Luminophonics experiment: A user study on visual sensory substitution device

Shern Shiou Tan, Tomas Maul, Neil R Mennie

Loss of vision is a severe impairment to the dominant sensory system. It often has a catastrophic effect upon the sufferer, with knock-on effects to their standard of living, their ability to support themselves, and their care-givers lives. Research into visual impairments is multi-faceted, focusing on the causes of these debilitating conditions as well as attempting to alleviate the daily lives of affected individuals. One of the methods is through the usage of sensory substitution device. Our proposed system, Luminophonics, focuses on visual to auditory cross modalities information conversions. A visual to audio sensory substitution device a type of system that obtains a continual stream visual inputs which it converts into corresponding auditory soundscape. Ultimately, this device allows the visually impaired to visualize the surrounding environment by only listening to the generated soundscape. Even though there is a huge potential for this kind of devices, public usage is still minimal (Loomis, 2010). In order to promote the adoption from the visually impaired, the overall performance of these devices need to be improved in terms of soundscape interpretability, information preservation and listening comfort amongst other factors. Luminophonics has developed 3 type of prototypes, which we have used to explore different ideas pertaining to visual to audio sensory substitution. In addition to these, one of the prototypes has been converted to include depth information using time of flight camera. Previously, an automated measurement method is used to evaluate the performance of the 3 prototypes (Tan, 2013). The results of the measurement cover the effectiveness in terms of interpretability and information preservation. The main purpose of the experiment reported herein, was to test the prototypes on human subjects in order to gain greater insight on how they perform in real-life situations.

Luminophonics experiment: A User Study on Visual Sensory Substitution Devices

Shern Shiou Tan[1], Tomas Henrique Bode Maul[1], and Neil Russell Mennie[2]

[1]School of Computer Science, The University of Nottingham Malaysia Campus
[2]School of Psychology, The University of Nottingham Malaysia Campus

ABSTRACT

Loss of vision is a severe impairment to the dominant sensory system. It often has a catastrophic effect upon the sufferer, with knock-on effects to their standard of living, their ability to support themselves, and their care-givers lives. Research into visual impairments is multi-faceted, focusing on the causes of these debilitating conditions as well as attempting to alleviate the daily lives of affected individuals. One of the methods is through the usage of sensory substitution device. Our proposed system, Luminophonics, focuses on visual to auditory cross modalities information conversions. A visual to audio sensory substitution device a type of system that obtains a continual stream visual inputs which it converts into corresponding auditory soundscape. Ultimately, this device allows the visually impaired to visualize the surrounding environment by only listening to the generated soundscape. Even though there is a huge potential for this kind of devices, public usage is still minimal (Loomis et al., 2010). In order to promote the adoption from the visually impaired, the overall performance of these devices need to be improved in terms of soundscape interpretability, information preservation and listening comfort amongst other factors.

Luminophonics has developed 3 type of prototypes, which we have used to explore different ideas pertaining to visual to audio sensory substitution. In addition to these, one of the prototypes has been converted to include depth information using time of flight camera. Previously, an automated measurement method is used to evaluate the performance of the 3 prototypes (Tan et al., 2013). The results of the measurement cover the effectiveness in terms of interpretability and information preservation. The main purpose of the experiment reported herein, was to test the prototypes on human subjects in order to gain greater insight on how they perform in real-life situations.

Keywords: Image Processing, Computer Vision, Auditory Display, Image Sonification

INTRODUCTION

A sensory substitution device is a type of system that converts information from one input sensor to another output sensor while preserving the important functionality of the original sensor. Luminophonics is the outcome of a research project that focuses only on visual to auditory sensory substitution (VASS). Using an input device such as a camera, visual signals are captured and then sent to a processing unit. The processing unit will then generate the corresponding soundscapes based on different visual-to-auditory conversion algorithms. This algorithm map properties in the visual signal to the corresponding audio properties. The output soundscape is then transferred to an audio output device such as speaker or a pair of headphones.

One of the main motivation behind Luminophonics is to promote the usage of such sensory substitution devices for the visually impaired. This category of people commonly suffers from vision loss that is beyond repair with medication or visual correction. According to the World Health Organization (2014), there are 285 million people who are visually impaired worldwide with 39 million completely blind. Most importantly, about 90% of the world's visually impaired live in low-income settings. Visual to auditory sensory substitution device is a great approach for such people due to its characteristic of being low cost (Maidenbaum et al., 2014) and the fact that it does not involve any surgical procedure. By using such system, they can visualize their surroundings by listening to and interpreting the soundscape which contributes towards allowing them to lead a normal life.

We have built a total of 3 main prototypes to examine the details of cross-modalities conversions.

The prototypes implement different algorithms and parameterizations and therefore function differently from each other. In order to measure their relative effectiveness, in the past we have applied automated methods to measure the relative performances of the 2D prototypes in terms of interpretability and also the information preservation of the conversion. Although these methods provide a fast way to evaluate sensory substitution devices, conducting experiments on human subjects is still essential because the quantitative method does not cover other crucial subjective elements, such as human reactions towards the system, sound preferences, feasibility in real-life situation, amongst others. These Psychology-based factors can only be measured in a controlled environment with human subjects.

Apart from providing rudimentary colour and shape vision, our 3D prototypes grants users the ability to perceive depth through stereoscopic vision. Human make use of the depth information frequently in our daily activities especially during the navigation. There are not many visual to auditory sensory substitution system that integrates depth in their conversion process. One of the system that uses depth information is See ColOr developed by Bologna et al. (2007) where a stereoscopic camera is used to estimate depth by triangulation. Our approach differs from others, since we include depth information by implementing a time-of-flight (TOF) depth camera on top of our prototype. TOF is a special technique to capture the surrounding producing a depth map. This camera operates by modulating visible and near-infrared radiation in each pixel which will then measures the amplitude, offset and phase of the received radiation Lange and Seitz (2001). By using the measurement, a depth map can be constructed in real-time. Thanks to this depth map, we can extend one of our prototypes with a new and critical type of information.

PROTOTYPES

To date, Luminophonis consists of 4 different prototyppes. All four prototypes have been designed with different ideas especially the latest prototype which incorporates depth information in the conversion process. This experiments reported in this paper were designed to measure the effectiveness of each prototype in the scenario of navigation. On top of that, the experiments can answer the question of whether depth information is indeed important by comparing the results of our 2D and 3D prototypes.

Prototype 1

This prototype was the first Luminophonics sensory substitution system developed. As demonstrated in (Tan et al., 2010), this system defers from many other similar system because Prototype 1 employs image processing methods from Bradski (2000) to transform the input image before generating the soundscape. Input image frames are simplified using K-means clustering to reduce the noise and also bring out the salient regions. A technique called connected-component labeling is used to create multiple blobs out of the simplified regions. Soundscapes are then generated by processing the blobs from left to right. This swiping technique picks up the properties such as colour, location and size. The visual cues differentially affect audio properties which in turns will be combined to form a soundscape. The colour of the blobs is associated with sound timbre, whereas intensity of the colour is associated with sound pitch.

Prototype 2

This prototype is similar to Prototype 1 which uses a swiping technique. However, Prototoype 2 swipes from top to bottom where the topmost blob will be heard in the soundscape first and the bottom most blob will be heard last. Addition to that, Prototype 2 makes use of differential stereo volume to translate horizontal blob positions into sound properties. For each blob, if the volume at the left is higher then the blob is situated slightly to the left or vise-versa. On the other hand, if the user hears the sound of the blob equally in both ear, the blob is located directly in the middle region. For other properties like colour and shade, Prototype 2 react like Prototype 1.

Prototype 3

Prototype 3 departs from using image processing to simplify the image frames. Prototype 3 revisit a popular visual to auditory sensory substitution device called vOICe by Meijer (1992). Meijer uses very simple technique in his system for auditory image representations where image frames are pixelated using a simple filter. The advantage of using this technique consist of processing speed and the fact that a decent amount of visual information is preserved. One of the main disadvantage is that the user needs more time to learn how to interpret the resulting soundscapes. Prototype 3 improves on vOICe mainly by

incorporating colour information in the representation. By applying this enhancement, Prototype 3 has more information because vOICe only uses grayscale image. From the results reported in Tan et al. (2013), Prototype 3 scores better compared to vOICe in terms of information preservation and also interpretation.

Prototype 4

Depth is one of the fundamental types of low-level information extracted by our visual system. Thanks to stereopsis, we are able to perceive depth and differentiate near and far objects. It is not common for current sensory substitution system to incorporate depth information due to the issue of cacophony. One of the underlying goals of the research in this paper was to incorporate depth information into one of our 2D prototypes, given how crucial depth is to navigation. From the results of our performance measurement (Tan et al., 2013), Prototype 3 ranked as the best out of all of three 2D prototypes. Therefore, we have decided to incorporate depth information into Prototype 3. A time-of-flight camera by DepthSense was used as the depth camera of choice in our setup. Using the resulting depth map, this prototype slices the image into 3 regions distinguished by depth, with the near region ranging from 0 to 3m, mid region ranging from 3m to 5m, and the far region for anything beyond 5m. When using this prototype, the users need to manually select which region they want to focus on. From the experiment experiment reported here, it is possible to conclude whether whether this type of depth implementation is ideal for navigation purposes.

EXPERIMENT DESIGN

Human use vision to perform various functions including object recognition, navigation, assessing distance and balance control. Unfortunately, visual to auditory sensory substitution devices are still lacking behind real human vision. Most sensory substitution devices developed so far attempt to tackle only a limited scope of visual functionalities. Possible causes of this situation, include cross-modal functional incompatibilities and differences in sensory bandwidth. Therefore, instead of testing the prototypes against real life scenario, we focused the experiment design on navigational-based tasks.

The experiments were designed with the intention to compare all 4 of our prototypes in the domain of navigation. The main questions we hoped to answer include:-

- how important depth information is in helping to navigate

- how long do users need to learn how to use the system

- how each prototype performs when applied to real-life navigation based scenarios

Objectives

In general, there are a few main objectives this experiment wanted to achieve. With the introduction of depth prototype, there is a need to find out whether the incorporation of depth did indeed improve the usability of Luminophonics devices. This experiment aimed to provide valuables insight on the implementation of depth by comparing the performance of the prototype against the other prototypes that did not use depth. Moreover, user feedback was deemed equally important in obtaining a deeper understanding of this type of prototype.

During navigation, it is suggested that human do generate memory maps (both spatial and episodic memory) to help them maneuver themselves (Burgess et al., 2002). If the human subject do indeed develop memory maps during the course of navigation, then journeys involving paths or regions experienced previously would be easier, which in turn should shorten traveling time. In one of our experiments, every human subject in order to complete a task, needs to move forward to attain some goal and then return back to the starting point, in order for us to measure the time difference of both trips. The outcome of this experiment allows us to determine which prototypes are better in helping the users in generating memory maps.

Before the experiment, the users went through a one-time tutorial session on how to use the Luminophonics devices. Users were taught the basic functionalities of the devices and also how to use the system to navigate. This procedure was intentionally designed so that no prototype was unfairly advantaged relative to the other prototype. Without a tutorial session, users would need to learn how to use the system during experimental tasks and therefore the performance of each prototype would depend additionally on its specific learning curve.

Apart from the above mentioned memory effect, the experiment aimed to measure the effectiveness of different prototypes in directly guiding the subject through their journey. In real life, we use vision to avoid obstacles and navigate along a winding road. In our experiment, there will be several obstacles were placed around the navigation course for the users to evade by listening to the soundscape from the prototype. An effective prototype will guide users correctly hence reduce the number of the obstacles knocked.

Navigation Course

Figure 1. Front View of Experiment Site

Figure 2. Rear View of Experiment Site

A fixed navigation course with obstacles is designed for every human subject to walk through. This course was similar to the path used in the experiment conducted by SeeColor by Bologna et al. (2009). Both Figure 1 and Figure 2 shows the front view and the rear view of the experiment site. From the figures, it is apparent that the experiment was setup in an indoor flat classroom. The purpose of using an indoor venue instead of an outdoor field was partly due to the consistency of lighting condition. In a controlled environment with consistent lighting throughout the day, the differences experienced by users is minimized for all test cases. Audio interference can also be reduced so that the users can listen closely only to the generated soundscape. On top of that, indoor room is safer to the human subjects when they are performing their tasks blindfolded.

Referring to Figure 3, the red wavy line indicates the path the user needed to walk along. Obstacles were placed next to, and all along, this path. The green flag indicates the end of the path. When the user reaches the green flag, they were required to perform a task involving object recognition.

As seen at Figure 1, a 90m long path was drawn with yellow floor-tape stuck on to the floor. Yellow colour was chosen because it is easier to be detected and recognized for the first-time user. Barricade floor-tape with alternating red and white stripe were taped on both left and right side of the yellow tape. This tape was intended as an indicator for users when they were not focusing at the correct angle. The yellow path guided users from the origin to the target point, where they were expected to perform a specific task. After finishing the task, users were required to return back to the origin. Time was measured for each direction the user navigated, i.e: from origin to target, and from target back to origin.

A total of 10 paper boxes were placed randomly throughout the path. They acted as obstacles that the users needed to avoid. For every box that was knocked was counted which in turn affected the result of user and also the performance of the prototype in use. Besides walking and avoiding obstacle, at the end of the path the user needed to select a balloon of a specific colour designated by the experimenter. The subject was required to select one out of three coloured balloons (i.e. red, yellow or blue). This task was designed to measure the object recognition functionality of each prototype. The time needed for each user to complete this task was also recorded.
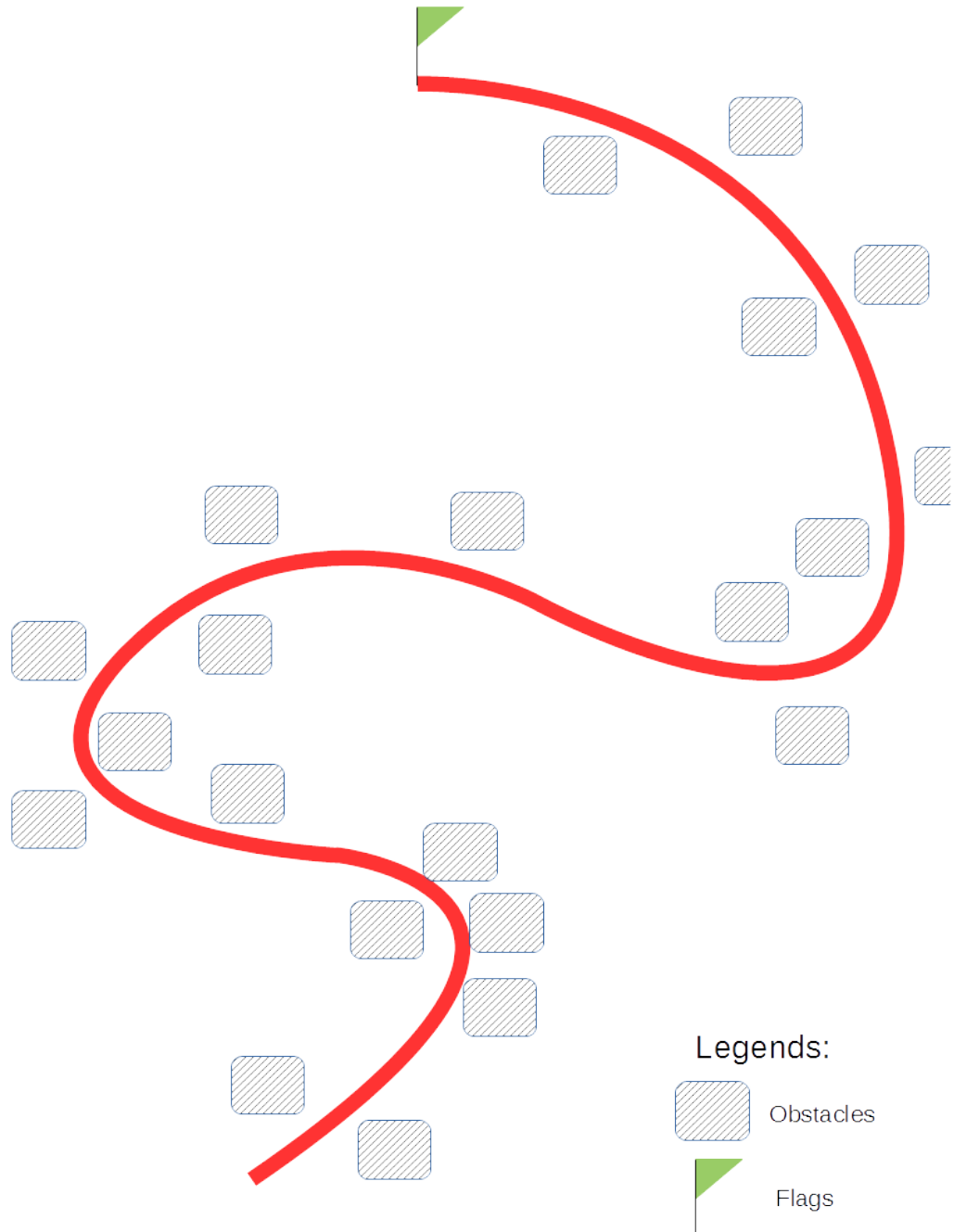
Figure 3. Course Path Design

Experiment Process

There were a total of 4 prototypes available to be tested by experimental subjects. Every subject needed to use all of the prototypes in order to produce an unbiased result. Because there were 24 possible sequence combinations, a minimum of 24 subject was required for the experiment. After learning the basic operation of different prototypes, subjects were randomly assigned to a unique sequence they needed to complete later. For example, if the sequence is 4213, the user started with Prototype 4, followed by Prototype 2, 1 and 3. Since the experiment course was the same for every prototype, implementing this approach enabled us to counterbalance the study. This reduced any sequence bias introduced via learning and memory (i.e. prototypes later in a sequence benefit from transfer learning obtained from prototypes earlier in a sequence).

The whole experiment was carried out in 5 working days where every participant was allocated a one hour slot to complete all 4 prototypes. 28 subjects signed up before the week, however only 16 attended the actual experimental session. For each session, there was a experimenter to handle the experiment. All of the subjects have normal vision. Before the start of the session, a suer was given a 4 number sequence. After that he/she was given a brief tutorial on how each prototype worked. The session started after the subject told the experimenter that he/she was ready. Following the sequence give, the subject completed the course journey using one prototype after another. At the end, subjects were given a feedback form. Besides the quantitative performance results such as time taken and number of obstacles knocked down, the feedback forms were used in order to capture insights pertaining to subjective aspects of prototype performance. Questions pertaining to the emotional response and opinion of users were asked and recorded.

RESULT AND DISCUSSION

Every human subject was allocated an hour for each session, and no subject exceeded this allocated time. However, there were 2 out of 16 users ended the session prematurely. Both of them showed signs of discomfort when using the devices. They were encouraged to go through part of the experiment but they took longer than usual to make decisions during navigation. Apparently, they were not able to distinguish sound patterns within the generated soundscapes. A large number of objects were knocked on their trials and they were not able to follow the yellow path correctly. If we can extrapolate, this suggests that there might exist a group of users, amongst the visually impaired, that is unable to use this kind of device on the first attempt. This group of users may be perceiving soundscapes in a way that is deeply different from the average user. This difference may contribute towards difficulties in recreating images and spatial relationships from soundscapes. Without specially tailored training material and tutorials targetted at this group of users, an excessively large amount of additional time to self-learn this kind of device might be necessary. When rolling out this sensory substitution device to the public, it is advisable to formulate training programs that can cater for both groups of users.

Navigation Result

Table 1 hows the average travel time in seconds separated by different prototypes labelled using P1 as Prototype 1, P2 as Prototype 2, P3 as Prototype 3 and P4 as Prototype 4. The navigation results are grouped into 3 different values where the first row shows the average total time traveled, the second row shows the average time taken from the origin to the target and finally the third row shows the average time taken for user to return from the target back to the point of origin.

	P1	P2	P3	P4
Total	293.0714	337.9286	360.9286	366
From Origin	183.9286	199.2143	217.7857	254.3571
Back to Origin	109.1429	138.7143	143.1429	111.6429

Table 1. Average Travel Time

There was no statistically significant difference between groups as determined by one-way ANOVA $(F(3,52) = 1.9497, p = 0.1376)$. However, the average travel time results show that Prototype 1 is better than the other 3 prototypes, in the sense that its total average task completion time was less than 300s,

while other prototypes required at least 10% more time. For the total time taken to travel, Prototype 3 and Prototype 4 obtained very close results with 360s for Prototype 3 and 360.9286 for Prototype 4.

The results are contrary to our original expectation, since from our previous tests with our own performance measurement tool (Tan et al., 2013), Prototype 3 fared the best in terms of information preservation and also interpretability. However, the results were consistent with the subjective opinions expressed by users, where most users felt that Prototype 1 was faster when generating soundscapes from frames, which in turn facilitated quick decision making during navigation. Prototype 1 takes 0.5s generating a soundscape from a frame, whereas Prototype 2 takes slightly less than 1s to perform the same task. As for the depth variant of Prototype 3, it takes at least 2s for soundscape generation. In a navigation scenario, normal human performance relies on quick judgment rather than a lengthy thought process before making the next decision (i.e. taking the next step) (Findlay and Gilchrist, 2003). Naturally human does gaze fixation to guide our movement to appropriate the landing target. On average, we fixate our gaze 2 steps ahead in a short period of time (about 800ms - 1000ms) before the limb is placed (Patla and Vickers, 2003). Similarly, during the experiment, subjects tend to move the camera left-right or up-down to emulate gaze fixation in real life which helps in increasing the accuracy of their next action (Mennie et al., 2007).

All prototypes demonstrated that return times were considerably reduced compared to the time taken for users to travel from the origin to the target. Prototype 4 showed a 56.1% decrease in time and is the prototype with the most time reduction out of the four prototypes. Apart from a significant reduction in return time, users of this prototype also tended to knock fewer obstacles on their way back to the origin. Prototype 1 came in at second with 40.67% time reduction. Prototype 2 and Prototype 3 yielded similar results, with 30.37% and 34.28% time reductions respectively.

This suggests that users create mental images/maps, on their first trip from origin to target, by listening to the soundscapes being generated while navigating. On their return trip, they presumably rely less on soundscapes because they can use the mental images/maps created on their previous trip as a guidance. It is not clear why Prototype 4 has such an improvement compared to its counterpart without depth information. However, the creation of mental images tends to be directly correlated to the information contained inside the soundscape. Prototype 1 converts relatively less visual information and therefore it helps in mental image recreation. From user feedback, mental images/maps are retained only for each individual prototype period. Subjects need to recreate a new mental image for the next test. This might be due to the fact that users are inexperienced in using this type of device. With proper training and exposure, we believe that users can be taught how to recreate rich mental images/maps from soundscapes with more information and how to prolong these mental images/maps.

Balloon Recognition Result

At the end of the path, users were asked to choose a specific balloon based on colour. Skills to locate the position of the balloon and also to distinguish balloons and other objects were required in order to perform this task. The table 2 shows the average time (in seconds) the users took to choose the correct balloon.

	P1	P2	P3	P3 with Depth
Average Time	78.92857	80.78571	72.71429	83.28571

Table 2. Average Balloon Recognition Time

The results from this aspect of the experiment do not exhibit an obvious pattern as with the navigational aspect. This could be due to the simplistic nature of the test, given that the main goal of the experiment focused on navigation rather than object recognition. In the future we could conduct an experiment to test all four prototypes focusing on object recognition as described in our "Swiping with Luminophonics" experiment (Tan et al., 2010). Comparing the four prototypes, users of Prototype 3 took 72s on average to choose the correct balloon while for the other prototypes they took 78.93s or more. Evolving from vOICe (Meijer, 1992), we did indeed expect Prototype 3 to be the best performer in this case, since it converts relatively more visual information than the other prototypes. In a task that requires more visual information like balloon recognition, Prototype 3 performs better compared to the results it obtains from the navigational task.

Depth Implementation

From these results, we can see that the prototype that incorporates depth information (i.e. P4) using a time-of-flight camera, does not result in a significant improvement over the other three prototypes. It performs worst both in terms of the navigation and recognition tasks. From our survey of the users who went through the experiments, most of them felt that our depth implementation did not provide a significant advantage over the other prototypes. In fact, some of them were even confused during the navigation course while using the prototype.

There are 2 possible reasons behind this poor result. The first one is due to our particular method for incorporating depth. Our method of letting users choose depth range during navigation not only caused confusion but also interrupted their flow of spatial reasoning. As mentioned earlier, during navigation users rely on snap judgments and quick reflexes, therefore prototypes which can provide faster refresh rates will perform better. This implementation slows down users' thought processes and creates additional confusion on top of the usual sound cacophony generated by VASS devices. The second reason, consists of our particular depth camera selection. Time-of-flight cameras do give us depth information but the QQVGA resolution (160x120) of the depth map is too limited for this type of application.

Even though our results are apparently discouraging towards depth information in VASS devices, we still believe that depth can be useful, and that it is a case of finding the right approach for incorporating this type of information. Additional effort needs to be put into researching new ways of converting and integrating different types of visual information, including depth. It is possible that new and effective ways for incorporating depth information will involve both advanced artificial intelligence and cognitive science, leading to methods that maximize information conversion, whilst reducing sound confusion and user interaction.

CONCLUSION

In summary, experimental designs based on human subjects are crucial for measuring the performance of sensory substitution devices. The experiment reported herein presents us with several valuable insights. Generally, it confirms that our understanding of this type of device is still basic. With this basic knowledge, the prototypes we build can typically only perform well in terms of one function. From our results, a faster prototype performs better in terms of navigation but often performs poorly at object recognition due to insufficient information. Trade-offs need to be made in order to create a good sensory substitution device unless we have a better understanding on how human brains work. In spite of our negative results pertaining to depth, we remain firm in the belief that future sensory substitution devices need to make use of this type of information. More work needs to be done to design better devices and algorithms that can adequately exploit the data from current depth cameras. Before visual to auditory sensory substitution devices can be widely adopted, there are a few major hurdles that researchers need to overcome, which include a better understanding on how the brain interprets visual information, more sophisticated hardware and also better algorithms that can convert visual information more effectively.

ACKNOWLEDGMENTS

REFERENCES

Bologna, G., Deville, B., and Pun, T. (2009). Blind navigation along a sinuous path by means of the see color interface. In Mira, J. M., Ferrández, J. M., Álvarez, J. R., de la Paz, F., and Toledo, F. J., editors, *IWINAC (2)*, volume 5602 of *Lecture Notes in Computer Science*, pages 235–243. Springer.

Bologna, G., Deville, B., Pun, T., and Vinckenbosch, M. (2007). Transforming 3d coloured pixels into musical instrument notes for vision substitution applications. *Journal on Image and Video Processing*, 2007(2):8–8.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Burgess, N., Maguire, E. A., and O'Keefe, J. (2002). The Human Hippocampus and Spatial and Episodic Memory. *Neuron*, 35(4):625 – 641.

Findlay, J. and Gilchrist, I. (2003). Active Vision: The Psychology of Looking and Seeing.

Lange, R. and Seitz, P. (2001). Solid-state time-of-flight range camera. *Quantum Electronics, IEEE Journal of*, 37(3):390–397.

Loomis, J., Wiener, W. R., Welsh, R. L., and Blasch, B. B. (2010). Sensory substitution for orientation and mobility: what progress are we making? sidebar 1.1 (pp. 7–10) to chapter 1, perceiving to move and moving to perceive: control of locomotion by students with vision loss by david a. guth, john j. rieser, and daniel h. ashmead (pp. 3–44). *Foundations of Orientation and Mobility (History and Theory)*, 1.

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

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

Mennie, N., Hayhoe, M., and Sullivan, B. (2007). Look-ahead fixations: anticipatory eye movements in natural tasks. *Experimental Brain Research*, 179:427–442.

Patla, A. E. and Vickers, J. N. (2003). How far ahead do we look when required to step on specific locations in the travel path during locomotion? *Experimental Brain Research*, 148:133–138.

Tan, S. S., Maul, T. H. B., Mennie, N., and Mitchell, P. (2010). Swiping with luminophonics. In *Cybernetics and Intelligent Systems (CIS), 2010 IEEE Conference on*, pages 52–57. IEEE.

Tan, S. S., Maul, T. H. B., and Mennie, N. R. (2013). Measuring the performance of visual to auditory information conversion. *PloS one*, 8(5):e63042.

World Health Organization (2014). Visual impairment and blindness.