



Champs de Markov multirésolution et algorithme multigrille pour l'estimation de mouvement

E. Mémin* & P. Pérez†

* LAMSI, Université de Bretagne Sud, rue Montaigne, B.P. 1104, 56014 Vannes, France (e-mail: memin@iu-vannes.fr)

† IRISA/INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France (e-mail: perez@irisa.fr)

RÉSUMÉ

Nous nous intéressons dans cet article à l'estimation de champs denses de déplacement entre deux images consécutives d'une séquence. Cet estimateur multirésolution, formulé dans un cadre markovien, s'exprime comme l'optimisation globale d'une fonction d'énergie. Celle-ci intègre des estimateurs robustes permettant d'une part d'extraire les discontinuités spatiales du champ à estimer et d'autre part de s'affranchir de données aberrantes. Définie dans un contexte multirésolution afin d'accéder de façon incrémentale aux mouvements de grande amplitude, la minimisation est menée au moyen d'un algorithme multigrille très efficace.

1. INTRODUCTION

De très nombreux problèmes en analyse d'images sont exprimés comme la recherche d'un optimum global d'une fonction d'énergie décrivant les différentes interactions entre les variables du système. Les champs markoviens offrent en particulier un cadre mathématique cohérent et unifié où les techniques d'estimation bayésiennes usuelles (critère du MAP) conduisent à minimiser une telle fonction.

Dans ce cadre, nous nous intéressons à l'estimation de champs denses de déplacements entre deux images consécutives d'une séquence. L'estimateur que nous décrivons dans cet article s'exprime comme la minimisation d'une fonction d'énergie globale non-linéaire. Cette fonction d'énergie est de façon classique composée d'un terme reliant les données observables (les gradients spatio-temporels de la luminance) aux vecteurs à estimer et d'un terme spécifiant les propriétés spatiales du champ de vecteurs vitesse. Ces deux termes mettent en jeu des fonctions de coût issues de la statistique robuste [2, 4]. Ces fonctions permettent d'extraire les discontinuités spatiales du champ à estimer et de s'affranchir de données aberrantes.

L'expression de cet estimateur dans un contexte multirésolution permet de mesurer les déplacements de grande amplitude. Les techniques multirésolutions employées pour l'estimation du flot optique [3, 6], s'appuient généralement sur des pyramides d'images gaussiennes $f^k(t) = \{f^k(s, t), s \in S^k\}$, $k = 0$ (résolution fine) ... $k = N$ (résolution grossière), construites par sous-échantillonnages et filtrages successifs des images originales $f(t) = \{f(s, t), s \in S\}$, $f(s, t)$ désignant la luminance à l'instant t du point s de la grille S des pixels. Une fois créées, ces pyramides sont utilisées

ABSTRACT

In this paper we address the intricate issue of recovering (long range) apparent velocity field between consecutive frames of an image sequence. Within the Bayesian estimation framework, we design a global objective function to be minimized. This energy function incorporates robust estimators. A multiresolution definition of this differential estimation method aims at accessing long range displacements in a coarse-to-fine incremental way. As for the associated successive minimizations, they are processed through a very efficient deterministic multigrid relaxation algorithm.

pour estimer de manière incrémentale les champs de déplacements : les grands déplacements sont estimés aux plus faibles résolutions et raffinés ensuite aux résolutions les plus fines. Ce processus de "raffinage" successif est mené en estimant un champ incrémental $d\mathbf{w}^k \in \Omega^k$ défini sur S^k , étant donné le champ "principal" \mathbf{w}^k issu de l'estimation à la résolution précédente $k + 1$.

L'estimation du flot optique conduit alors à résoudre $N + 1$ problèmes d'optimisation qui peuvent néanmoins s'avérer difficiles. Afin, d'éviter l'emploi de techniques d'optimisation stochastique extrêmement coûteuses d'un point de vue calculatoire, la minimisation est conduite à l'aide d'un algorithme de relaxation multigrille efficace et rapide proposé dans [7].

2. ESTIMATION MULTIRÉSOLUTION ROBUSTE DU FLOT OPTIQUE

Dans le cadre de l'estimation du MAP, le modèle markovien utilisé débouche sur le problème d'optimisation :

$$\widehat{d\mathbf{w}}^k = \arg \min_{d\mathbf{w}^k \in \Omega^k} H_1^k(\mathbf{w}^k + d\mathbf{w}^k; f^k) + \alpha H_2^k(\mathbf{w}^k + d\mathbf{w}^k),$$

avec :

$$H_1^k \triangleq \sum_{s \in S^k} \rho_1 \left[\nabla f^k(s + dt\mathbf{w}_s^k, t + dt) \cdot d\mathbf{w}_s^k + f_t^k(s, t) \right],$$

$$f_t^k(s, t) \triangleq f^k(s, t) - f^k(s + dt\mathbf{w}_s^k, t + dt),$$

$$H_2^k \triangleq \sum_{\langle s, r \rangle \in C^k} \rho_2 \left[\left\| (\mathbf{w}_s^k + d\mathbf{w}_s^k) - (\mathbf{w}_r^k + d\mathbf{w}_r^k) \right\| \right].$$



\mathcal{C}^k est l'ensemble des paires de sites voisins de S^k selon un 4-voisinage. Le premier terme de l'énergie est un terme d'adéquation au modèle d'observation qui repose sur une version recalée de l'équation de contrainte du mouvement apparent (ECMA) [3, 6]. Le deuxième terme est un *a priori* contextuel sur la solution. Ces deux termes mettent en jeux des estimateurs ρ_* [1, 2, 4, 8] issus de la statistique robuste [9].

Ces fonctions, comparables aux Φ -fonctions introduites dans [4], permettent de prendre en compte les discontinuités spatiales ou les grandes déviations par rapport au modèle d'observation. Elles sont en fait, moins pénalisantes vis à vis des données "hors modèles" que les estimateurs quadratiques.

Il est possible de montrer [1, 2, 4] que l'optimisation de ces estimateurs peut être écrite sous la forme d'un problème dual d'optimisation :

$$\min_x \sum_i \rho(x_i) = \min_{x,z} \sum_i [z_i x_i^2 + \psi(z_i)],$$

mettant en jeux des variables auxiliaires z_i . L'avantage majeur de cette formulation est qu'elle conduit à un algorithme d'optimisation par *moindres carrés pondérés itérés* [8] :

- les z_i étant fixés, estimer les variables x_i revient à un problème classique de moindres carrés pondérés ;
- les x_i étant figés, les poids optimaux sont alors donnés par $z_i = \frac{\rho'(x_i)}{2x_i}$ [1, 2, 4].

Notons que la robustesse de l'estimateur tient au fait que $\rho(x)$ est choisi tel que $\phi(x) = \frac{\rho'(x)}{2x}$ soit strictement décroissante de \mathbb{R}^+ sur $[0, M \triangleq \phi(0) < +\infty]$. Plus la valeur de x est importante plus le poids optimal correspondant décroît (l'influence de x devenant dès lors de plus en plus faible).

Dans notre cas, les variables auxiliaires sont soit des variables auxiliaires d'observation (en rapport avec la formulation duale de H_1^k), soit des variables auxiliaires de discontinuité (découlant de la formulation duale de H_2^k). Les premières que nous noterons $\delta^k = \{\delta_s^k, s \in S^k\}$ atténuent les effets des régions où l'ECMA n'est pas valide (régions d'occlusion, régions fortement texturées, gradients photométriques importants,...). Les secondes, notées $\beta^k = \{\beta_{sr}^k, \langle s, r \rangle \in \mathcal{C}^k\}$ marquent les discontinuités spatiales du champ de déplacement. Elles sont situées sur la grille duale (interstitielle) de S^k et empêchent un effet de lissage spatiale à travers les frontières de mouvement.

Les deux termes de la fonction d'énergie duale au niveau k s'écrivent¹ :

$$\begin{aligned} \bullet \mathcal{H}_1^k(\mathbf{w}^k + d\mathbf{w}^k, \delta^k; f^k) &\triangleq \sum_{s \in S^k} [M_1 \delta_s^k \\ & \left[\nabla f^k(s + dt\mathbf{w}_s^k, t + dt) \cdot d\mathbf{w}_s^k + f_t^k(s, t) \right]^2 + \psi_1(M_1 \delta_s^k)], \\ \bullet \mathcal{H}_2^k(\mathbf{w}^k + d\mathbf{w}^k, \beta^k) &\triangleq \sum_{\langle s, r \rangle \in \mathcal{C}^k} \left[M_2 \beta_{sr}^k \left\| (\mathbf{w}_s^k + d\mathbf{w}_s^k) \right. \right. \\ & \left. \left. - (\mathbf{w}_r^k + d\mathbf{w}_r^k) \right\|^2 + \psi_2(M_2 \beta_{sr}^k) \right]. \end{aligned}$$

Même ainsi formulée, l'optimisation reste difficile. Pour cette raison l'approche multigrille proposée dans [7] a été mise en œuvre à chaque niveau de résolution.

1. Les poids sont ici ramenés dans $[0, 1]$

3. OPTIMISATION MULTIGRILLE DÉTERMINISTE

Le schéma d'optimisation que nous considérons dans cette section consiste en une exploration hiérarchique contrainte de l'espace des configurations Ω^k en termes de sous-espaces emboîtés :

$$\Omega^{k,L} \subset \Omega^{k,L-1} \subset \dots \subset \Omega^{k,1} \subset \Omega^{k,0} = \Omega^k. \quad (1)$$

où $\Omega^{k,l}$ est l'ensemble des champs d'incrément constants sur des blocs de taille $2^l \times 2^l$ partitionnant la grille S^k . Notons $\mathcal{B}^{k,l} \triangleq \{\mathcal{B}_n^{k,l}, n = 1, \dots, N_{k,l}\}$ cette partition, avec $N_{k,l} = |S^k|/4^l = |S|/4^{l+k}$. Les configurations "contraintes" de $\Omega^{k,l}$ sont en fait équivalentes à des champs réduits d'incrément définis sur la grille réduite, $S^{k,l} \triangleq \{1, \dots, N_{k,l}\}$, associée à la partition $\mathcal{B}^{k,l}$. Notons $\Gamma^{k,l}$ cet ensemble de configurations réduites et $\Phi^{k,l}$ la bijection de $\Gamma^{k,l}$ dans $\Omega^{k,l}$. L'optimisation dans $\Omega^{k,l}$ est équivalente à la minimisation d'une nouvelle fonction d'énergie :

$$\begin{aligned} \mathcal{H}^{k,l}(\mathbf{w}^k, d\mathbf{w}^{k,l}, \delta^k, \beta^k; f^k) &\triangleq \\ \mathcal{H}^k(\mathbf{w}^k + \Phi^{k,l}(d\mathbf{w}^{k,l}), \delta^k, \beta^k; f^k) &\quad (2) \end{aligned}$$

définie sur $\Gamma^{k,l}$ (les variables auxiliaires et les données restant définies sur la grille S^k).

Le problème d'optimisation se ramène alors à chaque niveau de résolution à une séquence d'optimisations sur des grilles successivement réduites :

$$\operatorname{argmin}_{\substack{d\mathbf{w}^{k,l} \in \Gamma^{k,l} \\ \delta^k, \beta^k}} \mathcal{H}^{k,l}(\mathbf{w}^k, d\mathbf{w}^{k,l}, \delta^k, \beta^k; f^k), \quad l = L, \dots, 0. \quad (3)$$

Ces différentes optimisations sont effectuées selon une stratégie descendante (*coarse-to-fine*) par l'algorithme des moindres carrés pondérés itérés. L'estimé final au niveau $l+1$ est projeté au niveau l (de façon naturelle par $[\Phi^{k,l}]^{-1} \circ \Phi^{k,l+1}$) et est utilisé comme initialisation de la relaxation déterministe à ce niveau. Ce schéma est itéré jusqu'à ce que le niveau le plus fin soit atteint (se référer à [7] pour de plus amples détails).

3.1. Dérivation de l'énergie multigrilles

Il est possible de montrer que l'énergie au niveau l comporte deux termes similaires à ceux de \mathcal{H}^k : un terme d'adéquation au modèle d'observation et un terme contextuel. L'expression de ces deux termes est donnée ci-dessous pour une résolution k donnée. Par soucis de lisibilité cet indice sera omis. Notons que les calculs intermédiaires peuvent être trouvés dans [11].

□ **Terme d'adéquation des données** : en se définissant les matrices suivantes :

$$F_{xy}^l(n)^T \triangleq \begin{bmatrix} \nabla f(s_1 + dt\mathbf{w}_{s_1}, t + dt) \\ \vdots \\ \nabla f(s_{4^l} + dt\mathbf{w}_{s_{4^l}}, t + dt) \end{bmatrix} = \begin{bmatrix} \mathbf{F}_x^l(n)^T \\ \mathbf{F}_y^l(n)^T \end{bmatrix},$$

$\mathbf{F}_t^l(n) \triangleq [f_t(s_1, t) \dots f_t(s_{4^l}, t)]^T$ ainsi que $\Delta_n^l(\delta) \triangleq \operatorname{diag}(\delta_{s_1}, \dots, \delta_{s_{4^l}})$, ce terme s'écrit sous la forme suivante :

$$\begin{aligned} \mathcal{H}_1^l(\mathbf{w}, d\mathbf{w}^l, \delta; f) &= \\ \sum_{n \in S^l} \left[M_1 \|F_{xy}^l(n) d\mathbf{w}_n^l + \mathbf{F}_t^l(n)\|_n^2 + \Psi_1^l(\delta_{\mathcal{B}_n^l}) \right], \end{aligned}$$

avec la notation $\|\mathbf{X}\|_n^2 \triangleq \mathbf{X}^T \Delta_n^l(\delta) \mathbf{X}$ et $\Psi_1^l(\delta_{B_n^l}) \triangleq \sum_{s \in B_n^l} \psi_1(M_1 \delta_s)$.

□ **Terme contextuel:** nous noterons $C_n^l \triangleq \{ \langle s, r \rangle \in C : \langle s, r \rangle \subset B_n^l \}$ l'ensemble des paires de sites voisins incluses dans le bloc B_n^l et $C_{n,m}^l \triangleq \{ \langle s, r \rangle \in C : s \in B_n^l, r \in B_m^l \}$ l'ensemble des paires "à cheval" sur les blocs B_n^l et B_m^l . L'ensemble C^l désignera l'ensemble des paires de sites voisins de la grille réduite S^l munie d'un 4-voisinage. Le terme contextuel est alors :

$$\mathcal{H}_2^l(\mathbf{w}, d\mathbf{w}^l, \beta) = \mathcal{H}_2(\mathbf{w}, \beta) + M_2 \sum_{\langle n, m \rangle \in C^l} \left[\|\mathbf{d}\mathbf{w}_n^l - \mathbf{d}\mathbf{w}_m^l\|^2 B_{n,m}^l(\beta) + 2(\mathbf{d}\mathbf{w}_n^l - \mathbf{d}\mathbf{w}_m^l) \cdot \mathbf{W}_{n,m}^l(\beta) \right],$$

avec :

$$\mathbf{W}_{n,m}^l(\beta) \triangleq \sum_{\langle s, r \rangle \in C_{n,m}^l} \beta_{sr} (\mathbf{w}_s - \mathbf{w}_r) = [U_{n,m}^l(\beta), V_{n,m}^l(\beta)]^T$$

et $B_{n,m}^l(\beta) \triangleq \sum_{\langle s, r \rangle \in C_{n,m}^l} \beta_{sr}$.

3.2. Minimization de l'énergie

À chaque niveau l , la minimisation est effectuée par l'algorithme des moindres carrés pondérés itérés. Cet algorithme consiste en une minimisation alternée des différents champs de variables. Dans un premier temps, le champ $d\mathbf{w}^l$ étant figé, les valeurs optimales des variables auxiliaires sont données explicitement par $z^* = \frac{\rho'(x^l)}{2Mx^l}$. Une fois les poids calculés et figés, la fonction \mathcal{H}^l est minimiser selon $d\mathbf{w}^l$. La minimisation de cette fonction quadratique est équivalente à la résolution d'un système linéaire. Un algorithme de relaxation du type Gauss-Seidel conduit à des solutions itératives. Ces solutions sont précisément décrites dans [11].

4. RÉSULTATS EXPÉRIMENTAUX

Nous présentons ici des résultats obtenus sur la séquence "Mobi" (Fig. 1). Cette séquence TV présente des mouvements de grande amplitude générés par un panoramique de la caméra et différents objets mobiles. Les estimateurs robustes choisis sont respectivement celui de Leclerc [10] pour le modèle d'observation et celui de Geman & McClure pour le terme contextuel [5] :

$$\rho_1(x) = 1 - \exp\left(-\frac{x^2}{\sigma_1}\right), \quad \rho_2(x) = \frac{x^2}{\sigma_2 + x^2}.$$

L'estimation a été menée pour 3 niveaux de résolution et 5 niveaux de grilles. Les mêmes paramètres ont été fixés aux différents niveaux de résolution et de grilles (le paramètre de régularisation $\alpha \frac{M_2}{M_1} = 300$ et $\sigma_1 = 35$, $\sigma_2 = 0.15$).

La figure 1 représente les champs de déplacement et les discontinuités spatiales seuillées ($seuil = 0.95$) estimés aux différents niveaux de résolution pour plusieurs niveaux de grilles.

Les résultats obtenus sont prometteurs. Visuellement le champ estimé est proche du mouvement apparent réel et la position des discontinuités spatiales s'adapte bien aux frontières des objets en mouvement. L'algorithme multigrille permet de converger rapidement vers des solutions de bonne qualité. Nous pouvons remarquer que ce modèle permet, en outre, d'extraire aux niveaux grossiers les grandes structures des frontières du mouvement apparent.

5. CONCLUSION

Dans cet article, nous avons présenté un estimateur de mouvement multirésolution. L'utilisation de cet estimateur conduit à un ensemble multirésolution de problèmes d'optimisation. Chacun est réexprimé en un problème dual avec variables auxiliaires qui peut être appréhendé par un algorithme déterministe de moindres carrés pondérés itérés. Les performances de cet algorithme sont accrues par l'utilisation de techniques multigrilles associées à une exploration hiérarchique contrainte des espaces de configurations. Une convergence rapide vers des estimées de bonne qualité est ainsi obtenue.

D'autre part cette méthode permet d'extraire à un faible coût, au niveau des grilles grossières, les grandes structures de discontinuité du flot optique. Ces structures devraient pouvoir être utilisées plus avant pour guider une caractérisation plus globale des frontières de mouvement.

6. RÉFÉRENCES

- [1] G. AUBERT, M. BARLAUD, L. BLANC-FÉRAUD, and P. CHARBONNIER. *Deterministic edge-preserving regularization in computed imaging*. Technical Report 94-01, I3S, Université de Nice-Sophia Antipolis, January 1994.
- [2] M.J. BLACK. *Robust incremental optical flow*. PhD thesis, Yale university, 1992.
- [3] W. ENKELMANN. Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences. *Computer Vision, Graphics, Image Processing*, 43:150-177, 1988.
- [4] D. GEMAN and G. REYNOLDS. Constrained restoration and the recovery of discontinuities. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(3):367-383, March 1993.
- [5] S. GEMAN and D.E. McCLURE. Statistical methods for tomographic image reconstruction. In *Bull. ISI, Proc. 46th Session Int. Statistical Institute. Vol. 52.*, 1987.
- [6] F. HEITZ and P. BOUTHEMY. Multimodal estimation of discontinuous optical flow using Markov Random Fields. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 15(12):1217-1232, Dec. 1993.
- [7] F. HEITZ, P. PÉREZ, and P. BOUTHEMY. A multiscale minimization of global energy functions in some visual recovery problems. *CVGIP: Image Understanding*, 59(1):125-134, 1994.
- [8] P.W. HOLLAND and R.E. WELSCH. Robust regression using iteratively reweighted least squares. *Com. Stat. Theor. Meth.*, A6:813-826, 1977.
- [9] P.J. HUBER. *Robust Statistics*. Wiley, 1981.
- [10] Y. G. LECLERC. Constructing simple stable descriptions for image partitioning. *Int. Journ. of Computer Vision*, 3:73-102, 1989.
- [11] E. MÉMIN and P. PÉREZ. Multiresolution Markov random field and multigrid algorithm for discontinuity preserving estimation of the optical flow. In *SPIE conf. on Neural, Morphological and Stochastic Methods in Image and Signal Processing*, 1995.

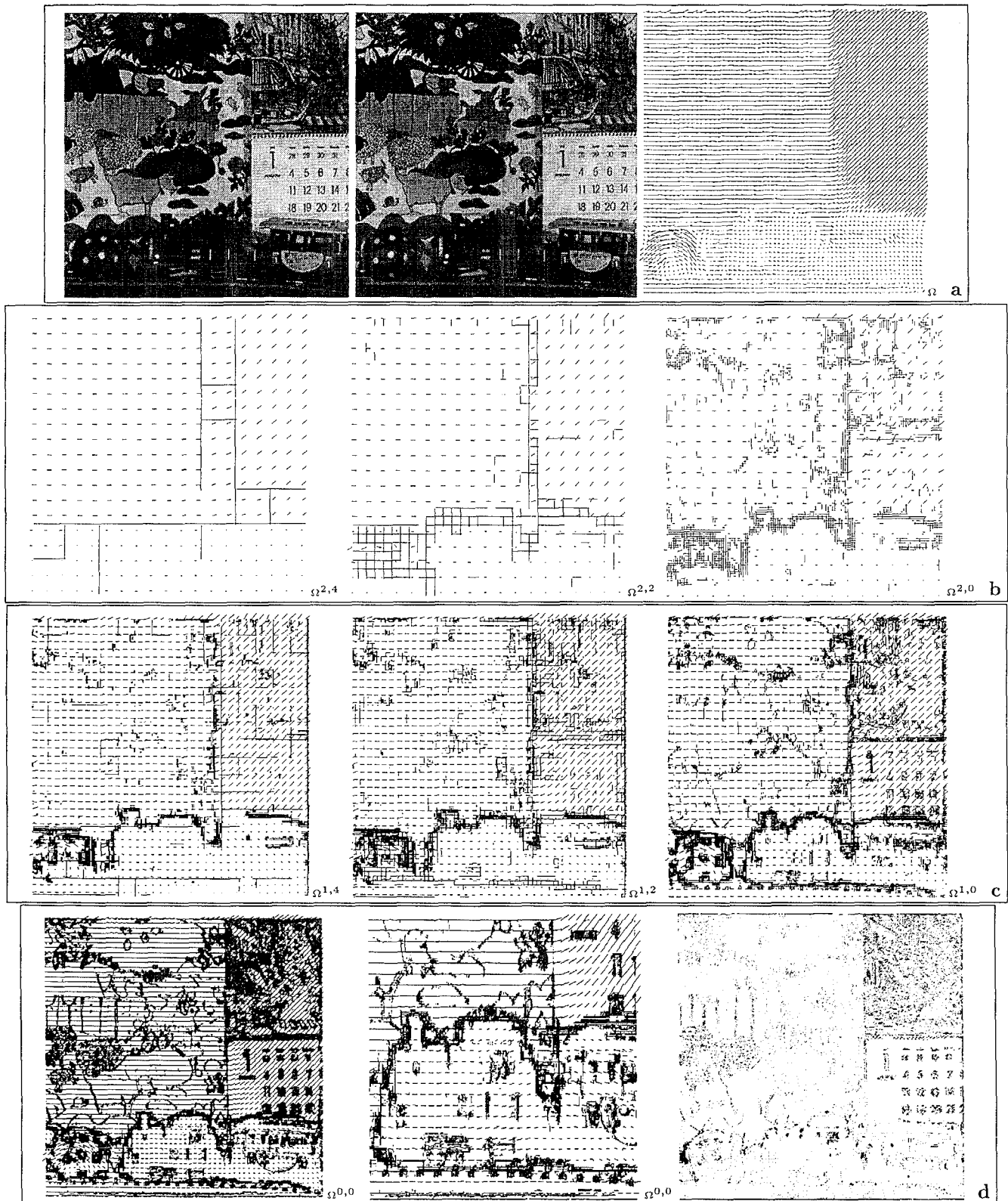


Figure 1: (a) deux images de la séquence "Mobi" (512×512) et le champ de vecteurs final pour Ω – (b) et (c) discontinuités spatiales et champs de vecteurs pour les niveaux de résolution k ($k = 2$ [Fig. b], $k = 1$ [Fig. c]) et pour les sous-espaces $\Omega^{k,l}$, $l = 4, 2, 0$ – (d) discontinuités spatiales et champ de vecteurs estimé pour Ω ($k = 0, l = 0$), zoom sur le train et image des variables d'observation pour $k = 0, l = 0$.