Development of the Automatic Human Tracking Method by Integration of Pedestrian Behavioral Model

歩行者行動モデルを統合した人物自動追跡手法の開発

NAKANISHI Wataru / 中西 航

PhD Student
Regional Planning and Information Lab.
Dept. of Civil Engineering, The University of Tokyo
nakanishi@trip.t.u-tokyo.ac.jp

Background

- Importance of microscopic understanding of pedestrian behavior for sophisticated space design and flow control
  - Detail passenger flow leads to good station improvement works

- Automatic human tracking is difficult under situations that people are occluded by others or close to each other

Background and Purpose

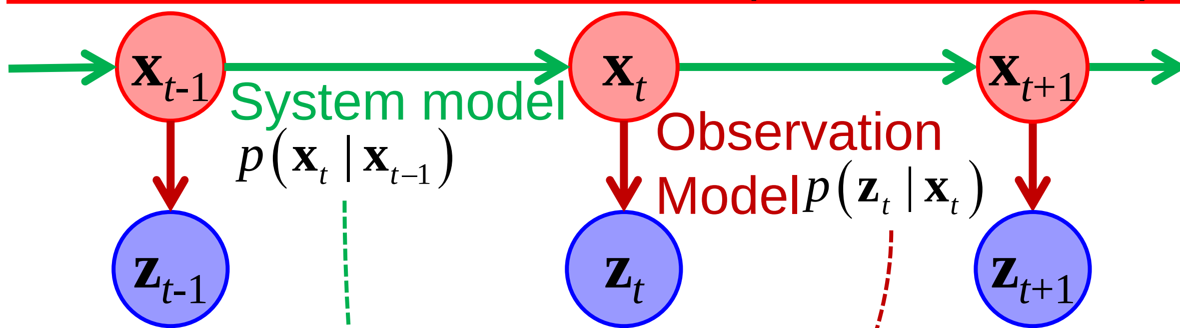
- Approach to understanding pedestrian behavior
  - Sensing and imaging technology
    ::Human tracking using color and range information
  - Modeling and simulation
    ::Microscopic behavior model considering interaction

- To achieve multiple human tracking under complex situations by integrating observation and simulation as data assimilation process

General Space State Model

- Effective way to stochastically combine observations and simulations

$\mathbf{x}_t$: State Vector at time t, cannot be observed
: Human position and shape



$\mathbf{z}_t$: Observation Vector at time t, can be observed
: 3D Range and color

- After observation $\mathbf{z}_t$, estimating $\mathbf{x}_t$ by MAP estimation
::Optimal $p(\mathbf{x}_t | \mathbf{z}_{1:t})$

$$p(\mathbf{x}_t \mid \mathbf{z}_{1:t}) \propto p(\mathbf{z}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{z}_{1:t-1})$$
$$= \boxed{p(\mathbf{z}_t \mid \mathbf{x}_t)} \int \boxed{p(\mathbf{x}_t \mid \mathbf{x}_{t-1})}\, p(\mathbf{x}_{t-1} \mid \mathbf{z}_{1:t-1})\, d\mathbf{x}_{t-1}$$

Particle Filter

$$p\left(\mathbf{x}_t \mid \mathbf{z}_{1:t}\right) = p\left(\mathbf{z}_t \mid \mathbf{x}_t\right) \int p\left(\mathbf{x}_t \mid \mathbf{x}_{t-1}\right) p\left(\mathbf{x}_{t-1} \mid \mathbf{z}_{1:t-1}\right) d\mathbf{x}_{t-1}$$

$p\left(\mathbf{x}_{t-1} \mid \mathbf{z}_{1:t-1}\right)$

Approximate p.d.f. by weighted particles

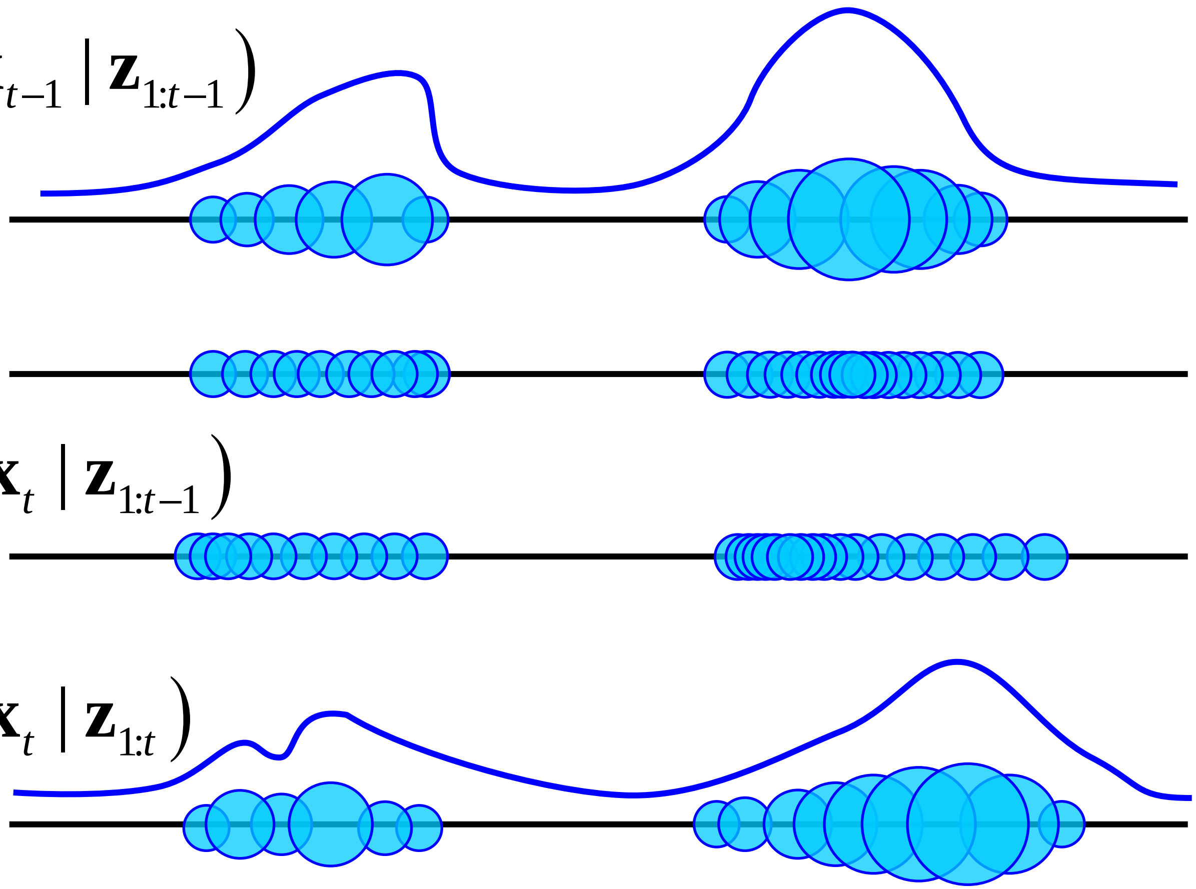
Resampling

$p\left(\mathbf{x}_t \mid \mathbf{z}_{1:t-1}\right)$

Move particles by $p\left(\mathbf{x}_t \mid \mathbf{x}_{t-1}\right)$

$p\left(\mathbf{x}_t \mid \mathbf{z}_{1:t}\right)$

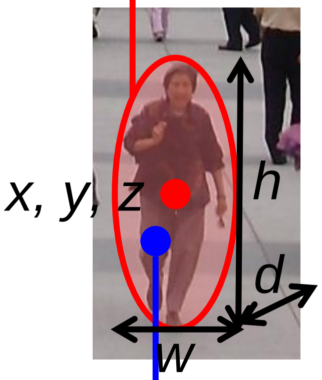
Weight particles by $p\left(\mathbf{z}_t \mid \mathbf{x}_t\right)$



Expected value of this particles: Estimated value of $\mathbf{x}_t$

State Vector / Observation Vector

One ellipsoid



x$_t$: State Vector at time *t*, cannot be observed
: Human position and shape

- Defined as one ellipsoid

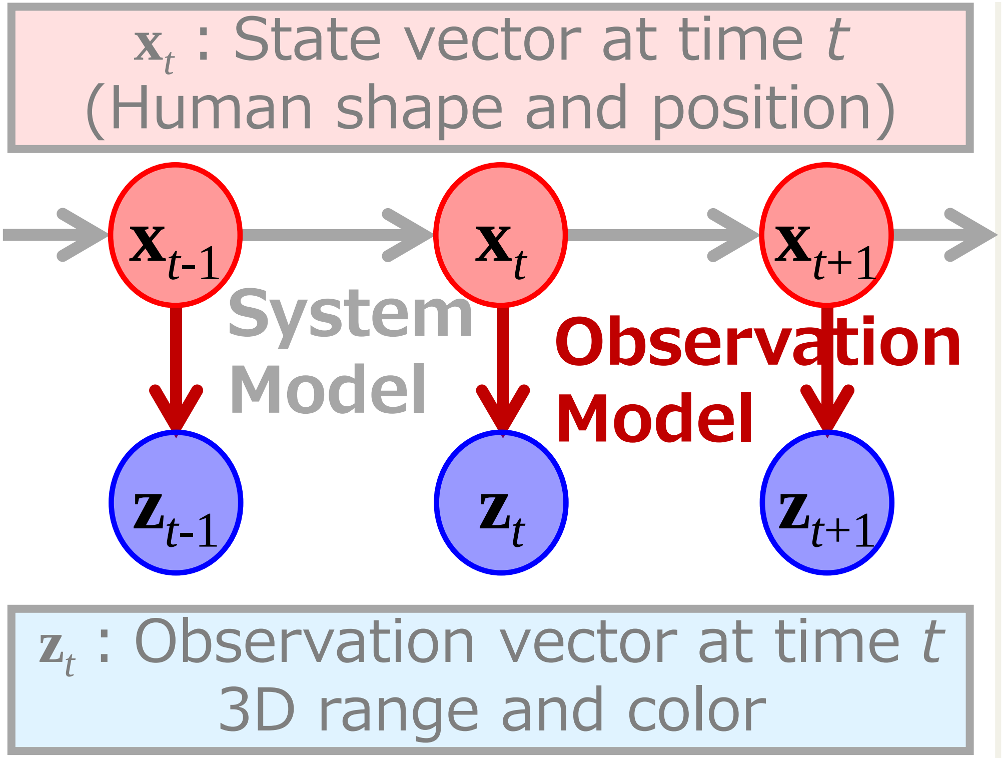
$$\mathbf{x}_t = (x_t, y_t, z_t, w_t, h_t, d_t)$$

z$_t$: Observation Vector at time *t*, can be observed
: 3D Range and color

- Defined as 3D coordinates (*X*, *Y*, *Z*) and color values (*r*, *g*, *b*) at each pixel

$$\mathbf{z}_{ij,t} = \left(X_{ij,t}, Y_{ij,t}, Z_{ij,t}, r_{ij,t}, g_{ij,t}, b_{ij,t} \right)$$

Observation Model $p(\mathbf{z}_t \mid \mathbf{x}_t)$



Probability distribution
of observation vector
on condition of state vector

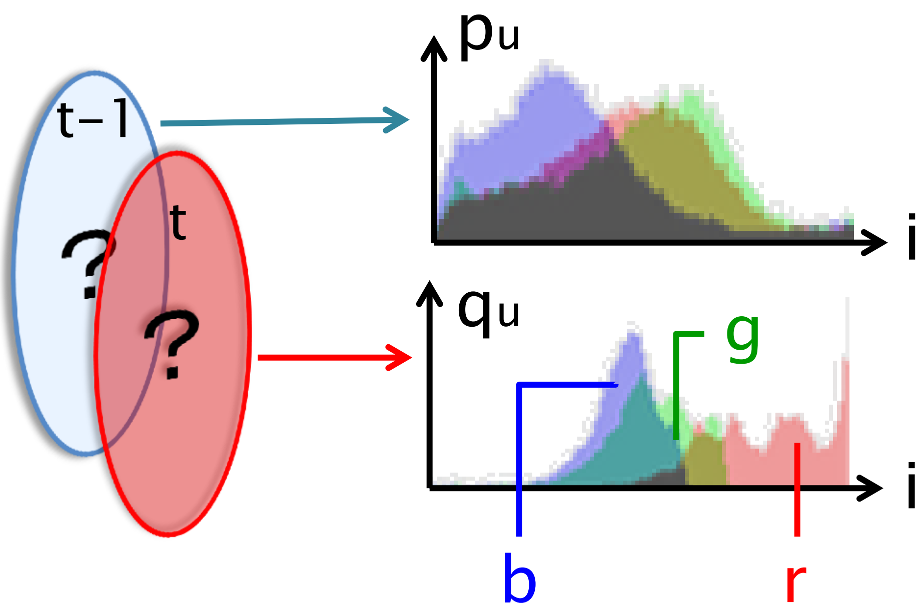
→
Corresponding to existing
tracking method

- Model from both color and range observation

$$p(\mathbf{z}_t \mid \mathbf{x}_t) = \underbrace{p\left(\mathbf{z}_t^{color} \mid \mathbf{x}_t\right)}_{\text{Color model}} \cdot \underbrace{p\left(\mathbf{z}_t^{range} \mid \mathbf{x}_t\right)}_{\text{Range model}}$$

Color Observation Model

- Color Histogram Matching between successive frames
- Compare similarity of histograms of two ellipsoids



Bhattacharyya coefficient

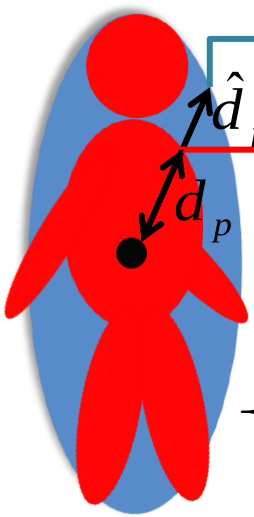
$$\rho = \sum_{i=0}^{255} \sqrt{p_u q_u}$$

- Product of each color, *r*, *g*, and *b*

$$p\left(\mathbf{z}_t^{color} \mid \mathbf{x}_t\right) = \prod_{r,g,b} \sum_{i=0}^{255} \sqrt{p_u q_u}$$

Range Observation Model

- Shape Matching between estimated ellipsoid and observed points



$$\left(\frac{x-x_t}{w/2}\right)^2+\left(\frac{y-y_t}{h/2}\right)^2+\left(\frac{z-z_t}{d/2}\right)^2=1$$

$$\left(X_{ij,t}, Y_{ij,t}, Z_{ij,t}\right)$$

Distance from center of ellipsoid to
::observed point d_p
::surface of ellipsoid $\hat{d}_p$

Evaluation equation

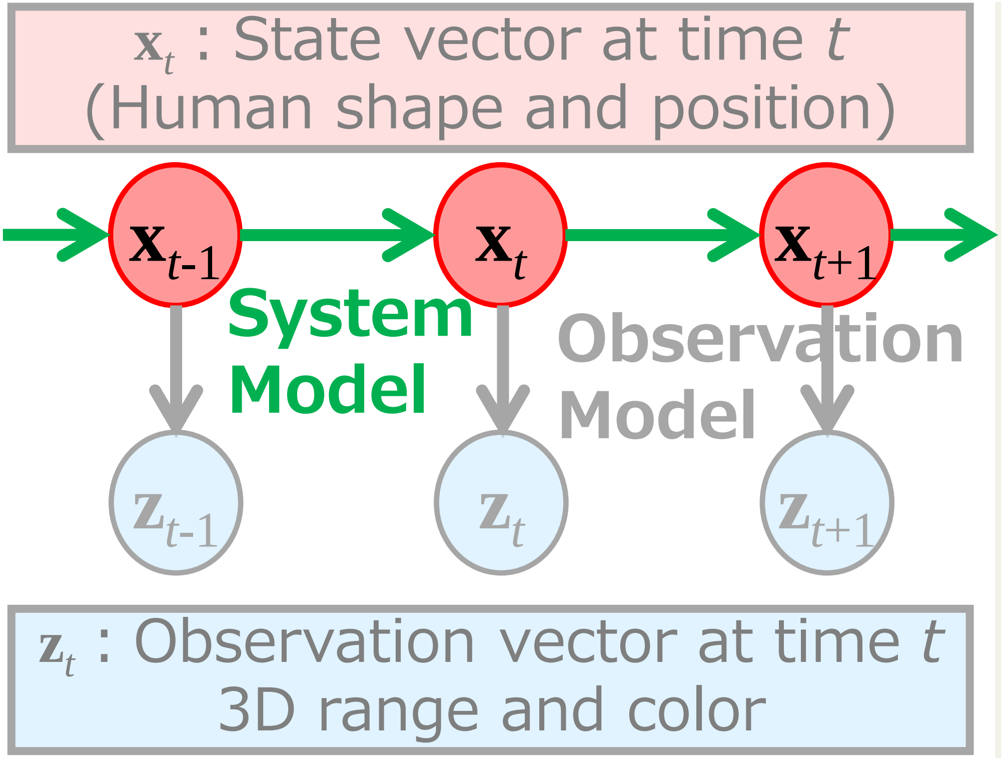
$$1-\frac{1}{I}\sum_p\left(d_p-\hat{d}_p\right)^2$$

(I: # of observed points)

- p: Observed points inside ellipsoid are calculated

$$p\left(\mathbf{z}_t^{range} \mid \mathbf{x}_t\right)=1-\frac{1}{I}\sum_p\left(d_p-\hat{d}_p\right)^2$$

System Model $p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$



Time-serious change of state vector

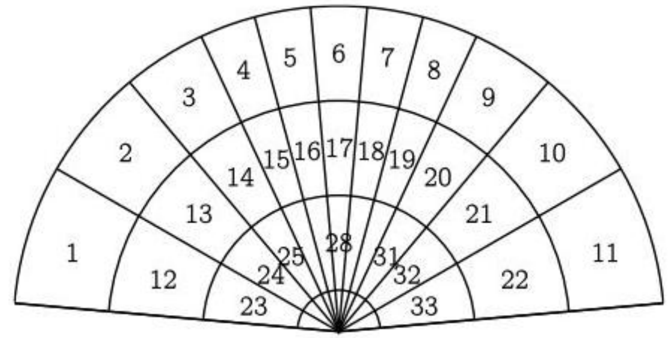
→

Corresponding to pedestrian behavioral model

- Most of existing tracking method regard time-serious change as random walk or uniform motion
- Introducing non-linear behavioral model to tracking method is feature of this research

System Model $p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$

- Discrete choice model by Robin *et al*. (2009)
- 33 choice set for each pedestrian for next step



- Utility function:

$$
\begin{aligned}
V_{vdn} = {} & \left.\begin{aligned} & \beta_{\text{dir\_central}} \text{dir}_{dn} I_{d,\text{central}} \\ & + \beta_{\text{dir\_side}} \text{dir}_{dn} I_{d,\text{side}} \\ & + \beta_{\text{dir\_extreme}} \text{dir}_{dn} I_{d,\text{extreme}} \end{aligned}\right\} \quad \text{keep direction} \\
& \left.\begin{aligned} & + \beta_{\text{ddist}} \text{ddist}_{vdn} \\ & + \beta_{\text{ddir}} \text{ddir}_{dn} \end{aligned}\right\} \quad \text{toward destination} \\
& \left.\begin{aligned} & + \beta_{\text{dec}} I_{v,\text{dec}} (v_n / v_{\max})^{\lambda_{\text{dec}}} \\ & + \beta_{\text{accLS}} I_{n,\text{LS}} I_{v,\text{acc}} (v_n / v_{\text{maxLS}})^{\lambda_{\text{accLS}}} \\ & + \beta_{\text{accHS}} I_{n,\text{HS}} I_{v,\text{acc}} (v_n / v_{\max})^{\lambda_{\text{accHS}}} \end{aligned}\right\} \quad \text{free flow acceleration} \\
& \left.\begin{aligned} & + I_{v,\text{acc}} I^{\text{L}}_{d,\text{acc}} \alpha^{\text{L}}_{\text{acc}} D_{\text{L}}^{\rho^{\text{L}}_{\text{acc}}} \Delta v_{\text{L}}^{\gamma^{\text{L}}_{\text{acc}}} \Delta \theta_{\text{L}}^{\delta^{\text{L}}_{\text{acc}}} \\ & + I_{v,\text{dec}} I^{\text{L}}_{d,\text{dec}} \alpha^{\text{L}}_{\text{dec}} D_{\text{L}}^{\rho^{\text{L}}_{\text{dec}}} \Delta v_{\text{L}}^{\gamma^{\text{L}}_{\text{dec}}} \Delta \theta_{\text{L}}^{\delta^{\text{L}}_{\text{dec}}} \end{aligned}\right\} \quad \text{leader—follower} \\
& \left. + I_{d,\text{C}} \alpha_{\text{C}} e^{\rho_{\text{C}} D_{\text{C}}} \Delta v_{\text{C}}^{\gamma_{\text{C}}} \Delta \theta_{\text{C}}^{\delta_{\text{C}}} \right\} \quad \text{collision avoidance}
\end{aligned}
$$

Unknown while tracking
→omit this terms

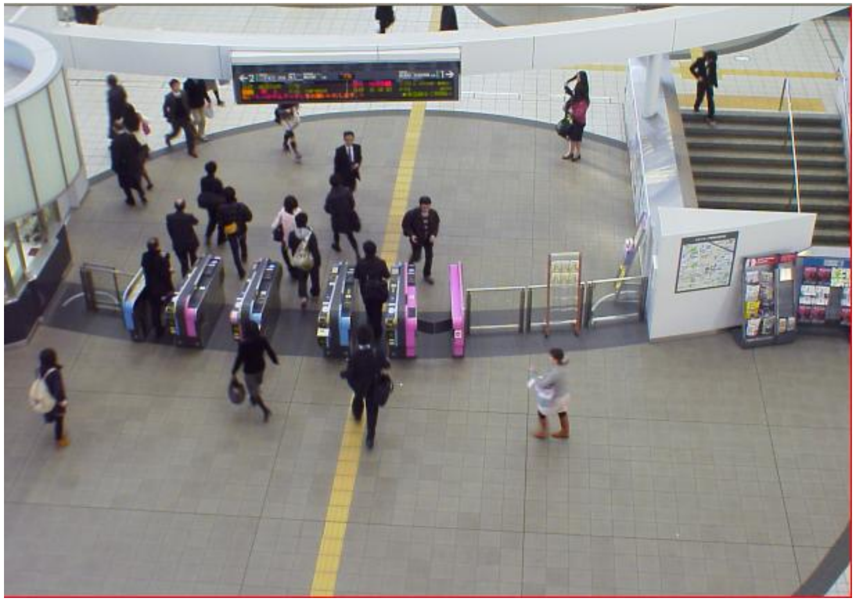
Interaction
between pedestrians

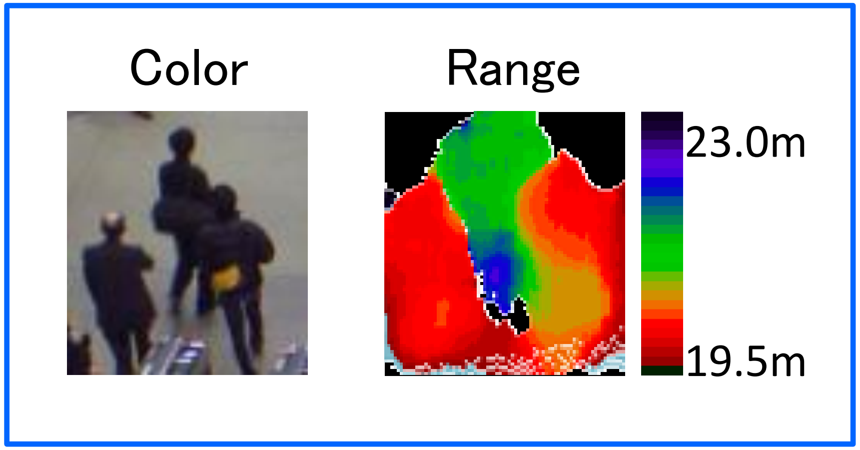
- Cross nested logit, nests for velocity and angle

Application

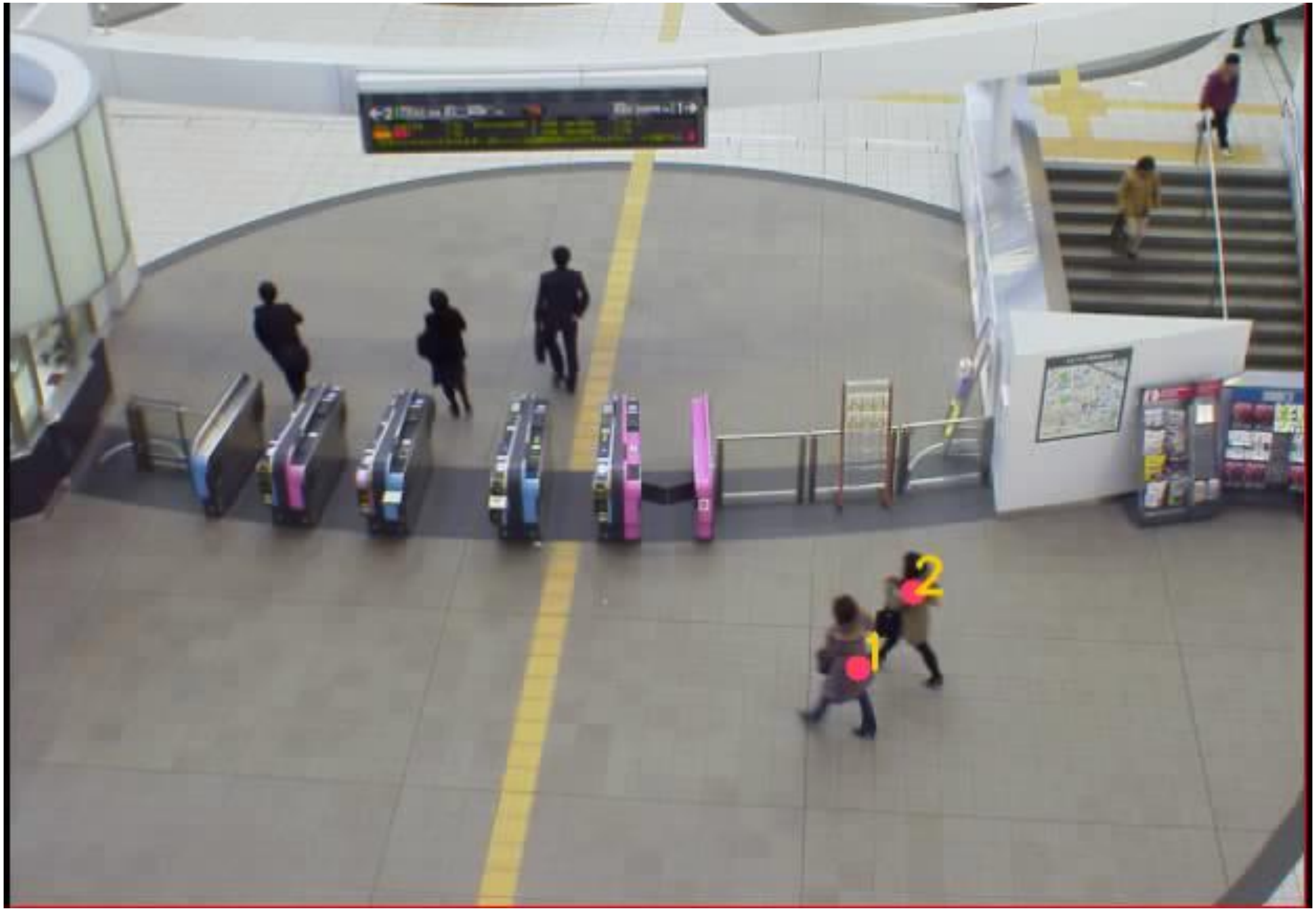
- Commuter rush hour at Tama-Plaza Station, about 30km west from central Tokyo
  - Passengers move interdependently

- Condition of stereo video camera
  - Frame rate 7.5 fps
  - Baseline 1m



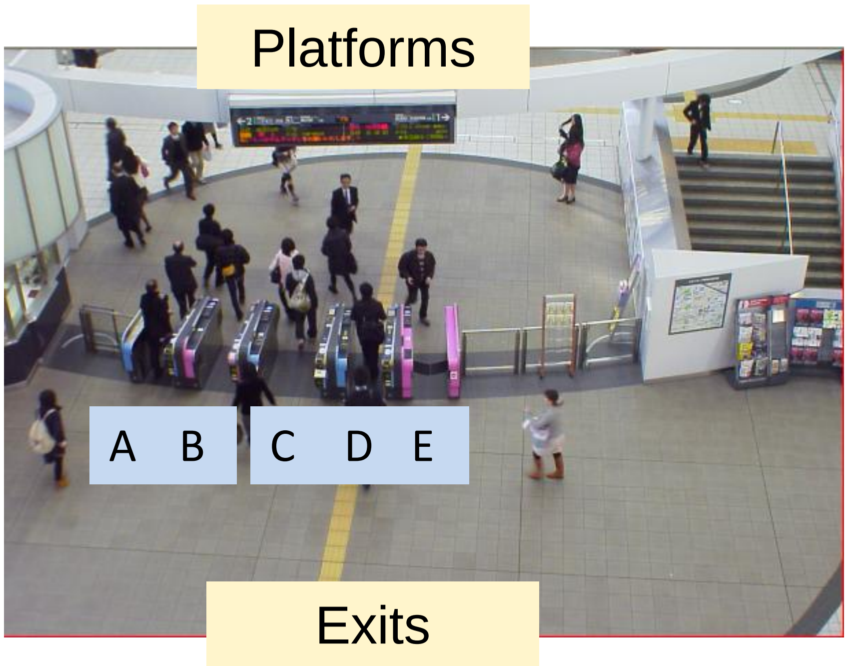


Application

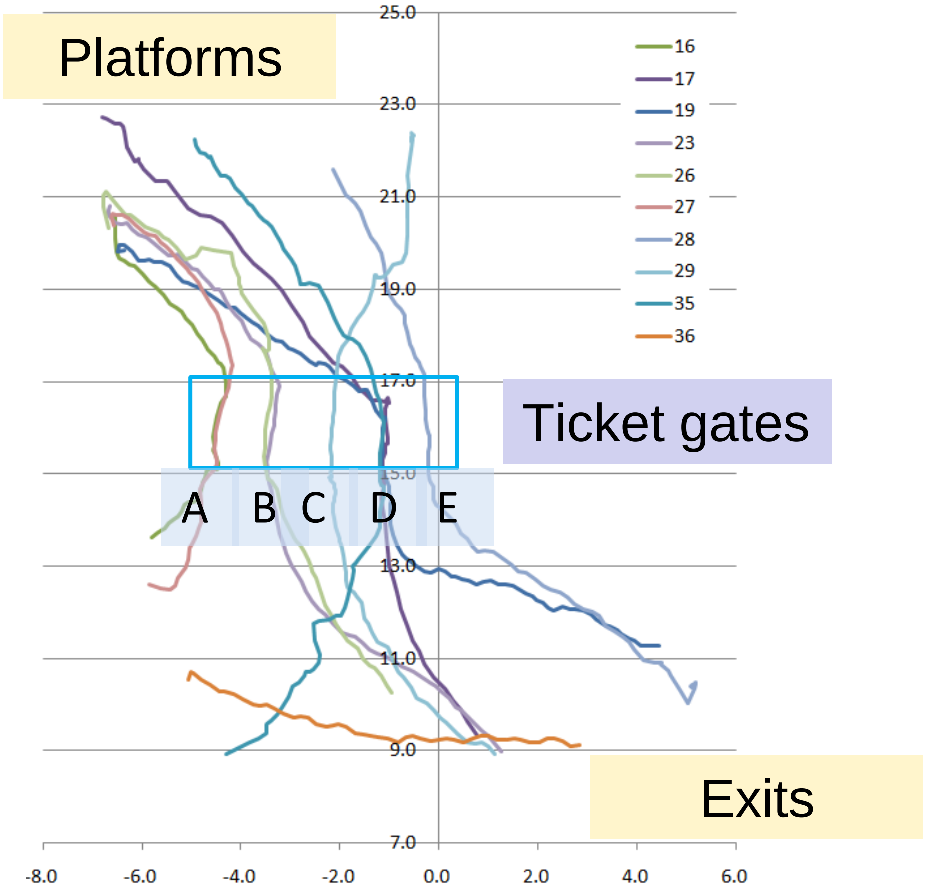


Application

- Automatic acquisition of
  - Position and velocity
  - Passenger flow
  - Ticket gate choice



Part of passenger flow data



How to improve model?

- What is better observation model?
- What is better system model?

- From tracking result, optimize likelihood…

$$
\begin{aligned}
p(\mathbf{z}_{1:T}) &= p(\mathbf{z}_T \mid \mathbf{z}_{1:T-1})\, p(\mathbf{z}_{1:T-1}) \\
&= \prod_{t=1}^{T} \boxed{p(\mathbf{z}_t \mid \mathbf{z}_{1:t-1})} \\
\log p(\mathbf{z}_{1:T}) &= \log\left\{ p(\mathbf{z}_T \mid \mathbf{z}_{1:T-1})\, p(\mathbf{z}_{1:T-1}) \right\} \\
&= \sum_{t=1}^{T} p(\mathbf{z}_t \mid \mathbf{z}_{1:t-1})
\end{aligned}
$$

$$
\begin{aligned}
p(\mathbf{x}_t \mid \mathbf{z}_{1:t}) &= \frac{p(\mathbf{z}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{z}_{1:t-1})}{\boxed{p(\mathbf{z}_t \mid \mathbf{z}_{1:t-1})}} \\
&= \frac{p(\mathbf{z}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{z}_{1:t-1})}{\boxed{\int p(\mathbf{z}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{z}_{1:t-1})\, d\mathbf{x}_t}}
\end{aligned}
$$

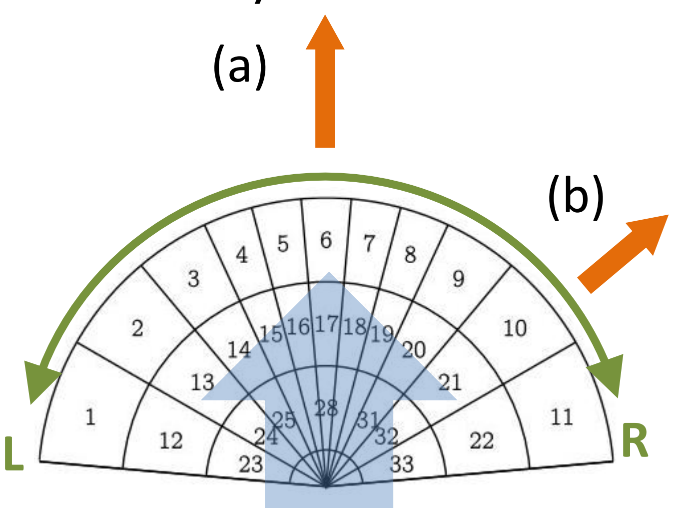
$$
(\propto p(\mathbf{z}_t \mid \mathbf{x}_t)\, p(\mathbf{x}_t \mid \mathbf{z}_{1:t-1}))
$$

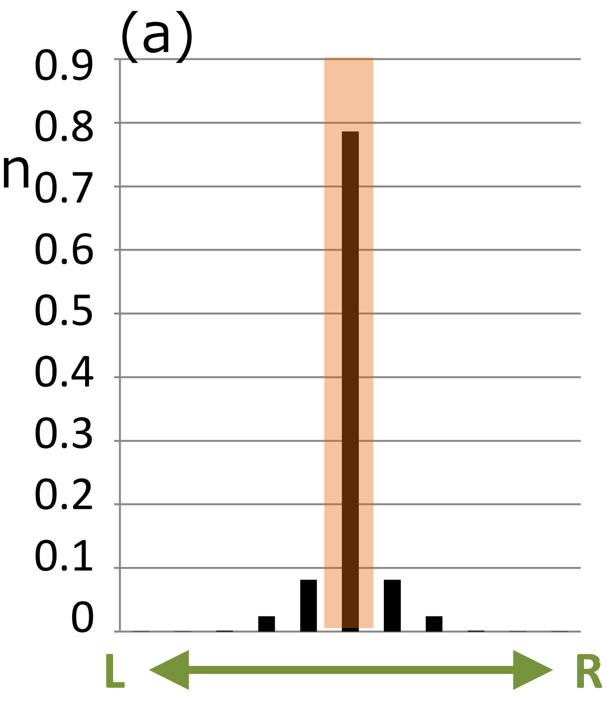
- For any parameter ϑ in the model, we can optimize $l(\theta) = \log p(\mathbf{z}_{1:T} \mid \theta) = \sum_{t=1}^{T} \log p(\mathbf{z}_t \mid \mathbf{z}_{1:t-1}, \theta)$
- We also have to consider model structure itself

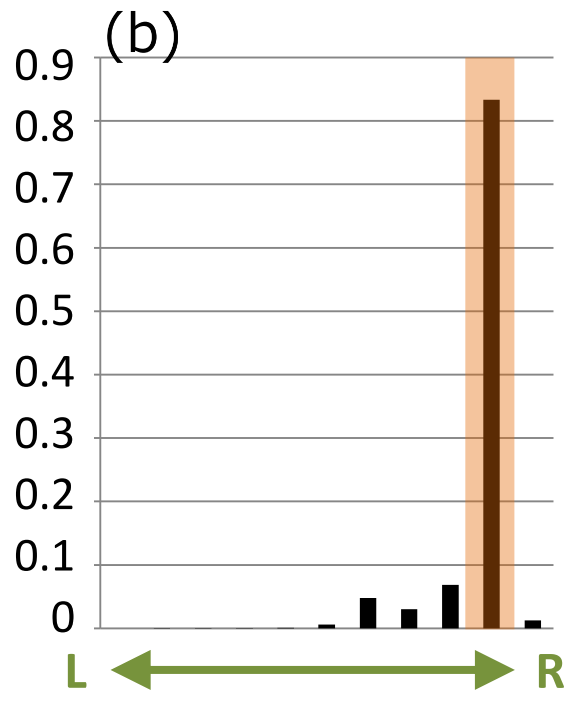
Consideration of System Model

- Problems of Simulation model
  - In the model from Robin *et al.*, the term “toward destination” makes much effect than other terms, although we cannot know destination while tracking
  - Generally in actual situation we don’t decide our destination every one second or shorter time step, even if the model reproduce their behavior correctly

Destination is set
30m away in each direction







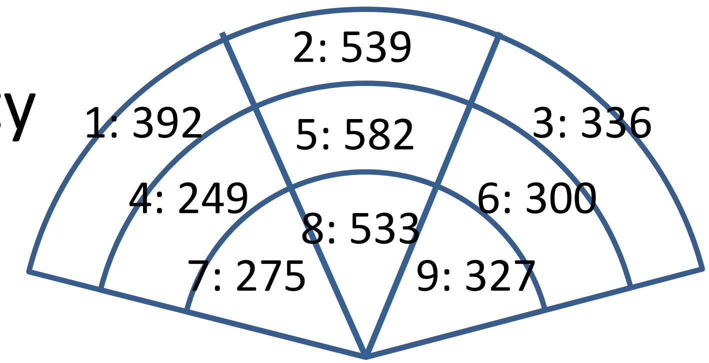
Consideration of System Model

- Consideration of general state space model
  - From viewpoint of observation, pedestrian position fluctuates if observation rate is high enough
  - However observation cannot be smoothed because there is observation model
- System model is not the same as simulation model
  - System model needs to describe only probability distribution of next step on condition of present situation
- This may bring about knowledge also for simulation model
  - Actual structure of human behavior (time step and choice etc.)
  - Relationship between condition of field and human behavior

Trial (Setting)

- Dataset: Manually acquired at station (mentioned before)
- Time step: 0.27[s]
- Choice set: 3 for angle, 3 for velocity
- Form: MNL

Total: 3533



$$V_i = \beta_{dist} dist_i$$ Distance to ticket gate (chosen in the future)

$$+\beta_{angle} angle_i$$ 30 if candidate 1, 3, 4, 6, 7 and 9 otherwise 0

$$+\beta_{acc} I_{acc} \left(\frac{v_i}{vmax} \right)^{\lambda} + \beta_{dec} I_{dec} \left(\frac{v_i}{vmax} \right)^{\lambda}$$

I_{acc} = 1 if candidate 1 to 3
I_{dec} = 1 if candidate 7 to 9
v_{max} = 4.87m/s
λ = 2.4 (from previous research)
v_i : speed vector present

$$+\beta_{leader} leader_i$$

$$+\beta_{collider} collider_i$$

Sum of distance between each candidate *i* and leader/follower *j*, which are within $5v_i$ from present position and if angle between v_i and v_j are within 90° then *j* is leader, otherwise collider

Trial (Result)

- Both leader and collider are positive?
  - There is no difference?
  - Setting of "*collider*" is not good?
- Destination is not significant enough
  - Difference is slight for each candidate?
  - Passengers never change ticket gate choice within this time step?
- Velocity is dominant

	estimated value	t-value
β_{dist}	-0.336	-1.71
β_{angle}	-0.0193	-16.6
β_{acc}	-11.6	-10.3
β_{dec}	4.10	5.71
β_{leader}	1.01	5.80
$\beta_{collider}$	1.19	5.89

- How to set parameters of <u>system model</u> from this result?
  - Maybe needs smoothing (e.g. interpolation) for any point between candidates

Future Works

- Setting system model and running tracking
- Comparison between some conditions
  - Tracking result
  - Feeding back to parameter values and model structure
- Development of the manner to segment condition of field, and to detect structural changes
  - On-line estimation of parameters (e.g. dynamic parameter)
  - On-line detection of structural changes