

Caractérisation d'objets sur des images acquises par drone recherchant des défauts pour la maintenance aéronautique

Julien Miranda¹

Stanislas Larnier¹

Matthieu Claybrough¹

¹Donecle, IoT Valley, 201 Rue Pierre et Marie Curie, 31670 Labège, France
{julien.miranda ; stanislas.larnier ; matthieu.claybrough}@donecle.com

Résumé

Nous nous proposons de décrire nos travaux relatifs à la détection et la reconnaissance d'objets sur des images acquises par drone automatisé pour la maintenance aéronautique. La solution proposée pour l'analyse des images distingue l'étape de détection des objets d'intérêt de celle de leur identification. L'article expose les méthodes employées pour reconnaître les éléments détectés, et les résultats obtenus par le système proposé.

1 Introduction

Afin de fiabiliser la maintenance aéronautique qui repose principalement sur des inspections visuelles, des solutions automatiques se sont développées avec des robots de surface [1] et des robots roulants [2]. Donecle développe des drones automatisés dédiés à l'inspection visuelle de la surface extérieure des avions. Le drone est capable de naviguer de manière autonome par rapport à l'avion en intérieur ou extérieur en employant un positionnement laser. Il a la capacité d'éviter les obstacles. Une caméra embarquée acquiert les images à analyser. Des algorithmes de traitement d'images permettent la détection et la catégorisation des défauts. L'ensemble de la chaîne d'acquisition et de traitement est illustrée par la Figure 1. Les images haute résolution acquises par le drone couvrent la surface extérieure de l'avion et sont complétées par des données issues des capteurs du drone (position dans l'espace, distance de la caméra à la surface, etc.).

Les inspections de ce type sont réalisées sur des avions de la famille A320, durent moins d'une heure et aboutissent à l'analyse de plusieurs centaines d'images, à l'issue de laquelle un nombre raisonnable de défauts sont présentés à l'opérateur pour validation.

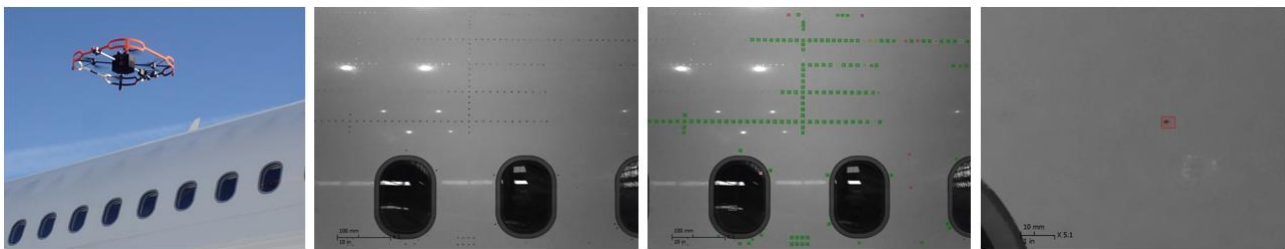


Figure 1: De gauche à droite : vol du drone, image acquise, extraction de régions d'intérêt, détection d'un défaut potentiel.

2 Extraction de régions d'intérêt

Le défi que représente la détection automatisée de défauts sur la surface d'un avion se traduit par plusieurs considérations dans le cadre de la vision par ordinateur : la taille des défauts (de plusieurs centimètres à quelques millimètres) et leur diversité (impact foudre, fuite, enfoncement) constituent des éléments délicats tant pour leur détection que pour leur identification. La Figure 2 présente des exemples de défauts ainsi que d'autres objets détectés (vis, rivet, marquage signalétique) naturellement présents sur la surface d'un avion.

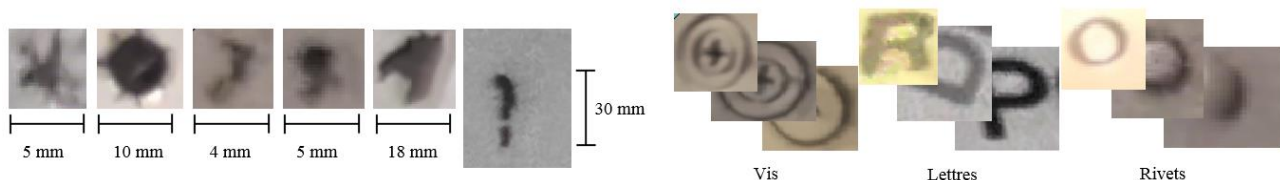


Figure 2: Exemples d'éléments détectés. A gauche : défauts et échelle, à droite : exemples d'autres objets annotés.

L'extraction de régions d'intérêt (ROI) est la première étape dans le processus de détection de défauts. Il faut détecter les zones de l'image susceptibles de contenir des éléments pertinents pour les opérations de maintenance. La méthode utilisée consiste à appliquer un seuillage adaptatif de l'intensité des pixels, avec une taille de fenêtre adaptée à la taille des objets recherchés.

Cette méthode est pertinente car on peut a priori donner un ordre de grandeur de la taille des défauts recherchés, ce qui donne une signification physique aux paramètres des algorithmes de détection. L'utilisation d'ortho-images permet

en outre de réécrire certains algorithmes de traitement d'image classiques afin d'introduire la modalité physique (notion de taille ou prise en compte d'autres données issues des capteurs embarqués sur chaque drone) sous forme de pondération.

3 Reconnaissance d'objets

Une fois les ROIs déterminées, l'enjeu majeur exposé est l'identification des éléments qui y sont contenus, au sein d'un ensemble d'objets connus (classes). Une approche pour la classification consiste à extraire des valeurs caractéristiques, les descripteurs, de ces sous images que constituent les ROIs. Ces valeurs sont utilisées comme données d'entrées par une machine à vecteurs de support (SVM) pour déterminer la classe de l'élément présent dans la région concernée. La mise en place d'un tel système implique une procédure d'annotation des données (indispensable lors la phase d'apprentissage) : cette étape est réalisée par les opérateurs de la maintenance aéronautique, experts dans la reconnaissance des éléments susceptibles d'apparaître à la surface d'un avion.

Le travail exposé dans cette section décrit quelques combinaisons de descripteurs classiques ainsi que le système prédictif par apprentissage qu'ils nourrissent pour l'application évoquée.

Pour le choix des descripteurs, une première approche consiste à décrire les éléments par valeurs correspondant à des notions intuitives pour un observateur humain : leur forme (convexité de l'élément, circularité, etc.), leur taille (diamètre, aire) ou leur couleur (couleur dominante, texture, etc.).

Ce type de descripteurs est cependant sensible aux variations de l'échelle, de l'angle de prise de vue de l'objet ou encore à la composition de la lumière. Deux solutions sont alors possibles pour dépasser cette limite : utiliser des descripteurs dits « invariants », ou coupler les données visuelles avec d'autres informations disponibles pouvant, par exemple caractériser la taille réelle des éléments observés, ou fournir des informations sur les conditions de luminosité lors de l'acquisition. La première option sera développée ici. Ainsi ce type de descripteurs, robustes à différentes perturbations par construction, ont été testés.

Les **moments de Hu** [3] constituent l'un des ensembles de descripteurs les plus utilisés pour l'analyse d'image. Ils sont construits à partir des moments de l'image, auxquels sont appliquées des transformations visant à les doter de propriétés d'invariance en translation, rotation et échelle. Pour chaque image, on obtient alors les sept moments de Hu.

La **transformée de Fourier-Mellin**, proche d'une transformée de Fourier double en coordonnées polaires, est un opérateur mathématique permettant de construire des familles de descripteurs invariants en translation, rotation et homothétie. Ainsi, à partir de la transformée \hat{f} , il est possible [4] de construire la famille suivante :

$$\Phi' = \left\{ \frac{|\hat{f}(1,1)|^2}{|\hat{f}(0,1)|^2}, \dots, \frac{|\hat{f}(2,1)|^2}{|\hat{f}(0,1)|^2}, \frac{|\hat{f}(2,2)|^2}{|\hat{f}(0,2)|^2}, \dots, \frac{|\hat{f}(K,V)|^2}{|\hat{f}(0,V)|^2} \right\}$$

Une approche répandue consiste à utiliser des descripteurs calculés à partir de points d'intérêt extraits de l'image. Cette dernière catégorie de descripteurs nécessite un paramétrage adapté pour la détection des points d'intérêt ainsi qu'une mise en forme des données pour l'interfaçage avec le système SVM utilisé, au moyen de sacs de mots visuels (*Bag Of Words*) [5].

Les descripteurs **SIFT** (*Scale Invariant Feature Transform*) présentent de nombreux avantages pour les applications envisagées, comme leur robustesse aux transformations affines, aux variations des conditions d'illumination et au bruit. Ils font partie des descripteurs produisant les meilleurs résultats dans les études comparatives [6].

Une autre famille de descripteurs repose sur les points d'intérêt : les **SURF** (*Speed Up Robust Features*) décrits dans [7]. Ces descripteurs sont obtenus par le calcul de la transformée en ondelette de Haar circulaire dans un rayon donné avec facteur d'échelle lié au point d'intérêt. Les opérations suivantes sont similaires à celles nécessaires au calcul de descripteurs de type SIFT. Les réponses en ondelettes sont ensuite pondérées par une gaussienne, et la matrice obtenue est alors divisée en sous régions.

Les deux familles de descripteurs basées sur l'extraction de points d'intérêts évoquées précédemment sont très utilisées. Il existe néanmoins une alternative récente : les descripteurs **ORB** (*Oriented FAST and Rotated BRIEF*) qui produisent des résultats comparables pour un temps de calcul réduit. Ce type de descripteur s'appuie sur la méthode FAST (*Features from Accelerated Segment Test*) permettant d'extraire les points d'intérêt, et procède au calcul de descripteurs en utilisant l'algorithme BRIEF (*Binary Robust Independent Elementary Features*) [8].

La démarche a été de définir la structure du système d'apprentissage, puis les combinaisons de descripteurs les plus adaptées au problème. Plusieurs familles de descripteurs et combinaisons ont été testées. Le paramétrage du classificateur SVM a été fait par validation croisée sur l'ensemble des combinaisons de descripteurs, avec l'utilisation d'un noyau gaussien pour l'obtention de l'espace de re-description.

(I)	Descripteurs simples	(IV)	SIFT	(VII)	I + II + III
(II)	Moments invariants de Hu	(V)	SURF	(VIII)	I+VI
(III)	Descripteurs de Fourier-Mellin	(VI)	ORB	(IX)	I+III+VI

4 Résultats et perspectives

Les tests ont été menés sur une base de données composée de plusieurs centaines d'éléments, répartis en cinq classes. Selon le principe de validation croisée, le jeu de données a été partitionné aléatoirement en jeux de données d'apprentissage et de validation pour tester les performances du système sur des éléments jamais rencontrés, et ce plusieurs fois à partir du jeu de données initial. Les résultats présentés dans le Tableau 1 ont été obtenus en validation.

Descripteurs	I	II	III	IV	V	VI	VII	VIII	IX
Taux de succès	35%	47%	67%	72,3%	69,4%	61%	59,4%	66,5%	86%

Tableau 1: Taux de succès en validation pour l'identification des objets (5 classes, 1654 éléments)

Ces résultats permettent de valider l'utilisation combinée de plusieurs types de descripteurs dans notre cadre d'application : les résultats de ces combinaisons dépassent l'utilisation séparée de chacun. De plus, les meilleurs résultats sont obtenus par combinaison de descripteurs n'ayant pas de construction commune, ce qui montre leur complémentarité.

Un tel modèle prédictif utilisant des combinaisons de descripteurs est ainsi déjà utilisé par les opérateurs de la maintenance aéronautique utilisant la solution Donecle pour détecter et cartographier les défauts potentiels en générant des rapports d'inspection, suivant le modèle de l'extrait en Figure 3.




Position	Context	Object	Result
 FR67 STGR08R		 13 x 13 mm	Def (Impact)

Figure 3: Extrait d'un rapport d'inspection généré automatiquement

Nous avons ici présenté un procédé d'analyse d'image basé sur la séparation des opérations de détection et des opérations de reconnaissance des objets. Cette approche permet une modularité des deux actions, mais prive ces modules d'informations pourtant pertinentes : contexte pour la reconnaissance d'objet et nature de l'objet pour la détection. De nombreuses approches utilisées actuellement, en particulier celles basées sur l'utilisation de réseaux de neurones profonds, proposent un système permettant une porosité entre les deux actions. Néanmoins ces approches sont souvent exclusivement basées sur les données (principe de l'émergence d'un modèle) et en sont donc dépendantes.

La notion d'ortho image constitue une approche originale : chaque acquisition se compose d'une photo accompagnée de nombreuses données complémentaires (position de la camera par rapport au modèle, luminosité etc.). En exploitant cette richesse, il sera possible d'aller plus loin en ce sens afin d'introduire les données physiques comme une réelle modalité supplémentaire, associée à l'image en tant qu'objet visuel. En effet, si l'invariance des descripteurs à l'échelle gomme toute notion de taille physique des objets dans l'image, certains objets peuvent au contraire être caractérisés grâce à leurs dimensions (impacts foudre de taille très réduite, ou éléments de fixation de taille précisément connue. Cette piste, dont l'exploration sera rendue possible par la forte intégration de la chaîne d'acquisition et de traitement, pourra permettre d'accroître les performances, la fiabilité et l'évolutivité de la solution proposée.

Bibliographie

- [1] M. Siegel, P. Gunatilake and G. Podnar, "Robotic assistants for aircraft inspectors," *IEEE Instrumentation & Measurement Magazine*, vol. 1, no. 1, pp. 16-30, 1998.
- [2] I. Jovancevic, S. Larnier, J.-J. Orteu and T. Sentenac, "Automated exterior inspection of an aircraft with a pan-tilt-zoom camera mounted on a mobile robot," *Journal of Electronic Imaging*, vol. 24, no. 6, 2015.
- [3] M. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, pp. 179-187, 1962.
- [4] Y. Sheng and L. Shen, "Orthogonal Fourier-Mellin moments for invariant pattern recognition," *Journal of Optical Society of America*, pp. 1748-1757, 1994.
- [5] C. Grana, B. Daniele and C. Rita, "A Fast Approach for Integrating ORB Descriptors in the Bag of Words Model," Modena, Italy, 2013.
- [6] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the International Conference on Computer Vision*, 1999.
- [7] H. Bay, A. Ess, T. Tuytelaars and L. V. Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008.
- [8] E. Rublee, V. Rabaud, K. Konolige and G. Bradski., "ORB: an efficient alternative to SIFT or SURF," in *IEEE International Conference on Computer Vision*, 2011.

N.B. : les performances indiquées ne sont pas représentative de la solution commercialisée.