

---

# 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, F-13228 Marseille cedex 9  
christophe.palmann@gmail.com  
{sebastien.mavromatis, jean.sequeira}@univ-amu.fr*

---

*RÉSUMÉ. Cet article 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 a unified manner with a new type of primitives; then, we set the link between the existence of 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.29.29-49 © 2012 Lavoisier

### **Extended abstract**

This paper introduces a new approach to characterize similar areas of two 2D point clouds. This is performed through the definition of a similarity measure that takes into account several criteria such as invariance by rotation, outlier elimination, and one-dimensional structure enhancement.

Image analysis and pattern recognition often use 2D point clouds as processed representations of images (e.g. these points being the gradient highest value locations, or Laplacian zero-crossing ones). When using images of different modalities (e.g. multispectral and radar ones for “Remote Sensing” applications), this binary representation can be a useful bridge between these two images to register them.

Several algorithms have been designed to match point clouds (e.g. ICP for Iterative Closest Point, or RANSAC for RANdom SAmple Consensus,) but none of them takes advantage of underlying one-dimensional structures, which is a “knowledge element” we have to take into account in most practical cases. The approach we propose here integrates this “knowledge element” in the matching process.

Let us consider two locations, one in each point cloud, and a neighborhood of these locations. Our goal is to evaluate how much these neighborhoods are similar, up to a rotation. The most important outcome of this evaluation is the characterization of landmark pairs that can be used for globally registering the two point clouds, without any a priori knowledge on the registration transformation.

Our main contributions are the integration of various one-dimensional structure representations into a unified formalism, and the design of a robust estimator to evaluate the common information related to these structures.

We first design a primitive model that can take into account all the possible configurations of points related to one-dimensional structures. These configurations depend on the way the original structures are modified by the noise induced by sensors and preliminary processes. In this paper, we show how to detect such one-dimensional structures and how to calculate the parameters of the corresponding primitives.

We then study the distribution of primitive orientation differences when these primitives respectively belong to the neighborhood of two given locations, each of them in a different point cloud. If these areas can be roughly derived one from the other by using a rotation, then a peak appears in the distribution histogram. This peak corresponds to the rotation angle value, but we do not focus on this value: what we are interested in is the characterization of such a peak in the histogram because it means that these two areas carry common information.

In this paper, we propose a theoretical study as well as a set of experimental tests to robustly characterize a peak emergence. One of the key points of our approach is the control of the probability of having a false positive detection: in order to avoid such wrong results, we discuss the opportunity of handling this characterization

through the study of histogram modes (instead of histogram peaks) and their inter-relations, and finally we propose a new estimator based on this discussion.

The last section is dedicated to results. We illustrate our approach with a few examples based on the problem of image registration: images are from different modalities (radar ESR2 vs multispectral Landsat 7, radar ESR2 vs panchromatic Landsat 7); we propose in each case a scheme to obtain the corresponding binary images, and we show pairs of corresponding locations (with their neighborhood) obtained using our algorithm.

## 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 faisons 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 entier) : 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 voisinages), cette mesure permettant de quantifier l'information commune aux deux voisinages associés.

## 2. État de l'art

De nombreuses méthodes traitent de la mise en correspondance de nuages de points. Parmi ces méthodes, une sous-classe nous intéresse plus particulièrement ; il s'agit de celles qui introduisent un élément de connaissance supplémentaire dans le processus d'identification à travers une hypothèse généralement vérifiée (et qui correspond à un élément perceptuel d'évaluation de la ressemblance des sous-nuages) : il existe dans les nuages des structures linéiques sous-jacentes, qui se retrouvent sous la forme d'agglomérations (sous des formes diverses) de points.

En effet, les nuages de points issus de cas réels 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, qui ont la propriété d'être communes aux nuages dans une large mesure. 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 (cf. figure 1). Par ailleurs, ces structures peuvent être localement mises en correspondance par des transformations rigides (rotation + translation 2D).

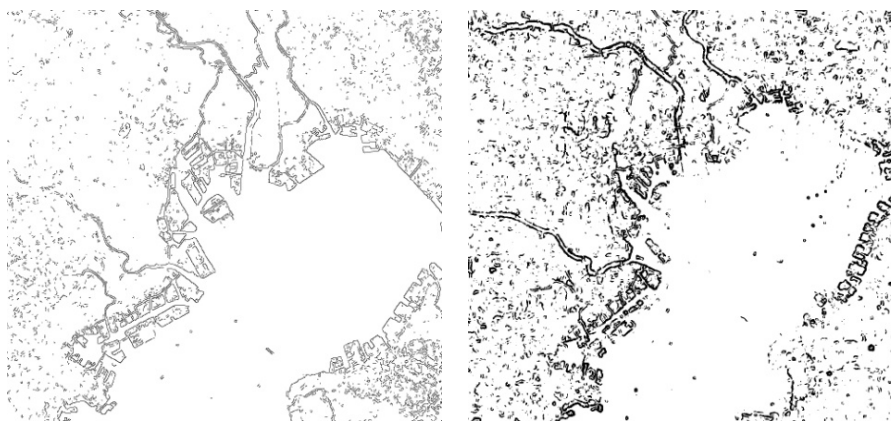


Figure 1. Nuages de points 2D, contenant des objets de différentes complexités

Ces remarques sont importantes, puisque ces structures sont en fait 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 utile à la définition d'une mesure de similarité entre leurs sous-parties (sous-nuages).

Parmi les exemples de méthodes reposant sur de telles structures linéiques, nous pouvons mentionner les « Chain codes » présentés par Li *et al.* (1995). Le principe à la base des Chain codes est d'encoder séparément chaque structure linéique : pour un point sélectionné sur une composante connexe, l'encodage consiste à utiliser la

direction du prochain voisin. Une structure linéique est donc représentée par une séquence d'entiers, dont une unité correspond à un angle de  $45^\circ$ . Plus exactement, les auteurs « dérivent » une courbe à partir de cette séquence, qui sera corrélée aux courbes représentant d'autres structures linéiques pour décider de leurs mises en correspondance. Dans Chen *et al.* (2007), on peut trouver un exemple d'application des Chain codes dans l'identification automatique d'aéroports.

Un certain nombre de méthodes utilisent l'approximation polygonale pour représenter des structures linéiques (une telle approximation peut être effectuée par exemple par des méthodes comme celles de Salotti (2001), ou Ray et Ray (1995). Un exemple concerne les travaux de Zhao et Chen (2004). Dans cette publication, une structure linéique est décrite comme une séquence de segments orientés de même longueur, qui approchent celle-ci comme une ligne brisée. Plus exactement, les auteurs tiennent compte de l'orientation relative d'un segment par rapport aux autres segments qui composent la ligne brisée : celle-ci est alors représentée par une matrice  $n \times n$ , où la ligne  $i$  correspond à la distribution des différences d'orientations entre le  $i$ ème segment et les  $n-1$  autres segments de la ligne brisée. La décision d'apparier deux lignes brisées se fait simplement en seillant la distance entre les deux matrices qui les représentent. Une autre approche intéressante concerne « la distance d'édition », introduite par Wagner et Fisher (1974). Cette distance était liée à l'origine à la correction orthographique, mais elle a été depuis adaptée à la résolution d'autres problèmes, comme les ASM (« Attributed String Matching »), qui correspondent à des chaînes composées de primitives simples (pour les polygones, il peut s'agir de leurs segments ou bien de leurs sommets...). L'idée à la base de la distance d'édition est de découvrir la séquence de modifications de coût minimal qui permette de transformer une chaîne  $A$  en une chaîne  $B$ , avec comme modifications autorisées la suppression, l'insertion et la substitution. La distance ainsi obtenue permet la décision de mise en correspondance de deux polygones. On peut trouver l'utilisation d'une telle méthodologie dans Kaygin et Bulut (2002), et dans Tsai et Yu (1985) où les auteurs retiennent les orientations ainsi que les longueurs des segments de polygones.

Enfin, il existe des techniques qui consistent à approcher des structures linéiques par des primitives de haut niveau, *i.e.* par des courbes paramétrées (en général différentiables au moins deux fois) : il s'agit alors de réaliser les mises en correspondance des structures par l'intermédiaire de leurs représentations continues. D'une certaine manière, nous pouvons interpréter ces approches comme un passage du discret au continu (de telles approximations sont possibles grâce par exemple à des techniques de type contours actifs, introduits par Kass *et al.* (1988)). Dans cet esprit, la méthode de Cui *et al.* (2009) consiste à décrire une courbe par l'intégrale de la courbure en valeur absolue, qui forme une signature invariante par similitude (rotation + homothétie + translations) ; de plus, les auteurs se préoccupent du problème de recaler partiellement deux courbes. Manay *et al.* (2006) utilisent eux aussi la notion d'intégrale pour la mise en correspondance de courbes sous similitude ; on retrouve des approches similaires avec Kong et Kimia (2001) et Sebastien *et al.* (2003). L'avantage principal des représentations par intégrale est leur moindre sensibilité au bruit. Par ailleurs, une sous-classe importante des approches de type « passage au continu » concerne les descripteurs CSS

(« Curvature Scale Space »), introduits par Mokhtarian et Mackworth (1992), Mokhtarian (1995), et Mokhtarian et Bober (2003). L'idée de ces méthodes est d'utiliser un invariant aux transformations affines très simple lorsque l'on dispose de courbes paramétrées : les points d'inflexion, qui correspondent aux endroits de la courbe où la courbure devient nulle et change de signe. Parmi les très nombreuses méthodes dérivant de l'algorithme CSS original, citons par exemple les travaux de Cui *et al.* (2007) où les auteurs appliquent une version améliorée de la méthode sur des images prises de capteurs aéroportés.

Une condition préalable à l'utilisation de ces méthodes est d'être en mesure d'obtenir des nuages de points à partir desquels les structures linéiques peuvent être extraites de façon idéale. Or dans la pratique, cela est loin d'être le cas, et la question de la stabilité des descripteurs devient primordiale (si une structure linéique est légèrement déformée, alors sa description ne doit être que légèrement modifiée). Mais d'autres critères doivent aussi être remplis pour qu'un descripteur puisse être utilisé efficacement : l'unicité (deux primitives différentes doivent avoir des descriptions différentes) et l'indépendance (sous-entendu des attributs qui forment la description) ; de plus, nous voudrions que les primitives soient uniformément réparties dans les nuages, et plus difficile encore, que ces primitives puissent être extraites quels que soient les modes d'acquisition et le contexte de la scène étudiée... Dans la pratique, il est évidemment très difficile de remplir toutes ces conditions à la fois. Manquer à l'une d'entre elles provoque bien souvent un manque de robustesse, où de nombreux faux-positifs peuvent émerger (structures appariées à leurs mauvaises contreparties dans l'autre nuage).

Dans la suite de cet article, nous proposons donc une mesure de similarité entre sous-nuages de points reposant sur des structures linéiques qui réponde à de telles contraintes pratiques.

### 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 (cf. figure 2).

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 l'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 des valeurs discrètes dans

$[0^\circ, 180^\circ[$ ). Ces primitives forment alors un dénominateur commun entre les différentes structures linéiques possibles (fines, épaisses, lacunaires, bruitées...).



Figure 2. De gauche à droite : structure fine, structure épaisse, structure non connexe, présence de bruit

Le processus d'extraction de telles primitives consiste, pour un point donné  $P$  d'un nuage, à caractériser un ensemble de points qui s'organisent de façon linéaire dans le voisinage de  $P$ . Nous utilisons pour cela un ensemble de masques circulaires binaires, qui chacun représente une orientation discrète particulière (cf. figure 3). Les masques sont caractérisés par deux paramètres :

- le rayon  $R_m$ , qui définit la taille du voisinage autour d'un point courant que nous estimons pertinente dans l'analyse d'une orientation locale ;
- l'épaisseur  $E$  de la « ligne d'orientation », utile à la prise en compte d'alignements évanescents.

L'orientation de chaque masque est une valeur discrète comprise dans l'intervalle  $[0^\circ, 180^\circ[$ . Le nombre d'orientations est noté  $n'$  ; ce nombre est directement relié au rayon  $R_m$  des masques (progression linéaire).

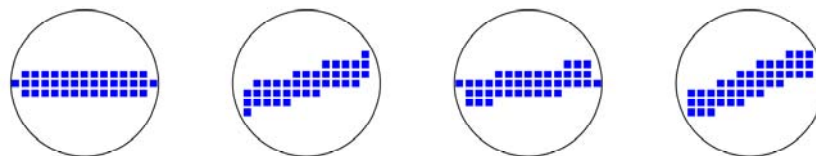


Figure 3. Exemple des premiers éléments d'un ensemble de masques binaires ( $R_m = 7$ , et  $E = 3$ )

Le masque ayant intercepté le plus de points dans un voisinage  $R_m$  permet d'identifier une orientation locale centrée sur  $P$ . Si le nombre de points interceptés est supérieur à un seuil  $S_{\min}$ , alors une primitive potentielle est détectée. La sélection finale des primitives suit alors un processus itératif : la primitive de meilleur score est sélectionnée, puis les suivantes sont sélectionnées à leur tour à condition qu'elles soient situées à une distance supérieure à un seuil  $D_{\min}$  d'une primitive déjà sélectionnée. Le paramètre  $D_{\min}$  permet d'éviter la redondance de la même information à un niveau local, qui générerait la représentation des structures linéiques.

Enfin, les orientations de chaque primitive sont affinées en reprenant les points interceptés par le meilleur masque, et en procédant par exemple à une régression linéaire ou bien une analyse en composantes principales (ACP cf. figure 4). Les orientations correspondent alors à un ensemble discret de  $n$  valeurs dans  $[0^\circ, 180^\circ[$ , où  $n > n'$ .

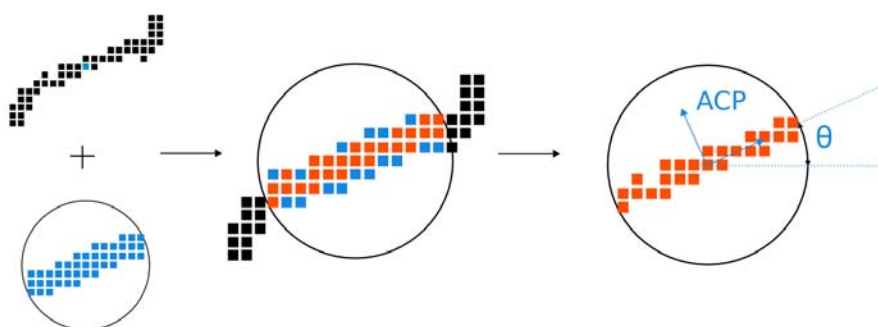


Figure 4. Calcul d'une orientation locale plus précise par ACP, grâce aux points du voisinage ayant intercepté le masque de meilleur score

Pour de plus amples informations sur cet algorithme, nous invitons le lecteur à consulter la référence (Palmann *et al.*, 2009). De même, nous suggérons la lecture des travaux de (Desolneux *et al.*, 2000), où l'on évalue la probabilité pour qu'un



alignement de points (entre autres formes géométriques) soit dû au hasard : si cette probabilité est suffisamment faible, alors un alignement est détecté.

L'étude d'un grand nombre de relations géométriques entre primitives 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.

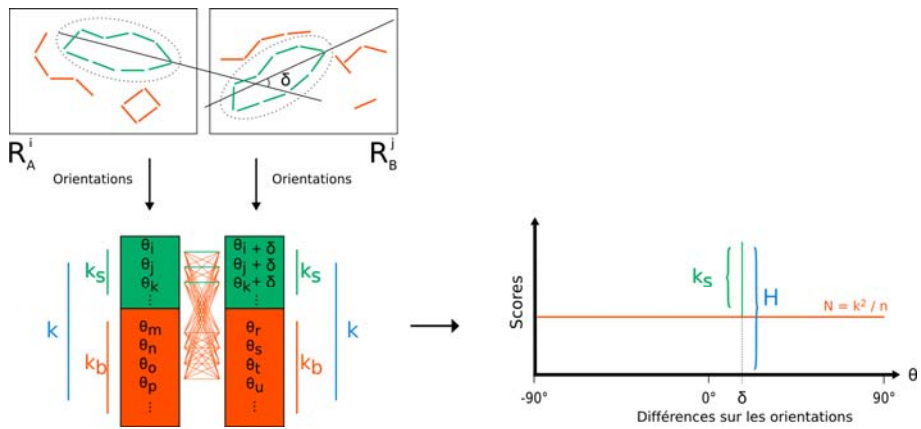


Figure 5. Deux régions possédant des primitives communes entourées en pointillés ; histogramme des différences d'orientations prises deux à deux

Plus précisément, supposons que nous ayons deux régions  $R_A^A$  et  $R_B^B$ , chacune dans deux nuages de points A et B (voir figure 5). 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 pointillés) ; ainsi,  $k = k_s + k_b$ . Par ailleurs, 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 (cf. figure 5).

Il se forme alors en  $\delta$  un maximum  $H$  dont nous pouvons donner approximativement la valeur :

$$H = k_s + N \quad (1a)$$

$$\text{soit } H = (1 - \varepsilon).k + N \quad (1b)$$

où  $N = k^2/n$  correspond à la valeur moyenne du bruit et où  $\varepsilon = k_b/n$  et s'interprète comme le taux de contamination. On remarquera que  $(1-\varepsilon)$  correspond au taux d'information commune entre deux sous-nuages, puisque  $(1-\varepsilon) = k_s/k$  : 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 ( $\varepsilon = k_b/k$ ).

#### 4.1. Étude expérimentale

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

---

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

différences d'orientations est présenté dans la figure 6. Comme on peut le constater, le maximum est bien situé en  $10^\circ$ , et ceci sans ambiguïté.

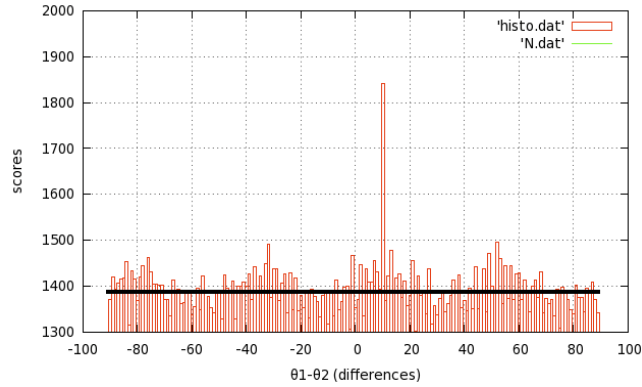


Figure 6.  $k = 500$ ,  $n = 180$  ( $1^\circ$ ),  $\varepsilon = 0\%$ . Maximum correctement placé en  $10^\circ$

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 7. 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.

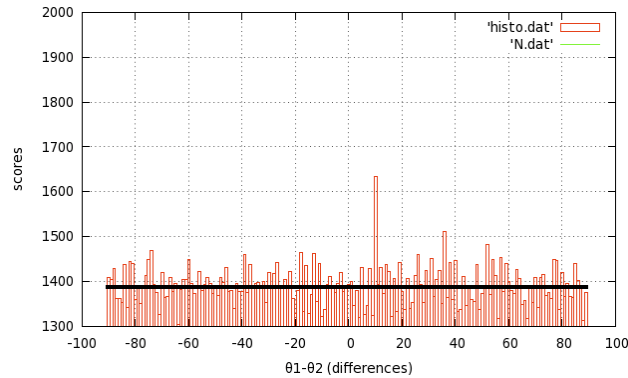


Figure 7.  $k = 500$ ,  $n = 180$  ( $1^\circ$ ),  $\varepsilon = 50\%$ . Maximum correctement placé en  $10^\circ$

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 8. 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.

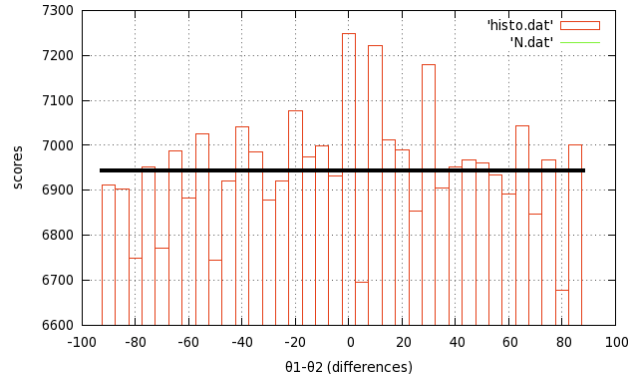


Figure 8.  $k = 500$ ,  $n = 36$  ( $5^\circ$ ),  $\varepsilon = 50\%$ . Obtention d'un faux-positif en  $0^\circ$

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 dans la section suivante). Les deux histogrammes sont représentés dans la figure 9. 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 sous-nuages, nous ne travaillons que sur de petits sous-ensembles de primitives.

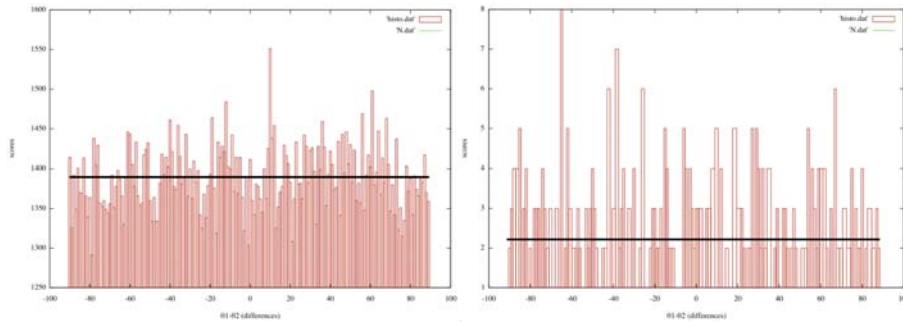


Figure 9.  $k = 500$  (à gauche) et  $k = 50$  (à droite),  $n = 180$  ( $1^\circ$ ),  $\varepsilon = 80\%$ .  
Maximum correctement placé en  $10^\circ$  dans le cas  $k = 500$  ; obtention d'un faux-positif en  $-65^\circ$  dans le cas  $k = 50$

#### 4.2. Étude théorique

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, Huttenlocher, 1990, Stewart, 1995). Pour cela, nous nous concentrons sur l'évolution des occurrences de  $n$  cellules<sup>2</sup> de l'histogramme lorsque celles-ci sont incrémentées uniformément ; notre approche est de type probabiliste, et nous considérerons le calcul d'une occurrence comme l'équivalent d'un tirage d'une expérience aléatoire.

Soit donc l'une de ces  $n$  cellules ; il n'existe que deux issues possibles pour cette cellule lors du calcul d'une différence d'orientations (*i.e.* un tirage) : soit le score de cette cellule est incrémenté, soit il ne l'est pas. Comme nous nous intéressons à la formation du bruit, et que celui-ci possède une distribution uniforme, le score d'une cellule quelconque est bien entendu incrémenté avec une probabilité constante de  $1/n$  à chaque tirage. Une information pertinente pour nous est de savoir quelle est la probabilité pour que le score d'une cellule quelconque soit égal à «  $s$  » au bout de  $k^2$  tirages ; nous sommes en présence d'un schéma de Bernoulli (un succès correspondant à une incrémentation), et la loi que nous recherchons est une loi binomiale de paramètres  $(k^2, 1/n)$ . Il vient alors immédiatement que la probabilité pour que le score d'une cellule quelconque soit supérieur ou égal à une valeur  $H$  est :

$$P(\text{score} \geq H) = \sum_{s=H}^{k^2} \frac{k^2}{s} \cdot \left(\frac{1}{n}\right)^s \cdot \left(\frac{n-1}{n}\right)^{k^2-s} \quad (2)$$

Cette expression peut-être interprétée comme une mesure du risque pour qu'un pic issu du bruit supplante un pic de hauteur  $H$  ; autrement dit, nous mesurons le risque d'apparition d'un faux-positif. Cette expression reste néanmoins d'une utilisation difficile puisqu'elle n'a pas d'expression analytique. Pour remédier à ce problème, nous avons procédé à deux simplifications. Tout d'abord, puisque dans la pratique  $k^2 > 50$  et  $n > 10$ , la loi binomiale peut être remplacée par une loi de Poisson (Saporta, 2006). Ensuite, nous pouvons nous intéresser non pas au risque d'apparition d'un faux positif, mais à un majorant de celui-ci. Si un tel majorant est suffisamment proche de la valeur réelle du risque, alors ce majorant peut être assimilé au risque lui-même. La borne de Chernoff (Stewart, 2009) nous permet de répondre à cet impératif. Ainsi, en utilisant ces deux simplifications, et en remplaçant  $H$  par la formule (1b) (voir début de section), nous avons obtenu la formule suivante :

$$P(\varepsilon, k, n) = e^{-\left((1-\varepsilon) \cdot k + \frac{k^2}{n}\right) \cdot \log\left(\frac{k^2}{n \cdot ((1-\varepsilon) \cdot k + \frac{k^2}{n})}\right) + (1-\varepsilon) \cdot k} \quad (3)$$

---

2. Nous rappelons que  $n$  correspond aussi à la précision des orientations.

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 des approches expérimentales et théoriques ont donc 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 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 (3) 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 10, 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 noire représente alors les précisions requises sur les orientations en fonction de leur nombre pour une identification robuste du signal recherché.

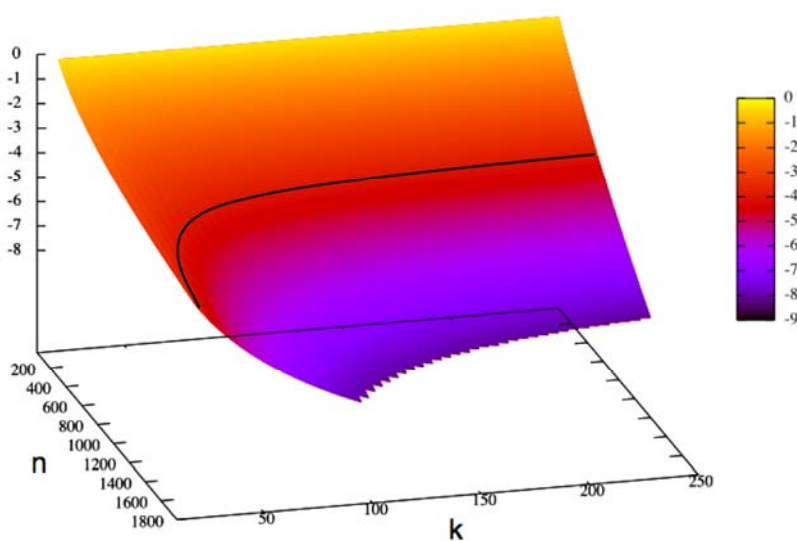


Figure 10. Probabilité d'obtenir un faux-positif (échelle logarithmique) en fonction du nombre d'orientations  $k$  et de leur précision  $n$ , pour un taux de contamination  $\varepsilon = 80\%$ . La courbe noire correspond à la résolution de l'équation  $P(\varepsilon, k, n) = P_0$  avec  $P_0 = 10^{-4}$

Par ailleurs, l'émergence d'un pic isolé, caractéristique d'une information commune, peut être perturbée par le fait que les primitives en correspondance n'ont 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  $\delta$ . 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 (cf. figure 11).

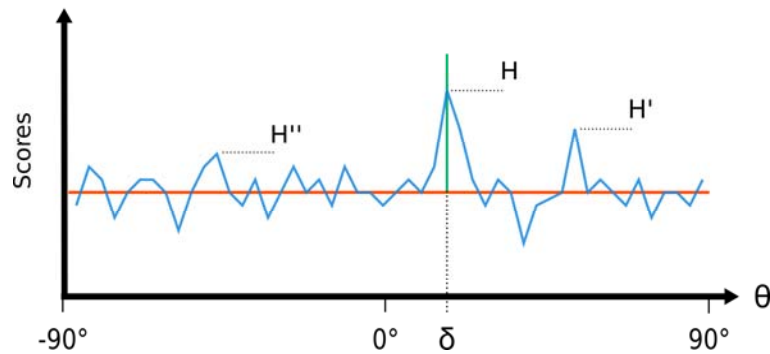


Figure 11. Localisation imparfaite du signal recherché (H), et formation de maxima locaux (H' et H'')

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 :

$$\alpha = \frac{|H - \bar{H}|}{H} \quad \text{avec} \quad \bar{H} = \frac{H' + H''}{2} \quad (4)$$

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 (filtre médian par exemple) 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. Notre méthode se situe en aval de ce processus d'extraction, ce qui permet d'adapter son utilisation à de nombreuses situations différentes ; la seule contrainte requise est d'être en mesure de produire des nuages contenant une information commune suffisante.

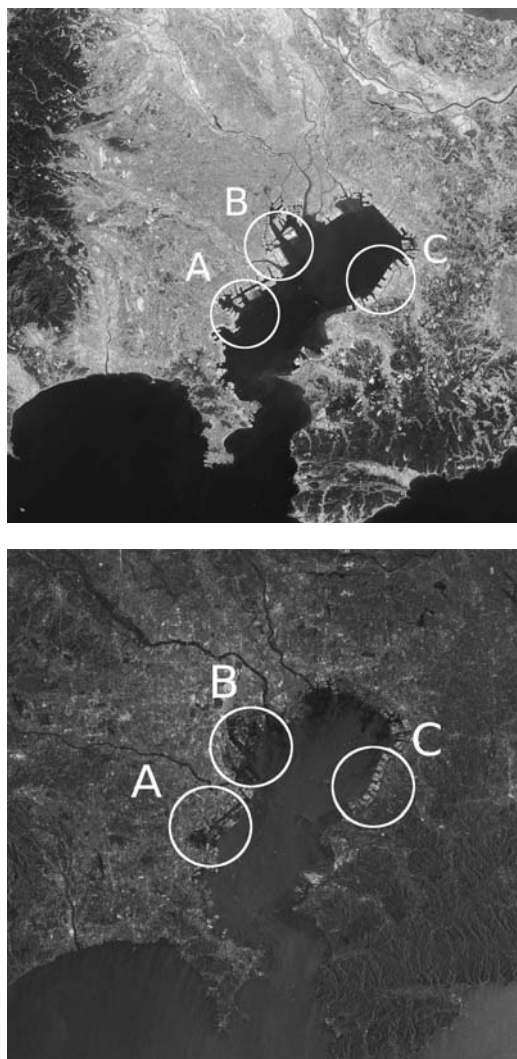
Dans les deux exemples ci-après, nous utilisons deux images radar (ERS2) avec une image multispectrale et une image panchromatique (Landsat 7), dont les résolutions sont proches l'une de l'autre (30 m environ) ; la première paire d'images concerne des images acquises au-dessus de la baie de Tokyo, tandis que la seconde concerne des images acquises au-dessus de Pékin (voir figures ci-après). Pour ces deux paires, nous avons utilisé un processus d'extraction générique afin d'obtenir des nuages de points avec structures linéiques. Ainsi, pour les images radar, nous avons utilisé l'opérateur de Touzi (Touzi *et al.*, 1988) spécialement dédié à ce type de modalité ; pour l'image multispectrale, nous avons tout d'abord utilisé une analyse en composantes principales (ACP) afin de produire des images en niveaux de gris contenant une quantité maximale d'information (au sens de la variabilité des pixels), puis nous avons utilisé l'opérateur d'extraction de contours de Canny-Deriche (Deriche, 1987). Pour l'image panchromatique, nous avons omis le traitement par ACP, puisque celle-ci était déjà en niveaux de gris.

Ce type de traitements favorise l'émergence de frontières situées sur de fortes variations de niveaux de gris : dans la première paire d'images nous privilégions l'émergence de la ligne de côte (la figure 1 correspond à deux sous-parties de ces nuages de points), tandis que dans la seconde paire nous privilégions l'émergence des lacs. D'autres traitements auraient pu être utilisés : par exemple, puisque la seconde paire d'images concerne une scène en milieu urbain, nous aurions pu extraire le réseau routier afin de produire des nuages de points pertinents.

Afin d'identifier des zones supportant une information commune, chaque image a été « divisée » régulièrement en une grille de positions espacées de 150 pixels. En chacune de ces positions, une zone est décrite par un cercle de rayon égal à 150 pixels ; ainsi, il existe un recouvrement partiel des zones de recherche. Pour le processus d'extraction des primitives, nous avons utilisé les paramètres suivants :  $R_m = 6$ ,  $E = 3$ ,  $S_{\min} = 0.7$  et  $D_{\min} = 4$ . Pour chaque paire d'images, nous avons retenu les trois paires de sous-nuages ayant produit les meilleures valeurs de similarité. Les résultats sont directement représentés dans les images suivantes, où l'on peut



constater des appariements corrects situés sur les zones géographiques qui ont été privilégiées par les traitements effectués en amont.



*Figure 12. Baie de Tokyo ; image radar ERS2 (en haut)  
et image multispectral Landsat 7 (en bas)*

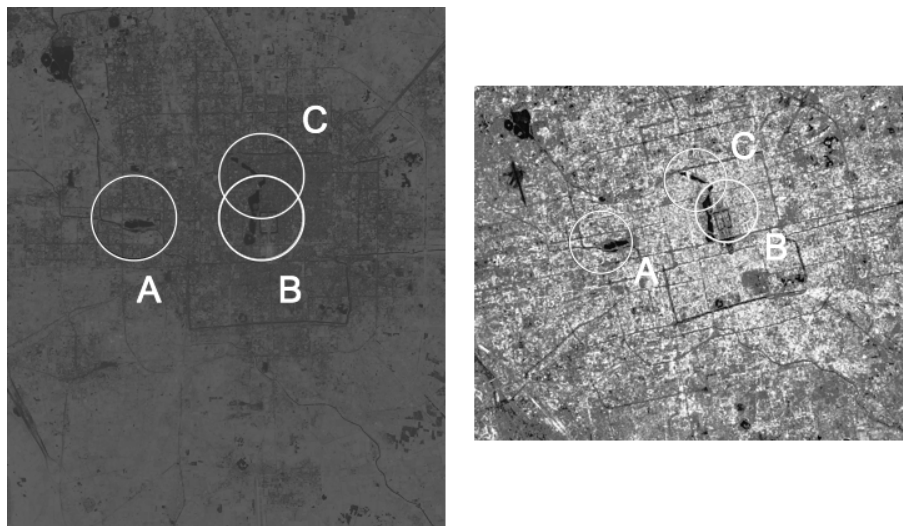


Figure 13. Ville de Pékin ; image panchromatique Landsat 7 (à gauche),  
et image radar ERS2 (à droite)

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 images en niveaux de gris précédentes. 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 nuages 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., Silbermann M. (1986). Mouvements et traitement d'images. *Traitement du Signal*, vol. 3, n° 6, p. 281-290.

- Chen Y., Sun K., Zhang J., Lin Z. (2007). Automatic recognition of airport in remote sensing images based on improved methods, *International Symposium on Multispectral Image Processing and Pattern Recognition*. vol. 6786, n° 2, p. 45.1-45.8.
- Cui M., Femiani J., Hu J., Wonka M., Razdan A. (2009). Curve matching for open 2D curves. *Pattern Recognition Letters*, vol. 30, n° 1, p. 1-10.
- Cui M., Wonka P., Razdan A., Hu J. (2007). A new image registration scheme based on curvature scale space curve matching. *The Visual Computer: International Journal of Computer Graphics*, vol. 28, n° 8, p. 607-618.
- 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.
- Desolneux A., Moisan L. and Morel J.-M. (2000). Meaningful Alignments. *International Journal of Computer Vision*, vol. 40, n° 1, p. 7-23.
- 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.
- 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.
- Kass M., Witkin A., Terzopoulos D. (1988), Snakes: Active contour models. *International Journal of Computer Vision*, vol. 1, n° 4, p. 321-331.
- Kaygin S., Bulut M.M. (2002). Shape recognition using attributed string matching with polygon vertices as the primitives, *Pattern Recognition Letters* vol. 23, n° 1-3, p. 287-294.
- Kong W., Kimia B.B. (2001). On solving 2d and 3d puzzles using curve matching. *IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, p. 583-590.
- Li H., Manjunath B., Mitra S. (1995). A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, vol. 4, n° 3, p. 320-334.
- Manay S., Cremers D., Hong B.W., Yezzi A.J., Soatto S. (2006). Integral invariants for shape matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence* vol. 28, n° 10, p. 1602-1618.
- Mokhtarian F., Bober M., (2003). *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*. Dordrecht, Kluwer Academic Publishers.
- Mokhtarian F. (1995). Silhouette-based isolated object recognition through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, n° 5, p. 539-544.
- Mokhtarian F., Mackworth A.K. (1992). A theory of multi-scale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, n° 8, p. 789-805.

- 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.
- Ray B.K., Ray K.S (1995). A new split-and-merge technique for polygonal approximation of chain coded curves. *Pattern Recognition Letters*, vol. 16, n° 2, p. 161-169.
- Salotti M. (2001). An efficient algorithm for the optimal polygonal approximation of digitized curves. *Pattern Recognition Letters*, vol. 22, n° 2, p. 215-221.
- Saporta G. (2006). *Lois de probabilité discrètes d'usage courant. Probabilités, analyse des données et statistique*. Paris, Editions Technip, p. 34.
- Sebastian T.B., Klein P.N., Kimia B.B. (2003). On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, n° 1, p. 116-125.
- 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.
- 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.
- Tsai W.H., Yu S.S. (1985). Attributed string matching with merging for shape recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, n° 4, p. 453-462.
- Wagner R.A., Fisher M.J. (1974). The string-to-string correction problem. *Journal of ACM*, vol. 21, n° 1, p. 168-173.
- Zhao Y., Chen Y.Q. (2004). Connected Equi-Length Line Segments For Curve And Structure Matching. *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, n°6, p. 1019-1037.

**Christophe Palmann** est titulaire depuis 2011 d'un Doctorat de l'Université Aix-Marseille dans le domaine du traitement d'images. Ses recherches ont porté sur les estimateurs robustes et leurs applications au recalage d'images de modalités différentes, en particulier les images issues de la télédétection. Il a notamment développé ses algorithmes en collaboration avec l'IRSA (« Institut of Remote Sensing Applications », Pékin), institut de recherche appartenant à l'Académie Chinoise des Sciences. Ses domaines d'intérêt portent également sur les applications en vision industrielle.

**Sébastien Mavromatis** est maître de conférence à l'Université d'Aix-Marseille depuis 2003. Il est membre de l'équipe SimGraph du LSIS (UMR CNRS 7296). Il s'intéresse à l'intégration et à la visualisation d'informations extraites d'images d'origines diverses et/ou formant un volume important de données, ainsi qu'aux possibilités de simulation à partir de ces informations. Ses travaux s'intègrent dans de nombreux projets industriels qui concernent, par exemple, l'analyse de séquences d'images pour la vidéo surveillance ou pour la simulation de scènes sportives (projet SimulFoot). Il collabore avec l'IRSA (Institute of Remote Sensing Applications) sur l'intégration d'informations issues de données de télédétection.

***Jean Sequeira** est Professeur à l'Université Aix-Marseille depuis 1991. Il est diplômé de l'École Polytechnique (1977) et de l'École Nationale Supérieure des Télécommunications (1979). Il est également titulaire d'un diplôme de Docteur Ingénieur (1982) et d'un Doctorat d'État (1987). Avant d'entrer à l'Université, il a occupé un poste de chercheur au Centre Scientifique d'IBM France (1981-1991). Son activité de recherche depuis une trentaine d'année couvre divers domaines de l'imagerie numérique. Il est également Professeur à l'IRSA (Institute of Remote Sensing Applications – Académie Chinoise des Sciences) et à l'ISIG (Institut Supérieur d'Informatique et de Gestion - Burkina Faso). Il est actuellement le chef de l'équipe SimGraph du laboratoire LSIS (UMR CNRS 7296 à Marseille).*

