Spatiotemporal models for motion planning in human populated environments

Tomáš Vintr, Sergi Molina, Gregorz Cielniak, Tom Duckett, Tomáš Krajník

Abstract— **In this paper we present an effective spatio-temporal model for motion planning computed using a novel representation known as the temporary warp space-hypertime continuum. Such a model is suitable for robots that are expected to be helpful to humans in their natural environments. This method allows to capture natural periodicities of human behavior by adding additional time dimensions. The model created thus represents the temporal structure of the human habits within a given space and can be analysed using regular analytical methods. We visualize the results on a real-world dataset using heatmaps.**

I. Introduction

As service robots are supposed to operate in human-populated environments, they have to be able to navigate in ever-changing worlds with dynamics defined by human movements and actions. Moreover, service robots are expected to be helpful to the humans and therefore need to predict human behavior in order to navigate to places populated by people who require its services. For example, a mobile information terminal should offer information about a lunch menu at a cafeteria entrance just before the people appear for lunch. Another example is a conference guide robot that has to provide directions to certain events at locations where people commonly gather before the events. In both cases, the robot needs to be at the right place just before the people appear there, so it does not have to struggle navigating in a moving crowd. These locations and times can be provided by an expert, but a truly intelligent robot should be able to alter and refine these based on its experience, or learn them itself by observing people and creating spatio-temporal models of their activity patterns.

Several approaches to spatiotemporal models for human populated areas have been proposed. The approach in [1] and [2] extends a classic occupancy grid by adding a temporal model for each cell. The system models the cell occupancies by detected humans using a set of harmonic functions that capture long-term patterns of the changes observed. The resulting representation can then predict the presence of humans in each cell for any given time.

In [3] the space is also represented as an occupancy grid and the model considers the probability functions associated with the neighboring cells. This function predicts the probability of the most likely direction of dynamic objects traversing the modeled cell. In [4] this concept is extended and a continuous model estimated by Expectation Maximization is proposed. Moreover, [4] suggests to use such a probabilistic directional map to navigate robots through the space in accordance with the directions chosen by the humans. A similar approach for model generation is proposed in [5], where the direction of traversal through each cell is obtained using an input-output hidden Markov model connected to the neighboring cells.

In [6], the authors argue that continuous models are of better use for robot navigation than discrete ones. Their dataset is generated by a range-finder sensor, which identifies not only the occurrence of humans and other objects, but allows to determine which parts of the space are empty. Using this method the two categories of observations (empty and occupied) allow to employ a classification method with a sigmoid-based output. The following method [7] speeds up creation of the model by using an elegant combination of kernels and optimization methods. The speedup achieved by combination of these methods allows to recalculate the model on a frequent basis, which allows to keep the model updated despite the environment change.

Unlike the aforementioned works, which are aimed primarily at modeling the spatial distribution of objects or predicting their motion in the short term, we aim to build a model capable of long-term prediction of people and object presence. We propose a method that allows to capture the natural daily and weekly periodicities of human presence (imposed by daily habits of going to work, cleaning, etc) in different areas of space. Our approach proposes to extend continuous spatial representations by adding several dimensions representing different periodicities in time. In particular, we transform every time periodicity into two new dimensions that form a circle in its $2D$ subspace and use them to extend the representations that model $2D$ or $3D$ space. The model is then built using traditional clustering over the resulting spatio-temporal hyperspace, which efficiently represents both the structure of the space (given by environment) and time (given by the human habits). Unlike [2] that can identify and represent multiple periodicities, we decided to model only the daily one for the sake of simplicity.

II. Method Description

To extend a continuous spatial representation we have to create new time dimensions. We transform the time values t into two new dimensions

$$\begin{aligned} t_c &= \cos 2\pi\omega t, \\ t_s &= \sin 2\pi\omega t, \end{aligned} \tag{1}$$

Artificial Intelligence Center, Czech Technical University tomas.vintr@fel.cvut.cz

The work has been supported by the Czech Science Foundation project 17-27006Y and EU H2020 project 732737.

where ω corresponds to the chosen periodicity. The measured vectors $\chi = (x,y,t)$ are transformed into vectors $\mathbf{x} = (x,y,t_c,t_s)$, which generate a new $4D$ hyperspace. The set of vectors $\tau = (t_c,t_s)$ lies on a circle in its $2D$ subspace.

For the purposes of creating the $4D$ continuous model we have chosen from the list of well-known clustering methods [8] the Gustafson–Kessel Algorithm [9]. This algorithm iteratively minimizes the objective function

$$J_{GK}(\mathbf{X};\mathbf{U}_{GK};\mathbf{C}) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2}, \tag{2}$$

where $\mathbf{X} = \{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$, $\mathbf{X} \subset \mathbf{R}^{dim}$, is a set of objects, $\mathbf{C} = \{C_1,\ldots,C_c\}$ is a set of clusters, $\mathbf{U}_{GK} = (u_{ij})$ denotes a fuzzy partition matrix with conditions

$$\sum_{j=1}^{n} u_{ij} > 0, \forall i \in \{1,\ldots,c\}, \tag{3}$$

$$\sum_{i=1}^{c} u_{ij} = 1, \forall j \in \{1,\ldots,n\}, \tag{4}$$

where $\mathbf{c}_i$ is a cluster prototype of a cluster C_i, and $m > 1$ is a fuzzifier. $d_{ij} = \left\|\mathbf{x}_j - \mathbf{c}_i\right\|_M$ is the cluster-specific Mahalanobis distance, where the distance is defined for every cluster as

$$d^2(x_j,C_i) = (x_j - c_i)^T \Sigma_i^{-1}(x_j - c_i), \tag{5}$$

and Σ_i is the covariance matrix of a cluster C_i. The update equations for the covariance matrices are

$$\Sigma_i = \frac{\Sigma_i^*}{\sqrt[dim]{det(\Sigma_i^*)}}, \tag{6}$$

where

$$\Sigma_i = \frac{\sum_{j=1}^{n} u_{ij}(\mathbf{x}_j - \mathbf{c}_i)(\mathbf{x}_j - \mathbf{c}_i)^T}{\sum_{j=1}^{n} u_{ij}}. \tag{7}$$

The gained pairs C_i, Σ_i are than understood as the parameters of the set of distributions that generate the dataset.

III. Evaluation

To evaluate the presented approach, we used data about people presence collected at one of the corridors of the School of Computer Science at the University of Lincoln. Data collection was performed by a mobile robot equipped with a Velodyne 3d laser range-finder. The robot was placed at a T-shaped junction so that its laser range-finder was able to scan the three connecting corridors simultaneously. To detect and localize people in the 3d point clouds provided by the scanner, we used an efficient and reliable person detection method [10]. Since we needed to recharge the robot occasionally, we did not collect data about people presence on a 24/7 basis, but we recharged the robot batteries during nights, when the building is vacant and there are no people in the corridors. Thus, our dataset spanned from early mornings to late evening over several weekdays. Each day contains approximately 28000 entries, which correspond to hundreds of people walking or standing in the monitored corridors.

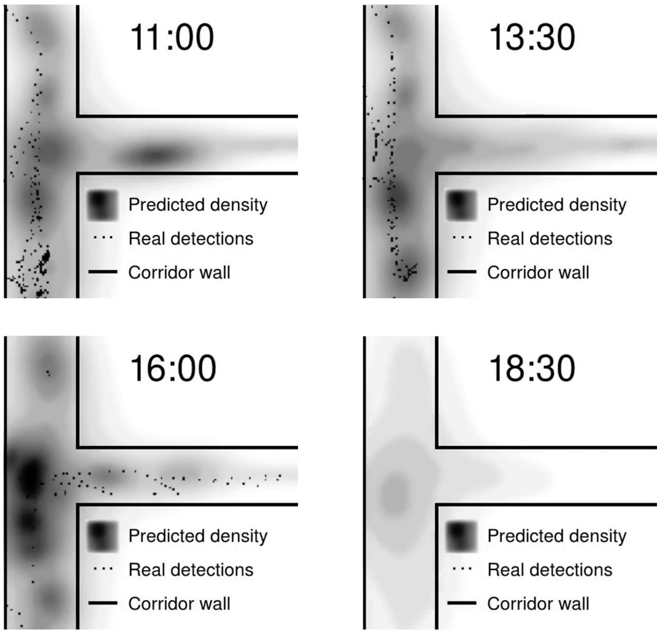

Fig. 1. Screen grab from the video available at [11].

As a natural period of time to prove this concept we have chosen one day. Parameters for clustering algorithm were chosen as: fuzzifier $m = 2$, number of clusters $k = 30$.

To evaluate the model qualitatively we have chosen visualization in the form of a video, see Figure 1. The video visualises measured positions of people, and the probabilistic distribution of their presence generated by our model. Every frame of the video projects one time frame. The time step is chosen to be 5 minutes, and the whole video comprises two days of measurements. Time values are placed in the form of text on the bottom left part of the video screen. It is obvious that the model outputs correlate with the human behavior in the selected space.

IV. Conclusion

We have presented a novel approach to model dynamic human-populated environments for robots that are supposed to navigate around and provide services to the people there. The model is based on projections of the observed time to several new dimensions derived from the natural human behavior.

We have tested this method on a real dataset from the corridors at a university collected over several days. We have shown that using such a projection of the space-time it is possible to use regular analytical methods to create a functional model.

In future work we will create a dataset consisting of a data from a longer period of time and the periods of human behavior will be automatically derived from the data using FreMEn [2]. As was shown in our previous work [12], it is not possible to expand the hyperspace indefinitely, because the chosen clustering method starts to be mathematically unstable. Therefore, it will be necessary to select only the most influential periodicities in the data.

REFERENCES

[1] T. Krajnik, J. P. Fentanes, G. Cielniak, C. Dondrup, and T. Duckett, "Spectral analysis for long-term robotic mapping," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3706–3711.

[2] T. Krajník, J. P. Fentanes, J. M. Santos, and T. Duckett, "Fremen: Frequency map enhancement for long-term mobile robot autonomy in changing environments," *IEEE Transactions on Robotics*, 2017.

[3] T. Kucner, J. Saarinen, M. Magnusson, and A. J. Lilienthal, "Conditional transition maps: Learning motion patterns in dynamic environments," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 1196–1201.

[4] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts, and A. J. Lilienthal, "Enabling flow awareness for mobile robots in partially observable environments," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1093–1100, 2017.

[5] Z. Wang, R. Ambrus, P. Jensfelt, and J. Folkesson, "Modeling motion patterns of dynamic objects by iohmm," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 1832–1838.

[6] S. T. OCallaghan and F. T. Ramos, "Gaussian process occupancy maps," *The International Journal of Robotics Research*, vol. 31, no. 1, pp. 42–62, 2012.

[7] F. Ramos and L. Ott, "Hilbert maps: scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, 2016.

[8] R. Kruse, C. Döring, and M.-J. Lesot, "Fundamentals of fuzzy clustering," *Advances in fuzzy clustering and its applications*, pp. 3–30, 2007.

[9] D. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in *Decision and Control including the 17th Symposium on Adaptive Processes, 1978 IEEE Conference on*. IEEE, 1979, pp. 761–766.

[10] Z. Yan, T. Duckett, N. Bellotto, *et al.*, "Online learning for human classification in 3d lidar-based tracking," 2017.

[11] T. Krajník, "Warped hypertime representations for long-term mobile robot autonomy," 2018. [Online]. Available: https://youtu.be/4SW4j7DDxYE

[12] T. Vintr, L. Pastorek, and H. Rezankova, "Autonomous robot navigation based on clustering across images," *Research and Education in Robotics (EUROBOT) 2011*, pp. 310–320, 2011.