An Egocentric Qualitative Spatial Knowledge Representation Based on Ordering Information for Physical Robot Navigation

Thomas Wagner[1] and Kai Hübner[2]

[1] Center for Computing Technologies (TZI),
Universität Bremen, D-28359 Bremen
twagner@tzi.de

[2] Bremen Institute of Safe Systems (BISS),
Universität Bremen, D-28359 Bremen
khuebner@tzi.de

Abstract. Navigation is one of the most fundamental tasks to be accomplished by many types of mobile and cognitive systems. Most approaches in this area are based on building or using existing allocentric, static maps in order to guide the navigation process. In this paper we propose a simple egocentric, qualitative approach to navigation based on ordering information. An advantage of our approach is that it produces qualitative spatial information which is required to describe and recognize complex and abstract, i.e., translation-invariant behavior. In contrast to other techniques for mobile robot tasks, that also rely on landmarks it is also proposed to reason about their validity despite insufficient and insecure sensory data. Here we present a formal approach that avoids this problem by use of a simple internal spatial representation based on landmarks aligned in an *extended panoramic representation* structure.

1 Introduction

Navigation is one of the most fundamental tasks to be accomplished by robots, autonomous vehicles, and cognitive systems. Most successful approaches in the area of robot navigation like potential fields (see [10] and [7]) are based on allocentric, static maps in order to guide the navigation process (e.g. [9]). This approach has an intuitive appeal and gains much intuition from cognitive science: the *cognitive map* (a good recent overview [16]). The main purpose is to build up a precise, usually allocentric, quantitative representation of the surrounding environment and to determine the robot's position according to this allocentric, quantitative map.

One difficulty results from the fact that the same spatial representation serves as a basis for different tasks often with heterogeneous requirements. For example, more abstract reasoning tasks like planning coordinated behavior, e.g., *counterattack* and *double pass*, and plan recognition usually rely on more abstract,

D. Nardi et al. (Eds.): RoboCup 2004, LNAI 3276, pp. 134–149, 2005.

qualitative spatial representations. Generation of qualitative spatial descriptions from quantitative data is usually a difficult task due to uncertain and incomplete sensory data. In order to fit heterogeneous requirements, we should be able to represent spatial qualitative description at different levels of granularity, i.e., invariant according to translation and/or rotation and based on different scalings.

Based on recent results from cognitive science (see, e.g., [30]), we present a formal, egocentric, and qualitative approach to navigation which overcomes some problems of quantitative, allocentric approaches. By the use of ordering information, i.e., based on a description of how landmarks can shift and switch, we generate an *extended panoramic representation* (EPR). We claim that our representation in combination with path integration provides sufficient information to guide navigation with reduced effort to the vision process. Furthermore the EPR provides the foundation for qualitative spatial descriptions that may be invariant to translation and/or rotation.

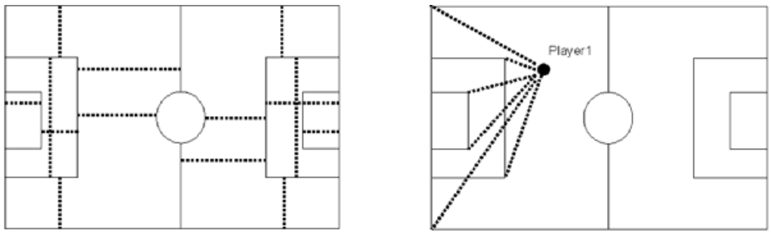
Since our approach abstracts from quantitative or metrical detail in order to introduce a stable qualitative representation between the raw sensor data and the final application, it can for example be used in addition to the well-elaborated quantitative methods.

2 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. [4]). Objects can be described either in ontology-based languages (e.g., OWL [21]) or constraint-based languages (e.g., [8]). Both types of representations allow for the representation of knowledge at different levels of abstraction according to the domain and task specific requirements. 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.

2.1 Allocentric and Egocentric Representations

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. In contrast, 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 and all other objects in the environment, whereas the number of



(a) Allocentric relations (b) Egocentric relations

Fig. 1. Allocentric vs. egocentric spatial relations

relations in egocentric representations can be significantly smaller (see Fig. 1)[1]. An interesting phenomenon, when looking into the didactic literature about, e.g., sports [12] we often find that (tactical and strategic) 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 sections 3.3 and 4 for more details).

3 Related Work

3.1 Cognition: Dynamic, Egocentric Spatial Representations

The fact that even many animals (e.g., rodents) are able to find new paths leading 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 that humans are able to perform (beginning with [26] and [14]). Nevertheless, recent results in cognitive science provide strong evidence for a different view ([30] among many others). Instead of using an allocentric view-independent map, humans and many animals build up a dynamic, view-dependent egocentric representation. Although the allocentric interpretation of the *cognitive map* seems to differ radically from the egocentric representation theory, both theories can account for many observations and differ mainly in two points: The allocentric, *cognitive map*-interpretation assumes that the spatial representation is view-independent and that therefore viewpoint changes do not have any influence on the performance of, e.g., spatial retrieval processes. Many recent experiments provide evidence for the opposite, they show that viewpoint changes can significantly

[1] For reasons of clarity not all allocentric relations are drawn in diagram 1(a).

reduce performance in terms of time and quality (e.g. pointing errors) (among others, [28] and [29]). The second main difference is concerned with the dynamic of the underlying representation. The egocentric interpretation assumes that all egocentric relations have to be updated with each egocentric movement of a cognitive system. The underlying assumption of a sophisticated series of experiments done by Wang ([28] and [29]) was that spatial relations have to remain stable in an allocentric, *cognitive map* independent from egocentric movements. When errors arise, e.g., because of path integration, 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 and have been confirmed in a series of differently designed experiments[2], e.g., [3] and [5].

3.2 Robot Navigation

Navigation and localization is the most fundamental task for autonomous robots and has gained much attention in the robotic research over the last decades. While several earlier approaches addressed this problem qualitatively [9], e.g., topological maps ([11], [15], [1]), more recent approaches focus very successfully on probabilistic methods. Famous examples are RHINO [23], MINERVA [22] and more recently [25]. Currently, the most promising techniques for robust mobile robot localization and navigation are either based on Monte-Carlo-Localization (MCL) (see [18] for RoboCup-application and the seminal paper [24]) or on various extentions of *Kalman-filters* (e.g., [13]) using probabilistic representations based on quantitative sensory data. MCL is based on a sample set of postures; the robot's position can be estimated by probabilities which allow to handle not only the *position tracking*- and the *global localization* problem but also the challenging *kidnapped robot* problem of moving a robot without telling it.

Furthermore, probabilistic methods based on quantitative data play a crucial role in handling the mapping problem, i.e., the SLAM-problem[3]. Very much the same is true for many robotic approaches to navigation, e.g., potential fields for avoiding obstacles by following the flow of superposed partial fields in order to guide the robot to a goal position (see [10] and [7] for a RoboCup-application) based on quantitative data.

According to the *spatial semantic hierarchy* (SSH) [9], these approaches try to address the problems related to robot navigation on the *control level*. Besides the strong computational resource requirements, they usually do not address the problem of generating a discrete, qualitative spatial representation which for

[2] Nevertheless, these results do not allow the strict conclusion that humans do not build up an allocentric cognitive map. On the contrary, e.g., Easton and Sholl [3] have shown that under very specific conditions it is possible to build up allocentric maps. Nevertheless, these results indicate, that under more natural conditions human navigation relies on egocentric snapshots and a dynamic mapping between these.

[3] This term is also directly connected to a set of algorithms addressing exactly this problem (e.g., [2]).

(a) Concrete Panorama (b) Abstract Panorama

Fig. 2. Panorama-views

instance is required at more abstract levels, e.g., for describing complex coordinated tactical and strategic behavior both on individual- and on team level.

3.3 The Panorama Approach

The concept of panorama representation has been studied extensively in the course of specialized sensors (e.g., omnivision, see, e.g., [31]). We present an extended approach based on the panorama approach by Schlieder ([20] and [19]).

A complete, circular panorama can be described as a 360^o view from a specific, observer-dependent point of view. Let P in Fig. 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. 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, [20] introduced a formal model of a panorama.

Definition 1 (Panorama). *Let $\Theta= \{\theta_1, \ldots, \theta_n\}$ be a set of points $\theta_i \in \Theta$ 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 Fig. 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). As the panorama is a cyclic structure 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 $dl(vp, lm_1, lm_2)$[4]. Based on this representation, [19] also developed an efficient qualitative navigation algorithm.

The panorama representation has an additional, more important property: it is invariant with respect to rotation and translation. But evidently, not ev-

[4] Short for $direct - left(viewpoint, landmark_1, landmark_2)$.

ery 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 information (e.g., distance) is required in some situations. In the following section, we show how the panorama can be extended in a way that more detailed ordinal and metric information can be introduced.

4 An Extended Panorama Representation

Instead of building an allocentric map we provide an egocentric snapshot-based approach to navigation. The most fundamental difference between both approaches is that an egocentric approach strongly relies on an efficient, continuous update mechanism that updates all egocentric relations in accordance with the players' movement. In this section we show that this task can be accomplished by strict use of a simple 1D-ordering information, namely an extended qualitative panorama representation (EPR).

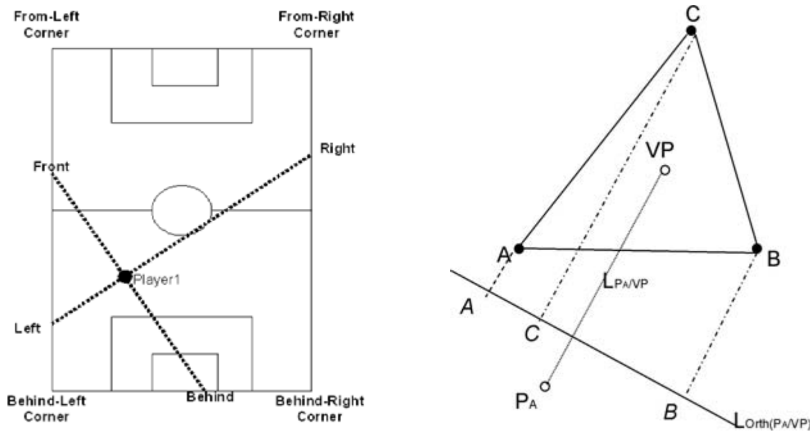
This update mechanism has to be defined with respect to some basic conditions:

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

Due to the nature of ordering information, this task has to be divided into two subtasks: (1) updating within a given frame of reference (short notation: FoR), i.e., the soccer field and (2) updating of landmark representations from an external point of view, e.g., the penalty area. In section 4.1 we briefly discuss the key properties of the first task in relation to ordering information from a more theoretical point of view, whereas in section 5 these aspects are investigated in more detail. In section 4.2 we describe the theoretical framework underlying the mapping- and update-mechanism for egocentric views on external landmarks.

4.1 Within a Frame of Reference

A crucial property of panoramic ordering information is that it does not change as long as an agent stays within a given FoR, i.e., the corners of a soccer field, do not change unless the player explicitly leaves the field (see Fig. 3(a)). So in order to use ordering information for qualitative self-localization we have to introduce an egocentric FoR. But even with an egocentric FoR the location within this



(a) Use of Egocentric Frame of Reference

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

Fig. 3. FoR and Triangle panorama

FoR can only be distinguished into a few different qualitative states (e.g., ego-front between front-left and front-right corner of the field, see Fig. 3(a)). This way of qualitative self-localization is too coarse for many domains as well as for the different RoboCup-domains. In section 5 we demonstrate in more detail how angular distances can be used to overcome this problem[5].

A perhaps even more important property of spatial locations within a given FoR is that they can be used as a common FoR for the position of different landmarks in relation to each other (e.g., the position of the penalty area can be described in within-relation to the soccer field). This property is especially important for an egocentric snapshot-based approach to navigation since it provides the common frame that is required to relate different snapshots to each other (for a more detailed discussion see [27]).

4.2 Updating Outside-Landmark Representations

In a re-orientation task we can resort the knowledge about the previous position of a player. Therefore we concentrate on an incremental updating process, based on the following two 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 position P_{t-1} of an agent A in relation to L at time step $t-1$ is known. The EPR (LP_T) of a triangle configuration can then be defined as follows (see also Fig. 3(b)):

---

[5] An additional approach is to introduce more landmarks that are easy to perceive or to introduce additional allocentric FoR when available (e.g., north, south, etc.).

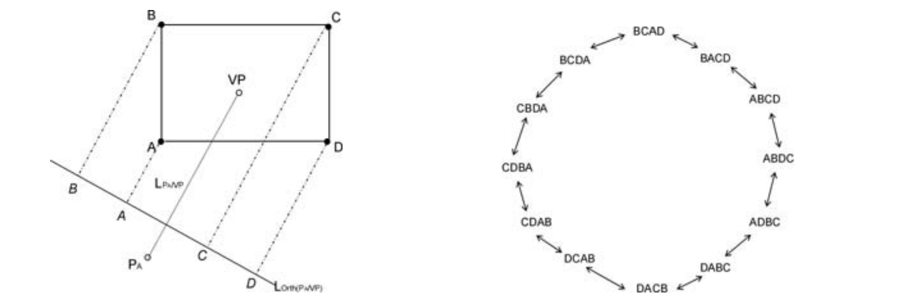
Definition 2 (Triangle Landmark Panorama). *Let P_A denote the position of an agent A and $C_{T(ABC)}$ the triangle configuration formed by the set of points A, B, C in the plane. The line $L_{P_A/VP}$ is the line of view from P_A to* VP*, with* VP *being a fixed point within $C_{T(ABC)}$. Furthermore, $L_{Orth(P_A/VP)}$ 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 describe 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 EPR 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:*
$(CAB), (ACB), (ABC), (BAC), (BCA), (CBA)$

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 view 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 (see Fig. 3(b)). This property holds in general for triangle configurations but not, e.g., for quadrangle configurations (except for some special cases as we will see below). 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[6]. A natural extension seems to be applying the same idea to quadrangles (see Fig. 4). 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. With this approach, the space around a quadrangle would be separated into ten areas and therefore it would be more expressive than the more simple triangle panorama. It can be shown that eight of the resulting triangle landmark panoramas (one for each triangle of the quadrangle) can be transformed into quadruple tuples that result when we transform, e.g., a rectangle directly into a landmark panorama representation (e.g., the given tuple ((BCA)(CDA)) can be transformed into (BCDA) without

[6] The triangle configuration can be applied generally to any triplet of points that form a triangle - also to solid objects. The connecting lines pictured in Fig. 3(b) and 4(a) are used to explain the underlying concept of position exchange (transition).



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

(b) Circular representation of panoramic ordering information for parallelograms

Fig. 4. Parallelogram panorama

loss of information)[7]. The expressiveness of the other two landmark panoramas is weaker: they have to be described as a disjunction of two quadruple tuples. 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 one (see Fig. 4(a)).

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 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 $C_{P(ABCD)}$. 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 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 describe the location of the observer position qualitatively. A 360^o movement can be split into twelve different states:

Observation 2. *(Parallelogram Landmark 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:*

[7] The detailed proof will take too much space. However, 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 (multiplied 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 (for more details on the constraint representation see section 3.3).

$((BCAD), (BACD), (ABCD), (ABDC), (ADBC), (DABC),$
$(DACB), (DCAB), (CDAB), (CDBA), (CBDA), (BCDA))$

The two presented landmark panoramas can be mapped flexibly onto landmarks that can be found in natural environments like a penalty area. 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 us to build up nesting representations which support different levels of granularity according to the requirements of the domain.

5 Implementation

According to the described scenarios, the EPR is meant to be a qualitative fundament for tasks that are important for mobile robot exploration. Due to the oversimplification of the four-legged league RoboCup scenario (i.e., no penalty area and no goal area to move around), the latter outside-case described in section 4.2 does not find capital application here, but we claim that it will show its features in more complex scenarios which offer a larger number of landmarks to move around. Here, we will show some experimental extraction of EPR sequences to practically point up the idea presented in section 4.1 and the basic idea of building panoramic ordering information from the image data.

For our first experiments, we use the *RobotControl/SimRobot* [17] simulation environment for the simulation of one four-legged robot. This tool is shared with the GermanTeam, which is the German national robotic soccer team participating in the Sony four-legged league in the international RoboCup competitions. The EPR concept presented is not proposed to be restricted to this special domain, as discussed. The tool supports simulated image retrieval and motion control routines that are easy to use and portable to physical robots, while it is possible to encapsulate the EPR and adapted image feature extraction in distinct solutions, letting other modules untouched.

5.1 Visual Feature Extraction

In order to expediently fill the EPR with information, the recognition of landmarks is necessary. Usually, the robot's viewing angle of 57.6^o degrees is not sufficient to get a reasonably meaningful EPR with the feature extraction of goals and flags supported by the *RobotControl* tool (see [18] for a description of these features).

Even if the scene is regarded from one goal straight to the other, there are just three landmarks that can be found. On the other hand, the standard configuration of all landmarks as can be seen in Fig. 5 is of an unfavorable kind for the EPR. The landmarks build a convex structure that the robot can never

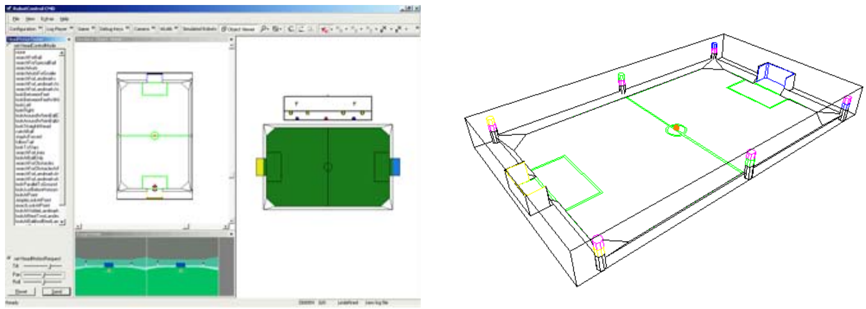


Fig. 5. Simulation environment of the GermanTeam (left); the standard four-legged league field configuration (right)

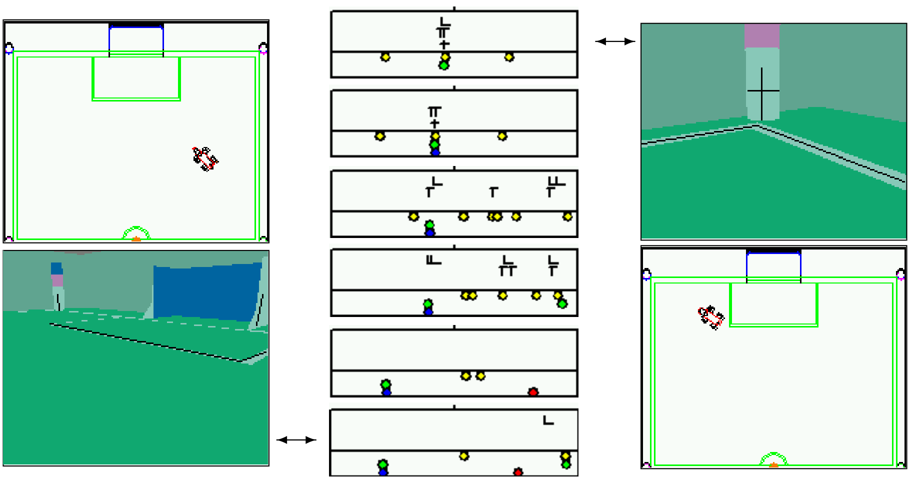


Fig. 6. Landmarks for the EPR. Center column: Landmarks extracted (for six representation between given start position (left) and goal position (right): "L" for L-junctions, "T" for T-junctions, "X" for X-junctions; horizontal lines (circles), vertical lines (squares), goals (light triangles) and flags (dark triangles))

leave, thus the ideal EPR will never allow to reason about the environment by permuted landmarks (see section 4.1). Thus, we further introduced the symmetry line operator proposed by Huebner [6] to extract 2D field lines as additional features from the image data. The method is simple, robust, and works without plenty of parametrization. Additionally, it offers the opportunity to test the approach with natural landmarks (lines) instead of artifacts (colored beacons). After processing the images, lines are distinguished from curves and represented by their start and end point in the image (see Fig. 6).

These lines can be put into the EPR by adopting these points or the center point, for example. Anyway, a classification of edge types is more efficient with respect to the subsequent need of recovering landmarks. To support the panorama with a broader range of landmark types which ideally are points on the field, we classify each pair of lines extracted from an image into different line pair types. In our experiment, we extracted L-junctions, T-junctions and X-junctions (see Fig. 6) representing the additional landmarks that are used for the EPR.

5.2 Qualitative Representation

The simulated environment for the experiment corresponds to the standard four-legged league field configuration with lines instead of the sideboards. One robot is instructed to move a certain path presented by a given sequence of EPRs. Using the EPR representation and a qualititative conversion of the feature angles, we can establish a qualitative EPR sequence of detected landmark configurations for a path. Some samples of such sequences might look like the following, corresponding to the EPR of Fig. 6:

```
[(T_JUNC,VERY_FAR);(L_JUNC,SAME);(T_JUNC,SAME);(X_JUNC,SAME);]
[(T_JUNC,VERY_FAR);(X_JUNC,SAME);(FLAG,SAME);(T_JUNC,SAME);]
[(T_JUNC,FAR);(FLAG,SAME);(L_JUNC,CLOSE);(T_JUNC,MEDIUM);
 (T_JUNC,MEDIUM);(L_JUNC,SAME);(L_JUNC,CLOSE);]
[(FLAG,FAR);(L_JUNC,SAME);(L_JUNC,CLOSE);(T_JUNC,MEDIUM);
 (L_JUNC,SAME);(T_JUNC,CLOSE);(L_JUNC,MEDIUM);(T_JUNC,SAME);]
[(FLAG,FAR);(GOAL,FAR);]
[(FLAG,FAR);(GOAL,FAR);(L_JUNC,CLOSE);]
```

Each of these ordering sequences corresponds to a snapshot-like qualitative description of the robot's location during the path. E.g., in the first sequence, there are four ordered and classified landmarks that are additionally described by their qualitative angular distance to the previous landmark. Caused by the panoramic representation, the first "T"-junction is very far (VERY_FAR) displaced from the previous landmark (the "X"-junction in this case). Including qualitative angular distances like VERY_FAR also allows to convert this angular representation to a number of qualitative location descriptions (e.g., according to the first sequence, "The X-junction is very far LEFT of the T-junction." or "The X-junction is same RIGHT of the T-junction.").

As also can be seen in this example, the line landmarks appear and disappear frequently in the robot's view. This is caused by the landmark feature extraction working on insufficient simulated image data. We are optimistic that real images are more comfortable for the extraction of lines, because they are not supposed to be fragmented like those in simulated images. Although this is error-prone in this regard, we claim to deal with this problem using the EPR. The representation can generally be useful for this re-orientation task, where the 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.

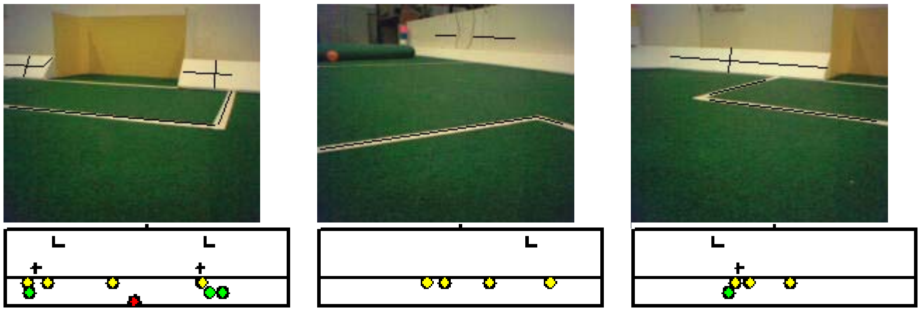


Fig. 7. Landmarks for the EPR on real images. Top row: image data and extracted field / border lines. Bottom row: Landmarks extracted

The same panoramic representation is additionally used in our simulation soccer team *Virtual Werder*. Although sensor problems are neglectable since the world model is more comprehensive and detailed, it provides a simple and intuitive interface for the generation of qualitative descriptions.

5.3 Experiments on Real Images

Finally, some experiments have been made to test the proposed feature extraction and EPR construction on real images (see Fig. 7)[8] using one Sony AIBO ERS-7 model inside a common four-legged league scenario. Without plenty of adaptation, the results are as good as those in the simulation examples. Problems appearing by the line extraction technique (e.g. side walls as lines, lines found over horizon, optional grouping of lines to handle occlusions) will be addressed in future work to increase robustness and performance.

6 Conclusion and Future Work

Navigation, localization, planning, and reasoning for physically grounded robots imposes strong but heterogeneous requirements on the underlying spatial representation in terms of abstraction and precision. 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. The qualitative EPR is dynamic in a predictable way for outside landmarks as stated in the two observations described above. This representation, however, provides also interesting properties for navigation inside fixed landmarks (e.g., navigating within a room).

Besides the re-orientation task mentioned in the last section, the landmark panorama can help to focus perception in a qualitative self-allocation task. Dur-

[8] The difference of size in the corresponding images is caused by the different image sizes between the old AIBO model ERS-210 and the new ERS-7.

ing the transition of one panorama landmark into another exactly one position change is performed. Therefore, in this case the perception of further landmarks is without any use for updating the qualitative position of the agent. Additionally, the panorama landmark representation is not only useful for position updating but also for re-orientation without knowledge about the previous position. The perception of a partial landmark panorama of a triangle configuration is sufficient to provide us with two hypotheses about the current position. In order to validate which hypothesis holds we just have to find out where another landmark appears in the panoramic structure. Additionally, a landmark panorama provides a stable basis for qualitative, spatial descriptions (e.g. left of, right of), since it is, obviously, sensitive to rotation but invariant to transition, it is also interesting for several outstanding applications based on qualitative information.

Although a detailed analysis of the relation to the recent cognitive results is out of the scope in this paper, we want to mention that the EPR shows several properties which are observed in recent experiments: e.g., translation tasks seem to be performed more easily and accurately than rotation tasks.

Several tasks remain to be done. We are currently extending our landmark-based (re-)orientation vision module so that it is not only able to track EPRs but also allows active snapshot-based navigation (first results are available). Thereby we implement the concept of outside-landmarks that formally describes how landmarks can shift and switch during movement (see section 4.2). This should also allow to detect the geometric structure of previously unseen objects. After validating our extended panorama representation in the RoboCup-domain, we consider to transfer this method of the EPR into an omnidirectional vision module for mobile robot tasks.

Acknowledgements

The presented work is being funded by the German Research Foundation (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

1. D. Busquets, C. Sierra, and R. L. De Mantaras. A multiagent approach to qualitative landmark-based navigation. *Autonomous Robots*, 15(1):129–154, 2003.
2. H. Durrant-Whyte, S. Majumder, S. Thrun, M. de Battista, and S. Schelling. A bayesian algorithm for simulaneous localization and map building. In *Proceedings of the 10'th International Symposium on Robotics Research (ISRR'01))*, Lorne, Australia, 2001. AAAI Press/MIT Press.
3. 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.

4. K. Erol, J. Hendler, and D. 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.
5. 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.
6. K. Huebner. *A Symmetry Operator and its Application to the RoboCup.* 7th International Workshop on RoboCup 2003 (Robot World Cup Soccer Games and Conferences). Lecture Notes in Artificial Intelligence. Springer Verlag, 2004. To appear.
7. S. Johansson and A. Saffiotti. Using the Electric Field Approach in the RoboCup Domain. In A. Birk, S. Coradeschi, and S. Tadokoro, editors, *RoboCup 2001: Robot Soccer World Cup V*, number 2377 in LNAI, pages 399–404. Springer-Verlag, Berlin, DE, 2002.
8. U. John. Solving large configuration problems efficiently by clustering the conbacon model. In *IEA/AIE 2000*, pages 396–405, 2000.
9. B. Kuipers. The spatial semantic hierarchy. *Artificial Intelligence*, 119(1-2):191–233, 2000.
10. J. C. Latombe. *Robot Motion Planning*, volume 18. Kluwer Academic Press, 1991.
11. T. S. Levitt and D. T. Lawton. Qualitative navigation for mobile robots. *Artificial Intelligence*, 44:305–360, 1990.
12. M. Lucchesi. *Coaching the 3-4-1-2 and 4-2-3-1.* Reedswain Publishing, edizioni nuova prhomos edition, 2001.
13. C. Martin and S. Thrun. Online acquisition of compact volumetric maps with mobile robots. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2002.
14. J. O'Keefe and A. Speakman. *The Hippocampus as A Cognitive Map.* Oxford, Clarendon Press, 1978.
15. T. J. Prescott. Spatial representation for navigation in animats. *Adaptive Behavior*, 4(2):85–125, 1996.
16. A. D. Redish. *Beyond the Cognitive Map - From Place Cells to Episodic Memory.* The MIT Press, Cambridge, Mass., 1999.
17. T. Roefer. *An Architecture for a National RoboCup Team*, pages 417–425. RoboCup 2002: Robot Soccer World Cup VI. Lecture Notes in Artificial Intelligence. Springer Verlag, 2003.
18. T. Roefer and M. Juengel. Vision-Based Fast and Reactive Monte-Carlo Localization. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-2003)*, pages 856–861, Taipei, Taiwan, 2003.
19. C. Schlieder. Representing visible locations for qualitative navigation. pages 523–532. 1993.
20. C. Schlieder. Ordering information and symbolic projection, 1996.
21. M. K. Smith, C. Welty, and D. L. McGuinness. Owl web ontology language guide. W3c candidate recommendation 18 august 2003, 2003. http://www.w3.org/TR/owl-guide/.
22. S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, , and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. *International Journal of Robotics Research*, 19(11):972–999, 2000.

23. S. Thrun, A. Bücken, W. Burgard, D. Fox, T. Fröhlinghaus, D. Henning, T. Hofmann, M. Krell, and T. Schmidt. *AI-based Mobile Robots: Case Studies of Successful Robot Systems*, chapter Map learning and high-speed navigation in RHINO. MIT Press, 1998.
24. S. Thrun, D. Fox, W. Burgard, and F. Dellaert. Robust monte carlo localization for mobile robots. *Artificial Intelligence*, 128(1-2):99–141, 2000.
25. S. Thrun, D. Hähnel, D. Ferguson, M. Montemerlo, R. Triebel, W. Burgard, C. Baker, Z. Omohundro, S. Thayer, and W. Whittaker. A system for volumetric robotic mapping of abandoned mines. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2003.
26. E. C. Tolman. Cognitive maps in rats and men. *Psycholoical Review*, 55:189–208, 1948.
27. T. Wagner, C. Schlieder, and U. Visser. An extended panorama: Efficient qualitative spatial knowledge representation for highly dynamic enironments. *IJCAI-03 Workshop on Issues in Desgning Physical Agents for Dynamic Real-Time Environments: World Modeling, Planning, Learning, and Communicating*, pages 109–116, 2003.
28. R. F. Wang. Representing a stable environment by egocentric updating and invariant representations. *Spatial Cognition and Computation*, 1:431–445, 2000.
29. R. F. Wang and E. S. Spelke. Updating egocentric represenations in human navigation. *Cognition*, 77:215–250, 2000.
30. R. F. Wang and E. S. Spelke. Human spatial representation: Insights from animals. *Trends in Cognitive Science*, 6(9):176–182, 2002.
31. 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.