



HAL
open science

Gyroscope visuel sphérique basé mélange de potentiels photométriques

Guillaume Caron, Fabio Morbidi

► **To cite this version:**

Guillaume Caron, Fabio Morbidi. Gyroscope visuel sphérique basé mélange de potentiels photométriques. *Reconnaissance des Formes, Image, Apprentissage et Perception*, RFIAP, Jun 2018, Marne-la-vallée, France. hal-01836032

HAL Id: hal-01836032

<https://hal.science/hal-01836032>

Submitted on 12 Jul 2018

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.

Gyroscope visuel sphérique basé mélange de potentiels photométriques

Guillaume Caron

Fabio Morbidi

Université de Picardie Jules Verne, Laboratoire MIS EA4290, Amiens
{guillaume.caron, fabio.morbidi}@u-picardie.fr

Résumé

Dans cet article, nous présentons un nouveau gyroscope visuel omnidirectionnel pour robot mobile et volant. Le gyroscope estime l'orientation 3D d'un robot embarquant une caméra sphérique en comparant l'image courante avec celle acquise à une pose de référence. Cette comparaison se fait à l'aide d'une nouvelle mesure de similarité entre images : le Mélange de Potentiels Photométriques. Considérée au sein d'une optimisation non-linéaire classique, cette mesure offre un domaine de convergence étendu et estime l'attitude précisément et de façon robuste, comme le montrent les résultats pour une caméra à deux lentilles fisheye montée sur l'effecteur d'un bras robotique et un drone à voilure fixe.

Mots Clef

Vision omnidirectionnelle, gyroscope, recalage direct

Abstract

In this paper, we present a new direct omnidirectional visual gyroscope for mobile and flying robots. The gyroscope estimates the 3D orientation of a robot embedding a spherical camera by comparing the current image with that acquired at a reference pose. Comparison is done thanks to a new image-similarity measure : the Mixture of Photometric Potentials. Considered in a classical nonlinear optimization, the latter measure offers an extended convergence domain and provides accurate and robust attitude estimates, as shown by results got for a twin-fisheye camera mounted on the end-effector of a robot arm and on a fixed-wing flying robot.

Keywords

Omnidirectional vision, gyroscope, direct registration

1 Introduction

1.1 Contexte et état de l'art

Les caméras apportent beaucoup d'informations sur la scène observée et sont largement utilisées en robotique par leur taille et leur masse réduites, pour un coût modéré. Estimer l'orientation d'une caméra à partir de l'image qu'elle acquière (requête) par rapport à une image de référence est un problème classique en robotique [1]. Les compas visuels estiment le cap du robot mobile embarquant

la caméra, les gyroscopes visuels calculent l'attitude complète (angles de roulis, tangage et lacet), alors que l'odométrie visuelle exploite une séquence d'images complète entre la référence et la requête, pour estimer la pose du robot (position et orientation).

Les caméras omnidirectionnelles étendent considérablement le champ de vue du robot qui l'embarque. Les images omnidirectionnelles peuvent être obtenues en montant un système catadioptrique [2] ou une lentille fisheye [3] sur un boîtier de caméra. Les miroirs et les lentilles peuvent aussi être combinées pour obtenir une vue sphérique complète par une seule matrice photosensible autour d'un point de vue [4]. Ce type de montage permet d'obtenir une vue sphérique sans motorisation et avec un nombre minimal de lentilles.

De récents travaux ont été menés sur de nouvelles représentations de l'image (Spherepix) pour son traitement [5] et pour l'estimation d'attitude [6].

La conception de gyroscopes visuels pour des caméras omnidirectionnelles est un sujet assez fouillé en robotique mobile mais les méthodes existantes [7, 8, 9] reposent toutes sur des primitives géométriques extraites des images (ex. : la projection dans l'image de droites 3D parallèles, les points de fuite, des primitives SIFT) pour estimer l'attitude de la caméra.

Cependant, l'extraction et le suivi de ces primitives requièrent du traitement d'image complexe, sensible au bruit et aux changements d'illumination. Ces méthodes basées primitives géométriques reposent, de plus, sur des hypothèses à propos de la structure de la scène observée, les limitant à certains environnements, par exemple à des environnements construits, riches en droites parallèles [7, 8]. Pour pallier à ces problèmes, les intensités de tous les pixels de l'image, interprétée comme un signal 2D, peuvent être exploitées pour estimer la pose de la caméra. Considérer tous les pixels apporte, de plus, une redondance d'information menant à une robustesse et une précision accrues. Ces approches sont appelées *denses*, directes ou globales, en fonction du contexte, et deux classes en ont émergé : les méthodes basées apparence [10, 11], dans le domaine de l'image, et les méthodes basées analyse harmonique dans le domaine fréquentiel pour l'estimation d'attitude [12] et de mouvement rigide dans l'espace [13].

Enfin, des méthodes hybrides ont récemment émergé, tel que SVO [14] (Semidirect Visual Odometry), qui bénéficie

cient à la fois des avantages des méthodes basées primitives géométriques et des méthodes denses.

1.2 Contribution, organisation et notations

Dans cet article, nous proposons un gyroscope visuel dense reposant sur la nouvelle représentation des intensités d’une image sphérique qu’est le *Mélange de Potentiels Photométriques* (MPP). Cette représentation se base sur [15], où des mélanges de gaussiennes photométriques 2D sont exploitées pour étendre le domaine de convergence de l’asservissement visuel photométrique. A partir de la forme analytique simple du MPP, une mesure de similarité entre images est construite. Cette mesure peut-être optimisée par un algorithme classique de type Gauss-Newton ou Levenberg-Marquardt, potentiellement combiné à un M-Estimeur pour plus de robustesse.

Le gyroscope visuel proposé présente quelques propriétés remarquables. Il est précis, même quand les images sphériques sont de faible définition, et facile à régler, grâce à l’unique paramètre, pour une définition d’image donnée, qui contrôle la largeur des potentiels photométriques (fonctions gaussiennes dans cet article) équivalent au “degré d’influence” de chaque pixel sur ses voisins. Contrairement à la majorité des gyroscopes visuels existants, il fonctionne aussi bien dans des environnements structurés que non-structurés et ne demande aucun traitement d’image (les gradients de l’image ne sont pas à calculer). De plus, les expérimentations faites à l’extérieur en conditions réelles suggèrent qu’il est robuste aux variations d’éclairage et aux occultations partielles.

Dans la suite de l’article, la partie 2 présente le modèle mathématique de caméra sphérique considéré dans ce travail. La partie 3 décrit comment représenter une image sphérique par un mélange de potentiels photométriques. Dans la partie 4, la fonction de coût basée MPP est introduite et exploitée pour construire le gyroscope visuel. Les résultats d’expérimentations, en conditions réelles, menées avec une caméra Ricoh Theta S montée sur l’effecteur d’un bras robotique et sur un drone à voilure fixe sont présentés et discutés dans la partie 5. Enfin, la partie 6 synthétise les principales contributions de l’article et indique quelques pistes de travaux futurs.

Notations : \mathbb{R}^n indique l’espace euclidien de dimension n , $\mathbb{R}^{m \times n}$ l’ensemble des matrices de dimensions $m \times n$, $\mathbf{I}_{n \times n}$ la matrice identité de dimensions $n \times n$ et $\text{SO}(3)$, $\text{SE}(3)$ le groupe spécial orthogonal 3D et le groupe spécial Euclidien, respectivement. Pour une matrice \mathbf{A} carrée et non-singulière, \mathbf{A}^{-1} indique son inverse et $\det(\mathbf{A})$ son déterminant. Enfin, $\|\mathbf{x}\|$ est la norme euclidienne de $\mathbf{x} \in \mathbb{R}^n$, $[\mathbf{z}]_{\times}$ la matrice anti-symétrique associée au vecteur $\mathbf{z} \in \mathbb{R}^3$, et \triangleq l’égalité par définition. \diamond

2 Modélisation de caméra sphérique

Certaines caméras à lentille fisheye se sont montrées équivalentes, approximativement, à un système catadioptrique central [16]. Leur fonction de projection peut donc se re-

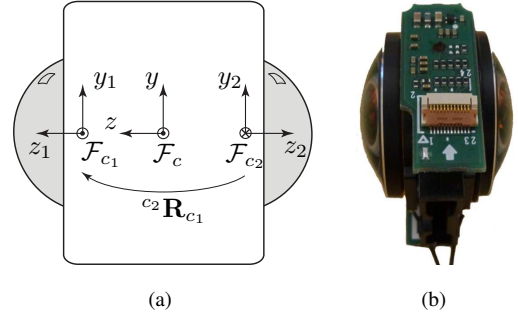


FIGURE 1 – La caméra double-fisheye Ricoh Theta : (a) Représentation schématique (vue de dessus) : les lentilles, de repères \mathcal{F}_{c_1} et \mathcal{F}_{c_2} , sont grisées. Dans cet article, \mathcal{F}_c coïncide avec \mathcal{F}_{c_1} et les origines de \mathcal{F}_{c_1} et \mathcal{F}_{c_2} sont confondues ; (b) Vue de dessus de la caméra sans boîtier.

présenter à l’aide du modèle de projection centrale unifié, faisant intervenir deux projections successives [17], la première sur une sphère unitaire et, la seconde, sur le plan image par projection perspective.

Cet article se concentre sur les caméras sphériques combinant deux lentilles fisheye identiques, orientées dans des directions opposées mais alignées (ex. : les caméras Ricoh Theta et Samsung Gear 360). Deux jeux de paramètres intrinsèques sont considérés, un pour chaque lentille fisheye, $\mathcal{P}_{c_j} = \{\alpha_{u_j}, \alpha_{v_j}, u_{0_j}, v_{0_j}, \xi_j\}$, $j \in \{1, 2\}$, où α_{u_j} et α_{v_j} sont des facteurs d’échelle (axe horizontal et axe vertical de l’image) du plan image normalisé au plan image numérique (coordonnées en pixels), u_{0_j} et v_{0_j} sont les coordonnées du point principal (pixels), et ξ_j est la distance entre le centre de la sphère et le deuxième centre de projection (perspective) de la caméra j (Fig. 1).

Enfin, les *paramètres extrinsèques* décrivent la rotation 3D entre les axes optiques z_1 et z_2 des lentilles fisheye. En utilisant la représentation axe-angle d’une rotation dans l’espace 3D sous la forme d’un vecteur unitaire $\mathbf{w}_{1,2} \in \mathbb{R}^3$ et un angle $\theta_{1,2}$, les paramètres extrinsèques sont les éléments du vecteur $\mathbf{r}_{1,2} = \theta_{1,2} \mathbf{w}_{1,2}$. Grâce aux formules de Rodrigues, la matrice de rotation ${}^{c_2}\mathbf{R}_{c_1}$ entre le repère de la première lentille fisheye \mathcal{F}_{c_1} et le repère de la seconde, \mathcal{F}_{c_2} , se calcule à partir de $\mathbf{w}_{1,2}$ et $\theta_{1,2}$ (et vice-versa).

Les dix paramètres intrinsèques et extrinsèques de la caméra sphérique peuvent être calculés simultanément à partir de points extraits d’images d’une mire d’étalonnage connue. Une fois que les paramètres de la caméra ont été calculés, en supposant que le repère \mathcal{F}_c de la caméra coïncide avec \mathcal{F}_{c_1} , l’élévation de l’image sur la sphère unitaire s’exprime par rapport à \mathcal{F}_{c_2} via ${}^{c_2}\mathbf{R}_{c_1}$, et se projette, grâce aux paramètres \mathcal{P}_{c_j} , $j \in \{1, 2\}$, dans la zone de l’image correspondant à la lentille fisheye qui le perçoit. La projection sphérique d’un point 3D ${}^c\mathbf{X} = [{}^cX, {}^cY, {}^cZ]^T$, exprimé dans \mathcal{F}_c , en un point

${}^c\mathbf{X}_S = {}^c\mathbf{X}/\|{}^c\mathbf{X}\|$ appartenant à l'image sphérique, cela nous permet de nous restreindre à une sphère unitaire, sans perte de généralité. Plus précisément, nous pouvons admettre que ${}^c\mathbf{X}_S$ repose directement à la surface de la sphère du modèle de projection centrale unifiée.

Dans cet article, une image sphérique est représentée par un ensemble de points (pixels) *répartis uniformément*, au sens géodésique, sur la surface de la sphère. L'ensemble discret de points ${}^c\mathbf{X}_{S_i}$, $i \in \{1, 2, \dots, P\}$ est uniformément réparti sur la surface de la sphère quand il est construit à partir d'un icosaèdre convexe régulier (un polyèdre de vingt triangles équilatéraux et douze sommets), dont les faces subissent N subdivisions, N étant proportionnel à la définition désirée de l'image sphérique [18]. A noter que les sommets d'un icosaèdre convexe régulier sont régulièrement distribués sur la sphère unitaire circonscrite au polyèdre : cette propriété reste valide pour un icosaèdre dont les faces ont été divisées récursivement. Pour N niveaux de subdivision, le polyèdre correspondant possède $P = \frac{1}{2}(20 \times 4^N) + 2$ sommets et 20×4^N facettes.

3 MPP

L'objectif est de concevoir un gyroscope visuel qui estime l'orientation relative entre une image sphérique I_S , acquise par la caméra à l'instant t et une image sphérique de référence I_S^* . Cette orientation relative se représente par un axe et un angle combinés en un vecteur $\mathbf{r} = \theta \mathbf{w}$, tel que $\|\mathbf{w}\| = 1$. Puisque \mathbf{r} minimise la différence entre $I_S(\mathbf{r})$ et I_S^* , ces dernières peuvent être comparées par la *Somme des Carrés des Ecartés* (SCE), donnant la fonction de coût :

$$C_{\text{SCE}}(\mathbf{r}) = \|\mathcal{I}_S(\mathbf{r}) - \mathcal{I}_S^*\|, \quad (1)$$

dans laquelle les intensités $I_S(\mathbf{r})$ et I_S^* , des images en niveaux de gris, aux points ${}^c\mathbf{X}_{S_i}$, $i \in \{1, 2, \dots, P\}$, ont été empilées en vecteurs $\mathcal{I}_S(\mathbf{r})$, $\mathcal{I}_S^* \in \mathbb{R}^P$, respectivement. Tout en étant intuitive et facile à relier aux degrés de liberté (DDL) de rotation de la caméra (dans le cas des images hémisphériques [19]), $C_{\text{SCE}}(\mathbf{r})$ est connue pour présenter, en pratique, un domaine de convergence étroit autour de son minimum. C'est un inconvénient majeur qui justifie l'introduction d'une nouvelle représentation d'image sphérique. Nous proposons alors d'adapter la notion de mélange de gaussiennes photométriques 2D (initialement proposé pour l'asservissement visuel de caméra perspective [15]), pour que notre gyroscope visuel présente un domaine de convergence étendu. De ce fait, une couverture quasi-complète de $\text{SO}(3)$ sera atteinte (cf. partie 5).

Rappelons que la *fonction mélange* $G(\mathbf{x})$ de m fonctions de densités de probabilités (fdp) gaussiennes hétéroscédastiques s'écrit [20, Ch. 3] :

$$G(\mathbf{x}) = \sum_{i=1}^m w_i \phi(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \quad (2)$$

où

$$\phi(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma}_i)}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right), \quad (3)$$

indique la fdp gaussienne d'un vecteur aléatoire $\mathbf{x} \in \mathbb{R}^n$ de moyenne $\boldsymbol{\mu}_i \in \mathbb{R}^n$ et de matrice de covariance définie positive $\boldsymbol{\Sigma}_i \in \mathbb{R}^{n \times n}$, $i \in \{1, \dots, m\}$, et,

$$\sum_{i=1}^m w_i = 1, \quad w_i \geq 0, \quad i \in \{1, \dots, m\}. \quad (4)$$

De cette définition, $G(\mathbf{x})$ de (2) est la *combinaison convexe* de m fdp gaussiennes avec les pondérations w_1, \dots, w_m . En s'inspirant de (2) et en étendant la représentation de [15] aux images sphériques ($n = 3$), on définit le *Mélange de Potentiels Photométriques* au point ${}^c\mathbf{X}_{S_g}$ par :

$$G_S({}^c\mathbf{X}_{S_g}, I_S) = \sum_{i=1}^P \bar{\mathcal{I}}_S({}^c\mathbf{X}_{S_i}) \frac{1}{\lambda_g^3 (2\pi)^{3/2}} \exp\left(-\frac{(\text{D}^c\mathbf{X}_{S_i})^2}{2\lambda_g^2}\right), \quad (5)$$

où $\text{D}^c\mathbf{X}_{S_i} = \arccos({}^c\mathbf{X}_{S_g}^T {}^c\mathbf{X}_{S_i})$ est la distance géodésique entre ${}^c\mathbf{X}_{S_g}$ et ${}^c\mathbf{X}_{S_i}$ sur la sphère unitaire \mathcal{S}^2 et $\bar{\mathcal{I}}_S({}^c\mathbf{X}_{S_i}) \geq 0$ décrit l'*intensité normalisée de l'image* au point ${}^c\mathbf{X}_{S_i}$ tel que $\sum_{i=1}^P \bar{\mathcal{I}}_S({}^c\mathbf{X}_{S_i}) = 1$ (cf. équation (4)). Le *paramètre d'extension* $\lambda_g > 0$ est réglable pour ajuster la largeur des gaussiennes de (5).

Avant d'établir l'algorithme du gyroscope visuel basé MPP (cf. partie 4), il est important de noter deux différences majeures entre les équations (2) et (5) :

- Dans (5), on pose $\boldsymbol{\mu}_i = {}^c\mathbf{X}_{S_i}$ et $\boldsymbol{\Sigma}_i = \lambda_g^2 \mathbf{I}_{3 \times 3}$, $\forall i$, c'est-à-dire que l'on se restreint à un *mélange isotrope homoscédastique* [20, Sect. 3.3]. Le carré de la distance de Mahalanobis $(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)$ de (3) se réduit au carré d'une distance géodésique pondérée, dans notre cas.
- Alors que l'intuition s'apparente aux probabilités et aux statistiques, il n'y a rien de stochastique dans la représentation de l'image sphérique dans (5). Le fait que $G_S({}^c\mathbf{X}_{S_g}, I_S)$ puisse représenter une fdp n'est pas exploité.

Remarque 1 (Généralité de la représentation du MPP)
Notons que (5) est une instance d'une représentation d'image sphérique plus générale. En fait, les gaussiennes de l'équation (5) sont utilisées pour régler le pouvoir d'attraction (ou le rayon d'influence) de chaque pixel ${}^c\mathbf{X}_{S_g}$ de l'image. Elles peuvent être remplacées par d'autres fonctions noyau (à support infini) considérées en statistiques non-paramétriques [21]. \diamond

4 Gyroscope visuel basé MPP

Soit ${}^c\mathbf{R}_{c^*} \in \text{SO}(3)$, la matrice de rotation *inconnue* entre l'image de référence I_S^* et l'image courante I_S que notre

gyroscope visuel vise à estimer. En adaptant (5), on introduit le MPP au point ${}^{c^*}\mathbf{X}_{Sg}$:

$$G_S({}^{c^*}\mathbf{X}_{Sg}, I_S, {}^c\mathbf{R}_{c^*}) \triangleq \sum_{i=1}^P H_S({}^c\mathbf{X}_{Sg}, {}^c\mathbf{X}_{Si}, I_S, {}^c\mathbf{R}_{c^*})$$

$$= \sum_{i=1}^P \bar{\mathcal{I}}_S({}^c\mathbf{X}_{Si}) \frac{1}{\lambda_g^3 (2\pi)^{3/2}} \exp\left(-\frac{(\mathbf{D}^c\mathbf{X}_{Si})^2}{2\lambda_g^2}\right), \quad (6)$$

où ${}^c\mathbf{X}_{Sg} = {}^c\mathbf{R}_{c^*} {}^{c^*}\mathbf{X}_{Sg}$, ${}^c\mathbf{X}_{Si} = {}^c\mathbf{R}_{c^*} {}^{c^*}\mathbf{X}_{Si}$ et $\|{}^{c^*}\mathbf{X}_{Sg}\| = \|{}^{c^*}\mathbf{X}_{Si}\| = 1$. Remplaçons ensuite la fonction de coût SCE de (1) par :

$$\mathcal{C}_{MPP}(\mathbf{r}) \triangleq \|\mathcal{C}_{MPP}(\mathbf{r})\| = \|\mathbf{G}_S(I_S, {}^c\mathbf{R}_{c^*}(\mathbf{r})) - \mathbf{G}_S(I_S^*)\|, \quad (7)$$

où $\mathbf{G}_S(I_S, {}^c\mathbf{R}_{c^*}(\mathbf{r}))$, $\mathbf{G}_S(I_S^*) \in \mathbb{R}^P$ empilent les valeurs de $G_S({}^{c^*}\mathbf{X}_{Sg}, I_S, {}^c\mathbf{R}_{c^*}(\mathbf{r}))$ et, respectivement, de $G_S(I_S^*)$ des P pixels des images sphériques. ${}^c\mathbf{R}_{c^*}(\mathbf{r})$ indique que la matrice de rotation est calculée à partir du vecteur ‘‘axe-angle’’ \mathbf{r} . Un estimé $\hat{\mathbf{r}}$ de \mathbf{r} peut s’obtenir par minimisation de (7) à l’aide de l’algorithme de Gauss-Newton. En initialisant cet algorithme itératif avec $\mathbf{r}^{(0)}$ (itération 0), le calcul de $\mathbf{r}^{(k+1)}$ à partir de $\mathbf{r}^{(k)}$ nécessite de calculer ${}^{c^{(k+1)}}\mathbf{R}_{c^*}$ en composant les matrices de rotation déterminées à partir de $\mathbf{r}^{(k)}$ et de l’incrément $\Delta\mathbf{r}^{(k)}$, c’est-à-dire :

$${}^{c^{(k+1)}}\mathbf{R}_{c^*} = {}^{c^{(k+1)}}\mathbf{R}_{c^{(k)}}(\Delta\mathbf{r}^{(k)}) {}^{c^{(k)}}\mathbf{R}_{c^*}(\mathbf{r}^{(k)}), \quad (8)$$

avec $k \in \{0, 1, \dots\}$. D’autre part, $\Delta\mathbf{r}^{(k)}$ s’exprime à partir d’un développement limité de Taylor au premier ordre de $\mathcal{C}_{MPP}(\mathbf{r})$ en $\hat{\mathbf{r}}$:

$$\mathcal{C}_{MPP}(\hat{\mathbf{r}}) \simeq \mathcal{C}_{MPP}(\mathbf{r}^{(k)}) + \frac{\partial \mathcal{C}_{MPP}(\mathbf{r}^{(k)})}{\partial \mathbf{r}^{(k)}} \Delta\mathbf{r}^{(k)}. \quad (9)$$

Ensuite, en posant (9) égale à zéro, on a :

$$\Delta\mathbf{r}^{(k)} = -\gamma \left[\frac{\partial \mathcal{C}_{MPP}(\mathbf{r}^{(k)})}{\partial \mathbf{r}^{(k)}} \right]^\dagger \mathcal{C}_{MPP}(\mathbf{r}^{(k)}), \quad (10)$$

où $[\cdot]^\dagger$ est la pseudo-inverse à gauche (de Moore-Penrose) et γ est un gain positif. Ainsi, l’équation (8) permet de mettre à jour l’orientation estimée de la caméra à chaque itération k jusqu’à convergence. $\mathbf{G}_S(I_S^*)$ étant constant, la matrice jacobienne de dimensions $P \times 3$ de (10) est :

$$\frac{\partial \mathcal{C}_{MPP}(\mathbf{r}^{(k)})}{\partial \mathbf{r}^{(k)}} = \frac{\partial \mathbf{G}_S(I_S, {}^{c^{(k)}}\mathbf{R}_{c^*})}{\partial \mathbf{r}^{(k)}}. \quad (11)$$

En appliquant la règle des dérivées en chaîne à (11) et en factorisant l’expression, on obtient :

$$\frac{\partial \mathcal{C}_{MPP}(\mathbf{r}^{(k)})}{\partial \mathbf{r}^{(k)}} = \begin{bmatrix} \vdots \\ \left(\sum_{i=1}^P \frac{\partial H_S}{\partial {}^{c^{(k)}}\mathbf{X}_{Sg}} \right) \frac{\partial {}^{c^{(k)}}\mathbf{X}_{Sg}}{\partial \mathbf{r}^{(k)}} \\ \vdots \end{bmatrix}, \quad (12)$$

où les paramètres de H_S de (6) ont été omis par souci de compacité.

Il reste à exprimer les deux dérivées partielles de l’opérateur de droite dans l’équation (12).

En supposant ${}^{c^{(k)}}\mathbf{X}_{Si} \neq {}^{c^{(k)}}\mathbf{X}_{Sg}^T$, la dérivée partielle de H_S par rapport à ${}^{c^{(k)}}\mathbf{X}_{Sg}$, un vecteur 1×3 , est donnée par :

$$\frac{\partial H_S}{\partial {}^{c^{(k)}}\mathbf{X}_{Sg}} = \frac{\bar{\mathcal{I}}_S({}^{c^{(k)}}\mathbf{X}_{Si})}{\lambda_g^5 (2\pi)^{3/2}} \mathbf{D}^{c^{(k)}}\mathbf{X}_{Si} \exp\left(-\frac{(\mathbf{D}^{c^{(k)}}\mathbf{X}_{Si})^2}{2\lambda_g^2}\right) \cdot \frac{{}^{c^{(k)}}\mathbf{X}_{Si}^T}{\sqrt{1 - ({}^{c^{(k)}}\mathbf{X}_{Sg}^T {}^{c^{(k)}}\mathbf{X}_{Si})^2}},$$

où ${}^{c^{(k)}}\mathbf{X}_{Sg} = {}^{c^{(k)}}\mathbf{R}_{c^*} {}^{c^*}\mathbf{X}_{Sg}$, ${}^{c^{(k)}}\mathbf{X}_{Si} = {}^{c^{(k)}}\mathbf{R}_{c^*} {}^{c^*}\mathbf{X}_{Si}$.

La dérivée partielle de ${}^{c^{(k)}}\mathbf{X}_{Sg}$ par rapport à $\mathbf{r}^{(k)}$, une matrice 3×3 , s’obtient à partir de la formule bien connue reliant le mouvement d’un point 3D à celui de la caméra, en considérant une vitesse de translation *nulle* :

$${}^{c^{(k)}}\dot{\mathbf{X}}_{Sg} = -\boldsymbol{\omega}_c \times {}^{c^{(k)}}\mathbf{X}_{Sg} = [{}^{c^{(k)}}\mathbf{X}_{Sg}]_\times \boldsymbol{\omega}_c,$$

où $\boldsymbol{\omega}_c \in \mathbb{R}^3$ est la vitesse angulaire du repère de la caméra \mathcal{F}_c , équivalente à l’incrément $\Delta\mathbf{r}^{(k)}$ en suivant les notations de cette partie. En conclusion, on a :

$$\frac{\partial {}^{c^{(k)}}\mathbf{X}_{Sg}}{\partial \mathbf{r}^{(k)}} = [{}^{c^{(k)}}\mathbf{X}_{Sg}]_\times.$$

Remarque 2 (Traitement d’image sphérique) *Le MPP est frugal en calculs puisqu’il n’implique aucun traitement d’image. Par exemple, contrairement aux approches directes existantes, les gradients de l’image ne sont pas à calculer par traitement d’image, tâche difficile, en particulier pour une image sphérique [5].*

Remarque 3 (Algorithme de Levenberg-Marquardt)

Pour améliorer les performances de convergence, l’algorithme de Gauss-Newton peut-être remplacé par celui de Levenberg-Marquardt. Dans ce cas, le calcul de l’incrément (10) devient :

$$\Delta\mathbf{r}^{(k)} = -\gamma (\mathbf{J}^T \mathbf{J} + \nu \text{diag}(\mathbf{J}^T \mathbf{J}))^{-1} \mathbf{J}^T \mathcal{C}_{MPP}(\mathbf{r}^{(k)}),$$

où $\mathbf{J} \triangleq \partial \mathcal{C}_{MPP}(\mathbf{r}^{(k)}) / \partial \mathbf{r}^{(k)}$, $\text{diag}(\mathbf{J}^T \mathbf{J})$ est la matrice diagonale construite à partir des éléments sur la diagonale de $\mathbf{J}^T \mathbf{J}$ et $\nu > 0$ est le facteur d’amortissement. \diamond

5 Validation expérimentale

5.1 Etalonnage de la caméra sphérique

Le gyroscope visuel proposé est évalué pour la caméra *Ricoh Theta S*. Les images sphériques sont réalisées par deux lentilles fisheye montées dos à dos, capturant la lumière dans un champ de vue de 360° (Fig. 1). La lumière est redirigée vers les deux matrices photosensibles à l’aide de

deux prismes à 90° . Dans les expérimentations, des images de 1280×720 pixels sont acquises au format double-fisheye¹ (deux images hémisphériques côte-à-côte, comme en Fig. 2(b)). Un masque constant, réalisé manuellement, est appliqué aux images pour n'en considérer que les parties utiles pour le gyroscope, c'est-à-dire les deux "disques" des images fisheye.

Pour l'étalonnage, la Ricoh Theta est modélisée comme un système stéréoscopique avec deux caméras fisheye partageant le même centre optique. Des images de mire à quatre plans sont acquises pour l'étalonnage (Fig. 2(a)). Une variante de l'algorithme d'étalonnage de [22] a été mise en oeuvre en fixant la translation entre les deux caméras fisheye à zéro et en estimant uniquement la rotation. Cela a mené à $\mathcal{P}_{c_1} = \{577, 77; 576, 11; 958, 66; 316, 90; 1, 99\}$, $\mathcal{P}_{c_2} = \{567, 90; 565, 17; 321, 55; 319, 48; 1, 94\}$, et $\mathbf{r}_{1,2} = [-0,008; 3,132; -0,011]^T$ rad (équivalent à une rotation de $179,45^\circ$ autour d'un axe très proche de y_1).

5.2 Bras robot

Afin d'avoir des rotations pures autour du centre optique, la caméra est montée sur l'effecteur d'un robot Stäubli TX-60 à 6 DDL au sein d'une pièce de $10,05 \text{ m} \times 7,03 \text{ m} \times 2,70 \text{ m}$ éclairée par des tubes au néon. L'implémentation dans la librairie ViSP² de l'algorithme de Tsai et Lenz a permis de réaliser l'étalonnage extrinsèque de la caméra par rapport à l'effecteur du robot, c'est-à-dire de calculer ${}^e\mathbf{M}_c \in \text{SE}(3)$, la transformation rigide entre le repère caméra \mathcal{F}_c et le repère effecteur \mathcal{F}_e . Six acquisitions de mire d'étalonnage ont été faites à partir de six poses différentes (Fig. 2(b)), menant à :

$${}^e\mathbf{M}_c = \begin{bmatrix} -0,0136 & 0,9997 & 0,0222 & -0,0401 \\ 0,0142 & -0,0220 & 0,9996 & 0,0000 \\ 0,9998 & 0,0139 & -0,0139 & 0,2372 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

où la translation est exprimée en mètres.

Dans les expérimentations, 5 Positions d'Acquisitions (PA) sur le même axe horizontal, de telle sorte que le centre optique de la caméra soit à 60 cm au dessus de la base du robot Stäubli, sont considérées. A chaque PA, la caméra subit une rotation de 360° autour de la verticale, avec un pas de $2,5^\circ$, donnant 144 images. En PA2, une seconde série d'acquisitions est réalisée en considérant 94 orientations 3D distinctes qui ne violent pas les contraintes mécaniques du robot. La distance entre PA0 et PA1, PA2, PA3 et PA4 est de 40, 120, 240 et 400 cm, respectivement. Les codeurs des articulations du robot Stäubli fournissent une vérité terrain très précise pour la validation du gyroscope visuel.

Estimation de cap (1 DDL). Afin d'étudier la forme de la fonction de coût C_{MPP} et pour évaluer facilement l'impact du paramètre d'extension λ_g et du niveau de subdivi-

1. Le jeu de données complet, intitulé SVMIS, est disponible au lien : http://mis.u-picardie.fr/~g-caron/pub/data/SVMIS_dataset.zip

2. <https://visp.inria.fr>

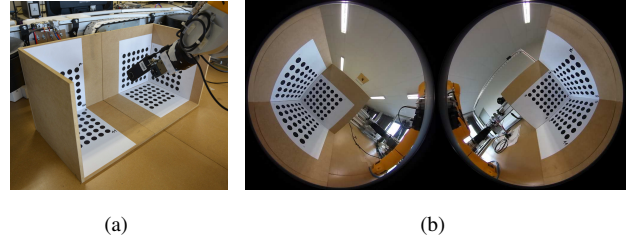


FIGURE 2 – Etalonnage de la caméra Ricoh Theta tenue par le bras robotique : (a) La mire d'étalonnage (quatre plans à disques noirs); (b) Image double-fisheye de la mire.

sions N sur C_{MPP} , la rotation pure de la caméra autour d'une seule axe est considérée, l'axe (vertical) x de \mathcal{F}_c .

L'étude se concentre sur PA2. L'impact d'un estimateur robuste (M-estimateur) sur $C_{\text{MPP}}(r)$ est aussi évalué en y intégrant une fonction de pondération (fonction de Cauchy [23, Sect. 4.8]), dont les détails sont omis par soucis de compacité et parce que très classiques. Dans un soucis d'uniformité de présentation des résultats, $C_{\text{MPP}}(r)$ est normalisée entre 0 et 1 (notée $\bar{C}_{\text{MPP}}(r)$) : en effet, $\max(C_{\text{MPP}}(r))$ varie significativement en fonction de λ_g et N (jusqu'à un facteur 3).

La Figure 3 montre la largeur du domaine de convergence de $\bar{C}_{\text{MPP}}(r)$ pour différentes valeurs de N et de λ_g . La Figure 3(a) représente la fonction de coût basée MPP pour

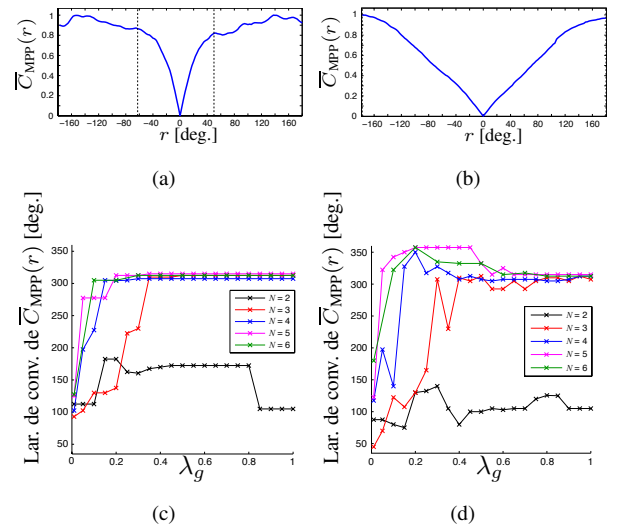


FIGURE 3 – [1 DDL Stäubli] Fonction de coût $\bar{C}_{\text{MPP}}(r)$ (tirets noirs verticaux : limites du domaine de convergence) : (a) $N = 2$, $\lambda_g = 0,01$, sans M-estimateur; (b) $N = 5$, $\lambda_g = 0,3$ avec M-estimateur. (c), (d) Largeur du domaine de convergence de $\bar{C}_{\text{MPP}}(r)$ en fonction de λ_g pour $N \in \{2, \dots, 6\}$ sans et avec M-estimateur, respectivement.

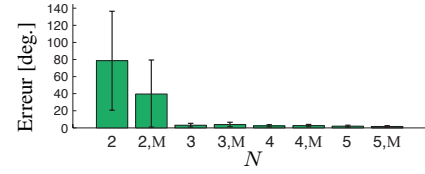
$N = 2$ et $\lambda_g = 0,01$ sans M-estimateur et la Figure 3(b) pour $N = 5$ et $\lambda_g = 0,3$ avec M-estimateur. La largeur du domaine de convergence est de 115° , $312,5^\circ$ et 360° , respectivement (pointillés tracés à la verticale). Pour $\lambda_g = 0,01$, la fonction de coût basée MPP est identique à la fonction de coût basée SCE (1).

Les largeurs du domaine de convergence de $\bar{C}_{MPP}(r)$, et donc de $C_{MPP}(r)$, la normalisation pour l’affichage n’ayant aucun impact sur cette largeur, sont indiquées aux figures 3(c)-3(d) en fonction de λ_g pour $N \in \{2, 3, \dots, 6\}$ (les valeurs considérées pour λ_g sont marquées d’une croix). Pour $\lambda_g > 1$, $\bar{C}_{MPP}(r)$ ne change plus.

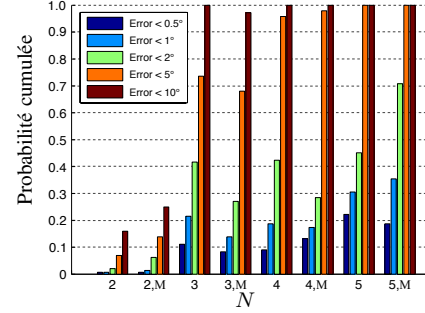
A partir de la Figure 3, on peut conclure que :

- Plus N est grand, plus $C_{MPP}(r)$ est lisse et plus le domaine de convergence est étendu,
- Plus λ_g est grand, plus le domaine de convergence est étendu (cependant, cet effet disparaît pour $\lambda_g > 0,5$ et $N > 2$),
- Le M-estimateur présente un “effet de linéarisation” de $C_{MPP}(r)$ (l’effet est plus clair quand λ_g est grand). En conséquence, le domaine de convergence s’étend (particulièrement pour $\lambda_g \in [0, 2, 0, 4]$). Cependant, pour de faibles valeurs de N , le M-estimateur perturbe $C_{MPP}(r)$, ce qui peut mener à une dégradation des résultats (cf. fortes discontinuités en Figure 3(a)),
- En conditions idéales (ex. : sans bruit), $C_{MPP}(r)$ est parfaitement symétrique par rapport à l’origine.

Dans la première série d’estimations, toujours en PA2, l’algorithme de Levenberg-Marquardt (avec $\gamma = 1$) est utilisé (critère d’arrêt de l’optimisation fixé à un seuil de 10^{-6} sur la stabilité du résidu). La précision d’estimation et la complexité de notre gyroscope sont évalués en Figure 4. Dans ces expérimentations, λ_g est fixé à $0,325$, $\gamma = 1$, et deux rotations initiales sont considérées, $r^{(0)} \in \{0, \pi\}$, pour sélectionner celle qui minimise $C_{MPP}(r)$. Cette stratégie élémentaire est suffisamment efficace pour éviter un potentiel minimum local très éloigné de l’optimum global, grâce à la largeur du domaine de convergence de $C_{MPP}(r)$. La Figure 4(a) reporte la moyenne et l’écart-type de l’erreur d’estimation angulaire. Le gyroscope est implémenté en C++ et exécuté sur un MacBook Pro pourvu d’un microprocesseur Intel Core i7 à 2,4 GHz et de 8 Go de RAM. Les temps de calculs moyens pour $N \in \{2, 3, 4, 5\}$ sont de $0,03s$, $0,50s$, $6,00s$ et $80,00s$, respectivement. Le M-Estimateur n’allonge pas le temps de calcul de façon significative. La Figure 4(b) complète la Figure 4(a, *haut*) en affichant la distribution cumulée empirique obtenue à partir de l’amplitude de l’erreur d’estimation angulaire pour $N \in \{2, 3, 4, 5\}$. Les barres verticales de la Figure 4(b) montrent le taux d’estimations dont l’erreur est inférieure à $0,5^\circ$, 1° , 2° , 5° et 10° . Par exemple, si une erreur de 5° est tolérée, environ 75% des estimations sont acceptables pour $N = 3$ et environ 95% pour $N = 4$. Ce dernier résultat est remarquable puisque, pour $N = 4$, le gyroscope utilise juste l’équivalent de deux images de $37 \text{ pixels} \times 37$



(a)



(b)

FIGURE 4 – [1 DDL Stäubli] Analyse de l’erreur d’estimation pour $N \in \{2, 3, 4, 5\}$ (a) Moyenne et écart-type de l’erreur d’estimation avec (“M”) ou sans M-estimateur; (b) Distribution cumulée empirique.

pixels (une par lentille fisheye).

Evidemment, plus N est grand, meilleure sera la précision du gyroscope, mais cela se fera au détriment du temps de calcul. En effet, puisque $P = \frac{1}{2}(20 \times 4^N) + 2$ (cf. partie 2), le temps de calcul augmente de façon exponentielle en fonction de N .

Le gyroscope s’est, enfin, montré robuste aux occultations partielles et occasionnelles causées par les mouvements du bras robotique.

Une dernière étude à 1 DDL a été réalisée pour évaluer l’impact de la *translation* sur la précision d’estimation du gyroscope. L’algorithme Levenberg-Marquardt est exécuté avec M-estimateur (comme en Figure 4) pour estimer les angles entre l’image 71 (l’image de référence) acquise en PA0 et toutes les autres images acquises en PA0, PA1, ..., PA4. La Figure 5 reporte la moyenne et l’écart-type des erreurs d’estimation angulaire en fonction de la distance (croissante) à PA0 et pour $N \in \{3, 4, 5\}$. La translation a un impact négligeable sur les erreurs d’estimation. Seul détail, pour $N = 5$ le gyroscope apparaît plus sensible aux translations. Pour $N \in \{3, 4, 5\}$, l’erreur d’estimation (pour les 5 PA) est de $2,95^\circ$, $2,49^\circ$, $1,69^\circ$, respectivement.

Estimation d’attitude (3 DDL). Le gyroscope a aussi été appliqué à l’estimation des trois DDL de rotation de la Ricoh Theta. Pour ce faire, seul le PA2 est considéré et λ_g est fixé à $0,275$, $\gamma = 1$, et $r^{(0)} = \mathbf{0}$ initialise l’algorithme de Gauss-Newton (sans M-estimateur). Ce dernier s’est avéré apporter de meilleurs résultats pour 3DDL

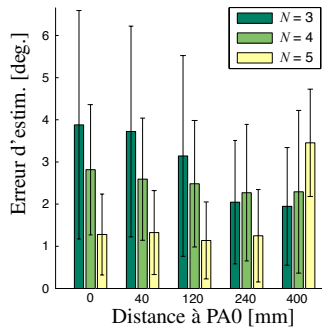


FIGURE 5 – [1 DDL Stäubli] Moyenne et écart-type de l’erreur en fonction de la distance à PA0 et pour $N \in \{3, 4, 5\}$.

que notre implémentation de l’algorithme de Levenberg-Marquardt. Le Tableau 1 reporte les statistiques d’erreur d’estimation $\|\mathbf{r} - \hat{\mathbf{r}}\|$ sur les 94 images pour $N \in \{3, 4, 5\}$, et les temps de calcul moyens correspondant, ainsi que le nombre d’itérations de l’algorithme d’optimisation.

Drone à voilure fixe. Dans cette dernière expérimentation, la Ricoh Theta a été fixée à un drone Parrot Disco et a enregistré un fichier vidéo au format MP4, aussi incluse dans le jeu de données SVMIS, à 30 images par secondes (Fig. 6(b)). Le Disco est un drone d’envergure d’1,15 m, à propulsion mono-rotor, alimenté par batterie, pour une masse de 750 g.

La position surélevée de la caméra à l’avant du drone permet de percevoir l’environnement en minimisant les occultations dues au drone et sans trop perturber son centre de gravité. Le Disco a été piloté au-dessus d’un champ proche de la ville d’Amiens pour un vol de 6’41’’ durant lequel il a parcouru 4,2 km, atteint l’altitude de 106 m et une vitesse au sol maximale de 85,9 km/h (Fig. 6(a)).

En absence de vérité terrain fiable (les données de la centrale inertielle embarquée dans la Ricoh Theta se sont révélées trop imprécises pour une étude quantitative), nous avons décidé d’effectuer une évaluation qualitative de la méthode. Pour ce faire, l’image de référence a été choisie à 4’41’’ de la vidéo, instant où le drone est à la fois horizontal et central dans la zone parcourue durant le vol. La Figure 6(c) montre une représentation équirectangulaire de cette image de référence à pleine définition pour en faciliter la visualisation. La Figure 6(d) montre une image courante (à 5’11’’ de la vidéo) et la Figure 6(e) la même image mais

| N | 3 | 4 | 5 |
|-------------------------------|-------|------|-------|
| Erreur moyenne [deg.] | 7,55 | 4,15 | 3,69 |
| Ecart-type de l’erreur [deg.] | 3,8 | 1,77 | 1,72 |
| Temps de calcul moyen [s] | 0,38 | 3,92 | 57,16 |
| Nombre moyen d’itérations | 11,10 | 7,83 | 7,66 |

TABLE 1 – [3 DDL Stäubli] Statistiques de $\|\mathbf{r} - \hat{\mathbf{r}}\|$, temps de calcul et nombre d’itérations moyens du gyroscope sur 94 images.

dont la rotation estimée, pour $N = 3$, par rapport à l’image de référence, a été compensée. Cette compensation est satisfaisante puisque les Figures 6(e) et 6(c) ont un aspect très similaire de ligne d’horizon et de position du soleil et ce, malgré la distance entre l’image courante et celle de référence de 108 m, pour une altitude respective de 58,3 m et 50,1 m. Un niveau de compensation comparable est observé dans 88% des images de la séquence de 6’41’’ dont une vidéo est disponible au lien³ en bas de page. Les mauvais résultats (12% des images) sont attribuables à des minimum locaux dans la fonction de coût MPP en présence de grandes translations. Il est à noter que ces résultats ont été obtenus pour un drone évoluant en environnement naturel non structuré avec une initialisation et un paramétrage fait exactement comme dans la partie 5.2 où un environnement structuré intérieur est considéré.

6 Conclusion et perspectives

Cet article propose un gyroscope visuel direct mettant en oeuvre une nouvelle représentation compacte d’image sphérique : le *Mélange de Potentiels Photométriques* (MPP). Les démonstrations expérimentales faites avec une caméra double-fisheye, montée sur deux plateformes robotiques, indiquent que le gyroscope proposé est précis, facile à régler et robuste aux occultations partielles.

Avec un temps de calcul inférieur à 1 seconde par image pour $N = 3$, notre gyroscope est prometteur mais n’est pas encore intégrable directement dans des applications temps-réel. Cependant, nous sommes confiants qu’une implémentation sur du matériel spécifique (ex. : GPU) dépassera cette limitation.

Enfin, nous poursuivons ces travaux en étudiant des méthodes de réglage adaptatif des paramètres λ_g et N (potentiellement à l’aide de mesures inertielles) ainsi que d’autres fonctions monotones pour étendre la fonction de coût basée MPP afin de proposer une convergence globale.

Remerciements

Ce travail est financé dans le cadre du projet INTERREG VA FMA ADAPT “Assistive Devices for empowering disAbled People through robotic Technologies”, adapt-project.com. Le programme FMA est un programme de Coopération Territoriale Européenne qui vise à financer des projets de coopération ambitieux dans la région frontalière entre la France et l’Angleterre. Le programme est financé par le Fonds Européen de Développement Régionale (FEDER).

Références

- [1] D. Scaramuzza and F. Fraundorfer. Visual Odometry - Part I : The First 30 Years and Fundamentals. *IEEE Rob. Autom. Mag.*, 18(4) :80–92, 2011.
- [2] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional Imaging with Hyperboloidal Projection. In *Proc. IEEE/RSJ*

3. http://mis.u-picardie.fr/~g-caron/pub/data/SVMIS_video_fr.zip

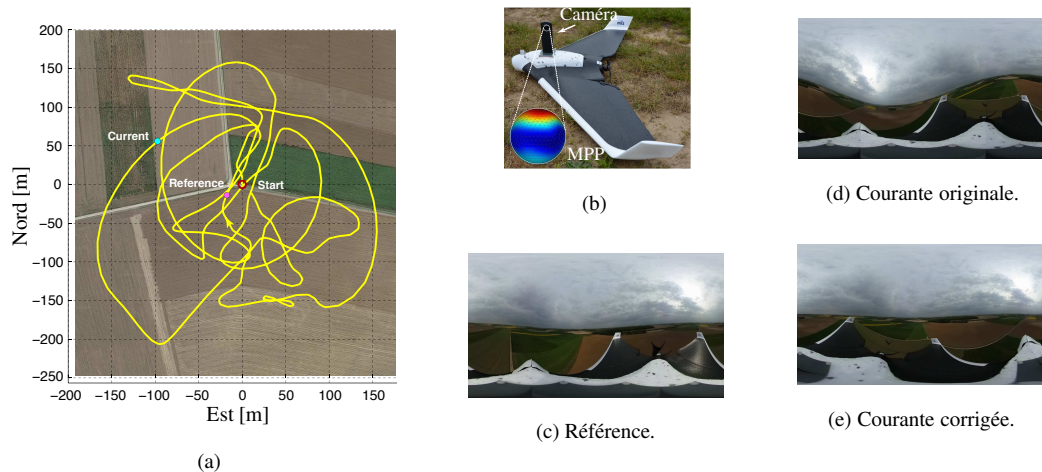


FIGURE 6 – (a) Trajectoire GPS du drone Disco (repère géodésique local); (b) Ricoh Theta embarquée sur le drone; (c) Représentation équirectangulaire de l’image de référence (cf. le point rose en (a)); (d) représentation équirectangulaire de l’image courante sans correction (cf. le point bleu en (a)); (e) Représentation équirectangulaire de l’image courante après correction en utilisant l’attitude estimée par le gyroscope visuel.

- Int. Conf. Intel. Robots Syst.*, volume 2, pages 1029–1034, 1993.
- [3] K. Miyamoto. Fish Eye Lens. *J. Opt. Soc. Am.*, 54(8) :1060–1061, 1964.
- [4] S. Li. Full-View Spherical Image Camera. In *Proc. IEEE Int. Conf. Pattern Recogn.*, volume 4, pages 386–390, 2006.
- [5] J.D. Adarve and R. Mahony. Spherpix : a Data Structure for Spherical Image Processing. *IEEE Robot. Autom. Lett.*, 2(2) :483–490, 2017.
- [6] H.-E. Benseddik, H. Hadj-Abdelkader, B. Cherki, and S. Bouchafa. Camera Rotation Estimation using 3D Mesh Surfaces Representation of Spherical Images. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 2514–2520, 2016.
- [7] G.L. Mariottini, S. Scheggi, F. Morbidi, and D. Prattichizzo. An accurate and robust visual-compass algorithm for robot-mounted omnidirectional cameras. *Robot. Autom. Syst.*, 60(9) :1179–1190, 2012.
- [8] J.-C. Bazin, C. Demonceaux, P. Vasseur, and I. Kweon. Rotation estimation and vanishing point extraction by omnidirectional vision in urban environment. *Int. J. Robot. Res.*, 31(1) :63–81, 2012.
- [9] D. Churchill and A. Vardy. An Orientation Invariant Visual Homing Algorithm. *J. Intell. Robot. Syst.*, 71(1) :3–29, 2013.
- [10] F. Labrosse. The Visual Compass : Performance and Limitations of an Appearance-Based Method. *J. Field Robot.*, 23(10) :913–941, 2006.
- [11] D. Scaramuzza and R. Siegwart. Appearance-Guided Monocular Omnidirectional Visual Odometry for Outdoor Ground Vehicles. *IEEE Trans. Robot.*, 24(5) :1015–1026, 2008.
- [12] A. Kyatkin and G. Chirikjian. Pattern Matching as a Correlation on the Discrete Motion Group. *Comput. Vis. Image Und.*, 74(1) :22–35, 1999.
- [13] A. Makadia, C. Geyer, and K. Daniilidis. Correspondence-free structure from motion. *Int. J. Comput. Vision*, 75(3) :311–327, 2007.
- [14] C. Forster, Z. Zhang, M. Gassner, M. Werlberger, and D. Scaramuzza. SVO : Semidirect Visual Odometry for Monocular and Multicamera Systems. *IEEE Trans. Robot.*, 33(2) :249–265, 2017.
- [15] N. Crombez, G. Caron, and E. Mouaddib. Photometric Gaussian Mixtures based Visual Servoing. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 5486–5491, 2015.
- [16] X. Ying and Z. Hu. Can We Consider Central Catadioptric Cameras and Fisheye Cameras within a Unified Imaging Model? In *Proc. Eur. Conf. Comp. Vis.*, pages 442–455, 2004.
- [17] C. Geyer and K. Daniilidis. A Unifying Theory for Central Panoramic Systems and Practical Implications. In *Proc. Eur. Conf. Comp. Vis.*, pages 445–461, 2000.
- [18] S. Li and Y. Hai. A Full-View Spherical Image Format. In *Proc. IEEE Int. Conf. Pattern Recogn.*, pages 2337–2340, 2010.
- [19] G. Caron, E. Marchand, and E. Mouaddib. Photometric visual servoing for omnidirectional cameras. *Auton. Robot.*, 35(2) :177–193, 2013.
- [20] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley, 2000.
- [21] D.M. Titterton, A.F. Smith, and U.E. Makov. *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons, 1985.
- [22] G. Caron and D. Eynard. Multiple Camera Types Simultaneous Stereo Calibration. In *Proc. IEEE Int. Conf. Robot. Automat.*, pages 2933–2938, 2011.
- [23] P.J. Huber. *Robust Statistics*. John Wiley & Sons, 1981.