

RECALAGE IMAGE SPOT / CARTE ROUTIERE

Michel ROUX, Jaime LOPEZ-KRAHE, Henri MAÎTRE

TELECOM PARIS, Département IMAGES
46 rue Barrault, 75634 PARIS Cedex 13, FRANCE
tél : (33 1) 45 81 72 04, fax : (33 1) 45 81 37 94

RÉSUMÉ : cet article présente une méthode originale pour le recalage d'une carte et d'une image SPOT, basée sur la mise en correspondance de primitives de haut niveau : les zones urbaines et les carrefours. La mise en correspondance est réalisée suivant le principe de génération et propagation d'hypothèses : les zones urbaines, peu nombreuses dans les images, permettent d'engendrer des hypothèses de déformation entre les images, chaque hypothèse est ensuite propagée et évaluée grâce aux nouveaux appariements qu'elle a engendrés.

I. INTRODUCTION

Le recalage d'images multi-sources est une opération fondamentale en télédétection. En particulier, le recalage d'une image satellitaire sur une carte est la première étape de nombreuses applications : mise-à-jour de cartes, navigation assistée, création de bases de données cartographiques,...

Le recalage est généralement réalisé par un opérateur humain, et les méthodes automatiques qui ont été proposées ne permettent que le recalage d'images représentant approximativement la même scène [6,15]. La recherche d'une petite scène provenant d'une source dans une vue très large provenant de la seconde source est un problème sensiblement différent, pour lequel peu de solutions ont été proposées jusqu'à présent [2,9].

Cet article présente une méthode originale pour le recalage d'une carte et d'une image SPOT. Cette méthode est basée sur la mise en correspondance de primitives de haut niveau : les zones urbaines et les carrefours. Elle a été appliquée avec succès à la localisation de petites cartes (10km x 10km) dans une grande image SPOT panchromatique de 3000 x 3000 pixels (i.e. 30km x 30km). L'utilisation de structures de haut niveau, plutôt que des primitives élémentaires comme les segments, a permis de développer une méthode efficace de mise en correspondance basée sur le principe de génération et propagation d'hypothèses : les zones urbaines, peu nombreuses dans les images permettent d'engendrer des hypothèses de déformation entre les images, chaque hypothèse est ensuite propagée et évaluée grâce aux appariements qu'elle a engendrés entre carrefours et agglomérations des deux images.

Dans le chapitre 2 de cet article, sont présentées les diverses procédures qui permettent d'extraire les zones urbaines et les carrefours des images SPOT et des cartes numérisées. Le chapitre 3 décrit le processus de mise en correspondance des primitives. Les résultats obtenus sont présentés dans le chapitre 4.

II. DÉTECTION DES PRIMITIVES

II.1 Images SPOT

II.1.1 Choix des images

Pour des applications cartographiques, trois éléments sont primordiaux pour le choix des images SPOT, ce sont le mode d'analyse, l'angle de prise de vue et le niveau de prétraitement.

Mode d'analyse : la détection des routes dans les images SPOT nécessite la résolution la plus fine, laquelle est donnée par le mode panchromatique P avec 10m/pixel.

Angle de prise de vue : la visée verticale permet d'éviter les problèmes de perspective et assure la même taille pour chaque pixel.

Niveau de prétraitement : le niveau 1A, qui n'a pas été rééchantillonné, est le mieux adapté au traitement des images. Les niveaux 1B et 2A peuvent être utilisés pour effectuer un recalage automatique, mais l'obtention de ces images à partir de l'image de niveau 1A suppose la connaissance à priori des données ancillaires du satellite SPOT. Par ailleurs la dégradation du signal dû au rééchantillonnage rend les opérations de détection des primitives moins efficaces.

II.1.2 Les zones urbaines

Dans les images SPOT, les zones urbaines apparaissent comme des régions fortement texturées, de type "poivre et sel", qui s'opposent aux zones relativement homogènes correspondant principalement aux champs et aux forêts. La méthode d'extraction qui a été développée s'inspire des travaux de Serendero [19], basés sur les outils élémentaires de la morphologie mathématique [20].

Les pics clairs et sombres qui composent ces zones texturées sont mis en évidence par la différence entre la fermeture et l'ouverture morphologiques appliquées à l'image originale. Un filtre séquentiel alterné (i.e., une succession de fermetures et d'ouvertures de taille croissante) permet de regrouper ces pics dans des zones homogènes et d'éliminer les lignes isolées. Les zones urbaines sont obtenues ensuite par seuillage de l'image. Le résultat de cette procédure est une image binaire, où chaque objet représente une agglomération.

II.1.3 Les routes

Dans les images SPOT, le réseau routier apparaît principalement comme un réseau de lignes claires sur fond sombre de largeur généralement inférieure ou égale à 3 pixels. Contrairement aux images aériennes, où la largeur des routes, lorsqu'elle est supérieure à 3 ou 4 pixels, permet l'utilisation de méthodes basées sur la détection de contours *antiparallèles* [17] ou sur la détection de texture [13], la plupart des méthodes d'extraction de routes sur les images SPOT utilisent des détecteurs de lignes [4].

Parmi les différents opérateurs dédiés à l'extraction de réseau linéaire fin, nous avons choisi le *Duda Road Operator* (DRO). Cet opérateur a été associé à de la programmation dynamique pour effectuer du suivi de route sur des images aériennes de faible résolution [7] et sur des images SPOT [12].

II.2 Cartes numérisées

II.2.1 Numérisation

Diverses études ont montré l'adaptation de l'imagerie SPOT aux applications cartographiques à l'échelle du 1 : 100 000 dans les régions industrialisées, l'échelle du 1 : 50 000 pouvant être envisagée dans des régions faiblement urbanisées et au parcellaire grossier [3,11,14].

Pour notre application, nous avons numérisé des cartes IGN à l'échelle du 1 : 100 000 (série Verte), à l'aide d'une caméra Eikonix avec un taux d'échantillonnage de 50 pixel/cm, ce qui correspond à une résolution de 20m/pixel au sol. La numérisation a permis d'obtenir 2 fichiers images :

- une image de luminance codée sur 1 octet.
- une image en couleur codée sur 3 octets.

Les procédures d'extraction d'éléments cartographiques doivent nécessairement être adaptées à la norme de représentation cartographique utilisée, ainsi les méthodes présentées dans les deux sections suivantes ont été développées pour traiter des cartes IGN à l'échelle du 1 : 100 000, et devront être revues pour traiter des cartes élaborées avec d'autres conventions.

II.2.2 Les agglomérations

Sur ces cartes, les agglomérations ne sont pas représentées par un unique objet surfacique, mais chaque bâtiment ou chaque groupe de bâtiments est représenté par un bloc noir. La détection des agglomérations est réalisée par regroupement des blocs noirs proches les uns des autres. La principale difficulté lors de la détection est de séparer les agglomérations et les noms.

A partir d'une image binaire, obtenue par seuillage de l'image de luminance, une première opération consiste à éliminer les traits par une ouverture morphologique, ce qui assure la séparation des différents objets présents dans la carte : bâtiments, lettres et symboles.

Les noms des agglomérations sont ensuite repérés par détection des groupes d'objets alignés horizontalement, la méthode pouvant être étendue à la détection des mots ayant une orientation quelconque. Les objets restants sont regroupés suivant un critère de distance, chaque groupe étant considéré comme une agglomération. Ce processus de détection est illustré par la figure 9.

II.2.3 Les routes

Sur les cartes IGN, les routes sont représentées par deux traits parallèles, un code de couleur sert en outre à les distinguer suivant leur importance. Pour les cartes à l'échelle du 1 : 100 000, ce code est le suivant : autoroutes (rouge), routes principales (orange), routes secondaires (jaune), autres routes (blanc).

a. Les routes principales et secondaires.

Une analyse de la couleur permet dans une première étape d'extraire séparément le réseau rouge et orange et le réseau jaune. Une ouverture morphologique est nécessaire pour éliminer le bruit, en particulier les courbes de niveau qui sont également représentées en orange. Le réseau obtenu est peu bruité. Cependant la densité de ces routes principales dans certaines régions est trop faible pour se limiter à ce seul réseau.

b. Les autres routes.

La largeur d'une route sur les cartes est fixe le long de la route et dépend de l'importance de la route. La méthode d'extraction utilisée est inspirée de celle présentée par Nagao *et al.* [16]. En masquant dans l'image de la carte les réseaux colorés déjà extraits, le filtrage par un élément linéaire de longueur l pixels, dans les 4 directions (0° , 45° , 90° , 135°), permet d'extraire le réseau des routes de largeur l pixels.

Sur les cartes IGN à l'échelle 1 : 100 000 numérisées avec une résolution de 50 pixel/cm, la détection a été appliquée en cumulant les résultats obtenus pour $l = 1$ pixel et $l = 2$ pixels. Le réseau obtenu est très bruité, à cause notamment des symboles et des caractères présents dans la carte. Cependant les post-traitements (cf section II.3.2) effectués pour obtenir une représentation vectorielle du réseau (ébarbulation et approximation linéaire) vont permettre d'éliminer une part considérable de ce bruit.

II.3 Représentation vectorielle

Le passage à une information symbolique de haut niveau est facilité par le passage d'une représentation pixel à une représentation vectorielle des éléments détectés dans les images. Les procédures de vectorisation présentées dans ce chapitre ont été appliquées de la même façon aux données extraites de l'image SPOT et à celles de la carte numérisée.

II.3.1 Les agglomérations

Pour la vectorisation des zones urbaines, nous avons choisi de représenter chaque partie convexe d'une agglomération par un disque de même surface.

Une première étape consiste donc à décomposer en parties convexes les objets de l'image binaire représentant les agglomérations de l'une ou l'autre image. La morphologie mathématique permet d'effectuer cette décomposition structurelle : une succession d'érosions permet d'isoler chaque élément convexe, qui est ensuite reconstruit par l'opération inverse de dilatation [1].



Figure 1. Décomposition d'un objet en régions convexes

Ensuite chaque objet convexe ayant une surface supérieure à un seuil donné (équivalent à 200 pixels SPOT panchromatique) est représenté par une primitive *agglomération*. Cette primitive agglomération est définie par un disque dans l'espace associé à l'image traitée. Ce disque est centré sur le centre de gravité de l'objet représenté, et sa surface est égale à celle de l'objet.

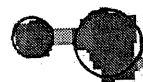


Figure 2. Représentation vectorielle des agglomérations

II.3.2 Les réseaux

La vectorisation consiste à représenter chaque réseau obtenu par un ensemble de segments de droite.

A partir de l'image binaire représentant le réseau extrait de l'image SPOT ou de la carte, une opération de squelettisation fournit un réseau ayant 1 pixel de large [5]. Une procédure de suivi de ligne permet ensuite d'isoler chaque chaîne de pixels et d'éliminer les *barbules* trop courtes [8]. L'approximation linéaire des chaînes a été réalisée suivant le principe de la corde [1]. Ce processus de vectorisation et l'efficacité des post-traitements pour éliminer une partie importante du bruit sont illustrés par la figure 10.

II.4 Les carrefours

La détection des carrefours est réalisée à partir de la représentation vectorielle des réseaux extraits de l'image SPOT et de la carte numérisée. Pour cela, 3 types d'objets sont déterminés successivement : les *chemins* (groupes de segments alignés), les *jonctions* (intersections de deux chemins), les *carrefours* (groupement perceptuel de jonctions). La suite de

ce chapitre va nous permettre de préciser les principes mis en place pour rechercher ces primitives dans les images.

II.4.1 Détection des chemins

Un chemin est défini comme un groupe de segments alignés deux à deux. Le critère d'alignement de deux segments s_1 et s_2 fait intervenir le segment intermédiaire s_i qui relie les deux extrémités les plus proches de s_1 et s_2 . Les segments s_1 et s_2 sont considérés comme alignés s'ils vérifient les propriétés suivantes :

- la longueur de s_i est inférieure à un seuil donné l_{max} .
- l'angle de s_i est compris entre les angles de s_1 et s_2 .

Les segments intermédiaires qui ont été créés sont des segments virtuels, qui seront également utilisés pour la détection des jonctions. Cependant ils ne doivent être en aucun cas considérés comme le prolongement des segments réels, leur localisation étant imprécise.

II.4.2 Détection des jonctions

Les jonctions correspondent aux intersections des chemins déterminés précédemment. Deux types de jonctions sont recherchés :

- jonction en X : intersection de deux segments, réels ou virtuels, provenant de deux chemins distincts.

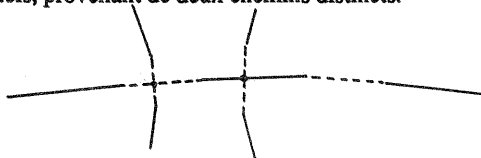


Figure 3. Détection des jonctions en X

- jonction en T : le segment réel correspondant à l'une des extrémités d'un chemin aboutit à un segment réel ou virtuel d'un autre chemin.

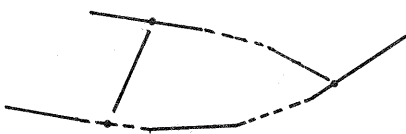


Figure 4. Détection des jonctions en T

Une jonction est définie par un point du plan, intersection des deux chemins, et par les segments réels qui interviennent lors de sa détection. S'il s'agit un segment virtuel, ce sont les deux segments réels qui l'entourent qui sont retenus.

II.4.3 Détection des carrefours

Un carrefour est un groupement perceptuel de jonctions. La détection des carrefours est déterminée par les trois règles suivantes :

C1 : un carrefour est réduit à une seule jonction s'il n'y a pas d'autre jonction suffisamment proche.

C2 : soient σ_i et σ_j deux jonctions,

$$\text{si } \text{dist}(\sigma_i, \sigma_j) \leq d_{max}$$

alors σ_i et σ_j appartiennent au même carrefour

C3 : soient σ_i , σ_j et σ_k trois jonctions,

si σ_i et σ_j appartiennent à un carrefour c_p

et si σ_j et σ_k appartiennent à un carrefour c_q

alors les carrefours c_p et c_q sont réunis.

La condition C2 est suffisante pour que deux jonctions appartiennent à un même carrefour, mais elle n'est pas nécessaire : grâce à la règle de transitivité (C3), deux jonctions peuvent appartenir à un carrefour sans vérifier la condition C2.

Formellement un carrefour est défini par un disque dans l'espace associé à l'image et par un ensemble de segments. Le disque est centré sur le centre de gravité des jonctions qui composent le carrefour, son rayon est égal à la distance maximale entre son centre et les jonctions, majoré d'une valeur δ_c fixe, rendant compte de l'incertitude sur la localisation des jonctions, et l'ensemble de segments est la réunion des ensembles associés à chacune des jonctions. La figure 5 donne un exemple de détection de carrefours.

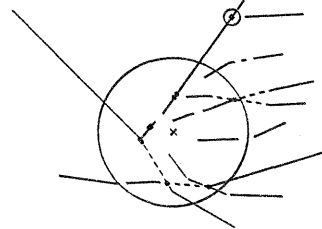


Figure 5. Détection de 2 carrefours l'un regroupant plusieurs jonctions, l'autre une seule.

Nous avons présenté dans ce chapitre les différentes procédures développées pour extraire les agglomérations et les carrefours des images SPOT et des cartes numérisées, et pour en donner une représentation vectorielle. Ces primitives sont à la base de la méthode de recalage des images, qui est décrite dans le chapitre suivant.

III. RECALAGE

La solution que nous proposons est basée sur la mise en correspondance des primitives par une technique de *génération et propagation d'hypothèses*.

III.1 Les Primitives

Pour la mise en correspondance, nous disposons, pour chacune des deux images, de trois ensembles de primitives : les segments de routes, les carrefours et les agglomérations. Nous précisons ici le formalisme utilisé pour décrire ces primitives sous forme vectorielle.

Les éléments de la carte seront décrits dans l'espace euclidien E , lié au repère Oxy . Ceux de l'image SPOT seront décrits dans un espace euclidien F , lié au repère $O'XY$.

- S_E : ensemble des segments de la carte.

$$S_E = \left\{ s_i = (p_i, p'_i), i = 1..N_S / p_i, p'_i \in E \right\}$$

pour chaque segment $s_i = (p_i, p'_i)$:

p_i, p'_i extrémités du segment

- C_E : ensemble des carrefours de la carte.

$$C_E = \left\{ c_i = (p_i, r_i, A_i), i = 1..N_C / p_i \in E, r_i \in \mathbb{R}^+, A_i \subset S_E \right\}$$

pour chaque carrefour $c_i = (p_i, r_i, A_i)$:

p_i centre du disque associé au carrefour

r_i rayon du disque associé au carrefour

A_i ensemble des segments composant le carrefour

- V_E : ensemble des agglomérations de la carte.

$$V_E = \left\{ v_i = (p_i, r_i), i = 1..N_V / p_i \in E, r_i \in \mathbb{R}^+ \right\}$$

pour chaque agglomération $v_i = (p_i, r_i)$:

p_i centre du disque associé à l'agglomération

r_i rayon du disque associé à l'agglomération

Les primitives de l'image SPOT ont une définition semblable à celle des primitives de la carte. S_F , C_F et V_F sont respectivement les ensembles des segments, des carrefours et des agglomérations détectés dans l'image SPOT.

III.2 Modèle de déformation

Le recours à une méthode par génération et propagation d'hypothèses pour déterminer automatiquement les appariements de primitives est lié à la possibilité de modéliser la déformation entre les deux images par une transformation paramétrée.

Le choix de ce modèle de déformation dépend de la qualité des amers. Dans le cas de points très précis désignés manuellement, il est possible d'utiliser des transformations localement déformables et d'interpoler la déformation entre les deux images à partir de ces points de contrôle. Cependant, pour une méthode automatique, le modèle de transformation doit permettre de rendre compte de la cohérence globale des amers obtenus. C'est pourquoi il est nécessaire d'utiliser des modèles rigides de transformations, qui seront calculés par une approximation sur les points des deux images mis en correspondance. La rigidité de la transformation peut servir à éliminer les amers ayant un résidu trop important. Des modèles de plus en plus souples pourraient être envisagés au fur et à mesure que la précision sur la localisation des amers augmente.

Nous avons retenu comme modèle de déformation entre une carte routière et une image SPOT panchromatique de niveau 1A une application affine du plan (i.e. transformation polynomiale de degré 1). Nous noterons Φ la famille des applications affines de E dans F , chaque transformations de Φ est déterminée par 6 paramètres :

$$\forall \phi \in \Phi, \forall p = (x, y) \in E, \quad \phi(p) = (X, Y) \in F$$

avec

$$\begin{aligned} X &= a_1x + a_2y + a_3 \\ Y &= b_1x + b_2y + b_3 \end{aligned} \quad (1)$$

Ce modèle ne reste valable que pour des petites zones, typiquement 10km x 10km, avec une faible dénivellation [14].

L'ensemble des transformations accessibles Φ est réduit grâce à la connaissance approximative des données de la prise de vue de l'image SPOT. En posant :

$$\begin{aligned} \text{tg}(\theta_1) &= \frac{b_1}{a_1} & \text{tg}(\theta_2) &= -\frac{a_2}{b_2} \\ k_1 &= \sqrt{a_1^2 + b_1^2} & k_2 &= \sqrt{a_2^2 + b_2^2} \end{aligned} \quad (2)$$

θ_1 et k_1 (respectivement θ_2 et k_2) représentent alors l'angle de rotation de l'axe Ox (respectivement Oy) lié à la carte et le coefficient d'homothétie le long de cet axe (voir figure 6).

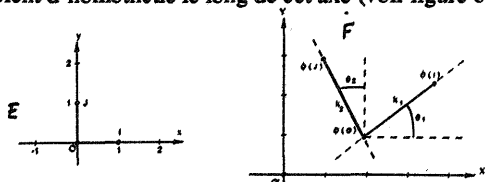


Figure 6. Transformation polynomiale de degré 1

Le rapport d'échelle approximatif k_{app} entre les deux images, calculé à partir du pas d'échantillonnage utilisé lors de la numérisation de la carte, et la différence d'orientation des prises de vues permettent de limiter la recherche aux transformations vérifiant :

$$\begin{aligned} 0.8 k_{app} &\leq k_i \leq 1.2 k_{app} \quad \text{pour } i = 1, 2 \\ \theta_{app} - 30^\circ &\leq \theta_i \leq \theta_{app} + 30^\circ \quad \text{pour } i = 1, 2 \end{aligned} \quad (3)$$

Ces contraintes serviront également à réduire le nombre des hypothèse propagées et d'éliminer des solutions aberrantes.

III.3 Mise en correspondance des primitives

III.3.1 Présentation du problème

Etant donné deux ensembles de primitives, l'un contenant les agglomérations et les carrefours détectés dans la carte ($P_E = V_E \cup C_E$), l'autre ceux détectés dans l'image SPOT ($P_F = V_F \cup C_F$), le problème est de déterminer l'ensemble des appariements d'une primitive de la carte avec une primitive de l'image SPOT, c'est-à-dire un sous-ensemble de $P_E \times P_F$ (implicitement cette formulation sous-entend qu'il est possible d'apparier un carrefour d'une image avec une agglomération de l'autre image).

Il est important de remarquer qu'il ne s'agit pas ici d'un simple problème de mise en correspondance de points comme pour d'autres problèmes de traitement des images (stéréo, constellation). Il ne s'agit pas en particulier de la recherche d'un isomorphisme entre sous-ensembles, une primitive d'une image pouvant très bien être appariée à plusieurs primitives de l'autre image.

Le principe de la mise en correspondance présentée dans cet article est basé uniquement sur la position relative des primitives et ne fait pas intervenir de relations sémantiques entre ces primitives. Par exemple, le fait que deux agglomérations de la carte sont reliées par une route impliquent que les primitives correspondantes de l'image SPOT doivent l'être également, aurait pu être pris en compte. Cependant, les problèmes de bruit et de parties cachées dans les images SPOT interdisent de telles considérations, qui pourront par contre être utilisées ultérieurement pour compléter l'extraction du réseau routier dans l'une ou l'autre image.

Dans ce chapitre, la définition d'un carrefour est réduite à un disque donné par son centre et son rayon. L'ensemble des segments associés à chaque carrefour n'est pas utilisé dans cette étape. Nous verrons plus loin (chapitre III.4), comment cette information complémentaire servira à affiner le recalage.

En suivant le formalisme présenté précédemment, ce problème de mise en correspondance se décompose en deux sous-problèmes :

- Pour toute partie A de $P_E \times P_F$, déterminer la meilleure transformation ϕ_A .
- Puis déterminer le meilleur ensemble A parmi toutes les parties de $P_E \times P_F$.

Le premier problème se résout algébriquement exactement si l'on choisit de prendre ϕ_A comme une transformation de Φ optimisée aux moindres carrés sur les couples de A . La transformation ϕ_A retenue minimise, pour toutes les fonctions ϕ appartenant à Φ , la moyenne des erreurs quadratiques sur les éléments de A .

En appelant $\delta_\phi(a)$ le résidu de la transformation ϕ pour l'amer a :

$$\forall \phi \in \Phi, \forall a = (c, d) \in P_E \times P_F, \quad \delta_\phi(a) = \text{dist}(\phi(c), d)$$

la transformation ϕ_A vérifie alors :

$$\sum_{a \in A} \delta_{\phi_A}(a)^2 = \min_{\phi \in \Phi} \sum_{a \in A} \delta_\phi(a)^2 \quad (4)$$

L'existence et l'inversibilité de ϕ_A sont déterminées par la composition des éléments de A . Soient E_A et F_A les ensembles des primitives qui appartiennent respectivement aux plans E et F et qui participent à A :

$$\begin{aligned} E_A &= \left\{ c \in P_E / \exists d \in P_F : (c, d) \in A \right\} \\ F_A &= \left\{ d \in P_F / \exists c \in P_E : (c, d) \in A \right\} \end{aligned}$$

Théorème 1 : il existe une application affine ϕ_A , approximation aux moindres carrés (AMC) sur les éléments de A , si et seulement si E_A contient au moins 3 points non-alignés.

Théorème 2 : l'application affine ϕ_A , AMC sur les éléments de A , si elle existe, est inversible si et seulement si F_A contient au moins 3 points non-alignés.

La recherche du meilleur ensemble A parmi toutes les parties de $P_{E \times P_F}$ est un problème combinatoire plus complexe, il peut être résolu en associant un coût à chaque partie de $P_{E \times P_F}$ et en recherchant la partie ayant le coût minimal.

III.3.2 Fonction de coût

Il s'agit ici de définir une fonction *coût* sur l'ensemble des parties de $P_{E \times P_F}$, qui soit minimale pour l'ensemble optimal recherché :

Soit A une partie de $P_{E \times P_F}$, vérifiant les conditions des théorèmes 1 et 2, nous pouvons lui associer une transformation ϕ_A , l'AMC sur les couples de A . La fonction *coût* que nous proposons est la somme de deux termes. Le premier rend compte des erreurs commises par la transformation ϕ_A pour chacun des éléments de A . Le second favorise les appariements en pénalisant d'un coût élémentaire les primitives de la *carte* n'ayant pas été appariées. Finalement, le coût associé à l'ensemble A prend la forme suivante :

$$\text{coût} (A) = \frac{\sum_{a \in A} \delta_{\phi_A}(a)^2 + \sum_{c \in E_A} r_c^2}{\text{card} (A)} \quad (5)$$

où \bar{E}_A est le complémentaire de E_A dans E .

La division par $\text{card} (A)$ permet les appariements multiples pour une primitive de la carte ou de l'image SPOT.

III.3.3 Génération et propagation d'hypothèses

Le nombre de sous-ensembles de $P_{E \times P_F}$ (2^{NM} , avec N et M variant de 50 à 100 pour des scènes de 10km x 10km, cf tableau 1) n'autorise pas le calcul de chacun d'entre eux, une méthode par génération et propagation d'hypothèse permet de réduire la combinatoire du problème.

Un amer étant défini comme l'appariement d'un point de la carte et d'un point de l'image SPOT, à chaque couple de primitives appariées est associé un amer défini par les centres des disques de chaque primitives, que ce soit une agglomération ou un carrefour.

Le principe de la méthode de génération et propagation d'hypothèse est le suivant :

Génération : deux amers quelconques n'ayant pas de point commun fournissent une transformation initiale ϕ_0 , qui est la *similitude* unique définie par ces deux amers (une similitude est la composée d'une rotation, d'une homothétie et d'une translation, i.e. $a_1 = b_2$ et $a_2 = -b_1$ dans les équations (1)).

Propagation : cette transformation ϕ_0 va permettre de déterminer de nouveaux appariements, en acceptant les couples de primitives (c, d) tels que la distance du point $\phi_0(c)$ au point d soit inférieure à un seuil donné. Avec ces amers une nouvelle transformation ϕ_1 sera calculée. ϕ_1 est l'*application affine* définie par une AMC sur les amers. Cette procédure est itérée jusqu'à la convergence de l'ensemble des amers et de la transformation associée.

Evaluation : le coût de l'ensemble final d'amers est calculé.

Toutes les hypothèses possibles étant engendrées et propagées, celle conduisant à l'ensemble d'appariements ayant le coût le plus faible est conservée.

Remarque : il est possible, dans notre application, d'initialiser le processus de propagation par une similitude, ce qui ne nécessite que 2 amers, car la transformation affine finale attendue entre l'image SPOT et la carte est proche d'une similitude. C'est-à-dire que les angles de rotations θ_1 et θ_2 sont peu différents, de même pour les coefficients d'homothétie k_1 et k_2 (voir figure 6), typiquement :

$$| \theta_1 - \theta_2 | \leq 10^\circ \quad \text{et} \quad \frac{\max (k_1 , k_2)}{\min (k_1 , k_2)} \leq 1.1 \quad (6)$$

Ce processus de génération et de propagation d'hypothèses étant défini, il convient de préciser quelles sont les primitives qui seront utilisées pour engendrer, propager et évaluer les hypothèses.

III.3.4 Choix d'une stratégie

L'optimisation du processus de mise en correspondance, tant au niveau de la stabilité que du temps de calcul, est déterminée par le choix des primitives pour chacune des trois phases : génération, propagation et évaluation des hypothèses. L'utilisation d'un type de primitive lors de l'une de ces trois phases dépend de la stabilité de cette primitive, du nombre de ces primitives et de la précision de leur localisation.

Génération : le nombre d'hypothèses engendrées étant proportionnel au carré du nombre de primitives de chaque image utilisées dans la phase de génération, il est souhaitable de restreindre cette phase aux primitives les plus stables. D'autre part, il est important que les hypothèses engendrées permettent de récupérer d'autres appariements de primitives, ce qui nécessite un positionnement relativement fiable des primitives, et entraîne l'élimination des primitives trop grosses, par exemple les villes trop importantes.

Propagation : les primitives utilisées pour propager les hypothèses sont déterminantes à la fois pour le temps de calcul nécessaire à chaque propagation et pour la qualité du recalage final. Supprimer des primitives pour accélérer le processus risque de perturber la stabilité de la mise en correspondance.

Evaluation : les primitives sur lesquelles reposent l'évaluation des hypothèses doivent permettre de rendre compte, via le calcul du coût défini par la formule (8), de la qualité du recalage obtenu pour chaque hypothèse propagée. A notre avis, les primitives utilisées pour évaluer une hypothèse doivent nécessairement avoir été utilisées pour propager cette hypothèses.

Afin d'appuyer le processus de mise en correspondance sur les primitives les plus stables et de limiter les temps de calcul, la stratégie retenue ici consiste à engendrer les hypothèses avec les seules agglomérations et de les propager et de les évaluer avec les deux types de primitives (agglomérations et carrefours).

Remarque : nous avons accordé la même importance aux primitives extraites de la carte et de l'image SPOT, que ce soit les agglomérations ou les carrefours. D'autres stratégies peuvent être envisagées où les primitives de l'une ou l'autre image auraient un rôle prépondérant.

Diverses solutions sont envisageables pour réduire les temps de calcul. Pour notre application, nous avons choisi de ne pas propager les hypothèses pour lesquelles la transformation initiale ϕ_0 ne vérifie pas les restrictions apportées sur les paramètres de la transformation recherchée (voir équations (3)). Il ne s'agit pas ici d'une optimisation de la méthode, mais d'une simple élimination, avant propagation, des hypothèses les plus mauvaises.

III.4 Calcul de la transformation finale

Après l'étape de mise en correspondance, nous disposons d'un ensemble d'appariements de primitives provenant de la carte et de l'image SPOT. Le recalage calculé à l'aide de la transformation ϕ associée à l'ensemble d'appariements reste grossier du fait du manque de précision sur la localisation exacte d'une primitive, que ce soit une agglomération ou un carrefour (voir le recalage intermédiaire sur la figure 11). D'où la nécessité de calculer des amers précis à partir des appariements de primitives. Nous proposons ici un schéma global pour le calcul d'un amer précis pour chaque appariement d'un *carrefour* de la carte avec un *carrefour* de l'image SPOT.

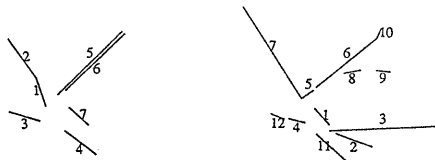


Figure 7. Deux carrefours appariés

L'amer précis associé à un couple de carrefours est calculé après le recalage local de ces deux carrefours, ce recalage étant effectué en appariant les segments associés à chacun des deux carrefours. Soient A et B ces deux ensembles de segments, le problème est alors de déterminer les couples de segments de $A \times B$ qui correspondent à des appariements corrects. Comme pour la mise en correspondance des carrefours et des agglomérations, il ne s'agit pas de la recherche d'un isomorphisme entre un sous-ensemble de A et un sous-ensemble de B : un segment d'une image peut correspondre à plusieurs segments de l'autre image.

Le calcul des amers comprend trois étapes :

- Présélection des appariements suivant l'écart angulaire.
- Sélection des appariements corrects de segments par une technique de programmation dynamique.
- Calcul de l'amer à partir des couples de segments appariés : dans chaque image, on calcule le centre d'inertie des intersections des segments sélectionnés.

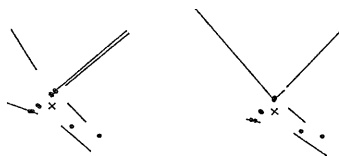


Figure 8. Calcul d'un amer

L'application affine finale ϕ_f est calculée par une AMC sur les amers déterminés par ce recalage local des carrefours. La rigidité du modèle de déformation permet d'éliminer successivement les amers ayant un résidu trop grand et d'éviter des erreurs grossières. Par contre, cette rigidité ne permet pas de prendre en compte les problèmes de dénivellation et de perspective (voir figure 12).

III.5 Evaluation du recalage

L'évaluation du recalage final est réalisé en comparant la transformation ϕ_f à une transformation calculée à partir de points de contrôle choisis manuellement.

Soient $(p_i, q_i)_{i=1..n}$ les amers obtenus par notre méthode de recalage automatique et ϕ_f la transformation finale associée. Soient $(P_i, Q_i)_{i=1..N}$ les points de contrôle repérés manuellement et ϕ_0 la transformation affine calculée par une AMC sur ces points. La mesure de l'estimation du recalage est la moyenne des résidus entre les transformés des points P_i , désignés manuellement, par les applications ϕ_f et ϕ_0 .

$$e = \frac{\sum_{i=1}^N \text{dist}(\phi_f(P_i), \phi_0(P_i))}{N} \quad (7)$$

Le choix de cette mesure, relative à la transformation ϕ_0 , permet de prendre en compte le type de transformation paramétrée qui a servi de modèle, ce qui ne serait pas le cas si la mesure était calculée à partir des écarts entre les points $\phi_f(P_i)$ et les points Q_i .

IV. RESULTATS

Les résultats présentés concernent tout d'abord l'application de notre méthode au recalage de différents couples image SPOT/carte routière, extraits d'une même scène SPOT, chaque image couvrant une zone de 10km x 10km. Ensuite le même principe de *génération et*

propagation d'hypothèses a été appliqué à la recherche d'une petite carte (10km x 10km) dans une grande image SPOT (30km x 30km). La scène couvre une partie de l'Alsace et de la Forêt Noire, entre Colmar, Mulhouse et Freiburg.

IV.1 Recalage simple

Le tableau 1 rassemble les résultats du recalage de 5 scènes différentes. Il précise pour chaque scène le nombre d'agglomérations et de carrefours extraits de chaque image, le temps de calcul sur Vax 3000-400 de la phase de mise en correspondance des primitives, le nombre final d'amers retenus et l'écart moyen.

La précision du recalage est limitée pour deux raisons. La première est la précision de la carte elle-même sur laquelle des erreurs de plusieurs dizaines de mètres sur le tracé des routes est possible. La seconde cause d'imprécision provient des diverses approximations utilisées tout au long du processus de recalage, en particulier lors de la mise sous forme vectorielle des réseaux extraits des deux images. Une précision plus grande pourrait être obtenue en retournant au niveau *pixel*, cependant la nature différente des images ne permettrait certainement pas un recalage subpixel comme cela peut être envisagé pour des images provenant d'une même source.

IV.2 Localisation d'une petite carte

Le problème est de déterminer la position d'une petite carte de 10km x 10km dans une grande image SPOT de 30km x 30km (i.e., 3000x3000 pixels). Le même principe de génération et propagation d'hypothèses a été utilisé pour résoudre ce problème. Cependant le nombre important de primitives, en particulier de carrefours, a nécessité la modification du processus de mise en correspondance, afin de limiter les temps de calcul : seules les agglomérations ont été utilisées lors de la mise en correspondance des primitives, que ce soit pour la génération, la propagation ou l'évaluation des hypothèses. La transformation ϕ associée au meilleur ensemble d'appariements d'agglomérations sert ensuite à déterminer les couples de carrefours pouvant être appariés. La transformation finale ϕ_f est calculée sur ces couples de carrefours de la même façon que précédemment.

Cette procédure a été testée pour localiser plusieurs cartes dans l'image SPOT. Les résultats sont présentés dans le tableau 2 et la figure 13 montre le recalage du réseau et des agglomérations extraits de deux cartes dans l'image SPOT de 3000x3000 pixels.

L'image SPOT comprend 126 agglomérations et 883 carrefours, et les 5 cartes sont celles présentées dans le tableau 1.

V. CONCLUSION

Nous avons présenté dans cet article une méthode fiable et rapide pour le recalage d'une image SPOT et d'une carte routière. Sa principale originalité réside dans l'utilisation conjointe de deux types de primitives de haut niveau : les zones urbaines et les carrefours.

Par rapport à une méthode antérieure [18] qui n'utilisait que les carrefours comme primitives de mise en correspondance, méthode qui donnait de bons résultats en milieu rural, mais rencontrait des problèmes de stabilité et de temps de calcul pour des scènes au réseau routier dense, l'utilisation conjointe des zones urbaines et des carrefours a permis d'accélérer et de stabiliser le processus de mise en correspondance.

L'application au problème difficile qu'est la recherche d'une petite zone de la carte dans une grande scène SPOT nous a permis de montrer l'efficacité et la stabilité de notre méthode.

Remerciements : ces travaux ont été réalisés grâce au soutien financier du CNES et de Dassault Electronique.

scène	image SPOT		Carte		temps CPU	amers	écart moyen
	A	C	A	C			
1	11	78	11	69	18s	7	29m
2	10	93	12	65	24s	11	26m
3	11	98	12	73	35s	18	28m
4	9	108	4	56	4s	12	28m
5	14	74	8	60	16s	9	30m

TABLE 1. Recalage de 5 scènes
A : agglomérations, C : carrefours

References :

- [1] Chassery J.-M., Montanvert A. : *Géométrie discrète en analyse d'images*, Traité des Nouvelles Technologies, Série Images, Editions Hermès, 1991.
- [2] Chen H., Huang T. : *Maximal Matching of Two Three-Dimensionnal Point Sets*, 8th International Conference on Pattern Recognition, Paris, 1986, 27-31 Oct, pp 1048-1050.
- [3] Denis P., Baudoin A. : *Applications topographiques de SPOT à l'IGN-France*, 16th ISPRS, Kyoto, 1988, Commission IV, pp 106-115.

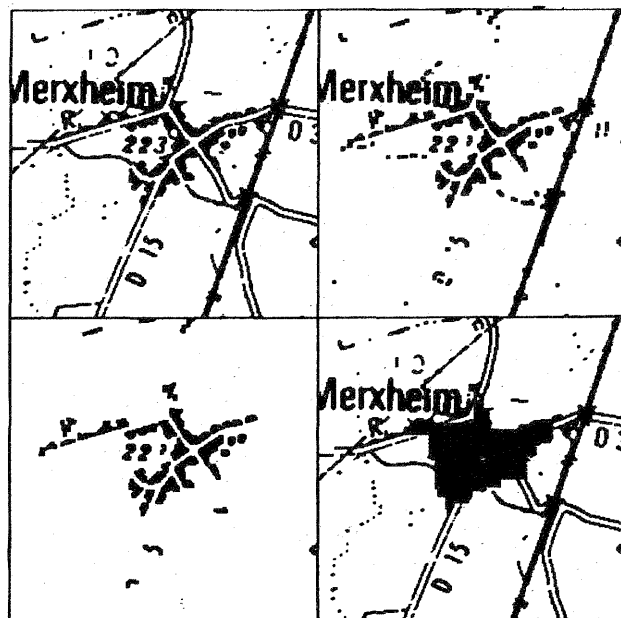


Figure 9. Détection des villes sur les cartes (© IGN)
a. image binaire d'une carte, b. ouverture morphologique
c. élimination des mots et des gros objets
d. silhouette de l'agglomération

scène	temps CPU	amers	écart moyen
1	1mn 25s	6	53m
2	1mn 35s	10	45m
3	1mn 40s	13	39m
4	14s	12	42m
5	40s	10	30m

TABLE 2. Localisation des 5 cartes

- [4] Destival I. : *Mathematical morphology applied to remote sensing*, Acta Aeronautica, Vol. 13, No 6/7, pp 371-385, 1986.
- [5] Deutsch E. S. : *Thinning algorithms on rectangular, hexagonal and triangular arrays*, CACM, Vol. 15, No 9, pp 827-837, 1972. No 4, July 1981, pp 412-424.
- [6] Faugeras O., Price K. : *Semantic Description of Aerial Images Using Stochastic Labeling*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-3, No 6, Nov 1981, pp 633-642.

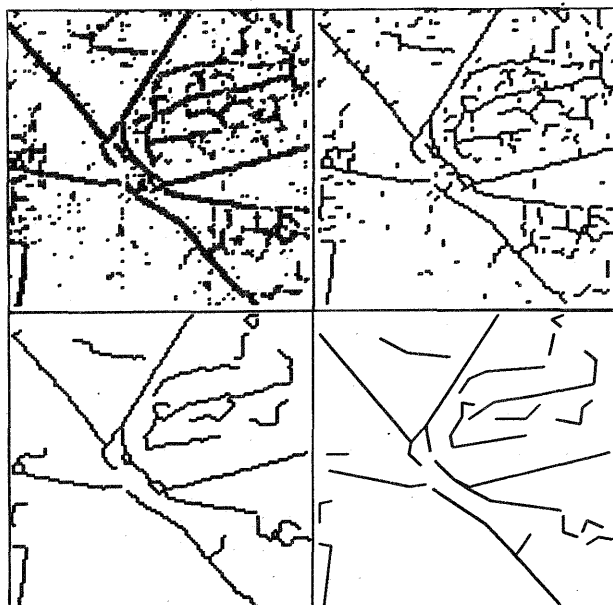


Figure 10. Vectorisation des réseaux
a. image binaire du réseau, b. squelette
c. ébarbulation, d. approximation linéaire

a	b
c	d

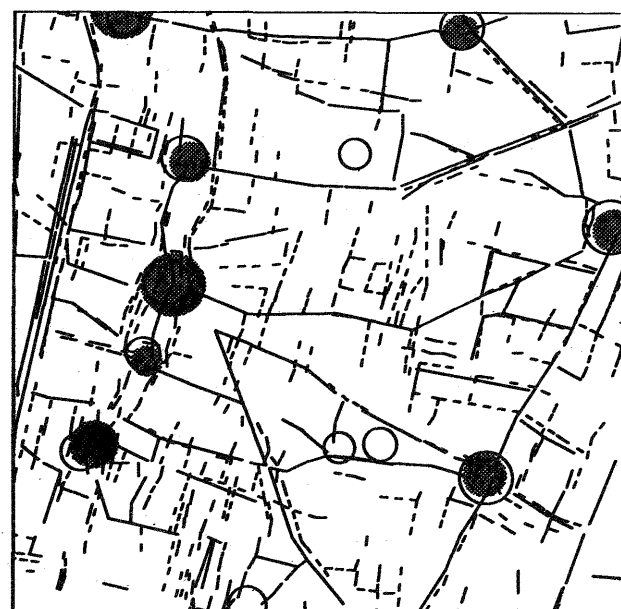


Figure 11. Recalage intermédiaire des réseaux

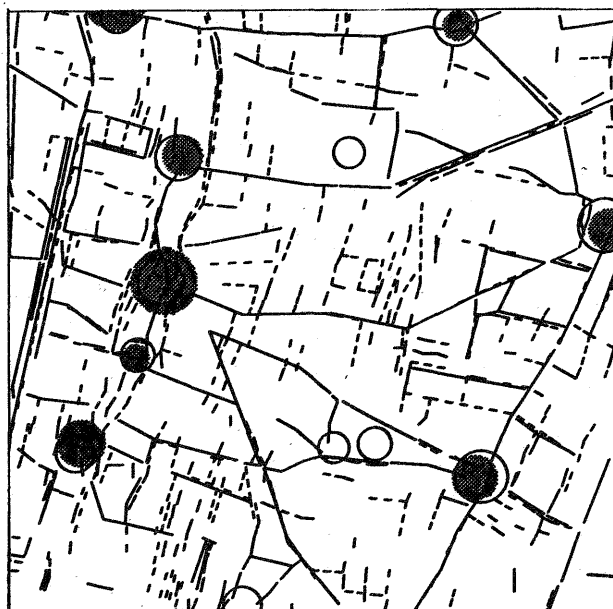


Figure 12. Recalage final des réseaux

- [7] Fischler M. A., Tenenbaum J. M., Wolf H. C.: *Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique*, Computer Graphics and Image Processing, Vol. 15, pp 201-223, 1981.
- [8] Giraudon G.: *A Efficient Edge Following Algorithm* 5th Scandinavian Conference on Image Analysis, Stockholm, June 1987.
- [9] Goldgof D., Huang T., Lee H.: *Feature Extraction and Terrain Matching*, IEEE Computer Vision and Pattern Recognition, Ann Arbor, June 1988, pp 899-904.
- [10] Goshtasby A.: *Image Registration by Local Approximation Methods*, Image and Vision Computing, Vol. 6, No 4, November 1988, pp 255-261.
- [11] Guban D., Dowman I.: *Topographic Mapping from SPOT Imagery*, Photogrammetric Engineering and Remote Sensing, Vol. 54, No 10, October 1988, pp 1409-1414.
- [12] Izraelevitz D., Carlotto M.: *Extracting Road Networks from Low Resolution Aerial Imagery*, SPIE, Vol. 1001, Visual Communication and Image Processing, 1988.
- [13] McKeown D., Denlinger J.: *Cooperative methods for road tracking in aerial imagery*, IEEE Computer Vision and Pattern Recognition, Ann Arbor, June 1988.
- [14] Michaelis M.: *A Geometrical Analysis of SPOT Data*, Colloque SPOT 1, Utilisation des images, bilan, résultats, Paris, Nov. 87, pp 1235-1243.
- [15] Medioni G.: *Matching of a map with an aerial image*, Proc of 6th ICPR, Munich, Germany, Oct. 19-22, 1982.
- [16] Nagao T., Agui T., Nakajima M.: *An automatic road vector extraction method from maps*, 9th ICPR, Rome, Nov. 1988, pp 585-587.
- [17] Nevatia R., Babu K. R.: *Linear feature extraction and description*, Computer Graphics and Image Processing, Vol. 13, 1980, pp 257-269.
- [18] Roux M., Lopez Krahe J., Maitre H.: *Matching Multi-Source Images : SPOT image / Geographic Map*, Eusipco 90, September 18-21, 1990, Barcelona, Spain, pp 1035-1038.
- [19] Serendero M.-A.: *Extraction d'informations symboliques en imagerie SPOT : réseaux de communication et agglomérations*. Thèse de Doctorat, présentée le 9 Déc. 1989, Université de Nice.
- [20] Serra J.: *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.



Figure 13. Localisation de petites cartes
 Superposition du réseau et des agglomérations de deux cartes
 localisées dans l'image SPOT de 3000x3000 pixels (© SPOT IMAGE)