

AIX-MARSEILLE UNIVERSITÉ

MÉMOIRE DE MASTER

Optimisation de la vision artificielle bio-inspirée par exploration saccadique de l'environnement

Auteur:

Pierre ALBIGÈS

Superviseur:

Laurent PERRINET

Un mémoire présenté à

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

en vue de l'obtention du diplôme de

MASTER DE NEUROSCIENCES, SPÉCIALITÉ INTÉGRATIVES ET COGNITIVES

et réalisée au sein de

Institut de Neurosciences de la Timone

Durant la période : 12/03/2018 - 08/06/2018

AIX-MARSEILLE UNIVERSITÉ

Résumé

Faculté des Sciences, département de Biologie

Master de Neurosciences

Master de Neurosciences, spécialité Intégratives et Cognitives

Optimisation de la vision artificielle bio-inspirée par exploration saccadique de l'environnement

by Pierre ALBIGÈS

Contents

1	Introduction	1
2	Matériel et méthodes	3
3	Résultats	4
4	Discussion et perspectives	5
	Bibliography	6
A	Figures	7
B	Code source et documents complémentaires	8

1 Introduction

Au cours de l'histoire évolutive et sous la pression de la sélection naturelle, tous nos systèmes perceptifs ont tendu (et tendent encore) vers une optimisation de leurs performances, en fonction de nos besoins et de nos ressources. L'ensemble de notre système visuel, de la rétine jusqu'aux aires cérébrales les plus associatives, a ainsi évolué pour arriver au fonctionnement à la fois rapide et efficace qu'on lui connaît aujourd'hui.³

La variabilité de l'acuité (la précision avec laquelle les stimuli visuels pourront être analysés) au sein de notre champ visuel (portion de l'espace observée par un oeil immobile) est l'un des produits de cette pression évolutive. On peut ainsi séparer le champ visuel en deux parties : la vision centrale et la vision périphérique.³

La vision centrale est soutenue anatomiquement par la fovéa, une région rétinienne comprenant exclusivement des cônes. Cette composition, couplée à une forte densité de photorécepteurs permet à cette région de présenter l'acuité visuelle la plus importante, ainsi qu'une bonne perception des couleurs.³

La composition et la densité en photorécepteurs de la rétine soutenant la vision périphérique change avec son excentricité par rapport à la fovéa, mais elle comprends majoritairement des bâtonnets. En conséquence, l'acuité visuelle et la perception des couleurs dans la vision périphérique diminuent avec la distance de la fovéa, mais on peut y observer une importante sensibilité aux variations de luminance et de fréquence spatiale.³

Cette variabilité des caractéristiques de notre système visuel, et notamment de son acuité, permet de fortement réduire la quantité d'informations que doivent traiter les réseaux nerveux, passant d'un flux arrivant à la rétine estimé à 10^8 bits/s à une sortie par le nerf optique estimé à 10^2 bits/s.^{1,3,4}

Mais cette optimisation du flux d'informations présente au moins un inconvénient majeur. Une description précise d'un stimulus visuel ne peut être réalisée avec une certitude élevée que dans une partie très réduite du champ visuel (environ 2 degrés chez l'Humain).

Ainsi lors de l'exploration visuelle de son environnement, un agent va devoir réaliser une suite de

mouvements oculaires brefs afin de placer les régions visuelles d'intérêt dans sa vision centrale et ainsi pouvoir en réaliser des descriptions précises. Par exemple, l'observation passive d'une scène va impliquer la réalisation de 2 à 4 saccades par seconde.^{2,3}

La sélection attentionnelle et motrice de la cible à décrire fait intervenir un réseau complexe d'aires cérébrales et corticales et nécessite l'intégration d'influences *top-down* comme *bottom-up*.³

Nous pouvons donc décrire la vision comme un processus actif, impliquant des notions de perception, de traitement de l'information et d'actions motrices.³

Depuis les débuts de l'intelligence artificielle dans les années 60, l'un des domaines phares de son développement a été la vision artificielle. L'objectif est de s'inspirer, voir de mimer les systèmes biologiques afin de permettre aux systèmes informatisés d'accéder à la compréhension de leur environnement. La modélisation de l'activité du système visuel a ainsi connu l'application de nombreuses méthodes, à diverses échelles et niveaux de complexité.³

Les modèles à carte de saillance permettent par exemple de reconstruire l'influence qu'ont les signaux *bottom-up* sur l'orientation du regard. Ces modèles décrivant chaque point de l'espace visuel par une valeur, ceux qui ressortent le plus de l'environnement sont considérés comme portant l'intérêt le plus grand pour le système, et attirent donc le regard. Lorsqu'une région est explorée, sa saillance devient ensuite nulle.³

Les modèles de saillance présentent deux inconvénients majeurs. Leurs prédictions sont meilleures que le hasard sans être exactes, notamment car elles ne prennent pas en compte la recherche de cible. Leur deuxième inconvénient est que ces modèles considèrent l'ensemble de l'environnement visuel à la fois et avec la même valeur, ce qui est loin de ce qui se déroule dans les systèmes biologiques.³

Ces modèles permettent non seulement d'améliorer les performances des systèmes informatisés, mais aussi de mieux appréhender les zones d'ombre dans nos connaissances du fonctionnement du système visuel.³

2 Matériel et méthodes

3 Résultats

4 Discussion et perspectives

Bibliography

- [1] Philip Kortum and Wilson S. Geisler. “Implementation of a foveated image coding system for image bandwidth reduction”. In: *SPIE Proceedings* 2657 (1996), pp. 350–360. ISSN: 0277786X. DOI: [10.1117/12.238732](https://doi.org/10.1117/12.238732).
- [2] Richard J. Krauzlis, Laurent Goffart, and Ziad M. Hafed. “Neuronal control of fixation and fixational eye movements”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1718 (2017), p. 20160205. ISSN: 0962-8436. DOI: [10.1098/rstb.2016.0205](https://doi.org/10.1098/rstb.2016.0205). URL: <http://rstb.royalsocietypublishing.org/lookup/doi/10.1098/rstb.2016.0205>.
- [3] John S. Werner and Leo M. Chalupa, eds. *The new visual neurosciences*. MIT Press. 2014, p. 1675. ISBN: 9780262019163.
- [4] Li Zhaoping. *Understanding vision : theory, models and data*. Oxford Uni. 2014, p. 383. ISBN: 9780199564668.

A Figures

B Code source et documents complémentaires

L'ensemble du code source du modèle sous forme de ipython notebooks, de ce rapport au format \LaTeX ainsi que de l'ensemble des autres documents issus de ce travail (dont les notes personnelles) sont entièrement disponibles **en ligne** ou en contactant directement l'**auteur**.