Chapitre 5

Géométrie et calibration des caméras

5.1. Introduction

Dans ce chapitre nous allons brièvement décrire le modèle géométrique associé au processus de saisie d'image à l'aide d'une caméra. Ce modèle est caractérisé par un certain nombre de paramètres que nous allons estimer par étalonnage (ou calibration, ou calibrage). Les paramètres d'une caméra ou paramètres intrinsèques seront estimés en même temps que les paramètres extrinsèques, soit les paramètres qui lient un repère associé à la caméra au repère associé à l'objet étalon. Nous allons étudier en détail un modèle projectif de caméra. Nous allons voir par la suite qu'il existe un modèle simplifié – affine – qui peut être utile dans certaines configurations. Nous allons ensuite proposer une technique de calibration d'une caméra linéaire (une barette CCD).

Il est utile de signaler que lorsqu'on calibre une caméra, on s'intéresse à la modélisation de l'ensemble de la caméra et du convertisseur analogique/digital. Si on change de convertisseur, ou de caméra, il faut alors recalibrer l'ensemble [16].

Ensuite nous allons décrire un capteur comportant deux caméras (capteur stéréoscopique) et nous allons préciser quelques propriétés de ce capteur qui nous permettrons de faciliter la tâche de mise en correspondance. Un capteur composé de deux caméras est un capteur tri-dimensionnel puisqu'il nous permet, par *triangulation passive* de calculer la distance d'un objet au capteur. Nous allons étudier également un capteur tri-dimensionnel à *triangulation active* ainsi qu'une technique d'étalonnage d'un tel capteur.

5.2. Le modèle géométrique d'une caméra

Nous allons caractériser le modèle d'une caméra à l'aide de deux transformations :

1. une projection qui transforme un point de l'espace (3D) en un point image (2D) et

2. une transformation d'un repère *métrique* lié à la caméra à un repère lié à l'image.

5.2.1. La projection perspective

Soit un point O dans le plan image appelé point principal et soit une droite perpendiculaire au plan image passant par O, l'axe optique. Soit un point F placé sur l'axe optique à une distance f du plan image. Le point F est le centre de projection et f est la distance focale. On peut placer le centre de projection devant ou derrière le plan image, dans ce qui suit nous allons le placer derrière, comme sur la figure 5.1.

Un point B se projette dans le plan image le long d'une droite passant par B et F. Choisissons un système de coordonnées attaché à la caméra : le plan $x-y$ de ce repère est parallèle au plan image et l'axe des z est confondu avec l'axe optique. L'origine de ce repère se trouve en F. Soient x, y et z les coordonnées du point B dans le repère qu'on vient de décrire. Dans ce même repère, les coordonnées de la projection de B dans le plan image sont :

$$\begin{aligned} x' &= fx/z \\ y' &= fy/z \\ z' &= f \end{aligned}$$

On peut écrire cette transformation sous forme matricielle :

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \qquad [5.1]$$

Nous avons donc adopté les coordonnées homogènes. Les coordonnées homogènes de B sont $(x, y, z, 1)$ et celles de b (sa projection) sont :

$$\begin{pmatrix} sx' \\ sy' \\ sz' \\ s \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Les coordonnées cartésiennes de b sont sx'/s, sy'/s et sz'/s.

5.2.2. Transformation caméra/image

Les points image sont mesurés en *pixels* dans un repère bi-dimensionnel $u-v$ associé à l'image, figure 5.1. Afin de pouvoir écrire la matrice de transformation du repère caméra au repère image nous devons introduire les paramètres suivants : u_0, v_0 et w_0 sont les coordonnées de F, dans le repère image (mesurées en pixels), k_u est le facteur d'échelle vertical (pixels/mm) et k_v est le facteur d'échelle horizontal. En effet, les pixels d'une caméra sont rarement carrés. La transformation du repère caméra au repère image s'écrit (pour le point b) :

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & 0 \\ 0 & k_v & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix}$$

C'est une transformation affine représentant un changement d'échelle, une rotation et une translation. La composante w étant toujours nulle, on peut ignorer la troisième ligne et écrire cette transformation sous la forme d'une matrice 3×4. Cette transformation représente une application linéaire de l'espace projectif vers le plan projectif :

$$\mathbf{K} = \begin{pmatrix} -k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad [5.2]$$

et

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \mathbf{K} \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix}$$

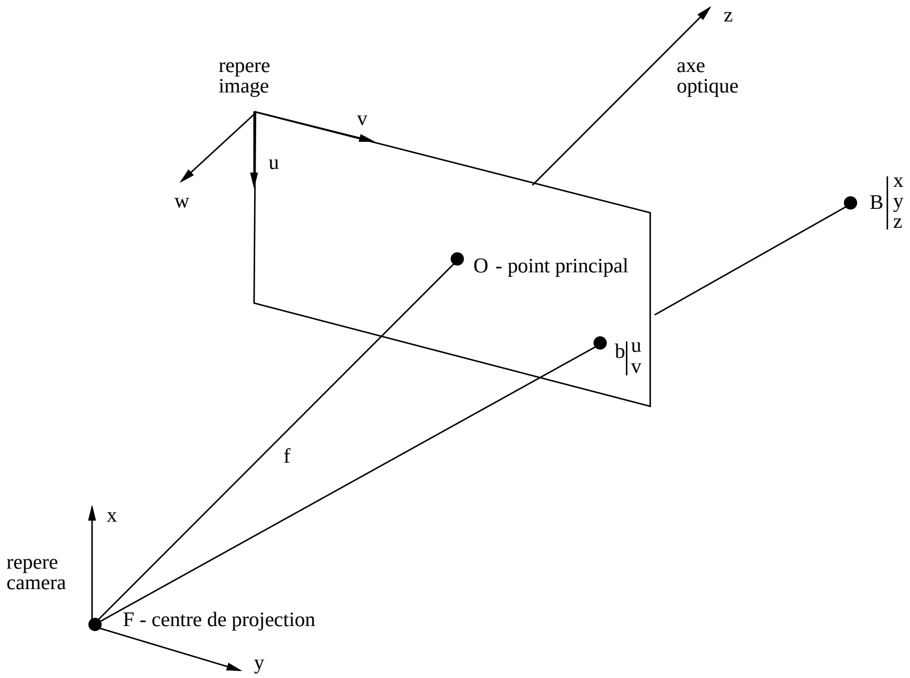

Figure 5.1. *Le modèle géométrique d'une caméra.*

5.2.3. Les paramètres intrinsèques

En multipliant les matrices **K**et **P**(projection perspective suivie d'une transformation affine) nous pouvons écrire les équations du modèle géométrique de la caméra, soit la relation entre les coordonnées caméra $(x\ y\ z)$ du point B et les coordonnées image $(u\ v)$ du point b :

$$\left\{ \begin{array}{lcl} u & = & -k_u f \frac{x}{z} + u_0 \\ v & = & k_v f \frac{y}{z} + v_0 \end{array} \right. \qquad [5.3]$$

En effet, le produit **KP** est :

$$\begin{pmatrix} -k_u & 0 & u_0/f & 0 \\ 0 & k_v & v_0/f & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix}$$

En multipliant tous les coefficients de la matrice par f (ce qui ne change pas le résultat puisque les coordonnées homogènes sont définies à un facteur multiplicatif près) on obtient :

$$\mathbf{I}_c = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad [5.4]$$

qui est une application linéaire de l'espace projectif vers le plan projectif exprimant la transformation perspective :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \mathbf{I}_c \begin{pmatrix} x \\ y \\ z \\ s \end{pmatrix}$$

Ce modèle comporte 4 paramètres : $\alpha_u = -k_u f$, $\alpha_v = k_v f$, u_0 et v_0. Ce sont ces paramètres qui vont être estimés par calibrage. Notons que la distance focale ne peut être calculée explicitement. En introduisant des coordonnées caméra *sans dimension* telles que :

$$\begin{aligned} x_c &= x/z \\ y_c &= y/z \\ z_c &= 1 \end{aligned}$$

on peut maintenant écrire la relation entre les coordonnées image et les coordonnées caméra :

$$\begin{cases} u = \alpha_u x_c + u_0 \;\; avec \;\; \alpha_u < 0 \\ v = \alpha_v y_c + v_0 \end{cases} \qquad [5.5]$$

soit sous forme matricielle :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{C} \begin{pmatrix} x_c \\ y_c \\ 1 \end{pmatrix}$$

avec :

$$\mathbf{C} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [5.6]$$

On peut maintenant décomposer la matrice $\mathbf{I}_c$ en une transformation affine caméra-image (soit la matrice $\mathbf{C}$) et une transformation projective :

$$\mathbf{I}_c = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

5.2.4. Les paramètres extrinsèques

Afin de déterminer les paramètres du modèle de la caméra, nous allons placer devant la caméra une mire (un objet étalon) : un ensemble de points dont les coordonnées sont parfaitement connues dans un repère de la mire qui est différent du repère caméra, figure 5.2. Chaque point de la mire se projette dans l'image et on mesure ses coordonnées dans le repère image. La transformation mire/image se décompose donc en une transformation mire/caméra suivie d'une projection et suivie enfin d'une transformation caméra/image. La transformation mire/caméra se compose d'une rotation et d'une translation :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

La transformation rigide (rotation et translation) peut s'écrire sous la forme d'une transformation homogène :

$$\mathbf{A} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix} \qquad [5.7]$$

Cette matrice représente une transformation rigide (3 rotations et 3 translations) : ce sont les paramètres extrinsèques. La matrice inverse est donnée par l'expression suivante :

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{R}^{\,t} & -\mathbf{R}^{\,t}\mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix} \qquad [5.8]$$

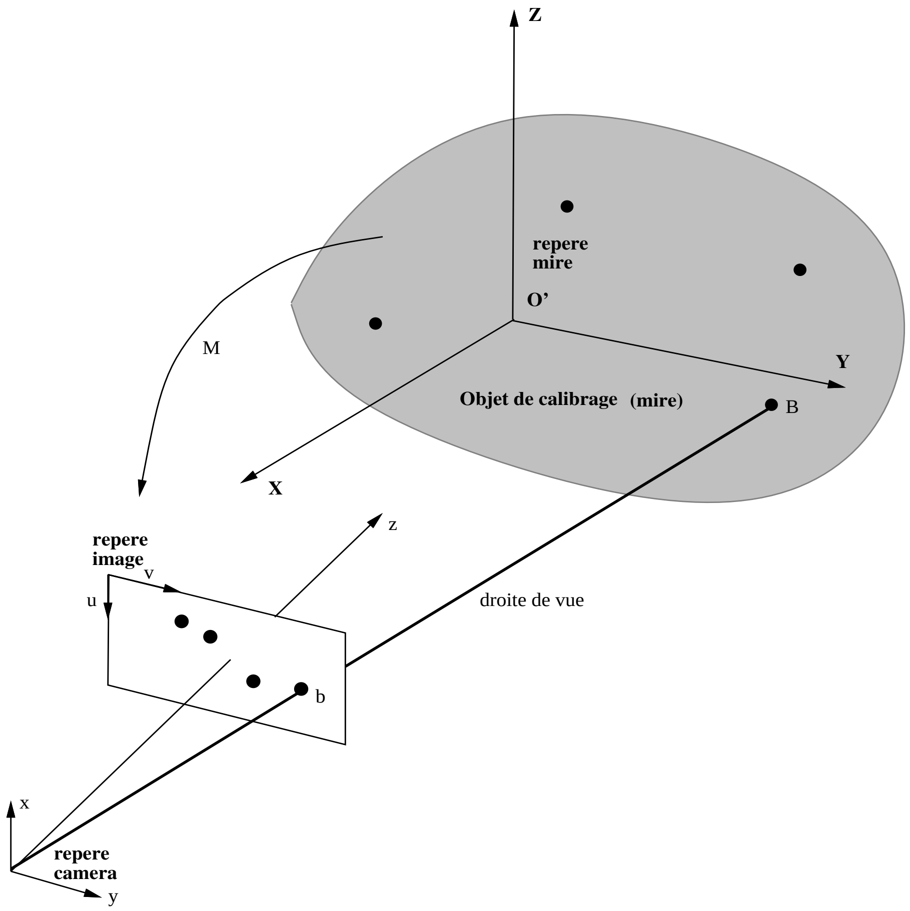

Figure 5.2. *Le principe de calibrage d'une caméra. Un point B dont les coordonnées sont exprimées dans le repère mire se projette en b dont les coordonnées sont exprimées dans le repère image.*

5.2.5. La transformation mire/image

Nous pouvons maintenant écrire la transformation mire/image sous la forme d'une matrice 3×4 appelée matrice de projection perspective et qui peut se décomposer comme suit :

$$\begin{aligned}
\mathbf{M} &= \mathbf{I}_c\mathbf{A} \\
&= \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} && [5.9] \\
&= \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} && [5.10] \\
&= \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix} && [5.11] \\
&= \begin{pmatrix} \alpha_u\mathbf{r}_1 + u_0\mathbf{r}_3 & \alpha_u t_x + u_0 t_z \\ \alpha_v\mathbf{r}_2 + v_0\mathbf{r}_3 & \alpha_v t_y + v_0 t_z \\ \mathbf{r}_3 & t_z \end{pmatrix} && [5.12]
\end{aligned}$$

Dans cette formule la matrice $\mathbf{A}$ a été écrite sous une forme plus compacte en utilisant la notation $\mathbf{r}_1 = (r_{11}\; r_{12}\; r_{13})$:

$$\mathbf{A} = \begin{pmatrix} \mathbf{r}_1 & t_x \\ \mathbf{r}_2 & t_y \\ \mathbf{r}_3 & t_z \\ \mathbf{0} & 1 \end{pmatrix}$$

$\mathbf{M}$ est la matrice de projection perspective et elle peut, en général, s'écrire sous la forme suivante :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \qquad [5.13]$$

Dans cette formule X, Y et Z sont les coordonnées d'un point B de la mire dans le repère mire. Cette matrice peut également s'écrire sous une forme plus simple :

$$\mathbf{M} = \begin{pmatrix} \mathbf{m}_1 & m_{14} \\ \mathbf{m}_2 & m_{24} \\ \mathbf{m}_3 & m_{34} \end{pmatrix}$$

avec $\mathbf{m}_i = (m_{i1} \;\; m_{i2} \;\; m_{i3})$. En identifiant $\mathbf{M}$ avec $\mathbf{I}_c\mathbf{A}$, en tenant compte des propriétés d'orthonormalité de la rotation et en remarquant que l'on doit obtenir une valeur négative pour α_u, on obtient un ensemble d'équations qui permettent de calculer les paramètres intrinsèques et extrinsèques en fonction des coefficients de $\mathbf{M}$. On a :

$$\left\{ \begin{array}{lcl} \mathbf{r}_3 & = & \mathbf{m}_3 \\ u_0 & = & \mathbf{m}_1 \cdot \mathbf{m}_3 \\ v_0 & = & \mathbf{m}_2 \cdot \mathbf{m}_3 \\ \alpha_u & = & -||\mathbf{m}_1 \wedge \mathbf{m}_3|| \\ \alpha_v & = & ||\mathbf{m}_2 \wedge \mathbf{m}_3|| \\ \mathbf{r}_1 & = & 1/\alpha_u(\mathbf{m}_1 - u_0\mathbf{m}_3) \\ \mathbf{r}_2 & = & 1/\alpha_v(\mathbf{m}_2 - v_0\mathbf{m}_3) \\ t_x & = & 1/\alpha_u(m_{14} - u_0 m_{34}) \\ t_y & = & 1/\alpha_v(m_{24} - v_0 m_{34}) \\ t_z & = & m_{34} \end{array} \right. \qquad [5.14]$$

Afin de trouver les paramètres intrinsèques on doit donc :

1. estimer les coefficients de la matrice de projection $\mathbf{M}$ et

2. extraire les paramètres de la caméra à partir de ces coefficients grâce aux formules données par l'equation [5.14].

5.2.6. Un modèle légèrement amélioré

Les équations précédentes montrent que l'on peut déterminer les 10 paramètres intrinsèques et extrinsèques à partir des coefficients de la matrice de projection perspective, soit 11 paramètres, cette matrice étant définie à un facteur multiplicatif près. Ceci implique qu'il y aurait un 11^e paramètre associé à la caméra. Lequel ? Il pourrait s'agir de l'angle entre l'axe des u et l'axe des v qui a été supposé égal à 90^0. Si ces deux axes ne sont pas perpendiculaires on peut introduire comme 11^e paramètre θ, l'angle entre ces deux axes, figure 5.3. La relation entre le repère orthonormé $(u - v)$ et le nouveau repère $(u' - v')$

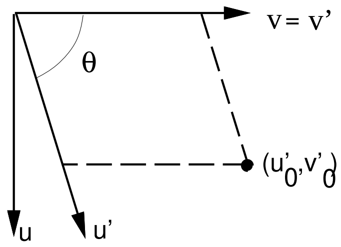

Figure 5.3. *Une image avec des axes non perpendiculaires.*

est :

$$\begin{cases} u' &= u \sin\theta + v \cos\theta \\ v' &= v \end{cases}$$

La matrice donnée par l'équation [5.6] devient dans ces conditions (on divise par $\sin\theta$) :

$$\mathbf{C}' = \begin{pmatrix} \alpha_u & \alpha_v \cot\theta & u'_0 \\ 0 & \alpha_v / \sin\theta & v'_0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [5.15]$$

avec:

$$\begin{cases} u'_0 &= u_0 + v_0 \cot\theta \\ v'_0 &= v_0 / \sin\theta \end{cases}$$

La matrice $\mathbf{M}$ s'écrit maintenant :

$$\mathbf{M} = \begin{pmatrix} \alpha_u \mathbf{r}_1 + (\alpha_v \cot\theta)\mathbf{r}_2 + u'_0\mathbf{r}_3 & \alpha_u t_x + (\alpha_v \cot\theta)t_y + u'_0 t_z \\ (\alpha_v / \sin\theta)\mathbf{r}_2 + v'_0\mathbf{r}_3 & (\alpha_v / \sin\theta)t_y + v'_0 t_z \\ \mathbf{r}_3 & t_z \end{pmatrix}$$

Les formules de l'équation [5.14] permettant d'exprimer les paramètres intrinsèques en fonction des coefficients de la matrice de projection perspective sont maintenant un peu plus complexes.

On obtient immédiatement u'_0 et v'_0 :

$$\begin{aligned} u'_0 &= \mathbf{m}_1 \cdot \mathbf{m}_3 \\ v'_0 &= \mathbf{m}_2 \cdot \mathbf{m}_3 \end{aligned}$$

En effectuant le produit vectoriel $\mathbf{m}_2 \wedge \mathbf{m}_3$ on obtient :

$$\frac{\alpha_v}{\sin\theta} = ||\mathbf{m}_2 \wedge \mathbf{m}_3||$$

ainsi que :

$$\mathbf{r}_2 = \frac{\mathbf{m}_2 - (\mathbf{m}_2 \cdot \mathbf{m}_3)\mathbf{m}_3}{||\mathbf{m}_2 \wedge \mathbf{m}_3||}$$

Le produit scalaire $\mathbf{m}_1 \cdot \mathbf{m}_2$ nous permet d'obtenir le cosinus de l'angle entre les deux axes :

$$\cos\theta = \frac{(\mathbf{m}_1 \cdot \mathbf{m}_2) - (\mathbf{m}_1 \cdot \mathbf{m}_3)(\mathbf{m}_2 \cdot \mathbf{m}_3)}{||\mathbf{m}_2 \wedge \mathbf{m}_3||^2}$$

On effectue ensuite les produits vectoriels $\mathbf{m}_1 \wedge \mathbf{m}_3$ et $\mathbf{m}_1 \wedge \mathbf{m}_2$ et on obtient :

$$\alpha_u = \frac{||\cos\theta(\mathbf{m}_1 \wedge \mathbf{m}_2) - (\cos\theta v_0' - u_0')(\mathbf{m}_1 \wedge \mathbf{m}_3)||}{||\frac{\alpha_v}{\sin\theta}\cos\theta\mathbf{m}_3 - u_0'\mathbf{r}_2||}$$

$$\mathbf{r}_1 = \frac{1}{\alpha_u}(\mathbf{m}_1 - u_0'\mathbf{m}_3 - \frac{\alpha_v}{\sin\theta}\cos\theta\mathbf{r}_2)$$

Le vecteur de translation est égal à :

$$\begin{aligned} t_z &= m_{34} \\ t_y &= \frac{\alpha_v}{\sin\theta}(m_{24} - v_0' t_z) \\ t_x &= \frac{1}{\alpha_u}(m_{14} - u_0' t_z - \frac{\alpha_v}{\sin\theta}\cos\theta t_y) \end{aligned}$$

5.3. Calibrage d'une caméra

Utilisant l'équation [5.13] on peut écrire les coordonnées image d'un point de la scène. On obtient donc :

$$u = \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \qquad [5.16]$$

$$v = \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \qquad [5.17]$$

Notons au passage que ces deux équations décrivent la droite passant par le centre de projection et le point image (u, v), *dans le repère mire*, soit la droite passant par F et b sur la figure 5.2. Cette droite s'appelle également *la droite de vue* associée avec un point image de coordonnées (u, v).

Pour évaluer les coefficients de la matrice $\mathbf{M}$, il suffit d'écrire ce système d'équations pour les points de la mire (pour lesquels on mesure la projection dans l'image). Chaque point (X_i, Y_i, Z_i) se projetant en (u_i, v_i) fournit deux équations. Ces équations sont linéaires par rapport aux coefficients de la matrice. Il faut donc au moins 6 points pour déterminer les 12 coefficients de la matrice $\mathbf{M}$ qui nous intéresse.

Les équations [5.16] et [5.17] peuvent se réécrire comme une combinaison linéaire des m_{ij} :

$$\begin{aligned} X_i m_{11} + Y_i m_{12} + Z_i m_{13} + m_{14} - & \\ -u_i X_i m_{31} - u_i Y_i m_{32} - u_i Z_i m_{33} &= u_i m_{34} \end{aligned} \quad [5.18]$$

$$\begin{aligned} X_i m_{21} + Y_i m_{22} + Z_i m_{23} + m_{24} - & \\ -v_i X_i m_{31} - v_i Y_i m_{32} - v_i Z_i m_{33} &= v_i m_{34} \end{aligned} \quad [5.19]$$

On obtient donc $2n$ équations pour n points et on peut écrire ces équations sous forme matricielle (les indices indiquent la taille des matrices) :

$$\mathbf{K}_{2n\times 11}\, \mathbf{x}_{11} = \mathbf{u}_{2n} \quad [5.20]$$

Soit, plus précisément :

$$\begin{pmatrix} & & & & & \vdots & & & & & \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_iX_i & -u_iY_i & -u_iZ_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_iX_i & -v_iY_i & -v_iZ_i \\ & & & & & \vdots & & & & & \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \end{pmatrix} =$$

$$\begin{pmatrix} \vdots \\ u_i m_{34} \\ v_i m_{34} \\ \vdots \end{pmatrix}$$

5.3.1. Utilisation de la contrainte $m_{34} = 1$

Le système défini par l'équation [5.20] est un système homogène : afin d'obtenir une solution non triviale il faut fixer un des coefficients m_{ij}. On choisit $m_{34} = 1$, ce qui revient à diviser tous les coefficients de la matrice par m_{34} et à déterminer les paramètres de la caméra à un facteur près. On peut remarquer que m_{34} n'est autre que la composante en z du vecteur de translation entre le référentiel mire et le référentiel caméra et qu'on peut facilement s'arranger pour que cette composante ne soit pas nulle. Avec ce choix, la solution de l'équation [5.20] est donnée par :

$$\mathbf{x}_{11} = (\mathbf{K}^t\mathbf{K})^{-1}\mathbf{K}^t\ \mathbf{u} \tag{5.21}$$

Cette solution n'est autre que la solution au sens des moindres carrés d'un système d'équations linéaires lorsque le nombre d'équations est supérieur au nombre d'inconnues. En effet, on peut écrire l'équation [5.20] sous la forme suivante :

$$\mathbf{K}\,\mathbf{x} - \mathbf{u} = \mathbf{e}$$

$\mathbf{e}$ représente un vecteur erreur. Autrement dit, on cherche une solution approximative. La meilleure solution $(\mathbf{x})$ est celle qui minimise le module du vecteur erreur. On cherche donc $\mathbf{x}$ tel que $||\mathbf{e}||^2 = \mathbf{e}^t\mathbf{e}$ soit minimum. On a :

$$\begin{aligned} \mathbf{e}^t\mathbf{e} &= (\mathbf{Kx} - \mathbf{u})^t(\mathbf{Kx} - \mathbf{u}) \\ &= \mathbf{x}^t\mathbf{K}^t\mathbf{Kx} - \mathbf{u}^t\mathbf{Kx} - \mathbf{x}^t\mathbf{K}^t\mathbf{u} + \mathbf{u}^t\mathbf{u} \\ &= \mathbf{x}^t\mathbf{K}^t\mathbf{Kx} - 2\mathbf{x}^t\mathbf{K}^t\mathbf{u} + \mathbf{u}^t\mathbf{u} \end{aligned}$$

En différenciant par rapport à X on obtient :

$$\mathbf{K}^t\mathbf{Kx} - \mathbf{K}^t\mathbf{u} = 0$$

La valeur de $\mathbf{x}$ qui minimise l'erreur est bien celle fournie précédemment. La matrice $(\mathbf{K}^t\mathbf{K})^{-1}\mathbf{K}^t$ est appelée pseudo-inverse de $\mathbf{K}$. La matrice $(\mathbf{K}^t\mathbf{K})^{-1}$ est une matrice 11×11 symétrique et positive. En pratique on utilise des techniques numériques (élimination Gauss-Jordan, décomposition LU ou QR ou décomposition en valeurs singulières) pour calculer une solution approximative aux moindres carrés pour l'équation [5.20]. Voir [156], pages 528–539 pour résoudre cette équation.

Il faut toutefois remarquer que le fait de choisir $m_{34} = 1$ introduit un inconvénient. En effet, les paramètres du modèle de la caméra ne sont calculés qu'à un facteur multiplicatif près, soit $t_z = m_{34}$.

5.3.2. Utilisation de la contrainte $||\mathbf{m}_3|| = 1$: méthode Faugeras-Toscani

En calculant explicitement les coefficients de $\mathbf{M}$ en fonction des coefficients des matrices qui la composent ($\mathbf{I}_c$ et $\mathbf{A}$) on trouve en particulier (voir équation [5.14]) :

$$m_{31} = r_{31} \quad m_{32} = r_{32} \quad m_{33} = r_{33}$$

D'autre part il est facile de vérifier qu'on a : $r_{31}^2 + r_{32}^2 + r_{33}^2 = 1$. On obtient donc :

$$||\mathbf{m}_3||^2 = m_{31}^2 + m_{32}^2 + m_{33}^2 = 1$$

Dans ce qui suit nous montrons comment on peut calculer la matrice $\mathbf{M}$ en tenant compte de cette contrainte. Ce résultat est dû à Faugeras & Toscani [57, 58].

L'équation [5.20] peut maintenant se réécrire :

$$\mathbf{B}_{2n\times 9}\mathbf{x}_9 + \mathbf{C}_{2n\times 3}\mathbf{x}_3 = 0 \qquad [5.22]$$

avec :

$$\mathbf{B}_{2n\times 9} = \begin{pmatrix} & & & & \vdots & & & & \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_i \\ & & & & \vdots & & & & \end{pmatrix}$$

$$\mathbf{C}_{2n\times 3} = \begin{pmatrix} & \vdots & \\ -u_i\,X_i & -u_i\,Y_i & -u_i\,X_i \\ -v_i\,X_i & -v_i\,Y_i & -v_i\,X_i \\ & \vdots & \end{pmatrix}$$

On décompose par ailleurs $\mathbf{x}_{11}$ en deux inconnues (cette fois-ci on intègre m_{34} comme inconnue) :

$$\mathbf{x}_9 = \begin{pmatrix} \mathbf{m}_1 & m_{14} & \mathbf{m}_2 & m_{24} & m_{34} \end{pmatrix}^t$$

$$\mathbf{x}_3 = (\mathbf{m}_3)^t$$

Le critère à minimiser est le même que précédemment :

$$Q = ||\mathbf{B}_{2n\times 9}\mathbf{x}_9 + \mathbf{C}_{2n\times 3}\mathbf{x}_3||^2 \qquad [5.23]$$

avec la contrainte :

$$||\mathbf{x}_3||^2 = 1$$

Le critère peut donc s'écrire de la façon suivante :

$$Q = ||\mathbf{B}\mathbf{x}_9 + \mathbf{C}\mathbf{x}_3||^2 + \lambda(1 - ||\mathbf{x}_3||^2)$$

En développant on obtient :

$$Q = \mathbf{x}_9^t\mathbf{B}^t\mathbf{B}\mathbf{x}_9 + \mathbf{x}_3^t\mathbf{C}^t\mathbf{C}\mathbf{x}_3 + \mathbf{x}_9^t\mathbf{B}^t\mathbf{C}\mathbf{x}_3 + \mathbf{x}_3^t\mathbf{C}^t\mathbf{B}\mathbf{x}_9 + \lambda(1 - \mathbf{x}_3^t\mathbf{x}_3) \qquad [5.24]$$

En imposant que les dérivées partielles par rapport à $\mathbf{x}_9$ et $\mathbf{x}_3$ soient nulles, on obtient les deux équations suivantes :

$$\begin{aligned} \mathbf{B}^t\mathbf{B}\mathbf{x}_9 + \mathbf{B}^t\mathbf{C}\mathbf{x}_3 &= 0 \\ \mathbf{C}^t\mathbf{C}\mathbf{x}_3 + \mathbf{C}^t\mathbf{B}\mathbf{x}_9 - \lambda\mathbf{x}_3 &= 0 \end{aligned}$$

D'où on obtient :

$$\begin{aligned} \mathbf{x}_9 &= -(\mathbf{B}^t\mathbf{B})^{-1}\mathbf{B}^t\mathbf{C}\mathbf{x}_3 \\ \mathbf{D}\mathbf{x}_3 &= \lambda\mathbf{x}_3 \\ \mathbf{D} &= \mathbf{C}^t\mathbf{C} - \mathbf{C}^t\mathbf{B}(\mathbf{B}^t\mathbf{B})^{-1}\mathbf{B}^t\mathbf{C} \end{aligned}$$

Finalement, en substituant dans [5.24] on obtient pour Q :

$$Q = \mathbf{x}_3^t\mathbf{D}\mathbf{x}_3 = \lambda\mathbf{x}_3^t\mathbf{x}_3 = \lambda$$

On peut remarquer que $\mathbf{D}$ est une matrice symétrique et positive 3×3. Elle a donc des valeurs propres réelles et positives. $\mathbf{x}_3$ est un vecteur propre de $\mathbf{D}$ associé à la valeur propre λ. Pour minimiser le critère il faut donc calculer les valeurs propres de la matrice $\mathbf{D}$, choisir la plus petite valeur propre (car c'est elle qui minimise le critère), calculer le vecteur propre qui lui est associé, soit $\mathbf{x}_3$, le normaliser, et finalement calculer $\mathbf{x}_9$. Les coefficients de $\mathbf{M}$ sont fournis par $\mathbf{x}_3$ et $\mathbf{x}_9$. Etant donné que le signe de $\mathbf{x}_3$, vecteur propre, n'est pas défini, on a deux solutions, $\mathbf{M}$ et $-\mathbf{M}$. On peut en choisir une parmi ces solutions en utilisant le fait que l'objet de calibrage se trouve devant la caméra et non pas derrière la caméra. Dans ce cas on doit avoir $m_{34} = t_z > 0$.

5.3.3. Détermination des paramètres du modèle

A partir des coefficients de la matrice $\mathbf{M}$ et grâce aux équations [5.14] on peut déterminer les paramètres intrinsèques et extrinsèques de la caméra. Nous allons nous concentrer sur les paramètres intrinsèques car ils caractérisent la caméra indépendamment de la mire utilisée pour le calibrage. Quant aux paramètres extrinsèques ils caractérisent la position et l'orientation d'une caméra par rapport à un objet (un ensemble de points 3D). La détermination de ces paramètres peut se faire grâce aux équations [5.14] ou par d'autres méthodes qui seront décrites au chapitre 8. Comme nous allons le voir, la détermination des position et orientation relatives entre une caméra et un objet est un problème central en vision et en robotique.

Revenons maintenant aux paramètres intrinsèques et notons le caractère non linéaire des équations [5.14]. En particulier, une petite erreur sur $\mathbf{m}_3$ a une influence considérable sur le calcul des paramètres de la caméra. Puget & Skordas [157] ont calculé l'erreur commise sur les paramètres de la caméra en fonction des erreurs commises sur les coefficients de la matrice $\mathbf{M}$.

L'erreur commise sur le produit scalaire de deux vecteurs $\mathbf{v}$ et $\mathbf{w}$ s'écrit :

$$\Delta(\mathbf{v} \cdot \mathbf{w}) = \Delta\mathbf{v} \cdot \mid \mathbf{w} \mid + \mid \mathbf{v} \mid \cdot \Delta\mathbf{w}$$

avec $\mid \mathbf{v} \mid = (\mid v_x \mid, \mid v_y \mid, \mid v_z \mid)^t$. L'erreur commise sur u_0 peut donc s'écrire :

$$\Delta u_0 = \Delta\mathbf{m}_1 \cdot \mid \mathbf{m}_3 \mid + \mid \mathbf{m}_1 \mid \cdot \Delta\mathbf{m}_3$$

En pratique on peut estimer l'erreur commise sur les m_{ij} en fonction des erreurs commises sur les mesures (les coordonnées dans l'image des points de la mire). Puget & Skordas ont calibré *plusieurs fois* la même caméra en plusieurs positions et ont obtenu des variations sur les valeurs de u_0 et v_0 allant jusqu'à 20 pixels ! Ils proposent d'améliorer cette précision en calibrant plusieurs fois.

Soit $\mathbf{M}^i$ la matrice obtenue avec la méthode Faugeras-Toscani pour la $i^{ème}$ position caméra/mire. Pour chaque position on obtient un jeu de paramètres intrinsèques. Les paramètres optimaux sont obtenus en minimisant un critère de la forme (pour N positions distinctes) :

$$C = \frac{1}{N} \sum_{i=1}^{i=N} (u_0 - u_0^i)^2$$

La valeur moyenne des u_0^i minimise ce critère :

$$u_0 = \frac{1}{N} \sum_{i=1}^{i=N} u_0^i \qquad [5.25]$$

Il y a des expressions équivalentes pour les autres paramètres intrinsèques.

5.4. Modélisation des distorsions

Jusqu'à présent nous avons supposé un modèle linéaire : l'image d'un point est obtenue en intersectant le plan image avec la droite passant par ce point et par le centre de projection. En réalité la lumière ne voyage pas toujours le long d'une droite et l'équation [5.5] doit être remplacée par :

$$\left\{ \begin{array}{lcl} u & = & \alpha_u x_c + u_0 + \Delta u \\ v & = & \alpha_v y_c + v_0 + \Delta v \end{array} \right. \qquad [5.26]$$

Dans cette équation on a rajouté un terme de correction qui peut être non linéaire. Les nombreuses études sur ce sujet ont montré qu'une correction radiale est suffisante dans la plupart des cas [170] :

$$\left\{ \begin{array}{lcl} \Delta u & = & K_1 r^2 (u - u_0) \\ \Delta v & = & K_1 r^2 (v - v_0) \end{array} \right.$$

avec $r^2 = (u - u_0)^2 + (v - v_0)^2$. Le modèle de la caméra a maintenant 5 et non plus 4 paramètres intrinsèques, à savoir α_u, α_v, u_0, v_0, et K_1. Comment peut-on estimer le nouveau modèle d'une caméra ? La réponse n'est pas simple car on ne peut plus linéariser le système d'équations comme nous l'avons fait en l'absence des distorsions. En effet, le système d'équations s'écrit maintenant de la façon suivante :

$$\left\{ \begin{array}{lcl} \vdots & & \\ (u_i - u_0) - \alpha_u \frac{r_{11}X_i + r_{12}Y_i + r_{13}Z_i + t_x}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z} - K_1 r^2 (u_i - u_0) & = & 0 \\ & & \\ (v_i - v_0) - \alpha_v \frac{r_{21}X_i + r_{22}Y_i + r_{23}Z_i + t_y}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + t_z} - K_1 r^2 (v_i - v_0) & = & 0 \\ \vdots & & \end{array} \right.$$

Pour résoudre ce système d'équations il faut faire appel à des techniques d'optimisation non linéaire et le résultat dépend du choix des valeurs initiales des paramètres. Une solution consiste à évaluer les paramètres (sauf pour K_1) utilisant, par exemple, une des techniques linéaires décrites auparavant et à les utiliser par la suite comme valeurs initiales.

Une autre solution consiste à déterminer la correction non linéaire indépendamment du modèle de la caméra. Une telle solution a été recemment proposée dans [28]. Soient (u, v) la position d'un point image et (u', v') la position corrigée :

$$\begin{cases} u' &= u - \Delta u(u, v) \\ v' &= v - \Delta v(u, v) \end{cases}$$

Les termes de correction Δu et Δv sont fonction de la position du pixel dans l'image. La figure 5.4. montre un exemple de distorsion. Les cercles illustrent les pixels tels qu'ils apparaissent dans l'image, (u et v dans la formule précédente) tandis que les carrés illustrent les positions corrigées de ces pixels (u' et v'). Le vecteur de correction $(\Delta u \, \Delta v)$ est illustré par un segment fléché.

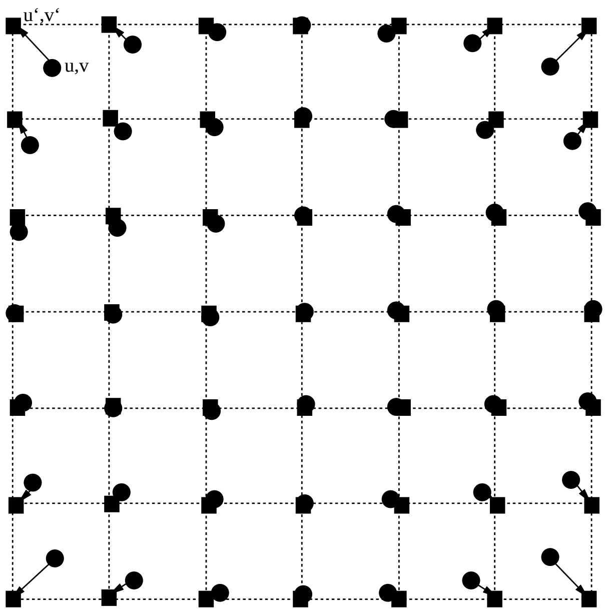

Figure 5.4. *Un exemple de distorsion. Les pixels "ronds" doivent être corrigés pour obtenir les pixels "carrés".*

5.5. Mise en œuvre du calibrage

Il est utile de rappeler que la calibration d'une caméra consiste en réalité en la calibration de l'ensemble objectif optique, caméra, convertisseur analogique/digital. Toute variation d'ouverture, de mise au point ou de focale modifie la configuration géométrique du capteur. Il est donc important que ces paramètres ne varient pas lors du calibrage. Bien évidemment, si l'on modifie un de ces paramètres il faut recalibrer l'ensemble du capteur. Les mêmes remarques s'appliquent au convertisseur analogique digital : tout changement concernant les paramètres du convertisseur entraîne un changement géométrique du capteur.

En outre, on doit disposer d'un objet de calibrage fournissant des "points" dont la position doit être connue avec une très bonne précision dans le repère de l'objet (repère de calibrage). Disons que, pour obtenir un bon calibrage, cette précision doit être inférieure ou égale au 1/10 de millimètre. La figure 5.5. montre l'objet de calibrage que nous utilisons au LIFIA[1] Les points de calibrage sont les sommets d'un ensemble de carrés formant une mire plane. Ce motif a été choisi pour que sa projection dans l'image puisse être mesurée avec une grande précision. L'origine se trouve au sommet du premier carré en haut et à gauche de l'image. Une table à déplacement micrométrique (1/100 millimètre dans notre cas) munie d'une équerre assure un mouvement perpendiculaire au plan de la mire. On obtient ainsi des coordonnées tri-dimensionnelles des points de mesure dans un repère lié au plan de la mire.

Le traitement de l'image a pour rôle de détecter avec précision les projections des points de la mire. Ces points étant, dans notre cas, les sommets de carrés alignés, l'image d'un sommet sera obtenue en intersectant des droites passant par les arêtes des carrés. Ce principe est illustré sur la figure 5.6. La projection d'un point de calibrage est obtenue en calculant l'intersection de deux droites passant "au mieux" par les arêtes des projections de carrés alignés. Un exemple permettra de mieux comprendre ce principe.

La figure 5.7. montre une image de la mire de calibrage (gauche) ainsi que les segments de droites extraits de cette image (droite). Pour extraire ces segments de droite nous avons tout d'abord extrait les contours qui ont été ensuite

[1]Laboratoire d'Informatique Fondamentale et d'Intelligence Artificielle. Unité associée au CNRS (URA 394), 46 avenue Félix Viallet, 38000 Grenoble.

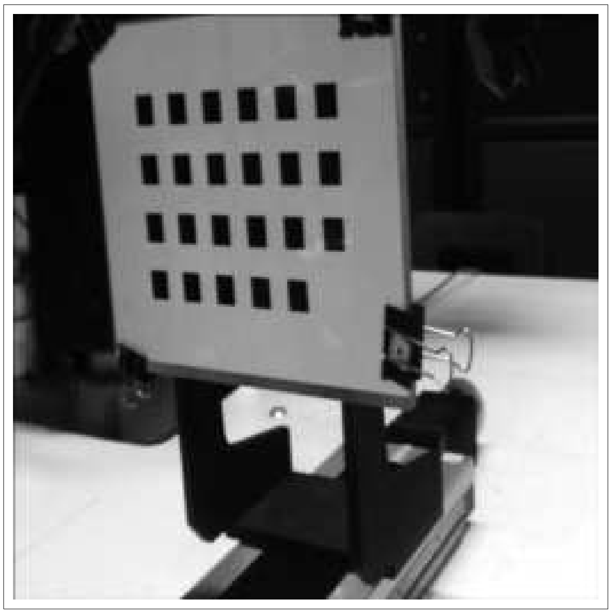

Figure 5.5. *Une vue générale de l'objet de calibrage utilisé. Une mire plane est fixée sur une équerre elle-même fixée sur une table à déplacement micrométrique.*

approximés par des segments de droite (voir chapitre 3). Une méthode basée sur la transformée de Hough (voir chapitre 3) nous a ensuite permis d'extraire deux familles de droites parallèles, familles montrées sur la figure 5.8. Chacune de ces familles est ensuite divisée en un ensemble de segments de droites colinéaires. Finalement, tous les points de contours appartenant au même ensemble de droites colinéaires sont approximés par une droite au sens des moindres carrés. On obtient une "grille", chaque arête de cette grille approximant les points de contours formant un ensemble d'arêtes colinéaires. Les points finalement retenus comme points de projection de la mire sont les intersections des arêtes de cette grille. La figure 5.9. (gauche) montre la grille ainsi obtenue.

Pour calibrer avec les méthodes que nous avons décrites il faut disposer de points de calibrage non coplanaires. Pour cela, on déplace la mire perpendiculairement à son plan sans bouger le capteur. Une fois qu'on a mesuré les projections de plusieurs positions de la mire et qu'on a ainsi obtenu plusieurs centaines de correspondances mire/image, on peut estimer la matrice de pro-

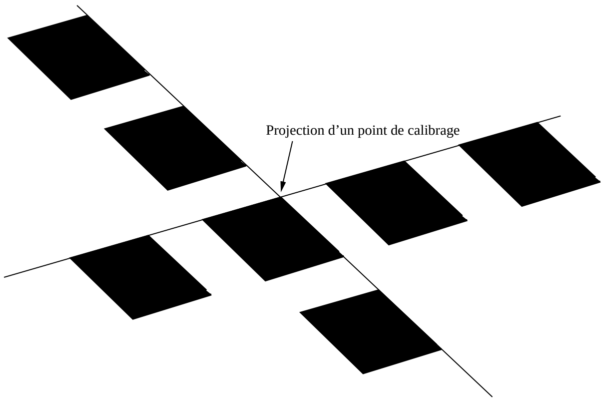

Figure 5.6. *L'image d'un point de calibrage est obtenue en intersectant deux droites passant au mieux par un ensemble d'arêtes alignées.*

jection perspective. La figure 5.9. (droite) montre la superposition de la grille avec des points obtenus en projetant les points de la mire grâce à la matrice de projection qu'on vient de calculer.

5.6. Caméra affine

Jusqu'à présent nous avons étudié le modèle projectif de caméra. Il est cependant utile dans certains cas de considérer un modèle simplifié. On remplace alors la transformation projective 3D/2D par une transformation affine. Le modèle affine de caméra est valide lorsque la taille des objets que l'on observe et faible par rapport à la distance moyenne de ces objets au centre de projection de la caméra.

Soit le schéma de la figure 5.2. et les mêmes notations que celles qui ont été utilisées auparavant. Nous designons par x et y les coordonnées d'un point de l'image dans le repère de la caméra. Ce point s'obtient en projetant un point de la scène dont les coordonnées sont exprimées dans le repère de la scène (ou de calibrage). Les équations de la projection peuvent s'écrire :

$$x = \frac{\mathbf{r}_1 \cdot \overrightarrow{O'B} + t_x}{\mathbf{r}_3 \cdot \overrightarrow{O'B} + t_z} \qquad [5.27]$$

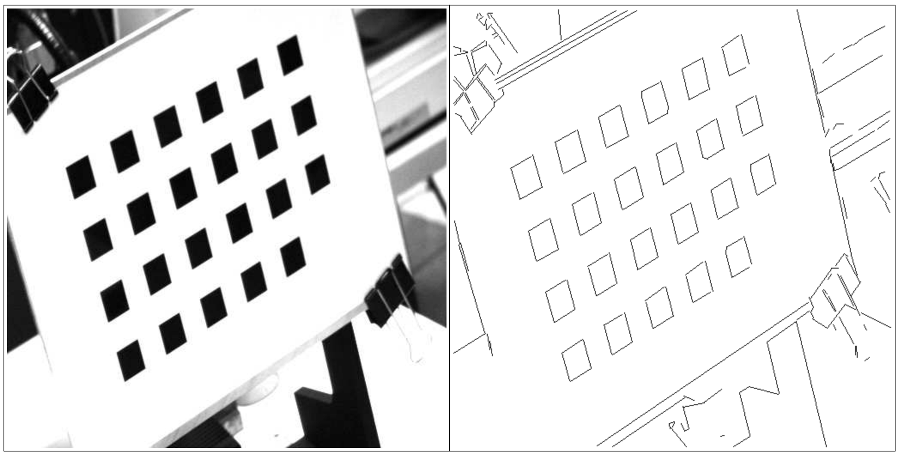

Figure 5.7. *L'image de la mire et les segments de droite extraits de cette image.*

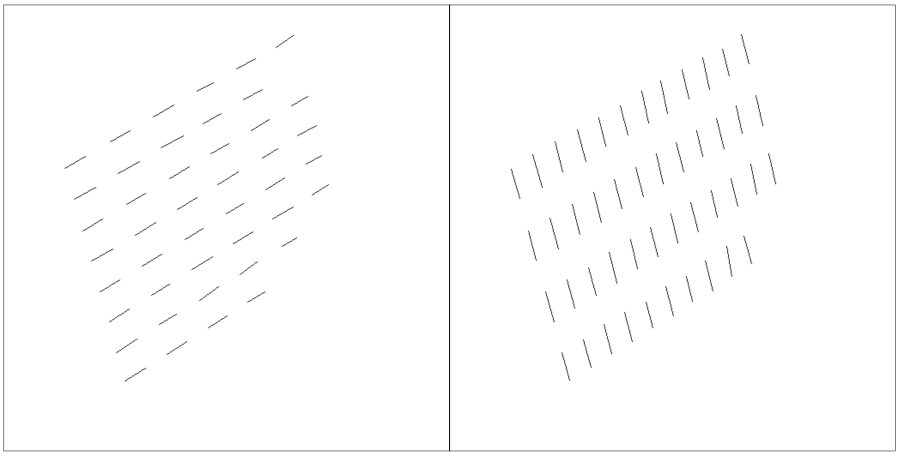

Figure 5.8. *Deux familles de droites parallèles détectées parmi l'ensemble des segments de droites extraits de l'image.*

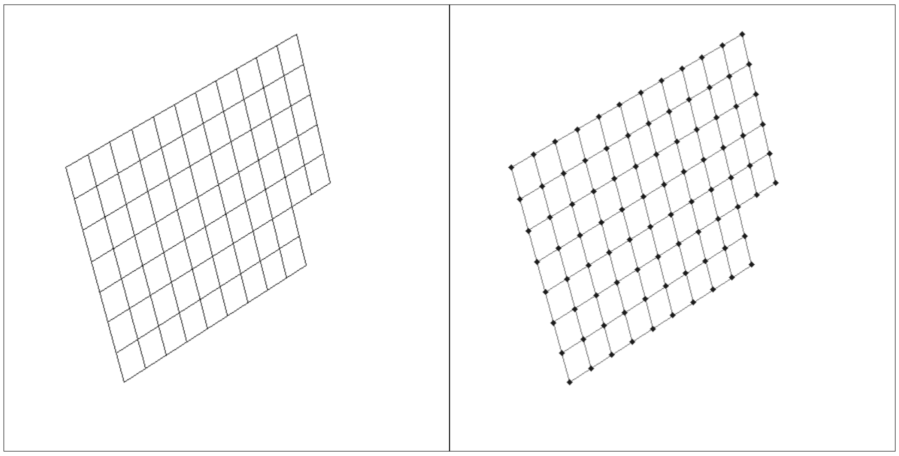

Figure 5.9. *Les sommets de cette grille sont les points image servant au calibrage. L'image de droite montre la superposition sur cette grille des points de la mire projetés grâce à la matrice de projection perspective obtenue par calibration.*

$$y = \frac{\mathbf{r}_2 \cdot \overrightarrow{O'B} + t_y}{\mathbf{r}_3 \cdot \overrightarrow{O'B} + t_z} \quad [5.28]$$

Divisons ces deux expressions par t_z. On peut écrire la projection sous la forme suivante :

$$x = \frac{s\mathbf{r}_1 \cdot \overrightarrow{O'B} + x'}{1 + \varepsilon} \quad [5.29]$$

$$y = \frac{s\mathbf{r}_2 \cdot \overrightarrow{O'B} + y'}{1 + \varepsilon} \quad [5.30]$$

- x' et y' sont les coordonnées de la projection de l'origine O$'$ du repère scène,
- $s = 1/t_z$ est un facteur d'echelle et
- $\varepsilon = \mathbf{r}_3 \cdot \overrightarrow{O'B}/t_z$ est un rapport qui est petit si l'objet est distant de la caméra ou si la taille de l'objet observé est petite par rapport à la distance à la caméra.

On peut donc obtenir un modèle affine de caméra en négligeant la valeur de

ε par rapport à 1 :

$$\frac{1}{1+\varepsilon} \approx 1$$

Avec cette approximation le modèle de caméra qu'on obtient est le suivant :

$$\begin{aligned} x - x' &= s\mathbf{r}_1 \cdot \overrightarrow{O'B} \\ y - y' &= s\mathbf{r}_2 \cdot \overrightarrow{O'B} \end{aligned}$$

On peut maintenant introduire à nouveau les coordonnées des points de l'image dans le repère image. On a :

$$\begin{pmatrix} u - u' \\ v - v' \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 \\ 0 & \alpha_v \end{pmatrix} \begin{pmatrix} x - x' \\ y - y' \end{pmatrix}$$

En substituant dans l'équation précédente on obtient successivement :

$$\begin{pmatrix} u - u' \\ v - v' \end{pmatrix} = \begin{pmatrix} s\alpha_u & 0 \\ 0 & s\alpha_v \end{pmatrix} \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad [5.31]$$

$$= \begin{pmatrix} a_u & 0 \\ 0 & a_v \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad [5.32]$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad [5.33]$$

$$= \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad [5.34]$$

Les paramètres intrinsèques d'une caméra affine sont a_u et a_v qui ont la même signification géométrique (à un facteur d'échelle près) que α_u et α_v dans le cas d'une caméra projective. On peut remarquer, que dans le cas d'une caméra affine on ne peut plus parler de centre de projection. Les paramétres extrinsèques sont constitués par une matrice de rotation.

Pour calibrer une telle caméra il suffit de disposer d'une mire de calibrage suffisamment loin de la caméra pour annuler les effets de perspective, et de choisir un des point de la mire comme origine du repère de la scène. Pour une correspondance mire/image i on obtient les deux équations suivantes :

$$\begin{pmatrix} X_i & Y_i & Z_i & 0 & 0 & 0 \\ 0 & 0 & 0 & X_i & Y_i & Z_i \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{pmatrix} = \begin{pmatrix} u_i - u' \\ v_i - v' \end{pmatrix} \qquad [5.35]$$

On a donc besoin, au minimum, de 4 points de correspondance (3 points courants et l'origine) pour calibrer une caméra affine. On obtient facilement les paramètres intrinsèques et extrinsèques de cette caméra :

$$\begin{aligned} a_u &= ||\mathbf{a}_1|| \\ a_v &= ||\mathbf{a}_2|| \\ \mathbf{r}_1 &= \mathbf{a}_1/||\mathbf{a}_1|| \\ \mathbf{r}_2 &= \mathbf{a}_2/||\mathbf{a}_2|| \\ \mathbf{r}_3 &= \mathbf{r}_1 \wedge \mathbf{r}_2 \end{aligned}$$

Remarquons pour finir que le modèle affine peut également s'écrire sous la forme d'une matrice 3×4 :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & u' \\ a_{21} & a_{22} & a_{23} & v' \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

5.7. Caméra linéaire (barette CCD)

Dans beaucoup d'applications de la vision par ordinateur les caméras linéaires (barettes CCD) peuvent remplacer les caméras matricielles qu'on vient d'étudier. L'inspection utilise abondamment ce type de capteurs car on a en même temps besoin de précision et de vitesse de calcul. Le signal fourni par une caméra linéaire est plus simple donc plus rapide à traiter que les images 2D. De plus, une barette CCD peut avoir jusqu'à 4096 pixels à un prix raisonnable.

Les techniques de calibrage des caméras matricielles qu'on vient de décrire sont nécessaires mais pas suffisantes pour calibrer une caméra linéaire [93], [94].

Dans cette section nous proposons une solution en deux étapes pour calibrer une telle caméra. La première étape utilise les mêmes techniques que les caméras matricielles. La deuxième étape utilise un objet de calibrage spécialement conçu à cet effet en combinaison avec le birapport, un invariant projectif qui a déjà été utilisé en vision, [126].

La meilleure façon d'imaginer une caméra linéaire est de penser à une caméra matricielle pour laquelle une seule ligne de pixels est active. Un point de l'espace se projette sur cette ligne suivant l'équation [5.16]. Par ailleurs, le point de l'espace se projetant sur la ligne image est contraint d'appartenir dans un *plan de vue*, soit le plan défini par le centre de projection et par la ligne de pixels, figure 5.10. L'équation du plan de vue peut être écrite comme suit :

$$X = pY + qZ + r \qquad [5.36]$$

En substituant l'équation [5.36] dans l'équation [5.16], en renommant les

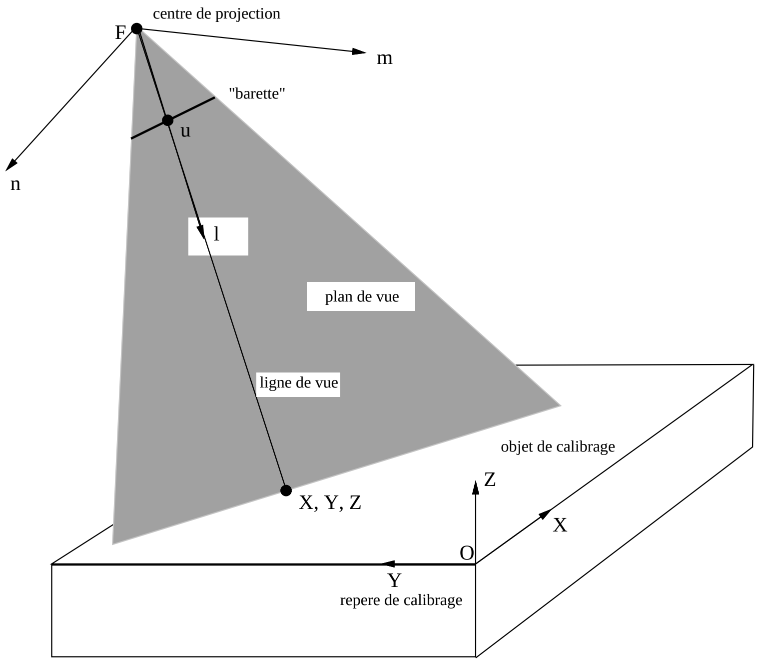

Figure 5.10. *Une vue générale d'une caméra linéaire, un objet de calibrage et les systèmes de coordonnées qui leur sont associés.*

variables et en notant que l'équation [5.16] est définie à un facteur multiplicatif près, on obtient :

$$u = \frac{n_1 Y + n_2 Z + n_3}{n_4 Y + n_5 Z + 1} \qquad [5.37]$$

Les équations [5.36] et [5.37] définissent le modèle de la caméra linéaire, autrement dit elles décrivent l'équation de la droite passant par le centre de projection et un pixel, dans le repère de calibrage. Cette droite est la *droite de vue*.

Ce modèle a 8 paramètres : n_1, n_2, n_3, n_4, n_5, et p, q, r. Le problème de calibrage de la caméra linéaire est donc le problème d'estimation de ces paramètres. Ce problème peut donc se décomposer en deux étapes :

- estimer les paramètres n_1, n_2, n_3, n_4, n_5. Si au moins 5 correspondances sont disponibles, ce problème est un problème d'optimisation linéaire similaire au cas de la calibration matricielle.

- estimer p, q, r. Ces paramètres peuvent être estimés si et seulement si on possède des points de l'espace appartenant au plan de vue.

5.7.1. Calibration en deux étapes

On commence par décrire la structure de l'objet de calibrage utilisé. La raison du choix de cette structure sera justifiée plus bas. L'objet de calibrage est montré sur la figure 5.11. Il consiste en quatre droites coplanaires, D_1, D_2, D_3, D_4. Les trois premières sont parallèles entre elles et la quatrième fait un angle aigu avec la direction des trois autres. Les équations de ces droites sont données dans le repère de calibrage qui est défini comme suit : l'axe des X est confondu avec D_1 et l'axe des Y est perpendiculaire à D_1. L'origine se trouve quelque part le long de D_1. L'axe des Z est perpendiculaire au plan contenant l'objet de calibrage. Dans ce repère et dans le plan $Z = 0$ les équations des quatre droites s'écrivent :

$$\begin{array}{rcl} Y = & 0 & (D_1) \\ Y = & \alpha & (D_2) \\ Y = & \beta & (D_3) \\ Y = & \gamma X + \delta & (D_4) \end{array}$$

les paramètres α, β, γ, et δ sont fixés et ils déterminent la structure de l'objet de calibrage.

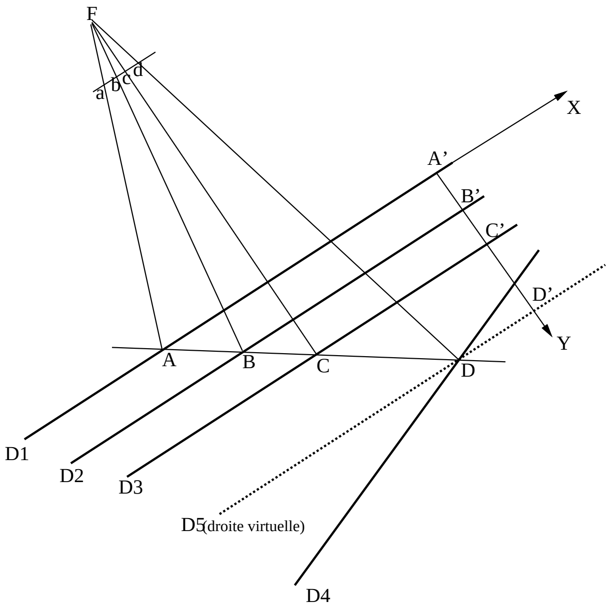

Figure 5.11. *La structure de l'objet de calibrage et la façon dont il est observé par la caméra linéaire.*

Lorsque la caméra linéaire "observe" cet objet elle "voit" quatre points, a, b, c, d qui sont les projections de A, B, C, D soient les intersections du plan de vue avec les quatre droites D_1, D_2, D_3, D_4, figure 5.11. Remarquez que l'on ne doit pas connaître les positions de ces points le long des droites formant l'objet. Quelle que soit l'orientation du plan de vue par rapport à l'objet de calibrage, les coordonnées Y et Z des points A, B, C sont connues dans le repère de calibrage. On peut donc établir des correspondances du type $\{Y_A, Z_A, u_a\}$, $\{Y_B, Z_B, u_b\}$ et $\{Y_C, Z_C, u_c\}$ pour chaque position de l'objet par rapport à la caméra. En faisant bouger l'objet de calibrage dans les directions Y et Z on peut établir trois nouvelles correspondances dans chacune des nouvelles positions. Chaque correspondance i vérifie l'équation [5.37] qui peut être écrite sous la forme suivante :

$$Y_i n_1 + Z_i n_2 + n_3 - u_i Y_i n_4 - u_i Z_i n_5 = u_i \qquad [5.38]$$

Pour n telles correspondances on obtient ($n \geq 5$) :

$$\begin{pmatrix} Y_1 & Z_1 & 1 & -u_1 Y_1 & -u_1 Z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_n & Z_n & 1 & -u_n Y_n & -u_n Z_n \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \qquad [5.39]$$

Il s'agit d'un système formé de n équations linéaires en 5 inconnues de la forme : $AX = B$. La solution est du même type que l'équation [5.21].

Afin d'estimer p, q, et r on doit avoir à notre disposition un ensemble de points appartenant au plan de vue dont on connaît les coordonnées dans le repère de calibrage. La structure de l'objet de calibrage est telle que l'on peut calculer pour des points de l'espace se trouvant sur cet objet et dans le plan de vue leurs coordonnées 3D dans le repère de calibrage.

Rapelons la définition du birapport de quatre points colinéaires :

$$\begin{aligned} \text{birapport}_{A,B,C,D} &= (A, B, C, D) \\ &= (CA/CB)/(DA/DB) \end{aligned}$$

CA étant la distance algébrique de C à A. La propriété fondamentale du birapport est qu'il est invariant par projection perspective ou orthographique. On a donc :

$$(A, B, C, D) = (a, b, c, d)$$

Considérons une droite *virtuelle* D_5 parallèle et coplanaire avec les droites D_1, D_2, D_3 et passant par le point D. Ces quatre droites parallèles (D_1, D_2, D_3 et D_5) intersectent l'axe des Y en A', B', C' et D'. Compte tenu de la propriété d'invariance du bi-rapport on a :

$$(A, B, C, D) = (A', B', C', D')$$

et on obtient :

$$(C'A'/C'B')/(D'A'/D'B') = (ca/cb)/(da/db)$$

Si la caméra linéaire observe simultanément D_1, D_2, D_3 et D_4 (en réalité la caméra ne voit que les points A, B, C et D) on peut alors calculer le bi-rapport r, $r = (a, b, c, d)$ dans l'image et on peut alors déterminer la position du point

D' le long de l'axe Y, soit $Y = \lambda$. En effet on a :

$$\begin{aligned} r &= (C'A'/C'B')/(D'A'/D'B') \\ &= \left(\frac{\beta}{\beta-\alpha}\right) / \left(\frac{\lambda}{\lambda-\alpha}\right) \end{aligned}$$

et on obtient :

$$\lambda = \frac{\alpha\beta}{r\alpha + (1-r)\beta}$$

Les coordonnées du point D seront donc données par l'intersection des droites D_4 et D_5 (qui est une droite virtuelle) :

$$\begin{aligned} Y &= \gamma X + \delta \\ Y &= \lambda \\ Z &= 0 \end{aligned}$$

Le point D appartient forcément au plan de vue et sa position dans le repère de calibrage peut être calculée à partir de mesures images mais sans calibrer la caméra. Encore une fois, en déplaçant l'objet de calibrage dans les directions Y et Z on peut déterminer un ensemble de points appartenant tous au plan de vue, chaque point j de cet ensemble vérifiant l'équation du plan de vue :

$$Y_j p + Z_j q + r = X_j$$

En combinant ces équations pour k points on obtient :

$$\begin{pmatrix} Y_1 & Z_1 & 1 \\ \vdots & \vdots & \vdots \\ Y_k & Z_k & 1 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \qquad [5.40]$$

Il s'agit une fois de plus d'un système d'équations linéaires à trois inconnues.

5.8. Capteurs stéréoscopiques passifs

Nous sommes maintenant en mesure d'étudier un capteur stéréoscopique. Il s'agit tout simplement d'utiliser deux caméras qui "observent" la même scène. On récupère ainsi deux projections de chaque point de la scène. Considérons le schéma de la figure 5.12. qui représente deux caméras. A chaque caméra est associé un repère. Soit P un point de la scène et soient p et p′ les deux projections dans les images de gauche et de droite. Nous pouvons donc écrire l'équation de la droite passant par le centre focal (F) de la caméra de gauche et le point p. De même on peut écrire l'équation de la droite passant par F′ et p′. Le point de la scène P se trouve à l'intersection de ces droites. Afin de pouvoir calculer cette intersection et donc de déterminer la position de P il faut pouvoir exprimer les deux équations (de la droite Fp et de la droite F′p′) *dans le même repère.*

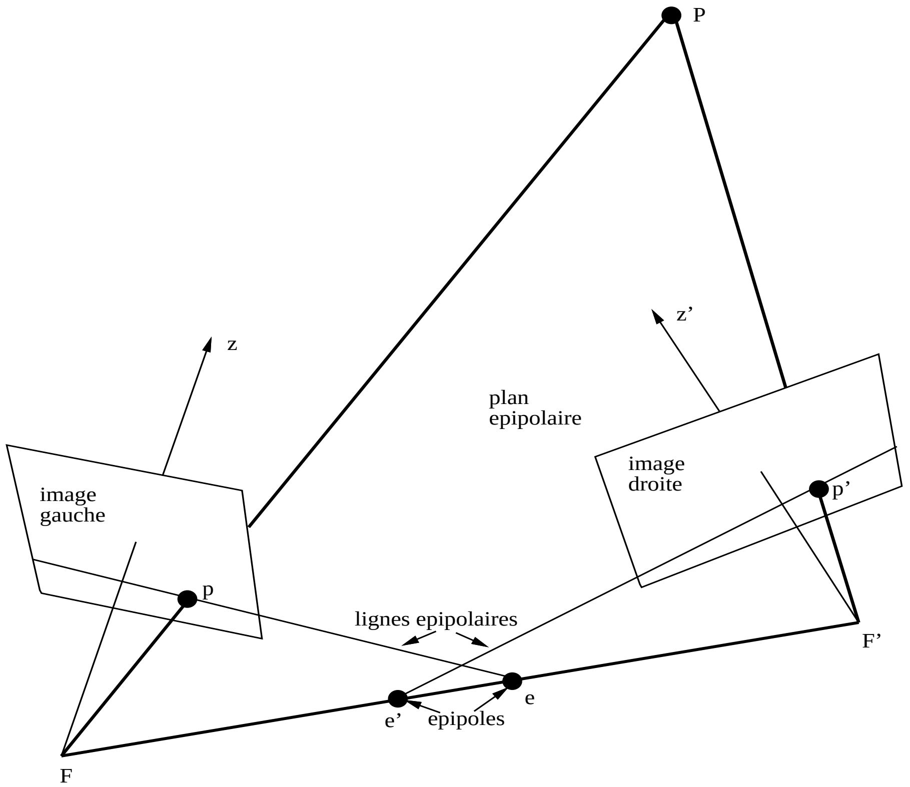

Figure 5.12. *La géométrie d'un capteur stéréoscopique passif.*

5.8.1. Calibrage stéréoscopique

Le calibrage stéréoscopique consiste donc à déterminer la matrice de transformation entre le repère caméra gauche et le repère caméra droite. Soit A_s

cette matrice qui est une matrice composée d'une matrice de rotation et une matrice de translation du même type que la matrice A de l'équation [5.7]. Cette matrice est également représentée sur la figure 5.13. Cette figure suggère également la démarche à suivre pour calibrer un tel capteur, soit déterminer A_s :

- on commence par calibrer chaque caméra par rapport à une mire unique. Ceci nous fournit les coefficients de deux matrices, $\mathbf{M}$ et $\mathbf{M}'$.
- on extrait ensuite les paramètres intrinsèques et extrinsèques de chaque caméra.
- à l'aide des paramètres extrinsèques on construit deux matrices, $\mathbf{A}$ et $\mathbf{A}'$, la première étant la transformation du repère mire au repère de la caméra gauche, la deuxième étant la transformation du repère mire au repère de la caméra droite.
- on calcule ensuite $\mathbf{A}_s$:

$$\mathbf{A}_s = \mathbf{A}'\mathbf{A}^{-1}$$

Cette matrice s'écrit comme auparavant :

$$\mathbf{A}_s = \begin{pmatrix} r_{11} & r_{12} & r_{13} & b_x \\ r_{21} & r_{22} & r_{23} & b_y \\ r_{31} & r_{32} & r_{33} & b_z \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1 & b_x \\ \mathbf{r}_2 & b_y \\ \mathbf{r}_3 & b_z \\ \mathbf{0} & 1 \end{pmatrix}$$

Le vecteur $\mathbf{b} = (b_x b_y b_z)^t$ est le vecteur allant de F$'$ à F. Il représente les coordonnées de F dans le repère de la caméra de droite, figure 5.12.

5.8.2. Relation droite-gauche

Nous allons maintenant établir une relation simple entre un point de l'image de gauche et un point de l'image de droite. Dans ce qui suit nous allons nous placer dans le repère caméra et non pas dans le repère image. Rappelons qu'à partir d'un pixel image de coordonnées u et v on peut facilement déduire ses coordonnées caméra (équation [5.6]) :

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{C}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \qquad [5.41]$$

et on a une expression similaire pour la caméra droite :

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{C}'^{-1} \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix}$$

Exprimons maintenant un point P de la scène à la fois dans les deux repères. Soient (X, Y, Z) ses coordonnées dans le repère gauche et (X', Y', Z') ses coordonnées dans le repère droit. Les deux systèmes de coordonnées sont reliés par la formule :

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = \mathbf{A}_s \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Les coordonnées de p, la projection de P dans l'image de gauche sont $(x, y, 1)$ où $x = X/Z$ et $y = Y/Z$. On a également p$'$, la projection dans l'image de droite de coordonnées $(x', y', 1)$. x' et y' peuvent s'écrire :

$$x' = \frac{X'}{Z'} = \frac{r_{11}X + r_{12}Y + r_{13}Z + b_x}{r_{31}X + r_{32}Y + r_{33}Z + b_z}$$

$$y' = \frac{Y'}{Z'} = \frac{r_{21}X + r_{22}Y + r_{23}Z + b_y}{r_{31}X + r_{32}Y + r_{33}Z + b_z}$$

En remarquant que $X = xZ$ et $Y = yZ$ et avec la notation $\mathbf{p} = (x\ y\ 1)^t$ on peut simplifier ces formules :

$$x' = \frac{Z\,\mathbf{r}_1 \cdot \mathbf{p} + b_x}{Z\,\mathbf{r}_3 \cdot \mathbf{p} + b_z} \qquad [5.42]$$

$$y' = \frac{Z\,\mathbf{r}_2 \cdot \mathbf{p} + b_y}{Z\,\mathbf{r}_3 \cdot \mathbf{p} + b_z} \qquad [5.43]$$

On peut donc exprimer la position d'un point de l'image de droite en fonction de *son correspondant* dans l'image de gauche, des *paramètres du capteur* et de la *profondeur* Z du point P.

5.8.3. La contrainte épipolaire

En éliminant Z entre les équations [5.42] et [5.43] on obtient une relation linéaire entre x' et y' :

$$(b_z\,\mathbf{r}_2 \cdot \mathbf{p} - b_y\,\mathbf{r}_3 \cdot \mathbf{p})\,x' + (b_x\,\mathbf{r}_3 \cdot \mathbf{p} - b_z\,\mathbf{r}_1 \cdot \mathbf{p})\,y' = b_x\,\mathbf{r}_2 \cdot \mathbf{p} - b_y\,\mathbf{r}_1 \cdot \mathbf{p} \qquad [5.44]$$

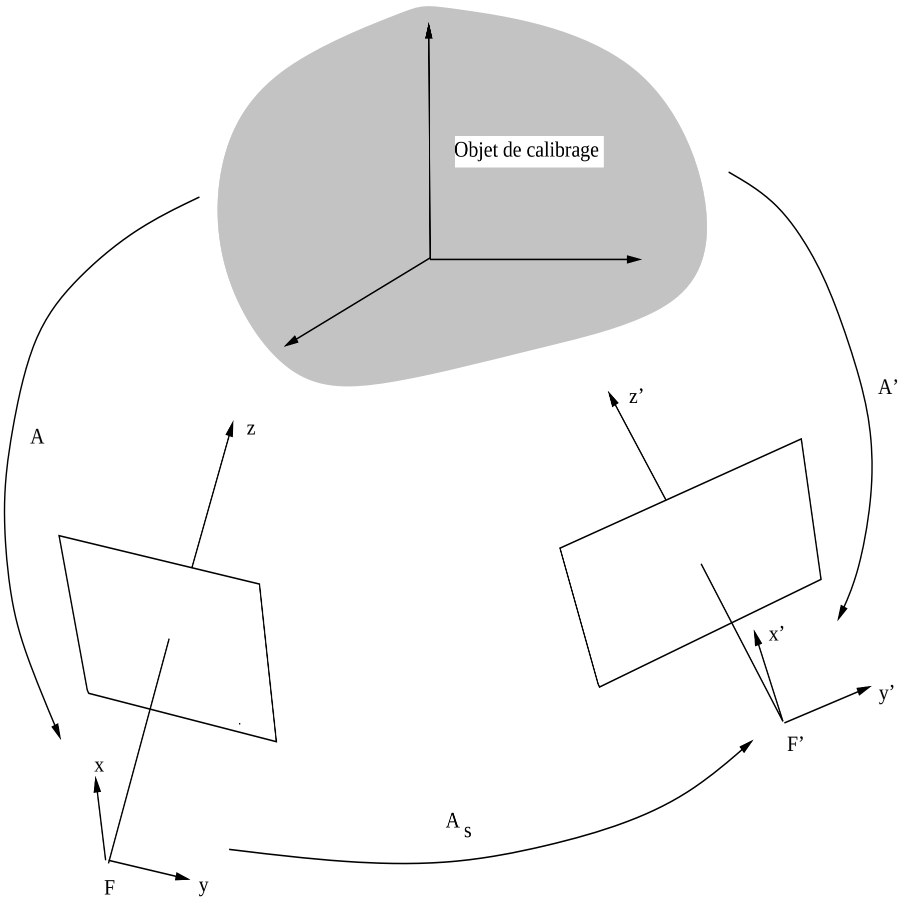

Figure 5.13. *Le calibrage de deux caméras formant un capteur stéréo à partir d'une seule mire.*

Cette équation décrit le lieu des points de l'image de droite pouvant correspondre à un point p de l'image de gauche : c'est une ligne qui s'appelle la ligne épipolaire droite. Il y a pour chaque point de l'image de gauche une telle ligne épipolaire (et réciproquement pour chaque point de l'image de droite il y a une ligne épipolaire gauche). Il est facile de remarquer que toutes les lignes épipolaires droites forment un faisceau. Le point commun de toutes ces lignes s'appelle l'épipôle droit et il s'obtient facilement en prenant $Z = 0$ dans les équations [5.42] et [5.43]. On obtient les coordonnées de l'épipôle droit, e$'$ dans le repère de droite :

$$\left\{ \begin{array}{lcl} x'_e & = & \frac{b_x}{b_z} \\ y'_e & = & \frac{b_y}{b_z} \end{array} \right.$$

L'interprétation géométrique est immédiate. En effet le centre de projection F de la caméra gauche a comme coordonnées dans le repère de droite $(b_x b_y b_z)^t$ (le vecteur translation de la transformation gauche-droite). L'épipôle droit n'est autre que la projection de F dans l'image de droite. On peut également définir un épipôle gauche. La figure 5.14. montre les deux faisceaux de lignes épipolaires.

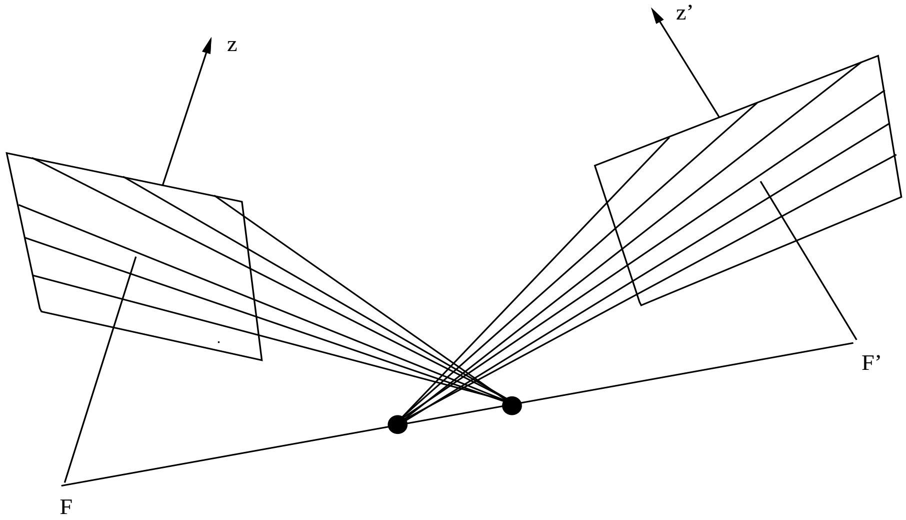

Figure 5.14. *Les deux faisceaux de lignes épipolaires et les deux épipôles.*

5.8.4. Rectification épipolaire

Il existe une configuration géométrique du capteur particulièrement intéressante : c'est lorsque les deux caméras sont disposées de telle façon que leurs axes optiques (z et z') sont parallèles et lorsque la droite FF' est confondue avec les axes horizontaux (y et y') des deux caméras. La matrice de transformation gauche/droite devient :

$$\mathbf{A}_b = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad [5.45]$$

Dans ce cas l'équation d'une ligne épipolaire droite devient tout simplement (en substituant dans l'équation [5.44]) :

$$x' = x$$

Autrement dit, dans ce cas les lignes épipolaires sont parallèles. A un point de l'image de gauche se trouvant sur la ligne x correspond la ligne x' de l'image de droite. Dans ce cas les deux épipôles se trouvent à l'infini. Cette configuration est illustrée sur la figure 5.15.

Les lignes épipolaires jouent un rôle fondamental en vision stéréoscopique, comme nous allons le voir au chapitre 6. En effet, lorsqu'on cherche pour un point de l'image de gauche un correspondant dans l'image de droite, on peut limiter cette recherche le long de la ligne épipolaire correspondante. Il est donc important d'avoir les expressions mathématiques les plus simples pour ces épipolaires. On ne peut pas toujours disposer les deux caméras de façon que leur matrice gauche-droite associée soit réduite à une translation le long de FF'.

Cependant on peut *rectifier* les deux images, soit faire subir à chaque image une transformation de façon à obtenir une paire d'images stéréo coplanaires et parallèles à la droite passant par les centres de projection des deux caméras, [8] [118]. La position des caméras pour des images rectifiées est telle que :

- les deux centres de projection (F et F') restent inchangés. De cette façon on pourra reconstruire la scène à partir des images rectifiées,

- les axes $\mathrm{F}y$ et $\mathrm{F}'y'$ des caméras dans la position rectifiée sont parallèles et confondus avec la droite FF'.

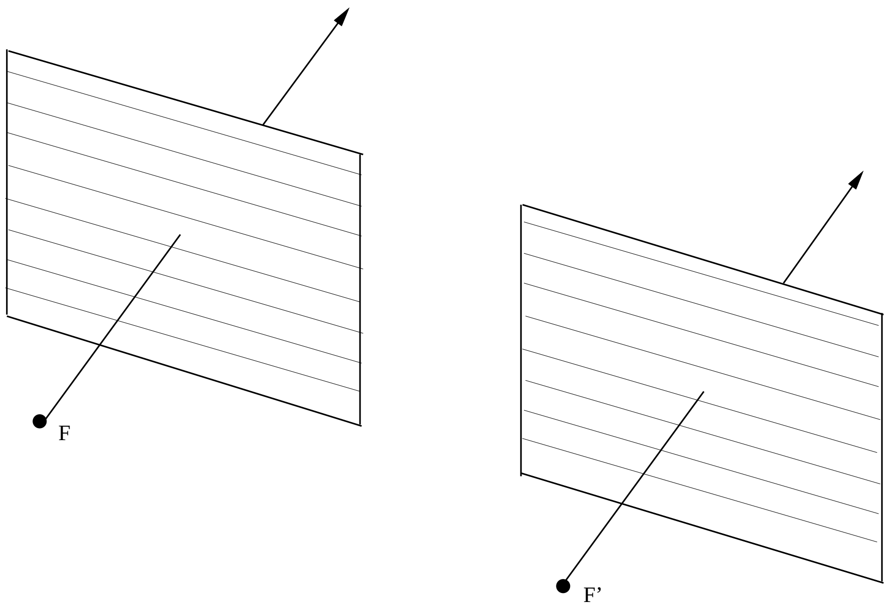

Figure 5.15. *La rectification épipolaire ramène le capteur stéréo dans une configuration telle que les lignes épipolaires sont parallèles aux axes Fy et $F'y'$. Dans cette configuration les deux épipôles sont rejetés à l'infini.*

La figure 5.16. illustre le principe de la rectification épipolaire. Un point P de la scène se projette en p et p′ sur les images initiales et en q et q′ sur les images rectifiées. Comme les centres de projections restent les mêmes, on peut reconstruire le point P, soit à partir de p et p′, soit à partir de q et q′. On peut noter que les paramètres intrinsèques de la projection ne sont pas modifiés. Ce qui change est la relation spatiale entre le repère de calibrage et la caméra. La transformation entre les deux positions (position initiale d'une caméra et position rectifiée de la même caméra) est donc une rotation dans l'espace.

Afin de rectifier une paire d'images stéréo, il faut donc appliquer à chacune des caméras une transformation spatiale. Les centres de projection restant inchangés, cette transformation est une rotation, figure 5.17.

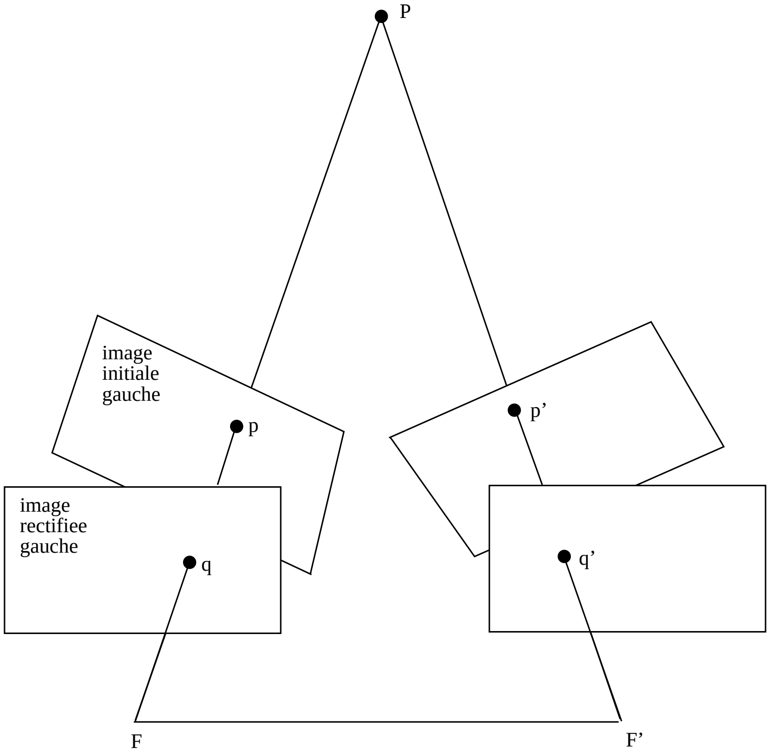

Figure 5.16. *Le principe de rectification épipolaire.*

Si on se place dans le repère de calibrage, soit le repère $OXYZ$ de la figure 5.17., la calibration du capteur stéréo nous fournit les transformations

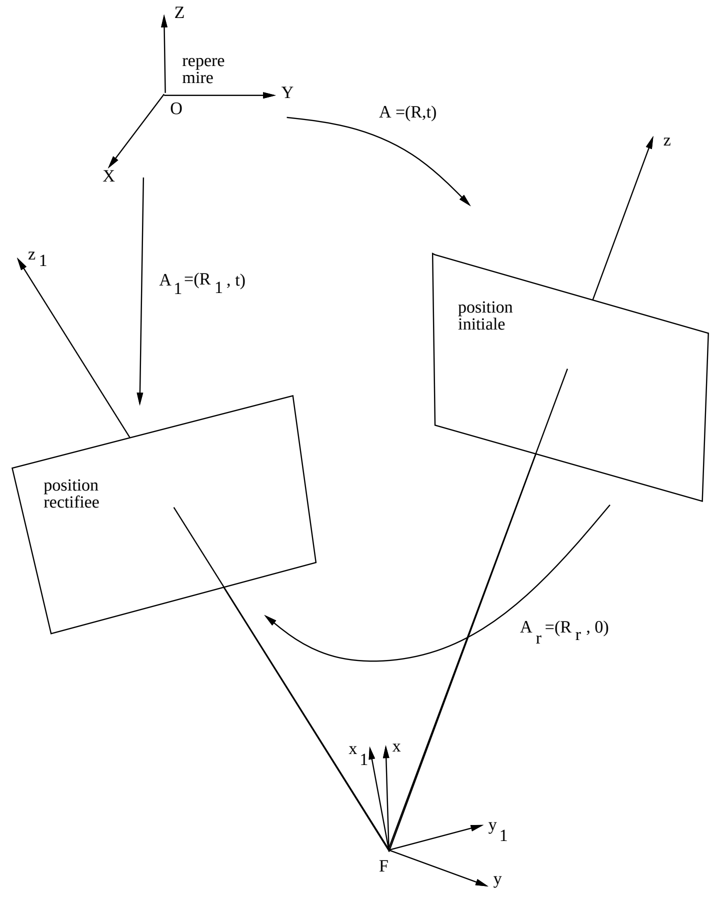

Figure 5.17. *Une image rectifiée s'obtient en appliquant à la caméra correspondante une rotation spatiale.*

$\mathbf{A}$ et $\mathbf{A}'$ du repère de calibrage vers les repères des caméras gauche et droite respectivement. Chacune de ses transformations se compose d'une rotation et d'une translation :

$$\mathbf{A} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}$$

$$\mathbf{A}' = \begin{pmatrix} \mathbf{R}\,' & \mathbf{t}' \\ \mathbf{0} & 1 \end{pmatrix}$$

On cherche à déterminer la transformation du repère de calibrage au repère de la caméra gauche dans sa position rectifiée. Soit $\mathrm{F}x_1y_1z_1$ le repère de la caméra gauche dans sa position rectifiée.

Les coordonnées de F dans le repère de calibrage s'obtiennent en remarquant que F est l'origine du repère de la caméra de gauche. On aura donc :

$$\mathbf{R}\,\overrightarrow{OF} + \mathbf{t} = 0$$

ainsi qu'une expression similaire pour la caméra de droite :

$$\mathbf{R}\,'\overrightarrow{OF}' + \mathbf{t}' = 0$$

et on obtient ainsi la direction de la droite FF$'$:

$$\begin{aligned} \overrightarrow{FF'} &= \overrightarrow{OF'} - \overrightarrow{OF} \\ &= \mathbf{R}\,^{-1}\mathbf{t} - \mathbf{R}\,'^{-1}\mathbf{t}' \end{aligned}$$

Cette droite n'est pas modifiée lors de la rectification. Choisissons les axes des caméras rectifiées. Pour la caméra de gauche ces axes sont $\mathrm{F}x_1$, $\mathrm{F}y_1$ et $\mathrm{F}z_1$. L'axe $\mathrm{F}y_1$ doit être parallèle à la droite FF$'$. On obtient donc pour le vecteur directeur de cet axe :

$$\mathbf{j}_1 = \frac{\overrightarrow{FF'}}{||\overrightarrow{FF'}||}$$

Les deux autres axes doivent être dans un plan perpendiculaire au vecteur $\mathbf{j}_1$. On peut par exemple choisir un de ces axes, soit l'axe optique, dans un plan contenant O (l'origine du repère de calibrage), F et F$'$. Le vecteur directeur de l'axe optique sera :

$$\mathbf{k}_1 = \frac{\mathbf{K}_1}{||\mathbf{K}_1||}$$

avec :

$$\mathbf{K}_1 = (\overrightarrow{OF} \wedge \overrightarrow{OF'}) \wedge \overrightarrow{FF'}$$

Le sens de $\mathbf{K}_1$ peut être déterminé par la contrainte :

$$\mathbf{K}_1 \cdot \mathbf{K} > 0$$

où $\mathbf{K}$est le vecteur directeur de l'axe optique de la caméra gauche dans sa position initiale. L'axe restant sera défini par :

$$\mathbf{i}_1 = \mathbf{j}_1 \wedge \mathbf{k}_1$$

La transformation du repère mire au repère de la caméra gauche rectifiée est donc :

$$\mathbf{A}_1 = \begin{pmatrix} \mathbf{R}_{\ 1} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}$$

avec :

$$\mathbf{R}_{\ 1} = \begin{pmatrix} \mathbf{i}_1 & \mathbf{j}_1 & \mathbf{k}_1 \end{pmatrix}$$

la translation étant la même que pour la transformation $\mathbf{A}$.

La rectification de l'image gauche est obtenue en appliquant une rotation au repère initial de la caméra gauche pour la ramener dans la position rectifiée :

$$\mathbf{R}_{\ r} = \mathbf{R}_{\ 1}\mathbf{R}^{\ -1}$$

Les coordonnées rectifiées s'obtiennent en appliquant aux coordonnées initiales cette rotation :

$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \mathbf{R}_{\ r} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Bien sûr, le même raisonnement s'applique à la caméra de droite. Les deux images rectifiées étant parallèles, la transformation de rectification de la caméra de droite est :

$$\mathbf{R}_{\ r}^{\ \prime} = \mathbf{R}_{\ 1}\mathbf{R}^{\ \prime -1}$$

Une fois que les deux images ont été rectifiées on a une géométrie épipolaire telle qu'elle a été décrite au début de ce paragraphe. La relation gauche/droite devient celle décrite par la transformation donnée par l'équation [5.45] :

$$\begin{aligned} \mathbf{A}_b &= \mathbf{A}_1'\mathbf{A}_1^{-1} \\ &= \begin{pmatrix} \mathbf{R}_{\ 1} & \mathbf{t}' \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_{\ 1} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \end{aligned}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

avec $b = (b_x^2 + b_y^2 + b_z^2)^{1/2}$.

5.9. Capteurs stéréoscopiques actifs

Nous pouvons combiner une caméra avec une source de lumière afin de mesurer les coordonnées tri-dimensionnelles de points sur la surface d'un objet. Un faisceau laser (lumière monochromatique) éclaire une scène, plus précisément éclaire une *tranche* de la scène qui correspond à l'intersection du plan lumineux et de la scène. La figure 5.18. illustre un capteur actif. Une caméra observe cette tranche de lumière qui se projette dans l'image. A un instant donné la seule information disponible dans l'image est la projection de cette tranche de lumière. Un point P de la scène se trouvant sur cette tranche se projette dans l'image en p.

Dans ce qui suit on va voir comment on peut calculer les coordonnées de P dans le repère mire connaissant :

- les coordonnées u et v de sa projection dans le plan image (après correction éventuelle de la distorsion) ;
- l'équation du plan de lumière dans le repère scène et
- la matrice de calibrage $\mathbf{M}$ de la caméra utilisée, en supposant que le même repère mire est utilisé pour calibrer la caméra et le capteur actif.

Comme nous allons le voir il est important de choisir le repère mire de façon que l'axe des Z soit perpendiculaire au plan de travail.

Le principe de calibrage qui sera décrit par la suite est dû à Bolles, Kremers et Cain, [23]. Les équations décrivant la droite passant par P, p et F sont (dans le repère mire) :

$$\begin{aligned} (m_{11} - um_{31})\,X + (m_{12} - um_{32})\,Y + (m_{13} - um_{33})\,Z &= -(m_{14} - um_{34}) \\ (m_{21} - vm_{31})\,X + (m_{22} - vm_{32})\,Y + (m_{23} - vm_{33})\,Z &= -(m_{24} - vm_{34}) \end{aligned}$$

Ce sont exactement les mêmes équations que les formules [5.16] et [5.17]. A ces équations on rajoute l'équation du plan de lumière (qu'on ne connaît pas

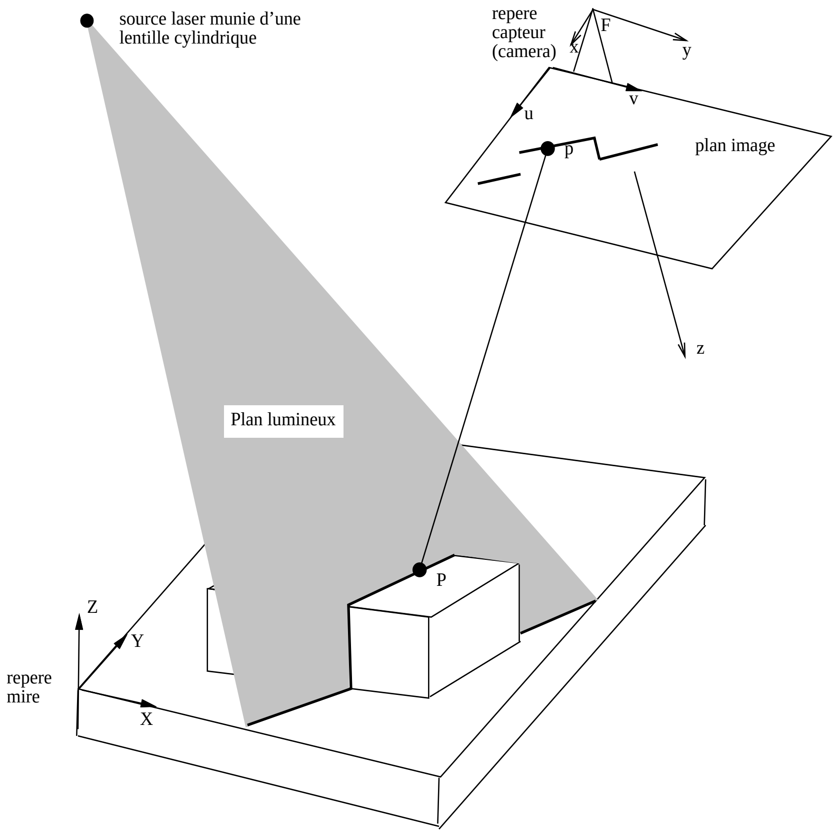

Figure 5.18. *Un capteur actif muni d'une caméra et d'un plan de lumière.*

encore) :

$$b_1X + b_2Y + b_3Z + b_4 = 0 \qquad [5.46]$$

Ces trois équations peuvent se combiner en une notation matricielle.

$$\mathbf{N}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} -b_4 \\ -(m_{14} - um_{34}) \\ -(m_{24} - vm_{34}) \end{pmatrix}$$

où :

$$\mathbf{N} = \begin{pmatrix} b_1 & b_2 & b_3 \\ m_{11} - um_{31} & m_{12} - um_{32} & m_{13} - um_{33} \\ m_{21} - vm_{31} & m_{22} - vm_{32} & m_{23} - vm_{33} \end{pmatrix}$$

En inversant on peut donc calculer les coordonnées X, Y et Z de P. Cependant les calculs risquent d'être lourds car ils faut inverser une matrice 3×3 pour chaque point. Il existe une autre solution qui consiste à inverser la matrice 3×3 symboliquement (utilisant un système de calcul formel – *Maple*, *Mathematica*, etc.) et à la multiplier par le vecteur de droite. Le calcul de X, Y et Z peut alors se mettre sous la forme suivante :

$$\begin{pmatrix} sX \\ sY \\ sZ \\ s \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \qquad [5.47]$$

Ce qui est équivalent à :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{N}^{-1} \begin{pmatrix} -b_4 \\ -(m_{14} - um_{34}) \\ -(m_{24} - vm_{34}) \end{pmatrix}$$

La matrice $\mathbf{T} = (t_{ij})$ est la matrice de calibrage du capteur actif. Ses coefficients dépendent de la matrice de calibrage de la caméra et de l'équation du plan de lumière. Les coefficients de $\mathbf{T}$ s'obtiennent facilement par identification et sont (les m_{ij} ont été divisés par m_{34}) :

$$\begin{aligned} t_{11} &= (b_4m_{22} - b_2m_{24})m_{33} + (b_3m_{24} - b_4m_{23})m_{32} + (b_2m_{23} - b_3m_{22}) \\ t_{12} &= (b_2m_{14} - b_4m_{12})m_{33} + (b_4m_{13} - b_3m_{14})m_{32} + (b_3m_{12} - b_2m_{13}) \\ t_{13} &= (b_2m_{13} - b_3m_{12})m_{24} + (b_4m_{12} - b_2m_{14})m_{23} + (b_3m_{14} - b_4m_{13})m_{22} \end{aligned}$$

$$
\begin{aligned}
t_{21} &= (b_1 m_{24} - b_4 m_{21}) m_{33} + (b_4 m_{23} - b_3 m_{24}) m_{31} + (b_3 m_{21} - b_1 m_{23}) \\
t_{22} &= (b_4 m_{11} - b_1 m_{14}) m_{33} + (b_3 m_{14} - b_4 m_{13}) m_{31} + (b_1 m_{13} - b_3 m_{11}) \\
t_{23} &= (b_3 m_{11} - b_1 m_{13}) m_{24} + (b_1 m_{14} - b_4 m_{11}) m_{23} + (b_4 m_{13} - b_3 m_{14}) m_{21}
\end{aligned}
$$

$$
\begin{aligned}
t_{31} &= (b_4 m_{21} - b_1 m_{24}) m_{32} + (b_2 m_{24} - b_4 m_{22}) m_{31} + (b_1 m_{22} - b_2 m_{21}) \\
t_{32} &= (b_1 m_{14} - b_4 m_{11}) m_{32} + (b_4 m_{12} - b_2 m_{14}) m_{31} + (b_2 m_{11} - b_1 m_{12}) \\
t_{33} &= (b_1 m_{12} - b_2 m_{11}) m_{24} + (b_4 m_{11} - b_1 m_{14}) m_{22} + (b_2 m_{14} - b_4 m_{12}) m_{21}
\end{aligned}
$$

$$
\begin{aligned}
t_{41} &= (b_2 m_{21} - b_1 m_{22}) m_{33} + (b_1 m_{23} - b_3 m_{21}) m_{32} + (b_3 m_{22} - b_2 m_{23}) m_{31} \\
t_{42} &= (b_1 m_{12} - b_2 m_{11}) m_{33} + (b_3 m_{11} - b_1 m_{13}) m_{32} + (b_2 m_{13} - b_3 m_{12}) m_{31} \\
t_{43} &= (b_2 m_{11} - b_1 m_{12}) m_{23} + (b_1 m_{13} - b_3 m_{11}) m_{22} + (b_3 m_{12} - b_2 m_{13}) m_{21}
\end{aligned}
$$

5.9.1. Calibrage d'un capteur actif

La calibration du capteur actif consiste donc à calculer les valeurs numériques des coefficients de $\mathbf{T}$. Or nous avons déjà calculé les coefficients m_{ij}. Il nous reste à calculer l'équation du plan de lumière.

Pour cela on procède de la façon suivante. *On place devant le capteur des objets parallélipipédiques dont la hauteur est connue dans le repère de la mire de calibrage.* Un point P_i appartenant à une face de l'objet à la hauteur $Z = H_i$ a comme coordonnées mire X_i, Y_i et H_i et il se projette dans l'image en u_i, v_i. X_i et Y_i peuvent être déterminés grâce aux équations [5.16] et [5.17] qui peuvent s'écrire :

$$
\begin{aligned}
(m_{11} - u_i m_{31})\ X_i + (m_{12} - u_i m_{32})\ Y_i &= -(m_{13} - u_i m_{33})\ H_i + m_{14} - u_i m_{34} \\
(m_{21} - v_i m_{31})\ X_i + (m_{22} - v_i m_{32})\ Y_i &= -(m_{23} - v_i m_{33})\ H_i + m_{24} - v_i m_{34}
\end{aligned}
$$

Par ailleurs, le point P_i appartient au plan de lumière. En divisant par b_3 on obtient une relation linéaire entre b_1, b_2 et b_4 :

$$b_1 X_i + b_2 Y_i + b_4 = -H_i \qquad [5.48]$$

Pour n points appartenant à des faces à différentes hauteurs on obtient n

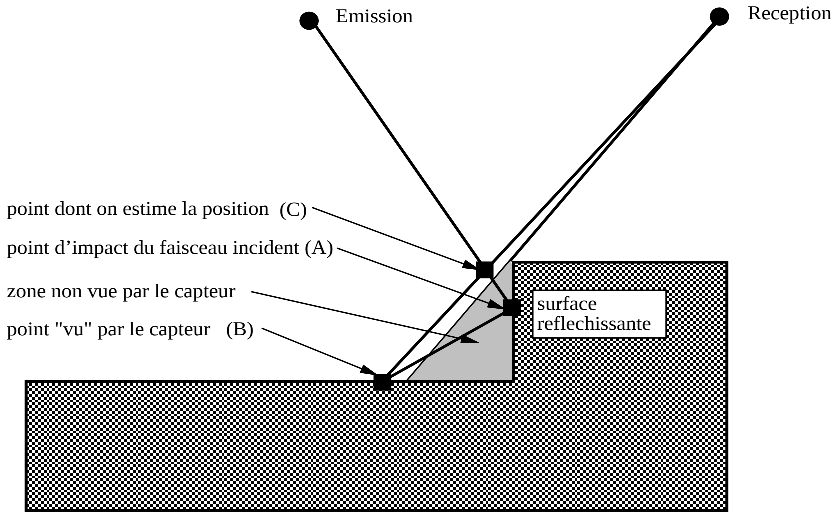

Figure 5.19. *Une configuration montrant un cas typique de la limite d'utilisation d'un capteur actif.*

équations qu'on peut écrire sous forme matricielle :

$$\begin{pmatrix} X_1 & Y_1 & 1 \\ \vdots & \vdots & \vdots \\ X_n & Y_n & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_4 \end{pmatrix} = \begin{pmatrix} -H_1 \\ \vdots \\ -H_n \end{pmatrix} \qquad [5.49]$$

Ceci peut s'écrire sous forme matricielle :

$$\mathbf{A}_{n\times 3}\mathbf{b}_3 = \mathbf{c}_n$$

Les paramètres du plan de lumière sont la solution aux moindres carrés de l'équation [5.49], soit :

$$\mathbf{b} = (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\mathbf{c}$$

5.9.2. Limites du capteur actif

L'avantage du capteur actif par rapport à un couple de caméras est évident : on évite le problème combinatoire de la mise en correspondance en "allumant" les points et en calculant directement leur position dans un repère tri-dimensionnel. Cependant ce capteur n'a pas que des avantages.

Le premier inconvénient est qu'on doit bouger soit l'objet soit le capteur lui-même pour pouvoir faire l'acquisition d'une carte dense $Z = f(X, Y)$ où f est un tableau rectangulaire correspondant à la projection au sol de la zone vue par le capteur.

Le deuxième inconvénient est difficilement contournable car il s'agit de propriétés intrinsèques des surfaces qui sont nuisibles à la bonne mesure. La figure 5.19. montre un objet dont la surface est réfléchissante : le plan de lumière éclaire un pan de l'objet qui ne peut être vu par la caméra. Cependant le rayon de lumière se réfléchit pour finalement éclairer un autre point de l'objet qui, lui, est vu par la caméra. Une inspection de la figure montre clairement que le point pour lequel on calcule la position ne se trouve pas sur la surface de l'objet. La figure permet également de comprendre pourquoi certains points d'un objet ne peuvent être mesurés : soit ils sont vus par la caméra mais ils ne sont pas éclairés par le plan lumineux, soit ils sont éclairés mais ne peuvent être vus.

Un troisième inconvénient est illustré sur la figure 5.20. En effet, le faisceau laser étant conique, il y a une imprécision de mesure lorsque le faisceau rencontre une discontinuité sur l'objet analysé.

Enfin, un quatrième inconvénient est illustré sur la figure 5.21. Lorsque le faisceau fait avec la normale de la surface observée un angle d'incidence proche de 90^0, l'énergie du faisceau réfléchissant qui repart vers la caméra est faible. Dans ce cas la mesure peut être entachée d'erreurs.

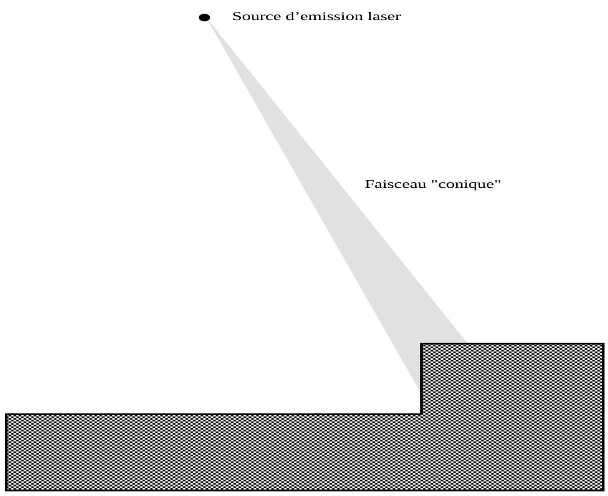

Figure 5.20. *Le faisceau laser étant conique, il y a des problèmes d'imprécision lorsque le faisceau rencontre une discontinuité de l'objet observé.*

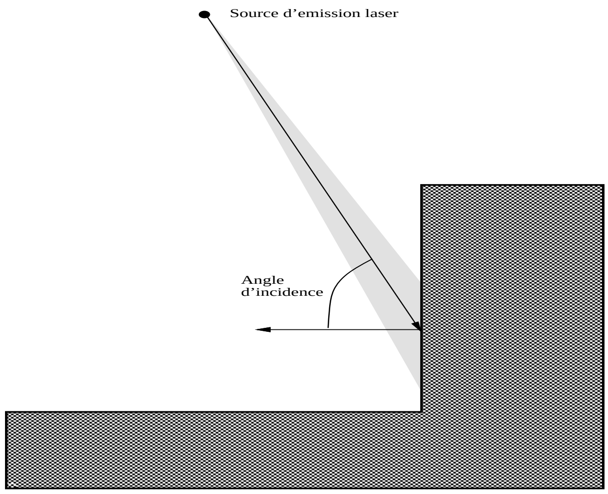

Figure 5.21. *Le faisceau laser a un angle d'incidence proche de* 90^0*, mesuré par rapport à la normale de la surface mesurée.*