Occlusion-based Trajectory Estimation for Pedestrians using LiDAR sensors

Takayuki Saki, Akihito Hiromori, Hirozumi Yamaguchi, and Teruo Higashino

Graduate School of Information Science & Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka, Japan
{t-saki, hiromori, h-yamagu, higashino}@ist.osaka-u.ac.jp

Abstract. LiDAR sensors may provide an alternative solution for surveillance purposes because LiDAR sensors do not record privacy sensitive data unlike cameras used in the mainstream methods. Since the methods also extract features of body parts to detect pedestrians based on laser points, the complex clustering process is required for the huge amount of the laser point data. This paper proposes an alternative way to derive pedestrian trajectories using LiDAR sensors. Unlike the existing approaches, the proposed method does not utilize foreground points derived from pedestrians but utilizes background points representing objects such as walls, obstacles and so on. Such background points cannot be observed in the case that a pedestrian is crossing in front of the points. The occlusion of the points indicates the existence of the pedestrian. We estimate the trajectory of the pedestrian by aggregating such several occlusions that occurred by the same pedestrian. The method measures time not to observe the background points and estimates pedestrian trajectories from the occlusions by simple mathematical calculations on several laser points. The simulation results showed that the proposed method can detect more pedestrians located far from sensors and in rather crowded situations compared with an existing scheme based in point clouds.

Keywords: LiDAR sensors · Pedestrian · Tracking · Trajectory Estimation

1 Introduction

GPS is commonly used for a wide range of location-dependent applications. However, it is not appropriate for surveillance purposes because mobile terminals with GPS capability like smartphones have to report their positions. Since it may pose scalability problems, visual-based methods are widely used for such purpose as shown in a survey [2]. Recently, LiDAR sensors may provide an alternative solution for the purpose because LiDAR sensors do not record privacy sensitive data unlike cameras used in the visual-based methods. They periodically transmit laser pulses toward a fan-shaped area with a signal transmission range of a few tens of meters and then receive reflected signals from the surrounding objects. Thus, the distance to those objects can be calculated based on the time of flight (TOF) of the measurement signals. The distance measurement error is usually within a few centimeters, which is almost the best accuracy among existing commercial sensing devices. The crowd tracking capability of LiDAR sensors can be seen in [5] where it is shown through a field experiment that 100 people in an exhibition hall were successfully detected by a combination of three LiDAR sensors. The methods categorized into this category also extract features of body parts to detect pedestrians from laser points as same as the visual-based methods. Although precise data can contribute to detecting pedestrian positions and trajectories with high accuracy, the complex clustering process is required for the huge amount of the data. In addition, the methods can not detect pedestrians in some cases even if they are in the sensing area. Although a commercial LiDAR sensor transmits 1080 lasers in 270 ° on the sensing region, only several lasers hit a pedestrian located in far from the sensor and the pedestrian cannot be detected by the existing method because the pedestrian should be represented as a point cloud in the existing methods. Also, a point cloud might not be constructed either in the case where there are many pedestrians because pedestrians hide some other pedestrians from sensors and only a part of the pedestrian can be seen from the sensors. Thus, the existing methods do not always detect pedestrians correctly even in the sensing area since not only sensor deployment but also pedestrian mobility prevent to construct point clouds for pedestrians.

We propose a different scheme to detect a trajectory of a pedestrian in order to improve the feasibility of the LiDAR-based method. The proposed method does not utilize foreground points derived from pedestrians but utilizes background points representing objects such as walls, obstacles and so on. Such background points cannot be observed in the case that a pedestrian is crossing in front of the background points. The occlusion of the points indicates the existence of the pedestrian. We estimate the trajectory of the pedestrian by aggregating such several occlusions that occurred by the same pedestrian. The method measures time not to observe the background points, and estimates pedestrian trajectories from the occlusions by simple mathematical calculations on a part of laser points instead of whole laser points. Thus, the proposed method can derive pedestrian trajectories from a small number of laser points since the proposed method does not require point clouds representing human bodies. The simulation results showed that the proposed method can detect more pedestrians located far from sensors and in rather crowded situations compared with an existing scheme based in point clouds.

2 Related Work

Detecting pedestrians have been studied for surveillance, social security, construction planning and so on. Many approaches have been proposed to detect pedestrians by analyzing image data [2]. These visual-based surveillance methods analyze images obtained by cameras and derive features of pedestrians in the images by subtracting the background, measuring features of remain parts and detecting pedestrians. Deep learning-based methods [4] are also introduced recently for detecting pedestrians with high accuracy. The method can help detection algorithms with high complexity to run faster. The performance of the methods also depends on much larger datasets and preparing appropriate datasets might be bottleneck points to deploy the system. Even though modern approaches can improve accuracy, these methods are time-consuming and are not suitable for real-time pedestrian detection.

Infrared laser provides an efficient alternative for detecting pedestrians. LIDAR sensors can emit eye-safe laser beams to measure the distances of nearby pedestrians. Zhao et al. [1, 5] detect pedestrians by using a number of laser-range sensors on the ground level to monitor pedestrians ’ feet. Similar attempts are conducted in Ref [3] to estimate pedestrians by setting laser-range sensors to target the waists of pedestrians. These methods calculate features of human parts from point data obtained by LiDAR sensors and require that pedestrians should be represented as a point cloud. In this paper, we propose a different scheme to detect pedestrians by estimating pedestrians based on a part of laser points so that our method can detect pedestrians located far from sensors and in congested situations.

3 Modeling pedestrian trajectories and occlusions of background points

We propose a method to detect pedestrians and estimate their trajectories from occlusion time that pedestrians hide background points from LiDAR sensors. LiDAR sensors transmit laser beams radially and can measure the distances for each beam direction. The background points can be modeled as laser points without any pedestrian in the

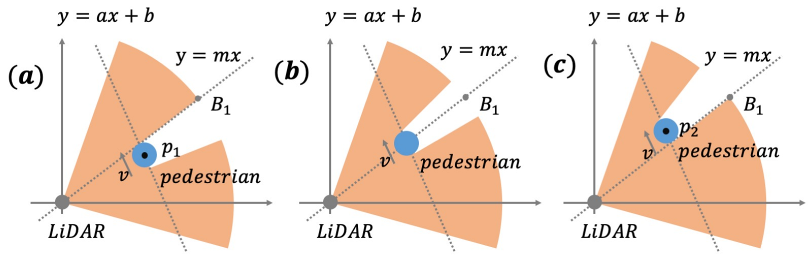


Fig. 1: Relationship between a pedestrian and a background point

sensing area. When a pedestrian moves there, it hides some background points since it prevents laser beams from LiDAR sensors. This occlusion indicates the existence of the pedestrian, and the duration of the occlusion depends on its trajectories and velocity. Our method estimates a trajectory and velocity for a pedestrian by utilizing three occlusions occurred by the pedestrian. At first, we explain how to model the relationship between pedestrian trajectories and occlusions that occurred by pedestrians.

A pedestrian can be modeled as an object with shoulder width that of the average. Accordingly, we define a pedestrian as a cylinder with a radius of r. r can be estimated by applying the existing methods when a pedestrian enters the sensing area occasionally. We assume that pedestrians move straight among a sufficiently small time slot ts. Thus, the trajectory of a pedestrian can be represented as a line $y = a^{ts}x + b^{ts}$ in time slot ts. Each occlusion time caused by a pedestrian depends on its mobility. We associate the occlusion time of background points with a, b and v to represent the relationship between the occlusion and trajectory.

At first, we explain how to calculate an occlusion time of a background point. Fig.1 shows a geographic relationship between a pedestrian and a background point. The sensor is located at $(0, 0)$, and the pedestrian moves along a line $y = ax + b$ with velocity v. B_1 is a background point on a line $y = mx$. The orange region represents the sensing area. Fig.1-(a) shows when the pedestrian begins to hide B_1. p_1 is the coordinate of the pedestrian at the time. The coordinate can be calculated by the following equations.

$$\begin{cases} y = mx - r\sqrt{1 + m^2} & \text{(1a)} \\ y = ax + b & \text{(1b)} \end{cases}$$

$$p_1 = \left(\frac{b + r\sqrt{1 + m^2}}{m - a}, \frac{mb + ar\sqrt{1 + m^2}}{m - a} \right) \tag{2}$$

Fig.1-(b) shows that the pedestrian is hiding B_1 from the sensor. After that, the sensor can see B_1 as shown in Fig.1-(c). p_2 is the coordinate of the pedestrian and can be calculated by the following equations.

$$\begin{cases} y = mx + r\sqrt{1 + m^2} & \text{(3a)} \\ y = ax + b & \text{(3b)} \end{cases}$$

$$p_2 = \left(\frac{b - r\sqrt{1 + m^2}}{m - a}, \frac{mb - ar\sqrt{1 + m^2}}{m - a} \right) \tag{4}$$

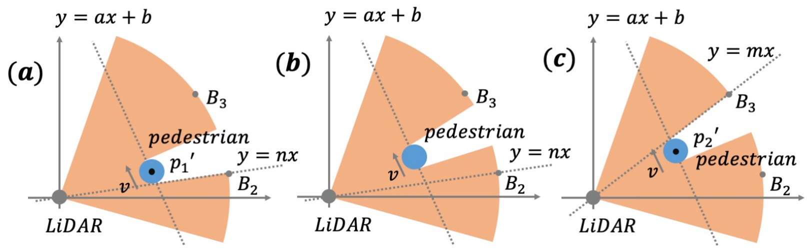


Fig. 2: Relationship between a pedestrian and two background points

Since the distance can be derived from the two coordinates, the occlusion time of a background point t are derived from the following equations.

$$t = \frac{2}{v}\sqrt{\frac{(a^2+1)(m^2+1)r^2}{(a-m)^2}} \tag{5}$$

We can see that the occlusion time t depends on only a and v according to Eq. (5).

Next, we explain how to calculate traveling time between two continuous background points. Fig.2 shows a geographic relationship between a pedestrian and two background points. B_2 and B_3 are background points on lines $y = nx$ and $y = mx$, respectively. Fig.2-(a) shows when a pedestrian stops hiding B_2. p'_1 is the coordinate of the pedestrian and can be calculated by the following equations.

$$\begin{cases} y = nx + r\sqrt{1+n^2} & \text{(6a)} \\ y = ax + b & \text{(6b)} \end{cases}$$

$$p'_1 = \left(\frac{b - r\sqrt{1+n^2}}{n-a}, \frac{nb - ar\sqrt{1+n^2}}{n-a}\right) \tag{7}$$

Fig.2-(b) shows when a pedestrian moves between B_2 and B_3. Among this time, B_2 and B_3 are not hidden by the pedestrian. After that, B_3 cannot be observed from the sensor. p'_2 is the coordinate of the pedestrian and can be calculated by the following equations.

$$\begin{cases} y = mx - r\sqrt{1+m^2} & \text{(8a)} \\ y = ax + b & \text{(8b)} \end{cases}$$

$$p'_2 = \left(\frac{b + r\sqrt{1+m^2}}{m-a}, \frac{mb + ar\sqrt{1+m^2}}{m-a}\right) \tag{9}$$

Since the distance can be derived from the two coordinates, the traveling time t' is also calculated by the following equations.

$$t' = \frac{\sqrt{dx^2 + dy^2}}{v} \tag{10}$$

$$dx = \frac{b + r\sqrt{1+m^2}}{m-a} - \frac{b - r\sqrt{1+n^2}}{n-a} \tag{11}$$

$$dy = \frac{mb + ar\sqrt{1+m^2}}{m-a} - \frac{nb - ar\sqrt{1+n^2}}{n-a} \tag{12}$$

As shown in Eq. (10), the traveling time of two background point t' can be represented by a combination of a, b and v. This relationship shows that three occlusions of background points are necessary to estimate the trajectory of a pedestrian. In the following section, we estimate pedestrian trajectories based on this relationship.

4 Trajectory Estimation

In section 3, we explain how to estimate a pedestrian trajectory based on a set of three occlusions of background points that occurred by the same pedestrian. We represent a set of N pedestrians as $P = \{P_1, P_2, ...P_N\}$. A pedestrian P_n moves along $tra^{n,ts}$: $y = a^{n,ts}x + b^{n,ts}$ with velocity $v^{n,ts}$ in a time slot ts. Our goal is to estimate a pedestrian trajectory for pedestrian P_n as $a_o^{n,ts}$, $b_o^{n,ts}$ and $v_o^{n,ts}$ which represent a small line segment in a time slot ts. We deploy a LiDAR sensor and assume that it can measure M background points $B = \{B_1, B_2, ..., B_M\}$ per θ degree to observe occlusions caused by pedestrians. The LiDAR sensors can know that there is an occlusion when the measured distance becomes shorter. An observed occlusion duration in a time slot ts is represented as $O_{o_m}^{m,ts}$ at a background point B_m. We also represent an interval time between two occlusions as I_{o_i,o_j}. At first, we find occlusions of background points occurred by the same pedestrian in a time slot ts . Although the sensor can measure distances and know that occlusions occur, we do not know who cross the background points and cause the occlusions. Also, some pedestrians hide the same background point together, the occlusion duration includes two or more pedestrians' crossing and cannot be used to estimate trajectories. In order to know who causes the occlusions and eliminate such overlapped occlusions, we calculate a set of parameters $\hat{a}$, $\hat{b}$ and $\hat{v}$ for a combination of three occlusions. If the three occlusions are occurred by the same pedestrian, $\hat{a}$, $\hat{b}$ and $\hat{v}$ should be the same. Based on this characteristic, we can find a combination of occlusions caused by the same pedestrian P_n. A set of parameters $\hat{a}$, $\hat{b}$ and $\hat{v}$ also means a trajectory and is derived as an output a_o^{ts}, b_o^{ts} and v_o^{ts} in the time slot ts.

In order to estimate efficiently, we calculate $\hat{a}$ and $\hat{v}$ first instead of calculating three parameters at once. We calculate $\hat{a}$ and $\hat{v}$ for all pairs of observed occlusions of two background points since we cannot know which pedestrians cause occlusions. $\hat{a}$ and $\hat{v}$ can be calculated for different two occlusions $O_{o_i}^{i,ts}$ and $O_{o_j}^{j,ts}$ occurred in different background points B_i and B_j from by the following equations derived Eq. (5).

$$\begin{cases} O_{o_i}^{i,ts} = \dfrac{2}{\hat{v}}\sqrt{\dfrac{(\hat{a}^2+1)(m_i^2+1)r^2}{(\hat{a}-m_i)^2}} & \text{(13a)} \\ O_{o_j}^{j,ts} = \dfrac{2}{\hat{v}}\sqrt{\dfrac{(\hat{a}^2+1)(m_j^2+1)r^2}{\left(\hat{a}-m_j\right)^2}} & \text{(13b)} \end{cases}$$

where m_i, m_j are the gradients of lines from the sensor to each background point. This calculation is applied to all pairs of observed occlusions of two background points $\forall O_{o_i}^{B_i,ts}, O_{o_j}^{B_j,ts} (1 \leq i, j \leq M)$ and derives a set of $\hat{a}$ and $\hat{v}$ in the time slot ts.

Next, we calculate $\hat{b}$ based on $\hat{a}$ and $\hat{v}$. $\hat{b}$ can be calculated from two occlusions by Equations (14). If three occlusions $O_{o_i}^{i,ts}$, $O_{o_j}^{j,ts}$ and $O_{o_k}^{k,ts}$ are occurred by the same pedestrian, three values calculated from a pair of two different occlusions as $\hat{b}$ should be the same. At first, we find a set of two occlusions that derive the same pair of $\hat{a}$ and $\hat{v}$ and calculate $\hat{b}$ from three different occlusions among the set. If the values are the same, we can know that the three occlusions are occurred by the same pedestrians and

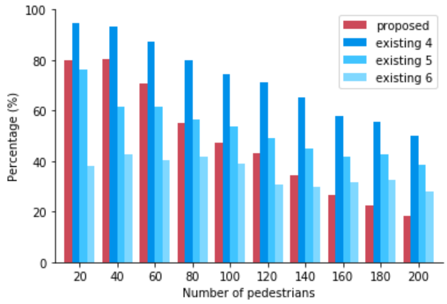


Fig. 3: Percentage of Trajectories

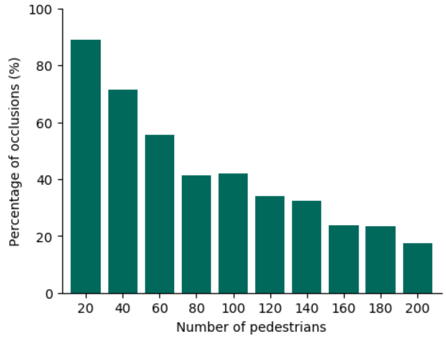


Fig. 4: Percentage of occlusions caused by one pedestrian

its trajectory a_o^{ts}, b_o^{ts} and v_o^{ts} can be obtained as calculated $\hat{a}$, $\hat{b}$ and $\hat{v}$.

$$I_{o_i,o_j} = \frac{\sqrt{dx^2 + dy^2}}{\hat{v}} \tag{14}$$

$$dx = \frac{\hat{b} + r\sqrt{1 + m_j^2}}{m_j - \hat{a}} - \frac{\hat{b} - r\sqrt{1 + m_i{}^2}}{m_i - \hat{a}} \tag{15}$$

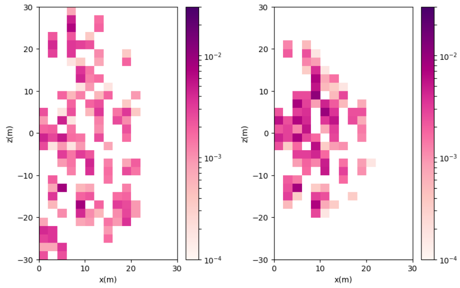
$$dy = \frac{m_j\hat{b} + \hat{a}r\sqrt{1 + m_j^2}}{m_j - \hat{a}} - \frac{m_i\hat{b} - \hat{a}r\sqrt{1 + m_i{}^2}}{m_i - \hat{a}} \tag{16}$$

where m_i, m_j are the gradients of lines from the sensor to each background point.

5 Experiments

We conducted several experiments to evaluate our proposed method through simulations. We deploy one LiDAR sensor that can measure 720 distances equally in 180 degrees on the sensing area in the simulations. The sensor can measure distances up to 30 meters. We also assume that half of the pedestrians move left to right and the other half of the pedestrians move right to left with average velocity 1 m/s in the front of the sensor. Our method can estimate a trajectory only for a pedestrian who causes three more occlusions within a time slot (5 seconds in the experiments). On the other hand, we modeled that an existing method constructs a point cloud and derives a position for each pedestrian. Since the method requires several foreground points to detect features of human bodies, we assume that the existing method can derive a trajectory for a pedestrian if the number of observed foreground points is more than a threshold for 1 second in a time slot. We compare the proposed method with the existing method with a different number of pedestrians.

Fig.3 shows the average percentages of estimated trajectories in 10 time slots. The number of pedestrians is varied from 20 to 200. Red bars show the percentages of our proposed method and blue bars show the percentage of the existing method with different thresholds. For example, ”existing 4” means that the existing method requires 4 foreground points to detect a pedestrian. As shown in Fig.3, the existing method with 4 points achieved the highest performance in all cases. However, it might contain the wrong detection because the threshold is small for a human body. Our method can derive more trajectories than the existing method with 5 points and 6 points when the number of pedestrians is fewer than 60. The results of our method and the existing



(a) Proposed **(b)** Existing5

Fig. 5: Locations of Estimated Pedestrians

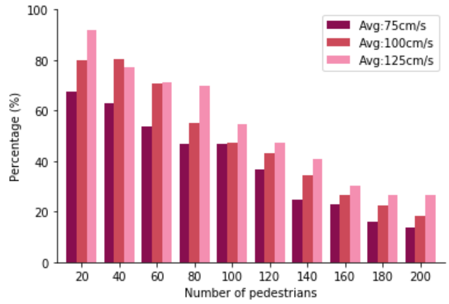


Fig. 6: Percentage of Trajectories

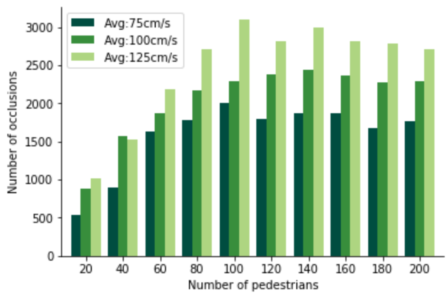


Fig. 7: Percentage of occlusions caused by one pedestrian

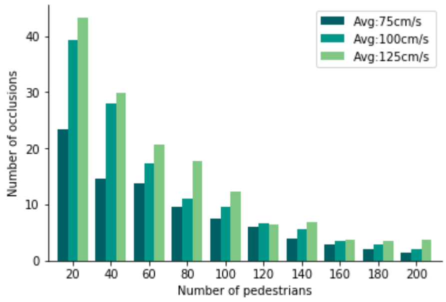


Fig. 8: Average number of occlusions caused by one pedestrian per pedestrian

method with 5 points are almost the same when the number of pedestrians is 80. When the number of pedestrians is more than 160, the results of our method become lower than that of 6 points. We can see that our method is more robust than the existing methods when the density is moderately high. However, both of the methods cannot derive pedestrians in crowded situations because point could not be constructed for almost people for existing methods and occlusions caused by one pedestrian cannot be observed for the proposed method. To investigate this in detail, we also calculated the percentage of occlusions derived from one pedestrian to all occlusions (Fig.4). Fig.4 shows that the number of occlusions caused by one pedestrian becomes smaller when the number of pedestrians increases and the density of the pedestrians increases. Since the occlusions derived from one pedestrian are only 20 percent of all occlusions in the case of 200 pedestrians, the theoretical best performance of the method is also expected to be about 20 percent since the performances of our method follow this trend. Therefore, the performance of the proposed method is reasonable even though the performances of the method are lower than that of the existing methods. It also means that we have to utilized occlusions overlapped by two and more pedestrians to increase estimation accuracy in the proposed scheme.

Also, we investigated the positions of estimated pedestrians. Figures 5 show the positions of estimated pedestrians when the number of pedestrians is 120 in the sensing area as heat maps. The sensor is located at $(0, 0)$. We can see that the existing method based on point clouds can detect only the position of pedestrians within about 20m radius from the sensor. On the other hand, our proposed could detect pedestrians at a distance of more than 24m from the sensor.

Furthermore, we have another experiment in order to investigate the effect of velocities. The average velocities are varied from 75cm/s to 125cm/s. The average percentages of the number of estimated pedestrians show shown in Fig. 6. The lighter color bars show the results of our proposed method with faster velocity. The results of the fastest velocity are better than that of smaller velocities in almost all cases. Fig. 7 and Fig.8 also show the number of all occlusions in these cases and the average number of occlusions caused by one pedestrian per pedestrian, respectively. According to these figures, more occlusions are observed with larger velocities because pedestrians cross many background points and generate many occlusions in a time slot in those cases. Thus, the performances with faster velocities are better then that of slower velocities. From Fig. 7, we can see that the number of pedestrians and the number of all occlusions increase until 100 pedestrians and become stable in the cases from 120 pedestrians to 200 pedestrians. Although there are many occlusions, the number of occlusions overlapped by two or more pedestrians increases and the number of occlusions caused by one pedestrian decreases as shown in Fig.8. Therefore, the performances of the proposed method with different velocities decrease because the proposed method needs more than 3 occlusions derived from the same pedestrian to estimate trajectories.

From above results, although the performance of the proposed method is not good in crowded situations as well as the existing methods, we can see that our method has different characteristic and can detect pedestrians in different situations such as density, positions and so on. We think that it is better to combine two different schemes to achieve better estimation accuracy.

6 Conclusion

We have proposed an alternative scheme to derive pedestrian trajectories using a part of data from LiDAR sensors. The method measures time not to observe the background points and estimates pedestrian trajectories from the occlusions by simple mathematical calculations on a part of laser points instead of whole laser points. The method also aggregates occlusions that occurred by the same pedestrian to derive its trajectory. Since the proposed method detects pedestrians based on three occlusions for a pedestrian, the proposed method has different characteristics. The simulation results showed that the proposed method can detect more pedestrians located far from sensors and in rather crowded situations compared with an existing scheme based in point clouds.

Acknowledgement

This work is supported in part by a I-O DATA Foundation KA-015.

References

1. Cui, J., Zha, H., Zhao, H., Shibasaki, R.: Robust tracking of multiple people in crowds using laser range scanners. In: 18th International Conference on Pattern Recognition (ICPR'06). vol. 4, pp. 857–860 (Aug 2006). https://doi.org/10.1109/ICPR.2006.1017
2. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: An evaluation of the state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence **34**(4), 743–761 (April 2012). https://doi.org/10.1109/TPAMI.2011.155
3. Fod, A., Howard, A., Mataric, M.A.J.: A laser-based people tracker. In: Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292). vol. 3, pp. 3024–3029 vol.3 (May 2002). https://doi.org/10.1109/ROBOT.2002.1013691
4. Zhang, Y., Zhou, D., Chen, S., Gao, S., Ma, Y.: Single-image crowd counting via multi-column convolutional neural network. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 589–597 (June 2016). https://doi.org/10.1109/CVPR.2016.70
5. Zhao, H., Shibasaki, R.: A novel system for tracking pedestrians using multiple single-row laser-range scanners. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans **35**(2), 283–291 (March 2005). https://doi.org/10.1109/TSMCA.2005.843396