



**HAL**  
open science

# Extraction d'information contextuelle d'images et de séquences d'images pour en proposer une visualisation pertinente à un utilisateur dédié

Sébastien Mavromatis

► **To cite this version:**

Sébastien Mavromatis. Extraction d'information contextuelle d'images et de séquences d'images pour en proposer une visualisation pertinente à un utilisateur dédié. Vision par ordinateur et reconnaissance de formes [cs.CV]. Aix Marseille University, 2019. tel-02465825

**HAL Id: tel-02465825**

**<https://amu.hal.science/tel-02465825v1>**

Submitted on 4 Feb 2020

**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.

## **Habilitation à Diriger les Recherches**

rédigée et présentée par

**Sébastien MAVROMATIS**

---

### **Extraction d'information contextuelle d'images et de séquences d'images pour en proposer une visualisation pertinente à un utilisateur dédié**

---

**Soutenance le mercredi 11 septembre 2019**

#### **Composition du jury :**

Christine Fernandez-Maloigne	Université de Poitiers / Xlim	Rapporteuse
Franck Marzani	Université de Bourgogne / ImViA	Rapporteur
Jean Meunier	Université de Montréal / DIRO	Rapporteur
Saïda Bouakaz	Université de Lyon (Claude Bernard) / LIRIS	Examinatrice
Valérie Burdin	IMT Bretagne / ITI	Examinatrice
Mohamed Quafafou	Aix Marseille Université / LIS	Examineur
Jean Sequeira	Aix Marseille Université / LIS	Tuteur

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Curriculum Vitæ</b>	<b>8</b>
1.1 État Civil	8
1.2 Diplôme	9
1.3 Responsabilités collectives	10
1.4 Enseignement	11
1.4.1 Investissement	11
1.4.2 Service d'enseignement	12
1.4.3 Projet pédagogique	12
1.5 Encadrement de thèses de doctorat	16
1.5.1 Thèses soutenues	16
1.5.2 Thèses en cours	22
1.5.3 Devenir des docteurs	23
1.5.4 Encadrement de stages au Master Recherche SIS	24
<b>2 Projets de recherche</b>	<b>25</b>
2.1 Projets structurants et pérennes	25
2.1.1 Reconstruction 3D de scènes à partir de séquences vidéo	25
2.1.2 Analyse et recalage d'images de modalités différentes	27
2.1.3 Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef	28
2.2 Projets spécifiques développés dans un cadre industriel	30
2.2.1 Rendu des inter-réflexions diffuses dans le cadre d'un scénario « walk-through »	30
2.2.2 Caméra Intelligente pour la Surveillance Autonome	31
2.2.3 Détection d'objets dans des images acquises par un sonar multi-faisceaux	32
2.2.4 Caractérisation du mouvement des yeux et de la tête lorsque celle-ci est soumise à des sollicitations externes	33
2.3 Projets développés dans un cadre international	35
2.3.1 Portugal	35
2.3.2 Burkina Faso	35
2.3.3 Chine	36

2.3.4	Algérie . . . . .	36
2.3.5	Pérou . . . . .	36
2.4	Projet émergent . . . . .	37
<b>3</b>	<b>Valorisation de la Recherche</b>	<b>40</b>
3.1	Brevet . . . . .	40
3.2	Licence logicielle . . . . .	41
3.3	Maturation de projet . . . . .	41
3.3.1	SimulFoot . . . . .	41
3.3.2	PhotoMap . . . . .	42
<b>4</b>	<b>Quelques travaux de recherche significatifs</b>	<b>44</b>
4.1	Reconstruction 3D de scènes à partir de séquences vidéo . . . . .	45
4.2	Analyse et recalage d'images de modalités différentes . . . . .	62
4.3	Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef . . . . .	75
<b>5</b>	<b>Projet scientifique</b>	<b>100</b>
	<b>Publications</b>	<b>104</b>
	Revue, Livres et Brevets . . . . .	104
	Conférences . . . . .	105
	Thèse . . . . .	108
<b>6</b>	<b>Annexes</b>	<b>109</b>
	<b>Annexe A How Do Expert Soccer Players Encode Visual Information to Make Decisions in Simulated Game Situations?</b>	<b>110</b>
	<b>Annexe B Method and system for estimating a similarity between two binary images</b>	<b>118</b>
	<b>Annexe C Embedding image processing intelligent algorithms</b>	<b>133</b>

# Introduction

Ce mémoire présente les travaux de recherche que j'ai développés au cours de ces vingt dernières années dans le domaine de l'image, essentiellement en analyse d'images et en reconnaissance des formes, mais aussi dans le domaine de la visualisation scientifique, en particulier dans le cadre immersif de la « Réalité Virtuelle » et de la « Réalité Augmentée ».

Plus précisément, je me suis intéressé à l'extraction de l'information dans les images et à la communication de cette information de manière pertinente et ergonomique à un utilisateur. Comme nous aurons l'occasion de le voir par la suite, cela ne saurait se faire de manière cohérente sans aborder la représentation de la connaissance, qui constituera le support permettant de véhiculer cette information, de la capitaliser et de l'utiliser pour enrichir cette connaissance. Cela a constitué le fil conducteur de mes travaux et se trouve également, dans un cadre nouveau et plus large, à l'origine du projet scientifique que je souhaite développer au cours de ces prochaines années.

Mes premiers travaux de recherche ont été développés au cours de ma thèse, soutenue en fin 2001, et ont porté sur l'analyse de la texture. D'une certaine manière, cela s'apparentait fortement au problème posé ainsi « comment peut-on extraire, dans une image, de l'information d'une région présentant une certaine cohérence ? » et qui serait alors suivi de la question « comment peut-on utiliser cette information pour reconnaître une entité que je souhaite détecter dans cette image ? ». Au-delà des théories et des algorithmes permettant d'analyser les images et d'y reconnaître des « formes » (au sens large), il s'agit tout d'abord de définir le cadre dans lequel cela s'inscrit car, même si c'est une évidence, nous n'aborderons pas la résolution d'un problème de la même manière suivant que nous voulons découvrir la présence d'une pathologie dans une image médicale, évaluer l'évolution de l'urbanisme à partir d'images satellite, ou caractériser une activité au moyen d'une caméra de vidéosurveillance. Cela m'a donc amené très tôt à m'intéresser à ce cadre qui peut être associé à ce que l'on appelle « la représentation de la connaissance », et que j'ai essayé de mettre en place dans chacun des projets de recherche que j'ai mis en place par la suite.

De ce fait, dès ma prise de fonction comme Maître de Conférences à l'Université de la Méditerranée, en 2003, tout en participant à différents projets de recherche, en codirigeant des travaux de thèses, et en mettant en place des projets financés, j'ai toujours essayé – même si je n'ai pas, à proprement parler, développé de recherche sur ce sujet – de mettre en place un

cadre de « représentation de la connaissance » pour mieux y situer mes travaux.

Les concepts de données, d'information et de connaissance font l'objet de discussions souvent « vives » et mon but n'est absolument pas de prendre part à ce débat, d'autant plus que – comme je l'ai dit précédemment – je n'en suis pas un spécialiste et n'ai fait aucun travail de recherche spécifique dans ce domaine. Mon but est simplement de donner un cadre scientifique cohérent et « productif » à l'ensemble de mes travaux, qu'ils soient passés ou futurs.

S'il n'y a pas de difficulté à définir « les données » – on peut dire qu'il s'agit du « matériau brut dont on dispose » - cela peut être un peu plus délicat pour ce qui concerne les concepts d'information et de connaissance. Je me suis donc appuyé sur des travaux scientifiques classiques et reconnus pour ce qui est de « l'information » et également sur des travaux scientifiques, mais aussi philosophiques pour ce qui est de « la connaissance ».

Le XX<sup>ième</sup> siècle a vu le développement de la théorie de l'information avec en particulier les travaux de Renyi, de Kolmogorov et de Shannon, et bien sûr, la définition de l'entropie (même si ce terme a été introduit un peu plus tôt – en 1865 – dans un cadre davantage lié à la physique par Rudolf Clausius). D'une certaine manière, l'information peut être définie comme étant la « représentation minimale » (ou l'essence) d'un ensemble de données, indépendamment de ce que ces données sont supposées représenter. La frontière avec « la connaissance » n'est pas aussi nette qu'elle ne le paraît et on parle alors « d'information contextuelle » lorsqu'on ne s'intéresse qu'à la partie de l'information relative à domaine spécifique.

En ce qui concerne « la connaissance », je m'appuierai fortement sur la vision philosophique classique de ce concept (telle qu'elle est apparue avec Berkeley, puis Hume, Kant, Schopenhauer, ... et « l'école allemande » qui l'a associée au « Weltanschauung » ou « représentation du monde »). Cette vision du concept de connaissance se trouve également au centre d'un domaine des sciences de l'information, qui s'intitule « l'intelligence compétitive » et qui aborde des problèmes dont la complexité est liée à l'hétérogénéité des éléments du système dans lesquels on souhaite les résoudre. Pour « définir » la connaissance, je reprendrai l'énoncé de Markus Gabriel, philosophe allemand, professeur à l'Université de Californie Berkeley : la connaissance est la représentation du monde, le monde étant tout ce qui « existe », à la fois de manière tangible (et qu'on appelle l'univers) et de manière intangible.

Pour ce qui nous concerne, dans la composante « tangible » du monde, il y a bien sûr tous les éléments physiques entrant en jeu dans le problème à traiter, et qu'il faut définir précisément en fonction du contexte (on retrouve ici les notions d'espace de représentation et d'espace d'interprétation de la « reconnaissance des formes » - dont la définition est capitale pour déterminer la manière d'aborder le problème, de choisir les algorithmes et même les théories à utiliser). Mais il y a aussi la composante intangible, et qui est tout aussi importante : il s'agit par exemple du savoir-faire de l'expert, de l'ergonomie, des capacités perceptuelles de l'utilisateur ...

Sans être en aucune manière un spécialiste du domaine, cette manière d'aborder les problèmes m'a permis d'intégrer dans mes travaux de recherche, en particulier dans le cadre de projets comportant des développements effectués avec des partenaires industriels, des éléments qui m'ont permis de faire émerger des solutions davantage en adéquation avec les besoins, exprimés ou pas. J'essaierai, tout au long de ce mémoire, de donner des liens sur cette problématique pour chacun des travaux de recherche.

Cela m'a permis également de définir un nouveau projet global de recherche qui, tout en étant dans la continuité de ce que j'ai fait au cours de ces vingt dernières années, établit un tremplin vers les domaines du « big data » et du « machine learning », et en particulier du « deep learning ». Mais tout cela sera précisé par la suite.

La période où j'ai commencé à travailler sur l'image (milieu des années 90) correspond à la période où l'image s'est « démocratisée ». Je pourrais exprimer cela différemment en disant que jusqu'au milieu des années 90, les ordinateurs qui traitaient ou produisaient les images étaient onéreux et « peu puissants » (aussi bien en mémoire qu'en capacité de traitement) et les images (les données dans le cas de l'analyse d'images) étaient également onéreuses et pas toujours facilement accessibles. Néanmoins, le socle des théories et des algorithmes associés était déjà bien établi. D'une certaine manière, l'accessibilité au plus grand nombre des moyens matériels dans ce domaine a ouvert le champ à de nombreuses applications pour lesquelles la résolution des problèmes passait par la maîtrise des théories et des algorithmes existants, mais aussi, et surtout, par la manière de poser ces problèmes, puis de choisir et d'agencer les algorithmes pertinents pour les résoudre (même si, bien évidemment, il fallait améliorer ces algorithmes, en trouver de nouveaux...). Cela est, s'il était nécessaire de le préciser, la raison pour laquelle j'ai attaché une certaine importance à la manière de poser le problème, et donc de représenter la connaissance relative à ce problème.

### Des travaux de recherche et un fil conducteur

Les travaux de recherche que j'ai développés au cours de ces vingt dernières années ont été prioritairement axés sur les aspects théoriques (au niveau informatique) des problèmes à résoudre, ce qui est normal pour une activité de recherche en informatique. Néanmoins, j'ai attaché une importance non négligeable aux composantes « développements » et « applications » de ces problèmes, et ce en relation avec leur composante théorique : les développements permettent d'évaluer quantitativement les avancées théoriques, et ils fournissent un support de réflexion, mais aussi de dialogue avec les experts du domaine ; quant aux applications, elles permettent d'appréhender l'environnement du problème à résoudre pour pouvoir être « en résonance » avec les utilisateurs et ainsi, avancer de manière utile et efficace. Ce travail de compréhension puis de représentation de la connaissance est souvent difficile à effectuer (de par l'hétérogénéité des éléments à prendre en compte), mais il permet de mettre en place un cadre efficace pour aborder les questions théoriques.

Dans les prochaines pages, je ferai état, assez brièvement, de ce lien à ce « fil conducteur » à travers une vision chronologique de certains de mes travaux, de manière à faire apparaître

la cohérence globale de ces derniers. Par la suite, je reprendrai l'ensemble de mes travaux et les présenterai dans un cadre quantitatif et exhaustif (objectifs du projet, partenaires, financements, résultats, publications et brevets ...).

J'ai tout d'abord travaillé, au cours de ma thèse (soutenue en 2001), sur l'analyse de la texture et ma contribution a essentiellement consisté à approfondir et à prolonger (avec une dimension « directionnelle ») le concept « d'extrema locaux d'ordre ». Cette avancée a permis une discrimination plus efficace, et surtout plus rapide, de textures visuellement très proches. J'ai également développé une approche de visualisation scientifique permettant d'interagir avec l'espace de texture : cette composante de mon travail n'était peut-être pas des plus classiques (habituellement, on s'intéresse davantage à la pertinence des opérateurs fournissant des indices discriminants dans l'espace de représentation de la texture), mais c'est certainement celle qui m'a le plus passionné. En effet, on peut représenter « informatiquement » la texture (à travers un ensemble d'indices associés à un espace de représentation, dans la problématique classique de la « Reconnaissance des Formes »), mais le seul critère permettant de décider si deux textures sont différentes est le critère « visuel », et à ce niveau-là, l'introduction de l'être humain dans la « boucle » de « perception – décision » devient incontournable. C'est lorsque j'ai véritablement compris l'importance de cette « boucle » que je me suis focalisé sur ces aspects « visualisation » et « interaction ».

Le projet Simulfoot illustre tout particulièrement l'intérêt que j'ai porté, et que je porte toujours, dans le cadre de l'analyse d'images et de la visualisation scientifique, à l'analyse de la connaissance chez les experts du domaine abordé et à la représentation de cette connaissance au niveau informatique. J'ai participé à la création de ce projet en l'an 2000, j'ai développé des travaux de recherche et j'ai codirigé deux thèses dans le cadre de ce projet, et j'ai également participé à l'encadrement du développement d'un logiciel destiné à la commercialisation par la cellule de valorisation d'Aix-Marseille Université (ValorPaca puis SATT Sud-Est), et ce jusqu'en 2016. L'objectif de ce projet était de fournir aux entraîneurs, joueurs et arbitres de football, un outil interactif leur permettant, à partir de situations réelles (séquences vidéo de match) ou imaginées (schémas tactiques), d'observer et d'analyser des séquences de jeu, puis d'effectuer des simulations, le tout en se plaçant à divers endroits du terrain (à la place de joueurs, à la place de l'entraîneur, au-dessus du terrain ...) pour comprendre les prises de décisions des joueurs (et les aider à s'améliorer dans ce domaine). Les travaux scientifiques que j'ai développés pour le projet Simulfoot ont porté sur l'analyse de la couleur, la segmentation d'images, la reconnaissance de configurations (arrangement de lignes), la détection d'ellipses, le suivi (tracking) et l'interaction avec l'utilisateur (en lien avec le domaine de l'ergonomie cognitive).

En fait, le thème qui m'a le plus intéressé ici est celui des relations, au niveau perceptuel, entre la couleur et l'espace de représentation. Lorsqu'un spectateur regarde un match de football, il reconnaît immédiatement le terrain, et ce quel que soit son « état » : qu'il y ait ou non des joueurs, voire même d'autres éléments (orchestre, banderoles publicitaires ...), qu'il y ait de la pelouse ou pas, que les lignes soient bien visibles ou pas, et même lorsqu'il existe des zones



fortement contrastées (partie dans l'ombre et partie en plein soleil). Le schéma perceptuel du spectateur est intéressant à analyser et à essayer de comprendre pour pouvoir le « reproduire » afin de caractériser automatiquement la zone de jeu (car c'est là, et seulement là, que les joueurs seront détectés, suivis ...). Le spectateur tient compte à la fois de cohérence spatiale et de la cohérence « couleur », mais que signifie la cohérence « couleur » lorsqu'on associe un pixel vert clair (pelouse dans la lumière), proche du blanc (ou du turquoise) à un pixel vert foncé (pelouse dans l'ombre), proche du noir (ou du bleu marine)? L'idée que j'ai développée (pour l'analyse de l'espace « couleur », sachant j'ai travaillé également sur sa relation avec le domaine spatial) a consisté à développer une fonction « potentiel » 3D (dans l'espace couleur HLS) s'appuyant sur la distribution des pixels (dans cet espace) et à chercher des « chemins pertinents » reliant certaines régions de cet espace à travers cette fonction potentielle. Cette idée peut être généralisée et approfondie pour analyser « l'Espace de Représentation » en « Reconnaissance des Formes » et cela fait également partie de mon futur projet scientifique.

En 2008, j'ai participé à la mise en place d'un partenariat avec l'Académie Chinoise des Sciences dans le domaine de la Télédétection, et je m'y suis impliqué activement au cours de ces dix dernières années. Ce projet, formalisé par une convention entre l'IRSA (Institute of Remote Sensing Applications) à Pékin et l'Université de la Méditerranée à Marseille, nous a permis de collaborer avec cet institut dans le domaine de l'image, et de chercher des solutions à divers problèmes d'analyse d'images dans le cadre d'applications effectives dans les domaines des risques naturels (séismes), de l'urbanisme (indices de végétation) et de l'agriculture (contrôle des cultures). L'un des principaux problèmes que nous avons abordés et dans lequel je me suis particulièrement investi est celui du recalage d'images de modalités différentes, que l'on peut inclure dans un cadre plus global qui est celui de l'évaluation de la similarité entre deux régions situées dans des images de modalités différentes. Ces travaux, qui ont fait l'objet d'une thèse que j'ai codirigée, ont été reconnus sous la forme d'un brevet international intitulé « Similar Imaging ». Je vais décrire dans les prochaines lignes la problématique posée et son intérêt au niveau applicatif, les idées fortes sous-tendant ces travaux de recherche et la solution mise en œuvre.

Comme je l'ai indiqué au début de cette introduction, nous pouvons à présent disposer de volumes considérables de données, et c'est particulièrement le cas en télédétection. Mais un gros volume de données ne draine pas en général un gros volume d'information; en fait, le plus souvent, des volumes de données importants permettent d'extraire un complément d'information, complément souvent minime, mais qui peut se révéler déterminant. C'est la raison pour laquelle on cherche à « capitaliser » l'information au sein d'un SIG (Système d'Information Géographique), et ce en associant des images à un élément donné du SIG et en traitant ces images pour en extraire l'information. De nombreux algorithmes ont été développés pour mettre en correspondance (recaler) ces différentes images lorsqu'elles représentent le même type de données (par exemple la réponse multispectrale d'un terrain sur un même nombre de bandes), mais ce n'est plus le cas lorsque ces données sont de modalités différentes. J'ai donc travaillé sur la conception d'un tel algorithme.

Les pixels de deux images de modalités différentes (par exemple, une image optique multispectrale et une image radar) ne contiennent pas le même type d'information (c.-à-d. la même information contextuelle); en conséquence, il n'est pas possible, dans un cadre général, d'utiliser directement le contenu des pixels pour effectuer le recalage. En revanche, la distribution spatiale de l'information dans chacune des images peut, dans certains cas, être utilisée comme « élément d'accrochage » entre ces deux images, d'autant plus que même si l'information renvoyée par chacune des images est différente, la « base » dont sont issues les images est la même (le terrain dans le cas d'imagerie satellitaire, les structures anatomiques en imagerie médicale...). Cela nous a amené à penser qu'une mise en correspondance des images segmentées, ou plus exactement, des images des contours devrait permettre d'effectuer ce recalage : cela est partiellement vrai puisqu'en regardant ces images des contours, nous sommes capables (au niveau perceptuel) de les mettre en correspondance, mais nous devons constater que nombre des contours dans les deux images sont différents. La question qui s'est posée alors a été : comment sommes-nous capables (toujours au niveau perceptuel) de séparer les « contours analogues » des « contours différents » pour les mettre en correspondance. Après avoir réfléchi à cette question, et y avoir répondu, nous avons défini des primitives représentant au mieux les éléments de connaissance se trouvant dans la réponse à la question et de mettre en œuvre un processus analogue à celui que nous développons « personnellement ». Mon objectif, ici, n'était pas de décrire un algorithme original, mais de montrer le processus qui nous a permis de concevoir un tel algorithme, toujours en se référant à l'appréhension de la « connaissance », comme support de « modélisation » permettant d'extraire l'information des images.

J'illustrerai une dernière fois cette démarche, ou cet intérêt dans la manière d'analyser, de prendre en compte et de représenter tous les éléments d'un problème (y compris ceux qui sont intangibles), à travers la description d'un problème posé dans le cadre d'une collaboration avec la société Airbus Helicopters. Il y a quatre ans de cela, nous avons mis en place une collaboration avec cette société. Dans ce cadre, une thèse a été soutenue en février 2019 et je codirige actuellement deux thèses en convention CIFRE. La première thèse avait pour objectif d'apporter une aide au pilote en vue de son atterrissage sur une plateforme dédiée (qu'on appelle hélipad). Les hélipads doivent satisfaire certaines contraintes de forme, de motif et de couleur, mais cela n'est pas toujours le cas (même si une norme a été définie et est souvent respectée), et surtout, ils peuvent être situés dans des environnements différents (très dépouillés ou au contraire très bruités).

Un hélipad est constitué d'une plateforme octogonale, à l'intérieur de laquelle se trouve un cercle (ou plutôt une couronne circulaire) qui entoure un H. Habituellement, la plate-forme est verte, le cercle est jaune et le H est blanc. Détecter une telle plate-forme si elle respecte toutes les normes et si elle est située dans un environnement uniforme est un problème très simple d'analyse d'images, mais dans ce cas-là, le pilote n'a pas véritablement besoin d'aide! En revanche, si la plateforme se trouve en environnement urbain (ou complexe), s'il y a d'autres éléments lui ressemblant ou si elle est partiellement occultée, cela est beaucoup plus difficile à résoudre, et, naturellement, ces situations sont également à prendre en compte.

De même, on doit être capable de détecter un tel hélipad aussi bien lorsqu'il se trouve au loin que quand il est assez près.

La recherche d'un algorithme pour résoudre ce problème de « détection d'hélipad » passe par la recherche d'invariants suffisamment robustes pour prendre en compte toutes ces situations. Pour cela, nous nous sommes inspirés, une nouvelle fois, du « modèle perceptuel humain ». Que se passe-t-il lorsque nous observons une scène dans laquelle un hélipad peut se trouver ? Nous balayons la scène à différents niveaux de résolution, en cherchant si une telle « forme » (au sens général de la Reconnaissance des Formes) apparaît de manière évidente ou s'il faut la rechercher sous une forme plus petite. Lorsque nous effectuons cette recherche, pour chaque fenêtre, nous essayons de détecter des éléments caractéristiques nous permettant de décider si on est sûr qu'il n'y a pas d'hélipad dans la fenêtre, ou s'il existe une possibilité qu'il y en ait un. Enfin, sur les fenêtres restantes (où il y a une possibilité d'hélipad), nous mettons en œuvre tout un ensemble de tests permettant d'éliminer cette possibilité ou au contraire de caractériser la présence d'un hélipad et de donner sa position). La dernière question en suspens était de trouver un test qui permette, dans une fenêtre donnée, de ne pas risquer de « manquer » un hélipad tout en étant assez sûr de ne pas avoir trop de faux positifs. Après une analyse approfondie, nous avons décidé que cet invariant suffisamment robuste aux variations des conditions et des paramètres de la plate-forme était un arc d'ellipse (le cercle « épais » entourant le H se transforme en ellipse par transformation projective, et s'il est partiellement occulté, il devient un arc d'ellipse). Ceci étant posé, nous avons étudié les différents algorithmes de détection d'arc d'ellipse, ce qui nous a amené à en concevoir un autre, plus adapté au problème posé.

La liste des projets que j'ai mis en œuvre, que j'ai gérés, et dans lesquels je me suis fortement impliqué est beaucoup plus importante que ce que j'ai pu évoquer précédemment ; et sur ces quatre projets, du reste, je n'ai donné aucun détail ni aucun élément d'ordre scientifique. Tout cela sera décrit par la suite. Mon intention, dans cette introduction, était simplement de faire état, à travers quatre exemples caractéristiques, de ce qui m'a motivé et a constitué le fil conducteur sur lequel sont venus s'articuler les travaux de recherche que j'ai développés.

Dans les toutes prochaines pages, je présenterai mon activité professionnelle (recherche, enseignement et administration) au cours des vingt dernières années (sous forme de « CV détaillé »). Cela sera suivi par le détail de tous les projets auxquels j'ai participé, en précisant tous les éléments (rôle joué, travaux de recherche, financements, publications . . .). Et enfin, je présenterai mon futur projet scientifique, qui se trouve dans la continuité de ce que j'ai fait jusqu'à présent, avec une composante « apprentissage » liée à l'autre versant de la représentation de la connaissance. Quelques articles caractéristiques seront également mis en copie à la fin de ce manuscrit.

# 1 Curriculum Vitæ

## 1.1 État Civil

**Sébastien MAVROMATIS**

Né le 20 août 1971

Nationalité française

Marié, six enfants

Maître de conférences depuis septembre 2003  
au Laboratoire d'Informatique et Systèmes (LIS - UMR 7020)

**Adresse professionnelle :**

Polytech Marseille / LIS - équipe Mod

Case 925

Campus de Luminy

13288 MARSEILLE cedex 09

Tél : +33(0) 491 828 550

Fax : +33(0) 491 828 551

Mob : +33(0) 625 832 398

**Web :** <http://www.lis-lab.fr>

**Adresse e-mail :** [sebastien.mavromatis@univ-amu.fr](mailto:sebastien.mavromatis@univ-amu.fr)

**Pages personnelles :** <http://sebastien.mavromatis.p.luminy.univ-amu.fr>

## 1.2 Diplôme

**Thèse de doctorat** de l'Université de la Méditerranée, mention mathématiques et informatique, obtenue avec une mention très honorable et félicitations du jury.

**Sujet :** Analyse de texture et visualisation scientifique

**Soutenu le :** 14 décembre 2001

**Jury composé de :**

- Françoise Prêteux,
- Georges Stamon,
- Rémy Bulot,
- Jean-Christophe Olivo Marin,
- Hubert Ripoll,
- Jean Sequeira,
- Jean-Marc Boï .

**Mots-clés :** analyse de texture, visualisation, segmentation.

**Résumé :** Ce travail introduit la notion d'extrema d'ordre directionnel et montre comment une texture peut-être représentée de manière robuste et efficace par une surface de type  $z = f(x, y)$ . Le principe consiste à attribuer à chaque pixel des caractéristiques texturales et directionnelles de la région centrée en celui-ci. Les définitions de suite *a priori* non régulière de voisinages, d'ordre limite, d'extremum local d'ordre  $k$ , d'extremum local d'ordre maximum  $k$  sont introduites. Ces concepts sont enrichis par les notions de direction et de multiplicité qui permettent de définir un extremum local d'ordre maximum directionnel et un extremum local d'ordre maximum multiple.

A l'aide de ces notions, il est possible de caractériser les propriétés texturales d'une image. Cette caractérisation présente l'avantage de permettre à un expert d'interagir efficacement avec le processus de segmentation. Cette ouverture est importante car la segmentation d'une image dépend largement de l'application dans laquelle elle s'intègre, et il est donc indispensable que l'expert puisse conserver un contrôle sur celle-ci. Ceci a apporté des solutions, parfois classiques, à tous les problèmes posés dans ce cadre. Cela a permis de montrer la faisabilité de cette approche, en même temps que de souligner l'intérêt qu'elle présente.

**Aspects originaux :**

- Définition du concept d'extremum local d'ordre maximum multiple
- Extraction des attributs directionnels des textures
- Interaction avec l'espace de texture
- Outils de visualisation scientifique

### 1.3 Responsabilités collectives

- **Responsable de la 5<sup>e</sup> année** de la formation d'ingénieur en informatique de Polytech Marseille - depuis 2018.
- **Responsable de la spécialité ReVA** (Réalité virtuelle et Applications) de la formation d'ingénieur en informatique de Polytech Marseille - depuis 2015.
- **Membre de l'équipe pédagogique** du parcours GIG (Géométrie et Informatique Graphique) du master informatique de la Faculté des Sciences - depuis 2018.

#### Auparavant

- **Représentant de Polytech Marseille pour le projet Henri Fabre**<sup>1</sup>  
Echanges sur la mise en place d'une offre de formation autour de la réalité virtuelle et la réalité augmentée - 2016 et 2017.
- **Co-responsable de la spécialité MIRA** (Multimédia, Image, Réalité virtuelle et Applications) à Polytech Marseille - de 2010 à 2015
- **Responsable de l'équipe SimGraph du laboratoire LSIS** (UMR CNRS 7296) - en 2014.
- **Chargé de la formation continue à Polytech Marseille.** Représentant de l'école au sein de la commission VAE de l'Université - de 2008 à 2013.
- **Membre de l'équipe pédagogique** de l'option « Imagerie Numérique » du Master SIS Pro - 2004 à 2011.
- **Responsable du parcours « Image » du Master Recherche SIS** - 2005 à 2010.
- **Membre de comités de sélection** à l'Université de la Méditerranée : en 2009, poste ESIL MCF 0770; et en 2010, poste ESIL MCF 1575.
- **Responsable de la spécialité « Imagerie Numérique »** à l'ESIL - 2004 à 2010.
- **Membre de jurys de thèse**
  - Jinnian Wang, Caractérisation des sols par l'analyse d'images hyperspectrales en télédétection. Jury : Gilles Gesquiere, Franck Marzani, Sebastien Mavromatis, William Puech, Jean Sequeira, Sophie Viseur - octobre 2014.
  - Yuxi Wang, Recalage d'images de modalités et de géométries différentes. Jury : Renyi Liu, Qinjiu Tian, Zetian Ye, Xingfa Gu, Sébastien Mavromatis, Jean Sequeira - octobre 2012.

---

1. Le projet Henri-Fabre est un projet public-privé permettant aux partenaires de mettre en synergie leurs expertises, leurs ressources, leurs compétences, leurs projets et de structurer un réseau avec l'ensemble des acteurs de l'innovation et de la recherche. Le projet se déploie sur les filières de l'aéronautique, de l'énergie, du médical et du naval. - <http://www.projet-henri-fabre.com>

— Xiaomin Che, Bassins versants et simulation de crues. Jury : Samir Akkouche, Jean-Marc Boï, Denis Friboulet, Jean Pierre Jessel, Sébastien Mavromatis, Jean Sequeira - novembre 2005.

## 1.4 Enseignement

### 1.4.1 Investissement

J'effectue l'essentiel de mon service d'enseignement à Polytech Marseille (anciennement ESIL, Ecole Supérieure d'Ingénieurs de Luminy avant la fusion des écoles d'ingénieurs d'Aix Marseille Université). Lorsque j'ai été recruté en 2003, la formation dans le domaine de « l'Imagerie Numérique » existait sous la forme d'une « coloration » en cinquième année pour l'équivalent de 80h équivalent TD. J'ai dans un premier temps créé plusieurs enseignements qui sont venus compléter la formation. En 2010, j'ai créé une option de 5<sup>e</sup> année qui s'articulait autour du multimédia et de l'image MIRA (Multimédia, Image, Réalité virtuelle et Application). La création de cette option a eu lieu suite à la fusion entre l'ESIL et l'IAAI (Institut des Applications Avancées de l'Internet). J'ai ainsi eu l'opportunité de proposer une spécialité avec une nouvelle équipe pédagogique comportant à la fois des collègues de l'ESIL et de l'IAAI. Nous avons pu collectivement mettre en place MIRA en prenant en compte à la fois les besoins du marché de l'emploi et les compétences de chacun. La moitié des membres du département informatique ont été fortement impliqués dans cette réflexion. Aujourd'hui, l'option MIRA a évolué en une spécialité du diplôme d'ingénieur en informatique de Polytech Marseille. Cette spécialité, nommée ReVA (Réalité Virtuelle et Augmentée), représente près de 400 heures réparties sur 2 semestres. L'école propose ainsi une formation reconnue autour de la Réalité Virtuelle et Augmentée qui permet à nos élèves d'obtenir des stages et emplois de qualité en adéquation avec leur projet professionnel.

Je suis également membre de l'équipe pédagogique du parcours GIG (Géométrie et Informatique Graphique) du master informatique de la faculté des sciences d'Aix-Marseille Université. A ce titre, j'ai participé à la création de ce parcours qui a ouvert ses portes en septembre 2018.

Auparavant, je me suis impliqué dans le cadre du master SIS (Sciences de l'Information et des Systèmes). J'ai été responsable, pendant cinq ans (2005-2010), de l'option « Imagerie Numérique » du Master SIS Recherche et je me suis particulièrement investi en assurant un travail administratif, pédagogique et d'animation de la recherche. De plus, j'ai fait partie de l'équipe pédagogique du Master SIS Professionnel et j'y ai enseigné un cours de « Vision Industrielle et Télédétection » (60h eq. TD de 2005 - 2010).

Enfin, c'est avec beaucoup de passion et d'enthousiasme que j'ai enseigné à deux reprises (en 2008 et 2009) au Burkina Faso lors de missions pédagogiques supportées par l'Agence Universitaire de la Francophonie. J'ai effectué chaque année 48h de cours dans le domaine de la programmation graphique.

### 1.4.2 Service d'enseignement

J'enseigne à tous les niveaux du cycle d'ingénieur ainsi qu'en seconde année des classes préparatoires intégrées à l'école. Pendant plusieurs années, j'ai également créé et donné des cours mutualisés entre le département informatique et le département biomédical. Le tableau ci-dessous présente de façon synthétique l'essentiel de mon service d'enseignement des quatre dernières années à Polytech Marseille.

Intitulé	Année	Eq. TD
Algorithmique et Programmation Python	PEIP 2 <sup>ème</sup> année	48h
Langages de l'internet	PEIP 2 <sup>ème</sup> année	20h
Introduction à l'Informatique	INFO 3 <sup>ème</sup> année	20h
Projets – Programmation Orientée Objet	INFO 4 <sup>ème</sup> année	32h
Analyse d'Images	INFO 4 <sup>ème</sup> année	24h
Synthèse et Visualisation	INFO 4 <sup>ème</sup> année	24h
Projets – spécialité ReVA (4 par an)	INFO 4 <sup>ème</sup> année	12h
Analyse d'Images Avancée	INFO 5 <sup>ème</sup> année	32h
Réalité Virtuelle	INFO 5 <sup>ème</sup> année	16h
3D sur le web	INFO 5 <sup>ème</sup> année	32h
Projets Industriels – spécialité ReVA (4 par an)	INFO 5 <sup>ème</sup> année	12h

Depuis de nombreuses années, mon service d'enseignement dépasse les 192 heures statutaires et représente un nombre d'heures complémentaires significatif (souvent bien au delà des 100h eq. TD), et ce pour deux raisons principales : d'une part, les besoins du département informatique de l'école sont importants et il me semble tout naturel de participer à répartition de la charge d'enseignement avec mes collègues. D'autre part, Polytech Marseille propose 8 spécialités d'ingénieur. C'est une palette de choix très large et il est indispensable que des enseignants du département informatique participent aux enseignements du cycle préparatoire intégré afin de donner de la visibilité à la formation en informatique. J'enseigne ainsi près de 70 heures en cycle préparatoire.

### 1.4.3 Projet pédagogique

Depuis deux ans, j'ai mis en place le projet pédagogique que je vais présenter succinctement dans cette section. J'ai obtenu des financements et les premières réalisations voient le jour.

Pour mettre en place ce projet, j'ai répondu, en tant que porteur de projet, à différents appels à manifestation d'intérêt pour l'aménagement de nouveaux espaces pédagogiques dans les campus de l'Université. Ceux-ci représentent un enjeu important pour répondre aux demandes des étudiants de bénéficier d'espaces adaptés pour un travail individuel ou collectif (espaces informels) et pour favoriser de nouvelles pratiques pédagogiques dans lesquelles les étudiants sont plus actifs (espaces formels). Le projet pédagogique que je soutiens et que je porte a pour objet le développement de contenus numériques pour une immersion



3D interactive du public. Ce projet anticipe des évolutions pédagogiques très probables dans l'enseignement supérieur mais aussi secondaire, évolutions rendues possibles par la démocratisation des technologies numériques.

### **Motivation du projet**

Le secteur de la formation connaît de profondes mutations et doit s'adapter à un nouvel écosystème davantage numérisé. Face à cette révolution numérique, le comportement des élèves, leur façon de consommer l'information, d'apprendre, de se former a fortement évolué. Beaucoup d'enseignants ont expérimenté l'efficacité de courtes vidéos, directement disponibles sur la toile (donc du « monde réel et moderne » pour nos élèves), intégrées dans le cadre d'une approche pédagogique globale. Avec des supports soigneusement sélectionnés en appui, en quelques minutes, l'information (qualitativement et quantitativement) transmise avec efficacité aux élèves est grandement améliorée par rapport à des pratiques plus « magistrales ». Si l'interactivité initiale est d'abord entre le support vidéo et l'enseignant qui s'appuie dessus, cette interactivité déclenche souvent, une plus grande participation des étudiants. Ces démarches méritent d'être développés mais nous proposons d'aller un cran plus loin et d'immerger l'élève-ingénieur dans un univers 3D dans lequel il sera plus acteur.

Parmi les récentes avancées technologiques, la réalité virtuelle occupe une place de premier choix et permet d'envisager des utilisations fluides du monde numérique (3D, nouvelles interfaces naturelles, gestion de l'information par le geste, mobilité, réseaux sociaux...). C'est une opportunité unique qui permet de proposer une pédagogie par l'action, engageante, immersive, interactive, collaborative, ludique et amusante.

### **Objectifs poursuivis**

Ce projet a pour objectif de créer un espace immersif multi-usages, multi-publics, configurable essentiellement par logiciel (« comme on change de DVD ») qui aspire de surcroît à une ouverture (bi-directionnelle) vers l'extérieur de l'Université (élèves du secondaire, grand public, monde socio-économique). Pour que cet espace soit fonctionnel et réponde aux attentes deux points sont essentiels :

- L'accueil d'une trentaine de personnes simultanément
- Et la capacité de réaliser une projection immersive 3D d'une taille d'au moins 200 pouces.

Ces caractéristiques ont motivé une demande de financement qui concerne donc la construction d'un lieu dédié au dispositif qui est en cours de discussion au sein de l'école.

Ce projet propose ainsi une rupture à la fois avec les lieux d'immersion habituels (qui accueillent généralement un ou deux participants à la fois et qui sont généralement dédiés à la recherche) mais aussi avec la nature des supports multimédias utilisés aujourd'hui auprès

des apprenants. En effet, il s'agira de diffuser des supports multimédias immersifs. Ces supports permettront de proposer un contenu multimédia 3D-immersif comme, par exemple, la visite d'une nanoparticule, d'un virus, d'un dispositif médical, d'un bâtiment, de données massives ... et ainsi d'immerger l'apprenant dans un univers 3D dans lequel il sera plus acteur. Un travail sur la création et la diffusion de ce type de contenu est en cours de réalisation à Polytech et notamment au département informatique par des élèves de dernière année du cycle d'ingénieur.

### Contacts académiques

Plusieurs contacts académiques ont été pris et les retours sont positifs. Je peux notamment citer :

- Le Centre de Réalité Virtuelle de Luminy : collaboration/mise à disposition du moteur de rendu ICE 3D.
- Des laboratoires de recherche (première source de contenus) : un premier tour de table a permis de s'assurer de l'intérêt de nombreux laboratoires. Ils nous fournissent de la matière. Nous la valorisons sous forme de fichiers 3D ( -> Polythèque 3D ). En retour, nous leur fournissons ces fichiers 3D communicants qu'ils pourront alors projeter avec des moyens très légers. Pour plus d'impact, nous pourrions occasionnellement les accueillir dans la salle immersive.
- Le réseau Polytech : plusieurs écoles du réseau sont intéressées par s'associer au projet et à la démarche de création de contenus.
- L'ESPE (École Supérieure du Professorat et de l'Éducation) a été approché. Cette expérience pourrait être le support à des travaux de recherche.
- Le Rectorat s'est dit intéressé par l'idée et par les contenus notamment dans la perspective d'applications futures en secondaire (complétion de Eduthèque).

### Résultats attendus

Ce projet aspire à être attractif et à innover pédagogiquement. Les bénéficiaires d'un tel dispositif pourraient être :

- Les élèves-ingénieurs : tous les élèves mais à commencer par ceux de la filière informatique (réalité virtuelle) et particulièrement, pour leurs projets, ceux de la déclinaison HuGO (qui intègre des élèves ingénieur en situation de handicap) de la filière informatique qui expérimentera de nouvelles pratiques pédagogiques appelées à être généralisées si elles s'avèrent efficaces.
- Des publics d'autres composantes.

- Des représentants de la société civile : On reste dans un objectif pédagogique, d'acculturation d'une thématique mais on vise là d'autres publics.
- Des scolaires (accueil des scolaires et leurs enseignants).

Les résultats attendus sont triples :

- Sur le plan pédagogique (connaissances et compétences développées chez les apprenants) : il s'agit d'apporter de la motivation, de développer une compétence pour mieux faire le lien avec le réel et aider les élèves à s'inscrire dans un projet global et complexe.
- Sur les pratiques des enseignants : il est attendu des pratiques plus interactives avec les élèves et aussi le développement de nouvelles pratiques permettant aux élèves de valoriser leur recherche.
- Pour l'école, l'Université, les partenaires : il s'agit de développer un outil pédagogique et communicant porteur d'une image dynamique et d'ouverture.

### **Indicateurs de réussite**

Afin d'évaluer la pertinence du projet plusieurs indicateurs de réussite seront mis en place assez tôt (t0+18 mois) :

#### **Technique**

- Choix de la solution interactive (capteurs, intégration, moteur de rendu3D, hardware),
- Validation de cette solution interactive en 100 pouces,
- Dimensionnement pour sa mise en œuvre en 200 pouces ou plus,
- Projection immersive 3D en 200 pouces de diagonale pour 30 élèves min.

#### **Contenus**

- Définition du type de licence,
- Lancement du premier tour de table,
- Abondement de la Polythèque de 12 supports 3D interactifs minimum à créer à partir des sources de données,
- Et de (>24) vidéos 3D.

#### **Satisfaction**

- Evaluation de la satisfaction des élèves (en tant qu'élèves),
- Evaluation (externe, sociologues) de l'engouement des enseignants à utiliser l'outil pour leur cours,
- Des partenaires.

### Premiers résultats

Ce projet est soutenu par Polytech Marseille et par l'Université. Un premier budget d'environ 40000 euros a été alloué et a permis l'achat de matériel afin de développer une première version du projet (video-projecteur laser, large écran de projection, ordinateur ...).

Deux développements informatiques ont été initiés pour permettre de réaliser un démonstrateur. Le premier porte sur la visualisation et la manipulation de scène en 3D Stéréoscopique et le second sur la définition d'un format de fichier ouvert permettant de stocker l'ensemble des ressources nécessaires relatives au contenu pédagogiques.

### Personnel impliqué

L'équipe qui œuvre à ce projet est composée d'enseignants chercheurs de Polytech Marseille, de personnel s'occupant de la logistique (aménagement de salle, fonctionnement du matériel ...), d'un ingénieur de recherche qui joue le rôle de soutien pédagogique et de coordonnateur.

Les élèves de l'école sont également impliqués dans ce projet à travers leur participation au développement logiciel ainsi que par leur participation à la création de contenus.

## 1.5 Encadrement de thèses de doctorat

J'ai co-dirigé huit thèses. Cinq d'entre elles ont été soutenues et trois sont en cours.

### 1.5.1 Thèses soutenues

J'ai encadré cinq thèses de doctorat en co-direction avec le professeur Jean Sequeira.

A. Le Troter, « **Mise en correspondance d'images et de modèles : application à la reconstruction 3D de scènes sportives** », soutenue en septembre 2006.

**Jury composé de :** Christian Roux, Rachid Deriche, Maciej Orkisz, Jean-Marc Boï, Jean Sequeira, Sébastien Mavromatis.

**Mots-clés :** Analyse d'image couleur, détection de région d'intérêt, extraction et mise en correspondance d'images, reconstruction automatique.

**Résumé :** Cette thèse s'inscrit dans le cadre du projet SimulFoot dont l'objectif est l'analyse de scènes de football. Nos problématiques concernent l'analyse de séquences d'images, la modélisation géométrique et la visualisation, ainsi que la simulation et l'ergonomie cognitive. Nos principales recherches se sont focalisées sur des processus de traitement d'images pour la reconstruction automatique d'un modèle 3D d'une scène à partir d'une seule image. Une

première étape permet la détection de la région d'intérêt (le terrain) par une analyse colorimétrique. Nous illustrons notre méthodologie à travers l'étude de la distribution des pixels dans l'espace HLS. En complément, nous proposons une nouvelle façon de structurer, au niveau perceptuel, l'espace des couleurs CIELab. Une seconde étape propose une approche originale d'extraction (par transformées de Hough) des amers de la région d'intérêt (droites et ellipses) et leur mise en correspondance avec des éléments remarquables du modèle 3D.

### **Publications :**

- [30] Arnaud LE TROTTER, Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. "Automatic Landmark Detection and Validation in Soccer Video Sequences". In : *Computer Vision and Graphics*. Springer Netherlands. Warsaw, Poland, 2007, p. 774-779. DOI : 10.1007/1-4020-4179-9\_112.
- [31] Arnaud LE TROTTER, Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. "Segmentation d'images en plages de couleurs dominantes". In : *TAIMA'05 Ateliers francophones sur le Traitement et l'analyse de l'information, méthodes et applications*. Hammamet, Tunisie, 2005.
- [32] Arnaud LE TROTTER, Sébastien MAVROMATIS et Jean SEQUEIRA. "Soccer field detection in video images using color and spatial coherence". In : *Image Analysis and Recognition*. Springer Berlin Heidelberg. Porto, Portugal, 2004, p. 265-272. DOI : 10.1007/978-3-540-30126-4\_33.
- [33] Arnaud LE TROTTER, Sébastien MAVROMATIS et Jean SEQUEIRA. "Automatic selection of a region of interest in 3D scene images : application to video captured scenes". In : *Advanced Concepts for Intelligent Vision Systems (ACIVS)*. Brussels, Belgium, 2004.
- [34] Hubert RIPOLL, Arnaud LE TROTTER, Jean BARATGIN, Sébastien MAVROMATIS, M FAISOLLE, F ZMILSONY, Gérald POPLU, JP PETIT et Jean SEQUEIRA. "The interest of simulation for research and training in sport : the example of football". In : *Proceedings of the Third International Sport Sciences Days, The analysis of elite performance in its contextual environment (INSEP)*. Paris, France, 2004.
- [35] Hubert RIPOLL, Arnaud LE TROTTER, Jean BARATGIN, Sébastien MAVROMATIS, M FAISOLLE, F ZMILSONY, Gérald POPLU, Jean-Philippe PETIT et Jean SEQUEIRA. "Un simulateur pour l'entraînement en football". In : *Journées Internationales des Sciences du Sport (INSEP)*. Paris, France, 2004.

---

L. Lluçia, « **Suivi d'objets à partir d'images issues de cameras mobiles non calibrées** », soutenue en février 2011.

**Jury composé de :** Dominique Faudot, Samir Akkouche, Laurent Lucas, Marc Daniel, Jean Sequeira, Sébastien Mavromatis.

**Mots-clés :** Analyse d'image couleur, mise en correspondance d'amers, reconstruction automatique.

**Résumé :** Cette thèse s'inscrit dans le cadre du projet SimulFoot dont l'objectif est l'analyse de scènes de football. Nos problématiques concernent l'analyse de séquences d'images, la modélisation géométrique et la visualisation, ainsi que la simulation et l'ergonomie cognitive. Nos principales recherches se sont focalisées sur l'automatisation du processus de recalage par la sélection automatique des amers dans les images et leur mise en correspondance avec ceux du modèle. Une approche originale et robuste de détection de faisceaux de droite est présentée et permet d'atteindre nos objectifs. Dans une seconde partie, à partir d'hypothèses réalistes dans notre cadre applicatif, nous proposons une estimation du calibrage de la caméra à partir d'une matrice d'homographie. Enfin, cette thèse qui a été réalisée dans un cadre particulier (CIFRE + projet de maturation) a fait l'objet d'une forte contribution au développement du logiciel SimulFoot Entraînement.

**Publications :**

- [25] Bachir ZOUDJI, Hubert RIPOLL, Ludovic LLUCIA, Sébastien MAVROMATIS et Jean SEQUEIRA. "SimulFoot : un outil technologique pour l'analyse et la formation tactique en football". In : *Visions prospectives du professionnalisme sportif en Algérie*. 2009.
- [26] Ludovic LLUCIA, Sébastien MAVROMATIS, Sébastien PERROTTE, Paulo DIAS et Jean SEQUEIRA. "Camera location and aperture characterization using the transformation between a 2D plane and the image captured by the camera". In : *Image Analysis and Recognition*. Springer Berlin Heidelberg. Povia de Varzim, Portugal, 2008, p. 385-394. DOI : 10.1007/978-3-540-69812-8\_38.
- [28] Hubert RIPOLL, Ludovic LLUCIA, Sébastien MAVROMATIS, Bachir ZOUDJI et Jean SEQUEIRA. "SimulFoot entraînement, un outil de simulation pour les entraîneurs et les joueurs". In : *3ème Colloque International Football & Recherches*. Valenciennes, France, 2008.

---

Ch. Palmann, « **Contribution au recalage d'images de modalités différentes à travers la mise en correspondance de nuages de points. Application à la télédétection** », soutenue en juin 2011. **Jury composé de :** Jean-Marie Nicolas, Jocelyn Chanussot, Christian Roux, Gu Xingfa, Frédéric Bouchara, Jean Sequeira, Sébastien Mavromatis.

**Mots-clés :** Fusion de données, nuages de points, mesure de similarité, recalage, reconnaissance de formes, télédétection.

**Résumé :** La fusion de données acquises à partir de sources différentes est aujourd'hui une problématique fortement émergente tant aux niveaux méthodologique qu'applicatif. L'obtention d'une information plus riche est en effet profitable dans de nombreuses situations

pratiques. Or les données correspondent souvent à de grands volumes à traiter, dans lesquels l'information est redondante, et ne sont généralement pas présentées de façon uniformisée. Dans un tel contexte, l'intégration automatique des données au sein d'un même modèle réclame le développement d'outils appropriés pour être réalisée. Dans le cas de capteurs de natures différentes (modalités différentes), les images ne sont généralement pas situées dans un même référentiel spatial, ce qui empêche une exploitation simultanée. Il est donc nécessaire que des images subissent des transformations géométriques afin d'être superposées (recalées) sur d'autres images. Deux difficultés majeures rendent ce processus de recalage difficile à réaliser automatiquement. Tout d'abord, la représentation de l'information n'est pas la même d'une modalité à l'autre; de plus, les éléments d'une même scène n'ont pas forcément les mêmes expressions dans des images de modalités différentes. C'est pourquoi ce type de recalage reste à l'heure actuelle un domaine de recherche très actif. Dans cette thèse, nous avons apporté notre contribution à la résolution de ce problème, à travers la mise en correspondance de nuages de points rattachés à chaque image. Le cadre applicatif a été la télédétection.

### Publications :

- [7] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "Mesure de similarité entre sous-parties de nuages de points 2D". In : *TS. Traitement du signal* 29.1-2 (2012), p. 29-49. DOI : 10.3166/ts.29.29-49.
- [8] Christophe PALMANN, Sébastien MAVROMATIS, Mario HERNÁNDEZ, Jean SEQUEIRA et Brian BRISCO. "Earth observation using radar data : an overview of applications and challenges". In : *International Journal of Digital Earth* 1.2 (2008), p. 171-195. DOI : 10.1080/17538940802038317.
- [20] Sébastien MAVROMATIS, Christophe PALMANN et Jean SEQUEIRA. "Characterization of Similar Areas of Two 2D Point Clouds". In : *Advances in Visual Computing*. Crete, Greece : Springer Berlin Heidelberg, 2012, p. 509-516. DOI : 10.1007/978-3-642-33191-6\_50.
- [22] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "A new geometric invariant to match regions within remote sensing images of different modalities". In : *SPIE Remote Sensing*. International Society for Optics et Photonics. Prague, Czech Republic, 2011, p. 818003-818003. DOI : 10.1117/12.898158.
- [23] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "Mesure de similarité entre sous-parties de nuages de points 2D". In : *TAIMA'11 Ateliers francophones sur le Traitement et l'analyse de l'information, méthodes et applications*. Hammamet, Tunisie, 2011.
- [24] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "A new approach for registering remote sensing images from various modalities". In : *SPIE Europe Remote Sensing*. International Society for Optics et Photonics. Berlin, Germany, 2009, p. 74770C-74770C. DOI : 10.1117/12.830248.

- [27] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. “SAR image registration using a new approach based on the generalized hough transform”. In : *ISPRS 2008*. Beijing, China, 2008.
- 

F. Graglia, « **Réalisation d'une maquette virtuelle interactive. Application a la modélisation des caractéristiques lumineuses d'une cabine d'hélicoptère** », soutenue en novembre 2012.

**Jury composé de :** Saïda Bouakaz, Jean-Pierre Jessel, Franck Marzani, Gilles Gesquière, Jacques Delacour, Jean Sequeira, Sébastien Mavromatis.

**Mots-clés :** Rendu, physiquement réaliste, 3D, lancer de rayons, photon mapping, illumination globale, temps réel.

**Résumé :** Lors du développement d'un produit industriel, il est nécessaire de visualiser à différentes étapes du processus la maquette numérique du produit à des fins de validation et de mesures. Cette thèse a pour objectif d'améliorer la qualité et le réalisme des simulations 3D dans le contexte d'un processus de production. Ces simulations impliquent souvent des scénarios de type « walkthrough », avec une géométrie fixe mais un déplacement libre de l'observateur. Nous nous intéresserons donc plus précisément aux modèles et aux méthodes de rendu physiquement réaliste de scènes complexes pour ce type de scénario. Durant l'étape de rendu, l'utilisateur doit pouvoir mesurer précisément la luminance énergétique émise par un point ou une zone donnée, ainsi que modifier en temps réel la puissance des sources lumineuses. A travers l'étude approfondie du photon mapping, nous allons montrer comment réduire la complexité de l'étape de rendu afin d'améliorer les performances du processus dans le cadre d'un scénario « walkthrough ». Nos travaux s'intéressent notamment aux relations existant entre la géométrie et la carte de photons, aux structures accélératrices utilisées, ainsi qu'à la réduction du biais induit par la méthode du photon mapping. Nous montrerons enfin les résultats obtenus sur des scènes industrielles complexes en les comparant notamment à des algorithmes physiquement réalistes consistants.

**Publications :**

- [21] Florian GRAGLIA, Jean SEQUEIRA et Sébastien MAVROMATIS. “Real-time Walkthrough Rendering - Speed Improvement of Photon Mapping Algorithm”. In : *Proceedings of the International Conference on Computer Graphics Theory and Applications (VISIGRAPP 2012)*. Rome, Italy, 2012, p. 291-298. DOI : 10.5220/0003817702910294.

Les travaux développés ont fait l'objet d'un dépôt auprès de l'Agence de Protection des Programmes et une licence logicielle a été cédée à la société OPTIS.

---



P. Zoppitelli, « **Détection et caractérisation d'entités en mer à partir d'un hélicoptère** », soutenue en février 2019.

**Jury composé de :** Saïda Bouakaz, Sylvie Treuillet, Franck Marzani, Marc Daniel, Nicolas Bélanger, Jean Sequeira, Sébastien Mavromatis.

**Mots-clés :** Reconnaissance de formes, suivi d'objet, vol autonome.

**Sujet :** Dans le contexte défini (la détection et caractérisation d'entités en mer), trois phases se distinguent, chacune correspondant à un moment de la phase d'approche. Lors de la première étape qui se passe au début de la phase d'approche, le pilote distingue un objet démarquant de la mer et décide de se diriger vers lui. Durant cette étape, les caractéristiques principales de l'objet ne peuvent être distinguées, les seules propriétés de l'objet que l'on peut inférer sont la petitesse de celui-ci dans l'image (première apparition dans l'image) et sa dissemblance par rapport à la mer (sinon aucune détection n'est possible). Une fois que l'hélicoptère s'est avancé vers sa destination (deuxième étape), des éléments de formes et de structures apparaissent : la forme d'un bateau ou l'agencement particulier d'un ensemble de plateformes pétrolières. Il est encore trop tôt pour donner une définition précise de la nature de l'entité mais son caractère remarquable est certain. Enfin, lorsque l'hélicoptère s'est fortement rapproché, les éléments visuels constitutifs de l'objet deviennent apparents. La détection et la reconnaissance de chacune de ses entités demandent d'introduire une connaissance spécifique à la classe d'objets à reconnaître et sont des problèmes entiers. Dans le cadre des missions SAR ces éléments peuvent être le pont d'un bateau, le mât du bateau, le câble rattaché au treuil, les opérateurs travaillant sur le pont du bateau, les vagues, etc . . . Dans le contexte des missions de transport sur plateformes offshore, ces éléments peuvent être : une zone de poser, un hélicoptère posé, un symbole particulier placé sur l'hélicoptère, une grue, une tour, etc. Dans le cadre de cette thèse les problématiques privilégiées seront celles portant sur le transport vers les plateformes offshore avec une insistance particulière sur la détection d'hélicoptère.

### **Publications :**

- [15] Pierre ZOPPITELLI, Sébastien MAVROMATIS, Jean SEQUEIRA, Guillaume ANOUFA, Nicolas BELANGER et Francois-Xavier FILLIAS. "Embedding intelligent image processing algorithms : the new safety enhancer for helicopters missions". In : *44th European Rotorcraft Forum - ERF 2018*. Delft, Netherlands, 2018.
- [17] Pierre ZOPPITELLI, Sébastien MAVROMATIS et Jean SEQUEIRA. "Ellipse Detection in Very Noisy Environment". In : *WSCG*. Pilsen, Czech Republic, mai 2017.

Un brevet a été publié le 11 janvier 2018 par Airbus sur le travail réalisé (numéro 20180009546). L'intitulé est « Device for assisting the piloting of a rotorcraft, an associated display, and a corresponding method of assisting piloting ». Les inventeurs sont FX. Filiat, P. Zoppitelli, N. Belanger, M. Salesse-Lavergne.

### 1.5.2 Thèses en cours

Je co-encadre actuellement trois thèses de doctorat en financement CIFRE.

N. Jin, « **Vers une évaluation robuste, précise et « temps réel » des rapports entre le mouvement des yeux et ceux de la tête lorsque celle-ci est soumise à des sollicitations externes : Mise en œuvre dans le cadre du programme STABILITAS de Synapsys** », soutenance prévue en juin 2020.

**Encadrants :** Jean Sequeira, Sébastien Mavromatis.

**Sujet :** Une caméra vidéo assujettie à un masque permet de capturer une image (dynamique) de l'œil et de son environnement immédiat. L'œil est mobile dans son orbite, en 3D, à la fois en rotation autour de deux axes relatifs « horizontaux » et « verticaux » – mouvement qui est associé à une translation dans l'image – mais aussi autour d'un troisième axe, perpendiculaire aux deux premiers, et c'est ce qu'on appelle le mouvement « torsionnel » de l'œil – mouvement qui est associé à une rotation autour du centre de la pupille dans l'image. L'objectif de ce travail de recherche est de caractériser ces mouvements et de les mettre en rapport avec les mouvements de la tête qui les ont induits (à travers la perception des accélérations par les canaux semi-circulaires de l'oreille interne), cette caractérisation devant être robuste, précise et effectuée au rythme de l'acquisition des images par la caméra vidéo. Un contrat de confidentialité a été signé dans le cadre de cette thèse et, à ce stade, les publications sont bloquées.

---

R. Girard, « **Vers un système de reconstruction 3D temps réel de type SLAM couplé à la projection d'une lumière structurée pour l'aide au pilotage en hélicoptère** », soutenance prévue en janvier 2021.

**Encadrants :** Jean Sequeira, Sébastien Mavromatis.

**Sujet :** Le problème que nous abordons dans cette thèse est celui de la mise en œuvre du SLAM (Simultaneous Localization and Mapping) – dont l'objectif est le créer un modèle de l'environnement, sans disposer de connaissance sur celui-ci, et de localiser le capteur dans cet environnement – sur des hélicoptères dotés d'une seule caméra (l'acquisition stéréoscopique – bi-caméras – n'est pas envisageable du fait de la forme et de la compacité d'un hélicoptère, et de la distance d'acquisition), mais à laquelle on peut adjoindre un dispositif de projection de mire (lumière structurée).

**Publications :**

[13] Rémi GIRARD, Sébastien MAVROMATIS, Jean SEQUEIRA, Nicolas BELANGER et Guillaume ANOUFA. "A vision-based assistance key differentiator for helicopters autonomous

scalable missions”. In : *20th Towards Autonomous Robotic Systems Conference, TAROS*. London, England, 2019.

---

A. Monneau, « **Appontage automatique pour hélicoptère** », soutenance prévue en janvier 2021.

**Encadrants :** Nacer M’Sirdi, Sébastien Mavromatis.

**Sujet :** Ce travail de recherche vise à développer un système autonome permettant l’appontage d’un hélicoptère d’un tonnage significatif sur une frégate en mouvement. L’objectif de ce système est d’établir une trajectoire jusqu’au point de rendez-vous entre l’hélicoptère et la plateforme d’appontage. Le calcul de la trajectoire utilise l’estimation de la position relative de l’hélicoptère par rapport à la plateforme. Cette estimation devra être suffisamment précise et robuste à toute panne simple, par conséquent il s’agira de fusionner des informations issues d’un jeu de senseurs utilisant différentes technologies. Afin d’affiner la stratégie d’appontage et de faciliter la prise de décision, une partie de la Recherche sera consacrée aux prédictions des mouvements du pont de bateau et à cette estimation de position relative. Enfin, la dernière étape consistera à implémenter une loi de contrôle robuste mais avec un objectif de performances ambitieux, asservissant l’hélicoptère sur cette trajectoire.

**Publications :**

[14] Antoine MONNEAU, Kouider Nacer M’SIRDI, Sébastien MAVROMATIS, Guillaume VARRA, Marc SALESSE et Jean SEQUEIRA. “Adaptive Prediction for Ship Motion in Rotorcraft Maritime Operations”. In : *5th CEAS Conference on Guidance, Navigation and Control - EuroGNC19*. Milan, Italy, 2019.

---

### 1.5.3 Devenir des docteurs

- Pierre Zoppitelli a été embauché en CDI par un sous-traitant d’Airbus et continue à travailler sur le projet Eagle.
- Florian Graglia a été embauché en CDI par la société OPTIS à l’issue de sa thèse.
- Christophe Palmann, après un post-doctorat à l’IRSA (Chinese Academy of Science – Beijing), a été embauché en qualité de responsable « Traitement d’images » en CDI par la société Uratek.
- Ludovic Lluçia a été embauché en qualité d’ingénieur R&D en CDI par la société SII.

- Arnaud Le Troter a été embauché en qualité de Post-Doc pour une durée de 18 mois par la société PIXXIM. Il est aujourd'hui ingénieur de recherche CNRS et travaille au centre de résonance magnétique biologique et médicale à Marseille.

### 1.5.4 Encadrement de stages au Master Recherche SIS

J'ai encadré 3 stages de master recherche. A l'issue de leur master recherche, les trois étudiants (T. Bonaccorsi, A. Nisan et Ch. Palmann) ont poursuivi en thèse au sein du pôle image du laboratoire LSIS.

- Ch. Palmann, « Analyse d'images RSO », juin 2007.
- A. Nisan, « Modèles Numériques de Terrain à partir d'images stéréoscopiques ASTER », juin 06.
- T. Bonaccorsi, « Détection des situations anormales sur des scènes acquises en caméra fixe », juin 05.

## 2 Projets de recherche

Ce chapitre présente les projets de recherche que j'ai mis en oeuvre et gérés, souvent avec l'aide et la participation d'autres chercheurs du laboratoire, au cours de ces quinze dernières années. Naturellement, je ferai état de mon niveau d'implication lorsque celui-ci sera secondaire au sein du projet. En fait, plutôt que d'énumérer une liste de projets, j'ai préféré montrer comment ils ont jalonné mon parcours au sein du laboratoire, et m'ont permis, tout en faisant de la recherche fondamentale, de travailler en partenariat avec des industriels et en collaboration avec des instituts étrangers.

La première section est la plus importante, et je reviendrai dessus dans un autre chapitre (dédié aux aspects scientifiques fondamentaux), car elle montre comment, à travers trois grands projets, aux multiples facettes, j'ai pu mettre en oeuvre une véritable « ligne de recherche » qui s'inscrit dans le temps. Au delà de cette ligne forte, je souhaitais aborder tout un ensemble de problèmes connexes, plus spécifiques voire associés à un verrou technologique, et qui ne s'inscrivaient pas dans ce « cadre général » : c'est la raison pour laquelle j'ai initié, au cours du temps, un ensemble de projets plus limités, essentiellement en partenariat avec des industriels, et c'est l'objet de la section suivante. Dans la troisième section, je présente un ensemble de collaborations internationales auxquelles j'ai participé activement. Enfin, dans la dernière section, je donne quelques éléments sur un projet émergent, qui a débuté effectivement depuis plus d'un an, et qui fait partie de mon projet scientifique pour les années à venir.

### 2.1 Projets structurants et pérennes

#### 2.1.1 Reconstruction 3D de scènes à partir de séquences vidéo

- **Date :** 2000 - ... en cours
- **Partenaires :** Laboratoire des Sciences de l'Information et des Systèmes (LSIS), Institut des Sciences du Mouvement (ISM Marey), Université d'Aveiro (Portugal), Société SII, SATT Sud-Est

- **Financement** : bourse de thèse du ministère, bourse CIFRE, agence universitaire de la Francophonie (AUF), ValorPACA, SATT Sud-Est. Le montant des financements est de l'ordre de 500 000 euros.
- **Résultats** : analyse colorimétrique d'une image, détection et mise en correspondance d'amers, calibrage de caméra, maturation de projet, développement de deux prototypes logiciels, licence logicielle.

Pendant ma thèse de doctorat, j'ai été ponctuellement employé comme agent contractuel à la faculté des sciences du sport de Luminy à Marseille. Mon travail consistait essentiellement à « monter des manipulations » pour les chercheurs en psychologie cognitive. Concrètement, il s'agissait de développer des programmes informatiques qui permettaient d'afficher un série d'image, d'enregistrer les réponses des sujets, de mesurer les temps de réponses . . .

Cette fonction m'a permis d'engager des discussions avec les chercheurs autour des environnements virtuels et de la possibilité de créer du contenu pour leurs expérimentations. Nous avons très rapidement convergé vers le souhait de développer un tel environnement autour de scènes issues du domaine du sport. L'enthousiasme était bien présent mais la tâche ambitieuse dans la mesure où il ne s'agissait pas « seulement » de développement mais bien d'un projet de recherche pluridisciplinaire et ambitieux. Le projet SimulFoot est alors né de cette réflexion sur l'utilisation de ces nouvelles technologies pour l'analyse de situations et de comportements dans le domaine des sports collectifs, et particulièrement en football [34, 35, 36, 38, 39]. Il est important de noter qu'il ne s'agissait pas de développer un logiciel de création d'images réalistes (pour produire la visualisation la plus esthétique possible) mais d'apporter une contribution foncière sur l'analyse du rôle de l'image dans la communication, avec une dimension ergonomique incontournable.

Les verrous scientifiques à résoudre dans le cadre de ce projet sont nombreux. Il est possible de citer, entre autres, les problèmes d'analyse d'images et de séquences vidéo, de modélisation géométrique et d'animation, mais aussi de visualisation réaliste, statique et dynamique, ainsi que de simulation et de communication homme-machine (ergonomie cognitive).

Le projet SimulFoot a permis de créer une forte activité aussi bien en matière de recherche que de relation avec les entreprises (clubs sportifs, entreprises du monde du sport, sociétés de service) et de valorisation industrielle des travaux.

J'ai co-encadré deux thèses dans le cadre de ce projet. La première était financée par une allocation de recherche du ministère de la recherche. Elle est intitulée « Mise en correspondance d'images et de modèles : application à la reconstruction 3D de scènes sportives ». Nous avons apporté des solutions à la détection de « régions d'intérêts » dans des images couleurs ainsi qu'à l'extraction et la mise en correspondance d'information permettant le calibrage de la prise de vue [30, 31, 32, 33]. Les travaux concernant la segmentation et l'analyse d'images couleurs sortent largement du contexte du projet SimulFoot. La méthodologie proposée permet à travers l'étude de nuages de points dans un espace couleur de proposer une segmentation

pouvant prendre en compte différentes problématiques comme la perception, l'impression, la transmission ...). La seconde thèse intitulée « Suivi d'objets à partir d'images issues de caméras mobiles non calibrées » a été financée par une convention industrielle de formation par la recherche (CIFRE) avec la société SII. Cette thèse a été menée simultanément à un projet de maturation que je présenterai dans le paragraphe 3.3 consacré aux projets de maturation auxquels j'ai participé. Nous avons essentiellement apportés une contribution scientifique à la mise en correspondance d'information permettant le calibrage de la prise de vue [26, 5]. Une contribution significative a également été apportée à la structuration des travaux menés dans le projet SimulFoot aussi bien en informatique qu'en psychologie cognitive [25, 28] afin de permettre le développement d'un prototype logiciel réalisé par la société SII.

Parallèlement à ces travaux scientifiques, deux projets de maturation logicielle et un projet de création d'entreprise ont été initiés. Je ferai le point sur ces aspects dans le chapitre 3.3 dédié à la valorisation de mes travaux scientifiques.

### 2.1.2 Analyse et recalage d'images de modalités différentes

- **Date** : 2008 - ... en cours
- **Partenaires** : Institute of Remote Sensing and Application (IRSA - Chine) - Yukon River Inter-Tribal Watershed Council (YRITWC - Alaska) - Institut Supérieur d'Informatique et de Gestion (ISIG - Burkina Faso) - Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
- **Financement** : bourse de thèse du ministère, agence universitaire de la Francophonie. Le montant du financement est de l'ordre de 100 000 euros.
- **Résultats** : caractérisation d'une mesure de similarité entre sous-nuages de points. Brevet international.

Cette thématique est née de contacts privilégiés avec trois entités étrangères qui sont l'IRSA (Institute of Remote Sensing and Application - Chine), l'YRITWC (Yukon River Inter-Tribal Watershed Council - Alaska) et l'ISIG (Institut Supérieur d'Informatique et de Gestion - Burkina Faso). Il n'y a pas de lien de collaboration directe entre ces institutions mais un axe de recherche fort sur l'exploitation d'images acquises par des satellites et, en particulier, la nécessité de recalculer ces images.

Ces recherches ont été financées par une bourse du ministère qui a permis le recrutement d'un doctorant. A travers ces travaux, nous avons apportés une contribution scientifique forte par la caractérisation d'une mesure de similarité entre sous-parties de nuages de points 2D. Une seule hypothèse, généralement vérifiée sur des nuages de points issus de cas réels est posée : ceux-ci possèdent des groupes de points qui s'organisent en structures linéiques. Nous avons ainsi proposé une mesure de similarité invariante par rotation, ainsi qu'un algorithme permettant de la calculer. Ces travaux ont donné lieu à des publications [7, 8, 20, 22, 23, 24, 27] et à un dépôt de brevet international [2, 6].

Les résultats obtenus m'ont permis de prolonger ces travaux au Burkina Faso. Dans le cadre de deux financements de l'agence universitaire de la Francophonie, j'ai pu participer à la mise en place d'une collaboration entre l'ISIG, AMU et le LSIS dans le domaine de l'enseignement et de la recherche sur le thème du contrôle des écosystèmes à partir d'images satellitaires.

J'ai également contribué à la mise en place d'une collaboration avec l'IRSA qui m'a permis de prolonger ces travaux et a donné lieu à des publications communes [1, 3, 18]. Nous avons récemment proposé une méthodologie pour extraire un « indice de végétation » à partir d'images satellites en environnement urbain. Je reviendrai plus longuement sur cette collaboration dans le chapitre 2.3.3 présentant la collaboration mise en place en Chine.

### 2.1.3 Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

- **Date :** 2015 - ... en cours
- **Partenaires :** Airbus Helicopters - Laboratoire d'Informatique et Systèmes (LIS)
- **Financement :** bourses CIFRE. Le montant du financement pour un accompagnement CIFRE est de l'ordre de 75 000 euros.
- **Premiers résultats :** Méthodologie pour la détection de points de poser.

J'ai établi un partenariat avec la société Airbus Helicopters qui se concrétise par trois conventions industrielles de formation par la recherche (CIFRE). Une de ces thèses a été soutenue le 01 février 2019 et la soutenance des deux autres est prévue pour le mois de janvier 2021.

Les travaux engagés adressent essentiellement deux préoccupations de la société : l'aide au pilotage et le vol autonome, le vol autonome pouvant être vu comme l'aide ultime au pilote. Sur ce sujet du vol autonome, la phase d'atterrissage est particulièrement critique. Deux disciplines scientifiques complémentaires entrent en jeu pour tenter d'apporter une solution à un atterrissage automatique. D'un côté, la vision par ordinateur permet de reconnaître l'environnement, tandis que de l'autre côté, l'automatique prend en charge le contrôle et le maintien de l'objet volant. Le partenariat qui a été établi entre Airbus et le LIS permet de traiter ces deux aspects du problème.

Une première thèse a démarré sur le sujet en octobre 2015. Les préoccupations étaient essentiellement du domaine de la vision par ordinateur. Le sujet de la thèse est « Détection et caractérisation d'entités en mer à partir d'un Hélicoptère ». Cette thèse a été soutenue le 01 février 2019. Les travaux développés ont permis de définir une méthodologie pour la détection de point de poser. Ces travaux ont fait l'objet de deux publications [15, 17] et d'un brevet déposé par la société Airbus Helicopters. Au delà de l'aspect scientifique, la méthodologie développée a été implémentée et testée en vol sur le système Eagle d'Airbus (système optronique proposant un calculateur et plusieurs optiques). La solution proposée pour la détection automatique de points de poser a pu montrer tout son potentiel dans des situations représentatives en étant à la fois efficace dans la détection, précise, et robuste.



Ces résultats très appréciés m'ont permis d'obtenir le financement de deux autres thèses dès le mois de janvier 2018. Un sujet consiste à poursuivre les travaux engagés en proposant un couplage vision-pilotage. Le titre de cette thèse est « Appontage automatique pour hélicoptère ». La manœuvre d'appontage est sans conteste une des manœuvres les plus difficiles à réaliser pour un pilote. En effet, le pilote doit établir et suivre une trajectoire jusqu'au point de posé (stratégie d'appontage) qui lui permette d'apponter dans les meilleures conditions de sécurité. Pour cela il doit anticiper le mouvement du pont de bateau et gérer précisément la position relative de l'hélicoptère par rapport à l'aire d'appontage (qui est mouvante), tout en respectant les limitations imposées par la machine. La charge de travail du pilote est par conséquent très élevée. Rendre automatique cet appontage est un challenge très ambitieux pour l'automatique, le captage d'informations et le pilotage.

L'appontage automatique d'un hélicoptère consiste à lever plusieurs verrous scientifiques et technologiques :

- Le captage d'informations (concernant plusieurs engins mobiles, la houle, le vent ...) par la vision, le traitement d'images et la fusion de données;
- L'identification et la modélisation de systèmes en interaction avec celui que nous pilotons;
- La définition d'une trajectoire adaptative ajustée en temps réels en fonction de la situation de pilotage;
- L'observation robuste et la prédiction (en présence de perturbations aléatoires) d'autres engins mobiles ou systèmes avec lesquels l'hélicoptère interagit;
- L'asservissement à cette trajectoire malgré les interactions non contrôlées avec d'autres systèmes (houle, bateaux, rafales de vent, ...).

Les premiers travaux engagés dans ce sens sont très encourageants. Une première publication a été acceptée pour présentation à la conférence « 5th CEAS Conference on Guidance, Navigation and Control (EuroGNC) » [14].

Le troisième sujet de recherche porte sur la reconstruction 3D temps réel de type SLAM couplé à la projection d'une lumière structurée. Aujourd'hui un intérêt grandissant se porte sur les applications visant à reconstruire l'environnement 3D d'un l'hélicoptère. Pour un hélicoptère disposant d'un système capable d'effectuer de la reconstruction 3D et de la localisation (SLAM : Simultaneous Localizing and Mapping), le champ des applications possibles deviendrait très vaste. La problématique d'intégration du système de caméras sur l'hélicoptère représente une contrainte pour l'option stéréovision. Afin de rendre cette technologie éligible à tous les types d'appareils, il est important de concevoir une approche « monoculaire » du problème. Les limites des approches de SLAM monoculaire proposées dans la littérature associées aux contraintes du monde de l'hélicoptère ne permettent pas de proposer une solution. De plus, il sera impossible de disposer de caméras produisant de la stéréoscopie comme en photogrammétrie, ni d'une caméra de profondeur. L'utilisation d'une lumière structurée projetée sur l'environnement peut être une solution innovante aux limites du SLAM monoculaire. Cette

## 2.2. Projets spécifiques développés dans un cadre industriel

---

piste est actuellement privilégiée pour ces travaux. On s'intéressera à ce que pourra apporter la reconstruction 3D lors d'une phase d'atterrissage et il sera pertinent d'étudier le cas d'un atterrissage prévu et le cas d'un atterrissage non prévu.

Une première publication a été soumise à la conférence « 20th Towards Autonomous Robotic Systems Conference (TAROS) » [13].

## 2.2 Projets spécifiques développés dans un cadre industriel

### 2.2.1 Rendu des inter-réflexions diffuses dans le cadre d'un scénario « walkthrough »

- **Acronyme du projet :** Virtu'ART
- **Date :** 2009 - 2015
- **Partenaires :** Société EUROCOPTER - Société EADS - Société OPTIS - L'institut des sciences du mouvement (ISM Marey) - Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
- **Financement :** Fonds unique interministériel (FUI) - Labellisation pôle PEGASE. Le montant du financement pour le laboratoire était de l'ordre de 200 000 euros utilisés pour le fonctionnement et recrutement d'un doctorant.
- **Résultats :** Algorithme de rendu 3D temps réel des inter-réflexions diffuses. Licence logicielle. Maturation logicielle.

Le projet Virtu'ART a consisté en la simulation de l'éclairage d'une cabine de pilotage en environnement immersif. Virtu'Art a été financé par le FUI (fonds unique interministériel) et labellisé par le pôle PEGASE (Astronautique) lors de la campagne 2008. Le projet a débuté en janvier 2009 sur une période de trois ans, avec en particulier le financement d'un contrat doctoral (thèse de Florian Graglia). Ce projet met en jeu cinq partenaires qui sont la société OPTIS, la société EADS, la société Eurocopter, l'institut des sciences du mouvement Marey (ISM Marey) et le laboratoire des sciences de l'information et des systèmes (LSIS).

La société Toplink innovation s'est occupée du montage global du dossier. J'ai pris en charge le montage de ce projet pour la partie LSIS. Le travail a été ponctué de nombreuses réunions pour proposer un projet cohérent qui conviennent à toutes les parties.

L'objectif de ce projet était de produire une maquette virtuelle de cabine d'hélicoptère permettant d'appréhender de manière réaliste (au sens physique du terme, et non pas esthétique) les effets et sensations lumineuses par rapport à différents éclairages et avec différents matériaux. Le calcul temps réel des inter-réflexions diffuses est un verrou important pour véhiculer l'information pertinente lors de la visualisation de la maquette numérique. La thèse de Florian Graglia, que j'ai co-encadré, apporte une solution efficace et permet de calculer une image physiquement réaliste qui véhicule cette information notamment concernant le rendu des matériaux de la maquette.

## 2.2. Projets spécifiques développés dans un cadre industriel

---

La société OPTIS a souhaité inclure les travaux réalisés dans son offre logicielle. Nous avons alors convenu de rédiger une publication incluant uniquement certains aspects du travail effectué et de conclure un contrat d'acquisition de licence logicielle. Certains aspects de ces travaux ont ainsi pu être présentés lors de la conférence GRAPP (International Conference on Computer Graphics Theory and Applications)

citegraglia2012real. Concernant le travail relatif à la licence logicielle nommée PhotoMap, j'ai travaillé en collaboration avec la SATT Sud-Est et la société OPTIS à la rédaction de la description du travail sous licence.

Quelques temps plus tard, en début 2015, les travaux réalisés dans ce projet ont fait l'objet d'un contrat de maturation en collaboration avec la société AMD. Dans ce cadre, j'ai accompagné le développement de la technologie et son transfert vers une plateforme propriétaire d'AMD. Le travail a consisté d'une part à assurer un service d'assistance aux développeurs d'AMD au niveau algorithmique et d'autre part à les aider à restructurer une partie de leur « pipeline graphique » afin de pouvoir utiliser sur GPU les cartes de photons prétraitées par PhotoMap.

### 2.2.2 Caméra Intelligente pour la Surveillance Autonome

- **Acronyme du projet :** CISA
- **Date :** 2009 - 2011
- **Partenaires :** société Kaolab - société Nexvision - Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
- **Financement :** Fonds unique interministériel (FUI) - Labellisation pôle OPTITEC. Le montant du financement était de l'ordre de 100 000 euros utilisés pour le fonctionnement et recrutement d'un post-doctorant.
- **Résultats :** Méthode de calibrage d'une caméra « fish-eye ». Etude de brevabilité.

Le projet CISA a consisté en l'analyse de scènes à travers une « Caméra Intelligente pour la Surveillance Autonome ». CISA était un projet financé par le FUI (fonds unique interministériel) et labellisé par le pôle OPTITEC (photonique et imagerie) lors de la campagne 2008, et qui a débuté en mars 2009 sur une période de deux ans, avec en particulier le financement d'un post-doctorant.

Ce projet a mis en jeu trois partenaires qui étaient la société Kaolab, la société Nexvision et le laboratoire des sciences de l'information et des systèmes (LSIS). L'objectif de ce projet était de permettre la détection d'activités suspectes (vidéosurveillance) au moyen d'une caméra dite « 360° ». Le principal verrou technologique était le calibrage de cette caméra. Nous avons mis en place une méthodologie permettant de calibrer une optique fish-eye quelconque et ainsi d'obtenir une « carte de profondeur » de la scène filmée.

Je me suis occupé du montage de ce dossier. Mon travail a consisté à rédiger la partie scientifique de la description du projet, à établir la fiche financière détaillée et à définir les différentes

## 2.2. Projets spécifiques développés dans un cadre industriel

---

tâches du projet et leurs livrables. Afin de proposer un dossier de qualité, plusieurs réunions ont été organisées entre les partenaires. Naturellement, j'ai également rédigé le rapport scientifique final.

Mes travaux n'ont pas donné lieu à des publications car l'enjeu pour les entreprises partenaires était stratégique. Néanmoins, suite à ce projet, une étude de brevetabilité a été réalisée par l'intermédiaire de la SATT Sud-Est. J'ai participé à la rédaction de cette étude et assisté à diverses réunions préparatoires. Bien que cette étude ait conclu à la brevetabilité des travaux, la DGA n'a pas donné son accord immédiat et nous n'avons finalement pas déposé ce brevet.

### 2.2.3 Détection d'objets dans des images acquises par un sonar multi-faisceaux

- **Acronyme du projet :** CAT-SURVEYOR
- **Date :** 2012 - 2014
- **Partenaires :** Société Subsea Tech - Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
- **Financement :** RAPID (DGA). Le montant du financement était de l'ordre de 100 000 euros utilisés pour le fonctionnement et recrutement d'un ingénieur de recherche.
- **Résultats :** Détection de cible dans des images sonar multifaisceaux. Trophée de l'innovation au salon Euronaval.

La surveillance maritime et sous-marine d'infrastructures sensibles telles que ports militaires ou de commerce, zones de mouillage, terminaux pétroliers, centrales nucléaires est un enjeu majeur dans le cadre de la lutte contre les menaces asymétriques. D'autre part, l'augmentation des trafics illicites, de l'immigration clandestine et des pollutions marines ont renforcé le besoin de surveillance des zones côtières. Enfin le coût et les risques liés à la plongée humaine dans tous les secteurs, militaires, sécurité civile, industriels et scientifiques, et notamment en conditions hostiles (conflits, eaux froides, mauvaise visibilité, etc.) demandent de nouvelles solutions pour l'observation sous-marine, avec des moyens plus économiques, plus performants et plus réactifs que la plongée conventionnelle. C'est dans ce contexte que le projet CAT-Surveyor est né. L'objectif était de développer un système de détection et d'identification automatique de cibles sous-marines quelconques en profondeurs modérées (<100 m) utilisant principalement un sonar d'imagerie très haute résolution de dernière génération, embarqué sur un drone de surface (USV) du type catamaran.

Ce projet a été financé dans le cadre des projets RAPID-DGA et a impliqué initialement la société SubSea Tech, la société Kaolab et le laboratoire des sciences de l'information et des systèmes (LSIS). J'ai pris en charge le montage de ce dossier pour ce qui concerne le LSIS. Mon travail a consisté à rédiger la partie scientifique de la description du projet, à établir la fiche financière détaillée et à définir les différentes tâches du projet et leurs livrables. J'ai également rédigé le rapport scientifique final. Le financement obtenu a permis, en plus d'un « classique budget » de fonctionnement, d'embaucher pour une durée de 18 mois un ingénieur

## 2.2. Projets spécifiques développés dans un cadre industriel

---

de recherche.

Suite au dépôt de bilan de la société Kaolab quelques mois après le démarrage du projet, il a été nécessaire de redéfinir les différents « work packages » et naturellement d'obtenir l'aval de la DGA concernant la nouvelle mouture proposée avec deux acteurs au lieu de trois. La restructuration a été validée et le projet a démarré avec beaucoup de retard.

Nous avons néanmoins proposé une solution à la détection automatique de cible dans des images de type sonar multi-faisceaux. J'ai encadré, sur une période de 18 mois, un ingénieur qui a développé la méthodologie que nous avons proposé. Le travail réalisé a été mis à disposition de la société SubSea Tech pour réaliser des présentations techniques. Ces présentations ont été réalisées lors de salons commerciaux à l'UDT (Underwater Defense Technology) en à Hambourg 2013 et à Euromaritime à Paris 2013. La maquette du Cat-Surveyor, réalisée au 1/6ième a été exposée au Salon Milipol à Paris en 2013, au salon Oceanology à Londres en 2014. Le projet a reçu le trophée de l'innovation sur le thème de la sécurité maritime au salon Euronaval en 2014.

### 2.2.4 Caractérisation du mouvement des yeux et de la tête lorsque celle-ci est soumise à des sollicitations externes

- **Acronyme du projet :** STABILITAS
- **Date :** 2017 - ... en cours
- **Partenaires :** Société Synapsys - Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
- **Financement :** Bourse CIFRE. Le montant de l'accompagnement demandé est de l'ordre de 50 000 euros.

Ce travail de recherche a essentiellement une finalité industrielle mais, pour débloquer les verrous technologiques qui lui sont associés, il est nécessaire d'explorer certaines voies de recherche, non pas en utilisant simplement les algorithmes existants et en les améliorant, mais en travaillant sur certains axes fondamentaux. Je vais présenter ici le contexte dans lequel s'inscrit ce travail et les objectifs à atteindre en terme de travail de recherche.

Le programme STABILITAS a notamment pour objectif la conception d'un masque muni d'une caméra vidéo et de capteurs de mouvements. Un tel masque pourrait être utilisé par un personnel non expert, voire dans certains cas par le patient lui-même. En suivant des protocoles spécifiques, un tel masque permettrait la collecte de données de diagnostic et la réalisation de certains traitements thérapeutiques sans erreur ou biais.

L'objectif de ce travail de recherche est ainsi de caractériser ces mouvements et de les mettre en rapport avec les mouvements de la tête qui les ont induits (à travers la perception des accélérations par les canaux semi-circulaires de l'oreille interne), cette caractérisation devant être robuste, précise et effectuée au rythme de l'acquisition des images par la caméra vidéo.

## 2.2. Projets spécifiques développés dans un cadre industriel

---

L'oeil est mobile dans son orbite, en 3D, à la fois en rotation autour de deux axes relatifs « horizontaux » et « verticaux » – mouvement qui est associé à une translation dans l'image – mais aussi autour d'un troisième axe, perpendiculaire aux deux premiers, et c'est ce qu'on appelle le mouvement « torsionnel » de l'oeil – mouvement qui est associé à une rotation autour du centre de la pupille dans l'image.

Dans les lignes suivantes, je vais énoncé un ensemble de problèmes devant être résolus pour que le produit concerné puisse être certifié et commercialisé :

- En matière de précision, il est nécessaire d'établir une corrélation fine entre l'image acquise (2D) et les mouvements respectifs de la tête et de l'œil (tous deux en 3D), car ce sont les mouvements relatifs de ces derniers que l'on doit extraire de l'acquisition vidéo (et non pas uniquement ceux obtenus par l'analyse des images de la séquence); une modélisation géométrique du système « tête –œil » est à envisager pour établir cette corrélation.
- La reconnaissance de points de repère « fixes » (amers) sur le visage, à proximité de l'oeil, est nécessaire mais cette reconnaissance, au niveau de précision requis, n'est pas simple à mettre en œuvre; en effet, si l'on connaît certains invariants du visage, ceux-ci peuvent disparaître, ou être légèrement déplacés, à la fois à travers des effets de reflet sur la peau mais aussi parce que le masque, mal fixé par le patient, peut être légèrement mobile; une voie de recherche consisterait à ne pas utiliser un nombre limité et « sûr » d'amers mais un ensemble d'amers plus important ayant une fiabilité variable et qui seraient traités à travers un processus statistique de mise en correspondance de points (méthodes ICP, ou RANSAC, par exemple).
- La rotation de l'oeil est directement corrélée à celle l'iris, mais les marques présentes dans ce dernier peuvent apparaître ou disparaître au cours de sa rotation; là aussi, un traitement statistique peut être efficace mais d'autres pistes peuvent apporter une solution plus précise, comme l'utilisation de la Transformée de Hough, ou même la caractérisation de la texture.
- La précision requise peut nous amener à travailler au niveau subpixelique; plusieurs pistes sont envisagées pour cela, comme, par exemple, l'utilisation de la cohérence temporelle (une fréquence d'acquisition importante devrait permettre d'obtenir des suites partielles de positions identiques ou presque, et de transformer cela en une évolution « continue » associée à des positions plus précises); nous pouvons aussi nous appuyer sur l'introduction de connaissance et le suivi de certaines formes au moyen de contours actifs (ces contours étant contraints par la connaissance dont nous disposons sur l'entité détournée).
- En termes de robustesse, le principal défi à relever est celui de l'hétérogénéité des conditions d'acquisition des séquences vidéo : en effet, il est important de produire un résultat et que celui-ci soit fiable (et précis) quels que soient la forme des yeux, la couleur de l'iris, le grain de la peau, la conformation des cils et des sourcils, le battement des paupières, la présence de maquillage, ... et l'on doit également être capable de

## 2.3. Projets développés dans un cadre international

---

caractériser des situations non valides (masque ne couvrant pas correctement la région des yeux, ou trop incliné)

- Enfin, il faut intégrer le fait que le masque ne sera pas nécessairement bien positionné (possibilité de biais, de dérive, de rotation, ...).

Ce travail est financé par une convention industrielle de formation par la recherche (CIFRE). La thèse a démarré en juin 2017 et le travail réalisé jusqu'à présent nous conforte dans la proposition d'une approche de reconstruction 3D pilotée par modèle. Deux publications sont en cours de rédaction pour l'année à venir.

## 2.3 Projets développés dans un cadre international

### 2.3.1 Portugal

Il s'agit de la première collaboration qu'il m'ait été donné de mettre en place dans la carrière d'enseignant chercheur. Alors que je venais d'être recruté comme Maître de Conférences (en 2003) et que des relations informelles existaient déjà avec l'Université d'Aveiro, j'ai eu l'occasion d'aller à Aveiro et d'échanger avec un jeune enseignant chercheur de cette université, M. Paulo Dias, qui avait les mêmes centres d'intérêt que moi, en particulier en lien avec le projet Simulfoot. Paulo Dias est venu à Marseille, nous avons échangé régulièrement, et j'ai donc pris l'initiative de constituer et de déposer un dossier de candidature dans le cadre des AUI luso-françaises, de manière à ce que nous puissions disposer de fonds pour effectuer des missions au Portugal (et en France pour les chercheurs portugais de l'Université d'Aveiro). J'ai pu obtenir ce support financier sur la période 2005-2007. Ce support a permis l'échange de chercheurs (dont des doctorants de l'équipe) et le partage de travaux de recherche. Ces travaux ont fait l'objet de publications communes dans des conférences internationales [26, 29].

### 2.3.2 Burkina Faso

Il s'agit d'un projet que je n'ai pas mis en place mais dans lequel je me suis impliqué. Le Professeur Marc Daniel a initié un projet de coopération avec l'ISIG (Institut Supérieur d'Informatique et de Gestion) au Burkina Faso (à Ouagadougou et à Bobo Dioulasso), et ce projet a été repris ensuite par le Professeur Jean Sequeira pour y créer une activité commune dans le domaine de l'imagerie numérique. Des liens ont ainsi été tissés entre cet institut et la composante à laquelle j'appartiens, qui s'appelait l'ESIL à ce moment-là, et qui s'appelle Polytech à présent. J'ai participé à différents niveaux à cette collaboration, en allant enseigner à l'ISIG à Ouagadougou, par exemple, ou en co-encadrant des stages de Master, l'une des idées fortes que nous avons essayé de mettre en place étant l'utilisation de drones (et donc l'analyse d'image, la reconnaissance des formes et la modélisation géométrique) pour traiter des problèmes tels que la gestion de l'agriculture ou des troupeaux en Afrique sub-saharienne. Cette

## 2.3. Projets développés dans un cadre international

---

coopération s'est située sur les années 2009-2012.

### 2.3.3 Chine

Là aussi, je n'ai pas été à l'origine de la collaboration mise en œuvre entre l'IRSA, institut de l'Académie Chinoise des Sciences traitant de télédétection (l'IRSA est devenu le RADI) et notre laboratoire (c'est le Professeur Jean Sequeira qui en est à l'origine) mais j'y ai pris une part très active, et ce depuis 2008, et mon implication dans cette collaboration est allée en croissant.

Cela a commencé par la co-direction d'une thèse (Jean Sequeira étant l'autre co-directeur) – il s'agit de la thèse de Christophe Palmann – sur une problématique qui nous intéressait (Christophe Palmann a eu une bourse du Ministère) et qui intéressait aussi l'IRSA (Institute for Remote Sensing Applications), à savoir « Le recalage d'images de modalités différentes ». Ces travaux ont, par ailleurs fait l'objet de plusieurs publications, mais aussi d'un brevet international (j'en discuterai davantage dans les chapitres suivants).

Ainsi, j'ai reçu au laboratoire et j'ai travaillé avec plusieurs doctorants et post-doctorants chinois faisant partie de l'IRSA (Mrs. Wang Yuxi) mais aussi de l'Université de Zhejiang, à Hangzhou (Du Zhenhong) et depuis deux ans, avec le Professeur Wei Song, dans le cadre d'une collaboration avec l'Université de l'Océan à Shanghai (Shanghai Ocean University). J'ai été co-auteur de plusieurs publications dans ce cadre-là, dont la dernière avec Mrs. Wei Song (une autre communication a été publiée dans une conférence internationale en 2016 – avec une doctorante du RADI, et un article vient d'être proposé pour publication dans une revue internationale, avec un autre doctorant du RADI).

### 2.3.4 Algérie

La collaboration mise en place avec l'Algérie est récente mais s'inscrit dans une dynamique intéressante. Elle a démarré de manière informelle avec la venue du Professeur Salim Ouchtati, de l'Université de Skikda en 2013. Depuis, nous avons fortement interagi et nous nous sommes impliqués dans la participation à la recherche dans le domaine de l'imagerie dans cette université. La dernière action de collaboration à laquelle j'ai participé avec cette université est la participation au comité scientifique d'une conférence qui sera organisée à Annaba à la fin juin 2019.

### 2.3.5 Pérou

Cette collaboration s'est mise en place en 2016, suite à une mission du Professeur Jean Sequeira à l'Université Catholique San Pablo, à Arequipa. J'ai pris en charge la totalité du montage du projet au niveau administratif, et, depuis l'été 2018, un accord-cadre a été signé entre cette université et Aix-Marseille Université. Au-delà de cet accord-cadre, j'ai participé aux discussions et à la mise en œuvre de deux projets de recherche, l'un sur la restauration de vieux



journaux (journaux datant de l'indépendance du Pérou, et dont on souhaite extraire le texte, les images, etc.) et l'autre sur l'analyse de l'évolution de la forêt amazonienne à partir d'images de télédétection (satellites et drone); pour ces deux projets, j'ai participé à la rédaction du dossier de demande de fonds auprès des instances péruviennes, en attendant de lancer, en 2019, un projet STIC-Amsud.

## 2.4 Projet émergent

Les projets décrits dans les sections précédentes sont des projets effectifs, avec ou sans financement, avec des partenaires industriels ou institutionnels, en France ou à l'étranger; et certains de ces projets sont encore « actifs ». Dans cette section, je vais présenter diverses actions que j'ai pu mener au cours de ces 20 dernières années de recherche qui m'ont permis d'acquérir une compétence dans le domaine de la Réalité Virtuelle et Augmentée et qui constituent un socle pour le démarrage d'un nouvel axe fort de recherche.

Dès mes premières années dans la recherche, j'ai eu l'opportunité de m'investir dans ce domaine. J'ai participé activement à plusieurs projets parfois sur des aspects d'enseignement, parfois sur des aspects recherche. Ce parcours commence donc très tôt dans ma carrière. J'ai tout d'abord modestement contribué à la création du Centre de Réalité Virtuelle de la Méditerranée (CRVM) en assistant à plusieurs réunions techniques concernant le choix de l'infrastructure (Graphe de PC, mode projection, système de tracking, offre logicielle). Par la suite, dans le cadre du projet Virtu'ART (voir section 2.2.1), j'ai eu l'opportunité de collaborer avec le CRVM et d'apporter une contribution à la synthèse d'image et plus particulièrement au rendu réaliste des inter-réflexions diffuses [21]. Ces travaux se sont naturellement intégrés dans un des objectifs principaux du centre dont une des vocations est d'améliorer l'interprétation des phénomènes physiques et d'étudier les conséquences sensorimotrices et cognitives de l'immersion en utilisant les flux visuels comme moyen de stimuler ou de perturber le sujet humain.

Dans le cadre des problématiques posées par le projet SimulFoot (voir section 2.1.1), plusieurs travaux sont à inclure dans le cadre de la Réalité Virtuelle :

- Un premier axe de recherche développé dans ce projet concerne l'étude des mécanismes cognitifs mis en place par des joueurs experts lors d'une tâche de prise de décision. Le mode de visualisation des scènes de jeu est alors un facteur essentiel dans la mise en place de ces mécanismes. En particulier, j'ai proposé l'utilisation de différentes abstractions lors de la création des images de synthèse. Ainsi, j'ai produit un rendu 3D des scènes de jeu avec différents niveaux de réalisme (environnement plus ou moins "dépouillé", symbolisme des joueurs, choix du point de vue, ...) qui a permis de mieux comprendre ces mécanismes. Il s'agit à nouveau de s'intéresser aux flux visuels afin de mieux comprendre les mécanismes cognitifs du sujet humain [9].
- Je me suis également impliqué dans l'étude de la communication entre l'utilisateur et l'environnement virtuel, ce qui nous a amené à proposer à l'utilisateur d'agir sur

sa perception de l'environnement. A titre d'exemple, cela lui permet d'explorer son environnement en tournant la tête afin de pouvoir "regarder" autour de lui à l'aide de la souris ou bien en modifiant son point de vue à partir un menu. Ces opérations posent des questions quant à la façon de procéder (clavier, souris, menu, capteur ...) mais également quant à l'impact sur les processus cognitifs mis en oeuvre pour la compréhension de la scène. Il est, par exemple, légitime de se demander si l'analyse d'une scène de jeu est plus "efficace" à partir d'un point de vue externe à la scène (vue des tribunes, vue du dessus, vue du terrain, ...) ou bien à partir une vue "subjective" (à sa place sur le terrain) [39].

J'ai souhaité donner une dimension formation et enseignement à mon projet de Réalité virtuelle et augmentée. Je pense que mêler recherche et enseignement est très profitable aussi bien pour dynamiser les travaux de recherche que pour former des futurs ingénieurs innovants et créatifs. C'est pour cela que j'ai choisi de créer une spécialité de la formation d'ingénieur en informatique qui s'articule autour de la Réalité Virtuelle et Augmentée (spécialité ReVA du département informatique de Polytech Marseille). Cela m'a permis de proposer des projets de fin d'études dans ces domaines et je me suis tout particulièrement intéressé aux différents modes d'interaction en Réalité Virtuelle. J'ai collaboré avec le professeur Marc Daniel sur ce sujet, en particulier sur la manipulation d'un morceau de surface NURBS à l'aide d'environnements immersifs « léger » (Casques de visualisation, capteurs magnétiques, capteurs infra-rouge, accéléromètres ...) dans le but de mener une réflexion sur les usages de la Réalité Virtuelle en modélisation géométrique. Ces travaux ont fait l'objet d'un article intitulé "Réalité virtuelle et manipulation de surfaces B-splines" [19], d'une présentation en mai 2018 à la journée "Modélisation Géométrique et Réalité Virtuelle" organisée à l'ENSAM Paris et d'une présentation en novembre 2018 au colloque annuel d'innovation pédagogique de l'Université d'Aix-Marseille (CIPE).

Concernant plus particulièrement la Réalité Augmentée, c'est à travers l'analyse de scène que j'ai tout d'abord abordé ce domaine. Que ce soit par le projet SimulFoot et les problématiques de segmentation d'images, de recalage d'images, de pose de caméra mais également par « l'enrichissement » des images de football avec des marquages virtuels (pour l'entraînement ou pour l'arbitrage) ; ou bien à travers mes travaux sur le recalage d'images, j'ai pu acquérir un savoir faire essentiel aux problématiques de la réalité augmentée. En prolongement de ces sujets, je suis en train de mettre en place une collaboration à la fois sur le plan de la recherche et sur celui de l'enseignement avec le RADI (institut de l'Académie Chinoise des Sciences dédié aux applications de la télédétection dans le cadre de Digital Earth) [18] avec deux objectifs principaux. Le premier concerne l'exploitation de modèles de terrains, en incluant une analyse temps réels et une communication ergonomique des « résultats » (pollution, minerais, végétaux, ...) dans un environnement de Réalité Augmentée. Ce premier objectif est à un niveau de maturité suffisant pour envisager à la fois des travaux de recherche fondamentale mais également des développements dans le cadre de projets importants (projet "Google

troisième génération" mis en oeuvre par le RADI). Le second objectif est de proposer une méthodologie de "rendu réaliste" pour la Réalité Augmentée. Il s'agira d'utiliser l'information extraite des images pour visualiser conjointement mondes réel et virtuel et ainsi de ne produire plus qu'une seule réalité visuelle. Il est question ici d'être capable d'évaluer à partir d'images l'information relative à l'éclairage d'une scène réelle et d'extraire des paramètres comme, par exemple, l'intensité lumineuse, la couleur, l'emplacement afin de proposer un rendu 3D d'objets virtuels à la scène conformément à ces paramètres.

Au delà de l'émergence d'un projet Réalité Virtuelle ou Augmentée que je compte mener, je voudrais faire état de deux actions à court terme pour l'année 2019. Je viens de déposer une demande de financement CIFRE en partenariat avec la société E2VR. Il s'agit d'un co-encadrement avec le professeur Mohamed Quafafou et le sujet porte sur l'analyse de données pour la personnalisation et la recommandation en environnement virtuel (V-commerce). Je suis également membre du comité d'organisation des journées du GDR IGRV (Informatique Graphique et Réalité Virtuelle) qui auront lieu du 13 au 15 novembre 2019 au Centre des Congrès du Parc Chanot à Marseille. Je suis membre du comité scientifique et en charge des relations avec les sponsors de l'évènement.

## 3 Valorisation de la Recherche

L'adéquation entre mes thématiques de recherche et les projets que j'ai menés m'a permis de valoriser les résultats scientifiques obtenus. Outre les publications scientifiques, mes actions de valorisation ont été menées de trois manières : le brevet, la licence logicielle et la maturation de projet.

### 3.1 Brevet

J'ai activement participé au côté de la structure de valorisation de l'Université (SATT sud-est) à la valorisation de mes travaux sous forme de brevet. Ce travail est différent de celui de rédaction d'articles scientifiques. Le mode de description d'une invention est bien spécifique et le travail avec le cabinet chargé de rédiger les brevets a été très intéressant mais également assez chronophage. Trois travaux ont été concernés : la mise en correspondance de nuages de point, l'analyse d'image sonar à deux dimensions et le calibrage de caméra.

Les travaux réalisés dans le cadre de la thèse de Christophe Palmann sur la mise en correspondance de nuages de point ont donné lieu à un dépôt de brevet intitulé « Procédé et système pour estimer une similarité entre deux images binaire » (brevet n° 115710 déposé le 2 mars 2011). L'invention concerne un procédé et un système pour estimer la ressemblance entre deux images de modalités éventuellement différentes. Plus particulièrement, l'invention permet de caractériser une similarité entre deux images binaires. Elle permet notamment un recalage d'images acquises dans les domaines de télédétection, l'imagerie médicale, la vision industrielle. Le brevet a été étendu à l'international sous les références CN103688289A, US20140044362, WO2012117210A1.

Les travaux, réalisés dans le cadre du projet CAT-SURVEYOR, relatifs à la détection de cibles à partir d'un sonar d'imagerie multifaisceaux ont fait l'objet d'une étude de brevabilité. L'invention s'intitule : « Procédé automatique d'identification d'une ombre générée par une cible réelle dans une image à deux dimensions issue d'un sonar ». La société SubSeaTech, qui s'était

dite intéressée par l'acquisition de ce brevet, n'a pas donné suite. Le brevet a été finalement abandonné car le domaine d'application a été jugé trop restreint.

Enfin, les travaux, réalisés dans le cadre du projet CISA, relatifs au calibrage d'une caméra 360° (fish-eye) ont également fait l'objet d'une étude de brevetabilité. L'accord de la DGA concernant l'autorisation de dépôt de brevet n'a pas été donnée et le dépôt a été abandonné.

### 3.2 Licence logicielle

Les travaux réalisés dans le cadre du projet Virtu'ART ont donné lieu à un contrat de licence entre l'Université de la Méditerranée et la société OPTIS portant sur l'accélération de la phase de calcul d'illumination d'une scène modélisée en 3D. L'algorithme transféré propose une technologie de « photon mapping par polygone » massivement parallèle (PhotoMap). L'aspect massivement parallèle de la méthode permet sa programmation sur les unités de calcul des cartes graphiques et permet ainsi d'augmenter ses performances applicatives. Les codes sources ont fait l'objet d'un dépôt auprès de l'APP (Agence pour la Protection des Programmes). Le travail que j'ai réalisé a consisté en une réécriture des codes sources pour le dépôt et à la rédaction d'une documentation technique pour la cession de licence.

Une licence logicielle exclusive a également été négociée avec la société My Coach Sport. Cette licence porte sur les travaux menés dans le cadre du projet SimulFoot et fait suite à la maturation de projet réalisée en 2014 et 2015. Le domaine d'exploitation est celui du sport.

### 3.3 Maturation de projet

La maturation de projet consiste à amener des résultats de la recherche à un stade de développement suffisamment avancé pour permettre leur transfert vers l'industrie ou la création d'entreprises. Ceci n'est possible que si les résultats de recherche sont suffisamment aboutis pour être transformés en « solution ».

J'ai travaillé sur trois projets de maturation. Deux d'entre eux concernent mes travaux sur l'analyse et la reconstruction 3D de scènes sportives (Projet SimulFoot) et le troisième sur une technologie de « photon mapping par polygones » massivement parallèle (Projet Virtu'Art).

#### 3.3.1 SimulFoot

En 2007, le projet SimulFoot a été sélectionné par la structure de valorisation ValorPACA (aujourd'hui SATT Sud-Est) comme « projet de maturation prioritaire ». L'objectif de cette maturation était de développer un logiciel dédié à la formation des joueurs. Les utilisateurs visés étaient les formateurs des centres de formation. Les fonctionnalités attendues étaient la création de scénarii d'entraînement sous la forme de séquences 3D animées et annotées.

Pour cela, il s'agissait de modéliser en 3D des scènes sportives à partir d'images issues de flux vidéo de caméras non calibrées mais contenant des repères ou amers propres au sport en question. Le sport ciblé était le football. Cette application a été développée par la société SII (agence d'Aix-en-Provence). Mon rôle a été de transférer les algorithmes de traitement d'image vers SII sous la forme d'une « API SimulFoot » et de conseiller les ingénieurs de SII. J'ai également contribué avec le professeur Hubert Ripoll à la définition de l'IHM de façon à ce que l'interface soit adaptée aux besoins spécifiques d'une application dédiée à l'entraînement des joueurs pour les clubs. L'application a été livrée en septembre 2010. Elle fonctionnait sur « tablette PC » et a été présentée au centre national de formation de Clairefontaine lors du colloque « Football et Recherche ». Nous avons eu des retours très positifs des entraîneurs et joueurs. La commercialisation de ce produit n'a malheureusement pas pu eu lieu faute « d'entreprise porteuse ». Ce projet a été financé par ValorPACA (100 000 euros) et par la société SII (financement d'une bourse CIFRE avec un accompagnement de 50 000 euro et mise à disposition d'un ingénieur pour le développement).

En 2015, la SATT Sud-Est m'a à nouveau sollicité pour créer une nouvelle version du logiciel. L'objectif initial était de mettre à jour la version créée en 2007. Toutefois, l'évolution rapide des périphériques et des technologies entre 2007 et 2015, nous a amené à choisir une ré-écriture complète de l'application. C'est la société Inotekk qui a été choisie pour le développement. J'ai à nouveau assuré le suivi du développement et le transfert des algorithmes. Après des discussions autour de la création d'une start-up autour de ce produit, c'est finalement la société « My Coach Sports » qui a acquis une licence d'exploitation exclusive de la technologie en février 2017. Le montant investi par la structure de valorisation a été de l'ordre de 150 000 euros.

Ce travail a été passionnant et parfois frustrant. Il est très valorisant d'amener des travaux fondamentaux jusqu'au développement d'une application réelle. Dans ce cas, Il ne s'agissait pas de « simplement transférer » des travaux issus de la recherche dans une technologie ou une application existante mais de créer un nouveau produit, une nouvelle activité autour de ces recherches pluridisciplinaire en informatique et en sciences cognitives. Cependant, un produit n'est rien sans sa commercialisation et c'est cette étape qui a été tardivement au rendez-vous.

#### 3.3.2 PhotoMap

En 2013, J'ai été contacté par la société AMD qui avait « repéré » les travaux réalisés lors de la thèse de Florian Graglia (projet Virtu'ART) dans le « catalogue technologique » de la SATT-Est. Après plusieurs rencontres, nous avons décidé de lancer un projet de maturation pour étudier la possibilité d'intégrer ces travaux dans le moteur de rendu propriétaire d'AMD (Pro Render). Les algorithmes ont été repris par les ingénieurs d'AMD pour les tester et les adapter à l'architecture des cartes professionnelles d'AMD. Mon rôle a été d'expliquer les algorithmes et les optimisations possibles. J'ai également fourni le code source qui avait été écrit pour la

### **3.3. Maturation de projet**

---

licence cédée à la société OPTIS (voir section 3.2). Ce projet a été financé à la fois par la SATT et par AMD (50 000 euros chacun). De façon un peu cavalière, il n'y a pas eu de cession de licence à la fin du projet.

## 4 Quelques travaux de recherche significatifs

Dans ce chapitre, je vais présenter les contributions scientifiques apportées à mes trois projets principaux que j'ai présenté à la section 2.1. Ces trois projets portent sur :

- La reconstruction 3D de scènes à partir de séquences vidéo
- L'analyse et recalage d'images de modalités différentes
- L'analyse, la reconstruction et l'aide au pilotage à partir de caméras embarquées sur un aéronef

Une section est dédiée à chacun de ces projets. Les deux premiers ont déjà fait l'objet de nombreuses publications. Plutôt que de plagier ces documents, j'ai choisi d'inclure pour chaque projet, une publication pertinente dans sa version originale. Sur le sujet de la reconstruction 3D de scènes, il s'agit d'un chapitre de livre de la collection IC2 traitement du signal et de l'image [5] et pour l'analyse et le recalage d'images d'une publication dans la revue Traitement du Signal [7].

Concernant le troisième thème, les premiers résultats ont été obtenues dans le cadre du travail de thèse de Pierre Zoppitelli qui a été soutenue le 01 février 2019. Je vais reprendre ces résultats en les plaçant dans un cadre méthodologique qui permet de faire évoluer le travail réalisé pour répondre à des contextes relativement différents comme, par exemple, le vol en environnements urbains ou le traitement de scènes de nuit.



## 4.1 Reconstruction 3D de scènes à partir de séquences vidéo

### Table des matières

<b>Chapitre 1. Reconstruction 3D de scènes sportives</b> . . . . .	5
Sébastien MAVROMATIS, Jean SEQUEIRA	
1.1. Introduction . . . . .	5
1.2. Sélection automatique de la région d'intérêt . . . . .	6
1.2.1. Intérêt et caractéristiques d'une région d'intérêt . . . . .	6
1.2.2. Segmentation de l'espace couleur . . . . .	6
1.2.3. Cohérence spatiale . . . . .	9
1.3. Extraction de primitives par Transformée de Hough . . . . .	9
1.3.1. Détection d'arcs d'ellipses . . . . .	10
1.4. Mise en correspondance primitives/modèle . . . . .	11
1.4.1. Faisceaux de droites . . . . .	12
1.5. Conclusion . . . . .	15
1.6. Bibliographie . . . . .	16

# Chapitre 1

## Reconstruction 3D de scènes sportives

### 1.1. Introduction

Il est commun aujourd'hui d'utiliser des applications faisant intervenir l'analyse d'image, la modélisation géométrique ou la vision par ordinateur pour la diffusion d'évènements sportifs. Ces applications proposent, par exemple, l'annotation de ces séquences ou encore la sélection automatiquement des moments forts d'un évènement afin de permettre son accès a posteriori [YAN 08, EKI 03, REN 09]. Parmi ces applications, celles qui permettent la reconstruction en trois dimensions de scènes sportives sont encore très rares. Ceci s'explique par les nombreux verrous scientifiques induits par des conditions d'acquisition très variées (éclairage, apparence du fond, tenues des joueurs ...) ou encore l'absence d'information liée aux paramètres d'acquisition qui entraîne l'utilisation d'un flux vidéo "brut" (paramètres de la caméra inconnus). De plus, l'utilisation de ces applications lors d'évènements en direct ne s'envisage que si la reconstruction peut-être réalisée en temps réel (éventuellement en "léger différé"). Pour pallier ces difficultés, des systèmes complexes sont proposés. Ils imposent l'installation de plusieurs caméras dédiées ce qui, au delà d'un coût important, est parfois difficile à mettre en oeuvre et ne permet pas la reconstruction automatique espérée [SU13, RED13, ORA13]. Certains projets de recherche se focalisent sur l'automatisation du processus de reconstruction à partir de caméras non-calibrées et proposent des solutions innovantes [TRI13, LET 06, LLU 11, KIL 07, GRA 07].

Ce chapitre présente les étapes clés du processus de reconstruction en se focalisant sur certains "points difficiles". Une première partie traite de l'analyse d'images couleur en vue de sélectionner automatiquement la surface de jeu. Dans une seconde

partie, l'extraction du marquage dans l'aire de jeu sera abordé en s'appuyant sur l'utilisation de la transformée de Hough. Une troisième et dernière partie traite de la mise en correspondance des primitives extraites des images avec celles du modèle de la scène.

### 1.2. Sélection automatique de la région d'intérêt

#### 1.2.1. Intérêt et caractéristiques d'une région d'intérêt

La région d'intérêt (ROI) d'une scène sportive est naturellement l'aire de jeu. Très souvent, cette aire de jeu peut-être caractérisée comme étant à la fois une région de teinte globalement uniforme et la plus grande composante "presque connexe" de l'image considérée (figure 1.1).

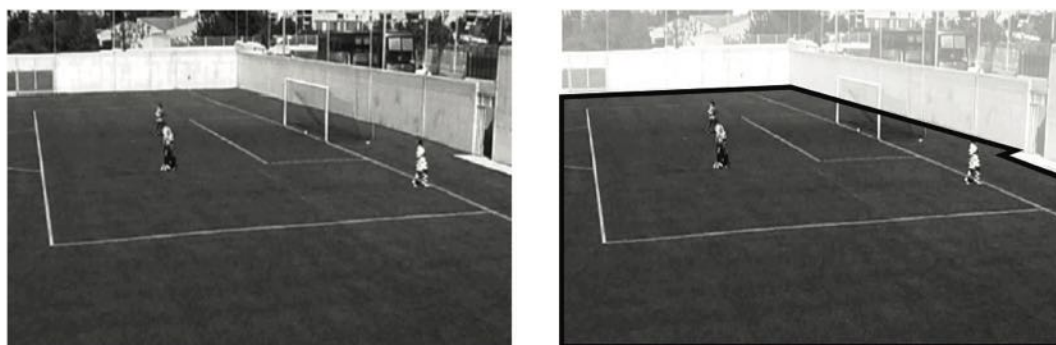
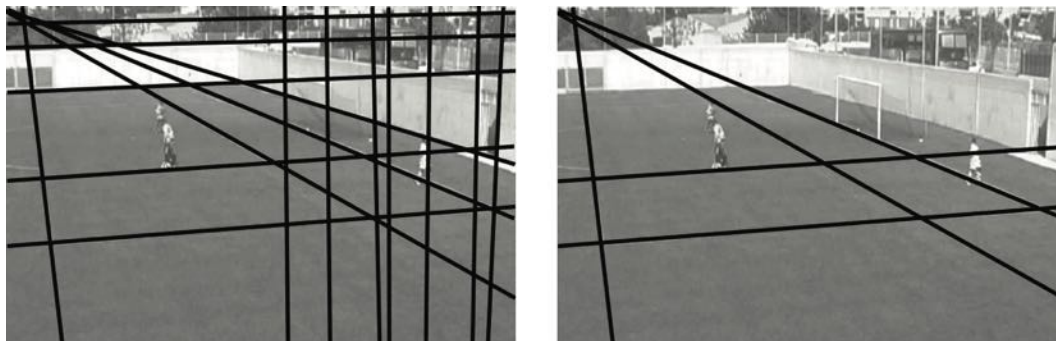


Figure 1.1. ROI d'une scène de football

La caractérisation d'une région d'intérêt est une étape importante du processus de reconstruction. En effet, il est fréquent d'observer dans les images à traiter "l'environnement" de l'aire de jeu : tribunes, piste d'athlétisme, personnel de sécurité, publicités, etc . . . Comme le montre la figure 1.2, la connaissance de la ROI permet une validation immédiate des "éléments" présents sur l'aire de jeu.

#### 1.2.2. Segmentation de l'espace couleur

La détermination de la ROI qui est essentiellement définie par la notion de "cohérence de couleur" peut paraître triviale. En fait, l'homogénéité de couleur n'est qu'apparente. En effet, la représentation discrète de la couleur, peu importe l'espace de représentation choisi, fournit des composantes discrètes qui ne sont pas suffisamment semblables. De nombreux auteurs se sont attachés à proposer une approche permettant de caractériser la région de l'espace couleur qui supporte cette homogénéité [DAI 95, LUC 01, LEF 00, VAN 00]. L'intérêt de l'utilisation d'un espace tridimensionnel pour segmenter une image couleur a été plusieurs fois démontré [CAR 95].



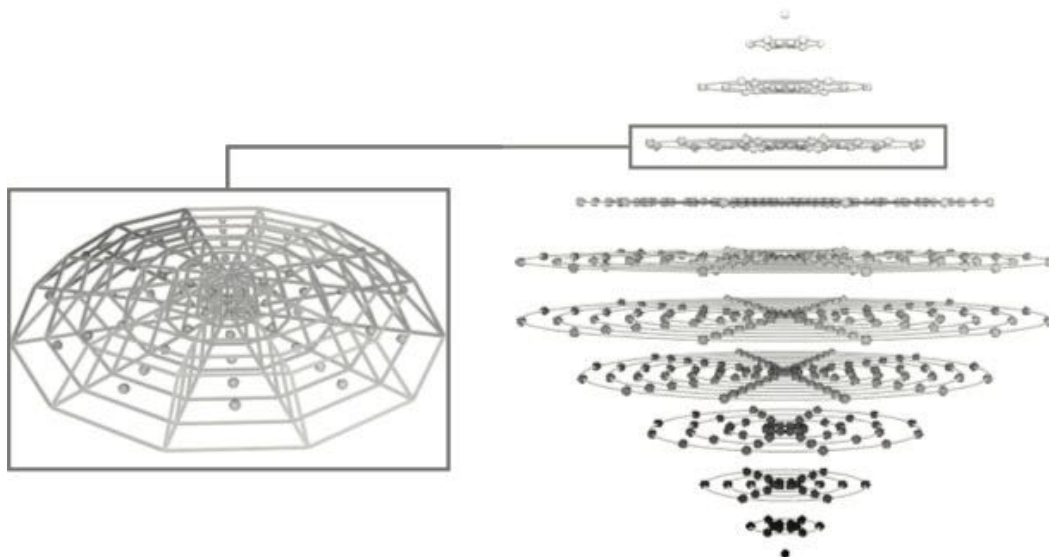
**Figure 1.2.** *Sélection automatique du marquage du terrain sans ROI (gauche) et avec ROI (droite)*

Dans cette partie, l'espace de représentation HLS sera choisi pour illustrer une méthodologie de sélection entièrement automatique des attributs colorimétriques de la région d'intérêt.

La méthodologie présentée ici s'apparente, par de nombreux aspects à celle de fenêtre de Parzen [PAR 62]. Elle consiste à regrouper dans l'espace de représentation les couleurs similaires, et séparer entre eux les nuages de couleurs foncièrement différentes, de manière à isoler et extraire autant de nuages "denses" qu'il y a de régions de couleur dominantes dans l'image à analyser : c'est ce que nous appellerons le critère de cohérence de couleur. L'efficacité de cette approche réside dans la représentation continue de la distribution des couleurs des pixels de l'image [LET 04]. Le contexte informatique engendre une représentation discrète des espaces couleurs. Cette discrétisation induit des difficultés quand à l'analyse de la distribution des couleurs associées aux pixels d'une image.

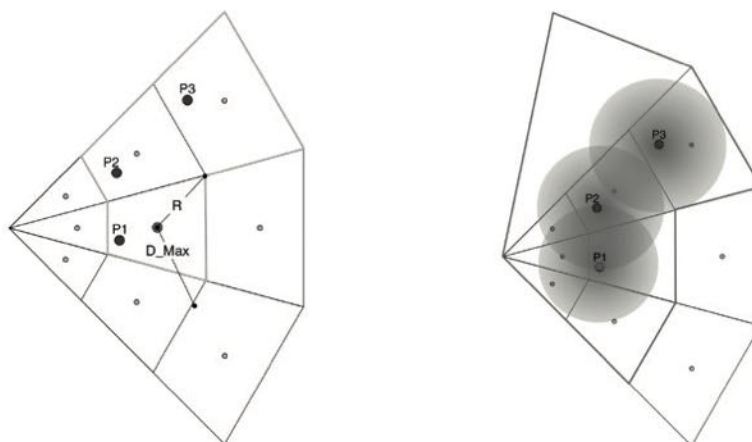
Une première idée consiste à calculer l'histogramme des couleurs de l'image et à ne conserver que les occurrences les plus nombreuses. Evidemment le résultat obtenu est peu satisfaisant. Néanmoins, c'est à partir de cet histogramme qu'il est possible de construire une représentation continue et efficace. Dans un premier temps, il s'agit de morceler l'espace des couleurs en cellules représentant grossièrement une même couleur (figure 1.3).

Dans un second temps, l'utilisation d'une fonction potentiel dans cet espace grossier permet de définir un espace continu de représentation des couleurs. La fonction potentiel a pour objet d'établir des liens entre points voisins mais aussi entre régions proches (par exemple, différents "verts") tout en séparant les nuages caractéristiques de chaque couleur dominante dans l'image. Pour cela, chaque point de l'espace couleur (correspondant à un pixel de l'image) est considéré comme une source de potentiel. La fonction potentiel peut-être une simple fonction affine décroissance de la distance à ce point, qui a pour valeur 1 en ce point et dont le support est un disque de



**Figure 1.3.** A gauche, une tranche de l'espace HLS. A droite, une partition de cet espace en 554 cellules

rayon  $D_{max} = R * (1 + \epsilon)$  où  $R$  représente le "rayon" de la cellule considérée et  $\epsilon$  une valeur d'absorption.



**Figure 1.4.** Influence des points  $P_1, P_2, P_3$  sur les cellules voisines

Les régions de couleur dominantes conservées sont obtenues par seuillage dans l'espace HLS modifié par la fonction potentiel. Le seuil peut-être obtenu automatiquement en segmentant l'histogramme cumulé des valeurs de potentiels. Dans le cas où plusieurs régions de couleurs dominantes sont présentes dans l'image, elle peuvent-être

identifiées à l'aide d'une propagation à partir des relations de voisinage qui existent dans la partition grossière choisie pour l'espace HLS.

La figure 1.5 présente le résultat d'une sélection entièrement automatique sur une scène de football. Ce résultat fait apparaître la région recherchée et il permet également de mettre en évidence le marquage de l'aire de jeu qu'il sera nécessaire d'extraire en vue d'une reconstruction 3D.



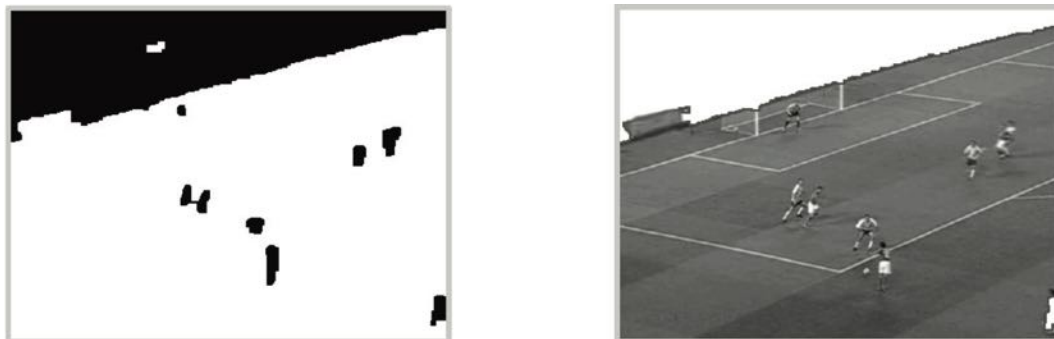
**Figure 1.5.** A droite, en blanc, les pixels retenus par l'analyse colorimétrique

### 1.2.3. Cohérence spatiale

Bien que la sélection de la ROI s'appuie essentiellement sur l'utilisation de la notion de couleur, il est nécessaire de tout de même prendre en compte la cohérence spatiale de cette région. A l'issue du traitement basé sur la notion de couleur, le masque binaire obtenu peut éventuellement contenir de petites régions associées à des éléments de même couleur dans l'image (panneau publicitaire, banderole, drapeau ...). La ROI est considérée comme "la plus grande composante connexe de l'image". En réalité, la région est "presque connexe" : il reste de petits "isthmes" ou "détroits" qu'il faut combler sans pour autant "connecter" des éléments extérieurs au terrain. De plus, certains éléments dans la ROI ne sont pas de couleur identique à la ROI, par exemple, le marquage de l'aire de jeu n'est évidemment pas de la même couleur ! Il est naturel de s'appuyer sur les opérateurs élémentaires de morphologie mathématique [SER 88] pour supprimer ces éléments. Après un remplissage de la plus grande région, la ROI est sélectionnée (figure 1.6).

### 1.3. Extraction de primitives par Transformée de Hough

Il existe de nombreuses méthodes permettant d'extraire des éléments caractéristiques d'une image. Il n'est pas question, dans cette partie, de proposer une revue complète de ces approches. En ce qui concerne, les scènes sportives, le marquage



**Figure 1.6.** *A gauche, l'image binaire de la figure 1.5 après fermeture et ouverture. A droite, la ROI sélectionnée.*

de l'aire de jeu est souvent constitué de segments de droites et d'arcs d'ellipses. Il est alors naturel d'envisager l'utilisation de la transformée de Hough [ILL 88] pour caractériser ces primitives géométriques simples.

La détection de droites à l'aide de la transformée de Hough est un cas d'école. Néanmoins, la détection peut s'avérer délicate tout particulièrement lorsque les segments sont sécants, bruités ou en partie occultés. Dans les scènes sportives, il est fréquent de détecter des droites épaisses et occultées en partie par des joueurs. Dans ce cas, il est primordial de sélectionner pertinemment parmi les points caractéristiques retenus les couples qui participent au vote dans l'espace de Hough. La figure 1.7 présente ce type d'image et montre que la détection du marquage du terrain est efficace.



**Figure 1.7.** *A gauche, les résultats de l'analyse colorimétrique. A droite, la détection des droites après suppression des joueurs*

### 1.3.1. Détection d'arcs d'ellipses

La détection d'ellipses ou d'arcs d'ellipses est une opération plus délicate que la détection de droites. Certains utilisent des ellipses variables s'adaptant aux données

selon divers critères [MAT 00, KAN 01] avec plus ou moins de succès. Cette partie montre comment utiliser la transformée de Hough dans ce but.

L'équation cartésienne la plus générale d'une ellipse s'exprime au moyen de six paramètres reliés par une relation de proportionnalité, ce qui signifie qu'une ellipse se caractérise par cinq paramètres effectifs. Il n'est pas envisageable d'utiliser une transformation de Hough de "5 à 1" pour une telle détection pour des raisons de complexité (espace des paramètres de dimension 5), de manque de précision et d'homogénéité de l'espace des paramètres. De nombreux travaux se sont appuyés sur la décomposition de l'espace de Hough en plusieurs sous-espaces [YUE 89, PAO 93, AGU 95, HO 95].

Une propriété caractéristique des tangentes d'une ellipse est : "Soient  $M_1$  et  $M_2$  deux points de l'ellipse,  $T_1$  et  $T_2$  les tangentes en ces points,  $I$  le point d'intersection de  $T_1$  et  $T_2$ ,  $J$  le milieu de  $M_1M_2$ , alors la droite  $(IJ)$  passe par le centre  $O$  de l'ellipse". En utilisant cette propriété, il est possible de découpler la recherche en deux recherches, l'une en dimension 2 et l'autre en dimension 3. A partir de cette décomposition, les principales étapes de la détection sont :

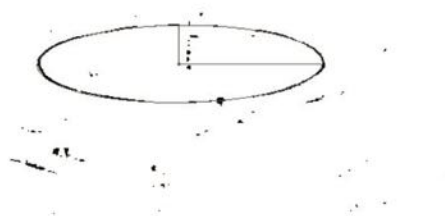
- Caractérisation des couples (point, tangente) pertinents : il s'agit, pour chaque point, de déterminer, à partir de ses voisins, s'il est possible de lui associer une "tangente estimée".
- Recherche du centre des ellipses : à partir de l'ensemble des couples (point, tangente) ainsi retenus et à partir de la propriété évoquée précédemment, il est possible de détecter le centre des ellipses dont les arcs apparaissent sur l'image.
- Calcul des paramètres  $a, b, \theta$  : l'ellipse ayant été recentrée, il existe une expression en fonction de l'angle  $\theta$  fait par son grand axe avec l'axe  $Ox$  et des paramètres  $a$  et  $b$  de son expression normalisée ( $a$  et  $b$  étant les demi-longueurs du grand axe et du petit axe).
- Calcul de l'arc d'ellipse : ayant détecté une ellipse, l'étude de la distribution des points associés à celle-ci permet d'extraire les arcs d'ellipse correspondants.

Les figures 1.8 et 1.9 montrent les résultats obtenus. Sur la figure 1.8, l'ellipse détectée est formée par quatre arcs d'ellipses.

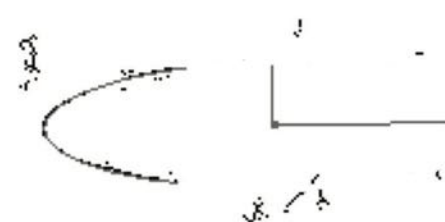
### 1.4. Mise en correspondance primitives/modèle

La mise en correspondance d'éléments caractéristiques entre une image et un modèle géométrique est un processus particulièrement délicat car il implique la comparaison entre des éléments de natures différentes : les pixels de l'image et les éléments du modèle. Dans la reconstruction de scènes sportives, on trouve naturellement des travaux visant à mettre en correspondance le marquage de l'aire de jeu, souvent des droites et des ellipses comme introduit dans la partie 1.3, avec ces mêmes éléments définis dans le modèle de la scène à reconstruire [OKU 04, FAR 05, HAY 05, THO 07].





**Figure 1.8.** Détection du "rond central"



**Figure 1.9.** Détection de l'arc des 18 mètres

Les images à traiter sont une représentation partielle du modèle. Ainsi, les éléments détectés forment souvent un sous-ensemble très restreint de l'ensemble des primitives définies dans le modèle. La mise en correspondance de ces éléments avec ceux du modèle peut-être guidée par l'utilisation de mesures de similarité basées sur la connaissance du modèle : positions relatives des droites, distances, dimensions. Ce type de solution est trop fortement liée aux caractéristiques du modèle pour pouvoir être exploité dans un cadre plus général. Néanmoins, une caractéristique commune à beaucoup de terrains de sport est l'utilisation de lignes parallèles comme "marques". Il est intéressant de retrouver cette information parmi les éléments détectés afin de l'injecter dans l'étape la mise en correspondance de ces éléments avec ceux du modèle. La partie suivante propose une méthode géométrique de sélection des faisceaux de droites dans une image.

### 1.4.1. Faisceaux de droites

La détection de faisceaux de droites est souvent apparentée à la détection de points de fuite dans la littérature. Un problème intrinsèque à la recherche des points de fuite dans une image est lié à l'espace de recherche de ces derniers. Ce problème peut-être

écarté par l'introduction de méthodes basées sur l'utilisation de la sphère de Gauss comme espace de recherche [BAR 83, MAG 84]. Elles utilisent le plus souvent des approches de type accumulation, ce qui entraîne d'autres difficultés (discrétisation, recherche de maxima, etc.) [SHU 99]. Ainsi, pour être robustes, ces méthodes impliquent des calculs lourds, l'ajout de connaissances ou encore l'utilisation de modèles d'incertitude. De plus, les méthodes de type accumulation font l'hypothèse implicite que le nombre d'occurrences du phénomène observé (droites qui forment un faisceau) est important ; hypothèse qui n'est pas garantie dans les images issues de scènes sportives. En effet, ces images sont des projections du modèle (le terrain) qui est souvent constitué de deux groupes de droites parallèles (groupes formés pas peu de droites).

Il est possible d'éviter l'utilisation d'espace d'accumulation à l'aide de propriétés géométriques. Identifier les différents faisceaux de droite d'une image est alors possible en deux temps. Dans un premier temps, il s'agit de déterminer tous les triplets de droites ; dans un second temps, ces triplets sont fusionnés pour former l'ensemble des faisceaux.

En s'appuyant sur le cercle de Thalès, il est possible d'énoncer le corollaire suivant : **Corollaire (Thalès)** : Soient trois droites  $D_1, D_2$  et  $D_3$ . Soit  $A$  et  $A_1, A_2$  et  $A_3$  ses projetés orthogonaux respectivement sur  $D_1, D_2$  et  $D_3$ . Si le cercle formé par  $A_1, A_2$  et  $A_3$  passe par  $A$  alors c'est un cercle de Thalès. Ainsi, les droites  $D_1, D_2$  et  $D_3$  sont sécantes en un même point.

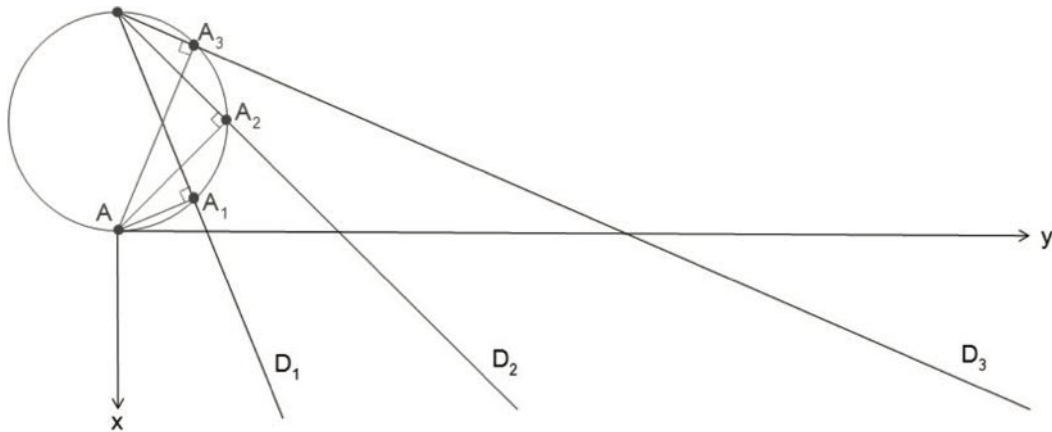


Figure 1.10. Le point  $A$  et ses projetés forment un cercle

Déterminer si trois droites sont concourantes revient alors, d'après le corollaire de Thalès, à vérifier si un point  $A$  et ses projetés sur ces trois droites sont sur un même cercle ;  $A$  étant un point quelconque de l'image qui n'appartient à aucune droite. En effet, si  $A$  appartient à une droite (ou s'il est "proche" d'une droite) alors le corollaire de Thalès est nécessairement vrai (trois points forment un cercle) même dans le cas de trois droites non concourantes. Pour éviter ces cas dégénérés, le point  $A$  peut, par

exemple, être choisi dans l'image de telle sorte que ce soit le point ayant la distance maximale avec son projeté orthogonal le plus proche (sur une des droites).

Vérifier que le point  $A$  et ses projetés sont cocycliques n'est pas complexe mais peut-être source de problèmes numériques. Afin d'éviter le calcul explicite des paramètres du cercle, il est possible d'utiliser le théorème de Ptolémée qui implique uniquement le calcul de distance entre des points (projetés orthogonaux de  $A$  sur les différentes droites de l'image (ou dans un voisinage proche)).

**Théorème de Ptolémée :** un quadrilatère convexe est inscriptible si et seulement si le produit des longueurs des diagonales est égal à la somme des produits des longueurs des cotés opposés.

Il est intéressant de remarquer que les points de fuite infinis sont "naturellement" traités. En effet, le cercle de Thalès formé par un ensemble de droites (participant à un point de fuite infini) correspond à un "cercle plat" et elles vérifient donc le théorème de Ptolémée.

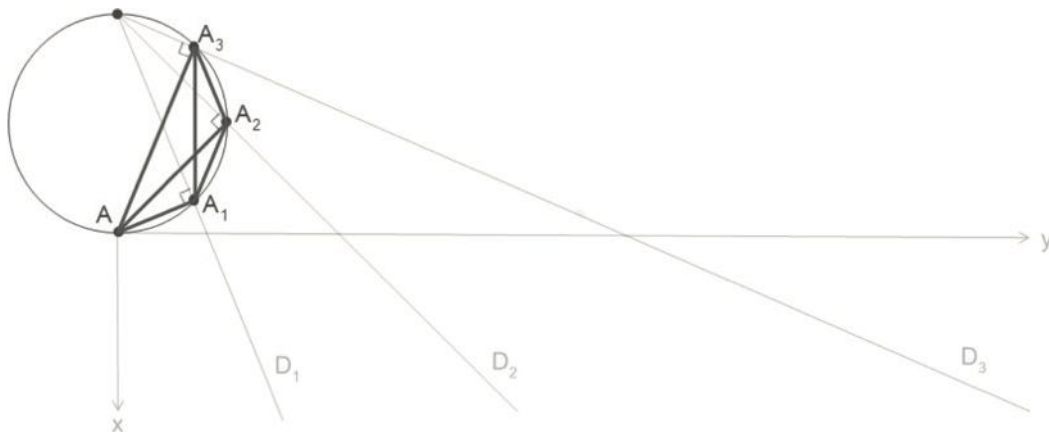


Figure 1.11. Le quadrilatère  $AA_1A_2A_3$  est inscriptible.

Afin d'identifier les triplets de droites participant au même faisceau, il est nécessaire de déterminer un critère de similarité. Une idée intuitive consiste à comparer les cercles de Thalès définis par ces triplets de droites. Mais ceci nécessite le calcul des paramètres de ces cercles qu'il a été possible d'éviter. Une autre idée serait d'utiliser un espace d'accumulation, pour faire émerger les cercles de Thalès équivalents mais cela a peu de sens au regard du nombre limité de cercles présents dans les images. En raisonnant sur les droites communes entre triplets, l'identification des triplets à d'éventuels faisceaux est possible. En effet, si un triplet de droites concourantes n'a de droite commune avec aucun des autres triplets, alors ce triplet correspond à un faisceau composé uniquement de trois droites. Sinon, si un triplet de droites concourantes

a deux droites en commun avec d'autres triplets, alors chacune des droites de ces triplets participe au même faisceau. Dans la mesure où le nombre de triplet est restreint, il est possible de simplement combiner les triplets entre eux pour vérifier ces deux propriétés et donc les fusionner le cas échéant.

La figure 1.12 illustre le résultat de la détection de deux faisceaux de droites dans l'image binaire présentée. Une droite issue d'une erreur de détection est ainsi rejetée. En plus, d'apporter une information supplémentaire aux droites, cette étape permet donc également de "valider" l'ensemble des "primitives candidates" à l'appariement.



**Figure 1.12.** *A gauche, l'ensemble des droites détectées. A droite, deux faisceaux sont retenus et la droite tracée en noir est exclue.*

L'information associée aux faisceaux de droites permet généralement de classer les droites candidates à l'appariement en deux groupes. Ceci facilite grandement l'étape d'appariement dans la mesure où le problème se ramène alors à rechercher parmi un ensemble réduit de droites parallèles "qui est qui?".

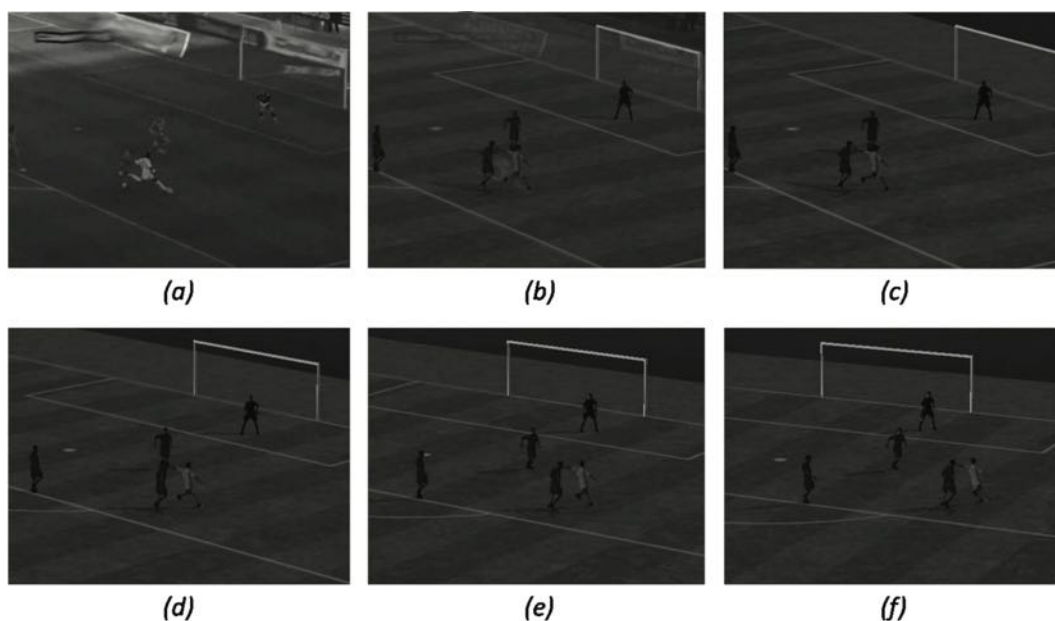
### 1.5. Conclusion

La segmentation de séquences issues de scènes de sports collectifs reste un verrou majeur dans la mise en place d'un dispositif de reconstruction totalement automatique. Ces séquences contiennent un fond et un éclairage variable, du flou de mouvement, des distorsions liées aux caméras, ce qui explique aisément les nombreuses difficultés rencontrées. L'identification et le suivi de plusieurs joueurs sur le terrain reste également très difficile, à la fois à cause des difficultés liées à la segmentation mais aussi de la présence d'occultations multiples tout au long de la séquence. En ce qui concerne la diffusion télévisée, la valeur ajoutée apportée par le suivi de joueurs ne justifie pas le coût important lié à l'installation de caméras spécifiques et à l'emploi d'opérateurs dédiés même si des informations comme le positionnement des joueurs ou les distances parcourues tout au long d'une rencontre sont des données intéressantes pour une analyse post-match.

Ce chapitre a montré qu'il est possible d'automatiser certaines étapes du processus de reconstruction. Les méthodes qui sont proposées sont robustes, automatiques et

fonctionnent en "temps réel". Elles ont été mises en oeuvre dans un prototype logiciel qui a été diffusé et utilisé par plusieurs centres d'entraînement de football [POP 08, PET 08, SFP13]. Néanmoins, il est difficile d'envisager une reconstruction totalement automatique sans une infrastructure matérielle qui proposerait plusieurs vues d'une même surface de jeu afin de régler par exemple le problème des occultations entre joueurs. Bien que ces systèmes existent déjà, ils n'utilisent que trop peu d'algorithmes performants et proposent "uniquement" une reconstruction supervisée d'une courte séquence d'image.

La figure 1.13 montre un mouvement circulaire, à une hauteur constante, réalisé à partir de la position de la caméra : la première image (a) est celle issue de la caméra, l'image (b) est un fondu entre cette dernière et celle de synthèse (c), i.e l'image reconstruite en 3D. Les images (d), (e) et (f) sont des images de synthèse. Ces images sont générées par une rotation séquentielle de la caméra virtuelle de  $-10^\circ$  autour de l'axe vertical qui passe par le projeté sur le terrain du centre de l'image.



**Figure 1.13.** *Reconstruction d'une scène réelle*

La Figure 1.14 est une illustration d'une situation réelle traitée par le logiciel SimulFoot Entraînement [SF13] : une image est sélectionnée à partir d'une vidéo (2a), puis elle est reconstruite de façon réaliste (2b) et symbolique (2c). Cette représentation virtuelle peut alors être visionnée de différents points de vue, tels que ceux des joueurs (2d). Elle peut également être rejouée telle qu'elle s'est déroulée ou avec une issue différente.

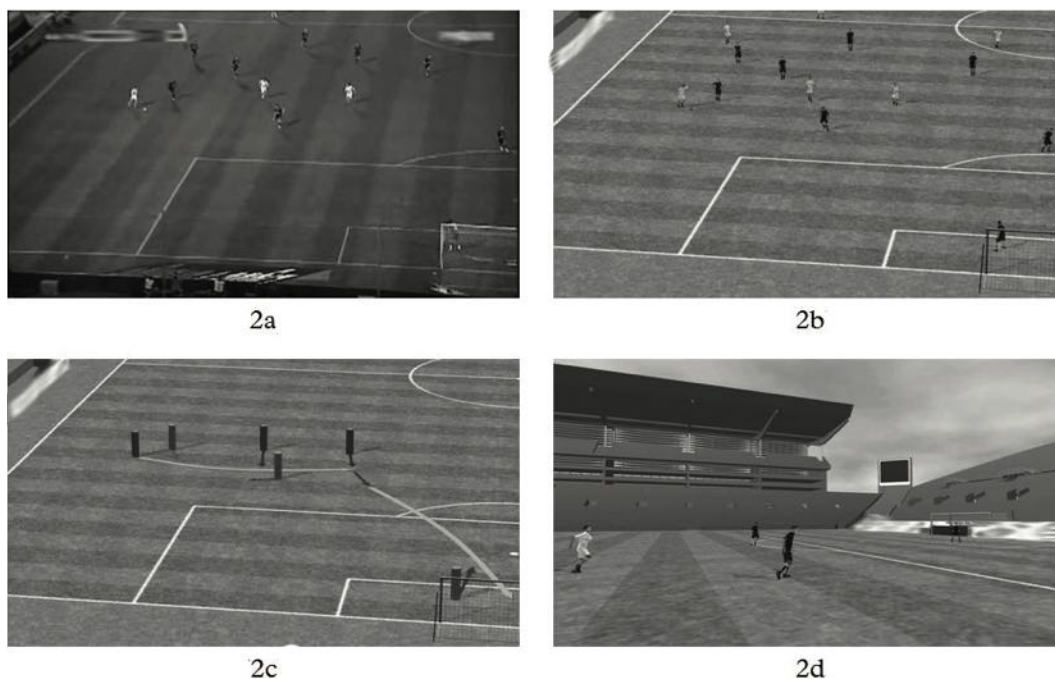


Figure 1.14. Différentes vues d'une reconstruction

### 1.6. Bibliographie

- [AGU 95] AGUADO A., MONTIEL M., NIXON M., « Ellipse Extraction via Gradient Direction in the Hough Transform », *Proc IEE 5th International Conference on Image Processing and its Applications*, p. 375—378, July 1995, Organisation : IEE Address : Edinburgh, UK.
- [BAR 83] BARNARD S. T., « Interpreting Perspective Images », *Artificial Intell.*, vol. 21, p. 435–462, 1983.
- [BUS 05] BUSIN L., VANDENBROUCKE N., MACAIRE L., POSTAIRE J., « Color Space Selection for Unsupervised Colour Image Segmentation by Analysis of Connectedness Properties », *International Journal of Robotics and Automation*, vol. 20, n°2, p. 70–77, 2005.
- [CAR 95] CARRON T., Segmentations d'images couleur dans la base teinte-luminance-saturation : approche numérique et symbolique, PhD thesis, Chambéry, Grenoble, 1995, Th. : électronique, électrotechnique, automatique.
- [DAI 95] DAILIANAS A., ALLEN R. B., PAUL E., « Comparison of Automatic Video Segmentation Algorithms », *In SPIE Photonics West*, p. 2–16, 1995.
- [EKI 03] EKIN A., TEKALP A. M., « Automatic Soccer Video Analysis and Summarization », *IEEE Trans. on Image Processing*, vol. 12, p. 796–807, 2003.
- [FAR 05] FARIN D., HAN J., DE P. H. N., « Fast camera-calibration for the analysis of sports sequences », *In Proceedings IEEE International Conference on Multimedia and Expo (ICME)*, p. 482–485, 2005.

- [GRA 07] GRAU O., HILTON A., KILNER J., MILLER G., SARGEANT T., STARCK J., « A Free-Viewpoint Video System for Visualisation of Sport Scenes », *Motion Imaging*, vol. 116, n°5–6, p. 213–219, SMPTE, May 2007.
- [HAY 05] BERNARD HAYET J., PIATER J. H., VERLY J. G., « Fast 2D model-to-image registration using vanishing points for sports video analysis », *In : ICIP 2005. Proc. of IEEE Int. Conf. on Image Processing*, p. 417–420, 2005.
- [HO 95] HO C.-T., CHEN L.-H., « A fast ellipse/circle detector using geometric symmetry », *Pattern Recognition*, vol. 28, n°1, p. 117-124, 1995.
- [ILL 88] ILLINGWORTH J., KITTLER J., « A survey of the Hough transform », *Comput. Vision Graph. Image Process.*, vol. 44, n°1, p. 87–116, Academic Press Professional, Inc., août 1988.
- [KAN 01] KANATANI K.-I., OHTA N., « Automatic Detection of Circular Objects by Ellipse Growing », *Memoirs of the Faculty of Engineering, Okayama University*, vol. 36, n°1, p. 107–116, décembre 2001.
- [KIL 07] KILNER J., STARCK J., HILTON A., GRAU O., « Dual-Mode Deformable Models for Free-Viewpoint Video of Sports Events », *3DIM '07 : Proceedings of the Sixth International Conference on 3-D Digital Imaging and Modeling*, Washington, DC, USA, IEEE Computer Society, p. 177–184, 2007.
- [LEF 00] LEFÈVRE S., HOLLER J., VINCENT N., « Real time temporal segmentation of compressed and uncompressed dynamic colour image sequences », *In International Workshop on Real Time Image Sequence Analysis*, p. 56–62, 2000.
- [LET 04] LE TROTTER A., MAVROMATIS S., SEQUEIRA J., « Soccer field detection in video images using color and spatial coherence », *ICIAI - International Conference on Image Analysis and Recognition*, vol. ICIAI 2004 - Vol 2, p. 265–272, 2004.
- [LET 06] LE TROTTER A., Mise en correspondance d'images et de modèles : application à la reconstruction 3D de scènes sportives, PhD thesis, Université Aix-Marseille II, september 2006.
- [LLU 11] LLUCIA L., Suivi d'objets à partir d'images issues de caméras mobiles non calibrées, PhD thesis, Université Aix-Marseille II, february 2011.
- [LUC 01] LUCHESE L., MITRA S. K., « Color image segmentation : a state-of-the-art survey », *Proceedings of the of the Indian National Science Academy (INSA-A)*, vol. 67, p. 207–211, 2001.
- [MAG 84] MAGEE M. J., AGGARWAL J. K., « Determining vanishing points from perspective images. », *Computer Vision, Graphics, and Image Processing*, , n°2, p. 256-267, 1984.
- [MAT 00] MATEI B., MEER P., « Reduction of bias in maximum likelihood ellipse fitting », *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 3, p. 794-798 vol.3, 2000.
- [OKU 04] OKUMA K., LITTLE J. J., LOWE D. G., « Automatic rectification of long image sequences », *Asian Conference on Computer Vision*, 2004.
- [ORA13] « ORAD - 3D Replay », <http://www.orad.tv/products/3dreplay>, 2013, Dernière visite le 31/01/2013.

- [PAO 93] PAO D. C. W., LI H. F., JAYAKUMAR R., « A decomposable parameter space for the detection of ellipses », *Pattern Recognition Letters*, vol. 14, n°12, p. 951-958, 1993.
- [PAR 62] PARZEN E., « On Estimation of a Probability Density Function and Mode », *The Annals of Mathematical Statistics*, vol. 33, n°3, p. 1065–1076, 1962.
- [PET 08] PETIT J.-P., RIPOLL H., « Scene Perception and Decision Making in Sport Simulation : A Masked Priming Investigation », *International Journal of Sport Psychology*, vol. 39, n°1, p. 1–19, 2008.
- [POP 08] POPLU G., RIPOLL H., MAVROMATIS S., J. B., « How do expert soccer players encode visual information in order to make decisions in simulated game situations ? », *Research Quarterly for Exercise and Sport*, vol. 79, n°3, p. 392–398, 2008.
- [RED13] « Red Bee Media - Piero », <http://www.redbeemedia.com/piero/piero>, 2013, Dernière visite le 31/01/2013.
- [REN 09] REN R., JOSE J., « General Highlight Detection in Sport Videos », HUET B., SMEATON A., MAYER-PATEL K., AVRITHIS Y., Eds., *Advances in Multimedia Modeling*, vol. 5371 de *Lecture Notes in Computer Science*, p. 27-38, Springer Berlin Heidelberg, 2009.
- [SER 88] SERRA J., Ed., *Image analysis and mathematical morphology. Volume 2. , Theoretical advances*, Academic Press, London, San Diego, New-York, 1988, Autre(s) tirage(s) : 1992.
- [SF13] « SimulFoot », [http://sebastien.mavromatis.perso.luminy.univ-amu.fr/SimulFoot/docs/assets/SimulFoot\\_08.pdf](http://sebastien.mavromatis.perso.luminy.univ-amu.fr/SimulFoot/docs/assets/SimulFoot_08.pdf), 2013, Dernière visite le 31/01/2013.
- [SFP13] « SimulFoot », [http://valorpaca.univmed.fr/IMG/fiches~marketing/fiche~marketing\\_simulfoot.pdf](http://valorpaca.univmed.fr/IMG/fiches~marketing/fiche~marketing_simulfoot.pdf), 2013, Dernière visite le 31/01/2013.
- [SHU 99] SHUFELT J. A., « Performance Evaluation and Analysis of Vanishing Point Detection Techniques », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, n°3, p. 282-288, IEEE Computer Society, 1999.
- [SPO13] « Sportstec », <http://www.sportstec.com>, 2013, Dernière visite le 31/01/2013.
- [SU13] « Sport Universal », <http://www.sport-universal.com>, 2013, Dernière visite le 31/01/2013.
- [THO 07] THOMAS G. A., « Real-time camera tracking using sports pitch markings », *J. Real-Time Image Processing*, vol. 2, n°2-3, p. 117-132, 2007.
- [TRI13] « TricTrac », <http://www.multitel.be/trictrac/?mod=1>, 2013, Dernière visite le 31/01/2013.
- [VAN 00] VANDENBROUCKE N., Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football, PhD thesis, Université de Lille 1, 2000.
- [YAN 08] YAN F., CHRISTMAS W., KITTLER J., « Layered Data Association Using Graph-Theoretic Formulation with Application to Tennis Ball Tracking in Monocular Sequences », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, n°10, p. 1814 - 1830, oct. 2008.



20 Vidéo et TVHD 3D

[YUE 89] YUEN H. K., ILLINGWORTH J., KITTLER J., « Detecting partially occluded ellipses using the Hough transform », *Image Vision Comput.*, p. 31-37, 1989.

## 4.2 Analyse et recalage d'images de modalités différentes

---

# Mesure de similarité entre sous-parties de nuages de points 2D

Christophe Palmann – Sébastien Mavromatis – Jean Sequeira

*Projet SimGraph*  
LSIS UMR CNRS 6168 – ESIL case 925  
163, av. de Luminy  
13288 Marseille Cedex 9  
{palmann,mavromatis,sequeira}@univmed.fr

*RÉSUMÉ.* Cette communication porte sur la caractérisation d'une mesure de similarité entre sous-parties de nuages de points 2D. Cette mesure est définie à partir d'une hypothèse généralement vérifiée sur des nuages de points issus de cas réels : ceux-ci possèdent des groupes de points qui s'organisent en structures linéiques, et qui apparaissent en même temps dans les différents nuages. Après avoir défini des primitives qui permettent une représentation unifiée de ces structures, nous montrons le lien qui existe entre la présence d'une information commune entre sous-nuages et la distribution des relations géométriques entre leurs primitives. Nous donnons alors une mesure de similarité invariante par rotation, ainsi qu'un algorithme permettant de la calculer.

*ABSTRACT.* This communication focuses on the characterisation of a similarity measure between parts of 2D point clouds. This measure is defined thanks to the use of a general knowledge about real point clouds : they share a large amount of one dimensional structures. These structures can be represented into an unified manner with a new type of primitives ; then, we set the link between the existence of a common information between parts of point clouds and the geometric relations of their primitives. Thus, we define a similarity measure that is rotationally invariant, and an algorithm to compute it.

*MOTS-CLÉS :* Nuages de points, mesure de similarité, recalage, reconnaissance de formes.

*KEYWORDS:* Point clouds, measure of similarity, registration process, pattern recognition.

---

DOI :10.3166/TS.27.p1-n © 2010 Lavoisier, Paris

2 Traitement du signal. Volume 27 – n° 4/2010

### 1. Introduction

Le recalage de deux nuages de points dans un espace de dimension quelconque est une problématique majeure en reconnaissance des formes, en particulier dans le cadre de l'analyse d'images. Il s'agit de caractériser une transformation géométrique mettant en correspondance les points de ces deux nuages sachant que ces derniers, modulo la transformation recherchée, ne sont pas absolument identiques mais perturbés par du bruit. Cela peut être énoncé sous la forme suivante : certains points sont « communs aux deux nuages » (information commune) et, dans chacun de ces deux nuages, le sous-ensemble complémentaire de celui de ces points « communs » est formé de points « spécifiques à ces derniers » (que l'on peut qualifier de bruit si l'on se réfère à l'information commune).

Dans cet article, nous nous intéressons au cas où ces nuages de points se trouvent dans un espace 2D. Ce cas particulier est néanmoins très important car il permet d'aborder, d'une manière générale, le problème du recalage d'images lorsque celles-ci sont représentées – sous forme binaire – par leurs points caractéristiques (points de contours, par exemple).

La plupart des problèmes de recalage de nuages de points 2D supposent que les deux nuages se correspondent à travers une transformation rigide (translation et rotation) – avec, parfois, certains éléments de déformation complémentaires (à travers une homographie, par exemple). Nous ferons ici une hypothèse moins restrictive en supposant que la correspondance à travers une transformation rigide est seulement locale (et non plus globale sur les deux nuages tout entiers) : cela nous permet de prendre en compte, en particulier, des déformations complexes, telles que celles produites par le relief sur des images aéroportées en télédétection par exemple.

Notre objectif ne sera plus, alors, de caractériser une transformation rigide, mais de mettre en évidence des paires d'amers sous la forme de positions, dans chacune des images, dont les voisinages respectifs se correspondent à une transformation rigide près : cela nous amènera à définir une mesure de similarité entre deux localisations (à travers leurs voisinage), cette mesure permettant de quantifier l'information commune aux deux voisinages associés.

Des méthodes classiques traitent de la mise en correspondance de nuages de points. C'est le cas en particulier des algorithmes de type ICP « Iterative Closest Point » (Besl et McKay, 1992) ou encore RANSAC « RANdom Sample Consensus » (Fischler et Bolles, 1981) : ces approches permettent d'effectuer une mise en correspondance des nuages à un niveau global, mais pas de quantifier la ressemblance entre deux localisations dans ces nuages, et, par ailleurs, ils montrent certaines limites lorsque les données sont bruitées.

En fait, la comparaison de deux sous-nuages de points (voisinages respectifs de deux positions, chacune dans un nuage) est rendue difficile par la présence de bruit au sein de ces sous-nuages. Nous avons donc décidé, pour apporter une solution à ce problème, d'introduire un élément de connaissance dans le processus d'identification à travers une hypothèse supplémentaire (généralement vérifiée et qui correspond à un élément perceptuel d'évaluation de la ressemblance de ces sous-nuages) : nous supposons qu'il existe des structures linéiques sous-jacentes, qui se retrouvent sous la forme d'agglomérations (sous des formes diverses) de points, et que nous représenterons sous une forme unifiée. C'est à travers des primitives qui caractériseront localement ces structures linéiques que nous pourrons procéder à l'identification (ou plutôt à la mesure de similarité) de ces deux sous-nuages (ou voisinages).

### 2. Caractérisation d'une information commune entre nuages

Dans un premier temps, nous allons parler d'information commune entre nuages de points, alors qu'il s'agit en fait de sous-nuages de points. Dans l'absolu, nuages et sous-nuages représentent la même entité, mais dans la problématique à laquelle nous nous intéressons, cela est complètement différent car nous supposons que nous ne nous appuyons pas sur un volume de données très important (pour les sous-nuages).

De même, comme l'avons indiqué précédemment, nous ne nous intéressons pas directement à un problème de recalage mais à la similarité entre régions, à une transformation près. Néanmoins, pour des raisons de compréhension de notre démarche, nous démarrerons sur un problème de recalage entre nuage de points, et nous convergerons ainsi vers la méthode que nous proposons.

La figure 1 ci-dessous représente deux nuages de points 2D associés à une même scène.

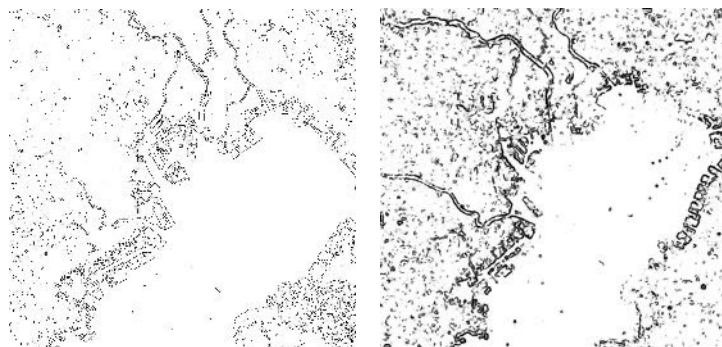


Figure 1. Nuages de points 2D, contenant des objets de différentes complexités.

4 Traitement du signal. Volume 27 – n° 4/2010

Les points que contiennent ces deux nuages ne sont pas distribués uniformément mais forment des objets de complexités différentes. En particulier, il existe des sous-ensembles de points qui s'organisent en structures linéiques. Ces structures linéiques ont la propriété d'être communes, dans une large mesure, aux deux nuages. En effet, bien que celles-ci puissent apparaître sous une forme plus ou moins altérée, ces structures sont bien rattachées à la même information de par leurs formes globales et leurs positions relatives. Elles peuvent être recalées par une transformation rigide (rotation et translation 2D).

Ces remarques sont importantes, car au-delà de l'exemple ci-dessus (figure 1), elles concernent la plupart des paires de nuages issus de situations réelles : des structures linéiques apparaissent dans ces derniers ; ces structures linéiques leur sont, en grande partie, communes et sont l'expression la plus générale d'une information commune entre nuages de points, indépendamment de la manière dont ceux-ci sont obtenus. L'existence de telles structures correspond donc à une connaissance qui sera utilisée pour définir une mesure de similarité entre leurs sous-parties (sous-nuages).

### 3. Modélisation de structures linéiques par de nouvelles primitives

Les structures linéiques sur lesquelles nous souhaitons nous appuyer peuvent apparaître sous des formes assez diverses à un niveau local (voir figure 2) :

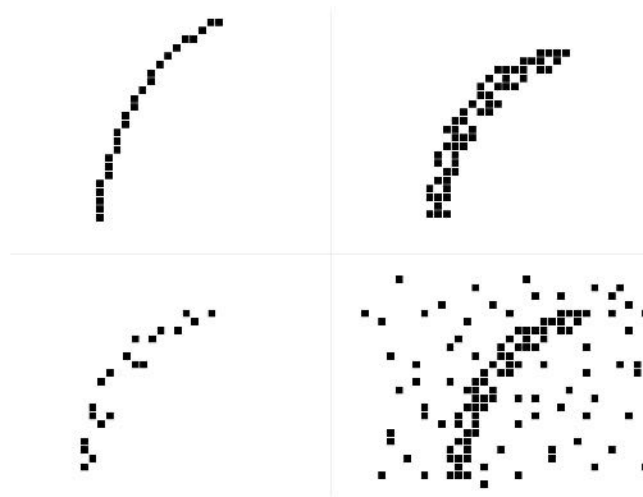


Figure 2. En haut, de gauche à droite : structure fine, structure épaisse ; en bas, de gauche à droite : structure non-connexe, présence de bruit.

Nous cherchons à représenter cette information de manière uniforme à l'aide d'une modélisation adaptée. Nous ne souhaitons pas représenter ces structures linéiques dans leur intégralité (du reste, une telle structure peut apparaître entièrement dans un nuage et partiellement dans l'autre, ce qui empêcherait l'identification), mais, plus simplement par leurs propriétés locales. Or, ces propriétés sont essentiellement une position et une orientation (on pourrait aussi associer une propriété de courbure, mais on complexifierait de trop la représentation – et on verra par la suite que ce niveau de représentation est suffisant).

La caractérisation locale d'une structure linéique sera donc représentée par sa position et l'orientation de l'alignement de points autour de cette position. Cela nous amène à définir des primitives sous la forme d'un triplet  $((x,y),\theta)$ , avec  $(x,y)$  donnant la position d'une primitive et  $\theta$  son orientation ( $\theta$  prenant ses valeurs parmi  $n$  éléments discrets dans  $[0^\circ,180^\circ]$ ). Ces primitives forment alors un dénominateur commun entre les différentes expressions possibles de structures linéiques (fines, épaisses, lacunaires, bruitées, ...). Des algorithmes existent pour réaliser facilement l'extraction de telles primitives (par exemple Palmann *et al.* -2009-).

L'étude d'un grand nombre de relations géométriques entre primitives, à la manière d'une transformée de Hough (1962), peut nous permettre de caractériser la présence d'une information commune entre deux sous-nuages. Plus précisément, cette caractérisation peut être faite grâce à la distribution des différences d'orientations de primitives. Le lien entre cette distribution et la présence d'une information commune entre deux sous-nuages est présenté dans la section suivante.

### 4. Distribution des différences d'orientations

L'objet de cette section est d'étudier la distribution des différences d'orientations de primitives. Cette étude a été faite de façon expérimentale en contrôlant les paramètres de différentes simulations et de façon théorique.

Si deux sous-nuages se correspondent à travers une rotation d'angle  $\delta$ , alors les orientations de primitives en correspondance sont séparées par le même angle ; en fonction de la proportion de primitives en correspondance, le nombre d'occurrences en  $\delta$  est donc plus ou moins important : il s'agit du signal recherché pour caractériser la présence d'une information commune.

Plus précisément, supposons que nous ayons deux régions  $R_i^A$  et  $R_j^B$ , chacune dans deux nuages de points A et B (voir figure 3). Supposons aussi que ces deux régions soient en correspondance par le biais d'une rotation d'angle  $\delta$ . Ces deux régions contiennent chacune  $k$  primitives<sup>1</sup> :  $k_b$  sont des primitives propres à chaque région, et  $k_s$  sont des primitives communes aux deux régions (entourées en

---

<sup>1</sup> La généralisation à deux listes de primitives de tailles  $k_1$  et  $k_2$  est cependant aisée, en posant  $k^2=k_1.k_2$

6 Traitement du signal. Volume 27 – n° 4/2010

pointillés). L'histogramme des différences d'orientations des primitives comporte  $n$  valeurs possibles, nombre qui est directement dépendant de la précision des orientations. Chaque fois qu'une différence d'orientations concerne deux primitives en correspondance, celle-ci prend la valeur  $\delta$ . Les autres différences se répartissent aléatoirement ; il est même possible qu'elles valent  $\delta$ , mais de manière « accidentelle » dans ce cas-là (voir figure 3).

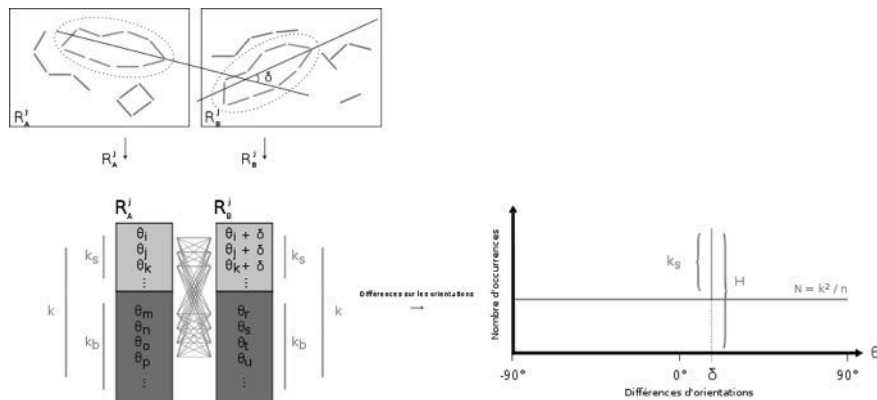


Figure 3. Deux régions possédant des primitives communes entourées en pointillés; histogramme des différences d'orientations prises deux-à-deux.

L'émergence d'un maximum  $H$  dans l'histogramme des différences d'orientations est donc caractéristique de la présence d'une information commune entre deux sous-nuages. De plus, ce maximum apparaît indépendamment de la valeur de l'angle de rotation entre les deux sous-nuages, s'il existe.

C'est l'émergence de ce pic qui nous permettra d'évaluer si les deux sous-nuages ont une similarité forte ou pas. Mais cette émergence, qui est présentée initialement comme une propriété évidente, n'est pas simple à caractériser de manière robuste et dépend en fait des paramètres suivants : nombre de primitives  $k$ , précision des orientations  $n$ , et taux de contamination  $\varepsilon$ , que nous définissons comme le rapport du nombre de primitives propres sur le nombre de primitives dans chaque région ( $\varepsilon = k_b/k$ ).

Nous illustrons tout d'abord le lien entre la robustesse et ces paramètres de façon expérimentale. Pour cela, nous tirons aléatoirement deux listes d'orientations en contrôlant la valeur de chaque paramètre séparément ; les orientations en correspondance sont décalées d'un angle  $\delta$  fixé arbitrairement à  $10^\circ$ . Dans la première simulation, les paramètres ont été fixés avec les valeurs suivantes :  $k=500$ ,  $n=180$  (précision  $1^\circ$ ), et  $\varepsilon=0\%$  (*i.e.* toutes les orientations sont en correspondance deux à deux et décalées d'un angle  $\delta$ ). L'histogramme des différences d'orientations

est présenté dans la figure 4. Comme on peut le constater, le maximum est bien situé en  $10^\circ$ , et ceci sans ambiguïté.

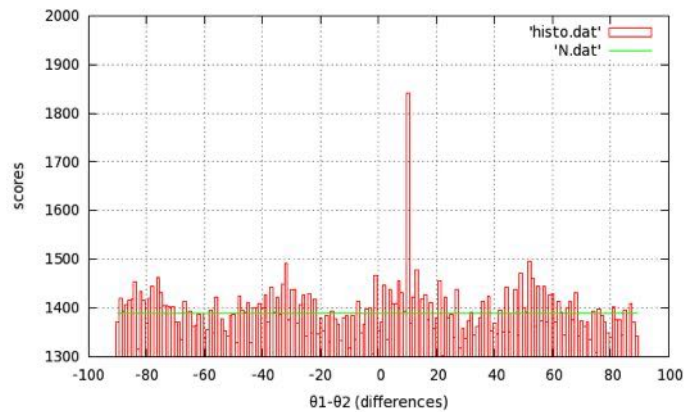


Figure 4.  $k=500$ ,  $n=180$  ( $1^\circ$ ),  $\varepsilon=0\%$ . Maximum correctement placé en  $10^\circ$ .

Dans la simulation suivante, nous fixons  $\varepsilon=50\%$  (seule la moitié des orientations sont décalées deux à deux d'un angle  $\delta$ ),  $k$  et  $n$  restant identiques par rapport au cas précédent (500 et 180 respectivement). L'histogramme des différences d'orientations est présenté dans la figure 5. Le maximum est toujours correctement placé en  $10^\circ$ , mais émerge de façon moins évidente que dans le cas précédent ; l'augmentation du taux de contamination n'est donc pas favorable à l'identification du signal recherché, comme cela était prévisible.

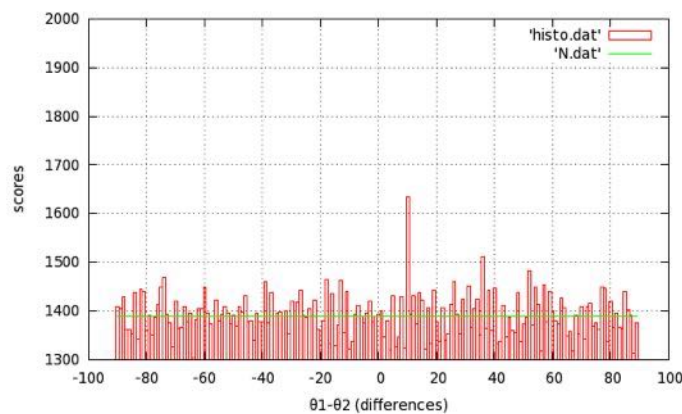


Figure 5.  $k=500$ ,  $n=180$  ( $1^\circ$ ),  $\varepsilon=50\%$ . Maximum correctement placé en  $10^\circ$ .



### 8 Traitement du signal. Volume 27 – n° 4/2010

Dans la simulation suivante, nous fixons cette fois  $n=36$  (précision égale à  $5^\circ$ ),  $k$  et  $\varepsilon$  restant identiques par rapport au cas précédent (500 et 50% respectivement). L'histogramme des différences d'orientations est présenté dans la figure 6. Cette fois-ci, le maximum est situé en  $0^\circ$  et non  $10^\circ$  : un faux-positif a été obtenu. La dégradation de la précision des orientations est défavorable à l'identification du signal pertinent.

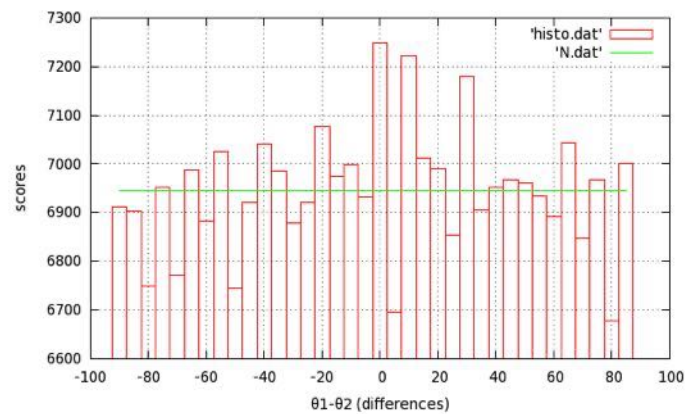


Figure 6.  $k=500$ ,  $n=36$  ( $5^\circ$ ),  $\varepsilon=50\%$ . Obtention d'un faux-positif en  $0^\circ$ .

Enfin, dans la dernière simulation, nous souhaitons illustrer l'influence du nombre de primitives sur l'émergence du signal recherché. Pour cela, nous procédons à deux simulations où  $k=500$  puis 50, et où  $n=180$  et  $\varepsilon=80\%$  à chaque fois (les valeurs particulières de  $n$  et de  $\varepsilon$  ont été fixées à l'aide de résultats théoriques que nous présentons tout de suite après). Les deux histogrammes sont représentés dans la figure 7. Dans le cas  $k=500$ , le maximum est correctement situé en  $10^\circ$ , tandis que lorsque  $k=50$ , à nouveau nous obtenons un faux-positif. Un faible nombre de primitives ne favorise donc pas l'émergence du signal recherché. Ce cas est intéressant puisque nous pouvons nous trouver dans cette situation dans la pratique : en effet, lorsque nous cherchons à établir la similarité entre voisinages, nous ne travaillons que sur de petits sous-ensembles de primitives.

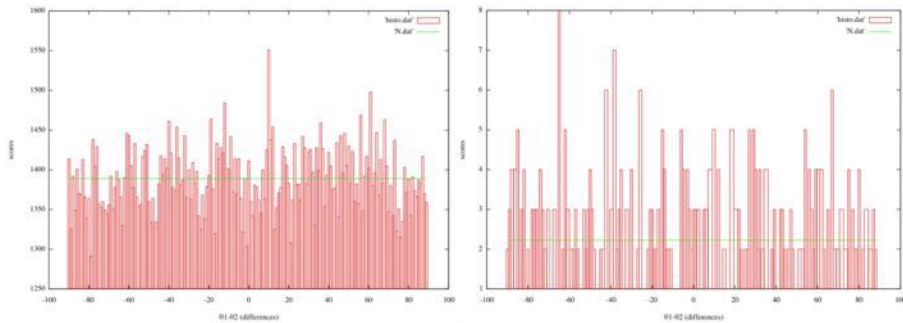


Figure 7.  $k=500$  (à gauche) et  $k=50$  (à droite),  $n=180$  ( $1^\circ$ ),  $\varepsilon=80\%$ . Maximum correctement placé en  $10^\circ$  dans le cas  $k=500$  ; obtention d'un faux-positif en  $-65^\circ$  dans le cas  $k=50$ .

Nous avons par ailleurs étudié de façon théorique l'influence du nombre de primitives  $k$ , de la précision des orientations  $n$ , et du taux de contamination  $\varepsilon$ , sur l'émergence du signal pertinent. En effet, il est possible de donner un majorant très proche de la probabilité  $P$  pour qu'une cellule issue du bruit ait un score supérieur à celui du signal recherché (et donc de fausser la caractérisation d'une information commune), en considérant qu'une telle cellule est incrémentée avec une probabilité constante et que son score suit une loi binomiale (Grimson et Huttenlocher -1990-, Stewart -1995-). Nous avons obtenu la formule suivante :

$$P(\varepsilon, k, n) = e^{\left( (1-\varepsilon) \cdot k + \frac{k^2}{n} \right) \cdot \log\left( \frac{k^2}{n \cdot \left( (1-\varepsilon) \cdot k + \frac{k^2}{n} \right)} \right) + (1-\varepsilon) \cdot k} \quad [1]$$

Cette probabilité est strictement décroissante en fonction de  $k$  et  $n$  et est strictement croissante en fonction de  $\varepsilon$ . Comme cela était attendu, les conclusions de ces deux approches ont été les mêmes. L'étude théorique nous a de plus permis d'apprendre que le risque d'obtention d'un faux-positif, à taux de contamination fixé, décroît plus rapidement en fonction de  $n$  que de  $k$ .

Dans la section suivante, les résultats de cette première étude vont être utilisés dans la définition de la mesure de similarité.

**5. Définition et mise en œuvre d'une mesure de similarité entre nuages de points**

La première étape consiste à éliminer le risque statistique en utilisant l'équation [1] établie lors de l'étude théorique de la robustesse. En effet, en fixant un taux de contamination  $\varepsilon$  arbitrairement grand pour faire face aux situations les plus défavorables, il est possible de déterminer  $n$  en fonction de  $k$  pour que la probabilité d'obtenir un faux-positif soit inférieure à un seuil  $P_0$ . Dans la figure 8, nous avons représenté à titre d'exemple la probabilité d'obtenir un faux-positif en fonction de  $k$  et  $n$  pour un taux de contamination  $\varepsilon = 80\%$ . Dans cette même figure, nous avons fixé  $P_0$  à  $10^{-4}$ ; la courbe verte représente alors les précisions requises sur les orientations en fonction de leur nombre pour une identification robuste du signal recherché.

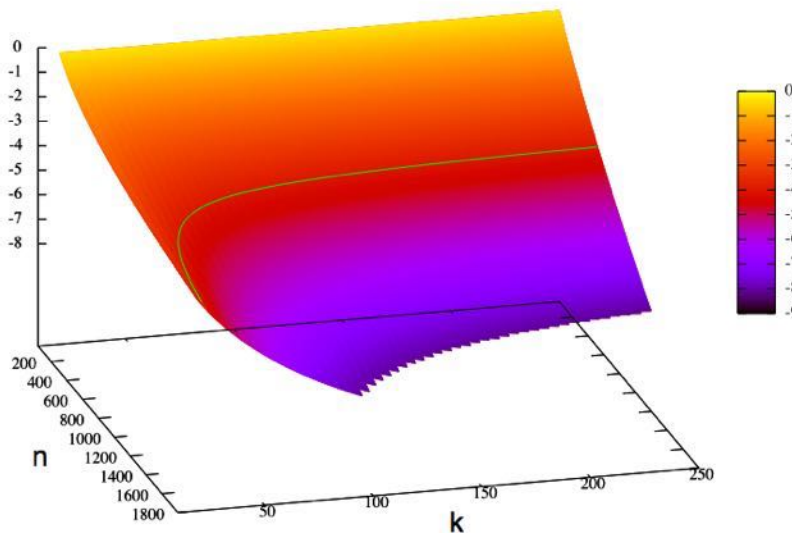


Figure 8. Probabilité d'obtenir un faux-positif (échelle logarithmique) en fonction du nombre d'orientations  $k$  et de leur précision  $n$ , pour un taux de contamination  $\varepsilon=80\%$ . La courbe verte correspond à la résolution de l'équation  $P(\varepsilon,k,n)=P_0$  avec  $P_0=10^{-4}$ .

Par ailleurs, l'émergence d'un pic isolé, caractéristique d'une information commune, peut être perturbé par le fait que les primitives en correspondance n'aient pas leurs orientations exactement séparées d'un angle  $\delta$ . Il en résulte que le signal recherché n'apparaît plus sous la forme d'un pic, mais sous la forme d'un « mode » s'appuyant sur un ensemble de valeurs voisines autour de  $\delta$  (voir figure 9).

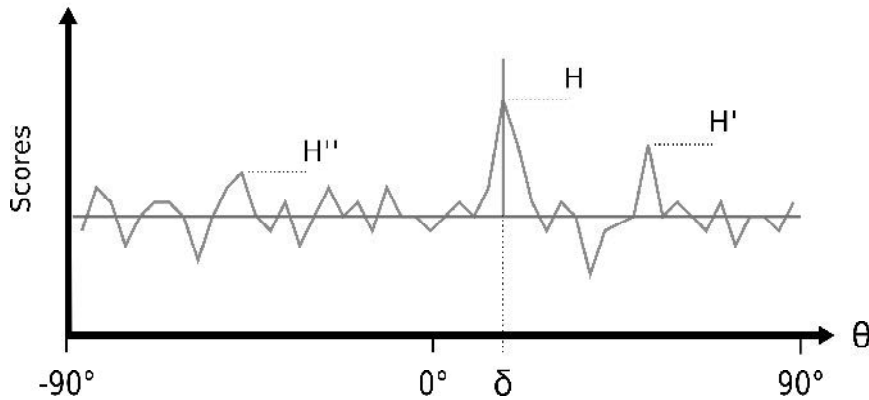


Figure 9. Distribution réaliste des différences d'orientations (distribution « accidentée») : localisation imparfaite du signal recherché H.

La conséquence de cette dispersion autour de la valeur  $\delta$  est que le signal recherché est affaibli et tend à se confondre avec des maxima locaux formés par le bruit. La « non-ambiguïté » du signal recherché est alors de première importance dans la décision de la présence d'une information commune entre deux régions. La prise en compte de l'écart relatif entre le meilleur mode H, et les second et troisième meilleurs modes H' et H'' constituerait une caractérisation plus adéquate. Cela nous a amené à définir le critère suivant :

$$\alpha = \frac{|H - \bar{H}|}{H} \text{ avec } \bar{H} = \frac{H' + H''}{2} \quad [2]$$

où  $\alpha$  est un réel compris entre 0 et 1 : plus sa valeur est proche de 1, et plus la présence d'une information commune est avérée.

Il reste alors à définir une stratégie pour identifier les trois meilleurs modes. Formellement, un mode est défini comme un groupement de cinq cellules dont la cellule centrale est d'un score supérieur aux quatre autres cellules. Après application d'un filtre passe-bas sur l'histogramme pour éliminer les artefacts indésirables, les modes sont repérés et triés en fonction de la somme des scores des cinq cellules qui les composent. Les valeurs H' et H'' correspondent aux scores des cellules centrales des second et troisième meilleurs modes. La valeur H est donnée par la fusion entre la cellule centrale du meilleur mode, et sa voisine de gauche ou de droite selon que l'une ou l'autre somme des scores donne un résultat supérieur.

### 6. Application à l'identification de régions similaires dans des images de modalités différentes

Nous illustrons l'utilisation de notre approche à propos du problème du recalage d'images de modalités différentes. L'analyse d'images, de quelque modalité qu'elles soient, permet d'obtenir des images binaires assimilables à des nuages de points 2D qui caractérisent la localisation d'une information pertinente. Les deux images présentées ci-dessous (figure 10) ont été acquises sur la baie de Tokyo par des capteurs de natures différentes (multispectrale et radar) ; les deux images possèdent la même résolution, soit 30m environ. Une image binaire a été produite pour chaque image, à l'aide d'un opérateur d'extraction de contours spécifique à la modalité d'acquisition (voir Touzi *et al.* -1988- pour l'image radar et Deriche -1987- pour l'image multispectrale, qui a par ailleurs subit une Analyse en Composante Principale afin d'être convertie en niveaux de gris) ; les nuages de points présentés dans la figure 1 correspondent à des parties de ces images binaires. Trois paires de régions ont pu être identifiées grâce à eux : celles-ci sont directement représentées dans les images ci-dessous.

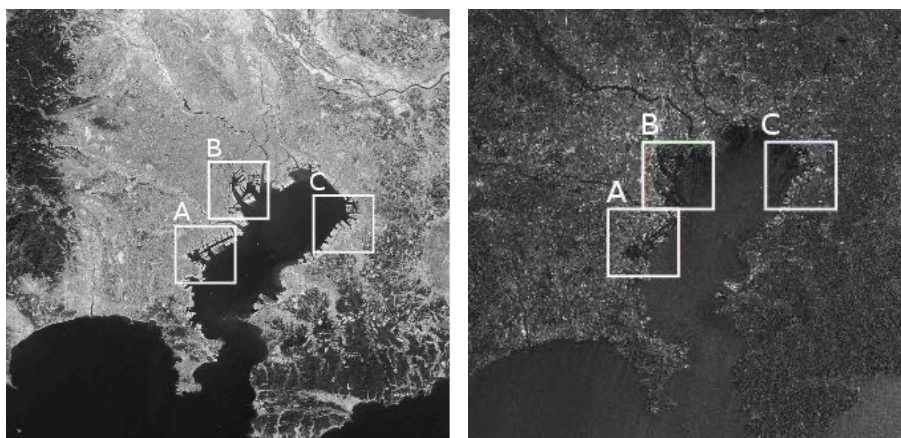


Figure 10. Baie de Tokyo ; image radar AMI-ERS2 (en haut) et image multispectrale Landsat 7 (en bas).

Remarque : dans le cadre de cette application, une mesure de similarité introduite par Bornard *et al.* (1986), et réputée dans l'analyse de textures, a été utilisée sur les précédentes images en niveaux de gris. Malgré l'intérêt de cette approche, il n'a pas été possible d'atteindre des objectifs fixés, les processus physiques menant à la formation de telles images étant très différents. Dans le cas de

la multimodalité, l'utilisation d'un ensemble de localisations liées à des événements pertinents semble plus indiquée.

### 7. Conclusion

Le recalage de nuage de points 2D concerne de nombreux problèmes d'imagerie ; définir une mesure de similarité à la fois robuste et invariante à certaines transformations géométriques est par conséquent d'un intérêt de premier ordre. Les résultats obtenus sur plusieurs jeux de données ont confirmé, au-delà de sa validation théorique et expérimentale, la pertinence et l'efficacité de la méthode que nous proposons.

### Bibliographie

- Besl P., McKay N. (1992). « A Method for Registration of 3-D Shapes », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, n°2, p. 239-256.
- Bornard G., Gauthier JP., Silberman M. (1986). « Mouvements et traitement d'images », *Traitement du Signal*, vol. 3, n°6, p. 281-290.
- Deriche R. (1987). « Using Canny's criteria to derive a recursively implemented optimal edge detector », *International Journal of Computer Vision*, vol. 1, n°2, p. 167-187.
- Grimson W.E.L., Huttenlocher D.P. (1990). « On the sensitivity of the hough transform for object recognition », *IEEE transactions on pattern analysis and machine intelligence*, vol.12, n°3, p. 255-274.
- Fischler M., Bolles R. (1981). « Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography », *Communications of the ACM*, vol. 24, n°6, p. 381-395.
- Hough P.V. (1962). « Method and Means for Recognizing Complex Patterns », *U.S. Patent 3069654*.
- Palmann C., Mavromatis S., Sequeira J. (2009). « A new approach for registering remote sensing images from various modalities », *Proceedings of the SPIE*, vol. 7477, n°1, p. 0C-0C.11.
- Touzi R., Lopes A., Bousquet P. (1988). « A statistical and geometrical edge detector for SAR images », *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, n°5, p. 1054-1060.
- Stewart C.V. (1995). « MINPRAN: A new robust estimator for computer vision », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, n°10, p. 925-938.

### 4.3 Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronaf

Les travaux que je vais présenter dans cette section ont débuté il y a un peu plus de trois ans et s'inscrivent dans la problématique du vol autonome (section 2.1.3). Dans un premier temps, je vais décrire l'approche méthodologique que j'ai mise en place et les solutions apportées. Je vais insister sur les éléments qui m'ont amené à proposer cette méthodologie plutôt que sur aspects techniques qui sont largement décrits dans la thèse de Pierre Zoppitelli. Ensuite, je mettrai en perspective ces travaux à travers les thèses CIFRE d'Antoine Monneau et de Rémi Girard qui ont débutées en janvier 2018.

#### 4.3.1 Problème et objectifs

Je vais reformuler le problème qui a été posé (voir le sujet de thèse de P. Zoppitelli – section 1.5.1, page 21), puis préciser les objectifs. Il est demandé que le système embarqué (disposant de plusieurs caméras) soit capable :

- de détecter une plateforme offshore alors que l'hélicoptère est loin de celle-ci
- de la suivre à mesure qu'on s'en approche et qu'elle devient davantage discernable
- de repérer ensuite le lieu de poser (un helipad)
- d'« accrocher » ce lieu de poser jusqu'à ce que l'hélicoptère se pose (ou reste en stationnaire au dessus)
- et de pouvoir signaler au pilote, lors de toutes ces étapes, qu'il y a eu détection (ou accrochage) ou pas.

Autant le problème est relativement simple à formuler, autant les objectifs sont difficiles à préciser.

L'objectif est double et s'exprime en termes de performance et de robustesse :

- performance : assurer une détection même dans des cas très difficiles
- robustesse : assurer que le résultat est valide

Pour des raisons évidentes (cas extrêmes, par exemple), il est quasi impossible de quantifier ces objectifs de manière appropriée. Cependant, il est intéressant de se placer dans un contexte « performance – robustesse » réaliste et le faire évoluer jusqu'à ce qu'il atteigne un niveau d'intérêt élevé : cette remarque sera l'une des clés de la démarche méthodologique.

#### 4.3.2 Domaine des situations et domaine des solutions

Il est possible de visualiser le problème et les objectifs en termes ensemblistes. Le domaine des situations peut être représenté comme un ensemble  $S$ . Pour une robustesse  $R$  donnée, nous

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

représenterons une classe d'algorithmes (ou un algorithme paramétré)  $A$  pour résoudre un sous-ensemble de situations avec cette robustesse-là comme un ensemble  $S_{A,R}$  inclus dans le précédent. Lorsque l'on fait évoluer la robustesse, cet ensemble rétrécit jusqu'à atteindre au minimum un point pour une robustesse de 100% (i.e. pour une séquence vidéo donnée, on a trouvé un algorithme qui résout le problème). Si l'on représente ces ensembles en 3D où la troisième dimension est la robustesse, alors les classes d'algorithmes (ou l'algorithme) associées à des sous-ensembles de situations correspondent à des nappes de la forme  $R = f(x, y)$ , avec  $R = 0$  au delà d'une certaine zone.

#### 4.3.3 Pavage de l'espace des situations généricité, paramétrage et taille des pavés

On ne peut pas associer de manière bijective un algorithme à une situation – cela ferait concevoir une infinité d'algorithmes. Ceci étant, un algorithme efficace sur une situation donnée devrait l'être sur des situations voisines, mais on remarque que si on utilise un même algorithme sur un sous-ensemble trop vaste de situations, la robustesse diminue de manière importante avec la taille du sous-ensemble.

Il va donc falloir utiliser des algorithmes génériques, qui soient paramétrés et dont le paramétrage soit effectué automatiquement de manière dynamique. Ces pavés ne doivent pas être trop petits (car on doit ne pas trop en avoir) ni trop grands (pour conserver leur efficacité).

La performance s'exprimera en termes de la plus grande couverture possible du domaine total.

Le recouvrement d'une zone de  $S$  par plusieurs pavés est important tant pour la performance que pour la robustesse. Une stratégie de niveau supérieur doit être mise en place, par la suite, pour exploiter efficacement (en les « combinant ») les résultats des différents algorithmes associés à une zone donnée.

Cette vision « globale » n'implique pas nécessairement une démarche de caractérisation préalable du type de situation : elle peut aussi être obtenue (en partie) par l'activation des algorithmes « a priori les plus pertinents » et en évaluant la fiabilité du résultat.

#### 4.3.4 Choix des premiers pavés

Les caméras embarquées produisent un flot de données. L'objectif est d'extraire l'information « pertinente » (i.e. en rapport avec ce que l'on cherche) des données pour établir un lien avec la « connaissance » et c'est cela qui permettra l'accrochage et le suivi.

En première intention, on distingue deux types de connaissance :

- la connaissance « a priori » qui sera délimitée en réfléchissant à ce qui nous, en tant qu'être humains, nous permet de comprendre où est le lieu de poser (je regarde une



### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

image – ou une vidéo et je me pose la question « quels sont les éléments qui, rapidement d’abord puis de manière plus fiable ensuite, me font trouver ce que je cherche? »)

- la connaissance « acquise par apprentissage »

Dans un premier temps, nous n’utiliserons que la connaissance « a priori », mais il est possible qu’ensuite, pour améliorer le processus, nous utilisions également, à un niveau plus fin, la connaissance « par apprentissage » (pour ajuster les paramètres des algorithmes, par exemple).

Les premiers pavés ont été choisis en fonction de leur importance dans le processus global, de la pertinence des situations, et des caractéristiques qui ont semblé les plus marquantes.

Etant donnée la spécificité du problème et le fait que la phase la plus importante est celle où le système se calait sur l’hélicoptère, nous supposons tout d’abord que l’hélicoptère est visible sur l’image (les problèmes, plus simples, de détection au loin des plateformes off-shore seront traités ultérieurement). Nous allons chercher à comprendre comment nous appréhendons cette scène et pourquoi nous en reconnaissons les éléments qui nous intéressent.

Sur une scène de ce type, par exemple :

- il y a des éléments de forme caractéristiques : forme 1, forme 2, ...
- ces éléments sont transformés par homographie
- il y a des « couleurs » dominantes sur des plages de couleur spatialement cohérentes
- il y a des éléments fortement texturés
- ... et vraisemblablement d’autres « éléments de connaissance a priori » que nous étudierons plus tard

Dans un premier temps, nous nous intéressons aux caractéristiques de forme, non pas pour trouver précisément ces formes mais pour les localiser.

#### 4.3.5 Initialisation et « tracking »

Le « tracking » se caractérise par trois types de traitements :

- l’initialisation (au début, où lorsqu’il y a eu décrochage)
- le suivi à proprement parler (d’une image par rapport aux précédentes)
- la validation du suivi (ou le repérage d’un décrochage) pour chaque nouvelle image

On parle souvent de « tracking » pour désigner la phase de suivi. En fait, il existe de nombreuses méthodes de suivi qui sont très efficaces et il sera possible « d’adapter » une de ces méthodes ou bien de développer une nouvelle approche si cela semble plus pertinent. Mais la phase-clé pour automatiser et valider le processus, c’est l’analyse d’images correspondant à l’initialisation : cette phase est déterminante, à la fois pour détecter automatiquement ce que l’on cherche (plateforme, hélicoptère, H, ...), pour initialiser le processus de « tracking », et pour disposer d’un élément complémentaire de contrôle du « tracking ».

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

Dans un premier temps, nous consacrerons donc tous nos efforts à cette phase d'analyse d'image.

#### 4.3.6 Détection du centre d'ellipses associé au centre de cercles dans le modèle

Il existe très souvent « - dans le modèle » un cercle ou, encore mieux, un ensemble de cercles concentriques, visibles sous la forme d'ellipses « quasi-concentrique (mais elle pourront être considérées comme concentriques dans le cadre qui nous intéresse) », et dont le centre est le centre de la plateforme. La détection de ce centre sera un élément déterminant et le premier résultat recherché consiste à estimer ce centre potentiel. Ainsi, par exemple, si ce centre est trouvé, cela permettra de caractériser la zone précise de recherche et la caractérisation d'éléments de validation, comme, par exemple, la présence d'un H sera alors très rapide et très efficace.

Le schéma proposé pour trouver ce centre potentiel est le suivant :

- une suite de procédures, pour certaines classiques, qui produisent des résultats satisfaisants et qui pourront être ensuite remplacées, si besoin est, par des procédures plus sophistiquées - un ensemble de paramètres qui, dans premier temps, ont une valeur donnée manuellement; ces paramètres seront ensuite ajustés automatiquement
- une architecture algorithmique simple mais que nous pourrons chercher par la suite à optimiser pour réduire la complexité.

Le cercle qui entoure le H de l'hélicoptère est une couronne dont la couleur est bien différenciée de celle du fond (même lorsqu'elle est recouverte d'un filet). On peut donc aisément la caractériser par les deux cercles qui la limitent. Et il arrive parfois que la plateforme elle-même soit circulaire (de même centre). De plus, dans le contexte qui nous intéresse, il y a peu (et, en général, pas) de cercles tracés dans la scène (et qui soient de dimension similaire) si ce n'est ceux-ci.

Sur les images acquises, ces contours circulaires apparaîtront alors sous forme d'ellipses.

La première tâche consistera à chercher ces ellipses dans l'image. En fait, il n'est pas nécessaire d'obtenir explicitement ces ellipses mais seulement de leur centre. Pour cela, nous avons fait largement évoluer un algorithme développé dans les premières années du projet SimulFoot [30]. Cet algorithme, que je vais décrire dans les sous-sections suivantes, présente plusieurs avantages :

- il sépare le calcul du centre de l'ellipse (deux paramètres) de celui des trois autres paramètres (qu'il calculerait ultérieurement ... mais que l'on ne recherchera pas ici)
- il permet de caractériser des ellipses mais aussi des arcs d'ellipses (la zone de poser peut être partiellement masquée)
- la présence de plusieurs ellipses concentriques rend plus aisée, dans cette méthode, la détection de leur centre

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

- l'existence de petites ellipses (petites structures circulaires sur la plateforme, par exemple) ne perturbe pas la détection de celles-ci (leur centre sera affecté d'une probabilité de caractérisation très faible par rapport à celui des « grandes ellipses » car cette probabilité dépend du nombre de points qui ont contribué à le trouver)
- l'existence d'une très grande ellipse (ligne d'horizon, par exemple) ne perturbe en rien le processus car, son centre étant hors de l'image, il ne sera pas comptabilisé.

#### Propriété géométrique utilisée pour détecter le centre des ellipses

La méthode proposée repose sur une propriété géométrique des ellipses : étant donnés deux points d'une ellipse, alors la droite passant par le milieu du segment les joignant et par le point d'intersection de leurs tangentes, passe aussi par le milieu de cette ellipse.

A présent, si nous considérons trois points de l'ellipse, alors, en utilisant cette propriété, il est possible d'associer ces points deux par deux, ce qui produit trois couples de points (avec leurs tangentes) qui fourniront trois droites, et ces trois droites seront concourantes au centre de l'ellipse (figure 4.1). C'est cette propriété que nous utilisons pour détecter le centre des ellipses en lesquelles sont projetés les cercles du modèle.

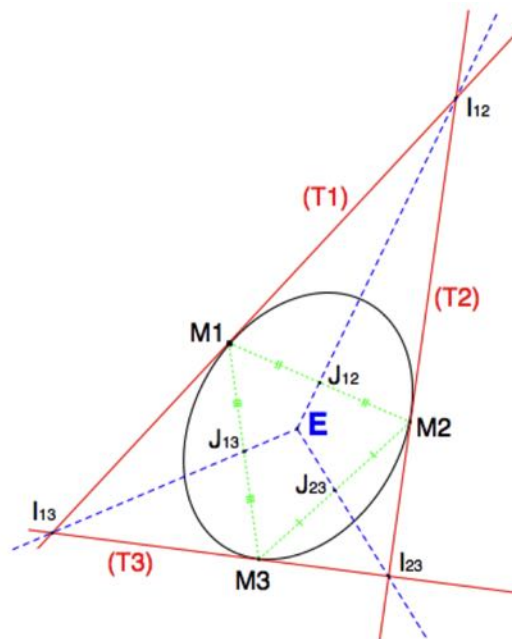


Fig. 4.1 – Le point  $E$  est le centre de l'ellipse.

#### Principe général de l'algorithme

L'algorithme utilise la Transformée de Hough avec comme modèle paramétré - dont nous cherchons les occurrences dans l'ensemble des données - non pas une ellipse (comme expli-

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

qué précédemment) mais un « centre d'ellipse ». On recherche donc des candidats potentiels pour être « centre d'ellipse » en s'appuyant sur des « données spécifiées », on accumule ces candidats potentiels dans un espace de « votes », appelé « Espace de Hough », et le vote majoritaire va permettre de désigner le centre d'ellipse « élu ».

Les « données spécifiées » sont des informations extraites de l'image sous la forme de couples  $(M_i, T_i)$  où  $M_i$  est un point supposé appartenir à une courbe ayant pour tangente  $T_i$  (comme les couples  $(M_i, T_i), \dots$  de la figure 4.1).

Si nous considérons trois couples  $(M_i, T_i)$ ,  $(M_j, T_j)$  et  $(M_k, T_k)$ , ils ne correspondent pas nécessairement à trois points situés sur une même ellipse. Dans ce cas, les trois droites  $D_{ij}$ ,  $D_{jk}$  et  $D_{ki}$  n'ont aucune raison d'être concourantes : un triplet (de couples) de ce type ne produira aucun vote.

Lorsque ces trois droites sont concourantes, cela pourra être parce que les trois points sont sur une même ellipse (qui sera l'une des ellipses qui nous intéressent, ou une petite ellipse se trouvant par ailleurs dans l'image), ou cela sera dû au hasard (avec une probabilité très faible).

Ainsi, en accumulant les centres de ces ellipses potentielles dans l'Espace de Hough, nous aurons du bruit, des petits maxima locaux associés aux petites ellipses (peu de triplets de couples participeront à chacune d'entre elles), et un maximum global associé au centre des ellipses concentriques recherchées, et qui doit être le centre de l'hélicoptère (où nous pourrions chercher plus finement le H, la cohérence de la couleur, ... et valider ainsi le résultat). Cette approche se généralise « naturellement » à la recherche de plusieurs hélicoptères (analyse plus fine, mais dans le même esprit, de l'Espace de Hough).

Nous avons donc à résoudre plusieurs problèmes :

- extraire de l'image tous les couples potentiels  $(M_i, T_i)$
- associer ces éléments en triplets pertinents
- définir l'Espace de Hough et la manière d'accumuler les « votes »
- analyser cet espace pour en extraire le point correspondant à un maximum

A cela, il faut ajouter un problème de complexité algorithmique étroitement lié à la robustesse du processus : en effet, si la caractérisation des couples  $(M_i, T_i)$  ou le choix des triplets de couples  $(M_i, T_i)$  ne sont pas pertinents, énormément de bruit est produit dans l'espace de Hough, ce qui engendre plus de calculs et noie le signal dans trop de bruit.

#### **Extraction des informations de type $(M_i, T_i)$**

Nous cherchons à extraire des points de l'image qui seraient des points de contour et pour lesquels on saurait caractériser la tangente au contour. Je rappelle que nous nous intéressons aux caractéristiques de forme (en particulier, une forme de couronne elliptique) dans le but de la localiser. Nous utiliserons d'autres types d'information (couleur, texture, ...) à des fins de

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

validation ultérieure.

Les cinq étapes listées ci-dessous sont donc relativement classiques mais sont proposées dans l'objectif de conserver au mieux l'information à chaque étape et d'obtenir ainsi de façon progressive les primitives  $(M_i, T_i)$  recherchées.

- 1 IMAGE EN NIVEAU DE GRIS.** Tout d'abord, l'image couleur est transformée en une image en niveau de gris. Les coefficients seront classiquement 0.30 pour la composante rouge, 0.59 pour la verte et 0.11 la bleue ... mais ils pourraient être différents pour mettre en évidence une couleur donnée (par exemple le jaune – ou mieux, le rouge – si on sait que la couronne est jaune).
- 2 IMAGE EN NIVEAU DE GRIS « RÉGULARISÉE ».** Nous appliquons un filtre de Nagao à l'image en niveaux de gris. Cela permet d'éliminer la plupart des variations situées dans les plages uniformes ou en limite de deux plages uniformes : cela correspond au cas de la plateforme (couleur uniforme mais dont certains pixels peuvent avoir une valeur différente – reflet, ...) mais aussi à la surface de l'eau parfois.
- 3 IMAGE BINAIRE.** Nous cherchons des points de contours pertinents, et pour cela nous procédons en deux étapes, dont le résultat final sera une image binaire : tout d'abord, nous appliquons un filtre passe-haut et ensuite nous produisons une image binaire par seuillage de l'image filtrée (La valeur du seuil est obtenue automatiquement de manière à ne conserver que les points correspondant aux  $P$  plus grandes valeurs,  $P$  étant proportionnel à la taille de l'image).
- 4 IMAGE BINAIRE NETTOYÉE - Etape 1.** L'image binaire obtenue fait apparaître les ellipses mais également de nombreux autres points, non pertinents (structure métallique sur la plate-forme, vagues et reflets sur l'eau, etc.). Les points « parasites » sont caractérisés par une densité surfacique très importante, ce qui n'est pas le cas des points associés aux ellipses. Il est facile de supprimer ces zones nettement plus denses. Il est également intéressant de supprimer tous les points isolés (caractérisés par le fait que tous leurs voisins « sont à 0 »).
- 5 IMAGE BINAIRE NETTOYÉE - Etape 2.** Les segments de droites jouent un rôle important dans la compréhension d'une image car ils se réfèrent à des éléments rectilignes de la scène 3D, vraisemblablement créés par l'homme et qui doivent donc correspondre à des composants d'une certaine structure (organisée, ...). En conséquence, il est important de les détecter et de les répertorier. Cependant, si on analyse l'image selon différents modes – i.e. si l'on ne s'intéresse qu'à un certain type d'information – alors l'information ne correspondant pas au mode en question aura un effet perturbateur, et c'est le cas ici. On recherche donc ces segments de droites (on les garde « en réserve » pour des traitements ultérieurs) et pour ne pas perturber ce processus (de recherche de centre d'ellipses pertinentes), on élimine les points proches de ces segments. Une « Transformée de Hough » permet de détecter efficacement ces segments.
- 6 TANGENTES.** Nous allons chercher, pour chaque point de l'image binaire ainsi nettoyée, s'il existe suffisamment de points « alignés dans une certaine direction » (autour de ce

point) pour considérer qu'ils seraient sur une même courbe que celui-ci : la direction associée à ce nuage de points correspondra alors à celle de la tangente. A cette étape, plusieurs solutions existent.

L'une d'entre elles consiste à exploiter l'image binaire nettoyée en recherchant des nuages de points à peu près alignés et suffisamment discriminants pour en déduire qu'il existe une tangente et rechercher cette tangente par l'application « d'un ensemble de masques directionnels ». Cette approche a été initiée dans la thèse de Christophe Palmann (section 1.5.1, page 18). Cependant, dans ce cadre, la précision sur l'orientation des tangentes obtenues n'est pas suffisante et induit un traitement supplémentaire a posteriori. Néanmoins cette approche a été conservée parmi les solutions possibles car elle fournit de bons résultats expérimentaux.

Une autre solution, plus classique est de considérer l'image en niveaux de gris comme « support » des points de l'image binaire et d'utiliser un opérateur classique d'estimation du gradient pour estimer les tangentes. Dans ce cas, un traitement supplémentaire concernant l'orientation des tangentes est à prévoir. Une solution exploitant des composantes connexes est proposée dans la thèse de Pierre Zoppitelli et fournit de bons résultats (section 1.5.1, page 21).

Nous disposons ainsi d'un ensemble de points, et à chacun de ces points est associé une tangente. Ces points ne sont pas sur des droites, ni dans des zones denses uniformément remplies. En d'autres termes, ces points sont sur des structures linéiques (courbes); de plus ils ne sont pas immédiatement voisins et ils sont en nombre peu élevé.

#### 4.3.7 Accumulation de tous les centres potentiels

L'enjeu, à présent, est de trouver un (des) centre(s) d'ellipses à partir de cet ensemble de couples (point, tangente).

Rappelons la propriété des ellipses illustrée par la figure 4.1.

Etant donnés deux points  $M_i$  et  $M_j$ , et leurs tangentes respectives  $T_i$  et  $T_j$ . Soient  $I_{ij}$  le point d'intersection des tangentes  $T_i$  et  $T_j$  et  $J_{ij}$  le milieu du segment  $[M_iM_j]$ . Alors la droite  $D_{ij}$  passant par  $I_{ij}$  et  $J_{ij}$  passe aussi par le centre de l'ellipse.

Les trois couples  $(M_i, T_i)$ ,  $(M_j, T_j)$  et  $(M_k, T_k)$ , permettront, en les prenant deux par deux, d'obtenir trois droites  $D_{ij}$ ,  $D_{jk}$  et  $D_{ki}$ . Si les trois points appartiennent à l'ellipse et que leurs tangentes sont effectivement les tangentes à l'ellipse en ces points, alors les trois droites  $D_{ij}$ ,  $D_{jk}$  et  $D_{ki}$  sont concourantes.

Nous prenons alors tous les triplets possibles  $(M_i, T_i)$ ,  $(M_j, T_j)$  et  $(M_k, T_k)$ , en imposant certaines règles (contrôle de la distance entre les primitives, orientation des primitives, ...) et nous calculons les trois points d'intersection des droites  $D_{ij}$ ,  $D_{jk}$  et  $D_{ki}$  prises deux à deux.

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

Si le triangle ainsi obtenu est petit (sa surface est calculée très simplement avec un produit vectoriel), cela signifie que les trois droites sont concourantes (aux erreurs d'évaluation près, d'où le triangle) et nous utilisons ce résultat pour incrémenter l'espace de Hough (l'espace de Hough est simplement l'image discrétisée à un niveau plus grossier, pour avoir la position approximative du centre); nous détectons la position du point ayant obtenu le plus d'accumulations; et enfin, si besoin est, nous pouvons ensuite, reprendre les tangentes ayant produit le centre (trouvé de manière grossière) pour obtenir le centre de manière plus précise (la complexité est fortement réduite puisqu'il n'y a plus, pour les calculs, que les triplets de tangentes associés à ce centre).

Le maximum obtenu donnera la position du centre des ellipses concentriques les plus caractéristiques, et donc de l'héliport.

#### 4.3.8 Ajustement des paramètres en fonction des situations

Il est légitime de se poser la question de l'ajustement automatique des paramètres. En effet, l'algorithme de détection d'ellipse proposé s'appuie sur un ensemble de paramètres, essentiellement liés à la dimension spatiale de l'image (ou plutôt de la fenêtre – i.e. partie de l'image) traitée. Dans ce cas, une solution élégante et efficace consiste non pas chercher à ajuster au mieux les paramètres en fonction des configurations mais plutôt à « ramener » l'image dans l'espace des paramètres.

Dans un premier temps, nous avons cherché quel était l'ordre de grandeur de la taille de fenêtre permettant la détection la plus efficace (les images trop petites ne permettent pas une bonne détection, mais cela est aussi cas pour des images trop grandes). Dans un deuxième temps, nous avons cherché quel était le jeu de paramètres optimal pour cette taille de fenêtre. Et au final, nous ne traitons que des images de cette taille, avec ce jeu unique de paramètres, après avoir fait une remise à l'échelle de l'image initiale à la taille de cette fenêtre. D'ailleurs, cette opération peut être effectuée dans les deux sens aussi bien pour réduire le nombre de pixels de l'image (figure 4.5 – lorsque l'héliport est très proche) que pour augmenter ce nombre (figures 4.6, 4.7, 4.8, en prenant une sous-image de l'image initiale – lorsque l'héliport est très loin). Cette opération permet une détection très efficace lorsque l'ellipse a une dimension comprise entre la moitié et le sixième de la taille de la fenêtre. En conséquence, il suffit qu'une fenêtre couvre l'ellipse dans ces conditions pour qu'elle soit détectée.

Afin de traiter une image dans laquelle l'héliport peut être de taille très variable, il est possible de gérer l'image globale à travers un ensemble hiérarchisé de fenêtres de tailles différentes. Si nous ne disposons d'aucune connaissance sur l'image (ce n'est pas le cas en général, mais là, nous nous plaçons en « pire cas »), nous pouvons lancer une exploration systématique à tous les niveaux de vision, en partant du niveau global (de l'image) et en descendant ensuite à des sous-niveaux, en divisant, par exemple, l'image en 2 dans chaque direction (4 sous-images), puis en 3 (9 sous-images), ...

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

Cette exploration systématique peut être un peu lourde mais elle reste dans des temps de calcul acceptables, surtout lorsque certains traitements sont déportés sur des architectures dédiées ou lorsque l'on dispose de plusieurs processeurs en parallèle. Elle présente l'avantage de faire une détection à tous niveaux d'observation, sur la totalité de la fenêtre et sans aucune connaissance sur le contenu de l'image.

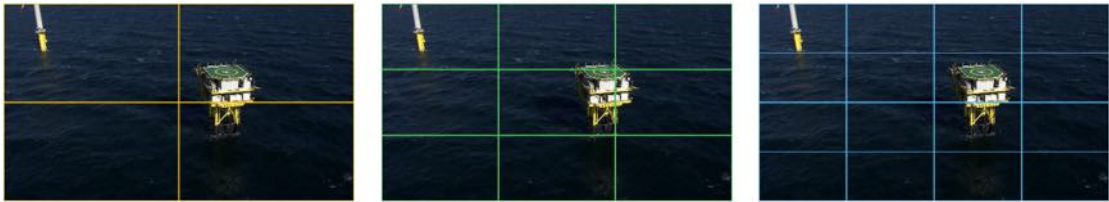


Fig. 4.2 – Les niveaux de vision 2, 3 et 4

Il arrive souvent que l'on dispose de connaissance sur l'image ; alors cette connaissance peut être utilisée pour accélérer le processus et se focaliser directement sur les zones d'intérêt. Ainsi, par exemple, si nous savons que la plateforme à rechercher se trouve sur un bateau (figure 4.3), on peut rapidement repérer la mer (pour l'éliminer) et lancer le processus sur un jeu de « petites » fenêtres choisies dans un ordre « pertinent » dans la zone où est censé être le bateau.



Fig. 4.3 – Image originale

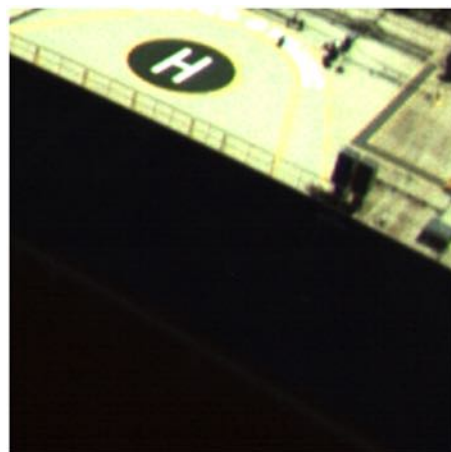


Fig. 4.4 – Fenêtre d'analyse pertinente

#### 4.3.9 Résultats visuels

##### Données de test

Les images en vol ont été tardivement disponibles. Dans un premier temps, c'est le type d'image présenté ci-dessous qui a été utilisé pour tester la méthodologie proposée. La locali-



### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

sation d'une potentielle zone de poser est correcte sur toutes les images de test qui nous ont été fournies. Le point rouge sur chaque image indique cet emplacement.



Fig. 4.5 – Présence d'un filet



Fig. 4.6 – L'hélicoptère est petit



Fig. 4.7 – Nombreuses structures



Fig. 4.8 – Nombreuses structures

#### Données en vol

Les images en vols sont sensiblement différentes des données de test. La colorimétrie est différente, les images sont souvent floues car le système optique subit le tremblement de l'appareil. Néanmoins, dès 1500 mètres, la localisation d'une potentielle zone de poser est correctement détectée. Les figures 4.9 et 4.10 illustrent le résultat obtenu sur ces images.



Fig. 4.9 – Approche classique



Fig. 4.10 – Approche haute

Les tests sur les données en vol ont été menés à partir d'images extraites de six séquences vidéos d'environ une minute chacune ce qui représente environ 8000 images. Il est très rare que la localisation de la zone de poser ne soit pas correctement détectée (la validation de la

détection est réalisée visuellement). Lorsque c'est le cas (moins de 1% des cas), il s'agit d'un nettoyage trop sévère de l'image binaire qui sert à l'obtention des primitives ( $M_i, T_i$ ) (section 4.3.6, page 81).

Il est important de noter qu'à cette étape aucune information temporelle n'est utilisée, il s'agit de détections systématiques. Les tests réalisés permettent de dire que le système est capable de s'initialiser de façon très efficace dans les situations qui nous ont été fournies (à différents moments du vol et donc à différentes distances de la cible ainsi que sous différents angles d'approche).

#### 4.3.10 Détection fine de la couronne elliptique

Dans les sections précédentes, j'ai montré qu'il était possible de désigner efficacement une région de l'image susceptible de contenir une zone de poser. Ce résultat est précieux car il permet de se focaliser sur une région d'intérêt de taille « contenue » et d'envisager de procéder à traitements « locaux » et « intelligents ». A partir de ce résultat, pour valider qu'une zone de poser est bien présente, nous allons valider la présence d'une couronne elliptique, même si elle est partiellement occultée. Dans la suite du document, je vais présenter l'approche choisie sans décrire tous les algorithmes dans le détail mais en insistant sur les points les plus critiques.

##### **Bande elliptique d'intérêt**

Plutôt que de chercher directement la couronne dans la région désignée, nous allons tout d'abord affiner la zone de recherche en définissant une bande elliptique d'intérêt (une ellipse avec une épaisseur) qui contiendra cette couronne. Il sera alors beaucoup plus simple de détecter les deux ellipses recherchées.

La désignation d'un centre potentiel d'ellipse nous permet de dire qu'il y a une accumulation de « centres potentiels » pour une ellipse dans la région désignée et que certaines primitives sont organisées suivant un arc d'ellipse . . . ou pas! Nous allons donc chercher à faire apparaître ces arcs d'ellipse, et s'ils apparaissent, alors le centre détecté sera celui de notre région d'intérêt.

##### Recherche d'ensemble de points appartenant à un « ensemble cohérent » (ellipse)

Pour atteindre cet objectif, nous avons mis en place une méthode basée sur la recherche de composante fortement connexes au sein d'un graphe. Le graphe utilisé ici est un graphe non orienté et pondéré. Les sommets sont les primitives (point – tangente). Les arêtes représentent la corrélation entre deux sommets pour la caractérisation du centre de l'ellipse. Initialement, le graphe est formé de  $k$  composantes connexes (où  $k$  est le nombre de primitives), toutes les arêtes potentielles sont mises à 0 (si une arête est à 0, elle ne relie pas les sommets correspondants). Et à chaque fois qu'un triplet de points contribue au centre, nous incrémentons de 1 chacune des trois arêtes joignant les points deux à deux. Après avoir parcouru tous les

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

triplets, les arêtes auront des poids dont la signification est la suivante : ce poids correspondra au nombre de « troisième primitive » ayant contribué (avec ces deux primitives) à l'obtention du centre. Cela représentera la « qualité » du lien entre deux primitives relativement à leur « participation à une même ellipse dont on connaît le centre ».

Nous étudions la connexité de ce graphe initialement (i.e. après sa création), puis nous l'étudions à nouveau, successivement, après obtention d'un graphe partiel par décimation (on supprime l'arête de poids de plus faible et ainsi de suite). Lorsque, après décimation, une composante connexe apparaît, on teste son nombre de sommets et on l'élimine si celui-ci est petit. A un niveau de seuil donné, nous obtenons seulement quelques composantes connexes ayant chacune un nombre de points suffisant. Cette étape nous permet de nous focaliser sur les arcs d'ellipses qui seraient « de bons candidats » et, en particulier, d'éliminer tous les artefacts.

Il est intéressant de remarquer que les composantes connexes de ce graphe ne sont pas, en général, des composantes connexes sur l'image (par exemple, deux arcs opposés et totalement distincts peuvent avoir contribué de manière conjointe à la création d'une ellipse – i.e. en appartenant à une ellipse donnée – comme le montre la figure 4.11).

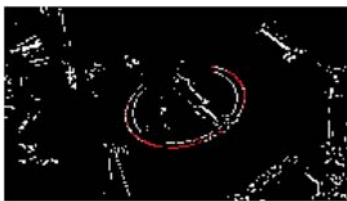


Fig. 4.11 – Une composante connexe du graphe

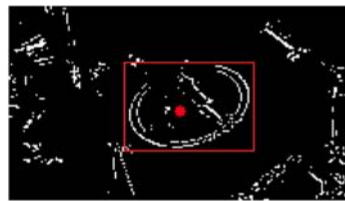


Fig. 4.12 – La boîte englobante des primitives retenues



Fig. 4.13 – Le centre ajusté

Après avoir caractérisé les arcs d'ellipse en utilisant la corrélation entre deux points (corrélation liée à leur contribution à la construction d'une ellipse), nous allons nous appuyer sur la cohérence spatiale (et d'orientation des primitives) pour effectuer cette tâche.

#### Caractérisation des arcs d'ellipse par composantes connexes

Nous disposons d'un ensemble de « points - tangentes » (les primitives présentes dans l'image) et nous souhaitons faire émerger un ensemble de composantes connexes (au sens spatial du terme, cette fois-ci). Les primitives ne sont pas jointives - au sens de la 4-connexité ou de la 8-connexité - mais elles seront considérées comme « voisines » et contribuant à la connexité si elles vérifient un ensemble de contraintes qui sont détaillées dans les lignes suivantes.

Pour cela, nous partons d'un point, qui sera le point de départ d'une nouvelle composante connexe, et nous cherchons à accroître cette composante connexe en lui adjoignant, pour autant que cela est possible, un nouveau point à l'une de ses extrémités. L'adjonction de ce nouveau point se fait par optimisation d'un critère sur l'ensemble des points situés dans le voisinage de cette extrémité : actuellement, nous choisissons le point le plus proche parmi

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

ceux qui vérifient certaines contraintes.

Tout d'abord, nous définissons un « secteur de recherche » qui est un secteur de disque tel que :

- le rayon a pour la valeur  $D$  (par exemple, dans notre fenêtre d'analyse,  $D = 7$  pixels)
- l'origine du secteur est l'extrémité considérée
- un de ses côtés est supporté par la tangente de la primitive associée à l'extrémité considérée
- son autre côté est situé dans la concavité de l'extrémité
- et l'angle  $\alpha$  de ce secteur est assez petit (inférieur à  $20^\circ$  par exemple)

Ainsi, un point peut-être raccroché à l'extrémité considérée si :

- il appartient au secteur de recherche
- l'angle formé entre la tangente de la primitive associée au point et la tangente de la primitive associée à l'extrémité est petit (inférieur à  $20^\circ$  par exemple)

Lorsqu'on ne peut plus agglomérer de points « libres » aux extrémités de la composante connexe, on repart d'un nouveau point libre pour démarrer la recherche d'une nouvelle composante connexe. Et nous poursuivons ainsi jusqu'à ne plus avoir de points libres. A la fin de cette étape, nous disposons d'un ensemble de composantes connexes et nous ne conservons que celles qui ont le plus de points.

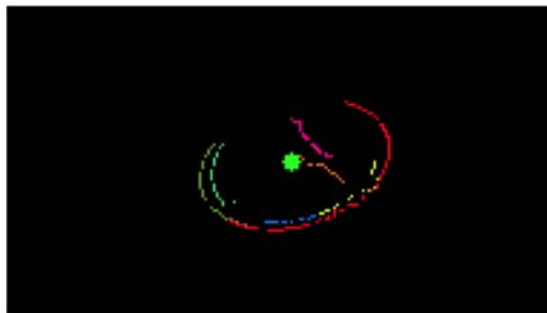


Fig. 4.14 – Chaque couleur désigne une composante connexe.

Nous disposons à présent d'un ensemble de composantes connexes : certaines composantes connexes seront associées à un arc d'ellipse centré au « centre potentiel » et d'autres pas. Nous devons donc déterminer si c'est le cas et calculer alors les coefficients de cette ellipse (si ce n'est pas le cas, nous éliminons la composante connexe correspondante).

#### Calcul des ellipses

Nous décrivons ici la méthode de calcul, méthode qui suppose que nous connaissons le centre de l'ellipse et que les points de cette composante sont sur une même ellipse.

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

Une ellipse centrée s'écrit sous la forme  $ax^2 + bxy + cy^2 = 1$  (on pourrait mettre un autre jeu de coefficients proportionnels du type  $\lambda ax^2 + \lambda bxy + \lambda cy^2 = \lambda$ , mais on aurait au final trois paramètres indépendants). A chaque triplet de points  $(M_1, M_2, M_3)$  d'une composante connexe on associe un jeu de coefficients  $(a, b, c)$ , par simple résolution d'un système linéaire de trois équations à trois inconnues.

Nous inscrivons ces coefficients  $(a, b, c)$  – pour tous les triplets (valides) possibles – dans trois tableaux  $A, B$  et  $C$  et nous calculons les moyennes et écarts-types de ces valeurs,  $E(a)$  et  $\sigma(a)$ ,  $E(b)$  et  $\sigma(b)$ ,  $E(c)$  et  $\sigma(c)$ .

Ainsi, il est aisé de supprimer les points aberrants (par exemple, ceux dont l'écart à la moyenne est supérieur à l'écart-type pour l'une des valeurs  $a, b$  ou  $c$ ) et de recommencer le processus sur les points restants; nous pouvons itérer le processus jusqu'à ce que l'écart type soit inférieur à une valeur donnée ou que le nombre de points restants soit inférieur à un nombre donné. Si l'on ne converge pas au niveau de l'écart-type et du nombre de points, la composante est éliminée; sinon, on conserve le jeu de paramètres  $(E(a), E(b), E(c))$  et on l'associe à la composante connexe.

#### Repérage des arcs d'ellipses connectant les primitives (point – tangente)

Nous arrivons à présent à l'étape d'étiquetage des composantes connexes associées chacune à un arc d'ellipse (nous avons vu dans la section précédente que celles qui ne produisaient aucun arc ou dont le nombre de points résiduels était trop petit, étaient éliminées).

Il est tout à fait possible que pour deux arcs d'une même ellipse, on ait des coefficients légèrement différents; et pourtant les composantes connexes correspondantes doivent être associées à une même ellipse.

Pour résoudre ce problème (et donc celui de la détermination du nombre d'ellipses et de l'étiquetage des arcs), nous utilisons une démarche de « fusion - actualisation ».

Nous partons de l'hypothèse que la plus grande composante connexe (en nombre de points) doit produire les coefficients les plus proches de la réalité. En fait, ce n'est pas le nombre de points mais le nombre de « triplets valides » (au sens vu précédemment) qui va être pris en compte pour choisir la « première » composante connexe (ainsi un long segment représentant une zone de très faible courbure ne sera pas prioritaire par rapport à un petit arc « bien marqué »).

Nous procédons pour cela de la manière suivante :

- Nous prenons la plus grande composante connexe (au sens établi précédemment) et à laquelle soit associée une ellipse
- Pour chacune des autres composantes connexes, nous testons si les points les composants appartiennent à cette ellipse (avec un certain coefficient de confiance lié au caractère discret des données) ... et si c'est le cas pour un fort pourcentage de ces points (par exemple 95%, pour autoriser l'existence de petites dérives ou de « léger outliers »),

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

alors on considère que cette composante connexe est liée à l'ellipse.

- On agglomère toutes les composantes connexes associées à cette ellipse (ce qui produit un nouvel ensemble de points, plus important mais « non connexe »)
- On ré-évalue les paramètres de l'ellipse associée sur cet ensemble de points.

... et on recommence sur la « plus grande » des composantes connexes « non traitée ».

A la fin de cette étape, nous disposons d'un certain nombre d'ellipses, et d'arcs qui leurs sont associés. Une ellipse moyenne est calculée par « proximité » des paramètres en privilégiant l'orientation du grand axe. les deux autres paramètres ( $a$  et  $b$ ) ont nécessairement une variabilité plus grande du fait qu'il peut s'agir de l'ellipse interne ou de l'ellipse externe.

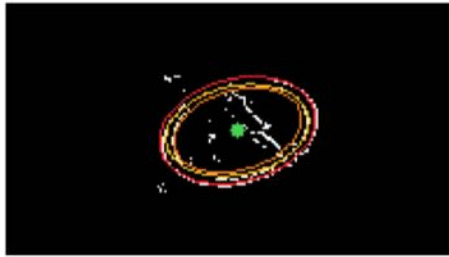


Fig. 4.15 – Ellipses retenues



Fig. 4.16 – Ellipse moyenne

Naturellement, si, tout au long du processus (i.e. lors de cette étape ou lors d'une étape précédente), le nombre de composantes connexes ou d'ellipses venait à être nul, cela signifierait que la région ne contient rien de pertinent (pour ce qui nous intéresse) et nous l'éliminerions (suppression progressive des faux positifs).

A partir de l'ellipse moyenne obtenue, nous définissons une bande elliptique d'intérêt dans laquelle nous allons pouvoir caractériser la couronne de façon très précise.

#### Détection de la couronne elliptique

Pour détecter finement la couronne elliptique, nous allons étudier des profils d'image issus d'un parcours de la couronne « quasi elliptique » (obtenue par « offset » de l'ellipse moyenne). J'illustrerai cette recherche par un parcours radial. Dans le cas d'une ellipse très plate, un parcours orthogonal sera mieux adapté et la méthode reste la même.

#### Balayage radial dans le voisinage de la double ellipse

Nous savons à présent dans quelle partie de l'image, suivant un rayon donné (i.e. demi-droite issue du centre de l'ellipse moyenne), nous devons chercher si les deux points caractéristiques de la couronne - i.e. le point d'entrée (ellipse interne) et le point de sortie (ellipse externe) –

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

existent et, si c'est le cas, quelles sont leurs positions respectives. Pour chaque orientation de la demi-droite (nous avons fait varier l'angle degré par degré, de 0 à 359° - ce choix est arbitraire et peut très bien être modifié), nous considérons un segment (de cette demi-droite) centré sur le point d'intersection de cette demi-droite avec l'ellipse moyenne et recouvrant potentiellement les deux points caractéristiques : nous analysons le profil correspondant (i.e. courbe de variation des niveaux de gris le long du segment) et nous en déduisons ce couple de points s'il existe (nous validons systématiquement un couple de points et non pas un seul point qui appartiendrait à l'ellipse interne ou à l'ellipse externe - i.e. un segment de couronne est présent ou pas).

La longueur du segment peut être trouvée empiriquement ou peut être évaluée précisément. Nous avons pris (empiriquement) une valeur de 12 pixels; cela dit, il est possible non seulement de trouver la valeur optimale de cette longueur mais même de l'adapter à la position sur l'ellipse (l'épaisseur de l'ellipse varie en fonction de l'angle) : en effet, connaissant les rayons des cercles externes et internes, et disposant du birapport de quatre points situés sur un diamètre de ces cercles, nous pouvons, par symétrie, déduire le rapport des rayons projetés. En fait, ce nombre de pixels (choisi empiriquement) n'est pas fixe et dépend linéairement de la distance du centre au point de l'ellipse moyenne. Ainsi, soit  $C$  le centre,  $M$  un point de l'ellipse moyenne,  $f$  un réel (0.5 par exemple),  $M_{int} = C + (1 - f) * CM$  et  $M_{ext} = C + (1 + f) * CM$  : on étudie alors les variations sur le segment  $[M_{int}, M_{ext}]$ .

Comme il peut y avoir du « bruit », nous n'étudions pas directement le profil (i.e. évolution sur le segment discret) mais la moyenne de cinq valeurs voisines (valeur du pixel et de deux pixels de part et d'autre) prises orthogonalement en chacun de ces points.

Les figures 4.17 et 4.18 illustrent deux moments du balayage. Les figures 4.19 et 4.20 montrent le profil de chaque segment. En abscisses, on retrouve donc chaque pixel du segment considéré, et en ordonnées, la valeur moyenne des niveaux de gris.

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

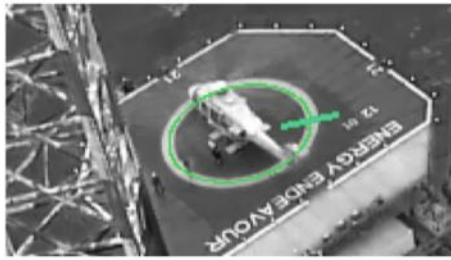


Fig. 4.17 – Le segment d'analyse traverse la couronne



Fig. 4.18 – Le segment d'analyse ne traverse pas la couronne

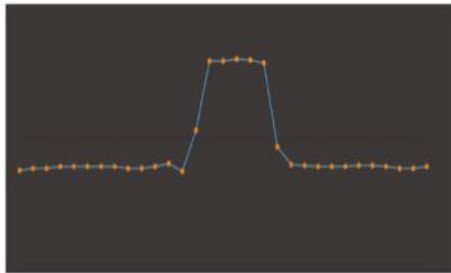


Fig. 4.19 – Profil des niveaux de gris lorsque la couronne est traversée

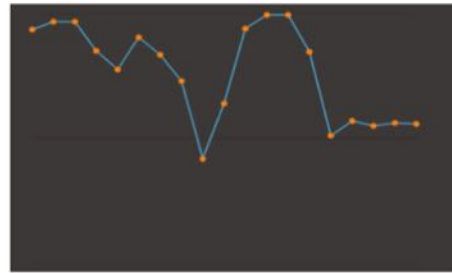


Fig. 4.20 – Profil des niveaux de gris lorsque la couronne n'est pas traversée

#### Caractérisation sur un segment des points d'entrée et de sortie de la couronne

La couronne est plus claire que les points situés à l'intérieur, et aussi, en général, que les points situés à l'extérieur : en effet, sa couleur jaune (sur l'image couleur) correspond au maximum d'énergie dans les deux bandes les plus « énergétiques » (le vert et le rouge). Cela apparaît dans la courbe (moyennée orthogonalement), avec des valeurs basses à l'intérieur, puis des valeurs hautes dans la couronne, puis des valeurs plus basses à nouveau, à l'extérieur. C'est ce que nous allons chercher à caractériser ici (pour savoir si nous sommes dans une telle situation et quels sont les points de transition). Parfois, certains éléments perturbateurs peuvent modifier la forme locale alors que la forme globale met en évidence une telle situation. Une analyse locale devient difficile et peut ne pas être pertinente, comme le montre les figures 4.21 à 4.24.



Fig. 4.21 – Le segment d'analyse traverse la couronne



Fig. 4.22 – Le segment d'analyse traverse partiellement la couronne



### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

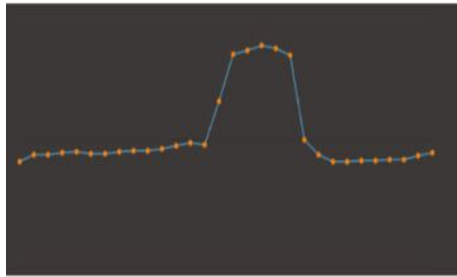


Fig. 4.23 – Profil des niveaux de gris lorsque la couronne est traversée

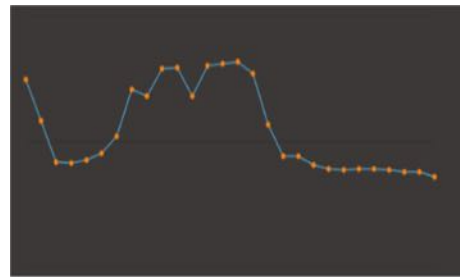


Fig. 4.24 – Profil des niveaux de gris lorsque la couronne est partiellement traversée

Pour éviter cet écueil, nous allons régulariser la courbe en l'intégrant (i.e. en la sommant). Les figures 4.25 et 4.26 montrent les courbes intégrales des deux courbes représentées par les figures 4.23 et 4.24 (ces courbes sont toujours croissantes puisque les courbes originelles sont à valeurs positives).

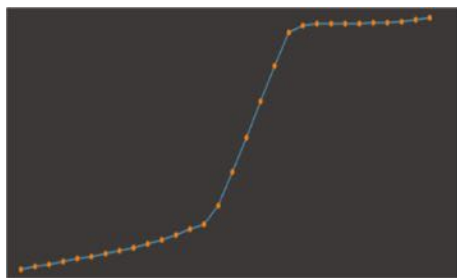


Fig. 4.25 – Intégration du profil (4.23)

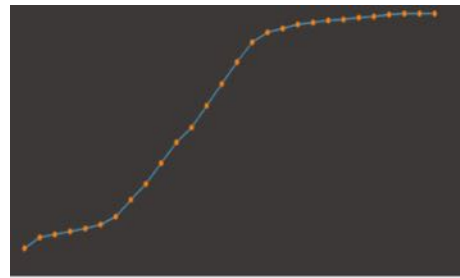


Fig. 4.26 – Intégration du profil (4.24)

Les courbes obtenues ont la même forme (malgré « l'accident » présent dans la deuxième courbe originelle). Cette forme est la suivante : une croissance faible, suivie d'une croissance forte, suivie d'une croissance faible à nouveau. Et cette forme est caractéristique du passage par une bande plus lumineuse. Ainsi, dans l'exemple initial (figure 4.19 : passage dans la couronne, et 4.20 : passage au niveau de l'hélicoptère), les courbes initiales fournissent les courbes intégrales de formes foncièrement différentes (figure 4.27 et 4.28).

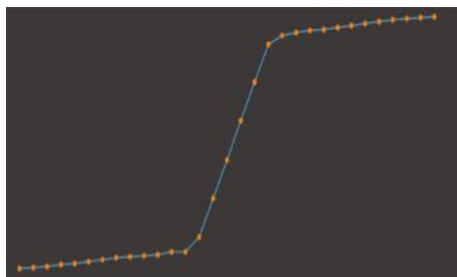


Fig. 4.27 – Intégration du profil (4.19)

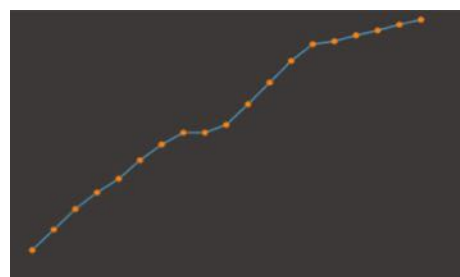


Fig. 4.28 – Intégration du profil (4.20)

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

Ce type de forme « croissance faible – croissance forte – croissance faible » peut-être caractérisé en segmentant la courbe sommée en trois segments par un processus de « division » (split) en évaluant le signe de l'angle entre les deux premiers segments puis entre le deuxième et le troisième : si le premier signe est positif et le deuxième négatif, alors nous sommes dans le cas recherché, et les extrémités du deuxième segment correspondent aux bords de la couronne.

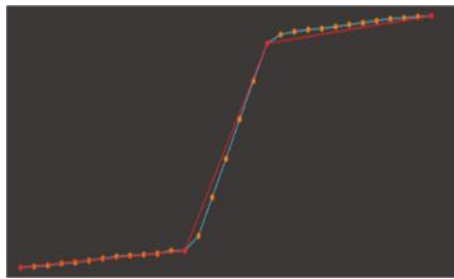


Fig. 4.29 – En rouge, la courbe segmentée

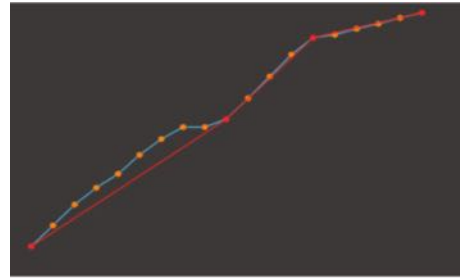


Fig. 4.30 – En rouge, la courbe segmentée

Nous obtenons ainsi deux listes de points, celle des points détectés comme appartenant à l'ellipse interne, et leurs analogues sur l'ellipse externe (ces deux listes ont le même nombre de points). Mais certains points peuvent correspondre à des artefacts (e.g. hublot arrière de l'hélicoptère détecté comme une zone « sombre » semblable à la luminosité de la zone autour du « point de poser » et hélicoptère « clair »). De nombreuses méthodes sont disponibles pour calculer les ellipses passant au plus proche d'un ensemble de points. Une approche simple consiste à lancer l'algorithme sur l'ensemble de points initial, de supprimer les points éloignés à une distance supérieure à un seuil. L'algorithme est relancé jusqu'à ce que le nombre de points reste inchangé entre deux itérations. Les résultats obtenus sont très satisfaisants comme le montre la figure 4.32 sur une image prise en vol.

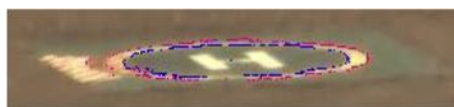


Fig. 4.31 – Les points retenus lors du parcours

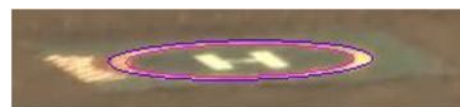


Fig. 4.32 – Les ellipses résultats

L'intérêt de la méthode décrite ici est multiple :

- Elle permet de donner une validation définitive de l'existence de la couronne entourant le « point de poser » (quand on a trouvé cette double couronne, on est « sûr » qu'il s'agit bien du point de poser et on peut éventuellement actionner d'autres processus de validation).
- Une manière d'assurer une validation supplémentaire consiste à évaluer le birapport des quatre points d'intersection d'un diamètre avec ces deux ellipses.

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

- Elle permet de caractériser les zones éventuelles d'occultation et donc de « comprendre » la scène : il y a deux ellipses mais aussi des zones invalidées qui permettront de lancer un processus d'analyse dans ces zones (présence d'un hélicoptère, ...) comme le montre la figure 4.34.
- Elle est très rapide



Fig. 4.33 – Les points retenus lors du parcours



Fig. 4.34 – Les arcs d'ellipses résultats

#### 4.3.11 Développement d'une plate-forme logicielle

Nous avons développé (Jean Sequeira, Pierre Zoppitelli et moi même) une plateforme logicielle qui reflète la méthodologie proposée et qui permet de disposer d'un outil modulaire et évolutif comme base expérimentale. C'est à partir de ce développement que le transfert vers le système Eagle a été réalisé. Cette plate-forme est développée en C++ et utilise la bibliothèque de fonctions OpenCV pour les fonctionnalités de base (entrées/sorties, filtrage, seuillage, ...).

Les objectifs d'un tel outil sont multiples :

- Disposer d'une implémentation effective sur un ordinateur banalisé permettant la détection et le suivi de l'hélicoptère lorsque celui-ci est caractérisable
- Permettre un portage aisé de cette implémentation dans une architecture dédiée
- Caractériser les modules « critiques » (en termes de temps de calcul) pour les déporter sur un support dédié et partager ainsi les ressources (entre ordinateur de contrôle et processeurs spécialisés)
- Permettre l'intégration de nouveaux modules pour la résolution de problèmes spécifiques (abord de la plateforme de nuit, approche rasante, brouillard, environnement complexe, ...)
- Permettre l'intégration de nouveaux modules dans le but d'accroître la robustesse du processus de détection et de suivi (à travers la couleur, la texture, les formes, ...)
- Permettre l'intégration des éléments de détection préalable, alors que la plateforme ne peut pas être encore caractérisée en tant que telle (vision de « très loin »)
- Permettre le passage progressif, à travers un logiciel en permanence opérationnel, d'une version semi-automatique avec faible intervention humaine à une version entièrement automatisée.

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

---

Notre démarche a consisté à étudier le problème le plus caractéristique (ou censé apparaître le plus fréquemment), et à le traiter de manière modulaire et paramétré, puis à le programmer sous la forme de l'enchaînement de ces modules sous l'égide d'une « structure d'états » (présente dans le programme) dont une image existe de manière externe dans un « fichier de configuration ». Ensuite, pour traiter un problème « assez différent », nous pouvons utiliser (ou pas) les modules existants, et éventuellement d'autres modules : ces différents cas sont alors simplement associés à des « fichiers de configuration » différents. Cela induit de disposer d'un ensemble (restreint) de « fichiers de configuration » associés à autant de classes de situations (vision de nuit, plateforme complexe, plateforme sur un bateau, ...).

Grâce à ce dispositif, il est possible d'équilibrer à la demande la partie gérée dans le système (ordinateur « maître » qui joue le rôle de superviseur) et la partie déportée dans une architecture dédiée, d'étudier les performances des divers modules indépendamment, et de maîtriser ainsi parfaitement la performance globale du processus (i.e. en sachant ce qu'il faut déporter, optimiser, ... et s'il faut éventuellement plusieurs processeurs, ...). Ce dispositif permet de se mettre, de manière opérationnelle, dans chacune des classes de situations et de faire fonctionner le processus de détection de plateforme (et pas uniquement d'ellipse) en y associant éventuellement d'autres fonctionnalités (du type couleur, forme, ... recherche d'obstacles, ...).

Pratiquement, le choix (purement interactif, par le pilote) de la « classe de situations » n'est pas du tout contraignant et est même naturel (le pilote sait s'il vole de nuit ou pas ... !). Mais, surtout, sans changer quoi que ce soit à l'architecture du programme, on peut le faire évoluer (toujours sous sa forme opérationnelle) à travers le « manager » (ou « superviseur ») qui peut aisément détecter de manière automatique la « classe de situations », de manière à rendre le programme entièrement automatique (ou quasi-automatique) ... en toutes situations.

Je vais très rapidement présenter cette plate-forme à travers trois illustrations : un extrait de fichier de configuration, un extrait de code et un exemple de fichiers résultats.

#### **Extrait du fichier de configuration**

Le fichier de configuration est écrit au format YAML. Ce format est reconnu pour sa grande lisibilité. Il est également « supporté » par OpenCV ce qui permet de ne pas créer de dépendances supplémentaires avec d'autres bibliothèques de fonctions. Les traitements à appliquer ainsi que les paramètres associés définissent l'état du système. Ils sont mentionnés dans le fichier de configuration. Il est ainsi possible d'activer tout ou partie d'un module simplement à partir de ce fichier (voir figure 4.35).

#### **Exemple d'utilisation**

L'utilisation des différentes fonctionnalités est « standardisée ». Un « fichier de version » décrit une chaîne de traitements et une gestion spécifique des entrées-sorties rend aisé

### 4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef

```
100 ##### TRACKING #####
101 ##
102 ## TRACKING : TRACKING_ON / TRACKING_OFF
103 ##
104 #####
105
106 TRACKING : TRACKING_ON
107
108 ##### SUBDIVISION CONFIGURATION #####
109 ##
110 ## SIZE_SUB_IMAGE : 250
111 ##
112 ## SUBDIVISION_METHOD : NO_SUBDIVISION / SUBDIVISION_1_4_9_16
113 ##
114 ## NO_SUBDIVISION -> one subImage which constitutes the whole image
115 ##
116 ## SUBDIVISION_1_4_9_16 -> 1_4_9_16 division procedure
117 ## if no depth is set : depth = 4
118 ## to set the depth : SUBDIVISION_DEPTH = 2
119 ##
120 #####
121
122 SUBDIVISION_METHOD : SUBDIVISION_1_4_9_16
123 SUBDIVISION_DEPTH : 3
124
125 SIZE_SUB_IMAGE : 350
126
127 ##### VERSIONS CONFIGURATION #####
128 ##
129 ## COMPUTATION_METHOD : VERSION_1 /
130 ##
131 #####
132
133 COMPUTATION_METHOD : VERSION_1
134
135 ##### VERSION_1 CONFIGURATION #####
136 ##
137 ## -----PREPROCESSING-----
```

Fig. 4.35 – Extrait du fichier de configuration

l'interprétation des résultats (voir figure 4.36).

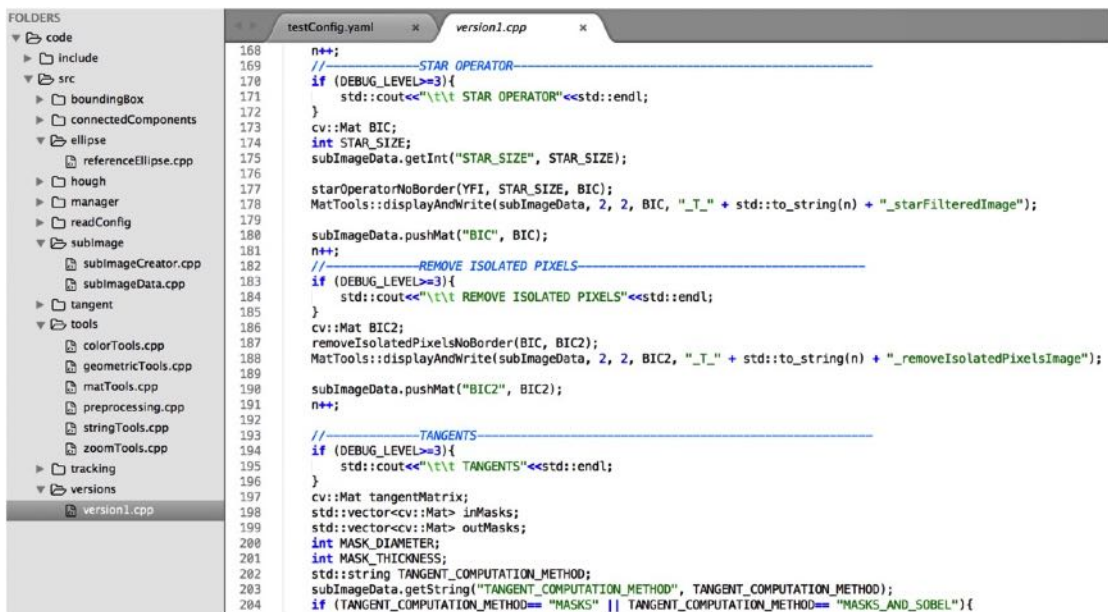


Fig. 4.36 – A gauche, l'organisation du code source. A droite, un extrait de code source

#### Exemple de fichiers en sortie

Chaque module permet de générer des fichiers en sortie afin de maîtriser au mieux le processus. Les images ci-dessous présentent les résultats de différents modules : traitement d'images en niveau de gris, binarisation, calcul des composantes connexes, calcul des primitives, extraction d'un centre potentiel, validation ... (voir figure 4.37)

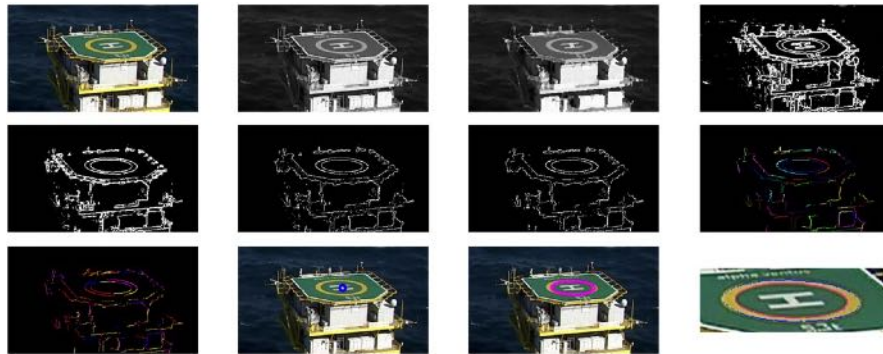


Fig. 4.37 – Images résultats à différentes étapes

#### 4.3.12 Prolongement des travaux

Les travaux que j'ai présenté dans cette section sont les premiers d'une collaboration avec la société Airbus Helicopters qui s'articule sur du long terme autour de la thématique du vol autonome.

La méthodologie proposée répond aux objectifs de performance et robustesse qui ont été initialement posés. Elle a été éprouvée sur tous les jeux de données qui ont été fournis (8000 images issues de six types d'approches et au moins autant images extraites de données récoltés ici et là) et s'est montrée très efficace (moins de 1% de non-détection, aucune fausse détection). Le portage vers la plateforme embarquée Eagle est en cours de développement par les ingénieurs d'Airbus.

Le travail réalisé a donc été très apprécié et a permis le démarrage de deux nouvelles thèses CIFRE qui vont bénéficier des résultats obtenus :

- La première sur le sujet de l'appontage automatique dont l'objectif est de proposer une approche mixe automatique/vision en vue de permettre l'appontage entièrement autonome d'un drone-hélicoptère sur un bateau. Ce travail, débuté en janvier 2018, a déjà fait l'objet d'une publication [14] et une seconde publication vient d'être soumise à la cinquante huitième édition de la conférence « IEEE Conference on Decision and Control ».
- La seconde traite de la mise en oeuvre d'une méthodologie de type SLAM (Simultaneous Localization and Mapping) – dont l'objectif est le créer un modèle de l'environnement, sans disposer de connaissance sur celui-ci, et de localiser le capteur dans cet environne-

### **4.3. Analyse, reconstruction et aide au pilotage à partir de caméras embarquées sur un aéronef**

---

ment – sur des hélicoptères dotés d’une seule caméra mais à laquelle on peut adjoindre un dispositif de projection de mire (lumière structurée). Ce travail a également débuté en janvier 2018 et une publication intitulée « A vision-based assistance key differentiator for helicopters autonomous scalable missions » a été acceptée à la vingtième édition de la conférence « TAROS (Towards Autonomous Robotic Systems) ».

## 5 Projet scientifique

Au cours de ces vingt années de travaux de recherche, j'ai approfondi différents aspects de l'analyse d'images, de la reconnaissance des formes et de la visualisation scientifique avec pour objectif d'essayer d'extraire l'information des données image de manière à la présenter sous une forme pertinente à un utilisateur. Cela m'a rapidement amené à une première conclusion, très importante : si l'information globale extraite des données est un élément essentiel de nombreuses disciplines (compression d'images, filtrage linéaire, analyse spectrale, communications et réseaux . . .), l'approche doit être totalement différente dans le cadre qui m'intéresse, que j'ai présenté dans ce mémoire, et que j'ai résumé dans la première phrase ci-dessus. En effet, dans la plupart des cas (si ce n'est dans tous les cas), il s'agit d'un utilisateur qui appréhende le problème à traiter avec une vision spécifique et qui a des objectifs spécifiques; pour les mêmes données, dans un autre contexte (et un autre « type d'utilisateur »), le problème sera abordé de manière différente au niveau scientifique.

Je reprendrais donc cette première phrase et la changerait légèrement, en y ajoutant deux mots, mais ces deux mots sont lourds de sens puisqu'ils caractérisent une orientation forte de ma démarche scientifique : « . . . avec pour objectif d'essayer d'extraire l'information contextuelle des données image de manière à la présenter sous une forme pertinente à un utilisateur dédié. ». Cela m'a donc amené à une deuxième conclusion, qui était la nécessité de travailler sur la représentation de la connaissance explicite. Ainsi, dans chacun des projets que j'ai mis en oeuvre, avant d'essayer de l'aborder à travers certaines « théories scientifiques », j'ai cherché à formaliser l'appréhension de l'utilisateur, ses objectifs, son monde fonctionnement; et cela a été systématiquement le point de départ de ma démarche.

Ainsi, un pilote d'hélicoptère ne percevra pas la scène de la même manière qu'une autre personne, et c'est ce qui lui permettra de décider plus rapidement et de façon plus pertinente de l'action à mener; ainsi, un entraîneur de football n'aura pas la même appréhension d'une séquence vidéo qu'une autre personne; ainsi, un médecin « spécialiste » se focalisera rapidement sur ce qui pertinent par rapport à sa spécialité, ce que ne pourra faire un médecin « généraliste » qui, lui, aura peut-être une vision globale plus pertinente.



---

Naturellement, nous retrouvons tout cela dans un abord plus scientifique de ces questions. Ainsi, la « Reconnaissance des Formes », dans son volet « classique » et « explicite » consiste à définir un « Espace de Représentation », un « Espace d'Interprétation » et à chercher une fonction (dont on voudrait qu'elle soit une application !) du premier vers le second. La première difficulté consiste à définir ces espaces, et c'est ce à quoi je me suis intéressé, la plupart des travaux publiés étant plutôt associés à la recherche de cette fonction (classification ...). En ce qui concerne l'extraction de l'« information contextuelle », nous pouvons faire état de la « Morphologie Mathématique » où, plutôt que de chercher « à ne pas perdre d'information » (filtrage linéaire, surtout lorsque l'on souhaite maintenir la réversibilité), on cherche à séparer différentes visions de l'information puis à appréhender tout cela globalement. Et c'est également le cas de tout ce qui concerne l'analyse multi-échelle.

Je voudrais enfin, même si c'est une évidence pour tous, préciser l'évolution de l'environnement technologique entre il y a vingt ans et maintenant, car, comme nous allons le voir (en particulier pour ce qui concerne les données disponibles), cela a une incidence sur mon projet scientifique. Il y a vingt ans de ça, il était possible de se procurer des images et des séquences vidéo, mais leur accessibilité n'était pas toujours simple (formats propriétaires dans le médical, images « tuilées » en télédétection ...) et il n'était pas possible d'accéder à de véritables bases de données d'images, comme c'est le cas aujourd'hui. Par ailleurs, les dispositifs techniques de visualisation se limitaient souvent à un écran alors qu'aujourd'hui, il est possible d'accéder à des outils technologiques plus performants (pour produire un environnement immersif, par exemple) et disponibles à un coût raisonnable.

J'en viens donc à mon projet scientifique à proprement parler, et je vais le décliner sous différents volets :

- Approche méthodologique : il s'agit en fait de ma démarche scientifique globale
- Domaines scientifiques : il s'agit des « théories » s'incluant dans ma démarche méthodologique et que je compte approfondir
- Outils technologiques : il s'agit des nouveaux dispositifs permettant d'aller plus loin dans ma démarche globale, avec de nouveaux problèmes d'ordre scientifique
- Projets : il s'agit de la poursuite des projets importants en cours et de l'initiation de nouveaux projets (actuellement en gestation)

## **Approche méthodologique**

Comme j'ai pu l'évoquer précédemment, je me suis intéressé, à travers l'imagerie numérique, aux processus cognitifs humains, ce qui m'a permis, par analogie avec ces processus cognitifs, de trouver des solutions efficaces aux problèmes que j'ai abordés. Cela étant, tous mes travaux de recherche étaient purement associés à l'analyse d'images, à la reconnaissance des formes et à la visualisation, et non à la « Représentation de la Connaissance ». Je souhaite à présent formaliser davantage cet aspect pour donner un cadre méthodologique à mes travaux en imagerie, sachant que ce domaine – lié à la connaissance explicite – a déjà fait l'objet de

---

nombreux travaux de recherche par ailleurs.

On distingue chez l'être humain deux niveaux de connaissance (même si la frontière entre les deux n'est pas aussi marquée que cela), celle qui nous a été enseignée ou que l'on est capable d'exprimer (explicite) et celle qu'on acquiert par l'expérience sans pouvoir la formaliser (implicite). En l'état actuel, et malgré des avancées importantes dans le domaine de l'apprentissage automatique (Machine Learning), un socle de « connaissance explicite » semble incontournable; mais à travers celle-ci, on n'arrive jamais (ou rarement, et seulement dans des cas simples) à prendre en charge toutes les situations possibles. L'apprentissage vient alors en complément pour aborder et résoudre ces problèmes « à la marge » et qui, dans la pratique, sont très importants. L'existence de volumes de données gigantesques, qui sont parfois structurés à travers des bases de données (Big Data), doit permettre d'aller plus loin dans la résolution de ces problèmes par apprentissage, en particulier à travers une approche largement publiée, qui est le « Deep Learning ». Je souhaite donc également avancer dans cette voie.

Enfin, toujours par rapport à cette problématique, je souhaite étudier les rapports entre ces deux niveaux de connaissance et voir comment ils peuvent s'enrichir mutuellement.

Dans cette section, je me suis situé à un niveau « au-dessus » de l'analyse d'images, de la reconnaissance des formes et de la visualisation; je n'ai jamais développé de travaux de recherche dans ce niveau-là; mais je pense que de nombreux verrous dans le domaine de l'image peuvent être cassés si on les aborde de manière scientifique à travers cette démarche, et cela constitue le coeur de mon projet scientifique.

## **Domaines scientifiques**

Si l'on revient au niveau de l'image, alors, au-delà des prolongements naturels de mes travaux actuels, je souhaite m'investir dans certains domaines scientifiques que j'ai déjà abordés de manière plus ou moins approfondie.

C'est le cas en particulier de la « Réalité Virtuelle » et de la « Réalité Augmentée ». On a pu voir, à travers différentes « expériences » (participation au projet d'environnement immersif du CRVM, projet Virtu'Art, enseignement et recherche à Polytech) que la « Réalité Virtuelle » était au coeur de mes préoccupations depuis de nombreuses années. Quant à la « Réalité Augmentée », l'idée était présente dans les différents projets que j'ai développés, que ce soit le projet SimulFoot avec le « marquage » des éléments clés et l'ajout d'information pertinente extraite d'une séquence vidéo, ou dans les autres projets. En fait, je souhaite intégrer les deux – ce qui nécessite de l'analyse d'image et de la reconnaissance des formes mais également un recalage de formes 3D (le tout « en temps réel ») – pour intégrer dans une scène observée « naturellement » des éléments additionnels purement virtuel.

Dans le domaine de l'analyse d'images, je souhaite prolonger mes travaux sur le recalage

---

d'images dans des cas plus complexes et plus généraux (changement d'échelle, déformations plus complexes comme dans le cas d'images de télédétection sur un relief montagneux ...).

Enfin, les gros volumes de données en télédétection (nombre croissant de satellites, drones) doivent permettre de produire des images plus fines, et de plus haute résolution grâce à ce que l'on appelle la « super-résolution », démarche qui consiste à intégrer l'information de plusieurs images d'une résolution donnée pour créer une image de résolution supérieure. Il s'agit d'un domaine de recherche important et qui vient dans le prolongement de mes travaux (pour produire une image par super-résolution, on doit pouvoir recaler ces images entre elles).

## **Outils technologiques**

Même si, dans cette section, on sort quelque peu du domaine purement scientifique, il est important de noter le rôle de nouvelles technologies qui deviennent techniquement acceptables, de coût abordable et qui entreront dans le cadre de notre quotidien dans les prochaines années. Deux de ces technologies m'intéressent tout particulièrement : la « Réalité Virtuelle » et les drones : j'envisage de développer plusieurs de mes futurs projets autour de ces technologies.

## **Projets**

Quelques nouveaux projets sont en gestation mais je souhaite avant tout m'articuler sur l'existant, qui fonctionne très bien (aux niveaux humain, relationnel et logistique) et où il y a matière à de nouveaux travaux de recherche dans le cadre de la thématique sur laquelle je travaille. C'est le cas, en particulier, pour les projets avec Airbus (où nous avons trois thèses en CIFRE – l'une d'elle ayant été soutenue récemment) mais aussi des projets (sur les satellites et les drones) avec des partenaires étrangers (en Chine, au Pérou, en Algérie).

Enfin, je pense que le support de Polytech – enseignement et recherche, projets d'étudiants – sera propice aux travaux sur la Réalité Virtuelle et sur la Réalité Augmentée.

En conclusion, les travaux de recherche que j'ai développés au cours de ces vingt dernières années m'ont donné un recul sur la manière d'aborder les problèmes d'imagerie, à travers la notion de connaissance. J'ai très peu travaillé sur cette notion mais j'ai l'intention de m'y consacrer ; et c'est essentiellement dans cette direction que je compte m'orienter. Dans ce cadre-là, je souhaite aborder certains domaines scientifiques – évoqués dans les lignes précédentes – à travers de nouveaux outils technologiques et en m'appuyant sur la base de projets (avec des industriels et à l'étranger) existants et qui ont un fonctionnement pérenne.

# Publications

## Revue, Livres et Brevets

- [1] Xiaojiang LI, Qingyan MENG, Weidong LI, Chuanrong ZHANG, Tamas JANCZO et Sébastien MAVROMATIS. “An explorative study on the proximity of buildings to green spaces in urban areas using remotely sensed imagery”. In : *Annals of GIS* 20.3 (2014), p. 193-203. DOI : 10.1080/19475683.2014.945482.
- [2] Sébastien MAVROMATIS et Jean SEQUEIRA. “Method and system for estimating a similarity between two binary images”. Brev. WO Patent 2012117210A1. 2014.
- [3] Xiaojiang LI, Qingyan MENG, Xingfa GU, Tamas JANCZO, Tao YU, Ke WANG et Sébastien MAVROMATIS. “A hybrid method combining pixel-based and object-oriented methods and its application in Hungary using Chinese HJ-1 satellite images”. In : *International Journal of Remote Sensing* 34.13 (2013), p. 4655-4668. DOI : 10.1080/01431161.2013.780669.
- [4] Sébastien MAVROMATIS et Jean SEQUEIRA. “3D Reconstruction of Sport Scenes”. In : *3D Video : From Capture to Diffusion*. John Wiley & Sons, Inc., 2013, p. 405-420.
- [5] Sébastien MAVROMATIS et Jean SEQUEIRA. “Reconstruction 3D de scènes sportives”. In : *Vidéo 3D : Capture, traitement et diffusion*. Hermès - Lavoisier, 2013, p. 217-232.
- [6] Sébastien MAVROMATIS et Jean SEQUEIRA. “Procédé et système pour estimer une similarité entre deux images binaires”. Brev. EP Patent PCT/FR2012/050,438. 2012.
- [7] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. “Mesure de similarité entre sous-parties de nuages de points 2D”. In : *TS. Traitement du signal* 29.1-2 (2012), p. 29-49. DOI : 10.3166/ts.29.29-49.
- [8] Christophe PALMANN, Sébastien MAVROMATIS, Mario HERNÁNDEZ, Jean SEQUEIRA et Brian BRISCO. “Earth observation using radar data : an overview of applications and challenges”. In : *International Journal of Digital Earth* 1.2 (2008), p. 171-195. DOI : 10.1080/17538940802038317.
- [9] Gérald POPLU, Hubert RIPOLL, Sébastien MAVROMATIS et Jean BARATGIN. “How do expert soccer players encode visual information to make decisions in simulated game situations?” In : *Research quarterly for exercise and sport* 79.3 (2008), p. 392-398. DOI : 10.1080/02701367.2008.10599503.

- [10] Sébastien MAVROMATIS, Jean-Marc BOÏ, Rémy BULOT et Jean SEQUEIRA. "Texture analysis using directional local extrema". In : *Machine Graphics & Vision International Journal* 13.3 (2004), p. 289-302. ISSN : 1230-0535.
- [11] Gérald POPLU, Jean BARATGIN, Sébastien MAVROMATIS et Hubert RIPOLL. "What kind of processes underlie decision making in soccer simulation? An implicit-memory investigation." In : *International Journal of Sport and Exercise Psychology* 1.4 (2003), p. 390-405. DOI : 10.1080/1612197X.2003.9671727.
- [12] Sébastien MAVROMATIS et Jean SEQUEIRA. "3D medical imaging : from 2D images to 3D models". In : *Journal of Medical Informatics & Technologies* 3 (2002), p. 7-15.

## Conférences

- [13] Rémi GIRARD, Sébastien MAVROMATIS, Jean SEQUEIRA, Nicolas BELANGER et Guillaume ANOUFA. "A vision-based assistance key differentiator for helicopters autonomous scalable missions". In : *20th Towards Autonomous Robotic Systems Conference, TAROS*. London, England, 2019.
- [14] Antoine MONNEAU, Kouider Nacer M'SIRDI, Sébastien MAVROMATIS, Guillaume VARRA, Marc SALESSE et Jean SEQUEIRA. "Adaptive Prediction for Ship Motion in Rotorcraft Maritime Operations". In : *5th CEAS Conference on Guidance, Navigation and Control - EuroGNC19*. Milan, Italy, 2019.
- [15] Pierre ZOPPITELLI, Sébastien MAVROMATIS, Jean SEQUEIRA, Guillaume ANOUFA, Nicolas BELANGER et Francois-Xavier FILLIAS. "Embedding intelligent image processing algorithms : the new safety enhancer for helicopters missions". In : *44th European Rotorcraft Forum - ERF 2018*. Delft, Netherlands, 2018.
- [16] Dongmei HUANG, Yan WANG, Wei SONG, Jean SEQUEIRA et Sébastien MAVROMATIS. "Shallow-water Image Enhancement Using Relative Global Histogram Stretching Based on Adaptive Parameter Acquisition". In : *24th International Conference on Multimedia Modeling - MMM2018*. Bangkok, Thailand, fév. 2018. DOI : 10.1007/978-3-319-73603-7\_37.
- [17] Pierre ZOPPITELLI, Sébastien MAVROMATIS et Jean SEQUEIRA. "Ellipse Detection in Very Noisy Environment". In : *WSCG*. Pilsen, Czech Republic, mai 2017.
- [18] Zhang JIAHUI, Sébastien MAVROMATIS, Qingyan MENG, Jean SEQUEIRA et Zhang YING. "Using mathematical morphology on lidar data to extract information from urban vegetation". In : *IGARSS*. Beijing, China, juil. 2016.
- [19] Marc DANIEL, Cédric GUYOT, Sébastien MAVROMATIS et Arnaud POLETTE. "Réalité virtuelle et manipulation de surfaces B-splines". In : *Journées du GTMG*. Strasbourg, France, mar. 2012, p. 121-126.

- 
- [20] Sébastien MAVROMATIS, Christophe PALMANN et Jean SEQUEIRA. "Characterization of Similar Areas of Two 2D Point Clouds". In : *Advances in Visual Computing*. Crete, Greece : Springer Berlin Heidelberg, 2012, p. 509-516. DOI : 10.1007/978-3-642-33191-6\_50.
- [21] Florian GRAGLIA, Jean SEQUEIRA et Sébastien MAVROMATIS. "Real-time Walkthrough Rendering - Speed Improvement of Photon Mapping Algorithm". In : *Proceedings of the International Conference on Computer Graphics Theory and Applications (VISIGRAPP 2012)*. Rome, Italy, 2012, p. 291-298. DOI : 10.5220/0003817702910294.
- [22] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "A new geometric invariant to match regions within remote sensing images of different modalities". In : *SPIE Remote Sensing*. International Society for Optics et Photonics. Prague, Czech Republic, 2011, p. 818003-818003. DOI : 10.1117/12.898158.
- [23] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "Mesure de similarité entre sous-parties de nuages de points 2D". In : *TAIMA'11 Ateliers francophones sur le Traitement et l'analyse de l'information, méthodes et applications*. Hammamet, Tunisie, 2011.
- [24] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "A new approach for registering remote sensing images from various modalities". In : *SPIE Europe Remote Sensing*. International Society for Optics et Photonics. Berlin, Germany, 2009, p. 74770C-74770C. DOI : 10.1117/12.830248.
- [25] Bachir ZOUDJI, Hubert RIPOLL, Ludovic LLUCIA, Sébastien MAVROMATIS et Jean SEQUEIRA. "SimulFoot : un outil technologique pour l'analyse et la formation tactique en football". In : *Visions prospectives du professionnalisme sportif en Algérie*. 2009.
- [26] Ludovic LLUCIA, Sébastien MAVROMATIS, Sébastien PERROTTE, Paulo DIAS et Jean SEQUEIRA. "Camera location and aperture characterization using the transformation between a 2D plane and the image captured by the camera". In : *Image Analysis and Recognition*. Springer Berlin Heidelberg. Povia de Varzim, Portugal, 2008, p. 385-394. DOI : 10.1007/978-3-540-69812-8\_38.
- [27] Christophe PALMANN, Sébastien MAVROMATIS et Jean SEQUEIRA. "SAR image registration using a new approach based on the generalized hough transform". In : *ISPRS 2008*. Beijing, China, 2008.
- [28] Hubert RIPOLL, Ludovic LLUCIA, Sébastien MAVROMATIS, Bachir ZOUDJI et Jean SEQUEIRA. "SimulFoot entraînement, un outil de simulation pour les entraîneurs et les joueurs". In : *3ème Colloque International Football & Recherches*. Valenciennes, France, 2008.
- [29] Sébastien MAVROMATIS, Paulo DIAS et Jean SEQUEIRA. "3D reconstruction of soccer sequences using non-calibrated video cameras". In : *Image Analysis and Recognition*. Springer Berlin Heidelberg. Montreal, Canada, 2007, p. 1254-1264. DOI : 10.1007/978-3-540-74260-9\_111.

- [30] Arnaud LE TROTTER, Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. "Automatic Landmark Detection and Validation in Soccer Video Sequences". In : *Computer Vision and Graphics*. Springer Netherlands. Warsaw, Poland, 2007, p. 774-779. DOI : 10.1007/1-4020-4179-9\_112.
- [31] Arnaud LE TROTTER, Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. "Segmentation d'images en plages de couleurs dominantes". In : *TAIMA'05 Ateliers francophones sur le Traitement et l'analyse de l'information, méthodes et applications*. Hammamet, Tunisie, 2005.
- [32] Arnaud LE TROTTER, Sébastien MAVROMATIS et Jean SEQUEIRA. "Soccer field detection in video images using color and spatial coherence". In : *Image Analysis and Recognition*. Springer Berlin Heidelberg. Porto, Portugal, 2004, p. 265-272. DOI : 10.1007/978-3-540-30126-4\_33.
- [33] Arnaud LE TROTTER, Sébastien MAVROMATIS et Jean SEQUEIRA. "Automatic selection of a region of interest in 3D scene images : application to video captured scenes". In : *Advanced Concepts for Intelligent Vision Systems (ACIVS)*. Brussels, Belgium, 2004.
- [34] Hubert RIPOLL, Arnaud LE TROTTER, Jean BARATGIN, Sébastien MAVROMATIS, M FAISOLLE, F ZMILSONY, Gérald POPLU, JP PETIT et Jean SEQUEIRA. "The interest of simulation for research and training in sport : the example of football". In : *Proceedings of the Third International Sport Sciences Days, The analysis of elite performance in its contextual environment (INSEP)*. Paris, France, 2004.
- [35] Hubert RIPOLL, Arnaud LE TROTTER, Jean BARATGIN, Sébastien MAVROMATIS, M FAISOLLE, F ZMILSONY, Gérald POPLU, Jean-Philippe PETIT et Jean SEQUEIRA. "Un simulateur pour l'entraînement en football". In : *Journées Internationales des Sciences du Sport (INSEP)*. Paris, France, 2004.
- [36] Sébastien MAVROMATIS, Jean BARATGIN et Jean SEQUEIRA. "Analyzing team sport strategies by means of graphical simulation". In : *ICISP 2003*. Agadir, Morocco, 2003.
- [37] Sébastien MAVROMATIS et Jean-Marc BOI. "Texture analysis and scientific visualization". In : *SPIE Electronic Imaging 2003*. Santa Clara, California, 2003, p. 115-124.
- [38] Sébastien MAVROMATIS, Jean BARATGIN et Jean SEQUEIRA. "Reconstruction and simulation of soccer sequences". In : *MIRAGE*. Nice, France, 2003.
- [39] Hubert RIPOLL, Christophe AUBERT et Sébastien MAVROMATIS. "Mechanisms involved in the change of point of view into a 3D-image simulation in sport". In : *Xith European Congress of Sport Psychology*. Copenhagen, Denmark, 2003.
- [40] Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. "Texture analysis using directional local extrema". In : *International Conference on Computer Vision and Graphics (ICCVG)*. Agadir, Morocco, 2002.
- [41] Sébastien MAVROMATIS, Jean-Marc BOI, Rémy BULOT et Jean SEQUEIRA. "Toward the characterization of directional texture classes". In : *Proceedings of the 12th Portuguese Conference on Pattern Recognition*. Zakopane, Poland, 2002.

- [42] Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. “Tissue differentiation by using texture analysis”. In : *Conference on Biomedical Engineering (BIOENG)*. Istanbul, Turkey, 2001.
- [43] Sébastien MAVROMATIS, Jean-Marc BOI et Jean SEQUEIRA. “Medical image segmentation using texture directional features”. In : *Conference Proceedings Of The 23rd Annual International Conference Of The IEEE Engineering In Medicine And Biology Society*. T. 3. IEEE. Istanbul, Turkey, 2001, p. 2673-2676. DOI : 10.1109/IEMBS.2001.1017333.

## **Thèse**

- [44] Sébastien MAVROMATIS. “Analyse de texture et visualisation scientifique”. Thesis. Marseille, France : Université de la Méditerranée, déc. 2001.



## 6 Annexes

J'ai choisi de mettre en annexe trois documents qui viennent compléter la présentation de mes travaux.

Le premier est un article qui illustre mon attachement à la pluridisciplinarité. Il s'agit de résultats auxquels j'ai contribué dans le cadre du projet SimulFoot et qui mettent en évidence le lien privilégié entre l'information véhiculée par l'image et les processus cognitifs.

Le second document est le brevet « Similar Imaging » qui porte sur un estimateur de similarité entre deux images binaires et qui a été publié à l'international. Ce document montre un travail très différent de la rédaction d'articles scientifiques.

Le troisième est un article récent (2018) co-écrit avec différents membres d'Airbus Helicopters. Il s'agit de la première publication scientifique d'Airbus sur la définition d'un système d'aide au pilotage basé sur le traitement d'images.

**A How Do Expert Soccer Players Encode  
Visual Information to Make Decisions  
in Simulated Game Situations?**

## How Do Expert Soccer Players Encode Visual Information to Make Decisions in Simulated Game Situations?

Gérald Poplu, Hubert Ripoll, Sébastien Mavromatis, and Jean Baratgin

*The aim of this study was to determine what visual information expert soccer players encode when they are asked to make a decision. We used a repetition-priming paradigm to test the hypothesis that experts encode a soccer pattern's structure independently of the players' physical characteristics (i.e., posture and morphology). The participants were given either realistic (digital photos) or abstract (three-dimensional schematic representations) soccer game patterns. The results showed that the experts benefited from priming effects regardless of how abstract the stimuli were. This suggests that an abstract representation of a realistic pattern (i.e., one that does not include visual information related to the players' physical characteristics) is sufficient to activate experts' specific knowledge during decision making. These results seem to show that expert soccer players encode and store abstract representations of visual patterns in memory.*

*Key words:* abstraction, collective games patterns, expertise, repetition priming

There is a large body of research on sport expertise aimed at identifying the processes underlying experts' skills. Most of the studies highlighted the importance of perceptual processes. Several sport psychologists have tried to determine whether the perceptual skills of sport experts are related to the optometric and physical characteristics of the visual system (Williams & Grant, 1999). They concluded that experts' perceptual skills are not based on visual system differences as compared to novices. On the other hand, results of many studies show that sport experts are better than novices at recalling and recognizing information specific to their skill domain (Allard, Graham, & Paarsalu, 1980; Garland & Barry, 1991; Williams, Davids, Burwitz, & Williams, 1993; Williams, Davids, & Williams, 1999).

For example, Allard and colleagues (1980) showed that expert basketball players were better than novices at recognizing and recalling structured basketball game patterns (i.e., in which the organization abided by game rules). Garland and Barry (1991) and Williams and colleagues (1993) obtained similar results for expert football players and expert soccer players, respectively. In light of these results, Williams and colleagues (1999) suggested that skilled perception depends on an enhanced sport-specific knowledge base that enables experts to encode and retrieve stored information more effectively.

Similar results have been reported in other domains such as chess (Chase & Simon, 1973), music (Sloboda, 1976), and bridge (Charness, 1979). Sport psychologists have concluded that sport experts and experts in domains requiring cognitive involvement (e.g., chess) implement similar perceptual processes (Allard & Burnett, 1985).

Gobet and Simon's (1996) template theory, an extension of the chunking theory (Chase & Simon, 1973), provides a strong framework to explain expertise in these domains. It claims that during practice experts acquire domain-specific knowledge. According to Gobet (2001), this specific memory is a rich network of links connecting productions and schemas that become accessible through perceptual chunks. A chunk is defined as long-term memory information grouped in some meaning-

---

Submitted: March 25, 2003

Accepted: December 12, 2004

Gérald Poplu, Hubert Ripoll, and Sébastien Mavromatis are with the Sport Science Faculty—ESIL-LSIS at Aix-Marseille Universities. Jean Baratgin is with the Sport Science Faculty at Aix-Marseille Universities.

ful way so it can be remembered as a single unit. The interaction between perceptual chunks and high-level knowledge explains how, by recognizing and encoding perceptual chunks, experts rapidly access this network of knowledge to understand the subtleties of the pattern and rapidly make decisions (Gobet, 1998).

Thus, perceptual chunking of relevant visual information is a process underlying experts' skills. In their study on chess expertise, Chase and Simon (1973) explored the substrate of visual information that composes such chunks. By contrast, this visual information still remains to be determined in team sports, particularly soccer. The aim of the present study was to explore the visual information soccer experts encode when asked to make a decision. More precisely, we attempted to determine whether the players' physical characteristics, such as posture and morphology, are relevant to decision making, independent of the pattern's structure.

### *Primacy of the Pattern's Structure*

In soccer, both natural and simulated visual scenes are complex. To guide their behavior efficiently in such complex visual environments, experts must be able to select visual information relevant to them, while ignoring other irrelevant information (Davis, Welch, Holmes, & Shepherd, 2001). In other words, expert players most likely use abstract and schematic representations of visual scenes (Intraub, 1997). Here, the concept of *abstraction* refers to a change of representation that simplifies the initial representation by hiding details and preserving desirable properties (Saitta & Zucker, 2001). According to Hochberg (1981), parts of a visual scene are not equally represented in our perceptual system. Specific features of the visual world are fitted into a schematic map of the spatial layout. This schema incorporates detailed attended information and vague global information, such as gross size and shape (Intraub, 1997). Thus, we hypothesized that expert soccer players would identify a game pattern by encoding an abstract representation of the soccer scene and unintentionally disregard irrelevant visual information.

Visual soccer scenes contain at least two kinds of information related to the game: (a) the pattern's structure, which depends on the players' relative positions, orientations, and directions, and (b) the players' physical characteristics. Visual information related to the pattern structure informs about the players' current relational organization, while information related to the physical characteristics indicate the players' current actions.

Some results have shown that experts' superiority over novices on recall and recognition tasks involving domain-specific stimuli, whether schematic (Garland & Barry, 1991) or realistic (Williams et al., 1993), is linked to the pattern's structure. When the pattern structure obeys the

game rules, experts are better than novices. In the opposite case, however, expert performance is close to novice performance. This suggests that the pattern's structure outweighs players' physical characteristics in determining expert behavior. To simplify the visual scene, experts should be able to ignore these physical characteristics when making a decision, insofar as we can assume the visual information is not required for decision making. More recently, Williams and North (in press) showed that expert soccer players recognized a play configuration from the relational information (i.e., the general team structure), while nonexperts need to see the superficial display of players' body orientation and movement. Thus, we assumed expert soccer players would encode a game pattern's structure independently of players' physical characteristics to decide the next play.

To test this hypothesis, we used the repetition-priming paradigm elaborated by Poplu, Baratgin, Mavromatis, and Ripoll (2003). Generally, a repetition-priming paradigm consists of two phases (Schacter, 1987). During the priming phase, participants are given a certain number of items (e.g., soccer game patterns) to perform a particular task (e.g., solving a problem). Then during the test phase, they perform another task that can be either similar to or different from the first. Two kinds of items are presented: (a) primed items are presented during the priming phase; (b) unprimed items are new, not encountered during the priming phase. Repetition priming effects are evaluated by comparing the performance on primed items with that on unprimed items (Tulving & Schacter, 1990).

According to McNamara (1992), spreading-activation models can explain priming effects. In these models, memory is conceived of as a network of interconnected memory traces. Retrieval consists of activating the relevant memory trace in the network. Then, the activation spreads to all connected memory traces, decaying with respect to distance to activated trace and time. The accumulated residual activation facilitates their subsequent retrieval (McNamara, 1992).

Decision making during the priming phase can activate an expert's specific knowledge about the current game pattern. When primed situations are presented during the test phase, the residual activation can facilitate retrieval of relevant information and allows for faster decisions. However, if the visual information that allows experts to access their specific knowledge is not available, then no priming effect should be observed. In this case, no specific information stored in memory can be activated in either the priming or test phase.

To test the structure-primacy hypothesis, we used two kinds of presentation (see Figure 1). For the first, called original, the stimuli were digital photographs taken during soccer practice. For the second, called abstract, the same stimuli were used, but the players' physical

characteristics were deleted and replaced by three-dimensional (3D) crosses. To validate our hypothesis, experts must benefit from a priming effect in abstract presentation. In this case, the results should show that the abstract representation of a realistic pattern is sufficient to activate the specific knowledge experts need to make a decision.

## Method

### Participants

Four groups of players ( $N = 48$ ) volunteered to participate in the experiment. Each participant provided informed consent. Two groups consisted of 12 male expert soccer players, all members of a French semiprofessional club (Division Three). These players were considered to be experts in that they had been playing soccer regularly (i.e., at least four sessions and one match per week) in a club for more than 10 years. The other two groups consisted of 12 male students from the Faculty of Sport Sciences of Marseilles who did not play team sports in a

club. The participants' mean ages were 22.1 years ( $SD = 1.7$ ) for experts and 21.1 years ( $SD = 1.9$ ) for novices.

### Equipment and Materials

The experiment was conducted on a Macintosh G4 800 MHz biprocessor (Apple, Inc., Cupertino, CA) using software developed in C-Language for MacOS 9. The stimuli were presented on a 160 x 160 cm screen from a video projector placed at a distance of 3 m. The image on the screen was 105 x 85 cm, with a 20° viewing angle.

The stimuli were digital photographs (hereafter called source stimuli) and 3D graphical representations that were made by new 3D reconstruction software (Mavromatis, Baratgin, & Sequeira, 2003), which transforms a two-dimensional video frame into a 3D representation from a source stimulus. This reconstruction makes it easy to modify information. For the original presentation, the stimuli were digital photos taken during soccer practice. The digital camera was placed 2 m high at the centerline 2 m away from the playing area. Its viewing angle covered the right half of the soccer pitch. The stimuli for the abstract presentation were 3D representations of the source stimuli. The viewpoint was the same



**Figure 1.** Examples of stimuli used in the present experiment. The two kinds of pictures were: (a) original video frame, and (b) three-dimensional abstract representation without the physical characteristics of the players (crosses are shown in place of the players).

as in the source stimuli. Players' relative positions and orientations were the same; however, their physical characteristics were deleted and replaced by 3D crosses.

The patterns were selected with the help of national level professional coaches using the following criteria: (a) the patterns had to be offensive game situations with 10 players (equal number of offense and defense players), (b) each game situation had to offer at least two good solutions (equal probability). Sixty different patterns were selected and randomly divided into two categories. The first was called P and consisted of 40 primed patterns; the second was UP and consisted of 20 unprimed patterns.

### Procedure

Each group of experts and novices was confronted with a different kind of presentation. The procedure was the same in the two kinds of presentation and consisted of three phases: familiarization, priming, and test. During priming, participants were presented with problem situations. During the test phase they were again presented with problem situations, but some were primed and others were new or unprimed.

*Training Phase.* This relatively brief phase (10 trials) was designed to familiarize the participants with the task and the experimental materials.

*Priming Phase.* First, a preparation message ("Ready!") was displayed on the screen for 1 s. As soon as the message disappeared, a pattern was shown. The instructions were to make a quick and accurate decision concerning the next action of the player in possession of the ball. The participants could choose from three possible actions (shoot, pass, or dribble). To respond, they clicked on the mouse to record their response time and immediately say aloud what response they chose. During this phase, 40 P configurations were presented at random, for 5 s each (regardless of whether the participant responded). The presentation time was the same on each trial.

*Test Phase.* After the priming phase, the experiment continued in the same manner (same task and procedure). However, the presentation time for each trial varied, because the patterns disappeared from the screen as soon as the participant responded. During this phase, 40 patterns were presented in random order. Twenty were P patterns (already presented during priming), and 20 were UP patterns (never seen).

### Design

The data were analyzed using factorial analyses of variance in which condition (prime and unprimed) was a within-participant factor, and expertise (expert vs. novice) and kind of presentation (original vs. abstract) were between-participant factors. The dependent vari-

ables were response time (ms) and response consistency (score). Priming effects were evaluated by comparing the performances obtained for P and UP conditions in the test phase. Response consistency was achieved if participants gave the same response on the priming and test phases to the same pattern. Thus, response consistency enabled us to check response stability for a given pattern and to see if the participants responded in a random manner. Response accuracy was not a relevant measure for the needs for our study, because each game pattern had to offer at least two good solutions (equal probability). Therefore, the comparison between participants' responses for a particular pattern and a standard solution determined by professional coaches was not possible. Participants' response consistency was compared with random responding using Student's *t* tests. Significant effects were followed up using the Newman-Keuls procedure, when appropriate. An alpha level of .05 was used for all statistical tests.

## Results

### Response Time

Experts ( $M = 1,367$  ms,  $SD = 508$ ) were generally faster than novices ( $M = 2,115$  ms,  $SD = 760$ ),  $F(1, 44) = 19.42$ ,  $p < .05$ ,  $\eta^2 = .12$ ,  $\eta^2_p = .31$ . The effect of presentation factor was significant,  $F(1, 44) = 11.55$ ,  $p < .05$ ,  $\eta^2 = .07$ ,  $\eta^2_p = .21$ . Participants' response time was longer in the original presentation ( $M = 2,030$  ms,  $SD = 844$ ) than in the abstract ( $M = 1,453$  ms,  $SD = 811$ ). A condition main effect was found,  $F(1, 44) = 14.43$ ,  $p < .05$ ,  $\eta^2 = .003$ ,  $\eta^2_p = .25$ . In addition, there was an interaction between expertise and condition,  $F(1, 44) = 6.09$ ,  $p < .05$ ,  $\eta^2 = .0004$ ,  $\eta^2_p = .12$  (see Figure 2). Post hoc tests revealed that experts' response time was longer in condition UP ( $M = 1,459$  ms,  $SD = 550$ ) than in P ( $M = 1,276$  ms,  $SD = 472$ ). On the other hand, novices' response time was statistically the same for P ( $M = 2,096$  ms,  $SD = 771$ ) and UP ( $M = 2,134$  ms,  $SD = 770$ ) conditions.

### Response Consistency

Experts ( $M = 15$ ,  $SD = 1.4$ ) were more consistent than novices ( $M = 14.1$ ,  $SD = 1.3$ ) regardless of the kind of presentation,  $F(1, 44) = 4.87$ ,  $p < .05$ . No presentation main effect was found,  $F(1, 44) = 0.28$ ,  $p > .05$ . There was no interaction between expertise and presentation,  $F(1, 44) = 0.54$ ,  $p > .05$ . Moreover, experts,  $t(11) = 5.15$ ,  $p < .05$ ;  $t(11) = 6.03$ ,  $p < .05$ , and novices,  $t(11) = 5.75$ ,  $p < .05$ ;  $t(11) = 5.40$ ,  $p < .05$ , did not respond in a random manner. This suggests that all the participants accomplished the task correctly.

## Discussion

The goal of this research was to explore the visual information expert soccer players encode when they have to complete a three-alternative (i.e., shoot, pass or dribbles) forced choice task. We used a priming paradigm to test the hypothesis, according to which expert soccer players encode a game pattern's structure independently of the players' physical characteristics. Several important results were obtained. First, experts were faster and more consistent than novices; these results are coherent with those reported in the literature on expertise. Second, only experts benefited from priming effects. Finally, experts benefited from priming effects independent of the presentation type (i.e., abstract or original). In other words, experts were faster, whatever the level of abstraction, when the stimuli were previously presented (i.e., P) than when they were not (i.e., UP).

In his work on chess player expertise, de Groot brought out two important facts. First, expert players demonstrated superiority in making effective decisions, and, second, they were capable of quickly extracting relevant cues from game situations (Gobet, 1993). Based on these results, decision making speed and effectiveness are the most salient properties of cognitive expertise. Results from several studies on sports expertise have supported these observations. For example, Poplu, Laurent, Baratgin, and Ripoll (2000) showed that expert basketball players were faster than novices at judging the

similarity between two basketball game patterns. More recently, Poplu, Baratgin, Mavromatis, and Ripoll (2003) showed that expert soccer players were faster than novices at making decisions in simulated game situations.

Thus, if we consider de Groot's observations, which can be explained by the template theory (Gobet & Simon, 1996), it follows that expert soccer players should be faster than novices at making decisions about simulated soccer situations. Indeed, if expert soccer players have specific knowledge they can quickly access by recognizing perceptual chunks, then they should be faster than novices at making decisions. Accordingly, by showing that experts were faster than novices, our results are in line with the literature on expertise.

Another important finding was that only experts benefited from priming effects. According to McNamara (1992), the consistent observation of priming effects suggests they are the product of memory-retrieval mechanisms. Numerous studies explain these effects in terms of spreading activation (Anderson, 1983; Collins & Loftus, 1975). This concept is mainly used in the connectionist approach to memory, in which knowledge is organized into networks containing nodes that symbolize information and arcs that represent connections, thus spreading the activation from node to node. Although the details differ between theories, the main idea is that retrieving a piece of information from memory amounts to activating its internal representation. The activation then spreads to other network-associated information, and so on

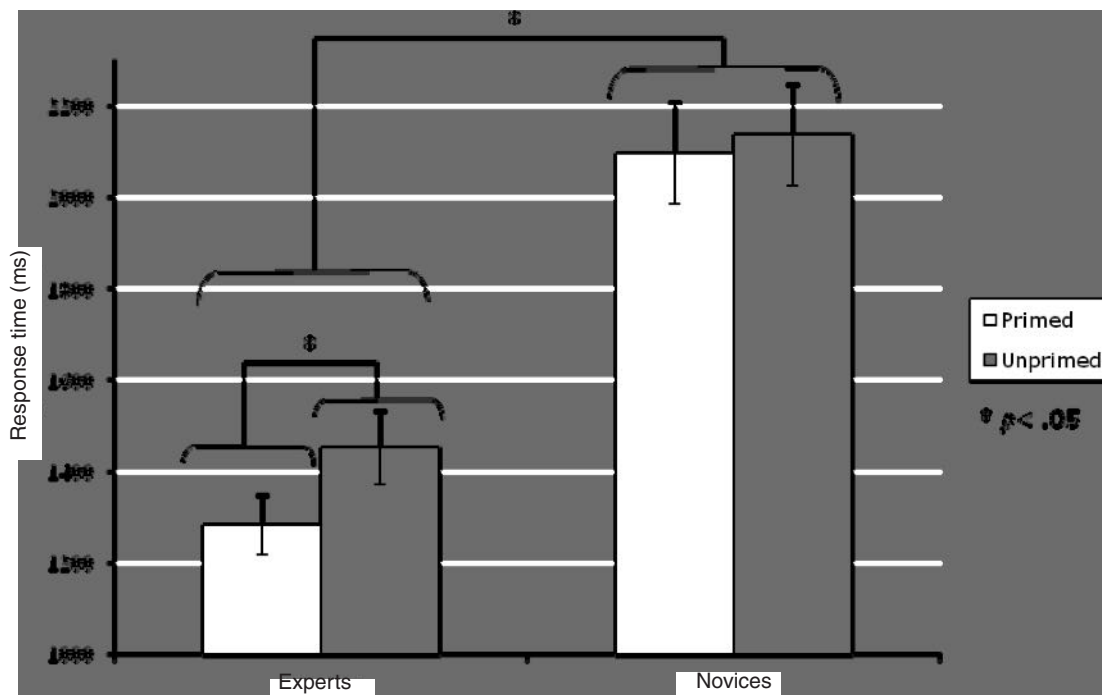


Figure 2. Mean response time and standard error as a function of expertise and condition.

(Costermans, 2001). The accumulated residual activation facilitates subsequent retrieval of that information (McNamara, 1992). In other words, priming occurs when mental representations have been recently reactivated. Moreover, Ripoll, Baratgin, Laurent, Courrieu, and Ripoll (2001) developed a neurocomputational model with architecture like a connectionist network to account for expert basketball players' performance in a similar task. The result supported the hypothesis that expert players' memory is organized into such a network. Based on this idea, decision making during the priming phase in the present study can be assumed to activate knowledge about the current game pattern. During the test phase, when primed situations were presented, the residual activation facilitated retrieval of relevant information and, therefore, allowed for faster decisions. Only the experts benefited from a priming effect, because they had specific knowledge (perceptual and conceptual) about soccer.

Finally, the most important result showed that the experts benefited from priming effects on both the original and abstract presentations. This means that visual information allowing experts to access their specific knowledge was available in the original and abstract stimuli. Moreover, the visual information modifications between these stimuli, namely, abstraction of the players' physical characteristics, did not disrupt the experts' behavior. This suggests that an abstract representation of a realistic pattern is sufficient to activate experts' knowledge when making a decision.

By showing that expert soccer players do not need to see physical characteristics to make decisions, the results tend to validate the structure-primacy hypothesis. Moreover, the results showed that response time varied as a function of abstraction level. Experts responded faster to the abstract presentation than the original presentation. Insofar as less information was available in the abstract presentation, the quantity of information processed can explain the difference in experts' performance. This could mean that experts process all information available in each presentation. More precisely, it could mean that experts process visual information about players' posture and morphology when it is available.

Interpretation of this finding could either support or contradict the structure-primacy hypothesis. Indeed, players' physical characteristics could be considered as either relevant or useless visual information. It was assumed here that visual information was useless because of the following findings: (a) the experts benefited from a priming effect independent of the abstraction level, and (b) their response consistency did not decrease when the amount of available information decreased. In other words, the experts' behavior was not disrupted when visual information was removed.

The availability of this information seemed to spread the experts' attention (Davis et al., 2001) and

thereby increase the time required to take in relevant information and make a decision. Accordingly, these results were not inconsistent with the structure-primacy hypothesis. The present experiment seems to show that expert soccer players encode and store abstract representations of visual patterns in long-term memory. These representations may retain some critical components of the pattern structure (Hochberg, 1981; Intraub, 1997), such as players' relative position and orientation but not their physical characteristics. Future research could investigate the properties of these abstract representations more precisely.

These results are also of applied interest in that they provide some answers to the question: What kind of representation should be used in simulations? It seems unnecessary to use realistic visual scenes. Abstract representations seem to be as informative as realistic ones for expert soccer players. Thus, one can use either a realistic representation or an abstract (less complex) one. These results support coaches' use of a black board, which is an extreme form of abstraction.

## References

- Allard, F., & Burnett, N. (1985). Skill in sport. *Canadian Journal of Psychology, 39*, 294–312.
- Allard, F., Graham, S., & Paarsalu, M. E. (1980). Perception in sport: Basketball. *Journal of Sport Psychology, 2*, 14–21.
- Anderson, J. R. (1983). *The architecture of cognition*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology, 33*, 1–16.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407–428.
- Costermans, J. (2001). *Les activités cognitives* [Cognitive activities]. Paris: De Boeck.
- Davis, G., Welch, V. L., Holmes, A., & Shepherd, A. (2001). Can attention select only a fixed number of objects at a time? *Perception, 30*, 1227–1248.
- Garland, D. J., & Barry, J. R. (1991). Cognitive advantage in sport: The nature of perceptual structures. *American Journal of Psychology, 104*, 211–228.
- Gobet, F. (1993). *Les mémoires d'un joueur d'échecs* [Memories of a chess player]. Fribourg, Switzerland: Éditions Universitaires.
- Gobet, F. (1998). Expert memory: A comparison of four theories. *Cognition, 66*, 115–152.
- Gobet, F. (2001). *Is experts' knowledge modular?* Paper presented at the 23rd Meeting of the Cognitive Science Society, Mahwah, NJ.
- Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*, 1–40.



- Hochberg, J. (1981). On cognition in perception: Perceptual coupling and unconscious inferences. *Cognition, 10*, 127–134.
- Intraub, H. (1997). The representation of visual scenes. *Trends in Cognitive Sciences, 1*, 217–222.
- Mavromatis, S., Baratgin, J., & Sequeira, J. (2003). *Toward the design of a simulator to analyze team sport strategies*. Paper presented at the Mirage Congress, Rocquencourt, France.
- McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1173–1190.
- Poplu, G., Baratgin, J., Mavromatis, S., & Ripoll, H. (2003). What kind of processes underlie decision making in soccer simulation? An implicit memory investigation. *International Journal of Sport and Exercise Psychology, 1*, 390–405.
- Poplu, G., Laurent, E., Baratgin, J., & Ripoll, H. (2000). Modalités d'encodage perceptif de configurations de jeu de basket-ball [Perceptual encoding of basketball game configurations]. *Proceedings of the International Congress of the French Society of Sport Psychology* (pp. 290–291). Paris: INSEP.
- Ripoll, H., Baratgin, J., Laurent, E., Courrieu, P., & Ripoll, T. (2001). Mechanisms underlying the activation of knowledge bases in the identification of basketball play configurations by expert and non-expert players. In A. Papaioannou, M. Goudas, & Y. Theodorakis (Eds.), *In the dawn of the new millennium: Proceedings of the 10th World Congress of Sport Psychology*, (Vol. 2, pp. 283–285). Thessaloniki, Greece: Christodoulidi Publications.
- Saitta, L., & Zucker, J. D. (2001). A model of abstraction in visual perception. *Applied Artificial Intelligence, 15*, 761–776.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501–518.
- Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28*, 1–16.
- Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301–306.
- Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1993). Cognitive knowledge and soccer performance. *Perceptual and Motor Skills, 76*, 579–593.
- Williams, A. M., Davids, K., & Williams, J. G. (1999). *Visual perception and action in sport*. London: E. & F. N. Spon.
- Williams, A. M., & Grant, A. (1999). Training perceptual skill in sport. *International Journal of Sport Psychology, 30*, 194–220.
- Williams, A. M., & North, J. S. (in press). Some constraints on recognition performance in soccer. In D. Araujo, H. Ripoll, & M. Raab (Eds.), *Perspectives on cognition and action*. New York: Nova Science Publishers, Inc.

### Authors' Notes

---

This research was supported by a grant from the French Ministry of Research (Research Fellowship ACI Cognitive, #90). The authors would like to thank Eric Laurent, Jean-Philippe Petit, and Jean-Daniel Zucker for their comments on the research presented in this paper. Please address all correspondence concerning this article to H. Ripoll, ESIL-LSIS Université de la Méditerranée, 163 avenue de Luminy CP 910, 13288 Marseille cedex 09, France.

E-mail: Hubert.ripoll@univmed.fr

## **B Method and system for estimating a similarity between two binary images**



US009171228B2

(12) **United States Patent**  
**Mavromatis et al.**

(10) **Patent No.:** **US 9,171,228 B2**  
(45) **Date of Patent:** **Oct. 27, 2015**

(54) **METHOD AND SYSTEM FOR ESTIMATING A SIMILARITY BETWEEN TWO BINARY IMAGES**

(2013.01); G06T 2207/10032 (2013.01); G06T 2207/20221 (2013.01); G06T 2207/30181 (2013.01)

(75) Inventors: **Constantin Mavromatis**, Marseilles (FR); **Christophe Palmann**, Marseilles (FR); **Jean Sequeira**, Roquevaire (FR)

(58) **Field of Classification Search**  
None  
See application file for complete search history.

(73) Assignees: **UNIVERSITE D' AIX-MARSEILLE**, Marseilles (FR); **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE**, Paris (FR)

(56) **References Cited**  
U.S. PATENT DOCUMENTS

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 120 days.

2002/0054210 A1\* 5/2002 Glier et al. .... 348/149  
2004/0228502 A1\* 11/2004 Bradley et al. .... 382/100  
2009/0316240 A1\* 12/2009 Hara ..... 359/33  
2010/0322522 A1\* 12/2010 Wang et al. .... 382/218

OTHER PUBLICATIONS

(21) Appl. No.: **14/002,518**

International Search Report (PCT/ISA/210) mailed on Aug. 14, 2012, by the French Patent Office as the International Searching Authority for International Application No. PCT/FR2012/050438. Written Opinion (PCT/ISA/237) mailed on Aug. 14, 2012, by the French Patent Office as the International Searching Authority for International Application No. PCT/FR2012/050438. Palmann et al., "A new approach for registering remote sensing images from various modalities", Proceedings of SPIE, Jan. 2009, vol. 7477, pp. 74770C-1 to 74770C-11, XP055004186.

(22) PCT Filed: **Mar. 2, 2012**

(86) PCT No.: **PCT/FR2012/050438**  
§ 371 (c)(1),  
(2), (4) Date: **Sep. 18, 2013**

(87) PCT Pub. No.: **WO2012/117210**  
PCT Pub. Date: **Sep. 7, 2012**

(Continued)

(65) **Prior Publication Data**  
US 2014/0044362 A1 Feb. 13, 2014

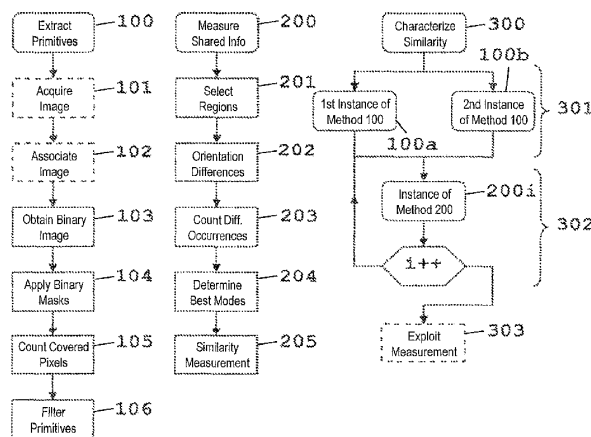
*Primary Examiner* — Hadi Akhavanik  
(74) *Attorney, Agent, or Firm* — Buchanan Ingersoll & Rooney PC

(30) **Foreign Application Priority Data**  
Mar. 2, 2011 (FR) ..... 11 51710

(57) **ABSTRACT**  
The invention relates to a method and a system for estimating the resemblance between two images of optionally different modalities. More particularly, the invention makes it possible to characterize a similarity between two binary images according to a formula that allows registration of images acquired in the fields of teledetection, medical imaging, and industrial vision.

(51) **Int. Cl.**  
**G06K 9/00** (2006.01)  
**G06K 9/62** (2006.01)  
**G06T 7/00** (2006.01)  
(52) **U.S. Cl.**  
CPC ..... **G06K 9/6202** (2013.01); **G06T 7/0028**

**14 Claims, 4 Drawing Sheets**



(56)

**References Cited**

OTHER PUBLICATIONS

Shams et al., "Image Registration in Hough Space Using Gradient of Images", 9<sup>th</sup> Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications, Dec. 2007, pp. 226-232, XP055004590.

Shams et al., "Gradient Intensity: A New Mutual Information-Based Registration Method", IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2007, pp. 1-18, XP031114655.

Gottesfeld, "A Survey of Image Registration Techniques", ACM Computing Surveys, Dec. 1992, pp. 325-376, XP002942558.

\* cited by examiner

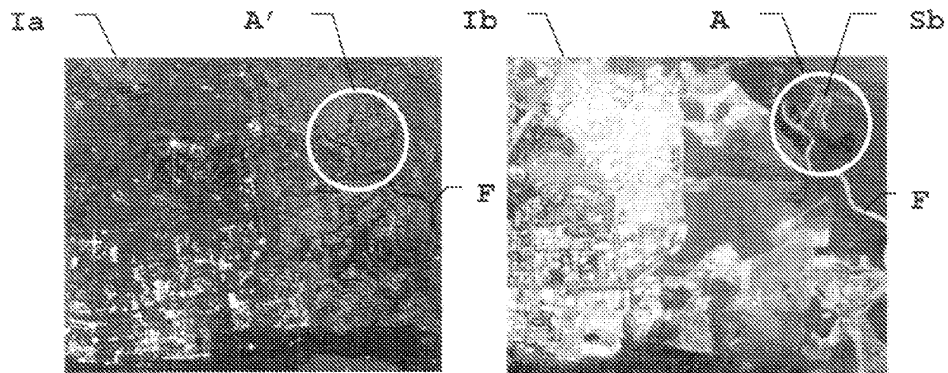


FIG. 1a

FIG. 1b

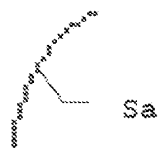


FIG. 2a

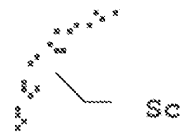


FIG. 2c

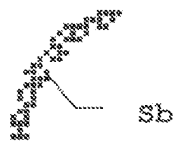


FIG. 2b



FIG. 2d

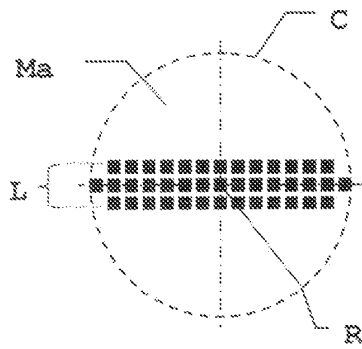


FIG. 3a

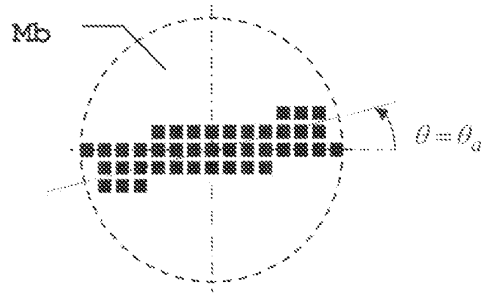


FIG. 3b

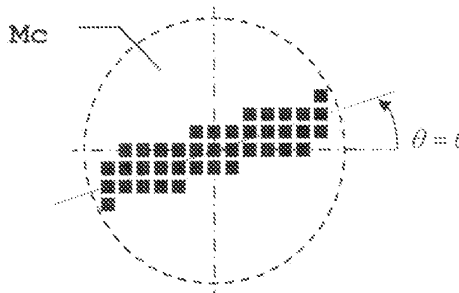


FIG. 3c

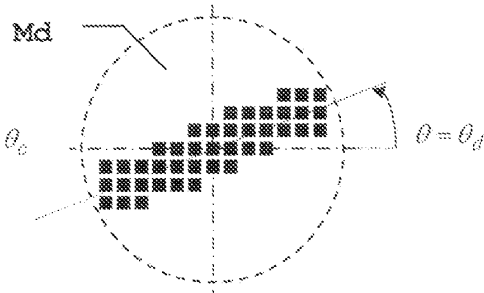


FIG. 3d

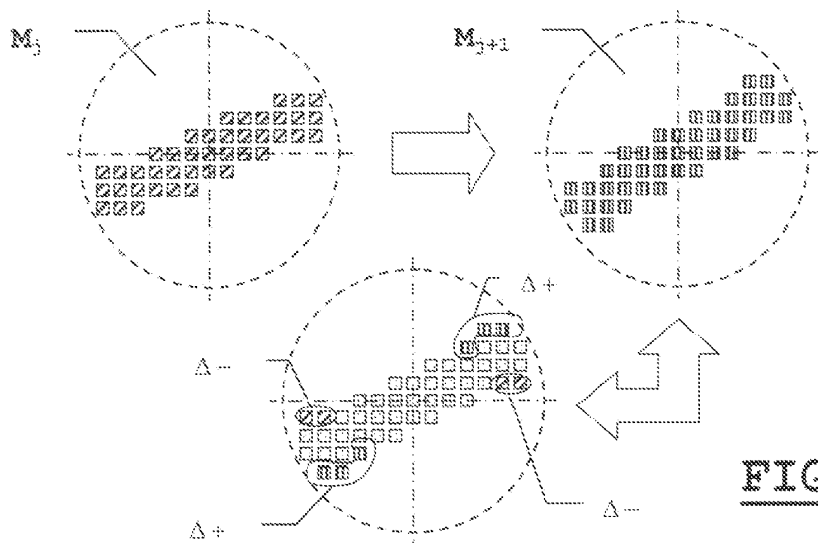
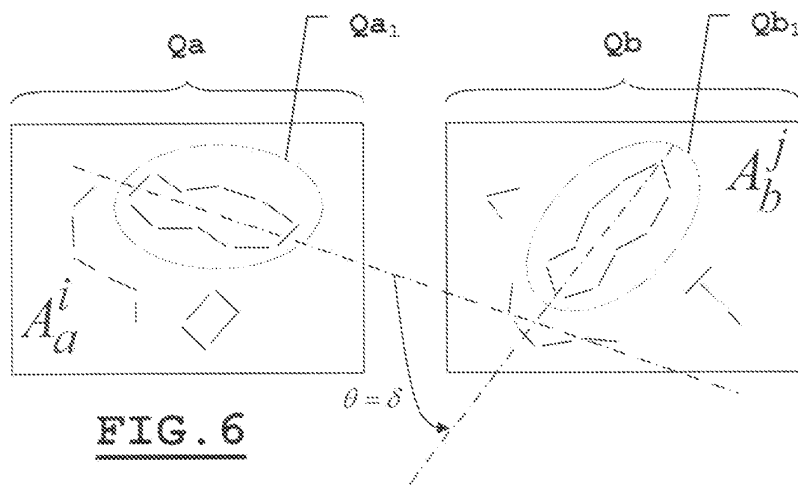
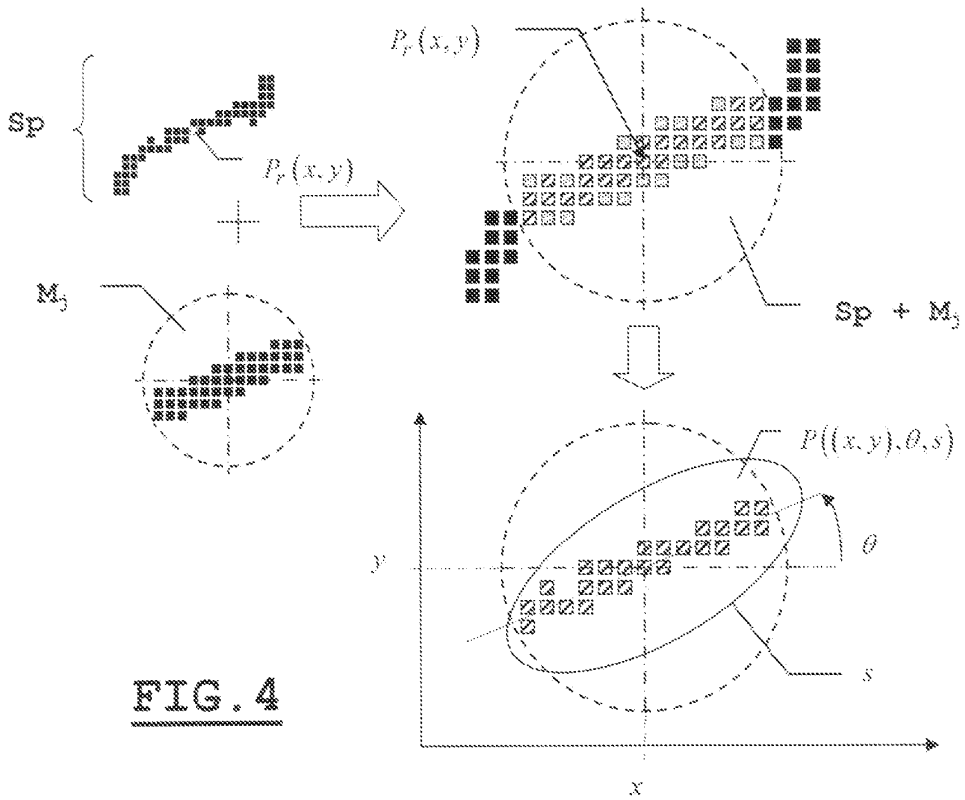


FIG. 5



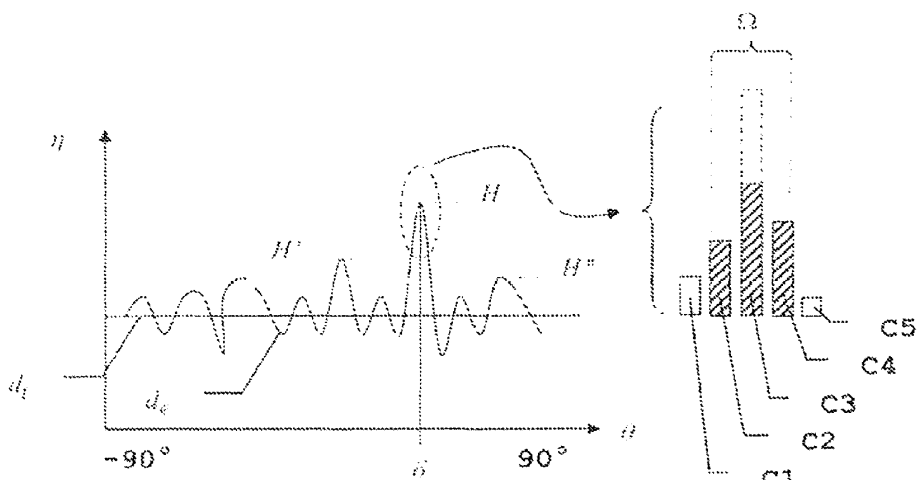


FIG. 7

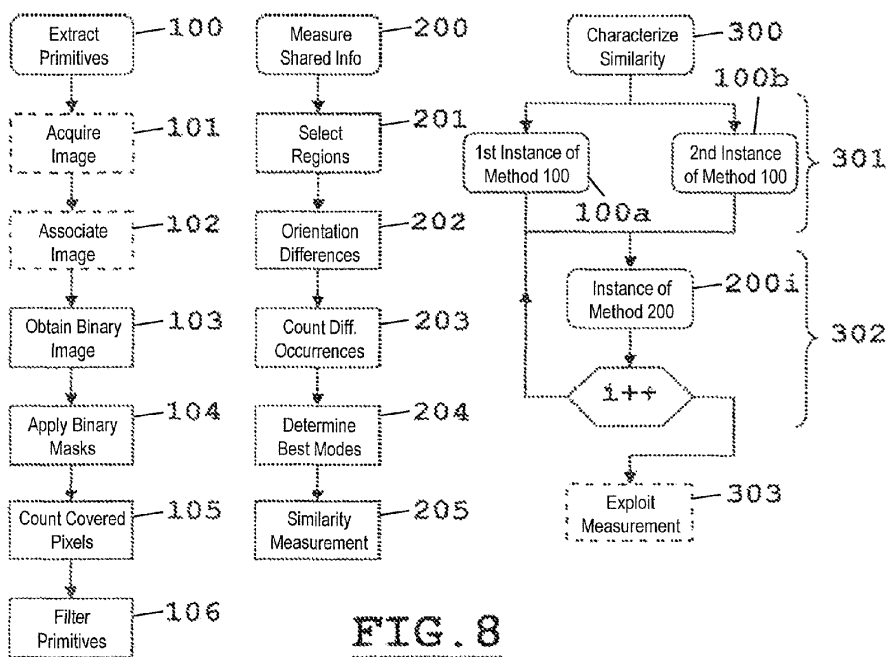


FIG. 8



1

## METHOD AND SYSTEM FOR ESTIMATING A SIMILARITY BETWEEN TWO BINARY IMAGES

The invention relates to a method and associated system for estimating the resemblance between two images optionally with different modalities. More particularly, the invention makes it possible to characterize a similarity between two binary images.

An image conveys a significant quantity of information. It is thus and for example possible to use images acquired at different moments to assess the evolution of a particular piece of information. These images are generally of the same modality.

The term “modality” refers to a physical characterization associated with a data capture. Thus, two images with different modalities contain different pieces of information and are often supported by different representations. For example, the pixels of a radar image are represented by complex numbers that correspond to the response to an emitted wave, said response in turn depending on parameters related to said wave. The pixels of a multispectral optical image are alternatively vectors—having as many components as there are bands—whereof each component represents the emission of light in a given frequency band.

In remote sensing, it is essentially radar images or multispectral images that are manipulated. In medical imaging, images coming from scanners or Nuclear Magnetic Resonance imaging devices are used. In industrial vision applications, it is common practice to use visible images or infrared images. Irrespective of the modality of the concerned images, to assess the evolution of a piece of information, it is generally necessary to recalibrate the images, i.e., to match them through a geometric transformation that makes it possible to determine, for each pixel of a first image, the pixel of a second image to which it corresponds.

The problem of recalibrating images of the same modality has been widely studied. Solutions are currently in use or published, although their respective effectiveness and precision vary. Examples of these solutions in particular include the ICP (Iterated Closest Point) and RANSAC (RANDOM Sample Consensus) methods, which process any matching of two clouds of points.

The ICP-type methods are based on an iterative pairing-estimation process: the points of the first cloud are paired with those of the second cloud according to a spatial proximity criterion, then all of the pairs thus produced make it possible to estimate a transformation. However, the process does not make it possible to differentiate between what is shared by the two clouds and what is not. These methods of course converge—sometimes at the expense of a large number of iterations—but to a large extent integrate noise in the determination of the matching. They therefore lack precision. Consequently, all image recalibrations using the obtained estimates are relatively imprecise.

RANSAC-type methods make it possible to highlight outliers, but do not always converge. Generally, these stochastic methods consist of evaluating a quantity of recalibrations of clouds of points done by randomly drawn geometric transformations. More specifically, these methods generally consist of an iterative process that seeks to estimate local transformations connecting small numbers of pairs of points drawn randomly. Contrary to ICP methods, these methods succeed in discriminating shared information. However, to maximize the likelihood of estimating the desired transformation, it is necessary to estimate a large number of local transformations, in particular if the quantity of information

2

shared between two clouds is small. These methods are therefore relatively ineffective when the shared information portion is small relative to the noise. However, this is clearly the case in the field of remote sensing in particular.

Other techniques further exist for measuring the similarity of clouds of points: for example, geometric invariants. According to these techniques, the objects are no longer simple points, but are described in a complex manner, the description not depending on the desired type of geometric transformation. To that end, it is necessary to formalize such objects. The latter then acquire the status of “primitive”. It is further necessary to define a metric for measuring the resemblance of two primitives. In general, the geometric invariants must satisfy different constraints to be able to be used correctly. Aside from the invariants relative to the desired type of transformation, the primitives must meet criteria of uniqueness (two different primitives must have different descriptions), stability (if a primitive is slightly deformed, then its description is only slightly modified), independence (understood as meaning of the components that form the description vector). The use of these techniques further requires that said primitives be uniformly distributed in the clouds, and still more difficult, that the primitives be able to be extracted irrespective of the available capture means and irrespective of the capture conditions of the studied scene. In practice, it is nearly impossible to meet all of the imposed constraints.

The problem of recalibrating images with different modalities remains particularly open and sensitive. In fact, there is no known method or system making it possible to assess a similarity or to perform a recalibration of images when the latter have different modalities. The main pitfall lies in the fact that it is difficult to compare the content of pixels that refer to different entities.

Improving the recalibration of images with different modalities is of major interest in particular in the field of remote sensing, where many satellites, or even satellite constellations, have been—or will soon be—launched, each satellite having sensors of different natures. It is therefore desirable to be able to integrate the information extracted from each of these data sources into the same model, which requires prior recalibration.

In the field of remote sensing, there is also a general interest in a particular entity designated on a monitor allowing the display of a first image with a given modality. An effort is then made to find that entity in a set of available images, regardless of their respective modalities, even when said entity does not explicitly appear within other images. This analysis currently requires manual, even tedious processing, or the estimates resulting from that processing are relatively imprecise or contradictory.

The invention makes it possible to respond to all of the drawbacks raised by the known solutions. It in particular makes it possible to precisely quantify the resemblance between a region of an image of a first modality and a region of a second image independent of its modality. The implementation of the invention makes it possible to automate the analysis of multiple images using an analysis system suitable for that purpose. The invention is furthermore particularly high-performing to estimate the resemblance between two images of a same modality.

To that end, a method is provided for characterizing a similarity between two binary images, said method being carried out by a processing unit of a system for analyzing similarities of images further including storage means cooperating with said processing unit to store binary images. According to the invention, the method includes a step for extracting a first set of primitives from a first binary image and

3

a second set of primitives from a second binary image, the first and second binary images being stored beforehand by the system. Each primitive is associated with a pixel of a binary image and is characterized by the position of said pixel and by the orientation of an orientation line passing through the latter. Such a method further includes a step for measuring the presence of a common piece of information within said binary images and delivering a similarity indicator  $\epsilon^*$  therefor, said indicator being calculated by the processing unit as the relative deviation between an overall maximum  $H$  and the average of two other local maximums  $H'$  and  $H''$ , i.e.,

$$\epsilon^* = \frac{|H - \bar{H}|}{2} \text{ where } \bar{H} = \frac{H' + H''}{2},$$

said maximums being determined from a distribution of the different orientations of pairs of primitives respectively resulting from two extracted sets of primitives.

According to one preferred embodiment, the extraction step may consist of:

- obtaining a binary image stored by the system;
- applying a plurality of binary masks to each pixel making up said binary image, the binary masks each consisting of an orientation line whereof the thickness and orientation are predetermined and which describes a diameter of a circular window whereof the radius is also predetermined;
- counting, for each binary mask applied to a pixel—called referring pixel—the number of pixels adjacent to said referring pixel that are covered by said binary mask;
- defining, then storing as primitive associated with a referring pixel, a triplet  $P((x,y)\theta,s)$  comprising the position  $(x,y)$  of the referring pixel in the binary image; the largest number  $s$ , called the score, of adjacent pixels covered by one of the binary masks applied to the referring pixel; as well as the orientation  $\theta$  of the orientation line of said binary mask.

To avoid considering isolated points or insignificant areas of a binary image, the invention provides that a primitive may only be stored if its score is greater than or equal to a predetermined minimum number  $s_{min}$ .

Furthermore, to improve the robustness of the method for a given noise level, the invention provides that only the primitives whereof the referring pixels are distant by a minimum and predetermined number of pixels  $D_{min}$  can be stored.

To optimize the calculation times for the calculations done by the processing unit, or even to maximize the likelihood of identifying a piece of information shared by two binary images, the measurement step carried out by the processing unit may preferably consist of:

- determining a region of predetermined size in the first and second binary images to thus form a pair of regions;
- counting and storing the occurrences of the different orientations for all pairs of primitives respectively extracted from said pair of regions;
- calculating the similarity indicator  $\epsilon^*$  between the two regions of the pair.

Optionally, a method according to the invention may include a prior step for associating and storing said first and second binary images, respectively, with two images stored or acquired by the image similarity analysis system.

According to one embodiment, a method according to the invention may further include a step for recalibrating two stored or acquired images or even identifying a remarkable element located in the two images from the calculated simi-

4

ilarity indicator  $\epsilon^*$ . In that case, such a method may include a step for displaying said recalibrated images and/or the remarkable element using a man-machine interface suitable for the image similarity analysis system.

According to a second subject-matter, the invention relates to a system for analyzing the similarity of binary images including a processing unit cooperating with storage means, said processing unit implementing a method for characterizing a similarity between two binary images according to the invention.

According to a third subject-matter, the invention relates to a system for recalibrating acquired or stored images. Such a system includes means for storing and/or acquiring images, said means cooperating with a processing unit, which in turn further cooperates with storage means for implementing the method—according to the invention—making it possible to recalibrate two stored or acquired images or even to identify a remarkable element located in the two images from the calculated similarity indicator  $\epsilon^*$ .

The processing unit of such a system may additionally deliver the two recalibrated images or the identified remarkable image to a man-machine interface.

To communicate the similarity measurement to an operator, whatever the system according to the invention, the latter may further include a man-machine interface for delivering the similarity indicator  $\epsilon^*$ .

Furthermore, to increase its interactivity, the processing unit of a system according to the invention may include a man-machine interface for selecting one image from among several and/or a particular region thereof.

Other features and advantages will appear more clearly upon reading the following description and examining the accompanying figures, in which:

FIGS. 1a and 1b respectively show two images of different modalities;

FIGS. 2a to 2d describe linear structures of a binary image; FIGS. 3a to 3d describe different configurations of a binary mask making it possible to extract, according to the invention, a primitive from a binary image;

FIG. 4 describes the application of a binary mask according to one embodiment of the invention;

FIG. 5 illustrates the steps for extracting a primitive according to one preferred embodiment of the invention;

FIG. 6 describes an example of matching primitives respectively extracted from regions of two binary images;

FIG. 7 illustrates an example of the determination of maximums of a distribution of differences of orientations of primitives to develop a similarity indicator according to the invention;

FIG. 8 shows a summary of the steps of a method according to the invention for measuring the similarity of two images.

FIGS. 1a and 1b respectively describe a radar image 1a and an optical image 1b. These figures illustrate a first example embodiment of the invention. The images 1a and 1b correspond to a same geographical zone. The latter information may for example be deduced from their respective acquisition parameters. Nevertheless, two images never match perfectly. This can be seen through the two images 1a and 1b described in connection with FIGS. 1a and 1b. It is therefore necessary to recalibrate said images to be able to compare what is comparable, i.e., to match the pixels of the two images to the same entities. For example, in the image 1b, it is possible to visually detect, in the top right (region A—above the river F—embodied by a white circle in FIG. 1b), a structure S that does not appear in the image 1a. The corresponding region A' in the latter nevertheless contains complementary information that may be important for the operator. The invention

5

allows the automatic and precise location, on the image 1a, of the analogue A' of the region A previously identified in the image 1b. By extension, and insofar as the method for estimating the similarity according to the invention is carried out completely automatically by a fitting system, this method may be applied to all images—regardless of their respective modalities—corresponding to that geographical zone. It is for example possible to designate a region of interest (for example the structure S previously mentioned in the image 1b) on a first image, and an analysis system according to the invention can automatically go through a satellite image bank to extract all images therefrom having close localization parameters. A region corresponding to said region of interest may then be embodied or designated in each image.

The implementation of the invention presumes a prior step that consists of associating each image with a binary image in the form of a cloud of points.

The image subsequently makes it possible to estimate the similarity between two binary images, or even to perform a recalibration thereof.

A priori, no shared information exists regarding the content of the matching pixels in two images, in particular if they are of any different modalities. The only link between the two images is formed by the actual medium (geographic for remote sensing) of the data. According to the invention, it is not relevant to take an interest in the information contained in each pixel, but rather in the distribution of the relevant information (whatever it may be) in those images.

To associate a binary image with an image, various methods exist that are generally respectively dedicated to a particular modality. A binary image is in a way the expression of a relevant piece of information present in an image in the form of points.

A radar image is often noisy. This type of noise is generally referred to as “speckle”, or “tavelure” or “chatoiemment” in French. For example, to associate a binary image with a radar image, it is known to use the Touzi filter to segment the image. Thresholding the segmented image makes it possible to obtain a binary image having edge points of visible regions on said image.

The pixels of the multispectral image do not express a scalar value, but express a set of scalar values (as many as there are bands). Said pixels are therefore represented by a vector in a space whereof the dimension is the number of bands. Another manner of approaching this representation problem consists of considering that one has not a multispectral image, but as many scalar images (called “gray level”) as there are bands and which one wishes to reset to produce a scalar image that would be the best possible recomposition or, at the very least, that which integrates the maximum amount of information from the scalar images associated with each band. To achieve this result, it is for example known to consider a vectorial space with dimension  $\phi$  ( $\phi$  being the number of bands) of the centered values of the pixels (considered to be vectors) of the multispectral image. It is then for example possible to carry out a Principal Component Analysis of this cloud of points where each point is a vector associated with each pixel. A “Principal Direction” (associated with the largest unique value) is then obtained that corresponds to the best linear combination of the  $\phi$  initial bands. The result produces a “gray level” image that may be processed with filtering, variation detection (passage by zero of the Laplacian, for example) and thresholding operators to produce a binary image.

Other techniques may be implemented to associate an image of a given modality with a binary image.

6

Once this prior step is carried out, the invention makes it possible to assess the resemblance between two binary images in the form of clouds of points through an invariant measurement by rigid transformation and scale change. The invariance relative to this type of transformation is particularly interesting, as the invention may respond to most recalibration situations between satellite images of different modalities in particular.

The implementation of a method according to the invention, by a processing unit of a suitable analysis system, consists of seeking, a rigid transformation (translation and rotation), if one exists, that best matches two clouds of points, and assessing the level of information shared by those two clouds of points.

Irrespective of the modality of an image, the points resulting from the associated binary image are based on a physical support whereof the characteristics are linear. In the field of remote sensing, these linear characteristics for example result from riverbanks, field edges, or coasts, or road edges, etc. Locally, the structures may be well marked, or may be noisy, or may even have lacunar aspects.

As an example, FIGS. 2a to 2d illustrates typical linear structures that may be encountered in particular in the field of remote sensing.

Thus, FIG. 2a describes a linear structure Sa—substantially in the shape of an arc of circle—that is ideal or perfect: fine, without artifact or intersection with other linear structures.

FIGS. 2b to 2d illustrate more realistic situations. Thus, FIG. 2b shows a thicker structure Sb than the previous one that may be interpreted as an ambiguous localization structure.

FIG. 2c describes an unconnected structure, even though the organization of the points appears visually obvious. FIG. 2d describes a fixed structure surrounded by points resulting from noises that may be interpreted as false positives.

One preferred implementation of the invention is based on the “Hough Transform”, which consists of looking for the possible occurrences of transformations causing certain subsets of points of the first binary image to match other subsets of points of the second binary image. The Hough space is then a space with three dimensions that are the parameters of the transformation, namely the angle of rotation and the two components of the translation.

Applying the Hough transform even though a small quantity of information may be shared between two binary images makes it very sensitive to noise, and its use may quickly prove ineffective.

To avoid this drawback, the invention consists of introducing a structuring element—or knowledge element—that reflects the structure of a binary image, but in no way restricts the use of the method.

Thus, one subject-matter of the invention consists of defining a primitive making it possible to depict, in a unified and equivalent manner, all of the local distribution possibilities of points associated with linear structures as described as examples by FIGS. 2a to 2d. These primitives, once extracted, will for example make it possible to perform an automatic recalibration.

A primitive according to the invention “naturally” eliminates the aberrant points and treats the structure of the clouds of points homogeneously. Such a primitive may correspond to a local cloud of points comparable to a small curve arc characterized by its position (x and y coordinates) in the image, its orientation  $\theta$  and its relevance or score s. Such a primitive P is denoted  $P=((x,y)\theta,s)$ .

To characterize a similarity between two binary images, a method according to the invention includes a step for respectively extracting a plurality of primitives from each binary image.

To extract the primitives from a binary image, the invention provides a preferred embodiment in which a plurality of binary masks is defined in the form of a circular window C comprising a straight line segment L (or orientation line) whereof the line thickness El is predetermined and expressed in number of pixels. FIGS. 3a to 3d respectively describe such a mask. Thus, FIG. 3a describes a circular window C, with radius R including a line L. The radius R thus defines the vicinity around a referring point, center of the circular window C. The thickness El of the orientation line L corresponds to a threshold over the Euclidean distance between an element and the discrete line connecting two ends of a discrete circle. For FIGS. 3a to 3d, the thickness El is equal to 1, i.e., a central edge plus two edges on either side. The radius R is equal to 7. The total number of elements for each mask depends on the radius R and the thickness El as well as the orientation  $\theta$  of the orientation line,  $\theta \in [0^\circ, 180^\circ]$ . FIGS. 3a to 3d describe examples of masks Ma, Mb, Mc and Md, respectively, for which the radius R is equal to 7, the thickness El is equal to 1 and the orientation  $\theta$  varies by  $0^\circ$  for FIG. 3a, at  $\theta = \theta_b$  for FIG. 3b, at  $\theta = \theta_c$  for FIG. 3c and at  $\theta = \theta_d$  for FIG. 3d. Thus, the number of elements is 41 for the mask Ma described by FIGS. 3a and 39 for the mask Md described relative to FIG. 3d.

To extract the primitives from a binary image, a system according to the invention successively applies each of the n' available binary masks for each point of the cloud of the binary image. As illustrated in FIG. 4, each mask is centered on a referring point  $P_r(x,y)$  whereof the coordinates x and y determine the position within the binary image.

The point  $P_r$  belongs to a piece  $Sp$  of a linear structure. The application of the binary mask  $M_j$  consists of counting the points of  $Sp$  situated in the vicinity of  $P_r$  that intercept the elements of the mask  $M_j$  (j comprised between 1 and n'). Thus, in FIG. 4, the set  $Sp+M_j$  indicates, in black, the points of  $Sp$  that are not covered by the mask  $M_j$ . The elements of the mask not intercepted by the points of  $Sp$  are indicated by squares with a gray texture. The points of  $Sp$  that intercept the elements of  $M_j$  are shown by crosshatched squares. It is the latter that are counted by the processing unit of a system according to the invention to determine whether the orientation line of the mask is relevant in light of the linear structure to which a referring point belongs. The mask that counts the most elements intercepted by the structure determines the best orientation line. The processing unit thus extracts a primitive  $P = ((x,y), \theta, s)$  associated with the referring point  $P_r$  determined by the coordinates (x,y) of said referring point, the orientation  $\theta$  of the orientation line of the mask and the number s of intercepted elements. The best orientation line associated with a referring point of a linear structure substantially constitutes the equivalent of the tangent at that point. To further improve the precision of the selected orientation  $\theta$ , the invention provides that it is for example possible to perform a Principal Component Analysis applied on the pixels having been intercepted by the best mask. The primitive is then in turn characterized by said discrete orientation thus obtained. The orientations  $\theta$  can then assume a number n of discrete values greater than the number n' of available binary masks.

A system according to the invention thus determines, for any point of the binary image, the primitive having collected the best score s.

To be able to optimize the calculation time of the processing unit of such a system, the invention provides a preferred embodiment of the step for extracting primitives. This

embodiment is described relative to FIG. 5. The counting of the respective elements of the set of binary masks may only require a single passage over the entire circular window. In fact, the counting of a mask may be deduced from that of the previous mask, iteratively, by only testing the added elements and the deleted elements to go from one mask to the next. Thus, according to FIG. 5, the elements of a mask  $M_j$  are shown using crosshatched squares. The elements of the following mask  $M_{j+1}$  are shown by squares with vertical lines. The view at the bottom of FIG. 5 describes the elements shared by the two masks using gray squares, the elements  $\Delta+$  added by the mask  $M_{j+1}$  using squares with vertical lines and the deleted elements  $\Delta-$  using crosshatched squares. According to this embodiment, the score of the mask  $M_{j+1}$  is then defined by the score of the mask  $M_j$  to which said elements  $\Delta+$  are added that are intercepted by the linear structure covered by the window and from which the elements  $\Delta-$  also intercepted are deleted. The gain in complexity is thus estimated at approximately 60%.

The invention also provides that only the primitives having a score above or equal to a threshold  $S_{min}$  can be kept and stored. This alternative makes it possible to avoid having to consider isolated points or insignificant areas of the binary image.

The invention also provides that a system according to the invention can account for a parameter corresponding to a minimum distance  $D_{min}$  between two extracted primitives. In other words, only the primitives having obtained the best score s from among those associated with referring points situated at a distance at least equal to  $D_{min}$  can be kept and stored by such a system. This minimum distance is a parameter of the process for extracting the primitives from a binary image according to the invention, for a given robustness and a given noise level.

Whatever the favored embodiment, the invention translates a cloud of points into a set of uniformly distributed primitives characterizing the presence and local orientation of underlying linear structures.

A system for estimating the similarity using the method according to the invention includes means for storing at least two sets of primitives respectively relative to two binary images.

Consider two binary images A and B and the sets  $Qa$  and  $Qb$  of primitives respectively extracted from said binary images. A single rigid transformation (translation, then rotation) associating  $Qa$  with  $Qb$  exists. Let  $t_x$ ,  $t_y$  and  $\alpha$  be the parameters making it possible to go from  $Qa$  to  $Qb$ . In fact, said transformation lies in the composition of a rotation by an angle  $\alpha$  and a translation in the plane  $(t_x, t_y)$ . According to the principle of the Hough transform, the sets of parameters (each set corresponding to an occurrence of the rigid transformation) accumulate in a dimension 3 space that is the Hough space.

The equations connecting the respective coordinates  $(z_a$  and  $z_b)$  and orientations  $(\theta_a$  and  $\theta_b)$  of two primitives coming from two binary images A and B to the parameters of the rigid transformation are expressed in complex notation by the relationship:

$$Z_a = e^{i\theta_a} \cdot z_b + t \text{ where the complex number } t = t_x + i \cdot t_y \text{ represents the translation in the plane and } \alpha = \theta_a - \theta_b.$$

The implementation of the invention now consists not of seeking the value  $\alpha$  that would correspond to the angle of rotation of the transformation, but of seeking whether, considering two regions (one  $A_a'$  in an image A and the other  $A_b'$  in the other image B as described as an example in FIG. 6), and taking all pairs of primitives of those two images, a signal (a particular value of  $\alpha$ ) does or does not emerge. This pro-

cessing may alternatively be done by the processing unit of a system according to the invention on all of the binary images.

If a signal emerges, this means that the implementation of a method according to the invention has detected the existence of a subset of primitives of the first window that is found in the second window to within a rotation by an angle  $\alpha$ .

The emergence of the signal that will characterize the presence of shared information between two windows concretely corresponds to the presence of a maximum marked in the distribution of the different orientations.

According to the invention, the parameters favoring the emergence of this signal, or on the contrary those which favor the presence of false positives, which correspond to maximums unfavorably formed by matches of irrelevant primitives, are primarily the number of primitives in each region (or binary image), the number of discrete orientations that can be attributed to some primitives, and the contamination rate, which corresponds to the ratio of the number of primitives with no match out of the number of primitives in each region.

Thus, when the number of primitives  $k$  increases, the emergence of the relevant signal that characterizes shared information is favored. The same is true when the number  $n$  of discrete orientations that can be attributed to the primitives increases.

On the other hand, when the proportion  $\epsilon$  of primitives between two regions or images with no match, also called "contamination rate", increases, the emergence of false positives is favored.

To guarantee the robustness of the characterization of shared information between two regions, it suffices to invert a risk equation  $r(\epsilon, k, n)$  such that:

$$r(\epsilon, k, n) = e^{(1-\epsilon)k + \frac{k^2}{n}} \log \left( \frac{k^2}{n(1-\epsilon)k + \frac{k^2}{n}} \right)^{(1-\epsilon)k}$$

Thus, by positing  $r(\epsilon_0, k, n) = p_0$ , where  $\epsilon_0$  corresponds to an arbitrarily large contamination rate (management of unfavorable situations) and where  $p_0$  corresponds to an arbitrarily low risk of appearance of false positives (robustness), it is possible to identify the appropriate number  $n$  of discrete orientations based on the number of primitives  $k$ . Thus, it is possible to parameterize the analysis system according to the invention to monitor the precision of the orientations of primitives and therefore the reliability of the characterization of shared information between two regions of two binary images.

To describe one embodiment of the step of a method for measuring the presence of shared information within binary images by counting occurrences of differences of orientations of pairs of primitives respectively extracted from the two binary images, let us consider two regions of images  $A_a^i$  and  $A_b^j$  as presented by FIG. 6.

This figure illustrates two regions and  $A_a^j$  and  $A_b^j$  for which two sets of primitives  $Q_a$  and  $Q_b$  have been extracted. This illustration also illustrates the existence of two subsets of primitives  $Q_{a1}$  and  $Q_{b1}$  that appear to match and whereof the orientations are shifted by a value  $\delta$ , i.e., the discretized value of  $\alpha$  relative to the degree of precision of the orientations of primitives, i.e., the discrete value closest to the angle of rotation  $\alpha$  that exists between the two binary images.

FIG. 6 describes the translation of two clouds of points that partially match through a rigid transformation. This transformation can also consist of a change of scale (scale factor), since the latter does not influence the difference of the angles.

To characterize the emergence of the sought signal, a method according to the invention integrates the following constraints:

the possibility that a "conspiracy" of the noise exists to form local maximums that would hinder the identification of a relevant maximum;

the possibility that a certain variability exists in the differences of orientation of matching primitives, and that the sought signal is not perfectly localized.

Said FIG. 7 makes it possible to describe one example of the formation of occurrences  $\eta$  in the Hough space making it possible to characterize information shared between two regions. In theory, the distribution of the occurrences for any  $\theta$  should have a quasi-uniform distribution  $d$ , of the Dirac type clearly identifying the signal  $\delta$ . The reality generally produces an experimental distribution  $d_e$  that may give rise to the emergence of local maximums  $H'$  and  $H''$  characterizing false positives that may hinder the identification of the relevant signal.

To take the local maximums into account, the invention provides—according to one preferred embodiment—that the system can measure a relative deviation  $\epsilon^*$  between the overall maximum  $H$ , and the average of two other best local maximums, denoted  $H'$  and  $H''$  as presented as an example in FIG. 7. Thus, if the best maximum clearly stands out from the other maximums, not only does the invention make it possible to characterize shared information between two regions, but additionally it eliminates some ambiguity as to the relevance of that characterization.

Within the meaning of the invention,

$$\epsilon^* = \frac{|H - \tilde{H}|}{2} \text{ where } \tilde{H} = \frac{H' + H''}{2}$$

The closer the value of  $\epsilon^*$  is to 1, the more the presence of shared information between the two regions is proven.

To be able to take into account the second aforementioned constraint related to the imperfect localization of the sought signal, the invention provides a preferred mode for calculating  $\epsilon^*$ . According to this alternative, the signal is recomposed and distributed over several adjacent cells of the histogram to obtain the value  $H$ . FIG. 7 makes it possible to illustrate the notion of "mode"  $\Omega$ , which consists of a grouping of five adjacent cells  $c1$  to  $c5$ , the central cell  $c3$  of which has a higher score than the other four cells.

The method according to the invention then consists of detecting all of the modes and ranking the latter by decreasing order out of the total score of the five cells making them up. The best mode (or that which obtains the best score) is linked to the relevant sought signal. To improve this discrimination, the invention provides that the processing unit of a system according to the invention calculates said score by adding the respective scores of the central cell with its left or right neighbor (the crosshatched cells  $c3$ ,  $c2$  or  $c4$  according to FIG. 7). The score of the best mode thus developed is assigned the size  $H$ . The scores of the central cells of the other two best modes make it possible to obtain the values of  $H'$  and  $H''$ .

The invention also provides, to guarantee the regularity of the histogram of the Hough space and thus facilitate the extraction of the modes, for optionally applying a low-pass filter to eliminate the unwanted artifacts while preserving the appearance of the histogram. Preferably, said filter is a median filter. Other filters may alternatively be used.

FIG. 8 makes it possible to synthesize the main steps of a method 300 for characterizing a similarity between two

11

binary images. Such a method **300** is carried out by the processing unit of an image analysis system according to the invention and adapted accordingly.

Such a method includes a first step **301** for extracting at least two sets of primitives  $Q_a$  and  $Q_b$  respectively associated with two binary images A and B associated with images of any modalities. This step results from the implementation of at least two instances **100a** and **100b** of a method **100** for extracting a set of primitives from a binary image as previously studied and which is further the subject-matter of a summary description below. The method **300** includes a step **302** for measuring the presence of shared information within the binary images A and B by counting the occurrences of differences of orientation of pairs of primitives respectively extracted from the two binary images. This step results from the iterative implementation of at least one instance **200i** of a method **200** for measuring the presence of shared information within two binary images as previously studied and further being the subject-matter of a summary description below. Having one or more measurements of similarities between two images or regions resulting therefrom, the method **300** can include a step **303** aiming to exploit said at least one measurement. As an example, the step **303** may consist of recalibrating two images of possibly different modalities from which the binary images result and the similarity of which has been measured. This step may also consist of displaying said recalibrated images using a suitable man-machine interface of a system according to the invention. Alternatively, the step **303** may consist of identifying a landmark or a remarkable and precisely localized element in both images. Reciprocally, the “non-detection” of shared information may also be very important in certain applications: presence of a camp, or movable structures, etc. The similarity measurement **302** is then used to characterize **303** the absence of a sought structure.

To extract—**100**—the primitives from a binary image, a method according to the invention includes a first step **103** for obtaining a binary image in the form of a cloud of points. This binary image is stored beforehand by storage means of the system according to the invention, said storage means cooperating with the processing unit of said system implementing the method **300**. The method **100** includes a step **104** consisting of applying a plurality of binary masks to each pixel making up the binary image. As previously studied, said binary masks may each consist of an orientation line whereof the thickness and orientation are predetermined and describing the diameter of a circular window whereof the radius is also predetermined. To extract the primitives, the method **100** includes a step **105** to:

count, for each binary mask applied to a pixel, called referring pixel, the number of neighboring pixels of said referring pixels that are covered by said binary mask;  
define, as primitive of said referring pixel, a triplet comprising the position of the referring pixel in the binary image, the largest number of neighboring pixels covered by one of the applied binary masks, and the orientation of the orientation line of said binary mask.

Optionally and preferably, the method **100** includes a step **106** for refining the list of primitives extracted from a binary image, that list being obtained and stored by the storage means in **105**.

Step **106** may consist of only keeping in memory the primitives whereof the number of neighboring pixels of a referring pixel and covered by a binary mask is greater than or equal to a predetermined minimum number  $s_{min}$ .

12

This step may alternatively or additionally consist of only keeping the primitives of referring pixels separated by a minimum and predetermined number  $D_{min}$  of pixels.

Furthermore, the method **100** for extracting the primitives from a binary image may include a prior step **102** aiming to associate (therefore develop through the processing unit of the system according to the invention) said binary image with an image irrespective of its modality. This image may be acquired or selected from among several, for the implementation of capture means or a man-machine interface suitable for the system according to the invention.

The method **200** for measuring the presence of shared information within the two binary images in turn consists of a first step **201** in which the processing unit considers two regions respectively within the two binary images. This step may consist of an automatic and autonomous selection done by the processing unit or may result from the assistance of an operator using a man-machine interface suitable for the system. Alternatively, a region may encompass all of a binary image. Considering a region of the binary image amounts to selecting a subset of primitives. This selection is possible because a primitive—within the meaning of the invention—includes the coordinates of the referring point associated with said primitive.

The method **200** includes a step **202** in which the processing unit obtains the difference of the orientations of all of the pairs of primitives (one primitive being taken from the first region and the other from the second region).

In **203**, the processing unit counts the occurrences of said differences. As an example, the step **203** may consist of filling out a one-dimensional table stored by the system. A distribution of the occurrences of the differences of orientation of the selected primitives is thus obtained as illustrated by FIG. 7, as an example.

A method **200** according to the invention further includes a step **204** in which the processing unit—for example by going through such a one-dimensional table—determines, according to one preferred embodiment, the three best modes H, H' and H'' resulting from the distribution of the counted differences of orientation.

Lastly, in **205**, the similarity measurement is developed by the processing unit by calculating and storing the similarity indicator  $\epsilon^*$  according to the techniques previously studied, for example

$$\epsilon^* = \frac{|H - \tilde{H}|}{2} \text{ where } \tilde{H} = \frac{H' + H''}{2}.$$

The invention has been described through one preferred example embodiment (remote sensing) relative to the processing of images of different modalities of the radar or optical (multispectral) types. The invention may also be implemented to match other types of images: for example, satellite images for which one wishes to know the level of resemblance with airborne images of the “aerial photography” type.

The invention also applies to other fields including medical imaging, where the matching for example of images from scanners and MRI corresponding to different acquisitions in terms of densities for the former and hydrogen atom levels for the latter. It is also possible to implement the invention to measure the similarity between images of modalities other than ultrasonography (ultrasound imaging) or functional imaging.

13

Industrial vision and video protection are also fields for which the recalibration of multimodal images is sensitive. For example, quality control banks use several acquisition modalities to find defects. The invention is naturally relevantly applicable in these fields.

Irrespective of the multiple fields of application, it should be noted that the invention further produces excellent results in the particular case where the modalities of the images are identical. These are particularly relevant in the context of situations considered to be difficult such as, for example, when the lighting conditions of a scene evolve greatly from one shot to the next, for example in remote sensing or industrial vision.

The invention has been described through example applications for which a single binary image is associated with an image of a given modality. Said binary image is thus assumed to be the "best representative" of said image: a binary image resulting from a multispectral image may be obtained from the best recombination of the bands  $\phi$  thereof using a Principal Component Analysis method (i.e., recombination associated with the Principal Direction). However, there may be interesting information to be exploited on the subsequent "Directions". In that case, it is no longer a single binary image that is produced, but multiple binary images. This is even more relevant if one uses "Independent Component Analysis" instead of using "Principal Component Analysis". This technique in fact causes the emergence not of a "recomposition" that best synthesizes the information contained in the different bands, but of several compositions, all of great importance and capable of revealing several aspects of overall information. Two images may thus respectively be associated with one or more binary images. The invention makes it possible to find information shared by the two initial images, by iteratively measuring the similarity of pairs of binary images formed from one of the binary images associated with the first image and one of the binary images associated with the second image. At the end, the invention makes it possible to determine the match maximizing the similarity measurement, for example.

Furthermore, the invention has preferably been described relative to a rigid transformation optionally made up of a scale factor (change of scale). The invention also applies to projective transformations (perspective, etc.) and complex transformations, i.e., continuously evolving transformations. This is in particular the case for airborne images (taken using airplanes, drones or airships) for which the relief (and therefore the variability of the altitude) plays a significant role in the recalibration transformation, this being different at various points of a same image. Nevertheless, this complex transformation may be locally approximated by a more elementary transformation. Thus, the invention makes it possible to associate enough landmarks to characterize not an overall transformation, but several local transformations that may then be "interpolated".

The invention claimed is:

1. A method for indicating a similarity between two binary images  $A_a^i, A_b^j$ , said method being carried out by a processing unit of a system for analyzing similarities of images further including a storage device cooperating with said processing unit to store binary images, wherein the method comprises:

a step for extracting a first set of primitives from a first binary image  $A_a^i$  and a second set of primitives from a second binary image  $A_b^j$ , the first and second binary images being stored beforehand by the system, each primitive being associated with a pixel of a binary image

14

and being characterized by the position of said pixel and by the orientation of an orientation line passing through the pixel;

a step for measuring the presence of a common piece of information within said binary images and delivering a similarity indicator  $\epsilon^*$  therefor, said indicator being calculated by the processing unit as the relative deviation between an overall maximum H and the average of two other local maximums H' and H'', i.e.,

$$\epsilon^* = \frac{|H - \bar{H}|}{2} \text{ where } \bar{H} = \frac{H' + H''}{2},$$

said maximums being determined from a distribution of the different orientations of pairs of primitives respectively resulting from two extracted sets of primitives.

2. The method according to claim 1, wherein the extraction step comprises:

obtaining a binary image  $A_a^i, A_b^j$  stored by the system; applying a plurality of binary masks to each pixel of said binary image  $A_a^i, A_b^j$ , the binary masks each consisting of including an orientation line whereof the thickness and orientation are predetermined and which describes a diameter of a circular window whereof the radius is also predetermined;

counting, for each binary mask applied to a pixel—called a referring pixel—the number of pixels adjacent to said referring pixel that are covered by said binary mask;

defining, then storing as a primitive associated with a referring pixel, a triplet  $P((x,y)\theta,s)$  comprising the position  $(x,y)$  of the referring pixel in the binary image; the largest number s, called the score, of adjacent pixels covered by one of the binary masks applied to the referring pixel; as well as the orientation  $\theta$  of the orientation line of said binary mask.

3. The method according to claim 2, wherein a primitive is only stored if its score is greater than or equal to a predetermined minimum number.

4. The method according to claim 2, wherein only the primitives whereof the referring pixels are distant by a minimum and predetermined number of pixels are stored.

5. The method according to claim 1, wherein the measurement step carried out by the processing unit comprises:

determining a region of predetermined size in the first and second binary images to thus form a pair of regions  $A_a^i, A_b^j$ ;

counting and storing the occurrences of the different orientations for all pairs of primitives respectively extracted from said pair of regions;

calculating the similarity indicator  $\epsilon^*$  between the two regions of the pair.

6. The method according to claim 5, further comprising a prior step for associating and storing said first and second binary images, respectively, with two images stored or acquired by the image similarity analysis system.

7. The method according to claim 6, further comprising a step for recalibrating two stored or acquired images or for identifying a remarkable element located in the two images from the calculated similarity indicator  $\epsilon^*$ .

8. The method according to claim 7, further comprising a step for displaying said recalibrated images and/or the remarkable element using a man-machine interface suitable for the image similarity analysis system.

9. A system for analyzing the similarity of binary images including a processing unit cooperating with a storage device,

wherein said processing unit implements a method for indicating a similarity between two binary images according to claim 1.

10. A system for recalibrating images acquired or stored by a device for storing and/or acquiring an image, said device cooperating with a processing unit, which in turn further cooperates with a storage device, wherein said processing unit implements a method according to claim 6.

11. The system according to claim 10, wherein said processing unit delivers two recalibrated images or a remarkable element identified and localized in the two images to a man-machine interface.

12. The system according to claim 9, further including a man-machine interface for delivering the similarity indicator  $\epsilon^*$ .

13. The system according to claim 9, further including a man-machine interface for selecting one image from among several and/or a particular region thereof.

14. The system according to claim 10, wherein the acquisition device acquires images of different modalities.

\* \* \* \* \*



# **C** Embedding image processing intelligent algorithms

## EMBEDDING INTELLIGENT IMAGE PROCESSING ALGORITHMS: THE NEW SAFETY ENHANCER FOR HELICOPTERS MISSIONS

Pierre Zoppitelli, Pierre.Zoppitelli@Airbus.com, Airbus

Sébastien Mavromatis, Sebastien.Mavromatis@univ-amu.fr, Aix Marseille Université

Jean Sequeira, Jean.Sequeira@univ-amu.fr, Aix Marseille Université

Guillaume.Anoufa, Guillaume.Anoufa@Sogeti.com, Sogeti HighTech

Nicolas Belanger, Nicolas.Belanger@Airbus.com, Airbus

François-Xavier, Francois-Xavier.Filias@Airbus.com, Airbus

### Abstract

Over the last two decades, image processing technologies rapidly emerged from the shadows to become one of the most important field of interest in computer science. Although image analysis is a hot topic in the automobile industry and for some aircraft applications (drones, airplane, space probes...), the certification of any vision based autopilot system for helicopter missions remains an ongoing challenge. Indeed, such a system would be required to perform complex missions with a high success rate while possibly facing adverse weather conditions. However the rapid increase of processing power, the development of image analysis algorithms, as well as the miniaturization of high resolution cameras, are allowing new technical solutions for autonomous flight. Facing this new technological deal, helicopters manufacturers can no longer ignore that vision based systems are about to become a key enhancer for versatile rotorcraft missions.

Airbus Helicopters is committed to put the safety of its aircrafts at the highest standards. For this purpose, Airbus Helicopters has initiated the development of advanced systems integrating many disciplines like sensor acquisition, scene understanding, situation awareness, and artificial intelligence. As a contribution to this company objective, the EAGLE project (Eye for Autonomous Guidance and Landing Extension) was launched 2 years ago to develop a generic optronic platform facilitating integrations of algorithms for different applications. The system aims to improve safety and reduce the pilots' workload during flights in oil and gas and SAR missions. This paper presents the latest results that have been obtained by Airbus Helicopters and the LIS-lab in the development of a landing platform detector in the frame of this project. We will first introduce the general methodology applied for the determination of the platform position. The approach is hierarchical and based on a collection of hints to determine, refine and validate suitable locations for the presence of a helipad. We will then present the strategy for the selection of regions of interest. The aim is both to determine the right size of portion of the image to be analyzed, and to enable the real-time adaptation of the selection and sequencing of the regions to be explored.

This article will then detail the methods used to determine the areas susceptible to contain a landing platform. The algorithm mainly relies on flat ellipse detection as it is the most visible feature of a helipad seen from long distances. An adaption of the Hough transform proved to be the most reliable method in the specific case of very flat ellipses. A validation step using many other properties and visual clues performs the verification of the presence of the helicopter landing platform in the research areas delimited by the obtained ellipses. Having presented the algorithm for the detection of a helicopter landing platform, we will discuss some approaches to increase the system's accuracy, robustness and integrity (detection of failure). In particular, safety and certification considerations are used to select the human-machine interfaces and overall design of the system. A result section will show how this system demonstrated to be capable of detecting landing platforms from a distance of 1500 meters and tracking it without interruption until the landing phase. Last but not least, this paper will introduce an open view of identified image processing technologies continued path for the upcoming years. Our vision of this technological field as a mandatory brand new core competency to be strengthened within Airbus Helicopters, and the way we intend to build up the necessary ecosystem with Airbus' other business units, will be the epilogue of the article.

## 1. OVERVIEW

### 1.1. Introduction

As drones' autopilots are becoming more and more complex, the field of possible missions is getting larger and larger: drones may now deliver packages, they monitor large portions of territorial waters, also spread fertilizers over crops and do many other different and difficult tasks. It would be natural to expect that helicopters would follow the same path towards completely autonomous missions. However, many challenges proper to the helicopter Industry needs be solved: some missions tend to be extremely demanding (search and rescue, transportation to offshore platform) and safety and certifiability concerns are of the utmost importance.

To achieve this goal, a complex system of sensors is required. Indeed, a drone has to be able to assess a situation before making any decision. Among the possible sensors we may cite: day and night cameras, infrared sensors, LIDAR, RADAR, and GPS. There is also the solution to have the sensors in a ground station instead of on-board the aircraft. In the frame of the EAGLE project the solution chosen was to have a set of cameras mounted on the helicopter. It has the advantage to be cheaper than LIDAR and RADAR solutions while not requiring the installation of specific systems on a ground station. GPS cannot be used alone as it requires the communication with other systems and lacks the precision to handle most missions.



Fig 1: The EAGLE system.

### 1.2. General methodology

Our algorithm is based on human perception. It is divided into many processes with limited scope. Each process has a single task with one clearly defined objective. At each step of the algorithm we look for visual properties verified by the desired object in order to remove all candidates who do not have those properties. We continue this process until only the desired objects remain. It is very important that none of the tasks remove a correct candidate. This approach has several advantages: it is always possible to know at which point the algorithm failed, it is easy to improve (we know the expected results at each step), we need limited data to start working (we may add some steps to tackle new difficult problems) and it is easier to prove that the algorithm has found a correct solution.

A landing platform which follows the norm described in the CAA (Civil Aviation Authority) CAP437 must have a thick yellow circle painted on a green platform; moreover, a white "H" is painted at the centre of the circle. The authorized lengths of each one of these elements are described in this document. We can assert that there are a few important visual properties: the circle which will be seen as an ellipse from afar, the presence of the H at the centre of the circle, the homogenous regions of the same colour. We can also expect that the landing platform will be on an offshore rig which will be heavily contrasted compared to the uniform sea landscape.

It would be too time consuming to search for all visual clues in the whole image; instead, we focus on the most dominant visual clues which allow us to identify interesting regions or portions of the image to which we apply more compute-intensive algorithms.

The proposed methodology for the detection of a landing platform in sea background includes the following steps: first, we locate potential oil platforms, then in these areas we will search for ellipses to define very precise regions of interest and eventually we validate the presence of the helipad with all visual clues available (homogenous colours, presence of the H, etc...). This approach mimics how the human vision operates; first, we search for something that looks like the desired object, then we focus our attention into it, and finally we determine if it is indeed the sought object.

## 2. THE DETECTION OF THE LANDING PLATFORM

### 2.1. Related work

Several papers about helipad detection techniques have been published for the autonomous landing of drones [12-15]. However, none of these addresses the specificities of the landing of helicopters. In particular, many common hypotheses made by the authors do not hold for the application domain we are targeting:

- They assume the helipad is viewed directly from above and not masked by other structures
- They do not take into consideration safety and certification requirements of helicopter systems
- They assume the videos are of good quality and taken in favourable lighting conditions.

As all these limitations cannot be disregarded for our application, we need to look for new solutions. These innovative methods should be able to detect helipad in difficult conditions such as grazing angle views, very far helipad and low quality images.

Since a helipad consists of an H within a circle and circles are transformed into ellipses by perspective projection, the largest visual clue of the presence of the helipad will be a flat ellipse. We chose to decompose our landing platform detector algorithm into two steps: first we look for potential helipad candidates by searching ellipses in the image; then, we verify for each candidate ellipse if there is a helipad at this location.

To detect potential helipad candidates, several techniques, based on ellipse detection, are available. They can be classified into several categories:

- Curve fitting techniques
- Arcs grouping methods
- Active contour techniques
- Hough transforms approaches

Let us detail briefly the advantages and drawbacks of each category of solutions.

Curve fitting methods are very efficient, but they fail when there is a large number of points of the processed image which do not belong to the ellipse. The most used technique is based on the least square optimization technique [2]. Several advances have increased the resistance to outliers; however, the limitation still exists.

The adjacent arcs and lines grouping method consists of linking edges to produce arcs and then merging arcs belonging to the same ellipse. It was first described by Kim et al. [11]. Jia et al. [1] improved this method by using a projective invariant that allows them to quickly assess if two arcs have to be merged. These techniques produce good results even for very complex images with a multitude of ellipses. However they fail when the ellipses are heavily occluded, dashed or extremely flat.

Active contours techniques let evolve an elliptic curve under some constraints. One major problem of these methods is their inability to deal with complex images: the contours may converge to suboptimal solutions.

The methods relying on the Hough transform ([5], [10]) map the input image into a parametric space and then take the peaks in this space as candidates for the ellipse. The principle of the Hough transform is to explore all minimal combination of points that produce a parametric model and count how often each model appears. For example, when used to detect ellipses, five points are required to determine a unique ellipse and by exploring all combination of five points we may see which ellipses are recurrent and thus which ellipses are in the image. In practice, a five point Hough Transform is very inefficient and solutions had to be found to reduce the complexity of the approach ([10], [6]).

Hough transforms are very effective even on the most difficult conditions. They can produce good results with partial, flat or dashed ellipses as it was proven in [16]. On the other hand, they are compute-intensive and difficult to parametrize even though several advances have reduced these shortcomings.

To solve our problem, we chose the Hough transform approach, as it is the only one robust enough to address the specificities of the helicopter domain. Furthermore, several techniques are available to reduce processing time:

- Reduction of algorithm complexity by separating the computation of the parameters [6],
- Hardware specific parallelization techniques (GPU/FPGA) ([3], [4]),
- Data reduction by using a multi-resolution or sub sampling approach ([9], [16]).

The following sections will describe the proposed helipad detector algorithm. They will comprise: the methodology adopted for the selection of the portion of the image to be analysed, the detection

of potential helipad candidates through the research of ellipses, and, the checking of the presence of the helipad at the location of each candidates.

## 2.2. Selection of the regions of interest

In this section, we present the method to select the right portion of the image to be more deeply analyzed. We first describe an algorithm to extract the oil & gas platform from the image, and then we propose a method to select the region of interest based on many factors like computation time and previous results of the detection algorithm.

To reduce the portion of the image to be analyzed we remove the sea background in the image. The segmentation procedure is based on texture analysis and on two properties of the sea texture in the image. The sea is preponderant in the image and it is lightly contrasted compared to the sought entities (oil & gas platforms). Haralick proved that the texture of an image could be described by the statistical relationships between neighbor pixels [8]. He proposed a set of features to extract some visual properties from those statistical relationships; those features include contrast, homogeneity, variance, correlation and many others. Our algorithm uses the contrast texture descriptor and an automatic segmentation method to precisely locate the offshore platforms (see Fig 2).



Fig 2: The detected platforms using Haralick features.

However, once the helicopter is getting close to the oil & gas rigs, the platforms can take a huge portion of the image which would make the exhaustive research of landing platforms difficult. Instead of spending a lot of time and resources to produce a solution capable of detecting all kind of platforms in all kind of images, we developed an

algorithm which only detects the landing platform on very specific conditions; and we adapt (by cutting or resizing) the input image to fit those criteria. Basically, we transform a difficult problem into a set of smaller and easier problems. More precisely, we have an algorithm that can only detect a landing platform whose size is comprised between one half and one sixth of an image with a fixed width and height; we decompose the portion of the image encompassing the rig in a multitude of sub-images which will be then resized to fit the above criteria. By using a multi-resolution approach, the algorithm is able to find both small and big landing platforms in the image.

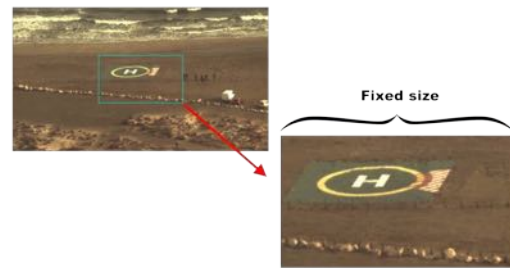


Fig 3 : The resizing of the sub-image to fit the criteria.

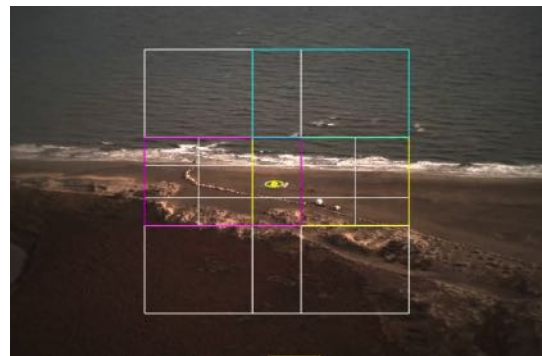


Fig 4: The multi-resolution approach. The input image is divided into many overlapping sub-images.

The multi-resolution approach provides several advantages. We do not have to adjust the parameters to handle specific edge cases: every problem comes down to solving the same simple problem. Another desirable feature is the bounded computation time; this is due to the fixed size of sub-images to be examined. Finally, one major advantage of the solution is to allow the parallelization of those processes: the computation of one sub-image is independent to the one of another which massively reduces the overall computation time on parallel architecture.

Even though the computation time has been significantly reduced thanks to the previously described techniques, we still cannot process all the sub-images at each frame. We therefore define a priority list which organizes the list of sub-images based on previous results of the detection algorithm. For example, if a sub-image produces a positive result it is sent to the top of the priority list and the detection will be done at each frame. On the other hand, if we are certain that there isn't anything interesting in the sub-image, it will be sent to the bottom of the priority list, and the computations on this area will be done at a lower frequency than the video frame rate.

### 2.3. The selection of potential helipad through the research of ellipses

In this section, the techniques used to locate the region of interests likely to contain a landing platform will be described. First, will be detailed the techniques used to extract the pixels likely to belong to the circle of a helipad. Then the Hough transform and the techniques used to obtain the candidate ellipse centre will be presented. Lastly, the methodology for the reconstruction of the ellipses will be described.

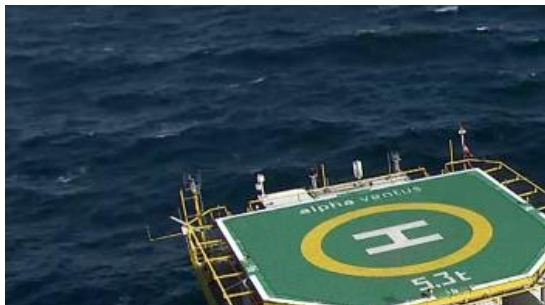


Fig 5: Input sub image used for the testing.

Several image pre-processing techniques are used before applying the Hough transform to improve the signal to noise ratio. They are all based on visual properties of the helipad. These processes include:

- A high-pass filter to extract the edges of the image as we look for the contours of the ellipse,
- A texture-based filter, since the helipad is homogenous around the circle,
- A colour separation scheme, as all points of the circle of the helipad have the same

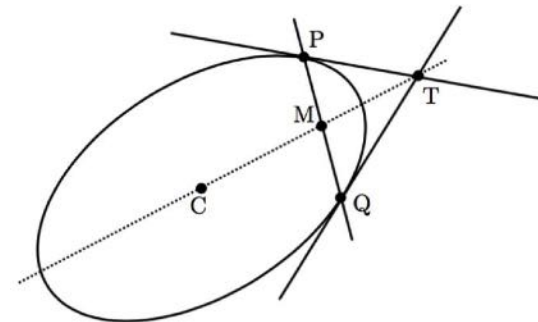
colour,

- A speckle noise filter to remove sparse points.



Fig 6: The binary image.

For each Hough transform, we need a function that maps a set of points to a model. Instead of using five points to compute all five parameters of the ellipse at once (which is costly), we chose to use only three points and compute only the centre of the ellipse. Let us describe the property of tangent lines of conics that we employ to compute the ellipse centre.



7: A property of ellipses. C, M and T are on the same line.

If P and Q are two points of an ellipse with M the middle of [PQ] and T the intersection of the tangents passing through P and Q, then the centre C of the ellipse is on the line passing through M and T.

We may combine this property with a Hough transform to compute the ellipse centres of the image. We can improve furthermore this method to increase the signal to noise ratio. When we apply this property to two points, we assume that they belong to the same ellipse which is most likely not the case. In order to check that the computed centres are valid, we may use three

points instead of two and verify that the three lines intersect close to a unique point (the centre).

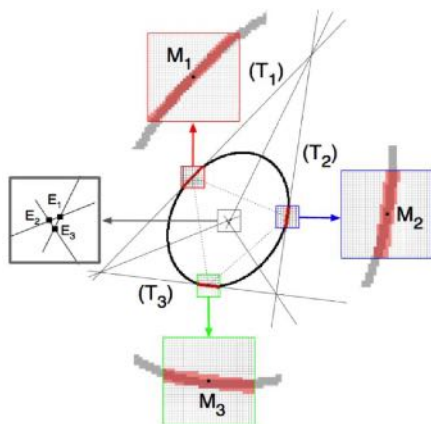


Fig 8: When three points belong to a same ellipse, the three lines form a very small triangle close to the centre.

We may further refine the centres by looking at the orientation of the arcs around the tangent. If three arcs belong to the same ellipse, each one of them should face the computed centre.

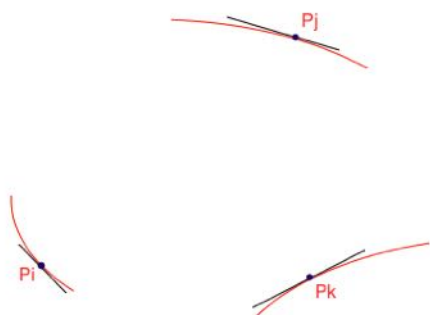


Fig 9: Example of badly oriented arcs.

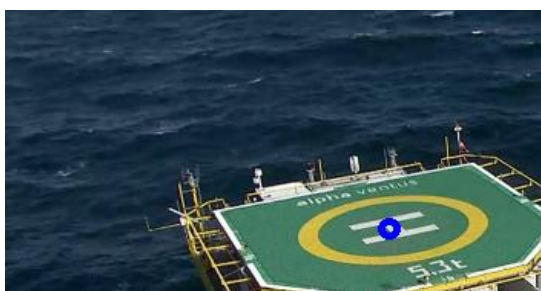


Fig 10: Computed centre using the Hough transform.

Once the ellipse centres are detected, we need to find the remaining parameters which are the two axes and the angle. To produce the desired ellipses, we group points that voted together for

the centre while also grouping those that are close spatially. The main idea is that if almost every triplet of points from one subset produced the same centre, then they are likely to be part of a unique ellipse. A method of ellipse fitting is applied to each subsets of points to produce the same number of ellipses. We then merge the ellipses that have similar parameters.

The selection of the best candidate ellipse is based on two criteria: the completeness of the ellipse and the resemblance to an ellipse. To evaluate the first criterion, we compute the expected number of points of the ellipse (if it were complete) and we compare it to the total number of points used in the ellipse fitting process. The second descriptor is computed by evaluating the dispersion of the pixels around the ellipse.

After the best ellipses are selected, we need to verify if they belong to a helipad. These validation steps are detailed in the next section.



Fig 11: The computed best ellipse.

## 2.4. The validation of the candidate helipads

In the previous section, we presented the methods used to obtain a set of regions of interest through the research of ellipses in the image. Since the portion of the image to be analyzed is heavily reduced after this first phase of image processing, we may perform some compute-intensive processing. The clues used to perform the validation of the landing platform are:

- Detection of the inner and outer contours of the circle of the helipad
- Verification of the coherence of the colors of each region of the helipad
- Detection of the H inside the circle

The first validation step of the helipad is the detection of the elliptic ring of the helipad. From the areas of the ellipses previously detected, we can define an enlarged zone in which we can look

for the inner and outer contours of the helipad elliptic ring. The collected points of these boundaries can then be processed with curve-fitting algorithms based on the least-square method. We may then check that the detected ellipses correspond to those of a helipad: they should have the same center and orientation.

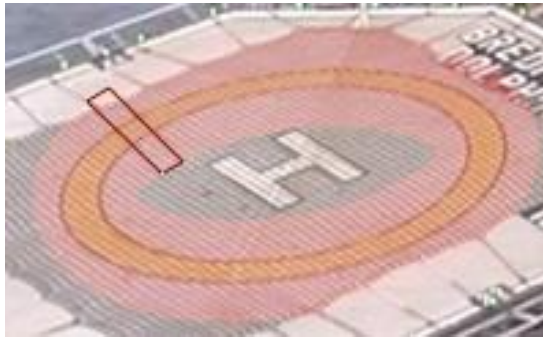


Fig 12: The enlarged research zone in light red. And a specific research zone in orange.



Fig 13: The inner and outer ellipses of the helipad elliptic ring.

In the following validation step, we check the coherence of the colors of the different areas of the helipad. Thanks to the previously identified contours, we can define 3 different zones: the inner part of the circle, the outer part and the circle itself.

We then verify that each of these areas has homogeneous color distribution and that the color inside the circle is different from those of the other areas.

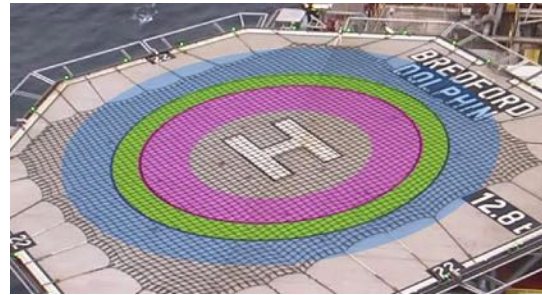


Fig 14: The three regions that need to be examined.

The last step of our validation process is the detection of the "H" in the inner area of the circle. However, depending on the distance of the camera to the helipad, the "H" may be more or less visible and sometimes barely detectable. In order to make the best use of available information and get the most reliable method of helipad detection, we chose to have two different methods for the validation of the presence of the "H". Each one tailored to the size of the helipad in the processed image.

When the « H » is barely visible, we rely on an implicit detection method based on a skeletonization process and the usage of 2D shape descriptors (perimeter, surface, number of holes, etc). This is due to the fact that we cannot extract the complete H structure and we can only assess the global shape of the H.

When the « H » is clearly visible, we determine its boundaries with line detection techniques. We then verify that the detected lines can be classified into two sets of parallel lines and that when joined they form the shape of an H.



Fig 15: The binary image of the H.



### 3. IMPROVING THE OVERALL SAFETY BY INTELLIGENT DESIGN SOLUTIONS

The final objective of the system Airbus is developing is to enable the safe automatic landing of a helicopter on an offshore platform. As for any system, Airbus Helicopters follows ARP4754 methodology to describe the system in terms of functions and analyse the consequences of function failures. Wrong helipad position determination can lead to a catastrophic event, and for this reason the functions developed require very stringent reliability performance or mitigations factors to insure the safety of the system.

To provide the most reliable autonomous landing system we propose a set of principles that we divide into three categories: multiple independent validation steps within the algorithm, automatic real-time verification of the results outside of the algorithm and lastly validation and supervision through interaction with the pilot.

Within the algorithm, the methodology to decrease the chance of producing a false alarm relies on several independent validations steps mentioned in the section 2.1.4. We will quickly remind all those requirements. The first one is to have two concentric circles on an area. Then, the colours of the circle and the surrounding area should match that of a landing platform. Furthermore, the inner area of the circle should have an H. Thus, for a false detection to be validated, the object needs to have all the properties of a helipad while not being a helipad which is extremely unlikely.

We also propose several solutions to automatically detect when the system fails to detect the desired objects. A generic 2D tracking system provides an independent method to predict the position of the helipad in the screen at each frame. By comparing the results of the 2D tracking with those of the landing platform detector, it is possible to verify that the system is indeed tracking the desired object. By extension, we may know in real-time when the detection starts to produce incorrect results. Additionally, the AFCS verifies that the real-time evolution of the computed relative position of the helipad is coherent with the 3D motion of the Helicopter. This enables another independent verification of the provided results.

To reduce the criticality level of the functions implemented, our system includes a human machine interface that mitigates the risk of false detections. A symbology is presented to the pilot that enables him to validate in real-time the results

given by our algorithm. Additionally, the pilot only can trigger the coupling of the system with the AFCS, which means that the automatic landing mode will be only activated when the pilot is certain that the provided information is reliable.

### 4. TESTING THE HELIPAD DETECTION ALGORITHM ON DIFFICULT IMAGES

In the following section, we will present the results obtained with the detection algorithm on images chosen to accurately represent the most difficult conditions our system is expected to face. These include: rapidly moving cameras, vibrations, and very low contrast.

Five images were used. For each one of them we present the results obtained in the full image and in a zoomed image of the helipad with the inside and outside contours displayed.

The first three images were taken from videos provided by offshore operators as the cameras of our system were still being specified. The last two images were captured by the EAGLE system on flight. We will quickly describe the difficulties of each image presented.

Let us first describe the images provided by our offshore partners. The first image was taken during an approach for an oil & gas transport mission. The image is blurry due to the rapid movement of the camera. The sensor quality is low and the helipad has a much lower contrast than the rest of the offshore rig. The second image has the worst overall contrast as the lighting conditions were extremely poor during the recording. The third one was taken during the maintenance of wind turbines; its difficulty is much lower than the two previous images. However, we can see that the colours of the helipad have shifted to the green.

The last two images were taken with the EAGLE system. They were taken from very far (above one km) with a low angle camera and consequently the vibrations are much more perceptible while the colours have less variations. The fourth image was taken with a forbidden approach direction landmark in front the helipad; this severely reduces the portion of the elliptic ring to be detected. The last image was taken during an advance towards to the landing platform. The approach angle is estimated to be under 7° while the helipad was at a distance farther than 1500m.



Fig 16: Detection results in the full image.



Fig 17: Inside and outside contours for each previously shown image.

We can see (Fig 16) that our algorithm is able to detect the helipad even for the most difficult conditions. We can also note (Fig 17) that the detection has a high accuracy. The detected contours of the helipad are very precise despite the blurry and little contrasted images.

## 5. CONCLUSION AND PERSPECTIVES

Significant progress in the detection of helipads has been achieved in the frame of the EAGLE project to allow safe helicopters automatic landing on offshore platforms.

It has been shown that techniques and algorithms already available for the detection of norms compliant helipads have shortcomings and that innovative computer vision methods had to be developed to enable a robust automatic detection of helipads in representative environmental conditions. Specific difficulties had to be addressed:

- Detection of helipads with grazing angle views and adverse weather conditions,
- Real time and hardware specific constraints,
- Compliance to stringent safety requirements and certification rules of the rotorcraft industry.

The innovation of our methodology relies on a set of requirements or criteria that a platform should abide to. First we look for ellipses in the image as it is the most visible property of the landing platform. Then we search for the other visual features at each potential location of the helipad. Each criterion, when observed, reinforces the probability that the detected object is indeed the researched landing platform. Consequently, relying on a high number of criteria makes wrong detection hardly possible.

Our system was evaluated with videos obtained by the EAGLE optic system to provide the most representative conditions of real operations through tests which demonstrated remarkable detection performance:

- Efficient detection from distances above 1500m with a 7° approach angle,
- High precision of 3D pose estimation, differences to the ground truth lower than 3% of the distance to the helipad,
- Every helipad, once detected, is tracked until the landing of the helicopter.

Having validated the system performance on conditions representative of real operations, the next step will be to test the system on real-time situations. This requires the integration of the various algorithms on a platform linked to the optronic system, namely: the ellipse detection, the image stabilization software, the generic 2D tracker and the supervisor. A human machine interface is also being developed and a direct link with the automatic flight control system is to be implemented. It will enable the execution of automatic landings with minimal interactions with the pilot.

To continue our investigations in this domain and ensure that our system is effective even in the worst possible environmental conditions, we will test our algorithms with data generated by a 3d

graphic engine. Indeed a simulation environment will enable the assessment of our system's performance over a wider range of environmental and flight conditions at a low cost and a fast paced development.

Recent breakthroughs in image processing are providing new perspectives for flight safety. We can imagine, in the near future, systems combining information from different sensors (GPS, IMU, computer vision) for automatic ship deck landing, emergency landings, hoist operations, collision avoidance and many others.

Computer vision is expected to have more and more applications in the near future as progress in algorithms and hardware processing power are expected to continue at an ever increasing pace.

In the automotive industry, computer vision is already very present in different ways: autonomous vehicles, emergency braking systems, automatic parking...

As the helicopters industry is now very mature and few breakthroughs are expected in the mechanical or structural domains, the development of systems, and in particular computer vision-based systems, will become key differentiating features.

It is expected that the coupling of these and other systems to the automatic flight control system will revolutionize the way helicopters and drones will fly in the future. For this reason, it is of the utmost importance for Airbus Helicopters to secure a leading role in the development of new vision systems and their software.

## 6. REFERENCES

- [1] JIA, Qi et al. (2017). « A Fast Ellipse Detector Using Projective Invariant Pruning ». In : IEEE Transactions on Image Processing 26.8, p. 3665–3679.
- [2] FITZGIBBON, Andrew, Maurizio PILU et Robert B. FISHER (1999). « Direct Least Square Fitting of Ellipses ». In : IEEE Trans. Pattern Anal. Mach. Intell. 21.5, p. 476–480.
- [3] BRAAK, Gert-Jan van den et al. (2011). « Fast Hough Transform on GPUs : Exploration of Algorithm Trade-Offs ». In : Advanced Concepts for Intelligent Vision Systems. Springer Berlin Heidelberg, p. 611–622.
- [4] TAGZOUT, Samir, Karim ACHOUR et Oualid DJEKOUNE (2001). « Hough transform algorithm for FPGA implementation ». In : Signal Processing 81.6, p. 1295–1301.

- [5] HOUGH, P. V. C. (1959). « Machine Analysis of Bubble Chamber Pictures ». In : Conf. Proc. C590914, p. 554–558.
- [6] YUEN, H.K., J ILLINGWORTH et J KITTLE (1989). « Detecting partially occluded ellipses using the Hough transform ». In: Image and Vision Computing 7.1, p. 31–37.
- [7] ZHANG, Si-Cheng et Zhi-Qiang LIU (2005). «A robust, real time ellipse detector» In Pattern Recognition 38.2, p.273– 287.
- [8] Haralick, Robert & Shanmugam, K & Dinstein, I. (1973). Textural Features for Image Classification. IEEE Trans Syst Man Cybern. SMC-3. 610-621.
- [9] CHIEN, Chung-Fang, Yu-Che CHENG et Ta-Te LIN (2011). « Robust ellipse detection based on hierarchical image pyramid and Hough transform ». In: Journal of the Optical Society of America A 28.4, p. 581.
- [10] TSUJI et MATSUMOTO (1978). « Detection of Ellipses by a Modified Hough Transformation ». In: IEEE Transactions on Computers C-27.8, p. 777–781.
- [11] KIM, Euijin, Miki HASEYAMA et Hideo KITAJIMA (2002). « Fast and Robust Ellipse Extraction from Complicated Images ». In: Proceedings of the IEEE International Conference on Information Technology and Applications.
- [12] Lee, Sewon & Jang, Jinwon & Baek, Kwangryul. (2012). « Implementation of vision-based real time helipad detection system ». International Conference on Control, Automation and Systems. 191-194.
- [13] Cosimo Patruno, Massimiliano Nitti, Ettore Stella, and Tiziana D’Orazio 2017 « Helipad detection for accurate UAV pose estimation by means of a visual sensor » International Journal of Advanced Robotic Systems
- [14] Prakash, R Om & Chandran, Saravanan. (2016). “Autonomous Robust Helipad Detection Algorithm Using Computer Vision”. 10.1109 / ICEEOT 2016 7755163.
- [15] A. Rodriguez-Ramos, C. Sampedro, H. Bavlé, Z. Milosevic, A. Garcia-Vaquero and P. Campoy, "Towards fully autonomous landing on moving platforms for rotary Unmanned Aerial Vehicles," 2017 International Conference on Unmanned Aircraft Systems (ICUAS), Miami, FL, USA, 2017, pp. 170-178.
- [16] P. Zoppitelli, S. Mavromatis and J. Sequeira (2017). “Ellipse detection in very noisy environment”. In : Proceedings of the Computer Science Research Notes [CSRN] 2703 ISSN 2464-4617, International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2017 p. 31–40.
- [17] BASCA, C.A., M. TALOS et R. BRAD (2005). « Randomized Hough Transform for Ellipse Detection with Result Clustering ». In: EUROCON 2005 - The International Conference on "Computer as a Tool"