

Affordance-Map : A Map for Context-Aware Path Planning

Lasitha Piyathilaka and Sarath Kodagoda

Centre for Autonomous Systems (CAS), University of Technology, Sydney

Jayaweera.M.Piyathilaka@student.uts.edu.au

Abstract

‘Context-awareness’ could be one of the most desired fundamental abilities that a robot should have when sharing a workspace with humans co-workers. Arguably, a robot with appropriate context-awareness could lead to a better human robot interaction. In this paper, we address the problem of combining context-awareness with robotic path planning. Our approach is based on affordance-map, which involves mapping latent human actions in a given environment by looking at geometric features of the environment. This enables us to learn human context in an given environment without observing real human behaviours which themselves are a non-trivial task to detect. Once learned, affordance-map allows us to assign an affordance cost value for each grid location of the map. These cost maps are later used to develop a context-aware global path planning strategy by using the well known A* algorithm. The proposed method was tested in a real office environment and proved our algorithm is capable of moving a robot in a path that minimises the distractions to human co-workers.

1 Introduction

Thanks to recent advances in robotic technologies, robots are slowly approaching daily lives of humans. As such, the ability of a robot to perceive and understand its operating environment is a key requirement in many of the scenarios. Often in robotics, context awareness refers to the ability of robots to sense and react according to the environment. Particularly, understanding human context is paramount for a robot that works alongside humans. Therefore, human context often involves recognizing probable human activities in a given environment. However, even recognizing a simple human activity is considered to be non-trivial due to large intra-activity

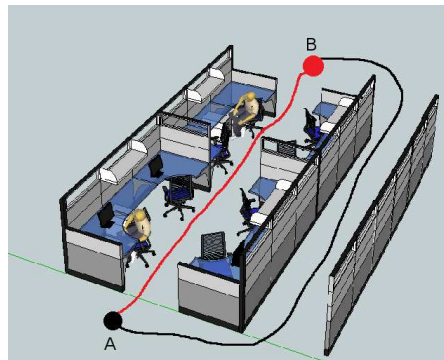


Figure 1: A robot with proper context-awareness will select a path along corridors rather than the shortest path

variations associated with human activities [Piyathilaka and Kodagoda, 2013b].

In order to react according to context of the environment, a mobile robot is often required to alter its path more often. Therefore, context-aware path planning is a fundamental ability that every robot should acquire in order to navigate in a human work space. However, most traditional path planning algorithms are based on two major factors, obstacle avoidance and minimizing cost associated with robot’s movements. In other words, these robots try to move from start position to the goal location on the shortest possible path while avoiding obstacles on its way. Although this path planning strategy is more suitable for a industrial setting, such an approach is often perceived as inappropriate for a robot that works alongside with humans, as human factors are not considered in path planning.

The need for context-aware path planning can be explained by the scenario shown in Fig. 1. In this example, a robot is required to move from point A to B. A robot that uses a traditional path planning strategy would select the shortest path between A and B, which is going along the human work space. Therefore in this case, a robot that navigates on the shortest path may

distract human workers. However, a robot with context-aware path planning capability would choose an alternative path along the corridor minimizing obstruction to human activities.

In recent years, a few researchers have considered associating human context with robotic path planning [Sehestedt *et al.*, 2010; Bennewitz *et al.*, 2005]. These approaches try to learn human context through observations. Therefore, first the robot is required to observe and track people for a considerable amount of time before using that knowledge for path planning. On the other hand, the robot could learn the human context, only if humans were observed in the environment.

Our approach for context-aware path planning is based on the concept of affordance-map. Affordance map allows us to learn human context in a given environment without seeing any humans. Affordance is defined as ‘action possibilities that are readily perceivable by an actor in a given environment’ [Greeno, 1994; Grabner *et al.*, 2011; Piyathilaka and Kodagoda, 2014]. The rationale here is that, it is possible to learn basic human action possibilities by only looking at geometric features of the environment as humans arrange objects to support their activities. For example, chairs and sofas support the activity ‘sitting’ and they are physically designed to support that affordance. Therefore, the action possibility ‘sitting’ could be learnt by simply detecting sittable locations in the environment. Although these affordances could be learnt by detecting human activities in a given environment, in our proposed method we used virtual skeleton models to learn possible affordances. This allowed us to learn affordances even in environments where human were not observed.

Following section briefly explains the major steps involved in our approach to build the affordance map. First the 3D map of the environment is obtained and converted to a 3D distance map. Then at the learning phase, virtual human models are placed on 3D CAD models of furniture downloaded from a public dataset to learn parameters of the affordance model. Later, at testing phase, virtual human skeleton models are placed across the 3D environment to search for locations that support the given affordance. Finally, affordance-map is created by calculating affordance value for each grid cell that represents how likely that location support the given affordance. In addition, each grid cell is consisted of orientation information of the human skeleton model. This information embedded in the affordance map is then used to assign context cost value for each grid cell. Finally, experiments were carried out in a real office environment to explore the applicability of affordance-map for context-aware path planning.

2 Related Works

There are many publications that explored various robotic path planning strategies. However, only a few researchers have previously considered global path planning with appropriate context awareness.

In [Sehestedt *et al.*, 2010] researchers developed social aware path planning strategy using a robot that learns human motion patterns based on sampled Hidden Markov Models. Then they utilized these models for path planning based on Probabilistic Roadmap. However their approach need to continuously observe real humans in the environment before using it for social aware path planning. Therefore, such a robot would require a considerable amount of time to learn human context in a new environment. On the other hand, human motion patterns could be affected by the presence of the robot itself. However our approach for context-aware path planning does not require to observe real humans. Instead it learns human behaviours by looking at geometric features of the environment.

In an attempt to build context-aware path planning strategy, some other researchers have used an off-line learning procedures to obtain motion models of individual people in an office using HMM [Bennewitz *et al.*, 2003]. Once a person is identified, a model is used to predict that persons future position in order to perform the ‘give way’ action . Thus, the model is used to improve local replanning and does not influence global planning. However such a behaviour could lead to more complex behaviours and robot might interfere with some one else when deployed in crowded environments.

In [Luber *et al.*, 2012] researchers tried to build socially aware navigation strategy using objective criteria such as travel time or path length as well as subjective criteria such as social comfort. As opposed to model-based approaches they posed the problem as an unsupervised learning problem. They learnt a set of dynamic motion models by observing relative motion behaviour of humans found in publicly available surveillance data sets. Again this strategy requires the human to be present and detected in the environment. Further, they have only focused on human like collision avoidance rather than global path planning.

3 Affordance Detection

In order to build a social cost map, possible affordances in the given environment need to be detected and mapped. Affordance-map is consisted of likelihood value assigned to each grid cell. This is achieved by modelling the 3D environment, virtual humans and their interactions which can be illustrated by the example shown in the Fig.3. It can be seen from the Fig. 3 ‘proper sitting’ of the virtual humans model can be described by placing

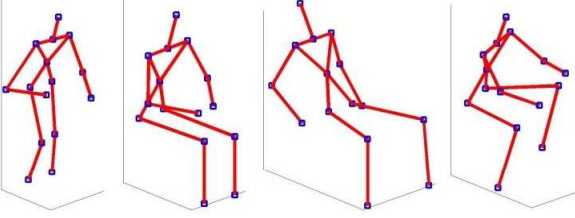


Figure 2: Virtual human models extracted from human activity detection data set

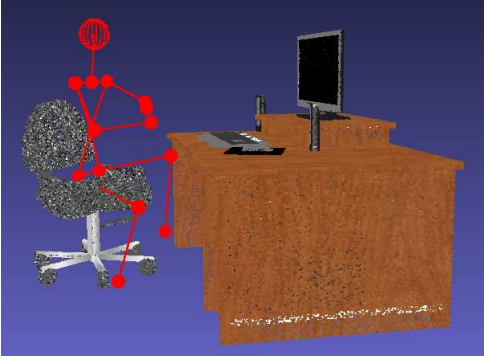


Figure 3: Virtual human models sitting near a computer

the spine supporting the vertical part of the chair while thighs supporting the horizontal part of the chair. Also human model should not go through any 3D point of the chair. Therefore, if each point of the chair is converted in to a 3D distance field then each point of the virtual human model should lie with a specific distance value. These distance features together with occupancy values are used to learn and detect affordances as explained in following sections.

3.1 Environment Model

The environment model is consisted of 3D Distance Transform Map $DT(\mathbf{x})$ and 3D Occupancy Map $OC(\mathbf{x})$, where \mathbf{x} is any 3D position of the environment. The 3D Distance Transform (DT) is a shape representation that indicates the minimum distance from a point in the environment to the closet occupied voxel. In our approach, we calculated 3D Distance Transform by using the occupied voxels of 3D point clouds, OC . The distance transform map $DT(\mathbf{x})$ of the occupancy grid map OC can be generated using an unsigned distance function (1), that represents Euclidean distance from each location \mathbf{x} of the environment to the nearest occupied voxel in $OC(\mathbf{x})$.

$$DT(\mathbf{x}) = \min_{O_j \in OC} |O_j - \mathbf{x}| \quad (1)$$

3.2 Virtual Human Skeleton Model

Instead of observing real humans in the environment, the proposed algorithm uses virtual humans to model interaction between the environment and the human. Although many human poses could be observed in a given environment, a very few of them actually influence the robot’s context-aware navigation strategy. For example, most frequently observed human pose in an office environment is sitting beside an office desk and working with a computer. Therefore, a service robot operating in an office environment should minimize the distractions caused to those people working beside office tables. In this scenario, human pose model can be selected from the activity ‘working with the computer’ to generate the social cost map. In our experiments, human skeleton models were extracted from a human activity detection dataset [Sung *et al.*, 2012] and used them to build the affordance-map. Fig. 2 shows the frequently observed human pose models in the dataset. Each human model is a skeleton with 15 joint body positions in 3D. Given these 3D points of the human skeleton H_l , we can move it across the given environment using the rigid body transformations of translation and rotation. Then we can effectively map each human skeleton model to the coordinate system of the environment using (2), where $X_k = (x_k, y_k, z_k, \theta_k)$ is the position and orientation of the skeleton’s torso in the world coordinate system and $R_z(\theta_k)$ is the rotational matrix about z axis. It is to be noted that only rotation about z axis is considered here.

$$H_w(X_k) = [x_k, y_k, z_k]^T + R_z(\theta_k).H_l \quad (2)$$

3.3 Human Environment Relationship

Once the models for environment and the human have been built, the next step is to model the relationship between them. This is achieved through two geometric features, namely distance features and collision features. Selection of these features are motivated by two facts. First is the proximity of objects for effective interactions and the second is to prevent collisions or intersects with occupied voxels of the environment.

Distance features are obtained by moving the human model across the voxels in the environment and calculating distance measure for each and every skeleton points of the human model. Once the environment is modelled by (1), we can effectively calculate distance features of a human skeleton with location and orientation $X_k = (x_k, y_k, z_k, \theta_k)$ by (3), where n is the number of 3D points in the skeleton.

$$[d_1, d_2, \dots, d_n] = DT(H_w(X_k)) \quad (3)$$

In the same way, we can check for any collisions of a skeleton at location and orientation, X_k by (4). In case of a collision c_i becomes 1 and 0 otherwise.

$$[c_1, c_2, \dots, c_n] = OC(H_w(X_k)) \quad (4)$$

There after, the collision check at X_k in the map can be converted into a probability value using (5).

$$P(C|X_k) = 1 - \frac{\sum_{i=1}^n c_i}{n} \quad (5)$$

Then the affordance map for ‘workable area’ in the given environment map can be calculated by using the distance features, d_i and collision features. Each cell in the affordance map represents the likelihood of that place being a ‘workable area’. Finally, given a virtual skeleton with location and orientation X_k , the underlying Likelihood, $P(A_k|X_k, \lambda)$ of that location being a ‘workable area’ can be calculated by (6).

$$P(A_k|X_k, \lambda) = P(C|X_k) \cdot \prod_{i=1}^n \frac{1}{\sigma_i \sqrt{2\pi}} e^{-(d_i - \mu_i)^2 / 2\sigma_i^2} \quad (6)$$

The Parameters of the above likelihood function $\lambda = \{\mu_i, \sigma_i\}$, for each distance measure can be estimated via training as explained below.

First human pose models are selected according to the environment that new social cost map need to be built. Then 3D models of furniture that support each virtual pose model are selected from synthetic datasets of ‘Google 3D Ware House’. For example, to learn ‘sit-able’ affordance, ‘sitting’ human skeleton models can be used from the human activity detection dataset [Sung *et al.*, 2012] and furniture like sofas and chairs can be used to train parameters of (6). Finally, distance features are recorded by manually placing human skeleton on the selected furniture, and estimating $[\mu_i, \sigma_i]$ using training samples.

4 Affordance-Map for Context-Aware Path Planning

This section explains how affordance map can be used for context-aware path planning in an office environment. For this research, context-aware path planning is defined as the path that minimizes the interferences caused to people working in an office cell. The Context-aware path planning strategy involves avoiding work spaces of humans and planning paths through the humans work spaces if and only if the human presence is not detected. The robot can use affordance map for global path planning as it provides location information of human work spaces. For this research we define this path as ‘Global Path with Minimal Distractions’, (GPMD). However, if this path is too long then the robot needs seek alternative path that goes through the nearest unoccupied workspace.

4.1 A* for Path Planning

The A* algorithm has been in use for robotic path planning for many years [Sehestedt *et al.*, 2010; Dechter and Pearl, 1985]. It utilizes best-first search strategy to search the least cost path that lies in the D-dimensional collision free space of the robot.

The algorithm searches for a path using a priority queue, where each node x in the path is sorted according to the cost function $f(x)$. Therefore highest priority is given to the nodes with least cost. Generally $f(x)$ is the sum of two main components, $g(x)$ which is the shortest collision free path from start to goal by Euclidean distance and $h(x)$, which is used as a heuristic estimate of the length of the path.

In order to do minimum distraction path planning we can simply incorporate a social cost $g_S(x)$, to the total cost function such that A* algorithm will seek a collision free path while minimizing interference to humans. Depending on the affordance model selected social cost $g_S(x)$ for each grid cell can be assigned according to the affordance likelihood given by (6). For example, for an office environment we can select ‘sitting and working’ as the affordance model and assign high social cost values to the areas where affordance likelihood is high. In such cases social cost is given by (7).

$$g_S(x) \propto P(A_k|X_k, \lambda) \quad (7)$$

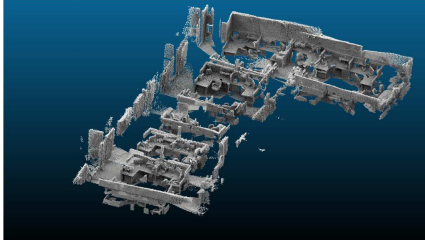
Consequently, $f(x)$ in the standard algorithm can be replaced by the cost function $F(x)$. In some scenarios it is appropriate to have additional weight term w for social cost function $g_S(x)$ in (7) to enable the robot to choose between shortest path and alternative minimal distraction path. This can be incorporated as,

$$F(x) = g(x) + w * g_S(x) + h(x) \quad (8)$$

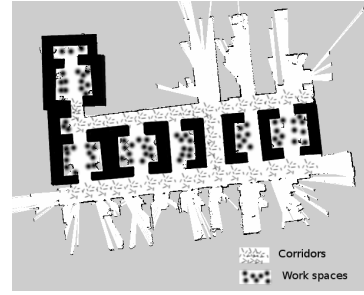
If w is set to 1, the robot prefers shortest path whilst any number between 0 and 1 denotes the combination of shortest path and a context-aware path.

4.2 Detecting Human Presence

Detecting human presence is vital for a robot that does context-aware path planning because the human context defines the role of the robot in most scenarios. Although a number of human detecting algorithms are available in literature, detecting humans while the robot is in move is considered to be non-trivial [Sehestedt *et al.*, 2010]. One reason for this challenge is the absence of a priori knowledge about the presence of human poses and their best viewing directions. On the other hand, how to react to the human presence depends on the context of the location. For instance, if a human is detected in a corridor, the robots can simply treat him/her as an obstacle



(a) 3D Point Cloud Map



(b) 2D Laser Map

Figure 4: Maps of the office environment

and plan to avoid the obstacle (could be a dynamic obstacle). Whereas, if the human presence is detected in an office space, then the robot needs to plan a path to minimize interference to humans.

However, these problems can be effectively addressed by using the information embedded in the affordance-map. Firstly, it contains information about human activities like ‘sitting and working’. Secondly, the affordance-map contains information about the locations and poses of possible human workers. This can be used for active human search, which will eventually lead to a higher human detection rate.

4.3 Context-aware Path Planning Algorithm

Algorithm 1 summarises the major steps involved with context aware path planning. First A* algorithm is used with affordance map to calculate shortest path, PS and GPMD path PA as explained in the Section 4.2. Path threshold T_p defines when to use GPMD. The path threshold can be set according to the urgency of the task that is being carried out by the robot using a high level planner. If the difference of travel distance between these two alternative paths is small, then the robot chooses the GPMD path which is reasonable. If the difference of travel distance is large, the robot starts to move along the shortest path while looking for humans. If a human is detected, then the robot alters its path by recalculating a new path to minimize the interference caused to the human.

Algorithm 1: Context-Aware Path Planning

Data: Affordance map
Result: Context-Aware Path

- 1 $PS \leftarrow$ Calculate the shortest Path;
- 2 $PA \leftarrow$ Calculate the Global Path with Minimal Distractions (GPMD) using Affordance-map;
- 3 $T_p \leftarrow$ Path Threshold;
- 4 **if** $(len(PA) - len(PS))/len(PS) \leq T_p$ **then**
- 5 | Follow PA
- 6 **else**
- 7 | **while** *Detect people OR Goal do*
- 8 | | Follow PS
- 9 | **end**
- 10 | Assign Affordance cost values to all locations within the sensor range and update the cost map;
- 11 | $PS_{new} \leftarrow$ Calculate the new shortest path ;
- 12 | $PS \leftarrow PS_{new}$;
- 13 | **GoTo** Line 7;
- 14 **end**

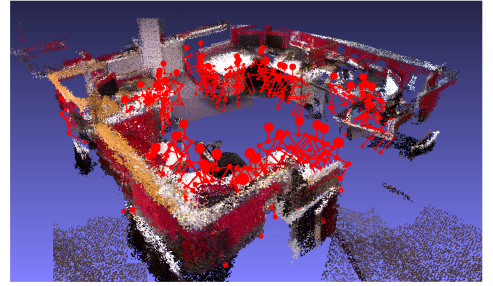
5 Experiments

In order to experimentally verify the effectiveness of the proposed algorithm, a series of path planning experiments were conducted in an office environment. First, 3D point cloud map of the environment was built using a state of the art RGBD SLAM (Simultaneous Localization and Mapping) algorithm [Dryanovski *et al.*, 2013] and a depth camera mounted on a mobile robot. Fig. 4a shows the 3D point cloud and Fig. 4b shows 2D Laser map of the environment with corridors and working benches.

Human pose model extracted from the activity ‘sitting and working with a computer’ is selected to detect affordances. This is justified by the fact that ‘sitting and working with a computer’ is the most frequently observed activity in the given environment. First, the map is voxelized into 10cm x 10cm x 10cm grids and distance fields and occupancy map are built. Then for

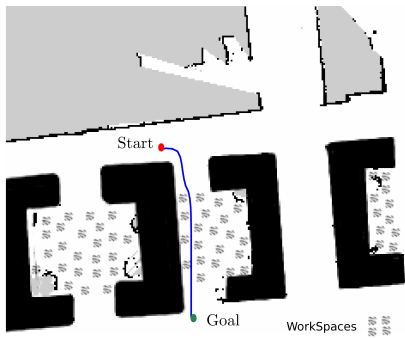


(a) 2D Affordance Likelihood Map

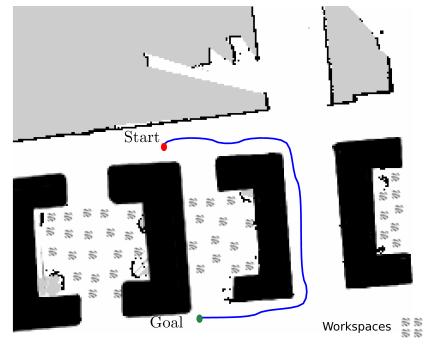


(b) 3D Skeleton Map

Figure 5: Learned Affordance-Map for the office environment

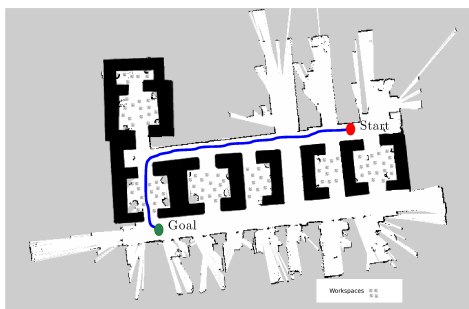


(a) Shortest path with A*

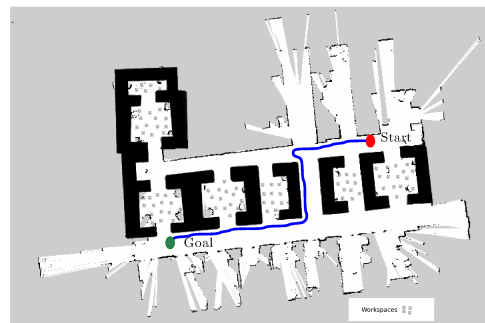


(b) 'Global Path with Minimal Distractions', GPMD

Figure 6: Experiment 1. Context-aware path planner selects a longer path to minimize the distractions to humans work spaces.

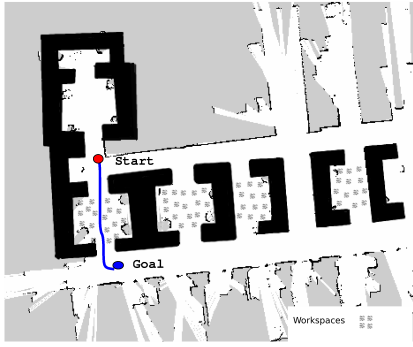


(a) Shortest path with A*

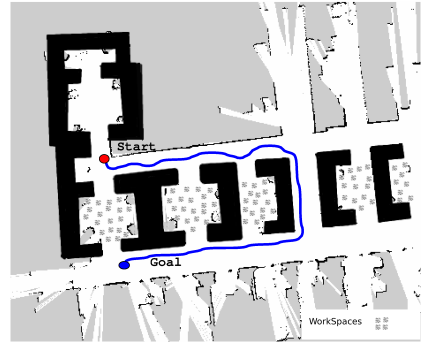


(b) 'Global Path with Minimal Distractions' (GPMD)

Figure 7: Experiment 2. Affordance-map based 'Global Path with Minimal Distractions' (GPMD) selects the only path that always lies along corridors.



(a) Shortest Path with A*



(b) Global Path with Minimal Distractions' (GPMD)

Figure 8: Experiment 3. Minimum distraction path along corridors is too long and shortest path goes through human work space.

each grid location $P(A_k|X_k, \lambda)$ is densely calculated by moving the virtual human model across the 3D map. The 2D and 3D representations of the affordance map is shown in Fig. 5a and Fig. 5b. High 'workable' affordances are recorded near work benches as can be seen from the Fig. 5a. A 3D representation of the affordance map with skeletons is shown in Fig. 5b. It is clear from these results that the proposed algorithm is sufficiently capable of learning tested affordance in the given environment. More importantly, the learned affordances are more or less realistic to real human behaviours as can be seen from the 3D affordance map in Fig. 5b.

The affordance map is then used for context-aware path planning. As the robot is required to avoid 'working spaces' of humans in the office environment, we used A* algorithm with the cost function given by (8). The values for the affordance cost, $g_S(x)$ are assigned in the range of 0-10 with proportional to the affordance likelihood 'sitting and working' for each grid cell. The value of w in (8) is set to '1' if minimum distraction path is required and set to '0' if just the shortest path is required.

In experiment 1, robot is placed in the corridor near a work station and requested to plan a path to a goal location in the other side of the work station as shown in Fig. 6. As depicted in Fig 6a, A* planner has calculated the shortest path through a human workspace whereas the GPMD based planner has chosen a path that avoids workstations as shown in Fig. 6b. In this case the difference of travel distances between these two paths is less than the chosen Path Threshold T_p in algorithm 1. Therefore, the robot selects affordance-map based path as the context-aware path which eventually causes lesser distractions to humans.

In experiment 2, robot is placed in the right corner of the map and asked to plan a path to a goal location in the opposite corner of the map as shown in Fig. 7.

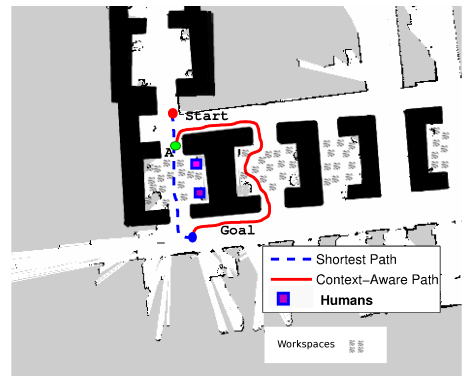


Figure 9: Experiment 3. Minimum distraction path along corridors is too long. Hence robot tries to find a path through an empty working space.

There is only one possible path along corridors and all other paths contain atleast one working space. According to the results, it can be seen that the shortest path calculated by the A* planner goes through a human work space as shown in Fig. 7a. In contrast affordance-map based planner has selected the only path that always lies in corridors. Since both of these paths have very similar travel distances the robot chooses the path based on affordance-map which causes minimal distractions to humans.

In experiment 3, the robot is placed at far left corner of the map near a workstation and requested to plan a path to a location in the other end of the workstation as shown in Fig. 8. The shortest path is shown in Fig. 8a and minimum distraction path based on affordance-map is shown in Fig. 8b. Note that there is a considerable difference in travel distances between the shortest path

and the minimum distraction path that lies along corridors. Therefore, according to the algorithm 1, robot starts to move along the shortest path while looking for human presence as shown in Fig. 9. To make the human detection more efficient, the robot only looks at locations where ‘sitting and working’ affordance likelihood is high. This leads to a huge computational savings as it does not look for humans in each and every location. At location ‘A’ of the shortest path the robot detects its first human. Then robot assigns affordance cost values to the locations which are within sensory range of the human detector. With this new cost map the robot plans a new path and follows it until it reaches its goal location as shown in Fig. 9. Although this new path is going through a human work space robot doesn’t alter its path as humans are not detected in this region.

Finally, these experiments showed that context-aware path planning can be done in an office environment efficiently using the proposed affordance-map. This map provides many additional information which a traditional grid based map fails to provide for a successful context-aware path.

6 Conclusions

In this paper, we introduced a context-aware path planning method for robots, centered around human context in an indoor environment. We also showed how a dense 3D point cloud can be converted into a more informative semantic map called ‘affordance-map’ which consists of virtual human models. This affordance-map is then used for path planning in an office environment. The experiments showed the proposed context-aware path planner is capable of avoiding human work spaces and hence contributing to less distractions to humans.

Our future work involves, using social path planning strategy for non invasive human activity detection [Piyathilaka and Kodagoda, 2013b] [Piyathilaka and Kodagoda, 2013a].

References

[Bennewitz *et al.*, 2003] Maren Bennewitz, Wolfram Burgard, and Sebastian Thrun. Adapting navigation strategies using motions patterns of people. In *Robotics and Automation, 2003. Proceedings. ICRA’03. IEEE International Conference on*, volume 2, pages 2000–2005. IEEE, 2003.

[Bennewitz *et al.*, 2005] Maren Bennewitz, Wolfram Burgard, Grzegorz Cielniak, and Sebastian Thrun. Learning motion patterns of people for compliant robot motion. *The International Journal of Robotics Research*, 24(1):31–48, 2005.

[Dechter and Pearl, 1985] Rina Dechter and Judea Pearl. Generalized best-first search strategies and the optimality of a. *J. ACM*, 32(3):505–536, July 1985.

[Dryanovski *et al.*, 2013] Ivan Dryanovski, Roberto G Valenti, and Jizhong Xiao. Fast visual odometry and mapping from rgb-d data. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 2305–2310. IEEE, 2013.

[Grabner *et al.*, 2011] Helmut Grabner, Juergen Gall, and Luc Van Gool. What makes a chair a chair? In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1529–1536. IEEE, 2011.

[Greeno, 1994] James G. Greeno. Gibson’s affordances. 1994.

[Jiang *et al.*, 2013] Yun Jiang, Hema Koppula, and Ashutosh Saxena. Hallucinated humans as the hidden context for labeling 3d scenes. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2993–3000. IEEE, 2013.

[Luber *et al.*, 2012] Matthias Luber, Luciano Spinello, Jens Silva, and Kai Oliver Arras. Socially-aware robot navigation: A learning approach. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 902–907. IEEE, 2012.

[Piyathilaka and Kodagoda, 2013a] Lasitha Piyathilaka and Sarath Kodagoda. Gaussian mixture based hmm for human daily activity recognition using 3d skeleton features. In *Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on*, pages 567–572. IEEE, 2013.

[Piyathilaka and Kodagoda, 2013b] Lasitha Piyathilaka and Sarath Kodagoda. Human activity recognition for domestic robots. In *Field and Service Robotics Conference*, pages 567–572. Springer, 2013.

[Piyathilaka and Kodagoda, 2014] Lasitha Piyathilaka and Sarath Kodagoda. Active visual object search using affordance-map in real world : A human-centric approach. In *Control, Automation, Robotics and Vision, ICARCV, The 13th International Conference on*. IEEE, 2014.

[Sehstedt *et al.*, 2010] Stephan Sehstedt, Sarath Kodagoda, and Gamini Dissanayake. Robot path planning in a social context. In *Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on*, pages 206–211. IEEE, 2010.

[Sung *et al.*, 2012] Jaeyong Sung, Colin Ponce, Bart Selman, and Ashutosh Saxena. Unstructured human activity detection from rgb-d images. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 842–849. IEEE, 2012.