Egocentric Qualitative Spatial Knowledge Representation for Physical Robots

Thomas Wagner
Center for
Computing Technologies (TZI)
University of Bremen
twagner@tzi.de

Ubbo Visser
Center for
Computing Technologies (TZI)
University of Bremen
visser@tzi.de

Otthein Herzog
Center for
Computing Technologies (TZI)
University of Bremen
herzog@tzi.de

Abstract

Although recent (physical) robots have powerful sensors and actuators their abilities to show intelligent behavior is often limited. One key reason is the lack of an appropriate spatial representation. Spatial knowledge plays a crucial role in navigation, (self- and object-)localization, planning and reasoning for physical grounded robots. However, it is a major difficulty of most existing approaches that each of these tasks imposes heterogeneous requirements on the representation. In this paper, we propose an egocentric representation, which relies on 1-D ordering information that still provides sufficient allocentric information to solve navigation and (self- and object) localization tasks. Furthermore, we claim that our approach supports an efficient, incremental process based on a simple 1D-representation. We conclude with a more abstract qualitative spatial representation.

Introduction

One major difficulty is that each of the navigation and localization tasks imposes heterogeneous requirements on the representation. While the purpose of the navigation- and localization process is to build up a precise, in the majority of cases allocentric static representation, planning and reasoning require more abstract, often egocentric representations. In this paper, we propose an egocentric representation which relies on 1-D ordering information, that provides sufficient allocentric information to solve the mentioned tasks: an *egocentric, extended panorama*. Furthermore, we claim that our approach supports an efficient, incremental process based on a simple 1D-representation for dynamic self-localization tasks and conclude with a more abstract qualitative spatial representation.

While qualitative knowledge representation was one of the central topics of AI research almost two decades ago, qualitative spatial knowledge representation has gained strong interest especially in the last decade resulting in many promising approaches which have been used successfully in several areas of applications (e.g., traffic monitoring, geographical information systems (Fernyhough and Hogg 2000), (Cohn and Hazarika 2001)). However, these approaches have been rarely used in robotics.

It was shown that, based on a precise allocentric representation, physical environments can be adequately described in terms of qualitative spatial knowledge i.e., topological, metric and ordinal descriptions (Kuipers 2000), (Yeap and Jefferies 1999). For all these types of qualitative knowledge, expressive representations with powerful inference mechanisms have been developed (popular examples can be found in: for metric representations (e.g., (Clementini *et al.* 1997)), for topological representations (e.g., (Renz and Nebel 1999)) and for ordinal representations (e.g. (Frank 1996))). However, the requirements for dynamic environments based on egocentric input data differ significantly from those in allocentric environments with almost perfect information.

Motivated from recent results in the cognitive science (see section „*Related Work*" for details) we propose in this paper a flexible, egocentric spatial representation, which is based on ordering information: an *egocentric, extended panorama* (EEP). We argue that the proposed sets of EEP provide a reliable basis for re-orientation and are expressive enough to represent a large set of relevant spatial descriptions at different levels of granularity, which are required to describe complex, coordinated behavior in real-time environments. Additionally, we motivate that an EEP can be maintained efficiently and in a flexible manner.

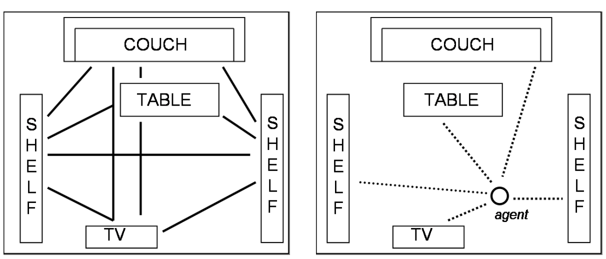
The rest of the paper is organized as follows: in the next section we investigate the problem and the resulting requirements in more detail. In section three we have a look on the previous work, focusing on cognition and panorama representations. In section four we finally introduce our *extended panorama* model. We discuss advantages and disadvantages of our approach and conclude with a brief a outlook with respect to future work.

Motivation

Modeling complex behavior imposes strong requirements on the underlying representations. The representation should provide several levels of abstraction for activities as well as for objects. For both types of knowledge, different representations were proposed and it was demonstrated that they can be used successfully. Activities can, e.g., be described adequately with hierarchical task networks (HTN) which provide clear formal semantics as well as powerful, efficient (planning-) inferences (see e.g. (Erol *et al.* 1994)). Objects can be described either in ontology-based languages (e.g., OWL (Smith *et al.* 2003)) or constraint-based languages (e.g., (John 2000)). Both types of representations allow for the representation of knowledge at different levels of abstraction according to the specific requirements of the domain/task. In physically grounded environments, the use of these techniques requires an appropriate qualitative spatial description in order to relate the modeled behavior to the real world. In the following three subsections we investigate the key problems and as a consequence, the requirements on qualitative spatial representations in order to model complex physically grounded behaviors. As an example scenario we use the *RoboCup* domain.

Allocentric and Egocentric Representations

During the modeling of behavior patterns we have to decide whether we would like to use an egocentric or an allocentric representation. In an egocentric representation, spatial relations are usually directly related to an agent by the use of an egocentric *frame of reference* in terms like, e.g., *left, right, in front, behind*. As a consequence, when an agent moves through an environment, all spatial relations need to be updated. On the other hand, representations based on an allocentric frame of reference remain stable but are much harder to acquire. Additionally, the number of spatial relations which have to be taken into account may be much larger because we have to consider the relations between each object α and all other objects in the environment, whereas the number of relations in egocentric representations can be significantly smaller.



(a) Allocentric relations (b) Egocentric relations

Figure 1: Allocentric vs. egocentric spatial relations

The decision whether to use an allocentric or an egocentric frame of reference has to be made with respect to the additional domain-specific aspects. An interesting phenomenon, when looking into the didactic literature about, e.g., sports (Lucchesi 2001) we often find that (strategic and tactical) knowledge is described in both, egocentric and allocentric terms, whereas, e.g., the literature about driving lessons strongly relies on purely egocentric views. At least one of the reasons are that the latter representation seems to provide better support for acting directly in physically grounded environments, since perception as well as the use of actuators are directly based on egocentric representations. In addition, egocentric representations provide better support for rotation and translation invariant representations when used with a qualitative abstraction (see the next section for more details). A clear disadvantage is that an egocentric spatial descriptions $SpDes_{\Theta_1}$ of an agent Θ_1 cannot be communicated to and used by another agent Θ_2 either without a transformation from $SpDes_{\Theta_1}$ into $SpDes_{\Theta_2}$ or the use of an intermediate allocentric model, which provides a common frame of reference.

Although the recent literature about spatial reasoning seems to favor allocentric representations, we claim, that the computational model of an *egocentric extended panorama*, which is mainly based on an egocentric representation, provides clear advantages.

Related Work

Our approach is motivated by both, cognitive aspects and AI-based representations. Therefore, we divide this section in two subsections accordingly.

Cognition: Dynamic, Egocentric Spatial Representations

The fact that even many animals (e.g., rodents) are able to find new paths to familiar objects seems to suggest that spatial relations are encoded in an allocentric static *„cognitive map"*. This almost traditional thesis is supported by many spatial abilities like map navigation and mental movement humans are able to perform (many literature about this topic has been published beginning with (Tolman 1948); another popular example is: (OKeefe and Speakman 1978)). Wang and Spelke (Wang and Spelke 2002) argue that these observations and results can be explained in two different ways, either by updating the egocentric movement in an allocentric map using movement information, or in contrary, that humans build a complete egocentric representation and are able to map between different views. Recent studies provide strong evidence for the latter thesis. The predictions of both do not differ essentially, they assume that allocentric information is crucial for many spatial tasks and will therefore be provided in either way. The underlying assumption of the sophisticated series of experiments[1] done by Wang (Wang 2000) and (Wang and Spelke 2000) was that the relations remain stable in an allocentric, *cognitive map* independent from egocentric movements, i.e., although the process of path integration is error-prone, the allocentric relations will not change, given that they are based on a cognitive map. When errors arise because of path integrations, the error rate (*configuration error*) should be the same for all allocentric relations; otherwise they rely on an egocentric representation. The results indicate clear evidence for egocentric representations. These results have been confirmed in a series of differently designed experiments e.g. (Roskos-Ewoldsen *et al.* 1998), (Sholl and Nolin 1997), (Easton and Sholl 1995) and (Garsoffky *et al.* 2002).

[1] Most experiments choose a spatial updating task which refers to *„the ability of a moving person to mentally update the location of a target initially seen, heard, or touched from a stationary point."* (in (Loomis *et al.* 2002) pp.335 c.).

AI: Dynamic, Egocentric Spatial Representations

A lot of approaches of spatial reasoning focus on the representation of large scale space. Large scale space can for instance be defined as *„space (that) is a significantly larger scale than the observations available at an instant"* (Kuipers and Levitt 1988). Based on these complex representations, different powerful inference methods have been developed which allow for reasoning on cardinal directions (Frank 1996), distance representations (e.g., (Clementini *et al.* 1997)) and topological representations (e.g., (Cohn *et al.* 1997)) on qualitative spatial representations. Most of these approaches, however, are based on allocentric views with precise information.

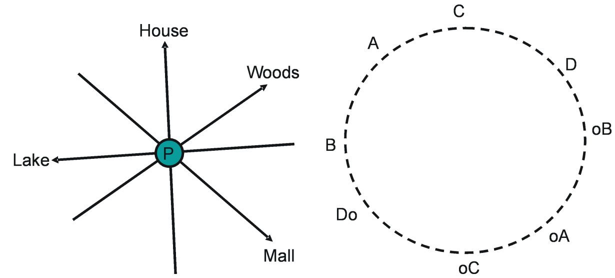
As a result, knowledge-based systems *with* spatial reasoning focus on domains like geographical information systems (GIS). When addressing domains with egocentric views, ranging from sports (e.g., soccer, football) to traffic, the lack of a precise, allocentric world model makes it difficult to apply these approaches. Instead we have to combine missing as well as uncertain knowledge. In order to build more abstract qualitative representations on different levels of granularity, we first have to handle these problems. Instead of building an allocentric world model, we propose an approach that directly relies on an egocentric (and therefore dynamic) representation. Schlieder ((Schlieder 1996), (Schlieder 1993)) proposed the *panorama* approach to qualitative spatial reasoning. This approach is based on a simple qualitative abstraction: the use of 1D-information (i.e., ordering information) in order to represent egocentric 2D-scenarios.

The Panorama Approach The concept of panorama representation has been studied extensively in the course of specialized sensors (e.g., omnivision, see, e.g., (Zheng and Tsuji 1992)). We present an extended approach based on the panorama approach.

A complete, circular panorama can be described as a 360^0 view from a specific, observer-dependent point of view. Let *P* in figure 2(a) denote a person, then the panorama can be defined as the strict ordering of all objects: *house, woods, mall, lake*. This ordering, however, does not contain all ordering information as described by the scenario 2(a). 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* (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 1: (Panorama) Let $\Theta = \{\theta_1, \ldots, \theta_2\}$ be a set of points $\theta \in \Theta$ and $\Phi = \{\phi_1, \ldots, \phi_n\}$ 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.

As a compact shorthand notation we can describe the panorama in figure 2(b) as the string $< \; A, C, D, Bo, Ao, Co, Do, B \; >$. Standard letters (e.g., A) describe reference points, and letters with a following *o* (e.g., Ao) the opposite side (the tail side). Because of the cyclic structure of the panorama the complete panorama has to be described by n strings with n letters, with n being the number of reference points on the panorama. In our example, the panorama has to be described by eight strings. Furthermore, the panorama can be described as a set of simple constraints $dl(vp, lm_1, lm_2)$[2]. Based on this representation, Schlieder (Schlieder 1993) also developed an efficient qualitative navigation algorithm.



(a) Concrete panorama (b) Abstract panorama

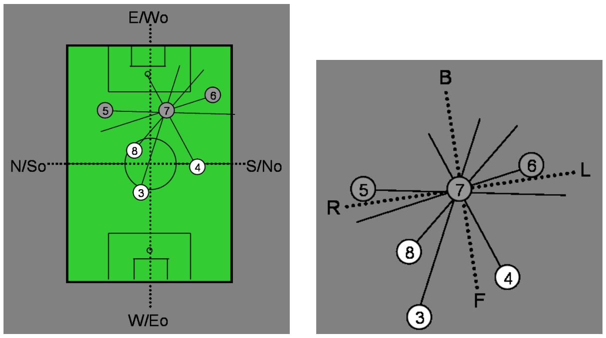
Figure 2: Panorama-views

Applying the panorama representation to competitive and cooperative scenarios for instance, we can infer groups of competitive and/or cooperative agents, given that we are able to distinguish them physically, i.e the basic panorama has to be differentiated slightly to $P_{basic} =< O_{pponent}, O_{wn}, OC_{dl} >$, with $O_{pponent} \in P$ and $O_{wn} \in OP$ the set of observed points and OC_{dl} a set of *direct-left*-ordering constraints.

The panorama representation has an additional, more important property: one of the major difficulties in planning complex coordinated behaviors in dynamic, physical environments is to identify stable situations. As a consequence, we would like to describe behavior on different levels of granularity according to the time scale and we want to identify situations that are similar enough to apply planned behavior. Therefore, spatial representations should abstract from irrelevant details. One way to achieve this is to use a representation which is invariant with respect to rotation and translation, like the panorama. Applied to the *RoboCup* domain, a translation- and rotation-invariant representation would allow to describe a behavior that is independent from the (exact) location of an agent on the field and that is on the other hand invariant to the orientation of the agent. Figure shows a situation with a concrete attack of the white team. Describing behavior for this situation relies on the orientation of the configuration of observed points but does not rely on the specific position according to the length of the field, i.e., even if the same configuration is found ten meters behind the current position, the same models of behavior can be applied. A more appropriate representation for this scenario is described in figure 3(b).

But evidently, not every behavior can be described in such an abstract manner. In order to model complex, coordinated behaviors, often more detailed ordinal information is involved. Additionally, different metric (e.g. distance) in-

[2] Short for $direct{-}left(viewpoint, landmark_1, landmark_2)$.



(a) Allo-centric view (b) Ego-centric, invariant View

Figure 3: Pseudo-ordinal information

formation is required in some situations. In the following section, we show how the panorama can be extended so that more detailed ordinal and metric information can be introduced.

Extended Panorama Representations

In order to use an egocentric spatial representation some inherent problems have to be solved: how can an agent update his position without the use of an allocentric map? How can spatially grounded egocentric knowledge be communicated and exchanged so that coordinated behavior becomes possible? The answers to these questions have to be given with respect to some basic conditions:

1. Updating has to be efficient since egocentric spatial relations change with every movement, i.e., the updating process itself and the underlying vision process.
2. The resulting representation should provide the basis for qualitative spatial descriptions at different levels of granularity.
3. The resulting representation should provide different levels of abstraction, i.e., rotation and/or translation invariance.
4. The process of mapping egocentric views should rely on a minimum of allocentric, external information.

Our answers to these questions will be given in two parts. In the first section we investigate the questions for agents moving between landmarks and in the second section we consider the case of agents moving within landmarks (e.g., rooms). In the last subsection we show how this approach can be extended according to different levels of granularity.

An Egocentric Frame of Reference

The most direct way to extend the panorama representation is to introduce an egocentric frame of reference. Therefore, we introduce the heading of an agent into the panorama (assuming that an agent at least knows his relative orientation according to his individual sensors). Defining the heading as *front*, we can easily introduce the *back/behind* side as the opposite direction and furthermore left and right as a 90^0 angle distance from the heading. Figure 3(b) shows an example leading us to the following *extended panorama* for player 7: $< 4, 5o, L, 6, 8o, 3o, B, 4o, 5, R, 6o, 8, 3, F >$. Although this representation does not provide exact ordinal information, it allows to infer an ordinal interval for a given landmark θ_1 according to its ego-centric orientation $OrdEgo_4 = \{left, right, front, behind\}^3$. In our example, we know that player 5 and the opposite of player 4 are between the back- and right side. Therefore, a panorama with an egocentric frame of reference can be defined as $P_{ERP} = \{DLM, Ref_{ego}, OC\}$ (ERP) with DLM as a set of possibly *dynamic* landmarks, an ordinal ego-centric reference system Ref_{ego} and a set of ordering constraints OC. Since this representation integrates only additional ordinal information, it is still invariant according to translation and rotation and hence this representation is abstract enough to provide spatial information which can be applied for numerous different situations[4]. In the following, we assume that each agent uses an egocentric frame of reference either within and outside a landmark.

Outside Landmarks: (Re-)Orientation

The use of egocentric representations relies on an efficient way to update spatial relations. In contrast to allocentric representations, spatial relations in egocentric representations have to be updated not only because of changes in the environment but also because of the movement of the agent itself. Therefore, the use of egocentric representations strongly relies on an efficient mechanism for path integration and re-orientation, i.e., to detect changes in relation to a previously observed view. Since the exchange of spatial information is a crucial requirement in multiagent systems, we need a mechanism to transform egocentric representations into each other. Motivated by cognitive observations (see section), we propose an incremental approach that uses geometric (and therefore allocentric) properties of given landmarks in a qualitative panorama representation. The resulting representation builds the foundation for re-orientation and path integration as well as for communicating spatial knowledge.

In a re-orientation task we can resort the knowledge about the previous position of an agent. Therefore, we concentrate on an incrementally updating process, based on the following assumptions:

1. It is known that the configuration of perceived landmarks $A, B, ... \in L$ either form a triangle- or a parallelogram configuration (e.g. either by vision or by use of background knowledge).
2. The positions P_{t-1} of an agent A in relation to L at time step $t-1$ is known.

The panorama landmark representation (LP_T) of a triangle configuration can be defined as follows:

[3] Different egocentric reference systems may be used at different granularities according to the requirements of the domain.

[4] The more fine-grained the introduced egocentric frame of reference is, e.g., 8 instead of 4 directions, the more precise the representation, but without any new quality (sensory) information.

Definition 2: (Triangle Landmark Panorama) Let P_A denote the position of an agent A and $C_{T(ABC)}$ the triangle configuration formed by the set of points A, B, C in the plane. The line $L_{P_A/VP}$ is the line of vision from P_A to *VP*, with *VP* being a fixed point within $C_{T(ABC)}$. Furthermore, $L_{Orth(P_A/VP)}$ be the orthogonal intersection of $L_{P_A/VP}$. The panoramic ordering information can be described by the orthogonal projection $P(P_A, VP, C_{T(ABC)})$ of the points ABC onto $L_{Orth(P_A/VP)}$.

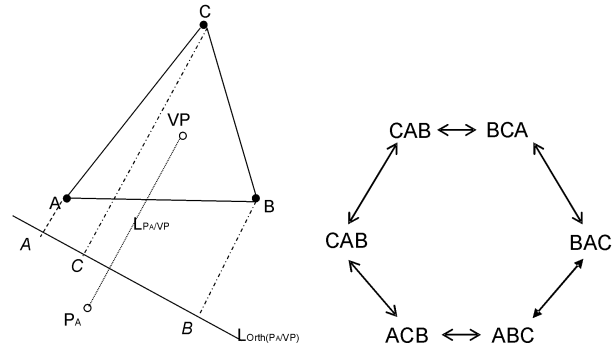
Therefore, moving around a triangle configuration $C_{T(ABC)}$ results in a sequence of panoramas which qualitatively describes the location of the observer position. A 360^o movement can be distinguished in six different qualitative states:

Observation 1: (Triangle Landmark Panorama Cycle) The landmark panoramic representations resulting from the subsequent projection $P(P_A, VP, C_{T(ABC)})$ by counter-clockwise circular movement around *VP* can be described by the following ordered, circular sequence of panoramas: $(CAB), (ACB), (ABC), (BAC), (BCA), (CBA)$

When we observe the panorama (BAC) of a triangle configuration $C_{T(ABC)}$ and the agent is moving counter-clockwise, the next panorama to appear is (BCA). The ordered sequence is circular in the sense that when the agent moves counter-clockwise, starting at (CBA) the next panorama is the first in observation 4.1: (CAB). The diagram 4(b) illustrates this property. Moving in the other direction just reverses the ordering in which the landmark panoramas will appear.

These properties hold for all triangle configurations: For each landmark panorama the landmark panorama directly left as well as at the right differ in exact two positions that are lying next to each other (e.g., $(ABC), (BAC)$ differ in the position exchange between A and B. These position changes occur exactly when the vision line $L_{P_A/VP}$ intersects the extension of one of the three triangle lines: L_{AB}, L_{AC}, L_{BC}. Starting with a given line (e.g., L_{AB}) and moving either clock- or counter-clockwise, the ordering of line extensions to be crossed is fixed for any triangle configuration. This property holds in general for triangle configurations but not e.g. for quadrangle configurations (except for - for some special cases as we will see in the next subsections). Since (almost) each triplet of landmarks can be interpreted as a triangle configuration, this form of qualitative self-localization can be applied quite flexibly with respect to domain-specific landmarks. The triangle landmark panorama, however, has (at least) two weaknesses: The qualitative classification of an agent's position into six areas is quite coarse and, triangle configurations are somewhat artificial constructs that are rarely found in natural environments when we consider solid objects[5]. A natural extension seems to be applying the same idea to quadrangles.

[5]The triangle configuration can be applied generally to any triple of points that form a triangle - also to solid objects. The connecting lines pictured in the graphics 4(a) and 5(a) are used to explain the underlying concept of position exchange (transition)



(a) Triangle panorama construction by projection (result here: (ACB))

(b) Circular representation of panoramic ordering information for triangles

Figure 4: The triangle panorama

The most direct approach is to interpret a quadrangle as a set of two connected triangles sharing two points by a common line so that each quadrangle would be described by a set of two triangle panoramas (e.g. ((BCA)(CDA))). With this approach, the space around a quadrangle would be separated into ten areas and it would therefore be more expressive than the simple triangle panorama. It can be shown that eight of the resulting triangle landmark panorama (one for each triangle of the quadrangle) can be transformed into quadruples. We simply transform e.g. a rectangle into a landmark panorama representation (e.g. the above given tuple ((BCA)(CDA)) can be transformed into (BCDA) without loss of information). The expressiveness of the other two landmark panoramas is weaker: they have to be described as a disjunction of two quadruple tuples[6]. Since the expressiveness is weaker and the landmark panorama representation of a quadruple tuple panorama representation is much more intuitive we focus on the latter.

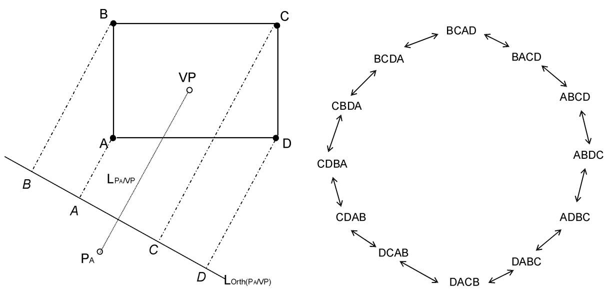
Definition 3: (Parallelogram Landmark Panorama) Let P_A denote the position of an agent A and $C_{P(ABC)}$ the parallelogram configuration formed by the set of points A, B, C, D the plane. The line $L_{P_A/VP}$ is the line of vision from P_A to *VP*, with *VP* being a fixed point within $C_{P(ABCD)}$. Furthermore $L_{Orth(P_A/VP)}$ be the orthogonal intersection of $L_{P_A/VP}$. The landmark panoramic ordering information can then be described by

[6]The detailed proof will take too much space here. But the basic proof idea is quite straightforward: each panorama transition happens because of the intersection of the landmarks' line extensions with the line of vision of the moving agent, so the number of disjoint lines (multiplication by 2, since each line is intersected twice) specifies the number of transitions and therefore the number of distinguishable areas. The loss of expressiveness of two of the triangle tuples can be explained in the same way: assume that the quadrangle $ABCD$ is defined by the two triangles ABC and ADC sharing the diagonal AC. Position changes of the points B/D cannot be distinguished since they happen in two different triangles, which are not in relation to each other. Alternatively, we can show that the number of resulting ordering constraints is smaller.

the orthogonal projection $P(P_A, VP, C_{P(ABCD)})$ of the points $ABCD$ onto $L_{Orth(P_A/VP)}$.

Moving around a parallelogram configuration $C_{P(ABCD)}$ also results in a sequence of landmark panoramas which describes the location of the observer position qualitatively. A 360^o movement can be split into twelve different states:

Observation 2: (Parallelogram Panorama Cycle) The panoramic landmark representations resulting from the subsequent projection $P(P_A, VP, C_{P(ABCD)})$ by counter-clockwise circular movement around *VP* can be described by the following ordered, circular sequence of panoramas: $((BCAD), (BACD), (ABCD), (ABDC), (ADBC), (DABC), (DACB), (DCAB), (CDAB), (CDBA), (CBDA), (BCAD))$



(a) Parallelogram panorama construction by projection (result here: (BACD))

(b) Circular representation of panoramic ordering information for parallelograms

Figure 5: Parallelogram panorama

The two presented landmark panoramas can be mapped quite flexibly onto landmarks that can be found in natural environments. While solid objects often form rectangle configurations, irregular landmarks can be used in combination as a triangle configuration, since this approach is not strictly restricted to point-like objects. An interesting extension is to build up more complex representations by using landmark configurations as single points in larger landmark configurations. This allows to construct nesting representations which support different levels of granularity according to the requirements of the domain (e.g., in the traffic domain: location relative to a single house vs. location relative to a complete block).

Generally, the panorama representation, which results from monitoring landmarks can be useful for applications in several ways: in a re-orientation task an agent knows at least to some extent where it has been. Based on this information, the circular panorama landmark representation can tell us which hypotheses are plausible according to previous information. Given the last landmark panorama according to a triangle configuration Tr_{ABC} was (ACB) and the currently perceived landmark panorama seems to be (BCA). We can conclude that this perception has a high probability to be wrong. Without taking odometry data into consideration, we know that the agent would have missed two panorama landmarks: (ABC) and (BAC)- which is highly implausible for most scenarios. Panorama landmark information may also be used in exactly the other direction, in order to validate odometry data. Furthermore, the landmark panorama can help to focus perception in a qualitative self-allocation task. In the transition of one panorama landmark into another exactly one position change is performed. Given the landmark panorama (ABC) was seen in the triangle configuration Tr_{ABC} and the agent is moving clockwise it is known that BC will change position. Therefore, the perception of A is without any use for updating the qualitative position of the agent. Furthermore, the panorama landmark representation is not only useful for position updating but also for re-orientation without knowledge about the previous position. Just the perception of the partial landmark panorama AB of a triangle configuration Tr_{ABC} provides us with two hypotheses about the current position: (CAB) and (ABC). In order to validate which hypothesis holds we just have to find out whether C appears on the left side of A or on the right side of B.

Finally we have to mention that a landmark panorama provides a stable basis for qualitative, ordinal, spatial descriptions (e.g. left of, right of), since it is, clearly, sensitive to rotation but invariant to transition.

Within Landmarks: Spatial Frames of Reference

Moving within spatial landmarks is interpreted in a specific way as moving within a spatially restricted area, e.g., a room or a building. The properties of ordering information, i.e., panorama representation, differ significantly from the case where an agent is moving outside of landmark areas. While in the latter case the landmark panorama changes in accordance with the agents movement, the panoramic, ordering information remains stable in spite of a change of viewpoint caused by movement. This allows us to interpret the configuration of points as domain-specific, allocentric frames of reference (FoR). In the soccer domain, different landmarks can be used intuitively do define a rectangle configuration, e.g., our own or the opposite goal, the left or right intersection point of the mid-line with the outer line[7]. In egocentric views, the reference systems can be easily defined in terms of physical properties of the agent itself. Since the reference system is not based on perceptional input, it can be defined in any way as long as it preserves a 90^0 distance between the basic four directions. Since the angular distances between the marks of an egocentric view have equal distance by definition (e.g., 90^0 distance between front and left, behind and right), it is not necessary to explicitly represent them. For domain-specific allocentric frames we generally lose information because the angular distance depends on the properties of the domain. Choosing an appropriate FoR should be done carefully since it is necessary to perceive at least two reference points to retrieve information about the current position based strictly on ordering information. Referring to the example in figure 3(a), we describe the pseudo-allocentric ordinal panorama for player 7 as $< 4o, N, 3o, 8o, 6, E, 4, S, 8, W, 6o, 5 >$ while using our own and the opponent goal as fixed landmarks. Given that

[7] In the physical RoboCup leagues: SSL, MSL and *Sony Legged League* there exists a fixed set of labeled landmarks.

we have chosen the opponent half of the field as the desired FoR, the above described problem would have arisen.

Although a domain-specific allocentric panorama does not provide a real allocentric view, it nevertheless provides the foundation for communication about spatial content and therefore for coordinated behavior. On this basis, player 7 can tell e.g. another player where he wants to pass a ball or where he expects a player to move. Furthermore, together with an egocentric FoR, an allocentric FoR significantly increases the precision of the ordering information just by introducing additional landmarks.

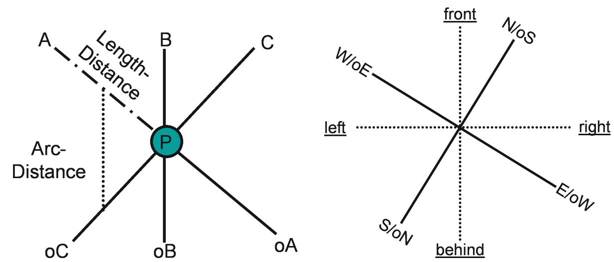
Seeing some points of a chosen FoR provides an agent usually only with coarse information about the position relatively to the used FoR. Much better results can be generated when the panorama representation is enhanced with quantitative or qualitative angle information[8].

Extending Panorama Representations

Egocentric and allocentric panoramas describe *ordinal intervals* (e.g. player 4 is between east and south), increasing the numbers of reference points minimizes the *ordinal intervals* and therefore increase ordinal precision. This approach is somehow limited because all fixed reference points have to be defined in a domain-dependent way. Therefore, it can be difficult to define a large set appropriate fixed reference points, especially, since reference points should be easy to perceive from different locations for all agents that participate in a coordinated behavior. Furthermore, they should have a similar distance between each other in order to provide a uniform granularity.

Given the case that fixed reference points cannot be defined, absolute angle distance between two possibly dynamic points (i.e. agents) can be used to increase ordinal precision instead (see figure). Angle distance should be mapped on an absolute qualitative scale with equal distance. Given a fixed angle distance of 30^0 we are able to distinguish the location of an object, e.g. *north/west(NW)* and *west/north(WN)* instead of somewhere between *north* and *west*. If we are able to distinguish between 30^0 angle distances, we can even classify, e.g., between *W, WNW, NW, NNW, N*. The choice between the two approaches depends on the characteristics of domains and on the quality of the perceptional input. As both approaches result in the same panorama representation, it is even possible to use both approaches simultaneously depending on the specific situation. Another way to adapt an *extended panorama* to specific situations can be achieved by using different reference systems on different levels of granularity. For more abstract planning phases, e.g., in HTN-planning or in situations with imprecise information, more coarse information can be handled with restricted reference systems. On the other hand, in detailed planning phases or in situations where more detailed information is required, a more finely grained reference system can be chosen.

[8]This information will, in most cases, be generated anyway by the vision process during the calculation of the panoramic representation.



(a) Distances in a panorama (b) A hybrid approach

Figure 6: Different combinations of EPO- and PAPO-panoramas

Depending on the specific domain and the specific situation, different variations and combinations of panoramic representations can be built.

Discussion and Future Work

Navigation, localization, planning and reasoning for physical grounded robots imposes strong but heterogeneous requirements on the underlying spatial representation in terms of abstraction and precision. One crucial property is the dependency on allocentric information, e.g., to communicate knowledge and to coordinate behaviors. In contrast to many other approaches to this topic which try to generate *allocentric* maps, we proposed a new *egocentric* approach based on recent results from cognition. Allocentric information is retrieved by a the qualitative description of landmarks based on ordering information. Two cases have to be considered: inside and outside of landmarks. While the qualitative panorama representation of the latter is dynamic in a predictable way as stated in the two observations described above, the panorama representation inside a landmark (e.g., a room) is stable and can therefore be used as an allocentric frame of reference. We assume that panorama ordering information can be generated by the vision process with limited effort, since it relies on pure geometric, and not on any domain-specific knowledge like color and texture (and also not on distance information). Nevertheless, we do not state that these qualities of information are generally not relevant We believe instead that they can be integrated into the extended panoramic representation EEP in a straightforward fashion and be that they can be combined in a flexible way. Furthermore, these representations provide abstractions in terms of translation, and/or rotation invariance which allows us also to identify stable and similar situation.

Although a detailed analysis of the relation to the recent cognitive results is out of the scope of this paper, we would like to mention that the extended panorama representation shows several properties which are observed in recent experiments: e.g., translation tasks seem to be performed more easily and accurately than rotation tasks. A panorama representation will behave very much in the same manner because of the translation invariance. Additionally, this approach may help to explain when participants why build a new egocentric representation instead of classifying it into a given one. Taking the general concept of similarity as complexity of transition, as, e.g., proposed in (Hahn *et al.* 2003), we can use the defined panorama cycles to determine the complexity of movement around a set of objects. Furthermore, the panoramic representation relies on the same kind of information: purely geometric properties as, e.g., proposed by (Wang and Spelke 2002) (among others). A detailed analysis is in preparation.

Several tasks remain to be done. Currently, we are developing a re-orientation vision module based on panorama representation and test it with the Sony Aibos. Perhaps even more promising as a result of an easier and more robust feature extraction, we try to integrate a panorama representation in an omnivision module. Although we use a panorama representation already in our simulation league team, we have to prove that it will also work with real vision data in physical grounded environments. Additionally, we need a more complete theoretical analysis, especially for the *within* landmark scenario.

Acknowledgement

The presented work is being funded by the Deutsche Forschungsgemeinschaft (DFG) within the project *Automatic Plan Recognition and Intention Recognition of Foreign Mobile Robots in Cooperative and Competitive Environments* as part of the Priority Research Program SPP-1125 *Cooperative Teams of Mobile Robots in Dynamic Environments*.

References

Eliseo Clementini, Paolino Di Felice, and Daniel Hernandez. Qualitative representation of positional information. *Artificial Intelligence*, 95(2):317–356, 1997.

A G Cohn and S M Hazarika. Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae*, 46(1-2):1–29, 2001.

Anthony G. Cohn, Brandon Bennett, John Gooday, and Nicholas Mark Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *GeoInformatica*, 1(3):275–316, 1997.

R.D. Easton and M.J. Sholl. Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21(2):483–500, 1995.

Kutluhan Erol, James Hendler, and Dana S. Nau. HTN planning: Complexity and expressivity. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, volume 2, pages 1123–1128, Seattle, Washington, USA, 1994. AAAI Press/MIT Press.

Cohn A.G. Fernyhough, J. and D. Hogg. Constructing qualitative event models automatically from video input. *Image and Vision Computing*, 18:81–103, 2000.

Andrew U. Frank. Qualitative spatial reasoning: Cardinal directions as an example. *International Journal of Geographical Information Science*, 10(3):269–290, 1996.

B. Garsoffky, S. Schwan, and F.W. Hesse. Viewpoint dependency in the recognition of dynamic scenes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(6):1035–1050, 2002.

U. Hahn, N. Chater, and L.B. Richardson. Similarity as transformation. *Cognition*, 87:1–32, 2003.

Ulrich John. Solving large configuration problems efficiently by clustering the conbacon model. In *IEA/AIE 2000*, pages 396–405, 2000.

B. J. Kuipers and T. Levitt. Navigation and mapping in large scale space. *AI Magazine*, 9(2):25–43, 1988.

Benjamin Kuipers. The spatial semantic hierarchy. *Artificial Intelligence*, 119(1-2):191–233, 2000.

J.M. Loomis, Y. Lippa, R.L. Klatzky, and R.G. Golledge. Spatial updating of locations specified by 3-d sound and spatial language. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(2):335–345, 2002.

Massimo Lucchesi. *Choaching the 3-4-1-2 and 4-2-3-1*. Reedswain Publishing, edizioni nuova prhomos edition, 2001.

J. OKeefe and A. Speakman. *The Hippocampus as A Cognitive Map*. Oxford, Clarendon Press, 1978.

J. Renz and B. Nebel. On the complexity of qualitative spatial reasoning. *Artificial Intelligence*, 108:69–123, 1999.

B. Roskos-Ewoldsen, T.P. McNamara, A.L. Sheldon, and W. Carr. Mental represenations of large and small spatial layouts are orienation dependent. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24(1):215–226, 1998.

C. Schlieder. Representing visible locations for qualitative navigation. pages 523–532. 1993.

C. Schlieder. Ordering information and symbolic projection, 1996.

M.J. Sholl and T.L. Nolin. Orientation specifity in representations of place. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23(6):1494–1507, 1997.

Michael K. Smith, Chris Welty, and Deborah L. McGuinness. Owl web ontology language guide. W3c candidate recommendation 18 august 2003, 2003. http://www.w3.org/TR/owl-guide/.

E.C. Tolman. Cognitive maps in rats and men. *Psycholoical Review*, 55:189–208, 1948.

R.F. Wang and E.S. Spelke. Updating egocentric represenations in human navigation. *Cognition*, 77:215–250, 2000.

R.F. Wang and E.S. Spelke. Human spatial representation: Insights from animals. *Trends in Cognitive Science*, 6(9):176–182, 2002.

R.F. Wang. Representing a stable environment by egocentric updating and invariant representations. *Spatial Cognition and Computation*, 1:431–445, 2000.

Wai K. Yeap and Margaret E. Jefferies. Computing a representation of the local environment. *Artificial Intelligence*, 107(2), 1999.

J.Y. Zheng and S. Tsuji. Panoramic representation for route recognition by a mobile robot. *International Journal of Computer Vision*, 9(1):55–76, 1992.