

MOBILE, REAL-TIME SIMULATOR FOR A CORTICAL VISUAL PROSTHESIS

Horace Josh, Benedict Yong and Lindsay Kleeman

*Department of Electrical and Computer Systems Engineering, Monash University, Wellington Road, Clayton, Australia
Monash Vision Group, Monash University, Clayton, Australia*

Keywords: Visual prosthesis, Mobile simulator, Visual cortex, Visuotopic mapping, Phosphene, Bionic vision.

Abstract: This paper presents a mobile, real-time simulator system for a cortical visual prosthesis, making use of current neurophysiological models of visuotopy. This system overcomes fundamental limitations of current simulator systems which include simplified visuotopic mapping and the lack of mobility, limiting use in open and untethered environments. A visual prosthesis simulator provides a useful demonstration and research platform for a bionic vision system. It can be used to simulate the visual results of such an implant, as well as aid in the development of algorithms and techniques that would most suitably present information to a patient. Cortical visual prostheses work by electrically stimulating the visual cortex, the part of the brain primarily responsible for vision, and eliciting visual perceptions known as ‘phosphenes’. The simulator’s main function is to translate a scene provided by a camera sensor into a low resolution form that closely mimics the phosphene pattern produced by a cortical visual prosthesis. Preliminary psychophysics testing has suggested that in some situations it can be advantageous to have four different levels of intensity rather than two. It was also found that there is a learning effect associated with continued use of the system which would need further psychophysics study.

1 INTRODUCTION

A study conducted in 1968 showed that electrical stimulation of the visual cortex of a human brain resulted in the elicitation of bright spots of light, called ‘phosphenes’, in the visual field of the subject (Brindley and Lewin, 1968). Supporting results were also found in (Dobelle and Mladejovsky, 1974; Dobelle et al., 1976; Bak et al., 1990). Further studies (Humayun et al., 1996; Veraart et al., 1998) have shown that it is also possible to generate phosphenes via electrical stimulation of the retina and optic nerve. These early studies provided a basis for widespread research into the development of functional visual prostheses.

A visual prosthesis, also often referred to as a ‘bionic eye’, is an implantable biomedical device that aims to restore vision to the blind. The core component of these devices is an array of electrodes, driven by specialised electronics. The electrodes inject electrical current into a particular section of the patient’s visual pathway in order to generate an ‘image’ in the visual field.

The term visual pathway refers to the path that signals take from the retina in the eye where they are generated to the primary visual cortex at the back of the brain. Light that is incident on photoreceptors in the retina, a layer of cells at the back of the eye, results in the generation of signals. These signals are passed through the optic nerve and Lateral Geniculate Nucleus (LGN) before arriving at the primary visual cortex (V1), which is at the back of the brain. From V1, signals diverge to subsequent levels of visual cortex where higher level processing takes place. In a blind individual, parts of the visual pathway may not function. Therefore, visual signals do not reach the visual cortex. A successful prosthesis would bypass these inoperative sections in order to deliver signals to V1.

The Australian Research Council funded a new collaborative research initiative in 2009 to develop a functional visual prosthesis. One of the two proposals accepted for this initiative was by a Monash University led team of researchers, now known as the Monash Vision Group (MVG) (Monash Vision Group, 2010). Established in 2010, the MVG aims to develop a visual prosthesis

(Monash Bionic Eye) centred on a cortical implant, making use of approximately 600 electrodes.

As research grows in this new area of bionics, there is a great need for simulation or visualisation of the possible results of such an implant. Bionic eye simulators serve as good platforms for researchers to investigate the effectiveness of implemented algorithms, tune parameters, and realise the importance of certain parameters prior to actual clinical trials. The simulators would be used most in psychophysical trials – trials involving normally sighted individuals attempting to complete tasks with the limited vision provided by a simulator. However, the simulators would also be of use to the general public for educational purposes and to handle the expectations of families and friends of potential patients. Input to the system is in the form of an image or image stream. This image data goes through processing that transforms it into a representation that attempts to mimic the elicitation of phosphenes through electrode stimulation. The processed image data is then stored and/or displayed on a screen for viewing by the user.

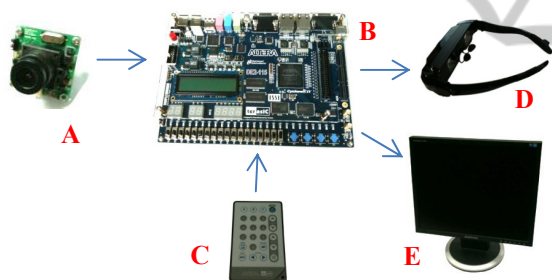


Figure 1: Main components of our simulator system: A) CMOS Camera B) FPGA Development Board C) IR Remote D) Head-Mounted Display E) External Monitor.

Many visual prosthesis simulators have already been developed and some of the more recent work is found in (Van Rheede et al., 2010; Zhao et al., 2010; Fehervari et al., 2010; Srivastava et al., 2009; Chen et al., 2005). Nevertheless, there are some significant limitations that arise in their implementations. The majority of these simulators perform their image processing on a computer using image processing libraries and so are often limited to use within an area close to a stationary computer. Depending on the complexity of processing and the available processing power of the equipment in use, these systems may sometimes suffer from latency and frame rate issues. In the case of simulators for cortical visual prostheses, visuotopic mapping – the mapping of electrode placement on the visual cortex to elicitation of phosphenes in the visual field, has

often been overlooked or used simplified models.

Our system aims to address the shortcomings of currently implemented systems. In comparison to other cortical FPGA based systems (Fehervari et al., 2010; Srivastava et al., 2009), our system is very mobile and has been used to do untethered preliminary psychophysics testing. Our simulator is based on a Field Programmable Gate Array (FPGA) system implementation. FPGAs are microchips that offer extremely dense amounts of electronically reconfigurable logic gates. FPGA systems offer the advantages of low latency, highly parallel implementation and the ability to integrate with large numbers of external devices through the high availability of peripheral interface pins. Figure 1 shows the main components of our simulator system. A CMOS camera captures a stream of image data, which is then processed on an FPGA development board and finally displayed on a head-mounted display and optionally on an external monitor as well. An infra-red remote control interface is used to enable/disable the various functions. A more detailed description of the system components is provided in Section 2.

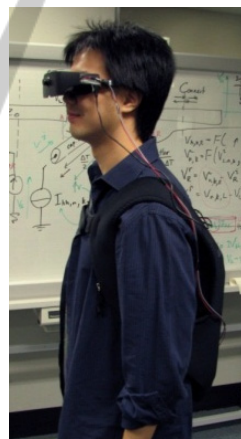


Figure 2: Integrated system.

2 SYSTEM SETUP

As shown in Figure 1, our system is comprised of the following main components: a camera for acquiring images, an FPGA development board for performing all image processing functions and visuotopic mapping, a head-mounted display as well as optional external monitor for display of the resulting image stream, and finally an infra-red remote control for toggling of functions.

The camera that we have chosen to use is a low cost CMOS camera (Sparkfun Electronics CM-26N/P)

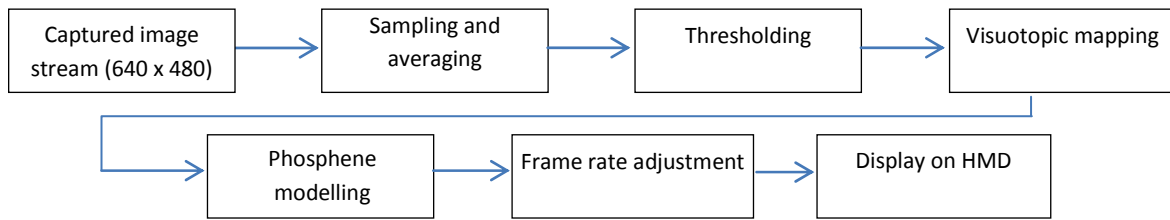


Figure 3: Flowchart of main functions of the system.

which has an analogue signal output. It captures images at a resolution of 640 x 480 pixels, at a frame rate of 59.94Hz and has a viewing angle of 70°. Reasons for choosing this particular camera include low cost, small physical size, switchable PAL/NTSC output, and the simplicity of a three wire power/signal connection which also allows for longer cable lengths.

At the centre of our system, we have a Terasic DE2-115 FPGA development board, which is based on an Altera Cyclone IV EP4CE115F29C7 FPGA chip. We chose this development board for its low cost, lower power consumption, high logic element and on-chip memory count, wide range of available peripheral devices and I/O pins, and our familiarity with its design and operation.

An infra-red remote, that comes standard with the DE2-115, was utilised for capturing user input. It provides a simple and easy way of toggling and controlling all implemented functions.

For display of the final output, we have chosen a head-mounted display (HMD) unit (Vuzix iWear VR920), sometimes referred to as virtual reality goggles. This HMD offers a 640x480 pixel display resolution with a viewing angle of 32°. The VR920 was chosen for its low cost, compatible resolution, lightweight design, and its ability to take an analogue VGA signal as its input. Since our system outputs video via a VGA port, we were able to use a simple passive splitter cable to provide dual output (HMD as well as an external monitor).

For our system to be mobile, all hardware needed to be integrated into a neat, wearable package. We achieved the result shown in Figure 2. The majority of components are fastened inside a hard plastic laptop casing, which is then placed in a neoprene laptop bag with cables running to the camera and HMD that the user is wearing. A 12V rechargeable lithium-ion battery pack is used to power the system.

3 SYSTEM IMPLEMENTATION

The flowchart shown in Figure 4 outlines the

implementation of the main functions of our system. A high resolution image stream (640x480 pixels) is captured by the CMOS camera, which is delivered to the DE2-115 development board via a standard NTSC analogue connection. After decoding of the NTSC signal is complete, the pixels are sampled and averaged. The sampled data is thresholded in order to simulate possible limitations of electrode stimulation. A pre-generated visuotopic mapping lookup table is then used to determine the placement of the phosphenes on the output display. A discrete Gaussian falloff profile is used to simulate the physiological phenomena of a phosphene dot in the visual field. Before output on the screen, the frame rate of the system can be set in real-time in order to simulate varying stimulation frequencies of electrodes. A more detailed explanation of these main system features is given in Subsections 3.1, 3.2, 3.3, 3.4, and 3.5.

Furthermore, features such as edge detection, histogram assisted threshold selection, and dead electrode assisted simulation, have been implemented in order to allow for evaluation of the effects of such image processing techniques on the perception of the provided low resolution data (Subsection 3.6).

All processing performed on the image stream from the camera is implemented using Verilog hardware description language. Unlike conventional code that is written for execution on a processor that runs at a specific clock speed, Verilog describes the way logic gates are to be arranged and connected and so is compiled into a synthesisable logic solution that can be either synchronous (operate with reference to a clock), asynchronous (without reference to a clock) or a mixture of the two. A Verilog solution was chosen due to the ability to create functions that can run in parallel, resulting in a low latency real-time system.

3.1 Visuotopic Mapping

Early physiological research (Schwartz, 1977; Wandell et al., 2007) proved that ‘points’ in the visual field correspond to specific locations on the

visual cortex, inferring a ‘map’ or transfer function between visual field points and the visual cortex. Furthermore, that map is mostly continuous in that neighbouring points in the visual field correspond with neighbouring points on the visual cortex. The map or transfer function which describes the translation of points between the visual cortex to its corresponding points on the visual field is known as the visuotopic map.

Due to the physiological non-linear properties of the visual cortex, the visuotopic map is also non-linear and ‘distorted’. In humans, the phenomenon known as cortical magnification describes how a small region at the centre of the visual field, known as the fovea, corresponds with a much larger area of the visual cortex (Horton and Hoyt, 1991; Duncan and Boynton, 2003). Early work by Schwartz (1977) indicated an approximation to the mapping by a ‘log-polar’ representation, where linear points on the visual cortex correspond to eccentrically logarithmic and angularly linear points in the visual field. The foveal region is represented this way as a dense packing of points in the centre of the visual field which corresponds to a disproportionately larger region on the visual cortex. Also important to note is that the visual cortex is spread over both halves of the brain with the left visual cortex corresponding with the right visual hemifield and vice versa, due to cross-over of the optic nerves. (Bear et al., 2007).

Mathematical models that came from this include the Monopole model (defined from the ‘log-polar’ observations) (Schwartz, 1977; Polimeni et al., 2006; Schira et al., 2010), the Wedge-Dipole model (adds a second parameter to Monopole model to account for curvature in the periphery region of the visual cortex) (Balasubramanian et al., 2002; Polimeni et al., 2006) and more recently the Double-Sech model (adds a shear function to the Wedge-Dipole model to account for changing local isotropy as well as increasing accuracy of mapping at higher levels of visual cortex V2, V3) (Schira et al., 2007; Schira et al., 2010).

As the implant is anticipated to consist of a linear array of electrodes, the resulting phosphene pattern would not be linear but rather follow this log-polar mapping. It would be useful and more accurate to model the output visualisation based off a mathematical model of the visuotopic mapping. Since the implant is expected to be placed in the primary visual cortex V1 and closer to the foveal side of the visual cortex, the Monopole model was chosen to model the output visualisation as it was mathematically simpler and still provides reasonable accuracy.

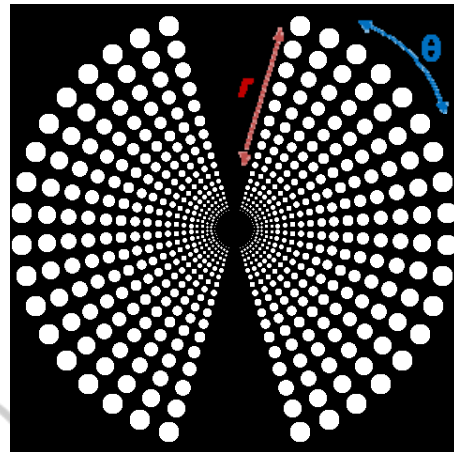


Figure 4: Resultant visual field of implemented visuotopic map.

The Monopole equation (1) describes the left visual cortex ‘ w ’ as a complex function of the right visual hemifield ‘ z_w ’. ‘ \mathbb{C} ’ is the set of complex numbers, and ‘ k ’ is a dilation factor constant.

$$w = k \log(z_w + a) \in \mathbb{C} \quad (1)$$

Visual field z_w can be represented as a complex exponential where r represents the eccentricity and θ represents polar angle.

$$z_w = r e^{i\theta} \in \mathbb{C} \quad (2)$$

Rearranging the Monopole equation describes visual field z_w as a function of visual cortex w .

$$z_w = e^{\frac{w}{k}} - a \in \mathbb{C} \quad (3)$$

The electrode array of the implant was assumed to be a linear array placed on the visual cortex closer to the foveal region. The visuotopic map was created using MATLAB and ported over to the FPGA for use as a large lookup table. Approximate values were used for the Monopole equation parameters, which are reasonably consistent with the various values used in the literature: $k=15$, $a=0.7$. (Polimeni et al., 2006; Schira et al., 2007; Fehervari et al., 2010). The exact dimensions and intended locations of the implant are still not known, the eccentricity and polar angle were limited to an 18×18 linear array on the visual cortex that cover the following values on: $r=[10,40]$, $\theta=[-0.8(\frac{\pi}{2}), 0.8(\frac{\pi}{2})]$. This only represents the left visual cortex, corresponding with the right visual hemifield. The 18×18 array was duplicated for the right visual cortex, creating another array on the left visual hemifield. This

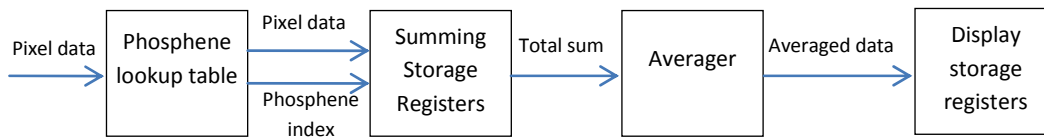


Figure 5: Averaging sampler implementation.

produces a total electrode count of 648. These assumptions were taken to make better use of the limited screen resolution of the head-mounted display while remaining realistic to the ‘log-polar’ mapping of the visual cortex. However, new maps can be simply regenerated on MATLAB to accommodate any changes to this and implemented into our system. The resultant visual field of our implemented map is shown in Figure 4.

3.2 Averaging Sampler

Figure 5 outlines our averaging sampler implementation. After NTSC decoding, the image stream from the camera is made available one pixel at a time in a sequential fashion. As each pixel arrives at the sampling section of the system, its X & Y pixel count values are compared against the mapping lookup table. This lookup table stores the corresponding phosphene index number for each pixel within the central 480 x 480 window of the full camera view. Pixels not belonging to a phosphene are assigned number zero. Once the phosphene index number is determined, the pixel is sampled by adding to a storage register that corresponds to that particular phosphene index number. This process repeats until all pixels have been sampled. Finally, an average is performed on all of the storage registers according to the number of pixels that are within each phosphene, and the results are stored in a separate set of storage registers.

3.3 Thresholding

Various studies (Brindley and Lewin, 1968; Dobbelle and Mladejovsky, 1974; Schmidt et al., 1996) have shown that the modulation of phosphene brightness is possible using a number of different techniques. However, there is some ambiguity in the possible number distinguishable brightness levels.

Our system takes an optimistic approach at simulation of this property, having the option to display at 2, 4 or 8 levels of intensity or greyscale. Since our system uses 10-bit storage registers for pixels, the full greyscale intensity range is 0 to 1023. This range is divided evenly in order to create bands of intensity for 2, 4 and 8 level modes. Results of 2

and 4-level thresholding are shown in Figure 6. It is often difficult to perceive the results of the system in a static image form, therefore we encourage you to view the videos we have listed in the appendix.



Figure 6: Thresholding: full resolution image (top), 4-level image (bottom left), binary image (bottom right).

To avoid high frequency oscillation between intensity bands, a hysteresis feature was included. Two threshold values are used to define changes between intensity bands, instead of one value. When a phosphene’s intensity is between the two thresholds, no change occurs. Figure 7 shows how hysteresis reduces the oscillation problem.

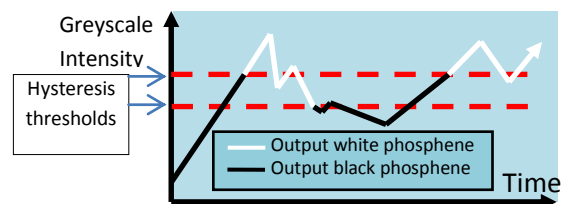


Figure 7: Binary thresholding with hysteresis.

3.4 Phosphene Modelling

Stimulation of each electrode on the implant will produce a phenomenon in the patient’s visual field known as a phosphene, whose appearance is somewhat similar to a bright spot of light (Brindley and Lewin, 1968). Rather than simply using square pixels that perfectly line up with each other, we attempted to model the output visualisation based on

what phosphenes would approximately look like.

In the literature, one common approach is to model the phosphene using a 2D Gaussian mask. (Chen et al., 2009). The 2D Gaussian function is based on the standard distribution curve, except in two dimensions instead of one. This creates the appearance of a round ‘spot’ where the centre of the spot has the highest intensity value with the intensity values decreasing radially towards the outside edge of the spot, following the standard distribution curve. A comparison between a phosphene with and without the Gaussian function applied is shown in Figure 8.

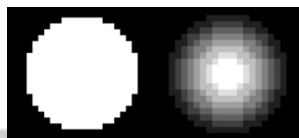


Figure 8: Phosphene modelling: without Gaussian function (left), with Gaussian function (right).

3.5 Frame Rate Reduction

The ability of a person to detect motion is very important when it comes to mobility exercises in low resolution vision. A key factor that would limit one’s ability to detect motion in the immediate environment is the lack of temporal resolution. It is expected that the temporal resolution of electrode stimulation achievable by the Monash Bionic Eye may be in the range of 5-15 frames per second. In order to simulate this temporal resolution and investigate the possible implications it may have on a patient’s ability to move around, we have implemented a frame rate reduction function. The output frame rate of our system can be changed in real-time. Our system has 8 different discrete frame rates available for selection (1, 2, 4, 8, 10, 15, 30 and 60 frames per second). Variable frame rate is achieved by holding the stored frame output data for the specific period of the chosen frame rate.

3.6 Extra Functions

Additional functions have been implemented in our system, such as edge detection histogram assisted threshold selection and dead electrode simulation. These features have not been evaluated in the preliminary testing we present in this paper; however, would be of importance for future psychophysical research we intend to carry out. Figure 9 demonstrates edge detection and dead electrode simulation.

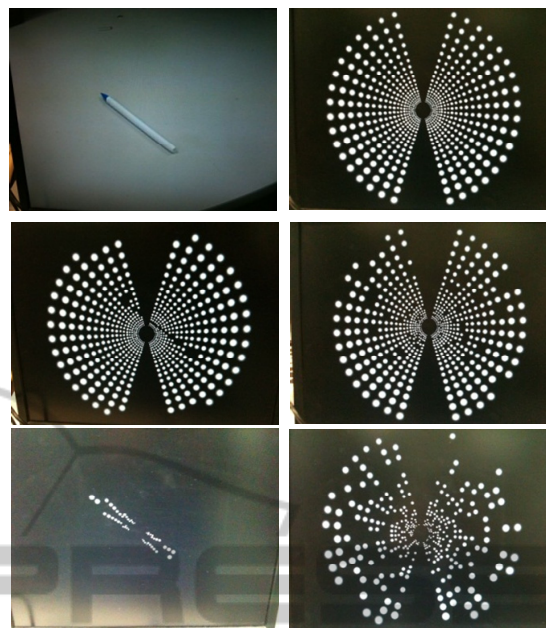


Figure 9: Edge detection: full resolution (top left), binary thresholding (middle left), edge detection (bottom left). Dead electrode simulation: 0% (top right), 10% (middle right), 50% (bottom right).

4 EXPERIMENTAL SETUP

After the hardware was built, two different psychophysical preliminary testing experiments were devised by the authors and conducted by a number of Monash Vision Group staff and post-graduate students as volunteers. These experiments were not formalised clinical trials, but rather preliminary trials to test the effectiveness of the system and to examine the effective difference between the modes and parameters set on the system on the end user. The two experiments were a mobility based obstacle avoidance walking maze test, as well as a sit-down contrast discrimination hand-eye co-ordination chessboard placement test.

4.1 Maze Test

In this test, there were 7 test subjects (6 male, 1 female). The maze test involved subjects walking through a course while avoiding obstacles. The obstacles were large cardboard boxes and office chairs with wheels. The placement of the obstacles was randomised within the maze area and 5 different configurations of obstacle layout were developed, one for each mode tested and kept consistent between subjects. Subjects were not allowed to see

the obstacle layout before each test. The starting point was around the corner from the main rectangular maze area, and the end point was at a table at the far end wall of the maze. There is a small black box on the table and the test ends when the subject finds and picks up the box.



Figure 10: Maze Test obstacle layout.

For the test, both time to completion and number of collisions were recorded for all subjects. Subjects were allowed to touch the obstacles in the maze so only unintentional collisions were counted. The 5 modes tested were a control (full resolution, full colour), 4-level thresholding (full frame rate), binary thresholding (full frame rate) and reduced frame rate at 15 fps and 4 fps (both with 4-level thresholding). Subjects were given 2 minutes accommodation time just before the test for each mode where they could adjust to using the system around a cardboard box and two chairs placed away from the actual maze area. Subjects were also given a minimum of 5 minutes break in between each test.

4.2 Chessboard Test

In this test, there were 7 test subjects (6 male, 1 female) and were the same subjects as used in the previous Maze Test. The task required subjects to sit down at a table with a chessboard in front of them and 16 chess pieces (8 black, 8 white) placed in a random pile to the left of the chessboard. The objective was for the subjects to sort and place any black coloured pieces on any white square in the bottom half of the chessboard, and the white pieces on black squares in the top half of the chessboard.

For the test, both time to completion and number of mistakes were recorded for all subjects. For a piece to be considered as correctly placed, at least half of it had to be over the right square. Another aspect to this experiment was to test for learning effects that come from repeated usage of the system. As such, the non-control modes tested were repeated 3 times in this order (all at full frame rate): control (full resolution, full colour), binary thresholding, 4-level thresholding, binary, 4-level, binary, 4-level. Before the testing, subjects were asked to attempt the task without wearing the system in order to familiarise themselves with the task itself. The testing was conducted in a single session, with a minimum 1 minute break in between each test.



Figure 11: Chessboard Test finished example.

5 RESULTS AND DISCUSSION

5.1 Maze Test

Figure 12 is a graph that details the time to completion (in seconds) for each mode, averaged over the 7 subjects. The order of the modes reflects the order that the subjects were tested in. The error bars show the standard error. 2-way, paired T-Tests were conducted between the control time and each of the non-control modes, as well as between the 4-level thresholding full frame rate mode and the other 3 modes (binary and both reduced frame rates).

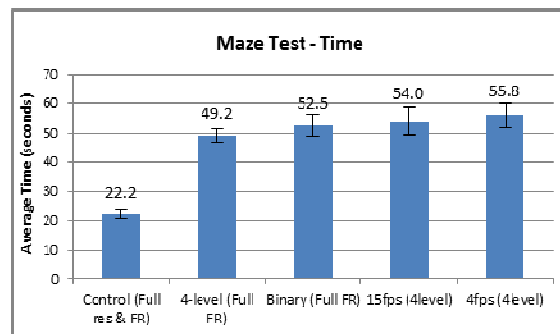


Figure 12: Maze Test - mode vs. average time (seconds).

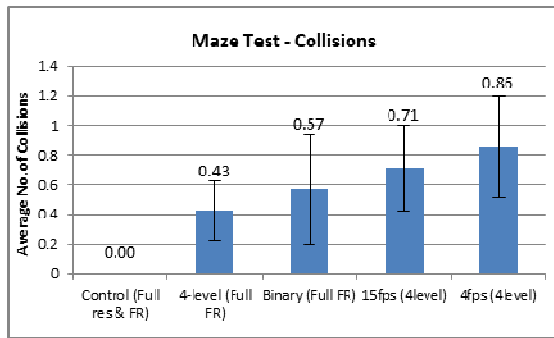


Figure 13: Maze Test - mode vs. average no. of collisions.

The times taken for all the non-control modes were significantly higher ($p < 0.05$ for all) than the time for the control. The binary and reduced frame rate modes were slightly longer than the 4-level thresholding full frame rate mode, but all the non-control modes were within the statistical margin of error ($p > 0.05$ for all).

Figure 13 details the number of collisions for each mode, averaged over the 7 subjects. The error bars show the standard error. The average number of collisions was very low, due to a few of the subjects not colliding with anything in any of the modes, but the binary thresholding and reduced frame rate modes had more collisions on average than the 4-level thresholding full frame rate.

5.2 Chessboard Test

Figure 14 details time to completion (in seconds) for each mode, averaged over the 7 subjects. The order of the modes reflects the order the subjects were tested in and shows how the same modes were tested repeatedly 3 times to examine learning effects. The error bars show the standard error. 2-way, paired T-Tests were conducted between the control time and each of the non-control modes, as well as between the binary and 4-level thresholding for each pair of repeated tests (eg. 1st binary with 1st 4-level).

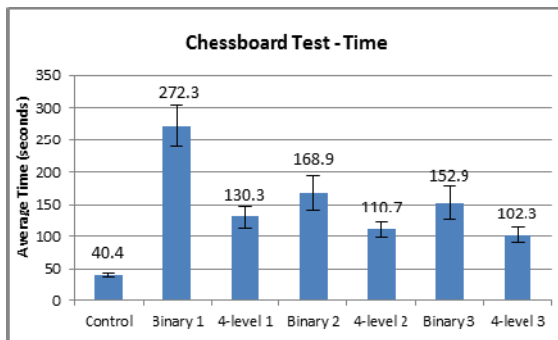


Figure 14: Chessboard Test - mode vs. average time (sec).

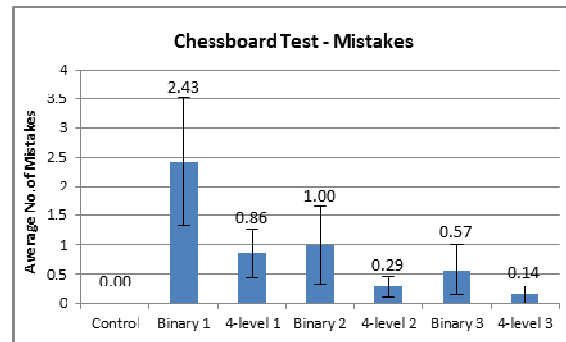


Figure 15: Chessboard Test - mode vs. average mistakes.

The times taken for all non-control modes were significantly longer than control mode ($p < 0.05$ for all). The times taken for the binary modes were significantly longer than the 4-level thresholding for the same repeated number of trial ($p < 0.05$ for the 1st and 2nd pairs of tests, $p = 0.063$ for the 3rd pair). The times for all modes decreases with increasing number of repeated tests.

Figure 15 details the number of mistakes for each mode, averaged over the 7 subjects. The error bars show the standard error. The average number of mistakes was quite low due to some subjects not making any mistakes. The trend however clearly looks similar to the Chessboard Time graph with decreasing number of mistakes with repeated trials.

5.3 Discussion

The Maze Test results show that subjects take much longer to finish the test in any of the non-control modes compared to the control, and that although the binary and reduced frame rate modes took slightly longer to complete than the 4-level full frame rate mode, the difference was not significant. This trend is also shown in the average number of collisions, but the standard error is very large.

From observations made while building and testing the system, reduction in colour depth and frame rate does increase the difficulty of most general tasks including navigational and obstacle avoidance tasks. Possible reasons for this not being made clear in the results are that the maze area was fairly small and straightforward so the task could be completed in a relatively short amount of time, and the number of test subjects was low, presenting a relatively large error. Also, the obstacles used in this test were large and obvious and so subjects may not have benefitted a lot from an increased colour depth and frame rate. Another problem could be the order of the modes in which the subjects were tested was made consistent and that the 'harder' modes were

tested later. A learning effect just from repeated testing, even with the changing obstacle placement and accommodation time between tests, could cause a decrease in times for the later tested 'harder' modes and hence reduce differences between them and the earlier test 'easier' modes.

For the Chessboard Test, the results demonstrate that the binary modes were significantly longer than the 4-level thresholding modes for each repeated test. The results also show that there is a clear downwards trend with increasing number of tests for both modes. The average number of mistakes also shows these trends, that the binary has more mistakes than the 4-level and that both modes decrease over repeated testing, however the standard error is very large. The reason the tests were completed much faster on 4-level compared to binary is likely because this test is based primarily on contrast discrimination and the extra levels of grey available on the 4-level allow the subjects to be able to tell the difference between the dark and light chess pieces and chessboard as well as the grey table more rapidly. This shows that different tasks may benefit differently from various modes. A significant learning effect was evident as times and mistakes would decrease with repeated testing, probably leading to an eventual plateau point where times do not get much faster. It is apparent that as people keep repeating a task they are unfamiliar with, they will improve at it. There should be no reason why it is not the same when using a visual prosthesis simulator, or even a patient with a visual prosthesis implant itself.

5.4 Limitations of the System

Although, our system uses a physiologically based model for mapping of phosphenes, it does not represent the gaze-locked nature of a cortical implant. In the case of a real cortical visual prosthesis, the patient will not be able to focus on different points of the visual field with eye movements. In our system however, the user is able to scan the presented pattern voluntarily. To overcome this limitation, an eye-tracker would be required to allow the system to move the pattern along with the movement of the user's eyes, therefore 'locking' the gaze at a specific point (usually at the center) in the presented pattern.

6 CONCLUSIONS AND FUTURE WORK

This paper has presented a simulator for a cortical visual prosthesis. By addressing fundamental limitations in current simulator systems through its portability, and physiologically based phosphene mapping, the system has met expectations and makes a good platform for investigation, improvement and tuning of algorithms for use with a visual prosthesis. The completion of preliminary psychophysical testing has shown that the number of greyscale intensities has a significant effect on results for certain tasks. It was also found that a learning effect is present with repeated trials which will need to be addressed in future work with broader and more rigorous sets of psychophysical testing. It is hoped that valuable insight can be gained and used to improve the implementation of future visual prosthesis devices.

ACKNOWLEDGEMENTS

Monash Vision Group is funded through the Australian Research Council Research in Bionic Vision Science and Technology Initiative (SR1000006). The authors would like to thank the members of Monash Vision Group that participated in the trials and all those that shared their valuable opinions and advice. The authors would also like to thank Grey Innovation for help with the physical layout of the integrated simulator system.

REFERENCES

- Bak, M., Girvin, J. P., Hambrecht, F. T., Kufts, C. V., Loeb, G. E., Schmidt, E. M., 1990. Visual sensations produced by intracortical microstimulation of the human occipital cortex. *Medical & Biological Engineering & Computing*, vol. 28, pp. 257-259.
- Balasubramanian, M., Polimeni, J. R., Schwartz, E. L., 2002. The v1-v2-v3 complex: quasiconformal dipole maps in primate striate and extra-striate cortex. *Neural Networks*, vol. 15, iss.10, pp1157-1163.
- Bear, M. F., Connors, B. W., Paradiso, M. A. 2007. *Neuroscience: Exploring the Brain*. Lippincott Williams & Wilkins. Baltimore, 3rd edition.
- Brindley, G. S., Lewin, W. S., 1968. The sensations produced by electrical stimulation of the visual cortex. *Journal of Physiology*, vol. 196, pp. 479-493.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679-698.

- Chen, S. C., Hallum, L. E., Lovell, N. H., Suaning, G. J., 2005. Visual acuity measurement of prosthetic vision: a virtual-reality simulation study. *Journal of Neural Engineering*, vol. 2, pp. S135-S145.
- Chen, S. C., Suaning, G. J., Morley, J. W., Lovell, N. H., 2009. Simulating prosthetic vision: i. visual models of phosphenes. *Vision Research*, vol. 49, pp. 1493-1506.
- Dobelle, W. H., Mladejovsky, M. G., 1974. Phosphenes produced by electrical stimulation of human occipital cortex, and their application to the development of a prosthesis for the blind. *Journal of Physiology*, vol. 243, pp. 553-576.
- Dobelle, W. H., Mladejovsky, M. G., Evans, J. R., Roberts, T. S., Girvin, J. P., 1976. 'Braille' reading by a blind volunteer by visual cortex stimulation. *Nature*, vol. 259, pp. 111-112.
- Dowling, J. A., Maeder, A. J., Boles, W., 2004. Mobility enhancement and assessment for a visual prosthesis. *Proceedings of SPIE Medical Imaging 2004: Physiology, Function, and Structure from Medical Images*, vol. 5369, pp. 780-791.
- Duncan, R. O., Boynton, G. M., 2003. Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, vol. 38, pp. 659-671.
- Fehervari, T., Matsuoka, M., Okuno, H., Yagi, T., 2010. Real-time simulation of phosphene images evoked by electrical stimulation of the visual cortex. *Neural Information Processing*, vol. 6443, pp. 171-178.
- Horton, J. C., Hoyt, W. F., 1991. The representation of the visual field in human striate cortex: a revision of the classic holmes map. *Archives of Ophthalmology*, vol. 109, pp. 816-824.
- Humayun, M. S., de Juan, E., Dagnelie, G., Greenberg, R. J., Propst, R. H., Phillips, D. H., 1996. Visual perception elicited by electrical stimulation of retina in blind humans. *Archives of Ophthalmology*, vol. 114, pp. 40-46.
- Lee, J. S. J., Haralick, R. M., Shapiro, L. G., 1987. Morphologic Edge Detection. *IEEE Journal of Robotics and Automation*, vol. 3, pp. 142-156.
- Monash Vision Group, 2010. Monash vision direct to brain bionic eye. Viewed 11th July, 2011, <<http://monash.edu.au/bioniceye>>.
- Polimeni, J. R., Balasubramanian, M., Schwartz, E. L., 2006. Multi-area visuotopic map complexes in macaque striate and extra-striate cortex. *Vision Research*, vol. 46, pp. 3336-3359.
- Schira, M. M., Wade, A. R., Tyler, C. W., 2007. Two-dimensional mapping of the central and parafoveal visual field to human visual cortex. *Journal of Neurophysiology*, vol. 97, pp. 4284-4295.
- Schira, M. M., Tyler, C. W., Spehar, B., Breakspear, M., 2010. Modeling magnification and anisotropy in the primate foveal confluence. *PLoS Computational Biology*, vol. 6, iss.1, pp. 1-10.
- Schmidt, E. M., Bak, M. J., Hambrecht, F. T., Kuffa, C. v., O'Rourke, D. K., Vallabhanath, P., 1996. Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. *Brain*, vol. 119, pp. 507-522.
- Schwartz, 1977. Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetics*, vol. 25, pp. 181-194.
- Srivastava, N. R., Troyk, P. R., Dagnelie, G., 2009. Detection, eye-hand coordination and virtual mobility performance in simulated vision for a cortical visual prosthesis device. *Journal of Neural Engineering*, vol. 6, pp 1-14.
- Van Rheede, J. J., Kennard, C., Hicks, S. L., 2010. Simulating prosthetic vision: optimizing the information content of a limited visual display. *Journal of Vision*, 10(14):32, pp. 1-15.
- Veraart, C., Raftopoulos, C., Mortimer, J. T., Delbeke, J., Pins, D., Michaux, G., Vanlierde, A., Parrini, S., Wanet-Defalque, M., 1998. Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Research*, vol. 813, pp. 181-186.
- Wandell, B. A., Dumoulin, S. O., Brewer, A. A., 2007. Visual field maps in human cortex: review. *Neuron*, vol. 56, pp. 366-383.
- Zhao, Y., Lu, Y., Tian, Y., Li, L., Ren, Q., Chai, X., 2010. Image processing based recognition of images with a limited number of pixels using simulated prosthetic vision. *Information Sciences*, vol. 180, pp. 2915-2924.

APPENDIX

- Vid.1) www.youtube.com/watch?v=oAxaNloHVVHg
 Vid.2) www.youtube.com/watch?v=2byh1qQfWGQ
 Vid.3) www.youtube.com/watch?v=gIVrnsk04LA