

ANALYSE DU FLOT OPTIQUE DANS LES SEQUENCES D'IMAGES PAR RESEAU D'AUTOMATES

Olivier BERNIER

e-mail : bernier@lannion.cnet.fr
FRANCE TELECOM C.N.E.T.
LAB/RIO/TNT Technopole Anticipa
2, avenue Pierre Marzin, 22307 LANNION CEDEX

Résumé

La méthode d'analyse du flot optique dans les séquences d'images animées que nous proposons est basée sur l'utilisation d'un réseau d'automates. Cette méthode permet d'introduire naturellement l'existence de frontières dans le flot optique, dues à la présence de plusieurs objets animés de mouvements différents dans la séquence.

Abstract

We propose in this communication a method to analyze the optical flow in sequences of images, using a network of automata. This method enables us to introduce in a natural way frontiers in the optical flow, caused by the existence of objects with different movements in the sequence.

1 INTRODUCTION

La plupart des algorithmes de calcul du flot optique existants, par exemple les méthodes de gradient [HS81], ou de gradient généralisé [TC94], ne tiennent pas compte de l'existence possible de ruptures du mouvement apparent dues aux frontières entre objets différents (voir par exemple [HS93]). Un certain nombre de techniques pour pallier ce problème ont été proposées: utilisation de contraintes de lissage orienté [NE86] [Ais89], ou calcul de flots optiques multiples superposés [SM94] par exemple. Aucune solution ne semble encore satisfaisante.

La difficulté principale vient du fait que l'estimation du flot optique et sa segmentation sont interdépendantes, et ne peuvent s'effectuer de manière consécutive. Nous proposons dans cet article une nouvelle technique permettant d'introduire des frontières dans le calcul du flot optique. Dans cette technique, basée sur l'utilisation d'un réseau d'automates, ces frontières sont déterminées conjointement avec celui-ci. Nous al-

lons maintenant détailler les différentes étapes de l'algorithme proposé: pré-traitement et réseau d'automates.

2 PRE-TRAITEMENT

La première phase consiste en l'obtention d'indices locaux du flot optique. Nous avons choisi comme indices la projection de ce flot sur le gradient local de luminosité. Désirant une bonne précision de détection, même pour des mouvements importants, nous nous sommes écartés des techniques différentielles et avons utilisé un appariement des structures locales des images. Celles-ci sont définies comme l'approximation parabolique des images au voisinage des points considérés, dans la direction du gradient de luminosité.

Pour s'affranchir en partie des problèmes de bruit et de discrétisation de l'image, il est nécessaire de procéder à une régularisation de celle-ci. Afin de ne pas perdre de précision dans la localisation des contours (forts gradients), nous avons réalisé une régularisation



adaptée. Quatre masques directionnels de détection de contours donnent l'intensité et la direction du gradient. La direction est obtenue par interpolation entre la direction du masque répondant le mieux et ses voisins. La régularisation est obtenue par convolution avec un masque de lissage directionnel adapté à cette direction, avec le degré de sélectivité directionnelle lié à l'intensité du gradient local. La figure 1 donne un exemple d'intensité de gradient de luminosité obtenu pour la première image de la séquence classique "car-phone".

3 LE RESEAU D'AUTOMATES

Le principe de l'algorithme est alors de déterminer, à partir des indices de mouvement obtenus précédemment, potentiellement bruités, un flot optique cohérent incluant ses frontières. Il semble naturel de formuler ce problème de manière variationnelle, mais les frontières sont difficilement représentables sous cette forme, car elles ont, par nature, des valeurs discrètes (présence ou absence).

Pour une formulation variationnelle, on identifie ce problème à celui de la minimisation d'une énergie ayant des variables continues (flot en chaque pixel), mais aussi discrètes (présence de frontière). L'énergie est définie ainsi :

$$\begin{aligned} \mathcal{E} &= a \sum_i (1 - c_i) \|\vec{M}_i - \vec{V}_i\|^2 \\ &+ a \sum_i c_i [\vec{V}_i \cdot \vec{N}_i - \vec{M}_i \cdot \vec{N}_i]^2 \\ &+ b \sum_i c_i [\vec{V}_i \cdot \vec{G}_i - \vec{M}_i \cdot \vec{G}_i]^2 \\ &+ c \sum_i c_i [\vec{V}_i \cdot \vec{G}_i - b_i]^2 \end{aligned} \quad (1)$$

$$\vec{M}_i = \frac{1}{P_i} \sum_j \nu_{ij} (1 - F_i) \vec{V}_j \quad (2)$$

$$P_i = \sum_j \nu_{ij} (1 - F_i) \quad (3)$$

où les sommes s'entendent sur tous les pixels de l'image, ν_{ij} représente les relations de voisinage entre pixels, c_i est un booléen indiquant l'existence d'un indice local de mouvement, b_i est cet indice. \vec{V}_i et F_i sont respectivement le flot optique et ses frontières. \vec{G}_i et \vec{N}_i représentent les vecteur unitaires parallèle et orthogonal au gradient de luminosité au point i . Les paramètres a , b et c donnent les poids relatifs des termes de lissage orthogonalement au gradient de luminosité, de

lissage parallèlement au gradient et de différence avec l'indice local de mouvement. \vec{M}_i s'interprète comme la moyenne des vecteurs voisins qui ne sont pas des frontières, et P_i comme le nombre de ces voisins. Pour simplifier les calculs, nous avons restreint le voisinage de chaque pixel aux quatre plus proches voisins.

Les techniques différentielles ne pouvant être utilisées, nous introduisons alors une technique d'optimisation qui approxime l'algorithme du recuit simulé, le recuit simulé en champ moyen (voir par exemple [YK94]). Dans notre cas, cet algorithme est équivalent à un réseau d'automates continus synchrones (un par pixel), chacun étant connecté à ses quatre plus proches voisins. On obtient comme itération de l'automate, pour le flot :

$$\begin{aligned} \vec{V}_i^{k+1} &= \vec{V}_i^k - \alpha(1 - c_i) [\vec{V}_i^k - \vec{M}_i^k] \\ &- \alpha c_i [\vec{V}_i^k \cdot \vec{N}_i - \vec{M}_i^k \cdot \vec{N}_i] \vec{N}_i \\ &- \beta c_i [\vec{V}_i^k \cdot \vec{G}_i - \vec{M}_i^k \cdot \vec{G}_i] \vec{G}_i \\ &- \gamma c_i [\vec{V}_i^k \cdot \vec{G}_i - b_i] \vec{G}_i \end{aligned} \quad (4)$$

$$\vec{M}_i^k = \frac{1}{P_i^k} \sum_j \nu_{ij} (1 - \phi_j^k) \vec{V}_j^k \quad (5)$$

$$P_i^k = \sum_j \nu_{ij} (1 - \phi_j^k) \quad (6)$$

Dans cette équation, α , β et γ sont les équivalents de a , b et c . \vec{V}_i^k représente l'estimation à l'itération k du flot au point i , ϕ_i^k l'estimation de la probabilité de présence d'une frontière. Tous ces termes ont une interprétation naturelle et découlent directement de l'énergie \mathcal{E} . L'évolution de ϕ_i^k obtenue est plus complexe. Pour plus d'efficacité, nous lui substituons l'évolution heuristique suivante :

$$\phi_i^{k+1} = \delta \left[(2\phi_i^k - 1) + dc_i^k - \frac{P_i^k}{4} dd_i^k \right] + \lambda G \quad (7)$$

$$dc_i^k = \frac{dv_i^k}{\rho_c \|\vec{V}_i^k\| + \tau_c} \quad (8)$$

$$dd_i^k = \frac{\rho_d \|\vec{V}_i^k\| + \tau_d}{dv_i^k} \quad (9)$$

$$dv_i^k = \sqrt{\frac{1}{P_i^k} \sum_j \nu_{ij} \|\vec{V}_j^k - \vec{V}_i^k\|^2} (1 - \phi_j^k) \quad (10)$$

dv_i^k s'interprète facilement comme l'importance des écarts entre le flot local et celui des points voisins qui



FIG. 1 - Première image de la séquence “Mobile and calendar” et gradient de luminosité obtenu

ne sont pas des frontières, dc_i^k est le rapport entre dv_i^k et un seuil dépendant linéairement du flot local induisant la création des frontières, dd_i^k est le rapport entre un seuil du même genre et dv_i^k induisant la destruction des frontières. Le facteur $\frac{P_i^k}{4}$ introduit une modulation de la destruction des frontières suivant la topologie locale de celles-ci (nombre de voisins frontières). G est l'intensité du gradient local de luminosité, utilisé pour faciliter la création de frontières aux niveau des contours de l'image. Le premier terme a pour rôle de forcer ϕ_i^k vers 0 ou 1. Les paramètres δ , λ , ρ_c , ρ_d , τ_c et τ_d sont des paramètres ajustables.

L'observation de l'existence de nombreuses solutions minimisant cette énergie (minima locaux ou globaux multiples) nous oblige à guider la convergence de l'algorithme, en faisant varier les paramètres du réseau d'automates au cours de son évolution. On obtient alors un algorithme en plusieurs phases, chaque phase étant associée à une énergie donnée, s'interprétant de manière directe en terme d'évolution de frontières. Les valeurs normales des paramètres α , β , γ et δ , indépendantes des séquences à traiter, sont les suivantes :

$$\begin{aligned}\alpha &= 0.9 \\ \beta &= 0.8 \\ \gamma &= 0.1 \\ \delta &= 0.1\end{aligned}$$

Le paramètre λ est lié à l'intensité du gradient de lumi-

nosité dans l'image. Une valeur de 0.2 semble donner des résultats corrects pour différentes séquences.

Quatres phases sont nécessaires pour obtenir de bons résultats. La première est une phase de propagation des informations locales sans prise en compte des frontières. Elle est réalisée avec $\beta = 0$, $\gamma = 0.9$, $\delta = 0$, $\lambda = 0$. La deuxième phase est une phase de lissage des données, β et γ revenant à leurs valeurs normales. La troisième phase est une phase de création de frontière, avec tous les paramètres à leur valeur normales, ρ_d et τ_d étant nuls, ρ_c et τ_c décroissant d'une valeur élevée à une valeur faible. Pour la quatrième et dernière phase, la phase de destruction des frontières ρ_d et τ_d croissent de 0 à leur valeur finale alors que ρ_c et τ_c restent figés à cette même valeur. Les deux valeurs initiales de ρ_c et τ_c ont été choisies dans tous nos essais à 0.1.

4 CONCLUSION

La figure 2 montre les frontières obtenues pour le début de la séquence classique “carphone”, ainsi que le flot obtenu, au voisinage du visage. Pour ces résultats, nous avons utilisé les paramètres suivants pour la troisième phase :

$$\begin{aligned}\rho_c &= 0 \\ \tau_c &= 0.02\end{aligned}$$

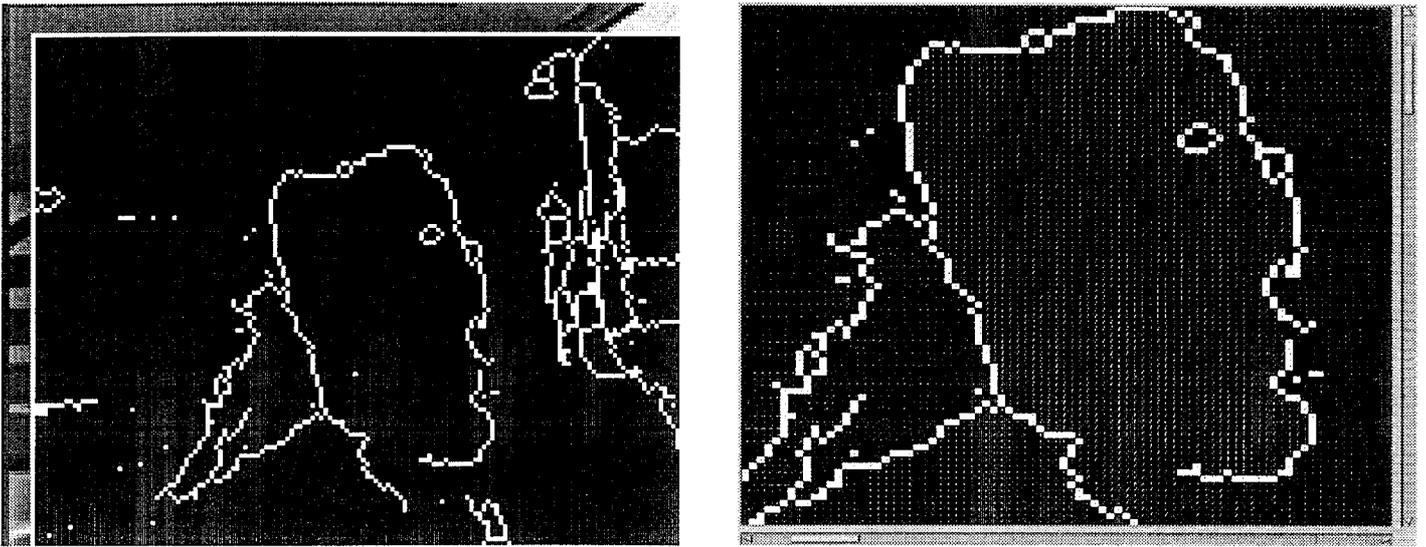


FIG. 2 - Frontières obtenues et détail du mouvement au voisinage du visage

et pour la quatrième :

$$\rho_d = 0.025$$

$$\tau_d = 0.125$$

Le nombre d'itération étant de l'ordre de quelques centaines par phases.

La qualité du résultat est bonne mais semble limitée par la précision des indices obtenus. Le prétraitement choisi ne donne pas suffisamment d'information pour pouvoir différencier totalement le mouvement de la personne par rapport au fond de la voiture. Un des problèmes est l'ajustement des paramètres sensibles (principalement ρ_d et τ_d) qui indiquent la sensibilité des frontières aux variations de mouvement. Ceux-ci doivent être suffisamment faibles sans descendre en dessous du bruit du aux indices locaux de mouvement.

Références

- [Ais89] J. Aisbett. Optical flow with an intensity-weighted smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-11(5):512-522, May 1989.
- [HS81] B.K.P. Horn and B.G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185-203, 1981.
- [HS93] B.K.P. Horn and B.G. Schunk. Determining optical flow: a retrospective. *Artificial Intelligence*, 59:81-87, 1993.
- [NE86] H.H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(5):565-593, September 1986.
- [SM94] M. Shiwaza and K. Maze. Multiple optical flow - fundamental constraint equations and a unified computational theory for detecting motion transparency and motion boundaries. *Systems and Computers in Japan*, 25(5):987-1005, 1994.
- [TC94] T.R. Tsao and V. Chen. A neural scheme for optical flow computation based on Gabor filters and generalized gradient method. *Neurocomputing*, 6(3):305-325, 1994.
- [YK94] A.L. Yuille and J.J. Kosowsky. Statistical physics algorithms that converge. *Neural Computation*, 6(3):341-356, 1994.