MULTISOURCE SONIFICATION FOR VISUAL SUBSTITUTION IN AN AUDITORY MEMORY GAME: ONE, OR TWO FINGERS?

Juan Diego Gomez
C.S Department
University of Geneva, Switzerland
juan.gomez@unige.ch

Guido Bologna
C.S Department
University of Geneva, Switzerland
guido.bologna@unige.ch

Benoît Deville
C.S Department
University of Geneva, Switzerland
benoit.deville@unige.ch

Thierry Pun
C.S Department
University of Geneva, Switzerland
thierry.pun@unige.ch

ABSTRACT

The *See ColOr* project aims at developing a mobility system for blind persons based on image color sonification. Within this project the present work addresses the optimal use of auditory-multi-touch interaction, and in particular the matter of the number of fingers needed for efficient exploration. To determine the actual significance of mono and multi-touch interaction onto the auditory feedback, a color matching memory game was implemented. Sounds of this game were generated by touching a tablet with one or two fingers. A group of 20 blindfolded users was tasked to find color matches into an image grid represented on the tablet by listening to their associated color-sound representation. Our results show that for an easy task aiming at matching few objects, the use of two fingers is moderately more efficient than the use of one finger. Whereas, against our intuition, this cannot be statistically confirmed in the case of similar tasks of increasing difficulty.

1. INTRODUCTION

Multi-touch technologies and table-top tangible user interfaces bring the promise of intuitiveness and ease of communication to the human-machine interaction field. Those strengths, together with an auditory feedback, might be targeted to the visual impaired population to attempt, for instance, at creating new mobility aids. The *See ColOr* project particularly aims at developing a mobility system for blind persons based on image color sonification [1].

We present an experiment carried out by means of a new sonification module of the *See ColOr* mobility aid for visually impared users. The key idea behind this aid is the mapping of colors into musical instrument sounds. The new module allows the user to explore an entire video frame captured by a stereoscopic camera with her/his fingers touching a small tablet. A benefit of this particular interaction is that it makes it possible to determine the relationship of picture objects with respect to color and depth. Compared to the previous developed perceptive module [1] that allows the user to receive sonified data about a small number of contiguous points [1], the new module now provides the relationships between distant points.

Throughout the work presented here, we attempt at assessing the objective difference on sound localization's ability of blindfolded people on a small tablet. We evaluate two cases: multi-touch mode (bi-manual) and mono-touch mode (one finger). Hopefully, the results of this study might be roughly extrapolated to achieve an initial guideline on how to perform efficient object localization via tactile-audio exploration for blind users.

Intuitively, people are inclined to believe that multi-touch interfaces are more helpful than mono-touch ones to interact with tangible technologies, since more information of the screen can be accessed at the same time. This intuition seems to make sense since it could be also said that the number of fingers is directly proportional to the rate of information transfer between the human and the system. Therefore, our belief is that using more than one finger represents a notable improvement on audio-based localization of objects in a scene. Since this assumption cannot hardy be generalized yet, we would like to set out this discussion by testing our supposition on a specific matching game that links the multi-touch interaction with the exploration of environments (video images) via audio.

Our multi-touch system is part of the development of general-purpose mobility assistance for visually impaired individuals, *See ColOr*. Before using the system, we are looking to establish the best parameters toward a proper usability and an optimal human-machine interaction. Therefore, we developed a simple game following the basic idea behind *See ColOr*. The goal is to match same colored squares into an image grid represented on a tangible tablet by mean of tactile exploration and auditory feedback. Our main goal is thus to assess the user performance in a multi-touch mode against mono-touch mode and define an objective difference if any.

2. STATE OF THE ART

2.1. Vision substitution based on audition

Several authors proposed special devices for visual substitution by the auditory pathway in the context of real time navigation. The *K Sonar-Cane* combines a cane and a torch with ultrasounds [14]. Note that with this special cane, it is possible to perceive the environment by listening to a sound coding depth. *TheVoice* is another experimental vision substitution system that uses auditory feedback. An image is represented by 64 columns of 64 pixels [13]. Every image is processed from left to right and each column is listened to for about 15 ms. In particular, every pixel gray level in a column is represented by a sinusoidal wave sound with a distinct frequency. High frequencies are at the top of the column and low frequencies are at the bottom.

Capelle *et al*. proposed the implementation of a crude model of the primary visual system [9]. The implemented device provides two resolution levels corresponding to an artificial central retina and an artificial peripheral retina, as in the real visual system. The auditory representation of an image is similar to that used in “TheVoice” with distinct sinusoidal waves for each pixel in a column and each column being presented sequentially to the listener.

TheVIBE [20], is a visuo-auditory substitution system invented by Sylvain Hanneton and currently developped in collaboration with Gipsa-Lab. This system converts video-streaming into auditory streaming. The sound generated by *TheVibe* is a weighted summation of sinusoidal sounds produced by virtual "sources", corresponding each to a "receptive field" in the image.

Gonzalez-Mora *et al*. developed a prototype using the spatialization of sound in the three dimensional space [15]. The sound is perceived as coming from somewhere in front of the user by means of head related transfer functions (HRTFs). The first device they achieved was capable of producing a virtual acoustic space of 17x9x8 gray level pixels covering a distance of up to 4.5 meters.

As none of the works quoted above attempts at providing a substitute to color information an important visual perceptual property is likely to be missed. The visual information and surrounding environments are mainly conformed by colored entities, so that the ability to properly interpret them could be diminished as colors are left out. Moreover, approaches which account for depth information are few as yet. A certain perception of the actual distance of the surrounding objects may also help the user to build a more accurate occupancy representation of the environment. Thus, methods like Gonzalez-Mora [15] could be complemented by fusing the spatialization of sound with depth information to better orientate blind persons.

2.2. Multi-touch technologies

Mobile devices with multi-touch capabilities are becoming increasingly common, largely due to the success of the Apple iPhone, iPad and iPod Touch. While there have been some advances in touch-screen accessibility for blind people, touch-screens remain inaccessible in many ways. Unfortunately, as discussed by McGookin et al. [16], the creation of accessible modifications and enhancements for touch-based devices is lagging behind the breakneck pace of mainstream development.

Touch-screens pose an especially daunting challenge for blind or visually impaired users. In place of familiar devices like the keyboard, screen computing generally offers a uniform, featureless surface. The trend of emulating existing GUI's while using the finger as a mouse can translate existing problems into an even more difficult context. On top of this, mobile systems often feature unique or restructured interface layouts. The end result is a confounding environment for accessibility programs [17].

In addition, touch-screen is becoming more popular not only for home appliances but also for so many kinds of machines in public facilities (e.g., ATM Bank Machine, Automatic Ticket Machines). However, most of visually impaired individuals have problems with the difficult usability of a touch-screen display. The reason for this problem is the fact that neither uneven screens nor embossed designs are used in this technology; therefore, the visually impaired people cannot define the location of GUI parts, such as buttons, links to other pages and menu lists. A potential solution for this problem is that instead of pure tactility, using the auditory feedback would help visually impaired users to access these equipments.

At the best of our knowledge, no works in the literature are entirely dedicated to merge/evaluate tangible interfaces and auditory feedback as mobility aid systems. However, the work carried out in [18] could be likely the closest approach. In fact, the authors also study the ability of sound-based localization into a scene for blind people by implementing a memory matching game using animal’s sounds. This work however doesn’t perform any evaluation on the optimal use of auditory-multi-touch interaction and its scope doesn’t cover the mobility assistance.

3. SEECOLOR

Perhaps, one of the most important aspects of the reality that surrounds us, which is imperceptible to the blind people, is color. The real world we live in is populated by colored objects. Not surprisingly, colors are extremely important to either locate objects in space and identify them. The *See ColOr* [1] mobility aid for visually impaired individuals transforms aimed-at regions of a color video image into sound sources represented by spatialized musical instruments. The association of sounds with colors is achieved thanks to an efficient quantization of the regions’ HSL representation (Hue, Saturation, Luminance). The main propose of *See ColOr* is to provide blind users with a capability of perception of the environment in real time.

At the highest level of abstraction, the *See ColOr* offers several units (Local Module, Alerting system and Global Module) [20] integrated into a multi-module architecture. In this architecture, the alerting system is designed to detect obstacles while the local module sonifies the central part of the image. Moreover, the Global Module allows the user to explore the current image scene, as a whole.

Figure 1: Decomposition of the *See ColOr* mobility aid into three distinct modules (colored engines). The input (stereoscopic camera) and output (headphones) peripherals are shown at either side of the schema.

Thanks to this global module the user will be able to discover the main components of an image and determine their relative position while a global perception of the environment will be built in the user's mind. This will be achieved by means of a small touchpad that will activate the auditory representation of the image being taken by the camera in real time. The prototype allows the user to navigate both indoor and outdoor spaces by listening to the sound of the references points chosen by her/him self via touching. Notice also that the final *See ColOr* prototype presents a sonification method that renders not only colors but also depth information thanks to the use of a stereoscopic (or time-of-flight) camera. In the *See Color* prototype, depth is represented by sound duration; the more prolonged the sound, the farther the touched point.

3.1. Sonification

Once a spatial-to-pixel mapping between the touchpad and the image is established, a user can explore any part of the image by touching on pad. A selected pixel is represented in HSL format and then encoded by labeling it with the index of a corresponding sound based on the following empirical quantization [3] of the hue variable H:

1. Oboe for red ($0 \leq H < 1/12$)
2. Viola for orange ($1/12 \leq H < 1/6$)
3. Pizzicato violin for yellow ($1/6 \leq H < 1/3$)
4. Flute for green ($1/3 \leq H < 1/2$)
5. Trumpet for cyan ($1/2 \leq H < 2/3$)
6. Piano for blue ($2/3 \leq H < 5/6$)
7. Saxophone for purple ($5/6 \leq H \leq 1$)

In addition, a linear mixture of two instruments whose gains depend on the proximity of the two closest ranges is used. For instance, if H lies between yellow and green, the resulting sound is a linear combination of pizzicato violin and flute. Moreover, in order to represent the degree of purity of the color, the saturation variable S is rendered by a sound pitch, which is assigned depending on four saturation values by means of four different notes:

1. C for ($0 \leq S < 0.25$)
2. G for ($0.25 \leq S < 0.5$)
3. B flat for ($0.5 \leq S < 0.75$)
4. E for ($0.75 \leq S \leq 1$)

Finally, when the luminance L is rather low (i.e. $L<0.5$), the pixel is represented by mixing the resulting sound from the above H-S encoding with a double bass, using one of the four possible notes (C, G, B flat or E) depending on the luminance level. The double bass is replaced by a singing voice (using the same possible notes) if $L > 0.5$. However, if the luminance L of the pixel is close to zero (the perceived color is black) only the double bass is retained. Similarly, when the luminance of the pixel is almost one, the resulting sound is an unmixed singing voice.

3.2. Sound spatialization

The ability to perceive sound in space is denoted as spatial perception. Since the outer ears can be either directly implemented or modeled with digital filters [10], by filtering a digitized sound source, one can potentially place sounds anywhere in the virtual auditory soundfield, through a head-phone listener.

The spectral filtering of a sound source before it reaches the ear drum is called the Head-Related Transfer Function (HRTF) [2]. The HRTF is a function that characterizes how an ear receives a sound from a point in space. It is possible to artificially spatialize a sound by empirically measuring the HRTFs of a given individual. Microphones are placed near the eardrums, while speakers are placed at various locations around the person; pseudorandom noise sounds are then played. As it reaches the eardrums, the recorded sound takes into account all the effects of the HRTFs for that person. By comparing the original and recorded waveforms, the subject's HRTFs can be computed for all the different positions of the sound source. This is achieved by using digital filter algorithms that transform the spectrum of the sound source in the same way it would have been transformed under normal spatial hearing.

We have only implemented spatialization of sound on the azimuth plane (left to right) because the spatialization of elevation represents a much more difficult problem [12]. The user, then, will be able to perceive the azimuth spatial location of any point over the images. The images have been divided into 25 vertical bands (see figure 2). Each of those bands has a particular sound location in the azimuth plane. Note that the choice of 25 divisions per image is due to the fact that we use the CIPIC database, for which 25 positions on the azimuth plane are recorded [11]. Figure 2 offers a graphical schema of this idea.

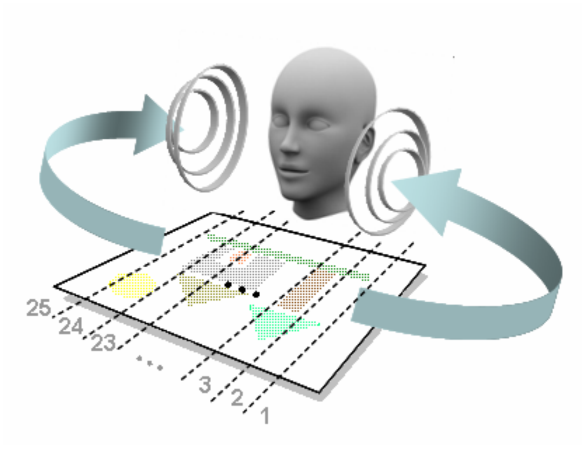

Figure 2: An user in front of an image sectioned into 25 parts. First and last parts are totally sonified on the left and right side, respectively. Mid parts are sonified in the central of the soundfield (half left and half right). Sounds are perceived by the user through headphones (see fig 1).

4. PREVIOUS SEE COLOR EXPERIMENTS

In the first step of the *See ColOr* project, we performed several experiments with six blindfolded persons who were trained to associate colors with musical instrument sounds [4]. As shown in Figure 3, the participants were asked to identify major components of static pictures presented on a special paper lying on a T3 tactile tablet representing pictures with embossed edges, http://www.rncb.ac.uk/page.php?id=872. Specifically, this tablet makes it possible to determine the coordinates of a contact point. When one touched the paper lying on the tablet, a small region below the finger was sonified and provided to the user. Color was helpful for the interpretation of image scenes, as it lessened ambiguity. As an example, if a large region "sounded" cyan at the top of the picture it was likely to be the sky. Finally, all participants to the experiments were successful when asked to find a bright red door in a picture representing a churchyard with trees, grass and a house.

Figure 3: Example of an embossed picture on the T3 tactile tablet.

The work described in [5] introduced an experiment during which ten blindfolded individuals participants tried to match pairs of uniform colored socks by pointing a head mounted camera and by listening to the generated sounds. Figure 4 illustrates an experiment participant observing a blue sock. The results of this experiment demonstrated that matching similar colors through the use of a perceptual (auditory) language, such as that represented by instrument sounds can be successfully accomplished.

Figure 4: A blindfolded subject observing a blue sock.

In [7] the purpose was to validate the hypothesis that navigation in an outdoor environment can be performed by "listening" to a colored path. As shown by Figure 5, we performed an experiment during which ten blindfolded participants and a blind person were asked to point the camera toward a red sinuous path painted on the ground and to follow it for more than 80 meters. Results demonstrated that following a sinuous colored path through the use of the auditory perceptual language was successful. A video entitled "The *See ColOr* project" illustrates several experiments on

http://www.youtube.com/guidobologna.

Figure 5: A blindfolded individual following a red sinuous path using the See ColOr guidance.

5. EXPERIMENTS AND RESULTS

The game in the experiments used pictures representing grids of colored squares. The user's goal was to find out the color matching squares. The pictures were represented onto a multi-touch pad, as shown in Figure 7. The only guidance to reach the goal was given by the sound of each color which was activated by the user's finger contact. The spatial position of each finger on the touchpad was also spatialized in terms of sound on the azimuth plane. In other words, finger contacts over the left, respectively the right part of the pad (grid) produced sounds originating from the left, respectively the right; fingers touches around over the center were played “half” left and “half” right.

In order to vary the level of complexity, we used for this experimental game a set of grid images with 2x2, 3x3 and 4x4 squares. Thus, each level had an exact number of correct matches of 2, 4 and 8 respectively, leaving for the 3x3 grid, one unpaired square. The entire game just ended as the participant had successfully completed those three levels of difficulty. For each grid (level), the colors of the couples were chosen based on an equal-spaced sampling over 360 degrees of the Hue variable domain in the HSL (hue, saturation and luminance) cylindrical-coordinate representation [3]. Sampling the color space uniformly is just to avoid repeated colored pairs into a same grid. Once a color was selected for a couple, a random position into the grid was assigned.

5.1. Experiments

We conducted experiments on 20 blindfolded persons tasked to play the game until complete the three levels. Every participant was given around 10 minutes of initial training as it was the first time they tried it out. Moreover, four markers were attached to the touchpad indicating the middle of the edges of the squared sensitive area; this gave users a measurable notion of the central position up, down, right and left of the image.

Figure 6: The *SeeColOr*-global-module touchpad on which images are represented. The circled areas are the four markers indicating the middle of the edges of the sensitive area.

During the first part of the test, participants tackled the game passing through all the levels, being allowed to interact with the system just using one finger in a mono-touch strategy. In the next phase, the game ran once again from the initial level with a new set of images so that the participant could play again in a bi-manual mode (using one finger per hand); this permitted the multi-touch evaluation.

For each participant we measured the elapsed time to complete each game stage (2x2, 3x3 and 4x4 grids) using one and two fingers, separately. Although rarely, some participants chose mismatched couples as well as pairs previously chosen. However, besides the fact that those cases seldom occurred, the evaluation did not target the memory capability so no relevance was given to this matter. And hence each level was successfully completed as soon as the maximum number of couples per grid was reached, regardless of mismatches. The actual impact of mono and multi touches on the matching’s ability (speed) was only taken into account. Therefore, participants were neither refuted nor confirmed on their responses while gaming, as they were given total freedom to perform the task.

Game participants were blindfolded for this experiment and were given a high quality headset to correctly perceive the spatialization of the sounds [2]. They were not allowed to see the images before or during the test; a two minutes break was given between fulfilled levels, as well as between the mono and multi-touch sessions.

Figure 7: A blindfolded individual playing the memory matching game of colors.

5.2. Results

In this section, we show the results on matching's ability over a multi-touch pad, based on auditory feedback, of the blindfolded users. As mentioned in the previous section various measurements were collected from the experiment. We however consider more relevant to determine the impact of mono and multi-touch exploration, to focus the results on the global time taken by the participants to finish up a game level in both cases. Indeed, this data could give an indication of the advantages of one method over the other, if any. The times of the participants while going through the three levels of the proposed game are shown in Table 1.

test	***Monotouch***			***Multitouch***		
	2x2	***3x3***	***4x4***	***2x2***	***3x3***	***4x4***
1.	23	316	900	10	289	491
2.	7	94	184	4	72	293
3.	9	132	367	13	199	498
4.	38	164	300	10	47	229
5.	19	185	341	17	113	398
6.	20	103	386	21	95	380
7.	11	211	405	6	202	426
8.	30	205	298	10	218	368
9.	14	232	426	7	200	411
10.	28	142	356	12	135	421
11.	8	158	423	9	157	450
12.	25	109	290	11	100	285
13.	33	183	325	28	170	338
14.	12	152	298	12	165	300
15.	20	190	352	19	171	401
16.	18	161	301	10	120	360
17.	30	170	289	12	85	582
18.	19	175	362	15	92	597
19.	8	198	402	10	120	605
20.	26	117	538	16	193	785

Table 1: Time (in seconds) spent by 20 blindfolded participants at reaching the goal of the memory matching game of colors through all its levels (2x2, 3x3 and 4x4) using one and two fingers.

To better understand the results obtained from the experiments we performed a paired t-test on the results obtained with one finger and two fingers. The paired t-test assesses whether the means of two series of experiments are statistically different from each other. By setting a ***new hypothesis*** $\boldsymbol{H_0}$ on the equality of the two series' averages, the t-test for the 2x2 game rejected $\boldsymbol{H_0}$ at a confidence level equal to 99%. Thus, for this game level the use of two fingers resulted significantly more efficient than the use of one finger. In the other games the t-tests failed to reject $\boldsymbol{H_0}$, thus there is no statistically significant advantage to use two fingers instead of one. Table 2 depicts the average time (and standard deviations) spent by participants during the games, while the results of the t-test are shown in table 3.

level	***Monotouch***	***Multitouch***
2x2	19.785 (10.024)	12.142 (6.261)
3x3	170.428 (58.949)	154.428 (65.493)
4x4	378.5 (163.295)	377.714 (80.315)

Table 2: Mean times (in seconds) to achieve the game's goal for each level, between parentheses the standard deviations.

level	***t-test conclusion***	***p-value***
2x2	Reject H_0	0.0087
3x3	Fail to reject H_0	0.1764
4x4	Fail to reject H_0	0.9821

Table 3: T-test results. Notice that only p-values < 0.01 can reject the herein-related equality hypothesis $\boldsymbol{H_0}$.

6. DISCUSSION

After the experiment, participants were asked for their own impressions on the distinction between doing the mono-touch and the multi-touch test. Quite oppositely to our intuition, 15 out of 20 participants described as irrelevant the question of using either one or two fingers: they invested almost the same effort at reaching the game's goal during both experiment. Furthermore, the others 5 participants found less useful two fingers than only one at the highest level of the game (4x4 grid). Based only on participants' feelings, it could be roughly said that the more elements are to be found in the scene, the less useful multi touches are.

This thesis was quite opposite to our intuition since it was intuitively hypothesized that two fingers should have been more functional due to the increased information rate that could be accessed. Results show that there is indeed an improvement when using two fingers at the first level of difficulty (2x2 grids). Whereas for the higher levels however, the t-test concluded that the equality of the average time spent to achieve the task (one finger versus two fingers) cannot be excluded: no clear advantage in using two fingers has been shown for localizing many elements so our primary intuition can not be confirmed.

As a comparison with the human visual system, vision is a fugitive and dynamic phenomenon. The central area of the human retina has the best resolution for approximately two degrees. Since our eyes are very frequently moving to analyze our environment or a given picture, and by analogy if we consider that our eyes play the role of a single pointing device, it is worth wondering whether a pointing device such as a finger would be sufficient and necessary to mimic in a crude manner the human visual system. The results obtained during our experiments suggest that the improvement factor when using two fingers could be small or negligible for medium/difficult tasks. Is training the key parameter that will allow individuals to improve the time required to achieve a difficult task by means of two fingers?

7. CONCLUSIONS

In this article, a global perception module for image exploration merging the usability of multi touch applications with auditory feedback has been introduced. This module is a component of the *See ColOr* system which aims at developing a mobility aid for blind and visually impaired individuals. As an attempt at determining the actual impact of mono and multi touch interaction onto the auditory feedback for sound localization's ability into a scene, a memory matching game of colors was implemented.

Experiments with 20 blindfolded participants were conducted in this investigation. These participants were tasked to find color matches into a grid of 2x2, 3x3 and 4x4 squares, using both one and two fingers in order to objectively measure the respective required times. A statistical analysis of the results of this experiment showed that for the cases of 2x2 grids, two fingers significantly improved the results with respect to one finger. In the case of 4x4 and 3x3 grids the t-test failed to confirm the hypothesis that two fingers are also more efficient. This last result could mean that for localizing many elements in an image, it might be more functional to use mono-touch rather than multi-touch.

8. REFERENCES

[1] Deville, B., Bologna, G., Vinckenbosch, M., and Pun, T. See Color: Seeing colours with an orchestra. Lalanne, D. and Kohlas, J. Eds. Lecture Notes in Computer Science, vol. 5440. Springer- Verlag, Berlin, Heidelberg, 2009, 251-279.18.

[2] Ordonez, C. Navarun, G. Barreto, A. Sound spatialization as a navigational aid in virtual environments. Proc. 6th CSI Conference –Orlando Fl, July 14-18, 2002: pag. 297-302.

[3] Gomez, J., Bologna, G., Pun, T., Color-Audio encoding Interface for Visual Substitution: See ColOr Matlab-based Demo. Proc. The 12th International ACM SIGACCESS Conference on Computers and Accessibility, Oct 25-27: pag. 216-218.

[4] Bologna, G., Deville, B., Pun T., and Vinckenbosch, M. Transforming 3D coloured pixels into musical instrument notes for vision substitution applications. Eurasip J. of Image and Video Processing, A. Caplier, T. Pun, D. Tzovaras, Guest Eds., (2007), Article ID 76204, 14 pages (Open access article).

[5] Bologna, G., Deville, B., Vinckenbosch, M., Pun, T. A Perceptual Interface for Vision Substitution in a Color Matching Experiment. In Proc. Int. Joint Conf. Neural Networks, Part of IEEE World Congress on Computational Intelligence, June 1-6, 2008, Hong Kong.

[6] Bologna, G., Deville, B. and Pun T. On the use of the auditory pathway to represent image scenes in real-time. Neurocomputing, 72, 839-849. Electronic publication: 2008, Article DOI: http://dx.doi.org/10.1016/j.neucom.2008.06.020 (2009).

[7] Bologna, G., Deville, B. and Pun T. Blind navigation along a sinuous path by means of the See ColOr interface. In Proceedings of the 3rd international Work-Conference on the interplay between Natural and Artificial Computation: Part II: Bioinspired Applications in Artificial and Natural Computation (Santiago de Compostela, Spain, June 22 - 26, 2009).

[8] Cronly-Dillon, J., Persaud, K., and Gregory, R.P.F. The Perception of Visual Images Encoded in Musical Form: a Study in Cross-Modality Information. In Proc. Biological Sciences (1999) 266, 2427–2433.

[9] Capelle, C., Trullemans, C., Arno, P., and Veraart, C. A real time experimental prototype for enhancement of vision rehabilitation using auditory substitution. IEEE T. Bio-Med Eng., 45 (1998), 1279–1293.

[10] Bregman, A. S. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press, 1990.

[11] Algazi, V.R., Duda, R.O., Thompson, D.P., and Avendano, C. The CIPIC HRTF atabase. In Proc. WASPAA'01, New Paltz, NY (2001).

[12] Begault, R. 3-D Sound for Virtual Reality and Multimedia. Boston A.P. Professional (1994), ISBN: 0120847353.

[13] Meijer, P.B.L. An experimental system for auditory image representations. IEEE Trans. Bio. Eng., 39 (2) (1992), 112–121.

[14] Kay, L. A sonar aid to enhance spatial perception of the blind: engineering design and evaluation. The Radio and Electronic Engineer, 44 (1974), 605–627.

[15] Gonzalez-Mora, J.L., Rodriguez-Hernandez, A., Rodriguez-Ramos, L.F., Dfaz-Saco, L., and Sosa, N. Development of a new space perception system for blind people, based on the creation of a virtual acoustic space. In Proc. IWANN’99 (1999), 321–330.

[16] McGookin, D., Brewster, S., Jiang, W. Investigating touchscreen accessibility for people with visual impairments. In Proceedings of the 5th Nordic conference on Human-Computer Interaction: Building Bridges, 298-307. (2008)

[17] MacKenzie, I. S., Soukore_, R. W. Text Entry for Mobile Computing: Models and Methods, Theory and Practice. Human-Computer Interaction, 17(2), 147. (2002)

[18] Keijiro U., Masamitsu, T., Yusuke, F. The Evaluation of Visually Impaired People’s Ability of Defining the Object Location on Touch-screen. Proc. The 12th International

ACM SIGACCESS Conference on Computers and Accessibility, Oct 25-27: pag. 116-118.

[19] Bologna, G., Deville, B., Gomez, J. and Pun T. Toward local and global perception modules for vision substitution. Neurocomputing, In Press, Corrected Proof, Available online 20 October 2010, ISSN 0925-2312, DOI: 10.1016/j.neucom.2010.07.022. http://www.sciencedirect.com/science/article/B6V10-518TDPG-2/2/28d78e81d7f798c73b8b8ec389f9708f)

[20] B. Durette, N. Louveton, D. Alleysson, J. Hérault (2008) Visuo-auditory sensory substitution for mobility assistance: testing TheVIBE. Workshop on Computer Vision for Applications for the Visually Impaired, 10th European Conference on Computer Vision (ECCV'2008), Marseille.