4C-03

Tracking Method of Multiple Users Group in a Crowd for Autonomous Guiding Robot

Bin ZHANG† Tomoaki NAKAMURA† Masahide KANEKO†

The University of Electro-Communications†

1. Introduction

Tracking multiple particular users is one of the most challenging work for an autonomous guiding robot. Especially for the robot which works in a crowd, it is more difficult to distinguish the users to be guided with the other pedestrians. However, recent research works [1] cannot track the particular users group well in a crowd. In this paper, we proposed a new tracking method which can find out the multiple users to be guided and track them robustly. The face information of the users to be guided is automatically initialized at the beginning, which is used for recognizing the particular users. The tracking system is realized by using Markov chain Monte Carlo (MCMC) [1] based particle filter.

2. Methods

The overview of our system is shown in Fig. 1.

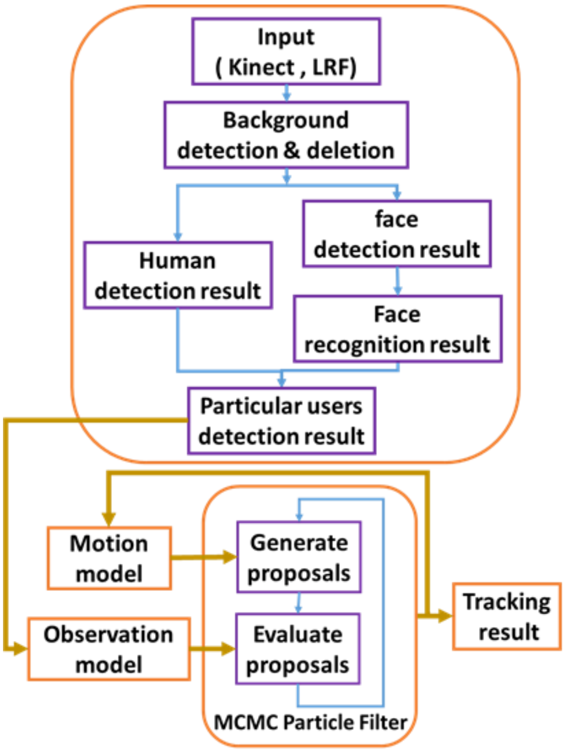

Fig. 1 Overview of the proposed system.

Tracking Method of Multiple Users Group in a Crowd for Autonomous Guiding Robot
†Bin ZHANG (UEC), Tomoaki NAKAMURA (UEC), Masahide KANEKO (UEC)

2.1 Background detection and deletion

To decrease the miss detections of humans, the background subtraction is usually used. Subtraction Stereo is also used to detect the moving object. However, these methods may not work for a mobile robot since the background keeps changing if the robot moves, or the humans may keep static. Thus, we detect out the background area in the robot's view in real time, and restrict the human detection process to the foreground area. For the indoor working environment, the walls, floor and ceiling are considered to be background. We detect out the background by using plane detection. The planes are considered to have the same normal direction, and the parallel planes can be divided by using the distance information. The shallow problem mentioned in [2] can also be solved as the floor plane is deleted before detection process.

2.2 Particular users' group detection

The humans are detected by using joint-HOG features [3] and feet ellipse fitting. The particular users group is distinguished by using Local Binary Pattern (LBP) histogram matching [4] based face recognition. Notice that the face information of the users to be guided is automatically initialized at the beginning. The face areas are detected by using Haar-like features, and only when the face regions are detected, the face recognition process works.

2.3 MCMC particle filter based people tracking

The tracking work is realized by using MCMC based particle filter. The tracking problem can be formulated as finding the maximum posterior. We set multi-people configuration as X, which represents the actual location of the people we are tracking, and the observations as Z, which represents the users' detection results. Our aim is to find the maximum posterior of $p(X_t|Z_{1:t})$ following the Equation (1).

$$p(X_t|Z_{1:t}) \propto p(Z_t|X_t) \int p(X_t|X_{t-1}) p(X_t|Z_{1:t}) dX_{t-1} \tag{1}$$

Here, $p(Z_t|X_t)$ means observation likelihood, which is calculated by the overlap ratio between human detection results and proposals; $p(X_t|X_{t-1})$ means evolution, which is designed by Equation (2); and $p(X_t|Z_{1:t})$ means the prior, which is the posterior for the previous time step.

Copyright ©2015 Information Processing Society of Japan. All Rights Reserved.

$$X_t = X_{t-1} + w \qquad (2)$$

Here, w is a process noise, and we set it as Gaussian noise. The evolution process is set like this since the people almost stay in the same place in the robot coordinate, considering the robot and the users group to be guided move in the similar speed.

3. Experiments

The robot that we used in our experiment is shown in Fig.2. The POINEER 3-DX is used as our mobile platform. Two LRF sensors (UTM-30LX) are set in front and back of the robot (height:32cm). A backward Kinect is set on the robot on the height of 70cm.

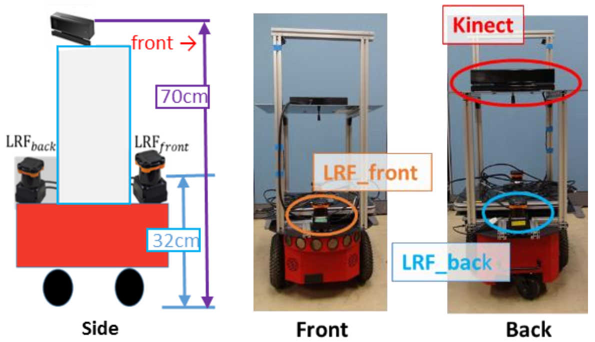

Fig.2 The guiding robot.

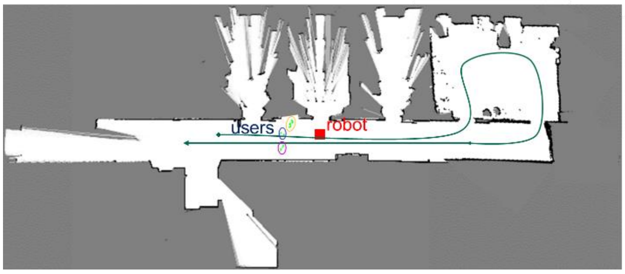

Fig.3 The global map and guiding route.

In the experiment, we guide a group with three users to get through a passage, to get into a room in a building, as shown in Fig.3. The global map is given and the route is planned beforehand by giving some sub-goal points. In each frame, we use the plane detection to detect the floor, walls and ceiling planes as the background area. The human detection is processed by combining the ellipse fitting results of two legs of human beings with the joint-HOG human detector. Our system automatically initialize the users' face information (5 face pictures detected by Haar-like features for each user), and use them for the face recognizer process. The whole users' recognition process is shown in Fig.4. The recognition results are used as the input of the observation model. The tracking results are shown in Fig.5. Three particular users were well tracked, and miss tracking occurred only when one of them disappeared for a long time. Other pedestrians are rejected by face recognition. We also calculated the users' location in the global map as shown in Fig. 3.

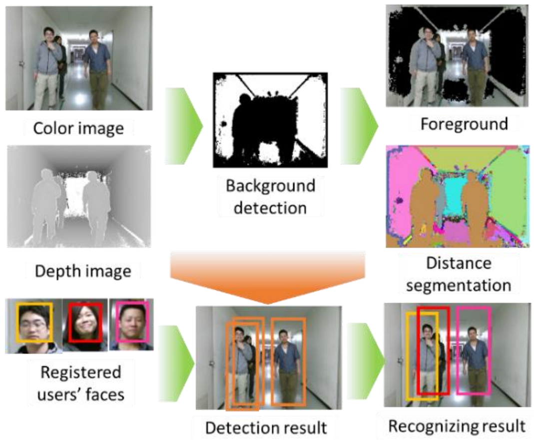

Fig.4 Detection and recognition of users' group.

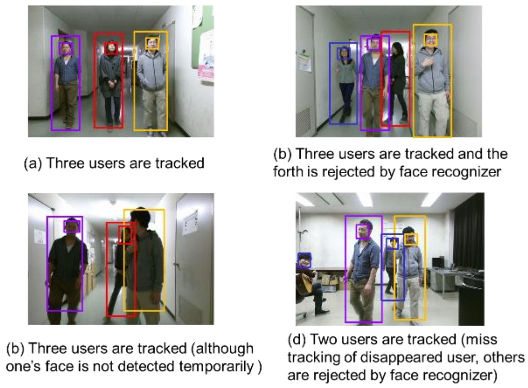

Fig.5 The tracking results of users' group.

4. Conclusion

We proposed a tracking method of multiple users group for an autonomous guiding robot in this paper, and it is proved working well in a crowded environment, although some miss recognition results exist. Future work will be focused on optimizing the algorithm to realize real time tracking of the users' group, and to improve the accuracy of particular users' detection and recognition.

References

[1] W. Choi, C. Pantofaru, and S. Savarese : "A general framework for tracking multiple people from a moving camera," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 35, no. 7, pp.1157-1591, 2013.

[2] M. Shibata, T. Ubukata, K. Terabayashi, A. Moro and K. Umeda : "Real-time human detection based on range image segmentation," Journal of the Robotics Society of Japan, vol. 32, no. 6, pp.558-568, 2014.

[3] T. Mitsui, Y. Yamauchi, and H. Fujiyosh : "Human detection by two stages adaBoost with joint HOG," IEICE Trans. Information and Systems, vol. J92-D, no. 9, pp.1951-1601, 2009.

[4] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen : "Tracking human under occlusion based on adaptive multiple kernels with projected gradients," IEEE Trans. Multimedia, vol. 15, no. 7, pp.1602-1615, 2013.

Copyright ©2015 Information Processing Society of Japan.
All Rights Reserved.