



HAL
open science

Multi-level implementation of hierarchical neural models

Kévin Géhère

► **To cite this version:**

Kévin Géhère. Multi-level implementation of hierarchical neural models. Bioengineering. Sorbonne Université, 2018. English. NNT : 2018SORUS335 . tel-02865333

HAL Id: tel-02865333

<https://theses.hal.science/tel-02865333>

Submitted on 11 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

présentée pour obtenir le grade de

DOCTEUR SORBONNE UNIVERSITÉ
délivré par l'Université Pierre et Marie Curie

Spécialité: Sciences de l'ingénieur

par

KÉVIN GÉHÈRE

sous la direction de RYAD BENOSMAN

Titre:

IMPLÉMENTATION MULTI-NIVEAUX DE MODÈLES HIERARCHIQUES NEURONAUX

soutenue publiquement le 14 décembre 2018

JURY

Ryad .B Benosman	Sorbonne Université	Directeur de thèse
Bruno Gas	Sorbonne Université	Examineur
Jean Lorenceau	Sorbonne Université	Examineur
Laurent Perrinet	Université Aix-Marseille	Rapporteur
Sylvain Saighi	Université de Bordeaux	Rapporteur



THÈSE

présentée pour obtenir le grade de

DOCTEUR SORBONNE UNIVERSITÉ
délivré par l'Université Pierre et Marie Curie

Spécialité: Sciences de l'ingénieur

par

KÉVIN GÉHÈRE

sous la direction de RYAD BENOSMAN

Titre:

IMPLÉMENTATION MULTI-NIVEAUX DE MODÈLES HIERARCHIQUES NEURONAUX

soutenue publiquement le 14 décembre 2018

JURY

Ryad .B Benosman	Sorbonne Université	Directeur de thèse
Bruno Gas	Sorbonne Université	Examineur
Jean Lorenceau	Sorbonne Université	Examineur
Laurent Perrinet	Université Aix-Marseille	Rapporteur
Sylvain Saighi	Université de Bordeaux	Rapporteur

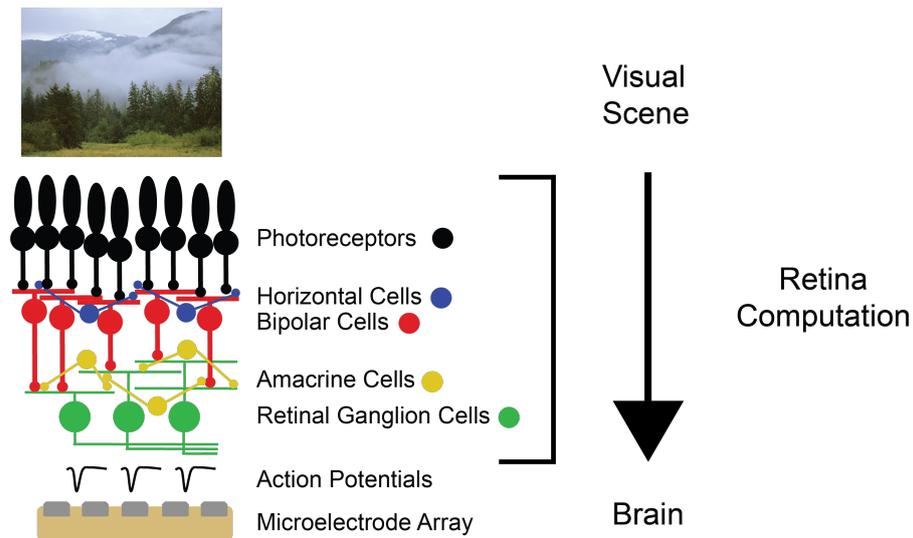
"By seeking and blundering we learn."
—Johann Wolfgang von Goethe

INTRODUCTION

An interesting thought experiment is to visualize yourself spending a normal day without having access to information from our eyes. Only one day of being blind. It is quite staggering to see that the list of activities that one would be able to achieve in normal times drastically decreases and that a large number of activities that would be intuitive and inexpensive in time literally become complicated or even impossible to overcome. Unfortunately, despite the handicap that it generates, blindness, even partial, is a disease that we still do not know how to treat because the functioning of the retina is still poorly understood. During my thesis, I had the opportunity to understand the functioning of the retina better and to collaborate within projects proposing solutions to the major problem of blindness.

Computer science and biology were the driving force behind most of my work and this chapter is presenting the context and the links between each part of this work. This introduction is giving a global overview to this Ph.D. For a detailed presentation of the scientific context, reader should be referred to the introduction section of each chapter. Each chapter is providing a comprehensive state-of-the-art in which my work is related to its field of application. The last chapter is dedicated to a discussion around the developed tools and the conclusions that this Ph.D. brings to event-based computation and to our understanding of biology.

The retina is a hierarchical arrangement of neurons that convert incoming light information into understandable electrical information for the brain, which it receives via the optic nerve. This conversion chain is particularly complex but could be seen as a three layers system. Photoreceptors compose the first layer and for catching the light, cones are sensitive to color spectrum distributions while rods allow achromatic vision in low light conditions. The second layer is responsible for the distribution of the information to the third layer and involves two types of cell. Horizontal cells regulate the light information by inhibiting groups of photoreceptors while bipolar cells measure the increase (“on” cells) or decrease (“off” cells) of the light caught by groups of photoreceptors and propagate the information using threshold crossing. There are roughly two distinct forms of horizontal cells and ten, for the bipolar cells. The third layer is responsible for sending the information to the brain and involves two types of cell, the amacrine and the ganglion cells. Amacrine cells are connected to groups of bipolar cells and are similar to horizontal cells as they regulate the signal of their input while ganglion cells are connected to



Borrowed from <https://www.bsse.ethz.ch/bel/research/electrophysiology-and-neuroscience/retinal-investigations.html>

optical nerve and are the output of the retina. There are roughly thirty distinct forms of amacrine cells and ten for ganglion cells. This chain is especially interesting with regards to its output signal, because the retina does not send images to the brain. The retina outputs action potentials, also called spikes, which are short burst of electrical signals in the order of μs .

Spikes are not triggered at fixed time intervals, but asynchronously, each cell being independent of each other. By generating information only when a change occurs locally in the light signal coming from the scene, the retina is able to effectively compress the visual information, leading to low latency in the feedback loop to a visual stimulus. We can however observe a remaining activity in cells even though no light is perceived. This phenomenon, called spontaneous activity, related to how cells are producing action potentials and shows that information should be driven by homogeneous groups instead of isolated cells.

Silicon retina sensors were developed initially designed to replicate how the biological retina converts light into an understandable electrical signal for the brain. We operate on these sensors and associated platforms to find methods able to mimic and explain how the brain's process is working and reach the brain's cost-efficiency to solve daily tasks. However algorithms build to be biologically plausible have to be biologically compatible by being used on real-world data. Using an algorithmic approach taking into account the asynchronous nature of the information and based on an event-based visual flow developed in the lab, I spent the first months of my thesis estimating how it was possible to reconstruct the orientation of an object in the scene by only relying on spikes and to estimate how far this method was efficient and less expensive in computation time compared to integrative approaches based on traditional vision algorithms. During

this study we sought to highlight that the use of specific cell types is able to increase the accuracy of the result, which led us to set up a spikesorting solution. Although several algorithms exist with relative efficiency to differentiate the information coming from each of the cells, no method was able to solve this problem on the fly with reasonable reliability. These methods use post-processing algorithms requiring heavy and long computation and often require the intervention of human operator. This process prevents us from designing close loop experiments capable of adapting stimulation in real time during the experiment. I worked for more than a year on an approach closer to the raw information of the retina in order to answer this problem. The stated objective of this work is to lead to the implementation of a chip that can start from basic information and provide real-time, latency-free interpretation of the scene.

In collaboration with the *Sight Again* project to use retinal implants to substitute the damaged photoreceptor layer due to *Retinis Pigmentosa* (the loss of peripheral vision) or *Dry Age-Related Macular Degeneration* (the loss of vision in the center of the visual field), I treat data to measure the capabilities of the implants: number of cells activated, type of cells, axons' propagation speed and cells' firing rate. Thanks to the treatment of the ex-vivo tests this study was able to show the relevance of its approach and thus to set up an in-vivo experiment on non-human primates which reinforced those preliminary results. Despite the complexity of obtaining a large number of non-human primates for experiments, those results have contributed to substantial progress, allowing the study to currently undergo first clinical trials on humans.

During the third year of my thesis, an interaction with specialists in psychophysics lead to the discovery of a problem previously unknown to me: the characterization of color blindness among individuals. Approximately 8% of men and 0.5% of women of Northern European descent are color blind, however this disease contains a large number of different malformations in the distribution of cells responsible for color detection that does not require the same correction. Although there are already methods to characterize color blindness, they are highly subjective and depend solely on the patient's estimation. In order to provide a robust solution, our study implemented a light spectrum scanning that provided a relevant measurement of the pupil response. The pupil is able to perform dilation and contraction to regulate incoming light using two different nerve marks, the variations of which characterized the capability of the photoreceptors to perceive light. Looking at the subjects' pupil response, our study aims to highlighting a fast, accurate and fully automated characterization of the perception of the visual spectrum of the subject to provide an appropriate correction. Taking into account the current progress of the project, this document will focus on the progress of data processing

for the first part of the healthy population that was recruited for the study because of the colorblind population has not yet been recruited for testing. Those results aim to validate the relevance of stimulation via the characterization it brings from a healthy subject, regards the information resulting from each part of the stimulus. Preliminary results of this study have highlighted the biometric aspect of the information extracted from the subjects' pupil response due to the intrinsic distribution of photoreceptors for each individual. Although related to the initial field of study by methodology, this new aspect opens up a second field of study for this project and this document will introduce in detail the major challenges of biometrics as well as the part that such a device can play in biometric state of the art.

PUBLICATIONS AND PATENTS

JOURNALS

P.-H. Prévot, K. Gehere, H. Akolkar, M. A. Khoei, K. Blaize, O. Oubari, P. Daye, M. Lanoë, S. Dalouz, P. Langlois, E. Esposito, V. Forster, E. Dubus, N. Wattiez, E. Brazhnikova, C. Nouvel-Jaillard, Y. LeMer, J. Demilly, C.-M. Fovet, P. Hantraye, M. Weissenburger, H. Lorach, E. Bouillet, M. Deterre, R. Hornig, G. Buc, J.-A. Sahel, G. Chenegros, P. Pouget, R. Benosman, S. Picaud. *"Visual perception of single pixel stimulations with aphotovoltaic subretinal prosthesis in non-human primates"*, Natural Biomedical Engineering, in review.

G. Haessig[#], K. Gehere[#], R. Benosman, *"Spikes decoding spikes : an event-based framework for real-time unsupervised spike sorting"*, In preparation.

[#] These authors contributed equally.

K. Gehere and R. Benosman, *"Event-based approach to extract motion in retina"*, In preparation.

PATENTS

J. Lorenceau, K. Gehere. and R. Benosman, *"Dilation and contraction of the pupil as a reliable biometric"*. In preparation.

ACKNOWLEDGMENTS

Au cours de ces trois années de thèse, j'ai pu rencontrer et échanger avec des personnes particulièrement brillantes et qui pour quelques unes sussitent pour moi une certaine forme d'admiration.

J'aimerais prendre le temps ici de remercier ces gens qui m'ont aidé. En premier lieu j'aimerais remercier mon maître de thèse Ryad Benosman pour m'avoir fait découvrir un nouveau champs d'activité en traitement du signal capable de poser des problématiques particulièrement intéressantes. Et plus que ça de me l'avoir fait découvrir au sein de l'équipe S9. Je ne sais pas exactement comment l'exprimer clairement, mais évoluer au sein des membres de cette équipe a été un gain inestimable. Aussi bien en terme d'ouverture d'esprit, que dans l'ouverture aux autres de manière générale, l'approche de la résolution d'un problème et notamment sur la manière de transmettre la connaissance via l'enseignement.

J'aimerais également rémercier Serge Picaud, responsable de l'équipe S7, pour l'opportunité qu'il m'a offert d'être formé au sein de son équipe et pour les échanges scientifiques que nous avons vu. Grâce à ces échanges, j'ai énormément appris sur les difficultés pour deux domaines de communiquer de par la simple barrière du vocabulaire technique qui peut renvoyer via les mêmes mots vers des concepts différents.

J'aimerais également remercier les membres de l'équipe S7 qui m'ont accueilli et aidé et plus particulièrement Sami Dalouz qui m'a formé aux expériences ex-vivo dont j'envie beaucoup l'habilité et Paul-Henri Prévot pour son incroyable gentillesse et sa disponibilité.

Pour finir j'aimerais particulièrement remercier Himanshu Akolkar pour les précieux conseils qu'il m'a donné et que j'ai tenté de suivre, en espérant ne pas avoir été un trop mauvais padawan.

Ps: j'aimerais ajouter mes remerciements à mon actuel voisin Omar Oubari pour absolument rien, uniquement parce qu'il me l'a demandé et que comme chacun sait je possède un sens de l'humour particulier.

CONTENTS

I VISUAL FLOW ON BIOLOGICAL RETINA	
1	EVENT-BASED APPROACH TO EXTRACT MOTION IN RETINA 3
1.1	Introduction 3
1.2	Methods and Materials 4
1.2.1	Recording 4
1.2.2	Algorithm 5
1.2.3	Rejection 6
1.3	Results 7
1.4	Discussion 9
II SPIKE SORTING FROM BENCHMARK TO BIOLOGICAL RETINA	
2	SPIKES DECODING SPIKES 13
2.1	Introduction 13
2.2	Methods and Materials 14
2.2.1	Event generation 15
2.2.2	Feature extraction and clustering 15
2.2.3	Classification 16
2.3	Results 16
2.3.1	Model description 16
2.3.2	Benchmarking 19
2.4	Discussion 24
3	VISUAL PERCEPTION WITH A PHOTOVOLTAIC SUBRETINAL PROSTHESIS 27
3.1	Introduction 27
3.2	Methods and Materials 28
3.2.1	Animal models 28
3.2.2	Animal preparation 28
3.2.3	Retinal imaging 29
3.2.4	Multifocal ERG 29
3.2.5	Stimulation system 29
3.2.6	Behavior 29
3.2.7	Retinal tissues 30
3.2.8	Retinal explant culture 30
3.2.9	Multi electrode array recordings 30
3.2.10	Histology 31
3.2.11	Human retina <i>ex vivo</i> recordings 32
3.3	Results 32
3.3.1	Ex-vivo 32
3.3.2	In-vivo 33
3.4	Discussion 36
3.5	Supplementary Figures 38

III PUPIL MODEL AS A BIOMETRIC

4 DILATION AND CONTRACTION OF THE PUPIL AS A RELI-
ABLE BIOMETRIC 47

4.1 Introduction 47

4.2 Methods and Materials 48

4.2.1 Light stimuli 48

4.2.2 Database 49

4.2.3 Algorithm 49

4.2.4 Classification 52

4.3 Results 52

4.4 Discussion 54

Conclusion

IV APPENDIX

A EVENT-BASED WATERMARKING 67

A.1 Introduction 67

A.2 Hypothesis 68

A.3 Method 68

A.4 Known weaknesses 71

BIBLIOGRAPHY 73

Part I

VISUAL FLOW ON BIOLOGICAL RETINA

EVENT-BASED APPROACH TO EXTRACT MOTION IN RETINA

Event based sensors are used in computer science to perform neuro-morphic computation, algorithms inspired by the way nature works to solve a problem. We propose in this study to demonstrate the relevance of event based silicon retina sensors to develop algorithms able to directly be applied on real biological retina data, by using a key algorithm of computer science as the optical flow. From an algorithm developed on the model we achieve robust and accurate results on real data with an average error of 3.5° of angular reconstruction.

1.1 INTRODUCTION

The characterization of motion is essential to understand the visual scene through the interaction between the objects present in the scene. Motions represent an important part of the dynamic of the scene and allow a mechanism of prediction to safely move. Motion is fairly a low level information, allowing to build a more complex representation of what is taking place in the visual scene. Tasks that appear to us as elementary, such as objects segmentation, comparisons and links between them, are always preceded by motion extraction. In 1986, Georgopoulos, Schwartz, and Kettner, 1986 experimented the relation between a visual stimuli and the processing chain that leads to a reaction. This study shows that the activity in the primates' motor cortex could be locally correlated to the needed direction to reach the target. Meaning the brain is able to select an adequate population of cells according to a visual target. Beyond the fact that this conclusion seems natural, it appears that the brain is so dependent on motion that several parts are dedicated to perform different motion tasks. In consequence, motion information that come from the retina needs to be as efficient as possible. Several studies proposed methods to reconstruct information from moving light stimuli, using the output signal of the retina. Berry et al., 1999 showed variations of delay in "off" cells when responding to moving stimuli. These variations lead to a start of prediction of the information and switching the role of the cell, from spatial gradient extraction to expectation. They also lead to an error of position, while allowing to prevent a delay of reaction in the rest of the chain. This effect seems to only be visible on repeated stimuli, as a mechanism of adaptation, according to the study of Marre et al., 2015 which succeeded to reconstruct position information of random moving bar according to the activity. According to the studies

of Butts et al., 2007 and Akolkar et al., 2015, the time precision in the retinal code contains an important amount of information and this time precision allows the extraction of robust features from the input stimulus. Following the work of Chichilnisky and Kalmar, 2003, the study of Frechette et al., 2005 proposed a method to reconstruct the velocity of the input stimulus, using correlation between spikes trains (Reichardt, 1961).

However this method required prior knowledge of the size of a trial and used a fixed time-window that limits the range of speeds that can be detected.

In this study, we proposed a solution to extract the orientation and the speed of an object, by using only the most recent local retinas' activity. The method is based on an event-based visual flow algorithm introduced by Benosman et al., 2014 and led to an implementation on a neuromorphic hardware (Giulioni et al., 2016). We show with this approach, a robust on-line algorithm, which does not need to be trained or fed with knowledge from input stimuli, but only constrained by the receptive field mapping.

1.2 METHODS AND MATERIALS

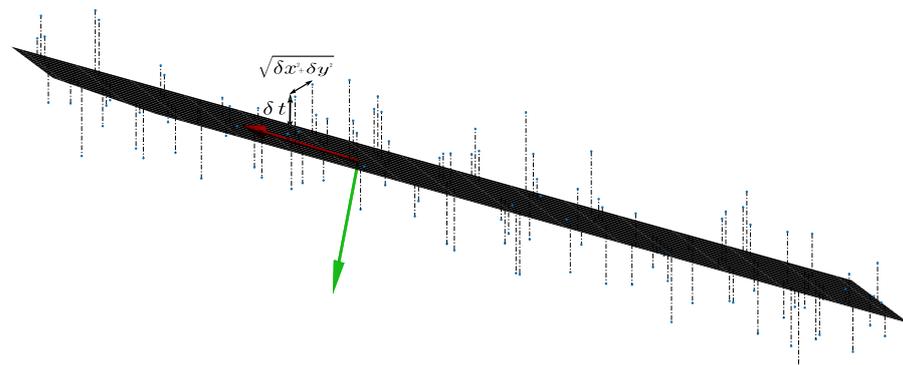


Figure 1.1 – The plane fitting elicits the plane that minimizes the distance between each event of the cloud. For each event projected on this plane, if the ratio between the distance of the nearest event and the error of time is greater than the estimated speed of the plane, the event is rejected. The estimated speed is given by the projection of normal, $\frac{1}{\sqrt{\alpha^2 + \beta^2}}$.

1.2.1 Recording

Data come from *ex-vivo* adult Long-Evans 8 weeks male rat retina. The recording is performed using multi-electrodes array (MEA) composed by 252 electrodes (16 by 16 without corner electrodes). The size of each electrode is $10\mu\text{m}$ in diameter and $60\mu\text{m}$ center-to-center. The

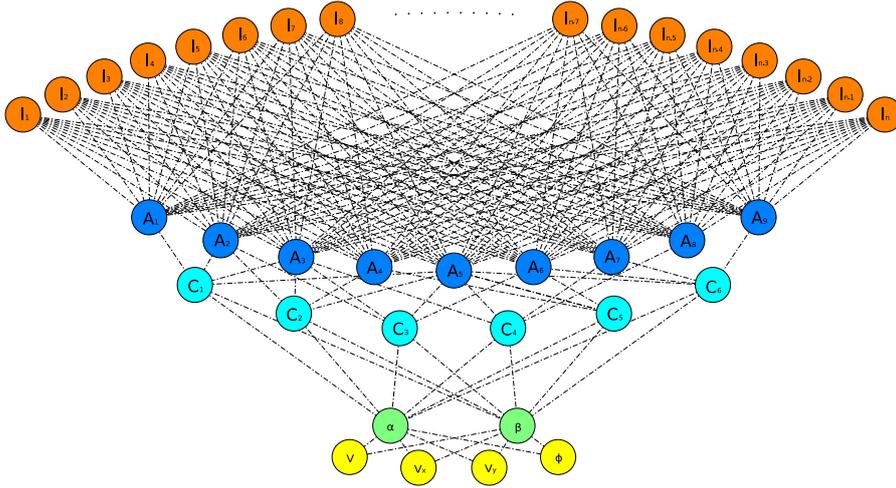


Figure 1.2 – The visual flow algorithm explicit as a graph of subprocesses levels (See 1.2 for details). Level 1 represents for each node the input information coming from cells. The level 1 is fully connected to the level 2. Level 2 represents each combination of scalar product of information coming from level 1. Level 3 represents operations as $x_1 - \frac{x_2 x_3}{x_4}$. Level 4 represents operations as $\frac{x_1 x_2 - x_3 x_4}{x_1 x_5 - x_6 x_4}$.

output signal is analog converted to digital with a sampling rate of 20kHz and 16 unsigned bits precision.

After sorting, 93 cells were found and 62 seemed to react to the considered light stimulus. All the cells are divided into 7 cell types and the reacting pool is composed by four of them. Each cell is considered as an independent event-based sensor that fired action potentials when an event occurred around it.

The light stimulus consists in a moving black and white grating at constant speed ($1000\mu\text{m}\cdot\text{s}^{-1}$) with 8 different orientations which covers all the retina. A bar is $333\mu\text{m}$ width on the retina. The stimulus is displayed at 60Hz for an elementary displacement of $16.65\mu\text{m}$.

1.2.2 Algorithm

Let's assume an event as a triplet as presented in 4.1:

$$ev_k = \{t_k, x_k, y_k\} \quad (1.1)$$

Where for each ev_k , t_k represents the timestamp of the event and (x_k, y_k) its spatial coordinates. These three elementary data allow an

efficient characterization of the local and global dynamic of an event stream and allow the study of its evolution.

$$\begin{cases} z_1 = \alpha x_1 + \beta y_1 + \gamma \\ z_2 = \alpha x_2 + \beta y_2 + \gamma \\ \vdots \\ z_n = \alpha x_n + \beta y_n + \gamma \end{cases} \quad (1.2)$$

$$\begin{pmatrix} \sum_i w_i x_i^2 & \sum_i w_i x_i y_i & \sum_i w_i x_i \\ \sum_i w_i x_i y_i & \sum_i w_i y_i^2 & \sum_i w_i y_i \\ \sum_i w_i x_i & \sum_i w_i y_i & \sum_i w_i \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \sum_i w_i z_i x_i \\ \sum_i w_i z_i y_i \\ \sum_i w_i z_i \end{pmatrix} \quad (1.3)$$

Movement is a variation of the position across time of an object in the scene. When an event stream is describing a movement, the close neighborhood of each event could be seen as a noisy measurement of the local motion. Thus, computing the plane of each local cloud of event allows the extraction of the local motion's features 1.1. By describing The problem as the following 1.2, it becomes possible to extract the optimum coefficients of each plan by minimizing 1.3. Where α , β and γ are the coordinates of the plan and w_i a weight factor on the reliability of each event. This problem is solved by inverting a 3x3 matrix and the figure ?? shows the connection between the cells to achieve this inversion. To decrease the impact of noisy events on the reconstruction, two approaches have been considered. The first one was to build the average plane by iteratively remove the most distant events and rebuild the plane on the rest. The second one was to use the weight factor w_i of 1.3 to add a constraint function to the cloud of events. To constrain the space, a squared invert function is applied, meaning that the further the events are the less they impact the reconstruction, while a delay function is applied to the time (1.4).

$$\begin{cases} W_{o_i} = \frac{(t_{ref} - t_i) x_{ref_i}^2}{x_i^2} \\ W_{r_i} = \frac{x_{ref_i}^2}{(t_{ref} - t_i) x_i^2} \end{cases} \quad (1.4)$$

Where W_o increases the weight on old events and W_r on recent events. t_{ref} is the timestamp of the last event and x_{ref_i} a weight that represent the minimum confident radius of neighborhood according to the size of the receptive field for each cell i .

1.2.3 Rejection

To establish if a generated plane is reliable or not, a threshold method is proposed. Using the coefficients of the plane, we are able to

project the expected timestamps according to the used neighborhood and compare it to the original timestamp. According to the computed speed, if the delay between the original and projected timestamp is bigger than the delay to cross the nearest point in the neighborhood, the timestamp is rejected (See 1.1). A plan is accepted under a threshold condition based on the delay. Presented results use a threshold value at fifty percent, meaning that at least half of the used timestamp needs to be accepted. We choose this flexible threshold as a compromise between reliability and responsiveness.

$$\text{Bool}_i = |\alpha x_i + \beta y_i + \gamma - t_i| \leq \Delta x_i \sqrt{\alpha^2 + \beta^2} \quad (1.5)$$

1.3 RESULTS

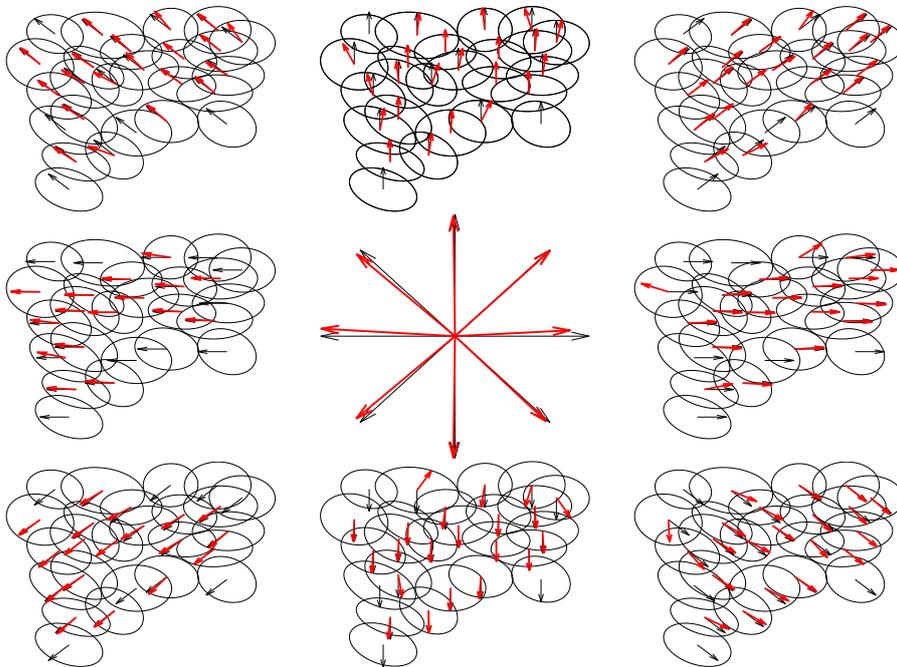


Figure 1.3 – The reconstruction of the eight orientations projected on the retina. On all these figures the dark arrows represent the real orientation and the red arrow the estimated one.

The visual flow results presented in this study come from an adaptation of the method proposed in Benosman et al., 2014. Eight computing rules are compared aiming to find the most robust ones, but also the most trustable biologically speaking. Events are selected according to a fixed region of interest and a time window. The first column of 1.1, shows the reconstruction results using the output of the retina. The method performs well to reconstruct half of the angles. Due to the complex geometry of the grid and small temporal window the error of reconstruction is quite important.

Table 1.1 – This table presented the reconstruction results of the orientations according to all conditions tested in this study. The position projected in a matrix combined with an event temporal window (SpTn), the real positions according an elliptic neighborhood combined with an event temporal window (SfTn), then combined with only the last event at each position (SfTf), and two different weights, increased on old events (SfTwo) and increased on recent events (SfTwr). A weight condition is also tested on the positions with a $\frac{1}{x^2}$ function (SwTf, SwTwo, SwTwr). The event temporal window is a complex solution to implement with retinal signal due to spike burst. The size of the window depends of the signal. The increasing of weight on recent event showed instabilities. On SfTwr, the error was twice bigger and more on several angles and the significant deviation. On SwTwr, the deviation is low but 10 times less planes were elicited. This shows that these solutions are not sustainable. The same problem could be seen on SwTf, with a low deviation leads to few planes elicited. This one shows that applying a constraint on the space is not enough if no constraints are applied on delays.

Error	SpTn	SfTn	SfTf	SwTf	SfTwo	SwTwo	SfTwr	SwTwr
0°	5.89°	7.21°	3.73°	3.30°	4.23°	3.05°	3.61°	2.94°
45°	12.11°	7.41°	5.35°	2.72°	5.10°	3.27°	6.31°	4.18°
90°	11.55°	10.49°	3.10°	5.10°	4.61°	2.23°	2.24°	1.64°
135°	8.83°	7.19°	5.85°	3.70°	6.88°	5.17°	4.23°	2.64°
180°	20.10°	6.89°	4.33°	3.91°	5.42°	3.75°	3.92°	5.12°
225°	10.50°	6.40°	5.60°	3.81°	4.56°	3.94°	4.06°	4.03°
270°	7.76°	6.13°	5.62°	5.15°	5.50°	3.94°	4.61°	2.98°
315°	9.27°	8.24°	4.81°	2.28°	4.92°	3.42°	15.65°	10.52°
\bar{X}	10.75°	7.56°	4.51°	3.54°	5.11°	3.49°	4.43°	3.38°
\tilde{X}	9.89°	6.48°	3.93°	3.32°	4.46°	3.03°	3.43°	2.92°
σ_X	4.01°	7.52°	4.10°	1.66°	5.22°	2.78°	7.55°	1.73°
$E_{v_{out}}$		9.4%	5.0%	1.6%	8.0%	7.3%	4.4%	0.5%

To improve the results, the method switched to a space invariant version and the use of events' window. Instead of extracting the N closest neighbors in a circle range, neighbors are elicited in an ellipse according the covariance of the space. The event window uses a decay to elicit events that are recent enough and select in these events the N most recent. As showed on the second column of 1.1, using this method decreased the error and made it more homogeneous. But the size of the event window remains a problem that highly depends on the signal. This motivated the restriction of the use of only the

last event for each position (column three, 1.1). Despite good results, this method leads to an increase of rejected planes, due to an equal weight on each point to perform the reconstruction. Unstable recent events or too old events are able to unbalance the computation of the plane. In 1.3, the weight factor w_i is able to increase or decrease the influence of each selected event in the reconstruction. We compared five weighting conditions. The column four of 1.1 used a weight based on neighbor's distance and no weight on time. Results are as stable as using equal contributions with a low variance, but generated output events decrease significantly. Both column five and six of 1.1 show the impact of adding a weight on time, increased on old events while column seven and eight show the impact of increasing the weight on recent events. Columns six and eight combined the time weighting with the neighborhood based weighting presented on column four, while columns five and seven presented results using only the time weighting. As expected, the weight on distance decreases the variance, however old events seem to be more stable than recent ones to build planes regarding to the number of generated output events that drop significantly. The error for the angle 315° increases significantly while using a weight based on recent events, however this could be specific to the data regarding to the results of others methods.

1.4 DISCUSSION

In this study we built a robust and accurate angular estimation algorithm using biological data with an average error of 3.5° . We show the relevance of using an event based silicone retina sensor as a reliable model to develop vision algorithm able to be applied on real biological retina data. From the biological retina data that we had, the accuracy of the angular reconstruction was the only part able to be analyzed due to the objects are moving at single-speed. The presented method shows that the time inclination of a cloud of events carries the movement information. However, in order to obtain a unique solution estimating this inclination, we apply a constraint on the normal norm to this event cloud. Although the computed orientation remains not impacted, the speed distribution is fixed. The algorithm is able to estimate the speed ratio between two speeds, however a speed reference value is needed to re-scale the speed and compute an accurate absolute value.

Part II

SPIKE SORTING FROM BENCHMARK TO BIOLOGICAL RETINA

SPIKES DECODING SPIKES : AN EVENT-BASED FRAMEWORK FOR REAL-TIME UNSUPERVISED SPIKE SORTING

With the increase of multi-electrode array size, spike sorting algorithms are often overwhelmed by the quantity of data to process. It is then impossible to have meaningful results in real-time for applications such as closed-looped experiments. Furthermore, there is almost always the need of an external human operation to control and validate results. Here, we show that neuromorphic computation can yield to very good results in real-time unsupervised spike-sorting tasks. Our results demonstrate that considering the time as the most valuable information in signals helps extracting coherent information from noisy data. Comparison between the proposed method and state-of-the-art algorithms shows that event-driven computation allows to significantly reduce computation time while increasing efficiency, reaching up to 90% of recognition rate on real data. We anticipate this work to open new horizons for embeddable real-time devices for closed-loops applications and low-cost performance analysis of in-vivo data.

2.1 INTRODUCTION

The use of conventional sampling methods implies a stroboscopic acquisition of information (unknown to biological systems) at a low sampling frequency. Thus, this method is unable to describe the full dynamics of a given signal. Recent neuroscience findings show that this temporal precision is found in subcortical areas, like the lateral geniculate nucleus (LGN) (Liu et al., 2001; Reinagel and Reid, 2000) and the visual cortex (Mainen and Sejnowski, 1995). The last decade has seen a paradigm shift in neural coding. It is now widely accepted that precise timing of spikes open new profound implications on the nature of neural computation (Rieke et al., 1999; Maass, 1999). The information encoded in the precise timing of spikes allows neurons to perform computation with a single spike per neuron (Thorpe, 1990). Initially supported by theoretical studies (Thorpe, Delorme, and VanRullen, 2001), this hypothesis has been later confirmed by experimental investigations (Johansson and I. Birznieks, 2004; R.S., Panzeri, and Diamond, 2001).

Here, we present a novel approach to spike sorting, which is compatible with ultra low latency and low power neuromorphic hardware

technologies (Chicca et al., 2014). In particular, we exploit advances made in both mixed signal Analog/Digital VLSI technology and computational neuroscience with brain-inspired spiking neural processing devices to build sophisticated real-time event-based processing systems (Neftci, 2013; Indiveri, Corradi, and Qiao, 2015; Serrano-Gotarredona, 2009). We show how precise timing of spikes allows the introduction of a novel, fast and reliable biologically plausible solution to the problem of spike sorting directly from the high temporal properties of spikes.

The developed architecture is particularly adapted on a variety of existing neuromorphic spiking chips such as the SpiNNaker (Furber et al., 2014), TrueNorth (Merolla et al., 2014) or LOIHI (Davies et al., 2018) neural chips. More specific neuromorphic hardware, such as the 256 neurons ROLLS chip (Qiao et al., 2015), can also be used.

State of the art spike sorting methods relies on prior knowledge of the dynamic of the signal for discriminant features extraction. Spike sorting algorithms mainly rely on three steps : spike detection, feature extraction and classification. The first step often employs an automatic spike detection method (threshold detection in most of the case (Rossant et al., 2015)). In the second step, a set of features is computed, *for example* using mixture of Gaussian kernels (Khadir et Al.Kadir, Goodman, and Harris, 2014), wavelets transformation (Quiroga, Nadasdy, and Ben-Shaul, 2004) or PCAs (Rossant et al., 2015). Then, the extracted features are assigned to cell types by learning algorithms, such as (un)supervised clustering.

This work aims to show that it is possible to extract and classify, in real time, the dynamic of the signal. Previous work on event-based vision algorithm (Lagorce et al., 2016) shown that the event-based paradigm shift yield to a better conceptualization of each new piece of information, with respect to past activity, both in spatial an temporal neighborhood. In this paper, we will first show a neuromorphic, event-driven approach for features extraction and spike sorting tasks, allowing a real-time pipeline.

The performances of the algorithm were tested both with a benchmark introduced by Wild et al., 2012, in-vivo (Henze et al., 2009) and ex-vivo recordings and compared to state-of-the-art algorithms (Rossant et al., 2015; Kadir, Goodman, and Harris, 2014; Quiroga, Nadasdy, and Ben-Shaul, 2004).

2.2 METHODS AND MATERIALS

Here we detail how, from an analog signal, the proposed algorithm generates events, extracts features, clusters them in a hierarchical model (2.2) and then classifies unknown examples.

2.2.1 Event generation

The sampled signal (48 kHz for the benchmark, 20 kHz for ex-vivo recordings) is first sent into a spike generator to (artificially) convert it in a event-based signal. Then, the sampled signal is filtered by a high-pass, 3rd order, Infinite Impulse Response (IIR) filter for DC filtering and low frequencies noise rejection. Then, the filtered signal is compared to a bench of levels. Each time the signal crosses a level, a spike is triggered, both positive if the signal is increasing while crossing the level, or negative if the signal is decreasing. A linear interpolation between two samples is used for better time precision. Each event $ev(i)$ can be seen as the following triplet :

$$ev(i) = \{t_i, c_i, p_i\}, i \in \mathbb{N} \quad (2.1)$$

where t_i is the timestamp of the event, c_i its channel (the crossed level) and p_i its polarity, $p_i \in \{-1, 1\}$ (*i.e.* positive or negative).

2.2.2 Feature extraction and clustering

A spatio-temporal point cloud is formed with these events, representing the spike spatial distribution and dynamic behavior. Because this point cloud contains information about the spikes dynamic and amplitude, we introduced the event context S_i of the event $ev(i)$ to convey information about surrounding activity just before time t_i . Let $N(c_i)$ be a windowing neighborhood with length $2R + 1$ (in channels) around the event's channel c_i :

$$N(c_i) = \{c_i + c \mid \text{abs}(c) \leq R\} \quad (2.2)$$

Then, this event's context S_i is defined as :

$$S_i = \exp\left(\frac{t_i - t_{\text{ref}}^{N(c_i)}}{\tau}\right) \quad (2.3)$$

where t_{ref}^c is the last spike time at the c^{th} channel.

When a context has been computed, it is compared to a bank of context, or centers. The most-closely matching center will then generate an output event. First, we do have a set of N initial contexts $m_k, k \in \llbracket 1, N \rrbracket$ where m_k takes the same form as S_i in 2.3. These contexts are initialized with the first N incoming contexts. More formally,

$$m_k = S_k \quad k \in \llbracket 1, N \rrbracket$$

Then, they are clustered using an online Iterative Inverted Weight Kmeans (*IIWKmeans*, Barbakh and Fyfe, 2008). This clustering method was preferred to KNNs (Cover and Hart, 1967) or KMeans (MacQueen,

1967) because of its independence towards the initial conditions and its ability to work online. For each incoming context S , we define its nearest center m_{k^*} :

$$m_{k^*} = \underset{k}{\operatorname{argmin}}(\|S - m_k\|) \quad (2.4)$$

where m_k is the k^{th} center. The update rule is then shown by:

$$m_{k^*}(t_s + 1) = m_{k^*} - \zeta a_{ik^*}(S - m_{k^*}) \quad (2.5)$$

with:

$$a_{ik^*} = -(n+1)\|x_i - m_{k^*}\|^{n-1} - n\|x_i - m_{k^*}\|^{n-2} \sum_{j \neq k^*} \|S - m_{k^*}\| \quad (2.6)$$

ζ and n are two learning parameters able to constrain the update rule. This process is summarized in algorithm 1.

Once the learning is over (*i.e.* all the centers converged), each context can be associated to a particular center m_k . So, the input stream of events is transformed into an output stream of centers activation :

$$ev_{\text{out}} = [c_i, t_i, k_i] \quad (2.7)$$

where k_i is the index of the matching center m_{k^*} .

At this point, for noise (isolated context) rejection, it is possible to implement context rejection based on some thresholding among the distances and so prevent emitting a new event if the match between the context and the center was not strong enough.

2.2.3 Classification

The recognition by itself is done by comparing online signatures to the trained ones. The distance between two histograms used here is the Bhattacharyya (Bhattacharyya, 1946) distance :

$$d^b(H_1, H_2) = -\log \sum_i \sqrt{\frac{H_1(i)}{\operatorname{card}(H_1)} \frac{H_2(i)}{\operatorname{card}(H_2)}}$$

2.3 RESULTS

2.3.1 Model description

We will here give an overview of the proposed algorithm (see 2.2). Detailed explanations and equations can be found in the Method section.

When the input signal changes, our algorithm asynchronously generates events forming a spatio-temporal point cloud representing the

Algorithm 1 Hots online update rule for one layer

Ensure: $\{t_i ; p_i ; l_i\}$ as the timestamp, polarity and level crossed of each new incoming event

procedure HOTS(t_i, p_i, l_i) ▷ Receive a new incoming event

Compute current context S_i

$t_{past}(p_i, l_i) \leftarrow t_i$ ▷ update past

$S_i \leftarrow \exp \frac{-(t_i - t_{past})}{\tau}$ ▷ compute around l_i

Compute distances

for each centers M_k **do**

$d_k \leftarrow -\log \left(\sum_k \sqrt{\frac{M_k \cdot S_i}{\text{card}(M_k) \cdot \text{card}(S_i)}}} \right)$

end for

Update nearest center

$k^* \leftarrow \underset{k}{\text{argmin}}(d_k)$

$M_{k^*} \leftarrow \text{update following eq (2.5)}$

end procedure

signals' dynamical behavior. The sampled signal is first sent into a spike generator to (artificially) convert it in a event-based signal. As in silicon retinas (Posch, Matolin, and Wohlgenannt, 2011), the variation of the signal triggers an event generation. In a more detailed approach, the sampled signal is first filtered by a high-pass, 3rd order, Infinite Impulse Response (IIR) filter for DC filtering and low frequencies noise rejection. Then, the filtered signal is compared to a bank of levels. Each time the signal crosses a level, a spike is triggered, both positive if the signal is increasing while crossing the level, or negative if the signal is decreasing (2.1a). A linear interpolation between two samples is used for better time precision.

Each event $ev(i)$ can be seen as the following triplet :

$$ev(i) = \{t_i, c_i, p_i\}, i \in \mathbb{N} \quad (2.8)$$

where t_i is the timestamp of the event, c_i its channel (the crossed level) and p_i its polarity, $p_i \in \{-1, 1\}$ (*i.e.* positive or negative).

Considering a stream of events, we can associate for each event a description of its spatio-temporal neighborhood. This description will be referred as context, generated by convolving an exponential kernel (2.1b) to the most recent activity on the surrounding of the incoming event (2.1c). These contexts are then clustered using an online iterative clustering method (Barbakh and Fyfe, 2008). Once this learning phase done (*i.e.* the centers are determined and represent the dynamics of the learned signal), we extract the response of the network for different spikes. Then, events from unknown spikes are fed in the network, and their signatures compared to the learned ones for classification.

Input data is presented to the network. Contexts are built with the convolution with an exponential kernel of time constant τ_1 and considering a spatial neighborhood of sidelength $(2R_1 + 1)$. Then, these contexts are clustered into N_1 centers. When a cluster is matched,

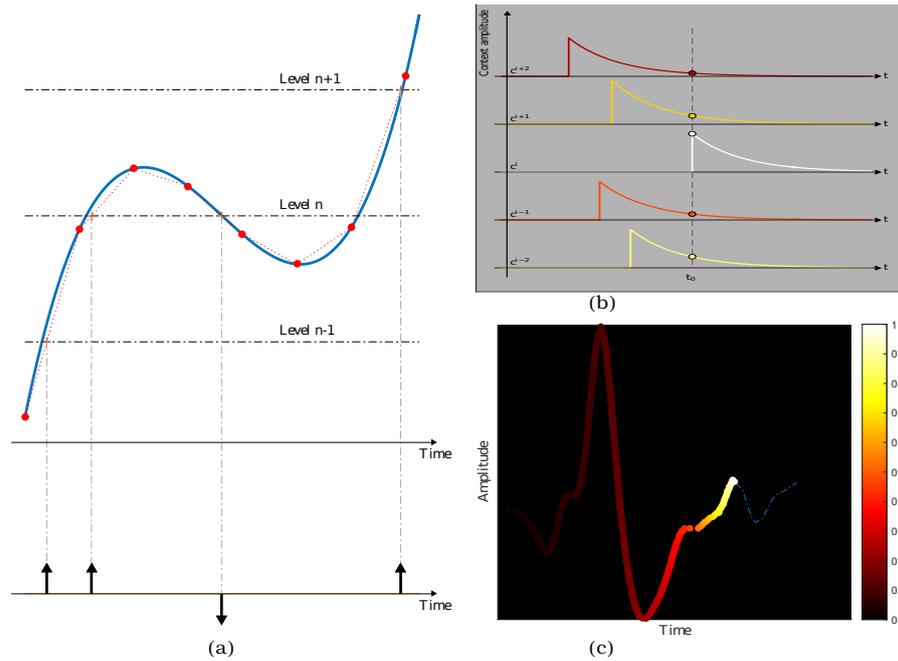


Figure 2.1 – Generation of a context. (a) Level crossing for spike generation. The sampled signal (red dots) is interpolated. Each time the linearly interpolated signal (dashed red line) crosses a level (horizontal dark dashed-dotted lines), a spike is generated, either positive or negative depending of the raising or falling behavior of the signal. (b) Generation of a context. When a spike occurs at channel c_i , all the values of the exponential decays among a spatial neighborhood are taken ($i \in \llbracket i - r, i + r \rrbracket$, r being the current radius value) and are our base features. (c) Exponential time decay among a spike. The value of the exponential kernel is color-coded : brighter pixels represent higher values than darker ones.

an output event is produced, constituting the output of Layer 1. This output is of the same type as its input, as shown in 2.8 and 2.7. Thus, the same process applied to Layer 1 can be applied to layer 2 using different parameters for space-time integration (τ_2, R_2, N_2).

- As stated before, each layer l is only characterized by 3 parameters :
- τ_l , the time constant of the exponential kernel,
 - R_l , the size of the neighborhood,
 - N_l , the number of centers to be learned by the clustering algorithm.

The output of the last layer can be used as features for shape classification. The training of the recognition algorithm consists in two distinct steps. In the first one, learning data is presented in order to learn the centers computed as described in the previous section. Then, the same inputs are provided to the network and an histogram of centers activation (signatures) is built for each different class, using rather TSNE projection or offline KMeans clustering. These histograms

represent the number of centers activated during the presentation of the example, independently of its spatial position, and are discriminant enough to allow proper classification.

2.3.2 Benchmarking

The proposed algorithm was tested with a database of artificially generated data from simulating extracellular signals recorded with a single electrode. Nine real spikes were manually picked from extracellular tungsten micro-electrode recorded during Deep Brain Stimulation operation from the sub-thalamus nuclei. They are then corrupted with noise coming from a superimposition of activity of many distant neurons in the brain (Wild et al., 2012).

2.3.2.1 Noise level

The noise level n_l was defined as the reciprocal value to the signal-to-noise ratio (SNR) (Wild et al., 2012) :

$$n_l = \text{SNR}^{-1} = \left(\frac{A_{\text{signal}}}{A_{\text{noise}}} \right)^{-1}$$

where A_{signal} is the root mean square value from all the extracted spikes, and A_{noise} the one for the rest of the signal.

2.3.2.2 Performance rating

To compare our method to state of the art spike sorting algorithms, we used the Adjusted Mutual Information. This measure provides information about how good the clustering is. Unlike the recognition rate (trace of the confusion matrix divided by the sum of all the elements within the confusion matrix), this measure is adjusted by luck, meaning that a zero value stands for luck and a 1 value for a perfect classification. Details and properties can be found in (Vinh, Epps, and Bailey, 2010).

2.3.2.3 Results

In this section, we will present results of the proposed algorithm on the benchmarking database (Wild et al., 2012). This database contains 32 sets of 9 artificially generated spikes, corrupted with noise. The training of the algorithm was done with one full set of the intermediate noise-level. The learned centers were then used for building signatures and recognition for all the 31 remaining sets. The parameters used for this benchmark are summarized in table 2.1.

2.4a shows the response among the centers of the fifth and last layer. The dynamic of the signal is well captured and clusters respond to local patterns of the input signal. Using high dimensionality reduction

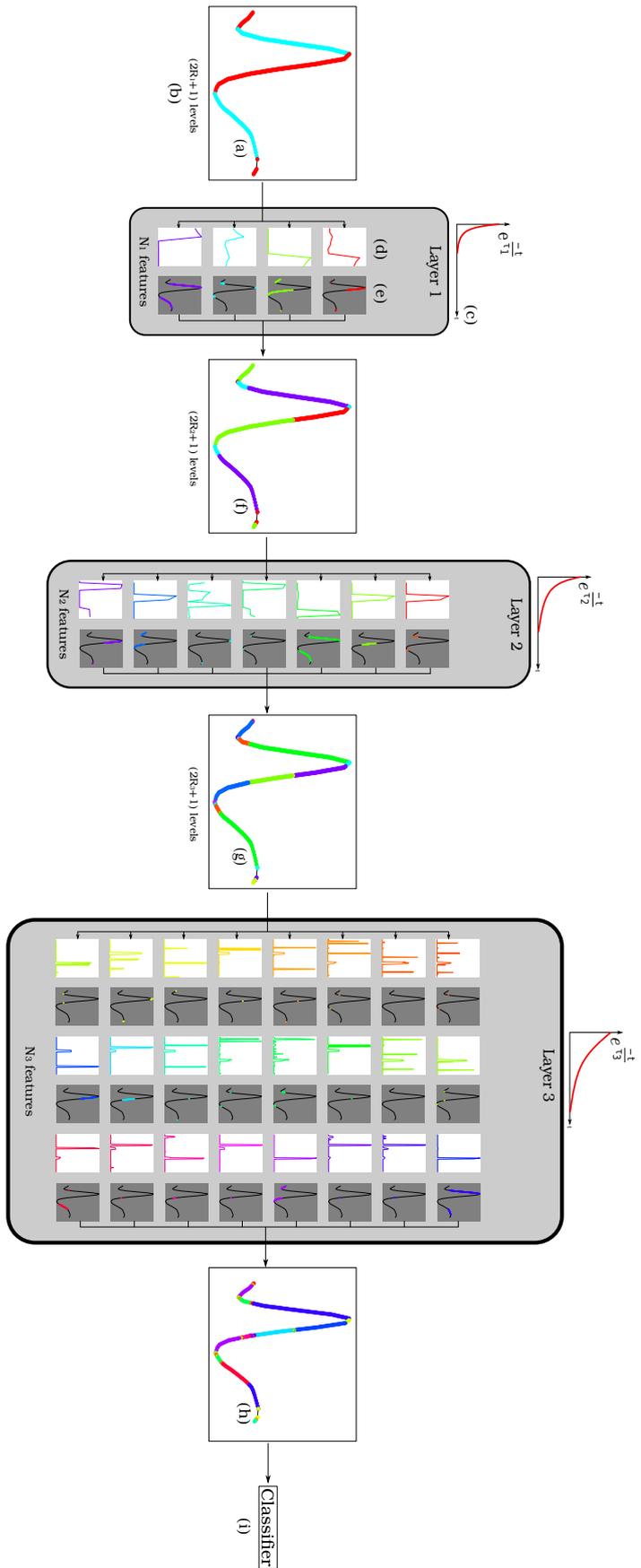


Figure 2.2 – This figure shows the structure of the proposed hierarchical algorithm. (a) The input signal is labeled by its polarity and sent to the first layer of the algorithm. (b) The radius of the considered neighborhood and (c) the exponential decay, building the context computed for each incoming event. (d) This incoming context is compared to a set of centers representing the N_1 features and fitting gradually the local behavior of the input signal. (e) These features are used to label each incoming event to build (f) the output signal for the layer 1. (f,g) By changing the size of the considered spatiotemporal context, the output of a layer could be used to feed the next layer and extract new features, that are combinations of the previous ones. By adapting, step by step, the size of the context of each incoming event, the algorithm combines, layers after layers, local features to extract global features. (i) Then a classifier is able to use the selectivity of these features to distinguish a signal from another. *Inspired from Lagorce et al., 2016*

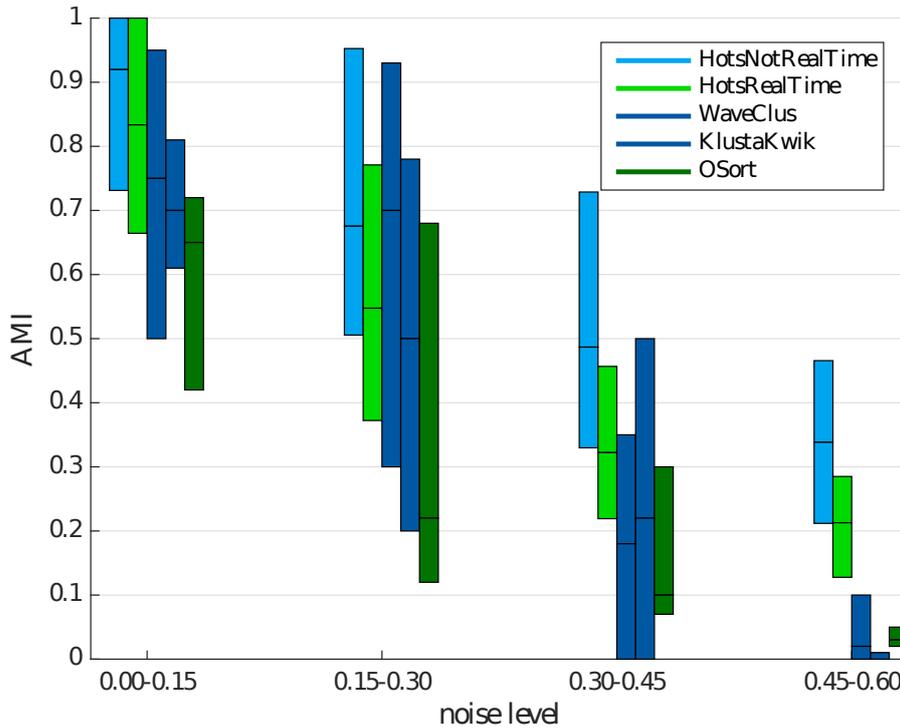


Figure 2.3 – Recognition rate (AMI) with median and standard deviation versus noise level and comparison to others methods. Blue bars report classification rates after layer 5, green ones after layer 3. Our method (blue and green bars) performs better than others : standard deviation is lower (median is represented in black middle line). For low noise conditions, our median is higher. For noise level between 0.15 – 0.3, we are a bit under WaveClus, but over all the other ones. We can also see that in noisy conditions (noise level > 0.45) our algorithm is significantly better than all the others. 3 Layers run real-time.

(TSNE, Maaten and Hinton, 2008) for data visualization, the projection of the signatures for some testing examples in two dimensions shows the ability of the algorithm to distinguish different types of spikes (2.4b). 2.3 presents the AMI score for the presented algorithm compared to state-of-the-art methods. It has to be taken in consideration that the only previous method working in real time is the OSort algorithm. The HotsL3 and HotsL5 presented here are for the output of the third and fifth layer of the hierarchical model. The classification score after layer 5 overcomes all the other methods. After layer 3, only for intermediate noise level (0.15 – 0.30) are we slightly below Waveclus. For noisy dataset (0.45 and more), the proposed method overcomes drastically all the others. This shows the very good immunity of the proposed method toward noise. Regarding the required time to process all these data, our algorithm performs real time for 1-3 layers on a standard desktop computer (Intel Core i7-4790 CPU @ 3.60GHz,

Table 2.1 – Model parameters for the artificial benchmarking

	L_1	L_2	L_3	L_4	L_5
$\tau(\mu s)$	25	50	75	100	125
$R(p\lambda)$	2	4	8	16	32
N_c	4	11	39	147	562
ζ	5.10^{-5}	5.10^{-5}	5.10^{-5}	5.10^{-5}	5.10^{-5}

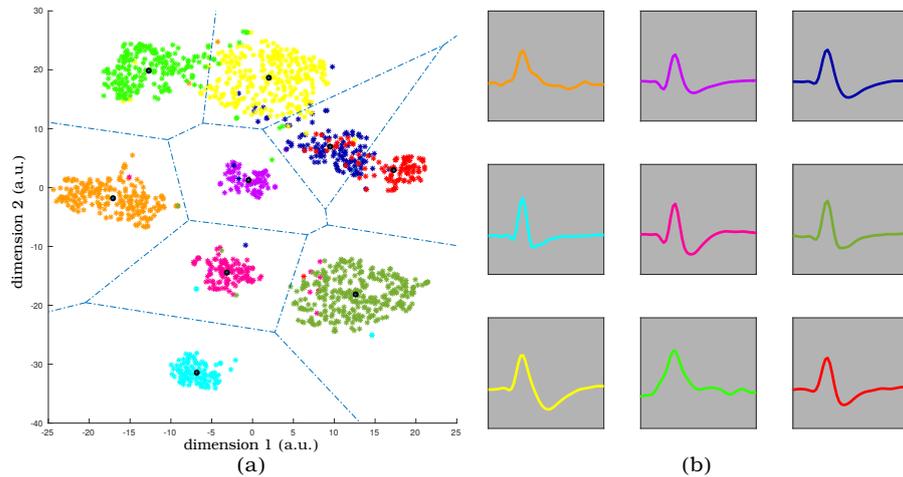


Figure 2.4 – Classification for the benchmarking data. (a) Voronoi diagram for the projected signatures in two dimensions and corresponding ground truth in color. Mostly, the signatures contain enough information to ensure good discriminant classification. (b) The associated temporal shapes. We can notice that our algorithm extracts signatures that are distinguishable one from each other all thereby allows good classification.

16GB of RAM, running Debian 8.5 Jessie). Figure 5 represents the required time for 1-5 layers.

2.3.2.4 Extension to *in-vivo* and *ex-vivo* data

Good results on semi-artificial dataset leads us to benchmark the proposed method on real data. Buzaki’s lab developed new techniques for simultaneous recordings of *in-vivo* (patch clamp) and *ex-vivo* (multi electrode array) data (Henze et al., 2000; Harris et al., 2000). Data from the CA1 hippocampal region of anesthetized rats (Henze et al., 2009), containing both intra and extracellular recordings, was used. The extracellular tetrodes (Recce and O’keefe, 1989) are made of four 13- μm polyimide-coated nichrome wires. The intracellular glass micropipette filled mainly with potassium solution. The extracellular data contains responses from two different cells (type A and type B), whereas the *in-vivo* data target only one (type A). So, the intracellular data can be used as a ground truth to validate the sorting on extracel-

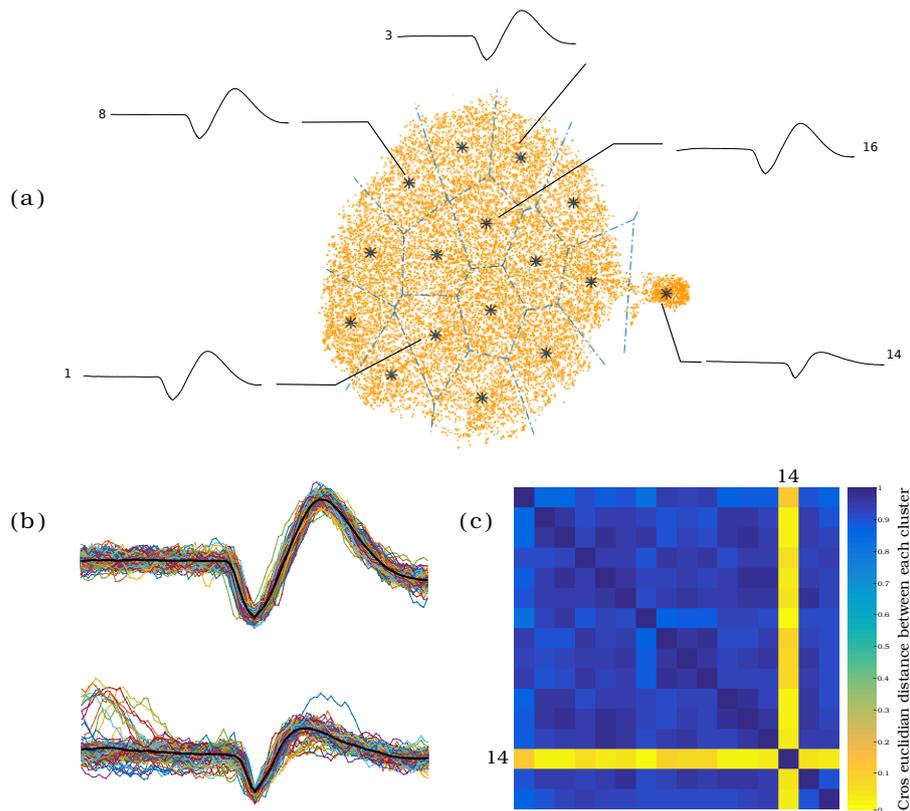


Figure 2.5 – Results with data from ex-vivo rat retina. (a) TSNE projection with output of the last layer. It is possible to distinguish two major types of cells (c) Clustering is able to find several possible new cells, which can be merged together. (c) A simple cross Euclidian distance able of merging each shape to estimate the best number of classes (the darked the closer). (b) The two found classes with mean shape (dark line) and some examples on top (color)

lular data. Without fine tuning of the parameters, the use of the same parameters as in previous section leads up to 90% recognition rate.

The dataset was split into 10 subsets, randomly picked, with around 30 % of type 1 cell. One subset is used for learning the clusters, and the 9 others to test the algorithm. The experience was run for 100 times. Table 2.2 shows the results of these runs. Our algorithm outperforms state-of-the art real-time spike, and is slightly lower than the best method reported on this dataset, which is not real-time. Computation time (on a Intel Core i7-4790 CPU @ 3.60GHz, 16GB of RAM, running Debian 8.5 Jessie) is about 25 seconds for one set of 80 seconds of data, which is more than 3 times faster than the real-time

Finally, we used the proposed algorithm to process data from ex-vivo primate and rat retinas recorded with a Multi-Electrode Array (MEA). This grid of electrodes records neuronal activity, feed it through an analogical to numerical converter (ADC), filtered by an high-pass filter to remove noise and offset. During the learning, the

Table 2.2 – Recognition rate on in-vivo data. The bolted algorithm are the only one working in real-time. Our algorithm performs better than state of the art online spike-sorting, and is slightly below the known best algorithm reported on this dataset, which is not real-time.

Method	minimum	maximum	average
Shahid, Walker, and Smith, 2010	N/A	N/A	66%
Werner et al., 2016	N/A	N/A	82.1%
Gasthaus, 2008	N/A	N/A	Fscore 0.91%
Haessig et. Al (proposed method)	84.75%	90.22%	87.33% - Fscore 0.89%

noise level of each electrode is evaluated to put a dynamic threshold (generally three times the standard deviation) able to separate spikes from noise. When a spike is detected, it is processed by the algorithm as an new unknown example.

The sorting method uses two main steps: separating shapes and evaluating the propagation speed of the spikes. Our algorithm is able to deal with the first issue. The second one is tackled with a method introduced by Li et al., 2015 that uses an axon on the grid to infer the timing between two electrodes and then the distance between them. 2D projection of recordings from a rat retina, after traveling through the proposed method with 3 layers, are shown in 2.5a. The clusters are different enough to be linearly separable and so authorize good classification. No ground truth was available for these recordings.

2.4 DISCUSSION

We presented a method being able to extract and decode in real time spiking activity, outperforming state of the art algorithms on standard benchmarks. In the framework of this article, the input signal is first sampled and then converted into events. Intrinsically, this approach cannot reflect the signal’s exact dynamic. The linear interpolation was a first step toward time precision, but we are still limited by the sampling period. Further work will implement a spike generation directly from the analog signal, and then removing the drawbacks induced by the Analog-to-Digital converters. All this work was targeted for mono-electrode recordings and settles fundamentals of such data processing. A generalization to multi-electrode setups need to be developed.

We expect the most significant impact of our model to be in the field of brain-machine interfaces. Today’s interfaces face severe limitations imposed both by the conventional sensors front-end (which produce large amounts of data with fixed sample rate), and the classical Von Neumann computing architectures (which are limited by

the memory bottleneck and require high power and high bandwidth to process continuous stream of data). The merging field of neuromorphic engineering has produced efficient event-based sensors, that produce low-bandwidth data in continuous time, and powerful parallel computing architectures, that have co-localized memory and computation and can carry out low-latency event-based processing. This technology promises to solve many of the problems associated with conventional sampling. However, the progress so far has been chiefly technological (Corradi and Indiveri, 2015), whereas related development of event-based models and signal processing algorithms has been comparatively lacking (with few notable exceptions). This work elaborates on an innovative model that can fully exploit the features of event-based sensors.

VISUAL PERCEPTION OF SINGLE PIXEL STIMULATIONS WITH A PHOTOVOLTAIC SUBRETINAL PROSTHESIS IN NON-HUMAN PRIMATES

Blindness occurs following photoreceptor degeneration in retinal dystrophies and age-related macular degeneration. Electrical activation of the residual retinal circuit has provided useful artificial visual perception to blind patients. Unfortunately, the resolutions of the current retinal prostheses have not yet enabled face recognition, text reading or independent locomotion due to low electrode numbers or lack of distinct percepts upon neighboring electrode stimulations. Here, we provide evidence for single electrode resolution (100 μm pitch) by the near-infrared-sensitive photovoltaic retinal prosthesis Prima in the non-human primate degenerated retina *ex vivo* and *in vivo*. This visual perception obtained below radiation safety limits with this prosthesis containing hundreds of electrodes has paved the way towards its clinical evaluation starting on dry atrophic age-related macular degeneration.

3.1 INTRODUCTION

Recent clinical trials have shown that retinal prostheses can restore some visual perception through electrical stimulation of the remaining inner neurons in the degenerated retina (Brandli et al., 2016). Psychophysics studies in sighted subjects had shown that 600 pixels or distinct percepts are required to restore autonomous locomotion, face recognition and text reading even in a slightly eccentric position (Cha, Horch, and Normann, 1992; Cha et al., 1992; Fornos et al., 2008; Sommerhalder et al., 2003; Sommerhalder et al., 2004). Accordingly, retinal prostheses were designed to progressively reach such spatial resolution by increasing the number of pixels: starting from few to more than a thousand electrodes (Stingl et al., 2017; Da Cruz et al., 2013). However, visual perception has not yet reached the expected visual acuities, suggesting an absence of distinct individual percepts for neighboring pixels. Indeed, in all these retinal prostheses, electrical currents generated at stimulating electrodes have to return to a distant ground. Modeling studies suggested that this configuration with a distant ground produces a poor confinement of the electrical field whereas a ground grid surrounding individually all stimulating electrodes would greatly improve the resolution and contrast of pattern stimulations (Joucla and Yvert, 2009; Bendali et al., 2015). Using

such an architecture, a photovoltaic near-infrared (NIR) sensitive device, demonstrated a high spatial resolution in rodents (Lorach et al., 2015). Each pixel is independently powered by two NIR sensitive photodiodes connected in series between the central electrode and the surrounding hexagonal ground grid (3.1a). NIR light can be delivered through the pupil to stimulate this completely wireless and modular implant, allowing a less invasive surgical procedure while leveraging eye movement object scanning.

3.2 METHODS AND MATERIALS

3.2.1 *Animal models*

All non-human primates in this study were *Macaca fascicularis* and originated from Mauritius. Experiments were performed at the Institut du Cerveau et de la Moelle Epiniere (ICM, Paris) and the Molecular Imaging Research Center (MIRCent- CEA, Fontenay aux Roses) or donated for the *ex vivo* study by Sanofi (Chilly-Mazarin). All experiments were ethically approved by the French “Ministère de l’Education, de l’Enseignement Supérieur et de la Recherche” under the project references APAFIS#6795-201609201505827 v2 and APAFIS#5929.

3.2.2 *Animal preparation*

Three non-human primates (NHP, *Macaca fascicularis*) were used in this study, all male (7-8 years old, 6.5-7.5 kg). They were chair-trained and familiarized with the laboratory and trainers prior to studies. A titanium head post was surgically implanted on top of the NHP skulls with titanium screws. This implantation was realized in sterile conditions. The surgery for the prosthesis implantation was performed under general anesthesia. After a complete 23-gauge vitrectomy, a retinal bleb was created via subretinal injection of BSS around 4 mm away from the fovea. An endodiathermy of 3 mm close to or between the upper retina vessels was realized before the 2.5 mm retinotomy so that the delivery system with the implant could be presented perpendicular to the retinotomy. This tool was inserted into a 2 mm sclerotomy, 3 mm from limbus, to insert the implant subretinally. Perfluorocarbon liquid (PFC) was injected to flatten the detached retinal tissue and stabilize the implant under the retina. The laser coagulation was performed at the border of the retinotomy (OPHTHALAS 532 EYELITE, Alcon, VITRA-LASER 532 nm GREEN, Quantel Medical). PFC was then removed within the fluid-air exchange. Finally, silicone oil was injected in the eye to ensure retinal reattachment and removed 4 weeks after implantation because it is known to alter the eye refractivity (Smith, Smith, and Wong, 1990).

3.2.3 *Retinal imaging*

Optical coherence tomography (OCT), infrared, blue reflectance imaging were performed using SPECTRALIS (HRA+OCT, Heidelberg Engineering, Heidelberg). Fundus color photographs were captured with the Smartscope camera (Optomed, Oy, Oulu, Finland).

3.2.4 *Multifocal ERG*

Multifocal electroretinography (mfERG) was performed using the RETImap system coupled with the scanning laser ophthalmoscope (SLO) with a binary m-sequence technique (Roland Consult, Germany). The test pattern was in a 30 deg diameter field for 37 hexagons without distortion. Recording was performed at a 59.8 Hz frame rate with 1 kHz (983.5 μ s) sampling rate. Evaluation was for the first-order kernel.

3.2.5 *Stimulation system*

Our stimulation system, derived from (Lorach et al., 2015), included a single-mode NIR laser (880 nm, 2.5 W, Laser 2000) and a visible light source (590 nm), which beams were merged together thanks to a dichroic mirror onto a Digital Micro-mirror Device (DMD, DLI innovations) to generate visual patterns. We mounted the optical system on a slit lamp (Zeiss SL-130). Three NIR non-reflective lenses were placed after the DMD to create a collimated image of the DMD pattern.

3.2.6 *Behavior*

The NHPs were trained daily and were rewarded for making saccadic eye movements from a central square fixation point to a peripheral and circular spot. Once central fixation was achieved for 300 ms on a square spot (300 μ m), a peripheral stimulation was presented within less than 9 ms while removing the central target. The peripheral stimulation consisted in a circle (300 μ m by default or less depending on the near infrared stimulation size) presented for 210 ms. Due to the optical constraints, the left eye was used for eye-tracking while stimulating the implanted right eye. Daily training involved immediate rewards for proper center-out saccades in the adequate peripheral direction in the visual field. The entire visual field was divided in 72 angles. Trials (100-200) presented in a pseudorandom order with visible targets were interleaved with 20% NIR-stimulation trials and control trials (10% no DMD patterns or 10% no light). Default NIR stimulations consisted of 6 flashes lasting 10 ms every 40 ms such that this stimulation ended (210 ms) before the average latency

required for saccade initiation by visible stimuli. Nine implant stimulation locations were used in a 3x3 array with a distance between the most distant stimulations generating a 5° of visual angle. Control trials consisted in the disappearance of the fixation target without the presentation of the peripheral stimulation. In half of the controls, the DMD was set to the complete OFF position, whereas in the other half the light source was switched off but the DMD was set for presenting a pattern in the tested visual field to investigate potential uncontrolled reflections. The reward was delivered either immediately after the saccade at a proper location or, for controls, upon maintained fixation at the centre for 300 ms after the central target disappearance. The same applied for visible or NIR stimulations. To calculate the distance for the peripheral target presentation, we used the reported correspondence $200 \mu\text{m} \cdot \text{deg}^{-1}$ in non-human primates (Lapuerta and Schein, 1995).

3.2.7 *Retinal tissues*

NHPs received a lethal dose of pentobarbital. Eyeballs were removed and placed in sealed containers for transport with CO² independent medium (ThermoFisher scientific), after transcorneal puncture of the eye with a sterile 20-gauge needle. Retinas were then isolated and dissected either for histology or for retinal explant culture.

3.2.8 *Retinal explant culture*

Retinal tissues were preserved in CO² independent medium (ThermoFisher scientific). To produce a blind NHP retinal model, the retina was transversely sectioned to remove the photoreceptor layer as previously reported for transplantation of photoreceptor layers (Silverman and Hughes, 1989). A square piece of retinal tissue (2x2 mm²) was cut at its four corners to flatten it down photoreceptor side on a flat pork skin based gel block. A razor blade cut 100 – 110 μm deep from the vitreal surface was performed to isolate the inner retina. These inner retinal tissues were maintained in culture for 12 to 36 hours on polycarbonate transwell (Corning) in Neurobasal + B27 medium.

3.2.9 *Multi electrode array recordings*

Retinal explants were transferred onto a cellulose membrane preincubated with polylysine (0.1%, Sigma) overnight. A NIR-sensitive photovoltaic implant was inserted on the cellulose membrane with implant electrodes facing the photoreceptor side of the retinal explant. Once on a micromanipulator, the retinal explant was gently pressed against a Multi electrode array (MEA256 100/30 iR-ITO; Multi-

Channel Systems, Reutlingen, Germany), the retinal ganglion cells facing the MEA recording electrodes. The retinal tissue was continuously perfused with Ames medium (Sigma-Aldrich, St Louis, MO) bubbled with 95% O₂ and 5% CO₂ at 35 °C at a rate of 8 mL.min⁻¹. Spontaneous spiking activities from retinal ganglion cells (RGCs) were recorded after filtering the signals from each electrode below 300 Hz (Butterworth filter, 2nd order). Spikes were attributed to individual cells using an in-house spike sorting algorithm. Spike velocity were then calculated according to the protocol previously reported (Li et al., 2015). The implant was stimulated through the MEA using an NIR laser (2.5 W, 880 nm Laser 2000). This laser was identical to the ones used by Pixium Vision during dry electro-optical characterizations. Output light intensities were calibrated to range from 0.1 mW.mm⁻² to 5 mW.mm⁻² while NIR patterns were created using a Digital Micro Mirror Device (DMD, DLI innovations). To define whether NIR-elicited spikes were resulting from bipolar cells or direct RGC activation; the glutamate receptor blockers, LAP-4 (50 µM) and CNQX (100 µM), were added to the AMES medium. RGC activity was defined as the average number of action potentials in a 40 ms window following the stimulation artifact. The latter lasted during the stimulation pulse and prevented spike detection within that period. The cell response threshold was set at level exceeding the cell spontaneous activity by three-fold.

3.2.10 Histology

Retinal tissues were fixed overnight at 4 °C in 4% (wt/vol) paraformaldehyde in phosphate buffered saline (PBS; 0.01M, pH 7.4). The tissue was cryoprotected in successive solutions of PBS containing 10%, 20% and 30% sucrose at 4 °C and embedded in OCT (Labonord, Villeneuve d'Ascq, France). Retinal sections were permeabilised for five minutes in PBS containing 0.1% Triton X-100 (Sigma, St. Louis, MO) and rinsed. For immunolabelling, retinal fragments were incubated in a blocking solution in PBS (1% bovine serum albumin (Eurobio, Les-Ulis, France), 0.1% Tween 20 (Sigma)) for 1 h at room temperature. They were then incubated for one night at 4 °C with primary antibodies in a blocking solution. The antibodies used were a monoclonal antibody directed against mouse G0α (1:200, Chemicon, Darmstadt, Germany) and a polyclonal rabbit cone arrestin (Luminaire junior, LUMIj, 1:20,000). The fragments were rinsed and then incubated with secondary antibodies: donkey anti-mouse IgG and donkey anti-rabbit IgG antibodies conjugated to AlexaTM₄₈₈ and AlexaTM₅₉₄, respectively (1:500, Molecular Probes, Invitrogen, Eugene, Oregon) for 1 h at room temperature. Cell nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI), which was added during the final incubation period. Sections were rinsed then mounted with Fluorsave reagent (Permafluor) on a mi-

crosscope slide. Confocal microscopy was performed on an Olympus FV1000 laser-scanning confocal microscope.

3.2.11 *Human retina ex vivo recordings*

To further control the ability of photovoltaic implants to activate human retinal neurons, we tested their efficacy on a post-mortem human retina. The eye was dissected 8 hours after the time of death and retinal explants were cultured for 48 h without photoreceptor ablation. Despite holes in the inner nuclear layer, the tissue structure appeared well preserved (3.3a,b). Full field measurements were performed using an implant with 100 – μm -wide pixels. 3.2c,d, show a raster plot and post-stimulus time histogram (PSTH) of 120 stimulations at 1 ms and 5 $\text{mW}\cdot\text{mm}^{-2}$ and 3.2f,g, show the evolution of the number of action potentials with pulse duration and stimulation intensity for 2 cells from this human retina. Less action potentials were observed using the human retina compared to the NHP retinas' recordings. This is likely explained by the lack of oxygen and nutrients during the postmortem interval before dissection and culture.

3.3 RESULTS

3.3.1 *Ex-vivo*

After adapting the fabrication process of these photovoltaic implants to obtain a clinical-grade device (1 mm or 1.5 mm wide, 30 μm thick, 140 or 100 μm electrode pitch), we here investigated their efficacy and spatiotemporal resolution in the degenerated non-human primate (NHP) retina. The implant efficacy was first assessed *ex vivo* on a NHP retinal model of photoreceptor degeneration. We generated this blind model by slicing off the retina in its thickness using a vibratome to eliminate the photoreceptor layers (3.1b, 3.1c). Pieces of such blind NHP retina were kept in cultures for few days and recorded on a multielectrode array (MEA, 3.1a). Retinal ganglion cells (RGCs), the retinal output, were recorded while stimulating the photovoltaic implant located on the bipolar cell side in place of the removed photoreceptors, bipolar cells are presynaptic to RGCs. RGCs generated spikes in response to full-field pulses of NIR light (880 nm) at a low NIR irradiance (0.48 $\text{mW}\cdot\text{mm}^{-2}$, 2 retinas, 166 cells) for 4 ms flashes. When varying the flash duration (1 to 10 ms), peak responses were reached at 1.9 ms duration for full field stimulations with 1 $\text{mW}\cdot\text{mm}^{-2}$ (2 retinas, 166 cells). With the shortest stimuli, responses were observed with latencies of around 20 ms after the stimulation artefact (3.5 $\text{mW}\cdot\text{mm}^{-2}$, 4 retinas, 99 cells). Application of synaptic blockers (LAP-4, CNQX) led to the disappearance of these RGC responses, indicating that this spiking activity was due to bipolar

cell activation rather than direct RGC stimulation (3.1h) in agreement with a previous report in rodents (Mathieson et al., 2012). However, direct RGC activation cannot be completely excluded because the stimulation artefact could dissimulate rapid spike initiation. Full field stimulations also elicited RGC spikes in a postmortem human retina (3.3).

Because vision requires pattern recognition with high spatial resolution, we characterized the response to single electrode activations (3.1d,e,f). RGC activation was obtained upon higher intensities ($4\text{mW}\cdot\text{mm}^{-2}$) and longer stimulation durations (5.4 ms) (3 retinas, 236 cells, 3.1g), parameters which remain below the radiation safety limits for chronic use (Lorach et al., 2016). To further assess spatial resolution of the retinal stimulation by the implants, we mapped the electrical receptive fields (eRFs) of RGCs using single electrode stimulations (4 ms, 10 Hz, $5\text{mW}\cdot\text{mm}^{-2}$) in a pseudo-random manner. 3.1g,h illustrate the eRF of a RGC with a single stimulating electrode (100 μm pitch) while its neighbors generate no response. Out of the 22 recorded cells, 9 cells responded to only one single electrode (100 μm pitch), 2 cells responded to 2 electrodes when stimulated independently and no cells responded to 3 or more electrodes. Similarly, with a larger pixel size (140 μm), on 77 recorded cells from 3 retinas, 41 cells responded to a single pixel, 8 cells to 2 separate units and 0 cells to 3 or more units. To define RGC types activated by these NIR stimulations, we measured their spike propagation velocity as previously described and found velocities ($0.56 \pm 0.1\text{m}\cdot\text{s}^{-1}$, N=99) consistent with that of midget RGCs (Li et al., 2015). Their single electrode receptive fields are in agreement with the small dendritic field of midget cells in the recorded perifoveal area. This selective single electrode activation of midget RGCs provided the first evidence for high resolution in the blind non-human primate retina.

3.3.2 *In-vivo*

To further translate these photovoltaic implants to clinical trials, we developed the surgery to introduce them in the subretinal space on living NHPs (*Macaca fascicularis*, P1-3, 3.2a). The consecutive retinal detachment from the underlying retinal pigment epithelium and choroid induced photoreceptor degradation in the retinal area facing the implant as indicated by the disappearance of photoreceptor layers (ONL) on *in vivo* OCT scans (3.2b and 3.4). Further histological examinations and multifocal electroretinograms confirmed photoreceptor degeneration (3.5, 3.6) in agreement with a previous study in normal rats (Lorach et al., 2015). This local degeneration of photoreceptors in the implant area produces a blind spot modeling retinal diseases with photoreceptor degeneration. It enabled us to assess if NIR activation of the photovoltaic implant can activate the corresponding

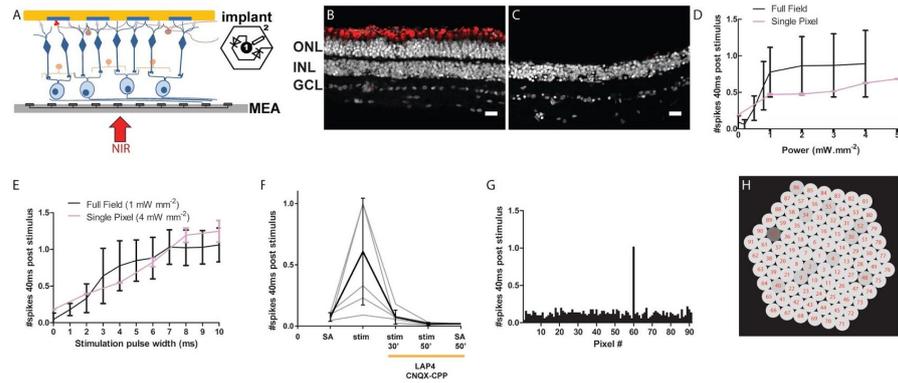


Figure 3.1 – RGC responses to full field and single electrode activations on an *ex vivo* blind NHP retina. (A) RGC recording configuration on a multielectrode array (MEA) with the implant (yellow) above bipolar cells. Inset: the wiring diagram for a single implant unit, consisting of two photodiodes connected in series between the active central electrode (1) and the surrounding ground grid (2). (B,C) Histological sections of a control NHP retina (ONL outer nuclear layer, INL inner nuclear layer, GCL ganglion cell layer) (B) and a retina after vibratome sectioning the photoreceptor layers (C). The removal of photoreceptors is illustrated by labeling cone photoreceptors (Cone arrestin in red) and ON bipolar cells (G0 α protein in white), with DAPI nuclear staining of the cell layers (white), scale bar 20 μm . (D) RGC responses to full-field stimulations (black) and single implant electrode stimulations (red) while varying near-infrared illumination power with 4 ms pulses (error bars are standard deviations) (E) Thresholds measurements after full-field stimulations (black) and single pixel stimulations (red) with varying near-infrared stimulation pulse widths at 1 and 4 $\text{mW}\cdot\text{mm}^{-2}$ respectively (error bars are standard deviations). (F) Spiking frequency during a pharmacological test showing the suppression of the implant-elicited activity by synaptic blockers. "SA" stands for spontaneous activity and "stim" stands for tests involving 120 stimulations at 2 Hz, 4 $\text{mW}\cdot\text{mm}^{-2}$ using an implant with 100 μm pixels. (G) Spike histogram of a RGC showing an increase in its spiking frequency only for pixel 60 (10 ms and 4 $\text{mW}\cdot\text{mm}^{-2}$). (H) Electrical receptive field of the RGC recorded in (G) after stimulation of individual implant electrodes (100 μm pitch).

blind retinal area *in vivo*. Three NHPs underwent the surgery with an implant positioned parafoveally with an electrode pitch, 140 μm for P1 and 100 μm for P2-P3. These animals were trained to perform a saccade detection task to demonstrate perception (3.2c). After a maintained central fixation for 300 ms on a square spot, a peripheral stimulation was presented within less than 9 ms while removing the central target. After an initial training period on a computer screen, the test was conducted on a modified slit-lamp allowing visible and NIR pattern stimulations using a Digital Micro-mirror Device (DMD)

(3.7). Due to optical constraints, the left eye was used for eye-tracking while stimulating the implanted right eye. Daily training involved immediate rewards for proper center-out saccades in the correct peripheral direction of the visual field, divided in 36 or 72 angles (3.2c). Experiments consisted of 100 to 200 pseudo-random trials including controls (10%: DMD in OFF position; 10%: light sources OFF). For controls, the reward was supplied for maintained fixation at the center for 300 ms after the disappearance of the central fixation target. Out of 345 control tests, only 6 failed; 4 of them corresponding to eye blinks. The remaining 2 saccades had neither the latency, nor the spatial accuracy towards the peripheral control position set on the DMD. Under visible stimulations, the three animals performed successful saccades at a rate higher than 90% except at the implant location (3.2c and 3.8). The locations of these blind spots finely defined with 2° test pitch (3.8) were consistent with the implant position defined on the fundus and OCT images (cf 3.2a,b). The visual field covered by the implant ranged from 9 to 18° for P1 in the lower part of his visual field and 5 to 14° for P2 and P3 in the upper left part of their vision.

The high rate of correct trials under visible light enabled us to interleave occasional (< 10%) stimulation trials of NIR light pulses (880 nm) without generating animal frustration. NIR stimulations were targeted either on the implant or away from the implant for controls. One animal, P2, responded to NIR stimulation in the direction of the implant with repeated successful saccades whereas NIR control stimulations away from the implant induced no saccade (3.2c). Modulating NIR intensities indicated that the maximum responses to implant stimulation were reached at 300 – 400 $\mu\text{W}\cdot\text{mm}^{-2}$ with a threshold at 200 $\mu\text{W}\cdot\text{mm}^{-2}$ (3.2d). The NIR sensitivity of the implant *in vivo* was therefore well below radiation safety limits for chronic use (Lorach et al., 2016). At the maximum response, the saccade reproducibility and response latencies were comparable to those of natural visible light perception (3.2f and 3.9). We then modified the other parameters of the NIR spot stimulation: size, number of flashes and flash duration. To further reduce the possibility for an eventual natural perception of the NIR light in these experiments, the ambient light was turned on during the tests and the NIR intensity kept at 800 $\mu\text{W}\cdot\text{mm}^{-2}$ (3.8right). Decreasing the spot size showed that the response maximum and smallest latencies were reached for a 175 μm spot size on the implant (3.2e). This spot size correspond to the spot required to fully cover one implant unit (100 μm) regardless of the respective implant position. This single electrode visual perception is well in agreement with our single electrode RGC responses *ex vivo* (3.1). Three flashes (10 ms) were sufficient to generate systematic perceptions whereas the reproducibility of the responses decreased with stimulations below 4 ms (3.9) in agreement with our *ex vivo* measurements during single electrode activation. These behavioral results

indicate that the photovoltaic subretinal implant can stimulate the NHP degenerated retina to elicit a reliable behavioral motor response indicative of a phosphene perception with stimulation parameters as low as $200 \mu\text{W}\cdot\text{mm}^{-2}$ NIR intensity, a single electrode spot size or three 10 ms flashes.

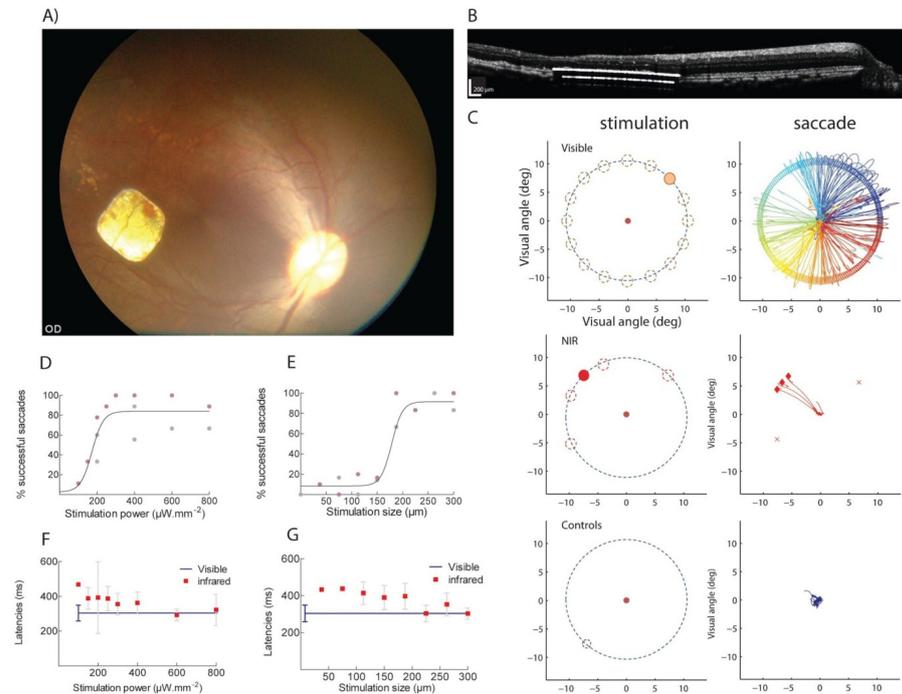


Figure 3.2 – **Single electrode activation in a behaving monkey.** (A) Eye fundus of P2 showing the photovoltaic implant. (B) *In vivo* OCT image illustrating the loss of photoreceptor layers above the implant of P2. (C) Each row represents the different types of stimulation: visible, near-infrared and controls. After central fixation (red dot, 300ms), a peripheral stimulation (half a degree) is presented every 2° in the periphery (orange and red discs). The NHP generates saccades in all directions except in the upper left part of the visual field corresponding to the implant position while with NIR stimulation (red circles), the NHP generates saccades only in this area. For controls (no light or DMD off), no saccade was observed. (D-G) Spot size and intensity dependences of the NIR stimulations. Percentages and latencies of successful saccades towards the implant when varying either the NIR power (spot size: $300 \mu\text{m}$) or the stimulation spot size (illumination power: $800 \mu\text{W}\cdot\text{mm}^{-2}$). Grey and purple dots represent two different days of experiment. The horizontal line represents the averaged latency following visible stimulations.

3.4 DISCUSSION

One NHP, P1, was not responding to the NIR stimulations and we were not able to measure artefacts upon NIR stimulations suggesting

that we were not able to activate this implant (140 μm pitch) *in vivo* (3.10b) while P2 showed clear artefacts upon NIR stimulations (3.10c). In P3, the 100 μm unit implant was properly activated as indicated by the large NIR-induced artefacts (3.10d). The lack of saccades for P3 may result from the perceptual differences between electrical stimulations and natural stimuli. With other retinal prostheses, implanted patients often acknowledged having difficulties to identify electrically elicited phosphenes as visual stimulations. Another explanation could lie in the rapid plasticity affecting the visual cortex in response to a lesion. Indeed, cortical areas corresponding to a blind retinal lesion were found to be rapidly reprogrammed to respond to other regions surrounding the scotoma (Gilbert and Wiesel, 1992; Heinen and Skavenski, 1991; Calford et al., 2003; Giannikopoulos and Eysel, 2006). This plasticity, which aims at producing a perceptual filling-in of the blind spot, can occur in days (Murakami, Komatsu, and Kinoshita, 1997). However, it remains unclear why one NHP would have had such a fast filling-in while the other had not. Cortical plasticity was similarly reported in patients with active retinal implants (Castaldi et al., 2016), and it likely enhanced their implant-visual perception by recruiting larger cortical circuits for their restored visual field.

These experiments have demonstrated that the NIR-sensitive, fully wireless photovoltaic implant can activate the blind non-human primate retina *ex vivo* and *in vivo* at intensities below radiation safety limits with a single 100 μm electrode resolution. This high resolution relies on the ground grid configuration confining electrical stimulations, which could have major implications for the development of other brain/machine interfaces. The Prima wireless design allows rapid and simplified surgery while providing gaze compatibility. This study has contributed to the acceptance for human clinical trials in dry age-related macular degeneration by the French and American agencies. Simulations indicate that thanks to its single electrode resolution, these implants could enable patients to reach a theoretical visual acuity 3/60, which is the threshold for blindness as defined by the World Health Organization (Mariotti, 2010). This therapeutic strategy could therefore provide the first treatment for geographic atrophy, the advanced form of dry age-related macular degeneration. The clinical implant with its 378 electrodes or a combination of several implants could thus provide blind patients autonomous locomotion, face recognition and text reading.

3.5 SUPPLEMENTARY FIGURES

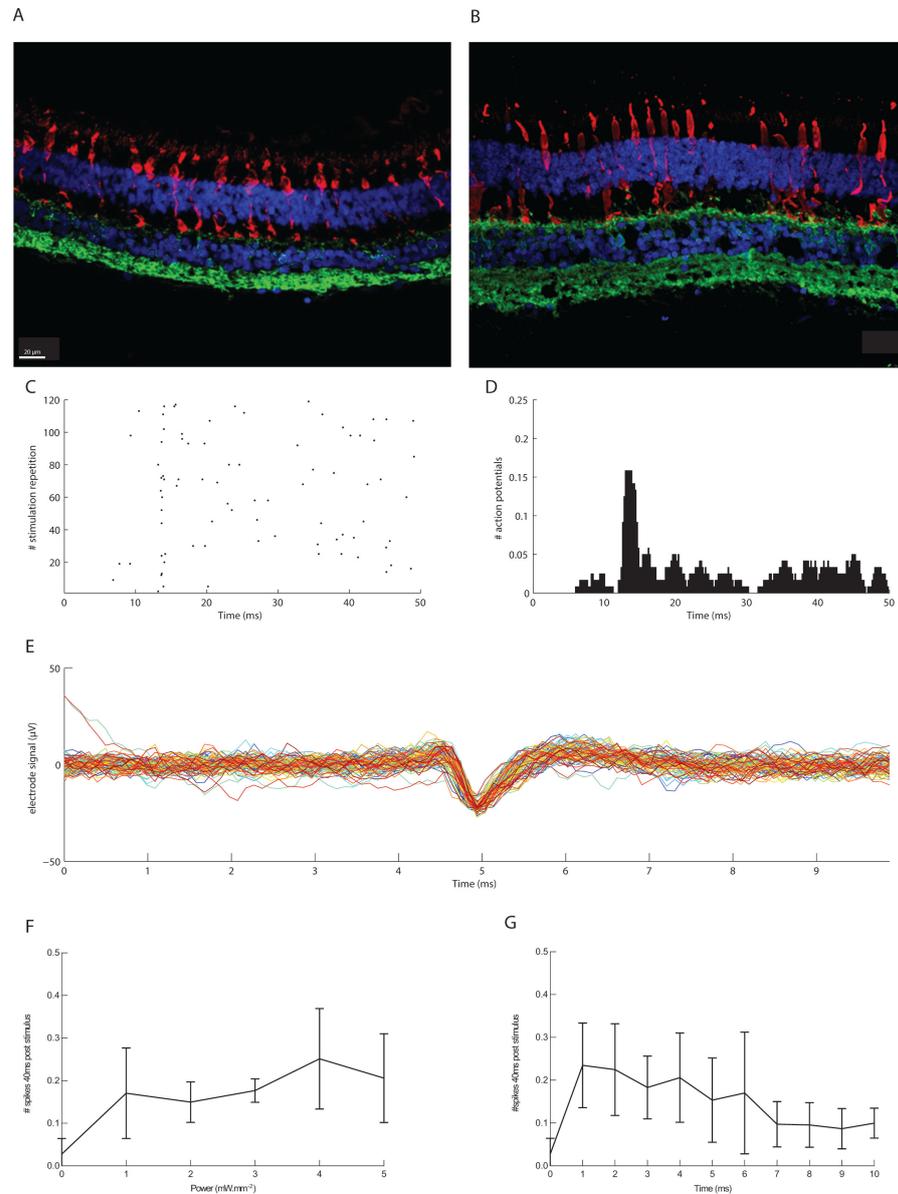


Figure 3.3 – Human retina stimulation. Cross sections of a non-human primate retina (A) and human retina (B). Cones were marked using cone arrestin (red), bipolar cells with G0 α (green), and cellular bodies with DAPI (blue). (C,D) Raster plot and PSTH of 120 full-field stimulations at 1 ms and 5 mW.mm⁻². Full-field measurements were performed using an implant with 100 μ m pixels. (E) Shape of the action potentials recorded. (F,G) Evolution of the number of action potential in a 40 ms window post full-field stimulus following with varying near-infrared power at 4 ms and stimulation pulse width at 5 mW.mm⁻² respectively (error bars are standard deviations).

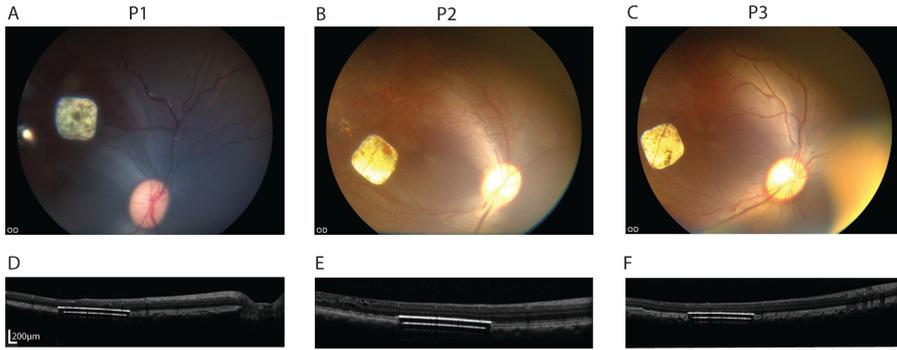


Figure 3.4 – **Non human primates imagery.** (A-C) Fundus images two weeks post implantation for P1-3 (implants are 1.5 mm in size). (D-F) OCT images of the implant below P1-3 retinas.

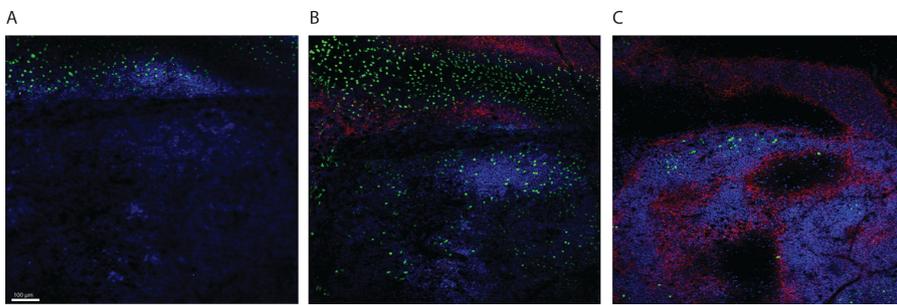


Figure 3.5 – **Photoreceptor degeneration above the implant.** (A-C) Confocal images of a NHP retina at different depth from the implant. The implant is located on the bottom part of the pictures. Cones were marked using cone arrestin (green), bipolar cells with G0α (magenta), and cellular bodies with DAPI (blue). There were 5 weeks between implantation and implant/retina retrieval.

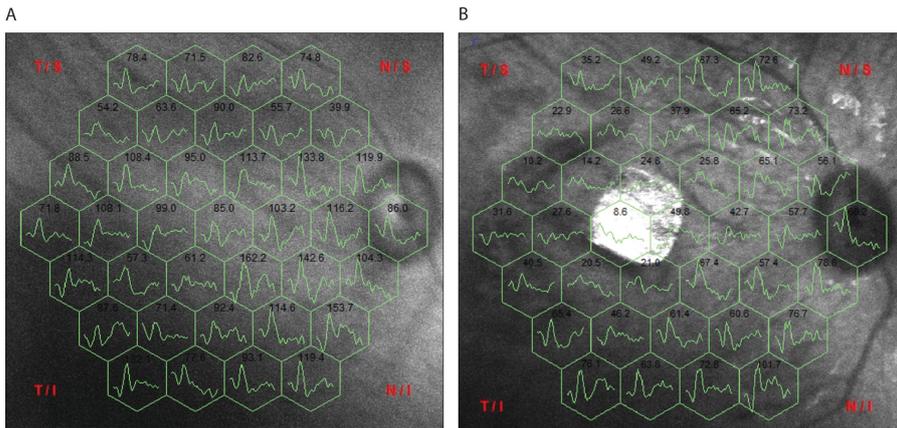


Figure 3.6 – **Photoreceptor degeneration.** (A) Multifocal electroretinogram of a non-human primate retina before implantation (B) electroretinogram of at the same location three weeks after implantation

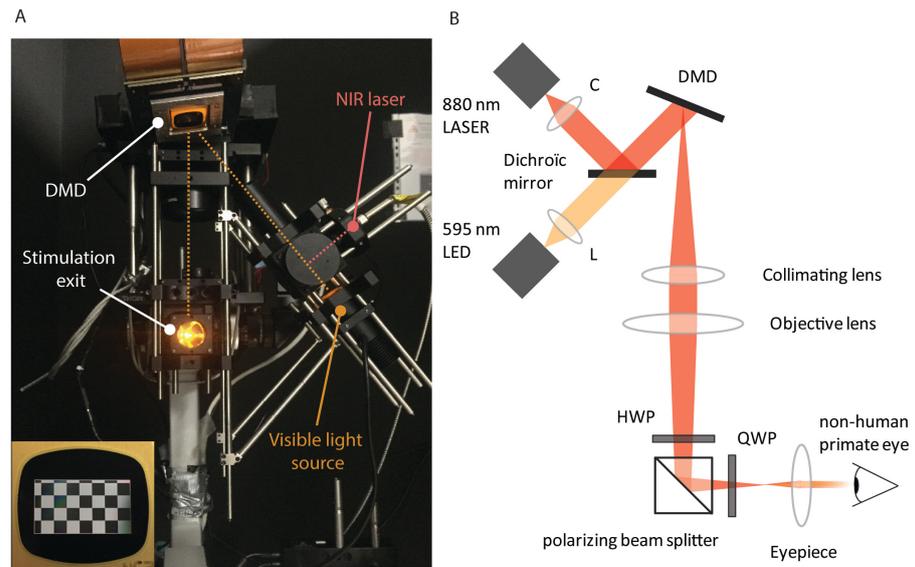


Figure 3.7 – **Stimulation apparatus.** (A) Picture and (B) schematic of the stimulation apparatus. The stimulation system included a single-mode NIR laser (880 nm) and a visible light source (590 nm). The two illuminations were merged together thanks to a dichroic mirror to illuminate a Digital Micro-mirror Device (DMD, DLI innovations, example of picture shown in inset) to form the patterns. We mounted the optical system on a slit lamp (Zeiss SL-130). Three near-infrared non reflective lenses were placed after the DMD to create a collimated image of the DMD pattern. HWP: Half wave plate, QWP: Quarter wave plate, PBS: Polarizing beam splitter, C,L: Collimators.

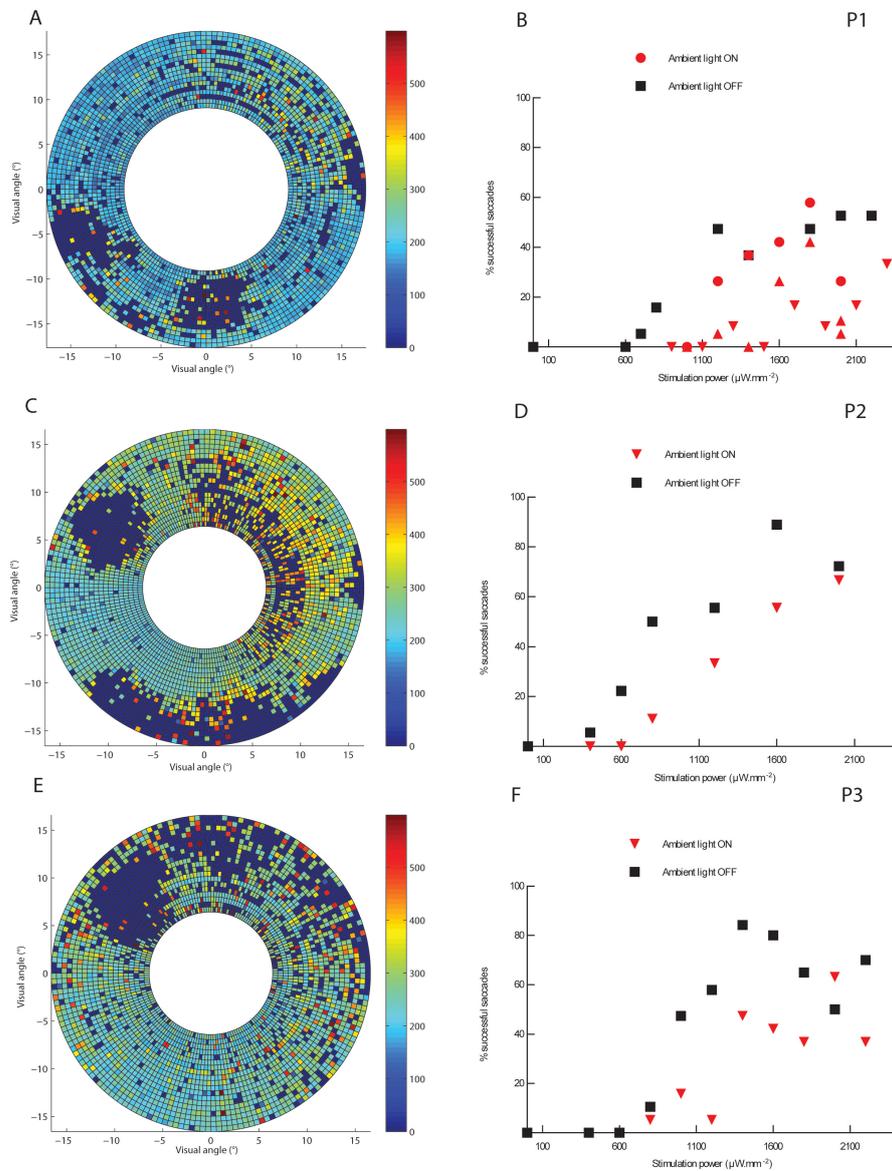


Figure 3.8 – **Visual fields and natural NIR vision.** (A,C,E) Visual fields of P1-3 respectively. The different blind spots correspond to the implant, the surgery and the optic nerve. The latencies of the saccades are depicted through the colors. (B,D,F) Near-infrared natural vision for P1-3 respectively, with and without ambient red light. The ambient red light was aimed at creating a red background for the near-infrared stimulation, lowering the contrast.

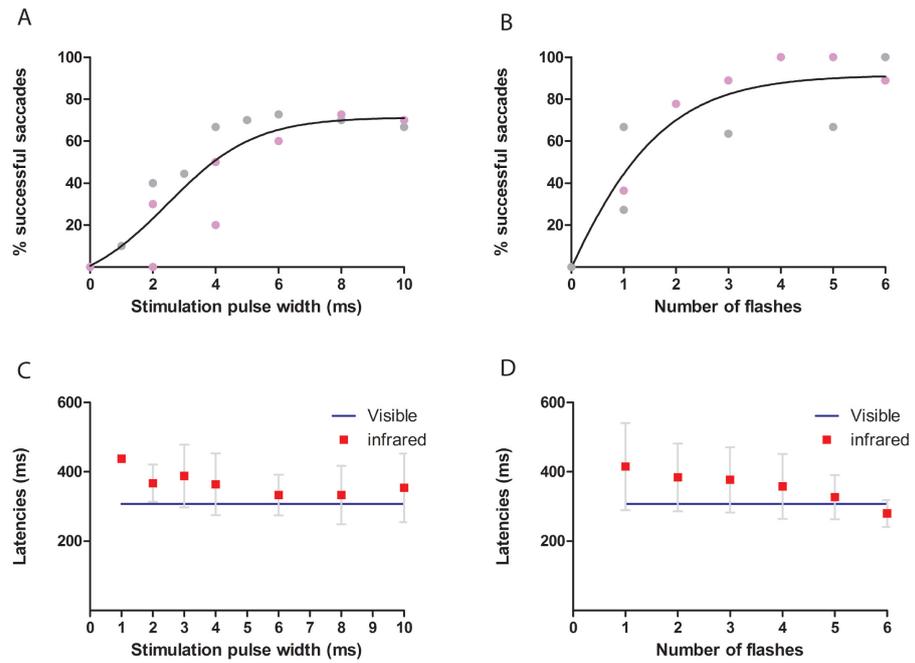


Figure 3.9 – **Implant saccades analysis.** Percentage of successful saccades towards the implant (A,B) and latencies (C,D) when studying stimulation pulse widths and number of flashes. Nine implant stimulation locations were used in a 3x3 array. Stimulations were separated by 5 degrees of visual angle. Grey and purple dots represent two different days of experiment, the horizontal line represent average response time following visible stimulations (error bars are standard deviations).

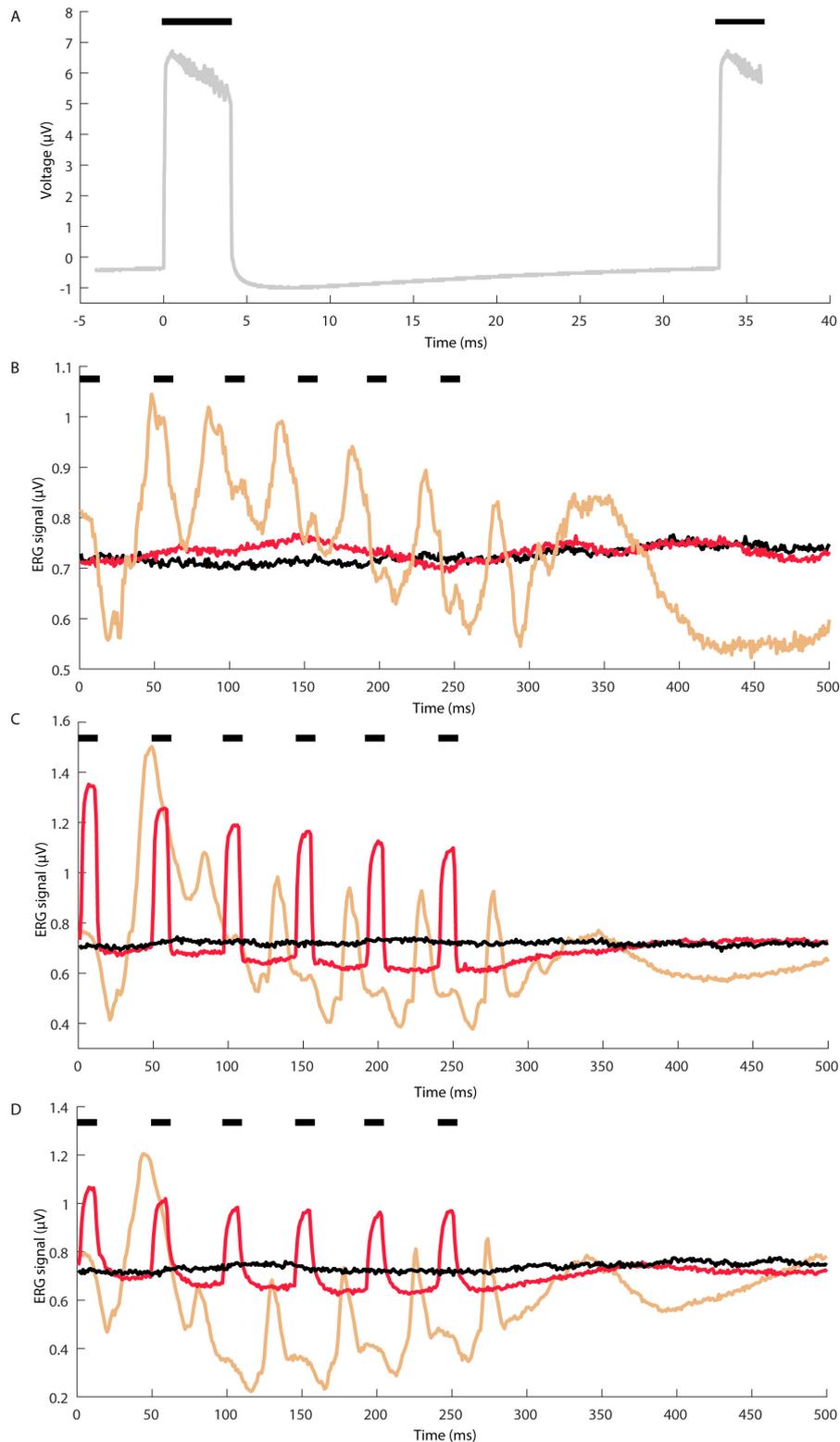


Figure 3.10 – **Implant electro optical characterization.** (A) Voltage measurement above a single pixel stimulation at $5 \text{ mW} \cdot \text{mm}^{-2}$ (B, C, D) electroretinograms of P1, P2 and P3 respectively showing retinal responses following six flashes of visible illumination (595 nm - orange curve) and NIR illumination ($2.4 \text{ mW} \cdot \text{mm}^{-2}$, red curve) in the implanted eye. NIR flashes did not induce any response in the control eye ($2.4 \text{ mW} \cdot \text{mm}^{-2}$, black curve). All flashes are represented by black rectangles.

Part III

PUPIL MODEL AS A BIOMETRIC

DILATION AND CONTRACTION OF THE PUPIL AS A RELIABLE BIOMETRIC

The problem of pirating data is more present in daily life than ever and biometric technology seems to be an effective answer. State of the art shows however that even this technology has weaknesses to take into account, especially face recognition which has been very popular in these recent years. The liveness detection seems to be the more effective countermeasure against most of the threats, however for face recognition several spoofing techniques stay available. Based on recent studies that show the biometric potential of the pupil, our work proposes a robust biometric solution able to be paired with face recognition. Our method relies on a live response to an active light stimulus which is able to provide an effective liveness detection and achieve 0.94 adjusted mutual information in recognition using a hidden markov model of the pupil behavior.

4.1 INTRODUCTION

Nowadays, it is easier than ever to steal and spoof personal data due to technological progress and the variety of available attacks. Biometric technology is now critical for personal, national and global security, making the search for hacks all the more desirable (Evans et al., 2015). Progress in synthetic or transformed voices can show certain limits of common recognition systems (Jin et al., 2008, Faundez-Zanuy, 2004), highlighting the need for effective countermeasures. Despite the iris and fingerprint being the most robust biometrics (Bednarik et al., 2005), they require dedicated and accurate sensors, in addition to being unpopular with the general public (Prabhakar, Pankanti, and Jain, 2003). Facial and speech recognition algorithms are able to work with regular sensors and could easily be used in daily life, they are however less robust (Bednarik et al., 2005). State of the art biometric technology shows several weaknesses, mainly, liveness detection (Smith, Wiliem, and Lovell, 2015, Wen, Han, and Jain, 2015, Tirunagari et al., 2015, Garcia and Queiroz, 2015, Sanchez et al., 2015, Sizov et al., 2015). Facial recognition algorithms are frequently spoofed by images, video sequences or 3-D models that circumvent liveness detection. Algorithms that only take into account 2-D textures or projection of a regular camera sensor can be fooled by a simple picture of the subject (Wen, Han, and Jain, 2015). Algorithms that are using a grid projection or a stereo-camera to evaluate depth as a way to avoid picture spoofing, can also be fooled using 3-D models (Erdogmus and

Marcel, 2013, Erdogmus and Marcel, 2014). A strategy to improve the accuracy of less robust methods is to combine multiple biometrics (Gragnaniello et al., 2015).

(Villalobos-Castaldi and Suaste-Gómez, 2013) shows that the pupil light reflex could be used as a countermeasure to liveness fooling and that pupil dynamics could be a reliable biometric. Using a model of natural eye movements (Komogortsev, Karpov, and Holland, 2015) is a good countermeasure for static representation or artificial 3-D models, however video sequences or masks could bypass it. A live pupil's response to light stimulus (Czajka, 2015) provides a further safety countermeasure to artificial models by triggering spontaneous pupils oscillations.

In this study, we present a framework that uses a light stimulation of the subject's eye eliciting a pupil's response to provide a reliable biometric based on dilation/contraction of the pupil that could be easily paired with other biometrics via widely accessible eye-tracker technologies (Cheng and Vertegaal, 2004). We measure the variations of the subject's pupil area using a scan of the visible spectrum from red to the purple at controlled luminance. Dilation and contraction are controlled by the sympathetic and the parasympathetic circuits that are directly relying on the retina (Wang and Munoz, 2015) therefore the distribution of the photoreceptors' types and their connections with the ganglion cells. The distribution of photoreceptors is highly specific (Williams, 2011) and sustainable (Weinrich et al., 2017) for each individual, providing specific behaviors in the light reflex mechanism of the pupil. These variations make use of the subject's sensitivity to the visible spectrum as a specific footprint.

4.2 METHODS AND MATERIALS

4.2.1 *Light stimuli*

As presented on figure 4.1, a light stimulus is sent to the subject's eye and the pupil's area variation is recorded by an eye-tracker (eyelink 2 sr research). The stimulus is a scan of the visible spectrum across time as described in figure 4.1. The wavelength of the light source switches from purple to red and from red to purple using the variation of only one primary RGB component at a time (see figure 4.2).

While the wavelength of the light source is changing, its luminance is controlled by the function represented by the curve L on figure 4.2. To study the influence of the luminance itself, the same stimulus is displayed colorless, by scanning from black to white. The eye tracker records the signal at 60Hz and we perform a z-score normalization and a level crossing to extract each significant change. A level is chosen as a variation of 0.01 on a z-scored signal.

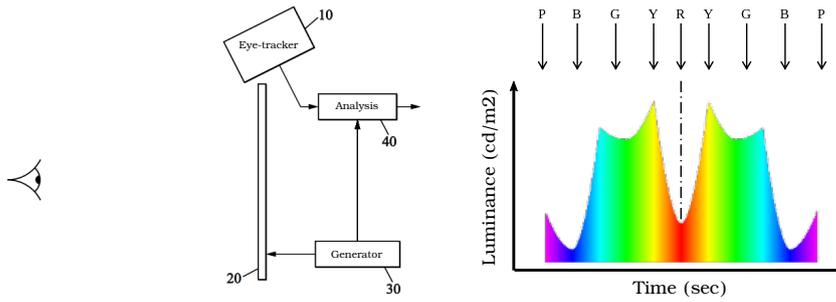


Figure 4.1 – This figure describes the experimental setup and the performed scan of the visible spectrum. The light source 20 represented by a screen shows the stimuli to the subject’s eye. Its dilation/contraction is recorded by the eye-tracker 10 and sent to the algorithm 40 using a processing tool 30. We describe a scan of the visible spectrum as follow: the light source smoothly shifts its wavelength from purple, to blue, to green, to yellow, to red and by the same way from red to purple. The luminance is controlled to stay constant according to the calibration of the display screen.

4.2.2 Database

The database of 6 subjects was recorded over a month. Each subject’s pupil was recorded three days per week with one day break between each day and between 2 and 4 trials of 45 seconds were recorded each day. A trial is the record of the subject’s pupil when the complete stimulation presented in figure 4.2 is performed. The eye tracker calibration was performed before each trial.

4.2.3 Algorithm

The subject’s pupil dilation/contraction is processed as a mono-dimensional signal that varies across the time, according to the wavelength and the luminance of the visible spectrum. The wavelength, or the color, of a light source and its luminance are used as a way to characterize and contextualize those variations (See 4.2.4). The Red, Green, Blue code (RGB) is used to describe each color and the luminance is measured in $\text{cd}\cdot\text{m}^{-2}$.

We choose in this study to work using the event based paradigm to analyze the signal as it allows an effective representation of the changes in the signal. Lets assume an event in a signal as a triplet as presented in equation 4.1:

$$ev_k = \{t_k, x_k, p_k\} \quad (4.1)$$

Where for each k , t_k represents the timestamp of change in the dilation/contraction of the subject’s pupil, x_k the value and p_k the type

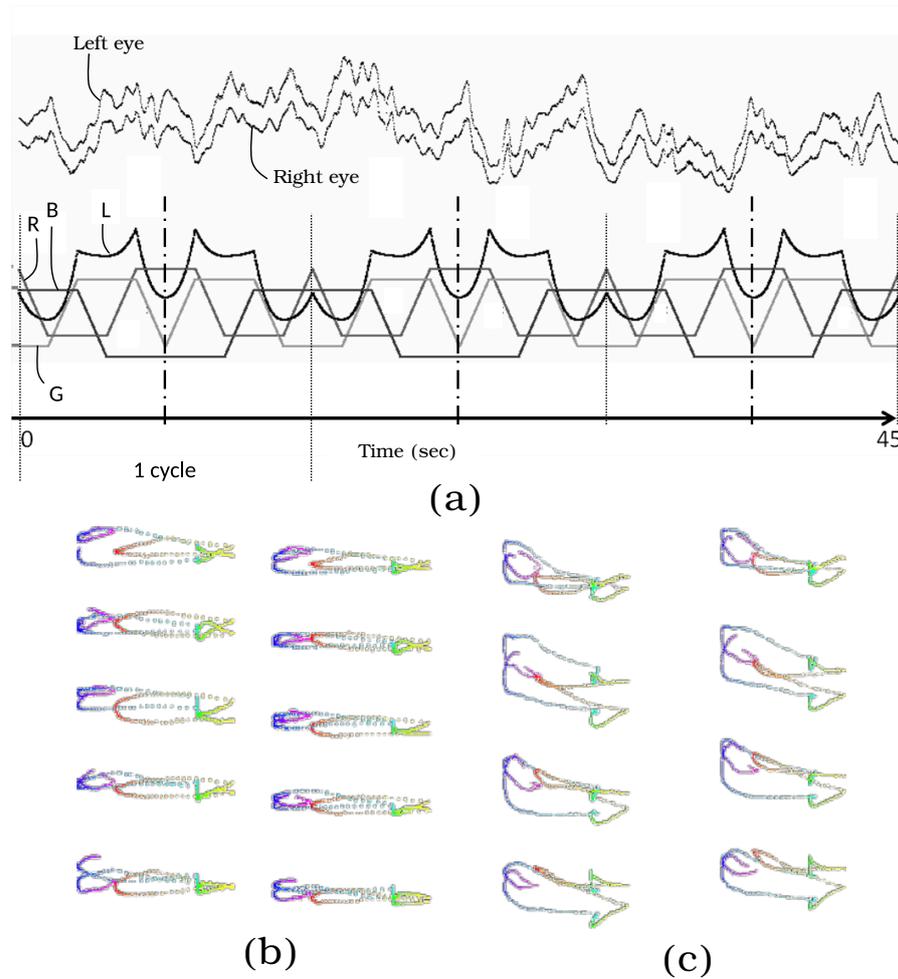


Figure 4.2 – (a) Those six curves represent the course of a recording. The top two curves represent the value of the subject’s pupils area across the time. Curves labeled R, G and B represent how the wavelength of the light source is controlled across time while the curve labeled L represents its luminance. Dotted dashed lines mark the half of a trial while dashed lines mark the beginning and the ending of the trials. For two subjects (b) and (c), the left curves represent the variations of the subject’s left pupil across 1 cycle in function of the luminance, while the right, the subject’s right pupil. We can see that a same subject’s pupil can vary for each cycle, while staying stable enough to be differentiated from another subject’s pupil.

of behavior of this change, called polarity. We based our descriptor on the concept of time surface introduced by Lagorce et al., 2016 and adapted to mono-dimensional signal by Haessig et al., 2019 that describes local spatio-temporal patterns around an event. A time surface, generated for each incoming event, is a matrix that contains the delay between the current event and its neighborhood stored using their

positions and polarities. An exponential kernel is applied to those delays according the equation 4.2:

$$TS_i = e^{-\frac{(t_k - t_i)}{\tau}} \quad (4.2)$$

Where t_k represents the timestamp of the current event and t_i , the timestamp of the neighborhood. While using the same online learning rule presented in *Haessig et al., 2019*, we introduce a sparse propagation criteria to improve both feature extraction and speed of the algorithm. This criterium is explained in the algorithm 2 and could be summarized as: each time a time surface is generated and attributed to a cluster, if the current event is not part of the region of the previous propagated event, or the current event is associated to a different cluster of the previous propagated event, then the current event is propagated.

Algorithm 2 Sparse propagation criteria in the Hots online routine for one layer from *Haessig et al., 2019*

Ensure: $\{t_i ; p_i ; x_i\}$ as the timestamp, polarity and position of each new incoming event
 lastp $\leftarrow -1$
 lastNeighborhood $\leftarrow \emptyset$
procedure SPARSE PROPAGATION(t_i, p_i, x_i) \triangleright Receive a new incoming event
 Compute current context S_i
 $t_{past}(p_i, x_i) \leftarrow t_i$ \triangleright update past
 $S_i \leftarrow \exp \frac{-(t_i - t_{past})}{\tau}$ \triangleright compute around x_i
 Compute distances
 for each centers M_k **do**
 $d_k \leftarrow -\log \left(\sum_k \sqrt{\frac{M_k \cdot S_i}{\text{card}(M_k) \cdot \text{card}(S_i)}}} \right)$
 end for
 Propagation
 $k^* \leftarrow \underset{k}{\text{argmin}}(d_k)$
 if $x_i \ni \text{lastNeighborhood}$ **OR** $k^* \neq \text{lastp}$ **then**
 lastNeighborhood \leftarrow around x_i
 lastp $\leftarrow k^*$
 Propagate(t_i, k^*, x_i)
 end if
end procedure

The algorithm is using two layers and achieves real-time analysis of the signal by needing less than half a second to process a one minute trial. This computation time was performed with a single thread Matlab code on a regular desktop computer (Intel Core i7-6700 CPU @ 3.40GHz, 32GB of RAM, running Debian 9.3 Stretch). The algorithm's parameters are registered on table 4.1.

Table 4.1 – This table contains the parameters for the HOTS algorithm to perform the presented results.

\	τ	radius	input polarities	output polarities
Layer 1	250 ms	15	2	14
Layer 2	500 ms	60	14	13

4.2.4 Classification

In their work, Lagorce et al., 2016 introduced the signature of an example as an histogram of polarities. In the case of the pupil, we decided to focus the signatures on the transition between each polarity and build a transition matrix per color displayed that brings a signature close to an Hidden Markov Model (HMM) of the subject's pupil behavior. These complex signatures allow a full description of the subject's pupil behavior by describing stable states as well as unstable states.

The signature of each subject's pupil is a succession of polarities across time. To sort and link these polarities to the projected colors, a 3-D transition matrix is built for each trial of each subject, where the third dimension represents a displayed color and the two first dimensions represent the occurrence of a polarity to follow another, normalized by the amount of polarities while this color is displayed. Figure 4.3 shows that some polarities seem to be color specific, have a role of color transition or even be a stable state for some colors. This representation leads to an effective characterization of the subject's pupil and distinction of a subject from another.

4.3 RESULTS

Details of the methods introduced here can be found in sections 4.2.3 and 4.2.4. We recorded a database of 6 subjects to perform our proof of concept, with twenty examples per subject, and trained a two layers HOTS (Haessig et al., 2019) to achieve our results. The database was randomly divided in a learning group and recognition group, respectively 80%, 20% and we performed a cross-validation. Results are summarized in table 4.2 using the Adjusted Mutual Information (AMI) for reliability and the variance of the cross-validation: Unlike the recognition rate (trace of the confusion matrix divided by the sum of all the elements within the confusion matrix), this measure is adjusted by luck, meaning that a zero value stands for luck and a 1 value for a perfect classification. Vinh, Epps, and Bailey, 2010 highlighted the importance of correcting the recognition rate, especially for small

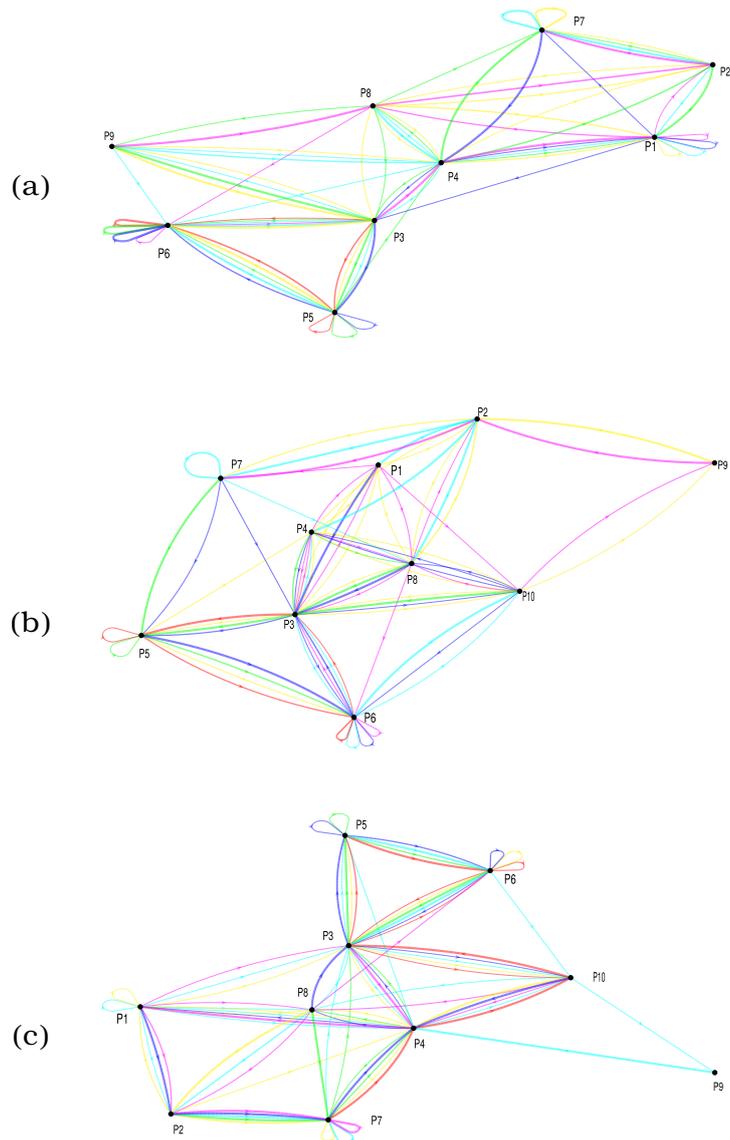


Figure 4.3 – We use a transition matrix to represent the variations of the subject’s pupil dilation/contraction across the presented colors. Here are 3 examples of this matrix, each graph representing the model of a trial for a different subject. Each color in this Hidden Markov Model (HMM) represents the transition between the polarities when this color is sent to the subject’s eye. We can see that some polarities never occur when specific colors are displayed as P_9 and P_8 for red and some polarities have a role of color transition as P_7 for magenta. Some colors seem to induce stable behaviors which are represented by internal loop with high transition value as P_7 for cyan and yellow and P_6 for blue. In each subjects’ graph we can see that some polarities have a common role as P_6 , P_3 and P_5 but with different weights and some others are very different as P_{10} . By computing the distance between these graphs, we are able to recognize a subject from the others.

Table 4.2 – This table contains the classification results of the database. The first two columns present the results for the hidden markov model as a signature, while the two last, for the histogram representation. We show that despite good results on layer 1, the histogram representation does not improve the AMI on layer 2 and maintains a high variance, while the hmm representation leads to a better representation of the information by both increasing the AMI and decreasing the variance.

\	hmm AMI	hmm Variance	hist AMI	hist Variance
Layer 1	0.55	0.057	0.77	0.058
Layer 2	0.94	0.015	0.79	0.045

databases, and provide this measure to fully use the interval of $[0, 1]$ and the values of the confusion matrix as the false positives. As it is showed in table 4.2 the histogram representation performed less than the hmm representation regarding both the AMI and the variance.

4.4 DISCUSSION

We presented in this work the analysis of a new light stimulation as a pupil's behavior based biometric and its proof of concept by achieving a 0.94 AMI on a database of 6 subjects. This preliminary result looks promising and encourages us to recruit more subjects in order to improve the light stimulation protocol. The size of the database does not yet allow us to draw conclusions about the distribution of information during the stimulation, including the highlighting of less significant parts. Moreover, the observed hysteresis in the relation between the projected luminance and the variation of the pupil dynamics suggests that it is difficult to guide an experimental change via the analysis of an isolated segment. The presentation of a condition in a long process and the presentation of this condition isolated could lead to very different experimental results and bring new conclusions. Under its present structure, a light stimulation of 45 seconds as a biometric is too long to be used in daily life while face recognition could be achieved in less than a second. Even though the pupil is a slow organ, by achieving a recognition in the order of the second, this method could be easily paired with a face recognition algorithm when a high security rate is required. By using the transition of the subject's pupil behavior according to the displayed color, the analysis is highly sensitive to blinks that lead to a reset of the light stimulation and increasing the delay to achieve recognition. The light stimulation could be improved by the isolation of the most reliable color transitions to reduce the period of the scan oscillation and reduce

the impact of a blink in the middle of a sequence. A lock-and-key approach could also be considered. The subject would choose a light sequence as key and its pupil would be the lock. In this configuration, the subject's pupil dynamics should aim to verify whether the expected behavior is matching. The HMM representation that we present in this work explore new possibilities to encode complex behaviors from local behaviors and characterize finely the impact of each part of the signal that could make this light stimulation useful in medicine, especially in ophthalmology, to diagnose dyschromatopsia and color deficits.

CONCLUSION

CONCLUSIONS AND DISCUSSION

Computer science and biology were the driving force behind most of my work. The recurrent question of my Ph.D. being: “how would the brain compute a specific task if neurons and in a broader way precise timing are to be used?”

Following this line of thought and using such hypothesis has shown to be a fruitful path. It is the scope of the research group I belong to, and the main topic of my supervisors’ work since more than two decades. In this Ph.D., we initiated a new line of research by going beyond what the team is used to do, that is to infer from pure thought using the hypothesis of precise timing how to compute observed or “thought to be” behavior of the brain.

In my Ph.D., we have decided to introduce a shift in our methodology by allowing ourselves to operate also on biological recorded spikes and not only using artificial spikes from silicon retinas or other sensors. The scope being to be able also to follow the methodology of more traditional fundamental sciences that collect data, derive models and infer the presence of a possible general scheme behind the observed phenomenon.

In my case the natural phenomenon were recorded behaviors possibly with associated neural activities. The scope being to transpose them into an understandable representation by a computer. In this case the “life line” of our thought was to always be able to reliably ensure the validity of our algorithms by matching results with recorded data and the initial hypothesis.

Silicon retinas and other neuromorphic sensors were initially developed to replicate some limited but core functionalities of the brain in a simplified manner and to some extent to match a certain level biological retinas. Operating directly on artificial data output from these sensors allows grasping the general problem to be solved, and explore new possibilities. The hope of the field of neuromorphic engineering is that this “silicon copy” of biology contains sufficient information to unveil some “truth” about the brain that we would be able to scale up in a second stage on more realistic data.

The work presented in section [i](#) aimed to demonstrate such a process. We considered a neuromorphic algorithm computing optical flow that has been initially developed for silicon retinas. We showed that the algorithm efficiently operates on the activity of ganglion cells. The method was able to provide low reconstruction errors despite the fact that silicon retinas do not fully reproduce all the information encoded by the biological retina and the complexity of its cells. Although

designed from a reduced space of information that assumes that every pixel of the retina is performing the same type of acquisition (change detection), results show that considering only a single channel of information that uses time for its encoding sheds light on a general principle that can be robustly transposed to any type of temporal information. We expect all precise timing event-based visual algorithms developed by the lab since its creation to be successfully operating on biological retina data. The real question becomes then, is the brain really computing in the same way? This is a hard question to answer, because it involves being able to record simultaneously from all visual brain areas if one wants to be certain.

This is of course doable, but this is far beyond the capabilities of our lab currently centered on recording the retina. I believe that in the near future once the team will acquire more physiology competence, we will be able to design the exact experiments required to go beyond this initial step. As for now, we can only refer to existing findings in the brain.

It is general knowledge that brains are highly recurrent, and we are deep believers that the type of work that should be pursued must be able to tackle dynamic and adaptive stimulation. We currently use multi electrode arrays that allow us to record the electrical activity of the retina while being stimulated by a video stream. These arrays allow us a global overview of the cells' activity. Whether it is at the retina or brain level, spike sorting is useful as it allows to cluster incoming information and assign activity to classes of cells. However currently this operation is costly and prevents us from setting up adaptive stimuli experiments. Existing efficient algorithms are slow and far from operating in real time. In section [ii](#), we introduced a different approach from what is being developed and used by the field.

Instead of relying on conventional sampled signals, we used an amplitude sampling methodology to map incoming data into the time domain. It allows an adaptive sampling of signals. If something changes fast or slow it will always be possible to extract sufficient samples to describe the incoming signal. It also allows naturally compressing data and removing at the lowest level the signal redundancies while increasing the signal-to-noise ratio. We used in this work a method developed by the lab called Hierarchy of temporal surfaces. The method has been developed to operate on any type of temporal information and it proved to be efficient in spike sorting.

The same method combined with a probabilistic methodology has also been used in chapter [iii](#) where we introduced another application of event based classification tackling a study of a pupil tracking behavior to several light pattern stimuli. We used an event-based approach to study the variations in the expansion and the contraction of the subject's pupil across a scan of the visual light spectrum at a dedicated

luminance. The subject's unique signature relies on the presence of pupil dynamics under specific conditions. Time surfaces seem to be robust descriptors of event based spaces and a valuable multipurpose front end of any type of event based computation.

In the presented experiment we combined time surfaces to describe the dynamics of the pupil to lead in a second stage to a probabilistic graph, each recording the dynamic behavior of subjects. Among the important information, we have shown that each subject had his own duration and transitions, encoding that remain stable across experiments and stimulus. It is reasonable to wonder if this type of signatures can be used in other applications, and to a general extent if they are a reliable solution to this form of dynamic problems. The idea behind event based computation is that time is the most important information, more precisely the time separating two events. It seems that preserving time between events is the main information needed to solve several computational problems in signal processing. The notion of time surface used all along the document provides perhaps a good hint that it is a key representation when solving this type of problems.

As a last observation, although imperfect in many aspects, I believe my Ph.D shows that neuromorphic engineering might be correct in stating that it has the possibility to explain part of the brain. I believe that the general stream of research initiated by our lab that introduced the use of time as a computation feature in several fields such as computer vision, robotics and prosthesis could not have happened without the silicon retinas. However, it seems that the field of engineering is facing the constant temptation of staying on the zone of comfort of what is known. Many papers tend to ignore time and use these sensors to generate frames by summing events, or using crude binary activity images. The scope is to avoid thinking in the time domain and recycle what their fields has been developing for the last decades. Although disappointing, this behavior can be explained by the fact that there is currently no easily programmable and available hardware able to make full use of the temporal properties of event based sensors. I am convinced that in the near future, the power of operating in the time domain, and writing algorithms that deal with every spike separately and locally will prevail as soon as such a hardware will exist. Promising work is ongoing, to cite a few, Intel's Loihi, Ibm's Truenorth, GML's Amsterdam, and many others. . .

I worked with biologists, neuroscientists, psychophysicists and computer scientists to understand problems linked to the eye and especially the retina to explore methods and solutions able to answer them. During my Ph.D, I was also taught biology and physiology that allowed me to perform dissections and recordings on rat's retinas.

I understand now, how difficult and time consuming it is to record useful data, and the endless amount of external parameters that can disrupt an experiment. As a final comment, it feels important to me to emphasize that this type of research will necessitate a new kind of engineers and not necessarily only bioengineers but most likely a more hybrid version of what is being trained currently.



[December 11, 2018 at 19:30 – classicthesis v4.6]

Part IV
APPENDIX

EVENT-BASED WATERMARKING

During my thesis, I work on a side project to find a method able to apply a robust watermark on event based data. Due to the asynchronous time of events, frequency analysis usually performed in this field is not possible and I aimed to evaluate a personal method able to achieve good invariant properties compatible with a asynchronous time based cloud of points. This work is on the appendix section to show what the method is able to achieve at the present moment, even through it is not finished yet. I demonstrated that the method is robust to time shift, to linearly time dilation and to the addition of a stationary blank noise on the timestamp values according some properties of the used key. At the end of this work, I exposed some possible attacks able to break the method and when it is possible some ideas to countermeasure them.

A.1 INTRODUCTION



Figure A.1 – Example of a visible watermark. An image is added to the original image, that make it unusable.

Figure A.2 – Example of an invisible watermark. The blue channel is modified but it is not visible on the RGB image.

A watermark is a print on a signal that allows its identification. It can be “visible” (fig A.1) and therefore destroy a small or big part of the signal, hence preventing its copy or fraudulent use. It can be “invisible” (fig A.2) and allows a proper use of the signal, but as it is marked one can still determine its origin. For instance, let us consider that A sells a product M to B and C. M is differently marked for B and C. Two months later, a third party D is found using product M. A proper watermarking makes it possible to find the source of

that leak and trace it back whether to A, B or C. The purpose of an effective watermarking is to be difficult to erase or get around. The ideal is a method which is difficult to crack for a third party but easy to control for a certified user. Robust methods to create invisible watermarking for audio and images use wavelet transforms to get a scalable frequency coding. No methods seem to exist to watermark a cloud of points without a sampling rate.

A.2 HYPOTHESIS

An event has a unique signature composed of a timestamp, a spatial position, a polarity and other additional information such as triggers, measurements, etc. An efficient watermarking must be robust to the different use of the signal and more importantly accessible no matter what the user is using the signal for. The events' additional information are too often discarded in many algorithms and can not make suitable candidates for watermarking. Nor can the polarity of an event as current event-base methods, for instance in optical flow computation, manage not to take it into account. The spatial position of an event is a suitable candidate for watermarking. However this would imply that the encryption is robust to any morphological transformations. The only remaining and best feature is the timestamp of an event which stands out to be the easier to control and yet the harder to corrupt.

In this study, the choice is to focus on a set of random pixels, which are not a part of the key, and use a key of delays on their timestamp signatures.

A.3 METHOD

The aim of the method is to add a local shift in the timestamps of a signal. This corruption leads to build a unique sequence of delays in the signal and use this uniqueness as a reliable authentication key.

Let assume T , a vector of timestamps as:

$$T := (t_i)_{0 \leq i \leq n-1} \quad (\text{A.1})$$

where $n \in \mathbb{N}$, and the transformation E as:

$$E := \{(I_i); (M_i); r\} \quad (\text{A.2})$$

$$E : T \rightarrow \tilde{T} \quad (\text{A.3})$$

where I is a set of index starting from the reference r and M a set of time shifts, called the message, that modify the signal for each I (See figure A.4). E leads to define the key as the duet K :

$$K = \{(I_i)_{1 \leq i \leq k}; (\Delta t_i)_{1 \leq i \leq k}\} \quad (\text{A.4})$$

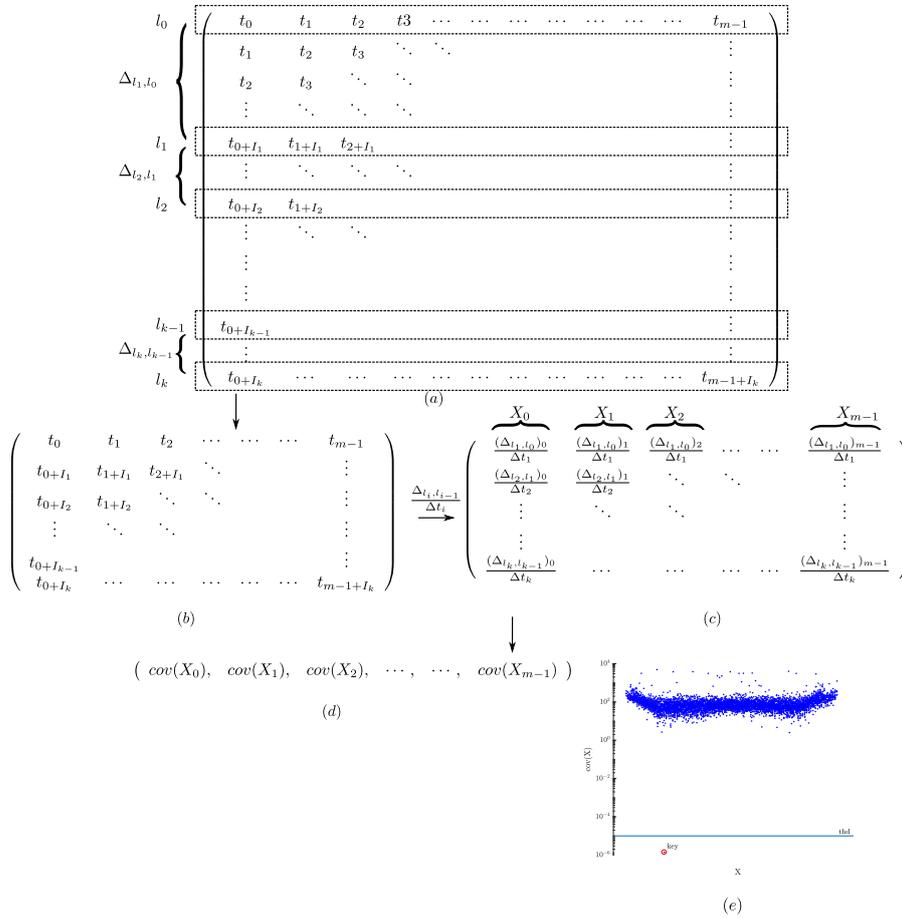


Figure A.3 – The key (Eq A.4) is a sequence of paired information as delays at expected indexations. In order to detect it, let see the vector t of size n (Eq A.1) as the matrix described in (a). A matrix with $I_k + 1$ lines and m column as $m = n - (I_k + 1)$. The first line, called l_0 , represents the beginning of each sequence that could contain the key. By beginning with this line, lines $l_i + j$ are extracted to become the submatrix (b) of size $(k + 1) * m$. Line by line are subtracted and for each subtraction, the expected delay divided each element that compose it, as show in (c). The matrix becomes $k * m$. The position of the key is detected using the covariance on each column of (c), with a confident threshold as showed by (d) and (e).

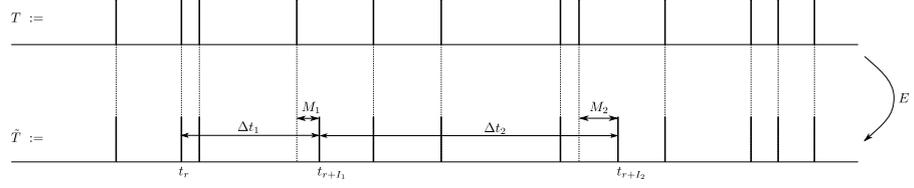


Figure A.4

where $k \in \mathbb{N}$ is the size of the message M and $r + I_k \in \tilde{T}$. The difference Δt_i represents the delay between t_{r+I_i} and $t_{r+I_{i-1}}$ with $I_0 = 0$.

This method is invariant to the following transformation:

$$b_i = t_i k_* + k_c + \lambda_i'' \quad (\text{A.5})$$

Where $k_c = k_1 k_2 + k_3$ represents a constant time shift, $k_* = k_2$ a constant acceleration/deceleration of the signal and $\lambda_i'' = \lambda_i k_2 + \lambda_i'$ an added random noise centered in μ , with a variance of σ^2 .

If $\forall i \in \mathbb{Z}$, $\lambda_i'' = 0$, then the following transformation guarantees the invariance to k_c and k_* :

$$\begin{cases} \phi_{1j} = b_{i+I_{j+1}} - b_{i+I_j} = k_*(t_{i+I_{j+1}} - t_{i+I_j}) = k_* \Delta t_j \\ \phi'_{1j} = \frac{\phi_{1j}}{\Delta t_j} = k_* \end{cases} \quad (\text{A.6})$$

By applying ϕ'_{1j} for each key members the variance of ϕ'_{1j} goes to zero (not strictly zero because of numerical encoding).

By considering the noise we get:

$$\begin{cases} \phi_{2j} = k_* \Delta t_j + \lambda_{i+I_{j+1}}'' - \lambda_{i+I_j}'' \\ \phi'_{2j} = k_* + \frac{\lambda_{i+I_{j+1}}'' - \lambda_{i+I_j}''}{\Delta t_j} \end{cases} \quad (\text{A.7})$$

By choosing $\forall j \in \mathbb{Z}$, $\Delta t_j \gg \lambda_{i+I_{j+1}}'' - \lambda_{i+I_j}''$, $\phi'_{2j} \rightarrow k_*$. By increasing the size of the message, the variance of ϕ'_{2j} decreases, however it still be linked to the variance of λ'' . Moreover, it increases the sensitivity of the key against corruption (See A.4).

To reduce this impact of the variance of λ'' , it is more reliable to look at the variance of the local average along the ϕ'_{2j} .

$$\begin{cases} \forall j \in \mathbb{Z}, \prod_L = \begin{cases} \frac{1}{L} & \text{if } 0 \leq j < L \\ 0 & \text{otherwise} \end{cases} \\ \phi_{3j} = \sum_{k=0}^{N-1} \phi'_{2k} \prod_{L-j-k} = k_* + \epsilon_j \end{cases} \quad (\text{A.8})$$

Let assume \prod_L an average filter of size L , ϕ_{3j} is the local average of b_j over the key. By applying a threshold on the variance of ϕ_{3j} , we are able to distinguish the key from the rest of the signal (See fig A.3).

A.4 KNOWN WEAKNESSES

- If the added noise is not stationary on the key, the reconstruction fails. (Answers: Add several keys at different positions helps)
- Truncate the signal to an other precision (ms for example) erases the key. (Answers: **unsolved**, use delays larger than the ms increase a lot the consuming time of the algorithm and make the uniqueness the key harder to ensure)
- Add/Remove events in the tracked sequence breaks it. (Answers: indexation could be relative at $\pm k$ but more than that we need to classify the level of degradation that the signal is able to ensure before the watermark falls)
- Add a second watermark to the signal could damage the previous one. (Answers: **unsolved**)

BIBLIOGRAPHY

- Akolkar, Himanshu, Cedric Meyer, Xavier Clady, Olivier Marre, Chiara Bartolozzi, Stefano Panzeri, and Ryad Benosman (2015). « What can neuromorphic event-driven precise timing add to spike-based pattern recognition? » In: *Neural computation*.
- Barbakh, Wesam and Colin Fyfe (2008). « Online clustering algorithms. » In: *International Journal of Neural Systems* 18.03, pp. 185–194.
- Bednarik, Roman, Tomi Kinnunen, Andrei Mihaila, and Pasi Fränti (2005). « Eye-movements as a biometric. » In: *Scandinavian conference on image analysis*. Springer, pp. 780–789.
- Bendali, Amel, Lionel Rousseau, Gaëlle Lissorgues, Emmanuel Scorsone, Milan Djilas, Julie Dégardin, Elisabeth Dubus, Stéphane Fouquet, Ryad Benosman, Philippe Bergonzo, et al. (2015). « Synthetic 3D diamond-based electrodes for flexible retinal neuroprostheses: Model, production and in vivo biocompatibility. » In: *Biomaterials* 67, pp. 73–83.
- Benosman, Ryad, Charles Clercq, Xavier Lagorce, Sio-Hoi Ieng, and Chiara Bartolozzi (2014). « Event-based visual flow. » In: *Neural Networks and Learning Systems, IEEE Transactions on* 25.2, pp. 407–417.
- Berry, Michael J, Iman H Brivanlou, Thomas A Jordan, and Markus Meister (1999). « Anticipation of moving stimuli by the retina. » In: *Nature* 398.6725, pp. 334–338.
- Bhattacharyya, Anil (1946). « On a measure of divergence between two multinomial populations. » In: *Sankhyā: the indian journal of statistics*, pp. 401–406.
- Brandli, Alice, Chi D Luu, Robyn H Guymer, and Lauren N Ayton (2016). « Progress in the clinical development and utilization of vision prostheses: an update. » In: *Eye and brain* 8, p. 15.
- Butts, Daniel A, Chong Weng, Jianzhong Jin, Chun-I Yeh, Nicholas A Lesica, Jose-Manuel Alonso, and Garrett B Stanley (2007). « Temporal precision in the neural code and the timescales of natural vision. » In: *Nature* 449.7158, pp. 92–95.
- Calford, Mike B, Layne L Wright, Andrew B Metha, and Vivian Taglianetti (2003). « Topographic plasticity in primary visual cortex is mediated by local corticocortical connections. » In: *Journal of Neuroscience* 23.16, pp. 6434–6442.
- Castaldi, Elisa, Guido Marco Cicchini, Laura Cinelli, Laura Biagi, Stanislao Rizzo, and Maria Concetta Morrone (2016). « Visual BOLD response in late blind subjects with Argus II retinal prosthesis. » In: *PLoS biology* 14.10, e1002569.

- Cha, Kichul, Kenneth Horch, and Richard A Normann (1992). « Simulation of a phosphene-based visual field: visual acuity in a pixelized vision system. » In: *Annals of biomedical engineering* 20.4, pp. 439–449.
- Cha, Kichul, Kenneth W Horch, Richard A Normann, and Duane K Boman (1992). « Reading speed with a pixelized vision system. » In: *JOSA A* 9.5, pp. 673–677.
- Cheng, Daniel and Roel Vertegaal (2004). « An eye for an eye: a performance evaluation comparison of the lc technologies and tobii eye trackers. » In: *Proceedings of the 2004 symposium on Eye tracking research & applications*. ACM, pp. 61–61.
- Chicca, E., F. Stefanini, C. Bartolozzi, and G. Indiveri (2014). « Neuromorphic electronic circuits for building autonomous cognitive systems. » In: *Proceedings of the IEEE* 102, pp. 1367–1388.
- Chichilnisky, EJ and RS Kalmar (2003). « Temporal resolution of ensemble visual motion signals in primate retina. » In: *Journal of Neuroscience* 23.17, pp. 6681–6689.
- Corradi, Federico and Giacomo Indiveri (2015). « A neuromorphic event-based neural recording system for smart brain-machine-interfaces. » In: *IEEE transactions on biomedical circuits and systems* 9.5, pp. 699–709.
- Cover, Thomas and Peter Hart (1967). « Nearest neighbor pattern classification. » In: *IEEE transactions on information theory* 13.1, pp. 21–27.
- Czajka, Adam (2015). « Pupil Dynamics for Iris Liveness Detection. » In: *IEEE Trans. Information Forensics and Security* 10.4, pp. 726–735.
- Da Cruz, Lyndon, Brian F Coley, Jessy Dorn, Francesco Merlini, Eugene Filley, Punita Christopher, Fred K Chen, Varalakshmi Wuyyuru, Jose Sahel, Paulo Stanga, et al. (2013). « The Argus II epiretinal prosthesis system allows letter and word reading and long-term function in patients with profound vision loss. » In: *British Journal of Ophthalmology*, bjophthalmol–2012.
- Davies, M. et al. (2018). « Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. » In: *IEEE Micro* 38.1, pp. 82–99. ISSN: 0272-1732. DOI: [10.1109/MM.2018.112130359](https://doi.org/10.1109/MM.2018.112130359). URL: doi.ieeecomputersociety.org/10.1109/MM.2018.112130359.
- Erdogmus, Nesli and Sébastien Marcel (2013). « Spoofing 2D face recognition systems with 3D masks. » In: *Biometrics Special Interest Group (BIOSIG), 2013 International Conference of the. IEEE*, pp. 1–8.
- Erdogmus, Nesli and Sebastien Marcel (2014). « Spoofing face recognition with 3D masks. » In: *IEEE transactions on information forensics and security* 9.7, pp. 1084–1097.
- Evans, Nicholas, Stan Z Li, Sebastien Marcel, and Arun Ross (2015). « Guest editorial: Special issue on biometric spoofing and countermeasures. » In: *IEEE Transactions on Information forensics and security* 10.4, pp. 699–702.

- Faundez-Zanuy, Macros (2004). « On the vulnerability of biometric security systems. » In: *IEEE Aerospace and Electronic Systems Magazine* 19.6, pp. 3–8.
- Fornos, Angélica Pérez, Jörg Sommerhalder, Alexandre Pittard, Avinoam B Safran, and Marco Pelizzone (2008). « Simulation of artificial vision: IV. Visual information required to achieve simple pointing and manipulation tasks. » In: *Vision research* 48.16, pp. 1705–1718.
- Frechette, ES, A Sher, MI Grivich, D Petrusca, AM Litke, and EJ Chichilnisky (2005). « Fidelity of the ensemble code for visual motion in primate retina. » In: *Journal of neurophysiology* 94.1, pp. 119–135.
- Furber, S.B., F. Galluppi, S. Temple, and L.A. Plana (2014). « The spinnaker project. » In: *Proceedings of the IEEE* 102.5, pp. 652–665.
- Garcia, Diogo Caetano and Ricardo L de Queiroz (2015). « Face-spoofing 2D-detection based on moiré-pattern analysis. » In: *IEEE transactions on information forensics and security* 10.4, pp. 778–786.
- Gasthaus, Jan A (2008). « Spike sorting using time-varying DIRICHLET process mixture models. » PhD thesis. Citeseer.
- Georgopoulos, Apostolos P, Andrew B Schwartz, and Ronald E Kettner (1986). « Neuronal population coding of movement direction. » In: *Science*, pp. 1416–1419.
- Giannikopoulos, Dimitrios V and Ulf T Eysel (2006). « Dynamics and specificity of cortical map reorganization after retinal lesions. » In: *Proceedings of the National Academy of Sciences* 103.28, pp. 10805–10810.
- Gilbert, Charles D and Torsten N Wiesel (1992). « Receptive field dynamics in adult primary visual cortex. » In: *Nature* 356.6365, p. 150.
- Giulioni, Massimiliano, Xavier Lagorce, Francesco Galluppi, and Ryad B Benosman (2016). « Event-based computation of motion flow on a neuromorphic analog neural platform. » In: *Frontiers in neuroscience* 10.
- Graganiello, Diego, Giovanni Poggi, Carlo Sansone, and Luisa Verdoliva (2015). « An investigation of local descriptors for biometric spoofing detection. » In: *IEEE transactions on information forensics and security* 10.4, pp. 849–863.
- Harris, Kenneth D, Darrell A Henze, Jozsef Csicsvari, Hajime Hirase, and György Buzsáki (2000). « Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. » In: *Journal of neurophysiology* 84.1, pp. 401–414.
- Heinen, SJ and AA Skavenski (1991). « Recovery of visual responses in foveal V1 neurons following bilateral foveal lesions in adult monkey. » In: *Experimental Brain Research* 83.3, pp. 670–674.
- Henze, D, KD Harris, Z Borhegyi, J Csicsvari, A Mamiya, H Hirase, A Sirota, and G Buzsáki (2009). *Simultaneous intracellular and extracellular recordings from hippocampus region CA1 of anesthetized rats.*

- Henze, Darrell A, Zsolt Borhegyi, Jozsef Csicsvari, Akira Mamiya, Kenneth D Harris, and György Buzsáki (2000). « Intracellular features predicted by extracellular recordings in the hippocampus in vivo. » In: *Journal of neurophysiology* 84.1, pp. 390–400.
- Indiveri, G., F. Corradi, and N. Qiao (2015). « Neuromorphic architectures for spiking deep neural networks. » In: *IEEE Electron Devices Meeting (IEDM)*, pp. 1–4.
- Jin, Qin, Arthur R Toth, Alan W Black, and Tanja Schultz (2008). « Is voice transformation a threat to speaker identification? » In: *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, pp. 4845–4848.
- Johansson, R.S. and I. I. Birznieks (2004). « First spikes in ensembles of human tactile afferents code complex spatial fingertip events. » In: *Nat Neurosci* 7, pp. 170–177.
- Joucla, Sébastien and Blaise Yvert (2009). « Improved focalization of electrical microstimulation using microelectrode arrays: a modeling study. » In: *PloS one* 4.3, e4828.
- Kadir, Shabnam N, Dan FM Goodman, and Kenneth D Harris (2014). « High-dimensional cluster analysis with the masked EM algorithm. » In: *Neural computation*.
- Komogortsev, Oleg V, Alexey Karpov, and Corey D Holland (2015). « Attack of mechanical replicas: Liveness detection with eye movements. » In: *IEEE Transactions on Information Forensics and Security* 10.4, pp. 716–725.
- Lagorce, Xavier, Garrick Orchard, Francesco Gallupi, Bertram E Shi, and Ryad Benosman (2016). « HOTS: A Hierarchy Of event-based Time-Surfaces for pattern recognition. » In:
- Lapuerta, Pablo and Stanley J Schein (1995). « A four-surface schematic eye of macaque monkey obtained by an optical method. » In: *Vision research* 35.16, pp. 2245–2254.
- Li, Peter H, Jeffrey L Gauthier, Max Schiff, Alexander Sher, Daniel Ahn, Greg D Field, Martin Greschner, Edward M Callaway, Alan M Litke, and EJ Chichilnisky (2015). « Anatomical identification of extracellularly recorded cells in large-scale multielectrode recordings. » In: *The Journal of Neuroscience* 35.11, pp. 4663–4675.
- Liu, Robert C., Svilen Tzonev, Sergei P. Rebrük, and K. D. Miller (2001). « Variability and information in a neural code of the cat lateral geniculate nucleus. » In: *Journal of neurophysiology* 86 6, pp. 2789–806.
- Lorach, Henri, Georges Goetz, Richard Smith, Xin Lei, Yossi Mandel, Theodore Kamins, Keith Mathieson, Philip Huie, James Harris, Alexander Sher, et al. (2015). « Photovoltaic restoration of sight with high visual acuity. » In: *Nature medicine* 21.5, p. 476.
- Lorach, Henri, Jenny Wang, Dae Yeong Lee, Roopa Dalal, Philip Huie, and D Palanker (2016). « Retinal safety of near infrared radiation in photovoltaic restoration of sight. » In: *Biomedical optics express* 7.1, pp. 13–21.

- Maass, Wolfgang (1999). « Computing with spiking neurons. » In: *Pulsed neural networks* 85.
- Maaten, Laurens van der and Geoffrey Hinton (2008). « Visualizing data using t-SNE. » In: *Journal of Machine Learning Research* 9.Nov, pp. 2579–2605.
- MacQueen, J. (1967). « Some methods for classification and analysis of multivariate observations. » In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Berkeley, Calif.: University of California Press, pp. 281–297.
- Mainen, ZF and TJ Sejnowski (1995). « Reliability of spike timing in neocortical neurons. » In: *Science* 268, pp. 1503–1506.
- Mariotti, Silvio Paolo (2010). *Global data on visual impairments*. World Health Organization. URL: www.who.int/blindness/GLOBALDATAFINALforweb.pdf.
- Marre, Olivier, Vicente Botella-Soler, Kristina D Simmons, Thierry Mora, Gašper Tkačik, and Michael J Berry II (2015). « High accuracy decoding of dynamical motion from a large retinal population. » In: *PLoS computational biology* 11.7, e1004304.
- Mathieson, Keith, James Loudin, Georges Goetz, Philip Huie, Lele Wang, Theodore I Kamins, Ludwig Galambos, Richard Smith, James S Harris, Alexander Sher, et al. (2012). « Photovoltaic retinal prosthesis with high pixel density. » In: *Nature photonics* 6.6, p. 391.
- Merolla, Paul A, John V Arthur, Rodrigo Alvarez-Icaza, Andrew S Cassidy, Jun Sawada, Filipp Akopyan, Bryan L Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, et al. (2014). « A million spiking-neuron integrated circuit with a scalable communication network and interface. » In: *Science* 345.6197, pp. 668–673.
- Murakami, Ikuya, Hidehiko Komatsu, and Masaharu Kinoshita (1997). « Perceptual filling-in at the scotoma following a monocular retinal lesion in the monkey. » In: *Visual neuroscience* 14.1, pp. 89–101.
- Neftci, E. et al. (2013). « Synthesizing cognition in neuromorphic electronic systems. » In: *Proceedings of the National Academy of Sciences*, pp. 3468–3476.
- Posch, Christoph, Daniel Matolin, and Rainer Wohlgenannt (2011). « A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS. » In: *Solid-State Circuits, IEEE Journal of* 46.1, pp. 259–275.
- Prabhakar, Salil, Sharath Pankanti, and Anil K Jain (2003). « Biometric recognition: Security and privacy concerns. » In: *IEEE security & privacy* 2, pp. 33–42.
- Qiao, Ning, Hesham Mostafa, Federico Corradi, Marc Osswald, Fabio Stefanini, Dora Sumislawska, and Giacomo Indiveri (2015). « A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses. » In: *Frontiers in neuroscience* 9, p. 141.

- Quiroga, R Quian, Zoltan Nadasdy, and Yoram Ben-Shaul (2004). « Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. » In: *Neural computation* 16.8, pp. 1661–1687.
- R.S., Petersen, S. Panzeri, and M.E. Diamond (2001). « Population Coding of Stimulus Location in Rat Somatosensory Cortex. » In: *Neuron* 32, pp. 503–414.
- Recce, M and J O'keefe (1989). « The tetrode: a new technique for multi-unit extracellular recording. » In: *Soc Neurosci Abstr.* Vol. 15. 2, p. 1250.
- Reichardt, Werner (1961). « Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. » In: *Sensory communication*, pp. 303–317.
- Reinagel, Pamela and R. Clay Reid (2000). « Temporal Coding of Visual Information in the Thalamus. » In: *Journal of Neuroscience* 20, pp. 5392–400.
- Rieke, Fred, Davd Warland, Rob de Ruyter van Steveninck, and William Bialek (1999). *Spikes: Exploring the Neural Code*. Cambridge, MA, USA: MIT Press. ISBN: 0-262-18174-6.
- Rossant, Cyrille, Shabnam N Kadir, Dan FM Goodman, John Schulman, Mariano Belluscio, Gyorgy Buzsaki, and Kenneth D Harris (2015). « Spike sorting for large, dense electrode arrays. » In: *Nat Neurosci*, pp. 634–641.
- Sanchez, Jon, Ibon Saratxaga, Inma Hernaez, Eva Navas, Daniel Erro, and Tuomo Raitio (2015). « Toward a universal synthetic speech spoofing detection using phase information. » In: *IEEE Transactions on Information Forensics and Security* 10.4, pp. 810–820.
- Serrano-Gotarredona, R. et al. (2009). « CAVIAR: A 45K Neuron, 5M Synapse, 12G Connects AER Hardware Sensory-Processing-Learning-Actuating System for High-Speed Visual Object Recognition and Tracking. » In: *IEEE Transactions on Neural Networks*, pp. 1417–1438.
- Shahid, Shahjahan, Jacqueline Walker, and Leslie S Smith (2010). « A new spike detection algorithm for extracellular neural recordings. » In: *IEEE Transactions on Biomedical Engineering* 57.4, pp. 853–866.
- Silverman, MS and SE Hughes (1989). « Transplantation of photoreceptors to light-damaged retina. » In: *Investigative ophthalmology & visual science* 30.8, pp. 1684–1690.
- Sizov, Aleksandr, Elie Khoury, Tomi Kinnunen, Zhizheng Wu, and Sébastien Marcel (2015). « Joint Speaker Verification and Antispoofing in the i-Vector Space. » In: *IEEE Transactions on Information Forensics and Security* 10.4, pp. 821–832.
- Smith, Daniel F, Arnold Wiliem, and Brian C Lovell (2015). « Face recognition on consumer devices: Reflections on replay attacks. » In: *IEEE Transactions on Information Forensics and Security* 10.4, pp. 736–745.

- Smith, RC, GT Smith, and D Wong (1990). « Refractive changes in silicone filled eyes. » In: *Eye* 4.1, p. 230.
- Sommerhalder, Jörg, Evelyne Oueghlani, Marc Bagnoud, Ute Leonards, Avinoam B Safran, and Marco Pelizzone (2003). « Simulation of artificial vision: I. Eccentric reading of isolated words, and perceptual learning. » In: *Vision Research* 43.3, pp. 269–283.
- Sommerhalder, Jörg, Benjamin Rappaz, Raoul de Haller, Angélica Pérez Fornos, Avinoam B Safran, and Marco Pelizzone (2004). « Simulation of artificial vision: II. Eccentric reading of full-page text and the learning of this task. » In: *Vision research* 44.14, pp. 1693–1706.
- Stingl, Katarina, Ruth Schippert, Karl U Bartz-Schmidt, Dorothea Besch, Charles L Cottrill, Thomas L Edwards, Florian Gekeler, Udo Greppmaier, Katja Kiel, Assen Koitschev, et al. (2017). « Interim results of a multicenter trial with the new electronic subretinal implant Alpha AMS in 15 patients blind from inherited retinal degenerations. » In: *Frontiers in neuroscience* 11, p. 445.
- Thorpe, Simon J., Arnaud Delorme, and Rufin VanRullen (2001). « Spike-based strategies for rapid processing. » In: *NEURAL NETWORKS* 14, pp. 715–725.
- Thorpe, Simon (1990). « Spike arrival times: A highly efficient coding scheme for neural networks. » In: *Parallel processing in neural systems*.
- Tirunagari, Santosh, Norman Poh, David Windridge, Aamo Iorliam, Nik Suki, and Anthony TS Ho (2015). « Detection of face spoofing using visual dynamics. » In: *IEEE transactions on information forensics and security* 10.4, pp. 762–777.
- Villalobos-Castaldi, Fabiola M and Ernesto Suaste-Gómez (2013). « A Video-Based Spatio-temporal Biometric Template Representation of the Spontaneous Pupillary Oscillations: A Pilot Experiment. » In: *Mexican Conference on Pattern Recognition*. Springer, pp. 165–174.
- Vinh, Nguyen Xuan, Julien Epps, and James Bailey (2010). « Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. » In: *The Journal of Machine Learning Research* 11, pp. 2837–2854.
- Wang, Chin-An and Douglas P Munoz (2015). « A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. » In: *Current opinion in neurobiology* 33, pp. 134–140.
- Weinrich, Tobias W, Michael B Powner, Aisling Lynch, Ravi S Jonnal, John S Werner, and Glen Jeffery (2017). « No evidence for loss of short-wavelength sensitive cone photoreceptors in normal ageing of the primate retina. » In: *Scientific reports* 7, p. 46346.
- Wen, Di, Hu Han, and Anil K Jain (2015). « Face spoof detection with image distortion analysis. » In: *IEEE Transactions on Information Forensics and Security* 10.4, pp. 746–761.
- Werner, Thilo, Elisa Vianello, Olivier Bichler, Daniele Garbin, Daniel Cattaert, Blaise Yvert, Barbara De Salvo, and Luca Perniola (2016). « Spiking Neural Networks based on OxRAM Synapses for Real-

- time Unsupervised Spike Sorting. » In: *Frontiers in Neuroscience* 10, p. 474.
- Wild, Jiri, Zoltan Prekopcsak, Tomas Sieger, Daniel Novak, and Robert Jech (2012). « Performance comparison of extracellular spike sorting algorithms for single-channel recordings. » In: *Journal of neuroscience methods* 203.2, pp. 369–376.
- Williams, David R (2011). « Imaging single cells in the living retina. » In: *Vision research* 51.13, pp. 1379–1396.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both L^AT_EX and L^YX:

<https://bitbucket.org/amiede/classicthesis/>

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

<http://postcards.miede.de/>

Thank you very much for your feedback and contribution.