

L'indexation des très grandes bases de données satellitaires

Henri Maître

GET / Télécom Paris – LTCI UMR 5141 CNRS - LTCI - Competence Centre on Information Extraction and Understanding for Earth Observation - 46 rue Barrault 75013 Paris

Henri.Maitre@Telecom-Paris.fr

Résumé : on présente un programme de recherche et une série de résultats initiaux pour indexer les très grandes bases de données issues des satellites de télédétection.

Mots clés : indexation, classification, détection, interprétation, reconnaissance des formes, bases de données, fouille, imagerie aérienne et satellitaire, télédétection

1 Introduction

Le traitement des images de télédétection, leur archivage, leur consultation ont pris, ces dernières années une importance exceptionnelle. A cela, plusieurs raisons :

- Tout d'abord la prolifération des satellites de télédétection. Pendant de nombreuses années, ces satellites étaient l'apanage d'une poignée de pays (URSS, USA, Canada, France, ...) à fort potentiel spatial. Ils se sont répandus maintenant à de nombreux autres pays : Inde, Brésil, Israël, Corée, Afrique du Sud, Chine. Par ailleurs, beaucoup de pays se sont regroupés pour opérer en commun de tels satellites.
- D'autre part l'amélioration des technologies d'imagerie, avec des capteurs à très forte intégration et des optiques à haute résolution permettant d'acquérir des images plus fines et plus grandes, avec des systèmes agiles permettant de réduire les survols sans intérêt (mers, déserts), au profit des zones importantes.
- Enfin l'amélioration des liaisons descendantes (du satellite vers la station de base), l'accroissement de l'archivage à bord et la multiplication progressive des stations de réception (réduisant les champs aveugles).

Corrélativement, les domaines d'application de l'imagerie de télédétection ont également considérablement crû : initialement concentrés sur les domaines privilégiés de la défense, de la météorologie et de l'agriculture, ils ont maintenant gagné d'importants marchés en aménagement du territoire (cartographie, étude d'impact, planification de grands travaux) ainsi qu'en surveillance des risques naturels éruptions volcaniques, glissements de terrain, inondations, incendies de forêts) et anthropiques (pollution marine, déboisement, croissance des mégalo-poles, etc.). L'irruption de l'image de télédétection dans les loisirs et le grand public (Google Earth, Géoportail IGN) ouvre une aire nouvelle de banalisation de l'image de télédétection qui devrait encore accélérer ce mouvement.

Les satellites de télédétection qui envoient des images vers la terre se comptent aujourd'hui par dizaine. Chaque satellite est capable de transmettre plusieurs fois par minutes des images de plus en plus grandes (SPOT 5 délivre des images de 25 000 par 25 000 pixels, QuickBird des images de 40 000 par 40 000 pixels). Chaque satellite délivre ainsi chaque année quelques téraoctets de données, qui, si elles ne sont pas toutes précieuses, méritent toutes d'être archivées pour des usages les plus variés. Traditionnellement, toutes ces images étaient systématiquement examinées par des experts qui déterminaient l'intérêt de chacune d'elles, en fonction de la couverture nuageuse et de la zone observée. Dûment annotées, ces images étaient ensuite archivées par des experts qui consignaient dans des bases de données les éléments indispensables pour les retrouver et les consulter. Cette tâche devient aujourd'hui hors des capacités des stations de réception et la nécessité de recourir à des techniques automatiques se fait nettement sentir. Néanmoins, supprimer l'opérateur humain dans l'analyse des images est une gageure que 40 ans de traitement d'image n'ont pas permis de résoudre. Il existe donc un important travail de recherche pour mettre en place les méthodes indispensables à cela. C'est l'objectif de cet exposé de montrer un ensemble de travaux qui jalonnent la piste de cette indexation automatique.

Les travaux qui sont présentés ici prennent place au sein du Cnes-Dlr-Enst CoC (Competence Centre in Image Analysis and Understanding for Earth Observation)¹, centre de recherche commun aux deux agences spatiales allemande et française et à l'Ecole des Télécommunications. Ce jeune centre de recherche a maintenant 2 ans et accueille une dizaine de chercheurs, surtout sur le site parisien, mais aussi à Munich et Toulouse.

¹ <http://www.coc.enst.fr>

2 Objectifs

2.1 Une recherche orientée vers l'utilisateur

L'objectif des recherches conduites est de mettre en place des techniques d'extraction de l'information des images satellitaires, permettant d'associer à chaque image des « résumés » très compactes qui serviront lors des recherches futures à diriger l'utilisateur vers les documents les plus pertinents pour sa quête. Quel utilisateur ? Là se situe l'un des enjeux importants du projet et qui le distingue notablement de ce que l'on est amené à développer dans d'autres problématiques des grandes bases de données d'images.

Le champ de la recherche dans les bases d'images peut être segmenté en deux grandes classes :

- les bases de données ouvertes, comme le web par exemple ou les archives personnelles d'un citoyen qui contiennent des documents de type très variés, d'origine souvent peu certaine et contenant des sujets à priori illimités,
- les bases de données thématiques (comme on en traite en imagerie médicale ou en astronomie) qui ne concernent qu'un champ précis de documents que l'on peut décrire pratiquement extensivement.

Il faut par ailleurs distinguer deux types de requêtes qui sont très différentes :

- les requêtes exactes (ou requêtes d'instances) qui visent à retrouver un objet précis (par exemple la Joconde ou Zinedine Zidane), dans des contextes variés ou sous des variations diverses ;
- les requêtes catégorielles qui visent à trouver des éléments, potentiellement très nombreux, d'une classe particulière d'objets (des voitures de sport, des scènes de fête populaire, etc.).

Ces deux types de requêtes conduisent aujourd'hui à des algorithmes de recherche assez différents et donc à des techniques d'indexation également différentes.

Dans le cadre de l'imagerie de télédétection, on est dans la situation d'une base de donnée thématique, et on s'intéresse aux deux types de requête, mais avec une très forte dominance des requêtes catégorielles. En effet, les recherches d'une situation particulière (telle ville, tel paysage) sont aujourd'hui bien traitées par les données ancillaires qui accompagnent toute image satellitaire et qui contiennent de façon précise les coordonnées (longitude et latitude) de la scène observée, la date de prise de vue, la résolution du capteur, permettant de résoudre simplement ces requêtes exactes sans faire appel à des structures de données complexes.

2.2 Une recherche très ouverte

Quelles sont donc les requêtes que nous anticipons ? Elles concernent des champs très variés d'application (agriculture, urbanisme, défense, prévision des risques naturels, surveillance de l'environnement, aménagement du territoire, écologie, etc.) et à ce titre ne sauraient être listées exhaustivement. Par ailleurs elles concernent les années à venir et nécessiteraient d'anticiper des usages encore très incertains. Des exemples permettent d'en cerner la forme :

- donner les forêts de chênes affectés d'une atteinte d'un parasite particulier ;
- retrouver les zones urbaines des mégapoles qui souffrent d'un habitat incontrôlé ;
- déterminer les configurations de terrains susceptibles d'accueillir un terrain d'aviation ou une station d'épuration d'eau ;
- proposer des tracés pour l'implantation de nouvelles lignes à haute tension (ou un gazoduc) desservant des zones éloignées des centre de production.
- localiser les zones de cultures maraîchères remplacées par des habitations ou des infrastructures urbaines.
- déterminer tous les ports de plaisance qui sont situés à proximité de plages et non loin d'un terrain de golf.

On voit à travers ces exemples que les requêtes prennent des formes très variées mais qu'elles réclament toutes une très lourde étape de traitement des images. Le traitement des images, au cours des quarante dernières années s'est beaucoup penché sur la résolution de problèmes semblables à ceux posés ci-dessus. Mais dans les approches adoptées alors, chaque problème était traité comme un objectif particulier du traitement et résultait de la spécialisation d'algorithmes choisis, paramétrés et combinés pour une recherche unique : algorithmes de détection, de filtrage, de classification, de reconnaissance. Nous souhaitons aujourd'hui prendre le problème de façon différente, afin que toutes ces questions puissent être abordées avec un jeu commun de primitives extraites une fois pour toutes et combinées dans des couches successives de représentation conduisant progressivement d'une représentation aussi fidèle que possible du signal (afin de pouvoir répondre à des questions non anticipées), jusqu'à une représentation riche en contenu sémantique (afin de pouvoir répondre à un interlocuteur humain s'adressant dans un langage très poche du langage naturel).

3 Les principes théoriques

3.1 La théorie de l'information appliquée à la fouille d'images

La démarche que nous allons suivre adopte un schéma proposé par M. Datcu (cf. Figure 1) qui considère l'image numérique satellitaire comme un signal à transmettre à un usager. L'approche naturelle est alors celle de la théorie de l'information qui consiste à extraire de l'image l'information pertinente contenue dans le message et de la coder de la façon la plus compacte pour en préserver le contenu tout en réduisant le volume. L'image initiale étant très volumineuse, il est illusoire d'arriver à des compressions très fortes (compatibles avec la consultation des très grosses bases de données) sans consentir une perte lors de la représentation. La théorie de l'information nous permet de choisir à chaque étape les solutions optimales et l'examen attentif du fonctionnement d'un expert de télédétection permet de fixer les compromis judicieux.

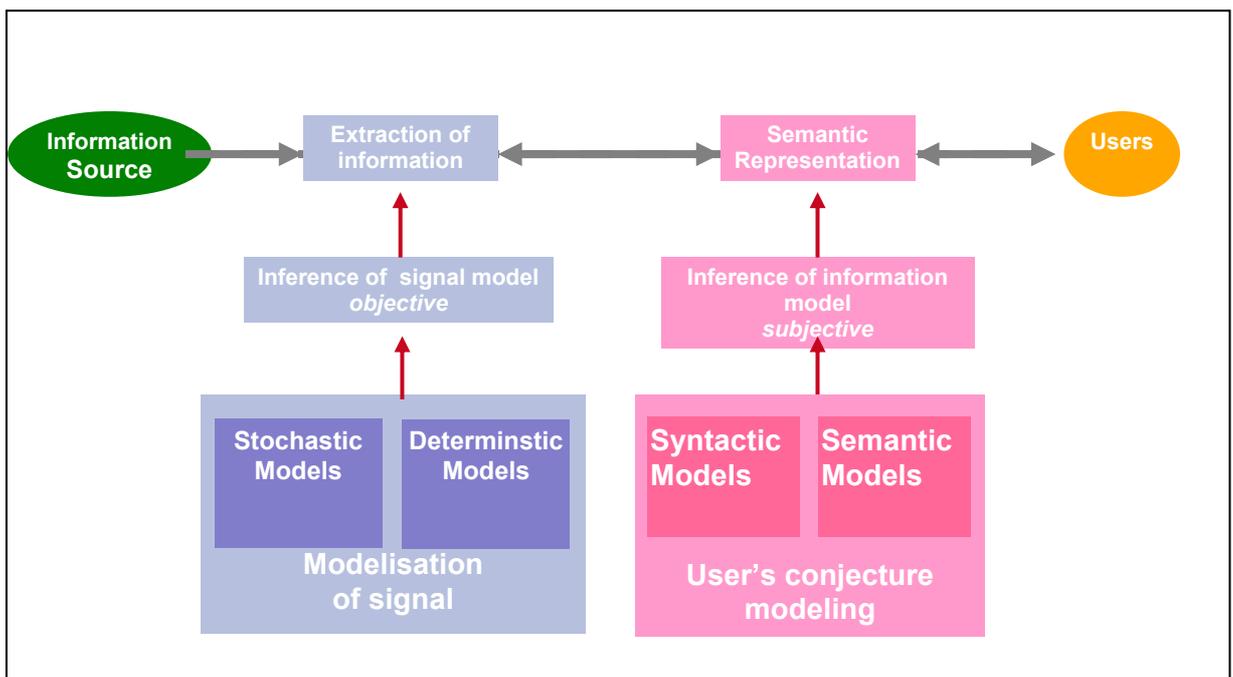


Figure 1 : l'image numérique, comme un signal en théorie de l'information (d'après M. Datcu). Dans le schéma ci-dessus, l'optimisation du canal se fait selon deux modes : une optimisation objective adaptée au contenu informationnel de l'image, une optimisation subjective, adaptée aux besoins de l'utilisateur.

3.2 Une description hiérarchique

Au premier niveau de représentation, nous choisissons ainsi d'extraire des primitives de bas niveau permettant de bien décrire les propriétés locales du signal. Pour des images à résolution de 3 à 10 m (comme les images SPOT par exemple), nous avons montré que les primitives de texture, prélevées régulièrement sur une maille assez serrée permettaient bien de rendre compte du contenu des images. Pour les images à résolution métrique et submétrique, (comme les images d'Ikonos, de QuickBird ou du futur capteur Pleiades), des primitives structurales semblent nécessaires, en complément des primitives de texture, pour rendre compte des zones à forte activité humaine.

3.2.1 Le choix des primitives

Les primitives de textures que nous avons sélectionnées par des comparaisons sur des bases de données sont de 3 types : des coefficients des décompositions en ondelettes de Gabor, des résultats de filtrage par des filtres miroirs en quadrature et des représentations réduites de Haralick issues de matrices de cooccurrence. Ces descriptions sont faites sur des fenêtres réduites (64 par 64 pixels, soit une cellule de 320 par 320 m par exemple sur des images de SPOT 5) Sur chaque cellule, un ensemble de 72 coefficients sont extraits. Par des techniques de sélection de primitives (méthode de Fisher par exemple), on en extrait un nombre réduit de coefficients (typiquement une vingtaine). C'est ce vecteur de primitives qui représentera au niveau le plus bas les images.

3.2.2 Le test des performances et le choix des paramètres

L'efficacité des descriptions par ces primitives est vérifiée sur de larges bases de données (annotées par des observateurs humains), qui permettent de qualifier leur capacité à discriminer des paysages différents. Un exemple de résultat est présenté sur le Tableau 1, pour un test sur 3 600 imageries thématiquement homogènes.

Table 1 – Tableau de confusion du classificateur. Les primitives choisies sont les 72 descripteurs de texture. La méthode de classification est un SVM à noyaux gaussiens avec apprentissage par validation croisée. La base de données est constituée de 3 600 imageries de 128 x 128 pixels issues d'images variées du satellite SPOT 5 à 5m de résolution. Les termes hors de la diagonale expriment le pourcentage d'erreur de reconnaissance

Vraie\Prédite (%)	ville	nuage	désert	champs	forêt	mer
ville	98.8	0	0	0.5	0	0
nuage	0	99.3	0.2	0	0	0
désert	0	0	99.0	0.3	0	0
champs	0.5	0.2	0.8	98.1	0.3	0.4

forêt	0	0.2	0	0	98.0	1.4
mer	0.7	0.3	0	1.0	1.7	98.2

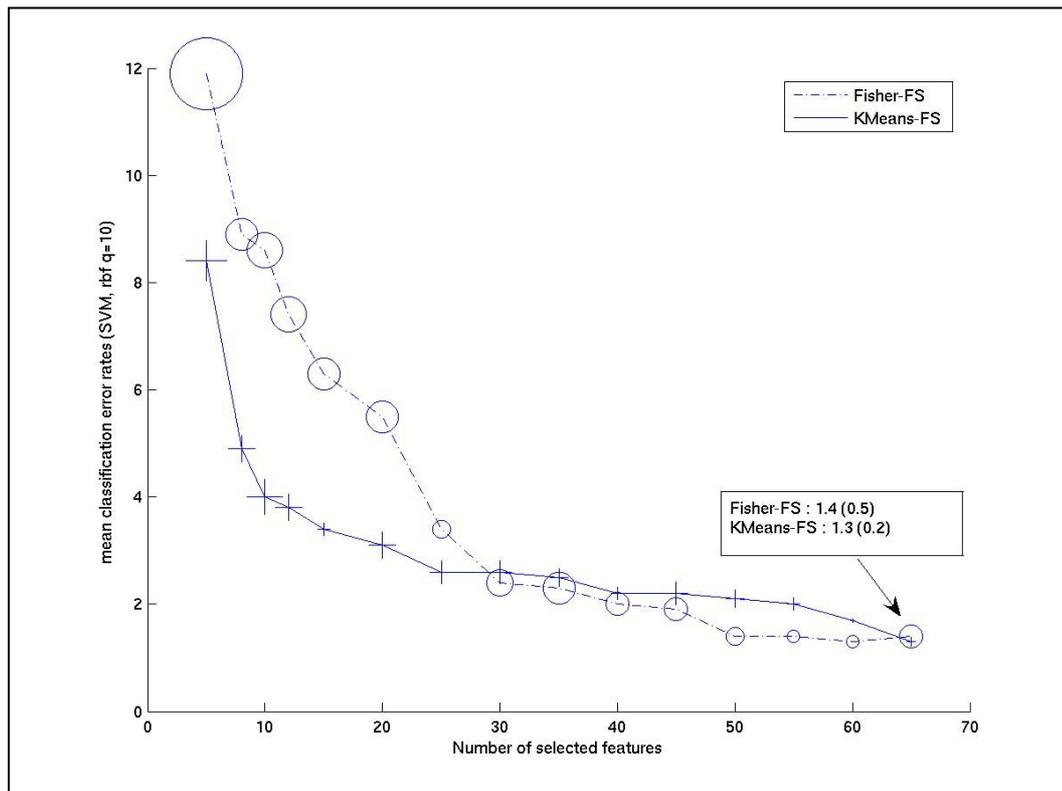


Figure 2 : sélection de primitives : comparaison de classificateurs. Deux méthodes sont comparées, la méthode de Fisher et les K-means. Partant de 70 primitives, on peut les réduire à une quinzaine sans dégrader considérablement les performances.

3.2.3 La réduction des primitives

C'est également avec des bases de données de ce type que l'on peut sélectionner les algorithmes les plus adaptés pour réduire le nombre de primitives de façon à obtenir des descriptions compactes des images (cf. Figure 2 et Tableaux 2 et 3).

Ces techniques de réduction de primitives s'appuient sur des principes clairs : elles visent une réduction du nombre de caractéristiques stockées tout en conservant (et parfois améliorant car on élimine des données redondantes qui interviennent comme du bruit) la performance de classification.

3.2.4 La classification

Juste au dessus de ce niveau élémentaire, les primitives sont regroupées par des classificateurs (cf. Tableau 4). Dans les cas simples, une classe est alors réduite à un nuage gaussien dans l'espace multidimensionnel des primitives. Dans les cas plus complexes, chaque classe est décomposée en mélanges de gaussiennes. A ce niveau, l'image est alors décrite par une collection d'étiquettes. Certaines correspondent à des entités connues portant un sens clair pour l'observateur humain. D'autres sont encore très ambiguës et ne prennent leur sens que par leur contexte. L'association spatiale d'étiquettes voisines permet de détecter dans l'image des zones à fort contenu sémantique. Cette démarche qui regroupe des primitives de niveau sémantique variable et des représentations spatiales de plus en plus étendues se décrit dans le schéma de la Figure 3 comme un cheminement dans le sens de la première diagonale.

Tableau 2 : les techniques de sélection de primitives supervisées et les principes de leur fonctionnement

	Type	Description
ReliefF	Filter	Score d'autant plus élevé que la caractéristique permet de discriminer les données de classes différentes.
Fisher	Wrapper	Analyse discriminante de Fisher.
RFE	Wrapper	Élimination récursive des caractéristiques de poids faible, à l'aide d'une SVM.
AROM	Wrapper	Approximation de la norme l0 des poids associés à chaque caractéristique, par une procédure récursive faisant intervenir une L1-SVM ou une SVM classique.

Tableau 3 : les techniques de sélection de primitives non supervisées et les principes de leur fonctionnement

	Type	Description
Schéma général (lourd)	Wrapper ou embedded (non implanté)	Recherche du sous-ensemble guidée par l'évaluation de la qualité d'une clusterisation des données. Utilisation d'algorithmes exploratoires (Greedy ou GA).
MIC	Filter	Utilisation d'une clusterisation KPPV des caractéristiques + choix d'un représentant par cluster.
kMeans-FS	Filter	Utilisation d'une clusterisation KMoyennes des caractéristiques +sélection de la caractéristique la plus proche du centroïde.
SVC-FS	Filter	Utilisation des vecteurs de support issus d'une classification 1-classe (clusterisation SVC)

Tableau 4 : les classificateurs « découvrent » des classes. Certaines sont attachées de façon biunivoque à des notions claires de l'interprétation de l'image (ici par exemple le cluster 1 qui correspond très bien à la ville). D'autres font apparaître des distinctions dans les classes proposées par l'expert, distinctions qu'il convient alors d'interpréter (c'est le cas de la classe « désert » qui se trouve distribuée dans le cluster5 et dans le cluster 8).

		Discovered clusters									
		cluster1	cluster2	cluster3	cluster4	cluster5	cluster6	cluster7	cluster8	cluster9	cluster10
CLASSÉS	Cloud	1	1	75	0	0	1	0	0	0	22
	Sea	0	30	0	69	1	0	0	0	0	0
	Desert	0	0	0	0	63	0	0	37	0	0
	City	99	0	0	0	0	0	0	0	1	0
	Forest	0	45	0	0	0	0	55	0	0	0
	Field	4	0	0	0	5	60	0	5	26	0

3.3 Interprétation de haut niveau

La description finale – sur laquelle nous travaillons encore beaucoup, consiste à retenir dans la couche supérieure de la description, les mots du langage naturel qui correspondent le mieux à la description de la scène. Ces mots devront appartenir à une structure logique partagée par les utilisateurs et permettant de faire les déductions logiques qui permettront de déduire les concepts absents des concepts utilisés lors de l’indexation. Une telle structure est connue sous le nom d’ontologie et donne un cadre général pour représenter la connaissance et échanger des morceaux de savoirs entre l’utilisateur, le concepteur et ... les images. Construire une ontologie pour l’imagerie satellitaire est aujourd’hui un challenge complexe pour notre projet.

Comment peut-on progressivement introduire de la connaissance experte dans le traitement de l’image ? Plusieurs pistes sont possibles, même si aucune n’a permis encore de franchir le « gap sémantique » qui nous sépare du résultat :

3.3.1 *Méthodes supervisés*

Elles peuvent être totalement supervisées et conduisent alors à une annotation manuelle assurée par un opérateur humain. Elles peuvent aussi être semi-supervisées, l’opérateur intervenant essentiellement pour donner au système de classification des « vérités » sur lesquelles s’effectuent les étapes de l’apprentissage. Elles peuvent enfin prendre la forme d’une boucle de pertinence dans laquelle l’utilisateur intervient pour modifier itérativement les classements d système

3.3.2 *Analyse du contexte des images*

Elles sont très utiles lorsque l’on dispose d’un riche contexte qui accompagne l’image : un titre, une légende, un texte d’accompagnement ou des données annexes comme c’est souvent le cas pour des données archivées sur un site web.

3.3.3 *Utilisation de bases de données annexes*

C’est un cas fréquent en imagerie satellitaire que l’on dispose d’informations très précieuses, archivées par ailleurs, dont on saura tirer profit pour « découvrir » la connaissance manquante. C’est le cas de :

- Corine Land Cover (apprentissage de classes et de catégories) base de données européenne (à moyenne résolution) ;
- Des cartes et des Systèmes d’Information Géographiques (SIG) qui fournissent un très large corpus de conventions de représentations et d’annotations ;
- Des bases de données qui collectent la plupart des grandeurs d’intérêt dans format spécifique.

3.3.4 Inférence de sémantique

Ce sont les techniques les plus complexes et qui n'ont pas encore fait toutes leurs preuves. Les plus classiques utilisent des techniques de modélisation bayésienne avec apprentissage supervisé. Les plus récentes s'inspirent des techniques de traitement des textes et mettent à profit des déductions de dépendances « latentes », enfouies dans des schémas de réalisation de type Dirichlet (modèle de Blei & Jordan). Enfin, des méthodes de raisonnement de type Intelligence Artificielle tirent profit de techniques déclaratives de représentation du savoir (soit selon le modèle classique des systèmes experts), soit dans le cadre plus nouveau de la déduction « ontologique ».

Enfin, spécificité du traitement de l'image, la possibilité de procéder à un véritable raisonnement spatial apparaît, qui introduit toutes les finesses du raisonnement humain sur la représentation de l'espace (notions d'imprécision inhérente au raisonnement subjectif, notion d'incertitude, de relativité, etc.)

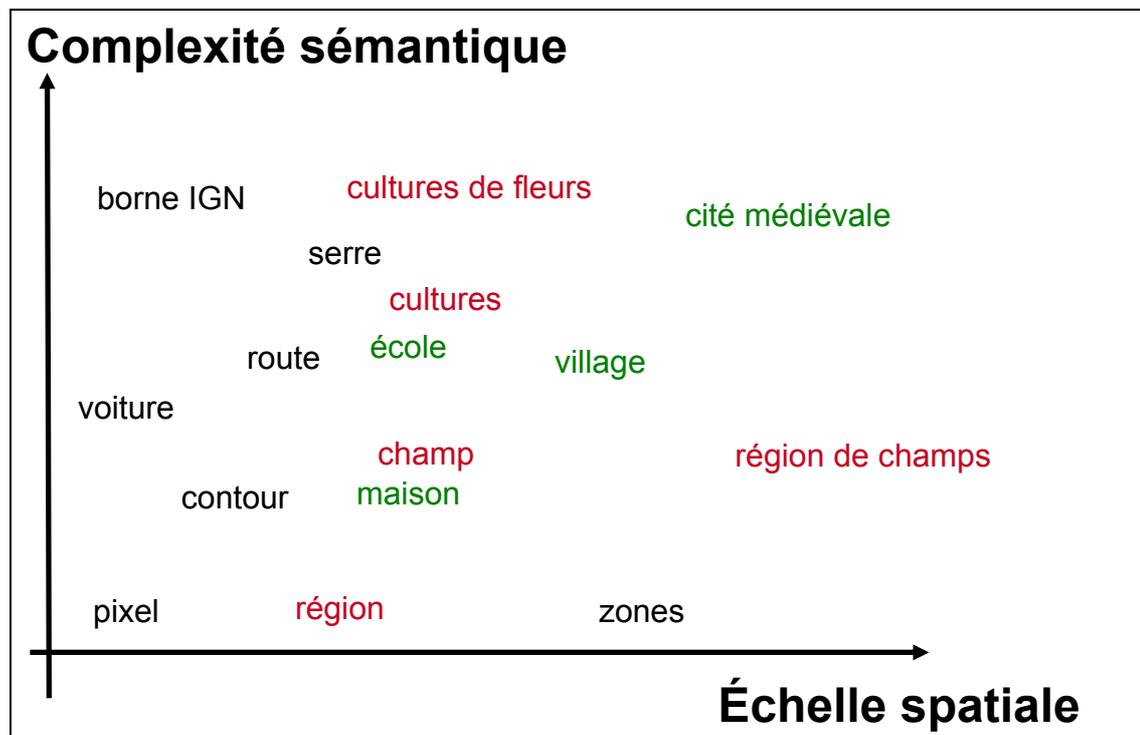


Figure 3 : L'analyse d'une image fait apparaître deux complexités de types différents : l'étendue spatiale, regroupant des objets complexes, introduit la première. L'interprétation sémantique associée à la reconnaissance des divers objets introduit la seconde. L'indexation de très grandes images devrait s'appuyer sur des descripteurs qui occupent le cadran supérieur droit, mais le traitement d'image ne sait bien faire que les traitements concernant le cadran inférieur gauche et – éventuellement – des propagations soit selon l'axe horizontal, soit selon l'axe vertical à partir de ce coin.

4 Conclusions

Le programme entrepris est ambitieux et loin d'être aujourd'hui abouti.

Plusieurs problèmes demeurent aujourd'hui sans solution. Le nombre de couches de la représentation hiérarchique est encore débattu. Le rôle et l'importance d'une ontologie est une question ouverte. Le rôle de l'interprète humain est très important mais sa place exacte n'est pas évidente, en particulier lorsque des critères théoriques d'optimisation (entropie, information mutuelle, complexité algorithmique ...) interviennent dans le schéma global. Par ailleurs, une expérimentation systématique des outils développés est indispensable, par des usagers extérieurs à leur définition et leur développement. Leur avis est primordial pour leur validation, mais cette étape n'a pas encore été entreprise.

Bibliographie

- J.B. Bordes et M. Roux. Detection of Roundabouts in satellite images. "ISPRS". Ankara (Turkey), février 2006.
- I. Bloch. "Fuzzy Spatial Relationships for Image Processing and Interpretation: A Review". *Image and Vision Computing*, Vol. 23, N. 2, pp. 89-110, février 2005.
- I. Bloch. "Spatial Reasoning under Imprecision using Fuzzy Set Theory, Formal Logics and Mathematical Morphology". *International Journal of Approximate Reasoning*, Vol. 41, N. 2, pp. 77-95, février 2006.
- M. Campedel, E. Moulines, H. Maître et M. Datcu. Feature Selection for Satellite Image Indexing. "ESA-EUSC: Image Information Mining". Frascati (Italy), octobre 2005.
- M. Campedel et E. Moulines. Modélisation de textures par sélection de caractéristiques. "CAp04". Montpellier, juin 2004.
- M. Campedel et E. Moulines. Classification et sélection automatique de caractéristiques de textures. "SFC04". Bordeaux, septembre 2004.
- M. Campedel et E. Moulines. "Classification et sélection automatique de caractéristiques de textures". *Revue des Nouvelles Technologies de l'Information (RNTI)*, Vol. C, N. 1, pp. 25-37, septembre 2004.
- M. Campedel et E. Moulines. "Classification et sélection de caractéristiques de textures". *Revue d'Intelligence Artificielle / RSTI (Hermès)*, Vol. 19, pp. 633-659, septembre 2005.
- M. Campedel et E. Moulines. Méthodologie de sélection de caractéristiques pour la classification d'images satellitaires. "CAP05". Nice - France, pp. 107-108, juin 2005.
- M. Campedel, E. Moulines et M. Datcu. Feature Selection for Satellite Image Indexing. "IGARSS'05". Séoul, Corée, juillet 2005.
- M. Ciucu et M. Datcu. Information Theoretical Approach for Searching Very Large Image Archives. "ESA EUSC Conference on Information Mining". Frascati (Italie), octobre 2005.
- M. Ciucu et M. Datcu. Incremental Grid Based Adaptive Vector Quantization for Indexing Very Large EO image archives. "ESA EUSC Conference on Information Mining". Frascati (Italie), octobre 2005.
- M. Costache, H. Maître et M. Datcu. Categorization based relevance feedback search engine for Earth observation images repositories. "IEEE IGARSS 2006". Denver (Colorado), juillet 2006.
- M. Costache, M. Lienou et M. Datcu. On Bayesian Inference, Maximum Entropy and Support Vector Machines Methods. "MaxEnt 2006". Paris, France, juillet 2006.
- M. Datcu, M. Soccorsi, D. Solimini et F. Del Frate. Rate Distortion Theory Method for EO Data Assessment and Model Selection. "ESA EUSC Conference on Information Mining". Frascati (Italy), octobre 2005.

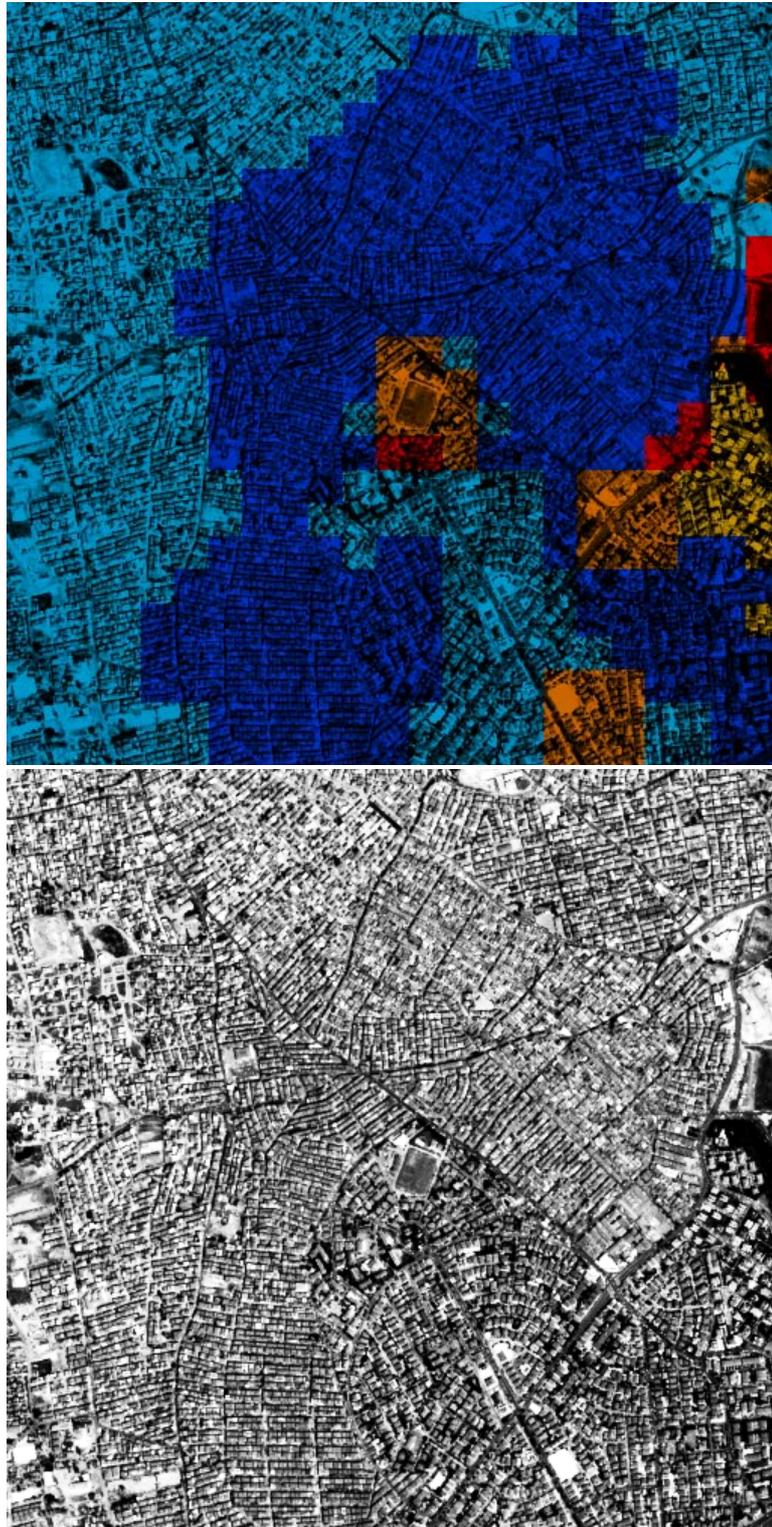


Figure 4 : La ville d'Istanbul telle que vue par SPOT 5 (Copyright CNES) et sa classification en fonction de l'organisation spatiale des quartiers (densité, orientation, présence de rues larges).

- M. Datcu, Seidel, K.; Walessa, M., Spatial information retrieval from remote-sensing images. I. Information theoretical perspective Geoscience and Remote Sensing, IEEE Transactions on , Volume: 36 Issue: 5 , Sept. 1998 Page(s): 1431 -1445
- M. Datcu, Melgani, F.; Piardi, A.; Serpico, S.B.; Multisource data classification with dependence trees Geoscience and Remote Sensing, IEEE Transactions on , Volume: 40 Issue: 3, March 2002 Page(s): 609 -617
- M. Datcu, Seidel, K.; Image information mining: exploration of image content in large archives Aerospace Conference Proceedings, 2000 IEEE , Volume: 3 , 18-25 March 2000 Page(s): 253 -264 vol.3
- R. Dehak, I. Bloch et H. Maître. "Spatial Reasoning with Incomplete Information about Relative Positioning". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, N. 9, pp. 1473 1484, septembre 2005
- A.Giros. Comparison of Partitions of Two Images for Satellite Image Time Series Segmentation. "IEEE-IGARSS". Denver (Colorado), juillet 2006.
- D. Gleich et M. Datcu. Wavelet based feature extraction for SAR data. "ESA EUSC Conference on Information Mining". Frascati (Italie), octobre 2005.
- L. Gueguen et M. Datcu. Spatio-temporal Textures Characterization Based on Information Bottleneck Principle. "ESA EUSC Conference on Information Mining". Frascati (Italie), octobre 2005.
- L. Gueguen, C. Le Men et M. Datcu. Analysis of Satellite Image Time Series based on Information Bottleneck. "MaxEnt 2006". Paris, pp. 367-374, juillet 2006.
- L. Gueguen et M.Datcu. Spatio-temporal structures characterization based on multi-information bottleneck. "ESA-EUSC 2006: Image Information Mining for Security and Intelligence". Madrid, novembre 2006.
- L. Gueguen et M. Datcu. The Model Based Similarity Metric. "DCC". Snowbird, Utah, USA, pp. 382, mars 2007.
- P. Heas et M. Datcu. "Modelling trajectory of dynamic clusters in image time-series for spatio-temporal reasoning". *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, N. 7, pp. 1635- 1647, novembre 2005.
- C. Iorga et M. Datcu. Survey on Methods for Feature Extraction of SAR data, ESA EUSC Conference on Information Mining. "ESA EUSC Conference on Information Mining". Frascati (Italie), octobre 2005.
- F. Kurz et M. Datcu. On the problematic of data integrity for communication of large EO Repositories and Information Mining Systems. "ESA EUSC Conference on Information Mining". Frascati (Italie), octobre 2005.
- I. Kyrgyzov, H. Maître et M. Campedel. Combining Clustering Results for the Analysis of Textures of SPOT5 Images. "ESA-EUSC: Image Information Mining". Frascati (Italy), octobre 2005.
- M. Lienou, H. Maître et M. Datcu. Is it possible to automatically produce a CORINE Land Cover map from a single SPOT image? "ESA EUSC". Madrid, Espagne, novembre 2006.
- B. Luo, J.F. Aujol, Y. Gousseau, S. Ladjal et H. Maître. Characteristic scale in satellite images. "IEEE ICASSP-06". Toulouse, France, Vol. 2, pp. II 809- II 812, mai 2006.
- G. Piella, M. Campedel et Béatrice Pesquet-Popescu. Adaptive wavelets for image representation and classification. "EUSIPCO'05". septembre 2005.
- M. Schroder, M.; Rehrauer, H.; Seidel, K.; Datcu, M.; Spatial information retrieval from remote-sensing images II. Gibbs-Markov random fields Geoscience and Remote Sensing, IEEE Transactions on , Volume: 36 Issue: 5 , Sept. 1998 Page(s): 1446 -1455
- H. Tang, H. Maître et N. Boujemaa. Similarity measure for satellite images with heterogeneous contents. "URBAN - 2007 - IEEE/ISPRS Joint Workshop on Remote Sensing and data fusion over urban areas". Paris (France), avril 2007.
- J.C. Tilton, Marchisio, G.; Koperski, K.; Datcu, M.; Image information mining utilizing hierarchical segmentation Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International , Volume: 2 , 24-28 June 2002 Page(s): 1029 -1031 vol.2

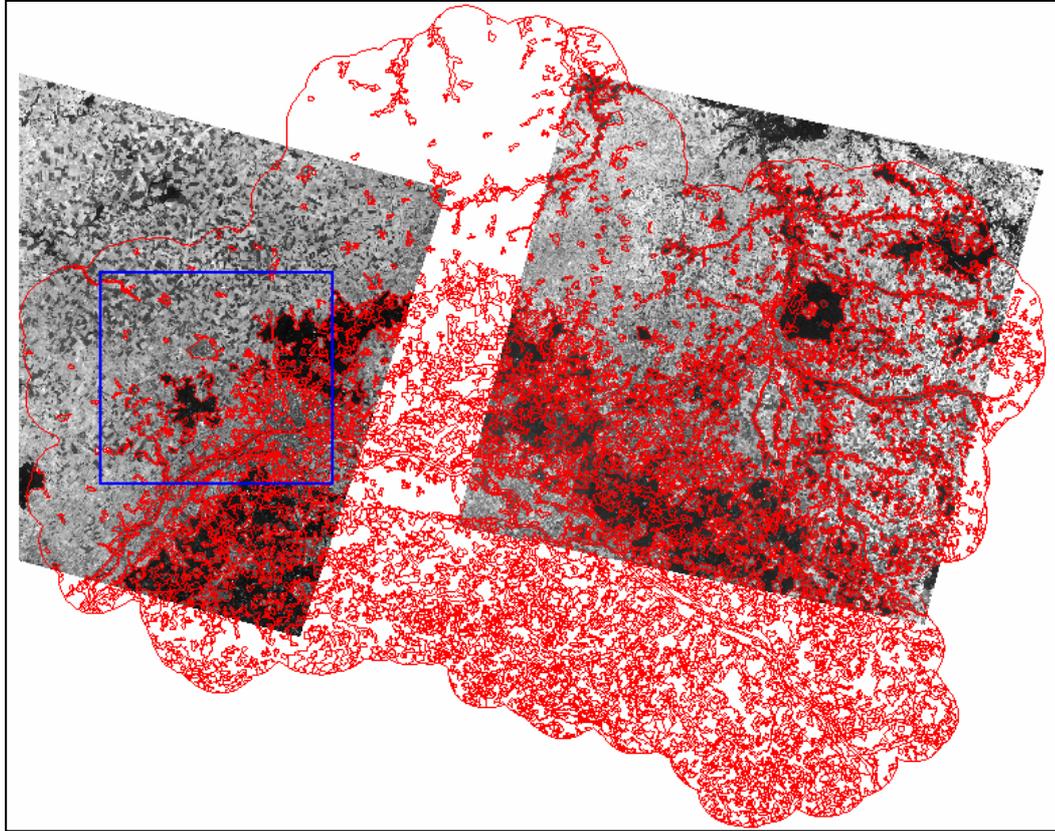


Figure 5 : superposition d'une carte Corine Land Cover de recensement des couvertures du sol et de 2 scènes SPOT5 (région du Loiret Copyright CNES)