Proc. of International Joint Conference on Artificial Intelligence (IJCAI-05)
Edinburgh, UK, August 2005. In press

Maintaining Coherent Perceptual Information Using Anchoring

Amy Loutfi, Silvia Coradeschi and Alessandro Saffiotti
Center for Applied Autonomous Sensor Systems
Örebro University
Örebro, Sweden 701-82
www.aass.oru.se

Abstract

The purpose of this paper is to address the problem of maintaining coherent perceptual information in a mobile robotic system working over extended periods of time, interacting with a user and using multiple sensing modalities to gather information about the environment and specific objects. We present a system which is able to use spatial and olfactory sensors to patrol a corridor and execute user requested tasks. To cope with perceptual maintenance we present an extension of the anchoring framework capable of maintaining the correspondence between sensor data and the symbolic descriptions referring to objects. It is also capable of tracking and acquiring information from observations derived from sensor-data as well as information from a priori symbolic concepts. The general system is described and an experimental validation on a mobile robot is presented.

1 Introduction

Consider a scenario where a mobile robot is patrolling a corridor. Its task is to discover new objects and gather information about them using both traditional modalities such as vision and sonar and also non-traditional modalities such as an electronic nose. A user is able to monitor the robot, suggest actions to improve perceptual performance and request the robot to perform specific tasks. Tasks are requested from the user using a symbolic representation consisting of natural language concepts. While no requests by the user are given, the robot autonomously prioritises tasks alternating between patrolling the corridor and inspecting objects or simply waiting on stand-by.

An important facet to this scenario, and scenarios of a similar nature, is the ability to maintain coherent perceptual information. This means that the system should be able to maintain the correspondence between the symbolic representation of objects (requests from the user) and the perceptual data that refers to them. The system should collect information from different sensing modalities working concurrently and/or sequentially and correctly attribute that information to its internal representation of the object. The system should include the preservation of object consistency, so that new information about previously seen objects is also correctly attributed. All the while, the insertion and/or removal of objects from the environment needs to be considered and accounted for. From a cognitive perspective, the maintenance of perceptual information is an integral part of the binding problem [Blackmore, 2003]. That is, how the conjunction of properties are represented, ranging from the binding of shape and colour in detecting blue triangles or red squares to the binding that must occur between and within the senses: "the way the smell and touch and sight of the sandwich in your hand all seem to belong to the same object" [Blackmore, 2003] (p.250).

For robotic systems, an important ingredient for maintaining perceptual information is an internal structure to store a representation of an object. Such an internal structure needs to satisfy a series of requirements. On one hand, perception management [Ronnie *et al.*, 2003], an extension of sensor management [Adrian, 1993], is required to integrate high-level information about the state of the object in order to make situation dependent decisions, direct control and select sensing actions. On the other hand, tracking and data association are also fundamental ingredients necessary to propagate information about coherent perceptions over time [D.Schulz *et al.*, 2003]. A third requirement, especially in the context of a cognitive robot, involves processes, which can create and maintain the link between high-level information (e.g. symbols) and low-level percepts.

In this paper, we show how the concept of anchoring can be used as a tool to confront some of the issues relating to perception management. The anchoring framework [Coradeschi and Saffiotti, 2000] aims at defining a theoretical basis for grounding symbols to percepts originating from physical objects. We present an extension of the framework that is able to cope with perception management considering multi-sensing resources and temporal factors. The pivot of this extension is the use of anchors as internal representations of objects that integrate symbolic and perceptual information across time and across sensing modalities. Our ultimate aim is to shed light on this important aspect of cognitive robotics as applications extend into realistic environments involving human-robot interactions and lifelong acquisition of knowledge.

We begin our discussions with a description and review of the anchoring framework in Section 2. We then present a modification of the anchoring framework in Section 3 which is adapted for the task of perception maintanance. We proceed by detailing the robotic system architecture used in our experiments. Section 5 gives the implementation details and provides a performance example of the complete system. The paper concludes with a summary of the presented work.

2 The Anchoring Framework

To date, the anchoring framework presented by [Coradeschi and Saffiotti, 2000] has mainly been considered in the context of abstraction of perceptual information. This includes association of words to visually recognised objects [Knoblauch *et al.*, 2004], managing of dynamic object anchoring for high-level reasoning [Chella *et al.*, 2004] and preliminary works considering people tracking applications [Kleinehagenbrock *et al.*, 2002]. A good overview of the range of applications of the anchoring concept is found in 2004 Robotics and Autonomous Systems special issue on anchoring symbols to sensor data [Coradeschi and Saffiotti, 2003].

Before we present our extension of the framework, we summarise here the basic elements of the computational theory of anchoring. See [Coradeschi and Saffiotti, 2000] for a full account. The theory considers an autonomous system that includes a symbol system and a perceptual system, and it focuses on the problem of creating and maintaining a correspondence between symbols and percepts that refer to the same physical object. The main ingredients of anchoring are the following [Coradeschi and Saffiotti, 2000]:

- A *symbol system* including: a set $\mathcal{X} = \{x_1, x_2, \ldots\}$ of individual symbols (variables and constants); a set $\mathcal{P} = \{p_1, p_2, \ldots\}$ of predicate symbols; and an inference mechanism whose details are not relevant here.
- A *perceptual system* including: a set $\Pi = \{\pi_1, \pi_2, \ldots\}$ of percepts; a set $\Phi = \{\phi_1, \phi_2, \ldots\}$ of attributes; and perceptual routines whose details are not relevant here. A percept is a structured collection of measurements assumed to originate from the same physical object; an attribute ϕ_i is a measurable property of percepts, with values in the domain D_i.

The symbol system manipulates individual symbols, like 'cup-21', which are meant to denote physical objects and associates each individual symbol with a set of symbolic predicates, like 'red', that assert properties of the corresponding object. The perceptual system generates percepts and associates each percept with the observed values of a set of measurable attributes. The task of anchoring is to create, and maintain in time, the correspondence between individual symbols and percepts that refer to the same physical objects.

The symbol-percept correspondence is reified in an internal data structure α, called an *anchor*. Since new percepts are generated continuously within the perceptual system, this correspondence is indexed by time. It is important that the connections are dynamic, since the same symbol may be connected to new percepts every time a new observation of the corresponding object is acquired.

At every moment t, $\alpha(t)$ contains: a symbol, meant to denote an object; a percept, generated by observing that object; and a signature, a collection of property values meant to provide the (best) estimate of the values of the observable properties of the object.

3 Anchoring for Perception Management

In this section, we present our extension to the anchoring framework. What makes this contribution unique is that we extend the applications of anchoring beyond its traditional use of data abstraction to also include percepts from different modalities that may be accessible at different times.

To effectively present our extension, we refer to the three abstract functionalities defined by [Coradeschi and Saffiotti, 2000]. used to manage anchors, namely, **Find**, **Reacquire**, and **Track**. The functionalities have been developed for the consideration of top-down approaches for information acquisition (i.e. imposed a priori symbolic concepts). In this paper, we revise these existing functionalities to include bottom-up approaches so that anchors can be created by perceptual observations derived from interactions with the environment. Bottom-up approaches have previously been considered on anchoring frameworks [Knoblauch *et al.*, 2004] but never in conjunction with top-down approaches. We advocate the presence of both approaches, in particular for robotic systems interacting with a human user. To accomplish this, an additional **Acquire** functionality is introduced.

3.1 Creation of Anchors

The creation of anchors can occur in both a top-down and bottom-up fashion. Bottom-up acquisition is driven by an event originating from a sensing resource (e.g. the recognition of a segmented region in an image) when perceptual information which cannot be associated to any existing anchor is perceived. Top-down acquisition occurs when a symbol needs to be anchored to a percept, such a call may originate from an external user or a top-level module (e.g. planner).

Acquire Initiates a new anchor whenever a percept is received which currently does not match any existing anchor. It takes a percept π, and return an anchor α defined at t and undefined elsewhere. To make this problem tractable, a priori information is given with regards to which percepts to consider. In bottom-up acquisition, a randomly generated symbol is attributed to the anchor. Furthermore, information about the object and its properties are included into the world model used by the planner, in this way the object can be reasoned about and acted upon.

Find Takes a symbol x and a symbolic description and returns an anchor α defined at t (and possibly undefined elsewhere). It checks if existing anchors that have already been created by the **Acquire** satisfy the symbolic description, and in that case, it selects one. Otherwise, it performs a similar check with existing percepts (in case, the description does not satisfy the constraint of percepts considered by the **Acquire**). If a matching percept is found an anchor is created. Matching of anchor or percept can be either partial or complete. It is partial if all the observed properties in the percept or anchor match

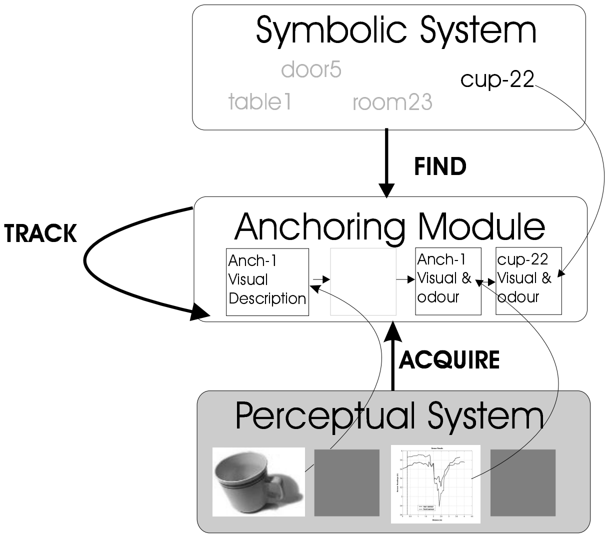

Figure 1: Graphical illustration of the extended anchoring functionalities where bottom-up and top-down information is possible and different sensing modalities are used.

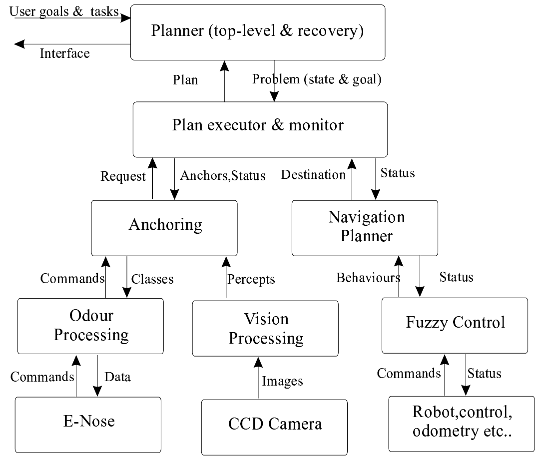

Figure 2: Overview of the robotic system which uses the anchoring module. Arrows indicate the flow of information.

the description, but there are some properties in the description that have not been observed.

3.2 Maintenance of Anchors

At each perceptual cycle, when new perceptual information is received, it is important to determine if the new perceptual information should be associated to existing anchors. The following functionality addresses the problem of tracking objects over time. In this extension, we include the previous **Reacquire** functionality as an integral part of the **Track** and make no special distinction for it.

Track The track functionality takes an anchor α defined for $t-k$ and extends its definition to t. The track assures that the percept pointed to by the anchor is the most recent and adequate perceptual representation of the object. We consider that the signatures can be updated as well as replaced but by preserving the anchor structure we affirm the persistence of the object so that it can be used even when the object is out of view. This facilitates the maintenance of information while the robot is moving as well as maintaining a longer term and stable representation of the world on a symbolic level without catering to perceptual glitches.

3.3 Deletion of Anchors

By having an anchor structure maintained over time, it is possible to preserve the perceptual information even if the object is not currently perceived (caused by the object being out of view and/or by the inaccuracy in the measurement of perceptual data). The challenge is to determine if the association of new percepts is justified or whether certain anchors should be removed. Mechanisms for destroying anchors when the corresponding object has been removed need to be in place. This is a difficult problem, because conceptually it is not clear when it is appropriate to remove anchors from the system. Anchors could be removed if they are not relevant for the current task, because the object to which it refers has been physically removed from the environment or the reliability of the perceptual information has expired. Anchors may also need to be removed if they have been associated to invalid perceptual data such as sensory glitches. We currently adopt simple solutions in which objects that are not perceived when expected decrease in a "life" value of the respective anchor. When the anchor has no remaining life, the anchor is removed. The converse could be implemented where anchors are created with initially lowlife values and persistent percepts increase its life value. The decreasing life of anchors is shown in Figure 4. A more adequate strategy to handle the maintenance of anchors may also be to include a "long term" memory where anchors may be stored for future use.

3.4 Integration of the Functionalities

The event-based functionalities are now restricted to the **Find** and **Acquire** while the **Track** functionality is regularly called. Figure 1 shows an overview and an example of the framework and its functionalities. In the example, anchors are created bottom-up from the visual percepts of a cup. Later, additional features of that object are required, for example, the olfactory property. These features are stored in the anchor. When a top-down request is sent to the anchor module to find a cup with matching properties denoted by the symbol "cup-22", the **Find** functionality anchors the symbol to the perceptual data.

As seen in the figure, properties can be collected at different time points using different modalities. Even when certain perceptual properties are updated, such as the smell property, which may change over time, other perceptual properties are maintained. Conversely, if the visual percepts of an anchor is replaced, the smell property previously obtained is not lost. In this way, the anchor is used to compensate for any dynamically changing features of an object. Furthermore, the perceptual description of anchors can be accessed by the planner

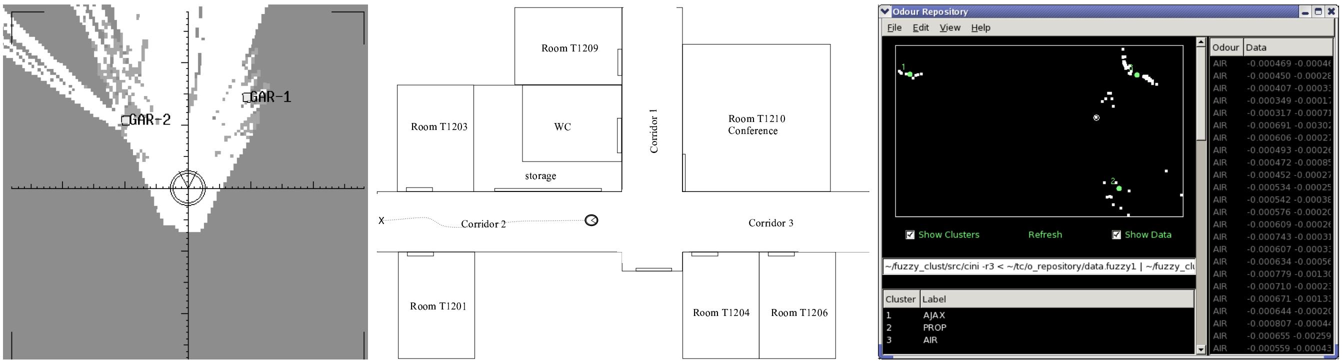

Figure 3: *(Left)* The local space shows the detected objects, the robot is located in the center of the space. Grey areas in the space are regions that are unexplored. Objects are represented by their visual percepts and are placed within the local space. In this screen, the robot identifies 2 garbage cans, denoted by the percepts *Gar-1* and *Gar-2*. *(Middle)* The a priori map of the office environment to be patrolled. The robot starting position is shown at point X and a path of the robot is denoted by the dotted line. *(Right)* The olfactory interface.

to reason about perceptual knowledge. In certain cases, this may result in specific calls to perceptual actions in order to disambiguate between similar objects.

4 The System Architecture

In this section, we present our own instantiation of the extended anchoring framework discussed. We begin our description at the sensor level and proceed to the higher levels which include a planner and user interfaces. An overview of the robotic system is given in Figure 2.

4.1 Sensing modalities and control

Our physical robot is a Magellan Pro compact robot and in the experiments we use its infrared sensors, sonars and tactile sensors. In addition, a CCD camera is mounted on the robot and the robot is able to recognise pre-defined signatures of objects using standard visual techniques. Perhaps the most novel of sensors on the robot is an electronic nose. The electronic nose consists of 32 conducting black polymer sensors, pattern recognition and classification components, and an electronic repository of odours stored in the on-board computer in the robot. The classification algorithm uses the odour repository as a training set for online recognition of new odours. To navigate the robot, a collection of basic behaviours is used. These behaviours are based on fuzzy control techniques and can be combined and reasoned about through use of a behaviour planner (B-plan) explained in [Saffiotti *et al.*, 1995].

4.2 Anchoring Module

The anchoring module, besides creating and maintaining the anchor data structure, also serves a secondary purpose to function as a flag between top-level tasks given by a planner and low-level sensor data. Although not immediately evident this function is important in order to co-ordinate perceptual processes such as a smelling action which may take up to several minutes. In such a case, specific calls to perceptual actions e.g. "Smell gar-22" are generated from the top-level planner, which then translates into several behaviours being activated e.g. "Go near to Gar-22, Touch Gar-2" and calls to an odour server which activates the respective pumps on valves on the nose. The odour classification provides the smell description e.g. "gar-22 smells ethanol" and the perceptual information is updated. Here the anchoring module "polices" each event signalling to the respective modules when certain process need to be activated and when they have reached completion.

In our instance of the anchoring module, the track functionality is achieved by performing a fuzzy matching algorithm between newly created percepts and previously stored anchors in order to partly deal with sensor noise. However, the purpose of the anchoring framework allows different strategies for the track functionality. A part of the future aims to integrate more advanced solutions into the existing platform.

4.3 Planner

PTLplanner [Karlsson, 2001] is a planner for partially observable domains with uncertainty (probabilities). It searches in a space of epistemic states, or e-states for short, where an e-state represents the agent's incomplete and uncertain knowledge about the world at some point in time. Although much of the planning component is standard, it is still worth emphasising that the planner can reason about perceptive actions, such as looking at or smelling an object. The consequence is that calls to perceptual actions may be made in order to gather more information about the environment.

4.4 Interfaces

In addition to the computations mentioned in the previous section for control, perception and autonomy, the system also has a number of processes for displaying the internal state of the robot as well as its current model of the external world. This model of the external world includes locally perceived objects and a gridmap of the environment built from sensor data. This is shown in Figure 3. In the left window, the local view of the robot shows the incoming percepts of the vision

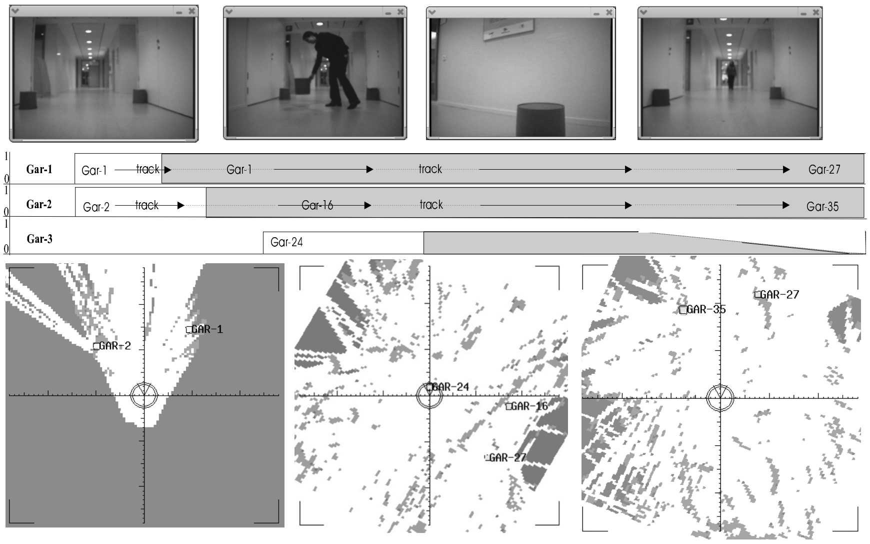

Figure 4: The top row shows the camera images at different time points, the middle row shows the activity at the anchoring level. Grey bars indicate anchors with olfactory properties. The bottom row shows the corresponding local perceptual space given the changing representation of visual percepts.

and spatial modules. In the center figure, a map of the environment is shown. Additional interfaces include a live feed of the images viewed from the CCD camera and an olfactory interface which shows clusters and data points for a loaded repository (Figure 3 (*Right*)). Certain points can be interactively selected to generate new plans for perceptual actions. It also shows the symbolic representation preserving the electronic perception of odours used to classify new odours. More information about the olfactory interface and the categorisation of odours can be found in [Loutfi and Coradeschi, 2004; 2005].

5 Experiments

The general experiment is performed in a series of corridors. In each corridor there may be several objects, in this case garbage cans. The robot automatically toggles between the task of patrolling the corridor, inspecting objects and waiting for commands from the user. Patrolling the corridor involves moving from corridor to corridor in a discovery for new objects and recognition of previous objects. When an inspect is invoked, the robot visits each object collecting the odour property. The inspect is usually autonomously invoked when new objects are detected.

The purpose of the experiment is to evaluate the ability of the extended anchoring framework to maitain an internal representation of the objects in the corridors for an extended period of time. Throughout the autonomous activity of patrolling the corridor, a user may interrupt tasks by requesting the acquisition of specific objects. Object requests can be given by using the image feed from the camera and directly selecting a region in the screen. The sensory signature of the object will be matched against current anchors and current execution of the patrol will be interrupted to include a plan to visit and inspect the selected object. Object can be requested by giving to the robot a sample of the smelling object and request to find similar objects. Finally top-down requests can be given to find objects by giving to the system a symbolic description of the object. The anchor that most matches the description will be returned.

5.1 Results

The a priori information given to the system consists of a rough map of the environment, shown in Figure 3 *(Middle)*, and a repository of interesting objects, namely garbage cans placed outside offices. The robot patrolled the corridors for a period of 4 days, with intermittent breaks during the day and longer breaks during the evening for charging the batteries. At any given time, garbage cans would be removed, displaced, or added into the environment. The total distance covered by the robot is approximately 1.2 km without the inclusion of the extra movement caused by smelling actions and over 70 odour samples were collected.

The local space of the robot together with the visual image from the camera as well as the creation, deletion and updating of anchors is depicted in Figure 4. The figure contains four snapshots throughout an experimental run described as follows:

- Scene 1 - The robot begins patrolling the corridor, two visual percepts are detected and two anchors denoted by **Gar-1** and **Gar-2**, are created. An inspect is performed and both anchors obtain olfactory properties, shown in the Figure by the grey colouring. Since the anchors are created in a bottom-up fashion their labels are arbitrary.
- Scene 2 - As the robot continues its patrol, another object is inserted into the environment at a later time. Note however, that the previous two anchors are still maintained by the track functionality. Although the local space shows only the current percepts, the anchoring

module updates the link between the anchor **Gar-1** and the percept *Gar-27*. A new anchor is also created for the third object denoted by **Gar-3** with visual percept *Gar-24*.

- Scene 3 - The robot approaches the object in order to acquire its odour property and the result is stored in the corresponding anchor. Some time later, the object is removed from the environment. The life of the anchor slowly decreases when an expected percept is no longer detected.
- Scene 4 - The anchor is removed from the system and unless it is perceived again, its properties cannot be accessed by the find functionalities described above.

This scenario shows how the anchoring module is used to create an internal structure which can then maintain the perceptual coherence of objects, considering each object has both spatial and olfactory properties. Even when visual properties of anchors are being updated, the stored smell property remains until a new odour character is acquired by the next inspect action. The previous odour character is then stored in the odour repository.

6 Conclusion

"Take a coin, toss it, and catch it again in your hand...you see a single object fly up in the air, twist over and over, and land in one piece on your hand. Bits don't fly off. The silver doesn't depart from the shape, and the shape doesn't lag behind the motion." [Blackmore, 2003] (p.244).

Maintaining perceptual coherence of objects over extended periods of time involves the management of perceptual information from different sensing sources, tracking over time and maintaining object persistency. In this paper we showed how a modified anchoring framework could be used as a tool to satisfy these requirements. Experiments on a mobile robot were performed where a robot used both spatial and olfactory sensors to monitor an office environment over an extended period of time.

The problem of perception management is far from being solved. In this paper, we have "scratched its surface" by recognizing the need to consider this problem, highlighting important issues that arise on an embedded system and presenting a first implemented solution.

Acknowledgments This work has been supported by: Vetenskapsrådet, and by ETRI (Electronics and Telecommunications Research Institute, Korea) through the project "Embedded Component Technology and Standardization for URC(2004-2008)". The authors would like to thank Mathias Broxvall and Lars Karlsson for their contribution to the experimental work.

References

[Adrian, 1993] R. Adrian. Sensor management. In *Proc of AIAA/IEEE Conference on Digital Avionics System*, pages 32–37, 1993.

[Blackmore, 2003] S. Blackmore. *Consciousness: An Introduction*. Hodder & Stoughton, Oxford, 2003.

[Chella *et al.*, 2004] A. Chella, S. Coradeschi, M. Frixione, and A. Saffiotti. Perceptual anchoring via conceptual spaces. In *Proc. of the AAAI-04 Workhsop on Anchoring Symbols to Sensor Data*, Menlo Park, CA, 2004. AAAI Press. Online at http://www.aass.oru.se/~asaffio/.

[Coradeschi and Saffiotti, 2000] S. Coradeschi and A. Saffiotti. Anchoring symbols to sensor data: preliminary report. In *Proc. of the 17th American Association for Artificial Intelligence Conf. (AAAI)*, pages 129–135, 2000.

[Coradeschi and Saffiotti, 2003] S. Coradeschi and A. Saffiotti, editors. *Robotics and Autonomous Systems, special issue on Perceptual Anchoring*. Elsevier Science, 2003.

[D.Schulz *et al.*, 2003] D.Schulz, W. Burgard, D. Fox, and A. Cremers. People tracking with mobile robots using sample-based joint probabilistic data association filters. *I. J. Robotic Res.*, 22(2):99–116, 2003.

[Karlsson, 2001] L. Karlsson. Conditional progressive planning under uncertainty. In *Proc. of the 17th Int. Joint Conferences on Artificial Intelligence (IJCAI)*, pages 431–438, 2001.

[Kleinehagenbrock *et al.*, 2002] M. Kleinehagenbrock, S. Lang, J. Fritsch, F. Lmker, G. A. Fink, and G. Sagerer. Person tracking with a mobile robot based on multi-modal anchoring. In *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication (ROMAN)*, pages 423–429, 2002.

[Knoblauch *et al.*, 2004] A. Knoblauch, R. Fay, U. Kaufmann, H. Markert, and G. Palm. Associating words to visually recognized objects. In S. Coradeschi and A. Saffiotti, editors, *Anchoring symbols to sensor data. Papers from the AAAI Workshop. Technical Report WS-04-03*, pages 10–16. AAAI Press, Menlo Park, California, 2004.

[Loutfi and Coradeschi, 2004] A. Loutfi and S. Coradeschi. Forming odour categories using an electronic nose. In *Proc. of European Conference in Artificial Intelligence (ECAI 2004)*, pages 119–124, 2004.

[Loutfi and Coradeschi, 2005] A. Loutfi and S. Coradeschi. Improving odour analysis through human robot interaction. In *To Appear Proc of the IEEE International Conference on Robotics and Automation (ICRA 2005)*, Barcelona, Spain, 2005.

[Ronnie *et al.*, 2003] L. Ronnie, M. Johansson, and N. Xiong. Perception management - an emerging concept for information fusion. *Information Fusion*, 4(3):231–234, 2003.

[Saffiotti *et al.*, 1995] A. Saffiotti, K. Konolige, and E. H. Ruspini. A multivalued-logic approach to integrating planning and control. *Artificial Intelligence*, 76(1-2):481–526, 1995.