



UNIVERSITÉ
LAVAL

GLO-4001/7021 INTRODUCTION À LA ROBOTIQUE MOBILE

Capteurs Visuels

(Automne 2019)

Philippe Giguère

Pourquoi on s'intéresse à la vision en robotique mobile?

- Mot-clef toujours élevé dans les conférences robotiques
- Humains s'en servent (preuve que ça fonctionne)
- Capteur passif
- Retourne beaucoup d'information
 - millions pixels
 - position objet
 - couleur
 - texture
- Grande portée
- Peu coûteux
- Environnements sont conçus pour employer la vision
 - panneaux indicateurs, feux de circulation, etc.

Vision : problème difficile

- 40 ans de travail, résultats mitigés
- Utilise 50% des ressources du cerveau humain
 - indique grande complexité du problème
- Toujours pas de solution générale
 - image 2D → monde 3D n'existe pas (tout à fait) encore
 - nécessite une connaissance préalable de la taille des objets

Vision : utilité en robotique mobile

- S'en sert pour :
 - localisation avec repères visuel
 - classification des obstacles
 - fermeture de boucle, etc...
- Interpréter des parties de l'image, pas toute la scène
 - *active vision* (mécanique, traitement image, fovéa)

œil agile

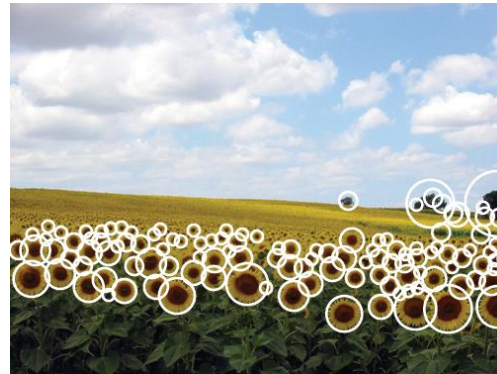
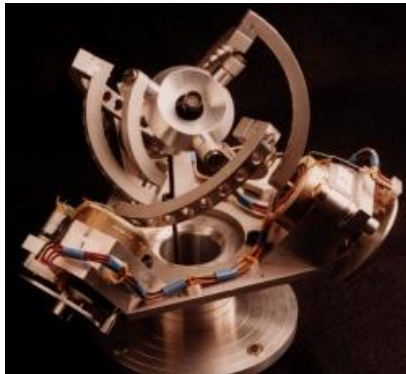
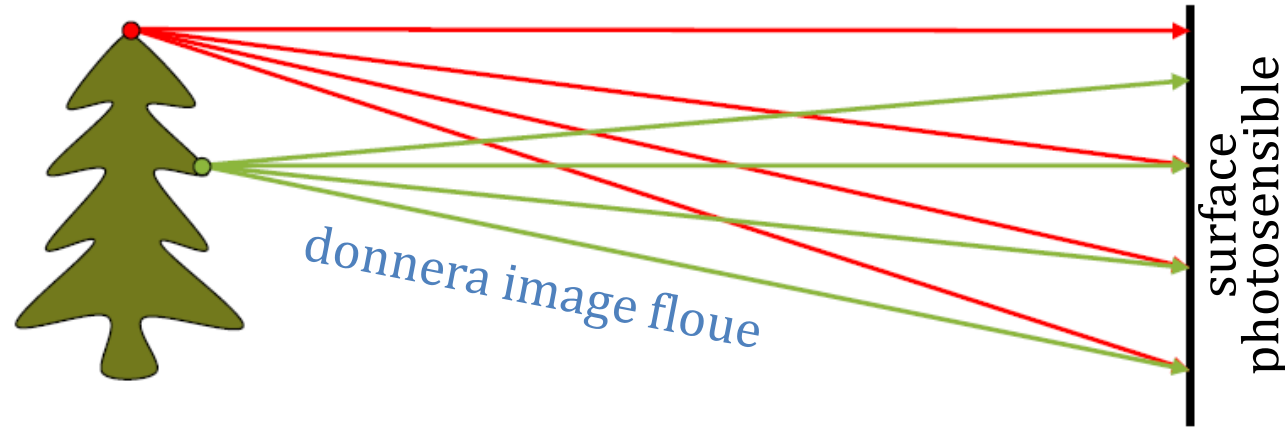


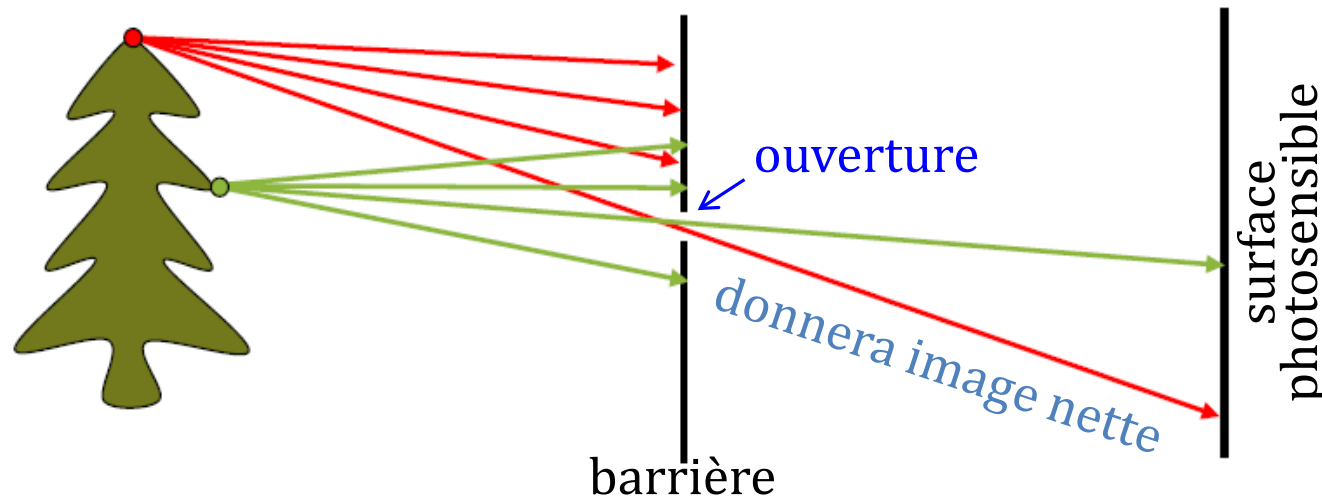
Image : processus de capture

- Pas juste avoir une surface photosensible :



plusieurs rayons
incidents par
« pixel »

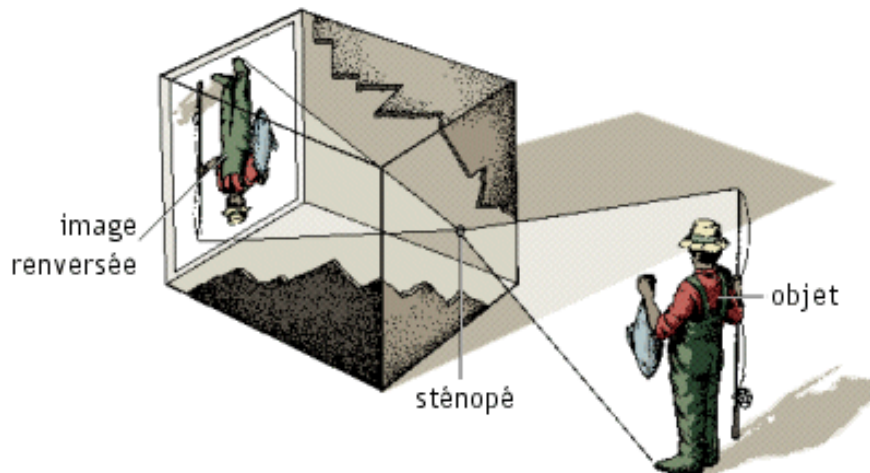
- Il faut bloquer la majorité des rayons :



un seul rayon
incident par
« pixel »

Modèle caméra à sténopé (*pinhole*)

- Approximer notre caméra + lentille comme une boîte avec un seul petit trou



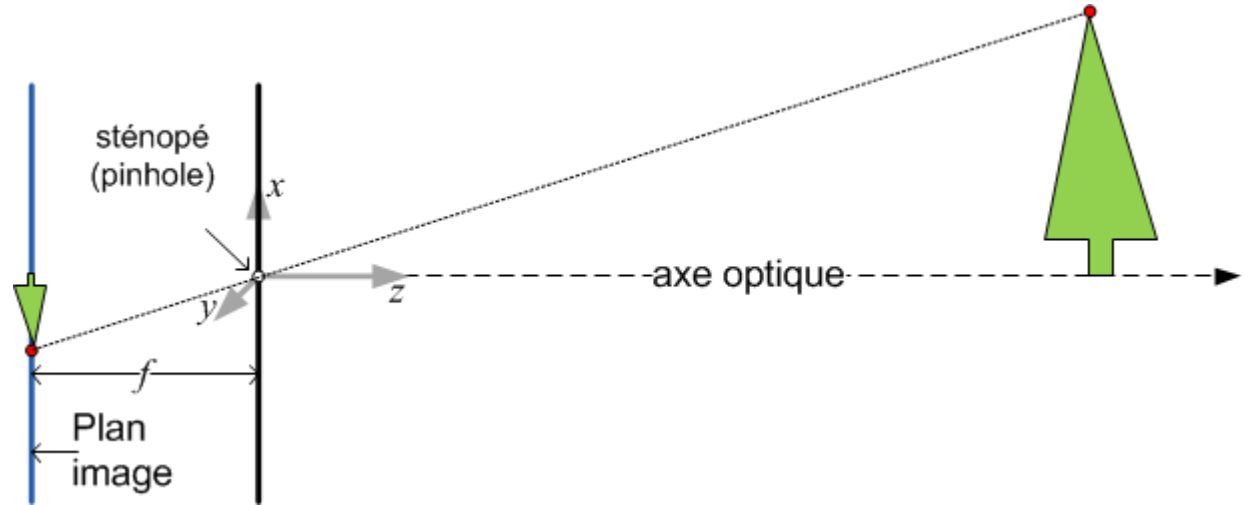
Principe de la chambre noire
Source : Encyclopédie Microsoft Encarta.



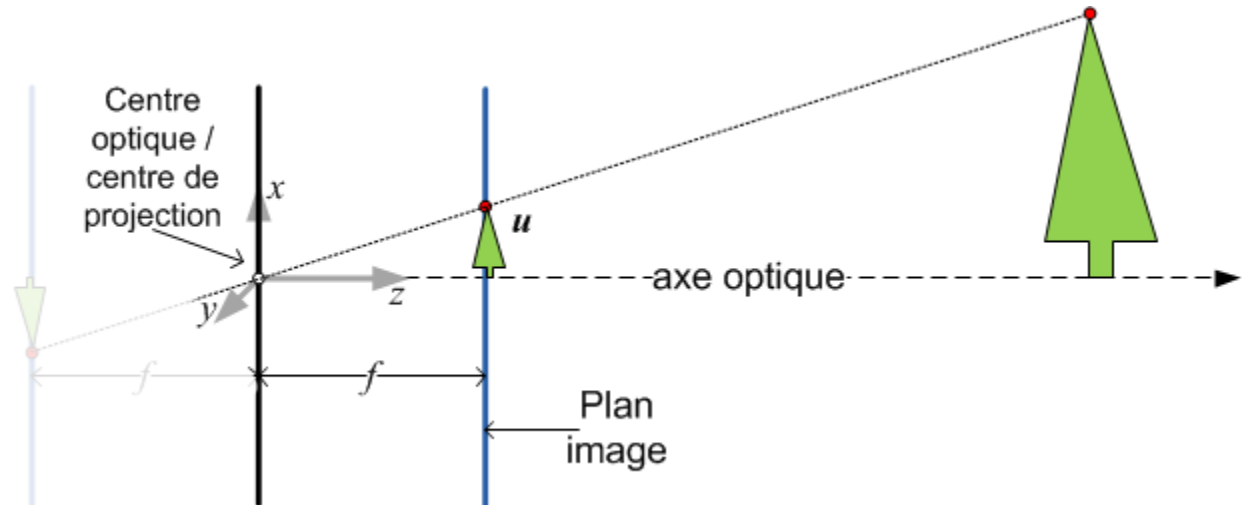
papier photo + caméra sténopé

Modèle de caméra sténopé

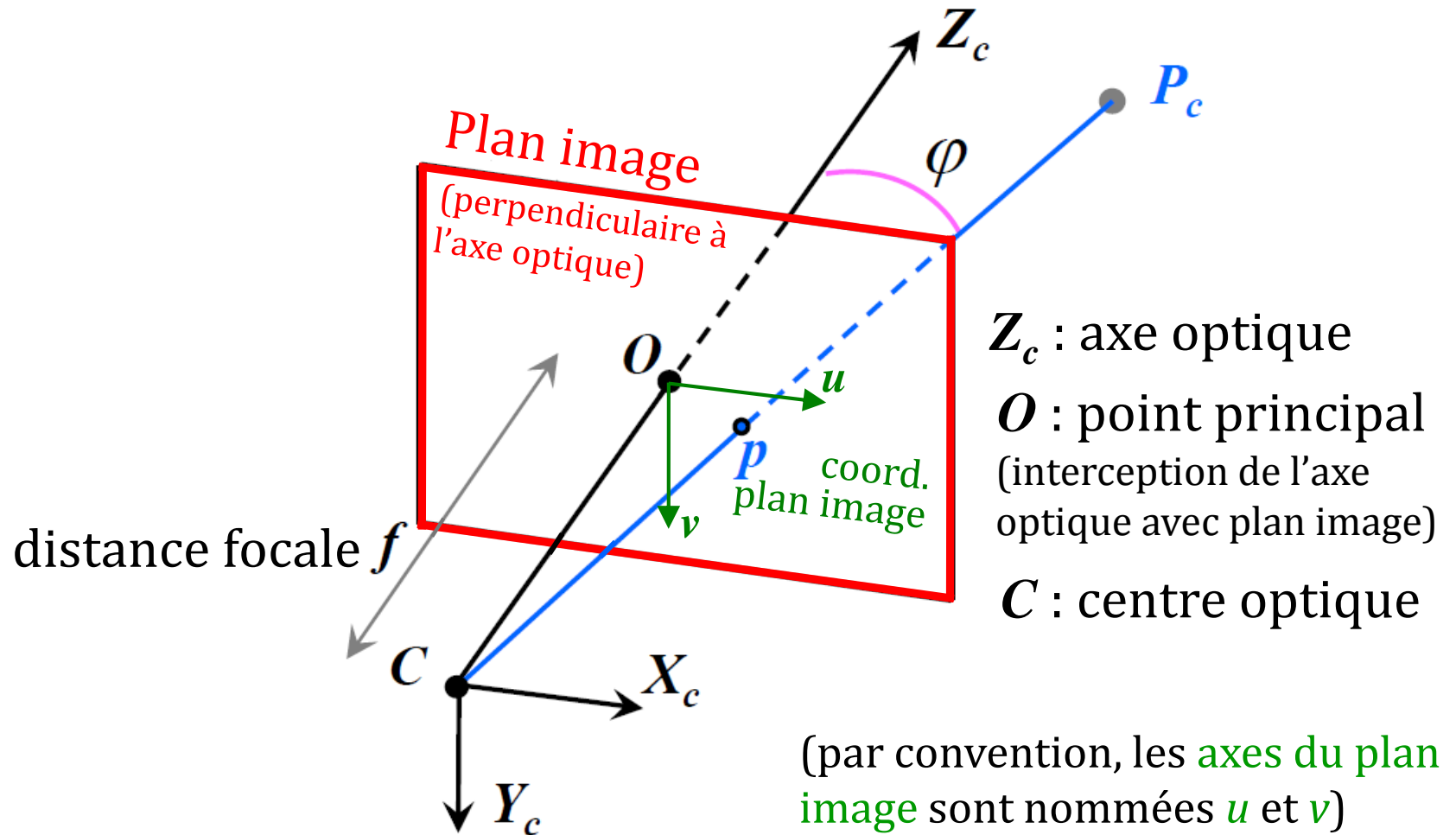
Caméra physique :
image formée à l'envers



Déplacer le plan image
en avant du trou :
image à l'endroit (mais
non-réalisable physiquement)

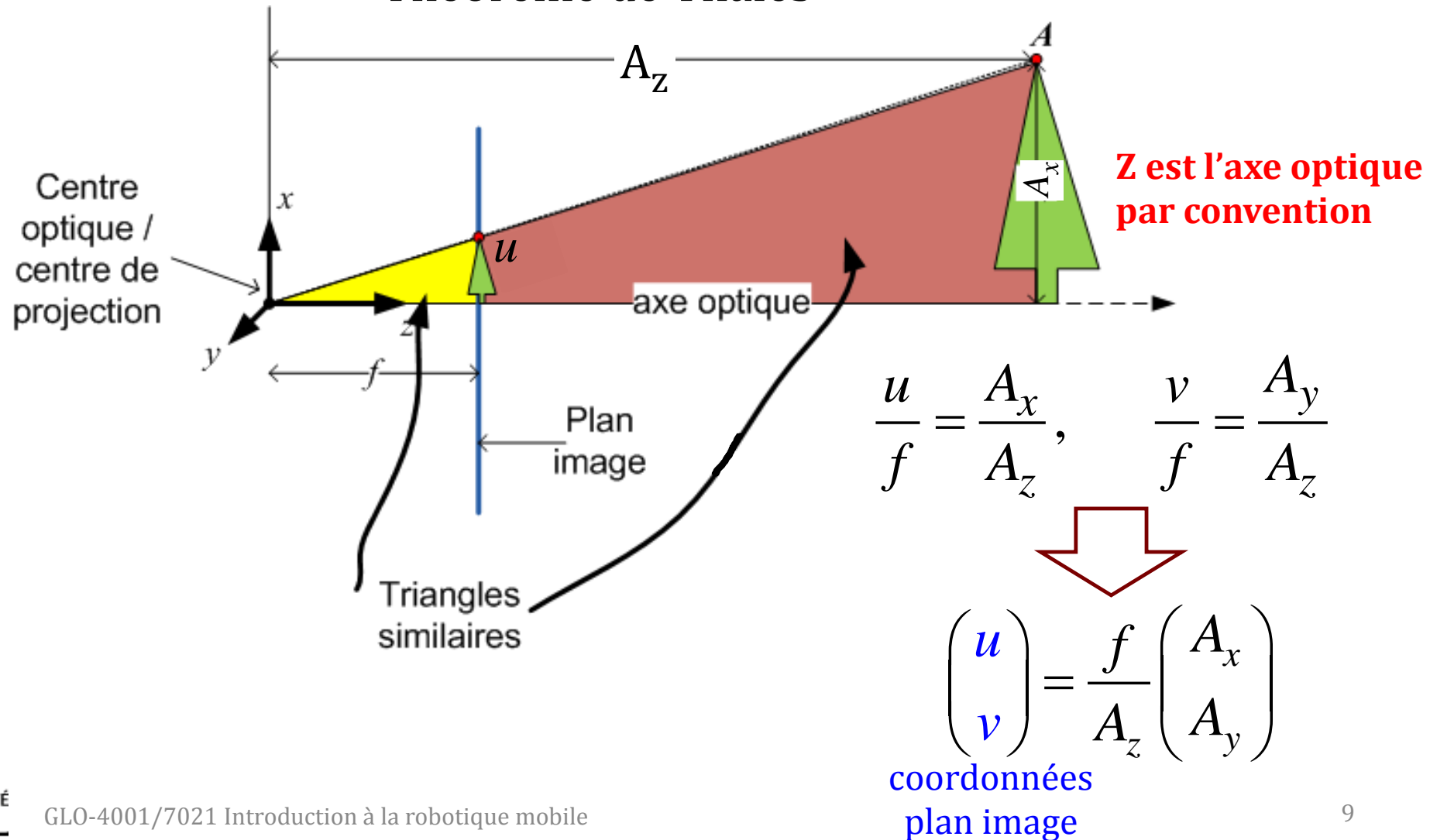


Nomenclature caméra perspective



Caméra perspective : génération d'une image

Théorème de Thalès



Exemple de projection

- Vous avez un point situé aux coordonnées*
($A_x=3$, $A_y=0$, $A_z=20$), en mètres
- La distance focale f de la caméra est de 50 *mm*
- À quelles coordonnées (u,v)** du plan image, en *mm*, ce point apparaîtra-t-il?

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{A_z} \begin{pmatrix} A_x \\ A_y \end{pmatrix} = \frac{50 \text{ mm}}{20 \text{ m}} \begin{pmatrix} 3 \text{ m} \\ 0 \text{ m} \end{pmatrix} = \begin{pmatrix} 7.5 \\ 0 \end{pmatrix} \text{ mm}$$

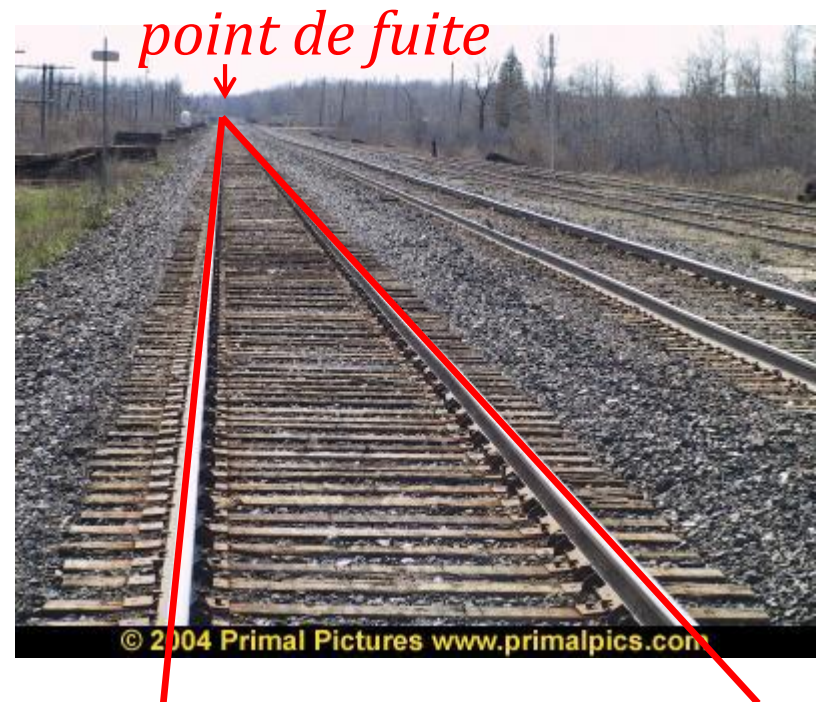
- (Notez comment les unités en *m* s'annulent, et ne dépendent que des unités *mm* de la focale)

*coordonnées de la caméra, avec centre optique C à $(0,0,0)$ et axe optique = Z_c

**en supposant que l'origine de (u,v) est au point principal O

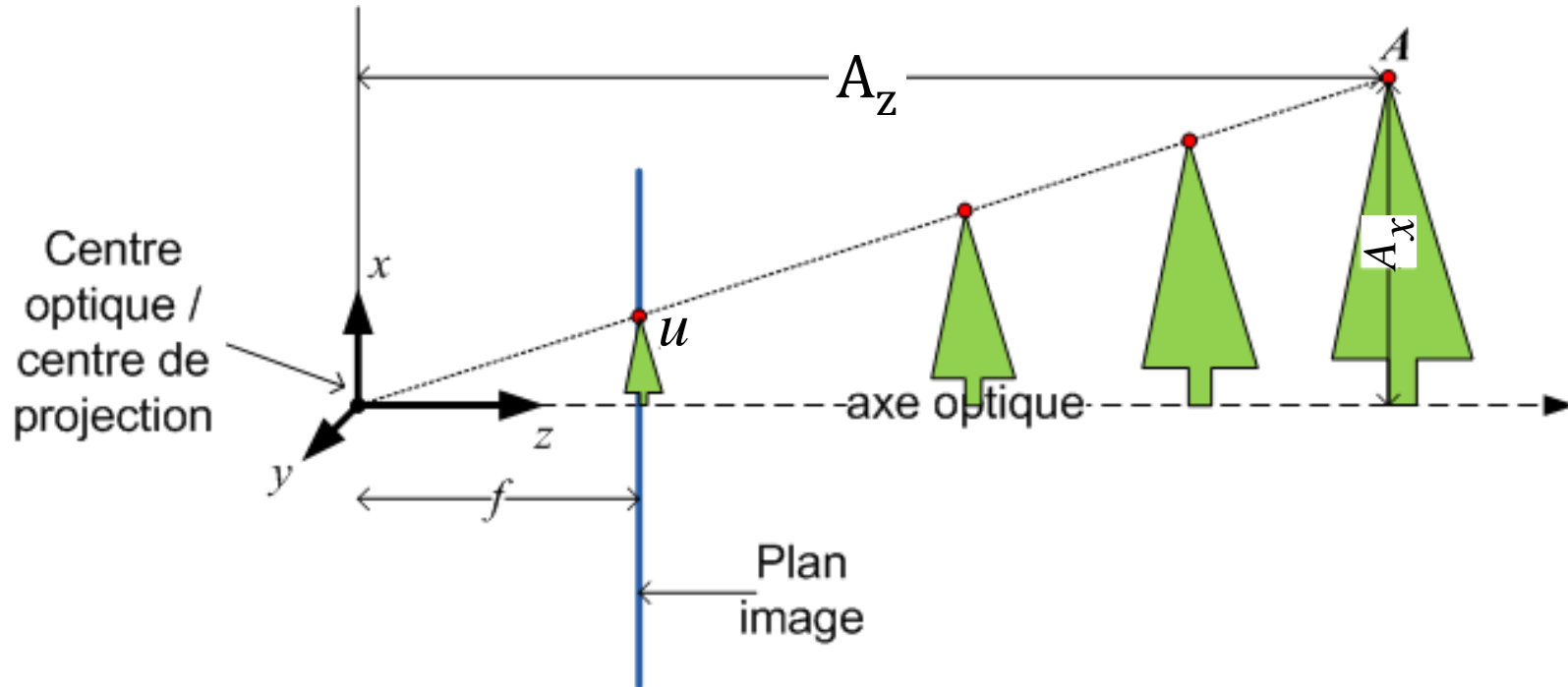
Caméra perspective : lignes

- Une ligne droite dans le monde apparaît comme une ligne droite dans le plan image.
- Des lignes parallèles vont se joindre au point de fuite (*vanishing point*)



Caméra 2D : perte de 3D

- Problème mal posé (plusieurs solutions pour l'inverse)



Jouer des tours...



Capteur numérique : plan image en pixel

Pour une image numérique, l'origine (u,v) est souvent le coin supérieur gauche.

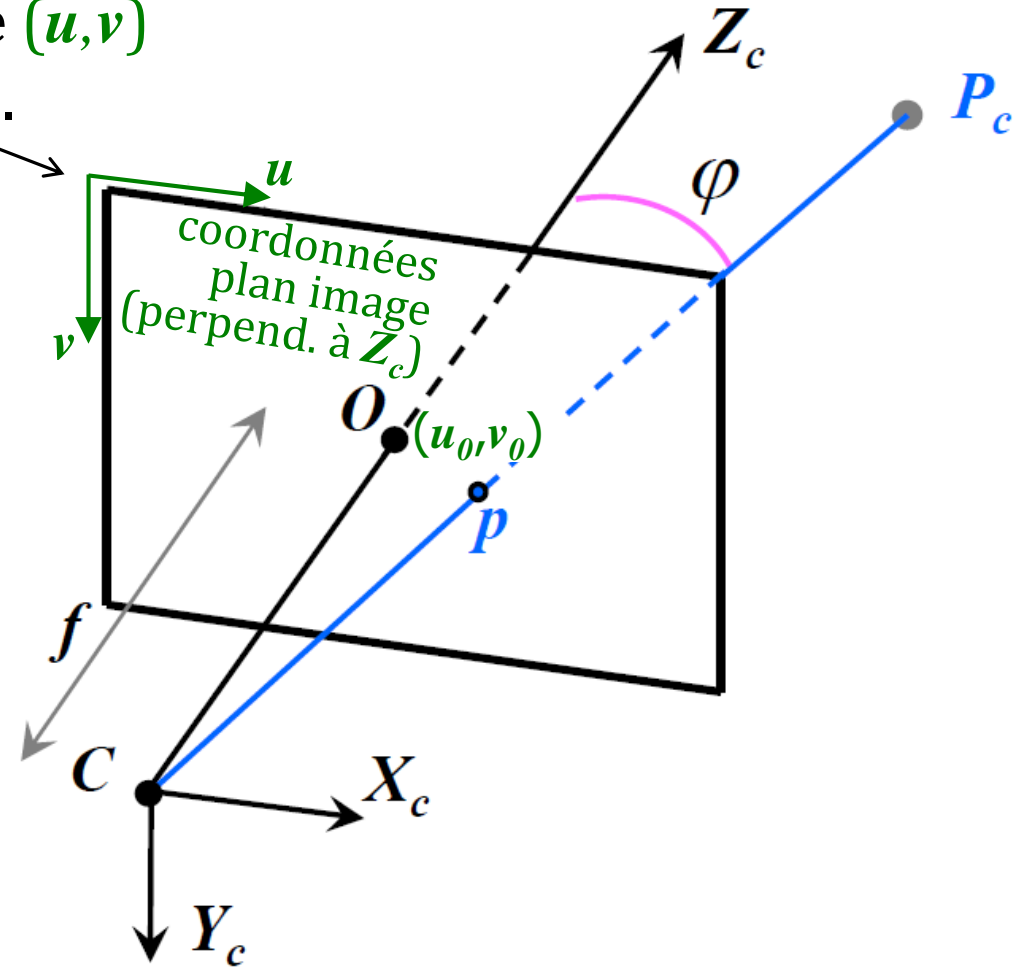
Le point principal O sera situé à la coordonnée (u_0, v_0) .

En tenir compte dans les équations :

$$u = u_0 + f \frac{A_X}{A_Z} \quad v = v_0 + f \frac{A_Y}{A_Z}$$

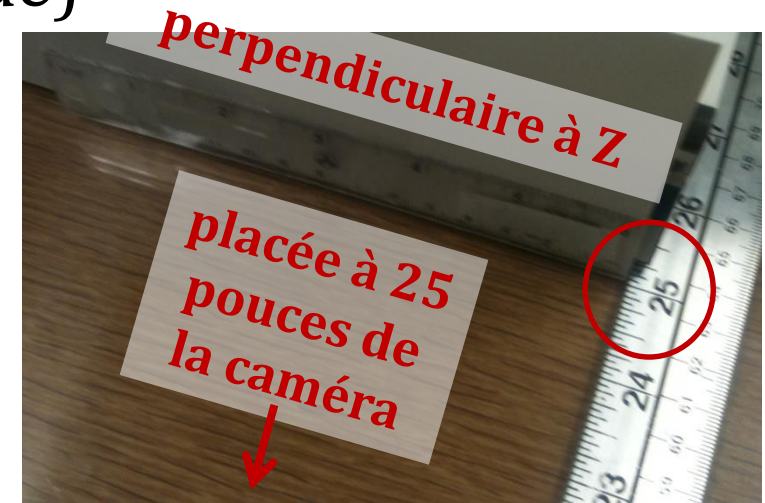
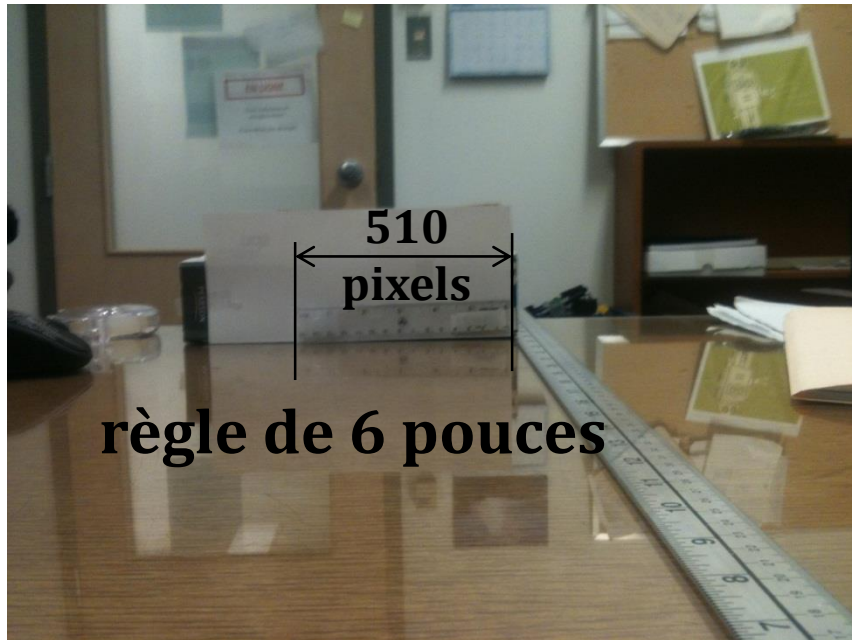
(avec f défini en pixel)

(Note matlab : les fonctions comme `imagesc` ou `imshow` vont avoir les axes intervertis, et l'axe vertical est inversé)



Valeur de la focale f

- Valeur *précise* change d'une caméra à l'autre
- Sera constante (sauf si zoom optique)
- On l'identifie via **calibration**
- Calibration *rudimentaire** :



$$\frac{510 \text{ pixels}}{f} = \frac{6 \text{ pouces}}{25 \text{ pouces}}$$

$$f = \frac{25}{6} 510 \text{ pixels} = 2125 \text{ pixels}$$

Note: on assume ici des pixels carrés sur la cellule

Coordonnées homogènes & Généralisation du modèle de projection

Math peu élégante

- Caméra est une fonction non-linéaire en A_z :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{A_z} \begin{pmatrix} A_x \\ A_y \end{pmatrix}$$

- Chercher quelque chose plus élégant :

$$A' = PA$$

A' : coordonnée plan image

A : coordonnées point 3D

P : modèle de caméra

Coordonnées homogènes

- Représentation point 2D avec 3 composantes:

$$\mathbf{x} = (x_1, x_2, x_3) \leftrightarrow (x_1/x_3, x_2/x_3)$$

$$(3, 2) \rightarrow (3, 2, 1) \text{ ou } (6, 4, 2) \text{ ou } \dots$$

- Représentation point 3D avec 4 composantes:

$$\mathbf{x} = (x_1, x_2, x_3, x_4) \leftrightarrow (x_1/x_4, x_2/x_4, x_3/x_4)$$

$$(2, 5, 7) \rightarrow (2, 5, 7, 1) \text{ ou } (6, 15, 21, 3) \text{ ou } \dots$$

(utile pour rotations et translations, à venir)

homogène \leftrightarrow **cartésien**

Caméra perspective en coord. homogène

- \sim : signifie **identique à un facteur d'échelle**

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

position dans
le plan image

coordonnées
d'un point dans
l'espace 3D

$$\begin{array}{ccc} A' & \sim & P \\ & & \text{(matrice intrinsèque)} \\ \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} & \sim & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \end{array}$$

coordonnées
homogènes

Dérivation de la matrice P perspective

$$\begin{array}{ccccccc} & \text{Thalès} & & \text{factorise} & & \text{échelle} & \\ \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} & = \begin{pmatrix} fA_x / A_z \\ fA_y / A_z \\ 1 \end{pmatrix} & = & \frac{f}{A_z} \begin{pmatrix} A_x \\ A_y \\ A_z / f \end{pmatrix} & \sim & \begin{pmatrix} A_x \\ A_y \\ A_z / f \end{pmatrix} & = & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \\ 1 \end{pmatrix} \end{array}$$

Caméra perspective : exemple

- Exemple

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \\ 1 \end{bmatrix}$$

coordonnées
homogènes

Point dans l'espace à $[4 \ 6 \ 2]^T \rightarrow [4 \ 6 \ 2 \ 1]^T$

focale $f = 0.1$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 10 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 20 \end{bmatrix}$$

Passage d'**homogène** en **cartésien** dans plan image $[4 \ 6 \ 20]^T \rightarrow [0.2 \ 0.3]^T$

Autres types de projection de caméra

- Perspective faible:

- distance moyenne \bar{z}

satellite  altitude 20,000 km

$$A' \sim P A$$

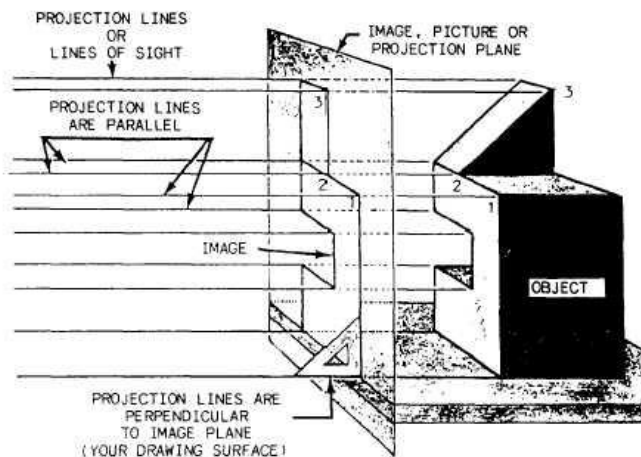
$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \bar{z}/f \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \\ 1 \end{bmatrix}$$

distance A_z n'intervient plus



- Orthographique :

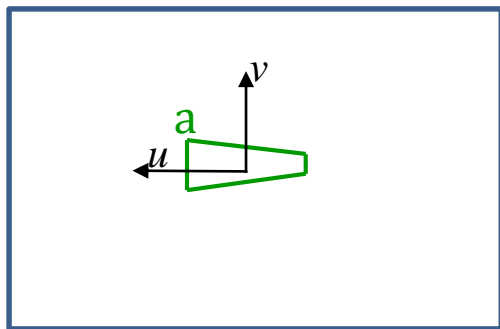
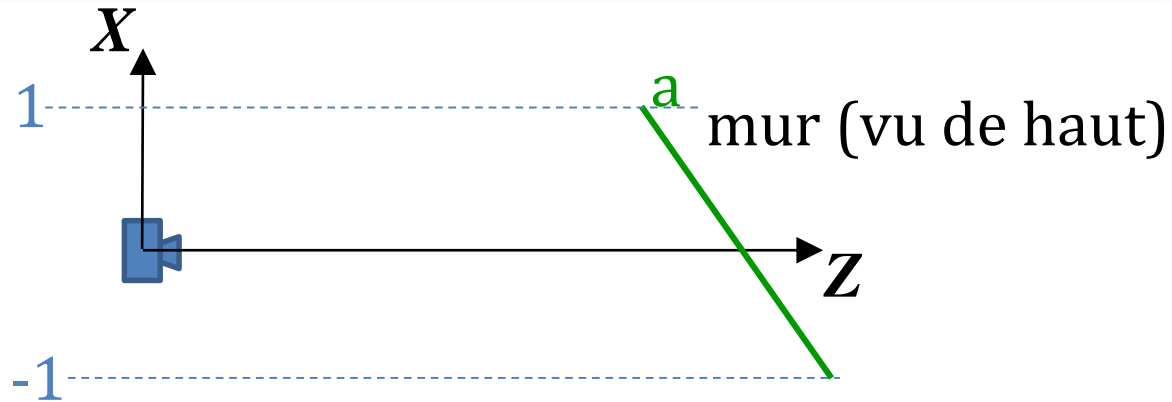
- rayons parallèles



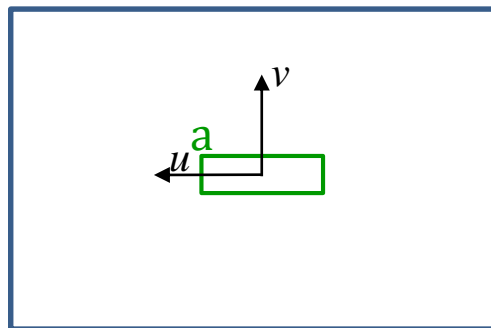
$$A' \sim P A$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \\ 1 \end{bmatrix}$$

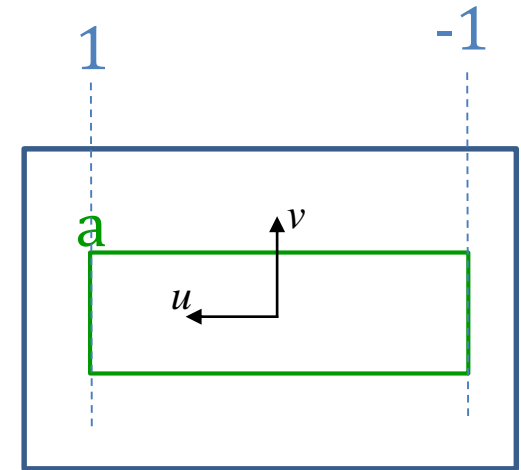
Exemples de projection



Perspective



Perspective faible



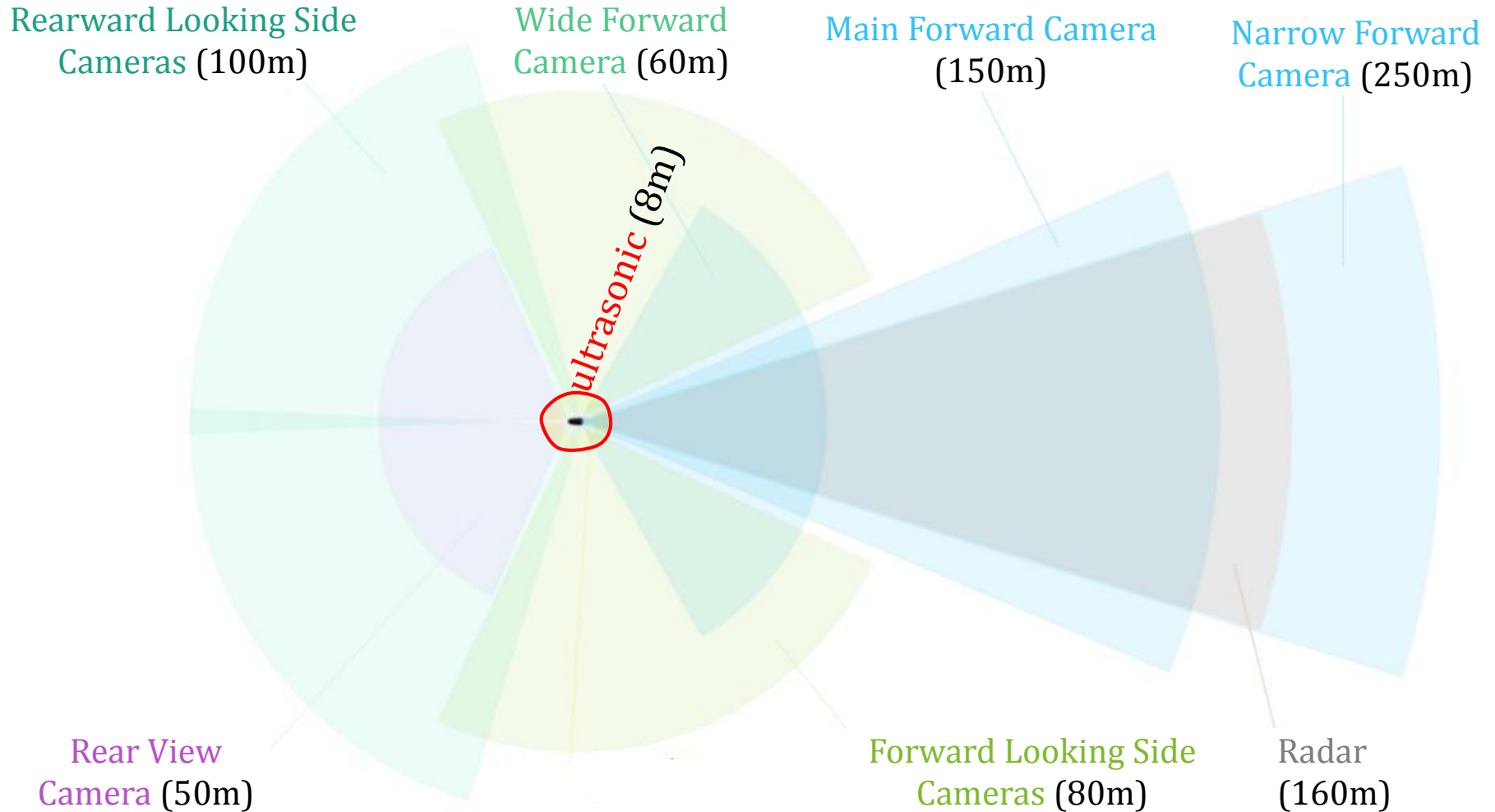
Orthographique

Capteurs sur Stanley (2005)



- 5 LiDAR
 - 1 caméra
 - 1 radar
 - 1 GPS
 - 1 centrale inertielle
 - odométrie des roues
-
- Portée de détection de surface navigable :
 - 22 m par LiDAR (vitesse max 25 m/h)
 - 70 m par caméra (vitesse max 35 m/h)
 - Caméra entraînée par apprentissage auto-supervisé

Capteurs sur Tesla (2018)



Localisation en 2D par caméra

Hypothèses simplificatrices

- Le robot est sur un plancher plat
- L'axe optique Z est parallèle au sol
- L'axe X de la caméra est parallèle au sol
- Tous les points de repère sont à la hauteur de la caméra du robot (sur une ligne horizontale passant par le point principal)

Localisation par caméra (en 2D)

Vous avez une carte avec balises l_i

pastilles de couleur

sur un mur

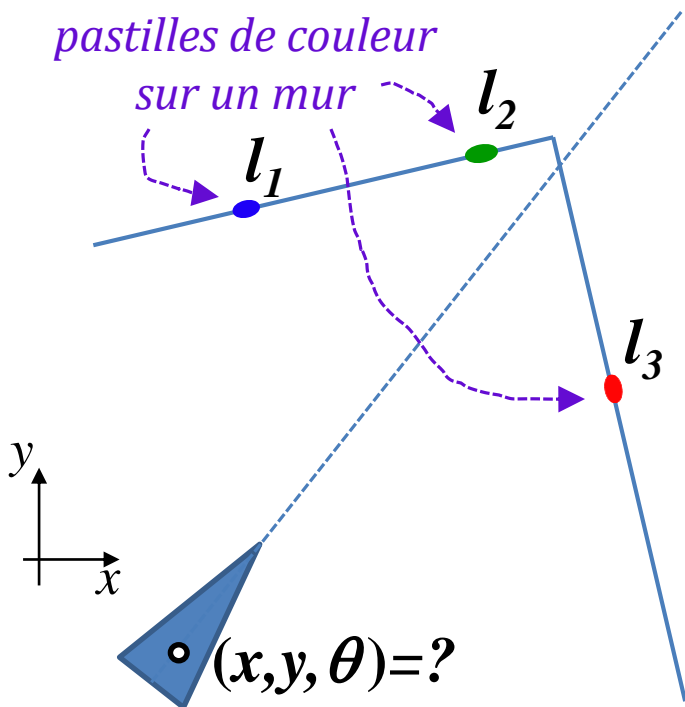
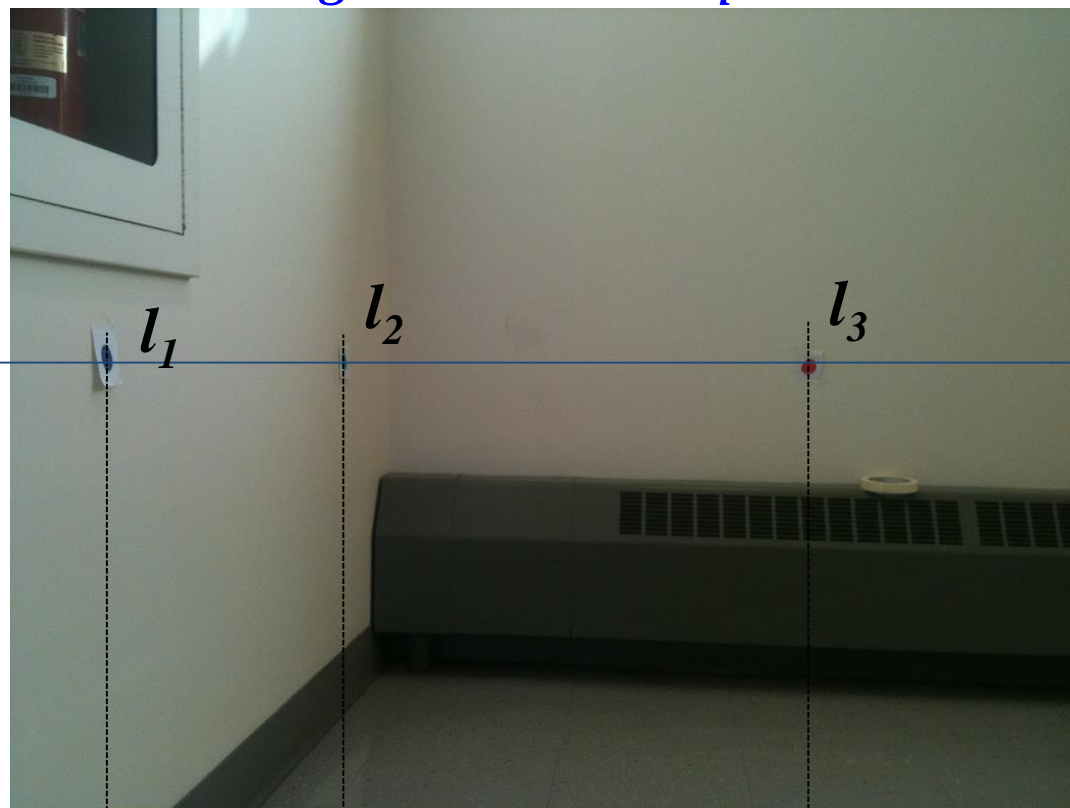


image de 512x384 pixels



On connaît la position de l_1, l_2, l_3

Robot a une pose (x, y, θ) inconnue

Sa caméra prend une photo

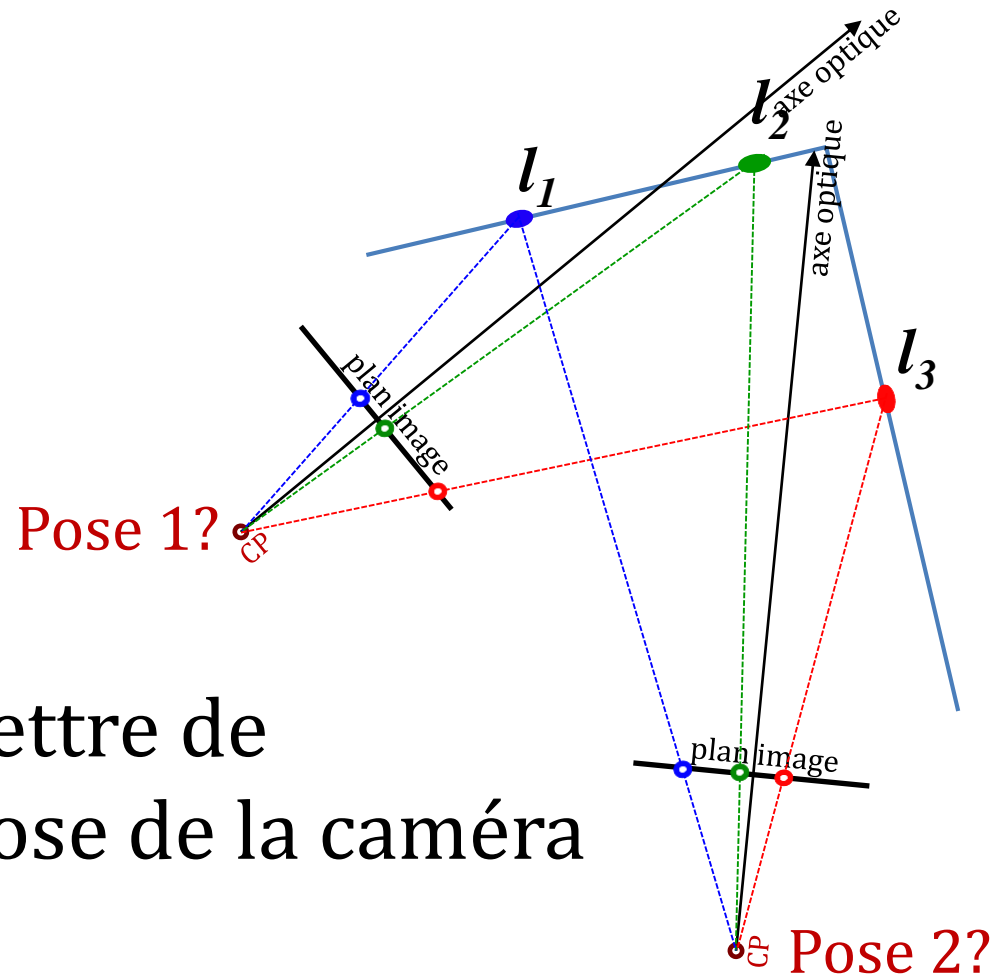
Retrouve les points de repères visuels dans l'image (*data association problem*)

i.e. position horizontale en pixel de l_1, l_2, l_3 pour ce problème en 2D

On veut retrouver x, y et θ à partir des positions en pixels de l_1, l_2, l_3

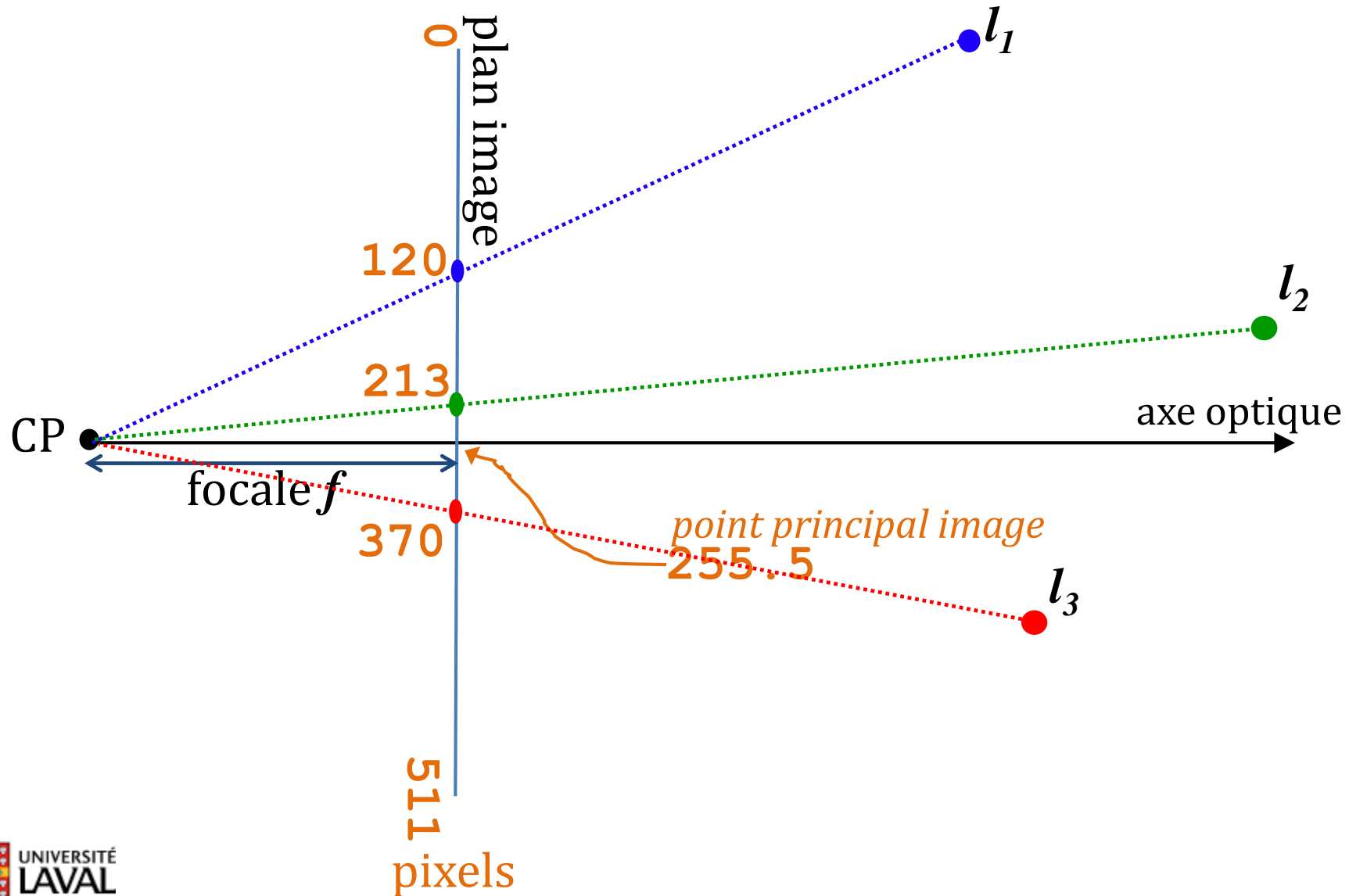
Localisation par caméra (en 2D)

- Se fier sur le fait que chacune des poses génère une image différente

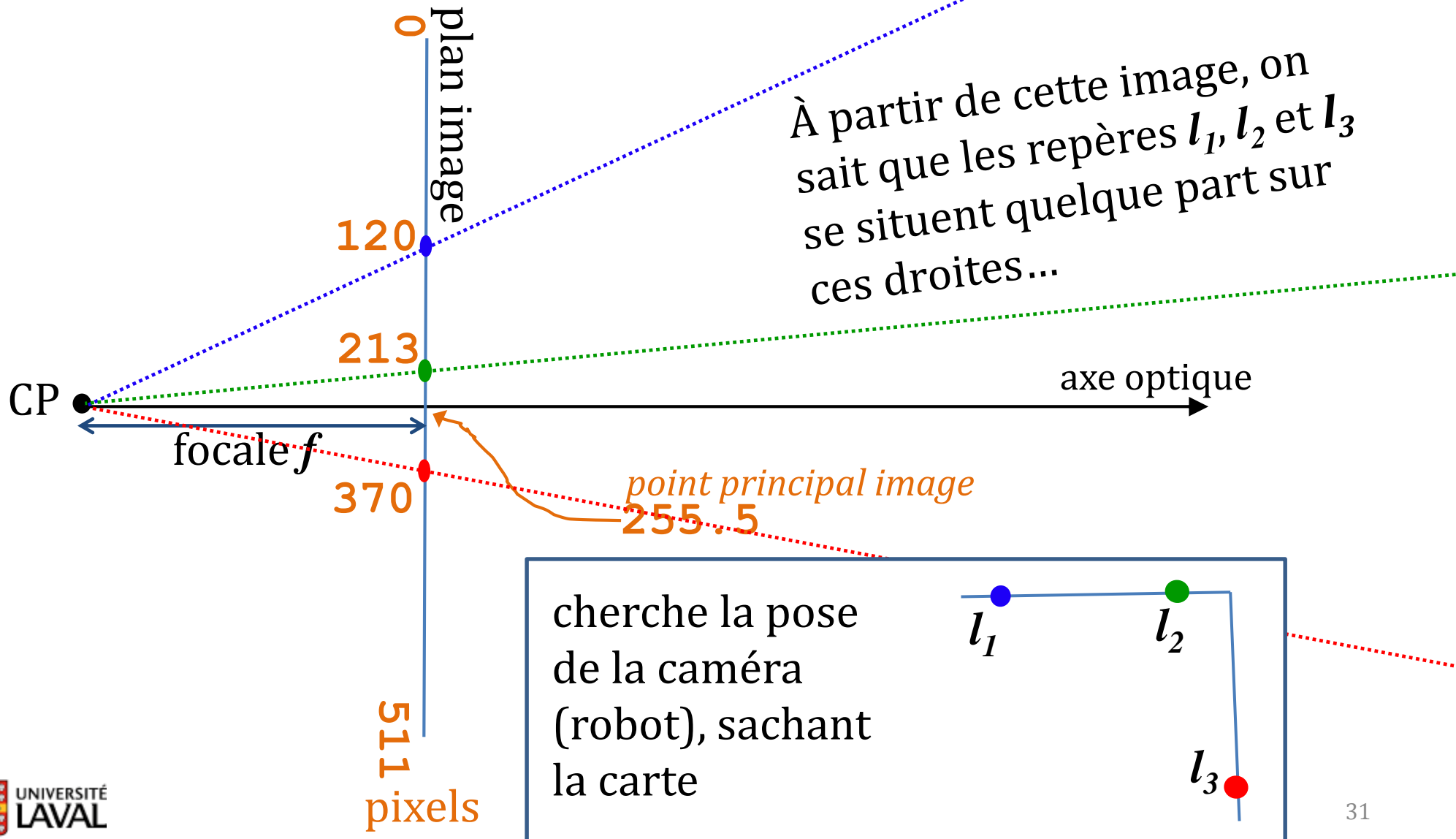


- Va nous permettre de retrouver la pose de la caméra

Caméra à sténopé : prise de la photo

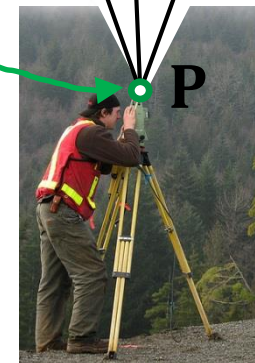
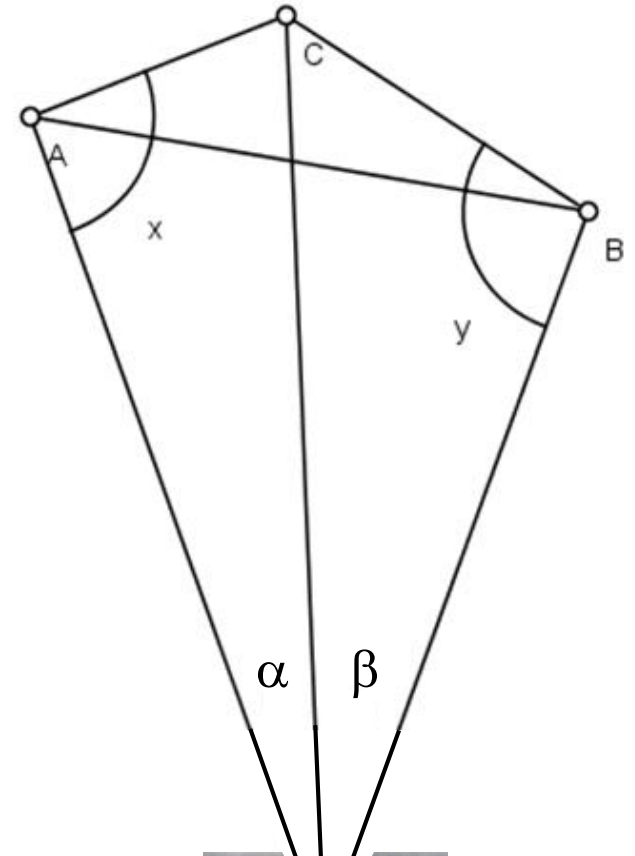


Caméra à sténopé : prise de la photo

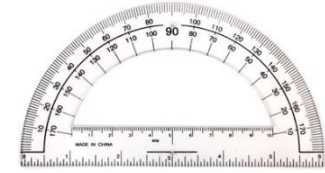


Le problème Snellius-Pothénot

- Étudié en 1615 en arpentage
- Soit 3 points, A B et C, avec coordonnées connues sur une carte
- Soit deux angles mesurés à **P**
 - α et β
- **Trouver position de P**
- Commençons par trouver α et β ...

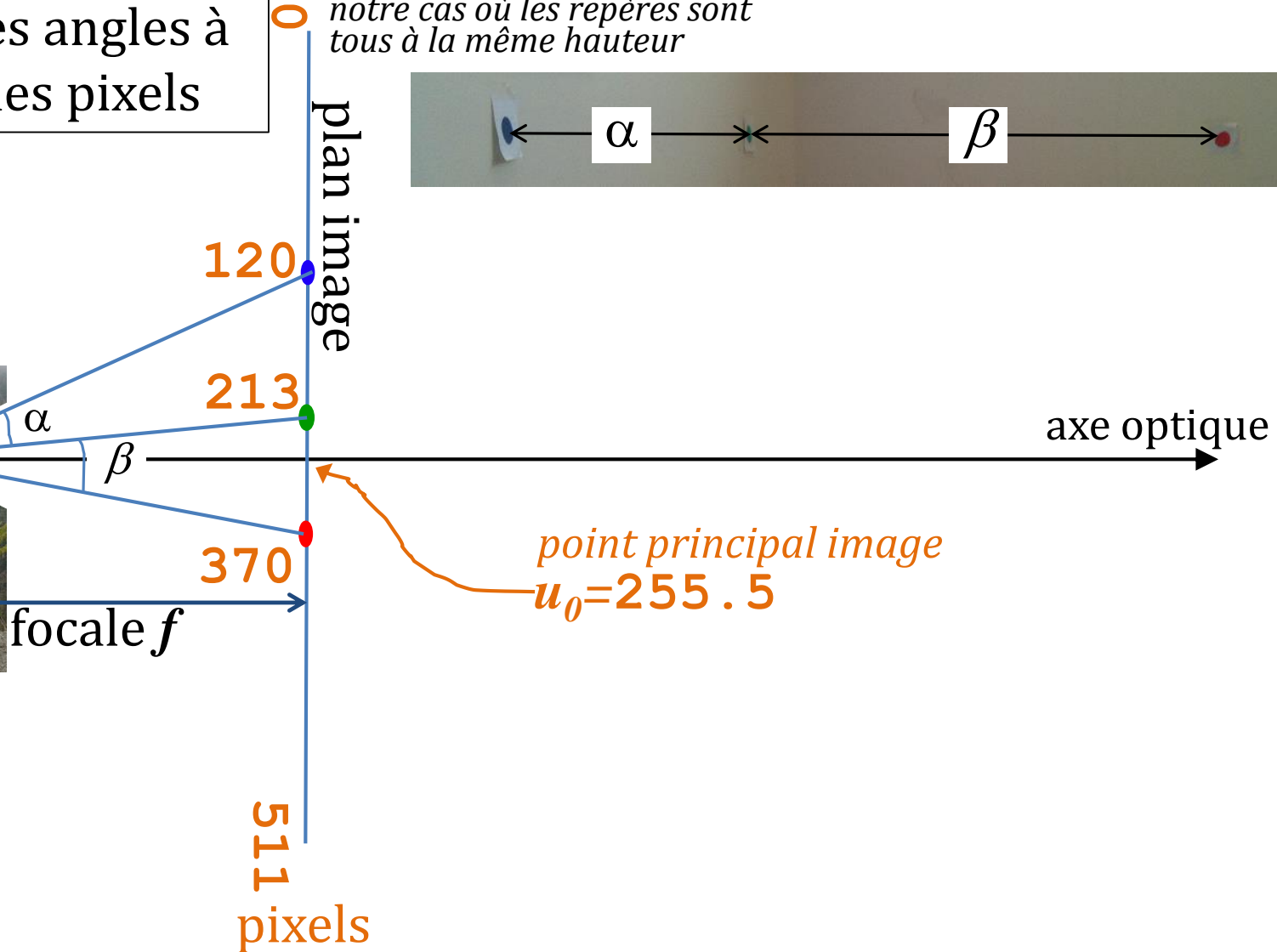
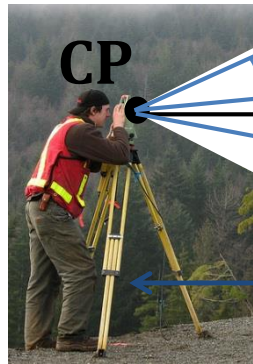


Caméra == rapporteur d'angles

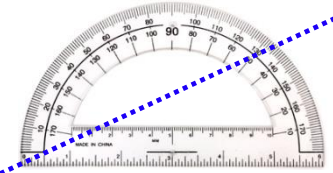


Calcul des angles à partir des pixels

Note : valide seulement pour notre cas où les repères sont tous à la même hauteur



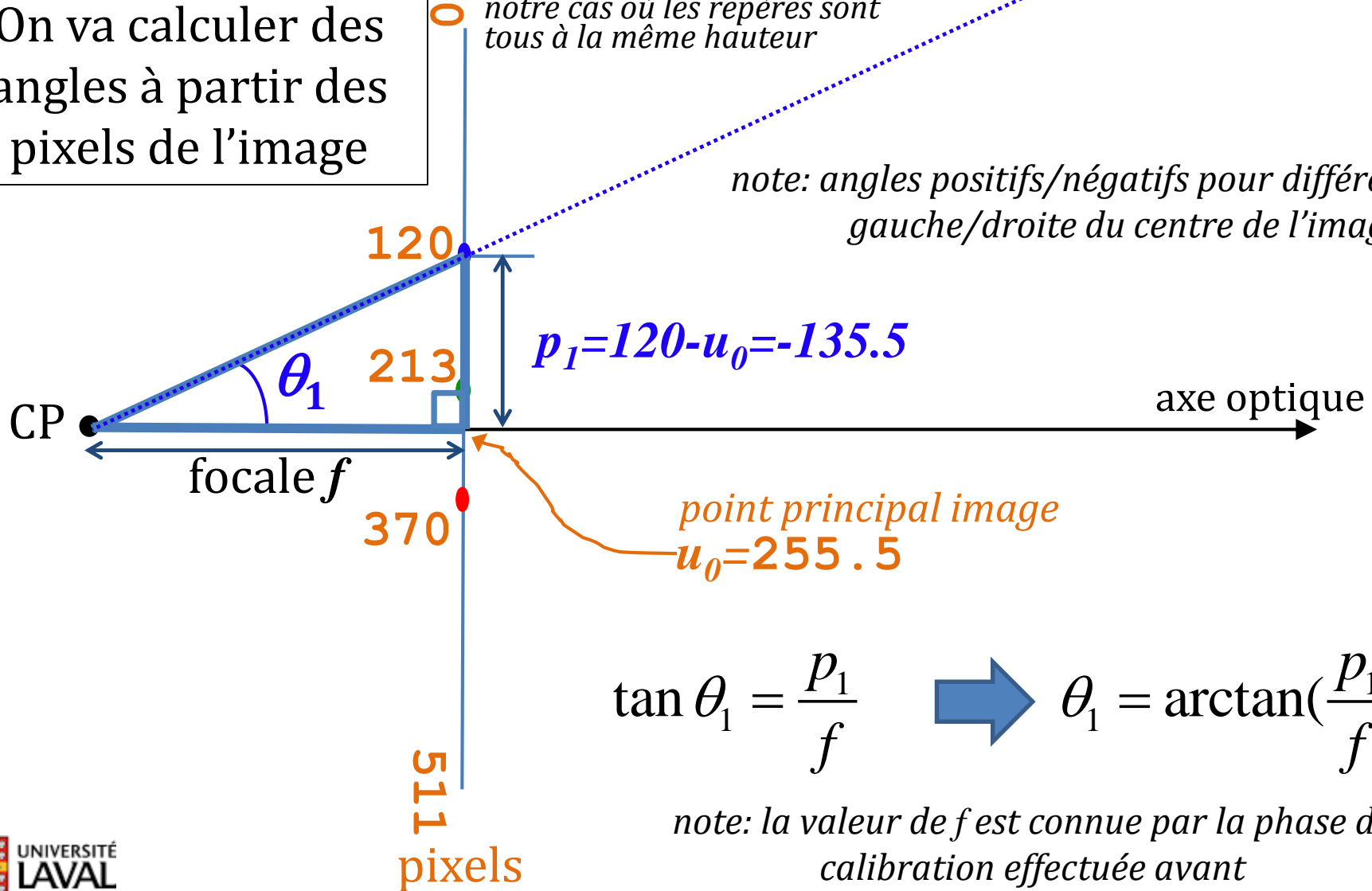
Caméra == rapporteur d'angles



On va calculer des angles à partir des pixels de l'image

Note : valide seulement pour notre cas où les repères sont tous à la même hauteur

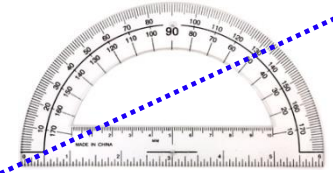
note: angles positifs/négatifs pour différentier si à gauche/droite du centre de l'image



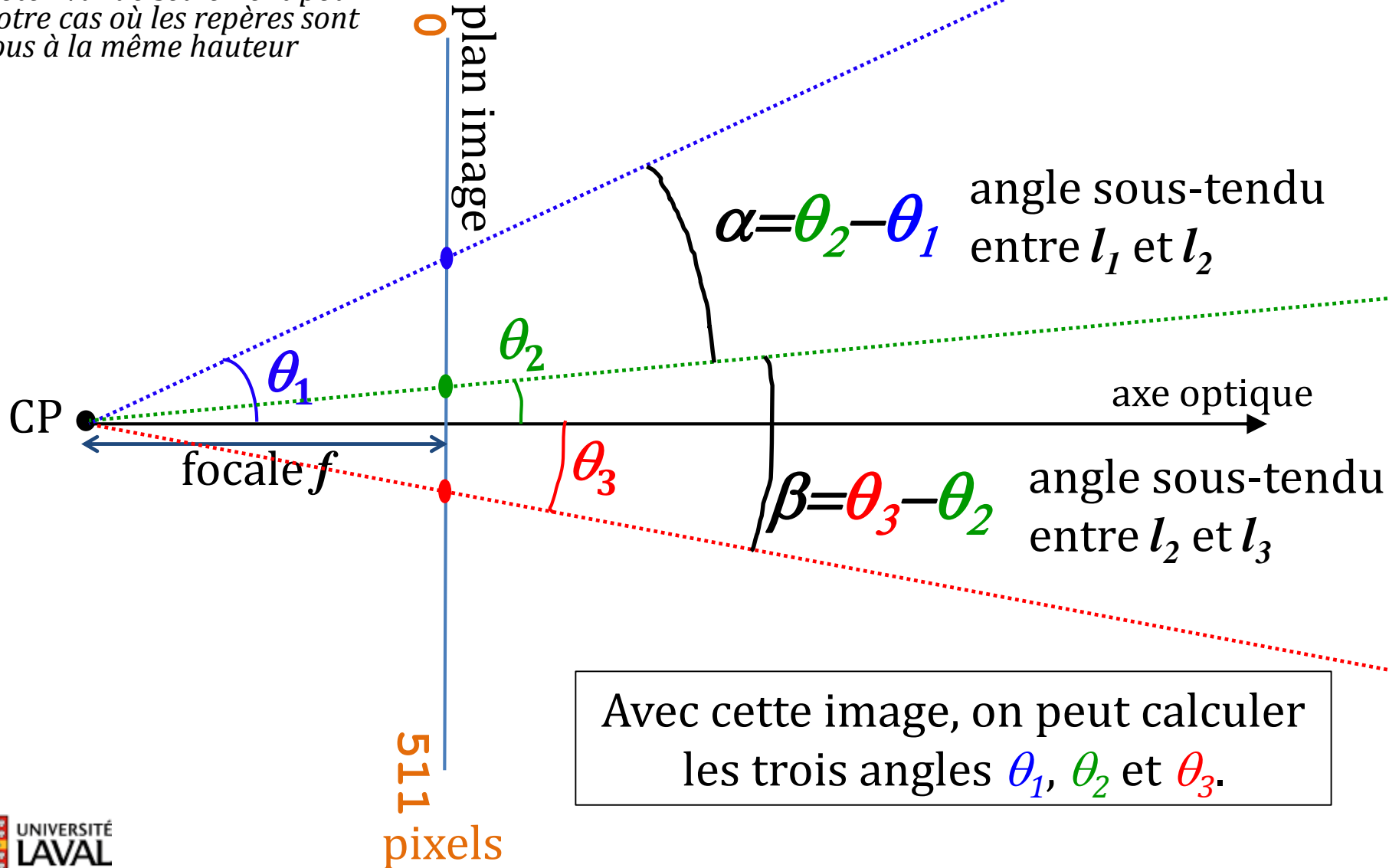
$$\tan \theta_1 = \frac{p_1}{f} \quad \Rightarrow \quad \theta_1 = \arctan\left(\frac{p_1}{f}\right)$$

note: la valeur de f est connue par la phase de calibration effectuée avant

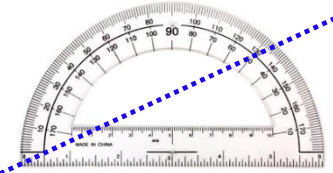
Caméra == rapporteur d'angles



Note : valide seulement pour notre cas où les repères sont tous à la même hauteur

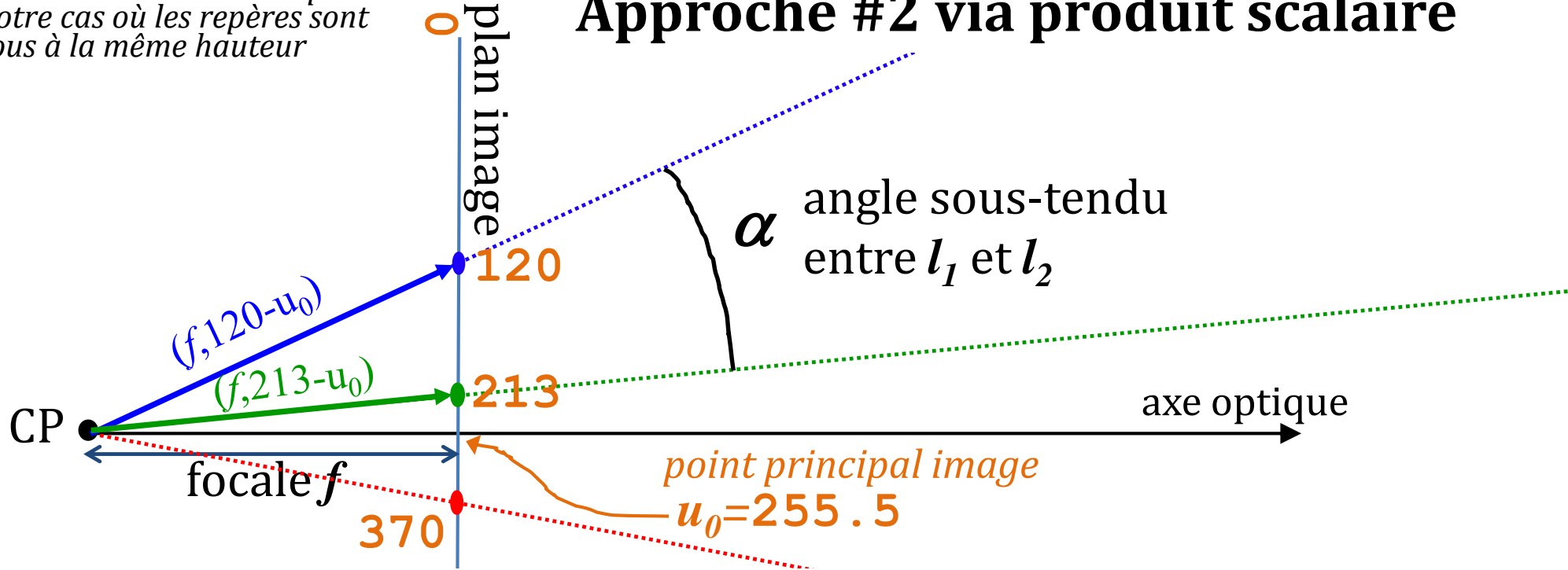


Caméra == rapporteur d'angles



Note : valide seulement pour notre cas où les repères sont tous à la même hauteur

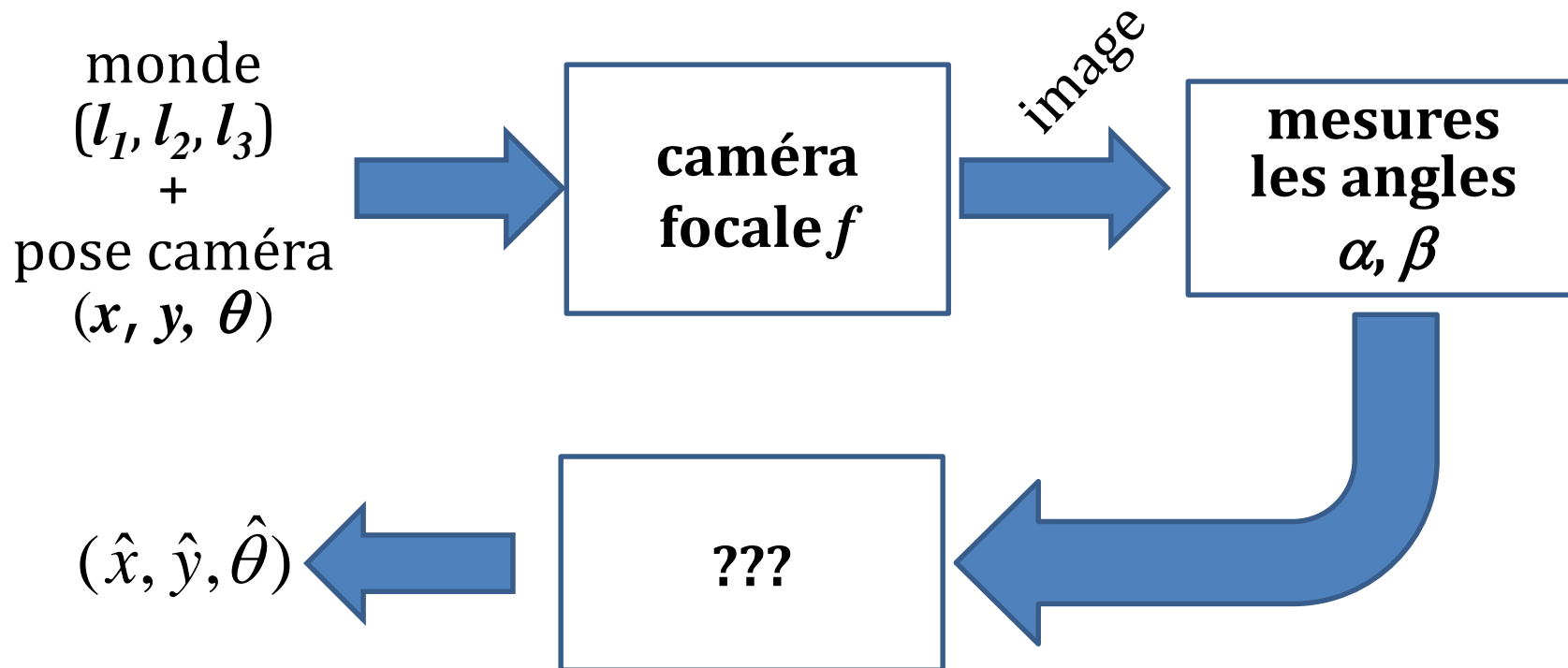
Approche #2 via produit scalaire



$$\alpha = \cos^{-1} \left(\frac{(f, -135.5) \cdot (f, -42.5)}{\|(f, -135.5)\| \|(f, -42.5)\|} \right) = 0.1590 \text{ rad}$$

511
pixels

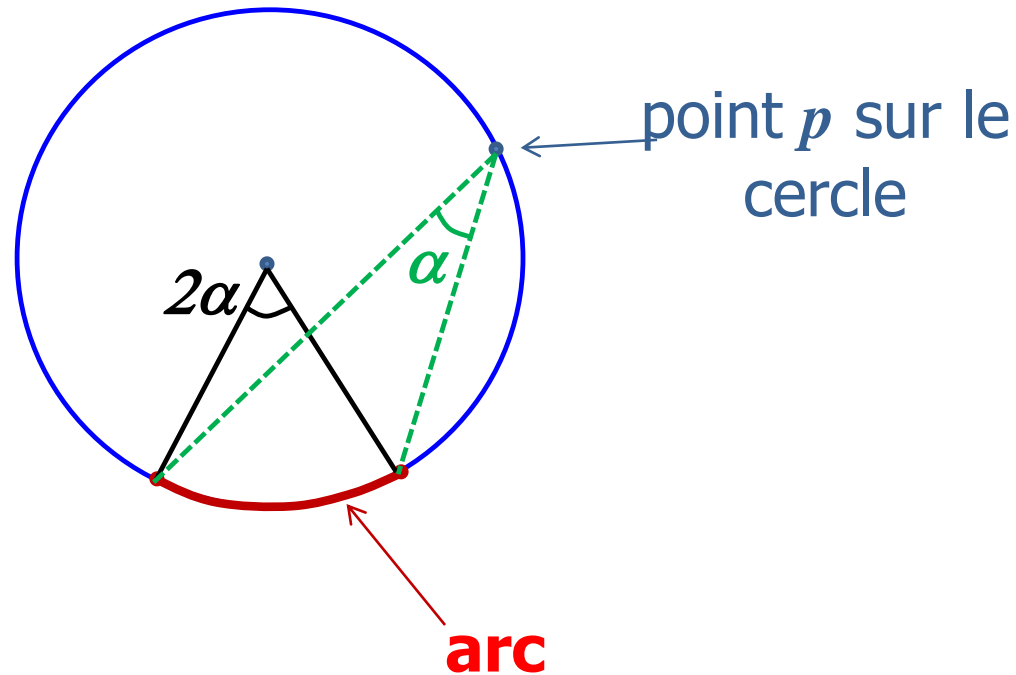
Localisation 2 D



Comment estimer la pose (x, y, θ) de la caméra?

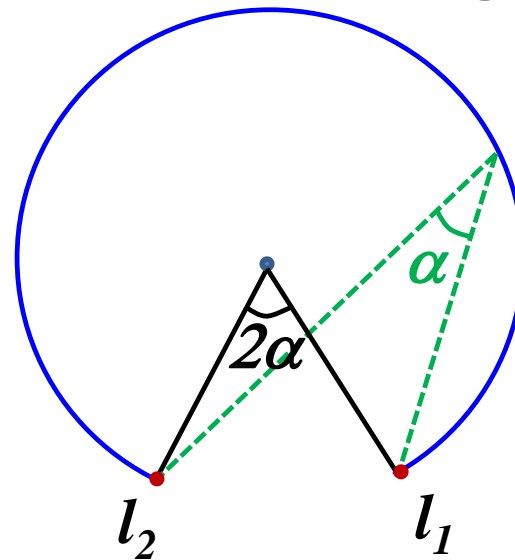
Théorème de l'angle inscrit et de l'angle au centre

- Pour un cercle, l'angle au centre mesure le double d'un angle inscrit interceptant le même **arc**.



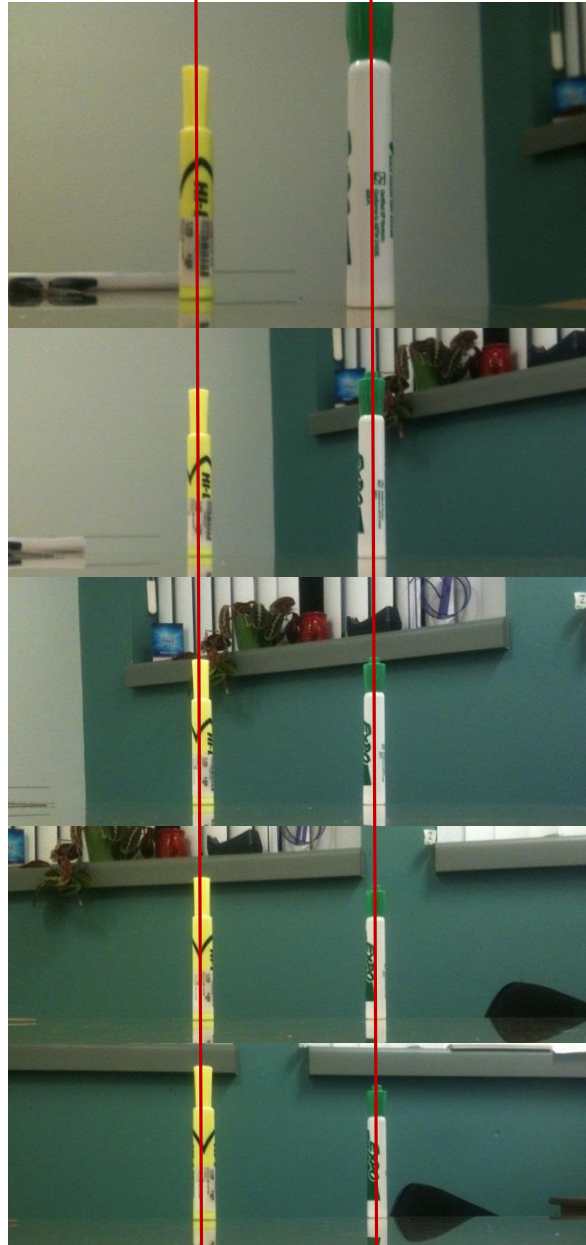
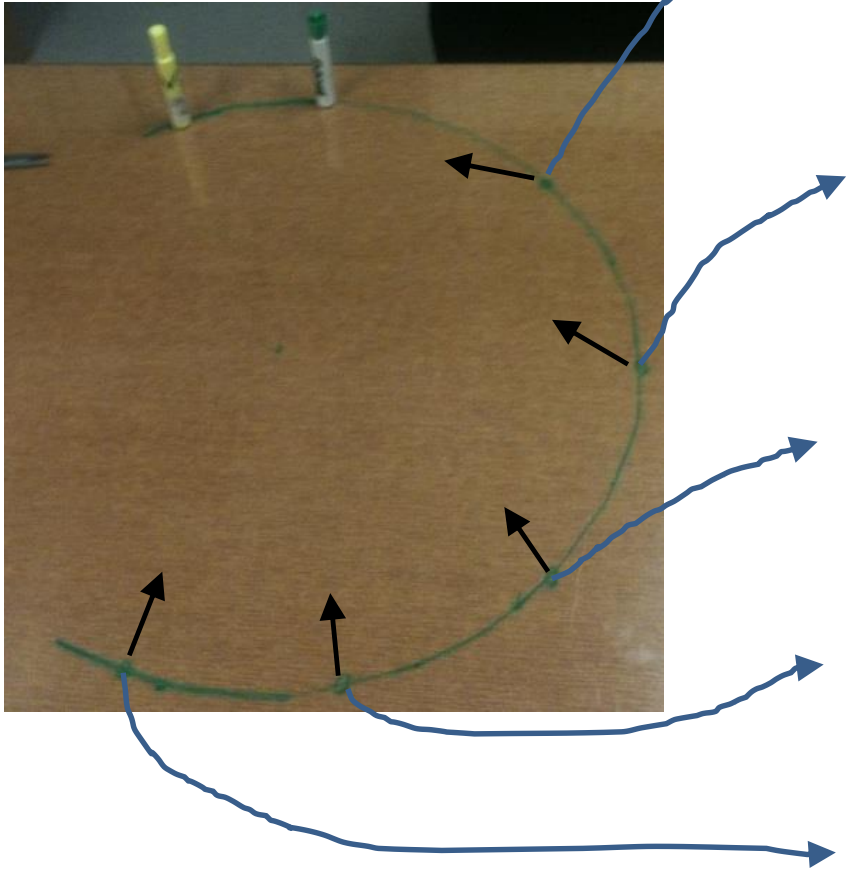
Corollaire : cercle unique pour l_1, l_2 et α

- Je connais la position de l_1, l_2 sur ma carte
- Je connais α , à partir de l'image



- On peut identifier un cercle unique avec ces trois paramètres : contrainte mathématique
- La caméra se trouve quelque part sur l'arc de cercle bleu

Exemple avec caméra sur un bureau



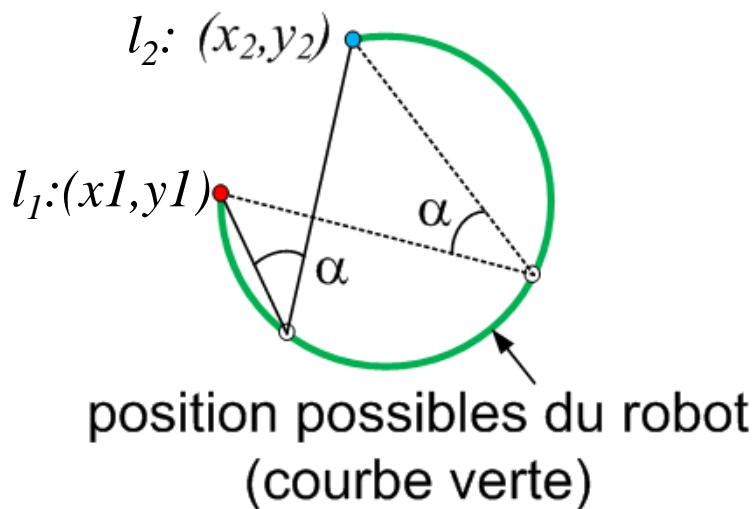
même angle α

(faire attention ici, car il faut toujours calculer les angles à partir du centre de l'image. Pour des fins de simplifications, je ne regarde ici que la distance entre les marqueur)

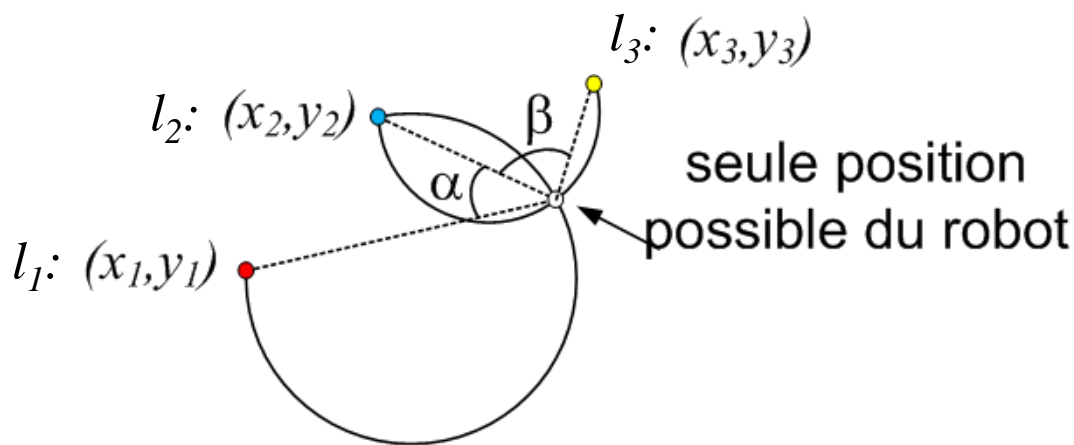
Localisation 2D par triangulation (angles)

- Angles α et β entre pt. de repères dans l'image de la caméra
- Connait position des points de repères l_1, l_2 et l_3

Deux repères, un angle α .

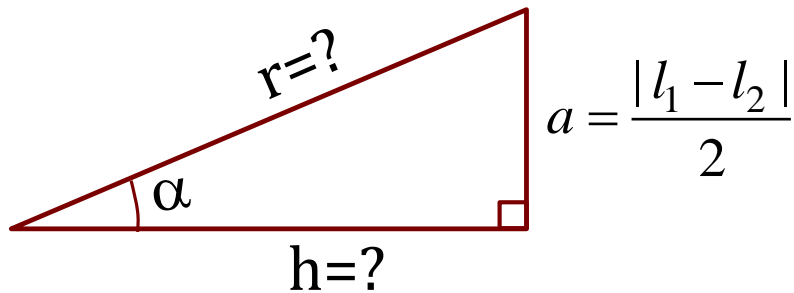


Trois repères, deux angles α, β .



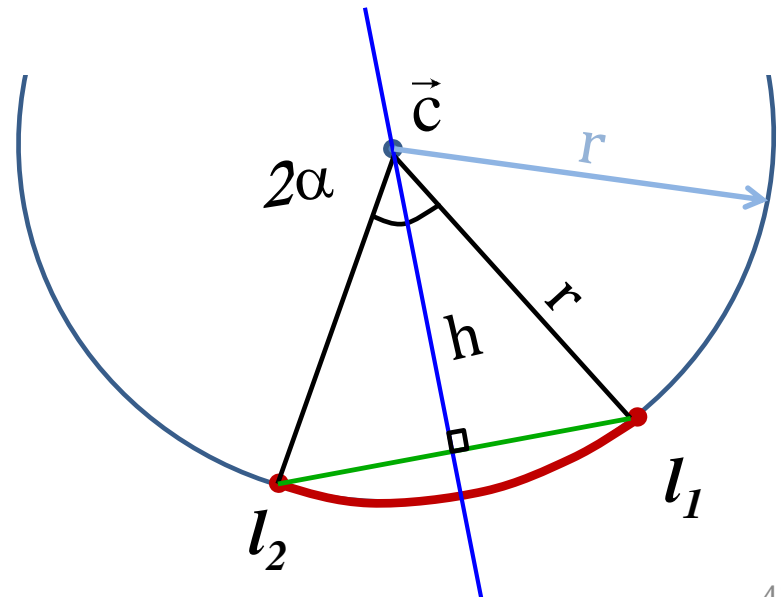
Identifier le cercle pour l_1, l_2 et α

- Un cercle se paramétrise par
 - position du centre $\vec{c} = (c_x, c_y)$ et rayon r
- Corde $l_2 l_1$
- \vec{c} est situé sur la médiatrice de la corde
- Triangle rectangle



$$r = \frac{a}{\sin \alpha} = \frac{|l_1 - l_2|}{2 \sin \alpha}$$

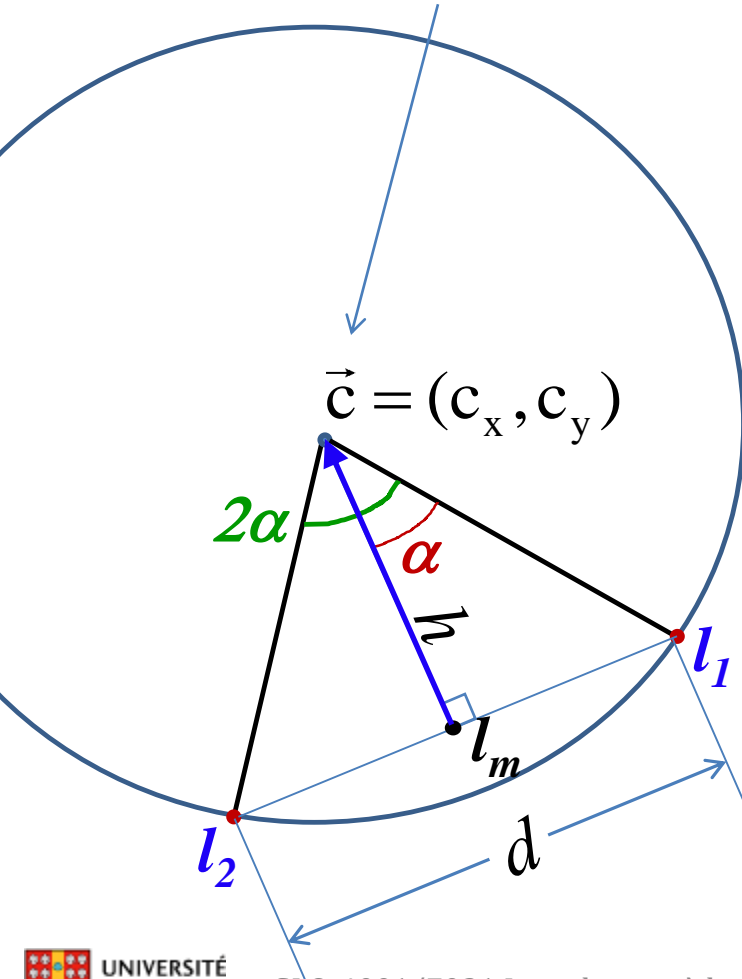
$$h = \frac{a}{\tan \alpha} = \frac{|l_1 - l_2|}{2 \tan \alpha}$$



Trouver le centre d'un des cercles

On connaît l_1, l_2 et α .

On cherche c



ratio altitude/base : $\frac{h}{d} = \frac{1}{2 \tan \alpha}$ car $\frac{d/2}{h} = \tan \alpha$

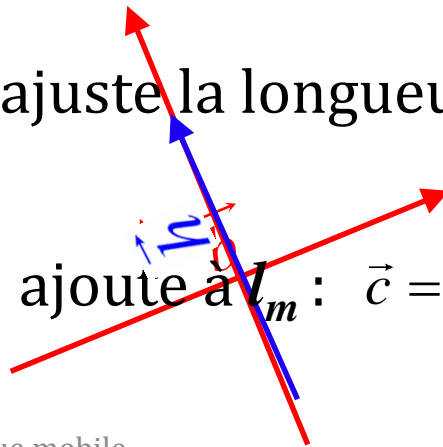
point milieu l_m : $\vec{l}_m = \frac{\vec{l}_1 + \vec{l}_2}{2}$

vecteur de la base : $\vec{b} = \vec{l}_1 - \vec{l}_2$

pivote de 90° : $\begin{bmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{bmatrix} \vec{b} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{b}$

ajuste la longueur: $\vec{h} = \left(\frac{1}{2 \tan \alpha}\right) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{b}$

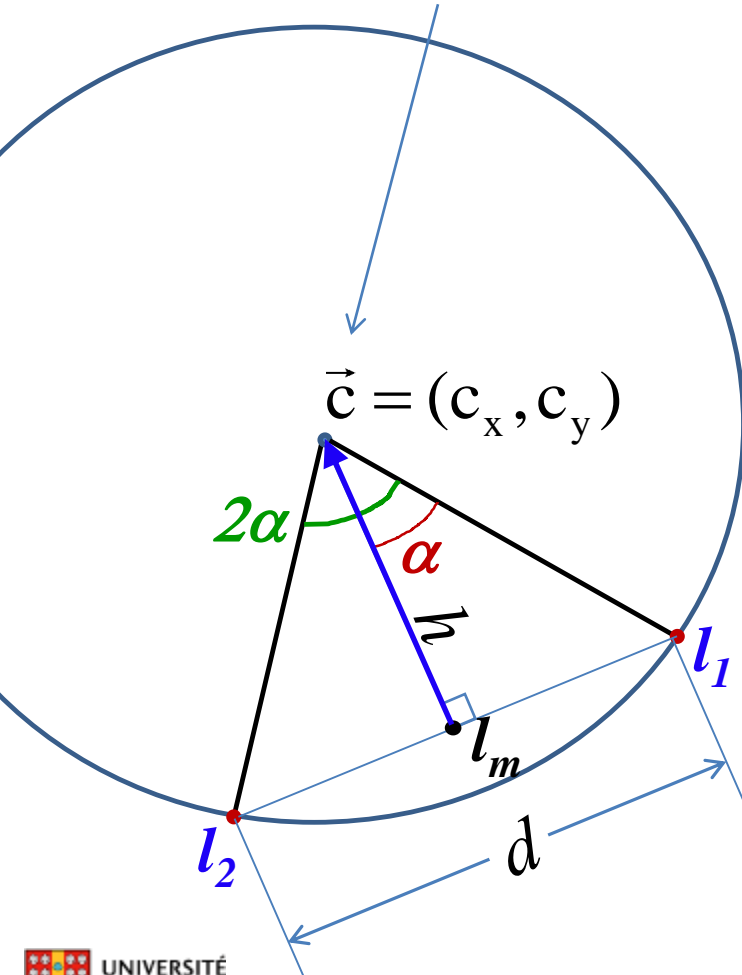
ajoute à l_m : $\vec{c} = \vec{l}_m + \vec{h} = \vec{l}_m + \left(\frac{1}{2 \tan \alpha}\right) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{b}$



Trouver le centre d'un des cercles

On connaît l_1, l_2 et α .

On cherche c



ratio altitude/base : $\frac{h}{d} = \frac{1}{2 \tan \alpha}$ car $\frac{d/2}{h} = \tan \alpha$

point milieu l_m : $\vec{l}_m = \frac{\vec{l}_1 + \vec{l}_2}{2}$

vecteur de la base : $\vec{b} = \vec{l}_1 - \vec{l}_2$

pivote de 90° : $\begin{bmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{bmatrix} \vec{b} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{b}$

ajuste la longueur: $\vec{h} = \left(\frac{1}{2 \tan \alpha}\right) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{b}$

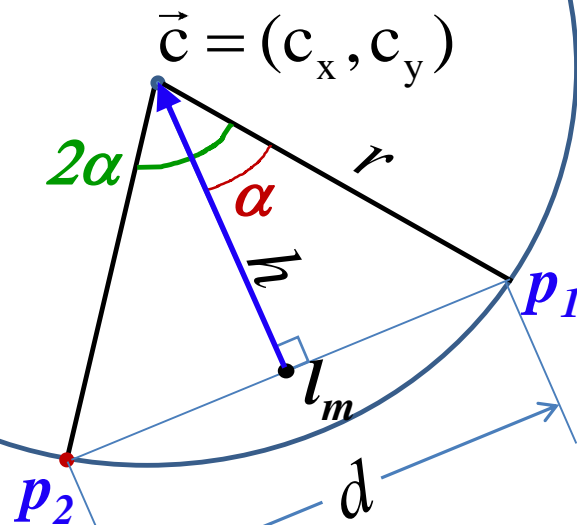
ajoute à l_m : $\vec{c} = \vec{l}_m + \vec{h} = \vec{l}_m + \left(\frac{1}{2 \tan \alpha}\right) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} (\vec{l}_1 - \vec{l}_2)$

Trouver le centre d'un des cercles

$$\vec{c} = \vec{l}_m + \left(\frac{1}{2 \tan \alpha}\right) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} (\vec{l}_1 - \vec{l}_2)$$

en matlab

```
c = [0 -1; 1 0] * (p1-p2) ./ (2*tan(alpha)) + 0.5 * (p2+p1);
```



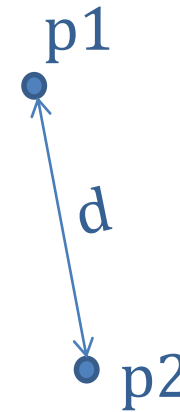
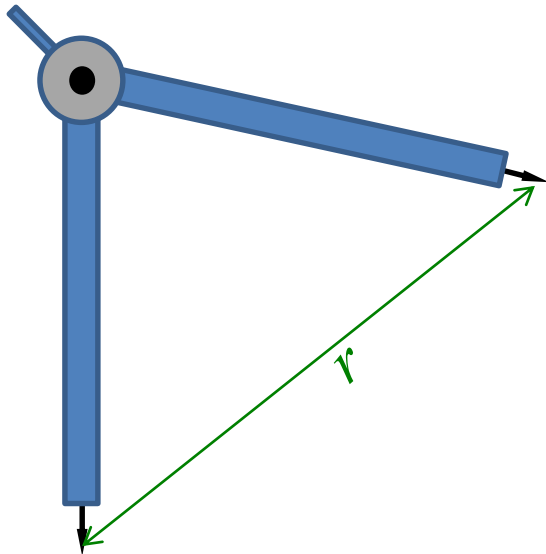
faire attention à l'ordre des points p_i et l'orientation des matrices :

```
p1=[0.2 1.2]';
```

note: j'ai choisis p ici car l ressemble trop au chiffre « un » en lettrage courrier new.

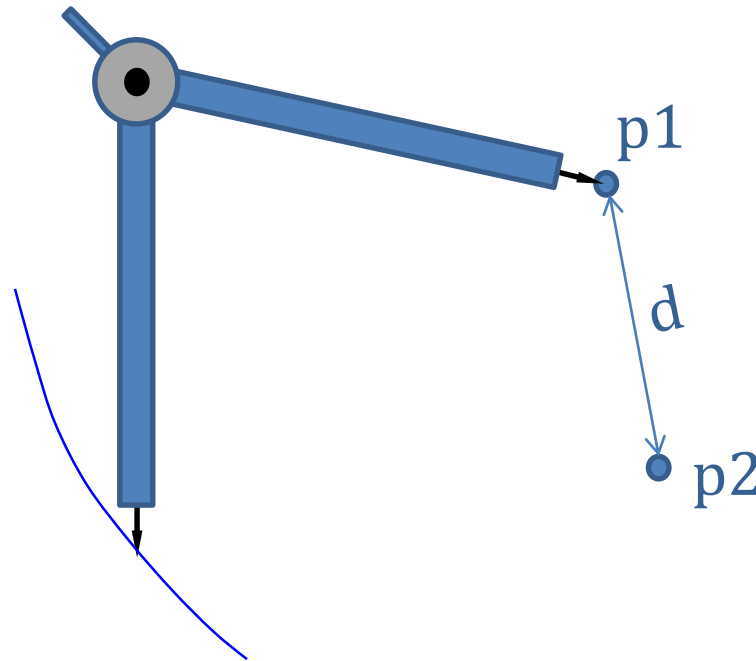
Sur papier

- Compas + règle
- Problème localisation, pour trouver centre du cercle $r = \frac{d}{2 \sin \alpha}$



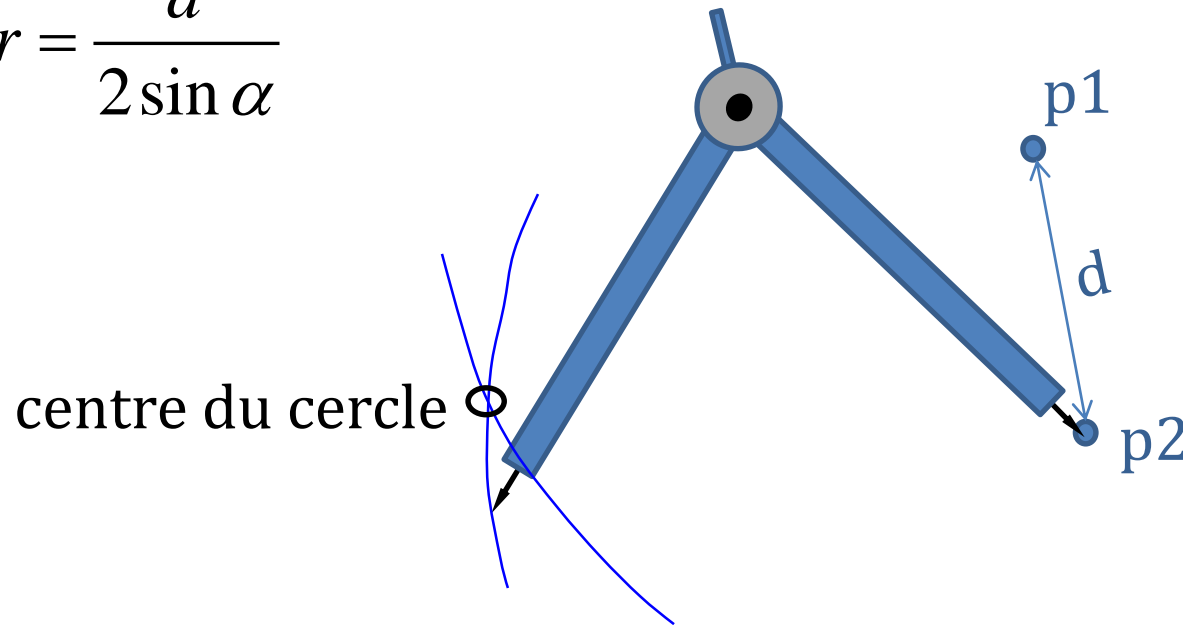
Sur papier

- Compas + règle
- Problème localisation, pour trouver centre du cercle $r = \frac{d}{2 \sin \alpha}$



Sur papier

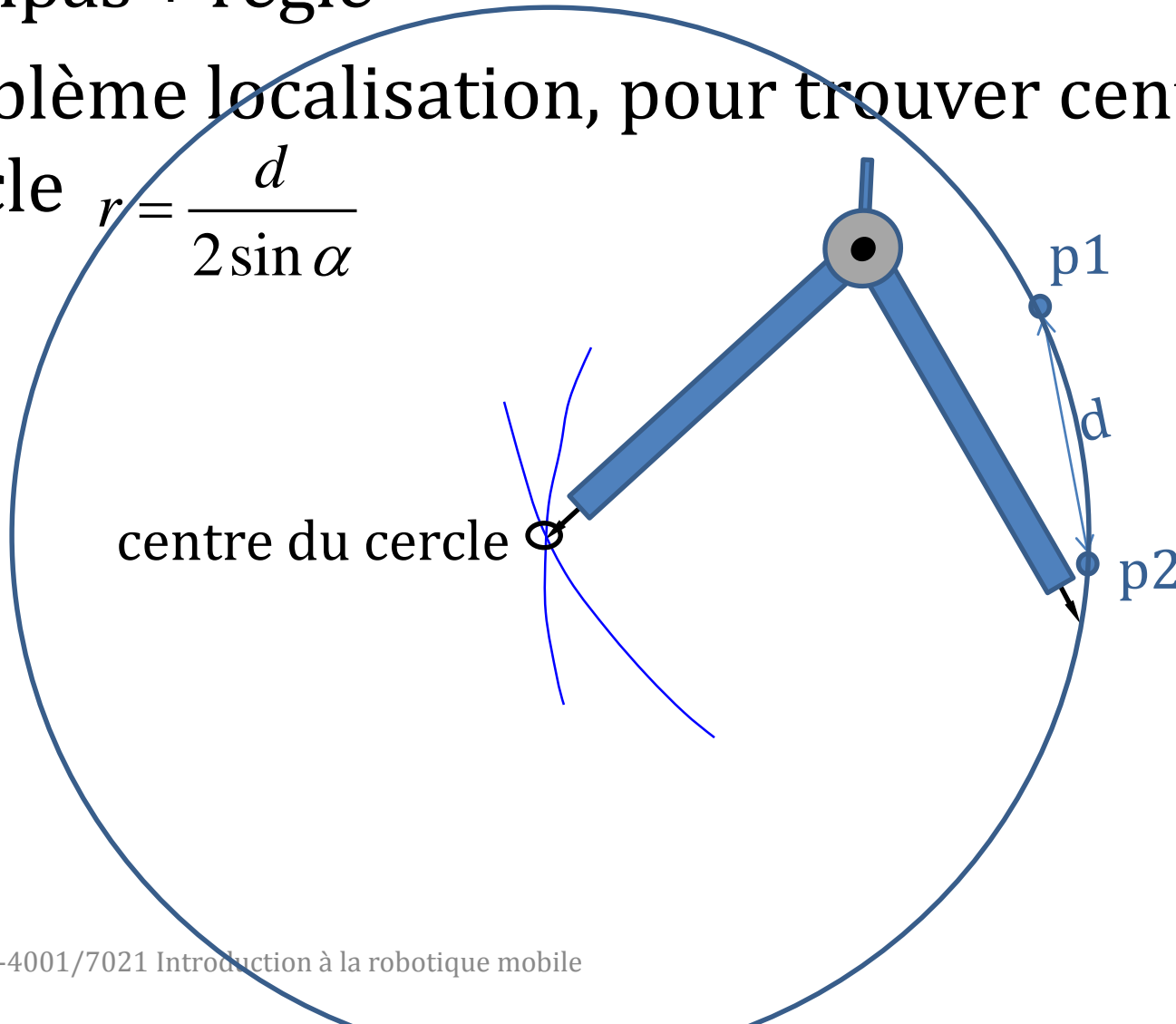
- Compas + règle
- Problème localisation, pour trouver centre du cercle $r = \frac{d}{2 \sin \alpha}$



Sur papier

- Compas + règle
- Problème localisation, pour trouver centre du cercle

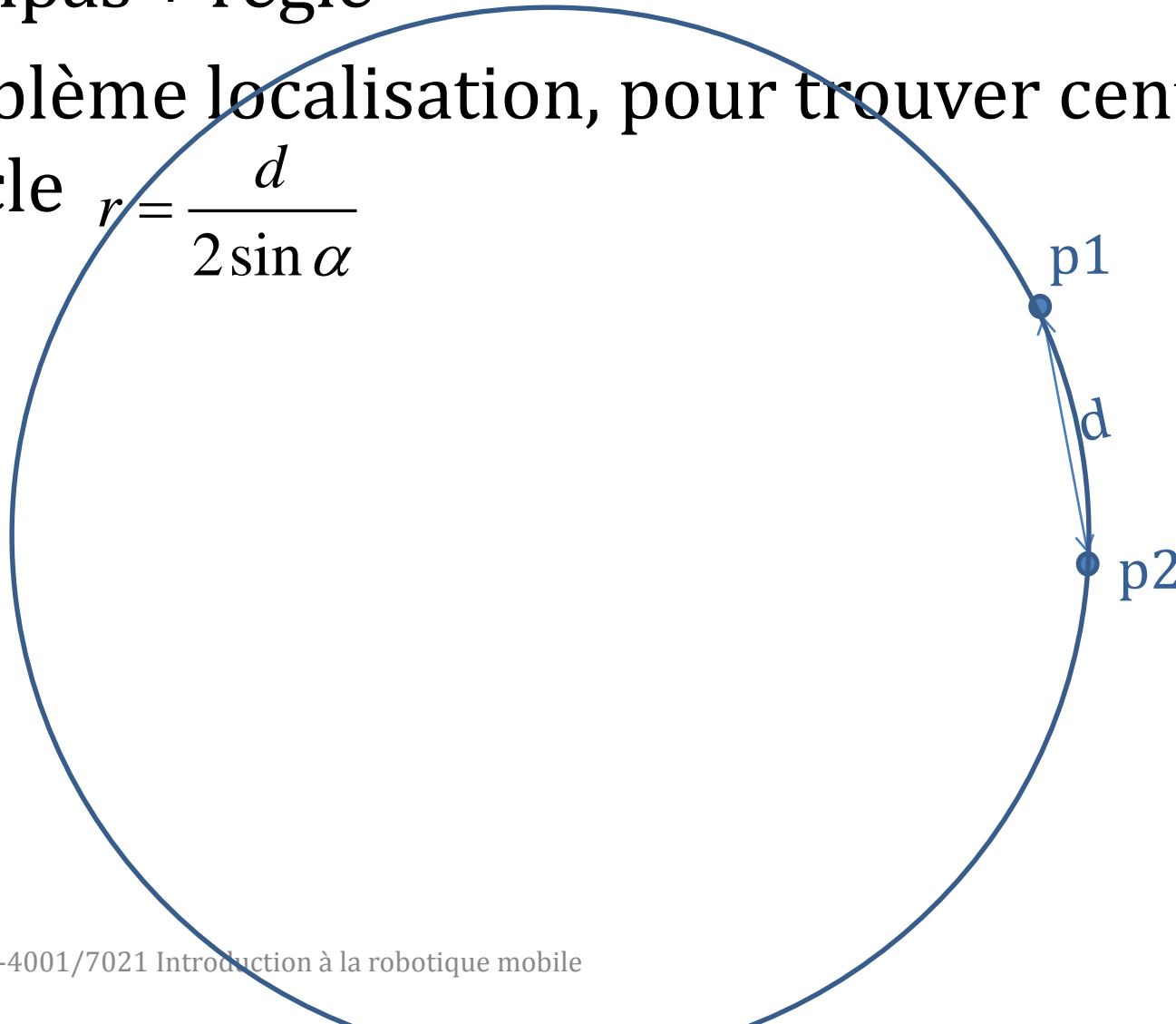
$$r = \frac{d}{2 \sin \alpha}$$



Sur papier

- Compas + règle
- Problème localisation, pour trouver centre du

cercle $r = \frac{d}{2 \sin \alpha}$



Solutions

Trouver la position du robot à partir de l_1, l_2, l_3, α et β ?

- Approche géométrique
 - Règle et compas

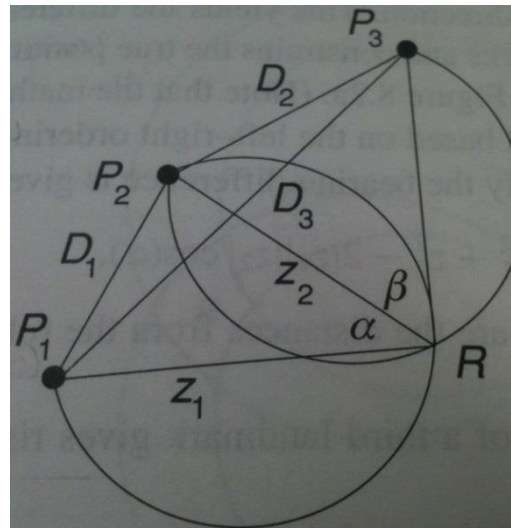


à l'examen!

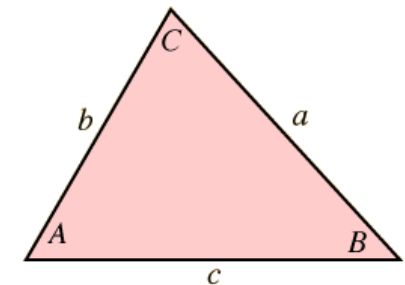
- Satisfaction de contraintes non-linéaires (p.246 manuel)

loi des cosinus

$$D_1^2 = z_1^2 + z_2^2 - 2|z_1||z_2|\cos(\alpha)$$
$$D_2^2 = z_2^2 + z_3^2 - 2|z_2||z_3|\cos(\beta)$$
$$D_3^2 = z_1^2 + z_3^2 - 2|z_1||z_3|\cos(\alpha + \beta)$$



$$c^2 = a^2 + b^2 - 2ab \cos C$$

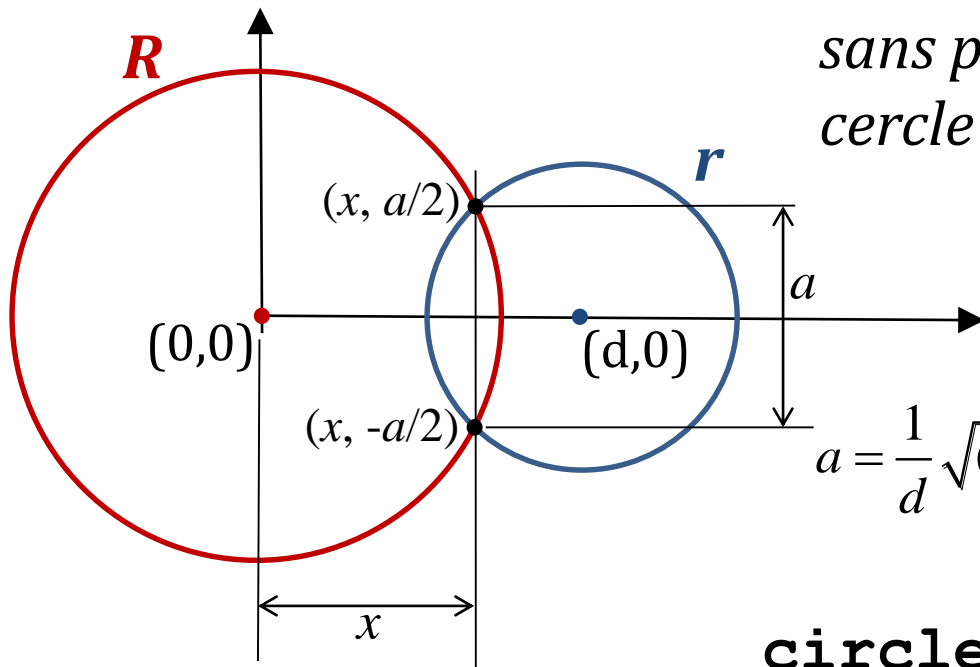


- Recherches itératives
- *Newton-Raphson*

A Comprehensive Study of Three Object Triangulation, Cohen & Koss, Proceedings of the SPIE, Volume 1831, p. 95-106 (1993)

Approche géométrique

- Trouver le centre et rayon des deux cercles
- Utiliser les formules de l'intersection de deux cercles <http://mathworld.wolfram.com/Circle-CircleIntersection.html>



sans perte de généralité, si vous avez un cercle à (0,0) et l'autre à (d,0)

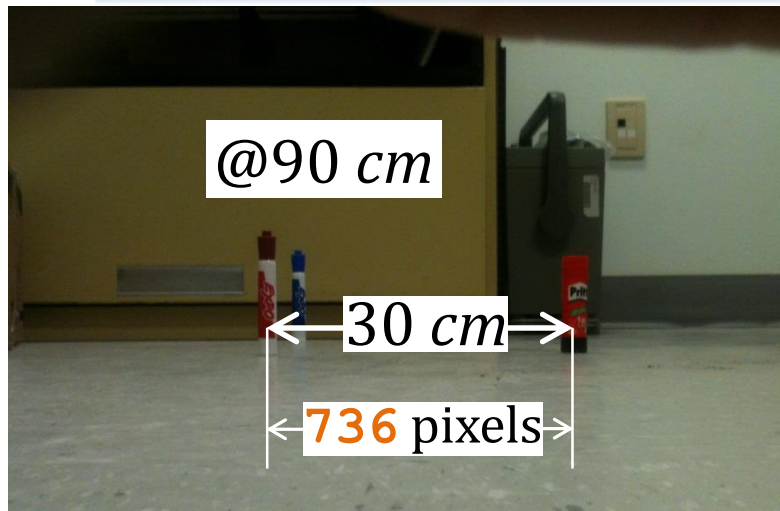
$$x = \frac{d^2 - r^2 + R^2}{2d}$$

$$a = \frac{1}{d} \sqrt{(-d + r - R)(-d - r + R)(-d + r + R)(d + r + R)}$$

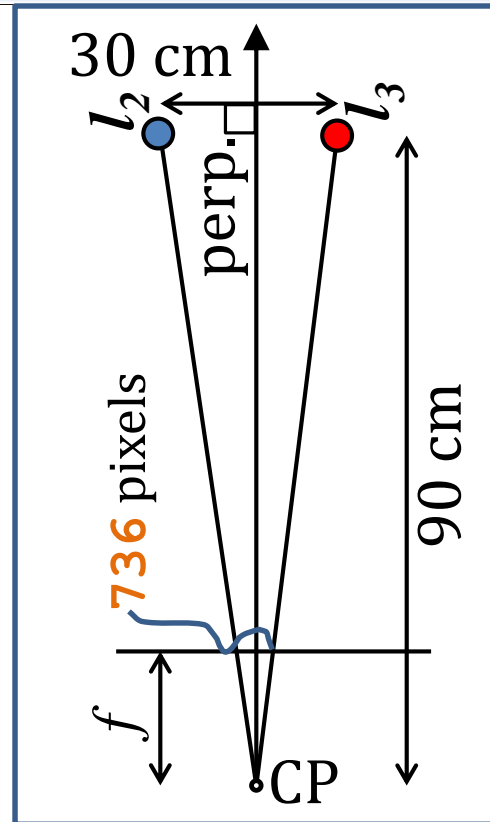
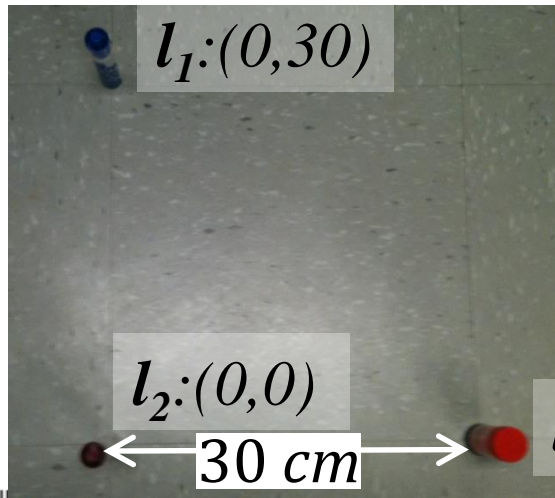
Encore mieux! Google
circleintersect.m par Peter Kovési

ou **circirc** dans mapping toolbox de matlab

Exemple localisation : calibration de f



vue de haut

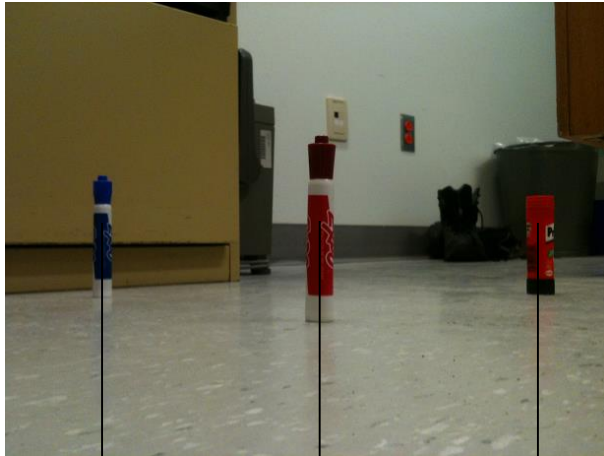


(Notez que cette étape n'est faite qu'une seule fois dans un système robotique, au préalable)

$$\frac{f}{736 \text{ pixels}} = \frac{90 \text{ cm}}{30 \text{ cm}}$$

$$f = \frac{90}{30} 736 \text{ pixels} = 2208 \text{ pixels}$$

Exemple localisation : calcul des angles



328

1066

1815

← Position horizontale des objets dans l'image

- Image de 2048 pixels horizontaux
- Centre : $u_0=1024,5$ (normalement on doit calibrer cette valeur aussi!)
- Distance focale $f = 2208$
- Trouve centre des objets

(Rappel : parce que le robot se déplace à plat, localisation 2D)

- Calcul les trois angles θ_1 , θ_2 et θ_3 :

$$\theta_1 = \arctan\left(\frac{328 - 1024.5}{f}\right) = -0.3056 \text{ rad}$$

$$\theta_2 = \arctan\left(\frac{1066 - 1024.5}{f}\right) = 0.0188 \text{ rad}$$

$$\theta_3 = \arctan\left(\frac{1815 - 1024.5}{f}\right) = 0.3438 \text{ rad}$$

- Calcul les angles α et β :

$$\alpha = \theta_2 - \theta_1 = 0.3244 \text{ rad}$$

$$\beta = \theta_3 - \theta_2 = 0.3250 \text{ rad} \text{ ou}$$

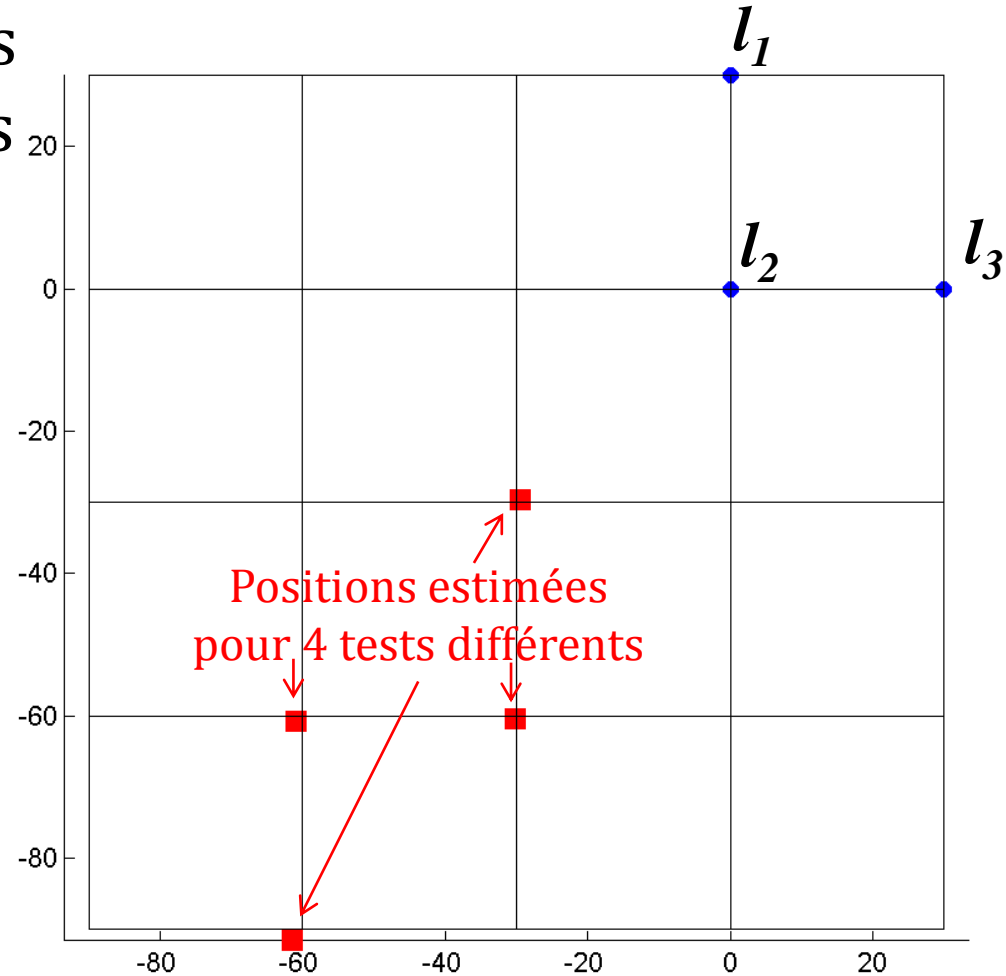
$$\alpha = \cos^{-1}\left(\frac{(f, -696.5) \cdot (f, 41.5)}{\|(f, -696.5)\| \|(f, 41.5)\|}\right)$$

Similaire pour β

- Résout avec l'une des méthodes mentionnées auparavant

Exemple de localisation : résultats

- Les **positions calculées** sont très près des positions réelles (coins des tuiles)
- Mais la précision diminue grandement à mesure que l'on s'éloigne des points de repère ou pour certaines positions



Geometric Dilution of Precision (GDOP)

Position A Position B

