 0.5 | 0.5 |
| Average distance traveled by robot, m | 28.2 | 24.1 | 38 | 38 | 19.1 | 28.9 |
| Total false positives number | 41 | 49 | 0 | 0 | 14 | 0 |
| Number of experiments | 100 | 100 | 100 | 100 | 30 | 30 |
| Total accuracy score | 0.67 | 0.57 | 1 | 1 | 0.59 | 0.89 |

from the robot (between 20 cm and 1.8 m) and slowly walk a distance of 30 m. If the tracker lost a person the robot stopped moving.

Our experiments demonstrated a stable behavior of Algorithm 1 regardless of having clothes that completely cover human legs or allow to distinguish two separate legs. Figure 2c demonstrates an experiment of following a person in trousers that allow distinguish two legs. Figure 2e shows an experiment of following a person in a long dress that does not allow distinguish two legs. Figures 2d, f show examples of areas $LRF_{img}$, $LRF_p$, $C_{img}$, $C_p$ and $C_\tau$ for people in different clothes. The same experiments were conducted for the MIL tracker.

Table 1 also presents the results of 30 real world experiments. In these experiments the robot speed was set to 0.5 m/s. Total accuracy score was calculated using Eq. (5). Algorithm 1 demonstrated significantly better results of $\mu = 0.89$ than the MIL approach with $\mu = 0.59$. Moreover, while the MIL approach had false positives almost in 50% of the experiments, Algorithm 1 did not have any false positives in all 30 runs.

## 4 Conclusions and Future Work

In this paper, we presented a novel person tracking and following approach for autonomous mobile robots that are equipped with a 2D laser range finder (LRF) and a monocular camera. A LRF and a monocular camera data fusion improved a person tracking reliability for indoor environments. The proposed method does not impose restrictions on person's clothes and does not require a head or an upper body to be within a monocular camera field of view. This allows to employ our method for low height indoor robots. Our algorithm was implemented in the Gazebo simulator and validated with TIAGo Base mobile robot. In addition, we proposed a new algorithm performance estimation criterion as a total accuracy score, which is a function of false positives number and traveled distance by a person and by a robot. The algorithm performance was compared with the MIL tracker performance using the proposed accuracy score in the Gazebo simulation

and in real world experiments with the TIAGo Base robot and demonstrated significantly better results than the MIL tracker.

In this work, the algorithms were tested only in static environments with a single walking person in a scene. As a part of our future work, we plan to extend our algorithm to the case of multiple static and dynamic objects in a scene, to increase a number of participants and clothing variety in real world experiments and to validate the algorithm at a broad variety of human walking speeds.

**Acknowledgements.** The reported study was funded by the Russian Foundation for Basic Research (RFBR) according to the research project No. 19–58-70002. Special thanks to PAL Robotics for their kind professional support with TIAGo Base robot software and hardware related issues.

## References

1. PMB-2 technical specifications. http://pal-robotics.com/wp-content/uploads/2016/07/PMB-2-Datasheet.pdf
2. Tiago base. http://wiki.ros.org/Robots/PMB-2
3. Abbyasov, B., Lavrenov, R., Zakiev, A., Yakovlev, K., Svinin, M., Magid, E.: Automatic tool for gazebo world construction: from a grayscale image to a 3D solid model. In: International Conference on Robotics and Automation (ICRA), pp. 7226–7232 (2020)
4. Arras, K.O., et al.: Range-based people detection and tracking for socially enabled service robots. In: Prassler, E., et al. (eds.) Towards Service Robots for Everyday Environments, pp. 235–280. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-25116-0_18
5. Babenko, B., Yang, M., Belongie, S.: Visual tracking with online multiple instance learning. In: Conference on Computer Vision and Pattern Recognition, pp. 983–990. IEEE (2009)
6. Bereznikov, D., Zakiev, A.: Network failure detection and autonomous return for PMB-2 mobile robot. In: International Conference on Artificial Life and Robotics (ICAROB 2020), pp. 444–447 (2020)
7. Bohannon, R.: Comfortable and maximum walking speed of adults aged 20–79 years: Reference values and determinants. Age Ageing **26**(1), 15–19 (1997)
8. Carballo, A., Ohya, A., Yuta, S.: Reliable people detection using range and intensity data from multiple layers of laser range finders on a mobile robot. Int. J. Soc. Robot. **3**, 167–186 (2011)
9. Chebotareva, E., Hsia, K.H., Yakovlev, K., Magid, E.: Laser rangefinder and monocular camera data fusion for human-following algorithm by PMB-2 mobile robot in simulated Gazebo environment. Smart Innovation, Syst. Technol. **187** (2020)
10. Chen, B.X., Sahdev, R., Tsotsos, J.K.: Integrating stereo vision with a CNN tracker for a person-following robot. In: Liu, M., Chen, H., Vincze, M. (eds.) ICVS 2017. LNCS, vol. 10528, pp. 300–313. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68345-4_27
11. Franěk, M.: Environmental factors influencing pedestrian walking speed. Percept. Mot. Skills **116**(3), 992–1019 (2013)

12. Guerrero-Higueras, Á.M., et al.: Tracking people in a mobile robot from 2D LIDAR scans using full convolutional neural networks for security in cluttered environments. Front. Neurorobot. **12** (2018)
13. Islam, M.J., Hong, J., Sattar, J.: Person-following by autonomous robots: a categorical overview. The Int. J. Robot. Res. **38**(14), 1581–1618 (2019)
14. Islam, M.M., Lam, A., Fukuda, H., Kobayashi, Y., Kuno, Y.: An intelligent shopping support robot: understanding shopping behavior from 2D skeleton data using GRU network. ROBOMECH J. **6**(1), 1–10 (2019). https://doi.org/10.1186/s40648-019-0150-1
15. Jiang, S., Li, L., Hang, M., Kuc, T.: An adaptive 2D tracking approach for person following robot. In: International Symposium on Computer Science and Intelligent Controls, pp. 147–151 (2017)
16. Kawarazaki, N., et al.: Development of human following mobile robot system using laser range scanner. Procedia Comput. Sci. **76**, 455–460 (2015)
17. Kim, H., et al.: Sensor fusion-based human tracking using particle filter and data mapping analysis in in/outdoor environment. In: International Conference on Ubiquitous Robots and Ambient Intelligence, pp. 741–744 (2013)
18. Koide, K., et al.: Monocular person tracking and identification with on-line deep feature selection for person following robots. Robot. Auton. Syst. **124** (2020)
19. Kristou, M., et al.: Target person identification and following based on omnidirectional camera and LRF data fusion. In: International Conference on Robot & Human Interactive Communication, pp. 419–424. IEEE (2011)
20. Lang, A., Vora, S., Caesar, H., Zhou, L., Yang, J., Beijbom, O.: Pointpillars: fast encoders for object detection from point clouds. In: Conference on Computer Vision and Pattern Recognition, pp. 12689–12697 (2019)
21. Lavrenov, R., Matsuno, F., Magid, E.: Modified spline-based navigation: guaranteed safety for obstacle avoidance. In: Ronzhin, A., Rigoll, G., Meshcheryakov, R. (eds.) ICR 2017. LNCS (LNAI), vol. 10459, pp. 123–133. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66471-2_14
22. Lavrenov, R.O., Magid, E.A., Matsuno, F., Svinin, M.M., Suthakorn, J.: Development and implementation of spline-based path planning algorithm in ROS/gazebo environment. Trudy SPIIRAN **18**(1), 57–84 (2019)
23. Lee, B.J., et al.: Robust human following by deep Bayesian trajectory prediction for home service robots. In: International Conference on Robotics and Automation, pp. 7189–7195. IEEE (2018)
24. Leigh, A., et al.: Person tracking and following with 2D laser scanners. In: International Conference on Robotics and Automation, pp. 726–733. IEEE (2015)
25. Magid, E., Lavrenov, R., Khasianov, A.: Modified spline-based path planning for autonomous ground vehicle. ICINCO **2**, 132–141 (2017)
26. Moskvin, I., Lavrenov, R.: Modeling tracks and controller for servosila engineer robot. In: Ronzhin, A., Shishlakov, V. (eds.) Proceedings of 14th International Conference on Electromechanics and Robotics "Zavalishin's Readings". SIST, vol. 154, pp. 411–422. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9267-2_33
27. Nakamori, Y., Hiroi, Y., Ito, A.: Multiple player detection and tracking method using a laser range finder for a robot that plays with human. ROBOMECH Journal **5**(1), 1–15 (2018). https://doi.org/10.1186/s40648-018-0122-x
28. Orita, Y., Fukao, T.: Robust human tracking of a crawler robot. J. Robot. Mechatron. **31**(2), 194–202 (2019)

29. Pages, J., Marchionni, L., Ferro, F.: Tiago: the modular robot that adapts to different research needs. In: International Workshop on Robot Modularity, IROS (2016)
30. Ren, Q., et al.: Real-time target tracking system for person-following robot. In: Chinese Control Conference, pp. 6160–6165 (2016)
31. Ronzhin, A., Saveliev, A., Basov, O., Solyonyj, S.: Conceptual model of cyberphysical environment based on collaborative work of distributed means and mobile robots. In: Ronzhin, A., Rigoll, G., Meshcheryakov, R. (eds.) ICR 2016. LNCS (LNAI), vol. 9812, pp. 32–39. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-43955-6_5
32. Safin, R., Lavrenov, R., Tsoy, T., Svinin, M., Magid, E.: Real-time video server implementation for a mobile robot. In: 11th International Conference on Developments in eSystems Engineering (DeSE), pp. 180–185. IEEE (2018)
33. Sato, Y., et al.: A maneuverable robotic wheelchair able to move adaptively with a caregiver by considering the situation. In: International Conference on Robot & Human Interactive Communication, pp. 282–287. IEEE (2013)
34. Simakov, N., Lavrenov, R., Zakiev, A., Safin, R., Martínez-García, E.A.: Modeling USAR maps for the collection of information on the state of the environment. In: 2019 12th International Conference on Developments in eSystems Engineering (DeSE), pp. 918–923. IEEE (2019)
35. Sung, Y., Chung, W.: Hierarchical sample-based joint probabilistic data association filter for following human legs using a mobile robot in a cluttered environment. IEEE Trans. Hum. Mach. Syst. **46**(3), 340–349 (2015)
36. Yan, Y., Mao, Y., Li, B.: SECOND: sparsely embedded convolutional detection. Sensors **18**, 3337 (2018)