Modèle computationnel d'attention pour la vision adaptative

Soutenance publique de thèse de doctorat

Matthieu Perreira Da Silva

10/12/2010

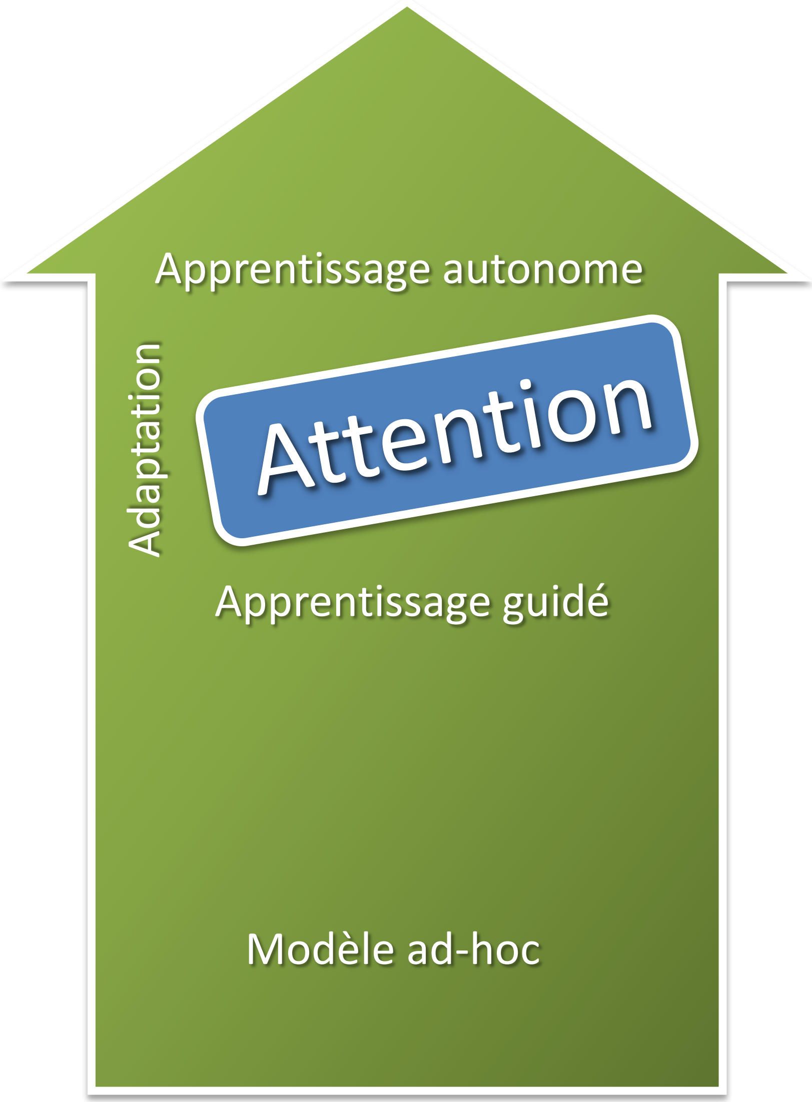
Sous la direction de :
Vincent COURBOULAY
Pascal ESTRAILLIER
Armelle PRIGENT

Vision, adaptation, attention…

VeriFinger 5.0 Algorithm Demo
File
Mode
View
Tools
Help
Main log
Stop
Scanner log
Identification results
Stop
Tetris
TETRIS
XNA
NEXT
SCORE
0
LEVEL
1
LINES
0
AutiSTIC GUI
Fichier
Aide
Démarrer
Arrêter
Pause
Pas à pas

```
Waiting for image(s)
VeriFinger 5.0 Algorithm Demo    Version 5.0.0.1
OS: Windows XP Professional, Service Pack 2. Version 5.1
SetMonitor: OK
0 scanners found
Please install scanner drivers f
C:\Program Files\Neurotechnologija\VeriFinger 5.0 Algorithm Demo\VFDemo.mdb
```

G	Result	Similarity	Record ID	Finger ID
134	Identit...	654	66	013_8_1



Apprentissage autonome

Adaptation

Attention

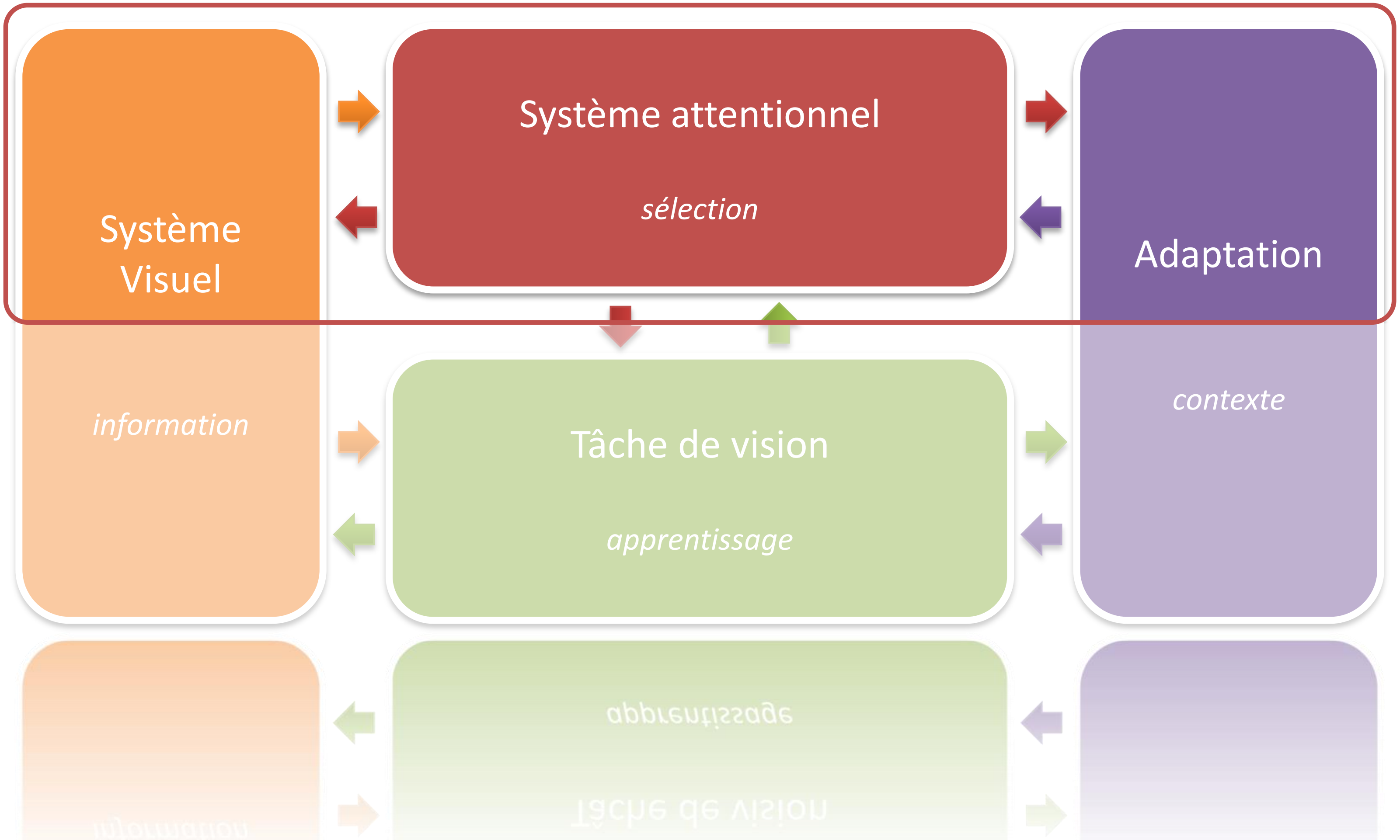
Apprentissage guidé

Modèle ad-hoc

Information acquise *via*

Vision, adaptation, attention…

Architecture théorique


Système attentionnel
sélection
Système Visuel
information
Adaptation
contexte
Tâche de vision
apprentissage

Au programme

L’attention

- Définition
- Cadre
- Contraintes
- Caractérisation

La théorie

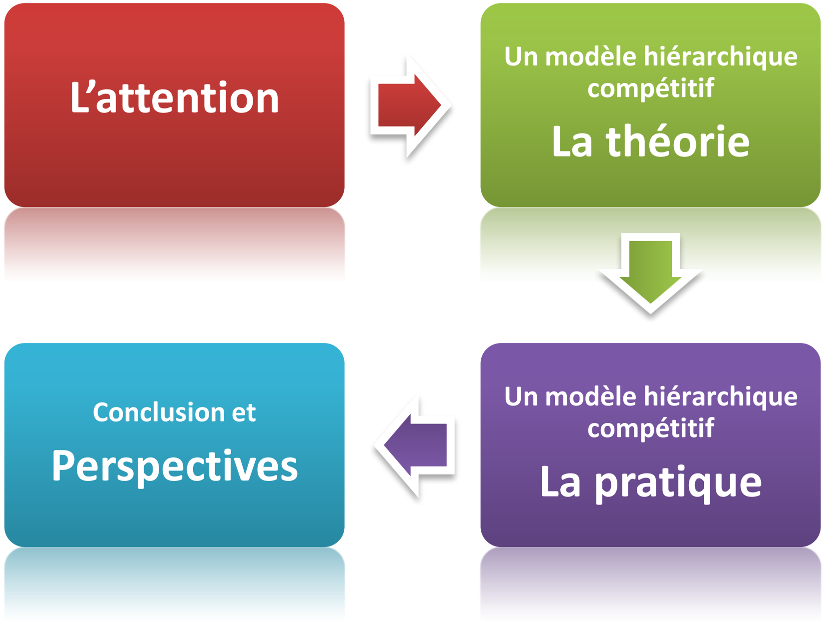
- Architecture
- Rapidité
- Dynamique
- Adaptation

La pratique

- Plausibilité
- Propriétés
- Applications

Conclusion et

Perspectives



L'attention

Un modèle hiérarchique compétitif

La théorie

Un modèle hiérarchique compétitif

La pratique

Conclusion et

Perspectives

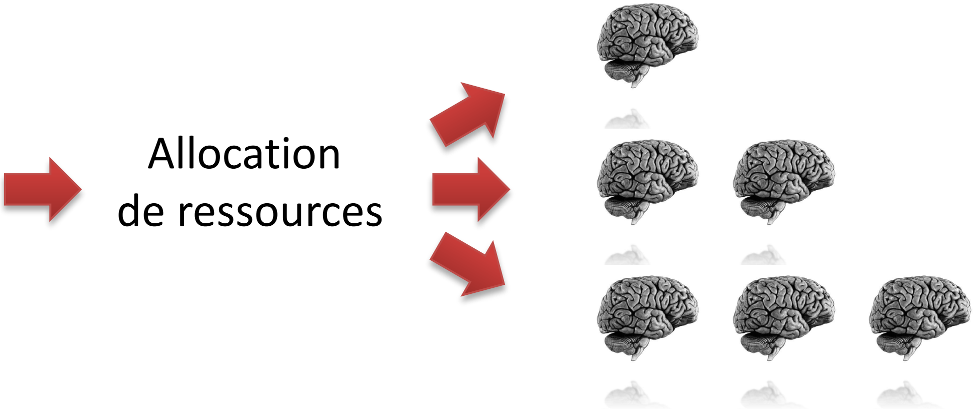
L'ATTENTION

DE LA DÉFINITION AU MODÈLE COMPUTATIONNEL

Attention : des définitions

- **Quoi ?**

Sélection

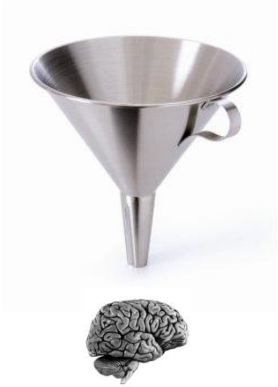


Allocation
de ressources

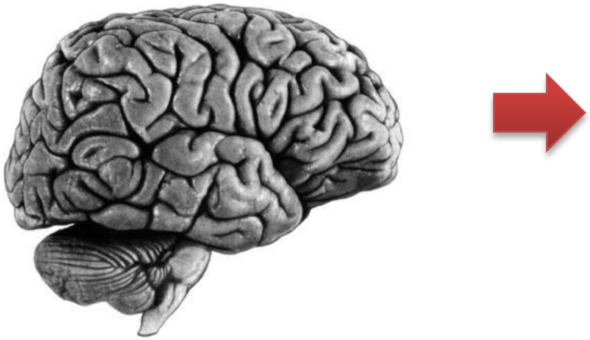
- Pourquoi ?
- De multiples facettes
- Attention ≠ ou = Saillance ?

Attention : des définitions

- Quoi
- **Pourquoi ?**



et / ou



Capacités limitées

[Broadbent1958]

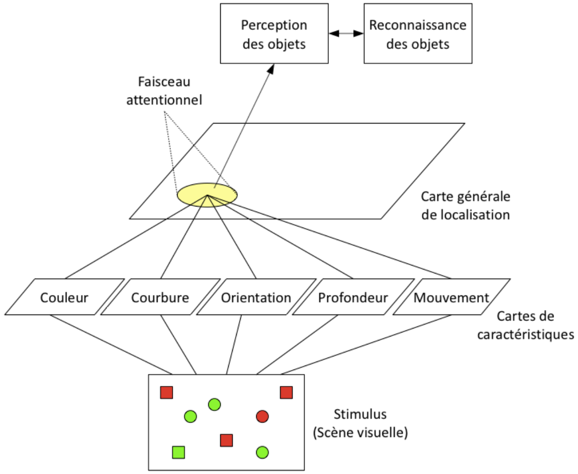
Objectif fonctionnel

[Allport1987]

- De multiples visages
- Attention ≠ ou = Saillance ?

Attention : des définitions

- Quoi
- Pourquoi
- **De multiples facettes**



Perception des objets

Reconnaissance des objets

Faisceau attentionnel

Carte générale de localisation

Couleur

Courbure

Orientation

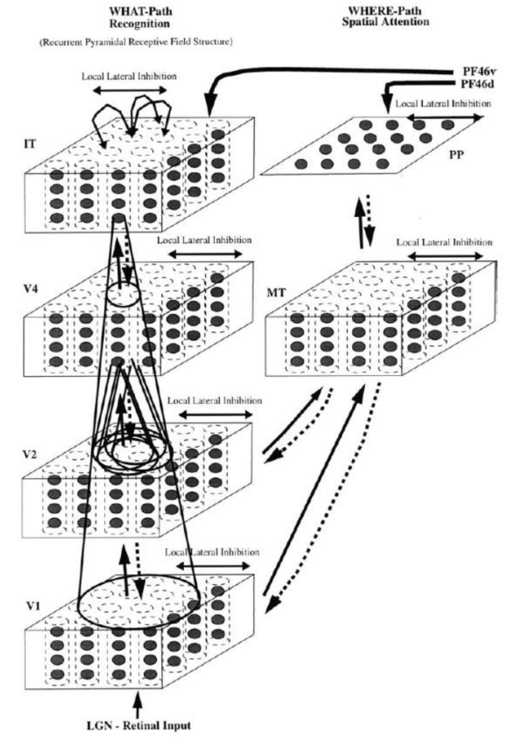
Profondeur

Mouvement

Cartes de caractéristiques

Stimulus (Scène visuelle)

et /ou


WHAT-Path
Recognition
(Recurrent Pyramidal Receptive Field Structure)
WHERE-Path
Spatial Attention
Local Lateral Inhibition
PF46v
PF46d
Local Lateral Inhibition
IT
PP
Local Lateral Inhibition
Local Lateral Inhibition
V4
MT
Local Lateral Inhibition
V2
Local Lateral Inhibition
V1
LGN - Retinal Input

Centralisée

Espace

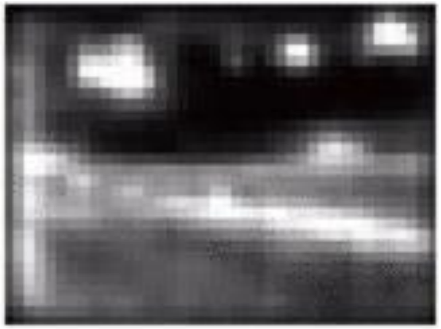
Objet

Distribuée

- Attention ≠ ou = Saillance ?

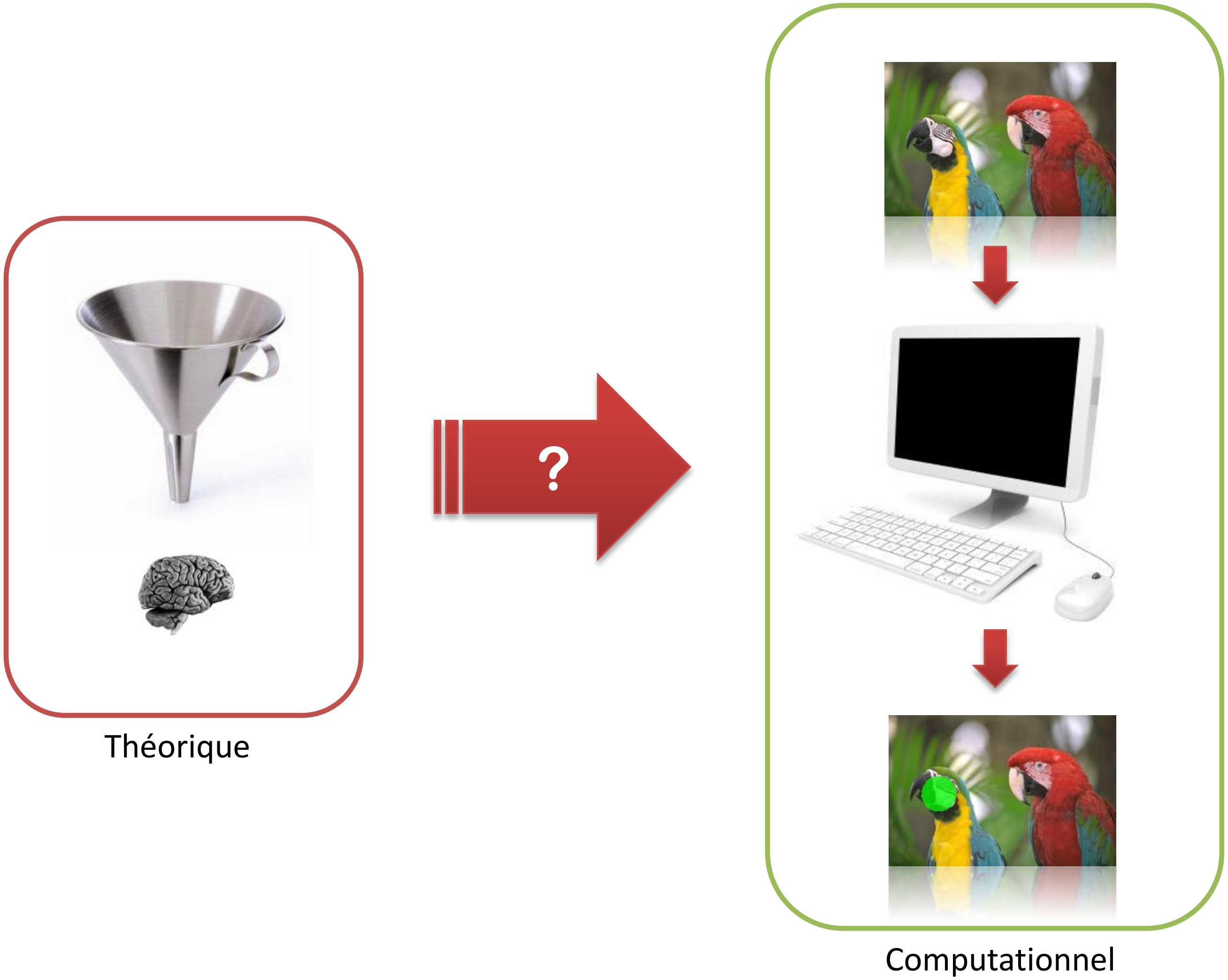
Attention : des définitions

- Quoi
- Pourquoi
- De multiples facettes
  - Ouverte ou couverte
  - Endogène ou exogène
  - Espace ou objet
  - Centralisée ou distribuée
- **Attention ≠ ou = Saillance ?**



Carte de saillance

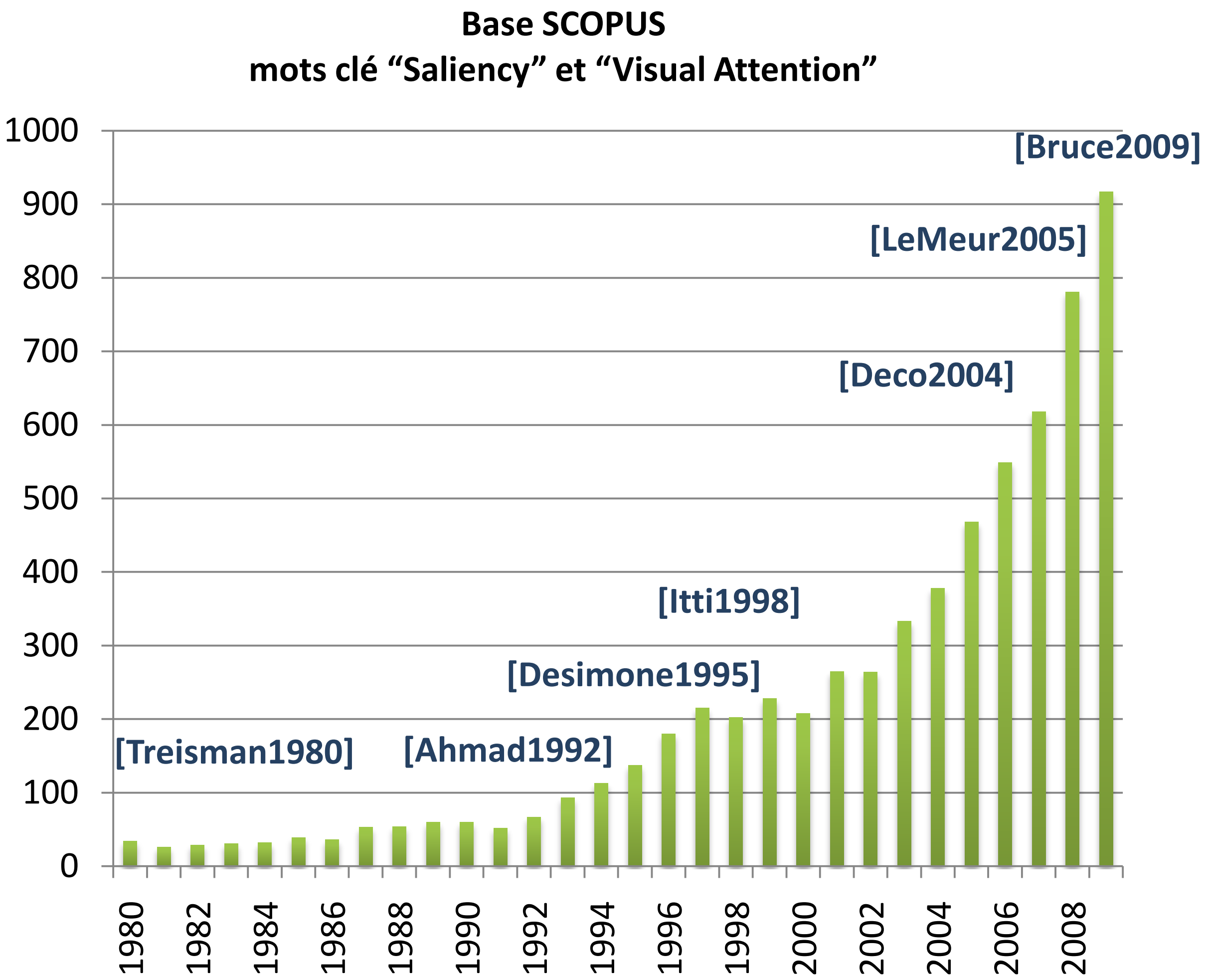
Vers un modèle computationnel d’attention


?
Théorique
Computationnel

Vers un modèle computationnel d’attention

- **Un intérêt grandissant**

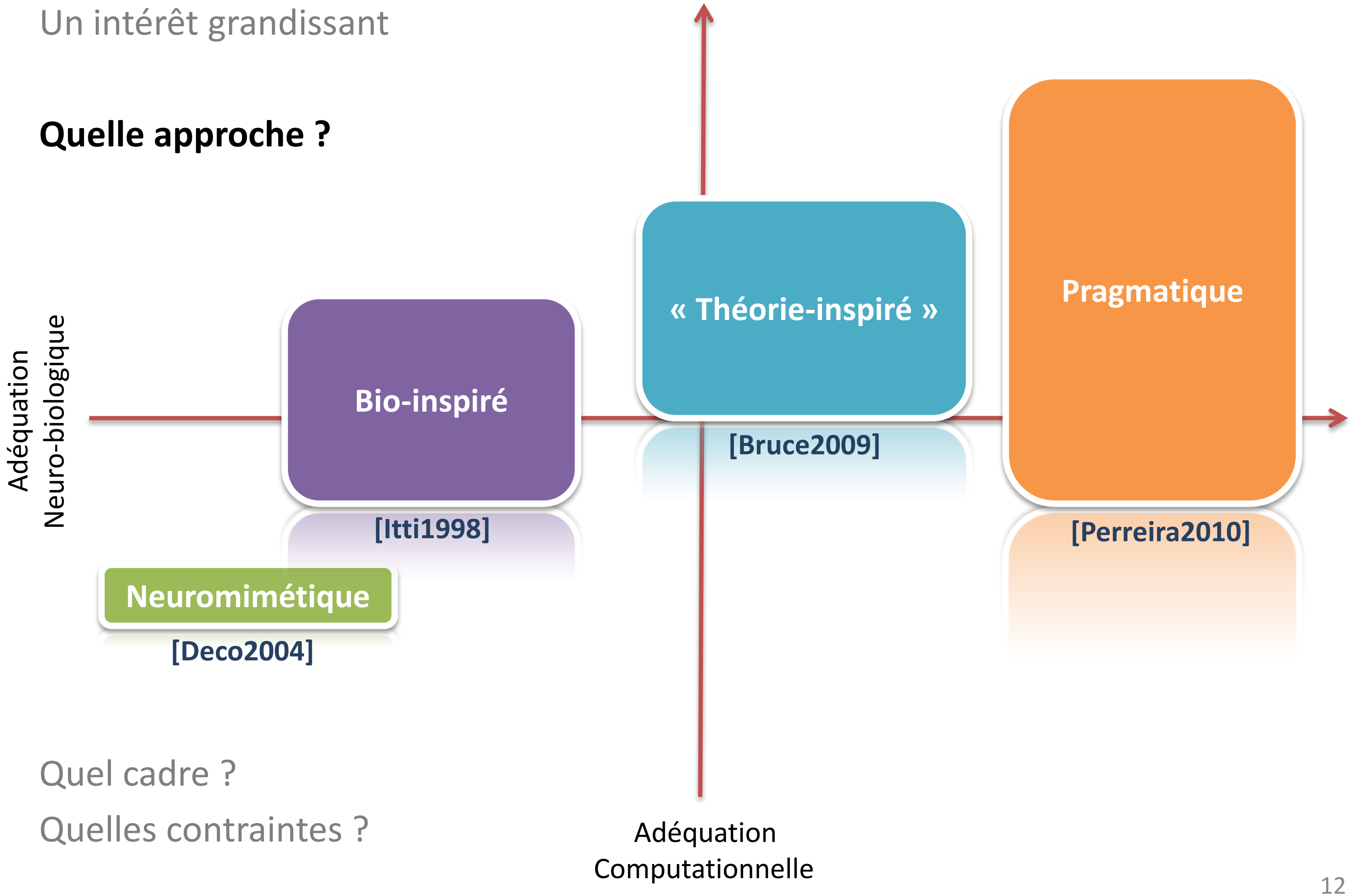
- Quelle approche ?
- Quel cadre ?
- Quelles contraintes ?



Base SCOPUS
mots clé “Saliency” et “Visual Attention”

Vers un modèle computationnel d'attention

- Un intérêt grandissant
- **Quelle approche ?**



Adéquation
Neuro-biologique
Bio-inspiré
[Itti1998]
Neuromimétique
[Deco2004]
« Théorie-inspiré »
[Bruce2009]
Pragmatique
[Perreira2010]
Adéquation
Computationnelle

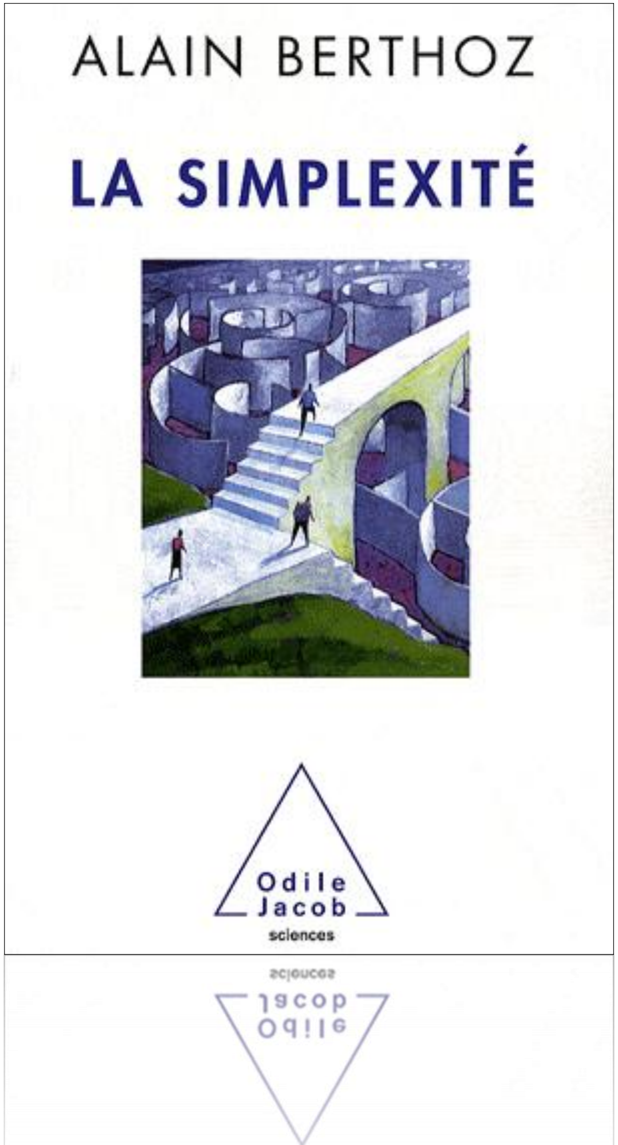
- Quel cadre ?
- Quelles contraintes ?

Vers un modèle computationnel d'attention

- Un intérêt grandissant
- Quelle approche
  - Pragmatique

- **Quel cadre : la Simplexité**
  - **→ Simplifier la complexité**
  - **→ Propriétés d'un système Simplexe…**

- Quelles contraintes ?


ALAIN BERTHOZ
LA SIMPLEXITÉ
Odile Jacob
sciences

Vers un modèle computationnel d'attention

- Un intérêt grandissant
- Quelle approche
  - Pragmatique
- Quel cadre
  - La Simplexité

- **Quelles contraintes ?**

	Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
Etude de l'attention	●●●	●	●	●	●	●●●
Ergonomie / publicité	●●●	●	●	●	●●	●●
Vision / interaction	**●●**	**●●●**	**●●**	**●●●**	**●●●**	**●●●**
CBIR	●	●	●●●	●●	●●●	●
Traitement d'image	●●	●●	●●	●●	●●●	●●

Contraintes simplexes *FAIRED* caractérisant l'adéquation d'un modèle à une application

Taxonomie des modèles computationnels

Distribués

+ compétition
+ gestion de la dynamique

- complexité
- ajout de caractéristiques

• Plutôt théoriques	[Spratling2004]
• Réellement computationnels	[Tsotsos2005]

Centralisés

+ efficacité computationnelle
+ facilement extensible

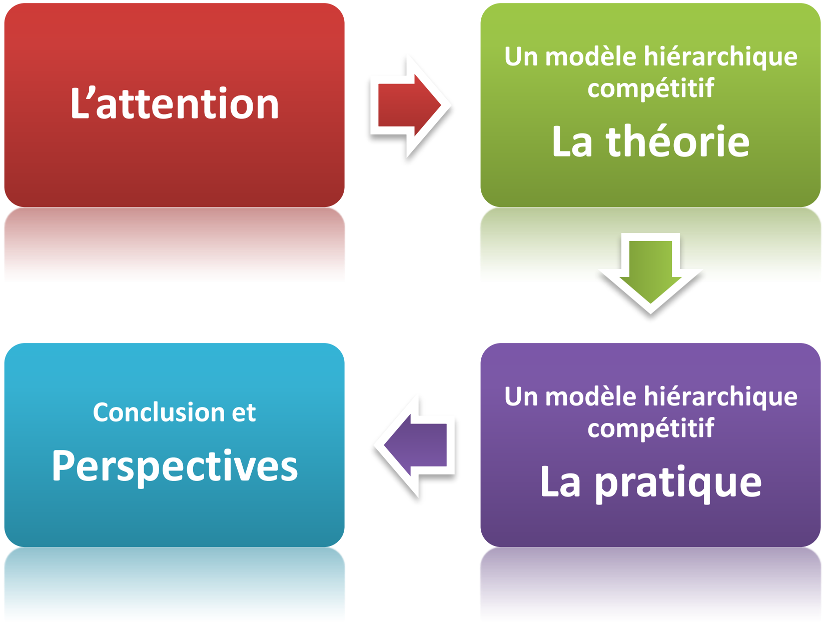
- gestion de la dynamique

• Hiérarchiques	[Itti1998]
• Statistiques / probabilistes	[Hamker2005]
• Théorie de l'information	[Mancas2007]
• Connexionnistes	[Mozer1998]
• Algorithmiques	[Aziz2009]

Quel modèle pour la vision adaptative ?

	Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
Objectif	●●	●●●	●●	●●●	●●●	●●●
Distribués	●●●	●●●	●●●	●	●	●●●
Hiérarchiques	●●	●●	●●	●●●	●●●	●
Statistiques	●●	●	●●	●●	●●	●
Théorie de l’information	●●	●	●●	●●	●●	●
Connexionnistes	●●	●●●	●●	●●	●●	●●●
Algorithmiques	●	●●	●●	●●●	●	●

Adéquation des différentes familles d’algorithmes avec les contraintes *FAIRED* caractérisant la vision par ordinateur



L'attention

Un modèle hiérarchique compétitif

La théorie

Un modèle hiérarchique compétitif

La pratique

Conclusion et

Perspectives

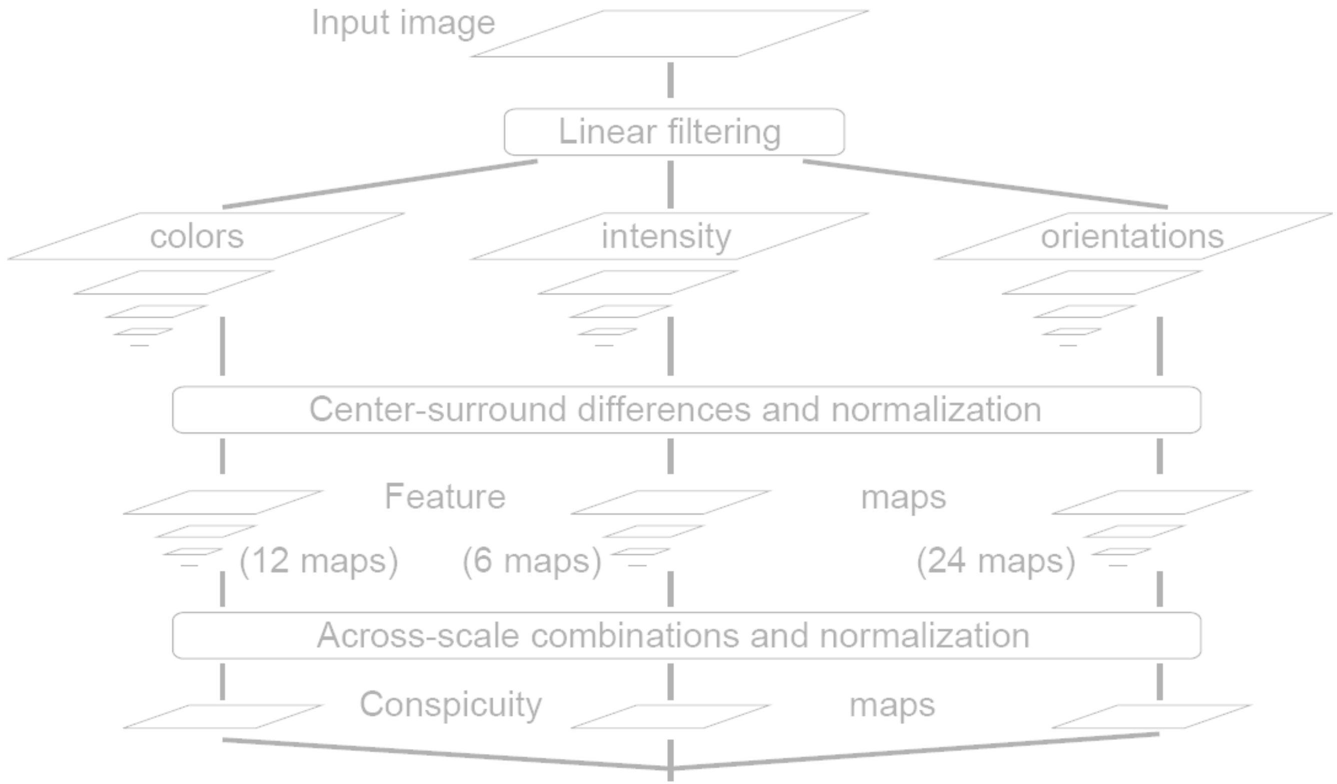
UN MODÈLE HIÉRARCHIQUE COMPÉTITIF

LA THÉORIE

Modèle d’attention visuelle

Hiérarchique

Distribuée



Input image

Linear filtering

colors

intensity

orientations

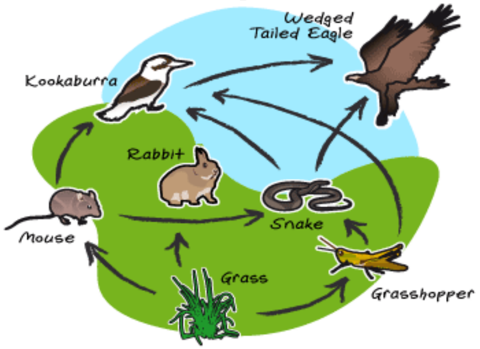
Center-surround differences and normalization

Feature maps

(12 maps) (6 maps) (24 maps)

Across-scale combinations and normalization

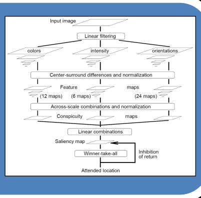
Conspicuity maps


Wedged Tailed Eagle
Kookaburra
Rabbit
Snake
Mouse
Grass
Grasshopper

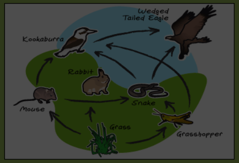
Compétition

Contributions

Approche hiérarchique


Input image
Linear filtering
colors
intensity
orientations
Center-surround differences and normalization
Feature maps
(12 maps)
(6 maps)
(24 maps)
Across-scale combinations and normalization
Conspicuity maps
Linear combinations
Saliency map
Winner-take-all
Inhibition of return
Attended location

Approche distribuée

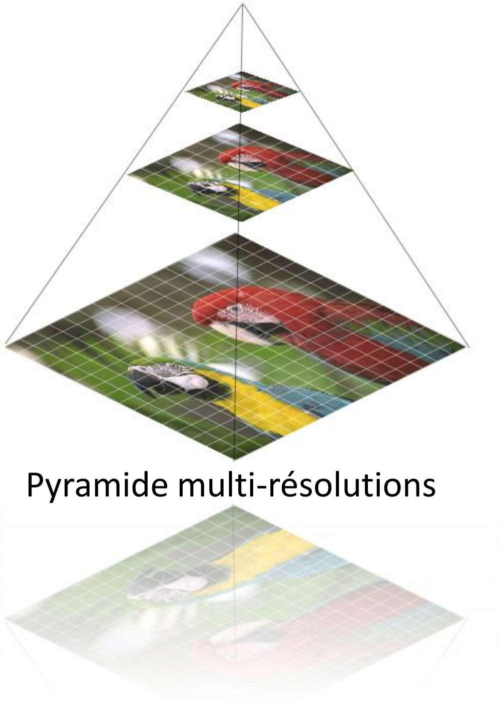

Wedged Tailed Eagle
Kookaburra
Rabbit
Mouse
Snake
Grass
Grasshopper

	Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
Objectif	●●	●●●	●●	●●●	●●●	●●●
Hiérarchiques	●●	●●	●●	●●●	●●●	●
Distribués	●●●	●●●	●●●	●	●	●●●

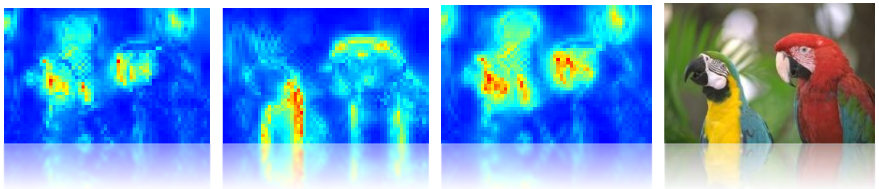
Fidèle	**Adaptable**	**Invariant**	**Rapide**	**Extensible**	**Dynamique**

Approche hiérarchique

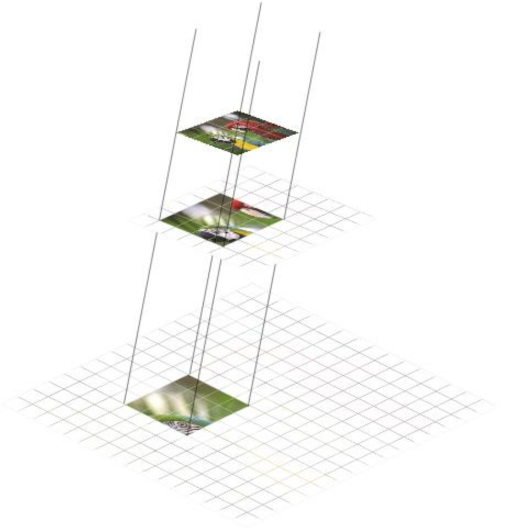
Plus vite…



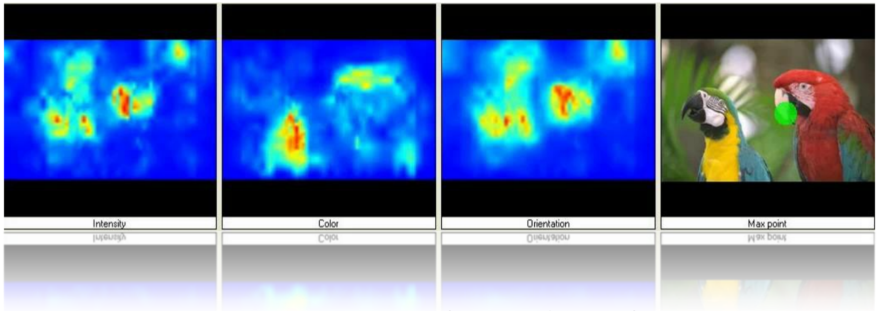
Pyramide multi-résolutions



Cartes de singularité



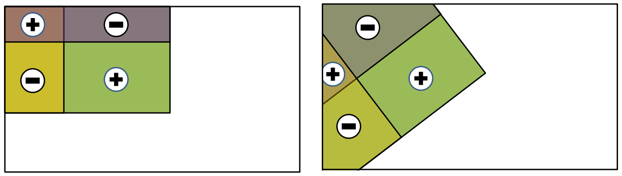
Colonne multi-résolutions


Intensity
Color
Orientation
Max point

Cartes de singularité avec flou rétinien

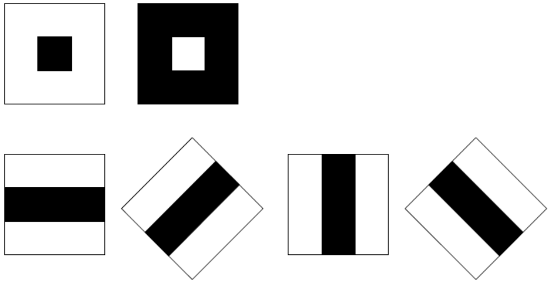
Approche hiérarchique

Plus vite…



Images intégrales

[Crow1984, Lienhart2002]

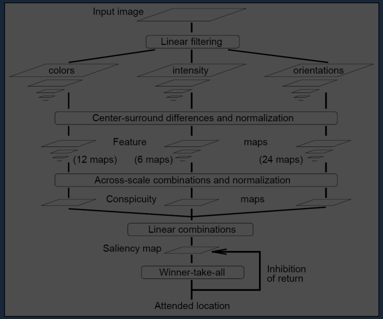


Filtres simplifiés « Harr like »

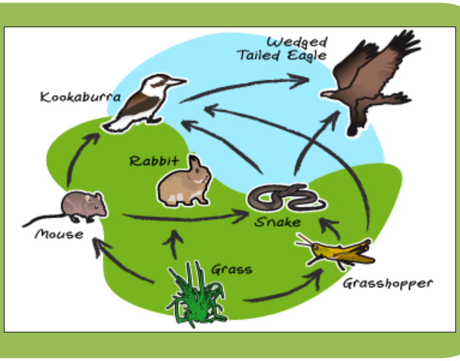
[Viola2002 , Frintrop2007]

	160x120	320x240	640x480
Nombre de niveaux	3	4	5
Pyramide	100 fps	28 fps	7 fps
Colonne (flou)	143 fps	35 fps	9 fps

Images par seconde sur PC quad core @ 2.8 GHz


Input image
Linear filtering
colors
intensity
orientations
Center-surround differences and normalization
Feature maps
(12 maps)
(6 maps)
(24 maps)
Across-scale combinations and normalization
Conspicuity maps
Linear combinations
Saliency map
Winner-take-all
Inhibition of return
Attended location

Approche hiérarchique

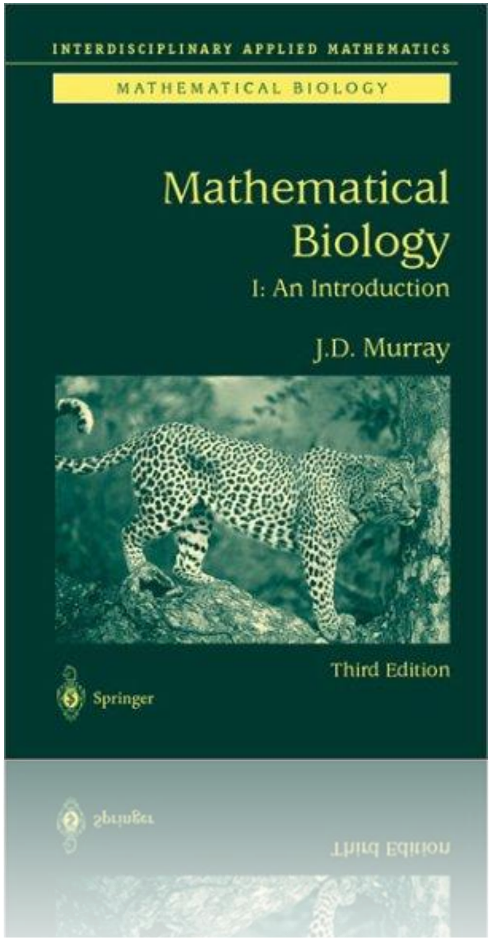

Wedged Tailed Eagle
Kookaburra
Rabbit
Snake
Mouse
Grass
Grasshopper

Approche distribuée

Pourquoi un modèle proies / prédateurs ?

- Un système dynamique…

- La compétitio… atégie de fusion


INTERDISCIPLINARY APPLIED MATHEMATICS

MATHEMATICAL BIOLOGY

Mathematical Biology

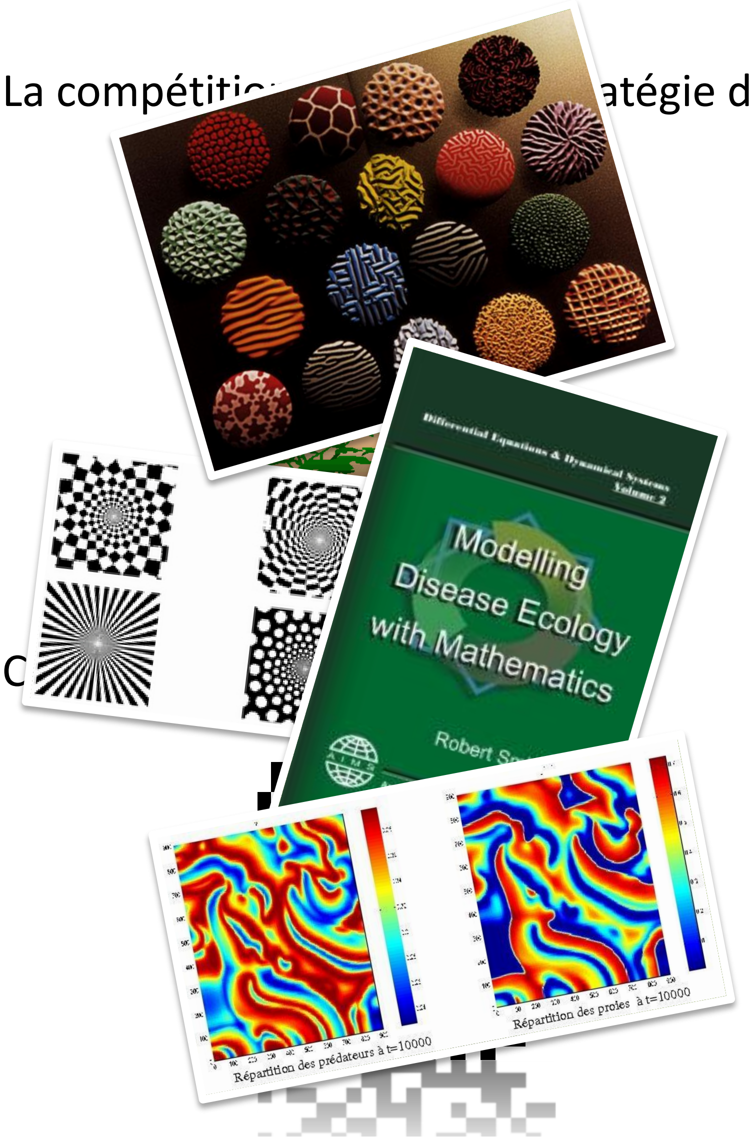
I: An Introduction

J.D. Murray

Third Edition

Springer

- C


Differential Equations & Dynamical Systems Volume 2

Modelling Disease Ecology with Mathematics

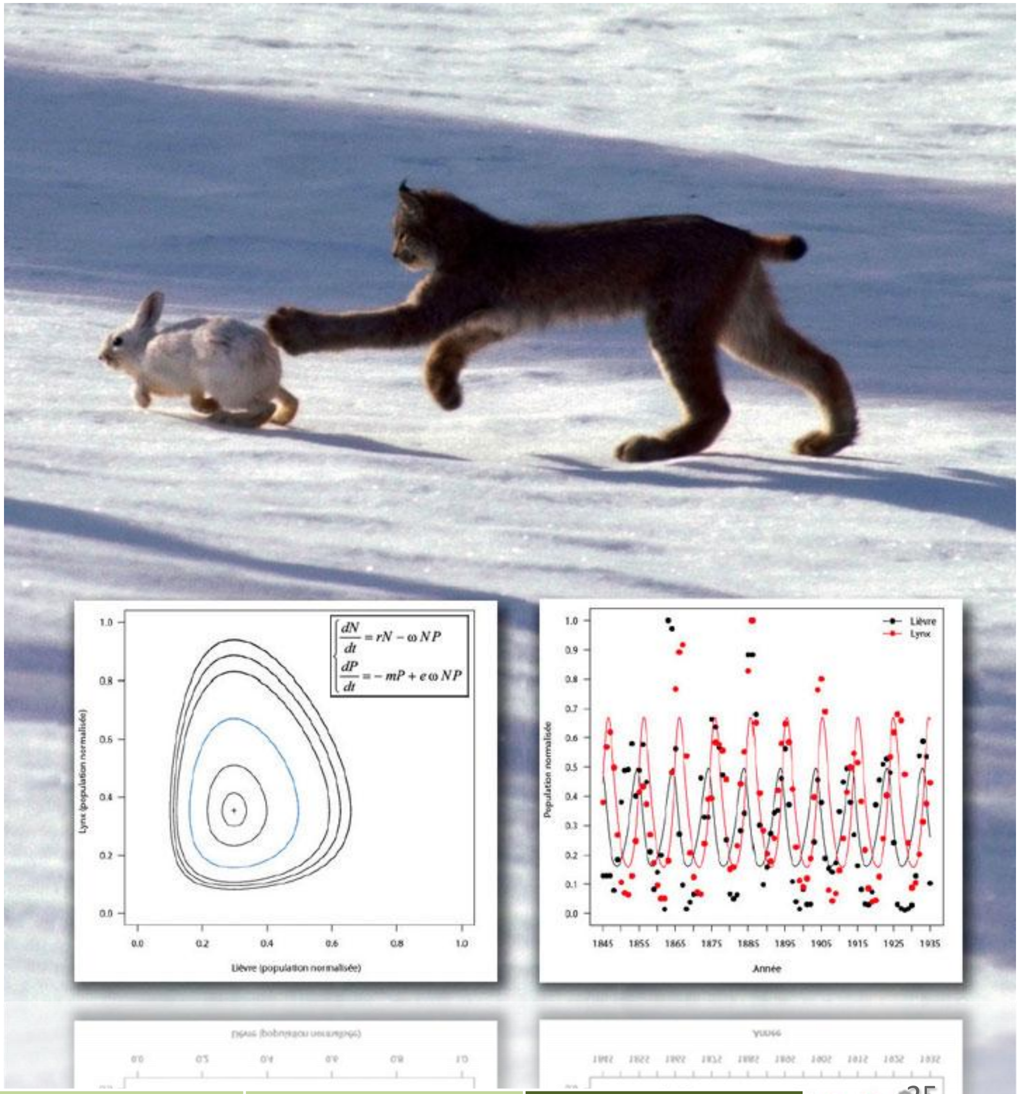
Robert S…

Répartition des prédateurs à t=10000
Répartition des proies à t=10000

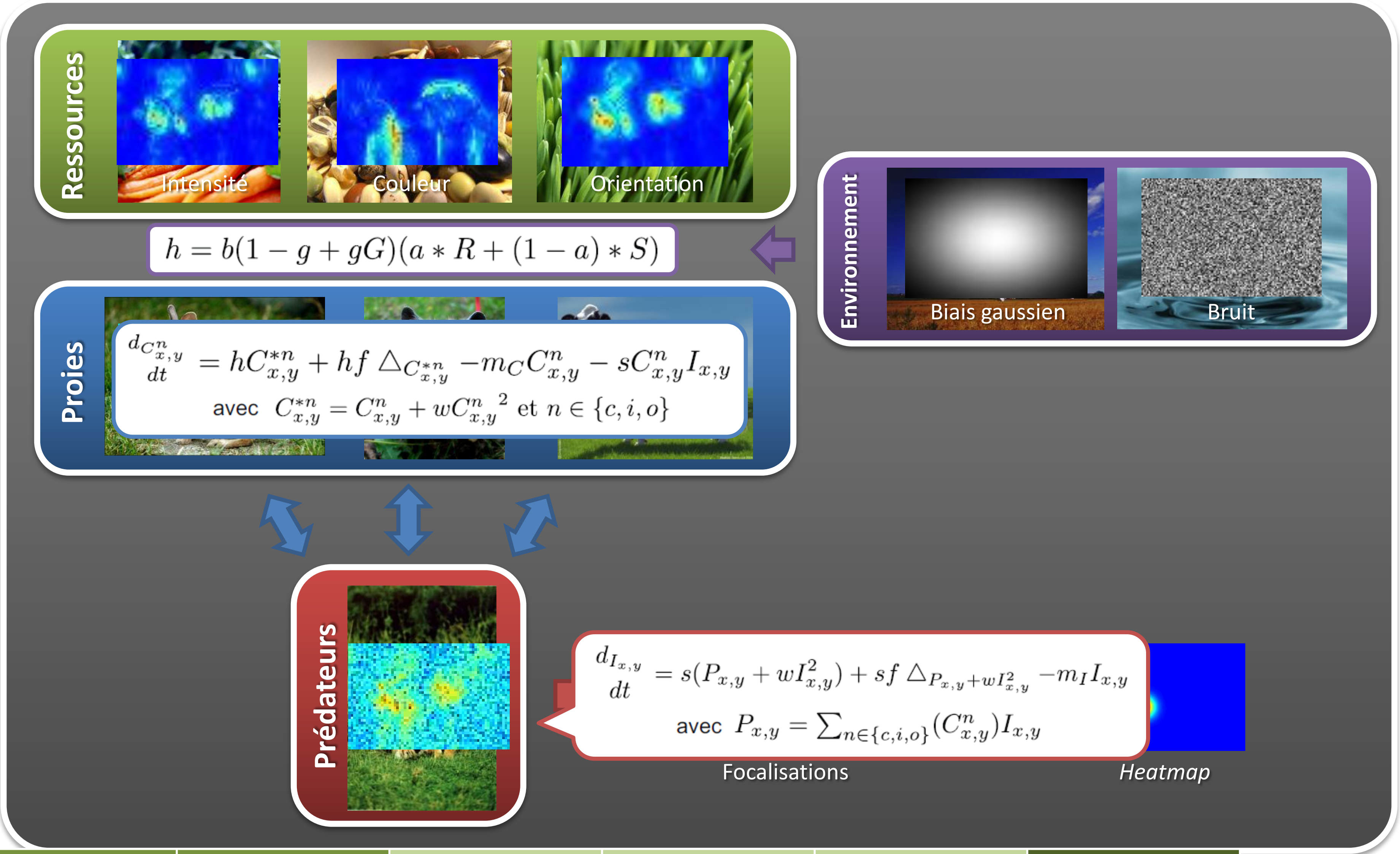
Comment ça marche ?

- Modèle de base (Volterra-Lotka)
  - Proies : $$\frac{dx}{dt} = x(\alpha - \beta y)$$
  - Prédateurs : $$\frac{dy}{dt} = -y(\gamma - \delta x)$$

En théorie les solutions sont périodiques



Comment ça marche ? Transposition à l'attention


Ressources
Intensité
Couleur
Orientation
Environnement
Biais gaussien
Bruit
Proies
Prédateurs
Focalisations
Heatmap

$$h = b(1-g+gG)(a*R+(1-a)*S)$$

$$\frac{dC^n_{x,y}}{dt} = hC^{*n}_{x,y} + hf\triangle_{C^{*n}_{x,y}} - m_C C^n_{x,y} - sC^n_{x,y}I_{x,y}$$

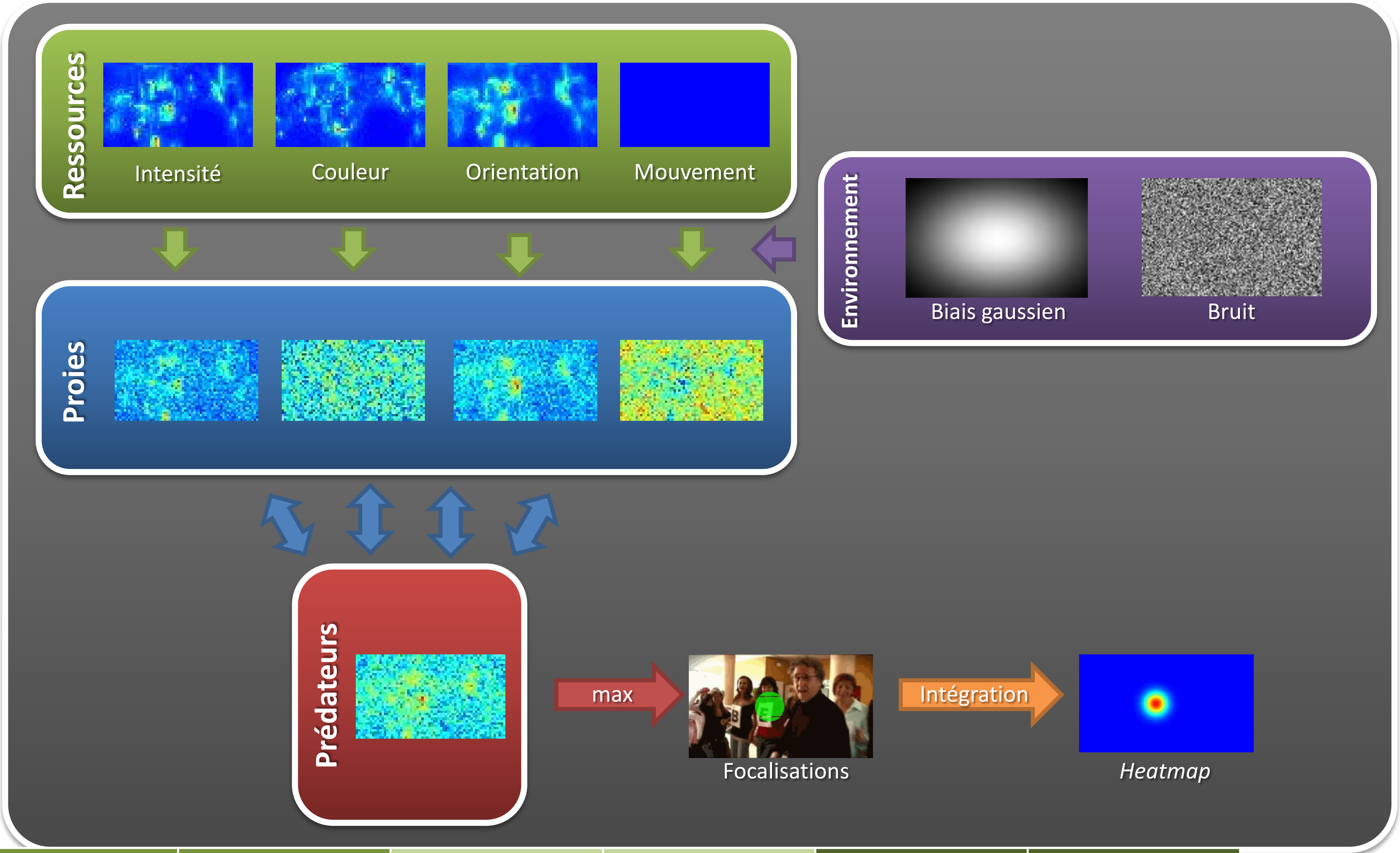
avec $C^{*n}_{x,y} = C^n_{x,y} + w{C^n_{x,y}}^2$ et $n \in \{c,i,o\}$

$$\frac{dI_{x,y}}{dt} = s(P_{x,y}+wI^2_{x,y}) + sf\triangle_{P_{x,y}+wI^2_{x,y}} - m_I I_{x,y}$$

avec $P_{x,y} = \sum_{n\in\{c,i,o\}}(C^n_{x,y})I_{x,y}$

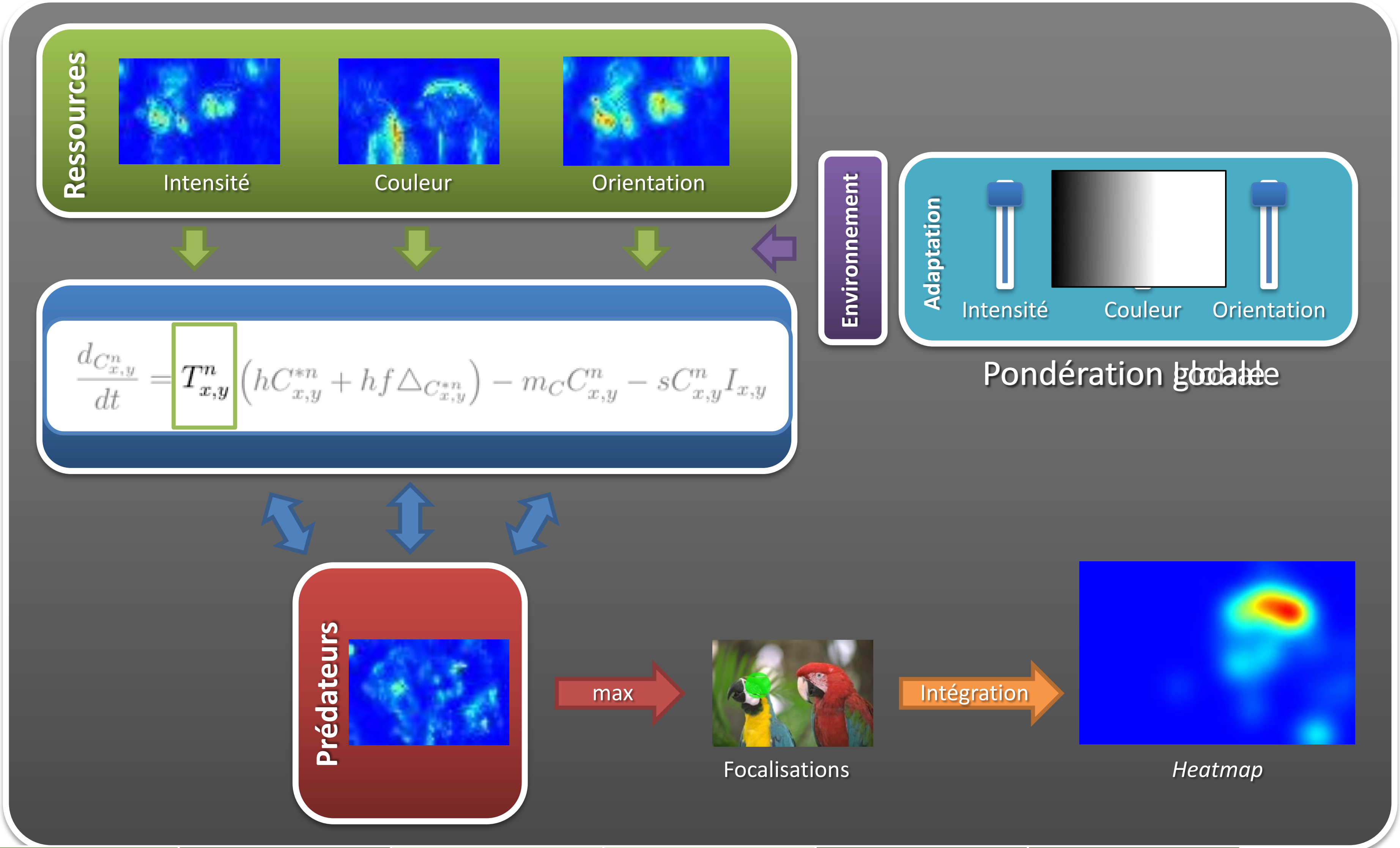
Extensibilité

Mouvement


Ressources
Intensité
Couleur
Orientation
Mouvement
Environnement
Biais gaussien
Bruit
Proies
Prédateurs
max
Focalisations
Intégration
Heatmap

Adaptation

Basique



Ressources

Intensité

Couleur

Orientation

$$\frac{dC^n_{x,y}}{dt} = \boxed{T^n_{x,y}}\left(hC^{*n}_{x,y} + hf\Delta_{C^{*n}_{x,y}}\right) - m_C C^n_{x,y} - sC^n_{x,y}I_{x,y}$$

Environnement

Adaptation

Intensité

Couleur

Orientation

Pondération globale locale

Prédateurs

max

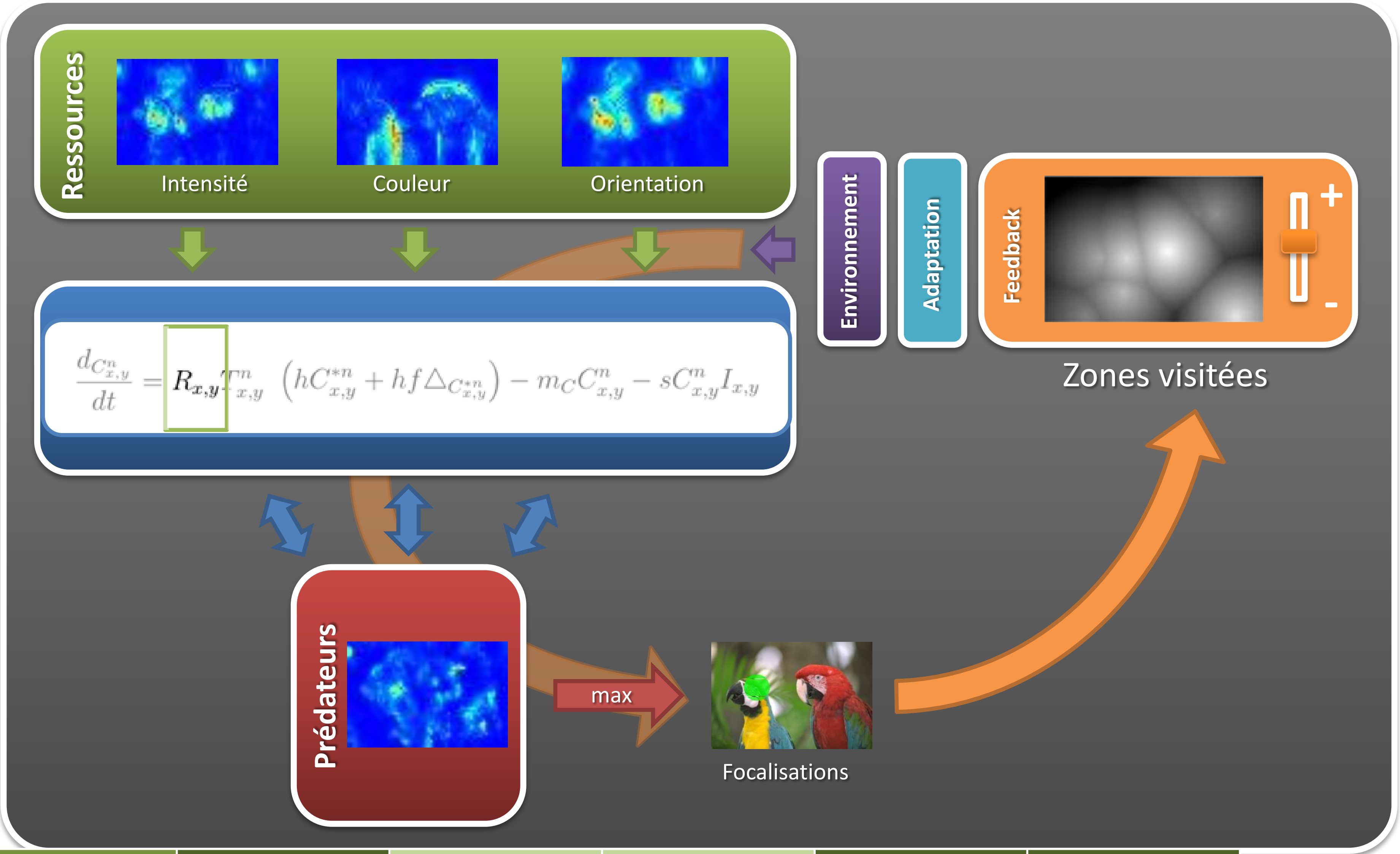
Focalisations

Intégration

Heatmap

Adaptation

Feedback



Ressources

Intensité

Couleur

Orientation

$$\frac{d_{C^{n}_{x,y}}}{dt} = \boxed{\boldsymbol{R_{x,y}}} T^{n}_{x,y} \left(hC^{*n}_{x,y} + hf\Delta_{C^{*n}_{x,y}}\right) - m_C C^{n}_{x,y} - sC^{n}_{x,y} I_{x,y}$$

Environnement

Adaptation

Feedback

+

-

Zones visitées

Prédateurs

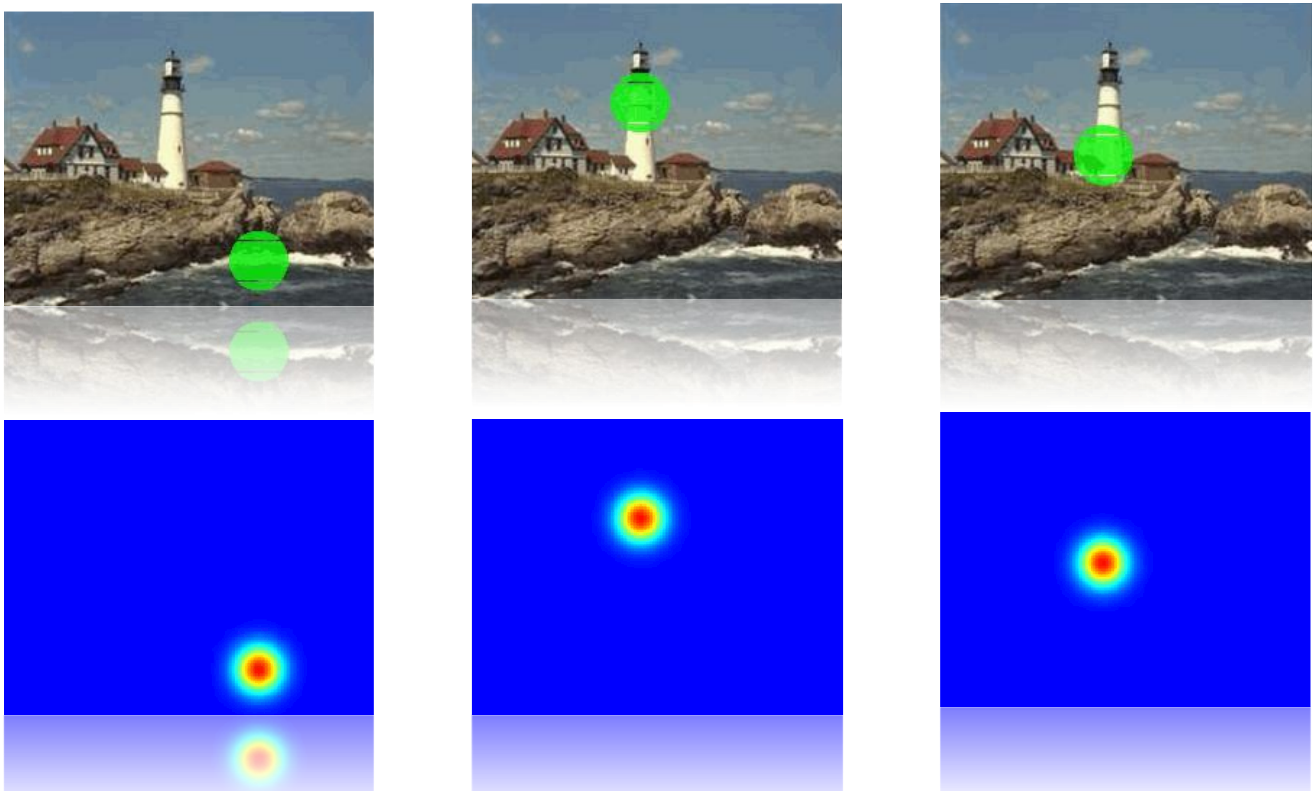
max

Focalisations

Adaptation

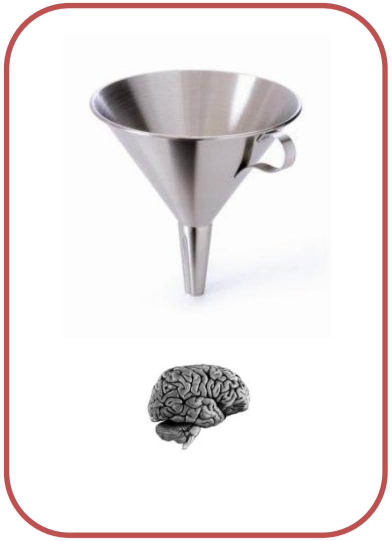
Feedback

- Critère « Exploration de l'espace » : $R(x,y) = \begin{cases} \frac{1+|F_{feedback}| \times M_{visites}(x,y)}{1+|F_{feedback}|} & si\ F_{feedback} \geq 0 \\ \frac{1+|F_{feedback}| \times (1-M_{visites}(x,y))}{1+|F_{feedback}|} & sinon \end{cases}$

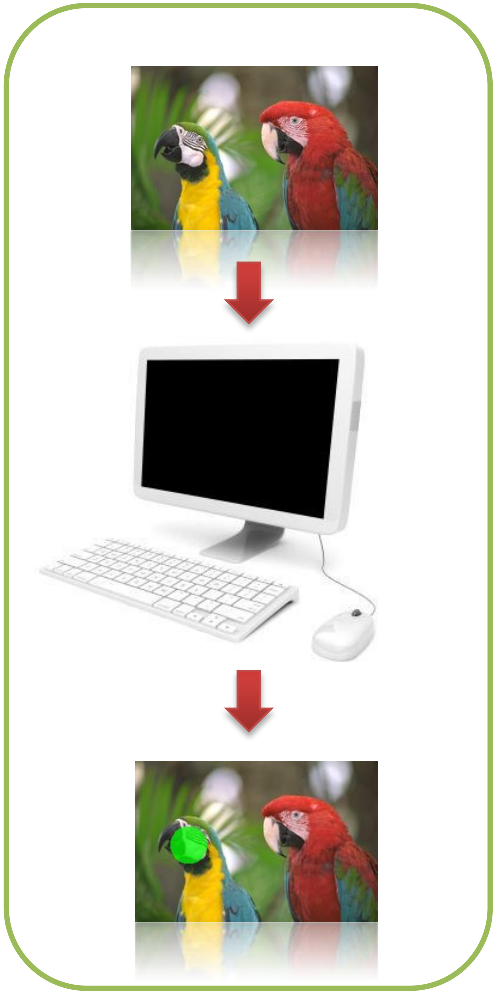


$F_{feedback} = -1.0$ $F_{feedback} = 0$ $F_{feedback} = 1.0$

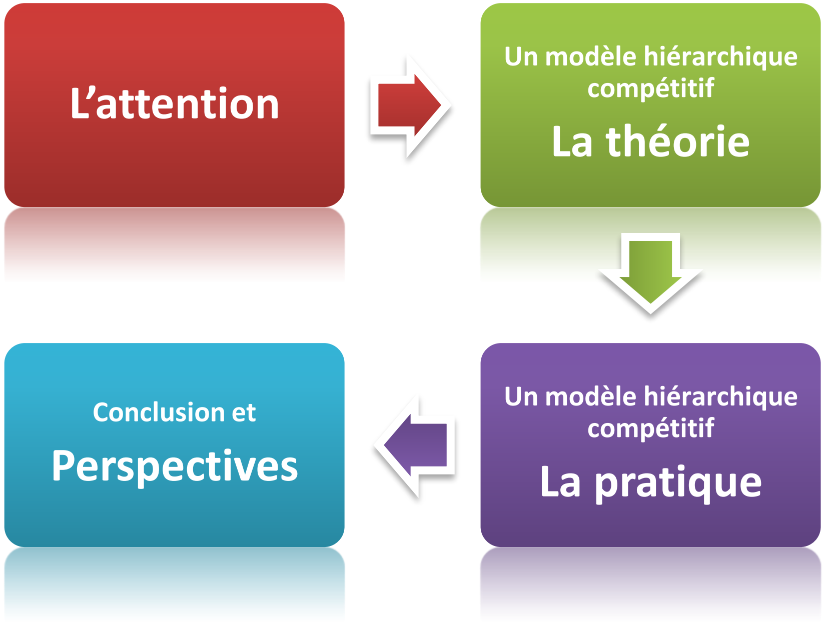
Un modèle computationnel d’attention



Théorique



Computationnel



L'attention

Un modèle hiérarchique compétitif
La théorie

Un modèle hiérarchique compétitif
La pratique

Conclusion et
Perspectives

UN MODÈLE HIÉRARCHIQUE COMPÉTITIF

LA PRATIQUE

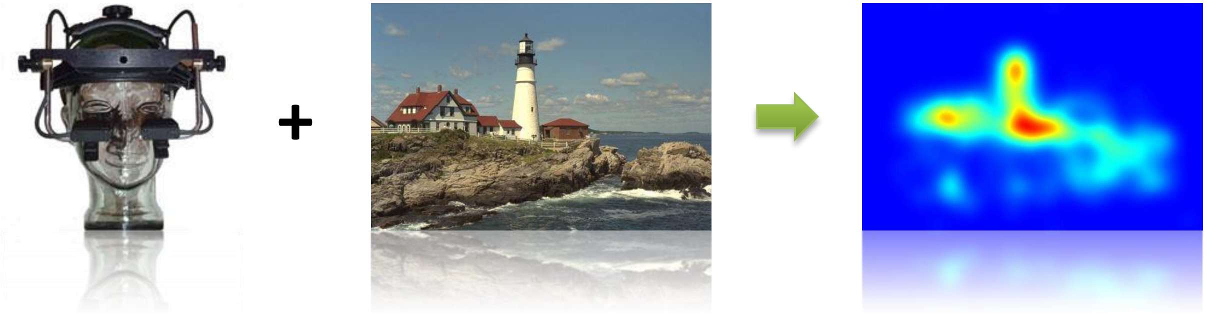
Méthodologie d'évaluation

Propriété	Mesure	Base
Plausibilité	Objective : Corrélation Divergence de Kullback-Leibler *Normalized Scanpath Salience*	Bruce, Le Meur
	Subjective : Scoring	Flickr
Stabilité	Moyenne des carte de proies / prédateurs	-
Reproductibilité	Corrélation, DKL, NSS	Bruce, Le Meur
Exploration	Taux de compression JPEG / PNG Somme des valeurs absolues des différences	Bruce, Le Meur
Dynamique	Proies / prédateurs : Temps de démarrage	Bruce, Le Meur
	Fixations : Déplacement moyen Temps de fixation moyen	

Plausibilité

Mesures objectives

- **Vérité terrain**

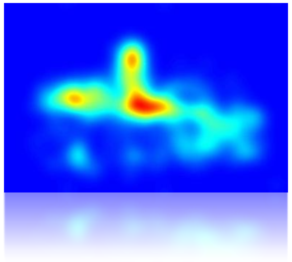


- Algorithmes évalués

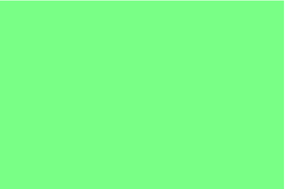
Plausibilité

Mesures objectives

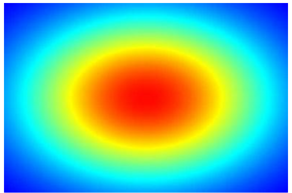
- Vérité terrain



- **Algorithmes évalués**

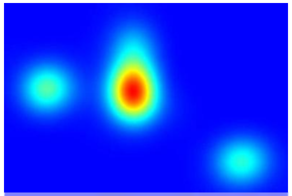


Constant

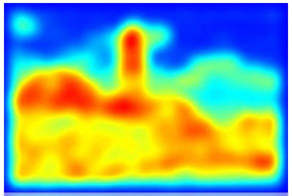


Gaussien

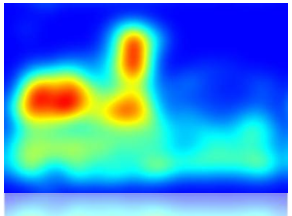
Naifs



Itti

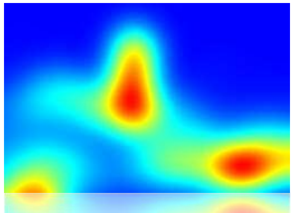


Bruce

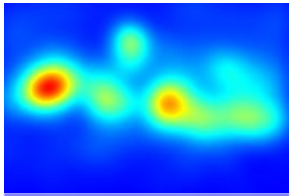


Le Meur

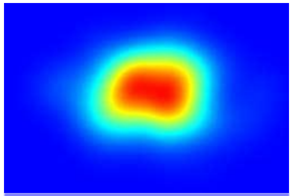
Etat de l’art



Hiérarchique



Défaut + rétine



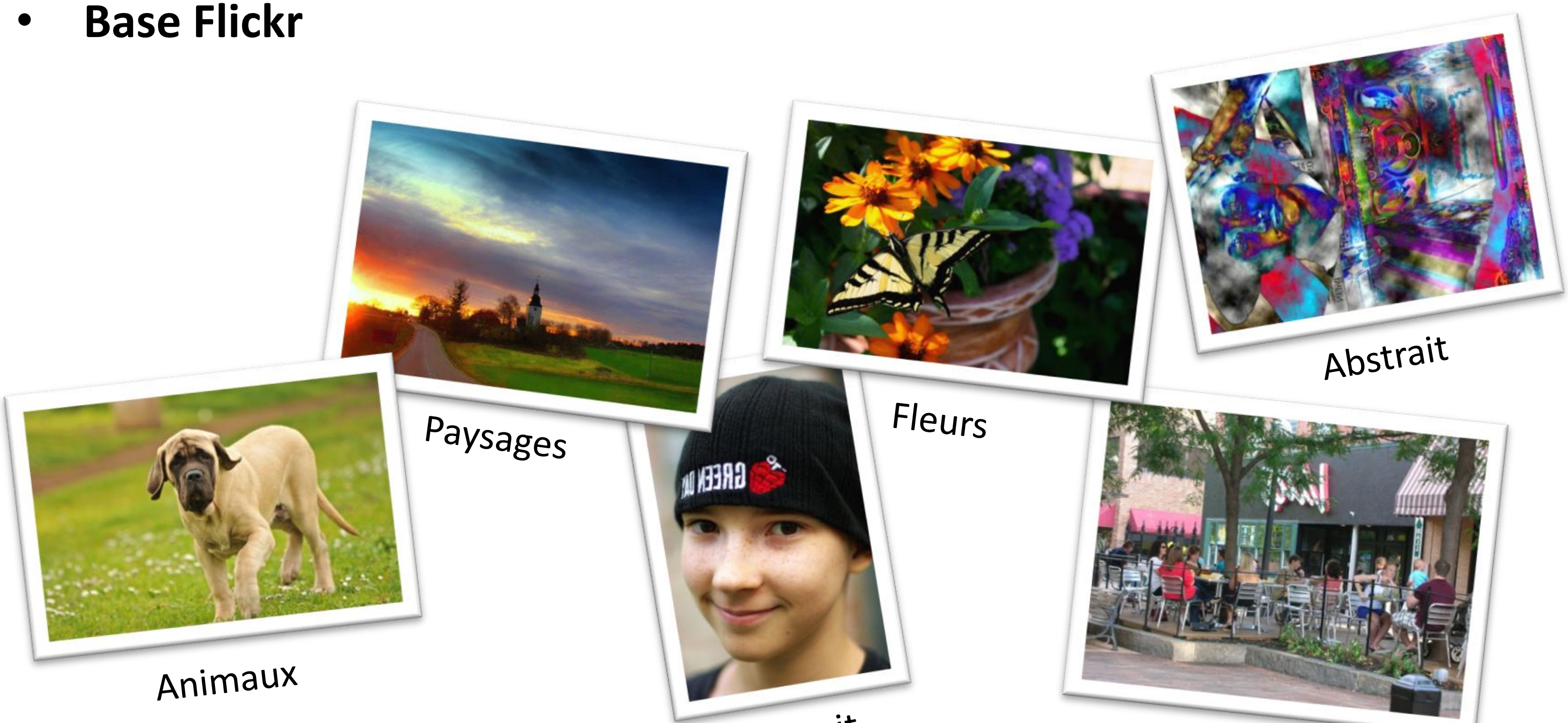
Plausibilité « maximale »

Perreira

Plausibilité

Evaluation subjective

- **Base Flickr**



Abstrait

Paysages

Fleurs

Animaux

Portrait

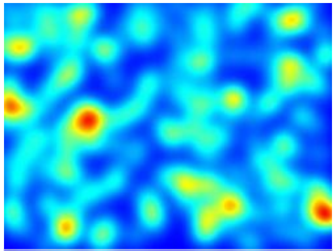
Ville

- Algorithmes
- Méthode

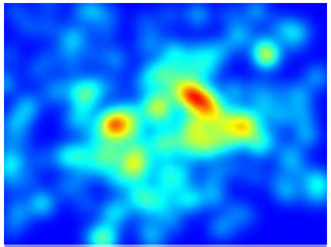
Plausibilité

Evaluation subjective

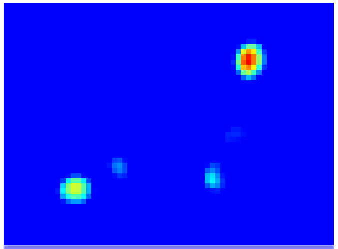
- Base Flickr
- **Algorithmes**
- Méthode



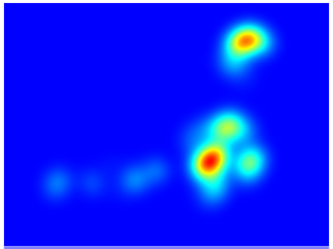
Aléatoire



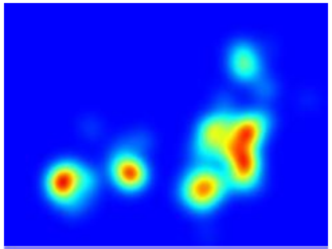
Aléatoire +
Biais central



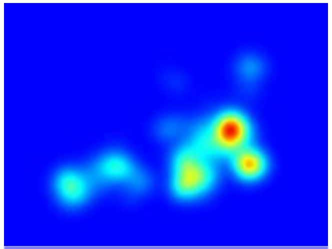
Itti



Itti +
Proies / Prédateurs



Proies / Prédateurs
+ Flou rétinien



Proies / Prédateurs

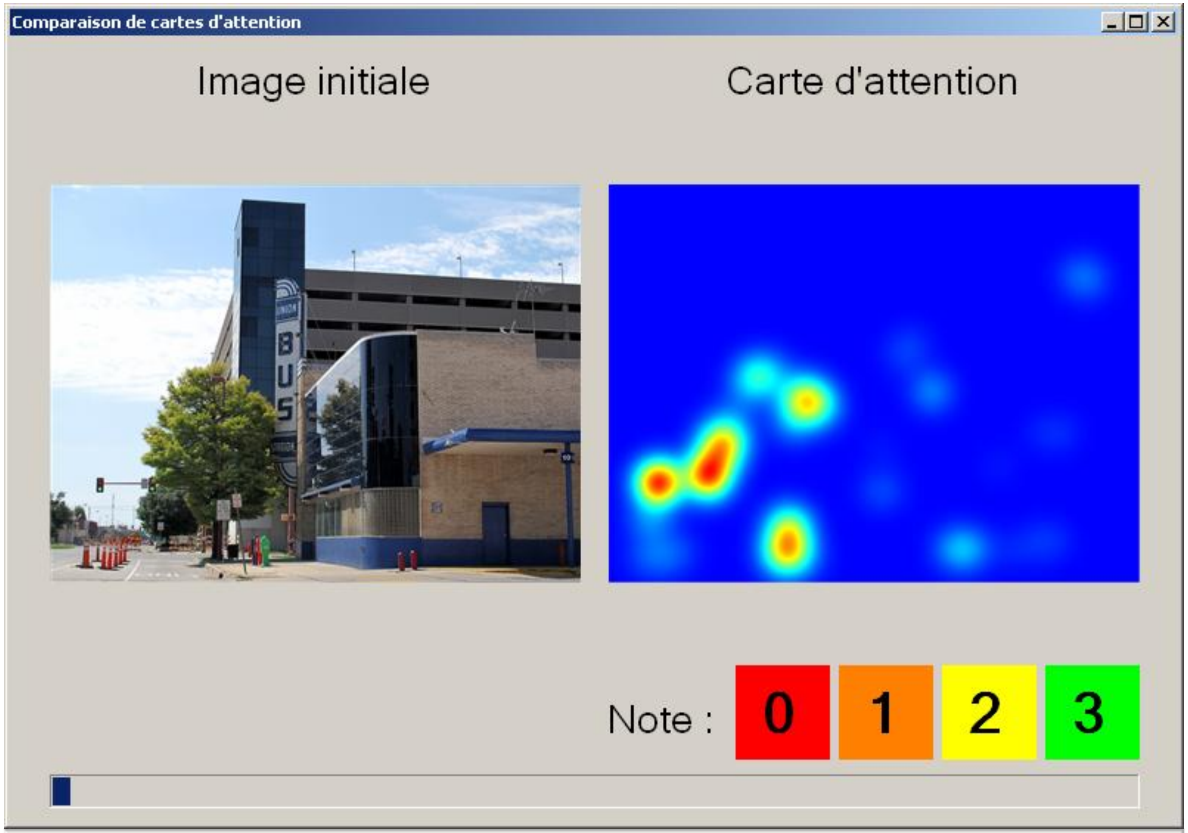
Plausibilité

Evaluation subjective

- Base Flickr
- Algorithmes

- **Méthode**

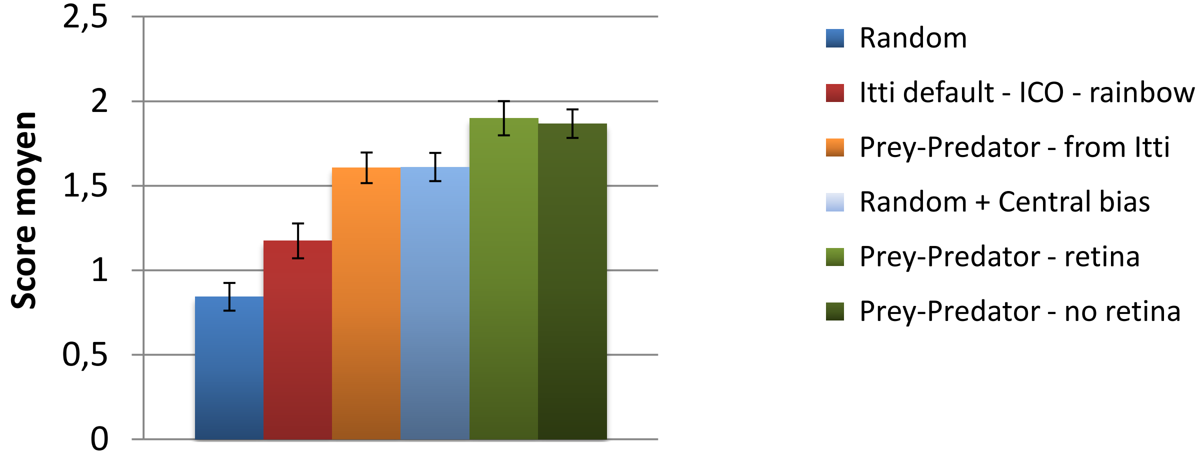
Scoring

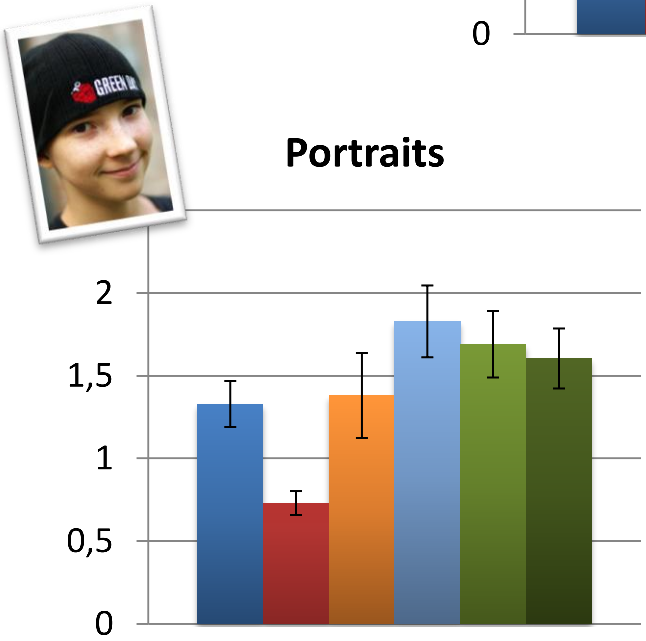

Comparaison de cartes d'attention
Image initiale
Carte d'attention
Note :
0
1
2
3

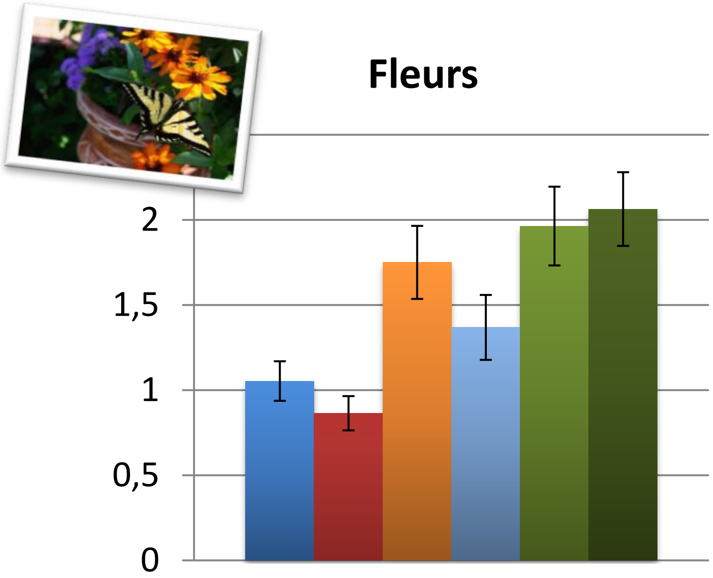
Plausibilité

Evaluation subjective

Résultats globaux

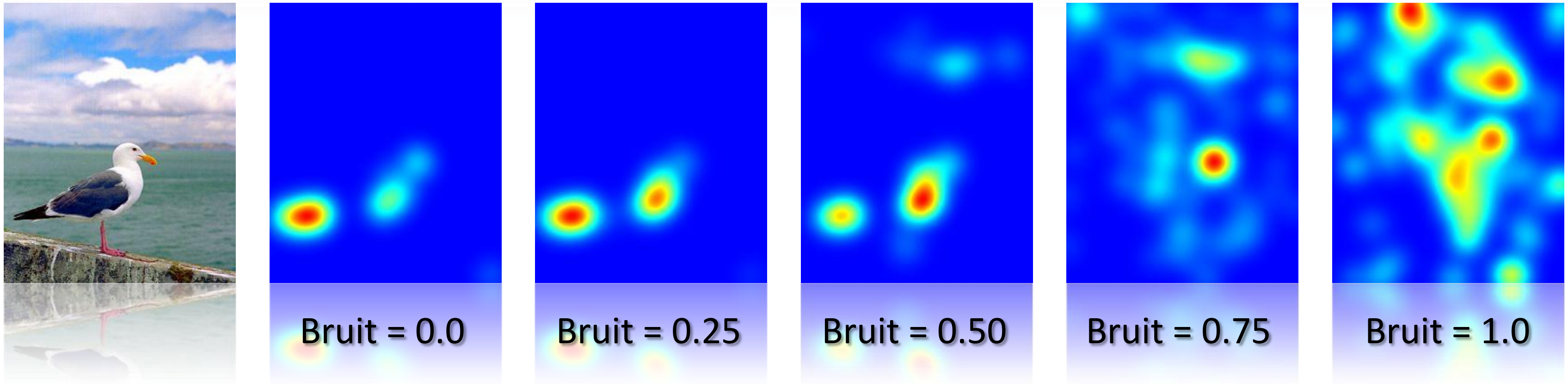






Propriétés du modèle

Paramètres	Plausibilité	Reproductibilité	Exploration	Dynamique
Flou rétinien	⬀	⬂	⬀	⇨
Biais central	⇧	⇨	⇧	⇩
Diffusion	⇨	⬂	⬀	⬂
Feedback	⇨	⇨	⇧	⬂
Bruit	⇧/⇩	⇩	⇧	⬀



Bruit = 0.0

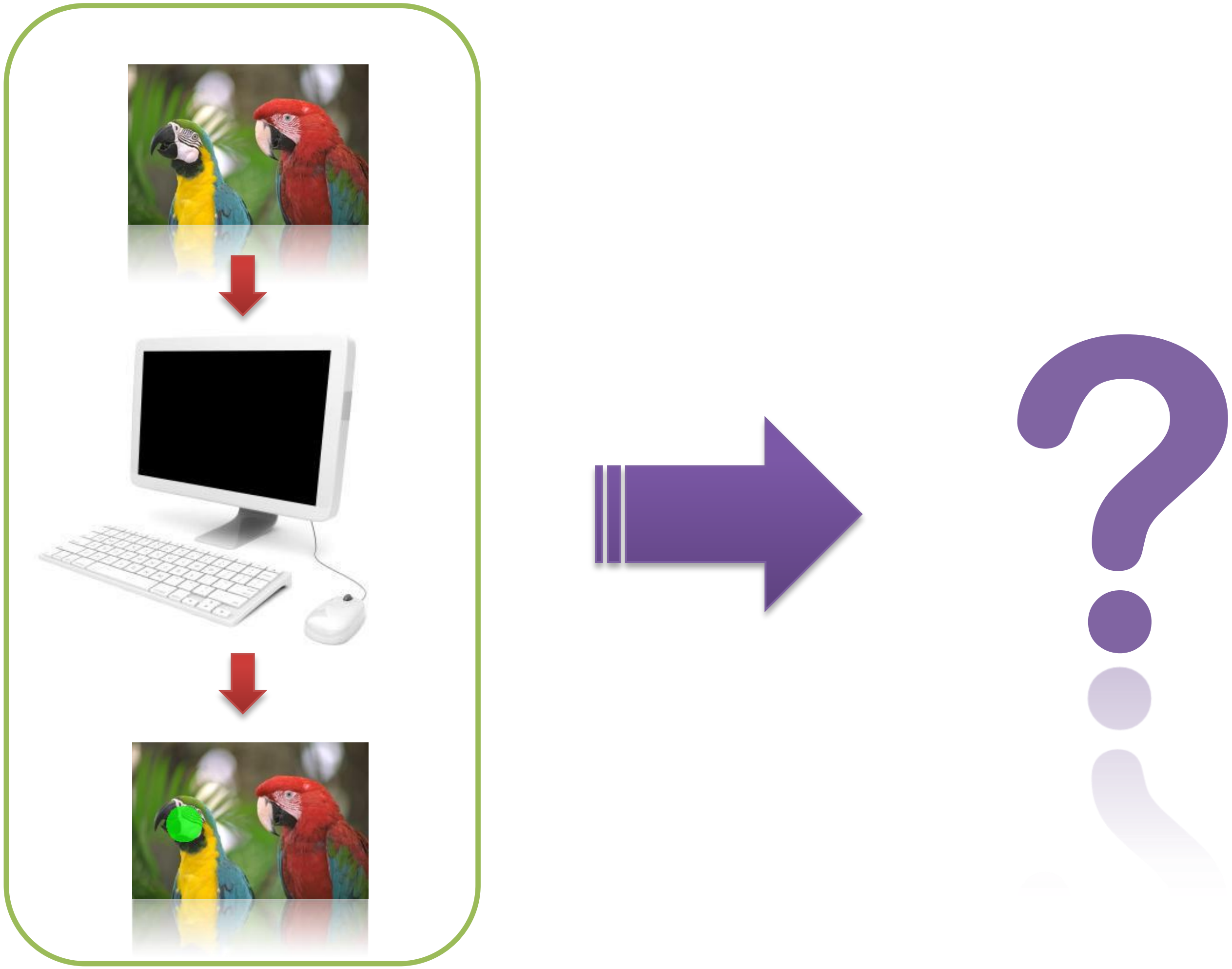
Bruit = 0.25

Bruit = 0.50

Bruit = 0.75

Bruit = 1.0

Des applications…



Des applications…

- Complexité des images

Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
○ ●	○ ○	○ ●	○ ○	○ ○	● ●

- Recadrage dynamique

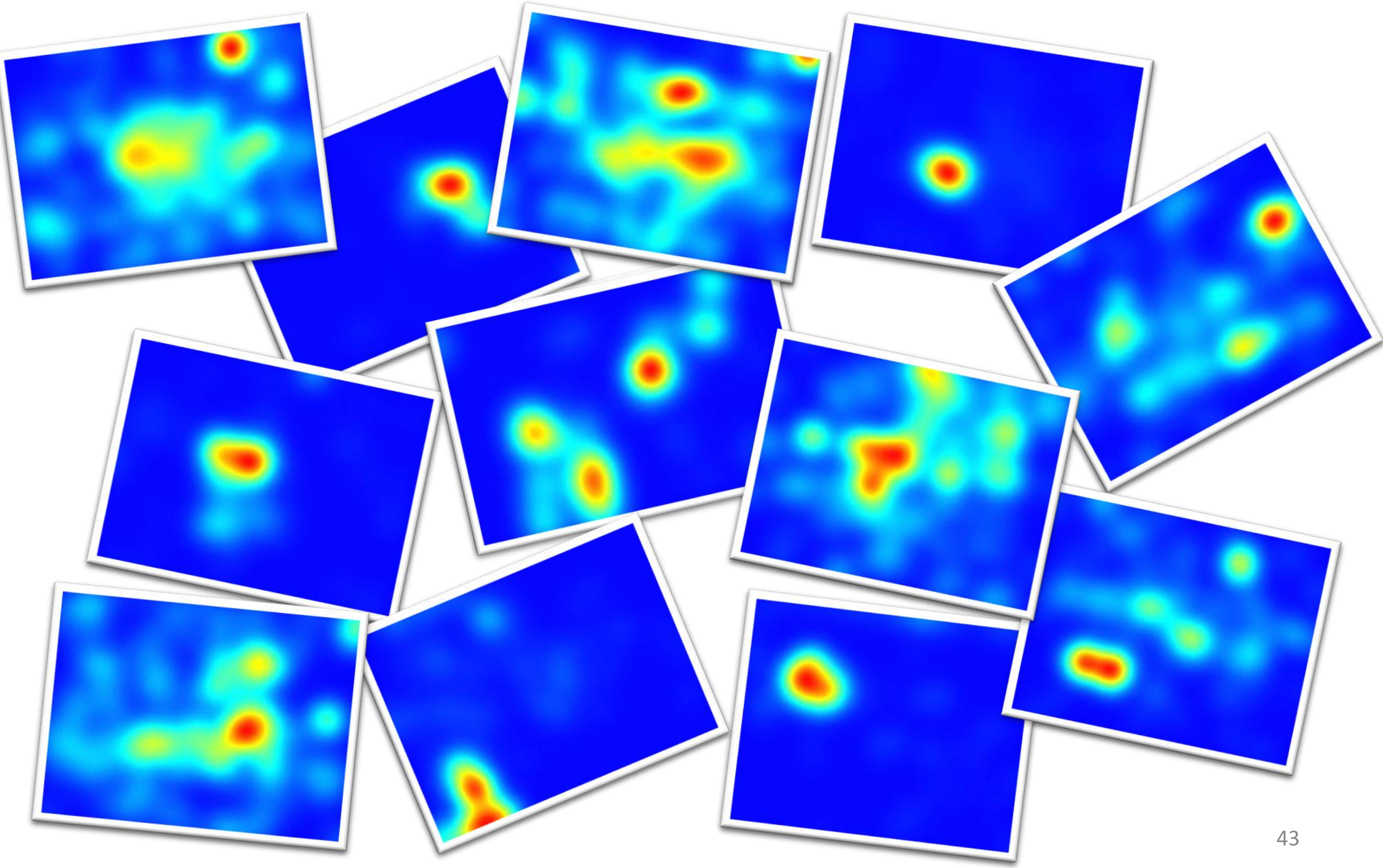
Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
● ●	● ●	○ ●	● ●	○ ●	● ●

- Segmentation de vidéo

Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
○ ●	● ●	○ ●	● ●	● ●	● ●

Applications

Complexité des images

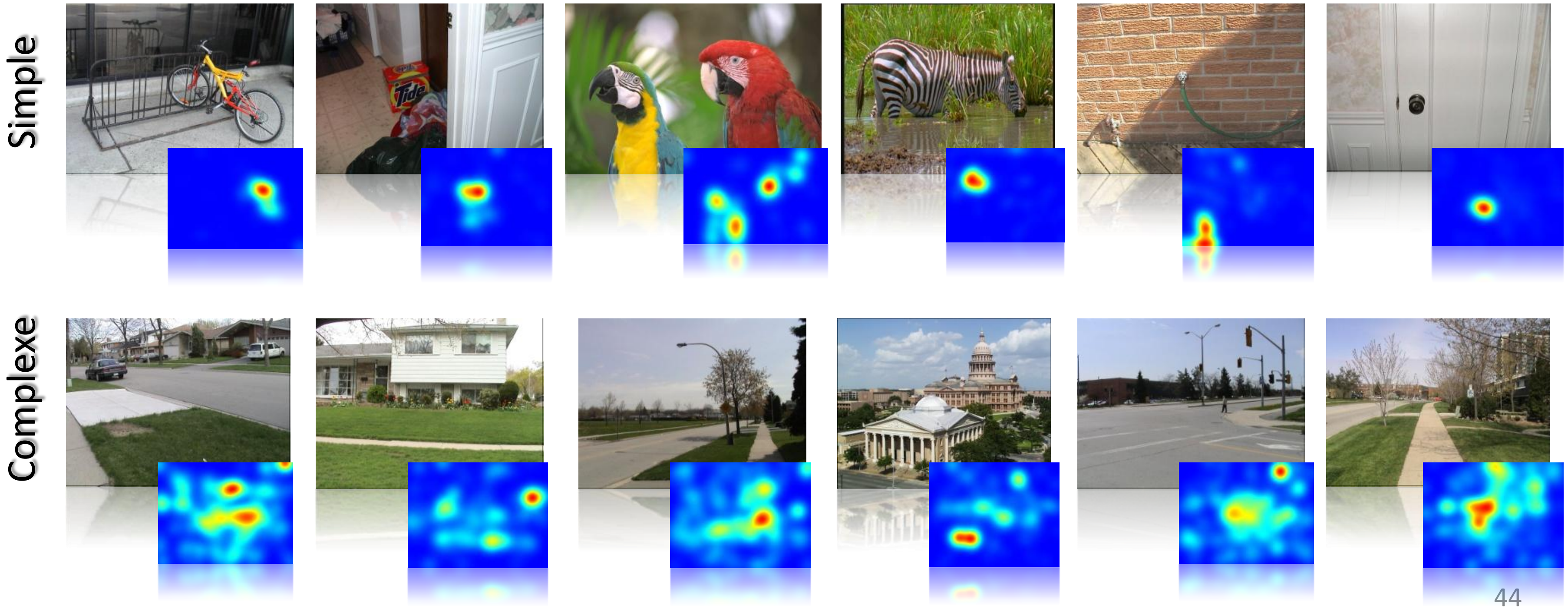


Applications

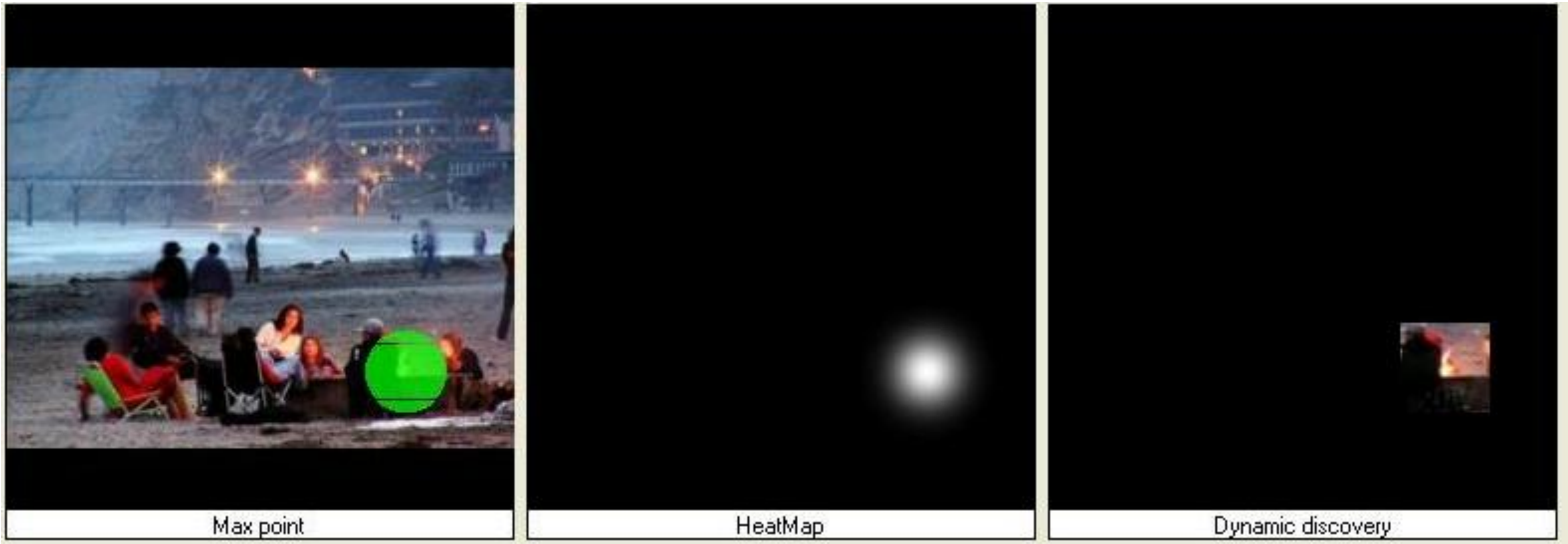
Complexité des images

- Deux méthodes d'estimation
  - Taux de compression des coordonnées des focalisations
  - Entropie de la longueur des saccades

- Combinaison des deux méthodes
  - Exemple après 300 itérations



Applications Recadrage dynamique d'images


Max point
HeatMap
Dynamic discovery

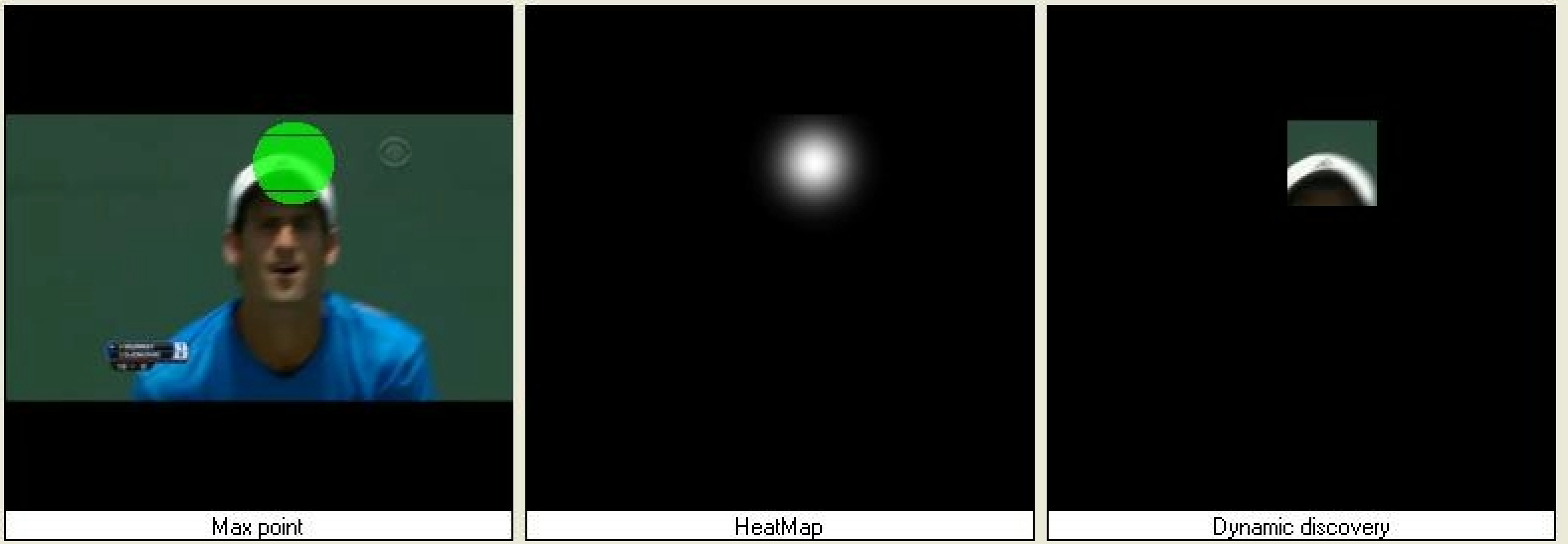
Feedback

Focalisation

Exploration

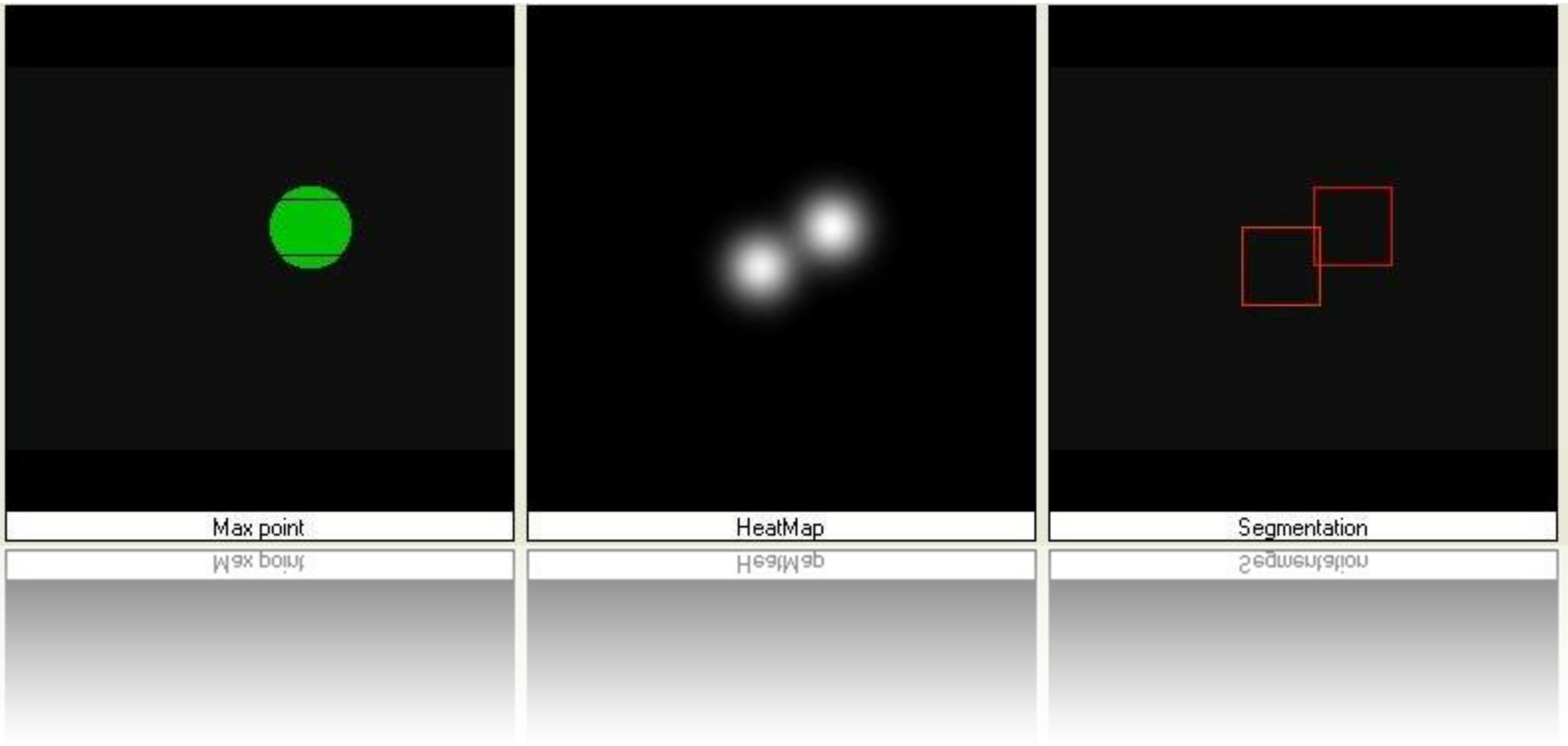
Applications

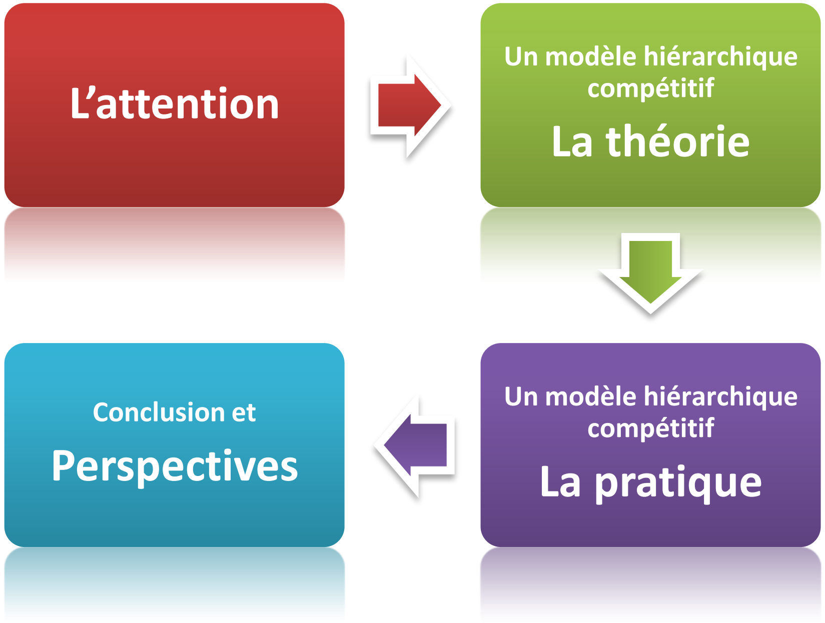
Recadrage dynamique de vidéo


Max point
HeatMap
Dynamic discovery

Applications

Segmentation de vidéo


Max point
HeatMap
Segmentation



L’attention

Un modèle hiérarchique compétitif
La théorie

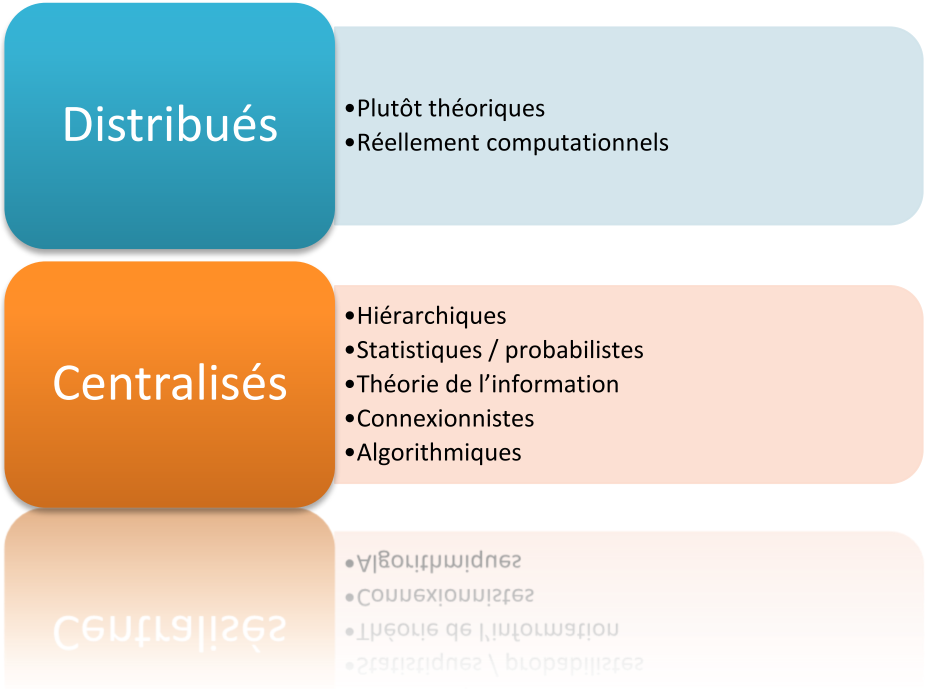
Un modèle hiérarchique compétitif
La pratique

Conclusion et
Perspectives

CONCLUSION ET PERSPECTIVES

Contributions

- **Une Taxonomie**



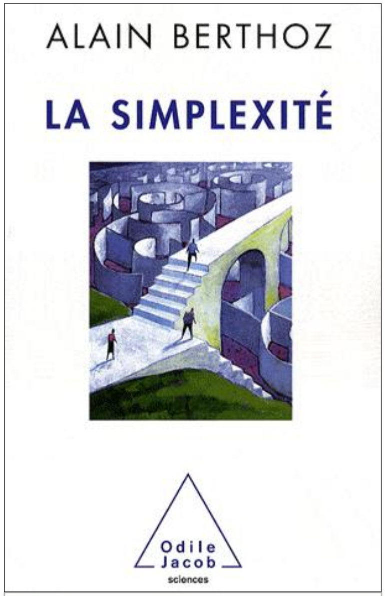
Distribués
- Plutôt théoriques
- Réellement computationnels

Centralisés
- Hiérarchiques
- Statistiques / probabilistes
- Théorie de l’information
- Connexionnistes
- Algorithmiques

- Un cadre *simplexe*
- Modèle computationnel Hiérarchique / Compétitif
- Caractérisation du modèle
- Applications du modèle

Contributions

- Une Taxonomie
- **Un cadre *simplexe***



ALAIN BERTHOZ

LA SIMPLEXITÉ

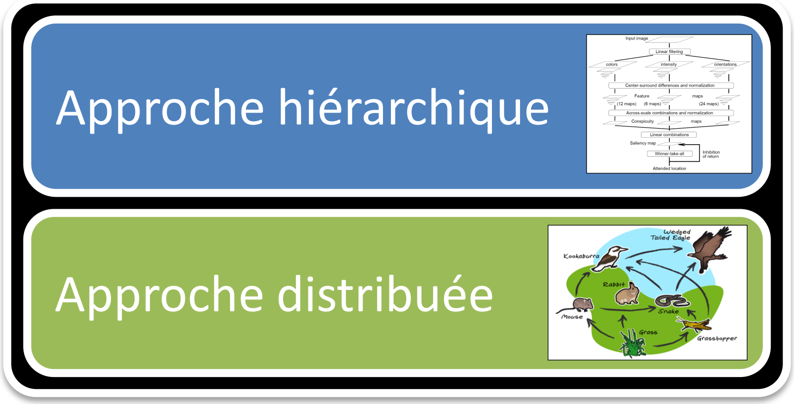
Odile
Jacob
sciences

- Modèle computationnel Hiérarchique / Compétitif
- Caractérisation du modèle
- Applications du modèle

Contributions

- Une Taxonomie
- Un cadre *simplexe*

- **Modèle computationnel Hiérarchique / Distribué**
  - Vitesse, flou rétinien et normalisation non paramétrique
  - Fusion d’informations hétérogènes par compétition
  - Dynamique des focalisations
  - Adaptation par rebouclage

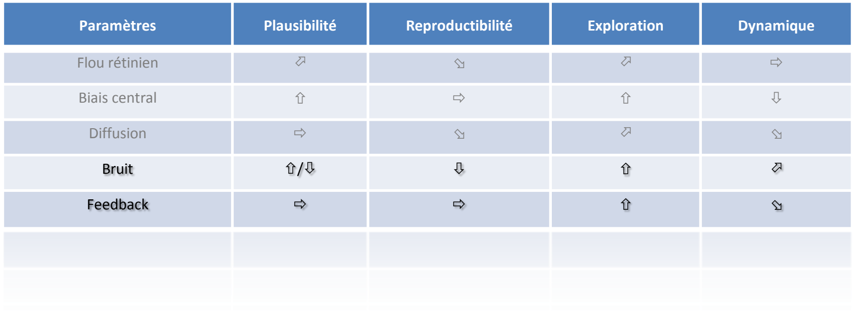

Approche hiérarchique

Approche distribuée
Input image
Linear filtering
colors
intensity
orientations
Center-surround differences and normalization
Feature maps
(12 maps)
(6 maps)
(24 maps)
Across-scale combinations and normalization
Conspicuity maps
Linear combinations
Saliency map
Winner-take-all
Inhibition of return
Attended location
Wedged Tailed Eagle
Kookaburra
Rabbit
Snake
Mouse
Grass
Grasshopper

- Caractérisation du modèle
- Applications du modèle

Contributions

- Une Taxonomie
- Un cadre *simplexe*
- Modèle computationnel Hiérarchique / Compétitif
- **Caractérisation du modèle**
  - Plausibilité
  - Stabilité
  - Reproductibilité
  - Exploration
  - Dynamique

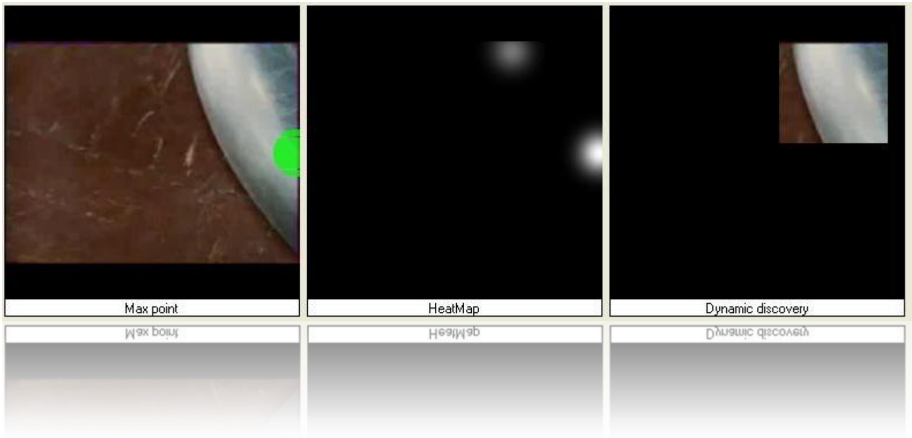


Paramètres	Plausibilité	Reproductibilité	Exploration	Dynamique
Flou rétinien	⬀	⬂	⬀	⇨
Biais central	⇧	⇨	⇧	⇩
Diffusion	⇨	⬂	⬀	⬂
Bruit	⇧/⇩	⇩	⇧	⬀
Feedback	⇨	⇨	⇧	⬂

- Applications du modèle

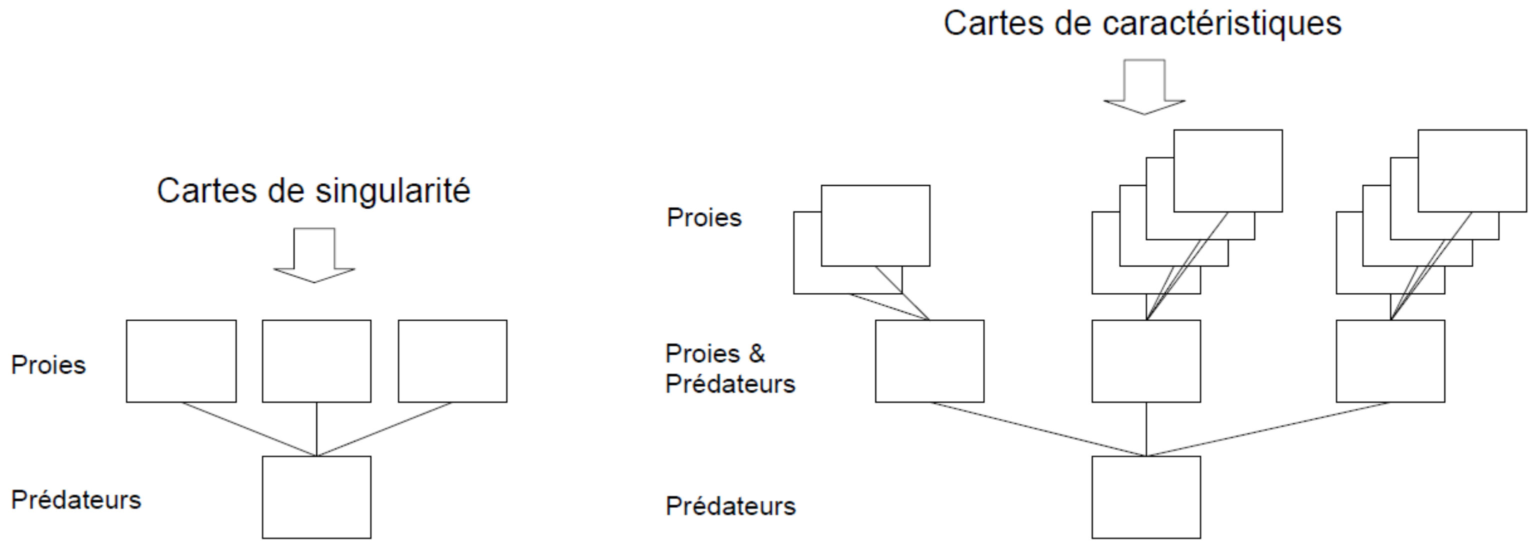
Contributions

- Une Taxonomie
- Un cadre *simplexe*
- Modèle computationnel Hiérarchique / Compétitif
- Caractérisation du modèle
- **Applications du modèle**
  - Complexité des images
  - Recadrage dynamique
  - Segmentation de vidéos


Max point
HeatMap
Dynamic discovery

Limites et Perspectives

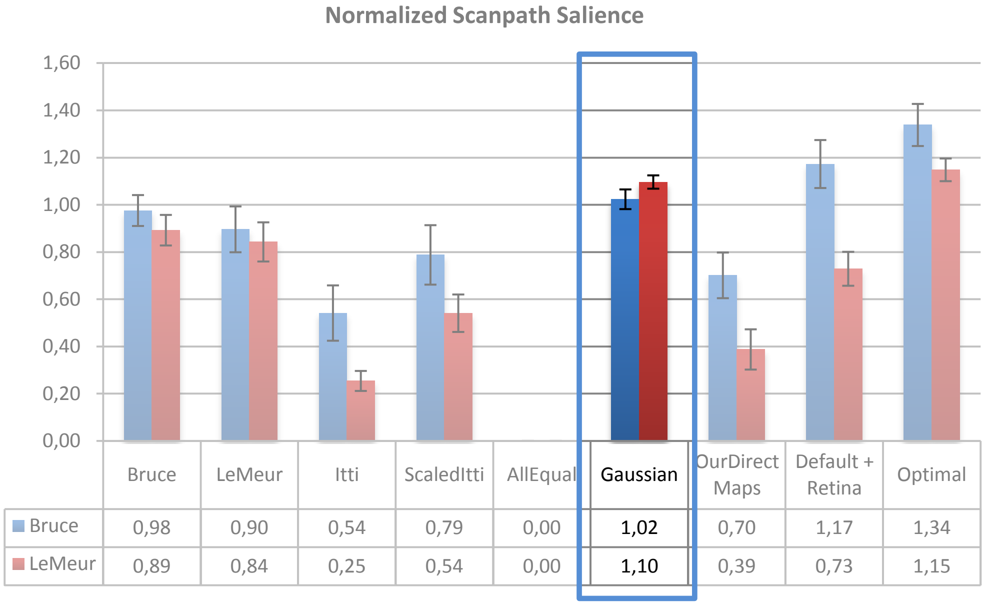
- **Modèle**
  - Adaptation → cartes de caractéristiques


Cartes de singularité
Proies
Prédateurs
Cartes de caractéristiques
Proies
Proies &
Prédateurs
Prédateurs

  - Priorisation des voies statiques et dynamiques [Bur2007,Marat2009]
- Evaluation
- Modèle de vision adaptatif

Limites et Perspectives

- Modèles
- **Evaluation**
  - Biais central : *Ordinal dominance score* [Bamber1975,Berg2009]
  - Vidéo / dynamique



Normalized Scanpath Salience

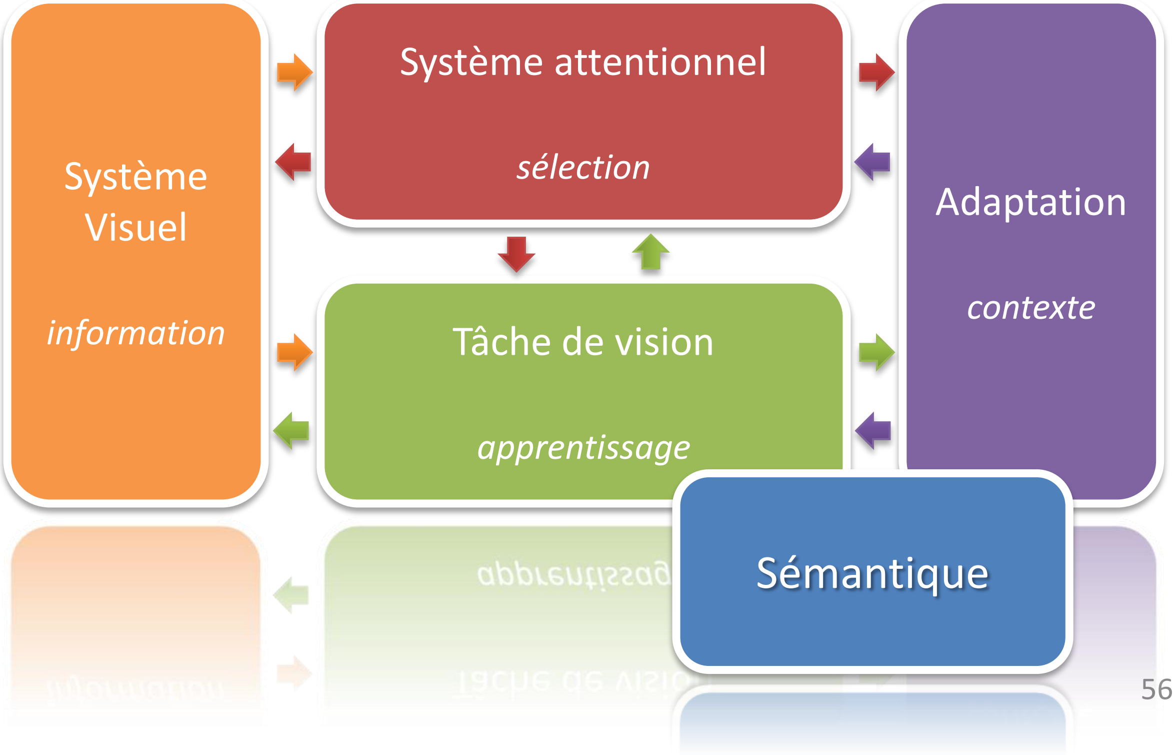
	Bruce	LeMeur	Itti	ScaledItti	AllEqual	Gaussian	OurDirect Maps	Default + Retina	Optimal
🟦 Bruce	0,98	0,90	0,54	0,79	0,00	1,02	0,70	1,17	1,34
🟥 LeMeur	0,89	0,84	0,25	0,54	0,00	1,10	0,39	0,73	1,15

- Modèle de vision adaptative

Limites et Perspectives

- Modèles
- Evaluation

- **Modèle de vision adaptative**
  - Représentation de l'information
  - Réutilisation
  - Mémorisation
  - Apprentissage
  - **Sémantique**


Système Visuel
information
Système attentionnel
sélection
Tâche de vision
apprentissage
Adaptation
contexte
Sémantique

Merci de votre *attention*…

ANNEXES

Publications

Revues

○ M. Perreira Da Silva, V. Courboulay, A. Prigent, P. Estraillier «Fast, low resource, head detection and tracking for interactive applications», PsychNology Journal, 2009, Volume 7, Number 3, 243-264

○ N. Rempulski, A. Prigent, P. Estraillier, V. Courboulay and M. Perreira Da Silva, «Adaptive Storytelling based on model-checking approaches», IJIGS, 2009, Volume 5, Number 2, 33-42

Conférences Internationales avec actes

▪ M. Perreira Da Silva, V. Courboulay, A. Prigent, P. Estraillier, «Evaluation of preys / predators systems for visual attention simulation», VISAPP 2010.

○ M. Perreira Da Silva, V. Courboulay, A. Prigent, P. Estraillier, «Real-time face tracking for attention aware adaptive games», ICVS 2008

○ M. Perreira Da Silva, V. Courboulay, A. Prigent, «Gameplay experience based on a gaze tracking system», COGAIN 2007

Conférences Nationales avec actes

▪ M. Perreira Da Silva, V. Courboulay, A. Prigent, P. Estraillier, «Système temps réel de simulation d’attention visuelle : application aux images et séquences d’images», RFIA 2010.

▪ M. Perreira Da Silva, V. Courboulay, A. Prigent, P. Estraillier, «Attention visuelle et systèmes proies / prédateurs», GRETSI 2009

▪ M. Perreira Da Silva, V. Courboulay, A. Prigent, P. Estraillier, «Adaptativité et interactivité : vers un système de vision comportemental», MajecSTIC 2008

○ A. Ould Mohamed, M. Perreira Da Silva, V. Courboulay, M. Ménard, «Intégration d’un modèle d’inattention dans une plateforme de rééducation adaptative pour enfants autistes», ASSISTH’2007

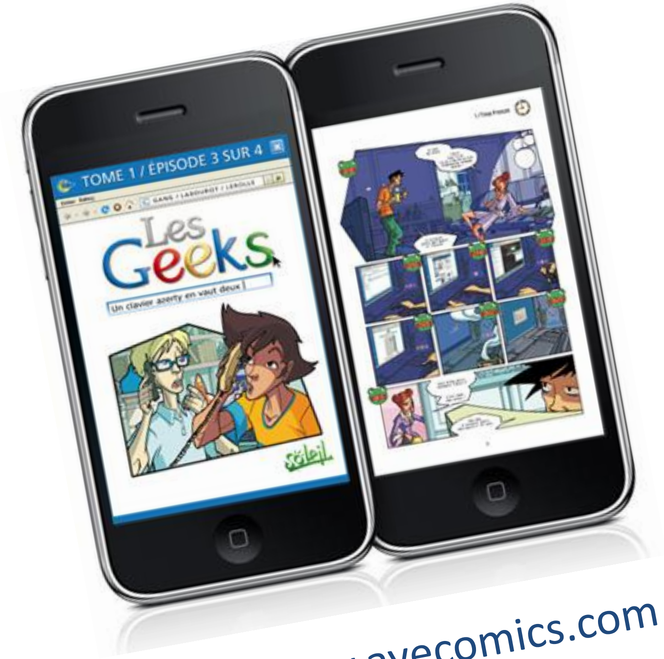
Vers plus d’applications…

	Fidèle	Adaptable	Invariant	Rapide	Extensible	Dynamique
Etude de l’attention	●●●	●	●	●	●	●●●
Ergonomie / publicité	●●●	●	●	●	●●	●●
Vision / interaction	●●	●●●	●●	●●●	●●●	●●●
CBIR	●	●	●●●	●●	●●●	●
Traitement d’image	●●	●●	●●	●●	●●●	●●

Vers plus d'applications…

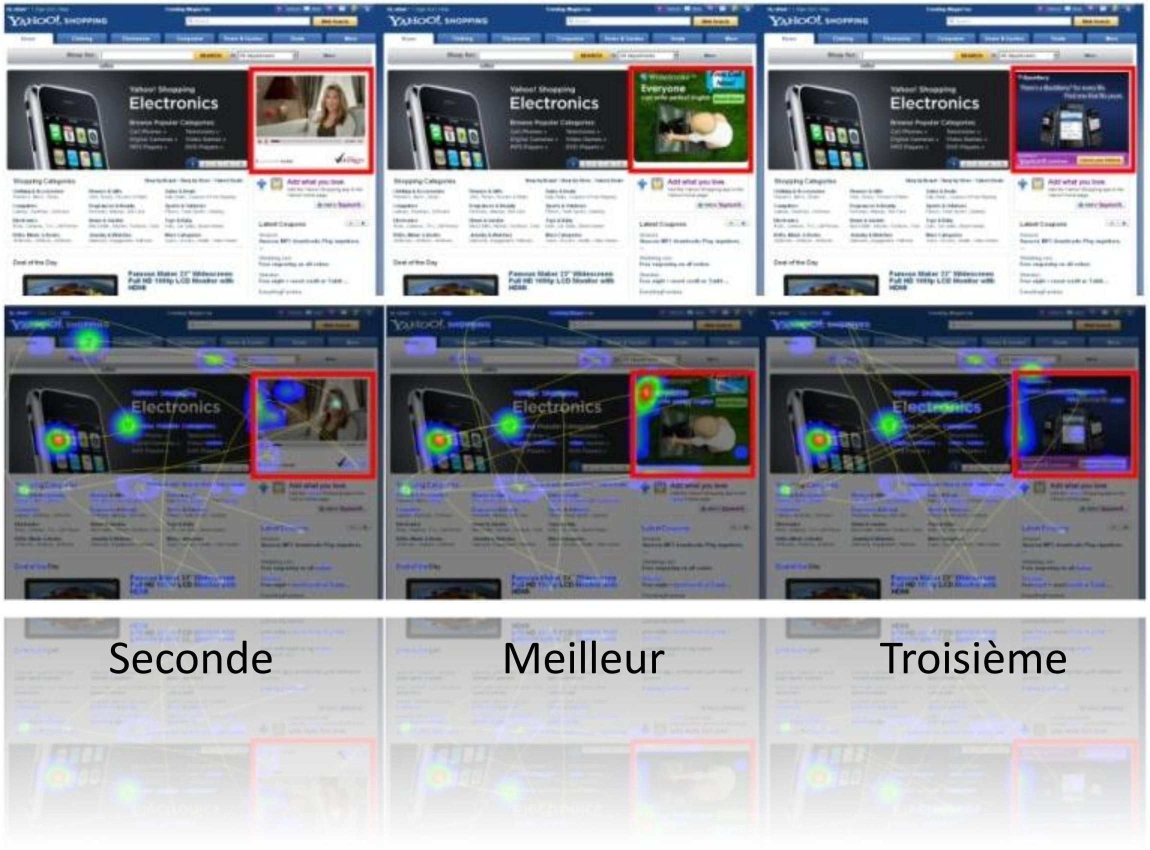
- **Ergonomie / publicité**

- Interaction
- CBIR



TOME 1 / ÉPISODE 3 SUR 4
Les Geeks
Un clavier azerty en vaut deux

http://www.avecomics.com



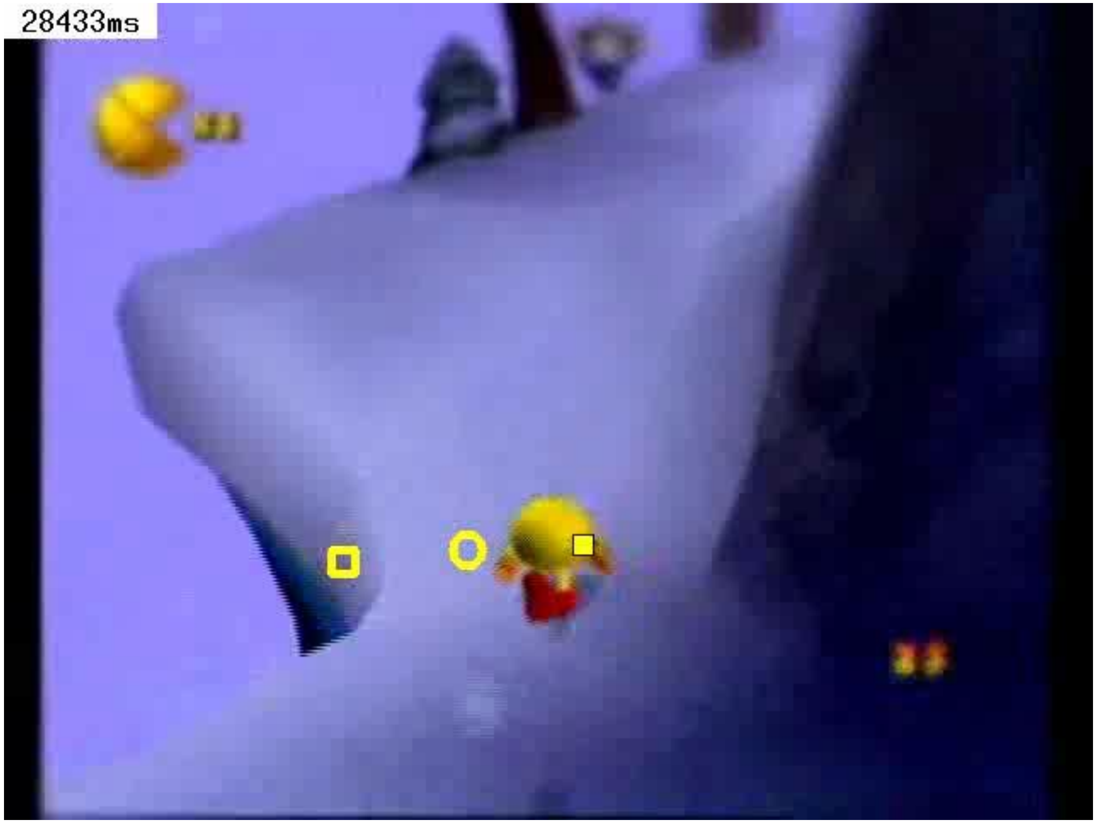
YAHOO! SHOPPING
Electronics

Seconde Meilleur Troisième

http://www.feng-gui.com

Vers plus d'applications…

- Ergonomie / publicité
- **Vision / interaction**
- CBIR

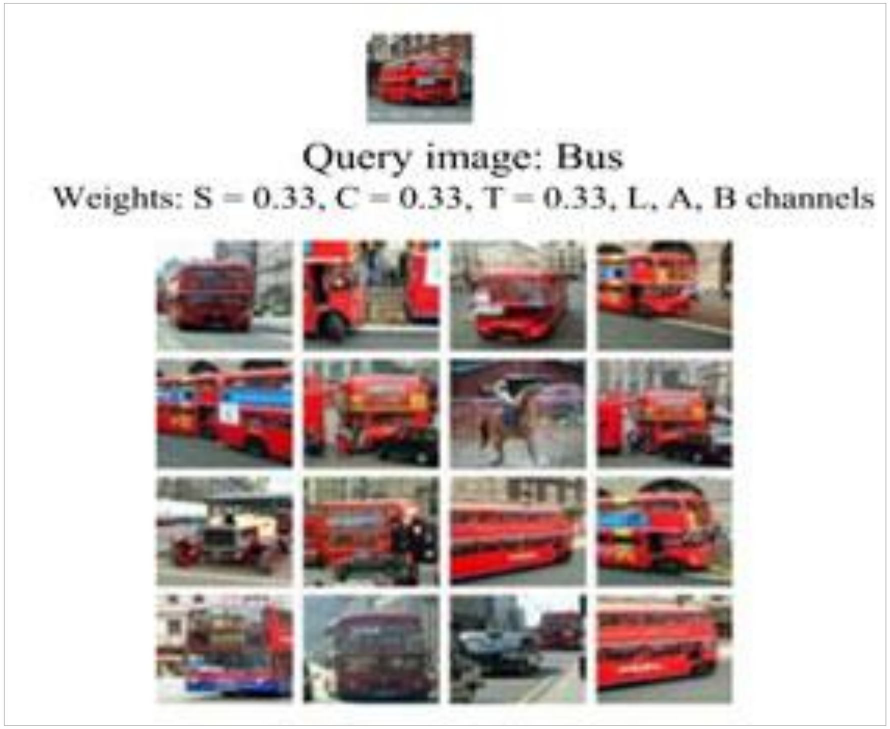


[Itti2003]

Vers plus d'applications…

- Ergonomie / publicité
- Vision / interaction

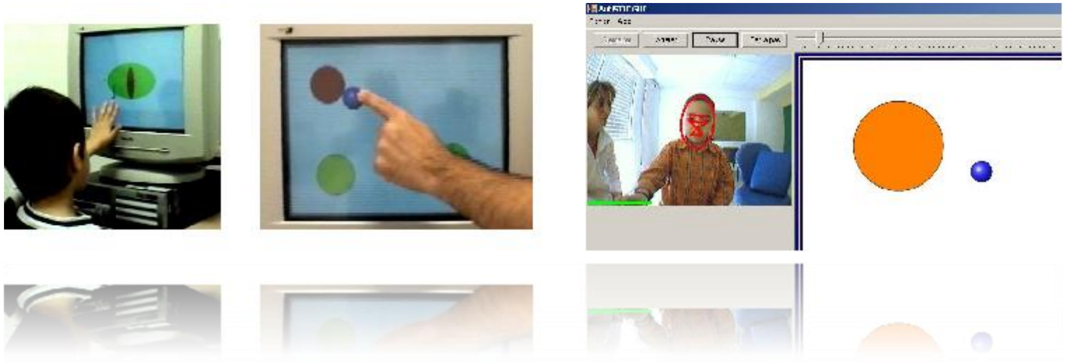
- **CBIR**


Query image: Bus
Weights: S = 0.33, C = 0.33, T = 0.33, L, A, B channels

[Bamidele 2004]

Pourquoi un modèle proies / prédateurs ?

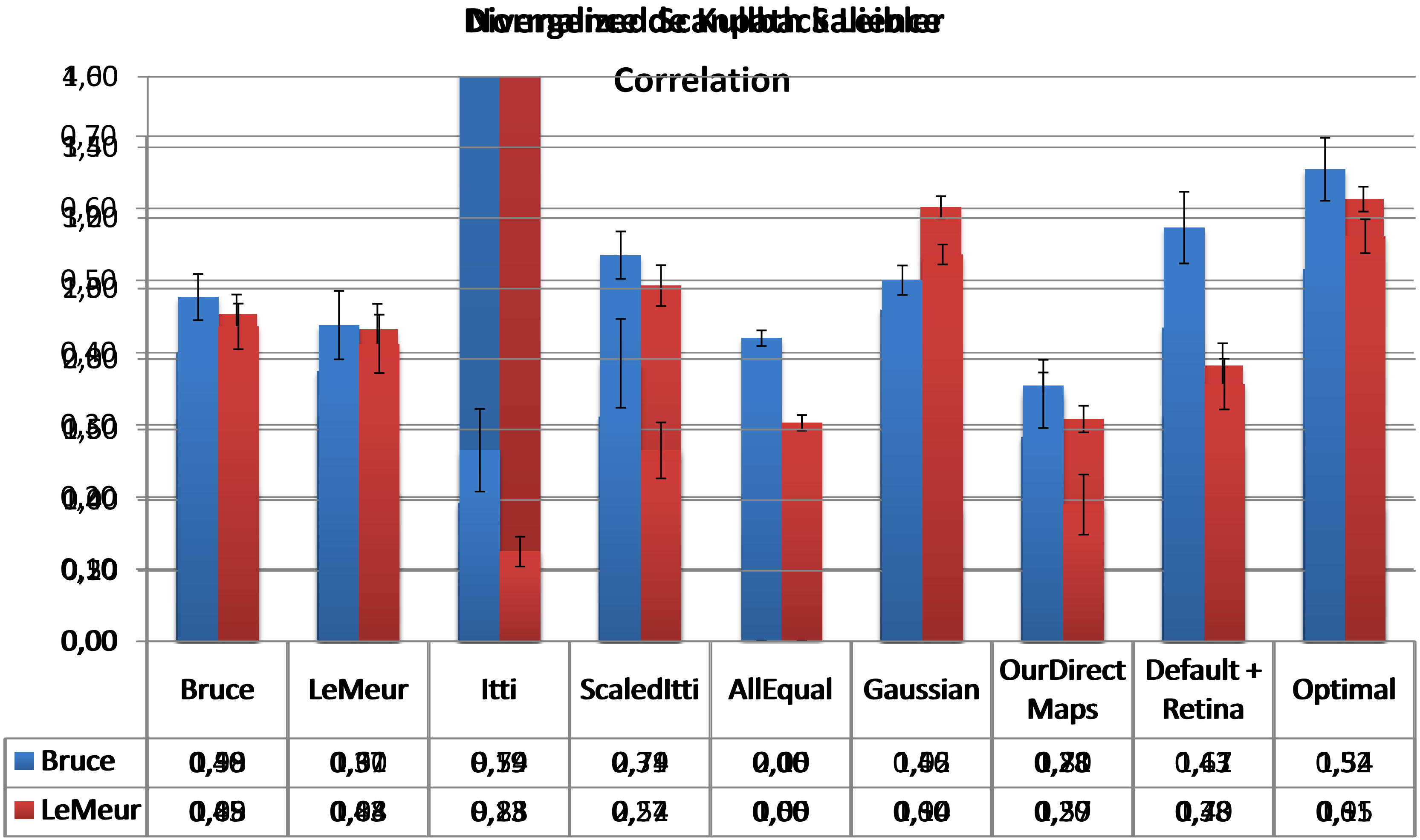
- 2006-2007 : projet Orange Autistic



- « Autism & computing » Mike Lesser et Dinah Murray
  - Monotropisme vs. Polytropisme [Murray2005]
  - « Mind as a dynamical system » [Lesser1998]

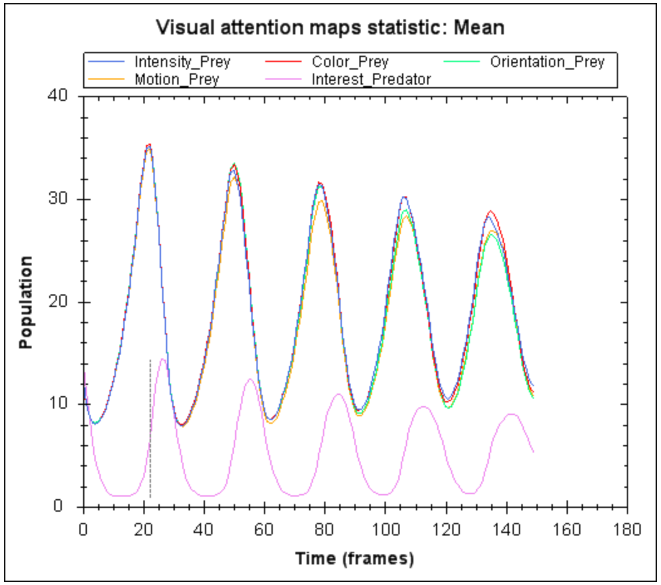
Plausibilité

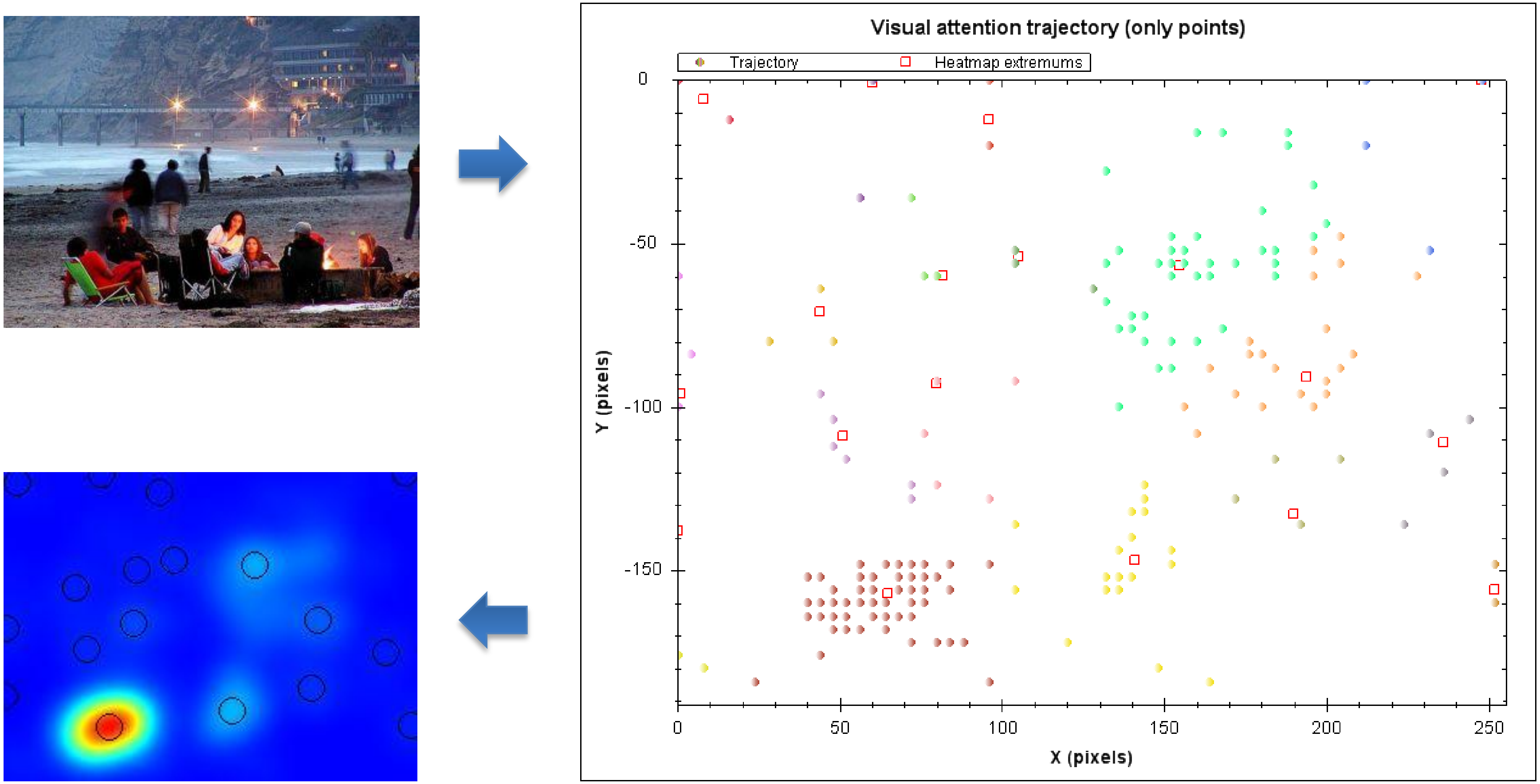
Mesures objectives



Stabilité

Stabilité ?



Clustering des focalisations