Applications of Virtual Reality for Visually Impaired People

TORRES-GIL, M. A., CASANOVA-GONZALEZ, O., GONZALEZ-MORA, J. L.
Department of Physiology
Universidad de La Laguna
Campus de Ciencias de la Salud S/N, 38071 Santa Cruz de Tenerife
SPAIN
mtorresg@ull.es ocasano@ull.es jlgonzal@ull.es http://nf.ull.es

Abstract: - This paper describes in detail the development and applications of a Virtual Reality Simulator for Visually Impaired People. It makes an auditory representation of the virtual environment, rendering the virtual world entirely through the hearing. The simulator has these main purposes: validation of auditory representation techniques, 3d sensor emulation for environment recognition and hardware integration, training of visually impaired users with these new auditory representation, and acoustic perception experiments aimed to improve the auditory rendering.

The interaction with the simulator is made by a 3d tracking system to locate user's head orientation and position. This means the user interaction is as natural as possible, all performed by just "walking through" the environment, and at the same time, the user perceives the environment through acoustic information.

Keywords: - Virtual reality, visually impaired, spatial sound, electromagnetic position tracker

1 Introduction

The main idea behind this virtual reality for visually impaired users lies on the extraction of distance information between the user and the virtual elements of the simulation and then the transformation of this information into sound. The process where this "depth" information is translated into an auditory representation is the **sonification** process. The origin of this idea started in the Virtual Acoustic Space project [1, 11, 12]. The literature reports several devices which offer auditory spatial information for the blind person's orientation and mobility [13]. Some teams have also explored the human's ability to recognize an object's shape from sound [14]. Further researching on this project and its continuations led to the development of the tools needed to perform virtual simulations about the sonification proccess. The current simulator, formerly called Virtual Reality Simulator for Sonification Studies or VRS3, is a tool that have been used in different research projects and it is constantly improved and upgraded.

Unlike typical Virtual reality simulators, the virtual environment is rendered using an acoustic representation instead of a visual representation. This simulator is aimed to visually impaired users, who mainly use hearing and touch to interact with the environment, so that is the reason on focusing in the hearing channel.

The goal in the development of this simulator is the creation of a tool for researching different approaches to represent three-dimensional information through sound. These sonification techniques have two applications: the interaction with virtual environments and the use in portable prototypes with environment detection features to give extra information about the real environment where the user is moving, as the techniques and prototypes developed and applied in the Virtual Acoustic Space project [1].

A detailed description of this simulator and its modules, the technical issues solved during the development and also some applications in different research projects will be addressed in this paper.

2 Previous work

Virtual reality for visually impaired people is not a new topic. Different approaches to this goal can be found in many published material and a lot of groups are working on enhancing the interaction with computer to blind people. Even multimodal systems have been developed, trying to give as much feedback as possible to enhance the immersion into the virtual reality, like the HOMERE system [2]. Other systems also combine the concept of acoustic virtual reality into mobility aid prototypes [8, 10], transforming spatial information into auditory cues to help guiding blind users in the real world.

This group developed his first virtual reality simulator as a tool in the scope of the project Virtual Acoustic Space III [3]. This first simulator was developed as a tool oriented to finding out the limitations of physiological capacities of the acoustic system. Through the use of this tool some limitations and problems led to the decision of redefining the idea behind the simulator and make if a more usable and

generic tool for newer experiments and researching. The main limitations to overcome were the following:
Tracker accuracy: as it is explained in detail in the next chapter, the 3D tracking device accuracy was very poor because the big workspace it was pretended to be used. A software based calibration and distortion correction looked like the approach to the solution.
Adaptability: the simulator has to be modular and configurable for multiple setups.
Scene formats and scene metadata: add more complex scene support and object handling.

The development of these features made a new simulator almost built from the scratch, and led to the actual virtual reality simulator VRS3.

3 System overview

The VRS3 simulator has been made using a modular design so it will perform in different setups. The whole system, represented in fig. 1 as a block diagram, is composed of a set of small modules, each one performing a different task. Not all the modules have to be working simultaneously in a given setup, thus for example the tracking system is only present in the simulator room, but a portable version controlled with standard input like keyboard and mouse won't need the tracking. The same happens with the sound processor, which can be a dedicated hardware like the PEGASO in VASIII [3] or run everything in software like the new sound processor developed for this simulator. The 3D simulation module is always present in any setup, performing the scene management, distance calculations, depth maps generation to feed the sound processors and user interaction.

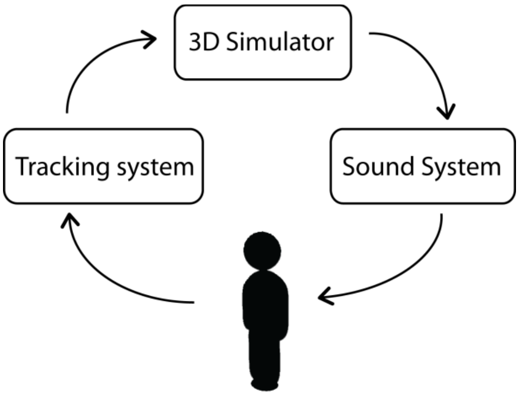

Fig. 1

The simulator is installed into a dedicated room that has been acoustically isolated and covered inside with panels for reverberation reduction. The current room is 4 meters width and 5.5 meters length, and the ceiling is about 3 meters height. The room has one access door and one window to monitor from outside, where the simulator computer is located. Fig. 2 shows both pictures from inside the room and outside.


Fig. 2

Inside the room the 3d tracking system is placed, including the antenna (the white ball seen in fig. 2), and the tracker processor. Fake walls, columns and furniture are placed also inside the room to match the virtual objects and add also the "touch" sense.

Outside the room in the control console the simulator computer and some sound hardware are placed. In front of the control console there is a window to allow the simulator operator watching what is going on inside. On fig. 3 a schematic top view of the simulator room and the control console is showing the layout of all the components.

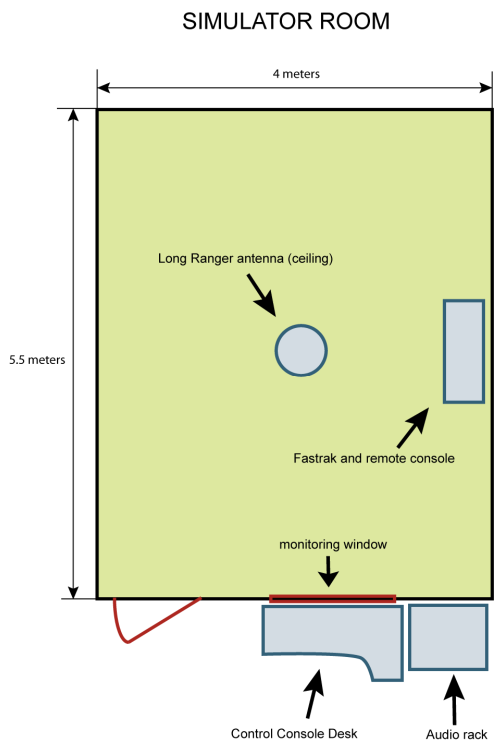

Fig. 3

As fig. 3 shows, the tracker antenna is placed in the middle on the room, hanging from the ceiling, and the remote console at one side of the room, holding the tracker hardware and a keyboard for remote management of the simulator computer. Outside the room the operator desk is located in front of the window and next to the audio rack. There is also a 32 inch screen for monitoring inside mounted on a mobile base so it can be placed in any position inside the room of placed outside if is not needed.

For simulation purposes and also to make graphical representations the simulator room had been modeled in 3d. In fig. 4 a render of this geometry gives an idea about the whole size of the room and how it looks like.

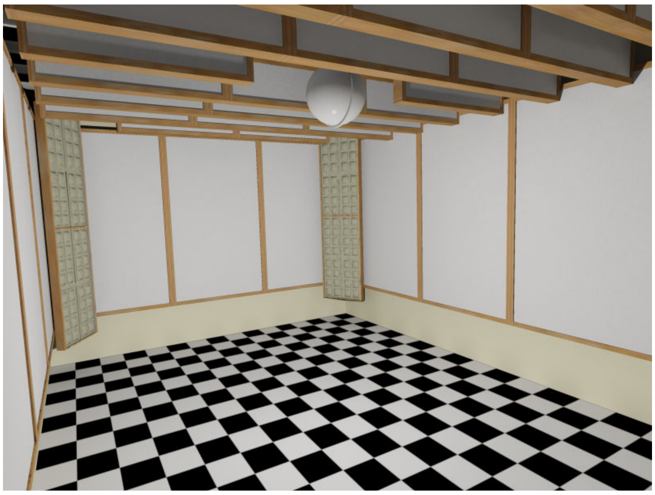

Fig. 4

The render in fig. 4 gives a better idea about the room size than fig. 2, where the camera couldn't take the whole room. The isolation panels in walls and ceiling, and also the diffusers in the corners attenuate the reverberation and echoes inside the room to facilitate the acoustic perception and also to isolate from outside noise.

3.1 The tracking subsystem

The main user input and channel of interaction is performed with the 3D tracker. The tracker allows locating the user's head position and orientation, so the user navigates through the virtual world just walking, moving and looking around. This approach makes the interaction more natural than a joystick or a mouse, and it is easier to learn for visually impaired users.

3.1.1 Fastrak

The tracking device used for the VRS3 is a magnetic field based tracker, the Polhemus Fastrak. This tracker provides absolute position and orientation for up to four probes and it's quite common in many virtual reality setups. We use the LongRanger antenna to cover the full simulator room, because the standard antenna gives a very tight space to interact. Both the fastrak device and the long ranger antenna can be seen in fig. 5.

Fig. 5

One small probe like the one seen in fig. 5 is attached to the diadem of the headphones used for gibing the auditory rendering. User interaction is performed all by wearing this set of headphones and tracker probe, reading the user's position and orientation and sending also the sounds, as seen in fig. 6.

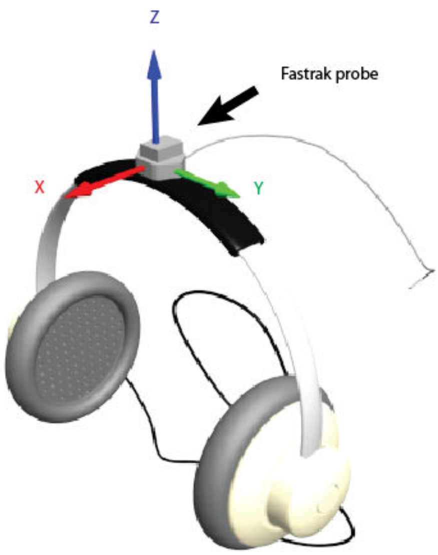

Fig. 6

3.1.2 Calibration

Magnetic based trackers have some well known problems. While they are quite accurate under ideal conditions, any metallic object present close to the antenna or the probes will distort the generated magnetic field and the measurements [4]. Due to the size of the simulator room and the presence of metallic structures in walls and ceiling, these measurement errors were too big to allow proper simulations. A calibration and correction technique was implemented in order to solve most of these errors and make the tracker more "usable".

The technique implemented to reduce the tracker error is a quite simple technique, similar to some tri-linear interpolation techniques commonly applied to magnetic field based trackers [4]. The implementation of this simple technique instead of trying a more accurate one focused on obtaining a compromise between correction performance and development effort. A technique capable of making the tracker suitable was preferred over solving the whole full accurate correction problem.

The first step to perform the calibration is to measure a grid of points at known positions and orientations. The idea behind this method is to interpolate these values to obtain the corrected measure at a given point. As measuring points is a time taking task, a small calibration grid was measured and then augmented in resolution by an offline spline interpolation. The distortion in the measurement is very noticeable when the measured grid is plotted, as shown in fig. 7 where a perfect regular squared grid measured in real coordinates gives a totally distorted result in the fastrak measured coordinates.

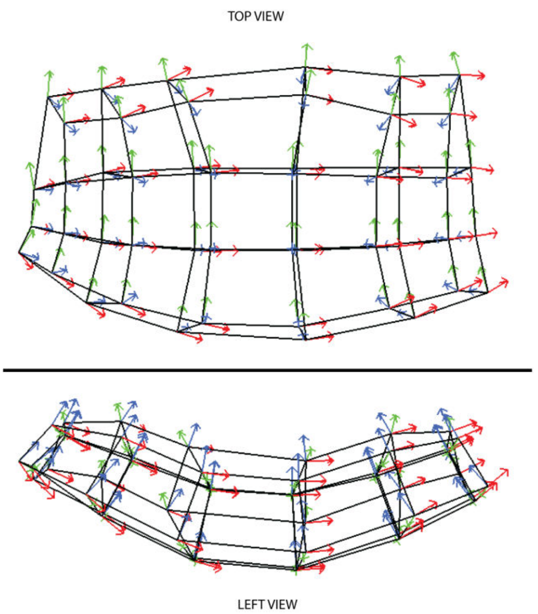

Fig. 7

This grid has each point 80 centimeters away from its neighbors. Together with the position errors, the orientation vectors are plotted. Even if all the measurement points had the same orientation, aligned orthogonally to floor and walls, the plotted vectors change a lot across the grid.

Because the measured grid is uniform in the real space but is distorted in the uncorrected space, a regular tri-linear interpolation is not valid straight. The implemented approach looks for two triangles into the calibration grid, and intersects them with a line that passes through the measured point to be corrected. Then the values are interpolated linearly in both triangle intersections and in the intersection line as shown in fig. 8.

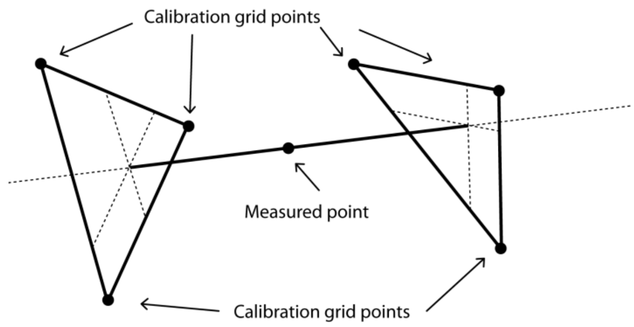

Fig. 8

The correction algorithm looks for a set of points in the correction grid where the current tracker measured point is contained. Each grid point has a "real" point values and orientation values associated, and these values are linearly interpolated to obtain the current measure correction. The calibration made a notary improvement in the accuracy of the tracker. Exhaustive measurements to validate the correction have not been done, but just some simple validation test. These tests show a very nice response from the system, removing most of the artifacts and distortions of the uncorrected system. The correction method will be refined and improved, but in its actual state allows the use of the simulator for practical applications. This method also corrected most of the tilt issues regarding orientation. Still the corrected values are not as smooth as desired, and is some areas of the room the incorrect tilt is still noticeable, but small enough to run simulations.

3.2 The 3D Simulator

The simulator uses a 3D engine to manage the virtual environment. The main task of the 3D engine is to generate the depth maps that a real 3D sensor would generate from a real scene. It also gives visual feedback about what is happening in the simulator, so the supervisor of the experiment or training session has all the information about the virtual scene the blind user is interacting with. The 3D engine has been implemented in C++ and OpenGL, and it is using the Fox Toolkit library for the user interface. The simulator currently runs on windows platforms but the source code is portable and can be compiled in Unix platforms also.

3.2.1 Scene management

The simulator actually loads scenes from 3DS mesh format, but is also planned to be updated to other 3D formats so the scenes can be designed in multiple 3D editors.

The simulator handles multiple scenes at the same time and manages the scenes as two groups. The background group will only give visual feedback and be used as references for the operator, not appearing in the depth maps, and the main group of scenes will appear into the depth maps, so it will be the group of scenes that the user is interacting with. This allows the operator to load reference objects, like the simulator room to be aware of the limits and walls, and virtual objects. This approach also gives the ability to split the elements of the scene to study how each one of them perform in the sonification process, like for example including or excluding the floor from the sonification.

The simulator room is currently being simulated with a 3d mesh with about 30000 polygons. All the elements in the model are separated and can be selected individually to generate or not sounds. A wireframe rendered version of this mesh is shown in fig. 9. Big flat surfaces have been subdivided to avoid artifacts in the depth map generation, as the distance calculation is calculated per vertex.

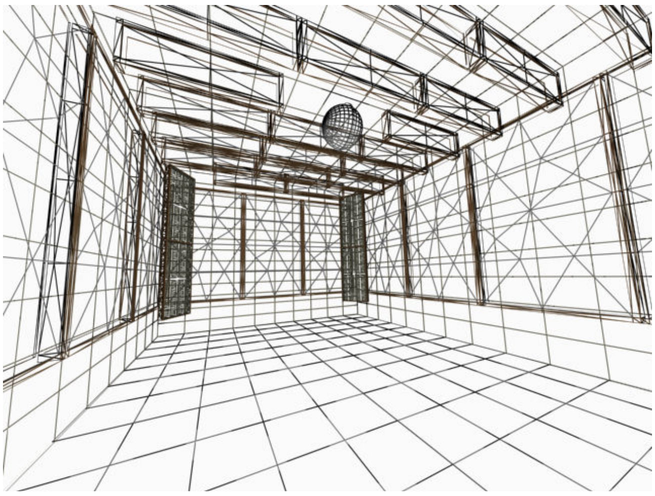

Fig. 9

3.2.2 Depth calculation

The depth maps generated by the simulator consists generally on a two-dimensional map that represents, instead of the color of the pixel, the distance at that pixel is from the camera.

These depth maps are calculated directly from the 3D mesh geometry used in the simulator, and these calculations are performed in the graphics hardware. This implementation works smooth and doesn't need a powerful computer to run the simulation, just a medium performance graphics card that supports OpenGL 2.0. The depth calculation algorithms are written in GLSL shading language, and new algorithms can be written easily without reprogramming the whole 3D simulator.

Fig. 10 shows a virtual camera and the depth map that it will catch from the scene it is capturing. Represented in grayscale, lighter values mean closer distance and darker values farther distance.

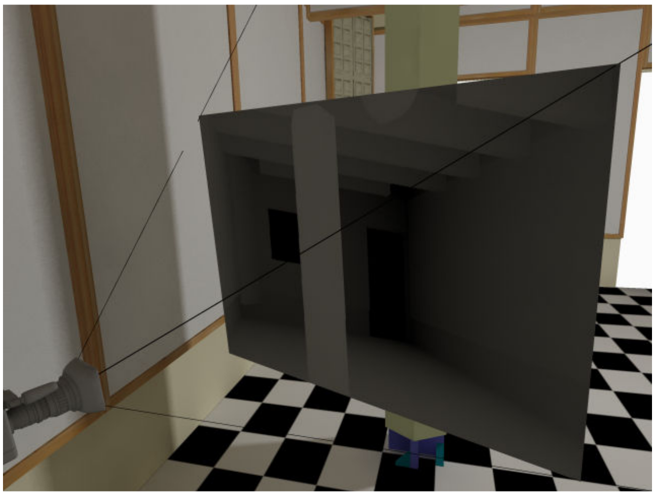

Fig. 10

3.2.3 Screen rendering

Screen rendering is used mainly for monitoring purposes and also for development and debugging. The simulator renders the scene viewed from the tracker's probe, the same field of view that the user has inside the room. Depth maps are shown also overlapped to the rendered scene in the simulator to allow the operator to check the matching between the virtual scene, the depth map generated, and the sonification of the scene itself. As shown in fig. 11, the depth map of two virtual columns inside the simulator room is overlapped to the scene and represents a 64x48 pixels size map with a pseudo-color depth scale.

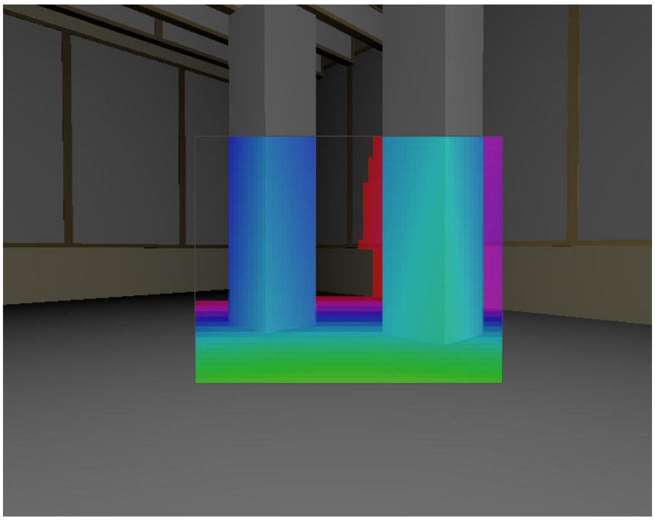

Fig. 11

The depth map rendered in fig. 11 has elements from the two columns and also the floor and the walls of the room. This map is clamped from half meter distance to 5 meters, so that is why the walls behind and the corner disappear from the depth map. In the following fig. 12 a 13x7 pseudo-color depth map of two columns also is displayed. In this case, the room model is not contributing to generate depth information, so walls floor and ceiling are not producing any sound. This is very useful if very simple experiments or training sessions have to be performed, so the user can focus on the objects without the presence of the room, like in a virtual void space.

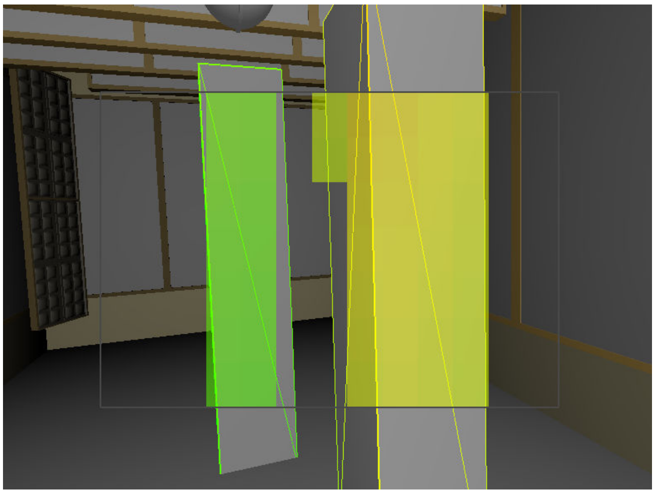

Fig. 12

Fig. 13 shows a 13x7 grayscale depth map, but in this case both the columns and the room are contributing to depth information. The clamping range in this case is from half meter to 15 meters, so that's why, unlike in fig. 11, all the elements are filling the depth map.

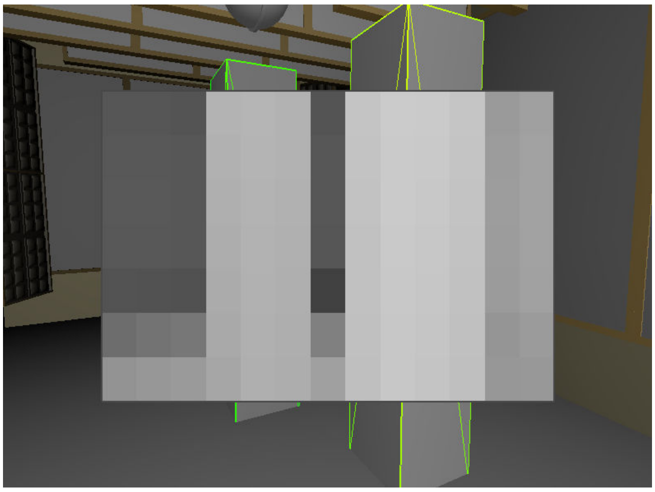

Fig. 13

The depth map resolution and clamping ranges are directly related to the sound generation and the sound banks used in the simulation.

The depth scale can be a grayscale or a pseudo-color scale just switching the depth rendering shader, and new shaders can be developed quickly to make different scales and filtering. This information is very useful when new sound collections or sonification strategies are being developed and tested.

For convenience to the operator of the simulator, the 3DConnexion SpaceNavigator control device shown in fig. 14 is now supported in the simulator, allowing the operator to navigate and test scenes quickly, without needing to turn on the tracker or even running the simulator in any other computer away from the simulation room.

Fig. 14

To add some virtual reality features also available to non visually impaired users the simulator has also in development a 3D stereo GLSL shader for 3D stereo representation. Current tests are being performed with a Vfx3D Head Mounted Display device but is planned to be tested with more modern

HMDs or with the new 3d monitors and glasses available in the market.

3.3 The sound renderer

The sound renderer module is in charge of transforming the depth maps into an auditory representation of the scene [5]. Each element of the depth map will have a matching spatialized sound that will help to locate it on the virtual acoustic space. The spatialitation of the sounds is made by applying custom measured Head Related Transfer Functions. The first stage of the sound module is to select and sort the sound elements, since usually the size of each depth map is too big to have a fast enough auditory representation. This first stage is called the sonification strategy, where the most important sounds of each frame are selected and transformed into a timed sequence.

The sonification strategy will lead to a sequence of sounds with exact timing between them and most of the time with overlap. The sound system will mix these sequences through the soundcard to deliver the auditory feedback to the simulator's user.

The current implementation of the simulator includes a software sound rendered module, performing the sound mixing and timing in real time, but previous versions communicate through RS232 connection with the hardware sound processor PEGASO developed in the Virtual Acoustic Space project. More detailed explanations about the sound generation and the spatialized sound can be found on the Literature about the Virtual Acoustic Space project [1, 11, 12].

4 Applications and studies

4.1 Perception

VRS3 has been used as a tool to carry on experiments and studies about auditory perception, including testing different approaches to the generation of a virtual acoustic space, changing parameters in the sounds and also in the algorithms.

For example, fig. 15 shows the simulator room during an experiment about slightly variations in the timing of the sounds mixing, where the user has to try and test different configurations to determine the ones that gives more accuracy and the ones that make a more comfortable sensation on the user.

The virtual scene presented in this case is very simple and consists only on one object, either a column or a wall, to compare performance in the representation of narrow objects and very wide objects.

Experiments on this field and focusing on making better auditory representations are well documented also in the literature and published in many conferences, like the interesting work of Hong Jun Song and Kirsty Beilharz [9].

Fig. 15

4.2 Sonification techniques

VRS3 is used also for validation purposes. Each time a new modification in the sonification of the environments, including new sounds, new sets of HRTFs or new sonification strategies and algorithms are carried on, are then tested in the simulator to fix bugs, look for sound artifacts and check if the auditory representation are performing as expected.

Different sonification strategies have been tested, from linear sequence of sounds, randomization of sounds order, object's edges sonification, etc…

The development about sonification techniques involves also testing sets of HRTF with different angular resolutions, processing sound at 16 or 24 bit depth, using different sample length for HRFT and for the processed sounds. Looking for a compromise between real time performance and sound accuracy determines the final parameters for sound processing.

4.3 Training

The simulator is used to train users of our mobility prototypes in order to get them used to the auditory perception and how to move around any environment with the new feedback before training with the real device in real situations. This involves experiments about performing location and navigation tasks, measuring times, accuracy of the subject recognizing the environment and mistakes. These training sessions also give very interesting feedback about how visually impaired people performs in common navigation and environment recognition tasks.

Fig. 16 shows a typical training setup, where the user has to complete some tasks, like identify the objects present in the scene and its positions, and also navigate through this scene to “touch them”.

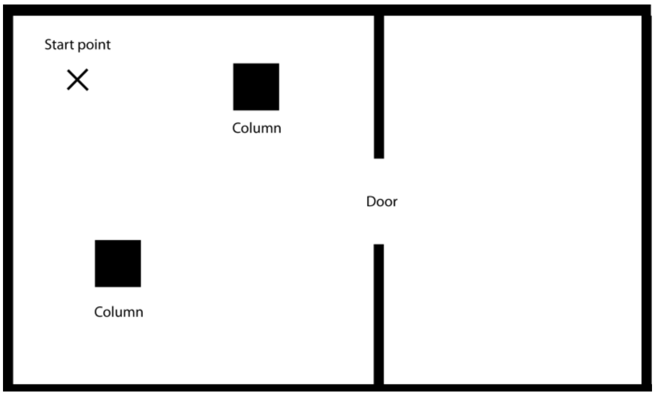

Fig. 16

These kinds of training environments are created both in the real and in the virtual world, using fake furniture and modeling that furniture in the simulator also. These double representations provides the auditory representation and also the tactile feedback that help the user achieving faster learning curves

Fig. 17 shows a picture of a user performing a training task, moving around fake walls and columns and trying to identify the whole environment.

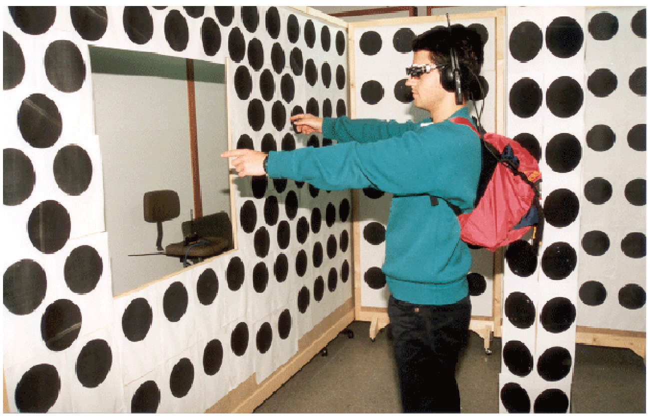

Fig. 17

In this kind of training sessions we measure completion times, collisions with undetected virtual objects, accuracy in the description of the locations and a log of the trajectories walked along the session. These training sessions are also performed with prototypes prepared to be used in the real world for navigation aiding.

4.4 Sensor emulation and hardware integration

The early testing, development and debugging of sensors and sound processing hardware with the simulator has been a task commonly performed in the last research projects performed by this group.

During some research projects, the simulator was used also to emulate some environment sensors, like a Time Of Flight 3D sensor used in the CasBlip project [6], to allow the developing and testing of the audio techniques before the real sensor was built and available.

Fig. 18

The final prototype, shown in fig. 18, was compound by a 3d distance sensor (glasses), the control and sound generation hardware, and a pair of headphones. This prototype has been used by different blind users and tested in indoor and outdoor real life situations. The sonification techniques applied in this prototype were developed and tested in the virtual reality simulator.

Currently the simulator is integrating an infrared Time of Flight based 3D camera to start running simulations of a new mobility enhancement device for visually impaired users in early development.

4.5 Conclusions and future work

The virtual reality simulator VRS3 has proven its utility and value through different research projects and experimentation. There is still a lot of work to do about visually impaired people integration into computer technologies. VRS3 is growing and some new features are planned to be implemented according to the needs of the researching. The tracking calibration will be refined to gain some extra accuracy. The scene management will implement individual object's metadata and occlusion algorithms to integrate other features in the simulator like different sound encoding for different object categories. Also the scene file types is planned to grow to most of the common 3D scene formats. The integration with a physics simulation engine is also a planned step in order to give the environments life, as they are right now mostly static. The generation of auditory images and the sonification process is also very suitable for improvements and enhancements. Real time convolution of the HRTF functions with the sounds is a feature planned to be added to the simulator. Also new methods to obtain the HRFT functions, like the artificial neural network method proposed by Ivan Bogdanov, Virgil Tiponut, Radu Mirsu [7].

Acknowledgements
The funds for the development of this tool had been promoted by the following national public projects:

- ULLAPD-08/01 granted by the Agencia Canaria de Investigación, Innovación y Sociedad de la Información.
- Ministerio de Industria, Turismo y Comercio (Proyecto AVANZA); TSI-020100-2008-337
- Ministerio de Ciencia e Innovación (Proyecto CICYT); TIN2008-06867-C02-01/TIN

References:

[1] Gonzalez-Mora, J.L., Rodriguez-Hernandez, A.F., Burunat, E., Chulani, H.M., Albaladejo, J.C. and Rodriguez-Ramos, *L.F. Touch, Blindness, and Neuroscience*, chapter Seeing the World by hearing: Virtual Acoustic Space, pages 371-383. 2003

[2] Lecuyer, A., Mobuchon, P., Megard, C., Perret, J., Andriot, C. and Colinot, J.P. 'HOMERE: a multimodal system for visually impaired people to explore virtual environments', *IEEE Virtual Reality Conference*, p.251. 2003

[3] Gonzalez-Mora, J. VASIII: Development of an interactive device based on virtual acoustic reality oriented to blind rehabilitation. *Jornadas de Seguimiento de Proyectos en Tecnologías Informáticas. Spain* 2003

[4] Volodymyr V. Kindratenko. A survey of electromagnetic position tracker calibration techniques. *Virtual Reality: Research, Development and Applications*, 5(3):169-182, 2000

[5] Rodriguez-Ramos, L., Chulani, H., Diaz-Saco, L., Sosa, N., Rodriguez-Hernandez, A. and Gonzalez-Mora, J. Image and sound processing for the creation of a virtual acoustic space for blind people. *Signal Processing and Communications*, pages 472-475, 1997.

[6] Sound Map Generation for a Prototype Blind Mobility System Using Multiple Sensors. CASBliP project. http://casblipdif.webs.upv.es

[7] Bogdanov, I., Tiponut, V., Mirsu, R., New Achievements in Assisted Movement of Visually Impaired in Outdoor Environments. WSEAS TRANSACTIONS on CIRCUITS and SYSTEMS, Issue 9, Volume 8, September 2009

[8] Tiponut, V., Haraszy, Z., Ianchis, D., Lie, I. Acoustic Virtual Reality Performing Man-machine Interfacing of the Blind. 12th WSEAS International Conference on SYSTEMS, Heraklion, Greece, July 22-24, 2008

[9] Jun Song, H., Beilharz, K. Spatialization and Timbre for Effective Auditory Graphing. Proceedings of the 8th WSEAS International Conference on Acoustics & Music: Theory & Applications, Vancouver, Canada, June 19-21, 2007

[10] Rajamäki, J., Viinikainen, P., Tuomisto, J., Sederholm, T. and Säämänen, M.. LaureaPOP Indoor Navigation Service for the Visually Impaired in a WLAN Environment. Proceedings of the 6th WSEAS Int. Conf. on Electronics, Hardware, Wireless and Optical Communications, Corfu Island, Greece, February 16-19, 2007

[11] Rodríguez-Ramos L.F., Chulani H.M., Díaz-Saco L., Sosa N., Rodríguez-Hernández A., González Mora J.L. "Image And Sound Processing For The Creation Of A Virtual Acoustic Space For The Blind People". Signal Processing and Communications, 472-475, 1997.

[12] González Mora, J.L., Rodríguez-Hernández, A., Rodríguez-Ramos, L.F., Díaz-Saco, L., Sosa, N. "Development of a new space perception system for blind people, based on the creation of a virtual acoustic space". Proceedings of the International Work Conference on Artificial and Natural networks. Vol. 2, pp 321-330. Springer. 1999

[13] Loomis, J. et al. "Navigation without Vision: Basic and Applied Research". Optometry and Vision Science, Vol. 78, No. 5, May 2001.

[14] Capelle, C.H. et al. "A real time experimental prototype for enhancement of vision

rehabilitation using auditory substitution". IEEE Transactions on Biomedical Engineering. Vol. 45, nº 10, pp 1279-1293. Oct. 1998.