



Les capteurs visuels : Modèle de caméra classique et Calibration

Diane Lingrand

`lingrand@essi.fr`

Cours destiné aux élèves de dernière année de cycle ingénieur
version provisoire 2002
mises à jour [http ://www.essi.fr/~lingrand/](http://www.essi.fr/~lingrand/)

Table des matières

1	Introduction	2
1.1	Capteur visuel ?	2
1.2	Les premiers appareils photographiques : les caméras sténopées.	2
1.3	Les caméras à CCD	5
2	A l'intérieur de la caméra	6
3	Modélisation d'une caméra	9
3.1	La caméra sténopée.	9
3.2	La projection perspective	11
3.2.1	La projection para-perspective	12
3.2.2	La projection orthographique	15
3.2.3	Plusieurs modèles en une équation	16
4	Raffinement de cette modélisation.	18
4.1	Les distorsions.	18
4.1.1	Distorsion radiale.	18
4.1.2	D'où proviennent les distorsions ?	18
4.1.3	Exemple de correction de distorsion	19
4.2	Modèle de lentille épaisse	20
5	Equation de projection dans un repère global : paramétrisation du mouvement	23
5.1	Mouvement rigide	23
5.2	Angles d'Euler	23
5.3	Exponentielle de vecteur	24
5.4	Représentation par quaternions	25
5.5	Analyse des diverses représentations	26
6	Notions de Calibration	27
6.1	Introduction	27
6.2	Comment calibrer ?	27
6.2.1	Déterminer la projection	27
6.2.2	Déterminer les paramètres	28
6.3	Conclusion	29
6.3.1	Définitions - Principes	30
6.3.2	Application à la vision par ordinateur	31
7	Bibliographie	32

Chapitre 1

Introduction

1.1 Capteur visuel ?

Un capteur visuel est un système qui effectue une projection d'un espace tridimensionnel vers un espace mono-, bi-, ou tri-dimensionnel, et mesure l'amplitude d'énergie lumineuse reçue. Il discrétise l'espace en éléments d'image (pixel dans le cas bidimensionnel, voxel dans le cas tridimensionnel). Les capteurs que nous allons considérer par la suite sont soit des capteurs bidimensionnels, soit des ensembles de capteurs monodimensionnels afin d'obtenir dans tous les cas des images bidimensionnelles. On peut classer les capteurs en plusieurs catégories : les capteurs photo-chimiques (systèmes biologiques, films photographiques, ...), les capteurs thermiques (thermopile, ...) et les capteurs photoélectriques (photo-diode, détecteurs à transfert de charge ou DTC, ...).

1.2 Les premiers appareils photographiques : les caméras sténopées.

Même si, au Vème siècle avant Jésus-Christ, des chinois découvrent la propagation linéaire de la lumière et observent une formation d'image sténopée, la première caméra fut construite à la Renaissance par des astronomes, ceci afin d'observer l'éclipse de 1544.

Cette caméra très simple (elle consiste en une boîte munie d'un trou et d'un film photographique) est encore utilisée actuellement à des fins artistiques et aussi à titre de collection. Vous trouverez un exemple de caméra de ce type en figure 1.2 ainsi que quelques photos artistiques résultantes en figure 1.3 suivantes.

Diverses améliorations eurent lieu, la première consista en l'ajout d'une lentille afin de réduire les effets de flou et les aberrations chromatiques. Cependant, l'ajout d'une lentille induit d'autres défauts que l'on peut atténuer en ajoutant divers dioptries. C'est ainsi qu'on arrive à des modèles de caméra de plus en plus compliqués. Les principaux défauts observés sont :

aberration sphérique : concerne les objets proches de l'axe optique. L'image d'un point situé sur l'axe optique est une tache sphérique due à la diffraction de la lumière. Sachant que les aberrations sphériques d'une lentille convergente et d'une lentille divergente se compensent, on pourra atténuer ce type d'aberration en ajoutant à une lentille convergente une lentille divergente ou bien en utilisant plutôt une lentille plan-convexe à la place d'une lentille bi-convexe.

aberration de coma (ou effet de comète) : concerne les objets paraxiaux. Comme pour l'aberration sphérique, l'image d'un point est une tache.

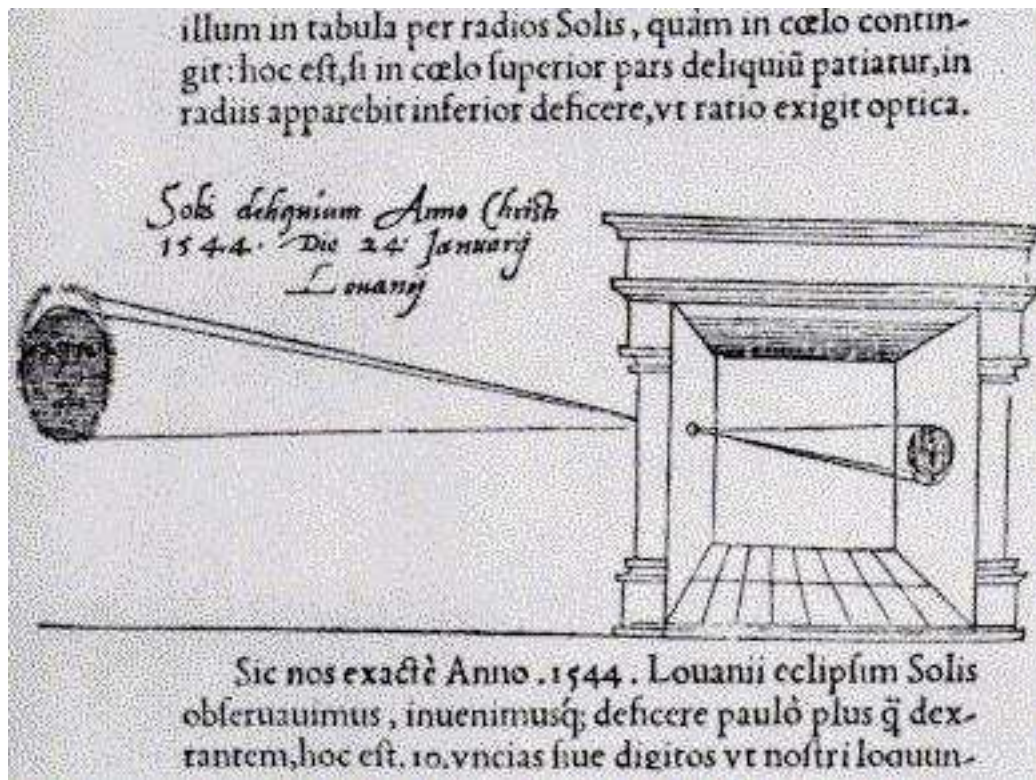


FIG. 1.1 – Schéma de caméra sténopée (dans *De Radio Astronomica et Geometrica*, par Gemma Frisius, 1545), utilisée pour l'observation de l'éclipse de soleil de 1544.

astigmatisme : la surface des lentille ne peut être considérée comme étant exactement sphérique; il en résulte une mauvaise convergence (objets légèrement flous).

distorsions : l'image d'une droite est une courbe (voir le paragraphe 4.1 pour plus de détails).

qui correspondent, d'une part, aux différents termes de la série du sinus de la loi de Snell-Descartes (voir 2.1) qui ont été négligés, d'autre part à la géométrie du dioptré de la lentille.



FIG. 1.2 – Un exemple de caméra sténopée fabriquée à partir d'une boîte de conserve.



FIG. 1.3 – Quelques photos obtenus à l'aide de caméras sténopées

1.3 Les caméras à CCD

Les caméras courantes du marché sont des caméras à base de barrettes à transfert de charge dite CCD (*Charged Coupled Device*). Ces barrettes sont des alignements de cellules photo-conductrices au silicium correspondant chacune à un élément d'image appelé "pixel" (*picture element*). En présence de lumière, l'impact des photons incidents fait apparaître localement des charges électriques qui sont piégées dans un puits de potentiel où elles s'accumulent pendant le temps qui sépare deux scrutations. A des instants donnés, les charges stockées (dépendant du nombre d'impacts de photons reçus) sont transférées simultanément sur les entrées parallèles d'un registre analogique à la sortie duquel l'image de la ligne analysée sera disponible. Après numérisation, chaque pixel possède une valeur de niveau de gris entre 0 et 255.

Le cas des caméras couleurs est un peu plus complexe. On distinguera quatre types de caméras couleurs :

- **filtres optiques colorés** : Kodak et Wratten utilisèrent des filtres colorés sur leurs caméras à niveau de gris. Ces manipulations étaient lourdes et le résultat nécessite des corrections à cause de la densité variable des filtres.
- **caméras couleur mono-CCD** : consistant en une superposition d'un masque de filtres sur la matrice CCD. De tels capteurs n'ont pas beaucoup de dynamique.
- **couleur tri-CCD** : il s'agit de séparer la lumière en trois faisceaux de longueurs d'onde différentes (rouge, vert, bleu) en utilisant un prisme puis d'envoyer chacun des faisceaux primaires sur un détecteur CCD sensible aux fréquences correspondantes. Une image est ainsi séparée en 3 images : une rougeâtre, une verdâtre et une bleuâtre. Plus cher, ce capteur demande une grande précision pour faire coïncider les 3 images (voir pour exemple la caméra Canon XL1, <http://www.canondv.com/xl1/prodinfo/ccd.html>)
- **couleur quadri-CCD** : à chaque pixel sont associées quatre cellules CCD différemment sensibles aux longueurs d'onde : une dans les fréquences proches du rouge, une pour le vert, une pour le bleu et une dernière afin d'améliorer la précision dans les fréquences proches du vert (le vert détient environ 60% de l'information dans une image).

Tous ces capteurs nécessitent une **gamma-correction** car l'intensité mesurée par les caméras est une fonction concave logarithmique de l'intensité réelle. Les couleurs apparaissent moins saturées et les coordonnées chromatiques sont inexactes. La fonction correctrice inverse est appelée gamma-correction.

Chapitre 2

A l'intérieur de la caméra

Les éléments optiques présents dans une caméra sont des lentilles, des miroirs, des prismes, des filtres, des polarisateurs et des mécanismes d'ouverture. Les dioptries sont régis par la loi de Snell-Descartes (voir figure 2.1) :

$$n_1 \sin(i_1) = n_2 \sin(i_2) \quad (2.1)$$

qui, dans les conditions d'approximation de Gauss (le système est centré, les centres des dioptries sont alignés, les milieux sont homogènes et les rayons paraxiaux), est approximée par la loi de Képler :

$$n_1 i_1 = n_2 i_2$$

Cette loi permet de déterminer le chemin de la lumière à travers les différents dioptries qui constituent une caméra.

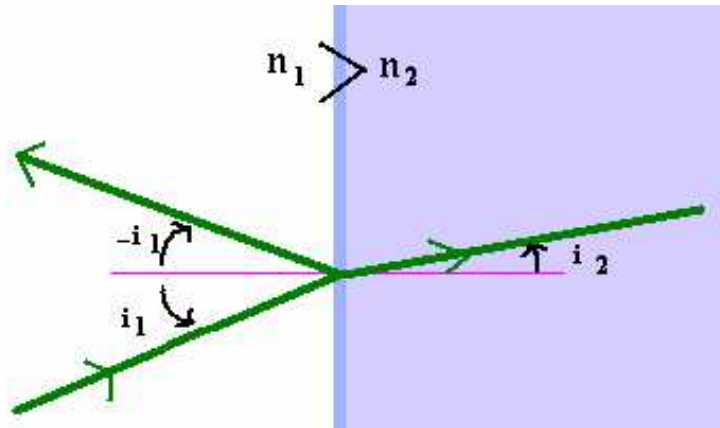


FIG. 2.1 – Réflexion et réfraction entre deux milieux, loi de Snell-Descartes.

Connaissant les équations générales de chaque composant optique de la caméra, on peut déterminer les équations de la caméra. Bien qu'on obtienne alors un modèle en fonction des éléments présents dans la caméra, on sait déjà que ce sont, d'une part, des équations approchées et d'autre part, qu'elles nécessitent une connaissance très précise des positions relatives de ces divers éléments.

Mais bien plus important, on obtient alors une modélisation de la caméra qui comprend un très grand nombre de paramètres et qui est une expression trop complexe pour être utilisée

numériquement. Un tel modèle est donc pratiquement inutilisable. D'autre part, il faut également noter qu'il est rare d'avoir accès au schéma interne d'une caméra (confidentialité commerciale, ...) ainsi qu'aux mécanismes mis en jeu lors de variations de zoom par exemple (il y a des mouvements complexes des dioptries les uns par rapport aux autres) ou de mise au point.

À titre d'exemple, le schéma interne de l'objectif SUMMILUX-R f/1.4/50mm de chez Leica (voir

http://www.leica-camera.com/r_system/pdf/50summ_e.pdf

) est reproduite sur la figure 2.2.

Ces multiples raisons nous conduisent à rechercher un modèle simplifié et utilisable d'une caméra qui représente son fonctionnement *vu de l'extérieur*. Cela est l'objet de la première partie de ce cours. Nous verrons ensuite qu'un tel modèle dépend de paramètres inconnus à priori et nous nous intéresserons à leur détermination : la calibration.

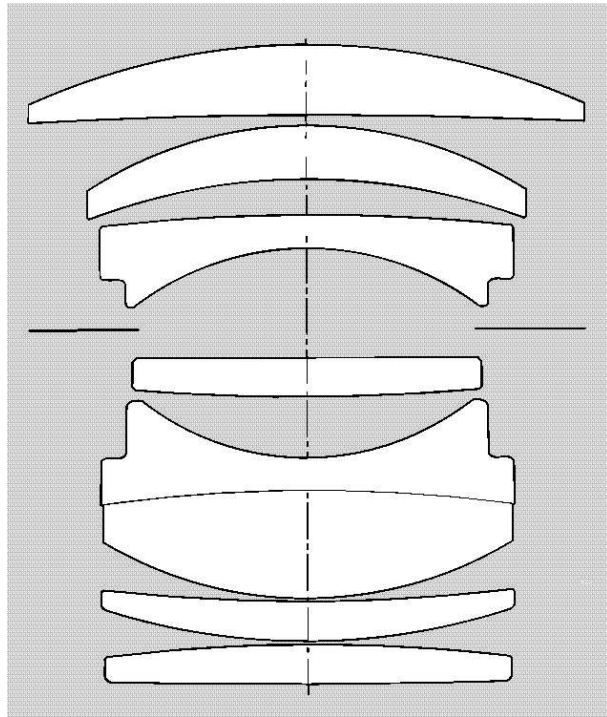
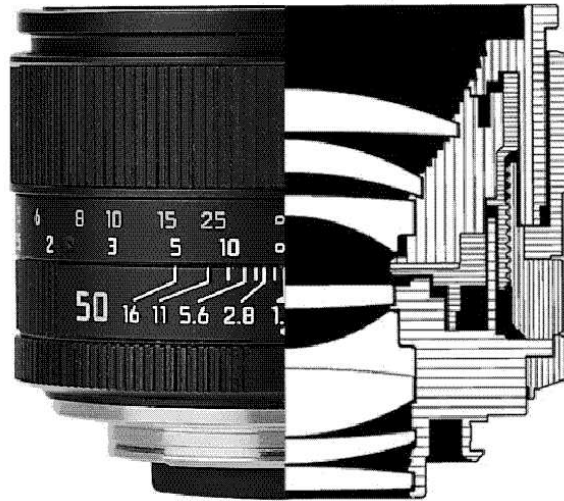


FIG. 2.2 – Le schéma interne de l'objectif "SUMMILUX-R f/1.4/50mm" de chez Leica

Chapitre 3

Modélisation d'une caméra

Nous nous intéressons maintenant au modèle le plus simple qui rende compte du comportement d'une caméra. En particulier, on s'attachera à trouver un modèle linéaire, ce qui nous permettra de le manipuler facilement à l'aide des outils d'algèbre linéaire et de dériver des algorithmes facilement implémentables. Une caméra étant une transformation d'un monde tridimensionnel vers un plan, on pense naturellement à une projection. Il reste alors à déterminer quel type de projection et avec quels paramètres.

Notations Une caméra effectue une projection perspective de centre C sur un plan rétinien Π . Afin d'exprimer cette projection dans les coordonnées de l'espace et de l'image, il est nécessaire de connaître le changement de repère entre la scène observée et le repère attaché à la caméra. Un point \mathbf{M} de l'espace a pour coordonnées (X, Y, Z) dans le repère \mathcal{R} de l'espace. Il se projette en un point \mathbf{m} de l'image a pour coordonnées (u, v) dans le repère image \mathcal{R}_i , déterminé par les axes horizontal et vertical de l'image ainsi que du coin supérieur gauche et (u', v') dans le repère ρ centré sur l'image. Ce repère image induit un repère espace \mathcal{R}' dont le troisième axe est donné par l'axe optique et le centre par le centre optique \mathbf{C} . Dans le repère image \mathcal{R}_i , le point C a pour coordonnées $(u_0, v_0, -f)$, où f est la distance focale (distance du plan rétinien au centre optique). Le point M de l'espace a pour coordonnées (X', Y', Z') dans le repère \mathcal{R}' . La transformation entre les repères \mathcal{R} et \mathcal{R}' est la composée d'une rotation \mathbf{R} et d'une translation \mathbf{t} .

Ces notations sont illustrées par la figure 3.1.

3.1 La caméra sténopée.

Le principe de la caméra sténopée est présenté sur la figure 3.2

Les équations du modèle sténopé, dans le repère centré sur l'image, s'écrivent :

$$-\frac{f}{Z} = \frac{u'}{X} = \frac{v'}{Y}$$

ce qui donne, dans le repère image \mathcal{R}_i :

$$-\frac{f}{Z} = \frac{u - u_0}{X} = \frac{v - v_0}{Y}$$

et qui s'écrit aussi sous la forme :

$$Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & u_0 \\ 0 & -f & 0 & v_0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

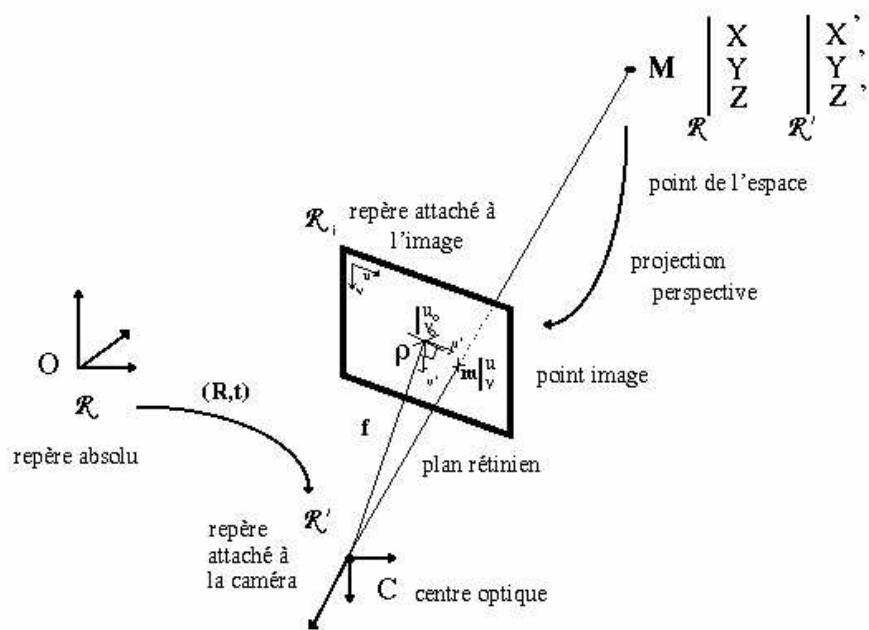


FIG. 3.1 – Le modèle sténopé

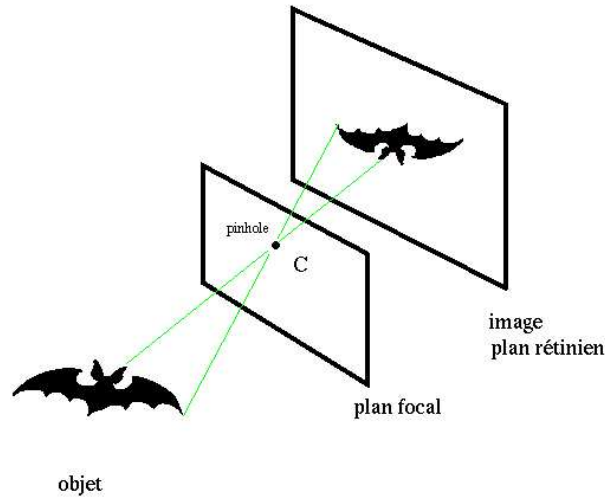


FIG. 3.2 – Formation d’une image avec une caméra sténopé

Pour obtenir un modèle expérimentalement valide, il va falloir ajouter des termes correctifs à ce modèle pour aboutir au modèle perspectif.

3.2 La projection perspective

La projection perspective (voir figure 3.3), qui est la projection la plus généralement validée, est représentée par l’équation suivante, dans le repère attaché à la caméra :

$$Z \mathbf{m} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{M} \quad (3.1)$$

Dans cette projection,

α_u et α_v représentent les tailles horizontale et verticale des pixels. On décompose parfois $\alpha_u = f k_u$ et $\alpha_v = f k_v$ afin d’exhiber que ces variables contiennent la focale. Lorsque α_u et α_v diffèrent, les pixels ne sont pas carrés mais rectangulaires. Il a été expérimentalement prouvé que le rapport $\frac{\alpha_u}{\alpha_v}$ est constant et, pour les caméras usuelles, de valeur 0,7.

u_0 et v_0 représentent la projection du centre optique sur la rétine, le centre de la zone fovéale.

γ mesure le défaut d’orthogonalité du plan des cellules CCD par rapport à l’axe optique.

Si γ n’est pas nul, les pixels, au lieu d’être rectangulaires, sont de la forme de parallélogrammes. Sa valeur est généralement très faible.

Le modèle projectif est le plus employé mais d’autres variantes ont été utilisées par des auteurs. Il s’agit des :

- projection para-perspective
- projection ortho-perspective
- projection affine, appelée aussi orthographique, perspective faible (“*weak perspective*”) ou orthographique graduée (“*scaled orthographic*”) selon son paramétrage.

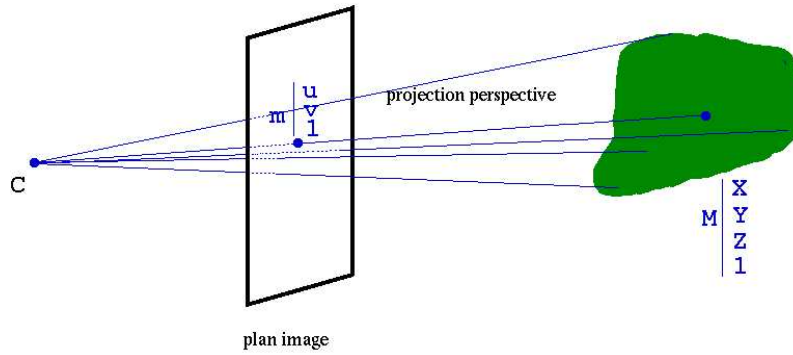


FIG. 3.3 – Principe de la projection projective

Ces types de projection diffèrent par le nombre de paramètres qu'ils comportent et par leur plus ou moins bonne adéquation aux données observés. Leur analyse est l'objet des paragraphes suivants.

Dans la suite, on se placera dans le repère \mathcal{R}' attaché à la caméra et on notera $\mathbf{M} = [X, Y, Z, \{0, 1\}]$ un point de l'espace dans ce repère.

3.2.1 La projection para-perspective

La projection para-perspective a été utilisée pour la première fois afin de résoudre un problème de reconstruction 3D à partir de textures. Elle fut ensuite réutilisée par divers auteurs. Le but de ce paragraphe est de montrer que cette approximation est à la fois une approximation géométrique et une approximation au premier ordre des équations de projection perspective.

Ce modèle est bien adapté à la focalisation d'attention autour d'un angle de vue particulier.

Approximation géométrique

Cette approximation peut être vue comme une approximation géométrique, c'est-à-dire en supposant que l'objet est loin de la caméra ou proche de l'axe optique. La projection est alors correctement approximée par la composée d'une projection parallèle sur un plan orthogonal à l'axe optique et d'une projection perspective. La direction de la projection parallèle dépend de la scène observée (ou de l'objet observé) et est donnée par la direction du rayon partant du centre optique de la caméra vers le centroïde de la scène, i.e. le point de profondeur moyenne. Le plan image fictif sur lequel sont projetés les points issus de la projection parallèle est parallèle au plan rétinien et passe par le centroïde. La figure 3.4 illustre cette situation qui consiste à regarder "en coin".

Notons $\mathbf{M}_0 = [X_0 Y_0 Z_0 1]^T$ le point de coordonnées moyennes (et donc de profondeur moyenne). Le plan image auxiliaire a pour équation $Z = Z_0$. Pour un point $\mathbf{M} = [X Y Z 1]^T$ de l'espace, notons $\mathbf{M}_a = [X_a Y_a Z_0 1]^T$ le point image auxiliaire correspondant.

Ecrivons la relation de parallélisme pour la projection parallèle :

$$\begin{cases} \mathbf{C M}_0 \parallel \mathbf{M}_a \mathbf{M} \\ \mathbf{M}_a \in \mathcal{P}(Z = Z_0) \end{cases} \Leftrightarrow \begin{cases} X_a = X - \frac{X_0}{Z_0} Z + X_0 \\ Y_a = Y - \frac{Y_0}{Z_0} Z + Y_0 \end{cases}$$

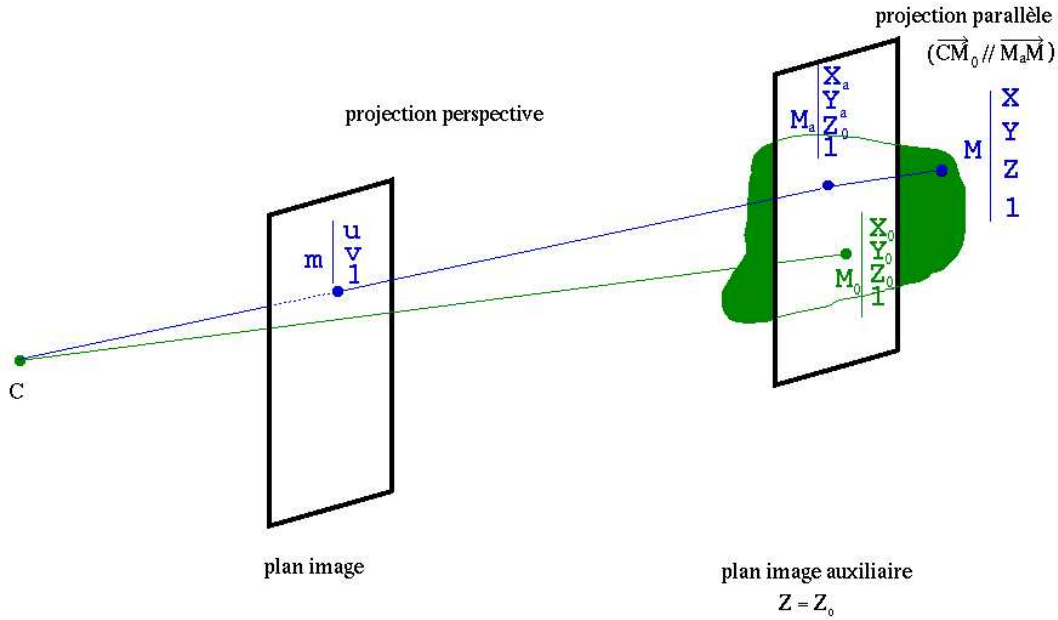


FIG. 3.4 – Principe de la projection para-perspective

Ecrivons maintenant les équations de projection perspective :

$$\begin{cases} u = \alpha_u \frac{X_a}{Z_0} + \gamma \frac{Y_a}{Z_0} + u_0 \\ v = \alpha_v \frac{Y_a}{Z_0} + v_0 \end{cases}$$

Soit, en remplaçant X_a et Y_a par leurs expressions :

$$\begin{cases} u = \frac{\alpha_u}{Z_0} X + \frac{\gamma}{Z_0} Y - \frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} Z + \alpha_u \frac{X_0}{Z_0} + \gamma \frac{Y_0}{Z_0} + u_0 \\ v = \frac{\alpha_v}{Z_0} Y - \frac{\alpha_v}{Z_0} \frac{Y_0}{Z_0} Z + \alpha_v \frac{Y_0}{Z_0} + v_0 \end{cases}$$

On remarque qu'on obtient alors un modèle linéaire en (X, Y, Z) (la profondeur Z n'apparaît plus au dénominateur) mais, en contre-partie, ce modèle comporte plus de paramètres que le modèle perspectif original, et ces paramètres dépendent de la scène observée. Sachant qu'il s'agit d'une approximation, on considérera que la variation des paramètres X_0 , Y_0 et Z_0 est faible par rapport aux variations des paramètres internes physiques de la caméra lorsque le mouvement est faible

En posant :

$$\begin{cases} \alpha_{u_{para}} = \frac{\alpha_u}{Z_0} \\ \alpha_{v_{para}} = \frac{\alpha_v}{Z_0} \\ \gamma_{para} = \frac{\gamma}{Z_0} \\ \beta_{u_{para}} = -\frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} \\ \beta_{v_{para}} = -\frac{1}{Z_0} \frac{\alpha_v Y_0}{Z_0} \\ u_{0_{para}} = u_0 + \alpha_u \frac{X_0}{Z_0} + \gamma \frac{Y_0}{Z_0} \\ v_{0_{para}} = v_0 + \alpha_v \frac{Y_0}{Z_0} \end{cases} \Leftrightarrow \begin{cases} \alpha_u = \alpha_{u_{para}} Z_0 \\ \alpha_v = \alpha_{v_{para}} Z_0 \\ \gamma = \gamma_{para} Z_0 \\ u_0 = u_{0_{para}} - \beta_{u_{para}} Z_0 \\ v_0 = v_{0_{para}} - \beta_{v_{para}} Z_0 \\ \frac{X_0}{Z_0} = \frac{\gamma_{para}}{\alpha_{u_{para}}} \frac{\beta_{v_{para}}}{\alpha_{v_{para}}} - \frac{\beta_{u_{para}}}{\alpha_{u_{para}}} \\ \frac{Y_0}{Z_0} = -\frac{\beta_{v_{para}}}{\alpha_{v_{para}}} \end{cases} \quad (3.2)$$

En sus des paramètres intrinsèques de la caméra interviennent $\frac{X_0}{Z_0}$ et $\frac{Y_0}{Z_0}$, paramètres représentant la direction $g = [X_0 \ Y_0 \ Z_0]^T$ du regard sur la scène, ainsi que Z_0 au titre de facteur d'échelle.

Les équations de la projection para-perspective s'écrivent alors sous la forme :

$$\begin{cases} u &= \alpha_{u_{para}} X + \gamma_{para} Y + \beta_{u_{para}} Z + u_{0_{para}} \\ v &= \alpha_{v_{para}} Y + \beta_{v_{para}} Z + v_{0_{para}} \end{cases}$$

On remarque que $(u_{0_{para}}, v_{0_{para}})$ représente le projeté dans le plan rétinien du point \mathbf{M}_0 de l'espace, selon une projection perspective et on note également les relations suivantes :

$$\begin{cases} \beta_{u_{para}} &= -\alpha_{u_{para}} \frac{X_0}{Z_0} - \gamma_{para} \frac{Y_0}{Z_0} \\ \beta_{v_{para}} &= -\alpha_{v_{para}} \frac{Y_0}{Z_0} \\ u_{0_{para}} &= u_0 - \beta_{u_{para}} \\ v_{0_{para}} &= v_0 - \beta_{v_{para}} \end{cases} \quad (3.3)$$

On obtient ainsi le modèle para-perspectif qui s'écrit matriciellement de la façon suivante :

$$\mathbf{m} = \begin{pmatrix} \alpha_{u_{para}} & \gamma_{para} & \beta_{u_{para}} & u_{0_{para}} \\ 0 & \alpha_{v_{para}} & \beta_{v_{para}} & v_{0_{para}} \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{M} \quad (3.4)$$

et qui comporte 2 paramètres supplémentaires par rapport au modèle perspectif, β_u et β_v . Cependant, si l'on se réfère aux relations 3.3, ces paramètres s'expriment en fonction des autres paramètres et de la direction du regard, $(\frac{X_0}{Z_0}, \frac{Y_0}{Z_0})$.

Approximation au premier ordre

D'un autre point de vue, on peut considérer que l'objet est loin de la scène et écrire la coordonnée de profondeur Z sous la forme :

$$Z = Z_0 (1 + \epsilon_Z + o(\epsilon_Z))$$

ainsi que les autres coordonnées :

$$\begin{aligned} X &= X_0 (1 + \epsilon_X + o(\epsilon_X)) \\ Y &= Y_0 (1 + \epsilon_Y + o(\epsilon_Y)) \end{aligned}$$

En écrivant les équations de projection projective, on obtient :

$$\begin{cases} u &= \alpha_u \frac{X_0}{Z_0} (1 + \epsilon_X + o(\epsilon_X)) \frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} + \gamma \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) \frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} + u_0 \\ v &= \alpha_v \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) \frac{1}{1 + \epsilon_Z + o(\epsilon_Z)} + v_0 \end{cases}$$

Or, en utilisant le développement limité suivant, au premier ordre au voisinage de 0 :

$$\frac{1}{1 + \epsilon} = 1 - \epsilon + o(\epsilon) \text{ où } \frac{o(\epsilon)}{\epsilon} \xrightarrow{\epsilon \rightarrow 0} 0$$

et en l'appliquant à $\epsilon_Z + o(\epsilon_Z)$, on fait disparaître les termes en Z du dénominateur, soit :

$$\begin{cases} u &= \alpha_u \frac{X_0}{Z_0} (1 + \epsilon_X + o(\epsilon_X)) (1 - \epsilon_Z + o(\epsilon_Z)) \\ &+ \gamma \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) (1 - \epsilon_Z + o(\epsilon_Z)) + u_0 \\ v &= \alpha_v \frac{Y_0}{Z_0} (1 + \epsilon_Y + o(\epsilon_Y)) (1 - \epsilon_Z + o(\epsilon_Z)) + v_0 \end{cases}$$

Au premier ordre, il vient :

$$\begin{cases} u &= \alpha_u \frac{X_0}{Z_0} (1 + \epsilon_X - \epsilon_Z) + \gamma \frac{Y_0}{Z_0} (1 + \epsilon_Y - \epsilon_Z) + u_0 \\ v &= \alpha_v \frac{Y_0}{Z_0} (1 + \epsilon_Y - \epsilon_Z) + v_0 \end{cases}$$

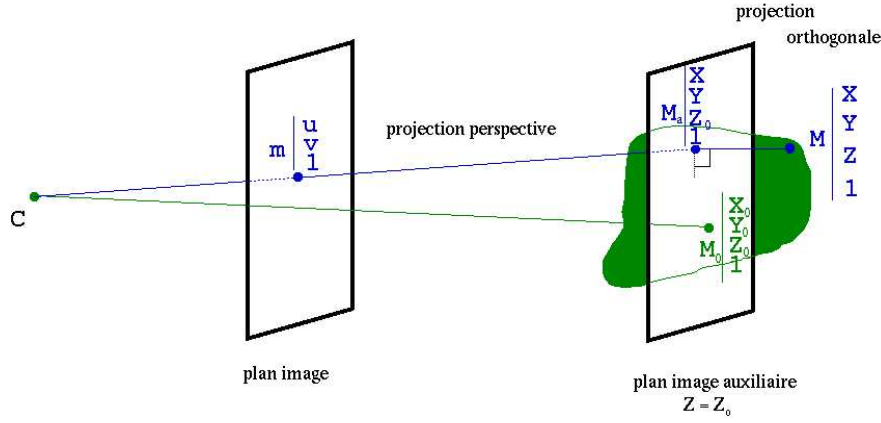


FIG. 3.5 – Principe de la projection orthographique

Or, toujours au premier ordre :

$$\begin{cases} \epsilon_X &= \frac{X}{X_0} - 1 \\ \epsilon_Y &= \frac{Y}{Y_0} - 1 \\ \epsilon_Z &= \frac{Z}{Z_0} - 1 \end{cases}$$

D'où les équations suivantes :

$$\begin{cases} u &= \frac{\alpha_u}{Z_0} X + \frac{\gamma}{Z_0} Y - \frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} Z + u_0 + \frac{\alpha_u X_0 + \gamma Y_0}{Z_0} \\ v &= \frac{\alpha_v}{Z_0} Y - \frac{1}{Z_0} \frac{\alpha_v Y_0}{Z_0} Z + v_0 + \frac{\alpha_v Y_0}{Z_0} \end{cases}$$

On retrouve ainsi les expressions linéaires en les coordonnées du point \mathbf{M} de l'espace mais dont les coefficients dépendent d'un point "milieu" de la scène \mathbf{M}_0 .

3.2.2 La projection orthographique

On rencontre également cette approximation sous les noms de perspective faible ("weak perspective") ou plus exactement orthographique graduée ("scaled orthographique"), la projection orthographique étant un cas particulier de cette dernière pour lequel on ne tient absolument pas compte de la profondeur. Nous utiliserons par la suite le terme de projection orthographique désignant l'ensemble des projections orthographiques graduées ou non.

Ce modèle est bien adapté à la vision fovéale.

Approximation géométrique

Il s'agit d'un cas particulier de l'approximation para-perspective dans le sens où la projection perspective est approximée par la composée d'une projection orthogonale sur un plan auxiliaire qui est perpendiculaire à l'axe optique de la caméra.

Plus concrètement, il s'agit d'une projection para-perspective pour laquelle la direction du regard et la même que celle de l'axe optique de la caméra.

La figure 3.5 illustre cette composée et précise les notations.

Les relations géométriques sont les suivantes :

$$\begin{cases} \mathbf{C} \mathbf{M}_0 \parallel \mathbf{M}_a \mathbf{M} \\ \mathbf{M}_a \mathbf{M}_0 \perp \mathbf{M}_a \mathbf{M} \\ \mathbf{M}_a \in \mathcal{P}(Z = Z_0) \end{cases} \Leftrightarrow \begin{cases} X_0 (X_a - X_0) + Y_0 (Y_a - Y_0) + Z_0 (Z_a - Z_0) = 0 \\ \frac{X - X_a}{X_0} = \frac{Y - Y_a}{Y_0} = \frac{Z - Z_a}{Z_0} \\ Z_a = Z_0 \end{cases}$$

Soit :

$$\begin{cases} X_a = X \\ Y_a = Y \\ Z_a = Z_0 \end{cases}$$

Après projection, il vient :

$$\begin{cases} u = \frac{\alpha_u}{Z_0} X + \frac{\gamma}{Z_0} Y + u_0 \\ v = \frac{\alpha_v}{Z_0} Y + v_0 \end{cases}$$

On remarque alors qu'on obtient des équations linéaires en les coordonnées X et Y mais, contrairement à l'approximation para-perspective, il n'a pas été nécessaire d'introduire des paramètres supplémentaires. Cependant, les paramètres intrinsèques dépendent, tout comme pour la projection para-perspective, de la profondeur moyenne de la scène.

Ce modèle est équivalent au modèle para-perspectif dans le cas où on regarde droit devant soi (vision fovéale), c'est-à-dire le cas où $X_0 = 0$ et $Y_0 = 0$.

Approximation à l'ordre 0

D'un point de vue équationnel, il s'agit d'un développement à l'ordre 0 en la profondeur Z du modèle perspectif, $Z = Z_0 + o(1)$:

$$\begin{cases} u = \alpha_u \frac{X}{Z_0} + \gamma \frac{Y}{Z_0} + u_0 \\ v = \alpha_v \frac{Y}{Z_0} + v_0 \end{cases}$$

Soit, en notant :

$$\begin{cases} \alpha_{u_{ortho}} = \frac{\alpha_u}{Z_0} \\ \alpha_{v_{ortho}} = \frac{\alpha_v}{Z_0} \\ \gamma_{ortho} = \frac{\gamma}{Z_0} \end{cases}$$

sous forme matricielle :

$$\mathbf{m} = \begin{pmatrix} \alpha_{u_{ortho}} & \gamma_{ortho} & 0 & u_0 \\ 0 & \alpha_{v_{ortho}} & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{M} \quad (3.5)$$

On retrouve bien le modèle orthographique, ce qui nous montre l'équivalence entre l'approximation géométrique et l'approximation à l'ordre 0.

L'avantage de cette projection est de nous procurer des équations affines ce qui simplifie énormément les calculs par la suite. Bien plus restreint que le modèle para-perspectif, en général, ce modèle est cependant bien réaliste pour des objets loins de la caméra où dont l'image est dans la zone fovéale (i.e. près de l'axe optique) et pour des mouvements laissant le plan rétinien invariant (translation parallèle au plan rétinien, rotation d'axe perpendiculaire au plan rétinien).

3.2.3 Plusieurs modèles en une équation

Le modèle para-perspectif a l'avantage de linéariser les équations de projection mais possède le désavantage d'introduire deux paramètres supplémentaires. Le modèle orthographique réunit les avantages en linéarisant les équations sans faire intervenir de paramètre supplémentaire.

Ces deux modèles font intervenir la scène par l'intermédiaire d'un point de profondeur moyenne mais conservent tous deux le facteur d'échelle (les objets proches paraissent plus grands que les objets lointains).

Cependant, le modèle para-perspectif reste tout de même plus précis que le modèle orthographique, car il tient compte des positions des objets en périphérie (vus sous un angle différent).

Nous proposons ici un modèle général de projection pouvant englober les modèles vus précédemment, en notant \hat{x} la variable générique x :

$$\kappa \mathbf{m} = \begin{pmatrix} \hat{\alpha}_u & \gamma & \lambda \hat{\beta}_u + \mu u_0 & (1 - \mu) \hat{u}_0 \\ 0 & \hat{\alpha}_v & \lambda \hat{\beta}_v + \mu v_0 & (1 - \mu) \hat{v}_0 \\ 0 & 0 & \mu & (1 - \mu) \end{pmatrix} \mathbf{M} \quad (3.6)$$

Dans ce modèle, on distingue deux types de paramètres :

- **Paramètres modaux** : λ et μ
- **Paramètres intrinsèques** : $\alpha_u, \alpha_v, \beta_u, \beta_v, u_0, v_0$ et γ

Les paramètres modaux $\lambda \in \{0, 1\}$ et $\mu \in \{0, 1\}$ déterminent le modèle de projection :

	λ	μ
Projection perspective pure	1	1
Projection orthographique	0	0
Projection para-perspective	1	0

Et nous récapitulons dans le tableau suivant les valeurs des paramètres intrinsèques pour les approximations de la projection perspective :

	projection perspective	projection para-perspective	projection orthographique
$\hat{\alpha}_u$	α_u	$\frac{\alpha_u}{Z_0}$	$\frac{\alpha_u}{Z_0}$
$\hat{\alpha}_v$	α_v	$\frac{\alpha_v}{Z_0}$	$\frac{\alpha_v}{Z_0}$
$\hat{\beta}_u$	0	$-\frac{1}{Z_0} \frac{\alpha_u X_0 + \gamma Y_0}{Z_0}$	0
$\hat{\beta}_v$	0	$-\frac{1}{Z_0} \frac{\alpha_v Y_0}{Z_0}$	0
$\hat{\gamma}$	γ	$\frac{\gamma}{Z_0}$	$\frac{\gamma}{Z_0}$
\hat{u}_0	u_0	$u_0 + \alpha_u \frac{X_0}{Z_0} + \gamma \frac{Y_0}{Z_0}$	u_0
\hat{v}_0	v_0	$v_0 + \alpha_v \frac{Y_0}{Z_0}$	v_0

Chapitre 4

Raffinement de cette modélisation.

4.1 Les distorsions.

Les distorsions linéaires ont déjà été prises en compte dans le modèle de projection perspective. Nous nous intéresserons donc ici aux distorsions non-linéaires dont les causes proviennent de défauts de géométrie du système optique : mauvais alignement des centres optiques des dioptrés, courbure non parfaite des lentilles, ...

4.1.1 Distorsion radiale.

En pratique, c'est la distorsion géométrique radiale symétrique qui a l'influence la plus importante. Les distorsions géométriques radiales les plus couramment observées sont de deux formes, comme illustré par la figure 4.1 :

- distorsions en forme de barillet
- distorsions en forme de coussinet

On découple généralement la distorsion en deux composantes radiale et tangentielle. Si on note ρ et θ les coordonnées polaires d'un point (x, y) de l'image, $\delta_\rho(\rho, \theta)$ et $\delta_\theta(\rho, \theta)$ les composantes radiale et tangentielle de la distorsion, $\delta_x(x, y)$ et $\delta_y(x, y)$, les composantes selon x et y de la distorsion, s'expriment au premier ordre selon :

$$\begin{cases} \delta_x(x, y) &= \cos(\theta) \delta_\rho(\rho, \theta) - \sin(\theta) \delta_\theta(\rho, \theta) \\ \delta_y(x, y) &= \sin(\theta) \delta_\rho(\rho, \theta) + \cos(\theta) \delta_\theta(\rho, \theta) \end{cases}$$

4.1.2 D'où proviennent les distorsions ?

Les principales causes de distorsions sont généralement au nombre de trois. La première concerne les défauts de courbure des lentilles constituant la caméra qui induisent des distorsions purement radiales. Cette déformation de la courbure est cependant utilisée dans le but d'avoir une luminosité constante sur toute la surface de l'image, afin d'éliminer l'effet de vignettage (luminosité forte au centre, plus faible en périphérie). Une approximation suffisante en pratique pour modéliser est donnée par :

$$\begin{cases} \delta_x(x, y) &= k_1 x (x^2 + y^2) \\ \delta_y(x, y) &= k_1 y (x^2 + y^2) \end{cases}$$

k_1 étant une constante dont le signe détermine le sens de distorsion ($k_1 > 0$ correspond aux distorsions en forme de coussinet ; $k_1 < 0$ aux distorsions en forme de barillet)

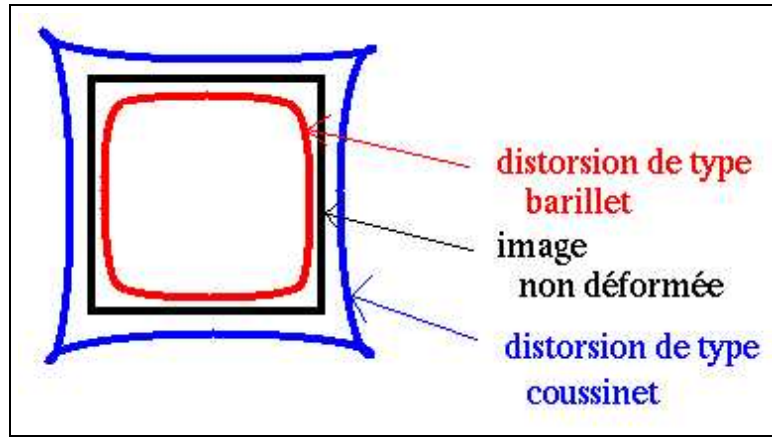


FIG. 4.1 – Les types de distorsions les plus couramment observés : barillet et coussinet

Une deuxième cause de distorsion concerne le mauvais alignement des centres optiques des lentilles à l'intérieur de la caméra qui induit une distorsion de décentrage. Une approximation est donnée par :

$$\begin{cases} \delta_x(x, y) &= p_1 (3x^2 + y^2) + 2p_2 xy \\ \delta_y(x, y) &= 2p_1 xy + p_2 (x^2 + 3y^2) \end{cases}$$

Et enfin, une dernière cause est l'inclinaison des lentilles les unes par rapport aux autres, phénomène équivalent à l'ajout d'un prisme mince, donnant lieu à des distorsions dites prismatiques. Une approximation est donnée par :

$$\begin{cases} \delta_x(x, y) &= s_1 (x^2 + y^2) \\ \delta_y(x, y) &= s_2 (x^2 + y^2) \end{cases}$$

La distorsion totale s'exprime alors au premier ordre comme la somme de ces distorsions, soit :

$$\begin{cases} \delta_x(x, y) &= k_1 x (x^2 + y^2) + p_1 (3x^2 + y^2) + 2p_2 xy + s_1 (x^2 + y^2) \\ \delta_y(x, y) &= k_1 y (x^2 + y^2) + 2p_1 xy + p_2 (x^2 + 3y^2) + s_2 (x^2 + y^2) \end{cases}$$

Ce modèle de distorsion comporte 5 paramètres. Si on envisage le problème dans sa globalité, il faut tenir compte de ces 5 paramètres supplémentaires. Dans notre cas, on préfère découpler les problèmes, sachant que ceux-ci sont décorrélés, et effectuer dans un premier temps la correction des distorsions afin d'appliquer ensuite un modèle de projection linéaire.

D'autre part, l'effet des distorsions étant minime dans la zone fovéale de l'image, on pourra la négliger si on se concentre sur cette zone.

4.1.3 Exemple de correction de distorsion

Dans les images suivantes, des droites ont été tracées afin de mettre en évidence la courbure de l'image du bâtiment de l'ESSI. Cette image a été prise avec un appareil photo standard et développée sur support numérique.

Le logiciel utilisé a été écrit par Frédéric Devernay et le principe consiste à rechercher dans une image des droites approximatives et de rechercher, en considérant uniquement les distorsions radiales, le coefficient k de la transformation à appliquer à l'image afin de redresser au mieux ces droites.



FIG. 4.2 – Image originale : des droites ont été superposées afin de mettre en évidence les distorsions

4.2 Modèle de lentille épaisse

Principalement utilisé dans le cas de la modélisation du zoom, le modèle de lentille épaisse nécessite de déterminer des paramètres (position des plans principaux, points focaux, positions d'entrée et de sortie des lentilles et position du centre de la lentille) avec un dispositif particulier et des instruments très précis, ce qu'il est impossible de faire pour les caméras du commerce comme discuté précédemment.

Heureusement, le modèle de lentille épaisse peut être ré-exprimé en un modèle sténopé dans le cas où la distance de l'objet à l'image est considérée comme variable et que l'image de l'objet reste dans un voisinage restreint de la rétine.

Cela nous permet de ne pas considérer le modèle de lentille épaisse dans la suite mais on considérera les paramètres intrinsèques comme étant des fonctions de la scène observée ce qui est déjà le cas pour les restrictions du modèle projectif. Nous tenterons par ailleurs de rester le plus possible dans la zone fovéale de l'image pour une meilleure adéquation de notre modèle.

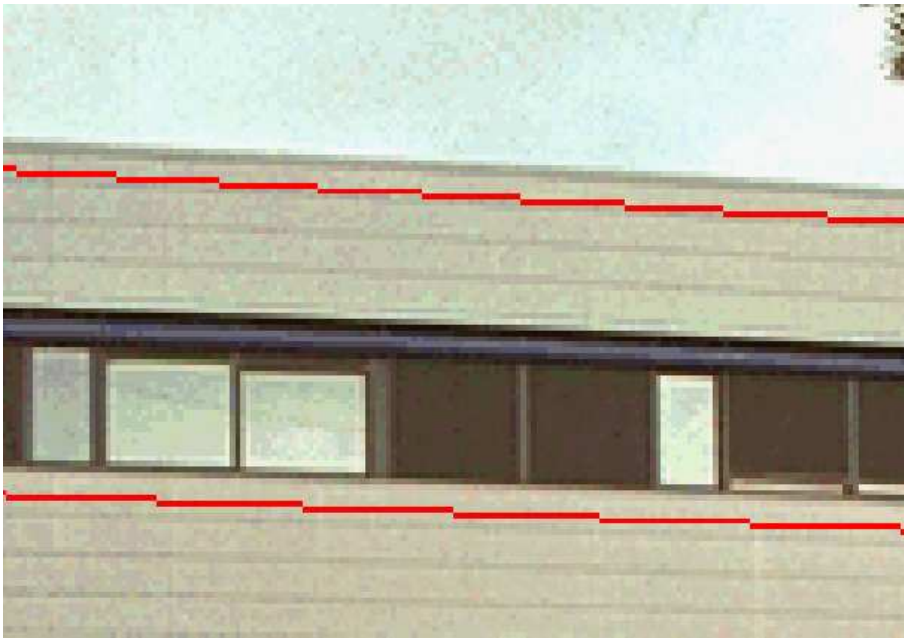


FIG. 4.3 – Zoom sur un détail



FIG. 4.4 – Image corrigée ($k=0.03$) : on voit ici que les droites tracées coïncident avec les arêtes dans l'image.

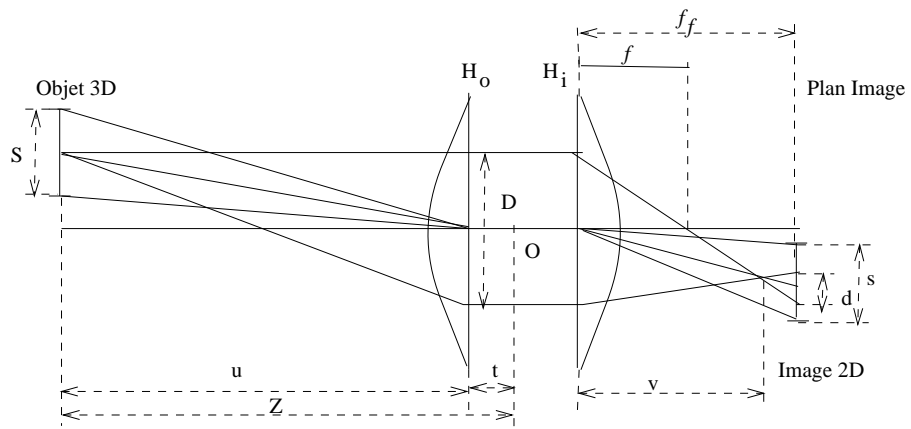


FIG. 4.5 – Modèle de lentille épaisse d'après Pahlavan et al. 1993

Chapitre 5

Equation de projection dans un repère global : paramétrisation du mouvement

Afin d'exprimer la projection effectuée par une caméra dans un repère quelconque de l'espace, on va s'intéresser à la transformation qui amène le repère attaché à la caméra \mathcal{R}_i à un tel repère. Nous allons ainsi passer en revue les différentes paramétrisations d'une telle transformation, ce qui nous sera également utile dans le cadre de l'étude du mouvement.

5.1 Mouvement rigide

On suppose que notre scène est rigide ou bien constituée d'objets rigides. La contrainte de rigidité impose que les déplacements d'une position \mathbf{M} vers une position \mathbf{M}' sont des translations, exprimées par un vecteur $\mathbf{t} = [t_0 \ t_1 \ t_2]^T$ et des rotations, exprimées par une matrice orthogonale positive \mathbf{R} , composées selon :

$$\mathbf{M}' = \mathbf{R} \mathbf{M} + \mathbf{t}$$

Une matrice de rotation \mathbf{R} dans l'espace est une matrice carrée (3x3) qui a pour propriété :

$$\mathbf{R}^{-1} = \mathbf{R}^T$$

ce qui fournit 6 contraintes indépendantes sur les coefficients de la matrice de rotation. Ainsi, \mathbf{R} ne dépend plus que de 3 paramètres. Il existe plusieurs façons pour choisir ces paramètres, ce qui est l'objet des paragraphes suivants.

5.2 Angles d'Euler

Cette paramétrisation est la plus naturelle car chacun des paramètres correspond à la rotation autour d'un des trois axes du repère, comme illustré par la figure 5.1.

La matrice de rotation s'exprime alors comme la composée de ces trois rotations dites canoniques :

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \zeta & -\sin \zeta \\ 0 & \sin \zeta & \cos \zeta \end{bmatrix}$$

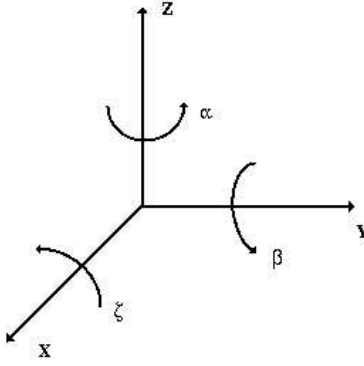


FIG. 5.1 – Les angles d'Euler

L'apparente simplicité de cette représentation cache la complexité de la relation qui permet de déterminer les paramètres α (lacet), β (tangage) et γ (roulis) à partir d'une matrice de rotation \mathbf{R} . Par exemple, l'ordre des matrices intervient d'où différentes définitions non compatibles (coordonnées sphériques, ...). Mais plus important encore, il faut un atlas de 4 cartes avec les paramètres d'Euler pour paramétriser l'ensemble des rotations exceptées celle pour lesquelles $\beta = \pm \frac{\pi}{2}$ et que l'on ne peut pas paramétriser ainsi.

Ainsi, un nombre important de rotations, que l'on ne veut pas s'interdire dans des applications de vision-robotique, ne sont pas représentables par les angles d'Euler.

Nous allons alors nous intéresser à d'autres représentations.

5.3 Exponentielle de vecteur

On note $\boldsymbol{\omega} = [\omega_0 \ \omega_1 \ \omega_2]^T$ le vecteur rotation dont la norme correspond à l'angle de rotation θ et s'exprime par $\theta = \|\boldsymbol{\omega}\|$, tandis que l'axe de rotation a pour vecteur directeur unitaire $\mathbf{u} = \frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|}$. L'opérateur produit vectoriel avec $\boldsymbol{\omega}$ s'écrit :

$$\tilde{\boldsymbol{\omega}} = [\boldsymbol{\omega}]_{\times} = \boldsymbol{\omega} \wedge . = \begin{bmatrix} \omega_0 \\ \omega_1 \\ \omega_2 \end{bmatrix} \wedge \begin{bmatrix} . \\ . \\ . \end{bmatrix} = \begin{bmatrix} 0 & -\omega_2 & \omega_1 \\ \omega_2 & 0 & -\omega_0 \\ -\omega_1 & \omega_0 & 0 \end{bmatrix}$$

La matrice de rotation \mathbf{R} s'exprime alors par :

$$\mathbf{R} = e^{\tilde{\boldsymbol{\omega}}} = e^{[\boldsymbol{\omega}]_{\times}}$$

On vérifie aisément que cette expression est celle d'une rotation :

$$\mathbf{R}\mathbf{R}^T = \mathbf{I}, \det(\mathbf{R}) = 1$$

Cependant, nous devons ajouter 3 cartes de paramètres pour représenter l'ensemble des rotations, y compris celles qui sont des retournements ($\theta = \pi$).

Si on développe cette expression en utilisant le développement en série de l'exponentielle, les puissances de $\tilde{\boldsymbol{\omega}}$ supérieures à 2 s'expriment en fonction de $(\mathbf{I}, \tilde{\boldsymbol{\omega}}, \tilde{\boldsymbol{\omega}}^2)$ (car on peut facilement montrer que $\tilde{\boldsymbol{\omega}}^3 = -\theta^2 \tilde{\boldsymbol{\omega}}$) et on obtient une représentation sous forme de trois termes, connue sous le nom de formule de Rodrigues, et dont les coefficients sont eux-mêmes des séries entières trigonométriques :

$$\mathbf{R} = \mathbf{I} + \frac{\sin \theta}{\theta} \tilde{\boldsymbol{\omega}} + \frac{(1 - \cos \theta)}{\theta^2} \tilde{\boldsymbol{\omega}}^2$$

soit :

$$\mathbf{R} = \mathbf{I} + \sin \theta \tilde{\mathbf{u}} + (1 - \cos \theta) \tilde{\mathbf{u}}^2$$

Compte-tenu de la nécessité de plusieurs cartes de paramètres pour cette représentation, on propose alors de choisir le vecteur rotation de la façon suivante :

$$\mathbf{r} = 2 \tan\left(\frac{\theta}{2}\right)\mathbf{u} \Leftrightarrow \theta = 2 \arctan\left(\frac{\|\mathbf{r}\|}{2}\right) \quad (5.1)$$

car on obtient alors une formule de même type que la formule de Rodrigues, mais qui, cette fois, est rationnelle :

$$\mathbf{R} = \mathbf{I} + \left[\frac{\tilde{\mathbf{r}} + \frac{1}{2}\tilde{\mathbf{r}}^2}{1 + \frac{\mathbf{r}^T \cdot \mathbf{r}}{4}} \right] \quad (5.2)$$

Les développements limités pour de faible rotation ($\theta \rightarrow 0$) ou des rotations proches des retournements ($\theta \rightarrow \pi$) s'écrivent de façon simple :

$$\begin{aligned} \mathbf{R} &= \mathbf{I} + \tilde{\omega} + \frac{1}{2}\tilde{\omega}^2 + o(\|\omega\|^2) \\ \mathbf{R} &= \mathbf{I} + \frac{2\tilde{\omega}^2}{\|\omega\|^2} + o\left(\frac{1}{\|\omega\|^2}\right) \end{aligned}$$

Le développement limité en $(\frac{1}{\|\omega\|})^2$ permet de modéliser le comportement de la modélisation de la rotation lorsque son angle se rapproche de π et donc que la tangente du demi-angle tend vers l'infini. Dans ce cas, on obtient effectivement l'expression d'un retournement d'axe \mathbf{u} :

$$\mathbf{R} = \mathbf{I} + 2\tilde{\mathbf{u}}^2$$

d'où :

$$\begin{aligned} \mathbf{R} \mathbf{x} &= \mathbf{x} + 2\mathbf{u} \wedge (\mathbf{u} \wedge \mathbf{x}) \\ &= 2(\mathbf{u} \cdot \mathbf{x})\mathbf{u} - \mathbf{x} \end{aligned}$$

Cette représentation nous permet donc de paramétrer toutes les rotations de l'espace y compris celles qui sont des retournements. Il est donc naturel, si l'on doit tenir compte de ces retournements, de réécrire \mathbf{r} en coordonnées homogènes sous la forme de \mathbf{q} :

$$\mathbf{q} = \left(\cos\left(\frac{\theta}{2}\right), \sin\left(\frac{\theta}{2}\right)\mathbf{u}\right) = (s, \mathbf{v})$$

L'équation 5.2 devient alors rationnelle :

$$\mathbf{R} = \mathbf{I} + 2s\tilde{\mathbf{v}} + 2\tilde{\mathbf{v}}^2$$

5.4 Représentation par quaternions

C'est à Hamilton que nous devons l'invention des quaternions au siècle dernier, précurseurs des vecteurs. Leur utilisation première fut l'expression de la rotation ce qui ne fut pas tâche aisée.

Un quaternion \mathbf{q} peut s'exprimer de deux façons différentes par :

- $\mathbf{q} = [\lambda_0 \ \lambda_1 \ \lambda_2 \ \lambda_3]^T$ où $\lambda_0, \lambda_1, \lambda_2$ et λ_3 , sont des scalaires
- $\mathbf{q} = (s, \mathbf{v})$ où s est un scalaire, \mathbf{v} un vecteur : $\mathbf{v} = [v_0 \ v_1 \ v_2]^T$

Le produit \times sur les quaternions est défini par :

$$\mathbf{q} \times \mathbf{q}' = (s s' - \mathbf{v} \cdot \mathbf{v}', s \mathbf{v}' + s' \mathbf{v} + \mathbf{v} \wedge \mathbf{v}')$$

le conjugué par :

$$\bar{\mathbf{q}} = (-s, \mathbf{v})$$

et la norme par :

$$\|\mathbf{q}\|^2 = \bar{\mathbf{q}}\mathbf{q} = \mathbf{q}\bar{\mathbf{q}} = (s^2 + \|\mathbf{v}\|^2, 0) = (\|\mathbf{q}\|^2, 0)$$

Dans le cas de la rotation, $\|\mathbf{q}\| = 1$.

La rotation d'un vecteur $\boldsymbol{\chi}$ par le quaternion \mathbf{q} s'exprime par :

$$\mathbf{R}\boldsymbol{\chi} = \mathbf{q} \times \boldsymbol{\chi} \times \bar{\mathbf{q}}$$

En notant θ l'angle de rotation et \mathbf{u} le vecteur directeur unitaire de l'axe de rotation, comme précédemment :

$$\begin{cases} s &= \cos\left(\frac{\theta}{2}\right) \\ \mathbf{v} &= \sin\left(\frac{\theta}{2}\right) \mathbf{u} \end{cases}$$

On peut remarquer ici que cela reste vrai pour $\theta = 0$ ou $\theta = \pi$, c'est-à-dire pour toutes les rotations.

Cette représentation est équivalente à la représentation précédente car la matrice de rotation s'exprime par :

$$\begin{bmatrix} \lambda_0^2 + \lambda_1^2 - \lambda_2^2 - \lambda_3^2 & 2(\lambda_1\lambda_2 - \lambda_0\lambda_3) & 2(\lambda_1\lambda_3 + \lambda_0\lambda_2) \\ 2(\lambda_1\lambda_2 + \lambda_0\lambda_3) & \lambda_0^2 - \lambda_1^2 + \lambda_2^2 - \lambda_3^2 & 2(\lambda_2\lambda_3 - \lambda_0\lambda_1) \\ 2(\lambda_1\lambda_3 - \lambda_0\lambda_2) & 2(\lambda_2\lambda_3 + \lambda_0\lambda_1) & \lambda_0^2 - \lambda_1^2 - \lambda_2^2 + \lambda_3^2 \end{bmatrix}$$

sous la contrainte $\lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1$, expression qui se simplifie en ¹ :

$$\mathbf{R} = \mathbf{I} + 2s\tilde{\mathbf{v}} + 2\tilde{\mathbf{v}}^2$$

qui correspond à une des formes de la formule de Rodrigues.

5.5 Analyse des diverses représentations

Le choix d'une représentation de rotation s'effectue selon l'usage qu'on désire en faire, afin d'obtenir les formes algébriques les plus simples et selon les valeurs d'angle utilisées afin d'éviter les points critiques. Cependant, d'un point de vue de calcul numérique, il a été démontré que la représentation de la rotation n'a pas d'influence numérique, sauf lorsque l'on s'approche du voisinage d'un point critique.

Généralement, dans la suite de ce cours, notre choix s'est porté sur l'expression de la matrice de rotation \mathbf{R} en fonction du vecteur rotation \mathbf{r} défini par l'équation (5.1), cette expression ayant l'avantage d'être en fraction de polynômes contrairement aux autres représentations proposées.

Ainsi, r_0, r_1, r_2, t_0, t_1 et t_2 , constituent les paramètres extrinsèques tandis que $\alpha_u, \alpha_v, \gamma, u_0$ et v_0 constituent les paramètres intrinsèques. Il faut cependant noter qu'ils ne constituent pas forcément un ensemble de paramètres indépendants mais sont reliés par des relations dites *équations de Kruppa*.

¹Démonstration à titre d'exercice (encore un !)

Chapitre 6

Notions de Calibration

Dans ce chapitre, pour des raisons de simplifications, on se place uniquement dans le cas de la projection perspective.

6.1 Introduction

Calibrer une caméra, c'est trouver à la fois :

ses paramètres intrinsèques : $(\alpha_u, \alpha_v, \gamma, u_0, v_0)$

ses paramètres extrinsèques : $(r_0, r_1, r_2, t_0, t_1, t_2)$

Ce qui revient à déterminer la matrice \mathbf{P} et le vecteur \mathbf{p} définit, dans le cas perspectif, selon :

$$Z \mathbf{m} = \underbrace{\begin{pmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{A}} \cdot (\mathbf{R} \quad \mathbf{t}) \mathbf{M} \quad (6.1)$$

$(\mathbf{P}=\mathbf{A}\mathbf{R}, \mathbf{p}=\mathbf{A}\mathbf{t})$

6.2 Comment calibrer ?

La calibration se fait 2 étapes. On va tout d'abord déterminer (\mathbf{P}, \mathbf{p}) à partir de correspondances entre points 3D connus et leurs projetés, connus, dans l'image (voir paragraphe 6.2.1). Ensuite, on va déterminer les paramètres intrinsèques et les paramètres extrinsèques à partir de (\mathbf{P}, \mathbf{p}) (voir paragraphe 6.2.2), où :

$$\lambda \mathbf{m} = \mathbf{P} \cdot \mathbf{M} + \mathbf{p}$$

avec $\mathbf{P} = \mathbf{A}\mathbf{R}$ et $\mathbf{p} = \mathbf{A}\mathbf{t}$

6.2.1 Déterminer la projection

Cette première étape dans la procédure de calibration a pour but la détermination du couple (\mathbf{P}, \mathbf{p}) . Or, on remarque que l'équation 6.1 s'écrit :

$$\begin{cases} 0 & = & -\lambda u + P_{00} X + P_{01} Y + P_{02} Z + p_0 \\ 0 & = & -\lambda v + P_{10} X + P_{11} Y + P_{12} Z + p_1 \\ 0 & = & -\lambda + P_{20} X + P_{21} Y + P_{22} Z + p_2 \end{cases}$$

Si on connaît les coordonnées de quelques points de l'espace ainsi que de leur projection, les équations précédentes vont former des systèmes linéaires permettant de déterminer les coefficients de \mathbf{P} et \mathbf{p} ainsi que les coefficients λ . Il suffit d'ailleurs de connaître 6 points seulement, à la condition que ces points ne soient pas coplanaires (écrivez le système et cherchez pourquoi).

En réalité, pour obtenir une meilleure précision, on prend plus de 6 points et on effectue une minimisation.

Pour la détermination des coordonnées des points, on utilise une mire de calibration (voir, par exemple, la figure 6.1) dont un modèle 3D est connu (on l'a fabriquée pour) et dont on mesure avec précision la position et l'orientation par rapport à celle de la caméra.

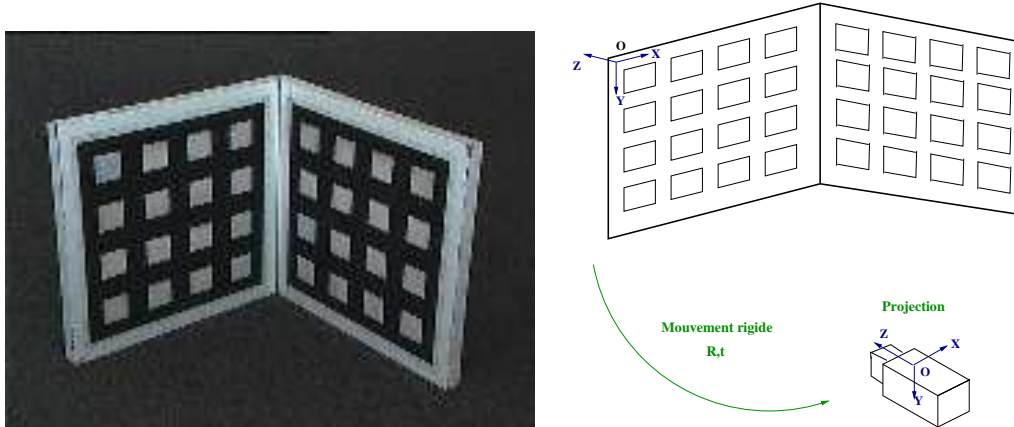


FIG. 6.1 – Mire de Calibration

6.2.2 Déterminer les paramètres

Considérons la matrice $\mathbf{K} = \mathbf{P} \cdot \mathbf{P}^T$, symétrique positive. Celle-ci s'exprime selon :

$$\mathbf{K} = \mathbf{A} \cdot \mathbf{A}^T = \begin{pmatrix} \alpha_u^2 + \gamma^2 + u_0^2 & u_0 v_0 + \gamma \alpha_u & u_0 \\ u_0 v_0 + \gamma \alpha_v & \alpha_v^2 + v_0^2 & v_0 \\ u_0 & v_0 & 1 \end{pmatrix}$$

et peut donc être paramétrisée de la sorte :

$$\mathbf{K} = \begin{pmatrix} b_u & b_c & u_0 \\ b_c & b_v & v_0 \\ u_0 & v_0 & 1 \end{pmatrix}$$

Contrairement à \mathbf{P} , \mathbf{K} dépend uniquement des paramètres intrinsèques, mais surtout, ce qui est intéressant, c'est que ces équations nous permettent justement de calculer les paramètres intrinsèques, et ce, sans ambiguïté car les coefficients α_u et α_v sont strictement positifs.

Ainsi, on obtient les relations inverses suivantes :

$$\left. \begin{array}{l} \alpha_u \geq 0 \\ \alpha_v \geq 0 \end{array} \right\} \Rightarrow \begin{array}{l} \alpha_u = \sqrt{b_u - u_0^2 - \gamma^2} \\ \alpha_v = \sqrt{b_v - v_0^2} \\ \gamma = \frac{b_c - u_0 v_0}{\alpha_v} \end{array}$$

Sachant que la matrice de rotation et le vecteur de translation s'expriment de la façon suivante :

$$\begin{cases} \mathbf{R} = \mathbf{A}^{-1} \mathbf{P} \\ \mathbf{t} = \mathbf{A}^{-1} \mathbf{p} \end{cases}$$

on retrouve alors les paramètres extrinsèques.

Une contrainte est que l'on doit avoir :

$$(\mathbf{P.P}^T)^{22} = \mathbf{K}^{22} = 1$$

6.3 Conclusion

L'intérêt de la calibration est de permettre de connaître avec précision la transformation de l'espace vers l'image.

Par contre, elle possède les inconvénients majeurs :

- de nécessiter d'une mire de calibration : c'est encombrant
- de ne pas permettre de faire varier les paramètres intrinsèques (focus, zoom, ...)

La suite au prochain cours ...

Annexe A : La géométrie projective : généralités et applications

La géométrie projective date du début du XX^{ème} siècle et on doit son utilisation massive en vision par ordinateur à Olivier Faugeras et Richard Hartley. Pour une bonne introduction, on conseillera la lecture de *Three-Dimensional Computer Vision : a Geometric Viewpoint* par Olivier Faugeras, MIT Press¹.

Cette géométrie permet l'unification des géométries Euclidienne et Affine qui en deviennent des cas particuliers.

6.3.1 Définitions - Principes

Un espace projectif \mathbb{P}^n peut être défini comme l'ensemble des droites de \mathbb{R}^{n+1} passant par 0. On note $\mathbb{P}^n = \mathcal{P}(\mathbb{R}^{n+1}) = \mathbb{R}^{n+1} \setminus \{0\}$.

Un point projectif x de \mathbb{P}^n est un vecteur de $(n + 1)$ éléments dont au moins une des composantes est non nulle :

$$\mathbf{x} = [x_1, x_2, \dots, x_{n+1}]^T \in \mathbb{P}^n \iff \exists i / x_i \neq 0$$

L'égalité entre deux vecteurs de \mathbb{P}^n est définie comme suit :

$$\mathbf{x}' = \mathbf{x} \iff \exists \kappa / \mathbf{x}' = \kappa \mathbf{x}$$

La base canonique de l'espace projectif \mathbb{P}^n est composée de $(n + 2)$ vecteurs (la base comporte un élément supplémentaire par rapport à \mathbb{R}^{n+1} à cause de la notion d'égalité² :

$$\begin{array}{c} [1, 1, \dots, 1]^T \\ [1, 0, 0, \dots, 0]^T \\ \vdots \\ [0, \dots, 0, \overset{\downarrow}{1}, 0, \dots, 0]^T \\ \vdots \\ [0, \dots, 0, 1]^T \end{array}$$

On définit une collinéation ou homographie \mathbf{c} comme une application linéaire inversible de \mathbb{P}^n vers \mathbb{P}^n . L'ensemble des collinéations définit le groupe projectif.

¹Ouvrage disponible à la bibliothèque de l'INRIA

²La démonstration de ceci est un très bon exercice

6.3.2 Application à la vision par ordinateur

En coordonnées homogènes, les points de l'espace à distance finie sont représentés par des vecteurs à 4 coordonnées $\mathbf{M} = [X \ Y \ Z \ 1]^T$ tandis que les points à l'infini sont représentés par $\mathbf{M} = [X \ Y \ Z \ 0]^T$. Les points de l'image à distance finie sont représentés par $\mathbf{m} = [x \ y \ 1]^T$ et ceux à l'infini par $\mathbf{m} = [x \ y \ 0]^T$. On obtient ainsi une représentation uniforme des points à l'infini (points de fuite), c'est-à-dire des directions de droite.

Le principe de dualité dans le plan projectif concerne les droites et points qui sont représentés et se manipulent de la même façon. Un point \mathbf{p} est défini comme l'intersection de deux droites \mathbf{d}_1 et \mathbf{d}_2 ; deux points \mathbf{p}_1 et \mathbf{p}_2 définissent une droite d :

$$\begin{aligned}\mathbf{p} &= \mathbf{d}_1 \wedge \mathbf{d}_2 \\ \mathbf{d} &= \mathbf{p}_1 \wedge \mathbf{p}_2\end{aligned}$$

Ainsi, toutes les opérations effectuées sur les points sont également applicables aux droites.

Une collinéation \mathbf{H} du plan vers le plan (homographie) est représentée par une matrice (3x3) telle que $\kappa \mathbf{m}' = \mathbf{H} \mathbf{m}$, où κ est le coefficient d'égalité projective. Par exemple, un changement de repère image représenté par la composée d'une rotation de matrice \mathbf{R} et translation de vecteur \mathbf{t} s'écrit $\mathbf{m}' = [\mathbf{R}|\mathbf{t}] \mathbf{m}$; les changements de base sont linéaires plutôt qu'affines grâce à l'introduction de la coordonnée supplémentaire.

Un autre intérêt de la géométrie projective est la linéarisation du modèle sténopé. En effet, l'équation de projection s'écrit maintenant :

$$Z' \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}$$

et la relation $\mathbf{M}' = \mathbf{R} \mathbf{M} + \mathbf{t}$ devient $\kappa \mathbf{M}' = [\mathbf{R}|\mathbf{t}] \mathbf{M}$ d'où :

$$\kappa \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} (\mathbf{R}|\mathbf{t}) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Chapitre 7

Bibliographie

vision, en général

- O. Faugeras “Three-Dimensional Computer Vision : a Geometric Viewpoint”, *MIT Press*

distorsions

- Thèse de F. Devernay, “Vision Stéréoscopique et Propriétés Différentielles des Surfaces”, Ecole Polytechnique, 1997

pinhole camera

- <http://www.photo.net/photo/pinhole/pinhole.htm>

ouvrages à la documentation de l’INRIA

- <http://www.inria.fr/dias/DOC-fra.html>

Base de cours sur Internet

- CV-Online <http://www.dai.ed.ac.uk/CVonline/geom.htm>