



Vision par ordinateur

Vision 3D *(première partie)*

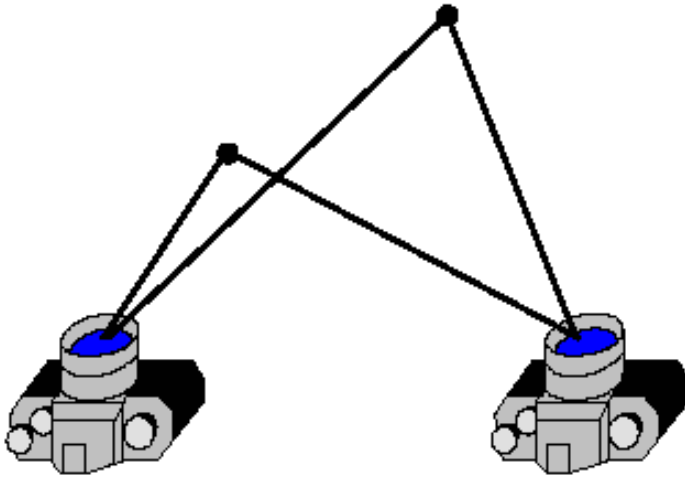
Alain Boucher – IFI



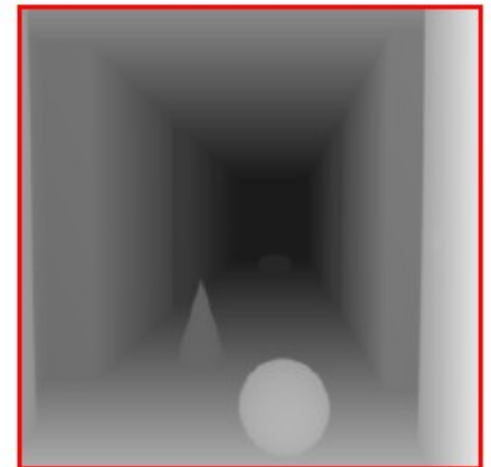
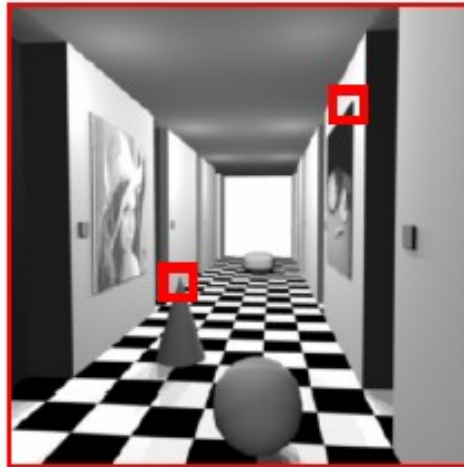
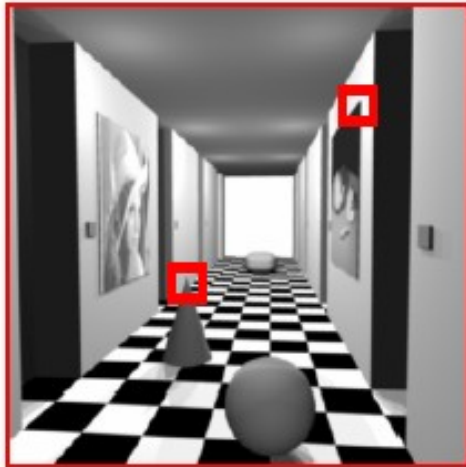
Introduction

- Même en calibrant une caméra, nous ne connaissons pas la profondeur des points de l'image
 - Tous les points sont connus à un facteur d'échelle près
 - Nous n'avons pas une réelle vue en 3D
 - *Cf cours Calibration*
- A chaque point correspond une droite 3D dans la scène
- En utilisant deux ou plusieurs vues, on peut évaluer la position 3D des objets
 - Intersection de 2 droites 3D provenant des deux caméras
 - C'est pour cette raison que l'humain a deux yeux pour voir
- On obtient une vue supplémentaire par l'ajout d'une caméra ou en déplaçant la même caméra

Vision stéréoscopique



Nous allons évaluer la
position d'un point dans
l'espace à partir de
plusieurs caméras



Carte des profondeurs

Vision stéréoscopique



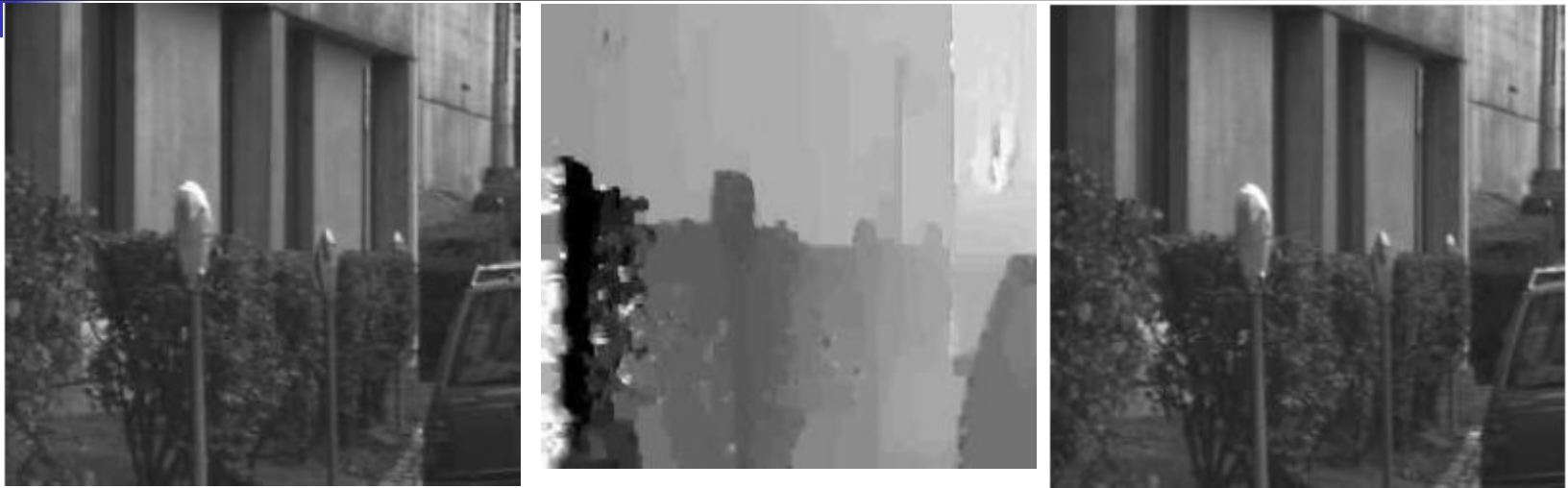
Vue gauche

Carte des profondeurs



Vue droite

Application : interpolation de vue



*Addition des
deux vues*



*Interpolation
des deux vues*



Application : Interaction 2D - 3D

- Aussi appelée *réalité augmentée*



- But
 - Intégration d'objets synthétiques 3D dans la vidéo
 - Doit être automatique !
- Donc...
 - La profondeur doit être récupérée automatiquement



Problèmes de la vision 3D

- **Calibration stéréoscopique**

- Géométrie épipolaire
- Orientation relative des caméras

- **Rectification d'images**

- Transformation d'une image pour simplifier la mise en correspondance

- **Mise en correspondance (appariement)**

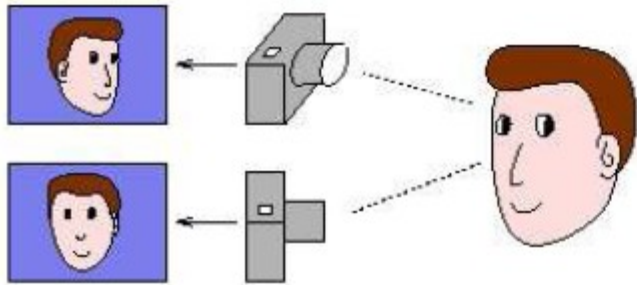
- Trouver les paires de points correspondantes dans chaque image

- **Reconstruction 3D**

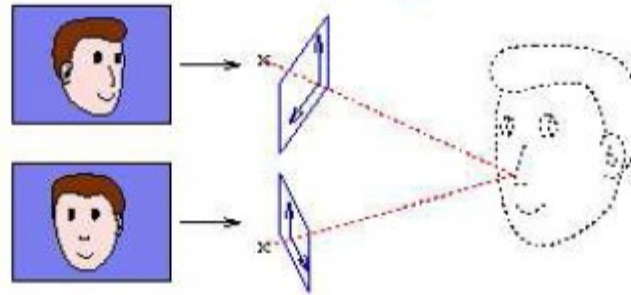
- Triangulation pour reconstruire un modèle 3D

Problèmes de la vision 3D

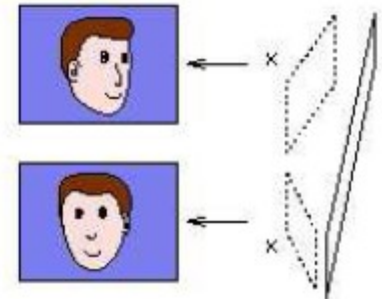
1. acquisition



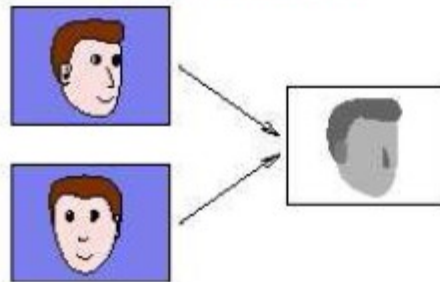
2. calibrage



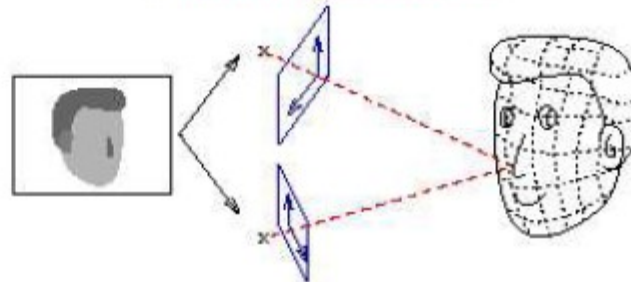
3. rectification



4. corrélation



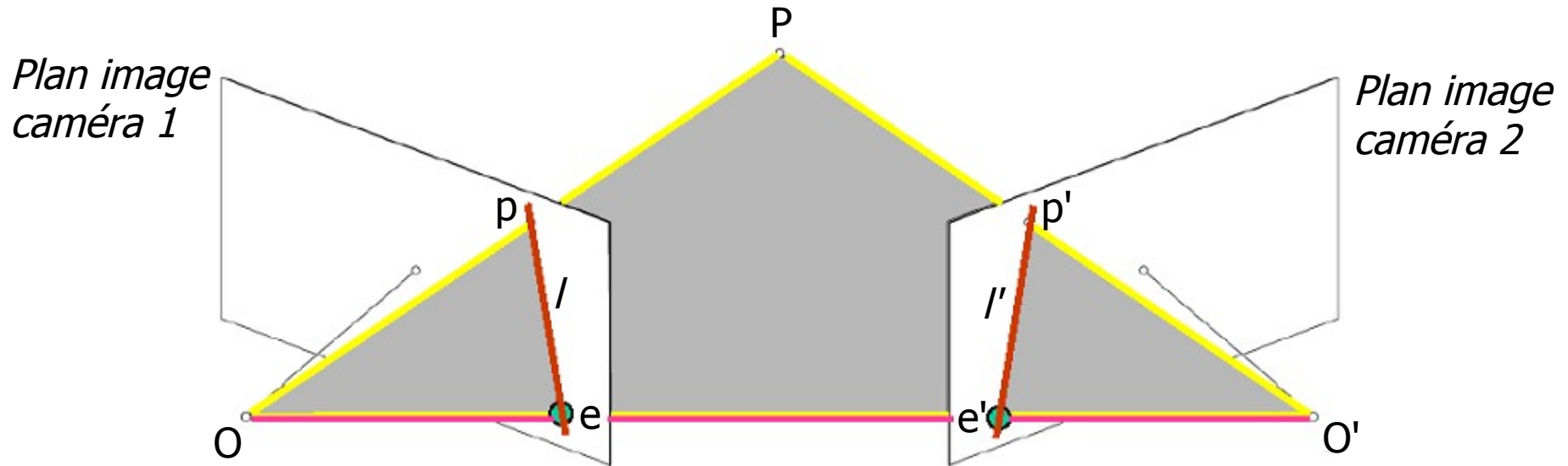
5. reconstruction...





Géométrie épipolaire

Géométrie épipolaire



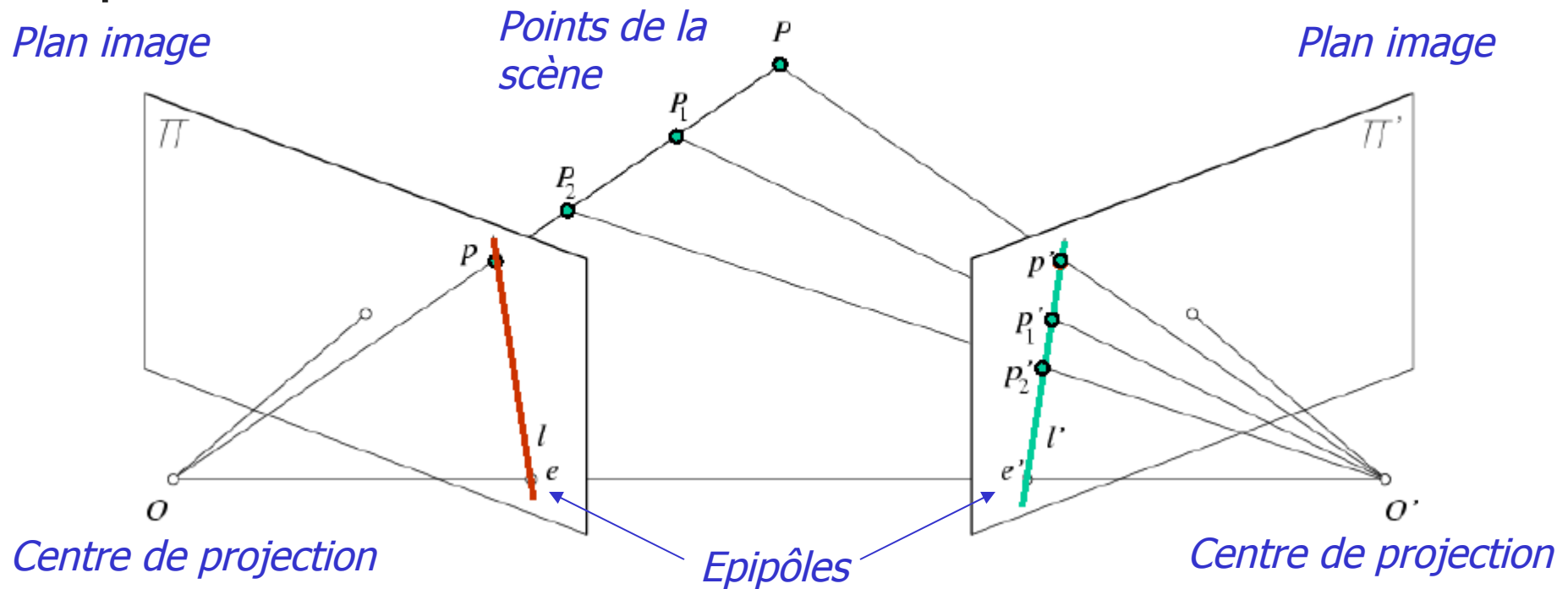
- 5 points dans le même plan : P , p , p' , O et O'
- les droites l et l' sont les **droites épipolaires** (conjuguées)
- les points e et e' sont respectivement les **épipôles** gauche et droit
- e est en fait O' vu dans l'image gauche
- e' est O vu dans l'image droite



Droites épipolaires et épipôles

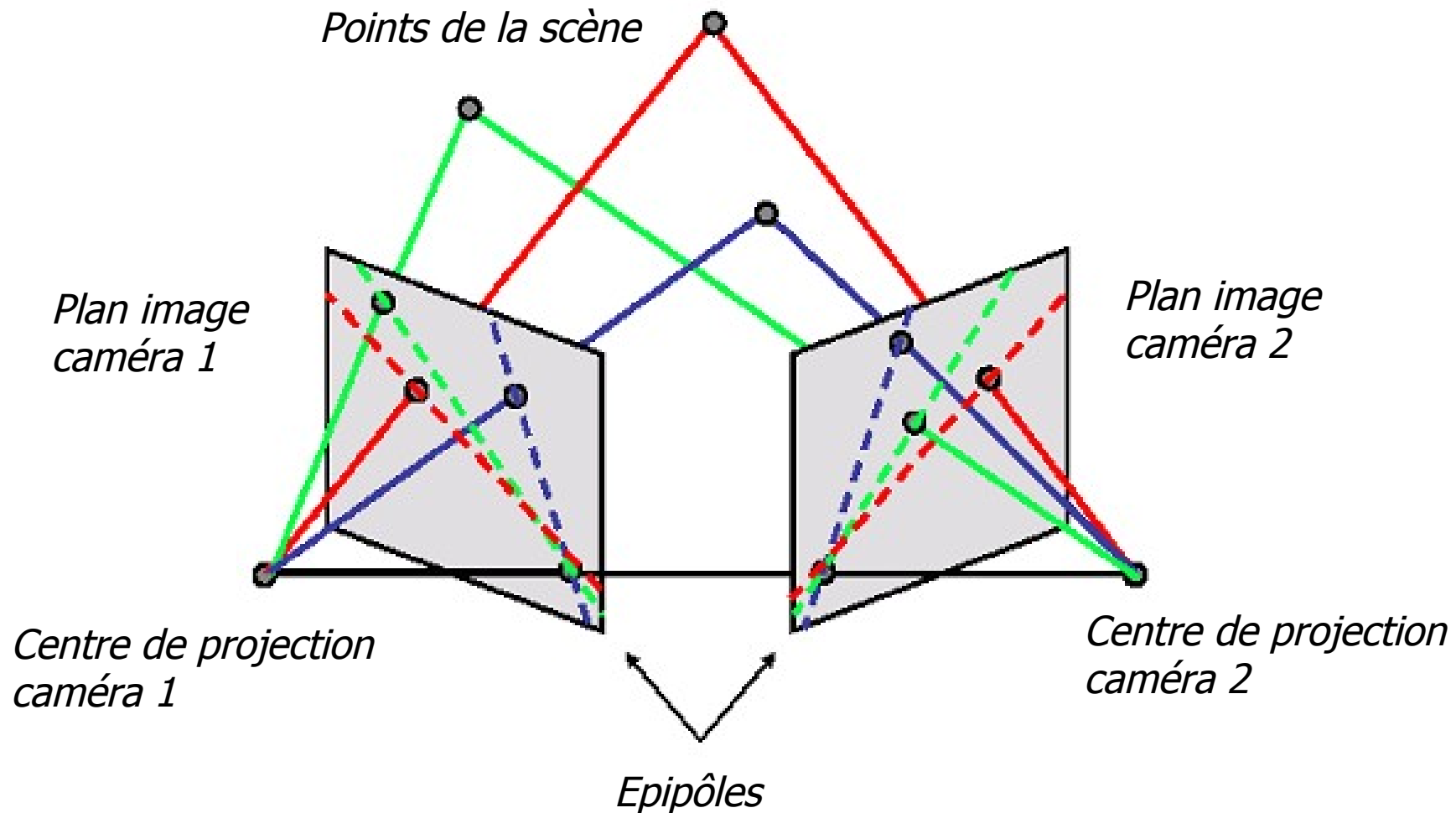
- Un **épipôle** est le centre de projection d'une caméra vue dans le plan image de l'autre caméra
 - Une seule droite épipolaire passe par chaque point des images (sauf aux épipôles)
- Une **droite épipolaire** est formé par l'épipôle et un point de l'image (multitude de droites épipolaires par image)
 - Les droites épipolaires passent toutes par les épipôles

Droites épipolaires et épipôles



Un point dans l'image gauche se situe sur la droite épipolaire correspondante dans l'image droite

Géométrie épipolaire

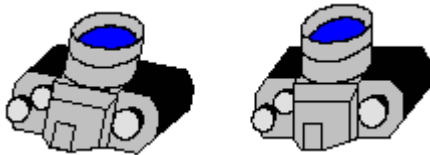




Géométrie épipolaire

- Le correspondant d'un point de l'image gauche dans l'image droite **est contraint sur la droite épipolaire**
 - Ce qu'on appelle la **contrainte épipolaire**
 - Hypothèse d'aucune autre "distorsion" que la projection perspective
 - Cela limite l'espace de recherche pour trouver le correspondant d'un point
- Lorsque les deux plans images des caméras sont parallèles :
 - les droites épipolaires sont parallèles
 - les épiholes sont à l'infini

Exemples de droites épipolaires





Paramètres de la géométrie

- **Problème :**
 - Connaître les épipôles et les paramètres de cette géométrie épipolaire
- **Solution :**
 - Calibration d'un système à deux caméras fixes
 - Connaître la transformation d'une caméra vers l'autre
- Nous souhaitons connaître la transformation qui permet de passer d'une caméra vers l'autre caméra
 - Correspondances entre les points des deux caméras
Correspondances entre les points des deux images



Cas des caméras calibrées

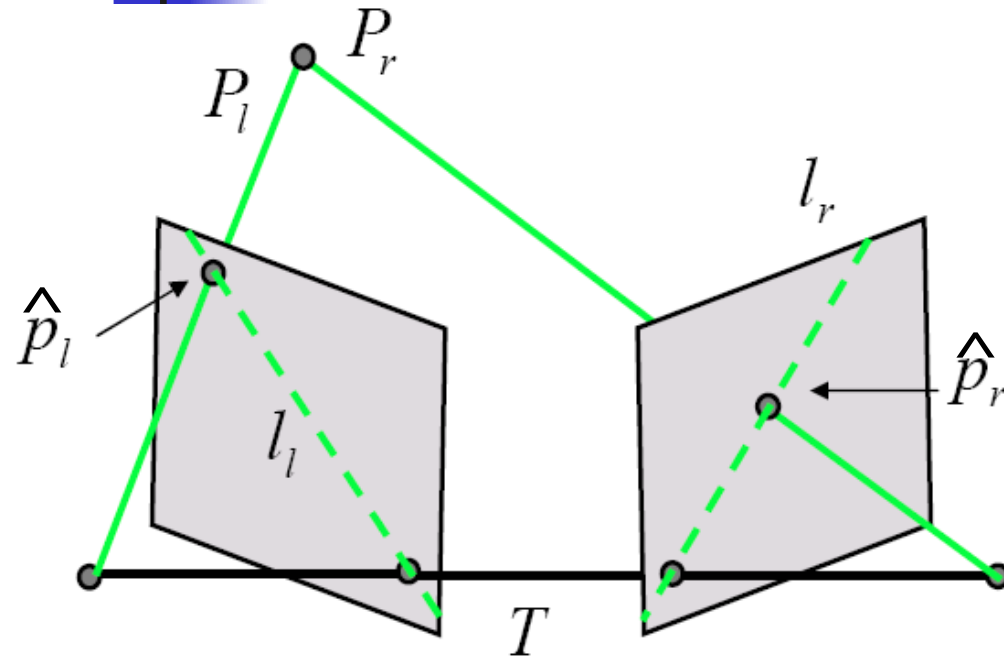
- Si nous connaissons les **matrices de calibration** de chaque caméra (*cf cours calibration*) :

$$P_i = M \cdot P_e = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \cdot P_e$$

- Alors nous pouvons calculer la **matrice essentielle (E)** pour ces deux caméras :
 - Relation entre les points de l'image définis en coordonnées du **référentiel caméra** et les droites épipolaires :

$$\hat{p}_l^T E \hat{p}_r = 0$$

Matrice essentielle E (cas calibré)



$$\hat{p}_l^T \left[T \times (R \hat{p}_r) \right] = 0$$

$$\hat{p}_l^T \underbrace{TR}_{\text{matrice essentielle}} \hat{p}_r = 0$$

matrice essentielle

$$\hat{p}_l^T E \hat{p}_r = 0$$

*Les droites épipolaires
sont données par :*

$$l_l = E \hat{p}_r$$

$$l_r = E^T \hat{p}_l$$

$\left. \begin{matrix} \hat{p}_l \\ \hat{p}_r \end{matrix} \right\}$ *Coordonnées des
pixels dans le
référentiel caméra*

Mire de calibration (cas calibré - E)

On connaît pour chaque point :

- sa position (x_1, y_1) en pixels dans l'image gauche
- sa position (x_2, y_2) en pixels dans l'image droite
- sa position réelle (x, y, z) en 3D dans l'environnement

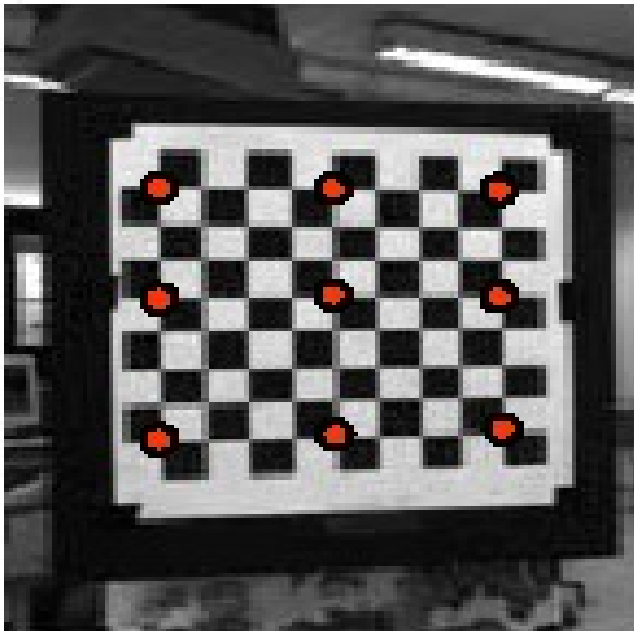


Image caméra gauche

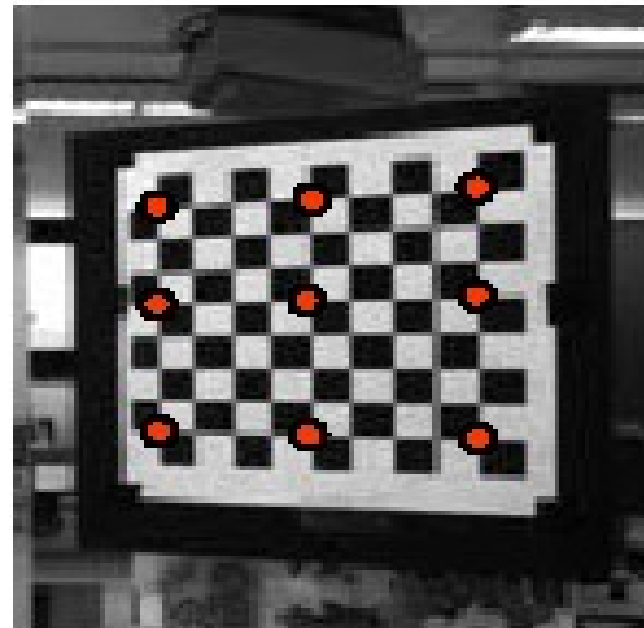


Image caméra droite



Cas des caméras non-calibrées

- **E** est une matrice singulière (rang 2) contenant 5 degrés de liberté
- Si nous ne connaissons pas les **matrices de calibration** de chaque caméra, alors nous n'avons pas la relation entre les coordonnées dans le *référentiel caméra* et celles dans le *référentiel image*
- Alors nous allons calculer la **matrice fondamentale (F)** pour ces deux caméras :
 - Relation entre les points de l'image définis en coordonnées du *référentiel image* et les droites épipolaires :

$$\bar{p}_l^T F \bar{p}_r = 0$$

Matrice fondamentale F (cas non-calibré)

$$\hat{p}_l^T E \hat{p}_r = 0$$

Dans le référentiel image :

$$\bar{p}_l = K_l \hat{p}_l$$

$$\bar{p}_r = K_r \hat{p}_r$$

$\left. \begin{array}{c} \hat{p}_l \\ \hat{p}_r \end{array} \right\}$ Coordonnées des pixels dans le référentiel caméra

K = Matrice de calibration de la caméra (*caméra* \rightarrow *image*)

M = Matrice de calibration de la caméra (*environnement* \rightarrow *image*)

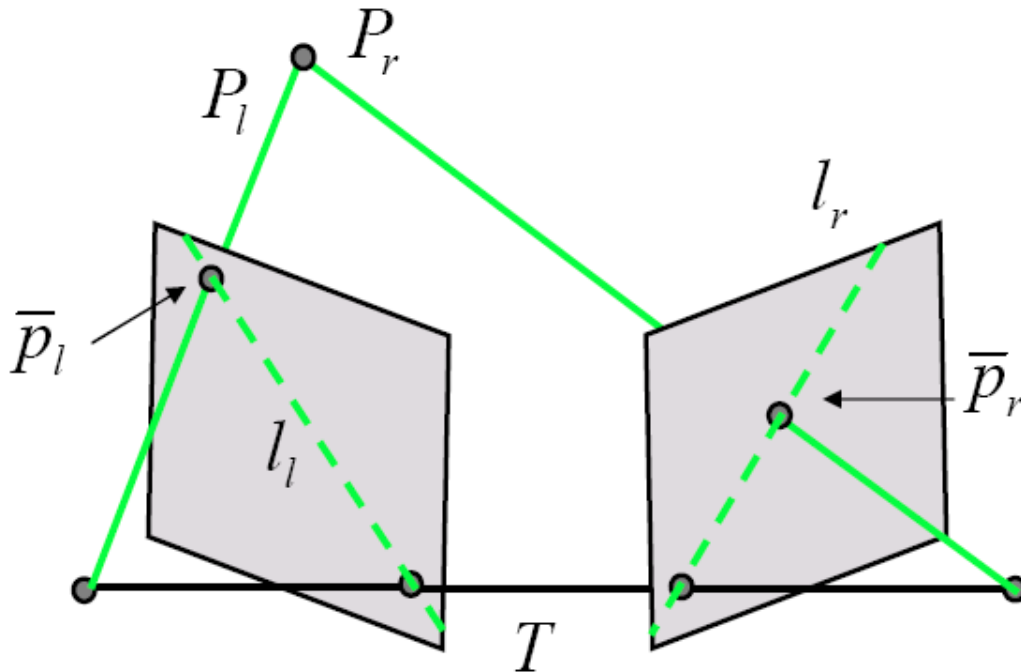
$$\bar{p}_l^T \underbrace{K_l^T E K_r}_{\text{matrice fondamentale } F} \bar{p}_r = 0$$

matrice fondamentale F

$$\bar{p}_l^T F \bar{p}_r = 0$$

$\left. \begin{array}{c} \bar{p}_l \\ \bar{p}_r \end{array} \right\}$ Coordonnées des pixels dans le référentiel image

Matrice fondamentale F (cas non-calibré)



$$\hat{p}_l^T E \hat{p}_r = 0$$

$$\bar{p}_l^T \underbrace{K_l^T E K_r}_{\text{matrice fondamentale } F} \bar{p}_r = 0$$

matrice fondamentale F

$$\bar{p}_l^T F \bar{p}_r = 0$$

*Les droites épipolaires
sont données par :*

$$l_l = F \bar{p}_r$$

$$l_r = F^T \bar{p}_l$$

$\left. \begin{array}{l} \bar{p}_l \\ \bar{p}_r \end{array} \right\}$ *Coordonnées des
pixels dans le
référentiel image*

Mire de calibration (cas non-calibré - F)

On connaît pour chaque point :

- sa position (x_1, y_1) en pixels dans l'image gauche
- sa position (x_2, y_2) en pixels dans l'image droite

Et c'est tout. On ne connaît rien sur sa position réelle (x, y, z) .

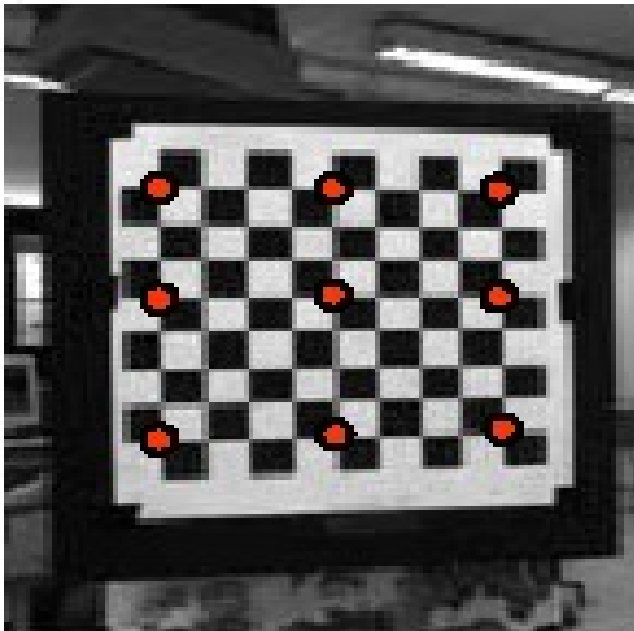


Image caméra gauche

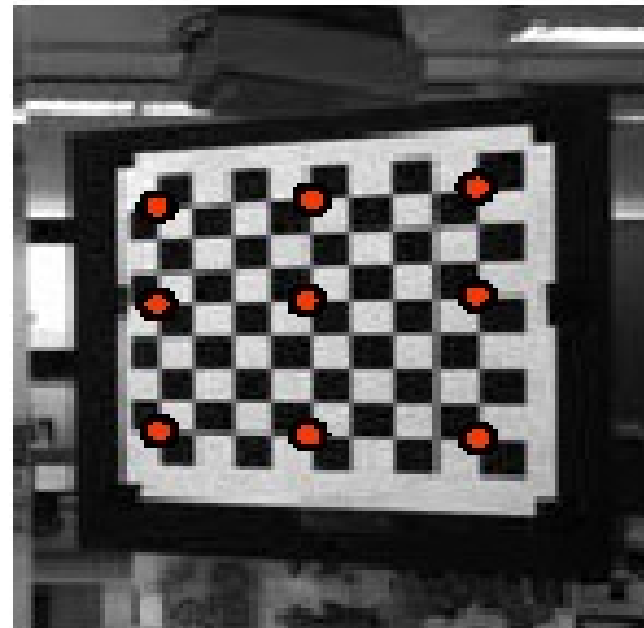


Image caméra droite



Estimation de la matrice fondamentale

- Procédure de ***calibration faible***
 - Calibration des deux caméras à la fois sans connaître et sans rechercher les paramètres extrinsèques et intrinsèques de chacune des deux caméras
 - i.e. estimer F sans connaître les deux K (ou les deux M)
- F : matrice 3x3 de rang 2 contenant 8 inconnues
 - Définition à un facteur d'échelle près
- Pour résoudre avec un ensemble (>8) correspondances :

$$\min_F \sum \left(\bar{p}_l^T F \bar{p}_r \right)^2 \quad \text{avec la contrainte de } \det(F) = 0$$
- Algorithme de décomposition en valeurs singulières
 - *SVD : Singular Value Decomposition*



Analogie avec un bébé

- Un bébé de quelques jours n'arrive pas à focaliser son regard sur un objet (*les yeux suivent dans des directions différentes*)
 - *Les deux yeux sont non-calibrés*
- Plus tard, il réussit à diriger ses deux yeux pour regarder un objet. Mais il n'arrive pas à attraper cet objet.
 - *Calibration des yeux dans le référentiel image seulement*
 - *Le cerveau connaît la matrice F du système*
- Enfin, il réussit à regarder un objet et à saisir cet objet
 - *Calibration des yeux dans le référentiel caméra (yeux)*
 - *Le cerveau connaît la matrice E du système*



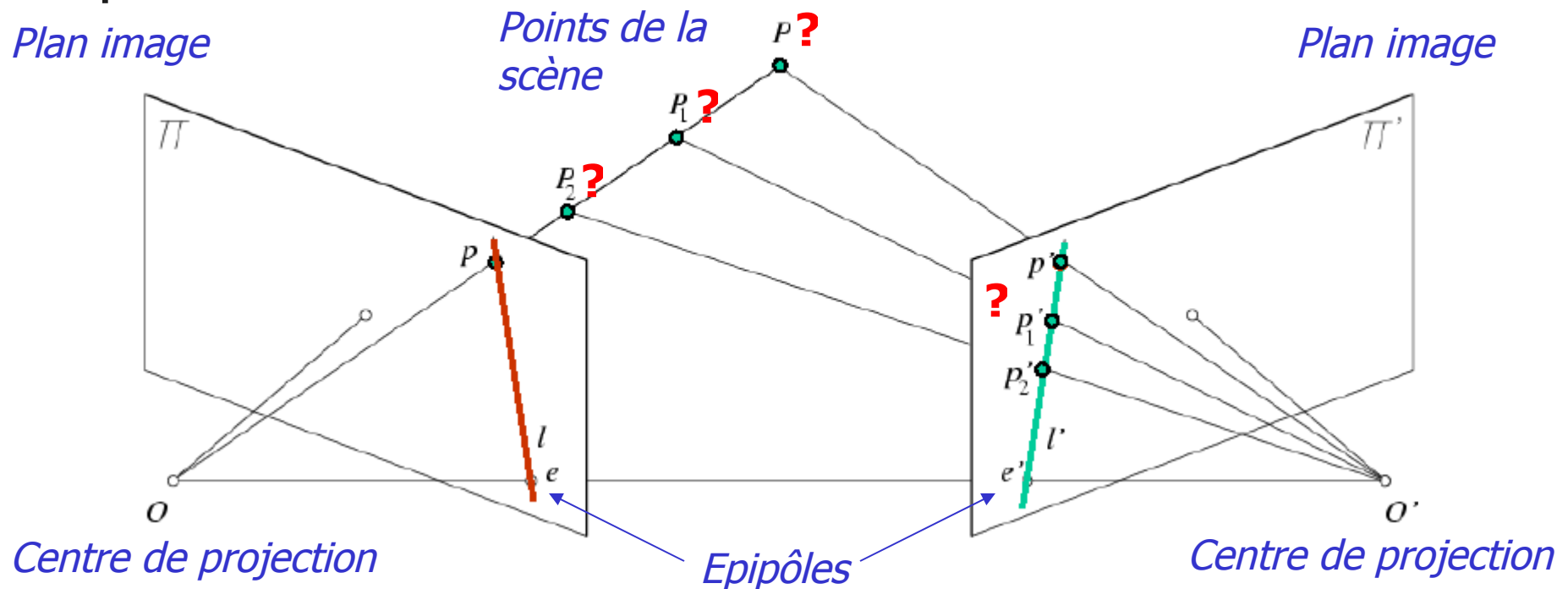
Mise en correspondance (appariement)



Problème de la correspondance

- Même si nous arrivons à calculer tous les paramètres de la géométrie épipolaire (matrices E ou F), nous n'avons pas automatiquement la correspondance entre les points
 - Sachant
 - (1) que p_l et p_r correspondent au même point dans la scène
 - (2) la matrice E ou la matrice F
 - Alors
 - nous pouvons retrouver la profondeur (z) réelle du point
- **Problème**
 - Retrouver la correspondance entre les paires de points p_l et p_r

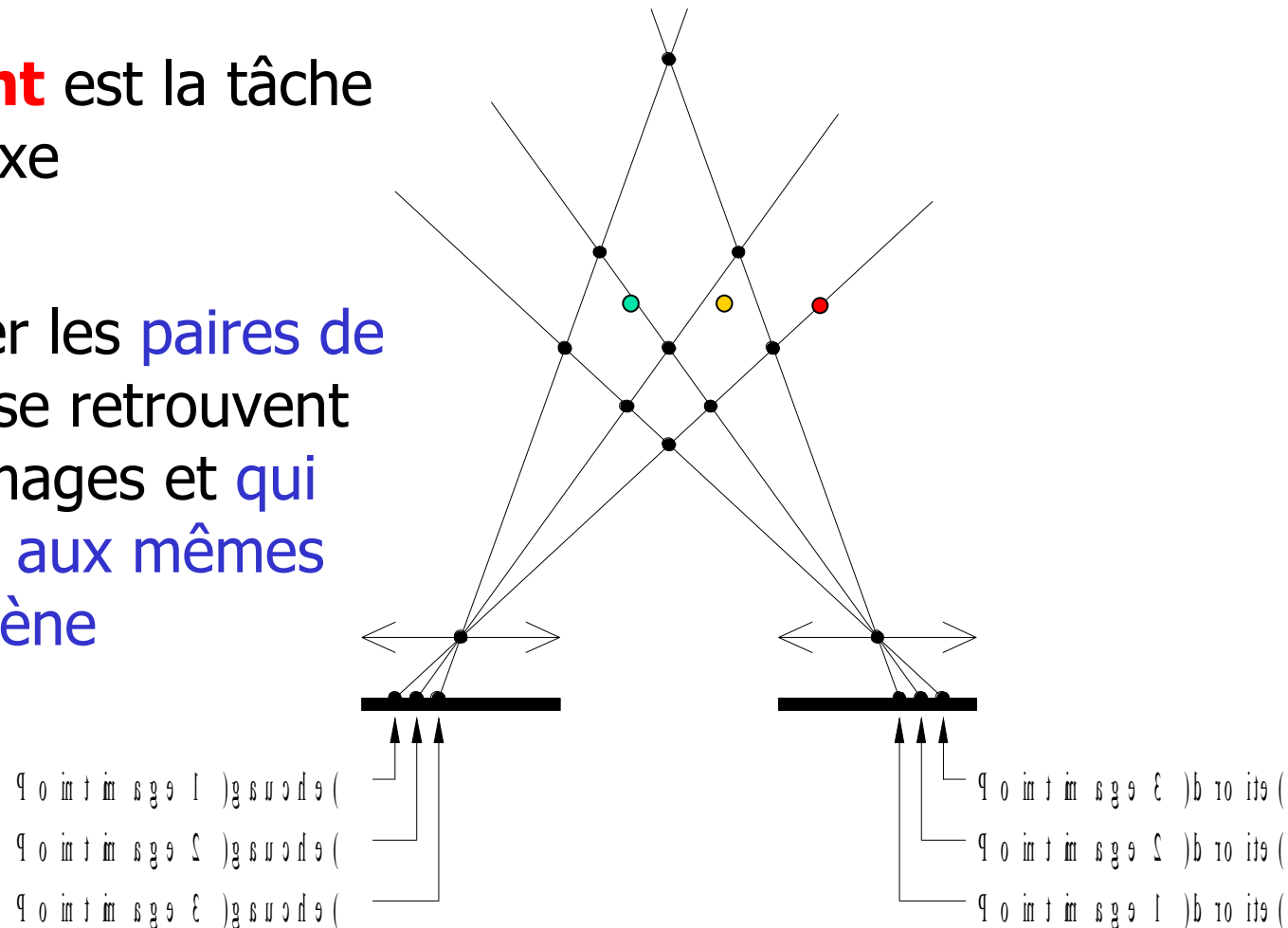
Problème de la correspondance



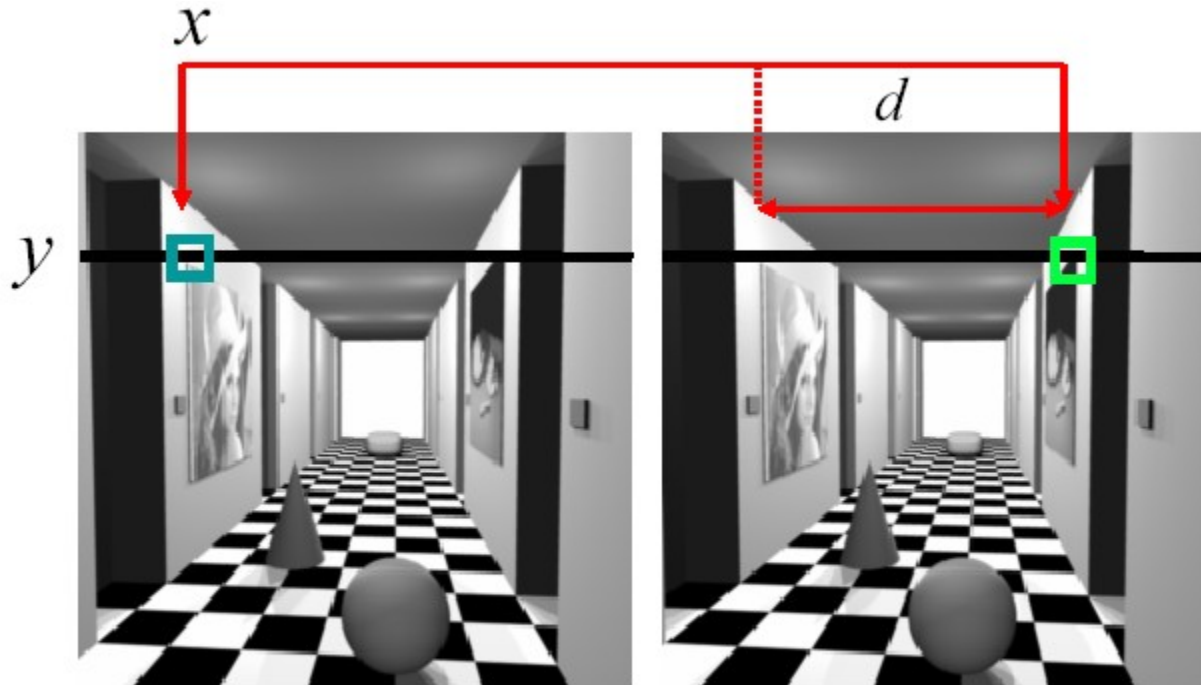
- A partir d'un point dans l'image gauche et de la géométrie épipolaire, quel est le bon point correspondant dans l'image droite ?
 - Problème de l'appariement ou mise en correspondance

Le problème de l'appariement

- L'**appariement** est la tâche la plus complexe
- Il faut identifier les **paires de primitives** qui se retrouvent sur les deux images et **qui correspondent aux mêmes points de la scène**



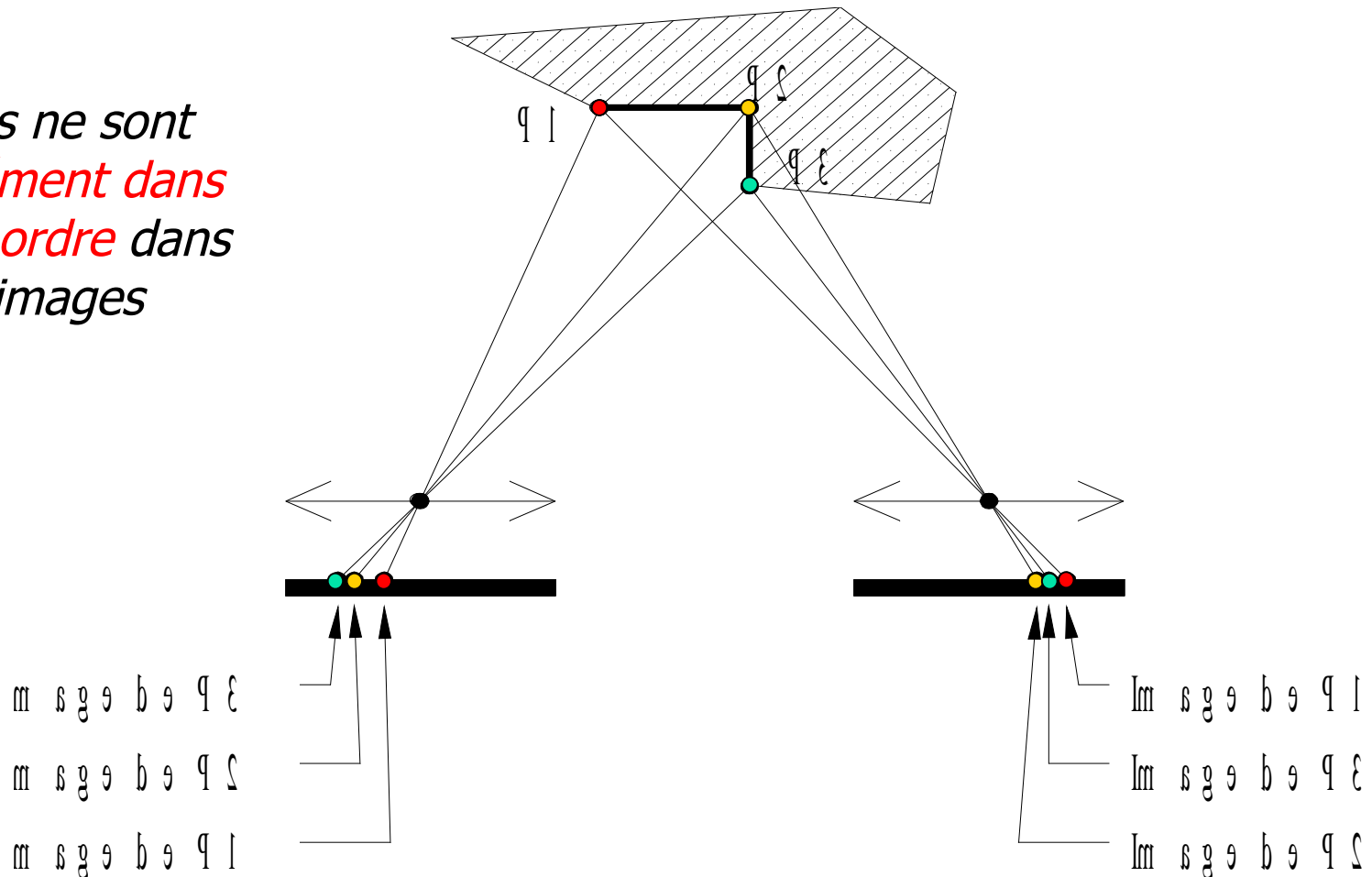
Correspondance entre primitives



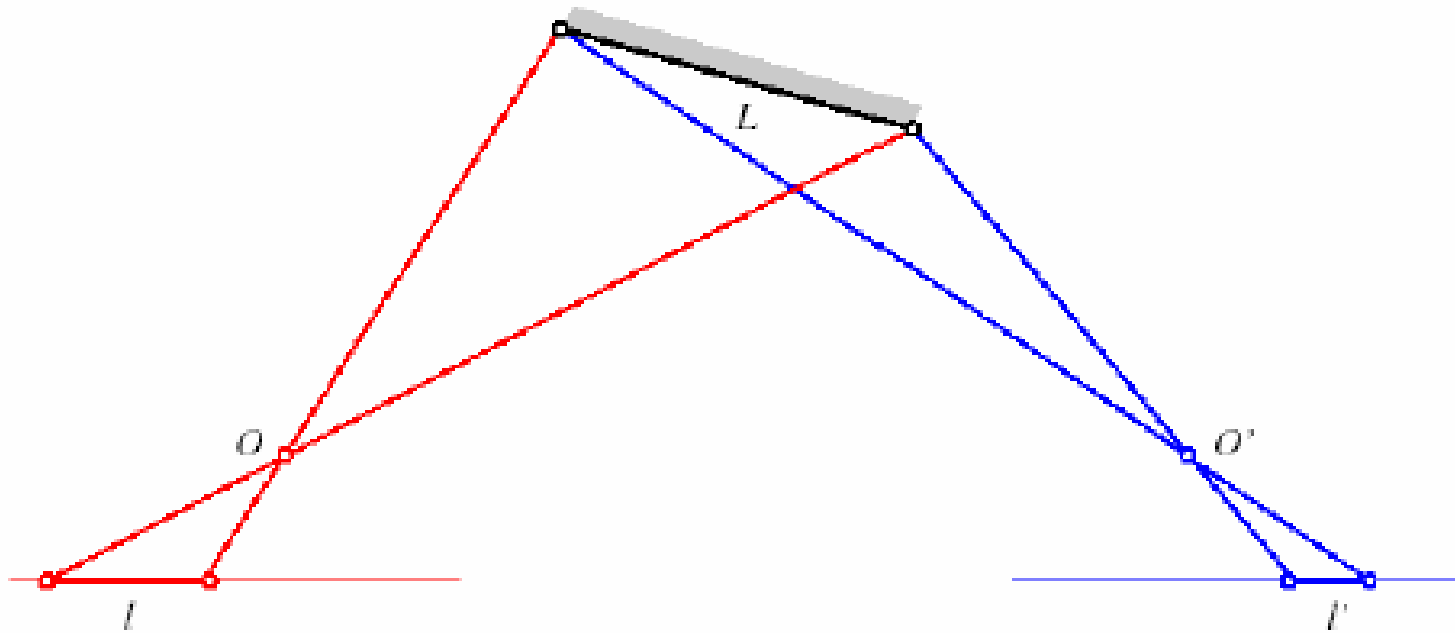
- Il faut trouver les paires de primitives correspondantes dans les deux images
- Chaque primitive est vue avec un angle et un contexte différent par les deux caméras

Difficultés de l'appariement

*Les points ne sont
pas forcément dans
le même ordre dans
les deux images*

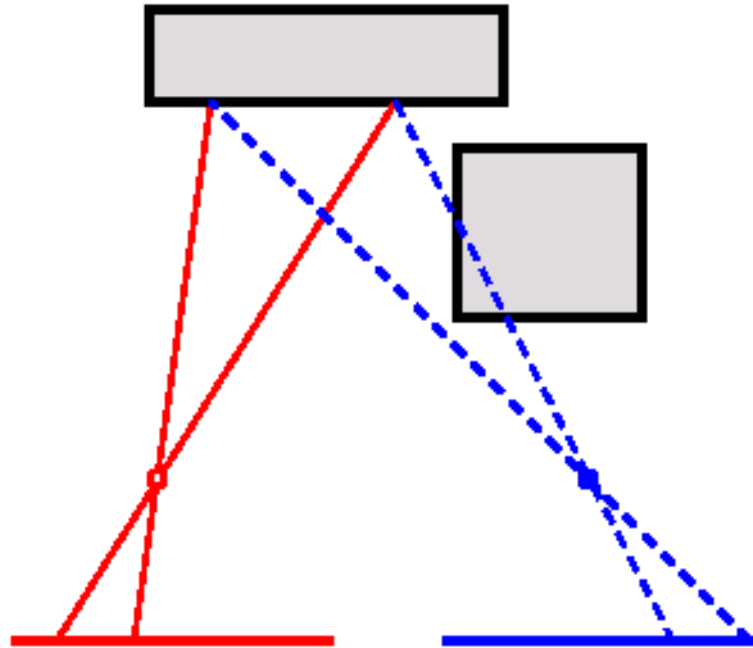


Difficultés de l'appariement



Les tailles et distances ne sont pas les mêmes d'une image à l'autre

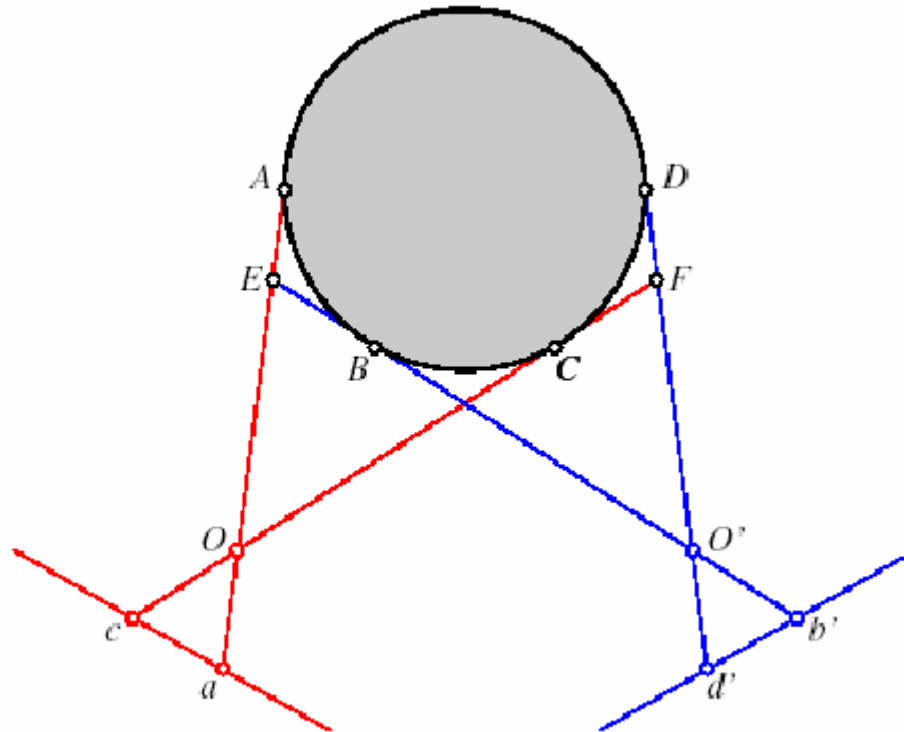
Difficultés de l'appariement



Occlusions : des objets ou parties d'objets sont cachés

La correspondance n'existe pas dans ce cas

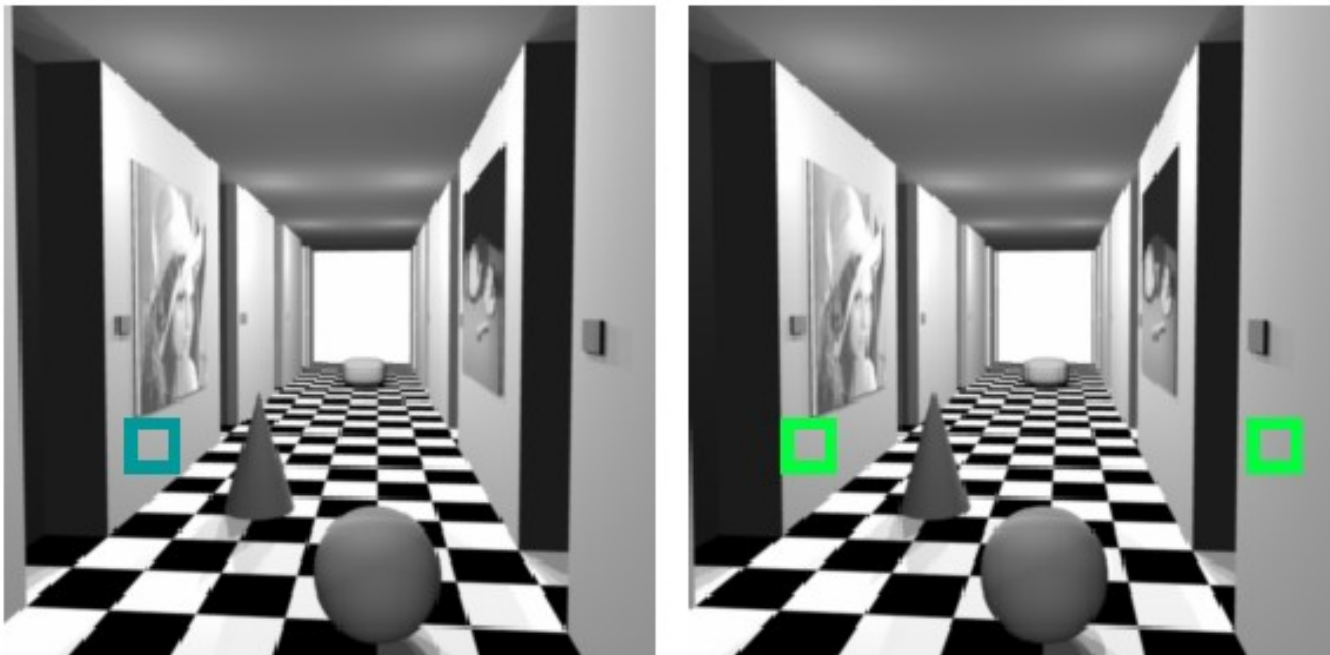
Difficultés de l'appariement



Dans le cas d'objets sphériques ou curvilignes, les arêtes perçues ne correspondent pas d'une image à l'autre

Difficultés de l'appariement

- Correspondance difficile (voire impossible) dans les zones faiblement texturées
- Ceci quelque soit l'approche utilisée





La mise en correspondance

- **Hypothèses :**

- Les points de la scène sont visibles (en général) des deux points de vue (pas toujours vrai)
- Les points correspondants sont similaires dans les images
 - valide surtout si les points sont beaucoup plus loin que la base de triangulation

- **Approches :**

- (1) Correspondance de caractéristiques*

- ensemble éparse de correspondances
 - exemple: lignes, coins, jonctions, ...

- (2) Corrélation de voisinage*

- permet un ensemble dense de correspondances



Matrices E ou F connues ou non ?

- **Si** on connaît la géométrie épipolaire (matrices E ou F)
Alors on recherche les correspondances le long des droites épipolaires
Sinon on doit rechercher des correspondances dans toute l'image et non pas que sur la droite épipolaire
- Dans le cas de caméras proches et/ou d'objets éloignés, la correspondance est recherchée d'abord dans le **voisinage proche**



Mise en correspondance (appariement)

Correspondance de caractéristiques



Correspondance de caractéristiques

- Il s'agit d'une correspondance qui n'est pas dense
 - C'est-à-dire pas beaucoup de points de correspondance
- Rechercher des primitives extraites des images :
 - droites et points d'arêtes (*perpendiculaires aux droites épipolaires*)
 - utile dans des environnements structurés comme l'intérieur de bâtiments
 - coins (détecteur de Harris, SIFT ou autres)
 - centre de cercles
- On peut aussi examiner la corrélation autour de ces caractéristiques



Correspondance de caractéristiques

- Trouver les points caractéristiques dans chaque image
- Fonction de distance basée sur la distance, l'orientation, les niveaux de gris, ...
- Pour chaque point caractéristique de l'image gauche, calculer la similarité avec tous les points caractéristiques de l'image droite
 - Trouver les paires de points (gauche, droite) qui maximise la similarité

Exemple basé sur les coins (Harris)



Image gauche



Image droite

Extraction des coins de Harris

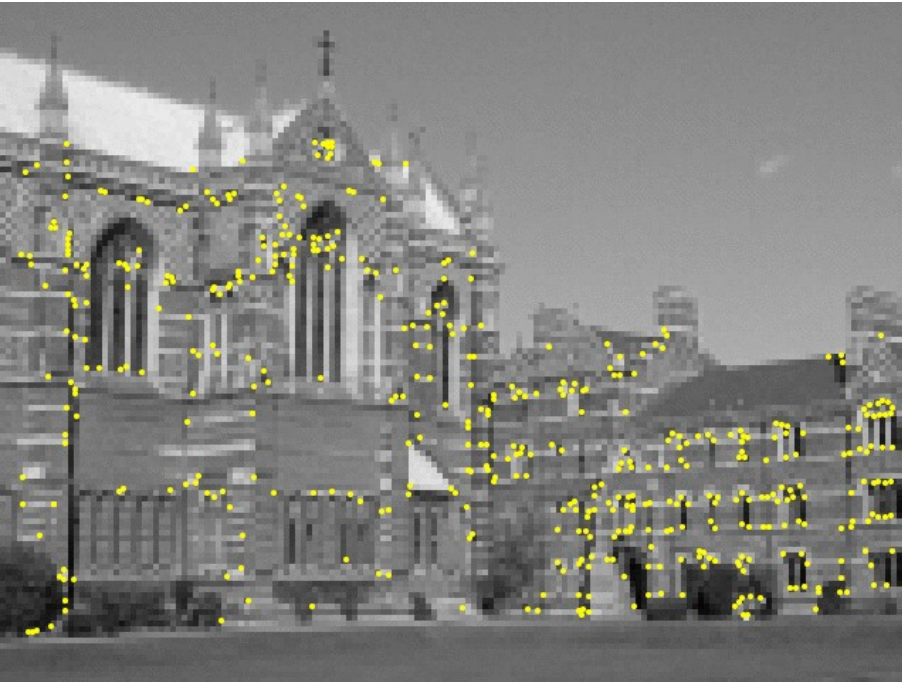


Image gauche

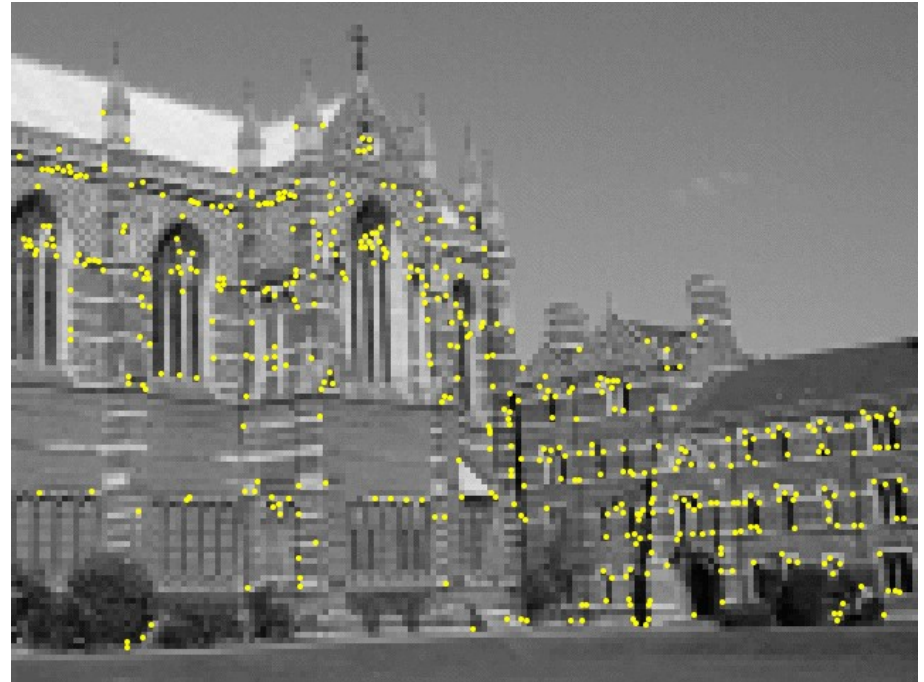


Image droite

Extraction des coins (détecteur de Harris) sur chaque image

- *environ 500 coins sur chaque image*

Problème : correspondances

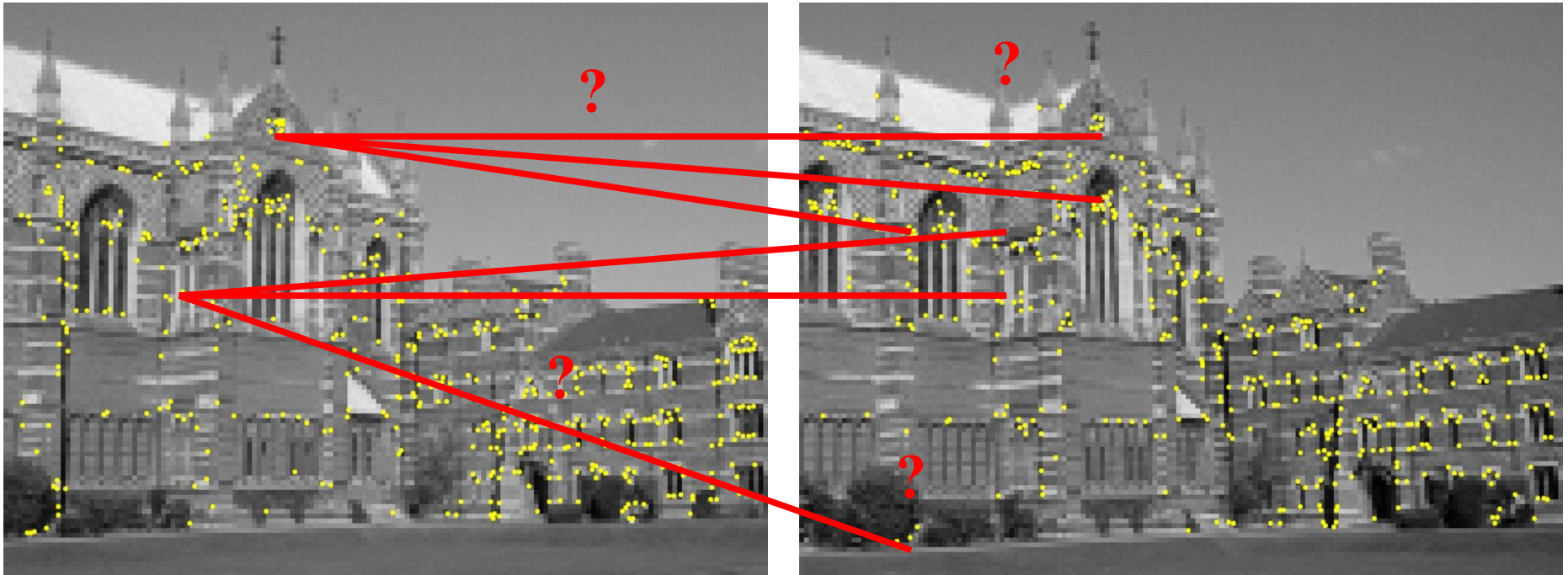


Image gauche

Image droite

Problème : trouver pour chaque point de l'image de gauche à quel point de l'image de droite il correspond

Correspondance trouvée

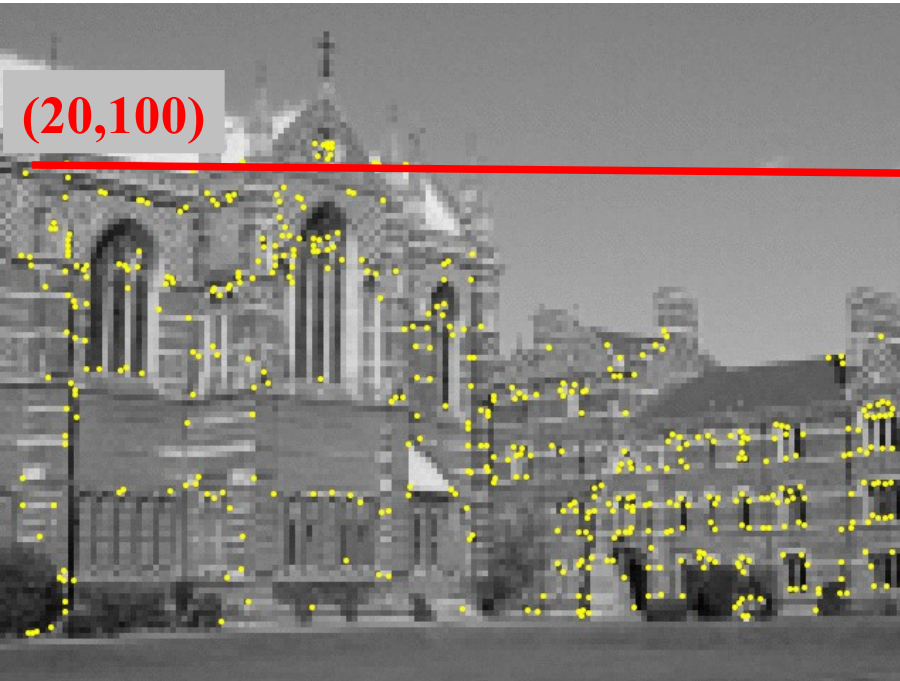


Image gauche

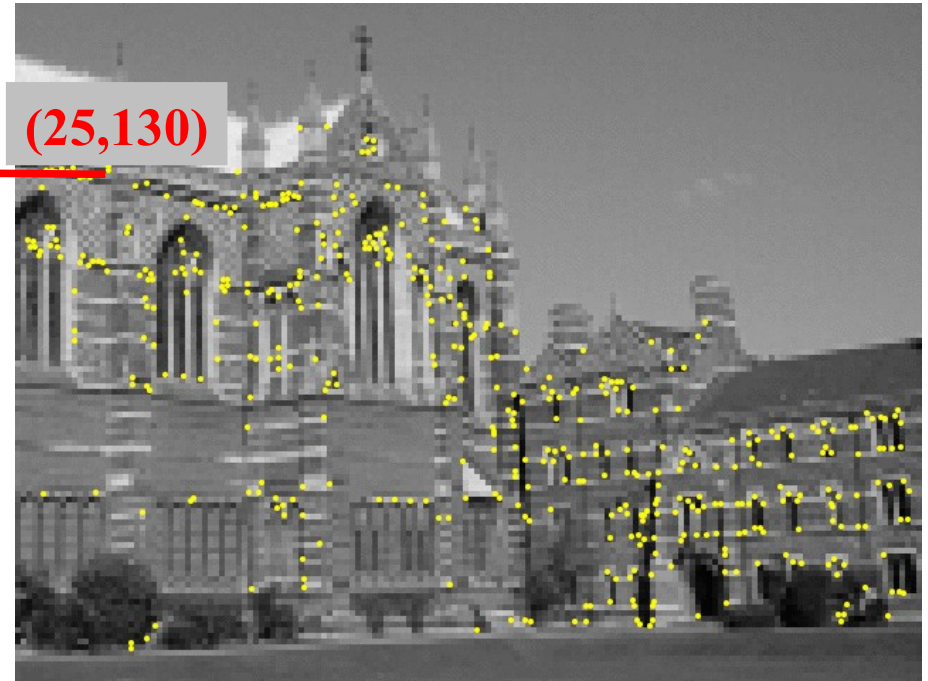
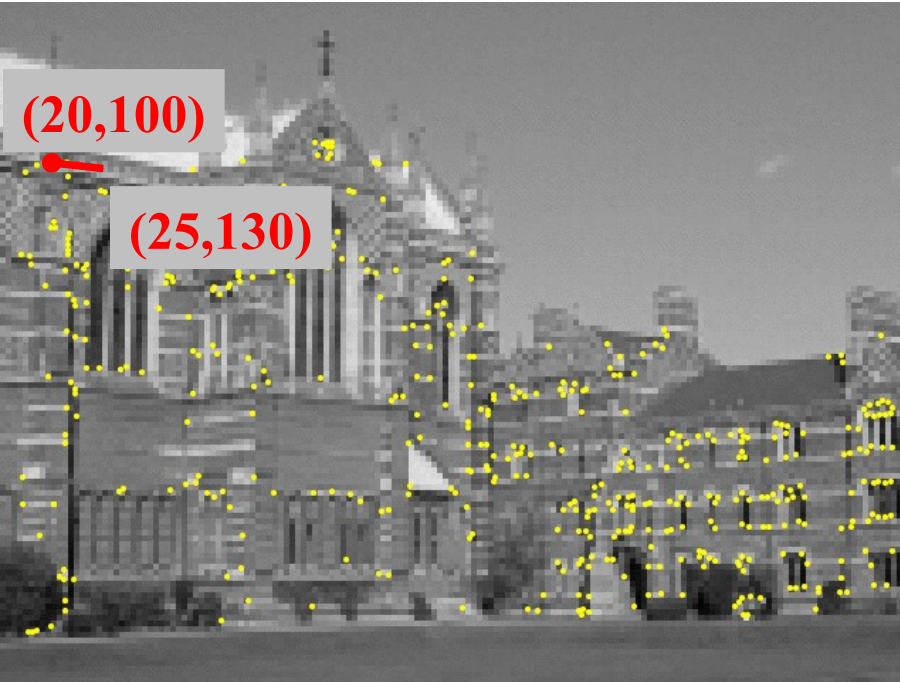


Image droite

Différentes fonctions existent pour trouver les correspondances entre les points de l'image gauche et ceux de l'image droite

Représentation des correspondances

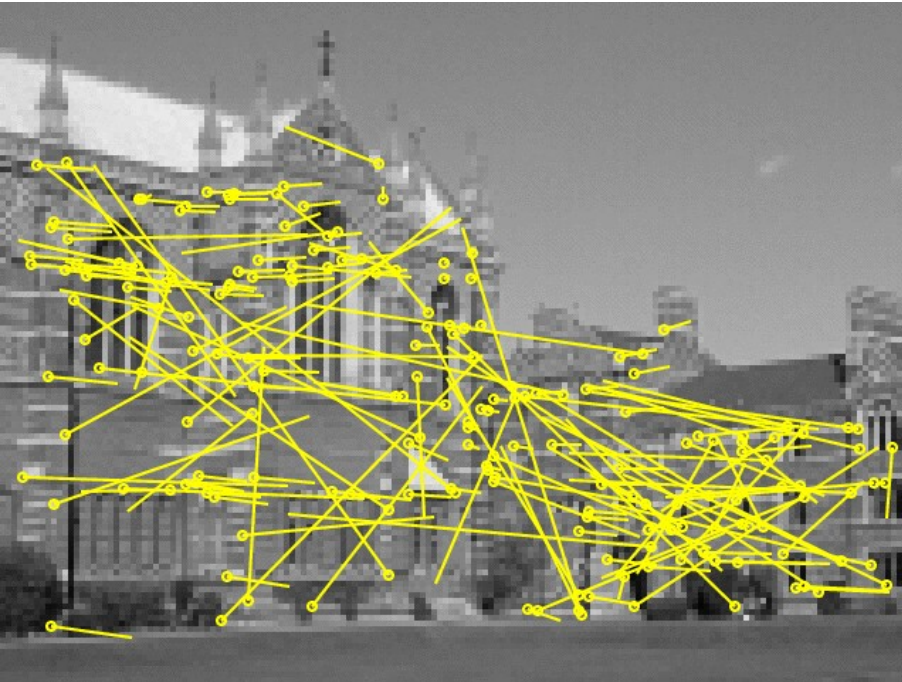


Pour représenter une correspondance, on trace une ligne sur une image montrant le déplacement du point

Exemple : ligne allant de $(20,100)$ à $(25,130)$ sur l'image gauche

Image gauche

Représentation des correspondances



Toutes les correspondances entre les deux images superposées sur l'image gauche (188 au total)

Le point à l'extrémité d'une ligne montre le point d'origine dans l'image gauche. L'autre extrémité d'une montre le point dans l'image droite.

Image gauche

Correspondances les plus fortes



Image gauche

On conserve seulement les correspondances les plus fortes (157 au total)

Ce résultat montre les relations 3D entre les deux images initiales

Ces correspondances sont alignées sur les droites épipolaires

Les lignes les plus courtes montrent les points proches de la caméra et les lignes les plus longues les points loins de la caméra



Correspondances et matrice fondamentale

- Si on ne connaît pas la matrice fondamentale F , alors nous recherchons les correspondances sur toute l'image
- Une fois les correspondances trouvées, nous avons suffisamment de points pour calibrer les deux caméras
 - Calibrage faible
 - Calcul de F sans mire de calibration



Références

(voir aussi la page web du cours)

- Marc Pollefeys, Multiple View Geometry (comp290-89 Spring 2003), University of North Carolina (USA).
 - Class 15 Rectification and structure reconstruction
<http://www.cs.unc.edu/~marc/mvg/course15.ppt>