FP6-IST-002020

COGNIRON

The Cognitive Robot Companion

Integrated Project

Information Society Technologies Priority

D 5.2.1
Cognitive Models of Space and Objects

Due date of deliverable: 31/12/2004
Actual submission date: 31/1/2005

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

Organisation name of lead contractor for this deliverable:
University of Amsterdam

Revision: Final
Dissemination Level: PU

Executive Summary

This report summarizes a number of key insights from the cognitive and neural sciences on how biological organisms represent and interact with the space and the objects around them, and describes similarities with and impact on spatial and object representation and processing in robotics and computer science. The report consists of two parts, one focusing on space and the other focusing on objects. In both parts, relevant experimental cognitive science theories are discussed next to corresponding robotic and computer science implementations. In a concluding section concrete implications for the Cogniron project are suggested.

Role of the deliverable “Cognitive models of space and objects” in Cogniron

Research area RA5, “Spatial cognition and multimodal situation awareness”, addresses the question of how an embodied system can come to a conceptualization of sensory and sensory-motor data, and generate plans and actions to navigate and manipulate in typical home settings. When trying to build a robot capable of “cognitive” tasks, it makes sense to look at the cognitive sciences for inspiration and guidance. Cogniron partners working on the problem of spatial representations have studied models for spatial representations from cognitive sciences and neurosciences that have relevance to cognitive robots.

Relation to the Key Experiments

This deliverable addresses basic research questions in RA5, “Spatial cognition and multimodal situation awareness”, which is mostly concerned with Key Experiment “Home Tour Scenario”. However, cognitive models of space and objects, as discussed in this deliverable, are used in all Key Experiments.

Contents

Cognitive Models of Space and Objects

Bram Bakker (UvA), Ben Kröse (UvA), Henrik Christensen (KTH),
Tony Belpaeme (VUB), Williams Paquier (LAAS), & Raja Chatila (LAAS)

1 Introduction

When trying to build a robot capable of "cognitive" tasks, it makes sense to look at the cognitive sciences for inspiration and guidance. In many ways, human and animal brains have to cope with the same problems as robot control systems, and much may be learned from the wealth of information about how biological brains work.

This report summarizes a number of key insights from the cognitive and neural sciences on how biological organisms represent and interact with the space and the objects around them. We do not intend to provide a complete review of what the cognitive and neural sciences have to say about space and objects; rather, we report on a limited number of highlights with particular relevance to cognitive robots.

The report has two parts, making up the next two sections: one focusing on space and the other focusing on objects. Within each part, experimental cognitive science theories are discussed next to robotic implementations that are inspired by or have similarities with those theories. Section 4 contains general conclusions.

2 The cognitive science of space representation and its application to robotics

2.1 Tolman's cognitive map

Tolman (1948) was one of the first and certainly among the most influential researchers providing evidence for and speculating about the brain's internal mechanisms that represent the organism's spatial environment. However, this type of work goes back at least to Trowbridge (1913), who used the term "imaginary maps".

Tolman and his colleagues investigated mazes with rats navigating through them while doing "reinforcement learning"; i.e. behaviorist operant conditioning experiments where rats could find food at certain positions in the maze. See figure 1 for an example of the type of maze used in these experiments. Reviewing a large number of experiments, Tolman came to the conclusion that simple stimulus-response descriptions of the type favored by contemporary behaviorist theory did not suffice to explain the rats' abilities. Instead, the rats must have learned "cognitive maps", allowing them to cope with unexpected changes in the environment almost immediately without any additional training and to generalize over mazes in a smarter way than can be explained by pure fixed route learning. Tolman (1948) informally defines a cognitive map as follows (p. 193):

> In a comprehensive map a wider arc of the environment is represented, so that, if the starting position of the animal be changed or variations in the specific routes be introduced, this wider map will allow the animal still to behave relatively correctly and to choose the appropriate new route.

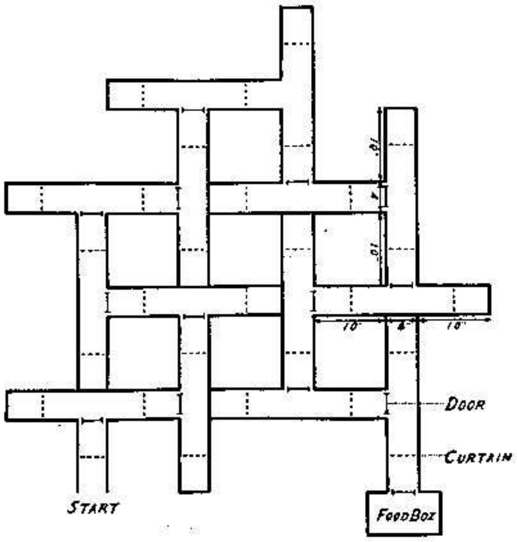

Figure 1: An example of the type of mazes used by Tolman and other researchers studying the spatial memory capabilities and cognitive maps of rats. In this maze, the rat always begins at the designated Start position, and is rewarded with food when it reaches the Food Box. From Tolman (1948).

Tolman's strongest arguments for the existence of a cognitive map in rats come from experiments where the rats were transferred from one position in the maze to another during a trial, and from experiments where the optimal path which the rat had learned was suddenly blocked. In both cases, almost immediately the rats tended to choose the path that was best given the new situation. They had not been explicitly rewarded to do that, thus showing that they had not simply memorized a sequence of actions but could generalize in a way that suggests an internal map of the maze.
Interestingly, Tolman describes cases (also reported by researchers much earlier than Tolman) where rats trained in a maze managed to push away the cover of a maze, climb out of the maze, run toward the food location directly, and climb back into the maze to eat the food. Again, this indicates that the rats have learned much more about their spatial environment than a single route or simple stimulus-response associations.
The idea of the cognitive map has been enormously influential, both in terms of follow-up work in the cognitive and the neural sciences, and in terms of inspiration for robotics. We examine examples of both below.

2.2 Neural realization of the cognitive map: hippocampus and place cells

The notion of the cognitive map has been criticized because for a long time it remained a purely theoretical construct: no hard neurological or physiological evidence, comparable to the type of physiological evidence available for conditioning, was available to support it.
However, O'Keefe and Dostrovsky (1971) reported on findings of neurons in the rat's hippocampus (in particular areas CA1 and CA3) that fire specifically when the rat finds itself in a specific place. Figure 2 shows the rat's hippocampal brain structure and its location within the rat brain; figure 3 shows how a typical place cell may respond when the rat visits the same area within its environment several times during an experimental run.
Place cells fire only in one specific place and not in others. That is, their receptive field is quite specific

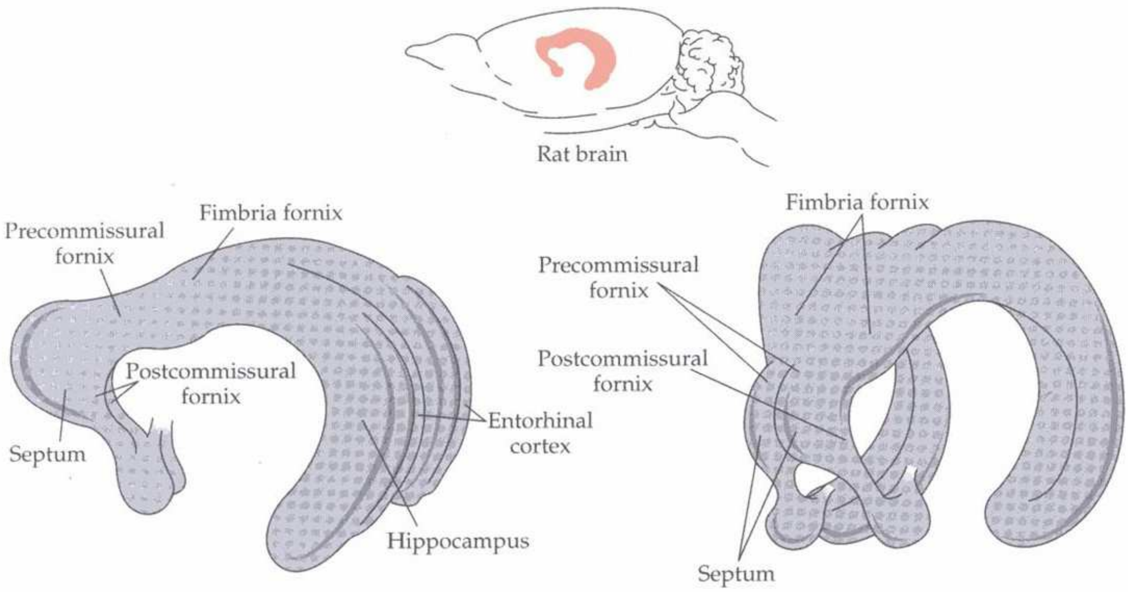

Figure 2: Side view and oblique view of the rat hippocampal brain structure (bottom), and its position within the entire rat brain (top).

and correlated with one spatial location. Other neurons in the same brain region fire correlated with other locations. In some cases, the firing patterns are not directly related to specific behavior. For example, they may fire both when the rat passes a location going one way and when it passes the same location going the other way. These so-called "place neurons" were quickly embraced as a possible instantiation of a cognitive map in the mammal brain (e.g. O'Keefe & Nadel, 1978). Implicit in most of these theories is the idea that a place cell functions as a kind of pointer ("you are here") in a larger, topographically organized *facsimile* representation of the environment.

Subsequent research has refined, and to some extent attenuated, this view of the hippocampus as realizing a cognitive map. First of all, it seems that place cells are not topographically organized: neighboring cells do not necessarily code for neighboring locations, and there is little evidence to suggest that there are direct linkages between cells coding for neighboring positions. Furthermore, firing of place cells tends to be correlated strongly with specific (visual, olfactory) cues in the environment, rather than with location per se. And despite evidence of firing patterns of place cells that are independent of specific behavior in strictly controlled randomized behavior experiments (Muller & Kubie, 1987), in more natural scenarios when there are regularities in behavior, it is hard to find firing patterns which are truly uncorrelated with behavior. Finally, even though there is little doubt that the hippocampus plays an important role in the rat's spatial memory, it is by no means its only role. In rats, and much more clearly also in humans, the hippocampus seems to be involved in a much broader variety of functions related to memory, in particular episodic memory (Eichenbaum, Dudchenko, Wood, Shapiro, & Tanila, 1999).

2.3 Systematic errors in spatial judgments

A great variety of experiments with human subjects has revealed that humans tend to make systematic errors when acting within and reasoning about space. In some cases, these errors suggest that the

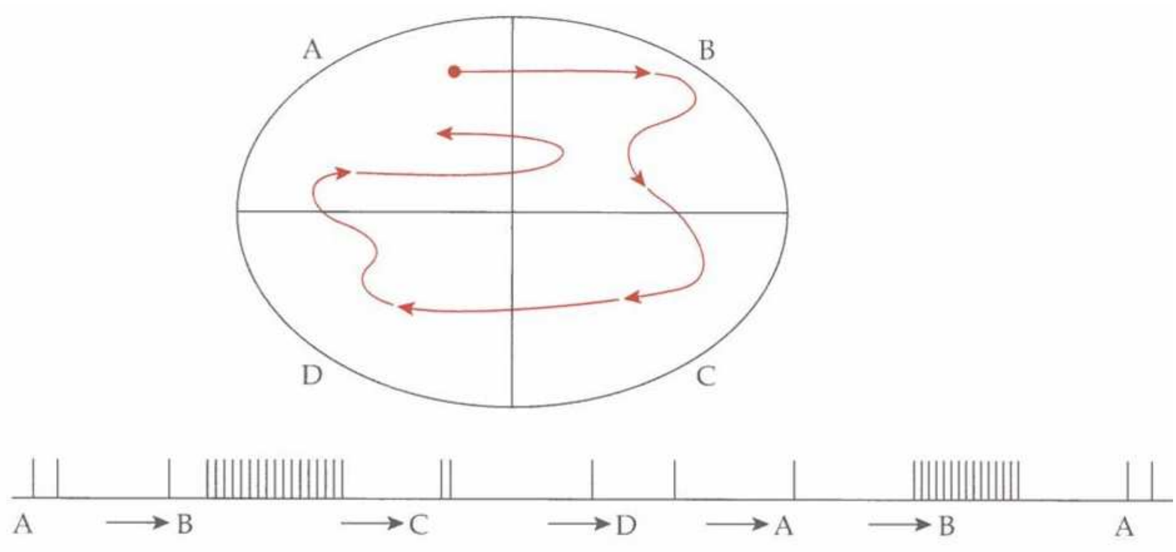

Figure 3: Typical firing patterns (below) from a place cell while the rat is running around its environment. Firing patterns are correlated with the particular place of the rat.

cognitive representations of space are organized in particular ways.
In one well-known study (Stevens & Coupe, 1978), students were asked to draw the relative positions of San Diego, California, and Reno, Nevada. A great majority of students indicate that San Diego lies to the west of Reno, whereas in fact it is the other way around, due to the fact that California cuts to the east as it cuts southward. This finding suggests that in some case memory of space is organized in a hierarchical fashion. Rather than remembering absolute locations of cities, people tend to remember cities as part of, and relative to, a state. Because California as a whole lies to the west of Nevada, people tend to group the cities within California as lying to the west of cities within Nevada.
Other studies reveal that the particular cognitive perspective taken by subjects influences spatial judgments. In one such study (Holyoak & Mah, 1982), one group of students was asked to imagine themselves on the east coast of the United States, and another group was asked to imagine themselves on the west coast. Next, both groups were asked to estimate the distance between east coast cities, as well as the distance between west coast cities. It turns out that the students given an east coast perspective overestimated the east coast distances relative to the west coast distances, and the students given a west coast perspective did the opposite. Thus, the particular perspective distorted the spatial judgments systematically, in a way that suggests that the area around the current vantage point appears larger than similar areas farther away.
Humans tend to describe and remember locations relative to prominent landmarks in the environment, such as when one says that one lives close to a particular department store. This appears to introduce a particular distortion of asymmetric distance: people estimate the distance from an ordinary building to a landmark to be smaller than from a landmark to an ordinary building (Sadalla, Burroughs, & Staplin, 1980).
Other systematic errors include the finding that distance of a path is estimated to be longer if there are more barriers, turns, or clutter, the finding that borders and rivers are straightened out in people's judgments, and that curves and angles are regularized to right angles.
In many of the cases described above, distances and directions appear to be distorted by "higher-

order" cognitive influences, such as the cognitive reference point, the hierarchical organization of one's memory, and the regularization of irregular but real properties of the environment. At the very least, this suggests that the cognitive representation of space is, to a large extent, not a purely metric affair, but is intertwined very much with cognitive factors related to interpretation and perspective.

2.4 The multimodal nature of human spatial representation

As is already apparent from the previous section, cognitive spatial representation does not involve solely metric information. An amalgam of different types of information appears to be involved in spatial reasoning and acting. However, from the variety of experiments and theories, two distinct types of cognitive representations for space emerge (Tversky, 1993).

In most situations, when environments are not known in great detail, cognitive representations of space appear to be more like a fragmented collection of different types of information than like an actual map. They may include representations of relational information, such as the information that a location is next to a specific landmark, visual information regarding specific landmarks, higher-level information about the layout of a particular city or cities in general, etcetera.

In certain specific situations, when environments are simple or well-learned, cognitive representations of space are more like actual maps as one would normally construe them. They capture spatial layouts more or less accurately, and preserve coarse spatial relations. Unlike the first type of representations, they allow generic perspective taking and reasoning, and relatively abstract spatial inference. However, like the first type of representations, they do not preserve metric information very accurately.

2.5 Robot cognitive maps

Within robotics, maps similar to and inspired by cognitive science's postulated cognitive maps are widely used, both as an intermediate representation for control generation, as a mechanism for geometric reasoning, and as a basis for reasoning and planning. Over time many different types of map representations have been considered including:

- Grids
- Geometric feature maps
- Topological maps
- Appearance models
- Conceptual spaces
- Hybrid maps

The grid representation is by far the simplest, in the sense that the representation is a homogeneous tessellation of space into a cells. Each cell can either be binary and simply denote free/occupied or it can be a probabilistic representation that is used from sensory reading using a sensor model. Grid models were in particular popularized by Moravec (1988), Elfes (1989) and in terms of probabilistic modelling by Thrun and Bücken (1996) and Fox, Burgard, and Thrun (1998).

Topological maps have been widely used for discrete mapping of the environment. The topological maps have been implemented as a graph of places in which each node is a topologically district region of space. Early work was reported by Kortenkamp and Weymouth (1994), and more recently it has been popularized by Choset and Burdick (2000) and Choset and Nagatani (2001) for modelling using generalized Voronoi maps. In addition, the discrete representation has been used for mapping in applications such as domestic navigation (Althaus & Christensen, 2003).

Feature maps are probably the most widely used model for robot systems in terms of points, lines (Jensfelt & Christensen, 2001), regions of constant depth (Leonard & Durrant-Whyte, 1992), etc. The use of such geometric maps has been used widely both for passive mapping and as part of simultaneous localization and mapping (SLAM) (Leonard & Durrant-Whyte, 1991; Csorba, 1997; Moutarlier & Chatila, 1989; Castellanos, Neira, & Tardos, 2001; Davidson & Murray, 2002). The basic premise here is to design a correspondence strategy for alignment of sensory readings with a prior defined map. The feature maps have the advantage that they are computationally efficient and space adaptive.

Appearance models try to model the world in terms of its direct sensory appearance. The appearance can both be in terms of visual imagery or as laser data aligned in the world frame. The advantage of the appearance models is that matching can be performed directly on the sensory information as for example performed by Gutmann and Schlegel (1996) and Crowley, Wallner, and Schiele (1998). In visual processing similar results have also been reported by Tell and Carlsson (1999) for object localization and alignment.

Conceptual spaces have been proposed in particular by Gärdenfors (2000) and exploited by Balkenius (1998) for robot navigation. The idea here is to embed cognitive labels and associated geometric features in high dimensional space where qualitative information is treated as yet another dimension. The representation has the advantage that it easily integrates new information, but at the expense of complex matching functions that must integrate qualitative and quantitative features into a single metric.

As mentioned in the previous sections, sophisticated spatial cognition goes beyond simple geometry, and extensive cognitive processing is required to integrate sensory information with qualitative world knowledge such as labels and associations. Integrating these into a single unified representation might not always be easy to accomplish. Consequently, research has also considered design of hybrid representations in which different characteristics can be utilized. One of the most comprehensive models presented is the Spatial Semantic Hierarchy (SSH) proposed by Kuipers (2000, 1977) and Kuipers and Byun (1991). This work is explicitly inspired by cognitive science insights regarding the multimodal nature of human spatial representation, described above. The model dates back more than 20 years but has recently been used in experimental studies. The SSH model integrates qualitative and quantitative information as two parallel (and related) hierarchies that are organized according to sensory, control, causal, topological and metrical layers. Each layer has a specific function in terms of reasoning and/or perception action integration. In addition each layer has a clear causal relation to other layers and there are specific cognitive functionality associated with each of the individual models. So far the SSH has been used for spatial navigation tasks for robots as seen for example in the home-tour scenario, but the model has not yet been explored for agents that actively interact with the environment for manipulation and general scale 3D modelling. A desirable property of the deployed model is that it has strong logic and control primitives associated that support reasoning about change etc.

2.6 Robot cognitive maps inspired by hippocampus and place cells

In a number of mobile robotics studies concerned with mapping and navigation (Mallot & Franz, 1999; Arleo, Smeraldi, Hug, & Gerstner, 2001; Mataric, 1991; Hafner, 2000; Guazzelli, Corbacho, Bota, & Arbib, 1998; Trullier, Wiener, Berthoz, & Meyer, 1997; Trullier & Meyer, 1998; Tapus, Ramel, Dobler, & Siegwart, 2004), references are made to and inspiration is found in the findings on the involvement of the hippocampus and the existence of place cells (see section 2.2). In some

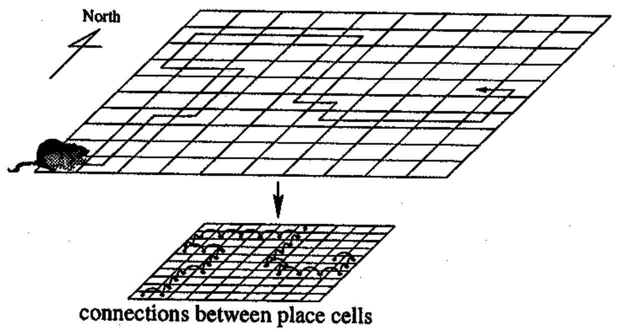

Figure 4: Neural model of the cognitive map based on the findings on the hippocampus and place cells (Trullier & Meyer, 1998). Note the isomorphic mapping from the real world to the neural structure.

cases, the neuroscience findings are merely used as a general justification for the use of internal, possibly topographically organized maps; in other cases, the hippocampus and place cells are explicitly modeled.
For instance, in Arleo et al. (2001), Guazzelli et al. (1998), and Trullier and Meyer (1998), place cells are explicitly modeled in simulated topological cognitive maps, and the maps are used for navigation by robotic agents. As is implicitly the case in cognitive neuroscience's theories, place cells are assumed to function as pointers, representing the organism's positions in the entire, topographically organized neural cognitive map (figure 4).
Mataric (1991) presents a more abstract model of the rat hippocampus and its role as a kind of cognitive map, within a behavior-based robotics context. This cognitive map is different from most conceptions of the cognitive map as described above. It consists of a network of nodes, each representing a registered landmark and corresponding movement (see figure 5). The current location of the robot is represented by one of the nodes being active. Activation spreads to other nodes, thus generating "expectations" about what will be perceived when the robot performs the corresponding movement. A goal location can also become active, and activation will similarly spread out to neighboring nodes. This spreading of activation depends on the physical distance between landmarks. Consequently, locally at each landmark suggestions can be made as to which direction to go to reach the goal most rapidly. Overall, this results in the robot choosing the globally shortest path to the goal. Determination of the current position, map building, and action selection are not separated into distinct functional components, but they are all combined in this single map. There is no central planning mechanism figuring out the optimal path to the goal; the navigation behavior emerges as the result of interacting local units.

2.7 The ecological approach

The cognitive map and other hypothesized information-processing brain mechanisms are examples of the type of models that are developed within the so-called "cognitivist" approach to spatial cognition. These models are typically derived from carefully crafted experiments where animals and humans are subjected to very artificial conditions with minimal sensory inputs, to determine how the limited information available through perception in these experiments is processed so as to arrive at internal

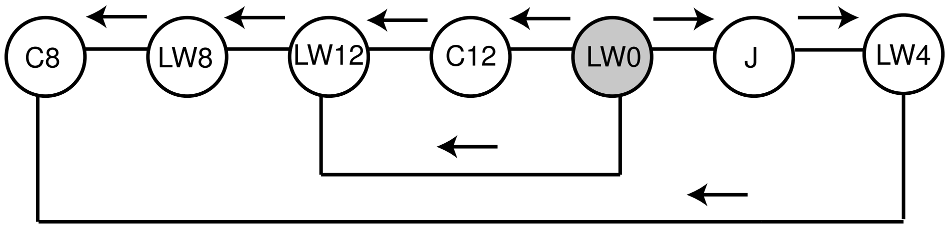

Figure 5: Cognitive map learned by a robot in a cluttered office environment (Mataric, 1991). Nodes indicate both places and movements; LW8 means Left Wall heading south, for instance. Arrows denote the spreading of activation from the goal (gray node).

representations of the environment.

This approach has led to what can be called a counter-movement, called the ecological approach. The ecological approach (Gibson, 1950, 1979; Turvey & Carello, 1981), or the Gibsonian or direct perception approach, emphasizes the richness of ecologically realistic information available through perception of the natural world. It emphasizes in particular the fact that the organism and external elements in the environment are virtually always in motion. The claim is that relevant properties of the environment can be picked up "directly" from the rich perceptual stream, and that this does not require extensive internal operations computing complex internal world models.

The prototypical example of this principle is the estimation of "time to contact" of an object approaching an organism in its visual field. The time to contact can be estimated directly from the proportion of the size of a retinal image caused by the approaching object and its rate of expansion on the retina, without the need for explicit estimation of the distance of the object.

More generally, the ecological approach focuses on the structure in changing patterns of light, optic flow, and how it affects and is affected by movement through the environment. See figure 6 for a schematic representation of the principle of optic flow. Through this focus on the dynamic nature of perception of the environment, the ecological approach distinguishes itself from most cognitivist work on perception of the environment, which focuses more on simple, static situations.

Furthermore, the ecological approach emphasizes, more than the cognitivist approach, the notion that the purpose of perception of the environment is not necessarily to construct internal representations of models of the external world, but to extract information for *action*. Therefore, research within the ecological approach is often aimed at discovering how particular information relevant for action can be extracted from the perceptual stream. The extraction of time to contact information, described above, is one example of such a study. In this context, information that is relevant for action is sometimes referred to as *affordances*. Affordances reflect the possibilities for actions within an organism-environment setting. For example, in a situation where a ball is thrown around, an affordance that may be extracted from the perceptual stream is the "catchability" of the ball.

In sum, the ecological approach to spatial cognition focuses much more than cognitivist approaches on the dynamic nature of perception of the environment, the richness of information in the perceptual stream, and the possibilities for action extracted more or less directly from the perceptual stream.

2.8 Behavior-based robotics and the ecological approach

The work of Mataric, presented in section 2.6, and others working within behavior-based robotics (e.g. Brooks, 1989, 1991a, 1991b), has important similarities with the ecological approach described

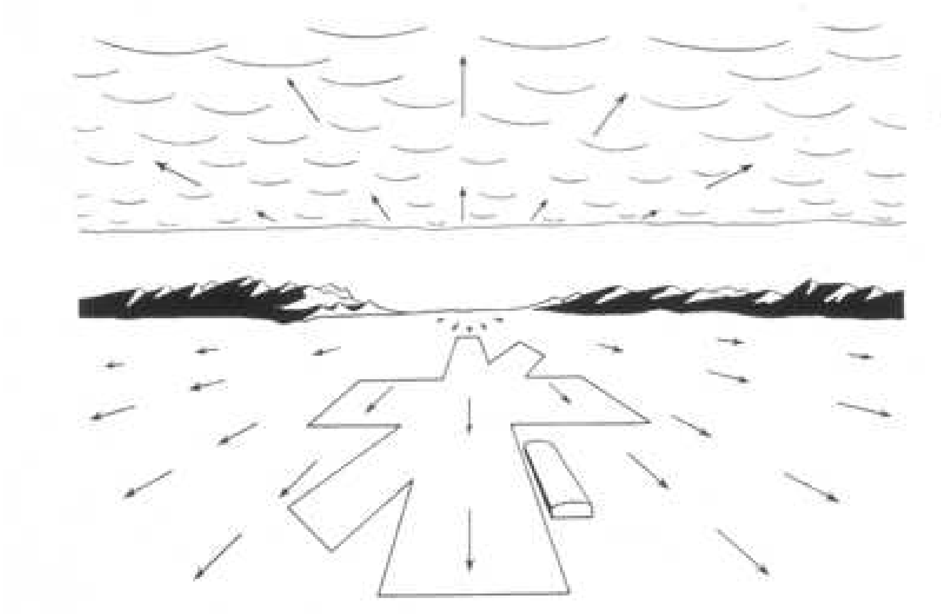

Figure 6: Schematic representation of the optic flow field of a plane or a bird flying straight ahead.

above. This was pointed out before by e.g. Clark (1997) and Duchon and Warren (1994).
Like the ecological approach, behavior-based robotics rejects extensive internal models of the environment, and it emphasizes the direct extraction of opportunities for action ("affordances") from perceptual information. The ecological approach ideas concerning optic flow and time-to-contact have been applied to mobile robots for obstacle avoidance and navigation by several groups (Nelson & Aloimonos, 1989; Duchon & Warren, 1994; Dev, Kröse, & Groen, 1997; Huber, Franz, & Blthoff, 1998).

3 The cognitive science of object recognition and its application to computer vision and robotics

3.1 Visual object recognition

Visual object recognition is a central topic in the cognitive sciences. The question addressed by research is to find out how “objects” are represented and recognized from the sole information coming in through the visual system. The approaches must account for the capacity of recognizing objects from different views and in variable conditions, but also for the property of generalization across different examplars of objects in a same category.
In the early eighties, there was very little evidence to support and validate explanations of how 3D objects and corresponding categories are extracted from visual information by humans (and primates). The main theory on computational models of object recognition was Marr and Nishihara’s (Marr, 1982; Pinker, 1985). Since then, a large number of studies from psychophysics (Jolicoeur, 1985; Biederman, 1987; Tarr & Pinker, 1989) and neurosciences (Perrett, Mistlin, & Chitty, 1987; Perrett et al., 1985, 1987; Goodale & Milner, 1992; Logothetis & Sheinberg, 1996; Tanaka, 1996), accomplished tremendous progress that lead to other computational models (Poggio & Edelman, 1990; Ullman & Basri, 1991).
Evidence from psychophysics and neuroscience tends to support the claim that representations are view-based and objects are represented by sets of local features that are related together, rather than structural 3D shape descriptions from a set of primitives (”geons”). However, both approaches have some shortcomings and there has been much controversy among the competing theories. An excellent review of this topic, from which this section borrows extensively, can be found in Tarr and Bülthoff (1998). Next, we present the two approaches and some arguments on the possibility of a mixed approach.

3.2 The structural approach

The structural approach is based on the idea that sets of features are systematically collected from the visual information and deterministically combined in a hierarchical manner to provide a 3D reconstruction of the observed scene. A central assumption is that instead of being view-based, the representation is object-based (Marr & Nishihara, 1978). This hypothesis seems indeed necessary to account for the stability of the representations across different views and for avoiding the combinatorial explosion that would occur if only view-based representations were used.
One main approach based on this stream of work is the Recognition-By-Components (RBC) model introduced by Biederman (Biederman, 1987; Hummel and Biederman, 1992). This model proposes a volumetric reconstruction of objects based on components related together by spatial relations that connect such volumes (figure 7).
This approach is, however, not very well supported by experimental evidence in psychophysics and neuroscience which tends to show that view-based representations are used by humans and higher primates.

3.3 The view-based approach

The view-based approach proposes that representations are sets of local features in different images or views of the same scene. These features will help the recognition process provided they are present in the perceived view (Riesenhuber & Poggio, 1999) (see figure 8).

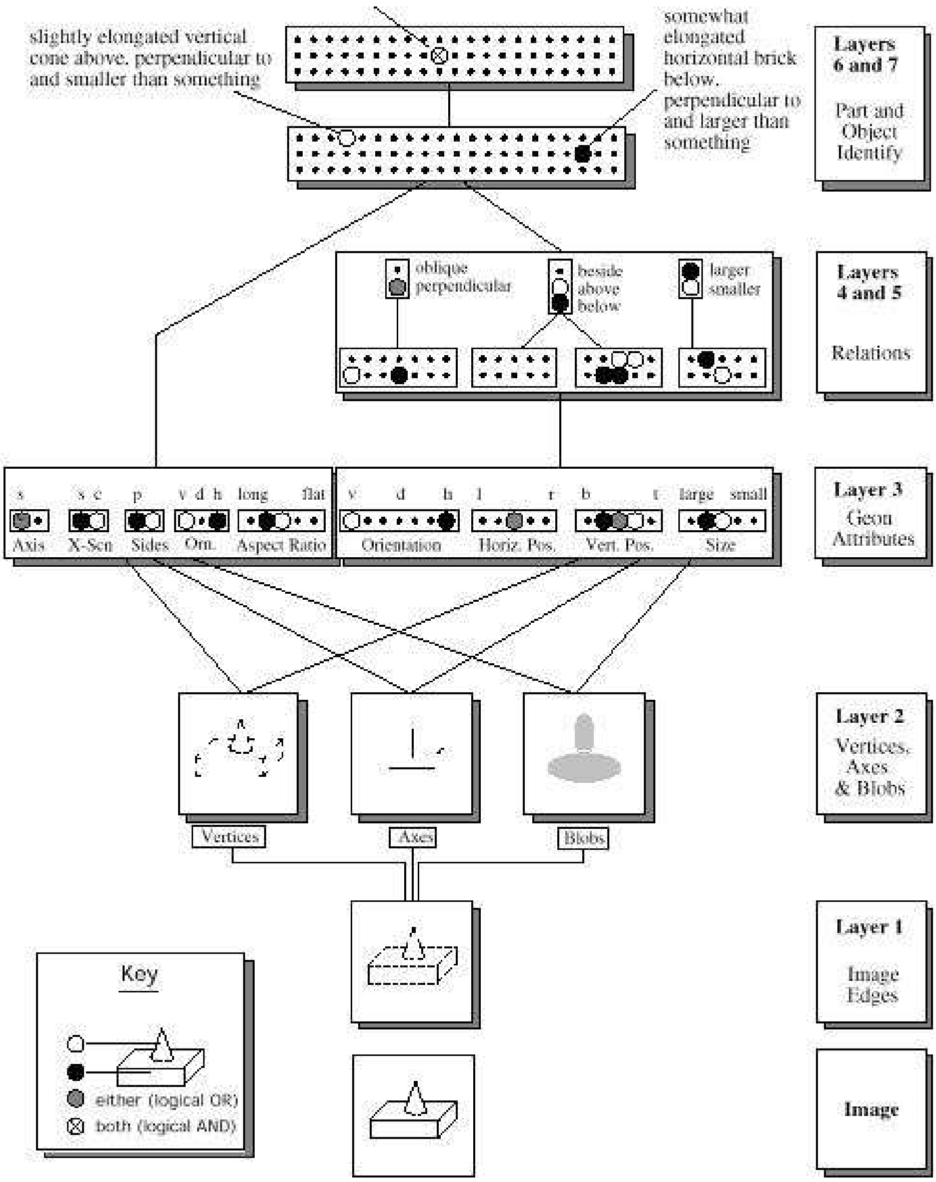

Figure 7: Dynamic Binding in a Neural Network for Shape Recognition (from Hummel and Biederman, 1992).

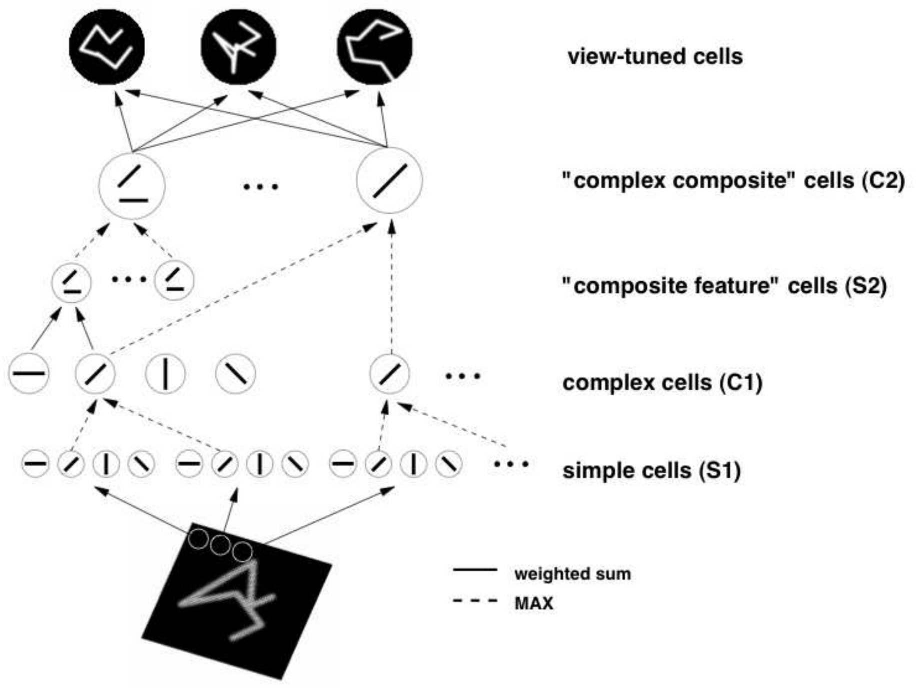

Figure 8: Example of the view-based models. It is an hierarchical model building complex cells from simple cells. The circuitry consists of a hierarchy of layers leading to greater specificity and greater invariance by using two different types of pooling mechanisms (Riesenhuber and Poggio, 1999).

However, view-based approaches must also account for generalization across different instances of a similar objects to extract the notion of class, not only the same object with different viewpoints.

3.4 Evidence for Mixed Models

In fact, both the structural approach and the image-based approach have problems to fully explain the object recognition process. Structural approaches are too deterministic. Image-based models have difficulties to explain class generalization and categorical representation, are sensitive to the views, and require matching algorithms and normalization mechanisms for coping with invariance. The latter approach can, however, be extended to enable interpolation across views and across exemplars, as well as temporal associations with associative memory mechanisms. Structural descriptions (not necessarily RBC) have the relative advantage that they can more easily account for categorization. Image or view-based representations, on the other hand, may be better suited for the recognition process for a given instance or object.

3.5 Computational models of object recognition

Computational models of object recognition reflect, to a large extent, the distinction between the structural approach and the image-based approach described above.
Structural decomposition methods, inspired by the structural approach, store an object as a set of simple geometric primitives joined together using spatial relationships. In its simplest form, objects

are represented as stick figures, but the primitives can be elaborated to represent cylinders, spheres, blocks, etcetera (Edelman, 1999). This method, although attractive because of its conceptual rigour and its invariance to changes under different viewing conditions, loses much of its attractiveness due to the fact that no reliable methods exist to extract structural primitives from raw images and because some objects, such as a shoe, are notoriously difficult to describe as a composition of primitives (for a critical assessment see Edelman, 1997).

The image-based approach provides inspiration for so-called memory-based methods, which simply store for each object a number of snapshots taken from different viewpoints. Another approach within the image-based approach uses geometric constraints (e.g. Lowe, 1987; Ullman, 1989). Here an object is represented by the coordinates of a small set of salient features and their relative geometrical positions. This method has proven to be highly invariant to changes in viewing conditions, and performs well as long as the stored features can be extracted from the raw image (Lowe, 2004).

A related class of object modelling methods within the class of image-based methods relies on feature spaces. An object is represented by a vector of feature values. The features can contain any type of information, with values being discrete or continuous. Examples of features are: hue, size, number of corners, etcetera. The features form a high-dimensional feature space, in which several statistical analysis techniques can be applied to unveil structure and to generalize over a number of examples. The features are usually easy to detect and extract from raw images, and the feature-based modelling combines detection, recognition and categorization in one elegant mechanism. However, structure and category properties are represented implicitly, rather than explicitly as the structural-decomposition approach. Also, it is not straightforward to extract higher-level information, such as a 3D-description or functionality, from a feature vector. Examples of this approach can be found in Schiele and Crowley (1996), Mel (1997) and Belpaeme, Steels, and Looveren (1998). The particular structure of a feature representation, e.g. a colour histogram or a point representation in the n-dimensional feature space, determines the ease and robustness of recognition. Connectionist solutions to object representation also fall within the feature space approach. Artificial neural networks for example cut up the feature space in order to seperate distinct objects, the network then serves to classify new perceptions of objects.

The feature space approach suffers from three problems. Firstly, it is difficult to find features for which two similar, but distinct, objects (such as the faces of two individuals) will be discriminable enough to distinguish between the two objects. At the same time the feature vector should be insensitive to object transformations or deformations. The compromise between sensitivity and stability of features is difficult, if not impossible to solve (Marr, 1982). Secondly, when learning from examples in a high-dimensional feature space, many examples are needed to delineate the region feature space that corresponds to each concept. The more dimensions, the more examples will be needed. If no structure is present in the examples (which might be extracted using statistical or dimensionality reduction techniques) this will prove to be problematic (Edelman, 1997). And thirdly, feature vectors typically do not contain information on abstract or high-level properties of the concepts, such as the manipulation of the object, its affordances or its structural constituents. However, hybrid approaches, combining feature vectors with extra information and hierarchical structure might provide a solution (e.g. Edelman, 1999).

3.6 Additional constraints on object recognition

Recent results in psychophysics have added new constraints on cognitive object recognition models. One of these constraints focuses on the time a human or a monkey needs to differentiate two visual categories. The other focuses on the need for attention to perceive changes in scenes.

Thorpe and others measured the time it takes for the human visual system to process a complex natural image (Thorpe, Fize, & Marlot, 1996). Recognition of familiar objects and scenes appears to be virtually instantaneous. They measured this processing time experimentally using two different approaches: behavioural measures such as reaction times, and event-related potentials (ERPs). They used a go/no-go categorization task in which subjects have to decide whether a previously unseen photograph, flashed on for just 20 ms, contains an animal. ERP analysis revealed a frontal negativity specific to no-go trials that develops roughly 150 ms after stimulus onset. The result of this study is that the visual processing needed to perform this highly demanding task can be achieved in under 150 ms.
This result gives an important constraint on the architecture and the processes need to accomplish a recognition task. The visual cortex is mainly used as a feedforward pipeline. The first wave of spikes carry all the information. A model called the Rank Order Coding describes this feed forward categorization process (Thorpe & Gautrais, 1998). SpikeNET is an implementation of this system, and it has been used for face recognition, optical flow recognition, etc.
A second constraint on object recognition models comes from studies on active vision and attention. Recent studies have shown that important changes in a visual scene (color, size, texture or shape of objects) can go undetected if the changes coincide temporally with the moment an eye saccade occurs (O'Regan, Rensink, & Clark, 1996; Rensink, O'Regan, & Clark, 1997). However, this effect can also be produced by a screen flicker for example, which tends to show that it is related to the visual transient produced by the saccade, not the saccade itself. This evidence appears to show that the attentional process plays an important role in detecting changes. The consequences of this result for cognitive models of object may be very important. According to these experiments, active vision and attentional mechanisms could considerably simplify the recognition mechanism in the view-based approach. It is not necessary to detect objects everywhere in the visual sensor, but one may move the camera in order to center the interest area of the image.

3.7 The ecological approach to object modeling

As explained above (sections 2.7 and 2.8), there has been a reaction against the standard, cognitivist view of the brain as an information-processing machine that processes limited amounts of sensory information with the goal of arriving at internal representations of the environment. This reaction is called the ecological or Gibsonian approach. Within robotics, Brooks (1990, 1991b), in his landmark articles, argued that robots could well do without elaborate internal representations. This also was a counter reaction to years of robot research which relied heavily on structure-based, geometrical and hierarchical representations of objects, research to which Brooks himself contributed for years (Brooks, 1994). That approach had led the field to use ever more simple experimental environments; these so-called blocks worlds were tailored to fit the limits of the computational models. Instead, Brooks claimed that roboticists should have been constructing computational models that could cope with the complexity of the actual world. The "reductionist" approach to object modelling was countered by Brooks and others by refraining from using any models at all. As Brooks emphasized: "The world is its own best model".
The crux of Brooks' ideas is that robots should be situated systems and they should have physical "grounding". The robot cannot be seen as an isolated system, but should be part of a context or an environment, possible interacting with other robots. The environment and other systems have a profound influence on the behaviour and representations of the robot. These ideas can be related to the views of the very first cybernetics researchers, such as Grey Walter and Ross Ashby. More recently, the dynamical systems approach seems to provide the theoretical underpinning for the ideas

first presented in the field of behaviour-based robotics (Clark, 1997).
The same ideas resonated in computer vision. Ballard (1991) and Aloimonos (1993), for example, took a fresh look at many outstanding problems in computer vision. They did not immediately argue for a radical change in computer vision where models and representation would be superfluous. But they did show that taking into account information from the environment could simplify the task of computer vision considerably. Computer vision algorithms up till then focused on still, achromatic images (confer the shape-from-shading approach or the model-based approach), and had been neglecting much information that could aid image understanding, e.g. taking different viewpoints, or using attention mechanisms such as mentioned in the previous section.

3.8 Multimodal object modelling

The strong reliance of object modelling on visual information reveals one of its largest problems: the fact that visual information alone is often not enough to go beyond object detection and object recognition. Visual information for example does not suffice to construct object categories and to make generalizations. Humans' object concepts on the contrary only rely partially on visual information: most object categories are defined by linguistic labelling or by functional properties.
Furthermore, in robotics, object detection and recognition is only one aspect of object modelling. Robots need to be able to do more than just recognize objects: they need to handle objects; for this they need to infer the use of objects, their functionalities and functional limitations. They also need to be able to communicate about objects. For this a two-way linguistic interaction is needed: the robot needs to be able to recall the object when it hears a word or a description, this is the interpretation part of the problem, and the robot needs to be able to produce a linguistic description for an object, this is the production part.
It is immediately clear that visual information will not suffice to solve these problems. The field of computer vision has not solved object modelling, and will not as they are lacking fundamental information to construct object models as needed by robotics. The solution calls for multi-model perception of objects, using not only visual perception but also haptic perception, active perception and other descriptors, extracted from contextual information. Linguistic input plays a crucial part in building object models, this is further elaborated in section 3.10.

3.9 Lessons from developmental psychology

Adults perceive objects as structured three-dimensional entities and have little or no problem assigning categories (such as animal, plant, inanimate, graspable) to novel objects. Children are quick to grasp the classification of objects, even though they still struggle with inclusion and exclusion of category membership; children typically overgeneralize categories, for example, classifying cats as dogs. Strangely enough this overgeneralization is not symmetric. French, Mareschal, Mermillod, and Quinn (2004) shows how children in general do not make the mistake of classifying dogs as cats.
The central question is "What is it that holds things together and affords a set of objects some coherence under a category name?" (Landau, Smith, & Jones, 1998, p. 19). Categorization can be either a bottom-up process, that is, categorization uses only perceptual information, or categorization can be a top-down process, thus relying only on conceptual information. Quine (1977) called the former *intuitive* classification and the latter *theoretical* classification. Children seem to be primarily using a bottom-up process, while adults rely more on top-down processing. Although it is instructive to introduce this dichotomy in conceptualization, it is obvious that most (if not all) adult conceptualization will rely on a combination of both bottom-up and top-down processes. Experiments with preschool

children show that children at the age of four already use conceptual knowledge of categories to make category judgements (Keil, 1994); in order words, after three to four years of bottom-up categorization, children gradually start using top-down categorization.
Experiments with children and adults have shown that adults during category judgements often rely on the function of artifacts, rather than trusting on appearance or shape of the artifact to assign it to a category (Bloom, 2000). So chairs, hammers, and spoons are not so much classified according to their shape, texture, or colour, but rather according to their function (Landau et al., 1998). Seen from a computational perspective this leaves us with a paradox: how can an intelligent system classify an object based on its function, if in order to judge the object's function it needs to first recognize it? Our knowledge of objects such as "dog" or "chair" goes beyond the shape or appearance of the objects. Shape and appearance similarity often is not a good criterion to determine the concept an object belongs to. Rather, adults rely on higher-level knowledge to assign objects to concepts. Children on the other hand rely on simpler mechanisms to assign objects to concepts, as do adults when they have no other information to rely on. Probably a lesson can be learnt from this: the conceptualization of an intelligent system could be bootstrapped using shape and appearance similarity, and could then gradually evolve to use other, top-down, characteristics of the objects. Language could play a crucial role in this second phase.

3.10 Language and conceptualization

Even at an early age, words seem to play an important role in learning object representations. Xu (2002) reports experiments where nine-month old babies were tested in keeping track of objects. Objects (a ball or a duck) were shown to the infants, together with a description of the objects ("Look Maggie a ball" or "Look a toy") and placed behind a screen. When the screen was lifted and the situation that was unveiled did not match the description just given, the babies showed surprise (measured by the time they looked at the objects). Seemingly the infants used the linguistic descriptions of the objects to keep track of what could be expected behind the screen. While doing so they were able to generalize across categories: they seemed to know that "toy" comprised both "ball" and "duck". These results suggest that language plays an important role in the acquisition of object concepts.
As mentioned before, studies show that children start to conceptualize by first relying on perceptual cues. From the age of nine months on they rely more on functional information and on linguistic labels to learn and refine exisiting concepts (Booth & Waxman, 2002). How exactly children learn concepts using lexical labels is not quite clear. What is known is that children start by overgeneralising linguistic concepts (for example by understanding the word "doggie" to include cats) after which they steadily refine the concept using negative and positive examples.
Steels and Belpaeme (2005) propose a computational model to acquire concepts through linguistic interactions. The crux of the model is that conceptualization and lexical labels associated with the concepts are learned simultaneously. During learning, the words influence concepts, and vice versa, the concepts influence the use of words. One of the strong points of this method, relevant to learning of abstract concepts, is that the language serves as a delineation of concepts in feature space without the need for preprogramming or specific instruction of the learner. Conceptualization is no longer based on perceptual cues, although perceptual information still plays a central role in perceiving objects, but the concepts are shaped according to their linguistic labels. Language is the channel through which concepts are transmitted from one agent to the other. One could even think of humans teaching concepts to robots using these linguistic interactions.

4 Conclusions

Cognitive science insights in the way humans and animals interact with and represent the space and objects around them has interesting similarities with, and in some cases a significant impact on the way robotics and artificial intelligence handles spatial and object representation and processing.
With regard to spatial representation and processing, the cognitive science concept of cognitive map has been particularly influential in inspiring a variety of different types of robotic cognitive maps. In recent years, neuroscience findings on the involvement of the hippocampal brain structure in organisms' cognitive maps has led to robot studies which explicitly refer to or model this knowledge. Findings on the multimodal (not solely metric) nature of human spatial knowledge have inspired multimodal spatial representation in robots, and the Cogniron project will pursue this line of research as well.
With regard to object representation and processing, the described approaches to human object representations and recognition support the idea to use mechanisms that combine view-based and structure-based aspects, and suggest active vision as a way to cope with combinatorial explosion problems. The approaches to object representation and recognition in the Cogniron project will develop similar approaches in robots and investigate their limitations. Furthermore, as in spatial processing there are strong arguments for assuming multimodal (not solely visual) information in human object processing, such as higher-level conceptual and linguistic information, and object processing algorithms in the Cogniron project may similarly exploit these types of multimodal information.

References

Aloimonos, Y. (Ed.). (1993). *Active perception.* Lawrence Erlbaum Associates.

Althaus, P., & Christensen, H. (2003). Automatic map acquisition for navigation in domestic environments. In *Icra-03.* Taipei.

Arleo, A., Smeraldi, F., Hug, S., & Gerstner, W. (2001). Place cells and spatial navigation based on 2d visual feature extraction, path integration, and reinforcement learning. In *Advances in neural information processing systems, NIPS 13.*

Balkenius, C. (1998). Spatial learning with perceptually grounded representations. *Robotics and Autonomous Systems*, *25*(3–4), 165–176.

Ballard, D. (1991). Animate vision. *Journal of Artificial Intelligence*, *48*, 57-86.

Belpaeme, T., Steels, L., & Looveren, J. van. (1998). The construction and acquisition of visual categories. In A. Birk & Y. Demiris (Eds.), *Proceedings of the 6th european workshop on learning robots* (p. 1-12). Berlin: Springer.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115-147.

Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: The MIT Press.

Booth, A., & Waxman, S. (2002). Object names and object functions serves as cues to categories for infants. *Developmental Psychology*, *38*, 948-957.

Brooks, R. A. (1989). A robot that walks: Emergent behaviors from a carefully evolved network. *Neural Computation, 1*, 253–262.

Brooks, R. A. (1990). Elephants do not play chess. *Robotics and Autonomous Systems*, *6*, 3-15.

Brooks, R. A. (1991a). Intelligence without reason. In *Proceedings of the 12th internal joint conference on artificial intelligence.* San Mateo, CA: Morgan Kauffman.

Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159.

Brooks, R. A. (1994). *Model-based computer vision.* Ann Arbor: UMI Research Press.

Castellanos, J. A., Neira, J., & Tardos, J. D. (2001). Multisensory fusion for simultaneous localisation and mapping. *IEEE Trans on Robotics and Automation*, *17*(6), 908–914.

Choset, H., & Burdick, J. (2000). Sensor-based exploration: The hierachical generalized voronoi graph. *The International Journal of Robotics Research*, *19*(2), 96–125.

Choset, H., & Nagatani, K. (2001). Topological simultaneous localisation and mapping: Towards exact localisation without explicit localisation. *IEEE-TRA*, *17*(2), 125–137.

Clark, A. (1997). *Being there: Putting mind, body, and world together again.* Cambridge, MA: MIT Press.

Crowley, J. L., Wallner, F., & Schiele, B. (1998). Position estimation using principal components of range data. In *Proc. of the ieee international conference on robotics and automation (icra'98)* (pp. 3121–3128). Leuven, Belgium.

Csorba, M. (1997). *Simultaneous localisation and mapping.* Unpublished doctoral dissertation, Univ of Oxford.

Davidson, A. J., & Murray, D. W. (2002). Simultaneous localisation and map-building using active vision. *IEEE Trans. on PAMI*, *24*(7), 865–880.

Dev, A., Kröse, B. J. A., & Groen, F. C. A. (1997). Navigation of a mobile robot on the temporal development of the optic flow. In *Proceedings of the 1997 IEEE/RSJ international conference on intelligent robots and systems, iros 1997.*

Duchon, A., & Warren, W. (1994). Robot navigation from a gibsonian viewpoint. In *Proceedings of the 1994 IEEE conference on systems, man and cybernetics.*

Edelman, S. (1997). Computational theories of object recognition. *Trends in Cognitive Sciences*, *1*(8), 296-304.

Edelman, S. (1999). *Representation and recognition in vision.* Cambridge, MA: The MIT Press, a Bradford book.

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., & Tanila, H. (1999). The hippocampus, memory, and place cells: Is it spatial memory or a memory space? *Neuron*, *23*, 209–226.

Elfes, A. (1989). Using occupancy grids for mobile robot perception and navigation. *Computer*, *22*(6), 46–57.

Fox, D., Burgard, W., & Thrun, S. (1998). Active markov localization for mobile robots. *Robotics and Autonomous Systems*, *25*, 195–207.

French, R. M., Mareschal, D., Mermillod, M., & Quinn, P. C. (2004). The role of bottom-up processing in perceptual categorization by 3- to 4-month-old infants: Simulations and data. *Journal of Experimental Psychology: General*, *133*(3), 382-396.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought.* Cambridge, MA: MIT Press.

Gibson, J. J. (1950). *The perception of the visual world.* Boston, MA: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*, 20-25.

Guazzelli, A., Corbacho, F., Bota, M., & Arbib, M. A. (1998). On robots and flies: Modeling the visual orientation behavior of flies. *Adaptive Behavior: Special Issue on Biologically Inspired Models of Spatial Navigation*, *6 (3/4)*, 435–471.

Gutmann, S., & Schlegel, C. (1996). Amos: Comparison of scan-matching approaches for self-localization in indoor environments. In *1st euromicro conf on adv. mobile robotics.*

Hafner, V. (2000). Learning places in newly explored environments. In J.-A. Meyer, A. Berthoz, D. Floreano, H. Roiblat, & S. Wilson (Eds.), *Proceedings of the sixth international conference on simulation of adaptive behaviour: From animals to animats. the int. society for adaptive behavior.*

Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology*, *14*, 328–352.

Huber, S., Franz, M., & Blthoff, H. (1998). On robots and flies: Modeling the visual orientation behavior of flies. *Robotics and Autonomous Systems*, *29*, 227-242.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*(3), 480-517.

Jensfelt, P., & Christensen, H. I. (2001). Pose tracking using laser scanning and minimalistic environmental models. *IEEE Transactions on Robotics and Automation*.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, *13*, 289-303.

Keil, F. (1994). Explanation, association, and the acquisition of word meaning. In L. Gleitman & B. Landau (Eds.), *The acquisition of the lexicon* (p. 166-196). Cambridge, MA: The MIT Press.

Kortenkamp, D., & Weymouth, T. (1994). Topological mapping for mobile robots using a combination of sonar and vision sensing. In *Proc. of the national conference on artificial intelligence (aaai-94).*

Kuipers, B. J. (1977). *Representing knowledge of large-scale space* (Tech. Rep. Nos. TR-418 (revised version of Doctoral thesis May 1977, MIT Mathematical Department)). MIT Artificial Intelligence Laboratory.

Kuipers, B. J. (2000). The spatial semantic hierarchy. *Artificial Intelligence*, *119*, 191–233.

Kuipers, B. J., & Byun, Y. (1991). A robot exploration and mapping strategy based on semantic hierarchy of spatial representations. *Robotics and Autonomous Systems*, *8*, 75–91.

Landau, B., Smith, L., & Jones, S. (1998). Object perception and object naming in early development. *Trends in Cognitive Sciences*, *2*(1), 19-24.

Leonard, J. J., & Durrant-Whyte, H. F. (1991). Simultaneous map building and localization for an autonomous mobile robot. In *Proc. of the international workshop on intelligent robots and systems* (Vol. 3, pp. 1442–1447). Osaka, Japan.

Leonard, J. J., & Durrant-Whyte, H. F. (1992). *Directed sonar sensing for mobile robot navigation.* Boston: Kluwer Academic Publisher.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, *19*, 577-621.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, *31*(355-395).

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

Mallot, H., & Franz, M. (1999). Biological approaches to spatial representation: a survey. In T. Dean (Ed.), *Proceedings of the 16th international joint conference on articial intelligence (IJCAI-99).* Morgan Kaufmann.

Marr, D. (1982). *Vision.* San Francisco, CA: W. H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B*, *200*, 269-294.

Mataric, M. J. (1991). Navigating with a rat brain: A neurobiologically-inspired model for robot spatial representation. In J.-A. Meyer & S. Wilson (Eds.), *Proceedings of the first international conference on simulation of adaptive behavior.* Cambridge, MA: MIT Press.

Mel, B. (1997). Combining color, shape and texture histogramming in a neurally-inspired approach to visual object recognition. *Neural Computation*, *14*, 5-24.

Moravec, H. P. (1988). Sensor fusion in certainty grids for mobile robots. *AI Magazine*, *9*(2), 61–74.

Moutarlier, P., & Chatila, R. (1989). Stochastic multisensory data fusion for mobile robot location and environement modelling. In *Isrr.*

Muller, R. U., & Kubie, J. L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience*, *7*, 1935–1950.

Nelson, R. C., & Aloimonos, J. (1989). Obstacle avoidance using flow field divergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11(10)*, 1102–1106.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely moving rat. *Brain Research*, *34*, 171-175.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map.* Clarendon Press.

O'Regan, J., Rensink, J., & Clark, J. (1996). "mud splashes" render picture changes invisible. *Investigative Ophthalmology & Visual Science*, *37:213*, 157-165.

Perrett, D. I., Mistlin, A. J., & Chitty, A. J. (1987). Visual neurones responsive to faces. *Trends in Neurosciences*, *10*(9), 358-364.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, Series B*, *223*, 293-317.

Pinker, S. (Ed.). (1985). *Visual cognition.* Cambridge, MA: MIT Press.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*.

Quine, W. (1977). Natural kinds. In S. Schwartz (Ed.), *Naming, necessity, and natural kinds* (p. 155-175). Cornell University Press.

Rensink, J., O'Regan, J., & Clark, J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373.

Riesenhuber, M., & Poggio, T. (1999). Separate visual pathways for perception and action. *Nature*, *2*(11), 1019–1025.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 516–528.

Schiele, B., & Crowley, J. (1996). Probabilistic object recognition using multidimensional receptive field histograms. In *Proceedings of the 13th international conference on pattern recognition (icpr 96)* (Vol. B, p. 50-54).

Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language. A case study for colour. *Behavioral and Brain Sciences*. (Accepted as target article. Available from http://arti.vub.ac.be/˜tony)

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, *13*, 422–437.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109-140.

Tapus, A., Ramel, G., Dobler, L., & Siegwart, R. (2004). Topology learning and recognition using bayesian programming for mobile robot navigation. In *Proceedings of the 2004 IEEE/RSJ international conference on intelligent robots and systems, iros 2004.*

Tarr, M. J., & Bülthoff, H. H. (1998). *Object recognition in man, monkey, and machine.* MIT Press.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, *21*, 233-282.

Tell, D., & Carlsson, S. (1999). View based visual servoing using epipolar geometry. In *Proc. 11th scandinavian conference on image analysis.*

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.

Thorpe, S., & Gautrais, J. (1998). Rank order coding: A new coding scheme for rapid processing in neural networks. In *Computational neuroscience: Trends in research, j. bower, ed, (plenum press, new york)* (p. 113-118.).

Thrun, S., & Bücken, A. (1996). Integrating grid-based and topological maps for mobile robot navigation. In *Proc. of the national conference on artificial intelligence (aaai-96)* (pp. 944–950). Portland, Oregon, USA.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208.

Trowbridge, C. C. (1913). Fundamental methods of orientation and imaginary maps. *Science*, *38*, 888–897.

Trullier, O., & Meyer, J. (1998). Animat navigation using a cognitive graph. In *From animals to animats 5: Proceedings of the fifth international conference on simulation of adaptive behavior.* Cambridge, MA: MIT Press.

Trullier, O., Wiener, S., Berthoz, A., & Meyer, J. (1997). Biologically-based artificial navigation systems: Review and prospects. *Progress in Neurobiology*, *51*, 483–544.

Turvey, M., & Carello, C. (1981). Cognition: The view from ecological realism. *Cognition*, *10*, 313–321.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A. Frank & I. Campari (Eds.), *Spatial information theory.* Springer.

Ullman, S. (1989). Aligning pictorial descriptions: an approach to object recognition. *Cognition*, *32*, 193-254.

Ullman, S., & Basri, R. (1991). Recognition by linear combinations of models. *IEEE Trans. Pattern Anal. Machine Intell.*, *13*, 992–1006.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, *85*, 223-250.