

Erreurs de calibration en réalité augmentée : une étude pratique

Javier-Flavio Vigueras, Gilles Simon, Marie-Odile Berger

► **To cite this version:**

Javier-Flavio Vigueras, Gilles Simon, Marie-Odile Berger. Erreurs de calibration en réalité augmentée : une étude pratique. 15ème congrès francophone Reconnaissance des Formes et Intelligence Artificielle - RFIA 2006, AFRIF / AFIA, Jan 2006, Tours, France. inria-00001130

HAL Id: inria-00001130

<https://hal.inria.fr/inria-00001130>

Submitted on 23 Feb 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.

Erreurs de calibration en réalité augmentée : une étude pratique

Calibration errors in augmented reality : a practical study

Javier-Flavio Vigueras

Gilles Simon

Marie-Odile Berger

LORIA - INRIA Lorraine - UHP Nancy 1

LORIA, Campus Scientifique, BP 239, 54506 Vandoeuvre-lès-Nancy, France
{vigueras,gsimon,berger}@loria.fr

Résumé

Dans cet article, nous évaluons la pertinence effective du modèle sténopé dans le cadre de la réalité augmentée (RA). Nous mesurons la précision obtenue avec ce modèle en situation réelle, avec des paramètres intrinsèques éventuellement entachés d'erreur. Nous utilisons pour cette étude des images réelles obtenues à partir de caméra de différentes qualités dont les déplacements sont contrôlés par ordinateur. La scène acquise est constituée d'une mire de calibration et de trois marqueurs placés à différentes distances de la mire. Ce protocole nous permet d'utiliser des critères d'évaluation spécifiques à la RA tels que l'erreur d'alignement et l'impression visuelle, en plus de l'erreur de positionnement plus couramment utilisée.

Mots Clef

Calibration de caméra, calcul de point de vue, réalité augmentée.

Abstract

This paper confronts some theoretical camera models to reality and evaluates the suitability of these models for effective augmented reality (AR). It analyses what level of accuracy can be expected in real situations using a particular camera model and how robust the results are against realistic calibration errors. An experimental protocol is used that consists of taking images of a particular scene from different quality cameras mounted on a 4DOF micro-controlled device. The scene is made of a calibration target and three markers placed at different distances of the target. This protocol enables us to consider assessment criteria specific to AR as alignment error and visual impression, in addition to the classical camera positioning error.

Keywords

Camera calibration, pose computation, augmented reality.

1 Introduction

Le positionnement (ou suivi) de caméra basé vision suscite un vif intérêt dans la communauté réalité augmentée (RA). L'atout majeur de ce type de méthodes est qu'elles reposent sur la minimisation directe de l'erreur d'alignement dans les images à augmenter, ce qui les rend généralement plus précises que d'autres procédés basés sur des capteurs (magnétiques, inertiels, GPS, etc.). Le modèle de caméra le plus souvent utilisé est le modèle sténopé, dont les paramètres intrinsèques sont le plus souvent calculés dans une étape préliminaire dite de calibration, puis considérés comme constants durant le processus de suivi.

Différentes méthodes peuvent être mises en œuvre pour calibrer une caméra (utilisation d'une mire de calibration, détection d'un ou plusieurs marqueurs dans une ou plusieurs images, calculs de points de fuite, etc.), mais toutes ne permettent pas d'obtenir la même précision. Des hypothèses simplificatrices comme celle du point principal au centre de l'image sont aussi souvent utilisées pour réduire le nombre de données requises par le système de calibration et / ou la sensibilité du système face au bruit présent sur ces données. Dans cet article, nous regardons quels peuvent être les effets de telles approximations dans un contexte de réalité augmentée, et proposons des recommandations pratiques pour obtenir une précision suffisante suivant le type d'application visé.

Plusieurs travaux ont été dédiés à l'analyse des effets d'une calibration erronée sur le calcul simultané du mouvement de la caméra et de la structure de la scène [20, 14, 1, 12, 10]. La plupart de ces travaux s'intéressent à l'étude théorique des effets d'une calibration erronée sur l'estimation du mouvement, en supposant que le modèle sténopé est parfaitement respecté par la caméra. Ces travaux ont pu expliquer certaines difficultés liées au positionnement de caméra : par exemple, Agapito et al. ont montré dans [2] qu'une erreur obtenue sur le point principal pouvait être compensée par une erreur sur les angles de rotation, sans que cela ne soit perceptible au niveau de la projection image. Cependant,

ces travaux ne remettent généralement pas en cause le respect du modèle sténopé par les caméras réelles. D'autre part, ils n'intègrent pas certains aspects pratiques et numériques de la calibration : les paramètres sont calculés à partir d'un ensemble d'images et de primitives particulières, et leur validité est souvent restreinte au volume où les mesures sont prises [15, 4]. Or, dans nombre d'applications, il peut être nécessaire d'obtenir un alignement précis hors du volume de mesure. Un exemple est donné en gynécologie [16], dans une application d'apprentissage de l'extraction d'un enfant au forceps : dans cet exemple, les parties invisibles d'un forceps, situées à l'intérieur d'un mannequin, sont affichées par-dessus le mannequin en utilisant un système de positionnement basé sur le suivi automatique de marqueurs rétro-réfléchissants, placés aux extrémités visibles du forceps.

Tout comme Bougnoux dans [3], bien que dans un contexte différent, nous analysons l'influence des erreurs de calibration en situation réelle. Nos résultats ont été obtenus avec deux caméras de qualités différentes, induisant plus ou moins de distorsion dans les images acquises : une caméra possédant une bonne optique et une simple webcam. Ces résultats montrent en particulier que fixer les paramètres intrinsèques n'a de sens que lorsque la distorsion radiale est corrigée, y compris pour des caméras présentant très peu de distorsion. Nos résultats ont été obtenus à l'aide d'une caméra fixée sur un dispositif de déplacement contrôlé par ordinateur, possédant 4 degrés de liberté. Une mire de calibration est utilisée pour calculer les paramètres de la caméra sur toute une séquence, en supposant les paramètres intrinsèques fixes ou variables, puis corrects ou erronés. Enfin, trois marqueurs, dont les positions 3D sont connues, sont placés à des distances plus ou moins éloignées de la mire. Ce protocole nous permet d'analyser l'influence de la calibration en RA, en nous basant sur trois critères : les points de vue de la caméra (les points de vue calculés sont confrontés aux commandes envoyées au dispositif de déplacement), l'erreur d'alignement dans les images (les positions projetées des marqueurs sont comparées aux positions détectées dans la séquence) et l'impression visuelle (des vidéos d'incrustation accompagnent l'article).

Le plan de l'article est le suivant : la section 2 présente plus en détail notre protocole expérimental ; en section 3, nous évaluons la pertinence de fixer les paramètres intrinsèques de la caméra avec le modèle sténopé classique, puis en tenant compte de la distorsion radiale ; la section 4 montre l'influence d'une calibration erronée en RA et nos conclusions sont résumées en section 5.

2 Protocole expérimental

Le dispositif permettant de contrôler les mouvements de la caméra est constitué d'une table en translation à 2 degrés de liberté (appelée table XY), de résolution 0.48 mm, sur laquelle se déplace une unité à deux axes de rotation (*pan-tilt unit* ou ptu), de résolution 0.013 deg (Fig. 1). Une caméra

est fixée à la ptu et une séquence de 753 images d'une scène calibrée est acquise à une résolution de 640×480 pixels. Dans cette séquence, la caméra suit un chemin en créneaux (fig. 1), et les commandes envoyées à la ptu sont choisies de manière à ce que la caméra vise à peu près le centre de la scène dans toutes les images. À trois reprises dans la séquence (au début, au milieu et à la fin), une rotation pure est appliquée à la caméra ayant pour effet un déplacement de la projection de la scène sur les bords de l'image.

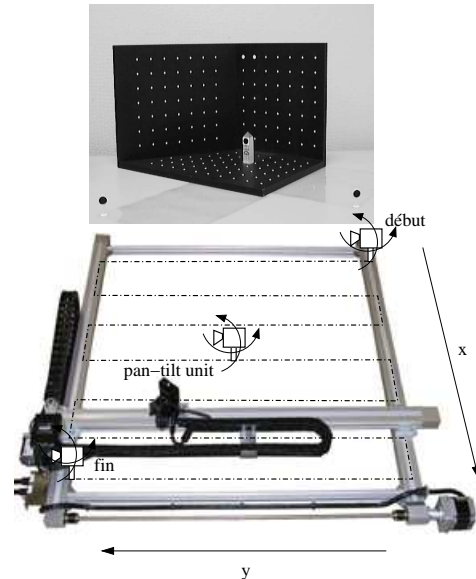


FIG. 1 – Le dispositif de contrôle à 4 degrés de liberté et la scène utilisée.

La scène est composée d'une mire de calibration de taille $39 \times 26 \times 26$ cm et de trois marqueurs dont les coordonnées 3D des centres ont été mesurées dans le repère de la mire (Fig. 1). La mire de calibration a une fonction double : elle permet d'une part d'obtenir les valeurs correctes des paramètres intrinsèques de la caméra (notre *gold standard*), et d'autre part de calculer les points de vue de la caméra à intrinsèques constants, éventuellement entachés d'erreur. Les points de vue calculés peuvent être comparés à la vérité terrain (commandes envoyées à la table XY et à la ptu) en utilisant une transformation hand-eye calculée au préalable. Un marqueur, appelé marqueur *mire* dans la suite, est placé à l'intérieur de la distribution des points de la mire. Deux autres marqueurs (marqueur *gauche* situé à gauche en haut de la figure 1 et marqueur *droite* situé à droite) sont disposés à une distance de 27 et (respectivement) 19 cm de la mire. Ces marqueurs, qui sont des disques noirs apparaissant sur fond blanc, sont suivis automatiquement d'une image à l'autre de la séquence en utilisant une simple technique de corrélation.

Enfin, pour ne pas limiter la portée de nos résultats à un seul type de caméra, des tests ont été réalisés en utilisant trois optiques de qualités différentes : une optique de bonne qualité (une caméra numérique Sony DFW-VL500 avec une grande focale), une optique de qualité moyenne (la

| Caméra | f | $\lambda_1 \cdot 10^7$ | $\lambda_2 \cdot 10^{13}$ | c_x | c_y | Erreur |
|----------|--------|------------------------|---------------------------|-------|-------|--------|
| Logitech | 824.4 | 1.492 | -1.187 | 236.4 | 298.7 | 2.46% |
| Sony 1 | 1113.1 | -1.417 | 2.480 | 303.0 | 290.2 | 1.56% |
| Sony 2 | 2325.2 | -0.675 | -0.153 | 311.2 | 312.5 | 0.95% |

TAB. 1 – Distances focales et paramètres de distorsion des optiques utilisées dans nos expérimentations.

même caméra avec une plus petite focale) et une optique de qualité médiocre (la webcam Logitech QuickCam). Les focales et paramètres de distorsion mesurées sur ces optiques (voir détails plus loin) sont données en table 1. La dernière colonne du tableau indique l’erreur relative due à la distorsion radiale obtenue dans le coin en-haut à droite de l’image.

3 Choix du modèle de caméra

Une hypothèse souvent utilisée en RA est que les paramètres intrinsèques de la caméra restent constants au cours du temps. Fixer les intrinsèques permet d’éviter les phénomènes de compensations qui peuvent se produire entre les paramètres intrinsèques et extrinsèques [2]. De telles compensations induisent généralement des fluctuations statistiques sur les paramètres calculés, ce qui peut générer un effet de tremblement au niveau des éléments virtuels ajoutés. Toutefois, il est important de vérifier que cette hypothèse est pertinente en pratique et qu’elle ne risque pas au contraire d’être source d’imprécision pour l’étape de positionnement. Dans cette section, nous montrons que fixer les paramètres intrinsèques est généralement source d’imprécision, sauf si la distorsion optique de la caméra est prise en compte.

3.1 Cas du modèle sténopé sans distorsion

Le modèle sténopé d’une caméra projette un point 3D (X, Y, Z) en un pixel de coordonnées (u, v) , suivant l’équation :

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & af & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R}|\mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

où f est la distance focale mesurée en taille horizontale d’un pixel, a le rapport d’aspect hauteur/largeur d’un pixel et (u_0, v_0) les coordonnées pixel du point principal. \mathbf{R} et \mathbf{t} sont respectivement la matrice de rotation et le vecteur de translation qui permettent de passer du repère du monde au repère de la caméra. Des travaux antérieurs ont montré que le calcul du rapport d’aspect est robuste aux changements de position et d’orientation de la caméra par rapport à la mire utilisée pour la calibration [18], ainsi qu’aux éventuels zooms et changements de focus de la caméra [5]. Dans la suite de l’article, nous considérons donc que ce paramètre est connu et constant durant le suivi de la caméra, et nous utilisons pour nos tests sa valeur donnée par le *gold standard* (voir la section 3.3).

| | | Gold std | Distorsion non corrigée | Distorsion corrigée |
|----------|-------|----------|-------------------------|---------------------|
| Logitech | u_0 | 341.4 | 334.7 ± 11.4 | 341.3 ± 2.24 |
| | v_0 | 261.3 | 273.0 ± 13.3 | 263.7 ± 1.48 |
| | f | 824.4 | 816.8 ± 16.8 | 825.8 ± 2.67 |
| Sony 1 | u_0 | 306.6 | 305.8 ± 19.0 | 304.3 ± 2.24 |
| | v_0 | 253.0 | 223.2 ± 15.0 | 253.9 ± 2.10 |
| | f | 1113.1 | 1141.3 ± 7.73 | 1114.0 ± 3.64 |
| Sony 2 | u_0 | 303.4 | 304.1 ± 38.1 | 302.5 ± 5.17 |
| | v_0 | 306.2 | 207.0 ± 35.6 | 306.9 ± 5.04 |
| | f | 2325.2 | 2429.8 ± 44.6 | 2326.8 ± 10.9 |

TAB. 2 – Moyennes et écarts-types des paramètres intrinsèques obtenus avec les trois optiques, sans et avec correction des distorsions.

Afin d’évaluer en pratique la pertinence de l’hypothèse des paramètres intrinsèques constants, nous avons réalisé une minimisation à 9 paramètres de l’erreur de reprojection classique :

$$err_f(\Phi, \mathbf{p}) = \sum_j (u_j - \hat{u}_j)^2 + (v_j - \hat{v}_j)^2,$$

où $\Phi = [f, u_0, v_0]^t$ est le vecteur des paramètres intrinsèques, $\mathbf{p} = [\alpha, \beta, \gamma, t_x, t_y, t_z]^t$ le vecteur des paramètres extrinsèques (angles d’Euler et translation), (u_j, v_j) sont les projections des centres des cibles de la mire et (\hat{u}_j, \hat{v}_j) les centres détectés. La méthode de Levenberg-Marquardt est utilisée pour minimiser itérativement cette fonction, avec une estimée initiale donnée par la méthode de Faugeras-Toscani [6].

La première ligne de la figure 2 montre les valeurs de u_0 , v_0 et f obtenues dans chaque image de la séquence, en utilisant la lentille de qualité moyenne. Les intervalles de confiance correspondant aux valeurs des paramètres plus ou moins trois fois leur écart type. Ces écarts-types sont calculés par propagation inverse d’un bruit gaussien d’écart-type 0.5 pixels supposé sur la position détectée des cibles de la mire [9]. On constate que les valeurs obtenues peuvent être très différentes d’une image à une autre : les valeurs de u_0 varient entre 260 et 364 pixels, celles de v_0 entre 177 et 265 pixels et celles de f entre 1122 et 1178 pixels. Pour u_0 et v_0 , il n’existe d’ailleurs aucune droite horizontale passant à l’intérieur de tous les intervalles de confiance.

Des résultats similaires ont été obtenus avec les optiques de qualité basse et élevée : la table 2 montre la valeur moyenne et l’écart-type des paramètres intrinsèques calculés sur la séquence pour les trois optiques. Ces valeurs peuvent être comparées aux valeurs du *gold standard*, obtenues selon la méthode précisée en section 3.3 : on constate que certaines valeurs moyennes sont très éloignées des valeurs attendues, y compris pour l’optique de bonne qualité.

3.2 Prise en compte de la distorsion radiale

La même expérience a été réalisée en ajoutant au modèle de caméra une approximation à l’ordre 4 de la distorsion radiale. Les coordonnées (\hat{u}', \hat{v}') de l’image corrigée sont

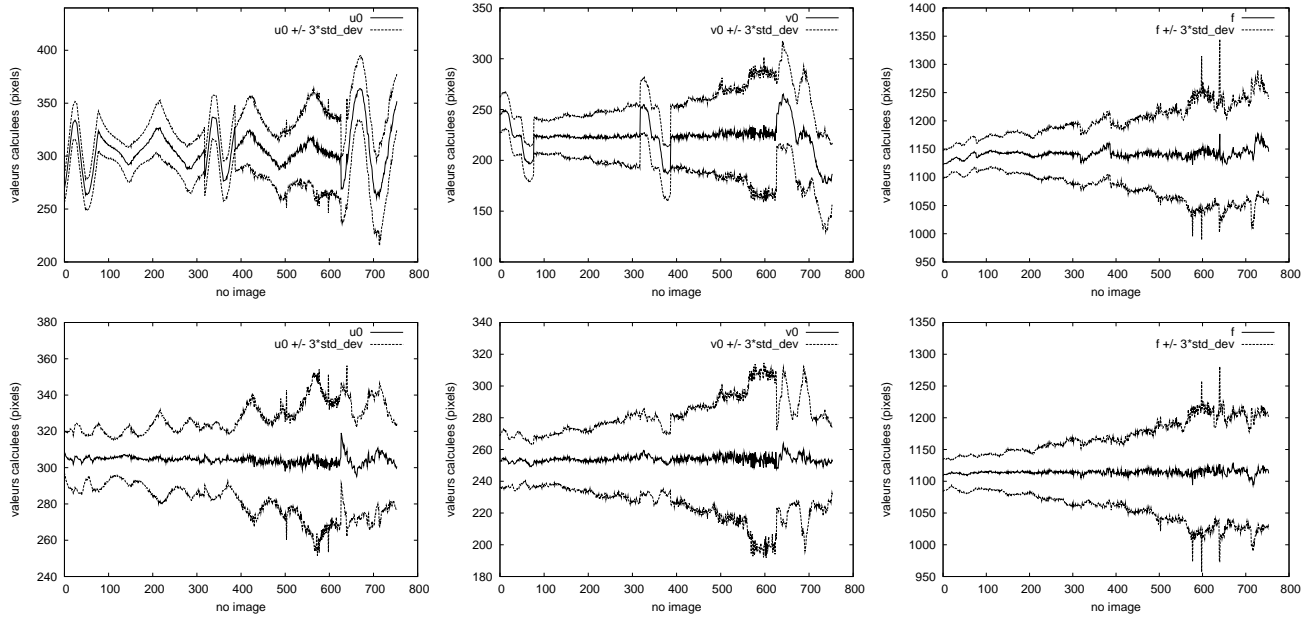


FIG. 2 – Valeurs des paramètres intrinsèques et intervalles de confiances, obtenus dans chaque image de la séquence, sans utiliser (première ligne) puis en utilisant (deuxième ligne) le modèle de distorsion.

obtenues à partir des coordonnées (\hat{u}, \hat{v}) observées, en appliquant la formule suivante :

$$\begin{bmatrix} \hat{u}' \\ \hat{v}' \end{bmatrix} = \begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} + (\lambda_1 r^2 + \lambda_2 r^4) \begin{bmatrix} \hat{u} - c_u \\ \hat{v} - c_v \end{bmatrix},$$

où $r = \sqrt{(\hat{u} - c_u)^2 + (\hat{v} - c_v)^2}$, λ_1 et λ_2 sont les premiers et le deuxième paramètres de correction de distorsion et (c_u, c_v) est le centre de distorsion. En prenant les paramètres de correction de distorsion donnés par le *gold standard* et en corrigeant les coordonnées des centres des cibles à partir de ces valeurs, le minimum de la fonction de coût err_f est atteint pour des valeurs de paramètres intrinsèques plus stables que précédemment (cf. deuxième ligne de la figure 2) : u_0 varie entre 297 et 320 pixels, v_0 entre 246 et 261 pixels et f entre 1094 et 1130 pixels. De plus, pour chaque paramètre nous obtenons une large bande de droites horizontales traversant tous les intervalles de confiance. Enfin, les valeurs moyennes obtenues avec les trois optiques sont à présent bien plus proche des valeurs attendues, et les valeurs des écarts-types sont considérablement réduites (Tab. 2).

Ces résultats sont en accord avec les résultats synthétiques présentés dans [7]. Dans cette étude, 60 points 3D sont tirés aléatoirement dans une sphère de rayon 14 unités. La caméra est placée à différentes distances du centre de la sphère (en moyenne à 70 unités) avec des orientations aléatoires. Les points 3D sont projetés dans l'image et du bruit gaussien ainsi qu'une distorsion radiale sont appliqués aux projections. Lorsque les paramètres intrinsèques sont calculés sans tenir compte de la distorsion, des valeurs incorrectes sont obtenues pour ces paramètres, avec des intervalles de confiance qui n'incluent pas les bonnes valeurs.

3.3 Validation du modèle retenu

Utiliser le modèle sténopé de caméra sans tenir compte de la distorsion radiale conduit donc généralement à des valeurs incorrectes de paramètres intrinsèques (compensations locales dues à la trop grande simplicité du modèle). Corriger la distorsion radiale permet au contraire d'obtenir des valeurs d'intrinsèques moins sensibles à la position de la caméra par rapport à la mire de calibration. Cependant, cela ne prouve pas que le modèle sténopé avec prise en compte de la distorsion radiale est suffisamment riche pour les applications de RA. Pour évaluer la pertinence de ce modèle en RA, nous mesurons à la fois l'erreur de positionnement 3D et l'erreur d'alignement 2D obtenues en utilisant ce modèle pour chaque image de la séquence.

Les paramètres intrinsèques et de distorsion utilisés dans cette validation sont obtenus à partir de 50 images de la mire de calibration. Les positions de caméra ont été choisies de manière à ce que la mire apparaisse en grand dans chaque image, avec ses trois faces bien visibles. Les paramètres retenus sont ceux qui minimisent l'erreur de reprojection des centres des cibles dans toutes les images :

$$\min_{a, \Phi, \Lambda} \sum_{i=1}^{50} err_f(\Phi, \mathbf{p}_i(\Phi)),$$

où a est le rapport d'aspect et $\Lambda = [\lambda_1, \lambda_2, c_u, c_v]^t$ est le vecteur des paramètres de distorsion. $\mathbf{p}_i(\Phi)$ est le point de vue de la caméra qui minimise err_f dans l'image i pour un vecteur donné Φ de paramètres intrinsèques. En utilisant ces paramètres pour calculer les points de vue de caméra sur les 753 images de la séquence, nous obtenons à la fois :
– des points de vue très proches de la vérité terrain sur

| | n | $\Delta T_x(\text{mm})$ | $\Delta T_y(\text{mm})$ | $\Delta \text{pan}(\text{deg})$ | $\Delta \text{tilt}(\text{deg})$ |
|----------|-----|-------------------------|-------------------------|---------------------------------|----------------------------------|
| Logitech | 222 | 0.001 | 0.000 | 0.020 | 0.060 |
| Sony 1 | 753 | -0.165 | -0.004 | -0.020 | 0.014 |
| Sony 2 | 480 | 0.000 | 0.000 | -0.011 | -0.027 |

TAB. 3 – Erreurs moyennes obtenues sur les paramètres du point de vue en utilisant des intrinsèques fixes et en corrigeant la distorsion radiale.

toute la séquence : l’erreur moyenne est indiquée en table 3 (les composantes T_x et T_y de la translation sont exprimées dans le repère de la table XY). Ces erreurs sont de l’ordre de la résolution de la table XY et de la PTU. Notons que le nombre n d’images utilisées pour le calcul des moyennes varie d’une optique à l’autre : suivant l’optique utilisée et la position de la caméra sur la table, la mire de calibration peut en effet apparaître trop petite, trop grande ou hors du champ de vision dans certaines images, ne permettant pas de calculer le point de vue ;

- des erreurs d’alignement très faibles pour les trois marqueurs (Fig. 3) : ces erreurs sont de l’ordre de 1 pixel et sont principalement dues aux erreurs d’incertitude dans la détection des marqueurs.

4 Influence de la calibration

Nous mesurons à présent la précision obtenue en pratique avec une distorsion corrigée mais une calibration imprécise, et un calcul de point de vue reposant sur des appariements 3D-2D de primitives quelconques. Ces résultats s’appliquent à n’importe quel système de positionnement basé modèle, qu’il repose sur la détection automatique de marqueurs artificiels placés dans la scène [11, 19, 16] ou sur l’apprentissage et la reconnaissance d’indices naturels déjà présents dans la scène [8, 13, 17].

La précision est mesurée en terme d’erreur de positionnement et d’alignement 2D. Dans cette section, la lentille de qualité moyenne est utilisée, mais comme seules les images corrigées sont considérées, l’étude est largement indépendante du niveau de distorsion présent dans les images originales.

4.1 Interprétation analytique

Une interprétation analytique de certains de nos résultats peut être obtenue à l’aide de l’équation du flot optique, étendue à la prise en compte d’intrinsèques variables. Lorsqu’une caméra subit un déplacement $(\dot{\alpha}, \dot{\beta}, \dot{\gamma}, \dot{t}_x, \dot{t}_y, \dot{t}_z)$ et une variation d’intrinsèques $(\dot{f}, \dot{u}_0, \dot{v}_0)$ entre deux images, le mouvement horizontal apparent \dot{x} d’un pixel de coordonnées $(x, y) = (u - u_0, v - v_0)$ centrées autour du point principal est donné par :

$$\dot{x} = \frac{x}{f} \dot{f} + \dot{u}_0 + y \dot{\alpha} - f \left(1 + \frac{x^2}{f^2}\right) \dot{\beta} - \frac{xy}{f} \dot{\gamma} - \frac{f}{Z} \dot{t}_x + \frac{x}{Z} \dot{t}_z,$$

où Z est la profondeur du point dans le repère caméra (une expression similaire est obtenue pour \dot{y}). Cette expression

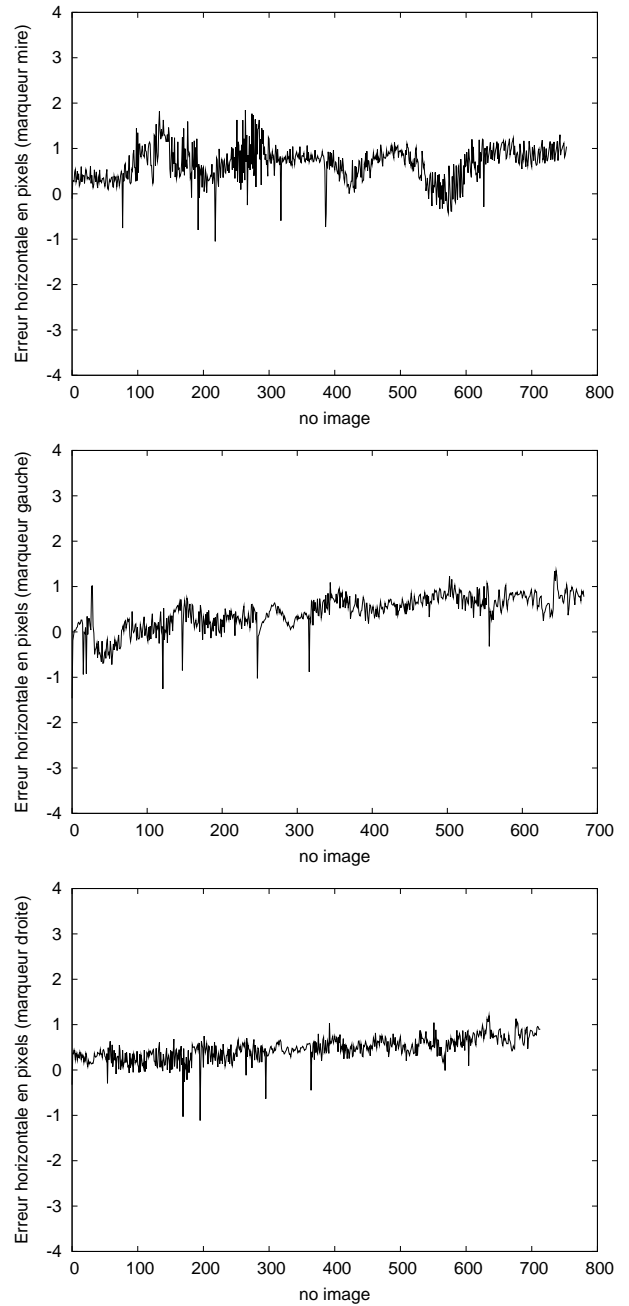


FIG. 3 – Les valeurs du *gold standard* minimisent l’erreur d’alignement à l’intérieur et à l’extérieur du volume de mesure.

peut être interprétée comme l’erreur d’alignement horizontale théoriquement obtenue lorsqu’un ou plusieurs paramètres p_i de la caméra sont erronés ($\dot{p}_i \neq 0$).

4.2 Perturbation des intrinsèques

Dans cette étude, nous voulons perturber les paramètres de calibration de la caméra, mais sans toutefois sortir des marges d’erreurs habituellement observées en situation pratique. Nous nous basons pour cela sur le type d’erreurs que nous obtenons lorsque la caméra est calibrée à partir

d'une seule vue d'une mire de calibration, cette procédure de calibration étant couramment utilisée.

La figure 4 montre la distribution des valeurs de (λ_1, λ_2) obtenues en minimisant l'erreur de reprojection err_f suivant Λ , Φ et \mathbf{p} dans chaque image de la séquence. Les valeurs de (λ_1, λ_2) correspondant au *gold standard* sont étiquetées L0; les valeurs que nous utilisons dans nos expérimentations sont étiquetées de L1 à L4 : elles correspondent aux valeurs obtenues au 5ème (L1), 95ème (L2), 10ème (L3) et 90ème (L4) centile des valeurs ordonnées de λ_1 . Ces valeurs ne sont pas seulement utilisées pour mesurer l'influence des paramètres de distorsion : elles nous permettent aussi d'établir des marges d'erreurs réalistes pour la distance focale et le point principal.

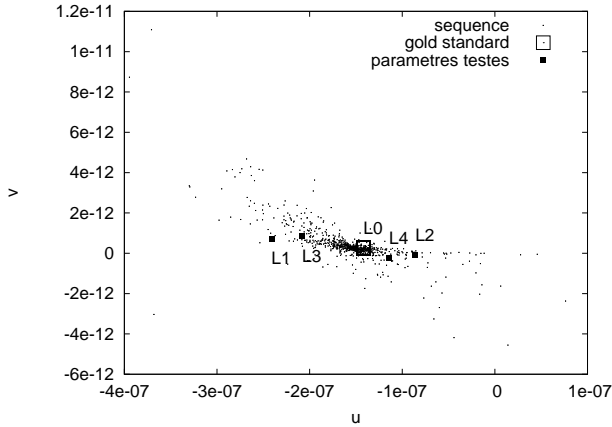


FIG. 4 – Valeurs de (λ_1, λ_2) obtenues dans chaque image de la séquence.

4.3 Influence du point principal

L'influence du point principal est évaluée en minimisant err_f suivant les six paramètres du point de vue, avec des valeurs de f et Λ issues du *gold standard*, et des valeurs de (u_0, v_0) modifiées. L'amplitude et la direction des erreurs introduites sur le point principal sont choisies de la manière suivante. La figure 5 montre l'ensemble des points principaux obtenus en minimisant err_f suivant Φ et \mathbf{p} dans chaque image de la séquence, en utilisant les paramètres de distorsion L0 à L4 (pour plus de visibilité, une image sur 10 seulement est représentée). Cela donne une idée des erreurs pouvant être obtenues avec une calibration mono-image et des paramètres de distorsion obtenus en pratique. Afin de tester des amplitudes et directions d'erreur variées sans toutefois obtenir un trop grand nombre de cas à présenter, nous avons choisi de faire évoluer le point principal sur une spirale représentative du nuage de points et centrée autour du point principal PP0 issu du *gold standard*. Les points principaux ainsi choisis sont étiquetés de PP1 à PP5 en figure 5. Remarquons que PP2 est exactement au centre de l'image, ce qui nous permettra d'évaluer la pertinence de l'hypothèse couramment acceptée du point principal à cette position.

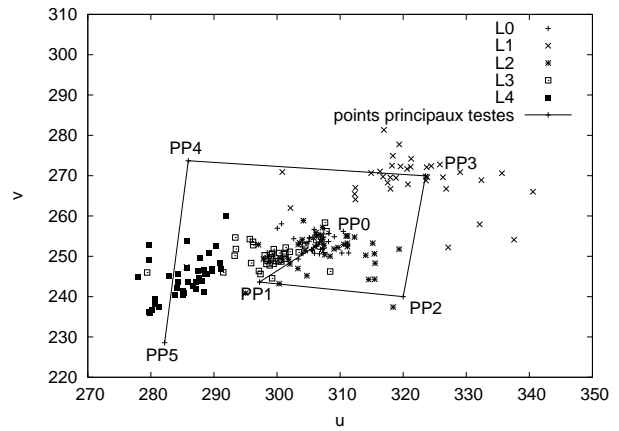


FIG. 5 – Points principaux obtenus en utilisant les paramètres de distorsion L0 à L4.

La figure 6 montre les erreurs d'alignement horizontal que nous obtenons sur les trois marqueurs en utilisant les points principaux PP1 à PP5. Ces erreurs sont négligeables pour le marqueur *mire* placé au centre du volume de mesure. Cela est dû au fait que les erreurs introduites sur le point principal sont compensées par des erreurs sur les angles β et α (pan et tilt) [2], de manière à minimiser l'erreur de reprojection des primitives utilisées pour le positionnement. Ces compensations apparaissent en table 4.

Cependant, les erreurs d'alignement sont plus importantes en dehors du volume de mesure : jusqu'à 4 pixels pour le marqueur *gauche* avec PP4 et jusqu'à 3 pixels pour le marqueur *droite* avec PP5 (Fig. 6). La forme sinueuse des courbes d'erreurs s'explique à l'aide de l'équation du flot optique : les erreurs de compensation entre u_0 and $\dot{\beta}$ introduisent des erreurs d'alignement horizontale $\dot{x} \approx u_0 - f(1 + \frac{x^2}{f^2})\dot{\beta}$. Comme l'erreur sur le point principal u_0 est constant le long de la séquence et la compensation observée sur $\dot{\beta}$ varie peu (voir la figure 7.(a)), la forme des courbes est principalement due aux rapprochements et éloignements successifs des marqueurs à la droite verticale passant par le point principal ($x^2 = (u - u_0)^2$ croît ou décroît, cf. Fig. 8.(a)).

Afin d'évaluer visuellement ces erreurs d'alignement, des cones virtuels dont les bases doivent être superposées aux marqueurs réels ont été projetés dans les images. Les résultats d'incrustation obtenus dans l'image 110 où les trois marqueurs sont bien visibles sont présentés en figure 6. Une vidéo montrant les incrustations obtenues sur toute la séquence avec le point principal PP5 (PP5.avi) peut être téléchargée à l'adresse <http://www.loria.fr/~gsimon/rfia2006.html>. Cette vidéo peut être comparée à la vidéo basée sur les paramètres de calibration issus du *gold standard* (GOLD.avi).

| | ΔT_x (mm) | ΔT_y (mm) | Δpan (deg) | $\Delta tilt$ (deg) |
|-----|-------------------|-------------------|--------------------|---------------------|
| PP1 | -0.22 | -0.29 | 0.45 | -0.49 |
| PP2 | -0.35 | -0.14 | -0.70 | -0.66 |
| PP3 | 0.88 | 0.22 | -0.85 | 0.90 |
| PP4 | 1.64 | 0.81 | 1.01 | 1.09 |
| PP5 | 0.46 | -1.00 | 1.20 | -1.30 |
| F1 | 23.46 | 1.05 | -0.03 | 0.03 |
| F2 | -31.66 | -1.41 | -0.01 | -0.01 |
| F3 | 39.22 | 1.75 | -0.03 | 0.04 |
| F4 | -47.41 | -2.11 | -0.00 | -0.03 |
| F5 | 54.98 | 2.45 | -0.04 | 0.05 |
| L1 | -4.91 | -0.18 | -0.04 | -0.02 |
| L2 | 0.85 | 0.29 | -0.01 | 0.01 |
| L3 | -1.66 | 0.17 | -0.03 | 0.00 |
| L4 | 0.23 | 0.04 | -0.02 | 0.01 |
| L5 | 7.73 | 1.00 | -0.04 | 0.09 |

TAB. 4 – Erreurs moyennes obtenues sur les paramètres du point de vue lorsque des points principaux (PP1-PP5), distances focales (F1-F5) et paramètres de distorsion (L1-L5) erronés sont utilisés.

4.4 Influence de la distance focale

Nous évaluons à présent la sensibilité du positionnement face à des erreurs sur la focale. Les distances focales F1=1128, F2=1093, F3=1138, F4=1083 et F5=1148 sont choisies suivant le même principe que précédemment, sauf que la sélection se fait ici dans une seule dimension. Les résultats sont présentés en figure 9 (F0=1113.1 est la valeur du *gold standard*). À nouveau, l'influence des erreurs sur la focale est imperceptible à l'intérieur du volume de mesure, en raison du phénomène de compensation bien connu qui a lieu entre la focale et la translation le long de l'axe optique (voir table 4 et figure 7.(b)). Cependant, les erreurs d'alignement sont plus importantes en dehors du volume de mesure. La forme des courbes d'erreur s'explique par le fait qu'une erreur sur la focale induit une erreur horizontale $\hat{x} \approx x(\frac{\hat{f}}{f} + \frac{t_z}{Z})$, dont la valeur absolue décroît (avec $|x|$) lorsque le marqueur se rapproche de la droite verticale passant par le point principal, mais aussi lorsque la caméra s'éloigne des marqueurs (accroissement de Z) : la figure 8.(b) montre l'évolution de $1/Z$ pour le marqueur *gauche*. Une vidéo montrant les résultats d'augmentation obtenus avec la focale F5 est disponible sur notre site web (F5.avi).

4.5 Influence de la distorsion radiale

L'influence de la distorsion radiale sur les erreurs d'alignement est présentée en figure 10. L0 à L4 sont les ensembles de paramètres de correction de distorsion dont le choix est présenté en section 4.2 et L5 correspond à des valeurs nulles pour λ_1 et λ_2 (distorsions non corrigées). L'influence des distorsions est négligeable à l'intérieur du volume de mesure, et aussi très faible à l'extérieur, sauf pour L1 qui correspond au couple (λ_1, λ_2) le plus éloigné du *gold standard* (cf. Fig. 4), et pour le cas des distorsions non corrigées. La vidéo correspondant à ce dernier cas est disponible sur le site (L5.avi). Corriger les distorsions

améliore donc généralement la précision de l'alignement de manière significative, même lorsque le calcul des paramètres de correction de distorsion n'est pas réalisé dans des conditions optimales (une seule vue de la mire placée dans une position et orientation arbitraires). Nous n'étudions pas ici de manière théorique les compensations possibles entre les paramètres de distorsion et les paramètres externes de la caméra. Dans nos expérimentations, elles concernent toutefois principalement la translation de la caméra le long de l'axe optique, comme en atteste le tablea 4.

5 Conclusion

Dans cet article, nous avons présenté un ensemble de résultats expérimentaux permettant d'évaluer la pertinence du modèle sténopé en pratique et l'influence des erreurs de calibration dans un contexte de RA. Nos conclusions sont les suivantes :

1. L'utilisation du modèle sténopé de caméra à intrinsèques fixes et distorsion radiale corrigée conduit à des résultats précis de positionnement 3D et d'alignement 2D : avec ce modèle, nous avons obtenu une excellente précision quelque soit l'optique utilisée, la position relative entre la caméra et la scène filmée, et la distance entre l'objet ajouté et le volume de mesure.
2. À l'inverse, l'utilisation du modèle sténopé sans correction de distorsion aboutit à des résultats imprécis, y compris pour des optiques de bonne qualité (moins de 1% d'erreur dans un coin de l'image). Corriger les distorsions s'avère particulièrement nécessaire lorsqu'on cherche à déterminer les paramètres intrinsèques de la caméra, pour éviter des ajustements locaux dus aux insuffisances du modèle. Durant le processus de positionnement de la caméra, cela affecte principalement la précision de l'alignement en dehors du volume de mesure. Des solutions matérielles et logicielles existent pour corriger la distorsion radiale à cadence vidéo, qui devraient être employées lorsqu'une précision élevée est requise en dehors du volume de mesure.
3. Fixer le point principal au centre de l'image ou utiliser des paramètres intrinsèques "imprécis" (obtenus à partir d'une seule vue de mire par exemple) ne nuit généralement pas de façon perceptible à la précision de l'alignement à l'intérieur du volume de mesure. Par contre, ces approximations se traduisent à l'extérieur du volume de mesure par un effet de glissement à basse fréquence des objets ajoutés, dû aux rapprochements et éloignements successifs de l'image de ces objets par rapport au point principal (en ce qui concerne les erreurs sur le point principal et la focale) et aussi de la position 3D de ces objets par rapport à la caméra (en ce qui concerne les erreurs sur la focale). Pour des applications qui requièrent une grande précision en dehors du volume de mesure, la calibration doit donc être réalisée avec soin, en utilisant de

préférence plusieurs images et une technique de type ajustement de faisceau.

Références

- [1] Steffen Abraham and Wolfgang Förstner. Calibration errors in structure from motion. In Paul Levi, Rolf-Jürgen Ahlers, Franz May, and Michael Schanz, editors, *DAGM-Symposium*, Informatik Aktuell, pages 117–124. Springer, 1998.
- [2] L. Agapito, E. Hayman, and I. Reid. Self-calibration of rotating and zooming cameras. *International Journal of Computer Vision*, 45(2), Nov. 2001.
- [3] S. Bougnoux. From Projective to Euclidean Space under any Practical Situation, a Criticism of Self-calibration. In *Proceedings of 6th International Conference on Computer Vision, Bombay (India)*, pages 790–796, January 1998.
- [4] D. Cobzas and M. Jägersand. A comparison of viewing geometries for augmented reality. In *Image Analysis, 13th Scandinavian Conference, SCIA 2003, Halmstad, Sweden, June 29 - July 2, 2003, Proceedings*, pages 501–508, 2003.
- [5] R. Enciso, T. Viéville, and O. Faugeras. Approximation du Changement de Focale et de Mise au Point par une Transformation Affine à Trois Paramètres. Rapport de recherche 2071, INRIA, 1993.
- [6] O. D. Faugeras and G. Toscani. Camera Calibration for 3D Machine Vision. In *Proceedings of International Workshop on Machine Vision and Machine Intelligence*, Tokyo, 1987.
- [7] G. Florou and Roger Mohr. What accuracy for 3d measurements with cameras? In *Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria*, volume I, pages 354–358. IEEE Computer Society Press, 1996.
- [8] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar, and N. Navab. Marker-less tracking for ar : A learning-based approach. In *Proceedings of International Symposium on Mixed and Augmented Reality, Darmstadt, Germany*, pages 295–304, September 2002.
- [9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN : 0521623049, 2000.
- [10] R.I. Hartley and R. Kaucic. Sensitivity of Calibration to Principal Point Position. In *Proceedings of 6th European Conference on Computer Vision, Trinity College Dublin (Ireland)*, June 2002.
- [11] H. Kato and M. Billinghurst. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. In *Proceedings of the 2nd International Workshop on Augmented Reality, San Francisco*, 1999.
- [12] J.M. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *Proceedings of 5th European Conference on Computer Vision, University of Freiburg (Germany)*, pages 158–174, 1998.
- [13] V. Lepetit, L. Vachetti, D. Thalmann, and P. Fua. Fully Automated and Stable Registration for Augmented Reality Applications. In *Proceedings of International Symposium on Mixed and Augmented Reality, Tokyo*, pages 93–101, June 2003.
- [14] J. Oliensis. A Critique of Structure from Motion Algorithms. Technical report, NECI, April 1997.
- [15] J. Park, B. Jiang, and U. Neumann. Vision-based Pose Computation : Robust and Accurate Augmented Reality Tracking. In *Proceedings of the 2nd International Workshop on Augmented Reality, San Francisco*, pages 3–12, 1999.
- [16] T. Sielhorst, J. Traub, and N. Navab. The ar apprenticeship : Replication and omnidirectional viewing of subtle movements. In *Proceedings of International Symposium of Mixed and Augmented Reality, Arlington, VA, USA*, 2004.
- [17] Iryna Skrypnyk and David G. Lowe. Scene modeling, recognition and tracking with invariant image features. In *Proceedings of International Symposium of Mixed and Augmented Reality, Arlington, VA, USA*, 2004.
- [18] Peter Sturm and S. Maybank. On Plane-Based Camera Calibration : A General Algorithm, Singularities, Applications. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina (USA)*, pages 432–437, 1999.
- [19] S. Vogt, A. Khamene, F. Sauer, and H. Niemann. Single camera tracking of marker clusters : Multiparameter cluster optimization and experimental verification. In *Proceedings of International Symposium on Mixed and Augmented Reality, Darmstadt, Germany*, 2002.
- [20] J. Weng, P. Cohen, and M. Herniou. Camera Calibration with Distorsion Models and Accuracy Evaluation. *IEEE Transactions on PAMI*, 14(10) :965–980, 1992.

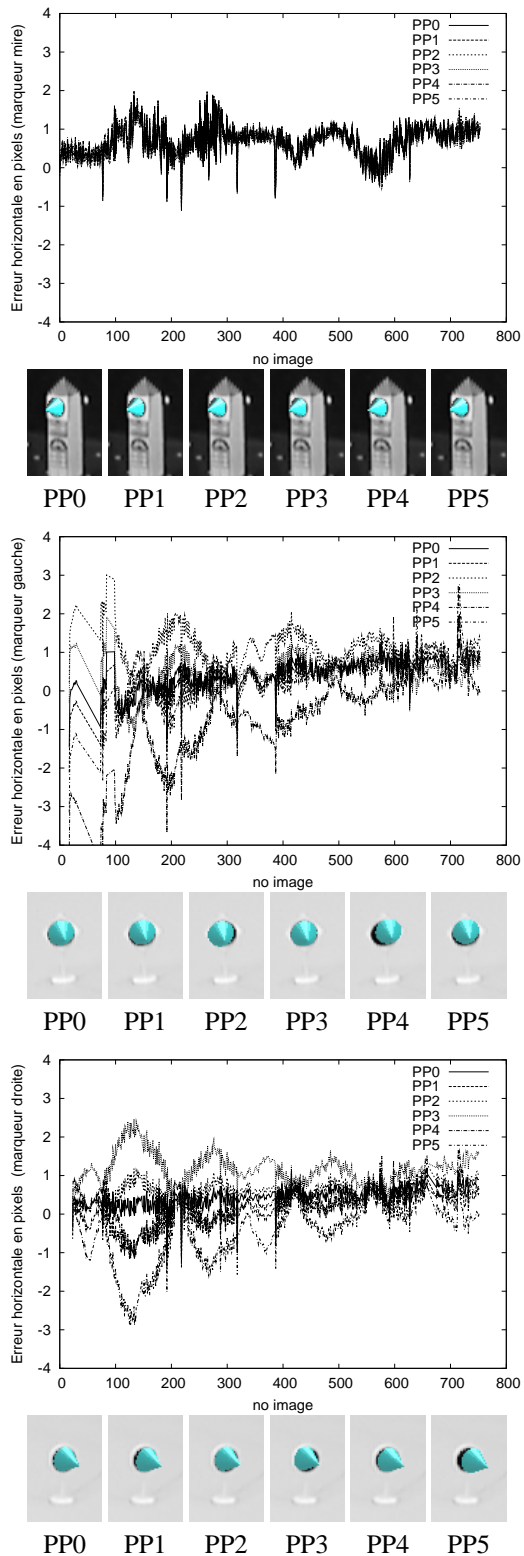


FIG. 6 – Influence du point principal sur l’alignement 2D.

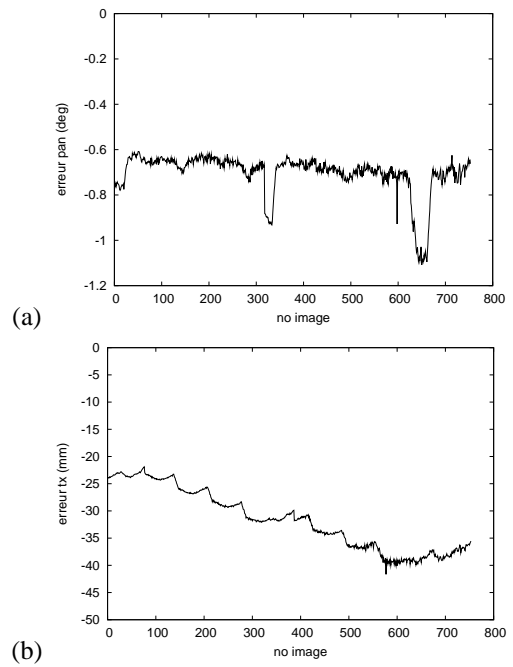


FIG. 7 – (a) Erreur sur β obtenue avec le point principal PP2. (b) Erreur sur T_x obtenue avec la distance focale F2.

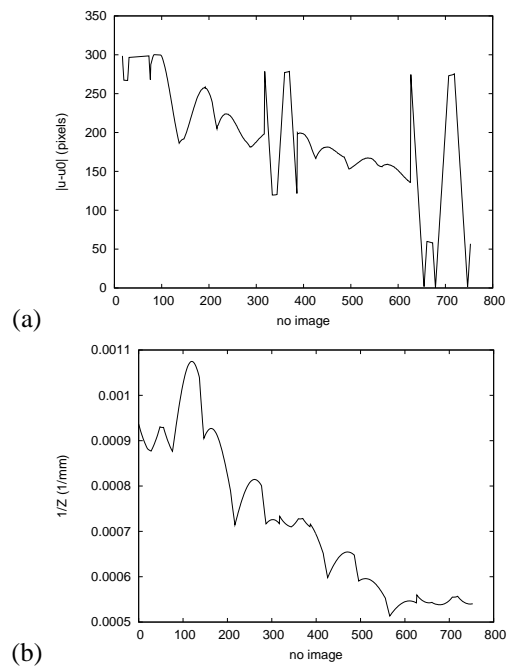


FIG. 8 – Évolution sur la séquence de (a) $|u - u_0|$ et (b) $1/Z$ pour le marqueur gauche.

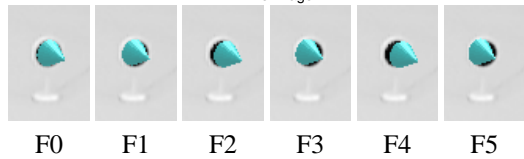
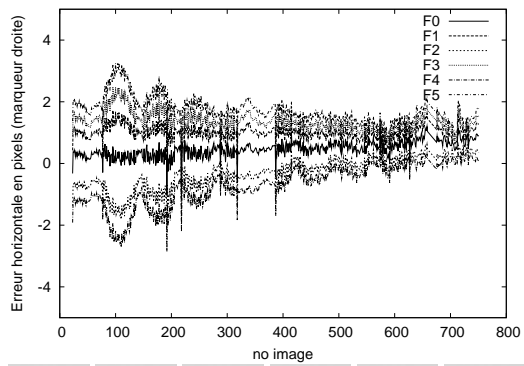
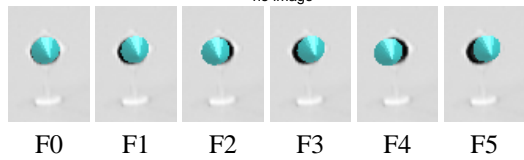
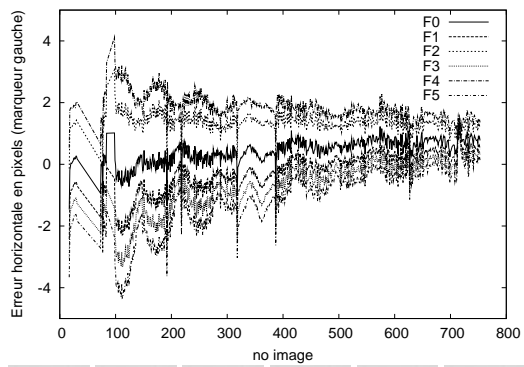
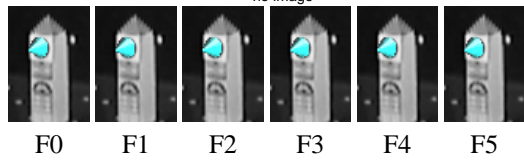
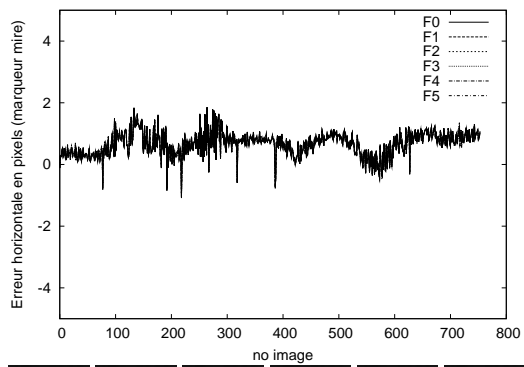


FIG. 9 – Influence de la distance focale sur l’alignement 2D.

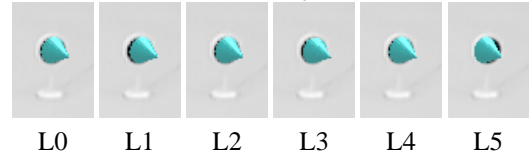
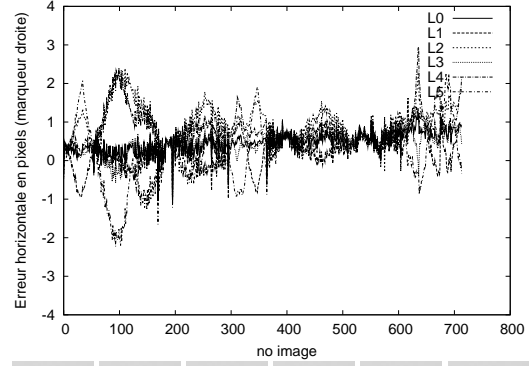
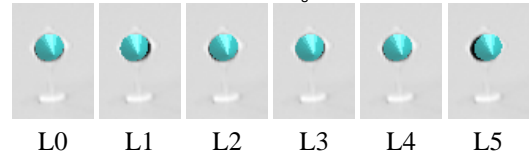
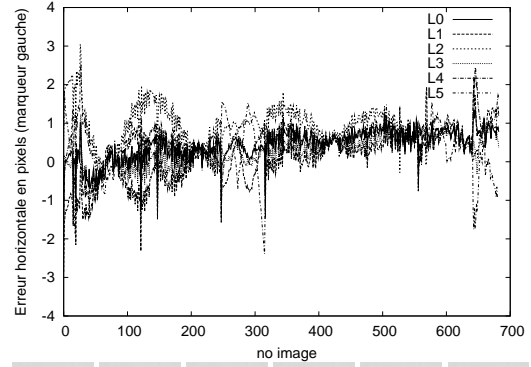
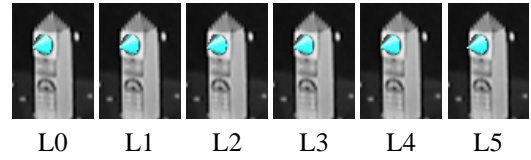
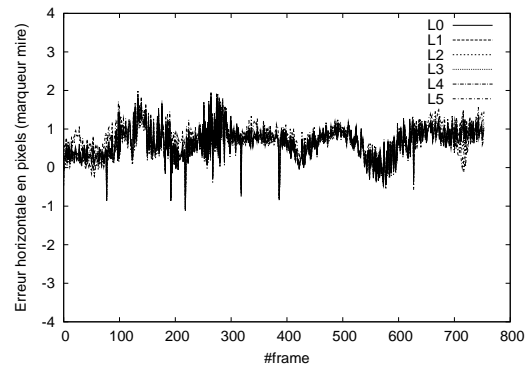


FIG. 10 – Influence des paramètres de distorsion sur l’alignement 2D.