

Système temps réel de simulation d'attention visuelle : application aux images et séquences d'images

M. Perreira Da Silva

V. Courboulay

A. Prigent

P. Estrailier

Laboratoire Informatique Image et Interaction

Université de La Rochelle, Avenue M. Crépeau, 17042 La Rochelle cedex 0, France mperreir@univ-lr.fr

Résumé

Cet article présente un démonstrateur de simulation d'attention visuelle incluant différentes améliorations du modèle de Laurent Itti [3]. Ce modèle utilise une approche basée sur les systèmes proies / prédateurs permettant de simuler une évolution temps réel du focus d'attention. Nous décrivons rapidement une méthode de génération de ces cartes de singularité, permettant le traitement temps réel sur un ordinateur standard tout en améliorant les performances finales sur la saillance. Enfin nous présentons notre démonstrateur qui mettra en avant les résultats obtenus sur des images fixes et sur des flux vidéos. Des comparaisons avec des algorithmes classiques et des résultats d'expériences oculométriques seront également proposées.

Mots Clef

Attention visuelle, proies / prédateurs, temps réel.

1 Introduction

L'approche que nous proposons ici permet de simuler l'évolution temporelle du focus d'attention en se basant sur une version modifiée des algorithmes classiques de [3]. Le système original propose une architecture comprenant des calculs *bas niveau* permettant d'extraire trois cartes de singularité (conspicuity maps). Ces cartes de singularité représentent l'importance respective des différents éléments de la scène observée par rapport aux principaux canaux perceptuels humains (intensité, couleur, orientation). Dans notre architecture, ces calculs bas niveau sont effectués en utilisant des algorithmes dérivés des travaux de [1] [2] eux même basés sur [3]. L'approche utilisée afin d'accélérer les calculs (*i.e.* l'utilisation d'images intégrales) a été reprise et étendue afin que toutes les cartes soient calculées en temps réel. Ces améliorations sont décrites dans la section 2.1.

La deuxième partie de l'architecture de [3] utilise un système *moyen niveau* permettant de fusionner les cartes de singularité puis de générer un parcours d'attention visuelle à partir d'un système de *winner takes all* et *inhibition of return*. Nous proposons de remplacer cette deuxième partie par un système proies / prédateurs afin d'introduire un paramètre temporel permettant de générer des

saccades, fixations et trajectoires de manière plus plausible.

Le choix des systèmes proies/prédateurs est justifié par les arguments suivants :

- les systèmes proies / prédateurs sont des systèmes dynamiques. Ils gèrent donc naturellement l'évolution temporelle de leur activité ce qui est primordial pour générer l'évolution temporelle du focus d'attention à partir d'une image fixe ou d'une séquence d'images.
- en l'absence d'objectif (flux d'information top-down ou *prégnance*), le choix d'une méthode de fusion des informations issues des différentes cartes de singularité est difficile à faire. Les systèmes proies / prédateurs proposent une solution permettant la mise en place d'un équilibre naturel entre ces différentes cartes *via* un mécanisme de compétition (*prédation*).
- les systèmes dynamiques discrétisés peuvent avoir un comportement chaotique. Bien que cette propriété ne soit généralement pas recherchée, c'est une qualité importante dans notre cas puisqu'elle permet l'émergence de parcours originaux, permettant d'explorer la scène visuelle, même dans des zones faiblement saillantes.

Dans la section suivante, nous présentons brièvement l'approche que nous avons utilisée pour ce démonstrateur.

2 Apports scientifiques

2.1 Génération temps réel de cartes de singularité

Une solution utilisant les images intégrales [8] afin de générer rapidement des cartes de singularité est proposée dans [2]. Cependant, l'auteur explique qu'elle n'a appliqué ces optimisations que sur les cartes d'intensité et de couleur. Pour notre système, afin d'obtenir des performances optimales, nous avons généralisé l'utilisation de cette technique à l'ensemble des calculs des cartes de singularité. Notre implémentation en C# de cet algorithme permet d'obtenir des performances temps réel pour des images de résolution 320x240 sur un ordinateur Compaq nc8430, équipé de 2Go de Ram et d'un CPU double coeur Intel Core T2400 à 1.83Ghz.

2.2 Simulation de l'évolution temporelle de l'attention visuelle

L'utilisation des systèmes proies / prédateurs à des fins de simulation de l'attention visuelle a été décrite dans [6], nous ne décrivons donc ici que les quelques détails nécessaires à la compréhension globale de notre modèle d'attention.

Les systèmes proies / prédateurs sont des systèmes d'équations utilisés pour simuler l'évolution et l'interaction de différentes colonies de proies et de prédateurs. Plus de détails sur ces systèmes peuvent être trouvés dans [5]. Dans notre système, nous nous sommes inspirés de [4] afin de représenter l'évolution temporelle de l'intérêt (*i. e.* du focus d'attention) pour des images et séquences d'images.

3 Démonstrations

3.1 Démonstrateur

Notre démonstrateur est constitué d'une application écrite en C# (figure 1). Celle-ci permet de simuler en temps réel l'évolution du focus d'attention visuelle de notre modèle à partir de différentes sources de données (image, séquence d'images, vidéo ou capture *live* à partir d'une webcam).

L'interface du démonstrateur permet de visualiser la sortie principale de l'algorithme (point de focalisation) ainsi que différentes cartes représentant les résultats intermédiaires (carte de caractéristique, cartes de singularité, population des différentes proies et prédateurs, etc.). Enfin l'application construit en temps réel au cours de la simulation une carte de saillance de la scène observée.

Il est également possible de modifier la majorité des paramètres du système, afin d'observer leur effet sur la simulation du processus d'attention visuelle.

3.2 Expérimentations

Dans une première partie de la démonstration, nous souhaitons mettre en avant les performances de différents algorithmes de génération de carte de saillance en comparaison du notre. Pour ce faire, notre démonstrateur permettra de choisir une image dans une base [7], à partir de cette dernière sera alors présentée un ensemble de cartes de saillance issue d'algorithmes de références, du notre et d'une vérité terrain à partir de mesure oculométriques. L'utilisateur pourra aussi choisir une image quelconque (sans présentation de vérité terrain).

Une autre possibilité qu'offrira la démonstration est d'avoir accès à une carte de saillance dynamique obtenu à partir d'un flux vidéo (webcam ou fichier vidéo). Le comportement de l'algorithme sera appréhendé visuellement par l'évolution des cartes de singularités et d'attention. Pour cette expérience, une carte de mouvement a été incluse directement en plus des trois autres cartes.

4 Perspectives

La simulation d'attention visuelle sur des flux vidéos est un travail en cours. Bien que techniquement déjà fonction-

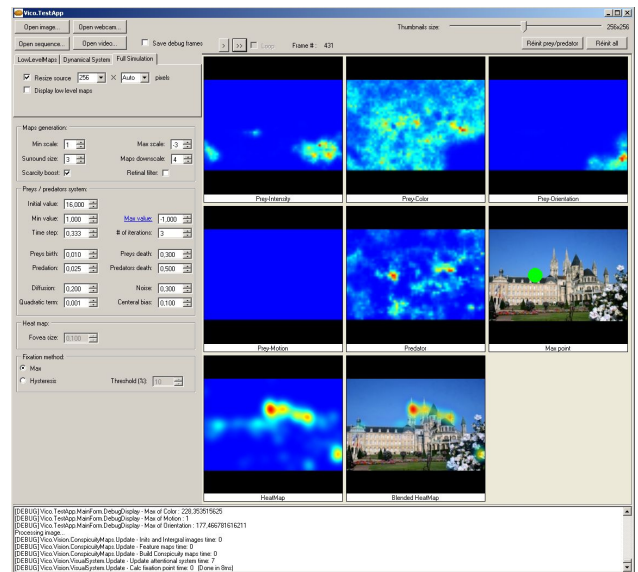


FIGURE 1 – Capture d'écran de l'application de démonstration.

nelle, cette partie du système reste encore à valider.

Références

- [1] Simone Frintrop. *VOCUS : A Visual Attention System for Object Detection and Goal-Directed Search*. PhD thesis, University of Bonn, 2006.
- [2] Simone Frintrop, Maria Klodt, and Erich Rome. A real-time visual attention system using integral images. University Library of Bielefeld, 2007.
- [3] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(11) :1254–1259, Nov 1998.
- [4] Mike Lesser and Dinah Murray. Mind as a dynamical system : Implications for autism. In *Psychobiology of autism : current research & practice*, 1998.
- [5] James D. Murray. *Mathematical Biology : I. An Introduction*. Interdisciplinary Applied Mathematics. Springer, 2002.
- [6] Matthieu Perreira Da Silva, Vincent Courboulay, Arnel Prigent, and Pascal Estrailier. Attention visuelle et systèmes proies / prédateurs. In *XXIIIe Colloque GRETSI - Traitement du Signal et des Image*, Dijon France, 09 2009.
- [7] John K. Tsotsos, Scan M. Culhane, Winky Yan Kei Wai, Yuzhong Lai, Neal Davis, and Fernando Nuflo. Modeling visual attention via selective tuning. *Artificial Intelligence*, 78(1-2) :507 – 545, 1995. Special Volume on Computer Vision.
- [8] P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision (IJCV)*, 57 :137–154, 2004.