Design and Development of an Indoor Navigation and Object Identification System for the Blind

Andreas Hub Joachim Diepstraten Thomas Ertl

Visualization and Interactive Systems Institute
University of Stuttgart, Universitätsstraße 38
70569 Stuttgart, Germany
{hub,diepstraten,ertl}@informatik.uni-stuttgart.de

ABSTRACT
In this paper we present a new system that assists blind users in orienting themselves in indoor environments. We developed a sensor module that can be handled like a flashlight by a blind user and can be used for searching tasks within the three-dimensional environment. By pressing keys, inquiries concerning object characteristics, position, orientation and navigation can be sent to a connected portable computer, or to a federation of data servers providing models of the environment. Finally these inquiries are acoustically answered over a text-to-speech engine.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: *User-centered design*, *Prototyping;* K.4.2 [**Social Issues**]: *Assistive technologies for persons with disabilities*

General Terms
Measurement, Design, Experimentation, Human Factors

Keywords
Indoor navigation, blind users, impaired vision, mobile computing

1. INTRODUCTION
When we asked a blind colleague at our institute about the most pressing navigation-related problems of the blind, he answered that the main problems are determining one's own position, determining head direction or movement direction, and the lack of information about important objects in the near and distant environment. In this context, any information about object features can be important. For indoor navigation tasks in particular, the unknown structure of the building is a huge problem. For example, after successfully reaching the main entrance of an unknown building – which in itself can be quite difficult – important information such as the number of floors in the building or the purpose of specific rooms or even of the building itself is usually not known.

More detailed questioning taught us that the blind especially have problems with stairs, all kinds of steps, elevators, revolving doors, and doors with an automatic opener, since they all may cause serious injuries. Furthermore, blind people prefer not to walk in the center of a room because there is normally no orientation line for their cane and because they try to avoid crowds of people. When walking in the center of a room or in a crowd, the orientation sense of a blind person is easily misled. Of great interest is the size of rooms and corridors, which is sometimes estimated by the blind with remarkable precision using the reflected sound of spoken words or whistled tunes. When a blind person has the possibility to explore a building for an extended time period, it is astonishing to hear their precise description of even large buildings. Most of the distances between important indoor landmarks are known as numbers of steps. Such landmarks can be nearly any type of object, ranging from doors and hallway intersections to photocopiers that are in the way or even really small things like light switches or barely noticeable thresholds.

To summarize, the blind want to know where they are and what kind of objects are in their environment, especially in front of them. Furthermore, any object feature can be of importance, depending on the current situation or intention. If one's own location is known, the next step is to plan how to get from there to other places. On the way to these destinations, the blind want to be informed particularly about dangers, critical situations and other persons.

At the same time, the development of portable computers has finally reached a level where available memory and processor speed are sufficiently powerful to resolve most of these challenging issues. We would therefore like to propose a new design concept to develop a solution for assisting blind users in unknown indoor environments. Our design makes use of the enhanced processing power of mobile devices, which allows real-time processing of 3D models of the environment. Additionally, input information from local sensors is used for object identification and for orienting users in the real world.

The paper is organized as follows. The subsequent section focuses on related work. We then describe our design and developments and summarize the results achieved up to now. Based on these current results, an outlook on future work is provided in the

ASSETS'04, October 18–20, 2004, Atlanta, Georgia, USA.

following section. The paper closes with an overall discussion and a final conclusion.

2. RELATED WORK

In general, navigation and orientation support systems for the blind can be divided into two main categories: outdoor and indoor systems. Several outdoor navigation systems have been proposed [13, 17, 22] or are commercially available [23]. Although they share common characteristics with indoor navigation systems, we will not go into further detail regarding these systems, as they often rely on technologies that are not available in indoor environments - for example by using GPS for tracking. Indoor navigation systems, on the other hand, are more challenging with respect to orienting a user, because a generic solution for this problem does not yet exist.

The increased popularity of wearable and small mobile computer devices has been accompanied by the development of solutions for aiding blind people. Sonnenblick [19] evaluated and implemented a large-scale indoor navigation system for blind individuals at the Jerusalem Center for Multi-Handicapped Blind Children. It relies on specially installed infrared beacons for orienting blind users and on a custom-built end-user device. In contrast, the “Chatty Environment” by Coroama [2] is focusing on technology that does not require special hardware, but rather equipment that is already generally available. This system uses standard 802.11 WiFi for positioning and PDAs as end-user devices, which increases the chances for a widespread use. However, the system does not really address common problems with WiFi positioning and its rather coarse precision. This is one of the main points in which our solution differs from Coroama, as we do not simply rely on large area sensors but use local object identification to get a higher accuracy for the current location of the blind user. Although other indoor location technologies exist [1, 10, 11, 15], we decided to use a WLAN based indoor location because it is the best and most cost effective solution available today. Furthermore, the latest research in this area shows promising improvements considering the accuracy of the system [5, 24].

3. DEVELOPMENT OF AN ORIENTATION ASSISTANT FOR THE BLIND

3.1 Concept

The basic aim of our research is to allow object identification for the blind and to improve their indoor navigation abilities using local sensor information in combination with 3D models of the environment. The structure of our concept is shown in an overview in Figure 1. The components of the architecture will be described in detail in the following subsections.

3.2 Color for object detection

Color is an important object feature for the blind, even tough this may not be easily understandable for people with normal vision. But when considering clothes, food, traffic lights or weather conditions it becomes obvious why the blind frequently talk about color - even those who have never seen any color. A lot of object features, such as the size of an object or its surface structure, are accessible to different senses. The problem with color for the blind is that this object feature is only available to the sense of vision.

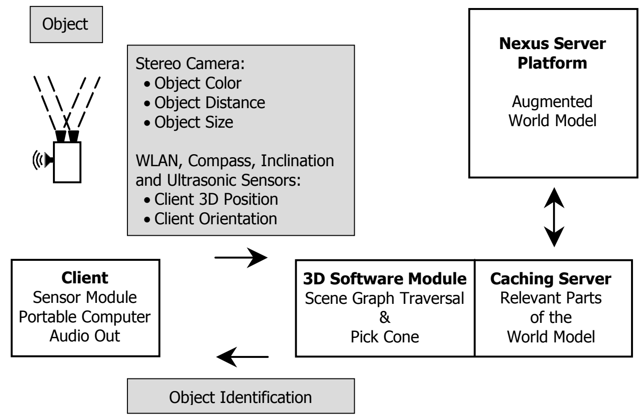

Figure 1: Matching between local sensor information and 3D models of the environment leads to object identification. Our architecture is realized as a location based client-server solution. The arrows describe data transfers between components.

In contrast, the size or the weight of objects are also accessible to our tactile sense. But there is no other means for the perception of color except seeing. Perhaps this is the reason why color is so important and also fascinating for the blind.

A lot of objects and materials have a characteristic color or their color is within a typical range, like for example the colors of skin, metals and fruit. This suggests the use of color for the detection of objects.

In 2001, the development of our first functional prototype of an orientation aid for the blind based on color detection was completed [8]. The basic idea for this device was to use color for object recognition for blind users. This device combined a laptop and an auto focus video camera. The camera allowed the (blind) user to search for objects in the user’s environment. By pressing a key on the laptop an image was taken by the camera. Hue information of the central part and a region growing algorithm were used for the segmentation of the object in the middle of the image.

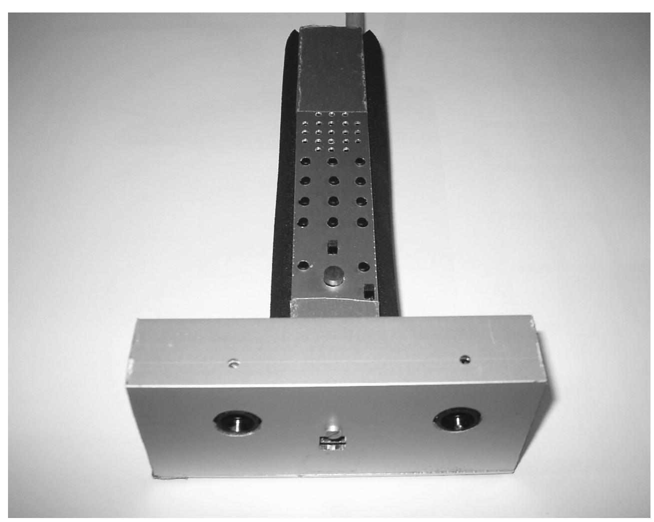

Figure 2: First prototype of a sensor module with two cameras, keyboard and loudspeaker.

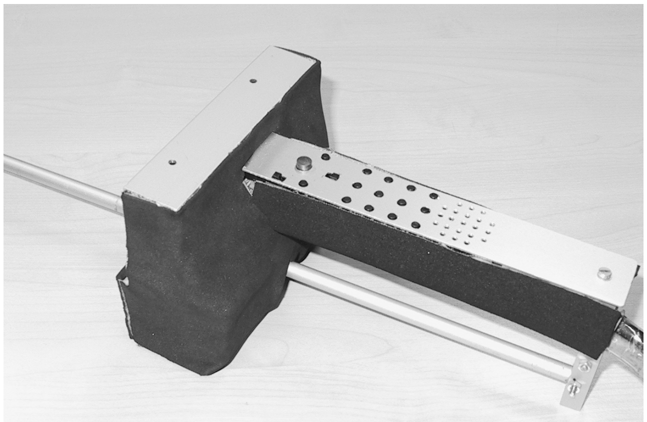

Figure 3: Extended sensor module with digital compass, inclination sensor and detachable blind man's stick.

For the detection of object colors through technical means an essential mechanism of human perception has to be considered. Color appearance of an object is dependent on the object's surroundings. This has been known by artists and painters for centuries [12]. Today, this phenomenon of human color perception is called color or chromatic induction [21] or described as simultaneous color contrast in older literature [6, 7]. On the basis of psychophysical experiments on test persons with normal color vision, a model of human color vision and a corresponding algorithm for color naming were developed. Both the model and the algorithm take color induction into account [9].

This algorithm was used to determine the color of a segmented object in the middle of the image. The result of the color measurement was transmitted to the blind user over a text-to-speech engine and a loudspeaker or earphone. Usability tests with two test persons, one totally blind and the other having only minimal rest vision, showed that color information is of great interest for blind users, in particular because the device allowed for the first time the determination of color names without direct contact to the object surface. For example, it was possible to determine the color of the sky or of things the blind persons could not or did not want to touch. The tests showed that such a color device can be quite helpful for sorting clothes or "looking" at the sky. However, using color information alone allows the identification of objects only in very few cases.

3.3 Object distance and object size

In the next step the lens position of the auto focus camera was used to measure the distance of the object in the center of the taken image and the object's color at the same time. With the information of the object distance and the pixel range of the segmented region it is possible to approximately calculate the real width and the height corresponding to the segmented region. Doing this, it was now possible to tell the blind user not only the color of objects but also the width, height and distance of the focused object over the text-to-speech engine.

This combination of object features made it easier for the blind user to identify a known object. However, the recognition of an arbitrary object remained still very difficult. Furthermore, there were other problems with the distance measurement based on auto focus. To determine the object distance via auto focus, it is necessary to use a lens with a long focal length. With such a lens,

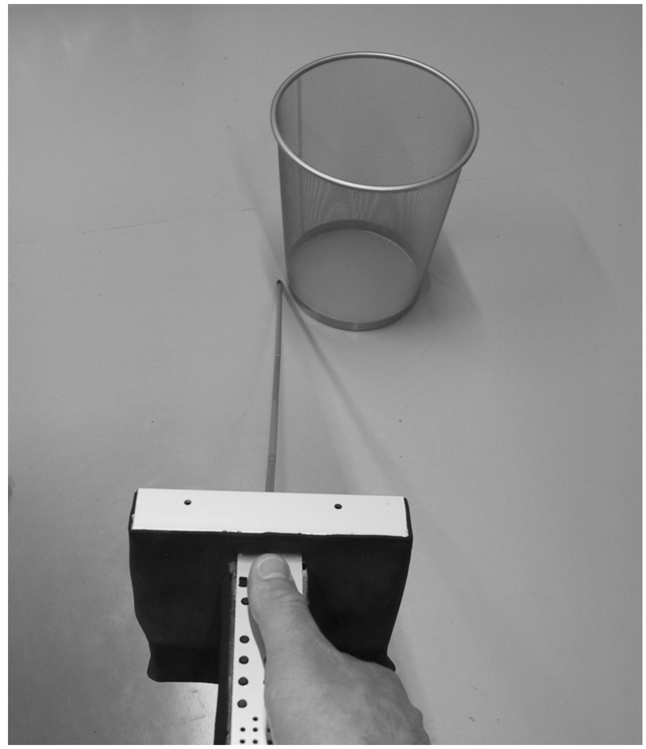

Figure 4: Only one hand is needed for the usage of the orientation assistant in combination with the blind man's stick.

only a small part of most nearby objects can be covered. Unfortunately, this is contrary to the blind users' desire, as blind users are mostly interested in objects lying in a range from 0.1m to 3m away. Other problems of auto focus systems are the time delay and the inexact distance measurement in complex scenes. Therefore, a new solution had to be found.

We then developed a sensor module consisting of a stereo camera (Figure 2). This new sensor module is equipped with a small keyboard so that the user can handle it like a combination of a flashlight and a cellular phone. The keyboard is integrated in a hand grip. The sensor module can be used for searching tasks in the 3D environment with the direction of the hand grip defining the search direction. The object in the imaginary extension of this direction is segmented in both images of the two cameras. The object distance is calculated from the pixel shift. After the distance is known, the width and height can also be calculated.

3.4 Positioning and direction sensors

To be able to use models of the environment in combination with local sensor information, we had to equip the sensor module with several positioning and direction sensors. Therefore, a WLAN card was integrated in the portable computer and the connected sensor module described above was extended by a digital compass and a 3D inclination sensor. For the measurement of the sensor module's distance above the floor we use an ultrasonic sensor. For safety reasons, the sensor module was equipped with a light and detachable blind man's stick, so that only one hand is necessary for using both the mobile orientation assistant and the stick. This cane is equipped with a hinge and can click into place below the hand grip. The blind user can change between two options. In the first option, the hand grip and the cane have the same direction (Figure 3).

The second option allows the user to click the cane downwards. In this configuration the user is always in contact with the floor through the cane, and the cameras “look” straight ahead (Figure 4). This development should ease the adaptation phase for blind people to this new technology, because they will not be missing the feeling of safety that can only be obtained from a blind man's stick.

3.5 Operating the orientation assistant

By pressing designated keys, different sequence and loudness options can be chosen and inquiries concerning the object features can be sent to the connected portable computer. After successful evaluation these inquiries are acoustically answered over a text-to-speech engine and a loudspeaker, which is also integrated into the hand grip.

3.6 Design of an Indoor Navigation System

As a first step, a virtual 3D environment model for a test scenario had to be created. We chose our department building as the testing ground. With the help of a professional 3D modeling software package, a model of the 1st floor of the building was created with an accuracy of about one millimeter. In addition, the interior of several rooms, including furniture, was added to the 3D environment (an example of such an interior room can be seen in Figure 5).

To process the 3D environment model data in real time on a portable computer, an appropriate 3D software module had to be developed. We use an ordinary Notebook for the software development (Fujitsu Siemens Amilo D 8820). However, its weight and its energy consumption do not make it an agreeable solution to be carried by blind users. Therefore, the software was transferred to a sub notebook with a weight of under 900 g (JVC MP-XP7250). We were able to demonstrate that the software module is running on this system close to real time demands.

The 3D software is based on a high-level scene graph API. A scene graph is an acyclic graph data structure storing a 3D scene in a hierarchical order. It organizes and controls the rendering of its constituent objects. A scene graph mostly consists of several types of nodes that either store grouping, transformation, object appearance or object shape information. When using a scene graph one gets the benefit of having a hierarchy that allows storing object affiliations. This becomes particular advantageous when trying to figure out what object the user is currently pointing at.

There are several commercial and non-commercial scene graph APIs available [3, 16, 18, 20]. We decided to use Open-SceneGraph [16]. The reason for this choice was that OpenSceneGraph is still under active development, as opposed to many other scene graphs. It offers a rich functionality and data import from several established 3D model formats. It also provides satisfactory performance even on weaker machines and it is available on all major platforms.

After loading the necessary 3D environment data from hard disk, a virtual camera is placed into the scene. The x/y position of the virtual camera corresponds to the estimated position of the user derived from the WiFi location system. The z (height) of the camera corresponds to the measured height of the ultrasonic sensor. At the moment, the compass and the 3D inclination sensor directly influence the virtual camera’s viewing direction.

Figure 5: Rendering of an example room modeled as 3D environment used in our system.

When a corresponding button on the device is pressed, the color of the front-most object, its estimated size, and distance to the current user position is read from the local sensors. At the same time a picking cone is emitted from the virtual camera in the scene graph into the current viewing direction. Then, all the objects in the scene graph that are hit by the pick cone are compared to the color and size information gathered from the sensors and the best match is returned. Each object in the scene graph can contain additional information, like warnings or detailed object descriptions, that can then be transmitted through the text-to-speech-engine to the blind user. The hierarchy of the hit object can be used to retrieve additional information. For example, if a door handle is hit, this relationship information is used not only to return that a door handle is hit, but also that this door handle belongs to a door leading into a neighboring room.

4. RESULTS

We have developed the hardware for a new type of orientation assistant for the blind. Its sensor module allows blind users to aim at locations in the real world where static objects or parts of the building are expected. The blind user can communicate with the device by means of a keyboard to send inquires to the connected mobile computer. These inquiries are answered via the text-to-speech-engine either through a loudspeaker or an earphone.

The orientation assistant is robust and can be handled easily even by elderly persons as the keyboard is nearly identical to a keyboard of a cellular phone. The integrated blind man’s stick ensures the safety and the feeling of safety of a normal white cane.

The software components implemented in the connected portable computer allow determination of the color, the distance and the size of objects in the real world environment of the blind user with close to real time performance. The sensor information can be used to identify objects by matching this information with data stored in the 3D environment models. The processing of these models is at present also possible with close to interactive speed.

5. FUTURE WORK

In the next step we will implement the matching algorithm for comparing the local sensor data with the data of the environment model. This algorithm should allow the identification of all static objects and parts of the building stored in the 3D model.

Furthermore, we will expand this algorithm in a way that allows the recognition of semi-static objects like chairs, computers or other things that can easily be moved but normally remain in the same room. We also started working on a route planning algorithm considering the special needs of the blind described above. As can be seen in the figures showing the prototypes, there is still a lot of room for improvements, especially concerning hardware optimizations such as improved ergonomics, reduced case size, chamfered edges, and reduced weight.

In the near future we plan to embed the device into the Nexus framework [4, 14]. The Nexus framework is a general location-aware application platform that is currently under development at the University of Stuttgart. The heart of the Nexus platform is the "Augmented World Model" that allows a general description of arbitrary physical real-world and virtual objects.

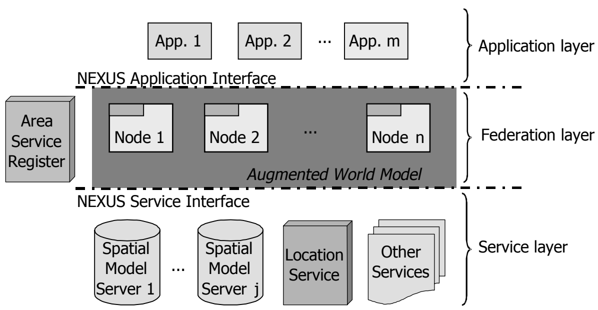

Figure 6: Overview of the Nexus architecture.

The information stored about these objects can be derived from anywhere and anytime through location dependent queries by the end user. Figure 6 shows an overview of the Nexus architecture where Nexus nodes mediate between applications and Nexus services. They are responsible for distributing information to the different services and for answering queries, thus providing an integrated and consistent view for applications. The Spatial Model Servers are currently used to store static objects, which can be used to assign additional data to locations. The location service, on the other hand, is responsible for the storing of mobile objects such as other users.

Due to the large amount of necessary data and the desire to have an adequate response time on the end user's client, we decided to introduce a sort of caching server between the blind user device and the Nexus server structure (Figure 1). This caching server tries to retrieve all the required data of the assumed user's location in the current building area. It then processes this data into chunks that can be handled by the device without introducing too much strain on the portable machine itself. This should ensure that it does not lose its real time capability.

6. DISCUSSION

Blind persons are able to move and navigate easily within known indoor environments with a speed that is sometimes very astonishing for people with normal vision. In contrast, orientation and navigation in unknown environments can be very difficult and also very dangerous for the blind. Based on the recent development of mobile computing devices it appears possible to develop systems that can assist blind persons in unknown and complex indoor environments.

For indoor scenarios, most landmarks and objects important to the blind can be stored in 3D models of selected buildings, for example government and public buildings or training centers for the blind. For outdoor scenarios this still remains an overwhelming task, given the available computing resources and the complexity of the environment.

In our opinion it is currently impossible to realize object identification of arbitrary objects using systems that are only based on image segmentation and image interpretation. We are therefore convinced that indoor object identification, orientation and navigation tasks for the blind can be optimized significantly by combining local multi-sensor information with global world models.

Since we work in close cooperation with blind persons, mobility teachers, and associations for the blind, we know the desire of the blind for an assistant that, in the next version, can be handled discreetly and easily.

By embedding the device into the Nexus architecture, several new possibilities besides navigating, orienting, and local object identification become available. For example, the system can indicate important orientation lines and distances between landmarks. Furthermore, it can provide warning messages to the user or even help to locate other users, colleagues, and friends.

When looking at the promising options we also have to consider some remarkable challenges. The energy consumption of the portable computer limits the usage to a maximum of five hours. Another issue are errors in measurement due to electromagnetic fields or strong temperature drifts. A general problem is to find a good compromise between the size of the device and its robustness. Also, there still is a general uncertainty as to how good a match between the virtual view in the 3D environment and the real view can be achieved. A similar problem arises when GPS is used in outdoor tracking.

The additional options provided by the Nexus platform are not yet exactly determined. It may be a problem to extract the really relevant information a blind user wants to receive, because Nexus it is not a service specially built for the blind. At the same time, the amount of information should not exceed a certain level so as not to annoy the user. On the other hand, blind persons have different requirements concerning safety and security issues. A system that will be accepted by blind users has to be extremely stable and reliable.

7. CONCLUSION

Most of the blind have the desire to stay independent as long as possible. This is also true for persons with impaired vision, a growing group as more and more people grow old. However, elderly persons are afraid to rely on new and unknown technology. Our concept of combining traditional aids with innovative technologies may also be a solution for persons suffering from other sensory deficits. Our solution will not only solve navigation and orientation tasks but can even provide additional value by accessing resources from the Nexus platform.

As a closing remark we must admit that there are still a lot of challenges that have to be resolved before a global deployment can be considered.

8. ACKNOWLEDGMENTS

This project is founded by the Deutsche Forschungsgemeinschaft within the Center of Excellence 627 "Spatial World Models for Mobile Context-Aware Applications". We also would like to thank Alfred Werner for his helpful input and patience.

9. REFERENCES

[1] Axess Inc, Active Tag™, http://www.axsi.com/products/rfidoverview.shtml.

[2] Coroama, V. The Chatty Environment - A World Explorer for the Visually Impaired. *Adjunct Proceedings of Ubicomp* (2003).

[3] Criterion Software Inc., Renderware. http://www.renderware.com.

[4] Dürr, F., Rothermel, K. On a Location Model for Fine-Grained Geocast. In: Dey, Anind K. (ed.); Schmidt, Albrecht (ed.); McCarthy, Joseph F. (ed.): *Proceedings of the 5th International Conference on Ubiquitous Computing: UbiComp '03* (Seattle, Oct. 12-15, 2003), Springer, 18-35.

[5] Ekahau, Inc. Ekahau Positioning Engine. http://www.ekahau.com/products/positioningengine/.

[6] Helmholtz, H. v. Handbuch der physiologischen Optik, Issue 2, 3rd Edition, Voss Hamburg Leipzig, 1911.

[7] Hering, E. Grundzüge der Lehre vom Lichtsinn, Handbuch der gesamten Augenheilkunde, Issue 3, 2nd Edition, Julius Springer Verlag, 1925.

[8] Hub, A. Von der Objektunabhängigkeit der Farbe im Landschen Versuch über ein Messverfahren der Farbbenennung zum Prototyp einer Orientierungshilfe für Blinde. *PhD Thesis, Technische Universität München*, 2001.

[9] Hub, A., Teufel, H. Messung von wahrgenommenen Farben in realen Szenen und Vorausberechnung von Objektfarben in verschiedenfarbigen Umgebungen. *In: Rahman Jamal, Hans Jaschinski (ed.): Virtuelle Instrumente in der Praxis, Begleitband zum Kongress VIP* (2003), Hüthig, 11-14.

[10] Hightower, J., Want, R., Borriello, G. SpotON: An Indoor 3D Location Sensing Technology Based on RF Signal Strength, *UW CSE 2000-02-02, University of Washington, Seattle*, 2000.

[11] Identec LTD., Census (RFID). http://www.identec.com/index.php?p=products/census/matrix.

[12] Lamb T., Bourriau J. Colour: Art and Science (Darwin College Lectures), Cambridge University Press, 1995.

[13] Luo, A., Zhang X.-F., Tao W., Burkhardt H. Recognition of Artificial 3-D Landmarks from Depth and Color – a first Prototype of Electronic Glasses for Blind People, Proceedings of the 11th Scandinavian Conf. on Image Analysis (SCIA'99), 1999.

[14] Nicklas, D., Großmann, M., Schwarz, T., Volz, S., Mitschang, B. A Model-Based, Open Architecture for Mobile, Spatially Aware Applications, *Proceedings Advances in Spatial and Temporal Databases, 7th International Symposium, SSTD*, 2001.

[15] Priyantha, N. B., Chakraborty, A., Balakrishnan, H. The Cricket Location-Support System, *Proceedings of the 6th ACM MOBICOM*, (2000), 32-43.

[16] OpenSceneGraph API http://www.openscenegraph.org/.

[17] Ross D. A., Blasch, B.B Wearable Interfaces for Orientation and Wayfinding, *Proceedings of the ACM ASSETS'00*, (2000), 193-200.

[18] Sillicon Graphics Inc., OpenGL Performer. http://oss.sgi.com/projects/performer/.

[19] Sonnenblick, Y. An Indoor Navigation System for Blind Individuals, *Proceedings of the 13th Annual Conference on Technology and Persons with Disabilities*, 1998.

[20] Sun Microsystem Inc., Java3D. http://java.sun.com/products/java-media/3D/.

[21] Shevell, S. K., Wei, J. Chromatic induction: Border Contrast or Adaption to Surrounding Light? *Vision Research Vol. 38 (11)*, (1998), 1561-1566.

[22] Strothotte, T., Petrie H., Johnson, V., Reichert L. MoBIC: User Needs and Premlimary Design for a Mobility Aid for Blind and Elderly Travellers, *Proceedings of the 2nd Tide congress*, (1995), 348-351.

[23] VisuAide Inc., VisuAide Trekker. http://www.visuaide.com/gpssol.html.

[24] Wang, Y., Jia, X., Lee, H.K. An Indoors Wireless Positioning System Based on Wireless Local Area Network Infrastructure, *Proceedings of the 6th International Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services*, 2003.