La caméra:modèle d'acquisition

Références utiles:

Sonka et al:

Sections: 3.4.2, 3.4.3, 11.2.1, 2.5

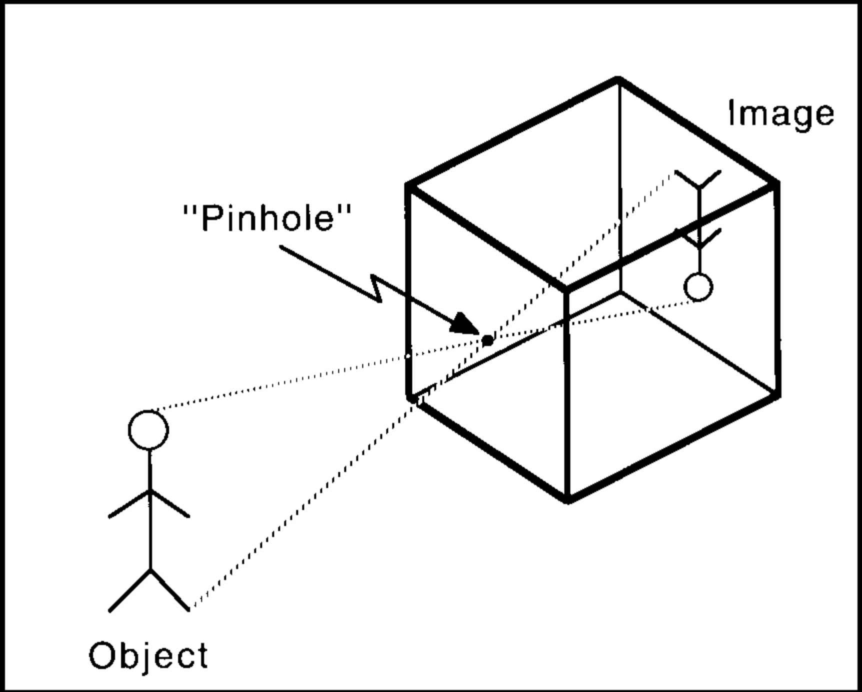
. Site commercial:

http://www.1394imaging.com/en/resources/whitepapers/

Patrick Hébert & Denis Laurendeau (Dernière révision : juin 2016)

Projection de perspective

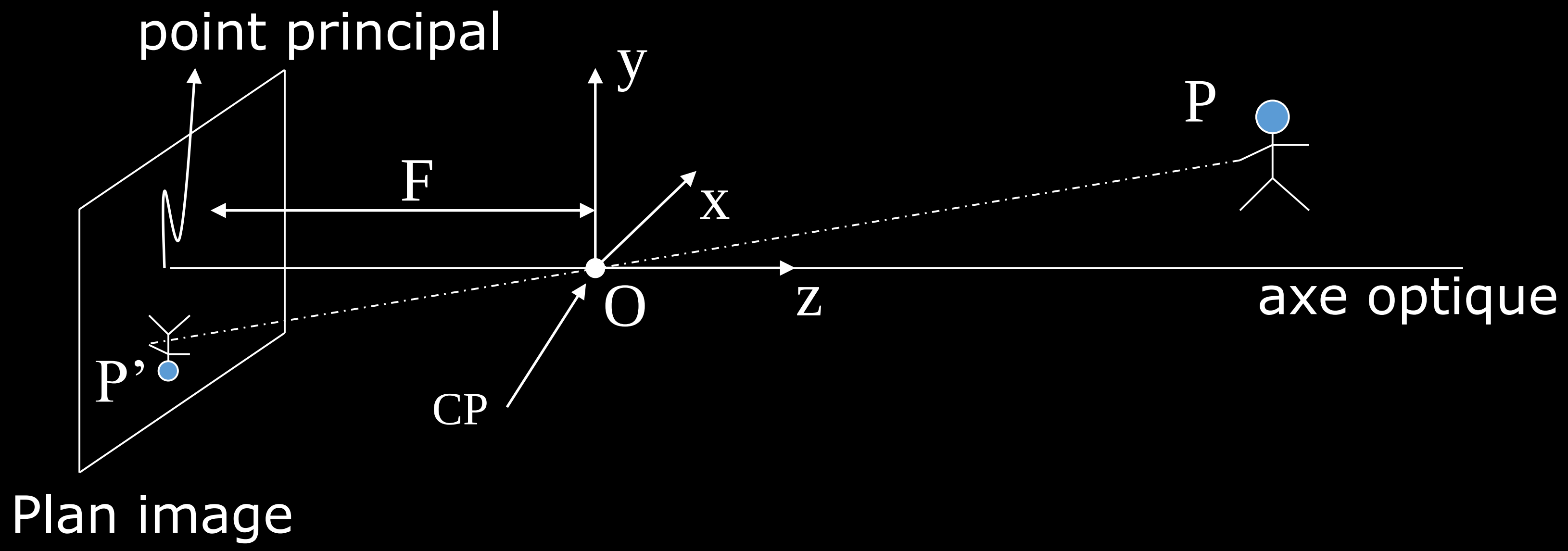
- Le modèle du sténopé (pinhole)
  - Brunelleschi (début 15ième siècle)



*tirée de Nalwa

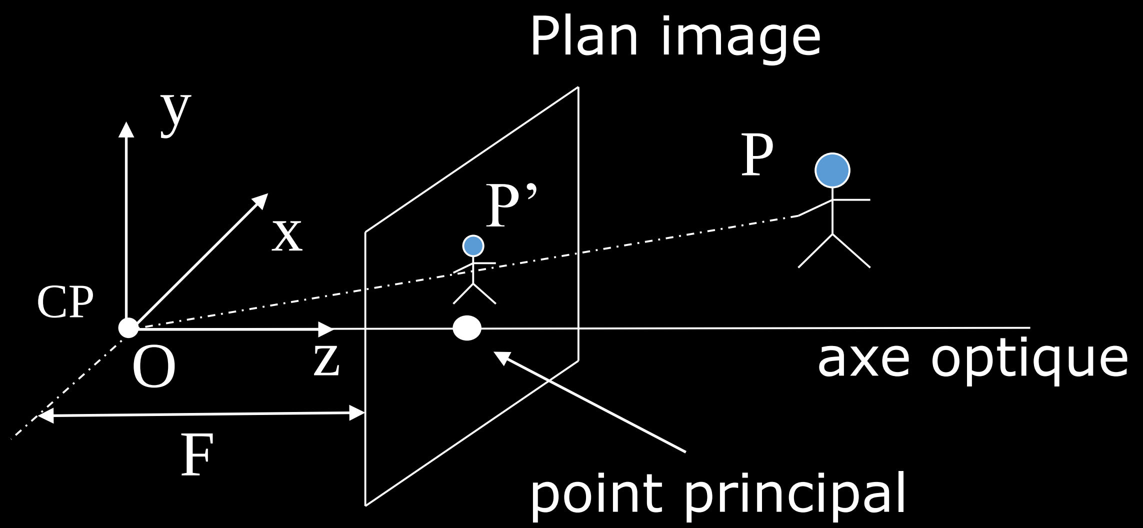
www.niepce.com

Le sténopé inverseur (à un repère de coordonnées)



- un seul rayon atteint chacun des points du plan image (profondeur de champ infinie)
- l'ouverture ne peut être infiniment petite

Le sténopé non-inverseur (à un repère de coordonnées)



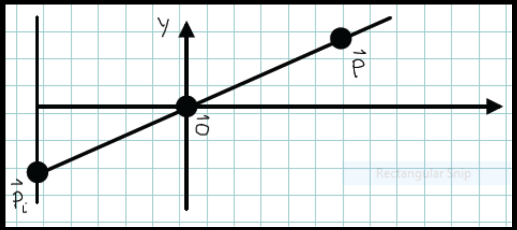
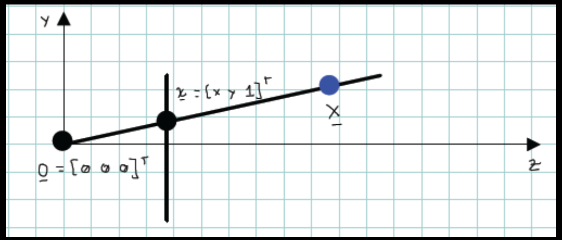
- plus commode sur le plan mathématique car aucune inversion d'image
- ne peut évidemment être réalisé en pratique

Les équations

- Deux Buts
  1. Prédiction de la position d'un point P' dans l'image à partir de P, connu dans le repère de la caméra (projection de perspective **directe**)

  2. Calcul de l'équation d'un rayon (projecteur) à partir d'un point image (projection de perspective **inverse**)

     - OP' = λ OP
     - équation paramétrique d'un projecteur

	Sténopé inverseur	Sténopé non-inverseur
Coordonnées de P’	$x' = -\frac{FX}{Z}$ $\quad y' = -\frac{FY}{Z}$	$x' = \frac{FX}{Z}$ $\quad y' = \frac{FY}{Z}$
Equation paramétrique d’un projecteur		
	$$\underline{P} = \underline{p_i} + \lambda \frac{(\underline{O} - \underline{p_i})}{\lVert \underline{O} - \underline{p_i} \rVert}$$	$$\underline{X} = \underline{O} + \lambda(\underline{x} - \underline{O})$$ $$\underline{X} = \lambda \underline{x}$$

Remarques

- L'équation de projection est une forme non-linéaire en z

- On souhaiterait une forme: x' = **P** X

- Comment? En passant par les coordonnées homogènes dans un espace de dimension
  3 + 1 = 4

Projection de perspective en coordonnées homogènes

$$\underline{\tilde{x}} = \underline{\underline{\tilde{P}}}\,\underline{\tilde{X}}$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{F} & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$\Rightarrow$

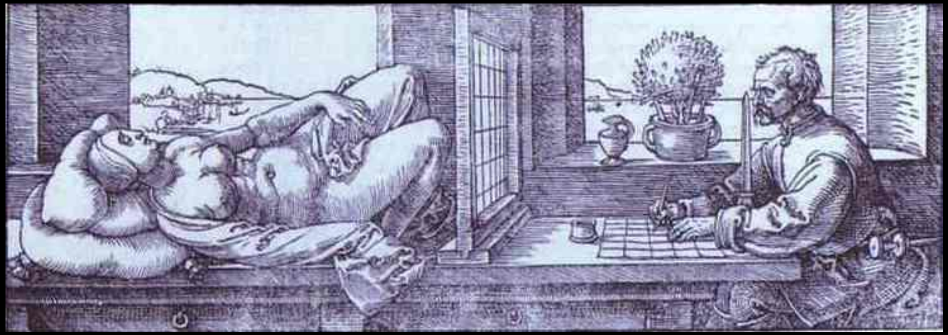
$$x' = X$$
$$y' = Y$$
$$z' = Z$$
$$w = \frac{Z}{F}$$

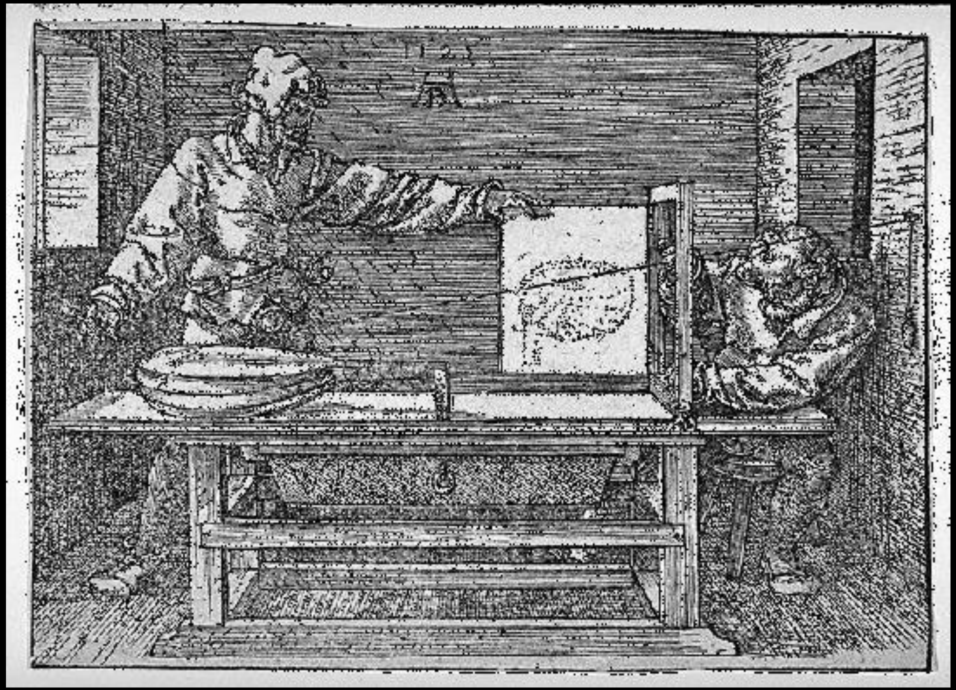
$\Rightarrow$

En coordonnées réelle

$$x' = \frac{FX}{Z}$$
$$y' = \frac{FY}{Z}$$
$$z' = F$$

Albrecht Durer – 16^{e} siècle





Définitions: *point de fuite* et *ligne d'horizon*

- Le point de fuite est un point de l'image (en projection de perspective) où toutes les droites parallèles selon une orientation en 3D, convergent. C'est donc l'image d'un point à l'infini: x = (a, b, 0). Ex: le couloir!
- Ce point est sur l'image à l'endroit où une droite de même orientation passe par le CP.
- Il existe un point de fuite par orientation de droites 3D.
- La ligne d'horizon est l'intersection du plan image et d'un plan, passant par le CP, qui est parallèle à un plan contenant un ensemble d'éléments observés - des droites d'un plan 3D par exemple.

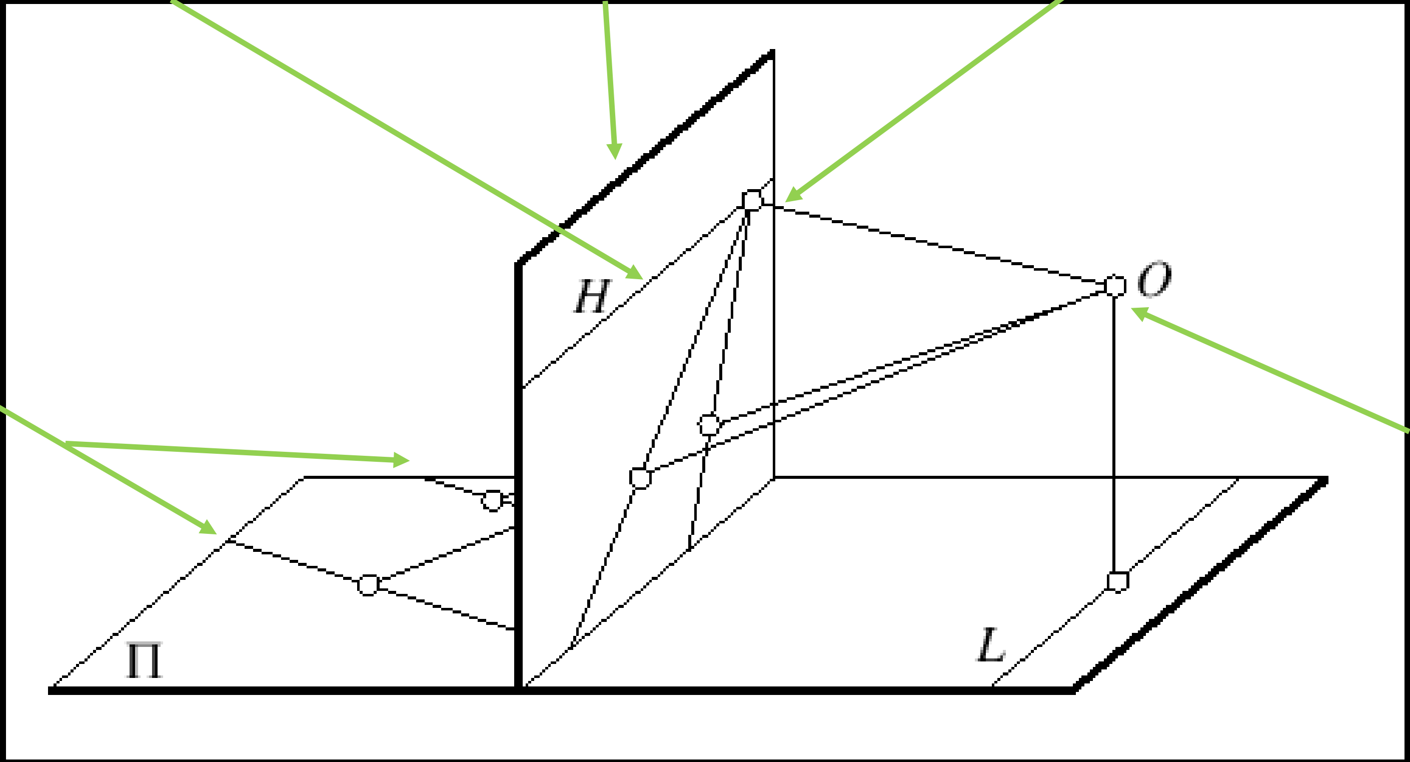
Illustration : *point de fuite* et *ligne d'horizon*

Ligne d’horizon

Plan image

Point de fuite

Groupe de droites parallèles dans un plan



CP

*tirée de Forsyth

Le point de fuite est **l'intersection** dans le plan image de droites **parallèles** dans le monde qui s'intersectent à **l'infini**

En coordonnées homogènes, les équations de deux droites parallèles dans un plan sont:

$$\underline{\tilde{l}} = [a \quad b \quad c]^t$$

$$\underline{\widetilde{l'}} = [a \quad b \quad c']^t$$

l'intersection de ces deux droites dans le monde est donnée par:

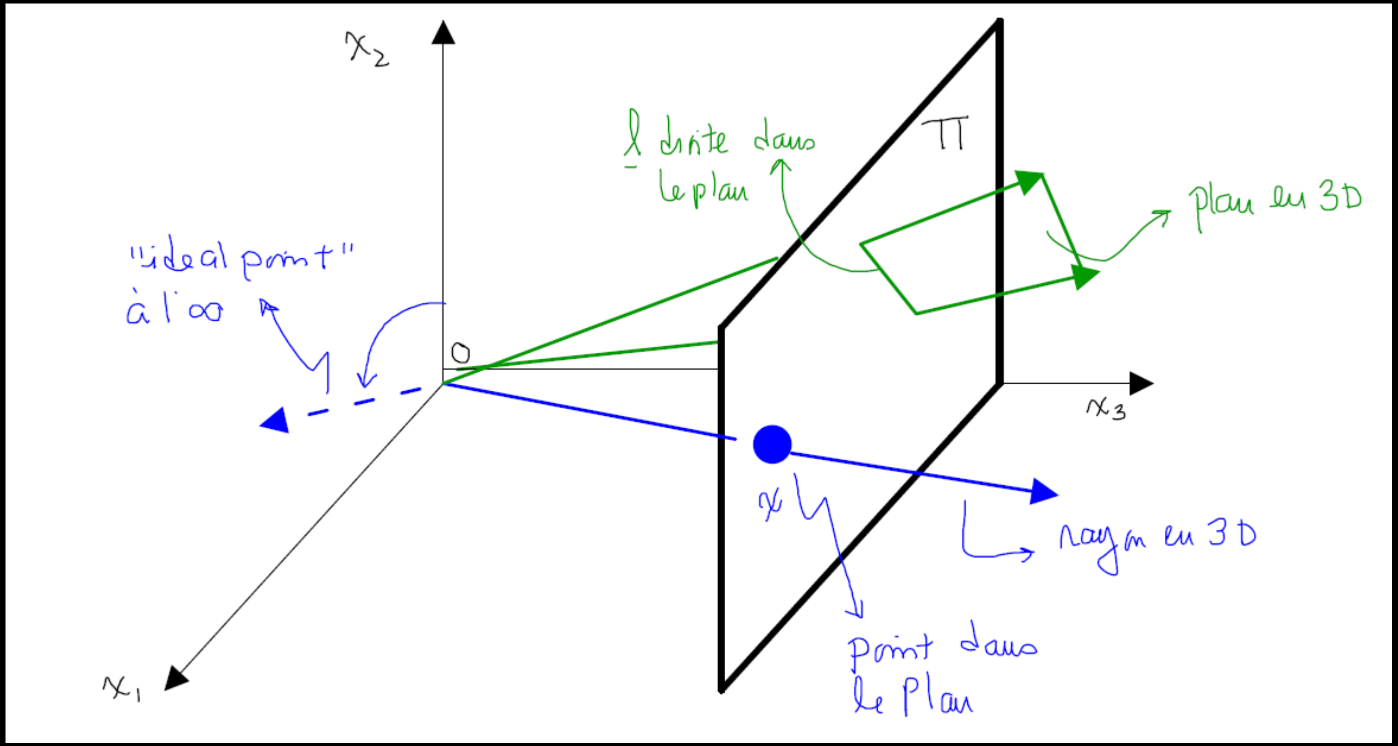
$$\underline{\tilde{l}} \times \underline{\tilde{l}}' = [(bc' - bc) \quad ac - ac' \quad 0]^t = [b \quad -a \quad 0]^t$$

$$\begin{bmatrix} b/0 \\ -a/0 \end{bmatrix}^t$$

qui est bien à l'infini
en coordonnées réelles

Représentation des points à l'infini

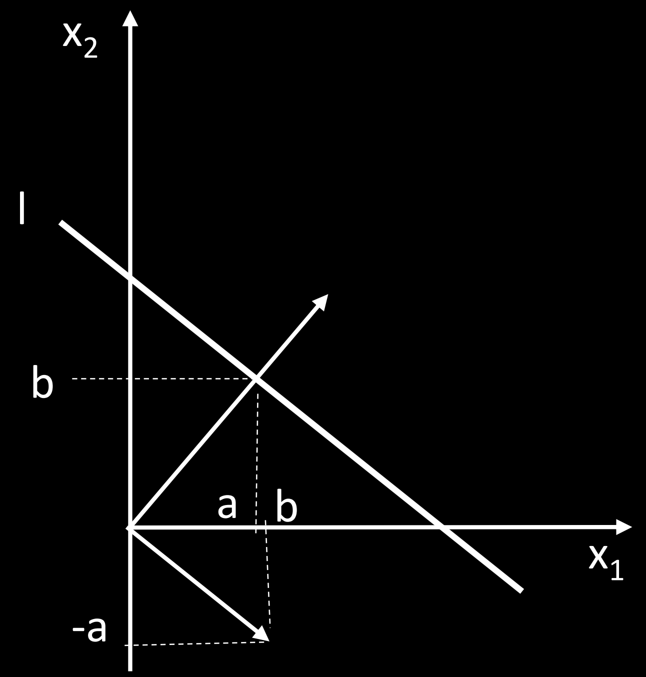
Considérons le plan projectif Π^2



- un rayon passant par l'origine de R^3 correspond à un point $\underline{x}$ dans Π^2
- un plan passant par l'origine de R^3 correspond à une droite $\underline{l}$ dans Π^2
- les droites dans le plan $x_1 - x_2$ de R^3 correspondent aux points à l'infini ("ideal points")
- le plan $x_1 - x_2$ de R^3 correspond à la droite à l'infini
- un point à l'infini a les coordonnées $[x_1 \;\; x_2 \;\; 0]^t$
- la droite $\underline{l}_\infty = [0 \quad 0 \quad 1]^t$ est appelée "droite à l'infini"

Les points à l’infini résident sur cette droite à l’infini

$$\underline{x}^t\,\underline{l} = [x_1 \quad x_2 \quad 0]\begin{bmatrix}0\\0\\1\end{bmatrix} = 0$$



Les droites suivantes intersectent la droite à l’infini

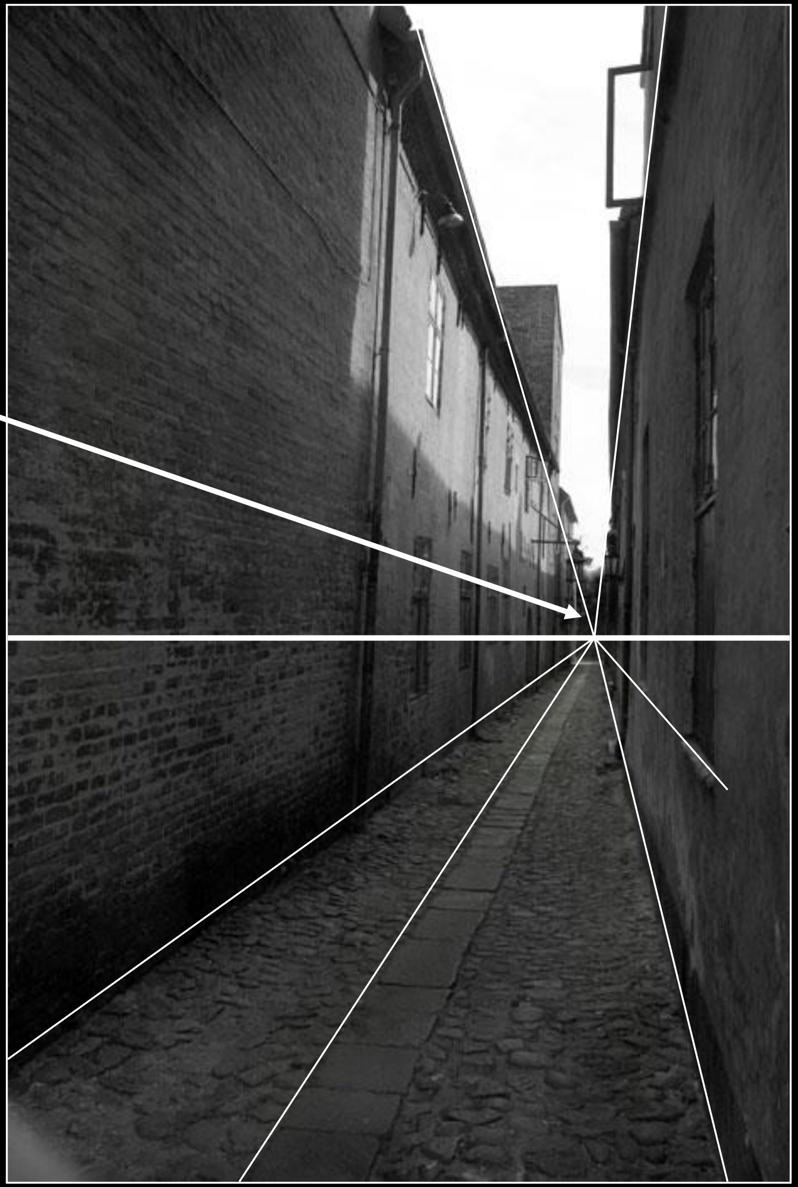
$$\underline{\tilde{l}} = [a \quad b \quad c]^t$$
$$\underline{l}' = [a \quad b \quad c']^t$$
$$\underline{l}_{\infty} = [0 \quad 0 \quad 1]^t$$

$\Rightarrow$

$$[b \quad -a \quad 0]^t$$

En notation non-homogène, [b -a] est un vecteur tangent à la droite [a b c] et perpendiculaire au vecteur [a b] normal à la droite. Quand la direction de la droite varie, le point à l’infini $[b \ \ -a \ \ 0]^t$ se déplace sur la droite à l’infini (dans le plan $x_1 - x_2$)

Cas familier : ensembles de droites parallèles dans le monde 3D qui se rencontrent dans l'image



Autre exemple

Photo T. Fuhrmann

Simplifications de la projection de perspective :

Projections affines

- Considérons $m = F/z$ comme un facteur d'échelle (magnification)
  - $x' = -m X$
  - $Y' = -m Y$
- Si la scène observée est relativement peu profonde selon z (Δz est petit), alors on considère m constant (perspective faible)
- Si z est environ constant dans la scène, alors on utilise un modèle de projection orthographique
  - $x' = X$ (après normalisation des coordonnées images telles que $m = -1$)
  - $y' = Y$

Représentations matricielles

Perspective

$$\underline{\underline{\tilde{P}}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{F} & 0 \end{bmatrix}$$

Perspective faible

$$\underline{\underline{\tilde{P}}}_f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{\bar{Z}}{F} \end{bmatrix}$$

Orthographique

$$\underline{\underline{\tilde{P}}}_o = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Généralisation : projection de perspective à plusieurs référentiels

Le modèle de sténopé non-inverseur (à un référentiel de coordonnées placé au centre de projection) pour modéliser la projection de perspective est intéressant, mais peu pratique:

- on doit exprimer les coordonnées des points objets dans le repère du sténopé
- ce repère n’est pas accessible (situé à l’intérieur du boîtier de la caméra)
- ce repère se déplace avec la caméra...à chaque mouvement de la caméra, on doit réexprimer les coordonnées objets dans le nouveau repère

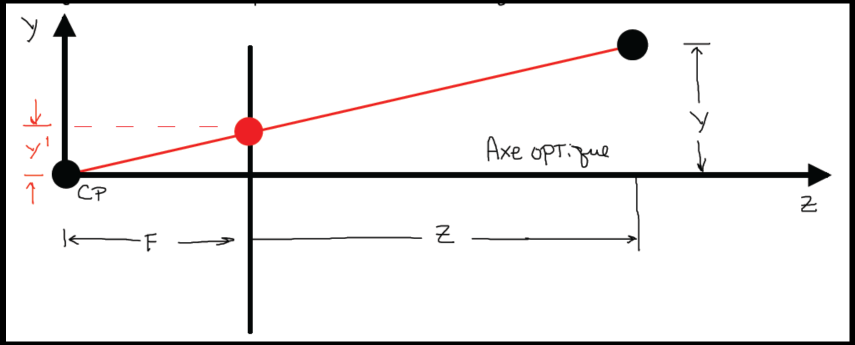
Généralisation : projection de perspective à plusieurs référentiels

Solution: généraliser le modèle en considérant:

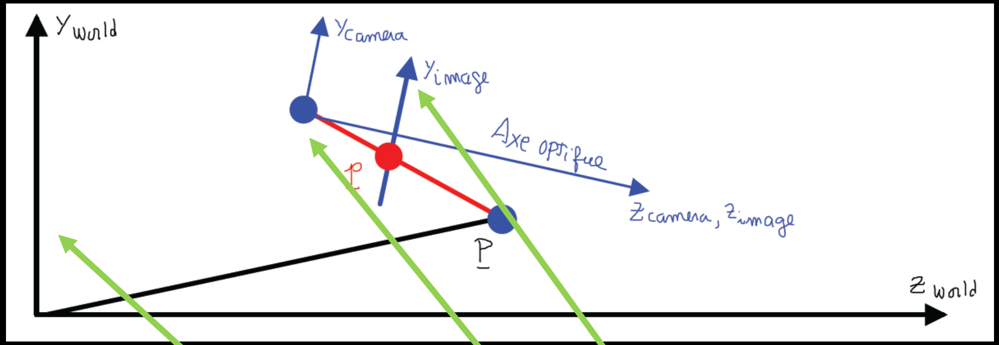
- un référentiel caméra pour exprimer les coordonnées des points images
- un référentiel monde (“world”) pour exprimer les coordonnées des points objets

Illustration en 2D

1 référentiel



3 référentiels

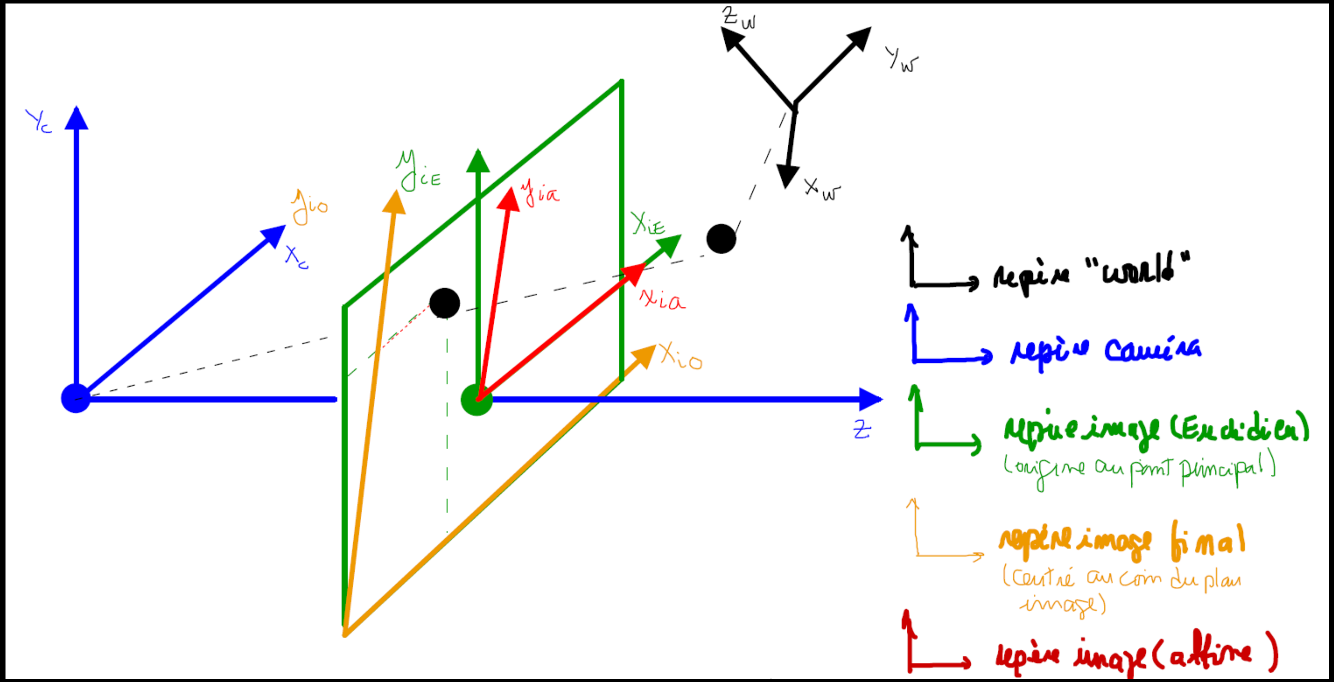


référentiel monde

référentiel caméra

référentiel image

Modèle complet



Repère world : repère dans lequel on exprime les points **"objet"**

Repère caméra : repère dans lequel on peut exprimer les points **"objet"** grâce à un changement de repère (rotation et translation) entre le repère "world" et le repère caméra. On peut aussi y exprimer les coordonnées des points image.

Repère "image" (euclidien) : repère situé dans le **plan image** et dans lequel sont exprimées les coordonnées des points image suite à la **projection de perspective**. Les axes de ce repère sont **orthogonaux**. **L'origine** est située au **point principal** (intersection entre l'axe optique et le plan image)

Repère "image" : même que le repère non-orthogonal mais avec l'origine au **coin** du plan image

Repère "image" (affine) : repère situé dans le plan image et dans lequel sont exprimées les coordonnées des points image suite à la **projection de perspective**. Les axes de ce repère ne sont **pas orthogonaux** parce qu'ils servent à modéliser un phénomène qui se rencontre dans les **caméras réelles** (surtout les caméras analogiques)

Equation de projection de perspective généralisée

Point image

$$\underline{\tilde{p}} = \begin{bmatrix} su \\ sv \\ sw \\ s \end{bmatrix}$$

Point objet

$$\underline{\tilde{P}} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Projection de perspective

$$\boxed{\underline{\tilde{p}} = \underline{\underline{\tilde{T}}}\,\underline{\underline{\tilde{S}}}\,\underline{\underline{\tilde{C}}}\,\underline{\underline{\tilde{P}}}_P\,\underline{\underline{\tilde{E}}}_{CW}\,\underline{\tilde{P}}}$$

$$\underline{\underline{\tilde{E}}}_{CW} = \begin{bmatrix} \underline{\underline{R}}^t_{3x3} & -\underline{\underline{R}}^t\underline{t} \\ \underline{0}^t_{3x1} & 1 \end{bmatrix}$$

Conversion monde -> caméras

$$\underline{\underline{\tilde{P}}}_P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/F & 0 \end{bmatrix}$$

Projection de perspective

$$\underline{\underline{\tilde{C}}} = \begin{bmatrix} 1 & -\cot\theta & 0 & 0 \\ 0 & 1/\sin\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

oblique entre les axes du capteur

$$\underline{\underline{\tilde{S}}} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Conversion mm -> pixel

$$\underline{\underline{\tilde{T}}} = \begin{bmatrix} 1 & 0 & 0 & u_0 \\ 0 & 1 & 0 & v_0 \\ 0 & 0 & 1 & -F \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Translation coin inférieur

La matrice $\underline{\underline{\tilde{E}}}_{CW}$ contient les **paramètres extrinsèques** de la caméra

Les paramètres extrinsèques décrivent la position et l'orientation de la caméra dans le repère monde (i.e. le monde "externe" à la caméra)

La matrice $\underline{\underline{\tilde{T}}}\,\underline{\underline{\tilde{S}}}\,\underline{\underline{\tilde{C}}}\,\underline{\underline{\tilde{P}}}_P = \begin{bmatrix} s_x F & -s_x F \cot\theta & u_0 & 0 \\ 0 & \dfrac{s_y F}{\sin\theta} & v_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ contient les **paramètres intrinsèques** de la caméra

Les paramètres intrinsèques renferment des propriétés **internes** de la caméra quelle que soit sa position et son orientation dans le repère monde

En posant $\alpha = s_x F$, $\beta = \dfrac{s_y F}{\sin\theta}$, $\gamma = -s_x F \cot\theta$ on peut écrire

$$\underline{\underline{\tilde{T}}}\,\underline{\underline{\tilde{S}}}\,\underline{\underline{\tilde{C}}}\,\underline{\underline{\tilde{P}}}_P = \begin{bmatrix} \alpha & \gamma & u_0 & 0 \\ 0 & \beta & v_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

La projection de perspective devient

$$\underline{\tilde{p}} = \begin{bmatrix} su \\ sv \\ sw \\ s \end{bmatrix} = \begin{bmatrix} \alpha & \gamma & u_0 & \cancel{0} \\ 0 & \beta & v_0 & \cancel{0} \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} \\ 0 & 0 & 1 & \cancel{0} \end{bmatrix} \begin{bmatrix} \underline{\underline{R}}^t & -\underline{\underline{R}}^t t \\ \underline{0}^t & 1 \end{bmatrix} \underline{\tilde{P}}_w$$

$$\boxed{\begin{gathered} \underline{\tilde{p}} = s\underline{\tilde{m}} = \underline{\underline{K}}_{3x3} \left[\underline{\underline{R}}^t \quad -\underline{\underline{R}}^t t \right]_{3x4} \underline{\tilde{P}}_w \\ \text{avec} \quad \underline{\tilde{m}} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad \underline{\underline{K}} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \end{gathered}}$$

Quelques remarques sur ce résultat

1- Si le point objet est le centre de projection,

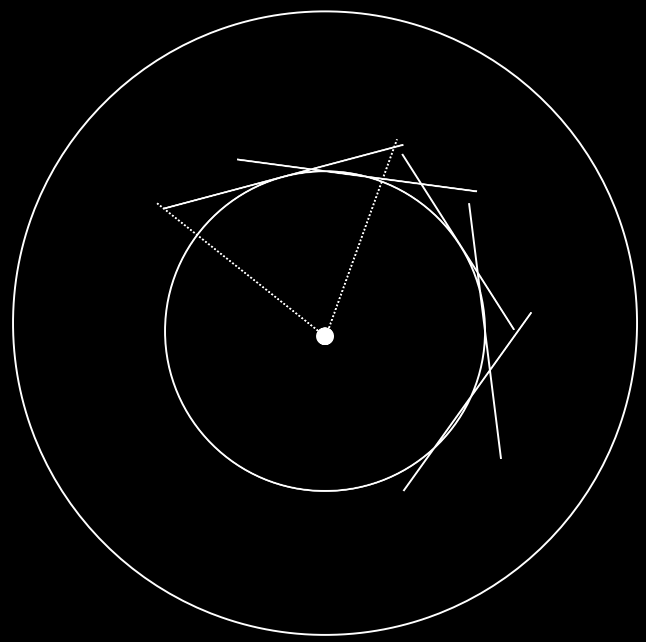
$$\underline{\tilde{P}}_w = \underline{\tilde{c}} = \begin{bmatrix} t \\ 1 \end{bmatrix}_{4x1}$$

l'image est $(0\ 0\ 0)^t$

2- Avec cette formulation de la projection de perspective, il faut **calibrer** expérimentalement les paramètres **intrinsèques** et **extrinsèques**

Autres projections

- sphérique: la rétine!
- cylindrique: des images panoramiques

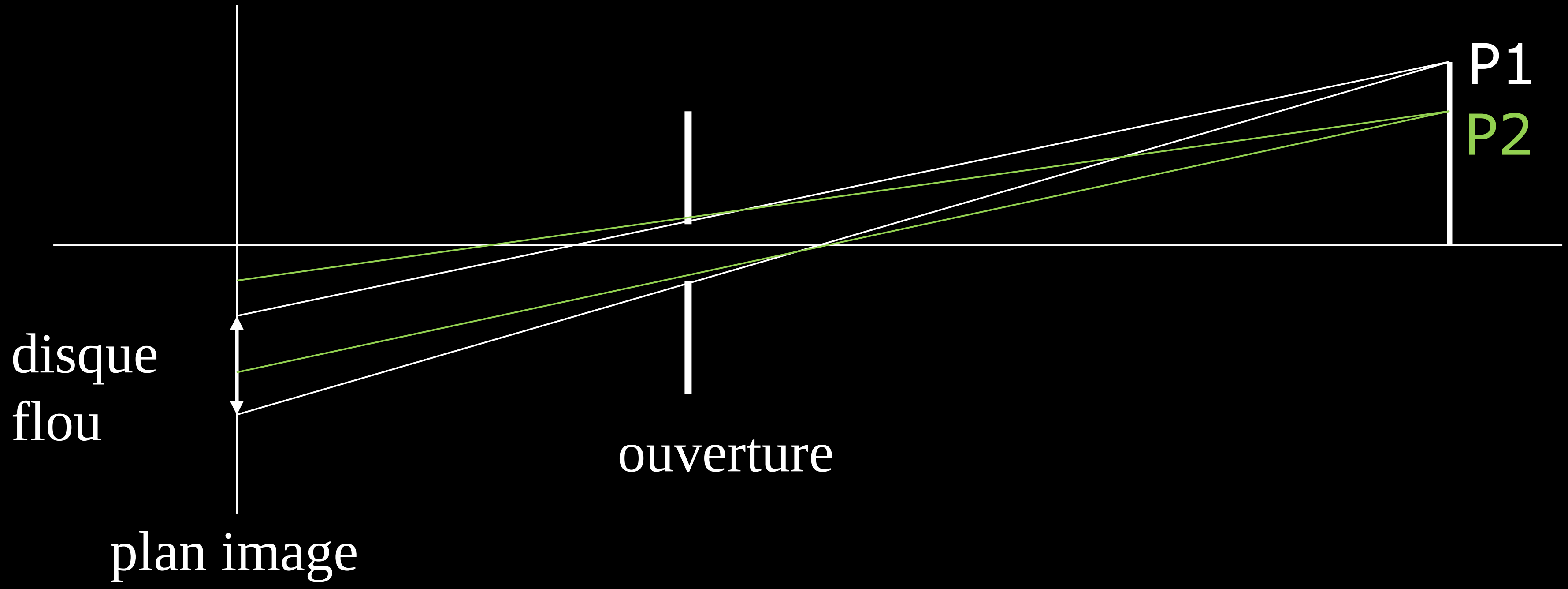


Ex: Quiktime VR

La caméra avec lentille

- Inconvénients du sténopé :
  1. Petite ouverture -> peu de lumière
  2. Effet de diffraction, i.e. courbure des rayons à cause des rebords d'objets opaques.
     - Phénomène expliqué par l’optique ondulatoire
     - La diffraction crée un flou. L'effet augmente si le diamètre de l'ouverture diminue.
  3. Si on augmente la taille de l'ouverture, la profondeur de champ diminue.

Flou créé par l'augmentation de l'ouverture



Exemple d’une image floue

http://www.yorku.ca/eye/lensfn6.htm

Solution: utiliser une lentille

- avantage: modèle équivalent au sténopé
- inconvénient: seuls des points à une distance donnée de la lentille sont au focus -> profondeur de champ limitée

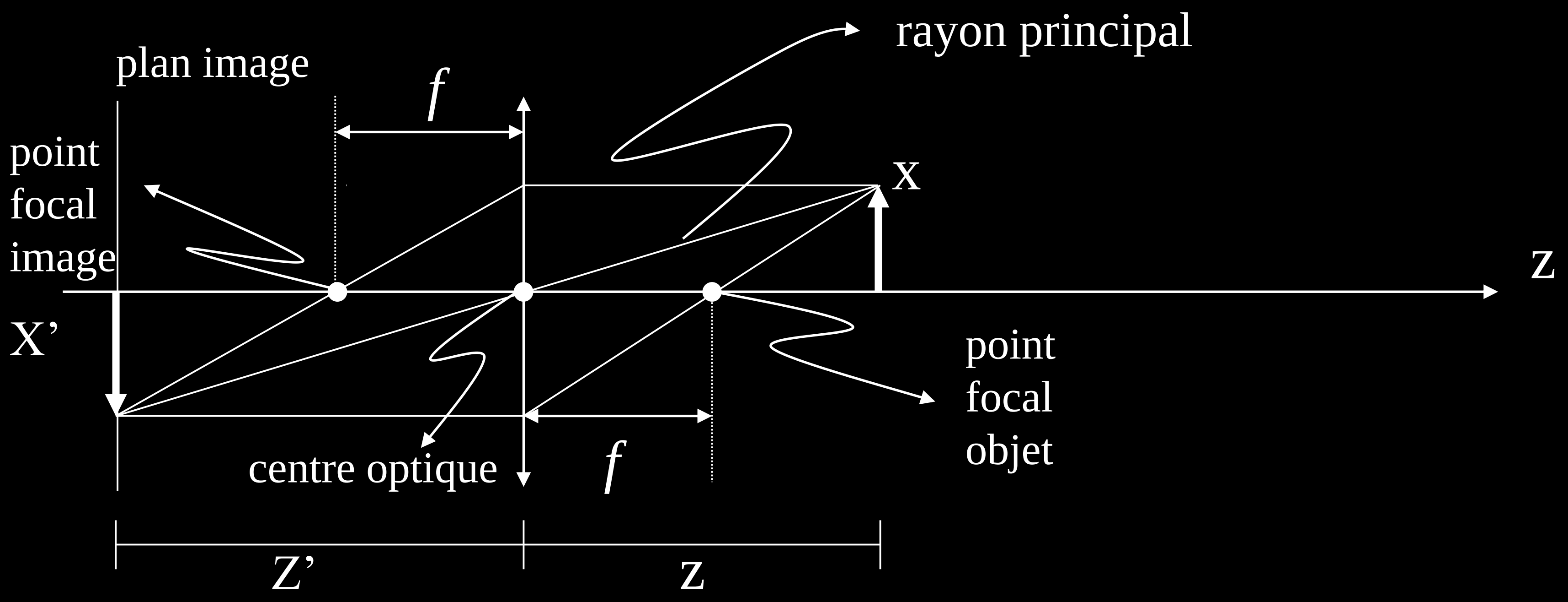
profondeur de champ: distance en z telle que le disque flou a un diamètre inférieur à 1 pixel (dépend de la taille de l'ouverture)

Modèle de lentille mince convergente

- On peut en faire la démonstration de son comportement optique par l'application de la loi de la réfraction (Snell-Descartes). Un rayon entrant réfracté sur la frontière droite de la lentille est immédiatement réfracté sur la frontière gauche

- Corollaires:
    - un rayon parallèle à l'axe optique passe par le point focal
    - le rayon principal (passe par le centre optique) n'est pas dévié

Schéma et équations



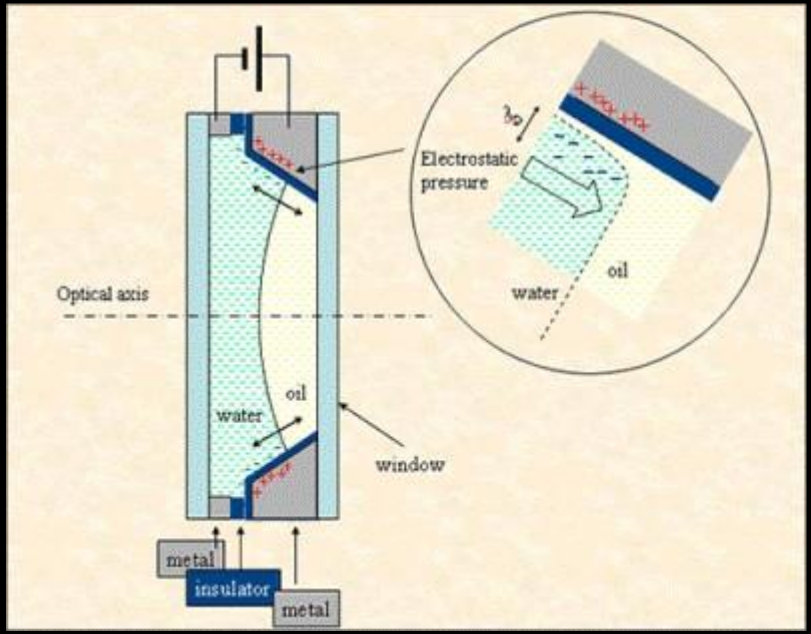
Magnification: $X'/X = f/(z-f) = (z'-f)/f$

Attention: notation différente de Sonka p. 82

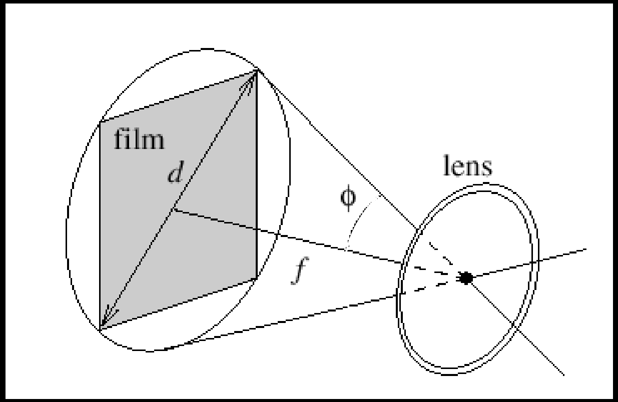
Ajustement mécanique d’une caméra

- Focus : Z’
- L’iris : l’ouverture affecte la prof. de champ, la quantité de lumière et la distorsion puisqu’elle est plus faible au centre de la lentille
- Zoom : Z’ et F

Lentille liquide pour téléphone
Cellulaire (Philips)



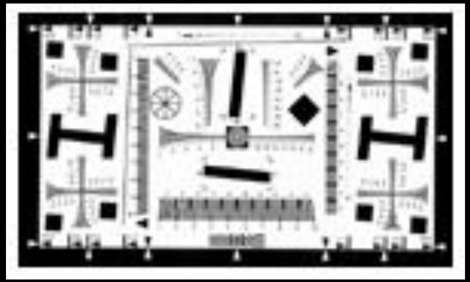
Terminologie



**tirée de Forsyth

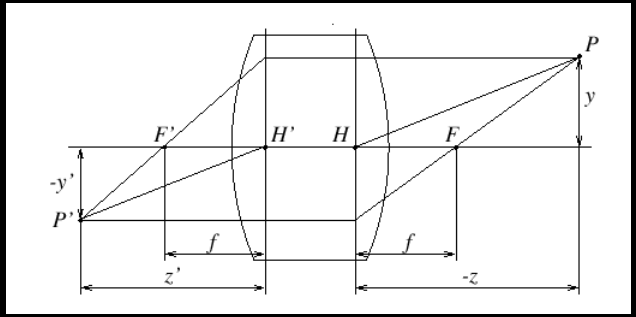
- *champ de vue*
- lentilles
  - Téléphoto, grand angle, télécentrique
- *f-number* d'une lentille: rapport f/diamètre
  *permet de comparer les lentilles de focales différentes
  ex: 2.8 (ouverture max), 4, 5.6, 8, 11 (les f-stops en photographie varient d’un rapport √2)
- *résolution*: distance minimum entre deux caractéristiques (objets) pour pouvoir les distinguer (1 pixel entre les deux),
  donc 2* taille d'un pixel
- *plage dynamique*: écart entre le plus petit et le plus grand niveau d'illuminance (exemple: 256 niveaux pour une image 8 bits)

Les High Dynamic Range images : cours de photographie algorithmique de J-F Lalonde



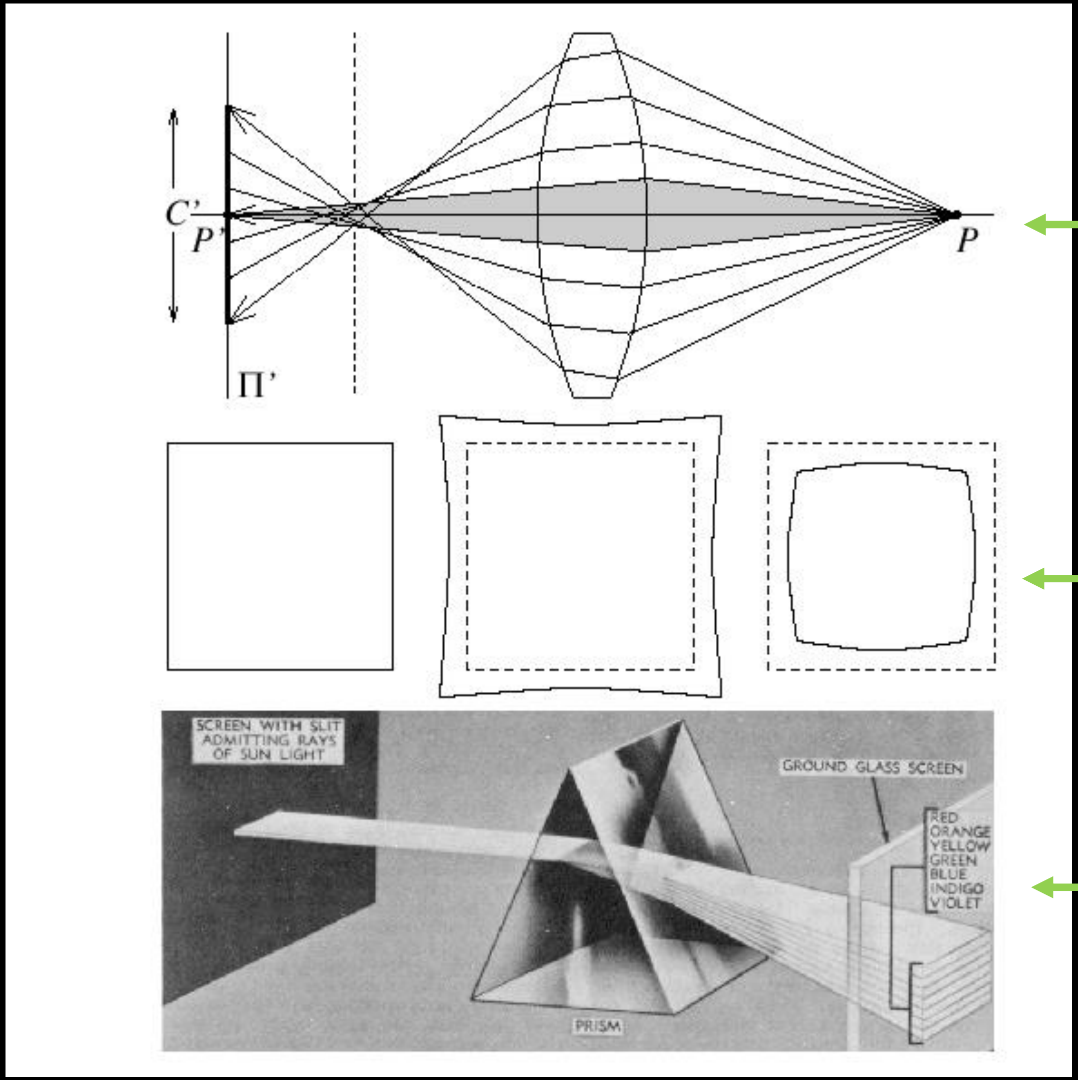
Lentilles réelles : terminologie

- *lentilles épaisses : elles présentent des imperfections par rapport aux lentilles minces ou au sténopé:*
  - *imperfections « optiques »*
    - *aberrations*: caractéristiques d'une lentille qui l'empêche de former l'image d'un point objet en un seul point du plan image. Un point a plutôt comme image une petite région floue.
    - *aberrations sphériques*: s'appliquent aux points qui devraient être imagés sur l'axe optique (netteté -, dépend de l'ouverture).
    - le "*coma*" est le type d'aberration pour les points hors axe
    - *aberration chromatique*: liée à la dépendance de l'indice de réfraction à la longueur d'onde (moins fort pour le rouge que le bleu)
  - imperfections « géométriques »
    - *distorsion radiale*: souvent importante pour les grands angles
    - indépendant de l'ouverture, n'affecte pas la netteté



*tirée de Forsyth

Lentilles réelles : illustration de quelques imperfections



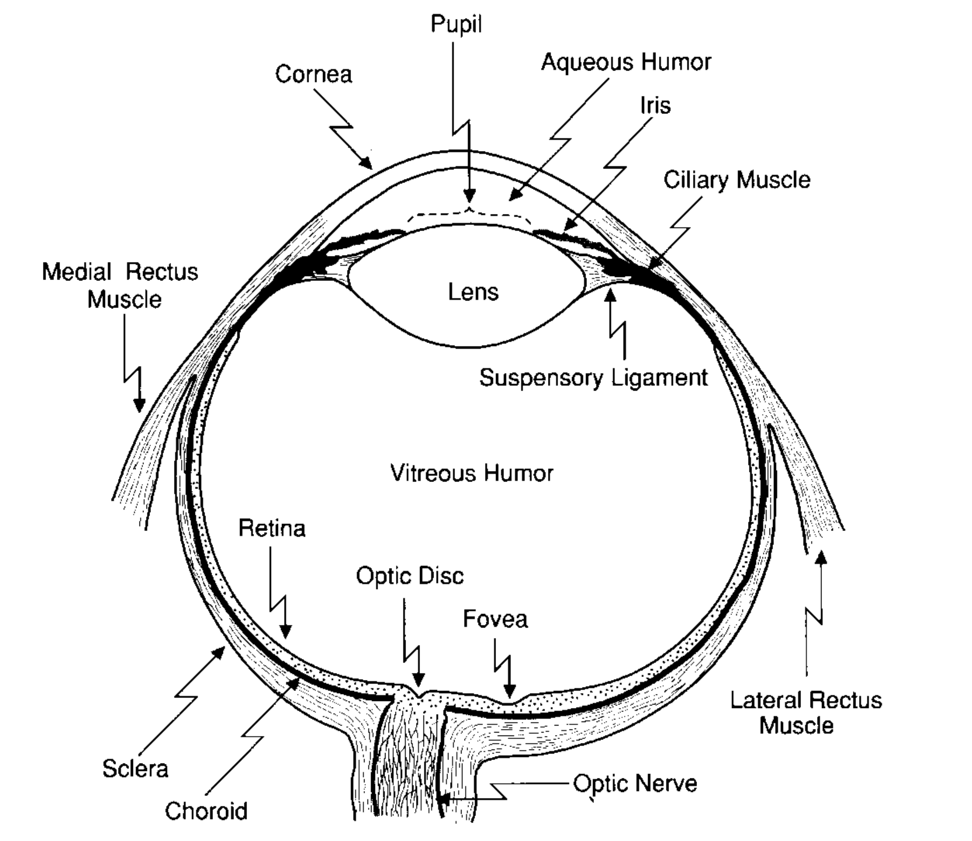
aberrations sphériques

distorsion radiale:
négative (pincushion)
positive (barrel-type)

aberration chromatique :
la cause, la diffraction

*tirée de Forsyth

L'œil humain (bref aperçu)



*tirée de Nalwa

*supporte jusqu'à 9 ordres de grandeur vs 2 à 4 pour un CCD pour la sensibilité à l'illuminance incidente

. champ de vue: 160(L) x 135(H) degrés
. résolution: 1/120 degrés
(30s d'arc car 1 degré = 60 min)
. f est environ 16 mm
. photorécepteurs
a) cônes: (couleur: 330-730 nm)
violet au rouge
3 types
environ 5 millions (*à 10 millions*)
b) bâtonnets:
environ 100 millions
utiles pour la faible lumière*
moins utiles pour la résolution spatiale malgré leur grand nombre
c) la fovée (haute densité de cônes)

Les systèmes d'acquisition en pratique (aperçu)

- appareil photo numérique
- appareil photo + scanner
- caméra + connexion USB ou 1394
- caméra + cartes d'acquisition (camera-link)
- Plusieurs liens commerciaux:
  http://www.1394imaging.com/en/resources/whitepapers/

Technologies de capteurs photosensibles

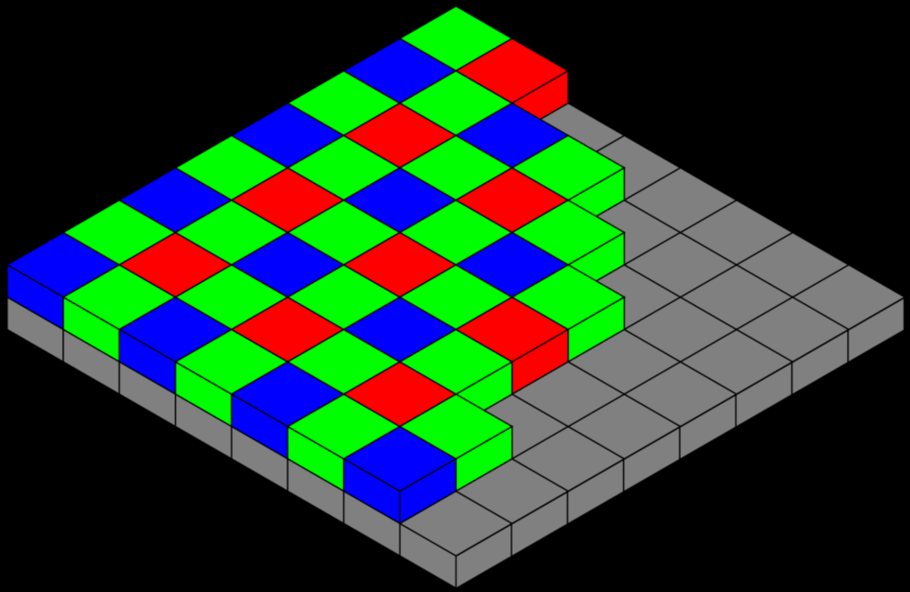
- Tube vidicon
- CCD
- CMOS
  [CVonline: Sensors and their Properties]

- le standard NTSC → le balayage interlacé
- standards modernes → le balayage progressif
- les sources de bruit (dark current, blooming, jitter, …)

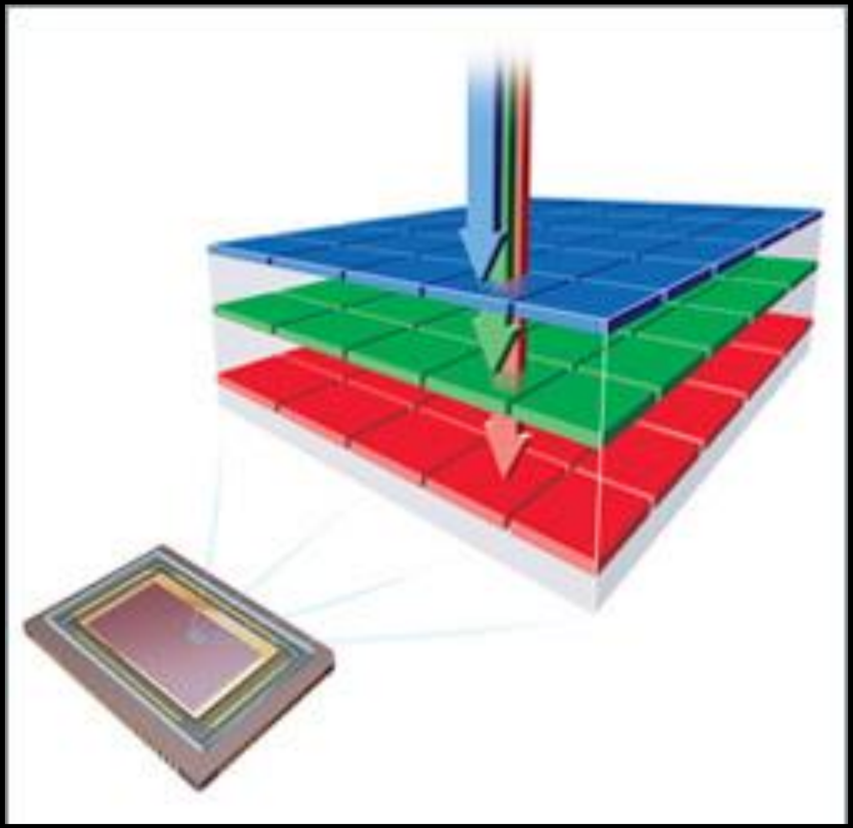
Capter la couleur

- 3 CCD ou filtre Bayer: lien sur filtres Bayer
- Le capteur Foveon:
- http://www.dpreview.com/news/0202/02021102foveonx3tech.asp

Bayer



Foveon



Sommaire des points essentiels

- Le sténopé: terminologie, un modèle, les équations (projecteur et prédiction d’un point image), les limites du modèle
- La caméra réelle: lentilles minces, les lentilles épaisses, aberrations et distorsion
- Connaissances générales sur les caméras et l’acquisition d’images (terminologie)