

Dynamique du traitement visuel dans la voie rétino-corticale : Multiplexage spatio-temporel de la luminance et de la chrominance *

William H.A. Beaudot

CSEM Centre Suisse d'Electronique et de Microtechnique SA,

Jaquet-Droz 1, CH-2007, Neuchâtel, Switzerland,

Tél/Fax: +41 38 205 251/763; E-mail: william.beaudot@csemne.ch

Résumé

L'inséparabilité spatio-temporelle est une des caractéristiques du traitement réalisé par le système visuel humain. Nous montrons comment cette propriété émerge de la structure du modèle de rétine achromatique de Beaudot-Hérault et comment elle engendre une perception de type *Coarse-to-Fine* dès la sortie du modèle rétinien. Cette analyse "coarse-to-fine" s'avère être par ailleurs le mode de traitement privilégié par le système visuel humain aussi bien de bas-niveau que de haut-niveau. L'extension du modèle à la perception de la couleur permet de rendre compte du multiplexage spatial des informations de luminance et de chrominance toutes deux présentes dans la voie parvocellulaire. La combinaison de ces deux mécanismes se traduit par une modulation spectrale et temporelle du traitement spatial. En combinant les différents canaux rétiens de façon appropriée, nous montrons enfin comment démultiplexer cet entrelacement d'informations afin de fournir les signaux qui devraient être exploités par les sous-systèmes corticaux spécialisés dans l'analyse de la luminance et de chrominance.

1 Introduction

La propriété de séparabilité spatio-temporelle dans la perception visuelle, dont la fonction s'exprimerait comme le produit $g(x, t) = f(x).h(t)$ d'un filtre purement spatial $f(x)$ et d'un filtre purement temporel $h(t)$, ne fait toujours pas l'unanimité [1, 2]. L'intérêt d'une inséparabilité spatio-temporelle dans le traitement visuel n'a toutefois pas encore été clairement mis en évidence pour justifier une remise en cause des théories basées sur la séparabilité spatio-temporelle. Comment la propriété d'inséparabilité peut-elle apparaître et en quoi est-elle utile ? En particulier, quelle est son influence sur le codage de l'information visuelle et la dynamique de son traitement ? Nous nous proposons dans cet article de montrer comment elle peut émerger d'un modèle rétinien et quelles en sont les conséquences sur le traitement des informations de luminance et de chrominance dans la voie rétino-corticale.

2 Le modèle de rétine achromatique de Beaudot-Hérault

Un modèle spatio-temporel de la rétine achromatique des vertébrés a été proposé dans [3]. Outre sa sensibilité aux contours et aux mouvements, le filtrage réalisé par la couche plexiforme externe (PLE) du modèle se caractérise par une inséparabilité spatio-temporelle. Cette propriété émerge de la structure du modèle qui "calque" l'architecture neuronale de la rétine des vertébrés : il consiste en deux réseaux résistifs et capacitifs (R-C) représentant respectivement, la couche des photorécepteurs et la couche des cellules horizontales, dont les constantes d'espace et de temps ont été fixées à des valeurs biologiquement plausibles pour la rétine des primates (respectivement 10 ms et 20 ms pour les constantes de temps). La différence entre ces deux couches produit la sortie des cellules bipolaires qui fournit la sortie des cellules ganglionnaires de type X. Le filtrage spatio-temporel ainsi réalisé est de type passe-bande et à variables non séparables (figure 1a).

La propriété qui résulte de cette inséparabilité spatio-temporelle est une dépendance temporelle de la fonction de transfert spatiale du traitement rétinien (figure 1c) : lorsqu'un échelon temporel de stimulation visuelle (par exemple entre deux saccades oculaires) est présenté en entrée du système, la sortie du modèle est le résultat d'un filtrage spatial d'abord passe-bas avec une fréquence de coupure élevée, qui converge au cours du temps vers un filtrage spatial passe-bande. Nous observons ainsi que le pic des basses fréquences spatiales apparaît 14 ms après le début de la stimulation lumineuse et que les hautes fréquences spatiales ne commencent à être discernées qu'à

*Publié dans les actes des 8ièmes Journées NEUROSCIENCES ET SCIENCES DE L'INGÉNIEUR, pp. 165-168, Marly-le-Roi, France, 6-9 Mai 1996.

partir d'environ 30 ms. Le modèle nous suggère aussi que l'extraction des hautes fréquences spatiales au niveau rétinien est sévèrement limitée pour une stimulation de durée inférieure à 30 ms.

La sortie de type Y du modèle est quant à elle fournie par convolution du signal issu de la PLE à travers la couche plexiforme interne (PLI) qui réalise un filtre temporel de type passe-haut. Le filtrage résultant est passe-bas spatialement et passe-bande temporellement (figure 1b), c'est-à-dire spatio-temporellement séparable comme le montre la figure 1d : la fonction de transfert spatiale de la voie Y reste passe-bas au cours du temps.

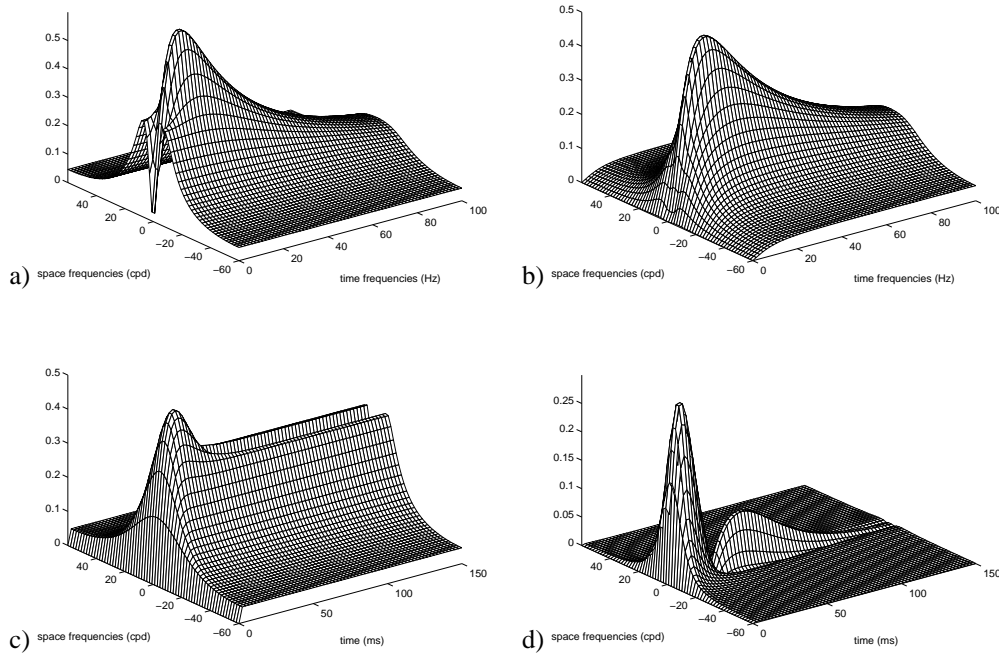


Figure 1: Spectres d'amplitude des fonctions de transfert spatio-temporelles (a,b) et évolutions temporelles de la fonction de transfert spatiale (c,d) pour une stimulation en échelon temporel pour le modèle de la couche plexiforme externe a) et c), et le modèle de la couche plexiforme interne b) et d).

Le modèle de rétine nous suggère donc qu'une perception "Coarse-to-Fine" pourrait résulter de l'un ou l'autre des mécanismes suivants, voire peut-être des deux à la fois :

1. de l'inséparabilité spatio-temporelle présente dans le filtrage passe-bande de la PLE et transmise à la voie X. L'information spatiale de luminance y serait donc multiplexée temporellement : la voie X ne serait pas purement tonique mais serait composée d'un canal tonique, plus lent, pour les hautes fréquences spatiales ("Fine channel") et d'une canal phasique, plus rapide, pour les basses fréquences ("Coarse channel").
2. l'extraction précoce par la PLI du signal phasique "coarse" transmis à la voie Y pour laquelle la conduction axonale est plus rapide que celle de la voie X, permettant ainsi une arrivée de l'information "coarse" avant celle de type "fine".

3 Extension du modèle à la perception de la couleur

Le modèle achromatique de la rétine fournit les canaux de types X et Y et montre un traitement temporel "Coarse-to-Fine" de l'information spatiale. Ce modèle achromatique peut être étendu à la vision des couleurs en considérant les sensibilités spectrales des cônes afin de modéliser les propriétés fonctionnelles de la voie parvo-cellulaire (P), la voie magno-cellulaire (M) étant quant à elle identique à la voie Y. Comme escompté d'après les caractéristiques des cellules à opposition de couleur que l'on trouve au niveau du corps genouillé latéral (CGL) (e.g., R^+ / G^- , B^+ / Y^-) [4], les propriétés chromatiques du stimulus modulent le champ récepteur des cellules bipolaires [5] et conduit à un multiplexage spatial des informations de luminance et de chrominance (figure 2a) : à une stimulation statique, les cellules P montrent un comportement spatial passe-bande aux variations de luminance et un comportement passe-bas aux changements de couleur isoluminants. En outre, l'inséparabilité spatio-temporelle, présente pour la luminance dans le modèle de la voie parvo-cellulaire, ajoute un multiplexage

temporel de l'information de luminance : le filtrage spatial de la luminance n'est plus que passe-bas pour les hautes fréquences temporelles. La figure 2b résume ce multiplexage des informations de luminance et de chrominance dans le domaine des fréquences spatio-temporelles.

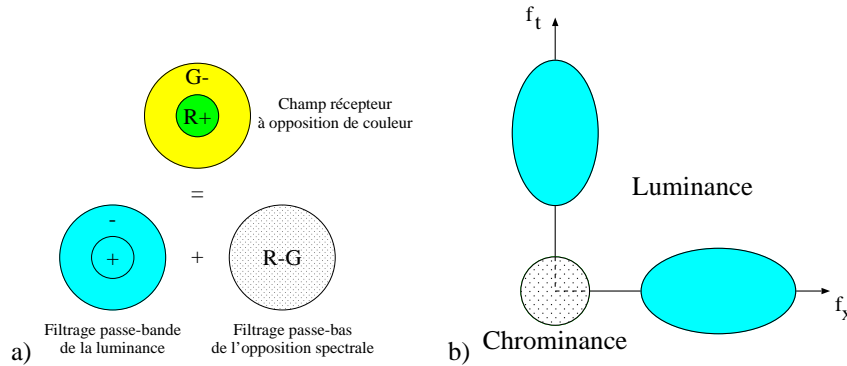


Figure 2: Multiplexage des informations de luminance et de chrominance : a) champ récepteur des cellules à opposition de couleur, b) représentation simplifiée de la fonction de transfert spatiale (f_x) et temporelle (f_t) du modèle chromatique de rétine.

L'hypothèse du démultiplexage de cette entrelacement d'informations suggère des interactions entre les différents canaux rétiniens. En combinant localement de façon additive ou soustractive des canaux parvocellulaires de signes opposés (e.g., G^+/R^- et R^+/G^-), et une inhibition de la part de la voie magno-cellulaire, on prédit l'existence d'au moins trois canaux fonctionnels qui devraient apparaître soit au niveau du CGL, soit au niveau des premières aires corticales visuelles (V1/V2) : 1) un canal phasique (temporellement passe-bande) et spatialement passe-bas, 2) un canal tonique (temporellement passe-bas) et spatialement passe-bande, tous deux dédiés à l'analyse de l'information de luminance mais de façon séparable spatio-temporellement et bien adaptés par exemple à l'analyse respective des ombres en mouvement rapide et des textures fixes ou en mouvement lent; et enfin 3) un canal spatio-temporellement passe-bas à opposition de couleur, peu affecté par la présence d'ombres, plus représentatif des objets et conduisant au phénomène d'induction de couleur.

4 Evidences psychophysiques

L'analyse "Coarse-to-Fine" de l'information de luminance présente dans les modèles rétiniens achromatique et chromatique s'avère être le mode de traitement privilégié par le système humain aussi bien en vision de bas-niveau qu'en vision de haut-niveau :

1. Watt a montré pour des tâches de discrimination simples que le système visuel humain opère effectivement selon une stratégie "Coarse-to-Fine", mais qui reposerait, selon lui, sur la sélection temporelle d'un filtre spatial dans un banc de filtres, et plus précisément sur l'extinction progressive des filtres de plus basses fréquences qui seraient utilisés plus spécifiquement pour des tâches de nature géométrique [6].
2. Afin de différencier le traitement visuel basé sur les basses *versus* hautes fréquences spatiales, Schyns et Oliva ont quant à eux proposé un protocole expérimental basé sur l'utilisation d'images dites "hybrides", composées de basses fréquences spatiales d'une scène réelle additionnées aux hautes fréquences spatiales d'une autre scène [7]. Ils ont montré à l'aide de ce protocole que pour une stimulation visuelle d'une durée de 30 ms, tout le spectre fréquentiel est intégré par le système visuel et utilisé pour identifier une scène mais aussi qu'une stratégie de type "Coarse-to-Fine", qui ne serait pas "câblée" physiquement, est privilégiée lorsque la tâche à accomplir requiert peu de ressources attentionnelles.

Le modèle de rétine n'est pas en désaccord avec l'hypothèse proposée par Watt reposant sur la désactivation au cours du temps des filtres de basses fréquences, mais l'analyse du modèle suggère plutôt qu'elle serait due à la nature spatio-temporelle des signaux atteignant ce banc de filtres, et à l'existence d'un gain supérieur pour les basses fréquences suffisant pour privilégier l'utilisation précoce de l'information "coarse" (lorsque celle-ci est pertinente vis-à-vis de la tâche à accomplir) devant celle de type "fine". Le modèle est aussi en accord avec les résultats de Schyns et Oliva en condition pré-attentive. Pour prendre en compte la nature non figée du système visuel, un troisième mécanisme nécessiterait l'inhibition de l'information "coarse" guidée par l'attention afin de permettre une sélection de l'information selon une séquence "Fine-to-Coarse", par exemple par l'intermédiaire d'une interaction d'origine attentionnelle entre les voies X et Y.

Pour le modèle chromatique, l'hypothèse du démultiplexage spatio-temporel des informations de luminance et de chrominance est en accord avec l'organisation fonctionnelle du cortex visuel chez le primate [8], c'est-à-dire comme un ensemble de voies de traitement parallèles pour les aspects spatiaux et temporels de la luminance et de la couleur. Afin de vérifier certaines hypothèses suggérées par ce modèle rétino-cortical, nous avons proposé un protocole expérimental basé sur une extension chromatique du paradigme des images "hybrides" qui permettrait de dissocier le traitement des voies magno- et parvo-cellulaires. Un logiciel produisant ces stimuli combinant des versions isoluminante/achromatique et basses/hautes fréquences de scènes différentes a été développé pour être exploité à des fins expérimentales à la fois sur le modèle de rétine et des sujets humains.

Dans le cadre de la reconnaissance visuelle, l'existence d'une telle stratégie "Coarse-to-Fine" implique qu'à la présentation d'un stimulus statique, les basses fréquences spatiales seraient d'abord perçues permettant ainsi une perception basée dans un premier temps sur la forme globale des objets et leurs relations spatiales (i.e. la scène ou le contexte), et qu'ensuite ce serait ses hautes fréquences spatiales qui seraient perçues favorisant ainsi dans un deuxième temps une perception basée sur les détails de la scène (i.e. les objets). L'existence d'une pré-activation par le contexte des concepts liés aux objets peut déjà être postulée, et motivée l'existence de mécanismes descendants ("top-down") pour la reconnaissance des objets. Le rôle de la couleur dans ce schéma reste encore à éclaircir.

5 Conclusion

La perception "Coarse-to-Fine" de l'information de luminance semble être une propriété intrinsèque du système visuel puisque mise en évidence aussi bien à bas niveau qu'à haut niveau [6, 7]. Elle pourrait émerger de la nature inséparable du traitement spatio-temporel de la rétine. L'extension du modèle à la couleur suggère aussi que la dynamique des voies parvo- et magno-cellulaires a une incidence sur le traitement cortical de la luminance et la chrominance. Ce codage par multiplexage temporel de l'information spatiale et spectrale nous incite à penser en outre que le contenu informationnel d'un potentiel d'action transmis le long du nerf optique devrait dépendre de sa position au sein du train d'impulsions émis en réponse à une stimulation lumineuse. Finalement, cette dynamique du traitement visuel pourrait être à l'origine de son efficacité : plutôt que fonctionner de façon strictement parallèle ou séquentielle, le système visuel exploiterait ces deux modes afin d'optimiser son traitement. Des travaux récents vont dans ce sens, et suggèrent même que le multiplexage pourrait fournir une solution au groupement perceptif alternative à celle basée sur des oscillations synchrones [9]. Cette approche dans la modélisation du système visuel permettrait donc de réduire le gouffre qui sépare encore la neurophysiologie de la psychologie cognitive, et de proposer des hypothèses testables expérimentalement.

Références

- [1] **D.H. Kelly & C.A. Burbeck.** Critical problems in spatial vision. *CRC Critical Reviews in Biomedical Engineering*, 10:125–177, 1984.
- [2] **B.G. Breitmeyer.** Parallel processing in human vision: History, review, and critique. In Brannan J. (ed) *Applications of parallel processing in vision*, pp. 37–78. Elsevier Science Publishers, 1992.
- [3] **W.H.A. Beaudot.** *Le traitement neuronal de l'information dans la rétine des vertébrés : un creuset d'idées pour la vision artificielle.* Thèse de Doctorat en Informatique, Institut National Polytechnique de Grenoble, France, Décembre 1994. (Accessible à <ftp://tirf.inpg.fr/pub/HTML/beaudot/thesis.html>)
- [4] **P. Gouras.** Colour coding in the primate retinogeniculate system. In Ottoson, D. & Zedi, S. (Eds) *Central and peripheral mechanisms of colour vision*, pp. 183–197. London: Macmillan, 1985.
- [5] **R.L. De Valois, D.M. Snodderly, E.W. Yund Jr & N. Hepler.** Responses of macaque lateral geniculate cells to luminance and color figures. *Sensory Processes*, 1:244–259, 1977.
- [6] **R.J. Watt.** Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *J. Opt. Soc. Am. A*, 4(10):2006–2021, 1987.
- [7] **P.G. Schyns & A. Oliva.** From blobs to boundary edges: Evidence for time and spatial scale dependant scene recognition. *Psychological Science*, 5(4):195–200, 1994.
- [8] **E.A. DeYoe, D.J. Felleman, D.C. Van Essen & E. McClendon.** Multiple processing streams in occipitotemporal visual cortex. *Nature*, 371:151–154, 1994.

- [9] **J.W. McClurkin & L.M. Optican.** Primate striate and prestriate cortical neurons during discrimination: I. Simultaneous temporal encoding of information about color and pattern. *J. Neurophysiol.*, 75(1):481–495, 1996.