

GENII, une plate-forme pour le développement d'applications d'interprétations d'image

H. ALLANIC¹, E. PETIT¹, F. LOPES¹

¹L.E.R.I.S.S., Université PARIS XII-Val de Marne, 61 Av. du Général de Gaule, 94010 CRETEIL cedex

allanic@univ-paris12.fr

Résumé – Le système d'interprétation présenté ici trouvera son utilité si l'on désire manipuler des informations structurelles et spatiales issues d'images 2D. En effet, ce type d'information se manipule plus facilement dans le domaine symbolique, sous la forme de relations entre nœuds, arcs et faces, que dans le domaine matriciel où il est difficile d'ajouter des informations structurelles aux pixels.

Abstract – The image understanding system presented here can be used to process structural and spatial data on 2D-pictures. This kind of data is better processed with symbols like nodes, edges and faces than with pixels on which it is hard to add structural informations.

1. Introduction

La vision par ordinateur associe le traitement d'images et l'intelligence artificielle. Une des difficultés majeures à laquelle est confronté cette discipline est de traiter des données incomplètes. Une scène 3D doit par exemple être analysée à partir d'images issues de capteurs 2D. Certains auteurs considèrent donc que l'interprétation de scènes est un problème mal posé [1] et ont recours à des méthodes de régularisation pour déterminer une solution unique [2].

Une autre stratégie consiste à exploiter les images d'une même scène issues de plusieurs capteurs 2D et à opérer une fusion d'informations [3]. On peut également apporter des informations complémentaires au moyen d'une description simplifiée de la scène issue de bases de données contenant des informations numériques ou symboliques. Les travaux fondés sur la fusion d'informations symboliques [4] ont ainsi conduit à de bons résultats dans l'interprétation d'images aériennes à des fins cartographiques.

Certains auteurs se sont également penchés sur une approche multi-spécialistes de la reconnaissance ([5], [6]). Ainsi, des systèmes experts sont développés afin de focaliser la recherche sur un type particulier d'objets (un « spécialiste » par objet). Ces systèmes modulaires permettent de décentraliser le processus de reconnaissance globale de la scène qui est ainsi réparti sur plusieurs bases de connaissances indépendantes. Un autre groupe de spécialistes peut valider la reconnaissance au niveau sémantique en exploitant des connaissances contextuelles et relationnelles attachées aux objets. Une scène complexe peut ainsi être reconstruite, par détection et validation progressive d'objets simples.

En règle générale, la stratégie d'analyse de l'image doit pouvoir sectoriser les objets dans l'espace et inférer leurs positions afin de maintenir la cohérence de la reconnaissance à mesure que le processus d'interprétation se poursuit. On peut à cet effet, utiliser un réseau sémantique [7] ; la scène est alors considérée comme un ensemble d'entités spatiales

entre lesquelles sont définies des relations de hiérarchie, de topologie, de distance ou d'orientation.

Le système que nous présentons dans cet article est fondé sur une modélisation proche des réseaux sémantiques, appelée Modèle Vectoriel Topologique (MVT) et qui correspond à une norme dans le domaine de la cartographie numérique [8]. Nous décrivons les principales caractéristiques de ce modèle et son intérêt pour l'interprétation d'images et montrons comment l'utiliser pour déduire une stratégie d'analyse d'image à partir de la description d'un objet émanant d'un expert.

2. Le Modèle Vectoriel Topologique

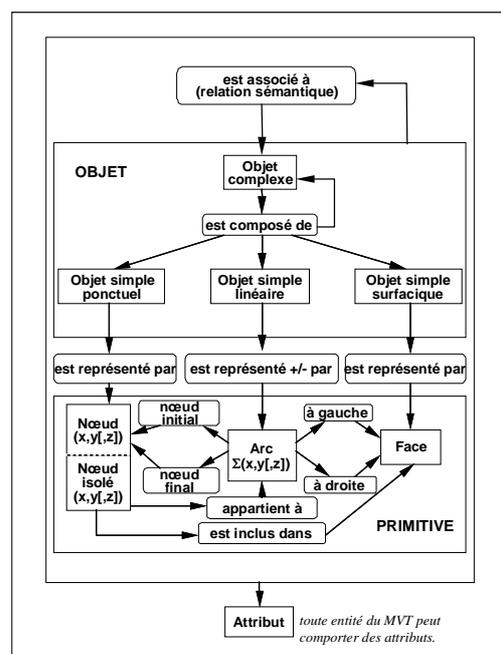


figure 1 : Modèle Conceptuel de Données « Vecteur Topologique »

Le **Modèle Vectoriel Topologique** (voir *figure 1*) est couramment utilisé en cartographie numérique [8]. Il permet de manipuler des informations sémantiques associées à des données géométriques. Il assure une cohérence formelle pour la composition ou la décomposition en éléments. Il permet de modéliser toute information géographique de la plus simple à la plus complexe.

Le MVT comporte deux types de données : les *Objets* qui donnent un sens au contenu de l'image (ville, route, ...) et les *Primitives* qui permettent d'en dessiner ce contenu (Nœuds, Arcs et Faces).

Ces données peuvent être liées entre elles par quatre types de relations :

- sémantiques (entre Objets) – « est associé à »,
- hiérarchiques (entre Objets) – « est composé de »,
- de représentations (Objets vers Primitives) – « est représenté par »,
- topologiques (entre Primitives) – voir *figure 1*.

3. Architecture proposée

L'architecture proposée ne correspond pas vraiment à un système d'interprétation d'images fonctionnel ; il s'agit plutôt d'une plate-forme de développement dont l'objectif est de faciliter le traitement des informations symboliques par leur intégration dans un réseau sémantique stocké de manière normalisé au moyen du M.V.T. Cet environnement fournit un noyau qu'on enrichit par des *Opérateurs* spécifiques de manière à construire l'application désirée modélisée sous la forme d'un *Gabarit*. L'idée est de fournir une architecture ouverte et facile à modifier puisqu'on a la maîtrise complète de l'ensemble des étapes d'interprétation de l'image.

3.1 Le Noyau

Comme illustré dans la *figure 2*, l'ensemble du système s'articule autour de 3 modules principaux (Stratégie, Synthèse et Extraction) qui manipulent trois formes de symboles : les Gabarits, les Extraits et les Instances).

Le module de *Stratégie* pilote le processus d'interprétation

complet en se fondant sur les *Gabarits* qui modélisent l'analyse à réaliser sur les données. Sa tâche principale est le lancement du module de *Synthèse* d'un symbole donné. Pour mener à bien cette synthèse, un ou plusieurs appels au module d'*Extraction* seront réalisés de manière à récupérer les informations pertinentes contenues dans les données à analyser (stockées sous formes d'*Instances*). Ces informations seront récupérées sous la forme d'*Extraits* qui seront ensuite analysés par le module de *Synthèse* dans le but de créer les *Instances* du symbole qui devait être interprété.

Ce découpage en trois modules est justifié par deux raisons. D'une part, le module de *Stratégie* permet de bien différencier l'ordonnancement des traitements par rapport à l'analyse des données. D'autre part, le traitement des informations (module d'*Extraction*), qui peut utiliser des algorithmes génériques, est séparé de la génération du résultat (module de *Synthèse*), qui doit par contre être spécifique à l'application réalisée.

3.2 Les Opérateurs

Ce sont les *Opérateurs* qui permettent de réaliser les traitements spécifiques à l'application d'interprétation désirée. Ces Opérateurs sont développés en se basant sur des classes C++ qui permettent l'interfaçage avec le noyau générique de la plate-forme. Il est aussi possible de dériver ces classes pour intégrer, dans le système d'interprétation, un ensemble de codes déjà existant. Il y a 5 catégories d'*Opérateurs* :

- L'*Opérateur de Stratégie* est utilisé pour ordonnancer les appels au module de synthèse afin de permettre l'analyse de l'ensemble des symboles à interpréter.
- C'est l'*Opérateur de Synthèse* qui permet de construire les *Instances* d'un symbole à partir de l'interprétation des *Extraits*.
- Ces *Extraits* sont générés par les *Opérateurs d'Extraction* qui exécutent généralement des algorithmes de traitements standards.
- les *Opérateurs d'Importation* servent à traduire les informations issues des capteurs en données symboliques, de bas niveau sémantique.

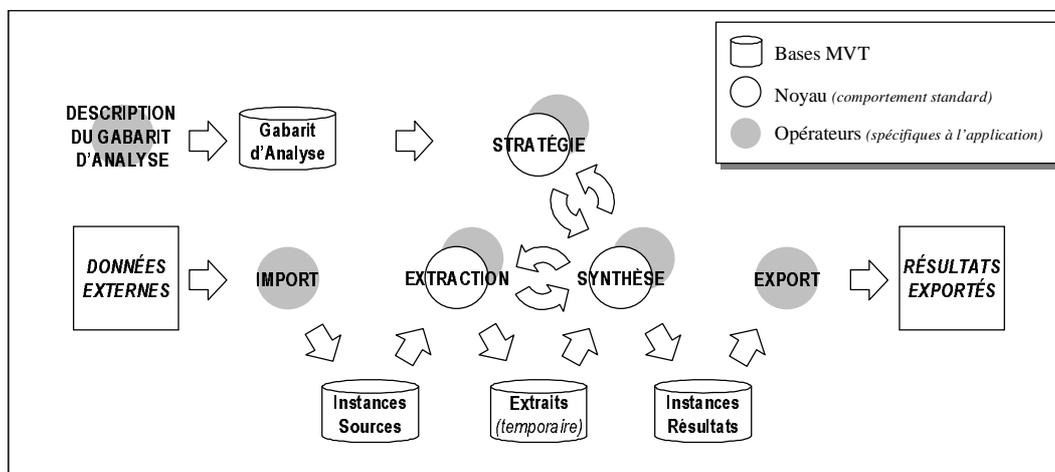


figure 2 : Architecture du système d'interprétation

- les *Opérateurs d'Exportation* servent à transformer les Instances en données ne respectant pas le format MVT (un exemple d'exportation est la construction d'une image bitmap permettant de visualiser sur écran le résultat d'une interprétation).

3.3 Les Gabarits

Le formalisme des Gabarits permet de décrire l'analyse à réaliser pour obtenir les instances du symbole à interpréter. Un Gabarit d'analyse se présente sous la forme d'un réseau sémantique où les symboles participant à l'interprétation sont reliés entre eux par les relations. Ce sont ces relations qui permettent de réaliser l'interprétation comme présenté dans la *figure 3*.

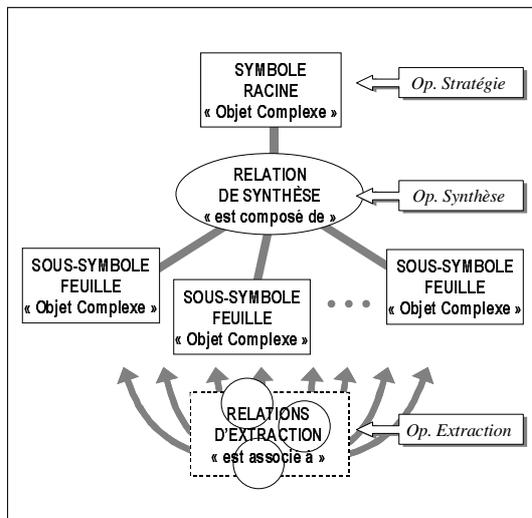


figure 3 : Synopsis d'un Gabarit d'Analyse.

Comme on peut le voir ci-dessus, le Gabarit d'Analyse utilise aussi les entités du MVT ; ce qui permet d'utiliser cette norme pour son stockage.

4. Un exemple illustratif

Précisons tout d'abord que le but de cet exemple n'est pas de concurrencer les systèmes d'interprétation déjà existants mais d'illustrer certaines caractéristiques intéressantes des applications que l'on peut générer à partir de la plate-forme proposée.

Nous avons validé notre approche sur un problème assez souvent repris dans la littérature : la reconnaissance de pistes d'aéroport sur des photographies aériennes ([9], [10]). Un des résultats majeurs de ce travail est de montrer que l'introduction d'une stratégie sémantique dans l'analyse d'une image permet d'aboutir à une interprétation robuste des objets modélisés bien que l'on se fonde sur une segmentation initiale de la scène très imprécise.

Pour réaliser cette interprétation, nous nous sommes basé sur un modèle assez simple : « sur une photographie aérienne, une piste d'atterrissage peut être décrite comme une zone rectangulaire de couleur claire et de longueur 1 à 2 km sur une largeur de 45 mètres environ ».

4.1 Importation du contenu d'une image

Par rapport au problème posé, les traitements choisis pour permettre l'importation vers le modèle M.V.T (*figure 5 – flèche n°1*) se divisent en deux étapes.

La première est d'effectuer une sur-segmentation en régions de la photographie de manière à ne pas perdre la forme générale de la piste et de donner aux pixels de chaque zone connexe obtenue la valeur de la radiométrie moyenne de cette zone.

L'importation effective correspond à la génération de deux types de symboles de bas niveau sémantique. D'un côté, les informations géométriques sont récupérées par la vectorisation des contours des zones connexes dont le résultat est une liste de symboles *Segment*. De l'autre, on crée une instance de *Région* pour chaque zone connexe de l'image segmentée et on y associe un attribut *Radiométrie* portant la valeur de la radiométrie moyenne sur cette région.

4.2 Élaboration des Gabarits

A partir du modèle présenté précédemment et sachant comment sont importées les données de l'image, on élabore 4 gabarits pour instancier les 4 symboles suivants :

- *Revêtement* – Ensemble de *Régions* ayant, en attribut, une radiométrie claire (*figure 5 – flèche n°2*),
- *Bordure droite* – Ensemble de *Segments* alignés sur 1 à 2 km (*figure 5 – flèche n°3*),
- *Rectangle* – Deux *Bordures* parallèles d'écart 45 mètres environ (*figure 5 – flèche n°4*),
- *Piste* – *Rectangle* recouvert à au moins 50% par du *Revêtement* (*figure 5 – les deux flèches n°5*).

Ces quatre gabarits peuvent être connectés entre eux pour résoudre le problème posé (voir *figure 4*).

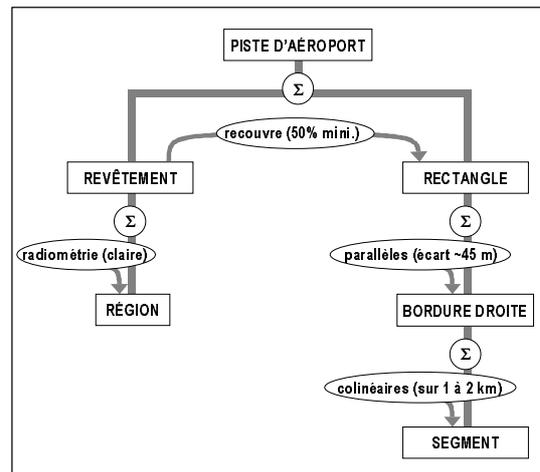


figure 4 : Gabarit complexe permettant d'interpréter une piste d'aéroport

4.3 Stratégie d'analyse guidée par le modèle

Le système Génii est mieux adapté à une interprétation guidée par le modèle. Ici, on sait dès le début du traitement que l'on cherche des pistes d'aéroport. L'implémentation de cette stratégie consiste simplement à effectuer un parcours

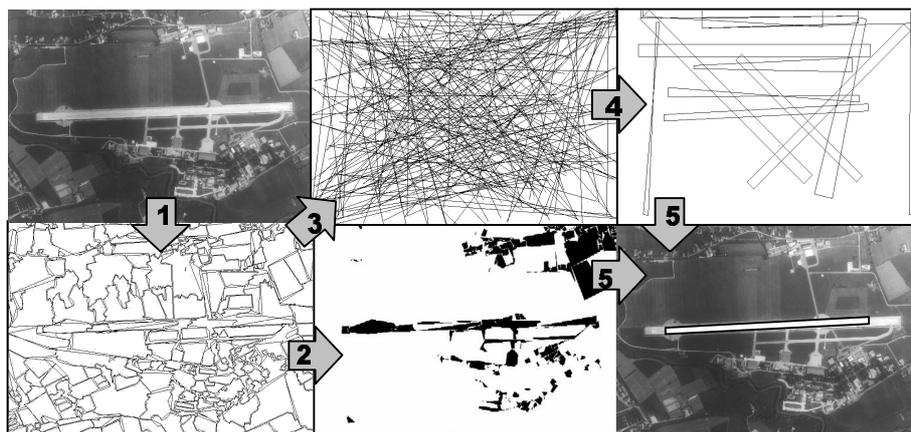


figure 5: exemple d'interprétation avec le gabarit « Piste d'Aéroport »

d'arbre en profondeur sur les gabarits élaborés. Ce parcours récursif est stoppé s'il existe des instances correspondant au symbole parcouru.

5. Conclusion et perspectives

Quand on analyse les résultats associés aux différentes étapes d'interprétation de l'exemple présenté *figure 5*, on constate trois problèmes :

- la segmentation et la vectorisation ont altérés la géométrie de la piste entraînant ainsi des contours très imprécis,
- l'information sur le revêtement (radiométrie claire) ne permet pas d'être certain que les instances de *Revêtement* appartiennent réellement à la piste,
- les opérateurs d'extraction utilisent des algorithmes trop simplistes qui ne donnent pas, individuellement, des résultats satisfaisants.

Dans cette application, on réussit à pallier ces problèmes grâce à deux possibilités offertes par la structure même des gabarits :

- la faiblesse d'un opérateur peut-être compensée par l'utilisation d'un autre opérateur à un niveau hiérarchique supérieur du gabarit complexe (exp. : l'opérateur « colinéaire » est enchaîné à l'opérateur « parallèle »),
- il est possible de réaliser des fusions de données pour augmenter la robustesse de l'interprétation (exp. : fusion des informations contours et régions par un opérateur de recouvrement pour obtenir les instances de « piste »).

Cette approche est prometteuse mais, pour la valider complètement, il faut étoffer la bibliothèque des opérateurs utilisables dans la plate-forme. Cette enrichissement passe en priorité par l'intégration d'un ensemble bien formalisé d'opérateurs travaillant sur les relations spatiales. En effet, c'est le domaine à privilégier de part la nature du Modèle Vectoriel Topologique ; de plus une étude à ce sujet a déjà été réalisée au sein du service à Thalès [11]. Un autre aspect à formaliser pour mieux l'intégrer dans la plate-forme est la prise en compte des imprécisions et des incertitudes.

Références

- [1] R. Horaud, New methods for matching 3D objects with single perspective views, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 9, pp. 401-412, 1987.
- [2] Alex P. Pentland, Perceptual organization and the representation of natural form. In *Readings in Computer Vision*, pages 680--699. Morgan Kaufmann Publishers Inc., 1987.
- [3] A. Michel, L. Benquet et P. Cornic, Détection d'objets complexes par fusion d'informations – Application à la détection d'aéroports dans les images satellitaires, 15^{ème} Colloque GRETSI sur le traitement du signal et des images, Juan-les-pins (France), 18-22 septembre 1995.
- [4] H. Moissinac, P. Guérin, I. Bloch and H. Maître, Image interpretation using symbolic data. Actes du 7^{ème} congrès « EUSIPCO » sur le traitement du signal, vol I, pp. 431-434, Edimburgh, 1994.
- [5] F. Sandakly, MESSIE-II : Contribution à la mise en œuvre d'une architecture à base de connaissances pour l'interprétation de scènes 2D et 3D. Thèse de doctorat, Université de Nice-Sophia Antipolis, 1995.
- [6] M. Thonnat, V. Clément, Pilotage de procédures de traitement d'images pour la description morphologique de galaxies, *Traitement du Signal*, volume 9 – n°5, 1991.
- [7] H. Niemann, G. Sagerer, S. Schröder and F. Kummert, ERNEST : A Semantic Network System for Pattern Understanding. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 12 (9) : 883-905, September 1990.
- [8] EDIGéO, Echanges de Données Informatisées dans le Domaine de l'Information Géographique, AFNOR, 1992.
- [9] D. M. McKeown, W. A. Harvey, et J. McDermott. « Rule-based Interpretation of Aerial Imagery », *PAMI*, vol. 7(N°5), pp. 570-585, September 1985.
- [10] L. Benquet, T. Quiguer, « Airport Recognition using Contextual Information », *EUSIPCO'94*, pp. 427-430, Sept 94.
- [11] V. EUDE, Modélisation spatio-temporelle floue pour la reconnaissance d'activités militaires, Thèse de l'université Paris 6 (LIP6), déc. 1998.