



Evaluation of a New Mobility Assistive Product for the Visually Impaired

Pablo Revuelta Sanz, Belén Ruiz Mezcuca, José M. Sánchez Pena, and Bruce N. Walker

Carlos III U. of Madrid, Av. Universidad 30, 28911, Leganés, Spain.

prevuelt@ing.uc3m.es

Carlos III U. of Madrid, Av. Universidad 30, 28911, Leganés, Spain.

bruiz@inf.uc3m.es

Carlos III U. of Madrid, Av. Universidad 30, 28911, Leganés, Spain.

jmpena@ing.uc3m.es

School of Psychology, GeorgiaTech, 654 Cherry St., 30332-0170, Atlanta, GA, U.S.A.

bruce.walker@psych.gatech.edu

ABSTRACT

In this study we present the results of evaluating the sonification protocol of a new assistive product aiming to help the visually impaired move more safely with the help of sounds organized in different cognitive profiles. The evaluation was carried out with 17 sighted and 11 visually impaired participants. The experiment was designed over both virtual and real environments and divided into 3 virtual reality based tests. Finally, four participants became experts by means of longer and deeper training, and they participated in a real life test and in a focus group at the end of the tests. Both quantitative and qualitative results were extracted, showing that the proposed system is able to help the users understand their surroundings via sounds. However, important limitations have been found in the sample used (some important demographic characteristics are strongly linked among them, limiting segregated analysis), the usability of the most complex profile, or even the special difficulties faced by the completely blind participants as compared to the sighted and low vision participants.

Indexing terms/Keywords

Assistive technology; blind; mobility; sonification; stereovision.

Academic Discipline And Sub-Disciplines

Engineering, Psychology, Experimental science.

SUBJECT CLASSIFICATION

Engineering Subject Classification;

TYPE (METHOD/APPROACH)

Experimental; Survey/Interview.

Council for Innovative Research

Peer Review Research Publishing System

Journal: INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol 13, No. 2

editor@cirworld.com

www.cirworld.com, www.ijctonline.com

INTRODUCTION

Mobility is to be “able to carry out a plan to get there” [1] and, thus, it is related to perception of the environment in which the navigation plan must be implemented. However, visually impaired people have extra problems compared to sighted individuals and, hence, they use different techniques and devices such as the white cane or the guide dog to get around.

With advances of technology, new assistive products have appeared, designed to help this heterogeneous group of users in their daily tasks. Among them, there is a range of assistive products specifically for orientation and mobility (O&M) [2]. These systems use various different strategies to convey the relevant information to the users, mainly via tactile and auditory interfaces. For a review of them, see [3].

Sonification is the use of sound to represent data. We can find many types of sonification, such as text-to-speech programs (converting text into audible speech), color readers (color into synthetic voice), Geiger counters (radioactivity into clicks), acoustic radars, MIDI synthesizers, etc. Sonification is, finally, a translation between two essentially different sets of data. It is a very active field, mostly focused on accessibility, but also on aesthetic and artistic goals, psychoacoustic research, improvement in data analysis or high-tech commercial devices, among others. Sound has been widely used in the assistive technology field to replace visual information, with plenty of documented success [4].

In this paper, we will evaluate the sonification protocol for a vision-to-sound system (described in [5]) designed as a mobility aid for blind users with an architecture shown in figure 1.

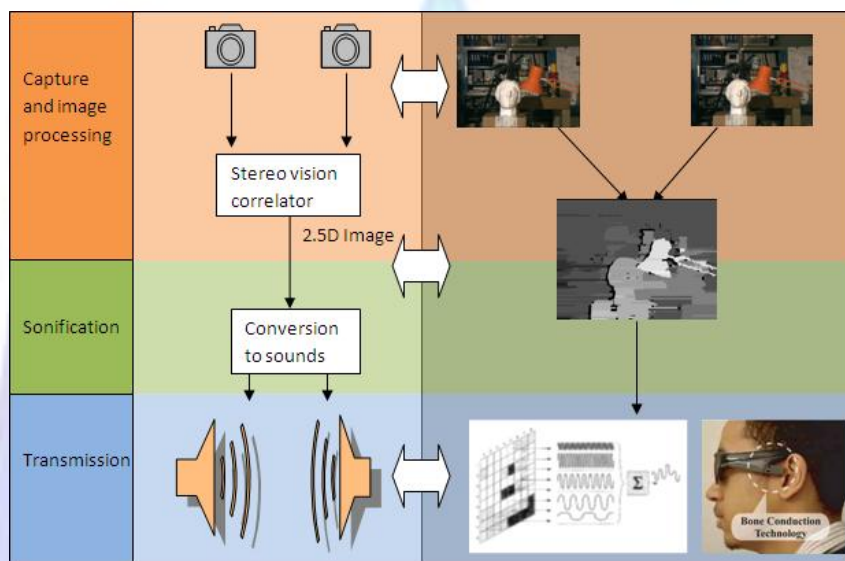


Fig 1: Block diagram of the system under test.

Although the complete system supports 7 different profiles of sonification, we tested only the 4 more complex ones. The complete set of sounds used in this evaluation, and the three dimensions of the space (from the user's point of view) is summarized in Table 1. As it can be seen, the different profiles implement increasing complexity in the vertical representation of the scene, using the double of pitches of the previous level.

Table 1. Profiles description.

Profile	Depth	Vertical	Horizontal
3	Loudness+ Low Pass Filter (LPF)	2 pitches in 1 notes and 2 different octaves	8 columns, 90° cone width, stereo +vibrato
4	Loudness+ LPF	4 pitches in 1 notes and 2 different octaves	8 columns, 90° cone width, stereo +vibrato
5	Loudness+ LPF	8 pitches in 2 notes and 4 different octaves	8 columns, 90° cone width, stereo +vibrato
6	Loudness+ LPF	16 pitches in 4 notes and 4 different octaves	8 columns, 90° cone width, stereo +vibrato

EXPERIMENTAL VALIDATION

In the validating process, the protocol designed to help the visually impaired considered the following hypotheses:

- H1: The sonification protocol helps in achieving mobility objectives.
- H1.1: The sonification protocol helps in detecting obstacles.

- H2: There are factors that modulate the usability of the system.
- H2.1 The daily use of the computer (factor COMP) helps in the understanding of a new sonification protocol.
- H2.2: The educational level (EDU) is positively related with the ease of learning a new sonification protocol.
- H2.3: The age (AGE) is inversely related with the utility of the protocol for mobility.
- H2.4: Visually impaired (VI) people encounter more problems to understand new sonifications, following the results of [6].
- H2.5: The efficiencies of the profiles (LEVEL) are different.

There are, hence, five a priori independent factors to be analyzed: use of computer, educational level, age, visual impairment and profile.

Method

The mobility assistive product was tested in two different ways: (1) with a virtual reality environment (VRE), where only the sonification is evaluated; and (2) in a real environment (RE), with four experts (with longer training) and different soft but opaque obstacles.

Participants. In the VRE test, 17 undergraduate and graduate students from the Georgia Institute of Technology, plus 11 volunteers from the Center for the Visually Impaired (CVI) of Atlanta (Georgia) participated in this study. The sample included 11 males and 17 females, with a mean age of 33.5 years, ranged 18-62. Among them, 13 were sighted, 10 had low vision and 5 were completely blind. All reported normal or corrected-to-normal hearing. In the RE, there were four participants, with an average age of 22 years old, 3 female and one male. All reported normal vision and hearing.

Apparatus for VRE. The evaluation of the sonification protocol was done via a virtual reality environment, directly sonified and transmitted to the participants through earphones.

The virtual environment was developed with the Unity3D engine (<http://unity3d.com/>), connected through the IServer program to the InterSense InertiaCube2 head tracker (<http://www.intersense.com/pages/18/11/>). The Unity3D rendered images are sent to the sonification program, written in C, and connected to the V-Stack (http://www.steinberg.net/en/support/unsupported_products/vstack.html) and Edirol HQ Hyper Canvas Synth (<http://www.roland.com/products/en/HQ-GM2/>). This synthesizer allows processing General MIDI 2 signals (<http://www.midi.org/techspecs/gm.php>) produced by the sonification program, whose correlative sounds are transmitted to the user by means of a pair of earphones.

Tests with VRE. Three different setups were designed in Unity3D, including two simple labyrinths and a room with some objects. The forward and backward movements were done manually in the keyboard of the computer in fixed steps each time the participant said "go forward" or "go backward". The direction change was signalled by spinning around in a chair, and detected by the head tracker.

In the first test, the users were asked to follow along the wall to their right as fast as possible, without missing the reference. If they said "lost", they were manually taken to the last position to continue performing the task, and this fact was written down. This labyrinth could be completed in 42 steps.

The second labyrinth presented a similar structure, but some objects were placed into the virtual space; some of them were taller than the avatar controlled by the user (such as columns), some objects were shorter. In this second space, the users were asked to follow the wall, to report each time they think they are close to an object and to say "lost" whenever they did not know where the wall was. This labyrinth could be completed in 50 steps.

Both labyrinths are shown in Figure 2.

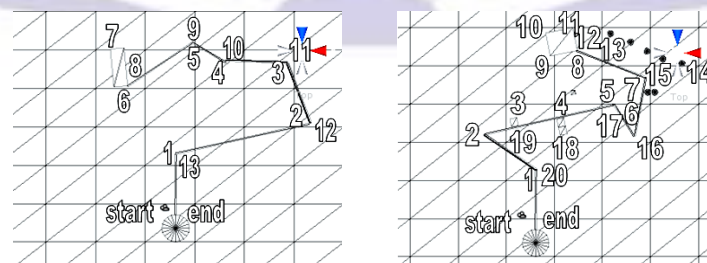


Fig 2: Top view of the labyrinths of test 1 (left) and test2 (right).

The labyrinths are divided into milestones which serve as check points. A time limit of 10 minutes was given to the users to perform the complete task. The obstacles in the second test were a column attached to the wall (milestone 3), a column at the left (4), a closed corner (6), a corridor of columns (13-15), a lower obstacle (18) and a second closed corner (19).

The third test was done in a virtual room, where the users were left free to move in any direction, with the instructions of reporting each time they perceived something, and describe it in terms of "low" or "high" obstacle. They had 20 minutes to virtually walk around the room. Sighted participants were also asked to draw a map of the room, with the found objects on it, marked as "L" (for lower obstacles) and "H" (for higher ones). The room is shown in Figure 3.

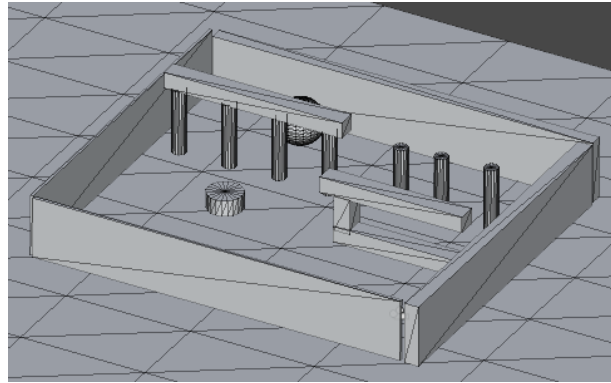


Fig 3: Perspective view of the virtual room (test 3). The starting point is the lower corner in the figure.

As it can be seen in Figure 3, in the room there are two sets of columns (called left and right sets), a horizontal obstacle composed by three sticks, a lower and cylindrical obstacle, and a sphere.

Procedure with VRE. Participants were briefly trained in one single level, which was assigned to them randomly (but according to their visual status (divided as sighted, low vision and blind), to cover all the cases as uniformly as possible). This training was done over static images (available over the online test of the protocol at http://163.117.201.122/validacion_ATAD_cerrada) for around 5 minutes, and then over the virtual environment, listening to some simple combinations of objects in front of them, for 10 more minutes. When they had difficulties to see the images being sonified, they were verbally described by the experimenter. In every case, participants could not see the screen after the training and the only feedback of the virtual reality was provided through the earphones.

After these three experiments, they were asked to complete a survey about the subjective perception of the training and the tests parts, as well as some demographic questions.

Apparatus for RE. The setup of the RE test consisted of a computer running the stereovision algorithm used to build the depth map of the scene and the sonification program. A couple of webcams, attached to a helmet, captured the scene which was transmitted through two USB cables to be processed. The produced sound was transmitted to the user through a pair of earphones.

Procedure with RE. The four students were trained between 5 and 6 hours (depending of the time they took for some exercises) in mobility and artificial vision tasks. The training included completing the VRE test twice, plus walking in the testing room, eyes opened, to learn to correlate the sound with the real objects, and a test in the same room (with a different configuration of obstacles than that of the eyes-opened step) measuring the number of detections of the different obstacles, the number of miss-detections, the number of “STOP” commands the experimenter had to say to avoid a contact, and the number of false positive detection of obstacles. The room was setup for this last experiment as shown in Figure 4.

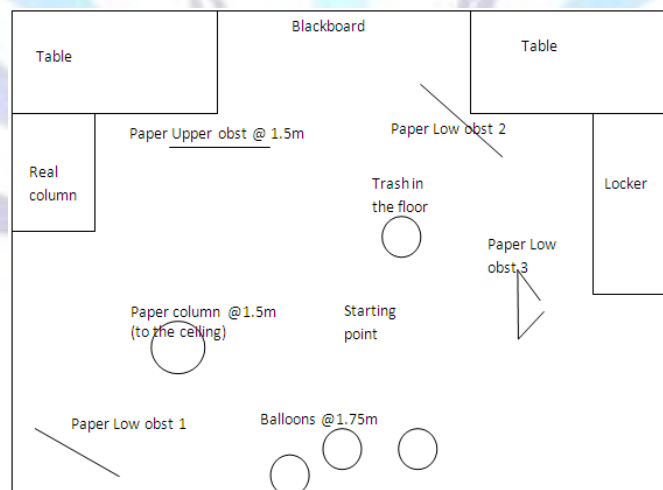


Fig 4: Room and obstacles organization.

The size of the (real) space was $5.1 \times 4.6 \text{ m}^2$.

The participants moved freely and blindfolded during 20 minutes around the room, with the experimenter behind them. After this, they were asked to complete a survey and participated in a focus group qualitatively discussing the experience, and identifying pros and cons of the real system.

RESULTS

Pre-processing of data and dependent variables

Initial analysis of the data gathered from the demographic questions of the survey raised the evidence of non-independency of some of the factors exposed in the previous section. The factors are coded as follows:

- AGE: the age in years, as of the day of the test.
- VI: three ordered values: sighted (1), low vision (2) and blind (3).
- EDU: four values: elementary school (1), high school (2), some college (3) and college degree or higher (4).
- COMP: five ordered values about the use of computers: never (1), rarely (2), once a week (3), once a day (4) and many times a day (5).

More in detail, we found the following Pearson correlation matrix shown in table 2.

Table 2. Pearson's correlation index and p-value.

Pearson correlation	AGE	VI	EDU	COMP
AGE	1	.752** (p<.001)	-.458* (p=0.014)	-.708** (p<.001)
VI		1	-.474* (p=0.011)	.637** (p<.001)
EDU			1	-.527* (p=.004)
COMP				1

* marks significant correlations at a level of .05.

** marks significant correlations at a level of .001.

These results are a consequence of the specific characteristics of participants from both the Georgia Tech and CVI pools. The first group is mostly composed by sighted individuals, aged between 18 and 26 years old, using the computer many times a day and with some college as minimum educational level. They had an average age of 21.5 years (range between 18 and 26), educational level of 3.4 and computer use of 4.9 over the same scale. The CVI participants are 51.9 years old in average with a range between 36 and 64 years old, with an average educational level of 2.7 and a computer use of 3.8 over a 5 levels Likert scale.

Use of computer, age and visual impairment are highly correlated and, thus, the analysis will be done over the use of computer, since it is the most descriptive variable (the age is quite variable and the visual impairment is so narrow with only 3 different values).

Test 1 results

First of all, we found that the 67.8% of the users achieved the end point and the 85.7% crossed the middle point in less than 10 minutes, both groups in less than 10 minutes. Highly significant results were found with a one-way ANOVA mean comparison between the use of computer and the position achieved after the 10 minutes ($F(3,24)=7.265$, $p=0.001$, shown in Figure 5).

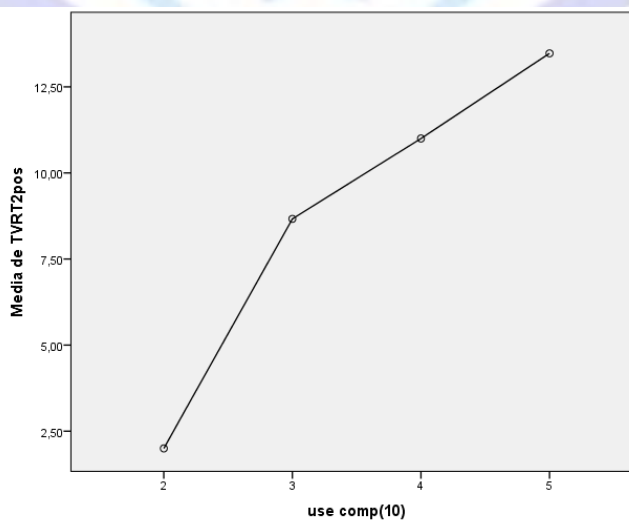


Fig. 5: Final position of the avatar (from 1 to 14) after 10 minutes or less compared to the use of computer.

Similar results were found when comparing the position against the visual impairment ($F(2,25)=9.239$, $p=.001$), the position against the educational level ($F(2,25)=10.755$, $p<.001$), and the time to finish against the visual impairment ($F(2,25)=4.146$, $p=.028$).

Another interesting result was found when comparing the time required to finish (if less than 10 minutes) against the profile. Significant results were found with a one-way ANOVA analysis ($F(3,24)=4.691$, $p=.01$). However, in the Figure 6, we see an interesting change of shape.

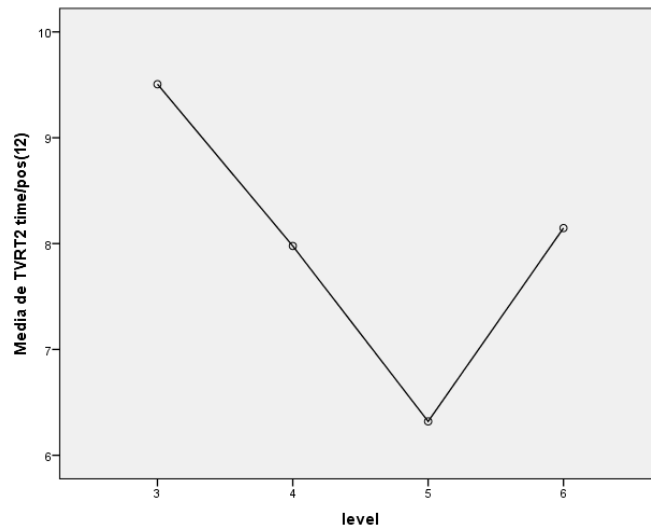


Fig. 6: Time (in minutes) required to finish compared to the profile.

Looking at the subjective perception of this test, the perception of ease of achieving the task was inversely correlated (but not statistically significant) with the use of computer, as seen in Figure 7 (Pearson non-significant correlation of -0.276 , $p=.155$).

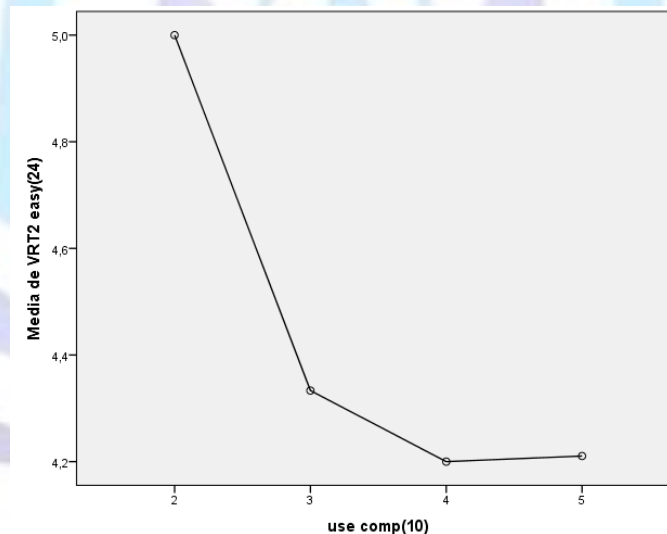


Fig. 7: Ease perceived in test 1 in terms of use of computer.

Similar relations were found when comparing the means of the perception of ease and the educational level ($F(2,25)=2.247$, $p=.147$).

When asked about the difficulty presented by corners, significant results were found with a one-way ANOVA with the profile level as factor ($F(3,24)=6.365$, $p=.005$). Not statistically significant by very interesting result was the relation between the ease perceived and the profile level, shown in Figure 8.

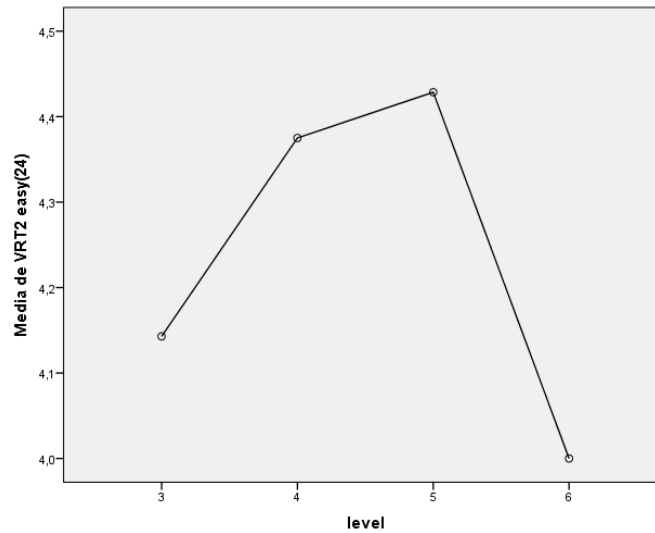


Fig. 8: Ease perceived against profile level.

Test 2 results

In the second tests, some more variables were measured: how many obstacles were detected and which ones.

Regarding the objective data, 32.1% of the participants achieved the end point in less than 10 minutes, and the 78.5% crossed the middle point in less than 10 minutes.

The descriptors of correct detection of the different obstacles are shown in table 3.

Table 3. Detection rates of different obstacles in the VRE.

Obstacle	Mean	Std. Dev
Attached column	.1071	.31497
Column	.3929	.49735
Corner1	.2222	.42366
Columns	.5600	.50662
LowBlock	.5882	.50730
Corner2	.5000	.52705

No significant results were found when comparing the achieved position after 10 minutes or less with the use of computer (ANOVA, $F(3,24)=1.037$, $p=.394$) or the profile level (ANOVA, $F(3,24)=.613$, $p=.613$). However, in this last case, we find a similar (and inverse) distribution than that shown in Figure 6, as shown in Figure 9.

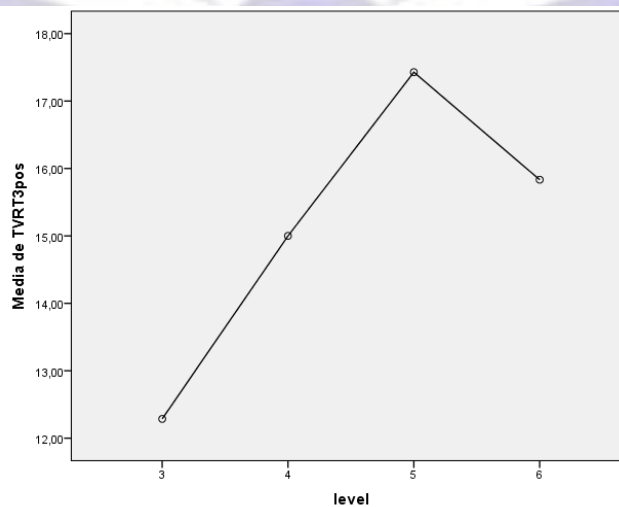


Fig. 9: Achieved position (ranged between 1 and 21) in terms of profile level.

The correct detection of obstacles also depended of the profile level following a similar pattern:

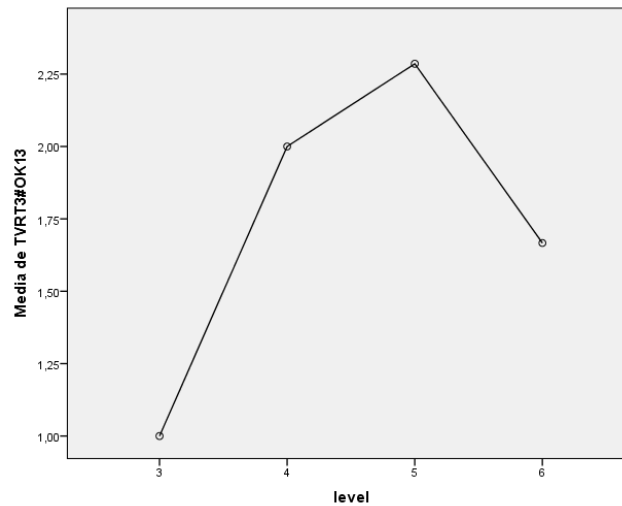


Fig. 10: Number of correctly detected obstacles in terms of the profile level.

These differences were, however, not significant ($F(3,24)=1.687, p=.196$).

Another interesting result was the distribution of probability of detection of the lower block in terms of the profile level, shown in Figure 11.

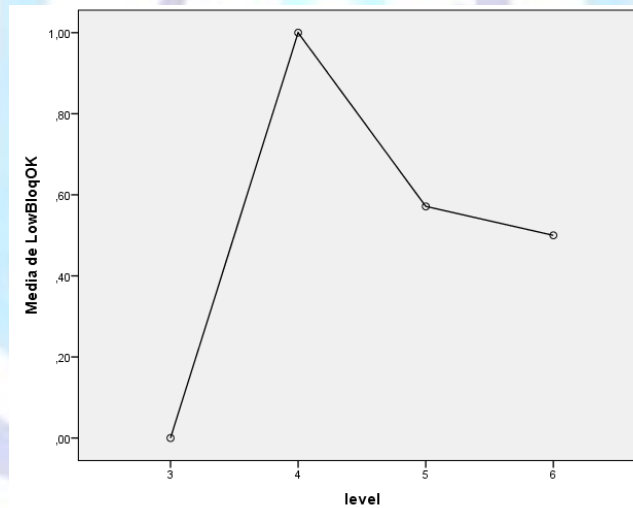


Fig. 11: Detection level (ranged between 0 –no detection- and 1 –detected-) of the lower block in terms of profile level.

The one-way ANOVA means comparison was not found to be significant ($F(3,13)=2.240, p=.132$).

We also found statistically significant differences between the number of correctly detected obstacles and the visual impairment ($F(2,25)=4.366, p=.024$), shown in Figure 12.

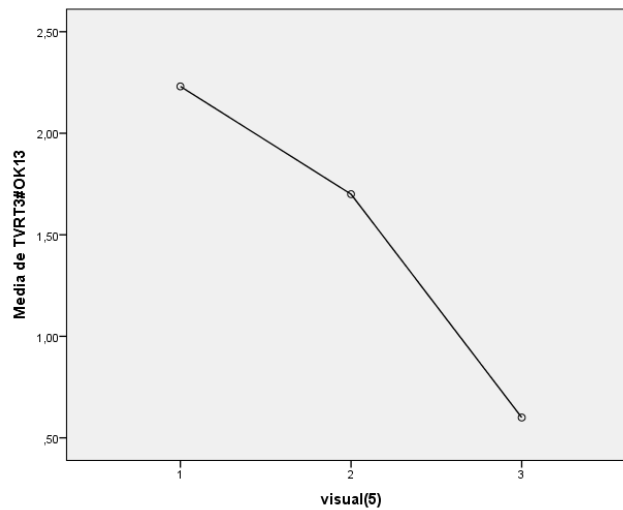


Fig. 12: Number of correctly detected obstacles in terms of the visual impairment.

Attending to the subjective results, no significant results were found with profile level, use of computer and visual impairment as factors, when the participants were asked for ease of detecting bigger or lower obstacles.

Once again, users perceived the ease of detecting obstacles differently in terms of the profile level, as shown in Figure 13.

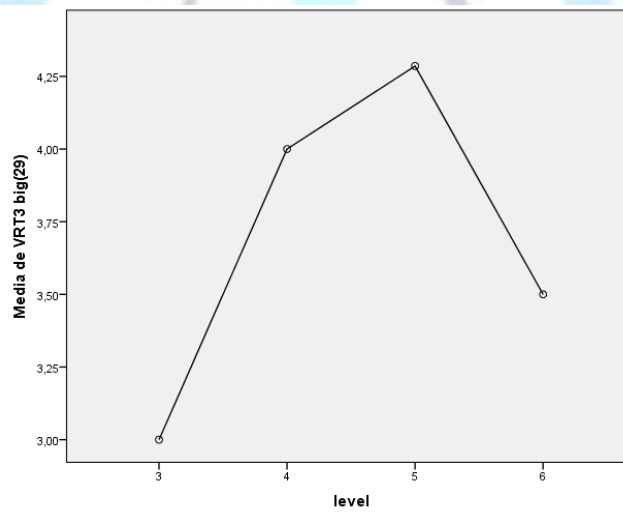


Fig. 13: Ease of perception of bigger obstacles in terms of the profile level used.

This difference was not significant ($F(3,24)=1.755$, $p=.183$).

Test 3 results

The room test provided some significant results. The Pearson correlation and one-way ANOVA analysis showed relevant correlations between detection of the right set of columns and the visual impairment (Pearson correlation = $-.599$, $p=.001$. ANOVA: $F(2,24)=6.778$, $p=.005$), the sphere detection and the same parameter (Pearson correlation = $-.466$, $p=.014$. ANOVA: $F(2,24)=4.689$, $p=.019$) and between the use of computer and the detection of the sphere (Pearson correlation = $.501$, $p=.008$. ANOVA: $F(3,23)=2.627$, $p=.075$, marginally significance).

When analyzing the subjective perception of the test, although none of the results were significant, some interesting relations can be appreciated. On the one hand, the blind people had to put more effort to produce a mental image of the room, as it can be seen in Figure 14.

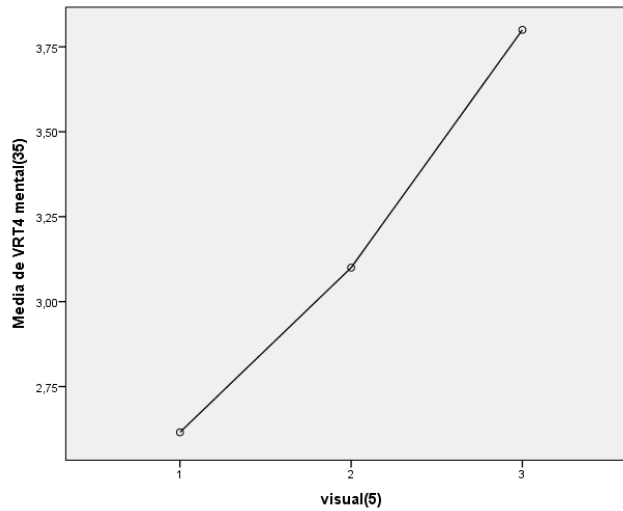


Fig. 14: Mental effort (ranged from 1 –no effort- to 5 –high effort-) against visual impairment.

In a similar way, the perception of being lost in the room was also related with the same factor, as shown in Figure 15.

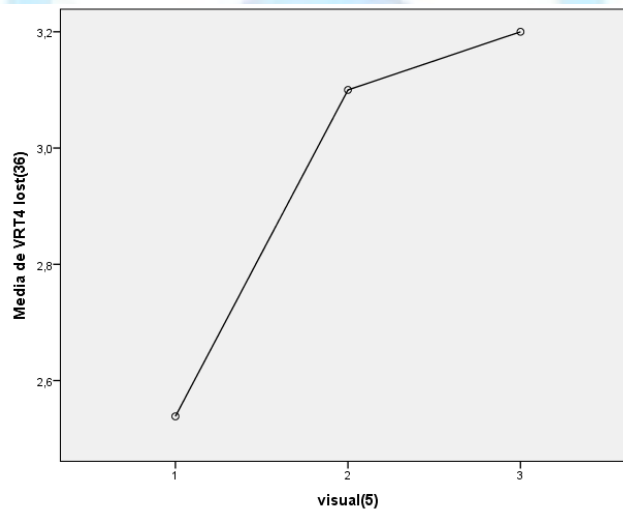


Fig. 15: Sensation of being lost (ranged from 1 –not at all- to 5 –every time-) against the visual impairment.

Experts results in the RE

Figure 16 summarizes the correct detections, miss-detections, false positives and “STOPs” (not included in miss-detections) needed to avoid crashes during the experiment of 20 minutes of blindfolded free walk around the room.

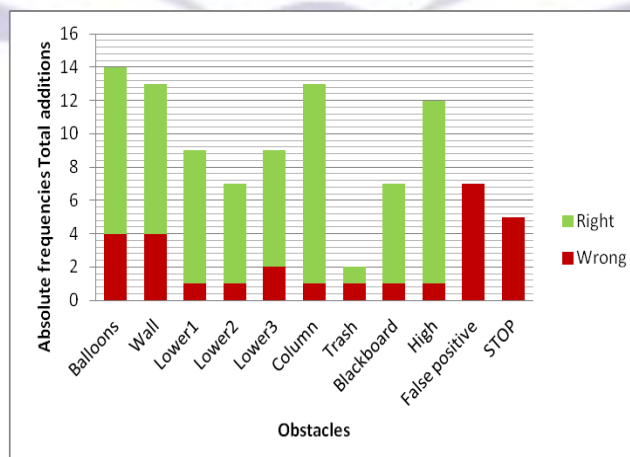


Fig. 16: Total number of right detections (in green) and miss-detections (in red) for each obstacle. False positives and STOPs are shown also in blue.

The number of errors (miss-detections, false positives and STOPS needed, all Wrong detections) and correct detections (Right) for each profile is shown in table 4.

Table 4. Detection rates of different obstacles in the real test.

	Profile 3	Profile 4	Profile 5	Profile 6
Right	11	17	17	25
Wrong	6	10	10	5
% errors	35.3	37.0	37.0	16.7

Focus group

A one hour discussion between the experimenter and the 4 experts took place to analyze subjective and qualitative aspects of the system and the perception of the whole process.

They pointed out the difficulties of getting oriented in the VRE, since no movement feelings (except rolling in the chair) were available. Moreover, the first time they did this test, it was found to be hard and with frequent feeling of being lost. In those situations, finding check points (corners) helped in orientation. However, they found it easy to follow the wall, with some complaints about the confusing effect of the corners.

Lower obstacles were found to be more difficult to detect, in both VRE and RE. Discussing about the differences between these two setups, the RE presented specific problems, such as the noise introduced by the stereovision algorithm (not always matching correctly the objects captured by the cameras).

The main complaint about the profile levels was done for level 6: too many sounds and sometimes hard to manage them mentally. The vibrato effect to detect laterality of objects sometimes distorted the interpretation when too many sounds were present, as one expert pointed out. In this way, another one suggested to cut some information before sonifying it, since there was irrelevant noise not usable for mobility but confusing in the whole understanding of the perceived sounds.

They talked and agreed about the utility of walking in the room eyes opened with the system as training for the final test blindfolded. They suggested that blind people could use known places to train the system, as well as verbally describing how their surroundings are.

It was a generalized opinion that confidence with the system increased with use, but some of them pointed out the necessity of more practice and use in everyday tasks, where it can be very useful.

Final evaluations

Final questions about the global process were asked to the participants, among which the fatigue of the global process, considerations about the length of the training, feelings of safeness and use of the white cane or the guide dog.

Statistically significant Pearson correlation was found between the educational level and the perception of the fatigue of the whole process (-.384, p=.044). This result is shown in Figure 17.

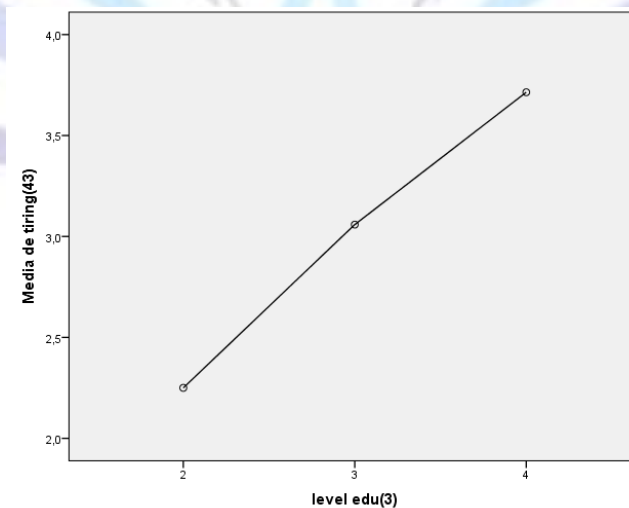


Fig. 17: Perception of fatigue (ranged between 1 –not tiring at all- and 5 –very tiring-) against the educational level.

Another interesting result is the distribution of the perception of fatigue versus profile level, shown in Figure 18.

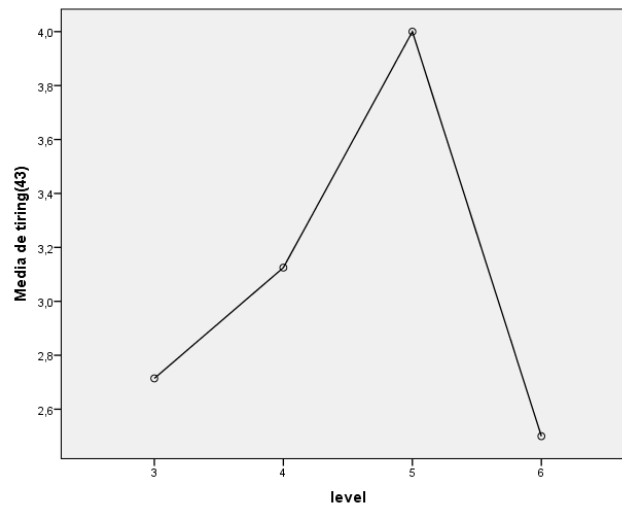


Fig. 18: Perception of fatigue (ranged between 1 –not tiring at all- and 5 –very tiring-) against the profile level.

An unexpected result comes from the comparison of the use of computer and the perception of safeness (marginally significant with a one-way ANOVA, $F(3,24)=2.707$, $p=.068$), shown in Figure 19.

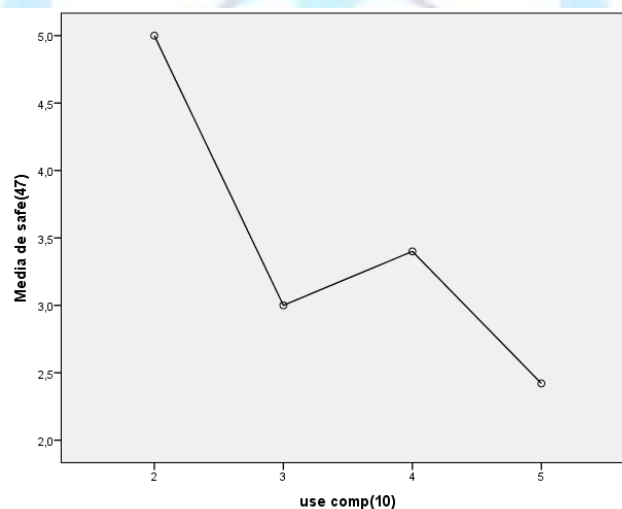


Fig. 19: Perception of safeness (ranged between 1 –not safe at all- and 5 –very safe-) against the use of computer.

Another marginally significant result (one-way ANOVA, $F(2,11)=3.378$, $p=.072$) is the relation between the intention of keep using the white cane or the dog guide against the visual impairment (with a positive Pearson correlation of .585 and $p=.028$).

The lower correlation was found between the visual impairment and the perception of tiredness (.030, $p=.878$).

DISCUSSION

First of all, we have to discuss the specific composition of the participants' pool available for this experiment, as well as its size. The highly correlation found between visual impairment and other theoretically independent variables such as use of computer, educational level or age is due to the composition of the clients and staff of the CVI, compared to the average student of the GeorgiaTech, much younger and used to use computers daily. This first condition forces us to unify hypotheses H2.1 to H2.4 with the exception of some specific results where we appreciated some internal differences.

Regarding the second point, 28 participants (and 4 experts) are not a very big set of subjects, and the quantitative data obtained should be contrasted with larger experiments. This one can be taken as a preliminary study of the usability and efficiency of the system proposed in [5].

We found that the 67.8% of the users achieved the end point in the first labyrinth and the 85.7% crossed the middle point in less than 10 minutes, both groups in less than 10 minutes, with a training of 15 minutes. Given that the sonification protocol presents some aspects which are not intuitive at all, this result can be taken as an evidence of the correctness of H1. The same hypothesis could be checked in the second test, although the efficiency decreased in the second labyrinth. This can be due to the following factors:

- The second labyrinth was a 20% longer than the first one, but the time available remained the same.



- The second labyrinth presented some obstacles, which had to be found (according to the verbal instructions to the users), and many users spent some time exploring their environment to decide whether some sounds represented an object or just the empty ground or the wall.

The amount of obstacles found is not so high, a mean of 2.93 and a standard deviation of 1.3. Given that there were 6 tagged obstacles (two of them narrow corners which were ignored by some participants), slightly less than the half of them were perceived as obstacles. More in detail, the column attached to the wall was only perceived as an obstacle by 3 participants. The rest of them simply dodged it as if it was the normal shape of the path. The limit of what can be considered as an obstacle and what is part of the background is not always clear. The low block and the columns were detected by the half of the participants that reached those points. According to these results, perceiving a flat wall seems to be easier than some objects (always tinier). The four experts (participants with longer training) had better results when using the real system, with error rates in detecting obstacles and walls between .167 and .37. The main conclusion of these different results is the importance of the training to be able to distinguish smaller and more subtle and fine elements (we would be entering in the artificial vision field). H1.1 can be then taken as correct, taking into account the dependency of the training.

We also found many evidences to support the general hypothesis H2 (see Figures 4-15 and 17-19 and the correlative statistic data), however, and due to the specific demographic composition of the subjects' pool, we cannot find specific differences between the COMP, AGE, EDU and VI factors. Taking them as a more or less homogeneous set (see table 2), we found evidences of the influence of these factors on the efficiency (see Figures 5 and 12) and subjective perception of its usability (Figures 7, 14, 15, 17 and 19). In all these cases, the higher is the familiarity with the computers (and the higher is the educational level), the better are the results. However, among these results, we find unexpected relations. Let's consider the fact that the lower is the use of computers makes people think the completion of the labyrinth was easier (Figure 7 and correlative statistical data). This does not match with the objective data about completing it ($F(3,24)=7.265$, $p=0.001$, and shown in Figure 5), and some conjectures can be proposed:

- There can be differences between the time perception of the people used to manage computers and/or doing exams and those who are not.
- The level of self-exigency could be modulated by the educational level.
- There can be differences in the conservatism and the notion of risk (and, thus, the difficulty to perform a task correctly) between young and older people.

Even if age, use of computer or visual impairment are somehow correlated in our dataset, sometimes we can find significant results focusing our attention in some specific factor. Attending to differences in terms of visual impairment, we find important data when we compare the amount of correctly detected objects in the second labyrinth ($F(2,25)=4.366$ and $p=.024$, Figure 12). Blind people (of our pool) encounter more difficulties to find objects. This fact is coherent with the subjective perception of the effort needed to perform this task (Figure 14, no significant relation) or the feeling of being lost (Figure 15, neither significant results).

The hypothesis H2.5, proposing a relation between the profile level and the efficiencies of the tests (and, maybe, the perception of usability by the users) could be tested independently of the demographic characteristics of the participants' pool, assigning levels equally to each visual category. In the first three tests, we find important changes of efficiency and subjective perceptions in terms of the profile level assigned. When comparing the time required to finish (if less than 10 minutes) against the profile in the first test, significant results were found ($F(3,24)=4.691$ and $p=.01$, Figure 6). The results of the second test (Figures 9 and 10) were found to be not significant but with the same shape. Significant or not, the performance in different aspects reach a maximum with the profile 5, with the exception of the detection of the lower block in the second test. This exception can be due to specific differences between users. It is important to notice that not every user achieved this block (57.1% of them faced the lower block, one of them with profile level 3, with profile level 4, 6 with profile level 5, 4 with profile level 6). These data should be taken carefully and tested more extensively in the future. This is not reproduced in the efficiency of the experts, but no conclusions about this fact can be extracted since the sample is the smaller one.

Profile level 5 is not only the most efficient. It is perceived as so (Figure 8 for test 1, Figure 13 for test 2, Figure 18 for general evaluation of the tiredness of the process).

The experts allowed us to reach new limits in the usability and efficiency of the system, since larger trainings permitted to them to get more familiar with the sonification and the specific problems of the real system. The detection rate of lower obstacles (see the "trash", in Figure 16) or plain surfaces (see "wall" or "balloons" in the same figure) produced most of the miss-detections. In contrast, columns and paper lower or high obstacles were easily detected by all of them every time they faced them.

When discussing about the mobility aspects of the system, they pointed out the ease of some tasks (such as following the wall) but at the same time, the problems to detect smaller obstacles in the ground. They also pointed the increase of confidence with the use, and the necessity of longer trainings and use.

The main problems were found in the complexity of profile level 6, and in the noise of the real system to produce reliable sonifications. The suggestion of cutting some information (the less relevant one) to increase the usability should be taken into account in further implementations.

Regarding the general evaluations, we can find again unexpected results, such as the direct correlation between educational level and the feeling of tiredness after the tests. People with higher educational level should be more prepared



for intellectual and sometimes boring tasks, but they manifested the higher rates of tiredness (Pearson correlation of $-.384$ and $p=.044$, Figure 17). The comparison of the safeness feeling with the use of computer gives us another unexpected result ($F(3,24)=2.707$ and $p=.068$, Figure 19): the group with better results (with higher rate of computer use) felt more unsafe than that with lower performances. The reasons can be related to those pointed out when discussing about the perception of ease of the labyrinth test.

Finally, as expected, blind people want to keep using the white cane or the guide dog even if they could use this system (4.8 over 5 as average response against 3 of the sighted group and 4.33 of the low vision group). Given that blind people use to encounter harder dangers in the middle height than in the bottom part [7], and the comment of the experts about the higher ease of detecting middle and higher obstacles, the system can increase the safeness in the travels of this collective.

The main limitations of this study (the demographic composition of the participant subgroups and the size of the sample) should be the first tasks to be performed in the future, with the goal of obtaining statistically stronger data to describe the use, efficiency and usability of the system.

Longer training should be deserved to evaluate the real potentials of the system, and more real life tests should also be designed and completed.

CONCLUSIONS

The sonification can help visually impaired people to perceive information about the spatial configuration of their surroundings. However, this translation must be trained.

In this study we have tested a new assistive product over a set of 28 people, sighted, with low vision and blind, trying to find its limitations and strengths.

We found that different combinations and structures of sounds can bring to differences in the efficiency and even in the usability of a sonification-based assistive product.

We also found relations with other aspects not specifically related with disabilities, such as educational level, use of computer in the daily life or the age.

Some simplifications can be done on the system (such as eliminating the most complex profile level, or "cleaning" the presented information) to give the user only the most relevant one.

The most efficient user of the tested assistive product is a young person, used to manage computers and with some college or a college degree educational level. People not matching these characteristics seem to face harder problems to understand or at least take the right decisions in mobility tests with this system. However, they perceive positively the system and its use makes them feel safer to move with some autonomy.

ACKNOWLEDGMENT

We would like to acknowledge the collaboration of the Center for the Visually Impaired of Atlanta, GA, for their support and helpful contacts. We would like to thank, as well, Georgia Tech and the Carlos III University for their facilities.

REFERENCES

- [1] C. Martinez, "Orientation and Mobility Training: The Way to Go." Texas Deafblind Outreach. 1998.
- [2] S. J. La Grow, "Orientation to Place." Center for International Rehabilitation Research Information and Exchange (CIRRIE), ed. vol. International Encyclopedia of Rehabilitation, pp. 1-8. 2010.
- [3] P. Revuelta Sanz, B. Ruiz Mezcuca, and J. M. Sánchez Pena, "ICTs for Orientation and Mobility for Blind People. A State of the Art." ICTs for Healthcare and Social Services: Developments and Applications. Isabel Maria Miranda and Maria Manuela Cruz-Cunha, eds. 2011.
- [4] B. N. Walker, A. Nance, and J. Lindsay, "SPEARCONS: Speech-based Earcons Improve Navigation Performance in Auditory Menus," Proceedings of the 12th International Conference on Auditory Display (ICAD 2006). pp.63-68, 2006.
- [5] P. Revuelta Sanz, B. Ruiz Mezcuca, and J. M. Sánchez Pena, "A Sonification Proposal for Safe Travels of Blind People." Proceedings of the 18th International Conference on Auditory Display (ICAD 2012) , pp. 233-234. 2012.
- [6] M. Pec, M. Bujacz, P. Strumillo et al., "Individual HRTF Measurements for Accurate Obstacle Sonification in an Electronic Travel Aid for The Blind." International Conference on Signals and Electronic Systems (ICSES 2008) , pp. 235-238. 2008.
- [7] P. Revuelta Sanz, B. Ruiz Mezcuca, and J. M. Sánchez Pena, "Users and Experts Regarding Orientation and Mobility Assistive Technology for the Blinds: a Sight from the Other Side." Proceedings of the AIRTech Int.Conference 2011. pp. 3-4. 2011.



Author biography



Pablo Revuelta Sanz, obtained his degree in Telecommunication Engineering (2006) at the Carlos III University of Madrid (Spain), obtained his Master's degree in 2008 and Ph.D. in 2013 from the same university.

He has published several conference papers and two book chapters focused on assistive products, image processing, sonification and political philosophy.



Belén Ruiz Mezcuca is Ph.D. in Physics from the Telecommunications School of the UPM (1995).

She is director of Informatic studies in the Carlos III University and lecturer in the Computer Science department of the Carlos III University. She is the head of the Spanish caption Center "Centro Español de Subtitulado y Audiodescripción".



José M. Sánchez Pena received his MS and PhD degrees in Telecommunication Engineering from the Polytechnic University of Madrid in 1988 and 1993, respectively.

His main research topics are assistive technologies, optoelectronic instrumentation and displays. He is currently a full professor and responsible for the Displays and Photonic Applications Group in the Electronic Technology Department. He is a senior member of IEEE.



Bruce Walker is an Associate Professor at Georgia Tech, in the School of Psychology and the School of Interactive Computing.

Particular research interests include sonification and auditory displays. Professor Walker teaches HCI, Sensation & Perception, Auditory Interfaces, and Assistive Technology. In addition to academic research leading to over 100 publications, he has worked and consulted on projects for NASA, state and federal governments, private companies, and the military.