# TRACKING PEDESTRIAN BY USING MULTIPLE LASER RANGE SCANNERS

Katsuyuki Nakamura[a], Huijing Zhao[a], Ryosuke Shibasaki[a]
Kiyoshi Sakamoto[b], Tomowo Ooga[c], Naoki Suzukawa[d]

[a] Center for Spatial Information Science, The University of Tokyo, JAPAN
[b] Frontier Service Development Laboratory, East Japan Railway Company, JAPAN
[c] Asia Air Survey Co.,Ltd, JAPAN [d] JR East Consultants Company, JAPAN
katsu@iis.u-tokyo.ac.jp

**Commision IV, WG IV/1**

**KEY WORDS:** Laser scanning, Tracking, Detection, Recognition, Acquisition, Systems, Application, Video

**ABSTRACT:**

Tracking pedestrians is an important topic in the field of pattern recognition and image understanding due to its various applications. Although video-based approaches have been studied for decades, restricted setting location, narrow viewing angle and limited resolution are the major obstacles to achieve higher accuracy and to apply to a wider variety of applications. In this paper, we propose a novel system of tracking pedestrians in a wide and open area using a network of single-row type laser range scanners. Laser range scanners are located in different places, scanning pedestrians' feet at a horizontal plane about 16cm above the ground. A pedestrians' walking model is defined, based on which a tracking algorithm is developed. An experiment is conducted in a railway station in Japan, which is used about 250,000 passengers a day. Pedestrians' trajectories are extracted and the extracted results are compared with real situations that are recorded using video cameras. It is concluded that this system has very high effectiveness.

## 1 INTRODUCTION

Human positioning data have been widely used in various fields such as architecture, disaster prevention and traffic engineering. Because of suchlike various applications, there are a lot of researches to measure pedestrians' trajectory. Great majority of these works focus on motion analysis through video image, e.g. Chia-Jung, 2004; Rosales, 2003; Ohta, 2003. On the other hands, video-based approaches have some disadvantages, such as narrow FOV (Field of View), limited resolution. Multiple cameras are always needed to cover a large area, to reduce occlusion and to solve crossing problem as well. However, it is practically difficult to digitally fuse the data of multiple cameras, which requires accurate calibration and complicated calculation between different perspective coordinate system. Therefore, only few systems have been applied to the measurement of high-density crowds in wide area, such as that of railway stations and exhibition halls.

In this research, we propose a novel system for tracking pedestrian over a wide and open area using a network of single-row type laser range scanners (briefly called laser scanner in the followings). We apply this system to measure passengers' flow in a railway station. The results are compared with real situations that are recorded using video cameras.

## 2 METHODS

In this section, we first briefly introduce the sensor system for pedestrians' tracking. A description to the tracking algorithm is given next, followed by an address to efficiency and accuracy assessment to the system.

Figure 1: Pedestrian Measurement by using Laser Scanner

### 2.1 Sensor

Single-row type laser range scanners produced by SICK corp. are employed. It measures range distances from the sensor to surrounding objects using the method of time-of-flight. Laser scanner has the advantages of direct measurement, high accuracy (average distance error is 4cm), wide viewing angle (180 degree) and long range-distance (maximum range distance of 30m). Moreover, it has high angle-resolution of 0.5 degree because of a little diffusion of laser beam. Frequency is 37.5 Hz and wave length of laser beam is 905 nm (near-infrared).

In this research, scanners are set on the floor for horizontal scanning at an elevation of about 16.3cm above the ground. Cross sectional data at the same horizontal level containing both moving and static objects are obtained in a rectangular coordinate system of real dimension. Figure 1 shows a laser scanner in experimental site.

In addition, we use multiple laser scanners to cover a large area, and to reduce occlusion as well. An integration of each scanner is conducted by using Hermart transformation that deals with a shift and a rotation. Transformation-parameters are calculated by matching the laser points of common objects such as wall and poles by manual operation. An interface that manually handles these procedures is implemented on the software, thus we can easily calibrate a number of laser scanners. A detailed description on registering multiple laser scanners can be found in Zhao and Shibasaki, 2001.

### 2.2 Pedestrian Detection

The flow of pedestrians' tracking is roughly divided into three parts: 1) Background subtraction, 2) Pedestrian recognition, 3) Pedestrian tracking. Procedure of how to compute these processes in this research is described as follows. A detailed description on following algorithms can be found in Zhao and Shibasaki, 2002.

**2.2.1 Background subtraction:** In each sampling angle of range scanning, a histogram is generated using range values from all range frames being examined. A peak value above a certain threshold is found out, which tells that an objects is continuously measured at the identical direction, i.e. the static objects. The background image is made up of the peak values at all sampling angles. As a result, by calculating the difference between range images to the background image, we can get only the moving objects.

**2.2.2 Pedestrian Recognition:** This part consists of the two processes: 1) clustering process that detects a pedestrian foot, 2) grouping process that groups two feet to have a pedestrian candidate.

Multiple points hit a pedestrians' foot because of high angle-resolution (0.5 degree). Therefore, close-by points are firstly grouped to one typical point using a centroid. We assume a number of points gathering within 10cm distance as a leg. Next, grouping process is conducted by grouping two detected feet within the 30cm distance. In this process, trajectory tracking is firstly conducted by extending the trajectories that have been extracted in previous frames, then looking for the seeds of new trajectories from the foot candidates that are not associated to any existing trajectories.

**2.2.3 Pedestrian Tracking:** When a normal pedestrian goes forward, one of the typical appearances is, at any moment, one foot swings by pivoting on the other. Two feet interchange their roles by landing and moving shifts so that the pedestrian steps forward. According to the ballistic walking model proposed by Mochon and McMahon, 1980, muscles act only to establish an initial position and velocity of the feet at the beginning half of the swing phase, then remain inactive throughout the rest half of the swing phase.

Here initial position refers to where swing foot and stance foot meets together. Let $v_L$ and $v_R$ be the speed, $a_L$ and $a_R$ be the acceleration, $p_L$ and $p_R$ be the position of left and right foot respectively, where both speed, acceleration and position are restricted to a horizontal plane, and relative to the two-dimensional global coordinate system that has been addressed in previous sections. In the case $|v_L| > |v_R|$, left foot swings forward by pivoting on right foot. At the beginning half of the swing phase, the left foot shifts from rear to initial position, and swings from standing still at an accelerated speed. Here the acceleration $|a_L|$ is a function of

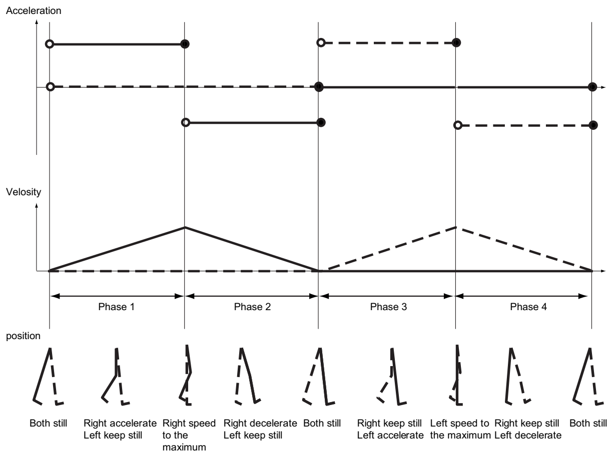



Figure 2: Pedestrian Walking Model

muscles strength. We define $|a_L| = f_L$(muscle strength). During the rest half of the swing phase, the left foot shifts from initial to front position, or swings with a negative acceleration from the maximum speed to standing-still status. Here the negative acceleration $|a_L|$ comes from factors other than left foot muscles.

We define $|a_L| = -f_L$(other forces). During the whole swing phase, the right foot keeps almost still, so it has $|v_L| \approx 0$ and $|a_L| \approx 0$. In the same way, we can deduce the speed and acceleration parameters when right foot swings forward by pivoting on left foot, where acceleration $|a_R| = f_R$(muscle strength),$|a_R| = -f_R$(other forces) at the beginning and end half of swing phase respectively, $|v_R| \approx 0$ and $|a_R| \approx 0$ during the whole swing phase. In this research, we simplify the pedestrian model by assuming that the acceleration and deceleration on both feet from either muscle strength or other forces ($|a_{L/R}|$) are equal and constant during each swing phase, and they have only smooth changes as the pedestrian steps forward. Figure 2 shows an example of the simplified pedestrian model.

As has been described in previous section, pedestrian model consists of three kinds of state parameters, position ($p_{L/R}$), speed ($v_{L/R}$), and acceleration ($a_{L/R}$). Position and speed vectors of each pedestrian change continuously shown in figure 2, while acceleration parameters change by swing phase in a discontinuous way. A discrete Kalman filter is designed in this research by dividing the state parameters into two vectors as follows.

$$\mathbf{s_{k,n}} = \Phi \mathbf{s_{k-1,n}} + \Psi \mathbf{u_{k,n}} + \omega \tag{1}$$

Where, $\mathbf{s_{k,n}}$ consists of the positions and speed vectors of both feet of pedestrian $n$ at range frame $k$, while $\mathbf{u_{k,n}}$ consists of the acceleration parameters. $\omega$ is the state estimation error.

$$\mathbf{s_{k,n}} = \begin{pmatrix} p_{L_{x,k,n}} \\ p_{L_{y,k,n}} \\ v_{L_{x,k,n}} \\ v_{L_{y,k,n}} \\ p_{R_{x,k,n}} \\ p_{R_{y,k,n}} \\ v_{R_{x,k,n}} \\ v_{R_{y,k,n}} \end{pmatrix} \tag{2}$$

$$\mathbf{u_{k,n}} = \begin{pmatrix} a_{L_{x,k,n}} \\ a_{L_{y,k,n}} \\ a_{R_{x,k,n}} \\ a_{R_{y,k,n}} \end{pmatrix} \tag{3}$$

Transition matrix $\Phi$ relates positions and speed vectors at previous time step to those of current one, while $\Psi$ relates acceleration values to the change in positions and speed vectors. They are defined as follows, where $\Delta t$ is the time interval between range frames. In this research, $\Delta t \approx 30$ms.

$$\Phi_{k,n} = \begin{pmatrix} 1 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{4}$$

$$\Psi_{k,n} = \begin{pmatrix} 0.5\Delta t^2 & 0 & 0 & 0 \\ 0 & 0.5\Delta t^2 & 0 & 0 \\ \Delta t & 0 & 0 & 0 \\ 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0.5\Delta t^2 & 0 \\ 0 & 0 & 0 & 0.5\Delta t^2 \\ 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & \Delta t \end{pmatrix} \tag{5}$$

In addition, the state vector $\mathbf{u_{k,n}}$ is predicted by identifying the swing phase. The discrete Kalman filter updates the state vector of $\mathbf{s_{k,n}}$ based on the measurements as follows.

$$\mathbf{m_{k,n}} = \mathbf{H}\mathbf{s_{k,n}} + \varepsilon \tag{6}$$

Where $\mathbf{m_{k,n}}$ denotes the measurements of pedestrian $n$ at time step $k$. $\mathbf{H}$ relate the state vector $\mathbf{s_{k,n}}$ to measurements $\mathbf{m_{k,n}}$. $\varepsilon$ represents the measurement error.

$$\mathbf{m_{k,n}} = \begin{pmatrix} p_{L_{x,k,n}} \\ p_{L_{y,k,n}} \\ p_{R_{x,k,n}} \\ p_{R_{y,k,n}} \end{pmatrix} \tag{7}$$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{8}$$

Based on these algorithms, the velocities of each pedestrians $n$ at range frame $k$ are firstly predicted by measurement $\mathbf{m_{k,n}}$. Next, a search area is defined by using velocity and direction. If foot candidates of the current frame are found inside the search area, the nearest ones are selected to compose the updated $\mathbf{m_{k,n}}$ .Otherwise, missing counter starts. If the missing counter is larger than a given threshold, e.g. 40 frames ( $\approx$ 2 sec ), then the tracing of the trajectory stops. Otherwise, the predicted $\mathbf{m_{k,n}}$ is employed as the updated one to update the state vector $\mathbf{s_{k,n}}$ and Kalman gain. The process continues until all the trajectories are traced.

(a) Video Image

(b) Laser Points

Figure 3: A snapshot of video image and laser points

### 2.3 Efficiency and Accuracy Assessment

Efficiency and Accuracy of the system is assessed to answer the following questions, what percentage of the pedestrians are measured by the sensor system (measurement ratio), what percentage of the pedestrians are successfully tracked using the tracking algorithm (success ratio), what is their relationship with the change of pedestrians' density. Laser points are manually overlapped with the video images that are taken from the ceiling. A comparison is conducted manually to find whether pedestrians on video images have corresponding laser points on their feet. Figure 3 shows a snapshot of video and laser points. The methods of overlapping and accuracy assessment are stated below.

**2.3.1 Overlapping:** In order to overlap a video and laser points, we need two requirements: 1) geometric registration, 2) temporal registration. In geometric registration, both images are required to be registered with each other so that they can have the same resolution and same coordinate system. In temporal registration, time-synchronization is correspondingly required. By satisfying both requirements, we can overlap these images.

Firstly, geometric correction of video image is conducted to satisfy the geometric requirements. This correction mainly removes a lens distortion shown as figure 3(a). Corner points of 60cm by 60cm tiles are used as reference points for geometric correction. Secondly, we set an "event" which can be recognized by both video and laser range image to satisfy the temporal requirement. In this event, we make a pump and dump a cardboard captured shown in figure3(b) and by using these event the laser time and video time are synchronized.

**2.3.2 Accuracy Assessment:** We assume that the accuracy is dependent on pedestrians' density $\mu$ and the number of sensors. Therefore, we set nine conditions as follows and conduct the accuracy assessment on each condition. The conditions are based on the number of sensors; three sensors, four sensors and six sensors; and on pedestrians' density; low-density, mid-density and high-density listed below.

- Low-density: $\bar{\mu} = 0.1$ [persons/m$^2$].
- Mid-density: $\bar{\mu} = 0.5$ [persons/m$^2$].
- High-density: $\bar{\mu} = 0.8$ [persons/m$^2$].

Where, $\bar{\mu}$ represents an average of pedestrian density $\mu$.

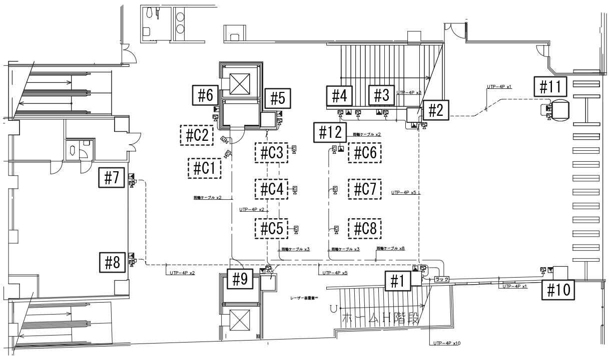



Figure 4: A Sensor alignment on a Japanese railway station, where #1 to #12 and #C1 to #C8 represent the laser and camera positions respectively.

## 3 EXPERIMENT

An experiment is conducted at the concourse of a railway station, which has a dimension of 30m by 20m. Sensors's alignment is shown in figure 4. In this experiment, we applied two types of sensors; video camera and laser scanner. Twelve laser scanners are set on the floor to cover most part of the concourse. Every scanner is controlled by computers that are physically connected by 10/100 Base LAN.

Then, we set eight video cameras on the ceiling to assess the proposed system. Figure 4 shows a concourse plan and a sensor alignment. In this plan, six cameras film most part of the crowded area by the nadir images (#C3 to #C8) and the rest two cameras film whole aspects by the diagonal images (#C1 and #C2).

## 4 RESULTS AND DISCUSSION

### 4.1 Pedestrian Detection

Although we used twelve laser scanners in this experiment, all of the sensors could not be integrated because network trouble was occurred to #8 and #10-#12. Therefore, pedestrian detection was conducted by using eight sensors excluding troubled four sensors. Figure 5 shows a result of pedestrian detection, where white circles show recognized pedestrians and white lines show those trajectories. White points near the wall show laser scanners. This image is captured when concourse was not crowded, but the proposed method could track 120 pedestrians simultaneously by the post processing.

Calibration of eight sensors could be conducted easily by using the implemented interface. Generally, the calibration of camera images require complex transformation, thus laser measurement system has an advantage on calibration. By implementing an automatic calibration, it may be to conduct calibration more easily. In addition, proposed method has an advantage on wide measurement area.

However, there are several examples where some pedestrians could not be detected or tracked due to the reasons of high pedestrians' density. First, grouping processes were failed at procedure of pedestrian recognition. For example, if density was getting higher about 0.8 [persons/$m^2$], there are failure examples where

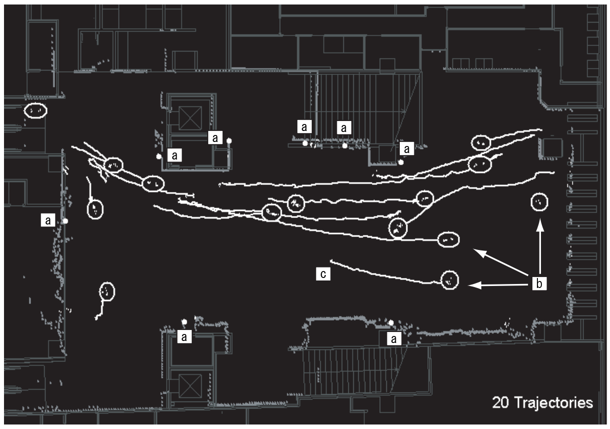



Figure 5: A Result of Pedestrian Detection. a: laser scanner, b: detected pedestrian, c: trajectory.

a leg was grouped neighbouring pedestrian because they are too close.

Moreover, in case of a pedestrian wearing a long skirt and pedestrian who have any bag or cart, pedestrian recognition tended to fail. In addition, sudden acceleration and sudden deceleration made tracking failures. For instance, tracking was failed when other pedestrian suddenly change his or her direction or stop to avoid other pedestrian coming from the opposite direction. This example is due to the reasons that a value used in Kalman filter between estimated a position to measured position surpasses the state estimation error.

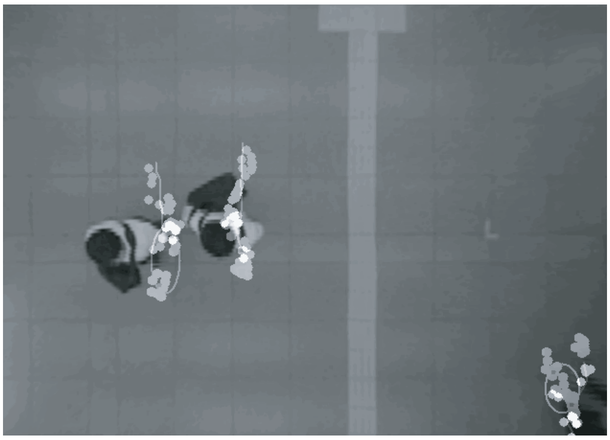
### 4.2 Movement-pattern Analysis

We analyzed the trajectories obtained by the above procedures to investigate availability for the station design. Figure 6 shows the oriented flow-lines and collision distribution at 5:30 pm using data of 50 second. Where, bright lines show passengers going to right (gate) from left (gate) and dark lines show the opposite flow lines. Dark zone means available area for measurement, gray zone shows an area covered by obstacle such as a wall and poles. In this image, white points show collisions; we assume a point that some pedestrians come close within 60cm as a collision. Moreover, we added another assumption restricted by directional vector angle that made a 180 degree ± 45 degree.

We can see the specific stratification between flows of rightward to leftward. In addition, almost all people going to the platform from the gate go to the escalator located on the left side. On the other hands, people going to the gate from central stair approach toward wall, because few people use a stair located on the center of concourse. Collisions were occurred in an area interfusing each opposite flows. We think that proposed method is very useful system for movement analysis.

### 4.3 Accuracy Assessment

Figure 7 shows overlapped results of geometrically corrected video images, laser points and detected trajectories. We conducted an accuracy assessment by using these data. At first, we calculated the measurement ratio at the nine patterns of density and number of sensors. Each pattern contains 50 images that was aggregated



(a) Low density: $\bar{\mu} = 0.1$

(b) Mid density: $\bar{\mu} = 0.4$

(c) High density: $\bar{\mu} = 0.8$

Figure 7: Overlapped results of corrected video images, laser points and detected trajectories

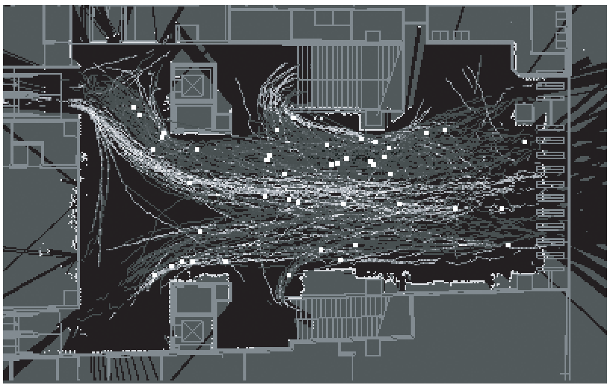


Figure 6: Oriented flow-lines and Collision-distribution

the time resolution from 20 [frames/sec] to 5 [frames/sec]. Obtained results are shown in Figure 8 and table 1, where the value inside of the case arc is measurement ratio of both feet.

By using six sensors, the measurement ratio was achieved approximately 100% in the case of $\bar{\mu} = 0.8$. In this case, 88% people were hit to both feet. Also, success ratio of pedestrian detection was approximately 100% under the environment of $\bar{\mu} = 0.4$ [persons/m$^2$]. However, we could not precisely determine a success ratio under the condition of the high density because of the complicated passenger flow. This is the future work.

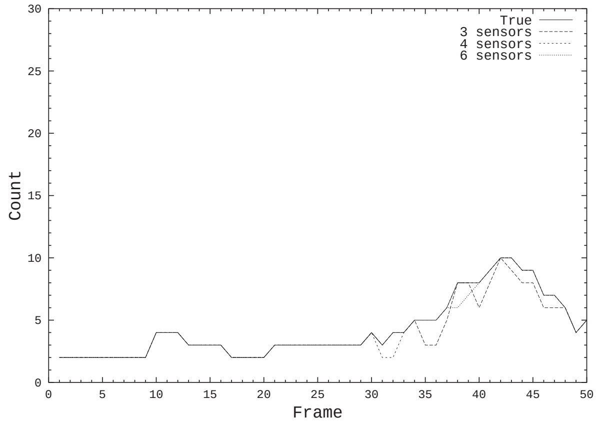
Table 1: Calculated A Measurement Ratio

| $\bar{\mu}$ | 3 seonsors | 4 sensors | 6 sensors |
|---|---|---|---|
| 0.1 | 99% (96%) | 100% (98%) | 99% (98%) |
| 0.5 | 94% (81%) | 98% (90%) | 99% (96%) |
| 0.8 | 91% (70%) | 95% (76%) | 99% (88%) |

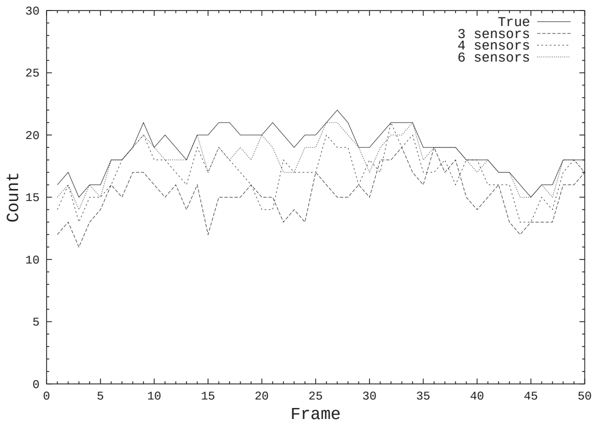
## 5 SUMMARY

In this paper, we proposed the novel system of pedestrian detection and tracking by using multiple laser range scanners. Proposed system could measure approximately 99% pedestrians in the case of $\bar{\mu} = 0.8$ [persons/m$^2$]. Also, success ratio of pedestrian detection was achieved approximately 100% under the environment of $\bar{\mu} = 0.4$ [persons/m$^2$]. Proposed system was shown that it could measure multiple pedestrians in the case of wide range and high-density area, comparing to methods by using video camera.

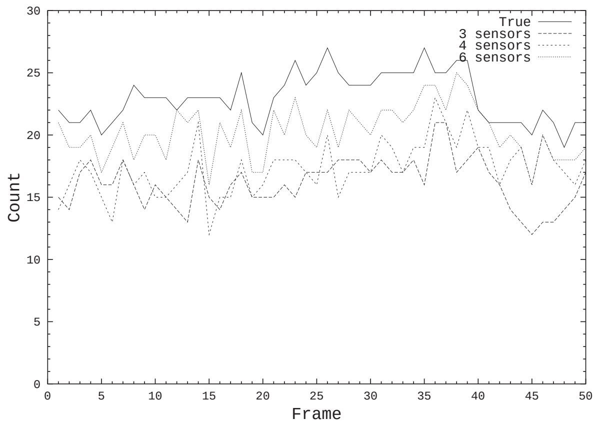
In future works, a hybrid measurement system fusing video cam-




(a) Low density: $\bar{\mu} = 0.1$




(b) Mid density: $\bar{\mu} = 0.4$




(c) High density: $\bar{\mu} = 0.8$

Figure 8: Transition of Measurement Ratio

era and laser scanners will be developed. Improvement and face identification of pedestrian will be able to be achieved.

## REFERENCES

Chia-Jung Pai and Hsiao-Rong Tyan and Yu-Ming Liang and Hong-Yuan Mark Liao and Sei-Wang Chen, 2004, Pedestrian detection and tracking at crossroas, *Pattern Recognition*, 37,pp.1025-1034.

Huijing Zhao and Ryosuke Shibasaki, 2003, Pedestrian Tracking using Multiple Laser Range Scanners, *Proceedings of Computers on Urban Planning and Urban Management*.

Huijing Zhao and Ryosuke Shibasaki, 2001. A Robust Method for Registering Ground-based Laser Range Images of Urban Outdoor Objects, *Photogram metric Engineering & Remote Sensing*, 67(10),pp.1143-1153.

Naoya Ohta, Masanobu Kiyama and Kenichi Yoshikawa, 2003, Pedestrian Detection and Identification Using Two Cameras(in Japanease), *IPSJ Journal*, 44(10), pp.2459-2468.

Romer Rosales and Stan Sclaroff, 1998, Improved Tracking of Multple Humans with Trajectory Prediction and Occlusion Modeling, *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*.

Simon Mochon and Thomas A.McMahon, 1980, Ballistic Walking, *Journal of Biomechanics*, 13, pp.49-57.