

Analyse multiéchelle de films pour la reconstruction du relief

LIONEL MOISAN

Résumé - Nous décrivons l'unique filtrage multiéchelle de films causal, local, et compatible avec la reconstruction du relief lors d'une observation selon un mouvement de translation rectiligne de la caméra, et montrons qu'il peut s'interpréter comme l'unique filtrage linéaire du mouvement de la caméra laissant le relief invariant.

Multiscale analysis of movies for depth recovery

Abstract - We describe the unique multiscale analysis of movies that is causal, local, and compatible with the depth recovery in the case of a straight movement of the camera, and we show that it can be interpreted as the unique linear filtering of the camera movement which preserves depth.

Abridged english version - Let us define a movie as a bounded map $u : \mathbb{R}^3 \rightarrow \mathbb{R}$, where $u(x, y, \theta)$ is the grey-value at point (x, y) of the image at time θ . If we observe in perspective a surface $Z(X, Y)$ by a straight movement $X = C(\theta)$, in such a way that a point $M(X, Y, Z)$ of the scene is projected at time θ onto $P(x = \frac{X-C(\theta)}{Z}, y = \frac{Y}{Z}, \theta)$ in the image referential, then we can hope to recover the depth Z of the scene from the movie by using the equation $Z = -\frac{1}{v}C'(\theta)$, where $v = -\frac{u_x}{u_x}$ is the velocity field on the movie. However, on the one hand the movie is often difficult to process as it is, on the other hand the speed of the camera, $V(\theta) = C'(\theta)$ is neither always known nor constant, so that it is necessary to regularize the movie in a way we study here as multiscale analysis.

A multiscale analysis is a family of operators $(T_t)_{t \geq 0}$ which, applied to an initial movie u_0 , leads to filtered versions $u(t) = T_t u_0$ at all scales t . It has been shown (cf [1]) that under reasonable hypotheses of locality, causality, and space/time/grey-level translation invariance, multiscale analyses of movies can be described by partial derivative equations of the type $\frac{\partial u}{\partial t} = F(D^2 u, Du, t)$, with initial condition $u(., 0) = u_0$.

In addition to these hypotheses, we constrain the analysis to satisfy three invariance properties in agreement with the depth recovery :

1. Any operator T on u which preserves the velocity field commutes with the analysis.
[v-Compatibility] : $\forall T : u \mapsto T(u, x, y, \theta), (\forall u, v[u] = v[Tu]) \Rightarrow (\forall t, T \circ T_t = T_t \circ T)$.
2. The analysis commutes with the superimposition of a uniform straight translation movement on the movie (cf [1]).
[Galilean Invariance] : $\forall u, t, \alpha, (T_t u) \circ B_\alpha = T_t(u \circ B_\alpha)$, where $B_\alpha(x, y, \theta) = (x - \alpha\theta, y, \theta)$.
3. The analysis does not depend on the focal of the camera.
[Zoom Invariance] : $\forall u, t, \lambda, (T_t u) \circ H_\lambda = T_t(u \circ H_\lambda)$, where $H_\lambda(x, y, \theta) = (\lambda x, \lambda y, \theta)$.

Theorem 1 *Under the axioms [v-Compatibility], [Galilean Invariance], [Zoom Invariance], the multiscale analysis satisfies, up to a rescaling, the unique partial differential equation*

$$\frac{\partial u}{\partial t} = tu_{\xi\xi} \quad \text{with } \xi = \left(-\frac{u_\theta}{u_x}, 0, 1\right) \text{ and } u_{\xi\xi} = [D^2u](\xi, \xi)$$

and for this equation, scale is covariant to θ .

Proposition 1 *The multiscale analysis $\frac{\partial u}{\partial t} = tu_{\xi\xi}$ diffuses each component of the movement (velocity $v = \frac{dP}{d\theta}$, acceleration $\frac{d^2P}{d\theta^2}, \dots, \frac{d^n P}{d\theta^n}, \dots$) in the same direction as u , i.e*

$$\frac{\partial}{\partial t} \left(\frac{d^n P}{d\theta^n} \right) = t \left(\frac{d^n P}{d\theta^n} \right)_{\xi\xi}.$$

In particular, v follows an intrinsic, polynomial and causal diffusion equation,

$$\frac{\partial v}{\partial t} = tv_{\xi\xi} = t(v_{\theta\theta} + 2vv_{\theta x} + v^2v_{xx}).$$

Moreover, we give another justification of the fundamental equation obtained in theorem 1. Let us say a movie $u(x, y, \theta)$ is ideal if we can find some functions $C(\theta)$, $Z(X, Y)$ and $U(X, Y)$ so that everywhere

$$u\left(\frac{X - C(\theta)}{Z(X, Y)}, \frac{Y}{Z(X, Y)}, \theta\right) = U(X, Y),$$

meaning that the movie u can be interpreted as a scene $Z(X, Y)$, $U(X, Y)$ (depth and lambertian luminosity) observed under camera movement $C(\theta)$.

Theorem 2 *The equation $\frac{\partial u}{\partial t} = tu_{\xi\xi}$ produces, from an ideal movie u_0 , a sequence of ideal movies $u(t)$, with the same depth interpretation as u_0 , but for which the underlying camera movement is a linearly filtered version of u_0 's one. This is the unique v-compatible multiscale analysis of movies that satisfies these properties ; moreover, the evolution of the movie interpretation is given by*

$$\frac{\partial Z}{\partial t}(X, Y, t) = 0 \quad ; \quad \frac{\partial C}{\partial t}(\theta, t) = t \frac{\partial^2 C}{\partial \theta^2}.$$

On définit un film comme une fonction u bornée de \mathbb{R}^3 dans \mathbb{R} , dont la valeur $u(x, y, \theta)$ représente le niveau de gris au point (x, y) de l'image à l'instant θ . Lorsqu'on observe en perspective une surface $Z(X, Y)$ selon un mouvement de translation rectiligne $X = C(\theta)$, de sorte qu'un point $M(X, Y, Z)$ de la scène observée se projette à l'instant θ en $P(x = \frac{X - C(\theta)}{Z}, y = \frac{Y}{Z}, \theta)$ dans le référentiel de l'image, on peut théoriquement reconstruire le relief Z de la scène à partir du film obtenu grâce à l'équation simple $Z = -\frac{1}{v}C'(\theta)$, v étant le champ des vitesses sur le film, déterminé par $v = -\frac{u_\theta}{u_x}$. Cependant, d'une part le film brut est souvent difficile à exploiter tel quel, parce que le calcul direct du champ des vitesses, très sensible au bruit et à l'estimation du gradient, est en général impossible, d'autre part la vitesse $V(\theta) = C'(\theta)$ de la caméra n'est pas toujours connue ou constante. Il devient alors nécessaire de faire subir au film un traitement régularisant préalable en exploitant la redondance de l'information qu'il contient sur le relief de la scène, ce que nous étudions ici sous la forme d'analyse multiéchelle.

Une analyse multiéchelle est une famille d'opérateurs $(T_t)_{t \geq 0}$ qui, appliquée à un film initial u_0 , conduit à des versions filtrées $u(t) = T_t u_0$ à des échelles t de plus en plus grossières. On sait (cf [1]) que sous des hypothèses raisonnables de localité, causalité, et invariance par translation d'espace, de temps et de niveau de gris, les analyses multiéchelles de films sont décrites par une équation aux dérivées partielles de la forme

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

avec la condition initiale $u(., 0) = u_0$, où F est une fonction continue et croissante par rapport à son premier argument.

Nous imposons de plus à l'analyse multiéchelle de vérifier des propriétés d'invariance en accord avec la reconstruction du relief. Lorsque la vitesse de la caméra $V(\theta)$ est connue ou constante, la reconstruction du relief est équivalente au seul calcul du champ des vitesses sur le film. Il est alors naturel d'imposer que tout opérateur sur les films préservant le champ des vitesses (et donc le relief) commute avec l'analyse, ce qui peut s'exprimer par l'axiome suivant :

Axiome 1 *L'analyse multiéchelle (T_t) vérifie [v-Compatibility] ssi tout opérateur $T : u \mapsto T(u, x, y, \theta)$, tel que $v[Tu] = v[u]$ pour tout film u , vérifie $T \circ T_t = T_t \circ T$ à toute échelle t .*

L'axiome suivant est l'invariance galiléenne classique, déjà utilisée dans [1], exprimant la commutativité entre l'analyse multiéchelle et la superposition d'un mouvement de translation rectiligne uniforme B_α sur le film :

Axiome 2 *L'analyse multiéchelle (T_t) vérifie [Galilean Invariance] ssi elle commute avec l'opérateur $B_\alpha(x, y, \theta) = (x - \alpha\theta, y, \theta)$, i.e. pour tout film u et pour tout réel $\alpha : (T_t u) \circ B_\alpha = T_t(u \circ B_\alpha)$.*

Enfin, pour que l'analyse multiéchelle ne dépende pas de la focale de la caméra, il convient d'imposer qu'elle commute avec toute homothétie spatiale :

Axiome 3 *L'analyse multiéchelle (T_t) vérifie [Zoom Invariance] ssi pour toute homothétie $H_\lambda(x, y, \theta) = (\lambda x, \lambda y, \theta)$, on a $(T_t u) \circ H_\lambda = T_t(u \circ H_\lambda)$.*

Ces trois axiomes, ajoutés aux axiomes élémentaires du filtrage multiéchelle (localité, causalité, invariance par translation d'espace, de temps et de niveaux de gris), déterminent une unique équation d'évolution.

Théorème 1 *Sous les axiomes [v-Compatibility], [Galilean Invariance] et [Zoom Invariance], l'analyse multiéchelle vérifie, à une renormalisation d'échelle près, l'unique équation aux dérivées partielles*

$$\frac{\partial u}{\partial t} = t u_{\xi\xi} \quad \text{avec } \xi = \left(-\frac{u_\theta}{u_x}, 0, 1\right) \text{ et } u_{\xi\xi} = [D^2 u](\xi, \xi) \quad (2)$$

et pour cette équation, l'échelle est covariante à θ .

Démonstration simplifiée : Nous allons tout d'abord donner quelques notations. Considérons une fonction $f(P(\theta), \theta) = f(x, y, \theta)$, où $P(\theta)$ est la trajectoire sur le film d'un point fixe de la scène. On peut définir la dérivée *totale* (ou dérivée selon le mouvement) de f par

$$\frac{Df}{D\theta} = \frac{d}{d\theta} f(P(\theta), \theta) = f_\theta + \nabla f \cdot \frac{dP}{d\theta} = f_\theta + v f_x.$$

Ainsi, l'hypothèse lambertienne de conservation de la luminosité de P , qui s'écrit $\frac{Dv}{D\theta} = 0$, permet de définir v en termes de l'image u par $v = -\frac{u_\theta}{u_x}$. On définit ensuite de manière naturelle l'accélération

$$\Gamma = \frac{d^2P}{d\theta^2} = \frac{Dv}{D\theta} = v_\theta + vv_x,$$

soit, sous forme développée :

$$\Gamma = -\frac{1}{u_x} \left(u_{\theta\theta} - 2\frac{u_\theta}{u_x}u_{\theta x} + \left(\frac{u_\theta}{u_x}\right)^2 u_{xx} \right) = -\frac{u_{\xi\xi}}{u_x}.$$

- Pour établir la preuve du théorème précédent, nous utilisons tout d'abord le fait que de l'axiome [v-Compatibility] découlent deux autres propriétés, à savoir [Morphological Invariance], qui traduit la commutativité de T_t avec la transformation $u \mapsto H(u)$ pour toute fonction H croissante (cf. [1]), et [Transversal Invariance], qui traduit l'invariance de l'analyse vis-à-vis de la coordonnée y [i.e. T_t commute avec l'opérateur $u(x, y, \theta) \mapsto u(x, y, \theta)f(y)$ pour toute fonction f injective].

Lemme 1 *Sous les axiomes [Morphological Invariance] et [Transversal Invariance], l'analyse multiéchelle peut s'écrire sous la forme*

$$\frac{\partial u}{\partial t} = u_x F(\Gamma, v, t) \tag{3}$$

Éléments de démonstration du lemme 1: Partant de l'équation (1) : $\frac{\partial u}{\partial t} = F(D^2u, Du, t)$, on montre tout d'abord en utilisant l'invariance transversale que F ne peut dépendre de u_{yy} . Dans un second temps, par un procédé inspiré de Giga et Goto [2] utilisant la croissance de F , on élimine toute dépendance en u_{xy} et $u_{y\theta}$. Le dernier terme selon y , u_y , disparaît alors naturellement en réutilisant l'axiome d'invariance transversale, et l'on aboutit ainsi à une équation du type $\frac{\partial u}{\partial t} = F(u_{xx}, u_{x\theta}, u_{\theta\theta}, u_x, u_\theta, t)$.

On sait alors (cf. [1]) que pour les termes du second ordre, la croissance de F et la morphologie restreignent les possibilités à des termes de la forme $[D^2u](\vec{a}, \vec{b})$, où \vec{a} et \vec{b} sont dans le plan orthogonal à Du , noté $(Du)^\perp$. Comme nous venons de le voir, l'axiome [Transversal Inv.] interdit de plus toute composante en y , de sorte que \vec{a} et \vec{b} doivent appartenir à la droite $(Du)^\perp \cap (x, \theta) = (\xi)$. On obtient ainsi l'unique terme (morphologique) admissible du second ordre, $\Gamma = -\frac{1}{u_x}[D^2u](\vec{\xi}, \vec{\xi})$, à un coefficient multiplicatif du premier ordre près. Pour réduire les termes restant du premier ordre, on réécrit $\frac{\partial u}{\partial t} = F(\Gamma, u_x, u_\theta, t)$ en $\frac{\partial u}{\partial t} = u_x G(\Gamma, v, t, u_x)$ et l'invariance morphologique montre alors que G ne peut dépendre de u_x , d'où le résultat annoncé.

Puisque l'axiome [v-Compatibility] supprime toute dépendance de l'analyse multiéchelle selon la coordonnée y , ce qui revient encore à dire que cette analyse opère indépendamment selon les sections $y = \text{cte}$, nous ignorerons souvent par la suite la coordonnée y et écrirons $u(x, \theta)$ pour $u(x, y, \theta)$, y étant supposé fixé.

- Pour achever la démonstration du théorème, on réduit séparément l'équation (3) par les axiomes [Galilean Invariance] et [Zoom Invariance], qui mènent respectivement aux expressions $\frac{\partial u}{\partial t} = u_x F(\Gamma, t)$ et $\frac{\partial u}{\partial t} = u_\theta F\left(\frac{\Gamma}{v}, t\right)$. Il suffit alors de remarquer que la seule forme commune à ces deux expressions est

$$\frac{\partial u}{\partial t} = -u_x \Gamma f(t) = u_\theta \frac{\Gamma}{v} f(t) = u_{\xi\xi} f(t)$$

et puisque cette évolution est définie à une renormalisation près de la variable d'échelle t , on peut choisir f de manière à ce que t soit homogène à un terme temporel de degré 1, ce qui est obtenu pour $f(t) = t$.

Proposition 1 *L'analyse multiéchelle définie par $\frac{\partial u}{\partial t} = tu_{\xi\xi}$ diffuse toutes les composantes du mouvement (vitesse $v = \frac{dP}{d\theta}$, accélération $\Gamma = \frac{d^2P}{d\theta^2}, \dots, \frac{d^n P}{d\theta^n}, \dots$) dans la même direction que u , i.e*

$$\frac{\partial}{\partial t} \left(\frac{d^n P}{d\theta^n} \right) = t \left(\frac{d^n P}{d\theta^n} \right)_{\xi\xi}.$$

En particulier, v vérifie une équation de diffusion intrinsèque, polynomiale et causale,

$$\frac{\partial v}{\partial t} = tv_{\xi\xi} = t(v_{\theta\theta} + 2vv_{\theta x} + v^2v_{xx}).$$

Démonstration simplifiée : • Pour établir ces propriétés, il est intéressant d'introduire le formalisme des crochets de Lie associés aux dérivées partielles, $\frac{\partial}{\partial x}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial t}$, qui commutent entre elles, et à la dérivée totale $\frac{D}{D\theta}$. Plus précisément, on établit sans difficulté :

$$\left[\frac{\partial}{\partial x}, \frac{D}{D\theta} \right] = \frac{\partial}{\partial x} \frac{D}{D\theta} - \frac{D}{D\theta} \frac{\partial}{\partial x} = v_x \frac{\partial}{\partial x} \quad ; \quad \left[\frac{\partial}{\partial \theta}, \frac{D}{D\theta} \right] = v_\theta \frac{\partial}{\partial x} \quad ; \quad \left[\frac{\partial}{\partial t}, \frac{D}{D\theta} \right] = v_t \frac{\partial}{\partial x}$$

Notons au passage que la notation $f_{\xi\xi} = [D^2 f](\xi, \xi)$ correspond en fait à l'opérateur

$$(\)_{\xi\xi} = \frac{D^2}{D\theta^2} - \Gamma \frac{\partial}{\partial x}.$$

Lemme 2 *Indépendamment de toute équation d'évolution,*

$$\left[\frac{\partial}{\partial t} - t(\)_{\xi\xi}, \frac{D}{D\theta} \right] = (v_t - tv_{\xi\xi}) \frac{\partial}{\partial x}. \quad (4)$$

• Si l'on applique le lemme précédent à u dans le cas de l'évolution $u_t = tu_{\xi\xi}$, le terme de gauche est nul et l'on obtient bien par conséquent la relation $\frac{\partial v}{\partial t} = tv_{\xi\xi}$. Puisque le terme de droite dans (4) est alors un opérateur identiquement nul, on obtient, pour toute fonction q sur l'image vérifiant $\frac{\partial q}{\partial t} = tq_{\xi\xi}$, l'égalité $\frac{\partial}{\partial t} \left(\frac{Dq}{D\theta} \right) = t \left(\frac{Dq}{D\theta} \right)_{\xi\xi}$. Par récurrence, la relation $\frac{\partial v}{\partial t} = tv_{\xi\xi}$ s'étend donc à toutes les dérivées totales successives de v , avec $\frac{D^{n-1}v}{D\theta^{n-1}} = \frac{d^n P}{d\theta^n}$.

Nous donnons maintenant une seconde justification de l'équation fondamentale obtenue dans le théorème 1. Nous dirons qu'un film $u(x, y, \theta)$ est idéal si l'on peut trouver des fonctions $C(\theta)$, $Z(X, Y)$ et $U(X, Y)$ telles qu'en tout point, on ait

$$u \left(\frac{X - C(\theta)}{Z(X, Y)}, \frac{Y}{Z(X, Y)}, \theta \right) = U(X, Y),$$

ce qui signifie que u n'est autre que l'observation de la scène $Z(X, Y), U(X, Y)$ (relief et luminosité lambertienne) par une caméra de trajectoire $C(\theta)$.

Théorème 2 *L'équation $\frac{\partial u}{\partial t} = tu_{\xi\xi}$ produit, à partir d'un film idéal u_0 , une séquence de films idéaux $u(t)$, dont l'interprétation du relief est identique à celle de u_0 , et pour lesquels le mouvement sous-jacent de la caméra est une version linéairement filtrée de celui de u_0 . C'est l'unique analyse multiéchelle de films v -compatible qui possède ces propriétés ; de plus, l'interprétation du film initial évolue selon les équations*

$$\frac{\partial Z}{\partial t}(X, Y, t) = 0 \quad ; \quad \frac{\partial C}{\partial t}(\theta, t) = t \frac{\partial^2 C}{\partial \theta^2}.$$

Shéma de démonstration : Grâce aux équations de passage du référentiel (X, Y, θ, t) de la scène au référentiel (x, y, θ, t) de l'image, données par $x = \frac{X-C(\theta,t)}{Z(X,Y,t)}$ et $y = \frac{Y}{Z(X,Y,t)}$, on peut identifier la différentielle de Z d'un référentiel à l'autre pour obtenir l'évolution de Z dans le référentiel de la scène :

$$\left(\frac{\partial Z}{\partial t}\right)_{\text{scene}} = \frac{\partial Z}{\partial X} \left[\frac{ZZ_t}{Z_x} - C_t \right].$$

Si une analyse multiéchelle vérifie les propriétés du théorème, le lemme 1 nous autorise à réécrire l'équation d'évolution en

$$\frac{\partial u}{\partial t} = u_\theta F\left(\frac{\Gamma}{v}, v, t\right) = u_\theta F\left(\frac{V'}{V}, -\frac{V}{Z}, t\right).$$

La dérivée totale de cette expression permet de lier $V_t, V, V', V'', Z_t, Z, Z_x$ aux dérivées partielles de F , et les conditions de préservation du relief ($(\frac{\partial Z}{\partial t})_{\text{scene}} = 0$) et de linéarité sur l'évolution de V conduisent alors naturellement aux équations cherchées, soit, à un changement d'échelle près, $u_t = tu_{\xi\xi}$ et $C_t = tC''$. Réciproquement, on vérifie aisément que l'évolution $u_t = tu_{\xi\xi}$ satisfait bien les propriétés annoncées.

REFERENCES

1. L.Alvarez, F.Guichard, P.-L.Lions, J.-M.Morel , "Axioms and Fundamental Equations of Image Processing". *Archive for Rational Mechanics*, vol. 123, 1993, pp. 199-304.
2. Y.Giga, S.Goto, *Motion of Hypersurfaces and geometric equations*, *J. Math. Soc. Japan*, 44, n.1, 1992, pp. 99-111.
3. O.Faugeras , "On the motion of 3D curves and its relationship to optical flow". *rapport INRIA*, n. 1183, mars 1990.
4. J.Serra, *Image analysis and mathematical morphology*, 1, Academic Press, 1982.

CEREMADE, Université Paris-Dauphine,
place du Maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France.