We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

2

Human Sensing in Crowd Using Laser Scanners

Katsuyuki Nakamura[1], Huijing Zhao[2],
Xiaowei Shao[3] and Ryosuke Shibasaki[3]
[1]*Central Research Lab., Hitachi, Ltd.*[*]
[2]*Peking University*
[3]*The University of Tokyo*
[1,3]*Japan*
[2]*China*

1. Introduction

Human sensing is a critical technology to achieve surveillance systems, smart interfaces, and context-aware services. Although various vision-based methods have been proposed (Aggarwal & Cai, 1999)(Gavrila, 1999)(Yilmaz et al., 2006), tracking humans in a crowd is still extremely difficult. Figure 1 shows a snapshot of a railway station during rush hour, which is one of the hardest examples of human sensing.

Suppose a camera is diagonally set up on a low-ceiling, like the one in Fig. 1. Significant occlusion tends to occur in crowded places because pedestrians severely overlap each other. As a result, sufficiently high sensing performance cannot be achieved. On the other hand, if a camera is positioned to take measurements looking straight down in order to reduce occlusions, the viewing angle is limited. Furthermore, covering large areas using multiple

Fig. 1. Snapshot of a railway station during rush-hour

_____

[*]This work was done while the first author was a Ph. D student in the University of Tokyo.

cameras is difficult due to the computational cost of data integration. These problems cannot be solved even using fisheye cameras or omni-directional cameras to expand the viewing angle. In addition, there are some cases in which cameras cannot be installed due to privacy concerns.

In this chapter, we propose a method to tackle these problems using laser scanners for human sensing in crowds. We especially focus on human tracking and gait analysis techniques. Our proposed method is well-suited to privacy protection because it does not use images but only range data. Moreover, because of the simple data structure of the laser scanner, we can easily integrate data even as the number of sensors increases, and real-time processing can be performed even when multiple sensors are used. Therefore, our method is especially suitable for crowd sensing in large public spaces such as railway stations, airports, museums, and other such facilities. That is to say, the above issues can be solved with our approach. We conducted an experiment of the proposed method in a crowded railway station in Tokyo in order to evaluate its effectiveness.

This chapter is structured as follows. Section 2 reviews existing research on human sensing. Section 3 proposes a method of tracking people in crowds using multiple laser scanners. Section 4 describes the gait analysis of tracked people. Section 5 is a performance evaluation in a crowded station.

2. Review

In this section, we briefly review the existing research on human sensing. The approaches are roughly classified into three types: vision-based, laser-based, and sensor fusion.

2.1 Vision-based approach

The first type is the vision-based approach using video cameras.

Much research has been done using this approach, although the number of people targeted for tracking has been relatively small. A well-known human detector was proposed by the article (Dalal et al., 2005). They used histograms of oriented (HOG) descriptors with a support vector machine (SVM). Felzenszwalb et al. extended this detector based on the deformable part-based model (Felzenszwalb et al., 2009). For tracking targets, mean shift trackers are widely used (Comaniciu et al., 2000).

In order to extend such approaches to track multiple targets in a crowd, we have to handle significant occlusion of each object. A typical solution is to utilize data association such as a Kalman filter, particle filter, or Markov chain Monte Carlo (MCMC) data association approach. Okuma et al. proposed a boosted particle filter that can track multiple targets by combining the AdaBoost detector and a particle filter (Okuma et al., 2004). Zhao et al. proposed a principled Bayesian framework that integrates a human appearance model, background appearance model, and a camera model (Zhao et al., 2008). The optimal solution is inferred by an MCMC-based approach. The result shows that up to 33 people are tracked in a complex scene. Kratz et al. models spatial-temporal motion patterns of crowds by using a hidden Markov model (HMM), and tracks individuals in extremely crowded scenes (Kratz & Nishino, 2010).

Other research on human sensing can be found in published surveys (Aggarwal & Cai, 1999)(Gavrila, 1999)(Yilmaz et al., 2006)(Szeliski, 2010).

2.2 Laser-based approach

The second approach is based on lasers. Fod et al. proposed laser-based tracking using multiple laser scanners (Fod et al., 2002). Their system measures a human's body, and tracks it by using a Kalman filter. More practical approaches for tracking people in crowds have been proposed by (Zhao & Shibasaki, 2005) (Nakamura et al., 2006). They measure pedestrians' feet to reduce the occlusion, and track individual pedestrians in crowds by recognizing their walking patterns. Experimental results showed that 150 people were simultaneously tracked with 80% precision in a railway station. Cui et al. combines laser-based tracking with a Rao-Blackwellized Monte Carlo data association filter (RBMC-DAF) to overcome tracking errors that occur when two closely situated data points are mixed (Cui et al., 2007). Song et al. have proposed a unified framework that couples semantic scene learning and tracking (Song, Shao, Zhao, Cui, Shibasaki & Zha, 2010). Their system dynamically learns semantic scene structures, and utilizes the learned model to increase the accuracy of tracking.

2.3 Sensor fusion

The third approach involves sensor fusion. Several techniques have been proposed to track multiple people by fusing the laser and vision approaches. Nakamura et al. used the mean shift visual tracker to support laser-based tracking (Nakamura et al., 2005). Cui et al. extended this approach by combining it with decision-level Bayesian fusion (Cui et al., 2008). Song et al. proposed a system of joint tracking and learning, which trains the classifiers that separate the targets who are in close proximately (Song, Zhao, Cui, Shao, Shibasaki & Zha, 2010). Trained visual classifiers are used to assist in laser-based tracking. Katabira et al. proposed an advanced air-conditioning system that combines laser scanners and wireless sensor networks (Katabira et al., 2006). The area that should be ventilated is determined by the humans' positions in the room and the temperature distribution.

2.4 Focus of this chapter

This chapter introduces the method of laser-based tracking and gait detection, which first emerged as a practical technique for sensing people in crowds that target more than a hundred people.

3. Laser-based people tracking

3.1 Sensing System

3.1.1 Human sensing using laser scanner

We use a laser scanner called SICK LMS-200. This sensor measures distance by using the time of flight (ToF) of laser light, and can also perform wide-area measurements (30-m distance). In addition, because the dispersion of the laser waves is minimal, resolution is high, and the angular resolution is $0.25°$ at the maximum. The wavelength of laser light is 905 nm

Fig. 2. Snapshot of human sensing using a laser scanner
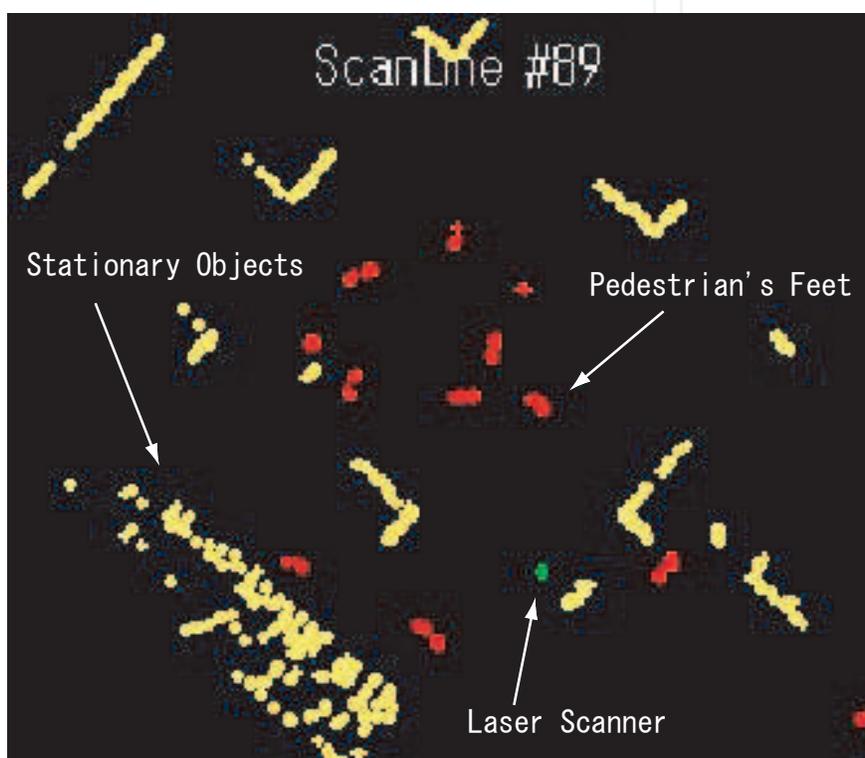
Fig. 3. Example of range data obtained with laser scanner

(near-infrared region), and it is a class 1A laser that is safe for peoples' eyes. The sampling frequency varies depending on the settings; it was 37.5 Hz in our case.

In the proposed method, a flat area about 16 cm off the floor is scanned with the sensors set on the floor. As a result, range data for ankles, including both static objects and moving objects, can be obtained. Figure 2 shows a sensing system, and Fig. 3 shows an example of the obtained range data.

3.1.2 Human sensing using multiple Laser scanners

We performed human sensing by using multiple laser scanners in order to minimize occlusions in the wide-area sensing. Suppose that each sensor obtains data at the same horizontal level; integration of multiple range data can be achieved using the following

Helmart transformation.

$$\begin{pmatrix} u \\ v \end{pmatrix} = m \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \tag{1}$$

where (x, y) represent a laser point in the local coordinate, (u, v) represent a transformed laser point in the global coordinate, m is a scaling factor, α is a rotation angle, and Δx and Δy are vectors shifted from the origin. These parameters are estimated by taking a visual correspondence using the rotation and shift of shared static objects (e.g. walls, pillars, etc.) measured by each sensor. The interface that performs this operation is built into the software. After the integration and synchronization, human tracking is conducted by the algorithms explained in Section 3.2.

3.2 Tracking algorithm

3.2.1 Tracking flow

Figure 4 illustrates the flow diagram of laser-based people tracking.
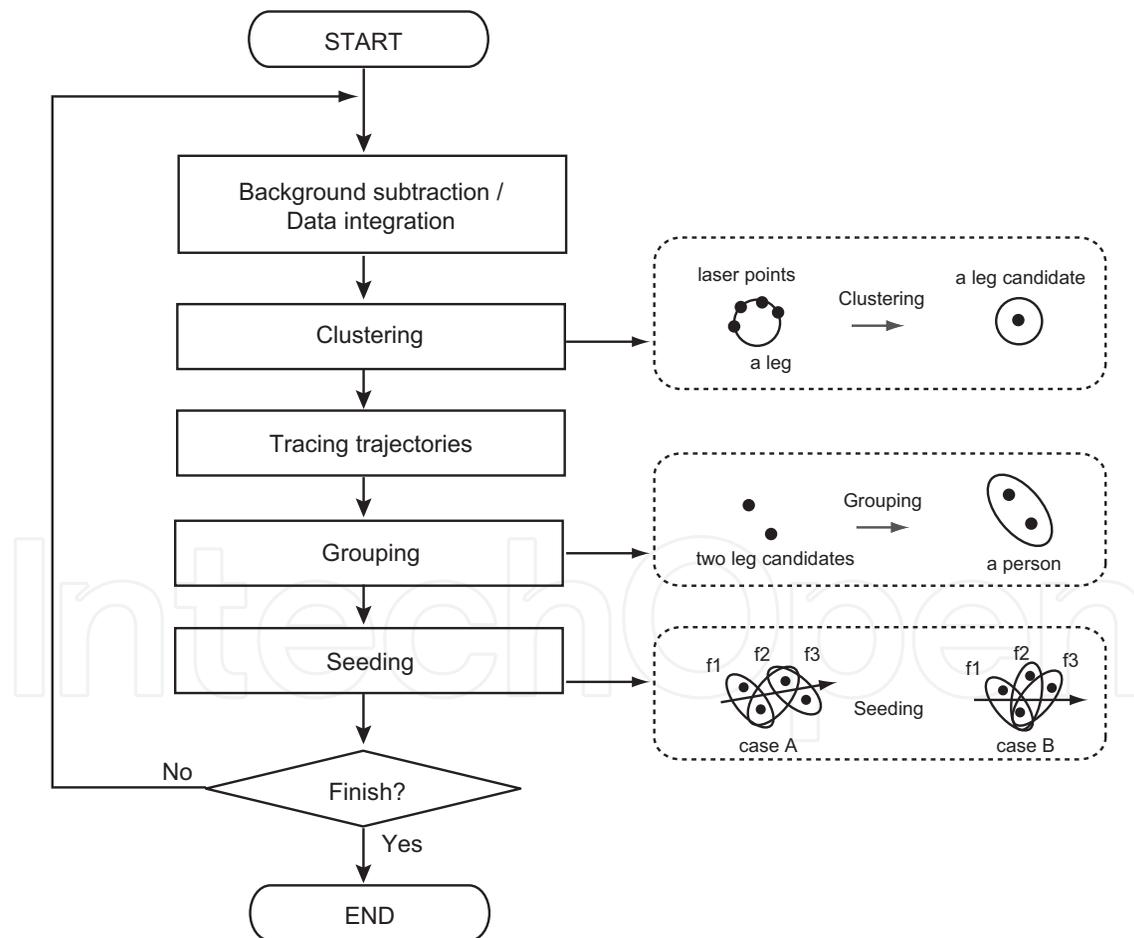
Fig. 4. Flow diagram of laser-based people tracking

First, background subtraction is conducted in each sensor in order to detect moving objects and integrate them to the global coordinates by using equation (1).

Second, several laser point strikes on the foot are clustered in order to extract one foot candidate. In this study, a group of points within a radius of 15 cm is clustered as a foot candidate. In practice, due to errors in the sensor calibration, there are several cases in which the foot of one human is not entirely within a cluster, or the feet of several humans are within the same cluster. However, such false positives or false negatives can be reduced during the subsequent stages, and there is no significant impact on tracking processing.

Third, the existing trajectories are extended to the current frame by using the Kalman filter. The details of this process are described in Section 3.2.2 to 3.2.4. By using a dynamic model of humans, the best foot candidate is integrated.

Last, if the foot candidate does not integrate into the existing trajectories, a new trajectory is created in the following steps of initial detection, and the initial state is set for the Kalman filter.

1. Grouping: When foot candidates not belonging to any of the existing trajectories are within 50 cm of each other, two foot candidates are grouped to create a human candidate. When there is a crowd, several human candidates can be created. Invalid human candidates are eliminated using the following seeding process.

2. Seeding: In consecutive frames, the candidates that satisfy the following two conditions are taken to represent the same human, and the connected centers of gravity for the two moving foot candidates represent a new trajectory. (a) At least one foot overlaps for a human candidate in consecutive frames (three or more frames). (b) The motion vector created by the other foot, which does not overlap, changes smoothly.

3.2.2 Walking model

When walking, pedestrians make progress by using one foot as an axis and moving the other foot. The two feet change roles alternately as they reach the ground and create a rhythmic walking motion. According to the ballistic walking model (Mochon & McMahon, 1980), muscle power acts when generating speed during the first half of the foot's movement, and the latter half of the foot's movement is passive. Figure 5 represents a simplified model of a pedestrian walking with attention given to the changes in the movement, speed, and acceleration of the feet.

In this research, the movement of the two feet is defined using four phases. Phase 1 is defined as going from a stationary state for both feet through the acceleration of the right foot alone, and to where the two feet are in alignment. Phase 2 is defined as when the right foot then decelerates and reaches the ground. In the same fashion, Phase 3 is defined as when the left foot accelerates, and Phase 4 is defined as when it decelerates.

The values v_L and v_R are the speed of the left and right feet respectively, a_L and a_R are their acceleration, and p_L and p_R are their positions. These variables are taken to have the values in the observed plane integrated using the process in Section 3.1.2.

Table 1 summarizes the transitions of state parameters in walking phases. When $|v_R| > |v_L|$, the right foot is in front, with the left foot serving as the axis. Here, the acceleration $|a_R|$ acting on the right foot can be taken to be a function of muscle power as based on the ballistic walking model. The authors define the acceleration of the right foot in walking phase 1 as $|a_R| = f_R \dot{v}$.
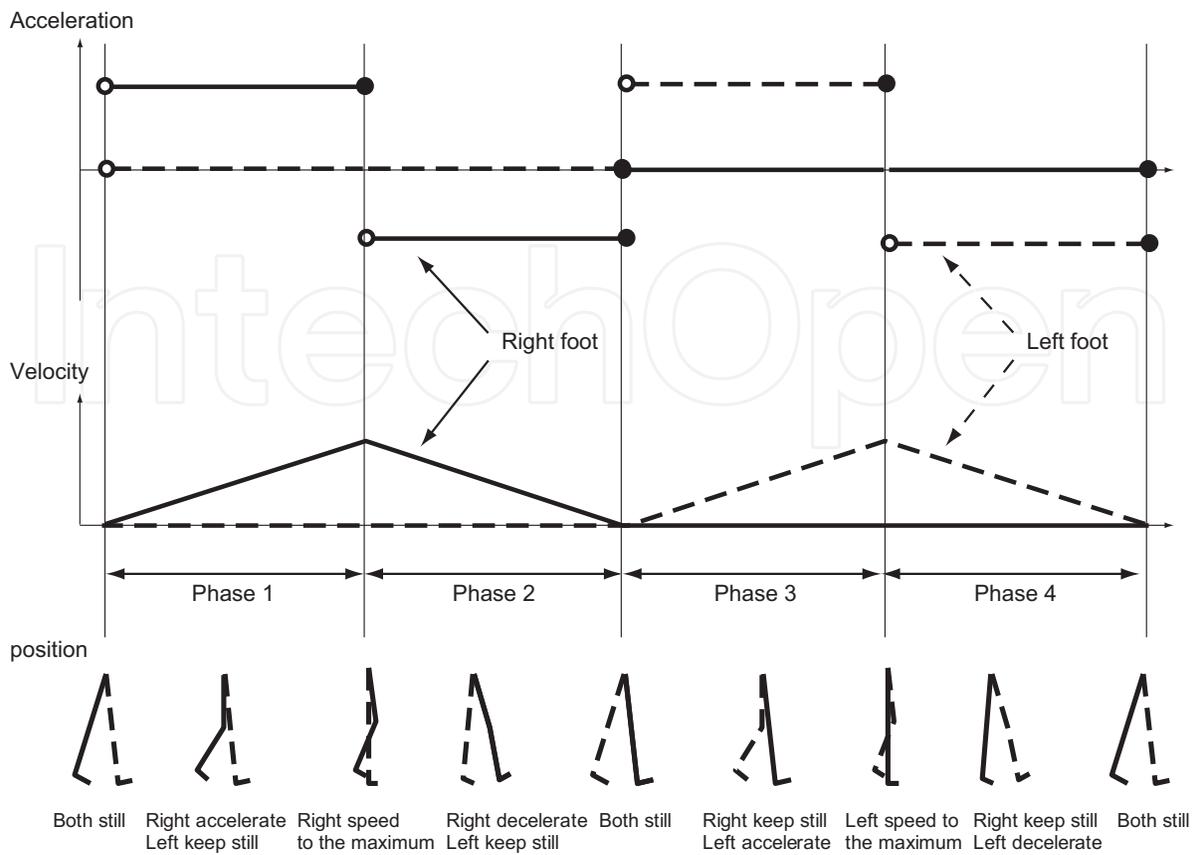
Fig. 5. Simplified walking model (top row: acceleration; middle row: velocity; bottom row: position)

Here, $f_{L/R}$ represents the acceleration function for the two feet defined in Equation (8), and $\dot{v}$ represents the unit direction vector. In walking phase 2, the right foot decelerates at a steady rate, and both feet are on the ground. The acceleration acting here has a negative value because of the effect of an external force other than muscle power. This is defined as $|a_R| = -f_R\dot{v}$. In walking phases 1 and 2, the left foot is virtually stationary, and thus, $|v_L| \approx 0$ and $|a_L| \approx 0$. When the right foot is the axis, the acceleration for the left foot in walking phases 3 and 4 for foot movement is $|a_L| = f_L\dot{v}$ and $|a_L| = -f_L\dot{v}$ and the velocity can be defined as $|v_R| \approx 0$ and $|a_R| \approx 0$. In the state in which both feet are on the ground, $|v_{L/R}| \approx 0$ and $|a_{L/R}| \approx 0$.

Phase	Phase1	Phase2	Phase3	Phase4
v_L	$|v_L| \approx 0$	$|v_L| \approx 0$	$|v_R| < |v_L|$	$|v_R| < |v_L|$
v_R	$|v_R| > |v_L|$	$|v_R| > |v_L|$	$|v_R| \approx 0$	$|v_R| \approx 0$
a_L	$|a_L| \approx 0$	$|a_L| \approx 0$	$|a_L| > 0$	$|a_L| < 0$
a_R	$|a_R| > 0$	$|a_R| < 0$	$|a_R| \approx 0$	$|a_R| \approx 0$

Table 1. Transitions of state parameters in the walking phases

3.2.3 Definition of Kalman filter

As was described in Section 3.2.2, the walking model proposed in this chapter has three state parameters: $v_{L/R}$, $a_{L/R}$, and $p_{L/R}$. As shown in Fig. 5, although the position and velocity

of each pedestrian vary continuously, the acceleration varies discretely depending on the phase of the foot movement. Thus, the state parameters are divided into two vectors, and the Kalman filter is defined based on the dynamics for moving objects.

$$\mathbf{s}_{k,n} = \Phi \mathbf{s}_{k-1,n} + \Psi \mathbf{u}_{k,n} + \omega \tag{2}$$

Here, $\mathbf{s}_{k,n}$ represents the state vector for the position $p_{L/R}$, and the velocity $v_{L/R}$ for both feet for pedestrian n at the measurement time k. The vector $\mathbf{u}_{k,n}$ represents the state vector for the acceleration $a_{L/R}$, and ω represents the system noise. The subscripts x and y for each element represent the spatial coordinates.

$$\mathbf{s}_{k,n} = \begin{pmatrix} p_{L_{x,k,n}} \\ p_{L_{y,k,n}} \\ v_{L_{x,k,n}} \\ v_{L_{y,k,n}} \\ p_{R_{x,k,n}} \\ p_{R_{y,k,n}} \\ v_{R_{x,k,n}} \\ v_{R_{y,k,n}} \end{pmatrix} \tag{3}$$

$$\mathbf{u}_{k,n} = \begin{pmatrix} a_{L_{x,k,n}} \\ a_{L_{y,k,n}} \\ a_{R_{x,k,n}} \\ a_{R_{y,k,n}} \end{pmatrix} \tag{4}$$

The transition matrix Φ and Ψ are related to the state vector $\mathbf{s}_{k,n}$ and $\mathbf{u}_{k,n}$ from the past frame $k-1$ to the present frame k. Here, Δt is the interval for observations, and in this study $\Delta t \approx 26$ milliseconds.

$$\Phi = \begin{pmatrix} 1 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5}$$

$$\Psi = \begin{pmatrix} \frac{1}{2}\Delta t^2 & 0 & 0 & 0 \\ 0 & \frac{1}{2}\Delta t^2 & 0 & 0 \\ \Delta t & 0 & 0 & 0 \\ 0 & \Delta t & 0 & 0 \\ 0 & 0 & \frac{1}{2}\Delta t^2 & 0 \\ 0 & 0 & 0 & \frac{1}{2}\Delta t^2 \\ 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & \Delta t \end{pmatrix} \tag{6}$$

Moreover, the state vector $\mathbf{u}_{k,n}$ is estimated by recognizing the walking phase using the algorithm 1. Here, $\cdot$ represents the inner product of the vector.

Algorithm 1 Predicting $\mathbf{u}_{k,n}$

1: **if** $|v_{L_{k-1,n}}| < |v_{R_{k-1,n}}|$ **then**
2:      **if** $(p_{R_{k-1,n}} - p_{L_{k-1,n}}) \cdot \dot{v}_{R_{k-1,n}} > 0$ **then**
3:          /* Right foot is the rear foot (Phase 1) */
4:          $a_{R_{k,n}} \leftarrow f_R \dot{v}_{L_{k-1,n}}$
5:          $a_{L_{k,n}} \leftarrow 0$
6:      **else**
7:          /* Right foot is the front foot (Phase 2) */
8:          $a_{R_{k,n}} \leftarrow -f_R \dot{v}_{L_{k-1,n}}$
9:          $a_{L_{k,n}} \leftarrow 0$
10:      **end if**
11: **else if** $|v_{L_{k-1,n}}| > |v_{R_{k-1,n}}|$ **then**
12:      **if** $(p_{L_{k-1,n}} - p_{R_{k-1,n}}) \cdot \dot{v}_{L_{k-1,n}} > 0$ **then**
13:          /* Left foot is the rear foot (Phase 3) */
14:          $a_{L_{k,n}} \leftarrow f_L \dot{v}_{L_{k-1,n}}$
15:          $a_{R_{k,n}} \leftarrow 0$
16:      **else**
17:          /* Left foot is the front foot (Phase 4) */
18:          $a_{L_{k,n}} \leftarrow -f_L \dot{v}_{L_{k-1,n}}$
19:          $a_{R_{k,n}} \leftarrow 0$
20:      **end if**
21: **end if**

The acceleration function $f_{L/R}$ is calculated with the equations below using the average step length $S_{L/R}$.

$$S_{L/R} = \frac{1}{N} \sum_{t=j+1}^{k} \left| p_{L/R,t,n} - p_{L/R,t-1,n} \right| \tag{7}$$

$$f_{L/R} = \frac{S_{L/R}}{(k-j+1)^2 \Delta t^2} \tag{8}$$

Here, N represents the number of walking phases recognized from frame j to k. Frame j is determined experimentally. Furthermore, the initial value for the acceleration is calculated based on the amount of movement in the new tracing.

The Kalman filter updates the state vector $\mathbf{s}_{k,n}$ using the following equation based on the observed value vector $\mathbf{m}_{k,n}$.

$$\mathbf{m}_{k,n} = \mathbf{H}\mathbf{s}_{k,n} + \epsilon \tag{9}$$

where, $\mathbf{H}$ represents measurement matrix, ϵ is measurement noise.

$$\mathbf{m}_{k,n} = \begin{pmatrix} p_{L_{x,k,n}} \\ p_{L_{y,k,n}} \\ p_{R_{x,k,n}} \\ p_{R_{y,k,n}} \end{pmatrix} \tag{10}$$

$$\mathbf{H} = \begin{pmatrix} 1\,0\,0\,0\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,1\,0\,0 \end{pmatrix} \tag{11}$$

3.2.4 Tracing trajectories using Kalman Filter

Figure 6 shows the flow diagram of the tracing trajectories using the Kalman Filter.

First, the walking phase is recognized by using the algorithm described in Section 3.2.3, and $\mathbf{u}_{k,n}$ is estimated. Then, $\hat{\mathbf{s}}_{k,n}$ and $\hat{\mathbf{m}}_{k,n}$ are predicted. Next, the foot candidates in a search area S_{area} from among the predicted vector $\hat{\mathbf{m}}_{k,n}$ are searched. If a foot candidate is detected in the search area, then it is taken to be the foot $\mathbf{m}_{k,n}$ of the pedestrian candidate, and the state vector $\mathbf{s}_{k,n}$ is updated. If several foot candidates are found, then the one with the smallest Euclid distance for $\hat{\mathbf{m}}_{k,n}$ is taken to be the foot $\mathbf{m}_{k,n}$ of the pedestrian candidate. If no foot candidates are found, then because of the possibility of occlusion, this is allowed for only a set period of time T_{thd}. In this instance $\mathbf{m}_{k,n}$ cannot be obtained, and so only the state vector and the error covariance matrix are predicted. If the set period of time T_{thd} is exceeded, then the target is considered lost and the search is canceled. Tracking is performed by repeating this process until all tracings are completed.

4. Gait analysis for tracked people

4.1 Gait features

Gait refers to the walking style of humans, and is defined by several parameters such as walking speed, stride length, cadence, step width, ratio of stance phase, and swing phase. These parameters are useful not only in clinical applications, but also in research focusing on human identification, gender recognition, and age estimation (Sarkar et al., 2005). Gait detection is actively studied in the field of computer vision; for example, Bobick and Johnson used the action of walking to extract body parameters instead of directly analyzing dynamic gait patterns (Bobick & Johnson, 2001), BenAbdelkader et al. analyzed the periodicity of the width of an extracted bounding area, and then computed the period using autocorrelation (BenAbdelkader et al., 2002).

Stride length and cadence (number of steps per minute) are generally considered to be the most important gait features. This is because these features are easy to measure by visual observation. However, laser scanners can achieve more detailed analyses that have shorter time intervals than those observed visually. In this research, we extracted step length rather than stride length, and cycle time (walking cycle) rather than cadence, as depicted in Fig. 7.

4.2 Gait detection by using spatial-temporal clustering of range data

Generally, the movements of pedestrians' feet are periodic. If we put all the laser points into the spatial-temporal domain, we can see that some periodic spiral patterns are generated, as shown in Fig. 8. The cross points of this spiral pattern correspond to the axes of the feet. Therefore, we can detect gait features by using nearby cross points.
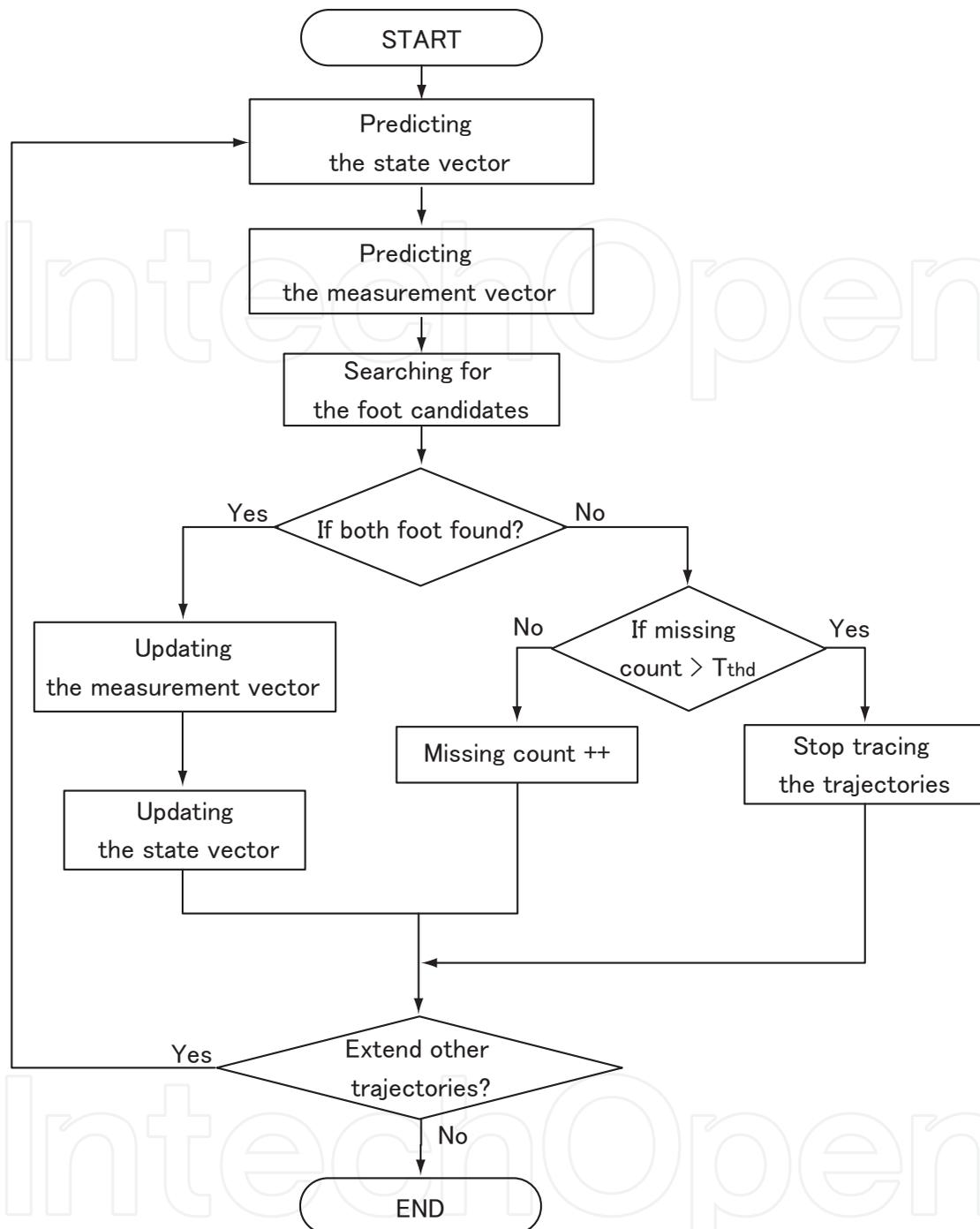
Fig. 6. Flow diagram of tracing trajectories using the Kalman filter

In this research, we used mean shift clustering (Comaniciu & Meer, 1999) to detect cross points. Mean shift is a well-known algorithm for finding the local maximum of an underlying density function. Here, a Gaussian kernel is used, where σ_s and σ_t stand for the kernel size in the space and time domains, respectively. We implement the mean-shift algorithm with $\sigma_s = 0.15$ m $\sigma_t = 0.5$ sec. The detected cross points are indicated by solid circles in Fig. 8. It can be seen that the cross points have been correctly detected. More details of this process can be found in our previous research (Shao et al., 2006).
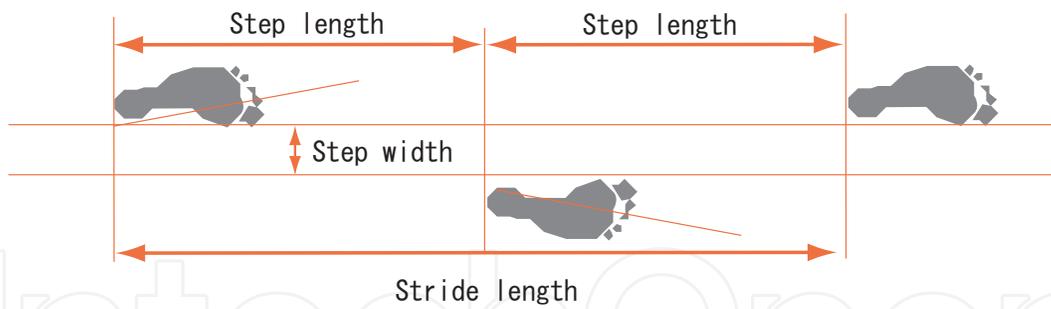
Fig. 7. Gait features
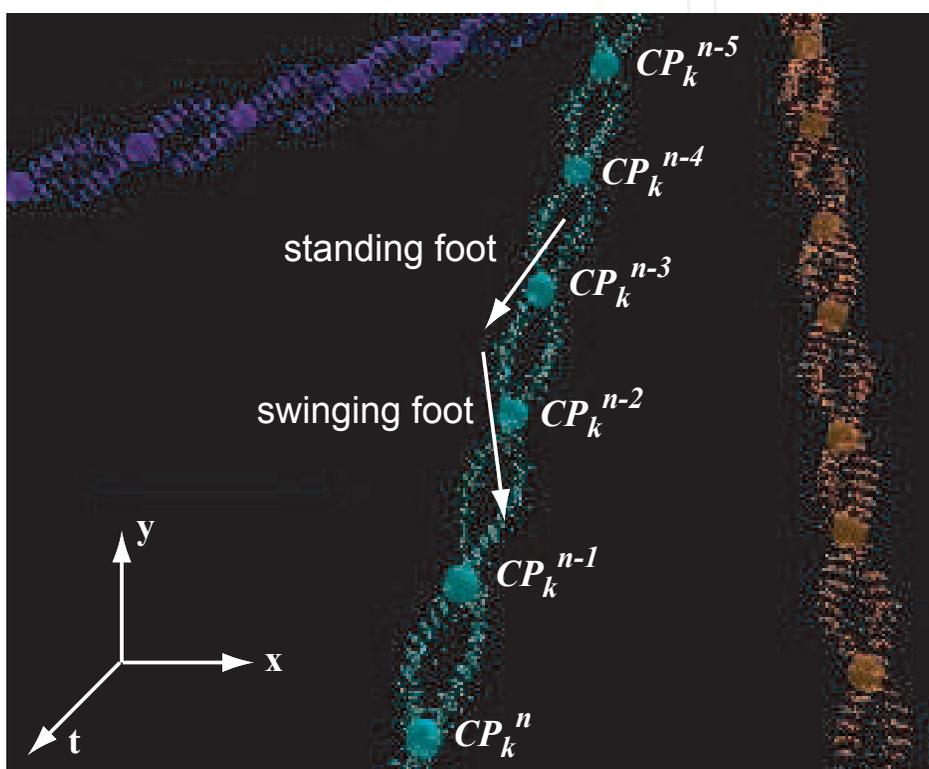
Fig. 8. Periodic spiral patterns of laser points in spatial-temporal domain. Detected cross points (CPs) representing axes of the feet are also marked by solid circles.

With the cross point $CP_k^n = (cx_k^n, cy_k^n, ct_k^n)$ for pedestrian k at measurement time n, step length s_k^n and cycle time ω_k^n can be computed by the following equations.

$$s_k^n = \left\{ (cx_k^n - cx_k^{n-1})^2 + (cy_k^n - cy_k^{n-1})^2 \right\}^{\frac{1}{2}} \tag{12}$$

$$\omega_k^n = ct_k^n - ct_k^{n-1} \tag{13}$$

Moreover, walking speed can be calculated by $v_k^n = s_k^n / \omega_k^n$. In this research, several cross points satisfying $v_k^n \leq 0.2\,\text{m/s}$ were eliminated, because a stationary human does not provide a spiral pattern in the spatial-temporal spaces, which leads to a detection error.

5. Experiment

5.1 Experimental conditions

We evaluated the effectiveness of the proposed method through an experiment conducted at a railway station in Tokyo, which is used by roughly 250,000 people per day. The station concourse is about 20 meters by 30 meters, and can hold over 150 passengers at a time.

Figure 9 shows a plane view of the concourse and the locations of the sensors. The shadow areas are indicated in the observation field. The darker the shadow, the greater the number of sensors observing the area. Eight laser sensors (#1 through #8) were set up around the most crowded area. Furthermore, in order to evaluate the proposed method with real-world conditions, the authors set up several cameras to obtain video. Six cameras (#C3 - #C8) were positioned on the ceiling to take video from directly above the concourse, and two video cameras (#C1 - #C2) were positioned to take video diagonally.
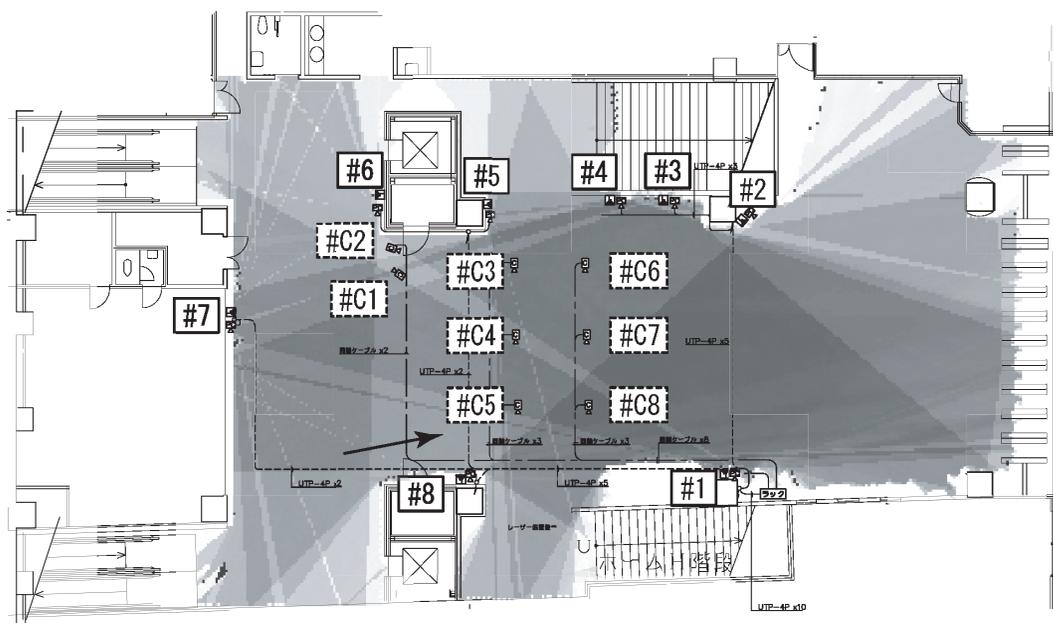
Fig. 9. Sensor alignments in a railway station, where #1 to #8 and #C1 to #C8 represent the position of laser scanners, and video cameras respectively. Shadow area shows the observation fields.

5.2 People tracking in crowd

Figure 10 shows the results of people tracking in crowds during rush hour. The red ellipses are recognized people, and yellow points are laser points[1]. Although significant occlusions occur, our proposed method can robustly track each pedestrian in the crowd. We found that a maximum of 150 people could be tracked at the same time and that tracking precision

---

[1] Calibration between cameras and lasers was done using Tsai's method (Tsai, 1987). Recognized people are approximated by a 170-cm height by 50-cm width, and back-projected to the image plane of the #C2 camera.

exceeded 80% during rush hour. The average pedestrian density at this time was roughly about 0.6 people/m^2.

The proposed method was more effective for tracking people in crowds in wide open areas than with vision-based methods. Because this method is also useful for protecting privacy due to using only range data, it can be used for sensing in areas where it is difficult to set up video cameras.

(a) frame 41

(b) frame 70

(c) frame 97

(d) frame 127
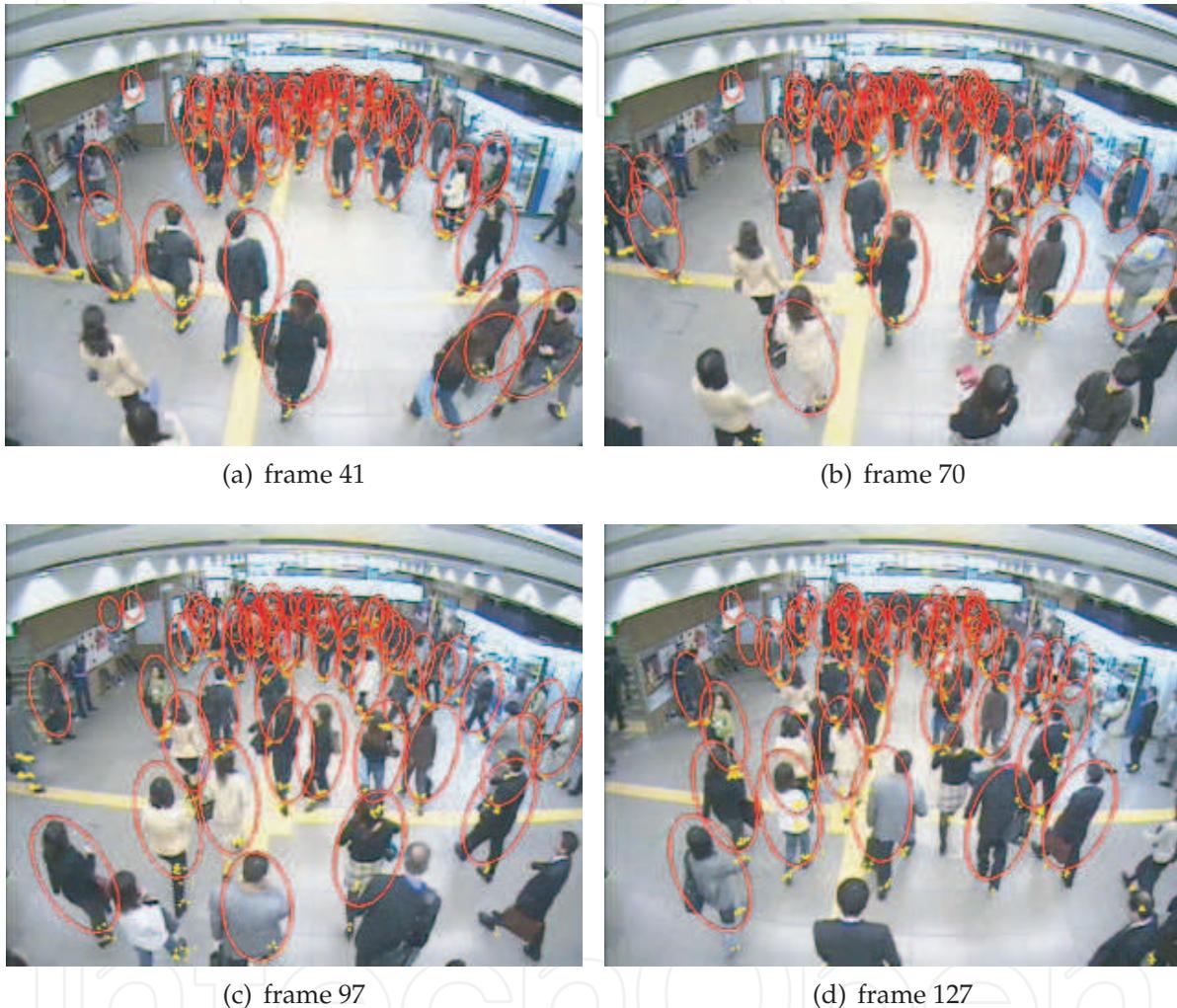
Fig. 10. Results of people-tracking in crowd during rush-hour

5.3 Gait analysis

Figure 11 plots the distribution of step length and cycle time of some different walking patterns. The mean and variance of step length, cycle time, and speed are listed in Table 2. The x-axis in Fig. 11 is the step length, and the y-axis is the cycle time. As we can see, pedestrians #2 and #3 have almost the same speed (average speed = 1.21 m/s and speed variance = 0.08 m/s) but they are well separated because they have different step lengths and cycle times. Also, pedestrian #4 walks stably, with a variance of step length of 3 cm and variance of cycle time of 40 ms. As another example, pedestrians #1, #2, and #3 have almost the same

cycle times, but they are separable because of their different step lengths. Pedestrian #5 was running. He walked very fast and then began to run. From Table 2 we can see that his cycle time is short and stable, with a mean of 0.33 s and a variance of 20 ms, but his step length varies greatly, from about 0.6 m to 1.5 m, because of the change from fast walking to running. In this experiment, step length and cycle time of each step were extracted from our walking model and employed with speed to analyze different walking patterns. The results demonstrate that different walking patterns have their own distributions in the step length to cycle time space, and useful information about their behavior can be obtained. More information on activity recognition using gait features can be found in the reference (Nakamura et al., 2007).
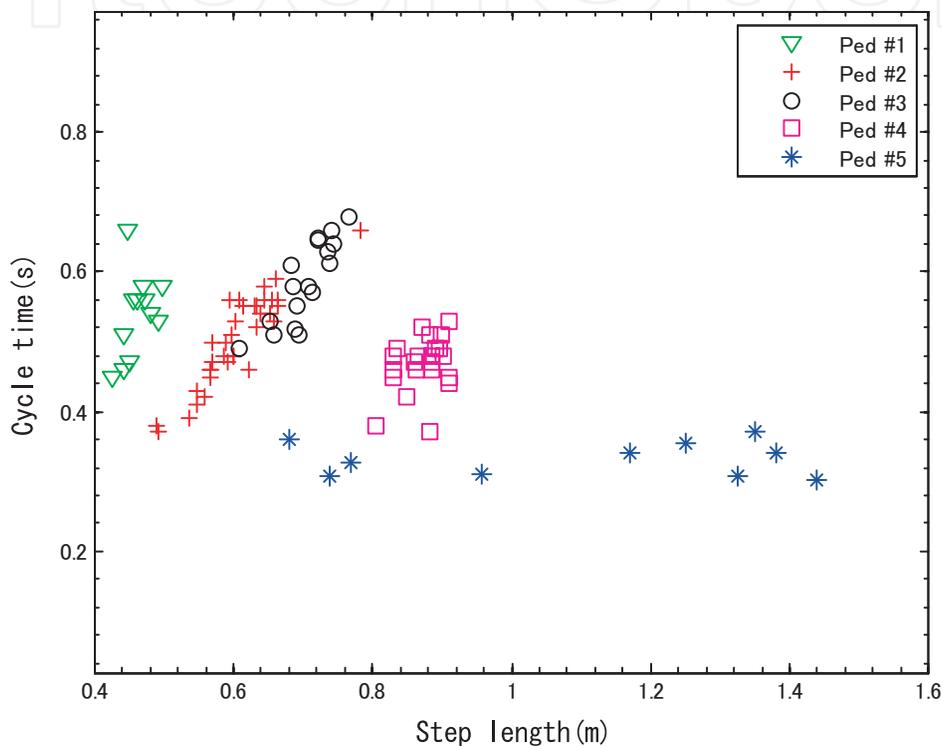
Fig. 11. Examples of detected gait features in different walking styles

Ped #	Step length (m)	Cycle times (s)	Speed (m/s)
1	0.46 ± 0.02	0.54 ± 0.06	0.87 ± 0.08
2	0.60 ± 0.06	0.50 ± 0.07	1.21 ± 0.08
3	0.70 ± 0.04	0.59 ± 0.06	1.21 ± 0.08
4	0.87 ± 0.03	0.47 ± 0.04	1.88 ± 0.17
5	1.11 ± 0.29	0.33 ± 0.02	3.34 ± 0.92

Table 2. Statistical features of different walking styles

5.4 Application to crowd-flow analysis

Our method can be applied to analyze both crowd flow and local activities. Figure 12 shows the results of visualizing crowd flow for one day. Blue lines indicate movement from right to left, and yellow lines are the opposite flow. The red points represent static people (e.g. moving at a speed < 0.3 m/sec), and white points show collision avoidance between two people (e.g.

two passengers get close to each other within 60 cm). Figure 13 shows a detected average number of train passengers during a day. We can see two peaks during the commuter rushes.
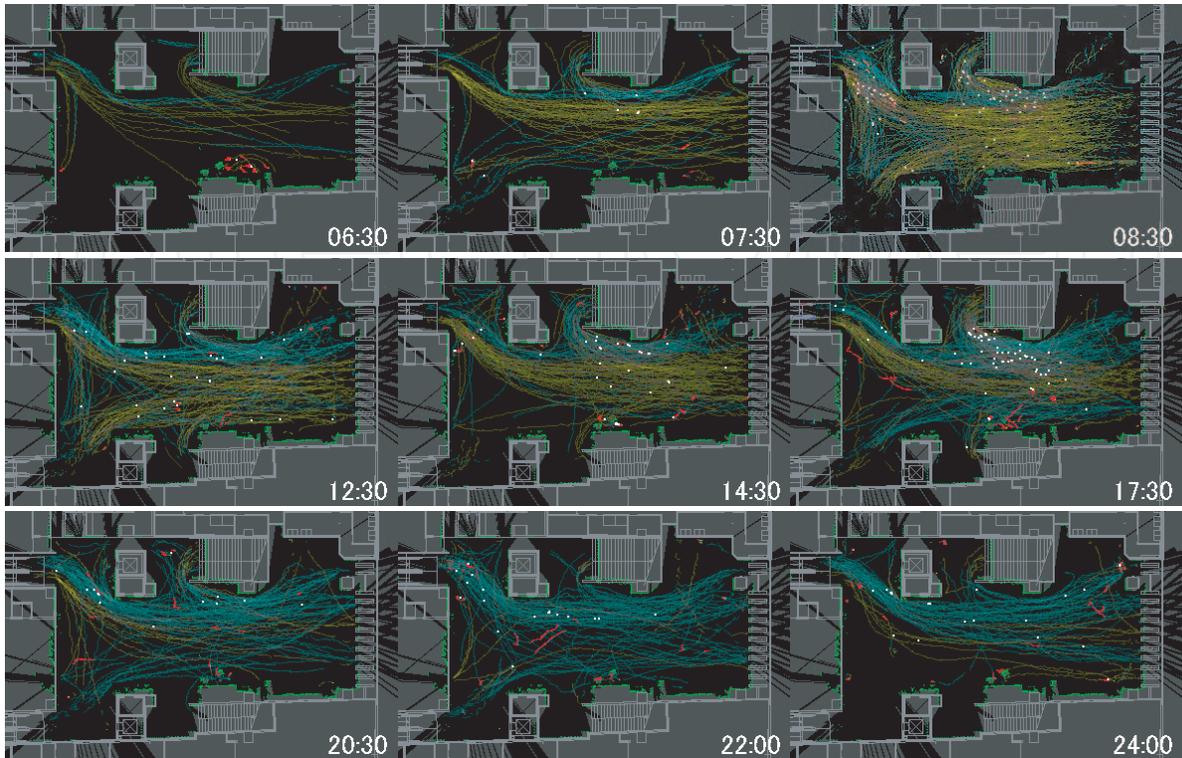

Fig. 12. Visualization results of crowd flow in one day. Blue lines are movement from right to left, and yellow lines are from left to right. Red points represent static people, and white points indicate collision avoidance between two people.
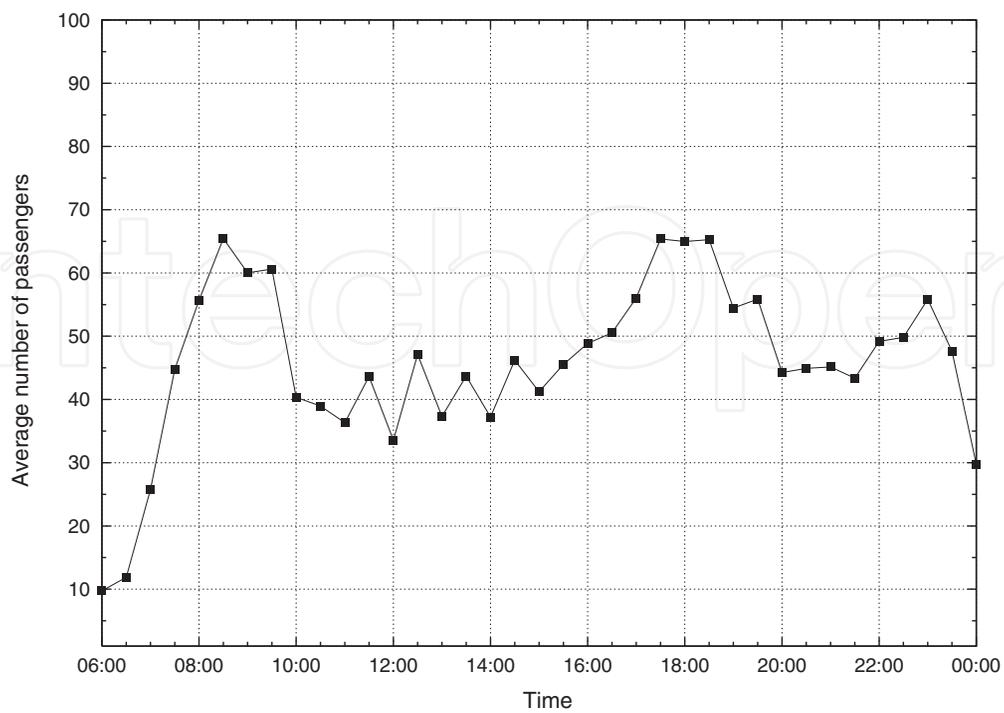

Fig. 13. Detected average number of passengers at a railway station in one day.

6. Conclusion

In this chapter, we described a method of human sensing in a crowd using multiple laser scanners. We evaluated the effectiveness of the proposed method for human tracking and gait analysis in a crowd through an experiment conducted at a large railway station. Our proposed method is well-suited to protect privacy because it does not use images but uses range data only. Therefore, our method is especially suitable for crowd sensing in public spaces such as railway stations, airports, and museums. We believe that this laser-based method is a necessary approach to complement vision-based methods that makes it possible to achieve a wider range of applications.

7. Acknowledgment

8. References

Aggarwal, J. K. & Cai, Q. (1999). Human Motion Analysis: A Review, *Computer Vision and Image Understanding* 73(3): 428–440.

BenAbdelkader, C., Cutler, R. & Davis, L. (2002). Stride and cadence as a biometric in automatic person identification and verification, *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (FGR)*, pp. 357–362.

Bobick, A. F. & Johnson, A. Y. (2001). Gait recognition using static, activity-specific parameters, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 423–430.

Comaniciu, D. & Meer, P. (1999). Distribution Free Decomposition of Multivariate Data, *Pattern Analysis & Applications* 2(1): 22–30.

Comaniciu, D., Ramesh, V. & Meeo, P. (2000). Real-Time Tracking of Non-Rigid Objects using Mean Shift, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 142–151.

Cui, J., Zha, H., Zhao, H. & Shibasaki, R. (2007). Laser-based detection and tracking of multiple people in crowds, *Computer Vision and Image Understanding* 106(2-3): 300–312.

Cui, J., Zha, H., Zhao, H. & Shibasaki, R. (2008). Multi-modal tracking of people using laser scanners and video camera, *Image and Vision Computing* 26(2): 240–252.

Dalal, N., Triggs, W. & Schmid, C. (2005). Human Detection using Oriented Histograms of Flow and Appearance, *Proceedings of the European Conference on Computer Vision (ECCV)*, Vol. 3952, Springer, pp. 428–441.

Felzenszwalb, P. F., Girshick, R. B., Mcallester, D. & Ramanan, D. (2009). Object Detection with Discriminatively Trained Part Based Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(9): 1–20.

Fod, A., Howard, A. & Matarić, M. J. (2002). Laser-Based People Tracking, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3024–3029.

Gavrila, D. M. (1999). The Visual Analysis of Human Movement: A Survey, *Computer Vision and Image Understanding* 73(1): 82–98.

Katabira, K., Zhao, H., Shibasaki, R. & Ariyama, I. (2006). Real-Time Monitoring of People Behavior and Indoor Temperature Distribution using Laser Range Scanners

and Sensor Networks for Advanced Air Conditioning Control, *Proceedings of the International Conference on Networked Sensing Systems (INSS)*, pp. BOF–11.

Kratz, L. & Nishino, K. (2010). Tracking with local spatio-temporal motion patterns in extremely crowded scenes, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 693–700.

Mochon, S. & McMahon, T. A. (1980). Ballistic Walking, *Journal of Biomechanics* 13: 49–57.

Nakamura, K., Shao, X., Zhao, H. & Shibasaki, R. (2007). Recognizing Non-Stationary Walking based on Gait Analysis using Laser Scanners (in Japanese), *IEEJ Transactions on Electronics, Information and Systems* 127(4): 537–545.

Nakamura, K., Zhao, H. & Shibasaki, R. (2005). Tracking Pedestrians using Laser Scanners and Image Sensors (in Japanese), *Proceedings of the Symposium on Sensing via Image Imformation (SSII)*, pp. 177–180.

Nakamura, K., Zhao, H., Shibasaki, R., Sakamoto, K., Ohga, T. & Suzukawa, N. (2006). Tracking pedestrians using multiple single-row laser range scanners and its reliability evaluation, *Systems and Computers in Japan* 37(7): 1–11.

Okuma, K., Taleghani, A., Freitas, N. D., Little, J. J. & Lowe, D. G. (2004). A Boosted Particle Filter : Multitarget Detection and Tracking, *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, pp. 28–39.

Sarkar, S., Phillips, J., Liu, Z., Vega, I. R., Grother, P. & Bowyer, K. W. (2005). The Human ID Gait Challenge Problem: Data Sets, Performance, and Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(2): 162–177.

Shao, X., Zhao, H., Nakamura, K., Shibasaki, R., Zhang, R. & Liu, Z. (2006). Analyzing Pedestrian's Walking Pattern Using Single-Row Laser Range Scanners, *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1202 – 1207.

Song, X., Shao, X., Zhao, H., Cui, J., Shibasaki, R. & Zha, H. (2010). An Online Approach: Learning-Semantic-Scene-by-Tracking and Tracking-by-Learning-Semantic-Scene, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 739–746.

Song, X., Zhao, H., Cui, J., Shao, X., Shibasaki, R. & Zha, H. (2010). Fusion of Laser and Vision for Multiple Targets Tracking via On-line Learning, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 406–411.

Szeliski, R. (2010). *Computer Vision : Algorithms and Applications*, Springer-Verlag New York Inc.

Tsai, R. Y. (1987). A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology using Off-the-shelf TV Cameras and Lenses, *IEEE Journal of Robotics and Automation* RA-3(4): 323–344.

Yilmaz, A., Javed, O. & Shah, M. (2006). Object tracking: A Survey, *ACM Computing Surveys* 38(4): 1–44.

Zhao, H. & Shibasaki, R. (2005). 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.

Zhao, T., Nevatia, R. & Wu, B. (2008). Segmentation and tracking of multiple humans in crowded environments., *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(7): 1198–1211.

Laser Scanner Technology

Edited by Dr. J. Apolinar Munoz Rodriguez

Laser scanning technology plays an important role in the science and engineering arena. The aim of the scanning is usually to create a digital version of the object surface. Multiple scanning is sometimes performed via multiple cameras to obtain all slides of the scene under study. Usually, optical tests are used to elucidate the power of laser scanning technology in the modern industry and in the research laboratories. This book describes the recent contributions reported by laser scanning technology in different areas around the world. The main topics of laser scanning described in this volume include full body scanning, traffic management, 3D survey process, bridge monitoring, tracking of scanning, human sensing, three-dimensional modelling, glacier monitoring and digitizing heritage monuments.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds