DOI: 10.1088/1741-2560/11/6/066002

Improved visual performance in letter perception through edge orientation encoding in a retinal prosthesis simulation.

F. Isabell Kiral-Kornek[1,2,3], Elma O'Sullivan-Greene[1,2], Craig O. Savage[1,2], Chris McCarthy[1,2,4,5], David B. Grayden[1,2,3,6], and Anthony N. Burkitt[1,2,3,6]

Affiliations: [1]NeuroEngineering Laboratory, Department of Electrical & Electronic Engineering, The University of Melbourne.
[2]Centre for Neural Engineering, The University of Melbourne.
[3]NICTA, c/- Department of Electrical & Electronic Engineering, The University of Melbourne.
[4]NICTA Canberra Research Laboratory, 7 London Circuit, Canberra ACT 2601.
[5]College of Engineering and Computer Science, Australian National University, Canberra, ACT 2601, Australia.
[6]Bionics Institute, 384-388 Albert St, East Melbourne VIC 3010, Australia.

Objective. Stimulation strategies for retinal prostheses predominately seek to directly encode image brightness values rather than edge orientations. Recent work suggests that the generation of oriented elliptical phosphenes may be possible by controlling interactions between neighboring electrodes. Based on this, we propose a novel stimulation strategy for prosthetic vision that extracts edge orientation information from the intensity image and encodes it as oriented elliptical phosphenes. We test the hypothesis that encoding edge orientation via oriented elliptical phosphenes leads to better alphabetic letter recognition than standard intensity-based encoding. *Approach.* We conduct a psychophysical study with simulated phosphene vision with 12 normal-sighted volunteers. The two stimulation strategies were compared with variations of letter size, electrode drop-out and spatial offsets of phosphenes. *Main results.* Mean letter recognition accuracy was significantly better with the new proposed stimulation strategy (65%) compared to direct greyscale encoding (47%). All examined parameters - stimulus size, phosphene dropout, and location shift - were found to influence the performance, with significant two-way interactions between phosphene dropout and stimulus size as well as between phosphene dropout and phosphene location shift. The analysis delivers a model of perception performance. *Significance.* Displaying available directional information to an implant user may improve their visual performance. We present a model for designing a stimulation strategy under the constraints of existing retinal prostheses that can be exploited by retinal implant developers to strategically employ oriented phosphenes.

1 Introduction

Since the discovery that electrical stimulation of the visual pathway can lead to visual perception by Foerster (1929) and a thorough investigation of this effect by (Brindley and Lewin, 1968), many research groups around the world have been working on the development of a bionic eye for vision impaired people (Humayun et al., 1999; Wyatt and Rizzo, 1996; Zrenner et al., 2011; Suaning et al., 1998). Current research of retinal stimulation techniques target two degenerative retinal diseases: retinitis pigmentosa (RP) and age-related macular degeneration (AMD). Typical of both conditions is the loss of photoreceptors, the light-sensitive cells in the eye. In the case of RP, it has been shown in many patients that the remaining cells, usually retinal ganglion cells, can be electrically stimulated with a chronically implanted array of electrodes close to the retina. Retinal stimulation has so far led to a visual percept in one patient with dry-form macular degeneration (Humayun et al. (1999)) in an acute experiment only.The goal in either case is to convert an input signal retrieved by a camera to electrical signals that are interpreted as visual input by the patient's brain. The resulting visual percepts are called phosphenes. Phosphenes evoked by electrical stimulation are often considered as uniformly round dots of white or yellow brightness in an idealized sense, but individual phosphenes have been shown to have varying intensity, shape, and sometimes color (Horsager and Fine, 2011; Wilke et al., 2011a). Current density has previously been tied to the perception of phosphene in retinal implants (Wilke et al., 2011b; Abramian et al., 2010). Analyses by Abramian et al. (2011) further support the use of finite element models in estimating ganglion cell activation.

Different stimulation strategies have been proposed that make use of the relationship between the stimulation parameters and the perceived visual impression. Most strategies rely on the assumption that round phosphenes of different intensity levels can deliberately be elicited by changing the amplitude or pulse rate of the stimulation current. These phosphenes can then be composed into an image, similar to how pixels on a computer screen are used. Due to the variety of reported phosphene perceptions, and the limited number of implantees, research is often done with simulated phosphenes that match descriptions of reported phosphenes.

In this work, we explore a different phosphene property, that of phosphene orientation in simulation. Different approaches to elicit oriented phosphenes are described by Horsager et al. (2010), Opie et al. (2013) and Savage et al. (2012). Horsager et al. (2010) hypothesize that the stimulation order can lead to differently oriented phosphenes. Psychophysical tests in patients show that the order in which electrodes are stimulated leads to different distinguishable percepts. Weitz et al. (2013) demonstrate in an *in vivo* experiment that different stimulus pulse shapes can influence the pattern of excited retinal ganglion cells. In finite element simulations, Savage et al. (2012) create stretched stimulation fields. The interaction between several adjacent electrodes that are being stimulated simultaneously leads to differently shaped areas of high current density in the simulated tissue. In particular, Savage et al. (2012) show that flanking a stimulating electrode with two return electrodes can lead to a current spread that resembles the shape of an ellipse. If current density is taken as an indicator for visual perception, this will also lead to the perception of differently shaped phosphenes.

Our objective here is to investigate the effect of including additional directional information using phosphene shapes. Thus, we construct directional phosphenes in a controlled manner using elliptically shaped phosphene simulations shown to normal sighted volunteers. Although our model does not account for physiological factors, such as axonal stimulation (that can lead to complex and distorted shapes rather than ideal rounded phosphenes), it gives us the opportunity to study the benefit of including additional directional information into the stimulation strategy and study the effects on a fundamental level. Moreover, while realistic phosphenes do have varying and complex shapes due to physiological factors, many will still contain a net vector direction, which could be exploited in future strategies (as suggested by Kiral-Kornek et al. (2013)) should our fundamental work in this paper indicate that directional information contains true perceptual advantages. While this simulation model cannot be expected to accurately predict patient outcomes, we expect that the relative performance between strategies would be similar in real implant patients, particularly with high acuity electrode arrays. Moreover, a device with a large number of individually controllable electrodes could potentially enable more controlled steering of current, which would make this strategy easier to implement, since defective electrodes could potentially be bridged due to the closeness of surrounding electrodes.

Chen et al. (2009) give a general overview of the different phosphene models that have been used in

simulations. It is commonly assumed that each electrode can elicit one round phosphene. In simulations that use this assumption, the intensity of a phosphene decays according to a two-dimensional Gaussian distribution, and the grid in which phosphenes are represented is often the same as that of the electrodes of the visual prosthesis.

In order to offer a basic comparison of the two stimulation strategies, we conduct a psychophysical experiment with simulated phosphene vision of letters. This stimulus (letters) was chosen, because of its widespread use as a visual performance measure in patients. Moreover, it inherently contains strong edges along which phosphenes can be oriented and, therefore, make pre-processing of the stimulus unnecessary.

For the remainder of this paper, we will assume that we can elicit a phosphene that is either round or elliptical in shape. Based on this model of stimulation, we propose and test a stimulation strategy that may utilize this ability to influence phosphene shape in order to improve accuracy on a letter recognition task. Specifically, we propose the use of oriented elliptical phosphenes that most closely align with a letter's edges in order to provide additional directional information.

In our experiment, two different distortion factors are investigated that may occur when wearing a retinal implant. Nonfunctional electrodes or electrodes that have an unsafely high stimulation threshold to evoke a percept are modeled with phosphene dropout.A phosphene location shift is introduced to model spatial deviations between the stimulation location on the retina and the perception of a phosphene in the user's visual field, as reported by Humayun et al. (2003b). In addition, we vary the sizes of the stimuli in line with the different sizes on a standard vision chart (British Standards Institute, 2003). This study is methodologically similar to a study performed by Zhao et al. (2011), in which the effect of different distortion parameters on Chinese character recognition was measured. We propose a new stimulation strategy that displays directional information.

The hypotheses we test are as follows. *Hypothesis 1:* Using elliptical phosphenes instead of uniformly round phosphenes to represent edges of letters results in a significantly increased percentage of correctly identified letters in a letter recognition task. *Hypothesis 2:* The degree that the directional stimulation strategy is better than the uniform stimulation strategy depends on dropout, location shift, and/or stimulus size, i.e., there are considerable two-way interactions between the strategy and distorting variables. *Hypothesis 3:* There exist considerable two-way interactions between the independent variables dropout, location shift, and stimulus size, i.e., the effect of one independent variable on letter recognition accuracy can depend on another independent variable.

The remainder of this paper is structured as follows. In Section 2, the generation of the stimuli that are used to simulate bionic vision is described. These stimuli are presented to study participants in a psychophysical experiment. The section further outlines the statistical analysis used. Results are presented, and analyzed in Section 3 and, from this, we propose a complete and a simplified model to assist the design of future stimulation strategies. In Section 4, we discuss implications and limitations of our results. We conclude with Section 5.

2 Methods

2.1 Stimulus generation

Two different simulated stimulation strategies are tested. In a conventional strategy, phosphene intensity simply reflects the intensity value in an image at the location at which the phosphene is assumed to be evoked. This strategy will be simulated using only round phosphenes and referred to as the "uniform" stimulation strategy, S_u. We compare this strategy to one in which the orientation of edges in the input image is represented using oriented elliptical phosphenes, called "directional" stimulation strategy, S_d.

2.1.1 Uniform stimulation strategy, S_u

The greyscale input image, $I(\mathbf{x})$, with $\mathbf{x} = (x, y)$ as the center position of the electrode, serves as the input image for the extraction of phosphene intensity. In addition, to minimize the impact of spatial noise, the input image is convolved with a two-dimensional Gaussian filter, $G(\mathbf{x})$, with standard deviation σ and a mean

of zero:

$$I(\mathbf{x}) = I_m(\mathbf{x}) * G(\mathbf{x}), \tag{1}$$

with

$$G(\mathbf{x}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2+y^2}{2\sigma^2}}. \tag{2}$$

The filter size is set to 20 by 20 pixels with a standard deviation of 6 pixels.

A grid of equidistant locations is set up based on electrode positions in a 16 by 16 electrode array. The intensity value at the center position, $\widehat{\mathbf{x}}$, of each phosphene, $P(\mathbf{x})$, is set to the value of $I(\widehat{\mathbf{x}})$. The intensity level is then quantized to either the number of stimulation bits available for conveying intensity information to the implant or the number of intensity levels a patient can perceive. In our experiment, we quantize to 6 bits, which is chosen to be higher than the number of levels between which patients were able to differentiate brightness (Humayun et al., 2003a). These operations result in a regular grid of phosphenes with different intensity values at the positions of the electrodes. An example of a stimulus created with the uniform stimulation strategy can be seen in Figure 1(a).

2.1.2 Directional stimulation strategy, S_d

For the directional stimulation strategy, S_d, the orientation of the letter outline is retrieved. This is done by filtering the input image using the Sobel edge detection algorithm (Duda et al., 1973). The edge orientation for each electrode position is estimated using the Matlab (version 2011b) function "regionprops". In our simulations, we limited the number of different orientations to eight. Assuming a stimulation strategy that relies on electric cross-talk as presented by Savage et al. (2012), this is the maximum perceivable number of phosphene orientations that can be produced in a rectangular regular grid of 3 x 3 electrodes. Elliptical phosphenes are created using a two-dimensional Gaussian distribution with different standard deviations along the two major axes. The phosphene orientation closest aligned with the edge orientation at this image position is chosen. They are subsequently normalized to the same intensity as a corresponding round phosphene at this specific location. An example of a stimulus created with the directional stimulation strategy can be seen in Figure 1(e).

2.1.3 Stimulus variables

Different stimulation sizes and two different distortions are modeled in the experiment: phosphene dropout, d and location shift of phosphenes. Different letter sizes, s, as appear on a standard Snellen chart to measure visual acuity (Snellen, 1863), are also simulated. Naturally, this also results in a smaller number of simulated phosphenes, see Figure 1(b) and 1(f). Phosphene dropout is specified as the number of actual phosphenes and is reported as percentage of the total number of possible phosphenes. The specified number of phosphenes are randomly selected and set to zero intensity. Figure 1(c) and 1(g) display a 30% dropout. The horizontal and vertical location shift, δx and δy, respectively, of each phosphene is controlled through the standard deviation σ_l and is calculated as

$$\begin{aligned} \delta x &= \sigma_l \sqrt{(2ln(\frac{1}{1-U_1}))} cos(2\pi U_2) \\ \delta y &= \sigma_l \sqrt{(2ln(\frac{1}{1-U_3}))} cos(2\pi U_4), \end{aligned} \tag{3}$$

with random variables U_n, $n \in \{1, 2, 3, 4\}$. Example of a location shift of $\sigma_l = 10$ pixels can be found in Figure 1(d) and 1(h). Each of these variables is tested using round phosphenes and oriented elliptical phosphenes for accentuation of edges.

2.2 Psychophysical experiments

A psychophysical experiment with simulated prosthetic vision is conducted to compare the stimulation strategies, S_u and S_d. Twelve participants were recruited. None of the participants had any previous exposure to simulated phosphene vision. Subjects had normal or corrected-to-normal vision, which was measured with

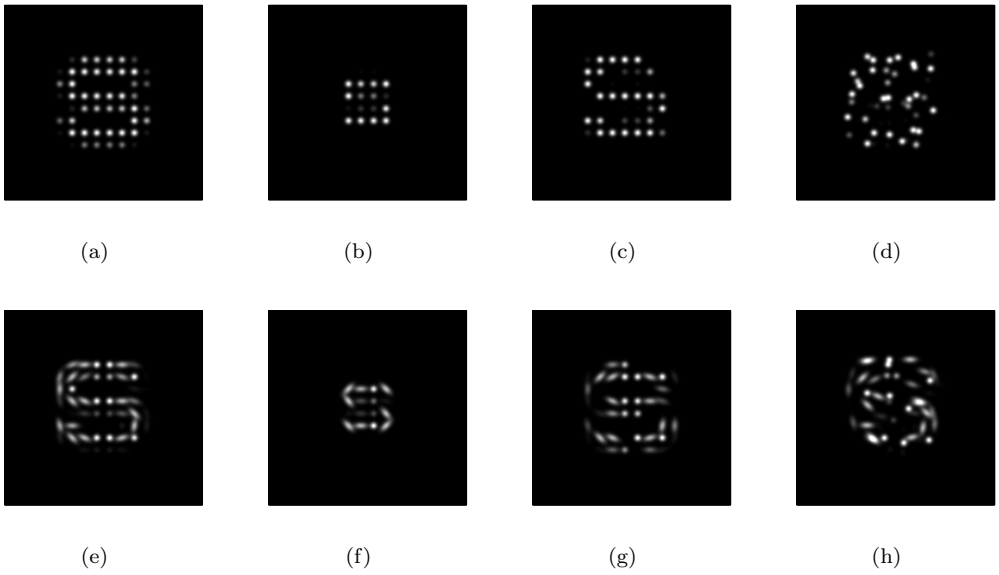


Figure 1: Visualization of the independent variables phosphene dropout, phosphene location shift, stimulus size and stimulation strategy. During the experiments stimuli are distorted with a combination of the independent variables. (a) to (d) uniform stimulation strategy S_u, (e) to (h) directional stimulation strategy S_d. (a), (e) no distortion, (b), (f) stimulation size $s = 1.5$ degrees of the visual angle, (c), (g) $d = 30\%$ dropout, (d), (h) location shift $\sigma_l = 10$ pixels.

the LandoltC test, using Freiburg Vision Test software "FrACT" (Bach, 1996) before the experiment was conducted. The study was approved by the University of Melbourne Human Research Ethics Committee under the approval number HREC 1238453.3.

Stimuli were presented on a Dell UltraSharp U2312HM 58.4 cm LED monitor. To keep a specified distance to the screen, each observer rested their head on a chin support that was fixed to a table. The viewing distance was 57 cm for every experiment. Stimuli were preceded and succeeded by a white noise mask image for 500 ms. The stimuli were presented for 160 ms, a time chosen to avoid the possibility of eye movements and image scanning but above the critical perception time established by Efron (1970).

Psychophysical experiments were conducted using the method of constant stimuli (Urban, 1910). For this, a limited number of stimuli were presented to each participant, using predefined levels for each independent variable. The four letters "N", "H", "R", and "S" were used. These letters were chosen as per the British Standard 4274-1:2003 (British Standards Institute, 2003) for clinical determination of distance visual acuity. The independent variables were stimulus size, phosphene dropout, and spatial phosphene shift. The dependent variable was the percentage of correctly identified letters. Tested stimulus sizes were chosen to cover 1.5, 2, and 2.5 degrees of the visual angle. Dropout was varied between 10 and 90 percent in four equidistant steps. The standard deviation, σ_l, of the spatial shift spanned from 0 to 20 pixels in four equidistant steps. These values were not chosen to reflect a realistic implant, but were determined in a pilot study to best reveal the underlying psychometric function for each variable.

For each combination of stimulation strategy and stimulus size, visual stimuli were presented to the participant in a random order. Each letter was presented the same number of times for each combination of variables.

After a short training session in which 8 stimuli were presented to the observers, the experiment was conducted in 6 blocks of 100 stimuli each with breaks after every 50 stimuli. During each block, the stimulus size and the stimulation strategy under test were kept constant. The order of stimuli in each block as well as the order of blocks differed between participants and were predetermined using general factorial design. Answers were typed on the number pad of a keyboard, the letters were associated with the numbers in the following way: N $\mapsto$ 8, H $\mapsto$ 4, R $\mapsto$ 6, and S $\mapsto$ 2. To facilitate this assignment, printed labels were attached to the screen in the according direction. Total time for an individual session was approximately 1.5 hours.

2.3 Statistical analysis

To analyze the recorded data, a regression analysis is performed using the statistical analysis software GenStat. In order to investigate the statistical properties and the existence of interactions between the independent variables, the generalized linear model M_c is used:

$$M_c : f_c = c_0 + \boldsymbol{\alpha}\mathbf{q}^{\mathrm{T}} + \boldsymbol{\beta}\mathbf{p}^{\mathrm{T}}, \tag{4}$$

with $(\cdot)^{\mathrm{T}}$ denoting vector transpose. The model consists of a constant c_0 and two weighting vectors of which $\boldsymbol{\alpha}$ describes one-way interactions and $\boldsymbol{\beta}$ two-way interactions. The vectors $\mathbf{q}$ and $\mathbf{p}$ are variable vectors for one- and two-way interactions, respectively. Since we record letter recognition rates, binomial data is measured, as the results are either correct or incorrect. To account for this nonlinear function in the linear model, the axes are transformed using the logistic link function $g(x) = (x/(100-x))$ before a weighted linear regression is performed (Jaeger, 2008). All variants are normalized to [0,1]. The influence of the variants (phosphene dropout, spatial phosphene jitter, stimulus size) and factors (participant ID, stimulation strategy) on the successful identification of a letter is analyzed using Student's t-test. Only the influence of first and second-order terms on the success rate f are modeled. Two-way interactions between the participant and other variables are not taken into account.

The weighting vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are model specific. The variable vectors capture all different interactions between stimulation strategy $S = \{S_u, S_d\}$, phosphene dropout, (subscript d), phosphene location shift, (subscript l), and stimulus size, (subscript s), with $\mathbf{q} = [q_S, q_d, q_l, q_s]$ and $\mathbf{p} = [p_{S*d}, q_{S*l}, q_{S*s}, q_{d*l}, q_{l*s}, \}_{d*s}]$, where q_{a*b} denotes the two-way interaction between the variables a and b.

3 Results

The recorded data is displayed in Figure 2. For each combination of size and stimulation strategy, the fraction of correctly identified letters is displayed in color as a function of phosphene dropout and phosphene location shift. Figure 3 displays the difference between the fraction of correctly identified letters for each stimulation strategy. Data is averaged over different letters and all participants. The results as depicted in Figures 2 and 3 show that, on average, letter recognition accuracy was increased using the directional strategy. For all sizes, the fraction of correctly identified letters generally drops with increasing phosphene dropout and location shift.

Figure 4 shows the difference in performance for strategies S_u and S_d. The bars indicate the standard error. Because performance was measured as a binary variable (correct, incorrect), the standard error was found to be more descriptive than the standard deviation, which spans almost the entire region between zero (incorrect answer) and one (correct answer). Averaged over 3600 trials for each stimulation strategy, the strategy S_u has a mean of 47% correctly identified letters, with a standard deviation of 0.50 and a standard error of 0.0083. The mean of stimulation strategy S_d is 65% with a standard deviation of 0.48 and a standard error of 0.0080. The p-value is ≤ 0.001. Thus, Hypothesis 1, which stated that there will be a significant performance increase with the directional strategy, is supported by the data.

The recorded data was modeled as outlined in Section 2.3. An initial analysis showed that participant is not a significant factor (p-values ranging from 0.02 to 0.95). Thus, we only focus on the other interactions. The fitted model parameters c_0, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, p-values, and standard error for the complete model M_c can be found in Table 1. An example of how to apply the model is given in Section 3.1.1.

Hypothesis 2 states that there are significant interactions between the stimulation strategy and the independent variables; stimulus size, phosphene dropout, and location shift. To test this, we look at the two-way interaction between the stimulation strategy and all influencing factors. With p-values larger than 0.05, we can not conclude that there are significant interactions between the stimulation strategy and the stimulus size. Hypothesis 2 is not supported for the interaction between stimulation strategy and stimulus size.

Hypothesis 3 states that the degree to which letter recognition accuracy is influenced by dropout, location shift, and image size will depend on significant two-way interactions between these independent variables. In support of Hypothesis 3, corresponding p-values, shown in the last rows in Table 1, indicate significance.

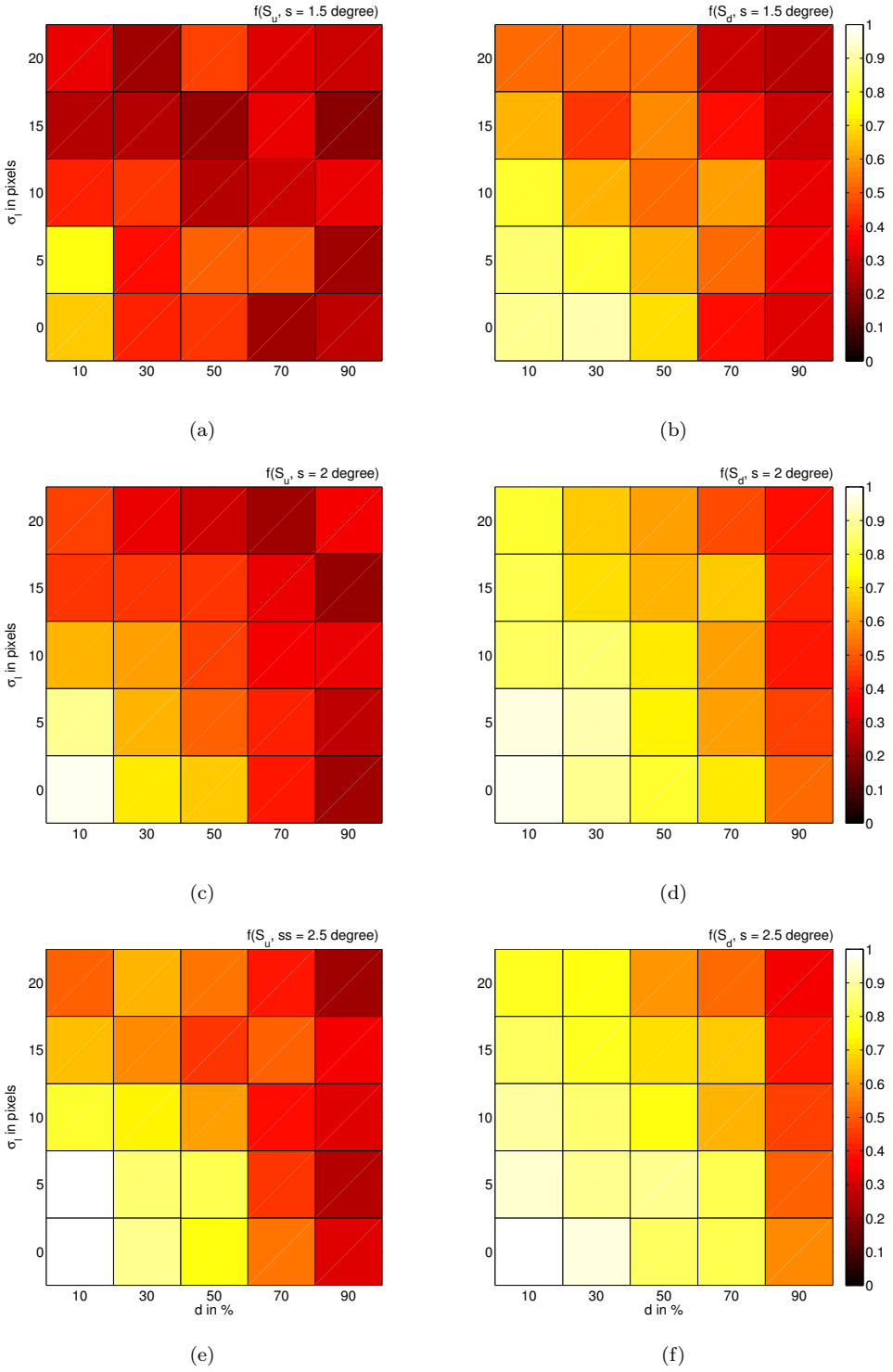


Figure 2: Recorded data. The color indicates the fraction of correctly identified letters f as a function of phosphene dropout d and phosphene location shift σ_l. Lighter shades equal better performance. (a) S_u, 1.5 degree visual angle, (b) S_d, 1.5 degree visual angle, (c) S_u, 2 degree visual angle, (d) S_d, 2 degree visual angle, (e) S_u, 2.5 degree visual angle, (f) S_d, 2.5 degree visual angle.

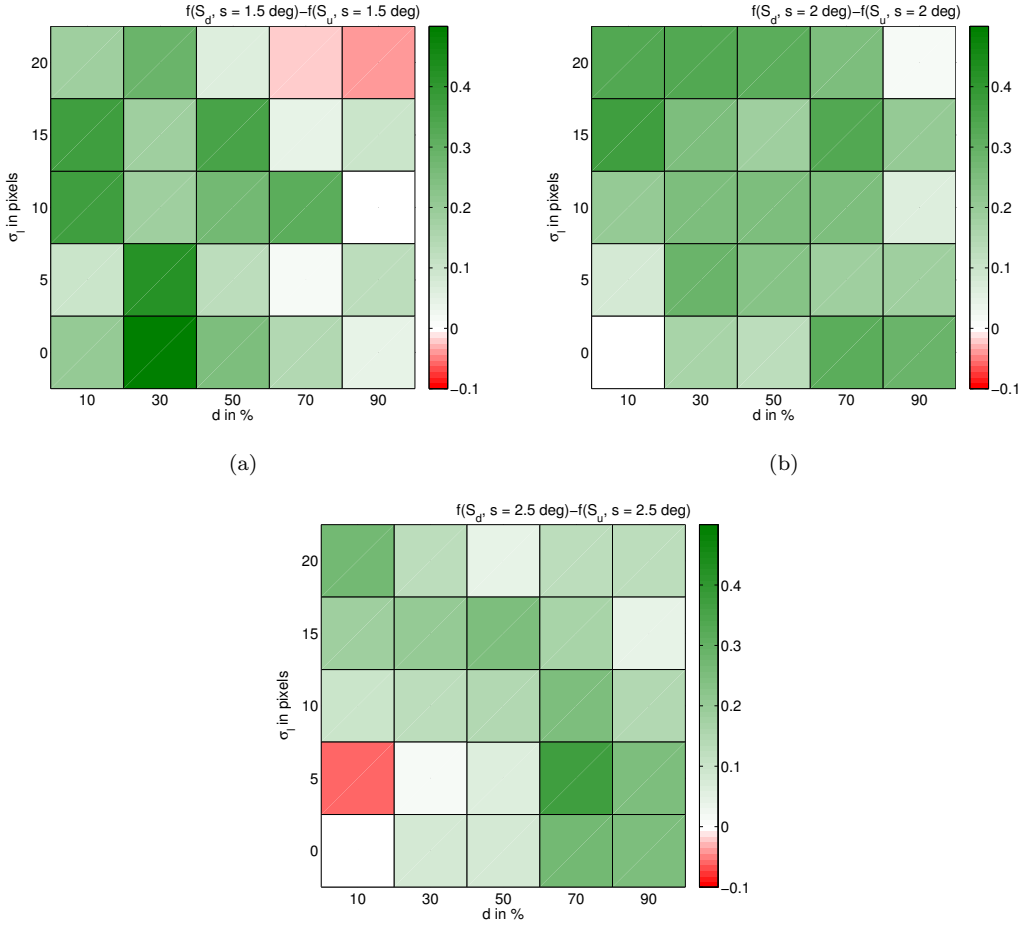


Figure 3: Difference of the fraction of correctly identified letters f between stimulation strategies S_u and S_d as a function of phosphene dropout d and phosphene location shift σ_l for (a) 1.5 degree visual angle, (b) 2 degree visual angle, (c) 2.5 degree visual angle. Lighter shades indicate a smaller difference between performance, red shades indicate that participants performed better with strategy S_u, green shades indicate that participants performed better with strategy S_d.

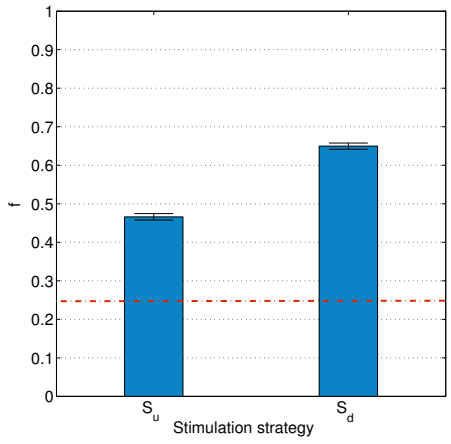


Figure 4: Mean fraction of correctly identified letters using stimulation strategies S_u and S_d. The dashed red line indicates the chance level of 25%. Standard error is displayed by error bars.

Parameter	Value	Standard error	p-value
c_0	0.832	0.124	< 0.001
α_{S_d}	1.414	0.145	< 0.001
α_d	-2.197	0.181	< 0.001
α_l	-1.879	0.175	< 0.001
α_s	2.010	0.175	< 0.001
β_{S_d*d}	-0.689	0.158	< 0.001
β_{S_d*l}	-0.324	0.154	0.035
β_{S_d*s}	0.055	0.132	0.680
β_{d*l}	2.316	0.228	< 0.001
β_{d*s}	-1.250	0.195	< 0.001
β_{l*s}	-0.709	0.189	< 0.001

Table 1: Estimated model parameters for complete model M_c, obtained using a regression analysis as described in Section 2.3.

3.1 Reduced model, M_r

Even though most of the two-way interactions are significant, they are not equally strong. We simplify our model to only include interactions that are significant, but also strong. Because all values are normalized, we refer back to Table 1 to have a closer look at the interaction strength. Values that are closer to zero indicate a less strong interaction between variants. This can be visualized by comparing the relative steepness between the linear fitted curves in Figure 5 that show the relationship between two variants.

Figures 5 (a)-(c) display the relationships between the stimulation strategies and variables phosphene dropout (d), phosphene location shift (l), and stimulus size (s). Figures 5 (d)-(f) display the relationship between the different distortion parameters phosphene dropout (d), phosphene location shift (l), and stimulus size (s). In each figure, the data is averaged over all variables, except than the ones displayed. Error bars indicate the standard error. A first-order polynomial fit is applied to the data. The resulting equations are given in each figure.

The interaction between stimulation strategy, image size and phosphene location shift have already been found to be insignificant. With model values of 0.055 and -0.324 (see Table 1), respectively, they would have also been the least strong interactions. This is followed by the interaction between the stimulation strategy and phosphene dropout with a model value of -0.689 and phosphene location shift and stimulus size with a value of -0.709. The stronger interactions between phosphene dropout, stimulus size and phosphene location shift indicate that these become less important as dropout rate increases. When operating at a dropout rate of 90%, neither stimulation size nor phosphene location shift influence the fraction of correctly identified letters substantially. By contrast, when operating with low dropout, the extent of phosphene location shift and the stimulus size are important for letter identification. In Figure 5 (d) and (e), this is reflected in the difference in steepness of the polynomial fit.

The model introduced earlier includes all two-way interactions. Following our analysis, we can simplify the model and only take interactions into account that, based on our previous analysis, are significant and strong. We dismiss the weaker two-way interactions between the stimulation strategy and the other independent variables and the interaction between phosphene dropout and stimulus size. Fitted model parameters for the resulting reduced model M_r can be found in Table 2.

Finally, Figure 6 shows the recorded data versus model predictions separately for the different models. A perfect model would place all circles on the diagonal. To compare the performance of the models, the root mean square (RMS) values are calculated. They can be found in Table 3. Based on our previous analysis, we know that there are significant interactions between distorting factors. We should, therefore, refrain from modeling the independent variables only individually and instead include at least the most important two-way interactions. Even though the reduced model M_r does not account for all interactions, the RMS is not substantially larger than in the complete model. Thus, it is a good compromise between simplicity and accuracy.

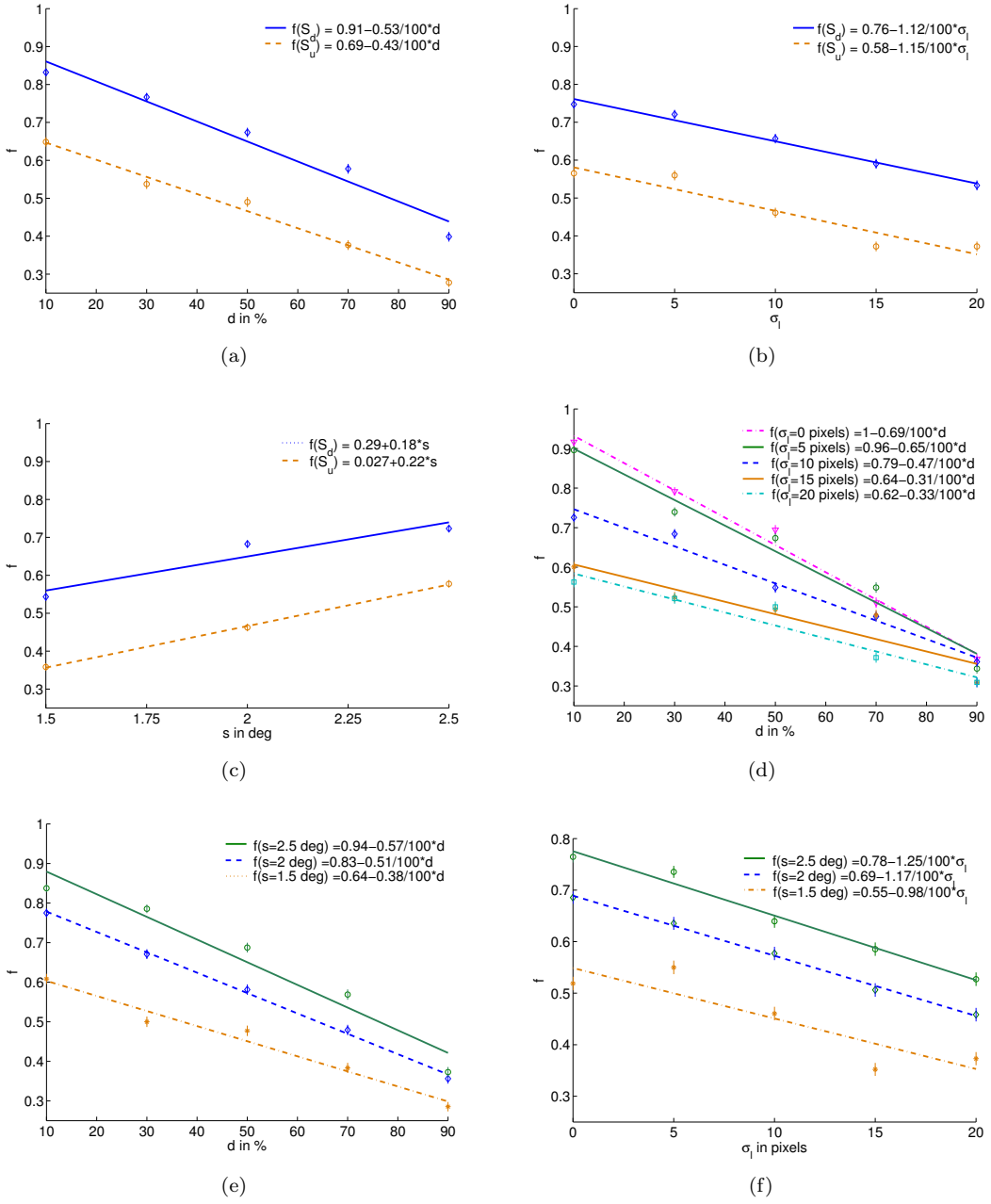
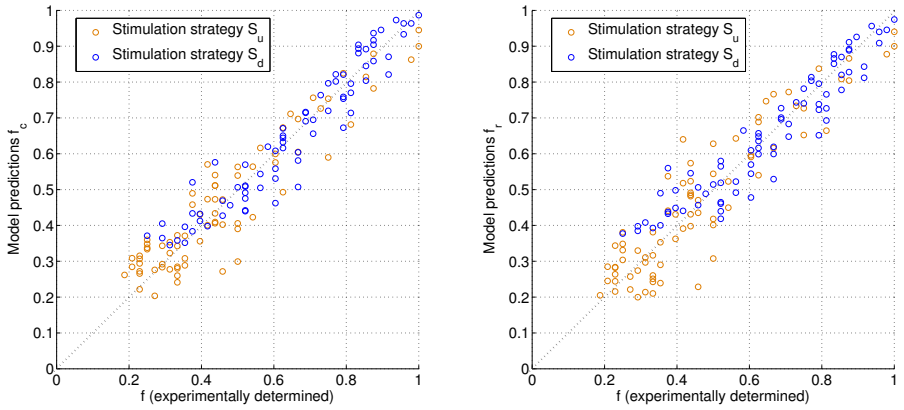


Figure 5: (a)-(c)Two-way interaction between stimulation strategy S_u ($\circ$) and stimulation strategy S_d ($\diamond$) and distortion parameters. Lines indicate the first order polynomial fit for stimulation strategy S_u (dashed orange line) and stimulation strategy S_d (solid blue line). Success as a function of phosphene dropout d (a), phosphene location shift σ_l (b), and stimulus size s (c). (d)-(f) Two-way interaction between different distortion parameters. (d) Success as a function of phosphene dropout d for spatial phosphene location shift values of 0 ($\triangledown$), 5 ($\circ$), 10 ($\diamond$), 15 ($*$), and 20 ($\square$) pixels. First order polynomial fit for phosphene location shift of 0 (dashed-dotted pink line), 5 (solid green line), 10 (dashed blue line), 15 (solid orange line), and 20 (dashed-dotted cyan line). (e) Success as a function of phosphene dropout d for stimulus sizes s that cover 2.5 degree ($\circ$), 2 degree ($\diamond$), and 1.5 degree ($*$) of the visual angle. First order polynomial fit for stimulus sizes that cover 2.5 degree (solid green line), 2 degree (dashed blue line), and 1.5 degree (dashed-dotted orange line) of the visual angle. (f) Success as a function of phosphene location shift σ_l for stimulus sizes s that cover 2.5 degree ($\circ$), 2 degree ($\diamond$), and 1.5 degree ($*$) of the visual angle. First order polynomial fit for stimulus sizes that cover 2.5 degree (solid green line), 2 degree (dashed blue line), and 1.5 degree (dashed-dotted orange line) of the visual angle.



(a) (b)

Figure 6: Evaluation of the models. Circles closer to the diagonal indicate a higher accuracy of the model. (a) Experimentally recorded fraction of correctly identified letters f versus model predictions for model M_c, f_c (complete model). (b) Experimentally recorded percentage of correctly identified letters f versus model predictions for model M_r, f_r (reduced model).

3.1.1 Example application of model

To retrieve the percentage of correctly identified letters, we need to transform the axis back to a linear axis. The percentage of correctly identified letters can be calculated using

$$f_m(\mathbf{q},\mathbf{p}) = 100 \cdot \frac{100e^{(c_0+\boldsymbol{\alpha}\mathbf{q}^T+\boldsymbol{\beta}\mathbf{p}^T)}}{1+e^{(c_0+\boldsymbol{\alpha}\mathbf{q}^T+\boldsymbol{\beta}\mathbf{p}^T)}},\ m \in \{c,r\}. \tag{5}$$

When using this model to estimate parameters, it needs to be noted that values of the vector $\mathbf{q}$ have been normalized to [0, 1]. Phosphene location shift has been tested between [0, 20], phosphene dropout between [10, 90] %, and stimulus size in degrees of the visual angle covered [1.5, 2.5]. Because the stimulation strategy can not be treated as a variant, the value needs to be set to $q_S = 0$ for stimulation strategy S_u and $q_S = 1$ for strategy S_d. We want to know the expected percentage of correctly identified letters for a phosphene dropout of 50% ($q_d = 0.5$), phosphene location shift variance of 10 pixels ($q_l = 0.5$), and a stimulus size of 2.5 degree ($q_s = 1$) for stimulation strategy S_d ($q_S = 1$) in model M_r. Thus, $\mathbf{q} = [1, 0.5, 0.5, 1]$. The vector $\mathbf{p}$ includes all interactions accordingly: $\mathbf{p} = [0.5, 0.5, 1, 0.25, 0.5, 0.5]$. The model parameters that

Parameter	Value	Standard error	p-value
c_0	1.187	0.107	< 0.001
α_{S_d}	0.8870	0.0527	< 0.001
α_d	-2.445	0.164	< 0.001
α_l	-2.231	0.144	< 0.001
α_s	1.567	0.122	< 0.001
β_{d*l}	2.101	0.221	< 0.001
β_{d*s}	-1.106	0.190	< 0.001

Table 2: Estimated model parameters for reduced model M_r, as described in Section 3.1.

Model	RMS
M_c: complete model	0.054
M_r: reduced model	0.060

Table 3: RMS values for the different models. While M_c is the most accurate model, it is also the most complex. M_r is a good compromise between accuracy and simplicity.

we insert into the equation can be found in Table 2: $c_0 = 1.187$, $\boldsymbol{\alpha} = [0.887, -2.445, -2.231, 1.567]$, and $\boldsymbol{\beta} = [0, 0, 0, 2.101, -1.106, 0]$. Using equation (5), we obtain a percentage correct of $f = 78.2\%$. If we use stimulation strategy S_u, $\mathbf{q} = [0, 0.5, 0.5, 1]$ and the percentage of correctly identified letters is $f_r = 59.6\%$.

4 Discussion

Traditional stimulation strategies (stimulation strategy S_u) for retinal prostheses use different intensity levels of phosphenes to convey information to the user. We have shown that including additional directional information may be beneficial. Previously, Savage et al. (2012) showed, using a finite element model, that it may be possible to stimulate in a way that deliberately changes the shape of a phosphene. Even if not changed deliberately, phosphenes often exhibit a net directional component, which could be exploited to transport additional information to the implantee.

Oriented phosphenes are simulated by displaying elliptical instead of uniformly round phosphenes. In a visual psychophysics experiment with simulated phosphene vision, we were able to show a significant improvement from 47% to 65% correctly identified letters in a letter recognition task. We have shown that the proposed stimulation strategy may be useful when dealing with low acuity, irregular spacing of phosphenes, or a large percentage of non-stimulating electrodes, all of which are known issues with current retinal prostheses.

While simulation does not entirely capture the perceptual experience of retinal implant recipients, the simulation methodology adopted provides an effective means of examining potential differences between proposed stimulation strategies under controlled conditions. The idealized phosphene models used have been balanced by the inclusion of known variables such as non-regular phosphene layouts, and phosphene dropout. While the results cannot precisely predict implanted patient performance on the same task, the results presented suggest there is some advantage to be gained by encoding directional information in the elicited phosphenes.

Study participants performed better in a letter recognition task when presented with directional information than without, even for smaller letter sizes. Figure 5(c) shows that we could decrease the stimulus size by as much as 63% while retaining the same percentage of correctly identified letters. A strategy that directly encodes edge orientation information may enable patients to read smaller print than otherwise possible and increase reading speed by showing more than one letter at a time to the implantee.

In order to explore the entire range of possible distortions, our simulations include distortion to an extent that is likely to be much higher than in an actual implant. Phosphene location shift is unlikely to persist to an extent that has been simulated here, because an initial mapping of the phosphene location and the input image can be conducted. A similar observation is true for phosphene dropout. As the main reason for phosphene dropout is defective electrodes or contacts, it is unlikely to occur as randomly as simulated in the presented experiment or to the extent we simulated. Additionally, the outcome of this stimulation strategy will depend strongly on the electrode array configuration and stimulation approach. Using a low-resolution grid and a stimulation method that relies on adjacent electrodes being stimulating and return electrodes, as described by Savage et al. (2012), a large dropout due to defective electrodes could impact the benefit of this stimulation strategy. However, for an uneven distribution of the defective electrodes, the proposed stimulation strategy can still be used in parts of the array, using the naive strategy where elliptical phosphenes cannot be elicited. The percentage of correctly identified letters for this hybrid strategy can be assumed to be somewhere in between that of the conventional and the directional strategy. With a growing resolution and density of electrode arrays, we believe it is likely that defective electrodes could be bridged.

Knowing that the same number of presented phosphenes can result in significantly better reading performance, a stimulation strategy may be employed that uses this information to reduce the power consumption of the implant, ensure longer battery life and, thus, give more comfort for the user. Strategies that use psychophysical data in order to limit the power consumption of an implant have successfully been implemented and tested in cochlear implants, as demonstrated by Nogueira et al. (2005) and Büchner et al. (2008). The model presented here can assist in designing a similar strategy for vision implants.

In this work, we focus on , low-resolution devices similar to those currently being used and tested with patients. It is to be expected that future devices will have a higher number of individually controllable electrodes. The derived model makes it possible to extrapolate the effectiveness of the proposed stimulation

strategy for higher resolution devices. It becomes clear that for implants that can elicit many individual phosphenes, the benefit is likely to be reduced. An example of an existing implant with 1500 separate electrodes has been developed by Zrenner et al. (2011). However, the electrodes of this device are driven by photodiodes and, thus, are not individually controllable. Psychophysical data for this device shows that even with 1500 photodiode-controlled electrodes the visual acuity (20/1000) was similar to that of a device with 60 independent stimulating electrodes (20/1260 (Humayun et al., 2012)). Moreover, a device with a large number of individually controllable electrodes could potentially enable more controlled steering of current, which would make this strategy easier to implement, since defective electrodes could potentially be bridged due to the closeness of surrounding electrodes.

Using letters, we tested the stimulation strategy with a very basic stimulus. With growing numbers of electrodes and improved visual acuity, examining other stimuli will become increasingly interesting. In order to benefit from the proposed stimulation strategy, it is likely that more complex stimuli will require pre-processing such as contrast enhancement of edges (Al-Atabany et al., 2010; Kwon et al., 2012) to facilitate edge detection. We note that the proposed stimulation method is not limited to edge orientations in the scene, and may also be combined with more sophisticated image processing methods to encode additional information relating to other scene properties such as surface orientation, or motion direction.

Another consideration is that the effects of simultaneous stimulation of electrodes have not yet been studied extensively. There may well be effects, such as cross-talk between electrodes, that force us to only stimulate a certain number of electrodes at the same time. The stimulation strategy we present has potential to turn electrode cross-talk, as presented by Savage et al. (2012), into useful information for the user by harnessing directionality of phosphenes.

While the ability to deliberately change phosphene directionality is yet to be shown clinically, psychophysical experiments with retinal implant patients have shown that stimulation of electrodes leads to a variety of different shapes (Horsager and Fine, 2011; Nanduri et al., 2008). The steadily growing number of electrodes may make a deliberate manipulation of shape more likely in future implants. The data presented in this paper suggests that directional information itself can be beneficial. Thus, it is possible that we can use the naturally occurring phosphene shapes with diverse manifestations, such as the reported doughnuts, arches, and wedges (Horsager and Fine, 2011; Nanduri et al., 2008), to convey at least some directional information to the user. A first attempt to do this in simulations has been presented by Kiral-Kornek et al. (2013).

5 Conclusion

Recent research has shown that electrical stimulation of the retina can lead to the perception of phosphenes. Those phosphenes have been recorded in patients with retinal implants and it has been shown that they can have complex shapes. In this paper, we presented a new approach to a stimulation strategy for retinal implants that makes use of differently shaped phosphenes. More specifically, the new strategy makes use of phosphene direction to convey additional information to the user. We compared this strategy to a more traditional method that only uses the intensity of phosphenes to encode information. In a psychophysical experiment, we were able to show that under controlled conditions with simulated phosphene vision the ability to recognize letters improved significantly from 47% with standard intensity-based encoding to 65% correctly identified letters with the proposed stimulation method. By varying the rate of phosphene dropout, phosphene location shift, and letter size, we have demonstrated that participants performed better under all circumstances using the stimulation strategy that displays directional phosphenes. Thus, we conclude that the direct encoding of directional cues may be of benefit for letter and symbol recognition with a retinal prosthesis.

We presented two models of different accuracy and complexity that can be used to anticipate the effects of phosphene dropout, phosphene location shift, and stimulus size for both of the presented stimulation strategies. While our model is only valid for static images and based on simulations, the results are promising and encourage further research into the area of oriented phosphenes. Future work will need to include testing the strategies under more realistic conditions, and in the context of activities of daily living. Encoding directional information may improve general shape discrimination and even mobility and orientation.

6 Acknowledgements

This research was supported by the Australian Research Council (ARC) through its Special Research Initiative (SRI) in Bionic Vision Australia (BVA). The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program. This work was supported by the Australian Federal and Victorian State Governments and the Australian Research Council through the ICT Centre of Excellence program, National ICT Australia (NICTA). Statistical support was provided by the Statistical Consulting Centre of the University of Melbourne. We thank all participants for volunteering their time.

References

M. Abramian, S. Dokos, J. W. Morley, and N. H. Lovell. Activation of ganglion cell axons following epiretinal electrical stimulation with hexagonal electrodes. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 6753–6756. IEEE, 2010.

M. Abramian, N. H. Lovell, J. W. Morley, G. J. Suaning, and S. Dokos. Activation of retinal ganglion cells following epiretinal electrical stimulation with hexagonally arranged bipolar electrodes. *Journal of Neural Engineering*, 8(3):035004, 2011.

W. I. Al-Atabany, M. A. Memon, S. M. Downes, and P. A. Degenaar. Designing and testing scene enhancement. *BioMedical Engineering OnLine*, 9(27), 2010.

M. Bach. The freiburg visual acuity test-automatic measurement of visual acuity. *Optometry & Vision Science*, 73(1):49–53, 1996.

G. S. Brindley and W. S. Lewin. The sensations produced by electrical stimulation of the visual cortex. *Journal of Physiology (Lond)*, 196:479 – 493, 1968.

British Standards Institute. BS 4274-1: Visual acuity test types. test charts for clinical determination of distance visual acuity. specification. British Standards, 2003.

A. Büchner, W. Nogueira, B. Edler, R.-D. Battmer, and T. Lenarz. Results from a psychoacoustic model-based strategy for the nucleus-24 and freedom cochlear implants. *Otology & Neurotology*, 29(2):189–192, 2008.

S. C. Chen, G. J. Suaning, J. W. Morley, and N. H. Lovell. Simulating prosthetic vision: I. visual models of phosphenes. *Vision Research*, 49(12):1493 – 506, 2009.

R. O. Duda, P. E. Hart, et al. *Pattern classification and scene analysis*, volume 3. Wiley New York, 1973.

R. Efron. Effect of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics*, 8(4):231–234, 1970.

O. Foerster. Beitraege zur pathophysiologie der sehbahn und der sehsphaere. *J. Psychol. Neurol., Lpz.*, 39: 463–485, 1929.

A. Horsager and I. Fine. The perceptual effects of chronic retinal stimulation. *Visual Prosthetics: Physiology, Bioengineering, Rehabilitation. New York, NY: Springer*, pages 271–300, 2011.

A. Horsager, R. J. Greenberg, and I. Fine. Spatiotemporal interactions in retinal prosthesis subjects. *Investigative Ophthalmology & Visual Science*, 51(2):1223–1233, 2010.

M. Humayun, J. Weiland, G. Fujii, R. Greenberg, R. Williamson, J. Little, B. Mech, V. Cimmarusti, G. Van Boemel, G. Dagnelie, et al. Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision research*, 43(24):2573–2581, 2003a.

M. S. Humayun, E. de Juan Jr., J. D. Weiland, G. Dagnelie, S. Katona, R. Greenberg, and S. Suzuki. Pattern electrical stimulation of the human retina. *Vision Research*, 39(15):2569 – 2576, 1999.

M. S. Humayun, J. D. Weiland, G. Y. Fujii, R. Greenberg, R. Williamson, J. Little, B. Mech, V. Cimmarusti, G. V. Boemel, G. Dagnelie, and E. de Juan Jr. Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Research*, 43(24):2573 – 2581, 2003b.

M. S. Humayun, J. D. Dorn, L. da Cruz, G. Dagnelie, J.-A. Sahel, P. E. Stanga, A. V. Cideciyan, J. L. Duncan, D. Eliott, E. Filley, et al. Interim results from the international trial of second sight's visual prosthesis. *Ophthalmology*, 119(4):779–788, 2012.

T. F. Jaeger. Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4):434–446, 2008.

F. Kiral-Kornek, C. Savage, E. O'Sullivan-Greene, A. Burkitt, and D. Grayden. Embracing the irregular: A patient-specific image processing strategy for visual prostheses. In *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE*, pages 3563–3566. IEEE, 2013.

M. Kwon, C. Ramachandra, P. Satgunam, B. W. Mel, E. Peli, and B. S. Tjan. Contour enhancement benefits older adults with simulated central field loss. *Optometry and vision science: official publication of the American Academy of Optometry*, 89(9):1374, 2012.

D. Nanduri, M. Humayun, R. Greenberg, M. McMahon, and J. Weiland. Retinal prosthesis phosphene shape analysis. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 1785–1788. IEEE, 2008.

W. Nogueira, A. Büchner, T. Lenarz, and B. Edler. A psychoacoustic "nofm"-type-speech coding strategy for cochlear implants. *EURASIP Journal on Aplied Signal Processing*, 18:3044–3059, 2005.

N. Opie, N. Lovell, G. Suaning, P. Preston, and S. Dokos. Current steering for high resolution retinal implants. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 2760–2763. IEEE, 2013.

C. O. Savage, F. I. Kiral-Kornek, and B. Tahayori. Can electric cross-talk be used to control perception of a retinal prosthesis patient? In *Proceedings on 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'12)*, pages 3013–3016, 2012.

H. Snellen. *Dr. H. Snellen's Probebuchstaben zur Bestimmung der Sehschaerfe*. H. Peters, 1863.

G. J. Suaning, N. H. Lovell, K. Schindhelm, and M. T. Coroneo. The bionic eye (electronic visual prosthesis): A review. *Australian and New Zealand Journal of Ophthalmology*, 26(3):195–202, 1998.

F. M. Urban. The method of constant stimuli and its generalizations. *Psychological Review*, 17(4):229, 1910.

A. C. Weitz, M. R. Behrend, N. S. Lee, R. L. Klein, V. A. Chiodo, W. W. Hauswirth, M. S. Humayun, J. D. Weiland, and R. H. Chow. Imaging the response of the retina to electrical stimulation with genetically encoded calcium indicators. *Journal of neurophysiology*, 109(7):1979–1988, 2013.

R. Wilke, V. Gabel, H. Sachs, K. Schmidt, F. Gekeler, D. Besch, P. Szurman, A. Stett, B. Wilhelm, T. Peters, et al. Spatial resolution and perception of patterns mediated by a subretinal 16-electrode array in patients blinded by hereditary retinal dystrophies. *Investigative Ophthalmology & Visual Science*, 52(8):5995–6003, 2011a.

R. Wilke, G. Moghadam, N. Lovell, G. Suaning, and S. Dokos. Electric crosstalk impairs spatial resolution of multi-electrode arrays in retinal implants. *Journal for Neural Engineering*, 8(4), June 2011b.

J. Wyatt and J. Rizzo. Ocular implants for the blind. *IEEE Spectrum*, 33(5):47 – 53, May 1996.

Y. Zhao, Y. Lu, C. Zhou, Y. Chen, Q. Ren, and X. Chai. Chinese character recognition using simulated phosphene maps. *Investigative Ophthalmology & Visual Science*, 52(6):3404–3412, 2011.

E. Zrenner, K. Bartz-Schmidt, H. Benav, D. Besch, A. Bruckmann, V. Gabel, F. Gekeler, U. Greppmaier, A. Harscher, S. Kibbel, et al. Subretinal electronic chips allow blind patients to read letters and combine them to words. *Proceedings of the Royal Society B: Biological Sciences*, 278(1711):1489–1497, 2011.