

Détection multirésolution des changements par analyse spectrale locale de Walsh-Hadamard

Olivier Desprez

Eric Petit

Laboratoire d'Etudes et de Recherches en Instrumentation Signaux et Systèmes (L.E.R.I.S.S.)
Université Paris XII Val de Marne, Avenue du Général de Gaulle, 94010 Créteil, France

RÉSUMÉ

Cet article traite du problème de la détection automatique des changements sur un couple d'images multitemporelles en imagerie dite de surveillance. Les limites des approches classiques nous ont conduit à développer un nouvel opérateur, fondé sur l'analyse en composantes principales d'attributs spatio-fréquentiels extraits de spectres bidimensionnels locaux de Walsh-Hadamard à échelles dyadiques décroissantes. L'opérateur est utilisé dans le cadre d'une stratégie d'analyse multirésolution, ce qui permet de détecter des changements d'ordre textural tout en réduisant de façon significative le problème d'occlusion régionale sans perte de résolution spatiale. La méthode proposée fournit une description hiérarchique des changements de la scène et peut en outre bénéficier d'une optimisation calculatoire par décomposition du couple d'images en pyramides passe-bas.

1 Introduction

Le problème de la détection des changements intervient dans le traitement des images acquises par des capteurs observant la même scène de façon répétitive. Dans un contexte de réactualisation de scène, le but est d'effectuer une comparaison non dirigée entre deux images acquises par le même capteur à des instants différents. Les changements détectés entre les images doivent être significatifs de l'apparition ou de la disparition d'objets dans la scène qui peut alors être mise à jour en fonction de son évolution temporelle. En imagerie dite de surveillance l'intervalle de temps séparant l'acquisition des deux images est faible, tout en restant supérieur à la cadence vidéo. La réactualisation s'effectue alors dynamiquement par scrutations périodiques de la scène (capteur en « état de veille »).

En réactualisation de scène, les méthodes fondées sur une approche iconique sont les plus classiques. Elles sont basées sur la représentation des images au niveau du pixel et supposent la quasi-invariance de la scène en dehors des zones de changement. Il faut donc mettre les deux images en parfaite correspondance spatiale et radiométrique, puis les comparer au moyen d'attributs de changement extraits sur la scène. La mise en correspondance s'effectue par application de l'opérateur d'intercorrélation bidimensionnelle ou par le calcul des coefficients d'un polynôme (d'ordre > 2), à partir du découpage de la scène en quadrilles appelées *cellules* [1-3]. Sur ces cellules Kawamura [3] extrait trois attributs (coefficient de corrélation, entropie et probabilité de haute

ABSTRACT

This paper deals with the problem of automatic temporal change detection on a pair of images of a scene. Because of the limitations of classical approaches, we developed a new operator based on a principal component analysis of spatio-frequency attributes extracted from Walsh-Hadamard local 2-dimensional spectra of decreasing dyadic scales. The operator is used within the framework of a multiresolution strategy. This allows to significantly reduce the problem of regional occlusion without loss of spatial resolution. The method we propose provides a hierarchic description of the scene changes. It can benefit of a time processing improvement thanks to a low-pass pyramidal decomposition of the original pair of images.

intensité) caractérisant la dissemblance du paysage urbain d'une scène photographique aérienne.



fig.1 : Un couple d'images multitemporelles (256x256 pixels)

L'image des différences (à droite) souligne les risques d'erreurs de détection : ombres portées à l'arrière-plan, bruits de contours, occlusion partielle du piéton, variations locales non significatives.

L'attribut de changement le plus utilisé est la variation locale des luminances obtenue par simple soustraction d'images. Il caractérise les changements de réflectance des structures de la scène à l'échelle du pixel ou d'une cellule. Un seuillage permet alors d'exhiber les zones de changement au sens de la luminance. Le bruit et les faibles contrastes régionaux entre les deux instants d'observation rendent le choix du seuil délicat (fig.1). Une solution consiste à considérer que le bruit est non corrélé spatialement, et à ne conserver comme pertinentes que les zones de l'image possédant un nombre minimum de pixels connexes au-dessus du seuil. Néanmoins cette méthode s'avère inefficace dans le cas où une région changeante présente une texture à hautes fréquences spatiales, ou lorsque le changement provient d'un léger déplacement d'une entité faiblement contrastée. Une solution alternative revient à considérer la dissemblance entre les distributions statistiques des luminan-

ces plutôt que la variation des luminances. Sur ce principe, Eghbali [4] utilise le test de Kolmogorov-Smirnov pour détecter les changements entre deux images multibandes du satellite LANDSAT. Hsu et col. [5] détectent les changements entre deux images consécutives d'une séquence, à partir de la modélisation locale des images par un polynôme fonction des coordonnées des pixels, et du rapport de vraisemblance introduit par Yakimovsky.

Une autre approche dite symbolique est fondée sur une représentation sémantique des images. Le principe consiste à mettre en correspondance des attributs extraits sur les entités détectés dans la scène par segmentation spatiale. Dès la fin des années 1970, Price [6] propose une méthode de détection au niveau symbolique, arguant son potentiel d'adaptation à une plus grande classe d'images. Les changements sont détectés après appariement régional à partir d'un ensemble d'attributs (luminance moyenne, taille en pixels, compacité, etc.), extraits après segmentation de type division de région. D'autres méthodes plus récentes sont issues de l'analyse du mouvement par segmentation spatio-temporelle. Kottke [7] propose une mise en correspondance de régions segmentées par un algorithme dérivé des K-means. Benois [8] s'affranchit du problème d'appariement en fusionnant des régions spatiales temporellement homogènes. Les méthodes symboliques permettent de prendre en compte le contexte spatial des pixels et limitent ainsi le risque d'occlusion des régions changeantes détectées. Néanmoins elles posent le problème du choix d'une méthode de segmentation suffisamment stable pour que la description symbolique de chaque image soit indépendante des fluctuations naturelles de la scène. Au risque d'erreur dans la comparaison (variabilité des attributs), s'ajoute le risque d'erreur dans la représentation.

Nous proposons dans cet article un nouvel opérateur pour la détection des changements, fondé sur l'analyse corrélative d'attributs spatio-fréquentiels extraits de spectres locaux.

2 Description de l'opérateur

La comparaison temporelle doit bénéficier d'attributs texturaux caractérisant la nature et l'organisation d'un groupe de pixels. Parmi les méthodes de caractérisation de textures, nous avons choisi l'analyse spectrale locale qui présente l'avantage de prendre en compte à la fois des caractéristiques spatiales et fréquentielles de l'image [9-10]. La transformation orthogonale retenue est la transformation bidimensionnelle discrète de Walsh-Hadamard. Elle est réelle et à faible complexité algorithmique. D'autre part l'ordonnement de ses vecteurs de base permet d'accéder à une information de séquence spatiale et de bénéficier ainsi des propriétés angulo-fréquentielles des transformations fondées sur des fonctions harmoniques [11-12].

La détection des changements à une échelle d'analyse N comprend trois étapes successives : l'extraction, l'analyse corrélative et le seuillage des attributs de dissemblance.

2.1 Extraction des attributs de dissemblance

On considère sur chaque image des blocs $b(m,n)$, de dimensions $N \times N$ ($N=2^p$, p entier), transformés sans recou-

vrement par W_N , noyau de Walsh-Hadamard d'ordre $\log_2(N)$. On obtient ainsi les blocs transformés $B(k,l)$ exprimés comme suit :

$$B(k,l) = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} b(m,n) W_N(k,m) W_N(l,n)$$

Le spectre énergétique associé $S(k,l) = |B(k,l)|^2$ est calculé ainsi que l'énergie totale du bloc $E = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} S(k,l)$.

La dissemblance entre un bloc $b(m,n,t_1)$ et son correspondant $b(m,n,t_2)$ est évaluée au moyen de trois attributs notés EQMI, DET et DAS.

L'Ecart Quadratique Moyen des Intensités (EQMI) :

$$EQMI(l_b, c_b) = \sqrt{\frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} [b(m,n,t_1) - b(m,n,t_2)]^2}$$

(où l_b et c_b sont les indices spatiaux du bloc dans la scène) Cet attribut est préféré à la Différence des Intensités Moyennes ($DIM(l_b, c_b) = |B(0,0,t_1) - B(0,0,t_2)|$) car il est plus sensible aux fortes variations locales d'intensité.

La Différence d'Ecart-Type (DET) traduit la variation d'homogénéité statistique entre les deux blocs et s'exprime comme suit :

$$DET(l_b, c_b) = \left| \sqrt{E_1 - S(0,0,t_1)} - \sqrt{E_2 - S(0,0,t_2)} \right|$$

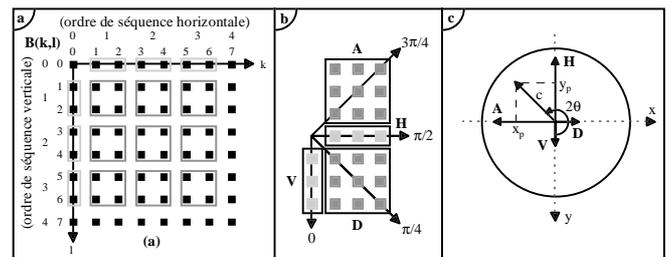


fig.2 : Création de filtres directionnels et calcul de l'orientation principale

Le troisième attribut est prélevé dans l'espace transformé des deux blocs. Les coordonnées (x_b, y_b) du barycentre du spectre d'énergie de chaque bloc et son correspondant sont calculées au moyen des expressions suivantes:

$$x_b = \frac{1}{E} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} k \times S(k,l) \quad y_b = \frac{1}{E} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} l \times S(k,l)$$

L'écart entre les logarithmes des distances ρ des barycentres énergétiques à l'origine du spectre fournit une mesure de la variation de périodicité des structures des 2 blocs. Néanmoins il ne mesure pas la variation d'orientation de ces structures qui constitue une information de dissemblance supplémentaire. Cependant, la recombinaison des composantes spectrales de Walsh selon la notion de séquence dans l'espace de Fourier indique que l'on peut procéder à une analyse angulo-fréquentielle de signaux 2D présentant une direction principale d'angle $\theta_p \in [0, \pi/2]$ (fig.2a). Des études ont montré qu'il était possible de lever l'ambiguïté directionnelle sur certaines séquences, et de créer à partir des composantes de l'espace transformé un ensemble de filtres invariants à la phase [13] (fig.2b). Nous proposons d'utiliser les réponses de ces filtres pour calculer l'orientation énergétique principale $\theta_p \in [0, \pi[$. Les modules et les directions multipliés par deux des filtres définissent un ensemble de vecteurs d'orientations dont la sommation fournit l'orientation principale du bloc (fig.2c). Ce calcul tient compte de la complémentarité d'une direction ($\theta = \theta \pm \pi$) et,

contrairement à la méthode proposée dans [14], toutes les réponses des filtres directionnels sont prises en compte. On obtient ainsi le vecteur d'orientation énergétique principale ($c, 2\theta_p$), tel que :

$$c = \sqrt{x_p^2 + y_p^2} \quad \theta_p = \begin{cases} \pi - 0,5 \arccos(y_p / c) & \text{si } x_p < 0 \\ 0,5 \arccos(y_p / c) & \text{sinon} \end{cases}$$

La norme c fournit une mesure de certitude sur θ_p comprise entre 0 et 1 que l'on peut traduire sous forme d'un intervalle de confiance : $\pm \pi(1-c)/2$. Nous pouvons alors écrire la variation d'orientation $\Delta\theta$ entre deux blocs comme :

$$\Delta\theta = \Delta\theta_p - \pi(2-c_1-c_2)/2 \quad (\text{avec } 0 \leq \Delta\theta \leq \pi/2)$$

Les cas d'indéfinition de $\Delta\theta$ ($\Delta\theta < 0$) compromettent une utilisation indépendante de ce paramètre. Ce problème est résolu par la création d'un attribut de Différence Angulo-Séquentielle (DAS) qui exprime la différence vectorielle entre les signatures spectrales (ρ, θ_p) des deux blocs comme :

$$\text{DAS} = \sqrt{(\log_2 \rho_1)^2 + (\log_2 \rho_2)^2 - 2(\log_2 \rho_1)(\log_2 \rho_2) \cos(\Delta\theta)} \\ (\text{avec } \Delta\theta = 0 \text{ si } \Delta\theta < 0)$$

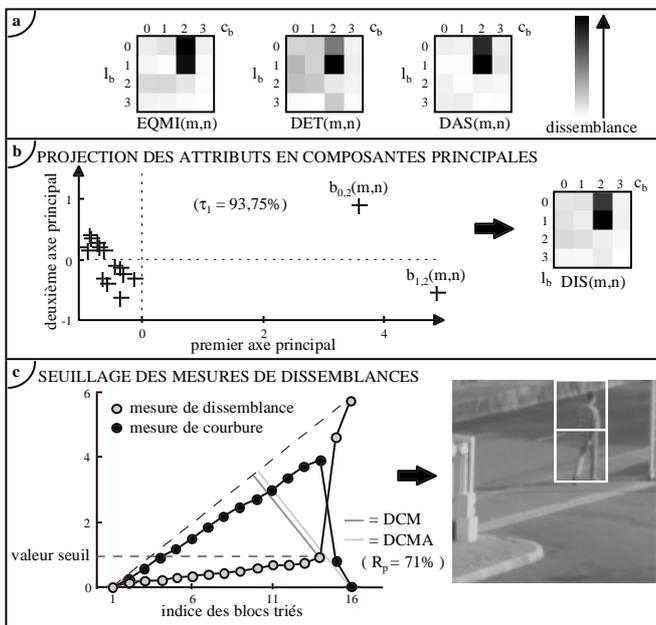


fig.3 : Détection des changements à l'échelle d'analyse N
Illustration des trois étapes nécessaires à la détection des changements pour l'échelle $N = 16$ sur la scène de la figure 1.

2.2 Analyse corrélative des attributs

A ce stade de traitement la scène multitemporelle de dimension $L \times C$ est représentée par 3 canaux d'attributs de dimensions $(L/N) \times (C/N)$ (fig.3a). Une analyse en composantes principales normées de ces canaux permet de synthétiser la dissemblance générale de chaque bloc en fonction de sa dissemblance relative sur chaque canal [15]. Le principe consiste à projeter le vecteur d'attribut de chaque bloc sur 3 axes dont les directions sont définies par les vecteurs propres associés aux valeurs propres de la matrice de corrélation des canaux. Les canaux étant corrélés, les composantes du premier vecteur propre sont de même signe, défini positif. La projection sur le premier axe principal P_1 (correspondant à la plus grande valeur propre) ordonne donc les blocs par ordre croissant des changements. La valeur de projection d'un bloc sur P_1 est d'autant plus importante que sa dissem-

blance est grande sur chacun des canaux (fig.3b). D'autre part, l'axe P_1 exprime le plus grand taux de variabilité τ des attributs ($\tau_i = \lambda_i / (\lambda_1 + \lambda_2 + \lambda_3) = \lambda_i / 3$). Le deuxième axe principal P_2 ne restitue pas la dynamique du changement ($P_2 \perp P_1$). Néanmoins un écart important par rapport à l'origine factorielle d'une valeur de projection sur P_2 est significatif d'une forte réponse du bloc concerné sur au moins un des canaux d'attributs. Ces considérations nous permettent d'unifier les 3 canaux sous la forme d'un seul attribut appelé « mesure de dissemblance » (DIS) qui s'exprime comme suit:

$$\text{DIS}(m, n) = \sqrt{(P_1(m, n) - \text{MIN}\{P_1\})^2 + P_2^2(m, n)}$$

Sur la fig.3b on constate que cet attribut assure une meilleure détection des blocs changeants dans la scène.

2.3 Seuillage des mesures de dissemblance

Le seuillage de $\text{DIS}(m, n)$ est effectué après étude de la distribution statistique des mesures de dissemblance [15]. Le principe consiste à créer un signal composé des mesures triées par ordre croissant puis à rechercher la mesure seuil correspondant à la Distance de Courbure Maximale (DCM). Cette méthode revient à analyser la fonction de répartition des mesures de dissemblance sur LC/N^2 intervalles. La validité du seuillage est évaluée par le rapport R_p de la DCM sur la Distance de Courbure Maximale Absolue (DCMA). Lorsque R_p est supérieur à une valeur seuil fixée (typiquement 50%) les changements sont jugés pertinents et les blocs détectés sont restitués (fig.3c).

3 Stratégies d'analyse multirésolution

3.1 Détection des changements avec *a priori*

Cette stratégie d'analyse est dédiée à la surveillance de scène. L'information *a priori* est l'échelle d'analyse de départ N_{\max} pour laquelle le capteur est en état de veille. N_{\max} correspond à la dimension caractéristique du type d'objet recherché compte tenu de la résolution spatiale de la scène. Les blocs changeants détectés à l'échelle N_{\max} sont analysés à l'échelle $N = N_{\max}/2$. Le processus d'« analyse-division » (fig.4) se poursuit tant que l'opérateur détecte des changements pertinents et que $N \geq N_{\min}$ (dimension minimale de bloc fixée). Cette approche permet une focalisation sur les objets d'intérêt, puis une optimisation de leur segmentation en travaillant à une résolution de plus en plus forte sur des régions de la scène de tailles de plus en plus réduites.

3.2 Détection des changements sans *a priori*

Cette stratégie d'analyse est adaptée au problème général de la réactualisation de scène. La détection débute à une échelle N_{\max} fixée. Lorsque tous les objets détectés à partir du niveau N_{\max} ont été analysés jusqu'au niveau N_{\min} , une nouvelle détection est lancée à partir de l'échelle d'analyse $N_{\max} = N_{\max}/2$, et seuls les nouveaux blocs changeants détectés sont considérés. La procédure se poursuit jusqu'à ce que $N_{\max} = N_{\min}$. Cette approche permet d'exhiber les changements de la scène de manière hiérarchique. D'autre part, l'analyse à une dimension $N/2$ des blocs rejetés à une di-

mension N peut permettre la récupération de composantes régionales et limiter encore le phénomène d'occlusion.

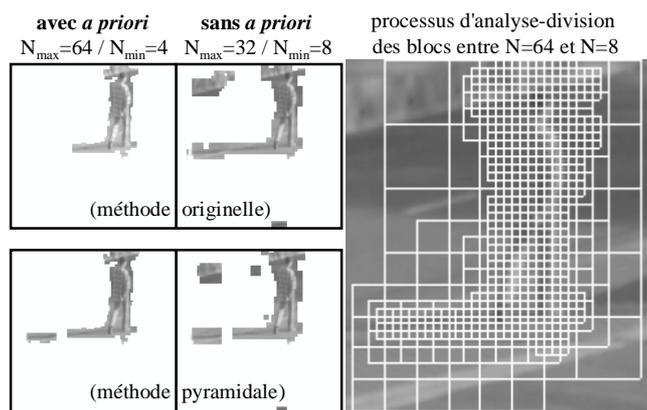


fig.4 : Résultats de détection pour les 4 configurations d'analyse

3.3 Détection pyramidale

Une autre approche consiste à déplacer une fenêtre d'analyse de taille fixe sur une sous-image de résolution croissante. Cette approche multirésolution par filtrage a été adaptée aux deux stratégies d'analyse : avec et sans *a priori* (fig.4-5). Compte tenu du calcul d'orientation sur $[0, 2\pi[$ la plus petite dimension fixe d'analyse est $N = 4$. Chaque image multitemporelle est décomposée en pyramide passe-bas gaussienne sur $[\log_2(N_{\max}) - 2]$ niveaux de résolution. La détection à l'échelle 4 commence sur le niveau de plus faible résolution. La diminution de l'échelle d'analyse se traduit à présent par le passage à l'image de niveau de résolution supérieur. La qualité de détection reste légèrement inférieure à celle de la méthode originelle, mais le coût en calculs est moindre, comme en témoignent les temps d'exécution en état de veille à la dimension $N_{\max} = 64$ (cf. tableau ci-dessous).

(scène de la fig.1)	méthode originelle	méthode pyramidale
avec <i>a priori</i> $N_{\max} = 64 / N_{\min} = 4$	21,37s	12,64s
sans <i>a priori</i> $N_{\max} = 64 / N_{\min} = 4$	28,72s	16,53s
$N_{\max} = 64 / N_{\min} = 64$	3,87s	0,43s

(programmation non optimisée sous MATLAB / PC Pentium 133MHz)

4 Conclusion

Nous avons présenté dans cet article un nouvel opérateur pour la détection des changements. Il est peu paramétré et se fonde sur l'analyse corrélative d'attributs spatio-fréquentiels. Les propriétés texturales de ces attributs garantissent l'atténuation du phénomène d'occlusion régionale jusqu'à de fines échelles d'analyse ($N > 2$). Cet opérateur a été implanté sous deux configurations d'analyse multirésolution dédiées aux applications spécifiques de réactualisation et de surveillance de scène. La diminution progressive de la dimension d'analyse permet la segmentation à forte résolution des régions changeantes détectées hiérarchiquement à plus faible résolution. L'approche proposée favorise l'extraction d'un objet physique changeant, spatialement hétérogène, en une seule région connexe. Le couplage avec une décomposition pyramidale passe-bas permet une optimisation calculatoire tout en approchant les performances de la méthode originelle.

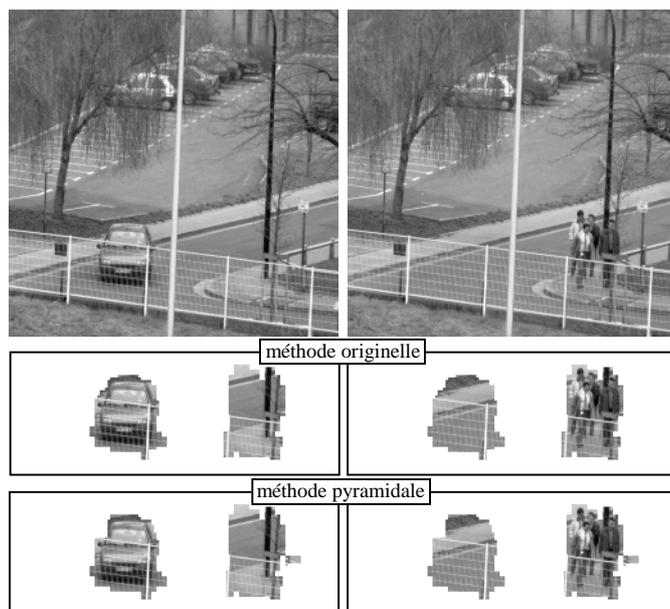


fig.5 : Détection des changements sur une scène CCD de 512x512 pixels. Les deux régions changeantes détectées avec *a priori* $N_{\max} = 128 / N_{\min} = 4$.

5 Références

- [1] Robert L. Lillestrand, "Techniques for Change Detection", IEEE trans. on Comp., vol.C-21, n°7, p.654 à 659, juil. 1972.
- [2] Ferdinand Bonn, Guy Rochon, "Précis de Télédétection", Presses de l'Université du Québec, vol.1 Principes et Méthodes, 1992.
- [3] Joseph G. Kawamura, "Automatic Recognition of Changes in Urban Development from Aerial Photographs", IEEE trans on SMC, vol.1, p.230 à 239, juil. 1971.
- [4] Hassan J. Eghbali, "K-S Test for Detecting Changes from Landsat Imagery Data", Proc. of the IEEE, p.17 à 23, 1979.
- [5] Y.Z. Hsu, H.-H. Nagel, G. Rekers, "New likelihood test methods for change detection in image sequences", CVGIP, vol.26, p.73 à 106, 1984.
- [6] Keith Price, Raj Reddy, "Change Detection and Analysis in Multispectral Images", 5^e congrès IJCAI, 1977, p.619 à 625.
- [7] Dane P. Kottke, Ying Sun, "Motion Estimation via Cluster Matching", IEEE trans on PAMI, vol.16, n°11, p.1128 à 1132, nov. 1994.
- [8] Jenny Benois, Dominique Barba, "Image segmentation by region-contour cooperation as a basis for efficient coding scheme", SPIE, vol.1818, p.1218 à 1229, 18-20 nov. 1992.
- [9] A. Baskurt, F. Peyrin, Y.M. Zhu, R. Goutte, "Analyse d'images par utilisation de spectres locaux", 13^e colloque GRETSI, p.1061 à 1064, Juans les Pins, 16-20 sept. 1991.
- [10] Y.M. Zhu, F. Peyrin, R. Goutte, M. Amiel, "Caractérisation de texture par l'analyse spectrale locale", 13^e colloque GRETSI, p.989 à 992, Juans les Pins, 16-20 sept. 1991.
- [11] Robert Jeansoulin, "Etude de la transformation de Walsh-Hadamard et applications au traitement numérique d'images", Thèse de l'université Paul-Sabatier de Toulouse, sept. 1976.
- [12] J.P. Crettez, "Contribution à la réduction des représentations d'images par des transformations orthogonales", Thèse de l'université Paris VI, déc. 1978.
- [13] J.P. Crettez, "Champs récepteurs et analyseurs de textures", 3^e congrès RFIA, p.383 à 394, Nancy, 16-18 sept. 1981.
- [14] J.P. Morichon, J.P. Dazelle, P. Larcher, "Classification directionnelle d'une empreinte digitale par les opérateurs d'Hadamard", ≥1989.
- [15] O. Desprez, "Analyse d'images multirésolution pour la réactualisation et la description hiérarchique de scène", Thèse de l'université Paris XII, à soutenir en 1997.

Remerciements

Ces travaux ont été effectués dans le cadre d'une Convention Industrielle de Formation par la Recherche (CIFRE) au sein du Groupe de Géographie Numérique et de Télédétection de la société THOMSON-CSF (Colombes, France).