

Vers un paramétrage local automatique d'opérateurs de vision précoce en imagerie aérienne oblique

Ronan Thomas^{1,2,3}. Georges Stamon¹. Véronique Serfaty². Raphaël Horak³

¹ Laboratoire CRIP5/
Systèmes Intelligents de Perception,
Université René Descartes,
Paris. France.

thor@math-info.univ-paris5.fr,
stamon@math-info.univ-paris5.fr

² Laboratoire DGA/DCE/CTA/
Géographie Imagerie Perception,
Arcueil. France.

Veronique.Serfaty@etca.fr

³ Thomson-CSF Detexis/
Division Electronique de Missile,
Trappes. France.

Raphael.Horak@detexis.thomson-csf.com

Cette thèse (bourse CIFRE) est réalisée dans le cadre d'une collaboration entre le laboratoire SIP-Université ParisV, le laboratoire GIP de la DGA et le laboratoire TNU de Thomson-CSF Detexis.

des opérateurs aux données image, et, de façon duale, l'adaptation des données image aux opérateurs.

Abstract

Many Computer Vision based navigation systems integrate external and internal data, such as camera parameters, to interpret their image features (edges, segments, regions, blobs, ...) by the use of geometrical models. However, this data is rarely exploited in low-level processing. Therefore, in aerial oblique view applications, we propose to guide, via this data, the parameter setting of some image processing operators which are sensitive to perspective effects. Several possible methods, ranging from adapting operators to image, and adapting image to operators, are presented.

Résumé

En imagerie aérienne oblique, beaucoup d'applications échouent à la suite d'une segmentation défaillante de l'image. Nous montrons que la projection perspective de la scène introduit des déformations géométriques et radiométriques variables empêchant un paramétrage a priori des opérateurs d'extraction global à l'image. Cependant, le paramétrage local optimal de ces opérateurs dépend directement des lois de perspective. Or, la plupart des systèmes automatiques de navigation aérienne à base de vision utilisent déjà, via les données instantanées de vol et de caméra fournies par le système, un modèle géométrique de la formation de l'image pour traiter leurs primitives extraites de l'image. Actuellement, ces informations sont très rarement exploitées dès la vision précoce (prétraitement de l'image, extraction des primitives). Il nous apparaît donc opportun, dans le cadre spécifique de nos applications en visée oblique, de guider localement certains opérateurs sensibles aux effets de perspective (échelle, déformations,...). A l'aide d'un modèle géométrique simple, nous proposons deux approches pour mieux gérer la segmentation : l'adaptation

Mots-Clefs

Vision précoce. Imagerie aérienne oblique. Modèle géométrique de scène. Projection perspective. Filtrage. Segmentation. Opérateurs adaptatifs. Multi-résolution.

I. Cadre Applicatif

Dans le cadre de la navigation et du guidage de porteurs aériens (autodirecteurs de missiles, drones,...), les industriels cherchent à optimiser la qualité de leurs systèmes de vision. Ceux-ci consistent essentiellement en des modules de segmentation (contours, régions, points chauds, ...) et de recalage automatique (mise en correspondance de primitives, corrélation, ...). Notre étude se restreint au cadre spécifique de scènes planes (aéroports,...). Par exemple (Figure 1), afin de rectifier sa trajectoire, le calculateur génère une vue synthétique du modèle cartographique de la scène selon la position estimée du porteur, et détermine le recalage entre les primitives de cette image, avec celles de l'image réelle fournie par le capteur. Le bon comportement du système dépend donc en grande partie de la qualité et de la pertinence des primitives segmentées.

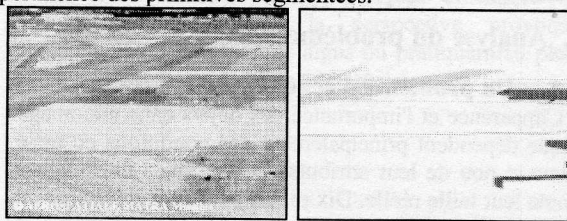


Figure 1 : Recalage d'une image FLIR avec une vue synthétique du modèle cartographique de la scène.

II. Motivations

Aujourd'hui encore, la segmentation des images aériennes reste un problème difficile. Les caractéristiques radiométriques et géométriques des structures contenues dans ces images sont difficiles à modéliser, car très

diverses. Aussi utilise-t-on des algorithmes très généraux, rapides, efficaces, souples et reconnus pour leur polyvalence sur une grande variété d'images (filtre optimal de Canny-Deriche [1], croissance de régions [2], ...).

De plus, et c'est ici le principal objet de nos recherches, la prise de vue sur ces types de porteur très véloces (>mach2), est de nature *oblique* voire frontale (Figure 1). Ceci offre un champ de perception maximal dans l'axe de déplacement et permet de mieux anticiper et contrôler à temps tout changement de direction. Contrairement à une vue verticale, la projection perspective de la scène dans le plan focal entraîne des disparités géométriques (résolution, échelle, déformations,...) et radiométriques (lissage des textures, baisse de contraste,...) éventuellement importantes sur les structures perçues, selon les conditions de prise de vue (site, ouverture, altitude,...) et la position dans l'image. La forte réactivité demandée du porteur oblige une détection très lointaine de l'objectif et donc une sensibilité des opérateurs aux structures éloignées donc petites et souvent faiblement contrastées. Un phénomène de sursegmentation est souvent inévitable. Une grande focale du capteur permet de mieux percevoir la cible et de limiter les variations de résolution mais augmente aussi les risques de décrochage. Les déformations géométriques comme l'écrasement des structures de la scène entraînent de toute façon certains problèmes d'extraction (Figure 3).

Or, la plupart des opérateurs utilisés en traitement d'image [2][3][4] ne s'adaptent pas localement aux effets dus à la perspective. Ces opérateurs travaillent généralement de façon "isotrope", c'est-à-dire de façon identique en tout point de l'image et dans toute direction, et ne modifient pas leurs paramètres (taille et géométrie du voisinage, seuils radiométriques, ...), faute d'information sur la résolution ou la déformation dans l'image. Ainsi, certains opérateurs peuvent localement s'avérer inefficaces [5]. Une sursegmentation est alors souvent nécessaire à l'extraction des structures mais entraîne un risque de surcharge temporelle du processus d'analyse ou d'erreur de recalage. La nécessité d'une échelle d'analyse *variable* apparaît donc en imagerie aérienne oblique.

III. Analyse du problème

A. Un problème d'échelle locale

L'apparence et l'importance des objets dans une image oblique dépendent principalement des conditions de prise de vue et non de leur attributs géométriques intrinsèques comme leur taille réelle. Dix pixels peuvent représenter un buisson ou un bâtiment.

Aussi, une propriété inhérente aux objets du monde est qu'ils n'ont un sens, dans une application ou une tâche donnée, qu'à une certaine échelle d'interprétation ou à un niveau de détails [6]. Par exemple, pour extraire un objet de classe route dans une image aérienne, il est difficile voire impossible d'analyser l'image à une résolution au sol de l'ordre du millimètre, ni du kilomètre. Le type d'information que l'on peut extraire d'une image est donc,

dans une large mesure, déterminée par l'adéquation de la taille (résolution) des structures de l'image et du domaine spatial d'analyse (taille, voisinage) de l'opérateur. S'il est parfois possible de connaître a priori la résolution des données, il n'est pas toujours simple de déterminer l'échelle d'analyse à utiliser pour une scène inconnue. C'est le cas des scènes tridimensionnelles où se produisent des effets de perspective sur la géométrie des structures et sur les textures.

B. Comportement visuel humain

On peut remarquer qu'un être humain adapte son analyse d'une scène 3D en fonction de la profondeur visuelle des objets. L'information d'échelle fournie par la vision binoculaire et par un processus cognitif permet une analyse scalaire localement adaptée à l'apparence des structures présentes dans l'image. Ainsi, au premier plan, l'opérateur interprète les données à plus grande échelle en lissant les informations trop détaillées (textures, irrégularités,...), puis progressivement, avec l'éloignement des objets, il affine sa zone d'analyse et ses seuils d'interprétation (précision, contraste, ...) afin d'interpréter correctement les données en perspective et en extraire les primitives les plus pertinentes. Il apparaît donc intéressant et justifié de prendre en compte ce comportement visuel et de l'injecter au processus de segmentation d'images aériennes en visée oblique au moyen d'un modèle perspectif de la scène.

C. Notre approche

Afin d'interpréter leurs primitives images, la plupart des systèmes de navigation aérienne à base de vision utilisent déjà un modèle géométrique de formation de l'image [4][7][8] via certaines données géométriques de vol ou de prise de vue fournies par le système (altimétrie, GPS, inertie, paramètres caméra, laser, ...). Ces conditions de prise de vue (CPDV) sont pourtant rarement utilisées dès le prétraitement de l'image et l'extraction des primitives [9]. Il nous apparaît donc opportun [5][10] d'utiliser ces CPDV à l'aide d'un modèle géométrique simple, permettant de guider localement le paramétrage de certains opérateurs sensibles aux effets de perspective. Nous nous intéressons donc à la contribution des CPDV au processus d'extraction de primitives en imagerie aérienne oblique.

Nous proposons principalement deux approches duales pour mieux gérer la segmentation : l'adaptation des opérateurs aux données image, et l'adaptation des données image aux opérateurs.

D. Cadre algorithmique

Devant l'étendue des techniques de segmentation existantes et dans le cadre applicatif considéré, nous avons restreint notre étude à des opérateurs de segmentation de type extracteur de contours par filtrage optimal de Shen-Castan [11] et par croissance de régions. Ce choix, a priori limitatif, permet néanmoins de mieux cerner et mesurer les

apports de nos recherches par une expertise avancée de ces techniques bien connues et reconnues.

La première partie de l'article expose plus précisément la problématique de la visée oblique et l'influence des effets de perspective sur l'efficacité de certains opérateurs. Puis, nous proposons un modèle géométrique fournissant les informations nécessaires au paramétrage local de ces opérateurs. Nous présentons ensuite quelques approches originales améliorant la segmentation. Enfin, nous concluons et dégageons certaines perspectives.

IV. Description et modélisation du problème

A. Les effets de la perspective

1. Sur l'apparence des structures dans l'image

La force de la gravitation a forgé l'aspect de la Terre, par l'érosion naturelle et d'autres phénomènes naturels. La majeure partie du relief s'est aplanie. Cependant, dans une moindre mesure, car souvent issues de l'activité humaine, apparaissent certaines structures verticales. Ainsi, les droites prépondérantes dans une scène naturelle sont parallèles ou perpendiculaires au sol.

Les structures significatives dans les applications de navigation aérienne sont des éléments cartographiques tels que routes, bâtiments, cours d'eau, forêts, ... Dans une image aérienne oblique, les disparités locales induites par la perspective sont de différentes natures.

- **Géométrique** : modification des propriétés géométriques intrinsèques des structures de la scène : changement de taille apparente, déformations des angles (parallélisme, ...), écrasement des structures, ...
- **Radiométrique** : résolution variable (texture, bruit...), effets atmosphériques éventuels (brume, diffusion de type Rayleigh...), atténuation du contraste, ...

De plus, l'importance de ces disparités diffère selon la position des structures dans l'image.

2. Sur les opérateurs

a) Sensibilité des opérateurs :

La plupart des outils de traitement d'image, de par leurs fonctions de filtrage, de caractérisation d'homogénéité ou de discontinuité, utilisent la notion de voisinage à travers une taille de fenêtre ou la valeur d'un paramètre (ex: variance d'une gaussienne). Or, la perspective de l'image ne permet pas d'établir une seule résolution des objets à extraire, donc de définir *a priori* les dimensions des domaines d'analyse des opérateurs. Aussi, la dépendance directe de l'apparence des objets aux CPDV limite l'efficacité d'une fenêtre fixe de calcul. Un voisinage de calcul trop restreint préserve les détails et le bruit des textures, alors qu'un voisinage trop important lisse ou élimine certains contours. L'anisotropie de l'image n'est donc pas prise en compte, ni dans la taille du voisinage, ni dans aucun seuil lié à une mesure géométrique dans l'image.

b) quelques opérateurs sensibles :

- **Filtrage du bruit** : L'importance et la nature du bruit (détails, textures,...) varient dans l'image car dépendent de la résolution locale de l'image.
- **Calcul du gradient** : Un voisinage fixe de calcul du gradient est localement inadapté aux tailles et aux détails apparents des structures. La « fenêtre de différenciation » dépend des résolutions verticale et horizontale de la scène. Les transitions entre structures sont plus étendues au devant de la scène, contrairement aux structures éloignées petites et aplaties. Les structures proches sont souvent texturées et sur-détaillées produisant une forte réponse aux filtres différentiels.
- **Extraction des contours** : Du premier au dernier plan de l'image, la composante verticale du gradient décroît contrairement à l'horizontale qui progressivement intègre toute l'information de contour ; ce phénomène est bien évidemment dû à l'écrasement des structures selon l'horizon. La norme du gradient diminue également en fonction de la distance ; cet effet est principalement dû à la baisse de résolution et surtout à certains effets atmosphériques. Il est donc impossible, par exemple, de choisir des seuils d'hystérésis *globaux à toute l'image*. De plus, la génération des chaînes de contours (extraction, chaînage, fermeture, sélection, polygonalisation) est souvent fondée sur des critères de longueur minimale, d'écart maximal exprimés dans la métrique image (pixel). Or ces mesures de critère sont généralement invalidés par la perspective car utilisés dans l'espace des pixels 2D qui ne reflète pas la réalité 3D de la scène. Il convient donc de définir, d'une part, l'espace métrique d'analyse des opérateurs exprimé en pixel, et d'autre part, l'espace métrique des structures de la scène exprimée en mètres (Figure 2). Ainsi, Figure 6, certaines chaînes parasites appartenant à une zone texturée du premier plan pourront être conservées avec une précision inutile, alors que d'autres chaînes plus courtes et éloignées appartenant à des vrais contours d'objets seront éliminées ou imprécises. De plus, la perspective provoque l'apparition d'angles très aigus ou pratiquement plats qu'il convient de détecter très précisément.

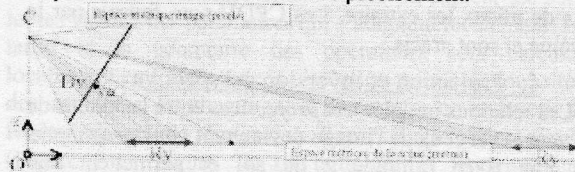


Figure 2 : Effet de perspective entre les métriques scène/image

- **Approche région** : Pour les mêmes raisons que celles invoquées pour le filtrage du bruit, les seuils radiométriques fondés sur des estimateurs de propriétés d'homogénéité doivent varier dans l'image

en raison de leur dépendance étroite à la résolution et au contraste locaux. Un seuil donné induit une sursegmentation sur les structures détaillées et contrastées du premier plan, alors que dans le fond de l'image, des structures plus faiblement contrastées fusionnent ou sont éliminées par le critère de taille minimale de région. Par contre, il faut noter (Figure 3) le meilleur comportement de l'extraction des structures écrasées aux angles aigus grâce à la fermeture intrinsèque des frontières et à la finesse d'agglomération des pixels (généralement un critère radiométrique).

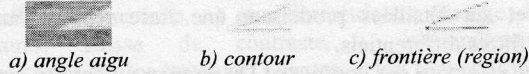


Figure 3 : Un avantage de l'approche région

3. Une solution géométrique

Le manque de connaissances sur la résolution et la géométrie locale des structures dans l'image est donc le principal défaut de ces opérateurs à paramétrage préfixé. Pourtant, il existe une valeur locale adéquate du paramétrage de ces opérateurs que nous pouvons directement relier à ces informations. Cela démontre que l'apport d'un contexte géométrique de la formation de l'image oblique via les CPDV doit guider localement et efficacement ces opérateurs. Notons que certaines modifications radiométriques dépendent également de cette géométrie.

B. Modèle géométrique

La position dans l'image et l'appréhension cognitive de la perspective permettent la bonne interprétation locale des données images. Aussi, nous modélisons l'aspect cognitif de la scène par un modèle géométrique de la prise de vue. L'analyse du problème nous a incités à choisir un modèle projectif simplifié de formation de la scène dans l'image à partir de modèles existants tels que [4][7][8]. De fortes hypothèses ont ainsi été posées. Le relief des scènes survolées est globalement plat car l'altitude de vol est supposée suffisamment grande devant la hauteur des structures de la scène. Nous avons ainsi modélisé la scène par un plan (Figure 4). Rappelons que la question n'est pas de reconstruire ou d'analyser les structures de la scène mais de mieux les extraire. Les CPDV sont fournis par le système et sont précis.

1. Projection et inversion perspective

Les relations initialement *projectives* entre les points de la scène 3D et ceux de l'image deviennent *homographiques* une fois fixée l'altitude z du sol et permettent une correspondance biunivoque plan à plan entre l'espace "image" et l'espace "scène" (hors points du ciel).

• Relations géométriques entre scène et image :

avec :

- $NbCol, NbLig$: taille de l'image (pix)
- $Site$: site de la camera
- Zc : altitude de la camera (m)

- f_u, f_v : focales selon u et v (pix)
 - C_u, C_v : centre image (pix)
 - O_u, O_v : angles d'ouverture caméra (rad)
 - $x, y, (z=constante connue)$: coordonnées d'un point de la scène (sol) relativement au repère capteur (m)
 - u, v : coord d'un point image (pix)
- avec $f_u = \frac{NbCol/2}{\tan(O_u/2)}$ $f_v = \frac{NbLig/2}{\tan(O_v/2)}$

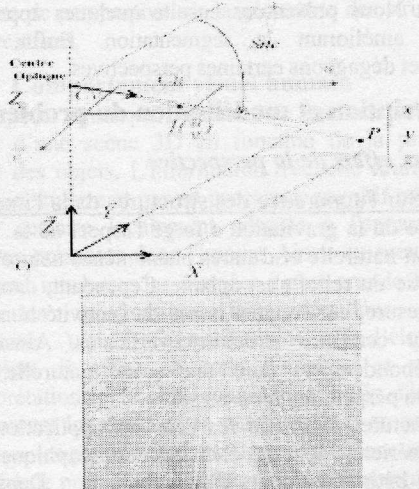


Figure 4 : Le modèle géométrique local d'une scène

Ligne de l'horizon : $V_{horiz} = f_v \tan(Site) + C_v$

De la scène $(x,y,z$ fixé) vers l'image (u,v) :

$$U(x, y, z) = \frac{f_u x}{y \cos(Site) + (z - Zc) \sin(Site)} + C_u$$

$$V(y, z) = \frac{f_v (y \sin(Site) - (z - Zc) \cos(Site))}{y \cos(Site) + (z - Zc) \sin(Site)} + C_v$$

De l'image (u,v) vers la scène $(x,y,z$ fixé) :

$$X(u, v, z) = \frac{f_u (C_u - u) (Zc - z)}{f_u (f_u \sin(Site) + (C_u - v) \cos(Site))}$$

$$Y(v, z) = \frac{(Zc - z) ((C_u - v) \sin(Site) - f_u \cos(Site))}{(C_u - v) \cos(Site) + f_u \sin(Site)}$$

Résolutions horizontale et verticale (ie la valeur métrique au sol) pour un pixel (u,v) :

$$R_x(u, v, z) = X\left(u + \frac{1}{2}, v, z\right) - X\left(u - \frac{1}{2}, v, z\right)$$

$$R_y(v, z) = Y\left(v - \frac{1}{2}, z\right) - Y\left(v + \frac{1}{2}, z\right)$$

Taille de fenêtre (W_u, W_v) d'une résolution au sol (R_x, R_y) en un point image (u,v) :

$$W_u(u, v, z, R_x) = U\left(X(u, v, z) + \frac{R_x}{2}, Y(v, z)\right) - U\left(X(u, v, z) - \frac{R_x}{2}, Y(v, z)\right)$$

$$W_v(v, z, R_y) = V\left(Y(v, z) - \frac{R_y}{2}, z\right) - V\left(Y(v, z) + \frac{R_y}{2}, z\right)$$

Calcul de la Ligne Vdr de résolution R_y mètres au sol ayant une projection dans l'image de Dv pixels :

Ceci revient à résoudre (Figure 2) :

$$Y\left(C_u, Vdr - \frac{Dv}{2}, Z_s\right) - Y\left(C_u, Vdr + \frac{Dv}{2}, Z_s\right) = R_y$$

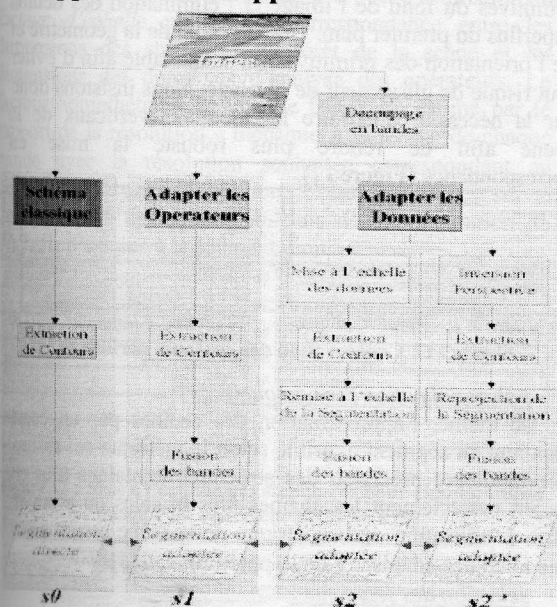
dont la solution est :

$$Vdr = C_v + f_v \tan(\text{Site}) + \frac{1}{2} \sqrt{D_v^2 + 4 \frac{f_v D_v (Z_c - Z_s)}{R_v \cos(\text{Site})^2}}$$

directions (verticale, parallèle, perpendiculaire, ...) : Le modèle fournit les angles apparents ou réels entre structures comme, par exemple, la direction verticale au sol ou la perpendiculaire à une droite au sol dans l'image.

Ce modèle bien qu'approximatif apporte un contexte suffisant pour guider localement des opérateurs. Ce modèle n'est, en fait, valide que pour les éléments plans de la scène, cependant la majorité des structures 3D restent dans des échelles comparables à celles fournies par le modèle. Une extension de ces travaux de recherches prévoit l'utilisation de modèles de relief plus riches. A ce stade de l'étude, nous ne nous intéressons qu'à démontrer l'adéquation nécessaire de la taille des opérateurs à l'échelle des données.

V. Approches développées



$s0$ =schéma classique ; $s1$ =adaptation du paramétrage des opérateurs ; $s2$ =mise à l'échelle des données ; $s2'$ = $s2$ + correction des distorsions géométriques

Figure 5 : Les approches développées

Deux voies sont proposées pour traiter les effets de perspective. La première approche $s1$ est d'adapter les opérateurs en ajustant dynamiquement leur paramétrage. L'approche complémentaire est d'adapter les données images et se divise en deux sous schémas. La technique $s2$ effectue une mise à l'échelle des données aux domaines d'analyse des opérateurs. La technique $s2'$ effectue en plus une correction géométrique par inversion perspective.

Ces approches ont été expérimentées selon deux modalités d'expérimentation : la première, naturelle, est de traiter l'image dans sa globalité, la seconde consiste à

effectuer un découpage en bandes horizontales. Le découpage retenu est effectué par recherche des lignes successives de l'image correspondant à des changements notables de résolution. Ce découpage autorise ainsi un traitement isotrope à l'intérieur de chaque bande, améliorant sa rapidité et sa simplicité.

A. Schéma classique proposé

Le schéma dénoté $s0$ consiste en une extraction classique de contours par filtrage de Shen-Castan [11] (β Shen), puis extraction des maxima locaux et chaînage par hystérésis (*HighHyst*, *LowHyst*), suivie d'une élimination des contours parasites (*LengthMin*) et d'une approximation polygonale selon un écart maximum de la chaîne au segment (*ApproxSeg*). La segmentation $S0$ sert ainsi d'étalon pour les autres approches (Figure 5). Pour toutes les approches proposées y compris $s0$, les seuils d'hystérésis sont calculés selon ce schéma : calcul du maximum *MaxGrad* de la norme du gradient sur l'image ou la bande image, puis affectation du seuil *HighHyst* à $MaxGrad/2$ et du seuil *LowHyst* à $MaxGrad/4$.

Figure 6, les contours selon $S0$ sont peu pertinents et imprécis au loin. En particulier, la piste n'est pas bien extraite. Des contours parasites apparaissent (1^{er} plan).

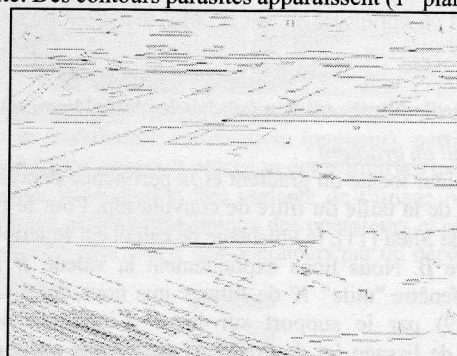


Figure 6 : Segmentation directe $S0$

(β Shen=0.8, *LowHyst*=6, *HighHyst*=17, *LengthMin*=20, *SegApprox*=3)

B. Adapter l'opérateur aux données image

1. Principe

Le principe général est le suivant : les données traitées sont les données images brutes issues du capteur ; ce sont ici les valeurs des paramètres qui s'adaptent localement. La taille et la géométrie des opérateurs sont calculés localement en fonction de la résolution minimale souhaitée des structures. En chaque pixel de l'image, on reprojette la fenêtre au sol de dimension R sur l'espace des pixels (Figure 2 et Figure 7). La taille de la fenêtre-image s'adapte ainsi à l'échelle apparente des structures de la scène. Une extension possible est de déformer la fenêtre de calcul en parallélogramme ou un cercle en ellipse.

2. Quelques développements

Remarque : Les développements suivants sont traités indépendamment les uns des autres.

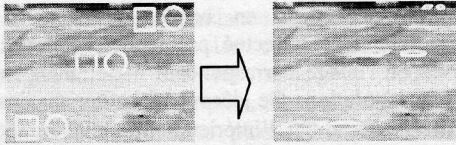


Figure 7 : S1=Adaptation du domaine spatial d'analyse des opérateurs à l'échelle des structures dans l'image

- Filtrage du bruit (par filtre médian) : Sur la Figure 8, on peut comparer les images issues d'une part d'un filtre médian à fenêtre constante et d'autre part, d'un filtrage adapté à la résolution des structures recherchées. La taille de fenêtre de lissage est ici recalculée localement par $W_v(v,z,ex:R_y=20m)$. La fenêtre s'étend sur plusieurs dizaines de pixels à l'avant-plan alors qu'elle devient nulle dans le lointain. Le bruit est éliminé pour les structures détaillées et texturées de l'avant-scène, tout en préservant les structures éloignées. Un tel compromis est impossible avec une fenêtre fixe de travail.



Figure 8 : filtrage médian (a:3x3pixels b:25x25m au sol)

- calcul du gradient :

La valeur locale du gradient et sa pertinence dépendent en partie de la taille du filtre de convolution. Pour le filtre optimal de Shen [11], le voisinage de calcul est fonction du paramètre β . Nous liions explicitement la valeur β à sa taille de fenêtre "utile" W de filtrage que nous définissons (Figure 9) par le support sur lequel l'intégrale de la fonction de lissage (et de gradient) atteint une proportion donnée K (ex:90%). Notons bien que ce calcul ne fait qu'estimer β en amont du filtrage et que celui-ci reste à support infini. Nous définissons localement W comme la valeur $W_y(v,z,R_y)$ du modèle géométrique, R_y étant l'échelle des structures recherchées. Grâce au découpage offrant une résolution peu variée dans chaque bande, nous extrayons les contours avec un β constant par bande.

$$f(x) = \frac{\beta}{2} e^{-\beta|x|}, \quad \int_{-W}^W f(x) dx = K \Rightarrow \beta = \frac{-\log(1-K)}{W}$$

avec W = largeur efficace du filtre (lissage et gradient)
 = échelle pixellique locale des structures = $W_y(v,z,R_y)$

Figure 9 : Le filtre de Shen et une mise en correspondance de sa largeur utile à la résolution locale de l'image oblique

On peut noter Figure 10 l'amélioration globale de l'extraction (détails, texture) par la sélection des zones de contours en fonction de l'adéquation taille de l'opérateur/échelle des données.

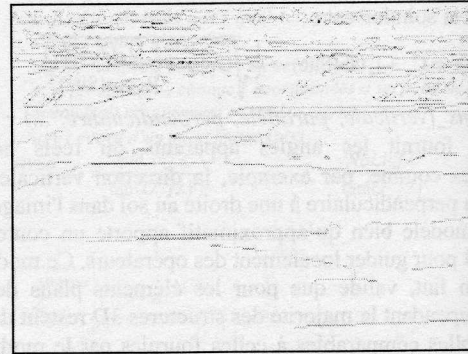


Figure 10 : segmentation s1 reconstruite (extraction avec filtre de Shen adapté par bandes de résolution)

- sélection des chaînes de contour:

Comme abordé au paragraphe IV.A.2, la sélection des contours par un critère de métrique au sol permet également le compromis entre sauvegarde et précision des primitives du fond de l'image et l'élimination des détails superflus du premier plan. Une diversité de la géométrie et de l'orientation des primitives est souhaitable afin d'éviter tout risque de glissement de recalage. Nous insistons donc sur la nécessité d'extraire les éléments verticaux de la scène afin de rendre plus robuste la mise en correspondance (Figure 11).



a) critère = longueur < 20 pixels b) critère = longueur < 20 m

Figure 11 : Conservation des contours pertinents

- polygonalisation des contours:

De même, l'approximation des chaînes par un écart maximal au segment exprimé selon la métrique-sol assure une meilleure précision des structures lointaines tout en lissant correctement les chaînes plus bruitées du premier plan. L'utilisation de la métrique de la scène assure ainsi une tolérance adaptative de l'approximation [5].

- sélection des segments :

Comme pour l'élimination des chaînes parasites, on peut adapter les critères de sélection des segments à la géométrie locale des structures (longueur, orientation).

L'approche proposée donne des résultats intéressants. Elle ne modifie pas ou peu la structure des opérateurs mais seulement leurs paramètres que l'on évalue localement. Les principaux gains sont la qualité du prétraitement qui élimine les contours parasites de l'avant-plan tout en sauvegardant les vraies structures. L'extraction des chaînes et des segments est plus homogène dans la pertinence et la précision. Une alternative à une taille de voisinage trop importante est de réduire localement la zone traitée de l'image comme le propose la seconde approche. La solution multi-bandes avec filtrage de Shen-Castan adapté

est néanmoins la plus séduisante par sa simplicité et rapidité d'implémentation.

C. Adapter les données image à l'opérateur

1. Principe général

L'approche duale proposée (Figure 5) est d'adapter, donc de modifier, les données en entrée des opérateurs. Le principe général ici est de conserver l'implantation des opérateurs sans modifier en aucune façon leurs structures de contrôle et de paramétrage. Les seules transformations autorisées s'appliquent sur les données-images brutes. On peut ainsi assimiler cette approche à un prétraitement. Une implémentation multi-bandes est présentée pour optimiser les traitements. Le voisinage-image des opérateurs est fixé.

Nous proposons ici deux méthodes.

2. Mise à l'échelle des données aux opérateurs

La première méthode consiste en un réajustement de l'échelle des données obliques sans tenir explicitement compte des phénomènes de distorsion ; cette estimation du *facteur d'échelle* permet la correspondance entre les dimensions des structures d'intérêt de la scène et les dimensions des domaines d'analyse des opérateurs. Le facteur homothétique de chaque bande est estimé par le rapport de la résolution « métrique » moyenne de chaque bande avec la résolution souhaitée correspondant au domaine d'analyse de l'opérateur. Cette opération est suivie de l'extraction des primitives, puis de la restauration de l'échelle d'origine à la segmentation.

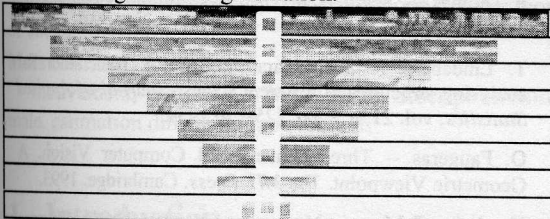


Figure 12 : Mise à l'échelle des structures pour chaque bande image à un opérateur de taille fixée

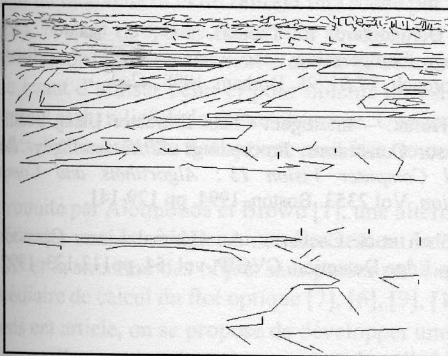


Figure 13 : Segmentation s2 reconstruite (bandes)

Cette méthode présente plusieurs avantages tels que la diminution du temps de calcul pour la plupart des bandes, le filtrage des détails inutiles de la scène, l'agrandissement

des détails lointains afin de satisfaire au domaine d'analyse fixé de l'opérateur.

3. Mise à l'échelle des données aux opérateurs avec correction des distorsions

La seconde méthode procède par une inversion de perspective restaurant *au mieux* la géométrie des données obliques. En complément du facteur d'échelle, les effets de distorsion sont alors pris en compte. Un redressement transforme l'image oblique en une pseudo ortho-image de la scène, (i.e. une vue verticale à l'infini de la scène). Cette technique courante en mosaïquage de vues [3], n'est pas appliquée, à notre connaissance, ni sur des vues frontales, ni en extraction de primitives. L'intérêt majeur de cette opération est la restauration de la géométrie et de l'isotropie des structures de la scène. Nous pouvons ainsi espérer garantir la validité des opérateurs 2D classiques de segmentation. Ses inconvénients sont la sensibilité de la reconstruction à la précision des CPDV, au caractère "plan" de la scène, et à l'éventuelle forte charge de mémoire et de calcul nécessaire.

La zone de redressement est déterminée par la profondeur maximale souhaitée ou les bornes de découpage de chaque bande. Chaque pixel redressé est interpolé ou moyenné selon le pas de résolution de l'image redressée.

a) Stratégies proposées

Traitement sur l'image "entière" : La Figure 14 illustre le gain en distribution (orientation, longueur), pertinence, détection (piste, coins) et précision des primitives extraites. Cependant, la résolution très variée de l'image oblique empêche un paramétrage unique des opérateurs et oblige à un adaptation plus locale des opérateurs par un découpage en bandes.

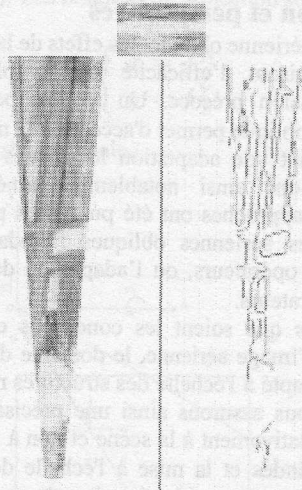


Figure 14 : Extraction de contours sur image redressée

Traitement multi-bandes : Nous proposons donc un traitement par bandes de l'image redressée adapté à leur résolution locale (Figure 15). Le paramétrage des

opérateurs est constant par bande par le découpage effectué et le pas résolutif de reconstruction choisi.

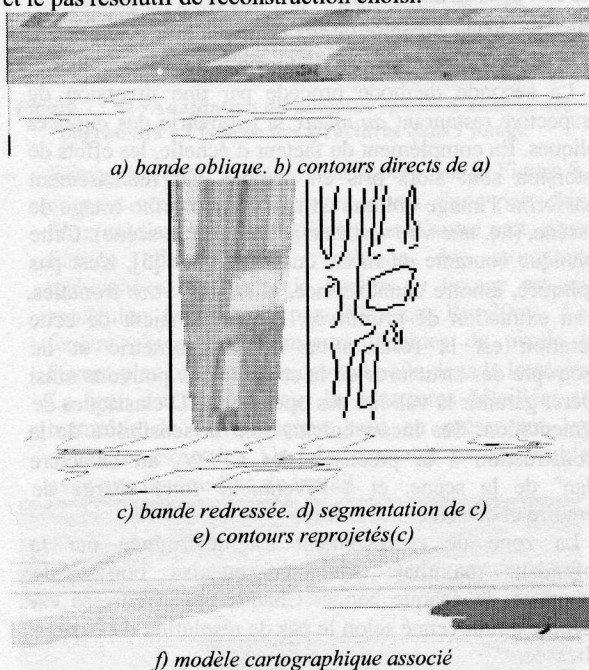


Figure 15 : approche par inversion perspective

La restauration de la géométrie et la mise à l'échelle améliorent l'extraction des primitives (Figure 15). En effet, la piste apparaît bien plus nette sur l'image redressée, facilitant son extraction. La reprojection des contours extraits (e) sont de meilleure qualité que ceux issus directement de la bande oblique (b).

VI. Conclusion et perspectives

En imagerie aérienne oblique, les effets de la projection perspective réduisent l'efficacité de la plupart des opérateurs de vision précoce. Un modèle perceptif de l'image en visée oblique permet d'accéder à la métrique de la scène autorisant une adaptation locale des opérateurs classiques qui ont ainsi notablement amélioré leur efficacité. Deux approches ont été proposées pour mieux traiter les images aériennes obliques : l'adaptation du paramétrage des opérateurs, ou l'adaptation des données d'entrée aux opérateurs.

Ainsi, quelles que soient les conditions connues de prise de vue de l'image aérienne, le domaine d'analyse de l'opérateur est adapté à l'échelle des structures recherchées dans l'image. Nous assurons ainsi une précision globale des primitives relativement à la scène et non à l'image. Le découpage en bandes et la mise à l'échelle des données entraîne généralement une réduction du volume de données à traiter et donc assure un temps de traitement aussi rapide qu'un schéma classique.

La technique multi-bandes sur images redressées nous semble prometteuse car rapide, simple et efficace, malgré

certaines difficultés qu'il reste à franchir. Il reste encore à évaluer l'ensemble de ces approches. La comparaison entre les différentes segmentations des méthodes présentées est en cours d'étude et se base sur des images de synthèse pour vérifier la précision et la pertinence des primitives extraites (précision relative à la métrique de la scène) sous différentes CPDV. Une solution hybride entre ces approches semble se dessiner.

D'autres opérateurs non cités ici ont été développés selon la même stratégie et ont révélé de bons résultats (réduction des effets atmosphériques, corrélation, extraction de cibles par morphologie mathématique, détection de mouvement et de changement,...).

Références bibliographiques

- [1] Deriche, R. -- Using Canny's criteria to derive a recursively implemented optimal edge detector. *IJCV*, 1987. vol.1(2), pp. 167-187.
- [2] J.P. Cocquerez and S. Philipp -- Analyse d'images : filtrage et segmentation. *Ed. Masson*, 1995.
- [3] W.K. Pratt. -- Digital Image Processing. *A Wiley-Interscience publication. 2nd edition*, 1991, ISBN 0-471-85766-1.
- [4] R. Horaud et O.Monga -- Vision par ordinateur, outils fondamentaux. 2^{de} édition. *Ed.Hermès*. Paris. *Traité des Nouvelles Technologies. Série Informatique*. 1995.
- [5] R Thomas. -- Segmentation d'images obliques guidées par le contexte. *AGIS'97*. Angers. déc 1997. pp257-260.
- [6] T. Lindeberg. -- Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, vol. 21, n°2, pp. 225-270, 1994.
- [7] O. Faugeras -- Three Dimensional Computer Vision. A Geometric Viewpoint. *The MIT Press*. Cambridge. 1993.
- [8] Haralick, R.M. -- Monocular Vision Using Inverse Perspective Projection Geometry: Analytic Relations. *CVPR '89*. Jun 1989, pp 370-378.
- [9] Y. Dieulangard, R. Thomas, and G. Stamon. -- FLIR Image Compression : the Accuracy of Analysis/Synthesis Approach. *VLBV'98*. Durham. 1998.
- [10] R. Horak. -- Intelligent Chain Extractor Using Scene and Sensor Constraints. *Proceedings SPIE Intelligent Robots and Computer Vision 13 : Algorithms and Computer Vision*. Vol 2353. Boston. 1994, pp 129-141.
- [11] J. Shen et S. Castan. -- An Optimal Linear Operator for Step Edge Detection, *CVGIP*, vol. 54, pp112-133. 1992