Proceedings of the 2006 IEEE/RSJ
International Conference on Intelligent Robots and Systems
October 9 - 15, 2006, Beijing, China

Bringing Together Human and Robotic Environment Representations – A Pilot Study

Elin A. Topp, Helge Huettenrauch, Henrik I. Christensen, and Kerstin Severinson Eklundh
School of Computer Science and Communication (CSC)
Royal Institute of Technology (KTH)
10044 Stockholm, Sweden
Email: {topp,hehu,hic,kse}@csc.kth.se

Abstract— **Human interaction with a service robot requires a shared representation of the environment for spoken dialogue and task specification where names used for particular locations are depending on personal preferences. A question is how such human oriented models can be tied to the geometric robotic models needed for precise localisation and navigation. We assume that this integration can be based on the information potential users give to a service robot about its working environment. We further believe that this information is best given in an interactive setting (a "guided tour") in this particular environment. This paper presents a pilot study that investigates how humans present a familiar environment to a mobile robot. The study is set up within our concept of Human Augmented Mapping, for which we assume an initial "guided tour" scenario to teach a robot its environment. Results from this pilot study are used to validate a proposed generic environment model for a service robot.**

I. INTRODUCTION

Service robots are often mobile platforms that provide assistance to humans. Thus, a basic competence for such a mobile robotic system is the ability to move from one place to another to provide its services. This requires navigation and localisation functionalities. Mobile robots can navigate on the basis of metric, often feature based maps, and they can build those maps autonomously while exploring an environment for the first time. Methods in robotics research are dealing with this issue of Simultaneous Localization and Mapping (SLAM) [7]. Humans have a topological, (partially) hierarchical, view on their environment [16]. To enable a service robot to perform tasks for users in arbitrary environments a spatial representation that is understandable for both, the robot and the user, is needed. Assuming an indoor environment such as a home or office building we mean by such a commonly understandable map representation that the robot's notion of the environment appears to be the same as the one the user might refer to. In other words, we need to build a common ground for interaction [4]. Supposedly the user knows the environment in question very well, while the robot is "added" to it. Thus, a common model is likely to depend on a personal view a potential user has on the environment. We propose to use this individual user perspective and understanding of the physical environment to "personalise" the robot's model. Provided that the robot has some general world knowledge, it is adapting the specific details based upon the user's preferences. We assume a scenario of a "guided tour" to be an appropriate way to "teach" a mobile service robot (equipped with the necessary sensor systems and functionalities) its environment and to test for the possible effects of such a user controlled personalisation of the robot's knowledge representation. The user can guide the robot around and name important places (e.g., rooms or large, mostly static, objects). Concurrently the robot can build a (metric) map of the environment. This map is augmented by the user's information which allows to integrate the robot's metric, feature-based map with the topological map representation of the user. Figure 1 illustrates a scene from such a guided tour.

Fig. 1. Illustration of a user showing the kitchen to her robot.

Central issues for the comprehensible representation of environments are a) the question of what strategies of presenting an environment would be used by different users, and b) how the given information can be incorporated into an environment model to actually satisfy the requirements for a shared representation.

In earlier work we introduced the concept of *Human Augmented Mapping* (HAM, [20], [21]), which allows us to subsume different aspects of Human-Robot Interaction (HRI) and robotic mapping. Different aspects of interaction as well as posture and positioning of subjects in relation to a robot were studied previously in our laboratory with a Wizard-of-Oz experiment [9]. In this case the experiment was also based on a "guiding the robot around" scenario, but the environment was limited to one room and the robot used in the study was controlled remotely.

The present paper describes a user study that investigates, how different users present a well known environment to a robot. The study is supposed to give implications on how a generic environment model can incorporate the user's information. Thus, we firstly explain our proposition of such a model together with the necessary background from psychological findings. Starting from this initial model we can demonstrate with results from a pilot study how it can be used in the context of the interactive presentation of a given environment.

A. *Outline of the paper*

The rest of this paper is organised as follows. We give an overview of related work and refer to hierarchical environment representations motivated from results in Cognitive Science and Psychology to propose a general high-level robotic environment representation in sections II and III respectively. Sections IV and V explain the design of the study and results from pilot trials, and in section VI we draw conclusions on the study setup and its results.

II. REPRESENTATION OF SPACE

In "The intelligent use of space" [11] Kirsh stated that in order to understand complex (human) models of an environment, we have to observe the interaction of the (human) agent with and within this environment. Based on those observations, corresponding *robotic* models can be obtained. Transferred to the interaction of two agents in and about a certain environment, observations from human-human interaction could be the basis for a general robotic environment model. We propose to use the following approach to achieve this adaptation of the robotic environment model. In our study a *user* and a *robot* are observed during their interaction when the user is showing an environment to the robot. The analysis of the observations will allow to better understand what robotic model can be used to build a "shared model" that both the user and the robot can refer to later.

Kyriakou *et al.* investigate in a study that uses a miniature robot on a table top street "map" how computer vision can be used to follow verbal guiding directions by having subjects guide the robot with commands like "follow this road to the station, then turn left" [15]. This is another, spoken dialogue based form of "guiding a robot" without actually being part of a collaboratively operating duo in a co-present, embodied interaction. Furthermore an important condition for such a setup is the a priori availability of a map representing the environment in question that contains all items a potential user considers important at the respective position. Since we wanted to learn about what users present in a given environment and how they do that, we consider such a setup too limiting for our purpose.

Kuipers *et al.* presented a mapping approach that represents the environment as a combination of global topological and local metric maps [14]. The main aspect of this work however is the handling of large scale maps, that can be achieved by representing the environment as local metric maps that are linked in a global, topological (and as such hierarchical) representation. Also in other approaches the segmentation of metric maps and/or organisation of them into hierarchies has been studied as part of research in SLAM, but primarily as a way to limit computational complexity [3], [18], [19, among others].

Techniques to interactive robotic mapping have been reported by Diosi *et al.* [5] as well as by Althaus and Christensen [1]. Diosi *et al.* obtain a purely metric spatial representation of an office environment by guiding a robot around and defining labelled rooms. Althaus and Christensen model the environment rather as a topological graph, but do not consider different levels of granularity in their representation. We believe that not only rooms are needed, but also a lower level of complexity has to be integrated in a topological model. This allows, for example, to integrate specific places (objects) into the specified areas.

A number of different theories on how spatial relations are acquired and represented by humans have been proposed throughout the years. In a comprehensive study on observable use of mental space models McNamara came to the conclusion that a *partially hierarchical model* supported his findings most appropriately [16]. Following these findings, we assume that users would not necessarily follow a hierarchical order when explaining the environment to the robot. Such a hierarchical order would be to first explain all rooms on a floor and then present all possible / important items within them. We expect users to present items and location in the order they are encountered during a tour. Also we assume, that according to personal preferences not all possible rooms or locations will be presented. Transferring the idea of a partially hierarchical human mental model to our guided tour implies that the assumed robotic environment model has to be able to handle spatial information given in arbitrary order. Thus we propose a hierarchical structure, that incorporates the required flexibility with generic entries on each level in which places can be represented. We express this assumption as well in a number of working hypotheses for the pilot study in section IV-C.

To incorporate other dimensions, particularly the *functionality*, the hierarchy needs to be extended. Galindo *et al.* [8] propose *Multi-Hierarchical Representations* to incorporate semantic information into their environment model used for mobile robotics. In their work two hierarchies, one conceptual, the other spatial, are linked through anchoring to enable reasoning. As stated previously, we assume a (partially) hierarchical representation of the environment, but do not incorporate any semantics so far.

Along with the functionality one issue is the personalisation [2] of a particular environment representation. From intuition one would expect that individuals have different preferences and ideas how to interpret and use their surroundings. However, we equally believe that these individual perspectives are based upon a (more or less vague) common understanding and agreement, e.g., on the semantics of the concept (room and term) "kitchen". We consider the fact that different users might give different information to the very same robot as an issue of future work. Our environment model though is flexible enough to model those individual differences within the same framework.

III. REPRESENTATION IN HUMAN AUGMENTED MAPPING

With our concept of *Human Augmented Mapping* [20] we can establish the link between a robotic map that enables the robot to navigate and the environment representation of a user (also referred to as "cognitive map" [12]). We use a graph based model of the environment, described in section III-A to incorporate the information that is given interactively into our framework. This framework uses a structure of different (robotic) map representation one of which is thus the high level graph representation described here. We further do not assume a full, initial environment model that allows the robotic system to instantiate content entities by autonomous exploration, but consider a more general, structural model that can be filled with personalised information and that can be revised if necessary.

A. A hierarchical graph structure

We model the environment by using a hierarchy of graphs. The main concepts we incorporate so far are *locations* (or places) and *regions*.

We define ***locations*** as

> *Specific positions/areas that can represent the position of large objects that are considered static.*

Such *locations* can for example be a closet, a refrigerator or a sofa.

A ***region*** is then

> *Any portion of space that is large enough to allow for different locations in it, or at least large enough to navigate in it.*

Typically this would be rooms or corridors or parts of those.

Regions are represented by local geometric features. They are internally linked by metrical connections. To maintain the hierarchical structure but allow for partially hierarchical representations as well, we assume a "generic *region*" in which we start the mapping process. With the "generic *region*" we can guarantee that all mapped areas are represented as a *region* on the respective level of the hierarchy. As soon as a *region* is assigned a name (label) by the user, it is stored together with the corresponding local representation that might already contain information on specific *locations*. When a specified *region* is left, and the adjacent area was not explored before, this "new" and yet unexplored *region* becomes a "generic *region*". Only specified *regions* are entities in the hierarchy that form a new branch from the respective level downward. This makes it possible to define a specific *location* in a *region*, that is not (yet) specified, for example, to point out the "entrance" in the "generic *region*" (e.g., the corridor in this situation).

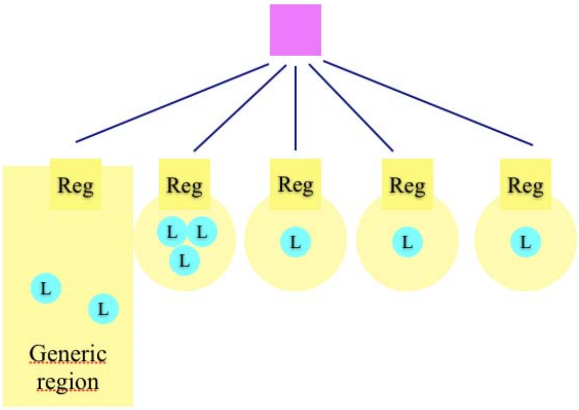

Fig. 2. Our hierarchical structure visualised in a simplified way. The "generic region" is used to store "locations" that were presented in a non-specified "region".

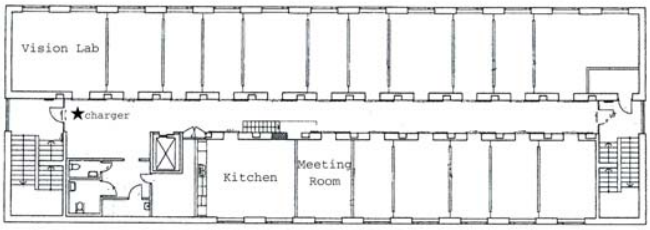

Fig. 3. The floor plan of our office environment on which the experiments took place. The star marks the starting point, where subjects encountered the robot

Figure 2 shows the model with two levels in the hierarchy, *regions* and *locations* and a top level, that is not incorporated so far, but could be a "floor", for example. Links between *regions* that form the topological structure are left out for clarity of the image.

IV. THE PILOT STUDY

We conducted a pilot study to test our proposed robotic environment model against the information on a specific office area given explicitly by a human user to a mobile robot. Additionally the pilot study serves as a proof-of-concept for a more comprehensive user study. The pilot study comprised trial runs with five subjects of about 45 minutes duration each. Within this time period the subjects spent about 20 minutes interacting with the robot, the rest of the time was used for instructions before and short interviews after the sessions. All participants received a cinema ticket as compensation for their participation.

A. Scenario

The scenario of the study was a "guided tour" through a portion of an office building. Figure 3 shows the floor plan with offices (not marked), the kitchen, the meeting room and the computer vision laboratory of our office building where the trials were conducted. Subjects were instructed to show the robot around in the environment so that it later could perform non-specified service tasks. The tasks were described generally as fetch-and-carry tasks (go and fetch something for the user). To do this the robot needed to have “seen” the respective *locations* (a more detailed description of the instructions and the technical realisation is given in sections IV-B.2 and IV-B.3).

B. Method

In the following section we explain our selection of subjects, the instructions given to them, and the methods used for data collection.

1) Subjects: As important precondition to our pilot study we assumed subjects to know the environment they would guide the robot around in. This assumption on user qualification and experience is important and based on the belief that potential users will ”add” service robots to their (to them already well known) homes and offices. Subjects were therefore recruited from the laboratory environment the experiments took place in. To require familiarity with the robot’s operation area is thus a design choice that differs from other human-robot interaction studies, where subjects often are invited into an unfamiliar or even ”simulated” environment. The deliberate choice however comes at the price of familiarity with robotic systems in our laboratory. To assure at least some variety in familiarity with robotic systems we selected our five subjects actively among the members of the Computer Vision and Active Perception Laboratory[1] that hosts a part of the Centre for Autonomous Systems[2] on our campus. The group of pilot subjects included one secretary (familiar with robots from films, presentations and frequent encounters in the office environment, but not familiar with their internals), three computer vision researchers, one of them somewhat familiar with the internals of robotic systems, and one robotics researcher from the field of robotic mapping. Thus, the participants represented the full range of robot expertise available at the laboratory. All subjects had been working in this particular office environment for about two years.

2) Instructions: Our subjects were given an instruction sheet that explained the task and the functionalities and abilities of the robot[3]. The task was to use a number of speech commands and explanations to make the robot follow and to point out “everything” that the subject considered important for the robot to know on the floor the experiments took place on. The time frame given to the subjects for the completion of their task was about 20 minutes (15 minutes for the guided tour and five minutes to test the robot’s “memory”). In the instruction none of the words *region, location, position* or *place* was named. We referred to *“everything, that you think the robot needs to know”, “whatever you pointed out before”, etc.*, so that subjects were completely free to decide, what they would present to the system and how they would name it. Neither did we give any example (e.g., “You can name for example the coffee maker”), to avoid priming the participants on items that a particular subject would not have considered important in the first place. Subjects were offered to ask for help before and during the actual experiment, and knew that they could abort the experiment at any time.

3) Technical realisation: The study was performed with a commercially available Performance PeopleBot by ActivMedia[4]. In a previous study this robot was used in a Wizard-of-Oz-setting [9], where the robot’s functionalities were remotely controlled or simulated by two experiment leaders. For the technical realisation of our study scenario we used a laser range data based tracking and following system [20], which has been extended to incorporate a metric laser range data feature based SLAM method [6] and an input option to label *regions* or *locations* with name tags. Basic platform control and access to the sensors and to a text-to-speech system (Festival[5]) are provided by the Player/Stage[6] software library.

The system represents labelled *regions* and *locations* in a simple graph structure that distinguishes between specifically labelled positions (“defined place”) and internal navigation nodes. The internal nodes are used to build a navigation graph on which the system can perform a graph search to plan a path to a previously named position [6], [13].

The verbal interaction of the user with the robot was still controlled by the experiment leader, i.e., utterances from the subject were interpreted by the experiment leader and labels of *locations* or *regions* as well as user commands were fed into the system via a graphical user interface (GUI) running on a laptop. This allowed us to reduce the trial’s complexity by, e.g. avoiding problems due to speech recognition failures. For verbal feedback though we used the text-to-speech system with precoded utterances, so that the robot could give spoken feedback about its own state and the task given to it (e.g., “`I will follow you`”, “`Stopped following`”, “`I think I have lost you`”, “`Stored <item>`”).

As the experiment took place on an entire floor of the building one experiment leader (the robot’s supervisor) followed the subjects to observe the experiment. To assure the subjects’ safety the system allowed switching from autonomous following based on the mentioned tracking approach to full remote control by invoking a soft joystick implementation.

We provided the robot with two different behavioural strategies for the labelling of either a *location* or a *region*. The choice which strategy to use was made by the experiment leader according to the definitions of *regions* and *locations*. If a *location* (including a “link” to a *region*, e.g., a doorway) was presented, the robot did not move and stated immediately, that it stored the given information. If on the other hand a *region* was presented, the robot stated that it needed to have a look around and performed a 360° turn before confirming the information.

---

[1] http://www.nada.kth.se/cvap

[2] http://www.nada.kth.se/cas

[3] The original instruction sheet can be found at http://www.csc.kth.se/∼topp/research.html

[4] http://www.activmedia.com

[5] http://www.cstr.ed.ac.uk/projects/festival/

[6] http://playerstage.sourceforge.net

4) Observation methods and data collection: By storing the data provided by the robot's sensory systems we could get a full "real time" (graphical) representation of each of the trials. Additionally we recorded the experiments with two digital video cameras. One of them recorded the robot's point of view. The other camera recorded an external perspective, accompanying the user and the robot. After their experiments our pilot subjects were asked to answer a number of questions on the experiment in a short interview. This interview was scripted with a list of prepared questions on the motivation for naming or not naming certain *locations* or *regions* and for the handling of the tour scenario. We were particularly interested in whether subjects had perceived the behaviour of the robot differently depending on what was pointed out (a *location* or a *region*) and what they thought about this difference.

C. Hypotheses

We wanted to study how different individuals present a known environment to a mobile robot and relate the resulting information to an environment model we consider appropriate in the context of Human Augmented Mapping. We assumed that humans do not necessarily follow a hierarchical structure when they present a known environment to a robot (see section II). Thus, we started out with a number of working hypotheses (WH) about the way subjects would present the *regions* and *locations* they considered relevant, as well as about the entities that would be named:

WH1: "users do not name all *regions* in the environment",
WH2: "users point out *locations* in *regions* they did not name before", and
WH3: "users point out *regions* without entering them".

We use these hypotheses to test whether the observations from the pilot study can validate our environment model. We did not formulate a specific hypothesis for the dependency "familiarity with robotic systems vs. way of explaining the environment to a robot" to explore this issue. Nevertheless we expected a robotic researcher particularly familiar with map representations to be more explicit than subjects not familiar with robotic environment representations. Further we speculated that the difference in the robot's behaviour would allow the subjects to "understand" the robot's internal processes, when storing either a *region* or a *location*. This goes along the line of argumentation of Sidner *et al.* [17], who found that a slight change in the feedback (nodding vs. not nodding) their robot gave during an experiment influenced the interaction with it, since it was easier to understand.

V. RESULTS OF THE EXPERIMENT

In this section we present the results from our pilot study. We are aware that the data set is small and consequently not entirely representative. However, it is possible to analyse the outcome of the experiments in terms of *occurrence* of different phenomena. Additionally, our observations and the subjective answers we obtained in the short interviews allow us to investigate how subjects reasoned about their strategy to show *regions* and *locations* and to improve the system for further studies.

As one outcome we gained an increased confidence that the methodology for conducting the pilot study actually can be applied to show the validity of our approach in getting information on individually different ways of building map representations in an interactive, joint process. Furthermore we believe that the soundness of our environment model seems to be supported by its demonstrated ability to handle the diverse situations observed. In table I we summarise the quantifiable results to give an overview over our observations and statements from the interview.

TABLE I
QUANTIFIABLE RESULTS FROM THE PILOT STUDY

Observation Subject	VR	VR	VR	SE	RR
Interaction time	22min	19min	11min	25min	24min
# *regions*	4	2	–	2	2
# *locations*^{I}	4	4	5	4^{II}	8^{III}
# *regions* w/o loc.	3	2	–	1	1
# loc. w/o *region*	3	4	5	2	3^{IV}
# *regions* w/o entering	1	2	1	1	–
Behaviour noticed	Yes	Yes	–	No	Yes
– appropriate	Yes	Yes	–	–	Yes
– appears smart	Yes	No	–	–	Yes

VR: Vision researcher, SE: Secretary, RR: Robotics Researcher

I: including *regions* that were only pointed to
II: including one small object (a salt shaker)
III: including one person and two doorways to respective rooms
IV: excluding doorways

A. Observations

All subjects but one used the full time frame to present the environment to the robot. The "tour" started for each experiment at one end of the corridor (see Figure 3), where the robot awaited its user at the charger.

All subjects took the robot into the kitchen, probably because this is a central room in our office environment, both from a topological, a functional, and a social point of view. However, the observed diversity in strategies to introduce the kitchen to the robot was quite large, ranging from the pure introduction of *the kitchen* over some combination of *specific locations in the kitchen and the kitchen itself* to *specific locations only*. Already from our small sample of data we can thus conclude that the variety of explicitly stated information that a robotic system in an interactive mapping process would have to cope with is large and needs to be handled by the robot's environment representation. More specifically, these differences in naming observed for the kitchen and its *locations* correspond to our expectations expressed in hypotheses `WH1` and `WH2`.

We also noted that none of the subjects named the corridor or hallway – leading toward and being traversed on the way to the kitchen – itself as a *region*, but all of them pointed out specific *locations* in it, which gives us further evidence for our hypotheses `WH1` and `WH2`. In a number of situations only the door to a *region* was pointed out, e.g., for the elevator or an office, confirming our expectation expressed in hypothesis `WH3`.

When asked about their strategy in the post-trial interview, most subjects stated that they had pointed out those *locations* or rooms they personally considered important. Other rooms or *locations* were therefore left out on purpose. In some cases the subjects stated that the time constraints given by the experiment leaders kept them from presenting further items to the robot. We see this as a sign of a strategy to personalise the robot's environment representation to personal needs and preferences while trying to adhere to the trial settings. A possible implication of this observation is to increase the time limit for the interaction with the robot, or to run multiple trial sessions with the same subject.

We asked all subjects that had presented a mixture of rooms and *locations* (four out of five), if they had perceived the difference in reaction of the robot (turning by 360° for a *region* vs. not turning for a *location*). Three out of those four answered that they had observed the difference in behaviour. All three stated that this behaviour seemed *appropriate* and/or made the robot *look smart*, since it obviously wanted "to understand its surroundings". Despite some technical problems all subjects expressed their satisfaction with the flow of interaction and communication as well as the robot's performance.

B. Particular situations

Even with the limited number of subjects we were able to observe some interesting strategies for the presentation of the environment. We relate the observations to statements from the short interviews where possible and try to establish the relationship and importance to our environment model.

1) Pointing out persons: In two cases subjects tried to point out a person. In one case the person was sitting at her desk and the robot was made to store the respective *location* (from where the robot had been shown the person) by the experiment leader. In the other case the subject reacted spontaneously to someone suddenly walking out of the elevator right in front of the robot. Here the robot was not made to store the person's name. This decision was made as the presentation of a person *in* her office can be interpreted as "presenting the room", but in the dynamic situation in front of the elevator it was not possible to tie any spatial concept to the person. An interpretation of the observed behaviour is that it is necessary to consider the influence of *spontaneous encounters* and *opportunities* which might prompt users to incorporate them into the interaction with the robot. For the presentation of "links to rooms" however, the proposed model is suited since it can incorporate connections between *regions*.

2) Explaining no rooms at all: One of our subjects concentrated only on *locations* (e.g., pigeon holes, coffee machine, refrigerator) and did not name any room (or other *region*). The subject considered if more imported for the robot to know about specific *locations* where it could and should *do* something, rather than how to get to a particular room. We see this as a strong evidence for hypothesis `WH2`. With the help of the "generic region" it is clearly possible to store the given information in the proposed model. The question though, in how far the robot should take the initiative to learn more about possibly detected delimited areas (rooms) remains to be answered.

3) Explaining doorways: We expected our subject with robotic research and mapping experience to be more precise and explicit than other subjects. This expectation could be confirmed by the fact that the doors to shown rooms were pointed out explicitly *when the robot was standing exactly in the door opening*. We could also observe that both named rooms were actually entered. Since only two rooms were presented during this experiment, we can of course not generalise, but we consider at least our expectations for the robotics researcher's strategy confirmed. On a more abstract level we might need to look for effects caused by different levels of understanding the robot's internals and provide plausible metaphors to be guiding the interaction design strategy accordingly. For our graph model the explicit information about doors (gateways) can be used to delimite *regions* as a clarifying addition to other methods of space segmentation.

C. General relation to our environment model

Our observations show that even with a small, rather homogeneous group of subjects different ways to show and explain the environment are to be expected and dealt with, depending on the individual view and use of particular items and rooms. We see these differences as a proof of concept for our proposed environment model (as introduced in section III-A) which we consider usable as a higher level model within a robotic hierarchy of maps that incorporates also a basic metric and an intermediate level topological representation.

A general assumption is that a given robotic system has the ability to perceive *regions* that are delimited from other *regions* autonomously. Such a method is currently investigated [21]. We also assume that we have a general knowledge model distinguishing between *regions* and *locations* and a dialogue model that uses this knowledge base. From the experiments we collected evidence on the strategy of users to point out a *region* by only showing the respective door leading into the *region* to be named. The detection of fix points and variations in the user's behaviour and spoken utterances could give a signal on the actual intention. These issues are subject to current investigations and will be incorporated in further studies. Summarising we believe that our model holds at least for the variety of strategies to present a known environment to a robot observed in our pilot study.

D. Interaction issues

During the pilot experiments we observed several issues of the technical realisation that had consequences for the actual interaction between subjects and the robot.

Despite the instruction of providing room for the robot after a "Follow-me"-command, subjects waited in front of the robot for its initial movement. This lack of manoeuvering space prevented in a few cases the robot to actually start the following of the user. We interpret this observation as a sign that the robot's verbal feedback to follow ("I will follow you") was not enough, to indicate that it would actually follow. From carefully studying the recorded interaction on video we concluded that the robot needs to indicate with a body (movement) gesture like turning toward the user that it has seen, heard, and understood the user and is ready to move [10]. A similar problem occurred, when subjects made the robot face something to "look at it" and wanted to continue the tour afterwards. We plan to make the robot turn back toward the user to indicate that it is ready to continue after storing a presented item.

VI. CONCLUSION AND FUTURE WORK

In this paper we presented two important aspects of our concept of Human Augmented Mapping, namely the environment representation of the robot and the interactive context that allows to build a shared mental model of an environment. We explained our approach to a robotic map representation and showed to what extend this representation holds in different situations within an interactive mapping process. A pilot study was conducted to investigate strategies of users to present a for them well known environment to a robot.

Despite the small number of subjects in the study we were able to observe various strategies to present a known environment to the robot in a "guided tour". Partially this diversity might be due to differing knowledge in robotics or the individual interest in the robot that our subjects had. However, we can state that all the different situations or strategies characterised in a number of hypotheses we formulated actually occurred at least once. The variety in presentation strategies we observed and the self reflecting comments on them showed us that there is a need for flexible representations, when a robotic system should be used and guided around by different users.

We got mostly positive feedback on the behaviour of the robot, especially on a "*region* observation strategy" we implemented to enable subjects to understand the internal processes of the robot to some degree. This assures us to keep such a behavioural strategy for further studies to allow subjects to understand more of the internal procedures of the robot.

The results from the pilot study encourage us to use the proposed setup in a more comprehensive user study and to investigate the applicability of the proposed environment model in a robotic framework in more detail.

VII. ACKNOWLEDGEMENTS

The work described in this paper was conducted within the EU Integrated Project COGNIRON ('The Cognitive Robot Companion' - www.cogniron.org) and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

We also thank Patric Jensfelt and John Folkesson for their technical help with the implementation of the system.

REFERENCES

[1] P. Althaus and H.I. Christensen. Automatic map acquisition for navigation in domestic environments. *in Proc. of the International Conference on Robotics and Automation,* September 2003, Taipei, Taiwan.

[2] J. Blom. Personalization: a taxonomy. *CHI'00 extended abstracts on Human factors in computing systems,* 313-314, ACM Press.

[3] M. Bosse, P. Newman, J. Leonard, S. Teller. SLAM in Large-scale Cyclic Environments using the *Atlas* Framework. *International Journal of Robotics Research,* 23(12):1113-1140, December 2004.

[4] H.H. Clark and S.E. Brennan. Grounding in Communication. *Teasley, Resnick and Levine (Eds.), Perspectives on socially shared cognition Reprinted in R.M. Baecker (Ed.), Groupware and computer-supported cooperative work: Assisting human-human collaboration.* San Mateo, CA, Morgan Kaufman, 1991

[5] A. Diosi, G. Taylor and L. Kleeman. Interactive SLAM using Laser and Advanced Sonar. *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2005, Barcelona, Spain.

[6] J. Folkesson, P. Jensfelt, and H.I. Christensen. Vision SLAM in the Measurement Subspace *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2005, Barcelona, Spain.

[7] J. Folkesson. Simultaneous Localization and Mapping with Robots. *Doctoral Thesis, School of Computer Science and Communication, Royal Institute of Technology (KTH),* Stockholm, Sweden, September 2005.

[8] C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J.A. Fernández-Madrigal, and J. González. Multi-Hierarchical Semantic Maps for Mobile Robotics. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* August 2005, Edmonton, AB, Canada.

[9] A. Green, H. Huettenrauch, K. Severinson-Eklundh. Applying the Wizard-of-Oz Framework to Cooperative Service Discovery and Configuration. *in Proc. of the 13th IEEE International Workshop on Robot and HumanInteractive Communication,* September 2004, Kurashiki, Okayama, Japan.

[10] A. Green, H. Hüttenrauch. Making a Case for Spatial Prompting in Human-Robot Communication. *Workshop on multimodal corpora: from multimodal behaviour theories to usable models, International Conference Language Resources and Evaluation* May 2006, Genoa, Italy

[11] D. Kirsh. The intelligent use of space *Artificial Intelligence,* 73:31–68, 1995.

[12] R.M. Kitchin. Cognitive Maps: What Are They and Why Study Them? *Journal of Environmental Psychology,* 14(1):1–19, 1994.

[13] G.-J. Kruijff, H. Zender, P. Jensfelt, and H.I. Christensen. Clarification dialogues in human-augmented mapping. *in Proc. of the ACM Conference on Human-Robot Interaction,* March 2006, Salt Lake City, UT, USA.

[14] B. Kuipers, J. Modayil, P. Beeson, M. MacMahon, and F. Savelli. Local Metrical and Global Topological Maps in the Hybrid Spatial Semantic Hierarchy. *in Proc. of the IEEE International Conference on Robotics and Automation,* April 2004, New Orleans, LA, USA.

[15] T. Kyriakou, G. Bugmann, and S. Lauria. Vision-based urban navigation procedures for verbally instructed robots. *Robotics and Autonomous Systems,* 51:69–80, January 2005.

[16] T.P. McNamara. Mental Representations of Spatial Relations. *Cognitive Psychology,* 18:87–121, 1986.

[17] C.L. Sidner, C. Lee, C.D. Kidd, N. Leash, C. Rich, Explorations in engagement for humans and robots. *Artificial Intelligence,* 166(1):140-164, 2005

[18] N. Tomatis, I. Nourbakhsh, and R. Siegwart. Hybrid Simultaneous Localization and Map Building: a Natural Integration of Topological and Metric. *Robotics and Autonomous Systems,* 44:3–14, 2003

[19] A. Tapus, G. Ramel, L. Dobler, and R. Siegwart. Topology Learning and Place Recognition using Bayesian Programming for Mobile Robot Navigation. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 2004, Sendai, Japan.

[20] E.A. Topp and H.I. Christensen. Tracking for Following and Passing Persons. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* August 2005, Edmonton, AB, Canada.

[21] E.A. Topp and H.I. Christensen. Topological Modelling for Human Augmented Mapping. *in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* October 2006, Beijing, China.