Max–Planck–Institut
für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 64

View–based vs. place–based navigation: What is recognized in recognition–triggered responses?

Hanspeter A. Mallot[1] & Sabine Gillner[2]

October 1998

The work described in this paper was done at the Max-Planck-Institut für biologische Kybernetik. Additional support was obtained Deutsche Forschungsgemeinschaft (DFG, grant number MA 1038/6-1). We are grateful to Heinrich H. Bülthoff and Sibylle D. Steck for intellectual and practical support. Preliminary versions of this paper have been presented at ECVP (Gillner and Mallot 1996) and ARVO (Mallot and Gillner 1997).

[1] AG Bülthoff, E–mail: hanspeter.mallot@tuebingen.mpg.de
[2] Abteilung Unfallchirurgische Forschung und Biomechanik, Universität Ulm, Helmholtzstr. 14, D-89 081 Ulm, E–mail: sabine.gillner@medizin.uni-ulm.de

This document is available as /pub/mpi-memos/TR-064.ps via anonymous ftp from ftp.kyb.tuebingen.mpg.de or from the World Wide Web, http://www.kyb.tuebingen.mpg.de/bu.html.

View–based vs. place–based navigation: What is recognized in recognition–triggered responses?

Hanspeter A. Mallot & Sabine Gillner

Abstract. The usage of landmark information in a route navigation task is investigated in a virtual environment. After learning a route, subjects were released at intermediate points along the route and asked to indicate the next movement direction required to continue the route. At each decision point, three landmarks were present, one of which was viewed centrally and two which appeared in the periphery of the visual field when approaching the decision point. Replacement of the landmarks in the test phase did not affect subjects' performance as long as the direction informations associated with each landmark and landmark position (left, central, right) during the learning phase were consistent. Only if landmarks were combined that carried conflicting movement informations, a reduced performance is observed. We conclude that local views and objects are recognized individually and that the associated directions are combined in a voting scheme. No evidence was found for a recognition of places as panoramic views or configurations of objects.

1 Introduction

One important source of information for navigation and spatial memory is provided by the external sensory signals obtained instantaneously at each position in space. This "local position information", i.e. the manifold of all sensor readings as a function of observer position and orientation, is the most general concept of allocentric, or landmark information. In vision, the local position information at one particular point is a view or "snapshot", i.e. a raw image.

Landmark information can be used in a number of different ways. We give a brief overview in terms of two largely independent dimensions: (i) the amount of image processing needed to extract the landmark from the sensory input and (ii) the function of a landmark in spatial behavior.

(i) Virtually no image processing (except, maybe, for normalization or bandpass filtering) is required in snapshot–based schemes (eg Cartwright and Collett 1982). Remembering only the pattern of black and white spots in an image without any higher level processing such as object recognition is already sufficient for a large number of navigation tasks; see Schölkopf and Mallot (1995) and Franz et al (1998a) for a view–based approach to cognitive maps. However, there is evidence for more sophisticated image processing being involved in mammalian navigation behavior. Cheng (1986), in rodents, and Hermer and Spelke (1994), in young children, have found that geometric information in images is a stronger cue than pure texture or contrast information. This indicates that some image processing has taken place to recover geometrical, i.e. depth cues from the images. Another image processing operation, the segmentation of the image into objects and the assignment of depth values to these objects is assumed in the theoretical approaches e.g., by Zipser (1985), O'Keefe (1991), Penna and Wu (1993), or Prescott (1995). Strategical selection of landmarks with respect to critical sections of a route has been demonstrated by Cohen and Schuepfer (1980) and by Aginski et al (1997). In summary, various types of landmark information ranging from snapshots to identified objects may co–exist in biological navigation systems.

(ii) The second dimension along which types of landmarks can be distinguished is landmark function. O'Keefe and Nadel (1978) distinguish guidance and direction, the latter of which is now usually referred to as "recognition–triggered response" (Trullier et al 1997), see figure 1. In guidance, movement is such that a certain configuration of landmarks is obtained. In the simplest case, this is just the central approach towards a landmark which is then often called a beacon. By keeping the image of a distant landmark at a fixed

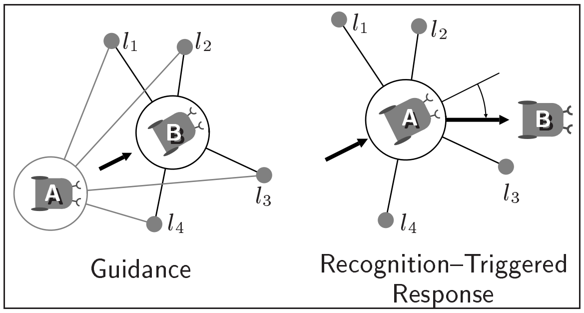


Figure 1: Two types of landmark function. The circles surrounding the vehicles symbolize the visual array of the respective position; $l_1, ..., l_4$ are landmarks. In *guidance* (left), the "snapshot" visible at position **B** has been stored. At a location **A**, movement is such that the currently visible snapshot will become more similar to the stored one. In *recognition–triggered response* (right) memory contains both a snapshot and an action associated with it. When the snapshot is recognized in **A**, an action such as a turn by some remembered angle is executed.

retinal position, straight walks with arbitrary direction can be produced; here, the global landmark provides some sort of compass information. A more general example of a guidance would be to move to a place where one landmark is straight ahead of the observer, a second is 90° to the left and a third landmark is at 90° to the right. By this token, guidances can be used to reach arbitrary places in open space. Examples include the Morris water maze task in rodents (Morris 1981), scene–based homing in insects (Cartwright and Collett 1982) and human place learning in virtual space (Jacobs et al 1998). In terms of the image processing classification, Cartwright and Collett (1982) suggest a snapshot scheme (see also Franz et al 1998b for a survey of scene–based homing schemes).

In guidance, spatial memory contains a desired snapshot or landmark configuration. The movement required to reach the place corresponding to this configuration is computed by comparing current and stored landmark positions. In recognition triggered response, memory contains also a second bit of information, namely an action to be performed when a place is reached, i.e. when a landmark configuration is recognized.

In the definition given by Trullier et al (1997), the term "place–recognition triggered response" implies that place recognition is independent of the oberserver's orientation or viewing direction and that prior to actually taking the local action, a standard orientation with respect to the place has to be obtained. Alternatively, one could assume a view–recognition triggered response, in which views, rather than places, are recognized. In honey–bees, Collett and Baron (1995) have shown that movement decisions can in fact be triggered by recognition of views.

In a previous paper (Gillner and Mallot 1998) we have presented evidence for recognition triggered responses in human subjects navigating through a virtual environment. It was shown that subjects returning to a given landmark are biased towards repeating the movement performed when last passing along that same landmark. This persistence seems to be independent of the currently pursued goal. While this behavior is rather stereotyped and may be classified as route knowledge, evidence of configuration knowledge and cognitive maps is simultaneously present in the same subjects. For a detailed discussion of the relation of route– and configuration knowledge in a unified framework (the view–graph approach) see Gillner and Mallot (1998) and Schölkopf and Mallot (1995).

What exactly is recognized in recognition–triggered responses: views or places? For the case of guidance, Poucet (1993) has argued that local views are mentally integrated into panoramic views which serve as a representation of the respective place. This representation will be independent of each local view and the observer's viewing direction. A similar conclusion has been drawn by Jacobs et al (1998) who had subjects find a place in a simulated arena surrounded by structured walls. In the recognition part of a recognition–triggered response, independence of observer orientation is not desirable, at least if the action triggered by recognition is a turning movement. If recognition were in fact independent of orientation, additional compass information would be required as a reference for such directional movements. In this paper, we will ask whether the recognition part of a recognition–triggered response concerns a plain view or snapshot of a scene, a panoramic view of a place, or a the landmark configuration of a place. The role of compass information, which would be required if actions were triggered by recognized places, but not if they were triggered by recognized views, has been addressed elsewhere (Steck and Mallot 1998).

We investigate the question of view–based vs. place–based direction memory by means of landmark transposition experiments in the "Hexatown" virtual environment (see Gillner and Mallot 1998 and section 2). The possibility of manipulating the environments by exchanging landmarks,

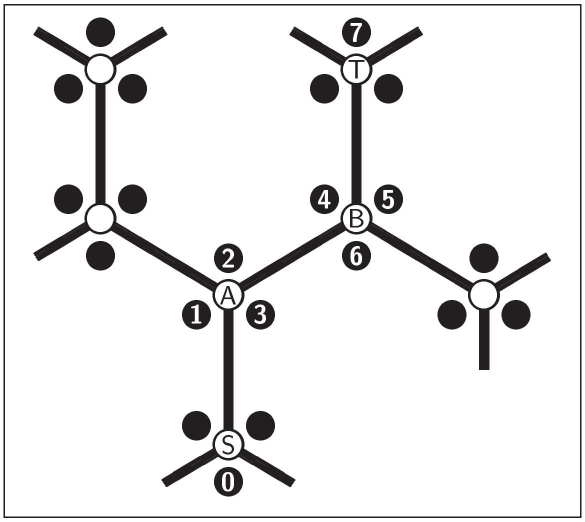


Figure 2: Street map of the virtual maze with 7 places and 21 views. The views numbered 1 – 6 are the ones used for the landmark transposition experiments. S marks the start and goal for the route being learnt; T marks the turning point. Excursions to the unnumbered places are allowed in the exploration phase but are counted as errors in the later parts of the experiment. The "compass" direction shown in the upper right is arbitrary and is included only for comparison with figure 4.

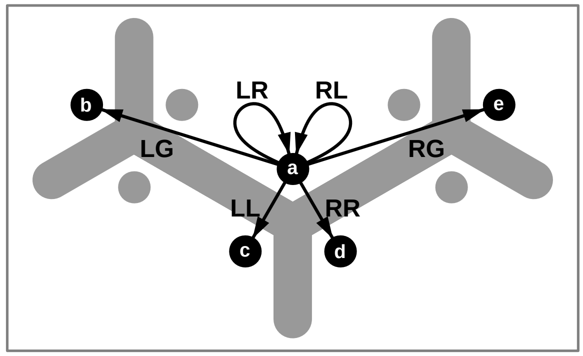


Figure 3: Possible movement decisions when facing the view marked a. L: turn left 60°. R: turn right 60°. G: go ahead to next place.

illumination, or the position of occluders is one of the biggest advantages of virtual reality technology (see van Veen et al 1997). The relation of experiments done in real and virtual environments has recently been reviewed by Péruch and Gaunet (1998).

2 The Hexatown Environment

A virtual town was constructed using Medit software and animated with a framerate of 36 Hz on a SGI Onyx RealityEngine[2] using IRIX Performer software. A schematic map of the town appears in figure 2. It is built on a hexagonal raster with a raster length (distance between two places) of 100 meters. At each junction, one object, normally a building, was located in each of the 120°–angle between the streets; so each place consisted of three objects. In the places with less than three incoming streets, dead ends were added instead, ending with a barrier at about 30 meters. The hexagonal layout was chosen to make all junctions look alike. In contrast, in Cartesian grids (city–block raster), long corridors are visible at all times and the possible decisions at a junction are highly unequal: going straight to a visible target or turning to something not presently visible. The whole town was surrounded by a distant circular mountain ridge which did not provide landmark information. It was constructed from a small model which was repeated periodically every 20 degrees.

Subjects could move about the town using a computer mouse. In order to have controlled visual input and not to distract subjects' attention too much, movements were restricted in the following way. Subjects could move along the street on an invisible rail right in the middle of each street. This movement was initiated by hitting the middle mouse button and was then carried out with a predefined velocity profile without further possibilities for the subject to interact. The translation took 8.4 seconds with a fast acceleration to the maximum speed of 17 meters per second and a slow deceleration. The movement ended at the next junction, in front of the object facing the incoming street. Similarly, turns could be performed in steps of 60 degrees by pressing the left or right mouse button. Again, the simulated movement was "ballistic", i.e., following a predefined velocity profile. Turns took 1.7 seconds with a maximum speed of 70 degrees per second and symmetric acceleration and deceleration.

Figure 3 shows the movement decisions that subjects could choose from. Each transition between two views is mediated by two movement decisions. When facing an object (e.g., the one marked "a" in figure 3), 60°–turns left or right (marked "L", "R") can be performed which will lead to a view down a street. If this is not a dead end, three decisions are possible: the middle mouse button triggers a translation down the street (marked "G" for go), while the left and right buttons lead to 60°–turns. If the street is a dead end, turns are the only possible decision. In any case, the second movement will end in front of another object.

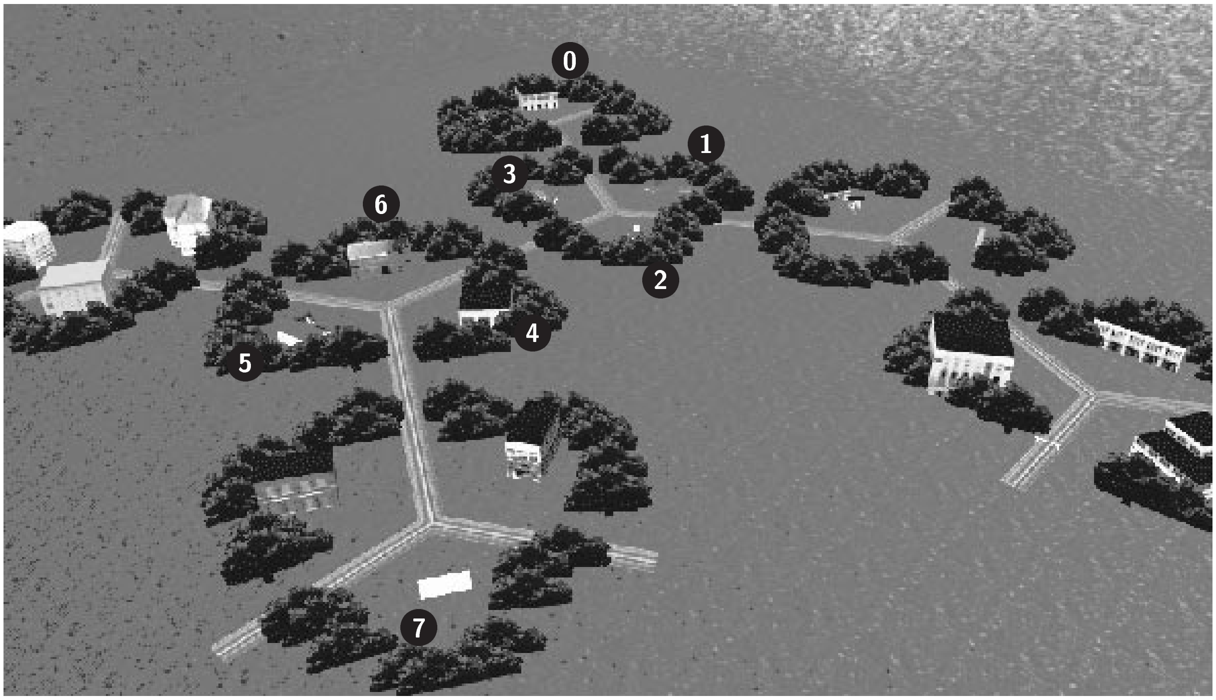


Figure 4: Aerial view of Hexatown. Note that the orientation is different from that of the street map appearing in figure 2. The numbers on black background are the view numbers. The aerial view was not available to the subjects. Object models are courtesy of Silicon Graphics, Inc., and Prof. F. Leberl, Graz.

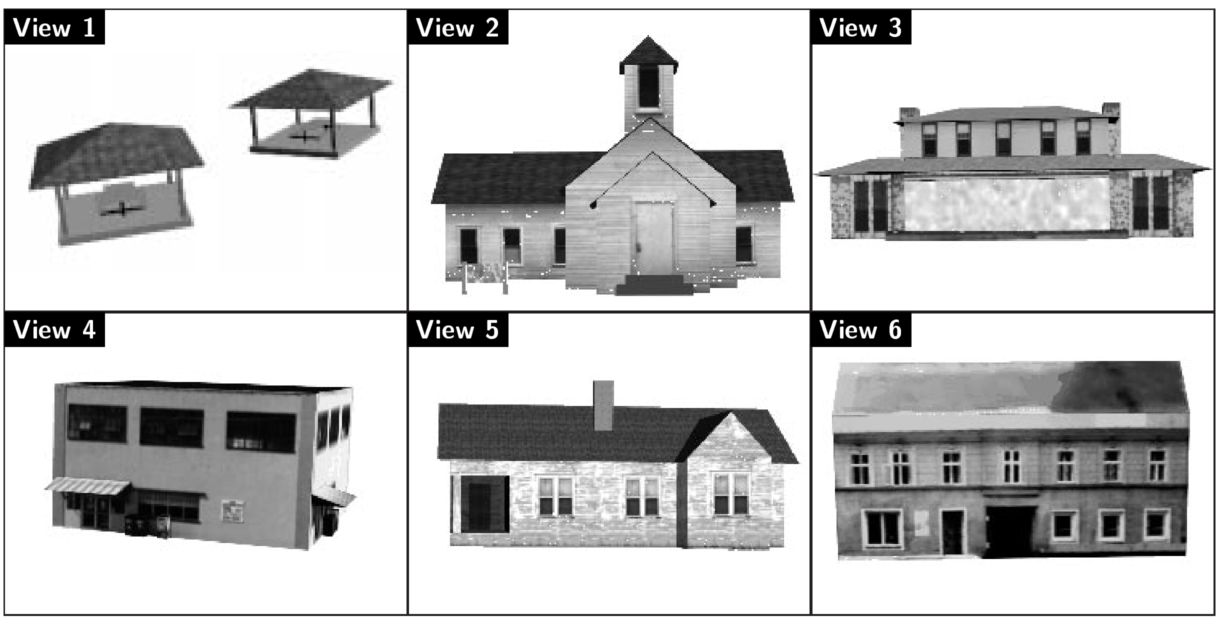


Figure 5: Frontal views of some objects used as landmarks in the Hexatown environment. The objects were located at place A or B in the maze (see Figs. 2, 4) and could be exchanged during the experiments.

An aerial view of Hexatown is shown in figure 4. Central views of the buildings playing a role in the experiments appear in figure 5.

A circular hedge or row of trees was placed around each junction with an opening for each of the three streets (or dead ends) connected to that junction. This hedge looked the same for all junctions and prevented subjects from seeing the objects at more distant junctions. The buildings were at a distance of 15 meters from the junction;

Table 1: Expected performance in the exchange experiment for four possible hypotheses of place recognition.

	C1: control	C2: within place	C3: consistent	C4: conflict
H1: landmark configuration	*max.*	*chance*	*chance*	*chance*
H2: set of landmarks	*max.*	*max.*	*chance*	*chance*
H3: frontal view only	*max.*	*max.*	*max.*	*max.*
H4: view voting	*max.*	*slight reduction*	*max.*	*reduced*

all three buildings were seen at once when passing the hedge and entering the place. The town was illuminated from the bright sky. Taken together, the visibility parameters were the same as in viewing condition 3 of Gillner and Mallot (1998).

3 Rationale of the Experiment

In recognition–triggered responses, recognition might apply to places or to local views. A place is defined either as a configuration of landmarks (structural description) or as the panoramic view visible from the place in question. A local view covers only a fraction of the visual array and its recognition does not necessarily imply the simultaneous recognition of the entire place where the local view occured. In order to distinguish between these two possibilities, we designed an experiment testing the question whether recognition–triggered response implies the recognition of the place where this response occurred. We trained subjects to learn one particular route in the maze as a chain of recognition–triggered responses. The route is marked by the letters $\mathsf{S} \rightarrow \mathsf{A} \rightarrow \mathsf{B} \rightarrow \mathsf{T} \rightarrow \mathsf{B} \rightarrow \mathsf{A} \rightarrow \mathsf{S}$ in figure 2. After training, individual landmarks were replaced in a number of different ways. These exchange conditions were chosen such that the recognition of places and views were affected to different degrees.

We illustrate the exchange conditions used for the approach of view 5 in place B (see figure 6). In all cases, the central view, view 5, would remain unchanged. Four exchange conditions were used in the experiments:

- C1 *control*: No exchanges were done here.
- C2 *within place*: Exchange of left and right peripheral views (i.e., $4 \leftrightarrow 6$). In the training phase, view 6 was either in the right or the central position; its occurence on the left side after mirroring does therefore not provide clear information. For view 4, the situation is different: it occured either on the left or the right side during training and correct turns were always in the direction of its position. Therefore, the information provided by view 4 after mirroring is in conflict with the information provided by the central view 5.
- C3 *across places, consistent*: The peripheral views are replaced by views from another place. The motion decision associated to these replacement views, when they were seen at the same peripheral position during learning, is in agreement with the movement decision associated to the central view. The replacement is: $4 \leftrightarrow 3$ and $6 \leftrightarrow 2$.
- C4 *across places, inconsistent*: As before, but this time, the central view and the replacement views have been associated with different movement decisions during learning. The replacement is: $4 \leftrightarrow 1$ and $6 \leftrightarrow 3$.

These exchange conditions affect the place or scene at which a movement decision has to be taken, to various amounts. In particular, four hypotheses concerning the stored place representation and the correspondingly expected outcome can be formulated:

- H1 *Landmark configuration (structural description or panoramic view):* Spatial memory could involve a structural description of places containing information on the full landmark configuration at each place. If movement decisions are triggered by recognition of these landmark configurations, performance should go down to chance level in exchange conditions C2, C3, and C4, since the landmark configuration is affected in all these conditions.
- H2 *Set of landmarks:* A place could be remembered by the set of landmarks defining it, irrespective of their configuration. In this case, we expect performance to be high in conditions C1 and C2 while performance should drop to chance level in conditions C3 and C4.
- H3 *Frontal view only:* If memory contains only frontal views, performance should be equally high in all exchange conditions.

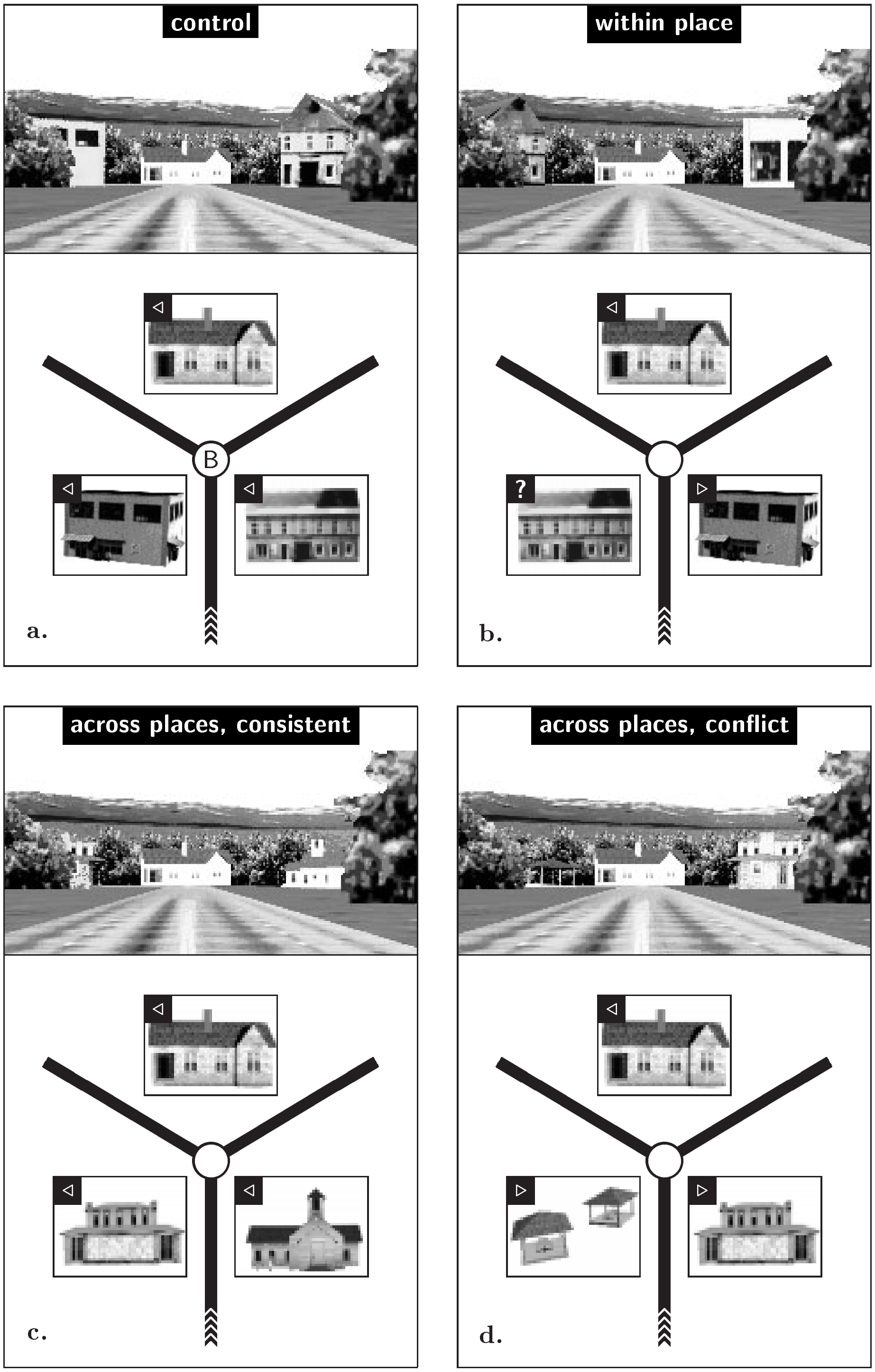


Figure 6: Exchange conditions used in the experiments. For illustration, the approach A → B is shown (release condition R3). ◁: This view in the current position has been associated with left turns during learning. ▷: same for right turns. ?: object did not occur in this position during training. **a.** Control condition without exchange. The place can be recognized as place B and the movement associated with all individual views is left. **b.** Exchange of peripheral landmarks within place. Both place recognition and view–movement associations might be affected. **c.** Consistent exchange across places. Place recognition is affected but view–movement associations are unequivocal. **d.** Conflicting exchange across places. Place recognition is affected and view–movement associations support different movement decisions.

H4 *View voting:* Finally, recognition might apply to views of individual objects, together with their position in the visual field. In this case, memory would contain items like "if view 2 is in the center, turn right" or "if view 1 is to the left, turn right". In this case, we expect that condition C3 should lead to high performance since direction information from all views is unanimous. In contrast, in condition C4 we expect a drop of performance to some level determined by the respective confidence given to the individual movement votes. In the mirroring condition C2, a small drop in performance can be expected since one of the exchanged landmarks (the right one in figure 6b) changes its directional information during replacement and is thus in conflict with the central view.

The expected experimental outcome for each of this four hypotheses is summarized in table 1.

4 Procedure

Experiments were performed using a standard SGI monitor with a visible image diagonal of 19 inch. Subjects were seated comfortably in front of the screen and no chin–rest was used. They moved their heads in a range of about 40 to 60 cm in front of the screen which results in a viewing angle of about 35 – 50°.

The experiment was run on 43 paid volunteers who were students at the University of Tübingen. Three participants realized and reported the landmark replacements. Their data have been excluded from the evaluation.

The experiment was done in three parts. In part 1, subjects were released facing view 0 (see figure 2). A printout of the view marked 7 in figure 2 was given to the subjects and they were instructed to learn the shortest possible way from 0 to 7 and back to 0. Path length was defined as the number of mouse–clicks or movement decisions, where turns are taken into account. In this first part of the experiment, subjects were allowed to explore the entire maze, i.e. they could leave the route. This part was terminated when the shortest possible route was found for the first time.

In the second part of the experiment, subjects were released at one of four positions along the route and transport towards the adjacent place was simulated. The release conditions were

R1: S → A(2): Release at place S and movement towards place A, facing view 2.

R2: B → A(1): Release at place B and movement towards place A, facing view 1.

R3: A → B(5): Release at place A and movement towards place B, facing view 5.

R4: T → B(6): Release at place T and movement towards place B, facing view 6.

In all cases, subjects were asked to continue the route initiated by the approach until reaching either place S or place T, whichever was reached first. This part of the experiment was repeated if the initial decision after releasement was incorrect.

The third part of the experiment was the actual test phase. Here, subjects were released as in the second part. After completing the approach to the adjacent place, they had to decide whether the correct route continued left or right. As always, movement decisions were performed by clicking the appropriate buttons of the computer mouse. In this test phase, however, no feedback was given to the subjects; i.e., after deciding left or right, the trial was terminated.

For each subject, 16 decisions were recorded corresponding to the 4 exchange conditions multiplied by the 4 release conditions (table 2). The sequence of decisions used for half of the subjects was (R4|C1), (R3|C1), (R2|C2), (R1|C4), (R4|C3), (R1|C1), (R3|C4), (R1|C3), (R3|C3), (R2|C4), (R4|C2), (R1|C2), (R4|C4), (R2|C3), (R3|C2), (R2|C1). For the other half of the subjects, the reverse sequence was used. This sequence was put together such that the release positions in subsequent trials were always different. No differences between the results from this sequence and the reverse sequence were found. The data will therefore be presented together.

The experiment was repeated in experiment 2 with a second group of subjects using a different initial arrangement of landmarks. With this control experiment, we attempted to exclude spurious results due to the selection and positioning of the landmarks in the map. The original and control arrangements appear in figure 7a,b.

5 Results

Altogether, 43 subjects took part in the experiments. The first learning phase, which was terminated when the subject had travelled the correct route without error for the first time, took 1 to 7 trials with an average of 2.6 trials. The number of wrong movement decisions (i.e., movements not reducing the number of mouse–clicks needed to reach the goal) occuring during the entire learning phase varied between 0 and 60 with an average of

Table 2: Overview of tests performed during the third part of the experiment. R1 – R4: release conditions. C1 – C4: conditions of landmark exchange. Approach direction is from below. For the control condition (left column), the letters A and B mark the decision place and the correct movement decision is given in the lower right corner.

	C1: control	C2: within	C3: consistent	C4: conflict
R1: S → A(2)	2; A; 1 3; *right*	2; 3 1	2; 5 4	2; 4 6
R2: B → A(1)	1; A; 3 2; *left*	1; 2 3	1; 4 6	1; 5 4
R3: A → B(5)	5; B; 4 6; *left*	5; 6 4	5; 3 2	5; 1 3
R4: T → B(6)	6; B; 5 4; *right*	6; 4 5	6; 1 3	6; 3 2

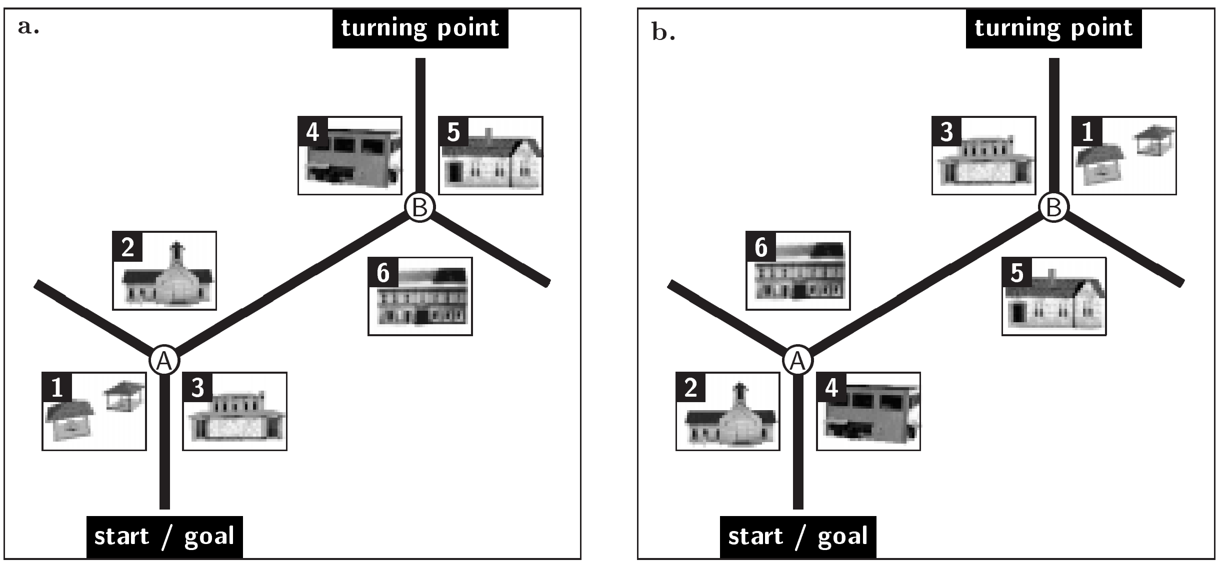
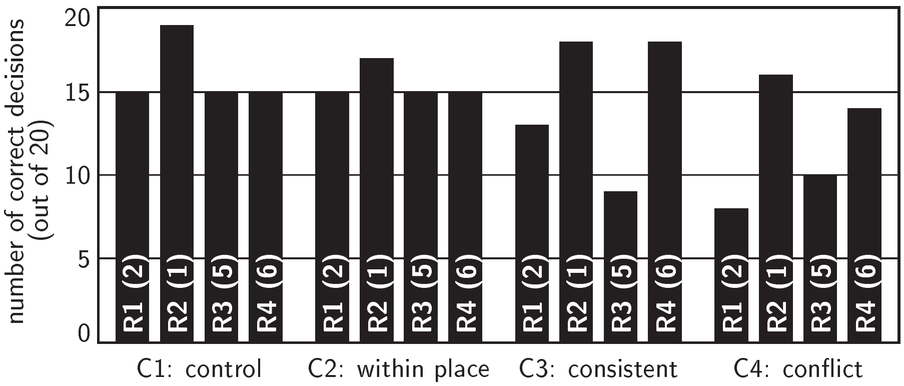


Figure 7: Landmark configuration in the training phase. **a.** Initial landmark layout used in Experiment 1. **b.** Reshuffled landmark layout used in Experiment 2.

10.7. In the second training phase (completion of route from a release point) most tasks were solved in the first trial. The highest number of repetitions necessary in the second phase was 4.

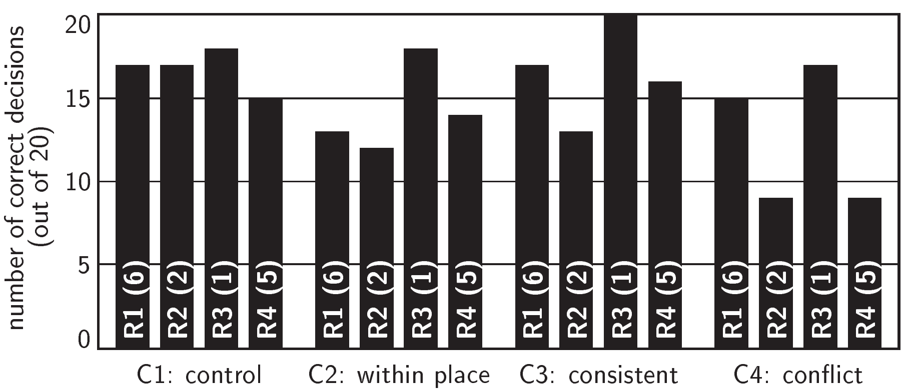
The data from experiment 1 (original landmark configuration as shown in figure 7a) appear in figure 8. In the histogram in the upper part, each column corresponds to one of the 16 test conditions listed in table 2. The height of each column shows the number of subjects choosing the correct movement decision, i.e. the movement decision suggested by the centrally viewed object. Twenty–two subjects participated in this experiment, two of which reported a change in landmark configuration in the test phase. These two subjects were excluded from the analysis in figure 8.

The first four columns show the control condition where no exchanges had been done. In this



	C2: within place	C3: consistent	C4 conflict
C1: original	$F(1,18) = 0.09$ $p = 0.76$	$F(1,18) = 1.10$ $p = 0.31$	$F(1,18) = 11.88$ $p = 0.003$ **
C2: within place		$F(1,18) = 0.51$ $p = 0.48$	$F(1,18) = 8.82$ $p = 0.01$ *
C3: consistent			$F(1,18) = 5.62$ $p = 0.03$ *

Figure 8: Results from experiment 1 (original landmark arrangement). *Top:* Number of correct decisions (in the sense of the centrally presented object). R1 etc: release condition; the number in brackets is the number of the central view. *Bottom:* Analysis of variance of number of correct decisions as a function of exchange condition. Data in condition C4 (conflict) differ significantly from the other conditions.



	C2: within place	C3: consistent	C4 conflict
C1: original	$F(1,18) = 0.20$ $p = 0.66$	$F(1,18) = 0.12$ $p = 0.74$	$F(1,18) = 7.81$ $p = 0.01$ *
C2: within place		$F(1,18) = 1.98$ $p = 0.18$	$F(1,18) = 3.04$ $p = 0.10$
C3: consistent			$F(1,18) = 6.74$ $p = 0.02$ *

Figure 9: Results from experiment 2 (reshuffled landmark arrangement). *Top:* Number of correct decisions (in the sense of the centrally presented object). *Bottom:* Analysis of variance of number of correct decisions as a function of exchange condition. Data in condition C4 (conflict) differ significantly from the other conditions.

condition, 80 % of the decisions were correct. Exchanging landmarks within one place (condition C2) had almost no effect. Consistent exchanges across places (condition C3) led to a reduction of the fraction of correct decisions to 73 %, which, however was not significant (see lower part of figure 8). Conflicting changes across places (condition C4) reduces the fraction of correct decisions to 60 %. As is shown by the analysis of variance in the lower part of figure 8, condition C4 differs significantly from all other conditions, whereas the pairwise differences between conditions C1, C2, and C3 are not significant.

The differences between the columns within one exchange condition reflect different saliences of the central landmarks. If view 1 appears in the center (release condition R3), subjects are more likely to decide in agreement with this central view. On the contrary, view 2 is often outvoted by the peripheral views.

In order to control for possible effects of the initial placement of landmarks, we repeated the experiment with the same landmarks arranged at different positions from the beginning of the experiment (figure 7b). Twenty–one subjects took part in this experiment, one of which reported changes of landmark configuration in the test phase. Again, this subject was excluded from further analysis.

The results from experiment 2 appear in figure 9. Presentation is as in figure 8. Note that the relation of release condition and centrally viewed landmark has changed due to the landmark reshuffeling. The results are well in line with those from experiment 1. As can be seen from the analysis of variance (lower part of figure 9), results in the conflict condition (C4) differs significantly from conditions C1 and C3, whereas differences between conditions C1, C2, and C3 are not significant. Performance in condition C2 is slightly reduced and the difference between C2 and C4 is not significant. Again, view 1 (now in release condition R3) leads to more correct decisions than view 2.

6 Discussion

The results indicate that recognition triggered response does not rely on structural descriptions or panoramic representations of places. The structure of places and even the selection of buildings making up a place can be destroyed without affecting recognition–triggered response. The only condition where a significant effect was found uses a novel combination of views (buildings) associated with conflicting directions during training. This result is consistent with the hypothesis of "view voting", but not with any of the other hypotheses formulated in Section 3. The slight reduction in performance found for exchange condition C2 in experiment 2 may also be expected from the view–voting hypothesis, since some conflict is involved in this condition as well. We therefore conclude that individual buildings or the snapshots taken from these buildings are the recognized landmarks in recognition–triggered response.

This result is well in line with the view–graph approach to visual navigation developed by Schölkopf and Mallot (1995). It states that local views of the maze together with their adjacencies are a sufficient representation of space. In the view–graph, views are connected if they can occur in immediate temporal sequence when exploring the maze. Views occuring in one place are not treated differently from views occuring in adjacent places as long as the temporal sequence constraint is satisfied. In this sense, the notion of a "place" does not exist in this view–based approach. Places can be recovered from the view–graph by more sophisticated analysis, however.

A second important result of the present study is that the directional votes of different views receive different weights. Directions associated with more salient views (such as the picknick huts of view 1) are more likely to be followed by the subjects. The same is true for view 6 (large greenish–yellow building) whereas views 2 and 5 seem to be less reliable. This effect remains after relocating all objects along the route (experiment 2), indicating that this salience depends on the objects themselves, not just on their position.

A third interesting result is that 40 out of 43 subjects did not report the landmark translocations. This is reminiscent of recent findings on change blindness (Simons and Levin 1997) where subjects fail to notice substantial changes to the currently watched scene. Note however, that in our experiment change detection requires a comparison between the current scene and a scene encountered several minutes ago. This scene is presumably represented in a long–term spatial memory, which makes our effect quite different from standard change blindness where working memory is affected.

References

Aginsky V, Harris C, Rensink R, Beusmans J, 1997 "Two strategies for learning a route in a driving simulator" *Journal of Environmental Psychology* **17** 317 – 331

Bülthoff H H, Foese-Mallot B M, Mallot H A, 1997 “Virtuelle Realität als Methode der modernen Hirnforschung”, in *Künstliche Paradise — Virtuelle Realitäten. Künstliche Räume in Literatur–, Sozial– und Naturwissenschaften*, Eds H Krapp, T Wägenbauer (München: Wilhelm Fink Verlag) pp 241 – 260

Cartwright B A, Collett T S, 1982 “How honey bees use landmarks to guide their return to a food source” *Nature* **295** 560 – 564

Cheng K, 1986 “A purely geometric module in the rat’s spatial representation” *Cognition* **23** 149 – 178

Cohen R, Schuepfer T, 1980 “The representation of landmarks and routes” *Child development* **51** 1065 – 1071

Collett T S, Baron J, 1995 “Learnt sensori–motor mappings in honeybees: interpolation and its possible relevance to navigation” *Journal of Comparative Physiology A* **177** 287 – 298

Franz M O, Schölkopf B, Mallot H A, Bülthoff H H, 1998a “Learning view graphs for robot navigation” *Autonomous Robots* **5** 111 – 125

Franz M O, Schölkopf B, Mallot H A, Bülthoff H H, 1998b “Where did I take that snapshot? Scene–based homing by image matching” *Biological Cybernetics* **79** 191 – 202

Gillner S, Mallot H A, 1996 “Place–based versus view–based navigation: Experiments in changing virtual environments” *Perception*, **25**(Suppl.) 93

Gillner S, Mallot H A, 1998 “Navigation and acquisition of spatial knowledge in a virtual maze” *Journal of Cognitive Neuroscience* **10** 445 – 463

Hermer L, Spelke E S, 1994 “A geometric process for spatial reorientation in young children” *Nature* **370** 57 – 59

Jacobs W J, Thomas K G F, Laurance H E, Nadel L, 1998 “Place learning in virtual space II: Topgraphical relations as one dimension of stimulus control” *Learning and Motivation* **29** 288 – 308

Mallot H A, Gillner S, 1997 “Psychophysical support for a view–based strategy in navigation” *Investigative Ophthalmology and Visual Science* **38**(Suppl.) 4683

Morris R G M, 1981 “Spatial localization does not require the presence of local cues” *Learning and Motivation* **12** 239 – 260

O’Keefe J, 1991 “The hippocampal cognitive map and navigational strategies” in *Brain and Space* Ed J Paillard (Oxford: Oxford University Press) pp 273 – 295

O’Keefe J, Nadel L, 1978 *The hippocampus as a cognitive map* (Oxford: Clarendon)

Penna M A, Wu J, 1993 “Models for map building and navigation” *IEEE Transactions on Systems, Man, and Cybernetics* **23** 1276 – 1301

Péruch P, Gaunet F, 1998 “Virtual environments as a promising tool for investigating human spatial cognition” *Cahiers de Psychologie Cognitive* **17** 881 – 899

Poucet B, 1993 “Spatial cognitive maps in animals: New hypotheses on their structure and neural mechanisms” *Psychological Review* **100** 163 – 182

Prescott T, 1996 “Spatial Representation for navigation in animals” *Adaptive Behavior* **4** 85 – 123

Schölkopf B, Mallot H A, 1995 “View–based cognitive mapping and path planning” *Adaptive Behavior* **3** 311 – 348

Simons D J, Levin D T, 1997 “Change blindness” *Trends in Cognitive Sciences* **1** 261 – 267

Steck S D, Mallot H A 1998 “The role of global and local landmarks in virtual environment navigation” Technical report 63, Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany

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

van Veen H A H C, Distler H K, Braun S J, Bülthoff H H, 1998 “Navigating through a virtual city: Using virtual reality technology to study human action and perception” *Future Generation Computer Systems* **14** 231 – 242

Zipser D, 1985 “A computational model of hippocampal place fields” *Behavioral Neuroscience* **99** 1006 – 1018