

Ecole Nationale Supérieure des  
Télécommunications

Calibrage des systèmes de vision

Lasmea UMR 6602 CNRS

Jean-Marc LAVEST  
lavest@lasmea.univ-bpclermont.fr

15 Mai 2005



## Table des matières

<b>Avertissement</b> . . . . .	13
<b>Introduction</b> . . . . .	15
<b>PREMIÈRE PARTIE. ÉTALONNAGE DES CAPTEURS DE VISION</b> . . . . .	17
<b>Chapitre 1. Étalonnage des Capteurs de Vision</b> . . . . .	19
Jean-Marc LAVEST, Gérard RIVES	
1.1. Introduction . . . . .	19
1.2. Formulation générale du problème de calibrage . . . . .	20
1.2.1. Formulation du problème . . . . .	20
1.2.1.1. Modélisation de la caméra et de l'objectif : Le modèle sténopé . . . . .	21
1.2.1.2. Formation d'images : la projection perspective . . . . .	22
1.2.1.3. Changement de repère Objet/Caméra . . . . .	23
1.2.1.4. Changement de repère Caméra/Image . . . . .	23
1.2.1.5. Changement de coordonnées dans le plan image . . . . .	24
1.2.2. Hiérarchisation des modélisations de projection de caméra . . . . .	25
1.2.3. Caméra projective . . . . .	25
1.2.4. Caméra à l'infini . . . . .	26
1.2.4.1. Projection orthographique . . . . .	26
1.2.4.2. Projection orthographique à l'échelle . . . . .	26
1.2.4.3. Projection perspective faible . . . . .	27
1.2.4.4. Projection Affine . . . . .	27
1.2.5. Expression générale . . . . .	27
1.2.5.1. Formulation générale du problème de calibrage . . . . .	28
1.3. Approche Linéaire . . . . .	29
1.3.1. Principe . . . . .	29
1.3.2. Remarques et commentaires . . . . .	31
1.4. Approche non linéaire photogrammétrique . . . . .	32

1.4.1. Modèle Mathématique . . . . .	33
1.4.2. Résolution du problème . . . . .	35
1.4.3. Calibrage Multi-images . . . . .	37
1.4.4. Auto-calibrage par ajustement de faisceaux (Bundle Adjustment)	38
1.4.4.1. Redéfinition du problème : . . . . .	38
1.4.4.2. Estimation de la redondance . . . . .	39
1.4.4.3. Solution à un facteur d'échelle près . . . . .	39
1.4.4.4. Conditions initiales . . . . .	40
1.4.5. Calcul de la Précision . . . . .	40
1.5. Résultats d'expérimentation . . . . .	41
1.5.1. Ajustement de Faisceaux pour un Objectif Classique . . . . .	41
1.5.1.1. Conditions initiales et expérimentales . . . . .	41
1.5.1.2. Séquence d'images classiques . . . . .	42
1.5.2. Cas spécifique des objectifs Fish-Eye . . . . .	44
1.5.2.1. Critère classique . . . . .	44
1.5.2.2. Distorsion nulle à $r_0$ . . . . .	44
1.5.2.3. Normalisation des coefficients de distorsion . . . . .	45
1.5.2.4. Expérimentations . . . . .	46
1.5.3. Calibrage des caméras sous-marines . . . . .	49
1.5.3.1. Rappels théoriques . . . . .	49
1.5.3.2. Expérimentations . . . . .	49
1.5.3.3. Le matériel . . . . .	49
1.5.3.4. Résultats dans l'air . . . . .	50
1.5.3.5. Étalonnage dans l'eau . . . . .	51
1.5.3.6. Relations entre étalonnage dans l'air et dan l'eau . . . . .	52
1.5.4. Calibrage des zooms . . . . .	54
1.5.4.1. Rappels des propriétés optiques . . . . .	54
1.5.4.2. Estimation du point principal . . . . .	56
1.5.4.3. Expérimentations . . . . .	56
1.6. Facteurs perturbants de la mesure . . . . .	58
1.6.1. Un système avant tout électronique . . . . .	58
1.6.2. Choix d'une mire 2D ou 3D . . . . .	58
1.6.3. Configuration spatiales des vues . . . . .	58
1.7. Détection subpixelaire des amers . . . . .	58
1.7.1. Accurate Computation of a cross . . . . .	58
1.7.2. Détection sub-pixellique des amers visuels taches photoréfléchis-	
sante. . . . .	60
1.7.2.1. Déformation géométrique. . . . .	61
1.7.2.2. Modélisation de la transition. . . . .	61
1.7.2.3. Exemple de détection . . . . .	62
1.7.3. Influence de la précision de détection, sur les paramètres de cali-	
brage . . . . .	63
1.8. Bibliographie . . . . .	63

Table des matières 11

**Chapitre 2. Index** ..... 67



## Avertissement

Une large partie de ce cours s'inspire de l'ouvrage "Perception visuelle par imagerie vidéo" sous la direction de M.Dhome 2003, paru dans la collection *Traitement du signal et de l'image* chez Hermes.

Je tiens ici à remercier les co-auteurs : M.Dhome, J.T. Lapresté et G.Rives pour le travail que nous avons réalisé ensemble et leur aide à la rédaction de ce cours.



## Introduction

La *Vision Artificielle* est un domaine d'investigation relativement récent. Son objectif essentiel est la perception et l'interprétation automatique de l'univers observé par un système muni d'une ou plusieurs caméras. Cet axe de recherche débouche sur un spectre de problèmes étonnamment large, la majorité d'entre eux n'étant pas actuellement résolus de manière fiable. Toutefois, bien qu'une théorie générale soit loin de voir le jour, des progrès significatifs ont été réalisés tant au plan théorique que méthodologique, ces dernières années.

Dans le monde visible, les images de luminance sont le résultat de deux processus physiques : le premier est lié aux propriétés de réflectance de la surface des objets observés, le second correspond à la projection de ces mêmes objets sur la plaque photo-sensible du capteur utilisé. D'un point de vue mathématique, il s'agit, pour interpréter la scène observée, de résoudre un problème inverse, c'est à dire d'inférer la géométrie de la surface (3D) des objets présents, à partir du contenu purement bidimensionnel de l'image ou des images enregistrées.

A travers ce cours de MASTER nous allons aborder deux problématiques classiques en vision artificielle :

- le calibrage des capteurs vidéo.
- les algorithmes de calcul de pose 3D.

Ainsi le lecteur trouvera une première partie consacrée au problème de l'*étalonnage* des capteurs vidéo ; nous détaillerons une approche d'estimation fine des paramètres intrinsèques d'une caméra vidéo, fortement inspirée du monde de la photogrammétrie. Cette approche est fondée sur l'interprétation d'images d'une mire d'étalonnage, cette dernière n'étant pas nécessairement connue avec une grande précision.

La deuxième partie de cet ouvrage concerne essentiellement l'estimation de la *position relative objet/capteur* en introduisant des connaissances a priori (modèle CAO

de l'objet). Le lecteur découvrira comment traiter le problème de la localisation d'un objet rigide observé par un système monoculaire. Le formalisme présenté est ensuite étendu pour appréhender des cas aussi divers que la localisation multi-oculaire et la calibration bras-œil, en passant par la recherche de la posture d'objets articulés du type bras robotiques.

PREMIÈRE PARTIE

## Étalonnage des capteurs de Vision



## Chapitre 1

# Étalonnage des Capteurs de Vision

### 1.1. Introduction

Le calibrage d'un système de vision consiste à déterminer une relation mathématique existant entre les coordonnées tridimensionnelles (3D) des points d'une scène et les coordonnées bidimensionnelles (2D) de ces mêmes points projetés, détectés dans une image.

La détermination de cette relation est une étape incontournable en Vision, en particulier pour la reconstruction, où il est alors nécessaire d'inférer une information tridimensionnelle à partir de données extraites de l'image (2D). En réalité le champ d'application est plus large et le calibrage s'avère nécessaire dès lors qu'il faut établir une relation entre l'image et le monde 3D : reconnaissance et localisation d'objets, contrôle dimensionnel de pièces, reconstruction de l'environnement pour la navigation d'un robot mobile.

L'analyse complète du calibrage d'un système de vision doit prendre en compte l'ensemble des phénomènes photométrique, optique et électronique présents sur toute la chaîne d'acquisition des images.

En général, un système permettant de calibrer une caméra est composé par :

- une mire de calibrage (grille ou objet étalon) généralement constituée de points de référence dont on connaît parfaitement les coordonnées 3D dans un repère propre à celle-ci ;

---

Chapitre rédigé par Jean-Marc LAVEST, Gérard RIVES.

- un système d’acquisition d’images pour la numérisation et la mémorisation des images de la mire ;
- un algorithme de mise en correspondance des points 2D détectés dans les images avec leurs homologues de la mire ;
- Un algorithme de calcul de la matrice de transformation perspective de la caméra, liant le repère associé à la mire de calibrage à celui de l’image.

Le problème du calibrage des caméras CCD se décompose généralement en une étude de calibrage *géométrique* (calcul de la matrice de projection) et un calibrage *radiométrique* (uniformité de la luminance dans l’image). Le premier problème a donné lieu à une littérature dense, le second est quant à lui moins étudié.

A travers ce chapitre, nous nous attacherons à poser de manière didactique le problème du calibrage géométrique d’un capteur vidéo et à le résoudre par des méthodes linéaires puis non linéaires. La littérature fait souvent état de calibrage *faible* ou *fort* ; la distinction se situe au niveau d’une estimation globale de la matrice de projection (cas faible) ou d’une estimation de chaque paramètre qui compose cette matrice (cas fort). La suite de cet exposé traitera du problème de calibrage *fort* des capteurs de vision.

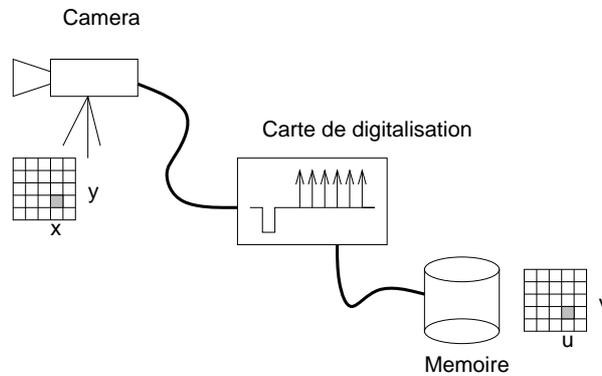
## 1.2. Formulation générale du problème de calibrage

Le processus de calibrage est un processus supervisé nécessitant souvent l’intervention de l’opérateur. C’est un processus statique qui s’effectue hors-ligne, avant l’utilisation de la caméra pour une tâche visuelle précise. Une fois la caméra calibrée, ses paramètres doivent rester fixes en cours d’utilisation. Chaque fois que l’on désire modifier la mise au point, la distance focale voire l’ouverture de l’objectif, la caméra devra être recalibrée. De très nombreux travaux, ces vingt dernières années, ont permis d’obtenir des méthodes plus ou moins complexes et donc plus ou moins précises permettant d’assimiler un système de vision à un ensemble métrologique. Le lecteur pourra se référer sur la base bibliographique suivante pour avoir une vue exhaustive des publications majeures dans le domaine : (<http://iris.usc.edu/Vision-Notes/bibliography>)

### 1.2.1. Formulation du problème

Soit un système d’acquisition d’images (figure 1.1).

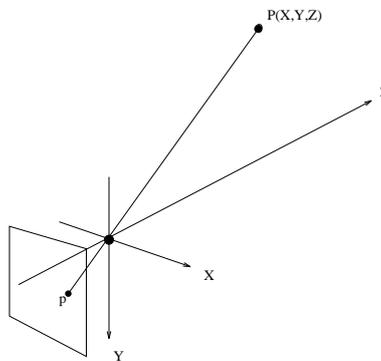
Calibrer le système revient à déterminer la transformation  $(R^3, R^2)$  qui permet d’exprimer sous une forme analytique le processus de formation d’images.



**Figure 1.1.** Système d'acquisition d'images.

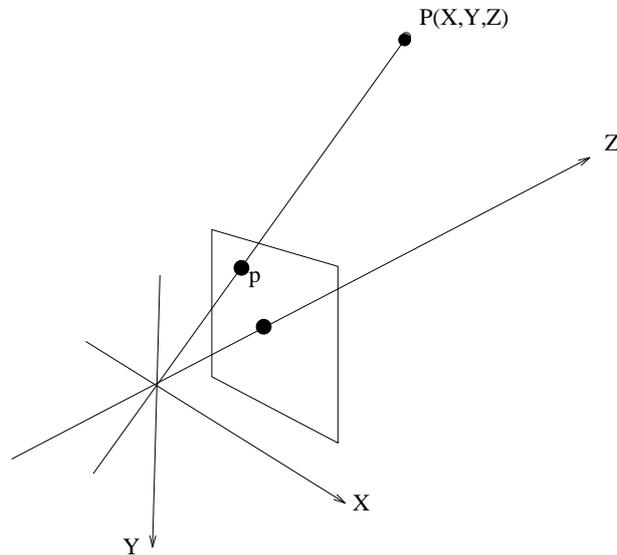
#### 1.2.1.1. Modélisation de la caméra et de l'objectif : Le modèle sténopé

L'optique géométrique classique voudrait qu'on utilise des modèles à lentilles épaisses ou à lentilles minces. La difficulté d'exprimer de façon simple les contraintes de vergence, nous pousse à recourir au modèle sténopé (*pin-hole model*) où tous les rayons passent par un seul et même point (le centre optique). La cellule photosensible (plan image) se situe à une distance  $f$  de ce centre et représente la distance focale ou tirage de l'objectif.



**Figure 1.2.** Géométrie d'un système de formation d'images

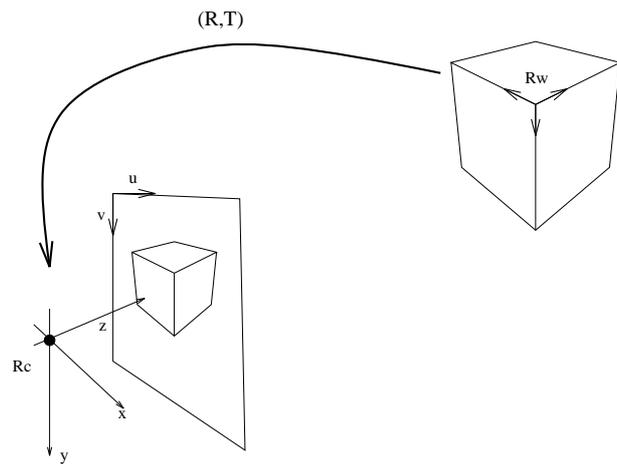
Notons que l'image obtenue est normalement inversée par rapport à la scène vue par l'œil. Pour pallier ce problème, on place artificiellement le plan image devant le centre optique (d'un point de vue physique, cet artifice est réalisé en lisant la matrice CCD de manière à obtenir l'inversion de l'image)



**Figure 1.3.** Plan image défini devant le centre optique

1.2.1.2. *Formation d'images : la projection perspective*

Dans la littérature, nous trouvons plusieurs types de projection d'image ; orthographique, orthographique à l'échelle, para-perspective et perspective. C'est cette dernière famille de projection qui retiendra notre attention pour sa meilleure adéquation à la réalité physique du capteur de vision.



**Figure 1.4.** Notation des différents repères

Dans ce paragraphe nous utiliserons les notations suivantes :

- $R_c$  : repère caméra (axe  $z$  confondu à l'axe optique) ;
- $R_w$  : repère lié à la modélisation de l'objet ;
- $(0, u, v)$  : repère Image (prenant en compte l'effet de digitalisation).

### 1.2.1.3. Changement de repère Objet/Caméra

Ce premier changement de repère permet d'exprimer les coordonnées de la mire ou de l'étalon, (positionné dans l'espace) dans le référentiel lié à la caméra. Comme on peut le noter, ce changement de repère s'exprime par une transformation composée d'une rotation et d'une translation.

$$P_c = (M_1) \cdot P_w \quad [1.1]$$

$$P_c = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad [1.2]$$

avec

- $P_w$  les coordonnées 3D d'un point de la mire, exprimées dans le repère de modélisation,
- $P_c$  les coordonnées du même point 3D exprimées dans le repère de la caméra.

Les matrices de rotation et de translation ( $R_{(3 \times 3)}$  et  $T_{(3 \times 1)}$ ) sont définies dans le repère de la *caméra*. Le couplage de  $(R, T)$  dans une même notation matricielle nécessite d'utiliser une notation en coordonnées homogènes.

### 1.2.1.4. Changement de repère Caméra/Image

Le passage du repère caméra au repère image est lié aux équations de projection perspective.

De manière classique, ces équations sont de la forme :

$$\begin{aligned} x &= X_c f / Z_c \\ y &= Y_c f / Z_c \\ z &= f \end{aligned} \quad [1.3]$$

En coordonnées homogènes, le système s'écrit :

$$\begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} P_c \quad [1.4]$$

Attention, les notations homogènes introduisent un facteur multiplicatif  $s$ , lors du passage  $(R^3, R^2)$  :

*Démonstration*

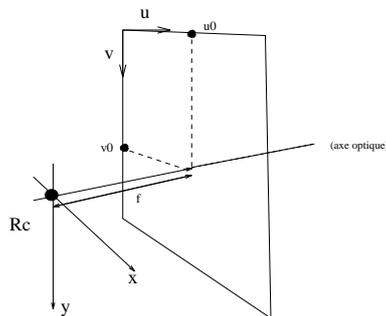
$$\begin{cases} sx = fX_c \\ sy = fY_c \\ s = Z_c \end{cases} \quad [1.5]$$

en substituant  $s$  :

$$\begin{cases} x = X_c f / Z_c \\ y = Y_c f / Z_c \end{cases} \quad [1.6]$$

#### 1.2.1.5. Changement de coordonnées dans le plan image

*Remarque* : Il est nécessaire d'inclure la différence de pas, selon les coordonnées  $x$  et  $y$ , relatives d'une part à la forme du pixel élémentaire sur la matrice CCD, d'autre part à la cadence d'échantillonnage du signal vidéo.



**Figure 1.5.** Système de coordonnées sur la matrice CCD

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} \quad [1.7]$$

soit encore, les équations classiques de changement de repère (Coordonnées Caméra / Coordonnées Pixel) :

$$\begin{cases} u = x/dx + u_0 \\ v = y/dy + v_0 \end{cases} \quad [1.8]$$

–  $(u_0, v_0)$  représentent les coordonnées (en pixel) dans l'image, de l'intersection de l'axe optique et du plan image (origine du changement de repère).

–  $(dx, dy)$  sont les dimensions respectivement selon  $x$  et  $y$  d'un pixel élémentaire de la matrice CCD. (voir figure 1.5).

### 1.2.2. Hiérarchisation des modélisations de projection de caméra

Ce paragraphe a pour but de représenter les différents modèles de projection que l'on retrouve dans la littérature. La stratification est sous-tendue à un certain nombre d'approximations et donc de simplifications du comportement réel inhérent aux rayons lumineux qui se propagent à travers les différentes lentilles des objectifs.

### 1.2.3. Caméra projective

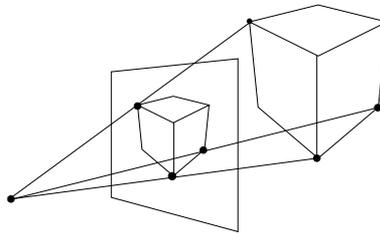


Figure 1.6. Caméra projective - Pin-hole

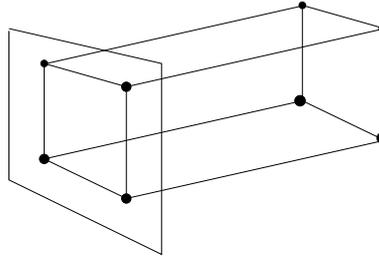
$$\begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad [1.9]$$

$$P_{Ima} = K [ I \quad , \quad 0 ] P_{Cam} \quad [1.10]$$

Tous les rayons lumineux se projettent dans l'image via un point unique, le centre optique.  $K$  représente la matrice de projection.

### 1.2.4. Caméra à l'infini

#### 1.2.4.1. Projection orthographique

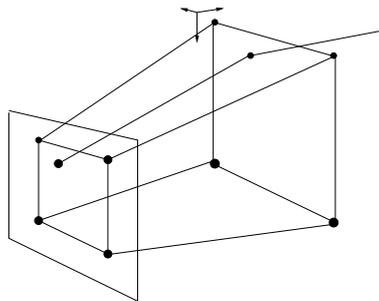


**Figure 1.7.** Caméra orthographique

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad [1.11]$$

Tous les rayons lumineux se projettent dans l'image selon un faisceau de droites parallèles. Il n'existe plus de notion de focale, de point de fuite, et donc plus d'effet perspectif dans la figure. Noter la représentation extrêmement simple de la matrice de projection.

#### 1.2.4.2. Projection orthographique à l'échelle



**Figure 1.8.** Caméra orthographique à l'échelle

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad [1.12]$$

La représentation précédente s'éloigne fortement en terme de réalisme des images réelles capturées par une caméra. On peut cependant se rapprocher d'une image conforme en introduisant une mise à l'échelle  $S$ .

Attention, dans cette représentation l'ensemble des points subit tout d'abord une projection orthographique, suivi d'une mise à l'échelle. (cela se voit notamment si on regarde la coordonnée  $z$  des points)

#### 1.2.4.3. Projection perspective faible

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad [1.13]$$

On utilise une double mise à l'échelle  $s_1$  et  $s_2$ .

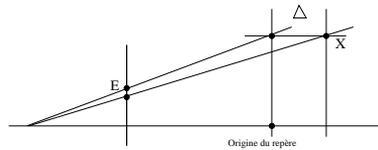


Figure 1.9. Caméra perspective faible

La Figure 1.9 montre l'erreur commise par ce type de projection par rapport à une projection perspective pure. Si la distance focale tend vers l'infini les deux sont équivalentes.

#### 1.2.4.4. Projection Affine

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} s_1 & d & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad [1.14]$$

#### 1.2.5. Expression générale

Le système complet de formation d'images s'exprime alors par la relation suivante :

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad [1.15]$$

où :

–  $(X_w, Y_w, Z_w)$  sont les coordonnées 3D d'un point de calibrage appartenant à l'étalon,

–  $(u, v)$  sont les coordonnées pixel 2D dans l'image de la projection de ce point,

–  $(u_0, v_0, f, dx, dy)$  sont appelés les *paramètres intrinsèques* de calibrage. Ils sont propres au système d'acquisition,

–  $(R_{(11, \dots, 33)}, T_{(x, y, z)})$  sont appelés *paramètres extrinsèques* de calibrage. Ils donnent la localisation de l'étalon dans le repère de la caméra, au moment de la prise de vue.

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = M_{int} M_{ext} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_{(3 \times 4)} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad [1.16]$$

$M$  est appelée **matrice de calibrage** du système.

Elle comporte 12 éléments, néanmoins nous pouvons la décomposer en :

– **5 paramètres intrinsèques** propres à la caméra.

On utilise généralement la matrice  $M_{int}$  sous la forme :

$$\begin{bmatrix} f/dx & 0 & u_0 & 0 \\ 0 & f/dy & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad [1.17]$$

En posant :

$$\left. \begin{array}{l} f_x = f/dx \\ f_y = f/dy \end{array} \right\} \rightarrow (f_x/f_y = dy/dx) \quad [1.18]$$

Le rapport  $dx/dy$  représente le rapport pixel ; à  $dx$  connu (fourni par le constructeur de la caméra) l'estimation des paramètres intrinsèques se ramène au calcul de **4 paramètres** ( $f_x, f_y, u_0, v_0$ ).

– **12 paramètres extrinsèques**, (9 pour la rotation ( $R_{11}..R_{33}$ ), 3 pour la translation ( $T_{(x, y, z)}$ )), *indépendants* de la caméra.

Soit un total de 16 paramètres.

#### 1.2.5.1. Formulation générale du problème de calibrage

Calibrer un système de vision, c'est être capable de déterminer tous les paramètres intervenant dans l'expression analytique de la formation d'une image explicitée en [1.15].

### 1.3. Approche Linéaire

#### Problème :

Étant donné un nombre  $n$  d'appariements 3D-2D  $(X_w, Y_w, Z_w; u, v)$  (expression [1.16]) entre une mire et son image, déterminer les 16 paramètres de formation d'images.

La résolution du problème de calibrage par une méthode linéaire présente l'avantage de ne pas demander de valeurs initiales des paramètres de calibrage pour trouver une solution. Nous verrons que ceci a un coût dans la pratique et que les performances d'une telle approche sont limitées. Faugeras et Toscani ([FAU 87]) ont proposé cette méthode dans le début des années 1980. Une approche similaire plus connue sous le nom de DLT (Direct Linear Transform) avait été introduite par la communauté photogrammétrique dans les années 1940.

#### 1.3.1. Principe

Soit la projection 3D-2D entre un point de l'étalon et son image :

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad [1.19]$$

$(m_{11} \dots m_{34})$  sont les 12 inconnues du système à résoudre.

En substituant  $s$ , il vient :

$$\begin{cases} u = \frac{m_{11}X_w + m_{12}Y_w + m_{13}Z_w + m_{14}}{m_{31}X_w + m_{32}Y_w + m_{33}Z_w + m_{34}} \\ v = \frac{m_{21}X_w + m_{22}Y_w + m_{23}Z_w + m_{24}}{m_{31}X_w + m_{32}Y_w + m_{33}Z_w + m_{34}} \end{cases} \quad [1.20]$$

avec les correspondances suivantes, tirées de l'expression matricielle [1.15] :

$$m_{1(1,2,3)} = \frac{f}{dx} R_{1(1,2,3)} + u_0 R_{3(1,2,3)}$$



En normant  $m_{3(1,2,3)}$ , on obtient le troisième vecteur de la matrice de rotation  $R_{3(1,2,3)}$ ,

$$T_z = m_{34} \quad [1.22]$$

$$\begin{pmatrix} m_{21} \\ m_{22} \\ m_{23} \end{pmatrix} \cdot \begin{pmatrix} m_{31} \\ m_{32} \\ m_{33} \end{pmatrix} = v_0 \quad [1.23]$$

$$\begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \end{pmatrix} \cdot \begin{pmatrix} m_{31} \\ m_{32} \\ m_{33} \end{pmatrix} = u_0 \quad [1.24]$$

$$\|m_{2(1,2,3)} - v_0 R_{3(1,2,3)}\| = f/dy \quad [1.25]$$

$$\|m_{1(1,2,3)} - u_0 R_{3(1,2,3)}\| = f/dx \quad [1.26]$$

$$(m_{2(1,2,3)} - v_0 R_{3(1,2,3)})/(f/dy) = R_{2(1,2,3)} \quad [1.27]$$

$$(m_{1(1,2,3)} - u_0 R_{3(1,2,3)})/(f/dx) = R_{1(1,2,3)} \quad [1.28]$$

Les 16 paramètres de formation d'images sont ainsi entièrement déterminés.

### 1.3.2. Remarques et commentaires

Cette méthode est très facile à implémenter. La résolution d'un système linéaire est peu coûteuse en temps de calcul. Cependant les résultats auxquels nous arrivons ne sont pas très stables. Par stabilité de la méthode, on entend son aptitude à fournir des résultats (sur les paramètres intrinsèques) semblables pour différentes attitudes de la mire.

Notons, que la décomposition précédente laisse entrevoir un problème au niveau de l'estimation de la rotation  $R$ . En effet, nous ne sommes jamais certain que :

$$R_{1(1,2,3)} \wedge R_{2(1,2,3)} = R_{3(1,2,3)} \quad [1.29]$$

En d'autres termes que la matrice de rotation  $R$  soit réellement conditionnée comme une matrice de rotation. Une solution palliative consiste alors à modifier la matrice de projection perspective pour lui adjoindre un terme supplémentaire qui peut s'interpréter comme la non orthogonalité de l'axe optique par rapport au capteur CCD et ce qui permet d'assurer l'orthogonalité des trois vecteurs de rotation..

En fait, la matrice de rotation n'est pas paramétrée de façon satisfaisante dans l'approche linéaire. Il conviendrait d'utiliser sa décomposition selon les angles d'Euler  $(\alpha, \beta, \gamma)$  respectivement autour des axes  $(x, y, z)$  du repère caméra. Il résulte, dans ces conditions, que le système à résoudre n'est plus *linéaire*. De plus, aucun phénomène de distorsion optique n'est pris en compte dans le processus de formation d'images.

Néanmoins cette méthode procure une très bonne estimation initiale à un processus d'optimisation non-linéaire.

#### 1.4. Approche non linéaire photogrammétrique

Dans cette section, nous abordons le problème du calibrage des caméras CCD selon le formalisme utilisé en photogrammétrie. Le modèle de projection utilisé pour le processus de formation d'images se réfère au modèle optique sténopé (c'est à dire, the *pin-hole* model), approximation du modèle optique de lentille mince. Cette approche se distingue de la précédente, d'une part par une modélisation précise des phénomènes de distorsion optique provoqués à la surface des lentilles et d'autre part par la mise en œuvre d'un processus d'optimisation non linéaire minimisant un critère, exprimé en pixel, de reprojction des points dans les images.

Les conventions de notations en photogrammétrie sont légèrement différentes de celles utilisées en vision artificielle. Entre autre, classiquement les données ne sont plus exprimées dans le repère de la caméra mais dans le référentiel du monde qui reste fixe quelle que soit la position de la caméra. Nous transformerons donc les notations photogrammétriques usuelles de manière à retrouver le formalisme propre à notre communauté.

Soient les notations suivantes :

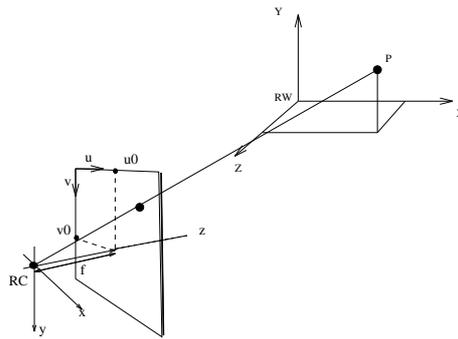
- $R_W-XYZ$  est un repère 3D direct. C'est le référentiel du monde qui servira entre autre de base de modélisation de l'objet.

- $o-uv$  est le référentiel 2D image, dont l'origine est située au coin haut-gauche de l'image.

- $R_C-xyz$  est le référentiel 3D de la caméra dont l'origine est au centre optique  $c$  et dont l'axe  $z$  est confondu avec l'axe optique.  $x, y$  sont respectivement parallèles à  $o-uv$ .

Les paramètres intrinsèques au capteur que nous devons déterminer sont : le point principal  $o - u_0v_0$ , la distance focale  $f$ , la taille des pixels de la matrice CCD ( $dx, dy$ ) ou leur rapport, et enfin les paramètres de distorsion optique introduits par l'objectif de la caméra.

Les paramètres extrinsèques sont la matrice de rotation  $\mathbf{R}$  ainsi que le vecteur de translation  $\mathbf{T}$  entre  $R_c-xyz$  et  $R_w-XYZ$ .



**Figure 1.10.** modèle sténopé, géométrie de l'image et systèmes de coordonnées.

La méthode décrite dans ce chapitre est strictement une approche au sens des moindres carrés ; nous allons essayer de minimiser des *erreurs de mesure* dans l'image, représentant la différence entre un point détecté dans l'image et le projeté du point 3D de la mire correspondant.

#### 1.4.1. Modèle Mathématique

Soit une projection perspective entre une image 2D et un objet 3D (sous une hypothèse *sténopé*). La relation entre une point de l'étalon et sa projection dans l'image est décrite par l'expression suivante :

$$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \lambda_i \left[ \mathbf{R} \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} + \mathbf{T} \right] \quad [1.30]$$

Où :

–  $(x_i, y_i, z_i)$  est un point image défini dans le Référentiel Caméra (cf figure 1.10 avec  $z_i \equiv f$ , c'est à dire , la distance focale de la caméra),

- $\lambda_i$  est un facteur d'échelle introduit lors du passage de  $R^3$  à  $R^2$
- $(X_i, Y_i, Z_i)$  sont les coordonnées du point de la mire définies dans le référentiel du monde  $W-XYZ$ ,
- $(T_x, T_y, T_z)$  est le vecteur de translation,
- $\mathbf{R}$ , la matrice de rotation exprimée dans le repère de la caméra et paramétrée selon les trois angles d'Euler :  $\alpha$  rotation autour de l'axe  $x$ ,  $\beta$  autour de l'axe  $y$ , et  $\gamma$  autour de l'axe  $z$  :

$$\mathbf{R} = \begin{pmatrix} \cos \gamma \cos \beta & \cos \gamma \sin \beta \sin \alpha - \sin \gamma \cos \alpha & \cos \gamma \sin \beta \cos \alpha + \sin \gamma \sin \alpha \\ \sin \gamma \cos \beta & \cos \gamma \cos \alpha + \sin \gamma \sin \beta \sin \alpha & \sin \gamma \sin \beta \cos \alpha - \cos \gamma \sin \alpha \\ -\sin \beta & \cos \beta \sin \alpha & \cos \beta \cos \alpha \end{pmatrix} \quad [1.31]$$

En éliminant  $\lambda_i$  dans (1.30) et en supprimant l'indice  $i$ , nous obtenons les expressions suivantes appelées *équations de colinéarité* en photogrammétrie :

$$\left. \begin{aligned} x &= f \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z} \\ y &= f \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z} \end{aligned} \right\} \quad [1.32]$$

Si nous exprimons  $(x, y)$  dans le système de coordonnées pixel de l'image nous obtenons :

$$\left. \begin{aligned} x &= (u + e_x - u_0)dx - do_x \\ y &= (v + e_y - v_0)dy - do_y \end{aligned} \right\} \quad [1.33]$$

Dans cette expression  $e_x, e_y$  sont les erreurs de mesure respectivement selon les coordonnées  $x$  et  $y$ , (c'est à dire, les corrections à apporter aux mesures pour qu'il existe une correspondance parfaite avec les données issues de la fonction de projection).  $do_x, do_y$  sont les composantes de distorsion optique qui se divisent en deux parties : distorsions *radiale* et *tangentielle*, (c'est à dire,  $do_x = do_{xr} + do_{xt}$  et  $do_y = do_{yr} + do_{yt}$ ). Nous introduisons ici, les deux formes communément utilisées en photogrammétrie [Ame84] :

$$\left. \begin{aligned} do_{xr} &= (u - u_0)dx(a_1r^2 + a_2r^4 + a_3r^6) \\ do_{yr} &= (v - v_0)dy(a_1r^2 + a_2r^4 + a_3r^6) \end{aligned} \right\} \quad [1.34]$$

$$\left. \begin{aligned} do_{xt} &= p_1[r^2 + 2(u - u_0)^2dx^2] + 2p_2(u - u_0)dx(v - v_0)dy \\ do_{yt} &= p_2[r^2 + 2(v - v_0)^2dy^2] + 2p_1(u - u_0)dx(v - v_0)dy \end{aligned} \right\} \quad [1.35]$$

où dans les expressions (1.33), (1.34), et (1.35),

- $u, v$  sont les coordonnées image dans le référentiel image,
- $u_0, v_0$  les coordonnées du point principal dans le référentiel image,
- $a_1, a_2, a_3$  les coefficients du polynôme qui modélise la distorsion radiale,
- $p_1, p_2, p_3$  les coefficients du polynôme qui modélise la distorsion tangentielle,
- $dx, dy$  représentent les facteurs d'échelle de la forme du pixel élémentaire,
- le paramètre  $r = \sqrt{(u - u_0)^2 dx^2 + (v - v_0)^2 dy^2}$ , est la distance radiale depuis le point principal. Étant donné que  $r$  peut prendre des valeurs importante (selon la taille de l'image),  $r^4$  et  $r^6$  deviennent parfois énormes ; l'expression (1.34) peut alors conduire à une instabilité numérique lors de l'estimation des différents paramètres. Un moyen de contourner cette difficulté est de récrire cette expression de la manière suivante :

$$\left. \begin{aligned} do_{xr} &= (u - u_0)dx(a_1(r^2 - r_0^2) + a_2(r^4 - r_0^4) + a_3(r^6 - r_0^6)) \\ do_{yr} &= (v - v_0)dy(a_1(r^2 - r_0^2) + a_2(r^4 - r_0^4) + a_3(r^6 - r_0^6)) \end{aligned} \right\} [1.36]$$

(1.36) suppose que la distorsion est nulle pour une distance radiale de  $r_0$ .

En substituant (1.33), (1.34) et (1.35) dans (1.32), et en posant  $f_x = \frac{f}{d_x}$  et  $f_y = \frac{f}{d_y}$ , nous obtenons le système suivant :

$$\left. \begin{aligned} u + e_x &= u_0 + do_{xr} + do_{xt} + f_x \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z} = P(\Phi) \\ v + e_y &= v_0 + (do_{yr} + do_{yt}) \frac{f_x}{f_y} + f_y \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z} = Q(\Phi) \end{aligned} \right\} [1.37]$$

où dans [1.37],  $\Phi$  est un vecteur de 15 paramètres, c'est à dire ,

$$\Phi = [u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, T_x, T_y, T_z, \alpha, \beta, \gamma]^T$$

#### 1.4.2. Résolution du problème

Reprenons les équations de colinéarités définies dans l'expression [1.37] :

$$\begin{pmatrix} e_x \\ e_y \end{pmatrix} = \begin{pmatrix} P(\Phi) - u \\ Q(\Phi) - v \end{pmatrix} = V(\Phi) \quad [1.38]$$

Le problème est à présent de déterminer la valeur de  $\Phi$  qui minimise

$$S = \sum_{i=1}^n (e_{x_i}^2 + e_{y_i}^2).$$

Dans (1.38),  $P(\Phi)$  et  $Q(\Phi)$  sont des fonctions non linéaires de  $\Phi$ , la minimisation est donc un problème d'optimisation non linéaire.

Un moyen de résoudre ce problème est de faire une linéarisation de [1.38] à partir d'une valeur initiale  $\Phi_0$  (généralement fournie par le résultats de la résolution linéaire du problème de calibrage décrite au paragraphe 1.3) et de calculer une correction  $\Delta\Phi$ . On ajoute alors  $\Delta\Phi$  à  $\Phi_0$  qui devient la nouvelle valeur initiale : le procédé doit être répété jusqu'à la convergence du système.

Soient  $n$  points 3D et leurs points correspondants dans l'image, on peut écrire sous forme matricielle le système de  $2 \times n$  équations linéarisées :

$$V(\Phi) = V(\Phi_0) - \sum_{i=1}^{15} \frac{\partial V}{\partial \Phi_i} \Delta\Phi_i \quad [1.39]$$

$$V = L - A \Delta\Phi \quad [1.40]$$

$L$  représente donc la valeur courante du critère et  $A$  la matrice jacobienne du système, autour du vecteur courant  $\Phi_0$ .

Soit la matrice de pondération des mesures <sup>1</sup>  $W$ ,

la résolution au sens des moindres carrés de (1.40) revient à estimer :

$$\Delta\Phi \in \mathbb{R}^{15} \min \left( V^T W V \right) \quad [1.41]$$

La solution de (1.41) est donnée par

$$\Delta\Phi = (A^T W A)^{-1} (A^T W L) \quad [1.42]$$

---

1. En général,  $W$  est une matrice diagonale  $2n \times 2n$ . Si toutes les mesures sont faites avec la même précision et qu'il n'existe pas de corrélation entre les paramètres, alors  $W$  est la matrice identité : c'est à dire,  $W = I$ .

**Démonstration :**

Soit  $\Omega = \mathbf{V}^T \mathbf{W} \mathbf{V}$ . A la solution du système, nous avons  $\frac{\partial \Omega}{\partial \Phi} = 0$ , c'est à dire ,

$$\frac{\partial \Omega}{\partial \Phi} = 2\mathbf{V}^T \mathbf{W} \frac{\partial \mathbf{V}}{\partial \Phi} = -2\mathbf{V}^T \mathbf{W} \mathbf{A} = 0, \implies \mathbf{A}^T \mathbf{W} \mathbf{V} = 0$$

En remplaçant  $\mathbf{V}$  à partir de son expression dans (1.40), l'équation ci-dessus devient :

$$\mathbf{A}^T \mathbf{W} (\mathbf{L} - \mathbf{A} \Delta \Phi) = \mathbf{A}^T \mathbf{W} \mathbf{L} - \mathbf{A}^T \mathbf{W} \mathbf{A} \Delta \Phi = 0$$

ce qui conduit à la solution de  $\Delta \Phi$  :

$$\Delta \Phi = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{W} \mathbf{L})$$

**1.4.3. Calibrage Multi-images**

Une des causes principale des mauvais résultats de calibrage provient des erreurs de mesure (quelles soient dans l'image aussi bien que sur la mire). Pour pallier ce problème, il est possible de combiner dans un même système plusieurs images provenant de la même caméra mais pour des positions spatiales (rotation et/ou translation) différentes. Dans ce cas, les paramètres intrinsèques du capteur sont les mêmes pour toutes les images et le calibrage estime le vecteur de paramètres suivant :

$$\Phi_{9+6m} = \left[ x_0, y_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \right. \\ \left. \dots, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m \right]^T$$

La matrice  $\mathbf{A}$  de (1.40) est alors de la forme :

$$\mathbf{A}_{2mn \times (9+6m)} = \left[ \begin{array}{c|ccc} & \mathbf{A}_{2n \times 6}^{11} & & 0 \\ & & \ddots & \\ \mathbf{A}_{2mn \times 9}^1 & & & \mathbf{A}_{2n \times 6}^{ii} \\ & & & & \ddots & \\ & 0 & & & & \mathbf{A}_{2n \times 6}^{mm} \end{array} \right]$$

où  $m$  est le nombre d'images et  $n$  le nombre de points par image. Le nombre total d'équations est  $(2mn)$  et le nombre total de paramètres est  $(9 + 6m)$ . La redondance  $r$  est donnée par  $r = 2mn - 9 - 6m$ , elle est bien plus importante que dans le cas de calibrage avec une seule image. L'approche multi-images, d'un point de vue expérimental, conduit à des résultats de qualité et permet d'assurer une grande reproductibilité de l'estimation des paramètres internes du capteur.

#### 1.4.4. Auto-calibrage par ajustement de faisceaux (*Bundle Adjustment*)

##### 1.4.4.1. Redéfinition du problème :

L'idée directrice part du constat suivant : les mires d'étalonnage de qualité sont difficiles à réaliser, la stabilité mécanique de l'ensemble pour une précision  $< 0,05mm$  ne s'obtient qu'avec des matériaux spécifiques et une mesure précise des points 3D utilisés est onéreuse. De plus la variabilité des champs angulaires selon les applications conduit à recourir à différentes mires adaptées aux conditions expérimentales.

L'auto-calibrage par ajustement de faisceaux est une approche multi-images qui permet une réestimation de la structure tridimensionnelle de la mire conjointement à l'estimation des paramètres classiques intrinsèques et extrinsèques du capteur. En d'autres termes, on étalonne la caméra et on reconstruit la mire simultanément.

Soient les équations de colinéarités :

$$\left. \begin{aligned} u + e_x &= u_0 + (do_{xr} + do_{xt})/dx \\ &+ \left( \frac{f}{dx} \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z} \right) = P(\Phi) \\ v + e_y &= v_0 + (do_{yr} + do_{yt})/dy \\ &+ \left( \frac{f}{dy} \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z} \right) = Q(\Phi) \end{aligned} \right\} [1.43]$$

Pour alléger l'écriture nous avons supprimé les indices, sans perte de généralité, mais les expressions de colinéarités s'entendent pour tous les points  $(X_i, Y_i, Z_i)$  des  $n$  points de la mire se projetant en  $(v_j, v_j)$  dans les  $m$  images.

Le vecteurs de paramètres à estimer, si l'on désire étalonner le capteur et calculer les coordonnées des points de la mire, prend alors la forme suivante :

$$\Phi_{9+6m+3*n} = \left[ x_0, y_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, \right]$$

$$\mathbf{X}^1, \mathbf{Y}^1, \mathbf{Z}^1, \dots, \mathbf{X}^n, \mathbf{Y}^n, \mathbf{Z}^n,$$

$$T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \dots, T_x^m, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m]^T$$

**Problème :** Trouver  $\Phi$  qui minimise  $\sum_{j=1}^m \sum_{i=1}^n (e_{x_{ij}}^2 + e_{y_{ij}}^2)$

#### 1.4.4.2. Estimation de la redondance

*Nombre d'inconnus :*

9 (paramètres intrinsèques) +  $3 * n$  (points de la mire) +  $6 * m$  (paramètres extrinsèques).

*Nombre d'équations :*

$2 * m * n$ .

La redondance du système  $r = 2 * m * n - (9 + 3 * n + 6 * m)$  est assurée sans trop de difficultés. Pour une mire de 11 points par exemple, un minimum de 3 images permet de surdéterminer le système.

#### 1.4.4.3. Solution à un facteur d'échelle près

*Paramètres intrinsèques*

De façon classique la matrice des paramètres intrinsèques est toujours déterminée à un facteur d'échelle près. L'introduction usuelle du facteur ( $dx = 1$ ) permet de fixer l'ensemble des paramètres en une unité arbitraire de pixels.

*Paramètres extrinsèques*

Dès lors qu'on estime à l'intérieur du processus les coordonnées tridimensionnelles des points de la mire, la géométrie extrinsèque du système est elle aussi fixée à un facteur d'échelle près. En effet il est toujours possible de trouver une mire plus volumineuse observée de plus loin qui donnerait strictement la même image.

Cette perte de métrique est sans grande importance pour l'étalonnage d'une caméra simple où seuls les paramètres intrinsèques représentent un intérêt pour l'utilisateur. Néanmoins, pour faciliter la convergence, deux points de la mire ne seront pas reconstruits et imposeront la métrique du système. Nous imposerons également qu'une des coordonnées d'un point quelconque parmi  $n$  reste fixe pour figer la géométrie extrinsèque globale de la scène reconstruite.

#### 1.4.4.4. Conditions initiales

Il est évident que l'optimisation d'un tel système non-linéaire requiert des conditions initiales dans le domaine de convergence. Nous montrerons dans la partie expérimentale que cette contrainte est satisfaite sans trop de difficultés dès lors que l'on a pris soin d'observer la mire sous des attitudes d'orientation différentes. Ce qui revient à assurer des angles de triangulation importants pour l'estimation des points de la mire.

Les valeurs de chaque terme du vecteur initial d'étalonnage pourront être fournies à partir de l'approche linéaire de calibrage ou bien plus simplement à partir des données constructeur pour les paramètres intrinsèques et d'un algorithme de localisation comme celui de D.Dementhon [DEM 95] pour les paramètres extrinsèques.

#### 1.4.5. Calcul de la Précision

A partir d'une estimation au sens des moindres carrés (1.40) et (1.41), il est possible de calculer une estimation du vecteur résiduel  $\mathbf{V}$  :

$$\hat{\mathbf{V}} = [\mathbf{A}(\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} - \mathbf{I}] \mathbf{L} \quad [1.44]$$

ainsi que l'estimation de l'erreur standard de poids unitaire (*standard error of unit weight*), qui représente une estimation a posteriori du bruit  $\sigma_0$  (scalaire) sur les points détectés dans l'image.

$$\hat{\sigma}_0^2 = \frac{\mathbf{V}^T \mathbf{W} \mathbf{V}}{N - P} \quad [1.45]$$

où  $P$  est le nombre total de paramètres estimés. La valeur de la matrice de *covariance* associée aux paramètres de  $\Phi$  est donnée par :

$$\mathbf{C}_\Phi = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \quad [1.46]$$

Ainsi il est possible de calculer pour chaque paramètre  $\phi_i$ , sa précision de détection (écart type) :

$$\hat{\sigma}_{\phi_i}^2 = \hat{\sigma}_0^2 c_{ii} \quad [1.47]$$

Un moyen de mesurer la fiabilité (qualité) d'un ajustement au sens des moindres carrés est de calculer la redondance relative du système [TOR 81], c'est à dire,

$$q = \text{tr} \left[ \mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \right] = \frac{r}{N} \quad [1.48]$$

où dans (1.45) et (1.47),  $r$  représente la redondance,  $N$  est le nombre total d'équations de mesure, et  $c_{ii}$  est le  $i^{\text{ème}}$  élément diagonal de la matrice de covariance  $\mathbf{C}_\Phi$ . Comme nous pouvons le noter, la redondance relative pour un calibrage multi-images est bien plus importante que pour un calibrage simple. Il résulte une plus grande fiabilité dans les résultats du calibrage multi-images,  $q_{\text{multi}} = \frac{r}{2mn} = 1 - \frac{9+6m}{2mn}$ , par rapport à la résolution simple-image,  $q_{\text{single}} = 1 - \frac{15}{2n}$ . Si  $m > 1$ , alors  $q_{\text{multi}} > q_{\text{single}}$ .

## 1.5. Résultats d'expérimentation

### 1.5.1. Ajustement de Faisceaux pour un Objectif Classique

Cette première partie présente un exemple d'étalonnage sur une caméra 1/2" équipée d'un objectif 10mm.

#### 1.5.1.1. Conditions initiales et expérimentales

Pour cette expérimentation, les coefficients des polynômes de distorsion ( $a_1, a_2, a_3, p_1, p_2$ ) sont initialisés à zéro. La mire ( $X_i, Y_i, Z_i, i \in [1, n]$ ) est mesurée grossièrement à quelques centimètres près ; notons que l'approche multi-images permet d'exploiter des mires planaires, ce qui facilite grandement leur fabrication et l'obtention d'un modèle grossier. Les positions initiales relatives caméra/mire ( $R_j, T_j, j \in [1, m]$ ) sont estimées par l'algorithme de Dementhon [DEM 95] appliqué aux objets plans.

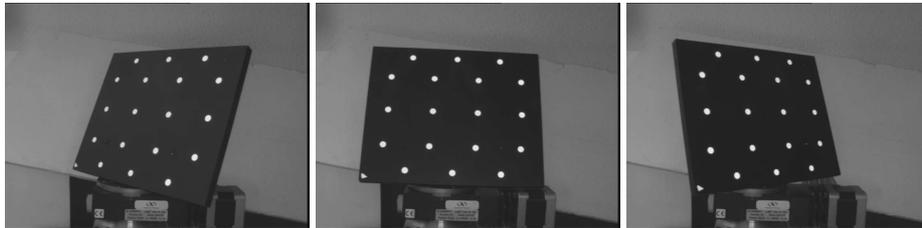
La position du point principal ( $u_0, v_0$ ) est placée au milieu de l'image et la focale ( $f_x, f_y$ ) est grossièrement estimée par la connaissance de la distance focale de l'objectif (10 mm par exemple) et la taille du pixel élémentaire de la matrice CCD (de 9 à 15  $\mu\text{m}$  selon les capteurs).

La série d'images analysée doit obligatoirement former un faisceau de vues convergentes afin de contraindre les angles de triangulation pour obtenir une reconstruction de la mire. Enfin l'approche s'appuie sur une détection de grande précision des points de la mire (voir [LAV 98]).

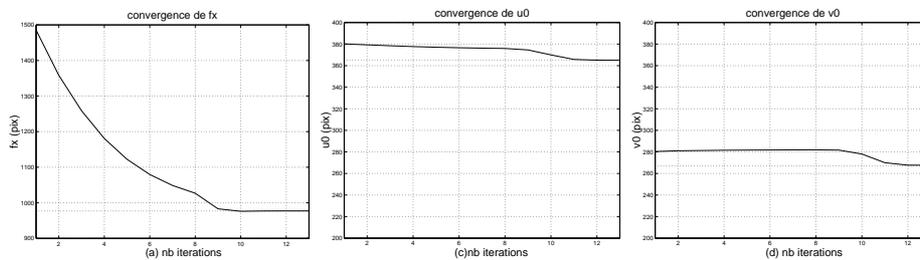
### 1.5.1.2. Séquence d'images classiques

Les images présentées sur la figure 1.11 montrent un échantillon des 15 vues prises pour calibrer une caméra équipée d'un objectif de 10mm. La mire est réalisée à partir d'une plaque plane et dotée de pastilles rétro réfléchissantes. Le dispositif de prise de vues intègre un éclairage annulaire haute fréquence qui permet d'obtenir des images des pastilles utilisées de bonne qualité quel que soit l'angle d'observation de l'objet.

Afin de mettre en évidence la convergence de l'algorithme, nous avons délibérément choisi de partir loin de la solution pour les paramètres internes de focale (1500 pixels au lieu de 1000) soit une erreur de 50%. En quelques itérations l'algorithme se stabilise autour d'un minimum qui conduit à des résidus de l'ordre de 0,025 pixel sur chaque coordonnée (voir tableau 1.1).



**Figure 1.11.** Séquence partielle des 15 vues prises pour l'auto-calibrage (768x576 pixels)

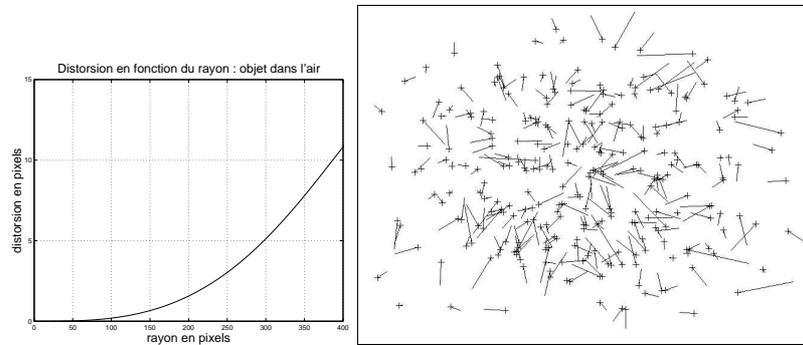


**Figure 1.12.** Convergence de  $f_x, u_0, v_0$ .

Les courbes (figure 1.12) montrent la convergence de  $f_x, u_0$  et  $v_0$  ; le minimum de l'expression [1.41] sera atteint au bout de 12 itérations.

#### Remarques :

– La stratégie de convergence ne pose pas de problème particulier pour cette série de données, mais nous verrons que ceci n'est pas toujours le cas pour les objectifs présentant une forte distorsion radiale. Pour ces images, tous les paramètres (intrinsèques, extrinsèques et coordonnées de la mire) sont estimés dès la première itération ;



**Figure 1.13.** Distorsion radiale (a) et Résidus à la convergence (b) (grossis 1000 fois) de l'ensemble des points de mesure (format 768x576 pixels).

caméra	: JAI M10	
objectif	: 10mm Ernitec	
carte de digitalisation	: Silicon Graphics	
algorithme	: Ajustement de faisceaux	
Nombre d'images	: 15	
Nombre de mesures	: 270	
Résidus ex moy et e-type (pix)	9.711e-05	2.555e-02
Résidus ey moy et e-type (pix)	2.480e-06	2.367e-02
$\sigma_0$ (pix)		2.917e-02

		$\sigma$
fx(pixel)	977.11	2.94e-01
fy(pixel)	977.50	2.86e-01
u0(pixel)	365.01	2.54e-01
v0(pixel)	267.64	3.75e-01
a1	+1.955e-07	3.11e-09
a2	+6.427e-14	4.82e-14
a3	-1.437e-18	2.17e-19
p1	+1.154e-06	5.42e-08
p2	-6.077e-07	9.04e-08

**Tableau 1.1.** Ajustement de faisceaux : Caméra Jai M10 et objectif 10mm

soit au total pour 15 images et une mire à 18 points :  $(11 + 15 * 6 + 3 * 18 = 155$  paramètres)

– Nous avons retracé à la solution les résidus de convergence ( $e_x, e_y$ ) (correspondant à l'expression (1.43)) . Chaque croix représente un point de mesure issu d'une des  $m$  images et le vecteur résidu associé subit un facteur multiplicatif de 1000 sur la figure 1.13b. La répartition homogène des vecteurs de résidu et leur orientation aléatoire traduit la non corrélation des erreurs à la solution.

### Souplesse d'utilisation

L'utilisation de l'approche décrite dans cet article s'avère très souple et permet de réaliser des étalonnages avec des mires modulables selon la distance focale de l'objectif utilisé. Les résultats de résidus obtenus en comparaison avec les approches classiques utilisant une mire mesurée avec précision, témoignent de la robustesse de l'algorithme.

### 1.5.2. Cas spécifique des objectifs Fish-Eye

Ce paragraphe aborde le cas spécifique des objectifs fish-eyes. La déformation principale engendrée par un objectif à courte focale est une déformation radiale. Plus l'objectif présente une distance focale petite et plus son champ angulaire d'observation est large. Les rayons convergeant sur la matrice CCD ont un angle incident sur le dioptre d'entrée de l'objectif de plus en plus important et s'éloignent par la même de l'hypothèse d'une optique paraxiale. Les rayons de courbure des lentilles induisent alors un phénomène radial prépondérant.

La stratégie pour modéliser les fortes déformations est intuitive ; elle consiste à augmenter l'ordre du polynôme de distorsion. Pour permettre la prise en compte de fortes distorsions radiales nous avons porté l'ordre du polynôme à 5. En revanche, plusieurs écriture du critère peuvent être envisagées.

#### 1.5.2.1. Critère classique

La formule [1.34], décrite précédemment, se transforme donc en

$$\left. \begin{aligned} do_{xr} &= (u - u_0)dx(a_1r^2 + a_2r^4 + a_3r^6 + a_4r^8 + a_5r^{10}) \\ do_{yr} &= (v - v_0)dy(a_1r^2 + a_2r^4 + a_3r^6 + a_4r^8 + a_5r^{10}) \end{aligned} \right\} \quad [1.49]$$

Le polynôme de distorsion se composera de cinq termes au lieu de trois. Comme le souligne Li dans [LI 94], étant donné que  $r$  peut prendre une valeur maximale de  $\frac{\sqrt{2}}{2}L$  (où  $L$  représente la taille de l'image)  $r^2 \dots r^{10}$  peuvent prendre des valeurs très importantes et l'expression [1.49] peut alors devenir numériquement instable.

#### 1.5.2.2. Distorsion nulle à $r_0$

Pour pallier cet inconvénient il est alors possible de récrire l'expression [1.49] de la façon suivante :

$$\left. \begin{aligned} do_{xr} &= (u - u_0)dx(a_1(r^2 - r_0^2) + a_2(r^4 - r_0^4) \\ &\quad + a_3(r^6 - r_0^6) + a_4(r^8 - r_0^8) + a_5(r^{10} - r_0^{10})) \\ do_{yr} &= (v - v_0)dy(a_1(r^2 - r_0^2) + a_2(r^4 - r_0^4) \\ &\quad + a_3(r^6 - r_0^6) + a_4(r^8 - r_0^8) + a_5(r^{10} - r_0^{10})) \end{aligned} \right\} \quad [1.50]$$

ou encore

$$\left. \begin{aligned} do_{xr} &= (u - u_0)dx \sum_{i=1}^5 (a_i(r^{2i} - r_0^{2i})) \\ do_{yr} &= (v - v_0)dy \sum_{i=1}^5 (a_i(r^{2i} - r_0^{2i})) \end{aligned} \right\} \quad [1.51]$$

L'expression [1.51] impose alors à la distorsion de prendre une valeur nulle pour une distance radiale  $r_0$  fixée. Il résulte dans cette écriture une variation de focale à la solution qui sera en première approximation égale à :  $\Delta f = \frac{f}{r_0} dr_0$

Néanmoins, [LAV 00a] montrent que cela ne résout en rien le fait que les coefficients aient des valeurs numériquement très faibles par rapport aux paramètres de focale par exemple. Cet aspect peut cependant être pris en compte par une troisième écriture du critère.

### 1.5.2.3. Normalisation des coefficients de distorsion

La normalisation consiste à compenser les valeurs importantes de  $r^n$  en normant toutes les distances par rapport à la focale  $f$ .

$$\left. \begin{aligned} (u - u_0)dx &\rightarrow (u - u_0)dx/f \\ (v - v_0)dy &\rightarrow (v - v_0)dy/f \end{aligned} \right\}$$

Il vient :

$$r'^2 = ((u - u_0)\frac{dx}{f})^2 + ((v - v_0)\frac{dy}{f})^2 = \frac{r^2}{f^2}$$

soit encore dans les expressions de distorsion :

$$\left. \begin{aligned} do'_{xr} &= (u - u_0)\frac{dx}{f}(a_1 r'^2 + \dots + a_5 r'^{10}) \\ do'_{xt} &= p_1[r'^2 + 2(u - u_0)^2 \frac{dx^2}{f^2}] + 2p_2(u - u_0)\frac{dx}{f}(v - v_0)\frac{dy}{f} \end{aligned} \right\} [1.52]$$

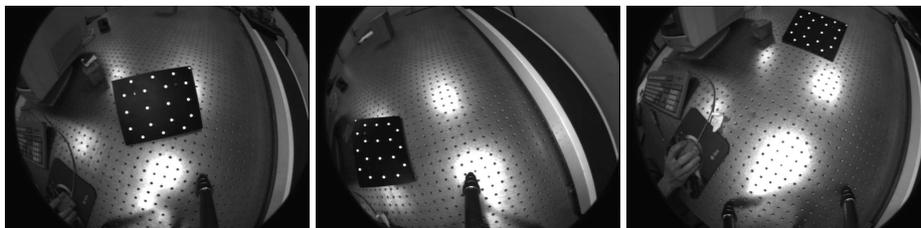
que nous pouvons récrire en mettant  $f$  en facteur :

$$\left. \begin{aligned} do'_{xr} &= \frac{1}{f}(u - u_0)dx \sum_{i=1}^5 (\frac{a_i}{f^{2i}} r^{2i}) \\ do'_{xt} &= \frac{1}{f}[ \frac{p_1}{f}(r^2 + 2(u - u_0)^2 \frac{dx^2}{f^2}) + 2\frac{p_2}{f}(u - u_0)dx(v - v_0)dy ] \end{aligned} \right\} [1.53]$$

Cette nouvelle expression présente une forme tout à fait semblable à celle définie en [1.34], au rapport de focale près. Le critère global prend alors la forme suivante :

$$\left. \begin{aligned} e'_x &= \frac{X}{Z} + do'_{xr} + do'_{xt} - (u - u_0)\frac{dx}{f} \\ e'_y &= \frac{Y}{Z} + do'_{yr} + do'_{yt} - (v - v_0)\frac{dy}{f} \end{aligned} \right\} [1.54]$$

L'intérêt de cette réécriture réside dans le fait que toute valeur importante de  $(r^{2i})$  dans le polynôme de distorsion se trouve compensée par  $(f^{2i})$ . De ce fait, et nous le verrons dans la partie expérimentale, les valeurs de  $(a_i)$  resteront proches de l'unité.

1.5.2.4. *Expérimentations*

**Figure 1.14.** *Exemple de vues (format : 768x576 pixels).*

Une partie de la séquence test à calibrer est présentée sur la figure (1.14). Nous pouvons noter que nous sommes en présence d'une très forte distorsion radiale. Une comparaison exhaustive du comportement des trois critères est analysée dans [LAV 00a], les résultats présentés dans le tableau 1.2 mettent en oeuvre l'écriture normalisée des expressions de colinéarité.

**Remarques générales :**

– La convergence de l'algorithme d'étalonnage requiert quelques précautions pour calibrer un capteur présentant une telle distorsion. En lâchant l'ensemble des paramètres dès la première itération, l'algorithme diverge systématiquement privilégiant une solution qui consiste à envoyer l'objet à l'infini. Ceci est d'autant plus vrai que nous partons loin de la solution finale puisque les coefficients de distorsion radiale et tangentielle sont initialisés à zéro.

– Pour contraindre les paramètres nous adoptons une démarche qui consiste à bloquer la focale et le point principal ( $f_x, f_y, u_0, v_0$ ) tant que le critère moyen n'est pas passé sous un seuil, fixé ici à 0,6 pixel. L'algorithme estimera donc dans un premier temps les distorsions, les localisations et la géométrie de la mire pour ensuite optimiser tous les paramètres dès que le seuil est franchi. Ce blocage de paramètre intervient au sein d'une procédure d'optimisation de Levenberg-Marquardt en agissant sur les dérivées du système. (valeur de départ de  $f_x, f_y$  fixée à 400 pixels, soit à 40% de la solution)

*Commentaires*

– Nous notons l'harmonisation de l'ordre de grandeur des coefficients de distorsion radiale sur le tableau 1.2. La variation entre les termes n'excède pas un facteur 10.

– la distorsion en bord d'image est supérieure à 200 pixels (figure 1.15).

– La décroissance du critère se fait sans grande difficulté. A partir de 0,6 pixel les 185 paramètres sont estimés en même temps. Toutes les courbes de décroissance présentent la même "cassure" au voisinage de la solution (0,25 pix) avant de plonger

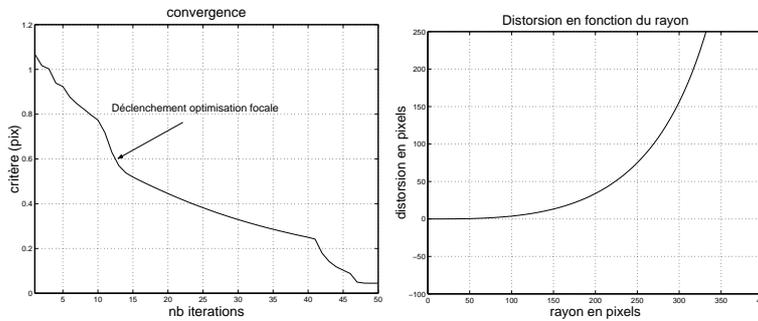


Figure 1.15. Convergence et distorsion du critère normalisé

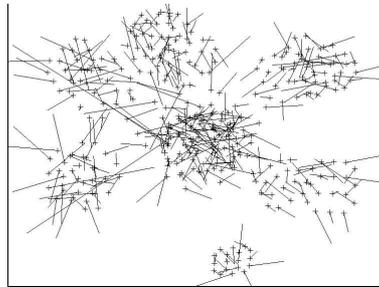


Figure 1.16. Résidus à la convergence (grossis 1000 fois)

vers la solution finale ; cette rupture n'est pas le fait d'un déblocage de paramètre de notre part, mais celui du choix de la stratégie de convergence de Levenberg-Marquardt.

- Les résidus à la solution sont de l'ordre de 0,05 pixels.
- Notons sur la figure 1.16 la présence de quelques *outliers* (norme de vecteur importante) qui pourront être supprimés par simple filtrage sur la valeurs des résidus.

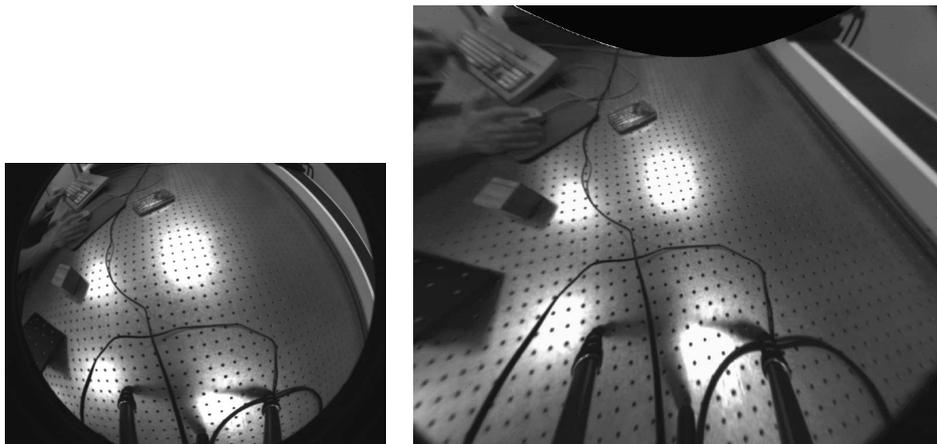
caméra	: JAI M300	
objectif	: 3.5mm TV-lens	
carte de digitalisation	: Silicon Graphics	
algorithme	: Ajustement de faisceaux	
Nombre d'images	: 20	
Nombre de mesures	: 360	
Résidus ex moy et e-type (pix)	9.631e-05	4.98e-02
Résidus e y moy et e-type (pix)	-1.309e-05	3.78e-02
Std error of unit weight (pix)		5.15e-02

		$\sigma$
fx(pixel)	326.89	1.82e-01
fy(pixel)	327.29	2.00e-01
u0(pixel)	396.67	1.78e-01
v0(pixel)	258.22	1.81e-01
a1	3.985e-01	4.51e-03
a2	1.892e-02	2.16e-02
a3	4.557e-01	4.54e-02
a4	-3.921e-01	4.22e-02
a5	2.261e-01	1.46e-02
p1	7.304e-06	1.27e-04
p2	1.253e-03	1.97e-04

Tableau 1.2. Étalonnage Fish-Eye. Critère normalisé

L'étalonnage des objectifs à courte focale s'appuie sur trois points majeurs : une détection précise de amers dans les images, un critère d'optimisation adapté et une stratégie de convergence qui lâchera progressivement les paramètres pour éviter l'écueil d'un minimum local.

L'approche d'auto-calibrage par ajustement de faisceaux présentée, offre l'avantage de réaliser très rapidement une mire adaptée à ces objectifs spécifiques. Sur les trois critères énoncés, deux retiennent plus particulièrement notre attention : le second qui consiste à fixer une distorsion nulle à une distance  $r_0$  du point principal, le troisième qui introduit une normalisation des coefficients de distorsion. En terme de résidus à la solution et pour une taille d'image constante les deux critères sont sensiblement équivalents, le second offre l'avantage de maintenir une taille d'image pratiquement constante avant et après correction de la distorsion. Quant au troisième il offre un conditionnement numérique plus favorable.



**Figure 1.17.** Image initiale (format 768x576 pixels) et corrigée (format 1168x976 pixels)

L'écriture d'un nouveau critère hybride reprenant les avantages des deux derniers doit apporter une solution satisfaisante à ce problème. La figure 1.17 montre une vue du capteur avant et après compensation des phénomènes de distorsion. La taille de l'image finale est augmentée de 400 pixels en ligne et colonne, et conduit à un champ d'observation réel de l'objectif de 120 degrés.

### 1.5.3. Calibrage des caméras sous-marines

#### 1.5.3.1. Rappels théoriques

Ce paragraphe aborde un certain nombre de notions relatives à l'étalonnage de caméras sous-marines. [LAV 00b] démontrent les liens de passage sur le changement de focale et la modification des courbes de distorsion entre l'utilisation du capteur vidéo dans l'air et dans l'eau. Il est alors possible d'étalonner le capteur dans l'air et de prévoir son fonctionnement dans un milieu d'indice quelconque.

Pour l'eau les lois suivantes doivent être vérifiées :

1/ Lorsque que la caméra est plongée dans l'eau, nous devons observer un facteur multiplicatif de 1.333 sur la valeur de la focale mesurée dans l'air.

$$f_{eau} = n_{eau} * f_{air} \quad [1.55]$$

2/ Soit  $u$  l'image distordue d'un point dans l'air et  $du$  la correction de distorsion à lui apporter pour obtenir le point de projection parfaite.

Soit  $u'$  l'image distordue du même point dans l'eau et  $du'$  la correction de distorsion à apporter pour obtenir le point de projection parfaite, alors :

$$1.333(u + du) = u' + du' \quad [1.56]$$

#### 1.5.3.2. Expérimentations

La partie expérimentale consiste à étalonner une caméra sous-marine dans l'air puis dans l'eau et d'analyser les résultats de calibrage compte tenu des relations théoriques exprimées au paragraphe précédent.

#### 1.5.3.3. Le matériel

Nous utilisons une caméra sous-marine expérimentale, réalisée autour d'un capteur CCD Sony. L'ensemble du dispositif optique est connu (indice, dimension et localisation de chaque dioptré).

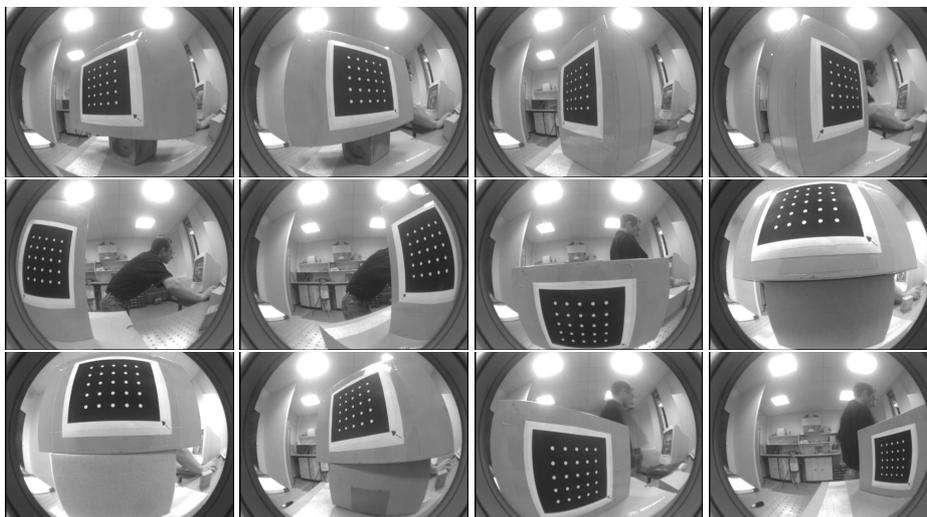


Figure 1.18. *Étalonnage dans l'air (768\*576 pixels)*

Résultats : Étalonnage dans l'air (indice = 1)		$\sigma$		
caméra	: Sony sous-marine	$f_x(\text{pix})$	375.65	3.39e-01
objectif	: 4mm	$f_y(\text{pix})$	375.81	3.06e-01
carte d'échantillonnage	: Silicon Graphics	$u_0(\text{pix})$	390.87	5.59e-02
algorithme	: Ajustement de faisceaux	$v_0(\text{pix})$	291.75	7.31e-02
Nombre d'images : 12		a1	6.63e-01	7.52e-03
Nombre de mesures : 283		a2	-1.15e-00	5.02e-02
Résidus moyen $e_x$ et $\sigma$ (pixel)		a3	6.83e+00	1.73e-01
	2.69e-05 4.15e-02	a4	-1.20e+01	2.65e-01
Résidus moyen $e_y$ et $\sigma$ (pixel)		a5	8.95e+00	1.56e-01
	1.49e-04 4.45e-02	p1	-1.02e-03	1.98e-04
		p2	1.22e-04	1.91e-04

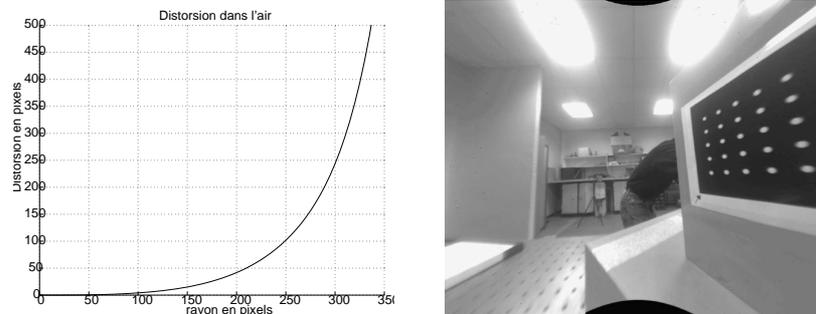
Tableau 1.3. *Ajustement de faisceaux dans l'air, sur caméra sous-marine*

#### 1.5.3.4. Résultats dans l'air

Nous avons calibrer le capteur à partir de 12 images présentées sur la figure 1.18. L'effet de distorsion radiale est important et le cercle noir visible en bord d'image provient de l'augmentation du champ angulaire de la caméra dans l'air. Il disparaîtra lors de l'immersion dans l'eau de cette dernière.

#### Notes :

– Nous avons utilisé une expression de la distorsion radiale en écriture normalisée, pour un polynôme d'ordre 5 (voir paragraphe 1.5.2.3).



**Figure 1.19.** Distorsion radiale et image corrigée (1168\*976 pixels)

– Le tableau (1.3) présente les résultats à la convergence du système. Les résidus sont de l'ordre de 0,04 pixels. L'algorithme se stabilise pour  $f_x = 376$  pixels et  $f_y = 376$  pixels, ce qui correspond à un champ angulaire d'observation de 110 degrés quand la distorsion est compensée.

– La distorsion radiale est représentée sur la figure 1.19a. Elle prend des valeurs importantes et au delà de 320 pixels (cercle noir) ne correspond plus à de mesures physiques prises dans les images. Pour représenter la vue compensée de la distorsion (Figure 1.19b) nous avons augmenté la taille de l'image de 400 pixels (lignes et colonnes). Cette image correspond à la sixième vue de la séquence (figure 1.18).

#### 1.5.3.5. Étalonnage dans l'eau

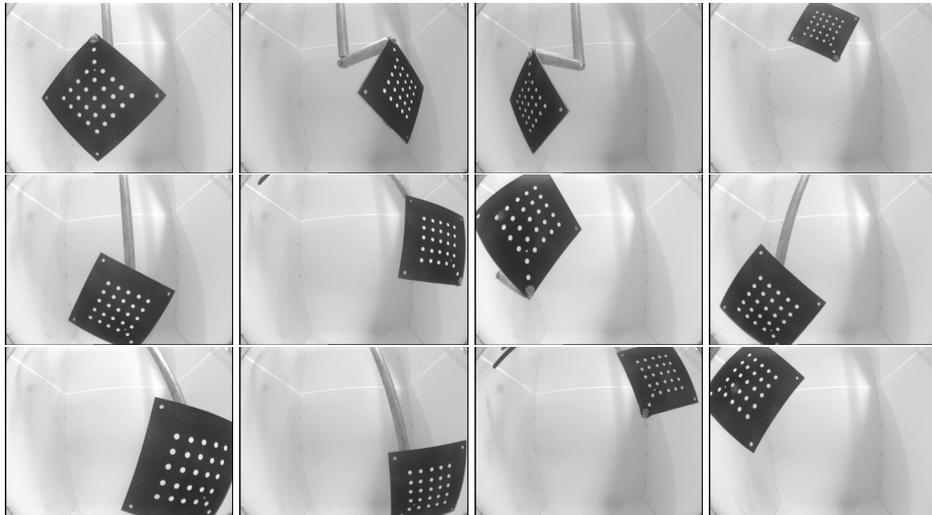
De façon similaire, nous avons calibré la caméra sous-marine dans l'eau à partir de l'analyse de 12 vues représentées sur la figure 1.20. Comme nous l'avons préalablement souligné le cercle noir en bord d'image a disparu et ceci tend à montrer qu'il y a bien modification du comportement interne du capteur au moment du changement d'indice. Notons enfin que le champ angulaire s'est fortement réduit et que la distorsion en bord d'image semble moins importante que dans l'air.

#### Notes :

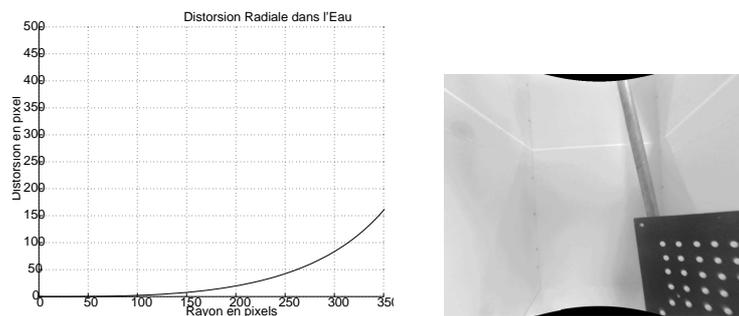
A la convergence la focale du dispositif sous-marin s'est stabilisée vers 500 pixels (au lieu de 376) ce qui conduit à un champ angulaire dans l'eau voisin de 90 degrés.

La courbe de distorsion radiale est tracée sur la figure 1.21a. Elle est bien moins importante que pour l'expérimentation dans l'air.

La figure 1.21b présente une des vues d'étalonnage après compensation des distorsions.



**Figure 1.20.** Séquence de vues pour le calibrage dans l'eau, (768\*576) pixels



**Figure 1.21.** Distorsion radiale et image sous-marine corrigée, (968\*776) pixels

### 1.5.3.6. Relations entre étalonnage dans l'air et dans l'eau

*Distance Focale* : Les lois de passage théoriques Air/Eau sont expérimentalement vérifiées :

La distance entre le point nodal image et la matrice CCD subit un facteur multiplicatif équivalent à l'indice dans lequel est plongée la caméra (voir [LAV 00b])

Le tableau (1.5) montre le rapport entre les distances focales dans l'air et dans l'eau. Si on intègre les incertitudes de détermination des focales estimées par le processus d'étalonnage, ce rapport est très voisin de 1.333.

Étalonnage dans l'eau (media index = 1.333) caméra sous-marine		
caméra	: Sony sous-marine	
objectif	: 4mm	
carte de digitalisation	: Silicon Graphics	
algorithme	: Ajustement de faisceaux	
Nombre d'images	: 13	
Nombre de mesures	: 325	
Résidus moyen $e_x$ et $\sigma$ (pixel)	1.36e-05	4.641e-02
Résidus moyen $e_y$ et $\sigma$ (pixel)	-3.17e-05	5.164e-02

		$\sigma$
fx(pixel)	499.12	5.51e-01
fy(pixel)	501.97	5.22e-01
u0(pixel)	391.81	9.27e-02
v0(pixel)	292.30	1.25e-01
a1	7.52e-01	1.06e-02
a2	-1.84e-00	7.22e-02
a3	9.88e-00	2.54e-01
a4	-1.73e+01	4.11e-01
a5	1.30e+01	2.57e-01
p1	-1.66e-03	2.32e-04
p2	1.62e-03	1.76e-04

**Tableau 1.4.** Ajustement de faisceaux. Caméra sous-marine

	f-air	f-eau	rapport (f-eau/f-air)
fx(pixel)	375.65	499.12	1.329
fy(pixel)	375.81	501.97	1.336

**Tableau 1.5.** Comparaison de la distance focale dans l'air et dans l'eau

Coordonnées du point principal  $(u_0, v_0)$  :

La position du point principal (intersection de l'axe optique et de la matrice CCD) est relativement stable entre les deux expérimentations (tableau (1.6)).

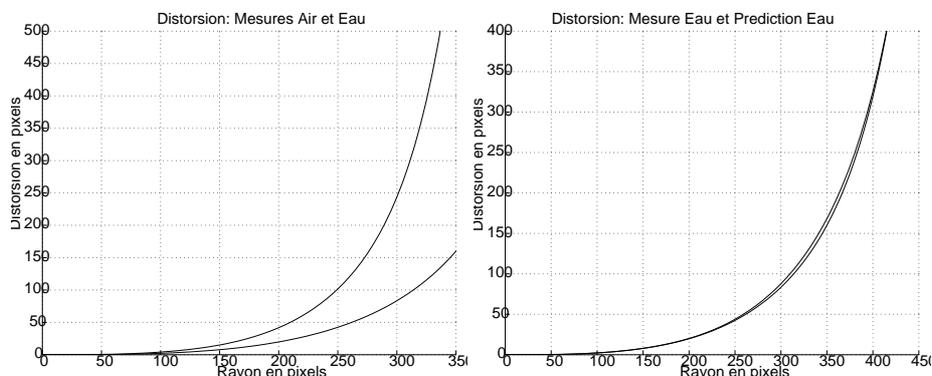
*Distorsion :*

La figure 1.22(a) montre une représentation conjointe des courbes de distorsion issues des campagnes de mesures dans l'air et dans l'eau. Nous pouvons noter l'importance de l'effet radial lors de l'expérimentation dans l'air.

Si on suppose que cette distorsion est majoritairement radiale les courbes doivent alors vérifier :

	Air	Eau
u0(pixels)	390.87	391.81
v0(pixels)	291.76	292.30

**Tableau 1.6.** Comparaison des coordonnées du point principal dans l'air et l'eau



**Figure 1.22.** Distorsions : (a) courbes tirées des mesures Air et mesures Eau, (b) superposition distorsion dans l'eau et prédiction de la distorsion dans l'eau à partir de la distorsion dans l'air

$$1.333(u + d_{air}(u)) = u' + d_{eau}(u')$$

La figure 1.23(b) montre la prédiction des courbes de distorsion dans l'eau à partir des données calculées dans l'air. Comme nous pouvons le noter la superposition est extrêmement bonne. Comme le champ de vue dans l'air est plus important que dans l'eau, la prédiction de la distorsion dans l'eau en bord d'image correspond à des mesures très précises dans l'air puisque cette partie de l'image est parfaitement visible. Ce qui n'est pas le cas pour les images dans l'eau, étant donné la difficulté expérimentale pour réaliser des vues correctes sur les bords d'images.

Il semble acquis que l'utilisation d'une caméra sous-marine puisse s'affranchir d'un étalonnage in-situ. En effet les prises de vues sous l'eau ne sont pas toujours facile à réaliser surtout si on doit faire attention à la nécessité de prise de d'images convergentes, uniformément réparties sur le capteur CCD. Un pré-étalonnage dans l'air est alors une solution élégante, et conduit à des résultats tout à fait intéressants.

#### 1.5.4. Calibrage des zooms

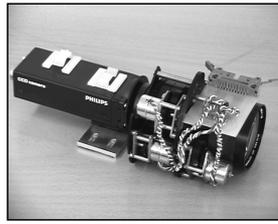
##### 1.5.4.1. Rappels des propriétés optiques

###### *Changement de focale*

Le changement de focale induit dans un zoom est réalisé par le déplacement d'un bloc mobile de lentilles convergentes. Cette modification de la configuration des dioptries provoque un grossissement apparent de l'image. Cependant la transformation subi par



**Figure 1.23.** *Distorsion : (a) Vue originelle (768x576 pix), (b) Non distordue (968x776pix) à partir des données dans l'eau, (c) non distordue à partir des données prédites depuis l'air*



**Figure 1.24.** *Caméra équipée d'un zoom*

l'image n'est une simple affinité et on montre [LAV 93] que les propriétés perspectives du dispositif sont largement modifiées.

Deux modifications majeures doivent être prises en compte :

- Modification de la distance focale ( $f_x, f_y$ ),
- Modification (souvent importante) de la distance entre le centre optique équivalent du dispositif et l'objet ( $T_z$ ). Ce phénomène dont l'origine est due au déplacement des points principaux du modèle optique épais, peut induire des déplacements axiaux d'une amplitude importante. (voir [LAV 93] pour plus amples détails)

Pour conclure l'action de zoom induit des modifications sur les paramètres intrinsèques et extrinsèques du vecteur de calibrage.

#### *Point principal*

D'un point de vue purement théorique, il conviendrait de considérer autant de points principaux indépendants que de configurations (zoom/mise au point) de l'objectif.

Dans la pratique, on note cependant que la précision de l'estimation du point principal décroît avec l'augmentation de la distance focale. En effet, plus la distance focale est grande et plus le modèle perspectif tend vers le modèle orthographique à l'échelle qui ne permet plus de déterminer  $(u_0, v_0)$ . (nous avons une compensation avec les paramètres  $T_x$  et  $T_y$ ).

Pour un zoom à fort grossissement, la solution consiste alors à fusionner le point principal avec le foyer d'expansion et de déterminer des coordonnées moyennes de  $(u_0, v_0)$  adaptées au zoom considéré [LI 96]. Cette approximation est d'autant plus valable que le phénomène de translation axiale du zoom est important.

#### 1.5.4.2. Estimation du point principal

Cette estimation se fait en analysant une séquence d'images en zoomant. Chaque point d'un même objet est suivi à travers la séquence et on estime la droite au sens des moindres carrés passant par chaque coordonnées (voir figure 1.25). Le calcul de l'intersection du faisceau de droites fournit le foyer d'expansion du zoom.

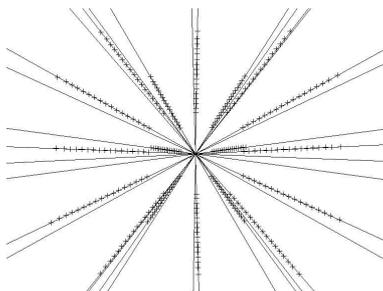


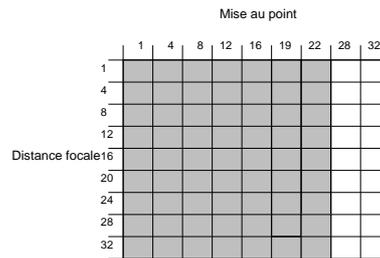
Figure 1.25. Détermination du foyer d'expansion

#### 1.5.4.3. Expérimentations



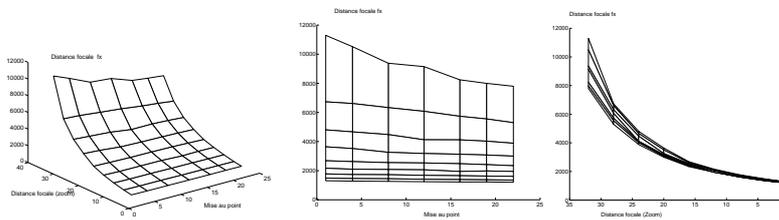
Figure 1.26. Tête active de l'Institut Royal de Technologie de Suède (KTH-Stockholm)

Le calibrage d'un zoom doit être réalisé pour l'ensemble des configurations de focale et mise au point utilisables par le dispositif. La figure 1.27 montre l'échantillonnage réalisé lors des expérimentations. Les amplitudes des déplacements sont normalisées.



**Figure 1.27.** Configurations testées (les zones grisées correspondent aux combinaisons zoom /mise au point réalisées durant les expérimentations).

Les figures montrent les résultats de calibrage de zooms qui équipent la tête active de KTH à Stockholm (Professeur JO.Eklundh) figure 1.26. Les courbes mettent en évidence l'effet couplé sur la focale d'un changement de zoom, mais également d'un changement de mise au point. Une analyse exhaustive des résultats est consultable dans [LI 95].



**Figure 1.28.** Distance focale

## 1.6. Facteurs perturbants de la mesure

### 1.6.1. *Un système avant tout électronique*

### 1.6.2. *Choix d'une mire 2D ou 3D*

### 1.6.3. *Configuration spatiales des vues*

## 1.7. Détection subpixelaire des amers

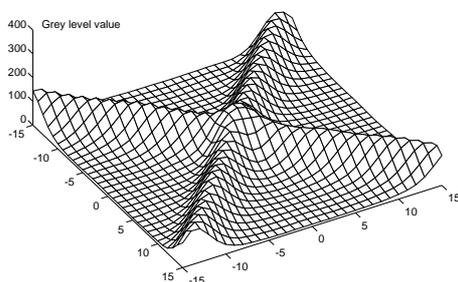
La recherche précise de la localisation des amers de calibrage est une étape cruciale dans un processus d'étalonnage. Nous montrerons que cette remarque prendra d'autant plus d'importance que nous nous approcherons des algorithmes d'auto-calibrage ;

Dans cette section nous développons plus précisément deux extracteurs sub-pixelique utilisés au laboratoire. Le premier qui date déjà un peu s'appuie sur une modélisation sous forme exponentielle de la réponse photonique d'une croix gravée sur une surface uniforme. Cette approche a été utilisée pour tous les résultats mettant en avant le cube du KTH. Les personnes intéressés par ce type de modélisation pourront facilement transformer l'expression analytique pour modéliser des points ou des croix inverses.

La seconde est à rapprocher de l'utilisation de pastilles photoréfléchissantes couplées à un éclairage annulaire haute-fréquence. En effet un tel système permet une observation optimale quel que soit l'incidence de l'angle d'observation mais ne tolère pas une modélisation de nature exponentielle de la réponse de la tâche.

### 1.7.1. *Accurate Computation of a cross*

In this paragraph we describe the special algorithm used to compute the coordinates of each intersection points of the KTH calibration cube (called cross position).



**Figure 1.29.** *Real grey level signal around a cross position*

Instead of using noisy edge detected information, the basic idea is to work directly with the grey level data to determine the cross coordinates. The software used to compute the sub-pixel localization was introduced by [PEU 93].

The author has found the expression of an analytical surface whose shape is really close to the photonic response detected around a cross position (see Figure 1.29).

Let us consider this expression  $F(x, y)$ ,

$$\begin{aligned} F(x, y) = & a + b * x + c * y \\ & + d_1^2 * \exp^{-e_1^2 * (x * \cos \Theta_1 + y * \sin \Theta_1 + l_1)^2} \\ & + d_2^2 * \exp^{-e_2^2 * (x * \cos \Theta_2 + y * \sin \Theta_2 + l_2)^2} \end{aligned}$$

- Coefficient  $a$  represents the mean grey level value of the background,
- Coefficients  $b, c$  take into account a regular slope according to the background grey level values. (Useful when cross surface is not parallel to the image plane)
- Coefficients  $(d_1, e_1, l_1, \Theta_1)$  and respectively  $(d_2, e_2, l_2, \Theta_2)$  are shape parameters of each cross branches.
- Coefficients  $d_1, d_2$  represent the branches height.
- Coefficients  $e_1, e_2$  represent the branches width.
- Coefficients  $\Theta_1, \Theta_2$  represent the branches orientation.
- Coefficients  $l_1, l_2$  represent the orthogonal distance from the branches to the local origin point  $(0, 0)$ .

Let us suppose that the cross position is approximatively found (using edge detection previously presented), then, the analytical surface  $F(x, y)$  can be fitted inside the grey level data by adjusting the function parameters. The problem can be summarized as following :

$$\mathbf{V} = (a, b, c, d_1, e_1, l_1, \Theta_1, d_2, e_2, l_2, \Theta_2)^T$$

such that

$$\sum_{x=-W_s}^{W_s} \sum_{y=-W_s}^{W_s} (I(x, y) - F(x, y))^2$$

is minimized. Where  $I(x, y)$  is the real grey level value at point  $(x, y)$  and  $W$ s the half size window of the local approximation.

Levenberg-Marquardt algorithm [PRE 92] has been used to solve this non-linear minimization. Even if the number of parameters to be fitted is important, straightforward tools can be used to start the process with an approximate value  $V'$  close to the solution. Usually, less than ten iterations per cross are needed to reach a stable solution.

### **How to compute the sub-pixel location from the analytical surface ?**

Performing this task doesn't not present any difficulty. According to  $F(x, y)$  formula, the cross center location  $(x_0, y_0)$  is computed by solving the following linear system :

$$\begin{cases} x_0 * \cos \Theta_1 + y_0 * \sin \Theta_1 + l_1 = 0 \\ x_0 * \cos \Theta_2 + y_0 * \sin \Theta_2 + l_2 = 0 \end{cases}$$

#### **1.7.2. Détection sub-pixellique des amers visuels taches photoréfléchissante.**

La détection précise de la projection d'une pastille photoréfléchissante sera obtenue par ajustement d'un modèle local paramétrique, capable de simuler la réponse du capteur grâce à un processus d'optimisation.

Cette modélisation comporte deux aspects.

Les pastilles utilisées, de forme circulaire, sont observées d'un point de vue quelconque. Leur projection dans le plan image prend donc un aspect elliptique.

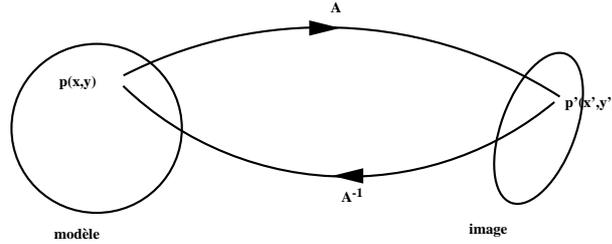
La modélisation doit permettre une déformation d'un modèle de luminance du support de la pastille, présentant a priori une symétrie de révolution, pour l'adapter à sa projection dans le plan image.

Nous supposons de plus que la luminance est approximativement constante respectivement au centre de la pastille et dans son voisinage extérieur proche. Ceci est obtenu grace à un éclairage annulaire centré autour de l'objectif de la caméra.

Il nous faudra donc estimer les deux niveaux de gris correspondants, ainsi que la forme de la transition entre ces derniers.

## 1.7.2.1. Déformation géométrique.

Cette partie repose sur la déformation par une transformation affine  $A$  du modèle de «rayon»  $R$  présentant a priori une symétrie de révolution de la tache observée vers sa forme courante elliptique.



$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} l_1 \cos \theta & l_2 \sin \theta & t_u \\ -l_1 \sin \theta & l_2 \cos \theta & t_v \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad [1.57]$$

En réalité, pour éviter d'avoir à interpoler entre les niveaux de gris des différents pixels physiques de la matrice CCD, nous procédons en sens contraire en déterminant la transformation inverse  $A^{-1}$  qui va ramener la tache de forme elliptique à un modèle circulaire.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 \cos \theta & -\lambda_2 \sin \theta \\ \lambda_1 \sin \theta & \lambda_2 \cos \theta \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} x' - t_u \\ y' - t_v \\ 1 \end{bmatrix} \quad [1.58]$$

après optimisation des paramètres  $t_u$  et  $t_v$  représenteront la position sub-pixellique du centre de notre tache elliptique.

Les autres paramètres représentent respectivement pour  $\lambda_1$  et  $\lambda_2$  les facteurs d'élongations du petit et du grand axe, et  $\theta$  est l'angle du grand axe avec l'axe des  $x$ .

## 1.7.2.2. Modélisation de la transition.

Le modèle de transition de luminance entre le centre et la périphérie de la tache est composée par la concaténation de 5 arcs de courbes (3 segments de droites et 2 arcs de cercles).

Les paramètres optimisés sont au nombre de 6 :

- niv-haut : niveau de luminance au centre de la tache ( $n_h$ );

- niv-bas : niveau de luminance à la périphérie ( $n_b$ ) ;
- h : localisation du milieu du front montant ;
- p : pente du front montant ;
- r1 : rayon du cercle haut ;
- r2 : rayon du cercle bas.

Sur les figures qui suivent on peut observer différents exemples de ce modèle pour diverses valeurs des paramètres.

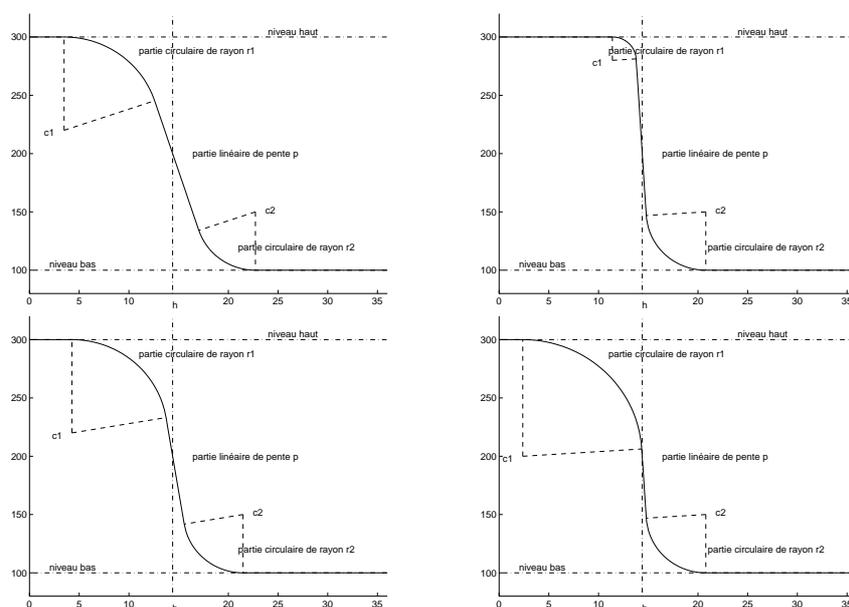


Figure 1.30. exemples du modèle de transition.

### 1.7.2.3. Exemple de détection

L'estimation, pour chaque amer visuel du vecteur de paramètres

$$(t_u, t_v, \lambda_1, \lambda_2, \theta, n_h, n_b, v_1, v_2, p)$$

est réalisé par l'algorithme d'optimisation de Levenberg-Marquardt. Nous obtenons ainsi, le vecteur de paramètres minimisant la somme des carrés des différences de niveau de gris entre les pixels de la tache elliptique dans l'image et le point correspondant (par la transformation  $A^{-1}$ ) du modèle théorique de la réponse du capteur. En général, après optimisation la différence moyenne est de l'ordre de 1 à 2 niveaux de gris.

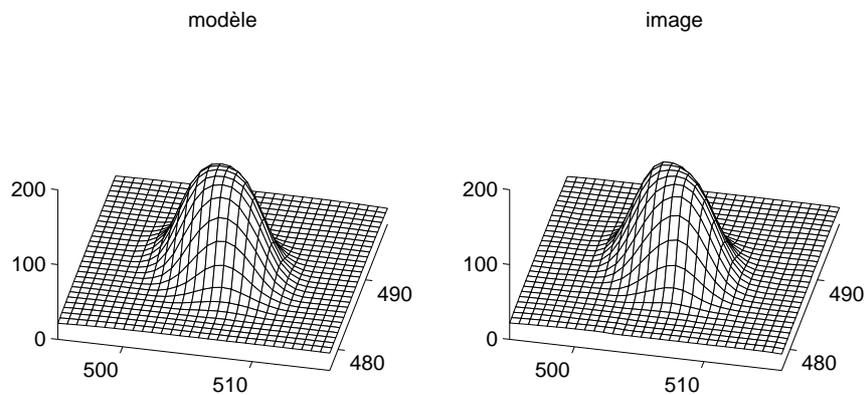


Figure 1.31. exemple de détection

### 1.7.3. Influence de la précision de détection, sur les paramètres de calibrage

#### Protocole

Soit une séquence de 11 images de synthèse représentant une mire de 11 points, observée à l'aide d'un objectif (focale 10mm). La séquence est calculée de manière à permettre l'observation de la mire sous des points de vues fortement distincts. Dans cette expérience, les points volumiques de la mire sont supposés sans erreur de métrologie. Les tableaux ci-dessous montrent les résultats d'étalonnage pour différentes valeurs de bruit (en pixel) introduit sur les coordonnées  $(u,v)$  des projections de chaque point de la mire.

L'analyse du tableau 1.7 montre qu'une détection sous-pixelique est indispensable pour réaliser un étalonnage de qualité. Au delà de 0.5 pixel les erreurs sur les paramètres intrinsèques commencent à devenir importantes.

## 1.8. Bibliographie

- [Ame84] American Society for Photogrammetry, Manual of Photogrammetry, 4th édition, 1984.
- [DEM 95] DEMENTHON D., DAVIS L., « Model-based object pose in 25 lines of code », *International Journal of Computer Vision*, vol. 15, p. 123, 141, June 1995.
- [FAU 87] FAUGERAS O., TOSCANI G., « Camera calibration for 3D Computer Vision », *Proc. of CVPR*, Tokyo, japan, 1987.

<i>bruit (u,v) 0.0 pix</i>			
$\sigma_0 = 0.0$ (pix)			
$fx(pix) = 1670.0$	$\sigma fx = 1.0e-9$		
$fy(pix) = 1671.0$	$\sigma fy = 1.0e-9$		
$u0(pix) = 391.0$	$\sigma u0 = 1.0e-9$		
$v0(pix) = 278.0$	$\sigma v0 = 1.0e-9$		

<i>bruit (u,v) 0.02 pix</i>		<i>bruit (u,v) 0.05 pix</i>	
$\sigma_0 = 0.02$ (pix)		$\sigma_0 = 0.05$ (pix)	
$fx = 1670.16$	$\sigma fx = 0.77$	$fx = 1670.58$	$\sigma fx = 1.94$
$fy = 1671.24$	$\sigma fy = 0.78$	$fy = 1671.79$	$\sigma fy = 1.96$
$u0 = 391.66$	$\sigma u0 = 0.56$	$u0 = 392.63$	$\sigma u0 = 1.41$
$v0 = 278.76$	$\sigma v0 = 0.63$	$v0 = 279.93$	$\sigma v0 = 1.58$

<i>bruit (u,v) 0.5 pix</i>		<i>bruit (u,v) 1.0 pix</i>	
$\sigma_0 = 0.51$ (pix)		$\sigma_0 = 1.06$ (pix)	
$fx = 1704.82$	$\sigma fx = 20.63$	$fx = 1837.14$	$\sigma fx = 51.85$
$fy = 1708.11$	$\sigma fy = 20.77$	$fy = 1842.94$	$\sigma fy = 52.00$
$u0 = 402.92$	$\sigma u0 = 14.70$	$u0 = 400.93$	$\sigma u0 = 33.58$
$v0 = 303.34$	$\sigma v0 = 16.53$	$v0 = 346.98$	$\sigma v0 = 40.94$

**Tableau 1.7.** Résultats d'étalonnage pour des erreurs croissantes sur les coordonnées images, de la projection des points de la mire

- [LAV 93] LAVEST J., RIVES G., DHOME M., « 3D Reconstruction by Zooming », *IEEE Trans. on Robotics and Automation*, vol. Vol. 9, n°Nr 2, p. pp 196-207, avril 1993.
- [LAV 98] LAVEST J., VIALA M., DHOME M., « Do we really need an accurate calibration pattern to achieve a reliable camera calibration », *Fifth European Conference on Computer Vision, ECCV98, Freiburg, Germany*, vol. Vol 1, p. pp 158-174, juin 1998.
- [LAV 00a] LAVEST J., DHOME M., « Comment calibrer des objectifs à très courte focale ? », *Conf. Reconnaissance des Formes et Intelligence Artificielle*, Paris, France, p. 81-90, 2000.
- [LAV 00b] LAVEST J., RIVES G., LAPRESTE J., « Underwater-camera calibration », *Proc. European Conference on Computer Vision*, vol. 2, Dublin, Ireland, p. 654-672, 2000.
- [LI 94] LI M., « Camera Calibration a head-eye system for Active Vision », EKLUNDH J., Ed., *Proc. of 3rd European Conf. on Computer Vision*, Stockholm, p. I :543-554, mai 1994.
- [LI 95] LI M., LAVEST J., « Some aspects of zoom lens calibration », *Technical Report KTH-NADA-CVAP*, , n°Nr 172, 1995.
- [LI 96] LI M., LAVEST J., « Some aspects of zoom lens calibration », *IEEE PAMI*, vol. Vol 18, n°Nr 11, 1996.

- [PEU 93] PEUCHOT B., « Camera Virtual Equivalent Model 0.01 Pixel Detector », *Computerized Medical Imaging and Graphics*, vol. Vol 17, n° 4/5, p. pp.289-294, juillet 1993.
- [PRE 92] PRESS W., TEUKOLSKY S., VETTERLING W., FLANNERY B., « Numerical Recipes in C, second edition », *Cambridge University Press*, 1992.
- [TOR 81] TORLEGÅRD K., « Accuracy Improvement in Close Range Photogrammetry », *Schriftenreihe, Wissenschaftlicher Studiengang Vermessungswesen, Hochschule der Bundeswehr München*, vol. Heft 5, septembre 1981.



## Chapitre 2

# Index

### A

ajustement de faisceaux 38  
auto-calibrage 38

### C

calibrage des caméras sous-marines 49  
  des zooms 54  
  faible 20  
  fort 20  
  géométrique 20  
  linéaire 29  
  Multi-images 37  
  photogrammétrique 32  
  radiométrique 20  
calibrage 19, 20  
caméra sous-marine 49  
chaîne d'acquisition 19  
conditions initiales 40

### D

distorsion optique 32

### E

équations de colinéarité 34  
erreur standard 40

### F

facteur d'échelle 39  
fi sh-eyes 44

### M

matrice CCD 25  
mire 19  
modèle sténopé 21

### N O

normalisation 45  
objectifs à courte focale 48

### P

paramètres extrinsèques 28  
  intrinsèques 28  
perte de métrique 39  
plan image 21  
polynôme de distorsion 44  
projection perspective 22

### R

repère Caméra 23  
  Image 23, 24  
  Modèle 23  
résidus de convergence 43

68 ENST 2005

**S**  
stratégie de convergence 42

système d'acquisition 20  
zoom 54

## Fiche pour le service de fabrication

**Auteurs :**

Jean-Marc LAVEST  
lavest@lasmea.univ-bpclermont.fr

**Titre du livre :**

Ecole Nationale Supérieurs des  
Télécommunications

Calibrage des systèmes de vision

Lasmea UMR 6602 CNRS

**Titre abrégé :**

*ENST 2005*

**Date de cette version :**

*13 mars 2005*

**Contact :**

- téléphone : 04 73 40 72 28
- télécopie : fax 04 73 40 72 62
- Mél : dhome@lasmea.univ-bpclermont.fr

**Logiciel pour la composition :**

- L<sup>A</sup>T<sub>E</sub>X, avec la classe ouvrage-hermes.cls,
- version 1.3, 17/09/2001.
- traité (option treatise) : Oui (*chapitres avec différents auteurs*)
- livre en anglais (option english) : Non (*par défaut en français*)
- tracé des limites de page (option cropmarks) : Non (*par défaut*)
- suppression des en-têtes de page (option empty) : Non (*par défaut*)

70 ENST 2005

- impression des pages blanches (option allpages) : Oui
- césures actives : voir la coupure du mot signal dans le fichier .log