



# Analyse multiéchelle de films.

F. Guichard

CEREMADE, Université Paris IX-Dauphine

## RÉSUMÉ

Nous définissons par une approche axiomatique une analyse multiéchelle de film. Nous prouvons que cette analyse est l'unique qui vérifie certaines propriétés d'invariance et de stabilité. Elle débouche sur trois types d'informations : un champ multiéchelle de vitesse apparente, un champ multiéchelle d'accélération apparente, et une certaine classification des trajectoires du film.

## 1 Introduction.

Quel est le but d'une analyse de film ? Brièvement nous voulons éliminer le bruit, "tracker" les trajectoires des objets, et donner une bonne estimation de leur vitesse. Ces tâches sont généralement effectuées en définissant un lissage du film à une certaine échelle considérée comme significative. Comme nous voulons définir une méthode générale, c'est à dire sans aucune connaissance a priori du type de bruit, de trajectoire, ou de taille des objets, la méthode doit être multiéchelle.

Nous associons alors au film original une analyse de celui-ci à toutes les échelles  $t \in \mathbb{R}^+$

Nous définissons un film comme une fonction bornée de  $\mathbb{R}^3$  :  $u(x, y, \theta)$ . (Le niveau de gris au point  $(x, y)$  au temps  $\theta$ ). On appelle analyse multiéchelle une famille d'opérateurs  $T_t$  qui transforment une fonction  $u$  en une famille de fonctions  $u(t) = T_t u$ . C'est à dire  $u(x, y, \theta, t) = (T_t(u))(x, y, \theta)$ .

En un sens  $u(t)$  est la version lissée de  $u$  contenant plus que les informations de "taille"  $t$ .

Cette approche est une extension au cas des films des théories de "scale-space" des images [11], [13], [20] qui ont conduit à associer une analyse multiéchelle à une équation aux dérivées partielles [2]. Mais, de part les équations auxquelles elle conduit, elle se rapproche des modèles de flot optique [15], [8].

## 2 Des principes pour l'analyse multiéchelle.

Dans l'optique de décrire quel est le "bon" opérateur  $T_t$ , nous imposons deux types de propriétés :

- des propriétés "d'architectures" qui donnent une certaine structure à l'analyse et qui imposent à l'analyse d'être un lissage.
- des propriétés géométriques qui imposent une invariance du traitement vis à vis de certaines transformations géométriques.

## ABSTRACT

We define by an axiomatic approach a multiscale analysis of movies. We prove that this analysis is the unique which satisfy some geometrical and stability properties. This analysis gives us three kinds of informations : a multiscale optical flow, a multiscale optical acceleration, and a classification of the trajectories of the objects of the movies.

### Principes d'architectures.

[Causality]

Notons  $B(x, r)$ , la boule de centre  $x$  et de rayon  $r$ .

- (i)  $\forall t, t'; (t \geq t') \exists T_{t,t'} \text{ tel que } T_t = T_{t,t'} T_{t'}$
- (ii)  $T_{t,t} = Id$
- (iii) soit  $u, v$  deux images si  $\forall y \in B(x, r)/x$   
 $u(y) > v(y)$  alors  $T_{t+h,t} u(x) \geq T_{t+h,t} v(x)$   
pour  $h$  suffisamment petit.

(i) et (ii) imposent une structure "pyramidale" pour la famille d'opérateurs  $T_t$  : si  $t' \leq t$ , on peut passer directement du film à l'échelle  $t'$  au film à l'échelle  $t$ .

(iii) implique que l'on effectue un lissage du film initial. C'est à dire que l'on s'interdit de rajouter des détails qui n'existent pas.

[Regularity]

Soit  $u(x, y, \theta) = u(x) = \frac{1}{2} \langle Ax, x \rangle + \langle p, x \rangle + c$  une forme quadratique de  $\mathbb{R}^3$ . Il existe une fonction  $F(A, p, x, c, t)$ , continue par rapport à  $A$ , telle que l'expression

$$(T_{t+\Delta t, t} u - u)(x) / (\Delta t)$$

tendent vers  $F(A, p, x, c, t)$ , quand  $\Delta t$  tends vers 0.

Cette hypothèse signifie que le lissage d'une image lisse se fait de façon régulière.

Propriétés géométriques.

[Transl. Inv.] Appelons  $B_{a,b,c}$  l'opérateur de translation,

$$B_{a,b,c} u(x, y, \theta) = u(x - a, y - b, \theta - c) \quad \text{alors}$$

$$T_t B_{a,b,c} u = B_{a,b,c} (T_t u)$$

[Affine Inv.]

Pour toute application lineaire  $A$  de la forme

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix} \text{ avec } a, b, c, d, e \in \mathbb{R},$$



il existe  $t'(t, A)$  tel que

$$A(T_t u) = T_t(Au)$$

Il est prouvé dans [2] que, sans perte de généralité on peut renormaliser les échelles de telle sorte que

$$t' = (ad - cb)^{1-\alpha} e^{\alpha t} \quad \text{for } \alpha \in [0, 1]$$

Nous définissons aussi un axiome plus faible que [Affine Inv.], en restreignant la propriété d'invariance aux rotations des deux premières coordonnées et aux homothéties.

[Euclidian Inv.] Pour toute application linéaire  $A$  du type

$$\begin{pmatrix} a \cos(b) & -a \sin(b) & 0 \\ a \sin(b) & a \cos(b) & 0 \\ 0 & 0 & c \end{pmatrix} \text{ où } a, b, c \in \mathbb{R}$$

il existe une échelle  $t'(t, A)$  telle que

$$A(T_t u) = T_t(Au)$$

[Morphological Inv.] Pour toute fonction  $h$  strictement monotone  $\mathbb{R}$  dans  $\mathbb{R}$

$$T_t h(u) = h(T_t u)$$

Cette propriété impose l'invariance de l'analyse par rapport à un changement de contraste [16]. Elle se justifie en remarquant la perception du mouvement ne doit pas dépendre de la luminosité ambiante. (le fait de porter des lunettes de soleil n'altère pas notre perception du mouvement des objets...)

Nous supposons enfin que l'analyse est invariante sous travelling, c'est à dire que le mouvement d'une image entière à vitesse constante  $v$  ne change pas l'analyse.

Appellons  $\vec{i}$  et  $\vec{j}$  les vecteurs unitaires spatiaux représentant l'axe des  $x$  et l'axe des  $y$ .

[Galilean Inv.] On définit  $B_v$  l'opérateur de translation uniforme

$$B_v u(x, y, \theta) = u(x - \langle v, \vec{i} \rangle \theta, y - \langle v, \vec{j} \rangle \theta, \theta)$$

### 3 Quels sont les modèles possibles?

Les principes d'architectures impliquent une équation au dérivées partielles.

**Théorème 1** (i) Sous les hypothèses [Causality] et [Regularity], l'analyse multiéchelle de film est "décrite" par l'équation

$$\frac{\partial u}{\partial t} = F(D^2 u, Du, u, x, t) \quad (1)$$

(ii) Si l'on suppose de plus [Transl. Inv.], on obtient

$$\frac{\partial u}{\partial t} = F(D^2 u, Du, u, t) \quad (2)$$

De plus la fonction  $F$  est croissante par rapport à son premier argument. ( $A \geq B$ , pour la relation d'ordre de matrices symétriques, implique  $F(A, p, c, t) \geq F(B, p, c, t)$ ).

Par "décrite" nous entendons que partout où la fonction  $u(x, t)$  est  $C^2$  alors elle vérifie l'équation précédente. Inversement cette équation admet une unique solution de viscosité si la fonction initiale  $u(x)$  est continue [4]. Une démonstration de ce théorème se trouve dans [2].

Nous allons maintenant restreindre le choix possible pour de telles fonctions  $F$ , en imposant les différents principes d'invariance.

**Les propriétés d'invariance déterminent les modèles possibles.** Introduisons quelques notations On note  $\nabla u = (u_x, u_y, 0)$  le gradient spatial du film  $u(x, y, \theta)$ . On associe au gradient  $Du = (u_x, u_y, u_\theta)$  deux autres vecteurs  $T^\perp$  et  $T^\pm$  définis par

$$T^\perp = \frac{1}{|\nabla u|} (-u_y, u_x, 0) \quad T^\pm = \frac{1}{|\nabla u| |Du|} (u_x u_\theta, u_y u_\theta, -(u_x^2 + u_y^2))$$

Quand  $\nabla u$  est non nul,  $(Du/|Du|, T^\perp, T^\pm)$  forment une base orthonormée de  $\mathbb{R}^3$ . Posons de plus,

$$\Gamma_1 = \langle (D^2 u) T^\perp, T^\perp \rangle, \quad \Gamma_2 = \langle (D^2 u) T^\perp, T^\pm \rangle, \\ \Gamma_3 = \langle (D^2 u) T^\pm, T^\pm \rangle.$$

$\Gamma_1$  est la dérivée seconde de  $u$  dans la direction  $Du^\perp$ ,  $\Gamma_3$  dans la direction de  $Du^\pm$ , et  $\Gamma_2$  la dérivée croisée dans ces deux directions.

On a alors la courbure spatiale des lignes de niveaux des images du film  $curv(u)$  donnée par la formule suivante

$$curv(u) = \frac{\Gamma_1}{|\nabla u|}$$

Nous introduisons un opérateur d'analyse du mouvement, "l'accélération apparente", donné par

$$accel(u) = \left( \frac{|Du|}{|\nabla u|} \right)^2 \left( \Gamma_3 - \frac{\Gamma_2^2}{\Gamma_1} \right) / |\nabla u|$$

**Théorème 2** Sous les propriétés [Causality], [Regularity], [Transl. Inv.], [Euclidian Inv.], [Morphological Inv.], et [Galilean Inv.] l'analyse multiéchelle de film est décrite par une équation du type :

$$\frac{\partial u}{\partial t} = |\nabla u| F(curv(u), accel(u), t) \quad (3)$$

Si l'on ajoute [Affine Inv.], l'analyse multiéchelle se restreint à l'équation

$$AMG \quad \frac{\partial u}{\partial t} = |\nabla u| curv(u)^{\frac{1-q}{3}} (sgn(curv(u)) accel(u)^q)^+ \quad (4)$$

pour un  $q \in [0, 1]$ .

Dans les précédentes formules nous prenons la convention que la puissance conserve le signe, c'est à dire  $a^q = |a|^q sgn(a)$ .

La démonstration de ce théorème est essentiellement basée sur des calculs algébriques qui réduisent la fonction  $F$  à la seule possibilité ci-dessus.

Notons que  $|\nabla u|^3 curv(u)$  et  $|\nabla u|^4 curv(u) accel(u)$  sont affines covariants de degré 2,2,0 et 2,2,2, par rapport aux coordonnées  $(x, y, \theta)$ .

De plus  $curv(u)$  et  $accel(u)$  sont invariants par changement des niveaux de gris de l'image.

### 4 Flot optique et accélération apparente.

Nous cherchons ici à donner une interprétation cinématique à l'opérateur différentiel  $accel(u)$ .

Nous supposons que le niveau de gris d'un point ne change pas durant son mouvement. Cette hypothèse ne suffit pas à extraire du film la vitesse des objets. En effet, il reste plusieurs façons de définir une vitesse apparente (flot optique). Toutefois,  $accel(u)$  prend le sens d'accélération pour une définition spécifique du flot optique.

**Définition d'une vitesse apparente.**

**Définition 1** On appelle vitesse apparente le vecteur  $V$  défini par :

$$V = \frac{Du}{\nabla u} (T^\perp \frac{\Gamma_2}{\Gamma_1} - T^\pm) \quad (5)$$

Et, si l'on pose  $v_1 = \langle V, \nabla u \rangle / (|\nabla u|)$ , la composante spatiale de  $V$  dans la direction du gradient spatial  $\nabla u$ , and  $v_2 = \langle V, T^\perp \rangle$  la composante spatiale de  $V$  orthogonale au gradient spatial, on a :

$$v_1 = -\frac{u_\theta}{|\nabla u|} \quad v_2 = \left(\frac{|Du|}{|\nabla u|}\right) \frac{\Gamma_2}{\Gamma_1} \quad (6)$$

La composante  $v_1$  est imposée par l'hypothèse que le niveau de gris d'un point ne change pas lors de son mouvement. En effet,

$$\langle V, Du \rangle = 0$$

et donc,  $\langle V, \nabla u \rangle = -u_\theta \Rightarrow v_1 = -\frac{u_\theta}{|\nabla u|}$

Le choix de la composante  $v_2$  correspond à imposer que le gradient ne change pas d'orientation lors du mouvement. Donc, si le film correspond à un objet en translation la vitesse apparente ainsi définie coïncide (où  $\nabla u$  est non nul) à la vitesse réelle de l'objet.

**accel(u) : une accélération apparente.**

**Proposition 1** Soit  $V$  le vecteur vitesse apparent définie précédemment (1), et  $v_1$  sa composante dans la direction du gradient spatial.

$$\begin{aligned} accel(u) &= -(v_1 \frac{\partial v_1}{\partial x} + v_2 \frac{\partial v_1}{\partial y} + \frac{\partial v_1}{\partial \theta}) \\ &= -\langle Dv_1, V \rangle = \langle D \langle V, -\nabla u \rangle, V \rangle \end{aligned} \quad (7)$$

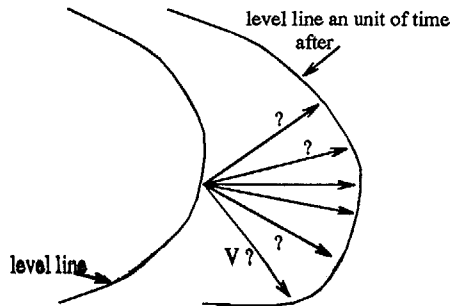
Cette formule montre que  $accel(u)$  est l'accélération dans la direction  $-\nabla u$ , et donc peut être appelé "accélération apparente"

**Proposition 2**

$$accel(u) = \frac{\langle (D^2u)V, V \rangle}{|\nabla u|} = |V|^2 curv_2(u) \quad (8)$$

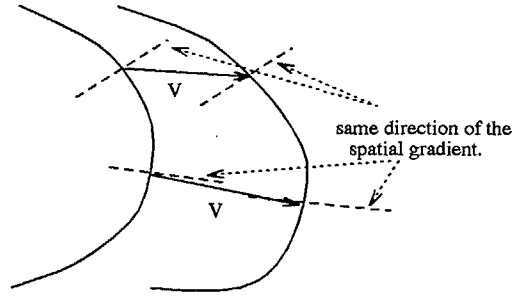
où  $curv_2$  est la courbure de la ligne de niveau incluse dans le plan généré par  $\nabla u$  et  $V$ .

N'importe quelle définition du flot optique correspond à un choix d'une direction de déplacement dans l'espace-temps sur les surfaces isophotes.



En principe, tous les vecteurs tracés si dessus sont possibles puisque ils font rester le point sur la surface isophote. Avec la définition (1), on choisie la direction qui ne change pas l'orientation du gradient spatial.

Alors,  $accel$  est proportionel à la courbure de la trajectoire choisie.



**5 Aspects numériques.**

Insistons seulement sur la difficulté de la discrétisation de l'opérateur  $accel(u)$ . En effet, l'intervalle de temps entre deux images d'un film est trop grand pour que l'on puisse discrétiser  $accel(u)$  par différences finies. Nous proposons un schéma semi-local basé sur l'égalité suivante

$$|\nabla u| (sgn(curv(u)) accel(u))^+ = \min_{v_-, v_+ \in \mathcal{W}}$$

$$\frac{1}{\Delta\theta^2} (|u(x-v_-\Delta\theta)-u(x)| + |u(x+v_+\Delta\theta)-u(x)| + \Delta\theta) \langle \nabla u, v_- - v_+ \rangle \quad (9)$$

où  $\mathcal{W}$  est l'ensemble des vecteurs vitesses possibles c'est à dire  $(\alpha, \beta, 1)$  pour  $\alpha$  et  $\beta \in \mathbb{R}$ .

Nous avons deux parties dans le second terme. La première  $(|u(x-v_-\Delta\theta)-u(x)| + |u(x+v_+\Delta\theta)-u(x)|)$  est la variation de niveau de gris du point  $x$  pour les vecteurs vitesses candidats :  $v_-$  (entre  $\theta - \Delta\theta$  et  $\theta$ ) et  $v_+$  (entre  $\theta$  et  $\theta + \Delta\theta$ ). Bien sûr cette variation doit être la plus petite possible, parce que le point est supposé ne pas changer de niveau de gris lors de son mouvement.

La deuxième partie  $|\langle \nabla u, v_- - v_+ \rangle|$  n'est rien d'autre que l'"accélération", ou la différence entre  $v_+$  et  $v_-$  dans la direction  $\nabla u$ .

La minimisation ne se fait pas en pratique sur l'ensemble des vecteurs possibles : si l'on sait au préalable que les vecteurs vitesses du objets du film sont bornés, on se limitera à minimiser sur les vecteurs ayant les mêmes bornes.

**6 Interpretation et Expérience.**

L'équation (4) peut s'interpréter comme une élimination des points selon leur accélérations. En effet

$$AMG \quad \frac{\partial u}{\partial t} = |\nabla u| curv(u)^{\frac{1-q}{3}} (sgn(curv(u)) accel(u)^q)^+$$

se compose de deux termes :  $|\nabla u| curv(u)^{\frac{1-q}{3}}$  qui correspond à un lissage des lignes de niveau des images (diffusion anisotrope sur chaque image). Et,  $accel(u)$  peut se voir comme la vitesse à laquelle se fait le lissage. Il n'y a pas de lissage lorsque  $accel(u) = 0$ , c'est à dire lorsque le point est en translation uniforme. En revanche le lissage est accéléré lorsque  $accel(u)$  est grand, c'est à dire lorsque le point a une grande accélération. Donc les trajectoires chaotiques ont tendance à disparaître très vite, alors que les trajectoires lisses se maintiennent plus longtemps en échelle. Et, lorsque l'échelle tend vers  $+\infty$  il ne reste plus que les points en translation uniforme.

Les expériences suivantes ont été faites sur une station HP730 (32MB). Le temps de calcul est pour une séquence de 10 images  $256*256$ , en limitant les vitesses à 8 pixels par images, d'environ 20 minutes.

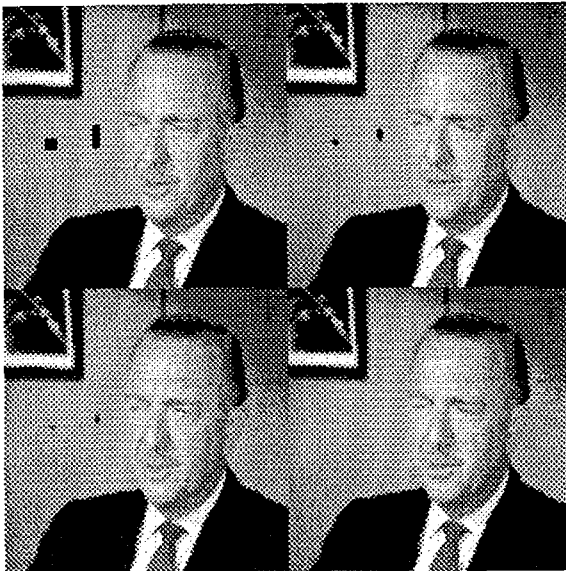


FIG. 1 - Image extraite d'une séquence, présentée à différentes échelles d'analyse. (la vitesse de déplacement de la tête est de l'ordre de 7 pixels par images.) Le bruit part.

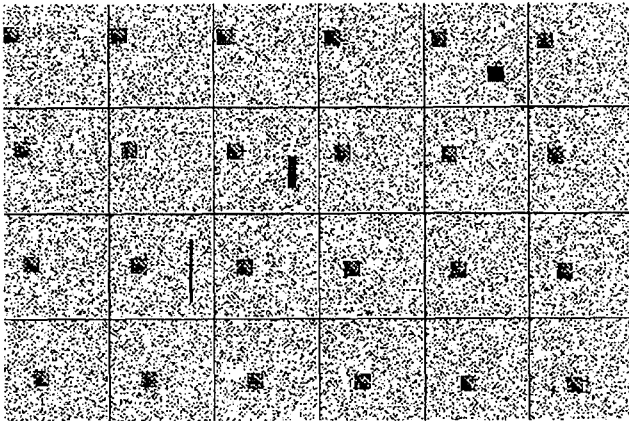


FIG. 2 - Film original (24images) d'un carré et d'un point en translation uniforme, (vitesse = 2 pixels/images). Bruit blanc à 30 %

## Références

- [1] E. H. Adelson and J.R. Bergen. *Spatiotemporal energy models for the perception of motion* J. Opt. Soc. Amer. A2(2) p. 284-299, 1985.
- [2] L. Alvarez, F. Guichard, P.L. Lions and J.M. Morel. *Axioms and fundamental equations of image processing. Report 9216, CEREMADE. Université Paris Dauphine, 1992.* Accepted at Arch. for Rat. Mech., July 1992.
- [3] J. Arnsperg *Optic Acceleration.* Proceeding of second Int. Conf. Comp. Vision. IEEE. 1988.
- [4] M.G. Crandall, H. Ishii and P-L. Lions. *User's guide to viscosity solution of second order partial differential equation.* Preprint, CEREMADE, 1990.
- [5] O. Faugeras. *Sur le mouvement des courbes tridimensionnelles et sa relation avec le flot optique.* C.R. Acad. Sci. Paris, p. 1279-1285, 1991, t.312, Série2.
- [6] J.J. Gibson. *The Perception of the Visual World.* Boston: Houghton Mifflin. 1950.
- [7] N.M. Grzywacz, A.L. Yuille. *The motion coherence theory.* Second conference on computer vision 1988, IEEE Computer Society Press.
- [8] D.Heeger. *Optical flow from spatiotemporal filters.* First conference on computer vision 1987, IEEE Computer Society Press.
- [9] E.C. Hildreth *The measurement of Visual Motion.* Cambridge: MIT Press, 1984.
- [10] Horn B.K.P. and B.G. Schunck. *Determining optical flow.* Artif. Intell. 17, p 185-203. 1981.
- [11] J.J. Koenderink. *The structure of images,* Biol Cybern. 50, p 363-370, 1984.
- [12] J.J. Koenderink. *Optic Flow,* Vis, Res. 26, p161-168, 1986.
- [13] D. Marr. *Vision.* Freeman and Co. 1982.
- [14] H.H. Nagel *On the estimation of the optical flow: relations between different approaches and some new results.* Artificial Intelligence 33: 299-324, 1988.
- [15] H.H. Nagel *Image sequences -ten (octal) years - from phenomenology towards a theoretical foundation.* Intern. J. Patt. Recog. Artif. Intell. 2(3): 459-483, 1988.
- [16] J.Serra *Image analysis and mathematical morphology.* Vol.1 Academia Press 1982.
- [17] A. Singh. *Image-flow computation: An estimation-theoretic framework, unification and integration.* PhD thesis, Departement of Computer Science, Columbia University, 1990.
- [18] M.A Snyder. *On the mathematical foundations of smoothness constraints for the determination of the optical flow and for surface reconstruction.* IEEE Transactions on Pattern Anal. and Mach. Intel, vol 13, No 11, November 1991.
- [19] A.Verri and T.Poggio. *Against quantitative optical flow.* Conference on computer vision 1987. IEEE Computer Society Press.
- [20] A.P. Witkin *Scale-space filtering.* Proc. of IJCAI, Karlsruhe, p 1019-1021, 1983.

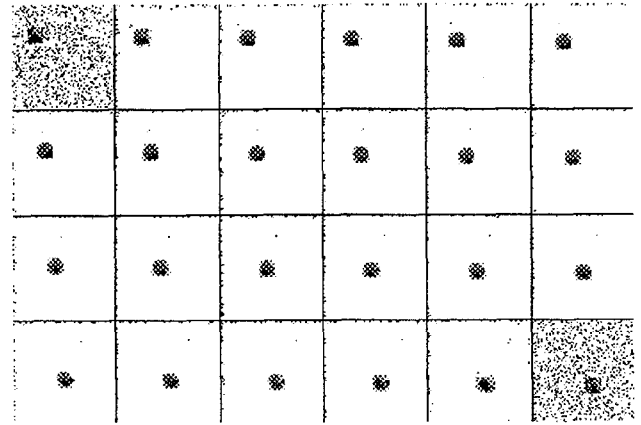


FIG. 3 - L'algorithme est appliqué en supposant que les vitesses ne sont pas plus grandes que 3 pixels par images. A une certaine échelle il ne reste plus que le carré et le point.