Vision through other senses: practical use of Sensory Substitution devices as assistive technology for visual rehabilitation

Shachar Maidenbaum*, Roni Arbel*, Galit Buchs, Shani Shapira & Amir Amedi

*Equal contribution
Department of medical Neuroscience, IMRIC & Edmond and Lili Safra brain center, ELSC
Hebrew University of Jerusalem
Jerusalem, Israel

Amira@ekmd.huji.ac.il
Brain.huji.ac.il

Abstract— **Visual-to-auditory Sensory Substitution Devices (SSDs) are non-invasive sensory aids that provide visual information to the blind via their functioning senses, such as audition. For years SSDs have been confined to laboratory settings, but we believe the time has come to use them also for their original purpose of real-world practical visual rehabilitation. Here we demonstrate this potential by presenting for the first time new features of the EyeMusic SSD, which gives the user whole-scene shape, location & color information. These features include higher resolution and attempts to overcome previous stumbling blocks by being freely available to download and run from a smartphone platform. We demonstrate with use the EyeMusic the potential of SSDs in noisy real-world scenarios for tasks such as identifying and manipulating objects. We then discuss the neural basis of using SSDs, and conclude by discussing other steps-in-progress on the path to making their practical use more widespread.**

Index Terms—**Sensory Substitution, Visual Rehabilitation** *(key words)*

I. Introduction

Blindness and severe visual impairments are a limiting condition affecting many millions worldwide, impairing their perception of the world and facing them with challenges in many everyday tasks. To address this problem the visual modality has been studied in depth but currently blindness is still at large incurable.

Our goal in the study of blindness is thus two-fold; to develop tools that will aid the blind in their every-day challenges, and in parallel to use these tools to learn about the neural mechanism in the absence of vision. Understanding these mechanisms is important, as it will aid the development of future vision-restoration procedures and tools, since the ability to see is not only a matter of the eyes, but also a matter of the brain.

One such approach is the use of Sensory Substitution Devices (SSDs). These are non-invasive human-machine interfaces which sense information via one modality and translate it into another modality through which the information is then transmitted to the user. In our context, this means sensing the raw visual information from the world via a camera, substituting it into auditory cues via a processing unit and then conveying the auditory information to the users via headphones so that his brain can re-interpret it and perceive the visual information.

The main goal of SSDs has been visual rehabilitation for the blind. Starting with the work of Bach-y-Rita, first with a tactile array and later on with an array of electrodes placed on the tongue (BrainPort,[1]), through the work of Meijer who created the visual-to-auditory SSD in widest use (The vOICe [2]), many attempts have been made to achieve this goal (see also [3-7]). However, despite their attempts SSDs are currently not widely used by the blind population and are considered to be mainly limited to controlled laboratory settings and research purposes due to reasons such as the cumbersomeness and cost of previous SSDs, limited information offered, failure to demonstrate behavioral results in the real world and finally clashes with traditional neuroscience theories limiting the potential of this approach in general.

In a recent review published in Neuroscience & Biobehavioral Reviews [8], Maidenbaum and Amedi attempted to analyze these reasons that have prevented their adoption for practical visual rehabilitation. They show that many of these reasons have changed for the better over the past few years, and that they may now be overcome. Here we will attempt to demonstrate some of these changes using the EyeMusic SSD (previously described in [9], see expansion in Methods), which offers the user whole scene information about shape, color and location. We added several features to this device to directly address some of these issues. Specifically:

A. Offered information:

The information conveyed by the EyeMusic is of the location shape and color of the whole-scene, allowing the user to not just identify a single item, but to understand the full context and scene in front of him. The use of the color parameter, which was lacking in veteran SSDs, is especially important as it plays an important role in scene segmentation and perception which has been shown many times in the past [10]. (see fig 1). Additionally, the new version of the EyeMusic now offers an increased resolution to enable the detection of more delicate visual features. See "Methods A & C" for expansion.

B. Cumbersomeness & Cost:

Historically, many SSDs have suffered from difficulties to procure, assemble and physically carry and use them. The new version of the EyeMusic can be freely downloaded and used unto iOS smartphone platforms (a notable previous similar exception is the veteran vOICe SSD which offers similar accessibility via Android platforms) thus limiting these problems (see fig 2). See "Methods B" for expansion.

C. Behavioral results:

Most previous behavioral results have been demonstrated under laboratory settings (see criticism [11]). As we will demonstrate here, EyeMusic can be used also in real world scenarios. See "Behavior & Results" for expansion.

Introduction summary

Thus, in this work we will describe the EyeMusic SSD, and demonstrate its use in real world scenarios for tasks otherwise impossible for the blind including object identification and manipulation in real and noisy environments such as a grocery store.

II. METHODS & MATERIALS

A. The EyeMusic Algorithm

The EyeMusic works by sequentially acquiring an image from the camera, creating an auditory pixel based waveform of the whole scene (known as a Soundscape), and conveying it to the user via headphones (See fig 1, and see [9] for expanded algorithm description).

The substitution processing includes the following steps. First the image is down sampled to the EyeMusic resolution (discussed in the next section). Then each pixel in the image is assigned an auditory value – it's y axis-value is translated into a musical note on a hexatonic scale (the higher the pixel the higher the note), and it's color value is translated into the Timbre of that note after being clustered to one of the following six –white, blue, red, yellow & green (Choir, brass instruments, Reggae organ, string instruments, and Rapman's reed organ, respectively) and black, represented by silence.

The image is then played from left to right with a sweep-line approach such that at any point in time a column of pixels is played simultaneously. The speed of the sweep-line can be adjusted by the user, and typically for new users will be around 2 seconds.

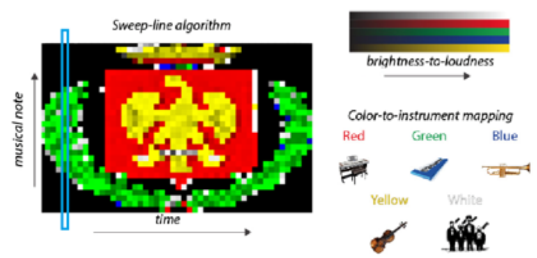

Figure 1: Illustrates the principles for the EyeMusic audio-to-visual transformation on the symbol of Palermo: y-axis to musical note in a hexatonic scale, x-axis to time via a sweep line method going from left to right, brightness to loudness and color to timbre (musical instrument) according to the mapping of white, blue, red, yellow & green to Choir, brass instruments, Reggae organ, string instruments, and Rapman's reed organ, respectively and black, represented by silence.

B. Setup & Mobile application

The EyeMusic does not require dedicated hardware, and can be downloaded for free (see fig 2, download at http://tinyurl.com/oe8d4p4). In an additional setup, used in the behavioral task presented here, the processing unit was a standard laptop, and the camera was mounted on a pair of sunglasses (see fig 2).

Figure 2: The EyeMusic setups. *(A) Mobile application setup (B) Head mounted glasses + laptop setup*

C. Increased Resolution

The resolution of the EyeMusic has been increased from 40X24 [9] to 50X30 while striving to preserve the pleasantness of the cues in order to allow perception of more complicated stimuli and more detailed visual information.

This increase in resolution is not trivial from a technical perspective as in order to accomplish it we had to add additional musical notes while avoiding note-combinations that would cause unpleasantness or masking. This included replacing the pentatonic scale with a carefully chosen hexatonic scale by adding F# to each octave for a total of 30 notes for each of the five colors, and a careful choice of Timbre.

Enlarging the resolution on the X axis requires additional columns, resulting in a cumbersome increase in the duration of each frame but was done to preserve the ratio between the two axis, chosen to comply with the current dimensions of retinal prosthesis (see discussion), and to increase the amount of information per frame.

D. Participants

This experiment included a congenitally blind individual, OB. OB underwent weekly EyeMusic training as part of a longitudinal study in our lab [12]. Accordingly, he already had several dozen hours of training with the low resolution EyeMusic algorithm and was able to identify visual features translated to soundscapes from 2D images.

This experiment was approved by the Hebrew University ethics committee and OB signed his informed consent.

III. Behavior & Results

A. Training for the real-life tasks

Generalization of the low-resolution EyeMusic principles to the new high-resolution device was intuitive and only required a short introduction to the new musical scale.

Training with the EyeMusic is not just about interpreting the soundscapes correctly. For example, unlike the sighted, congenitally blind individuals are not accustomed to move their heads in order to improve their sensory input. Another example is that in order to produce accurate reaching movement we will demonstrate here they need to develop the SSD equivalence of hand-eye coordination.

A familiarization session with 3D objects and the camera took place in the lab prior to the testing. In this session the participant was asked to identify colorful objects placed on a table in front of him, and to reach for a specific object. Explicit information was provided regarding basic visual principles such as perspective and the participant was asked to reach for a specific object without tactilely examining each of the objects on the table.

Upon grasping the desired object, OB was instructed to bring the grasped object closer to the camera, so that he could "hear" the object in greater detail for self-evaluation purposes.

B. Performance in the tasks

Congenitally blind participant OB was requested to use the EyeMusic for several real world tasks otherwise impossible for him. He was asked to discriminate different colors of vegetables, candy and bottled juice placed randomly on shelves in a grocery store. He was then asked to reach for the object of his preference, and to examine the grasped object in detail, to make sure he was correct. Finally, he was asked to recognize different notes of money. Emphasis was also put on keeping the camera parallel to the floor in order to improve orientation.

Figure 3: Demonstration of real world behavioral tasks using the EyeMusic SSD. *(A)Recognizing bottled juices (B) Recognizing candy (C) Recognizing fruit (D) Recognizing bank notes*

OB indeed successfully used the EyeMusic for all of these tasks (See video 1 and screenshots in fig 3).

It should be noted that for congenitally blind individuals, such tasks are largely inaccessible since many of the items are only discriminable by their colors; For example, different flavors of bottled juice are tactilely the same and are only discriminated externally by their color and label. Moreover these objects are often located in a random order on a shelf, so that the blind user has to be able to create a general understanding of the entire scene in order to guide his movement towards the desired object.

IV. DISCUSSION

A. General

We have demonstrated here that the new EyeMusic allows the congenitally blind to perform colorful object identification and reaching movements in real-life ecological environment.

These results join other recent behavioral results from our lab and others using various SSDs for orientation and Mobility, expression recognition and more [13-18]. However, it should be emphasized that these results were achieved only after a large amount of dedicated training and hard work, which at this point is still not scale-able enough.

Additionally, we have described how the use of a mobile application of the EyeMusic can help circumvent the problems of cumbersomeness, availability & cost, and how more information including parameters such as color from the whole-scene can be conveyed to the user in higher resolution.

Each of these separately, and all together demonstrate an example for some of the reasons we believe SSDs are now more mature and ready to be aimed at use for practical rehabilitation.

B. Whole Scene perception

In contrast to other visual aids, SSDs such as the EyeMusic give the user whole-scene information. For blind people this is different from their everyday information acquisition process, as they are used to tactilely examine objects separately in order to identify and grasp the desired object. Therefore, the ability to first perceive the gist of a scene, concentrate on a specific object, and then direct a movement to this object is not trivial for blind users, and OB reported the concept as being new to him.

Furthermore, the EyeMusic gives the blind the unique opportunity to self-evaluate their performance and correct their movement according to the online feedback they receive from the device. This feedback allowed OB to make sure that he indeed grasped the correct object, and if not replace it on the table and try again. The ability for self-evaluation is crucial for SSDs aimed for real-life use, as it allows the blind to be independent from an external examiner. Thus great emphasis was put not only on the accuracy of the movement and on grasping the right object, but also on being able to self-evaluate the performance and correct the mistake if one occurred. Hand-eye coordination improved during this training session, and self-evaluation process improved.

C. The auditory cues & Shared sensory space

Does the signal from the SSD interfere with other sensory inputs? Can it be integrated into the users' general sensory picture of the world?

Even though the device was tested in a noisy environment OB was able to both listen to the EyeMusic and communicate with the instructor at the same time. It appears than even though the EyeMusic is an auditory device, attention to other auditory inputs such as a conversation is preserved.

There have been several recent results showing that the information arriving via SSDs can be integrated with other senses [19] and even with motor control [20, 21], which match our participants abilities to close the sensory-motor loop and SSD-guide their actions using the EyeMusic.

D. The Neurobiological basis of using SSDs

We have demonstrated here the practical potential of SSDs, but what is happening "under the hood" in our brains? This question is especially important as according to accepted theories of neuroscience such as "critical periods" and "cross-modal plasticity" the visual parts of the brain should not be available for their original purpose, and thus it is unclear how such behavioral results could have been achieved.

To answer this question let us take a step back and look at the visual cortex. It is well established that the visual cortex is comprised of different functional areas, each processing different aspects of vision. E.g. the FFA shows preference for faces, VWFA for visual representation of language etc.

Surprisingly, several of these basic brain regions which were once considered "visual" were recently shown using SSDs to retain their function even without visual experience despite missing the critical periods. Thus, the LOC for tactile & SSD object location and perception, the VWFA for reading braille or reading via the vOICe, MT for non-visual motion (reviewed in [5, 8, 22]. Recently we have even shown that the EBA, which specializes in identification of body postures, is similarly recruited [23]. These results show that the visual information is processed also in the traditional "visual" areas even when received via other senses with SSDs.

These results have led to the hypothesis of the brain as task-oriented and sensory-modality independent, or in other words a "task machine" [5, 8]. The brain regions can still perform their specific task if they receive the relevant information, regardless of the sensory channel in which it was sensed. Thus, the lack of visual experience should not limit the task-specialization of the visual system, despite its use for various functions in the blind, and may still be able to retain its functional properties via other sensory-modalities. This is very encouraging for the potential for visual rehabilitation.

V. Future work & the steps to practical visual rehabilitation

As we have shown here, practical use of SSDs for visual rehabilitation is possible, but there yet remains work to make it available. We believe that the key steps towards this important goal must include the creation of structured training programs, dedicated virtual and real world training environments, the development of more new SSDs using state-of-the-art technology, and most importantly, the promotion of scientific efforts directly aimed at optimizing the visual rehabilitation process using SSDs, which has taken the back seat in the last few decades in favor of their use for research.

One possible direction we are currently proceeding in is to combine a post-operation system combining an SSD with retinal prostheses to create a combined device (which we call "Vision rehabilitation device"; VRD). This system will include a camera consistently capturing images of the surroundings and a processing unit. This unit will convert the visual information into (i) an auditory SSD representation and (ii) a neural stimulation conveyed by the prostheses' electrodes. Information about the surroundings would thus be received in parallel from the prostheses as well as from the SSD. In such a device, the SSD would serve as a "sensory interpreter" providing explanatory input for the visual signal arriving from the invasive device. This dual synchronous information is expected to significantly increase the rate of rehabilitation. At a later stage the SSD can be used to provide input beyond the maximal capabilities of the prostheses or capabilities it does not possess such as adding color information via EyeMusic. This concept is also applicable to cases of late sight-restoration.

VI. Conclusions

We believe that the time has come to return SSDs to their original goal of practical visual rehabilitation. As part of this, we introduced here the new version of the EyeMusic SSD which avoids problems with previous SSDs such as availability, cost & cumbersomeness, and demonstrated its use for several real world ecological tasks.

We have also briefly discussed the neurobiological basis for using SSDs despite potential problems suggested by traditional neuroscience and some other steps being taken on the way to practical use of these devices.

SUPPLEMENTAL MOVIE

Click here to download supplemental movie

ACKNOWLEDGMENT

This work was supported by a European Research Council grant to AA (grant number 310809); The Charitable Gatsby Foundation; The James S. McDonnell Foundation scholar award (to AA; grant number 220020284); The Israel Science Foundation (grant number ISF 1684/08).

REFERENCES

1Bach-y-Rita, P., *Brain mechanisms in sensory substitution*. 1972: academic Press.

2 Meijer, P.B.L., *An experimental system for auditory image representations.* Biomedical Engineering, IEEE Transactions on, 1992. **39**(2): p. 112-121.

3. Capelle, C., et al., *A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution*. Biomedical Engineering, IEEE Transactions on, 1998. **45**(10): p. 1279-1293.

4. Lenay, C., et al., *Sensory substitution: Limits and perspectives.* Touching for knowing, 2003: p. 275-292.

5. Reich, L., S. Maidenbaum, and A. Amedi, *The brain as a flexible task machine: implications for visual rehabilitation using noninvasive vs. invasive approaches.* Current Opinion in Neurology, 2011.

6. Maidenbaum, S. and A. Amedi, *Applying plasticity to visual rehabilitation in adulthood*, in *Plasticity in sensory systems*, L.R.H. Jennifer Steeves, Editor. 2012, Cambridge University Press.

7.Hanneton, S., M. Auvray, and B. Durette, *The Vibe: a versatile vision-to-audition sensory substitution device.* Applied Bionics and Biomechanics, 2010. **7**(4): p. 269-276.

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

9. Abboud, S., et al., *EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution.* Restorative neurology and neuroscience, 2014.

10. Yip, A.W. and P. Sinha, *Contribution of color to face recognition.* Perception-London, 2002. **31**(8): p. 995-1004.

11. Loomis, J.M. and R.L. Klatzky, *Navigating without vision: Basic and applied research.* Optometry & Vision Science, 2001. **78**(5): p. 282.

12. Striem-Amit, E., et al., *Reading with Sounds: Sensory Substitution Selectively Activates the Visual Word Form Area in the Blind.* Neuron, 2012. **76**(3): p. 640-652.

13. Ward, J. and P. Meijer, *Visual experiences in the blind induced by an auditory sensory substitution device.* Consciousness and cognition, 2010. **19**(1): p. 492-500.

14. Chebat, D.R., et al., *Tactile-'visual'acuity of the tongue in early blind individuals.* NeuroReport, 2007. **18**(18): p. 1901.

15. Chebat, D.R., et al., *Navigation with a sensory substitution device in congenitally blind individuals.* NeuroReport, 2011. **22**(7): p. 342.

16. Bologna, G., et al., *Toward local and global perception modules for vision substitution.* Neurocomputing, 2010.

17. Kim, J.K. and R.J. Zatorre, *Can you hear shapes you touch?* Experimental brain research, 2010. **202**(4): p. 747-754.

18. Reynolds, Z. and B. Glenney, *Interactive Training for Sensory Substitution Devices.* 2009.

19. Wright, T.D., et al., *Wheres Wally? Audio-visual mismatch directs ocular saccades in sensory substitution.* Seeing and Perceiving, 2012. **25**(s1): p. 61-61.

20. Levy-Tzedek, S., et al., *Cross-sensory transfer of sensory-motor information: visuomotor learning affects performance on an audiomotor task, using sensory-substitution.* Scientific Reports, 2012. **2**.

21. Levy-Tzedek, S., et al., *Fast, accurate reaching movements with a visual-to-auditory sensory substitution device.* Restorative Neurology and Neuroscience, 2012. **30**(4): p. 313-323.

22. Collignon, O., et al., *Sensory rehabilitation in the plastic brain.* Progress in brain research, 2011. **191**: p. 211.

23. Striem-Amit, E. and A. Amedi, *Visual Cortex Extrastriate Body-Selective Area Activation in Congenitally Blind People "Seeing" by Using Sounds.* Current Biology, 2014. **24**(6): p. 687-692.