

Long-Term Human Affordance Maps

R. Limosani¹, L. Yoichi Morales², J. Even², F. Ferreri², A. Watanabe², F. Cavallo¹, P. Dario¹ and N. Hagita²

Abstract—This paper presents a work on mapping the use of space by humans in long periods of time. Daily geometric maps with the same coordinate frame were generated with SLAM, and in a similar manner, daily affordance density maps (places people use) were generated with the output of a human tracker running on the robot. The contribution of the paper is two-fold: an approach to detect geometric changes to cluster them in similar geometric configurations and the building of geometric and affordance composite maps on each cluster. This approach avoids the loss of long term retrieved information. Geometric similarity was computed using a normal distance approach on the maps. The analysis was performed on data collected by a mobile robot for a period of 4 months accumulating data equivalent to 70 days. Experimental results show that the system is capable of detecting geometric changes in the environment and clustering similar geometric configurations.

I. INTRODUCTION

Robot autonomous navigation in dynamic environments has been subject of study for several years. Dynamic environments present a challenge given that robots usually rely in a static map for localization, nevertheless, there are works reported in the literature about successful robot systems in indoor environments [1], [2] and outdoor environments [3], [4]. In order to cope with the dynamics of the environment, there have been approaches to give maintenance to the maps and keep them updated (see [5], [6], [7], [8]). With updated accurate maps, robots can localize using state of the art approaches and perform autonomous navigation.

Even with the existence of accurate maps, robot social navigation in the presence of humans remains an active research topic. Human aware planners have been proposed in order to take into account the human partner [9]. For instance, an approach for human passing behaviors in office environments was presented in [10], the work in [11] described a socially inspired motion planning where the robot identified humans walking to the same goal. An approach to learn motion behaviors from humans was given in [12] and a work to generate typical activity patterns to build a coverage planner that minimizes interference probability with people was described in [13].

The work in this paper presents an autonomous robot equipped with odometry and a laser range finder that performs data collection in a daily basis. The robot navigates through an indoor office environment following a set of way-points provided in advance while avoiding moving obstacles. A geometric map using laser data and a human affordance

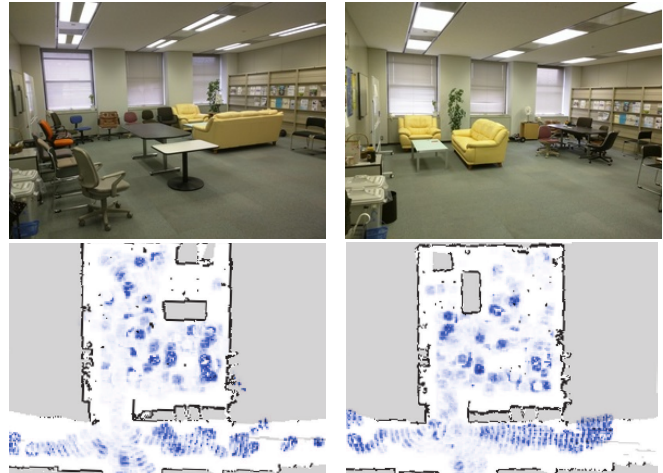


Fig. 1. Picture of an office open space in two different days. The space is used by people who in occasions change the layout. The bottom part shows their grid map representation and their correspondent affordance maps. It can be seen that place affordances change according to different geometric configurations.

map using tracked human data are built per day. In this work the concept of affordance map is used to describe how people use the environment and it is modeled as a *density map*. Using human collected data from several days an affordance density map can be built to represent the use of space in an environment.

Figure 1 shows an example of a typical office common space where people serve coffee and hold casual meetings. Figures on the top show pictures of two different days with different array configurations. Their geometric grid map representations with their respective affordance density maps are shown on the bottom. It can be seen that affordance densities are dependent on the geometry of the environment. When geometry changes, its affordance change. In case of geometric changes, the geometric map is built again. The problem in re-building the map is that accumulated affordance density information would be wrongly accumulated or totally lost .

This work proposes an approach to identify environmental geometric changes. Our approach is capable to look for similar past geometric configurations. If present geometry matches with a past stored configuration, information is retrieved and present earned geometric and affordance information is added to build composite maps of such configurations avoiding the loss of information due to re-building the map.

¹Raffaele Limosani is with the BioRobotics Institute, Scuola Superiore Sant'Anna, Italy r.limosani@sssup.it

²Luis Yoichi Morales is with the Intelligent Robotics and Communication Laboratories, Advanced Telecommunications Research Institute International, Japan yoichims@atr.jp

II. RELATED WORKS

A. Robots Operating in Dynamic Environments

There are many works about robots working in dynamic environments. Regarding detection and tracking, Anguelov et. al., in [14] treated the problem of learning models of non-stationary objects found in office-type environments. Wang et. al., in [15] proposed the detection and tracking of moving objects. They classified objects and mapped them into local and static moving object maps. The work in [16] proposed an approach to perform on-line map building for updating a previously available map where moving objects carried no information at the estimation of the posterior. Haehnel et. al., in [17] proposed an approach for map building in dynamic populated environments considering every measurement individually estimating a posterior about whether or not the data item was a dynamic object.

Finally, the method proposed in [18] focused on how to manage and model dynamic changes in the environment, rather than analyse the use of the environment by people.

B. Affordances

The concept of affordance was created and defined in the context of design and can be expressed as "An affordance is a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used" [19].

Recent works as [20] proposed the concept of affordance in a robotic navigation context as a *spatial affordance map* a model that represents human activity events as a rate function of a non-homogenous spatial Poisson process.

The aim of this work is not directly focused on the affordance map but on its correlation on the environment. Proxemics drives human interaction with surrounding objects and other people in an environment [21]. Proxemics between objects and people changes dynamically and is dependent on the context making it hard to keep track of. This means that the way of the use of an environment by people is not constant but changes as function of the current configuration of the space. As result, the use of an affordance map as a costmap for the navigation process of an autonomous robot has to be evaluated considering the geometric configuration of the environment.

The analysis of this work has been performed considering a particular case study using data acquired in a common space area of an office environment, used for rest and discussions by employees.

III. COMPOSITE MAP BUILDING

This section gives the definition and explains the process for building geometric-affordance composite maps.

The concept is shown in Figure 2. Geometric grid maps and affordance maps are built periodically. A geometric similarity check towards a data base of composite grid maps is done to determine if the configuration has already been seen.

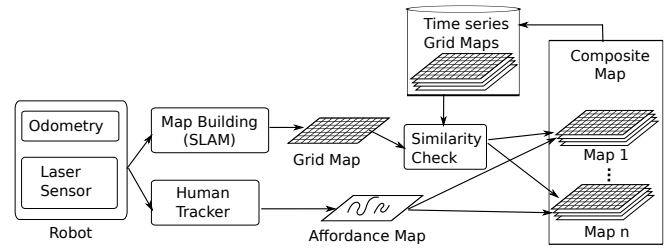


Fig. 2. Data flow diagram. Geometric and affordance maps are daily built, a symmetry check is performed to group the map in a cluster, if no symmetry is found a new cluster is created.

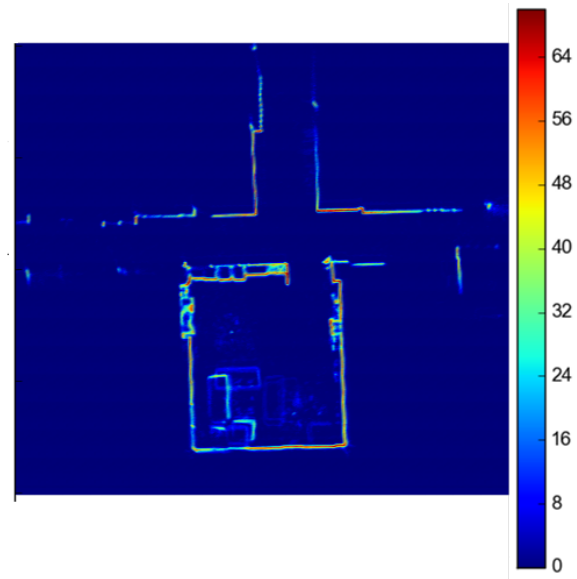


Fig. 3. Geometric Composite Map built from 70 maps taken in 4 months. It can be seen how movable objects (green to blue) have less density than static objects such as walls (red).

A. Geometric Composite Map

A composite map is a map built from multiple similar maps. Every map contributes information to form the composite map which accumulates the information taken in time. Geometric composite maps are built using multiple maps. The maps must share the same coordinate system and they have to be consistent. A geometric composite are built using the occupancy state of occupancy grid maps. Each cell in the map can have a state of free, occupied or unknown [22].

Taking several maps, a composite map can just be built by adding the number of times that a cell was occupied (or could also be represented as a probability).

An example of the composite map of 70 days of data with aligned consistent maps is shown in Figure 3. Movable objects have less density (green to blue) than static objects (red). It can be seen that static structures in the environment could be easily extracted.

B. Affordance Composite Map

Data acquired from laser sensor during the mapping activity of the robot were used to evaluate a density map representing areas of the environment most used by humans.



Fig. 4. Affordance Composite Map of 70 maps. Clearer values show higher probabilities of having humans in those locations.

Starting from an empty grid with the same dimension of the builded map, for each detected humans along the time of the process, a counter value related to each cell is updated; in particular, cells whose centroid is closer of $0.25cm$ to the tracked human was taken in account, modelling the human as a circle on the map.

Furthermore, starting from the same empty grid, a counter value related to the number of visits from the laser sensor is computed: for each ray a raycasting algorithm is used to evaluate which cells are crossed from the ray for at least $0.25cm$, which represents half of the length of the side of the cell.

Finally, the resulting value of each cell of the computed density map is the *number of humans*, meaning the number of times a human has been tracked in the cell, normalized by the *number of visits*. Since the robot is moving during the evaluation process, there is not guarantee that each area is evaluated the same number of times; normalizing the process by the number of the visits, this problem is overcome.

Cumulative Density Map: Daily density maps of consecutive days have been combined as sum of *number of humans* normalized by the sum of *number of visit*. Affordance composite map is shown in Figure 4.

Nevertheless, as introduced, the density map is strongly correlated with the current configuration of the environment. As a result, the overall addition of density maps would be a wrong description of the actual usage of the space of the environment by people.

In order to face this problem, a threshold base algorithm has been used to daily evaluate the differences between the current geometric map of the enviroment and the previous ones. Basing on this information, it could be autonomous chosen if the current affordance map has to cumulated

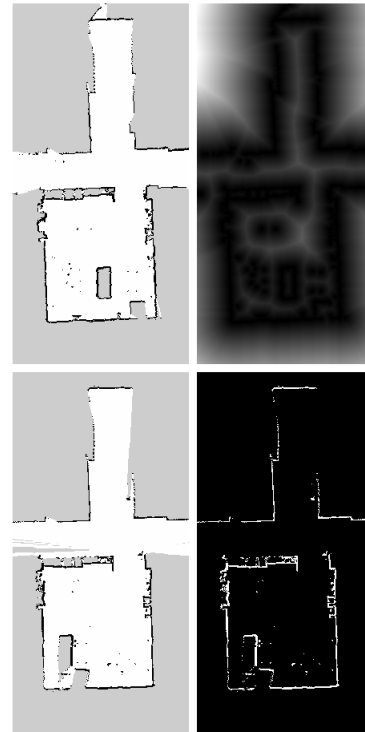


Fig. 5. On top left is the map from February 17th 2015 and on its right the corresponding distance field. On the bottom left is the map from January 13th 2015 and on its right is the result after binarization

with the old ones or has to be restarted related to the new configuration of the environment.

C. Similarity between maps

Considering two differents days, *dayX* and *dayY* with the their respective maps, *mapX* and *mapY*. The comparison method is as follows:

- compute *dmapX*, the distance field corresponding to *mapX*: for each cell determine the Euclidian distance to the nearest occupied cell. The distance is expressed in number of cells.
- binarize *mapY*, noted *bmapY*: apply a threshold on *mapY* occupancy grid to obtain 1 for the occupied cells and 0 for the rest
- multiply *dmapX* by *bmapY*:

```

1: for all i rows do
2:   for all j columns do
3:      $mapXmapY(i,j) = dmapX(i,j) * bmapY(i,j)$ 
4:   end for
5: end for

```

- the result, noted *mapXmapY*, is an occupancy grid where each cell represents how much its closest occupied cell moved from *dayX* to *dayY*

This procedure is applied on every combination of two maps. Note that both *mapXmapY* and *mapYmapX* are considered.

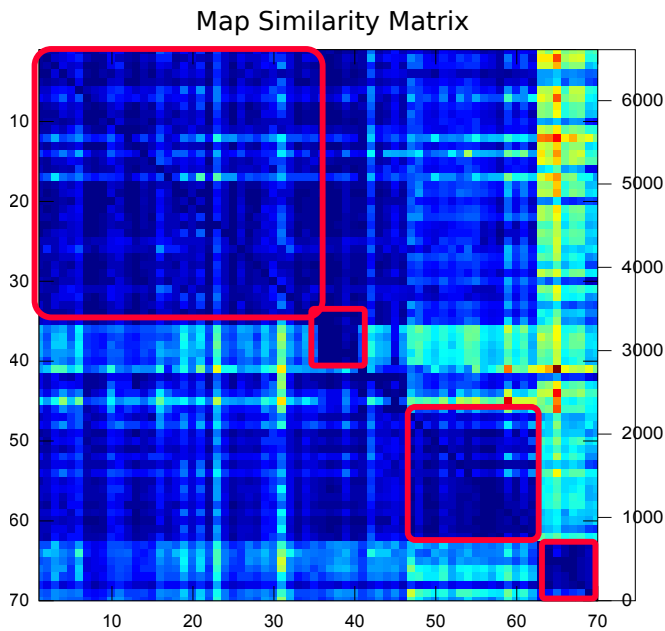


Fig. 6. Map similarity matrix.

Given a $mapXmapY$, we sum up all the distances greater than a certain threshold ct

```

1:  $cmapXmapY = 0$ 
2: for all  $i$  rows do
3:   for all  $j$  columns do
4:     if  $mapXmapY(i,j) > ct$  then
5:        $cmapXmapY += mapXmapY(i,j)$ 
6:     end if
7:   end for
8: end for

```

A value of 10 cells, (representing 0.5 m) for ct is selected as it can be considered to be a large enough variation to influence the way the space is used by people (see Figure 9). Figure 6 illustrates the results of this computation. Blue cells describe a small amount of change in the environment from one day to another, whereas red cells reflect an important difference between the two days. By utilizing a simple threshold approach on $cmapXmapY$ (1300), we cluster similar layouts of the room. Square blocks highlighted in Figure 6 represent maps similar to each other within periods of time.

D. Map Clustering

As introduced in the previous paragraph, similarity between maps was used to create clusters of maps:

- Given a list of n non-clustered maps, we compare the first one with the rest.
- To cluster similar maps we compare how close they are based on the distance between them. If the distance between maps X and Y is below a threshold (empirically set up to 1300) then they are considered similar and

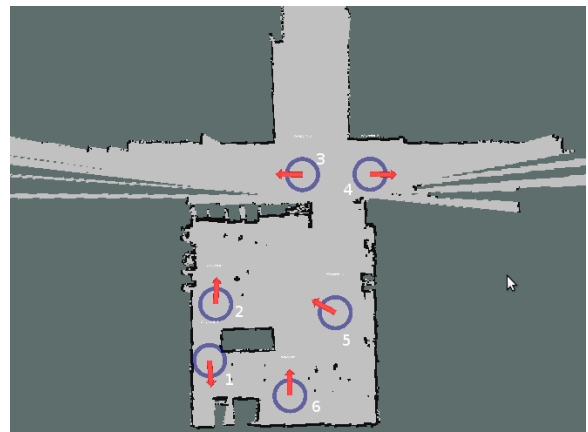


Fig. 7. Waypoints provided to the robot for logging everyday in the coffee open space.

grouped in the same cluster. If a map has no similarity match then it forms a cluster with a single map.

- The maps that are clustered together are removed from the list of unclustered maps and the process repeats until all the maps are clustered.

This method offered 70 % reliability for clustering similar layouts of the room. The ground truth was coded by a human who clustered similar layouts.

IV. EXPERIMENTAL PROCEDURE

We used a mobile platform mounted with a 2D Laser Range Finder sensor (UTM) to build maps of a coffee lounge in our office every day over a period of four months. Each map is represented by a 5 cm resolution occupancy grid generated using the slam gmapping package from ROS [23]. To ensure consistency in the coordinate frame, all maps were aligned with ICP to a reference map i.e. the first one. The initial map was generated by manually driving around the robot using a remote controller. To automatically generate the daily maps, we defined a list of fixed destinations in the environment and let the robot move autonomously (localization, path planning, and obstacle avoidance) towards each destination (see Figure 7). Laser scan data and odometry were logged during the run and the new map was created off-line afterwards. Every time the environment changed drastically e.g. furniture moved, appeared or disappeared, the robot would not localize itself properly and some of the destinations became invalid. In such cases we repeated the same procedure as for the initial map. In total, we collected 70 maps.

A. Human Tracker

This section explains how the laser scanner data is processed in order to detect and track people in the environment. The system is based on three different steps, which are implemented as three different modules:

- 1) a laser clustering module, which uses scan from lasers to detect candidates of single human legs; when two detect clusters are closest than a threshold value, the

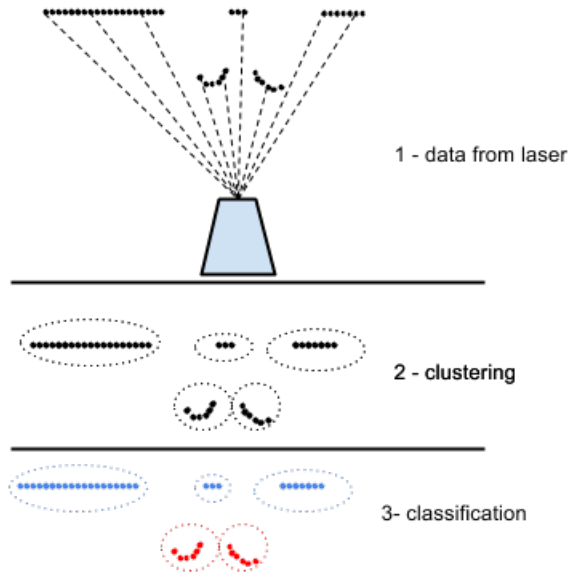


Fig. 8. Process applied to each scan taken with a laser sensor.

module publish a candidate of human as output; it is represented as a pose with covariance, the value of the pose is the midpoint between the centroids of the concerned clusters while the covariance is a fixed value based on the performances of the measurement;

- 2) an estimation module, that uses the human candidates generated from the laser clustering module and computes them over the time. Result are published at fixed rate;
- 3) a tracking module, which uses a particle filter algorithm to track humans;

1) *Data Clustering*: First we processed each scan of laser data grouping points into clusters. When two consecutive points are within 0.1 m of distance they are grouped as a cluster. A covariance matrix is built from the point distribution of each cluster and the eigenvalues of the matrix are computed to analyze the shape of the cluster. If the length of a cluster is smaller than 0.5 m and the cluster distribution does not represent a straight line (analyzing the ratio between the big and small eigenvalues) then it is determined that the cluster is a moving obstacle. This would be the case of a typical scan of the single human legs.

Starting from candidates of single human legs, a human candidate is created when two detected cluster are closer than a threshold. Human candidates are represented as pose with covariance.

2) *Gaussian Combination*: Human candidates are combined as a linear combination of gaussian functions, using the pose as the mean value of a normal distribution and the covariance matrix for the values of the variances along x and y axis. These functions are stored along the time in a dynamic vector and linearly combined using a *forgetting factor* which is a linear function of the time. Old functions,

meaning that the relative forgetting factor is under a threshold, are automatically removed. Using gaussian functions, the module is not bind to the clustering module's output, but it can use the output coming from other heterogeneous sensors and combine them.

3) *Human Detection and Tracking*: A particle filter is a common sample-based technique often used for Bayesian state estimation in robotics. In our implementation, particle filters were used to estimate four state variables (x, y, v, θ) for each entity being tracked. In the resampling step of the filter, we used the Sampling Importance Resampling method.

The key elements which define the behavior of a particle filter are the motion model used for propagating the particles and the likelihood model used for assigning weights to them.

Motion Model As has been observed in [24], the modeling of human motion presents difficulty because it is neither Brownian in nature, nor can it be modeled as a smooth linear function, since people may stop or change direction abruptly. Thus, as a compromise between the two, a Gaussian noise component is added to each particles v and θ values to capture the randomness of human motion.

Likelihood Model The likelihood value for each particles is directly estimated from the result of the Gaussian Combination module: the value is computed according to the position of the particle in the space. For each tracked entity, the sum of the weight of all the particles is computed in order to normalize the final values and keep the overall weight equal to one.

The weighted average of the particles is considered as the next updated position of a tracked human. Finally, the particles around each human are re-sampled according to their weights. A human entity is deleted from the system if the particles around the human exhibit very low confidence levels in terms of weights. The tracking system initializes a new human when the value of the resulting gaussian combination of previous module in the newest human candidates position is over an experimentally estimated threshold. New detections are simply matched against existing tracks considering a necessary minimum distance between two tracked people. Even if this procedure can create false negative estimations (e.g. if two people are very close in the map they are considered as a single tracked person), these errors would not compromise the final affordance map evaluation, since it is estimated as space occupied by people on the map.

V. RESULTS

The described process aims for describe the affordance map of environment considering the changes of configuration of the environment itself.

Analysis has been performed on 70 daily logs registered by autonomous robot used of maps maintenance of an office environment. Registration was performed at the same time of each day in order to have a common comparison between the usages of the environment in different days at the same time.

The overall affordance map resulted by the analysis is depicted in Fig. 4; as introduced, it is useless due to the

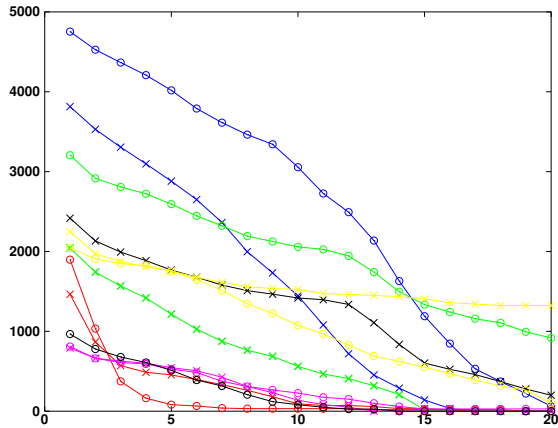


Fig. 9. $cmapXmapY$ as a function of ct for different pairs of maps. $cmapXmapY$ is drawn as circle and $cmapYmapX$ as crosses



Fig. 10. Maps of different configurations of the environment

changes of the environment during the overall building process.

Nevertheless, using the algorithm described different configurations have been recognized. Analyzing the affordance related to each configuration, it is more clear the influence of the environment on the use of the space by people; results are depicted in Fig. 11

VI. DISCUSSION

A. Limitations

The described analysis and all the processes, including tracking and mapping, have been performed using data gathered by 2D laser sensors; this limited the study, due to the presence of objects which are “invisible” for a laser

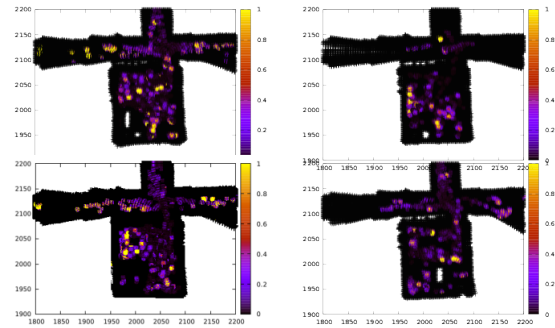


Fig. 11. Cumulative affordance maps on analyzed data, considering the current configuration of the environment.

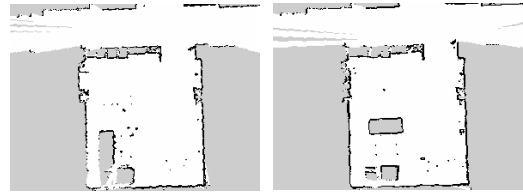


Fig. 12. Example of two layouts of the room that were wrongly clustered together.

scan, such as tables or chairs. Future works are planned in order to extend the analysis using 3D data, for example using depth cameras or 3D laser scanner, in order to improve both the tracking and the mapping processes, performing a deeper analysis on the correlation between affordance and configuration of the environment; as described in Sec. IV-A, the tracking process is not bound to a specific type of data, meaning that multiple sensors could be used at the same time.

The clustering method we used in this paper is sensitive to brief slight changes in the environment. For example, Figure 12 shows an example of two maps that were clustered together but have different layout. The problem of choosing a more strict threshold would be that clustering would become more sensitive to variation. Another limitation of the approach is that it is sensitive to inconsistencies in the exploration of the environment as the distance field of the map is function of occupied cells.

B. Usefulness of affordances maps

The information provided by the built of the affordance maps could be useful for different heterogeneous purposes, such as providing an *a priori* information for human trackers.

In the work carried out, the affordance map has been used to generate a new layer of global costmap used by the ROS *move_base* package of the ROS navigation stack [25]. As explained in [26] different layers can be merged to obtain a final costmap taking in account ad hoc information.

In particular, the affordance map created has been used as positive value on the cost, meaning that paths over area with high values of affordance (in other words, more used by people) have a greater cost rather than paths over small values of affordance. The final result aims for creating a

navigation behaviour more comfortable from the point of view of people that share the environment with the robot.

Using the same way of thinking, the affordance can be used as negative value on the cost, for tasks where robot has to interact with people, such as providing information.

The study of this work was centered in a single room, however, the complexity of modeling simple movement of furniture that change navigational affordances in time presents a complex task that requires further study. This work proposed to compute a distance index to detect changes in geometry. As future work authors plan to divide the map in regions to keep track of local geometrical changes where correlation between geometry and affordances could be computed.

VII. CONCLUSION

This paper presented a work on mapping the use of space by humans in long periods of time where daily affordance density maps were generated with the output of a human tracker running on the robot.

The contribution of the work is the proposal of an approach to cluster geometrically similar maps to build affordance composite maps on each cluster, this is, affordance maps with its geometric correspondent.

This approach avoids the loss of long term retrieved affordance information in the case of change in geometry. Geometric similarity was computed using a normal distance approach on the maps which showed to offer 70% of reliability. The analysis was performed on data collected by a mobile robot for a period of 4 months accumulating data equivalent to 70 days.

REFERENCES

- [1] W. Burgard, A. Cremers, D. Fox, D. Hanel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "The interactive museum tour-guide robot," in *Proc. of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, 1998.
- [2] W. Meeussen, E. Marder-Eppstein, K. Watts, and B. P. Gerkey, "Long term autonomy in office environments," in *ICRA 2011 Workshop on Long-term Autonomy*, IEEE, Shanghai, China: IEEE, 05/2011 2011.
- [3] Y. Morales, E. Takeuchi, A. Carballo, A. Aburadani, and T. Tsubouchi, "Autonomous robot navigation in outdoor cluttered pedestrian walkways," *Journal of Field Robotics*, vol. 26, no. 8, pp. 609–635, 2009.
- [4] R. Kümmerle, M. Ruhnke, B. Steder, C. Stachniss, and W. Burgard, "Autonomous robot navigation in highly populated pedestrian zones," *Journal of Field Robotics*, 2014.
- [5] A. Walcott-Bryant, M. Kaess, H. Johannsson, and J. Leonard, "Dynamic pose graph slam: Long-term mapping in low dynamic environments," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, Oct 2012, pp. 1871–1878.
- [6] B. Bacca, J. Salvi, and X. Cuf, "Long-term mapping and localization using feature stability histograms," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1539–1558, 2013. [Online]. Available: <http://dblp.uni-trier.de/db/journals/ras/ras61.html#BaccaSC13>
- [7] N. Carlevaris-Bianco and R. M. Eustice, "Long-term simultaneous localization and mapping with generic linear constraint node removal," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 1034–1041.
- [8] F. Pomerleau, P. Krüsi, F. Colas, P. Furgale, and R. Siegwart, "Long-term 3d map maintenance in dynamic environments," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, May 3 – June 7 2014, pp. 3712–3719. [Online]. Available: [bib/pomerleau.icra14.pdf](http://bib.pomerleau.icra14.pdf)
- [9] E. A. Sisbot, R. Alami, T. Simon, K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, and C. Nehaniv, "Navigation in the presence of humans," in *IEEE-RAS International Conference on Humanoid Robots*, 2005.
- [10] E. Pacchierotti, H. Christensen, and P. Jensfelt, "Design of an office-guide robot for social interaction studies," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, Oct 2006, pp. 4965–4970.
- [11] J. Müller, C. Stachniss, K. O. Arras, and W. Burgard, "Socially inspired planning for mobile robots in populated environments," in *International Conference on Cognitive Systems (CogSys)*, 2008, pp. 85–90.
- [12] M. Luber, L. Spinello, J. Silva, and K. Arras, "Socially-aware robot navigation: A learning approach," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, Oct 2012, pp. 902–907.
- [13] G. D. Tipaldi and K. O. Arras, "Please do not disturb! minimum interference coverage for social robots," in *IROS'11*, 2011, pp. 1968–1973.
- [14] D. Anguelov, R. Biswas, D. Koller, B. Limketkai, and S. Thrun, "Learning hierarchical object maps of non-stationary environments with mobile robots," in *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*, 2002, pp. 10–17.
- [15] C.-C. Wang and C. Thorpe, "Simultaneous localization and mapping with detection and tracking of moving objects," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Washington, DC, May 2002.
- [16] C.-C. Wang, C. Thorpe, and S. Thrun, "Online simultaneous localization and mapping with detection and tracking of moving objects: Theory and results from a ground vehicle in crowded urban areas," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Taipei, Taiwan, September 2003.
- [17] D. Hähnel, R. Triebel, W. Burgard, and S. Thrun, "Map building with mobile robots in dynamic environments," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2003.
- [18] D. Meyer-Delius, M. Beinhofer, and W. Burgard, "Occupancy grid models for robot mapping in changing environments," in *AAAI*, 2012.
- [19] D. A. Norman, *The Design of Everyday Things*, reprint paperback ed. New York: Basic Books, 2002.
- [20] G. Tipaldi and K. Arras, "I want my coffee hot! learning to find people under spatio-temporal constraints," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 1217–1222.
- [21] E. Hall, *The Hidden Dimension*, ser. A Doubleday anchor book. Anchor Books, 1990. [Online]. Available: <http://books.google.co.jp/books?id=zGYPwLj2dCoC>
- [22] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, June 1989.
- [23] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *Robotics, IEEE Transactions on*, vol. 23, no. 1, pp. 34–46, 2007.
- [24] A. Bruce and G. Gordon, "Better motion prediction for people-tracking," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2004.
- [25] E. Marder-Eppstein, E. Berger, T. Foote, B. P. Gerkey, and K. Konolige, "The office marathon: Robust navigation in an indoor office environment," in *International Conference on Robotics and Automation*, 05/2010 2010. [Online]. Available: http://www.ros.org/wiki/Papers/ICRA2010_Marder-Eppstein
- [26] D. V. Lu and W. D. Smart, "Towards More Efficient Navigation for Robots and Humans," ser. IROS, 2013.