

Addressing the Perceptual Limitations of Cortical Implants for Restoring Vision

Student: Jamin Wu

Student ID: 27025861

Supervisor: Dr Yan Tat Wong

Co-Supervisor: Dr Nicholas Price

Department of Physiology

Department of Electrical & Computer Systems Engineering

School of Biomedical Sciences, Monash University

Word Count: 7149 words

Contents

1	Introduction	7
2	What is a cortical visual prosthesis?	9
2.1	Overview of Cortical Visual Prostheses	9
2.2	Mechanism of Cortical Visual Prostheses	11
3	What do patients with cortical implants see?	12
3.1	Phosphenes Produced by Stimulating a Single Electrode	16
3.2	Phosphenes Produced by Stimulating Multiple Electrodes	20
3.3	Summary of the Perceptual Limitations of Phosphenes	21
4	How can we make what people see most useful?	23
4.1	The Role of Simulated Prosthetic Vision	23
4.2	A Brief Outline of Current Image Processing Approaches	25
4.3	The Limitations of Simulated Prosthetic Vision	30
4.4	Future Directions from Advances in Machine Learning	32
5	Conclusion	34
6	Aims & Hypotheses	35
7	Project Outline	36
7.1	Rationale	36
7.2	Brief Outline of Proposed Methods	36

List of Tables

1 Current progress of cortical visual prostheses. The number of electrodes
 places a hard upper bound on the resolution (and visual acuity) these
 devices can provide, so are noted here. 11

2 The participant demographics of studies which have looked at cortical
 phosphenes evoked *in vivo* in humans. 15

List of Figures

1	A modelled render of the Gennaris bionic vision headgear. Figure courtesy of Monash Vision Group.	10
2	A basic flowchart of the process of prosthetic vision compared to normal vision. Figure made by the author.	13
3	Spatial distribution of phosphenes mapped to the visual fields (left) in a patient implanted with an early 64-electrode array (right). Figure from Dobelle et al. 1979 [50]	17
4	Examples of different simulated renders of phosphenes. Figure from Chen et al. 2009. [64]	24
5	An example of an edge-detection algorithm with a CNN-driven object-filling algorithm to render a clean simulated phosphene image. Figure from [70]	26
6	An example of a saliency-based algorithm to highlight a region of interest of the image. Figure from [67]	27
7	An example of a transformative approach to image processing to re-encode information. Figure from [69]	28
8	An example of a temporal approach to conveying information through phosphenes, where electrodes are stimulated in succession instead of simultaneously. Figure taken from [56]	30
9	Basic overview of the processor and renderer combinations to be tested in this experiment. Figure made by the author.	38

List of Abbreviations

CNN Convolutional neural network

CPU Computer processing unit

ICVP Intracortical Visual Prosthesis

Declaration

The original concept for this project was devised by Dr Yan Tat Wong. Dr Yan Tat Wong, Dr Nicholas Price, and I (Jamin Wu) discussed and developed the research aims, hypotheses and methods for this project.

I conducted and wrote this literature review. Dr Yan Tat Wong, Dr Nicholas Price and I discussed the contents of this literature review. Dr Yan Tat Wong and Dr Nicholas Price provided feedback on concepts and drafts of this literature review. I would like to thank both of my supervisors for their continuous support.

I declare that the contents of this literature review are my own work unless explicit reference indicating otherwise has been made in the text. I declare that this work has not been submitted for the award of any other degrees or diplomas.

1 Introduction

Around 36 million people worldwide were blind in 2015. [1, 2] Many patients are left permanently blind when there is irreversible damage to the visual system such as from glaucoma or trauma. [3, 4] However these patients often still have portions of the visual system intact, which could be leveraged to restore vision artificially. **Cortical visual prostheses** are devices which aim to achieve exactly that - devices implanted on the brain which stimulate neurons to directly inject visual sensations into awareness. [5–7]

Studies in the late 20th century showed that electrically stimulating the brain could indeed produce spot-like sensations of light, known as **phosphenes**. [8–11] But these studies also revealed that we can exert only very limited control over what phosphenes look like, where they are located, and how they interact. [9, 12, 13] Of these studies, those conducted in small groups also demonstrated wide variation in the appearance of phosphenes between people. [9, 10] Therefore, whilst we now know we can produce these spots of light by stimulating the brain, it is not clear how to assemble these unyielding percepts into useful vision for all implantees. [14, 15]

To address this, new proposals of cortical visual prostheses are exploring the use of sophisticated computer processing to assemble phosphenes *strategically* so implantees can interpret them in useful ways. [7, 16] These methods generally focus on addressing two significant difficulties of prosthetic vision: its low resolution, and its lack of color and brightness levels, both of which are easy to simulate on graphical displays and test. [17–19] However it is unclear how well these image processing algorithms could cope with the much wider range of phosphene experiences reported by perceptual studies in humans. Thus, the primary issue facing these algorithms is how they can be made flexible to accommodate the large, uncontrollable variation in what phosphenes look like.

Recently, there have been remarkable advances in the ability of computers to *derive* image processing algorithms rather than having algorithms being programmed by hand. In this paradigm, computers iteratively "learn" and integrate patterns between

data inputs and outputs through a process known as **machine learning**. [20] By using machine learning methods on top of layered image processing architectures called **convolutional neural networks** (CNNs) , computers are able to achieve near-human performance in image processing feats such as handwritten digit recognition [21], segmenting images [22], and replicating image styles. [23] Machine learning principles could be similarly applied to prosthetic vision, using inferred patterns to derive image processing algorithms from performance rather than building algorithms around conceptions of what phosphenes look like. This could improve the flexibility of cortical implants for different perceptual experiences and improve the usefulness of cortical visual prostheses for future implantees.

In this review, we first briefly describe the relevant background of cortical visual prostheses in Section 2. We then discuss the empirical studies on the features and limitations of phosphenes produced by cortical implants in Section 3. Finally, we describe computational approaches which address these perceptual limitations and their implications for future research in Section 4.

2 What is a cortical visual prosthesis?

2.1 Overview of Cortical Visual Prostheses

A **prosthesis** is an artificial, implanted device which aims to restore a lost function to the human body. [24] A **visual prosthesis** is such a device which aims to restore vision to people with visual impairment. [25, 26] A **cortical visual prosthesis** specifically refers to a vision restoration device implanted on the cortical surface of the brain, as opposed to other portions of the visual system such as the eye, optic nerve or thalamus. [6] As a cortical prosthesis is implanted directly on the brain, implantees only require a functional cortical portion of the visual system; the device bypasses any damage to the eyes or nerves leading up to the brain. While visual prostheses at other locations in the visual pathway are also capable of producing visual sensations [27–30], their use-case, implementation and evoked vision differ from those produced by cortical prostheses and are outside the scope of this review.

Research into stimulating the brain to produce vision was first pioneered by Brindley & Lewin [8] and later Dobelle & Mladejovsky [9] in the late 20th century. Using rudimentary hardware, these early experiments showed that a temporary implant composed of a array of electrodes could stimulate the brain in an awake patient and make them see artificial sensations of light. [8, 9] These artificial sensations of light are known as **phosphenes**. Phosphenes are highly variable, but most often appear as dots of light likened to "a star in the sky". [9] These early successes in evoking phosphenes were instrumental in demonstrating the feasibility of cortical prostheses [13], which have adopted phosphenes as the fundamental building blocks of prosthetic vision.

In the past 50 years since, a small number of research groups have proposed modern cortical prostheses based on the principles of these early results. These include the Gennaris bionic vision system [31, 32], the Intracortical Visual Prosthesis (ICVP) Project [33], CORTIVIS [34] and the Orion Visual Cortical Prosthesis. [35] To illustrate what modern conceptions of a visual cortical prosthesis may look like, Figure 1 shows a simulated render of the headgear for the Gennaris bionic vision system.



Figure 1: A modelled render of the Gennaris bionic vision headgear. Figure courtesy of Monash Vision Group.

These devices, while still early in development, may eventually be an option for restoring a crude form of vision to patients who would otherwise be left permanently blind. The expectation is that these devices could provide gross light perception which might allow the recognition of basic forms and movement. [31, 32] Current technology cannot reproduce anything close to the trichromatic, approximately 15 million pixel resolution of the human eye [36], and as such, these devices are not yet a full replacement for vision. The development of cortical prostheses has been reviewed previously [37]; we briefly summarise the current progress of these devices in Table 1.

Device	Electrodes	Progress	References
Orion (Second Sight)	60 subdural surface electrodes	FDA-approved clinical trial ongoing 2018-2023, six patients implanted (public scientific results yet to be released).	[35, 37, 38]
Gennaris (Monash Vision Group)	Up to 473 penetrating micro-electrodes	Ethics approved to begin clinical trials, recruiting.	[31, 32, 39]
ICVP Project (Illinois Institute of Technology)	16 penetrating microelectrodes	Preclinical phase	[33]
CORTIVIS (Universidad Miguel Hernández)	100 penetrating microelectrodes	Preclinical phase.	[34]

Table 1: Current progress of cortical visual prostheses. The number of electrodes places a hard upper bound on the resolution (and visual acuity) these devices can provide, so are noted here.

2.2 Mechanism of Cortical Visual Prostheses

While the specific hardware of each device differs, the fundamental mechanism of these devices is similar. Patients must first undergo an operation to surgically implant an electrode array in the primary visual cortex at the back of the brain. [6] Historically, these were subdural surface electrodes sitting atop (but not penetrating) the brain. [8, 9] However, modern prostheses tend to opt for penetrating microelectrodes [32–34]

which are finer and can operate successfully at lower electrical current. [10, 13]

After implantation, the physical components of the system are in place. The system will need to be calibrated and tested before use to determine electrical stimulation thresholds and the spatial correspondence between electrodes and phosphenes in the visual fields. [31, 34]

When in use, an external camera (e.g. on glasses worn by the user) first captures an image. [6, 32] This image is transmitted to a portable processor, and is converted into electrical parameters for each electrode in the implanted array. Each electrode in the implanted array then delivers pulses of electrical charge into the brain based on its parameters, which electrically stimulates nearby neurons in the cortical tissue. Stimulating neurons in the visual cortex produces patterns of phosphenes which the patient can then perceive and interpret. [8–10] While these patterns may be difficult to decipher at first, it is expected that patients will eventually learn to match phosphene patterns to useful information such as letterforms. [34]

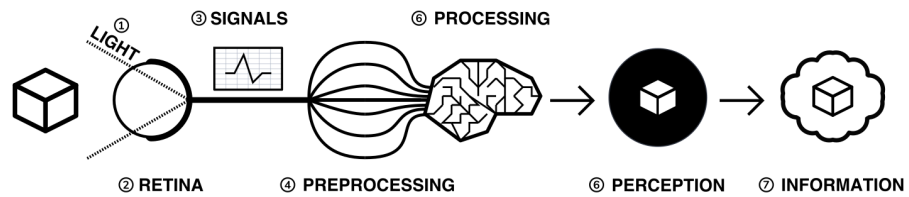
This process from camera image to neural stimulation loops continuously to produce a stream of images like frames of a video. Essentially, this system provides an artificial real-time link between environmental light and visual information; a link ordinarily present in natural vision, but not present in blindness. Figure 2 depicts the basic process of prosthetic vision in comparison to normal vision.

3 What do patients with cortical implants see?

Because cortical visual prostheses use phosphenes as the fundamental building block of prosthetic vision, it is imperative that we be able to compose multiple phosphenes into meaningful imagery. Whether phosphenes can be composed meaningfully depends on the visual and perceptual properties of phosphenes such as their size, color and interaction with other phosphenes.

Unfortunately, due to the technical and ethical issues surrounding stimulating peoples' brains, the number of studies characterising these properties of phosphenes

A) NORMAL VISION



B) PROSTHETIC VISION

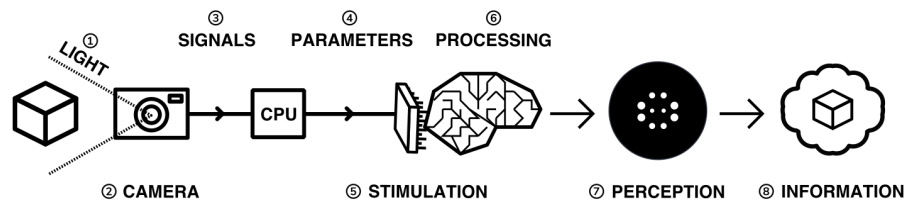


Figure 2: A basic flowchart of the process of prosthetic vision compared to normal vision.
Figure made by the author.

in humans is understandably small. Studies which characterise phosphenes evoked *in vivo* typically fall into two distinct groups:

1. Historical experiments with rudimentary hardware on noble volunteers [8, 10, 40],
or
2. Modern but conservative experiments in epilepsy patients who already have electrodes implanted for clinical monitoring. [11, 41–44]

As the pool of phosphene studies in humans is small and the demographics of these studies are skewed towards specific populations, we summarise the pertinent methodological features of each study in Table 2. These studies constitute the major perceptual evidence that a cortical visual prosthesis can produce vision, and point towards what type of vision might be possible.

There are several studies which also attempt to characterise phosphenes in non-human primates from trained behavioural responses. [45–48] While the qualitative perceptual information offered by these studies is limited, they provide some additional information about the spatial properties of phosphenes inferred from saccades (rapid eye movements).

Date	Reference	Setting	Electrodes	Parameters	Patient Demographic
1968	[8]	Acute	80 surface electrodes (array) on occipital cortex	Monophasic trains with 0.2ms pulses of unknown current (power 90mW) at 100Hz	1 patient blind from glaucoma and retinal detachment approx 1 year prior (female, 52 years)
1974	[9]	Acute	Variable number of surface electrodes on occipital cortex	Monophasic or biphasic trains with 0.25-2ms/phase pulses of up to 1-5mA at 30-200Hz for 1000ms	15 patients with cerebral tumours and partial visual field defects or normal sight (11 male, 4 female, 20-71 years)
1974	[40]	Acute	64 subdural surface electrodes (array) on occipital cortex	Biphasic trains with 0.5ms/phase pulses of up to 8mA at 50Hz for unknown duration	1 patient blind from congenital cataract in one eye and glaucoma and retinal detachment in the other for 28 years (male, 45 years); and 1 patient blind from trauma for 7 years (male, 28 years)
1978	[12]	Chronic {5½ years post implant}	Unknown	Unknown-phase trains with up to 2ms pulses of ?mA at 2-1000Hz for 1-8 pulses	Unknown
1976, 1979	[49, 50]	Chronic {unknown years post implant}	64 subdural surface electrodes (array) on occipital cortex	Biphasic trains with 0.25ms/phase pulses of 0.5-4.0mA at 50Hz for 500-1000ms	1 patient blind from trauma 10 years prior to implantation (male, 33 and 35 years)
1990	[10]	Acute	1-3 intracortical penetrating microelectrodes on occipital cortex	Biphasic trains with 0.2ms/phase pulses of up to 200µA at 100Hz for 100-1000ms	3 sighted patients with epilepsy (unknown demographic?)
1996	[13]	Acute	38 intracortical penetrating microelectrodes on occipital cortex	Biphasic trains with 0.2-0.8ms pulses of up to 80µA at 75-200Hz for 125-250ms	1 patient blind from glaucoma 22 years prior (female, 42 years)
1994, 1999	[51, 52]	Acute	Unknown number of surface electrodes on extrastriate visual cortex	Biphasic trains with 0.2ms pulses of 2-10mA at 50Hz for 5000ms	Unknown
2000	[53]	Chronic {21 years post implant}	64 (subdural?) surface electrodes (array) on on occipital cortex	Biphasic trains with 0.5ms/phase of (10-20V) at 30Hz for 1-50 pulses	1 patient blind from trauma 5 years prior to implantation (male, 62 years)
2000	[41]	Acute	Total 271 subdural surface electrodes on occipital cortex and adjacent areas across all subjects	Biphasic trains with 0.3ms pulses of 1-15mA at 50Hz for 5000ms	23 sighted patients with epilepsy (12 male, 11 female, 16-41 years)
2009	[43]	Acute	Total 50 subdural surface electrodes on 11 different visual areas across all subjects	Biphasic trains with 0.2ms pulses of 0.49-7mA at 200Hz for 300ms	10 sighted patients with epilepsy (6 male, 4 female, 19-67 years)
2016	[42]	Acute	1 or 2 subdural surface electrodes on V1 studied per subject	Biphasic trains with 0.2-1ms pulses of 0.2-5mA at 5-100Hz for 200-1000ms	4 sighted patients with epilepsy (3 male, 1 female, 24-40 years)
2017	[54]	Acute	Up to 16 subdural surface electrodes (array) on early occipital cortex per subject	Biphasic trains with 0.1ms/phase pulses of 0.3-4.0mA at 200Hz for 200-300ms	15 sighted patients with epilepsy (5 male, 10 female, 22-61 years)
2018	[55]	Acute	Up to 16 subdural surface electrodes (array) on early occipital cortex per subject	Biphasic trains with 0.1ms/phase pulses of 0.3-4.0mA at 200Hz for 200-300ms	8 sighted patients with epilepsy
2018	[56]	Acute	16 or 24 subdural surface electrodes (array) on early occipital cortex per subject	Biphasic trains with 0.1ms/phase pulses of 0.3-4.0mA at 200Hz for 50-300ms	4 sighted patients with epilepsy (all male, 20-54 years) and 1 patient blind 8 years prior, unspecified reason (female, 35 years)
2019	[44]	Acute	16 subdural surface electrodes (array) on occipital cortex	Biphasic trains with 1ms pulses of up to 11mA at 60Hz	1 patient with epilepsy and a partial visual field defect for 30 years from AVM haemorrhage (male, 45 years)

Table 2: The participant demographics of studies which have looked at cortical phosphenes evoked *in vivo* in humans.

3.1 Phosphenes Produced by Stimulating a Single Electrode

The most common result of stimulating a single electrode is a single phosphene characterised as a small, localisable dot of light likened to a star. [8, 9, 13, 41] Every study has, however, demonstrated large variation on this basic percept.

3.1.1 The Quality of a Phosphene

Being able to perceive qualities of light such as brightness and colour gives us richer, more specific information about the world. [57, 58] Composing images with phosphenes of different brightness and colour would enable us to mimic the richness of natural visual information.

It therefore seems promising that one of the most consistently reported features of phosphenes is that different levels of brightness *are* perceivable and even modifiable. The brightness of a phosphene reproducibly increases with stimulation amplitude, pulse duration and pulse frequency. [9, 12, 13, 40, 42, 53] An early study estimated up to 12 distinguishable levels of brightness by varying the stimulation amplitude of a surface electrode. [12]

The colour of phosphenes, however, is not as promising. Phosphene colours range from colourless to vividly coloured with large inter-individual variation. Some patients only report seeing white or colourless phosphenes. [8–10, 53] Others have reported a spectrum across almost every reportable colour and beyond to 'other-wordly' colours. [9, 10, 12, 13, 41, 43, 52] Sighted patients looking at a white background have also reported seeing black phosphenes, though this finding is not well reported elsewhere. [41] While it was previously speculated that blind patients saw colourless phosphenes due to long-term sight deprivation [9], this is not consistently the case, and coloured percepts have also been reported by a patient blind for 22 years. [13] Least promising is that colour is not consistently modifiable using different parameters of electrical stimulation, meaning phosphenes are most often randomly coloured. [12]

3.1.2 Spatial Properties of Phosphenes

Of great concern to cortical visual prostheses is how phosphenes are arranged in visual space, which may affect the shapes of patterns that can be formed by prosthetic devices. The visual cortex, as a sensory surface, is mapped retinotopically i.e. such that regions in the visual field which are next to each other are also next to each other on the cortex (though they may be distorted). [59, 60] Electrodes placed over visual cortex appear to follow this mapping, and relationships between adjacent electrodes are roughly conserved. [8, 50, 56] Figure 3 illustrates the mapping of a 64-electrode array to phosphene locations measured by perceptual testing.

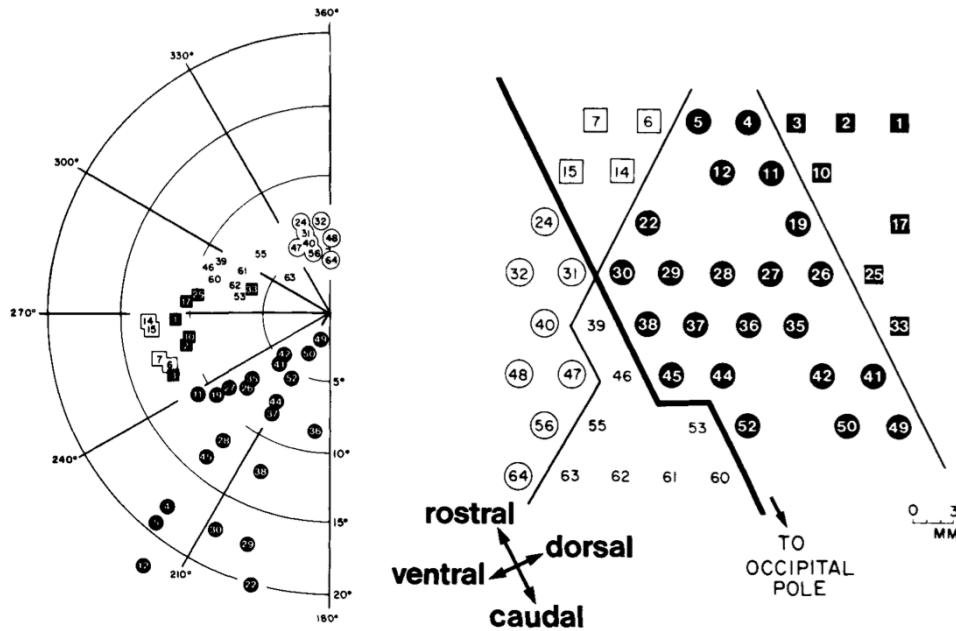


Figure 3: Spatial distribution of phosphenes mapped to the visual fields (left) in a patient implanted with an early 64-electrode array (right). Figure from Dobelle et al. 1979 [50]

However, while we can very grossly estimate positions of phosphenes in the visual field (especially in relation to the calcarine sulcus, below which phosphenes correspond to superior fields), the distortion of retinotopy on the visual cortex means *precise* mapping is not possible until post-implantation. In sighted patients, phosphene locations can be mapped with receptive fields in response to visual stimuli, to which they closely

correspond. [54–56] This is clearly not possible in blind patients, so phosphenes are often mapped by indicating directions or relative positions of pairwise phosphenes. [8, 13, 50, 56] The implication is that while we can roughly determine the quadrant of a phosphene in the visual field at implantation, we cannot know precisely where it is located until stimulation is trialled. In addition, sometimes stimulating one electrode produces more than one phosphene, which may be either adjacent or inverted about the horizontal meridian of the visual field. [8, 9, 13] This is most likely attributable to off-target stimulation of tissue across a sulcus, supported by observations that this phenomena occurs less severely with penetrating microelectrodes (which discharge less than surface electrodes). [13, 50]

The space a phosphene occupies in the visual fields varies with eccentricity and stimulation current. Early evaluations of phosphene size using various objects at arm's length [8, 9, 13, 40, 53] have generally been agreeable with more formal estimates using degrees of visual field. [10, 42, 54] Most phosphenes are 1-2° of visual field in diameter and range from 0.1-10° (a "grain of sago" to a coin at arm's length) . [8, 10, 54] The size of phosphenes depends on where they are located in the visual fields; more peripheral phosphenes are larger and reportedly have less distinct borders. [12, 42, 54] The variation of phosphenes with size is consistent with behavioural studies in non-human primates undergoing cortical stimulation [46] and the phenomenon of cortical magnification, where the central visual field is overproportionately represented on the surface of the brain. [61] Phosphenes also appear to increase in size with stimulation amplitude [12, 42, 54], though one early report of microelectrode stimulation also described instances where phosphene size decreased which have not been subsequently reproduced. [13]

While circular phosphenes are ubiquitous [8–10, 13, 41], other phosphene shapes have been reported. The most consistently reported shape other than circles are elongated elliptical or linear phosphenes oriented in horizontal, oblique or vertical orientations. [8–10, 12, 56] A few reports identify shapes ranging from triangles and stars, to checkerboards, to face or eye-like hallucinatory sensations. [41, 43] Often,

more abstract phosphenes appear on stimulation of later visual areas of the brain, which may not be relevant for prostheses targeting only primary visual cortex. [43] While phosphene shapes appear loosely related to the putative role of different brain regions [41], no studies have been able to deliberately control the shape of phosphenes.

Finally, phosphenes have been repeatedly shown to move with eye movements and have been likened to the movement of retinal afterimages. [8, 13, 40] It appears the whole map of phosphenes moves as multiple phosphenes maintain their relative positions after movement. [13, 40]

3.1.3 Temporal Properties of a Phosphene

Phosphenes generally appear synchronous with stimulation. [13, 56] It is difficult to measure the latency of percepts without also including motor reaction time, but studies comparing phosphene onset reaction times to auditory stimuli suggest that additional latency is minimal. [12] In multiple studies, phosphenes have been sporadically reported to persist for up to 20 minutes after stimulation ceased, particularly after a high-discharge stimulation prior. [8, 12, 13, 40] Perhaps paradoxically, phosphenes purposefully sustained by continuous stimulation demonstrate significant fading in as little as 15 seconds. [9, 13] The fading effect of phosphenes is also reflected over separate trials, where phosphenes progressively dim in each subsequent trial (though they "reset" the next day). [13]

On a shorter time scale, phosphenes elicited by surface stimulation may also flicker. The phosphene flicker produced by surface electrodes is fixed, fast, and asynchronous with hardware or physiological pulses. [8, 9, 12, 40, 53] This differs from the "flicker" produced by two separate successive stimulations, which disappears at stimulation frequencies of approximately 33Hz (though an overlying intrinsic flicker remains). [12] It is unknown whether flicker also occurs in stimulation with microelectrodes; of the few studies of stimulation with penetrating microelectrodes, flicker was not reported. [10, 13]

In summary, we can exert only very limited control over what individual phosphenes look like. Phosphenes are also highly variable, both between-individuals and between-

electrodes. While there are points of agreement between studies, such as the effect of stimulation current on brightness, other phenomena, such as colour and flickering, remain contentious. It remains unclear whether these disagreements are due to differences in stimulation parameters, hardware, participants or pathology.

3.2 Phosphenes Produced by Stimulating Multiple Electrodes

The appearance of images containing multiple phosphenes is fundamental to modern cortical prostheses as very little information can be transmitted through only a single electrode at once. [6, 37] The intention for cortical prostheses is to produce perceivable *patterns* which can be interpreted. The eventual hope is to approximate natural images with phosphenes used like pixels of a graphical display. Early chronic implants operated on this principle, albeit with very low resolution. [53]

However, the empirical evidence on *what* is perceived when multiple electrodes are stimulated is surprisingly scarce.

At the most simple level, two electrodes which produce individual phosphenes appear to also produce two separate perceivable phosphenes when stimulated simultaneously. [8, 9] Sometimes, the size of each phosphene decreases compared to individual stimulation, and the distance between phosphenes may increase. [55] When close together, these phosphenes may fuse together into a single percept. [8, 9] However, this is not always the case; in fact, dimmer phosphenes may not be perceived at all [9, 40, 55], though there is some evidence that increasing the stimulation amplitude may reintroduce the dimmer percept. [13] As a result, increasing the number of electrodes may not linearly increase the number of perceived phosphenes.

Several studies have characterised greater numbers of simultaneous phosphenes. Early evidence suggested that four-phosphene patterns (e.g. a square) could be recognised, but not reliably as spurious phosphenes appeared and some expected phosphenes were not perceived. [40] Another patient was able to perceive a six-phosphene vertical line. [13] Modern studies, however, have provided conflicting results. In one study in an epilepsy patient, five electrodes stimulated at once were

only able to produce two perceivable phosphenes that was not simply the aggregate of each of the five phosphenes. [56] In non-human-primates, stimulation of visual cortex simultaneously at two spacially distant points did not sum to an joint signal, further suggesting a separation of processing of simultaneous stimulation. [62]

The reasoning behind this difficulty is thought to be because cortical visual prostheses unselectively stimulate local regions of the brain. In normal primary visual cortex, neurons are typically selectively stimulated by specific image features such as the orientation of lines in the visual fields. [63] When electrodes instead unselectively stimulate neurons, the pattern of neural stimulation is unnatural and later visual areas may not immediately be able to decode the unrecognisable stimulus. [56]

Despite the difficulties of these temporary experiments, chronic studies suggest that patients are able to use this information usefully after a learning period. There are brief reports of a patient with a chronic implant being able to read phosphene patterns on a 64-electrode implant at 30 letters per minute, similar to Braille [49] Reports on a different patient from the same group described the ability to recognise symbols and letters at an estimated visual acuity of 20/1200 (seeing at 20 metres what could normally be seen at 1200 metres). [53] However, due to the absence of any further chronic studies of implants in blind patients, the upper limit to which people can learn to recognise phosphene patterns is unknown.

Given these limited studies of combinations of phosphenes, there is a tremendous gap between the current knowledge of phosphene patterns and the proposed mechanism of cortical prostheses. It is entirely unclear whether people can perceive phosphene patterns on the order of tens or hundreds, whether people can learn to perceive these patterns in useful ways, or to what degree these patterns may change.

3.3 Summary of the Perceptual Limitations of Phosphenes

The major issues surrounding the current literature on phosphenes are therefore:

- 1. There are no modern studies of phosphenes evoked in blind but otherwise-**

healthy patients, the primary demographic of cortical visual prostheses.

There are also scant chronic studies, none of which have been conducted with penetrating microelectrodes. [12, 49, 53] Emerging clinical trials will help resolve this issue, but until such studies bear fruit, our knowledge on what cortical prosthetic vision looks like may not be readily applicable to new devices.

2. **Phosphenes are highly variable.** Almost all features of phosphenes display uncontrollable variability, and the only two properties of phosphenes we have been shown to reliably control are phosphene brightness and size. [12] This variability permeates between electrodes, between patients and between studies. The heterogeneity, low sample size and skewed populations of the literature have made it difficult to distinguish the root cause of such variation.

3. **The interpretability of patterns formed by multiple phosphenes is unclear.** There is conflict amongst studies on whether multiple phosphenes at once can be integrated simultaneously, or whether people can learn to compensate for initial difficulties with interpreting phosphenes.

As a result, there is considerable uncertainty on exactly what visual sensations modern devices can give to implantees on a case-by-case basis.

4 How can we make what people see most useful?

Our ability to control the appearance of individual phosphenes and their patterns is clearly limited. In this section, we briefly review current literature on how images can be represented strategically in phosphene space to overcome these limitations.

4.1 The Role of Simulated Prosthetic Vision

Because of the difficulties of implanting electrode arrays, little research has been conducted on what methods of representing information in phosphene space are most useful *in vivo*. The only cortical implant which has been connected to a camera in humans was the Dobelle Implant in 2000. [53]. The Dobelle Implant used direct image processing techniques fitting of the software capabilities of the time, which essentially downsampled the camera image and directly mapped the brightness to implant electrodes. [53]. The group briefly entertained the idea of using edge-detection for more selective stimulation, but no subsequent studies reported the outcomes of this idea.

To allow the testing of new image processing algorithms in the absence of access to real implantees, research in image processing algorithms has largely moved to simulations of prosthetic vision. [64, 65] Simulated prosthetic vision is the primary vehicle through which most new image processing algorithms are tested. The features of simulated prosthetic vision have been reviewed previously. [64] Briefly, camera information is processed and rendered onto a head-mounted or other display as simulated phosphenes. Typically, these simulated phosphenes are rendered as the most commonly reported percept - white dots with a Gaussian blur filter applied. [64] In this way, phosphene "images" are displayed for the user with the aim to approximate the prosthetic vision of an implantee. Examples of these simulated renders are shown in Figure 4.

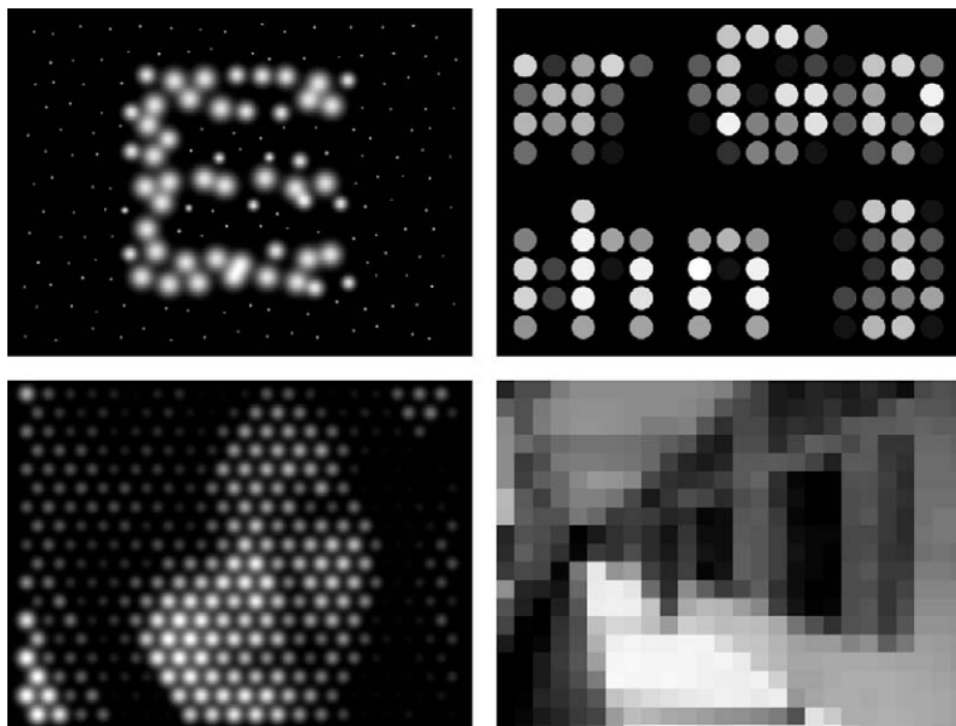


Figure 4: Examples of different simulated renders of phosphenes. Figure from Chen et al. 2009. [64]

4.2 A Brief Outline of Current Image Processing Approaches

4.2.1 Direct Methods

The prevailing paradigm of image processing for early cortical prostheses was to directly map camera images to a grid of electrodes as though they were superimposed. [13, 53] This produces a phosphene image like a mask full of holes placed on top of the original image. In this way, prosthetic vision began by attempting to emulate natural vision as closely as possible.

Such an approach may work with a large number of electrodes if all phosphenes could be interpreted correctly as "pixels". One study estimated that approximately 625 phosphenes would be sufficient to reach a visual acuity of 20/30, suitable for most general tasks [66].

However, there are several issues of direct methods when compared with the perceptual limitations of phosphenes: Direct methods, by virtue of keeping faithful to the original image, tend to produce simulated phosphene renders with large numbers of "on" phosphenes, particularly in well-lit environments. As the ability to interpret multiple phosphenes simultaneously is not well established, stimulating many electrodes at once may not produce the expected visual percept. Because of the high variability of phosphenes, it is also unlikely that the quality and spatial distribution of pixels of a transformed image could be reproduced as faithfully as intended.

Moreover, the quality of these methods very rapidly degrades once resolution drops. [67] No new implants are capable of producing 625 distinct phosphenes. [6] Without the resolution to support the interpretation of low-level features of directly processed images, images can be uninterpretable.

4.2.2 Edge-Based Methods

Edge-detection refers to image processing algorithms methods which highlight the edges of objects only. [68] Edges require less phosphenes at once and may reduce the amount of redundant information in an image. This is important when we can consider

that the number of perceivable phosphenes may not increase linearly with the number of stimulated electrodes. [55]

Edge detectors such as the Canny [68] edge detector are widely used. These edge detectors are able to detect fast pixel gradients in images, which typically occur at boundaries. If additional inputs to the processing algorithm are possible, then more sophisticated techniques can be used. For example, the use of a range camera or other depth sensing devices can be used to more intelligently find non-background edges. [69] State-of-the-art convolutional neural networks (CNNs) have also been applied to edge detection for prosthetic vision by semantic pixel labelling of images of rooms and determining edges by boundaries between walls. [70] An illustration of an edge-detection algorithm combined with a object-filling algorithm is shown in Figure 5.

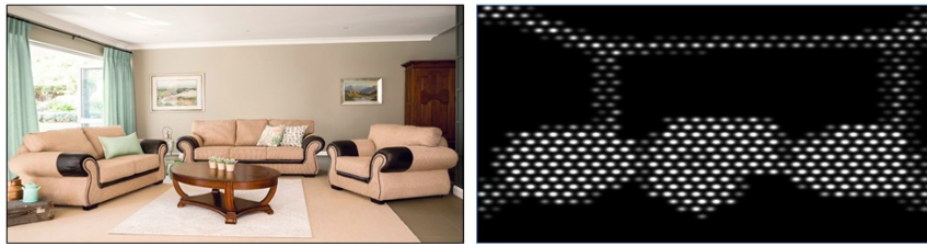


Figure 5: An example of an edge-detection algorithm with a CNN-driven object-filling algorithm to render a clean simulated phosphene image. Figure from [70]

The difficulty with edge-based methods is that edges easily degrade when resolution drops, similar to direct methods. [17] One method which aimed to resolve the fragility of edges combined edge-based methods with saliency-based methods to give greater form to objects. [71] Such hybrid methods may be more robust than the use of pure edges when faced with significant downsampling.

4.2.3 Saliency-Based Methods

As opposed to naively translating brightness values of camera images to electrode stimulation, saliency-based measures more intelligently identify the semantics of objects

in a scene. Using this semantic structure, the image can be divided or *segmented* into regions of interest which carry a common semantic meaning (e.g. "background" or "foreground"). [72]

With this approach, more deliberate differences between foreground and background can be made as depicted in Figure 6. [73] When applied to an image classification task, saliency-based methods improved the recognition accuracy of common objects. [67, 71]

Object detection neural networks have also been applied to highlight particular salient features of an image. [74] In these methods, powerful image classification algorithms are able to detect a specified object and solely highlight that object on the simulated phosphene render. While these methods were constrained to only specific objects, they demonstrate leverage of modern progressions in image processing to intelligently identify objects.

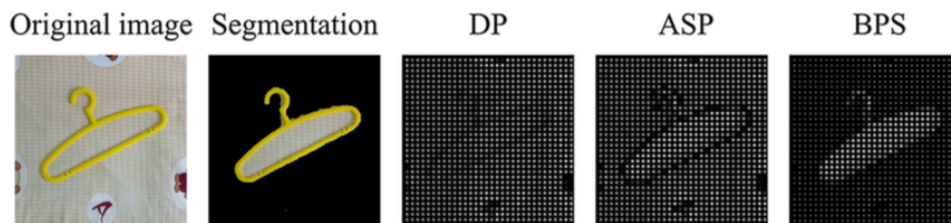


Figure 6: An example of a saliency-based algorithm to highlight a region of interest of the image. Figure from [67]

The methods help highlight what is most likely to be relevant in an image and suppress background, which may otherwise interfere. However the usefulness of this masking approach is importantly constrained by the implantee's ability to subsequently recognise what is being shown. While simulations have demonstrated the utility of this approach, the fidelity of the mask form when phosphenes are irregularly shaped and sized is not clear. While these methods are advantageous compared to direct methods in that irrelevant information may be reduced, it faces the same limitations that low-level forms may be obscured by phosphene distortions.

4.2.4 Transformative Methods

Recent advances in machine learning have meant that computers are now reaching human-level abilities for tasks such as image classification using deep learning methods. [20, 75] Because the processor in a cortical visual prosthesis has access to the full camera image (as opposed to the user, who can only see the phosphene version), the processor has more information available to interpret. Instead of expecting the user to interpret high-level information from degraded phosphene images, some interpretation could be relegated to the computer which can then intelligently re-encode the information in a deliberate manner.

Numerous patents have been filed for such a system. [76, 77] In these systems, important visual cues such as stairs, faces and bank notes are recognised by the computer, which can then remove unnecessary low-level detail and produce compact, abstract images that represent the *concept* of what is seen, not what is actually seen. For example, faces can be recognised and re-encoded as emoticons which cleanly fit in low-resolution space. [69] An example of this approach is shown in 7

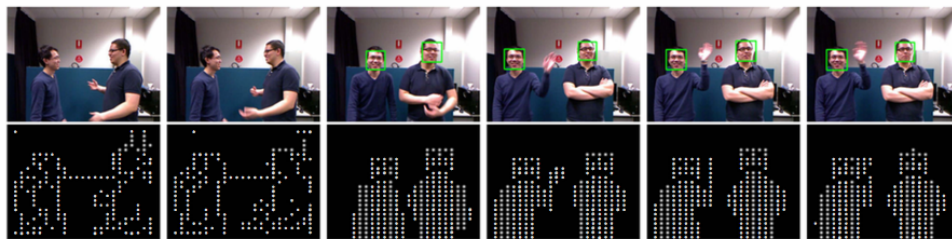


Figure 7: An example of a transformative approach to image processing to re-encode information. Figure from [69]

The chief benefit of such methods is that useful information can be communicated with less phosphenes. Since many low-level details (e.g. "is this bank note folded at the corner?") are not always relevant, the information burden to the user can be reduced to only what is necessary.

The perceptual issues facing these methods are that these typically rely on producing phosphene images that "mimic" real life (e.g. emoticons, which attempt to mimic

faces). [69] However, the perceptual distortions and variability of phosphenes make it unclear whether these mimics could be replicated and thus the ability to specifically evoke these "mimics" *in vivo* with phosphenes is not well established. What may appear cleanly represented in simulated phosphene space may be heavily distorted and even unrecognisable in real implantee settings. A potential rebuttal is that as these methods re-encode information at the bequest of the algorithm implementer, they could be optimised on a case-by-case basis for the particular phosphenes an implantee sees. Individual-level implementations of image processing algorithms are yet to be explored.

4.2.5 Temporal Methods

Given the challenges already faced by patients when trying to interpret multiple simultaneous phosphenes, some groups have begun to explore non-simultaneous methods of conveying patterned information. *Dynamic current steering*, where phosphene patterns are "traced" in quick succession rather than presented all at once have shown significant benefits for letter recognition tasks. [56, 78] Figure 8 illustrates the principle behind this approach. Patients, without prior training, were able to trace the path of phosphenes and interpret simple letterforms accurately. Unlike most of these other methods, dynamic current steering *has* been tested *in vivo* in patients with epilepsy with subdural surface electrodes and in fact arose out of perceptual difficulties noted by the investigators. [56] This highlights one of the issues with simulated prosthetic vision tests; fundamental perceptual differences of electrical neural stimulation may not be discovered and accounted for until tested in real patients.

These methods attempt to resolve the issues surrounding interpreting multiple simultaneous phosphenes by tapping into the brain's natural ability to interpret gross motion. [79] While temporal methods may be slower at conveying information per unit time, they have the advantage of not requiring simultaneous presentations of phosphenes, and requiring less current as only a limited number of phosphenes need to be conveyed for a single frame. It is clear, however, that the difference between temporal methods and form-based methods cannot be assessed in simulated prosthetic

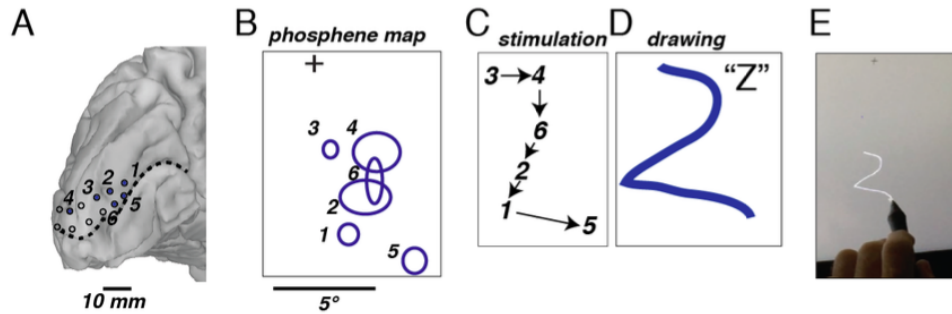


Figure 8: An example of a temporal approach to conveying information through phosphenes, where electrodes are stimulated in succession instead of simultaneously. Figure taken from [56]

vision as the differences between these two methods is based on differences in *in vivo* perception. More clinical trials are needed to establish whether temporal methods of information transfer as opposed to spatial methods are better suited for cortical prosthetic devices.

4.3 The Limitations of Simulated Prosthetic Vision

The crux of many of these methods rests on tests of simulated prosthetic vision, faces significant limitations for generalisability to *in vivo* implants:

1. **The possible difficulties *in vivo* of interpreting phosphenes are not accounted for.** One of the key unknowns in phosphene space are how well the brain can learn to decipher the unselective unnatural stimulation of visual cortex by multiple electrodes simultaneously. [15] Because our uncertainty in this area relates to the unnatural neural stimulation of the visual cortex, this cannot be assessed in a simulated setting. When phosphenes are simply shown on a display, a sighted subject naturally is able to process the displayed patterns making full use of the retinal and neural circuitry distal to the brain. This is most certainly not the case for an implantee. While normal-sighted subjects have often been shown to be able to recognise complex patterns with many simultaneous phosphenes in a simulated setting [65], it is still unclear whether this can be replicated in real

implantees.

2. **Phosphene simulation often does not account for all known properties of phosphenes.** The properties we have described above are rarely all accounted for. For example, temporal effects such as fading and accommodation are not implemented in most simulations. Additionally, most simulations render phosphene images as low-resolution greyscale images with uniform circular [67, 70, 80, 81] or hexagonal pixels. [82], though some studies have also incorporated biologically-based retinotopic distortion. [83, 84] Indeed, many simulated prosthetic vision algorithms approach the problem as chiefly one of low resolution, loss of colour and distortion. This does not accurately reflect the rich (but uncontrollable) perceptual experiences previously reported by *in vivo* experiments in shape, size, colour or flicker.
3. **Many simulations use regular and higher-resolution phosphene grids than have previously been achieved.** Sometimes, psychophysical experiments render on the order of a thousand phosphenes [67, 70, 73] This far outstrips the number of phosphenes which have been tested simultaneously *in vivo* so far (less than a hundred) [53], and also surpasses the estimates of capabilities of modern prostheses. [6] Ultimately, the ability to reliably evoke many cortical phosphenes regularly and of the calibre of many simulations has not been established.
4. **Most simulations do not specifically target cortical phosphenes.** Most advances in image processing methods for simulated prosthetic vision are targeted at retinal prostheses, for which there are already commercially available devices. [28, 85, 86] The lack of distinction between different biological methods of evoking phosphenes in some experiments of simulated prosthetic vision mean results may not be directly translatable.

The implication of these issues is that studies of image processing algorithms in simulated prosthetic vision are not flexible for different phosphene percepts. The methods we have described are dependent on being able to replicate the simulated

renders in *in vivo* implants, but little research has been conducted on this area. It is unclear how the image processing algorithms produced by studies of simulated prosthetic vision could be made flexible for the variability in phosphenes we have described in Section 3

4.4 Future Directions from Advances in Machine Learning

The remaining goal for better addressing the perceptual limitations for phosphenes is to find flexible ways to reconcile image processing algorithms in simulated phosphene space with the wide variability in what phosphenes look like. Image processing outside of prosthetic vision has experienced a wealth of improvements from advances in machine learning, where computers learn patterns from data without prior knowledge of those patterns. [20] Machine learning for image processing is most often applied to the training of **convolutional neural networks** (CNNs), which are layered architectures of image filters modelled after the function of physiological neurons. [75] CNNs have already discovered widespread use in image recognition tasks [87], video recognition tasks [88] and style transfer tasks, where images are modified to mimic the style of other images. [23]

The benefit of CNNs is that they produce image processing algorithms that are trainable from data without requiring explicit programming. This could be applied to the current gap between simulated prosthetic vision and perceptual limitations of phosphenes. By training CNNs on task performance data rather than programming algorithms to produce explicit patterns of phosphenes, CNNs could learn to process images in phosphene-agnostic ways. CNNs could also be tailored for individuals' perceptions of phosphenes, as the dependency of the algorithm shifts towards task performance tests rather than perceptual tests. This essentially positions the problem of deriving image processing algorithms from bottom-up (from phosphenes to algorithm) to top-down (from performance to algorithm).

To our knowledge, CNNs have not been applied in this way to derive image processing algorithms for prosthetic vision, and thus research on this topic is scant. While

CNNs and other neural network architectures have been applied to prosthetic vision to construct new algorithms [70, 74, 89], these have typically been object detection networks involved in segmentation-based processing. There has also been growing interest in using machine learning recognition algorithms for transformative techniques in prosthetic vision. [76] With continuous improvements in the capacity of hardware to support advanced processing [90], it is possible that emergent research in using machine learning to produce flexible, trainable algorithms could improve the utility of a cortical visual prosthesis.

5 Conclusion

Cortical visual prostheses are a promising enabling technology for the blind. As the development of these prostheses is still in early stages, it is still not precisely known what type of vision implantees will be able to achieve with a prosthesis. However, early studies of cortical stimulation seem to suggest that the visual percepts produced will be highly variable and difficult to control. Whilst there have been many progressions in image processing algorithms in simulated prosthetic vision, the ability of these algorithms to translate to the diverse range of *in vivo* phosphenes is unclear. This is because many of these algorithms are based on conceptions of the appearance of phosphenes which we do not know if we can replicate, particularly between different people. There is therefore a gap in the literature between how image processing algorithms could be derived flexibly in a phosphene-agnostic way. Modern advances in machine learning have shown outstanding achievements of layer image filter architectures, or convolutional neural networks, to learn image processing algorithms from data without explicit programming. As these methods have not been applied to prosthetic vision before, preliminary studies which assess the feasibility of using machine learning methods to address the perceptual limitations of phosphenes may be greatly informative.

6 Aims & Hypotheses

The aims and hypotheses of this project are:

Aim 1 To investigate the performance benefit of a CNN-based image processor trained on task-performance on a simple digit recognition task.

Hypothesis 1 A CNN-based image processor trained on task performance improves participants' accuracy on a digit recognition task under simulated prosthetic vision, compared to a brightness-based control.

Aim 2 To investigate the flexibility of a CNN-based image image processor trained on task-performance on a simple digit recognition task.

Hypothesis 2 A CNN-based image processor trained on task performance achieves greater performance consistency between different simulated prosthetic vision renders, compared to a brightness-based control.

7 Project Outline

7.1 Rationale

In this literature review, we have identified a gap in the literature; the absence of flexible image processing algorithms to address the perceptual limitations of phosphenes. We aim to approach this gap in the literature using methods from modern advances in machine learning applied to image processing with CNNs. As CNNs have experienced significant uptake within the image processing community, we believe that a preliminary study into the feasibility of CNNs for image processing in prosthetic vision would be informative. As no prior literature has addressed a similar top-down phosphene-agnostic approach before, we believe this topic is novel and can inform the future development of image processing algorithms for prosthetic vision.

7.2 Brief Outline of Proposed Methods

7.2.1 Planning Phase (May - Jun)

This experiment is comprised of a digit recognition task in six different conditions, composed of different combinations of a phosphene renderer (which aims to simulate the limitations of phosphenes we have addressed in Section 3) and an image processor (including the CNN-based model).

In this experiment, we will test 2 different simulated phosphene renderers with 3 different image processing methods.

The 2 image processing methods will be:

1. A brightness-based algorithm, which controls the brightness of phosphenes by their brightness of the image at the corresponding location
2. A trained CNN-architecture, which is first pretrained on a computational model before being further trained by task performance. First, the weights of a CNN image decoder network are trained on the original digit images to classify digits.

The weights of the image decoder network are then kept stable. Then, an image encoder network (which feeds into the selected phosphene renderer) is placed between the image and the decoder. The weights of the image encoder network are then trained to optimise the performance of the stable image decoder network, which now sees the phosphene render. This pretrained image encoder network is then used as the base image processor.

The 3 simulated phosphene renderers, informed by Section 3 of this literature review, will be:

1. A regular log-polar grid of 100 circular Gaussian-blurred phosphenes
2. A base grid comprising Render 1, with a random set of per-phosphene distortions of phosphene shape, size, location, colour, and multiplicity
3. A base grid comprising Render 2, with non-linear phosphene interactions (i.e. multiple simultaneous phosphenes produce dropout of nearby dimmer phosphenes)

These combinations are summarised in Figure 9

There are a total of 6 combinations of phosphene renderers and image processing methods. The regular-grid, brightness-based combination will serve as the control for this experiment.

7.2.2 Experimental Phase (Jul - Sep)

The full methods of this phase will be informed by progress during the planning phase in May and June.

10 normal-sighted participants will be recruited from students and staff at Monash University and informed consent will be obtained for a psychophysics study. Each participant will be allocated to the control combination of phosphene and renderer, and one of the remaining five combinations of phosphene renderer/image processing method, for a total of two participants per non-control combination.

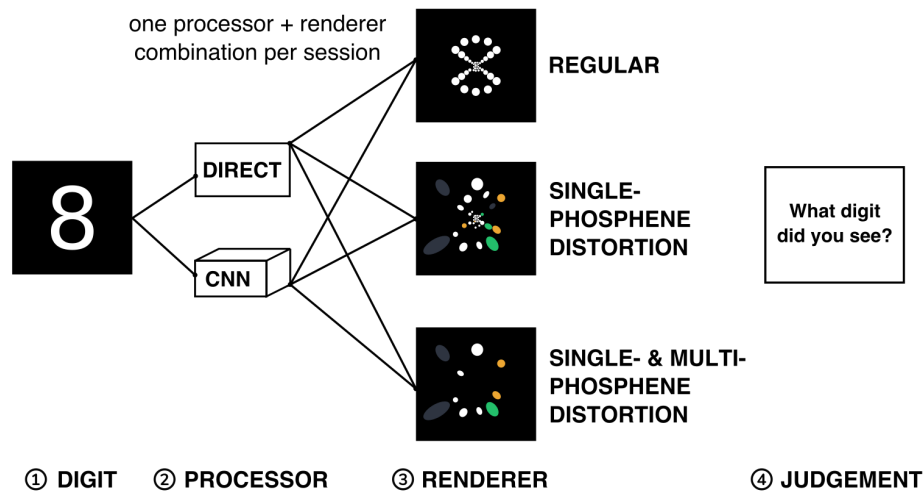


Figure 9: Basic overview of the processor and renderer combinations to be tested in this experiment. Figure made by the author.

Each participant will be introduced to a psychophysics room with a computer graphical display. Each participant will undergo a training and a testing phase. During the training phase, the participant is shown a single digit and is asked to input what they believe is the correct digit. They will receive feedback on their response with the intention to improve their ability to recognise patterns as would occur during the training phase for actual implantees. The participant will repeat this process for 30 trials of 20 digits, subject to pilot trials in early July. During this training phase, if the participant is allocated to the CNN group, the encoder will iteratively modify its network weights based on the correctness of the responses of the participant.

During the testing phase, the participant will undergo the same process of classifying a single digit, but without feedback. The accuracy of classification during the testing phase will be recorded. The classification accuracy of the CNN model with the brightness-based image processor will be evaluated with a t-test separately for each phosphene renderer to address Hypothesis 1. The classification accuracy of each image processor across renderers will be compared with an ANOVA to address

Hypothesis 2.

References

- [1] Bourne RRA, Flaxman SR, Braithwaite T, Cicinelli MV, Das A, Jonas JB, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *The Lancet Global Health*. 2017 Sep;5(9):e888–e897. Available from: [https://www.thelancet.com/journals/langlo/article/PIIS2214-109X\(17\)30293-0/abstract](https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(17)30293-0/abstract).
- [2] Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, et al. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *The Lancet Global Health*. 2017 Dec;5(12):e1221–e1234. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2214109X17303935>.
- [3] Lee DA, Higginbotham EJ. Glaucoma and its treatment: A review. *American Journal of Health-System Pharmacy*. 2005 Apr;62(7):691–699. Available from: <https://academic.oup.com/ajhp/article/62/7/691/5134357>.
- [4] Zachariades N, Papavassiliou D, Christopoulos P. Blindness after facial trauma. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*. 1996 Jan;81(1):34–37. Available from: <http://www.sciencedirect.com/science/article/pii/S1079210496801442>.
- [5] Normann RA, Greger BA, House P, Romero SF, Pelayo F, Fernandez E. Toward the development of a cortically based visual neuroprosthesis. *Journal of Neural Engineering*. 2009 May;6(3):035001. Available from: <https://doi.org/10.1088/1741-2560/6/3/035001>.
- [6] Lewis PM, Ackland HM, Lowery AJ, Rosenfeld JV. Restoration of vision in blind individuals using bionic devices: A review with a focus on cortical visual prostheses. *Brain Research*. 2015 Jan;1595:51–73. Available from: <http://www.sciencedirect.com/science/article/pii/S0006899314015674>.

- [7] Foroushani AN, Pack CC, Sawan M. Cortical visual prostheses: from microstimulation to functional percept. *Journal of Neural Engineering*. 2018 Feb;15(2):021005. Available from: <https://doi.org/10.1088%2F1741-2552%2F1502021005>.
- [8] Brindley GS, Lewin WS. The sensations produced by electrical stimulation of the visual cortex. *The Journal of Physiology*. 1968 May;196(2):479–493. Available from: <http://doi.wiley.com/10.1113/jphysiol.1968.sp008519>.
- [9] Dobelle WH, Mladejovsky MG. Phosphenes produced by electrical stimulation of human occipital cortex, and their application to the development of a prosthesis for the blind. *The Journal of Physiology*. 1974 Dec;243(2):553–576. Available from: <http://doi.wiley.com/10.1113/jphysiol.1974.sp010766>.
- [10] Bak M, Girvin J, Hambrecht DFT, Kufta CV, Loeb GE, Schmidt EM. Visual sensations produced by intracortical microstimulation of the human occipital cortex. *Medical and Biological Engineering and Computing*. 1990;28:257–259.
- [11] Bosking WH, Beauchamp MS, Yoshor D. Electrical Stimulation of Visual Cortex: Relevance for the Development of Visual Cortical Prosthetics. *Annual Review of Vision Science*. 2017 Sep;3(1):141–166. Available from: <http://www.annualreviews.org/doi/10.1146/annurev-vision-111815-114525>.
- [12] Rushton DN, Brindley GS. Properties of Cortical Electrical Phosphenes. In: Cool SJ, Smith EL, editors. *Frontiers in Visual Science*. Springer Series in Optical Sciences. Springer Berlin Heidelberg; 1978. p. 574–593.
- [13] Schmidt EM, Bak MJ, Hambrecht FT, Kufta CV, O'Rourke DK, Vallabhanath P. Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. *Brain*. 1996 Apr;119(2):507–522. Available from: <https://academic.oup.com/brain/article/119/2/507/382434>.
- [14] Fernández E, Pelayo F, Romero S, Bongard M, Marin C, Alfaro A, et al. Development of a cortical visual neuroprosthesis for the blind: the relevance of

neuroplasticity. *Journal of Neural Engineering*. 2005 Nov;2(4):R1–R12. Available from: <https://doi.org/10.1088%2F1741-2560%2F2%2F4%2Fr01>.

- [15] Beyeler M, Rokem A, Boynton GM, Fine I. Learning to see again: biological constraints on cortical plasticity and the implications for sight restoration technologies. *bioRxiv*. 2017 Mar;p. 115188. Available from: <https://www.biorxiv.org/content/10.1101/115188v1>.
- [16] Barnes N. The role of computer vision in prosthetic vision. *Image and Vision Computing*. 2012 Aug;30(8):478–479. Available from: <http://www.sciencedirect.com/science/article/pii/S0262885612000832>.
- [17] Buffoni LX, Coulombe J, Sawan M. Image Processing Strategies Dedicated to Visual Cortical Stimulators: A Survey. *Artificial Organs*. 2005;29(8):658–664. Available from: <http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1525-1594.2005.29104.x>.
- [18] Chang MH, Kim HS, Shin JH, Park KS. Facial identification in very low-resolution images simulating prosthetic vision. *Journal of Neural Engineering*. 2012 Aug;9(4):046012. Available from: <http://stacks.iop.org/1741-2552/9/i=4/a=046012?key=crossref.82f47f235753f422323435341cb2f4b0>.
- [19] Sharmili N, Swapna N, Ramakrishna G. Comparative analysis of image processing algorithms for visual prosthesis. In: 2017 International Conference on Communication and Signal Processing (ICCSP); 2017. p. 1120–1124.
- [20] Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. Deep learning for visual understanding: A review. *Neurocomputing*. 2016 Apr;187:27–48. Available from: <http://www.sciencedirect.com/science/article/pii/S0925231215017634>.
- [21] Cireşan D, Meier U, Schmidhuber J. Multi-column Deep Neural Networks for Image Classification. *arXiv:12022745 [cs]*. 2012 Feb;ArXiv: 1202.2745. Available from: <http://arxiv.org/abs/1202.2745>.

- [22] Ciresan D, Giusti A, Gambardella LM, Schmidhuber J. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. *Advances in Neural Information Processing Systems* 25. Curran Associates, Inc.; 2012. p. 2843–2851. Available from: <http://papers.nips.cc/paper/4741-deep-neural-networks-segment-neuronal-membranes-in-electron-microscopy-images.pdf>.
- [23] Gatys LA, Ecker AS, Bethge M. Image Style Transfer Using Convolutional Neural Networks. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE; 2016. p. 2414–2423. Available from: <http://ieeexplore.ieee.org/document/7780634/>.
- [24] Thurston AJ. Paré and Prosthetics: The Early History of Artificial Limbs. *ANZ Journal of Surgery*. 2007;77(12):1114–1119. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1445-2197.2007.04330.x>.
- [25] Weiland JD, Humayun MS. Visual Prosthesis. *Proceedings of the IEEE*. 2008 Jul;96(7):1076–1084.
- [26] Ong JM, da Cruz L. The bionic eye: a review. *Clinical & Experimental Ophthalmology*. 2012 Jan;40(1):6–17. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1442-9071.2011.02590.x>.
- [27] Humayun MS, de Juan E, Dagnelie G, Greenberg RJ, Propst RH, Phillips DH. Visual perception elicited by electrical stimulation of retina in blind humans. *Archives of Ophthalmology (Chicago, Ill: 1960)*. 1996 Jan;114(1):40–46.
- [28] Stingl K, Schippert R, Bartz-Schmidt KU, Besch D, Cottrill CL, Edwards TL, et al. Interim Results of a Multicenter Trial with the New Electronic Subretinal Implant Alpha AMS in 15 Patients Blind from Inherited Retinal Degenerations. *Frontiers in Neuroscience*. 2017;11. Available from: <https://www.frontiersin.org/articles/10.3389/fnins.2017.00445/full>.

- [29] Veraart C, Raftopoulos C, Mortimer JT, Delbeke J, Pins D, Michaux G, et al. Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Research*. 1998 Nov;813(1):181–186. Available from: <http://www.sciencedirect.com/science/article/pii/S0006899398009779>.
- [30] Panetsos F, Sanchez-Jimenez A, Diaz-de Cerio ER, Diaz-Guemes I, Sanchez FM. Consistent Phosphenes Generated by Electrical Microstimulation of the Visual Thalamus. An Experimental Approach for Thalamic Visual Neuroprostheses. *Frontiers in Neuroscience*. 2011;5. Available from: <https://www.frontiersin.org/articles/10.3389/fnins.2011.00084/full>.
- [31] Lowery AJ, Rosenfeld JV, Lewis PM, Browne D, Mohan A, Brunton E, et al. Restoration of vision using wireless cortical implants: The Monash Vision Group project. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2015. p. 1041–1044.
- [32] Lowery AJ, Rosenfeld JV, Rosa MGP, Brunton E, Rajan R, Mann C, et al. Monash Vision Group's Gennaris Cortical Implant for Vision Restoration. In: Gabel VP, editor. *Artificial Vision: A Clinical Guide*. Cham: Springer International Publishing; 2017. p. 215–225. Available from: https://doi.org/10.1007/978-3-319-41876-6_17.
- [33] Troyk PR. The Intracortical Visual Prosthesis Project. In: Gabel VP, editor. *Artificial Vision: A Clinical Guide*. Cham: Springer International Publishing; 2017. p. 203–214. Available from: https://doi.org/10.1007/978-3-319-41876-6_16.
- [34] Fernández E, Normann RA. CORTIVIS Approach for an Intracortical Visual Prostheses. In: Gabel VP, editor. *Artificial Vision: A Clinical Guide*. Cham: Springer International Publishing; 2017. p. 191–201. Available from: https://doi.org/10.1007/978-3-319-41876-6_15.
- [35] SecondSight. Second Sight Medical Products, Inc. Presents Positive Interim Results at the Fifth Annual BRAIN Initiative® Investigators Meeting;p. 2.

- [36] Deering MF. The Limits of Human Vision. In: 2nd International Immersive Projection Technology Workshop. vol. 2; 1998. p. 6.
- [37] Niketeghad S, Pouratian N. Brain Machine Interfaces for Vision Restoration: The Current State of Cortical Visual Prosthetics. *Neurotherapeutics*. 2019 Jan;16(1):134–143. Available from: <https://doi.org/10.1007/s13311-018-0660-1>.
- [38] SecondSight. Early Feasibility Study of the Orion Visual Cortical Prosthesis System - Full Text View - ClinicalTrials.gov;. Available from: <https://clinicaltrials.gov/ct2/show/NCT03344848>.
- [39] ANZCTR. The first safety study of a Bionic Vision Prosthesis in adults with complete and untreatable blindness; 2018. Available from: <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=373899>.
- [40] Dobelle WH, Mladejovsky MG, Girvin JP. Artificial Vision for the Blind: Electrical Stimulation of Visual Cortex Offers Hope for a Functional Prosthesis. *Science*. 1974 Feb;183(4123):440–444. Available from: <http://www.sciencemag.org/cgi/doi/10.1126/science.183.4123.440>.
- [41] Lee HW, Hong SB, Seo DW, Tae WS, Hong SC. Mapping of functional organization in human visual cortex: Electrical cortical stimulation. *Neurology*. 2000 Feb;54(4):849–854. Available from: <http://www.neurology.org/cgi/doi/10.1212/WNL.54.4.849>.
- [42] Winawer J, Parvizi J. Linking Electrical Stimulation of Human Primary Visual Cortex, Size of Affected Cortical Area, Neuronal Responses, and Subjective Experience. *Neuron*. 2016 Dec;92(6):1213–1219. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0896627316308467>.
- [43] Murphey DK, Maunsell JHR, Beauchamp MS, Yoshor D. Perceiving electrical stimulation of identified human visual areas. *Proceedings of the National Academy*

of Sciences. 2009 Mar;106(13):5389–5393. Available from: <https://www.pnas.org/content/106/13/5389>.

- [44] Collins KL, Sarma D, Hakimian S, Tsai JJ, Ojemann JG. Preserved evoked conscious perception of phosphenes with direct stimulation of deafferented primary visual cortex. *Epilepsy & Behavior Case Reports*. 2019 Jan;11:84–86. Available from: <http://www.sciencedirect.com/science/article/pii/S2213323218301270>.
- [45] Tehovnik EJ, Slocum WM, Carvey CE, Schiller PH. Phosphene induction and the generation of saccadic eye movements by striate cortex. *Journal of Neurophysiology*. 2005 Jan;93(1):1–19.
- [46] Tehovnik EJ, Slocum WM. Phosphene Induction by Microstimulation of Macaque V1. *Brain research reviews*. 2007 Feb;53(2):337–343. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1850969/>.
- [47] Tehovnik EJ, Slocum WM. Microstimulation of V1 delays visually guided saccades: a parametric evaluation of delay fields. *Experimental Brain Research*. 2007 Jan;176(3):413–424.
- [48] Tehovnik EJ, Slocum WM, Smirnakis SM, Tolia AS. Microstimulation of visual cortex to restore vision. *Progress in Brain Research*. 2009;175:347–375.
- [49] Dobelle WH, Mladejovsky MG, Evans JR, Roberts TS, Girvin JP. ‘Braille’ reading by a blind volunteer by visual cortex stimulation. *Nature*. 1976 Jan;259(5539):111–112. Available from: <http://www.nature.com/articles/259111a0>.
- [50] Dobelle WH, Turkel J, Henderson DC, Evans JR. Mapping the Representation of the Visual Field by Electrical Stimulation of Human Visual Cortex. *American Journal of Ophthalmology*. 1979 Oct;88(4):727–735. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0002939479906731>.

- [51] Allison T, Ginter H, McCarthy G, Nobre AC, Puce A, Luby M, et al. Face recognition in human extrastriate cortex. *Journal of Neurophysiology*. 1994 Feb;71(2):821–825. Available from: <https://www.physiology.org/doi/abs/10.1152/jn.1994.71.2.821>.
- [52] Puce A, Allison T, McCarthy G. Electrophysiological Studies of Human Face Perception. III: Effects of Top-down Processing on Face-specific Potentials. *Cerebral Cortex*. 1999 Jul;9(5):445–458. Available from: <https://academic.oup.com/cercor/article/9/5/445/310738>.
- [53] Dobelle WH. Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex. *ASAIO Journal*. 2000 Feb;46(1):3. Available from: https://journals.lww.com/asaiojournal/Fulltext/2000/01000/Artificial_Vision_for_the_Blind_by_Connecting_a.2.aspx.
- [54] Bosking WH, Sun P, Ozker M, Pei X, Foster BL, Beauchamp MS, et al. Saturation in Phosphene Size with Increasing Current Levels Delivered to Human Visual Cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2017;37(30):7188–7197.
- [55] Bosking WH, Foster B, Sun P, Beauchamp MS, Yoshor D. Rules Governing Perception of Multiple Phosphenes by Human Observers. *Neuroscience*; 2018. Available from: <http://biorxiv.org/lookup/doi/10.1101/302547>.
- [56] Beauchamp MS, Bosking W, Sun P, Foster B, Niketeghad S, Pouratian N, et al. Dynamic Electrical Stimulation of Sites in Visual Cortex Produces Form Vision in Sighted and Blind Humans. *bioRxiv*. 2018 Nov;p. 462697. Available from: <https://www.biorxiv.org/content/10.1101/462697v1>.
- [57] Solomon SG, Lennie P. The machinery of colour vision. *Nature Reviews Neuroscience*. 2007 Apr;8(4):276–286. Available from: <https://www.nature.com/articles/nrn2094>.

- [58] Vladusich T, Lucassen MP, Cornelissen FW. Brightness and Darkness as Perceptual Dimensions. *PLOS Computational Biology*. 2007 Oct;3(10):e179. Available from: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0030179>.
- [59] Fox PT, Miezin FM, Allman JM, Essen DV, Raichle ME. Retinotopic organization of human visual cortex mapped with positron- emission tomography. *Journal of Neuroscience*. 1987 Mar;7(3):913–922. Available from: <http://www.jneurosci.org/content/7/3/913>.
- [60] Engel SA, Glover GH, Wandell BA. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*. 1997 Mar;7(2):181–192. Available from: <https://academic.oup.com/cercor/article/7/2/181/318870>.
- [61] Born RT, Trott AR, Hartmann TS. Cortical magnification plus cortical plasticity equals vision? *Vision Research*. 2015 Jun;111:161–169. Available from: <http://www.sciencedirect.com/science/article/pii/S0042698914002351>.
- [62] Ghose K, Maunsell JHR. A Strong Constraint to the Joint Processing of Pairs of Cortical Signals. *Journal of Neuroscience*. 2012 Nov;32(45):15922–15933. Available from: <http://www.jneurosci.org/content/32/45/15922>.
- [63] Ben-Yishai R, Bar-Or RL, Sompolinsky H. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*. 1995 Apr;92(9):3844–3848. Available from: <https://www.pnas.org/content/92/9/3844>.
- [64] Chen SC, Suaning GJ, Morley JW, Lovell NH. Simulating prosthetic vision: I. Visual models of phosphenes. *Vision Research*. 2009 Jun;49(12):1493–1506. Available from: <http://www.sciencedirect.com/science/article/pii/S0042698909000467>.
- [65] Chen SC, Suaning GJ, Morley JW, Lovell NH. Simulating prosthetic vision:

- II. Measuring functional capacity. *Vision Research*. 2009 Sep;49(19):2329–2343. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042698909003253>.
- [66] Cha K, Horch K, A Normann R. Simulation of a phosphene-based visual field: Visual acuity in a pixelized vision system. *Annals of biomedical engineering*. 1992 Feb;20:439–49.
- [67] Li H, Su X, Wang J, Kan H, Han T, Zeng Y, et al. Image processing strategies based on saliency segmentation for object recognition under simulated prosthetic vision. *Artificial Intelligence in Medicine*. 2018 Jan;84:64–78. Available from: <http://www.sciencedirect.com/science/article/pii/S09333365716304195>.
- [68] Canny J. Readings in computer vision. In: *A Computational Approach to Edge Detection*. Morgan Kaufmann; 1987. p. 184–203.
- [69] Lui WLD, Browne D, Kleeman L, Drummond T, Wai Ho Li. Transformative Reality: Improving bionic vision with robotic sensing. In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. San Diego, CA: IEEE; 2012. p. 304–307. Available from: <http://ieeexplore.ieee.org/document/6345929/>.
- [70] Sanchez-Garcia M, Martinez-Cantin R, Guerrero JJ. Structural and object detection for phosphene images. *arXiv:180909607 [cs]*. 2018 Sep;ArXiv: 1809.09607. Available from: <http://arxiv.org/abs/1809.09607>.
- [71] Han T, Li H, Lyu Q, Zeng Y, Chai X. Object recognition based on a foreground extraction method under simulated prosthetic vision. In: *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*; 2015. p. 172–175.
- [72] Pal NR, Pal SK. A review on image segmentation techniques. *Pattern Recognition*. 1993 Sep;26(9):1277–1294. Available from: <http://www.sciencedirect.com/science/article/pii/003132039390135J>.

- [73] Guo F, Yang Y, Gao Y. Optimization of Visual Information Presentation for Visual Prosthesis. *International Journal of Biomedical Imaging*. 2018;2018:1–12. Available from: <https://www.hindawi.com/journals/ijbi/2018/3198342/>.
- [74] Macé MJM, Guivarch V, Denis G, Jouffrais C. Simulated Prosthetic Vision: The Benefits of Computer-Based Object Recognition and Localization: Computer Vision for Low Resolution Visual Implants. *Artificial Organs*. 2015 Jul;39(7):E102–E113. Available from: <http://doi.wiley.com/10.1111/aor.12476>.
- [75] Rawat W, Wang Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*. 2017 Jun;29(9):2352–2449. Available from: https://doi.org/10.1162/neco_a_00990.
- [76] Chichilnisky EJ, Greschner M, Jepsen L. SMART PROSTHESIS FOR FACILITATING ARTIFICIAL VISION USING SCENE ABSTRACTION. 2018;p. 27.
- [77] Li WH, Tang TJJ, Lui WLD. Going beyond vision to improve bionic vision. In: 2013 IEEE International Conference on Image Processing. Melbourne, Australia: IEEE; 2013. p. 1555–1558. Available from: <http://ieeexplore.ieee.org/document/6738320/>.
- [78] Spencer TC, Fallon JB, Shivdasani MN. Creating virtual electrodes with 2D current steering. *Journal of Neural Engineering*. 2018 Mar;15(3):035002. Available from: <https://doi.org/10.1088/1751-0755/15/3/035002>.
- [79] Grossman ED, Blake R. Brain Areas Active during Visual Perception of Biological Motion. *Neuron*. 2002 Sep;35(6):1167–1175. Available from: <http://www.sciencedirect.com/science/article/pii/S0896627302008978>.
- [80] McCarthy C, Walker JG, Lieby P, Scott A, Barnes N. Mobility and low contrast trip hazard avoidance using augmented depth. *Journal of Neural Engineering*. 2014 Nov;12(1):016003. Available from: <https://doi.org/10.1088/1751-0755/12/1/016003>.

- [81] Hu J, Xia P, Gu C, Qi J, Li S, Peng Y. Recognition of Similar Objects Using Simulated Prosthetic Vision. *Artificial Organs*. 2014;38(2):159–167. Available from: <http://www.onlinelibrary.wiley.com/doi/abs/10.1111/aor.12147>.
- [82] Chen SC, Lovell NH, Suaning GJ. Effect on prosthetic vision visual acuity by filtering schemes, filter cut-off frequency and phosphene matrix: a virtual reality simulation. In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. vol. 2; 2004. p. 4201–4204.
- [83] Josh H, Yong B, Kleeman L, Road W. A Real-time FPGA-based Vision System for a Bionic Eye. 2011;p. 8.
- [84] Josh H, Mann C, Kleeman L, Lui WLD. Psychophysics testing of bionic vision image processing algorithms using an FPGA Hatpack. In: *2013 IEEE International Conference on Image Processing*. Melbourne, Australia: IEEE; 2013. p. 1550–1554. Available from: <http://ieeexplore.ieee.org/document/6738319/>.
- [85] Luo YHL, da Cruz L. The Argus® II Retinal Prosthesis System. *Progress in Retinal and Eye Research*. 2016 Jan;50:89–107. Available from: <http://www.sciencedirect.com/science/article/pii/S1350946215000701>.
- [86] Markowitz M, Rankin M, Mongy M, Patino BE, Manusow J, Devenyi RG, et al. Rehabilitation of lost functional vision with the Argus II retinal prosthesis. *Canadian Journal of Ophthalmology*. 2018 Feb;53(1):14–22. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0008418217309729>.
- [87] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc.; 2012. p. 1097–1105. Available from: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.

- [88] Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L. Large-scale Video Classification with Convolutional Neural Networks; 2014. p. 1725–1732. Available from: https://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Karpathy_Large-scale_Video_Classification_2014_CVPR_paper.html.
- [89] Ge C, Kasabov N, Liu Z, Yang J. A spiking neural network model for obstacle avoidance in simulated prosthetic vision. *Information Sciences*. 2017 Aug;399:30–42. Available from: <http://www.sciencedirect.com/science/article/pii/S0020025517305819>.
- [90] Moore GE. Cramming More Components Onto Integrated Circuits. *Proceedings of the IEEE*. 1998 Jan;86(1):82–85. Available from: <http://ieeexplore.ieee.org/document/658762/>.