

# Human Aware Robot Navigation in Semantically Annotated Domestic Environments

Ioannis Kostavelis<sup>(✉)</sup>, Dimitrios Giakoumis, Sotiris Malassiotis,  
and Dimitrios Tzovaras

Centre for Research and Technology Hellas, Information Technologies Institute,  
6th Km Charilaou-Thermi Road, 57001 Thermi-Thessaloniki, Greece  
{gkostave,dgiakoum,malasiot,tzovaras}@iti.gr

**Abstract.** In the near future, the seamless human robot cohabitation can be achieved as long as the robots to be released in the market attain socially acceptable behavior. Therefore, robots need to learn and react appropriately, should they be able to share the same space with people and to adapt their operation to human's activity. The goal of this work is to introduce a human aware global path planning solution for robot navigation that considers the humans presence in a domestic environment. Towards this direction, hierarchical semantic maps are built upon metric maps where the human presence is modelled using frequently visited standing positions considering also the proxemics theory. During the human's perambulation within the domestic environment the most probable humans pathways are calculated and modeled with sequential, yet descending Gaussian kernel's. This way, the robot reacts with safety when operating in a domestic environment taking into consideration the human presence and the physical obstacles. The method has been evaluated on a simulated environment, yet on realistic acquired data modeling a real house space and exhibited remarkable performance.

**Keywords:** Human robot cohabitation · Safe navigation · Semantic mapping · Metric mapping · Path planning

## 1 Introduction

The seamless integration of robots in human inhabited environments demands the formation of strategies that allow them to navigate in a secure, appropriate and common manner among people. This provokes new research challenges, where the concept of safe navigation attains wider dimensions and aims to facilitate human-robot cohabitation beyond the established safety strategies [1]. In mobile robotics, mapping allows robots to construct a meaningful description of their surrounding that endows them with the capacity to accomplish high-level objectives. The emerging topic of human aware navigation deals with robots' operating in complex environments, while considering the convenience of the

humans being present. Therefore, the administration of space about humans and their reaction when interacting with each other need to be comprehended and modeled.

Environment modelling in terms of metric mapping comprises the cornerstone for the majority of the robotic applications. Specifically, in order for a robot to be able to navigate efficiently, a consistent metric map has to be built. In the last decades a lot of research has been conducted in the area of mobile robot navigation and mapping [2], yielding remarkable performance. With the purpose of accurately localizing themselves [3,4], mobile robots construct a consistent representation of the spatial layout of their working environment. The representative works described in [5,6] and [7] prove the necessity for an accurate representation of the robot's surroundings as well as the development of efficient mapping methods. More precisely, simultaneously localization and mapping (SLAM) provides solution to the problem, according to which a mobile robot placed at any unknown location in an unexplored area incrementally builds a consistent map of the environment while simultaneously determines its location within this map. Seeking to engineer an efficient solution to this problem, several successful research attempts have been carried out, an analytical summary of which is presented in a two-part review paper [8,9].

Albeit the advances in metric mapping, in order a robot to apprehend the environment in the way a human does, the formation of maps augmented by semantic attributes involving human concepts, such as types of rooms, objects and their spatial arrangement, is considered a compulsory attribute for the future robots. Semantic mapping [10] is a qualitative description of the robot's surroundings, aiming to augment the navigation capabilities and the task-planning, as well as to bridge the gap in human-robot interaction (HRI). A representative work where the semantic mapping is addressed with emphasis on HRI by using natural language is the one described in [11] and enables in the most direct way robots to socialize with humans. Moreover, it is an evidence that contemporary robots use to navigate in their environments by computing their pose within metric maps and, therefore, the vast amount of the semantic mapping methods reported in the literature use these metric maps to add semantic information on top of it [11,12].

The next step, after the construction of a map commonly apprehended by humans and robots, is the adoption of a human-aware behavior, while the robot performs navigation activities within the explored environment. However, during the robot's perambulation it is essential to consider navigation strategies that facilitate the convenience of the individuals being present and increase safety. One the one hand, safety can be achieved up to one extend using the on-board robot sensors for typical obstacle avoidance. One the other hand, the answer to the human's convenience during the robot's operation in a domestic environment stems from the social sciences, where the anthropologist Edward T. Hall [13] designated the Proxemics Theory. In accordance with this theory, the human comfort levels are influenced from their distance from other persons and, consequently, there are zones, i.e. proxemics zones, that determine the intimacy

equilibrium model among human-human interactions. Robotics science adopted the Proxemics theory to model the human-robot interactions according to which the robots should be capable of perceiving proxemics and adopt their behavior accordingly [14]. Specifically, the work in [15] presents a navigation strategy in populated environments based on the prediction of people’s movement and the level of discomfort as imposed by the proxemics theory. In a more contemporary solution, the work in [16] represent the social zones in terms of isocontours of an implicit function capable of describing complex social interaction. Such zones are shaped through non-linear probability functions which derive as solutions to a learning problem in the kernel space.

## 2 Proposed Methodology

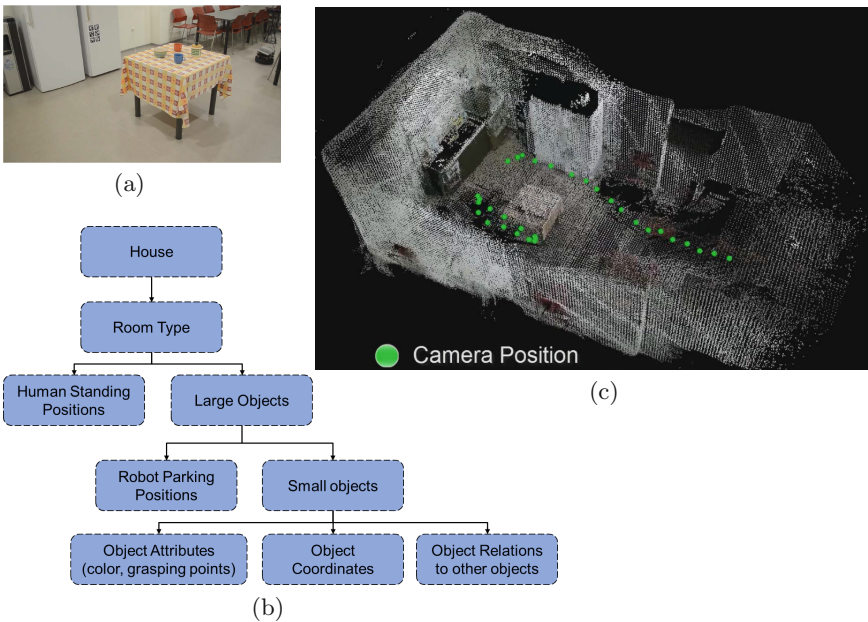
### 2.1 Outline

In this work, a dense 3D metric map is firstly constructed processing the input images acquired from a RGBD sensor mounted on a mobile robot. On the top of the metric map a hierarchical semantic knowledge base is formed encoding all the high level information that describe the domestic environment in a human compatible model. In the knowledge database, information about the location of the large objects, the supporting surfaces and the human frequently visited standing positions are stored. The direction of human motion is constantly calculated among consecutive frames and the distance from the human and the standing positions is computed. Then, by minimizing those two parameters, the most probable path to be followed given his/her current position and the defined standing positions is calculated using a Dijkstra algorithm. In course of the calculated human path, Gaussian kernels of descending amplitude are defined, the parameters of which follow the proxemics rules. During the robot global planning, all the static obstacles of the metric map are top down projected and together with the sequences of the Gaussian kernels are used to form the map on which the D\* Lite algorithm is executed to find an optimal path for the robot.

### 2.2 Metric and Hierarchical Semantic Mapping

The method we describe here uses a metric map in order to perform robot navigation activities. The metric mapping solution adopted, is the one presented in [17], yet enhanced in terms of memory and speed management in order to be operable for large scale mapping scenarios. This method utilizes the on-board RGBD sensor of the robot and performs incremental motion estimations along the robot’s travel. Specifically, visual features (SURF) are matched among consecutive frames and by using the corresponding depth information, 3D matched features are produced. This set of matched features are fed into a Random Sample Consensus (RANSAC) algorithm [18] and the robot’s motion estimation is calculated. Then, the robot estimated poses and the corresponding matched features are treated as graph, the optimization of which is performed using the

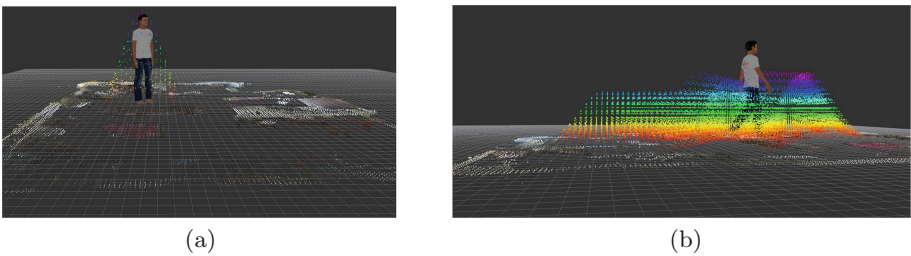
$g^2o$  optimization algorithm [19]. Upon the metric map, the hierarchical model that contains semantic information is constructed. The latter is tree structure that retains the relationships among objects and places, the object supporting surfaces as well as the human standing position describing explicitly the domestic environment in terms of human concepts. The hierarchical mapping comprises a connection among the metric map and the human oriented concepts providing thus the robot with the capacity to apprehend the environment in a human compatible manner. Specifically, the semantic information is stored in an XML schema and comprises the following structure: the house environment is organised in rooms, the room types consist of large objects and frequently visited standing positions, the large objects are related with the robot parking positions and with small objects. The small objects are organized in terms of their attributes, their grasping points and their relations to other objects. Additionally, the room borders are connected with the metric map to define their coordinates, something that is also determined for the large and the small objects as well as the human standing and the robot parking positions. An illustrative example of the metric and hierarchical semantic mapping components is presented in Fig. 1.



**Fig. 1.** (a) A reference image of a scene, (b) the hierarchical semantic model used for the annotation of the mapped environment, and (c) the 3D metric map

### 2.3 Human Motion Intention Modelling

Since all necessary information about the objects in the house has been modeled, this knowledge can be utilized from the robot to execute its navigation routines. Yet, the human presence needs to be considered during the robot's locomotion and, therefore, a human detection engine has been developed using the background information, by subtracting the current robot view from the 3D metric map. After certain optimization techniques on the subtracted frame, a rough human silhouette is extracted, the center of mass of which is utilized to calculate the intention of the human motion. In more detail, the direction of human motion is constantly calculated among consecutive frames and the distance from the human and the standing positions is computed. To model the short term human motion intention, the orientation deviation  $\mathbf{d}_r$  from the current human pose to all standing position is calculated and the distance deviation  $\mathbf{d}_l$  from the current human location to all the standing positions is also determined. Assuming that there are  $N$  human standing positions, the most probable one ( $P_i, i=1, \dots, N$ ) that the human will move towards to, can be determined by minimizing the  $P_{i=1}^N = \operatorname{argmin}(\alpha \mathbf{d}_r_{i=1}^N + \beta \mathbf{d}_l_{i=1}^N)$  where  $\alpha$  and  $\beta$  are regularization parameters that control situations where the modelled environment is congested, i.e. with many furniture where the user has to follow curved paths to reach a standing position. The minimized values of the criterion are sorted from the most probable to the less probable one. All the human paths among his/her current location and the standing positions are calculated using a Dijkstra algorithm. For each point (location in the map) of the calculated human path sequence, an oriented Gaussian kernel is centered, the parameters  $x$  and  $y$  of which, model the personal space of the human. The amplitude  $A$  of the Gaussian kernels that form each human path is reverse proportional to the values of criterion  $P_{i=1}^N$  indicating that the paths with less probability to be followed by the human have diminished weights. An example of the aforementioned strategy exhibited in the Fig. 2.



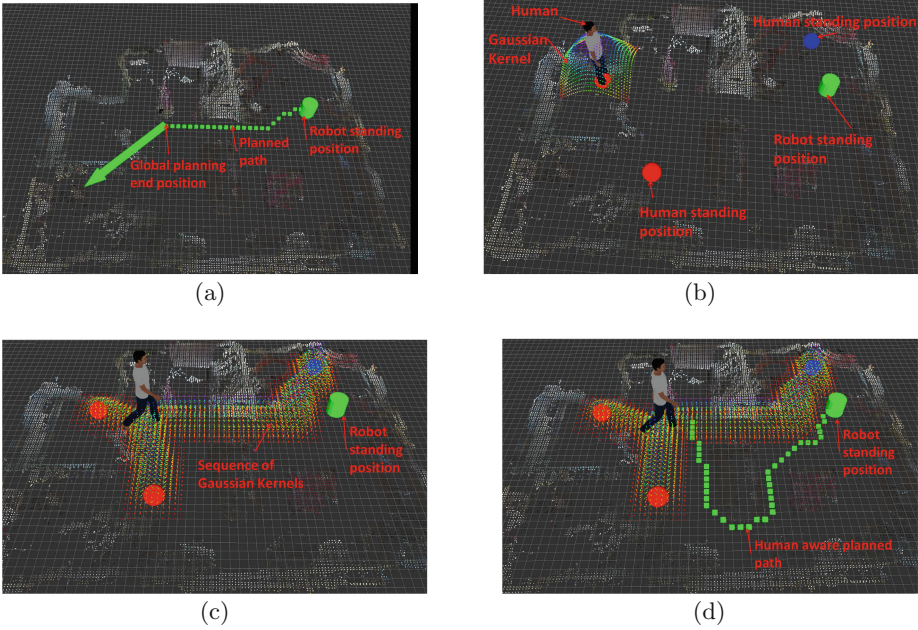
**Fig. 2.** (a) Still human modelled with a single Gaussian kernel and (b) walking human with predicted paths modeled as a sequence of Gaussian kernels with retaining various amplitude  $A$  for the different paths.

## 2.4 Human Aware Navigation

Given the metric map of the domestic environment and the modelling of the human motion intention, the robot should draw a path in order to reach to its target location. The paper in hand examines the occasion of the global path planning where the robot has to find an optimal path that conjugates its current position to a goal one. Although the global planners typically work on static maps, the proposed strategy incorporates dynamically updated metric maps by modelling the human presence in the environment. Specifically, during the robot global planning, all the static obstacles of the metric map are top down projected and together with the sequences of the Gaussian kernels are used to form the map. The weighs of the sequential Gaussian Kernels (i.e. their amplitude  $A$ ) are coded as weights in the metric map, while the respective points of the metric map are declared as lethal obstacles for the path planner. The utilize planning algorithm is the D\* Lite [20] and is executed to find an optimal path for the robot. Following the Proxemics theory, parameterizations of the human presence using Gaussian kernel varies in occasions where the human is relatively still, i.e. standing or sitting, with the occasions where the person is moving. Following the proxemics rules, the personal space of the human is within the range of [0.45–1.2]m; in our method in order to increase safety during robot operations and to avoid unintentional collisions, we selected different social radius to model the human presence and, thus, we set the human personal space to 0.8 and 1.2 for moving and standing still, respectively. This selection was found to be descent compromise among, human comfort and natural robot motion under social constrained environments. A graphical outline of the steps for the proposed method is illustrated in the Fig. 3.

## 3 Assessment of the Proposed Method

The proposed method has been evaluated on realistic acquired data from a regular domestic environment. Firstly, the metric mapping of the house has been performed along with the hierarchical semantic mapping using manual annotations. Three logically inferred human standing positions have been determined within this map i.e. in front of the kitchen, near the sofa in the living room area and near the fireplace. Figure 4(a) illustrates the reference image from the modelled house, while Fig. 4(b) illustrates the constructed 3D metric map along with the standing positions. A human model has been rendered within this environment and its motion has been simulated to arbitrarily perambulate within the house. The robot was placed on specific parking positions in the house and by monitoring the human presence navigation commands where passed. The 3D metric map is considered as a static one and was top down projected while the Gaussian kernels where dynamically updated during the human motion. Once a navigation command was passed to the robot, the Gaussian weights were incorporated along with the top down projected map to form an occupancy grid, on which the global path planner operates. Multiple evaluation scenarios have been performed proving the ability of the proposed method to guide the robot



**Fig. 3.** A graphical outline of the steps of the proposed method; (a) the robot planning without human presence, (b) the model personal area of the human using a Gaussian kernel, (c) the sequences of the Gaussian kernels from the human current position to the standing positions and (d) a human aware planned path

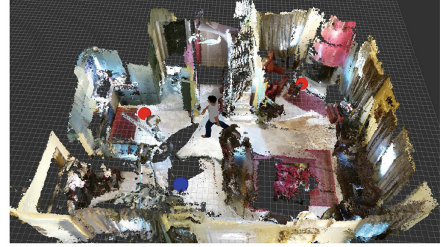
in a way that avoids unexpected collisions with human, selecting thus the optimal path. Exemplar instances of the conducted experiment are summarized in Fig. 4, where the ability of the propose method human aware path planning is illustrated.

## 4 Discussion

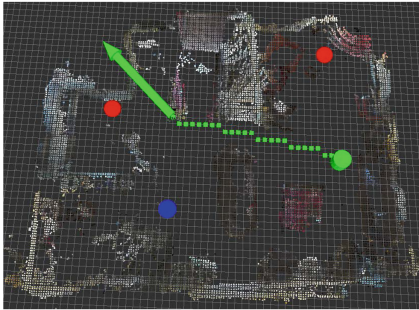
In this work, a human aware robot navigation method has been developed. The robot apprehends its environment in a human conceivable manner and, moreover, adapts its navigation policies in accordance to the humans presence and activities. The method has been evaluated on a simulated environment, yet on realistic acquired data modeling a real house space and exhibited remarkable performance, toward facilitating human comfort during human robot cohabitation, naturalness of the robot behavior, sociability and safety during robot navigation.



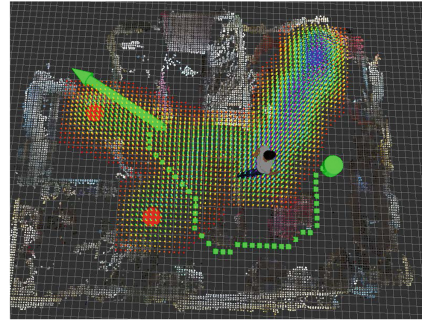
(a)



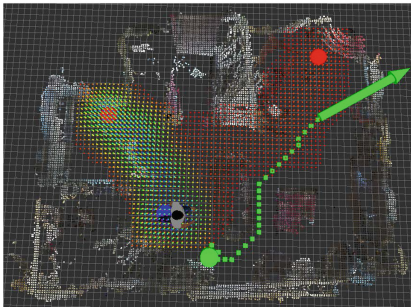
(b)



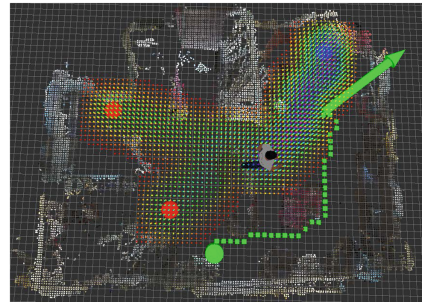
(c)



(d)



(e)



(f)

**Fig. 4.** The first row exhibits the reference image of the modelled environment and the 3D metric map along with the human standing positions. The rest two rows correspond to the experimental assessment of the methodology for various robot starting positions. On the left column the robot does not consider the human presence while on the right column the robot considers the human presence in order to draw a path.

**Acknowledgments.** This work has been supported by the EU Horizon 2020 funded project namely: “Robotic Assistant for MCI Patients at home (RAMCIP)” under the grant agreement with no: 643433.



## References

1. Mitka, E., Gasteratos, A., Kyriakoulis, N., Mouroutsos, S.G.: Safety certification requirements for domestic robots. *Saf. Sci.* **50**(9), 1888–1897 (2012)
2. Thrun, S., Burgard, W., Fox, D., et al.: *Probabilistic robotics*, vol. 1. MIT Press, Cambridge (2005)
3. Filliat, D., Meyer, J.A.: Map-based navigation in mobile robots: I. a review of localization strategies. *Cogn. Syst. Res.* **4**(4), 243–282 (2003)
4. Meyer, J.A., Filliat, D.: Map-based navigation in mobile robots: II. a review of map-learning and path-planning strategies. *Cogn. Syst. Res.* **4**(4), 283–317 (2003)
5. Jian, Y.D., Balcan, D., Panageas, I., Tetali, P., Dellaert, F.: Support-theoretic sub-graph preconditioners for large-scale slam. In: *International Conference on Intelligent Robots and Systems*, pp. 9–16. IEEE (2013)
6. Grisetti, G., Stachniss, C., Burgard, W.: Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Trans. Robot.* **23**(1), 34–46 (2007)
7. Hahnel, D., Burgard, W., Fox, D., Thrun, S.: An efficient fastslam algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In: *International Conference on Intelligent Robots and Systems*. vol. 1, pp. 206–211. IEEE (2003)
8. Durrant-Whyte, H., Bailey, T.: Simultaneous localization and mapping: part i. *Rob. Autom. Mag. IEEE* **13**(2), 99–110 (2006)
9. Bailey, T., Durrant-Whyte, H.: Simultaneous localization and mapping (slam): Part II. *Robot. Autom. Mag.* **13**(3), 108–117 (2006)
10. Kostavelis, I., Gasteratos, A.: Semantic mapping for mobile robotics tasks: A survey. *Robot. Autom. Syst.* **66**, 86–103 (2015)
11. Zender, H., Martínez Mozos, O., Jensfelt, P., Kruijff, G.J., Burgard, W.: Conceptual spatial representations for indoor mobile robots. *Robot. Autom. Syst.* **56**(6), 493–502 (2008)
12. Pronobis, A., Jensfelt, P.: Large-scale semantic mapping and reasoning with heterogeneous modalities. In: *International Conference on Robotics and Automation*, pp. 3515–3522. IEEE (2012)
13. Hall, E.T., Hall, E.T.: *The hidden dimension*, vol. 1990. Anchor Books, New York (1969)
14. Takayama, L., Pantofaru, C.: Influences on proxemic behaviors in human-robot interaction. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 5495–5502. IEEE (2009)
15. Rios-Martinez, J., Renzaglia, A., Spalanzani, A., Martinelli, A., Laugier, C.: Navigating between people: a stochastic optimization approach. In: *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2880–2885. IEEE (2012)
16. Papadakis, P., Spalanzani, A., Laugier, C.: Social mapping of human-populated environments by implicit function learning. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1701–1706. IEEE (2013)
17. Endres, F., Hess, J., Engelhard, N., Sturm, J., Cremers, D., Burgard, W.: An evaluation of the rgb-d slam system. In: *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1691–1696. IEEE (2012)
18. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**(6), 381–395 (1981)

19. Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: g2o: A general framework for graph optimization. In: 2011 IEEE International Conference on Robotics and Automation (ICRA), pp. 3607–3613. IEEE (2011)
20. Koenig, S., Likhachev, M.: Improved fast replanning for robot navigation in unknown terrain. In: IEEE International Conference on Robotics and Automation, 2002 Proceedings of the ICRA 2002, vol. 1, pp. 968–975. IEEE (2002)