An Integrated Robotic System for Spatial Understanding and Situated Interaction in Indoor Environments

Hendrik Zender[1] and **Patric Jensfelt**[2] and **Óscar Martínez Mozos**[3] and **Geert-Jan M. Kruijff**[1] and **Wolfram Burgard**[3]

[1]Language Technology Lab, German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany
[2]Centre for Autonomous Systems, Royal Institute of Technology, Stockholm, Sweden
[3]Department of Computer Science, University of Freiburg, Freiburg, Germany
[1]{zender, gj}@dfki.de, [2]patric@nada.kth.se, [3]{omartine, burgard}@informatik.uni-freiburg.de

Abstract

A major challenge in robotics and artificial intelligence lies in creating robots that are to cooperate with people in human-populated environments, e.g. for domestic assistance or elderly care. Such robots need skills that allow them to interact with the world and the humans living and working therein. In this paper we investigate the question of spatial understanding of human-made environments. The functionalities of our system comprise perception of the world, natural language, learning, and reasoning. For this purpose we integrate state-of-the-art components from different disciplines in AI, robotics and cognitive systems into a mobile robot system. The work focuses on the description of the principles we used for the integration, including cross-modal integration, ontology-based mediation, and multiple levels of abstraction of perception. Finally, we present experiments with the integrated "CoSy Explorer"[1] system and list some of the major lessons that were learned from its design, implementation, and evaluation.

Introduction

Robots are gradually moving out of the factories and into our homes and offices, for example as domestic assistants. Through this development robots will increasingly be used by people with little or no formal training in robotics. Communication and interaction between robots and humans become key issues for these systems.

A cornerstone for robotic assistants is their understanding of the space they are to be operating in: an environment built by people for people to live and work in. The research questions we are interested in concern spatial understanding, and its connection to acting and interacting in indoor-environments. Comparing the way robots typically perceive and represent the world with the findings from cognitive psychology about how humans do it, it is evident that there is a large discrepancy. If robots are to understand humans and vice versa, robots need to make use of the same concepts to refer to things and phenomena as a person would do. Bridging the gap between human and robot spatial representations is thus of paramount importance. Our approach addresses these questions from a viewpoint of cognitive systems, taking inspiration from AI and cognitive science alike.

We believe that a true progress in the science of cognitive systems for real world scenarios requires a multi-disciplinary approach. In this paper we present our experiences of integrating a number of state-of-the-art components from different disciplines in AI into a mobile robot system. The functionalities of our system comprise perception of the world (place and object recognition, people tracking, mapping, and self-localization), natural language (situated, mixed-initiative spoken dialogue), learning, and finally reasoning about places and objects.

The paper describes the principles we used for the integration of the "CoSy Explorer" system: cross-modal integration, ontology-based mediation, and multiple levels of abstraction of perception to move between quantitative and qualitative representations of differing granularity.

Related Work

There are several approaches that integrate different techniques in mobile robots that interact in populated environments. Rhino (Burgard *et al.* 2000) and Robox (Siegwart *et al.* 2003) are robots that work as tour-guides in museums. Both robots rely on an accurate metric representation of the environment and use limited dialogue to communicate with people. Also (Theobalt *et al.* 2002) and (Bos, Klein, & Oka 2003) present a mobile robot, Godot, endowed with natural language dialogue capabilities. They do not only focus on navigation, but rather propose a natural language interface for their robot. The main difference to our approach is that they do not capture the semantic aspects of a spatial entity.

Other works use integration of different modalities to obtain a more complete representation of the environment where the robot acts. (Galindo *et al.* 2005) present a mapping approach containing two parallel hierarchies, spatial and conceptual, connected through anchoring. For acquiring the map the robot is tele-operated, as opposed to our method that relies on an extended notion of human-augmented mapping. Other commands are given as symbolic task descriptions for the built-in AI planner, whereas in our system the communication with the robot is entirely based on natural language dialogue.

Robotic applications using the *(Hybrid) Spatial Semantic Hierarchy* (Beeson *et al.* 2007; MacMahon, Stankiewicz, &

[1]`http://www.cognitivesystems.org`

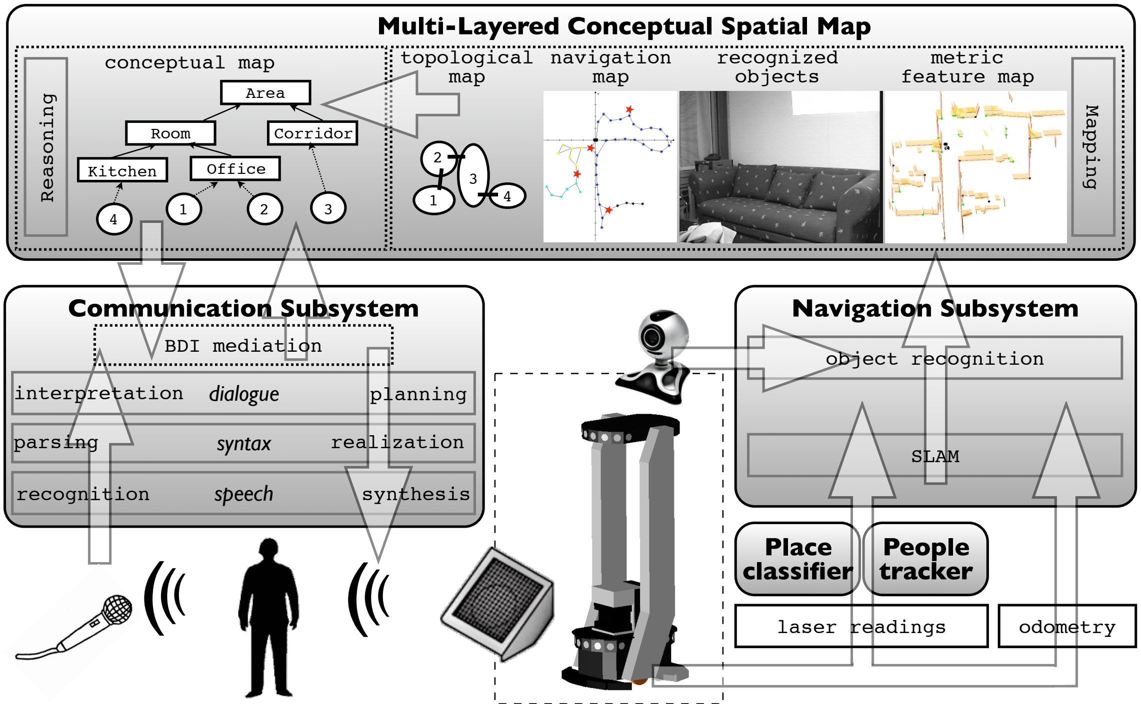

Figure 1: The information processing in the integrated “CoSy Explorer” system.

Kuipers 2006) also use different modalities for the integration of multiple representations of spatial knowledge. These approaches are particularly well-suited to ground linguistic expressions and reasoning about spatial organization in route descriptions. Compared to our implementation these approaches do not exhibit an equally high level of integration of the different perception and (inter-)action modalities.

Finally, the robot Biron is endowed with a system that integrates spoken dialogue and visual localization capabilities on a robotic platform similar to ours (Spexard *et al.* 2006). This system differs from ours in the individual techniques chosen and in the degree to which conceptual spatial knowledge and linguistic meaning are grounded in, and contribute to, the robot’s situation awareness.

System Integration Overview

In this section we give an overview of the different subsystems that our approach integrates. These subsystems will be explained in more detail in successive sections.

Fig. 1 sketches the connections between the different modalities implemented in our robot. The robot acquires information about the environment using different sensors. This information is used for object recognition, place classification, mapping and people tracking. All these perception components are part of the navigation subsystem, which uses the sensors for self-localization and motion planning.

The information is then used to create a multi-layered conceptual and spatial representation of the man-made environment the robot is acting in. Some of the information needed at the conceptual level to complete this representation is given by the user through spoken dialogue. The communication between the user and the robot supports mixed initiative: either the user explains some concepts to the robot, or it is the robot that poses questions to the user.

The complete system was implemented and integrated on an ActivMedia PeopleBot mobile platform. The robot is equipped with a SICK laser range finder with a 180° field of view mounted at a height of 30 cm, which is used for the metric map creation, for people following, and for the semantic classification of places. Additionally, the robot is equipped with a camera for object detection, which is mounted on a pan-tilt unit (PTU).

Just like in human-human communication where spoken language is the main modality, the user can talk to the robot using a bluetooth headset and the robot replies using a set of speakers. The on-board computer runs the Player software for control and access to the hardware, and the Festival speech synthesizer. The rest of the system, including the Nuance speech recognition software, runs on off-board machines that are interconnected using a wireless network.

Perception

Mapping and Localization To reach a high level of autonomy the robot needs the ability to build a map of the environment that can be used to navigate and stay localized. To this end we use a feature-based Simultaneous Localization And Mapping (SLAM) technique. The geometric primitives consist of lines extracted from laser range scans. The mathematical framework for integrating feature measurements is the Extended Kalman Filter. The implementation is based on (Folkesson, Jensfelt, & Christensen 2005).

Object Recognition A fundamental capability for a cognitive system interacting with humans is the ability to recognize objects. We use an appearance based method. Each object is modeled with a set of highly discriminative image features (SIFT) (Lowe 2004). The recognition is achieved with a Best-Bin-First (Beis & Lowe 1997) search approach for fast feature-matching between a new image and the object models. The system is limited to recognizing instances rather than classes of objects.

Place Classification As the robot navigates through the environment, the surroundings are classified into one of two semantic labels, namely `Room` or `Corridor`. The approach uses simple geometrical features extracted from laser range scans to learn a place classifier in a supervised manner (Martínez Mozos *et al.* 2006). This place classification relies on having a 360° field of view around the robot using two laser range finders. As the robot used here has only one laser scanner at the front covering a restricted 180° field of view, we follow (Martínez Mozos *et al.* 2006) and maintain a local map around the robot, which permits us to simulate the rear beams. The learning process can be carried out in a different environment (Martínez Mozos *et al.* 2007).

People tracking Keeping track of the people around the robot is important in an interactive scenario. We use a people tracking system that relies on laser range scans similar to (Schulz *et al.* 2003). People following is realized by sending the position of the robot's guide to the navigation system. The PTU is used to turn the camera towards the person the robot believes it needs to follow, thus providing a basic form of *gaze feedback*.

Language and Dialogue

Dialogue System Our system is endowed with a natural language dialogue system. It enables the robot to have a situated, mixed-initiative spoken dialogue with its human user (Kruijff *et al.* 2007). On the basis of a string-based representation that is generated from spoken input through a speaker-independent speech recognition software, the Combinatory Categorial Grammar (CCG) parser of OpenCCG (Baldridge & Kruijff 2003) analyzes the utterance syntactically and derives a semantic representation in the form of a Hybrid Logics Dependency Semantics (HLDS) logical form (Baldridge & Kruijff 2002). The dialogue system mediates the content from the speech input to the mapping or navigation subsystem in order to initiate the desired action of the robot or to collect pieces of information necessary to generate an answer. The answer string is then generated by the OpenCCG realizer and sent to a text-to-speech engine.

The user can use spoken commands to control the robot, e.g. for near navigation, initiating or stopping people following, or sending the robot to a specific location. Moreover, the user can augment the robot's internal map by naming objects and places in the robot's environment, and conduct a situated dialogue about the spatial organization with the robot.

Interactive Map Acquisition The multi-layered representation is created using an enhanced method for concurrent semi-supervised map acquisition, i.e. the combination of a user-driven supervised map acquisition process with autonomous exploration by the robot. This process is based on the notion of *Human-Augmented Mapping* (Topp & Christensen 2005). In our implementation, the map acquisition process is actively supported by the dialogue system.

The map can be acquired during a so-called guided tour in which the user shows the robot around and continuously teaches the robot new places and objects. During a guided tour, the user can command the robot to follow him or instruct the robot to perform navigation tasks. Our system does not require an initial complete guided tour. It is also possible to incrementally teach the robot new places and objects at any time the user wishes. With every new piece of information, the robot's internal representations become more complete. Still, the robot can always perform actions in, and conduct meaningful dialogue about, the aspects of its environment that are already known to it.

Following the approach in (Kruijff *et al.* 2006), the robot can also initiate a clarification dialogue if it detects an inconsistency in its spatial representation, illustrating the mixed-initiative capabilities of the dialogue system.

Multi-Layered Spatial Representation

Driven by the research question of spatial understanding and its connection to acting and interacting in indoor environments we want to generate spatial representations that enable a mobile robot to conceptualize human-made environments similar to the way humans do. Guided by findings in cognitive psychology (McNamara 1986), we assume that topological areas are the basic spatial units suitable for situated human-robot interaction. We also hypothesize that the way people refer to a place is determined by the functions people ascribe to that place and that the linguistic description of a place leads people to anticipate the functional properties or affordances of that place. In addition to accommodating the high level needs regarding conceptual reasoning and understanding, the spatial representation must also support safe navigation and localization of the robot. To this end we use a multi-layered spatial representation (Zender & Kruijff 2007) in the tradition of approaches like (Buschka & Saffiotti 2004) and (Kuipers 2000). Each layer serves an important purpose for the overall system (Fig. 2).

Layer 1: Metric Map The first layer comes from the SLAM component and contains a metric representation of the environment in an absolute frame of reference. The features in the map typically correspond to walls and other flat structures in the environment.

Layer 2: Navigation Map The second layer contains a navigation map represented by a graph. This representation establishes a model of free space and its connectivity, i.e. reachability, and is based on the notion of a *roadmap* of *virtual free-space markers* (Latombe 1991), (Newman *et al.* 2002). As the robot navigates through the environment, a marker (navigation node) is dropped whenever the robot has traveled a certain distance from the closest existing marker.

We distinguish two kinds of navigation nodes: place nodes and doorway nodes. Doorway nodes are added when

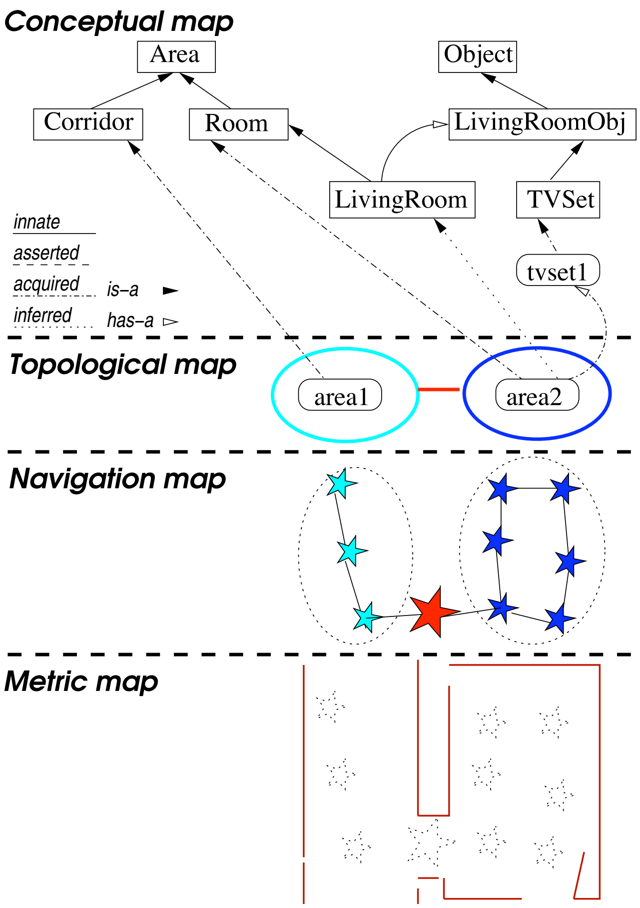

Figure 2: Multi-layered representation

the robot passes through a narrow opening, and indicate the transition between different places and represent possible doors. Each place node is labeled as `Corridor` or `Room` by the place classifier. As the robot moves we store the classification of the last N poses of the robot in a buffer. When a new node is added we compute the majority vote from this buffer to increase robustness in the classification.

Layer 3: Topological Map Previous studies (McNamara 1986) show that humans segment space into regions that roughly correspond to spatial areas. The borders of these regions may be defined physically, perceptually, or may be purely subjective to the human. Walls in the robot's environment are the physical boundaries of areas. Doors are a special case of physical boundaries that permit access to other areas. Our topological map divides the set of nodes in the navigation graph into areas based on the existence of doorway nodes.

Layer 4: Conceptual Map The conceptual map provides the link between the low-level maps and the communication system used for situated human-robot interaction by grounding linguistic expressions in representations of spatial entities, such as instances of rooms or objects. It is also in this layer that knowledge about the environment stemming from other modalities, such as vision and dialogue, is anchored to the metric and topological maps.

Our system uses a commonsense OWL ontology of an indoor environment that describes taxonomies (*is-a* relations) of room types, and typical objects found therein (*has-a* relations). These conceptual taxonomies are handcrafted, and only instances of concepts can be added to the ontology during run-time. The RACER reasoning system can *infer* information about the world that is neither given verbally nor actively perceived. The reasoner works on *acquired* (topological areas, detected objects, area classifications, etc.) and *asserted* (e.g. user says "This is the living room") knowledge gathered during interactive map acquisition together with *innate conceptual* knowledge represented in the office environment ontology. The conceptual map thus enables the robot to generate and resolve linguistic references to spatial areas in a way that accommodates the findings of (Topp *et al.* 2006): namely, that this reference varies from situation to situation and from speaker to speaker.

Experiments

To test the functionalities of our system we ran several experiments in which the robot learns its environment while interacting with a tutor. The experiments were conducted with two different PeopleBot mobile platforms at two different locations.

Before running the experiment, the system needs to have some initial knowledge. For one, the ontology representing the general knowledge about the environment. Furthermore, the classification of places is based on previous general knowledge about the geometry of rooms and corridors. Finally, the robot has to recognize different objects, such as couches or TV sets, using vision. Because we do instance recognition rather than categorization, the objects we want to be recognized must be presented to the robot beforehand.

One of the experiments is explained in more detail in this section. A video of it is available on the Internet.[2] Although the experiment was conducted non-stop, it can be divided into different situations which are explained next.

Place Classification The experiment starts in the corridor, where the user asks the robot to follow him through the corridor, entering a room. Using the method for laser-based place classification the robot correctly classifies the places along the trajectory (`Corridor` and `Room`, respectively) and updates its conceptual representation.

Clarification Dialogues Our door detector creates some false positives in cluttered rooms. Assuming few false negatives in the detection of doors, we get great improvements by enforcing that it is not possible to change room without passing through a door. If this happens, a clarification dialogue is initiated. To test this situation we put a bucket close to a table in the room creating an illusion of a doorway when using only the laser scanner as a sensor. The robot passes through this false doorway and comes back to a previously visited node. It then infers that there is an inconsistency in

[2] `http://www.dfki.de/cosy/www/media`

its spatial representation and initializes a clarification dialogue asking if there was a door previously. The user denies this fact and the corresponding layers in the representation are updated.

Inferring New Spatial Concepts Using the inference on our ontology the robot is able to come up with more specific concepts than the ones the laser-based place classification yielded. While staying in the room, the robot is asked for the current place and it answers with the indefinite description "a room", which is inferred from the place classification. Then the robot is asked to look around. This command activates the vision-based object detection capabilities of the robot. The robot detects a couch, and then a television set. After that, the user asks the robot for the name of the place. Because of the inference over the detected objects and places, the robot categorizes the place as a `Livingroom`.

Situation Awareness and Functional Awareness Here we show how social capabilities can be added to the system taking advantage of our spatial representation; e.g., the robot must behave appropriately when the user is opening a door. Continuing with the experiment, the user asks the robot to follow him while he approaches a doorway. The robot knows from the navigation map where the doorway is and keeps a long distance to the user when he is near the door. Keeping a long distance around doors is motivated by the fact that the user needs more space when opening or closing the door. It then continues following the user by again decreasing its distance to him when he has passed the door.

Improving the Human-Robot Communication and Understanding Finally, we show how we can achieve natural human-robot interaction. As an example, the robot is asked to go to the television. The robot then navigates to the node where the television was observed. The TV set is not a place, but people often indicate only the objects found in a place and assume that the place is known.

Lessons Learned

We believe that integrating different modalities lead to significant synergies in building up a more complete understanding of the spatial organization of an environment, particularly towards a semantic understanding. Moreover, we think that our work made technological progress on the basis of identifying and addressing scientific questions underlying cognitive systems which *understand*.

In addition to the synergies that integrating many components brings in terms of a more complete knowledge and more capabilities, integration also increases complexity and presents problems that arise from the fact that the real world is unpredictable to some extent.

In a scenario where the robot continuously interacts with a user and is facing her/him most of the time, the information content of the sensor input suffers as the user occupies a large part of the field of view. In our case, the camera was mounted on a pan-tilt unit and could have been used to actively look for objects and build a metric map using visual information while following the user. However, this conflicts with the use of the camera to indicate the focus of attention on the user. As a result, most of the time the camera only sees the user and not the environment. Therefore, we opted for giving the user the possibility to instruct the robot to "have a look around." The user's presence not only disturbs the camera-based object recognition but also the performance of the laser-data based place classification. In order to increase the reliability of the resulting classifications, we took two steps. First, a rear-view laser scanner is simulated by ray-tracing in the local obstacle map, and the simulated and the real laser scanner are used together as a 360^o laser-range finder. Second, for determining a robust classification of a navigation node we compute the majority vote of consecutive classifications around that node.

In addition to practical issues, like the ones explained previously, the experiments we run on real environment highlighted new necessities for the system. For example, spatial referencing needs to be improved in both directions of the communication and using several modalities. This would allow the user to indicate a specific object through, e.g., gesture or gaze direction when saying "This is X". This is also an issue when the robot asks "Is there a door HERE?".

Furthermore, experiments highlighted the need for non-monotonic reasoning, that is, knowledge must not be written in stone. Erroneous acquired or asserted knowledge will otherwise lead to irrecoverable errors in inferred knowledge.

When it comes to the natural language dialogue system, flexibility is a centerpiece for robotic systems that are to be operated by non-expert users. Such a free dialogue (as opposed to controlled language with a fixed inventory of command phrases) can be achieved by modeling the grammar of the domain in as much detail as possible. We are currently investigating how to exploit the benefit of corpora of experimentally gathered data from human-robot dialogues in the domestic setting (Maas & Wrede 2006).

Conclusions

In this paper we presented an integrated approach for creating conceptual representations that supports situated interaction and spatial understanding. The approach is based on maps at different levels of abstraction that represent spatial and functional properties of typical office indoor environments. The system includes a linguistic framework that makes for situated dialogue and interactive map acquisition.

Our work also shows there is a limit to certain engineering perspectives we took, and that there are further scientific questions we will need to address if we want to develop more advanced cognitive systems. Integration has played an important role in getting to this point: without a system running in realistic environments, the questions and answers would mostly have been purely academic.

Acknowledgments

This work was supported by the EU FP6 IST Cognitive Systems Integrated Project "CoSy" FP6-004250-IP.

References

Baldridge, J., and Kruijff, G.-J. M. 2002. Coupling CCG and hybrid logic dependency semantics. In *Proc. of the*

40th Annual Meeting of the Association for Computational Linguistics (ACL), 319–326.

Baldridge, J., and Kruijff, G.-J. M. 2003. Multi-modal combinatory categorial grammmar. In *Proc. of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*.

Beeson, P.; MacMahon, M.; Modayil, J.; Murarka, A.; Kuipers, B.; and Stankiewicz, B. 2007. Integrating multiple representations of spatial knowledge for mapping, navigation, and communication. In *Interaction Challenges for Intelligent Assistants*, AAAI Spring Symposium Series.

Beis, J. S., and Lowe, D. G. 1997. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. In *Proc. of the Conference on Computer Vision and Pattern Recognition (CVPR)*, 1000–1006.

Bos, J.; Klein, E.; and Oka, T. 2003. Meaningful conversation with a mobile robot. In *Proc. of the Research Note Sessions of the 10th Conference of the European Chapter of the Association for Computational Linguistics*, 71–74.

Burgard, W.; Cremers, A.; Fox, D.; Hähnel, D.; Lakemeyer, G.; Schulz, D.; Steiner, W.; and Thrun, S. 2000. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence* 114(1–2).

Buschka, P., and Saffiotti, A. 2004. Some notes on the use of hybrid maps for mobile robots. In *Proc. of the 8th Int. Conference on Intelligent Autonomous Systems*.

Folkesson, J.; Jensfelt, P.; and Christensen, H. 2005. Vision SLAM in the measurement subspace. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 30–35.

Galindo, C.; Saffiotti, A.; Coradeschi, S.; Buschka, P.; Fernández-Madrigal, J.; and González, J. 2005. Multihierarchical semantic maps for mobile robotics. In *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*.

Kruijff, G.-J. M.; Zender, H.; Jensfelt, P.; and Christensen, H. I. 2006. Clarification dialogues in human-augmented mapping. In *Proc. of the 1st ACM Conference on Human-Robot Interaction (HRI)*, 282–289.

Kruijff, G.-J. M.; Zender, H.; Jensfelt, P.; and Christensen, H. I. 2007. Situated dialogue and spatial organization: What, where…and why? *International Journal of Advanced Robotic Systems, special section on Human and Robot Interactive Communication* 4(1):125–138.

Kuipers, B. 2000. The Spatial Semantic Hierarchy. *Artificial Intelligence* 119:191–233.

Latombe, J. C. 1991. *Robot Motion Planning*. Boston, MA: Academic Publishers.

Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2):91–110.

Maas, J. F., and Wrede, B. 2006. BITT: A corpus for topic tracking evaluation on multimodal human-robot-interaction. In *Proc. of the Int. Conference on Language and Evaluation (LREC)*.

MacMahon, M.; Stankiewicz, B.; and Kuipers, B. 2006. Walk the talk: Connceting language, knowledge, and action in route instructions. In *Proc. of the 21st National Conference on Artificial Intelligence (AAAI)*.

Martínez Mozos, O.; Rottmann, A.; Triebel, R.; Jensfelt, P.; and Burgard, W. 2006. Semantic labeling of places using information extracted from laser and vision sensor data. In *IEEE/RSJ IROS Workshop: From Sensors to Human Spatial Concepts*.

Martínez Mozos, O.; Triebel, R.; Jensfelt, P.; Rottmann, A.; and Burgard, W. 2007. Supervised semantic labeling of places using information extracted from sensor data. *Robotics and Autonomous Systems* 55(5):391–402.

McNamara, T. 1986. Mental representations of spatial relations. *Cognitive Psychology* 18:87–121.

Newman, P.; Leonard, J.; Tardós, J.; and Neira, J. 2002. Explore and return: Experimental validation of real-time concurrent mapping and localization. In *Proc. of the IEEE Int. Conference on Robotics and Automation (ICRA)*.

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

Siegwart, R.; Arras, K. O.; Bouabdallah, S.; Burnier, D.; Froidevaux, G.; Greppin, X.; Jensen, B.; Lorotte, A.; Mayor, L.; Meisser, M.; Philippsen, R.; Piguet, R.; Ramel, G.; Terrien, G.; and Tomatis, N. 2003. Robox at Expo.02: A large scale installation of personal robots. *Robotics and Autonomous Systems* 42:203–222.

Spexard, T.; Li, S.; Wrede, B.; Fritsch, J.; Sagerer, G.; Booij, O.; Zivkovic, Z.; Terwijn, B.; and Kröse, B. J. A. 2006. BIRON, where are you? - enabling a robot to learn new places in a real home environment by integrating spoken dialog and visual localization. In *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*.

Theobalt, C.; Bos, J.; Chapman, T.; Espinosa-Romero, A.; Fraser, M.; Hayes, G.; Klein, E.; Oka, T.; and Reeve, R. 2002. Talking to Godot: Dialogue with a mobile robot. In *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*, 1338–1343.

Topp, E. A., and Christensen, H. I. 2005. Tracking for following and passing persons. In *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*.

Topp, E. A.; Hüttenrauch, H.; Christensen, H.; and Severinson Eklundh, K. 2006. Bringing together human and robotic environment representations – a pilot study. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Zender, H., and Kruijff, G.-J. M. 2007. Multi-layered conceptual spatial mapping for autonomous mobile robots. In Schultheis, H.; Barkowsky, T.; Kuipers, B.; and Hommel, B., eds., *Control Mechanisms for Spatial Knowledge Processing in Cognitive / Intelligent Systems*, Papers from the AAAI Spring Symposium, 62–66.