

Chauvin, A., Hérault, J. et Marendaz, C. Perception des scènes naturelles : effet des attracteurs visuels.
Laboratoire de Psychologie Expérimentale CNRS UMR 5105, BP 47, 38040 Grenoble Cedex 9.
Laboratoire des Images et des Signaux CNRS UMR 5083, 46 Av Félix Viallet, 38031 Grenoble.
Tel : 04 76 82 78 31, e-mail : chauvin@lis-viallet.inpg.fr

Nous cherchons à défendre l'hypothèse que les premières fixations seraient dirigées et prédictibles à partir du signal lumineux uniquement. Nous réalisons un modèle prédictif du comportement normal d'exploration que nous validons par une mesure comportementale des mouvements oculaires. Notre modèle ne prédit pas les invariants comportementaux mais un ensemble de variabilités comportementales qui constituent le plus petit ensemble commun à tous les comportements.

La catégorisation des scènes naturelles est réalisée très rapidement. En effet, le traitement visuel requis peut être réalisé en moins de 150 ms [Thorpe et al (1996)]. Cette rapidité de traitement implique premièrement que la catégorie de l'image est inconnue lors des deux premières saccades oculaires (La latence moyenne d'une saccade est de 135 ms) et deuxièmement que le traitement catégoriel est purement ascendant [Thorpe et Imbert (1989)]. La catégorisation rapide des scènes naturelles est « portée » par les composantes basses fréquences du spectre des images naturelles [Oliva (1995), Beaudot(1996)] et peu sensible à l'information de couleur donc cette première catégorisation rapide serait initiée par les voies magno-cellulaires achromatiques, basses fréquences et rapides. De plus une analyse plus poussée de la scène se fera selon l'échelle optimale de la scène qui dépend du contenu de la scène. La sélection de l'échelle dépend donc de la structure des objets analysés durant les premières fixations.

C'est pourquoi nous postulons que les premières saccades d'exploration d'une scène hors contexte sont soit aléatoires soit dirigées par l'information physique et dirigées par les basses fréquences. Les fixations suivantes s'adapteront à l'échelle optimale d'analyse de la scène et utiliseront les information de plus hautes fréquences des scènes naturelles.

Pour cela nous avons modélisé le traitement réalisé par les filtres corticaux visuels et nous prédisons les points et les régions d'intérêts en fonction du contenu fréquentielle des scènes. Dans un second temps nous avons recueilli les fixations oculaires lors de l'observation de ces mêmes scènes. Nous avons analysé les comportements individuels d'exploration ainsi que les régions délimités par les densités de fixations.

Les observations indiquent une bonne adéquation entre les régions prédites et les régions d'intérêt des sujets. Les points de fixations lors des explorations sont relativement bien corrélé avec les prédictions. Mais certains comportement atypique d'observation ne sont ni prédictible par notre modèle ni expliqué. De plus, la variabilité comportementale lors d'une observation n'est prédite que pour la position des fixations, mais rien ne permet de prédire le décours et les temps de fixations et l'ordre de leur occurrence. En faisant varier les consignes nous pouvons contraindre ces comportements et les « normalisés », mais cette approche n'a pas été développée.

Thorpe, S., Fize et al (1996). "Speed of Processing in the Human visual System". *Nature*, **381**, 520-522.

Oliva, A. (1995). Perception de scènes: Traitement fréquentiel du signal visuel aspects psychophysiques et neurophysiologiques. Thèse de l'institut national polytechnique de Grenoble.

Beaudot, W. H. A. (1996). Dynamique du traitement visuel dans la voie rétino-corticale : multiplexage spatio-temporel de la luminance et de la chrominance. *Reports from NSI 96*, 165-168.