

Egocentric Shape Representation of Landmark Configurations based on Static and Dynamic Ordering Information

Thomas Wagner

Intelligent Systems Department
University of Bremen

The Center for Computing and Communication Technologies (TZI)

Abstract

In this paper we show (a) how the shape of a 2D-landmark configuration can be encoded based on qualitative 1D-ordering information, (b) how relevant geometric shape properties of a landmark configuration (strictly based on ordering information) can be detected by a sequence of view-based snapshots. Furthermore we show how shape of landmark configurations supports view-based localization tasks specially in the face of erroneous and missing sensor information.

Introduction

Representation of shape plays an important role for natural cognitive systems in localization and navigation tasks ((Tolman 1948) and (Redish 1999)). Both tasks imply specific requirements on the representation of shape and the accessibility of the required inferences. Evidence has been found, especially for the former task, that animals (e.g., mouse (Tolman 1948), fish (Sovrano, Bisazza, and Vallortigara 2002)) as well as humans use geometric properties (e.g., in localization tasks the shape of landmarks configurations (e.g., (McNamara and Sheldon 2003))). A strong motivation for the use of geometric features in contrast (for instance) to more directly accessible visual features is the robustness and stability even under different radical (usually natural) environmental changes (either day/night or seasonal change) (Sovrano, Bisazza, and Vallortigara 2002). The dominant view, propagated from (Tolman 1948) to (Redish 1999)¹ is that spatial representation is encoded in an comprehensive, allocentric representation². This concept gains strong support from the *cognitive map*-research (for an overview again see (Redish 1999)). Nevertheless, more recent research shows that spatial representations are at least to some extent egocentric.

In this paper we show how the shape of a 2D-landmark configuration can be encoded based on qualitative 1D-ordering information and how relevant geometric shape

properties of landmark configurations can be detected by sequences of view-based snapshots, strictly based on 1-D information. Furthermore, we describe how the shape of landmark configurations support view-based localization tasks specially in the face of erroneous and missing sensor information.

The rest of the paper is organized as follows. In the second section we precisely define what we denote with *allocentric* and *egocentric* representations. Additionally, we shortly review the cognitive science literature which addresses the problem of egocentric shape representation (of landmark configurations) and localization based on egocentric spatial representations with respect to update functions. In the third section we shortly describe the state-of-the-art from an AI point of view. In the following section we describe how shape and geometric properties of a landmark configuration can be described in terms of qualitative and quantitative 1D spatial ordering information. In the final section we discuss the given approach and end up with an outlook on future work³.

Motivation

Various classifications for mammal navigation have been proposed (some prominent examples are (Easton and Sholl 1995), (McNamara and Sheldon 2003), (Redish 1999)). A classical classification separates navigation according to five different strategies mammals (especially rodents) choose from in order to reach a given platform (Redish 1999). In this paper we instead rely on the classification proposed by Wang and Spelke (Wang and Spelke 2002). Their classification consists of three different functional modules with a stronger focus on view-dependent (egocentric) navigation in contrast to the former ones⁴. They propose three modules:

path integration: realizes *dynamic update* (e.g., (Simons and Wang 1998), (Wang and Spelke 2000))

place recognition: *"template matching of viewpoint-dependent representations of landmarks"*⁵(e.g., (Mou

Copyright © 2010, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹More precisely, (Redish 1999) does not argue that spatial representation solely relies on allocentric (map-based) representations.

²In cognitive science literature often denoted as a *geocentric* representation.

³Due to page limitations the overall paper is dense/compact. All given *examples* are derived from a formal framework (for details see (Wagner 2006)).

⁴Their article appeared under the section *Opinion in Trends in Cognitive Science*

⁵(Wang and Spelke 2002), pp. 376 left column, bottom

and McNamara 2002))

reorientation system: "operates by congruence-finding on representations of the shape of the surface layout"⁶ (e.g., (Mou, Zhang, and McNamara 2004), (Mou et al. 2004))

Consistent with recent research in cognitive science (e.g., (Easton and Sholl 1995), (Wang and Brockmole 2003), (May 2004)) they propose an egocentric approach to mammal navigation⁷. In this paper we try to follow this general pattern. We describe an approach to egocentric navigation and show how the three modules can be realized based strictly on qualitative and quantitative 1-D spatial information.

Allocentric and egocentric approaches to navigation differ significantly from each other. They focus on different types of spatial relations, i.e., egocentric and allocentric relations. More precisely, egocentric and allocentric relations can be defined like follows. Let Θ denote the set of physical objects and A a cognitive agent with $A \notin \Theta$. Additionally, \mathcal{R}_{SP} denotes the set of spatial relations which are perceived by A_{pra} . A binary relation $R_{ego} \in \mathcal{R}_{SP}$ is an egocentric relation iff $R_{ego} \subseteq A_{pra} \times \alpha$, mit $\alpha \in \Theta$. Contrawise, a binary relation $R_{allo} \in \mathcal{R}_{SP}$ is called allocentric iff $R_{allo} \subseteq \beta \times \alpha$, mit $\beta, \alpha \in \Theta$. Therefore, a spatial representation \mathfrak{R} is called allocentric iff all relation $R_{allo} \in \mathcal{R}_{spa}$ are allocentric (and accordingly for an egocentric representation).

Furthermore, allocentric⁸ and egocentric approaches to navigation differ significantly from an operational point of view. Allocentric spatial representations are usually harder to retrieve due to the transformation process from egocentric sensor inputs into an allocentric representation (more details in the following sections). Although this task can be solved by use of different specialized modules e.g., by specialized cells (*head direction cells*) it still remains a complex task⁹. However, once generated an allocentric representation provides a stable *FoR* for localization tasks (i.e., egocentric as well as allocentric relations can be derived with limited effort). In contrast egocentric representations (i.e., (Schoelkopf and Mallot 1995) developed a concept for (egocentric) snapshots-based localization) can be used directly without any transformation. Instead it requires a continuous update of all relevant (egocentric) spatial relation in case of movement. Although the number of egocentric spatial relation in a given environment is strictly less in contrast to allocentric relations, complexity arises due to the continuous nature of the update process. An additional critical problem is the selection of appropriate *significant* egocentric snapshots and the mapping between different snapshots, e.g., in terms of *path integration*. Although it has been shown that even a simple pixel-based similarity measure can be sufficient in small environments under static conditions it still

⁶(Wang and Spelke 2002), pp. 376 left column, bottom

⁷They do neither claim that map-based navigation does not exist nor that the *cognitive map* is irrelevant in mammal navigation tasks.

⁸In this paper we use the term *allocentric* synonym to *map-based* although both terms have a different focus.

⁹In this paper we do not consider the role of different egocentric frames of reference (FoR). For more details see, e.g., (Pouget et al. 2002).

remains an open question how view-based navigation will under more dynamic conditions. In this paper we propose that a 1-D ordering representation of shape of landmark configurations provides an appropriate frame of reference that allows to give answers to the given problems.

Navigation based on ordering information

The first approach to qualitative navigation based on ordering information has been introduced by Levitt and Lawton (Levitt and Lawton 1990). They propose a genuine qualitative approach to navigation that does not rely on any additional quantitative spatial information and supports navigation strictly based on qualitative information¹⁰

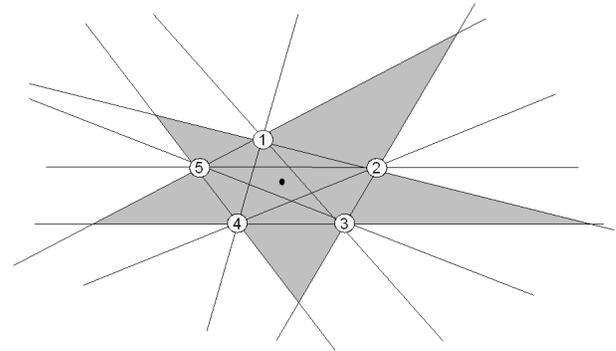


Figure 1: *Localization by Ordering Information (Adopted from Schlieder (Schlieder93))*

They introduced a three level spatial representation with different levels of abstraction. Only on the highest level spatial knowledge was represented strictly qualitative. Based on an omnidirectional sensor, they assumed that the egocentric 1-D ordering of landmarks which results from a round-view is sufficient to localize a robotic agent. Since the ordering does neither relies on metric distance information nor on metric angular information it can be assumed to be rather robust. However, although this idea has a strong intuitive appeal it does not hold in general. Given the position of the robotic agent is indicated by the black dot in figure 1, they assumed that the position can be described by the circular ordering $\dots, L_1, L_2, L_3, L_4, L_5, \dots$ of the landmarks. Figure 1 (adopted from Schlieder (Schlieder 1993)) is a simple counter example for this assumption. Each gray marked region is described by the same circular ordering $\dots, L_1, L_2, L_3, L_4, L_5, \dots$ and can therefore not be distinguished (for a detailed formal analysis see Schlieder). Therefore, the direct use of ordering information as it was proposed by Levitt and Lawton cannot be used for precise, qualitative localization in general.

Schlieder proposed a different spatial representation, the *panorama representation* which overcomes the described problems of Levitt and Lawton¹¹. In Schlieder's approach

¹⁰Nevertheless, their navigation architecture does not solely rely on qualitative navigation.

¹¹The concept of panorama representation has been studied extensively in the course of other domains like specialized sensors

a complete, circular panorama can be described as a 360° view from a specific, observer-dependent point of view. Let P in figure ?? denote a person, then the panorama can be defined by the strict ordering of all objects: *house, woods, mall, lake*. This ordering information, however, does not only contain all ordering information as described by the scenario ?. The *mall* is not only directly between the *woods* and the *lake*, but more specifically between the opposite side of the *house* and the *lake* (indicated by the tails of the arrows). In order to represent the spatial knowledge described in a panorama scenario, Schlieder (Schlieder 1996) introduced a formal model of a panorama.

Definition: (Panorama) Let $\Theta = \{\theta_1, \dots, \theta_n\}$ be a set of points $\theta \in \Theta$ and $\Phi = \{\phi_1, \dots, \phi_n\}$ the arrangement of $n-1$ directed lines connecting θ_i with another point of Θ , then the clockwise oriented cyclical order of Φ is called the panorama of θ_i .

The panorama approach provides a qualitative spatial representation which is strictly based on ordering information that also has been proven to be correct in any environment. Nevertheless the application in practical domains can be a rather difficult. The key problem to be solved is to determine the exact ordering with respect to opposite sides of landmarks. Unfortunately they can never be perceived directly but instead require quite precise quantitative data in order to calculate the ordering, especially in scenarios with many landmarks. In (Wagner 2006) Wagner propose an approach that does not rely on opposite landmarks but is restricted to localization tasks outside the convex hull of landmark configurations. We show how this approach and a quantitative extension in terms of relative (qualitative) speed and acceleration can also be used to describe and distinguish the shape landmark configurations. Additionally, we describe how qualitative shape is able to provide a stable and robust frame of reference for navigation tasks.

Qualitative Shape by Ordering Information

The use of qualitative ordering information is based on two fundamental assumptions:

1. three-dimensional landmarks can be interpreted as abstract(mathematical) points¹²
2. qualitative ordering information can be abstracted from quantitative sensor input

An ordering *snapshot* S is defined as a strictly ordered sequence of point-like objects perceived by a given cognitive agent. The construction process is described in figure 2. Assume a landmark configuration $\langle A, B, C \rangle$ and an observer P_Γ outside the convex hull of $\langle A, B, C \rangle$. Furthermore, assume a viewpoint SP within the convex hull of $\langle A, B, C \rangle$ which can be seen by P_Γ and the orthogonal $L_{Orth}(P_\Gamma/SP)$ of the line of sight from SP to P_Γ . The snapshot S is constructed by the projection of the landmarks A, B and C on

(e.g., omnivision, see, e.g., (Zheng and Tsuji 1992)).
¹²In practice the (re-)identification of three-dimensional landmarks can difficult specially with changing perspectives.

$L_{Orth}(P_\Gamma/SP)$ from the position of P_Γ .¹³

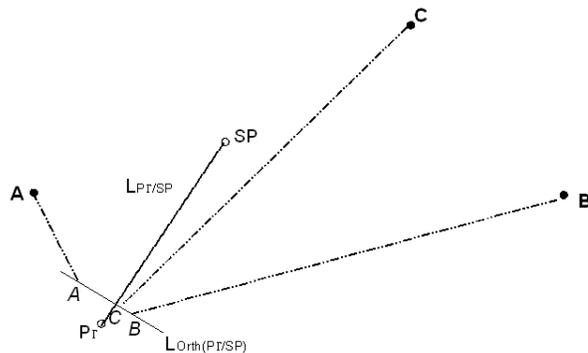


Figure 2: Construction of an ordering snapshot S based on a triple landmark configuration $\langle A, B, C \rangle$

At least three interesting questions are arising:

1. To which extent is the perception of an ordering snapshot determined by the (geometric) shape of the landmarks configuration?
2. Is it possible to define a simple similarity metric for ego-centric snapshots?
3. Can shape be described by sequences of ordering snapshots? - Additionally, can qualitative shape be used to guide the continuous update process?

First, we have to investigate the relations between position, landmark configuration shape and ordering perception. Assume landmark configuration LC_{qrps} with an rectangle layout based on the landmarks q, r, p and s . The relative position of an observer with respect to LC_{qrps} is given by the active ordering snapshot. Assume the position of the observer is given by Γ (indicated by the small circle in figure 3). The resulting ordering snapshot at the given position of Γ is given by $S^{LC_{qrps}} = \langle pqr s \rangle$.

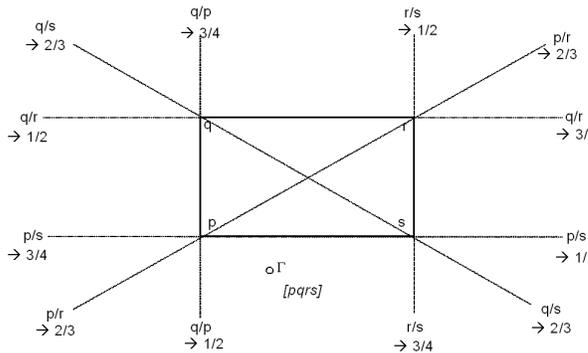


Figure 3: Landmark and position transitions based on a convex rectangle landmark configuration LC_{qrps}

¹³Alternatively, ordering information can be defined in terms of qualitative triangle orientations.

$S^{LC_{qrps}}$ will remain stable until Γ crosses either the line q/p or r/s . In the latter case $S^{LC_{qrps}}$ will change to $S^{LC_{qrps}} = \langle pqsr \rangle$ due to a landmark switch between landmark s and r (by crossing r/s). The circular navigation around LC_{qrps} (outside the convex hull) will therefore result in a sequence of ordering snapshots $S_{Seq} = S_1^{LC_{qrps}}, \dots, S_j^{LC_{qrps}}$.

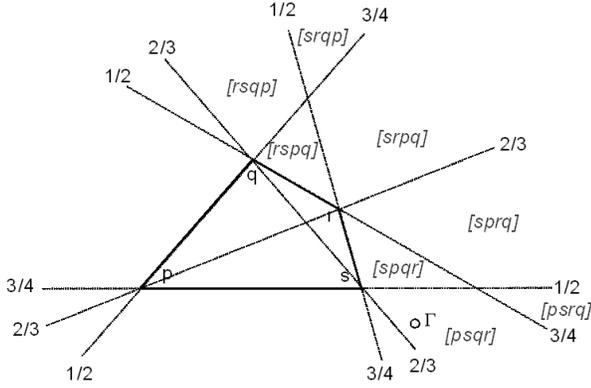


Figure 4: Landmark and position transitions of a non-parallel oriented quad-tuple landmark configuration

How does S_{Seq} depend on the specific shape of the landmark configuration? Figure 4 shows an example with a quad-tuple landmark configuration LC'_{qrps} with different shape and a different ordering sequence $S^{LC'_{qrps}}$. Again assume that the position of the observer (in figure 4) is given by the circular object denoted with Γ . The ordering snapshot at this position is given by $S^{LC'_{qrps}} = \langle psqr \rangle$. In contrast to LC_{qrps} in figure 3 there are two regions where the observer has a choice where to pass the line in (counter-clockwise) navigation. From $S^{LC'_{qrps}}$ observer Γ can either pass the line $3/4$ towards region $\langle spqr \rangle$ or towards region $\langle psrq \rangle$. The latter region denoted by $\langle psrq \rangle$ does not exist in figure 3. The additional region emerges as a result of the intersection of line q/r ($3/4$ in figure 3) with line p/s ($1/2$ in figure 3). Generally, each intersection of lines results in a new region but the possible locations of line intersections are not arbitrary. E.g., it is geometrically impossible to create a new region (intersection) at the opposite side line q/r and line p/s since two straight line can only intersect once¹⁴. In the case of quad-tuple landmark configurations nine different landmark configuration shapes can be distinguished by different S_{Seq} (figure 5).

The full sequence of ordering snapshots already contains redundant information with respect to the classification of different landmark configurations shapes. Even in the case when the observer is not able to distinguish the landmarks, the sequence of landmark positions transitions is already sufficient to determine landmark configuration shape. Given the observer walks counter-clockwise around a landmark

¹⁴At top of figure 3 an additional region has emerged $\langle srqp \rangle$ - For a full formal specification (with proofs and theorems) please refer to (Wagner 2006).

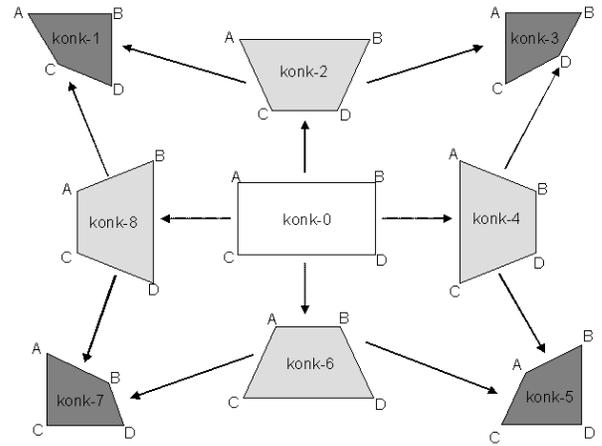


Figure 5: All differing convex landmark configuration shapes that can be distinguished by full sequences of ordering snapshots

configuration and observes three times the following sequence of landmarks transitions (without identification) between position $1/2$, $2/3$ and $3/4$, it is clear that the given landmarks form a parallelogram configuration. The shape of a landmark configuration can even be determined based on a sequence of limited ordering snapshot, i.e., the perception of transitions q/s , p/s , ... (see figure 3 for the full sequence) is also sufficient (in this case we do not need to know the ordinal position of the landmarks under transition, e.g., a switch between the first and second position). The sequence of possible transitions sequences as well as the sequence of all possible position transitions can be calculated in advance (see figure 6).

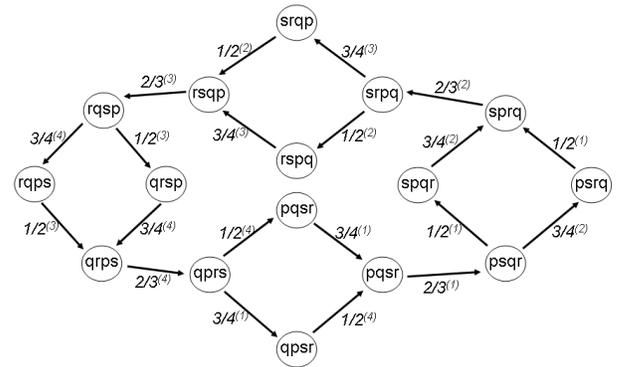


Figure 6: Complete ordering-transition-graph for counter-clockwise navigation

Discussion

In the beginning of the last section we asked three question according to the role and representation of shape (of landmark configurations) based on ordering information:

1. To which extent is the perception of an ordering snapshot

determined by the (geometric) shape of the landmarks configuration?

2. Is it possible to define a simple similarity metric for egocentric snapshots?
3. Can shape be described by a sequence of ordering snapshots? - Can it be used for continuous update?

The first question can be answered precisely: Ordering information, i.e., the egocentric perceived ordering of landmarks (*egocentric snapshot*) is tightly coupled to the shape of a given landmark configuration. Figure 4 describes all possible *egocentric snapshots* for every possible convex quad-tuple landmark configuration that can be distinguished strictly based on ordering information. Every possible variation is precisely (pre-)defined.

The answer to the second question is more complex. First, an *egocentric snapshot* implicitly defines a similarity measure in terms of the number equal ordering relations between two egocentric snapshots (e.g., number of equal *leftOf*-relations). Problems can arise since the number of equal relations does not directly contribute to metric distances, e.g., in figure 3 the distance between different regions/transitions can differ significantly and will therefore not always lead to an intuitive distance estimation. Notwithstanding, ordering provides a more stable measure than e.g., pixel similarity (see (Schoelkopf and Mallot 1995)). It has been shown that ordering perception strictly depends on the shape of the landmark configuration and therefore, it is less sensitive to environmental changes like lighting conditions or seasonal changes (in natural environments).

In turn, the answer to questions three is once again clear. Figure 4 describes precisely (also) all possible transitions either based on landmark identification or based on the given positional changes of landmarks (without identification). An advantage of ordering information is that it does not require a *real* continuous update since the changes are discrete and do not appear immediately. In addition, all transitions are known in advance with respect to the shape of the landmark configuration and the direction of navigation of the cognitive agent.

Although we believe that ordering information may play a crucial role in egocentric shape representation we also believe that other spatial information will also play an important role. E.g., in the process of ordering perception almost every sensor will additionally perceive the direction of landmark movement (either towards- or away from each other). Furthermore, absolute as well as relative speed of landmark movement in addition to acceleration (qualitative and quantitative) will allow for a more precise description of landmark configurations shape. Moreover this information can easily be combined with ordering snapshots.

References

- Easton, R., and Sholl, M. 1995. Object-Array Structure, Frames of Reference, and Retrieval of Spatial Knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition* Vol. 21(2):pp. 483–500.
- Levitt, T. S., and Lawton, D. T. 1990. Qualitative Navigation for Mobile Robots. *Artificial Intelligence* Vol. 44(3):pp. 305–360.
- May, M. 2004. Imaginal Perspective switches in remembered Environments: Transformation vs. Interference Accounts. *Cognitive Psychology* Vol. 48:pp. 163–206.
- McNamara, T., and Sheldon, A. 2003. Cognitive Maps and the Hippocampus. *Trends in Cognitive Science* Vol. 7(8):pp. 333–335.
- Mou, W., and McNamara, T. 2002. Intrinsic Frames of Reference in Spatial Memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* Vol. 28(1):pp. 162–170.
- Mou, W.; McNamara, T.; Valiquette, C.; and Rump, B. 2004. Allocentric and Egocentric Updating of Spatial Memories. *Journal of Experimental Psychology: Learning, Memory and Cognition* Vol. 30(1):pp. 142–157.
- Mou, W.; Zhang, K.; and McNamara, T. 2004. Frames of Reference in Spatial Memories Acquired from Language. *Journal of Experimental Psychology: Learning, Memory and Cognition* Vol. 30(1):pp. 171–180.
- Pouget, A.; Ducon, J.-C.; Torri, J.; and Bavelier, D. 2002. Multisensory Spatial Representations in Eye-Centered Coordinates for Reaching. *Cognition* Vol. 83:pp. B1–B11.
- Redish, A. 1999. *Beyond the Cognitive Map - From Place Cells to Episodic Memory*. The MIT Press, Cambridge, Mass.
- Schlieder, C. 1993. Representing Visible Locations for Qualitative Navigation. In: Qualitative Reasoning and Decision Technologies, N. Piera-Carrete & M. Singh (Eds.) CIMNE Barcelona. pp. 523–532.
- Schlieder, C. 1996. *Intelligent Image Database Systems*. Singapore: World Scientific Publishing. chapter Ordering Information and Symbolic Projection, pp. 115–140.
- Schoelkopf, B., and Mallot, H. A. 1995. View-Based Mapping and Path Planning. *Adaptive Behavior* Vol. 3(3):pp. 311–348.
- Simons, D., and Wang, R. 1998. Perceiving Real-World Viewpoint Changes. *Psychological Science* Vol. 9(4):pp. 315–321.
- Sovrano, V.; Bisazza, A.; and Vallortigara, G. 2002. Modularity and Spatial Reorientation in a Simple Mind: Encoding of Geometric and Nongeometric Properties of a Spatial Environment by Fish. *Cognition* Vol. 85:pp. B51–B59.
- Tolman, E. 1948. Cognitive Maps in Rats and Men. *Psychological Review* Vol. 55:pp. 189–208.
- Wagner, T. 2006. *Qualitative sicht-basierte Navigation in unstrukturierten Umgebungen*. Ph.D. Dissertation, Universitaet Bremen, TZI.
- Wang, R., and Brockmole, J. 2003. Human Navigation in Nested Environment. *Journal of Experimental Psychology: Learning, Memory and Cognition* Vol. 29(3):pp. 398–404.
- Wang, R., and Spelke, E. 2000. Updating Egocentric Representations in Human Navigation. *Cognition* Vol. 77:pp. 215–250.

Wang, R., and Spelke, E. 2002. Human Spatial Representation: Insights from Animals. *Trends in Cognitive Science* Vol. 6(9):pp. 176–182.

Zheng, J., and Tsuji, S. 1992. Panoramic pp. Representation for Route Recognition by a Mobile Robot. *International Journal of Computer Vision* Vol. 9(1):pp. 55–76.