

Approach Based on Geometric Shape of Pedestrian's Head to Shoulder Region for Human Tracking in High Density Crowd Using a 3D Laser Range Finder

Yuta Sampei and Mihoko Niitsuma

Department of Precision Mechanics, Chuo University, Tokyo, Japan
(Tel: +81-3-3817-1844; E-mail: sampei.yuta.lab@gmail.com, niitsuma@mech.chuo-u.ac.jp)

Abstract – This paper discusses the possibility of using the geometric shape of a pedestrian's head to shoulder region as a feature to identify individuals. The main aim is to track humans in a high density crowd using a 3D laser range finder (3D-LRF). One of the most critical problems of human tracking in a high density crowd is the occurrence of occlusion. Occlusion might occur because of the lack of features on tracking objects and the fragmentation of trajectories. When the trajectories are fragmented, they should be associated correctly in order to realize robust and accurate tracking. To improve the robustness of human tracking in a crowded environment, we consider features of a tracking object that can recognize individuals to be useful when combining trajectories fragmented by occlusion. In this paper, the geometric shape of a pedestrian's head to shoulder region was applied as a feature of a tracking object. The applicability of the feature to the tracking in dynamic environments was done through experiments using a 3D-LRF.

Keywords – human tracking, 3D-LRF, geometrical feature

1. Introduction

Many robots work in environments where people cannot access, such as factories and disaster sites. However, in recent years, robots living with and assisting people in daily living environments, the so-called autonomous mobile robots, are being actively researched. A localization function is required as one of the basic functions for autonomous mobile robots. However, more advanced functions, such as the ones that enable autonomous mobile robots to avoid obstacles and to approach a certain person and provide the person suitable services, are also needed. To realize these advanced functions, the robustness of human tracking is crucial.

Several methods for tracking humans using a laser range finder (LRF) have been studied. However, because conventional LRFs measure distances to objects by scanning the 2D plane, it is inevitable that occlusion greatly affects tracking accuracy. Therefore, even when multiple LRFs are used [1, 2], it is difficult to perform highly accurate tracking in high density crowds where three or more people enter a common sensing area.

One of the most critical problems of human tracking in a high density crowd is the occurrence of occlusion. Occlusion might occur because of the lack of features on

tracking objects and the fragmentation of trajectories. When the trajectories are fragmented, they should be associated correctly in order to realize robust and accurate tracking. To improve the robustness of human tracking in a crowded environment, we consider features of a tracking object that can recognize individuals to be useful when combining trajectories fragmented by occlusion. The feature needs to: avoid the occurrence of occlusion; be able to be stably extracted in a high density crowd; depend only on the individual. The head-to-shoulder signature (HSS) [3] uses the shape of peoples' head to shoulder for person recognition. Because in HSS occlusion is relatively unlikely to occur, it is considered a suitable feature in a high density crowd. This method requires a sensor capable of directly measuring 3D shapes of a person. A sensor that can cover a wide area is desirable. Therefore, in this study, a 3D-LRF with a wide horizontal viewing angle was used and the possibility of using HSS as a feature for person recognition is discussed. The main aim of the paper is to use a 3D-LRF to track humans in a high density crowd.

2. Feature Extraction Method

In this section, the authors describe the person extraction method of the HSS that was considered in this study. It was used as a feature to recognize individuals aiming to track humans in a high density crowd.

2.1 Person Extraction

The person extraction method proposed by Uenoyama et al. [2] was considered. A measurement distance histogram in each horizontal direction (for every 10 cm height) is generated from the scanned range data obtained by a 3D-LRF. Based on the histogram, the greatest distance, that is more frequent than a threshold, is determined as background.

After detecting the background, moving objects are detected by subtracting the background from a new range data. Assuming that the moving objects are only pedestrians, it is possible to extract persons.

2.2 Head-to-shoulder Signature Construction

The HSS construction example is shown in Fig. 1. Firstly, we select 40 cm from the top of the extracted point cloud of a person as the head-shoulders region. The selected point cloud is divided into a number of horizontal slices of equal thickness. A span, defined as the largest

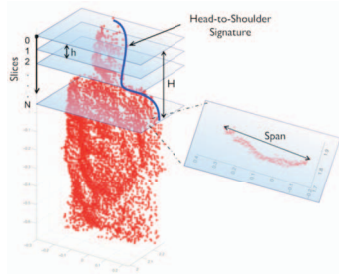


Fig. 1. HSS construction example [3].

distance that exists between any two points on each slice, is also calculated. The vertical change of the lateral width in the head-shoulder region expressed by the combination of the spans and the slice number is used as a feature to describe the person.

3. Span Measurement Experiment

The HSS was evaluated using Microsoft Kinect with limited action of subjects [3]. In this study, a 3D-LRF (UTM-30LX, Hokuyo Co. Ltd.) was used. The 3D-LRF has a wide horizontal viewing angle and a different point cloud density when compared to Microsoft Kinect. It is considered safe to use this type of sensor for humans since its light source is classified as Class 1 under IEC 60825-1. In addition, it can be applied to a wide dynamic environment.

A feature that is capable of recognizing individuals have to rely only on the individual. Therefore, this experiment evaluates the applicability of the HSS for tracking in dynamic environments using a 3D-LRF by comparing spans of a target person between resting and walking.

3.1 Experimental Conditions

The experimental environment is shown in Fig. 2. A subject faces the sensor at three different positions and rest for about 20 seconds in each position. Furthermore, the subject walks in the walking area for about 30 seconds. Considering that the region above the shoulder always enters the scanning range of the sensor during the observation, the sensor is installed horizontally 960 mm above the floor. The number of slices for HSS was set to five, taking the vertical angular resolution of the sensor into account.

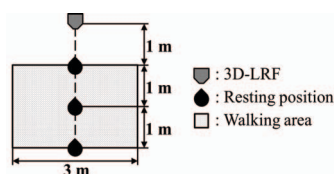


Fig. 2. Experimental environment.

3.2 Experimental Results and Discussion

The mean value and standard deviation of the measured spans of each slice were calculated. Since there was no significant difference on the measurement results for different resting positions, only the results for walking and

resting at 3m are shown in Table 1. The manually measured values were obtained using a tape measure. The slice number is set to 1 at the top and 5 at the bottom. As shown in Table 1, the variation is large when walking because the position and orientation of the body changes constantly with respect to the sensor, causing occlusion of one side of the shoulder and head. Furthermore, the variation of the slice 4 is significantly large because a mixture of cases where a shoulder is included and is not included in the slice is considered.

From the results, it became clear that the HSS depends on the relative relationship between the sensor and the human's position, and it was found that in order to use it as a feature, HSS should be calculated based on this relative relationship.

Table 1. Span Measurement results.

Slice No.	Manually measured value [mm]	Resting at 3 m [mm]	Walking [mm]
1	175.0	168.5 ± 21.4	214.6 ± 42.7
2	180.0	154.7 ± 14.4	232.9 ± 47.2
3	150.0	108.8 ± 27.1	205.6 ± 46.3
4	375.0	224.6 ± 50.0	319.0 ± 75.8
5	410.0	426.7 ± 25.7	388.8 ± 83.2

4. Conclusion

This paper discusses the possibility of using the geometric shape of a pedestrian's head to shoulder region as a feature to recognize individuals aiming to track humans in a high density crowd using a 3D-LRF. From the experiments, we found that the construction of a HSS using a 3D-LRF has a large dependency not only on the individual but also on the relative relationship between the sensor and the human's position.

In future works, the authors will propose a calculation method for HSS based on the relative relationship between the sensor and the human's position, and apply for human tracking.

References

- [1] K. Katabira, T. Suzuki, K. Nakamura, H. Zhao, R. Shibasaki, Y. Nakagawa. Tracking Pedestrians and Visualization of the Crowds-Flow using Multiple Single-Row Laser Range Scanners. *IPSJ SIG Notes. CIVM*, Vol.2007, No.31, pp.229–236, 2007.
- [2] N. Uenoyama, M. Niitsuma. Temporal Segmentation of Environmental Map Based on Changing Rate of Activity Areas and Improvement of Tracking Method using LRF. *Proc. of 2015 IEEE/SICE International Symposium on System Integration*, pp.57-62, 2015.
- [3] N. Kirchner, A. Alempijevic, A. Virgona. Head-to-shoulder signature for person recognition. *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1226-1231, 2012.
- [4] D. Brscic, T. Kanda, T. Ikeda, T. Miyashita. Person tracking in large public spaces using 3-D range sensors. *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 6, pp. 522-534, 2013.