Restorative Neurology and Neuroscience xx (20xx) x–xx
DOI 10.3233/RNN-150592
IOS Press

Integration and binding in rehabilitative sensory substitution: Increasing resolution using a new Zooming-in approach

Galit Buchs[a,b], Shachar Maidenbaum[b,c], Shelly Levy-Tzedek[d,e] and Amir Amedi[a,b,c,f,*]
[a]*Department of Cognitive Science, Faculty of Humanities, Hebrew University of Jerusalem, Hadassah Ein-Kerem, Jerusalem, Israel*
[b]*The Edmond and Lily Safra Center for Brain Research, Hebrew University of Jerusalem Hadassah Ein-Kerem, Jerusalem, Israel*
[c]*Department of Medical Neurobiology, Institute for Medical Research Israel-Canada, Faculty of Medicine, Hebrew University of Jerusalem, Hadassah Ein-Kerem, Jerusalem, Israel*
[d]*Recanati School for Community Health Professions, Department of Physical Therapy, Ben Gurion University of the Negev, Beer-Sheva, Israel*
[e]*Zlotowski Center for Neuroscience, Ben Gurion University of the Negev, Beer-Sheva, Israel*
[f]*Sorbonne Universités UPMC Univ Paris 06, Institut de la Vision Paris, France*

Abstract.

Purpose: To visually perceive our surroundings we constantly move our eyes and focus on particular details, and then integrate them into a combined whole. Current visual rehabilitation methods, both invasive, like bionic-eyes and non-invasive, like Sensory Substitution Devices (SSDs), down-sample visual stimuli into low-resolution images. Zooming-in to sub-parts of the scene could potentially improve detail perception. Can congenitally blind individuals integrate a 'visual' scene when offered this information via different sensory modalities, such as audition? Can they integrate visual information – perceived in parts - into larger percepts despite never having had any visual experience?

Methods: We explored these questions using a zooming-in functionality embedded in the EyeMusic visual-to-auditory SSD. Eight blind participants were tasked with identifying cartoon faces by integrating their individual components recognized via the EyeMusic's zooming mechanism.

Results: After specialized training of just 6–10 hours, blind participants successfully and actively integrated facial features into cartooned identities in 79 ± 18% of the trials in a highly significant manner, (chance level 10%; rank-sum $P < 1.55\text{E-}04$).

Conclusions: These findings show that even users who lacked any previous visual experience whatsoever can indeed integrate this visual information with increased resolution. This potentially has important practical visual rehabilitation implications for both invasive and non-invasive methods.

Keywords: Sensory substitution, vision rehabilitation, action-perception, motor control, active sensing

1. Introduction

Coherent perception of a full visual scene is reached through integration of its smaller visual details and sub-parts. Thus the ability to comprehend a complex whole from its parts is an important feature of the visual system in sighted people, and is a requirement that must be addressed in attempts for visual rehabilitation.

Additionally, the resolution that is available to blind users via most rehabilitation approaches, whether invasive, like implanted bionic-eyes (Chuang, Margo,

---

[*] Corresponding author: Amir Amedi, Tel.: +972 2 675 8331; Fax: +972 2 6758602; E-mail: amira@ekmd.huji.ac.il.

ISSN 0922-6028/15/$35.00

& Greenberg, 2014; Dagnelie, 2012; Weiland, Cho, & Humayun, 2011), or non-invasive methods, like Sensory Substitution Devices (SSDs) (Bach-y-Rita, Kaczmarek, Tyler, & Garcia-Lara, 1998), is currently very low, forcing these devices to significantly down-sample the visual stimuli into low-resolution images. This down-sampling process is generally a necessity as either due to the number of actuators the devices employ - an array of respectively 6 × 10 electrodes for the Argus II retinal implant (Humayun, Dorn, Cruz, Dagnelie, & Sahel, 2012), or 20 × 20 for the BrainPort SSD (Nau, Bach, & Fisher, 2013) – or to the inherent bandwidth of the sensory channel. For sight the input is received by 126 million photoreceptors (Meyer, 2002) while in contrast, for audition the input is absorbed by 3500 inner hair-cells (Kim, 1984).

Zooming-in to sub-parts of the scene could potentially significantly improve users' perception of details. But can blind individuals perform similar integration of a 'visual' scene when the input is provided via a different sensory modality, namely auditory soundscapes? Can congenitally blind individuals integrate visual information – perceived in parts - into larger percepts even though they previously never had any visual experience?

In this work we explore these questions using SSDs, non-invasive interfaces which translate information from one sense to another (Bach-y-Rita et al., 1998; Deroy & Auvray, 2012; Elli, Benetti, & Collignon, 2014; Proulx, Brown, Pasqualotto, & Meijer, 2014) and have seen extensive use, especially for research (Auvray, Philipona, O'Regan, & Spence, 2007; Collignon, Champoux, Voss, & Lepore, 2011; Haigh, Brown, Meijer, & Proulx, 2013; Renier & De Volder, 2010) but also to a lesser extent in practical real world situations (Meijer, 2015; Ward & Meijer, 2010). Specifically, we used the EyeMusic Visual-to-auditory SSD, which conveys whole-scene visual information via audition (see (Abboud, Hanassy, Levy-Tzedek, Maidenbaum, & Amedi, 2014, 2012) for more information) with a dedicated zooming-in mechanism developed for the purpose of this study (and subsequently integrated into the EyeMusic's software following the success of this experiment, see below). Users tapped a touch tablet to select and focus on a smaller sub-area within the image (referred to here as the "scanning window"), which is transformed into an audio signal with the EyeMusic's maximal resolution (in this experiment 24 × 40), thus ultimately enabling a finer-grained perception of the whole.

The idea of using a zooming-in method with SSDs is not entirely new. A zooming-in mechanism which sonifies a zoomed-in version of the center of the image instead of the full image was added to The vOICe (Meijer, 1992, 2015) SSD, but to the best of our knowledge has not been used for research in general or for exploring integration in particular. Several previous studies employed similar zooming-in mechanisms to compensate for resolution limitations. Maucher et al. (Maucher, Meier, & Schemmel, 2001) used a haptic display that moved on a large board and gave information of what was beneath it, thus avoiding the need of a large tactile display and minimizing scanning time though requiring dedicated haptic hardware. In the auditory modality, Arno and colleagues (Arno et al., 2001) used a pen on a tablet to obtain information of a window around the fixation point. In both studies the participants recognized one continuous shape with the help of the device but were not required to form a holistic complex representation from an assembly of its components.

Here, we tested the use of a zoom-in approach in increasing local resolution for recognition tasks of composite objects. For this purpose, users were presented with 10 cartoon characters demonstrating different emotions using cartooned facial expressions. Looking at the whole scene they could not discern the characters' specific features due to the low resolution, but by zooming-in on specific features they could be recognized (see Fig.1 and video). For a first simplified stage of testing this ability, users were presented with different cartoon figures, since for recognizing faces there is a need to both identify specific facial components and to create a holistic representation from them (Sergent, 1984; Yarbus, 1967).

2. Methods

2.1. Sensory substitution devices (SSDs)

SSDs are non-invasive visual rehabilitation devices that transform information usually conveyed by vision into audition and then convey it to the user. This approach relies upon the theory that the interpretation process that occurs in the brain does not require the information to come from a specific sensory input channel and can thus be processed as if coming from the eyes (Bach-y-Rita & W. Kercel, 2003; Maidenbaum, Abboud, & Amedi, 2014). Users at first have

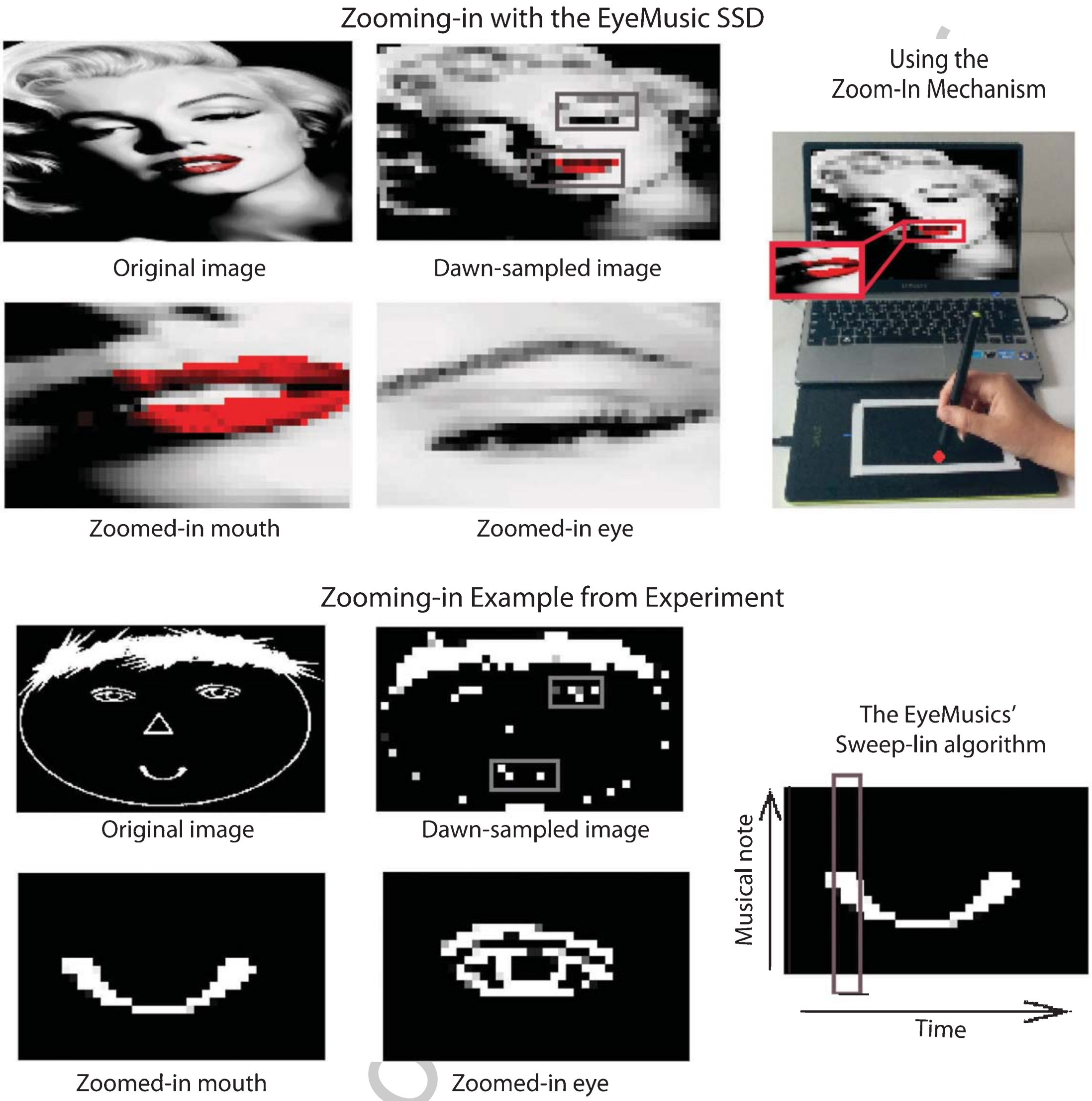

Fig. 1. Demonstration of the zooming-in concept: motivation & procedure.

to explicitly translate this information, but as they gain experience the process becomes automatic, and several late blind users report their perception as resembling to some extent ‘seeing’(Maidenbaum, Abboud, et al., 2014; Proulx, Stoerig, Ludowig, & Knoll, 2008; Ward & Meijer, 2010). SSDs can potentially aid blind individuals in numerous situations. For example they have recently been used for finding specific items in a cluttered room, navigation, obstacle detection and avoidance, object recognition and many more (Maidenbaum, Abboud, et al., 2014; Ward & Meijer, 2010).

2.2. *The EyeMusic*

The EyeMusic Visual-to-auditory SSDs, conveys visual information via audition (Abboud et al., 2014; S Levy-Tzedek et al., 2012). Shapes, location and color are converted into sound. X-axis information is conveyed via a left-to-right sweep-line while height is preserved through pitch, where high-pitched musical notes represent high locations in the image (see Fig. 1). Five different musical instruments are used to convey five colors: white, red, blue, yellow and green.

Black is represented by silence. Shades of colors can be differentiated through the volume.

The EyeMusic has been used before in several other projects, concerning the identification of shapes and objects (Abboud et al., 2014; Maidenbaum, Arbel, Buchs, Shapira, & Amedi, 2014) exploration of the meta-modal basis of number representation in the brain (Abboud, Maidenbaum, Dehaene, & Amedi, 2015), and swift cross-sensory transfer of spatial information between vision and SSDs even with minimal training in an implicit, unconscious manner (Shelly Levy-Tzedek et al., 2012).

It is important to note that the EyeMusic does not require any additional hardware beyond standard headphones and can be downloaded at no cost as an iOS App (http://tinyurl.com/oe8d4p4) and for Android on Google Play (http://tinyurl.com/on6lz4e), thereby increasing the availability and decreasing the cost for future potential users.

For this experiment, the EyeMusic program was adapted to include a "zoom" mode. In this mode, it was possible to select a specific area in a picture that was sonified at a higher resolution. The size of the scanning window around the selected point was fixed to 1/16 of the original picture, which was down-sampled to a resolution of 24×40 pixels (in the EyeMusic 2 it is down-sampled to a resolution of 30×50 pixels). For points close to the edge, the scanning window was snapped to the grid. The participants chose their points of interest by touching the desired location on the touch screen (see Fig. 1 and movie).

2.3. Equipment

Two touch tablets were used in this experiment: The intuos4, model number PTK-480, Wacom, and the BAMBOO PEN model number CTL-470, Wacom. In the first tablet, the picture occupied the entire screen. In the second, the active display area was marked with tape.

Users heard the audio stimuli via headphones.

2.4. Participants

A total of 8 blind individuals participated in this experiment, seven with congenital blindness and one who lost her eye sight at the age of one year.

The average age of the participants was 32 ± 6.2 years (mean ± SD). 7 of the participants had previous experience with the EyeMusic; the 8th had had an introduction to the device shortly before participating in the study.

2.5. Ethics

The experiment was approved by the ethical committee of the Hebrew university. All participants signed their informed consent.

2.6. Stimuli

We created ten cartoon figures that varied with respect to the shape of their eyes, nose and hair. There were three types of eyes: filled ovals, with eyebrows and baseless triangles (see supplementary materials Fig. 1, a-c). Likewise, there were four variations for the nose: a vertical straight line, a vertical straight line with a diagonal line at the edge, a triangle and no nose at all (see supplementary materials Fig. 1, d-f). The hair types were: regular hair, bangs, ponytails and a combination of bangs with ponytails (see supplementary materials Fig. 1, g-j). Any two cartoon characters varied in at least one component. Emotional expression of the cartoons was portrayed by the mouth. Four different emotions were represented: happiness, apathy, surprise and sadness. Each of the emotions was represented by a unique mouth shape (see supplementary materials Fig. 1, k-n).

Each of the components was designed so that it could fit into a single 'scanning window' (i.e., 1/16th of the full image).

2.7. Procedure

This experiment was composed of two parts: Specialized training followed by testing. The goal of the training phase was to instruct the trainee in the use of the zoom mechanism in recognizing the ten characters. In the second part, the trainees were tested on their recognition of the stimuli.

2.7.1. Training

All participants received approximately six hours of specialized training prior to testing. One participant (S3) underwent 5 additional hours of training due to technical issues with the tablet.

Participants were tasked with recognizing the faces of individual characters by combining their different features. Additionally, the concept of the scanning window was introduced through a demonstration of the principle with a paper frame that the trainees held and moved on the experimenter's face or on a three-dimensional model of a face. This enabled the participants to learn the concept of finding fixation points according to the component they want to hear.

Once the trainee grasped this idea, identifying the cartoon figures using the zoom mechanism was learned. The individual facial features were progressively introduced, separately and in gradually increasing combinations (see Fig. 2).

2.7.2. Testing

The combination of 10 characters with 4 possible emotions generated a pool of 40 possible stimuli. The test included a series of 20 cartoon figures, randomly selected from this pool of 40 for each participant. Each character was repeated twice and each emotion five times. Due to a technical error one of the subjects encountered the emotion *apathy* six times in the experiment, and the emotion surprise four times.

There was a short break between the first and the second set of ten cartoon figures. Upon presentation of each of the characters, the trainee was required to recognize it while using the zoom mechanism. Recognition was accomplished by specifying the name, the emotion and the components of the character (type of eyes, nose and hair). The recognition task was performed without any feedback from the experimenter (see movie). The time required to recognize each character was measured.

A total of 3 stimuli, out of the 160 stimuli presented to all 8 participants were disqualified for inclusion in the study, due to technical errors. In one case duration was not saved (s3, trial 7). In the other two cases (s8, trials 6&17) there was a program malfunction.

3. Results

All statistical analysis of the results was performed using a rank-sum test.

The participants' average success rate for a full correct answer was 64 ± 26% (mean ± SD). A full correct answer included naming the character correctly, recognizing the emotion and appropriately describing the components of the character. This success rate compares to a 2.5% level of chance, arrived at through the ability to name the character and recognize its emotion. Success in this necessitates identification of each of the facial features. Success rates were significantly higher than chance ($p < 1.55\text{E-04}$) (see Fig. 3A).

Also, in isolating the features that generated a full correct answer, a significant difference was found between the average success rates of the participants for a particular feature and the success rate predicted by chance level. In naming the character, the participants succeeded in 79+ 18%, as compared to a 10% level of chance ($p < 1.55\text{E-04}$) (see Fig. 3A). In recognizing the correct emotion, the average success rate was 82 +20% compared to a chance level of 25% ($p < 1.55\text{E-04}$) (see Fig. 3A). Additionally, the participants succeeded in appropriately describing all of the facial features of the characters (eyes, nose and hair types) in 80 ± 13% of the cases as opposed to a chance level of 2% ($p < 1.55\text{E-04}$) (see Fig. 3A).

The overall average success rate of the participants in recognizing the eyes was 89 ± 7%. The average success rate in recognizing the different types of eyes were respectively 85 +13%, 88 ± 8% and 100% for eyes shaped like a full oval, eyes with an eyebrow and eyes that look like a triangle without a base (see Fig. 3B).

The overall average success rate of the participants in recognizing the nose was 93 ± 7%. The average success rate in recognizing the different types of noses was 97 ± 9%, 81 +37%, 94 ± 18% and 94 ± 8% for a nose shaped like a vertical straight line, a vertical straight line with a diagonal at the edge, a triangle and no nose, respectively (see Fig. 3C).

The overall average success rate of the participants in recognizing the hair type was 96 ± 7%. The average success rate in recognizing the different types of hair was 99 ± 2%, 81 ± 37%, 87 ± 23% and 93 ± 18% for a regular hair type, bangs, ponytails and a combination of bangs with ponytails, respectively (see Fig. 3D).

For the different emotions, the average success rate in recognizing the different types: happiness, sadness, surprise and apathy was 90 ± 21%, 70 ± 29%, 71 ± 30% and 95 ± 9%, respectively (see Fig. 3E).

Time: On average, it took the participants 150 ± 51 seconds to provide an answer for the stimuli presented.

4. Discussion

Participants' success rates were highly significant above chance level for both component recognition and their integration into a holistic perception, thus demonstrating that this zoom-in method can indeed facilitate these two abilities.

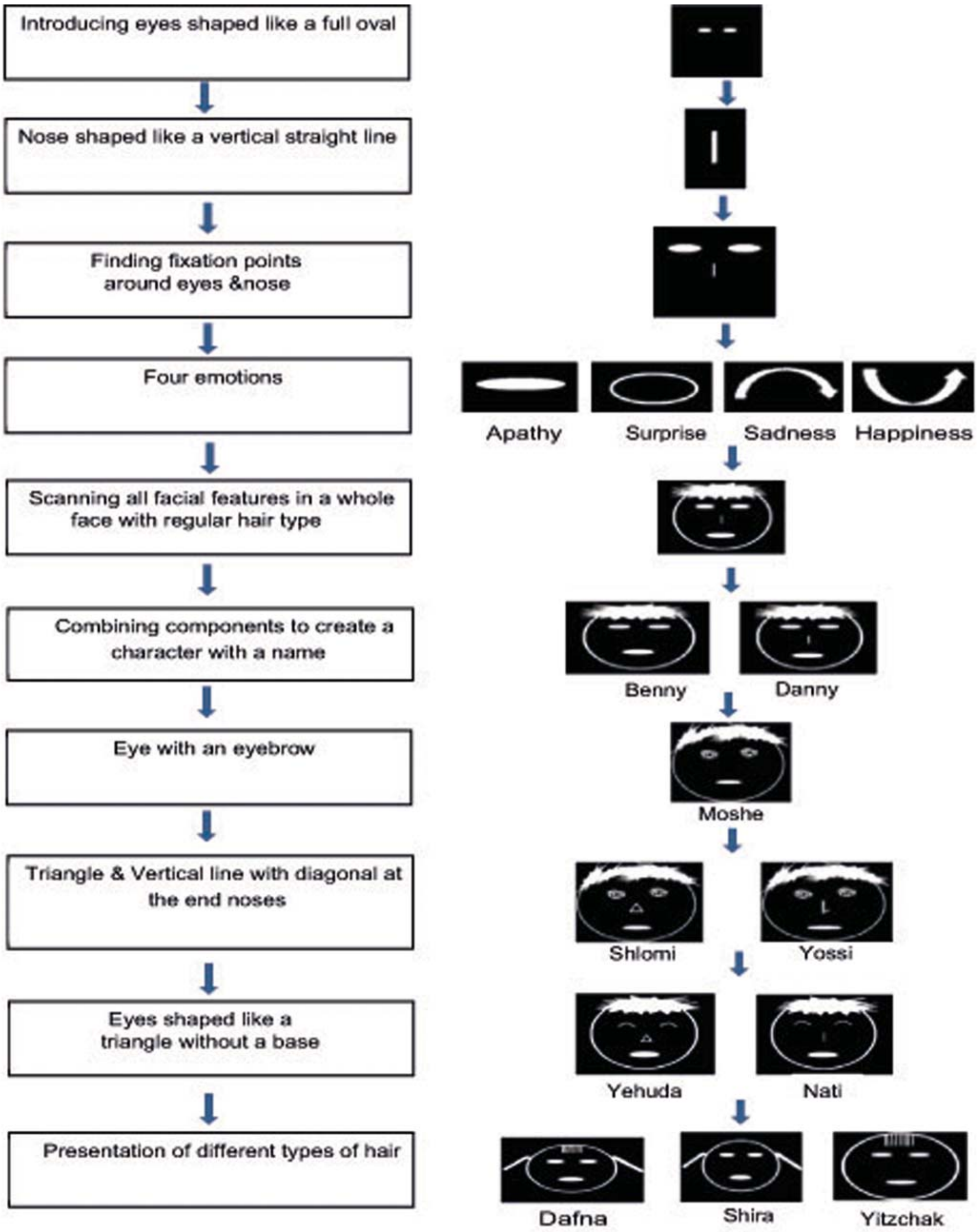

Fig. 2. Learning procedure.

The success rate of the participants in providing a fully correct answer was significantly higher than chance although there were noticeable differences in the identification of the different components thus proving that indeed the integration of individual components into a holistic representation is possible, and can further increase along with the increase in recognition of the individual components. This finding suggests that in their mind's eye blind people can indeed integrate visual information to create a whole percept out of its parts using a zoom-in method, even for users that did not have any visual experience whatsoever in their life. It suggests that the brain has the capacity to learn this ability later in life and via a com-

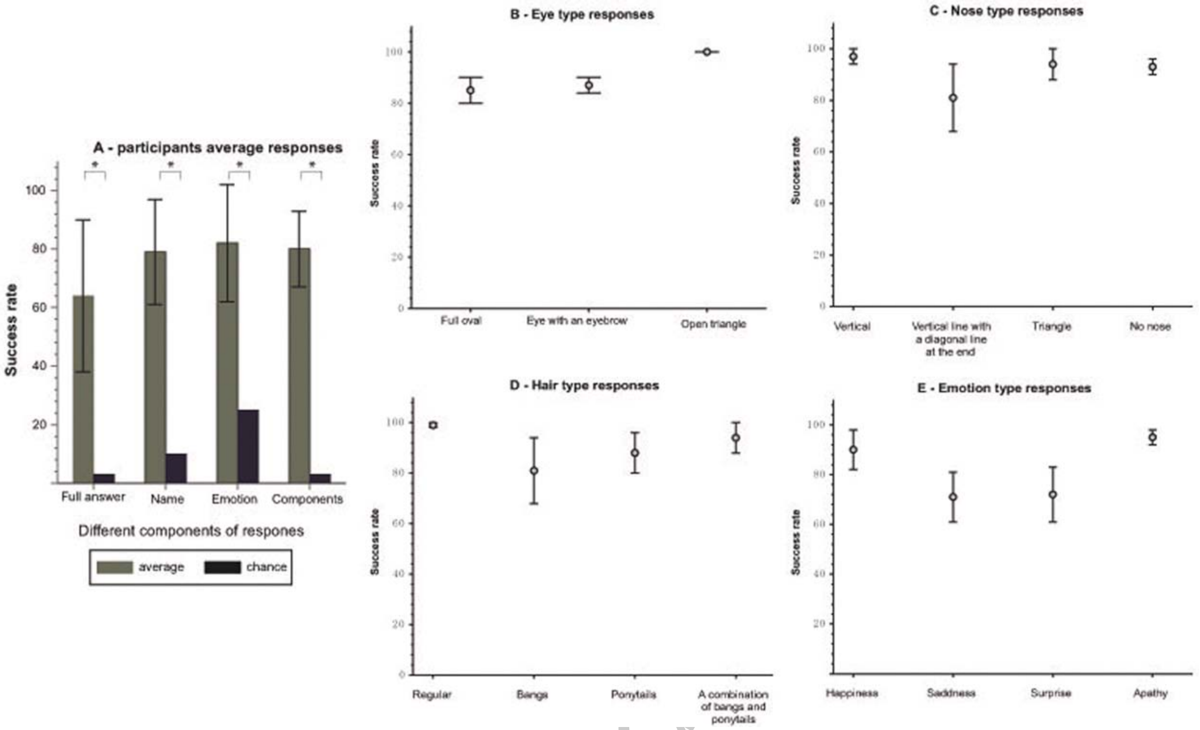

Fig. 3. Results – A: participants' average success rate B: eye type. C: nose type. D: hair type. E: emotion type.

pletely new sensory input – if it learns to use the input. It also suggests that these mechanisms have a strong multisensory compatibility.

The overall success rate of identification of all the components was significantly higher than chance, although there are noticeable differences between the success of participants in the identification of the various components. For the eyes, this may be due to the similarity of the full oval eye to that of the eye with an eyebrow. The nose comprised of a vertical line with a diagonal at the end was successfully identified in 81% of the cases compared to 90% for the other nose types. This misclassification was probably unduly influenced due to the relatively low number of repetitions in the total stimuli (appeared in 16 of the total 160 stimuli). For the hair, the participants recognized correctly bangs 81% of the time. It is important to note, that when misclassified, this hair type was perceived as a combination of ponytails with bangs. Therefore, this error isn't due to difficulties in identifying the bangs. These misclassifications can be overcome with additional training.

From a practical point of view we suggest this zoom-in-and-integrate approach to be extremely useful for both invasive and non-invasive approaches such as visual prostheses and SSDs, respectively. In addition to being helpful in raising input resolution, this method gives the user control of what to focus on in an active manner, thereby potentially benefiting from the advantages offered by active sensing (Saig, Gordon, Assa, Arieli, & Ahissar, 2012; Stiles, Zheng, & Shimojo, 2015). This joins previous recent results both in virtual environments (Maidenbaum, Abboud, Buchs, & Amedi, 2015) and in the real world (Ward & Meijer, 2010) demonstrating the increased potential of new SSDs for practical visual rehabilitation. We have recently showed that the addition of 'color' information into stimuli presented by the EyeMusic significantly enhances the 'visual' acuity (Shelly Levy-Tzedek, Riemer, & Amedi, 2014). This addition of color is being used also in other SSDs as well (D. Gomez, Bologna, & Pun, 2014; Hamilton-Fletcher & Ward, 2013). We anticipate that using both the 'zoom' option and added meaningful color information will further improve the acuity level that can be reached with the EyeMusic, and other SSDs, even for high level demanding task such as identification of other people's faces and important social information such as emotions and vantage point (such as where a person is looking).

More importantly, the ability to integrate parts into a whole image is critical for current visual retinal prostheses such as the Argus II, as the visual field of these devices is very limited, and in order to perceive even simple information, such as a single letter, their users must continually scan the visual scene and integrate the visual input signals into a whole (da Cruz et al., 2013). In fact many of the retinal prosthesis users find it very difficult to learn this integrating process (Stronks & Dagnelie, 2014) – while in our experiment all participants managed to learn this integrating mechanism. This suggests that SSDs can be a very important tool in the visual rehabilitation process of retinal prostheses patients as it can help them learn the skill of integrating features. This is in addition to other potential advantages of teaching these implant recipients to use SSDs such as increasing their resolution and adding color information when used in conjunction with the prosthesis (Maidenbaum, Abboud, et al., 2014; Sella, Reiner, & Pratt, 2014; Van Atteveldt, Murray, Thut, & Schroeder, 2014).

An additional benefit rising from this experiment is the non-visual use of a tablet. This is important in and of itself since touch displays are becoming a major part in our lives, and are currently still far from being accessible to blind and visually impaired users (Watanabe, Yamaguchi, & Minatani, 2015).

Future work will include the use of this approach on a larger participant pool, the use of more complicated stimuli and an adaptable zooming-in window size as well as tests of its use in the real world. These tests will also include late blind participants, whom we expect will have even higher performance scores thanks to their previous visual experience.

Supplementary material

The supplementary figure and video are available in the electronic version of this article. http://dx.doi.org/10.3233/RNN-150592.

Supplementary video: https://goo.gl/0mhTkh.

This video demonstrates the use of the zooming-in method to perceive components in higher resolution and integrating this information into a combined whole.

Acknowledgments

Authors wish to thank Eliana Harow and Idit Lozowick-Gabay for their help in running the experiment, and Sami Abboud for inspiring this work.

This work was supported by 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, The Helmsley Charitable Trust through the Agricultural, Biological and Cognitive Robotics Center of Ben-Gurion University of the Negev.

References

Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., & Amedi, A. (2014). EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution. *Restorative Neurology and Neuroscience*, *32*(2), 247-257. doi:10.3233/RNN-130338

Abboud, S., Maidenbaum, S., Dehaene, S., & Amedi, A. (2015). A number-form area in the blind. *Nature Communications*, *6*.

Arno, P., Vanlierde, A., Streel, E., Wanet-Defalque, M.-C., Sanabria-Bohorquez, S., & Veraart, C. (2001). Auditory substitution of vision: Pattern recognition by the blind. *Applied Cognitive Psychology*, *15*(5), 509-519. doi:10.1002/acp.720

Auvray, M., Philipona, D., O'Regan, J.K., & Spence, C. (2007). The perception of space and form recognition in a simulated environment: The case of minimalist sensory-substitution devices. *PERCEPTION-LONDON-*, *36*(12), 1736.

Bach-y-Rita, P., Kaczmarek, K.A., Tyler, M.E., & Garcia-Lara, J. (1998). Form perception with a 49-point electrotactile stimulus array on the tongue: A technical note. *Journal of Rehabilitation Research and Development*, *35*, 427-430.

Bach-y-Rita, P., & Kercel, W.S. (2003). Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences*, *7*(12), 541-546. doi:10.1016/j.tics.2003.10.013

Chuang, A.T., Margo, C.E., & Greenberg, P.B. (2014). Retinal implants: A systematic review. *British Journal of Ophthalmology*, *98*, 852-856. doi:10.1136/bjophthalmol-2013-303708

Collignon, O., Champoux, F., Voss, P., & Lepore, F. (2011). Sensory rehabilitation in the plastic brain. *Progress in Brain Research*, *191*, 211-231. doi:10.1016/B978-0-444-53752-2.00003-5

Gomez, D.J., Bologna, G., & Pun, T. (2014). See ColOr: An extended sensory substitution device for the visually impaired. *Journal of Assistive Technologies*, *8*(2), 77-94.

Da Cruz, L., Coley, B.F., Dorn, J., Merlini, F., Filley, E., Christopher, P., et al. (2013). The Argus II epiretinal prosthesis system allows letter and word reading and long-term function in patients with profound vision loss. *British Journal of Ophthalmology*, *97*(5), 632-636. doi:10.1136/bjophthalmol-2012-301525

Dagnelie, G. (2012). Retinal implants: Emergence of a multidisciplinary field. *Current Opinion in Neurology*, *25*(1), 67-75. doi:10.1097/WCO.0b013e32834f02c3

Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: A new perspective on sensory substitution. *Frontiers in Psychology*, *3*.

Elli, G.V, Benetti, S., & Collignon, O. (2014). Is there a future for sensory substitution outside academic laboratories? *Multisensory Research*, *27*(5-6), 271-291.

Haigh, A., Brown, D.J., Meijer, P., & Proulx, M.J. (2013). How well do you see what you hear? The acuity of visual-to-auditory sensory substitution. *Frontiers in Psychology*, *4*.

Hamilton-Fletcher, G., & Ward, J. (2013). Representing Colour Through Hearing and Touch in Sensory Substitution Devices. *Multisensory Research*, *26*, 503-532. doi:10.1163/22134808-00002434

Humayun, M.S., Dorn, J.D., Cruz, L., Dagnelie, G., & Sahel, J. (2012). Interim results from the international trial of Second Sight's visual prosthesis. *Ophthalmology*, *119*(4), 779-788. doi:10.1016/j.ophtha.2011.09.028.Interim

Kim, D. (1984). Functional roles of the inner and outer hair cell subsystems in the cochlea and brainstem. In C.I. Berlin (Ed.), *Hearing science: Recent advantages* (pp. 241-261). San Diego: College-Hill Press.

Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S., & Amedi, A. (2012). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device. *Restorative Neurology and Neuroscience*, *30*(4), 313-323. doi:10.3233/RNN-2012-110219

Levy-Tzedek, S., Novick, I., Arbel, R., Abboud, S., Maidenbaum, S., Vaadia, E., & Amedi, A. (2012). Cross-sensory transfer of sensory-motor information: Visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Scientific Reports*, *2*, 949. doi:10.1038/srep00949

Levy-Tzedek, S., Riemer, D., & Amedi, A. (2014). Color improves "visual" acuity via sound. *Frontiers in Neuroscience*, *8*, 1-7. doi:10.3389/fnins.2014.00358

Maidenbaum, S., Abboud, S., & Amedi, A. (2014). Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience & Biobehavioral Reviews*, *41*, 3-15.

Maidenbaum, S., Abboud, S., Buchs, G., & Amedi, A. (2015). Blind in a Virtual World: Using Sensory Substitution for Generically Increasing the Accessibility of Graphical Virtual Environments. In *IEEE VR*.

Maidenbaum, S., Arbel, R., Buchs, G., Shapira, S., & Amedi, A. (2014). Vision through other senses: Practical use of Sensory Substitution devices as assistive technology for visual rehabilitation. In *Control and Automation (MED), 2014 22nd Mediterranean Conference of* (pp. 182-187). IEEE.

Maucher, T., Meier, K., & Schemmel, J. (2001). An interactive tactile graphics display. In *Signal Processing and its Applications* (pp. 190-193). Kuala Lumpur. doi:10.1109/ISSPA.2001.949809

Meijer, P.B. (1992). An experimental system for auditory image representations. *IEEE Transactions on Bio-Medical Engineering*. doi:10.1109/10.121642

Meijer, P.B. (2015). seeingwithsound webcite.

Meyer, È. (2002). Retina implantm -a bioMEMS challenge, *98*, 1-9.

Nau, A., Bach, M., & Fisher, C. (2013). Clinical Tests of Ultra-Low Vision Used to Evaluate Rudimentary Visual Perceptions Enabled by the BrainPort Vision Device. *Translational Vision Science & Technology*, *2*(3), 1. doi:10.1167/tvst.2.3.1

Proulx, M.J., Brown, D.J., Pasqualotto, A., & Meijer, P. (2014). Multisensory perceptual learning and sensory substitution. *Neuroscience & Biobehavioral Reviews*, *41*, 16-25.

Proulx, M.J., Stoerig, P., Ludowig, E., & Knoll, I. (2008). Seeing "where" through the ears: Effects of learning-by-doing and long-term sensory deprivation on localization based on image-to-sound substitution. *PLoS ONE*, *3*(3). doi:10.1371/journal.pone.0001840

Renier, L., & De Volder, A.G. (2010). Vision substitution and depth perception: Early blind subjects experience visual perspective through their ears. *Disability and Rehabilitation. Assistive Technology*, *5*, 175-183. doi:10.3109/17483100903253936

Saig, A., Gordon, G., Assa, E., Arieli, A., & Ahissar, E. (2012). Motor-sensory confluence in tactile perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(40), 14022-14032. doi:10.1523/JNEUROSCI.2432-12.2012

Sella, I., Reiner, M., & Pratt, H. (2014). Natural stimuli from three coherent modalities enhance behavioral responses and electrophysiological cortical activity in humans. *International Journal of Psychophysiology*, *93*(1), 45-55. doi:10.1016/j.ijpsycho.2013.11.003

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology (London, England: 1953)*, *75*(2), 221-242. doi:10.1111/j.2044-8295.1984.tb01895.x

Stiles, N.R.B., Zheng, Y., & Shimojo, S. (2015). Length and orientation constancy learning in 2-dimensions with auditory sensory substitution: The importance of self-initiated movement. *Frontiers in Psychology*, *6*, 1-13. doi:10.3389/fpsyg.2015.00842

Stronks, H.C., & Dagnelie, G. (2014). The functional performance of the Argus II retinal prosthesis. *Expert Review of Medical Devices*, *11*(1), 23-30. doi:10.1586/17434440.2014.862494

Van Atteveldt, N., Murray, M.M., Thut, G., & Schroeder, C.E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, *81*(6), 1240-1253. doi:10.1016/j.neuron.2014.02.044

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

Watanabe, T., Yamaguchi, T., & Minatani, K. (2015). Advantages and Drawbacks of Smartphones and Tablets for Visually Impaired People——Analysis of ICT User Survey Results——. *IEICE TRANSACTIONS on Information and Systems*, *98*(4), 922-929.

Weiland, J.D., Cho, A.K., & Humayun, M.S. (2011). Retinal prostheses: Current clinical results and future needs. *Ophthalmology*, *118*(11), 2227-2237. doi:10.1016/j.ophtha.2011.08.042

Yarbus, A.L. (1967). *Eye movements and vision*. (L. A. Rigg, Ed.). New York: Plenum press.