Géométrie épipolaire et reconstruction 3D

Patrick Hébert & Denis Laurendeau (Dernière révision : juin 2017)

Introduction

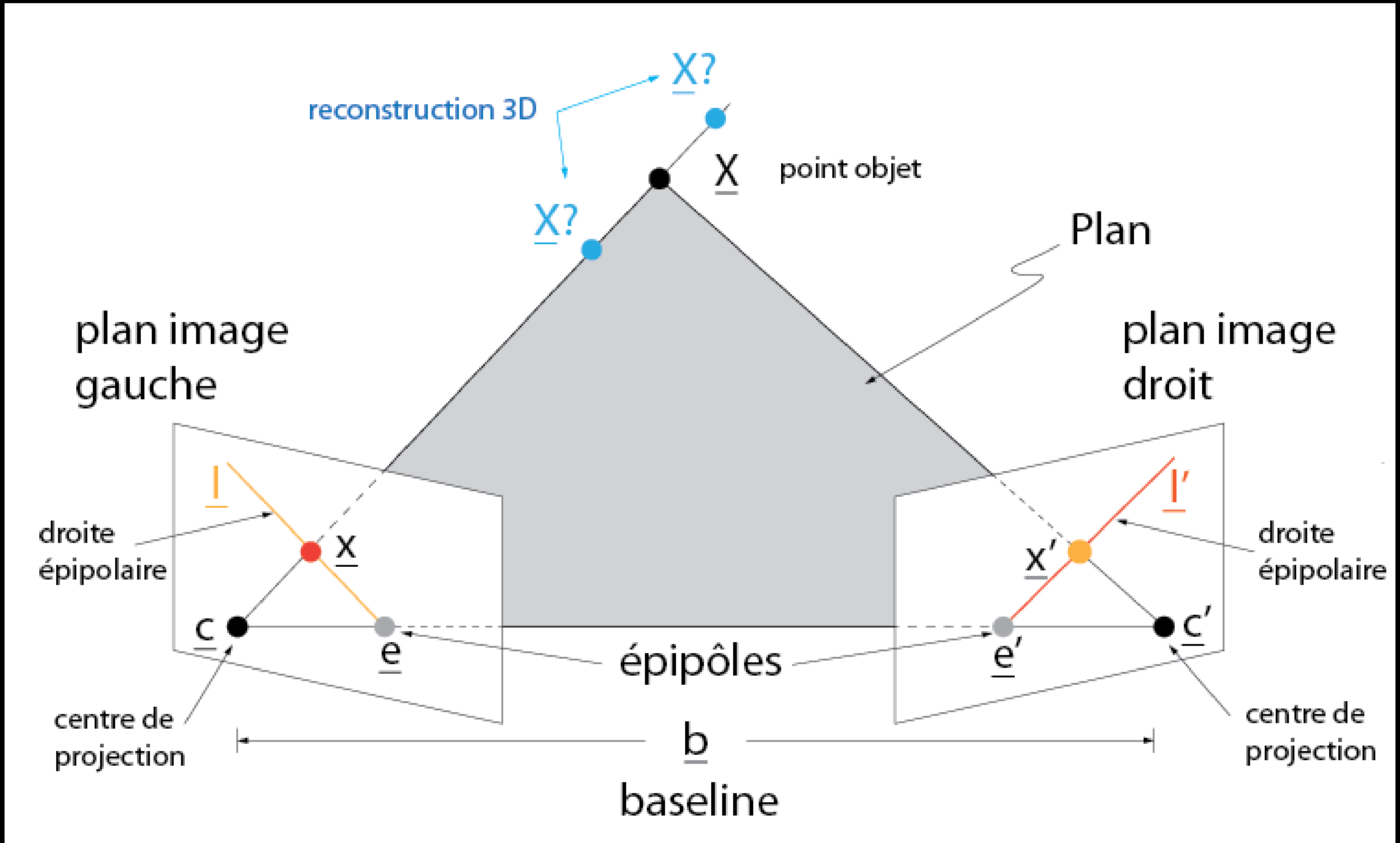
Cette partie du cours aborde les éléments suivants :

1. Géométrie épipolaire et stéréoscopie
2. Matrice fondamentale (et matrice essentielle)
3. Estimation de la matrice fondamentale (algorithme normalisé à 8 points)
4. Reconstruction 3D
5. Rectification d’images
6. Appariement stéréoscopique (brève introduction)

Géométrie épipolaire et stéréoscopie

Conditions générales de la géométrie épipolaire (1)

Deux sténopés non-inverseurs observant une scène ***quelconque***



Conditions générales de la géométrie épipolaire (2)

La géométrie épipolaire concerne la **géométrie projective *intrinsèque*** entre **deux vues** d'une **même scène** prises de **points de vue différents**.

La géométrie épipolaire dépend des ***caméras*** :

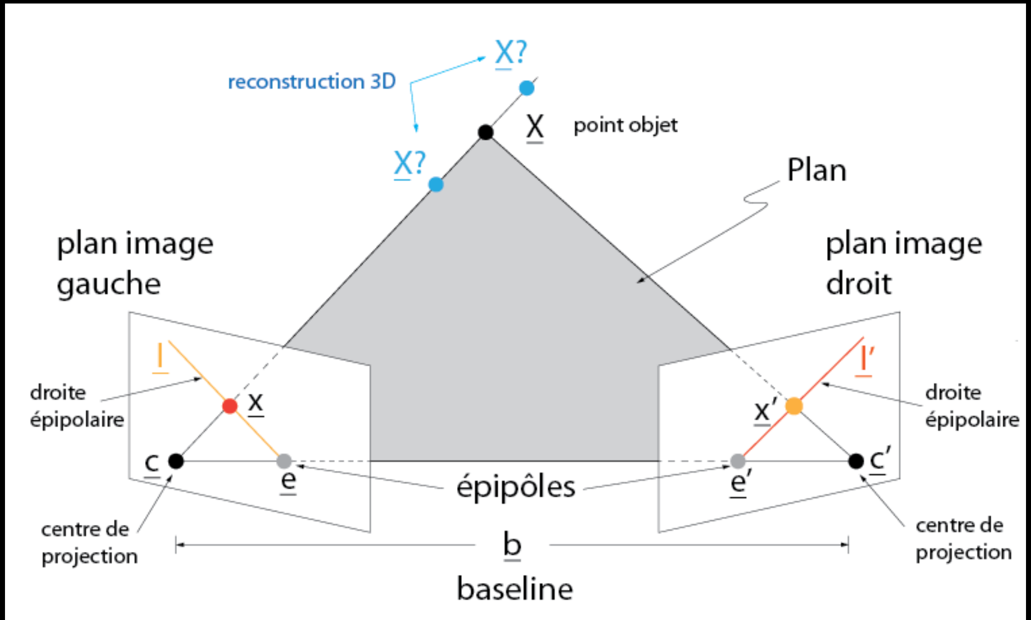
-paramètres ***intrinsèques***

-paramètres ***extrinsèques***

La géométrie épipolaire ne ***dépend pas*** :

-de la ***structure*** de la scène

-du ***contenu*** de la scène



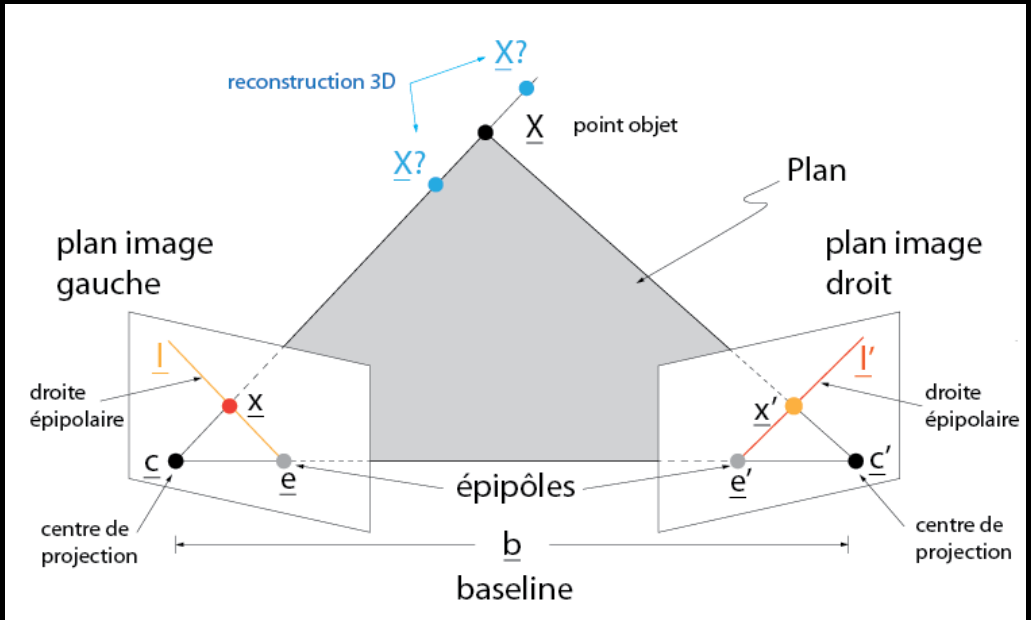
Conditions générales de la géométrie épipolaire (3)

La matrice fondamentale $\underline{\underline{\tilde{F}}}$ encapsule cette géométrie intrinsèque.

La matrice fondamentale est une matrice ***3 x 3*** de ***rang 2*** (i.e. deux colonnes sont linéairement indépendantes).

Si un point objet $\underline{\tilde{X}}$ donne naissance aux points image $\underline{\tilde{x}}$ et $\underline{\tilde{x}}'$, alors les points image satisfont l'équation suivante:

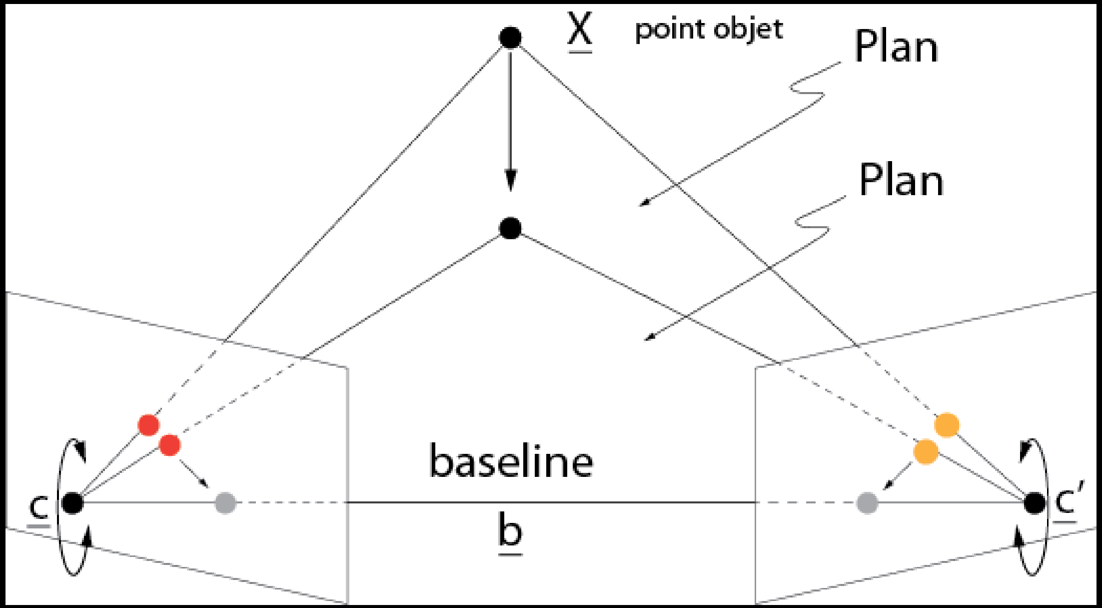
$$\boxed{\underline{\tilde{x}}'^T \underline{\underline{\tilde{F}}}\, \underline{\tilde{x}} = 0} \qquad (1)$$



Conditions générales de la géométrie épipolaire (4)

Bien que $\underline{\underline{\tilde{F}}}$ soit indépendante de la scène, elle peut être estimée à partir de **correspondances d'images** de points dans celle-ci (et rien d'autre).

La géométrie épipolaire entre 2 vues est la géométrie reliée à l'intersection des plans images avec **le faisceau de plans ayant le baseline comme axe de rotation**.



Le **baseline** est est la **droite** joignant les **centres de projection** des deux caméras.

Chaque **point** de la scène **sous-tend** un **plan** passant par le baseline (voir figure à droite)

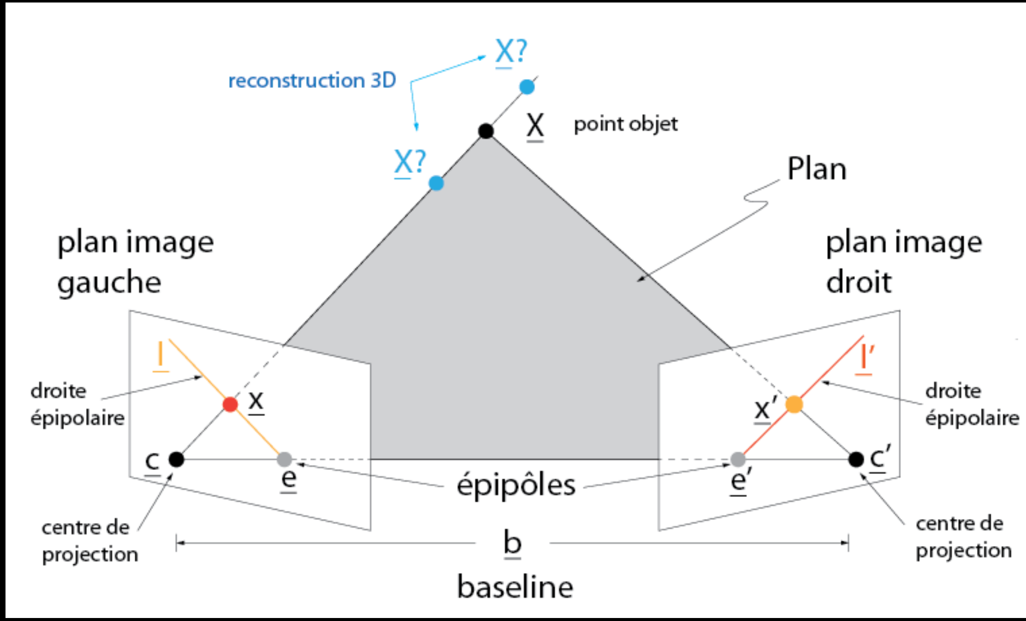
Quelques définitions

épipôles : points d'intersection de la droite reliant les centres de projection des deux caméras avec les plans image

baseline : droite reliant les centres de projection des deux caméras

plan épipolaire : plan passant par un point objet dans la scène et les centres de projection des deux caméras (il y a un plan épipolaire par point objet). Le plan épipolaire passe par les épipôles.

droites épipolaires : droites résultant de l'intersection du plan épipolaire avec les plans image des deux caméras. Dans chaque plan image, il n'y a qu'une seule droite épipolaire par point objet. Toutes les droites épipolaires dans un plan image passent par l'épipôle.



On remarque que quand les **axes optiques** des caméras sont **parallèles**, les **épipôles** sont reportés à **l'infini**.

Les **projections inverses** $\underline{c} - \underline{x}$ et $\underline{c}' - \underline{x}'$ sont des **droites** (rayons) s'intersectant au point **objet** $\underline{\underline{\tilde{X}}}$

Dérivation **géométrique** de $\underline{\underline{\tilde{F}}}$

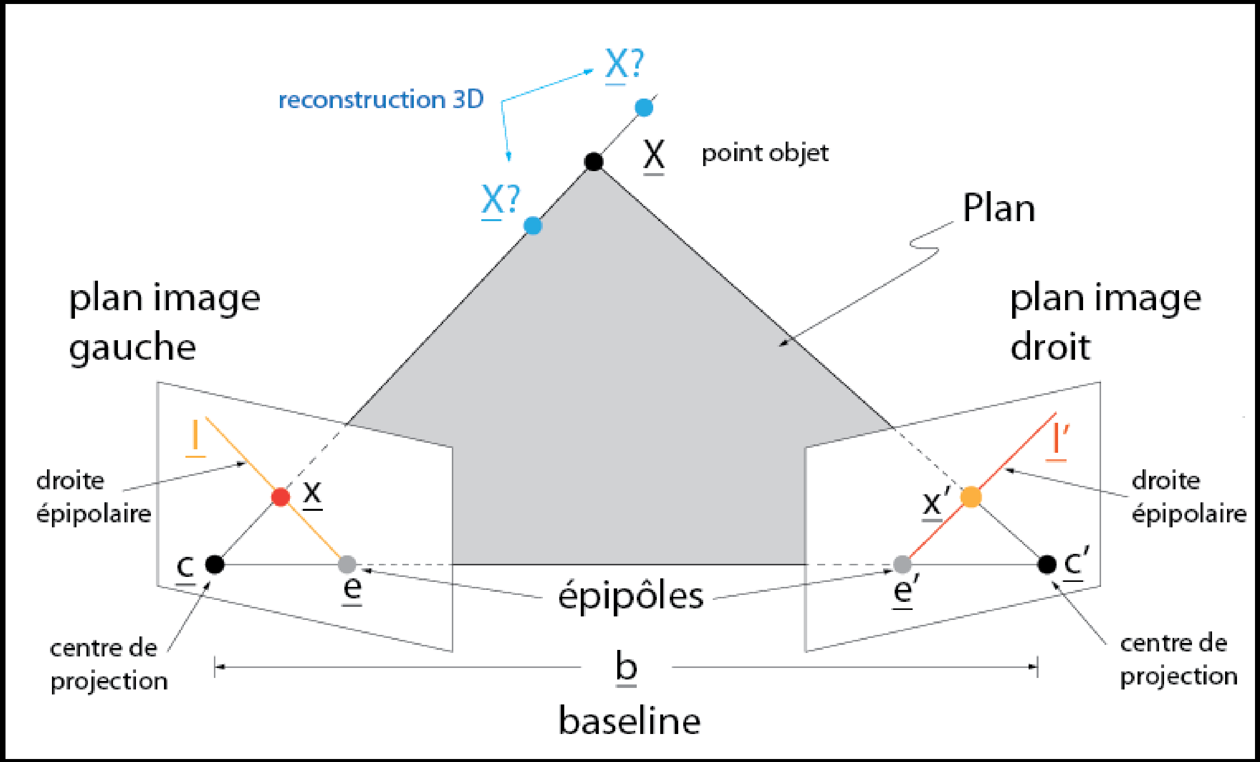
Supposons qu'on connaît seulement le point image $\underline{\tilde{x}}$. Où $\underline{\tilde{x}}'$ se trouve-t-il?

Avec la figure de droite, on sait que $\underline{\tilde{x}}'$ est dans le plan $\underline{c} - \underline{X} - \underline{c}'$.

Le plan $\underline{c} - \underline{X} - \underline{c}'$ intersecte le plan image de la caméra de droite sur la droite $\underline{l}'$ passant par l'épipôle $\underline{e}'$.

Par conséquent, le point $\underline{x}'$ appartient à la droite $\underline{l}'$.

De manière générale, il existe un **mapping** entre un **point image** $\underline{x}$ de la caméra de **gauche** et une **droite épipolaire** $\underline{l}'$ de la caméra de **droite**.



$$\underline{x} \rightarrow \underline{l}' \qquad (2)$$

Dérivation géométrique de la matrice fondamentale

Dérivation **géométrique** de $\underline{\underline{\tilde{F}}}$

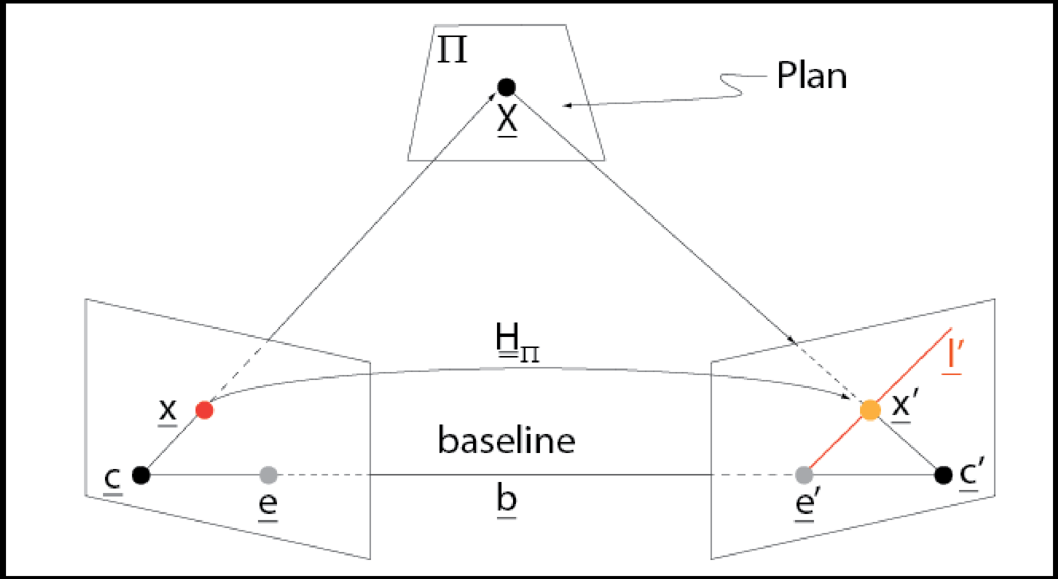
Etape 1: transfert d'un point via un plan dans l'espace

Supposons qu'un point objet $\underline{X}$ dans une scène repose sur un **plan** Π qui ne **passe pas** par les centres de projection $\underline{c}$ et $\underline{c}'$ des caméras (voir figure à droite)

Le rayon passant par $\underline{c}$ et le point image $\underline{x}$ de la caméra de gauche intersecte le plan Π au point objet $\underline{X}$.

Le point objet $\underline{X}$ est ensuite **projeté** sur le plan image de l'image de droite. L'image de $\underline{X}$ dans le plan image de la caméra de droite se trouve sur la droite épipolaire $\underline{l}'$.

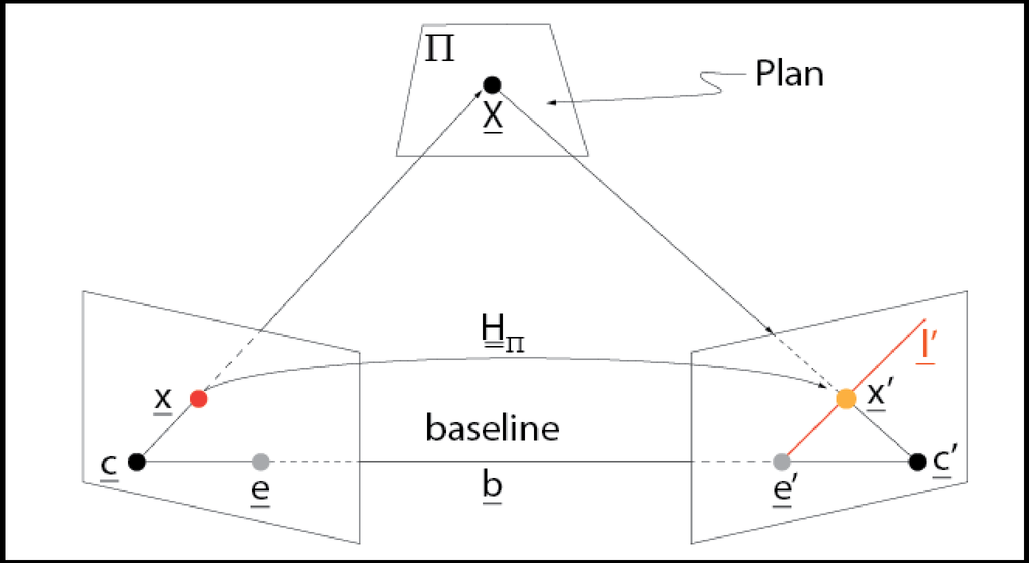
Il existe une homographie $\underline{\underline{H}}_{\Pi}$ qui effectue un mapping de $\underline{x}$ sur $\underline{x}'$.



Etape 2: construction de la droite épipolaire l’

Étant donné l’image $\underline{x}'$, la droite épipolaire $\underline{l}'$ est donnée par:

$$\underline{\tilde{l}}' = \underline{\tilde{e}}' \otimes \underline{\tilde{x}}' = [\underline{\tilde{e}}']_x \underline{\tilde{x}}' \qquad (3)$$



où $\otimes$ est le produit vectoriel et $[\underline{\tilde{e}}']_x$ l’expression matricielle du produit vectoriel:

$$[\underline{\tilde{e}}']_x = \begin{bmatrix} 0 & -e'_3 & -e'_2 \\ e'_3 & 0 & -e'_1 \\ e'_2 & e'_1 & 0 \end{bmatrix}$$

L’expression (3) vient de l’observation suivante. Soit $\underline{l}_i = \underline{u} \otimes \underline{v}$ une droite. Le triple produit $\underline{u}^t \ (\underline{u} \otimes \underline{v}) = \underline{v}^t(\underline{u} \otimes \underline{v}) = 0$ est nul. On conclut donc que $\underline{u}$ et $\underline{v}$ sont sur la droite $\underline{u} \otimes \underline{v}$ car $\underline{u}^t \underline{l}_i = \underline{v}^t \underline{l}_i = 0$ pour des points appartenant à une droite.

Etape 2: construction de la droite épipolaire l'

Grâce à l'homographie $\underline{\underline{\tilde{H}}}_{\Pi}$ on peut écrire pour $\underline{\tilde{x}}'$:

$$\underline{\tilde{x}}' = \underline{\underline{\tilde{H}}}_{\Pi}\underline{\tilde{x}} \qquad (4)$$

En combinant (3) et (4) on obtient:

$$\underline{\tilde{l}}' = [\widetilde{\underline{e}'}]_x \underline{\underline{\tilde{H}}}_{\Pi}\underline{\tilde{x}} \qquad (5)$$

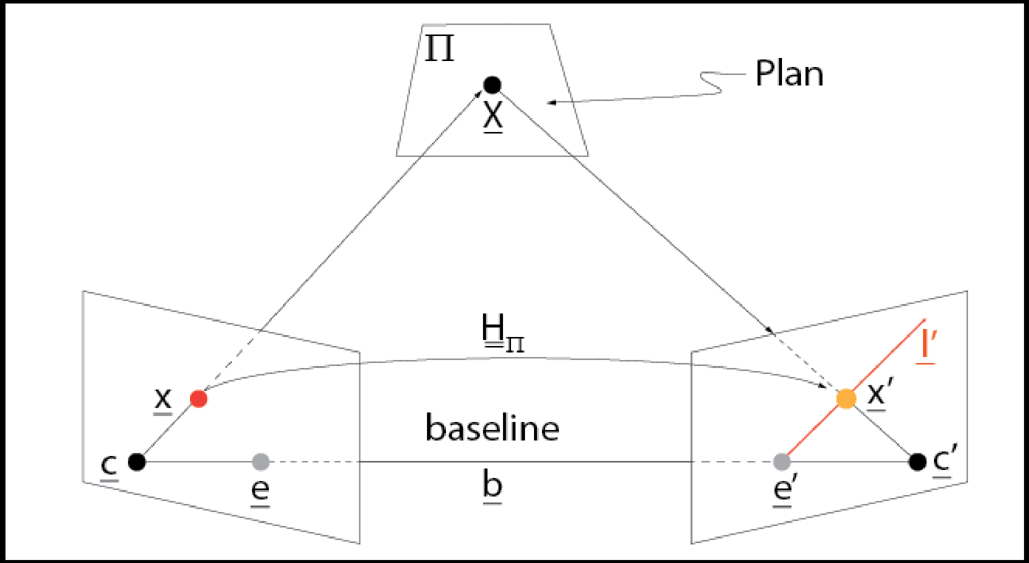
Comme $\underline{\tilde{x}}'$ est sur $\underline{\tilde{l}}'$, on a que:

$$\underline{\tilde{x}}'^t \underline{\tilde{l}}' = 0 \qquad (6)$$

En remplaçant (5) dans (6) on a finalement que:

$$\underline{\tilde{x}}'[\underline{\tilde{e}}']_x \underline{\underline{\tilde{H}}}_{\Pi}\underline{\tilde{x}} = \underline{\tilde{x}}'\underline{\underline{\tilde{F}}}\underline{\tilde{x}} = 0 \qquad (7)$$

où la **matrice fondamentale** de la paire de caméras est $$\boxed{\underline{\underline{\tilde{F}}} = [\underline{\tilde{e}}']_x \underline{\underline{\tilde{H}}}_{\Pi} \qquad (8)}$$



Remarques sur la matrice fondamentale $\underline{\underline{\tilde{F}}}$

Comme $[\underline{\tilde{e}}']_x$ est de rang 2 et $\underline{\underline{\widetilde{H}}}_\Pi$ est de rang 3, $\underline{\underline{\tilde{F}}}$ est de **rang 2**.

Comme $[\underline{\tilde{e}}']_x$ et $\underline{\underline{\widetilde{H}}}_\Pi$ ne dépendent que de l'arrangement géométrique des caméras, il en va de même pour $\underline{\underline{\tilde{F}}}$.

Dérivation algébrique de la matrice fondamentale

Dérivation **algébrique** de $\underline{\underline{\tilde{F}}}$

Cette dérivation exprime $\underline{\underline{\tilde{F}}}$ en fonction des **matrices caméras** $\underline{\underline{\tilde{P}}}$ et $\underline{\underline{\tilde{P}}}'$ (formées des paramètres intrinsèques et extrinsèques).

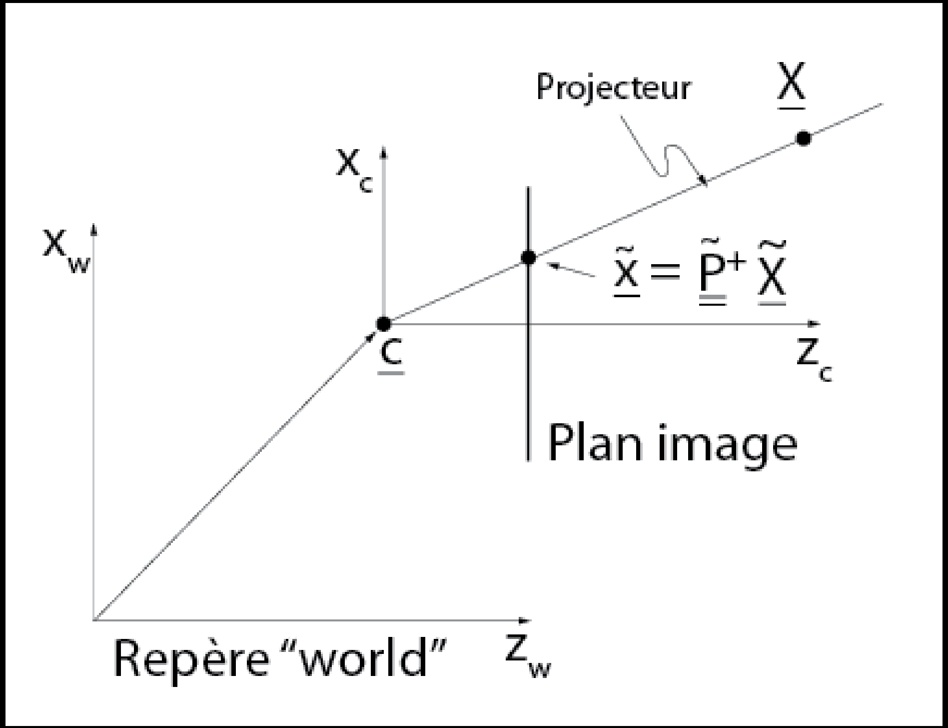
Par la projection de perspective, le point image $\underline{\tilde{x}}$ est obtenu de:

$$\underline{\tilde{x}} = \underline{\underline{\tilde{P}}}\,\underline{\tilde{X}} \qquad (9)$$

Dans le repère “world”, un point sur le projecteur est :

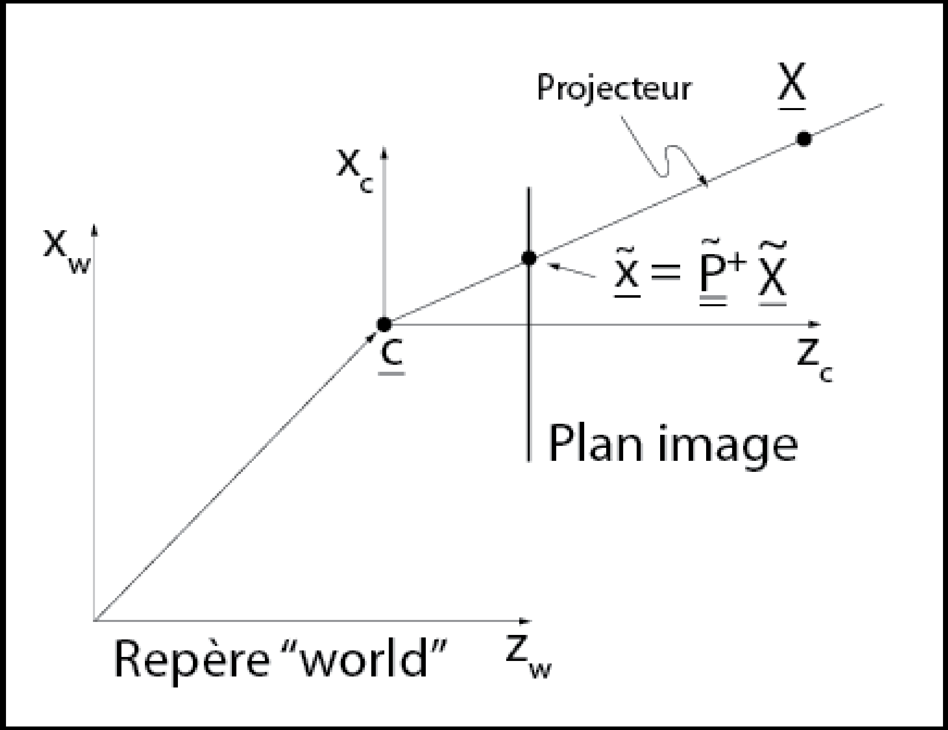
$$\underline{\tilde{x}}_w = \underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} \qquad (10)$$

où $\underline{\underline{\tilde{P}}}^{+}$ est la **pseudo-inverse** de $\underline{\underline{\tilde{P}}}$ (i.e. $\underline{\underline{\tilde{P}}}\,\underline{\underline{\tilde{P}}}^{+} = \underline{\underline{\tilde{I}}}$).



L'équation du projecteur passant par le point image et le centre de projection (les deux exprimés dans le repère “world”) est, en observant la figure de droite :

$$\underline{\tilde{X}}(\mu) = \underline{\tilde{c}} + \mu\left(\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} - \underline{\tilde{c}}\right) \qquad (11)$$



En développant (11), on obtient :

$$\underline{\tilde{X}}(\mu) = \mu\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} + (1-\mu)\underline{\tilde{c}}$$

Comme $\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}}$ contient déjà un facteur d'échelle on peut simplifier pour obtenir :

$$\underline{\tilde{X}}(\lambda) = \underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} + \lambda\underline{\tilde{c}} \qquad (12)$$

On a aussi que $\underline{\underline{\tilde{P}}}\underline{\tilde{c}} = \underline{\tilde{0}}$ (i.e. l'image du centre de projection est l'origine du repère camera qui est un point dont l'image est indéfinie)

Deux points sur le projecteurs (dans le repère “world”) sont :

$$\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} \quad \text{et} \quad \underline{\tilde{c}}$$

Les images de ces deux points sur le plan image de la caméra de droite sont :

$$\underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} \quad \text{et} \quad \underline{\underline{\tilde{P}}}'\underline{\tilde{c}}$$

La droite épipolaire passant par ces deux points est donnée par :

$$\underline{\tilde{l}}' = \underline{\underline{\tilde{P}}}'\underline{\tilde{c}} \otimes \underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} \tag{13}$$

Dans cette expression on remarque que (en revenant à (3), (4) et (5)) :

$$\underline{\tilde{l}}' = \underbrace{\underline{\underline{\tilde{P}}}'\underline{\tilde{c}}}_{\text{épipôle } \underline{\tilde{e}}'} \otimes \underbrace{\underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^{+}}_{\underline{\underline{\tilde{H}}}_{\Pi}}\underline{\tilde{x}} \tag{14}$$

image de $\underline{\tilde{c}}$ dans l'image de droite

homographie mappant $\underline{\tilde{x}}$ sur $\underline{\tilde{x}}'$

À partir de (14), on peut écrire :

$$\underline{\tilde{l}}' = [\underline{\tilde{e}}']_x\left(\underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^+\right)\underline{\tilde{x}} = \underline{\underline{\tilde{F}}}\,\underline{\tilde{x}} \tag{14}$$

Et donc :

$$\boxed{\underline{\underline{\tilde{F}}} = [\underline{\tilde{e}}']_x\underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^+} \tag{15}$$

ne dépend que des caméras et de leur arrangement

avec

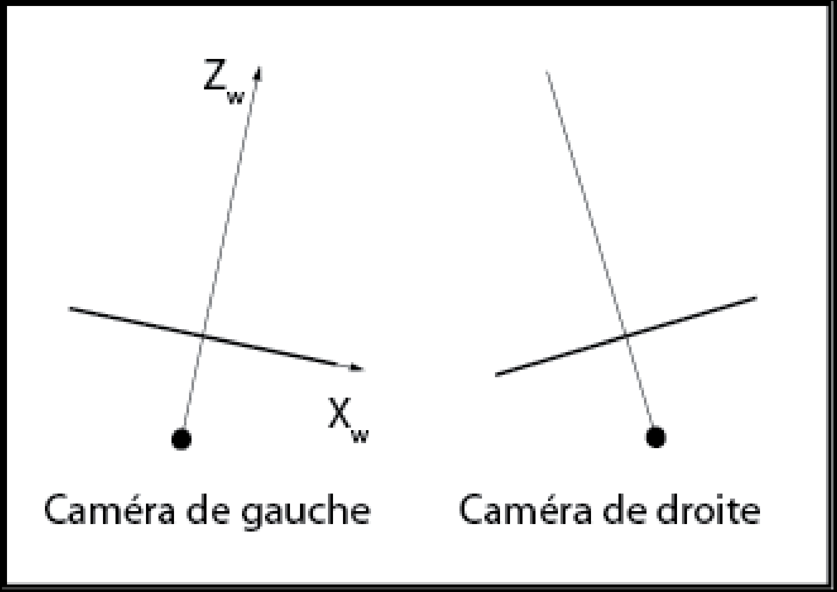
$$\underline{\underline{\tilde{H}}}_{\Pi} = \underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^+ \tag{16}$$

Matrice fondamentale pour une paire de caméras calibrées

Nous nous intéressons au cas de l'expression **algébrique** de la **matrice fondamentale** pour une **paire** de **caméras calibrées.**

Pour simplifier, le repère "world" est confondu avec le repère caméra de la caméra de **gauche** (tel que montré sur la figure). Dans ce cas, on peut écrire pour les matrices caméras :

$$\underline{\underline{\tilde{P}}} = \underline{\underline{\tilde{K}}}\left[\underline{\underline{\tilde{I}}}|\underline{0}\right] \qquad (17) \qquad\qquad \underline{\underline{\tilde{P}}}' = \underline{\underline{\tilde{K}}}'\left[\underline{\underline{\tilde{R}}}|\underline{\tilde{t}}\right] \qquad (18)$$



De (17) on peut écrire :

$$\underline{\underline{\tilde{P}}}^{+} = \begin{bmatrix} \underline{\underline{\tilde{K}}}^{-1} \\ \underline{0}^{t} \end{bmatrix} \qquad (19)$$

Si on revient à (14)

$$\underline{\tilde{l}}' = \underline{\underline{\tilde{P}}}'\underline{\tilde{c}} \otimes \underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} \qquad (14)$$

qu'on peut réécrire comme :

$$\underline{\tilde{l}}' = \left[\underline{\underline{\tilde{P}}}'\underline{\tilde{c}}\right]_x \underline{\underline{\tilde{P}}}'\underline{\underline{\tilde{P}}}^{+}\underline{\tilde{x}} \qquad (20)$$

on a que :

$$\underline{\underline{\tilde{P}}}'\underline{\tilde{c}} = \underline{\underline{\tilde{K}}}'\ \left[\underline{\underline{\tilde{R}}}|\underline{\tilde{t}}\right]\begin{bmatrix}\tilde{0}\\1\end{bmatrix} = \left[\underline{\underline{\tilde{K}}}'\ \underline{\tilde{t}}\right] \qquad (21)$$

Avec (21), (19) et (18) , on peut écrire pour la matrice fondamentale :

$$\underline{\underline{\tilde{F}}} = \left[\underline{\underline{\widetilde{K}}}' \ \underline{\tilde{t}}\right]_x \underline{\underline{\widetilde{K}}}' \ \underline{\underline{\tilde{R}}} \ \underline{\underline{\widetilde{K}}}^{-1} \qquad (22)$$

L’identité suivante :

$$[t]_x \underline{\underline{M}} = \underline{\underline{M}}^{-t} \left[\underline{\underline{M}}^{-1} \ \underline{t}\right]_x \qquad (23)$$

permet d’écrire diverses expressions équivalentes pour la matrice fondamentale :

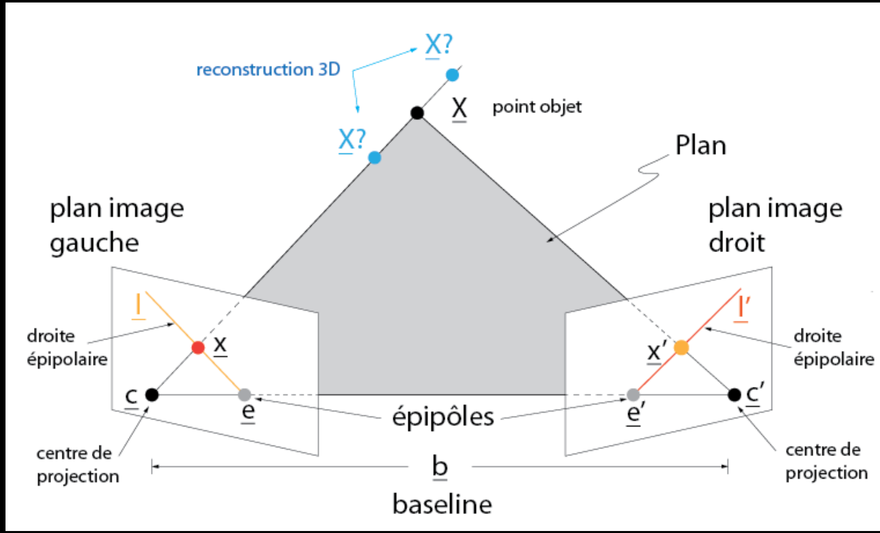
$$\underline{\underline{\tilde{F}}} = \left[\underline{\underline{\widetilde{K}}}' \ \underline{\tilde{t}}\right]_x \underline{\underline{\widetilde{K}}}' \ \underline{\underline{\tilde{R}}} \ \underline{\underline{\widetilde{K}}}^{-1} = \underline{\underline{\widetilde{K}}}'^{-t} \ [\underline{\tilde{t}}]_x \underline{\underline{\tilde{R}}} \ \underline{\underline{\widetilde{K}}}^{-1} = \underline{\underline{\widetilde{K}}}'^{-t} \ \underline{\underline{\tilde{R}}} \left[\underline{\underline{\tilde{R}}}^t \ \underline{\tilde{t}}\right]_x \underline{\underline{\widetilde{K}}}^{-1} = \underline{\underline{\widetilde{K}}}'^{-t} \ \underline{\underline{\tilde{R}}} \ \underline{\underline{\widetilde{K}}}^t \left[\underline{\underline{\widetilde{K}}} \ \underline{\underline{\tilde{R}}}^t \ \underline{\tilde{t}}\right]_x \qquad (24)$$

Propriétés de la matrice fondamentale

La matrice fondamentale satisfait la condition suivante:

$$\underline{\tilde{x}}'\, \underline{\underline{\tilde{F}}}\, \underline{\tilde{x}} = 0 \qquad (25)$$

De plus, si $\underline{\tilde{x}}'$ et $\underline{\tilde{x}}$ satisfont (25) ils sont **coplanaires** comme on peut le vérifier sur la figure de droite



La coplanarité est la condition à satisfaire pour qu'ils soient en correspondance stéréoscopique.

Ce résultat est important car, grâce à (25), il est possible de caractériser (et d'estimer) la matrice fondamentale uniquement à partir d'appariements stéréoscopiques valides sans connaître les matrices caméras de la paire stéréoscopique.

Formulé autrement, même si $\underline{\underline{\tilde{F}}}$ est indépendant du contenu de la scène, elle peut être estimée à partir de correspondances obtenues d'images de celle-ci.

Propriétés de la matrice fondamentale $\underline{\underline{\tilde{F}}}$

- **Propriété 1** : Si $\underline{\underline{\tilde{F}}}$ est la matrice fondamentale de $\left(\underline{\underline{\tilde{P}}}\,,\,\underline{\underline{\tilde{P}}}'\right)$, alors $\underline{\underline{\tilde{F}}}^t$ est la matrice fondamentale de $\left(\underline{\underline{\tilde{P}}}'\,,\,\underline{\underline{\tilde{P}}}\right)$.

- **Propriété 2** : $\forall\,\underline{\tilde{x}}$ dans l'image de gauche, la droite épipolaire dans l'image de droite est donnée par $\underline{\tilde{l}}' = \underline{\underline{\tilde{F}}}\,\underline{\tilde{x}}$. De plus, $\underline{\tilde{l}} = \underline{\underline{\tilde{F}}}^t\,\underline{\tilde{x}}'$ est la droite épipolaire de l'image de gauche correspondant à $\underline{\tilde{x}}'$.

- **Propriété 3** : $\forall\,\underline{\tilde{x}}$ autre que $\underline{\tilde{e}}$, la droite épipolaire correspondante $\underline{\tilde{l}}' = \underline{\underline{\tilde{F}}}\,\underline{\tilde{x}}$ contient l'épipole $\underline{\tilde{e}}'$. On a donc que $\underline{\tilde{e}}'^t\left(\underline{\underline{\tilde{F}}}\,\underline{\tilde{x}}\right) = \left(\underline{\tilde{e}}'^t\underline{\underline{\tilde{F}}}\right)\underline{\tilde{x}} = 0$. Comme en général $\underline{\tilde{x}} \neq \underline{\tilde{0}}$, alors $\underline{\tilde{e}}'^t\,\underline{\underline{\tilde{F}}} = \underline{0}$. On dit que $\underline{\tilde{e}}'$ est le “null vector” gauche de $\underline{\underline{\tilde{F}}}$. De même, $\underline{\underline{\tilde{F}}}\,\underline{\tilde{e}} = 0$. On dit que $\underline{\tilde{e}}$ est le “null vector” droit de $\underline{\underline{\tilde{F}}}$.

- **Propriété 4** : $\underline{\underline{\tilde{F}}}$ est une matrice homogène de rang 2 avec 7 degrés de liberté.

- **Propriété 5** : Si $\underline{\underline{\tilde{x}}}$ et $\underline{\underline{\tilde{x}}}'$ sont des points images en correspondance, alors $\underline{\underline{\tilde{x}}}'\,\underline{\underline{\tilde{F}}}\,\underline{\tilde{x}} = 0$.

- **Propriété 6** : $\underline{\tilde{l}}' = \underline{\underline{\tilde{F}}}\,\underline{\tilde{x}}$ est la droite épipolaire correspondant à $\underline{\tilde{x}}$. $\underline{\tilde{l}} = \underline{\underline{\tilde{F}}}^t\,\underline{\tilde{x}}'$ est la droite épipolaire correspondant à $\underline{\tilde{x}}'$.

UNIVERSITÉ LAVAL

- **Propriété 7** : $\underline{\underline{\tilde{F}}}\,\underline{\tilde{e}} = 0$ et $\underline{\underline{\tilde{F}}}^t\,\underline{\tilde{e}}' = 0$ (i.e. les épipôles sont les “null vectors” droit et gauche de $\underline{\underline{\tilde{F}}}$.

- **Propriété 8** : diverses expressions pour $\underline{\underline{\tilde{F}}}$:

matrices caméras **générales**:

$$\underline{\underline{\tilde{F}}} = [\underline{\tilde{e}}']_x \underline{\underline{\tilde{P}}}' \underline{\underline{\tilde{P}}}^+ \qquad (25)$$

matrices caméras **canoniques** ($\underline{\underline{\tilde{P}}} = \left[\underline{\underline{\tilde{I}}} \quad |\underline{\tilde{0}}\right]$ et $\underline{\underline{\tilde{P}}}' = \left[\underline{\underline{\widetilde{M}}} \quad |\underline{\widetilde{m}}\right]$) :

$$\underline{\underline{\tilde{F}}} = [\underline{\tilde{e}}']_x \underline{\underline{\widetilde{M}}} = \underline{\underline{\widetilde{M}}}^{-t}[\underline{\tilde{e}}]_x \quad \text{où} \quad \underline{\tilde{e}}' = \underline{\widetilde{m}} \quad \text{et} \quad \underline{\tilde{e}} = \underline{\underline{\widetilde{M}}}^{-1}\underline{\widetilde{m}} \qquad (26)$$

matrices caméras **non à l’infini** (i.e. focale non infinie $\underline{\underline{\tilde{P}}} = \underline{\underline{\widetilde{K}}}\left[\underline{\underline{\tilde{I}}}|\underline{\tilde{0}}\right]$, $\underline{\underline{\tilde{P}}}' = \underline{\underline{\widetilde{K}}}'\left[\underline{\underline{\tilde{R}}}|t\right]$) :

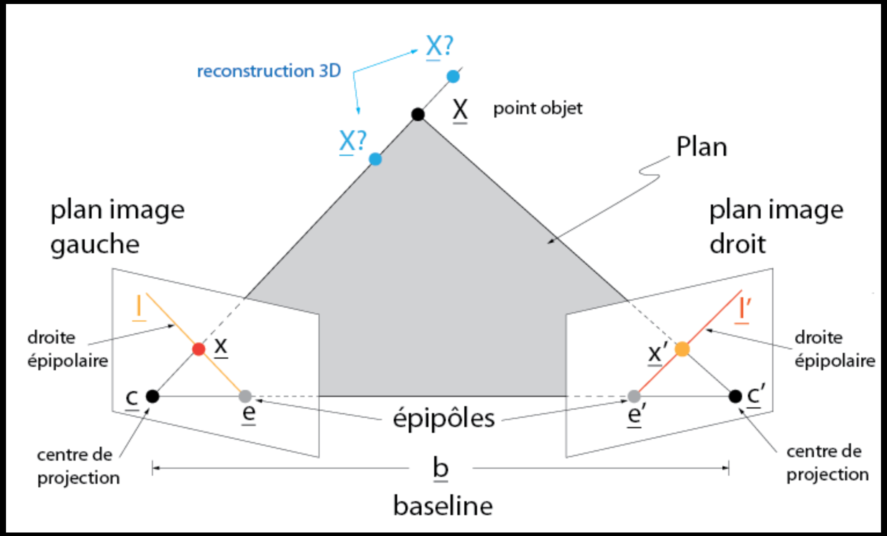
$$\underline{\underline{\tilde{F}}} = \underline{\underline{\widetilde{K}}}'^{-t}[\underline{\tilde{t}}]_x \underline{\underline{\tilde{R}}}\,\underline{\underline{\widetilde{K}}}^{-1} = \left[\underline{\underline{\widetilde{K}}}'\underline{\tilde{t}}\right]_x \underline{\underline{\widetilde{K}}}'\underline{\underline{\tilde{R}}}\,\underline{\underline{\widetilde{K}}}^{-1} = \underline{\underline{\widetilde{K}}}'^{-t}\underline{\underline{\tilde{R}}}\,\underline{\underline{\widetilde{K}}}^t\left[\underline{\underline{\widetilde{K}}}\,\underline{\underline{\tilde{R}}}^t\underline{\tilde{t}}\right]_x \qquad (27)$$

- **Propriété 9** : $\underline{\underline{\tilde{F}}}$ possède **7 degrés de liberté**

  \- $\underline{\underline{\tilde{F}}}$ est une matrice 3 x 3 et possède donc 9 composantes. Or, seulement 8 **rapports** entre les composantes sont indépendants parce que **l'échelle** n'est pas pertinente. $\underline{\underline{\tilde{F}}}$ vérifie aussi la condition $\det(\underline{\underline{\tilde{F}}}) = 0$ car la matrice est singulière. On a donc 9 degrés de liberté -2 contraintes = **7 degrés de liberté**.

- **Propriété 10** : $\underline{\underline{\tilde{F}}}$ est une corrélation, c'est-à-dire une opération qui effectue le mapping d'un point vers une droite. Par conséquent, un point $\underline{\tilde{x}}$ dans l'image de gauche définit une ligne droite $\underline{\tilde{l}}'$ dans l'image de droite.

  Si $\underline{\tilde{l}}$ et $\underline{\tilde{l}}'$ sont des lignes épipolaires **correspondantes** (voir figure) alors n'importe quel point $\underline{\tilde{x}}$ sur $\underline{\tilde{l}}$ est mappé sur $\underline{\tilde{l}}'$. Ceci signifie qu'il n'y a **pas** de mapping **inverse** et $\underline{\underline{\tilde{F}}}$ n'est **pas** une **corrélation propre** (qui devrait pouvoir être inversée).



La matrice **essentielle**

La matrice **essentielle** est une **spécialisation** de la matrice **fondamentale** au cas où les **points images** sont exprimés en **coordonnées normalisées**.

Historiquement, la matrice essentielle a été proposée **avant** la matrice fondamentale. La matrice essentielle possède **moins de degrés de liberté** que la matrice fondamentale et jouit de **propriétés additionnelles**.

Coordonnées normalisées

Soit une matrice caméra exprimée comme:

$$\underline{\underline{\tilde{P}}} = \underline{\underline{\tilde{K}}}\left[\underline{\underline{\tilde{R}}}|t\right] \tag{28}$$

et soit un point image obtenu par projection de perspective :

$$\underline{\tilde{x}} = \underline{\underline{\tilde{P}}}\underline{\tilde{X}} \tag{29}$$

Si la matrice de paramètres intrinsèques est connue (par calibrage par exemple), on peut écrire :

$$\underline{\hat{x}} = \underline{\underline{\tilde{K}}}^{-1}\underline{\tilde{x}} \tag{30}$$

avec (28), (29) et (30), on peut écrire :

$$\underline{\hat{x}} = [\underline{\underline{\tilde{R}}} | \underline{t}] \underline{\tilde{X}} \tag{31}$$

On dit que $\underline{\hat{x}}$ est exprimé en **coordonnées normalisées**.

On peut considérer ce cas comme celui d'une caméra pour laquelle la matrice des paramètres intrinsèques est la matrice identité $\underline{\underline{\tilde{I}}}$ (i.e. la matrice caméra est $\underline{\underline{\tilde{P}}} = [\underline{\underline{\tilde{R}}} | \underline{t}]$) ce qui revient à dire que la focale du sténopé est 1, qu'il n'y a aucune distorsion et que l'échelle est également unitaire.

La caméra $\underline{\underline{\tilde{K}}}^{-1}\underline{\underline{\tilde{P}}} = [\underline{\underline{\tilde{R}}} | \underline{\tilde{t}}]$ est appelée "caméra normalisée".

Considérons maintenant deux caméras normalisées :

$$\underline{\underline{\tilde{P}}} = \left[\underline{\underline{\tilde{I}}}|\underline{\tilde{0}}\right] \qquad\qquad P = \left[\underline{\underline{\tilde{R}}}|\underline{\tilde{t}}\right] \tag{32}$$

Dans ce cas, la matrice fondamentale est, selon la **propriété 8** :

$$\boxed{\underline{\underline{\tilde{F}}} = [\underline{\tilde{t}}]_x \underline{\underline{\tilde{R}}} = \underline{\underline{\tilde{R}}}\left[\underline{\underline{\tilde{R}}}^t \underline{\tilde{t}}\right]_x = \underline{\underline{\tilde{E}}}} \tag{33}$$

La matrice $\underline{\underline{\tilde{E}}}$ est appelée **matrice essentielle** de l'arrangement stéréo.

Supposons maintenant que $\underline{\tilde{x}}$ et $\underline{\tilde{x}}'$ soient en correspondance dans la paire stéréo. On définit $\underline{\underline{\tilde{E}}}$ comme:

$$\underline{\hat{x}}'^{t}\,\underline{\underline{\tilde{E}}}\,\underline{\hat{x}} = 0 \tag{34}$$

En remplaçant $\underline{\hat{x}}$ et $\underline{\hat{x}}'$ par leur expression (voir (30)) dans (34) on obtient :

$$\underline{\tilde{x}}'^{t}\,\underline{\underline{\widetilde{K}}}'^{-t}\,\underline{\underline{\tilde{E}}}\,\underline{\underline{\widetilde{K}}}^{-1}\,\underline{\tilde{x}} = 0 \tag{35}$$

Finalement :

$$\boxed{\underline{\underline{\tilde{E}}} = \underline{\underline{\widetilde{K}}}'^{t}\,\underline{\underline{\tilde{F}}}\,\underline{\underline{\widetilde{K}}}} \tag{36}$$

Propriétés de la matrice essentielle :

- La matrice essentielle $\underline{\underline{\tilde{E}}} = [\underline{\tilde{t}}]_x \underline{\underline{\tilde{R}}}$ compte **5 degrés de liberté** : 3 rotations et 3 translations, ce qui fait 6 degrés de liberté. Comme la matrice contient un facteur d'échelle arbitraire, le nombre de degrés de liberté se réduit à 5.

- Une autre propriété que nous énonçons sans la démontrer est qu'une matrice est une matrice essentielle si et seulement si **deux de ses valeurs singulières sont égales** et **la troisième est nulle** (preuve dans Hartley et Zisserman p. 258)

Estimation de la matrice fondamentale – considérations générales

En général, on n'utilise pas les expressions algébriques (27) pour **calculer** $\underline{\underline{\tilde{F}}}$, mais on utilise plutôt **des correspondances stéréoscopiques** dans les deux images pour **l'estimer**.

On a vu que :

$$\underline{\tilde{x}}'^t \underline{\underline{\tilde{F}}}\, \underline{\tilde{x}} = 0 \tag{37}$$

pour toute paire de correspondances $\underline{\tilde{x}}$ et $\underline{\tilde{x}}'$.

La **propriété** 7 stipule que $\underline{\underline{\tilde{F}}}$ possède 7 degrés de liberté. Si on dispose d'un nombre suffisant de correspondances, il est possible d'estimer $\underline{\underline{\tilde{F}}}$ avec l'équation (37).

Si, pour une correspondance $\underline{\tilde{x}}$ et $\underline{\tilde{x}}'$, on écrit :

$$\underline{\tilde{x}} = (x \quad y \quad 1)^t \qquad\qquad \underline{\tilde{x}}' = (x' \quad y' \quad 1)^t \tag{38}$$

En utilisant (37), pour chaque correspondance “i”, on peut écrire l’équation suivante :

$$x'\,x\,f_{11} + x'\,y\,f_{12} + x'\,f_{13} + y'\,x\,f_{21} + y'\,y\,f_{22} + y'\,f_{23} + x\,f_{31} + y\,f_{32} + f_{33} = 0 \tag{39}$$

On peut simplifier (39) en rassemblant les f_{ij} dans un vecteur:

$$\underline{f} = [f_{11} \quad f_{12} \quad f_{13} \quad f_{21} \quad f_{22} \quad f_{23} \quad f_{31} \quad f_{32} \quad f_{33}]^t \tag{40}$$

L’équation (39) devient :

$$[x'x \quad x'y \quad x' \quad y'x \quad y'y \quad y' \quad x \quad y \quad 1]\underline{f} = 0 \tag{41}$$

Si on considère “n” appariements, on peut écrire un **système d’équations linéaires homogène** pour lequel le vecteur f est l’inconnue à déterminer:

$$\underline{\underline{A}}_{nx9}\underline{f}_{9x1} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \end{bmatrix}_{nx9} \underline{f}_{9x1} = 0 \tag{42}$$

Le système (42) permet d’estimer $\underline{\underline{A}}$ à un facteur d’échelle près. Pour qu’une solution existe, $\underline{\underline{A}}$ doit être au moins de rang 8. Si le rang est exactement 8, la solution est unique et peut être obtenue par des méthodes de solution de systèmes linéaires (la solution est le générateur de l’espace nul droit de $\underline{\underline{A}}$). Ceci n’est valable que s’il n’y a **pas** de **bruit** dans les données images.

En général, les données images $\underline{\tilde{x}}$ et $\underline{\tilde{x}}'$ sont **bruitées** et le rang de $\underline{\underline{A}}$ est plus grand que 8 (il est de 9 car la matrice possède 9 colonnes). Il faut procéder par moindres carrés sur un système surdéterminé (i.e. plus d'équations que d'inconnues). L'approche est similaire à celle utilisée pour estimer une homographie.

La solution par moindres carrés pour $\underline{f}$ est le vecteur singulier correspondant à la plus petite valeur singulière de $\underline{\underline{A}}$, c'est-à-dire la dernière colonne de $\underline{\underline{V}}$ dans

$$svd(\underline{\underline{A}}) = \underline{\underline{U}}\,\underline{\underline{D}}\,\underline{\underline{V}}^t \tag{43}$$

Dans ce cas, la solution trouvée pour $\underline{f}$ minimise $\|\underline{\underline{A}}\underline{f}\|$ soumis à la condition $\|\underline{f}\| = 1$

Contrainte de **singularité** de $\underline{\underline{\tilde{F}}}$

Une propriété importante de $\underline{\underline{\tilde{F}}}$ est qu'elle est singulière de rang 2. De plus, les "null spaces" droit et gauche de $\underline{\underline{\tilde{F}}}$ sont générés par les épipôles (propriété 7)

La plupart des applications de $\underline{\underline{\tilde{F}}}$ reposent sur le fait qu'elle est de rang 2.

Or, en pratique, quand on estime $\underline{\underline{\tilde{F}}}$ par moindres carrés tel que discuté ci-dessus, il est rare que le résultat obtenu soit de rang 2.

Avant d'utiliser $\underline{\underline{\tilde{F}}}$, il faut s'assurer que son rang est bien 2.

La méthode la plus pratique de rendre $\underline{\underline{\tilde{F}}}$ de rang 2 consiste à corriger la solution obtenue pour $\underline{\underline{A}}$ par décomposition en valeurs singulières. Dans ce cas, $\underline{\underline{\tilde{F}}}$ est remplacé par $\underline{\underline{\tilde{F}}}'$ qui minimise la norme de Froebenius $\|\underline{\underline{\tilde{F}}} - \underline{\underline{\tilde{F}}}'\|$ sous la condition que $\det(\underline{\underline{\tilde{F}}}') = 0$.

Une façon d'effectuer cette opération est d'exploiter la décomposition en valeurs singulières à nouveau.

Soit $\underline{\underline{\tilde{F}}} = \underline{\underline{\tilde{U}}}\underline{\underline{\tilde{D}}}\underline{\underline{\tilde{V}}}^t$ la décomposition en valeurs singulières de $\underline{\underline{\tilde{F}}}$ où $\underline{\underline{D}} = diag(r,s,t)$ avec $r \geq s \geq t$. Alors

$$\underline{\underline{\tilde{F}}} = \underline{\underline{U}}\ diag(r,s,0)\ \underline{\underline{V}}^t \tag{44}$$

minimise $\|\underline{\underline{\tilde{F}}} - \underline{\underline{\tilde{F}}}'\|$.

Estimation de la matrice fondamentale avec l’algorithme à **8 points normalisé**

La méthode la plus pratique pour estimer $\underline{\underline{\tilde{F}}}$ à partir de “n” correspondances $\underline{\tilde{x}} \leftrightarrow \underline{\tilde{x}}'$ avec (43) est l’**algorithme à 8 points normalisé**.

En prenant certaines précautions assurant la **stabilité numérique** de la solution, cet algorithme à 8 points fonctionne généralement bien.

La clé de l’algorithme à 8 points réside en une normalisation adéquate des données (i.e. les paires de correspondances $\underline{\tilde{x}} \leftrightarrow \underline{\tilde{x}}'$) avant de construire le système (42) puis de solutionner par décomposition en valeurs singulières.

Algorithme à 8 points

Objectif :

En supposant que l'on dispose de $n \geq 8$ correspondances $\{\underline{\tilde{x}}_i \leftrightarrow \underline{\tilde{x}}'_i\}$, calculer $\underline{\underline{\tilde{F}}}$ telle que $\underline{\tilde{x}}'^t_i \underline{\underline{\tilde{F}}}\ \underline{\tilde{x}}_i = 0 \;\; \forall \;\; \underline{\tilde{x}}_i \leftrightarrow \tilde{x}'_i \;\; \in \; \{\underline{\tilde{x}}_i \leftrightarrow \tilde{x}'_i\}$.

Dans ce qui suit, nous allons détailler les étapes de l'algorithme à 8 points selon ce qui est proposé par ***Hartley et Zisserman***.

Algorithme à 8 points

- **Etape 1** : **Normalisation des données**. Cette opération consiste à transformer les données d'entrée $\underline{\tilde{x}}_i$ et $\underline{\tilde{x}}'_i$ de sorte que :

$$\underline{\tilde{\tilde{x}}}_i = \underline{\underline{\tilde{T}}}\,\underline{\tilde{x}}_i \qquad \underline{\tilde{\tilde{x}}}'_i = \underline{\underline{\tilde{T}}}'\,\underline{\tilde{x}}'_i \tag{45}$$

Les matrices $\underline{\underline{\tilde{T}}}$ et $\underline{\underline{\tilde{T}}}'$ consistent en une **translation** et une **mise à l'échelle**

La **translation** fait en sorte que **l'origine** des points en correspondance soit située en leur **centre de masse**.

La **mise à l'échelle** fait en sorte que la **distance RMS** (racine carrée de la moyenne des distances au carré) des points à l'origine des points en correspondance soit $\sqrt{2}$.

$$d_{RMS} = \sqrt{\frac{1}{n}\sum_{1}^{n} \|\underline{\tilde{x}}_{-} - \underline{\tilde{x}}_0\|^2} \tag{46}$$

Algorithme à 8 points

- **Etape 2** :

A) Calculer la matrice fondamentale $\underline{\underline{\tilde{\hat{F}}}}$ avec les correspondances normalisées $\{\hat{x}_i \leftrightarrow \hat{x}'_i\}_n \ (n \geq 8)$

$\underline{\underline{\tilde{F}}}$ est le vecteur singulier correspondant à la plus petite valeur singulière de $\underline{\underline{\hat{A}}}$ (voir (42) et suivantes) (i.e. la dernière colonne de $\underline{\underline{V}}$ dans $svd(\underline{\underline{\hat{A}}}) = \underline{\underline{U}}\,\underline{\underline{D}}\,\underline{\underline{V}}^t$)

B) Correction de : remplacer $\underline{\underline{\tilde{\hat{F}}}}$ par $\underline{\underline{\tilde{\hat{F}}}}'$ tel que $\det\left(\underline{\underline{\tilde{\hat{F}}}}'\right) = 0$ en se basant sur (44).

Algorithme à 8 points

- **Etape 3** : Dénormalisation

  Calculer (en utilisant $\underline{\underline{\widetilde{T'}}}$ et $\underline{\underline{\tilde{T}}}$ de (45)):

$$\underline{\underline{\tilde{F}}} = \underline{\underline{\tilde{T}}}'^{t} \underline{\underline{\tilde{\tilde{F}}}}' \; \underline{\underline{\tilde{T}}} \qquad (47)$$

  Ce résultat s'obtient comme suit :

$$\underline{\tilde{\tilde{x}}}_i = \underline{\underline{\tilde{T}}} \; \underline{\tilde{x}}_i \;\; , \;\; \underline{\tilde{\tilde{x}}}'_i = \underline{\underline{\tilde{T}}}' \; \underline{\tilde{x}}'_i \qquad (48)$$

$$\underline{\tilde{x}}'^{t}_i \underline{\underline{\tilde{F}}} \; \underline{\tilde{x}}_i = 0$$

$$\underline{\tilde{x}}_i = \underline{\underline{\tilde{T}}}^{-1} \underline{\tilde{\tilde{x}}}_i \;\; , \;\; \underline{\tilde{x}}'_i = \underline{\underline{\tilde{T}}'}^{-1} \underline{\tilde{\tilde{x}}}'_i \; \rightarrow \; \underline{\widetilde{x'}}^{t}_i = \underline{\widetilde{\widetilde{x'}}}^{t}_i \underline{\underline{\tilde{T}}}^{-1t} \qquad (49)$$

  En remplaçant (49) dans (48) on obtient (47)

$$\underline{\tilde{\tilde{x}}}'^{t}_i \underline{\underline{\tilde{T}}}'^{-1t} \underline{\underline{\tilde{F}}} \, \underline{\underline{\tilde{T}}}^{-1} \underline{\tilde{\tilde{x}}} \; \rightarrow \; \underline{\underline{\hat{F}}}' = \underline{\underline{\tilde{T}}}'^{-1} \underline{\underline{\tilde{F}}} \, \underline{\underline{\tilde{T}}}^{-1} \; \rightarrow \; \underline{\underline{\tilde{F}}} = \underline{\underline{\tilde{T}}}'^{t} \underline{\underline{\tilde{\tilde{F}}}}' \; \underline{\underline{\tilde{T}}}$$

Remarque importante :

L’algorithme décrit ci-dessus minimise **l’erreur algébrique**. Il existe des algorithmes pour minimiser **l’erreur géométrique** (Hartley et Zisserman p. 284)

Résultat complémentaire (sans démonstration):

Les matrices caméras correspondant à la matrice fondamentale d’un arrangement stéréo sont données par les expressions suivantes (Hartley et Zisserman p. 256)

$$\underline{\underline{\tilde{P}}} = \left[\underline{\underline{\tilde{I}}}|\vec{\tilde{0}}\right] \qquad \underline{\underline{\tilde{P}}}' = \left[[\underline{e}']_x\underline{\underline{\tilde{F}}}|\underline{e}'\right] \tag{50}$$