Human Tracking using Floor Sensors based on the Markov Chain Monte Carlo Method

Takuya Murakita, Tetsushi Ikeda, and Hiroshi Ishiguro
Department of Adaptive Machine Systems, Osaka University
E-mail: murakita/ikeda/ishiguro@ed.ams.eng.osaka-u.ac.jp

Abstract

The aim of this paper is to develop a human tracking system that is resistant to environmental changes and covers wide area. Simply structured floor sensors are low-cost and can track people in a wide area. However, the sensor reading is discrete and missing; therefore, footsteps do not represent the precise location of a person. A Markov Chain Monte Carlo method (MCMC) is a promising tracking algorithm for these kinds of signals. We applied two prediction models to the MCMC: a linear Gaussian model and a highly nonlinear bipedal model. The Gaussian model was efficient in terms of computational cost while the bipedal model discriminated people more accurately than the Gaussian model. The Gaussian model can be used to track a number of people, and the Bipedal model can be used in situations where more accurate tracking is required.

1. Introduction

We describe a human tracking system that is fast, accurate, and resistant to environmental changes, which is required in fields such as security, traffic, surveillance, and so forth. Tracking systems employing ultrasonic sensors, infrared sensors, vision sensors, or other commonly used conventional sensors have been applied to these kinds of applications; however, they have several disadvantages such as a large detection area, costly architecture, and susceptibility to disturbance.

Floor pressure sensors avoid these disadvantages. However, little research using the sensors has been reported. For example, The ORL Active Floor [1] developed by Addlesee et al. and The Smart Floor [2] created by Orr and Abowd are used for human identification based on footprint features. Yet their research interest is human identification rather than human tracking. Morishita et al. [3] take a research approach that more closely resembles ours in developing the High Resolution Pressure Sensor. Yet the area (2.0 m by 2.0 m) is too small to perform human tracking.

In contrast to these floor sensors, ours have simpler structures and cover enough area (37 square meters) to track people. Lossy signals of the sensors were processed by the Markov Chain Monte Carlo method, which implemented two prediction models: a generic linear model, called the Gaussian model, and a highly nonlinear model referred to as the bipedal model. We carried out four comparative experiments to examine the tracking performance of these two models.

The experiments revealed that the two models can perfectly track a person who walks alone at an error of 58 cm. The Gaussian model required much less computation, while the bipedal model discriminated people more finely. We can apply the Gaussian model to track a number of people and the bipedal model in situations where more accurate tracking is required.

2. The tracking system

We installed the InfoFloor system VS-SS-F (Vstone Corporation, Osaka, Japan) that tracks people with 1140 detection units called “blocks” (Figure 1a). A block is an 18 cm by 18 cm binary pressure sensor. An example of signals is shown as red squares in Figure 1b. The 20 seconds history of walking signals of a person is illustrated. Ideally, there must be alternating footsteps corresponding to each foot of a person; however, there are several missing footsteps because the sensors are insulated by thick carpets.

Figure 1. The system architecture
The adjustment of the sensitivity of the binary sensor is not easy since there are people of various weights. We have to handle this discrete signal.

One possible tracking algorithm for discrete and missing signals is the Kalman filter; however, the advantage of the filter is limited to linear or nearly linear systems. The filter is not effective for extremely nonlinear movements such as a rapid turn of an object, because predictions by the filter strongly depend on past observations.

Human walking can be considered to be a Markov process since a position at a time step strongly depends on the position just before the time step rather than on many past positions. A Markov Chain Monte Carlo method (MCMC) is the best tracking algorithm for these kinds of signals. The advantage of the MCMC also appears when some observations have been lost. If an observation at a time step is lost and no particles hit their prediction, the particles are re-sampled assuming an observation of uniform distribution. Then generally, the predicted distribution becomes very complicated that cannot readily be expressed with mathematical formulation. The distribution approximated by the samples is the most probable estimation of human position. Consequently, the MCMC is very resistant to signal loss; therefore, it is suitable for our tracking with simplified floor sensors.

3. Implementation of the MCMC

There are various ways of formulating MCMC [4]. In the field of visual processing, Isard and Blake formulated the CONDENSATION algorithm [5]. We followed the algorithm because floor sensor signals can be considered to be binary images.

The prediction model was modeled in two ways. One of them was a generic linear model, which we call the Gaussian model. It was formulated as follows:

\[ \mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{v}_t + \mathbf{w}_t \]  

(1)

where \( \mathbf{x}_t \) denotes the positions of particles at time step \( t \), \( \mathbf{v}_t \) represents velocities that were calculated employing the past six footsteps:

\[ \mathbf{v}_t = \frac{1}{6} \sum_{i=0}^{5} (\mathbf{x}_{t-i} - \mathbf{x}_{t-i-1}) \]  

(2)

\( \mathbf{w}_t \) is Gaussian noise that has a variance \( \sigma^2 \).

The observation model for the Gaussian model was

\[ p_j(\mathbf{O}_t | \mathbf{x}_t) = \sum_{j=1}^{M} \mathcal{N}(\mathbf{\theta}_j, \sigma_{obs}) \]  

(3)

where the set \( \mathcal{O}_t = \{\mathbf{\theta}_1, \ldots, \mathbf{\theta}_M\} \) denotes observed M blocks and \( \mathcal{N}(\mathbf{\theta}_j, \sigma_{obs}) \) is a normal distribution whose mean is the \( j^{th} \) block \( \mathbf{\theta}_j \) and variance \( \sigma_{obs} \).

The other prediction model was a bipedal model which is highly specialized for human tracking. As shown in Figure 4c, the bipedal model has four kinds of prediction based on the fact that the patterns of sensor activation by a person fall into four categories:

A. Activation by right foot only
B. Activation by left foot only
C. Activation by both feet with right foot forward
D. Activation by both feet with left foot forward

Model C and D have two square distributions that predict blocks neighboring the currently activated block. Model A and B have an additional distribution, which we call the composite Gaussian distribution. It predicts the most advanced footstep, which is about to appear.

Square distribution, on the lattice coordinates with a unit being a block, was formulated as follows:

\[ \begin{align*}
A & : 0,2,1,1,0,0,2,0 \\
B & : \pi^2, \pi^2, \pi^2, \pi^2, \pi^2, \pi^2, \pi^2, \pi^2 \\
C & : \rho^2, \rho^2, \rho^2, \rho^2, \rho^2, \rho^2, \rho^2, \rho^2 \\
D & : \rho^2, \rho^2, \rho^2, \rho^2, \rho^2, \rho^2, \rho^2, \rho^2 \\
\end{align*} \]  

(4)

Table 1 The transition condition

<table>
<thead>
<tr>
<th>Internal model at time ( t-1 )</th>
<th>Observation in ( \alpha ) domain</th>
<th>Observation in ( \beta ) domain</th>
<th>Weight of the particles</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>*</td>
<td>*</td>
<td>( \rho^2 ) ( \pi^2 )</td>
</tr>
<tr>
<td>B</td>
<td>*</td>
<td>( \pi^2 ) ( \rho^2 )</td>
<td>A</td>
</tr>
<tr>
<td>C</td>
<td>*</td>
<td>( \rho^2 ) ( \rho^2 )</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>*</td>
<td>( \rho^2 )</td>
<td>B</td>
</tr>
</tbody>
</table>

1) Tracking will be stopped after 12 times iteration. 2) Impossible transition

Figure 2. The bipedal model
Composite Gaussian distribution was represented by a vector $\mathbf{r}$. When the vector is converted to polar coordinates, its radius and angle have Gaussian noise with variances $\sigma_r^2$, $\sigma_\theta^2$, respectively. The mean of the radius is taken over from parent particle, and the angle from the direction of the vector calculated with equation 2.

Human walk was represented by a 13 dimensional particle:

$$s_t = \begin{bmatrix} x_1^r & y_1^r & \cdots & x_n^r & y_n^r & \cdots & x_1^\theta & y_1^\theta & \cdots & x_n^\theta & y_n^\theta & \mathbf{r} \end{bmatrix}^T$$  \hspace{1cm} (4)

where $x_i^r$ or $y_i^r$ represents a mean of blocks currently activated by the most advanced footstep. $y_i^\alpha$ and $y_i^\beta$ denotes a mean of predecessor footsteps that handle the case where sensors are activated by both feet. $y_i^\alpha$ and $y_i^\beta$ is ignored in the model A and B. $\mathbf{r}$ indicates human strides. The correspondence of notation $\alpha$ and $\beta$ to the left or right foot is dependent on initialization.

A combination of prediction using square and composite Gaussian distributions is applied to the elements of the particle. For example, in the case of model A:

$$x_{t+1}^\alpha = x_t^\alpha + A(x_t^\alpha)$$
$$x_{t+1}^\beta = x_t^\beta + \mathbf{r}$$

Then the same observation model as equation 3 was employed.

$x_t^\alpha$, $x_t^\beta$, $y_t^\alpha$, $y_t^\beta$ have internal weights $\pi^\alpha$, $\pi^\beta$, $\rho^\alpha$, $\rho^\beta$ respectively. The weights determine a total weight $\pi$ according to table 1. Internal models are changed based on observations as shown in Figure 4b.

4. Experimental results

4.1. Tracking performance and the required number of particles for the MCMC

We carried out four experiments to examine the effectiveness of floor sensors and the MCMC as shown in section 4.1 through 4.3. Firstly, tracking accuracy vs. walking speeds was examined.

10 people walked about a 30 meter test course as shown in green in Figure 1b. They did three times at arbitrary speeds, and we obtained 30 test samples. Then, tracking accuracy was regarded as number of tracking errors. The result is shown in Figure 3. The parameter of the Gaussian model is the standard deviation (S.D.) $\sigma$ of Gaussian noise. The bipedal model has two parameters: $\sigma_r$, $\sigma_\theta$ of the composite Gaussian noise as illustrated in Figure 2a. In the experiment, these two models were examined at three values of parameters as chosen based on preliminary experiments. The proportion of two parameters of the bipedal model was fixed to $\sigma_\theta = 3\sigma_r$, in experiment 4.1 and 4.2.

Secondly, tracking accuracy and number of particles were examined employing the same data set. Number of tracking errors was summed to absorb the variances of walking speeds. The result is separately shown in Figure 4 according to values of parameters.

4.2. Estimation errors of human positions

It is very difficult to measure positions of people correctly; therefore, we assumed that test subjects walk along the test course precisely with uniform walking speed, and that an error of position can be decomposed into two directions: a transversal error and a longitudinal error. As shown in Figure 5, the transversal error is a lateral aberration from the test course, and the longitudinal error is a deviation from uniform walking in the direction of travel.

We examined the errors of 10 test subjects and averaged them (Figure 5). However, the errors at the corners of the test course were excluded because of difficulty in defining the errors.

Figure 3. Tracking errors vs. walking speeds

Figure 4. Tracking errors vs. number of particles
4.3. Discriminating people in a crowded area

The tracking system often fails to track people in a crowded area. Assuming that the system would be applied to everyday use, these kinds of situations are inevitable; therefore, it is necessary to examine to what extent the system can discriminate. We settle the problem as shown in Figure 6a.

There are two persons. One of them stands still during the experiment. The other person walks close to the still person. If the system could discriminate them, the two persons are continuously tracked. However, if not, two trackers may track only one of the two people. We examined 30 test samples with a variety of gaps D. The result is shown in Figure 6b.

5. Discussion and conclusion

Figure 3 shows that the tracking errors became 0 if $\sigma \geq 3.0$ [blocks] for the Gaussian model and $\sigma_r \geq 2.0$ [blocks] for the bipedal model.

However, larger variances require more particles to properly approximate a prediction distribution. If the number of particles is not enough, the distribution becomes sparse and tracking errors occur. Figure 4 demonstrates that the minimum number of particles to perfectly track a person was 3000, at $\sigma = 3.0$ [blocks] and $\sigma_r = 2.0$ [blocks]. These values coincide with the result of former experiment, because the minimum number of particles is required if the prediction distribution is most proper.

Figure 5 shows that the system can track people with a mean error of about 20 cm. An error of a tracking system using vision sensors examined by Sogo et al. [6] was 0.17 m at its maximum. The S.D. of the errors was 0.0393 m. In terms of computational efficiency, the estimation precision of the floor sensors is not worse than that of the DOVS, although the evaluation methods differ.

Figure 6 shows that the floor sensors discriminated two people perfectly, if the gap D are larger than 1.42 m. Even if the interval is only 0.8 m, the bipedal model discriminated two people at nearly 90% accuracy. For every interval smaller than 1.4 m, the bipedal model discriminated more effectively than the Gaussian model.

Based on these experimental results, we could make the following conclusions:

(1) The system can perfectly track a person who walks alone.
(2) The system can track a person within 59 cm error.
(3) The system can independently track people at 90% accuracy, if they maintain a gap of more than 80 cm.

These facts indicate that the system will work in everyday use.

6. References


