

Détection multi-frames de satellites à faible rapport signal à bruit

Thomas DELAITE¹, David SAVARY MARTINEZ^{1,2}, Eric GLEMET¹, Baptiste LEVASSEUR¹, Jocelyn COUETDIC¹, Laurent MUGNIER³, Frédéric CASSAING¹

¹DTIS, ONERA, Université Paris Saclay
6 Chemin de la Vauve aux Granges, 91120 Palaiseau

²Image Processing and Computer Vision Erasmus Mundus Mater Program, entre Université de Bordeaux, France (UBx), Pázmány Péter Catholic University, Hungary (PPCU) et Universidad Autónoma de Madrid, Spain (UAM)

³DOTA, ONERA, Université Paris Saclay
29 Avenue de la Division Leclerc, 92322 Châtillon
thomas.delaite@onera.fr

Résumé – Une méthode de type *track before detect* a été développée pour optimiser le rapport signal à bruit de détection de cibles mobiles sur un fond fort. Cette méthode massivement parallèle exploite la récente disponibilité de composants rapides (caméras faible-flux sCMOS, calculateurs GPU) pour minimiser l’impact du fond. Cette méthode est illustrée par la détection optique de satellites en orbite terrestre moyenne.

Abstract – A *track before detect* method has been developed to optimize the detection signal-to-noise ratio for moving targets over a strong background. This massively parallel method makes use of the recent availability of fast components (low-light sCMOS cameras, GPU accelerators) to minimize the background contribution. The method is illustrated by the optical detection of satellites in medium Earth orbit.

1 Contexte et objectifs

L’ONERA s’équipe actuellement d’une station d’étude pour l’observation optique dédiée à la surveillance spatiale. L’objectif principal est d’analyser les différents contributeurs de l’ensemble de la chaîne, de l’instrument aux traitements de données, à la performance de catalogage des objets en orbite moyenne. Les performances découleront directement de la capacité à détecter des cibles mobiles peu lumineuses sur un fond stellaire, surtout en conditions dégradées (clair de lune,...).

La détection et le pistage de cibles mobiles est un sujet bien étudié. Les nombreuses méthodes de flot optique ou d’association plot-piste nécessitent un rapport signal à bruit (RSB) suffisant pour détecter les cibles dans la plupart des images. À l’opposé, les méthodes *track-before-detect* (TBD) cherchent à identifier des pistes complètes avant de décider d’une détection, mais nécessitent une connaissance du comportement des cibles et une grande capacité d’acquisition et de traitement [1].

En astronomie, les sources peu lumineuses sont observées avec des temps de pose longs. Pour la surveillance de l’espace, cela conduit à détecter les segments brillants formés par les objets mobiles [2]. Cependant, cette méthode ne permet pas de détecter les objets les plus faibles. Certains systèmes récents utilisent des méthodes TBD [3] mais sont encore peu diffusés dans la littérature.

On s’intéressera dans la section 2 au problème général de la détection, en aval d’un capteur d’image numérique, d’une cible mobile sur un fond fixe, éventuellement texturé, et à l’estima-

tion de sa position et de sa vitesse. Ce cas général sera ensuite décliné pour notre application en section 3 et les performances évaluées par simulation en section 4.

2 La détection de cibles mobiles

Après avoir introduit les modèles et notations, nous présentons une extension d’une méthode de maximum de vraisemblance puis l’appliquons à un mouvement rectiligne uniforme.

2.1 Modélisation des signaux

Le capteur utilisé est une caméra dans un plan conjugué des cibles. Sans perte de généralité, et pour des raisons de simplicité, on suppose que pendant le temps de mesure T_{total} , le flux continu d’images est échantillonné à temps discret avec un pas constant T , et on note k l’indice temporel. Toutes les coordonnées seront représentées en coordonnées cartésiennes bidimensionnelles dans ce plan image et dénotées de manière générique r . Le bruit de lecture, supposé gaussien, est supposé uniforme sur les pixels, d’écart-type σ_{lec} .

Les cibles sont supposées avoir une image de forme connue sur la caméra. Nous considérerons ici le cas de cibles non-résolues, de sorte que leur image soit la réponse impulsionnelle (RI) h du système optique ; ce modèle simple peut se généraliser. L’amplitude du signal est caractérisée par \dot{N} , l’intensité de la cible (en photoelectron.s⁻¹). La cinématique des cibles est supposée être un mouvement rectiligne et uniforme pendant le

temps total d'observation, caractérisée par une position \mathbf{r}_0 à un temps de référence noté k_0 et une vitesse \mathbf{v} dans le plan image. Le modèle instantané d'une cible est la convolution entre le flux ponctuel d'intensité \dot{N} et la RI h du système optique, noté

$$a\dot{s}(\mathbf{r}, t; \mathbf{r}_0, \mathbf{v}) = \dot{N} h(\mathbf{r} - (\mathbf{r}_0 + \mathbf{v}t)) \quad (1)$$

où s est la répartition du signal et a son intensité. Pour un temps d'exposition T inférieur ou équivalent au temps de passage d'une cible sur un pixel et une RI gaussienne de largeur quadratique moyenne σ_h , l'intégration de \dot{s} lors de la $k^{\text{ième}}$ pose s'approxime par

$$\frac{as(\mathbf{r}, k; \mathbf{r}_0, \mathbf{v})}{\dot{N}T} = \frac{\exp\left(-\frac{(\mathbf{r}_0 + (k+1/2)\mathbf{v}T - \mathbf{r})^2}{2\sigma_h^2}\right)}{C} \quad (2)$$

où C est une constante de normalisation.

Un fond F s'ajoute aux cibles, incluant le courant d'obscurité du détecteur et d'éventuels autres contributeurs (étoiles, fond de ciel,...). Ce fond, éventuellement texturé et supposé d'espérance constante pendant la durée totale de la mesure, s'estime classiquement par filtrage médian, pour ne pas supprimer le signal des cibles et apporter de la robustesse vis-à-vis de la distribution du bruit (outliers). Une fois le fond ôté, il reste dans les images sa fluctuation supposée poissonnienne.

De manière générale, l'image une fois le fond retiré peut être modélisée comme la somme des cibles mobiles présentes dans l'image et du bruit b ; pour une unique cible :

$$I(\mathbf{r}, k) = as(\mathbf{r}, k; \mathbf{r}_0, \mathbf{v}) + b(\mathbf{r}, k), \quad (3)$$

où on suppose que b est un bruit blanc de variance

$$\sigma_I^2(\mathbf{r}, k) = \mathbb{E}_k(F(\mathbf{r}, k)) + \sigma_{lec}^2(\mathbf{r}, k) \quad (4)$$

et de moyenne nulle, et a représente l'intensité de la source qui est répartie sur les pixels. Sous ces hypothèses, le modèle de cible se réduit à 5 paramètres dans un repère de référence du plan image : les 2 composantes de la position initiale \mathbf{r}_0 et de la vitesse \mathbf{v} , et l'intensité a .

2.2 Détection via le maximum de vraisemblance

La méthode utilisée repose sur une extension de la méthode Androméda de détection d'exo-planètes [4]. La première étape consiste à écrire la vraisemblance L des paramètres d'une cible; dans l'hypothèse d'un bruit gaussien (fond assez fort) :

$$L(\mathbf{r}_0, \mathbf{v}, a) \propto \exp\left\{-\sum_{k,r} \frac{[I(\mathbf{r}, k) - a s(\mathbf{r}, k; \mathbf{r}_0, \mathbf{v})]^2}{2\sigma_I^2(\mathbf{r}, k)}\right\} \quad (5)$$

Le valeur de \hat{a} qui maximise L peut être injecté dans le rapport de vraisemblance, et maximiser la vraisemblance revient à maximiser le rapport signal à bruit expérimental $\hat{r}_{s/b}$ défini comme :

$$\hat{r}_{s/b}(\mathbf{r}_0, \mathbf{v}) = \frac{\hat{a}(\mathbf{r}_0, \mathbf{v})}{\sigma[\hat{a}(\mathbf{r}_0, \mathbf{v})]} \quad (6)$$

$$\text{qui se réduit à } \hat{r}_{s/b}(\mathbf{r}_0, \mathbf{v}) = \frac{\sum_{k,r} \frac{I(\mathbf{r}, k) s(\mathbf{r}, k; \mathbf{r}_0, \mathbf{v})}{\sigma_I^2(\mathbf{r}, k)}}{\sum_{k,r} \frac{s^2(\mathbf{r}, k; \mathbf{r}_0, \mathbf{v})}{\sigma_I^2(\mathbf{r}, k)}}^{\frac{1}{2}} \quad (7)$$

$\hat{r}_{s/b}$ est effectivement une estimation du rapport signal à bruit théorique du signal a que l'on définit comme :

$$r_{s/b} = \frac{a}{\sigma[\hat{a}]} = a \times \prod_k \times \prod_r \frac{s^2(\mathbf{r}, k)}{\sigma_I^2(\mathbf{r}, k)}^{\frac{1}{2}}, \quad (8)$$

qui ne dépend pas des données, mais seulement des bruits et de la répartition exacte du signal (qui diffère du modèle gaussien par l'allongement de la trace des cibles rapides notamment), et sert de référence pour les performances de détections en 4.

2.3 Cas d'un mouvement rectiligne uniforme

Pour un modèle de cible arbitraire, il faudrait recalculer la corrélation données-modèle pour toute valeur de paramètre. Pour des modèles stationnaires, le calcul se simplifie [5]. Ici, la dynamique rectiligne uniforme permet d'écrire

$$s(\mathbf{r}, k; \mathbf{r}_0, \mathbf{v}) = s(\mathbf{r} - \mathbf{r}_0 - k\mathbf{v}T, k; 0, 0). \quad (9)$$

On peut donc calculer une seule fois la corrélation croisée

$$I_{corr}(\mathbf{r}, k) = \sum_{\mathbf{r}'} \frac{s(\mathbf{r}', k; \mathbf{r}, 0) I(\mathbf{r}', k)}{\sigma_I^2(\mathbf{r}', k)} \quad (10)$$

et la réutiliser en la translatant

$$\hat{r}_{s/b}(\mathbf{r}_0, \mathbf{v}) = \frac{\sum_k I_{corr}(\mathbf{r}_0 + k\mathbf{v}T, k)}{\sigma[\hat{a}(\mathbf{r}_0, \mathbf{v})]}. \quad (11)$$

Cette forme est extrêmement efficace en termes de calcul et permet d'envisager une recherche de maximum par discrétisation de l'espace des paramètres en temps contraint.

2.4 Méthodologie de détection

Le nombre de cibles étant inconnu, on procède en trois étapes. Tout d'abord, selon la précision voulue ou les contraintes calculatoires, on définit un pas d'échantillonnage sur les pixels et on discrétise l'espace des paramètres \mathbf{v} puis on calcule la carte de vraisemblance $\hat{r}_{s/b}$.

Ensuite, on recherche les maximums locaux supérieurs à un seuil défini en fonction d'un taux soit de détection pour un RSB de cible donnée $\tau_d r_{s/b}$, soit de fausse détection maximale $\tau_{fd} \sigma_0(\mathbf{r}_0, \mathbf{v})$. Les deux approches sont utilisées en section 4.

Enfin, on raffine la grille de quantification autour des maximums locaux pour gagner en précision, avec un coût calculatoire faible.

2.5 Temps de pose optimal

Au sein du temps total d'observation T_{total} , la durée T de chacune des M poses élémentaires est un paramètre libre tant que $T_{total} = MT$ pour conserver le flux total. Réduire T augmente M donc le bruit de lecture alors que l'allonger augmente la contribution du fond (donc son bruit) sur un pixel sans augmenter le signal une fois la cible partie. Un temps caractéristique important est donc le temps T_{pix} de passage de la cible sur

un pixel. En identifiant dans la variance du bruit le terme σ_{lec}^2 résultant de chaque lecture et un terme σ_{fond}^2 égal au flux en photo-électrons. s^{-1} du fond, on montre que le meilleur compromis T_{opt} s'approxime par

$$T_{opt} \approx \frac{2T_{pix}\sigma_h\sigma_{lec}^2}{\sigma_{fond}^2}. \quad (12)$$

Comme attendu, T_{opt} est proportionnel à T_{pix} , avec un coefficient traduisant le régime de bruit, croissant avec le bruit de lecture et décroissant avec le bruit de fond. L'évolution technologique (bruit de lecture réduit, rendement quantique accru) diminue ce coefficient qui peut même être inférieur à 1.

De nuit, pour un satellite en orbite MEO, et notre optique, les valeurs typiques sont $\sigma_h = 2pix$, $\sigma_{lec}^2 = 3e^{-pix^{-1}}$, $\sigma_{fond}^2 = 300e^{-.s^{-1}}$ et $T_{pix} = 0.05s$, $T_{opt} \approx 0.07s$ soit une fréquence de 14 images. s^{-1} .

Il est donc préférable, à temps d'observation T_{total} fixé, de disposer d'une pile de trames de faible durée T_{opt} plutôt que d'un nombre inférieur de trames si l'on souhaite rechercher de faibles cibles au sein d'un fond important. Ceci a d'importantes répercussions sur le dispositif, qui devra inclure une caméra rapide et d'importantes capacités de calcul ou de stockage. La récente disponibilité de caméras sCMOS rétro-éclairées, conjuguant grande vitesse de lecture et faible bruit, couplée aux capacités croissantes des processeurs graphiques renforcent désormais l'attractivité de la méthode TBD présentée.

3 Application à l'orbitographie

La méthode décrite en section 2 sera prochainement mise en œuvre au sein de la station CICLOPE en construction. Dans une première phase, on se limitera aux satellites d'altitude supérieure à 20 000 km, dont la vitesse apparente est inférieure à $v_{max}=40''/s$. Ces satellites, de dimension quelques mètres, sous-tendent un angle d'un dixième de seconde d'arc. Ils sont donc non résolus par un télescope de diamètre sub-métrique, d'autant plus que la turbulence atmosphérique limite la résolution angulaire et impose à la réponse impulsionnelle du télescope une forme gaussienne, de largeur à mi-hauteur typiquement supérieure à la seconde d'arc.

Les principaux paramètres de CICLOPE, incluant un tube optique Celestron RASA 14, une monture Astelco NTM-500, une caméra XIMEA MX377 et un serveur DELL R750xa avec trois processeurs nVidia A40, sont rassemblés en table 1. La procédure de balayage du ciel consiste à observer cycliquement une bande formée par la juxtaposition d'images acquises pendant $T_{total} \approx 10s$. Le déplacement des cibles est donc inférieur

TABLE 1 – Caractéristiques de la station CICLOPE.

Paramètre	Valeur
Diamètre du télescope	350 mm
Définition de la caméra	6K 6K pixels
Pas pixel sur le ciel	2,6''
Fréquence maximale de la caméra	46 trames/s

à 400'', soit moins d'un 3 000^{ème} d'orbite, portion sur laquelle l'hypothèse de linéarité est légitime.

Pour la mise en œuvre de l'algorithme décrit en section 2, le télescope est placé en suivi stellaire afin que le fond astronomique soit fixe. Les satellites apparaissent donc sous la forme de taches mobiles, de positions et vitesses inconnues, mais de largeur connue puisque la RI du télescope peut s'étalonner sur les étoiles présentes dans le champ. Par ailleurs, la vitesse maximale des cibles étant connue, la vraisemblance est calculée sur une grille de vitesses couvrant toutes les orientations et les amplitudes inférieures à v_{max} . Le temps pixel est de 0,1 s pour une vitesse de 26''/s, donc des temps de pose voisins seront utilisés selon le fond (cf § 2.5). Les satellites plus rapides, plus proches, seront aussi plus brillants. Ce sont des piles de 100 images de 37 Mpixels qui arriveront toute les 10 s.

L'algorithme décrit en section 2 a été implémenté en C++ et en CUDA. Les hypothèses étant indépendantes, les parties intensives du calcul sont facilement parallélisables.

4 Résultats

Dans l'attente de la station, une série de N images où se déplacent des cibles a été simulée. Le modèle 1 est intégré donc les cibles rapides ($T_{pix} < T$) forment des taches allongées, puis on calcule le vrai rsb défini en Éq. (8) de chaque cible en utilisant la répartition exacte du signal.

Un fond aléatoire sert de racine pour générer des images avec bruit de Poisson. Un bruit de lecture gaussien de moyenne nulle est ajouté. La figure 1 présente un scénario de détection en conditions proches du réel, pour une probabilité de fausse détection fixée à 0.05 (soit $1.5\sigma_0$). Les cibles de RSB inférieur ne sont pas détectées, mais tous les satellites de RSB supérieur sont détectés, sans fausse alerte.

La figure 2 représente une simulation synthétique pour l'estimation des performances de la méthode. On simule $N = 100$ images où une série de cible de RSB divers se déplacent. On calcule $\hat{r}_{s/b}$ en discrétisant l'espace des paramètres avec un pas de 1. Pour quantifier le taux de détection, le seuil de détection pour chaque cible est fixé à $0.7r_{s/b}$. Un algorithme de Hongrois [6] associe les cibles en entrée et les détections en sortie, et les valide si sont respectées les conditions

$$k\mathbf{r}_{vrai} - \mathbf{r}_0 k < 2 \quad \text{et} \quad k\mathbf{v}_{vrai} - \mathbf{v} k < 2, \quad (13)$$

soit une erreur de l'ordre du pixel en position et en vitesse. On répète la simulation un grand nombre de fois, on mesure les taux de fausses détections et de cibles détectées

$$\tau_{fd} = \frac{\text{fausses détections}}{\text{détectations}} \quad \text{et} \quad \tau_d = \frac{\text{détectations valides}}{\text{nombre de cibles}} \quad (14)$$

en fonction du $r_{s/b}$ théorique des satellites. La limite de la méthode en détection se situe autour d'un $r_{s/b}$ de 5, soit pour $N = 100$ images un RSB moyen de 0.5 sur chacune des images. En terme de fausses détection, on dépasse 10% autour de 8.

Les cartes $\hat{r}_{s/b}$ représentées en pseudo couleurs (fig. 1b et 2b) sont projetées sur l'espace des positions en suivant

$$projection(\hat{r}_{s/b}(\mathbf{r}_0, \mathbf{v}))(r) = \max_v(\hat{r}_{s/b}(\mathbf{r}_0, \mathbf{v})). \quad (15)$$

