Séance 2
Géometrie Epipolaire

(NB : - on note A^T la transposée
- $\wedge$ est le produit vect.
- par rapport au cours, les ~ ont disparu)

1) Géométrie projective

1.1) $\mathbb{P}^2$ (plan projectif)

- Point : $(u,v) \longrightarrow m=(u,v,1)^T$ à $\lambda \neq 0$ près (coord. homogènes)
- Direction : $u\vec{i}+v\vec{j} \longrightarrow m=(u,v,0)^T$ (point à l'∞)
- Droite $l=(a,b,c)^T$ avec $\boxed{m \in l \Leftrightarrow m^T l = 0}$

Propriétés :

i) m_1 et m_2 donnés, $m_1 \wedge m_2$ est la droite (m_1, m_2)

ii) l_1 et l_2 données, $l_1 \wedge l_2$ est leur intersection même si elles sont //.

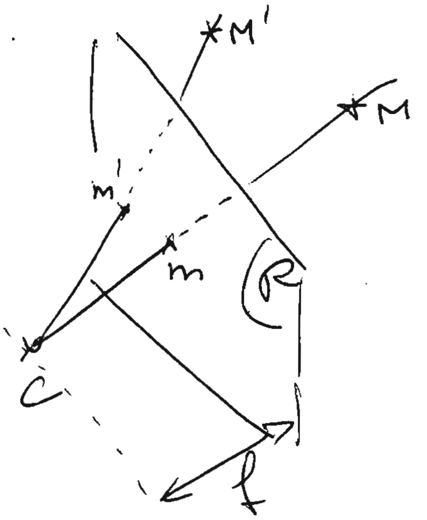
iii) $l_\infty = (0,0,1)$ est la droite à l'∞ $\left((u,v,0)^T \in l_\infty\right)$

1.2) $\mathbb{P}^3$ espace projectif

- Points $M=(X,Y,Z,T)^T$ à $\lambda \neq 0$ près
- $T=0 \rightarrow$ points à l'∞ (directions)
- $\pi=(a,b,c,d)^T$ plan $(M \in \pi \Leftrightarrow M^T\pi = 0)$
- Droites : coordonnées de Plücker (pas vu en cours)

2) Caméra

Le modèle le + simple (pinhole camera)



R : rétine

C : centre optique

M : point 3D

m : image 2D

f = distance focale

(i) Cas où $C=0$, $\mathcal{R}=\{z=1\}$ et le point $(X,Y,1)$ de $\mathcal{R}$ est numérisé en (X,Y) par le capteur.

Alors : $u = X/Z$, $v = Y/Z$ ce qu'on peut écrire

$$\underbrace{m}_{3\times1} = \underbrace{\left(I \middle| 0 \right)}_{3\times4} \underbrace{M}_{4\times1}$$ en coordonnées homogènes

(ii) Si maintenant $\mathcal{R}=\{z=f\}$ (focale $\neq 1$) alors

$m = PM$ avec $P = \left(\begin{matrix} f & & \\ & f & \\ & & 1 \end{matrix} \middle| 0 \right)$ car $u = f X/Z$, $v = f Y/Z$

(iii) Si on tient compte maintenant du changement de repère du capteur : $u = \alpha_u X/Z + u_0$, $v = \alpha_v Y/Z + v_0$

soit $P = \left(\begin{matrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{matrix} \middle| 0 \right)$ (NB : $\alpha_u = f/\text{larg. pixel}$, $\alpha_v = f/\text{haut. pixel}$)

(iv) Maintenant l'observateur se déplace → R rotation, t translation

$$\boxed{\begin{array}{l} m = PM \text{ avec } P = K[R|t] \rightarrow 6 \text{ param. extrinsèques} \\ K = \begin{pmatrix} \alpha_u & & u_0 \\ & \alpha_v & v_0 \\ & & 1 \end{pmatrix} \rightarrow 4 \text{ param. intrinsèques} \end{array}}$$

(NB : si pixels non rectangles (ie. ▱) alors un 5° param. intrins.)

(NB : on a négligé la distortion)

3) Relations utiles

i). Rayon de m :

– en coord. non homogènes pour M et avec $P = (A|b)$

$\lambda m = AM + b$

et donc $\boxed{M = \lambda A^{-1} m - A^{-1} b}$

– en particulier, on a (avec $\lambda=0$ et $\lambda=\infty$) :

$\boxed{\begin{array}{l} C = -A^{-1}b \quad \text{(coord. non homog.)} \\ A^{-1}m \text{ est la direction du rayon} \end{array}}$

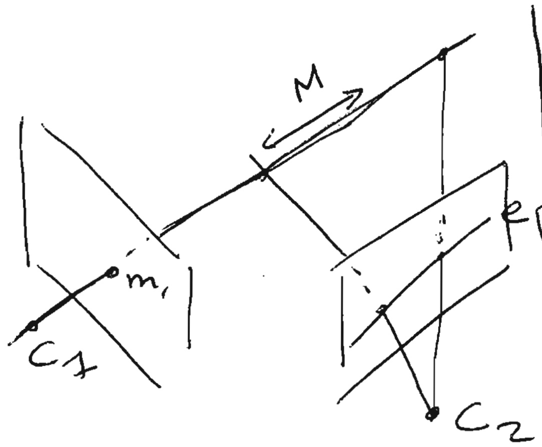
ii) Les points du plan ~~focal~~ focal (plan // à $\mathcal{R}$ passant par C) se projettent à l'∞, donc en $(u,v,0)^T$

→ la troisième ligne de P est l'équation du plan focal. En particulier :

$$\boxed{\text{la troisième ligne de } A \text{ est la normale à } \mathcal{R}}$$

4) Géometrie épipolaire

On a maintenant 2 caméras P_1 et P_2



m_1 étant donné, m_2 est nécessairement sur la projection du rayon C_1m_1 dans l'image 2 : c'est la droite épipolaire de m_1 dans 2

$$\boxed{m_1 \text{ connu} \longrightarrow m_2 \in ep_2(m_1) \text{"=" } P_2(C_1m_1)\text{"}}$$

Si $\boxed{P_i = (A_i|b_i)}$ alors :

- on cherche deux points du rayon C_1m_1, que l'on projette

$$\begin{cases} \times\ C_1 = \begin{pmatrix} -A_1^{-1}b_1 \\ 1 \end{pmatrix} \rightarrow e_2 = P_2C_1 = \boxed{-A_2A_1^{-1}b_1 + b_2 = e_2} \\ \qquad\qquad e_2 \text{ est l'épipôle } (\forall m_1,\ e_2 \in ep_2(m_1)) \\ \times\ M_\infty = \begin{pmatrix} A_1^{-1}m_1 \\ 0 \end{pmatrix} \rightarrow m_{2\infty} = P_2M_\infty = A_2A_1^{-1}m_1 \end{cases}$$

- on a alors $ep_2(m_2) = e_2 \wedge m_{2\infty} = [e_2]_\wedge A_2A_1^{-1}m_1$

en posant $\boxed{\begin{bmatrix} a \\ b \\ c \end{bmatrix}_\wedge = \begin{pmatrix} 0 & b & -c \\ -b & 0 & a \\ c & -a & 0 \end{pmatrix}}$

D'où finalement la relation symétrique :

$$\boxed{\begin{array}{l} m_2 \in ep_2(m_1) \Leftrightarrow m_1 \in ep_1(m_2) \Leftrightarrow m_2^T F_{21} m_1 = 0 \\ \text{avec } F_{21} = [e_2]_\wedge A_2A_1^{-1} \text{ (matrice fondamentale)} \end{array}}$$

I Mesure de photo consistance

- La plus courante est la corrélation croisée normalisée (NCC)

Soit W un voisinage autour des pixels :

$$\begin{cases} \langle I_i, m_i, I_j, m_j \rangle = \frac{1}{|W|} \sum_{w \in W} \left[(I_i(m_i+w) - \bar{I}_i(m_i)) \times (I_j(m_j+w) - \bar{I}_j(m_j)) \right] \\ \text{avec } \bar{I}_i(m_i) = \frac{1}{|W|} \sum_{w \in W} I_i(m_i+w) \end{cases}$$

puis $$\boxed{NCC(I_i, m_i, I_j, m_j) = \frac{\langle I_i, m_i, I_j, m_j \rangle}{\left(\langle I_i m_i I_i m_i \rangle \langle I_j m_j I_j m_j \rangle \right)^{1/2}}}$$

On a $-1 \leq NCC \leq 1$. Plus NCC est grande, plus les "textures" autour de m_i et m_j sont "similaires".

- Si deux images sont reliées par une transformation affine, $NCC = 1$
- Limitations :
  - invariance seulement par translation
  - pb de l'ouverture : ambiguïtés, par ex. le long d'un contour

II Points "saillants"

1- Points de Harris :

- G_σ gaussienne
- $H = \begin{pmatrix} (I_x)^2 * G_\sigma & (I_x I_y) * G_\sigma \\ (I_x I_y) * G_\sigma & (I_y)^2 * G_\sigma \end{pmatrix}$

tenseur de structure. (I_x, I_y dérivées)

Séance ~~4~~ 3

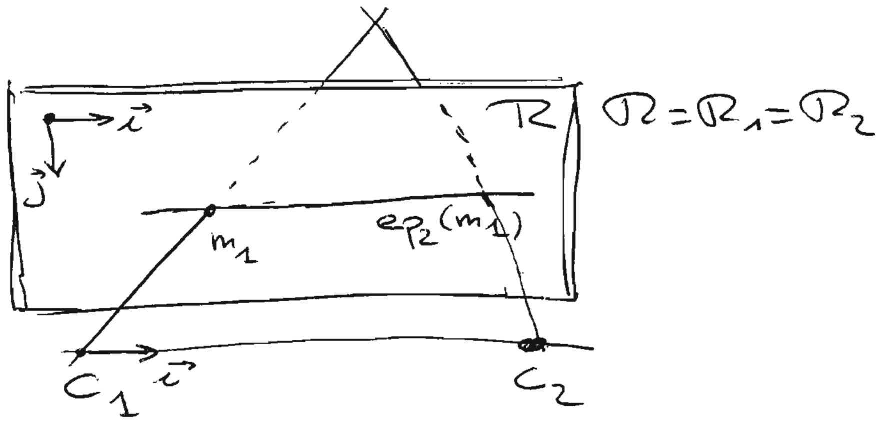
Rectification / Stéréo binoculaire / Graph Cuts

1) Rectification

On se place ici dans le cas de deux caméras de matrices $P_i = (A_i \mid b_i)$ connues

1.1) Configuration canonique

Supposons que les deux caméras partagent la même rétine, le même repère sur la rétine, et la direction $\overrightarrow{C_1C_2}$ comme horizontale de ce repère ($\overrightarrow{C_1C_2}$ est donc une des directions de la rétine).
Les épipolaires sont alors les horizontales !



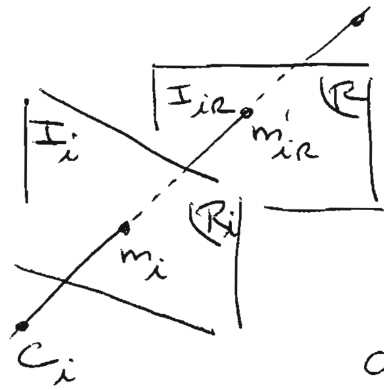
Une telle configuration est très confortable :

- (u_1, v_1) est associé à (u_2, v_1)
- on appelle $\boxed{d = u_2 - u_1 \text{ la disparité}}$
- apparier les images revient à estimer $d(u, v_1)$
- les fenêtres de corrélation se correspondent davantage.

On va voir qu'il est toujours possible (sauf cas dégénérés) de se ramener *a postériori* à ce cas canonique $\longrightarrow$ c'est la rectification

1.2) Principe

Il suffit de remarquer qu'on peut à postériori calculer l'image qui aurait été obtenue en changeant la rétine d'un appareil, et ce, sans connaître la position 3D des points



Le point M a donné naissance à $I_i(m_i)$ dans R_i. Avec R au lieu de R_i il aurait généré $\boxed{I_{iR}(m_{iR}) = I_i(m_i)}$

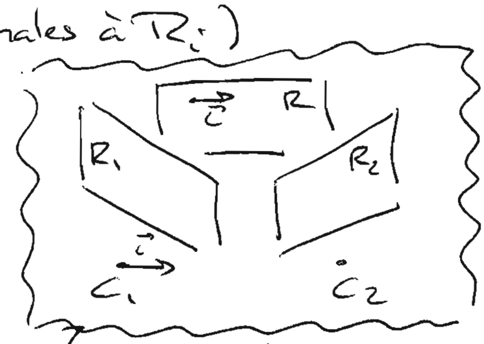
avec $\boxed{m_{iR} = (C_i m_i) \cap R}$

1.3) Calcul (NB : les points 3D sont ici en coord. non homogènes)

a) Plan R

On choisit pour R le plan (i) contenant $\overrightarrow{C_1C_2}$ et la direction $R_1 \cap R_2$ et (ii) passant par l'origine :

- $\vec{\imath} = \dfrac{\overrightarrow{C_1C_2}}{\|\overrightarrow{C_1C_2}\|}$ (Rappel : $C_i = -A_i^{-1} b_i$)
- $\vec{n}_i$ = 3ème ligne de A_i (normales à R_i)
- $\vec{\jmath}\,' = \vec{n}_1 \wedge \vec{n}_2$
- $\vec{k} = \vec{\imath} \wedge \vec{\jmath}\,' / \|\vec{\imath} \wedge \vec{\jmath}\,'\|$
- $\vec{\jmath} = \vec{k} \wedge \vec{\imath}$



R est alors le plan $\{M \cdot \vec{k} = 0\}$

b) Passage de $m_i = (u_i\, v_i\, 1)^T$ à $m_{iR} = (\alpha, \beta)$ dans R

Il suffit d'intersecter $(C_i m_i)$ avec R :

résoudre $\boxed{\begin{cases} M = \lambda A_i^{-1} m_i + C_i \\ M \cdot \vec{k} = 0 \end{cases}}$ en λ, puis en M

Alors $\boxed{\alpha = M \cdot \vec{\imath},\ \beta = M \cdot \vec{\jmath}}$

c) Projections

En réalité : (i) pour chaque image, on ramène (α,β) à des intervalles $[0,w]\times[0,h]$ (si les images initiales sont de dimension $w\times h$)

(ii) on remplit I_{iR} à partir de I_i. Il faut savoir passer de m_{iR} à m_i

Au total :

- Pour $i=1,2$
  - calculer $m_{iR}=(\alpha,\beta)$ pour $m_i=(0,0)\,(0,h)\,(w,0)$ et (w,h)
  - mémoriser les valeurs extrêmes $\alpha^i_{min}\ \alpha^i_{max}\ \beta^i_{min}\ \beta^i_{max}$
- $\beta_{min}=\min(\beta^1_{min}\,\beta^2_{min})$, $\beta_{max}=\max(\beta^1_{max}\,\beta^2_{max})$

  (NB : intervalle commun pour β)
- Pour $i=1,2$

  Pour $u_R=0\nearrow w$ et $v_R=0\nearrow h$

  - $\begin{cases}\alpha=\alpha^i_{min}+(\alpha^i_{max}-\alpha^i_{min})\dfrac{u_R}{w}\\ \beta=\beta_{min}+(\beta_{max}-\beta_{min})\dfrac{v_R}{h}\end{cases}$
  - $M=\alpha\vec{\imath}+\beta\vec{\jmath}$ (NB : point 3D !)
  - $m_i=A_iM+b_i=(u_i v_i w_i)^T$
  - $I_{iR}(u_R,v_R)\leftarrow I_i\left(\dfrac{u_i}{w_i},\dfrac{v_i}{w_i}\right)$

  (NB : - interpoler dans I_i
  - attention : on peut tomber en dehors de l'image)