

PROPOSITION DE SUJET DE THESE

Intitulé : Optimisation sous incertitudes avec contraintes physiques

Référence : **PHY-DOTA-2024-02**
(à rappeler dans toute correspondance)

Début de la thèse : 01/10/2024

Date limite de candidature : 01/06/2024

Mots clés :

Optimisation ; incertitudes ; robustesse ; algorithmes évolutionnaires ; multi/hyperspectral

Profil et compétences recherchées :

Master 2 ou école d'ingénieur avec compétences en statistiques, optimisation, machine learning

Présentation du projet doctoral, contexte et objectif :

Depuis quelques années, on constate un intérêt croissant pour la spécification et la conception d'imageurs multispectraux, qui permettent d'acquérir simultanément l'image d'une scène dans plusieurs (typiquement entre 2 et 10) bandes spectrales en infrarouge ou en visible. Les applications de ces instruments sont très variées : détection d'aéronefs, de drones, télédétection satellitaire... Lors de l'exploitation des données, les algorithmes de détection d'anomalies sont couramment utilisés pour rechercher des objets potentiels d'intérêt sur une image. Le principe consiste à analyser le signal lu sur les pixels spectraux de l'image et à distinguer deux classes : une classe majoritaire associée au fond, et une classe de pixels qui se démarque du signal de fond, les anomalies. Celles-ci peuvent ensuite être analysées de façon plus détaillée par d'autres méthodes, pour savoir si elles correspondent aux objets recherchés.

Ce problème s'écrit comme un problème d'optimisation où il faut choisir des bandes spectrales qui vont maximiser un critère associé à performance de détection des anomalies (aire sous la courbe probabilité de détection / probabilité de fausse alarme par exemple). Les variables du problème d'optimisation encodent les bandes spectrales.

Afin d'optimiser le choix des bandes spectrales pour des applications de type détection d'anomalies, il est important de tenir compte des incertitudes qui affectent certaines données, comme les conditions météorologiques, et qui conduisent à une variabilité des scènes observées, ainsi que de la diversité des objets susceptibles d'être rencontrés. Il faut donc évoluer d'une stratégie d'optimisation classique et déterministe vers une optimisation robuste, qui a pour but de concevoir des systèmes dont les performances sont peu sensibles aux fluctuations induites par les paramètres incertains. Elle s'appuie sur des mesures de robustesse pour quantifier les variations des objectifs dues aux incertitudes. On recherche ainsi le système qui possède la meilleure performance moyenne, ou bien celui dont la performance varie le moins, ou encore les deux à la fois.

Dans ce contexte, des algorithmes d'optimisation à base d'aléa ou stochastiques peuvent s'avérer être des alternatives intéressantes. En effet, par construction ces méthodes sont robustes aux variations des solutions échantillonnées. Par ailleurs, si les algorithmes présentent des propriétés théoriques d'invariance comme l'invariance par transformation monotone croissante de la fonction objectif ou l'invariance par transformation affine, cela permet de garantir une robustesse par rapport aux transformations associées de la fonction objectif [10]. Deux classes d'algorithmes stochastiques qui sont envisagés pour cette thèse présentent ces invariances : les Stratégies d'Évolutions [11] et notamment l'algorithme CMA-ES [10], ainsi que les algorithmes à Estimation de Distribution avec en particulier l'algorithme de Cross Entropy [12].

En s'appuyant sur les travaux de thèse de Florian Maire (prix de thèse DGA 2016) sur la prise en compte simultanée des variabilités spectrale et spatiale des objets d'intérêt et du fond pour la détection d'anomalies [1], et sur les avancées récentes en optimisation numérique stochastique, le ou la doctorant(e) sera amené(e) à proposer une méthodologie d'optimisation sous incertitudes qui permette la prise en compte de contraintes physiques et qui soit adaptée aux variables mixtes. Cette problématique, cruciale pour le dimensionnement des futurs systèmes opérationnels multispectraux, n'est encore que très peu abordée dans la littérature. En effet, si on note depuis une dizaine d'années un essor des méthodes de sélection de bandes qui prennent en compte l'application visée (classification pour la plupart) pour effectuer le choix, il existe peu de travaux qui prennent en compte les regroupements de bandes spectrales adjacentes et très corrélées, et encore moins qui tiennent compte de la diversité des objets d'intérêt et du fond.

Différents critères ont été proposés pour l'optimisation robuste sous incertitudes [2-3] : espérance, variance, quantile, superquantile (qui permet de quantifier le poids des événements situés au-dessus d'un quantile donné), mais il faudra les adapter à la problématique du choix de bandes spectrales, qui est un problème d'optimisation avec des variables discrètes et des contraintes physiques de faisabilité technologique. Il faudra de plus coupler de façon pertinente ces critères de robustesse à ceux de détection. On considèrera le critère de référence, le RX [4] et des variantes tirant profit des tests robustes [5]. Le critère de détection retenu doit en effet être assez simple pour permettre l'optimisation sous incertitudes.

Au-delà du choix du critère pour l'optimisation robuste, il s'agira d'étudier le couplage critère / stratégie d'optimisation. On pourra s'inspirer des approches basées sur la méthode CMA-ES qui voient le problème d'optimisation robuste comme un problème min-max (minimisation du pire cas) et appliquées avec succès pour des problèmes de pilotage automatique de navires [7-9].

Dans un deuxième temps, le calcul de critères du type quantile nécessitant un grand nombre d'évaluations des objectifs afin d'observer l'effet des variations des paramètres incertains autour de chaque solution considérée, il faudra proposer des estimateurs efficaces adaptés à la problématique du choix de bandes spectrales. Ces estimateurs devront permettre de bien explorer la diversité des solutions optimales concurrentes potentielles. On proposera des adaptations d'algorithmes stochastiques : pénalisation, modification des lois de génération, pour prendre en compte des contraintes qui traduisent les informations a priori associées à la forme, le nombre et la largeur des bandes spectrales. Les algorithmes envisagés sont CMA-ES (Covariance Matrix Adaptation Evolution Strategy) ainsi que des évolutions récentes de méthodes de type cross-entropy associées à des familles de loi en réseaux bayésiens [6].

La méthodologie sera appliquée sur une base de données simulées d'avions sur fond de ciel contenant plusieurs dizaines de milliers d'images dans différentes bandes spectrales infrarouge [1] et sur des données hyperspectrales réelles qui servent de benchmark dans la littérature. Pour cela, on évaluera les performances des combinaisons de bandes spectrales optimisées pour des algorithmes de détection par apprentissage profond, avec des architectures de type YOLO ou ResNet et ses variantes couplant CNN et Transformers, qui sont à l'état de l'art pour la détection de petits objets en infrarouge.

Bibliographie :

- [1] F. Maire, S. Lefebvre, Detecting aircraft in low-resolution multispectral images: specification of relevant IR wavelength bands, *J. of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (9), 4509-4523, 2015.
- [2] V. Gabrel, C. Murat, A. Thiele, Recent advances in robust optimization: An overview, *European J of Operational Research* 235, 471-483, 2014.
- [3] R. T. Rockafellar, Coherent approaches to risk in optimization under uncertainty, *Tutorials in operations research*, 3, 38–61, 2007.
- [4] I.S. Reed et X. Yu, Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution, *IEEE Trans. Acoustics Speech, Signal Process.*, 38, 1760-1770, 1990.
- [5] P.J. Huber, "Robust statistics. In *International encyclopedia of statistical science*" (pp. 1248-1251). Springer, Berlin, Heidelberg, 2011;
- [6] C. Musso, F. Dambreville and C. Chahbazian, "Filtering and sensor optimization applied to angle-only navigation," 2021 IEEE 24th International Conference on Information Fusion (FUSION), 2021;
- [7] Akimoto, Youhei, Yoshiki Miyauchi, and Atsuo Maki. "Saddle point optimization with approximate minimization oracle and its application to robust berthing control." *ACM Transactions on Evolutionary Learning and Optimization* 2.1 (2022): 1-32.
- [8] Edo, H., Miyauchi, Y., Maki, A., & Akimoto, Y. (2023, July). Trade-off Between Robustness and Worst-Case Performance in Min-Max Optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 1339-1347).
- [9] Conn Andrew R. and Vicente Luis N.. 2012. Bilevel derivative-free optimization and its application to robust optimization. *Optimiz. Methods Softw.* 27, 3 (2012), 561–577.
- [10] Hansen, Nikolaus, and Anne Auger. "Principled design of continuous stochastic search: From theory to practice." *Theory and principled methods for the design of metaheuristics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. 145-180.

- [11] Hansen, Nikolaus, Dirk V. Arnold, and Anne Auger. "Evolution strategies." Springer handbook of computational intelligence (2015): 871-898.
- [12] Rubinstein, Reuven. "The cross-entropy method for combinatorial and continuous optimization." Methodology and computing in applied probability 1 (1999): 127-190.

Collaborations envisagées : Equipe RandOpt INRIA/ CMAP

Laboratoire d'accueil à l'ONERA

Département : Département Optique et Techniques associées

Lieu (centre ONERA) : Palaiseau

Contact : Sidonie Lefebvre / Frédéric Dambreville

Tél. : 01 80 38 63 76 Email : sidonie.lefebvre@onera.fr

Directeur de thèse

Nom : Auger Anne

Laboratoire : RandOpt INRIA/CMAP

Tél. :

Email : anne.auger@inria.fr

Pour plus d'informations : <https://www.onera.fr/rejoindre-onera/la-formation-par-la-recherche>