Non-visual virtual interaction: Can Sensory Substitution Generically Increase the Accessibility of Graphical Virtual reality to the blind?

Shachar Maidenbaum[1] Amir Amedi[2]

ELSC, Hebrew University of Jerusalem, Israel

IMRIC, Hebrew University of Jerusalem, Israel

ABSTRACT

Most of the content of Graphical virtual environments is currently visual, severely limiting their accessibility to the blind population. While several steps improving this situation have been made in recent years they are mainly environment specific and there is still much more to be done. This is especially unfortunate as VR holds great potential for the blind, e.g., for safe orientation and learning. We suggest in this position paper that Visual-to-audio Sensory Substitution Devices (SSDs) can potentially increase their accessibility generically fashion by sonifying the on-screen content regardless of the specific environment while allowing the user to capitalize upon his experience from other use of the device such as in the real world. We will demonstrate the potential of this approach using several recent examples from literature and from our own work.

Keywords: Universal access; Sensory substitution; blind; Virtual reality; Accessibility

Index Terms: B.4.2 Input/Output Devices: Image display; H.1.2 User/Machine Systems; H.5.2 User Interfaces: Auditory (non-speech) feedback;

1 MAIN CLAIMS AND POSITION

- Sensory substitution devices can and should be used as another tool in the accessibility toolbox for graphical virtual information.
- While more complex and difficult to master they offer the blind user much more access to visual information.
- There use can be generic across environments, and extend into the real world as well.

2 MOTIVATION

Virtual environments (VEs) are an important part of our life, but as their content is mainly visual they present significant accessibility issues for the blind.

This is especially unfortunate as VEs hold great potential for them. For example, since transfer of orientation and mobility knowledge for the blind between virtual and real environments is well established [1, 2], these worlds can be used for of familiarization with an unknown environment virtually before going there physically, e.g., via Google-Earth. This has the benefit familiarizing them with the environment while avoiding the risks associated with such exploration in the real world. Importantly, this can be done alone at one's own leisure, sparing the cost and availability problems of personal trainers and ultimately increasing independence. Similarly, VEs hold the potential to be safely used for a wide range of additional purposes ranging from education and scientific research to social interaction and games.

There have been several previous efforts to overcome this challenge as traditional accessibility tools are not well fit for graphical, as opposed to textual, information. While these new attempts have significantly improved the level of accessibility [3-6] there is still much work to be done – especially in aspects of the conveyed information, required hardware, availability and genericity across environments.

Here we outline an approach which interfaces with any virtual environment either by using simple parameters that can be extracted from any virtual environment with minimal or no interface to the virtual world's code, using a consistent sonification scheme which users are familiar with from other tasks, including in the real world. Specifically we suggest that this can be achieved using a Sensory Substitution Device (SSD, [7, 8]).

SSDs transfer the visual information in the scene to the user via different senses. They rely on the ability of the brain to reinterpret this information coming through a different sense and process it in a similar fashion to as if it were coming from the original sense. Users at first have to focus and explicitly translate this information, but as they gain experience the process becomes automatic, and some late blind report their sensation as becoming similar to 'seeing' [8, 9]. These devices have been used over the past decade for a wide variety of tasks, including in noisy cluttered real-world environments (see advanced user-studies in and some examples at [9, 10]). On the down side, SSDs have several known disadvantages, mainly the difficulty of learning to use them properly. These disadvantages have been somewhat mitigated over recent years as discussed in depth in [8, 11].

The key difference between this approach and other virtual accessibility ones such as Computer-vision and tagging approaches is their focus on offering the user relatively "raw" visual information to interpret as opposed to offering descriptions or the names of tagged/recognized items.

In the following section we will discuss the virtual use of traditional mobility aids through this approach, following with the use of a minimalistic-SSD (EyeCane) and then with the use of a full complex SSD (EyeMusic).

3 TRADITIONAL AIDS TRANSLATED INTO VIRTUAL ENVIRONMENTS

Several steps have already been made in the direction of offering basic visual information to blind users instead of high-level descriptions using traditional mobility aids.

The main aid used by the blind population for navigation and mobility is the white-cane. Recent years have seen the creation of virtual versions of this device for the purpose of virtual mobility [1, 12].

While these devices indeed significantly boost accessibility, they are still limited by the same factors limiting the visual information conveyed by the white-cane in the real world. Another issue has been the cost and availability of the required haptic hardware.

Emails: [1]Shachar.maidenbaum@gmail.com, [2]amira@ekmd.huji.ac.il

3rd IEEE VR International Workshop on Virtual and Augmented Assistive Technology (VAAT) 2015
23 March, Arles, France
978-1-4673-6518-5/15/$31.00 ©2015 IEEE

4 THE EYECANE

The EyeCane is a minimalistic-SSD which augments the White-Cane by giving the user information about the distance to 1-3 single points via auditory or tactile cues. Its main additions to the information from the White Cane are additional distance (5m instead of 1.2m) and angles including protection from upper-limb obstacle. Importantly it is non obtrusive, i.e. unlike the white-cane which must make physical contact with an object for its detection leading to many blind users to limit its use for fear of tripping people or damaging fragile objects the EyeCane uses IR beams. It has been used recently for tasks such as navigation in real environments, obstacle detection and avoidance and more [13, 14].

A virtual version of the EyeCane was then used in virtual environments with users hearing the exact same auditory cues for tasks such as simple navigation and shape identification [15-17]. When exploring virtual-EyeCane users navigation patterns it was found that they are far more similar to those of sighted users then to those of virtual-White-Cane users – e.g. users walked through the center of rooms, walked down the center of corridors instead of staying near walls etc. [18].

Implementing a virtual version of the EyeCane is relatively simple, as all it requires is access to the distance to several single points, typically easy to extract from the 3D mesh, and convert them to a set of auditory cues without the need for any dedicated hardware. However this still requires the creation of a dedicated plug-in for each environment – i.e. requires someone with access to the environments code to create an environment-specific version.

Additionally while single-point distance is already enough to improve accessibility to a surprising degree it is still a very limited parameter, does not convey information such as color or complex shapes, and requires a long active-scanning process to cover the whole scene.

See https://www.youtube.com/watch?v=av2LlcjEQsY for a video example of using the virtual-EyeCane.

5 THE EYEMUSIC

The EyeMusic is a visual-to-auditory SSD, described in detail in [19]. It has been used for various tasks in both real-world and experimental conditions ranging from identifying simple shapes to navigating in noisy environments and finding items in them such as a red apple in a bowl of green ones or a specific flavoured bottle of juice among other bottles of identical shape in the juice section of a real supermarket. A mobile version is freely available online (see www.brainvisionrehab.com).

The EyeMusic receives images from file, camera, or as in this case from sequential screen-shots, capturing the visual information from the screen. This image is segmented to a 30*50 matrix, and each cell is assigned a value by its location and color. The HSL values are clustered into 6 colors (5+black which is silent), each represented by a different musical instrument. The location is represented by a rising blues scale of musical notes, chosen to avoid masking harmonics and clashes, such that the higher the Y value the higher the note. The X axis is converted into time within the sonification – i.e. when played the image will sound the columns in a sweep line from left to right, similar to the sonification scheme established by the veteran vOICe SSD [20]. For a more detailed explanations about the transformation see [19]. Frame-rate here was controlled by the user and varied, but was typically at 1-frame-per-second.

Importantly the EyeMusic can receive input from different sources and sonify them the same way – for example instead of using a webcam or smartphone camera as in the examples mentioned above it can capture the on-screen image and offer the user the visual information from a virtual environment without the need for any hardware beyond standard audio output (speakers/headphone).

In a poster that will be presented in the adjoining IEEE VR conference we will describe the first in a series of experiments exploring the use of the EyeMusic for virtual tasks otherwise impossible without vision. There, 5 congenitally blind users who had no previous experience with virtual environments but were veteran EyeMusic users in real-world situations accessed simple virtual environments where they had to navigate to doors based on their shape or color, and in a second task based on their context (i.e. find the door surrounded by trees and navigate to it). We found that these users were able to easily adapt to the idea of virtual environments and complete the tasks with near-ceiling success rate of 98.3±3.7% ($p<1.6E-8$, vs. control) for target recognition and 91.6±8.3% ($p<2E-6$, vs. control) for navigating correctly to the exit. They also reported a strong feeling of immersion within it (1-5 scale, 5 fully immersed, 3.6±1.1) and that they greatly enjoyed the experience 4±1 (1-5 scale, 5 enjoyable).

A control group of ten blindfolded participants without the EyeMusic could not perform the experiment. They reported great frustration at the tasks (2.6±0.9 for enjoyment, 2±1 for immersion), and failed at almost all trials with scores of 21%±13 for recognition of their target and 15.5±14% for navigating correctly to the exit – and reported doing so by guessing and moving randomly.

The blind users' verbal feedback focused both on its potential application to their lives, mainly for navigation, and on its potential for making games accessible. All participants agreed enthusiastically to participate in future experiments using this approach - *"Could you build me a model of my home?"*; *"Ok, now I want you to make the same game, just with monsters and shooting, and let me take it home with me"*. The control group on the other hand reported great frustration – "*These tasks are impossible! What's the point?!*", "*I'm just guessing. I have no idea what I'm doing or where anything is*".

See https://www.youtube.com/watch?v=swYJlcsRyrQ for a video example of a blind user performing these tasks.

6 DISCUSSION

In the previous sections we have demonstrated that virtual assistive technology can significantly boost the accessibility of virtual environments and enable blind users to perform tasks previously impossible to them such as perceiving graphical information from simple virtual environments and navigating in them.

Importantly, in examples 2-3 this was done without any dedicated hardware and in example 3 this was done via SSD in a generic manner without a direct software connection to a specific virtual environment, and by capitalizing upon the skills and experience of users from using the EyeMusic for other tasks.

The enthusiastic survey responses demonstrate the existence of the problem and the potential of this approach. The negative reactions from the control group demonstrate how frustrating the accessibility situation can be without assistive devices.

A known problem with SSDs is their output's unpleasantness. The results of the pleasantness survey and users’ enthusiasm for future use indicate that users found the current version acceptable.

A main potential criticism for this approach is the use of simple environments – will they extend into more complex ones? While not yet backed by concrete evidence, which is currently being gathered, initial results from more complex tasks indicate that this is indeed possible. Additionally, since users are able to use SSDs for complex tasks in noisy real world environments this ability should be able to extend into virtual environments as well.

We described here the use of this approach for the blind, but SSDs can also add additional multisensory dimensions to the sighted. E.g. sighted users could interact with dedicated VEs without the need for a large physical screen, or augment reality with additional information (see examples in [3]).

7 Conclusion

We have demonstrated here the potential of using SSDs for increasing the accessibility of graphical VEs to the blind by showing that blind users are able to capitalize upon their SSD experience to perceive the VEs and interact with them. Thus, while there is still a long way to go this approach has the potential to significantly boost the options for non-visual interaction with computerized VEs.

8 Acknowledgements

The European Research Council Grant (310809), the Israel Science Foundation (grant no. 1684/08), a James S. McDonnel Foundation scholar award (no. 220020284), The Edmond and Lily Safra Center for Brain Sciences (ELSC) Vision center.

References

[1] O. Lahav, "Improving orientation and mobility skills through virtual environment for people who are blind: past research and future potential," presented at the ICDVRAT 2012, Laval, 2012.

[2] L. B. Merabet, E. C. Connors, M. A. Halko, and J. Sánchez, "Teaching the Blind to Find Their Way by Playing Video Games," *PloS one,* vol. 7, p. e44958, 2012.

[3] J. Sánchez and A. Tadres, "Audio and haptic based virtual environments for orientation and mobility in people who are blind," 2010, pp. 237-238.

[4] G. R. White, G. Fitzpatrick, and G. McAllister, "Toward accessible 3D virtual environments for the blind and visually impaired," 2008, pp. 134-141.

[5] J. Wood, M. Magennis, E. F. C. Arias, T. Gutierrez, H. Graupp, and M. Bergamasco, "The design and evaluation of a computer game for the blind in the GRAB haptic audio virtual environment," *Proceedings of Eurohpatics,* 2003.

[6] S. Trewin, V. L. Hanson, M. R. Laff, and A. Cavender, "PowerUp: an accessible virtual world," in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, 2008, pp. 177-184.

[7] P. Bach-y-Rita, *Brain mechanisms in sensory substitution*: academic Press, 1972.

[8] S. Maidenbaum, S. Abboud, and A. Amedi, "Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation," *Neuroscience & Biobehavioral Reviews,* 2013.

[9] J. Ward and P. Meijer, "Visual experiences in the blind induced by an auditory sensory substitution device," *Consciousness and cognition,* vol. 19, pp. 492-500, 2010.

[10] S. Maidenbaum, R. Arbel, G. Buchs, S. Shapira, and A. Amedi, "Vision through other senses: practical use of Sensory Substitution devices as assistive technology for visual rehabilitation," 2014 2014.

[11] G. V. Elli, S. Benetti, and O. Collignon, "Is There a Future for Sensory Substitution Outside Academic Laboratories?," *Multisensory Research,* vol. 27, pp. 271-291, 2014.

[12] D. W. Schloerb, O. Lahav, J. G. Desloge, and M. A. Srinivasan, "BlindAid: Virtual environment system for self-reliant trip planning and orientation and mobility training," 2010, pp. 363-370.

[13] G. Buchs, S. Maidenbaum, and A. Amedi, "Obstacle Identification and Avoidance Using the 'EyeCane'," presented at the EuroHaptics, 2014.

[14] H. Maidenbaum , Abboud , Chebat , Levy-Tzedek, Amedi A, "The “EyeCane”, a new Electronic Travel Aid for the Blind: Technology, Behavior & Swift Learning," *Restorative Neurology and Neuroscience,* 2014.

[15] S. Maidenbaum, S. Levy-Tzedek, D.-R. Chebat, and A. Amedi, "Increasing Accessibility to the Blind of Virtual Environments, Using a Virtual Mobility Aid Based On the," *PloS one,* vol. 8, p. e72555, 2013.

[16] S. Maidenbaum, R. Arbel, S. Abboud, D. Chebat, S. Levy-Tzedek, and A. Amedi, "Virtual 3D shape and orientation discrimination using point distance information," 2012.

[17] S. Maidenbaum, D. R. Chebat, S. Levy-Tzedek, and A. Amedi, "Depth-To-Audio Sensory Substitution for Increasing the Accessibility of Virtual Environments," in *Universal Access in Human-Computer Interaction. Design and Development Methods for Universal Access*, ed: Springer, 2014, pp. 398-406.

[18] S. Maidenbaum, S. Levy-Tzedek, D. R. Chebat, R. Namer-Furstenberg, and A. Amedi, "The Effect of Extended Sensory Range via the EyeCane Sensory Substitution Device on the Characteristics of Visionless Virtual Navigation," *Multisensory Research,* vol. 27, pp. 379-397, 2014.

[19] S. Abboud, S. Hanassy, S. Levy-Tzedek, S. Maidenbaum, and A. Amedi, "EyeMusic: Introducing a “visual” colorful experience for the blind using auditory sensory substitution," *Restorative neurology and neuroscience,* 2014.

[20] P. B. L. Meijer, "An experimental system for auditory image representations," *Biomedical Engineering, IEEE Transactions on,* vol. 39, pp. 112-121, 1992.