

A Pedestrian Detection and Tracking Method for Robot Equipped with Laser Radar

Zhu Bin¹, Zhang Jian Rong², Wang Yan Fang³, Wu Jin Ping⁴

AI research center, School of Mechanical and Engineering
Jiangxi college of applied technology
Ganzhou, China

Abstract—In order to detect and track pedestrians in complex indoor backgrounds, a pedestrian detection and tracking method for indoor robots equipped with Laser radar is proposed. Firstly, The SLAM (Simultaneous Location and Mapping) technology is applied to obtain 2D grid map for a strange environment; then, Monte Carlo localization is employed to obtain the posterior pose of the robot in the map; then, an improved likelihood field background subtraction algorithm is proposed to extract the interesting foreground in changeable environment; then, the hierarchical clustering algorithm combining with an improved leg model is proposed to detect the objective pedestrian; at last, an improved tracking intensity formula is designed to track and follow the objective pedestrian. Experimental results in some complex environments show that our method can effectively reduce the impact of confusing scenarios which are challenges for other algorithms, such as the motion of the chair, the suddenly passing by person and when the objective pedestrian close to the wall and so on, and can detect, track and follow pedestrians in real time with high accuracy.

Keywords—Laser radar; likelihood field model; pedestrian detection; pedestrian tracking; simultaneous location and mapping

I. INTRODUCTION

Under the fast development of artificial intelligence, the research and application of intelligent service robot attracts more and more scholars and researcher's attention, and the application fields of robots cover all aspects of human life, such as restaurant service, Home-based services, shopping guide, accompanying dance and so on [1]. The main work of these robots is to interact with people, so the detection, tracking and avoidance of pedestrians are particularly important.

At present, pedestrian detection and tracking methods mainly depend on ordinary camera [2-3], infrared camera [4-7], laser radar and so on. Because the ordinary camera based method are easy to be affected by light, field size, shooting angle and so on, so these kinds of method are only suitable for the application scene with stable lighting and relatively fixed shooting position. The infrared camera based method can be employed in large workplace and is not affected by visible light and these methods can work normally even in dark environment. However, due to the technical limitation, many drawbacks exist, such as low signal-to-noise ratio, low contrast, low resolution, and the cost is high. Because the laser radar based methods have the advantage of high precision and low cost and is not affected by illumination changes, it is widely used. For the consideration of performance and price, 2D laser radar is always suitable. However, two-dimensional

plane information does not include the depth of the image and what's more, the longer the detection distance, the lower the resolution. Authors in [8-9] proposed 14 features of human leg, and used AdaBoost strong classifier to detect pedestrians. Authors in [10-11] proposed improved convolutional neural network to classify wheelchairs and human legs. However, in complex background, due to the lack of two-dimensional information, the classifier would output wrong results, such as the chair leg near the corner would be misjudged as a human leg. In order to reduce the interference of complex background, [12] proposed a background subtraction method, which reduces the interference of background in fixed laser radar and fixed scene, but can't be applied to mobile robot. Author in [13] firstly generate the local grid map in the environment, then match and align the grid map of the front and back frames to get the undetermined foreground. At the same time, assuming that the human leg corresponds to the minimum value of the laser radar distance histogram, the final foreground can be obtained by the operation of the laser radar distance histogram of the undetermined foreground and the laser radar distance histogram of the human lag. This method can be well applied in the relatively open environment such as corridor, but it can't deal with the situation that pedestrian is still and environment is complex.

In order to solve the problem of background interference in pedestrian detection and tracking, this paper propose a newly method, first, the environment map is constructed, and then the likelihood domain model is used to segment the foreground from the background; at last, the improved Kalman filter is used to track and follow the pedestrian in the complex background.

II. FLOWCHART OF OUR METHOD

As shown in Fig. 1, after the SLAM technology is applied to construct the environment map, the Monte Carlo localization is applied to determine the location of the robot in the map. Then, the foreground is extracted by likely likelihood domain model, the data cluster technology is employed to generate the steady background, which would enhance the accuracy of the foreground extraction. Then the foreground is judged whether the objective pedestrian exist, if the objective pedestrian is localized, then the robot will follow the pedestrian automatically. Otherwise, the robot will standstill or cruise randomly and waiting for the result of the next frame.

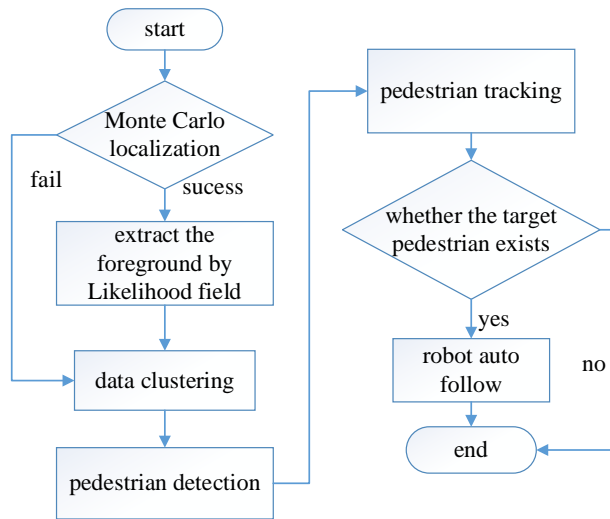


Fig. 1. The Flowchart of the Method.

III. LIKELIHOOD FIELD BACKGROUND DIFFERENCE

A. The Map Construction

The SLAM technology of mobile robot is first proposed by R.Smith, M.Self and P.Cheeseman [14], which locate the robot's pose when the robot move in an unknown environment, and then build a map according to the robot's pose, so as to achieve the purpose of Self-orientation and mapping at the same time. Grisetti G et al. [15] improved the traditional SLAM by rao-blackwellized particle filter, and proposed adaptive sampling to reduce particle loss, forming the current GMapping algorithm. GMapping is used to build environment map, because the method has the advantages of low requirement for the performance of the laser radar, low calculation consumption and high accuracy of the mapping.

Fig. 2 is the grid map of the experimental scene using our own device and GMapping technology. In the figure, the white part indicates that the robot can move in these areas; the black part indicates that the robot cannot move in these areas; the gray part means area still not be detected; numbers represent the position of where to operate background difference experiment.

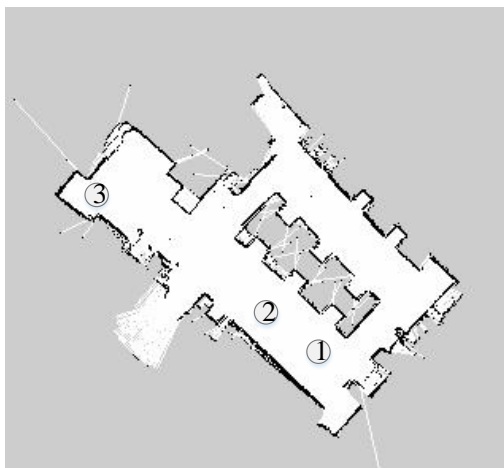


Fig. 2. The Grid Map of the Experimental Scene.

B. Monte Carlo Localization

Monte Carlo localization [16] is an algorithm which can be used to determine the position and direction for a robot in the grid map using odometer information and laser radar data. The algorithm first initializes a particle swarm in normal distribution using to standard mean and variance, then updates the pose of all particles in the particle swarm by the odometer data and the motion model, then obtains the importance weight of the particles by calculating the correspondence between the laser radar data and the map under the corresponding pose, and finally the maximum possibility rule is applied to resample the particle swarm, and the pose with the largest weight is treated as a posteriori pose. What's more, the random particles are usually added into the resampling step to recover the robot from global positioning failure and local optimal solution.

C. Likelihood Field Model

The likelihood field model is first applied to eliminate the uncertainty of the signal obtained by various sensors, the possibility of the value of signal intensity obtained by various sensors is employed to determine the final value of the signal, doing so, the output of the sensor would be more robust to the influence of the noise and the fluctuation of the environmental factors such as voltage, temperate, humidity and so on. In this paper, the likelihood field model is used to obtain a steady background, then under the circumstance that a constructed grid map and the position and direction of the robot is determined, the foreground can be extracted accurately.

The likelihood field model can be represented by a conditional probability distribution $p(z_t | x_t, m)$:

$$p(z_t | x_t, m) = \prod_{k=1}^K p(z_t^k | x_t, m) \quad (1)$$

Where, z_t is the measurement value at time t , $x_t = (x \ y \ \theta)^T$ is the pose of the robot, and m is the environmental map. Suppose that K measurement points are available for the data obtained by the laser radar sensor, and each measurement value obey noise independent distribution.

Where, k is the number of the measurement point and z_t^k is the measurement value of the k -th measurement point at time t .

But traditional likelihood field model do not consider the influence of the noise and the uncertainty of the environment in the construction of the background, to solve this problem, improved likelihood field model is proposed:

$$p(z_t^k | x_t, m) = z_{hit} P_{hit} + z_{rand} P_{rand} + P_{original} \quad (2)$$

$$P_{hit}(z_t^k | x_t, m) = e_{s_{hit}}(dist) \quad (3)$$

$$P_{rand}(z_t^k | x_t, m) = \frac{1}{z_{max}} \quad (4)$$

Where, $P_{original}$ represents the original likelihood value, it can be calculated using formula (1). P_{hit} is the likelihood value affected by noise, $dist$ is the Euclidean distance between the coordinate and the nearest obstacle on map m . It is considered that P_{hit} obeys the Gaussian distribution with the mean equal to 0 and variance equal to 1. P_{rand} is the likelihood value affected by objects randomly showing up. It is considered that the likelihood of obstacles detected by the measurement points obey an average distribution, which z_{max} is the maximum measurement distance. Based on the above two considerations, the likelihood of the objects detected in the map by the corresponding sensor, and z_{hit} , z_{rand} is the weight, respectively. Fig. 3 shows the relationship between likelihood and dist.

The value of likelihood represents the possibility of a surrounding point of the measurement point is a background, and it is related to the shortest distance $dist$ from the measurement point to the obstacle in the map. In this paper, a fixed threshold value $theta - pk$ is set. For each $p(z_t^k / x_t, m)$, when it is bigger than the threshold value, the scanned region is the background. When the calculated likelihood value is smaller than the threshold, it is considered that the scanned region include foreground. $theta - pk$ equal to 0.5, and the $dist$ equal to 0.12 in our experiment.

$$\begin{pmatrix} x_{z_t^k} \\ y_{z_t^k} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_{k,sens} \\ y_{k,sens} \end{pmatrix} + z_t^k \begin{pmatrix} \cos(\theta + \theta_{k,sens}) \\ \sin(\theta + \theta_{k,sens}) \end{pmatrix} \quad (5)$$

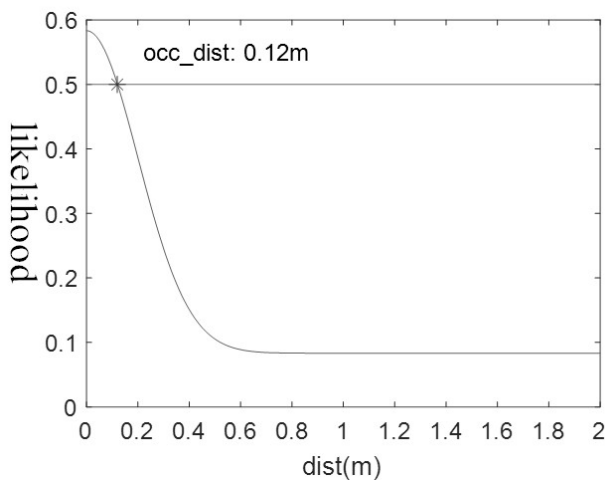


Fig. 3. The Relationship between Likelihood and Dist.

When the coordinates for each measurement point are obtained, the transformation for the world coordinate into the robot coordinate is needed. Where $(x_{k,sens} \ y_{k,sens})^T$ represents the position of laser radar sensor in robot coordinate system and $\theta_{k,sens}$ is the angle of the measurement point.

IV. PEDESTRIAN DETECTION AND TRACKING

A. Data Clustering

After obtained the foreground information taken from laser radar data, the clustering operation is then followed, the foreground is then classified into several class, and then is used to recognize. The state-of-the-art algorithms can be departed into three classes, they are hierarchical clustering algorithm, K-means clustering algorithm and DBSCAN clustering algorithm [13]. K-means clustering algorithm needs to know the number of clusters in advance, which is not suitable for the situation of unknown pedestrian and environment. DBSCAN is a density-based algorithm, which calculate the number of elements in a circle with a certain radius. As long as the number of elements exceeds the preset threshold, these elements in the circle are regarded as a class. The effect of DBSCAN algorithm depends on the radius of circle and the number of elements, and the calculation is huge. The hierarchical clustering algorithm first treats each data to be processed as a class, calculates the distance between classes, then compares the distance between adjacent classes, and merge points whose distance is less than the preset threshold value into a class.

The hierarchical clustering algorithm is selected in this paper, which does not need to specify the number of clustering results and has few restrictions. In traditional hierarchical clustering algorithm, the singular value would negatively affect the clustering results. To solve this problem, the laser radar foreground data is preprocessed firstly, directly removes the singular points which is far from the laser radar, and then conducts hierarchical clustering.

B. Pedestrian Detection

Although there are many extracted features related to human legs, [8] pointed out that only a few features will occupy a relatively large weight in the final trained AdaBoost detector, and most of the other features will have less weight. Adding these features may lead to over fitting.

On the basis of [8], the leg model is designed and can be seen in Fig. 4, in which two side points represent the cluster of each leg and center point represent the location of the pedestrian. D_k represents the distance between the head and tail elements of the corresponding cluster, and L_k represents the total length of the corresponding cluster. Finally, the detector gives the middle position of the pedestrian leg model as the pedestrian coordinate. The specific process of the pedestrian detector is shown in Fig. 5.

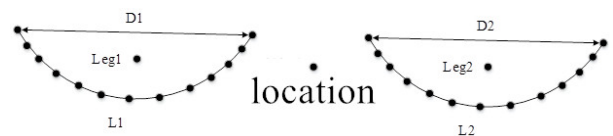


Fig. 4. The Leg Model used in our Paper.

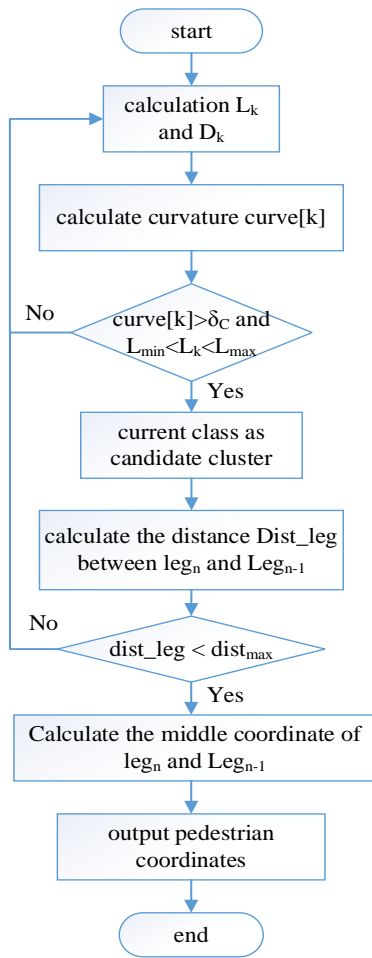


Fig. 5. The Flowchart of Pedestrian Detection.

C. Pedestrian Tracking

In order to solve the problem of one of leg is occluded temporarily, the pedestrian tracking algorithm as mentioned in [17-18] is applied, and rectified the tracking intensity formula.

$$cf_k = \begin{cases} cf_{k-1} + \log_{1.5}(cfi_k), & cfd = 0 \\ cf_{k-1} - 1.5^{cfd_k}, & cfi = 0, hide = 0 \\ cf_{k-1} - \log_{1.5}(cfd_k), & cfi = 0, hide = 1 \end{cases} \quad (6)$$

When a pedestrian suddenly lose in a scene, the history tracking information is used to judge whether the losing is caused by one of leg is occluded, if yes, the pedestrian is just occluded temporarily. In the actual experiment, it is found that the improved algorithm can effectively track multiple pedestrians, and robust in circumstances such as a pedestrian suddenly break in or leave or be occluded.

D. Automatically Following

The pedestrian is followed within 30cm in front of the robot as the following target until the tracking strength of the corresponding tracker is reduced to a certain extent or the following is given up artificially. After obtain the global coordinate of the objective pedestrian (x_p, y_p) and the robot

coordinate (x_r, y_r) , the next robot posture (x_g, y_g, yaw_g) is determined by direction of the line between the robot and the pedestrian, then the posture is sent to ROS (Robot Operating System) platform, and using the navigation function to complete the automatically following.

$$drp = \sqrt{(x_p - x_r)^2 + (y_p - y_r)^2}$$

$$drt = drp - 0.2$$

$$x_g = x_r + drt * ((x_p - x_r) / drp)$$

$$y_g = y_r + drt * ((y_p - y_r) / drp) \quad (7)$$

$$yaw_g = \arctan((y_p - y_r), (x_p - x_r)) \quad (8)$$

V. EXPERIMENTAL RESULTS AND ANALYSIS

The platform used in our experiment as can be seen in Fig. 6, is a wheel robot, which equipped with a Flash Lidar F4 laser radar which the scanning angle is 360, the angle resolution is 0.5 and the frame rate is 10FPS, two wheels with distance encoder, an industrial computer with Intel i7-3610 and using C++ in ROS to realize the algorithm.



Fig. 6. Experimental Platform.

A. The Experiment of Likelihood Field Background difference

The experimental results are shown in Fig. 7. When the robot is in the three locations as shown in Fig. 2. The black outline in the figure is the original obstacle of the map, the white fine points on the obstacle are the laser radar background data separated from the likelihood field, the points with thick white and black edges are the laser radar foreground data extracted from the likelihood field method, the black circle in the middle is the location of the robot, and it can be seen that the location of the objective pedestrian can be figured out in each image.

Using the method in [19], when the region detected by laser radar close to the obstacle in background, they would be treated as a background. Only when the region is far away from the background, they would be regarded as a foreground. But in Fig. 7, which applies our method, the foreground and the background can be classified correctly.

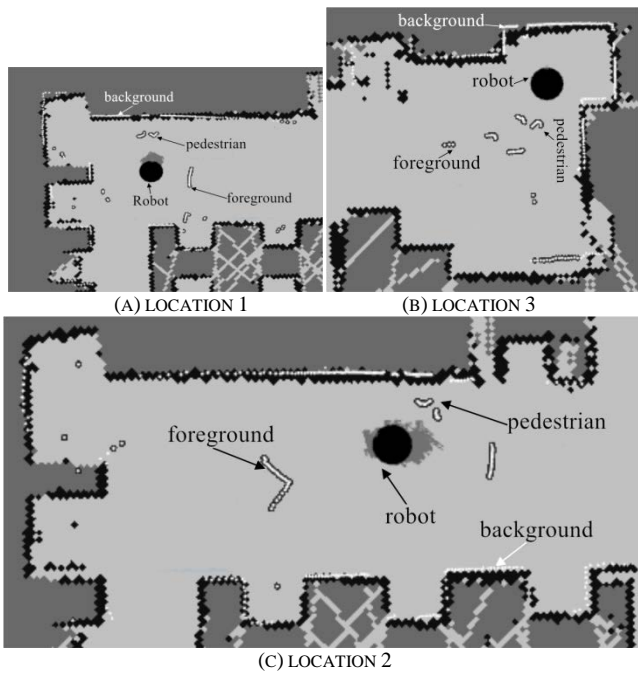


Fig. 7. The Result of Likelihood Field Background difference.

B. The Anti Interference Experiment

The method of [20] is used to pedestrian detection experiment. Compared the method with likelihood field background difference and the original method, it can be seen that our method can require better results in complex environments. Fig. 9 shows the experimental results, in which (a) is the actual scene, (b) is the result of our method (ours) and (c) is the result of original method (original).

Three scenes are used to test the algorithm, Fig. 8(a1) and Fig. 8(a3) is the circumstance that the pedestrian would easily be mistreated as background, Fig. 8(a2) is the circumstance that the leg of the chair would be mistreated as pedestrian. it can be seen in Fig. 8(b1), Fig. 8(b2), Fig. 8(b3), using original method, three wrong judgments are received in these three scenes. When using our method, we can receive the correct pedestrian detection results, in Fig. 8(c1) and Fig. 8(c3), the accurate location of the pedestrian is detected, and in Fig. 8(c2), the leg of the chair is not mistreated as pedestrian.

Experimental results show that after using likelihood field background difference, background and foreground can be classified correctly in complex environment, so as to reduce the interference of the complex environment to pedestrian detection.

C. The Experiment of Pedestrian Detection

The pedestrian experiments are done in two locations and with two different detection distances (50cm or 50-100cm). The experimental results can be seen in Fig. 9. The method can always detect the pedestrian in different locations and distances, which demonstrate the practicability of our method.

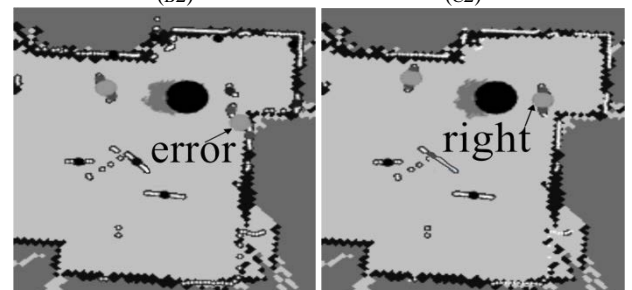
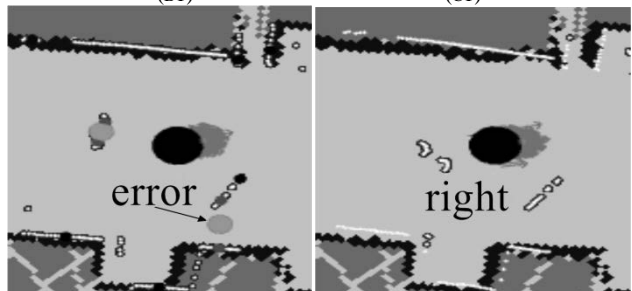
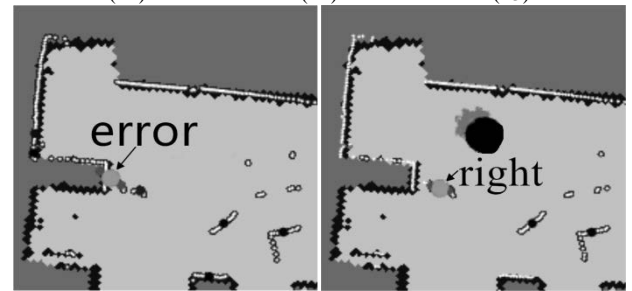


Fig. 8. The Result of Pedestrian Detection.

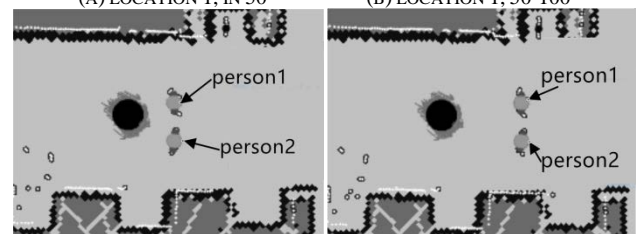
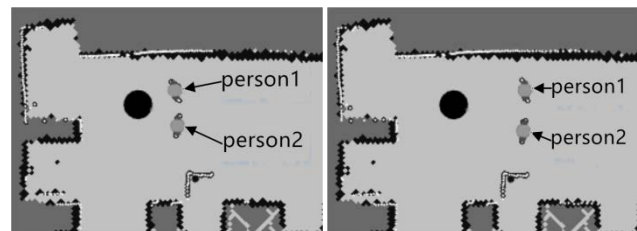


Fig. 9. The Result of Pedestrian Detection.

D. The Experiment of Pedestrian Tracking and Following

Our method is then used to pedestrian tracking and following, the experimental result can be seen in Fig. 10.

As shown in Fig. 10, black Pentagram and black circle represents the start and end points of the path of particle of the robot and the target pedestrian. The black thick line between the start point and the end point is the trajectory of the robot, and the white thick line is the trajectory of the target pedestrian. In order to test the stability of the proposed detection algorithm and tracking algorithm, a non-target pedestrian showed up, first move parallel to the target pedestrian, and finally surpass the target pedestrian and run counter to the target pedestrian. The track of non-target pedestrian is marked by white arrow, and the intermittent gray thick line represents the track of non-target pedestrian.

It can be seen that the trajectory of the robot and the target pedestrian basically coincides, and the emergence of non-target pedestrian does not affect the robot's follow-up to the target pedestrian. Therefore, the follow-up strategy in this paper can enable the robot to eliminate interference and track the target pedestrian continuously and stably.



Fig. 10. The Result of Pedestrian Tracking and Following.

VI. CONCLUSION

This paper proposes a pedestrian detection, tracking and following algorithm in complex environment using a mobile robot with laser radar. The algorithm firstly maps the environment and then extracts the foreground data by the likelihood domain model to reduce the interference of complex background, the hierarchical clustering algorithm is used to cluster the foreground data, and then the improved Kalman tracking algorithm is used to effectively track the multi pedestrians. Finally, the automatic tracking strategy proposed in this paper is used to effectively follow the target pedestrians in the known map environment. Experiment result shows that the whole system has high real-time performance and is not interfered by complex background, and has certain practical value. The future work would focus on the combination of the laser radar and machine vision.

REFERENCES

[1] S. Liu, L. Song, M. Gen, et al. "Indoor Pedestrian Tracking Method of Dancing Robot Based on Laser Radar," *Computer Engineering*, vol. 43, no. 6, pp. 247-252, 2017.

[2] G. Jiachen, X. Juan, Z. Hongfu, F. Hang, Z. Zhirong and X. Xiangqian, "Civil aircraft surface defects detection based on histogram of oriented gradient," 2019 IEEE 1st International Conference on Civil Aviation

Safety and Information Technology (ICCSIT), Kunming, China, 2019, pp. 34-38.

[3] P. Sombatpiboonporn, T. Charoenpong, A. Supasuteekul, C. Chianrabutra and K. Pattanaworapan, "Human Edge Segmentation From 2D Images By Histogram of Oriented Gradients and Edge Matching Algorithm," 2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP), Bangkok, Thailand, 2019, pp. 29-32.

[4] Wang G H, Liu Q, Zhuang J J. "Research on vehicle Infrared Pedestrian Detection method based on Local Features", *Acta Electronica Sinica*, 2015, 43(7): 1444-1448.

[5] Zhu C C, Xiang Z Y. "Infrared pedestrian detection based on gradient direction and intensity histogram", *Computer engineering*, vol. 40, no. 12, pp. 195-198, 2014.

[6] L. Li, F. Zhou, Y. Zheng and X. Bai, "Reconstructed Saliency for Infrared Pedestrian Images," in *IEEE Access*, vol. 7, pp. 42652-42663, 2019.

[7] T. Y. Han and B. C. Song, "Night vision pedestrian detection based on adaptive preprocessing using near infrared camera," 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), Seoul, 2016, pp. 1-3.

[8] Beyer L, Hermans A, Leibe B. "DROW: Real-Time Deep Learning based Wheelchair Detection in 2D Range Data", *IEEE Robotics and Automation Letters*, vol. 2, pp. 585-592, 2017.

[9] Y. Maalej, S. Sorour, A. Abdel-Rahim and M. Guizani, "Tracking 3D LIDAR Point Clouds Using Extended Kalman Filters in KITTI Driving Sequences," 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 2018, pp. 1-6.

[10] T. Tsai and C. Yao, "A Real-time Tracking Algorithm for Human Following Mobile Robot," 2018 International SoC Design Conference (ISOC), Daegu, Korea (South), 2018, pp. 78-79.

[11] Q. Ren, Q. Zhao, H. Qi and L. Li, "Real-time target tracking system for person-following robot," 2016 35th Chinese Control Conference (CCC), Chengdu, 2016, pp. 6160-6165.

[12] J. Zou, W. Yin, E. X. Wang, J. Wang and Y. Lu, "Human Motion Prediction Based on Visual Tracking," 2019 4th International Conference on Robotics and Automation Engineering (ICRAE), Singapore, Singapore, 2019, pp. 39-44.

[13] M. Shiomu, K. Shatani, T. Minato and H. Ishiguro, "How Should a Robot React Before People's Touch?: Modeling a Pre-Touch Reaction Distance for a Robot's Face," in *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3773-3780, Oct. 2018.

[14] H. Jo, H. M. Cho, S. Jo and E. Kim, "Efficient Grid-Based Rao-Blackwellized Particle Filter SLAM With Interparticle Map Sharing," in *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 2, pp. 714-724, April 2018.

[15] L. Chen, A. Yang, H. Hu and W. Naeem, "RBPF-MSIS: Toward Rao-Blackwellized Particle Filter SLAM for Autonomous Underwater Vehicle With Slow Mechanical Scanning Imaging Sonar," in *IEEE Systems Journal*, vol. 7, no. 14, pp. 3163-3170, 2019.

[16] F. Zhang, A. Cully and Y. Demiris, "Probabilistic Real-Time User Posture Tracking for Personalized Robot-Assisted Dressing," in *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 873-888, Aug. 2019.

[17] M. Hai, "A survey of big data clustering algorithms", *Computer Science*, vol. 43, pp. 380-383, 2016.

[18] K. S. S. Reddy and C. S. Bindu, "A review on density-based clustering algorithms for big data analysis," 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, 2017, pp. 123-130.

[19] H. Kivrak and H. Köse, "Social robot navigation in human-robot interactive environments: Social force model approach," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.

[20] A. Garrell, L. Garza-Elizondo, M. Villamizar, F. Herrero and A. Sanfeliu, "Aerial social force model: A new framework to accompany people using autonomous flying robots," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, 2017, pp. 7011-7017.