



HAL
open science

Une approche performante de suivi visuel pour les caméras catadioptriques

François Rameau, Désiré Sidibé, Cédric Demonceaux, David Fofi

► **To cite this version:**

François Rameau, Désiré Sidibé, Cédric Demonceaux, David Fofi. Une approche performante de suivi visuel pour les caméras catadioptriques. RFA 2012 (Reconnaissance des Formes et Intelligence Artificielle), Jan 2012, Lyon, France. pp.978-2-9539515-2-3. hal-00656553

HAL Id: hal-00656553

<https://hal.science/hal-00656553>

Submitted on 17 Jan 2012

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.

Une approche performante de suivi visuel pour les caméras catadioptriques

François Rameau

Désiré Sidibé

Cédric Demonceaux

David Fofi

Laboratoire Le2i UMR 5158 CNRS, Université de Bourgogne

12 rue de la fonderie, 71200 Le Creusot France

Résumé

Dans cet article, nous proposons une méthode performante permettant d'appliquer des algorithmes de suivi visuel à des images catadioptriques. Cette méthode est basée sur une représentation sphérique de l'image qui permet de prendre en compte les distorsions et la résolution non-uniforme des images catadioptriques. Les résultats expérimentaux proposés démontrent que les méthodes probabilistes et déterministes peuvent être adaptées de manière à suivre un objet avec précision dans une séquence d'images catadioptriques.

Mots Clef

Suivi visuel, caméra catadioptrique, Mean-Shift, filtre particulaire

Abstract

An effective technique for applying visual tracking algorithms to catadioptric image sequences is presented. The method is based on a spherical image representation which allows to take into account the distortions and non-linear resolution of catadioptric images. Experimental results show that both deterministic and probabilistic tracking methods can effectively be adapted in order to robustly track an object with a catadioptric camera.

Keywords

Visual tracking, catadioptric camera, Mean-Shift, particule filter

1 Introduction

Le suivi visuel est un sujet prédominant dans le domaine de la vision par ordinateur. Son utilisation est souvent nécessaire pour des applications telles que la vidéo surveillance, la robotique mobile, la réalité augmentée ou encore la compression vidéo [11]. La littérature comporte de nombreuses méthodes de suivi que nous pouvons dissocier en deux principaux groupes : les méthodes déterministes et les méthodes stochastiques. Les méthodes de suivi appartenant au premier groupe comme par exemple Kanade-Lucas-Tomasi (KLT) ou la méthode *mean-shift*, cherchent de manière itérative à maximiser une mesure de similarité locale entre un modèle de la cible et une région dans l'image courante. Les méthodes stochastiques utilisent quant à elles

une représentation d'état de l'objet en mouvement afin de modéliser sa dynamique. Parmi ces techniques nous pouvons notamment citer le filtre de Kalman et les filtres particuliers. Ces différentes méthodes sont couramment utilisées dans des séquences d'images perspectives, cependant nous ne pouvons pas directement les appliquer à des vidéos acquises à l'aide d'une caméra catadioptrique.

Un système de vision catadioptrique offre une vision panoramique de la scène jusqu'à 360°, rendue possible par l'association d'une caméra perspective ou orthographique et d'un miroir convexe. Les capteurs catadioptriques offrent un avantage certain dans beaucoup de domaines (comme la reconstruction 3D, la navigation de robot et la vidéo surveillance) en réduisant considérablement le nombre de caméras nécessaires à la réalisation de ces tâches. Cependant, comme nous l'avons déjà évoqué précédemment, les méthodes classiques de suivi visuel ne sont pas applicables de manière performante sur les images issues de ce type de capteur. Cette incompatibilité est principalement liée à la distorsion géométrique induite par l'utilisation d'un miroir. Le miroir implique également une forte sensibilité à l'illumination de la scène ainsi qu'une non-uniformité de la résolution de l'image qui rendent le suivi plus difficile. Ici, nous proposons une adaptation d'algorithmes de suivi visuel (la méthode *mean-shift* et le filtrage particulaire) à la géométrie des capteurs catadioptriques. Cet article est articulé en cinq parties principales : tout d'abord, une brève bibliographie à propos du suivi d'objet avec des capteurs catadioptriques est proposée. Dans la deuxième section, une méthode permettant une adaptation du voisinage à la géométrie du capteur est présentée. La troisième section rendra compte de la représentation d'une cible en utilisant un histogramme couleur. Ensuite, nous proposons dans la section 4 une adaptation de deux méthodes, l'algorithme *mean-shift* et le filtrage particulaire, pour les images catadioptriques. Enfin, la dernière section est dédiée aux résultats et aux conclusions.

1.1 Travaux antérieurs

Le suivi d'objet avec des capteurs catadioptriques ouvre la voie à de nombreuses applications, principalement dans le domaine de la surveillance et de la navigation robotique. Néanmoins, l'attention portée jusqu'à présent à ce sujet reste relativement limitée et très peu d'articles abordent

cette thématique.

Parmi les différentes approches permettant le suivi d'objet avec des capteurs catadioptriques, celle consistant à "rectifier" l'image est certainement la plus répandue. Cette méthode nécessite l'existence d'un Point de Vue Unique (PVU) afin d'obtenir une image perspective à partir de celle déformée. Le procédé permettant cette manipulation de l'image est décrit dans [2]. Suite à cette étape, la distorsion de l'image est réduite ce qui permet l'utilisation d'une méthode de suivi traditionnelle. Dans [18], un filtre particulière est ainsi utilisé sur une séquence d'image désanamorphosée (rectification de la déformation d'une image) tandis que dans [1] l'algorithme de suivi KLT est utilisé.

Le suivi de cible basé sur des connaissances a priori est également une méthode efficace souvent utilisée pour la navigation robotique, l'asservissement visuel et le suivi 3D. Elle nécessite des connaissances sur la géométrie de l'objet à suivre, tel que les plans ou les droites qui le composent. Le suivi de plan est essentiellement basé sur l'homographie. Par exemple, Mei *et al.* [13] utilise cette technique pour le suivi de modèle planaire dans des images catadioptriques. Cette approche nécessite cependant des conditions optimales pour être pleinement efficace, un déplacement court entre chaque image de la séquence est par exemple requis. Le suivi de droites fait intervenir une autre stratégie. Tout d'abord une détection de droites adaptée aux particularités des capteurs catadioptriques est appliquée sur l'image [3, 14], puis le suivi est effectué à l'aide d'un algorithme de suivi classique de droites [12].

La soustraction d'arrière plan est également une solution simple et efficace lorsque la caméra catadioptrique est statique car il n'est pas nécessaire de tenir compte de la distorsion de l'image. Boulton *et al.* [6] donnent un bon aperçu de l'utilisation de cette méthode pour la vidéo surveillance. Une autre approche consiste à adapter les méthodes conventionnelles de suivi sur la sphère unité. Cette adaptation permet une amélioration des résultats de suivi car le voisinage sera adapté aux particularités géométriques et à la non-uniformité de la résolution du capteur. Le concept principal est développé dans [8] avec l'utilisation de distance géodésique calculée sur la sphère. Cependant, seuls [5] et [16] utilisent cette représentation pour le suivi avec un filtre particulière. Nous présentons dans cet article un travail basé sur cette méthode, avec l'adaptation d'un filtre particulière et d'un suivi par *mean-shift*. Nous proposons également différentes améliorations, comme l'adaptation d'un noyau et l'utilisation d'un histogramme multiparties permettant une meilleure localisation spatiale ainsi qu'une plus grande robustesse aux variations d'échelle et aux fonds "encombrés" (forte similarité entre la cible et l'arrière plan). De plus, nous présentons une évaluation quantitative des algorithmes basée sur plusieurs type de mesures. De cette manière nous sommes plus en mesure d'apprécier les avantages offerts par notre approche.

2 L'adaptation du voisinage pour les images catadioptriques

Dans une image perspective, une région d'intérêt est généralement définie par deux paramètres (largeur et hauteur) dans laquelle le voisinage est uniforme. Nous obtenons alors un rectangle centré sur le pixel d'intérêt. Bien entendu, cet échantillonnage classique n'est pas adapté aux images catadioptriques car il considère la résolution uniforme sur l'ensemble de l'image et ne tient pas compte de la distorsion due à l'utilisation du miroir.

Nous pouvons modéliser les images acquises à partir d'une caméra respectant le PVU par un équivalent sphérique. Ce modèle de projection unifié a été introduit dans [9] et [4]. La projection sur la sphère permet de créer un voisinage adapté à la distorsion et à la non-uniformité de la résolution de l'image. Nous définissons ce voisinage par une distance géodésique calculée sur la sphère.

Les coordonnées sphériques sont définies de la manière suivante :

$$X = (\cos(\phi) \sin(\theta), \sin(\phi) \cos(\theta), \cos(\phi)), \quad (1)$$

où X est un point sur la sphère S^2 , ϕ est la latitude (comprise entre 0 et π) et θ la longitude (de 0 à 2π). Par conséquent, la localisation d'un point peut être définie à l'aide de 2 paramètres (θ, ϕ) comme nous pouvons le constater sur la figure 1. Nous déterminons le voisinage à partir d'un point

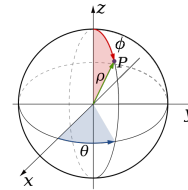


FIGURE 1 – système de coordonnées sphériques

central sur la sphère $X_{sph}(\theta, \phi)$ autour duquel nous définissons une plage de variation de θ et ϕ , respectivement $\delta\theta$ et $\delta\phi$:

$$N_S(X_{sph}) = \{X_{sph} = (\theta', \phi') \in S^2 \mid |\theta' - \theta| \leq \delta\theta \text{ et } |\phi' - \phi| \leq \delta\phi\}$$

De cette manière nous obtenons les coordonnées sur la sphère que nous reprojets ensuite sur le plan image afin d'obtenir les coordonnées des pixels concernés.

Comme nous pouvons le voir sur la figure 2, des fenêtres de dimension similaire sur la sphère auront des dimensions variables en fonction de leur position spatiale sur le plan image.

3 Représentation par des histogrammes couleur

Les méthodes de suivi visuel développées dans cet article sont basées sur les caractéristiques couleur de l'objet à

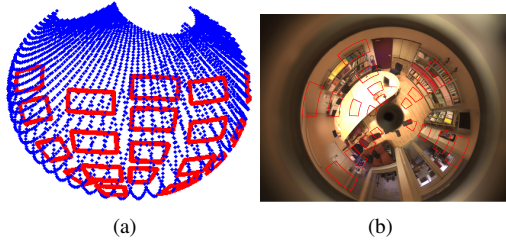


FIGURE 2 – Voisinage avec des valeurs fixes $\delta\theta=\pm 0.2$ et $\delta\phi=\pm 0.1$ (a) voisinage sur la sphère (b) voisinage projeté sur le plan image

suivre. Les histogrammes sont une méthode rapide et efficace pour représenter l'apparence d'un objet.

3.1 Calcul de la distribution colorimétrique

Nous calculons l'histogramme couleur représentatif de l'apparence d'un objet de la manière suivante :

$$q_u = C \sum_{i=1}^n k(\|x_i\|^2) \delta[b(x_i) - u], \quad u = 1 \dots m \quad (2)$$

où, m est la taille de l'histogramme, δ est la fonction de Kronecker, $b(x_i)$ est la classe de l'histogramme correspondant au pixel x_i , $\|x_i\|$ est la distance (dans notre cas une distance géodésique définie dans l'équation 3) entre le pixel d'intérêt x_i et le centre de la fenêtre, n est le nombre de pixels dans la fenêtre, k est le noyau et C est une constante de normalisation.

Si l'on considère deux points sur la sphère (X_{sph} and Y_{sph}), la distance géodésique (définie dans [8]) entre ces deux points peut être calculée par

$$\forall X_{sph}, Y_{sph} \in S^2 \quad d(X_{sph}, Y_{sph}) = \arccos(X_{sph} \cdot Y_{sph}) \quad (3)$$

3.2 Espace couleur

La modélisation des caractéristiques couleur d'un objet à l'aide d'un histogramme est une méthode rapide mais peu robuste aux changements d'illumination très fréquents avec les systèmes catadioptriques. Le choix d'un espace couleur moins sensible est donc déterminant pour permettre un suivi performant. Partant de ce constat, notre choix s'est porté sur l'espace couleur "transformed rgb", défini dans [17], qui est considéré comme robuste aux changements de lumière et d'intensité. Cette transformation repose sur la normalisation de chaque canal couleur de manière indépendante :

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} \frac{R - \mu_R}{\sigma_R} \\ \frac{G - \mu_G}{\sigma_G} \\ \frac{B - \mu_B}{\sigma_B} \end{pmatrix}, \quad (4)$$

σ et μ étant respectivement l'écart type et la moyenne du canal considéré.

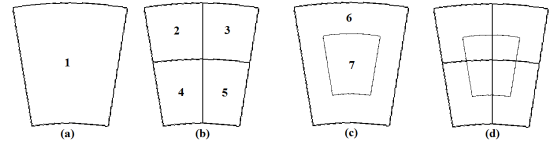


FIGURE 3 – Représentation multi-parties (a) région d'intérêt complète (b) division en 4 parties (c) division sensible aux changements d'échelle (d) représentation finale

La représentation d'un objet basée sur sa couleur ne considère pas l'organisation spatiale de la couleur de la cible, il est possible d'améliorer cette approche par l'ajout d'un noyau. Le noyau permet d'attribuer un poids à chaque pixel en fonction de sa position dans la région d'intérêt.

3.3 Le noyau

Afin de garantir une meilleure localisation de la cible, l'utilisation d'un noyau est fortement recommandée car il nous permet de prendre en compte des informations spatiales en plus de la distribution couleur d'un objet, une augmentation significative des performances est apportée. Son utilisation est d'ailleurs requise pour certains algorithmes de suivi tel que celui basé sur le *mean-shift*. Cependant, cet aspect n'a pas encore été abordé dans le cas d'image sphérique. Nous utilisons un noyau d'Epanechnikov pour obtenir une estimation non-paramétrique de la distribution couleur. Ce profil initialement conçu pour les images perspectives doit être adapté afin de pouvoir être défini correctement sur la sphère unité.

Le profil d'Epanechnikov est défini par

$$k(x) = \begin{cases} \frac{1}{2} C_d^{-1} (n+2)(1-x) & \text{si } x \leq 1 \\ 0 & \text{sinon} \end{cases} \quad (5)$$

où C_d est le volume de la sphère unité et n le nombre de dimension spatiale. Ici, nous souhaitons calculer ce profil directement sur la sphère unité. Nous calculons la distance d'un point avec le centre de la fenêtre par une distance géodésique.

3.4 Représentation multi-parties

Dans une séquence vidéo catadioptrique, l'échelle de la cible peut varier très rapidement. Il est par conséquent important de pouvoir gérer ce type d'inconvénient en adaptant au mieux la taille de la région d'intérêt. La représentation par histogramme multi-parties permet une meilleure localisation spatiale de la cible, ainsi qu'une meilleure gestion des changements d'échelle et une plus grande robustesse. Dans [10], les auteurs démontrent l'efficacité de cette méthode appliquée aux images perspectives. Au lieu de calculer un seul histogramme, cette approche consiste à en calculer 7 selon l'ajustement proposé dans la figure 3. L'histogramme numéro 1 (fig.3(a)) correspond à l'histogramme de la région d'intérêt entière. Dans un deuxième temps, la fenêtre sera scindée en quatre

parties, et un histogramme sera calculé dans chacune d'entre elles. Cette organisation permet d'obtenir une information sur la rotation de la cible. La dernière étape est la division de la fenêtre en deux parties (intérieure et extérieure) qui permet une meilleure adaptation à l'échelle.

Le calcul de la similarité entre deux histogrammes multi-parties q et p est donné par :

$$\rho(q, p) = \frac{\sum_{i=1}^N \beta[q^i, p^i]}{N}, \quad (6)$$

où, N est le nombre total d'histogrammes (7 histogrammes dans notre cas) tandis que q^i et p^i sont respectivement, les histogrammes de la i^e sous-partition de p et q . β est le coefficient de Bhattacharyya [7] représentatif de la similarité entre deux histogrammes :

$$\beta(q, p) = \sum_{u=1}^m \sqrt{q_u \cdot p_u} \quad (7)$$

Où m est le nombre de classes dans l'histogramme.

4 Algorithmes de suivi adaptés

4.1 Filtre particulière adapté

Le filtre particulière traditionnel utilise les coordonnées cartésiennes (x, y) afin de diffuser les particules directement sur le plan image. Dans l'approche proposée, les particules sont diffusées sur la sphère unité afin d'obtenir un comportement plus adapté aux particularités des capteurs catadioptriques. Nous pouvons alors définir le vecteur d'état par

$$S = [\theta, \phi, \dot{\theta}, \dot{\phi}, \delta\theta, \delta\phi], \quad (8)$$

où, les variables θ et ϕ sont les coordonnées du centre de la particule, $\delta\theta$ et $\delta\phi$ sont les dimensions de la fenêtre sur la sphère.

Les particules sont propagées selon le modèle dynamique suivant :

$$S_t = A \cdot S_{t-1} + W_{t-1}, \quad (9)$$

avec A la composante déterministe du modèle (ici, nous modélisons la dynamique de l'objet en mouvement par un modèle auto-régressif d'ordre 1) et W un bruit blanc gaussien.

4.2 Suivi Mean-Shift adapté

L'adaptation de l'algorithme (MS) *Mean-shift* [7] à la sphère unité nécessite une utilisation adéquate des coordonnées sphériques, notamment la gestion de la transition entre 0 et 2π . Le suivi utilisant *Mean-Shift* offre des résultats acceptables mais reste assujéti à de nombreuses contraintes comme l'occultation de la cible ou un changement trop important de son apparence. L'ajout d'un filtre de Kalman à l'algorithme MS est une bonne solution pour corriger ces problèmes. Lorsque le coefficient de similarité entre le modèle et la cible est faible, on peut considérer la cible comme étant perdue. Dans ce cas, le filtre de Kalman

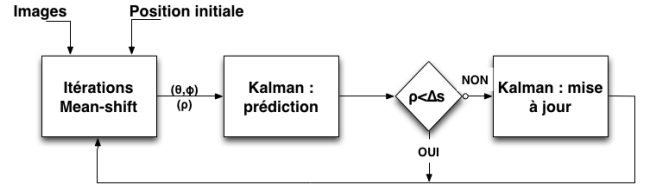


FIGURE 4 – Diagramme de fonctionnement Mean-Shift + Kalman

peut prendre le relais afin d'estimer la nouvelle position de l'objet en fonction de ces états précédents. Concrètement, après chaque itération de l'algorithme MS, nous calculons l'état prédit de l'objet en fonction de son état précédent. Après cette phase de prédiction, nous comparons le coefficient de similarité obtenu par la méthode MS avec un seuil (Δs). Si ρ est supérieur à Δs , nous considérons la mesure (calculée avec MS) comme suffisamment fiable, dans ce cas l'estimation offerte par le filtre de Kalman est mise à jour afin de corriger la prédiction (ce qui a également pour effet de lisser le suivi). Sinon la mise à jour n'est pas effectuée, ce qui signifie que la nouvelle position de l'objet est uniquement déterminée par la prédiction du filtre de Kalman. Le fonctionnement de la méthode est décrit par la figure 4.

Cependant, la trajectoire d'un objet sur une image catadioptrique sera différente de celle d'un objet sur une image perspective comme l'illustre la figure 5.

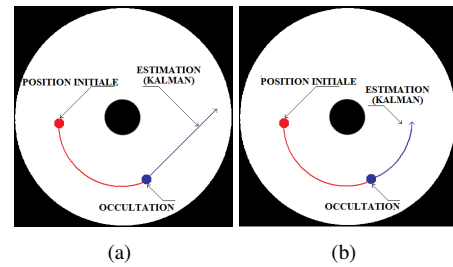


FIGURE 5 – Filtre de Kalman (a) Non-adapté (b) adapté

Ce qui signifie que l'estimation fournie par Kalman devra être faite sur la sphère unité afin de respecter la géométrie du capteur. Le vecteur d'état sera alors exprimé de la façon suivante :

$$X = [\theta, \phi, \dot{\theta}, \dot{\phi}] \quad (10)$$

Dans cet article les algorithmes basés sur la méthode MS incorporent un filtre de Kalman.

5 Expériences et résultats

Dans cette section, nous mettons en œuvre des expériences permettant de quantifier l'efficacité de nos méthodes de suivi. Au regard de l'absence de base de données concernant le suivi d'objet avec des capteurs catadioptriques, les

vidéos utilisées dans ce travail ont été directement acquises par nos soins dans différentes conditions afin d'offrir un large panel de situations : scène intérieure ou extérieure, objet ou humain en déplacement, avec ou sans occultation, caméra mobile ou statique... (voir tableau 1)

Afin d'apporter un bon aperçu des améliorations offertes par l'approche proposée, nous comparons les méthodes traditionnelles et adaptées. La *toolbox* utilisée pour la calibration est la "OCamCalib" de Davide Scaramuzza [15]. La résolution des images est de 640x480 pixels et la durée des séquences oscille entre 500 et 800 images.

5.1 Évaluation des performances des algorithmes de suivi

Il existe dans la littérature de nombreuses méthodes statistiques permettant l'évaluation des algorithmes de suivi. Les tests réalisés ici sont basés sur la création manuelle d'une vérité de terrain. Trois critères représentatifs de la précision des différentes méthodes sont utilisés : la superposition spatiale, la superposition temporelle, et la distance entre les centres.

Estimation des critères de superposition

La superposition spatiale correspond au pourcentage de surface commune $A(St, Gt)$ entre les deux boîtes englobantes (la fenêtre issue de l'algorithme de suivi St et la fenêtre obtenue par notre vérité de terrain Gt), illustré fig.6. Ces deux régions sont comparées de cette manière :

$$A(St, Gt) = \frac{\text{Surface}(Gt \cap St)}{\text{Surface}(Gt \cup St)} \quad (11)$$

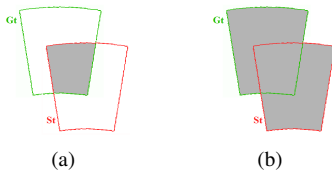


FIGURE 6 – (a)Surface($Gt \cap St$) (b)Surface($Gt \cup St$)

En ce qui concerne la superposition temporelle, un seuil est fixé arbitrairement afin de déterminer le pourcentage d'image où le suivi est considéré comme "correct" dans la séquence. Concrètement, on peut par exemple fixer un seuil de superposition spatiale minimum T_{Ov} à 20%, ce qui signifie que le suivi est correct si la superposition entre la vérité de terrain et le résultat de l'algorithme de suivi est supérieure à 20% :

$$TO(St, Gt) = \begin{cases} 1 & \text{if } A(St, Gt) > T_{Ov} \\ 0 & \text{if } A(St, Gt) < T_{Ov} \end{cases} \quad (12)$$

Ici, $TO(St, Gt)$ est le pourcentage d'images avec une superposition correcte (définie par le seuil T_{Ov}).

Distance entre les centres

La distance entre le centre de la fenêtre de la vérité terrain et le centre de la fenêtre obtenu par la méthode de suivi peut également être une information représentative de la précision de l'algorithme. Cette distance est calculée de manière euclidienne sur le plan image. Ce critère est certainement moins significatif que ceux présentés précédemment mais offre tout de même une information complémentaire dans l'évaluation proposée.

5.2 Résultats

L'ensemble des résultats obtenus (tableaux 2 et 3) nous permet de confirmer l'hypothèse de départ. En effet, l'adaptation des méthodes de suivi visuel dans les images catadioptriques apporte une précision supérieure aux méthodes traditionnelles. Notons tout de même que le filtre particulaire reste la méthode la plus fiable. Dans la première séquence, une personne marche autour d'une caméra fixe avec un éclairage naturel. Malgré l'apparente simplicité de cette vidéo nous remarquons de forts écarts de résultats entre les méthodes adaptées et les méthodes conventionnelles. Nous pouvons en déduire qu'une fenêtre de sélection rectangulaire n'est pas en mesure de sélectionner correctement des objets sur des images catadioptriques (voir figure 7(a)). La seconde séquence nous permet quant à elle d'étudier la robustesse de nos méthodes à d'important changements d'échelle (figure 7(b)). Des changements assez importants dans l'apparence de l'objet au cours du temps (surface réfléchissante) sont également notables. Dans cette séquence, les méthodes de suivi conventionnelles ne sont pas capables de suivre correctement l'objet durant la totalité de la séquence. Avec le filtre particulaire classique, la cible est d'ailleurs rapidement perdue (voir superposition temporelle dans le tableau 2). Dans ce cas notre estimation sur la superposition spatiale n'est effectuée que sur la partie de la séquence correctement suivie. A l'inverse, le filtre particulaire adapté permet de suivre correctement la cible dans presque la totalité de la séquence. Nous pouvons attribuer ces bonnes performances à l'utilisation d'un histogramme multi-parties (pour les méthodes adaptés) qui permet de mieux gérer les problèmes de changements d'échelle.

Le suivi MS adapté donne quant à lui une meilleure précision en ce qui concerne la localisation du centre de la cible. La séquence numéro trois fait intervenir de nombreux phénomènes typiques des capteurs catadioptriques, c'est à dire un fort changement de luminosité (éblouissement), une occultation et des changements d'échelle rapide avec une caméra mobile (figure 7(c)). Ici encore, les méthodes adaptées fournissent des résultats très convaincants en tous points comparativement aux méthodes conventionnelles. Finalement, la séquence 4 confronte les algorithmes de suivi à des occultations répétées sur un objet de taille réduite. Cette fois encore nous observons de meilleurs résultats avec notre approche en particulier en ce qui concerne le MS adapté. Notons que l'utilisation d'un noyau est un élément déterminant qui influe directement sur la qualité

du suivi. L'utilisation d'un noyau est d'ailleurs essentielle à l'algorithme MS, les performances atteintes par l'adaptation de cette méthode tendent à démontrer une parfaite adéquation de ce noyau sur la sphère. Pour mettre en exergue le gain de précision offert par l'emploi d'un noyau, une série de tests a été effectuée (voir tableau 4), ces expériences confrontent le filtre particulière avec et sans noyau. Pour montrer l'impact du noyau, la représentation multi-histogrammes n'a pas été utilisée (ce qui explique les scores moins importants et qui donne également un aperçu des améliorations apportées par l'utilisation de plusieurs histogrammes). Dans tous les cas de figures, un gain en terme de performance est apporté. De manière quantitative aussi bien que qualitative, notre méthode d'adaptation permet une meilleure gestion des changements d'échelles (toutes les séquences), des occultations (séquences 3 et 4) et des forts changements d'illumination (séquence 2).

6 Conclusion

Dans cet article, nous avons proposé une manière performante permettant l'adaptation des méthodes de suivi visuel basée sur une représentation sphérique de l'image et une modélisation de la cible avec un histogramme couleur (pondéré par un noyau adapté assurant une qualité de suivi élevée). Les résultats expérimentaux fondés sur des critères d'estimation tangible (dans plusieurs séquences et dans différentes conditions) ont permis de montrer une amélioration significative de la précision du suivi. Notre méthode d'adaptation permet de manière générale d'améliorer la robustesse aux occultations, aux changements d'échelle et aux changements d'illumination.

Ce travail peut être complété de différentes manières. La possibilité de suivre plusieurs cibles simultanément, l'utilisation d'une caractéristique de modélisation de l'apparence autre que la couleur (e.g. texture, forme) et un suivi utilisant un réseau de caméras peuvent être des pistes intéressantes.

Références

- [1] H. Bakstein and A. Leonardis. Catadioptric image-based rendering for mobile robot localization. In *Proceedings of the 6th International Conference on Computer Vision*, pages 1–6, 2007.
- [2] J. P. Barreto. General central projection systems, modeling, calibration and visual servoing. Technical report, University of Coimbra,, 2003.
- [3] J.P. Barreto and H. Araujo. Direct least square fitting of paracatadioptric line images. In *OMNIVIS*, page 78, 2003.
- [4] J.P. Barreto and H. Araujo. Geometric properties of central catadioptric line images and their application in calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 :1327–1333, 2005.
- [5] J.C. Bazin, K.J. Yoon, , I.S. Kweon, C. Demonceaux, and P. Vasseur. Particle filter approach adapted to catadioptric images for target tracking application. In *20th British Machine Vision Conference, (BMVC'09)*, 2009.
- [6] T.E. Boulton, R. Micheals, X. Gao, P. Lewis, C. Power, W. Yin, and A. Erkan. Frame-rate omnidirectional surveillance and tracking of camouflaged and occluded targets. In *Second IEEE International Workshop on Visual Surveillance*, pages 48–55, 1999.
- [7] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 :564–575, May 2003.
- [8] C. Demonceaux and P. Vasseur. Omnidirectional image processing using geodesic metric. In *Proceedings of the 16th IEEE International Conference on Image processing, ICIP'09*, pages 221–224, 2009.
- [9] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 45 :223–243, 2002.
- [10] E. Maggio and A. Cavallaro. Multi-part target representation for color tracking. In *Proc. Int. Conf. Image Processing, Italy*, pages 729–732, 2005.
- [11] E. Maggio and A. Cavallaro. *Video Tracking : Theory and Practice*. Wiley, February 2011.
- [12] E. Marchand and F. Chaumette. Fitting 3d models on central catadioptric images. In *IEEE Int. Conf. on Robotics and Automation, ICRA'07*, pages 52–58, Roma, Italy, 2007.
- [13] C. Mei, S. Benhimane, E. Malis, and P. Rives. Constrained multiple planar template tracking for central catadioptric cameras. In *British Machine Vision Conference*, pages 4–7, 2006.
- [14] C. Mei and E. Malis. Fast central catadioptric line extraction, estimation, tracking and structure from motion. In *IROS*, pages 4774–4779, 2006.
- [15] D. Scaramuzza. Omnidirectional camera and calibration toolbox for matlab. http://robotics.ethz.ch/~scaramuzza/Davide_Scaramuzza_files/Research/OcamCalib_Tutorial.htm, May 2009.
- [16] M. Taiana, J. Gaspar, J. Nascimento, R. Bernardino, and Pedro Lima. 3d tracking by catadioptric vision based on particle filters. In *RoboCup International Symposium*, 2007.
- [17] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9) :1582–1596, 2010.
- [18] X. Meng Z. Cao and S. Liu. Dynamic omnidirectional vision localization using a beacon tracker based on particle filter. In *Computer Vision, Xiong Zhihui (Ed.), ISBN : 978-953-7619-21-3*, pages 538–554, (2008).

TABLE 1 – Particularités des séquences utilisées

	Localisation	Type de caméra	Changement d'illumination	Occultation	Cible	Images/sec	Nombre d'images
Séquence 1	Extérieure	Non centrale	Non	Non	Personne	30	780
Séquence 2	Intérieure	Non centrale	surface réfléchissante	Non	Objet	30	649
Séquence 3	Extérieure	Non centrale	Éblouissement	Partiel	Personne	30	602
Séquence 4	Intérieure	Centrale	Non	Total	Objet	15	481

TABLE 2 – Résultats moyens obtenus avec un filtre particulaire conventionnel et avec la méthode adaptée

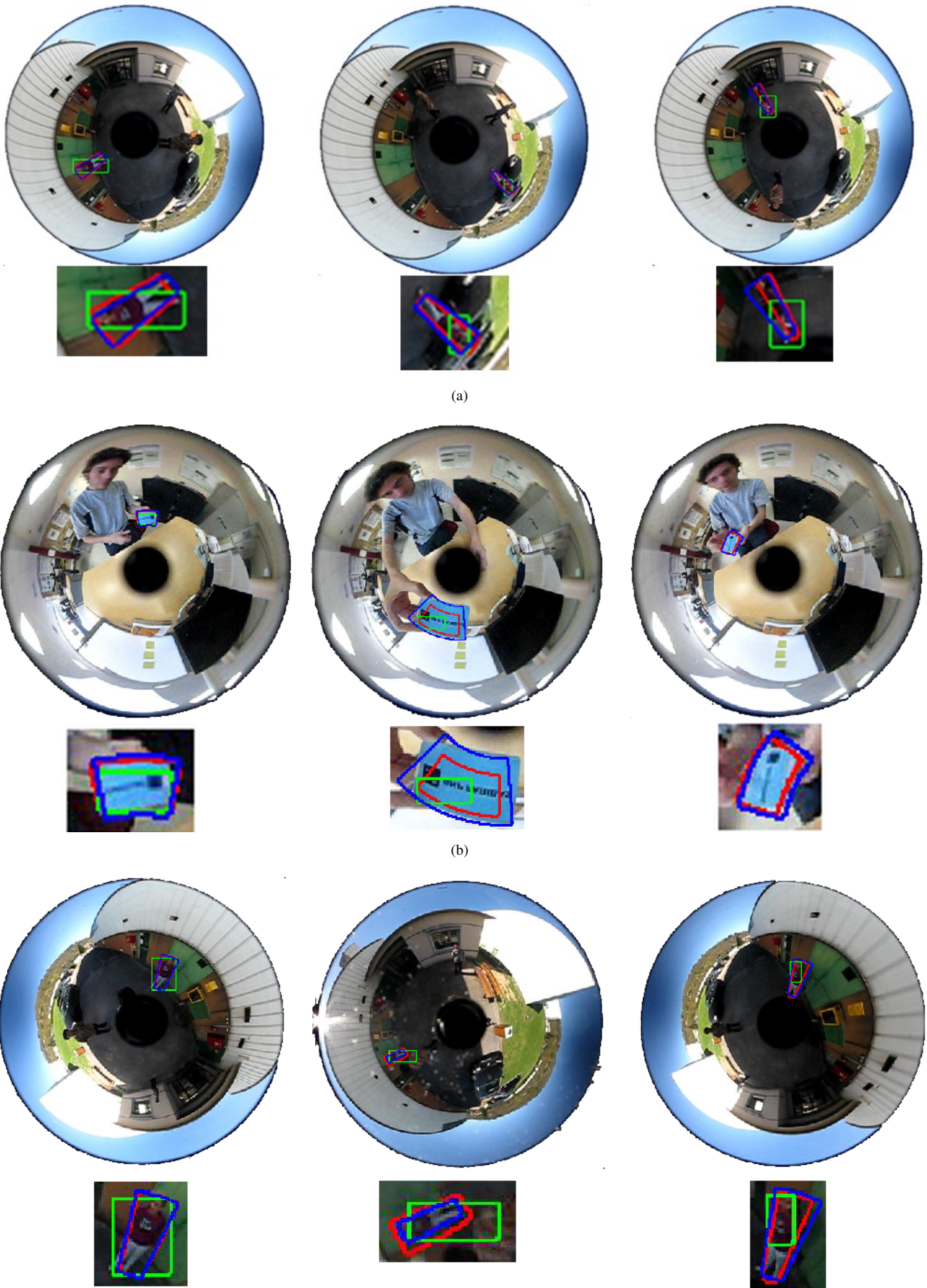
Séquence	Méthode	superposition spatiale	distance des centres	superposition temporelle
Sequence 1	Méthode conventionnelle	45%	5.7	100%
	Notre méthode	71%	2.6	100%
Sequence 2	Méthode conventionnelle	30%	5.5	44%
	Notre méthode	60%	3.8	99.7%
Sequence 3	Méthode conventionnelle	33%	7.7	88%
	Notre méthode	65%	4.8	99.5%
Sequence 4	Méthode conventionnelle	45%	6	96.8%
	Notre méthode	66%	4.1	97.3%

TABLE 3 – Résultats moyens obtenus avec l'algorithme *Mean-Shift* conventionnel et avec la méthode adaptée

Séquence	Méthode	superposition spatiale	distance des centres	superposition temporelle
Sequence 1	Méthode conventionnelle	32.4%	6.38	63.21%
	Notre méthode	64.5%	6.2	96.35%
Sequence 2	Méthode conventionnelle	42.5%	16.2	89%
	Notre méthode	47.79%	8.3	90%
Sequence 3	Méthode conventionnelle	40.55%	7.8	96%
	Notre méthode	61.8%	7.5	97.84%
Sequence 4	Méthode conventionnelle	34.72%	10.52	52.81%
	Notre méthode	67.67%	5.02	97.71%

TABLE 4 – Résultats moyens obtenus avec un filtre particulaire adapté, avec et sans noyau

Séquence	Noyau	superposition spatiale	distance des centres	superposition temporelle
Sequence 1	Sans	62,47%	4,93	98,7%
	Avec	65,67%	4,01	99,23%
	Gain	+3,2%	-0,92	+0,53%
Sequence 2	Sans	33,28%	4,46	80,7%
	Avec	40,26%	4,46	90,6%
	Gain	+6,98%	0	+9,9%
Sequence 3	Sans	54,37%	5,35	96%
	Avec	59,06%	5,04	99%
	Gain	+4,69%	-0,31	+3%



(a)

(b)

(c)

RFIA 2012

Lyon, 24-27 janvier 2012

FIGURE 7 – Aperçu des résultats de suivi visuel avec un filtre particulaire conventionnel (fenêtre verte), un filtre particulaire adapté (fenêtre rouge) et la vérité terrain (fenêtre bleue) (a) Séquence 1 (b) Séquence 2 (c) Séquence 3