

## ***APPROCHE CONNEXIONNISTE DE LA PERCEPTION DU MOUVEMENT***

---

**Claudio CASTELLANOS SÁNCHEZ,**

Doctorant en informatique à l'Université Henri Poincaré, Nancy, France  
Boursier mexicain du Conseil National de Science et Technologie (CONACyT, 113287/169076),  
México.

Tél.: + 33 3 83 59 20 56

Courriel : [castella@loria.fr](mailto:castella@loria.fr)

Site InterNet: <http://www.loria.fr/~castella>

**Bernard GIRAU,**

Maître de conférences en informatique à l'Université Nancy 2, Nancy, France

Tél. : + 33 3 83 59 30 92

Courriel : [girau@loria.fr](mailto:girau@loria.fr)

**Frédéric ALEXANDRE,**

Directeur de Recherche INRIA

Tél. : +33 3 83 59 20 53

Courriel : [falex@loria.fr](mailto:falex@loria.fr)

**Adresse professionnelle**

LORIA INRIA-Lorraine, équipe Cortex ★ Campus Scientifique BP 239 ★ 54506 Vandœuvre-lès-Nancy ★ France

**Résumé:** La perception du mouvement visuel modélisé par des modèles connexionnistes fournit différents axes de recherche pour le développement de modèles de perception-action en temps réel appliqués à la perception visuelle dynamique du mouvement. Nous présentons dans cet article les bases d'une approche connexionniste bio-inspirée avec une adaptation particulière des filtres de Gabor spatio-temporels que nous utilisons en vue d'une approche plus modulaire et fortement localisée pour la perception visuelle du mouvement.

**Summary :** The visual perception of motion modeled by connectionist models offers various areas of research for the development of models of real-time perception-action applied to the dynamic visual perception of motion. We present in this article the bases of a bio-inspired connectionist approach with a particular adaptation of the spatiotemporal Gabor filters that we use within a more modular and strongly localized approach for the visual perception of motion.

**Mots clés :** Perception du mouvement, réseau de neurones bio-inspirés, flot optique, filtres de Gabor.

**Keywords :** Perception of motion, bio-inspired neural network, optical flow, Gabor filters.

# Approche connexionniste de la perception du mouvement

## 1 – INTRODUCTION

Le traitement et l'interprétation de la perception du mouvement visuel a une importance croissante en robotique autonome. Les modèles connexionnistes y ont apporté leur puissance de généralisation et leur robustesse au manque d'information. Leur parallélisme intrinsèque conjugué à un traitement local nous fournit différents axes de recherche pour le développement de modèles de perception-action en temps réel appliqués à la perception visuelle dynamique du mouvement.

Nous proposons un rapide survol de l'état de l'art en neurophysiologie, des algorithmes proposés pour l'obtention du flot optique, et de quelques réalisations matérielles. Nous présentons ensuite les bases d'une approche connexionniste bio-inspirée avec une adaptation particulière des filtres de Gabor spatio-temporels. Finalement, nous envisageons une approche connexionniste plus modulaire et fortement localisée pour la perception visuelle du mouvement.

## 2 – ÉTAT DE L'ART

### 2.1 - Point de vue biologique

La perception du mouvement chez les animaux est indispensable pour survivre aux prédateurs en détectant leur forme 3D et leur profondeur d'une part, et en réalisant des jugements de coïncidence pour estimer la direction et la vitesse de déplacement d'autre part.

On sait actuellement que la détection et l'analyse du mouvement sont obtenues par une cascade d'opérations neuronales [Sekuler, 2002], à savoir: l'enregistrement des signaux de mouvement local dans des régions restreintes du champ visuel et leur intégration pour décrire de façon générale la direction et la vitesse des mouvements.

#### *A - Le parcours des signaux depuis la rétine*

De la rétine jusqu'au cortex cérébral de l'être humain, dix-sept régions différentes interviennent [Sunaert, 1999]. On peut organiser les principales en quatre étapes:

l'acquisition et la compression des signaux lumineux au niveau de la rétine, leur intégration binoculaire et temporelle au niveau du chiasma optique et des corps genouillés latéraux, respectivement; leur analyse corticale au niveau de l'aire visuelle primaire, et leur traitement secondaire dans les aires 20 et 21 du cortex temporal et l'aire 7 du cortex pariétal.

Dans tout ce chemin que parcourent les signaux lumineux on peut détecter deux voies [Rao, 2003]: la voie ventrale ou chemin cortical occipito-temporal (incluant les aires visuelles V1, V2 et V4 et l'aire inféro-temporale IT), qui passe par les cellules parvocellulaires, responsables de la perception des objets et de leur forme, et la voie dorsale ou chemin occipito-pariétal (incluant les aires visuelles V1, V2 et V3, l'aire temporelle moyenne -MT/V5- et l'aire pariétale postérieure PP), qui passe par les cellules magnocellulaires, responsables de la perception du mouvement et de l'espace. Nous nous sommes particulièrement intéressés à la voie dorsale et aux aires spécialisées dans l'analyse du mouvement.

La rétine répond fortement aux mouvements des contours, aux points lumineux et aux grilles, mais ses réponses ne sont sélectives ni à la direction ni à la vitesse de mouvement. Ces dernières caractéristiques sont issues des neurones de l'aire visuelle primaire.

#### *B - L'analyse du mouvement dans le cerveau*

Les zones du cerveau humain spécialisées dans la perception du mouvement sont [Sekuler, 2002]: l'aire visuelle primaire (V1) qui réagit au mouvement vacillant, l'aire temporelle moyenne (MT) et l'aire supérieure temporelle moyenne (MST) qui sont spécialisées dans la détection du mouvement cohérent autant que l'aire occipitale kynétique (KO) l'est pour les contours définis du mouvement, et, finalement, le sulcus supérieur temporel (STS) qui répond aux formes de mouvement dépendantes du mouvement biologique. Nous nous sommes notamment intéressés aux deux premières aires.

### **a - La détection de la direction et de l'orientation**

V1 et MT semble extraire les formes du flot global [Bremmer, 1999] et MST les formes particulières (radiales, rotationnelles et leurs combinaisons) [Lappe, 1996]. Et, comme MST reçoit une entrée vestibulaire depuis MT, il semble être impliqué dans la perception du mouvement propre.

La première analyse corticale est effectuée dans V1 en assurant la sensibilité aux contrastes lumineux par des champs récepteurs allongés. Organisée en six couches de bas en haut, et en bandes de gauche à droite [Alexandre, 1990], les neurones de cette aire sont organisés en colonnes d'orientation (qui tourne progressivement de 10° d'une colonne à l'autre) et de dominance oculaire. Ces neurones émettent principalement leurs prolongements dans la direction verticale et sont accordés à une direction préférentielle de mouvement [Hubel, 1962] effectuant une analyse locale de l'énergie du mouvement qualifiée de filtre. Ces « filtres » enregistrent la présence des caractéristiques fondamentales des objets bougeant autour de leurs champs récepteurs.

Ces cellules sélectives à l'orientation spatiale peuvent être modélisées par des filtres spatio-temporels [Adelson, 1985] et leurs champs récepteurs comme un produit des interactions inhibitrices et excitatrices dans l'espace et dans le temps.

Toutes ces cellules sont sélectives à une direction préférentielle, à une fréquence spatiale, peu sensibles à la composition de la longueur d'onde du stimulus du mouvement et limitées à une extraction locale du mouvement. Or, ces cellules ne voient qu'une partie de la scène, leur réponse est ambiguë car pour chaque direction il existe une famille de cellules sélectives répondant aux mêmes changements de luminosité spatio-temporelles qu'une cellule locale et qui génère la même réponse.

Pour autant, le mouvement local de l'image rétinienne est extrait par les neurones de V1 ayant un champ récepteur semblable à une petite fenêtre spatialement délimitée où ils peuvent enregistrer la présence de mouvement dans une direction déterminée. Ce traitement fortement localisé et récurrent est notre source d'inspiration vers un modèle bio-inspiré pour la détection et l'estimation du mouvement.

### **b - La détection de la direction et de la vitesse**

La perception visuelle du mouvement n'est pas complètement déterminée par les réponses locales dans le champ récepteur des neurones, elles sont de plus traitées pour obtenir la vitesse après les avoir collectées et combinées depuis V1 et en les regroupant dans MT. C'est cette combinaison des signaux qui résout l'ambiguïté des réponses individuelles des neurones.

MT intègre les signaux de mouvement issus de V1 sur des régions plus étendues de l'espace visuel, et construit une représentation rétinotopiquement organisée de l'espace visuel. Presque tous ses neurones sont sélectifs à la direction et à la vitesse d'un stimulus de mouvement qu'ils calculent de façon locale, explicite et non-ambiguë [Perrone, 2001]. Les champs récepteurs de ces neurones sont plus grands que ceux de V1 et augmentent avec l'excentricité. En outre, un neurone de MT répondra mieux si les signaux du mouvement sont renforcés par la luminosité, la couleur ou la texture. Par conséquent, ces neurones ont une certaine invariance car ils enregistrent l'information du mouvement par eux-mêmes. Ces neurones sont accordés à une direction préférentielle et activent un processus compétitif pour aboutir à un vote non-ambigu d'une vitesse particulière. Ce processus compétitif est dépendant de la fréquence spatiale, de la forme, du contraste et de la disparité des objets [Bradley, 1995].

La différence des réponses des neurones de MT et de ceux de V1 est que certains neurones de MT reçoivent en entrée les réponses des neurones de V1 avec différentes directions préférentielles et, comme une grande partie des neurones de MT répondent à une direction de mouvement donnée indépendamment de l'orientation du stimulus, cette aire élimine l'ambiguïté de la direction du mouvement. Ces neurones fonctionnent comme si leurs champs récepteurs étaient une combinaison de multiples sorties des neurones de V1 [Movshon, 1996].

Finalement, les neurones de MST reçoivent les réponses des neurones de MT. Ils sont sélectifs au mouvement complexe comme l'expansion et la contraction [Britten, 1998] et aussi sensibles à l'information non-rétinienne comme le mouvement des yeux.

D'autre part, la détection du contraste est suffisante pour l'identification de la direction du mouvement, de telle façon que les mécanismes visuels extrayant le mouvement sont étiquetés par la direction [Derrington, 1993].

## 2.2 - Point de vue mathématique

La détection et l'estimation du mouvement sont des analyses de type bas niveau. Par contre, sa segmentation et son interprétation sont des analyses de type haut niveau. Notre approche reste dans le premier type.

### A - Le flot optique

Le flot optique est « la distribution 2D de vitesses apparentes pouvant être associées à la variation de la luminosité des formes de l'image » [Horn, 1980]. Il cherche alors à apparier les pixels de même intensité.

L'estimation du flot optique a été proposée comme un pas de pré-traitement dans plusieurs algorithmes de vision de haut niveau [Galvin, 1999].

En général, pour obtenir le flot optique on récupère, d'abord, toute l'information localement disponible de la vitesse (en utilisant un petit voisinage spatio-temporel) pour tout le champ visuel en réalisant un pré-filtrage (passe-bas ou passe-bande) en vue d'extraire la structure d'intérêt du signal et de rehausser la proportion signal-bruit. Ensuite, on extrait des mesures (les dérivées spatio-temporelles ou les surfaces de corrélation, par exemple) et on les intègre soit par régularisation, soit par mise en correspondance, soit par le calcul des moindres carrés. Finalement, on propage la vitesse des régions disposant d'une information complète vers celles n'ayant qu'une information partielle ou nulle de façon à récupérer le champ de flot optique correct [Barron, 1994].

En accord avec la conservation de l'intensité lumineuse et son équation [Horn, 1980] :

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (1)$$

où  $I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t)$  est la luminosité sur le point  $(x, y)$  du plan de l'image et en supposant qu'elle reste constante au temps  $t$ , on peut se rendre compte des principes fondamentaux de quelques approches d'estimation du flot optique [Galvin, 1999]: par mise en correspondance, basées sur le gradient,

stochastiques et spatio-temporelles (fréquentielles). Nous nous sommes surtout intéressés à cette dernière approche.

### B - Des approches pour estimer le flot optique

Dans les approches de mise en correspondance, on suppose la conservation locale de la distribution d'intensité, laquelle est filtrée, tout d'abord, par un filtre passe-bande, puis traitée selon un schéma hiérarchique. Pour cela, on fait un appariement entre deux images consécutives d'une séquence donnée. Cette méthode définit un déplacement (lequel est une approximation de la vitesse) comme le meilleur ajustement parmi les régions voisines variant dans le temps [Anandan, 1989].

Dans les approches basées sur le gradient, on suppose, tout d'abord, la conservation de l'intensité de l'image dans un petit intervalle de temps et, en conséquence, on a une équation linéaire avec deux inconnues. Ensuite, on utilise le résultat d'une fonction de l'intensité de l'image plutôt que l'intensité de l'image elle-même. Finalement, on calcule la vitesse des dérivées spatio-temporelles des intensités de l'image en supposant que le domaine de l'image est continu (différentiable) dans l'espace et dans le temps [Nagel, 1987].

Les approches stochastiques (de relaxation) sont basées sur une contrainte concernant les discontinuités du mouvement. Elles manipulent les occlusions en modélisant explicitement les intensités discontinues par des processus stochastiques [Heitz, 1993] ou en utilisant les filtres de Kalman [Black, 1992].

Dans les approches fréquentielles, les fréquences spatio-temporelles sont mises en relation avec la vitesse du stimulus de mouvement, et le flot optique devient l'identification d'un plan d'énergie dans l'espace de la fréquence spatio-temporelle. On a trouvé que les mécanismes sensibles au mouvement basés sur l'énergie orientée spatiale et temporellement dans l'espace de Fourier pouvaient estimer le mouvement dans les endroits où les autres approches échouent. Ainsi, la détection du mouvement dans l'image revient ici à extraire l'orientation spatio-temporelle [Adelson, 1985 ; Heeger, 1987].

On connaît deux sous-approches, fondées sur la phase ou sur l'énergie. Dans le premier cas, on définit les composantes de la vitesse en termes de mouvement instantané au niveau de

la phase des contours de la sortie des filtres passe-bande de Gabor accordés à la vitesse [Fleet, 1990]. Dans le deuxième cas, ces filtres orientés spatio-temporellement peuvent être construits sur des filtres séparés, i.e. les filtrages spatiaux et temporels peuvent être appliqués séquentiellement à l'image d'entrée [Adelson, 1985]. Pour l'obtention de la vitesse, on aurait un schéma de codification de la population où la vitesse est exprimée comme le rapport des réponses des filtres d'énergie accordés aux différentes vitesses. En supposant une linéarité, les réponses de chaque filtre changent avec les différents contrastes mais leur rapport reste constant.

Cette approche est donc équivalente à l'approche de mise en correspondance [Adelson, 1986 ; Simoncelli, 1993]. L'utilisation des filtres spatio-temporels orientés qui sont sensibles à des stimuli en mouvement (comme les filtres de Gabor) a pour origine la recherche concernant la vision des mammifères, qui avait mis en évidence la présence de cellules simples dans l'aire corticale V1 [McLean-Palmer, 1985] et de cellules complexes dans l'aire MT [Newsome, 1983] se comportant comme des filtres passe-bande spatio-temporels. La réponse de ces cellules dépend de la vitesse d'un objet en mouvement ainsi que de sa forme et de ses caractéristiques fréquentielles [Hammond, 1981].

Parmi les avantages des méthodes énergétiques on pourrait énoncer [Spinei, 1998]: un lissage intrinsèque qui réduit les effets du problème d'ouverture, une meilleure robustesse au bruit et des résultats de qualité pour des séquences d'images naturelles. Par contre, un coût élevé en calcul et le manque de précision à proximité des frontières du mouvement en sont ses défauts.

Les principes de l'approche spatio-temporelle consistent à localiser le plan d'énergie non nul ou le plan de vitesse correspondant à un mouvement de translation 2D apparent sur un support spatial délimité par le biais de filtres spatio-temporels orientés.

### 2.3 – Modèles bio-inspirés

Étant donné que l'aire V1 est spécialisée dans la détection et l'estimation du mouvement visuel, et que MT et MST sont spécialisés dans sa segmentation et son interprétation, plusieurs chercheurs ont essayé de s'inspirer de leur

fonctionnement pour créer des modèles bio-inspirés.

Simoncelli et Heeger [Simoncelli, 1998] ont proposé un modèle séquentiel pondéré, rectifié et normalisé des aires V1 en avant et MT en aval.

Zemel et Sejnowski [Zemel, 1998] ont proposé un modèle bidirectionnel de MSTd segmentant les champs de mouvement des différents objets mobiles selon les formes cohérentes du mouvement de chaque objet. Tout d'abord, ils ont utilisé un pré-traitement tiré de Wang et al. [Wang, 1989], puis la modélisation de MT pour l'obtention du flot optique par l'algorithme de Nagel [Nagel, 1987], et finalement ils utilisent au niveau de MST une couche cachée qui est sélective aux formes.

Pack et al. [Pack, 2000] ont proposé un modèle qui distribue les cellules de MT et de MST en unités inhibées ou excitées avec une structure auto-récurrente pour MST en aval. Ils ont modélisé MT et MST en couples: MT → MSTd (calcul de la direction supposée, vitesse perçue, excitation aux petits mouvements et inhibition aux grands mouvements) et MT+ → MSTv (calcul du mouvement du fond par compétition, excitation et inhibition contraires au premier couple).

### 2.4 – Solutions matérielles

D'autre part, quelques chercheurs ont proposé des solutions matérielles comme Tsao et Chen [Tsao, 1994] qui ont proposé un schéma neuronal pour le calcul du flot optique fondé sur les filtres de Gabor et une méthode généralisée du gradient. Leur architecture VLSI est composée de trois couches tendant à estimer le mouvement comme un processus rétinien (et de LGN) effectué par un ensemble de filtres de Gabor lissant et rejetant le bruit de l'information mais conservant les composantes de la fréquence orientée spatialement dans les différentes directions et dans un canal déterminé (comme on peut voir dans les systèmes biologiques).

Stocker [Stocker, 2001] a proposé un système composé de deux réseaux: un pour la segmentation du mouvement et un autre pour la sélection du mouvement. Le premier cherche les discontinuités du mouvement et restreint le processus d'estimation collective du flot optique aux unités recevant l'information visuelle d'un même objet dans l'espace. Le

deuxième fournit une perception sélective du mouvement visuel en accord avec un mouvement préférentiel donné. Ces réseaux sont analogiques et récurrents. Le premier est de type « winner-take-all » (WTA) et le deuxième est entièrement connecté. Il modélise la perception visuelle du mouvement comme un problème d'optimisation.

Pirim [Pirim, 1998] a valorisé un système qui détermine l'existence et la localisation d'une zone en déplacement ainsi que sa vitesse et sa direction orientée. Le système a trois étapes: la première effectue un traitement temporel, la deuxième fait un traitement spatial suivant l'algorithme temps-échelle de Mallat [Pirim, 1998] et la troisième fait un calcul d'histogrammes pour trouver le maximum et les limites inférieures et supérieures de la zone du mouvement. Les paramètres de résolution spatiale et de la structure de l'image sont le résultat d'une analyse par des ondelettes de Gabor orientées décrites par Daugman [Daugman, 1988].

### 3 – APPROCHE RETENUE

En vue d'une réalisation bio-inspirée, connexionniste, fortement locale et embarquable sur un robot mobile qui suit une cible bougeant, nous proposons dans cet article une adaptation des filtres de Gabor spatio-temporels pour modéliser les cellules magnocellulaires qui peuvent être vues comme des senseurs de mouvement pour l'extraction des gradients spatio-temporels généralisés de l'intensité de l'image et de sa dérivée temporelle [Pollen, 1981].

#### 3.1 - Les fondements biologiques

Il y a deux types de calcul dans les cellules magnocellulaires: la différentiation temporelle et l'intégration spatiale, sachant que « la vision périphérique est presque totalement impliquée dans la détection des dérivées temporelles et une grande partie de la bande passante dans le nerf optique est utilisée par l'information des dérivées temporelles » [Mead, 1989].

Le mouvement visuel est ressenti comme des stimuli visuels de mouvement, notamment des contrastes de luminosité orientés spatialement, extraits par les différents champs récepteurs des cellules magnocellulaires dans le cortex visuel. Les opérations mathématiques réalisées par ces champs récepteurs peuvent être

simplifiées par des intégrations spatiales pondérées des signaux.

En voyant les cellules corticales comme des détecteurs de caractéristiques, quelques modèles permettent un principe d'attente et de comparaison [Adelson, 1985], c'est-à-dire, premièrement, on extrait les caractéristiques visuelles d'une séquence d'images et, deuxièmement, on calcule le mouvement visuel basé sur le déplacement spatial de ces caractéristiques.

Les champs récepteurs de type filtres spatiaux 2D de Gabor proposés par Marcelja [Marcelja, 1980] ont été étendus à des filtres spatio-temporels 3D et appliqués comme des analyseurs spectraux locaux spatio-temporels. Ces champs ont des fonctions pondérées de lissage distinctes des masques simples pour la détection de barres. En outre, ils sont passe-bande et sélectifs à l'orientation mais ils ne sont pas spécifiques aux caractéristiques. En conséquence, ils ne sont ni de bons analyseurs spectraux ni de bons détecteurs de barres ni de bons détecteurs de contours. Ce sont des champs récepteurs passe-bande sélectifs à l'orientation.

#### 3.2 - Le filtre de Gabor adapté

Soit  $I(x,y,t)$  la forme de l'intensité dans l'image variant dans le temps en supposant que chaque point a une luminosité invariante. Soit aussi  $I(x,y,t) = I(x - ut, y - vt)$  où  $(u,v)$  est le vecteur de mouvement d'une petite région de l'image, et où  $I(x,y)$  est la trame d'image de l'échantillonnage au temps  $t=0$ . Grâce à la supposition d'une fréquence d'échantillonnage suffisamment haute pour assurer la localité des détections de mouvement, nous pouvons faire l'hypothèse d'une vitesse locale instantanée constante.

En appliquant alors le filtre de Gabor orienté,  $G_\theta(x,y)$  avec  $0 \leq \theta < 2\pi$ , à  $I(x,y)$  nous obtenons:

$$\left. \frac{d}{dt} \int \int G(\hat{x}-\hat{u}, \hat{y}-\hat{v}) dx dy \right|_{t=0} = \frac{d}{dt} \int \int I(x-ut, y-vt) G(\hat{x}-\hat{u}, \hat{y}-\hat{v}) dx dy \Big|_{t=0} \quad (2)$$

où

$$\hat{x}=(x-\xi)\cos\theta-(y-\eta)\sin\theta, \hat{y}=(x-\xi)\sin\theta+(y-\eta)\cos\theta \quad (3)$$

avec  $(\xi, \eta) \in \Gamma$  un petit voisinage autour de  $(x,y)$ , et

$$\hat{u}=\frac{t-t'}{\tau-1}v_i\cos\theta, \quad \hat{v}=\frac{t-t'}{\tau-1}v_i\sin\theta \quad (4)$$

pour  $\tau$  images consécutives et une vitesse  $v_i$  supposée.

$(\hat{x},\hat{y})$  est l'endroit où le signal de Gabor orienté va être calculé et

$$G(x,y)=\exp\left(\frac{x^2+\gamma^2y^2}{\sigma^2}\right)\cos\left(2\pi\frac{x}{\lambda}+\phi\right) \quad (5)$$

est la fonction de réponse à l'impulsion du filtre de Gabor modélisant la fonction des cellules ganglionnaires magnocellulaires où  $\gamma$  est l'excentricité du champ récepteur et  $\sigma$  sa taille;  $\lambda$  la longueur d'onde et  $\phi$  la phase, tous d'après [Petkov, 1997].

La mesure obtenue par un seul filtre de Gabor ne peut pas déterminer le vecteur du mouvement 2D, il faut mettre un ensemble de filtres de Gabor ne différant qu'en orientation. Ils sont alors placés dans un vecteur appelé vecteur de senseurs de mouvement où chaque orientation est un senseur de mouvement.

De plus,

$$\frac{d}{dt}\left[\iint I(x-ut,y-vt)G(\hat{x}-\hat{u},\hat{y}-\hat{v})dxdy\right]_{t=0} = \frac{d}{dt}\left[\iint I(x,y)G(\hat{x}-\hat{u}-ut,\hat{y}-\hat{v}-vt)dxdy\right]_{t=0} \\ = -\left[\iint I(x,y)(\nabla G(\hat{x}-\hat{u},\hat{y}-\hat{v})\vec{v})dxdy\right]_{t=0} = -\vec{v}\cdot\left[\iint I(x,y)\nabla G(\hat{x}-\hat{u},\hat{y}-\hat{v})dxdy\right] \quad (6)$$

où  $\vec{v}=(u,v)$  est la vitesse 2D de la région  $\Gamma$ .

En notant

$$\vec{\Omega}_x(\hat{x},\hat{y})=-\vec{v}\cdot\iint I(x,y)\nabla G(\hat{x}-\hat{u},\hat{y}-\hat{v})dxdy, \\ \vec{\Omega}_x(\hat{x},\hat{y})=\iint I(x,y)\frac{\partial G_x}{\partial x}G(\hat{x}-\hat{u},\hat{y}-\hat{v})dxdy, \\ \vec{\Omega}_y(\hat{x},\hat{y})=\iint I(x,y)\frac{\partial G_y}{\partial y}G(\hat{x}-\hat{u},\hat{y}-\hat{v})dxdy \quad (7)$$

on peut récrire la dernière équation comme:

$$\vec{\Omega}_x(\hat{x},\hat{y})+u\vec{\Omega}_x(\hat{x},\hat{y})+v\vec{\Omega}_y(\hat{x},\hat{y})=0 \quad (8)$$

pour une vitesse  $v_i$  à un instant  $t$ . Voilà l'adaptation fréquentielle fondée sur l'énergie de l'équation 1.

Daugman [Daugman, 1989] avait suggéré que la fonction sélective de la direction des neurones dans MT est représentée de façon binaire par un réseau neuronal WTA. Dans ce cas, les neurones de MT sont bizarrement accordés à différentes vitesses et directions.

Nous appliquons simplement un WTA aux différentes orientations pour chaque vitesse supposée.

#### 4 - RÉALISATIONS

Nous avons testé notre version adaptée des filtres de Gabor spatio-temporels sur des séquences d'images naturelles et synthétiques. La figure 1 montre les filtres d'amplitude maximale obtenus dans chacune des 8 directions (en fixant les paramètres libres des équations 2, 3, 4 et 5) ainsi que leur somme vectorielle, dans le cas de séquences d'images synthétiques où une boule s'approche ou s'éloigne dans un couloir rectiligne.

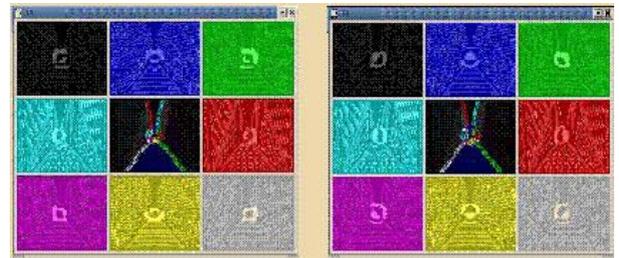


Fig. 1 - Perception du mouvement par filtres de Gabor adaptés. L'image du centre est la somme vectorielle des huit directions avec leur couleur respective et  $t=3, v \in \{-2,-1,0,1,2\}$  dans une image 2D de  $5 \times 5$  et une taille de pas de  $1 \times 1$ . Gauche : Boule s'éloignant. Droite : Boule s'approchant.

Nos travaux actuels consistent à inclure cette approche dans une architecture connexionniste utilisant un ensemble d'interactions latérales excitatrices et inhibitrices inspirées des colonnes corticales issues de V1, avec un schéma de connectivité fortement local et un rayon d'influence qui dépend de la réponse locale obtenue via l'équation 2. Notre but est d'arriver à proposer un modèle connexionniste complètement modulaire et fortement localisé en nous inspirant des résultats neurophysiologiques tirés des études du chemin des cellules magnocellulaires.

#### 5 - CONCLUSION

La perception du mouvement est vitale chez l'être vivant. D'après les résultats neurophysiologiques de la vision chez les primates, nous avons exposé que: la rétine joue un rôle de filtrage des signaux lumineux, les corps genouillés latéraux font une intégration binoculaire, V1 une analyse spatiale, MT et MST une analyse temporelle. La forte réduction de l'information depuis la rétine et la grande quantité des interactions latérales dans

les aires supérieures font émerger la dynamique de notre perception du monde. Ces différents niveaux de traitement et leurs rétroactions respectives illustrent les coopérations existant entre colonnes rétinotopiquement proches, et que nous exploitons dans nos architectures connexionnistes. À ce niveau, nous poursuivons nos travaux afin d'intégrer les différentes étapes de la voie dorsale que parcourent les signaux lumineux avec un point de vue modulaire, purement connexionniste et fortement localisé.

## BIBLIOGRAPHIE

- [Adelson, 1985] Adelson, E. H. and Bergen, J., Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(7):284-299 (1985).
- [Adelson, 1986] Adelson, E. H. and Bergen, J., The extraction of spatiotemporal energy in human and machine vision. In *Proc. IEEE Workshop on Visual Motion*, pages 151-156, Charleston (1986).
- [Alexandre, 1990] Alexandre, F., Une modélisation fonctionnelle du cortex: la Colonne Cortical. Aspects visuels et moteurs. PhD thesis, Henri Poincaré, Nancy, France (1990).
- [Anandan, 1989] Anandan, P., A computational framework and an algorithm for the measurement of visual motion. *International Journal Computer Vision*, 2:283-310 (1989).
- [Barron, 1994] Barron, J. J., Fleet, D. J., and Beauchemin, S. S., Performance of optical flow techniques. In *Proceedings Conference on Computer Vision and Pattern Recognition*, pp. 236-242, Champaigne (1994).
- [Black, 1992] Black, M. J., Robust incremental optical flow. PhD thesis, Yale University, Yale (1992).
- [Bradley, 1995] Bradley, D., Quian, N., and Andersen, R., Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, 373:609-611 (1995).
- [Bremmer, 1999] Bremmer, F. and Lappe, M., The use of optical velocities for distance discrimination and reproduction during visually simulated self-motion. *Experimental Brain Research*, 127(1):33-42 (1999).
- [Britten, 1998] Britten, K. H. and van Wezel, R. J., Electrical micro-stimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience*, 1:59-63 (1998).
- [Daugman, 1988] Daugman, J. G., Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169-1179 (1988).
- [Daugman, 1989] Daugman, J. G., Networks for images analysis: Motion and texture. In *IJCNN* (1989).
- [Derrington, 1993] Derrington, A. M. and Henning, G. B., Detecting and discriminating the direction of motion of luminance and color gratings. *Visual Research*, 33 (1993).
- [Fleet, 1990] Fleet, D. and Jepson, A. D., Computation of component image velocity from local phase information. *International Journal Computer Vision*, 5:77-104 (1990).
- [Galvin, 1999] Galvin, B., McCane, B., Novins, K., Mason, D., and Mills, S., Recovering Motion Fields: An Evaluation of Eight Optical Flow Algorithms. Technical report, Computer Science Department (1999).
- [Hammond, 1981] Hammond, P. and Reck, J., Influence of velocity on directional tuning of complex cells in cat striate cortex for texture motion. *Neuroscience Letters*, 19:309-314 (1981).
- [Heeger, 1987] Heeger, D. J., A model for the extraction of the image flow. *Journal of the Optical Society of America A*, 4(8):1455-1471 (1987).
- [Heitz, 1993] Heitz, F. and Bouthemy, P., Multimodal estimation of discontinuous optical flow using Markov random fields. *IEEE Transactions on Pattern Analysis and*

- Machine Intelligence, 15(12):1217-1232 (1993).
- [Horn and Schunk, 1980] Horn, B. K. P. and Schunk, B. G. Determining optical flow. Technical Report A. I. Memo No. 572, Artificial Intelligence Laboratory (1980).
- [Hubel, 1962] Hubel, D. H. and Wiesel, T. N., Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *Journal Physiology*, 160:106-154 (1962).
- [Lappe, 1996] Lappe, M., Bremmer, F., Pekel, M., Thiele, A., and Hoffmann, K. P., Optical flow processing in monkey STS: A theoretical and experimental approach. *The Journal of Neurosciences*, pages 6265-6285 (1996).
- [Marcelja, 1980] Marcelja, S., Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America A*, 70/11:1297-1300 (1980).
- [McLean-Palmer, 1985] McLean-Palmer, J., Jones, J., and Palmer, L., New degrees of freedom in the structure of simple receptive fields. *Investigative Ophthalmology and Visual Science Supplement*, 26:265 (1985).
- [Mead, 1989] Mead, C. A., editor (1989). *Analog VLSI and neural systems*. Addison-Wesley, Reading, MA.
- [Movshon, 1996] Movshon, J. A. and Newsome, W. T., Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *The Journal of Neurosciences*, 23:7733 (1996).
- [Nagel, 1987] Nagel, H. H., On the estimation of optical flow: Relation between different approaches and some new results. *Artificial Intelligence*, 33:299-324 (1987).
- [Newsome, 1983] Newsome, W. T., Gizzi, M. S., and Movshon, J. A., Spatial and temporal properties of neurons in macaque MT. *Investigative Ophthalmology and Visual Science Supplement*, 24:106 (1983).
- [Pack, 2000] Pack, C., Grossberg, S., and Mingolla, E., A Neural Model of Smooth Pursuit Control and Motion Perception by Cortical Area MST. Technical Report CAS/CNR-TR-99-023, Department of Cognitive and Neural Systems and Center for Adaptive Systems, Boston (2000).
- [Parker, 1998] Parker, A. J. and Newsome, W. T., Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience*, 21:227-277 (1998).
- [Perrone, 2001] Perrone, J. and Thiele, A., Speed skills: measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience*, 4(5):526-532 (2001).
- [Petkov, 1997] Petkov, N. and Kruizinga, P., Computational models of visual neurons specialized in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. *Biological Cybernetics*, 76:83-96 (1997).
- [Pirim, 1998] Pirim, P., Procédé et dispositif fonctionnant en temps réel pour le repérage et la localisation d'une zone en mouvement relatif dans une scène, ainsi que pour la détermination de la vitesse et de la direction du déplacement. Technical Report WO 98/05002, France (1998).
- [Pollen, 1981] Pollen, D. and Ronner, S., Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212:1409-1411 (1981).
- [Rao, 2003] Rao, H., Zhou, T., Zhuo, Y., Fan, S., and Chen, L., Spatiotemporal Activation of the Two Visual Pathways in Form Discrimination and Spatial Location: A Brain Mapping Study. *Human Brain Mapping*, 18:79-89 (2003).
- [Sekuler, 2002] Sekuler, R., Watamaniuk, S. N. J., and Blake, R., Motion Perception. *Steven's Handbook of Experimental Psychology*, 1:121-176 (2002).
- [Simoncelli, 1993] Simoncelli, E. P., Distributed analysis and representation of visual motion. PhD thesis, Massachusetts

Institute of Technology, Cambridge,  
M. A. (1993).

- [Simoncelli, 1998] Simoncelli, E. P. and Heeger, D. J., A Model of Neural Responses in Visual Area MT. *Visual Research*, 38(5):743-761 (1998).
- [Spinei, 1998] Spinei, A., Estimation du mouvement par triades de filtres de Gabor: Application au mouvement d'objets transparents. PhD thesis, Laboratoire des Images et des Signaux (1998).
- [Stocker, 2001] Stocker, A. A., Constraint Optimization Networks for Visual Motion Perception - Analysis and Synthesis. PhD thesis, Swiss Federal Institute of Technology Zürich, Zürich, Switzerland (2001).
- [Sunaert, 1999] Sunaert, S., van Hecke, P., Marchal, G., and Orban, G. A., Motion-responsive regions of the human brain. *Experimental Brain Research*, 127:355 (1999).
- [Tsao, 1994] Tsao, T.-R. and Chen, V., A neural scheme for optical flow computation based on Gabor filters and generalized gradient method. *Neurocomputing*, 6:305-325 (1994).
- [Wang, 1989] Wang, H. T., Mathur, B., and Koch, C., Computing optical flow in the primate visual system. *Neural Computation*, 1:92-103 (1989).
- [Watson, 1985] Watson, A. B. and Jr., A. J. A., Model of human visual motion sensing. *Journal of the Optical Society of America A*, 2(2):322-341 (1985).
- [Zemel, 1998] Zemel, R. S. and Sejnowski, T. J., A Model for Encoding Multiple Object Motions and Self-Motion in Area MST of Primate Visual Cortex. *The Journal of Neurosciences*, 18(1):531-547 (1998).