Virtual Reality (2007) 11:161–173
DOI 10.1007/s10055-006-0062-2

ORIGINAL ARTICLE

M²S maps: supporting real-world navigation with mobile VR

Xiaolong Zhang

Received: 1 February 2006 / Accepted: 13 November 2006 / Published online: 8 December 2006

Abstract Mobile devices are becoming increasingly integrated into our society. In addition to entertaining people with applications like pervasive games, mobile devices can also help to address cognitive challenges people face in the real world. This paper, by drawing on research findings from cognitive psychology and geography, explores a design to use mobile VR to help people overcome one cognitive barrier in navigation, which is to establish the correspondence between 2D spatial information found in maps and 3D entities they perceive from the real world. The design offers users multi-format, multi-scale, and semantic (M^2S) maps, ranging from 2D maps to 3D immersive environments, and helps users to connect 2D maps to the real world through 3D environments which are equipped with semantic representation and animation techniques. Consequently, users can apply various kinds of spatial knowledge, 2D or 3D, in understanding the real world as well as assisting in navigation. This research enhances the design repertoire of mobile VR, and suggests a way to integrate virtual environments into people's real-world life by examining the cognitive implications of 3D models on users' activities.

Keywords Virtual environments · Mobile computing · Multi-scale · Navigation · Spatial cognition

X. Zhang (✉)
College of Information Sciences and Technology,
The Pennsylvania State University, 307D IST Building,
University Park, PA 16802, USA
e-mail: lzhang@ist.psu.edu

1 Introduction

Mobile devices have become increasingly integrated into people's lives. Pervasive games bring gamers to the real world (Flintham et al. 2003; Magerkurth et al. 2005); mobile tourist information systems help people access location-aware information (Burigat and Chittaro 2005; Bruntsch and Rehrl 2005); education contents are delivered through handheld devices (Finn 2004); and social interactions are expanded by mobile applications (Dearman et al. 1929; Smith et al. 2005; Reid et al. 2004). Mobile devices can also help to address some cognitive challenges in the real world. For example, blind people can carry a mobile navigation system equipped with audio instructions to guide their wayfinding (Slavík and Berka 1999).

Most mobile navigation systems are based on the map metaphor. Although voice instructions are provided by some systems, visual information, which appears as 2D maps, still plays a dominant role in supporting navigation. However, people face some challenges in using 2D maps to guide actions in the 3D physical world. It is not easy to translate spatial structures on maps, which are presented in a birds-eye view, to objects in the real world, which are perceived through an immersive view. For example, Fig. 1 shows a scenario where a map is used to guide navigation at an intersection, at which a person has to choose among four possible directions. The map in Fig. 1a clearly shows the spatial relationship among different directions at this intersection, but it is hard to perceive this relationship from a real-world view (Fig. 1b). Matching 3D objects and structures in the real world with their 2D representations on maps is not easy. People need to

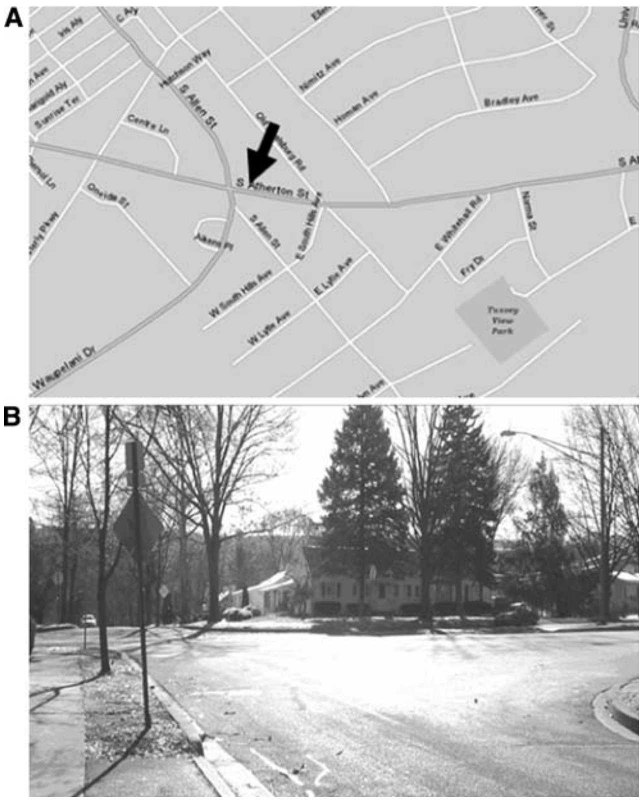


Fig. 1 Difficulty in matching a map and the real world. **a** Map of an intersection (from http://www.mapquest.com). **b** View of the interaction in the real world

find common objects as references to connect the physical world and the map

Street names are often used as a reference to link maps and the real world. However, sometimes people may find it a problem to use street names to connect maps to the real world. In some cities (e.g., London), it could be hard to find street signs. In some places, such as an intersection in a US town shown in Fig. 2, street signs are simply missing. In these situations, it is impossible to use a map without knowing current location.

Fig. 2 An intersection with missing street sign

One way to address the challenge in connecting a map to the physical world is to provide people with reference objects that can be easily discovered and recognized. For example, in Fig. 2, although street signs are not present, a large building at a corner is very distinct and can be used to help people understand their location. Some tourist maps offer people 3D drawings of important landmarks. However, these maps are usually designed for specific purposes (e.g., finding those famous landmarks) in a very small area. For general-purpose maps, it would be a daunting task to include hundreds or even thousands of 3D models. Actually, clustered 3D models may block street views and affect other tasks, such as route planning, which more rely on 2D spatial information.

While people may need different kinds of spatial information for navigation, 2D for route planning and 3D for object matching, it would be ideal if maps could be tailored for these different needs. Paper maps are static and cannot provide such on-demand services. For a computerized navigation system, however, integrating 3D models into 2D maps is feasible.

In this paper, we explore a design of mobile navigation systems which uses mobile VR to support real-world navigation with multi-format, multi-scale, and semantic (M^2S) maps. This M^2S approach, based on the principles of human spatial cognition, presents spatial information in traditional 2D maps, which people are familiar with, in 2D maps in perspective, which are aligned with people's movement direction in the real world, and in 3D immersive environments, which provide detailed information for decision-making and help people connect the real world with all other maps. Seeing maps in different formats, at different scales, and with different semantic implications, will help users construct a comprehensive understanding of the environment and improve their navigation performances.

The structure of this paper is as the following. The paper first reviews relevant literature, and then advances to the design of M^2S maps and the implementation of the M^2S map prototype. After the description of a user study, the paper discusses unsolved issues and the implications of this research for design.

2 Related research

Relevant research fits into four categories: human spatial cognition, navigation support in the real world, navigation support in virtual environments, and location-based services in mobile computing. Spatial cognition is of interest because it provides theoretical foundations for understanding human navigation behaviors. Research on navigation support in both the real world and virtual environments offers insights into what design problems should be considered and what approaches may or may not work. Because of our focus on navigation support with mobile devices, it is needed to examine relevant research in mobile computing.

2.1 Spatial cognition

It is widely accepted that spatial knowledge plays a key role in navigation (Downs and Stea 1973). Spatial knowledge has been argued to be stored and structured as cognitive maps (Tolman 1948) to represent the external environment—its structure, the entities in it, and their spatial relations (Golledge 1999). Spatial knowledge has three different forms: (1) landmark knowledge, which relates to distinct features of prominent objects, (2) route knowledge, which helps people formulate movement paths, and (3) survey knowledge, which provides a map-like overview of a region (Piaget and Inhelder 1967; Hart and Moore 1973; Thorndyke and Hayes-Roth 1982). People's spatial knowledge is often constructed as a hierarchy with different features of space at different levels (Stevens and Coupe 1978; Hommel et al. 2000), allowing effective access to necessary spatial knowledge in problem solving (Hommel et al. 2000; McNamara et al. 1989).

Navigation can be divided into different subtasks that require different kinds of spatial knowledge. Loomis and Beall (1998) argued that navigation includes such subtasks as formulating a route plan, which requires survey knowledge, and evaluating local spatial structures in route following, which needs route and landmark knowledge. Timpf and Kuhn (2003) argued that navigation is a hierarchical process including such spatial tasks as goal planning, strategy choosing, and moving. These tasks rely on spatial information at different scale levels, from the global view of a region to local details of routes.

To support navigation activities, tools in the real world largely focus on facilitating the acquisition of spatial information and the integration of different kinds of spatial information.

2.2 Navigation support in the real world

Maps are the most popular navigation tools in the real world, as external artifacts for spatial knowledge acquisition (Bagrow 1985) and as cognitive interfaces to external environments (Barkowsky and Freksa 1997). Being a map user, however, a person needs to establish element-to-element correspondence between objects on maps and entities in the real world and to recognize spatial structure correspondence between maps and the world (Newcombe and Huttenlocher 2000).

However, tying objects on 2D maps with 3D entities in the real world is not easy. They appear in different forms and are perceived through different cognitive processes, and these differences lead to significant cognitive loads in building a connection between maps and real-world environments (Thorndyke and Hayes-Roth 1982; Evans and Pezdek 1980; Levine et al. 1985). To address this issue, many tourist maps offer people 3D drawings of important landmarks. However, these designs still assume map users know how to identify their physical positions on maps and translate a route plan on maps into physical movement in the 3D world. Also, integrating 3D models into general-purpose maps may introduce new problems. For example, streets may be blocked by 3D models. People may find it difficult to plan routes.

Online map services address this issue with other approaches. Google maps allow users to put maps and satellite images together. Although such hybrid maps provide a way to connect the real world and maps, satellite images present real-world objects in a perspective different from the one people have in real life. People may still find it difficult to connect objects in satellite images with objects in the real world. Online services (e.g., http://www.mapquest.com) also provide people with turn-by-turn instructions. Drivers do not need to match maps with the real world as long as they can stay on course. They may not even need maps at all. However, such turn-by-turn instructions are only for a specific route, and do not provide landmark and survey knowledge to help the construction of comprehensive cognitive maps related to the surrounding environment, which is important to other spatial tasks, such as communicating with others about the environment.

Navigation systems based on Global positioning systems (GPS) move one step further. They give users voice instructions on turns, and also provide maps. In using such GPS-based systems, however, drivers largely rely on voice prompts and do not read maps often. Also, landmarks are seldom used in such systems.

While online map services and GPS-based navigation systems may be sufficient to drive to a specific place, they are not suitable for more exploration-based navigation, in which users need richer spatial information to better understand environments. For example, sometimes planned routes are not valid due to road constructions or the change of street patterns (e.g., from two-way to one-way). Then, drivers have to re-plan routes based on the road conditions. Under such situations, exploration probably is the best solution. Then, drivers would need comprehensive spatial information to guide navigation.

2.3 Navigation support in virtual environments

Research projects on navigation in virtual environments have different focuses: supporting locomotion control in virtual environments, supporting the presentation of spatial information, or supporting the construction of spatial structures. Our interest is in using mobile devices to present appropriate spatial information, so the focus here is on research with an emphasis on the presentation of spatial information.

Research on spatial cognition and navigation behaviors in the real world has a great influence on navigation design in virtual environments. It has been found that spatial cognition in virtual environments is similar to that in the real world (Ruddle et al. 1997; Wilson et al. 1997; Witmer et al. 1996). Theories on virtual navigation (Darken and Sibert 1996; Jul and Furnas 1997; Chen and Stanney 1999) suggested that navigation in a virtual environment could be a multiple-level process, similar to a person's behavior in the real world (Loomis and Beall 1998; Timpf and Kuhn 2003). Such similarities lead to the direct applications of design principles for real-world navigation support in virtual environments. Efforts have been made to help people access survey knowledge (e.g., overviews in 3D games), route knowledge (Elvins et al. 1997), and landmark knowledge (Vinson 1999; Pierce and Pausch 2004). Overview maps in virtual environments can also be automatically aligned with the direction of locomotion (Darken and Cevik 1999), giving users track-up maps, just like maps provided by GPS-based navigation tools in real life. Consideration has also been given to the improvement of the organization of spatial structures (Darken and Sibert 1996), based on a theory in architectural design: a better organized environment is easy to navigate (Lynch 1960).

Despite these efforts, some problems are still left unaddressed. It is a challenge to establish element-to-element and spatial structure correspondence between 2D overviews and 3D objects in immersive environments. Although some systems offered 3D scaled-down configurations (Leigh et al. 1996; Stoakley et al. 1995), these configurations usually appeared with a few predefined scales and did not allow users to fully control their sizes. Users may still face problems like observing necessary details of 3D objects or getting an adequate overview of the surrounding environment.

Furthermore, although it is known that different navigation subtasks, such as planning and movement assessment, require different kinds of spatial information (e.g., small scale maps with a broader view for planning and large scale maps with street details for assessment), the presentation of spatial information in most virtual environments usually does not reflect users' navigation tasks. Few designs consider providing users with spatial information at proper scales and with proper details according to what users are doing. Some research designed multi-scale tools to let users control at what scale to observe an environment (LaViola et al. 2001; Zhang and Furnas 2005). However, direct scale control may distract users from their primary navigation tasks due to the potential dramatic view changes in scaling.

2.4 Using mobile devices to support navigation

Mobile devices have been used to support navigation. Most often seen designs are GPS-enabled tourist information systems, which provide tourists with location-aware, supplemental information, such as historical, cultural, and social backgrounds, of points of interest (Burigat and Chittaro 2005; Bruntsch and Rehrl 2005; Krüger et al. 2004; Yue et al. 2005). Mobile devices can also be used to deliver multimodal information to assist people, such as the blind, who cannot use or have difficulty using conventional navigation tools (Slavík and Berka 1999).

Mobile devices have been used to enhance the presentation of different kinds of spatial information, such as landmarks (Sefelin et al. 2005; Goodman et al. 2004; May et al. 2003), route information (Kray et al. 2003; Narzt et al. 2003; Malaka and Zipf 2000), as well as maps. However, these designs usually focus on just one particular type of spatial information, and do not consider their integration, which is important to the construction of cognitive maps.

The advances of hardware and software have accelerated research on using mobile virtual environments to support navigation (Suomela et al. 2003). Some projects presented 3D models in a view separated from a view of 2D maps (Rakkolainen and Vainio 2003). Under such designs, users still need to figure out the relationship between 2D maps and 3D virtual environments. Also, having two separated views for 3D and 2D information on an already small screen on mobile devices further squeezes the screen space for each view, preventing users from obtaining sufficient context and content information. Some research projects take a mobile-reality approach to create a mixed reality with mobile devices (Goose et al. 2003; Narzt et al. 2004). While these projects consider 3D graphics, they do not focus on navigation support and do not consider the integration of 3D structures with 2D maps.

It should be pointed out that there are other streams of research on mobile devices and virtual environments or mixed reality, with focuses on different issues, such as navigation control in immersive virtual environments (Benini et al. 2002; Watsen et al. 1999; Marsden and Tip 2005) and pervasive games (Flintham et al. 2003). Although not directly related to our research, these projects offer useful insight into what contextual factors should be considered in design.

3 Using M^2S maps to enhance spatial cognition

To assist people in establishing correspondence between maps and the real world, we propose a design approach to present spatial information in a multi-format, multi-scale, and semantic way. This M^2S approach integrates 3D environments with 2D maps and considers the transition between different types of spatial information. In this section, we first present the theoretical foundation of the M^2S design, and then discuss the design and its features.

3.1 Theoretical foundation of M^2S approach

Existing research literature points out some ways to help people connect maps to the real world. First, while people need traditional 2D maps for planning routes, more concrete and detailed landmark knowledge would better facilitate the establishment of object correspondences (Raubal and Winter 2002; Elias and Sester 2002). Thus, in support of a navigation process with such subtasks as route planning, route following, and movement direction assessment, navigation tools should provide users with spatial information in multiple formats, including 2D maps, which can help route planning, maps aligned with moving direction, which can help route following, and 3D models, which can facilitate the assessment of movement and current location.

Second, route planning, route following, and movement assessment also require spatial information at different scale levels (Timpf and Kuhn 2003). For example, the global configuration of an area in 2D is usually required for route planning, but local details of streets are better for movement assessment. Thus, navigation tools should aggregate spatial information at multiple scale levels, from global to local, and provide spatial information at a scale appropriate to a particular task.

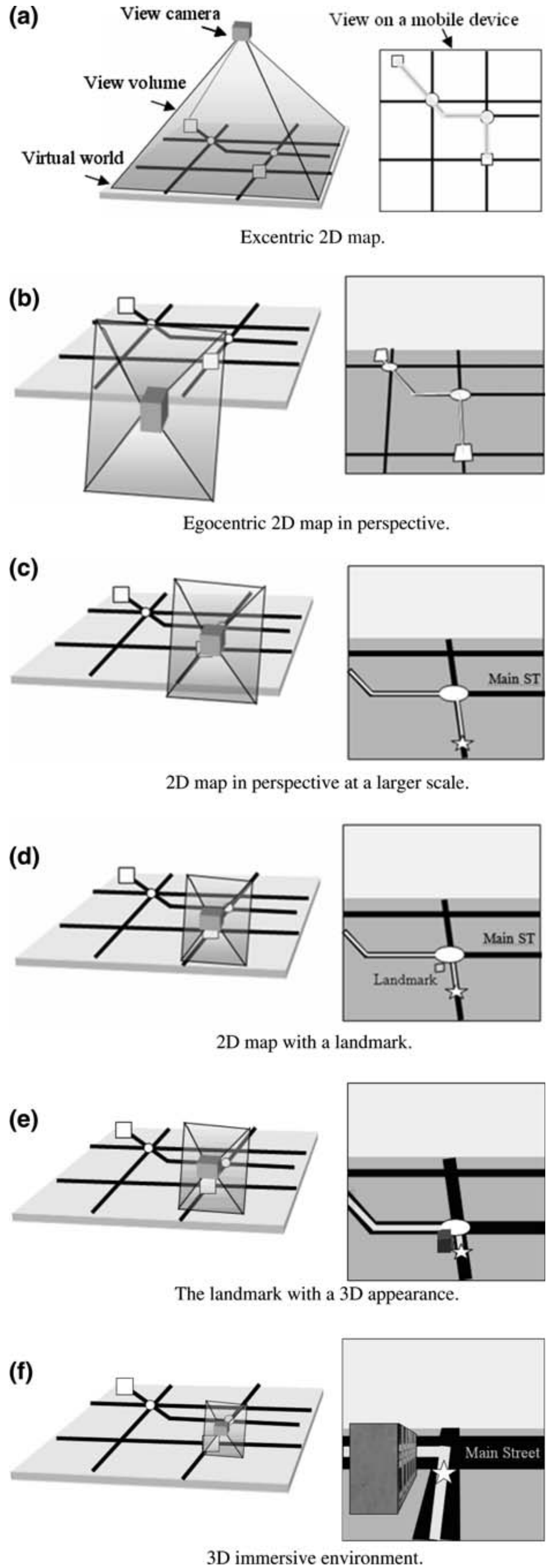
Furthermore, when spatial information is presented in different formats (e.g., 2D maps or 3D models), users need to know their relationship. Integrating spatial knowledge from maps with spatial knowledge from physical experience can be difficult, because in different forms of spatial knowledge, the same objects are represented in different ways and through different cognitive processes (Thorndyke and Hayes-Roth 1982; Evans and Pezdek 1980; Levine et al. 1985). Thus, navigation tools should also exhibit the semantic relationship between these presentation formats so that people can easily interpret them and tie them together.

In summary, to help people establish connections between maps and the real world, it would be desirable to have navigation tools that present spatial information in rich formats including 2D maps, 2D maps in perspectives, and 3D models, at different scales from global to local, and with semantic information that helps users see the transition between different formats of spatial information and at different sales. This multi-format, multi-scale and semantic approach lays the foundation for the design of M^2S maps.

3.2 Design of M^2S maps

Based on the above analysis, we developed an M^2S design seen in Fig. 3. This design provides users with four types of spatial information at various levels of details. Users can have 2D maps, 2D maps in perspective, 2D maps in perspective enhanced with 3D models, and immersive virtual environments with detailed 3D models. A map, 2D or 2D in perspective, can be scaled up and down so that users can obtain necessary spatial information appropriate to their tasks (e.g., route planning or route following). The transition between neighboring types, such as 2D maps and 2D maps in perspective, is designed to be continuous and semantic so that people can easily find common objects as references to connect them together, rather than see a sudden jump.

Figure 3 demonstrates what users can see in a simple navigation process, which includes route planning, route following, and movement direction choosing at intersections. Figure 3a shows an exocentric overview of a region for route planning. The left image illustrates the relationship between the position of the view camera and the region. The image on the right is a user's view on the screen of a mobile device. The view camera is high enough to give users an overview of the region and a route that connects a starting point to a



◀ **Fig. 3** M^{2}S maps. **a** Exocentric 2D map, **b** egocentric 2D map in perspective, **c** 2D map in perspective at a larger scale, **d** 2D map with a landmark, **e** the landmark with a 3D appearance, **f** 3D immersive environment

destination (squares) through some intersections, which are decision-making points (circles). When users start moving, they will see the 2D exocentric map changes to a 2D map in perspective (Fig. 3b). This map is egocentric and aligned with the navigator's movement direction. The view can be produced by simply changing the camera position and orientation. As they move, users will have a 2D map in perspective at a larger scale (Fig. 3c), which can be produced by reducing the view camera elevation. In this view, users will see objects increase in size and more map details.

Transitions between these views in Fig. 3 (the right images from Fig. 3a–c) can be made continuously by using 3D animation techniques. These views are produced by positioning the viewing camera to different locations. Moving the camera along a continuous function that interpolates above view camera locations will lead to an animated view transition. Techniques to develop such interpolation functions are mature (Parent 2001; Watt and Policarpo 2001). Most GPS-based navigation systems can provide exocentric and egocentric 2D maps, but only a few consider 2D maps in perspective and the smooth transition between views.

What distinguishes M^{2}S maps from existing designs are the views from Fig. 3d–f. When users approach an intersection and begin to make a decision about which direction to go, they will see a newly added landmark object (Fig. 3d). At first, the landmark appears as a 2D object, merely indicating its location. Moving closer to the decision-making point, the landmark will take a 3D shape, roughly mimicking its real-world appearance (Fig. 3e). Finally, when the user gets close to the decision-making point, the view camera is even lowered to generate immersive experiences, in which users find more details about the landmark and the surrounding environments (Fig. 3f).

3.3 Features of M^{2}S maps

As seen, M^{2}S maps provide users with spatial information in different formats. People can use exocentric 2D maps to obtain necessary survey knowledge for route planning, use egocentric 2D maps in perspective to examine more detailed information for route following, and use 3D objects in the immersive environment to connect maps and the real world.

In addition to formatting differences in the above views, M^{2}S maps present spatial information at different scales. The overview map is at a smaller scale to show the spatial configuration of the region and the whole route. Then, larger scale maps provide more detailed information and support route following. At decision-making intersections, spatial information is presented at the largest scale and with the most details.

Furthermore, semantic and smooth transitions are provided to help people tie different views together and build a more comprehensive understanding of the region. View changes are made smoothly by animating the position change of the view camera. Landmarks come with different appearances gradually, which semantically enrich spatial information presented. These different kinds of spatial information at different scale may help users to interpret connections between 2D objects (or structures) on maps and their 3D counterparts in the real world. Rich 3D entities in the immersive environment help users quickly establish the correspondence between the virtual environment and the real world. Smooth and semantic transitions between the virtual environment and maps allow users to tie 2D maps with the 3D virtual environment, as well as the real world.

The semantic representation of landmarks can be regarded as one kind of level of detail (LOD) techniques in 3D virtual environments, but with a different purpose. Similar to traditional LOD techniques, the semantic representation visualizes the same object with different geometric models based on a particular interaction parameter. However, traditional LOD designs are about computation efficiency (Puppo and Scopigno 1997). The semantic representation more emphasizes cognitive needs for spatial information by people. Our design not only renders the landmark with different geometries and textures, but also presents the same environment with different semantic abstractions. For example, the view transition from Fig. 3b, c includes not only the size change of objects, but also new information, such as landmark information, which helps users organize spatial objects in the area.

The design presented in Fig. 3 only shows what users can see from a starting position to a decision-making point. A navigation process may follow a route with several decision-making points. The route can be divided into individual segments, each starting from a known position and ending with a decision-making point. The design shown in Fig. 3 can then be applied in each segment.

Such a semantic integration and repetition of different types of spatial information at different scales will not only help people reach the destination quickly, but also facilitate the construction of cognitive maps about the region, which in the long run will benefit other spatial tasks, such as route re-planning. As discussed previously, the user's spatial knowledge appears in different forms, and is structured hierarchically. Research on behavior associated with the use of map shows that presenting the same spatial information in different ways can help people encode and organize spatial knowledge (Hommel and Knuf 2003). For example, the different appearances of a landmark in 2D and 3D may be more helpful than only showing its 2D location or its 3D shape. In particular, semantic transformation between its 2D and 3D appearances can help people understand the landmark from different aspects and in a coherent way. Furthermore, individual landmarks at each decision-making point are emphasized in different formats and scales. These landmarks may work together as snap shots of important objects on a route, and improve the construction of route knowledge (Elvins et al. 1997) and cognitive maps.

Another feature of the M^2S maps is its context-aware design. The transformation among different maps is mediated by a user's action context. The concept of integrating people's working contexts into the design of computing systems emerged in the 1990s (Schilit et al. 1994). Context is usually referred to as situational information which can be used to provide ‘‘task-relevant'' services (Abowd et al. 1999), by tailoring the presentation of necessary information, mediating services (Chalmers et al. 2004), or triggering actions (Schmidt et al. 1999). Context-aware design has been extensively explored in mobile computing by considering such issues as infrastructure factors, application factors, system factors, and geographic location factors (Rodden et al. 1998; Headon and Curwen 2002; Baus et al. 2002).

Most mobile navigation systems have focused on such contextual factors as a user's current location and the location's historical, cultural, and social backgrounds (Burigat and Chittaro 2005; Bruntsch and Rehrl 2005; Krüger et al. 2004; Yue et al. 2005). These designs have greatly expanded the design space of using mobile devices to supply context information, but most of context information in these designs do not concern navigation processes. Instead, contextual information is usually only about the final destination.

Our design considers such contexts as the types of navigation tasks. A user's navigation task may change often from goal planning, strategy choosing, and moving. These different tasks will require spatial information with a proper format (e.g., 2D maps or 3D models) and at a proper scale (e.g., a global view of a region or a detailed view of landmarks) (Loomis and Beall 1998; Timpf and Kuhn 2003). Design approaches based only on a user's location context are not suitable, because they do not provide users with relevant information in advance. For example, a tourist information system only needs to give tourists information when they are actually at a point of interest.

In navigation guidance, it is important to let users receive necessary information at the right place and right time. These timing and location issues are critical to our design. M^2S maps are intended to provide users with information at proper physical locations while leaving sufficient time for people to understand the information and take necessary actions. Which format to use in the presentation of spatial information is based on how far away a user is to a decision-making point and how long it may take the user to get there.

4 Implementation of an M^2S map prototype

An M^2S map prototype system was implemented on a Dell Axim x50v Pocket PC (Intel XScale 624MHz CPU, 64MB RAM, Intel 2700G graphics processor with 3D acceleration support). A Bluetooth GPS receiver was used to obtain information about user's actions. The operating system was Windows Mobile 2005. In this section, we discuss issues concerning system architecture, context information handling, and 3D graphics programming.

4.1 Architecture

Figure 4 shows the architecture of the prototype. The system has six modules. A GPS module provides a user's location and movement information. A 2D-map module and a 3D-model module supply different types of spatial information. A path module stores path information, including locations of decision-making points. A user-interface module is implemented to take input from users. The central component of the system is a context-determination module. It determines what tasks a user may carry out based on the user's position and movement and the user's distance to the next decision-making point, and then chooses an appropriate format of spatial information for display from the 2D-map or 3D-model module. The interface module supports user interaction and allows users to have a more flexible control on the formats of spatial information. For example, if needed, a user can temporarily change an immersive view to an overview map by pressing a button.

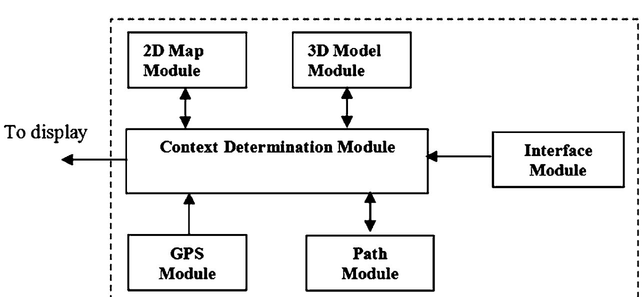


Fig. 4 System architecture of an M^2S map prototype

4.2 Action-context handling

The context factor is the time that a user needs to reach the next decision-making point. It mediates the format and scale of spatial information presentation. Because of the lack of theories on the division of navigation subtasks, we adopted an ad-hoc approach based on our experiences with commercial GPS-based navigation systems. Currently, many commercial navigation systems offer voice prompts to advise a driver about the next coming turn. These prompts come at different timings.

We surveyed several popular GPS-based navigation systems, including iGuidance, Mapopolis, and TomTom. We found out that timings vary from system to system and from speed to speed. It seems most systems hard-code timings for voice prompts based on driving speed and predefined sets of distance (e.g., 2 miles, 1 mile, 1,000 feet etc.). Only a few systems (e.g., CoPilot Live) allow users to adjust the timings relatively. Although we could not find an exact timing scheme agreed by most systems, we did notice a pattern of when voice prompts came out. Usually, voice prompts were provided when the next turn would be reached in about 1–2 min, 10–30 and 1–5 s.

Our design shown in Fig. 3 indicates four timings for presentation transformation: from 2D maps to 2D maps in perspective, from 2D maps in perspective to 2D maps with the location of a landmark, from a 2D landmark to a 3D-shape landmark, and from a simple 3D-shape landmark to a 3D landmark with a comprehensive appearance in an immersive environment. The first transition can be triggered by movement. Whenever a user starts moving, the presentation can be changed from a 2D map for route planning to a 2D map in perspective for route following. For the three transformation timings, we designed a timing scheme based on our observations on commercial GPS-based navigation systems.

The context-determination module works according to timing scheme in Table 1. The context-determination module obtains information about a user's action (e.g., location, movement speed, and movement orientation) from the GPS module, and the location of

Table 1 Action context and map formats

Action context	Map format
No movement	Exocentric 2D map
Movement and $t \geq 60$ s	Egocentric 2D map in perspective
Movement and $60 \text{ s} \geq t \geq 30$ s	2D map in perspective + 2D landmark
Movement and $30 \text{ s} \geq t \geq 10$ s	2D map in perspective + 3D landmark
Movement and $t \geq 10$ s	3D immersive environment

the next decision-making point from the path module. Then, with such information, the context-determination module can calculate the time, *t*, a user needs to reach the next decision-making point based on the distance to that point and the current speed. The user's current location and the time t are then used to retrieve maps or 3D models.

4.3 3D programming

One of the key issues in 3D programming is the choice of 3D graphics libraries. Some mobile virtual environments used the high-level virtual reality modeling language (VRML) for 3D modeling (Burigat and Chittaro 2005; Krüger et al. 2004). While VRML can simplify the construction of 3D scenes, it is weak in support of complicated user interactions, in particular time-critical interactions. Having a concern with this issue, we decided to use lower-level graphics libraries (GLs).

Graphics libraries for mobile devices are not as rich as those for desktop platforms. PocketGL (SundialSoft et al. 2005) and TinyGL (Bellard et al. 2005) are two subsets of OpenGL for Pocket PCs, but they seem outdated and limited. Other GLs are largely built for games on dedicated platforms, and can hardly be used for generic purposes.

To leverage available 3D acceleration hardware, we chose the Microsoft Mobile Direct3D, which, as a subset of Direct3D, provides strong support for 3D graphics. It can be integrated with GPS services under the .NET Compact Framework. We implemented the prototype with C#.

5 User study of the M^{2}S prototype

We conducted a study on the use of the M^{2}S prototype in a navigation process with just one decision-making point. The purpose of this study was to examine how M^{2}S maps could be used and what design issues users may have, so we adopted a more qualitative method which includes field observation and interview.

5.1 Subject and study settings

Two subjects were recruited. The pre-test questionnaires indicated that one subject had GPS experience, but the other did not. Both subjects were requested to use the prototype system to reach a destination. During the task, subjects needed to choose a direction among four possible options at the decision-making point. A researcher accompanied subjects during the navigation processes. Comments and behaviors of subjects in navigation were recorded.

Figure 5 captures four key frames of the M^{2}S maps subjects used. Figure 5a is an exocentric map with the whole route. Figure 5b is an egocentric 2D map in perspective which shows the decision-making point and the location of a landmark building in 3D. Figure 5c presents the landmark with a 3D shape, while Fig. 5d is

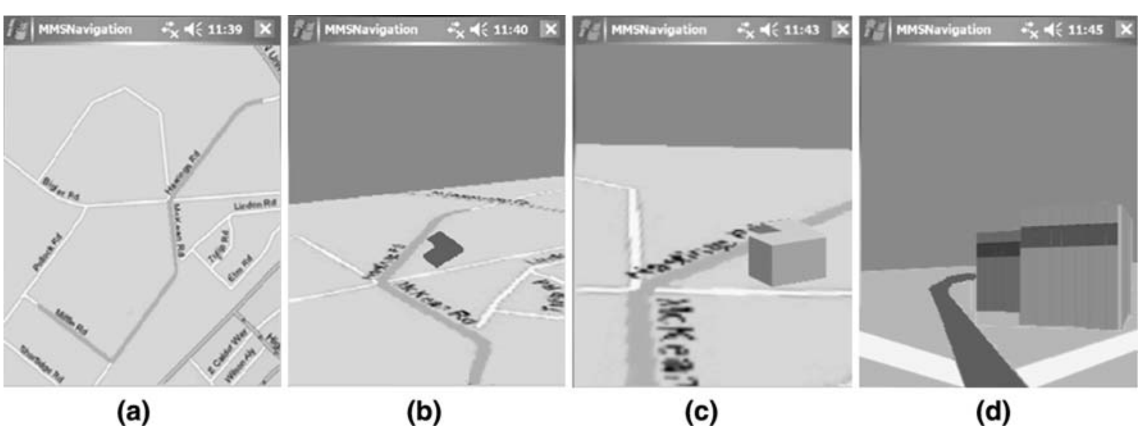


Fig. 5 Screen shots of M^{2}S maps from usability observation **a** 2D map, **b** 2D map in perspective with the landmark, **c** 2D map in perspective with the 3D landmark, **d** immersive view of the landmark

an immersive environment with more detailed information about the landmark.

After navigation, both subjects were interviewed. Interview questions focused on their feeling towards the system and their behaviors in navigation.

5.2 Field observation and interview

Both subjects successfully finished the task, but their comments and the behaviors were slightly different. Both of them seemed familiar with the 2D map presented for route planning in the beginning. The transition from the initial exocentric 2D map to the egocentric map as soon as subjects started moving surprised the subject without GPS experience, but not the subject with GPS experience. When the landmark appeared as a 3D model on the screen, the subject with GPS experience paused, and then moved on. The other subject did not stop at all. At the intersection, both subjects stopped and switched their views between the surrounding environment and the immersive scene on the screen. Eventually they both made the right choice.

During interviews, both subjects felt that the M^2S maps were a new design that gave them fresh experience. When asked about their perception of the effectiveness of this design, the subject with GPS experience said that using 3D models in navigation was impressive. The subject without GPS emphasized the automatic alignment of maps with the movement in the real-world as well as the gradual transformation between 2D and 3D.

Both subjects indicated that presenting the landmark in various formats was an innovative way to help their decision-making. The subject without GPS experience pointed out that the 3D model on top of a 2D map did a good job in helping to understand where the landmark building was on the map. The subject with GPS experience said that although he had no problem in using 2D maps, the 3D model could indeed simplify the way he used navigation tools. Instead of carefully reading street names on maps and looking for street signs in the real world, he could glance at the 3D model on the screen and quickly find its counterpart in the real world.

Some design issues were also raised. First, the choice of landmark needed to be carefully considered. When the subject without GPS experience was asked about why she stopped and what she was thinking at the intersection, she indicated the disagreement of the choice on the landmark. The landmark in the design was the largest building at the intersection (only one landmark was provided in our test), and was in front of subjects when they approached the intersection. The subject, however, preferred another building, which was also in close to the intersection, but was seen earlier and had a more colorful appearance than the landmark. Such a disagreement illustrates a more fundamental problem: the gap between a user's need and a designer's assumptions about what users may need. Such gaps exist in many designs (Fonseca and Martin 2005; Goble and Wroe 2004). Narrowing down such gaps is not trivial. One way to address this problem in our system could be to provide more landmarks so that users can find what they want. Another approach is to store multiple landmarks, which have different properties, and ask users to specify what kind of landmarks they want to use. The system would then only deliver landmarks that fit a specific user preference.

Furthermore, the subject with GPS experience expressed some concerns with the simplicity of the surrounding environment at the decision-making point. He indicated that at the intersection, he tried to find street names in the immersive environment and then compared them with the street signs in the real world. Our implementation was quite simple, and only presented roads and landmarks. Street names were deliberately omitted in the test in order to examine how 3D models, rather than other information such as street names, may help navigation. The subject said that when an immersive environment was provided, a richer 3D scene was expected. The over-simplified 3D scenes may impede the recognition and use of the landmark. While a richer 3D scene is possible, adding more objects may affect system performances, given the still limited computation resources in Pocket PCs.

5.3 System performance

The biggest concern with system performances is graphics rendering capabilities. We used a very simple 3D scene in the study because it was found that rendering complicated scenes (e.g., terrain and realistic textures) in real-time is still a challenge in the platform we used. Delay is significant and can affect the use of the system. Thus, we had to use a simple 3D scene in the study. To deal with this issue, more work is needed to improve the efficiency of 3D programming with the consideration of hardware characteristics. However, this is beyond the scope of this research.

6 Discussion

The contributions of this research lie in the following aspects. First, this research proposed a way to integrate

3D virtual environments into services that become more and more important to people's daily lives. The multi-format, multi-scale, and semantic presentations of spatial information could help to construct a comprehensive understanding of the real world, and facilitate spatial actions. Second, we presented a method to develop a light-weight navigation system, which can be implemented on Pocket PCs and smart phones based on Windows Mobile 2005. The penetration of Pocket PCs and smart phones into people's lives provides an opportunity to greatly expand the use of 3D virtual environments in our e-society. Furthermore, our approach implies a new way to use 3D virtual environments in the real world by going beyond traditional applications, such as scientific visualization, simulations, and games. Combined with other techniques and information, 3D environments could help to address cognitive issues by leveraging the volumetric appearances of 3D objects, 3D animation, and immersive experiences.

Our design has some limitations. It only works in areas where distinct landmarks can be easily found. In places without landmarks, immersive virtual environments may not help, and M^2S maps may not offer more benefits than other conventional navigation systems. Also, an M^2S map system may require more spatial information than other traditional navigation tools. With a traditional tool, such as a map, users only need to match street names in the real world with those from the tool. Using M^2S maps, the appearances of 3D landmarks need to be accurate. Otherwise, unmatched appearances, even in color, may confuse users.

There are some unaddressed issues in this research. First of all, the time values used to mediate the transformation of different presentations of spatial information is still ad-hoc and based on practical experiences. Whether such an approach would be sufficient remains a question. More research is needed to develop a theoretical understanding of the division of navigation subtasks. Furthermore, the current system architecture relies on Pocket PCs to provide needed computation powers and storage space and does not take advantage of increasingly richer network resources. Some research projects have explored hybrid approaches by including high-speed servers to provide necessary information to mobile devices. Accessing rich information through networks may help to build a more comprehensive and user-friendly system.

7 Conclusion

This paper discusses the use of virtual reality technology to enhance real-world navigation with mobile devices. Recent advances in hand-held devices make it possible to bring 3D into people's highly mobile lives. In this paper, we explored a design, M^2S maps, to provide users with multi-format, multi-scale, and semantic spatial information in support of navigation. We also presented the implementation of an M^2S system and a field study on the use of this system.

Future research efforts can be extended in different ways. First, we would like to continue to improve the design based on feedback from the user study. Second, we would also like to conduct a formal usability study to compare our design with traditional maps and commercial GPS navigation systems. Results from the usability study could provide more insight into the use of 2D maps and 3D models.

References

Abowd G, Dey A, Brown P, et al (1999) Towards a better understanding of context and context-awareness. In: Proceedings of the 1st international symposium on handheld and ubiquitous computing (HUC'99), pp 304–307

Bagrow L (1985) History of cartography. Precedent Publishers, Chicago

Barkowsky T, Freksa C (1997) Cognitive requirements on making and interpreting maps. In: Hirtle S, Frank A (eds) Spatial information theory: a theoretical basis for GIS. Proceedings of COSIT 97. Springer, Berlin Heidelberg New York, pp 347–361

Baus J, Krüger A, Wahlster W (2002) A resource-adaptive mobile navigation system. In: Proceedings of the 7th annual international conference on intelligent user interfaces, pp 15–22

Bellard, Fabrice, TinyGL, Retrieved December 20, 2005, http://www.fabrice.bellard.free.fr/TinyGL

Benini L, Bonfigli M, Calori L, Farella E, Ricc B (2002) Palmtop computers for managing interaction with immersive virtual heritage. In: Proceedings of EUROMEDIA, pp 183–189

Bruntsch S, Rehrl K (2005) Vienna-SPIRIT—combining inter-modal traveller information with onboard/offboard navigation services, GNSS. The European navigation conference, Munich

Burigat S, Chittaro L (2005) Location-aware visualization of VRML models in GPS-based mobile guides. In: Proceedings of Web3D, pp 57–64

Chalmers D, Dulay N, Sloman M (2004) A framework for contextual mediation in mobile and ubiquitous computing applied to the context-aware adaptation of maps. Personal Ubiquitous Comput 8:1–18

Chen JL, Stanney KM (1999) A theoretical model of wayfinding in virtual environments: proposed strategies for navigational aiding. Presence Teleoperators Virtual Environ 8(6):671–685

Darken R, Cevik H (1999) Map usage in virtual environments: orientation issues. In: Proceedings of IEEE virtual reality, pp 133–140

Darken RP, Sibert J (1996) Navigating large virtual spaces. Int J Hum Comput Interact 8(1):49–71

Dearman D, Hawkey K, Inkpen K (2005) Effect of location-awareness on rendezvous behaviour. In: Extended abstract of ACM CHI, pp 1929–1932

Downs R, Stea D (1973) Image and environment: cognitive mapping and spatial behavior. Aldine, Chicago

Elias B, Sester M (2002). Landmarks for routing—automatic identification, extraction and visualization. ISPRS/ICA joint workshop on multi-scale representations of spatial data, Ottawa

Elvins T, Nadeau D, Kirsh D (1997) Worldlets-3D thumbnails for wayfinding in virtual environments. In: Proceedings of the ACM UIST, pp 21–30

Evans GW, Pezdek K (1980) Cognitive mapping: knowledge of real-world distance and location information. J Exp Psychol Hum Learn Mem 6(1):13–24

Finn MVN (2004) The handheld classroom : educational implications of mobile computing. Aust J Emerg Technol Soc 2(1):21–35

Flintham M, Benford S, Anastasi R, et al (2003) Where on-line meets on the streets: experiences with mobile mixed reality games. In: Proceedings of ACM CHI, pp 569–576

Fonseca F, Martin J (2005) Play as the way out of the newspeak-tower of babel dilemma in data modeling. In: McLean E, Monod E (eds) The 26th international conference on information systems, Las Vegas, pp 11–20

Goble C, Wroe C (2004) The Montagues and the Capulets. Comparat Funct Genomics 5(8):618–622

Golledge R (1999) Human wayfinding and cognitive maps. In: Golledge R (ed) Wayfinding behavior: cognitive maps and other spatial processes. Johns Hopkins University Press, Baltimore, pp 5–45

Goodman J, Gray P, Khammampad K, Brewster S (2004) Using landmarks to support older people in navigation. In: Proceedings of mobile HCI, pp 38–48

Goose S, Wanning H, Schneider G (2003) Mobile reality: a PDA-based multimodal framework synchronizing a hybrid tracking solution with 3D graphics and location-sensitive speech interaction. Ubicomp, pp 33–47

Hart R, Moore G (1973) The development of spatial cognition: a review. In: Stea B, Downs R (eds) Image and environment. University of Chicago Press, Chicago, pp 226–234

Headon R, Curwen R (2002) Movement awareness for ubiquitous game control. Personal Ubiquitous Comput 6:407–415

Hommel B, Knuf L (2003). Acquisition of cognitive aspect maps. In: Freksa C, Brauer W, Habel C, Wender KF (eds) Spatial cognition III: routes and navigation, human memory and learning, spatial representation and spatial learning. Springer, Berlin Heidelberg New York, pp 157–173

Hommel B, Gehrke J, Knuf L (2000) Hierarchical coding in the perception and memory of spatial layouts. Psychol Res 64:1–10

Jul S, Furnas GW (1997) Navigation in electronic worlds. SIGCHI Bull 29(4):44–49

Kray C, Elting C, Laakso K, Coors V (2003) Presenting route instructions on mobile devices. In: Proceedings of IUI, pp 117–124

Krüger A, Butz A, Müller C, et al (2004) The connected user interface: realizing a personal situated navigation service. In: Proceedings of IUI, pp 161–168

LaViola J, Feliz D, Keefe D, Zeleznick R (2001) Hands-free multi-scale navigation in virtual environments. In: Proceedings of the symposium on interactive 3D graphics, pp 9–15

Leigh J, Johnson A, DeFanti T (1996) CALVIN: an immersimedia design environment utilizing heterogeneous perspectives. In: Proceedings of IEEE international conference on multimedia computing and systems, pp 20–23

Levine D, Warach J, Farah M (1985) Two visual systems in mental imagery: dissociation of ‘‘what’’ and ‘‘where’’ in imagery disorders due to bilateral posterior cerebral lesions. Neurology 35:1010–1018

Loomis JM, Beall AC (1998) Visually controlled locomotion: Its dependence on optic flow, 3-D space perception, and cognition. Ecol Psychol 10:271–285

Lynch K (1960) Image of the city. MIT Press, Cambridge

Magerkurth C, Cheok A, Mandryk R, Nilsen T (2005) Pervasive games: bringing computer entertainment back to the real world. Comput Entertain (CIE) 3(3)

Malaka R, Zipf A (2000) DEEP MAP—challenging IT research in the framework of a tourist information system. In: Proceedings of ENTER, pp 15–27

Marsden G, Tip N (2005) Navigation control for mobile virtual environments. In: Proceedings of mobile HCI, pp 279–282

May A, Ross T, Bayer S, Tarkiainen M (2003) Pedestrian navigation aids: information requirements and design implications. Personal Ubiquitous Comput 7:331–338

McNamara TP, Hardy JK, Hirtle SC (1989) Subjective hierarchies in spatial memory. J Exp Psychol Learn Mem Cogn 15:211–227

Narzt W, Pomberger G, Ferscha A, et al. (2003) Pervasive information acquisition for mobile AR-navigation systems. Fifth IEEE workshop on mobile computing systems & applications

Narzt W, Pomberger G, Ferscha A, et al (2004) A new visualization concept for navigation systems. Eighth ERC-IM workshop: user interfaces for all (UI4All), pp 440–451

Newcombe N, Huttenlocher J (2000) Making space: the development of spatial representation and reasoning. MIT Press, Cambridge

Parent R (2001) Computer animation: algorithms and techniques. Morgan–Kaufmann, San Francisco

Piaget J, Inhelder B (1967) The child’s conception of space. Norton, New York

Pierce J, Pausch R (2004) Navigation with place representations and visible landmarks. In: Proceedings of VRST, pp 173–180

Puppo E, Scopigno R (1997) Simplification, LOD, multiresolution: principles and applications. In: Proceedings of EUROGRAPHICS 16(3)

Rakkolainen I, Vainio T (2003) A 3D city info for mobile users. Comput Graph 25(4):619–625

Raubal M, Winter S (2002). Enriching wayfinding instructions with local landmarks. In: Egenhofer M, Mark D (eds) Geographic Information Science. Lecture Notes in Computer Science. Springer, Berlin Heidelberg New York, pp 243–259

Reid J, Lipson M, Hyams J, Shaw K (2004) Fancy a schmink?: a novel networked game in a caf. Adv Comput Entertain Technol 18–23

Rodden T, Cheverst K, Davies K, Dix A (1998) Exploiting context in HCI design for mobile systems. Workshop on human computer interaction with mobile devices

Ruddle R, Payne S, Jones D (1997) Navigating buildings in ‘‘Desk-Top’’ virtual environments: experimental investigations using extended navigational experience. J Exp Psychol Appl 3(2):143–159

Schilit B, Adams N, Want R (1994) Context-aware computing applications. In: Proceedings of the workshop on mobile computing systems and applications, pp 85–90

Schmidt A, Aidoo K, Takaluoma A, Tuomela U, van Laerhoven K, van de Velde W (1999) Advanced interaction in context. In: Proceedings of the 1st international symposium on handheld and ubiquitous computing (HUC’99), pp 89–101

Sefelin R, Bechinie M, Müller R, et al. (2005) Landmarks: yes; but which? Five methods to select optimal landmarks for a

landmark- and speech-based guiding system. In: Proceeding of mobile HCI, pp 287–290
Slavík P, Berka R (1999) Navigation system for blind users in mobile computing environment. In: Interactive applications of mobile computing, pp 53–60
Smith I, Consolvo S, LaMarca A, et al (2005) Social disclosure of place: from location technology to communication practices. Pervasive 134–151
Stevens A, Coupe P (1978) Distortions in judged spatial relations. Cogn Psychol 10:422–437
Stoakley R, Conway M, Pausch R (1995) Virtual reality on a WIM: interactive worlds in miniature. In: Proceedings of the ACM CHI, pp 265–272
SundialSoft, PocketGL, Retrieved December 20, 2005, http://www.sundialsoft.freeserve.co.uk/pgl.htm
Suomela R, Roimela K, Lehikoinen J (2003) The evolution of perspective view in WalkMap. Personal Ubiquitous Comput 7(5):249–262
Thorndyke P, Hayes-Roth B (1982) Differences in spatial knowledge acquired from maps and navigation. Cogn Psychol 14:560–589
Timpf S, Kuhn W (2003) Granularity transformations in wayfinding. Spat Cogn 77–88
Tolman EC (1948) Cognitive maps in rats and men. Psychol Rev 55:189–208
Vinson NG (1999) Design guidelines for landmarks to support navigation in virtual environments. In: Proceedings of the ACM CHI, pp 278–285
Watsen K, Darken R, Capps M (1999) A handheld computer as an interaction device to a virtual environment. Third immersive projection technology workshop
Watt A, Policarpo F (2001). 3D games: animation and advanced real-time rendering. Addison-Wesley, New York
Wilson PN, Foreman N, Tlauka M (1997) Transfer of spatial information from a virtual to a real environment. Hum Factors 39(4):526–531
Witmer BG, Bailey JH, Knerr BW, Parsons KC (1996) Virtual spaces and real world places: transfer of route knowledge. Int J Hum Comput Stud 45:413–428
Yue W, Mu S, Wang G, Wang H (2005) TGH: a case study of designing natural interaction for mobile guide systems. In: Proceeding of mobile HCI, pp 199–206
Zhang X, Furnas GW (2005) mCVEs: using cross-scale collaboration to support user interaction with multiscale structures. Presence Teleoperators Virtual Environ 14(1):31–46