

Estimation d'un mouvement de caméra

Georges Koepfler

Université Paris Descartes

MVA 2010-2011

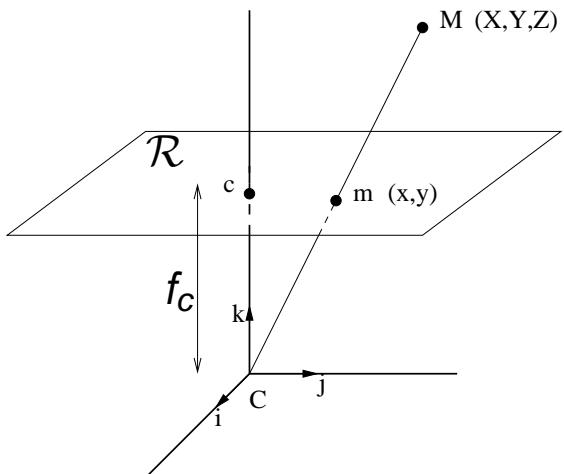
Plan :

1 Introduction

2 Modélisation

3 Approximations

4 Résultats



Caméra en C , axe optique k ;

$M = (X, Y, Z)$ dans (C, i, j, k) ;

(c, i, j) repère de \mathcal{R} ;

$m = (x, y)$ projection de M sur \mathcal{R} ;

$$x = f_c \frac{X}{Z}, \quad y = f_c \frac{Y}{Z}$$

Matrice de projection

M a comme coordonnées projectives

$$(X : Y : Z : 1) = (\lambda X : \lambda Y : \lambda Z : \lambda) ;$$

m a comme coordonnées projectives

$$(x : y : 1) = (\lambda x : \lambda y : \lambda) ;$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} f_c & 0 & 0 & 0 \\ 0 & f_c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathcal{P}} \begin{pmatrix} X/Z \\ Y/Z \\ 1 \\ 1/Z \end{pmatrix}$$

\mathcal{P} est la *matrice de projection* associée à la caméra :

$$m = \mathcal{P}M$$

- Paramètres *intrinsèques* : si le système de coordonnées de la caméra n'est pas (c, i, j) , mais (o, u, v) ,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} m_u & 0 \\ 0 & m_v \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u_o \\ v_o \end{pmatrix}.$$

- Paramètres *extrinsèques* : si le repère choisi dans \mathbb{R}^3 n'est pas (C, i, j, k) mais $(0, i', j', k')$, alors $M_0 = RM_C + t$ où R est une rotation et t une translation.

- Finalement

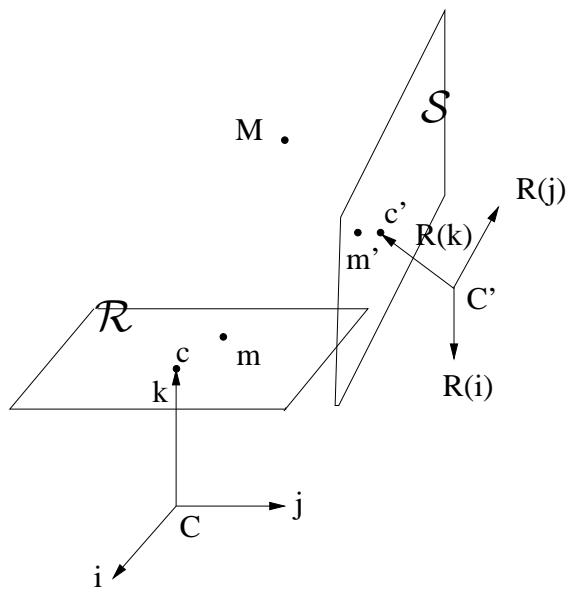
$$\mathcal{P} = \begin{pmatrix} m_u & 0 & u_o \\ 0 & m_v & v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_c & 0 & 0 & 0 \\ 0 & f_c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$

Calibrer une caméra consiste à déterminer la matrice \mathcal{P} .

Modélisation du mouvement de caméra

- Afin de simplifier, on ne va pas tenir compte des paramètres intrinsèques et extrinsèques ; en plus on suppose $f_c = 1$;
- Un déplacement rigide de \mathbb{R}^3 est une transformation D qui conserve les distances et les angles $D : \mathbb{R}^3 \rightarrow \mathbb{R}^3$
 $M \mapsto RM + t$
où R est une rotation et t une translation ;
- Composition de déplacements $D = D_1 \circ D_2$
pour $D_1 = (R_1, t_1)$ et $D_2 = (R_2, t_2)$:

$$D = D_1 \circ D_2 = (R_1 R_2, R_1 t_2 + t_1).$$



déplacement $D = (R, t)$,
 translation $t = \overrightarrow{CC'}$,
 rotation R , d'axe passant par C ,
 $M' = RM + t$.

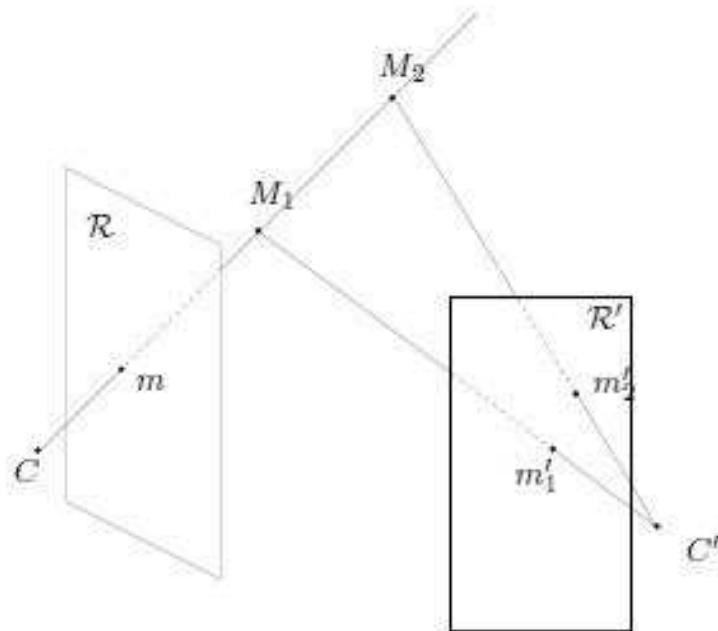
(C, i, j, k) transformé en
 $(C', R(i), R(j), R(k))$;

$S = D(R)$ de repère
 $(c', R(i), R(j))$
 où $\overrightarrow{C'c'} = R(k)$;

$m \in \mathcal{R}$ et $m' \in \mathcal{S}$ sont les projections de M

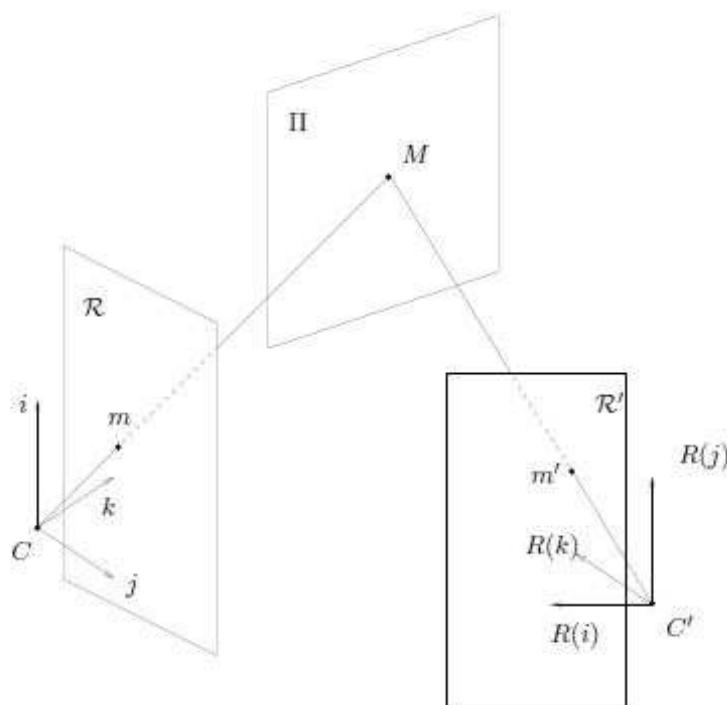
Influence sur les projections

- 1 Effet de parallaxe : la projection des points d'une *scène 3D fixe* donne lieu à des ambiguïtés du fait de la *profondeur*.
 Sauf
 - pour une scène plane ;
 - si la caméra a effectué une rotation pure.
- 2 Contrainte épipolaire : la projection m de M sur \mathcal{R} se trouve sur la ligne *ligne épipolaire*.
 C'est la droite passant par $e = \mathcal{R} \cap CC'$ et $\mathcal{R} \cap Cm'$.
 L'écriture algébrique de la contrainte épipolaire se fait grâce à la *matrice fondamentale*.

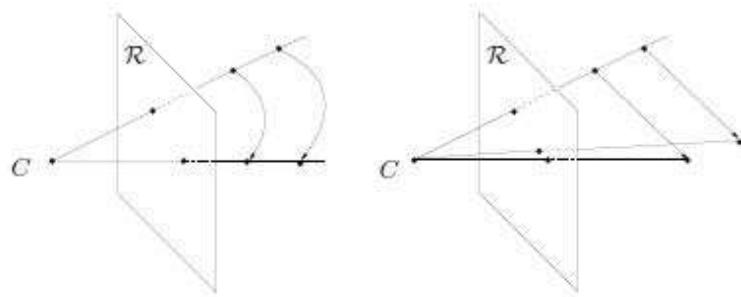


M_1 et M_2 ont le même point image dans \mathcal{R} .

Scène plane

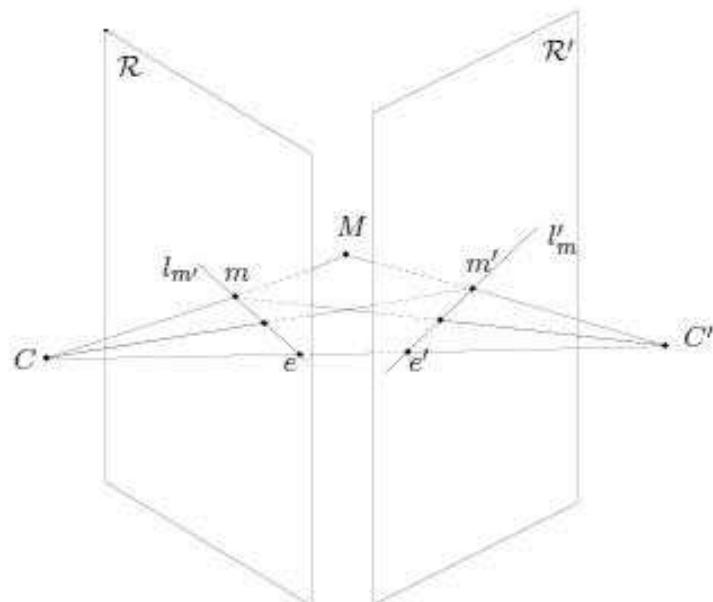


Le plan Π de l'espace ne passe pas par les centres optiques C et C' .
Il n'y a pas de phénomène d'occultation.



Après une rotation pure (à gauche) les points sont projetés sur le même point dans le plan image ;
après une translation pure (à droite) ont obtient deux points images.

Ligne épipolaire



La contrainte épipolaire.

Soit $\overrightarrow{CM} = Xi + Yj + Zk$,

$$R = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \text{ et } t = t_1i + t_2j + t_3k.$$

Comme $\overrightarrow{C'M} = X'R(i) + Y'R(j) + Z'R(k) = \overrightarrow{CC} + \overrightarrow{CM}$:

$$R \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = - \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} + \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

Or $m' = \left(\frac{X'}{Z'}, \frac{Y'}{Z'} \right) = (x', y')$ et $m = \left(\frac{X}{Z}, \frac{Y}{Z} \right) = (x, y)$.

Transformation plane (suite)

On obtient

$$x = \frac{X}{Z} = \frac{a_1x' + b_1y' + c_1 + t_1/Z'(x', y')}{a_3x' + b_3y' + c_3 + t_3/Z'(x', y')},$$

$$y = \frac{Y}{Z} = \frac{a_2x' + b_2y' + c_2 + t_2/Z'(x', y')}{a_3x' + b_3y' + c_3 + t_3/Z'(x', y')}$$

On définit ainsi une application

$$(x, y) = \varphi(x', y')$$

φ dépend des 6 paramètres de déplacement et de la profondeur $Z'(x', y')$.

Appariement de points

- L'image f est la restriction au plan \mathcal{R} : $\{Z = 1\}$ d'une scène 3D

$$F(X, Y, Z) = f\left(\frac{X}{Z}, \frac{Y}{Z}\right) = f(x, y) \quad \text{pour } Z \neq 0;$$

- L'image g est dans \mathcal{S} . Si $m' \in \mathcal{S}$ et $m \in \mathcal{R}$ sont deux points appariés alors

$$g(m') = f(m).$$

- D'où

$$g(x', y') = f(x, y) = f \circ \varphi(x', y')$$

$$= f\left(\frac{a_1x' + b_1y' + c_1 + t_1/Z'(x', y')}{a_3x' + b_3y' + c_3 + t_3/Z'(x', y')}, \frac{a_2x' + b_2y' + c_2 + t_2/Z'(x', y')}{a_3x' + b_3y' + c_3 + t_3/Z'(x', y')}\right)$$

$$f(x, y) = g(x', y') = g \circ \psi(x, y)$$

$$= g\left(\frac{a_1x + a_2y + a_3 - \langle t/Z(x, y), R(i) \rangle}{c_1x + c_2y + c_3 - \langle t/Z(x, y), R(k) \rangle}, \frac{b_1x + b_2y + b_3 - \langle t/Z(x, y), R(j) \rangle}{c_1x + c_2y + c_3 - \langle t/Z(x, y), R(k) \rangle}\right)$$

Approximation de la profondeur

Sous certaines hypothèses, on peut simplifier les équations qui définissent φ et ψ , grâce au résultat suivant.

Théorème

Pour $Z_{inf} \leq Z(x, y) \leq Z_{sup}$,

$$\text{si } \left| \frac{1}{Z_{inf}} - \frac{1}{Z_{sup}} \right| \|t\| C(L) \leq \varepsilon \text{ et } \frac{\|t\|}{Z_{inf}} C'(L) \leq \varepsilon$$

alors on peut remplacer $Z(x, y)$ par une constante Z_0 dans la définition de ψ avec une erreur bornée par ε .

De plus, $Z'(x', y')$ peut être remplacé par la même constante Z_0 dans φ .

Soit K , resp. K' le domaine rectangulaire borné où l'image f , resp g est définie ;
on note L la plus grande dimension de K et K' .

Hypothèses :

H1 : la variation de l'axe optique k et la translation dans la direction de k est faible, c.-à-d.

$$\left| \frac{1}{c_1 x + c_2 y + c_3 - \langle t/Z(x, y), R(k) \rangle} \right| \leq \frac{4}{3}.$$

H2 : on suppose que le flot optique est borné, c.-à-d. si $(x, y) = \varphi(x', y')$, resp. $(x', y') = \psi(x, y)$, alors

$$\max\{|x' - x|, |y' - y|\} \leq \frac{L}{2}.$$

Applications projectives

Sous l'**hypothèse de profondeur constante** $Z(x, y) = Z_0 = Z'(x', y')$, les déformations φ et ψ liant les images f et g sont des **applications projectives** sur \mathbb{R}^2 , associées aux matrices \mathcal{M}_φ et \mathcal{M}_ψ :

$$\mathcal{M}_\varphi = \begin{pmatrix} a_1 & b_1 & c_1 + \tilde{t}_1 \\ a_2 & b_2 & c_2 + \tilde{t}_2 \\ a_3 & b_3 & c_3 + \tilde{t}_3 \end{pmatrix} = R \begin{pmatrix} 1 & 0 & \langle \tilde{t}, R(i) \rangle \\ 0 & 1 & \langle \tilde{t}, R(j) \rangle \\ 0 & 0 & 1 + \langle \tilde{t}, R(k) \rangle \end{pmatrix} = RH$$

$$\mathcal{M}_\psi = \begin{pmatrix} a_1 & a_2 & a_3 - \langle \tilde{t}, R(i) \rangle \\ b_1 & b_2 & b_3 - \langle \tilde{t}, R(j) \rangle \\ c_1 & c_2 & c_3 - \langle \tilde{t}, R(k) \rangle \end{pmatrix} = R^{-1} \begin{pmatrix} 1 & 0 & -\tilde{t}_1 \\ 0 & 1 & -\tilde{t}_2 \\ 0 & 0 & 1 - \tilde{t}_3 \end{pmatrix} = R^{-1} \tilde{H}.$$

où $\tilde{t} = t/Z_0$.

Note : la décomposition RH , resp. $R^{-1}\tilde{H}$, est unique.

$$GP^2(\mathbb{R}) = \left\{ \phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \text{ telles que } \forall (x, y) \in \mathbb{R}^2, \right. \\ \left. \phi(x, y) = \left(\frac{\alpha_1 x + \beta_1 y + \gamma_1}{\alpha_3 x + \beta_3 y + \gamma_3}, \frac{\alpha_2 x + \beta_2 y + \gamma_2}{\alpha_3 x + \beta_3 y + \gamma_3} \right), \text{ avec } \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix} \neq 0 \right\}.$$

- 1 $GP^2(\mathbb{R})$ est un groupe pour la multiplication des matrices ;
- 2 l'application ϕ engendrée par une matrice M est identique à celle obtenue pour la matrice λM , pour tout réel non nul λ ;
- 3 un élément de $GP^2(\mathbb{R})$ dépend de 8 paramètres ;
- 4 un déplacement D ne dépendant que de 6 paramètres, il faut considérer un sous-ensemble de $GP^2(\mathbb{R})$;
- 5 ψ n'est pas l'inverse de φ dans le groupe projectif : $\mathcal{M}_\varphi^{-1} \neq \mathcal{M}_\psi$

Groupe des recalages

$$\mathcal{A} = \left\{ \varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \text{ telle que } \forall (x, y) \in \mathbb{R}^2 \right. \\ \left. \varphi(x, y) = \left(\frac{a_1 x + b_1 y + c_1 + A}{a_3 x + b_3 y + c_3 + C}, \frac{a_2 x + b_2 y + c_2 + B}{a_3 x + b_3 y + c_3 + C} \right), \right. \\ \left. \text{avec } R = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \in SO(3) \text{ et } t = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \in \mathbb{R}^3 \right\}.$$

$$\varphi \in \mathcal{A} \longleftrightarrow \text{déplacement } D = (R, t)$$

$$\varphi^{-1} \in \mathcal{A} \longleftrightarrow \text{déplacement } D^{-1} = (R^{-1}, -R^{-1}t)$$

(\mathcal{A}, \star) est le groupe des recalages :
la loi \star est déduite de la composition des mouvements rigides de \mathbb{R}^3 .

Groupe des recalages (suite)

- Soient φ_1 et φ_2 dans \mathcal{A} , associées respectivement aux déplacements $D_1 = (R_1, t_1)$ et $D_2 = (R_2, t_2)$. Alors $\varphi = \varphi_1 * \varphi_2$ est associée au déplacement $D = (R, t) = (R_1 R_2, t_1 + R_1 t_2)$ et

$$\mathcal{M}_\varphi = \mathcal{M}_{\varphi_1 * \varphi_2} = R_1 R_2 \begin{pmatrix} 1 & 0 & \langle \tilde{t}_1 + R_1 \tilde{t}_2, R_1 R_2(i) \rangle \\ 0 & 1 & \langle \tilde{t}_1 + R_1 \tilde{t}_2, R_1 R_2(j) \rangle \\ 0 & 0 & 1 + \langle \tilde{t}_1 + R_1 \tilde{t}_2, R_1 R_2(k) \rangle \end{pmatrix}$$

où $\tilde{t}_1 = t_1/Z_0$ et $\tilde{t}_2 = t_2/Z_0$.

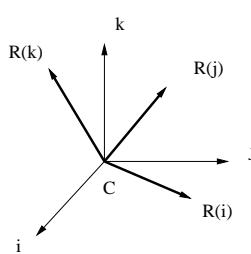
- Soit $f_1, f_2, f_3, \dots, f_n$ la séquence d'images. Si pour tout couple d'images successives, (f_i, f_{i+1}) , on connaît φ_i telle que $f_{i+1} = f_i \circ \varphi_i$ c.-à-d. $D_i = (R_i, t_i)$. Alors

$$f_n = f_1 \circ (\varphi_1 * \varphi_2 * \dots * \varphi_{n-1}) = f_1 \circ \varphi$$

où φ est l'application projective associée à $R = R_1 R_2 \dots R_{n-1}$ et $t = t_1 + R_1 t_2 + R_1 R_2 t_3 + \dots + R_1 R_2 \dots R_{n-2} t_{n-1}$.

Décomposition de la rotation

R



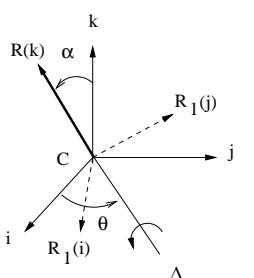
$$R = R_2 R_1$$

R_1 d'axe $\Delta \in (C, i, j)$
paramètres θ et α

$R_1 = R_{\theta, \alpha} = R_\theta^k R_\alpha^i R_{-\theta}^k$
déformation projective de f

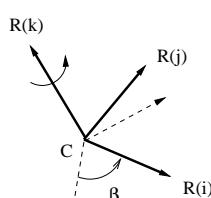
R_1

R_2



R_2 d'axe $R(k) = R_1(k)$
paramètre β

$R_2 = R_\beta = R_\theta^k R_\alpha^i R_\beta^k R_{-\alpha}^i R_{-\theta}^k$
rotation plane de $R_1(f)$



$$R = R_\beta R_{\theta, \alpha} = R_\theta^k R_\alpha^i R_\beta^k R_{-\theta}^k = R_\theta^k R_\alpha^i R_{-\theta}^k R_\beta^k = R_{\theta, \alpha} R_\beta^k.$$

Déplacement $D = (R, t) = (R_{\theta, \alpha} R_{\beta}^k, t)$ et $\tilde{t} = t/Z_0$.

La matrice \mathcal{M}_φ associée à φ s'écrit :

$$\begin{pmatrix} a_1 & b_1 & c_1 + \tilde{t}_1 \\ a_2 & b_2 & c_2 + \tilde{t}_2 \\ a_3 & b_3 & c_3 + \tilde{t}_3 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & < \tilde{t}, R(i) > \\ 0 & 1 & < \tilde{t}, R(j) > \\ 0 & 0 & 1 + < \tilde{t}, R(k) > \end{pmatrix} = R_{\theta, \alpha} R_{\beta}^k H,$$

$$(x, y) \in \mathbb{R}^2 : g(x', y') = f(\varphi(x', y')) = f(r_{\theta, \alpha} \circ s(x', y')).$$

- $r_{\theta, \alpha}$ transformation projective ;
- s similitude $R_{\beta}^k H$.

Paramètres du déplacement $(\theta, \alpha, \beta, A, B, C)$,
avec $(-A, -B, -C)$ coordonnées de \tilde{t} dans $(R(i), R(j), R(k))$.

Transformation des images (bis)

On a $g(x', y') = f(\varphi(x', y')) = f(x, y)$ et

$$\varphi(x', y') = \left(\frac{a_1 x' + b_1 y' + c_1 - \tilde{t}_1}{a_3 x' + b_3 y' + c_3 - \tilde{t}_3}, \frac{a_2 x' + b_2 y' + c_2 - \tilde{t}_2}{a_3 x' + b_3 y' + c_3 - \tilde{t}_3} \right).$$

Par ailleurs $f(x, y) = g(\psi(x, y)) = g(x', y')$ et

$$\psi(x, y) = \left(\frac{a_1 x + a_2 y + a_3 + A}{c_1 x + c_2 y + c_3 + C}, \frac{b_1 x + b_2 y + b_3 + B}{c_1 x + c_2 y + c_3 + C} \right).$$

On utilise les paramètres :

$$\tilde{t} = -A R(i) - B R(j) - C R(k) \text{ et } R = R_{\theta, \alpha} R_{\beta}^k.$$

paramètre	θ (radians)	α (radians)	β (radians)	A, B (long.f.)	C (long.f.)
valeurs	$]-\pi, \pi]$	$[0, 0.03]$	$[-0.05, 0.05]$	$[-0.09, 0.09]$	$[-0.03, 0.03]$

Pour α, β, A, B, C et L petits, le flot optique est approximé par trois termes indépendants :

$$\left\{ \begin{array}{l} x' - x \simeq -Cx + A + \beta y + \alpha x (\cos \theta y - \sin \theta x) \\ y' - y \simeq -Cy + B - \beta x + \alpha y (\cos \theta y - \sin \theta x) \end{array} \right.$$

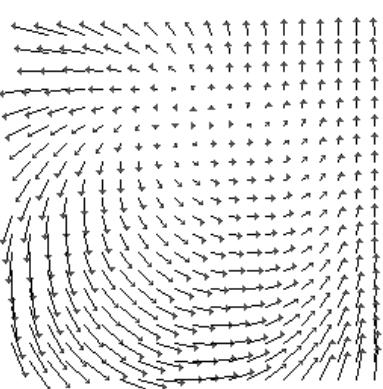
$\underbrace{}_{\tilde{t}}$
 $\underbrace{}_{R_\beta^k}$
 $\underbrace{}_{R_{\theta, \alpha}}$

Pour x et y petits, on peut négliger le terme quadratique : au centre de l'image on a une transformation affine.

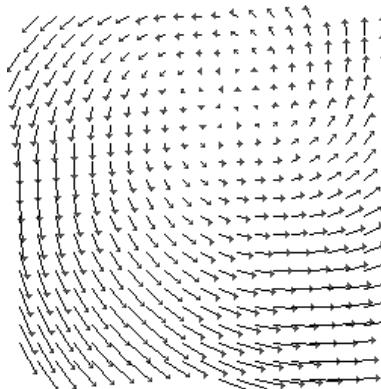
Exemple

$$\left\{ \begin{array}{l} x' - x \simeq -Cx + A + \beta y \\ y' - y \simeq -Cy + B - \beta x \end{array} \right.$$

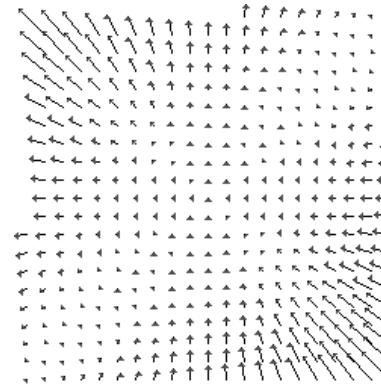
La similitude agit seule sur le centre de l'image.



flot du mouvement



similitude s



rotation $r_{\theta, \alpha}$

Paramètres : $\theta = 3\pi/4, \alpha = 5 \cdot 10^{-4}, \beta = 5, 1 \cdot 10^{-2}, A = 3, B = 2, \lambda = 0, 02$.

Estimation du flot optique

Algorithme **Motion2D** de BOUTHÉMY, ODOBEZ pour estimer les paramètres d'un modèle de mouvement 2D.

Différence entre deux images successives f et g associée au modèle de mouvement u_Θ au point (x, y) :

$$\text{DFD}_{(\Theta, \xi)}(x, y) = g((x, y) + u_\Theta(x, y)) - f(x, y) + \xi,$$

où Θ est l'ensemble de paramètres et ξ une variation de contraste. On estime les paramètres en minimisant

$$\sum_{(x, y) \in f} \rho(\text{DFD}_{(\Theta, \xi)}(x, y), \Gamma).$$

où ρ est l'estimateur robuste de Tukey :

$$\rho(t, \Gamma) = \begin{cases} \frac{t^2}{2} \left(\Gamma^4 - \Gamma^2 t^2 + \frac{t^4}{3} \right) & \text{si } |t| < \Gamma, \\ \frac{\Gamma^6}{6} & \text{sinon} \end{cases}$$

Estimation du flot optique (suite)

On utilise le modèle paramétrique suivant

$$u_\Theta(x, y) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ -a_2 & a_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} q_1 & q_2 & 0 \\ 0 & q_1 & q_2 \end{pmatrix} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix}.$$

Les six paramètres $(c_1, c_2, a_1, a_2, q_1, q_2)$ sont reliés au six paramètres de notre modèle $\alpha, \beta, \theta, A, B$ et C .

On obtient donc le déplacement de la caméra D entre deux images successives d'une séquence.

Grâce au groupe des recalages, on peut composer les déplacements et ainsi reconstituer le mouvement de la caméra au cours de la séquence.

Exemple 1

$$\theta = 0,8; \alpha = 0,0005; \beta = -0,051; A = 3; B = 1; C = -0,01.$$



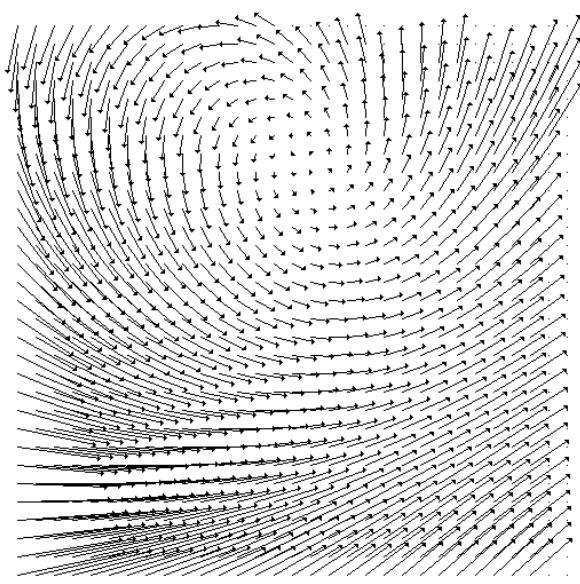
f



g1/2(f + g)

Exemple 1

	β	A	B	C	θ	α
initial	-0,051	2	1	-0,01	0,8	0,0005
estim.	-0,05113	2,812	0,075	0,98821	$\pi/4 = 0,785$	0,0004



flot $V1/2(f + g)$



1/2

Exemple 2



I_1

Exemple 3



Film incrustation

Film simulation

Références

Ces notes sont basées sur les documents suivants :

- C. Jonchery, PhD dissertation, *Estimation d'un mouvement de caméra et problèmes connexes*, ENS Cachan, November 2006.
- F. Dibos, C. Jonchery, G. Koepfler, *Camera motion estimation through planar deformation determination*, in *Journal of Mathematical Imaging and Vision*, vol 32, 1, p. 73-87, September 2008.
- F. Dibos, C. Jonchery, G. Koepfler, *Iterative Camera Motion and Depth Estimation in a Video Sequence*, in Jiang, X. and Petkov, N. (ed.) *CAIP 2009. LNCS*, vol. 5702, pp. 1028–1035, Springer, Heidelberg, September 2009.