

## Analyse de séquences d'images, Flot Optique

Pierre Maurel

Empenn, IRISA/INRIA

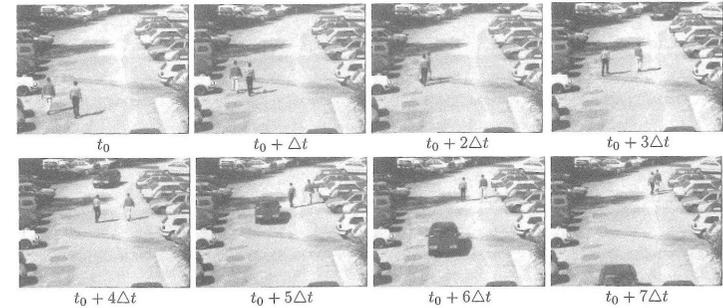
pierre.maurel@irisa.fr

[https://perso.univ-rennes1.fr/pierre.maurel/Prepa\\_Agreg/](https://perso.univ-rennes1.fr/pierre.maurel/Prepa_Agreg/)

1 / 27

## Séquences d'images, vidéo

- mêmes difficultés que les images statiques
- plus complexe : dimension temporelle, beaucoup d'informations (mouvement, profondeur, différents objets)

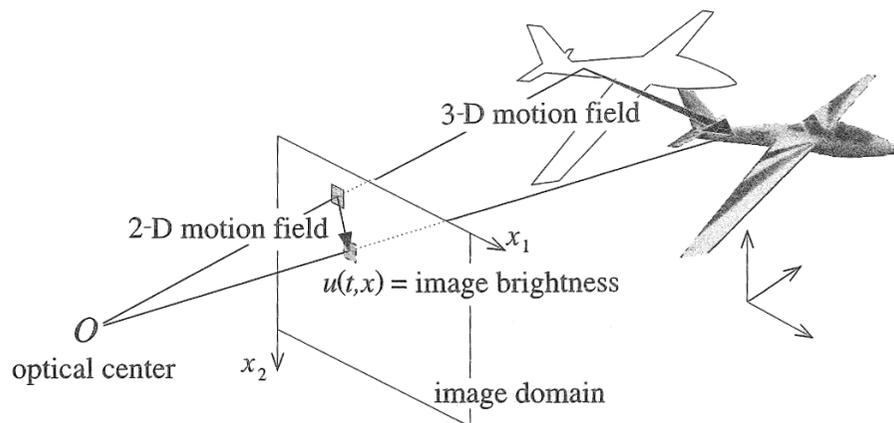


- arrière plan statique, trois objets en mvt, rôle important de l'échantillonnage temporel (nb images/sec)
- nature des mvts différents : voiture → trajectoire lisse, gens → chgts de direction, buisson → "aléatoire"
- qu'est ce qu'un mouvement ? → ombres ?
- profondeur : premier-plan, arrière-plan. Si caméra bouge : plus d'infos

3 / 27

## Séquences d'images, vidéo

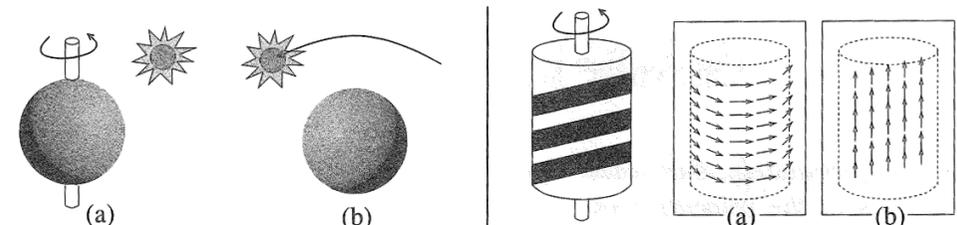
Dans une séquence d'images, les mouvements qu'on observe ne sont **qu'une projection** des mouvements réels



4 / 27

## Flot optique

- en général : impossible de mesurer le mouvement 2D ← projection sur le plan de l'image du mouvement 3D de la scène
- on perçoit uniquement le mouvement apparent, aussi appelé "flot optique"
- "apparent" → observable par des variations d'intensité dans l'image

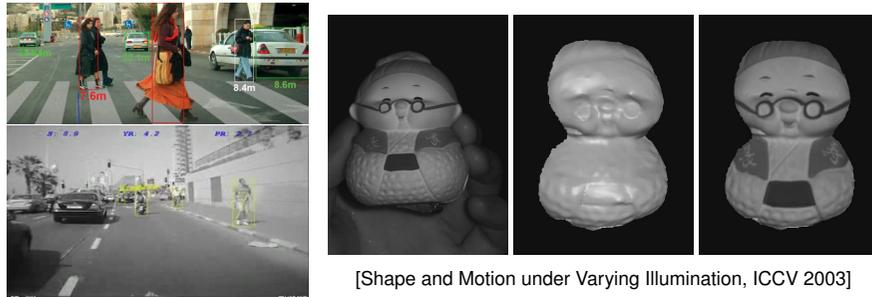


Cas où le mouvement réel est différent du mouvement apparent

6 / 27

## Flot optique

- quantitativement différents, mais partagent souvent les mêmes propriétés qualitatives (exple : mouvement des discontinuités)
- obtenir le flot optique → très utile : segmentation, détection de mouvement spécifique, anticipation, compression vidéo ...



[Shape and Motion under Varying Illumination, ICCV 2003]

[L.Zhang, B.Curlless, A. Hertzmann, S.Seitz]

[www.mobileye.com]

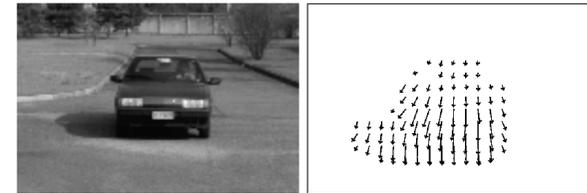
7 / 27

## Flot optique

- On cherche un champ de vecteurs représentant le mouvement de chaque point :

$$\sigma(t, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$(x, y) \mapsto (\sigma_1(t, x, y), \sigma_2(t, x, y)) \quad (1)$$



- un objet à la position  $(x, y)$  à l'instant  $t$  est à la position  $(x + dt \sigma_1(t, x, y), y + dt \sigma_2(t, x, y))$  à l'instant  $t + dt$
- permet de "suivre" un objet d'une image à l'autre
- donne un champ de déformation (défini en tout point) pour déformer une image en la suivante

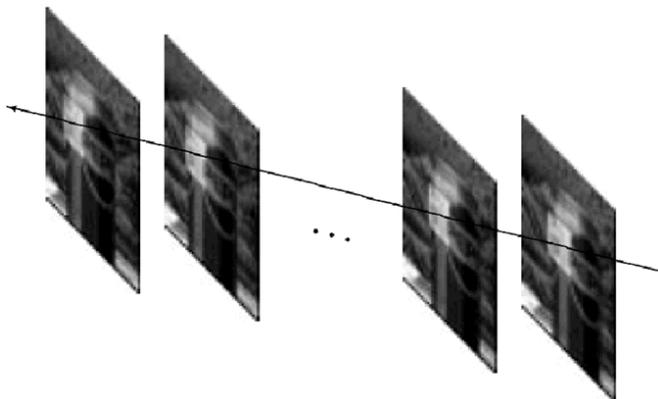
8 / 27

## Flot optique

### Principe de base

L'intensité d'un point reste constante le long de sa trajectoire, on cherche donc  $\sigma$  tel que :

$$u(t, x, y) = u(t + \Delta t, x + \Delta t \sigma_1(t, x, y), y + \Delta t \sigma_2(t, x, y))$$



9 / 27

## Flot optique

### Principe de base

L'intensité d'un point reste constante le long de sa trajectoire, on cherche donc  $\sigma$  tel que :

$$u(t, x, y) = u(t + \Delta t, x + \Delta t \sigma_1(t, x, y), y + \Delta t \sigma_2(t, x, y))$$

- hypothèse raisonnable :
  - pour de petits déplacements (changements d'illuminations faibles)
  - tant qu'il n'y a pas d'occlusions
- développement de Taylor au premier ordre :

$$u(t + dt, x + \Delta t \sigma_1(t, x, y), y + \Delta t \sigma_2(t, x, y)) =$$

$$u(t, x, y) + \Delta t \frac{\partial u}{\partial t} + \Delta t \sigma_1(t, x, y) \frac{\partial u}{\partial x} + \Delta t \sigma_2(t, x, y) \frac{\partial u}{\partial y} + \mathcal{O}(\Delta t^2)$$

On cherche donc  $\sigma$  qui vérifie :  $\frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle = 0$

10 / 27

## Flot optique, exercice

### Contrainte du Flot Optique

$$\frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle = 0$$

- Exemple : Soit  $u_0$  une image :  $u_0 : \mathbb{R}^2 \rightarrow \mathbb{R}$

On définit la séquence d'images  $u(t, x, y)$  par :

- $$\begin{cases} u(t + \Delta t, x, y) = u(t, x - 2\Delta t, y - \Delta t) \\ u(0, x, y) = u_0(x, y) \end{cases}$$
- $$\frac{\partial u}{\partial t} = -2\frac{\partial u}{\partial x} - \frac{\partial u}{\partial y} + \mathcal{O}(\Delta t)$$
- contrainte du flot optique (au 1er ordre) :  $(\sigma_1 - 2)\frac{\partial u}{\partial x} + (\sigma_2 - 1)\frac{\partial u}{\partial y} = 0$
- solution possible :  $\sigma = (2, 1)$  mais aussi  $\sigma = (2 + \alpha \frac{\partial u}{\partial y}, 1 - \alpha \frac{\partial u}{\partial x}), \forall \alpha$

correspond à une translation



11/27

## Flot optique

### Contrainte du Flot Optique

$$\frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle = 0$$

- $\frac{\partial u}{\partial t}$  : variation temporelle de l'intensité de l'image → **connu** (estimé)
- $\nabla u$  : variation spatiale de l'intensité de l'image → **connu** (estimé)
- $\sigma = (\sigma_1, \sigma_2)$  : flot optique → **deux inconnues**
- Problème : une seule équation, deux inconnues  $\sigma_1, \sigma_2$  → **problème mal posé**

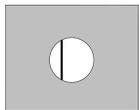
12/27

## Flot optique

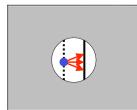
### Contrainte du Flot Optique

$$\frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle = 0$$

- "problème d'ouverture"

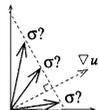


instant  $t$



instant  $t + dt$

- on ne connaît en fait que la composante du flot optique dans la direction de  $\nabla u$  :  $\langle \nabla u, \sigma \rangle = -\frac{\partial u}{\partial t}$ , Flot Normal

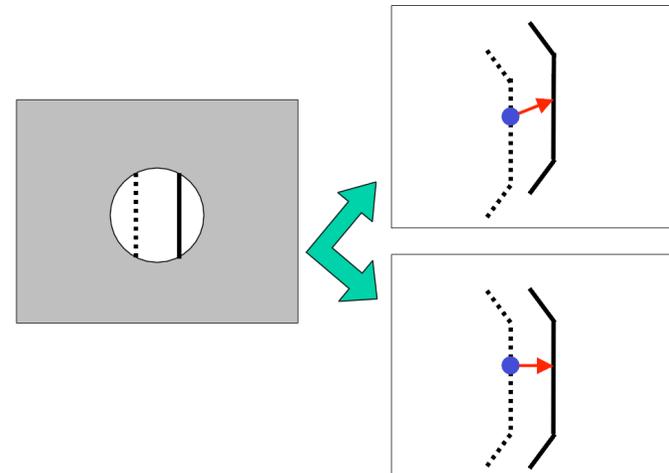


→ **besoin de contraintes supplémentaires**

13/27

## Problème d'ouverture

Agrandir "l'ouverture" peut, dans certains cas, aider :

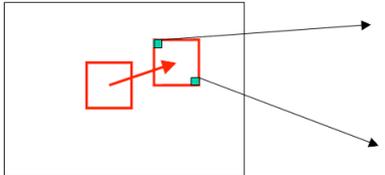


14/27

## Problème d'ouverture

### Lucas, Kanade (1981) :

On suppose  $\sigma$  constant sur un petit voisinage autour d'un point  $(x, y)$  :



$$\begin{cases} \frac{\partial u}{\partial t}(x_1, y_1) + \langle \nabla u(x_1, y_1), \sigma \rangle = 0 \\ \frac{\partial u}{\partial t}(x_2, y_2) + \langle \nabla u(x_2, y_2), \sigma \rangle = 0 \\ \dots \end{cases} \quad (2)$$

$$\underbrace{A}_{n \times 2} \underbrace{\sigma}_{2 \times 1} = \underbrace{b}_{n \times 1}$$

Problèmes :

- plus d'équations que d'inconnues : méthode des moindres carrés, on minimise  $\|A\sigma - b\|^2$
- choix du voisinage ? si trop petit : inutile. si trop grand : pas de sens.
- cas des zones uniformes

15/27

## Régularisation

### Horn, Shunk (1981) :

- On a vu qu'il fallait des contraintes supplémentaires. De la même manière qu'en restauration d'images on peut considérer le problème de minimisation (multiplicateur de Lagrange) :

$$\inf_{\sigma} J(\sigma) = \inf_{\sigma} \underbrace{\int_{\Omega} \left( \frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle \right)^2 d\Omega}_{\text{terme de fidélité}} + \lambda \left( \underbrace{\int_{\Omega} \|\nabla \sigma_1\|^2 d\Omega + \int_{\Omega} \|\nabla \sigma_2\|^2 d\Omega}_{\text{terme de régularisation}} \right)$$

- Même "problème" qu'en restauration d'images : régularisation quadratique trop pénalisante pour discontinuités
- Or les discontinuités du flot optique sont fondamentales pour l'analyse de séquences : détection d'objets, etc ...
- comme en restauration, on peut changer le terme de régularisation en :

$$\int_{\Omega} \Psi(\|\nabla \sigma_1\|) d\Omega + \int_{\Omega} \Psi(\|\nabla \sigma_2\|) d\Omega$$

16/27

## Régularisation de Horn et Shunk

- équation d'Euler-Lagrange :

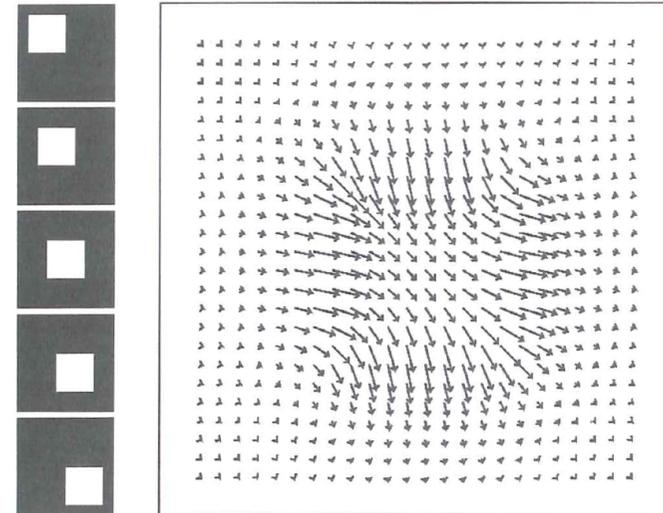
$$\begin{cases} \left( \frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle \right) \frac{\partial u}{\partial x} - \lambda \Delta \sigma_1 = 0 \\ \left( \frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle \right) \frac{\partial u}{\partial y} - \lambda \Delta \sigma_2 = 0 \end{cases} \quad (3)$$

- descente de gradient :

$$\begin{cases} \frac{\partial \sigma_1}{\partial t} = \lambda \Delta \sigma_1 - \left( \frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle \right) \frac{\partial u}{\partial x} \\ \frac{\partial \sigma_2}{\partial t} = \lambda \Delta \sigma_2 - \left( \frac{\partial u}{\partial t} + \langle \nabla u, \sigma \rangle \right) \frac{\partial u}{\partial y} \end{cases} \quad (4)$$

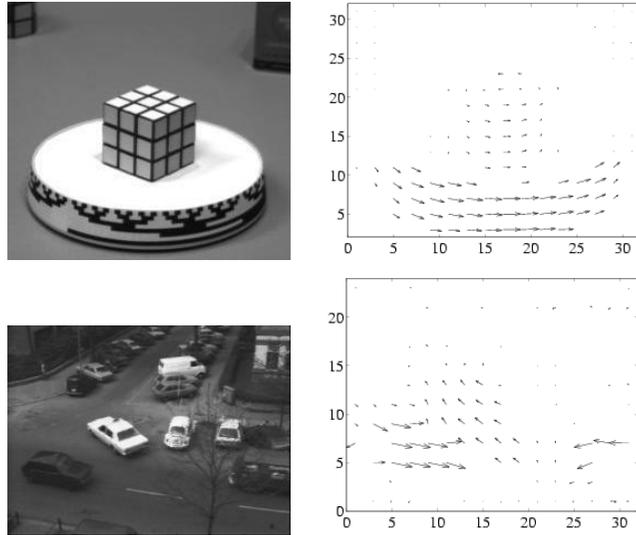
17/27

## Régularisation de Horn et Shunk : résultats



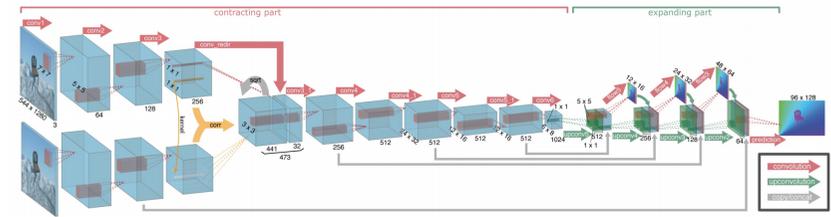
18/27

## Régularisation de Horn et Shunk : résultats



19 / 27

## Estimation du flot optique : Deep Learning



**Fig. 3 FlowNetC architecture** from [Dosovitskiy et al \(2015\)](#). The figure shows the full 15-layer architecture with contracting and expanding part. The gray skip connections between both parts allow for high-resolution predictions without passing high-resolution information through the bottleneck. We use the same architecture with fewer feature channels in our optical flow experiments. The DispNetCorr1D from [Mayer et al \(2016\)](#) (which we use for our disparity experiments in Section 5.4) implements the same idea, but with only a 1D correlation.

*What Makes Good Synthetic Training Data for Learning Disparity and Optical Flow Estimation?*, Mayer et al. 2018

<https://www.youtube.com/watch?v=g-peWxaQnQc>

20 / 27

## Exemple d'application



- but : reconnaître une action humaine à basse résolution (une personne = 30 pixels)

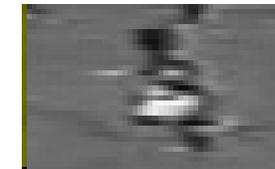


22 / 27

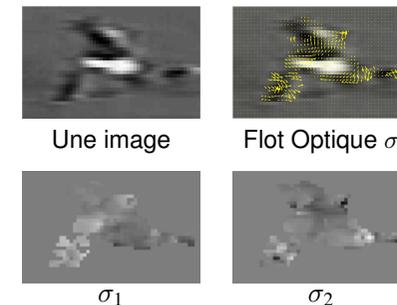
## Exemple d'application

"Recognizing Action at a Distance" [A.A. Efros, A.C. Berg, G. Mori, J. Malik]

- donnée :



- méthode basée sur le flot optique → plus d'informations que dans chaque image



23 / 27

## Exemple d'application



"Recognizing Action at a Distance" [A.A. Efros, A.C. Berg, G. Mori, J. Malik]

24 / 27

## Exemple d'application



"Recognizing Action at a Distance" [A.A. Efros, A.C. Berg, G. Mori, J. Malik]

25 / 27

## Exemple d'application

Conduite automatique (drones, voitures, ...)

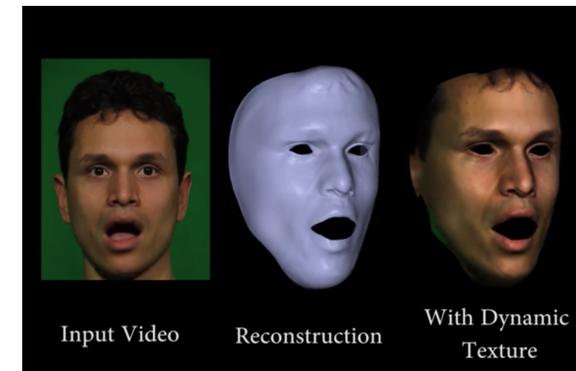


tesla

Attacking Optical Flow : <https://www.youtube.com/watch?v=FV-oH1aIdAI>

26 / 27

## Exemple d'application



<https://www.youtube.com/watch?v=2K6eZd3Ywss>

P. Garrido, L. Valgaerts, G. Wu, C. Theobalt, Reconstructing Detailed Dynamic Face Geometry from Monocular Video, In ACM Transactions on Graphics (Proc. of SIGGRAPH Asia)

27 / 27