



HAL
open science

Approximation du changement de focale et de mise au point par une transformation affine a trois parametres

Reyes Enciso, Thierry Viéville, Olivier Faugeras

► **To cite this version:**

Reyes Enciso, Thierry Viéville, Olivier Faugeras. Approximation du changement de focale et de mise au point par une transformation affine a trois parametres. [Rapport de recherche] RR-2071, INRIA. 1993. inria-00074601

HAL Id: inria-00074601

<https://inria.hal.science/inria-00074601>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Approximation du Changement
de Focale et de Mise au Point
par une Transformation Affine
à Trois Paramètres***

Reyes Enciso, Thierry Viéville

Olivier Faugeras

N° 2071

Octobre 1993

PROGRAMME 4

Robotique,
image
et vision



R ***apport
de recherche***

1993



Approximation du Changement de Focale et de Mise au Point par une Transformation Affine à Trois Paramètres

Reyes Enciso, Thierry Viéville
Olivier Faugeras

Programme 4 — Robotique, image et vision
Projet Robotvis

Rapport de recherche n° 2071 — Octobre 1993 — 32 pages

Résumé : Dans ce travail, on approxime dynamiquement le changement de focale et de la mise au point en suivant deux points quelconques entre deux images prises par la même caméra.

Plus précisément, on étudie la variation des paramètres intrinsèques lors d'un changement de la mise au point et du zoom. On démontre grâce à cette analyse, qu'un modèle de transformation affine à 3 paramètres est tout à fait suffisant, et qu'un modèle de transformation affine général ne se justifie pas, car la transformation à utiliser n'a mathématiquement que 5 paramètres et non 6 tandis que l'analyse physique du système montre que 3 paramètres suffisent. Expérimentalement, le modèle est justifié, car la précision est meilleure que 1.5 pixel dans la quasi-totalité des cas.

Par une méthode des moindres carrés on accède à des équations très simples qui nous permettent d'obtenir une précision dépendant du nombre des points suivis.

Mots-clé : Calibration des paramètres intrinsèques, Zoom, Mise au Point

(Abstract: pto)

A Three-Parameter Affine Approximation of Focus and Zoom Variations

Abstract: This work aims to develop a dynamic approximation to the variation in a lens' intrinsic parameters when the zoom and focus parameters are modified. This approximation is built by tracking two generic points in a monocular image sequence.

Our preliminary analysis demonstrates that a particular three-parameter affine model is sufficient to describe these modifications. The general affine model is not acceptable on a mathematical or physical level: the mathematical transformation to be used has only five parameters, instead of six, and an analysis of the physics reveals that three parameters are sufficient. Experimentally, this approximation is entirely valid, with the precision being better than 1.5 pixels in almost every case.

Using a least-squares method, we obtain very simple equations in which the precision of the estimate increases with the number of available correspondences.

Key-words: Intrinsic Parameters Calibration, Zoom, Focus

1 Introduction

Le problème de la calibration d'un capteur visuel statique est un sujet bien étudié [13, 14, 3]. Une revue de ces méthodes est donnée dans [15] et une présentation générale dans [5], nous n'y reviendrons pas.

Deux écoles principales se sont développées:

- celle de la **calibration "forte"**, où l'on calcule à partir d'un objet bien connu les paramètres *intrinsèques* (l'angle entre les deux axes du repère caméra θ , les facteurs d'échelle α_u, α_v , et le point principal (u_0, v_0)), et les paramètres *extrinsèques* (la translation et rotation du capteur par rapport au système robotique),
- celle de l'**auto-calibration** d'un système robotique où la même tâche est accomplie sans aucune connaissance "a priori" des objets observés et sans une connaissance complète du mouvement du système [6, 10, 16].

Dans le domaine de la Vision Active la calibration "forte" n'est pas envisageable, car les paramètres du système visuel varient en permanence et qu'il n'est pas facile de maintenir toujours un objet connu dans le champ visuel.

En se basant sur un modèle projectif de la géométrie du système il est possible de considérer une notion de calibration "faible", construite à partir de la connaissance de la géométrie épipolaire et de sa représentation sous forme de matrice fondamentale [6, 10, 9] qui nous permet de reconstruire un monde projectif.

De manière plus simple, dans le cas d'un système mécanique au comportement reproductible, on peut essayer de maintenir à jour une matrice de projection de la caméra quels que soient les paramètres du système. C'est le but de ce travail, où l'on va approximer dynamiquement le changement de mise au point et de focale en suivant deux points quelconques entre deux images prises par la même caméra.

Plus précisément, nous allons étudier la variation des paramètres intrinsèques lors d'un changement de mise au point et de focale. Nous démontrons grâce à cette analyse qu'un modèle de transformation affine à 3 paramètres

est tout à fait suffisant et qu'un modèle de transformation affine général à 6 paramètres ne se justifie pas, car la précision est meilleure que 1.5 pixels dans la plupart des cas.

Par une méthode des moindres carrés on déduit des équations très simples qui nous permettent d'obtenir une précision augmentant avec le nombre des points suivis.

2 Matrice des paramètres intrinsèques et modèle de lentille épaisse.

On utilise le *modèle standard sténopé* pour la caméra. Dans une position indexée par i , on suppose que la caméra réalise une transformation perspective parfaite de centre C_i (le centre optique de la caméra) à une distance f (distance focale) du plan rétinien.

On note que le modèle sténopé peut être utilisé dans le cas d'un objectif à lentille munie d'un zoom si la distance de l'objet à l'image peut être considérée comme étant variable et si l'on reste dans un voisinage restreint de la rétine [7]. C'est le cas de notre étude, car la métrique estimée sera locale à une région de la rétine, et locale à un objet, donc adaptée aux différentes positions des objets.

Dans le cas le plus général la **matrice de projection perspective** P 3×4 , qui lie les coordonnées 3D aux coordonnées pixelliques, peut se décomposer de la manière suivante [5] (à condition de satisfaire la contrainte $||[p_{31}p_{32}p_{33}]|| \neq 0$), de manière unique sous la forme:

$$P = A \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} D$$

où A est la **matrice des paramètres intrinsèques** et D la **matrice des paramètres extrinsèques** de la forme:

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

Ici, t est un vecteur 3×1 , 0 est le vecteur nul 1×3 , et R est une matrice 3×3 telle que $R.R^t = I$.

Dans cette décomposition, A est donc une matrice 3×3 triangulaire supérieure de la forme:

$$A = \begin{pmatrix} \alpha_u & -\alpha_u \cot \theta & u_0 \\ 0 & \alpha_v \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

faisant passer du repère rétinien normalisé (coordonnées normalisées m) au repère image (coordonnées pixelliques p), à travers la relation:

$$p = A.m$$

On peut décrire ainsi les 5 **paramètres intrinsèques**:

- α_u, α_v sont les facteurs d'échelle sur les axes. Ils sont reliés à la distance focale en mètres f_g par un facteur de proportionnalité :

$$\alpha_u = K_u \cdot f_g$$

$$\alpha_v = K_v \cdot f_g$$
- (u_0, v_0) est le point principal intersection de l'axe optique et du plan image.
- θ est l'angle relatif entre les deux axes du repère image (u, v) .

Afin d'étudier le modèle de variation des paramètres intrinsèques avec le changement de mise au point ou de focale considérons le **modèle d'une lentille épaisse** d'après [11], schématisé sur la Figure 1.

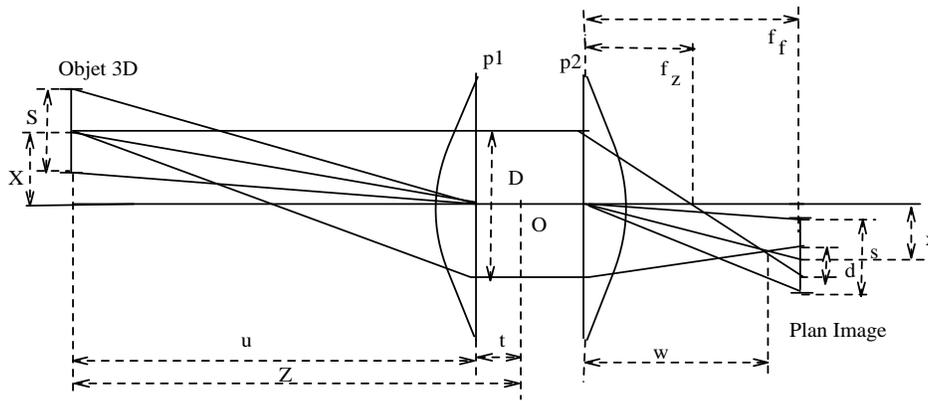


FIG. 1 - Géométrie d'une lentille épaisse

Soient:

- O le centre optique de la lentille,
- $p1$ et $p2$ les plans principaux,
- f_z la **distance focale** et f_f le **focus** (distance entre le plan $p2$ et le plan rétinien) exprimés en mètres,
- u la distance de l'objet 3D au plan principal $p1$, t la distance de $p1$ au centre optique, et finalement $Z = t + u$,
- w la distance du plan $p2$ au plan image,
- D le diamètre de l'ouverture de la lentille et d le diamètre du cercle flou,
- S la taille de l'objet 3D et s la taille de son image.

A l'aide de considérations géométriques simples et de la formule de conjugaison de la lentille $\frac{1}{f_z} = \frac{1}{u} + \frac{1}{w}$ on peut calculer le diamètre du cercle de flou (voir[11]):

$$d = D f_f \left(\frac{1}{f_z} - \frac{1}{u} - \frac{1}{f_f} \right).$$

En plus, comme il a été démontré par [1], ou obtenu par [4], on peut faire les hypothèses réalistes suivantes $Z \gg t$ et $Z \gg f_z$ qui correspond à faire l'**hypothèse d'un objet relativement lointain**, et $D \gg d$ qui correspond à l'hypothèse d'**une mise au point à peu près correcte**.

On a donc approximativement :

$$0 = \frac{1}{f_z} - \frac{1}{Z} - \frac{1}{f_f} \quad (1)$$

On peut alors écrire en considérant l'agrandissement angulaire K :

$$\frac{x}{f_f} = K \frac{X}{Z} \quad (2)$$

Où le grandissement de la lentille K est une fonction homographique de la distance focale [7], i.e :

$$K = \frac{K_a f_z + K_b}{K_c f_z + K_d} \quad (3)$$

et en première approximation K_c est négligeable.

Par ailleurs les quantités physiques f_f et f_z données en mètres ne sont pas directement mesurables, on n'accède qu'à des mesures *focus* et *zoom*, reliées à ces quantités par des gains et offsets inconnus :

$$\begin{aligned} focus &= g_f f_f + o_f \\ zoom &= g_z f_z + o_z \end{aligned} \quad (4)$$

En combinant (1), (2) , (3) et (4) et en posant $x = f_g \cdot \frac{X}{Z}$, on obtient :

$$\begin{aligned} f_g &= K.f_f \\ f_g &= \frac{(K_a.zoom + K_b.g_z - o_z.K_a)(focus - o_f)}{(K_c.zoom + K_d.g_z - o_z.K_c)g_f} \\ f_g &= \frac{(o_z.K_a - K_b.g_z)o_f}{(K_d.g_z - o_z.K_c)g_f} + \frac{(-o_z.K_a + K_b.g_z)}{(K_d.g_z - o_z.K_c)g_f}.focus - \frac{(K_a.K_d - K_c.K_b)g_z.o_f}{(K_d.g_z - o_z.K_c)^2.g_f}.zoom \end{aligned}$$

Ce développement de Taylor au premier ordre est valide, car K_c est petit tandis que la distance de mise au point ou focus varie relativement peu en pratique.

On arrive donc à un modèle **linéaire** pour f_g , (qui va porter le nom de " focale globale ") du type:

$$f_g = \delta.zoom + \beta.focus + \gamma$$

où, finalement nous devons tenir compte d'un facteur de gain, et d'un offset.

En reportant dans le modèle de variation de α_u et α_v on obtient :

$$\begin{cases} \alpha_u = K_u.(\delta.zoom + \beta.focus + \gamma) \\ \alpha_v = K_v.(\delta.zoom + \beta.focus + \gamma) \end{cases}$$

et en posant $k_u = \delta.K_u$, $k_v = \delta.K_v$, $\beta' = \frac{\beta}{\delta}$, $\gamma' = \frac{\gamma}{\delta}$, on obtient finalement :

$$\begin{cases} \alpha_u = k_u.(zoom + \beta'.focus + \gamma') \\ \alpha_v = k_v.(zoom + \beta'.focus + \gamma') \end{cases}$$

3 Influence du changement de focale et de mise au point sur la matrice de projection

On choisit d'étudier le problème dans un cas très particulier, en faisant les 2 hypothèses suivantes:

- **Le changement de mise au point ou de focale revient à une transformation affine** effectuée dans le plan image.
- **Le changement de mise au point n'affecte que la matrice des paramètres intrinsèques A .** Ceci est faux car la lentille "recule" lors d'une mise au point ou lors d'un zoom. Nous avons en fait $t'_3 = t_3 + K_f(f - f')$, où f est soit la quantité mesurée *focus*, soit la quantité *zoom*. Donc il reste à vérifier en pratique que l'on peut représenter cet effet à l'aide de la matrice A . On montre dans la dernière section la variation du centre optique.

La première hypothèse nous amène à introduire une **matrice de transformation affine** en coordonnées projectives: A_{aff} .

La nouvelle **matrice de projection perspective** P' s'écrit alors:

$$P' = A_{aff}.P$$

De la seconde hypothèse, on déduit la relation suivante:

$$A' = A_{aff}.A$$

(où A' représente la nouvelle matrice des paramètres intrinsèques).

La matrice A_{aff} est de la forme:

$$A_{aff} = \begin{pmatrix} m_{11} & m_{12} & a_0 \\ m_{21} & m_{22} & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Il vient alors la relation suivante:

$$A' = A_{aff}.A = \begin{pmatrix} m_{11} & m_{12} & a_0 \\ m_{21} & m_{22} & b_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_u & -\alpha_u \cot \theta & u_0 \\ 0 & \alpha_v \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha'_u & -\alpha'_u \cot \theta' & u'_0 \\ 0 & \alpha'_v \sin \theta' & v'_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'identification des coefficients conduit au système suivant :

$$\begin{cases} m_{11}.\alpha_u & = \alpha'_u \\ m_{21}.\alpha_u & = 0 \\ -m_{11}.\alpha_u \cot\theta + m_{12}.\alpha_v \sin\theta & = -\alpha'_u \cot\theta' \\ -m_{21}.\alpha_u \cot\theta + m_{22}.\alpha_v \sin\theta & = \alpha'_v \sin\theta' \\ m_{11}.u_0 + m_{12}.v_0 + a_0 & = u'_0 \\ m_{21}.u_0 + m_{22}.v_0 + b_0 & = v'_0 \end{cases}$$

où l'on voit, grâce à la deuxième équation, que $m_{21} = 0$, et que l'on peut calculer les nouveaux paramètres intrinsèques en connaissant A_{aff} et les anciens paramètres.

Nos transformations auront donc forcément la forme suivante:

$$A_{aff} = \begin{pmatrix} m_{11} & m_{12} & a_0 \\ 0 & m_{22} & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Elles engendrent un **sous-groupe des matrices de transformation affine**, qui est l'intersection des matrices triangulaires supérieures et des matrices de transformation affine. Ce sous-groupe préserve la direction de l'axe des X .

On note pour simplifier $\alpha_u = k_u.(f - df)$ et $\alpha_v = k_v.(f - df)$ où f est soit le *focus* soit le *zoom*, et df l'offset correspondant¹.

Selon notre modèle, on obtient:

$$\begin{cases} m_{11}.(f - df).k_u & = (f' - df').k'_u \\ m_{21}.(f - df).k_u & = 0 \\ -m_{11}.(f - df).k_u \cot\theta + m_{12}.(f - df).k_v \sin\theta & = -(f' - df').k'_u \cot\theta' \\ -m_{21}.(f - df).k_u \cot\theta + m_{22}.(f - df).k_v \sin\theta & = (f' - df').k'_v \sin\theta' \\ m_{11}.u_0 + m_{12}.v_0 + a_0 & = u'_0 \\ m_{21}.u_0 + m_{22}.v_0 + b_0 & = v'_0 \end{cases}$$

1. Dans le cas du *zoom* nous avons k_u, k_v comme décrit précédemment et $df_z = -(\beta'.focus + \gamma')$. Dans le cas du *focus* on a $k_u = \beta'.k_u$, $k_v = \beta'.k_v$, et $df_f = -\frac{zoom + \gamma'}{\beta'}$

donc dans le cas général:

$$\begin{cases} m_{11} & = \frac{(f' - df')}{(f - df)} \cdot \frac{k'_u}{k_u} \\ m_{12} & = \frac{(f' - df')}{(f - df)} \cdot \frac{k'_u}{k_v} \cdot \frac{(\cot\theta - \cot\theta')}{\sin\theta} \\ m_{21} & = 0 \\ m_{22} & = \frac{(f' - df')}{(f - df)} \cdot \frac{k'_v}{k_v} \cdot \frac{\sin\theta'}{\sin\theta} \\ m_{11} \cdot u_0 + m_{12} \cdot v_0 + a_0 & = u'_0 \\ m_{22} \cdot v_0 + b_0 & = v'_0 \end{cases}$$

4 Étude de l'invariance de k_u , k_v , df et θ par changement de mise au point et de focale

Les paramètres k_u , k_v et df étant des paramètres liés uniquement à la mécanique de la lentille, il paraît raisonnable de les considérer comme invariants lors d'un changement de mise au point. De plus, θ est en pratique très proche de $\frac{\pi}{2}$, donc on peut négliger ses variations ([9]). On étudie dans cette section la validité de ces hypothèses.

4.1 Invariance de l'angle θ

La matrice A_{aff} dans le cas général est:

$$A_{aff} = \begin{pmatrix} \frac{\alpha'_u}{\alpha_u} & \frac{\alpha'_u}{\alpha_v} \cdot \frac{(\cot\theta - \cot\theta')}{\sin\theta} & a_0 \\ 0 & \frac{\alpha'_v}{\alpha_v} \cdot \frac{\sin\theta'}{\sin\theta} & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Étudions pour une petite variation $d\theta$ de l'angle relatif entre les deux axes, comment se transforment les équations.

Soit l'angle relatif entre les deux axes sur la première image θ . On note $\theta = \frac{\pi}{2} + \epsilon$.

Et soit l'angle sur la deuxième image $\theta' = \frac{\pi}{2} + \epsilon + d\epsilon$. Donc:

$$\frac{\sin\theta'}{\sin\theta} = \frac{\sin(\frac{\pi}{2} + \epsilon + d\epsilon)}{\sin(\frac{\pi}{2} + \epsilon)} = \frac{\cos(\epsilon + d\epsilon)}{\cos\epsilon} = \cos(d\epsilon) - \tan(\epsilon) \cdot \sin(d\epsilon)$$

Par développements de Taylor on trouve:

$$\frac{\sin\theta'}{\sin\theta} \approx 1 - d\epsilon \cdot \epsilon + \frac{d\epsilon^2}{2} + \Phi$$

où Φ est d'ordre 4 en $d\epsilon$ et ϵ .

D'un autre côté,

$$\frac{(\cot\theta - \cot\theta')}{\sin\theta} \approx d\epsilon + \Phi$$

avec Φ d'ordre 3 en $d\epsilon$ et ϵ .

Il paraît raisonnable de conclure $m_{22} = \frac{\alpha_u'}{\alpha_v}$ et $m_{12} = 0$ dès que $d\epsilon$ ou ϵ est petit.

Si l'on fait donc l'hypothèse de l'invariance de θ on obtient:

$$A_{aff} = \begin{pmatrix} \frac{\alpha_u'}{\alpha_u} & 0 & a_0 \\ 0 & \frac{\alpha_v'}{\alpha_v} & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

avec 4 inconnues.

4.2 Invariance de k_u , k_v , df

Faisons maintenant une nouvelle hypothèse, à savoir que k_u , k_v , df (paramètres liés à la mécanique de la lentille et au système d'acquisition) ne sont pas modifiés lors du changement de mise au point.

Le système devient donc:

$$\left\{ \begin{array}{l} m_{11} \\ m_{21} \\ m_{12} \\ m_{22} \\ m_{11} \cdot u_0 + m_{12} \cdot v_0 + a_0 \\ m_{21} \cdot u_0 + m_{22} \cdot v_0 + b_0 \end{array} \right. \begin{array}{l} = \frac{(f' - df)}{(f - df)} \\ = 0 \\ = \frac{(f' - df)}{(f - df)} \cdot \frac{k_u}{k_v} \cdot \frac{(\cot\theta - \cot\theta')}{\sin\theta} \\ = \frac{(f' - df)}{(f - df)} \cdot \frac{\sin\theta'}{\sin\theta} \\ = u_0' \\ = v_0' \end{array}$$

Puis finalement, si l'on appelle $C_0 = \frac{(f'-df)}{(f-df)}$:

$$A_{aff} = \begin{pmatrix} C_0 & C_0 \cdot \frac{k_u}{k_v} \cdot \frac{(\cot\theta - \cot\theta')}{\sin\theta} & a_0 \\ 0 & C_0 \frac{\sin\theta'}{\sin\theta} & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

4.3 Invariance de k_u , k_v , df et l'angle θ

Si l'on se place dans le cas particulier où ces paramètres sont invariants :

$$A_{aff} = \begin{pmatrix} C_0 & 0 & a_0 \\ 0 & C_0 & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Deux remarques:

- Il s'agit d'une homothétie plus une translation, donc elle conserve les angles.
- On peut voir que l'on avait au départ 13 inconnues, et 6 équations. Le fait de supposer l'invariance de k_u , k_v , df et θ a fourni 5 relations supplémentaires, puisque $\theta = \theta'$ nous donne $\cot\theta = \cot\theta'$ et $\sin\theta = \sin\theta'$. Il nous reste donc $13-6-5 = 2$ indéterminées, que l'on retrouve bien exprimées sous la forme des 2 relations suivantes:

$$\begin{cases} C_0 \cdot u_0 + a_0 & = & u'_0 \\ C_0 \cdot v_0 + b_0 & = & v'_0 \end{cases}$$

4.4 Calcul de la transformation avec hypothèses d'invariance en effectuant le suivi de 2 points

Puisqu'il ne reste plus que 3 inconnues: C_0, a_0, b_0 , et que chaque point fournit 2 équations, il suffit de suivre deux points.

On considère deux points en coordonnées projectives $P_1 = (u_1, v_1, \xi)$ et $P_2 = (u_2, v_2, \xi)$ où $\xi \in \{0, 1\}$. On a $\xi = 1$ pour les points à distance finie et $\xi = 0$ pour les points à l'infini, c'est à dire les directions de droite. Notre calcul sera donc valable pour des points et des points à l'infini.

Soient P'_1, P'_2 leurs correspondants sur la deuxième image. Donc:

$$A_{aff}.P_1 = P'_1,$$

$$A_{aff}.P_2 = P'_2$$

nous donne les 4 équations suivantes:

$$\begin{cases} u'_1 &= C_0 \cdot u_1 + a_0 \cdot \xi \\ v'_1 &= C_0 \cdot v_1 + b_0 \cdot \xi \\ u'_2 &= C_0 \cdot u_2 + a_0 \cdot \xi \\ v'_2 &= C_0 \cdot v_2 + b_0 \cdot \xi \end{cases}$$

En soustrayant la troisième équation à la première, ainsi que la quatrième à la deuxième, il vient:

$$C_0 = \frac{u'_1 - u'_2}{u_1 - u_2} = \frac{v'_1 - v'_2}{v_1 - v_2} \Rightarrow \frac{v'_1 - v'_2}{u'_1 - u'_2} = \frac{v_1 - v_2}{u_1 - u_2}$$

où l'on voit bien que cette implication correspond au fait que les angles sont conservés par cette homothétie.

Et dans le cas des points à l'infini on obtient:

$$C_0 = \frac{u'_1}{v'_1} = \frac{u_1}{v_1} = \frac{u'_2}{v'_2} = \frac{u_2}{v_2}.$$

On doit donc suivre **au moins** un point à distance finie et une direction, ou deux points à distance finie, pour pouvoir calculer notre transformation.

5 Étude expérimentale de l'influence de la mise au point et de la focale sur les paramètres

Avant d'étudier expérimentalement notre modèle examinons la validité de nos hypothèses.

Les données proposées sont le résultat de 10 calibrations successives où l'on fait varier le *focus* mesuré dans l'intervalle $\{0.1, 0.2, \dots, 1.0\}$.

Nous avons aussi complétés ces résultats par un mesure de 5 calibrations successives en faisant varier la focale mesurée (*zoom*) entre $\{0.2, 0.4, 0.6, 0.8, 1.0\}$.

A chaque prise d'image on a calibré grâce à la mire :

- la valeur des paramètres intrinsèques: $\alpha_u, \alpha_v, u_0, v_0$,
- l'angle relatif entre les deux axes θ ,
- la position du centre optique dans le repère caméra,

ceci en utilisant la méthode développée par [12].

5.1 Stabilité des mesures des paramètres par rapport à la lumière dans la salle

Notre premier essai avait pour but de vérifier la stabilité des résultats de la calibration par rapport à la lumière présente dans la salle.

On peut noter dans la plupart des graphiques suivants que malgré la disparité des résultats, le comportement global des paramètres est le même en présence d'un fort éclairage (*avec*) que d'un plus faible (*sans*). Cette stabilité est liée à la méthode utilisée.

On peut remarquer les résultats suivants:

- On note $\theta = \frac{\pi}{2} + \epsilon$. On a représenté ϵ (Figure 2), et on peut noter que dans le cas avec lumière, on obtient des valeurs pour ϵ plus centrés autour de 0 et avec une plus petite variance. Dans la section suivante on a vérifié cette hypothèse en approximant aux moindres carrés.
- α_u et α_v décroissent linéairement (Figure 3).

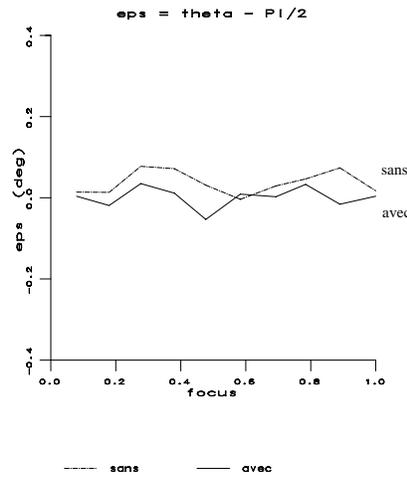


FIG. 2 - Évolution de ϵ par rapport à la présence ou absence de lumière

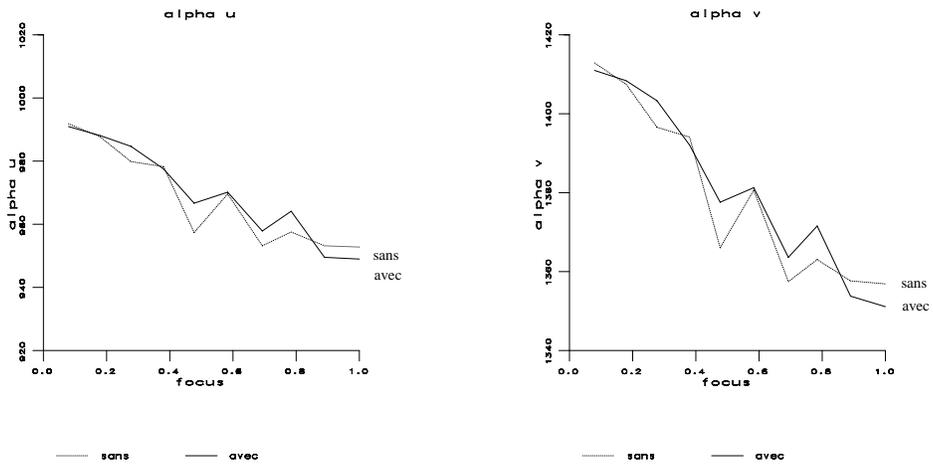


FIG. 3 - Évolution de α_u et α_v par rapport à la présence ou absence de lumière

- On a utilisé la même échelle pour représenter u_0 et v_0 . Le deuxième croît deux fois plus vite que le premier. (Figure 4).

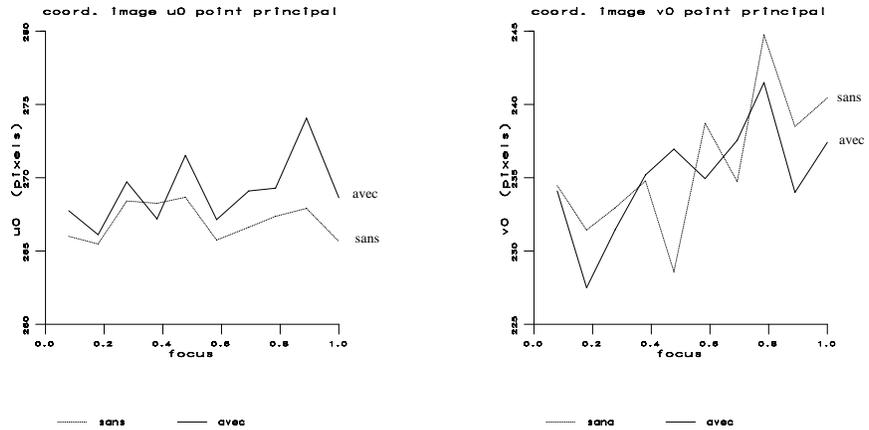
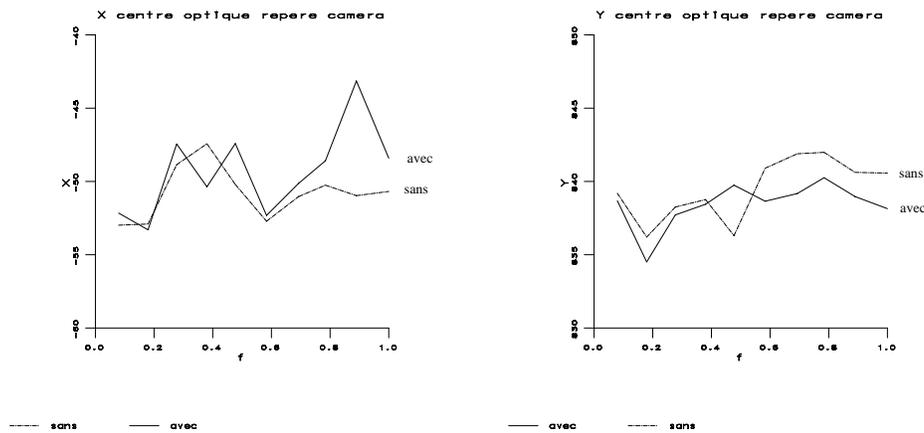


FIG. 4 - Évolution du point principal par rapport à la présence ou absence de lumière

- Le centre optique varie globalement de la même manière dans les deux cas (Figure 5).



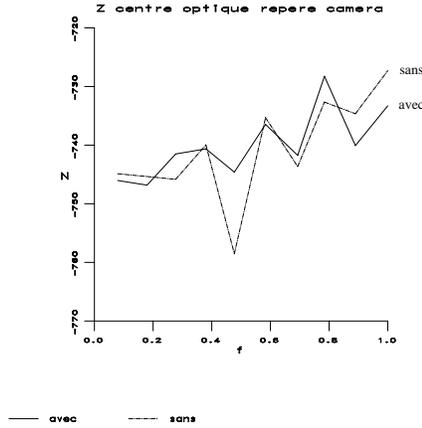


FIG. 5 - Évolution du centre optique par rapport à la présence ou absence de lumière

5.2 Précalibration du zoom et de la mise au point

Comme la lentille n'est calibrée ni au niveau du zoom, ni au niveau de la mise au point, il faut calculer les gains et les offsets de notre modèle.

On va donc pour cela mesurer α_u et α_v , en faisant varier le *focus* d'abord, puis dans une seconde série de mesures en faisant varier le *zoom*. On a représenté :

$$\begin{aligned} \alpha_u &= k_u \cdot focus \text{ avec } focus = f_f - df_f \\ \alpha_u &= k_v \cdot zoom \text{ avec } zoom = f_z - df_z \end{aligned}$$

De même pour α_v .

Les conclusions les plus intéressantes sont :

- Le rapport $\frac{\alpha_u}{\alpha_v}$ est invariant (la pente de la droite calculée par moindres carrés étant presque 0). Ce rapport est 0.702 pour le focus et 0.706 pour le zoom. Voir Figure 6.

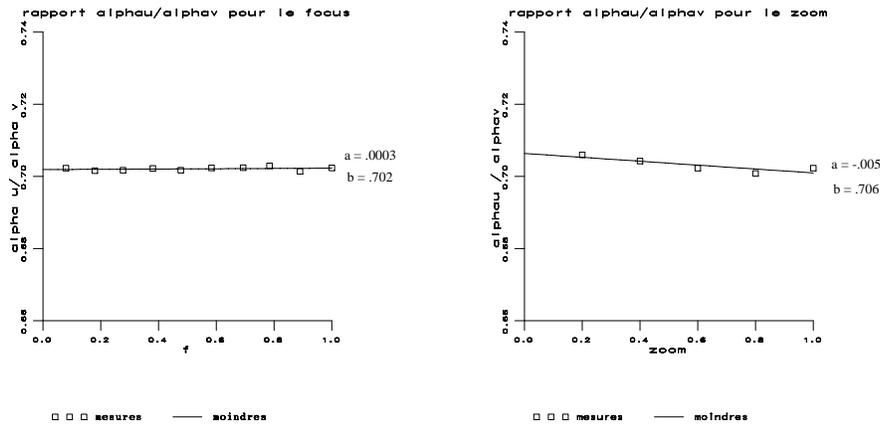


FIG. 6 - Rapport $\frac{\alpha_u}{\alpha_v}$ en fonction de la variation de la mise au point et du zoom

De plus, les données constructeur étant de 0.7, nous sommes fort proche de cette valeur théorique. Notons que ce résultat est valable dans le cas d'une variation soit de mise au point soit de zoom.

- On a mesuré les offsets df_f , donc **calibré la mise au point** (Figure 7).

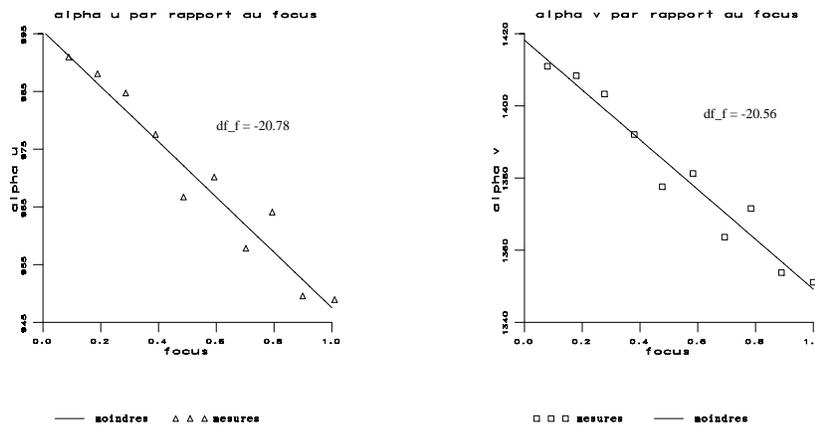


FIG. 7 - α_u et α_v par rapport à la variation de la mise au point

Par interpolation linéaire on peut écrire:

$$\begin{aligned}\alpha_u &= -47.9(f_f - df_f) + 995.43 \\ \alpha_v &= -68.97(f_f - df_f) + 1418.21\end{aligned}$$

Quand $\alpha_u = 0$, $f_f = 0$ on obtient:

$$\begin{aligned}df_f &\approx -20.78 \text{ pour } \alpha_u \text{ et} \\ df_f &\approx -20.56 \text{ pour } \alpha_v.\end{aligned}$$

Donc on peut conclure que df_f **n'est pas modifié lors d'un changement de mise au point**, et sa valeur est quasiment égale sur les deux axes comme attendu.

- On a **calibré le zoom** par la même méthode et $df_z \approx -0.04$, de même, **reste invariant sur les deux axes**. Voir Figure 8.

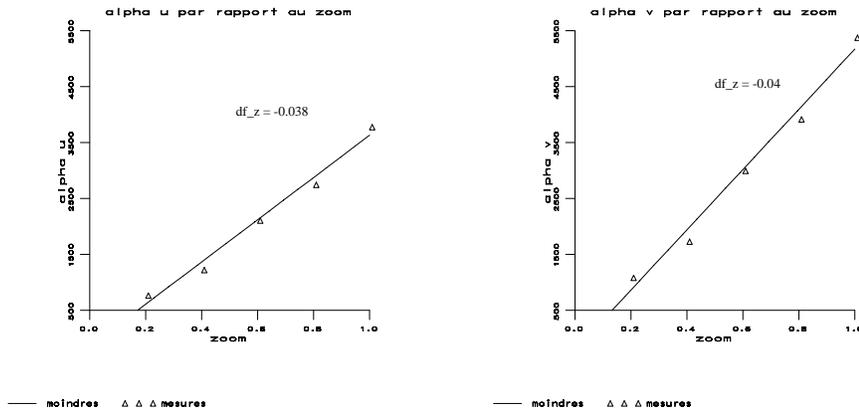


FIG. 8 - α_u et α_v par rapport à la variation du zoom

- On ne peut pas déduire grande chose à partir des graphiques sur le déplacement du point principal (u_0, v_0) (Figure 9), mais on constate néanmoins une **variation linéaire** de ces paramètres.

$$\begin{aligned}u_0 &= 3.68 * focus + 267 \\ v_0 &= 7.72 * focus + 230.9\end{aligned}$$

Cette variation n'est pas prévue par le modèle initial mais ce phénomène a été très précisément étudié par [18] et nous n'y reviendrons pas. Il y est établi qu'un modèle linéaire de variation des coordonnées du point principal en fonction de la mise au point est valide. Les variations de (u_0, v_0) en fonction du zoom sont plus complexes et nécessitent un modèle quadratique, mais nous n'aborderons pas ce point déjà étudié [18].

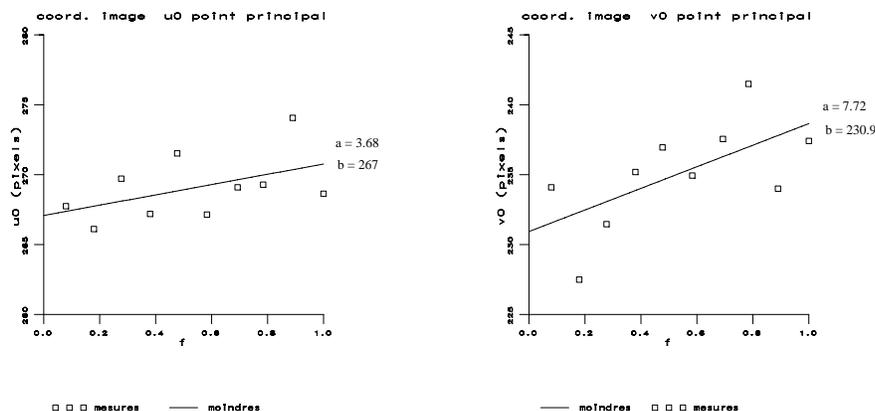


FIG. 9 - Évolution du point principal (u_0, v_0) en fonction de la mise au point

- L'angle θ est égal à $\frac{\pi}{2}$ à moins de 0,05 degrés près, dans le cas des mesures avec fort éclairage.
On note $\theta = \frac{\pi}{2} + \epsilon$. On a obtenu les droites d'approximation suivantes:

$$\begin{aligned} \epsilon &= 0.0025 * focus - 0.0004 \text{ avec forte lumière, et} \\ \epsilon &= 0.0074 * focus + 0.033 \text{ sans lumière.} \end{aligned}$$

Ceci est encore tout à fait valide pour les variations de focale.

Par conséquent cela **valide notre hypothèse d'invariance de l'angle** dans les deux cas et montre la plus grande disparité des mesures dans le cas du faible éclairage (Figure 10).

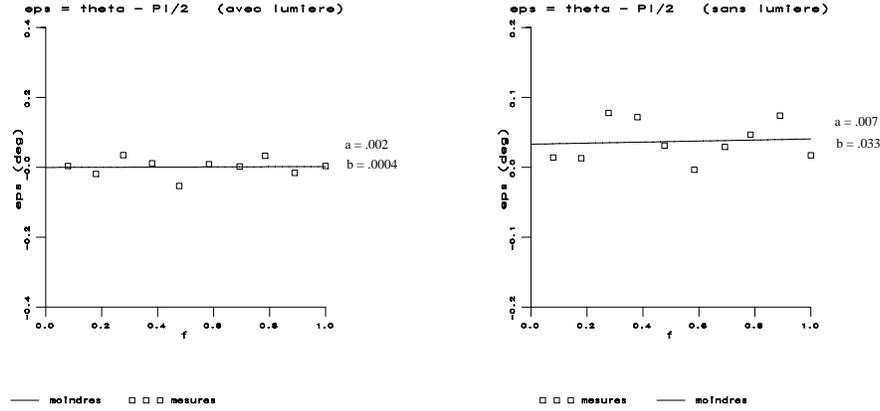


FIG. 10 - ϵ par rapport à la variation de la mise au point

- Le centre optique se déplace surtout le long de l'axe Z .
On a représenté les coordonnées du centre optique dans le repère caméra sur les trois axes X_{co} , Y_{co} , Z_{co} , et on a approximé par une droite aux moindres carrés, d'abord avec toutes les mesures, et après sans le point considéré aberrant (Figure suivante).
Le tableau suivant représente la pente des différentes droites obtenues:

Pente des droites aux moindres carrés		
Coord	Toutes mesures	Sans point aberrant
X_{op}	5.58	3.28
Y_{op}	2.35	0.66
Z_{op}	13.78	10.89

Donc on pourrait conclure que effectivement le centre optique se déplace surtout au long de l'axe Z au cours d'un changement de mise au point, bien que la variation en X et Y ne soit pas négligeable.

Par conséquent la mise au point doit vraisemblablement décaler le centre optique lors de sa variation, ce qui explique aussi la translation de u_0 et v_0 , comme attendu [18].

Un effet similaire a été observé pour les variations de focale.

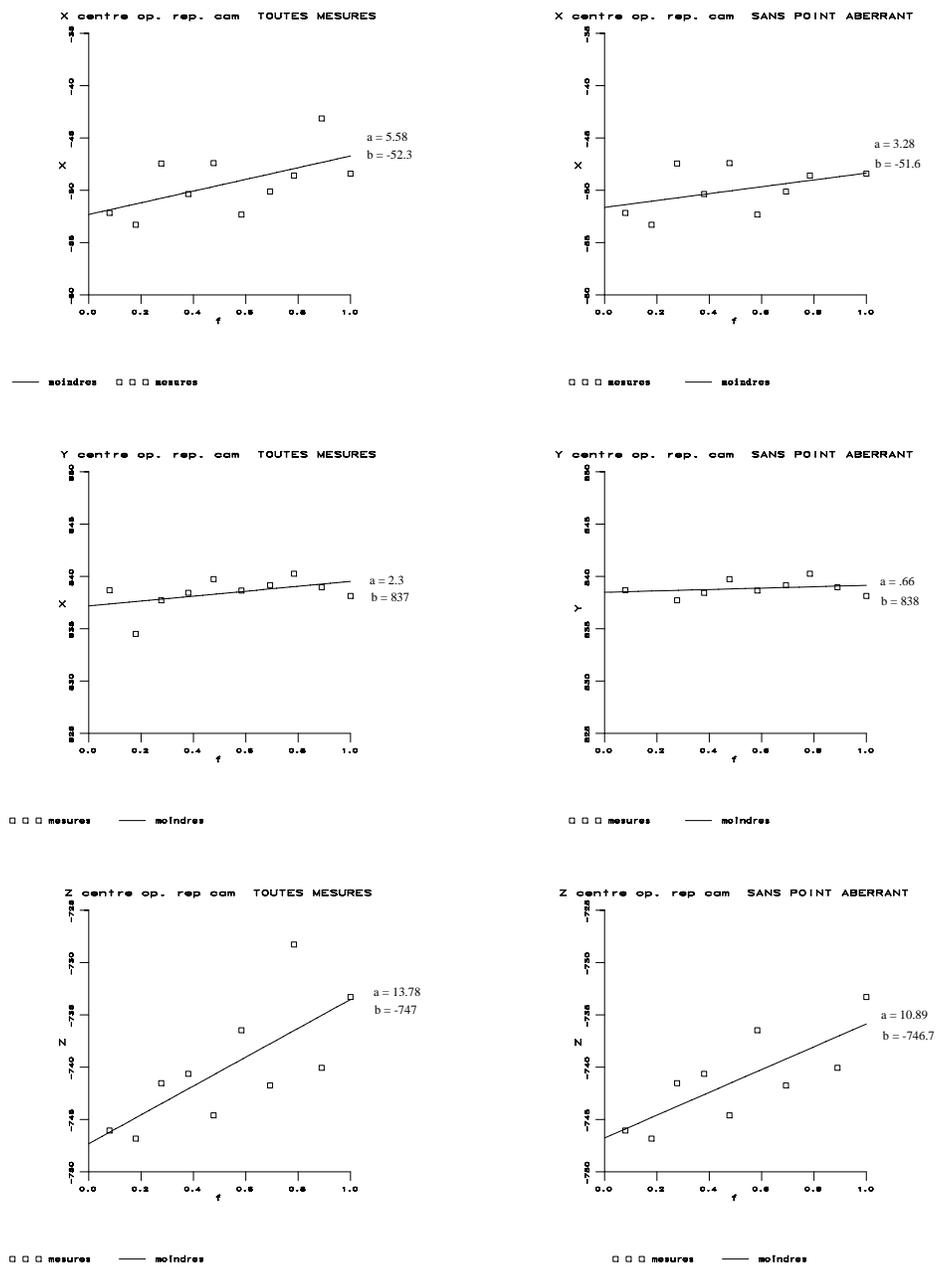


FIG. 11 - Centre optique avec toutes les mesures et sans le point aberrant

6 Résultats expérimentaux

Dans cette section on étudie comment estimer notre modèle lors du suivi de points d'intérêt.

6.1 Calcul de la matrice affine avec le nombre minimum de points

On utilise la méthode de [2] qui permet de mesurer la position de points d'intérêt avec une très grande précision, même dans le cas où l'image est floue. On extrait ces points sur deux images consécutives et on les apparie manuellement dans cette expérimentation.

A partir des hypothèses d'invariance et **en suivant deux points** on obtient les équations suivantes:

$$\begin{cases} C_0 &= \frac{u'_1 - u'_2}{u_1 - u_2} &= \frac{v'_1 - v'_2}{v_1 - v_2} \\ a_0 &= u'_1 - C_0 \cdot u_1 &= u'_2 - C_0 \cdot u_2 \\ b_0 &= v'_1 - C_0 \cdot v_1 &= v'_2 - C_0 \cdot v_2 \end{cases}$$

On a calculé la matrice de transformation noté *avec hypothèses*:

$$A_{hyp} = \begin{pmatrix} C_0 & 0 & a_0 \\ 0 & C_0 & b_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

D'un autre côté, on a calculé la matrice de transformation affine à 6 paramètres (noté *full matrix*) en **suivant 3 points**:

$$A_{full} = \begin{pmatrix} m_{11} & m_{12} & a_0 \\ m_{21} & m_{22} & b_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Le graphique suivant (Figure 12) montre l'évolution de l'erreur dans la localisation des points image à partir des matrices de transformation calculées par les deux méthodes: *avec les hypothèses* $m_{12} = 0$, $m_{21} = 0$, et $m_{11} = m_{22} = C_0$, et celle noté "*full matrix*" qui calcule les 6 paramètres de la transformation. Ceci à partir de deux correspondances pour la première

méthode et de 3 pour la deuxième.

On peut noter que la méthode *avec hypothèses* est plus robuste avec un petit nombre des points car elle estime moins de paramètres, et reste toujours à moins de 2 pixels d'erreur.

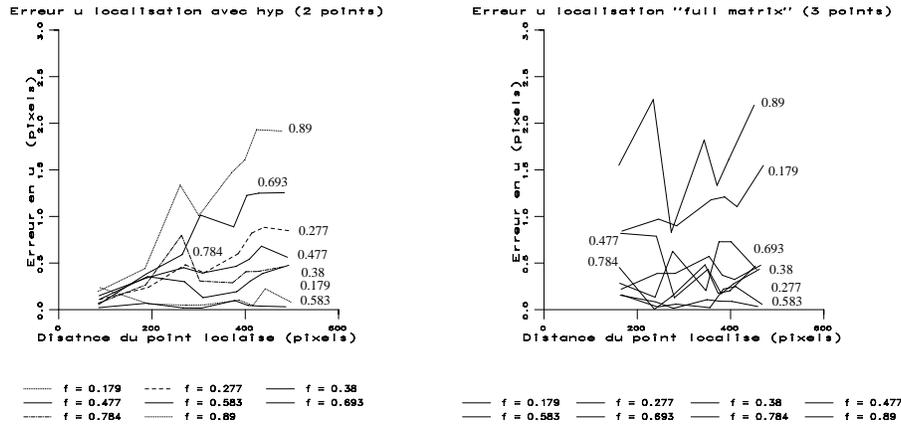


FIG. 12 - Évolution de l'erreur dans la localisation des points image par les 2 méthodes: avec les hypothèses d'invariance à partir de 2 points et "full matrix" à partir de 3 points de l'image

6.2 Minimisation de l'erreur aux moindres carrés

On note max le maximum des points disponibles ayant été appariés entre deux images, et

$\sum = \sum_{i=1}^{max}$ l'opérateur de sommation sur les quantités vectorielles u, v, u', v' .

Par la première méthode (A_{hyp}), en minimisant aux moindres carrés:

$$\sum_{i=1}^{max} (u'_i - [C_0 \cdot u_i + a_0])^2 + (v'_i - [C_0 \cdot v_i + b_0])^2$$

on obtient en dérivant par rapport aux trois paramètres (a_0, b_0, C_0):

$$\begin{cases} \sum u'_i u + \sum v'_i v &= C_0 \cdot [\sum u^2 + \sum v^2] + \sum u \cdot a_0 + \sum v \cdot b_0 \\ \sum u'_i &= C_0 \cdot \sum u + a_0 \cdot \sum 1 \\ \sum v'_i &= C_0 \cdot \sum v + b_0 \cdot \sum 1 \end{cases}$$

et finalement:

$$\left\{ \begin{array}{l} C_0 = \frac{\sum u' u \cdot \sum 1 + \sum v' v \cdot \sum 1 - \sum u' \cdot \sum u - \sum v' \cdot \sum v}{\sum u^2 \cdot \sum 1 + \sum v^2 \cdot \sum 1 - (\sum u)^2 - (\sum v)^2} \\ a_0 = \frac{\sum u' - C_0 \cdot \sum u}{\sum 1} \\ b_0 = \frac{\sum v' - C_0 \cdot \sum v}{\sum 1} \end{array} \right.$$

En minimisant au sens des moindres carrés par la deuxième méthode (A_{full}) à 6 paramètres:

$$\sum_{i=1}^{max} (u'_i - [m_{11}u_i + m_{12}v_i + a_0])^2 + (v'_i - [m_{21}u_i + m_{22}v_i + b_0])^2$$

et en calculant la dérivée partielle par rapport à chacun de nos paramètres ($m_{11}, m_{12}, a_0, m_{21}, m_{22}, b_0$) on obtient les équations linéaires suivantes:

$$\left\{ \begin{array}{l} \sum u' u = m_{11} \cdot \sum u u + m_{12} \cdot \sum u v + a_0 \cdot \sum u \\ \sum u' v = m_{11} \cdot \sum u v + m_{12} \cdot \sum v v + a_0 \cdot \sum v \\ \sum u' = m_{11} \cdot \sum u + m_{12} \cdot \sum v + a_0 \cdot \sum 1 \\ \sum v' u = m_{21} \cdot \sum u u + m_{21} \cdot \sum u v + b_0 \cdot \sum u \\ \sum v' v = m_{21} \cdot \sum u v + m_{22} \cdot \sum v v + b_0 \cdot \sum v \\ \sum v' = m_{21} \cdot \sum u + m_{22} \cdot \sum v + b_0 \cdot \sum 1 \end{array} \right.$$

que l'on résoud aisément de manière analytique.

Le graphique suivant montre l'évolution de l'erreur dans la localisation des points image à partir des matrices de transformation calculées par les deux méthodes avec 5 et 8 correspondances. (Figure 13).

On a aussi représenté les paramètres de la transformation affine calculés à partir de 2 et 3, 5 et 8 points par les deux méthodes décrites, (Figure 14, 15 et 16).

On peut facilement conclure que nos hypothèses $m_{12} = 0$, $m_{21} = 0$, et $m_{11} = m_{22} = C_0$, se vérifient, car avec la méthode A_{full} on a pu obtenir :

$$\begin{array}{l} m_{12} \approx 0 \text{ et } m_{21} \approx 0 \text{ à } 10^{-3} \text{ près, comme montré Figure 16, et} \\ m_{11} \approx m_{22} \text{ à } 10^{-2} \text{ près} \end{array}$$

comme montré Figure 15, tandis que avec la méthode A_{hyp} on obtient des résultats aussi précis, et quasiment égaux, (voir Figure 14, u 15, 16).

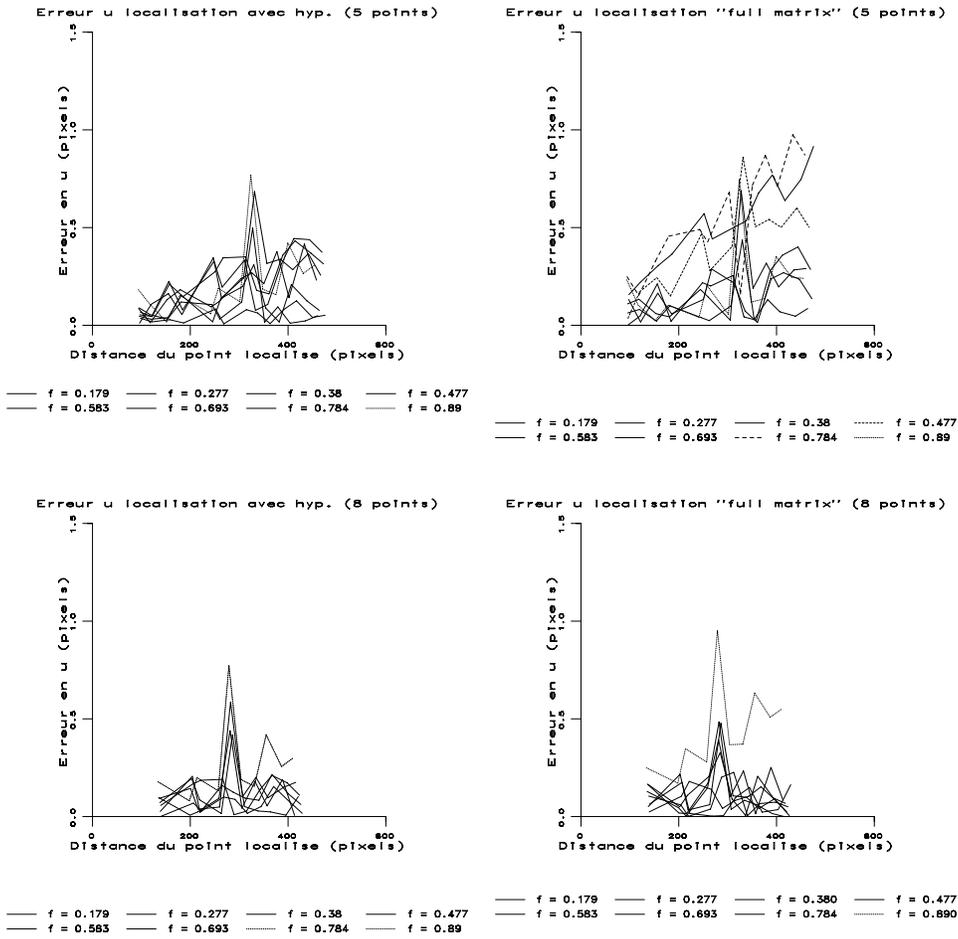


FIG. 13 - Évolution de l'erreur u dans la localisation calculé à partir de 5 et 8 points par moindre carrés

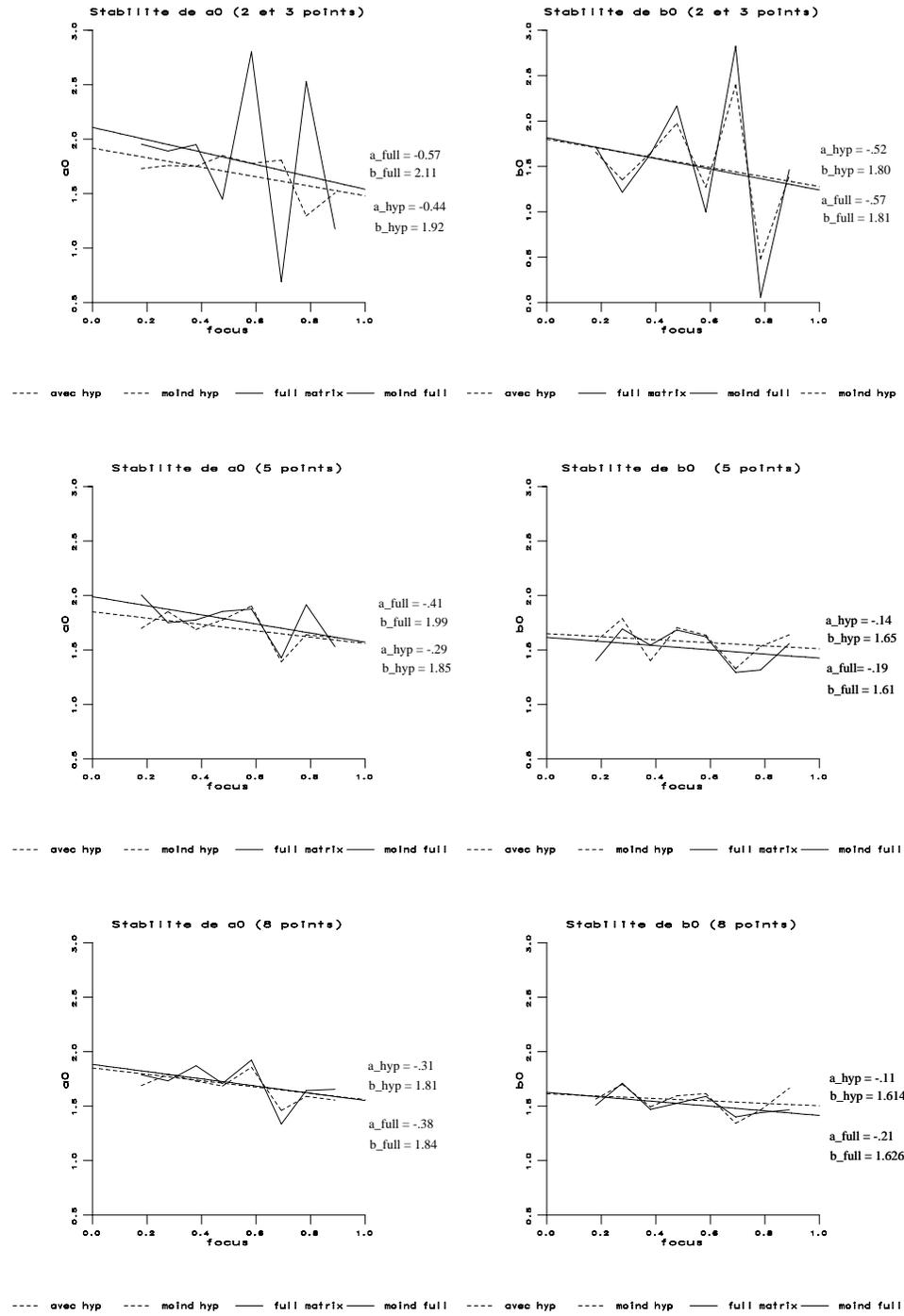


FIG. 14 - Évolution de a_0 et b_0 avec 2 et 3, 5 et 8 points

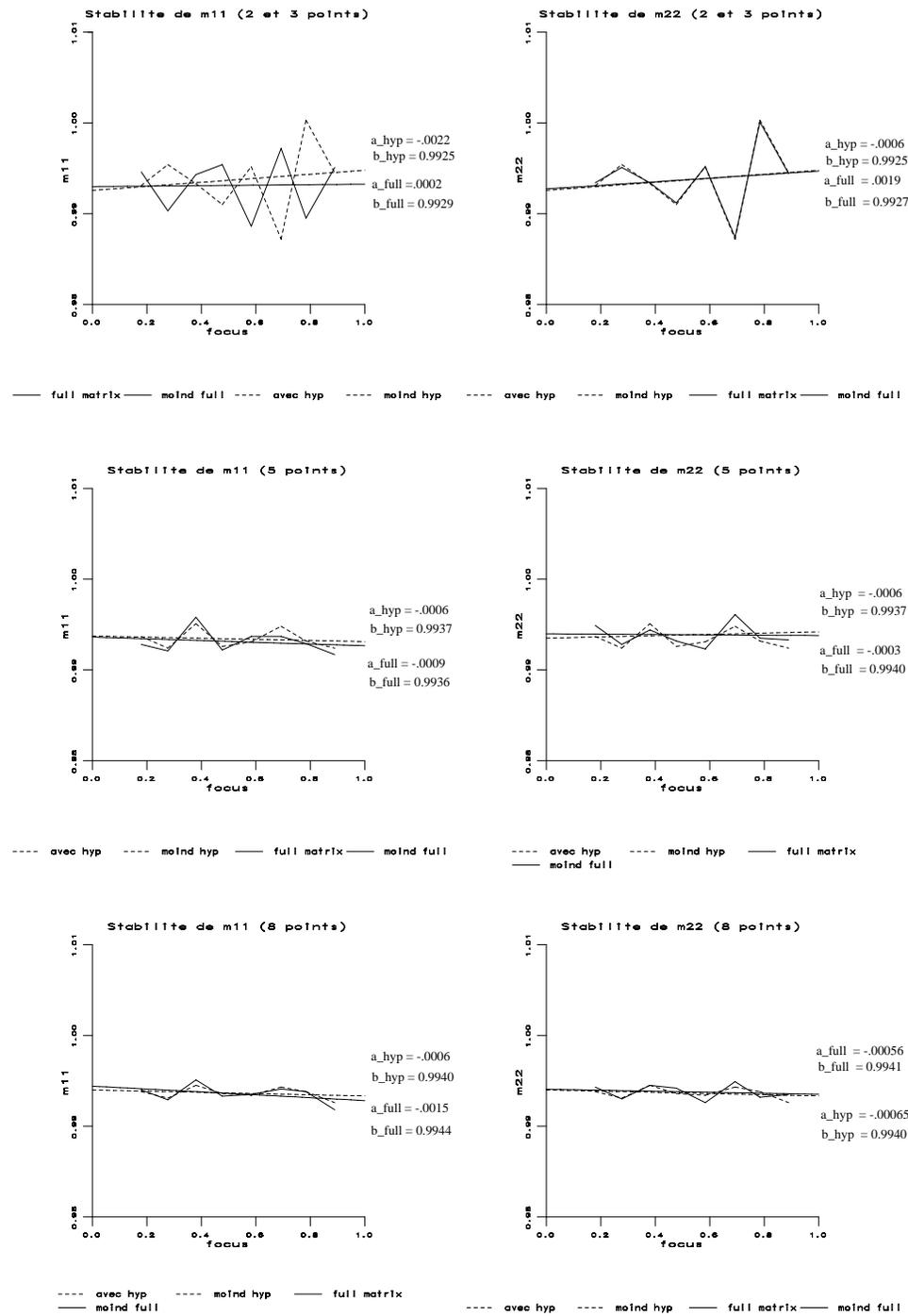


FIG. 15 - Évolution de m_{11} et m_{22} avec 2 et 3, 5 et 8 points

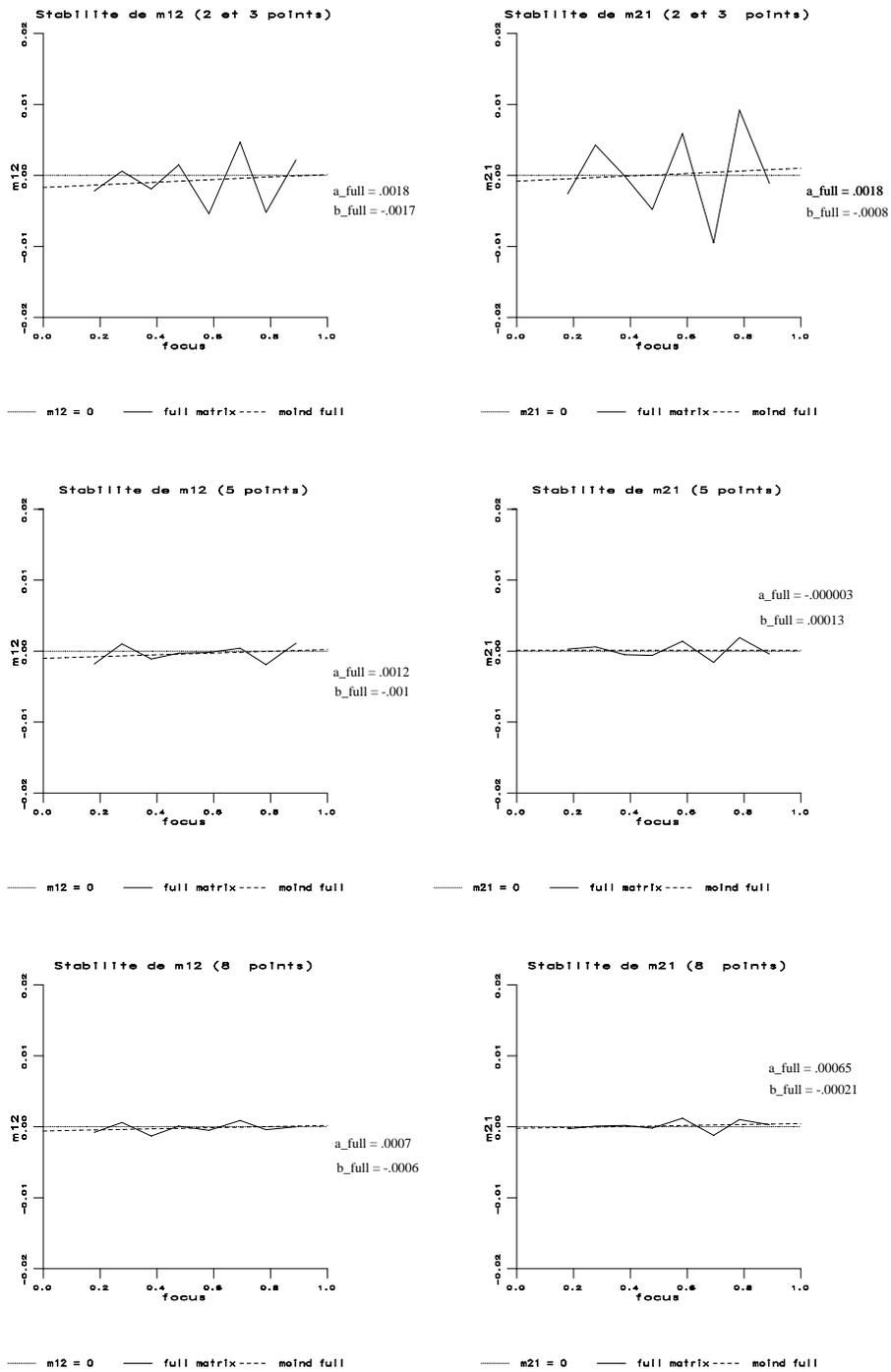


FIG. 16 - Évolution de m_{12} et m_{21} avec 2 et 3, 5 et 8 points

7 Conclusion

Ce travail préliminaire à une étude dynamique de l'autocalibration permet de cerner la complexité du problème et de prévoir comment aborder la construction effective d'algorithmes d'autocalibration.

En effet, nous avons montré deux choses : tout d'abord il est possible de *précalibrer* avec une assez forte précision une lentille dotée d'un zoom et

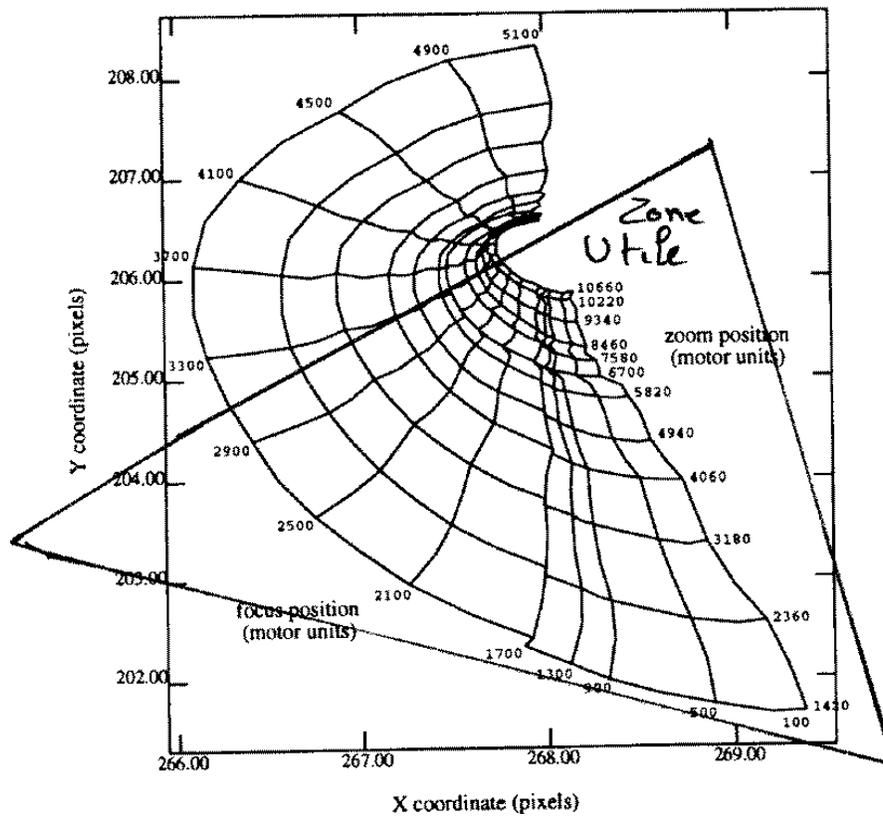


FIG. 17 - Évolution de u_0 et v_0 en fonction du zoom et de la mise au point [18]

d'un mécanisme de mise au point; de plus il est possible de suivre les modifications de ces paramètres, de *recalibrer* à partir de variations simples du modèle initial.

A partir des travaux de [18] dont nous donnons copie des résultats dans la figure 17, il est clair que le centre optique décrit un réseau de droites à peu près concourantes. Pour la partie qui nous intéresse les variations du centre optique en fonction de la mise au point sont approximativement linéaires.

En fait, bien que cet auteur trouve une variation assez complexe du point principal lors de changement de la mise au point, nos résultats montrent que pour la lentille considérée ici, la variation est linéaire.

Ceci est vraisemblablement relié à deux phénomènes : (1) la qualité mécanique de la lentille, et (2) la plage de variation dont nous avons besoin.

Un modèle réaliste de calibration d'une lentille est donc de la forme :

$$\begin{aligned}\alpha_u &= k_u \cdot (\text{zoom} + \beta' \cdot \text{focus} + \gamma') \\ \alpha_v &= k_v \cdot (\text{zoom} + \beta' \cdot \text{focus} + \gamma') \\ u_0 &= u_0^0 + u_0^1 \text{zoom} + u_0^2 \text{focus} \\ v_0 &= v_0^0 + v_0^1 \text{zoom} + v_0^2 \text{focus}\end{aligned}$$

et les acquisitions réalisées ici permettent de déterminer tous ces paramètres.

Par ailleurs, il est clair que ce modèle n'est pas valide pour une lentille particulière mais reflète des caractéristiques cohérentes avec ce que d'autres auteurs ont pu établir.

Ainsi, pour un autre système optique, Hervé Mathieu a t'il pu relever les données suivantes, en utilisant un objectif Canon $J8 \times 6B48$ de haute qualité avec une focale de $6mm$ à $48mm$:

focal	u_0	v_0	au	av	au/av
06	262.66	248.05	128.63	190.10	0.6766
07	252.76	249.89	xxx	xxx	xxx
08	256.33	249.07	126.85	189.60	0.6690
09	253.77	241.16	462.17	660.38	0.6998
10	258.86	251.62	556.26	793.67	0.7009
12	252.85	242.48	668.74	953.09	0.7016
14	251.40	242.41	783.06	1114.55	0.7026
16	252.73	265.28	868.10	1235.35	0.7027
25	244.85	291.51	1427.20	2033.12	0.7020
30	247.81	299.21	1742.43	2482.23	0.7018

dont l'analyse a montré les mêmes phénomènes de variation linéaire.

De même, comme nous l'avons signalé en introduction, le modèle sténopé est valide parce que nous l'utilisons de manière locale comme recommandé par [7]. Ainsi dans le cas de forte focale avons nous dû bien limiter l'étendue de la grille de calibration sur la rétine et ne pas lui faire couvrir tout le champ mais seulement une zone centrale ("fovéale") de façon à ce que la calibration reste valide.

Ainsi ce mécanisme va permettre de maintenir de manière approchée mais néanmoins assez précise, les paramètres intrinsèques d'une lentille d'un système actif. Ce module, opérationnel, a été implémenté sur la tête de l'INRIA et est utilisé dans d'autres travaux de vision active.

Par ailleurs, nous avons la possibilité *sans connaître explicitement les paramètres du système* de remettre à jour ce modèle de calibration étant donné : (a) un modèle initial, (b) un ensemble de correspondances. Bien que notre travail soit valide uniquement dans le cas où le système - immobile - observe une scène fixe et ne modifie que la focale et/ou la mise au point, il démontre néanmoins que certains cas particuliers valent la peine d'être exploités car ils correspondent à des algorithmes très simples et il est facile de ne déclencher ces algorithmes que dans les cas idoines.

Il va être très facile de généraliser ces travaux à des modèles plus complexes en considérant 2 ou 3 vues [9] où une séquence de vues [17]. Dans ce cas, les algorithmes potentiels promettant d'être hautement non-linéaires [8] et dotés d'un rayon de convergence assez faible (on doit être proche de la solution finale pour qu'il y ait convergence) [17], il sera d'autant plus important de se baser sur une estimée initiale précise, comme obtenue ici.

Références

- [1] M. Anderson. Range from out-of-focus blur. Technical Report LIU-TEK-LIC-1992:17, Linköping University, 1992.
- [2] T. Blaszk. Modélisation et caractérisation des variations d'intensité dans une image. Rapport de Stage de DEA,, Septembre 1992.
- [3] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37, 1971.
- [4] S. Das and N. Ahuja. A comparative study of stereo, vergence and focus as depth cues for active vision. In *IEEE Proc CVPR'93, New-York, June*, pages 194-199, 1993.
- [5] O. Faugeras. *Three-dimensional Computer Vision: a geometric viewpoint*. MIT Press, Boston, 1993.
- [6] O. Faugeras, Q. T. Luong, and S. Maybank. Camera self-calibration: Theory and experiment. In *2nd ECCV*, Genoa, 1992.
- [7] J. Lavest, G. Rives, and M. Dhome. 3D reconstruction by zooming. In *I.A.S.*, 1993.
- [8] Q. Luong, R. Deriche, O. Faugeras, and T. Papadopoulos. On determining the fundamental matrix: analysis of different methods and experimental results. Technical Report RR-1894, INRIA, Sophia, France, 1993.
- [9] T. Luong. *Matrice Fondamentale et Calibration Visuelle sur l'Environnement*. PhD thesis, Université de Paris-Sud, Orsay, 1992. PhD thesis.
- [10] S. Maybank and O. Faugeras. A theory of self-calibration of a moving camera. *The International Journal of Computer Vision*, 8, 1992.
- [11] K. Pahlavan, T. Uhlin, and J.-O. Eklund. Dynamic fixation. In *4th ICCV*, pages 412-419. IEEE Society, 1993.
- [12] L. Robert. *Stéréovision: de la mise en correspondance de courbes à l'analyse photogrammétrique de la scène*. PhD thesis, Université de Paris-Sud, Orsay, 1992. PhD thesis.
- [13] G. Toscani and O. Faugeras. Camera calibration for 3D computer vision. In *Proceedings of the International Workshop on Machine Intelligence, Tokyo*, February 1987.
- [14] R. Y. Tsai. An efficient and accurate calibration technique for 3D machine vision. In *IEEE Proc CVPR'86, Miami Beach, Fl., June*, pages 364-374, 1986.
- [15] R. Y. Tsai. Synopsis of recent progress on camera calibration for 3D machine vision. *Robotics Review*, 1:147-159, 1989.
- [16] T. Viéville. Autocalibration of visual sensor parameters on a robotic head. *Image and Vision Computing*, 1993. in press.
- [17] T. Viéville, Q. Luong, and O. Faugeras. Motion of points and lines in the uncalibrated case. submitted to IJCV, 1993.
- [18] R. Willsona and S. Shafer. What is the center of the image? In *IEEE Proc CVPR'93, New-York, June*, pages 670-671, 1993.



Unité de recherche INRIA Lorraine, Technôpole de Nancy-Brabois, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105,
78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS
Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
(France)
ISSN 0249-6399