



Sujet de stage M2R Sciences Cognitives

De la rétine binoculaire aux premières étapes du cortex visuel pour la perception visuelle tridimensionnelle : approches par modélisation et expérience en oculométrie

L'objectif du stage est de proposer un modèle computationnel prédictif des mouvements oculaires dans un contexte d'exploration de scènes tridimensionnelles en s'appuyant d'une part sur des modèles existants et déjà développés au laboratoire pour des scènes bidimensionnelles et d'autre part sur des données expérimentales des mouvements oculaires acquises en 2D et en 3D.

L'étude des facteurs visuels influençant les mouvements oculaires est un sujet de toute première importance dans le domaine des sciences cognitives, mais aussi de la vision par ordinateur où pour de nombreuses applications multimédia, les ingénieurs cherchent à estimer les zones les plus « importantes » dans les images, vidéos, ... pour adapter les contenus et ou les traitements.

La compréhension des mouvements oculaires est une des clés majeures pour comprendre notre perception visuelle c'est-à-dire comment nous explorons le monde visuel qui nous entoure pour interagir avec lui. En première approximation, les mouvements oculaires sont considérés comme une séquence spatiale (où ?) et temporelle (combien de temps ?) de fixations suivies de saccades. Durant les fixations, le regard est stabilisé sur une zone de la scène pour extraire l'information visuelle et les saccades sont les mouvements des yeux d'une fixation à une autre fixation. Grâce à l'usage de l'oculométrie (enregistrement du mouvement des yeux) il est possible d'approfondir les mécanismes sous tendant les mouvements oculaires. Les questions fondamentales sont entre autres de comprendre pourquoi le regard est attiré par telle ou telle région de la scène et pourquoi au bout d'un certain temps, le regard quitte cette région pour aller vers une autre région et comment est choisie cette nouvelle région. En se focalisant plutôt sur les aspects spatiaux, la question peut se reformuler ainsi : quelles sont les propriétés visuelles partagées par l'ensemble des régions qui attirent le regard (régions saillantes) ? Il est couramment admis que ces propriétés reflèteraient une modulation de facteurs visuels dits de bas-niveau (« bottom-up ») liés à la scène avec des facteurs dits de haut-niveau (« top down ») liés au sujet (idiosyncrasie), à sa motivation d'exploration, etc.

Pour traiter cette problématique, la méthodologie classiquement utilisée repose sur la modélisation (Fig c) et l'expérimentation en oculométrie pour recueillir sur un grand nombre de sujets et de scènes visuelles les positions spatiales des fixations oculaires (Fig a-b).

grenoble
images
parole
signal
automatique

Le modèle (Fig c) développé au laboratoire, cherche à prédire ces zones saillantes de l'image. Ce modèle a été réalisé dans un contexte de scènes visuelles 2D couleur, en tenant compte uniquement des facteurs « bottom-up » (thèse de Ho Phuoc Tien, 2010). Ces facteurs sont les contrastes en luminance à différentes bandes de fréquences spatiales et d'orientations. Ce modèle intègre une étape de modélisation de la rétine et implémente les caractéristiques connues des champs récepteurs des cellules du cortex visuel primaire.

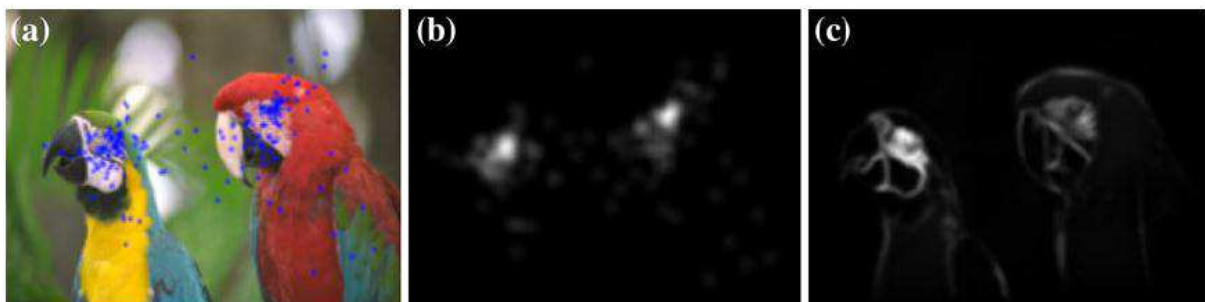


Figure : (a) Fixations oculaires (points) de plusieurs sujets pour cette même image, (b) Carte de densité des fixations oculaires correspondantes, (c) modèle prédictif des lieux saillants, c'est à dire ayant la plus forte probabilité d'attraction du regard.

Plus récemment, grâce aux nouveaux dispositifs de visualisation d'images 3D, nous avons développé au laboratoire (thèse Christophe Maggia, 2014) un modèle d'estimation de la profondeur à partir de l'indice binoculaire de la disparité rétinienne (le fait qu'en vision 3D, les champs visuels projetés sur les rétines gauche et droite diffèrent et que le décalage dépend de la profondeur de la zone fixée). De plus, les expériences en oculométrie menées dans le cadre de cette thèse, nous ont montré l'importance des indices monoculaires dans l'explication du mouvement des yeux réalisés pour explorer la scène.

Ainsi, nous souhaitons pouvoir concevoir et implémenter grâce à ce stage, un modèle prédictif de saillance visuelle 3D intégrant ces indices de perception de la profondeur aux facteurs « bottom-up », et ainsi mieux comprendre comment la perception de la profondeur interagit avec les facteurs 2D dans le déploiement de l'attention visuelle pour l'exploration de scènes visuelles 3D. Dans la littérature, il existe très peu de modèles computationnels réalisant cette intégration et de plus, aucun n'intègre les indices monoculaires.



Travail attendu :

- Revue de la littérature des modèles de saillance visuelle 3D, à partir de celle déjà réalisée par Christophe Maggia dans sa thèse.
- Extension et intégration des modèles proposés par T. Ho-Phuoc (saillance 2D), C. Maggia (estimation de la disparité binoculaire) et des travaux de la littérature pour l'estimation des indices monoculaires.
- Validation du modèle proposé grâce à une base expérimentale des mouvements oculaires recueillis en exploration de scènes 2D et 3D.

Extension possible :

Il est prévu de recruter en parallèle un étudiant (M1) pour monter une expérience utilisant l'oculométrie afin de suivre le regard de personnes visualisant des scènes stéréoscopiques (3D) via un écran 3D. Ces scènes intégreront des éléments visuels à fortes sémantiques, comme des personnages, connus pour attirer immédiatement le regard. Cela permettra d'étudier l'importance relative des visages et des indices de profondeur 3D.

Le stage proposé ici, plus orienté sur l'approche modélisation, pourra, en fonction de l'intérêt du candidat et de son investissement, comporter une partie sur l'expérience en oculométrie.

Une poursuite en thèse est naturellement envisagée sur cette problématique.

Compétences et/ou intérêt souhaités :

- Intérêt pour la modélisation computationnelle cognitive,
- Bases en traitement du signal et ou des images,
- Statistiques,
- Programmation en Matlab.

Rémunération : Oui

Encadrement :

- Nathalie Guyader (nathalie.guyader@gipsa-lab.grenoble-inp.fr)
- Simon Barthelme (simon.barthelme@unige.ch)
- Anne Guérin-Dugué (anne.guerin@gipsa-lab.grenoble-inp.fr)

grenoble
images
parole
signal
automatique

Lieu du stage : GIPSA-lab, département Images et Signal.

GIPSA-lab

Grenoble Campus
11 rue des Mathématiques - BP46
F-38402 SAINT MARTIN D'HERES Cedex
Tél. +33 (0)4 76 82 62 56
Fax +33 (0)4 76 82 64 26

www.gipsa-lab.fr
prenom.nom@gipsa-lab.fr

UMR5216

CNRS , Grenoble INP,
UJF, Stendhal