Interactive Tracking of Movable Objects for the Blind on the Basis of Environment Models and Perception-Oriented Object Recognition Methods

Andreas Hub Tim Hartter Thomas Ertl

Visualization and Interactive Systems Group
University of Stuttgart, Universitätsstraße 38
70569 Stuttgart, Germany

hub@vis.uni-stuttgart.de, tim.hartter@gmx.de, ertl@vis.uni-stuttgart.de

ABSTRACT
In previous work we have presented a prototype of an assistant system for the blind that can be used for self-localization and interactive object identification of static objects stored within 3D environment models. In this paper we present a new method for interactive tracking of various types of movable objects. The state of fixed movable objects, like doors, can be recognized by comparing the distance between sensor data and a 3D model. For the identification and model-based tracking of free movable objects, like chairs, we have developed an algorithm that is similar to human perception, based on shape and color comparisons to trained objects. Further, using a common face detection algorithm, our assistant system informs the user of the presence of people, and enables the localization of a real person based on interactive tracking of virtual models of humans.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: *User-centered design*, *Prototyping* I.4.6 **[Segmentation]**: *Edge and feature detection, Region growing, partitioning* I.4.7 **[Feature Measurement]**: *Feature representation* I.4.8 **[Scene Analysis]**: *Color, Depth cues, Object recognition, Shape, Stereo, Tracking* K.3.1 **[Computer Uses in Education]**: *Distance learning* K.4.2 [**Social Issues**]: *Assistive technologies for persons with disabilities*

General Terms
Algorithms, Measurement, Design, Experimentation, Human Factors

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

1. INTRODUCTION
Remarkable advances have been made in the development and technical optimization of assistant systems which provide blind and deafblind people web accessibility equal to that of others. Regarding mobility, however, there remain many challenges to equal information access for the sensory impaired. Several devices under development or on the market provide object and/or obstacle recognition. Most basic is the cane, advancing to devices based on ultrasound [21,32], laser [15] or other technologies which transform image data into sound [29]. Even more sophisticated are eyeglasses with an embedded camera to support recognition and naming of previously-trained faces [13]. The problem in most cases is that users of these systems often know neither their own position nor the position of specific objects within their environment. Localization becomes even more difficult when movable objects are involved. In such cases, positional estimations are little better than chance.

This self-localization and navigational system for the blind and partially-sighted is based upon detailed 3D models of indoor environments. These models, thus far restricting our system to indoor environments, are available for many public and private buildings or can be created in a reasonable time. In 2004 we presented a design for a new type of indoor navigation and object identification system for the blind [8]. The basic idea of this system combines 3D environmental models with information acquired from a local sensor module. The sensor module consists of a stereo camera, a 3D direction sensor, and a keyboard. By pressing keys, inquires can be sent either to the connected portable computer or to a platform of one or more servers distributing information about the immediate and more distant environment. Initial room location of the user can be determined using conventional WiFi installations. Precise localization of the user within rooms is accomplished using our self-localization method based on distance measurement of feature points on walls and appropriate adjustment of the building model [11]. Our system enables users to recognize objects from the 3D model by detecting the closest object in front of the user, and transmitting its name over a loudspeaker. In the near future it will also be possible to solve complex navigation inquiries using the "Nexus Platform", an open platform of several servers which allows access to many different applications and will include spatial-temporal features of the user's environment [23,24]. By pressing one of the keys on our system, it is also possible to switch the speech output to other languages. This offers the opportunity to playfully learn object names in foreign languages just by exploring one's own environment [10]. This feature might also be helpful for blind children while learning their native language.

Another of our developments is a combination of our object identification system with commercial portable Braille displays. This solution was designed to afford deafblind persons [9] equal access to environmental information. Use of the Braille display might also prove advantageous for blind users in situations when acoustical output is to be avoided.

During early usability tests of our first prototype it became clear that important objects can be missed if scanning of the environment is done in an unsystematic manner. It sometimes happened that the blind user unexpectedly came upon a critical area (stairs, e.g.) because the sensor module was pointed in another direction. To avoid such occurrences we augmented our indoor 3D models with "virtual navigation areas". These were conceptualized as simple rectangles and positioned in front of important places, such as building entries, room doors, hallway intersections, elevators, stairs, emergency exits, and restrooms. In reality there are no corresponding objects in these areas. However, when a blind person walks into a virtual navigation space, appropriate spoken information is provided, including the existence of banisters or landings, the number of steps, etc. [7].

The paper is organized as follows: The subsequent section focuses on related work. We then describe our design and developments and summarize our results thus far. We next provide an outlook on future work. The paper closes with an overall discussion and a final conclusion.

2. RELATED WORK

Navigation and orientation systems for blind and visual impaired people have recently become quite an active research area. Several different, interesting, and unique approaches exist, and it seems that all compete to be the best and most widely utilized. These diverse research approaches and the systems they generate can, for the most part, be classified into a few general categories. One category can be termed basic obstacle avoidance. It includes the NavBelt from Shoval et al. [30] which produces a 120-degree wide view ahead of the user, then translates the information into stereophonic acoustical sound. Another in this category is John Zelek's [34] work to extend the range of the walking cane. Two web cams, attached to the cane, stream stereographic images into the system. Incoming signals are translated into tactile feedback, expressed via vibrating buzzers inside a glove worn by the user. A similar approach has been described by Mecocci et al. [20]. They also use image analysis to produce a simple scene description with useful landmarks, but information is delivered acoustically to the user. Other approaches are more closely related to ours, since they also take advantage of large scale infrastructure equipment for determining the location of the user. Magatani et al. [19] use an optical beacon system, in which the user carries a beacon receiver on his shoulder. However, their system is only able to determine the user's position, not his directional orientation. Drishti from Ran et al. [27] also uses optical beacons for locating a user inside an indoor environment. Additionally, their system is equipped with a differential GPS sensor, affording both indoor and outdoor navigation. A spatial database containing detailed information about the indoor environment is used, similar to our modeled environments. They also provide the user's orientation, and their current position is accurate within 22 cm. Although their system provides both indoor and outdoor navigation, the user must use a voice command to switch between the two modes.

Optical systems are vulnerable to occlusion problems, however, representing a significant disadvantage. A different approach was proposed by Kulyukin et al. [14]. They use RFID tags and receivers, as well as a robotic assistant that navigates the user inside the building. Another system uses base plates equipped with RFID tags. Spoken text concerning the current location can be provided via a receiver on the cane. This system will be commercially available soon [22].

The increasing prevalence of WiFi, Bluetooth or similar network systems allows for the localization of persons in many areas on the basis of signal strength measurements combined with map information [2]. All these systems work well in new and unknown indoor environments. However, they rely on the provision and installation of specific infrastructures – in the case of optical tracking, very specialized and expensive sensors and transmitters. Unfortunately, this often prohibits large-scale deployment.

For outdoor navigation several systems have been proposed [5,17,18,28,31] or are commercially available [4]. They most often rely on the GPS navigation system and suffer different constraints than indoor navigation systems. Therefore, they are hardly comparable to our system.

One of the most advanced systems for face recognition is the iCARE interactions assistant for the blind, developed by Krishna and Black [13]. They use principal component analysis and images of distinctive facial features to discriminate between faces. When their small, eyeglass-integrated camera recognizes a face stored in a database, the name of the person can be presented over the loudspeaker. Like our system, theirs uses images and databases, however it cannot provide information about the distance and location of recognized persons.

3. TRACKING OF MOVABLE OBJECTS

3.1 Setup and Typical Test Scenario

We have thus far developed two different sensor modules enabling object identification and augmented indoor navigation for the blind within unknown and complex environments. Our first system was hand-guided and could be held like a combination flashlight and cell phone [8]. In our latest version a synchronized stereo camera (Bumblebee BB-Col-40, by Point Grey Research) is built into the front part of the bicycle helmet (Figure 1). This head-guided version allows hands-free operation, and its higher position offers a better overview than a sensor module at the typical hand level. The disparity images of the stereo camera are used for distance analysis, and in the case of free movable objects, also for segmentation purposes, described later in detail. An inertial sensor (MT9B, by Xsens) is mounted in the back of the helmet. This includes a 3D compass, a 3D gyroscope, and a 3D acceleration sensor. Inquiry options concerning objects or navigation advice, and adjustments such as volume changes, can be made via the external keyboard of a notebook computer (Samsung, x20, 1.6 GHz), or, in the case of the hand-guided sensor module, via the integrated keyboard of a cell phone.

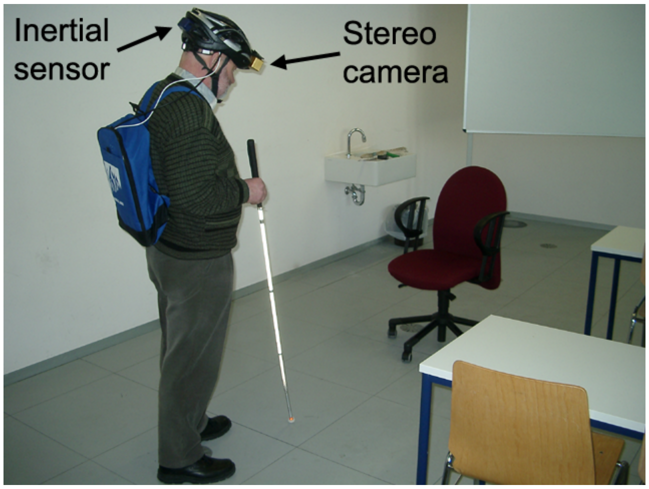

Figure 1: Typical test scenario for the tracking of movable objects. A stereo camera integrated in the front of a bicycle helmet is used for image analysis and distance measurement. The inertial sensor at the back of the helmet provides the current viewing direction. Distance and direction data are used for self-location by comparison with the 3D model, and for detection of differences between reality and model information.

A typical object identification procedure works as follows: First the user's current room is determined by means of a conventional WiFi system and a signal map showing the differences in the signal strength from various access points. Then user location within the room is determined by calculating the disparity between image-based distance measurement and the corresponding adjustment of the 3D model. The viewing direction of the user can be obtained from the 3D inertial sensor. Both location and directional data define a picking ray that enables the system to hit and recognize objects within the 3D model. Using a hierarchical scenegraph (OpenSceneGraph [26]) we can virtually navigate instantaneously within the 3D models following the real movements of the sensor, allowing real-time identification of modeled objects. The name and features of the closest object hit by the picking ray are acoustically announced to the user.

3.2 State of Partly-Fixed Movable Objects

For the recognition of the position of movable but partly-fixed objects, like doors, we work with a simple comparison of the measured distance using the stereo camera and the distance information provided by the building model. Then the distance to the object of interest is determined using the depth image calculated from the stereo images. Figure 2 shows a sequence of camera images on the left side, and the corresponding 3D model on the right side. If the distance of the targeted object (like the door on the left in the first image row) from the wall is equal to the distance from the wall in the model with the door closed, then the probability is high that the door is really closed. A corresponding message is then provided to the user acoustically.

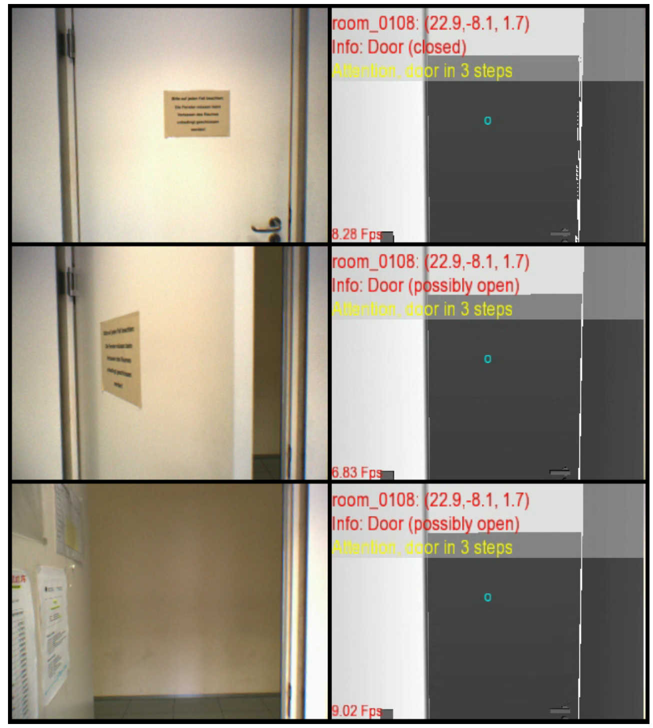

Figure 2: Sequence of a side by side synopsis of one camera image and the corresponding rendered 3D model. The comparison of real distances with virtual distances of the model suggests conclusions regarding the position of the door. Text printed on the right images – current location, the state of objects, and warnings – can be delivered to the blind user via acoustic or tactile means.

Figure 2 shows some of the text messages provided to the user: i.e., the room number and coordinates of the user (first text line), the closest object and features of interest (second line), additional warnings (third line), and other information about the objects and their location relative to the user.

Recognition of door position includes measurements of the distance from the user to the door, and/or from the user to the wall behind the door. Comparative measurements with a laser distancemeter (Disto^{TM} pro^{4}a, by Leica) have shown that the error function of the camera increases exponentially with distance. At a distance of one meter there is an error of about one decimeter. This error increases to about two meters at the distance range of ten meters. For distances typical of indoor navigation (below six meters), the camera operates with an error of about 0.2 meters. Additionally, measurement of the object's position is dependent upon detected contrast. The white wall behind the door in Figure 2 offers little contrast, increasing the probability of error. Statistically speaking, correct recognition of door position in this example would be unambiguous when the door is fully open or fully closed, and about 50% during the opening process.

Warnings can also be provided for objects that are not in the camera's line of vision. If there is a discrepancy between the image-based distance and the model information, as it is the case of a half-open or an open door (second and third image row) then a message can be given to the user that the door is "possibly open". Because of space constraints, not all acoustically-presented information can be visualized on the images.

3.3 Tracking of Free Movable Objects

When children are asked to draw an object they typically draw a rough shape and then fill it in with color. This simple but powerful example of human perception is comparable to our method for tracking free movable objects. This procedure enables our assistant system to determine the name, color and the current position of objects like chairs, which can be moved simultaneously in reality and virtually within our model.

Our algorithm for tracking of free movable objects is divided into three steps: training, recognition and model-based tracking. Figure 3 provides an overview of the algorithm. Details of the algorithm are described in subsequent paragraphs. Initially, a movable object is trained (first image on the left, Figure 3). Then, the object is moved to a new location (second row). If the object can be found again (row three) according to a stored object description a virtual model can be inserted in the 3D model (row four, right image). When the object is moved again the location of the corresponding virtual model can be updated interactively (row five).

Please note that the text information on the video screen shots of the 3D model show only a small part of the information that can be provided to the user. It should also be noted that the current frame rate, visible on the bottom left edge of these images, is not the frame rate under normal conditions. The process of recording and saving the real image and the rendered image to the hard disc reduces the frame rate significantly. Frame rates normally lie in a range between ten and thirty frames per second, depending on the complexity of the scene.

Training of Movable Objects

In the first step of our algorithm, the object of interest is trained according to its shape and its color. The corresponding object description and its name is stored in a database with shape images, color histogram information, and basic color terms. For the segmentation of the outline we use an alpha-blending between one of the color images and the disparity image of the stereo camera. The color image is multiplied by a factor of 0.9 and the disparity image is added with factor 0.1. The resulting image (Figure 4a) is used for the segmentation of the region of interest using a standard region-growing algorithm.

During the segmentation procedure the disparity information assures that only image parts with the correct depth are chosen. Otherwise objects in the background with the same color could be selected too. For the selected region an image mask is generated and stored in a database together with the pertinent color histogram and the basic color term of this region (Figure 4b).

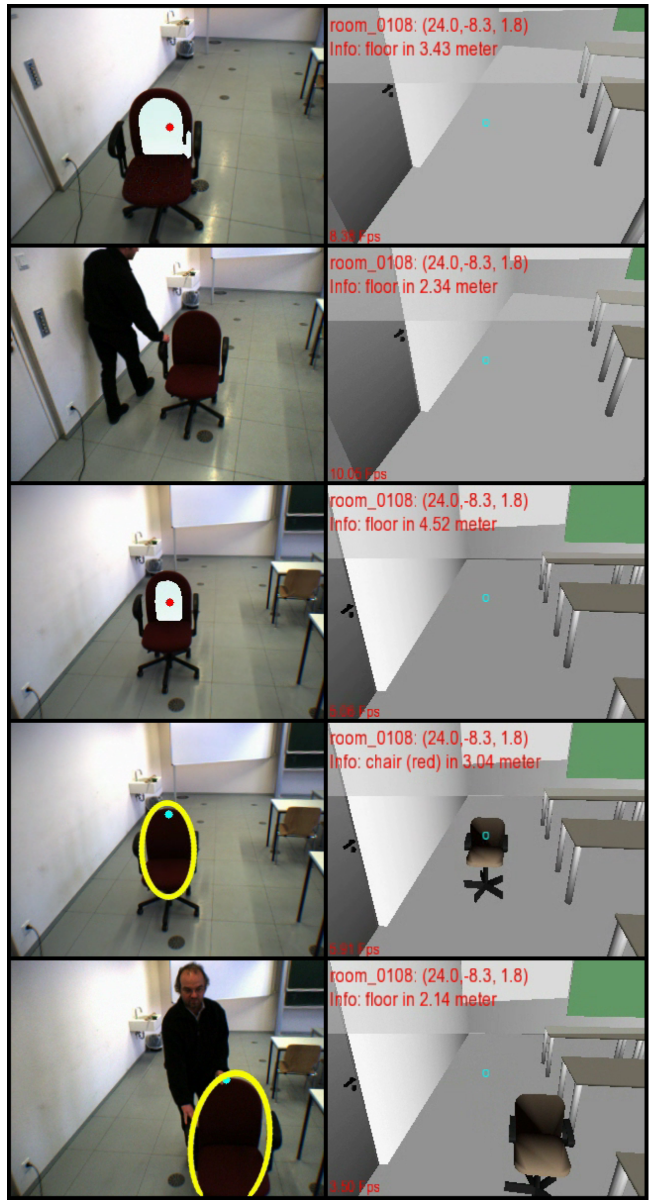

Figure 3: Training, recognition and tracking of a chair. The shape, color term, and color histogram of the seat, viewed from different positions, are stored in a database. When the chair is moved (row two) it can be recognized using these object descriptions (row three). A virtual chair model is inserted into the 3D model (row four, right image) and tracking this virtual object model (row five) allows further location and movement descriptions relative to other objects.

The color histogram consists of 16 hue-ranges of the HSV color model. The basic color term is determined with an algorithm that takes into account the region's color environment. This reflects human color perception [12], which is influenced by surrounding colors. To allow for increased independence of the viewing direction during the search, several image masks containing shape information can be stored for one object in the database.

Description-based Search and Object Recognition

In the second step, recognition, a search is initiated using camera images directed where objects of the database are expected. For this search we used the so-called CamShift-method (Continuously Adaptive Mean-Shift) of the Open Source Computer Vision Library (OpenCV) [25]. This method searches for a region with the most similar histogram values as the histogram of the object of interest. When this first test is positive a second test is initiated to assess the similarity between color terms. If this second test is also positive the shape of the objects is compared using the image masks. For this comparison the HuMoments-method of Intel's OpenCV Library is used [6]. This method involves seven special derivations of the direction, and has been shown to be independent of the rotation and the scaling of the image. When the last test has shown a good matching too, the name and the distance of the object is transmitted to the user. For this measurement the distance to the middle of the objects' top edge is used.

Model-based Tracking of Free Movable Objects

During the third step of the algorithm a virtual model of the movable object is inserted into the static 3D environment model. The position of this virtual object can be updated interactively according to the movements of the real object. If a virtual corresponding 3D model of this object is available, it can be stored with the object description in the database. Otherwise an abstract model or a simple bounding box can be used. When the real object is moved to a new location the corresponding location of the virtual object model can be updated interactively. This can be done by adding a new transformation node to the scenegraph, which shifts the virtual object to the new location.

This simultaneous tracking of the virtual object provides the blind user information about the real object, its features and its current location relative to the user's. Further, it warns of obstacles in the current environment close to real-time demands.

Restrictions, Possible Sources of Error, and Typical Durations

So far the recognition capability of the system is restricted to modeled or trained objects. That is to say, if objects which the system has not been trained to recognize are placed in a modeled environment, misleading information may be announced. The nature of this incorrect data depends upon the size of the novel object. For small items, like a briefcase or backpack, correct distance data is typically combined with the name of the closest modeled object, resulting in incorrect object identification. If the unknown object is large, like a movable poster panel, it may obscure large parts of a wall used for the self-positioning process. In this case the user's distance from the unknown object will be correct, but both the user's and the object's absolute position will be incorrect, and the object will be misnamed.

Bad lighting conditions can also lead to misinformation. Under normal office conditions (i.e., without extremely dark corners or direct sunlight), the stereo camera and the system's shutter and white balance algorithm function efficiently. Indeed, current usability tests have shown that it is advisable to turn office lights

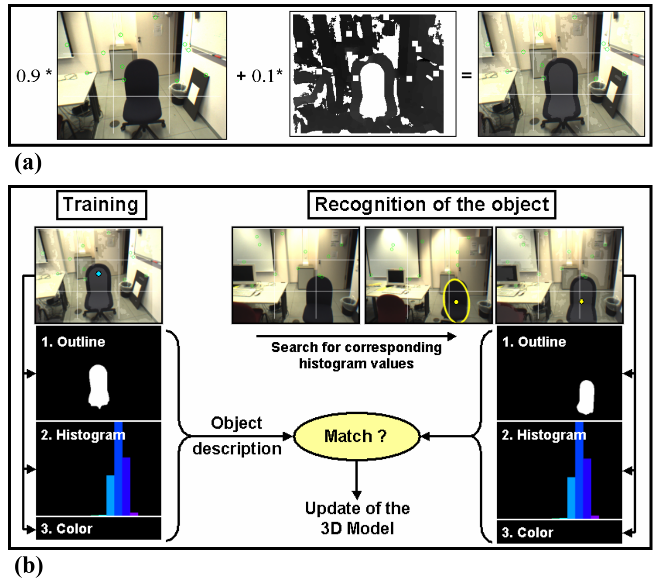

Figure 4: Training and search for a colored chair. The shape of the targeted object is extracted using alpha-blending of one of the color images and the disparity image of the stereo camera (a). The shape, the color histogram of the target region, and the region's color name are stored in a database. If the object is moved to a new location this description can be used to search for and track the object. In a successful matching process, the 3D model can be updated (b).

on to achieve optimum performance. This was a rather new experience for our blind subjects. However, since light switches are included in our model, it proved an easy adjustment.

Even when environmental objects have been modeled or trained, discrimination problems can occur. Since color and shape are the basis upon which objects are identified, similarly-shaped objects, especially those of the same color, cannot be distinguished from each other. A white metal table, for example, might be confused with a white wooden table. Also, identical objects of the same type cannot yet be detected. If there are, for example, two matching chairs close to each other, the search focus will change from one to the other while the camera is moved, but simultaneous identification is not possible.

The training process can be done within a few seconds by a person familiar with the procedure. During training the object should not be moved. Recognition accuracy is optimized by creating and saving several image masks of the same object. Imaging can be done from front, side, and back perspectives, for example, and saved within the object description. The more images saved, however, the more time it takes to train the object, the more computing power is necessary, and the more time it takes during use for the device to compare stored images and identify the object. Duration of the recognition process depends on the complexity of the scene, and as already mentioned, on lighting. Under good conditions (as shown in Figure 3) a typical recognition takes place within one or two seconds. Tracking of objects, then, is accomplished in close to real time.

3.4 Model-based Tracking of Persons

Our latest method focuses on the model-based tracking of persons. One constraint is that the face or upper body must be facing the camera more or less straight on. At present it does not discriminate between persons, but informs the user about the presence of persons and their location relative to his/her own. For the recognition of faces we use the face recognition algorithm of the OpenCV library [3] originally developed by Viola and Jones [33] and extended by Lienhart and Maydt [16]. This algorithm searches for image regions that contain what statistical model-based training has determined to be a typical facial pattern. Once a face is found we determine the distance to the user, and insert a virtual model of a person analogous to the movable objects in 3.3. Figure 5 shows a typical recognition sequence of a face, with screenshots of a tracking movie with a virtual model of a person. In the first image row on the left the face is recognized. It is possible to follow the recognized face as it moves around, and to measure its distance from the user using the disparity image. Both types of information, the detection of a man and the distance, can be transmitted to the user over the speech engine. An abstract model of a person is inserted into the virtual 3D model (second row, right image). The third line shows a screenshot of the approaching person, and the simultaneously approaching abstract object within the model. The small shift between real and virtual user positions, seen relative to the board in the background, is caused by small errors in image-based self-localization.

In normal office environments these errors are usually smaller than 0.5 meters. We should mention that for performance reasons it is not recommended at present to search for persons and movable objects at the same time. This would slow down the recognition performance significantly and prevent interactive tracking.

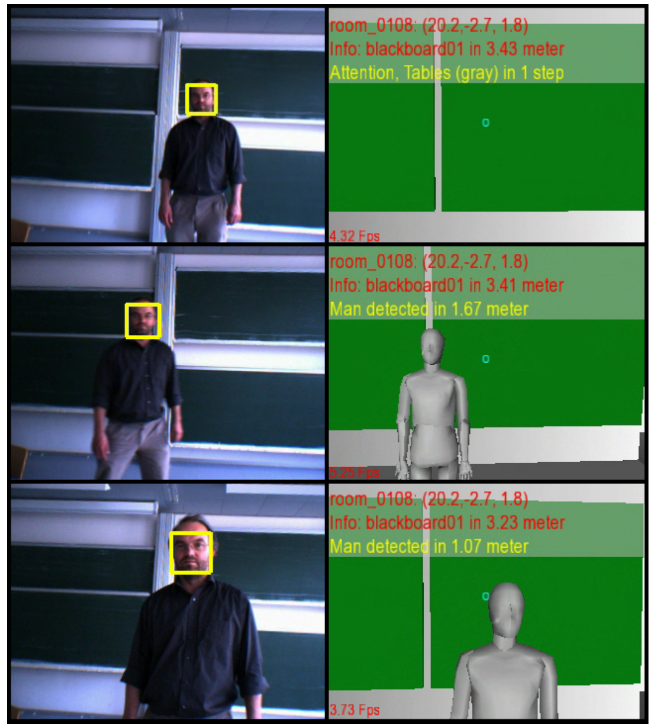

Figure 5: Model-based tracking of a person. Using a face detection algorithm a person can be detected in images (first row). The disparity information enables the insertion of a virtual person within the room model with acceptable accuracy (row two). Movements of the persons can be detected and transmitted to the user via acoustical or tactile means (row three).

4. RESULTS

We have developed a new method enabling object identification of free movable objects for the blind within unknown and complex indoor environments. On the basis of recognition algorithms and 3D environment models it is possible to track real moveable objects. Insertion of a corresponding virtual object follows the movement of the real object. If the velocity of the object does not significantly exceed normal human walking velocity, it is possible to track walking humans close to real-time demands.

Knowledge of the user's coordinates in the 3D model, and of the time-dependent location of movable objects or persons relative to all modeled objects enables our assistant system to solve information tasks heretofore possible only for a human escort. All objects present in the model or in the database of movable objects can be recognized and searched for by the blind user. Absolute coordinates of these objects, their features and distances from the user and other objects can be immediately transmitted acoustically to the user.

5. FUTURE WORK

The work presented in this paper is restricted to the limited resources of a conventional laptop. Under these conditions efficient tracking can be done only for single objects.

The next step will be to expand our system to a client-server solution. The features of such a solution were described in our previous work. By distributing extensive calculations among powerful servers, real-time processing of large environment models would be possible. Therefore, we will continue to embed our system into the Nexus framework. The Nexus framework is a general location-aware service platform currently under development at the University of Stuttgart. This approach will enable us to handle large databases with many objects. Further, extended computing power, especially that provided by programmable graphic cards, will make it possible to integrate multi-functional expert knowledge and other powerful recognition algorithms. This will allow such advances as face recognition even when faces are seen from the side, and the recognition of fast moving objects like cars.

In order to provide blind users accessibility to complex information, such as descriptions of crowd movements, usability tests must be done. These tests must involve a greater number of subjects, and should focus on software ergonomics. But optimization of hardware should also be considered. For normal users the hardware has to be reduced significantly in the next version, to be handled more discreetly and easily. Technical enthusiasts in our group of subjects are already willing to use our prototype during everyday activities, as bicycle helmets are not uncommon and relatively comfortable. They report that concern over the system's appearance is far outweighed by their gains in functional mobility and independence.

6. DISCUSSION

Blind persons are able to move and navigate easily within known environments with a speed that is sometimes astonishing. In contrast, their orientation and navigation in unknown environments can be both difficult and dangerous. Based on the recent development of mobile computing devices and optimized network connections, it appears possible to mitigate these mobility challenges. In our opinion it is currently impossible to achieve object identification of arbitrary objects using systems that are only based on image segmentation and image interpretation. We are therefore convinced that object identification, orientation and navigation tasks for the blind can be optimized significantly by combining local multi-sensor information with global world models.

For indoor environments, most landmarks and objects important to the blind can be stored in 3D models of selected buildings. Government and public buildings, and training centers for the blind, for example, are available or can be created in a reasonable period of time. The time required to set up the system within a multi-story building varies from a few days to several months, depending upon the complexity of the building/rooms and the desired level of detail. Training time is less for simple mapping and naming of walls and doors, and greater for 3D models which include door handles and light switches. Increased man power can speed up the process significantly. We have begun to experiment with automatic modeling, using a combination of 3D laser scanners, high resolution and panoramic cameras, developed by Biber et al. [1]. Thus far, semantic demands inherent in the naming process necessitate manual modeling.

Outdoor environments, however, still present quite a challenge to independent navigation for the blind, given the dearth of available computing resources and the complexity of the environment.

Overall, generic use of our system is fairly simple. Blind persons who use cell phones or computers are able to learn its basic features (i.e., object recognition, distance estimation, and the detection of humans) within minutes. Operation of more advanced features (like keyboard shortcuts and training of environmental objects) takes more time and practice, from several hours to several days. Learning time varies with the individual's motivation, patience and general aptitude, just as learning to use the more complex options on a cell phone.

By embedding our device into the Nexus architecture, several new possibilities become available. Besides navigating, orienting, and local object identification, the system can offer everyday services designed for sensory handicapped people, such as medical support. Furthermore, our device could provide time-dependant warning messages.

While considering potential research and development areas, we must also consider some remaining challenges. One issue is energy consumption of the portable computer, which currently limits its usage to a maximum of four hours. Another is the effect of electromagnetic fields and strong temperature drifts, which increase the errors of measurement. Further, a good compromise must be found between the size of the device and its robustness. And, although the Nexus platform provides countless additional options for assistant systems, it was not built exclusively to serve the blind. Extracting relevant information, without overwhelming the user with extraneous detail, may prove difficult.

7. CONCLUSION

Apart from the optimization work that must be done in the near future concerning system hardware and software ergonomics, limited usability tests have shown that our method of tracking movable objects, combined with environment model information, provides the basis for safe and independent navigation for blind and deafblind people. It offers real assistance to sensory impaired users, providing complex recognition and navigation tasks at home, at school, at the workplace, and during leisure time.

8. ACKNOWLEDGMENTS

This project is funded by the Deutsche Forschungsgemeinschaft within the Center of Excellence 627 "Spatial World Models for Mobile Context-Aware Applications". We would like to thank the blind persons and mobility teachers, who tested our system, especially our colleague Alfred Werner, for their patience and suggestions for improvements.

9. REFERENCES

[1] Biber, P., Fleck, S., Strasser, W. The Wägele: A Mobile Platform for Acquisition of 3D Models of Indoor Outdoor Environments. *9th Tübingen Perception Conference (TWK 2006)*, Tübingen, Germany, 66, 2006. http://www.twk.tuebingen.mpg.de/twk06/twk2006.pdf

[2] Bowen, C. L. III, Using Wireless Networks To Assist Navigation For Individuals With Disabilities. *California State University, Northridge Center on Disabilities' 21st Annual International Technology and Persons with Disabilities Conference,* Los Angeles, CA, USA, 2006. http://www.csun.edu/cod/conf/2006/proceedings/2931.htm

[3] Face detection. http://www.intel.com/technology/itj/2005/volume09issue02/art03_learning_vision/p04_face_detection.htm

[4] Gomez, L. Trekker Reviewed an Advanced GPS System for the Blind. *California State University, Northridge Center on Disabilities' 21st Annual International Technology and Persons with Disabilities Conference,* Los Angeles, CA, USA, 2006. http://www.csun.edu/cod/conf/2006/proceedings/2829.htm

[5] Hang, R., Black, J. The Use of a GPS-Mapping Program in Orientation and Mobility. *California State University, Northridge Center on Disabilities' 21st Annual International Technology and Persons with Disabilities Conference,* Los Angeles, CA, USA, 2006. http://www.csun.edu/cod/conf/2006/proceedings/ 2950.htm

[6] Hu, M.-K. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory,* Vol. 8, 179-187, 1962.

[7] Hub, A., Diepstraten, J., Ertl, T. Augmented Indoor Modeling for Navigation Support for the Blind. *Conference Proceedings: CPSN'05 - The International Conference on Computers for People with Special Needs,* Las Vegas, NV, USA, 54-59, 2005.

[8] Hub, A., Diepstraten, J., Ertl, T. Design and Development of an Indoor Navigation and Object Identification System for

the Blind. *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility ASSETS 2004,* Atlanta, GA, USA, Accessibility and Computing, No.77&78, 147-152, 2004.

[9] Hub, A., Diepstraten, J., Ertl, T. Design of an Object Identification and Orientation Assistant for the Deafblind. *Conference Proceedings: 6th DbI European Conference On Deafblindness,* Presov, Slowakia, 2005.

[10] Hub, A., Diepstraten, J., Ertl, T. Learning foreign languages by using a new type of orientation assistant for the blind. *Conference Proceedings: International Council for Education of People with Visual Impairment, European Conference,* Chemnitz, Germany, 339-341, 2005.

[11] Hub, A., Hartter, T., Ertl, T. Interactive Localization and Recognition of Objects for the Blind. *California State University, Northridge Center on Disabilities' 21st Annual International Technology and Persons with Disabilities Conference,* Los Angeles, CA, USA, 2006. http://www.csun.edu/cod/conf/2006/proceedings/2661.htm

[12] 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.

[13] Krishna, S., Black, J. iCARE Interaction Assistant: A Wearable Face Recognition Device to Facilitate Social Interaction. *California State University, Northridge Center on Disabilities' 21st Annual International Technology and Persons with Disabilities Conference,* Los Angeles, CA, USA, 2006. http://www.csun.edu/cod/conf/2006/proceedings/2893.htm

[14] Kulykukin, V., Gharpure C., DeGraw N. *Human-Robot Interaction in a Robotic Guide for the Visually Impaired*, AAAI Spring Symposium, 158-164, 2004.

[15] Laser long cane. http://www.vistac.com/VISTAC_Laser_Long_Cane.htm#Detailed%20Description

[16] Lienhart, R., Maydt, J. An Extended Set of Haar-like Features for Rapid Object Detection. *IEEE International Conference on Image Processing,* Rochester, NY, USA, Vol.1, 900-903, 2002.

[17] Loomis, J., Gooledge, R. G., Klatzky, R. L. Navigation System for the Blind: *Auditory Display Modes and Guidance.* Presence Vol.7 No.2, 192-203, 1998.

[18] 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 Conference on Image Analysis (SCIA'99),* Kangerlussuaq, Greenland, 1999.

[19] Magatani, K., Sawa, K., Yanashima K. Development of the Navigation System for the Visually Impaired by Using Optical Beacons. *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* Istanbul, Turkey, 488-1490, 2001.

[20] Mecocci, A., Lodola, R., Salvatore U., Outdoor scenes interpretation suitable for blind people navigation. *Proceedings of the 5th IEEE International Conference on Image Processing and its Applications,* Edinburgh, UK, 256-260, 1995.

[21] Miniguide. http://www.gdp-research.com.au/minig_1.htm

[22] Naviwalk. http://www.idfone.co.kr/

[23] Nexus. http://www.nexus.uni-stuttgart.de/index.en.html

[24] Nicklas, D., Großmann, M., Schwarz, T., Volz, S., Mitschang, B. A Model-Based, Open Architecture for Mobile, Spatially Aware Applications. *7th International Symposium on Spatial and Temporal Databases (SSTD 2001),* Redondo Beach, CA, USA, 117-135, 2001.

[25] Open Computer Vision Library Documentation. http://www.cs.bham.ac.uk/resources/courses/robotics/doc/index.php

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

[27] Ran, L., Helal, S., Moore, S. Drishti: An Integrated Indoor/Outdoor Blind Navigation System and Service. *Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04),* 23, Orlando, Florida, USA, 2004.

[28] Ross D. A., Blasch, B. B. Wearable Interfaces for Orientation and Wayfinding. *Proceedings of the 4th International ACM Conference on Assistive Technologies ASSETS 2000*, Arlington, VA, USA. 193-200, 2000.

[29] Seeing with Sound. http://www.seeingwithsound.com/

[30] Shoval, S., Borenstein, J., Koren, Y. Mobile Robot Obstacle Avoidance in a Computerized Travel for the Blind. *The IEEE International Conference on Robotics and Automation,* San Diego, CA, USA, 2023-2028*,* 1994.

[31] 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,* Paris, La Villette, France, 348-351, 1995.

[32] UltraCane. http://www.soundforesight.co.uk/product.htm

[33] Viola, P. and Michael J. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Kauai, HI, USA, 511-518, 2001.

[34] Zelek, J. The E. (Ben) & Mary Hochhausen Fund for Research in Adaptive Technology For Blind and Visually Impaired Persons. Research Project. Available from: http://www.soe.uoguelph.ca/webfiles/jzelek/cnib2.pdf.