

5—4 Pedestrian Navigation Based on 3D Map and Mobile Interaction

Hideaki Maehara¹

Koji Wakimoto² Satoshi Tanaka³
Sumio Usui⁴

Kazuo Kamata⁵

Information Technology
R & D Center
Mitsubishi Electric Corporation,
Graduate School of Engineering
Utsunomiya University

Information Technology
R & D Center
Mitsubishi Electric Corporation

Department of Information Science
Utsunomiya University

Abstract

Recently pedestrian navigation has been studied by many researchers, which provides pedestrians with a similar function of car navigation. On the other hand, a 3-dimensional (3D) map, which is one of major themes in machine vision research, has been utilized as a simulation tool in city and landscape planning, and other engineering fields. The 3D maps will give more intuitive information compared to conventional 2-dimensional (2D) ones. In this paper we first propose pedestrian navigation based on 3D maps, and describe technologies required and its use situations. Secondly we compare effectiveness of 2D and 3D maps for navigation by object search experiments under conditions: 3D maps with and without texture; display sizes corresponding to a mobile phone and PDA (Personal Digital Assistant). The experimental results show that 3D maps with texture on mobile phone display size, and 3D maps without texture on PDA display size are superior to 2D maps in search time and error rate. From these results, we can say the effectiveness of 3D maps in pedestrian navigation.

1 Introduction

Recently pedestrian navigation, which provides pedestrians with a similar function of car navigation, has been studied actively [4, 6, 16]. Moreover, a market research shows the strong need for geographic information service with a mobile phone [14]. Considering the present market situation, pedestrian navigation using mobile phones is more desirable than PDAs. However, mobile phones can not use a large display because of a physical limitation. In development of pedestrian navigation system, it will be one of important issues how a device shows geographic information effectively. Several studies related to the issue are reported, such as utilization of digital magnetic compass, voice and text description [2, 13]. On the other hand, a 3-dimensional (3D) map, which is one of major themes in machine vision research [7, 8, 15, 17], has been utilized as car navigation and city information service [9]. The 3D maps use modeling of the 3D form such as buildings and roads in a real world and are generated by

computer graphic technology. Therefore it is a particular feature for the 3D maps that a viewpoint for display can be set arbitrarily, while the viewpoint in conventional 2-dimensional (2D) maps is fixed in overlooking position. In 3D map representation, a scene of whole city from sky (bird view) or a similar scene as people look directly on ground can be generated. About 3D maps, the data related to the model of buildings and roads including usual 3D map information is called “geometry”, and the picture data, which indicates side view of buildings, is called “texture”. There are two types in 3D maps; 3D maps without texture and with texture. The former consists of geometry only. The latter includes both geometry and texture. The former is used for check of view from specified positions or propagation simulation of electric waves, and the latter is used for scene simulation of city and so on.

In this paper, we first propose pedestrian navigation based on 3D map [12]. We suppose 3D maps will provide more efficient representation of geometry information than conventional 2D maps, because 3D maps can generate a similar scene which users can see in front of them. From this point, we describe a concept and components of pedestrian navigation system based on 3D map. We also draw a picture of use interaction in information representation and operations.

Secondly we compare representation effectiveness of 2D and 3D maps for geometry information. We make an experiment that measures search time for objects predefined in a real scene with use of 2D and 3D maps. In order to make clear the effect of both texture in 3D maps and display size for representation, we make experiments on condition of 3D maps with/without texture and two display sizes corresponding to PDAs (Personal Digital Assistants) and mobile phones. The experimental result shows that 3D maps with texture and with a display of mobile phones and 3D maps without texture and with a display of PDAs can represent geometry information more effectively than 2D maps. We furthermore consider the effectiveness of pedestrian navigation based on 3D map with the use of experimental results. Our conclusion, however, is not the same as the car navigation system studied in [10], due to the difference of user situations.

¹ Address: 5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan. E-mail: maeh@isl.melco.co.jp

² Address: 5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan. E-mail: waki@isl.melco.co.jp

³ Address: 5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan. E-mail: stanaka@isl.melco.co.jp

⁴ Address: 5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan. E-mail: usui@isl.melco.co.jp

⁵ Address: 7-1-2 Yoto, Utsunomiya, Tochigi 321-8585, Japan. E-mail: kamata@is.utsunomiya-u.ac.jp

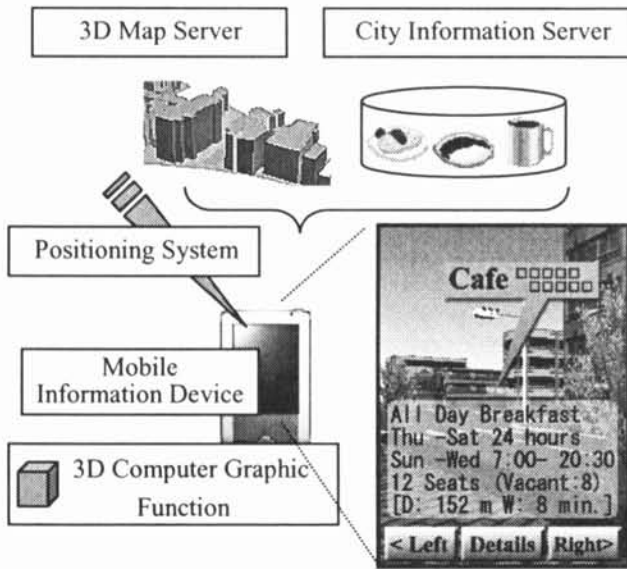


Fig. 1. Components of pedestrian navigation based on 3-dimensional map

2 Proposal of pedestrian navigation based on 3D map

2.1 Concept

In order to understand information shown in a map, users must find the correspondence between objects in a map and one in a real world. Since 2D maps represent a figure of a real world from upper viewpoint, there is obvious difference between the figure and the real world scene viewed from horizontal viewpoint. This is one of reasons why conventional 2D maps are not easy to understand. On the other hand, in 3D maps generated by computer graphics, it is a feature that a viewpoint can be set arbitrarily. Additionally, if the viewpoint is set to a user's present position, a generated image is similar to a scene that the user sees in front of him/her. Thus, since object matching in a 3D map and a real world is intuitive, we suppose it is easy to understand geometry information shown in 3D maps.

Furthermore, when city information is represented overlaying on 3D maps, users can understand it intuitively too. For example, if city information, such as shops and official services, is linked to a 3D map in advance, and is shown over the buildings in the 3D map, users can easily know where and what they can be served in their surrounding. In this point, we suppose that pedestrian navigation based on 3D map provides similar effect in information representation as AR (Augmented Reality) that electronically emphasizes and expands a real world [5].

We summarize the concept of the pedestrian navigation based on 3D map as following.

- A similar image as a scene that a user see in front of him/her can be presented by making a good correspondence between the position of the user and the viewpoint of the 3D map.
- Overlaying city information on a 3D map representation will have an effect of AR on a user.

2.2 Components

In this section, we describe components that are necessary to implement the pedestrian navigation proposed in the previous section (Figure 1).

(1) Mobile information apparatus

Users need to carry and operate a mobile information apparatus such as a PDA or a mobile phone. As we describe in the following sections, it is necessary to provide: (i) receiving function of 3D maps and city information; (ii) 3D computer graphic function; (iii) communication function with a positioning system.

(2) 3D map server

3D map data consists of two main components. One is "geometry" that has numerical data including 3D models of buildings and so on, and the other is "texture" that is picture data for mapping on surface of the geometry. The former is usually generated by 3D modeling software tools. There are some widely used data formats such as DXF, VRML and so on. The later plays a role of representing color and details of buildings. Typical formats used are GIF, JPEG and PNG. Mobile information apparatuses usually can not have secondary memory. The apparatuses, then, need get certain parts of a 3D map through communication network whenever they need them.

(3) City information server

In order to retrieve appropriate city information around a user two data for 3D map and city information should be linked well. Navigation for sightseeing around Kamakura city, for example, needs certain data: historical sites, places for shopping, public services, transportation, events and so on. City information, as well as 3D map, should be stored at a server, and must be sent to the apparatus through communication network.

(4) 3D computer graphic function

This function generates perspective projection at a mobile information apparatus using 3D map data and city information linked. For getting good image quality and usability, 3D computer graphic accelerator (hardware) should be equipped with the mobile information apparatus.

(5) Positioning system

In order to generate similar image that a user sees in front of him/her directly, present position of the user must be measured in around one meter accuracy. Recently GPS (Global Positioning System) is easily available, but there exists some restrictions such as difficulty in receiving GPS signal in high-density building areas. On the other hand, cell based positioning, that measures position by using information signals from plural cellular bases, is available indoors. This, however, dose not provide enough accuracy. With this background, other approaches have been studied, that use infrared beacons and a computer vision [1, 11]. For the proposed navigation, one of these robust and high accurate positioning system is necessary.

2.3 Examples of information representation and use scenarios

We show some examples of information representation and use scenarios for the proposed pedestrian navigation.

(1) Examples of information representation

1) 3D map representation (Figure 2)

A similar image that a user sees in front of him/her directly is shown at the first phase of the use. The user can adjust view range of the image by the operation of some buttons on an information apparatus.

2) Function selection (Figure 3)

The user can select functions by using hierarchical GUI (graphical user interface). Figure 3 shows an example of retrieving information about restaurants.

3) City information representation (Figure 4)

City information overlaid on a 3D map is represented: Location and menu for the restaurants retrieved are shown in Figure 4.

4) Route representation (Figure 5)

Overlaid figures on a 3D map show a route to the destination.

(2) Use scenarios

1) Information search

At lunchtime, a user turns on the navigation and executes "Restaurant Retrieval Function". At first an image similar to a scene that the user sees in front of him/her appears. Kinds of restaurants and their menus around the user are shown overlaying on the image.

2) Route representation

In a train, a user inputs an address of his/her destination. When the user gets out the train, a route from present location to the destination is shown automatically on a 3D map.

3) Walking simulation

One day before a business trip, a user inputs a destination and executes information retrieval function. Then an image of the nearest station from the destination is generated and overlays a route to the destination on a 3D map. By moving a viewpoint along the route, the user can execute walking simulation from the nearest station to the destination.

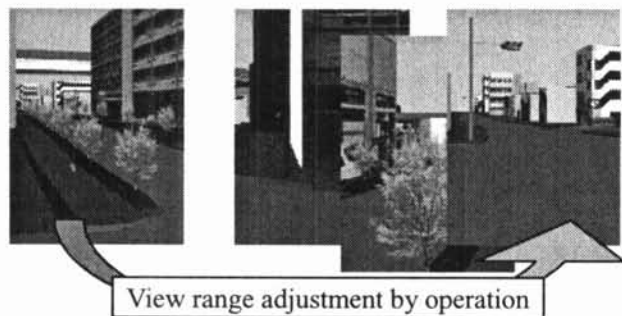


Fig. 2 Examples of 3-dimensional map display

3 Experiment

For the purpose that clarifies the effectiveness of proposed pedestrian navigation, we make a comparative experiment about representation efficiency of geographical information on 2D and 3D maps. In this experiment, we focus upon geographical information only, excluding city information, because city information provides another information to the user.

3.1 Specification of experiment

The use situation of pedestrian navigation is quite different from that of car navigation. One primary difference is that a user is able to use as much time as the user wants to recognize landmarks on the way to the destination. This is not the case of car navigation when a user drives. We set a task to search objects shown in a map from surrounding scene, that is different from the study of Karashima and Saitoh [10]. We measure time used to find out objects shown in maps from scenes, and compare search time in different conditions. From the results measured, we evaluate that a map which subjects take shorter time is more efficient to represent geographic information than another one which subjects take longer time.

In our experiment, we utilize a panorama viewer instead

of real scenes for the object search, that is a software tool using a 360-degree panorama photograph taken in a real world [3]. It generates scenes of arbitrary range at a position where the panorama photograph was taken. Subjects search the predetermined objects in a panorama photograph by operating the panorama viewer and point out it, with the use of the map that the object is marked. The relation between a real world search and the panorama viewer search is shown in the Appendix.

3.2 Data and conditions of experiment

Maps that are shown to subjects are conventional 2D maps and 3D maps. 3D maps are generated by perspective projection of 3D building models with view height corresponding to 1.6m in a real world. We used two kinds of 3D maps. One is a 3D map with texture that is generated with the use of color photographs of buildings. The other is a 3D map without texture. Accordingly there are totally three kinds of maps. Figure 6 shows the same object in different maps, where the object is hatched in each map. A 2D map covers 100m squares around the

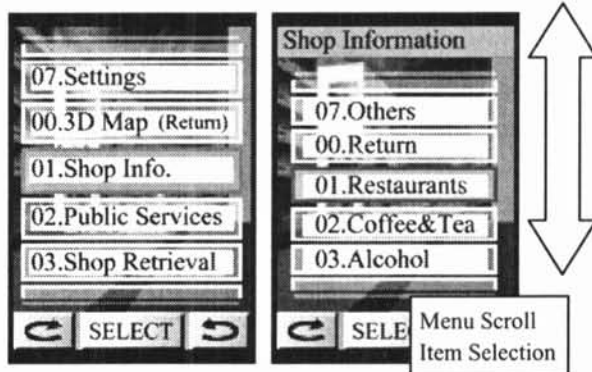


Fig. 3 Operation of hierarchical interface

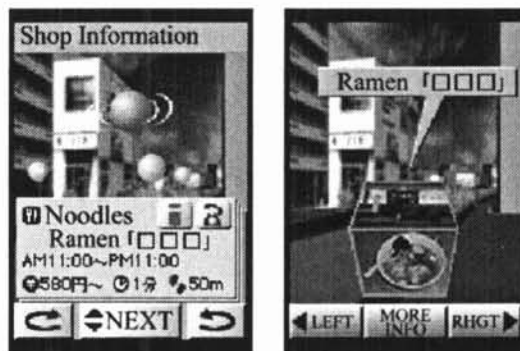


Fig. 4 Example of city information display

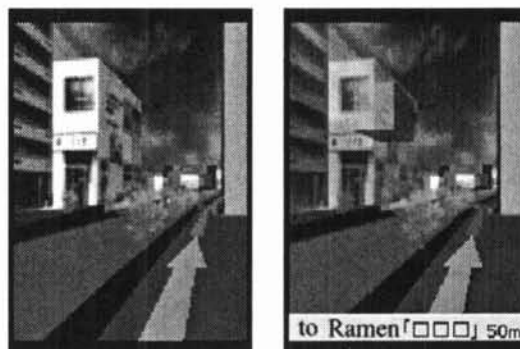


Fig. 5 Example of routes display for destination

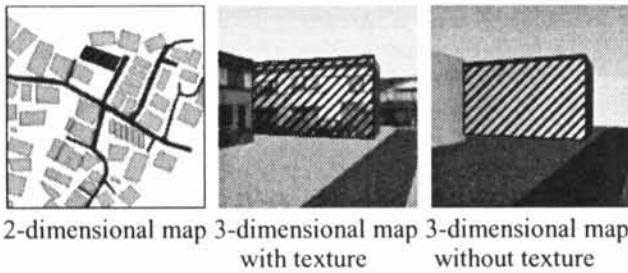


Fig. 6 Example of 2- and 3-dimensional maps

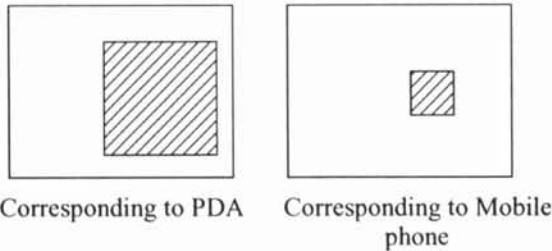


Fig. 7 Display image of maps

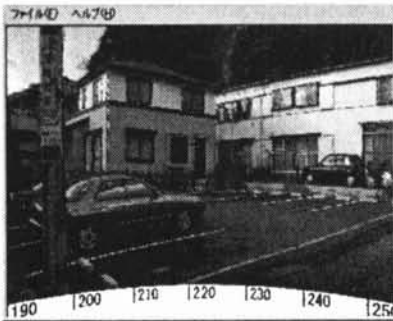


Fig. 8 Display example of panorama photograph

position where a panorama photograph is taken, and the object is always located upper half part of the map. The viewpoint for 3D maps corresponds to positions where a panorama photograph was taken, and the object is shown in horizontal center in the map. Moreover the height of the viewpoint is 1.6m from ground in a real world and view angle is 60 degrees, which are default values for VRML (Virtual Reality Modeling Language). We considered the effect of object and scene in map displayed. We selected three kinds of areas from central part of a city, residential area, shopping area and business area. Totally we prepared 6 maps, which consist of 2 maps from each area. Furthermore we determined one object (search target) in one map by the following conditions.

- Ground area of a building is more than 160m square.
- Distance from a position where a panorama photograph was taken is between 15m and 35m.
- In a 3D map, pixel number used by an object is more than 10% of total pixels.

Figure of each map is a square in shape on a computer screen. On the other hand, since we think representation size has effect on the target search efficiency. We prepared two representation sizes, as shown in Figure 7. One corresponds to a PDA display and the another one a mobile phone display. The former display size is 500 x 500 pixels and the latter 150 x 150 pixels. Accordingly there are

six map representation patterns.

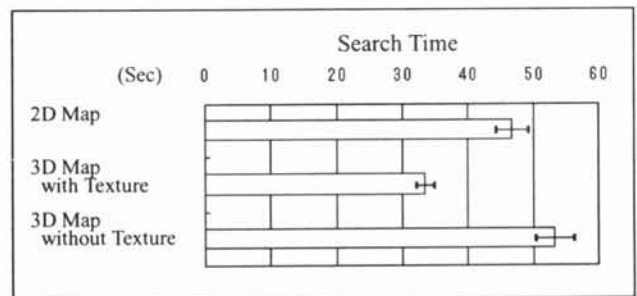
As other conditions, panorama photographs were given together with marks that indicate orientation in each 10 degree as shown in Figure 8. North is indicated by "0".

Subjects search objects and select one of the marks. Panorama photographs are 24-bit color images where the original photographs were taken at the height of 1.6m from ground. Initial view of the panorama viewer does not include a object. Subjects initially do not know which orientation they turn to on a map. The Panorama viewer represents panorama images on a 21 inch CRT display different from the display that shows maps.

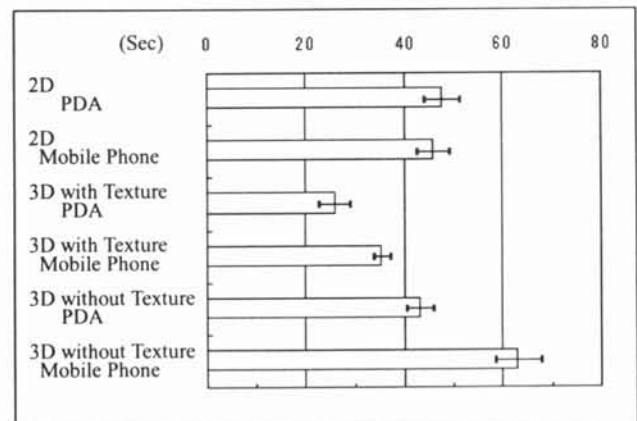
24 Subjects (males 17, females 7; university students 18, workers 6; normal vision) have joined the object search experiment. They all know nothing about the areas of maps. Their ages are between 20 and 40. Each subject is measured about 6 maps by 6 patterns (total measure times is $24 \times 6 = 144$). In the experiments we showed equal number of patterns to every subject. This means times of measuring about one map by same pattern is $4 (6 \times 6 \times 4 = 144)$.

3.3 Results

Figure 9 shows averages and standard errors of search time about 2D maps and 3D maps with/without texture. 3D maps with texture performed the shortest average search time, and 3D maps without texture performed the longest average search time as shown in Figure 9(a). Figure 9(b) shows average search time for different display sizes and maps. From these results, we can say PDA display size gave shorter search time compared to mobile display size significantly ($p < 0.05$). Figure 10 shows rates of cases where subjects pointed out an incorrect object ("errors") and when subjects could not point out any object

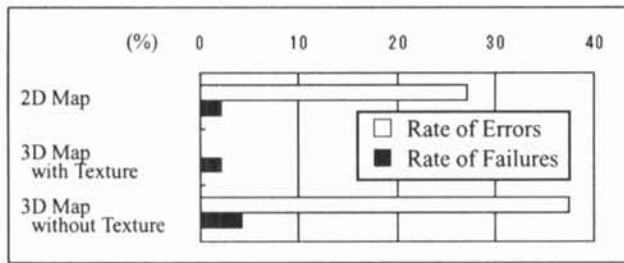


(a) Maps kinds

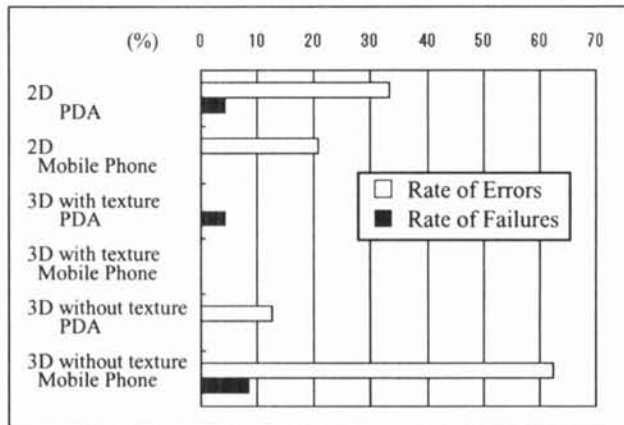


(b) Combination of map kinds and display sizes

Fig. 9 Comparison of search times



(a) Map kinds



(b) Combination of map kinds and display sizes

Fig. 10 Rates of error and failure in object search

(“failures”). Least errors and failures occurred in the case where 3D maps with texture were used and most ones occurred for 3D maps without texture (Figure 10(a)). Display size does not affect much on failures. The most failures occurred in the case of 3D maps without texture and with display size corresponding to mobile phones. There is no failure in the case of 2D maps with display size corresponding to mobile phones and 3D maps with texture and with display size corresponding to mobile phones. 3D maps without texture and with display size corresponding to PDAs give fewer errors than each of 2D maps with two display sizes (Figure 10(b)).

3.4 Discussion

Considering search time measured and number of errors and failures, we can say that 3D maps with texture are the most efficient one to represent geographic information. In addition to this, we can say that small display size affects much the efficiency. But 3D maps with texture on even display corresponding to mobile phones keep better than 2D maps. Although 3D maps without texture is inferior to 2D maps as shown in Figure 10(a), 3D maps without texture and with display size corresponding to PDAs give fewer errors and failures less than 2D maps. This means that 3D maps is superior to 2D maps in two cases; (i) with texture, (ii) without texture and with display size corresponding to PDAs. Because fundamental function of pedestrian navigation, as stated previously, is guidance of pedestrians, we can conclude that we can realize more effective pedestrian navigation by using 3D maps, which can represent geographic information to users in shorter time and correctly than using 2D maps.

4 Conclusions

In this paper, we proposed pedestrian navigation, which

represented geographic information to pedestrians, with the use of 3D maps instead of conventional 2D maps. We first described the concept and necessary components of the system and we also showed use scenarios including information representation. Moreover we have done an experiment in which subjects searched objects shown in maps from panorama photographs. As a result, we showed that 3D maps with texture on a display corresponding to mobile phones and 3D maps without texture on a display corresponding to PDAs were more efficient in representing geographic information than 2D maps. We furthermore concluded the effectiveness of pedestrian navigation based on 3D map. Selection of with/without texture and display size is an important specification that will be decided in a practical system design. The result, that shows even 3D maps without texture is more efficient than 2D maps in case of display size corresponding to PDAs, is one of important findings in system design considering system response time and communication cost. Prototyping of a mobile system is one of our future issues to develop real service.

Acknowledgments

In this study, we had useful advises from Professor Dr. Hirotaka Koike of Utsunomiya University and members of mobile computing group in Strategic Software Research Forum under auspices of International Information Science Foundation. We also thank students of Engineering Department in Utsunomiya University and researchers of Information technology R&D Center in Mitsubishi Electric Corporation for cooperation of experiments.

Appendix

Comparison between object search using a panorama viewer and one in a real world

We also have done an comparative experiment of measuring time in case of search using a panorama viewer and that of search in a real world standing at a position where the panorama photograph was taken. As shown on Table 1, setting representation of 6 maps with 6 different patterns (3 kinds of maps and 2 display sizes), and every 5 of 10 subjects (males 8, females 2; students 2, workers 8; normal vision) did the object search in each case of using a panorama viewer and a real world. From the results, it is shown that there is from 3 up to 6 times difference between their averages. We suppose it is because the view angle in the panorama viewer is smaller and operation load for pan is larger than human vision. However, as shown in Figure 11, correlation coefficient between two averages showed over 0.95. Therefore, even if we do the experiment using the panorama viewer, it is possible to evaluate their effectiveness quantitatively.

Table 1 Display patterns of maps at the experiment

Map No.	Kind of Map	Size
1	2D Map	PDA
2	3D Map with Texture	Mobile Phone
3	3D Map without Texture	Mobile Phone
4	3D Map without Texture	PDA
5	3D Map with Texture	PDA
6	2D Map	Mobile Phone

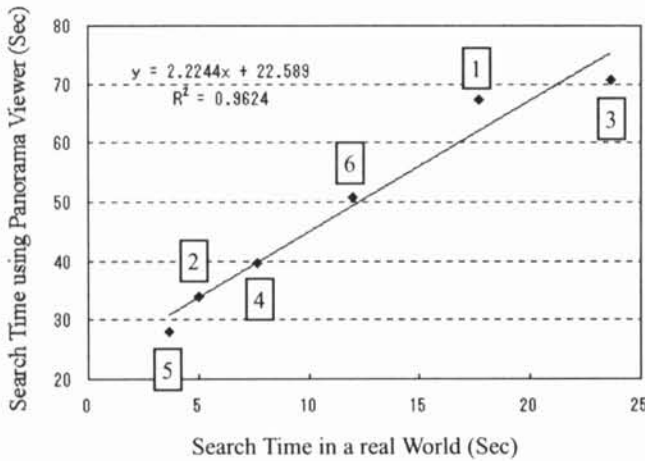


Fig. 11 Correlation between values using panorama viewer and values measured in real world

References

- [1] H. Aoki, B. Schiele and Alex Pentland: "Realtime personal positioning system for wearable computers", In The third Symposium on Wearable Computers, Los Alamitos, California, pp37-43, 1999.
- [2] A. Butz, J. Baus, A. Krüger, and M. Lohse: "A Hybrid Indoor Navigation System", In IUI2001: International Conference on Intelligent User Interfaces, New York, pp. 25-33, 2001.
- [3] S. E. Chen: "QuickTime VR - An image-based approach to virtual environment navigation", In Proceedings of SIGGRAPH 95, Los Angeles, CA, USA, pp.29-38, 1995.
- [4] B. Corona and S. Winter: "Datasets for Pedestrian Navigation Services", The AGIT Symposium, Salzburg, Austria, pp. 84-89, 2001.
- [5] S. Feiner, B. MacIntyre, T. Hollerer and A. Webster, "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment", The First International Symposium on Wearable Computer, Boston, MA, pp. 74-81, 1997.
- [6] A. Helal, S. Moore and B. Ramachandran, "Drishti: An Integrated Navigation System for Visually Impaired and Disabled", The 5th International Symposium on Wearable Computer, Zurich, Switzerland, 2001.
- [7] D. Huber, O. Carmichael, and M. Hebert: "3D map reconstruction from range data", Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '00), Vol. 1, pp. 891-897, 2000.
- [8] J. Iida and Kenji Yajima: "Generating 3D Urban Model From Airborne Laser Data", 3rd International Image Sensing Seminar on New Development in Digital Photogrammetry, Gifu, Japan, pp. 24-27, 2001.
- [9] T. Ishida: "Digital City Kyoto: Social Information Infrastructure for Everyday Life", Communications of the ACM (CACM), Vol. 45, No. 7, pp. 76-81, 2002.
- [10] M. Karashima and M. Saitoh: "A Study on the Comparison of the Usefulness between a 2 Dimensional Map and a 3 Dimensional Map", IPSJ SIGNotes Human Interface Abstract, No.082 - 005, 1998. (in Japanese)
- [11] Q. Ladetto, V. Gabaglio, B. Merminod: "Two Different Approaches for Augmented GPS Pedestrian Navigation", Int. Symposium on Location Based Services for Cellular Users (Locellus), 2001.
- [12] H. Machara and K. Wakimoto: "Mobile 3D Map: Proposal of 3D CG Navigation for Pedestrians", ITE Technical Report, 25(85), pp. 53-58, 2001. (in Japanese)
- [13] H. Maula: "Mobile GIS and Multimedia as Fieldwork Support in Helsinki Water", The Nordic GIS Conference, 2000.
- [14] Research and Development, Inc.: "About Next Mobile Phone", eM Report, 2001, <http://www2.rad.co.jp/eM/report/018ktai-core2001/index.htm> (in Japanese)
- [15] I. Suveg, G. Vosselman: "3D reconstruction of Building Models", Int. Archives of Photogrammetry and Remote Sensing, vol XXXIII, part B2, pp. 538-545, 2000.
- [16] K. Yachi, H. Ohkubo and K. Takeuchi: "Development of Pedestrian Information and Communication Systems (PICS)", 6th Annual World Congress on ITS, 1999.
- [17] H. Zhao, R. Shibasaki: "Reconstruction of Textured Urban 3D Model by Ground-Based Laser Range and CCD Images", IEICE Trans. Inf.&Syst., vol.E83-D, No.7, pp. 1429-1440, 2000.