Simulation-Based Behavior Tracking of Pedestrians in Partially Observed Indoor Environments

Arsène Fansi T.[⋆,∓], Vincent Thomas[⋆], Olivier Buffet[⋆], Fabien Flacher[∓], Alain Dutech[⋆]
[⋆]Université de Lorraine, INRIA - LORIA, Nancy, France
[∓]Thales Services, 20-22 rue Grange Dame Rose, 78140 Vélizy, France
{arsene.fansi-tchango, vincent.thomas, olivier.buffet, alain.dutech}@loria.fr
fabien.flacher@thalesgroup.com

ABSTRACT

Tracking and understanding moving pedestrian behaviors is of major concern for a growing number of applications. This problem, known as difficult, is more complex when the considered environment is not fully under sensory coverage. Classical approaches either focus on location estimation or attempt to build the relationship between possible activities in the environment and reason on it, which may turn out to be inadequate. In this paper, we propose an approach based on behavioral models from the situated artificial intelligence field, which aim to realistically reproduce human behaviors within complex environments. We focus on the special case of a single target and experimentally show that we are performing well even in case of long periods of occlusion.

Keywords

Behavior Tracking - Autonomous Agent-based simulations - Situated Artificial Intelligence - Particle filter

1. INTRODUCTION

The ability of using sensor networks to track and understand the behavior of moving human beings is of great importance for surveillance applications. When sensors are cameras, this implies retrieving behavioral information from image analysis. This is not a trivial task to perform, even for humans, and interpretation errors are common.

Moving pedestrians are usually driven by an inner motivation in relation with the activity they want to perform. Therefore, location and intent (motivation) are contextually dependent, and the knowledge of one may help predict the other. In the literature, people tracking has usually been addressed using Bayesian filters [2], which require the use of models representing the target's behavior. It is a common consensus that the more realistic the models used, the better the tracking results.

Nowadays, the use of simulators to generate realistic human behaviors in indoor environments has become very popular. Works from the situated artificial intelligence field have focused on designing virtual agent control architectures (behavior models) based on sensorimotor loops [1] whose purpose is to define, at any time, the action an agent will perform according to his perceptions, thus making him adaptive and autonomous in fulfilling a sequence of activities (subgoals) aiming at a particular goal.

In this paper, we consider integrating simulators implementing such architectures within a Bayesian filter (approximated by a particle filter) for the analysis of people intents. We focus on the single-target case and experimentally show that we are able to infer relevant information regarding the target behavior even with long periods of occlusion.

Appears in: *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), Lomuscio, Scerri, Bazzan, Huhns (eds.), May, 5–9, 2014, Paris, France.*

2. AGENT-BASED BEHAVIOR TRACKING

2.1 Process Overview

The proposed solution is represented in Fig. 1. The simulator is assumed to handle virtual-agent navigational features and object-agent interactions (e.g., escalators). It is used during the prediction step to estimate the belief regarding both the aimed location and the activity of the target.

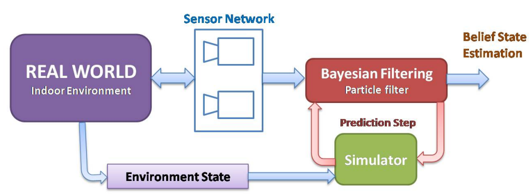


Figure 1: Process overview

Observations received from the sensors are typically noisy location data of the detected target. Furthermore, the simulator is fed with changes occurring in the real world (e.g., escalator failures, fire alerts) in order to preserve coherence with the simulated environment.

Next, we describe the models used in the filter.

2.2 Models

2.2.1 *System Dynamics*

Given a behavioral model, a state $\mathbf{x}_t$ of an agent contains all the attributes that are taken into account during the decision-making process. We assume that an agent cannot change the state of objects he may interact with. The system dynamics is fully encoded within the simulator. The latter takes as input the agent state $\mathbf{x}_t$ and the environment state $\mathbf{E}_t$ and modifies the agent's inner attributes, that is

$$\mathbf{x}_{t+1} \sim f(\mathbf{x}_t, \mathbf{E}_t),$$

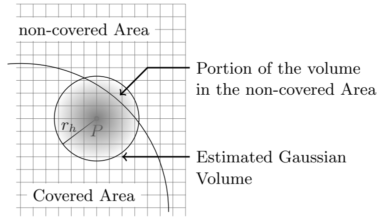


Figure 2: Approximation of φ: P is the agent position. The ratio is computed with respect to all cells in the considered volume.

where f is the implemented simulator stochastic function.

2.2.2 *Observation Model*

The data $\mathbf{z}_t$ received from the sensors depends on whether the agent is within a covered area or not. However, assuming a Gaussian noise with covariance matrix Q_v, the agent may still be undetected even within covered areas, especially when he is close to the boundaries of both areas. The probability φ of such an event can be approximated as the portion of the Gaussian (represented by a circle of radius r_h) belonging to the non-covered areas (Fig. 2). That is

$$\hat{\varphi} = \frac{\sum \text{prob. of region's cells in non-covered area}}{\sum \text{prob. of all region's cells}}. \quad (1)$$

Finally, we have

$$p(\mathbf{z}_t|\mathbf{x}_t) = \begin{cases} 1 & \text{if } \mathbf{z}_t = \varnothing \text{ and } unc(\mathbf{x}_t), \\ \hat{\varphi} & \text{if } \mathbf{z}_t = \varnothing \text{ and } \neg unc(\mathbf{x}_t), \\ \mathcal{N}_{0;Q_v}(\mathbf{u}_t) & \text{if not,} \end{cases}$$

where $\mathbf{u}_t = \mathbf{z}_t - h(\mathbf{x}_t)$, h is a function extracting the location data from $\mathbf{x}_t$, and $unc(\mathbf{x}_t)$ indicates that the agent ($\mathbf{x}_t$) is within a non-covered area. $\varnothing$ means that the agent is undetected.

3. EXPERIMENTAL EVALUATION

The simulator used is SE-Star, a Thales proprietary engine capable of modeling adaptive behaviors, navigations and interactions with objects. We conducted experiments in a virtual environment representing a two-floors subway station (Fig. 3). The station contains a train door, an ATM (in light green), a ticket machine (yellow), a beverage dispenser (brown), ticket barriers (white) and exit barriers (red). Also, it is equipped with a camera network set up to not completely cover the environment and configured to report noisy passenger location data.

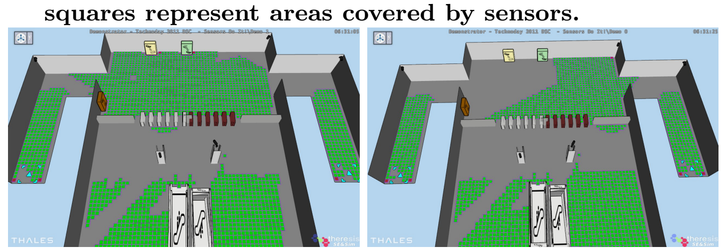
We consider two scenarios. In the first one, the cameras are set to cover areas occupied by the ATM, the ticket machine and the beverage dispenser. In the second scenario, these objects are no longer covered by the sensors (Fig. 3) in order to assess the robustness of our approach.

We consider that a passenger is motivated by three main goals: taking train, drinking or leaving. Also, he may initially own a certain amount of money and/or a valid ticket.

Such a passenger model has been designed in SE-Star in which the three motivations are represented by numerical attributes taking values in $[0, 2]$ and the simulator is in charge of their evolution. Also, the model includes two attributes representing respectively the the amount of money and the number of tickets owned.

We used 2000 particles and set r_h to be $0.5m$. The noise standard deviation is set to 0.8, 0.8 and $0.05m$ for the x, y and z coordinates respectively. Initially, a passenger has

Figure 3: Scenario 1 (left) and 2 (right): green squares represent areas covered by sensors.



30% chances to own a number of tickets (an amount of money) chosen uniformly in the interval $[1, 3]$ ($[1, 5]$) and nothing otherwise. For other attributes, we assume a Gaussian distribution $\mathcal{N}(0.75; 0.5)$.

For the evaluation, the criteria used are: **similarity** (is the goal/subgoal inferred from the filter the real one?), **robustness** (does the filter contain a valid hypothesis regarding the passenger goal/subgoal?) and the **mean square error** (goal/subgoal based) with respect to the passenger location. Results are reported in Table 1 and Table 2.

Table 1: Similarity and Robustness

	% Time uncov.	% Goal Sim.	% Goal Rob.	% Subgoal Sim.	% Subgoal Rob.
Sc. 1	32.7	94.31 (± 6.76)	99.17 (± 0.72)	83.26 (± 7.82)	95.02 (± 6.62)
Sc. 2	67.7	90.25 (± 12.27)	99.46 (± 0.26)	72.41 (± 12.23)	92.20 (± 9.42)

Table 2: Mean Square Errors

	General MSE	Goal-Based MSE	SubGoal-Based MSE
Sc. 1	3.08 (± 2.04)	2.35 (± 1.46)	2.27 (± 1.48)
Sc. 2	4.29 (± 3.48)	4.78 (± 3.45)	4.49 (± 3.74)

For scenario 1, in which the passenger is invisible 1/3 of the time, the algorithm has relatively high average scores with low standard deviations, regardless the criterion. In scenario 2, a degradation of performance can be observed. This is due to the fact that most interactions with objects (subgoal's satisfaction) do not occur under sensor coverage. However, despite these significant occlusions (2/3 of the time), the approach was still efficient and kept good estimates of the target's goal and subgoal.

4. CONCLUSION

In this paper we proposed a novel approach to track pedestrian behaviors in partially observed environments in which the tracked entity may spend some time in non-covered areas. We used a particle filter for inference purposes and relied on behavioral models from the situated artificial intelligence field as a basis for the behavior analysis, thus making our approach adequate for easily handling dynamic changes in the environment.

5. REFERENCES

[1] J.-A. Meyer. From natural to artificial life: Biomimetic mechanisms in animat designs. *Robotics and Autonomous Systems*, 22:3–21, 1997.

[2] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers. People tracking with a mobile robot using sample-based joint probabilistic data association filters. *Int. Journal of Robotics Research*, 22(2):99–116, 2003.