



# Reconstruction 3D automatique avec une caméra catadioptrique

Maxime Lhuillier

► **To cite this version:**

Maxime Lhuillier. Reconstruction 3D automatique avec une caméra catadioptrique. 2006, AFRIF-AFIA, pp.0, 2006. <hal-00091150>

**HAL Id: hal-00091150**

**<https://hal.archives-ouvertes.fr/hal-00091150>**

Submitted on 5 Sep 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.

# Reconstruction 3D automatique avec une caméra catadioptrique

## Automatic 3D Reconstruction with a Catadioptric Camera

Maxime Lhuillier

LASMEA-UMR 6602, UBP/CNRS

24 avenue des Landais, 63177 Aubière

Maxime.Lhuillier@univ-bpclermont.fr, maxime.lhuillier.free.fr

### Résumé

*Des méthodes pour l'estimation robuste et automatique du mouvement d'une caméra catadioptrique et de points de la scène observée sont proposées. Une première estimation de la géométrie complète est obtenue de façon robuste seulement à partir d'une connaissance très approchée des deux angles qui définissent le champ de vue. Cette approche contraste avec les travaux précédents, qui nécessitent un paramétrage du miroir et calculent seulement (jusqu'à présent) la géométrie des paires successives. Ensuite, la connaissance supplémentaire de la forme du miroir est introduite dans l'estimation. Les deux étapes sont rendues possibles grâce à l'introduction d'ajustements de faisceaux pour des caméras à centre unique et centre non-unique. Enfin, le système est présenté dans son ensemble et plusieurs longues séquences d'images sont reconstruites automatiquement pour montrer les qualités de l'approche.*

### Mots Clef

Caméra catadioptrique, caméras à centre unique et centre non-unique, reconstruction 3D, ajustement de faisceaux.

### Abstract

*Methods for the robust and automatic estimation of scene structure and camera motion from image sequences acquired by a catadioptric camera are described. A first estimate of the complete geometry is obtained robustly from a rough knowledge of the two angles which defines the field of view. This approach is in contrast to previous work, which required mirror parameterization and only calculated (until now) the geometry of image pairs. Second, the additional knowledge of the mirror shape is enforced in the estimation. Both steps have become tractable thanks to the introduction of bundle adjustments for central and non-central cameras. Finally, the system is presented as a whole, and many long image sequences are automatically reconstructed to show the qualities of the approach.*

### Keywords

Catadioptric Camera, Central and Non-Central Cameras, Structure from Motion, Bundle Adjustment.

## 1 Introduction

L'estimation robuste et automatique du mouvement d'une caméra perspective (calibrée ou non) tenue à la main et de points de la scène observée a été un sujet de recherche très actif [15, 6] durant la dernière décennie, et plusieurs systèmes ont été réalisés avec succès [3, 28, 26, 18]. D'un autre côté, beaucoup de travaux ont été effectués pendant une période similaire sur l'étude géométrique des caméras/systèmes catadioptriques, bien que de véritables résultats automatiques de reconstructions 3D omni-directionnels n'aient pas encore été publiés. Ces caméras ont un champ de vue large, comme les systèmes dioptriques (respectivement, catadioptriques) composés d'un fish-eye (respectivement, d'un miroir) placé devant une caméra perspective [12]. Cet article se focalise sur la reconstruction 3D avec un système catadioptrique, et est la version française de [19].

**Centre unique ou non-unique** Les caméras dites "à centre unique" sont les plus étudiées, et sont telles que tous les rayons lumineux qui s'en échappent s'intersectent en un point unique de l'espace (à gauche de la figure 1). Elles sont équivalentes à une caméra perspective à une distortion/transformation du plan image près, et pour cette raison, elles héritent de propriétés des caméras perspectives ainsi que de leurs méthodes de reconstructions 3D [13, 27]. Les autres caméras sont à centre non-unique (à droite de la figure 1).

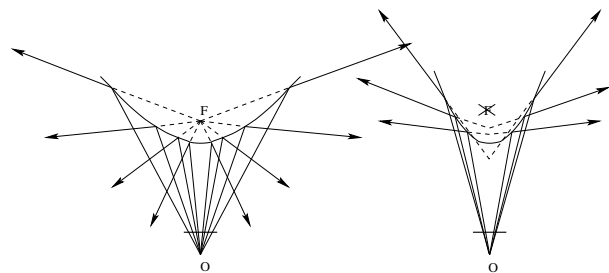


FIG. 1 – Exemples de caméras omni-directionnelles obtenues en alignant une caméra perspective (avec un centre de projection noté  $O$ ) et un miroir. À gauche : une caméra à centre unique est obtenue, car tous les rayons s'en échappant passent en un point unique  $F$ . À droite : une caméra à centre non-unique.

**Travaux précédents** Les plus proches travaux en rapport avec les caméras à centre unique sont [16, 4, 11, 21]. Dans [16], le calibrage et les matrices essentielles successives sont estimés pour une caméra catadioptrique parabolique (composée d’un miroir parabolique devant une caméra orthographique) à partir de points suivis dans une séquence d’images. L’initialisation du calibrage est donnée par une connaissance approchée de l’angle définissant le champ de vue. Un autre travail sur la même caméra [11] introduit la matrice fondamentale parabolique, qui définit une contrainte bilinéaire entre deux points appariés représentés dans un espace adéquat, et montre qu’il est possible d’estimer avec cela le calibrage et la matrice essentielle. Une comparaison [4] est faite entre des mouvements de caméras calibrées tous estimés de façon classique par la matrice essentielle : une caméra perspective (respectivement, omni-directionnelle) donne de meilleurs résultats pour des objets observés proches grâce à sa bonne résolution image (respectivement, distants grâce à son grand champ de vue). L’estimation d’un modèle de caméra à centre unique plus général est proposé par [21], en modélisant directement la fonction de projection et en généralisant l’estimation simultanée de la matrice fondamentale et un coefficient de distortion radiale [9]. Les paramètres de calibrage et les matrices essentielles successives sont estimés automatiquement à l’aides d’algorithmes robustes RANSAC [7] sur 9 et 15 points appariés et une technique de baquet adéquate [22].

Le sujet de ce papier étant la reconstruction 3D avec une caméra omni-directionnelle comportant un miroir, les caméras à centre non-unique obtenues avec une caméra perspective se déplaçant continûment sur une trajectoire connue [2, 29, 12] ne sont pas concernées ici. Ainsi, les travaux précédents se réduisent considérablement à [1, 23]. Étant donné un paramétrage du miroir, la caméra catadioptrique est approximée par une caméra à centre unique [23] : le centre est choisi comme celui qui minimise la somme des carrés des différences des angles entre les rayons réfléchis par la surface du miroir vers les points de la scène et les rayons approximant définis à partir du centre (unique) fictif. Une fois le modèle simplifié, la méthode automatique [21] est appliquée pour obtenir une première estimation de la géométrie de deux images. Finalement, la méthode de Levenberg-Marquardt est appliquée pour raffiner les paramètres des deux caméras à centre non-unique (le centre de symétrie radiale dans l’image, les paramètres de forme du miroir, l’orientation et la position de la caméra perspective, la position et l’orientation du miroir dans le repère de la caméra perspective, et la distance focale). L’erreur dans l’espace 3D est minimisée (au lieu de l’erreur de re-projection habituelle dans les images) car la fonction de projection de la caméra à centre non-unique n’est pas explicite. Ceci n’est pas idéal, car cette erreur 3D n’a pas de justification statistique, et son ordre de grandeur peut varier considérablement entre les points proches et distants des caméras. Dans un contexte différent, un système catadioptrique parabolique [1] est calibrée en supposant que le

miroir (parabolique) et les points 3D sont connus. Ceci est appliqué au calcul de pose temps réel pour un déplacement sur un plan horizontal dans une scène d’intérieur. Un calcul de pose plus précis est obtenu en remplaçant le modèle idéal de caméra orthographique par une caméra perspective.

**Méthodes proposées** La partie 2 décrit les contributions pour une reconstruction 3D automatique avec un modèle de caméra à centre unique : initialisation de la géométrie et ajustements de faisceaux. La géométrie est estimée de façon robuste à partir d’une connaissance grossière des 2 angles qui définissent le champ de vue. Seulement 5 points sont nécessaires en pratique dans l’étape d’estimation robuste par RANSAC [7] des paires d’images [25], sans méthode de baquets particulière. Cette approche simple contraste avec les travaux actuels les plus avancés [21, 23]. La partie 3 décrit les contributions pour un modèle de caméra à centre non-unique : initialisation à partir d’une première estimation avec le modèle à centre unique, et ajustement de faisceaux en prenant compte de la forme du miroir (sans supposer celui-ci parabolique). Dans les deux cas, les ajustements de faisceaux minimisent des erreurs images ou angles. De véritables ajustements de faisceaux n’ont pas encore été introduits et appliqués pour les caméras catadioptriques, bien que de telle méthodes sont reconnues essentielles pour améliorer la précision et la robustesse simultanément dans le cas de caméras perspectives [15]. Des expériences et conclusion sont proposées dans les parties 4 et 5. Enfin, la présentation du système de reconstruction 3D est complétée en annexe (détection des ellipses, mise en correspondance, calculs rapides de dérivées).

## 2 Caméra à centre unique

### 2.1 Modèle et notations

On définit le modèle catadioptrique à centre unique par son orientation  $R$  (une rotation), le centre  $t \in \mathbb{R}^3$  ( $\mathbb{R}$  est l’ensemble des réels), les deux exprimés dans le système de coordonnées de la scène, et une fonction de calibrage  $C : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ . Si  $X$  sont les coordonnées homogènes d’un point 3D tels que la dernière est positive, alors la direction du rayon de  $t$  à  $X$  dans le système de coordonnées caméra est donnée par  $d = R^T (I_3 - t) X$ . La contrainte sur le signe pour  $X$  est nécessaire pour choisir un rayon (demi-droite) parmi les rayons opposés définis par les directions  $d$  et  $-d$ , qui sont projetés en des points différents dans l’image omni-directionnelle. La projection  $p(X)$  de  $X$  dans le plan rétinien est finalement obtenue par  $p(X) = C(d)$ . Si  $X$  est dans le plan à l’infini, il y a deux projections :  $C(d)$  et  $C(-d)$ .

Le modèle à centre unique a aussi un axe de symétrie : l’axe des  $z$  du système de coordonnées de la caméra. L’image est alors comprise entre deux cercles concentriques, et  $C$  est telle qu’il y a une fonction positive  $r$  telle que

$$C(x, y, z) = r(\alpha(x, y, z)) \frac{1}{\sqrt{x^2 + y^2}} \begin{pmatrix} x \\ y \end{pmatrix}.$$

On introduit les notations  $r_{up}$  et  $r_{down}$  pour les rayons des deux cercles, l'angle  $\alpha = \alpha(x, y, z) = \alpha(X)$  entre l'axe des  $z$  et la direction du rayon, et les deux angles  $\alpha_{up}$  et  $\alpha_{down}$  qui définissent le champ de vue. On obtient  $r_{down} \leq r \leq r_{up}$  et  $\alpha_{up} \leq \alpha \leq \alpha_{down}$  comme cela est montré dans la figure 2. Les angles  $\alpha_{up}, \alpha_{down}$  sont approximativement connus et donnés par le constructeur du miroir, bien qu'ils dépendent de la distance focale inconnue de la caméra perspective et de la position relative entre la caméra et le miroir. La fonction  $r$  est précisée plus tard. Ce modèle nécessite la connaissance de la fonction de l'image vers le plan rétinien, qui associe deux cercles concentriques aux frontières de l'image omni-directionnelle (ellipses). Cette fonction est pré-calculée pour chaque image à partir d'une détection d'ellipse automatique (voir l'annexe), et est appliquée implicitement dans la partie 2 en faisant comme si la frontière des images omni-directionnelles sont les deux cercles concentriques. Un modèle de caméra à centre unique similaire a été proposé dans [24].

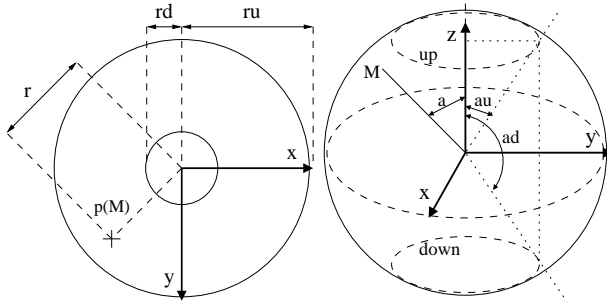


FIG. 2 – À gauche : 2 cercles concentriques de rayon  $r_{up}, r_{down}$  (abrégés  $ru, rd$ ) dans le plan rétinien. À droite : les angles  $\alpha_{up}, \alpha_{down}, \alpha(X)$  (abrégés  $au, ad, a$ ) par rapport à l'axe des  $z$  définissent le champ de vue pour un point  $X$  avec la relation  $\alpha_{up} \leq \alpha(X) \leq \alpha_{down}$ . Le cercle "up" (respectivement, "down") des rayons à droite est projeté par le modèle à centre unique sur le grand (respectivement, petit) cercle sur la gauche.

## 2.2 Initialisation de la géométrie

Une fois que les cercles frontières et la mise en correspondance entre les paires d'images consécutives ont été obtenues (voir l'annexe), la géométrie complète de la séquence peut être estimée. On calcule la géométrie par paire étant donné un calibrage approché, pour profiter des méthodes de reconstructions 3D classiques.

La fonction de calibrage  $r$  est définie par la fonction affine de l'angle  $\alpha$  telle que  $r(\alpha_{up}) = r_{up}$  et  $r(\alpha_{down}) = r_{down}$ . Les rayons de cercle  $r_{down}, r_{up}$  sont déjà estimés, et les angles approchés du champ de vue  $\alpha_{down}, \alpha_{up}$  sont donnés par le constructeur du miroir. Un tel choix suppose une résolution radiale image similaire pour toutes les directions de rayon, ce qui n'est pas inhabituel en pratique.

Une fois le calibrage approché  $C$  connu, une direction de rayon pour chaque point dans l'image est donnée par la fonction réciproque (explicite)  $C^{-1}$ . Des méthodes connues

[5, 25] sont ensuite appliquées pour obtenir une première estimation des matrices essentielles, et la géométrie de tous les triplets d'images consécutives par le calcul de pose de la 3-ième caméra avec des points appariés sur 3 vues et reconstruits par les 2 premières. Le raffinement de la géométrie par ajustement de faisceau est décrit dans la partie 2.3. La géométrie de chaque séquence d'images est estimée d'une façon hiérarchique [15] : une fois que la géométrie de deux sous séquences d'images  $1 \dots \frac{n}{2}, \frac{n}{2} + 1$  et  $\frac{n}{2}, \frac{n}{2} + 1, \dots n$  sont estimées, la seconde est exprimée dans le système de coordonnées de la première grâce aux 2 caméras en commun  $\frac{n}{2}, \frac{n}{2} + 1$ , et la séquence résultat  $1 \dots n$  est raffinée avec un ajustement de faisceaux.

## 2.3 Ajustements de faisceaux

Un ajustement de faisceaux améliore l'estimation des centres  $t_i$ , orientations  $R_i$  des caméras  $i$  et les positions  $X_j$  des points  $j$  par la minimisation d'un score : la somme des erreurs de reprojection avec les points appariés  $m_{ij}$ .

**Erreur image** A première vue, le score dans l'espace image omni-directionnel pourrait être

$$\sum_{i,j} \|C(R_i^T (I_3 - t_i) X_j) - m_{ij}\|^2$$

avec  $\|\cdot\|$  la norme euclidienne. Cependant, un problème pratique apparaît pour les points  $X_j$  voisins du plan à l'infini avec ce score : certains d'entre eux pourraient passer de l'autre côté de ce plan pendant certaines itérations de Levenberg-Marquardt (lorsque le signe de la 4ième coordonnée de  $X_j$  change), et ceci dégrade sévèrement le score. En fait, la projection dans l'image omni-directionnelle d'un point se trouvant exactement sur le plan à l'infini n'est pas un point, mais deux points  $C(d)$  et  $C(-d)$  (alignés et séparés par le centre des cercles concentriques), et seulement un de ces deux points est très proche du point détectés correspondant. Si aucune précaution n'est prise lorsqu'un point sur un rayon fixé traverse le plan à l'infini, sa projection change dramatiquement de  $C(d)$  à  $C(-d)$ , et le score global n'est plus significatif. Pour cette raison, on choisit un score continu sur le plan à l'infini :

$$\sum_{i,j} \min_{\epsilon=\pm 1} \|C(\epsilon R_i^T (I_3 - t_i) X_j) - m_{ij}\|^2.$$

**Erreur angulaire** Un autre score possible est la somme des carrés des angles entre la droite définie par les points  $X_j$  et  $t_i$ , et la droite définie par le point  $t_i$  et la direction  $C^{-1}(m_{ij})$ . Ces angles sont ceux entre  $C^{-1}(m_{ij})$  et  $R_i^T (I_3 - t_i) X_j$  dans le système de coordonnées de la  $i$ -ième caméra. Soit  $R_{ij}$  une rotation telle que  $R_{ij} C^{-1}(m_{ij}) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ , et  $\pi(x, y, z) = (x/z \ y/z)$ . Ces angles sont ceux entre  $R_{ij} R_i^T (I_3 - t_i) X_j$  et  $\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ . Le score est donc

$$\sum_{i,j} \|\pi(R_{ij} R_i^T (I_3 - t_i) X_j)\|^2$$

en approchant les angles par leurs tangentes. Cette approximation est acceptable car les angles sont petits pour les points proches d'une solution. On note aussi que

- ce score angulaire est indépendant du choix de  $R_{ij}$
- le problème du plan à l'infini du score image disparaît, car  $\pi(-x) = \pi(x)$
- les deux score restent valides sans la contrainte de signe sur  $X_j$

### 3 Caméra à centre non-unique

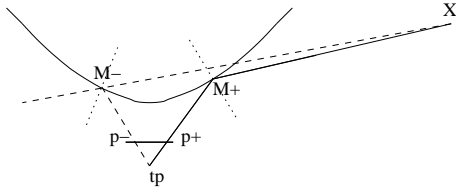


FIG. 3 – Le rayon qui passe par un point  $X$  et le centre de la caméra perspective  $\mathcal{P}$  est réfléchi par le miroir en un point  $M^+(t^p, X)$  et est projeté par la caméra perspective en  $p^+(X)$ . On définit aussi le point  $M^-(t^p, X)$  sur le miroir, projeté en  $p^-(X)$ .  $M^-(t^p, X)$  est comme  $M^+(t^p, X)$ , à ceci près que le rayon réfléchi de  $M^-(t^p, X)$  vers  $X$  pointe vers l'intérieur du miroir, pas l'extérieur. Les points  $M^+(t^p, X), p^+(X), M^-(t^p, X), p^-(X)$  sont abrégés  $M^+, p^+, M^-, p^-$  dans cette figure. On note que  $X, M^-, M^+$  ne sont pas nécessairement alignés.

#### 3.1 Modèle et notations

On définit le modèle de caméra à centre non-unique par l'orientation  $R$  (une rotation) et la position  $t \in \mathbb{R}^3$  du miroir dans le système de coordonnées de la scène, les paramètres intrinsèques  $K^p$ , orientation  $R^p$  et position  $t^p$  de la caméra perspective exprimées dans le système de coordonnées du miroir, et la surface du miroir.

La surface du miroir contient l'origine et est symétrique autour de l'axe des  $z$  des coordonnées du miroir. Soit  $M^+(t^p, X)$  la fonction qui donne le point de réflexion sur le miroir du rayon qui passe par les points  $t^p, M^+(t^p, X)$  et  $X$  dans cet ordre. On a  $t^p, X, M^+(t^p, X) \in \mathbb{R}^3$ , tous exprimés dans le système de coordonnées du miroir. Cette fonction n'a pas de forme explicite, et son calcul à partir de la surface du miroir est présenté en annexe.

Soit  $\tilde{X}$  les coordonnées (non-homogènes) d'un point fini  $X$ . Puisque  $R^\top(\tilde{X} - t)$  et  $K^p R^p \top(I_3 - t^p)$  sont respectivement les coordonnées miroir de  $X$  et la matrice de  $C^p$ , on en déduit que les coordonnées homogènes de la projection par le modèle à centre non-unique de  $X$  est

$$p^+(X) \equiv K^p R^p \top(M^+(t^p, R^\top(\tilde{X} - t)) - t^p).$$

La figure 3 montre une coupe du miroir et de la caméra perspective, avec les points  $t^p, X, p^+(X), M^+(t^p, X)$ .

Bien que ce modèle ne définisse pas directement la projection d'un point  $X_\infty$  dans le plan à l'infini, on peut la définir par passage à la limite. Soit  $X$  un point fini qui converge vers  $X_\infty$  tel que ses coordonnées  $x, y, z$  sont fixées et  $t$  converge vers 0. Deux limites de projections  $p^+(X)$  sont obtenues par la relation  $\tilde{X} = (x/t \ y/t \ z/t)$ , un pour chaque signe possible de  $t$ .

On note aussi que le paramétrage  $(R^p, t^p, R)$  n'est pas minimal à cause de la symétrie du miroir autour de l'axe des  $z$  : l'expression de  $p^+(X)$  est inchangée en remplaçant  $(R^p, t^p, R)$  par  $(R_z R^p, R_z t^p, R R_z^\top)$  et en appliquant la relation de symétrie  $R_z M^+(A, B) = M^+(R_z A, R_z B)$  pour toute rotation  $R_z$  autour de l'axe des  $z$ .

#### 3.2 Initialisation de la géométrie

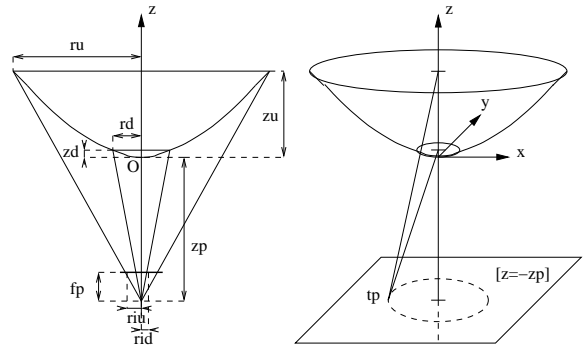


FIG. 4 – À gauche : notation utiles (et abrégées) pour l'estimation de  $z^p$  et  $f^p$ . À droite :  $t^p$  est située sur un cercle dans le plan  $z = -z^p$  étant donné  $\alpha$  et l'angle entre la direction de  $t^p$  vers les centres des grands et petit cercles.

**Paramètres perspectifs** Les paramètres  $K^p, R^p, t^p$  sont initialisés à partir des deux ellipses frontières détectées dans les images et de la connaissance du miroir. Ces ellipses sont les images des deux cercles frontières connus du miroir dans les plans  $z = z_{up}$  et  $z = z_{down}$  exprimés dans le repère miroir, avec  $z_{down} < z_{up}$ .

On fait les hypothèses suivantes : la caméra perspective pointe vers le miroir avec  $t^p$  dans le voisinage immédiat de l'axe de symétrie du miroir tel que  $R^p \approx I_3$ ,

$$t^p = (\epsilon_x^p \ \epsilon_y^p \ -z^p), \ 0 < z^p, |\epsilon_x^p| \ll z^p, |\epsilon_y^p| \ll z^p,$$

et les ellipses sont approchées par des cercles de rayon  $r_{up}^i, r_{down}^i$  et de centre  $c_{up}^i, c_{down}^i$ . De plus, on suppose que  $K^p$  est une matrice diagonale avec la distance focale  $f^p$ , le centre optique au milieu de l'image et des pixels carrés.

D'abord,  $f^p$  et  $z^p$  peuvent être estimés en supposant  $R^p = I_3$  et  $\epsilon_x^p = \epsilon_y^p = 0$  par les relations de Thalès  $r_{up}^i/f^p = r_{up}/(z^p + z_{up})$  et  $r_{down}^i/f^p = r_{down}/(z^p + z_{down})$  (à gauche dans la figure 4). L'estimation de  $z^p$  peut aussi être obtenue à partir de la loi de réflexion et de la connaissance de  $r_{up}, z_{up}, \alpha_{up}$ , ou directement en mesurant la distance entre le miroir et la caméra perspective. Une autre méthode [31] est encore possible pour l'estimation de  $f^p$ .

Ensuite,  $\epsilon_x^p$  et  $\epsilon_y^p$  sont estimés. Puisque les angles entre les directions pointant vers les centres des cercles délimitant le miroir est celui entre les vecteurs  $(K^p)^{-1}c_{up}^i$  et  $(K^p)^{-1}c_{down}^i$ , on connaît le rayon du cercle dans le plan  $z = -z^p$  où  $t^p$  est situé (à droite de la figure 4). Tout  $t^p$  sur ce cercle est possible grâce au paramétrage redondant  $(R^p, t^p, R)$ . L'hypothèse  $R^p \approx I_3$  fixe  $t^p$  de la façon suivante : des relations de projections

$$\begin{aligned} \lambda_u (-\epsilon_x^p \quad -\epsilon_y^p \quad z^p + z_{up})^\top &= R^p (c_{up}^{i\top} \quad f^p)^\top \\ \lambda_d (-\epsilon_x^p \quad -\epsilon_y^p \quad z^p + z_{down})^\top &= R^p (c_{down}^{i\top} \quad f^p)^\top, \end{aligned}$$

on a  $\lambda_u \approx \frac{f^p}{z_{up} + z^p} < \frac{f^p}{z_{down} + z^p} \approx \lambda_d$  et par soustraction  $(c_{up}^{i\top} - c_{down}^{i\top} \quad 0) \approx (\lambda_d - \lambda_u) (\epsilon_x^p \quad \epsilon_y^p \quad *)$ . On peut donc trouver  $\lambda > 0$  tel que  $(\epsilon_x^p \quad \epsilon_y^p)^\top = \lambda (c_{up}^i - c_{down}^i)$  à partir du rayon du cercle dans le plan  $z = -z^p$ . Finalement,  $R^p$  est estimé à partir des deux relations de projection.

**Structure et mouvement** Une première possibilité est d'identifier les notations  $t_i$ ,  $R_i$ ,  $X_j$  des deux modèles à centre unique et non-unique, dès qu'une approximation de la géométrie complète est donnée par la méthode de la partie 2. Le facteur d'échelle, en fait, n'est plus un paramètre libre comme dans le cas centre unique, car il est fixé par la taille du miroir. On choisit le facteur d'échelle initial pour  $t_i$  et  $X_j$  tel que la distance moyenne entre les positions consécutives du miroir sont physiquement plausibles comparées aux dimensions du miroir.

### 3.3 Ajustements de faisceaux

Un ajustement de faisceaux améliore l'estimation des positions  $t_i$  et orientations  $R_i$  du miroir, les coordonnées homogènes  $X_j = (X_j^x \quad X_j^y \quad X_j^z \quad X_j^t)$  des points  $j$ , par la minimisation d'un score : la somme des erreurs de re-projections avec les points appariés connus  $m_{ij}$ . Les explications de ce paragraphe sont similaires à celles de la partie 2.3.

**Erreur image** A première vue, le score dans l'espace image omni-directionnel pourrait être

$$\sum_{i,j} \|\pi(K_i^p R_i^{p\top} (M^+(t_i^p, R_i^\top (\tilde{X}_j - t_i)) - t_i^p)) - m_{ij}\|^2$$

avec  $\tilde{X}_j = (X_j^x/X_j^t \quad X_j^y/X_j^t \quad X_j^z/X_j^t)$ , les  $i$ -ièmes paramètres perspectifs  $K_i^p$ ,  $R_i^p$ ,  $t_i^p$ , et  $\pi(x, y, z) = (x/z \quad y/z)$ . Comme dans le cas à centre unique, la projection dans une image d'un point  $X_j$  qui traverse le plan à l'infini durant une itération de Levenberg-Marquardt change dramatiquement, et le score résultant devient non significatif. La solution à ce problème est le même pour le cas centre non-unique : on redéfinit un score qui est continu pour  $X_j$  dans le voisinage du plan à l'infini.

On introduit dans la figure 3 une sorte de point antipodal pour  $M^+$  défini par  $M^-$ , et on minimise le score

$$\sum_{i,j} \min_{\epsilon \in \{+, -\}} \|\pi(K_i^p R_i^{p\top} (M^\epsilon(t_i^p, R_i^\top (\tilde{X}_j - t_i)) - t_i^p)) - m_{ij}\|^2$$

Contrairement au choix simple  $M^-(t^p, X) = M^+(t^p, -X)$ , le choix fait ici est indépendant de l'origine du système de coordonnées miroir grâce à la relation  $M^-(t^p, X) = M^+(t^p, 2M^-(t^p, X) - X)$ . Les deux choix ont la même limite au voisinage du plan à l'infini et ont des projections images très différentes de celle de  $M^+$ , ce qui est nécessaire.

**Erreur angulaire** Soit  $t_{ij}$  et  $n_{ij}$  le point sur le  $i$ -ième miroir et la direction du rayon réfléchi, défini par le point image  $m_{ij}$  et la  $i$ -ième caméra perspective. Un score possible est la somme des carrés des angles entre la droite définie par les points  $t_{ij}$  et  $X_j$ , et la droite définie par le point  $t_{ij}$  et la direction  $n_{ij}$ . Ces angles sont ceux entre  $n_{ij}$  et  $R_i^\top (I_3 \quad -t_i) X_j - t_{ij} X_j^t$  dans le repère de  $i$ -ième miroir. En introduisant une rotation  $R_{ij}$  telle que  $R_{ij} n_{ij} = (0 \quad 0 \quad 1)$  comme dans le cas centre unique, on obtient le score

$$\sum_{i,j} \|\pi(R_{ij}(R_i^\top (I_3 \quad -t_i) X_j - t_{ij} X_j^t))\|^2.$$

Les 3 remarques de l'erreur angulaire dans la partie 2.3 sont encore correctes.

## 4 Expérimentations



FIG. 5 – De gauche à droite : le miroir Kaidan “360 One VR” et l'appareil photo Nikon “Coolpix 8700” (monté sur un monopode), le champ de vue (360° horizontalement et environ 50° au dessus et en dessous du plan horizontal, lorsque la caméra pointe vers le ciel), la silhouette du miroir et sa caustique pour  $z^p = 50\text{cm}$ .

### 4.1 Contexte

**Système catadioptrique à bas coût** Le manipulateur se déplace le long d'une trajectoire sur le sol avec le système catadioptrique monté sur un monopode, en alternant un pas en avant et la prise d'une photo. Un système catadioptrique à bas coût est utilisé, dédié à l'origine à la génération de panoramiques et rendu basé-image à partir d'une seule image de la scène (voir la figure 5). Les axes de symétrie de la caméra et du miroir ne sont pas exactement les mêmes (l'alignement des axes est réglé manuellement et vérifié visuellement). D'après le constructeur du miroir (et les tailles

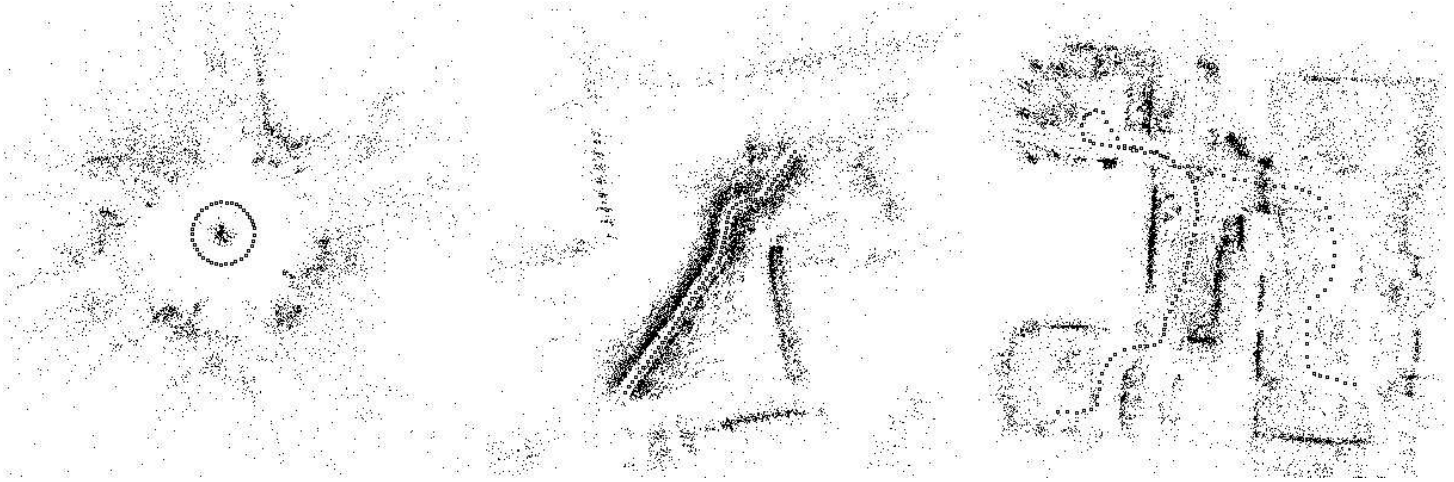


FIG. 6 – **En haut** : panoramiques obtenus à partir d'une image de chaque séquence. **En bas** : vues de dessus des reconstructions 3D avec la position des caméras (petits carrés) et des points 3D (points noirs). **À gauche** : Fontaine (38 vues, 5857 points). **Au milieu** : Chemin (54 vues, 10178 points). **À droite** : Maison (112 vues, 15504 points). Certains de ces points sont difficiles à reconstruire lorsqu'ils sont alignés (ou presque) avec les caméras à partir desquelles ils sont reconstruits. Ces résultats sont similaires pour les caméras à centre unique (avec une fonction  $r(\alpha)$  affine ou cubique) ou à centre non-unique.

relatives du miroir et de sa caustique), le système n'est pas à centre unique.

**Mouvements de caméras "habituels"** Le mouvement de caméra décrit dans le paragraphe précédent est proche d'un mouvement de translation, pour lequel la caméra perspective pointe vers une direction fixée (le ciel). Un tel mouvement est un mouvement critique pour la caméra à centre unique, de la même façon que pour une caméra perspective (les preuves sont très similaires). Comme plusieurs calibrages sont possibles pour un mouvement critique donné de la caméra catadioptrique, l'estimation complète des points de la scène et des caméras est impossible. D'un autre côté, ce serait plus facile de choisir un calibrage approché à fournir aux algorithmes de reconstruction 3D. La situation n'est pas claire pour les mouvements de caméra proches des mouvements critiques, et ceci doit être examiné expérimentalement.

## 4.2 Modèle à centre unique

Des panoramiques obtenus à partir d'images omni-directionnelles et des vues de dessus des reconstructions 3D sont montrées dans la figure 6.

**Fontaine** La séquence Fontaine est composée de 38 images d'une ville en arrière plan et d'une fontaine au centre d'un rond point à l'avant plan. Environ 50% des matrices essentielles estimées sont clairement erronées pour cette séquence, puisque les epipoles sont à peu près orthogonales au mouvement de la caméra comme cela est montré dans la figure 7. Cependant, le raffinement systématique de la

géométrie par ajustement de faisceaux après chaque étape de fusion de sous-séquences corrige toutes ces erreurs, et, comme prévu, le mouvement de caméra reconstruit est un mouvement circulaire régulier autour de la fontaine.

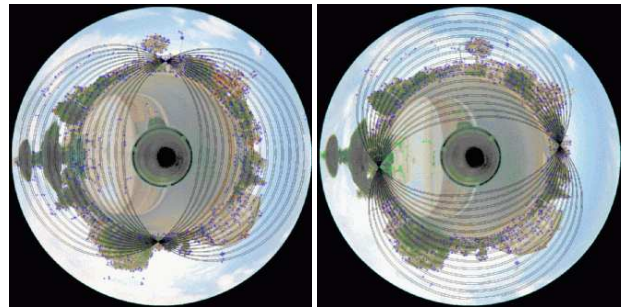


FIG. 7 – Courbes epipolaires pour la géométrie de deux paires d'images différentes de la séquence Fontaine. Les deux epipoles à gauche sont cohérents avec le véritable mouvement de la caméra autour de la fontaine car ils indiquent la direction du mouvement de la caméra. Ce n'est pas le cas pour ceux de droite, et de telles erreurs sont corrigées par les ajustements de faisceaux.

Des versions fermées et non fermées de cette séquence ont été reconstruites. La version non fermée est obtenue en dupliquant en fin de séquence la première image. L'écart entre les deux extrémités de la séquence non fermée est faible : la distance entre des centres de caméras est 0.01 fois le diamètre de la trajectoire, et l'angle de l'orientation relative  $R_{38}R_0^{-1}$  est  $0.63^0$ .

**Chemin** La séquence Chemin est composée de 54 images prise le long d’un petit chemin sur un sol plat, avec en arrière plan des immeubles, parkings et sapins. Toutes les matrices essentielles semblent être correctes d’après la position des epipoles.

**Maison** La séquence Maison est composée de 112 images prise dans une maison douillette, partant dans la salle de séjour, traversant l’entrée, faisant une boucle dans la cuisine, re-traversant l’entrée et entrant dans une chambre.

**Robustesse et stabilité de l’estimation de  $\alpha_{up}, \alpha_{down}$**

Toutes les reconstructions à centre unique précédentes sont obtenues avec une fonction de calibrage  $r(\alpha)$  affine définie avec les angles de champ de vue  $\alpha_{up} = 40^\circ, \alpha_{down} = 140^\circ$  donnés par le constructeur du miroir.

Deux propriétés de la méthode de reconstruction 3D à centre unique sont explorées dans ce paragraphe : la robustesse de l’initialisation avec des valeurs très grossières de  $\alpha_{up}, \alpha_{down}$ , et la possibilité de raffiner ces deux angles avec un polynôme de degrés 3 pour  $r(\alpha)$  (les 4 coefficients sont de nouvelles inconnues pour l’ajustement de faisceaux image). Puisque les mouvements du système catadioptrique sont presque des mouvements critiques, les résultats pour ces deux propriétés ne sont pas certains.

La figure 8 montre une bonne robustesse de la reconstruction 3D hiérarchique : des imprécisions importantes de  $\pm 20^\circ$  (respectivement,  $\pm 10^\circ$ ) degrés sont habituellement (respectivement, toujours) tolérées pour  $\alpha_{up}, \alpha_{down}$ . De plus, les angles estimés  $\hat{\alpha}_{up} = r^{-1}(r_{up})$  et  $\hat{\alpha}_{down} = r^{-1}(r_{down})$  sont raisonnablement stables pour chaque séquence et physiquement plausibles.

	$\alpha_u$	$\alpha_d$	$\alpha_u$	$\alpha_d$	$\alpha_u$	$\alpha_d$	$\alpha_u$	$\alpha_d$	$\alpha_u$	$\alpha_d$
I	40	140	20	160	60	120	20	120	60	160
F	46	143	45	146	46	145	45	159	(46)	(144)
C	43	138	43	138	54	131	35	139	42	131
M	45	143	(45)	(143)	45	143	43	141	45	143

FIG. 8 – Chaque paire de colonnes  $\alpha_u, \alpha_d$  donne les estimations d’angles de champ de vue pour chacune des 3 séquences Fontaine, Chemin et Maison. Ces estimations sont obtenues par reconstruction 3D avec le modèle à centre unique et les angles initiaux définis dans la ligne I. Les parenthèses indiquent une expérience qui échoue à cause d’angles initiaux trop différents de leur valeurs exactes. Les valeurs données entre parenthèses indiquent le résultat d’une autre expérience obtenue avec des angles initiaux deux fois plus proches de ceux donnés par le constructeur du miroir que dans la ligne I.

**Quelques détails techniques** Les temps de calculs pour des images de taille  $1632 \times 1224$  avec un P4 2.8GHz/800MHz sont d’environ 10s pour chaque calcul sur une image (détection de points, arêtes et cercles), 20s pour chaque calcul sur une paire d’images (mise en correspondance, matrice essentielle), et 5 min pour la reconstruction hiérarchique appliquée à la séquence Maison (soit environ 1 heure en tout). Environ 700-1100 points mis en correspondances vé-

rifient la contrainte epipolaire entre deux images omni-directionnelles consécutives.

Tous les ajustements de faisceaux sont implémentés par calculs creux et dérivées analytiques (voir l’annexe pour le cas de l’erreur image dans le cas centre non-unique). Les erreurs angulaires et images du modèle à centre unique sont respectivement utilisées pour les reconstructions hiérarchiques et l’estimation de  $\hat{\alpha}_{up}, \hat{\alpha}_{down}$ . Chaque ajustement de faisceaux est appliqué deux fois, avec entre deux une remise à jour des points intervenants dans l’optimisation.

**4.3 Modèle à centre non-unique**

Les méthodes à centre non-unique sont aussi appliquées pour introduire la connaissance du miroir dans l’estimation de la géométrie. La fonction qui définit la silhouette du miroir est un polynôme de degré 4, qui est estimé à partir de l’ombre montrée en figure 5 après rectification. Une fois que l’initialisation à centre non-unique est effectuée à partir de l’estimation à centre unique en mesurant  $z^p = 48cm$ , on applique d’abord l’ajustement de faisceaux angulaire car il est plus rapide et plus robuste aux conditions initiales, puis l’ajustement de faisceaux image. Tout comme dans le cas centre unique, l’utilisation de ce dernier permet d’obtenir une estimation au maximum de vraisemblance si on suppose que les points images sont normalement distribués autour de leurs positions exactes.

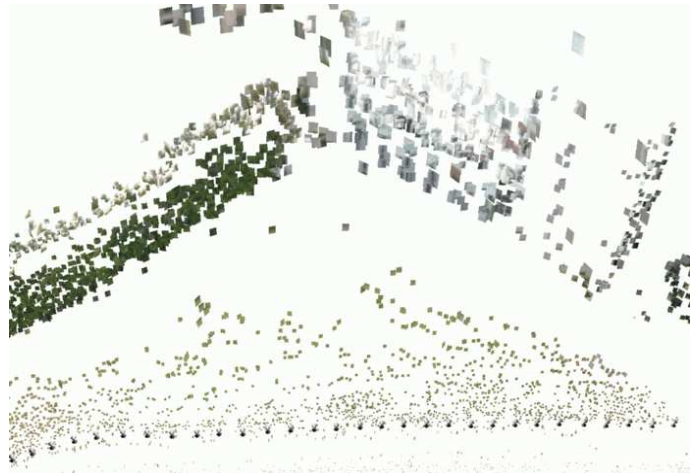


FIG. 9 – Vue locale de la reconstruction à centre non-unique du Chemin, avec une rangée de sapins à gauche, la façade d’un bâtiment à droite, et le sol. La taille des morceaux de plans carrés (dont la seule utilité est la visualisation) est choisie proportionnelle à la distance entre le point 3D reconstruit correspondant et la caméra la plus proche qui a contribué à sa reconstruction.

**Images réelles** Les reconstructions à centre non-unique pour la Fontaine, Chemin et Maison sont qualitativement similaires à celles montrées dans la figure 6. Une vue locale de la reconstruction du Chemin est aussi montrée dans la figure 9.

Des estimations des paramètres de caméra perspectives ont aussi été tentées en ajoutant des paramètres supplémentaires



à estimer dans un ajustement de faisceaux image final, en imposant plusieurs contraintes comme des distances focales/orientation/position fixées, communes à toutes les caméras ou indépendantes. La conclusion est que ces paramètres (ainsi que le facteur d'échelle 3D de la scène) sont difficiles à estimer précisément : ils évoluent trop lentement pendant l'optimisation à partir de différentes valeurs physiquement plausibles.

**Estimation du facteur d'échelle 3D de la scène** Des expériences de synthèse confirment la difficulté d'estimer le facteur d'échelle 3D de la scène, bien que ceci soit en principe possible avec le modèle à centre non-unique. On définit une reconstruction 3D "vérité terrain" par une scène de type Fontaine (un demi-tour seulement) avec pour rayon de trajectoire  $r_t$ , les orientations du miroir  $R_i$  perturbées autour de  $I_3$ , 20 caméras et 1000 points bien répartis dans l'espace 3D et les images. On ajoute d'abord un bruit gaussien avec  $\sigma = 1$  pixel aux reprojections des points 3D dans les images, et toutes les positions de miroirs  $t_i$  et points  $\tilde{X}_j$  subissent une homothétie de facteur  $s_{ini}$ . Ensuite, on applique successivement à la reconstruction 3D perturbée ainsi obtenue les ajustements de faisceaux angulaires et images, avec les paramètres de caméra perspective connus ( $z^p = 48cm$ ) tout en gérant les points inconsistants avec la géométrie courante. La table 1 montre l'erreur RMS (en pixels) et le rapport  $s_{rec}$  entre l'échelle 3D estimée de la scène et sa valeur exacte. On note que le RMS n'a pas de minimum clair, et que l'échelle 3D estimée est meilleure pour un petit rayon  $r_t$ .

	$r_t = 25cm$					$r_t = 2.5m$				
$s_{ini}$	0.5	.707	1	1.41	2	0.5	.707	1	1.41	2
$s_{rec}$	.95	.956	.990	.993	.995	.65	.812	.908	1.19	1.69
$rms$	.98	.974	.971	.970	.969	.96	.967	.967	.969	.970

TAB. 1 – Le rapport  $s_{rec}$  entre le facteur d'échelle estimée et sa valeur exacte est obtenu par des ajustements de faisceaux que l'on applique sur une reconstruction 3D (d'une scène synthétique), initialement perturbée par une homothétie et un bruit dans les images de  $\sigma = 1$  pixel. Les meilleures valeurs de  $s_{rec}$  sont obtenues pour la trajectoire avec le plus petit rayon  $r_t = 25cm$  (le résultat idéal est  $s_{rec} = 1$ ).

#### 4.4 Comparaisons entre modèles de caméras

Des comparaisons quantitatives entre reconstructions 3D réelles obtenues avec les deux modèles de caméras sont données.

**Consistance** La table 2 montre les consistances obtenues entre plusieurs reconstructions 3D et les points d'intérêts appariés donnés pour les 3 séquences selon plusieurs critères. Il y a 4 modèles de caméras : à centre unique avec la fonction de calibrage  $r(\alpha)$  linéaire et cubique, et à centre non-unique avec des paramètres de caméra perspective optimisés ou pas par ajustement de faisceaux (mêmes  $t^p$  et  $f^p$  pour toutes les caméras,  $R^p$  distincts).

Grâce à la connaissance du miroir, les consistances sont meilleures avec les modèles à centre non-unique : des amé-

liorations d'environ 5% (parfois 12% pour le Chemin) sont obtenues pour le nombre de points 3D et 2D compatibles avec la géométrie, avec des scores RMS légèrement inférieurs. Le nombre de paramètres estimés est le même pour les modèles à centre unique et non-unique lorsque les paramètres de calibrage ne sont pas raffinés.

	K	c-3d	nc-3d	c-2d	nc-2d	c-rms	nc-rms
F	n	5857	6221	31028	33262	0.84	0.78
F	y	5940	6223	32058	33876	0.78	0.76
C	n	10178	11312	50225	56862	0.87	0.82
C	y	10706	11313	54451	57148	0.78	0.79
M	n	15504	16432	75100	80169	0.83	0.78
M	y	15777	16447	77595	80311	0.76	0.76

TAB. 2 – Les consistances entre les géométries estimées et la donnée (appariements multi-vues) sont données pour les modèles à centre unique et non-unique, pour chacune des 3 séquences Fontaine, Chemin et Maison. La colonne K indique si l'expérience est faite avec ou sans raffinement des paramètres de calibrage. Chaque colonne propose une mesure de consistance : c-3d, c-2d, c-rms sont respectivement le nombre de points 3D reconstruits avec succès, le nombre de points 2d qui ont une erreur image de moins de 2 pixels, et le score image pour les modèles à centre unique. Les notations similaires nc-3d, nc-2d, nc-rms sont utilisées pour les modèles à centre non-unique.

**Différence** Les différences 3D entre les estimations à centre unique ( $t_i^c, \tilde{X}_j^c$ ) et non-unique ( $t_i^{nc}, \tilde{X}_j^{nc}$ ) sont calculées. Les différences de points et positions sont respectivement  $E_x^2 = \frac{1}{J} \sum_j \|S(\tilde{X}_j^c) - \tilde{X}_j^{nc}\|^2 / \|t_{i(j)}^{nc} - \tilde{X}_j^{nc}\|^2$  et  $E_t^2 = \frac{1}{I} \sum_i \|S(t_i^c) - t_i^{nc}\|^2$  avec  $S$  la similitude minimisant  $E_t$ ,  $I$  le nombre de caméras,  $J$  le nombre de points 3D, et  $i(j)$  l'index de la caméra la plus proche avec laquelle  $\tilde{X}_j^{nc}$  a été reconstruit. On obtient  $E_t = 0.79cm$ ,  $E_x = 0.027$  pour la Fontaine (respectivement,  $E_t = 2.6cm$ ,  $E_x = 0.017$  et  $E_t = 2.75cm$ ,  $E_x = 0.054$  pour le Chemin et la Maison). Les longueurs approximatives des trajectoires sont de 16, 42 et 22 mètres, respectivement.



FIG. 10 – Une image panoramique obtenue à partir de la séquence "contrôlée".

**Précision de la pose** Une séquence d'image (voir la figure 10) est enregistrée dans un environnement intérieur et contrôlé : le système catadioptrique se déplace sur un rail sur lequel on mesure sa position, dans une salle de dimensions environ  $7m \times 5m \times 3m$ . La trajectoire est une ligne droite de 1 mètre de long en translation, avec 6 poses équidistantes et alignées. On estime  $E_t^2 = \frac{1}{I} \sum_i \|S(t_i) - t_i^g\|^2$ , l'erreur de position, avec  $S$  la similitude minimisant  $E_t$  et  $t_i^g$  la position mesurée (supposée exacte) du miroir.

On obtient  $E_t^c = 1.1$ ,  $E_t^{nc} = 1.2$  millimètres : les deux modèles à centre unique et non-unique donnent un calcul de pose de précision voisine et de bonne qualité.

## 5 Résumé et conclusion

Les premières véritables méthodes automatiques de reconstructions 3D de scènes et de mouvements de caméra catadioptriques à partir d'une longue séquence d'images sont proposées dans ce papier, grâce à l'introduction d'ajustements de faisceaux adéquats (erreurs images et angulaires) et leurs méthodes d'initialisation de la géométrie, à la fois pour les modèles à centre unique et non-unique. Plusieurs expérimentations sur la robustesse, la précision et la comparaison entre les modèles sont effectuées. Le modèle à centre unique est une assez bonne approximation de la caméra catadioptrique testé ici, que l'on sait être à centre non-unique. De plus, le système de reconstruction 3D est décrit dans son ensemble. Bien que les résultats soient très prometteurs pour les applications futures (synthèse d'images, localisation ...), plus d'information est nécessaire pour une estimation précise des paramètres de calibrage de la caméra perspective et du facteur d'échelle 3D de la scène.

## Annexes

### Détection des ellipses frontières

La détection des ellipses intérieures et extérieures des images omni-directionnelles est un pré-traitement nécessaire aux méthodes de reconstruction 3D. Ces ellipses sont les projections par la caméra perspective des cercles frontières du miroir.

Les contours sont détectés, polygonisés en arêtes de longueur fixée, et les sommets résultats sont utilisés par des méthodes d'estimations RANSAC [7] puis Levenberg-Marquardt pour estimer une ellipse. Les sommets compatibles avec l'ellipse sont retirés de la liste de tous les sommets, et une autre ellipse est estimée de la même façon. On obtient ainsi simplement l'ellipse extérieure (grande) pour chaque image, contrairement à l'ellipse intérieure (petite) qui peut ne pas avoir suffisamment de sommets pour être estimée correctement. Les petites ellipses sont donc estimées par une méthode RANSAC en rajoutant la contrainte suivante : leur position et taille relative aux grandes ellipses sont les mêmes dans toutes les images.

En pratique, la grande ellipse se déplace parfois de quelques pixels entre images successives, et donc il faut réellement l'estimer pour chaque image. De simples estimations de cercles ont été utilisées ici, bien que des méthodes existent pour estimer des ellipses générales [8].

### Mise en correspondance initiale

La méthode de mise en correspondance classique de points d'intérêts par corrélation ne peut être appliquée directement aux images omni-directionnelles pour deux raisons : l'ambiguïté de la mise en correspondance causée par les motifs ou textures répétitifs dans une image, et les distortions géométriques entre portions/points appariés entre

deux images prises de deux points de vues différents. Les méthodes de mise en correspondances déjà publiées ne traitent que l'un de ces deux problèmes à la fois (par exemple, la relaxation [32] est utilisée pour les textures répétitives, ou les régions [20] pour les distortions géométriques dans les images omni-directionnelles).

Dans le contexte des mouvements de caméra habituels (les mouvements essentiellement de translation avec la caméra perspective pointant vers le ciel), on observe qu'une grande proportion des distortions est compensée par des rotations dans les images autour des centres des cercles frontières. Le détecteur de Harris [14] est encore une fois utilisé car il est invariant à ces rotations et car il a une bonne stabilité de détection. On compense aussi la rotation dans le voisinage des points détectés avant de comparer les voisinages de luminance de deux points à l'aide de la mesure de corrélation ZNCC (Zero Mean Normalized Cross Correlation). Pour limiter les erreurs de mise en correspondances causées par des textures répétitives, la méthode suivante est appliquée. D'abord, on essaye d'apparier les points d'intérêts d'une image avec eux même. Le résultat est une liste "réduite" de points qui retient ceux qui ne sont pas similaires aux autres selon ZNCC dans leur voisinage de recherche associés. Ensuite, les points des listes réduites de deux images différentes sont appariées en utilisant le même score de corrélation, voisinage de recherche, et seuils. Maintenant, le risque d'erreurs de mise en correspondance du aux textures répétitives est grandement réduit, mais la liste des mises en correspondances courante est très incomplète. Enfin, cette liste est complétée à l'aide d'une méthode de mise en correspondance par croissance de régions [17] : la majorité des pixels des deux images sont appariés progressivement à l'aide des contraintes de gradient de la disparité 2D et d'unicité, et deux points d'intérêts satisfaisant la fonction de mise en correspondance résultante entre les deux images sont ajoutée à la liste initiale.

Une telle liste de points d'intérêt est obtenue sans contrainte epipolaire (encore inconnue), et est adéquate dans le contexte d'images omni-directionnelles pour les méthodes d'estimations de la géométrie. Les seuls points d'intérêts suffisent dans ce contexte, contrairement au contexte précédent [18] (caméras non calibrées, champ de vue étroit, objets parfois faiblement texturés, reconstruction de surfaces lisses).

### Estimation et différentiation de $M^+(A, B)$

La minimisation de l'erreur image du modèle à centre non-unique par Levenberg-Marquardt dans la partie 3.3 nécessite des estimation et différentiation efficaces de  $M^+(A, B)$ . Cette fonction donne le point de réflexion sur la surface du miroir pour le rayon qui passe par les points  $A$  et  $B$ . Une fois que ces calculs sont effectués en  $(t^p, X)$  avec  $t^p$  le centre de la caméra perspective et un point  $X$  de la scène, tous les calculs de dérivées de la projection  $p^+(X)$  deviennent analytiques grâce au théorème de différentiation des fonctions composées. Les calculs sont très similaires

pour  $p^-(X)$ .

Le miroir est connu et est défini par le paramétrage cylindrique  $f(r, \theta) = (r \cos(\theta) \ r \sin(\theta) \ z(r)) \in \mathbb{R}^3$ . Étant donnés  $A, B \in \mathbb{R}^3$ , on recherche  $(r, \theta)$  tel que  $f(r, \theta)$  est le point de réflexion  $M^+(A, B)$  sur le miroir. On a alors

$$\vec{n}(r, \theta) \wedge \left( \frac{f(r, \theta) - A}{\|f(r, \theta) - A\|} + \frac{f(r, \theta) - B}{\|f(r, \theta) - B\|} \right) = 0$$

par la loi de réflexion, avec  $\vec{n}(r, \theta)$  la normale du miroir en  $(r, \theta)$ . Seulement 2 parmi ces 3 équations sont indépendantes : on recherche  $(r, \theta)$  tels que  $g(r, \theta, A, B) = 0$  avec  $g$  défini par

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & z'(r) \end{pmatrix} R(-\theta) \left( \frac{f(r, \theta) - A}{\|f(r, \theta) - A\|} + \frac{f(r, \theta) - B}{\|f(r, \theta) - B\|} \right)$$

et  $R(-\theta)$  est la rotation d'angle  $-\theta$  autour de l'axe des  $z$ . On obtient  $(r, \theta)$  à l'aide de la méthode de Gauss-Newton en minimisant  $\|g\|^2$ . Le théorème des fonctions implicites est ensuite appliqué à  $g$  et donne l'existence locale d'une fonction  $\mathcal{C}^1$  qui associe  $(r, \theta)$  à  $(A, B)$  et sa différentiation

$$\begin{pmatrix} d_A r & d_B r \\ d_A \theta & d_B \theta \end{pmatrix} = -(d_{(r, \theta)} g)^{-1} d_{(A, B)} g$$

en notant par  $d_X Y$  le Jacobien de la fonction  $Y(X)$  par rapport aux paramètres  $X$ . Les valeurs et différentiation de  $M^+(A, B) = f(r, \theta)$  sont alors obtenues par composition de fonctions et le théorème de différentiation des fonctions composées.

On note que tous les calculs de dérivées pour une erreur image sont essentiellement "analytiques", puisque la partie "numérique" est réduite à une seule utilisation de la méthode de Gauss-Newton pour une dimension faible (2).

## Références

[1] D. G. Aliaga, "Accurate Catadioptric Calibration for Real-time Pose Estimation in Room-size Environments," *ICCV'01*.

[2] H. Bakstein and T. Pajdla, "An Overview of Non-Central Cameras," *The Computer Winter Workshop'2001*.

[3] "Boujou," *2d3 Ltd*, <http://www.2d3.com>, 2000.

[4] P. Chang and M. Hebert, "Omnidirectional structure from motion," *OMNIVIS'00*.

[5] O.D. Faugeras, "Three-Dimensional Computer Vision - A Geometric Viewpoint," *MIT Press*, 1993.

[6] O.D. Faugeras and Q.T. Luong, "The Geometry of Multiple Images," *MIT Press*, 2001.

[7] M.A. Fischler and R.C. Bolles. "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Graphics and Image Processing*, vol. 24, no. 6, pp. 381-395, 1981.

[8] A.W. Fitzgibbon, M. Pilu and R.B. Fisher, "Direct Least Square Fitting of Ellipses," *IEEE TPAMI*, vol. 21, no. 5, 1999.

[9] A.W. Fitzgibbon, "Simultaneous Linear Estimation of Multiple View Geometry and Lens Distortion," *CVPR'01*.

[10] C. Geyer and K. Daniilidis, "A unifying Theory for Central Panoramic Systems and Practical Implications," *ECCV'00*.

[11] C. Geyer and K. Daniilidis, "Structure and Motion from Uncalibrated Catadioptric Views," *CVPR'01*.

[12] C. Geyer, T. Pajdla and K. Daniilidis, "Courses on Omnidirectional Vision," *ICCV'03*.

[13] J. Gluckman and S.K. Nayar, "Ego-Motion and Omnidirectional Cameras," *ICCV'98*.

[14] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," *Alvey Vision Conf.*, pp. 147-151, 1988.

[15] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," *Cambridge University Press*, 2000.

[16] S.B. Kang, "Catadioptric Self-Calibration," *CVPR'00*.

[17] M. Lhuillier and L. Quan, "Match Propagation for Image-Based Modeling and Rendering," *IEEE TPAMI*, vol. 24, no. 8, pp. 1140-1146, 2002.

[18] M. Lhuillier and L. Quan, "A Quasi-Dense Approach to Surface Reconstruction from Uncalibrated Images," *IEEE TPAMI*, vol. 27, no. 3, pp. 418-433, 2005.

[19] M. Lhuillier, "Automatic Structure and Motion using a Catadioptric Camera," *OMNIVIS'05*.

[20] J. Matas, O. Chum, M. Urban and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," *BMVC'03*.

[21] B. Micusik and T. Pajdla, "Estimation of Omnidirectional Camera Model from Epipolar Geometry," *CVPR'03*.

[22] B. Micusik and T. Pajdla, "Omnidirectional Camera Model and Epipolar Geometry Estimation by Ransac with Bucketing," *SCIA'03*.

[23] B. Micusik and T. Pajdla, "Autocalibration and 3D Reconstruction with Non-Central Catadioptric Cameras," *CVPR'04*.

[24] B. Micusik, "Two-View Geometry of Omnidirectional Cameras," PhD Thesis, Czech Technical University, 2004.

[25] D. Nister. "An efficient solution to the five-point relative pose problem," *CVPR'03*.

[26] D. Nister, O. Naroditsky and J. Bergen, "Visual Odometry," *CVPR'04*.

[27] T. Pajdla, T. Svoboda and V. Hlavac, "Epipolar geometry of central catadioptric cameras," *IJCV*, vol. 49, no. 1, pp. 23-37, 2002.

[28] M. Pollefeys, R. Koch and L. Van Gool, "Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters," *ICCV'98*.

[29] S.M. Seitz, "The Space of All Stereo Images," *ICCV'01*.

[30] B. Triggs, P.F. McLauchlan, R.I. Hartley and A. Fitzgibbon. "Bundle adjustment - a modern synthesis," *Vision Algorithms: Theory and Practice*, 2000.

[31] Y. Wu, H. Zhu, Z. Hu and F. Wu. "Camera Calibration from the Quasi-affine Invariance of Two Parallel Circles," *ECCV'04*.

[32] Z. Zhang, R. Deriche, O. Faugeras and Q.T. Luong, "A Robust Technique for Matching Two Uncalibrated Images through the Recovery of the Unknown Epipolar Geometry," *AI*, vol. 78, pp. 87-119, 1995.