RADBOUD UNIVERSITY NIJMEGEN

FACULTY OF SCIENCE

Biological motion for visual cortex induced phosphenes

Master Thesis 1 Medical Biology

Author:

Lynn Le

Daily supervisor: Dr. Yağmur Güçlütürk

Supervisor: prof. dr. Richard van Wezel

August 2018

Contents

List of Figures

Abstract

The active pursuit to develop cortical visual prostheses offers a promising future for blind people. Cortical implants with penetrating microelectrode arrays can generate visual percepts by electrically stimulating the visual cortex. The resulting percept of bright dots due to the stimulation of the visual pathway is called phosphenes. Biological motion perception occurs even with sparsely placed dots representing body parts, and it has been shown to not require local motion, making it a favorable notion for phosphene induction. Here, through psychophysical experiments of simulated phosphene vision in sighted participants, we investigated the performance in biological motion perception. We tested the performance of participants on a forward/backward task with different frame durations using a limited lifetime of the moving dot stimulations to eliminate local motion. Preliminary experiments show promising results supporting the development of a cortical visual prosthesis with functionalities that exploit the visual system's ability to perceive biological motion from sparse stimuli. Future work will involve development of sophisticated deep learning methods to automatically perform pose detection for biological motion representations, and testing them in augmented and virtual reality phosphene vision simulations, with the ultimate goal of integrating them in a cortical visual prosthesis under development.

1 Introduction

1.1 Visual prosthesis

Developing visual prosthesis is an approach to partially restore lost vision to blind individuals based on electrically stimulating intact neurons in the visual pathway. Implants can target various regions in the visual pathway including the retina (Weiland & Humayun, 2014), the optic nerve (Delbeke, Oozeer, & Veraart, 2003), the lateral geniculate nucleus (LGN) (Cudeiro & Sillito, 2006; Wiesel & Hubel, 1966), and the visual cortex (Brindley & Lewin, 1968). Each of these target regions described in figure 1 has their own advantages and disadvantages for visual prosthesis development. Although retinal, LGN and optic nerve implants allow the use of its compact dimensions, retinotopic organization, and physical separation of pathways specific to color and motion, there needs to be surviving populations of the retinal ganglion cells in the eyes and/or earlier parts of the visual pathway. Cortex-based artificial vision systems are implemented in the primary visual cortex (V1). Stimulation of the visual cortex is an attractive notion based on multiple factors. First, it is possible for higher number of electrodes to be inserted, potentially resulting in higher resolution. Secondly, the positioning of electrodes is more straightforward compared to the LGN or retina implants. Finally, these implants are suitable for all kinds of visual impairments, including those with late blindness due to retinal optic nerve diseases or injury.

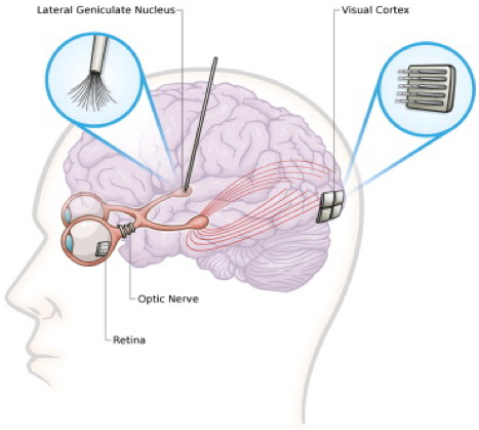


Figure 1: Electrical stimulation targets for visual prosthesis.

1.2 Background and related work

Neuroprosthetic devices are commonly designed to process information from the environment into the nervous system. An example of a device communicating with a specific group of neurons to successfully restore its functionalities, are cochear implants (Zeng, Rebscher, Harrison, Sun, & Feng, 2008). This demonstrates that scientists have a well understanding of the most fundamental requirements of neural prosthesis.

The concept of cortex-based artificial vision initiated with research on the cerebral cortex functional architecture. While studying epilepsy, Wilder Penfield and co-workers observed that electrical stimulation of the visual cortex surface, induced phosphenes (Mazzola, Isnard, Peyron, & Mauguière, 2011). Later on, researchers such as the group of Dobelle at the University of Utah and some other groups showed that electrode stimulation allowed blind subjects to perceive patterns (Dobelle, 2000; Bak et al., 1990; Pollen, 1975). Ever since these findings, efforts have been made to develop electrodes suitable for implants allowing the excitement of neurons selectively while eliciting a low amount of electrical currents.

The newer microelectrodes are spatially dense, which has the potential to permit higher resolution. There are other important attributes that must be taken into account while developing the optimal interface with neurons such as reaching the neurons of interest while minimizing tissue damage, minimizing the disruption of the subject's physical activity, avoiding unimportant signals that would induce unnecessarily high amount of current.

1.3 Current work

Neurons in the visual cortex are arranged into retinotopic maps of visual space, such that stimulation of neurons in a particular part of the brain will lead to a phosphene percept at a location in the visual field. The NESTOR project focuses on generating a direct-to-brain bionic vision system, which is based on multiple arrays leading to 1000 electrodes implanted into the visual cortex. The implementation of large number of electrodes theoretically gives a higher number and density of phosphenes, yielding better 'pixel resolution' in evoked artificial images. Being involved in prosthetic vision research, we seek to provide visual perception for blind people to help enable or enhance the performance of important and meaningful daily tasks. We want to do this by creating clear visual representations and taking away unnecessary signals form visual input as much as possible.

1.4 Biological motion and limited lifetime

A well-known concept is biological motion perception, which refers to the visual system's ability to recover information about a biological entity's motion from sparse input. Biological motion perception is therefore

a fascinating instance of the form-from-motion-effect (FFM), as well as part of social perception. This processing ability is part of the top-down processing in the visual system and is a very ecologically valid stimuli. Biological motion stimuli is physically complex, yet psychologically very important (de Xivry, Coppe, Lefèvre, & Missal, 2010).

It's been shown that for one to perceive biological motion, local motion is not required (Beintema & Lappe, 2002; Beintema, Georg, & Lappe, 2006). Beintema's research in 2006 showed that taking away the local motion by introducing limited lifetime of dots in a stimulus, subjects are still able to perceive whether the walker is walking forward or backward. The forward/backward task requires spatiotemporal integration and therefore is a suitable method to measure performance of biological motion perception.

1.5 Aim of this research

Although it has not been studied before, we expect that biological motion in the form of phosphenes will provide useful information for blind people since Beintema et. al. has proven that biological motion perception does not require local motion. We want to test this by observing whether test subjects are able to perform tasks that require spatial and temporal information integration when being presented with biological motion in the form of phosphenes stimuli. Next to it having the potential to provide high information, the concept of using sparse input for high information is favorable for efficient phosphene induction, reducing unnecessary input and saving the energy put into the electrodes. In this experiment, we replicate one of Beintema's experiments to see the performance of subjects doing the forward/backward task with limited lifetime stimuli along

different frame durations. Because it is not known yet whether all parts of the visual field will be able to be stimulated with the electrodes, we will test the performance with stimuli presented in the center and the lower left quadrant of the visual field. We compare the results of the dot-figure stimuli (from Beintema's experiment) with the phosphene-figure stimuli (made using a stickfigure and phosphene grid) presented with the different frame durations (40ms, 80ms, 120ms, 180ms, and 200ms).

2 Material and Methods

2.1 Set–up and stimuli property

Stimuli were displayed on a CRT monitor (49.53 x 47.244 x 49.276cm, 800 x 600 pixels) at a vertical refresh rate of 96 Hz, on a dark background. In all of the experiments, the head to monitor distance is 45 cm in a darkened room with a chinrest. The walker stimuli in all experiments were made with BioMotion toolbox. Using this toolbox, we generated the joint positions of a human, viewed from the side, walking on a treadmill at 0.625 cycles/sec with a fixation point on the middle of the screen. The stimulation presentation (2 seconds) and response data collection is done with the python library OpenCV2.

For the dot-figures (experiment 1 and 2), positioning on the screen is either center or in the lower left quadrant (LLQ). Let frame represent the coordinates of the stimulus on the monitor, in experiment 1 the coordinates of the walker's position is [0; 0; 0]. In experiment 2 the stimulus is presented in the lower left quadrant (LLQ) position of the screen, center coordinates + [-100; -75; 0]. For both experiments the center of the hips of the walker did not move horizontally. The width and height of the walker spanned a 5.52^o x 10.7^o visual angle. The starting position of the walker was randomized from trial to trial. Local motion signal was removed by removing points after 1 frame and by replacing them at random locations on the underlying stick figure. The possible point positions were distributed uniformly across the limb segments, with each segment defined by the line connecting the joints. The initial position of points on the walker was determined at random. 12 points on the walker were displayed per frame with each point containing a single-frame lifetime.

For experiments 3 (center) and 4 (LLQ), phosphene-figures are used as stimuli. Before applying the phosphene grid, stickfigures were made by coloring in the joint-connections of all segments with 5 pixel thick white lines. The stickfigure contains a head and is presented with a span of 5.52^o x 10.7^o visual angle. The starting position of the walker was randomized from trial to trial. The stickfigure is combined with a 1000 pixel phosphene grid, which introduces a limited lifetime of the points, removing local motion. The locations of the visible points are determined by the structure of the phosphene grid and the walker position.

2.2 Preliminary experiment

Since the concept we're testing is dependent on the settlement that biological motion perception does not require local motion, we first repeated the research that demonstrated it using a limited-lifetime stimuli of one (Beintema et al., 2006). We wanted to see whether we could also demonstrate that subjects are indeed capable of achieving high results in the forward/backward discrimination task, using a short limited-lifetime as shown in the research of Beintema et. al. in 2006. Since the stimuli was self-generated based on the paper, we first wanted to test whether our generated stimuli has the same effect as the original one. Therefore we ran the preliminary experiment, which was to test perception on biological motion with limited lifetime (without any phosphene figures).

2.3 The phosphene grid

The phosphene grid that was used to filter the original stimulus consisted of approximately 1000 phosphenes in a regular grid. Given that most reports of phosphenes mention roundish shapes with vague edges

(Niketeghad et al., 2019) we represented them as Gaussian blobs. For these experiments we simulated the phosphene vision as it would appear in ideal stimulation conditions without noise and variability, rather than more realistic conditions with electrode drop-out, size variability, shape variability, location variability and so on. This decision was made for practical reasons and since these variabilities exhibit differently for different individuals. Further study would be needed to implement a more comprehensive and realistic simulation of phosphene vision achieved with different versions of phosphene grids.

2.4 Procedure

The experimental set-up includes stimuli of 5 different frame durations (40ms, 80ms, 120ms, 160ms, and 200ms). The walker can be walking forward, and backward while facing the right or left side of the screen (original/mirror). Each of these 20 conditions, have 10 replicates each containing different starting coordinates. In a whole walking cycle of the walker, there are 133 frames. Ten starting positions is evenly distributed over the whole walking cycle giving a list: $\mathbf{x} = 0, 13.3, 26.6, 39.9, 53.2, 66.5, 79.8, 93.1, 106.4, 119.7$.

For all of number in x, a random number between 0 and 13 is added: $x_1 + 13R, x_2 + 13R, ..., x_10 + 13R$ with R being a random number between 0 and 1. This is done for each stimulus possibility per frame duration. The final conditions yield 200 different videos per experiment, giving 40 videos per frame duration (See table **??**).

Each experiment started with an instruction screen, and key press test. The screen has a red fixation cross in the center with a black background that remained throughout the trials. The sequence was then presented for a fixed duration (2 sec) after which the stimulus disappeared.

With a force choice task, subjects indicate their choice by pressing one of two possible keys (f or b). The next trial started upon the button response. Each experiment started off with 10 practice trials without feedback to the response. In all experiments, the direction in which the figure was facing and walking (left or right) was balanced over trials. The task is to determine whether the walker is walking forward or backward (relative to its facing direction). The source code of and details on the biological motion stimuli generation and experiment can be found at the GitHub repository *BiologicalMotionExperiment* of account *lelynn*.

2.5 Technical details

Each frame duration varying stimuli (40ms, 80ms, 160ms, 200ms) is generated by first calculating the frame rate of interest `framerate = 1000./[200 160 120 80 40]`, and then setting `myVideo.FrameRate = framerate`. Further calculations are performed with the assumption that there are 120 frames in total per video, and later on the amount of frames will be selected to present (depending on the frame duration). The file is saved with the frame duration included in the filename. Cropping the videos depending on its frame rate is done by:

- `fr = (60/myVideo.FrameRate)*(1:(120/(60/myVideo.FrameRate)))`
- `fr = fr+startfrno`

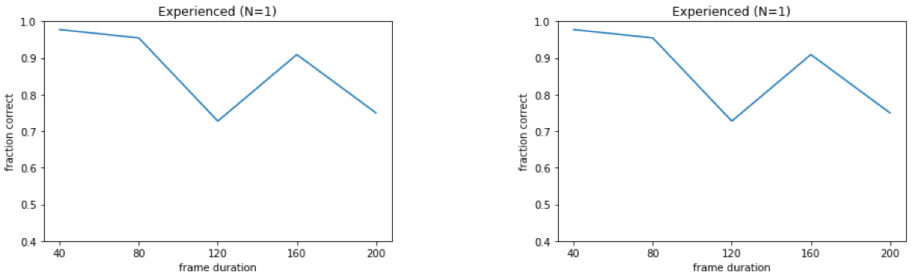
which ensures equally long stimuli of 2 seconds (where `startfrno` is defined as the starting frame number).

For each frame duration condition, the stimuli can walk forward or backward which is indicated by the variables 'Rev' and 'Fwd'. If the walker is walking forward, then `fr = fr` and the saved filename are initiated

with 'Fwd'. If the walker is walking backwards, then for every frame in the video `fr = 120 - fr` and the saved filename will start with 'Rev'.

For each 'Rev and 'Fwd' conditions per frame duration condition, there are original and mirrored images (walker facing left or right), which is indicated by 'Orig' and 'Mirror'. For the original facing direction, the frame remains the same `frame = frame` and the file is saved with the label 'Orig' in its filename. In its mirrored facing direction, each frame is multiplied element-wise by a vector in order to flip the image `frame = frame.* [-1, 1, 1]'`, and file is saved with the label 'Mirror' in its filename. The starting frame number is pseudo-randomized to avoid having recurrent starting position of the walker. This is done by defining the starting frame for each video by: `startfrno = ceil((0:9)*(133/10) + (rand*13)`, resulting in each video having a different starting position (See table **??**) and the filename is saved with the starting frame number.

For the stickfigure images, connections are made between the head to the center of the shoulders, from the center of the shoulders to the hips and from center to end of extremities. This is done by first calculating the average coordinates (in x and y) of the two shoulders, and also the two sides of the hip for each frame. When finding the averages, connections were made (For more detailed information, see `DotFigureGenerator.m` and `StickfigureGenerator.m` codes on the *BiologicalMotionExperiment* repository.



(a) **(b)**

Figure 2: Results of the pilot experiments. (A) Experienced subject has been more often exposed to biological motion. (B) Naïve subject has rarely or never been exposed to biological motion.

3 Results

3.1 Preliminary results

Before the full experiment was performed, we first wanted to see whether the replication of all conditions were correct, and whether out self-generated stimuli would cause similar effects as the one in Beintema's research. Figure 2 gives us the first indications of the effect of our self-generated stimuli on subject's performance. First, an experienced subject is tested and then naïve subjects are tested to see whether effects are similar. We see that for both experienced and naïve subjects, task performances decreases as frame duration increases.

3.2 Dot-figure center results

The present experiment is a replication of one of Beintema's research in 2006, where the role of local motion in biological motion is investigated across different frame duration (Beintema et al., 2006). The subjects need

to discriminate between figures walking (and facing) toward the left or the right, which requires spatial integration of visual information, since discrimination of whether the walker is walking forward or backward is not possible from just a single points and single frames (Beintema et al., 2006). In the present experiment, the forward/backward task experiment is repeated with the lowers possible limited lifetime (limited lifetime of 1) across the different frame durations in which we expect to have the same results as Beintema's study: the longer the frame duration, the lower the performance. Plotted in figure 3 is correct percentage as a function of frame duration for dot-figure stimuli presented on the center of the screen. It shows that indeed the percentage of correct responses in all subjects was strongly reduced with longer frame durations. This suggests that even with the lowest limited lifetime possible, biological motion is still perceivable indicating that motion signals are not a necessary factor.

3.3 Dot-figure LLQ results

Experiment 1 showed that our replication of Beintema's study worked and confirms that the local motion signals are not necessary for the perception of biological motion. However, we want to test whether this stimulation is also perceivable when not presented in the center of the visual field. This experiment is similar to experiment 1, however the stimuli is presented in the lower left quadrant of the monitor where subjects have to fixate their eyes in the middle. If specific fixation location relevant to stimulation area in the visual field is not necessary, it's expected to show similar results as experiment 1, where at lower frame durations subjects perform better than at higher frame durations. Figure 4 shows the mean

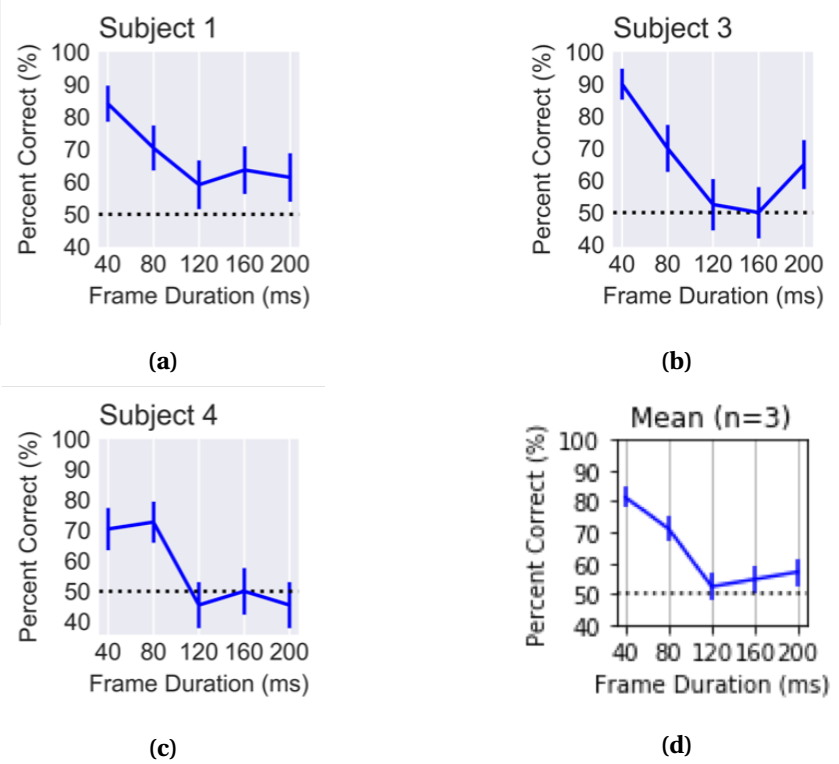


Figure 3: Results center dot experiment. Percentage of correct responses in a forward/backward identification task for dot-figure center stimuli (experiment 1) as a function of frame duration. Points and error bars show +/- standard error of the mean for (A) subject 1, (B) subject 3, (C) subject 4 and (D) the average of all subjects.

correct percentage as a function of frame duration for dot-figure stimuli presented in the lower left quadrant of the screen. Subjects still perform above chance level when shifting the dot-figure on the screen. The task appears to be slightly more difficult, however there is still the same curve present in this image.

3.4 Phosphene-figure center results

It has been found that V1 stimulation always induce stationary phosphenes in human subjects (Silvanto, Cowey, Lavie, & Walsh, 2005). NESTOR's

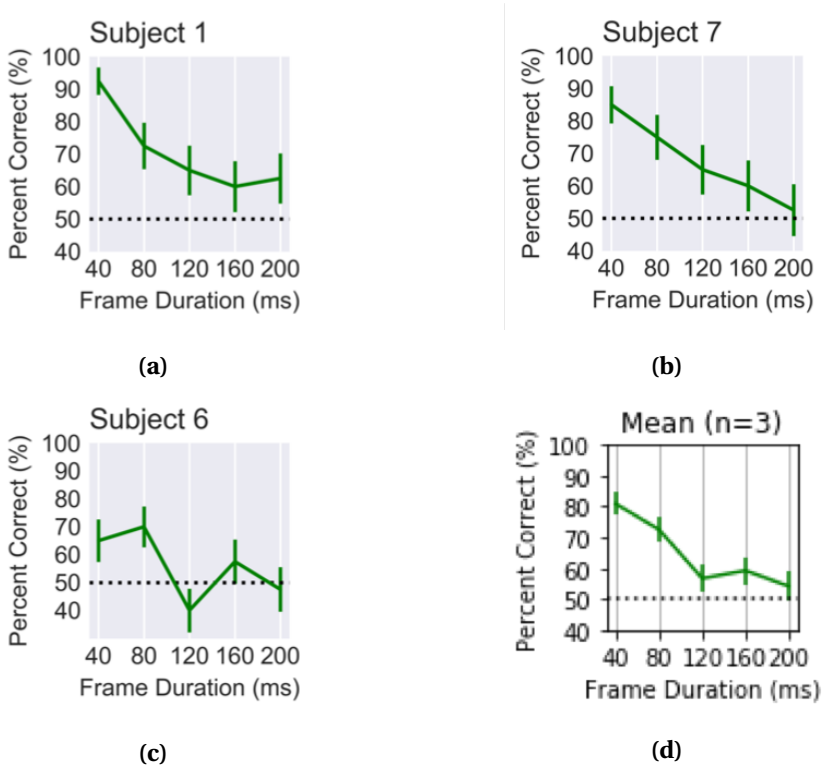


Figure 4: Results LLQ dot experiment. Percentage of correct responses in a forward/backward identification task for dot-figure LLQ stimuli (experiment 2) as a function of frame duration. Points and error bars show +/- standard error of the mean for (A) subject 1, (B) subject 7, (C) subject 6, and (D) the average of all subjects.

cortical implant will target the V1 with a certain frame rate (not yet known), which if fast enough, will be able to mimic motion. Since we do not know how fast and localized the induced phosphenes will appear and leave the perception of a patient, we want to test the performance of doing tasks from simulated phosphene patterns at different frame rates and with a form of limited lifetime. The limited lifetime in this case is different from the limited lifetime mentioned in experiments 1 and 2. Here, instead of the dots randomly reallocating, they enter a phosphene grid, which allows the dots to become visible only at specific areas. The phosphene

grid of 1000 pixels is thus placed on top of the generated stickfigure stimuli, which limits the stickfigure images to dots represented as phosphenes. As in the previous two experiments, the results of experiment 3 shows that the performance degraded with higher frame durations, this time with a less severe decrease (See figure 5). This suggests that simulated phosphene patterns provide useful information even with fewer frames per seconds (higher frame durations).

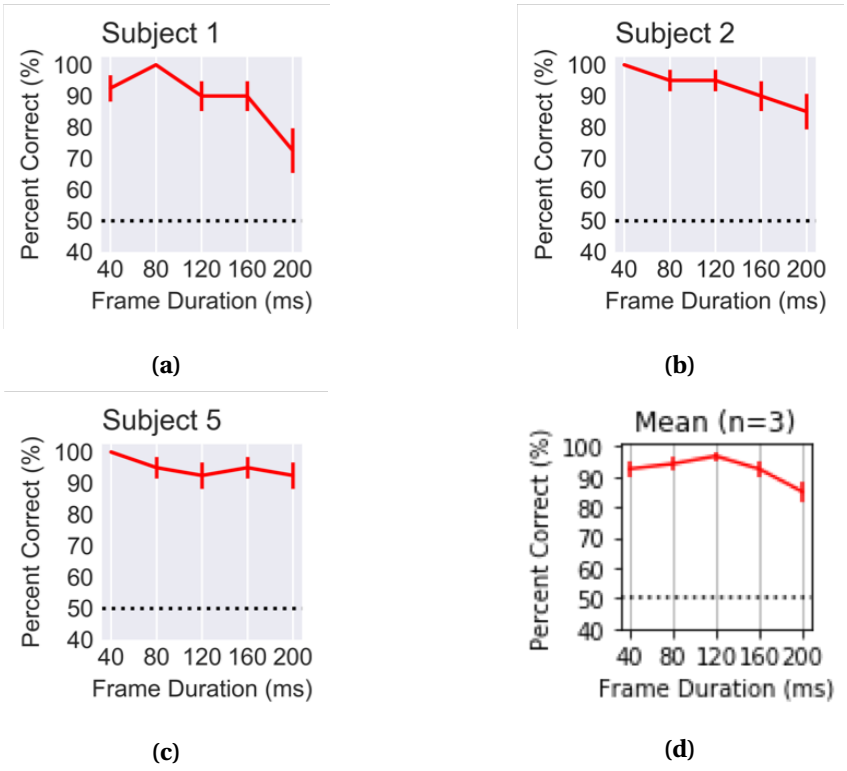


Figure 5: Results center phosphene-figure experiment. Percentage of correct responses in a forward/backward identification task for phosphene-figure center stimuli (experiment 3) as a function of frame duration. Points and error bars show +/- standard error of the mean for (A) subject 1, (B) subject 2, (C) subject 5, and (D) the average of all subjects.

3.5 Phosphene-figure LLQ results

As it isn't certain whether the microelectrode arrays will cover the entire visual cortex for all patients, we therefore carried out an experiment with phosphene-figure in the lower left quadrant. Since subjects performed well on the task with central phosphene-figure (experiment 3), we expect that this task will result in higher performance compared to LLQ dot-figures (experiment 2). The results of this task, shown in figure 6 testifies that subjects performed high above chance level when having to spatiotemporally integrate information with phosphene-figures presented in the lower left quadrant onto the visual field. On average, performance remained high even at longer frame duration.

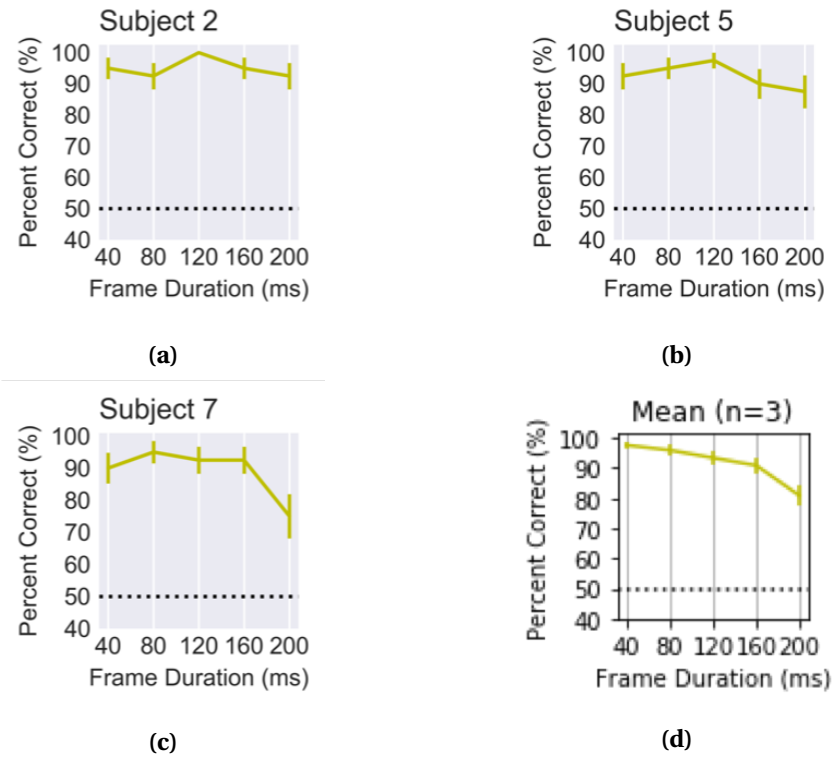


Figure 6: Results LLQ phosphene-figure experiment. Percentage of correct responses in a forward/backward identification task for phosphene-figure LLQ stimuli (experiment 4) as a function of frame duration. Points and error bars show +/- standard error of the mean for (A) subject 2, (B) subject 5, (C) subject 7, and (D) the average of all subjects.

4 Discussion and conclusion

Discussion We studied the use of phosphene representations in discriminating biological motion with limited lifetime in the forward/backward task. Our question was whether phosphene representations of biological motion could aid in obtaining spatial and temporal information from stickfigure walkers. To answer this question, we measured subjects' performance on the forward/backward discrimination tasks using stickfig-

ures in a phosphene grid as stimuli. We compared the performance of correct discrimination between results from experiments with dot-figures stimuli and phosphene-figures stimuli. Next to comparing the two types of limited lifetime stimuli, we also compared the presentation of the walkers on the center of the visual field and in the lower left quadrant of the visual field.

The effect of removing local motion from biological motion has earlier been studied by Beintema et al. in 2006. They found no effect of lifetime on the mean correct rate at any frame duration. Performance did not increase with longer lifetime, therefore motion signals do not play a role (Beintema et al., 2006). We performed the experiment with a single-frame limited lifetime (lowest lifetime possible) across different frame durations in experiment 1 and 2. In both experiments revealed high above chance level performance on frame durations 40ms and 80ms. Performance on higher frame durations degraded. For experiments 3 and 4, a phosphene grid reduced the local motion of a stickfigure, resulting in a phosphene-figure. The performance of subjects in these experiments was even higher (with all frame durations above chance level) than the performance with the dot-figure stimuli. Thus, we found that there was no evidence of the necessity of local motion information for performing a task that requires motion integration in either the dot-figure stimuli or the phosphene-figure stimuli.

Not only do our results so far confirm the that the role of local motion is not of great importance, we also see that it is not essential for the figure to be in the center of the visual field in order to obtain information from biological motion without local motion. Implanting microelectrode arrays into the visual cortex does not guarantee full coverage of the visual field. We wanted to test whether this stimuli is still informative when

presented onto a part of the visual field outside of the fixation area, provided the phosphene representations does not turn out to be in the center of the visual field. Experiments 2 and 3 are therefore carried out and subjects performing LLQ stimuli are shown to perform evenly well with center stimuli, as shown in figure 7.

In the phosphene-figure, there are more dots provided (1000 pixels) whereas the dot-figure contains 12 dots per frame only. Also, the limited lifetime in the phosphene-figures are different from the dot-figure such that the dot-figure stimuli consists of dots reallocating at random positions after each frame, whereas in the phosphene-figure, the dots are spatially relevant after each frame. It is likely that these factors caused our results. It is believed that the NESTOR cortical implant is able to provide many phosphenes percepts, however testing the percept of biological motion gives us insight of the information one could obtain from just a few dots. This notion shows how important a few dots can be in providing information.

Future perspectives A great deal of future investigation can be done including analyzing different viewing positioning of the walker, stimuli size, dots sizes, varying lifetimes, using real motion stimuli, and different positions in the visual field. An increasing amount of computer vision projects dealing with posture-detection have been developed in the recent years. In 2017, Zhe Cao's research presented efficient real-time algorithms to detect 2D pose of multiple people in images using a two-branch multi-stage convolutional neural network (Cao, Simon, Wei, & Sheikh, 2017). Their design can be used as inspiration for us to develop methods to send biological motion representations as input for the implant. Eventually we want to test the method using augmented and virtual reality phosphene vision simulations, with the ultimate goal of integrating

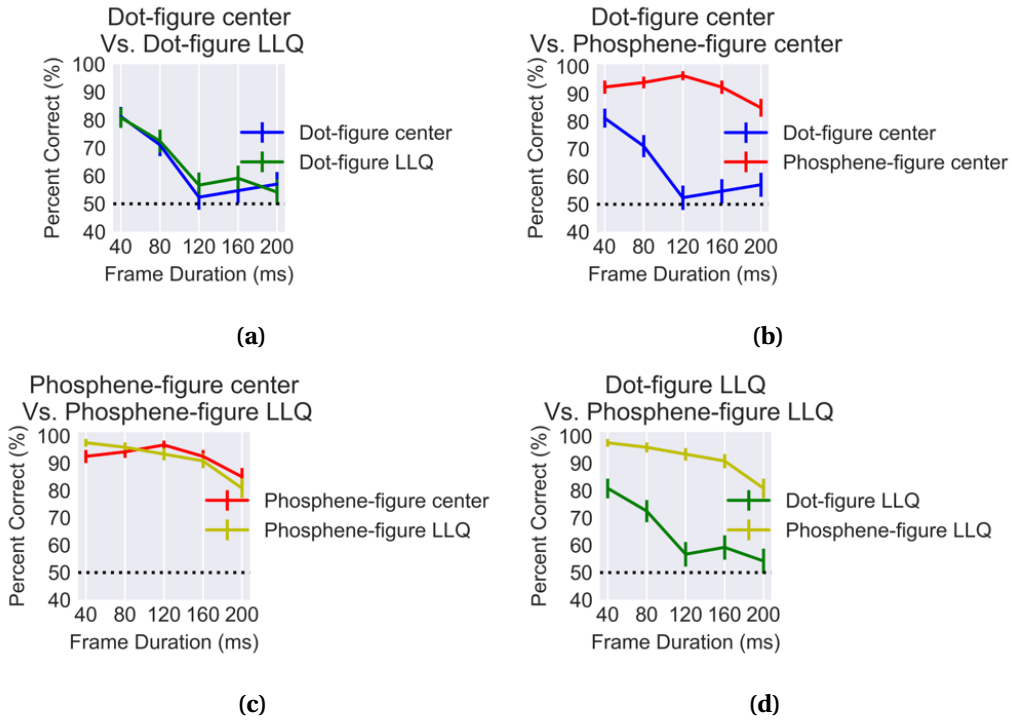


Figure 7: Comparison of the means over subject's results for all experiments. (A) The results of experiment 1 is compared to the results of experiment 2. (B) Results of experiment 1 is compared with experiment 3 results. (C) Experiment 3 result is compared with the result of experiment 4. (C) Experiment 2 result is compared with the result of experiment 4. Error bars represent +/- standard error of the mean.

the concept into the cortical visual prosthesis that is under development.

4.1 Conclusion

For the discrimination in the forward/backward task with phosphene-figures, there is no need for a local motion mechanism nor for it to be in the central of the visual field. The experiments show that using a stickfigure to translate into a phosphene simulation provides even better infor-

mation than biological motion stimuli. Regardless, just 12 dots already provides crucial information for human subjects suggesting that the concept of prosthetic vision likely to be very advantageous for blind people.

5 Acknowledgments

I would first like to thank my supervisor Dr. Yağmur Güçlütürk for her great supervision and steering me into the right direction. I would also like to thank Prof. Dr. Richard van Wezel for always having good and interesting advice and allowing me to do the project in his department. And lastly, thanks to the department of Biophysics for a peaceful atmosphere.

References

Bak, M., Girvin, J., Hambrecht, F., Kufta, C., Loeb, G., & Schmidt, E. (1990). Visual sensations produced by intracortical microstimulation of the human occipital cortex. *Medical and Biological Engineering and Computing*, *28*(3), 257–259.

Beintema, J., Georg, K., & Lappe, M. (2006). Perception of biological motion from limited-lifetime stimuli. *Perception & Psychophysics*, *68*(4), 613–624.

Beintema, J., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences*, *99*(8), 5661–5663.

Brindley, G. S., & Lewin, W. (1968). The sensations produced by electrical stimulation of the visual cortex. *The Journal of physiology*, *196*(2), 479–493.

Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 7291–7299).

Cudeiro, J., & Sillito, A. M. (2006). Looking back: corticothalamic feedback and early visual processing. *Trends in neurosciences*, *29*(6), 298–306.

Delbeke, J., Oozeer, M., & Veraart, C. (2003). Position, size and luminosity of phosphenes generated by direct optic nerve stimulation. *Vision research*, *43*(9), 1091–1102.

de Xivry, J.-J. O., Coppe, S., Lefèvre, P., & Missal, M. (2010). Biolog-

ical motion drives perception and action. *Journal of vision*, *10*(2), 6–6.

Dobelle, W. H. (2000). Artificial vision for the blind by connecting a television camera to the visual cortex. *ASAIO journal*, *46*(1), 3–9.

Mazzola, L., Isnard, J., Peyron, R., & Mauguière, F. (2011). Stimulation of the human cortex and the experience of pain: Wilder penfield's observations revisited. *Brain*, *135*(2), 631–640.

Niketeghad, S., Muralidharan, A., Patel, U., Dorn, J. D., Bonelli, L., Greenberg, R. J., & Pouratian, N. (2019). Phosphene perceptions and safety of chronic visual cortex stimulation in a blind subject. *Journal of neurosurgery*, *1*(aop), 1–8.

Pollen, D. (1975). Some perceptual effects of electrical stimulation of the visual cortex in man. *The nervous system*, *2*, 519–528.

Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (v1) activity gates awareness of motion. *Nature neuroscience*, *8*(2), 143.

Weiland, J. D., & Humayun, M. S. (2014). Retinal prosthesis. *IEEE Transactions on Biomedical Engineering*, *61*(5), 1412–1424.

Wiesel, T. N., & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of neurophysiology*, *29*(6), 1115–1156.

Zeng, F.-G., Rebscher, S., Harrison, W., Sun, X., & Feng, H. (2008). Cochlear implants: system design, integration, and evaluation. *IEEE reviews in biomedical engineering*, *1*, 115–142.

A Appendix

Frame duration	Forward	Backward	Original	Mirror
40ms	$\{x_1 + 13R_1,$ $x_2 + 13R_1,$ $...,$ $x_{10} + 13R_1\}$	$\{x_1 + 13R_2,$ $x_2 + 13R_2,$ $...,$ $x_{10} + 13R_2\}$	$\{x_1 + 13R_3,$ $x_2 + 13R_3,$ $...,$ $x_{10} + 13R_3\}$	$\{x_1 + 13R_4,$ $x_2 + 13R_4,$ $...,$ $x_{10} + 13R_4\}$
80ms	$\{x_1 + 13R_5,$ $x_2 + 13R_5,$ $...,$ $x_{10} + 13R_5\}$	$\{x_1 + 13R_6,$ $x_2 + 13R_6,$ $...,$ $x_{10} + 13R_6\}$	$\{x_1 + 13R_7,$ $x_2 + 13R_7,$ $...,$ $x_{10} + 13R_7\}$	$\{x_1 + 13R_8,$ $x_2 + 13R_8,$ $...,$ $x_{10} + 13R_8\}$
120ms	$\{x_1 + 13R_9,$ $x_2 + 13R_9,$ $...,$ $x_{10} + 13R_9\}$	$\{x_1 + 13R_{10},$ $x_2 + 13R_{10},$ $...,$ $x_{10} + 13R_{10}\}$	$\{x_1 + 13R_{11},$ $x_2 + 13R_{11},$ $...,$ $x_{10} + 13R_{11}\}$	$\{x_1 + 13R_{12},$ $x_2 + 13R_{12},$ $...,$ $x_{10} + 13R_{12}\}$
160ms	$\{x_1 + 13R_{13},$ $x_2 + 13R_{13},$ $...,$ $x_{10} + 13R_{13}\}$	$\{x_1 + 13R_{14},$ $x_2 + 13R_{14},$ $...,$ $x_{10} + 13R_{14}\}$	$\{x_1 + 13R_{15},$ $x_2 + 13R_{15},$ $...,$ $x_{10} + 13R_{15}\}$	$\{x_1 + 13R_{16},$ $x_2 + 13R_{16},$ $...,$ $x_{10} + 13R_{16}\}$
200ms	$\{x_1 + 13R_{17},$ $x_2 + 13R_{17},$ $...,$ $x_{10} + 13R_{17}\}$	$\{x_1 + 13R_{18},$ $x_2 + 13R_{18},$ $...,$ $x_{10} + 13R_{18}\}$	$\{x_1 + 13R_{19},$ $x_2 + 13R_{19},$ $...,$ $x_{10} + 13R_{19}\}$	$\{x_1 + 13R_{20},$ $x_2 + 13R_{20},$ $...,$ $x_{20} + 13R_{20}\}$

Table 1: Table showing each condition of the 200 videos in a trial.