



**HAL**  
open science

## Estimation de mouvement d'un drone à partir d'un capteur stéréo hybride

Damien Eynard, Pascal Vasseur, Cédric Demonceaux, Vincent Fremont

### ► To cite this version:

Damien Eynard, Pascal Vasseur, Cédric Demonceaux, Vincent Fremont. Estimation de mouvement d'un drone à partir d'un capteur stéréo hybride. ORASIS - Congrès des jeunes chercheurs en vision par ordinateur, INRIA Grenoble Rhône-Alpes, Jun 2011, Praz-sur-Arly, France. inria-00598255

**HAL Id: inria-00598255**

**<https://inria.hal.science/inria-00598255v1>**

Submitted on 6 Jun 2011

**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.

# Estimation de mouvement d'un drone à partir d'un capteur stéréo hybride.

Damien Eynard<sup>1</sup>

Pascal Vasseur<sup>2</sup>

Cédric Demonceaux<sup>3</sup>

Vincent Fremont<sup>4</sup>

<sup>1</sup> MIS, Université de Picardie Jules Verne, Amiens.

<sup>2</sup> LITIS, Université de Rouen, Rouen.

<sup>3</sup> Le2I, Université de Bourgogne, Le Creusot.

<sup>4</sup> Heudiasyc, Université de Technologie de Compiègne, Compiègne.

<sup>1</sup>damien.eynard@u-picardie.fr, <sup>2</sup>pascal.vasseur@univ-rouen.fr,

<sup>3</sup>cedric.demonceaux@u-bourgogne.fr, <sup>4</sup>vincent.fremont@hds.utc.fr

## Résumé

*Le mouvement et la vitesse sont deux paramètres essentiels à connaître lors de la navigation d'un drone en particulier durant les phases critiques d'atterrissage et de navigation. Dans cet article, nous présentons un système stéréoscopique mixte composé d'une caméra fisheye et d'une caméra perspective pour estimer le mouvement. Le mouvement est découplé en rotation et translation. La vue fisheye évalue cette rotation, l'attitude et la translation globale alors que la vue perspective contribue à déterminer l'échelle de la translation et apporte une information de précision. Contrairement aux systèmes stéréoscopiques classiques basés sur l'appariement de primitives, nous proposons un algorithme qui suit et exploite des points dans chacune des caméras indépendamment. Les points suivis dans chacune des caméras sont classés en deux types puis sont fusionnés : on distingue les points localisés sur le sol, dont la profondeur est connue et les autres points de l'environnement, dont la profondeur est inconnue. Nous pouvons alors estimer de façon robuste la translation et la vitesse en utilisant l'algorithme des deux points suivi d'un filtrage de Kalman. Nous montrons que cette approche est robuste et précise grâce à l'utilisation des vues hybrides. Des résultats de synthèse et expérimentaux sur des séquences réelles d'un drone démontrent la validité de cette approche.*

## Mots Clef

Drone, mouvement, mixte, stéréo.

## Abstract

*Motion and velocity are two of the most important parameters to be known for an Unmanned Aerial Vehicle (UAV) especially during critical maneuvers such as landing or steady flight. In this paper, we present mixed stereoscopic vision system made of a fish-eye camera and a perspective camera for motion estimation. Contrary to classical stereoscopic systems based on feature matching between cameras, we propose an algorithm which tracks and exploits points in each camera independently. The omnidirectional*

*view estimates the orientation of the motion while the perspective view contribute to estimate the scale of the translation and brings accuracy. By fusing points tracked in each camera and knowing the rotation between two consecutive positions, we can estimate robustly the translation and the velocity using the two-points algorithm. Then motion is filtered by Kalman filter to remove bad estimations. By experimental results on UAV, we validate the approach.*

## Keywords

UAV, motion, mixed, stereo.

## 1 Introduction

Les drones et micro drones font l'objet d'une attention particulière cette dernière décennie. Cette notion inclue la capacité à réaliser des manoeuvres complexes telles que le décollage et l'atterrissage. Il est donc indispensable d'estimer précisément et de façon robuste le mouvement et les vitesses. L'estimation autonome du mouvement est possible grâce à la mise en place de différents capteurs tels que le GPS, les radars, les INS. L'inconvénient majeur réside dans l'encombrement (poids, place), la consommation d'énergie de certains capteurs actifs ou encore le manque de précision.

Contrairement à ces systèmes, les capteurs de vision sont passifs, économes en énergie et peuvent déterminer de nombreuses informations de manière précise. La navigation dans un environnement 3D nécessite parfois l'utilisation de systèmes stéréoscopiques. La majorité des travaux existants utilisent les caméras perspectives [25, 21] mais ce type de caméras disposent d'un champ de vue limité, d'une ambiguïté de rotation/translation et de la possibilité de perte de points lors du suivi. Par contre, en combinant des caméras de types différents, il est possible d'associer leurs avantages. Les caméras fisheye présentent un large champ de vue et apportent une information d'orientation tandis que les caméras perspectives possèdent une résolution constante et conservent la notion de voisinage carré sur un modèle planaire.

C'est pourquoi nous proposons dans cet article une méthode basée sur un capteur stéréo hybride composé d'une vue fisheye et d'une vue perspective (fig.1) pour estimer le mouvement. Cette approche basée vision peut être utilisée seule pour estimer l'ensemble des paramètres de navigation tels que l'altitude [12], l'attitude [3, 9, 10, 11], les vitesses, la proximité des obstacles en temps réel ou l'ensemble de ces paramètres [24]. Mais cette approche peut également être utilisée conjointement à d'autres capteurs ou pour enrichir leurs données. Les principales méthodes d'estimation de mouvement sont basées flot optique [27], sur la décomposition des matrices fondamentale, essentielle ou d'homographie [15, 19] ou sur le découplage de l'estimation de la rotation et de la translation [4, 8]. D'autres travaux utilisent des systèmes hybrides pour estimer le mouvement. [2] propose une méthode de SfM basée sur un système de vision hybride composé d'un fisheye et d'une perspective, tous deux modélisés par une sphère. Ce système utilise une version adaptée de SIFT pour apparier les primitives entre les caméras. Le pré-traitement de cette version SIFT requiert un temps de calcul plus important et n'est donc pas adapté à la navigation temps réel.

Dans notre approche, nous supposons l'altitude connue et le plan segmenté par la méthode [12]. Nous proposons d'estimer le mouvement en découplant la rotation et la translation. La rotation peut être estimée par IMU ou par vision. Ensuite, des points sont suivis indépendamment dans chacune des caméras puis sont fusionnés pour déterminer la translation. Enfin un filtre de Kalman est appliqué sur les données brutes afin d'éviter les perturbations dues au bruit et aux erreurs de mesure.

En résumé, ce travail présente plusieurs contributions. Tout d'abord le système stéréoscopique hybride proposé permet d'estimer tous les paramètres de vol nécessaires à sa commande et son autonomie. Ensuite, ce système peut être employé pour d'autres tâches liées à des missions spécifiques comme par exemple la surveillance, la modélisation de scène. La seconde contribution concerne la possibilité de profiter des avantages des deux caméras (champ de vue et précision) tout en évitant une mise en correspondance compliquée du fait des images très différentes. Le mouvement est estimé à partir de deux types de points, les points sur le sol permettent d'estimer l'échelle de la translation et d'apporter la précision, puis les points de l'environnement apportent la robustesse de l'estimation. Nous montrerons sur des résultats de synthèse que l'utilisation d'un capteur hybride améliore également la robustesse au bruit.

La suite du papier s'articule comme suit. La partie II traite du système global et de sa modélisation. Dans la partie III, nous proposons d'estimer et de filtrer le mouvement des vues mixtes en utilisant et comparant différentes méthodes. Enfin, nous présenterons en partie IV les résultats de synthèse puis réels avec drone et une évaluation quantitative de la translation. Le travail proposé dans ce papier s'inscrit dans le projet ALTO. L'objectif global de ce travail est d'estimer les paramètres d'un drone tels que l'attitude

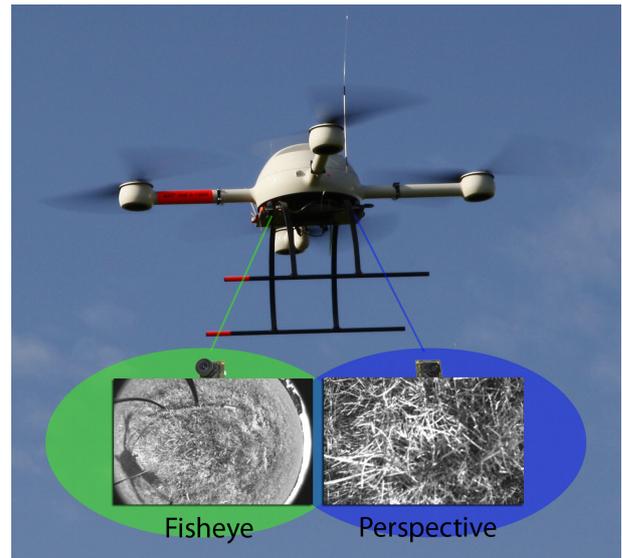


FIGURE 1 – Système hybride sur un drone quadri-rotor

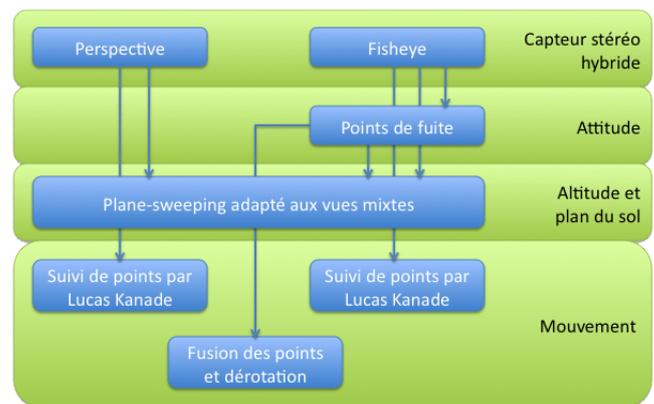


FIGURE 2 – Estimation des paramètres d'un drone à partir d'un capteur hybride

[11], l'altitude, la segmentation du sol [12] et le mouvement (voir fig.2).

## 2 Principe général du capteur hybride

### 2.1 Aperçu global

Nous proposons un système stéréo mixte composé d'une vue perspective et omnidirectionnelle afin de combiner les avantages de chacun de ces capteurs. La vue fisheye fournit l'information d'attitude et le mouvement global dans l'environnement alors que la vue perspective apporte la précision. Le problème principal, dû aux distorsions, consiste à trouver la relation entre la vue fisheye et perspective. Différentes approches sont possibles :

- Tout d'abord, connaissant les paramètres intrinsèques de la caméra fisheye, il est possible de transformer l'image pour obtenir l'équivalent perspectif. On peut ensuite ap-

plier une méthode de mise en correspondance. Cependant, cette approche nécessite différents traitements tels que l'appariement et l'interpolation. L'implémentation temps réel est alors difficile.

- D'autres travaux récents proposent d'utiliser la sphère unitaire pour représenter les images omnidirectionnelles. On peut ensuite appliquer les algorithmes de mise en correspondance. Cependant, comme précédemment, cette solution ne peut être implémentée en temps réel.
- Enfin, nous proposons une approche de fusion des points suivis sur chacune des caméras. La connaissance a priori de la segmentation du sol, de l'altitude [12] et de l'orientation nous permet d'estimer la translation à l'échelle millimétrique sans mise en correspondance entre les vues hybrides.

On suppose que les caméras mixtes regardent le sol segmenté noté  $\mathbf{G}$ . Soit  $\mathbf{R}_c$  et  $\mathbf{t}_c$ , la transformation rigide entre les caméras est connue par calibration. Nous allons démontrer qu'en utilisant à la fois les points de l'environnement majoritairement présents dans la vue fisheye et les points du sol présents principalement dans la vue perspective nous pouvons estimer la translation de façon robuste et à l'échelle métrique.

## 2.2 Modèles de caméras et calibration

Bien que les lentilles fisheye ne peuvent être classées comme capteurs à point de vue unique [1], [28] a démontré que nous pouvons tout de même utiliser un modèle sphérique unitaire pour la caméra fisheye. Pour calibrer notre capteur stéréo hybride, nous utilisons le travail de [6] basé sur l'asservissement visuel virtuel et permettant la calibration de capteurs hybrides de  $n$  caméras à  $n$  modèles.

## 3 Estimation de mouvement

Nous proposons dans cette section de calculer la translation  $\mathbf{t}$  à partir des points suivis indépendamment dans chacune des vues mixtes. Ces vues observent la même translation une fois la rotation découplée. Dans chacune des vues nous distinguons deux types de points :

- les points localisés sur le sol dont l'altitude et la segmentation sont supposées connus grâce à nos travaux antérieurs [12]
- les autres points de l'environnement dont la profondeur est inconnue

Les premiers apportent l'échelle et la précision tandis que les seconds améliorent la robustesse. Ces deux types de points issus de chaque vue sont ensuite fusionnés pour estimer la translation.

### 3.1 Mouvement du banc stéréo

Dans chacune des images, le suivi est réalisé par la méthode proposée par [5]. Chaque point  $\mathbf{x}_t$  exprimé dans le repère image est lié par une rotation  ${}^t\mathbf{R}_w$  au même point  $\mathbf{X}_w$  exprimé dans le repère monde :

$$\mathbf{x}_t = {}^t\mathbf{R}_w\mathbf{X}_w \iff \mathbf{X}_w = {}^t\mathbf{R}_w^{-1}\mathbf{x}_t \quad (1)$$

Puis, pour un couple de points  $(\mathbf{x}_t; \mathbf{x}_{t+1})$  correspondant à un mouvement rigide entre les instants  $t$  et  $t+1$ , on peut exprimer la rotation entre les deux repères images par :

$$\mathbf{x}_{t+1} = {}^{t+1}\mathbf{R}_w^t \mathbf{x}_t \quad (2)$$

on obtient la rotation du mouvement :

$$\mathbf{R}_{t+1} = {}^{t+1}\mathbf{R}_w^t \mathbf{R}_w^{-1} \quad (3)$$

Ces points sont reliés par un mouvement composé d'une rotation  $\mathbf{R}_{t+1}$  définie précédemment et d'une translation  $\mathbf{t}_{t+1}$  (voir fig.3). Notons  $d$  la profondeur d'un point  $\mathbf{x}_t$  à l'instant  $t$  (fig.3). On a :

$$\mathbf{x}_{t+1} \sim \mathbf{R}_{t+1}d\mathbf{x}_t + \mathbf{t}_{t+1} \quad (4)$$

### 3.2 Mouvement à partir des points du sol

Connaissant l'altitude et la segmentation du plan du sol, nous savons si les points suivis appartiennent au sol ainsi que leur profondeur (voir fig.3). Ainsi, le mouvement est estimé à l'échelle métrique et est défini comme suit :

$$(\mathbf{x}_{t+1} \times \mathbf{t}_{t+1}) = -d(\mathbf{x}_{t+1} \times \mathbf{R}_{t+1}\mathbf{x}_t) \quad (5)$$

On peut écrire l'eq.5 sous forme matricielle.  $\mathbf{A}\mathbf{t}_{t+1} = \mathbf{B}$  où  $\mathbf{x}_{t+1} = (x_{t+1}, y_{t+1}, z_{t+1})^T$ ,  $\mathbf{A}$  dans l'eq.6 et  $\mathbf{B}$  dans l'eq.7 :

$$\mathbf{A} = \begin{bmatrix} 0 & -z_{t+1} & y_{t+1} \\ z_{t+1} & 0 & -x_{t+1} \\ -y_{t+1} & x_{t+1} & 0 \end{bmatrix} \quad (6)$$

$$\mathbf{B} = -d(\mathbf{x}_{t+1} \times \mathbf{R}_{t+1}\mathbf{x}_t)^T \quad (7)$$

Dans le cas d'estimation de translation avec plus de deux points, on peut facilement obtenir  $\mathbf{t}_{t+1}$  par moindres carrés. Nous exprimerons le mouvement dans la vue perspective, dans la vue fisheye puis en mixte.

Tout d'abord, nous obtenons à partir des (eq.6 et 7) et des points localisés sur le sol d'altitude connue et observés dans la vue perspective les matrices  $\mathbf{A}_p$  et  $\mathbf{B}_p$ . Ensuite, nous obtenons à partir des (eq.6 et 7) et des points localisés sur le sol d'altitude connue et observés dans la vue fisheye les matrices  $\mathbf{A}_s$  et  $\mathbf{B}_s$ . Enfin dans le cas mixte, nous obtenons  $\mathbf{A}_m$  (eq.8) en combinant  $\mathbf{A}_s$  et  $\mathbf{A}_p$  puis  $\mathbf{B}_m$  (eq.9) en combinant  $\mathbf{B}_s$  et  $\mathbf{B}_p$ .  $\mathbf{A}_m$  et  $\mathbf{B}_m$  évaluent  $\mathbf{t}$  à partir de points situés sur le sol et observés dans chaque caméra.

Pour  $n$  points perspectifs et  $m$  points sphériques, la taille de  $\mathbf{A}_p$  est  $3n \times 3$ , de  $\mathbf{A}_s$  est  $3m \times 3$ , de  $\mathbf{B}_p$  est  $3n \times 1$ , de  $\mathbf{B}_s$  est  $3m \times 1$ , de  $\mathbf{A}_m$  est  $3(m+n) \times 3$  et de  $\mathbf{B}_m$  est  $3(m+n) \times 1$ .

$$\mathbf{A}_m = \begin{bmatrix} \mathbf{A}_s \\ \mathbf{A}_p \end{bmatrix} \quad (8)$$

$$\mathbf{B}_m = \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_p \end{bmatrix} \quad (9)$$

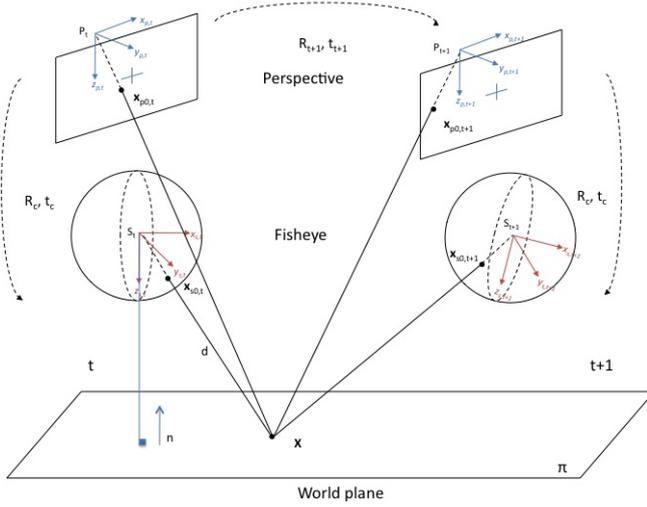


FIGURE 3 – Points 3D et mouvement.

### 3.3 Mouvement à partir des points de l'environnement

Dans le cas de points situés dans l'environnement, sans aucune connaissance de leur profondeur, la translation  $\mathbf{t}_{t+1}$  pour deux points [4] est représentée par l'eq.10 (voir fig.3). Comme énoncé précédemment, les points perspectifs et les points sphériques sont concaténés pour estimer la translation  $\mathbf{t}_{t+1}$ .

$$\mathbf{t}_{t+1} \cdot (\mathbf{x}_{t+1} \times \mathbf{R}_{t+1}\mathbf{x}_t) = 0 \quad (10)$$

### 3.4 Fusion des points 3D issus des vues mixtes

A partir des points du sol, la translation est estimée à l'échelle millimétrique. L'inconvénient de cette méthode est la forte sensibilité au bruit dans le cas d'un déplacement 3D. Cependant, à partir des points de l'environnement, le mouvement est estimé à l'échelle prêt mais l'estimation est plus robuste grâce à l'utilisation de la vue fisheye. Une des contributions de ce papier est la combinaison de ces deux méthodes afin d'obtenir une estimation robuste et à l'échelle millimétrique. La première méthode est définie par l'équation (12) et la seconde par l'eq. (13).  $\mathbf{C}_m$  est définie par l'eq. (11). En concaténant la matrice  $\mathbf{A}_m$  et  $\mathbf{C}_m$  en  $\mathbf{A}_f$ ,  $\mathbf{B}_m$  et 0 en  $\mathbf{B}_f$  nous obtenons l'eq. (14) résolue par les moindres carrés.

$$\mathbf{C}_m = \begin{bmatrix} \mathbf{R}_{t+1}\mathbf{x}_{s0,t} \times \mathbf{x}_{s0,t+1} \\ \vdots \\ \mathbf{R}_{t+1}\mathbf{x}_{sm,t} \times \mathbf{x}_{sm,t+1} \\ \mathbf{R}_{t+1}\mathbf{x}_{p0,t} \times \mathbf{x}_{p0,t+1} \\ \vdots \\ \mathbf{R}_{t+1}\mathbf{x}_{pn,t} \times \mathbf{x}_{pn,t+1} \end{bmatrix} \quad (11)$$

$$\mathbf{A}_m\mathbf{t}_{t+1} = \mathbf{B}_m \quad (12)$$

$$\mathbf{C}_m\mathbf{t}_{t+1} = 0 \quad (13)$$

$$\begin{bmatrix} \mathbf{A}_m \\ \mathbf{C}_m \end{bmatrix} \mathbf{t}_{t+1} = \begin{bmatrix} \mathbf{B}_m \\ 0 \end{bmatrix} \quad (14)$$

## 3.5 Filtre de Kalman

Une fois le mouvement estimé à l'échelle millimétrique, nous observons des discontinuités et des variations brutales lors de perturbations dans l'acquisitions des images. Nous proposons un filtre de Kalman linéaire [16] dans le but de réduire les mauvaises estimations et de lisser la trajectoire. L'état considéré correspond au vecteur de translation du mouvement i.e.  $\mathbf{x}_k = \mathbf{t}_{t+1}$ , modélisé par un système linéaire Gaussien donné par l'eq. 15 :

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{E}\mathbf{x}_k + \mathbf{v}_k \\ \mathbf{y}_k &= \mathbf{O}\mathbf{x}_k + \mathbf{w}_k \end{aligned} \quad (15)$$

où  $\mathbf{E} = \mathbf{I}_3$  est la matrice de transition d'état et  $\mathbf{O} = \mathbf{I}_3$  est le modèle d'observation, avec comme hypothèse une vitesse constante durant un intervalle échantillonné. Les vecteurs  $\mathbf{v}_k$  et  $\mathbf{w}_k$  correspondent respectivement à l'erreur du modèle et à l'erreur d'observation. Ces bruits sont supposés additifs, blancs de Gaussienne de moyenne nulle et les matrices de covariances sont pré-définies. Dans le but de garder la dynamique des mesures, une plus grande incertitude est attribuée aux mesures i.e.  $10^6$  qu'au modèle  $10^2$ . A partir de ces hypothèses, le filtre de Kalman consiste donc à prédire le vecteur translation  $\mathbf{t}_{t+1}$  et à obtenir une estimation lissée en utilisant une étape de mise à jour lorsqu'une observation est disponible. Ainsi, nous obtenons un vecteur translation estimé à partir des observations précédentes jusqu'à l'instant présent.

## 4 Résultats expérimentaux

Dans cette section nous comparons les méthodes présentées dans la section III. Une première partie traite des résultats théoriques pour valider notre méthode, une seconde traite des résultats expérimentaux sur drone. Dans chaque vue, les points sont suivis par la méthode [5] implémentée dans OpenCV.

### 4.1 Résultats de simulation

Pour chaque scénario, la répartition des points s'effectue selon le tableau 4.1. L'erreur introduite est un angle aléatoire appliqué sur les lignes de vue quel que soit le modèle de projection utilisé. L'erreur d'angle est calculée entre le vecteur réel de translation et le vecteur estimé. Pour l'erreur métrique d'estimation de translation, nous calculons la norme entre le vecteur translation réel et celui estimé. Notons que l'erreur d'entrée varie de  $0^\circ$  à  $1^\circ$  par pas de  $0.02^\circ$  et l'erreur de sortie est la moyenne de 1000 estimations.

**(Méthode 1) Mouvement à partir des points du plan - fig.4 et fig.5.** Dans le cas d'une estimation de mouvement à partir des points du sol, l'estimation issue de la vue sphérique est moins précise que sur la vue perspective tant pour

Vues	Nb. points sol	Nb. points env.
Hybride - fig. 4	60	60
Sphere - fig. 4,5	60	0
Perspective - fig. 4,5	60	0
Hybride - fig. 4,5	60	0
Sphere - fig. 4,5	30	30
Perspective - fig. 4,5	30	30
Hybride - fig. 4,5	60	60

TABLE 1 – Répartition du nombre de points en fonction du type et des vues

l'erreur angulaire que pour l'erreur de translation. Dans le cas hybride, le nombre de points est doublé grâce aux deux vues, l'estimation est alors plus robuste au bruit.

**(Méthode 2) Mouvement à partir des points de l'environnement - fig.4.** En estimant la translation à partir des points de l'environnement, la translation est obtenue à l'échelle. Nous quantifions l'erreur d'estimation à partir de l'erreur angulaire. La méthode d'estimation de mouvement à partir des vues hybrides utilisant les points de l'environnement n'est représentée que dans la fig.4 (pourpre). Cette méthode est globalement plus robuste à l'erreur que la méthode précédente.

**(Méthode 3) Fusion des points de l'environnement des vues mixtes - fig.4 et fig.5.** Dans le dernier cas, la fusion des deux classes de points permet d'améliorer la robustesse. Nous pouvons estimer la translation à l'échelle métrique avec une meilleure robustesse au bruit qu'en utilisant les points sur le sol. La fusion des points du sol et de l'environnement que ce soit en hybride ou dans les vues séparées permet une meilleure robustesse que les méthodes précédentes (voir fig.4 et fig.5). De plus, l'hybride apporte plus de points et augmente ainsi la robustesse.

Dans le cas réel, la vue fisheye apportera une majorité de points localisés dans l'environnement alors que la vue perspective fournira plutôt des points localisés sur le sol.

## 4.2 Résultats réels

Nous proposons d'estimer le mouvement par une méthode linéaire : les moindres carrés. Cette méthode présente l'avantage d'être robuste au bruit Gaussien mais sensible aux mesures aberrantes [22] que nous choisissons de rejeter par RANSAC [15]. Notre algorithme a ensuite été testé sur un drone quadri-rotor avec deux caméras uEye dont les images sont traitées a posteriori. L'approche vision est ici couplée à une centrale inertielle Xsens pour déterminer l'attitude et les rotations du mouvement entre deux instants alors que le plane-sweeping estime l'altitude. Dans chaque vue, le nombre de points suivis varie entre 50 et 200. Ensuite, nous supposons que la vue perspective pointe vers le sol, dont la segmentation est connue par plane-sweeping [12]. La fig 6 présente la trajectoire finale du mouvement estimé. Cette figure montre en rouge les données brutes,

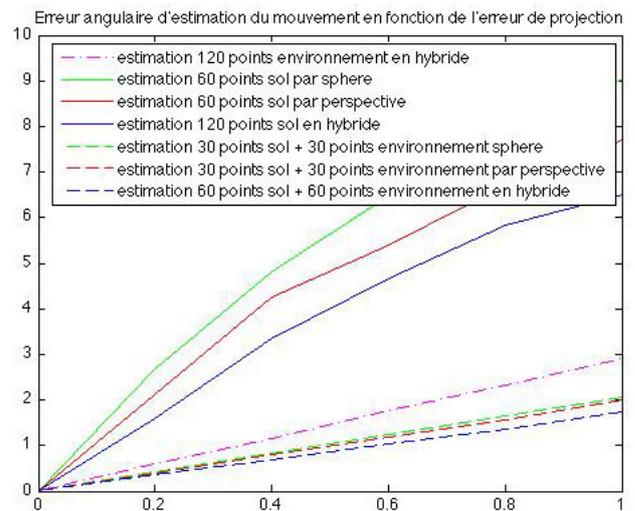


FIGURE 4 – Erreur angulaire entre la translation réelle et estimée en degrés. *En traits pleins, le mouvement est estimé dans la vue fisheye, perspective et mixte. En pourpre, les points de l'environnement sont utilisés pour estimer le mouvement. En pointillés, les points de l'environnement et les points sur le sol sont fusionnés pour estimer le mouvement dans la vue fisheye, perspective et les vues mixtes.*

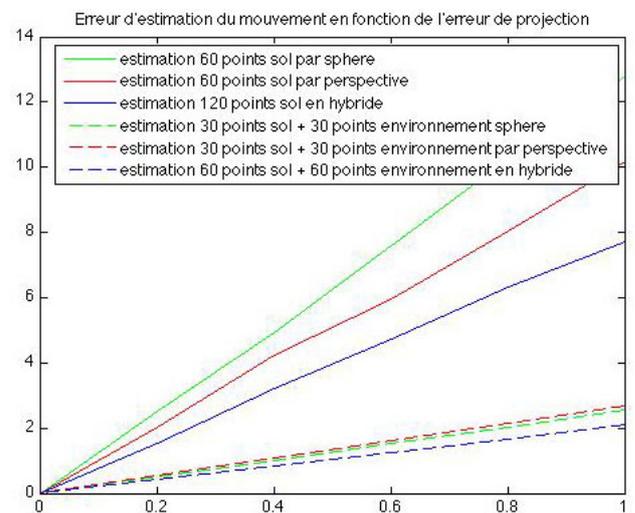


FIGURE 5 – Erreur de distance entre la translation réelle et estimée. *En traits pleins, le mouvement est estimé dans la vue fisheye, perspective et mixte. En pointillés, les points de l'environnement et les points sur le sol sont fusionnés pour estimer le mouvement dans la vue fisheye, perspective et les vues mixtes.*

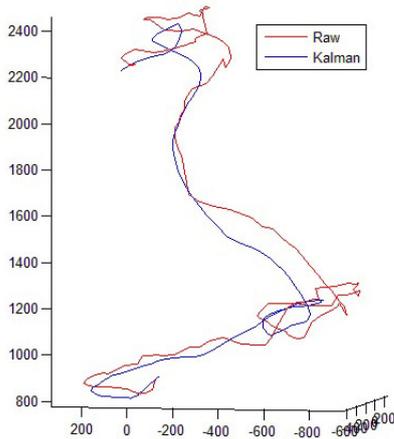


FIGURE 6 – Trajectoire 3D d'un drone.

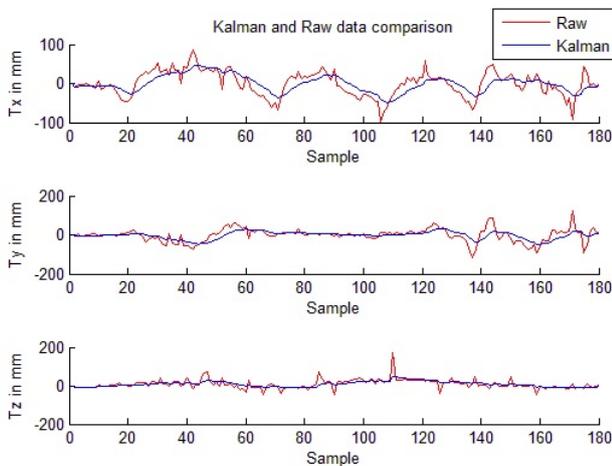


FIGURE 7 – Données brutes et données filtrées par Kalman.

sensibles au bruit et aux variations brutales alors qu'en bleu, les données sont filtrées et lissées par le filtre de Kalman.

La précision a été testée sur un robot mobile effectuant une translation le long de l'axe x. La distance réelle était de 2500mm, la distance estimée à partir des données brute est de 2352mm.

## 5 Conclusions et perspectives

Nous avons présenté dans ce papier un système stéréo hybride où des points sont suivis dans chacune des caméras. L'aspect collaboratif du capteur permet d'estimer l'attitude dans la vue fisheye puis d'estimer l'homographie définie entre les vues hybrides permettant d'estimer à la fois l'altitude et de segmenter le plan du sol par planesweeping. Ces informations sont ensuite utilisées pour estimer le mouvement. Dans la méthode proposée, nous supposons la rotation connue et estimons le mouvement à partir du suivi de deux types de point. Les points situés sur le sol apportent

l'information métrique de la translation et la précision. Les points situés dans l'environnement permettent une estimation robuste. La combinaison de capteurs hybrides ainsi que la fusion des deux types de points permettent à la fois d'améliorer la robustesse et d'obtenir l'échelle métrique. Les données sont ensuite filtrées par Kalman pour lisser et éliminer les mauvaises estimations.

La perspective de ce travail est une implémentation en vue de traitement temps réel et embarqué dans un drone.

## 6 Remerciements

Ce travail est soutenu et financé par le FEDER Européen et le projet ALTO de la Région de Picardie.

## Références

- [1] S. Baker, and S. K. Nayar, "A Theory of Single-Viewpoint Catadioptric Image Formation," *International Journal of Computer Vision*, 1999.
- [2] Y. Bastanlar, A. Temizel, Y. Yardimci, "Effective Structure-from-Motion for Hybrid Camera Systems", *International Conference on Pattern Recognition*, 2010.
- [3] J-C Bazin, I. Kweon, C. Demonceaux, P. Vasseur, "UAV Attitude Estimation by Vanishing Points in Catadioptric Images", *In Proceedings of IEEE International Conference on Robotics and Automation*, 2008.
- [4] J-C Bazin, I. Kweon, C. Demonceaux, P. Vasseur, "Motion Estimation by Decoupling Rotation and Translation in Catadioptric Vision", *CVIU*, 2009.
- [5] J-Y. Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm", *In Intel Corporation Microprocessor Research Labs*, 2000.
- [6] G. Caron, D. Eynard, "Multiple Camera Types Simultaneous Stereo Calibration", *In Proceedings of IEEE International Conference on Robotics and Automation*, 2011.
- [7] R. T. Collins, "A space-sweep approach to true multi-image matching," *In Proceedings of IEEE Computer Vision and Pattern Recognition*, 1996.
- [8] K. Daniilidis, H.H. Nagel, "The coupling of rotation and translation in motion estimation of planar surfaces", *In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1993.
- [9] C. Demonceaux, P. Vasseur, C. Pégard, "Robust Attitude Estimation with Catadioptric Vision," *In IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.
- [10] C. Demonceaux and P. Vasseur and C. Pégard, "Omnidirectional vision on UAV for attitude computation," *In Proceedings of IEEE International Conference on Robotics and Automation*, 2006.

- [11] C. Demonceaux and P. Vasseur and C. Pégard, "UAV Attitude Computation by Omnidirectional Vision in Urban Environment," In Proceedings of IEEE International Conference on Robotics and Automation, 2007.
- [12] D. Eynard, P. Vasseur, C. Demonceaux, V. Fremont, "UAV Altitude Estimation by Mixed Stereoscopic Vision", In Proceedings of IEEE International Conference on Intelligent Robots and Systems, 2010.
- [13] C. Geyer, K. Daniilidis, "Structure and motion from uncalibrated catadioptric views", Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
- [14] S. Hrabar and G. Sukhatme, "Omnidirectional vision for an autonomous helicopter," In Proceedings of IEEE International Conference on Robotics and Automation, 2004.
- [15] R. Hartley, A. Zisserman, "Multiple view geometry in computer vision", Cambridge University, 2nd edition, 2003.
- [16] Rudolf Emil Kalman, "A new approach to linear filtering and prediction problems", Transactions of the ASME - Journal of Basic Engineering, (82) :35-45, 1960.
- [17] J.W. Lee, S. You, U. Neumann, "Large motion estimation for omnidirectional vision", In Proc IEEE Workshop on Omnidirectional Vision, 2000.
- [18] M. Lhuillier, "Automatic scene structure and camera motion using a catadioptric system", Computer Vision and Image Understanding, 2008.
- [19] Y. Ma, S. Soatto, J. Kosecka, S.S. Sastry, "An Invitation to 3D Vision", Springer, 2004.
- [20] A. Makadia, C. Geyer, K. Daniilidis, "Correspondence-free Structure from Motion, accepted for publication in the International Journal of Computer Vision", 2007
- [21] K. Kanatani, "Detection the motion of a planar surface by line & surface integrals", In Computer Vision, Graphics, and Image Processing, 1985.
- [22] E. Malis, E. Marchand, "Méthodes robustes d'estimation pour la vision robotique", Journées nationales de la recherche en robotique, 2005.
- [23] C. Mei and P. Rives, "Single View Point Omnidirectional Camera Calibration from Planar Grids," In Proceedings of IEEE International Conference on Robotics and Automation, 2007
- [24] I. Mondragon , M. Olivares, P. Campoy, C. Martinez, L. Mejias (2010) Unmanned Aerial Vehicles UAVs attitude, height, motion estimation and control using visual systems. Autonomous Robots, 2010.
- [25] O. Shakernia, Y. Ma, T. J. Koo, S. Sastry, "Landing an Unmanned Air Vehicle : Vision based Motion Estimation and Nonlinear Control", Asian Journal of Control, 1999.
- [26] J. Shi and C. Tomasi, "Good features to track", IEEE Computer Society Conference on In Computer Vision and Pattern Recognition, 1994.
- [27] S. P. N. Singh and K. J. Waldron, "Motion Estimation by Optical Flow and Inertial Measurements for Dynamic Legged Locomotion", IEEE Conference on Robotics and Automation, 2005.
- [28] X. Ying and Z. Hu, "Can We Consider Central Catadioptric Cameras and Fisheye Cameras within a Unified Imaging Model," In European Conference on Computer Vision, 2004.
- [29] <http://www.youtube.com/watch?v=ubXzf0eLud4>, 2010