FP6-IST-002020

COGNIRON

The Cognitive Robot Companion

Integrated Project

Information Society Technologies Priority

D5.1.1

Methods for Hierarchical Probabilistic Representation of Space

Due date of deliverable: 31st December 2004
Actual submission date: 31st January 2005

Start date of project: January 1st, 2004 **Duration**: 48 months

Organisation name of lead contractor for this deliverable: EPFL

Revision: Final
Dissemination level: PU

Executive Summary

This report basically looks into the development of a new representational methodology, which may be useful for planning, navigation, manipulation and interaction related tasks, for a robot. The aim is to develop a multi-resolution probabilistic knowledge representation framework which can incorporate multi-modal information. Thus, it looks into the development of an extensible (general) representational methodology that can cope with the complexity in the environment. A general paradigm is presented and several approaches to the problem, taken by the individual partners, are reported. The report also relates the various approaches on the basic paradigm.

Role of "Hierarchical Probabilistic Representations of Space" in COGNIRON

Understanding and interpreting the environment with all its complexity is a very hard problem. None of the existing approaches to environment modeling cope with the vast amount of multi-modal information available from the environment. Further, all prior work in this area are done in the context of specific problem such as navigation, interaction etc. Intuitively, it is believed that humans do not have different representations for performing different tasks and that they use a single unified hierarchical representation for modeling the entire environment. As this project aims at developing robots into cognitive companions of human beings, this work package attempts to build a new hierarchical representational methodology for environment modeling, along the lines of what we know about us.

Relation to the Key Experiments

This report specifically addresses the problem of "hierarchical probabilistic representations of space". The work is primarily "situated" in the context of the "Home Tour Scenario" (Key experiment 1).

CONTENTS

Report on methods for Hierarchical Probabilistic Representations of Space

Shrihari Vasudevan (EPFL), Roland Siegwart (EPFL), Ben Kröse (UVA), Bram Bakker (UVA), Michel Devy (LAAS), Henrik I. Christensen (KTH)

1. Introduction and Problem Statement

Interpreting and understanding a scene from the environment beyond single object recognition is a hard task. Humans use various sensory cues to extract crucial information from the environment. This is processed in the cortex of the brain in order to obtain a high-level representation of what has been perceived. Intuitively, it appears that humans represent knowledge in a hierarchical fashion. With a view of having robots as companions of humans, we are motivated towards developing a knowledge representation system along the lines of what we know about us. While recent research has shown interesting results, we are still far from having concepts and algorithms that interpret space coping with the complexity of the environment.

This report basically looks into the development of a new representational methodology, which may be useful for planning, navigation, manipulation and interaction tasks. A multi-resolution framework is suggested as a possible solution. Thus, the primary theme of this report is perception, interpretation and representation on a multi-resolution framework.

2. Related Work

Most of the related research on formalizing levels of abstraction in literature can be found in cognitive science (e.g. hierarchical representation and reasoning with knowledge) and mobile robotics. While some works address specifically one of the two research areas, others contribute towards both as they use cognitive processes as an inspiration to mobile robotics research and do not view them separately. Space characterization and representation are very important from the Cognitive Science point of view, towards the quest of understanding the functioning of the human brain. In the engineering context, space representation is pivotal for fully understanding any system and to realize any form of intelligence. Thus, space representation plays a central role in the development of any cognitive and autonomous intelligent system.

The idea of cognitive maps (i.e. the human internal representation of space) was introduced for the first time by Tolman in [35]. Significant progress has been made since the seminal papers by Kuipers [16 & 17] where the cognitive maps are described as the body of knowledge representing large scale space. In his work, a "spatial semantic hierarchy" (SSH) is suggested which represents space at different levels of abstraction and attempts navigation using such a representation. The representation has 4 layers corresponding to sensorimotor information, view (sensory image) information with actions to represent transitions between views, a topological/place level and a metric level. This representation centers on the topological model for robust navigation. Metric information is used mainly in the context of optimization or disambiguation. The SSH model is explained in detail in [18] and also in [19]. An approach similar to the previous one can be found in [22] where a hierarchical multi-resolution space representation is addressed. Voicu uses landmarks and associations between them to construct a cognitive map of a large environment in [37]. The information from this cognitive map is then used for path planning and exploration. The authors of [30] use a hierarchical hidden Markov model (HHMM) to learn the route between two labs. The higher level states are the more abstract/distinct ones like corners and intersections. The lower level states represent intermediate positions.

Tapus et al have developed the fingerprint concept to a significant extent, in the context of navigation, as shown in [33]. The two most relevant contributions of this work are – (1) the fingerprint provides an efficient way to do place-characterization and (2) it also represents a method for compressing the huge amount of information that is required to represent a place, without loosing critical information about it. These advantages provided us with the inspiration to look to this approach as a means of achieving our objectives.

Tomatis et al in [36] model the environment as a hybrid map having both a global topological representation and several local metric representations (a metric map for each node in the topological map) to facilitate precise localization as and when necessary. The metric representation is made up of innumerable lines representing places. The topological map is a graph of several nodes. Once at a particular node (place), the robot uses the metric map to move to a particular position.

The work [14] demonstrates a hybrid topological-metric representation wherein the map is represented as a bidirectional graph of nodes. Each node itself is composed of a local metric map with a local reference frame associated with it. Edges in the graph represent coordinate transformations between two different reference frames. "FastSLAM" [25] is used to form the local – metric maps. The transformations between coordinate frames are realized by a set of particles (the transitions are represented as distributions which are then sampled to give the particles required).

In the context of task planning for navigation, Galindo et al have implemented efficient planning methods using a hierarchy of abstractions of space. In their work [12], they use "annotated and hierarchical graphs (AH Graphs)" to represent the environment at different levels of abstraction for efficient task planning. In essence, their methodology solves the problem for a highly abstracted view of the environment. This solution is then refined for more detailed representations of the environment, while discarding irrelevant entities at each level. Their work is based on a topological representation of the environment.

All the above mentioned works seem to capture a hierarchical representation of space with a navigational flavor in them. Semantic Networks (SN) provides a motivation to incorporate high-level reasoning to a more navigation specific representation – such as those suggested earlier. A SN represents the world as nodes and directed links denoting relationships between them. The work described in [31] uses this representation to interpret baseball-game video at a high level of abstraction using low level video information and a Bayesian belief network. The approach suggested here treats the problem in a similar manner i.e. using low level information to construct high level representations of the environment.

The problem of space representation draws inspiration from the state-of-art methodologies of human spatial modeling also. The hippocampus is the part of the brain that plays a pivotal role in spatial memory and cognition and its principal neurons are called place-cells [26]. Arleo et al in [4] model the hippocampal place-cell activity during spatial cognition and navigation. Both "allothetic" and "idiothetic" representations of the environment are collected and integrated to form a single stable representation of the environment. Place-cells (first introduced by O'Keefe in [27]) are activated only when the animal involved moves over a specific point in the environment. Using reinforcement learning and the fact that place-cells control neurons which produce locomotion signals, certain navigation tasks are performed and the corresponding place cell activity is modeled.

The work done by Brezetz et al in [6] bears close similarities to the initial steps used to realize the objective of [34]. It assumes the presence of an even ground with objects on it. It then uses range images and does segmentation to extract semantic information from the image. This information is further used to interpret spatial relationships between objects.

In summary, research efforts in spatial representation have been concentrated on navigation with negligible emphasis on adding semantics to such representations so as to deal with problems like interaction and manipulation. We hope to adequately address this issue in our work.

3. Proposed Frameworks / Individual Contributions

(1) **The General Idea**

The objective of this research work-package is to develop a multi-resolution probabilistic representation of space, capable of incorporating uncertain multi-modal information content. The representation must be scalable in terms of complexity and must be useable for navigation, interaction and manipulation for indoor robots. This subsection gives a generalized overview of the concept that is being envisioned. In the following subsections, the various approaches / methodologies that were conceived, towards achieving this objective, are elicited. This is followed by an explanation on how these approaches relate closely to each other.

Figure 1, given below, shows the general approach to the problem. At the lowest level of the hierarchy, multi-modal information (raw data) is obtained from a wide array of sensors. The information content may include laser / sonar data, vision, sound and even smells from the environment.

The raw data, from the sensors, at this level may be used to detect the presence of some higher level features (such as lines, contours and color blobs). This may be done by a process of imposing explicit models (for the features) or learning them. These features may in turn be used to form metric feature maps (geometric maps).

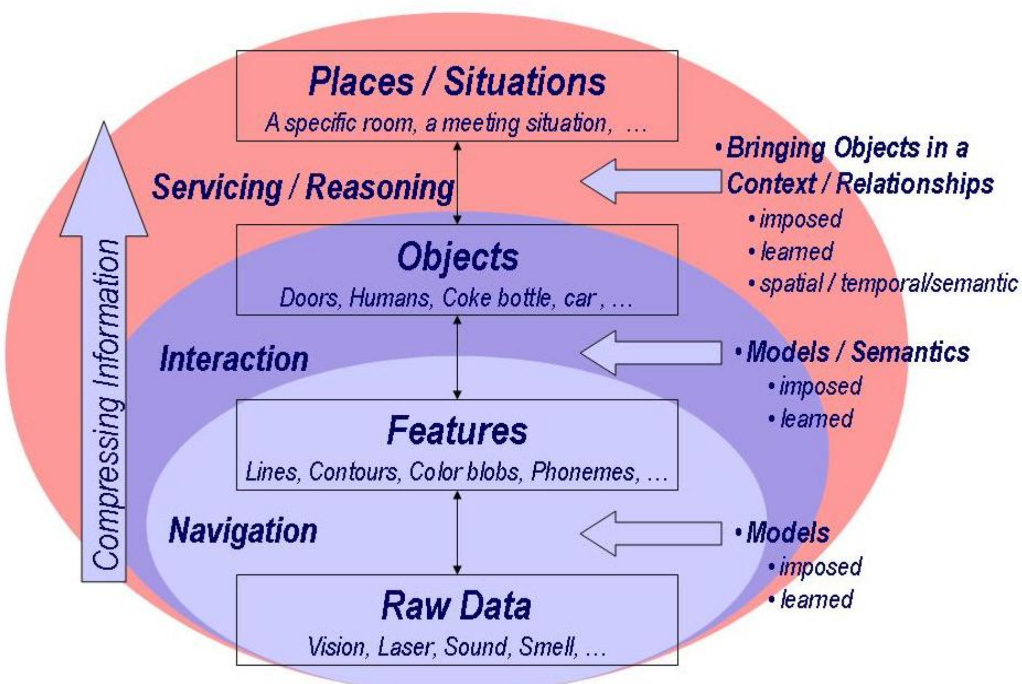


Figure 1: From Geometric to Cognitive Maps – a generalized overview of the concept.

Features can themselves be used to detect the presence of specific objects such as doors, humans, coke bottles and so on. A similar process of imposing or learning object models can be used for this purpose. However, as objects are associated with certain semantics, this may be used to facilitate the object detection. Apart from being the source of semantic knowledge in a representation, they may also serve to drastically improve place recognition.

Objects have associated features and semantics. These can give rise to relationships – of spatial, functional and temporal nature. It is these relationships which enable incorporation of semantics in the representation. The relationships and the associated semantics when incorporated into the representation, yields a model wherein objects are "situated" in a context, Such a modeling of space leads to the introduction of high level, semantic concepts of "places" or "situations".

As we go up the framework, information is compressed in order to get higher levels of abstraction. The "geometric" content of information at the higher levels is minimal, as is the "semantic" content of information at lower levels of abstraction. Both of them increase towards the other end of the hierarchy (with respect to their minimal levels). While navigation and manipulation related tasks require precision and thus tend to make extensive use of the information in the bottom of the hierarchy, reasoning and interaction based tasks require a great deal of semantics and thus will probably be more focused towards the top of the hierarchy. In general however, for executing any task, a good switching mechanism between the levels is warranted, so that it can access all the different types of information through the duration of the process.

(2) **The Fingerprint-OGM approach**

A fingerprint [20] is a circular list of features wherein, the ordering of features matches with the ordering of features around the robot. They are advantageous in that they can encode a large amount of place related information in a single sequence, thus providing a convenient means of place characterization. Further, they are particularly useful in the context of multimodal sensory input. Fingerprints are particularly useful for place recognition, as shown in [20 & 33].

A recent work, [34], suggests the "object graph model" (OGM) as a step towards the development of a multilevel cognitive probabilistic representation of space. Figure 2 shows a typical setup that may be encountered by a robot in the home-tour scenario. Figure 3 shows a diagrammatic representation of the OGM concept for a part of the environment. Objects and relationships between them could drastically increase the distinctiveness of a place. Thus, we would like to embed objects in the fingerprints (which is to say – integrate the OGM concept with the Fingerprint concept).

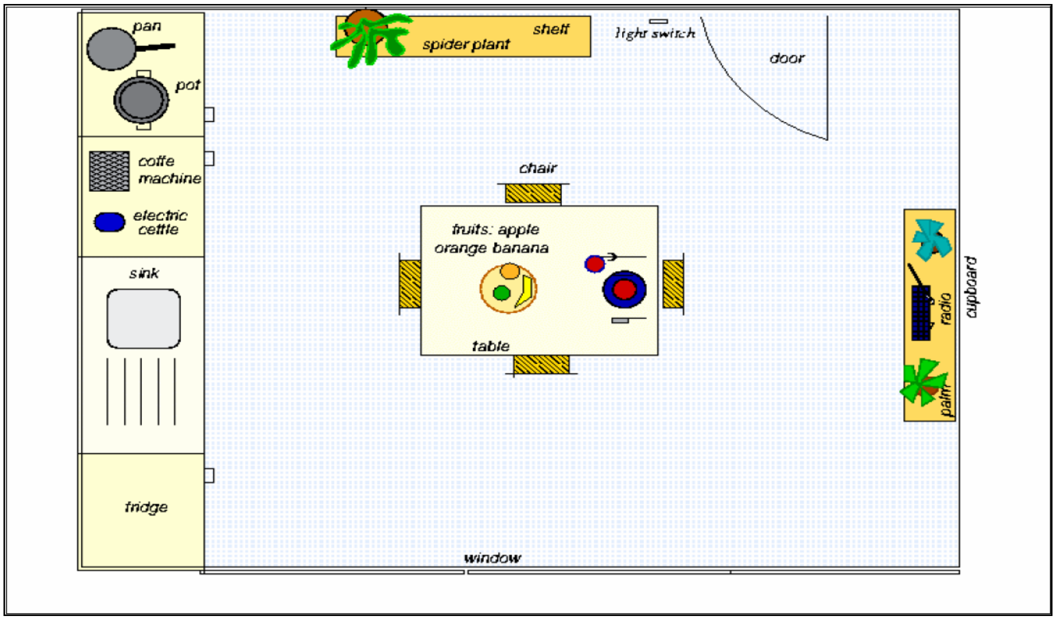


Figure 2: The Key Experiment – Robot Home Tour Scenario setup (taken from the key experiment specifications document)

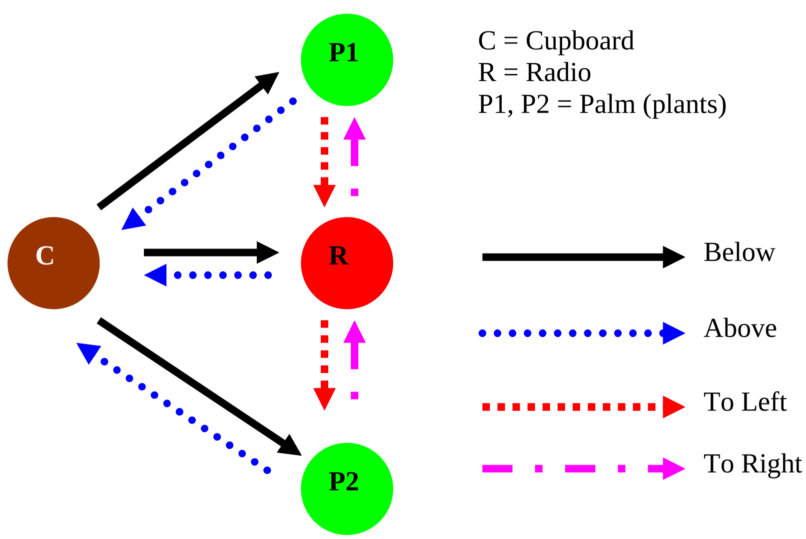


Figure 3: Example OGM representation for the cupboard area of the room.

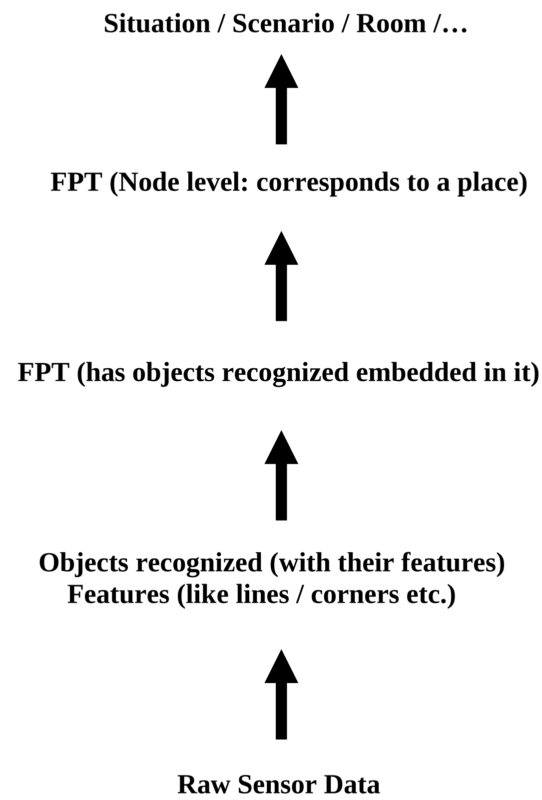


Figure 4: Hierarchical representation using fingerprints (FPT) and the OGM concept

Raw sensor data obtained from the sensors is processed to yield (1) features like lines, corners, color-blobs etc. and (2) objects, with their properties. Objects and their properties give rise to relationships (which may be spatial / functional or temporal) and thus encode the semantics of the environment. These features, objects and their interrelationships are collectively pooled into a single structure – the fingerprint, which basically denotes the "signature" of a location. Signatures of successive locations are "aggregated" to give rise to a node (corresponding to a place). A node is associated with an aggregate fingerprint of a set of "similar" fingerprints. A set of nodes may be used to define a region / room / situation / scenario / some such high level concept. To have precise localization at places of interest, a local geometric map will also be required. This may be constructed using methods such as those demonstrated in [2, 9, 11, 13, 22, 24, 36 & 38]. Figure 4 illustrates the hierarchical representation of space, using the fingerprint-OGM concept.

(3) **The "appearance-based" approach**

The central idea of this approach is that it makes sense to represent large environments at different resolutions or levels of abstraction simultaneously. Figure 5 shows the idea of a schematic representation of an environment at different levels of abstraction.

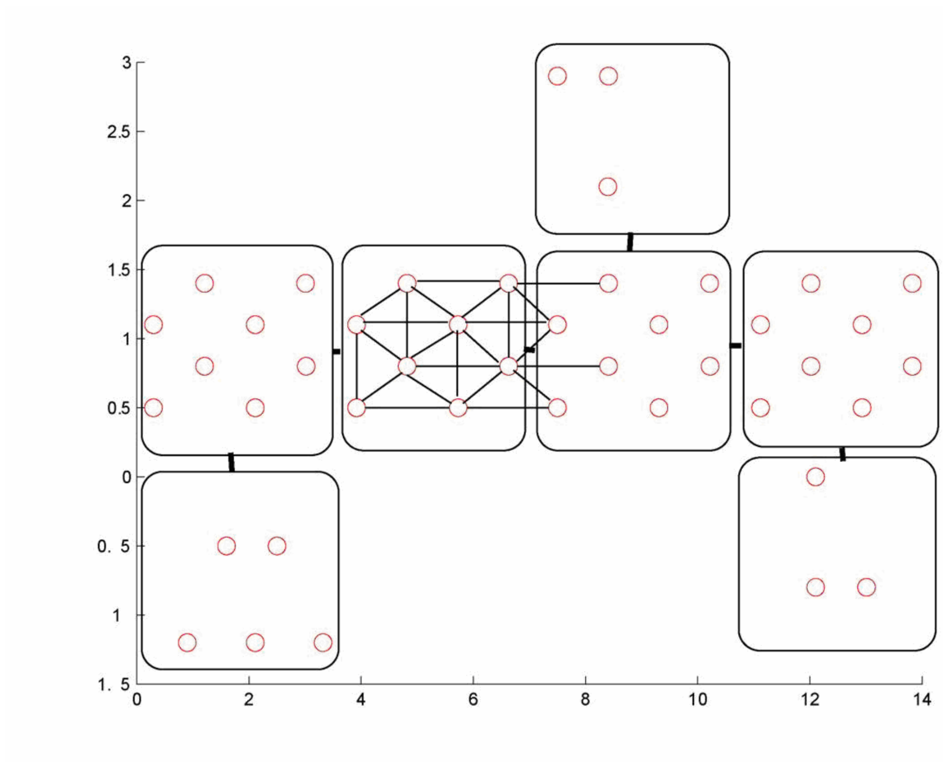


Figure 5: Schematic representation of an environment at different levels of abstraction. It comprises of an abstract, high-level topological map whose nodes represent low level maps, which are also topological maps here.

Low-level, local, detailed maps may be used to represent, for example, individual rooms in a large building in great spatial detail, without having to represent the spatial relationships to locations in other rooms. Such a low-level map may be used for navigation to precise target locations within an individual room, and to move to neighboring rooms, without having to worry about exact locations in other rooms. Higher-level, global, abstract, "conceptual" maps may be used to represent the entire building, for instance as a graph connecting rooms and corridors, without representing the exact spatial relationship of individual locations within rooms and corridors. Such a high-level map may be used to construct abstract plans to navigate from one room to another, without having to worry about exact spatial details. A hierarchy of maps may thus facilitate map building as well as subsequent planning based on the maps.

Another advantage of a hierarchy of maps is that it can facilitate the interaction of the robot with humans, because the elements in the higher-level map (e.g., the nodes in the graph) can be made to correspond to concepts that make sense to humans (rooms, corridors), instead of metric (x, y) coordinates that are not intrinsically meaningful to humans in office and home environments.

Thus, the human could instruct the robot to go the "kitchen" and because there is a node corresponding to the kitchen in the higher-level map, the robot would know where to go. After having constructed its high-level plan, the robot could even explain to the human that to go to the kitchen, it must first traverse the living room, then pass through the corridor, and finally enter the kitchen—because kitchen, living room, and corridor could all be nodes in the map. The hierarchy of maps provides a connection between this high-level plan and the actual low-level execution, as it allows the robot to plan the low-level details of the task through its lower-level maps.

This approach focuses not so much on the facilitation of robot-human interaction through hierarchical mapping, but instead on how planning for navigation can be much more efficient when hierarchical maps are used. In particular, the maps are interpreted as Markov Decision Processes (MDP's), and the path planning task as a dynamic programming problem ([5], [7] & [8]). Dynamic programming is attractive because it allows the system to efficiently plan optimal shortest-path policies for the entire state space, it can deal reliably with all kinds of noise in the execution of actions (i.e. stochastic state transitions), and it allows straightforward inclusion of cost factors other than distance traveled, such as energy consumption and obstacle avoidance. This work provides a demonstration of how the hierarchical approach leads to significant savings, in terms of the number of value function updates until convergence, when compared to using just one large, flat, low-level map. This advantage becomes more pronounced as paths must be planned to many possible target locations. This is particularly relevant with very large MDPs, i.e. the type of large, realistic environments in which domestic and office robots would eventually have to live. This computational advantage comes at the cost of some extra overhead to represent and coordinate the hierarchical system, and in some cases slightly longer paths to target locations.

The robot is equipped with an omni-directional camera. The method assumes a database of panoramic images which more or less cover the entire environment without any additional pose information. This database forms the basis for the "appearance-based" topological map. Given the database of images, from each image a set of distinctive local image features, SIFT features ([21]), are extracted (see figure 6). Based on matches between features in different images (see figure 7), the hierarchy of topological maps / MDP's is determined.

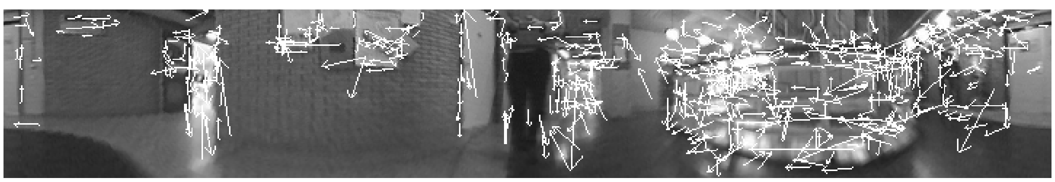


Figure 6: A panoramic image, annotated with positions of extracted SIFT features. Arrows indicate the scale and main orientation of the feature.

Figure 7: A pair of images from nearby locations in the environment (one of them shown in figure 6), annotated with lines indicating close matches between the descriptors in each of the images.

(4) **Alternative "schema" of the generalized hierarchical representation of space**

The following section looks at another very similar, generalized hierarchical representation of space. The overall concept is similar to both the previous approaches, but the schema of the representation (the hierarchical structure) has been envisioned in a slightly different fashion (illustrated in Figure 8).

The approach focuses on the acquisition of cognitive knowledge taking several different forms including geometrical, topological and semantic representations. Towards the construction of these representations, different forms of contextual knowledge are to be used – these include common properties like walls, the ceiling, predefined areas such as the corridor/kitchen etc. and predefined models of generic object classes. The knowledge base is to be general in that no specific configurations of areas or objects are to be used.

The built model is expected to be hierarchical having both geometrical and appearance based representations. The architecture of the envisioned representation is given below:

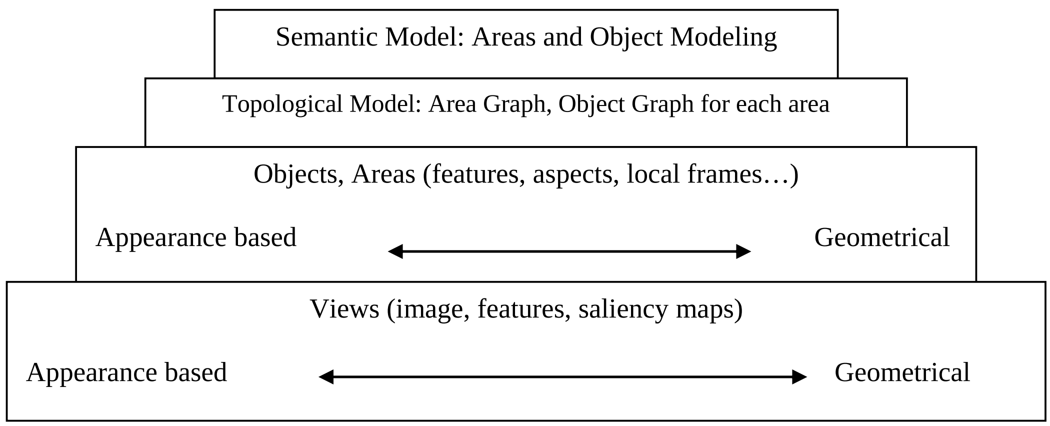


Figure 8: Alternative view of the Hierarchical representation of space

Appearance based representations may be view centered, area centered or object centered. View centered models record the sensor position, orientation and the image features that have been extracted, for each view. Area centered models represent the local reference frame, a list of objects that have been identified and descriptions of the area concerned, in terms of features like doors, walls, the ceiling and so on. Object centered descriptions of the environment detail the extracted image features, the aspect table, the convex hull and so on. Geometric representations include the classical stochastic map. Several independent maps on every area are linked together. Belief management through a Bayesian belief network scheme is required to propagate beliefs from low level features to high level labeling of an object or an area.

The cognitive knowledge is to be acquired online using four asynchronous visual processes. These processes include the acquisition and processing of panoramic images, focalized image acquisition on regions of interest, the reconstruction of 3D images and finally, a background process that interprets all this information on a global level and supervises the robots actions. Following image acquisition and processing, the images are time stamped and the view-centered representations are updated. Some of these steps will be achieved through an interaction with a user or some simple heuristics.

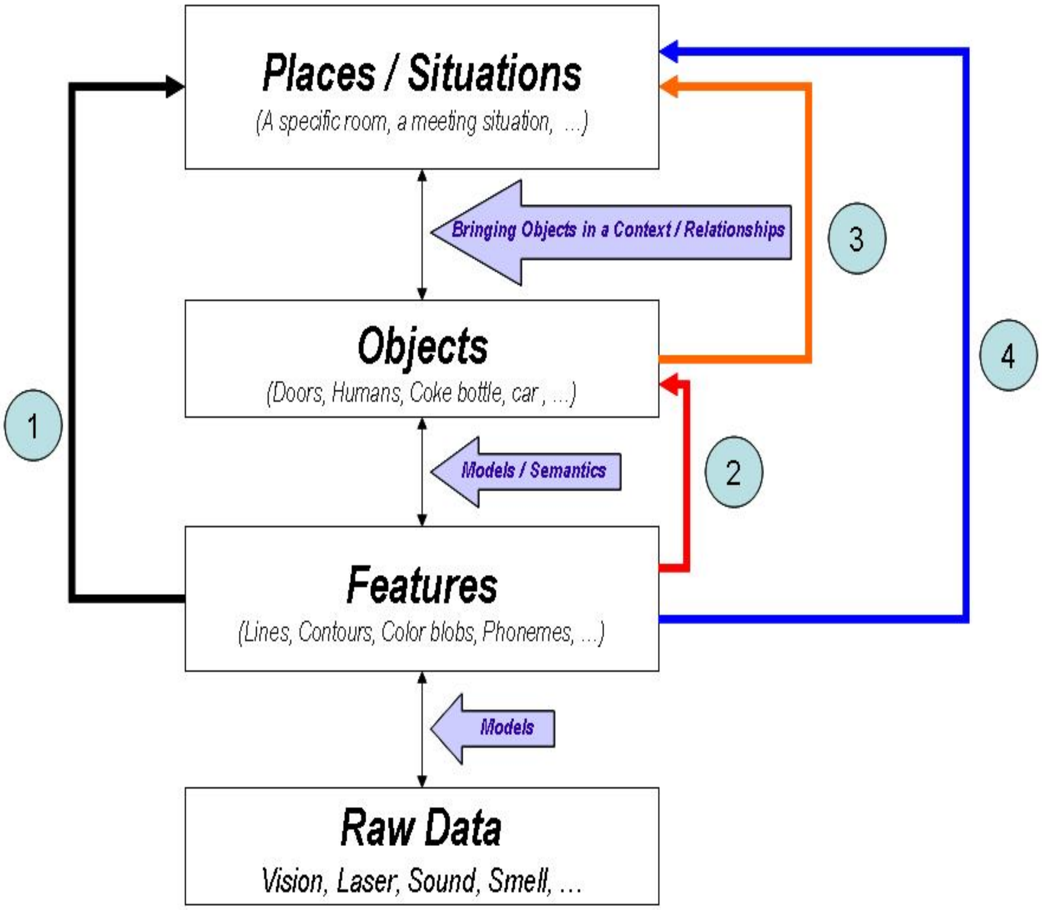
4. How do the approaches relate with one another?

The current state of research in this work package is witnessing a consensus on the overall strategy. Different partners are approaching the problem from different directions of the same overall schema of the representation. Each partner has a different "hierarchical" view of the problem. However, the general concept remains identical and is explained below.

From all the approaches reported above, the following common line of thought can be clearly seen. The representation is a probabilistic hierarchical structure at different resolutions of information abstraction. As the information becomes more abstract (we go up the hierarchy), it becomes more "compressed" and the "geometric" content in it reduces. Information becomes more semantic in nature, at the higher levels. Conversely, as the information content becomes more detailed (we go down the hierarchy), the "geometric" content in it increases dramatically.

While navigation and manipulation will require more detailed/precise information and thus will make extensive use of the lower levels of abstraction, tasks involving reasoning and interaction will probably require greater semantics and thus will focus more on the upper levels of the hierarchy. In general however, any task would require a good system to switch between the levels of the representation in order to use the most appropriate type of information it may require through its duration.

To understand how each partner is currently approaching the work-package, consider an adaptation of the general representation scheme suggested earlier, shown in figure 9. It shows the approach being followed by each partner in this research area. It also serves to underline the fact that all the approaches finally integrate into a single hierarchical probabilistic representation of space that is suited for navigation, interaction and manipulation – the central aim of this research area.



1 – Appearance based topological model (UVA / EPFL) (navigation)
2 – Object modeling (LAAS) (manipulation / interaction)
3 – Object Graph Models (EPFL) (interaction / manipulation)
4 – Geometric Maps (2D / 3D) (LAAS / EPFL / KTH) (navigation)

Figure 9: Adapted version of a generalized hierarchical representation methodology to show the approach of each partner

Lower level geometric representations (metric maps) are a common strength of each partner [1, 2, 9, 10, 11, 13, 23, 24, 38 & 39]. EPFL has also built its expertise in topological mapping [22, 33 & 36]. Work has already started on integrating geometrical information at the *object* level with the *place/situation* level through topological maps. UVA is putting its thrust on the appearance based approach [15, 28 & 29] and its current efforts establish the link between the *feature* and the *place / situation* levels through appearance based topological modeling. LAAS is concentrating its efforts on object modeling [32]. They form the link between the *feature* and *object* levels of the hierarchy. KTH, with its current focus on human augmented mapping [3] links the lower level metric representations (*feature* level) with higher level semantic models (*place/situation* level).

5. Conclusions and Outlook

The objective of this work-package is to develop a multi-resolution probabilistic representation of space that can accommodate uncertain multi-modal data. The said representation must be capable of handling tasks such as navigation, interaction, manipulation and other reasoning related tasks.

All involved partners have the same overall methodology towards approaching the problem. The representation will be hierarchical and will contain information at different levels of abstraction. Lower levels, with more detailed (geometric) information will prove to be useful in the context of navigation and manipulation. Higher levels of the representation will contain more information relevant to reasoning and interaction. However, each of these tasks may require information from multiple levels of the representation and thus an efficient connection between individual levels is warranted.

Current efforts have resulted in a consensus as far as the overall schema of the representation is concerned and each partner has made progress in approaching the problem. Future work will involve integrating “complimentary” efforts towards realizing the overall objective.

6. References

[1] Albalate, M. T. L., Devy, M., Marti, J.M.S., (2002), Perception planning for an exploration task of a 3D environment, International Conference on Pattern Recognition (ICPR'2002), Québec, Canada, 11-15 August 2002.

[2] Altermatt, M., Martinelli, A., Tomatis, N. and Siegwart, R. (2004) SLAM with Corner Features based on a Relative Map. In the proceedings of the IEEE/RSJ International Conference on intelligent Robots and Systems (IROS04), Sendai, Japan, September 2004.

[3] Althaus, P. and Christensen, H. I., (2003), Automatic Map Acquisition for Navigation in Domestic Environments, In the proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2003), Taipei, Taiwan. September 2003.

[4] Arleo A. and Gerstner W. , (2000), Spatial Cognition and Neuro-Mimetic Navigation: A Model of Hippocampal Place Cell Activity, Biological Cybernetics, 83:287-299.

[5] Bertsekas, D.P., (1995), Dynamic programming and optimal control, Athena Scientific Publishing company, Belmont, MA, USA.

[6] Brezetz, S. B., Chatila R., Devy M., (1994), Natural scene understanding for mobile robot navigation, In the proceedings of the IEEE International conference on Robotics and Automation, San Diego, USA.

[7] Briggs, A. J., Detweiler, C., Scharstein, D. and Vandenberg-Rodes, A., (2004), Expected Shortest Paths for Landmark-Based Robot Navigation, International Journal of Robotics Research, vol. 23 (7-8), pp. (717-728).

[8] Buhmann, J. M., Burgard, W., Cremers, A.B., Fox, D., Hofmann, T., Schneider, F.E., Strikos, J. and Thrun, S.,(1995), The Mobile Robot (RHINO), AI Magazine, vol. 16, no. 2, pp. 31-38.

[9] Bulata, H. and Devy, M., (1996), Incremental construction of a landmark-based and topological model of indoor environments by a mobile robot, In the proceedings of the IEEE International Conference on Robotics and Automation (ICRA'96), Minneapolis (USA), 22-28 April 1996, pp.1054-1060.

[10] Folkesson, J. and Christensen, H. I., (2004), Graphical SLAM – A self correcting map, In the proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2004), New Orleans, USA. April 2004.

[11] Folkesson, J. and Christensen, H. I., (2004), Robust SLAM, In the proceedings of the IFAC Symposium on Intelligent Autonomous Vehicles (IAV 2004), Lisbon, Portugal.

[12] Galindo, C., Madrigal, J.A.F. and Gonzalez, J. (2004), Improving Efficiency in mobile robot task planning through world abstraction, IEEE Transactions on Robotics, vol. 20, no. 4, August 2004.

[13] Hayet, J.B., Lerasle, F. and Devy, M. (2003), Visual landmarks detection and recognition for mobile robot navigation, in the proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2003), Madison (USA), 18-20 June 2003, vol. II, pp.313-318

[14] Kouzoubov, K., Austin, D. (2004) Hybrid topological/metric approach to SLAM, In the proceedings of the International Conference on Robotics and Automation (ICRA 2004), New Orleans, LA, USA

[15] Kröse, B.J.A., Vlassis, N., Bunschoten, R. and Motomura Y., (2001), A probabilistic model for appearance-based robot localization. Image and Vision Computing, 19(6):381-391, April 2001.

[16] Kuipers, B. J. (1978), Modeling Spatial Knowledge, Cognitive Science, 2: 129-153, 1978.

[17] Kuipers, B. J. (1983), The Cognitive Map: Could it have been any other way? In Spatial Orientation: Theory, Research and Application. Picks H.L. and Acredolo L.P. (eds.), New York. Plenum Press.

[18] Kuipers, B. J. (1996), A Hierarchy of qualitative representations for space, In the proceedings of the 10th International Workshop on Qualitative Reasoning (QR-96), Fallen Leaf Lake, California, USA.

[19] Kuipers, B. (2000) The Spatial Semantic Hierarchy, Artificial Intelligence 119: 191-233.

[20] Lamon, P., Tapus, A., Glauser, E., Tomatis, N. and Siegwart, R. (2003) Environmental Modeling with Fingerprint Sequences for Topological Global Localization. In the proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'2003), Las Vegas, U.S.A., October 2003, vol.3, 3781 - 3786.

[21] Lowe, D.G.,(2004), Distinctive image features from scale-invariant key-points, International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.

[22] Martinelli, A, Tapus A, Arras, K.O., and Siegwart. R. (2003), Multi-resolution SLAM for Real World Navigation, In the proceedings of the 11th International Symposium on Research Robotics, Siena, Italy.

[23] Martinelli, A. and Siegwart, R. (2004) Improving the SLAM Convergence with a Relative Map Filter. In the proceedings of the International Conference on Intelligent Autonomous Systems (IAS04), Amsterdam, March 2004.

[24] Martinelli, A., Svensson, A., Tomatis, N. and Siegwart, R. (2004) SLAM Based on Quantities Invariant of the Robot's Configuration. IFAC Symposyum on Inteligent Autonomous Vehicles (IAV04), Lisbon Portugal, July 2004.

[25] Montemerlo, M., Thrun, S., Koller D. and Wegbreit, B., (2002), Fastslam: A factored solution to the simultaneous localization and mapping problem, In the proceedings of the AAAI-2002, Vancouver, BC, July 2002.

[26] Moser E. I., Paulsen O., (2001), New excitement in cognitive space: between place cells and spatial memory, In Current Opinion in Neurobiology 2001, 11:745–751.

[27] O'Keefe J. and Dostrovsky J., (1971), The Hippocampus as a Spatial Map: preliminary evidence from unit activity in the freely moving rat. Brain Res 34: 171-175.

[28] Porta, J. M., Verbeek, J.J. and Kröse, B.J.A., (2003), Enhancing appearance-based robot localization using sparse disparity maps, In the proceedings of the IEEE International Conference on Intelligent Robots and Systems, pages 980-985, Las Vegas, USA, October 2003

[29] Porta, J. M. and Kröse, B.J.A., (2004), Appearance-based concurrent map building and localization, International Conference on Intelligent Autonomous Systems, IAS'04, pages 1022 - 1029. IOS Press, March 2004.

[30] Rohanimanesh K., Theocharous G. and Mahadevan S. (2000), Hierarchical Map learning for robot Navigation, In AIPS Workshop on Decision-Theoretic Planning, Breckenridge, Colorado.

[31] Shih H.C. and Huang C.L. (2003), A Semantic Network modeling for understanding baseball video, In the proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Hong Kong, China.

[32] Specht, A.R. and Devy, M., (2004), Surface-segmenting modified Ball-Pivoting-Algorithm, International Conference on Image Processing (ICIP'2004), Singapore, 24-27 October 2004.

[33] Tapus A., Tomatis N. and Siegwart R., (2004), Topological Global Localization and Mapping with Fingerprints and Uncertainty, In the proceedings of the International Symposium on Experimental Robotics, Singapore, June 2004.

[34] Tapus, A., Vasudevan, S. and Siegwart, R. (2005) Toward a Multilevel Cognitive Probabilistic Representation of Space. In the proceedings of the International Conference on Human Vision and Electronic Imaging X, part of the IS&T/SPIE Symposium on Electronic Imaging 2005.

[35] Tolman, E. C. (1948), Cognitive maps in rats and men, Psychological Review, 55:189-208.

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

[37] Voicu H. (2003), Building and Using a Hierarchical Representation of Space, In the proceedings of the International Joint Conference on Neural Networks (IJCNN), Portland, Oregon, USA.

[38] Wijk, O. and Christensen, H. I. , (2000), Localisation and navigation of a mobile robot using natural landmarks extracted from sonar data, Robotics and Automation Systems, vol. 31, no. 1-2, pp. 31-42, April 2000

[39] Wijk, O. and Christensen, H. I., (2000), Triangulation-based fusion of sonar data with application in robot pose tracking, IEEE Transactions on Robotics and Automation, vol. 16, no. 6, pp. 740-752, December 2000.