

# Spatial memories in place recognition

## **Dissertation**

der Mathematisch-Naturwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr.rer.nat.)

vorgelegt von  
**Stephan Lancier**  
aus Böblingen

Tübingen  
2015

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Prüfung:

08.04.2016

Dekan:

Professor Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Prof. Dr. Hanspeter Mallot

2. Berichterstatter:

Prof. Dr. Marc Ernst

## Veröffentlichungen

Lancier, S., Admard, M., Dorer, M., Gross, I.M., Link, A., Mallot, H.A. (2015). Spatial awareness: Beware of what is out of sight. Manuscript in preparation.

## Konferenzbeiträge

Lancier, S., Mallot, H.A. (2015). Likelihood estimation of places in local environments. Conference Abstract: European Conference on Visual Perception.

Lancier, S., Link, A., Admard, M., Mallot, H.A., (2014). Eccentricity of landmark configuration influences distribution of place recognition errors. *Perception*, ECVP Abstract Supplement, (43):49.

Lancier, S., Hofmeister, J., Mallot, H.A. (2014). Event-Related Potentials in the Recognition of Scene Sequences. *Cognitive Processing*, KogWis Abstract Supplement, 15(1):49-50.

Lancier, S., Halfmann, M., Link, A., Admard, M., Mallot, H.A. (2014). Distribution of place recognition errors depends on eccentricity of landmark configuration. Conference Abstract: Spatial Cognition.

Lancier, S., Halfmann, M., Gross, I.M., Mallot, H.A. (2014). How do different learning conditions influence place recognition. Conference Abstract: Tagung experimentell arbeitender Psychologen.

Halfmann, M., Lancier, S., Roth, S., Mallot, H.A. (2013). Depth of processing in human place recognition. Conference Abstract: 10th Meeting of the German Neuroscience Society.

Halfmann, M., Lancier, S., Mallot, H.A. (2013). Size vs Configuration of Visual Cues in Human Place Recognition. Conference Abstract: Bernstein Conference.

Lancier, S., Halfmann, M., Mallot, H.A. (2012). Mechanisms of human place recognition. Conference Abstract: Bernstein Conference.



## Acknowledgments

First of all I want to give thanks to my parents Karin and Bodo Lancier for the great support during my biological studies and doctoral thesis. I would like to thank Prof. Hanspeter Mallot to enable me to do this doctoral thesis at his Chair of Cognitive Neuroscience at the Department of Biology and of course for his great support. I am grateful for the inspiring discussions with all my colleagues, and also for the fun we had at social events as well as at work. I would like to thank my bachelor students Amanda Link, Isa-Maria Gross, Marie Admard, Marcel Dorer and Niklas Schulze for supporting the place recognition study. I would also like to thank my other bachelor students Julian Hofmeister and Simon Klesen for supporting the event-related potential study. This doctoral thesis was supported by the Bernstein Center of Computational Neuroscience Tübingen (Grant by Federal Ministry of Education and Research FKZ 01GQ1002), for that, thanks a lot.



## Zusammenfassung

Diese Doktorarbeit untersucht die Rolle von Raumgedächtnissen in der Ortserkennung. Die erste Studie fokussiert sich auf das Zusammenspiel zwischen dem räumlichen Arbeitsgedächtnis und der Ortskodierung im Langzeitgedächtnis. Des Weiteren wollen wir herausfinden wie die lokale Positionsinformation im räumlichen Arbeitsgedächtnis enkodiert ist. Unser Vorschlag ist, dass die Ortserkennung auf einer intermediären Repräsentation basiert, die die Distanz- und Richtungsinformation über navigations-relevante Objekte in der Umgebung beinhaltet. Dafür entwickelten wir einen experimentellen Aufbau mit dem die Verteilung von Entscheidungen gemessen wird, bei denen die Versuchspersonen einen Zielort wiedererkannt haben. Das experimentelle Ergebnis wurde verglichen mit den simulierten Daten des entwickelten Likelihood-Modells, welches die vom Arbeitsgedächtnis und der gespeicherten Ortskodierung repräsentierte wahrgenommene Distanz und Richtung mit einbezieht. Die Ergebnisse zeigen, dass das räumliche Arbeitsgedächtnis und die Ortskodierung auf einer kartenähnlichen Struktur beruhen. Das räumliche Arbeitsgedächtnis baut sich aus den metrischen Wahrscheinlichkeitsschätzungen der Umgebung auf. Allerdings sind die Distanzschätzungen nicht aus dem euklidischen Raum abgeleitet, sondern vermutlich hyperbolisch komprimiert. Wohingegen die Ortskodierungs-Schätzungen durch Triangulation eher wahrheitsgemäß sind.

Die zweite Studie adressiert die Frage, ob der Abruf von zeitlichem Kontext aus dem episodischen Gedächtnis, der entscheidend für die Integration von vielfachen Ansichten in der Ortserkennung ist, ein charakteristisches Ereignis-korreliertes Potential hervorruft. Hierzu modifizierten wir die klassische ‚alt/neu‘ Aufgabe, in dem wir den zeitlichen Kontext einer gelernten Abfolge von präsentierten Szenen abfragten. Die Ergebnisse zeigten, dass der Abruf von zeitlichem Kontext aus dem episodischen Gedächtnis durch eine Verzögerung der Latenz der P600 Amplitude gekennzeichnet ist, die in der zentral-parietalen Region lokalisiert ist. Wir vermuten, dass dieser Effekt mit der Integration von vielfachen Ansichten im räumlichen Arbeitsgedächtnis verknüpft ist.





## Abstract

This thesis investigates the role of spatial memories in place recognition. The first study focuses on the interplay between the spatial working memory and place codes in long-term memory. Further we want to figure out how the local position information is encoded in spatial working memory. Our proposal is that place recognition is based on an intermediate representation, containing distance and bearing information about the surrounding environmental cues. Therefore we designed an experimental setup, in which we measure the distributions of decisions, where the participants recognized a target location. The experimental results were compared with the simulated data of the formulated likelihood-based model which takes into account the perceived distance and bearing represented by the working memory and the stored place code. The results show that the spatial working memory and the place code are based on a map-like structure. Spatial working memory is built up by probabilistic metric estimates of the environment. However, distance estimates in spatial working memory are not derived from Euclidean space, but rather are assumed to be hyperbolically compressed. Whereas the place code estimates are assumed to be more veridical due to triangulation.

The second study addresses the question whether retrieval of temporal context from episodic memory, which is crucial for multiple view integration in place recognition, evokes a characterizing event-related potential. Concerning this we modified the classical ‘old/new’ task by inquiring the temporal context of a learned sequence of presented scenes. The results revealed that the retrieval of temporal context from episodic memory is signaled by a latency delay of the P600 amplitude located in the mid-parietal region. We assume that this effect is linked to the multiple view integration in spatial working memory.



# Contents

Acknowledgments . . . . .	v
Zusammenfassung . . . . .	vii
Abstract . . . . .	ix
<b>1 Introduction</b>	<b>1</b>
1.1 Single versus multiple views in place recognition . . . . .	1
1.2 Depth processing . . . . .	2
1.3 Spatial working memory . . . . .	4
1.4 A model of place recognition . . . . .	5
1.5 Place recognition in the hippocampal circuit . . . . .	7
1.6 Event-related potentials in episodic memory . . . . .	8
1.7 Purpose of the behavioral place recognition study . . . . .	9
1.8 Purpose of the sequence-related ERP study . . . . .	10
<b>2 Spatial awareness - Beware of what is out of sight</b>	<b>11</b>
2.1 General methods . . . . .	11
2.1.1 Participants . . . . .	11
2.1.2 Apparatus . . . . .	11
2.1.3 Virtual reality environment . . . . .	13
2.1.4 Procedure . . . . .	14
2.2 Experiment 1: Implicit versus explicit learning . . . . .	15
2.2.1 Methods . . . . .	15
2.2.2 Results . . . . .	16

2.2.3	Discussion . . . . .	17
2.3	Experiment 2: Landmark salience . . . . .	18
2.3.1	Method . . . . .	19
2.3.2	Results . . . . .	19
2.3.3	Discussion . . . . .	20
2.4	Experiment 3: Distance versus bearing . . . . .	20
2.4.1	Method . . . . .	21
2.4.2	Results . . . . .	21
2.4.3	Discussion . . . . .	22
2.5	Experiment 4: Dependence of place recognition on landmark configuration . . . . .	24
2.5.1	Method . . . . .	26
2.5.2	Results . . . . .	26
2.5.3	Discussion . . . . .	29
2.6	General Discussion . . . . .	30
2.6.1	Likelihood-based spatial representation . . . . .	31
2.6.2	Evaluation of the model . . . . .	35
2.7	Conclusion . . . . .	37
<b>3</b>	<b>Neural correlate of sequence-related retrieval of scenes</b>	<b>39</b>
3.1	Methods . . . . .	39
3.1.1	Participants . . . . .	39
3.1.2	Apparatus . . . . .	40
3.1.3	Procedure . . . . .	41
3.1.4	Stimuli . . . . .	43
3.1.5	Data analysis . . . . .	44
3.2	Results . . . . .	46
3.2.1	Psychophysics . . . . .	46
3.2.2	Event-related potentials . . . . .	48
3.3	Discussion . . . . .	55
3.4	Conclusion . . . . .	57

---

<b>References</b>	<b>66</b>
<b>A Appendix</b>	<b>67</b>
A.1 Heat maps of trajectories . . . . .	68
A.2 Bar graphs of landmarks within the field of view . . . . .	72
A.3 Procedure description . . . . .	74
A.4 Content form . . . . .	76
<b>List of figures</b>	<b>81</b>
<b>List of Tables</b>	<b>82</b>



# 1. Introduction

## 1.1. Single versus multiple views in place recognition

Recognition of familiar places requires a representation of memorized local position information, or place code of each place. This place code contains local position information visible in the current viewing direction but may also include information from previously observed, nearby angles (for review Madl et al., 2015). Single-view approaches relying exclusively on the cues available in the current field of view have been reviewed by Trullier et al. (1997). Depending on the amount of processing required to extract the relevant cues from the retinal image, various models can be distinguished ranging from edge-enhanced ‘snapshots’ (Cartwright and Collett, 1982), to ‘boundary vectors’, i.e. allocentric depth and bearing of boundaries and objects (Barry et al., 2006; Hartley et al., 2000), to the recognition of landmark objects (Janzen and Van Turenhout, 2004) or to spatial layouts (geometry of walls or landscapes) represented in the parahippocampal place area (Epstein, 2008). Among these different cues to place recognition, we define landmarks as salient, relevant for navigation, and reliably recognizable features or objects which are stored in spatial memory. Besides supporting binary decisions in place recognition, landmarks can also induce a reference frame for judging gradual deviations from a goal position. This function of providing a reference frame is particularly obvious in global landmarks, delimiting walls, or boundaries defined by object configurations (see Appleyard, 1969; Chrastil and Warren, 2014; Gillner et al., 2008; Steck and Mallot, 2000; Thorndyke and Hayes-Roth, 1982).

Multi-view models of place recognition can be useful to model view preferences in spatial recall or place representations of varying scale. They require an integration stage for the multiple views which can be thought of as an egocentric spatial working memory (Byrne et al., 2007; Giudice et al., 2013; Röhrich et al., 2014; Tatler and Land, 2011). The integration of multiple views can be solved by a view graph model in which views obtained at a place

are linked to each other by turning movements, whereas views of different places are linked by translation movements (Hübner and Mallot, 2007; Röhrich et al., 2014; Schölkopf and Mallot, 1995). In this thesis we will suggest that long-term place codes are built from such working memories, i.e. from integrated representations of multiple views, rather than from single, directed views alone.

## 1.2. Depth processing

Studies on the homing behavior of digger wasps (Tinbergen and Kruyt, 1938), honey bees (Cartwright and Collett, 1982) and wood ants (Durier et al., 2003) revealed that the recognition of places in insects is directly linked to the angular size and the geometrical configuration of landmarks. Cartwright and Collett (1983) proposed a snapshot memory for homing and place recognition based on a panoramic view of the environment. According to that theory a place is recognized if the current retinal image matches the directions to the landmarks and the angular size of each landmark stored in the snapshot of the target place. In one experiment the configuration of three landmarks was distorted after training in such a way that the angular sizes of the landmarks (equivalent to the distance) match at one place while the relative bearings of the landmarks match at another place. The honey bees preferred the place where the bearings matched those of the stored snapshot. In another two experiments the sizes of the three landmarks were either increased or decreased resulting in search areas of the bees centered also at the place where the bearings matched. This is generally taken as evidence that the honey bee place code is a snapshot (set of relative landmark bearings) rather than a local depth map (set of landmark distances).

Landmark bearing and distance were crossed in an experiment with human participants by Waller et al. (2000) in which participants searched for the goal at the place where the landmark distances matched rather than where the landmark bearings matched, especially if the goal location was enclosed by the landmark configuration. The reason for this discrepancy may be the limited field of view of human observers as compared to the nearly panoramic view of honey bees. As a consequence, in an enclosing configuration the match for each viewing direction will only provide partial information, whereas multiple view integration across head and body movements is required to exploit the complete visual sphere. Under pure rotations, multiple view integration might lead to a simple panorama as in the image-stitching function of a camera. If, however, small translations between view acquisitions



are also considered, image integration entails the representation of some sort of depth or correspondence, for instance a local map supporting also spatial updating (Giudice et al., 2013; Loomis et al., 2013). If information from this local map is stored as a place code, the reported matching by landmark distance is to be expected.

Rats returning to previously learned places in rectangular box-mazes confuse diagonally opposed but not adjacent corners of the maze. These ‘rotational errors’ occur even in the presence of visual features marking the target location or if box orientation is disambiguated by asymmetric wall coloring. These results and similar results from human studies led to the assumption of a pure geometric module for place recognition (Cheng, 1986; Hermer and Spelke, 1994; Learmonth et al., 2001; Margules and Gallistel, 1988). A simulation study by Stürzl et al. (2008) showed, however, that rotational errors are also predicted if place recognition is based on an edge rather than an intensity or color image, rendering the argument for a geometric place code elusive. Direct evidence for rodents using place codes based on image intensity or texture was presented by Graham et al. (2006) using a ‘flip-flop’ task in a kite-shaped room that contained sufficient geometric cues to find a goal location based on depth alone. The performance decreased if the wall texture frequently changed during training compared to those trained with constant wall texture, indicating an integrative use of both depth and textures cues. Again, simulation studies showed that the behavioral results can be explained by assuming a panoramic edge-image as the underlying place code (Cheung et al., 2008).

Evidence for the snapshot-based place recognition in humans in absence of localized cues and landmarks was presented by Gillner et al. (2008). Participants were navigating to target places within a circular room where the only local position information was provided by a homogeneous color gradient covering the walls. The performance of the participants decreased with color gradient modulation and wall distance as predicted by a simple snapshot-model minimizing the sum of squared RGB-differences over all pixels (Franz et al., 1998). Geometry-based models of place recognition were not able to explain the results. We take these various experimental results to indicate that place recognition can be based on a variety of different cues, including image texture, edges, and depth. The information source used in a given task may depend on the reliability of the local position information available.

### 1.3. Spatial working memory

Spatial awareness, i.e. the feeling of being at a given place and the knowledge of the current spatial context, is not based on instantaneous views alone, but is constructed and continuously updated over time. As a result, a spatial working memory is built integrating spatial information from various senses, viewing directions, and view-points into an amodal ‘spatial image’ (Avraamides et al., 2004; Byrne et al., 2007; Giudice et al., 2013; Loomis et al., 2013; Mellet et al., 2002). The content of spatial working memory is continuously updated using information from changing perceptual input and ego-motion (Farrell and Robertson, 1998; Rieser, 1989; Wolbers et al., 2008). If a percept is terminated, e.g. by leaving the field of view, a coarser representation will generally be maintained in the spatial image (Loomis et al., 2013; Tatler and Land, 2011). Mou et al. (2004) suggested that spatial updating depends on intrinsic axes of the environment and is generally more precise for rotational than for translational movements of the body. The usage of spatial content out of sight is crucial for preparing eye movements and motor actions, e.g. grasping an object out of sight, or planning routes using global landmarks currently not available, e.g. a salient building in a city that disappears from time to time.

The integration of linear and angular self-motion (e.g. from proprioception and optic flow) is an essential source of metric information (Cheung and Vickerstaff, 2010; McNaughton et al., 2006; Mittelstaedt and Mittelstaedt, 1980; Wolbers et al., 2007). This path integration can be thought of as the spatial updating of the start position of the current journey but will also enter computations for spatial updating of other environmental objects. Theoretically, it can be based on ego-motion cues alone (Mittelstaedt’s idiothetic cues), but is much improved by place or landmark recognition which allows to overcome error accumulation (Etienne et al., 2000; Etienne and Jeffery, 2004; Foo et al., 2005; Hübner and Mallot, 2007; Thrun et al., 2004). The maintenance of local metric knowledge thus involves interactions between spatial working and long-term memories, including the recognition of places from stored place-codes (Burgess, 2006; Giudice et al., 2013).

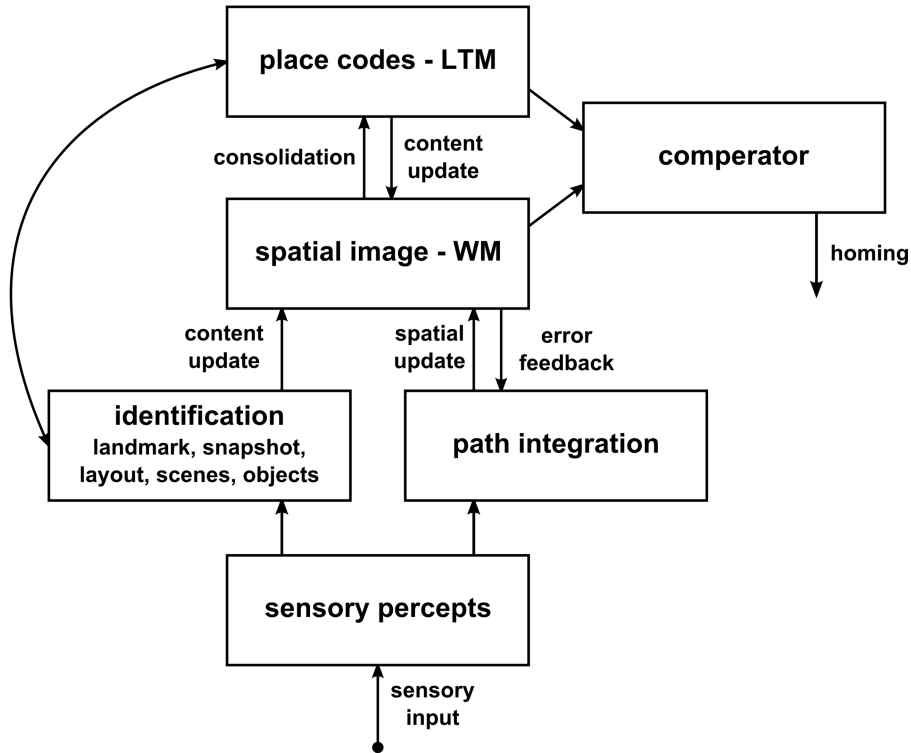
Another case of interaction between the spatial image and long-term memory is the formation of spatial expectations and the increase of the range of spatial awareness (Röhrich et al., 2014). This may even lead to the generation of mental images of distant places as in mental travel (Basten et al., 2012) and episodic foresight (Addis et al., 2007; Schacter and Addis, 2007; Schacter et al., 2007; Suddendorf and Corballis, 2007).

## 1.4. A model of place recognition

While physical space is continuous, mental representations of space are built of discrete entities of various granularity (Appleyard, 1969; Siegel and White, 1975; Tolman, 1948). In mazes, such entities may be sharply defined decision points (Gillner and Mallot, 1998; Kuipers, 1978), which are induced by the discrete structure of the environment and task. In open spaces, however, places (as a cognitive concept) often do not have well-defined boundaries and may cover a substantial spatial extent. Mathematically, places are therefore not just pairs of coordinates, but rather resemble a two-dimensional distribution of ‘belongingness’ to the place in question. Indeed, this is what is plotted as firing fields of place cells (O’Keefe and Dostrovsky, 1971), or error distributions (confusion areas) in homing experiments (Gillner et al., 2008; Jacobs et al., 1997, 1998; Morris, 1981). Such distributions arise naturally from the representation of places by discrete place codes, combined with some comparison operation between stored and current place information (Cartwright and Collett, 1983; Franz et al., 1998). This comparison operation involves the long-term memory place code and a position dependent representation. In this paper, we suggest that the latter is not just the sensory image (snapshot) but rather the content of the spatial image. Thus, place recognition will depend not only on currently visible cues, but also on terminated cues maintained in the spatial image by some updating process.

Figure 1.1 summarizes the assumed place recognition model. It consists of two processing streams extracting information about metric position and feature identification. These two streams resemble the locale and taxon systems of O’Keefe and Nadel (1978). They feed into a joint spatial working memory supporting spatial updating and spatial awareness. Long-term memory place codes are formed from this spatial image at discrete places by fixing a subset of working memory content in long-term memory. Place recognition is then achieved by a comparison operation acting the current spatial image and the long-term memory place codes. Further interactions between long-term memory and working memory exist, e.g. for memory-based updating and mental travel, as well as landmark-based reduction of error accumulation in path integration, etc., as discussed above.

Interactions between spatial working and long-term memories are complicated by the fact that the spatial image is largely ego-centric and will therefore change with body movements, whereas long-term memory place codes are fixed and independent of body pose. The comparison stage will therefore need to be invariant to the relative angle between the stored



**Figure 1.1:** Schematic illustration of functional interaction between perception, long-term memory (LTM) place codes and spatial image (WM = working memory). After the percept disappears the perceptual input is converted into a coarse spatial image version reduced to local position information. During a spatial relevant task the spatial image constantly receives spatial relevant input from percepts of different modalities and interacts with the long-term memory place codes initiated by identification processes (e.g. landmark recognition, scene recognition).

place code and the current spatial image. Simple schemes to achieve this can be based on labeling of place codes with independent compass information (Möller and Vardy, 2006, e.g.), comparison of ‘views’ each with a limited visual angle (Röhrich et al., 2014), or by using panoramic place codes with low angular resolution (Stürzl and Mallot, 2006). Behavioral evidence for the representation of local place orientation has been presented, e.g. by Mou and McNamara (2002) and Meilinger et al. (2015).

Note that the notion of place suggested here is similar to a vista space in the sense of Montello (1993), containing the space within the current perceptual horizon. Indeed, we suggest that the spatial image contains map-like information of an extended region, as is required for the assumed spatial updating processes. The long-term memory place

code, being a copy of working memory contents, may thus be fairly complex, including data structures such as local maps or the spatial layout of indoor environments. In the experiments described in this study, we will assume that the spatial image and place codes are local maps containing distance and bearings of landmarks at a fixed orientation.

## 1.5. Place recognition in the hippocampal circuit

The discovery of the place cells (O'Keefe and Dostrovsky, 1971), which inspired (O'Keefe and Nadel, 1978) to formulate the theory of the hippocampus as a 'cognitive map'. The cognitive map is proposed to contain metric relations of the environment. O'Keefe and Burgess (1996) showed that an extending of the geometrical dimensions of a rectangular box elongated or split the place fields in the corresponding direction. Thus, indicating that extrinsic cues like boundaries or landmarks influence the location and shape of place fields. The dependence of place fields on the eccentricity of the environmental cues can be explained by the activity of the boundary vector cells (Lever et al., 2009; Solstad et al., 2008) which fire maximal at a certain distance and direction from a landmark or boundary as postulated by the boundary vector model (Barry et al., 2006; Hartley et al., 2000).

In humans, Ekstrom et al. (2003) found place-responsive cells in the hippocampus which showed equivalent firing properties as place cells in rats. Ekstrom also found view-dependent cells mainly located in the parahippocampal region, which are proposed to extract distance and bearing information from specific landmarks. View-dependent cells can be divided in location-dependent view cells which respond to views of a landmark at a specific location, and location-independent views cells which respond to views of a landmark regardless of the current position of the observer. The function of the location-independent view cells could be related to the identification process described in figure 1.1, which triggers the interaction between the spatial image and long-term memory place codes. Whereas the response properties of a location-dependent view cell reflects the definition of a stored view at a place in the sense of the view graph model (Hübner and Mallot, 2007; Röhrich et al., 2014; Schölkopf and Mallot, 1995).

Episodic memory is thought to be involved in the recognition of places by memorizing and recollecting the spatio-temporal context of personally experienced events (Burgess et al., 2002). We assume that the recollection of spatio-temporal context of experienced episodes, i.e. views, plays a crucial role in multiple-view integration (Röhrich et al., 2014).

## 1.6. Event-related potentials in episodic memory

An event-related potential (ERP) is the averaged electroencephalogram (EEG) of time epochs, which are referenced to a certain stimulus or event. The signal to noise ratio is reduced by this averaging process since neural activity which is in phase with the event remains, whereas noise, which is not in phase with event, is canceled out. Many studies investigated ERPs associated to the recognition of objects and words. These studies mainly used an ‘old/new’ task scheme (Rugg and Nagy, 1989; Sanquist et al., 1980; Warren, 1980), in which participants learn a set of items, e.g. words, afterwards a set of new and previously learned items, i.e. old items, are presented. The participants are asked to subsequently recall whether the presented item was old or new. The recognition memory is proposed to be differentiated into two processing streams, the familiarity and recollection process. Recognition based on familiarity is fast-acting, relatively automatic and does not provide qualitatively information about the studied episode. By contrast the recollection process, which is related to the episodic memory, is conceived as slower-acting and more profound which gives rise to consciously accessible information about the prior occurrence of the studied item and the context of the occurrence, such as the location or color of the item (Rugg and Curran, 2007; Yonelinas, 2002). Evidences were found in studies with patients who had lesions restricted to the hippocampus (Aggleton et al., 2005; Holdstock et al., 2002), differential activity patterns of functional magnetic resonance imaging (fMRI) (Davachi et al., 2003; Eldridge et al., 2000), and dissociation of retrieval-related ERP effects (e.g., Düzel et al., 1997)

The ERP signature of recollection and familiarity can be dissociated by functional and topographic variables (Curran, 2000; Rugg et al., 1998). In a study by Rugg et al. (1998) participants performed either a ‘deep’ (generate a sentence with the presented word) or a ‘shallow’ (determine whether the first and last letters were in alphabetic order) task on each word studied. The learning phase was followed by a test phase, where the participants were asked to discriminate old words from new ones. The ERP of correctly recognized deeply studied words evoked a more positively extended wave 400-800 ms after stimulus onset (P600) compared to the ERP of correctly rejected new words. This effect is related to the recollection processing and is termed the ‘parietal old/new’ effect since the effects is maximal at the parietal electrodes. The ‘mid-frontal’ effect, however, is proposed to be characterized by a more negatively extended ERP wave in the time window 300-500 ms (N400) for correctly rejected new words in comparison to the ERP waves of the recognized shallow and deep studied words. According to this finding, recognition of deeply studied

words was based on both recollection and familiarity, whereas shallowly studied words were recognized on basis of familiarity. Similar findings have been reported for object and face recognition by Duarte et al. (2004) and Curran and Hancock (2007).

## 1.7. Purpose of the behavioral place recognition study

Using a simple environment with a goal location surrounded by four landmarks, we designed four behavioral experiments to test the following hypotheses: (i) visual homing and place recognition are based not just on the raw retinal snapshot, but on an intermediate representation (spatial image) integrating over multiple views and providing simple invariances. (ii) In particular, this working memory contains distance and bearing information about the surrounding landmarks and can thus be thought of as a local map.

Experiment 1 addresses the optimal learning method (incidental vs. explicit) to be used in the following experiments. Experiment 2 examines the influence of the salience of landmark colors on the performance of place recognition. In Experiment 3 we manipulated the size of the landmarks in order to test to what extent the participants use landmark bearings or landmark distances. The manipulation results in two possible goal locations, one where the angular sizes of the landmarks match the learned sizes and the other one where the bearings or angles between the landmarks match the angles as seen from the learned goal location, comparable to the Cartwright and Collett (1983) landmark manipulation experiment. As discussed, we expect that the participants will recognize the learned location at the distance-matching location. Our hypothesis that the spatial image metrically represents a local environment by encoding the distance and bearing information of landmarks will be examined in Experiment 4. We constructed two additional landmark configurations with different distributions of landmark distances from the goal. We expect that the homing error depends on the geometry of the landmark configuration. By checking homing performance for different starting points, we also analyze the effect of landmark familiarity, i.e. the visibility during encoding. In the discussion section, the results of experiment 3 and 4 will be compared to predictions from a simple maximum likelihood model.

## 1.8. Purpose of the sequence-related ERP study

We designed an experiment similar to the classic old/new task with the major difference that the knowledge about the temporal context is additionally inquired, i.e. the position of a learned item in a sequence of presented items. Further we used indoor scenes as study items to link this study to the integration of multiple views. Schyns and Oliva (1994) suggested that a coarse-to-fine process extracts a coarse description for scene recognition before finer information is processed. In this case the workload for recognizing a scene would not differ substantially from the workload required in object recognition. Thus, the resulting ERP correlates in scene recognition should be comparable to those found in object and word recognition (Curran, 2000; Curran and Hancock, 2007; Duarte et al., 2004; Rugg et al., 1998)

The hypotheses of this study are: (i) The detection of a distortion of a learned sequence evokes a characteristic ERP, which is dissociable from the ERP evoked in the presence of test scenes congruent with the sequence. We expect that the differentiation is characterized by the P600 wave in parietal region on the basis of the well studied parietal old/new effect, which reflects the retrieval of episodic memory content (Rugg and Curran, 2007; Yonelinas, 2002). (ii) Further we propose that subsequently correctly rejected non-target scenes, which are not sequence related, show a mid-frontal and parietal old/new effect compared to the ERPs of correctly recognized target scenes due to involvement of episodic memory content (Rugg et al., 1998).



## 2. Spatial awareness - Beware of what is out of sight

### 2.1. General methods

#### 2.1.1. Participants

In this study 123 healthy participants (61 females and 62 males) took part, which were recruited from the Eberhard-Karls University of Tübingen. At each of the experimental conditions twenty participants took part except for the parallelogram and peaked condition, which contained 16 participants each, and the salience condition, which contained 11 participants. After the participants had been informed about the procedure of the experiment in written form, the participants were asked to sign a consent form. The experiments lasted on average 45 minutes and all participants were paid 8 € per hour.

#### 2.1.2. Apparatus

The rendering and first-person perspective view of the virtual reality environment was realized by the OpenSceneGraph software. The environmental layout, the objects and textures were constructed with MultiGenCreator (MultiGenParadigm).

In the desktop experiments the participant were comfortably sitting at a distance of 80 cm in front of a 30" (44° horizontal field of view) monitor with a resolution of 2560 x 1600 pixel and a display refresh rate of 60 Hz. The movement (forward, backward, left and right sidestep) of the agent was implemented by keyboard buttons. Horizontal movement of the mouse induced a corresponding horizontal rotation of the viewing direction of the virtual

eye.

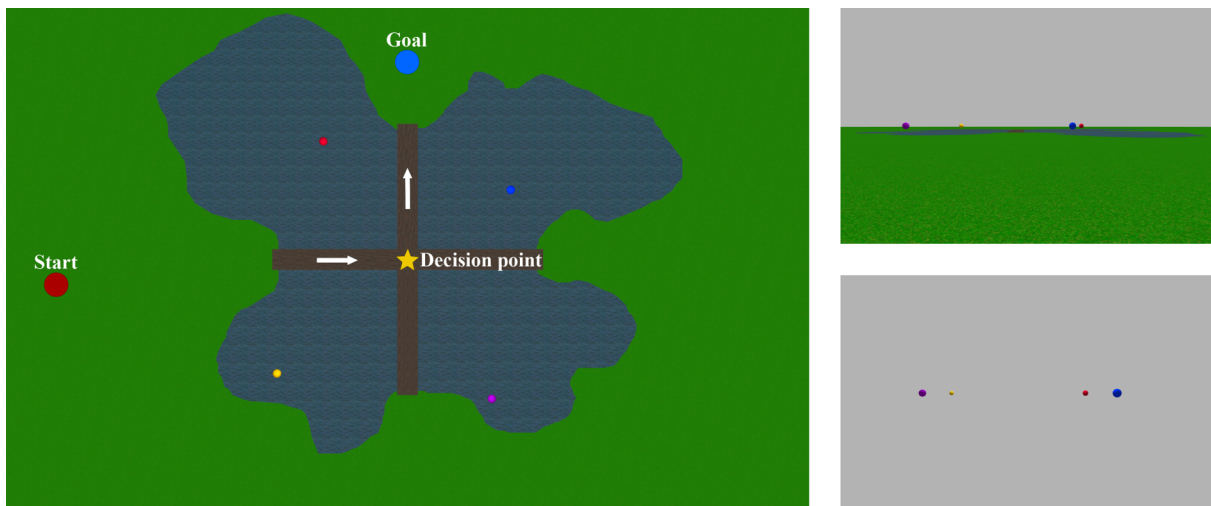


**Figure 2.1:** Setup of the HMD condition. The tracking target on the top of the HMD controlled the rotation of the virtual eye. The actual body direction was tracked by the body belt target, which was coupled with the translation direction of the virtual body (controlled by the mouse in the right hand of the participant).

The HMD condition, which displayed the virtual environment on a Head-mounted display (HMD), was carried out in the central part of a tracking room surveyed by six infrared cameras from ARTtrack & DTtrack System (ART). Two tracking targets with reflecting markers were used. One target was fixed at the top of the HMD to track the rotation of the head. The tracked head direction was coupled with the viewing direction of the virtual eye via LAN-interface (UCP communication) between tracking PC and virtual reality PC. The other target was mounted at the front side of a body belt (figure 2.1). This target made the actual body direction of the participant accessible, which coupled the body direction with the translation direction respectively the direction of the virtual body. The movement of the agent was induced by mouse buttons (forward and backward). We used the nVisor

SX60 HMD with a horizontal field of view of  $44^\circ$  to display the virtual reality (resolution 1280 x 1024 pixel per eye and display refresh rate 60 Hz). The HMD was connected to the virtual reality PC via a four meter cable. A rolling table with the virtual reality PC on top was positioned in a range of one meter next to participant. The participants were asked not to move around but only to rotate their body. The experimenter carefully looked after the participants that they did not walk over the cable. After every trial the experimenter aligned the participant for the new start location by verbal instructions. If the participant was ready, the experimenter started the trial.

For analysis we stored the coordinates of the position, the angle of the virtual eye (camera) direction and the time stamp each frame. The analysis was conducted using Matlab R2013a from Mathworks.



**Figure 2.2:** The bird's eye view of the virtual reality environment is illustrated on the left side including a task example of the learning phase. The participant had to navigate from the start location (red circle) to the goal location (blue circle) by crossing the bridge (white arrows). The crossing point of the bridge was always passed at least once per learning trial. The task of the test phase was to navigate to the crossing point of the bridge (decision point, yellow star). The first person view during the learning phase is illustrated on the top right. The whole environment was covered by ground fog during a fog trial as shown on the bottom right.

### 2.1.3. Virtual reality environment

The virtual reality environment consisted of a centrally located plus-shaped bridge (each arm 32.5 x 5 virtual meters in the following referred to meters) crossing a pond (figure 2.2).

The surrounding green area was flat and the sky was light grey. Both textures (grass and sky) had no detectable features as orientation points. Landmark information was provided by four colored spheres (red, blue, yellow and magenta) with a radius of 1 m, hovering in mid-air above the pod. The center of each sphere was on virtual eye level (1.8 m) to exclude that participants calculate the distance by the trigonometric relation between the eye level and an object (Ooi and He, 2007). The horizontal positions of the landmarks were varied in various experimental conditions. The only available depth cues were the texture gradient, occlusion, motion parallax and visual angle subtended by the landmarks. Around the landmarks a catch area with a radius of 10 m was implemented. Entering resulted in a teleport back to the starting point of the current trial. The forward and backward movement speed was 15 m/s and sideward 3 m/s. The virtual eye rotated  $0.1^\circ$  if the mouse cursor was moved by one pixel. Both the desktop experiments and the experiment with the head-mounted display were rendered with a horizontal virtual field of view of  $66^\circ$ .

#### 2.1.4. Procedure

Participants were briefed in written and verbal form about the procedure of the experiment. The experiment was divided into a learning phase and a test phase. The instruction was to navigate from alternating start points over a plus-shaped bridge crossing a pond to reach a goal location beyond the pond. Participants were informed that in some trials of the learning phase and in the whole test phase non-transparent ground fog would cover the bridge, the pond and also the horizon would disappear (figure 2.2). Thus their only orientation points for navigating on the bridge without falling into the water would be the hovering colored spheres. Participants completed 12 learning trials (4 with fog and 8 without) and 48 test trials. At the beginning of the learning trials the instruction for the following trial was displayed as a text. The instruction was to reach a hidden blue goal platform (radius 3 m and height 0.3 m) either on the left or right side beyond bridge. Thus, participants were always required to travel two branches of the plus-shaped bridge, including a turn by  $90^\circ$ , either left or right, at the bridge crossing point. The goal platform was located at a distance of 60 m from the crossing point, in the extension of the final bridge segment. When the virtual agent entered a catchment area of the goal location (radius 50 m), the hidden platform appeared and stayed visible for the current trial. A new trial was initiated when the virtual agent reached the goal platform. Trials started from the four cardinal directions (North, East, South and West) each one time with fog and two times without. Start locations were

varied both in distance and direction to the crossing point by adding a random jitter ( $\pm 10$  m in x and y coordinates), but the direction was never strictly aligned with the crossing point. There was no time limit during the learning phase, thus the participants could get familiar with the steering of the virtual agent and the virtual reality environment. The test phase followed immediately after the learning phase. Here, the task was to navigate to the crossing point of the bridge as exactly as possible, starting twelve times from each compass direction (as before, position jitter was added to the start position). Participants confirmed their decision by mouse click (decision point), which also initiated a new trial. If participants accidentally clicked already at the start location, the data point was excluded from the analysis. In the test phase the participants received no feedback about their performance and the environment was covered by ground fog during the entire phase of the experiment.

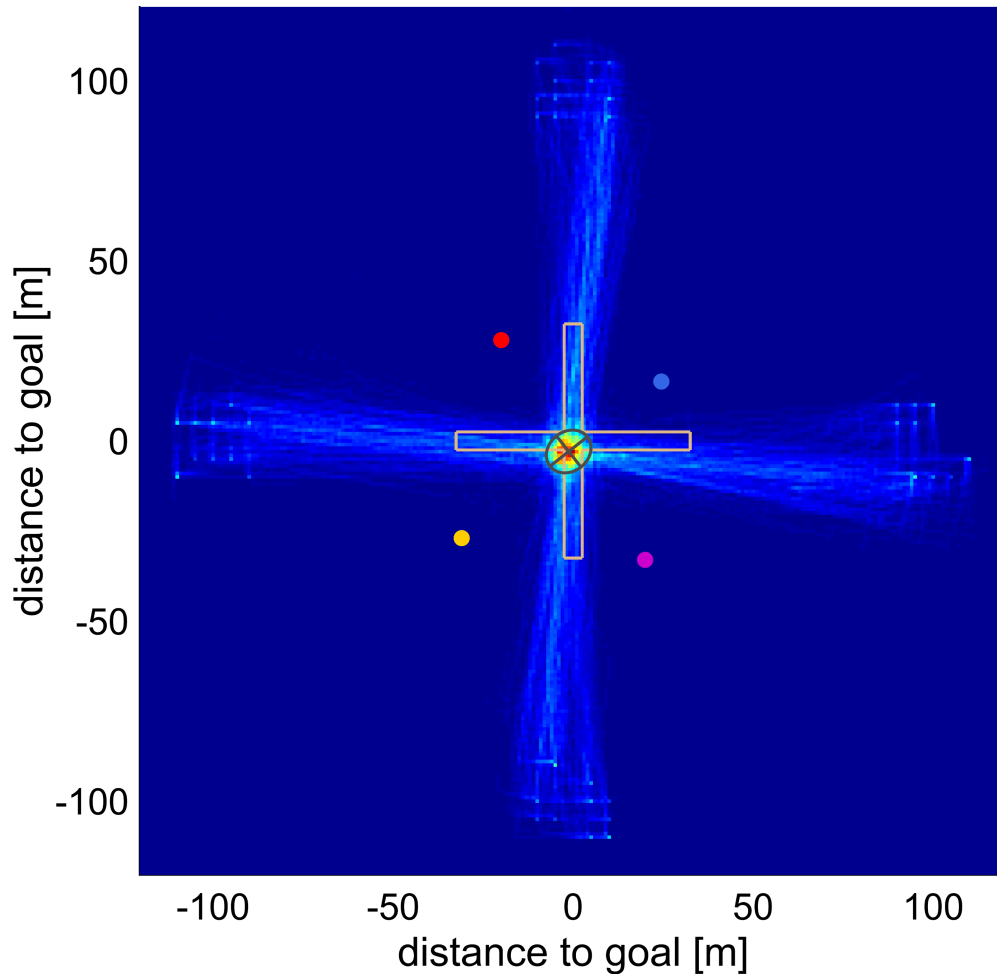
## 2.2. Experiment 1: Implicit versus explicit learning

This experiment examines the influence of incidental and explicit learning on place recognition to establish the learning scheme for the following experiments. The learning procedure described above was used as a control condition in which learning of the bridge crossing point occurred incidentally, i.e. while searching for the blue goal platform. We now added an ‘explicit’ condition in which subjects were told from the start that they would eventually be asked to navigate to the crossing point in the test phase. Except for this additional instruction, the procedure was unchanged.

### 2.2.1. Methods

The landmark configuration formed an irregular quadrangle with the landmark positions (in meters relative to the crossing point of the bridge) red (-20/28), blue (24.5/16.5), yellow (-31/-27) and magenta (20/-33), see figure 2.2.

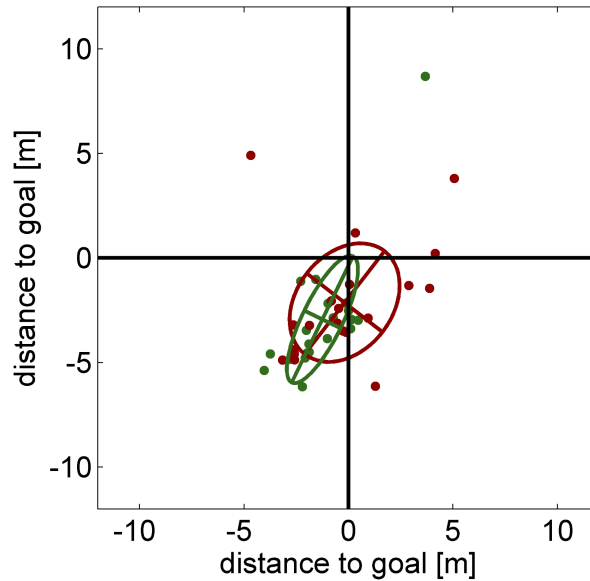
The control and explicit condition were carried out in a between subjects scheme with 20 participants in each condition (control: 8 females, 12 males; explicit: 13 females, 7 males).



**Figure 2.3:** Heat map of trajectories of individual trials. Data from the control condition is shown in which the participants learned incidentally.

### 2.2.2. Results

In total, 954 decision points for the control condition and 956 for the explicit condition have been analyzed. Figure 2.3 shows heat maps of the trajectories sampled in 1 x 1 m bins. A bin is incremented if a data point of a trajectory falls within the corresponding square. Data points of subsequent frames falling in the same square are not counted. Landmarks are shown as colored dots corresponding to their color in the experiments. The heat maps show that the trajectories are mostly straight, i.e. that the subjects directly approach their intended decision point from all starting points. The heat maps peak at the mean of the



**Figure 2.4:** Mean of the decision points of each participant for the control (green) and the explicit condition (red) including the error ellipses.

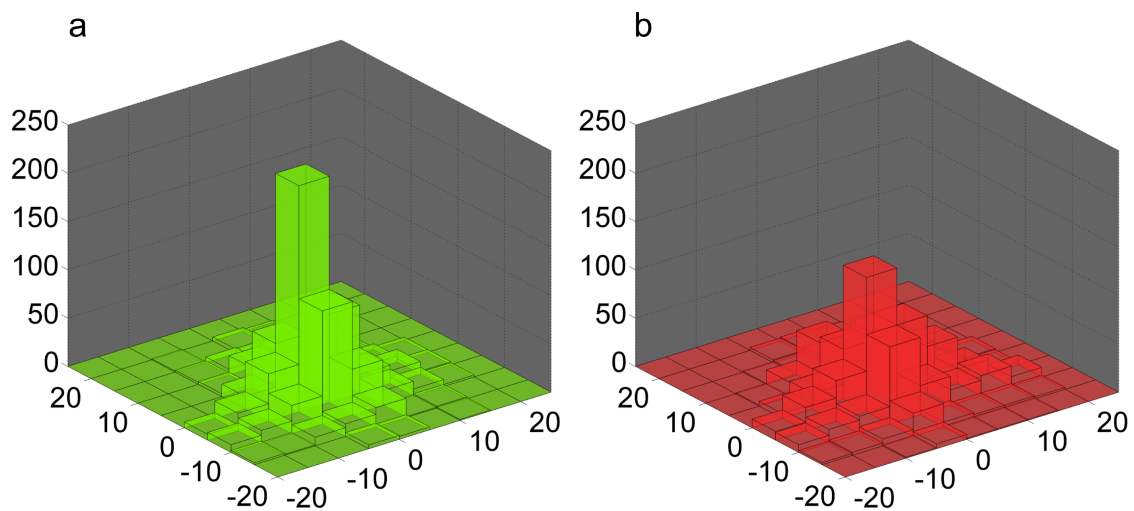
single decision point judgements (error ellipse), indicating that trajectories from different starting points run straight to the common goal.

Figure 2.4 shows the average decision points of the individual subject. One-sample Hotelling-t-squared test between the average decision points for each participant and the true crossing point confirmed a significant bias for both conditions (control:  $T^2(1,18) = 18.45$ ,  $p < .001$ ; explicit:  $T^2(1,18) = 11.73$ ,  $p < .013$ ). The mean bias of the reported decision point was  $(1.25, -2.93)$  for the control condition and  $(-0.19, -2.13)$  for the explicit condition. A two-sample comparison of these conditions did not reveal a significant difference ( $T^2(1,38) = 2.32$ ,  $p = .335$ ). The distributions in figure 2.4 appear to be more concentrated and less isotropic for the control group (green dots and error ellipse) as compared to the ‘explicit’ group. To test this, we cumulated the distributions of decision points over all subjects in each group and plotted them as histograms in figure 2.5. The statistical analysis of these histograms reveals a significant difference between the distributions confirming this assumption ( $\chi^2(49) = 147,61$ ,  $p < .001$ ).

### 2.2.3. Discussion

The results indicate that the place codes built in incidental and explicit learning differ,

presumably in the amount or type of local position information represented. Statistical errors are smaller in the control condition. On top of the statistical variation, subjects show a significant bias into South-West direction of the setup, which is probably related to asymmetries of the landmark configuration since the landmark in the South-West direction was the one with the largest distance to the decision point. This bias is the same for both conditions. For the following experiments, we decided to use the incidental learning scheme throughout because the more concentrated distributions of average decision point judgements allow to distinguish between various quantitative hypotheses for these experiments.



**Figure 2.5:** (a) 3D-histogram of the decision points for the control condition partitioned in  $5 \times 5$  m squares. (b) The distribution of decision points is less pooled in the explicit group.

### 2.3. Experiment 2: Landmark salience

If the reported systematic bias is elicited by the salience of the landmark coloring alone, a change of the coloring of the landmarks by rotating the color configuration would also lead to a corresponding rotation of the bias, since in this hypothetical case the landmark position would not matter. Thus, we rotated the coloring of the landmarks while the landmarks remained at the same position as in the control condition.



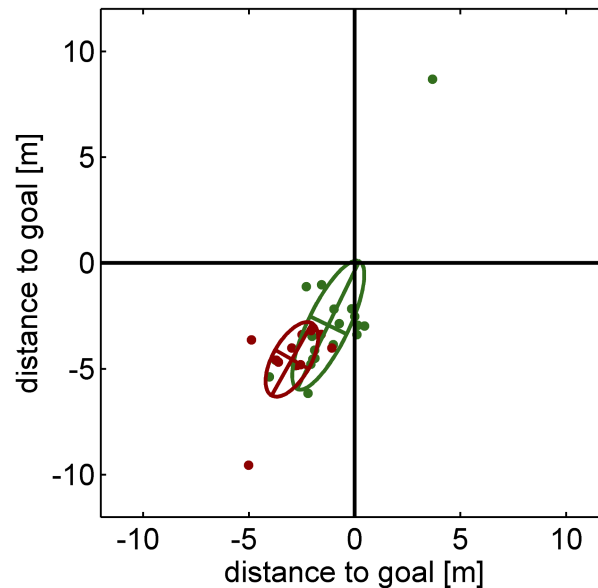
### 2.3.1. Method

As in Experiment 1 the landmark configuration formed an irregular quadrangle with the same landmark position, but in this experiment the coloring of the landmarks was rotated by  $180^\circ$  (magenta (-20/28), yellow (24.5/16.5), blue (-31/-27) and red (20/-33)). The procedure remained the same as in the control condition described above.

This experiment, named the ‘salience condition’, was carried out with 11 participants (6 females and 5 males).

### 2.3.2. Results

We analyzed 525 decision points for the salience condition. Figure 2.6 illustrates the mean decision points of each participant for the salience group (red) and the control group (green). The salience group showed a significant bias to the true crossing point in South-West direction tested by a one-sample Hotelling-t-squared test ( $T^2(1,9) = 85.85, p < .001$ ). The mean decision point was (-2.97/-4.53) and a two-sample Hotelling-t-squared test between the individual mean decision points of the salience and control condition revealed a significant difference ( $T^2(1,28) = 9.37, p = .02$ ).



**Figure 2.6:** Mean of the decision points of each participant for the control (green) and the salience condition (red) including the error ellipses.

### 2.3.3. Discussion

The rotation of the coloring of the landmarks showed that the salience of the colors had an effect on the decision point bias. However, the mean decision point was even more biased in South-West direction, which can not be explained by a color salience approach since the direction of the bias did not switch. We assume that the salience of the coloring has an influence on the systematical bias, but that the source of the bias is mainly based on geometrical properties of the environment. The coloring of the landmarks will stay constant in the following experiments, therefore, a experimental manipulation of definite geometrical properties, which result in a change of the distribution of decision points compared to other experimental conditions, is only related to these geometrical manipulations.

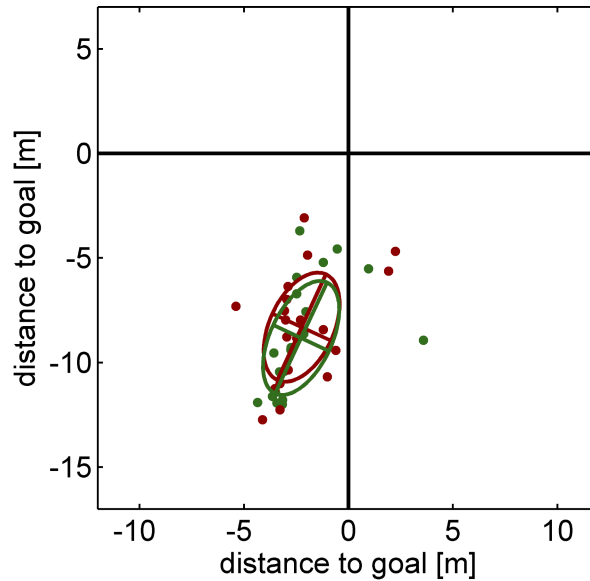
## 2.4. Experiment 3: Distance versus bearing

The purpose of this experiment was to investigate the relative role of bearing and distance information in the place code underlying homing and place recognition. To this end, the sizes of the landmarks were manipulated between the learning and test phase of the experiment. This manipulation renders the decision ambiguous since subjects might search either at the point where the bearing match, or at the point where the visual angles subtended by the landmarks match. The theoretical bias  $B_T$  under the assumption that the visual angles are matched is calculated as

$$B_T = \arg \min_x \sum_{i=4}^4 (\alpha_i(0) - \alpha_i(x))^2 = (-0.78 / -7.02), \quad (2.1)$$

where  $\alpha_i(0)$  is the visual angle of landmark  $i$  as seen from the center of the bridge during the learning phase, and  $\alpha_i(x)$  denotes the visual angle of landmark  $i$  in the test phase, as seen from vectorial position  $x$ . The theoretical bias under the assumption that the bearing are matched is of course zero.

This experiment was carried out in a desktop and in a HMD condition providing additional proprioceptive information on body pose that might allow better integration of local views. In particular, relative bearing exceeding the field of view should be perceived more accurately in the HMD condition. Details of the HMD presentation are described in the general methods section above.



**Figure 2.7:** Individual mean decision point judgements for the desktop (green) and HMD condition (red) including the error ellipses.

### 2.4.1. Method

The landmark configuration was the same as in Experiment 1, but the landmark sizes were manipulated between the learning and the test phase. The sizes of the two Northern landmarks (red and blue) were increased by 15 percent and the sizes of two Southern landmarks (yellow and magenta) were decreased by 15 percent. After the experiment the participants were asked if they noticed any change of the environment between the learning and test phases and if so, what kind of change had occurred. If they did not report the manipulation, they were asked directly for possible changes in landmark size. The experiment was carried out in a desktop and a HMD condition, with between subject design and group size of 20 participants for each condition (desktop: 10 females, 10 males; HMD: 9 females, 11 males).

### 2.4.2. Results

Overall 953 decision points were analyzed for the desktop condition and 958 for the HMD condition. None of the participants noticed the change of the landmark size even after the explicit inquiry. Figure 2.7 shows the average decision point judgements per subject for the

two conditions. The shapes and position of the error ellipses were almost identical as is confirmed by a Hotelling-test ( $T^2(1,38) = 0.45$ ,  $p = .803$ ).

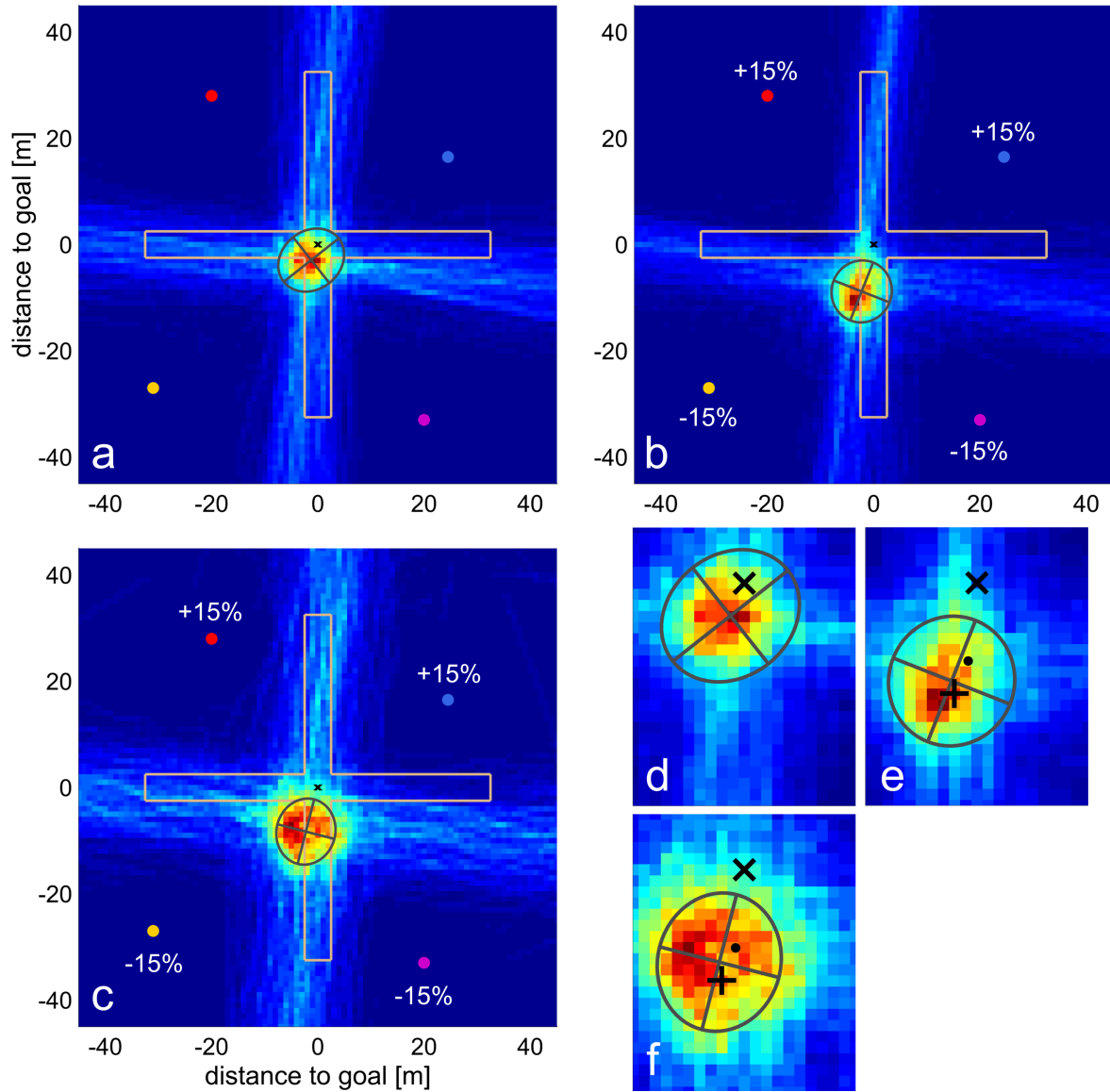
Both the desktop and HMD group showed a bias which significantly exceeds the bias found in the control condition of Experiment 1, as was confirmed by two-sample Hotelling tests (desktop:  $T^2(1,38) = 50.2$ ,  $p < .001$ ; HMD:  $T^2(1,38) = 41.16$ ,  $p < .001$ ). The mean bias of the decision point was  $(-2.27/-8.82)$  for the desktop condition and  $(-2.25/-8.3)$  for the HMD condition. This bias, however, is also different from the theoretically expected bias  $B_T$  as revealed by one-sample tests (desktop:  $T^2(1,18) = 15.03$ ,  $p = .005$ ; HMD:  $T^2(1,18) = 13.47$ ,  $p = .008$ ).

We therefore tested the assumption that the increase of the empirical bias as compared to the smaller bias found in the control condition of Experiment 1 ( $B_C = (-1.25/-2.93)$ ) is the result of the size manipulation. We therefore tested the distributions against the sum of the control bias  $B_C$  and the theoretical bias  $B_T$ . Indeed, the statistical comparison of the average decision point judgements with  $B_T + B_C = (-2.03/-9.95)$  did not reach significance for the desktop condition ( $T^2(1,18) = 6.89$ ,  $p = .073$ ) and only weak significance for the HMD condition compared to  $B_C$  ( $T^2(1,18) = 11.62$ ,  $p = .014$ ).

The Error ellipses in figure 2.8 show the distribution of single decision point judgements superimposed to the trajectory data. The true crossing point is marked by a black x. The zoomed-in heat map around the error ellipse center illustrates the South-West bias of the control group of Experiment 1 ( $B_C$ , figure 2.8a), and the larger biases found in Experiment 3 (figure 2.8b and c). The mean of the individual decision point judgements (center of the error ellipses) in Experiment 2 align with the sum of the control and theoretical bias (marked by a +-sign), not with the true decision point (x) or the pure theoretical bias (dot). As in the control condition the heat maps show that participants directly approach their intended decision point from all starting points in the desktop and HMD condition (see zoom-out view in figure A.1 and A.2).

### 2.4.3. Discussion

Our results show that in an enclosed landmark configuration humans prefer landmark size (visual angle) over relative landmark bearing (visual field position) in cue conflict situations. This is in line with the assumption that distance information is included in the place code and spatial image. Note however, that distance information which might also be obtained from



**Figure 2.8:** Heat maps of trajectories of individual trials. (a) Data from the control condition in Experiment 1, included here for comparison. The x in the marks the true bridge crossing point. (b) Desktop condition. (c) HMD condition. Zoomed-in heat maps with error ellipses of the decision point judgements from Experiment 1 (d), the desktop condition (e) and the HMD condition (f). The dot marks the theoretical bias  $B_T$  (equation (2.1)) and the plus marks the combined bias  $B_C + B_T$ . For further explanation see text.

motion parallax during approach to the decision point is not affected by our experimental manipulation. We therefore assume that the distance information in the spatial image is obtained from visual angle.

This result is in contrast to the results obtained in honey bees by Cartwright and Collett (1983) in which landmark distance was ignored in conflict situations. This may be related to the fact that in panoramic vision, landmark bearings spanning large visual angles can be judged more accurately than in limited field-of-view systems where such judgements need to rely on view integration. Note that in our enclosed landmark configuration, such large bearing angles will always occur. In limited field of view systems, depth information of individual landmarks is easier to obtain than inter-landmark bearings. Thus, we suggest that the human spatial image does contain depth information on which place recognition is based.

The bias of average decision point judgements found in Experiments 1, 2 and 3 may also be related to landmark depth since the center of gravity of the landmark configuration used in these experiments is not on the true decision point (bridge center). In the next experiment we address this possibility by testing various landmark configurations

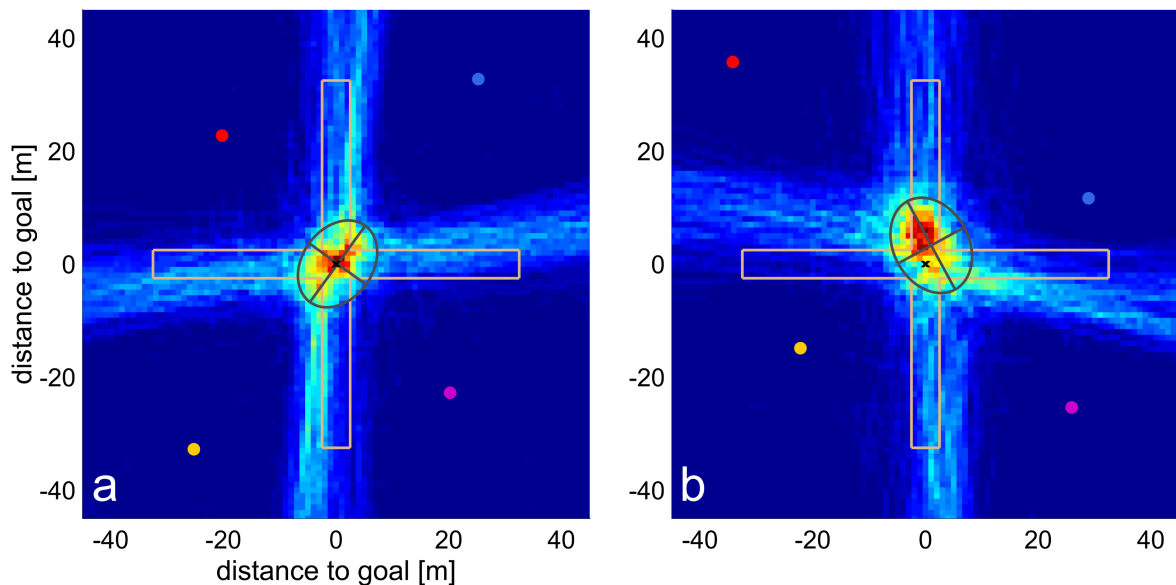
## **2.5. Experiment 4: Dependence of place recognition on landmark configuration**

If distance information is used in the place recognition process, various schemes for depth evaluation are possible. For example, the accuracy of bearing and distance information of a given landmark scales differently with landmark distance: bearing becomes less sensitive to small shifts of the viewpoint whereas discrimination thresholds for distance increase (Erkelenz, 2015; Foley, 1972; Gilinsky, 1951). Configurational features such as the center of gravity (O’Keefe, 1991) or the intersection of the diagonals might be well represented points with respect to which position could be evaluated. Depth distribution may also cause the position bias towards the most distant landmark found in Experiments 1, 2 and 3.

To clarify how depth information enters the place recognition process, we designed two additional landmark configurations, termed ‘parallelogram’ and ‘peaked’. For the parallelogram condition we expected no systematic bias due to the symmetry of the landmark distribution. However, the error ellipse should be elongated along the longer diagonal of

the parallelogram, since depth estimates towards the more distant landmarks should be less accurate. In the peaked condition, the distance of one landmark (the North-Western) from the crossing point was markedly increased with respect to the other three landmarks. If the systematic bias is caused by imbalanced distances from the landmarks to the crossing point, and therefore imbalanced accuracies of depth measurements, the systematic bias should occur in this configuration.

In this experiment, we additionally analyzed the decision points depending on the four different approach directions in which the place recognition will be based on different working memory contents. To this purpose, we counted the number of time-frames in which a given landmark was inside the field of view (i.e. displayed on the screen). For example, for the approach from the South, we expect that the northernmost landmarks have a longer visibility time than the Southern landmarks which will be passed well before reaching the bridge center. Therefore the place decision when approaching from the South will be determined mostly by the Northern landmarks and only to a lesser extent (mediated by visual working memory) from the Southern landmarks.



**Figure 2.9:** Heat maps of the trajectories for (a) parallelogram condition and (b) the peaked condition. Error ellipses (black) are elongated towards the more distant landmark objects (colored dots). The mean of the peaked group is systematically biased in the North direction from the center of the bridge (black cross).

### 2.5.1. Method

The parallelogram condition and the peaked condition differed only in the used landmark configuration. The landmark positions of the parallelogram configuration were point symmetrically (red (-20.25/22.75), blue (25.25/32.75), yellow (-25.25/-32.75) and magenta (20.25/-22.75), see figure 2.9a). The peaked configuration was an irregular quadrangle (red (-34.18/35.76), blue (28.94/11.66), yellow (-22.18/-14.86) and magenta (25.94/-25.36), figure 2.9b) in which one landmark was substantially more distant from the crossing point than the other three. The landmark configurations of all conditions were constructed such that the sum of the distances from each landmark to the crossing point of the bridge was equal in all configurations.

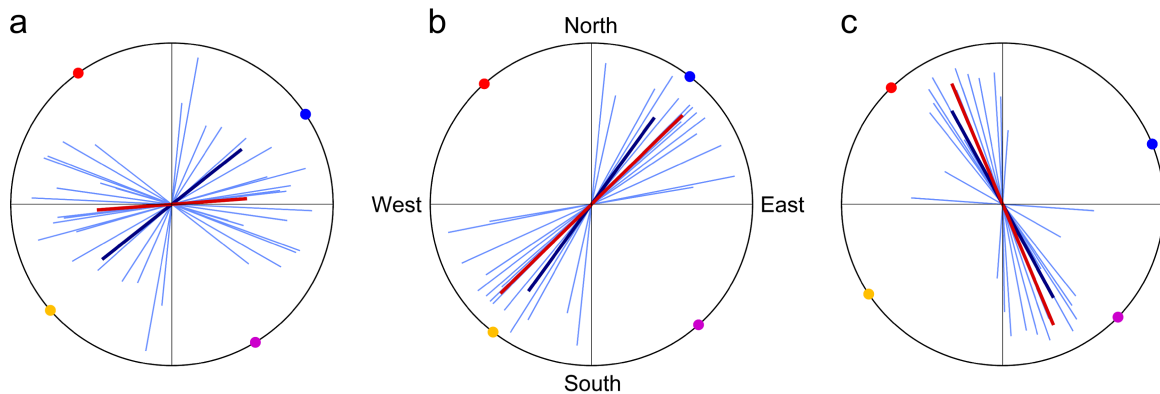
The experiment was carried out in a between subject design with 16 participants in each condition (parallelogram: 8 females, 8 males; peaked: 7 females, 9 males). Presentation was on the desktop setup.

### 2.5.2. Results

Figure 2.9 shows the results with the same conventions as figure 2.3 and 2.8. Again participants directly approach their intended decision point from all starting points (see zoomed-out view in figure A.3 and A.4). The error ellipse of the decision points in the parallelogram group (761 decision points) is elongated towards the more distant landmarks as expected. The mean decision point was not significantly different from the crossing point ( $T^2(1,15) = 0.94$ ,  $p = .655$ ) indicating that no position bias occurred in the landmark configuration (mean decision point (0.25/0.05)). In contrast, the irregular configuration in the peaked condition did result in a significant systematic bias of the mean decision point ( $T^2(1,15) = 37.95$ ,  $p < .001$ ) in the Northern direction, i.e. roughly, but not exactly towards the most distant landmark (mean decision point (0.99/3.37)). The shape of the error ellipse for the peaked group (754 decision points) was elongated towards the more distant landmarks as well as for the parallelogram condition.

The elongation direction of the error ellipses was analyzed in more detail by calculating individual error ellipses for each subject. The first eigenvectors of these ellipses were weighted with the numerical eccentricity of each ellipse and are plotted as light blue lines in figure 2.10. Circular means were calculated by averaging the weighted first eigenvectors and are shown as heavy red lines. Also included in figure 2.10 are the landmark bearings (color coded)

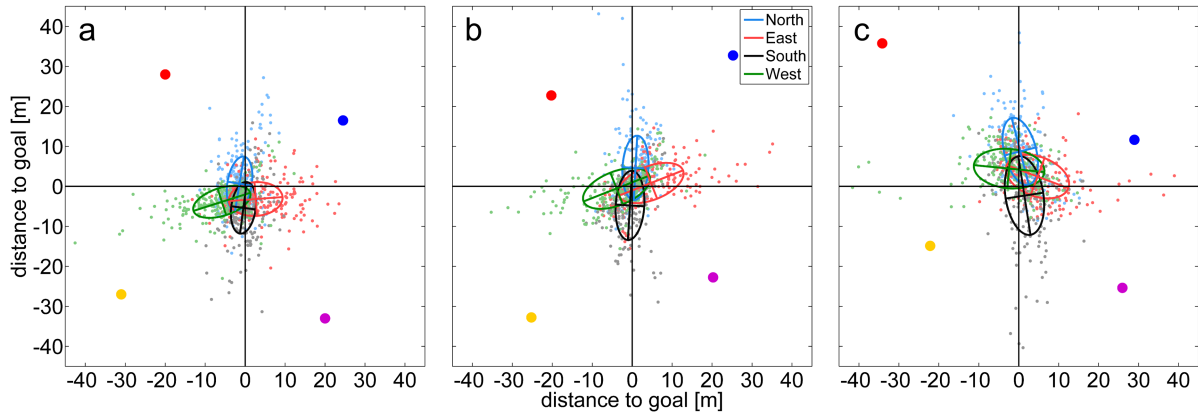




**Figure 2.10:** The circular relation between the landmarks (colored dots on the circle) and the main axes of the participant’s error distribution (light blue) are illustrated. The main axis overall decision points is drawn in dark blue and the circular mean of the single main axes in red. (a) Control condition of Experiment 1. The individual main axes are almost isotropically distributed. (b) Parallelogram condition, (c) peaked condition. In b and c, the main axes are focused in the direction of the most distant landmarks.

and the eccentricity-weighted first eigenvector of the overall error ellipses from figure 2.8a and 2.9, appearing as heavy blue line. Note that the data are restricted to a half circle and therefore appear reflected at the origin in figure 2.10. For analysis with circular statistics, they were converted to unimodality by doubling all angles (Batschelet, 1981). A circular V-test (CircStat toolbox, Berens, 2009) revealed a significant orientation preference of the individual main axes towards distant landmarks both for the parallelogram group and the peaked group (parallelogram: blue  $V(15) = 19.26$ ,  $p < .001$ ; yellow  $V(15) = 19.26$ ,  $p < .001$ ; peaked: red  $V(15) = 13.85$ ,  $p < .001$ ; magenta  $V(15) = 12.98$ ,  $p < .001$ ). The control group showed also significant p-values for the blue but not for the yellow landmark. However, the p-values of control group are less obviously as the parallelogram and distant group (control: blue  $V(19) = 6.49$ ,  $p = .039$ ; yellow  $V(19) = 3.61$ ,  $p = .164$ ).

The heat maps of the trajectories (figure 2.3, 2.8 and 2.9) clearly show that the approach direction of the participants is strongly target-oriented. The participants directly head to a small area, where they expect the crossing point. We split the set of all trials into groups for the four approach directions, where ‘North’ means approach from a Northern starting point, etc. Figure 2.11 shows scatter plots of all decisions, color coded for approach direction, together with the according error ellipses. Average decision points for each approach direction differ in the sense that the subjects stop short of the overlap regions of the four directional distributions, resulting in the characteristic four-leaved clover pattern. Two-sample Hotelling

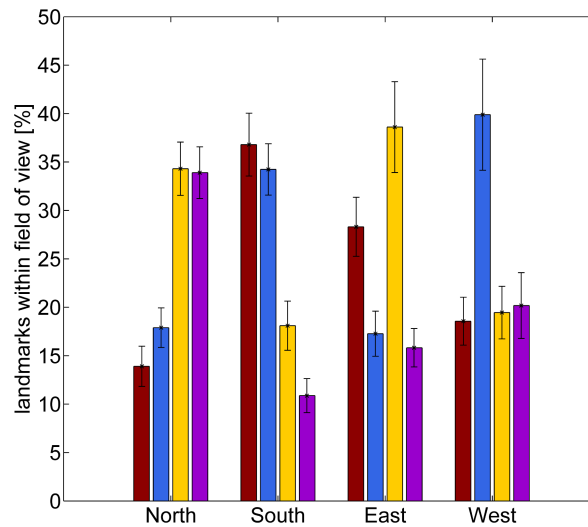


**Figure 2.11:** Decision points (small colored dots) and error ellipses separately displayed depending on the compass direction of the start location. Error ellipses are biased and elongated towards the approach direction for (a) the control, (b) parallelogram and (c) peaked condition.

tests of pairs of opposing approach directions revealed significant differences for all conditions and direction pairs (control: North vs. South  $T^2(1,38) = 24.6$ ,  $p < .001$ ; West vs. East  $T^2(1,38) = 30.29$ ,  $p < .001$ ; parallelogram: North vs. South  $T^2(1,30) = 17.24$ ,  $p = .001$ ; West vs. East  $T^2(1,30) = 15.28$ ,  $p = .003$ ; peaked: North vs. South  $T^2(1,30) = 19.98$ ,  $p < .001$ ; West vs. East  $T^2(1,30) = 10.49$ ,  $p = .013$ ).

The directions of the error ellipses for each approach direction are rotated in comparison to the overall error ellipses of the corresponding experimental condition (compare figure 2.11 to 2.8 and 2.9). Thus, the results clearly support the assumption that place decision depends on approach direction. In order to assess possible effects of the varying landmark visibility during the approach, we analyzed all individual approach trajectories and viewing directions from the point of entering a central area of 20 m around the bridge center. Within this area, only one landmark would be visible at any one frame. The percentage of image frames showing each landmark in the field of view for the four approach directions was calculated and averaged over all subjects from the control condition of Experiment 1 and both conditions of Experiment 3. The result appears in figure 2.12. Note that subjects were always able to turn on the spot to examine the whole visual environment and that such examinations are included in the data appearing in figure 2.12.

Four separate one-way Anovas were calculated for the four approach directions and revealed significant differences between the landmarks (North:  $F(3) = 50.72$ ,  $p < .001$ ; South:  $F(3) = 60.29$ ,  $p < .001$ ; East:  $F(3) = 28.96$ ,  $p < .001$ ; West:  $F(3) = 18.92$ ,  $p < .001$ ).



**Figure 2.12:** The colored bars in the graph (corresponding to the color of the landmark objects) show the mean proportion of landmarks in the field of view within a radius of 20 m around the crossing point depending on the approach direction (standard error in black).

A post-hoc analysis with Bonferroni correction revealed that the red and yellow landmarks were significantly longer in the field of view than the other landmarks for the North approach (yellow to blue and red  $p < .001$ ; magenta to blue and red  $p < .001$ ). When participants enter the central area from a given start location, their head direction is aligned to the approach direction (see figure 2.3, 2.8 and 2.9). As a result, the landmarks which require less turning to come into the field of view were the preferred landmarks used for visual orientation. The statistical analysis of the South and East start location resulted also in a significant preference for the landmarks which require less effort to extract visual information (South: red to yellow and magenta  $p < .001$ ; blue to yellow and magenta  $p < .001$ ; East: red to blue and magenta  $p < .001$ ; red to yellow  $p < .002$ ; yellow to blue and magenta  $p < .001$ ). For the West start location a significant preference could only be measured for the blue colored landmark as expected, but not for the magenta colored landmark which also required less movement of the virtual eye to look at (blue to all other  $p < .001$ ).

### 2.5.3. Discussion

The results of this experiment clearly show three effects of the two-dimensional landmark configuration on decision point choices. First, the decision point distributions are elongated

in the direction of the most distant landmarks. In the parallelogram condition, where two landmarks are equally distant, the decision point distribution aligns with the axis connecting these two landmarks. In the peak condition, the distribution is pulled towards the single most distant landmark, resulting in a more asymmetric shape. The distribution is also influenced by the other landmarks, but to a lesser extent.

Second, a systematic bias was found only for the peaked condition, in which case it is directed roughly but not exactly towards the most distant landmark. This finding is in line with the landmark-distance effect in the sense that symmetric landmark configurations will generate symmetric decision point distributions. These observations lead to the hypothesis that both the shape and the bias of the distribution of the decision points is defined by probabilistic weightings of the landmark positions. Therefore the proposed spatial image may represent a local environment as a map-like set of probabilistic measurements of the containing landmarks.

The analysis of the proportions of landmarks in the field of view showed evidence for an imbalanced acquisition of visual landmark information depending on the approach direction. Regarding the shift towards the start location of the error ellipses of each approach direction, this leads to our interpretation that the representation of landmarks in the spatial image depends not only on the probabilistic measurement achieved by perception, but also on their updating frequency. Thus the reliability of the local position information of the spatial image decreases with time and has to be renewed to recover the reliability.

## 2.6. General Discussion

In summary the experimental results show evidence for our hypothesis that place recognition is based on a spatial working memory integrating over multiple views, i.e. spatial image, which represents not only field of view information but also out of sight information. This spatial image builds up a map-like representation containing the distance and bearing information about surrounding environmental cues.

The results of the cue competition between distance and bearing information in Experiment 3 clearly points out that humans focus on the distance information when the goal location is enclosed by a landmark configuration, whereas a panoramic snapshot approach favors bearing information in a comparable rivalry task (Cartwright and Collett, 1983), supporting the multiple view approach. Moreover the fact that the distribution of decision

points in Experiment 4 depends on the geometrical array of landmarks determined by the distance of the landmarks to the goal location, indicates that local position information is incorporated into a local map weighted by probabilistic estimates. In addition, the observation of the systematic bias in the irregular configurations can only be explained by the involvement of configurational information.

The displacement of the error distributions of the decision points towards the approach directions can be explained, if landmarks in the local map are represented as probabilistic estimates, which reliability depends on the perceptual updating frequency. Thus, various unbalanced field of view updating of the landmarks results in distinct error distributions of the decision points depending on approach direction. At the same time this means that out of sight information is involved in the place recognition process.

One could argue that the shift of the decision points towards the approach direction is caused by human path integration, where undershooting or overshooting of traveled distance has been observed depending on the range of measured distances (Klatzky et al., 1990; Loomis et al., 1993; Riecke et al., 2002; Schwartz, 1999). But the results of Experiment 3 clearly figure out that participants rely dominantly on visual information, which led to the shift towards the theoretical goal location where the distances match, even in presence of proprioceptive information. Relying on path integration would not have resulted in such shift. Nevertheless the role of path integration in this study could be imprecise spatial updating of the self-position relative to the landmarks induced by range effects in magnitude estimation (Petzschner and Glasauer, 2011; Stevens, 1961; Teghtsoonian and Teghtsoonian, 1978) causing an variance increase of the probabilistic estimate of out of sight landmarks.

### 2.6.1. Likelihood-based spatial representation

Integration of sensory information can be modeled by estimating the maximum likelihood of the individual sensory estimates (Ernst and Banks, 2002; Ernst and Bühlhoff, 2004). Pickup et al. (2013) proposed a multi-view scene reconstruction model based on the probabilistic estimates of a set of cameras pointing to landmark locations of a viewed scene, but not taking into account the landmarks outside the field of view. The homing is modeled by comparing the current scene reconstruction with the stored reconstruction at the goal location to compute a map of the maximum likelihood estimate.

We share the opinion that the local position information may be coded by probabilistic

estimates of spatially relevant metric information, here landmarks. Thus we propose that not only view-dependent scenes are reconstructed but rather a local map of the probabilistic landmark estimates is stored as a place code in long-term memory. During homing the stored cognitive model of the environment is embedded into an egocentric reference system, i.e. the spatial image. This spatial image constantly updates the own position by visual and proprioceptive information as well as the local position information of the landmarks by perceived sensory inputs.

Our results clearly show that place recognition takes into account the position, i.e. distance and bearing, of all four landmarks, even if they are out of sight during the final approach. We therefore assume that the place code used for recognition consists of the four vectors  $\vec{l}_1, \dots, \vec{l}_4$  pointing from the goal location to the four landmarks. We also assume that these vectors are given relative to a fixed, geocentric ‘North’ direction induced in the experiment by the landmark configuration. Place recognition will rely on the comparison of this place code with the current measurements  $\vec{m}_1, \dots, \vec{m}_4$  of the landmark position, again taken in an egocentric, but geo-oriented sense. We assume that these measurements taken at a position  $\vec{x}$  are distributed normal with means  $\vec{\mu}_1, \dots, \vec{\mu}_4$  and covariance matrices  $\vec{C}_1, \dots, \vec{C}_4$ , where both the means and the covariance matrices will depend on  $\vec{x}$ . The goal is recognized if  $\vec{m}_i = \vec{l}_i$  for  $i = 1, \dots, 4$ , i.e. if the current measurements equal the stored place code. The probability density for this to happen at a location  $\vec{x}$  is

$$p\left(\vec{m}_1 = \vec{l}_1, \dots, \vec{m}_4 = \vec{l}_4 \mid \vec{x}\right) = \prod_{i=1}^4 \phi\left(\vec{l}_i; \vec{\mu}_i(\vec{x}), C_i(\vec{x})\right) =: L(\vec{x}), \quad (2.2)$$

where  $\phi\left(\vec{l}_i; \vec{\mu}_i, C_i\right)$  is the standard normal probability density function with mean  $\vec{\mu}_i$  and covariance matrix  $C_i$ . We denote this quantity by  $L(\vec{x})$  and observe that it is the likelihood of being at position  $\vec{x}$  when measuring the stored place code  $\vec{l}_1, \dots, \vec{l}_4$ . The model predicts that the mean of the recognized locations be  $x^* = \arg \max_x (L(\vec{x}))$  (maximum likelihood estimator) and that the error distribution should be proportional to  $L(\vec{x})$ .

In the model of equation (2.2), the place code as well as the means and the covariances of the measurements together with their dependence on position are all free parameters. We start by assuming that the place code is identical to the true landmark positions, as seen from the goal position. This may be justified by the fact that the subjects have ample motion parallax information when approaching the goal. For the mean values, an obvious choice would be to also set them to the true vectors from the current position to the goal,

$\vec{\mu}_i(\vec{x}) = \vec{l}_i - \vec{x}$ . In this case, however, the four distributions multiplied in equation (2.2) will all be centered at the veridical goal position  $\vec{x} = 0$  which will therefore be the maximum likelihood estimator for the recognized place.

Given that we find systematic deviations between the true and the recognized goal position in all but the parallelogram conditions of Experiment 4, we arrive at the unexpected conclusion that the measurements of landmark position done during the approach will systematically deviate from the landmark positions stored as a place code in long-term memory. This conclusion, however, is in line with studies by Philbeck and Loomis (1997) and Fukusima et al. (1997) who show that distance estimates in verbal reports as well as in blindfolded walking to a previously seen target (walking without vision) are roughly veridical, while measurements of Weber fractions in visually perceived distance can be explained by the Stevens' power function (Stevens, 1961). The value of the exponent decreases with an increasing range of tested distances and also outdoor environments have a decreasing effect on the exponent (Da Silva, 1985). Our experiment involved a large range of distances and took place in a virtual outdoor environment comparable to experimental conditions used by Gilinsky (1951). Gilinsky has been shown that visually perceived distances follow a hyperbolic compression law.

More specifically, we model the deviation between true and perceived landmark position by the 'Gilinsky-compression',

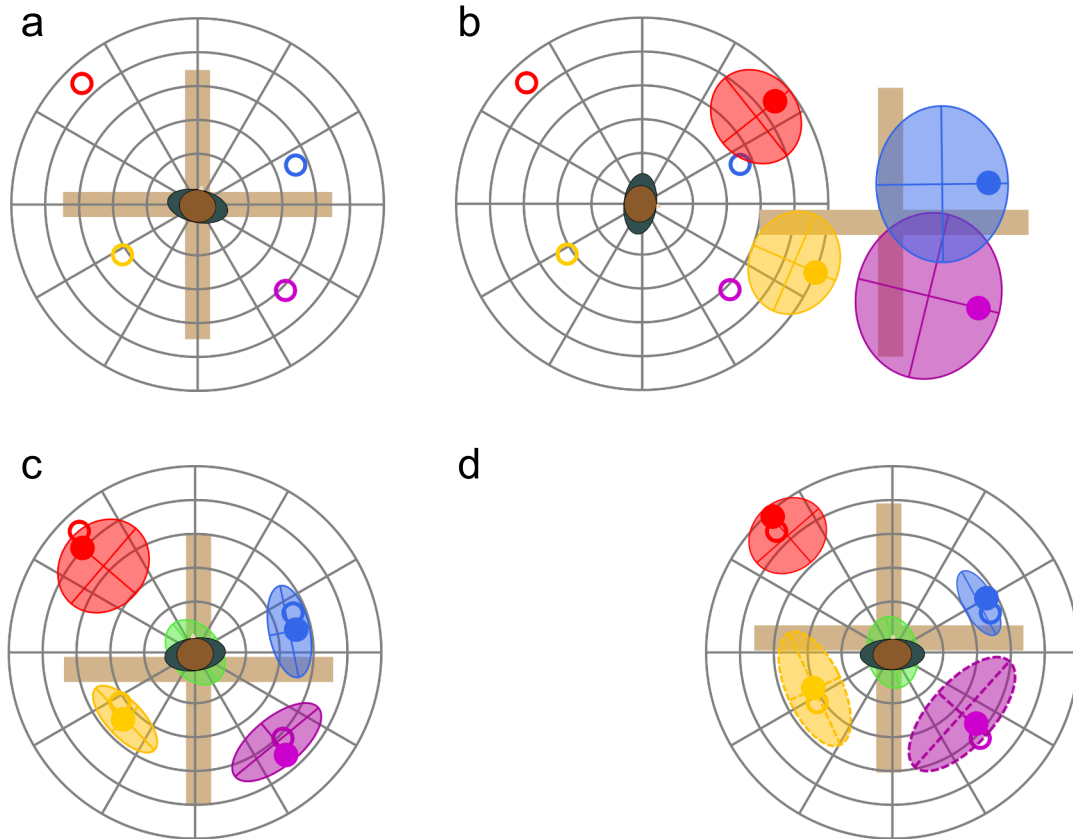
$$g_A(d) = \frac{Ad}{A+d} \quad (2.3)$$

for some suitable constant  $A$ , where  $d$  stands for each of the  $\|\vec{l}_i - \vec{x}\|$  values. In two dimensions we thus assume that when standing at a position  $\vec{x}$ , a landmark  $\vec{l}_i$  is actually perceived at:

$$\vec{\mu}_i := \frac{A(\vec{l}_i - \vec{x})}{A + \|\vec{l}_i - \vec{x}\|}, \quad (2.4)$$

i.e. at the same line of sight with perceived distance hyperbolically distorted. The constant  $A$  (the distance of 'infinity') was set to 90 m.

The covariance matrices of the individual landmark distributions were chosen such that one main axis of the error ellipse would be aligned with the current viewing direction towards this landmark,  $\vec{l}_i - \vec{x}$ . The variance in this direction would thus reflect the uncertainty of the depth estimate while the variance along the second main axis corresponds to the error



**Figure 2.13:** Schematic illustration of the model. (a) Participant stores the distance and bearing estimates of the landmarks  $\vec{l}_i$  (colored open circles) at the center of the bridge (goal location) as a place code. These estimates are veridical due to triangulation. (b) When approaching the goal location later in the test phase the perceived probabilistic distance and bearing estimates (colored error ellipses) of the landmarks (colored dots) are represented by the spatial image, whereas the distance estimates are hyperbolically compressed ( $\vec{\mu}_i$ ). (c) The homing is achieved by multiplying the non-euclidean probabilistic estimates of the spatial image at the veridical goal location  $\vec{x} = 0$ , which is the maximum likelihood estimator for the recognized place (green error ellipse). Note, in the case of an asymmetric configuration this always leads to a bias in direction of the most distant landmark since there is the largest mismatch to the true landmark position. (d) If the percept of a landmark is terminated, the position estimate  $\vec{m}_i$  is retained in the spatial image. We assume that the quality of the position estimates decreases with time. Thus, the weighting of the variances is correlated with the duration of landmarks out of sight. For example, a participant approaching from South direction looks more often at landmarks, which require less turning near the goal location relative to the approach direction (red and blue landmark). Thus, the variances for landmarks, which are more often out of sight, are increased (dotted error ellipses of the yellow and magenta landmark), while the others are decreased.



in bearing estimates. The covariance matrices thus have the form

$$C_i = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \lambda_d & 0 \\ 0 & \lambda_b \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \quad (2.5)$$

where  $\phi = \text{atan}(\vec{l}_i - \vec{x})$  is the turning angle towards the landmark, and  $\lambda_d$  and  $\lambda_b$  are the standard deviation in the depth and bearing estimates, respectively. We assume that with increasing distance, the distance error  $\lambda_d$  grows faster than the bearing error  $\lambda_b$  and model this effect by power functions of distance

$$\lambda_{id} = s_d \|\vec{l}_i - \vec{x}\|^4 \quad (2.6)$$

$$\lambda_{ib} = s_b \|\vec{l}_i - \vec{x}\|^2 \quad (2.7)$$

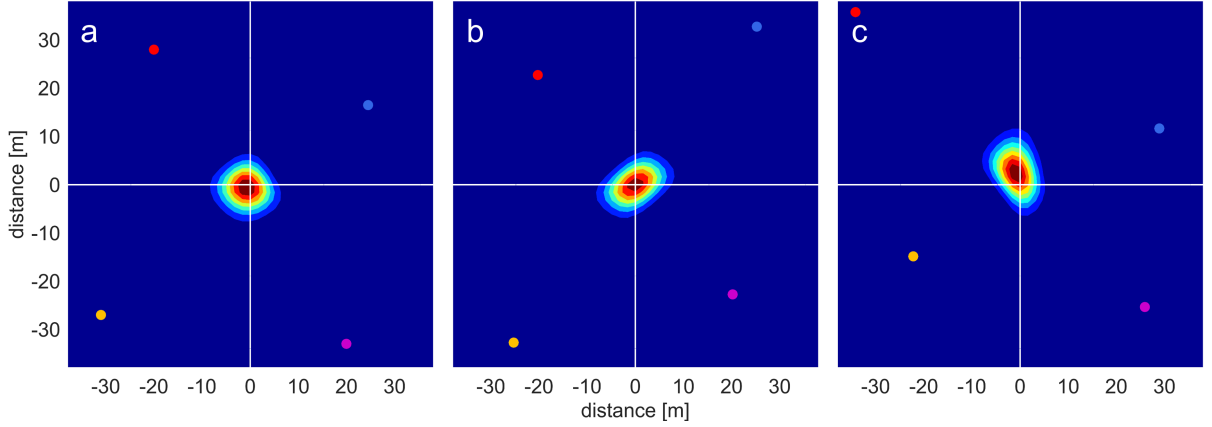
where the constants  $s_d$  and  $s_b$  are fixed for all landmarks. The powers are chosen such that the angle subtended by the distributions in equation (2.5) for each viewing position  $\vec{x}$  is identical, as is the relative distance error  $\Delta d_i/d_i$ . The angular errors of the different landmarks are assumed to be independent of each other, i.e. they are not due to errors about a common ‘North’ direction since only individual landmarks will be within the field of view during a final approach phase. The proposed likelihood-based model is evaluated by comparing the behavioral data collected in the three different landmark configurations with the simulated data of the model.

### 2.6.2. Evaluation of the model

As assumed above  $\lambda_b > \lambda_d$  should be true for near and  $\lambda_b < \lambda_d$  for far distance estimates. Therefore we chose  $s_d = 0.0001$  and  $s_b = 0.1$ . As a result,  $\lambda_b < \lambda_d$  is true for  $\mu < 31.63$  which equates to a distance of 48.76 m.

Figure 2.14 shows the simulated data of the model for the three different configurations which were used in the psychophysical experiments (control, parallelogram and peaked). The combined likelihood of the control configuration revealed a bias in South-West direction towards the most distant landmark and a circular shape. The error ellipse in the control condition is more elongated towards the yellow landmark in comparison to the simulated likelihood, but the analysis of the orientation of the main axes showed that the main axes of participants are scattered isotropically (figure 2.10). The combined likelihood in the

parallelogram configuration is centered over the goal location and elongated towards the more distant landmarks. The peaked configuration revealed again a bias toward the most distant landmark like in the control configuration and the shape is elongated towards the more distant landmarks comparable to the parallelogram configuration.



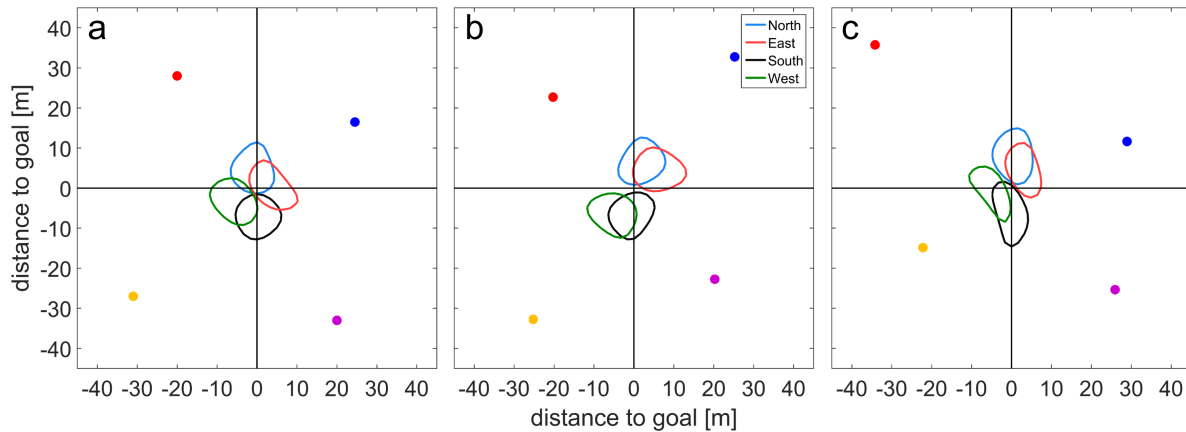
**Figure 2.14:** Combined likelihood of being at location  $\vec{x}$  if all four landmark positions (colored dots) are perceived for (a) the control, (b) parallelogram and (c) peaked configuration.

We simulated the approach effect of the experimental data, i.e. the bias and the elongation of the error ellipses towards the approach direction, by scaling  $s_d$  and  $s_b$  in equation (2.6) and (2.7) with the relative visibility durations  $v_i$  for each landmark configuration (see figures A.5, A.6 and A.7); the average visibility durations over all landmark configurations are reported in figure 2.12. The leaves the error unchanged if  $v_i$  equals 0.25 and is reduced for longer visibility durations

$$s_{di} = s_d (1 - (v_i - 0.25))^\alpha, \quad s_{bi} = s_b (1 - (v_i - 0.25))^\alpha. \quad (2.8)$$

The factor  $\alpha$  is set to 8. Note that the original parameters are reproduced if all  $v_i$  equal 0.25. Figure 2.15 shows that for all configurations the combined likelihoods are biased towards the approach direction. There is also a tendency that the orientations of the combined likelihoods turn towards their approach direction. The combined likelihoods qualitatively match the systematic biases of the experimental data as well as the shape and orientation of the error ellipses. The simulation of the approach effect showed analogous biases towards the direction of approach as observed in the experimental data. But the orientations of the combined likelihoods were not as clearly turned towards the approach direction as the error ellipses of the experimental data.

The simulations show that a configuration of landmarks can be represented in the spatial image as a probability function which is weighted by the distances to the landmarks. The appearance of the systematic bias in the experimental data can be traced back to comparison step during homing between the non-Euclidean source code of the spatial image and Euclidean place code in long-term memory. Further the simulation of the approach effect indicates that the visual updating frequency of landmarks determines their reliability in the spatial image.



**Figure 2.15:** Combined likelihood of being at location  $\vec{x}$  depending on the approach direction (shortcuts of the compass direction) for (a) the control, (b) parallelogram and (c) distant configuration.

## 2.7. Conclusion

We conclude that in a local environment the spatial image is a map-like representation built up by perceptual information and maintained local position information. The probabilistic distance estimates of the spatial image are derived from the non-Euclidean perceptual space, whereas estimates of the place codes are close to Euclidean space. In addition, the frequency of perceptual updating determines the reliability of the position estimates of environmental cues.



## 3. Neural correlate of sequence-related retrieval of scenes

This study focused on the retrieval-related ERPs concerning a previously learned sequence of scenes. Therefore participants learned a sequence of scenes embedded in a storyline, which told that they walk through a building by passing several hallways, which always led to a room. Participants were asked to remember the walked route. Afterwards a sequence of the same length was presented and the participants were asked to respond if the presented room was at same position in the sequence of visited places on their walkthrough. The visited rooms (following termed as target scenes) during the walkthrough could be presented in the recognition task either by matching the position of the corresponding hallway of the walkthrough or not. Mainly we are interested in the ERPs of correctly recognized target scenes matching the learned sequence and correctly recognized mismatching target scenes, expecting a dissociation of the ERPs in the P600 wave in the parietal region. Additionally we expect a mid-frontal effect and parietal old/new effect regarding the ERPs of correctly rejected non-target scenes and correctly recognized target rooms, which is related to the recollection process, i.e. episodic memory.(Rugg and Curran, 2007; Yonelinas, 2002)

### 3.1. Methods

#### 3.1.1. Participants

In this study 20 healthy participants (14 females and 6 males) took part, which were recruited from the Eberhard-Karls University of Tübingen. Each participant completed two experimental sessions at different days, which were at least spaced by 48 hours. One

experimental session lasted on average two hours, including half an hour of EEG preparation. All participants were paid 8 € per hour. After the participants had been verbally informed about the procedure of the experiment, the participants were asked to sign a consent form.



**Figure 3.1:** Experimental setup after EEG preparation. Participant was sitting in front of the experimental PC.

### 3.1.2. Apparatus

The study took place in dimly lit and electromagnetically shielded room. On an EEG recorder PC the amplified EEG data was stored at a sampling rate of 500 Hz as well as displayed for the experimenter (Vision Recorder software from Brain Products), but not visible for the participant. An experimental PC was used for the stimulus presentation of the procedure. Participants were comfortable sitting at a distance of 80 cm in front of a 19"

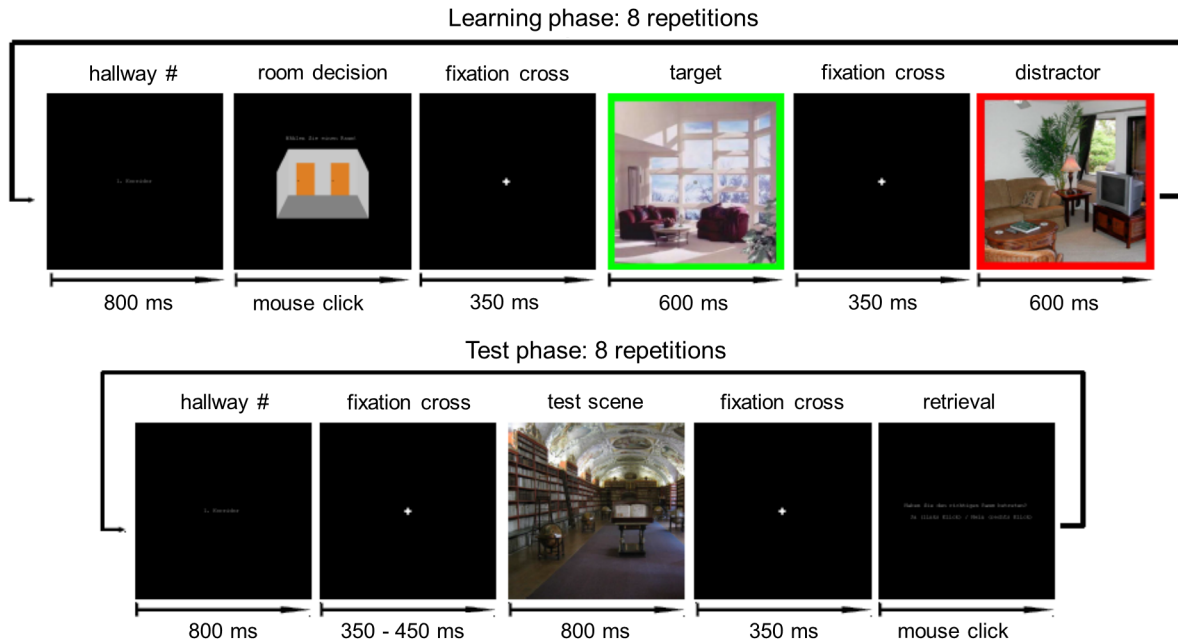
LCD monitor (visual angle  $27^\circ \times 21^\circ$ ) of the experimental PC with a resolution of 1280 x 1024 pixel and a display refresh rate of 60 Hz. The experimental PC was connected to an EEG amplifier (V-Amp 16 channel from Brain Products) by LPT ports to synchronize the sent trigger signals with the EEG data. The EEG was recorded from 16 active Ag/AgCl electrodes (actiCAP from Brain Products) placed after the 20-10 system on an elastic cap, which fitted adequately the circumference of the participant's head. The electrooculogram (EOG) was recorded from an electrode placed on the orbicularis oculi muscle below the right eye without restricting vision. The EEG data collection only started until the impedance of each electrode was reduced below 20 k $\Omega$ . All electrodes were referenced to a electrode placed on the left mastoid process. Participants were asked to put on an armband with an electrode touching the skin, which was connected to earth.

### 3.1.3. Procedure

The experiment consisted of 2 x 50 trials and three additional training trials in the first session. Each trial was divided into a learning phase and test phase (figure 3.2).

The learning phase was initiated by showing the lettering "learning phase" for 2.5 s. During the learning phase eight hallways each with two doors were shown, announced in advance by the number of the current hallway. In each hallway the participants had to choose one door by mouse click, which they wanted to pass. This decision had no impact on the further presentation, but was included to focus attention on the subsequent scene. The participants were asked to decide as fast as possible. After this decision two pictures of indoor scenes were presented for 600 ms. The first was the target scene, which participants had to detect in the test phase. The target scene was marked by a green frame (30 pixel wide). The second picture showed the distractor scene and was marked by a red frame (30 pixel wide).

The test phase followed immediately after all eight hallways had been presented and was initiated by showing the lettering "test phase" for 2.5 s. During the test phase, all hallways were tested with regard to their chronological order in the learning phase. The number of the current hallway was presented as a cue followed by the test scene, which was presented for 800 ms. The presentation of the test scene was temporally masked by randomly varying the duration of the previously shown fixation cross between 350 to 450 ms. In a 'Yes/No' task, participants were asked to hit the corresponding mouse button (left, "yes"; right, "no")



**Figure 3.2:** Schematic overview of the procedure. A single hallway presentation in the learning phase consisted of the presentation of the hallway number, the room decision display, the target scene (green frame) and the distractor (red frame). The participants learned a sequence of eight hallways. In the following test phase also a sequence of eight test scenes were presented. Participants had to recall, if the presented test scene matches the target scene of the corresponding sequence position in the learning phase, i.e. same hallway number.



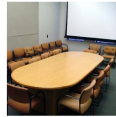




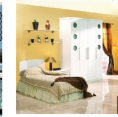



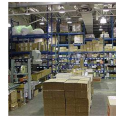
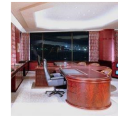



if the presented scene was the target scene they encountered in the corresponding hallway during the learning phase (figure 3.3). The participants could take their decision (click the mouse button) not before the retrieval screen appeared (1150 ms after test scene onset). Participants did not get any feedback of their performance, except for the training trials where they should get used to the task.

Fifty percent of the presented test scenes were sequence matching target scenes known from the learning phase (correct identity and position) and the other 50 percent were homogeneously distributed into distractor scenes of the same hallway (false identity and correct position) which always were from a different category as the target scene, new scenes (same or different category as target scene, each 50 percent) which were not presented in the learning phase (false identity and position) and target scenes which did not match the corresponding hallway (correct identity and false position). The same pictures did not occur in the same trial twice and the stack of all available pictures was only recycled if all pictures were shown



once.

Participants were asked to avoid blinks during the presentation of the indoor scenes as well as to fixate the cross at center screen before the onset of the scene presentation and to keep the fixation on the gray point at center of the pictures. Further they were asked to position the computer mouse on their right thigh in such way that their arm is in a relaxing position during the experiment. However, participants were told that these actions are necessary to avoid artifacts in the EEG measurement and that they could take a break between trials to relax.

hallway #	1	2	3	4	5	6	7	8
learning phase (targets only)								
test phase								
	●	●	●	●	●	●	●	●
"yes"	conjunction hit	fa	fa	conjunction hit	false hit	fa	conjunction hit	conjunction hit
"no"	miss	cr	cr	miss	disjunction cr	cr	miss	miss

**Figure 3.3:** Example of possible behavioral responses during one trial. The target scenes of the learning phase could be either in conjunction (green dot) or disjunction (yellow dot) with those of the test phase. A non-target scenes (red dot) could be a distractor, or a new scene from the same or different category as the corresponding target scene. The table below points out the results of the possible responses (false alarm = fa, correct rejection = cr).

### 3.1.4. Stimuli

We used a collection of 96 pictures (MIT Indoor Scene Recognition Database) in this study, which had a resolution 256 x 256 pixel and were scaled up to 700 x 700 pixel (visual angle 14.8° x 14.8°) for the presentation on the monitor against a black background. At the center of the pictures was a gray fixation point with a visual angle of 0.3° x 0.3°. The pictures were indoor scenes from eight categories each consisting twelve pictures (office, living room,

bedroom, kitchen, library, warehouse, hobby room and conference room). The model of the hallway (visual angle  $8.4^\circ \times 6.3^\circ$ ), the white lettering presentation screens and the fixation cross (visual angle  $0.6^\circ \times 0.6^\circ$ ) were also shown against a black background. The inscription above the hallway model inscription above was “Choose a room!”. The presentation screen of the hallway numbers was labeled by e.g. “hallway 1.”. The stimulus presentation was realized with Matlab R2013a from Mathworks and the Psychophysics Toolbox 3 (Brainard, 1997).

### 3.1.5. Data analysis

In the post analysis the EEG data was high-pass filtered at 0.01 Hz and low-pass filtered at 30 Hz using the Vision Analyzer from Brain Products. Only epochs of the test scene stimulus between 200 ms pre-stimulus to 1000 ms post-stimulus were taken into account in the further analysis. Recorded epochs where the absolute potential value of the difference between the maximum and minimum exceeded  $75 \mu\text{V}$  were defined as artifact containing trials and were excluded prior to averaging. The pre-stimulus interval of 200 ms was used for baseline correction. ERPs were analyzed from 9 electrode sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4), where effects of the experimental conditions were most evident and where previous studies which have used similar experimental design have reported such effects (Duarte et al., 2004; Iidaka et al., 2006; Rugg and Curran, 2007; Rugg and Nagy, 1989). Except for the analysis of the scalp maps using EEGLab 13.4.4b (Delorme and Makeig, 2004) where the other placed electrodes were additionally taken into account (F7, F8, T7, T8, P7, P8). The dipole source localization of the scalp maps was calculated by the brain electrical source analysis (BESA) developed by (Hoechstetter et al., 2004).

Note that all ERPs were referenced to the test scene onset and were averaged separately based on the behavioral response of the participant. We termed a correct rejection of a

<b>response</b>	<b>scene at correct position</b>	<b>scene at false position</b>	<b>distractor/novel</b>
yes	conjunction hit	false hit	false alarm
no	miss	disjunction correct rejection	correct rejections

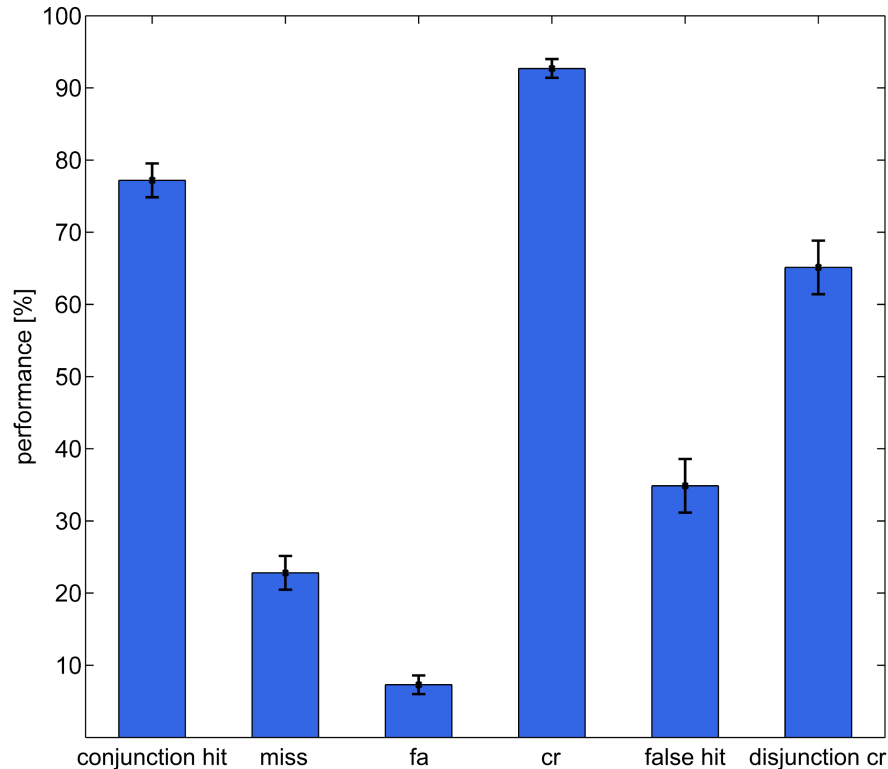
**Table 3.1:** Notation overview of the possible response conditions.

target scene, shown in the test phase, which did not match the corresponding hallway of the learning phase as disjunction correct rejection (disjunction cr) since the positions in the sequences were disjunctive. As a consequence, a positive answer (“yes”) to a target scene which did match the corresponding hallway was termed as conjunction hit. Whereas a positive answer to a target scene which did not match the corresponding hallway was termed as false hit (see table 3.1.5). The other possible behavioral responses were termed according to the standard convention (miss, correct rejection, false alarm).

The false alarm rates in this study were too low for adequate averaging (on average 12.2 artifact free trials) due to the high levels of accuracy (figure 3.4). Additionally three participants were excluded from the false hit grand average process since less than eleven artifact free trials were recorded.

Filtered EEG raw data was imported to the Matlab R2013a programming environment (MathWorks), where artifact rejection, baseline correction and averaging was processed. For statistical analysis the ERP data was divided into five time windows 0-240 ms, 240-380 ms, 380-660 ms, 660-800 ms and 800-1000 ms. The participant’s mean amplitudes were calculated for each time window. On this values we performed repeated measurement two-way Anova (electrode site x response condition). The significance levels (0.05) of multiple pairwise comparisons were Bonferroni-Holmes corrected. For the analysis of the latency of the grand average P600 waves we performed an ICA (independent component analysis) on the participant’s average waves of the conjunction hit and disjunction correct rejection condition with the Infomax algorithm (Bell and Sejnowski, 1995) implemented in EEGLab 13.4.4b.

The psychophysical data entered a repeated measurement one-way Anova in the case for the different conditions of the correct rejected non-target scenes presented as test scenes. These were the mean correct rejections (cr) of distractor scenes, new scenes from the same category as the target scene or new scenes from different categories. Additionally the mean conjunction hits were separately analyzed depending on the spacing to the correct position of the target scene presented in the learning phase. The significance level (0.05) of multiple pairwise comparisons were Bonferroni-Holmes corrected. The conjunction hit and disjunction correct rejection rate as well as the miss and false hit rate were compared by paired t-tests. All statistics were performed by using SPSS statistics 22 from IBM.

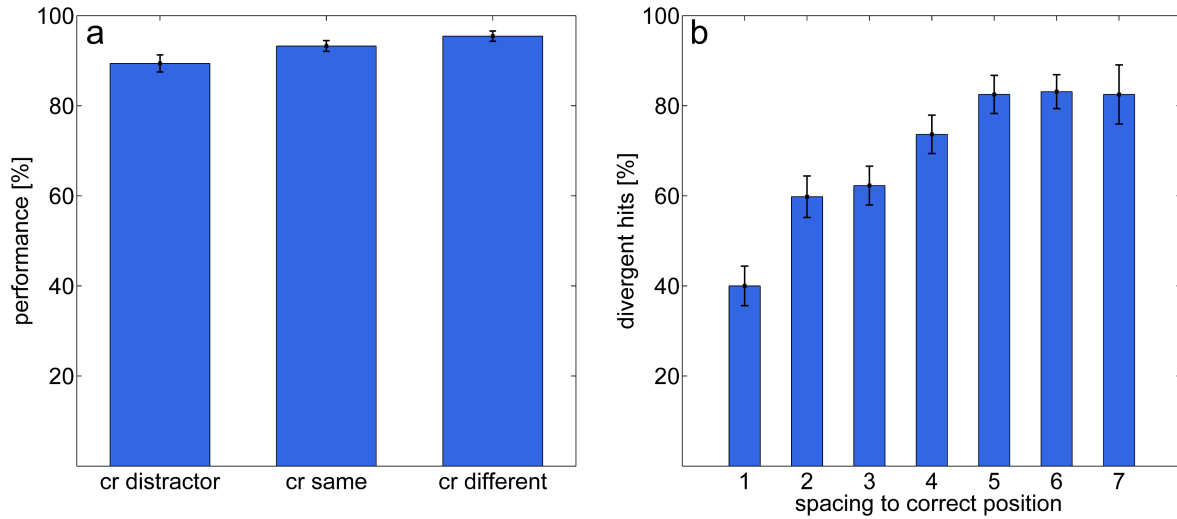


**Figure 3.4:** Bar graph of the mean behavioral responses of the participants including the SEM bars (black lines).

## 3.2. Results

### 3.2.1. Psychophysics

The behaviorally results showed a high accuracy for both the detection rate of conjunctive target scenes and non-target scenes (see figure 3.4 for overview of the behavioral responses). The miss and false alarm rate was accordingly low. The pairwise comparison of the participant's mean values of the conjunction hits and disjunction correct rejections revealed a significant difference ( $t(19) = 2.7$ ,  $p = .014$ ). Participants correspondingly reported more frequently false hits, which were unrecognized position errors of target scenes, compared to the miss rate, which represents the unrecognized target identities ( $t(19) = 2.7$ ,  $p = .014$ ). Note that the disjunction correct rejection rate and the false hit rate can reach at maximum the same rate as the conjunction hit condition since a position failure can only be detected



**Figure 3.5:** (a) Mean correct rejections depending on the presented test scene. A non-target scene could be a distractor scene, a new scene from the same scene category as the target scene or a new scene from different category. (b) Mean conjunction hits depending on the spacing to the correct position of the target scene presented in the learning phase. The SEM bars are drawn in black.

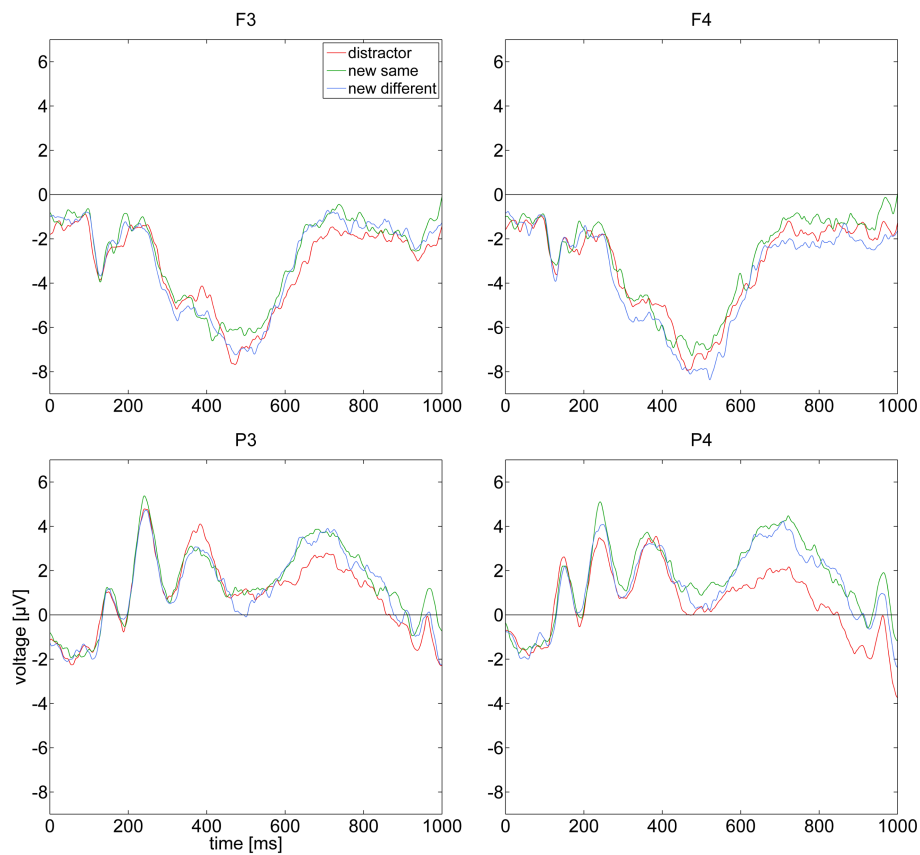
or missed if the identity of the scene was recognized. The other way around the minimum rate for both conditions is the miss rate.

Figure 3.5b shows that target scenes appearing at a false position were significantly more likely to be recognized as the spacing between their presentation in the learning and test sequence increased ( $F(6) = 28.68$ ,  $p < .001$ ). Especially when the spacing differed only by one position between both sequences, which resulted in significant pairwise comparisons to all other spacing distances. Also a spacing of two and three positions differed significantly from the rest. The comparison of position differences of four or more revealed significant differences to the other conditions. For clarification, the sample sizes of the different spacing distances  $n_i$  varied (for each participant:  $n_1 = 18$ ,  $n_2 = 24$ ,  $n_3 = 22$ ,  $n_4 = 22$ ,  $n_5 = 16$ ,  $n_6 = 8$ ,  $n_7 = 2$ ).

A closer look on the different possible types of correct rejections (figure 3.5a) by an one-way Anova revealed a significant difference ( $F(3) = 14.87$ ,  $p < .001$ ). The post-hoc analysis based on a Bonferroni-Holmes correction showed that all correct rejection conditions were significant different to each other (distractor to new conditions  $p < .01$  and new conditions to each other  $p = .029$ ).

### 3.2.2. Event-related potentials

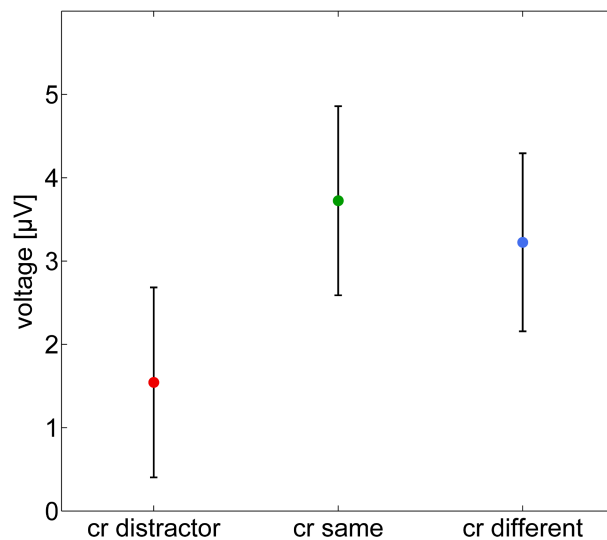
The ERP analysis was concentrated on the factor response condition and on the interactions between the factors electrode and response condition due to the fact that our purpose focus on the distinction between the ERP properties of the response conditions. Thus, we did not report in detail the statistical results of the factor electrode site, which showed in summary significant differences for all comparisons ( $p < .037$ ).



**Figure 3.6:** Grand average ERPs during encoding of distractor scenes (red), new scenes from same category (green) and different category (blue) as the target scene, which subsequently were correct rejected. Time axis is referenced to the test scene onset. The electrode sites are labeled above the corresponding ERP graph.

First we distinguished between the ERPs of the correct rejection conditions, which were the distractor correct rejection (mean epoch quantity (MEQ) = 70.95, standard error of the mean (SEM) = 5.17), new scenes from same (MEQ = 63.7, SEM = 4.5) and different (MEQ = 63.85, SEM = 4.35) category as the target scene. The two-way Anova of all time windows

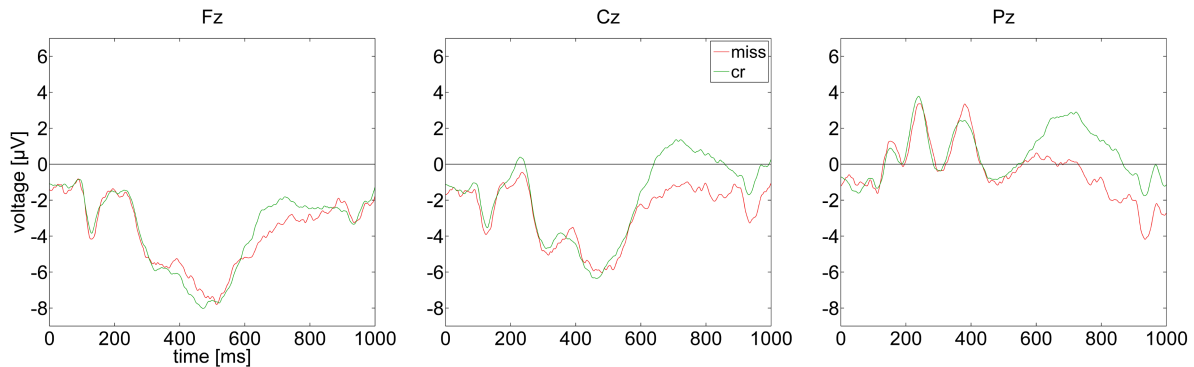
revealed no significant difference between the response conditions. However, the time window 660-800 ms revealed an interaction between electrode site and response condition ( $F(2,8) = 1.95$ ,  $p = .016$ ). Figure 3.6 shows that the correct rejection of a distractor scene evoked a less positive P600 amplitude at the parietal sites. Thus, we performed a more focused comparison of the participant's mean amplitudes for each parietal sites of the time window 660-800 ms. Only the repeated measurement one-way Anova of the P4 electrode site revealed a significant difference ( $F(2) = 6,32$ ,  $p = .004$ ). A Bonferroni-Holmes corrected pairwise comparisons showed that the distractor response condition differed significantly from both new response conditions, but not the new response conditions to each other. However, the three correct rejection conditions were averaged for further analysis since the difference was restricted to the P4 electrode site.



**Figure 3.7:** Mean amplitudes of time window 660-800 ms at the P4 electrode site for the correct rejection response conditions. The colored dots represent the mean voltage values evoked by subsequently correctly rejected distractor scenes (red), new scenes from same category (green) and different category (blue) as the target scene. SEM bars for each response condition are shown as black lines.

The response conditions miss ( $MEQ = 66.3$ ,  $SEM = 8.55$ ) and correct rejection are similar regarding the decision of the participant. In both cases the participants responded negative to a subsequently presented test scene even if the target scene was presented in the case of a miss. Figure 3.8 illustrates that the differences between the miss and correct rejection conditions in the time window 660-800 ms and 800-1000 ms continuously increases from the

frontal to parietal electrode sites. A two-way Anova revealed a significant interaction between the response conditions and the electrode sites for the time windows 660-800 ms ( $F(1,8) = 2.792$ ,  $p = .007$ ) and 800-1000 ms ( $F(1,8) = 2.962$ ,  $p = .004$ ) approving this observation. Additionally the time window 660-800 ms revealed significant difference between the response conditions ( $F(1,8) = 6.48$ ,  $p = 0.02$ ), but the other time windows did not significantly differ ( $F(1,8) < 1.92$ ,  $p > 0.18$ ).

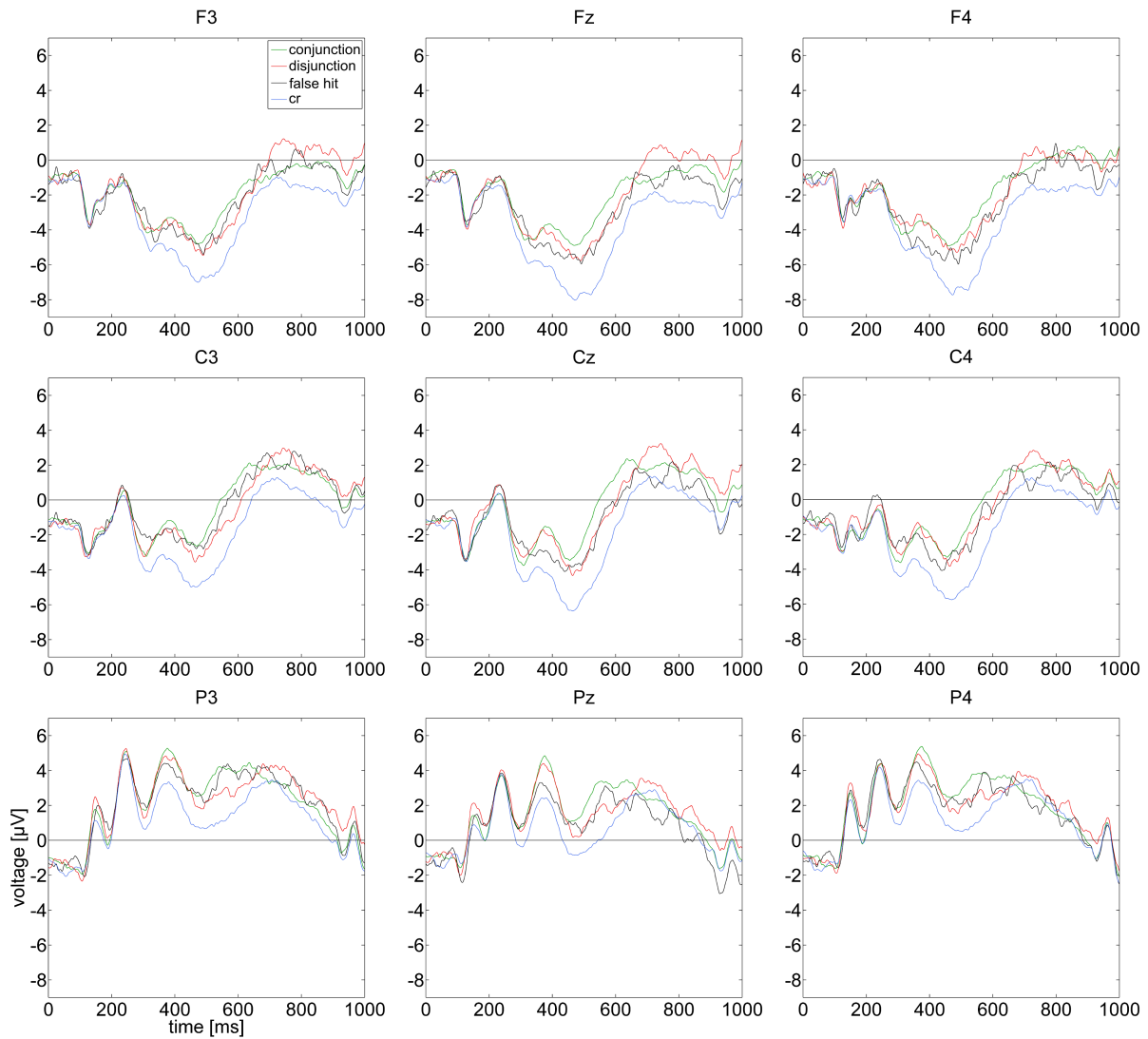


**Figure 3.8:** Grand average ERPs of subsequently missed conjunctive target scenes (red) and correct rejected (green) non-target scenes are illustrated for the electrode sites Fz, Cz and Pz which are labeled above the corresponding ERP graph. Time axis is referenced to the test scene onset.

We expected that the conjunction hit ( $MEQ = 234.45$ ,  $SEM = 15.56$ ) responses evoke a comparable ERP as the false hit ( $MEQ = 29.15$ ,  $SEM = 4.17$ ) responses since both responses correspond to the decision of the participant that the position of the target scene presented in the test phase was corresponding to the position in the learning phase. In figure 3.9 the grand average ERPs are shown for the subsequently correctly recognized conjunctive target scenes (green) and for the false hit condition (black), which were incorrectly recognized as being conjunctive. For analysis the available false hit averages (17 participants after artifact rejection) were compared to the conjunction hit averages of the corresponding participants by a two-way Anova for each time window. Both response conditions did not significantly differ in any time window ( $F(1,8) < 10.09$ ,  $p > .495$ ).

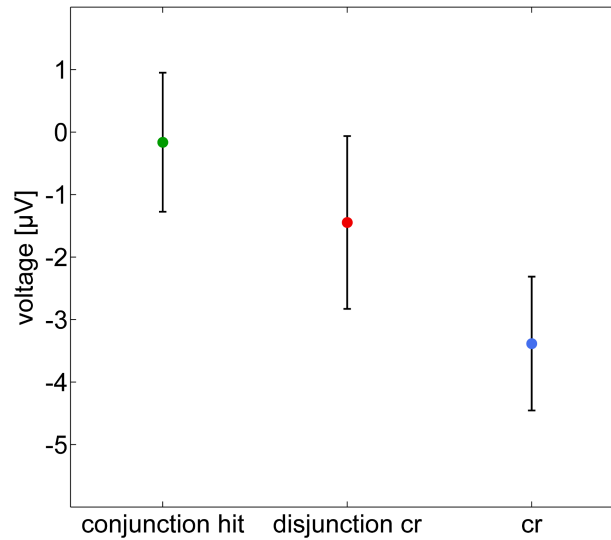
The response condition disjunction correct rejection (drawn in red,  $MEQ = 54.35$ ,  $SEM = 4.72$ ) and correct rejection (blue) are additionally illustrated in figure 3.9. A two-way Anova for each time window with the factor electrode sites and the factor response condition containing the grand averages of the conjunction hits, disjunction correct rejections and correct rejections revealed significant differences for all time windows except for the period 0-240 ms (240-380 ms:  $F(2,8) = 16.48$ ,  $p < .001$ ; 380-660 ms:  $F(2,8) = 24.91$ ,  $p < .001$ ; 660-





**Figure 3.9:** Grand average ERPs of subsequently correctly recognized conjunctive (green) and disjunctive (red) target scenes. Additionally the ERPs of disjunctive targets scenes which were incorrectly classified as being conjunctive, a so-called false hit (black), and the correct rejected non-target scenes (blue) are illustrated. Time axis is referenced to the test scene onset. The electrode sites are labeled above the corresponding ERP graph.

800 ms:  $F(2,8) = 4.69$ ,  $p = .015$ ; 800-1000 ms:  $F(2,8) = 4.93$ ;  $p = .012$ ). More specifically, the pairwise comparisons revealed that the conjunction hit condition significantly differed from the correct rejection condition for these time windows except for the period 660-800 ms. The disjunction correct rejection condition compared to the correct rejections revealed



**Figure 3.10:** Mean amplitudes of the conjunction hit (green), disjunction correct rejection (red) and correct rejection (blue) ERP in the time window 380-660 ms at the Cz electrode site are shown. SEM bars for each response condition are plotted in black.

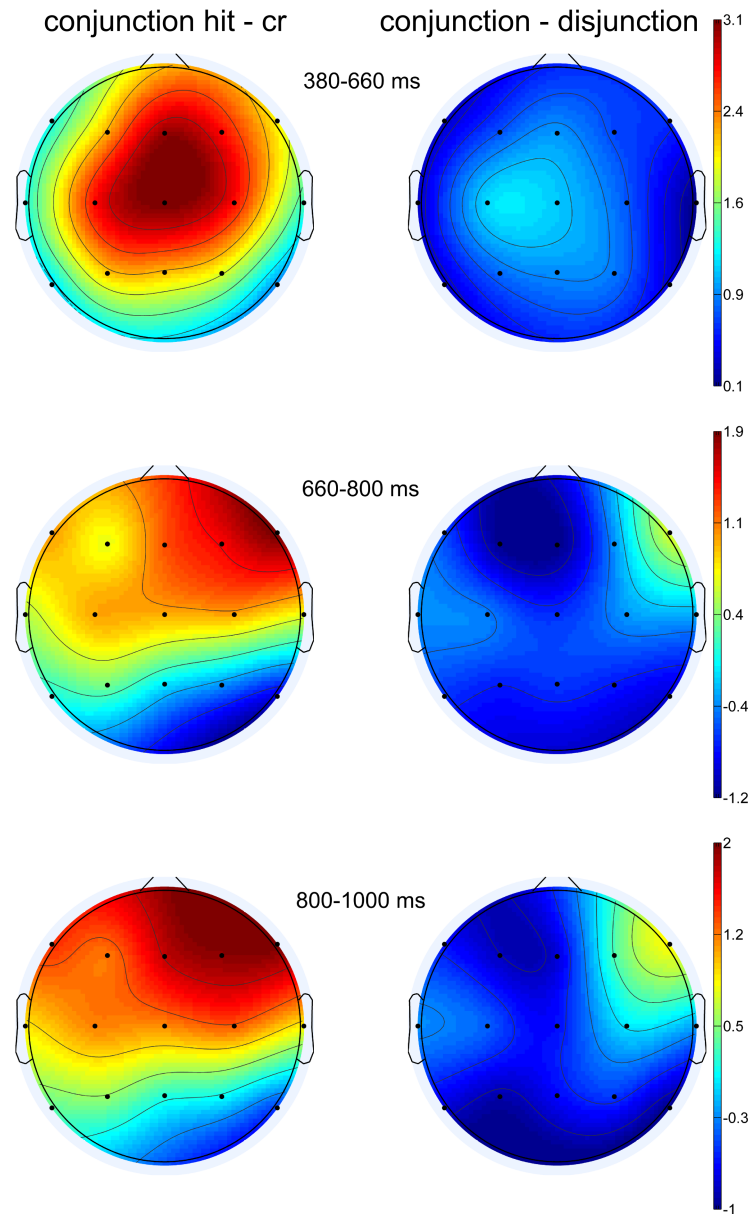
significant difference for the time windows from 240-380 ms to 800-1000 ms. The only significant difference between the conjunction hit and disjunction correct rejection condition was detected in the time window 380-660 ms. The significant interactions between the response conditions and the electrode sites in the time windows from 380-660 ms to 800-1000 ms (380-660 ms:  $F(2,8) = 1.89$ ,  $p = .021$ ; 660-800 ms:  $F(2,8) = 2.77$ ,  $p < .001$ ; 800-1000 ms:  $F(2,8) = 2.23$ ,  $p = .005$ ) indicated that some differences between the response conditions occurred only at specific electrode sites. Figure 3.11 shows the most potential differences between the conjunction hit and disjunction correct rejection condition were located at the C3 and Cz site in the time window 380-660 ms. In the same time window the difference between the conjunction hit and correct rejection condition was increased at the Fz and Cz electrode. Thus, each of these electrodes were analyzed separately by an one-way Anova and revealed significant results for the three comparisons (Fz:  $F(2) = 18.25$ ,  $p < .001$ ; C3:  $F(2) = 20.93$ ,  $p < .001$ ; Cz:  $F(2) = 25.07$ ,  $p < .001$ ). The pairwise comparison between the correct rejection and the other two conditions revealed significant difference at all analyzed

electrode sites ( $p < .004$ ). More interestingly, significant differences between the conjunction hit and disjunction correct rejection condition was revealed at the C3 ( $p = .022$ ) and Cz ( $p = .029$ ) electrode site (figure 3.10).

The further analysis of the Fz electrode site which showed increased differences in the later time windows (figure 3.11), revealed significant difference of the factor response condition (660-800 ms at Fz:  $F(2) = 9.57$ ,  $p < .001$ ; 800-1000 ms at Fz:  $F(2) = 9.57$ ,  $p = .002$ ), but no significant difference between the conjunction hit and disjunction correct rejection conditions was revealed by pairwise comparisons. However, the correct rejection conditions showed again significant differences to the other two conditions ( $p < .009$ ).

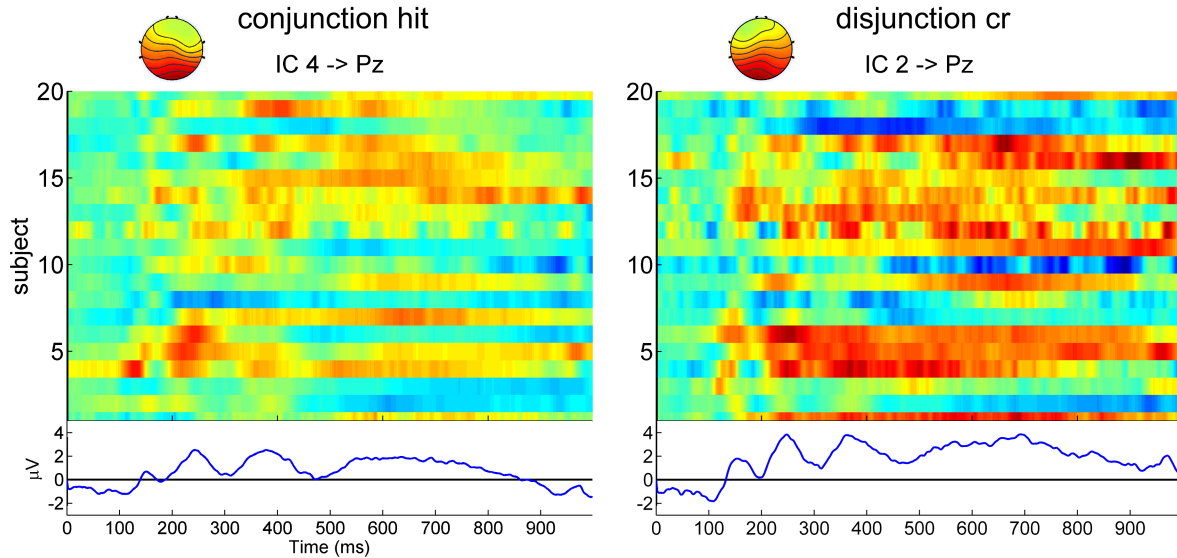
The difference between the conjunction hit and disjunction correct rejection in the time window 380-660 ms resulted in a latency difference of the P600 waves. The latency difference between both response conditions increased from the frontal to parietal electrode sites as seen in figure 3.9. At the Fz and Cz electrode site the P600 wave of the conjunction hit ERP peaked at 636 ms whereas the disjunction correct rejection ERP peaked at 742 ms resulting in a latency difference of 106 ms. The latency difference between the conjunction hit (556 ms) and disjunction correct rejection (682 ms) P600 amplitude at the Pz site amount to 126 ms.

We performed ICAs on the averages of the participants for the conjunction hit and disjunction correct rejection condition to identify the source activity of the P600 wave at the parietal electrode sites. The scalp maps in figure 3.13 show that for both response conditions the potential values increase from the frontal to parietal electrode sites in the time windows 380-660 ms and 660-800 ms, which are relevant for the latency difference of the P600 waves. Thus, we chose for both response conditions an independent component (IC) which adequately characterizes the distributions of the scalp maps and the extending of the P600 wave. The IC 4 described these criteria for the conjunction hit condition and the IC 2 for disjunction correct rejection condition. The scalp maps of the IC 4 and IC 2 in figure 3.12 showed the highest potential values in parietal region comparable to those observed in scalp maps of the grand averages (figure 3.13). The extending of the averaged ICs at the Pz electrode site (blue graph) in figure 3.12 also showed qualitatively similarities to the P600 wave of the grand averages of both response conditions. The participant's power spectra of the ICs at the Pz electrode site in figure 3.12 show strong variation in the P600 wave onset and the latency difference does not clearly emerge from this illustration. However, the mean peak latency of the IC 4 (conjunction hit) was at 570 ms whereas the IC 2 (disjunction cor-



**Figure 3.11:** Scalp maps of the difference ERPs from the time windows 380-660 ms, 660-800 ms and 800-1000 ms are illustrated. The scalp maps of the difference potential between the conjunction hit and disjunction correct rejection are plotted in the left column. The difference between the conjunction hit and correct rejection is shown in the right column. Color coding is scaled in  $\mu\text{V}$  and illustrated by the bars on the right. Electrode sites are marked by black dots.

rect rejection) peaked at 692 ms which resulted in a latency difference of 122 ms comparable to the measured ERP amplitude difference of 126 ms at the Pz electrode site.



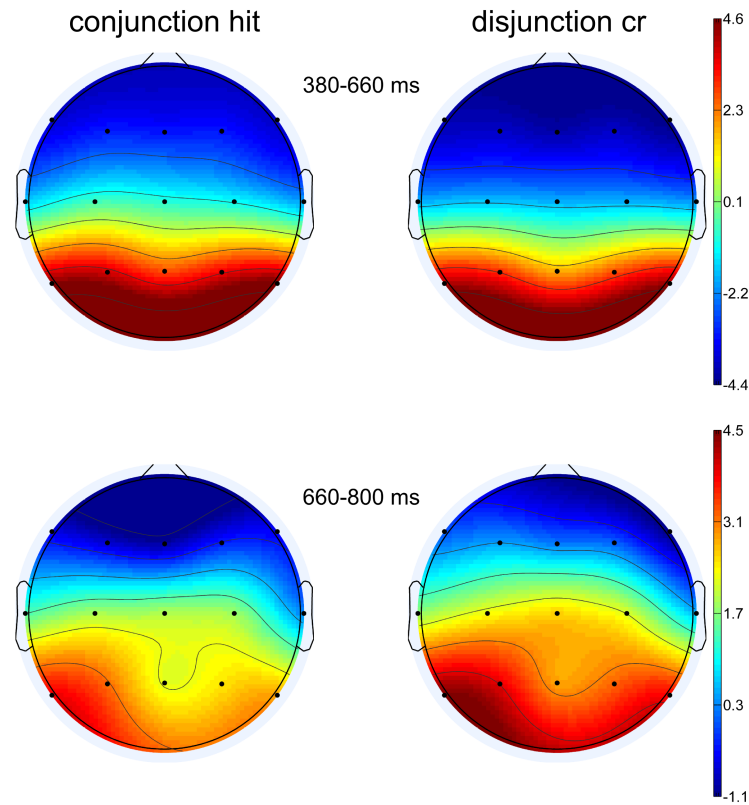
**Figure 3.12:** Overview of the ICA results. Properties of the IC 4 from the conjunction hit averages and IC 2 from the disjunction correct rejection averages at the Pz electrode site. The participant's power spectra of the ICs (same scaling of color coding) are illustrated and below the corresponding IC average referenced to the test scene onset. The corresponding scalp maps of the ICs are plotted left above the power spectra.

### 3.3. Discussion

The results of this study show evidence that the processing of sequence-related temporal context of memorized episodes is located in the mid-parietal region. Since the difference between the conjunction hit and disjunction correct rejection condition is maximal in the central region and the latency delay of the disjunctive P600 wave amplitude is maximal in the parietal region. Thus, indicating that the retrieval of temporal context of memorized episodes is characterized by the extending of the P600 wave in the mid-parietal region.

The correctly rejected non-target scenes showed in comparison to the target scene conditions the hypothesized mid-frontal old/new effect, but not the parietal old/new effect found by classical old/new studies for word and object recognition (Curran, 2000; Duarte et al., 2004; Düzel et al., 1997; Rugg et al., 1998). By contrast, the encountered differences between

the P600 wave of correctly recognized non-target and target scenes are located in the frontal region primarily in the right hemisphere. Nevertheless, correctly recognized non-target and target scenes evoke dissociable P600 waves indicating the involvement of episodic memory content in this study.



**Figure 3.13:** Scalp maps of conjunction hit and disjunction correct rejection ERPs from the time windows 380-660 ms and 660-800 ms are shown. The scalp maps of the conjunction hit ERPs are plotted in the left column and the disjunction correct rejection ERPs in the right column. Color coding is scaled in  $\mu\text{V}$  and illustrated by the bars on the right. Electrode sites are marked by black dots.

The decrease of the conjunction hit rate compared to the disjunction correct rejection rate is a result of the four times more less probability of the appearance of disjunctive target scenes in the recognition test. Therefore, participants had been expecting that a presented target scene in the recognition test was more likely congruent with the learned sequence, which indicates that the disjunctive target scenes required a deeper processing of the temporal context resulting in the P600 wave effect. However, the observed P600 wave effect reflecting the retrieval of episodic temporal context might more clearly emerge if

correctly recognized target scenes which involve retrieval of temporal context are compared with correctly recognized target scenes which do not involve any retrieval of temporal context, but episodic content.

The conjunction and false hit condition did not differ, which supports the assumption that especially detected sequence violation involves deeper processing of temporal context. The analysis of the spacing of the target scene in the recognition test to corresponding target scene in the learning phase shows that the detection of the target scene position in the sequence is identified more gradually as exactly since the detection rate increases with spacing.

The observed differences between the correct rejection and miss condition in the mid-parietal region during P600 wave are not obviously since both responses declared the presented test scene as being a non-target scene. However, in our experimental design the meaning of a miss is ambiguous. A miss could simply imply that a participant did not recognize the target scene, on the other hand it could imply that they recognized the target scene, but falsely estimated the target scene as being at a false position. Thus, the differences between both response conditions may reflect the falsely detected position errors.

The accuracy of correctly rejected distractor scenes was less precisely as the correct rejection conditions of new scenes, however also relatively high. The ERP of the correct rejected distractor scenes only differed from the new conditions at the P4 electrode. This result indicates that the distractor scenes interfere with the retrieval of the target scenes and the temporal context. Additionally we found no difference either in performance or in ERP extending between new scenes from the same category and new scenes from a different category as the target scene. By contrast, Curran (2000) found that new words sharing features with studied words were less likely correctly rejected as new dissimilar words and elicited a dissociable N400 wave. An explanation could be that the chosen categories are derived from the same general category (i.e. rooms), and therefore the global features (e.g. spatial layout, illumination, depth) were too similar in the sense of the proposed coarse-to-fine processing in scene recognition (Schyns and Oliva, 1994).

### 3.4. Conclusion

We conclude that the recognition of temporal context of memorized episodes is characterized by the ‘mid-parietal effect’ found in this study. This effect reflects not only the retrieval

of temporal content from episodic memory, but may be also directly linked to the multiple view integration in spatial working memory. Multiple view integration requires the decoding of temporal coherence of a set of retrieved views since views of places are experienced in the temporal context of their appearance during navigation.



# Bibliography

- Addis, D. R., Wong, A. T., and Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, 45(7):1363–77.
- Aggleton, J. P., Vann, S. D., Denby, C., Dix, S., Mayes, A. R., Roberts, N., and Yonelinas, A. P. (2005). Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia*, 43(12):1810–1823.
- Appleyard, D. (1969). Why buildings are known. *Environment and Behavior*, 1:131–156.
- Avraamides, M. N., Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (2004). Functional equivalence of spatial representations derived from vision and language: evidence from allocentric judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(4):801.
- Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O’Keefe, J., Jeffery, K., and Burgess, N. (2006). The boundary vector cell model of place cell firing and spatial memory. *Reviews in the Neurosciences*, 17(1-2):71.
- Basten, K., Meilinger, T., and Mallot, H. A. (2012). Mental travel primes place orientation in spatial recall. *Spatial Cognition VIII*, pages 378–385.
- Batschelet, E. (1981). *Circular Statistics in Biology*. Mathematics in Biology. Academic Press London.
- Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159.
- Berens, P. (2009). Circstat: a matlab toolbox for circular statistics. *Journal of Statistical Software*, 31(10):1–21.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10:433–436.
- Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, 10(12):551–7.

- Burgess, N., Maguire, E. A., and O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, 35(4):625–641.
- Byrne, P., Becker, S., and Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review*, 114(2):340–375.
- Cartwright, B. and Collett, T. (1982). How honey bees use landmarks to guide their return to a food source. *Nature*, 295:560–564.
- Cartwright, B. and Collett, T. S. (1983). Landmark learning in bees. *Journal of Comparative Physiology*, 151(4):521–543.
- Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, 23(2):149–178.
- Cheung, A., Stürzl, W., Zeil, J., and Cheng, K. (2008). The information content of panoramic images ii: view-based navigation in nonrectangular experimental arenas. *Journal of Experimental Psychology: Animal Behavior Processes*, 34(1):15.
- Cheung, A. and Vickerstaff, R. (2010). Finding the way with a noisy brain. *PLoS Computational Biology*, 6(11):e1000992.
- Chrastil, E. R. and Warren, W. H. (2014). From cognitive maps to cognitive graphs. *PLoS One*, 9(11):e112544.
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory and Cognition*, 28(6):923–938.
- Curran, T. and Hancock, J. (2007). The fn400 indexes familiarity-based recognition of faces. *Neuroimage*, 36(2):464–471.
- Da Silva, J. A. (1985). Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *The American Journal of Psychology*, 98(1):119–144.
- Davachi, L., Mitchell, J. P., and Wagner, A. D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences*, 100(4):2157–2162.
- Delorme, A. and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.
- Duarte, A., Ranganath, C., Winward, L., Hayward, D., and Knight, R. T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cognitive Brain Research*, 18(3):255–272.

- Durier, V., Graham, P., and Collett, T. S. (2003). Snapshot memories and landmark guidance in wood ants. *Current Biology*, 13(18):1614–1618.
- Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H.-J., and Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences*, 94(11):5973–5978.
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., and Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature*, 425(6954):184–188.
- Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., and Engel, S. A. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nature Neuroscience*, 3(11):1149–1152.
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Science*, 12(10):388–96.
- Erkelens, C. J. (2015). The extent of visual space inferred from perspective angles. *i-Perception*, 6(1):5–14.
- Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433.
- Ernst, M. O. and Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4):162–169.
- Etienne, A. S., Boulens, V., Maurer, R., Rowe, T., and Siegrist, C. (2000). A brief view of known landmarks reorientates path integration in hamsters. *Naturwissenschaften*, 87(11):494–498.
- Etienne, A. S. and Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, 14(2):180–92.
- Farrell, M. J. and Robertson, I. H. (1998). Mental rotation and automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1):227.
- Foley, J. M. (1972). The size-distance relation and intrinsic geometry of visual space: Implications for processing. *Vision Research*, 12(2):323–332.
- Foo, P., Warren, W. H., Duchon, A., and Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2):195–215.
- Franz, M. O., Schölkopf, B., Mallot, H. A., and Bühlhoff, H. H. (1998). Where did i take that snapshot? scene-based homing by image matching. *Biological Cybernetics*, 79(3):191–202.

- Fukushima, S. S., Loomis, J. M., and Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):86.
- Gilinsky, A. S. (1951). Perceived size and distance in visual space. *Psychological Review*, 58(6):460–482.
- Gillner, S. and Mallot, H. A. (1998). Navigation and acquisition of spatial knowledge in a virtual maze. *Journal of Cognitive Neuroscience*, 10(4):445–463.
- Gillner, S., Weiss, A. M., and Mallot, H. A. (2008). Visual homing in the absence of feature-based landmark information. *Cognition*, 109(1):105–22.
- Giudice, N. A., Klatzky, R. L., Bennett, C. R., and Loomis, J. M. (2013). Combining locations from working memory and long-term memory into a common spatial image. *Spatial Cognition and Computation*, 13(2):103–128.
- Graham, M., Good, M. A., McGregor, A., and Pearce, J. M. (2006). Spatial learning based on the shape of the environment is influenced by properties of the objects forming the shape. *Journal of Experimental Psychology: Animal Behavior Processes*, 32(1):44–59.
- Hartley, T., Burgess, N., Lever, C., Cacucci, F., and O’keefe, J. (2000). Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus*, 10(4):369–379.
- Hübner, W. and Mallot, H. A. (2007). Metric embedding of view-graphs. *Autonomous Robots*, 23(3):183–196.
- Hermer, L. and Spelke, E. S. (1994). A geometric process for spatial reorientation in young-children. *Nature*, 370(6484):57–59.
- Hochstetter, K., Bornfleth, H., Weckesser, D., Ille, N., Berg, P., and Scherg, M. (2004). Besa source coherence: a new method to study cortical oscillatory coupling. *Brain Topography*, 16(4):233–238.
- Holdstock, J., Mayes, A., Roberts, N., Cezayirli, E., Isaac, C., O’reilly, R., and Norman, K. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, 12(3):341–351.
- Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., and Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. spatial and temporal analyses using fmri and erp. *Cerebral Cortex*, 16(9):1349–1360.
- Jacobs, W. J., Laurance, H. E., and Thomas, K. G. (1997). Place learning in virtual space i: Acquisition, overshadowing, and transfer. *Learning and Motivation*, 28(4):521–541.
- Jacobs, W. J., Thomas, K. G., Laurance, H. E., and Nadel, L. (1998). Place learning in virtual space: ii. topographical relations as one dimension of stimulus control. *Learning and Motivation*, 29(3):288–308.

- Janzen, G. and Van Turenout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, 7(6):673–677.
- Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Doherty, S., and Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior*, 22(1):19–43.
- Kuipers, B. (1978). Modeling spatial knowledge. *Cognitive Science*, 2(2):129–153.
- Learmonth, A. E., Newcombe, N. S., and Huttenlocher, J. (2001). Toddlers’ use of metric information and landmarks to reorient. *Journal of Experimental Child Psychology*, 80(3):225–244.
- Lever, C., Burton, S., Jeewajee, A., O’Keefe, J., and Burgess, N. (2009). Boundary vector cells in the subiculum of the hippocampal formation. *The Journal of Neuroscience*, 29(31):9771–7.
- Loomis, J. M., Klatzky, R. L., and Giudice, N. A. (2013). Representing 3d space in working memory: Spatial images from vision, hearing, touch, and language. *Multisensory Imagery*, pages 131–155.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., and Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *Journal of Experimental Psychology: General*, 122(1):73.
- Madl, T., Chen, K., Montaldi, D., and Trapp, R. (2015). Computational cognitive models of spatial memory in navigation space: A review. *Neural Networks*, 65:18–43.
- Margules, J. and Gallistel, C. (1988). Heading in the rat: Determination by environmental shape. *Animal Learning and Behavior*, 16(4):404–410.
- McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M. B. (2006). Path integration and the neural basis of the cognitive map. *Nature Reviews Neuroscience*, 7(8):663–78.
- Meilinger, T., Frankenstein, J., Simon, N., Bühlhoff, H. H., and Bresciani, J.-P. (2015). Not all memories are the same: Situational context influences spatial recall within one’s city of residency. *Psychonomic Bulletin & Review*, pages 1–7.
- Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., and Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. *Cerebral Cortex*, 12(12):1322–1330.
- Mittelstaedt, M.-L. and Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften*, 67(11):566–567.

- Möller, R. and Vardy, A. (2006). Local visual homing by matched-filter descent in image distances. *Biological Cybernetics*, 95(5):413–430.
- Montello, D. R. (1993). Scale and multiple psychologies of space. *Spatial information theory a theoretical basis for GIS*, pages 312–321.
- Morris, R. G. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, 12(2):239–260.
- Mou, W. and McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(1):162.
- Mou, W., McNamara, T. P., Valiquette, C. M., and Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1):142–57.
- O’Keefe, J. (1991). An allocentric spatial model for the hippocampal cognitive map. *Hippocampus*, 1(3):230–235.
- O’Keefe, J. and Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, 381(6581):425–428.
- O’Keefe, J. and Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 34(1):171–175.
- O’Keefe, J. and Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press Oxford.
- Ooi, T. L. and He, Z. J. (2007). A distance judgment function based on space perception mechanisms: revisiting gilinsky’s (1951) equation. *Psychological Review*, 114(2):441–54.
- Petzschner, F. H. and Glasauer, S. (2011). Iterative bayesian estimation as an explanation for range and regression effects: a study on human path integration. *The Journal of Neuroscience*, 31(47):17220–17229.
- Philbeck, J. W. and Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):72.
- Pickup, L. C., Fitzgibbon, A. W., and Glennerster, A. (2013). Modelling human visual navigation using multi-view scene reconstruction. *Biological Cybernetics*, 107(4):449–64.
- Röhrich, W. G., Hardiess, G., and Mallot, H. A. (2014). View-based organization and interplay of spatial working and long-term memories. *PLoS One*, 9(11):e112793.
- Riecke, B. E., van Veen, H. A., and Bühlhoff, H. H. (2002). Visual homing is possible without landmarks: A path integration study in virtual reality. *Presence: Teleoperators and Virtual Environments*, 11(5):443–473.

- Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(6):1157.
- Rugg, M. D. and Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6):251–257.
- Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., and Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392(6676):595–598.
- Rugg, M. D. and Nagy, M. E. (1989). Event-related potentials and recognition memory for words. *Electroencephalography and Clinical Neurophysiology*, 72(5):395–406.
- Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., and Lindsley, D. B. (1980). Electroconvulsive signs of levels of processing: perceptual analysis and recognition memory. *Psychophysiology*, 17(6):568–576.
- Schacter, D. L. and Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions: Biological Sciences*, 362(1481):773–86.
- Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, 8(9):657–661.
- Schölkopf, B. and Mallot, H. A. (1995). View-based cognitive mapping and path planning. *Adaptive Behavior*, 3(3):311–348.
- Schwartz, M. (1999). Haptic perception of the distance walked when blindfolded. *Journal of Experimental Psychology: Human Perception and Performance*, 25(3):852.
- Schyns, P. G. and Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5(4):195–200.
- Siegel, A. W. and White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, 10:9–55.
- Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., and Moser, E. I. (2008). Representation of geometric borders in the entorhinal cortex. *Science*, 322(5909):1865–1868.
- Steck, S. D. and Mallot, H. A. (2000). The role of global and local landmarks in virtual environment navigation. *Presence: Teleoperators and Virtual Environments*, 9(1):69–83.
- Stevens, S. S. (1961). To honor fechner and repeal his law. *Science*, 133(3446):80–86.
- Stürzl, W., Cheung, A., Cheng, K., and Zeil, J. (2008). The information content of panoramic images i: The rotational errors and the similarity of views in rectangular experimental arenas. *Journal of Experimental Psychology: Animal Behavior Processes*, 34(1):1–14.

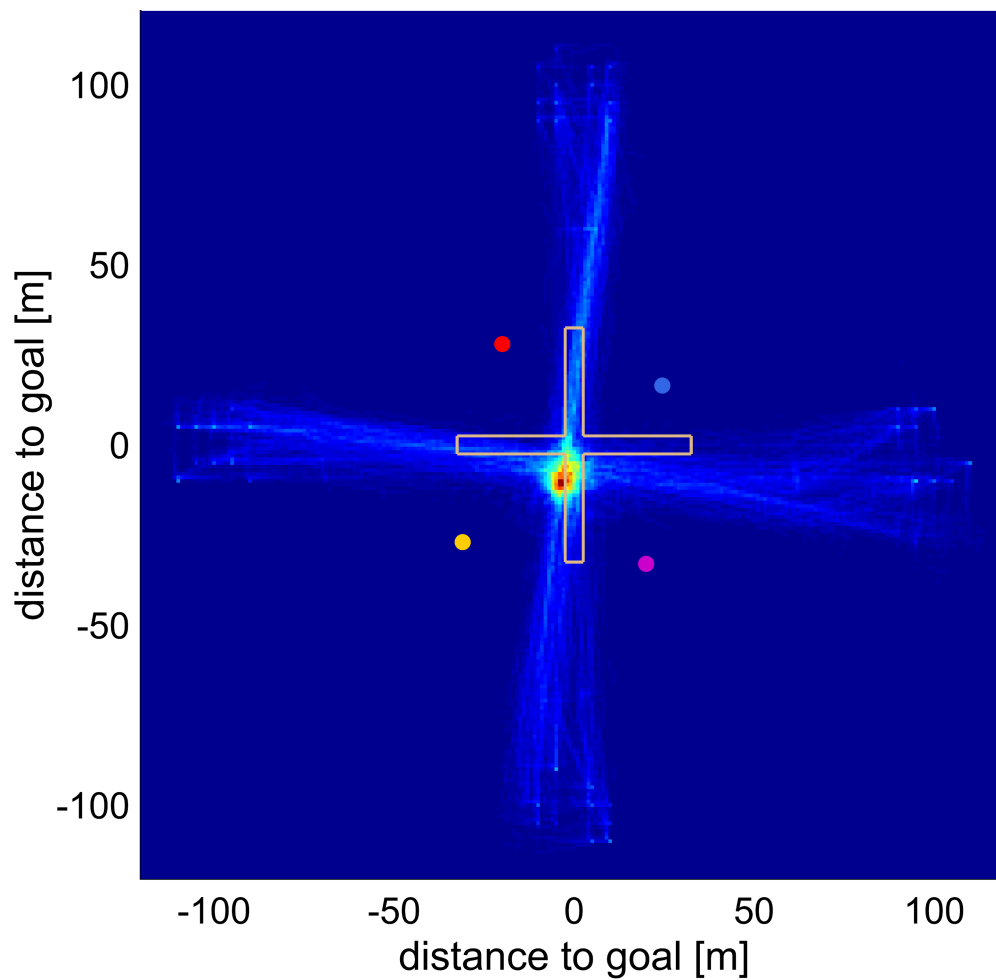
- Stürzl, W. and Mallot, H. A. (2006). Efficient visual homing based on fourier transformed panoramic images. *Robotics and Autonomous Systems*, 54(4):300–313.
- Suddendorf, T. and Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, 30(03):299–313.
- Tatler, B. W. and Land, M. F. (2011). Vision and the representation of the surroundings in spatial memory. *Philosophical Transactions: Biological Sciences*, 366(1564):596–610.
- Teghtsoonian, R. and Teghtsoonian, M. (1978). Range and regression effects in magnitude scaling. *Perception and Psychophysics*, 24(4):305–314.
- Thorndyke, P. W. and Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14(4):560–589.
- Thrun, S., Liu, Y., Koller, D., Ng, A. Y., Ghahramani, Z., and Durrant-Whyte, H. (2004). Simultaneous localization and mapping with sparse extended information filters. *The International Journal of Robotics Research*, 23(7-8):693–716.
- Tinbergen, N. and Kruyt, W. (1938). Über die orientierung des bienenwolfes (*philanthus triangulum* fabr.). *Zeitschrift für vergleichende Physiologie*, 25(3):292–334.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4):189.
- Trullier, O., Wiener, S. I., Berthoz, A., and Meyer, J.-A. (1997). Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology*, 51(5):483–544.
- Waller, D., Loomis, J. M., Golledge, R. G., and Beall, A. C. (2000). Place learning in humans: The role of distance and direction information. *Spatial Cognition and Computation*, 2(4):333–354.
- Warren, L. R. (1980). Evoked potential correlates of recognition memory. *Biological Psychology*, 11(1):21–35.
- Wolbers, T., Hegarty, M., Büchel, C., and Loomis, J. M. (2008). Spatial updating: how the brain keeps track of changing object locations during observer motion. *Nature Neuroscience*, 11(10):1223–30.
- Wolbers, T., Wiener, J. M., Mallot, H. A., and Büchel, C. (2007). Differential recruitment of the hippocampus, medial prefrontal cortex, and the human motion complex during path integration in humans. *The Journal of Neuroscience*, 27(35):9408–9416.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3):441–517.



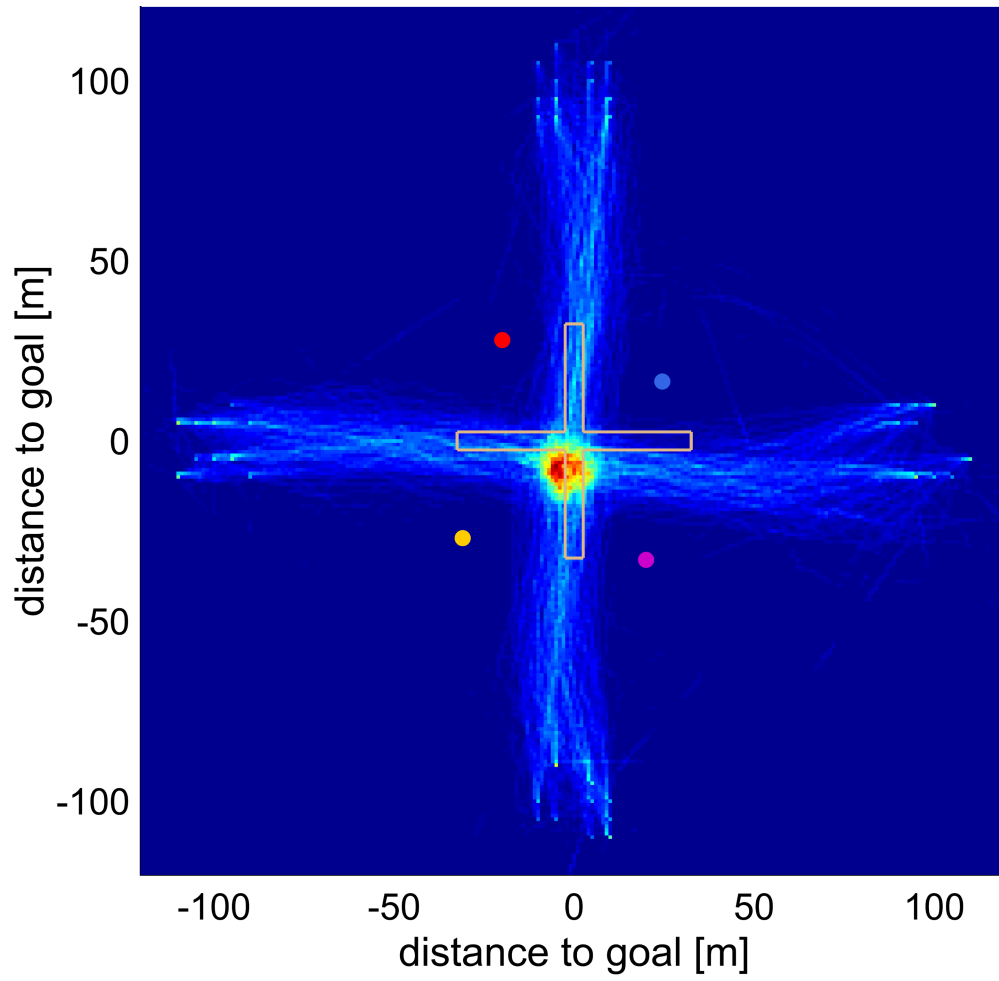


## A. Appendix

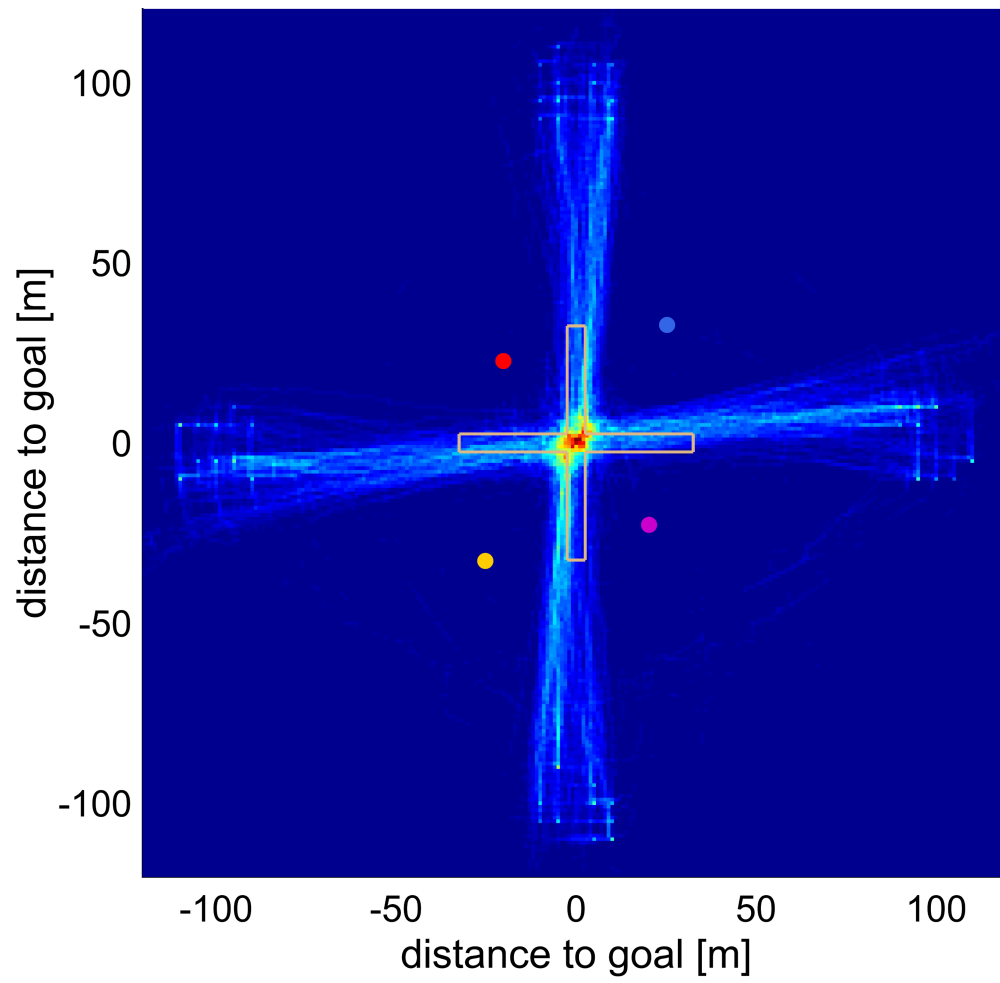
### A.1. Heat maps of trajectories



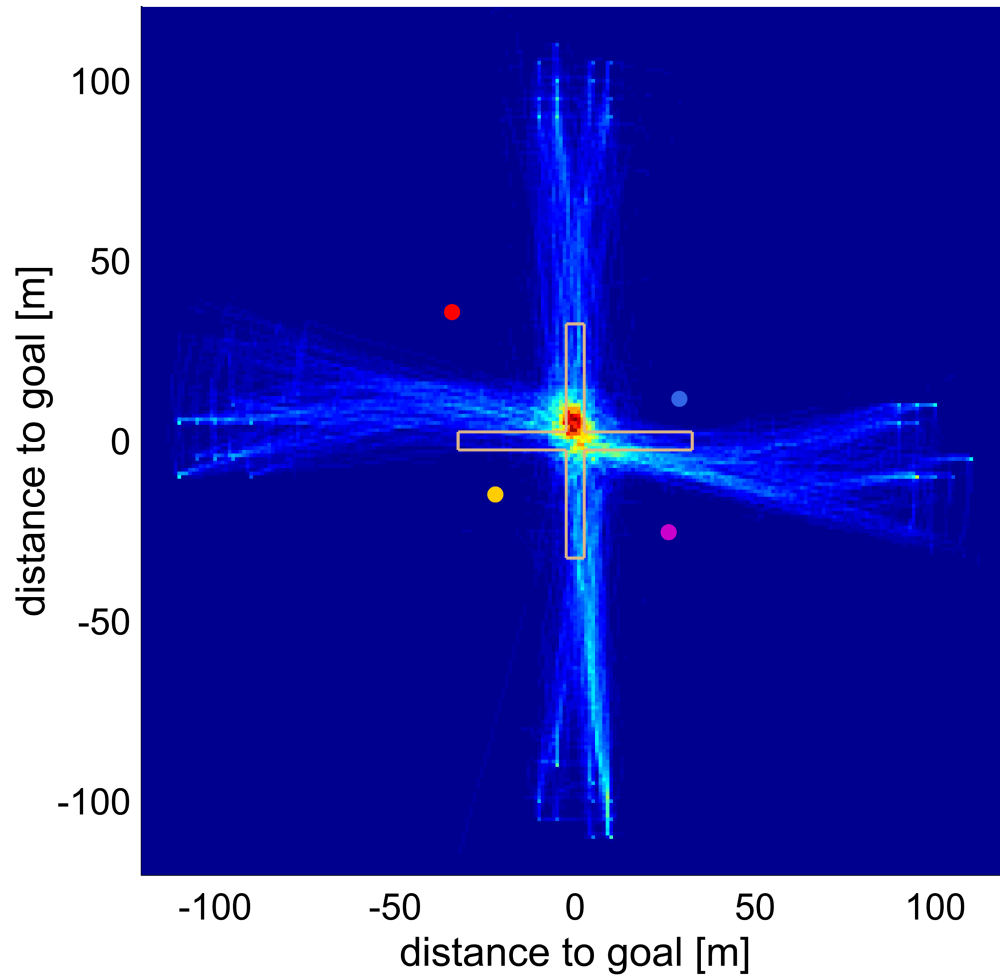
**Figure A.1:** Heat map of trajectories of individual trials. Data from the desktop condition in Experiment 3 is shown.



**Figure A.2:** Heat map of trajectories of individual trials. Data from the HMD condition in Experiment 3 is shown.

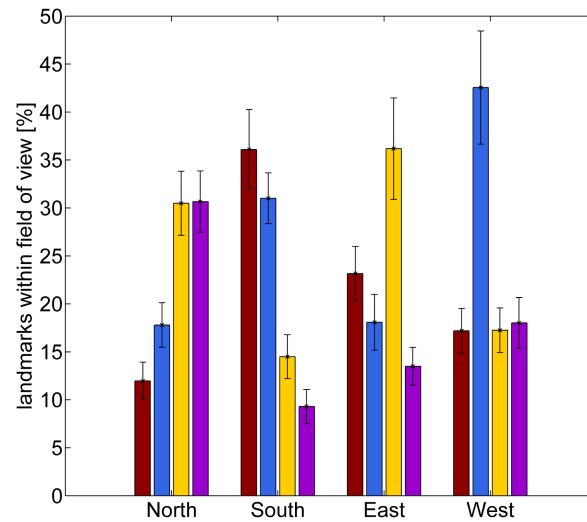


**Figure A.3:** Heat map of trajectories of individual trials. Data from the parallelogram condition in Experiment 4 is shown.

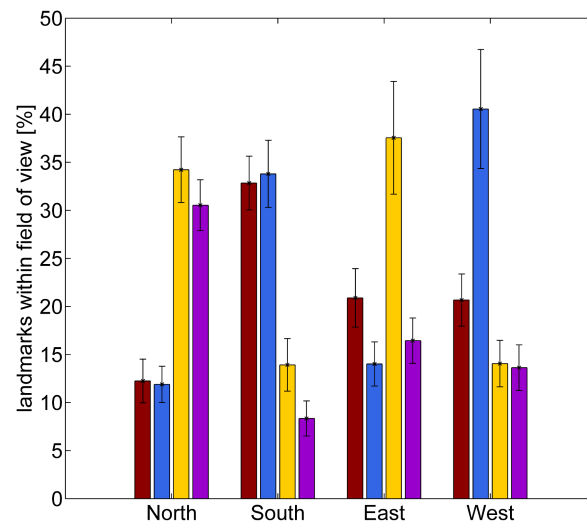


**Figure A.4:** Heat map of trajectories of individual trials. Data from the peaked condition in Experiment 4 is shown.

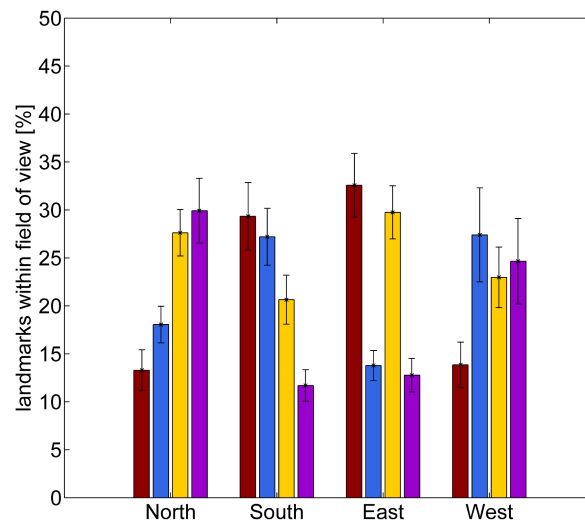
## A.2. Bar graphs of landmarks within the field of view



**Figure A.5:** Bar graph of landmarks within the field view during participants approaches in the control condition.



**Figure A.6:** Bar graph of landmarks within the field view during participants approaches in the parallelogram condition.



**Figure A.7:** Bar graph of landmarks within the field view during participants approaches in the peaked condition.

### A.3. Procedure description

#### Versuchsbeschreibung

##### **Virtuelle Umgebung**

Die Umgebung besteht aus einem kleinen See mit einer plusförmigen Brücke. Innerhalb des Sees befinden sich vier unterschiedlich gefärbte Ballons. Der Ort ist dafür bekannt, dass aus heiterem Himmel Bodennebel auftauchen kann und die gesamte Landschaft im Nebelschleier versinkt.

##### **Steuerung**

Die Bewegungsrichtung in der virtuellen Umgebung wird mittels der *wasd*-Keyboardtasten gesteuert. Mit *w* bewegt man sich vorwärts, mit *s* rückwärts, mit *a* bewegt man sich seitwärts nach links und mit *d* seitwärts nach rechts. Die Blickrichtung wird mit einer entsprechenden Bewegung der Maus verändert.

##### **Lernphase**

Sie haben die Aufgabe von unterschiedlichen Startpunkten aus, unterschiedliche Ziele auf der linken oder rechten Seite vom See zu erreichen. Sie müssen stets die Brücke überqueren und versuchen, dabei nicht ins Wasser zu fallen. Das gilt vor allem bei plötzlich auftretendem Bodennebel. Orientieren Sie sich an den Ballons. Diese sind Ihre einzige Orientierungshilfe bei auftauchendem Bodennebel.

##### **Versuchsphase**

Während dieser Phase bedeckt der Bodennebel ständig die Landschaft. Vor jeder Phase wird Ihnen mitgeteilt, welches Ziel Sie besuchen sollen. Versuchen Sie dieses Ziel so genau wie möglich zu lokalisieren.

##### **Allgemein**

Ihre Versuchsdaten werden für wissenschaftliche Veröffentlichungen verwendet. Deshalb führen Sie das Experiment bitte gewissenhaft und konzentriert durch.

**Figure A.8:** General procedure description of the behavioral place recognition study.



## Versuchsbeschreibung

### **Virtuelle Umgebung**

Die Umgebung besteht aus einem kleinen See mit einer plusförmigen Brücke. Innerhalb des Sees befinden sich vier unterschiedlich gefärbte Ballons. Der Ort ist dafür bekannt, dass aus heiterem Himmel Bodennebel auftauchen kann und die gesamte Landschaft im Nebelschleier versinkt.

### **Steuerung**

Die Bewegungsrichtung in der virtuellen Umgebung wird mittels Drehen ihres Körpers gesteuert. Mit der linken Maustaste bewegt man sich vorwärts, mit der rechten rückwärts. Die Blickrichtung wird mit dem Kopf gesteuert.

### **Lernphase**

Sie haben die Aufgabe von unterschiedlichen Startpunkten aus unterschiedliche Ziele auf der linken oder rechten Seite vom See zu erreichen. Sie müssen stets die Brücke überqueren und versuchen dabei nicht ins Wasser zu fallen, das gilt vor allem bei plötzlich auftretenden Bodennebel. Orientieren Sie sich an den Ballons, diese sind ihre einzige Orientierungshilfe bei auftauchendem Bodennebel.

### **Versuchsphase**

Während dieser Phase bedeckt der Bodennebel ständig die Landschaft. Vor jeder Phase wird ihnen mitgeteilt, welches Ziel Sie besuchen sollen. Sie bestätigen das Ziel mit Klicken auf das Mauselement. Versuchen Sie dieses Ziel so genau wie möglich zu lokalisieren.

### **Allgemein**

Ihre Versuchsdaten werden für wissenschaftliche Veröffentlichungen verwendet, deshalb führen Sie das Experiment bitte gewissenhaft und konzentriert durch.

**Figure A.9:** Procedure description of the HMD condition in Experiment 3 of the behavioral place recognition study.

## A.4. Content form

### Probandeninformation zum Experiment

Name des Versuchsleiters: Stephan Lancier

Sie werden an einem psychophysikalischen Experiment teilnehmen. Während des Experimentes werden Sie in einer virtuellen Umgebung navigieren und Ihre Navigationsleistung wird dabei gemessen.

Ihre persönlichen, im Rahmen dieser Studie erhobenen Daten werden für wissenschaftliche Publikationen - in anonymisierter Form - verwendet und am Lehrstuhl für Kognitive Neurowissenschaft gespeichert.

Die Teilnahme erfolgt freiwillig und kann zu jedem Zeitpunkt ohne Angabe von Gründen abgebrochen werden.

Tübingen, den \_\_\_\_\_  
Unterschrift

### Einverständniserklärung zur Teilnahme

Name der Versuchsperson (in Druckbuchstaben): \_\_\_\_\_

Ich erkläre mich bereit, an dem Experiment teilzunehmen.

Ich habe den Text der Probandeninformation und dieser Einverständniserklärung gelesen und verstanden. Aufgetretene Fragen wurden mir verständlich und vollständig beantwortet. Ich hatte ausreichend Zeit, Fragen zu stellen und mich für oder gegen eine Teilnahme zu entscheiden.

Ich erkläre mich damit einverstanden, dass meine persönlichen, während dieser Studie erhobenen, Daten im Rahmen von wissenschaftlichen Publikationen - in anonymisierter Form - veröffentlicht und am Lehrstuhl für Kognitive Neurowissenschaft gespeichert werden.

Tübingen, den \_\_\_\_\_  
Unterschrift

**Figure A.10:** General content form of the place recognition study.

### Probandeninformation zum Experiment

Name des Versuchsleiters: Marie Admard

Sie werden an einem psychophysikalischen Experiment teilnehmen. Sie werden während des Experimentes mittels eines HMD (head-mounted display) und eines Targets um den Bauch in einer virtuellen Umgebung navigieren und Ihre Navigationsleistung wird dabei gemessen.

Ihre persönlichen, im Rahmen dieser Studie erhobenen Daten werden für wissenschaftliche Publikationen - in anonymisierter Form - verwendet und am Lehrstuhl für Kognitive Neurowissenschaft gespeichert.

Die Teilnahme erfolgt freiwillig und kann zu jedem Zeitpunkt ohne Angabe von Gründen abgebrochen werden.

Tübingen, den \_\_\_\_\_  
Unterschrift

### Einverständniserklärung zur Teilnahme

Name der Versuchsperson (in Druckbuchstaben): \_\_\_\_\_

Ich erkläre mich bereit, an dem Experiment teilzunehmen.

Ich habe den Text der Probandeninformation und dieser Einverständniserklärung gelesen und verstanden. Aufgetretene Fragen wurden mir verständlich und vollständig beantwortet. Ich hatte ausreichend Zeit, Fragen zu stellen und mich für oder gegen eine Teilnahme zu entscheiden.

Ich erkläre mich damit einverstanden, dass meine persönlichen, während dieser Studie erhobenen, Daten im Rahmen von wissenschaftlichen Publikationen - in anonymisierter Form - veröffentlicht und am Lehrstuhl für Kognitive Neurowissenschaft gespeichert werden.

Tübingen, den \_\_\_\_\_  
Unterschrift

**Figure A.11:** Content form of the HMD condition in Experiment 4 of the place recognition study.

## Probandeninformation zum Experiment

Name des Versuchsleiters: Stephan Lancier

Sie werden an einem psychophysikalischen Experiment teilnehmen. Während des Experimentes werden Ihnen Bilder auf einem Bildschirm präsentiert, die Sie sich merken sollen, dabei wird Ihre Erinnerungsleistung gemessen. Für die Dauer des Experiments werden Sie an ein EEG angeschlossen, d.h. Sie werden eine Kappe mit Elektroden tragen und um den Widerstand zwischen Kopfhaut und Elektrode zu minimieren werden wir Elektrodengel auf Ihre Kopfhaut auftragen. Das EEG zeichnet während des Experimentes die elektrischen Potentiale an Ihrer Kopfhaut auf.

Ihre persönlichen, im Rahmen dieser Studie erhobenen Daten werden für wissenschaftliche Publikationen - in anonymisierter Form - verwendet und am Lehrstuhl für Kognitive Neurowissenschaft gespeichert.

Die Teilnahme erfolgt freiwillig und kann zu jedem Zeitpunkt ohne Angabe von Gründen abgebrochen werden.

Tübingen, den \_\_\_\_\_  
Unterschrift

## Einverständniserklärung zur Teilnahme

Name der Versuchsperson (in Druckbuchstaben): \_\_\_\_\_

Ich erkläre mich bereit, an dem Experiment teilzunehmen.

Ich habe den Text der Probandeninformation und dieser Einverständniserklärung gelesen und verstanden. Aufgetretene Fragen wurden mir verständlich und vollständig beantwortet. Ich hatte ausreichend Zeit, Fragen zu stellen und mich für oder gegen eine Teilnahme zu entscheiden.

Ich erkläre mich damit einverstanden, dass meine persönlichen, während dieser Studie erhobenen, Daten im Rahmen von wissenschaftlichen Publikationen - in anonymisierter Form - veröffentlicht und am Lehrstuhl für Kognitive Neurowissenschaft gespeichert werden.

Tübingen, den \_\_\_\_\_  
Unterschrift

Figure A.12: Content form of the ERP study.

# List of Figures

1.1	Schematic Illustration of functional interaction between perception, spatial image and place code . . . . .	6
2.1	HMD condition setup . . . . .	12
2.2	Bird's eye view of the virtual environment . . . . .	13
2.3	Heat map of the trajectories of the control condition . . . . .	16
2.4	Mean decision points of the control and explicit condition . . . . .	17
2.5	3D-histograms of the control and explicit condition . . . . .	18
2.6	Mean decision points of the control and salience condition . . . . .	19
2.7	Mean decision points of the desktop and HMD condition . . . . .	21
2.8	Heat maps of trajectories related to the manipulation of the landmark sizes .	23
2.9	Heat maps of trajectories of the parallelogram and peaked group . . . . .	25
2.10	The circular relation between the landmarks and the main axes . . . . .	27
2.11	Decision points and error ellipses depending on the approach direction . . . .	28
2.12	Bar graph of landmarks within the field view during participants approaches	29
2.13	Scheme of the model . . . . .	34
2.14	Combined likelihoods of the control, parallelogram and peaked condition . .	36
2.15	Combined likelihoods depending on the approach direction . . . . .	37
3.1	Experimental setup after EEG preparation . . . . .	40

3.2	Schematic overview of the EEG procedure . . . . .	42
3.3	Example of possible behavioral responses during one trial in the ERP experiment	43
3.4	Bar graph of the mean behavioral responses . . . . .	46
3.5	Mean correct rejections and means of spacing to correct position . . . . .	47
3.6	Grand average ERPs of the different correct rejection conditions . . . . .	48
3.7	Mean amplitudes of the correct rejection conditions in time window 660-800 ms at P4 . . . . .	49
3.8	Grand average ERPs of miss and correct rejection . . . . .	50
3.9	Grand average ERPs of conjunction hit, disjunction correct rejection, correct rejection and false hit . . . . .	51
3.10	Mean amplitudes of the conjunction hit, disjunction correct rejection and correct rejection ERP in the time window 380-660 ms at Cz . . . . .	52
3.11	Scalp maps of difference ERPs . . . . .	54
3.12	Overview of the ICA results . . . . .	55
3.13	Scalp maps of conjunction hit and disjunction correct rejection ERPs . . . . .	56
A.1	Heat map of the trajectories of the desktop condition . . . . .	68
A.2	Heat map of the trajectories of the HMD condition . . . . .	69
A.3	Heat map of the trajectories of the parallelogram condition . . . . .	70
A.4	Heat map of the trajectories of the peaked condition . . . . .	71
A.5	Bar graph of landmarks within the field view for the control condition . . . . .	72
A.6	Bar graph of landmarks within the field view for the parallelogram condition	72
A.7	Bar graph of landmarks within the field view for the peaked condition . . . . .	73
A.8	General procedure description of the place recognition study . . . . .	74
A.9	Procedure description of the HMD condition . . . . .	75
A.10	General content form of the place recognition study . . . . .	76
A.11	Content form of the HMD condition . . . . .	77



A.12 Content form of the ERP study . . . . . 78

# List of Tables

3.1 Notation overview of the possible response conditions . . . . . 44