Sensory Substitution for Visually Disabled People: Computer Solutions

ANTONIO FRANCISCO RODRÍGUEZ-HERNÁNDEZ[1], CARLOS MERINO[1], OSCAR CASANOVA[1], CRISTIÁN MODRONO[1], MIGUEL ÁNGEL TORRES[1], RAQUEL MONTSERRAT[1], GORKA NAVARRETE[1], ENRIQUE BURUNAT[2] and JOSÉ LUIS GONZÁLEZ-MORA[1].

Department of Physiology[1] and
Department of Psycobiology[2]
University of La Laguna
S. Cristóbal de la Laguna, Tenerife, The Canary Islands
SPAIN
afrguez@ull.es http://www.nf.ull.es

Abstract: - Sensory substitution can be defined as a technical-scientific discipline which aims to provide sensory disabled people with information they cannot acquire from the disabled sense through their intact senses. We present here our team's work in this R+D line for providing blind and severe visually impaired people with real time spatial and text environmental information through sounds. The objective is to model the real environment as a virtual space where the object's surface appears as if covered by small sound sources, which emit very locatable sounds in a continuous and near simultaneous way. It is based on the hypothesis that the brain, when provided with this highly rich spatial information, will generate a kind of visual-like perception of the surrounding world. In this paper we describe our approach to this field and the main results obtained, which have practical consequences in the field of sensory rehabilitation as well as on the theory of perception.

Key-Words: Sensory substitution; Blindness; Sonification; HRTF; Computer vision; Brain plasticity.

1 INTRODUCTION

Once particular surrounding information has been identified as relevant for the sensory disabled persons, the sensory substitution approach firstly focuses on finding the optimal way of representing that information through the person's remaining senses and subsequently implementing it [1]. Applications of this concept range from traditional substitution methods like the long cane and Braille and Sign languages to the most recent developments based on high technology for acquiring and presenting the information of interest. In our approach, we use three technologies: computer vision, virtual reality and 3D sound.

Sensory substitution relies on the fact that environmental information is on many occasions available for a person throughout different sensory modalities. This is particularly certain for the spatial information. So, for example, the egocentric location of an object can be known by vision, audition, touch and even olfaction and the sense of temperature. This suggests that the brain may manage spatial information in an amodal way, that is, independently of the sensory modality that provides the information [2].

Several neurophysiologic studies support this notion. So, in the inferior Colliculus Nucleus of the Barn Owl, an early processing station at the auditory pathway of this predator species with an accredited ability for detecting and localizing relevant sounds, it has been found a topographic distribution of the auditory neurons sensitive to the location of the sound [3]. This distribution, although yet not found in higher cortical levels nor in humans, reminds the retinotopic organization of the visual pathway, where contiguous neuronal areas process contiguous areas of the perceptual field, showing that a very important feature for the visual spatial perception can be also developed for the auditory sense, in the form of a kind of "auditory retina"

The Posterior Parietal Cortex of the brain contains different areas. These areas are the 7ª area, the Lateral Intraparietal area (LIP), the Temporal Medial Superior area (MST), the 7b area and the Intraparietal Ventral area. In particular, the LIP area receives a considerable amount of projections from visual areas [4, 5], what justifies the notion of the LIP as the "parietal visual field". Nevertheless, the LIP area has been later described as a receiver region of acoustic information which, in conjunction with the visual and the somatosensory one, contributes to generate the representational map of the three-dimensional space [6]. In this sense, it has been shown that the auditory response of the LIP neurons has the same preferential directionality than the visual response, what suggests that both sensory receptive fields, the auditory and the visual one, together with their respective sensory memory fields, overlap each other [7]. These multiple sensory inputs (visual, auditory, somatosensory and vestibular sense) to the LIP area are combined in a process of signal maximization for the coding of the spatial coordinates, what forms the basis of the surrounding spatial representations. In addition, these spatial representations at the posterior parietal cortex are related with high level neuronal cognitive activities, including attention.

Regarding sensory substitution, the cue point is to know whether the information required to carry out a particular perceptual task can be or not provided by one or more intact senses. In this sense, we have developed a series of prototypes for blind people's orientation, mobility and environmental perception, which provide the user with two types of real time information through sounds: information on the spatial volume occupied by the objects and surfaces located in front of him, and information on the written text present in the frontal scene (i.e., shop signs, advertisings, etc). The volume information is translated into a special sound code which is delivered through headphones in order to generate an auditory spatial representation coherent with the environment. The text information is presented as verbal spatial sound, as if a reading voice was coming from the area where the text is located. Then we follow a psychoacoustical approach in order to evaluate the users' perceptual response to the vision inspired acoustic stimulation. This work is complemented with the study of the disabled people's neurological substrate of the sensory substitution experience, through brain function registering techniques such as functional Magnetic Resonance Imaging (fMRI) and Event Related Potentials (ERP).

2 APPROACH

This section explains our approach to the problem of sensory substitution in blindness and severe visual impairment. We first pose our hypothesis regarding what particular information from the visual scene should, and then could, be acoustically provided to the listener so he will experience an auditory visual-like image of such scene. Next, the key methodological steps and the main results will be summarized. Finally we outline the current state of the on real time environmental text reader system.

2.1. The auditory code for the visual scene

A sighted person perceives images as series of light rays coming from every point of objects and surfaces inside his field of view. In a similar way, a person can also obtain a tactile sequential representation of the objects by touching them coordinate by coordinate. The perceived spatial image of an object comes from the acquisition of very significant spatial information, i.e., the subject centered spatial coordinates which are occupied by the objects. Following on from this idea, the following hypothesis can be posed: a blind person is exposed to sound rays radiating from an object's surface in such a way that his perceptual system can recover the whole set of relative spatial coordinates involved, similar to what happens in vision from perceiving light, will he be able to perceive some kind of crude visual-like image, similar perhaps to a 3D visual image containing mainly low spatial frequencies?

The literature reports several devices which offer auditory spatial information for the blind person's orientation and mobility (see [8] for a review on the earlier developments).These devices mainly provide auditory information that indicates the presence and location of the detected objects in order to avoid or to use them as landmarks when navigating. Dr. Kay's KASPA system is based on an ultrasonic sensor that calculates the obstacle distance from the time of flight of the emitted ultrasonic wave. This distance information is translated into a sound code which consists on a progressive change in the pitch of a pure tone as changes the detected distance. His system has later evolved to a version that includes wide-angle overlapping peripheral fields of view with a narrow central field superposed, what improves the blind subjects' auditory ability to resolve close objects [9]. Dr. Jack Loomis and his colleagues from UCSB were who first applied binaural technology for both representing the location of environmental landmarks and subsequently guiding the user's steps towards them. In this case, GPS information is translated into spatial verbal and non-verbal sound indications [10]. More recently, Dr. Tiponut and his team from the Polytechnic University of Timisoara have developed an integrated multisensory device which provides both information on the location of obstacles and a pilot signal to indicate the direction of the movement to a target. This information is coded as spatial sound obtained from binaural technology [11].

Some teams have also explored the human's ability to recognize an object's shape from both sound and tactile cues. Dr. Meijer's The Voice system directly translates the position and level of gray of the pixels composing the image from a video camera into a sound code where the vertical dimension is represented by the pitch of a pure tone, the horizontal dimension by both binaural cues and the presentation time, and the level of gray by the amplitude of the sound [12]. Dr. Capelle and colleagues from Catholic U. of Louvain have explored a similar shape codification by also attaching a pure tone to every pixel of the image from a video camera, although in this case the sound representing any activated pixel is emitted in a continuous way and only black and white levels of gray are considered [13]. Dr. Lakatos found that normally sighted subjects show considerable ability in recognizing alphanumeric characters whose patterns are outlined acoustically through the sequential activation of specific units in a speaker array[14]. Dr Hong and Dr. Beilharz, from the University of Sidney, find that the shape of two concurrent graphic lines can be gathered from an auditory representation based on mapping the x-axis to time and the y-axis to MIDI notes. The performance is improved when the concurrent audio streams are presented as independent separated virtual sound sources [15].

Regarding the tactile devices, it is mandatory to cite the work of Dr. Bach-y-Rita and colleagues who, in the 60s of the past century, introduced the concept of sensory substitution and explored the human ability to perceive object's shape and width from a bi-dimensional tactile projection of the scene onto the user's skin [16]. The image of single objects from a CCD camera is directly translated into a spatial pattern of vibro-tactile effectors which stimulate the corresponding coordinates on the user's skin. Later, their work derived toward the development of an electro-tactile interface placed on the tongue [17]. Drs. Segond, Weiss and Sampaio have recently explored its possibilities for perceiving shape and spatial cues for navigation [18].

The studies above suggest that spatial hearing, like vision and touch, has access to the mechanisms that give rise to the amodal spatial representations involved in the perception of the shape and width of the objects, that is, in the figure perception. In order to carry this argument one step further, we have explored what the perceptual effect of coding the environment with 3D sound is, i.e., coding the environment with sounds which are perceived coming from every occupied space coordinate.

In this vein, we have focused on whether an image of a unitary whole object, broadened in the space as in the visual experience, can be generated in the blind person from the appropriate auditory stimulus [19, 20, 21]. Figure 1 shows this approach: a sighted person sees the frame of a window as light rays coming from it. Then, a series of small loudspeakers are located occupying the same space location of the frame. We would expect the appropriate emission from that spatial configuration of loudspeakers to the blind person (third picture) to experience a spatial image of the frame which is spatially similar to the visual one.

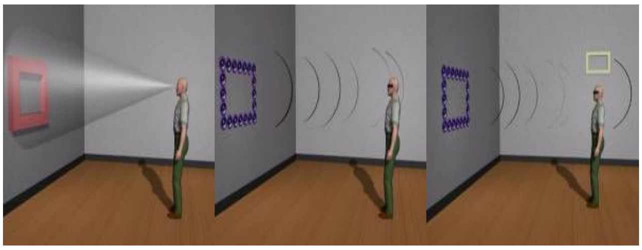

Fig.1. Visual to auditory sensory substitution

According to this scheme, we first explored blind people's ability to experience visual-like images from spatial patterns of real sound sources [19]. Figure 2 shows the experimental set-up consisting on a 6x6 array of small loudspeakers facing a point where the subject is placed to receive the emitted sound. Every loudspeaker is conveniently directed to the point where the subject's head will be placed, given that otherwise part of the higher spectral content of the sound would not reach the subject's ears, distorting the perceived elevation of the sound sources.

Fig.2. Loudspeakers array inside the acoustically prepared experimental room

These experiments were conducted with both blind and normally sighted people in an acoustically optimized environment. The set of loudspeakers configuring a particular spatial pattern is sequentially activated in order to test the subject's ability to recognize it as well as the spatial audio image being experienced. The presented figures are composed from, for example, the top and the bottom rows (two parallel horizontal lines), the left and the right columns (two parallel vertical lines), a C letter-form constructed with the left column and the top and bottom rows, the frame of a window, etc. The tests show that blind people can recognize the presented spatial patterns and clearly note its physical distribution, whilst referring that a kind of auditory spatial image extended in the space with the shape of the presented spatial pattern can be perceived.

Afterwards, a first augmented reality prototype was developed. This is a non portable laboratory prototype which first obtains the information of the spatial coordinates occupied by the objects in the scene, and then generates an auditory stimulus representing this information which is delivered to the user through headphones. This auditory stimulus is such that, in spite of being delivered through headphones, creates the illusion in the user that the previously detected object is covered by small emitting sound sources. This effect is obtained by combining computer vision techniques [22] for recognizing the environment with 3D sound techniques based on HRTF filtering (Head Related Transfer Function) [23, 24, 25], for creating the illusion of sound "externalization". This initial version sums up the philosophy of the subsequent prototypes, a brief description and the main results obtained are as follows:

2.1.1 The first device, Virtual Acoustic Space I: the validation of the idea.

VAS I consists of two subsystems: the first one for acquisition and analysis of the scene (visual subsystem), and the second one for conversion of the information into sounds and playing them back to the subjects (acoustic subsystem)[19].

Figure 2 shows a conceptual diagram of the technical solution we have chosen for the prototype development.

Two miniature cameras are fixed on both sides of a pair of conventional spectacles, which will be worn by the blind person using the system. Different computer vision algorithms are applied to the captured images, such as the detection of geometric features or stereovision, in order to obtain a depth map.

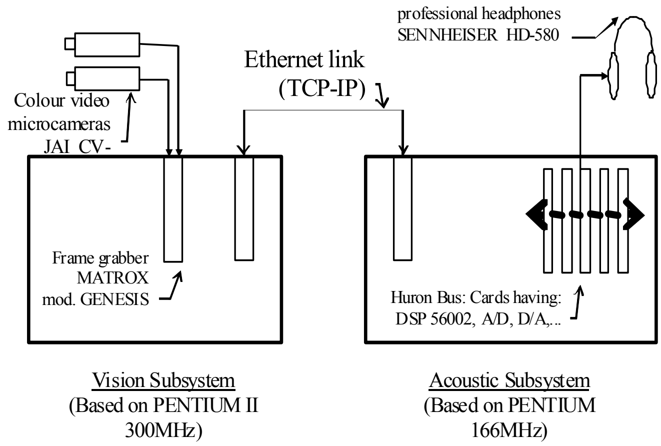

Fig.3. Prototype conceptual diagram

The acoustic subsystem then plays a random sequence of short sounds, one for each position provided in the depth map. Each sound has been previously "spatialized" so, in spite of being heard by headphones, it seems to come from a certain position in the environment.

The virtual sound generator uses the Head Related Transfer Function (HRTF) technique in order to obtain the spatial sounds [24]. For each position in space a set of two HRTFs are needed, one for each ear, so that the interaural time and intensity difference cues, together with the behavior of the outer ear are taken into account. In our case we are using a reverberating environment, so the measured impulse responses also include the information related to the echoes in the room. Individual HRTF's are measured as the responses of miniature microphones (placed at the entrance of the auditory channel) to a special measurement signal (MLS) [26]. Also the transfer function of the headphones is measured in the same way, in order to equalize its contribution.

Having measured those two functions, the HRTF and the Headphone Equalizing Data, properly selected or designed sounds can be filtered and presented to both ears, obtaining the same perception as if the sound source were placed in the position from where the HRTF was measured.

The sound selected for encoding the object's coordinates is a very short click without tonal quality. This type of sound is easily locatable in space and, given its short duration, it makes it possible to present a high number of coordinates within a short period of time. The perceptual effect of this stimulus could be described as hearing a large number of raindrops hitting the surface of a glass window. A field of view of 80° on the horizontal axis by 45° on the vertical axis is divided into in a number of x (horizontal), y (height) and z (distance) coordinates or stereopixels.

2.1.2. Results summary

The initial results were obtained in a controlled environment from a broad group of blind and visual impaired people [27], and they have been confirmed by multiple tests subsequently carried out with the more advanced versions of the device. These results support the hypothesis that, when using an auditory stimulus, as previously described, to represent large objects in a scene, it is possible to generate a perceptual experience in the user of a global and maintained presence of those different objects inside the field of view, which are perceived as occupying the space with their gross shape, dimensions and location. For example, two walls surrounding a path are perceived as sounding objects which are always present on both sides of the subject, with their vertical and depth dimensions. A central soundless space can be perceived, and the blind person can walk or "look" through it (Fig.2.).

Fig.4. A participant signals the limits of the auditory spatial image corresponding to a hole in the wall

In one of the tasks the blind person is asked to point to the figure and "draw it in the air" by moving the arm along the perceived extension of the sound image (Fig. 5).

Fig.5. The participant is asked to move her arm through the area where she is perceiving that the sound seems to be extended through

The arm is in a straight position and the person holds in the hand a magnetic sensor for registering the coordinates of the arm movement. In addition, a verbal description of the perceived extension of the sound image is collected for every figure.

Blind people can discriminate a line of sounds from a single point of sound, introducing the concept of persistent broadened sound. He or she can also recognize the horizontal, vertical or diagonal layout of this audible line by perceiving the extension of the spatial image of the line (Fig.6).

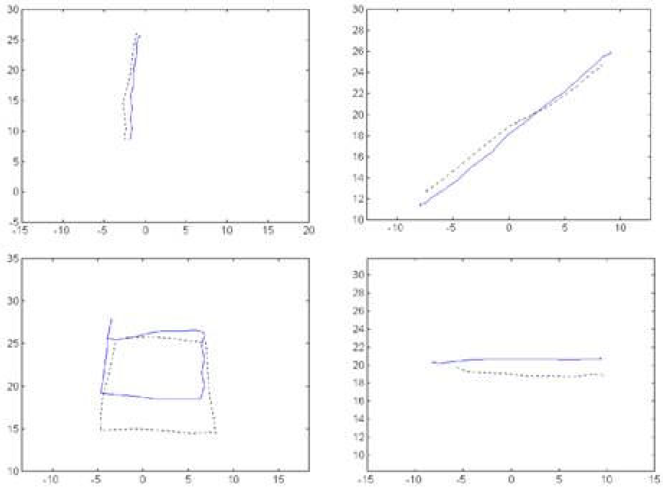

Fig.6. Register of the hand signaling of different figures from visual (continuous line) and auditory (dotted line) information (scale in inches)

These results are especially important in relation with the question of whether an auditory perception of objects similar to the visual one can be experienced or not. The scientific study of the spatial aspects of the auditory perception has mostly focused on the perception of isolated sound sources. This may be in relation with the fact that sounds are commonly perceived as coming from concrete isolated locations in the space. Nevertheless, sound sources are not single points but have a width. A spatial attribute of the spatial auditory image regarding to the "width of the sound" has been described, namely the sound broadness. Several acoustic parameters involved in this effect have been studied [28]. It is worth pointing out the references in these studies to the auditory spatial image that is experienced when hearing a multitude of close sound sources which are simultaneously radiating their respective sounds, for example, the rustling leaves and branches of a tree. Similarly, blind people who is asked about this question usually refers an auditory experience of sound broadness when perceiving the raindrops or the wind hitting diverse objects or surfaces in their surroundings. Somehow our approach operates on these considerations and the results support the notion of auditory figure perception.

Finally, it is worth pointing out the experience of a blind person who was able to detect and identify the different objects and surfaces presented to her when using the prototype inside an experimental room and without any previous knowledge of this environment (walls, a column, a window, the door and a small table). She was able to move between them, and to make a correct verbal and graphical description of the room and the relative position of every object and surface.

As regards the neurological substrate of this perceptual activity, preliminary results using the functional Magnetic Resonance Imaging technique (fMRI) have shown that spatial sound processing in blind people occurs more in occipital cortical brain regions than in sighted people [29]. This suggests that blind people recruits the brain's visual areas for spatial sound processing, which has important consequences when considering the blind person's potential ability to use sensory substitution devices. Many other studies support also the notion of brain intermodal sensory plasticity (see for example [30]).

2.1.3 Recent work

The latest version of the prototype has been developed within a recent EUFP6 project whose name is CASBliP (Cognitive Aid System for Blind People-www.casblip.com). The CASBliP device acquires environmental information from two different independent subsystems: 1) a time-of-flight infrared sensor placed on the frame of a pair of glasses which acquires distance information from the objects in the frontal scene inside a range of 0.5m to 5m for a horizontal plane at eye level; and 2) a pair of cameras placed at the top of a helmet. Segmentation and shape identification algorithms enable us to detect a moving object in a range of 5m to 15m. This information is presented in an auditory way basically according to the representation strategy outlined above. A series of tests have been conducted which show that a blind person can perform orientation and mobility tasks with a progressive improvement with learning (Fig.7).

Here it is shown the performance (measured in seconds by meter) of ten blind persons who were asked to navigate, relying only on the sensory substitution portable system, through a 14 m long route, i.e. a path with 4 pairs of soft obstacles of 180 cm height put up asymmetrically (Fig.8).

The participants had to locate every pair of columns (1 m of separation between them) and detect the gap through which they might move without touching or knocking the objects over, and then go on to the next pair of columns.

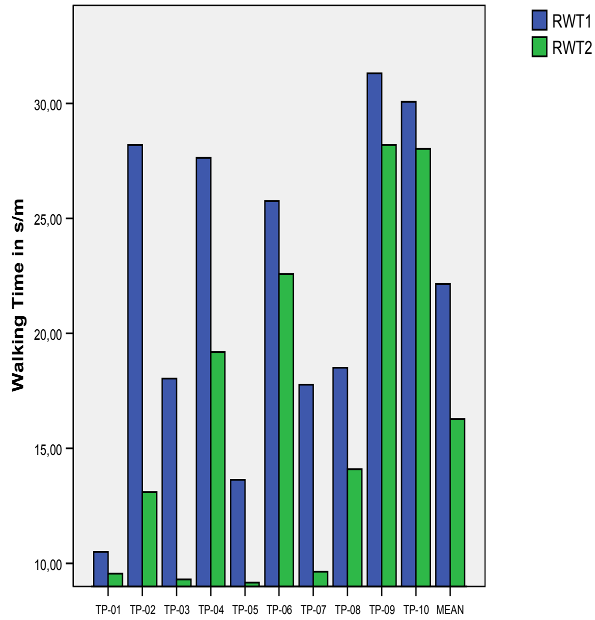

Fig.7. Relative Walking Time (s/m) of ten blind persons who were asked to navigate through the obstacles shown in Fig.5. In blue the results of the first run test (RWT1) and in green the results of a second run test (RWT2) after a short period of intensive training

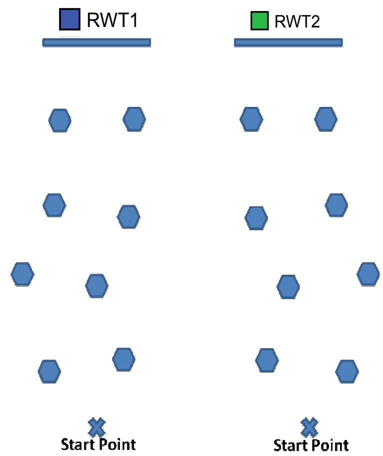

Fig.8. 14 m long route with 4 pairs of soft obstacles of 180 cm height put up asymmetrically. A flat surface indicates to the participant where the end of the course is. The differences between both arrangements try to avoid a possible influence of memory on the performance of the orientation and mobility task

The subjects had previously spent three sessions in order to become familiarized with the substitution stimulus, what required a mean time per subject of 75 min. After the first run test persons spent some time on the site with the mobility instructor to do training before they commenced a second test run, which one was carried out on a different arrangement of the obstacles. Relative Walking Time (RWT) scores from the post-training session show that results had significantly improved (t=4.36; p=0.002) [31].

A virtual reality simulator for blind and visually impaired people has been developed. It is called Virtual Reality Simulator for Sonification Studies, or VRS3 [32], and provides the user with a spatial auditory representation of the virtual environment previously designed. A 3D tracking system locates user's head orientation and position, so the user can "walk through" the virtual environment while he or she perceives the environment through auditory information (Fig.9).

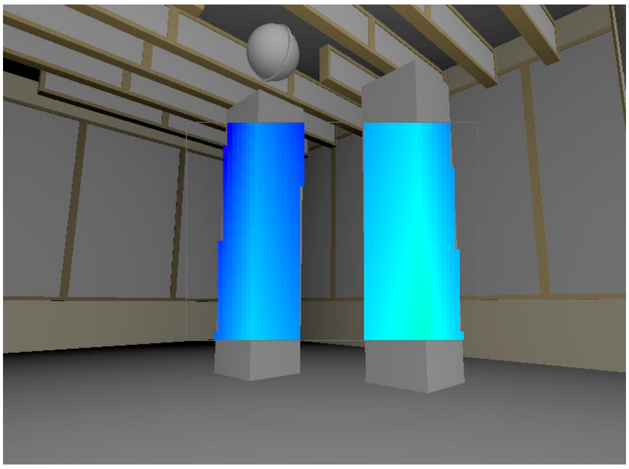

Fig.9. The depth map of two virtual columns inside the simulator room is overlapped to the scene and represents a 64x48 pixels size map with a pseudo-color depth scale

The simulator has these main purposes: validation of sonification techniques, 3d sensor emulation for environment recognition and hardware integration; also for training and auditory perception experiments. This simulator can recreate any simple or complex scene and present it to the user as a 3D sound world. Then, it allows the researcher to surpass the need of a "sensor system" for studying the perception of the auditory representation of the scenes. Concretely, we use it for defining the representation strategy, that is, the way the scene information is coupled to a sound code. Then the studies are oriented to get a better understanding on the perceptual effects of several significant acoustic parameters, as the interclick time interval, the sound reverberation level or the tonal colour of the click sound.

We also study the effect of training on the quality of the auditory spatial image of the scene experienced by the user, which implies some research on learning protocols of sensory perception as well as researching on individual or group differences (by sex, age, and so). In this way we have obtained preliminary results showing some advantages of more complexes (spectrally rich) sounds for distance localization using a set of real sound sources located from 50 cm till 6 meters in front of the subject (submitted for publication).

A robotic system has been developed that allows intensive measurement of both human and mannequin HRTFs in every spatial axe inside an 8 x 4 x 4 m width acoustically isolated room, with a spatial resolution up to 1º. This system allows getting massive sets of spatial filters from both subjects and mannequins. We are currently studying the effects of training on the precision of auditory localization when using both individual and generic virtual spatial sounds.

2.2 Development of a text reading system

The objective of this research line is to develop a scene text reading system for blind people. It is widely accepted that Optical Character Recognition (OCR) for scanned documents is no longer a problem. There are several commercial and open source OCR engines available, with recognition rates of over 95% for clean, scanned documents. Text recognition of scene text extracted from a video camera is a much harder problem and remains largely unsolved. There has been great interest in recent years in this field among research groups all around the world. Some applications are automatic indexing and cataloging of video libraries, road sign driver assistance, mobile phone document scanning, or visually impaired assistance systems, etc. Advances in digital cameras, computing power and modern computer vision techniques are making real-time text extraction and document processing from video images and its application on blind people assistance possible.

A system is being developed that detects, segments and tracks scene text such as shop signs, traffic signs, advertisements and billboards in nearly real-time (Fig.10).

For demonstration purposes, a simple communication module with an OCR engine and a voice synthesizer were integrated into the system [33].

Fig.10. Example of a text both detected and recognized by the portable version of the reading system (green letters inside the white box)

3 CONCLUSION

We present in this paper our R+D line in the field of sensory substitution for providing blind and visually impaired people with relevant visual environmental information through sounds. We mainly focus on exploring the hypothesis that perception of objects' spatial attributes like shape, width and location, can be experienced by the blind users in a gross visual-like way through hearing, whenever the appropriate acoustic representation of those spatial features is provided. Our results support this idea and encourage to going on in order to define optimal acoustic representations of the real scene information. In addition, a series of progressively more sophisticated prototypes has been developed with the aim of obtaining a portable device susceptible to be added to the existing arsenal of rehabilitation aids for orientation, mobility and perception of the environment.

Our group is currently continuing on with the improvement of the above-mentioned prototypes. We aim to develop an integrated portable prototype capable of acquiring, from the user's frontal scene, a robust 3D depth map segmented into objects, distant text information and both the identification and position of selected items, to immediately deliver this information as an adequate auditory representation based on spatial sound. The development of computer vision algorithms for video image segmentation, detection and labeling of the environment will enrich the depth map information provided by the 3D sensor. Concomitantly we study how different acoustic parameters affect the user's spatial auditory image of the scene, in order to optimize his or her auditory representation of it. The question of using individual versus generic or semi-personal HRTFS in order to achieve an appropriate spatial sound perception is still unsolved. In this sense, we are currently addressing the role of learning on the calibration of the auditory system to a non individual collection of spatial sounds.

Furthermore, the fact that blind people occasionally perceive spots of lights located at the spatial location of suddenly presented noises (which were reported as phosphenes in the decade of the 70's of the last century [34,35]) points to the fact that the brain can mixture sound and vision in a unknown way. Then we are currently researching to elucidate the neurological substratum of the phosphene phenomenon (visual perception elicited by sound stimulation), and preliminary results have already been reported [36].

We feel that the sensory substitution approach, when supported by the advances in high technology and a progressively better knowledge of the human brain's perceptual capabilities, opens up a wide field of applications in sensory rehabilitation.

References:

[1] Bach-y-Rita, P. "Brain mechanisms in sensory substitution". New York: Academic Press.

[2] Loomis, J. "Sensory replacement and sensory substitution: overview and prospects for the future". Converging technologies for improving human performance. Edited by Mihail C. Roco and William Sims Bainbridge. Kubler Academic Publisher, 2003.

[3] Takahashi, T. T., Keeler, C.H. "Representation of multiple sound sources in the owl´s auditory space map". Journal of Neuroscience. Vol 14, No 8, 4780-4793. August, 1994.

[4] Lynch J.C., Graybiel A.M. Lobeck L.J. "The differential projection of two cytoarchitecture subregions of the inferior parietal lobule of macaque upon the deep layers of the superior colliculus". J. Comp. Neurol. 235:241-254. 1985.

[5] Blatt G.J. Andersen R.A. Stoner G.R. "Visual receptive-field organization and cortico-cortical connections of the lateral intraparietal area (lip) in macaque". J. Comp Neurol. 299: 241-245. 1990.

[6] Mazzani P., Bracewell, R.M., Barash, S., Andersen, R.A. "Spatially tuned auditory responses in area LIP in macaques performing

delayed memory saccades to acoustic targets". J. Neurophysiol. 75: 1233-41. 1996.

[7] Andersen, R. A. "Multimodal representation of space in the posterior parietal cortex and its use in planning movements". Annual Review Neuroscience. Vol 20, 303-330. 1997.

[8] Kay, L. "Electronics aids for blind persons: an interdisciplinary subject". IEE review. IEE Proceedings, Vol. 131, Pt. A, No. 7, 559-576. September 1984.

[9] Kay L. Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar. J Acoust Soc Am. 2000 Jun;107(6):3266-75.

[10] Loomis, JM. Klatzky RL, Golledge RG. "Navigation without Vision: Basic and Applied Research". Optometry and Vision Science, Vol. 78, No. 5, May 2001.

[11] Virgil Tiponut, Zoltan Haraszy, Daniel Ianchis, Ioan Lie. "Acoustic Virtual Reality Performing Man-machine Interfacing of the Blind". 12th WSEAS International Conference on SYSTEMS, pp 345-349, Heraklion, Greece, July 22-24, 2008.

[12] Meijer, P.B.L. "An experimental system for auditory images representations". IEEE Transactions on Biomedical Engineering, 39; 112-121. 1992.

[13] 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.

[14] Lakatos, S. Recognition of complex auditory spatial patterns. Perception. Vol. 22. 363-374. 1993.

[15] Hong Jun Song, Kirsty Beilharz. "Spatialization and Timbre for Effective Auditory Graphing". Proceedings of the 8th WSEAS International Conference on Acoustics & Music: Theory & Applications, pp 18-26, Vancouver, Canada, June 19-21, 2007.

[16] Bach-Y-Rita, P. et al. "Visual substitution by tactile image projection". Nature, 1221, 963-964. 1969.

[17] Bach-y-Rita, P., Kaczmarek, K.A., Tyler, ME, García-Lara, J. "Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note". J Rehabil Res Dev. Oct; 35(4):427-30. 1998.

[18] Segond, H., Weiss, D., Sampaio, E. "Human spatial navigation via a visuo-tactile sensory substitution system". Perception. 2005; 34(10):1231-49.

[19] 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.

[20] 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.

[21] Rodríguez Hernández, A; Sosa, N, Rodríguez Ramos, L.F. , Chulani, H. Díaz Saco, L, y González-Mora, J.L.. "Percepción del entorno en personas ciegas a través de un estímulo sonoro espacial virtual generado por computador". Actas del Congreso Iberoamericano 3º de CAA, 1º de Tecnologías de Apoyo para la Discapacidad, Oct. 2000. pp 81-84. Ramón Cerés y Comité Organizador Editores.

[22] Faugeras. "Three dimensional computer vision. a geometric viewpoint". The MIT Press. 1993.

[23] Begault, D. R. "3-D sound for virtual reality and multimedia". 1994. AP Professional.

[24] Wightman, F., Kistler, D.J. "Headphone simulation of free field listening I: stimulus synthesis". Journal Acoustical Society of America. V. 85 858-867. 1989.

[25] Kenneth John Faller II, Armando Barreto, Navarun Gupta And Naphtali Rishe. "Enhanced modeling of head-related impulse responses towards the development of customizable sound spatialization". 4th WSEAS Int. Conf. on Computational Intelligence, Man-Machine Systems and Cybernetics, pp 82-87, Miami, Florida, USA, November 17-19, 2005.

[26] Rife, D., Vanderkooy, J. Tranfer-function measurement with maximum-length sequences. J. Audio Eng. Soc., Vol. 37, No. 6, 1989.

[27] González-Mora, J.L., Rodríguez-Hernández, A.F., Burunat E., Chulani, H.M, Albaladejo, J.C., Rodríguez-Ramos, L.F. "Seeing the world by hearing: Virtual Acoustic Space (VAS) a new space perception system for blind people." Ballesteros, S., & Heller, M. A. (Eds.). pp 371-383. UNED. 2004.

[28] Blauert, J., Lindemann, W. "Auditory spaciousness: some further psychoacoustic analyses".J. Acoust. Soc. Am. Aug; 80(2):533-42. 1986.

[29] Rodríguez Hernández, A.F., González Mora, J.L., Pujol, J., López García, J.A.,Muñiz Montes, J.R., Rodríguez Ramos, L.F., Chulani, H.M. Burunat, E. "Functional magnetic resonance imaging of spatial auditory perception in blind

and sighted human subjects". J Physiol. 548P O76. 2003.

[30] Sadato, N. Pascual-Leone, A. Grafman, J., Ibáñez, V., Daiber, M.P., Dold, G., Hallett, M. "Activation of primary visual cortex by braille reading in blind people". Nature. 380,526-527. 1996.

[31] Hans Kaltwasser. Internal Document CASBliP Project. Deutscher Blinden-und Sehbehinder-tenverband e.V. (DBSV), 2008.

[32] Miguel Ángel Torres Gil, Óscar Casanova González, and José Luis González-Mora. "Virtual reality simulator for sonification studies". Recent Advances in Computational Intelligence, Man-Machine Systems and Cybernetics, Proceedings of the 8th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics (CIMMACS '09), pp 112-116. El Puerto de la Cruz, Tenerife, December 14-16, 2009.

[33] Merino, C., Mirmehdi, M. "A Framework Towards Realtime Detection and Tracking of Text". In: Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2007), pp10–17. Sept 2007.

[34] Lessell, S. and Cohen, M.M. "Phosphenes induced by sounds". Neurology, 29(11): 1524-1526. 1979.

[35] Page, N.G.R., Bolger, J.P., Sanders, M.D."Auditory evoked phosphenes in optic nerve disease". Journal of Neurology, Neurosurgery and Psychiatry. V. 45, 7-12. 1982.

[36] Burunat E., Rodríguez-Hernández A. F., López-García J. A., Muñiz-Montes J. R., Pujol J., Rodríguez-Ramos L. F., Chulani H.M., González-Mora J. L. "Blind woman experiencing auditory evoked phosphenes: A functional magnetic resonance imaging (fMRI) study". FENS Abstr., A224.3, vol.2, 2004.

Acknowledgements

This R+D line has been carried out with funds from the Spanish MCyT, MCeI and MITyC, the Govern of Canarias, the Spanish National Blind Organization (ONCE), EU-FEDER Program, EUFP6-IST, Agencia Canaria de Investigación, Innovación y Sociedad de la Información (ULLAPD-08/01), Ministerio de Industria, Turismo y Comercio (Proyecto AVANZA), TSI-020100-2008-337; Ministerio de Ciencia e Innovación (Proyecto CICYT), TIN2008-06867-C02-01/TIN. We express our gratitude to Hans Kaltwasser, from DBSV, for sharing partial results of our joint work in the CASBliP Project and to Manuel Mauricio, from ULL, for the infographics of figure 1.