

Circuits neuromorphiques pour l'estimation du mouvement

Antonio Torralba et Jeanny Hérault

Laboratoire de Traitement d'Images et Reconnaissance de Formes,
INPG, 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France

RÉSUMÉ

Nous présentons une nouvelle famille de filtres sélectifs en vitesse plus simple que les filtres de Gabor utilisés dans certains algorithmes pour l'estimation du mouvement. Ces filtres sont réalisés par des réseaux analogiques avec connexions latérales asymétriques qui permettent de contrôler tous les paramètres du filtrage. Nous présentons une architecture pour l'estimation du mouvement mieux adaptée à une réalisation sur circuits intégrés que les méthodes basés sur les filtres de Gabor.

ABSTRACT

This paper presents a new family of velocity-tuned filters much simpler than classical Gabor filters used in some motion estimation algorithms. The filters are implemented with analogue networks with asymmetrical lateral connexions which allows a complete control of the filter parameters. We present a simple scheme for motion estimation based on analogue networks more efficient for VLSI realisation than methods based on Gabor filters.

1 Introduction

Ces dernières décennies, de multiples algorithmes pour l'estimation du mouvement à partir d'une séquence d'images ont été proposés [8]. Ces algorithmes peuvent être classés dans trois familles : 1) les méthodes basées sur les gradients spatiaux et temporels. En raison des problèmes associés au calcul des dérivées, ces méthodes sont très sensibles au bruit et au repliement de spectre. En raison du problème d'ouverture, il faut appliquer des techniques de régularisation coûteuses en temps de calcul. 2) Les méthodes de mise en correspondance qui cherchent à trouver les déplacements des points caractéristiques entre deux images successives. Ces techniques sont bien adaptées pour des grands déplacements mais le coût de calcul est très importants. 3) Les méthodes fréquentielles, basées sur les propriétés d'une image, avec mouvement uniforme, dans l'espace fréquentiel. Ces méthodes utilisent des filtres spatio-temporels sélectifs à différentes vitesses et orientations. Les filtres souvent utilisées sont des ondelettes de Gabor (filtres passe-bande spatio-temporels à bande étroite) [1, 6, 11]. Les sorties sont combinées de façon à estimer la vitesse. Ces méthodes sont robustes vis-à-vis du bruit et du repliement de spectre mais là aussi la simulation des filtres demande beaucoup de calculs. L'algorithme que nous présentons est basé sur des filtres de type passe-bas spatio-temporels inspirés des circuits neuromorphiques [5] permettant des réalisations sur des circuits intégrés analogiques pour des applications en temps réel. La caractéristique passe-bas à large bande des filtres que nous proposons, en comparaison avec les filtres de Gabor à bande étroite, permet de réduire le nombre de filtres nécessaires pour estimer le champ de vitesses. Ces filtres peuvent être simulés numériquement avec peu de calculs (4 produits et additions par pixel). Le temps de calcul pour l'obtention du flux optique, pour des images de 256^2 , est d'une seconde par image sur une machine SUN UltraSparc.

2 Filtres directionnels neuromorphiques

Les méthodes fréquentielles ont comme point de départ le fait qu'une image animée d'un mouvement uniforme a toute son énergie contenue dans un plan de Dirac dans l'espace des fréquences spatio-temporelles. La fonction d'intensité lumineuse pour une image en mouvement de translation, sans variation d'éclairement peut s'écrire de la manière suivante : $i(x, y, t) = \tilde{i}(x - v_x t, y - v_y t)$. Où $\tilde{i}(x, y)$ est la fonction d'intensité pour le motif statique, et v_x et v_y sont les composantes de la vitesse selon les axes x et y . Si l'on applique à cette fonction les transformées de Fourier pour la variable temporelle t , et pour les variables spatiales x et y , on trouve $I(f_x, f_y, f_t) = \tilde{I}(f_x, f_y) \delta(f_t + v_x f_x + v_y f_y)$ où $\delta(\cdot)$ est la distribution de Dirac. Ceci nous montre que le spectre du signal est contenu dans le plan de Dirac $f_t + v_x f_x + v_y f_y = 0$ [11]. Ce plan a une orientation qui est fonction de la vitesse et de la direction du mouvement (fig 1.a). L'objectif est d'échantillonner l'espace fréquentiel pour estimer l'orientation du plan de Dirac. La figure 1.b montre la disposition des filtres de Gabor utilisés par Heeger [6]. La figure 1.c montre la disposition de nos filtres à large bande. L'orientation de chaque filtre détermine la vitesse pour laquelle le filtre est le plus sensible.

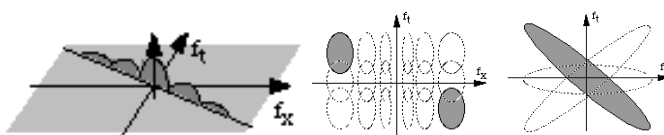


FIG. 1 — a) mouvement dans le domaine fréquentiel, et son traitement par b) une batterie de filtres de Gabor à bande étroite (neuf filtres), c) batterie de filtres à large bande (trois filtres) et allure de la coupe à -3dB.

Les filtres que nous proposons sont une alternative aux filtres de Gabor souvent utilisés dans certains algorithmes pour

l'estimation du mouvement. Ces filtres sont réalisés à l'aide de circuits électroniques neuromorphiques qui modélisent le comportement des réseaux neuronaux du système visuel [5]. La figure 2 montre la structure des filtres directionnels. La tension $i_{x,y}(t)$ est la valeur de l'entrée du noeud x, y . La tension aux bornes du condensateur, $o_{x,y}(t)$, est la tension de sortie. Chaque noeud est connecté à ses quatre voisins par des branches composées d'amplificateur (de gain a_x, a_y, b_x ou b_y) et d'une résistance, R . La résistance r connecte l'entrée au noeud de sortie. La sortie du filtrage est composée par l'ensemble des noeuds $o_{x,y}(t)$. La relation entre l'entrée et la sortie est caractérisée par la fonction de transfert du circuit.

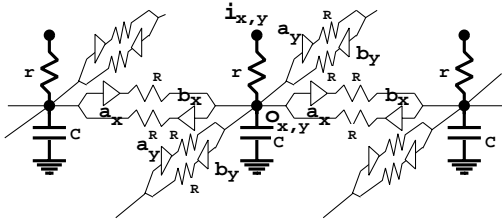


FIG. 2 — Noeud x, y du circuit neuromorphique 2D pour la réalisation d'un filtre sélectif en vitesse.

Si on applique la loi de courants de Kirchoff dans le noeud $o_{x,y}(t)$, on obtient :

$$i_{x,y} = [1 + 4\gamma] o_{x,y} - \gamma [a_x o_{x-1,y} + b_x o_{x+1,y} + a_y o_{x,y-1} + b_y o_{x,y+1}] + \tau \frac{\partial o_{x,y}}{\partial t} \quad (1)$$

où $\gamma = r/R$ et $\tau = rC$. Si l'on applique la transformée de Fourier, on obtient la fonction de transfert du circuit. Pour faciliter l'étude théorique, nous donnons le développement pour les basses fréquences spatiales. Ceci nous permet de simplifier la fonction de transfert [9, 10] :

$$H(f_x, f_y, f_t) \simeq \frac{H_0}{1 + 2 \frac{f_x^2}{\Delta B_x^2} + 2 \frac{f_y^2}{\Delta B_y^2} + j 2 \frac{(f_t + v_{x_0} f_x + v_{y_0} f_y)}{\Delta B_t}} \quad (2)$$

Les paramètres du filtre sont :

$$H_0 = 1 / [1 + \gamma [4 - a_x - b_x - a_y - b_y]] \quad (3)$$

$$\Delta B_x^2 = 1 / [H_0 \pi^2 \gamma (a_x + b_x)] \quad (4)$$

$$\Delta B_y^2 = 1 / [H_0 \pi^2 \gamma (a_y + b_y)] \quad (5)$$

$$\Delta B_t = 1 / [H_0 \pi \tau] \quad (6)$$

$$v_{x_0} = v_0 \cos(\theta) = \frac{\gamma}{\tau} (a_x - b_x) \quad (7)$$

$$v_{y_0} = v_0 \sin(\theta) = \frac{\gamma}{\tau} (a_y - b_y) \quad (8)$$

Nous prenons $a_x + b_x = a_y + b_y$ de façon à avoir $\Delta B_x = \Delta B_y = \Delta B$. L'étude de stabilité du circuit montre qu'il faut vérifier $a_x + b_x + a_y + b_y < 4$ [10]. H_0 est le gain du filtre pour les basses fréquences spatiales. ΔB est la largeur

de bande spatiale maximale du filtre. ΔB_t est la largeur de bande temporelle pour les basses fréquences spatiales. v_{x_0} et v_{y_0} sont les deux composantes du vecteur vitesse pour lesquelles l'énergie de la sortie du filtre est maximale. En fait, pour une image avec cette vitesse, l'énergie en entrée est contenue dans le plan $f_t + v_{x_0} f_x + v_{y_0} f_y = 0$, ceci rend minimal le dénominateur de l'expression 2. L'énergie en sortie du filtre est calculée par intégration spatiale du carré du signal $o_{x,y}$. Si l'entrée est une image avec une vitesse v_x, v_y , sa transformée de Fourier est $I(f_x, f_y, f_t) = \tilde{I}(f_x, f_y) \delta(f_t + v_x f_x + v_y f_y)$. La sortie peut être calculée comme ceci : $O(f_x, f_y, f_t) = I(f_x, f_y, f_t) H(f_x, f_y, f_t) = \tilde{I}(f_x, f_y) H(f_x, f_y, f_t) \delta(f_t + v_x f_x + v_y f_y)$

$= \tilde{I}(f_x, f_y) H(f_x, f_y, -v_x f_x - v_y f_y) \delta(f_t + v_x f_x + v_y f_y)$. La transformée inverse est : $o(x, y, t) = \tilde{o}(x - v_x t, y - v_y t)$, où $\tilde{o}(x, y)$ est la transformée inverse de $\tilde{O}(f_x, f_y) = \tilde{I}(f_x, f_y) H(f_x, f_y, -v_x f_x - v_y f_y)$. L'énergie de la sortie peut être calculée en considérant que l'entrée est un signal à spectre blanc, $\tilde{I}(f_x, f_y) = K$, où K est une constante. Dans ce cas, l'énergie en sortie peut être approximée par l'expression :

$$E = \iint w(x, y) |o(x, y, t)|^2 dx dy \simeq \frac{K H_0^2 \pi \Delta B^2 / \sqrt{2}}{\sqrt{1 + 1.2 \frac{\Delta v_x^2 + \Delta v_y^2}{\Delta v_0^2}}} \quad (9)$$

où $w(x, y)$ est la fenêtre d'intégration pour avoir une estimation locale de l'énergie. La fonction $w(x, y)$ vérifie : $\iint w(x, y) dx dy = 1$. $\Delta v_x = v_{x_0} - v_x$ et $\Delta v_y = v_{y_0} - v_y$ sont les différences entre la vitesse de l'image est la vitesse optimale du filtre. Le paramètre $\Delta v_0 = 2\Delta B_t / \Delta B$ est la sélectivité en vitesse. La principale différence entre le filtre que nous proposons et le filtre proposé par Delbrück [4] est que le circuit présenté ici permet un contrôle complet des paramètres du filtre $\Delta B, v_{x_0}, v_{y_0}$ et Δv , en fonction des composantes du circuit.

Supposons maintenant que l'image d'entrée présente une variation uniquement dans une direction de l'espace. Ceci peut être modélisé par une fonction d'intensité de la forme : $i(x, y) = g(\cos(\beta)x + \sin(\beta)y)$. Dans ce cas là, il sera uniquement perçu la composante du mouvement dans la direction de variation de l'image. Ceci est un cas extrême du problème d'ouverture. Si l'on suppose que dans la direction de variation le spectre est blanc, l'énergie de la sortie du filtre directionnel pour une image avec une vitesse $(v_x, v_y) = v(\cos(\alpha), \sin(\alpha))$ est :

$$E = \frac{K H_0^2 \pi \Delta B / 2}{\sqrt{1 + \frac{(v_0 \cos(\theta - \beta) - v_n)^2}{\Delta v_0^2}}} \quad (10)$$

avec $v_n = v(\cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)) = v\cos(\beta - \alpha)$. K dépend de l'énergie de l'entrée. L'énergie est fonction uniquement de v_n qui est la vitesse dans la direction de variation de l'image.

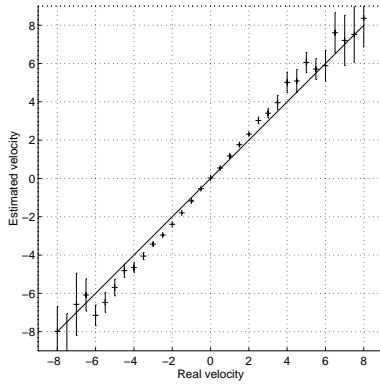


FIG. 3 — vitesse réelle et vitesse estimée pour un signal de spectre blanc avec bruit aditif. les vitesses sont en pixels par seconde. Les filtres sont sélectifs pour les vitesses $v_o = -3$, $v_o = 0$ et $v_o = 3$ pixels/seconde. Les lignes verticales représentent l'écart-type de l'estimation.

3 Estimation du mouvement

L'énergie de la sortie est fonction de la vitesse de l'image d'entrée. On peut combiner la sortie de plusieurs filtres pour estimer la vitesse du mouvement. Si l'on utilise trois filtres sélectifs pour les vitesses $v_o \angle \theta$, $v_o \angle \theta + \pi$ et $v_o = 0$. C'est à dire, deux filtres sélectifs pour les vitesses de module v_o mais dans directions opposées θ et $\theta + \pi$, et un filtre pour la vitesse nulle. Avec l'expression 9, on peut estimer la vitesse de l'entrée en combinant les énergies de ces trois filtres de la façon suivante :

$$\tilde{v}_\theta = \frac{v_o}{2} \frac{E_{v_o \angle \theta}^2 - E_{v_o \angle \theta + \pi}^2}{E_{v_o \angle \theta}^2 + E_{v_o \angle \theta + \pi}^2 - 2E_{v_o \angle \theta} E_{v_o \angle \theta + \pi} / E_{v_o = 0}^2} \quad (11)$$

où \tilde{v}_θ est la projection de la vitesse d'entrée sur la direction θ , $\tilde{v}_\theta = v \cos(\theta - \alpha)$. Les composantes v_x et v_y sont estimées avec cinq filtres ; trois pour la direction $\theta = 0$ et trois pour la direction $\theta = \pi/2$ (le filtre $v_o = 0$ est commun aux deux groupes). Cette expression est valide pour toutes les vitesses d'entrée, cependant, l'estimation est fiable uniquement pour des vitesses dans l'intervalle $[-v_o, v_o]$. La figure 3 montre l'estimation d'une composante de la vitesse et son écart-type pour différentes vitesses d'entrée.

Comment nous l'avons vu auparavant, dans le cas du problème d'ouverture, on peut estimer uniquement la composante de la vitesse dans la direction de variation de l'image. Si l'on prend l'expression de l'énergie donnée par 10 et si l'on combine les sorties de trois filtres comme précédemment, on obtient : $\tilde{v}_\theta = v_n / \cos(\theta - \beta)$. L'estimation donne la valeur de la projection de la vitesse v_n (vitesse apparente de l'image dans la direction de variation spatiale) dans la direction θ .

La présence du problème d'ouverture peut être détectée en utilisant d'autres directions d'analyse que $\theta = 0$ et $\theta = \pi/2$. La figure 4 montre la distribution des estimations pour quatre directions d'analyse (neuf filtres). Dans le cas du problème d'ouverture, les estimations sont distribuées sur une droite, sinon elles sont distribuées sur un cercle.

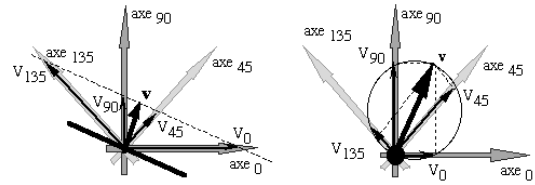


FIG. 4 — Distribution des estimations des composantes du vecteur vitesse. Chaque flèche correspond à une direction d'analyse par un groupe de trois filtres. À gauche la distribution pour une barre en mouvement, et à droite la distribution pour une impulsion de Dirac en mouvement.

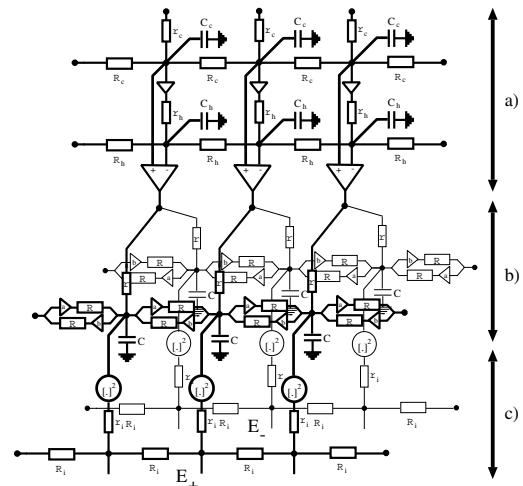


FIG. 5 — Architecture de l'algorithme d'estimation du mouvement pour des images avec une dimension spatiale. Il y a trois étapes de traitement : a) prétraitement de la rétine (filtrage passe bande), b) filtres sélectifs en vitesse (on montre uniquement les deux filtres v_o et $-v_o$) et c) calcul de l'énergie en sortie de chaque filtre avec des réseaux résistifs.

4 Estimation pour images réelles

Dans le cas des images réelles, il est nécessaire de préfiltrer les images pour garantir l'hypothèse de spectre blanc. Les études menées par d'Atick [2] montrent que les images réelles ont des spectres spatio-temporels de la forme $1/f^2$. Le traitement réalisé par la rétine dans le système visuel permet de compenser cette caractéristique du spectre. La rétine réalise un filtrage passe-haut dans la bande utile du signal d'entrée et une atténuation des hautes fréquences pour améliorer le rapport signal sur bruit. Dans notre algorithme nous avons utilisé le modèle de la rétine présenté dans [3] (figure 5). La sortie de la rétine est ensuite filtrée par la batterie de filtres sélectifs en vitesse que nous avons présentée. Le dernière étape du traitement consiste à calculer localement l'énergie en sortie des filtres. Ceci est réalisé par un réseau résistif appliqué au carré de la sortie de chaque filtre, la fenêtre d'intégration $w(x, y)$, est la réponse impulsionnelle de ce réseau. La figure 5 montre le système complet (1-D) pour l'estimation du mouvement. La figure 6 montre les résultats des simulations.

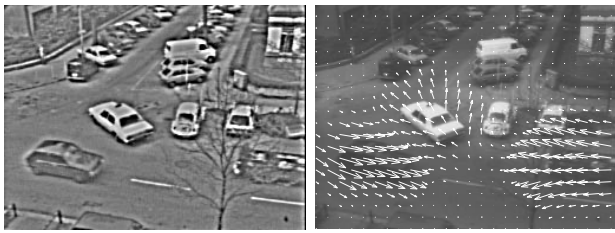


FIG. 6 — Résultats des simulations. À gauche, la sortie de la rétine (réhaussement des contrastes spatio-temporels). La sortie est à moyenne nulle. Le gris du fond représente la sortie nulle. À droite, le champ des vitesses estimées.

5 Intégration de l'énergie

L'intégration de l'énergie est l'étape qui peut régler le problème d'ouverture, car c'est la taille de la fenêtre d'intégration qui détermine la taille de la région d'analyse ; elle est aussi responsable du lissage du champ de vitesses. On a un double compromis, d'une part, éviter le problème d'ouverture demande des fenêtres d'intégration grandes, et d'autre part, pour respecter les discontinuités du champ de vitesses il faut une fenêtre d'intégration petite ! Pour résoudre ce dilemme nous proposons d'utiliser un réseau résistif régularisant avec processus de ligne [7]. La figure 7 montre les résultats sur une séquence composée de deux cercles en mouvement dans des sens opposés. La première estimation est réalisée avec une fenêtre d'intégration petite. Le problème d'ouverture apparaît et l'algorithme donne les vitesses normales (7.a). Dans le second cas (fig 7.b), on utilise une fenêtre d'intégration grande pour éviter le problème d'ouverture mais ceci produit une mauvaise estimation due à la proximité des deux cercles. L'utilisation du réseau avec processus de ligne permet d'intégrer l'énergie uniquement sur chaque cercle pour obtenir une estimation des vitesses correcte (7.d). Ici, les coupures du réseau résistif (fig 7.c) sont placées sur les contours des objets.

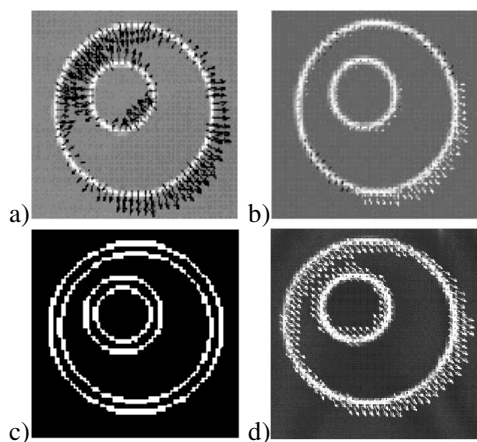


FIG. 7 — Résultats des simulations. a) Estimation avec une petite fenêtre d'intégration et b) avec une grande fenêtre d'intégration. c) Localisation des circuits ouverts sur le réseau de régularisation et d) champ de vitesses final.

6 Conclusion

Nous avons présenté une nouvelle méthode pour l'estimation du mouvement basée sur les circuits neuromorphiques. Les filtres présentés ont le double avantage de réduire la complexité de calcul et de permettre la réalisation sur des circuits intégrés pour des applications en temps réel. L'algorithme proposé utilise très peu de filtres et permet de traiter le problème d'ouverture. La précision des estimations et la robustesse au bruit sont comparables à celles obtenues avec des algorithmes plus complexes basés sur les ondelettes de Gabor.

Références

- [1] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2(2) :284–299, February 1985.
- [2] J. Atick and A. Redlich. What does the retina know about natural scenes ? *Neural Computation*, 4 :196–210, 1992.
- [3] W. Beaudot. *Le traitement neuronal de l'information dans la rétine des vertébrés*. PhD thesis, Laboratory TIRF, Grenoble, France, 1994.
- [4] T. Delbrück. Silicon retina with correlation-based, velocity-tuned pixels. *IEEE Trans. Neural Networks*, 4 :529–541, May 1993.
- [5] R. Douglas, M. Mahowald, and C. Mead. Neuromorphic analogue vlsi. *Annu. Rev. Neurosci.*, 18 :255–281, 1995.
- [6] D. J. Heeger. Model for the extraction of image flow. *J. Opt. Soc. Am. A*, 4(8) :1455–1471, August 1987.
- [7] J. Hutchinson, C. Koch, J. Luo, and C. Mead. Computing motion using analog and binary resistive networks. *IEEE Computer*, pages 52–63, March 1988.
- [8] A. Mitiche and P. Bouthemy. Computation and analysis of image motion : a synopsis of current problems and methods. *International Journal of Computer Vision*, 19(1) :29–55, 1996.
- [9] A. Torralba. Traitements spatio-temporels non linéaires dans un modèle de vision d'inspiration biologique. Dea signal-image-parole, Laboratoire TIRF, Grenoble, France., 1996.
- [10] A. Torralba and J. Héroult. From retinal circuits to motion processing : a neuromorphic approach to velocity estimation. In *ESANN'97*, 1997.
- [11] A. B. Watson and A. J. Ahumada. A look at motion in the frequency domain. Technical Report 84352, NASA, 1983.