ÉTUDE DE L'INFLUENCE DES PARAMÈTRES D'ACQUISITION ET DE TRAITEMENT DES IMAGES SUR LA PRÉCISION DE LA CORRÉLATION AUTOMATIQUE

Par ***P. ANTONIO , G. BORDES , E. BRETON , G. KUNTZ , C. LATRY , M. PAUSADER***
- CNES TOULOUSE FRANCE -

Résumé

La prochaine génération de satellites de télédétection (tels que Spot 5 et Landsat 7) sera partiellement dédiée à l'acquisition de couples stéréoscopiques à haute résolution spatiale. La méthode d'acquisition le long de la trace facilitera l'acquisition des paires stéréos tout en améliorant leur qualité . La corrélation automatique des couples stéréos sera de plus en plus utilisée pour automatiser la production des modèles numériques de terrain. L'élaboration d'un MNT nécessite un processus complexe et la qualité du produit final dépend d'un grand nombre de paramètres parmi lesquels :

- des paramètres liés au paysage
- des paramètres radiométriques
- des paramètres géométriques
- des paramètres liés aux traitement

Certains de ces paramètres peuvent être ajustés pour optimiser le processus et le résultat . Nous avons développé une méthode afin d'évaluer la sensibilité des différents paramètres. Nous avons simulé les images de différents paysages en utilisant les caractéristiques des instruments de Spot 5. La deuxième image du couple stéréo est obtenue en appliquant à l'image simulée une translation d'une fraction de pixel . Le processus d'appariement permet alors de calculer la valeur du décalage appliqué . L'effet des paramètres suivants a été testé :

- la fonction de corrélation
- l'interpolateur
- le rapport signal sur bruit
- la fonction de transfert de modulation
- la bande spectrale

Abstract

The next generation of high-resolution remote sensing satellites (such as Spot 5 and Landsat 7) will emphasize the acquisition of stereoscopic pairs . Along-track stereoscopy will enhance the quality of stereo pairs while easing their capture . The automatic matching of stereo pairs is currently used and will be more and more in order to produce digital terrain models in an operational way . The DTM extraction is a rather complex process encompassing several steps and the accuracy of the final product depends on a large number of parameters including :

- landscape parameters
- radiometric parameters
- geometric parameters
- ground processing methods

Some of these parameters can be tuned in order to optimise the matching process .To increase our knowledge about the sensitivity to the different parameters we have defined a method. We have simulated the images of various landscapes by using the characteristics of the Spot 5 instruments .The second image of the stereoscopic pair is obtained by applying to the simulated image a shift of a fraction of pixel . The stereomatching is then used in order to compute the value of the shift applied . The effect of the following parameters has been tested :

- the correlation function
- the interpolation function
- the signal to noise ratio
- the FTM value
- the spectral band

INTRODUCTION

Parmi les applications de la télédétection spatiale, l'utilisation de couples d'images stéréoscopiques prend une place grandissante . Le système SPOT permet opérationnellement l'acquisition de couples stéréoscopiques grâce au dépointage latéral de ses miroirs. Cette technique permet d'obtenir une grande diversité de conditions d'acquisition stéréoscopique , chacune des 2 images pouvant être obtenue avec un angle d'incidence variable de - 29° à + 29° par pas de 0,6° . On obtient ainsi des rapports B/H pouvant varier de 0 à 1,2 , les couples pouvant être acquis aussi bien en configuration symétrique que dissymétrique .

Par contre un inconvénient de ce mode d'acquisition tient au fait que les deux images ne peuvent être acquises au cours de la même séquence de prise de vues. Elles seront ainsi acquises à un ou plusieurs jours d'intervalle. Outre la complexité de programmation induite (car l'accessibilité d'une zone varie selon la latitude et le B/H recherché), l'acquisition à des dates différentes entraîne une dissemblance plus ou moins importante des images : conditions atmosphériques différentes (nuages, brumes, vents de sable, ...), ou

modification du paysage (conditions différentes de la végétation et du sol : humidité superficielle, neige , ...). Une fois qu'une première image de qualité correcte est acquise , il peut être éventuellement nécessaire d'effectuer plusieurs acquisitions avant d'obtenir la seconde (en cas de couverture nuageuse par exemple) , ce qui accroît encore le décalage temporel et donc la probabilité de dissemblance des deux images .

Pour pallier à ces difficultés , SPOT 5 propose un nouveau mode stéréoscopique : la stéréoscopie avant-arrière. Grâce à un dépointage longitudinal des lignes de visée des deux instruments les deux images seront acquises quasi-simultanément au cours de la même orbite (moins de deux minutes d'intervalle) , ce qui remédie aux défauts de diachronisme et simplifie la programmation .

L'une des applications importantes de la stéréoscopie est la génération de modèles numériques de terrain par corrélation automatique . La qualité des modèles numériques de terrain obtenus dépend comme nous allons le voir de nombreux paramètres . L'objet de ce travail qui est effectué dans le cadre des études d'avant-projet du satellite SPOT 5 est d'évaluer la sensibilité de la précision du résultat à la valeur de ces paramètres afin d'optimiser les instruments de prise de vue et les méthodes de traitement.

◆ PARAMÈTRES CONDITIONNANT LA PRÉCISION DES MODÈLES NUMÉRIQUES DE TERRAIN

Les modèles numériques de terrain sont obtenus à l'issue d'un processus complexe comprenant l'acquisition et le traitement des images d'une zone donnée . On peut identifier quatre familles de paramètres conditionnant la précision des modèles numériques de terrain . Ces quatre familles sont liées respectivement :

- au paysage observé
- aux instruments de prise de vue
- aux conditions géométriques d'acquisition
- à la méthode de traitement

Une liste non exhaustive de ces paramètres peut ainsi être établie :

- paramètres liés au paysage :
  . relief
  . texture
  . variabilité du paysage dans le temps
  . conditions atmosphériques de la prise de vue
  . présence de superstructures artificielles (bâtiments) ou naturelles (arbres)

- paramètres liés aux instruments de prise de vue :
  . taille du pas au sol
  . Fonction de transfert de modulation (FTM)
  . rapport signal sur bruit
  . méthode et taux de compression
  . nombre de bits de codage
  . bande spectrale

- paramètres liés aux conditions géométriques d'acquisition
  . B/H
  . dissemblances dues aux angles de prises de vue (effets stéréoradiométriques, dissymétrie du couple...)

- paramètres liés à la méthode de traitement :
  . opérateur de corrélation (fonction , calcul mono ou bidimensionnel , taille de la fenêtre , ...)
  . méthode de validation
  . fonction d'interpolation (recherche du maximum de corrélation)
  . modèle géométrique de la prise de vue
  . méthode de «remplissage des trous «

Il est nécessaire pour dimensionner le système d'évaluer l'influence de ces paramètres sur la qualité du résultat . Nous avons développé pour ce faire une méthode explicitée ci-dessous . Cette méthode a été appliquée selon deux axes : l'évaluation de la sensibilité des paramètres instrumentaux et des paramètres du traitement . Elle a été testée sur plusieurs types de paysages . L'influence des paramètres liés aux conditions géométriques d'acquisition n'est pas traitée dans cette communication .

◆ MÉTHODE DE MESURE DE LA SENSIBILITÉ DES PARAMÈTRES

Les images testées ont été obtenues en simulant finement le processus d'acquisition des instruments de SPOT 5 . Il est nécessaire pour ce faire d'utiliser des images sources plus résolues qui représentent le paysage réel acquis par un instrument parfait . Nous nous sommes servi d'images source trois fois plus résolues .

Un MNT est déterminé en mesurant pour chaque pixel image la valeur de la parallaxe stéréoscopique qui est proportionnelle à l'altitude du point considéré et au rapport B/H . Pour un terrain réel non plat cette parallaxe varie donc selon le pixel considéré . Afin de simplifier le problème , nous avons introduit une parallaxe fictive constante en translatant la première image d'une translation sub-pixellaire de valeur connue (un nombre entier de pixels de l'image source) . Le couple stéréoscopique est donc formé d'une image et de sa translatée . Les expériences consistaient à mesurer cette parallaxe ; la comparaison entre parallaxe introduite et parallaxe mesurée donne la précision du résultat. Les avantages de cette méthode sont multiples :

- maîtrise et connaissance parfaite de la parallaxe introduite
- mesure très précise de la qualité du résultat puisqu'il est possible d'effectuer des statistiques sur un nombre très important de pixels ; en d'autres termes puisque on mesure un grand nombre de fois la même chose , on peut donc observer la dispersion de cette mesure.
- la valeur de la parallaxe à mesurer est constante dans toute fenêtre de corrélation , ce qui garantit une meilleure corrélation et autorise l'utilisation de fenêtres de grande taille .
- la ressemblance des deux images est maximale puisque la deuxième est obtenue à partir de la première ; on s'affranchit en particulier des paramètres liés aux

conditions géométriques de prise de vue et de l'évolution du paysage dans le temps .

Par contre il ne faut pas oublier que les précisions obtenues sont alors théoriques et bien supérieures à ce qui sera obtenu dans les cas réels.
Plusieurs «paysages» ont été ainsi observés :

- un signal aléatoire monodimensionnel de spectre blanc
- des images à haute résolution acquises par le push-broom aéroporté du CNES dont les bandes spectrales sont celles de SPOT ; ces images (nommées par la suite Aix , Forêt et Landes) sont acquises sur des sites variés :

paysage périurbain (Aix en Provence)
forêt des Landes
zone agricole dans le département des Landes

Au cours de la procédure de simulation il est possible notamment :

- de faire varier le niveau de bruit (y compris le nombre de bits de codage)
- de faire varier la valeur de la FTM à la demie-fréquence d'échantillonnage
- d'introduire une translation subpixellaire sur la deuxième image.

◆ RÉSULTATS

▮ Sensibilité des paramètres de traitement

Un MNT est calculé en trois étapes : mesure des parallaxes , transformation de ces parallaxes en altitudes et «remplissage des trous" pour les zones non corrélées . La première étape peut être réalisée par corrélation automatique ; la deuxième étape est la modélisation géométrique , la troisième étape s'effectue par des méthodes d'interpolation. Nous n'étudions ici que la première étape .

La corrélation peut elle-même être décomposée en deux étapes : le calcul de la fonction de corrélation discrète et l'interpolation de cette fonction afin d'en déterminer le maximum local . Nous avons étudié ces deux étapes

- Influence de l'opérateur de corrélation

L'opérateur de corrélation sert à déterminer pour chaque point de l'une des deux images dite de référence la position de son homologue dans l'autre image . Le problème est donc bidimensionnel . Afin d'accélérer le traitement , on se ramène généralement à un problème monodimensionnel en utilisant un rééchantillonnage selon des lignes dites pseudo-épipolaires . La question ne se pose pas dans notre cas puisque nous introduisons seulement une parallaxe dans le sens des lignes Nos images sont donc naturellement en géométrie épipolaire .

Plusieurs fonctions de ressemblance ont été proposées dans la littérature . Nous en avons testé quatre :

- coefficient de corrélation linéaire :

$$CO3(x_p, y'_p) = \frac{1}{p}\sum_{i=0}^{i=p-1}\frac{(x_p(i)-\overline{x_p}).(y'_p(i)-\overline{y'_p})}{\sqrt{s^2(x_p).s^2(y'_p)}} = \frac{r_{xy'}}{s_x s_{y'}}$$

$$r_{xy'} = E\left[(x-\overline{x})(y'-\overline{y'})\right]$$

- somme des valeurs absolues des différences ;

$$CO1(x_p, y'_p) = \sum_{i=0}^{i=p-1}\left| x_p(i) - y'_p(i)\right|$$

Cette fonction mesure en fait la norme (dite «L1») du vecteur différence.

- somme des carrés des différences:

$$CO2(x_p, y'_p) = \sum_{i=0}^{i=p-1} \left(x_p(i) - y'_p(i) \right)^2$$

Celle-ci mesure la norme quadratique du vecteur différence (norme dite «L2»).

- corrélateur par transformée de Fourier :

$$CO6\left(x_p(k), y'_q(k)\right)(j) = \sum_{i=0}^{i=p-1} \left(\overline{x_p(i)} \cdot y_q(i+j) \right)$$

Les meilleurs résultats ont été obtenus avec le corrélateur classique (coefficient de corrélation linéaire) qui présente en outre l'avantage de ne pas être affecté par des écarts linéaires de radiométrie entre les deux images (qui peuvent être dus par exemple à l'utilisation de gains électroniques différents ou à des étalements de dynamique) .

- **Influence de l'interpolateur**

La phase précédente fournit une fonction de corrélation discrète au pas de un pixel . Cette précision est insuffisante mais peut être améliorée en interpolant le graphe de corrélation et en déterminant le maximum de la courbe interpolée . Il faut tout d'abord localiser un pic candidat (la fonction pouvant avoir plusieurs pics) , le valider puis utiliser une fonction d'interpolation .
La validation a été effectuée en invalidant les cas où le minimum discret est trouvé avec un coefficient de corrélation inférieur à un seuil donné et quand la valeur de la parallaxe diffère de plus de 0,5 pixel de la valeur théorique .

Pour ce qui est de la fonction d'interpolation appelée couramment interpolateur , plusieurs fonctions sont proposées dans la littérature . Nous en avons testé quelques-unes et nous avons également proposé un nouvel interpolateur appelé «modèle de pic « . Les fonctions d'interpolation testées sont les suivantes :

- fonction spline
- interpolateur par série de Fourier
- polynômes de Lagrange
- modèle de pic

Le modèle de pic est particulièrement adapté aux signaux dont le spectre est blanc (c'est à dire que l'amplitude des composantes fréquentielles est constante) . En effet on peut montrer que dans le cas d'un signal blanc la transformée de Fourier de la fonction d'autocorrélation du signal est proportionnelle au carré de la FTM . Comme nous utilisons pour générer les images simulées une modélisation analytique de la FTM (modèle représentatif des instruments HRV) , il est possible d'en déduire la forme réelle de la fonction de corrélation . Il suffit ensuite d'ajuster au mieux la fonction théorique sur la fonction discrète , cet ajustement étant lui-même réalisé par corrélation entre les deux courbes . Une fois l'ajustement effectué , on lit directement sur la fonction théorique la position du maximum

Le modèle de pic s'est révélé effectivement le plus efficace dans le cas du signal blanc . Par contre dans le cas des images réelles , il s'est révélé moins efficace . En effet on verra plus loin que le spectre des images réelles utilisées était loin d'être blanc . Si quelqu'un connaît un paysage dont le spectre est blanc veuillez nous le signaler ... . Toutefois dans le cas du signal blanc on a ainsi une évaluation de la précision de la corrélation proprement dite et aussi de la dégradation amenée par des interpolateurs non optimaux (cf. figure 1) .

Sur les images l'interpolateur le plus intéressant s'est révélé être le polynôme de Lagrange de degré 4 .

Influence des paramètres instrumentaux

Les paramètres suivants ont été testés :
- rapport signal sur bruit
- FTM
- bande spectrale

- **Sensibilité au rapport signal sur bruit**

Les différents bruits instrumentaux ont été simulés :

bruit de quantification, bruits électroniques et bruits liés aux détecteurs. Le rapport signal sur bruit (S/B) est calculé au maximum de la dynamique de codage. Nous avons fait varier le rapport signal/bruit entre 10 et une valeur maximum qui correspond au fait de n'avoir plus que le bruit de quantification. La figure 2 montre la valeur de l'écart-type en fonction de S/B pour les trois paysages réels et le tableau 1 indique le pourcentage de mauvaises corrélations. On observe une décroissance puis une stabilisation de l'écart-type quand S/B augmente ; le pourcentage de fausses corrélations décroît quand S/B augmente. La précision la meilleure est obtenue pour l'image d'Aix, puis pour la forêt, puis pour «landes». La valeur du plateau est atteinte plus rapidement pour AIX (S/B=100) que pour les 2 autres (S/B =150). Quand le plateau est atteint le pourcentage de fausses corrélations devient très faible. On observe donc que d'une part le résultat dépend du paysage et d'autre part pour les niveaux de S/B prévus pour Spot 5 (supérieur à 200) on est dans la zone de plateau pour les différents paysages.

- Sensibilité à la fonction de transfert de modulation

La FTM a été appliquée en utilisant un modèle analytique développé pour SPOT. Le paramètre variable est la valeur prise à la demie-fréquence d'échantillonnage. Cinq valeurs de la FTM ont été testées : 0,64 (valeur correspondant à l'intégration sur un détecteur parfait), 0,3 (objectif pour SPOT 5), 0,15 et 0,05. On observe également que l'effet de la FTM dépend du paysage. Le tableau suivant indique la FTM optimale pour chaque paysage.

Image	FTM optimale
Signal blanc	0,10
AIX	0,15
Forêt	0,30
Landes	0,64

Ce résultat étonnant a priori peut s'expliquer par l'effet du repliement de spectre. En effet les hautes fréquences parasites dues au repliement de spectre sont d'autant plus importantes que l'image sera plus riche en hautes fréquences. En observant les spectres des images à haute résolution à partir desquelles ont été produites les simulations, on constate effectivement que l'amplitude des hautes fréquences est la plus importante pour Aix, puis pour forêt puis pour «landes» (le bruit blanc représentant le cas extrême).

Cet effet a pu être confirmé en supprimant par traitement le repliement de spectre. Pour ce faire l'image source a été filtrée de manière à éliminer les composantes de fréquence supérieure à la demie-fréquence d'échantillonnage. On observe alors que la précision croît avec la valeur de la FTM et on peut mesurer la perte due au repliement (figure 3).

- Influence de la bande spectrale

Deux des bandes spectrales prévues pour Spot 5 ont été évaluées : le panchromatique et le proche infrarouge (B3). On observe également que le résultat dépend du paysage (figures 4 et 5). Pour les paysages forêt et «landes» le résultat est meilleur en proche infrarouge ; il est légèrement meilleur en panchromatique pour Aix. Ce résultat n'est pas étonnant, le proche infrarouge étant plus sensible dans les zones à fort couvert végétal. Il convient toutefois de préciser que les performances en rapport signal sur bruit et FTM seront moins bonnes pour les instruments de Spot 5 dans le proche infrarouge qu'en panchromatique.

◆ CONCLUSIONS ET PERSPECTIVES

La méthode développée permet de comprendre l'effet de différents paramètres sur la performance de corrélation automatique et de tester leur impact. Dans la perspective de la définition des instruments de Spot 5 on retiendra qu'il n'est pas possible de définir un instrument optimal pour tous les types de paysage mais que les valeurs prévues de 200 pour le rapport signal sur bruit et de 0,3 pour la FTM constituent un bon compromis. Les essais vont être poursuivis sur d'autres types de paysage.

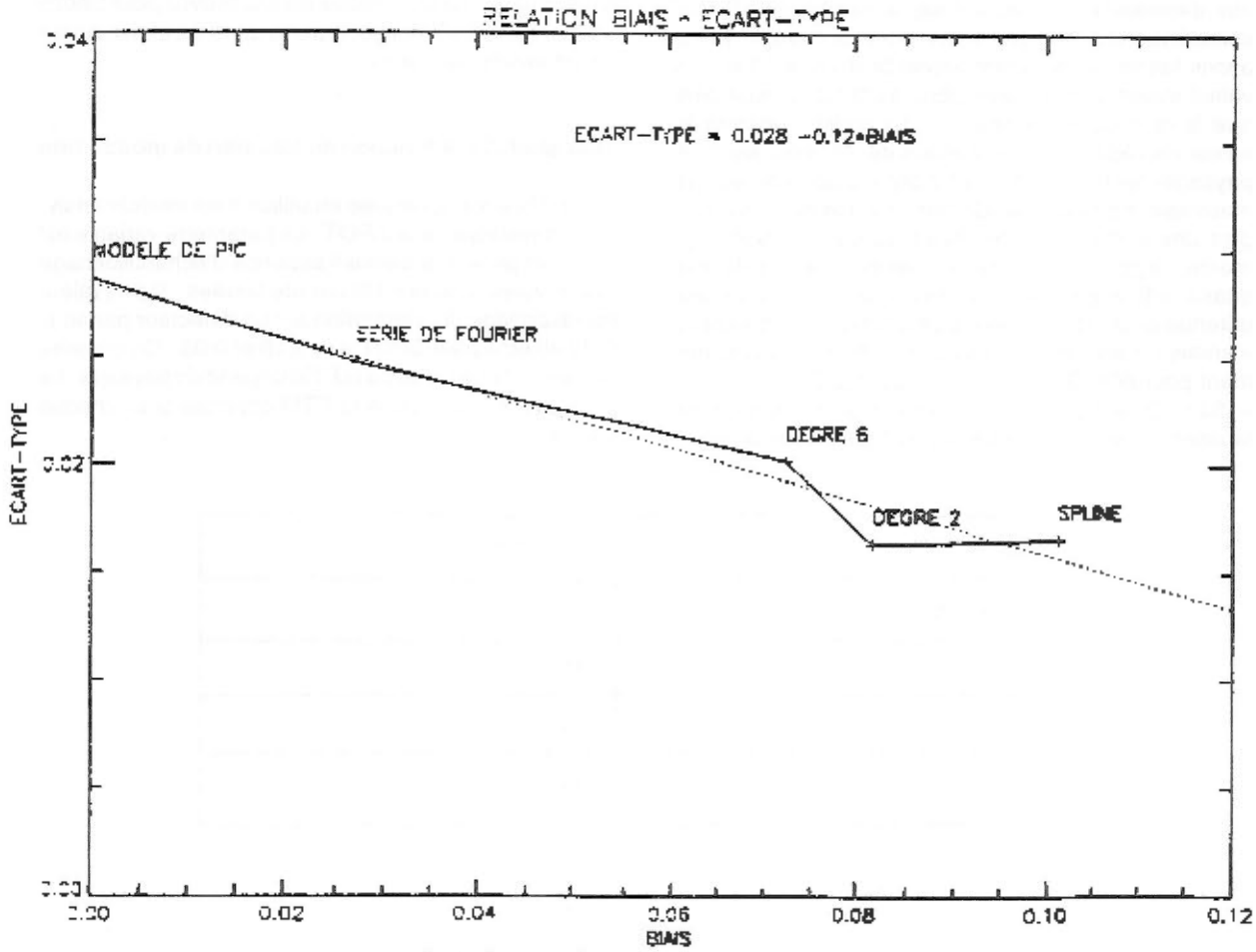

fig 1: Comparaison des interpolateurs

S/B	Aix	Forêt	Landes
10	13.7%	34.2%	96.9%
50	0	7.6%	36.4%
100	0	3%	20%
150	0	0	2.7%
200	0	0	9%
250 et plus	0	0	0

tableau 1 : Pourcentages de mauvaises corrélations selon le paysage

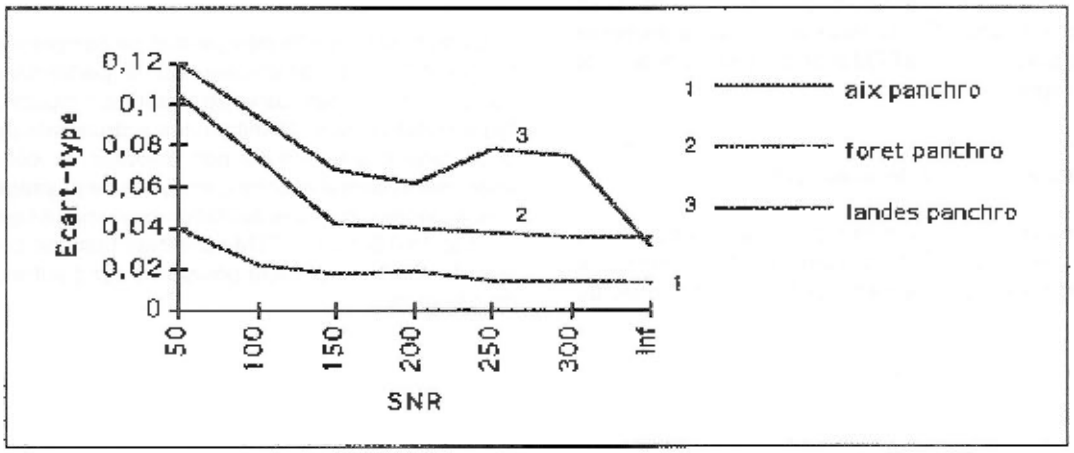

fig 2 : Influence du rapport signal sur bruit pour les trois paysages

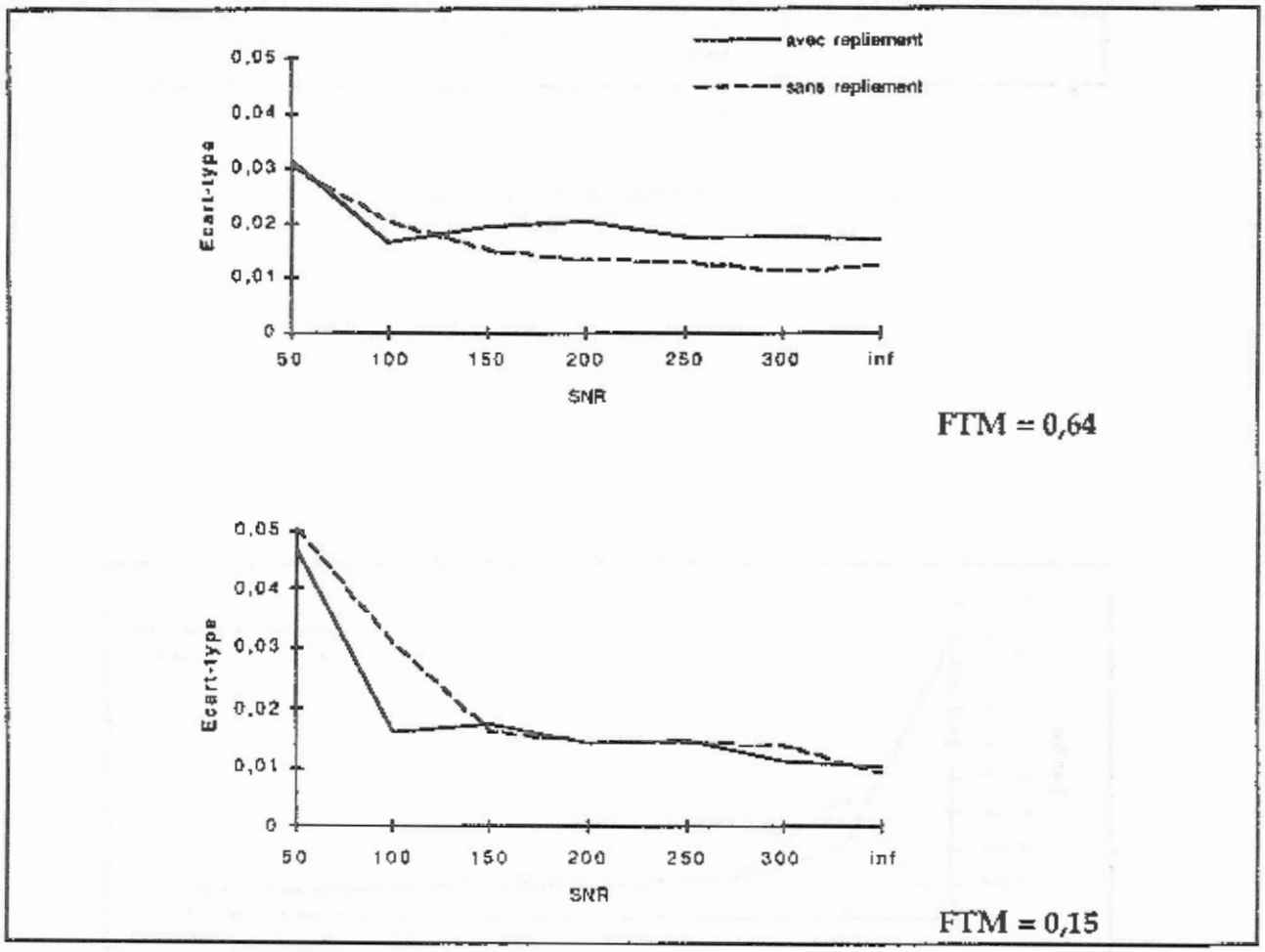

Fig. 3 : Influence du repliement de spectre pour "Aix" en panchromatique

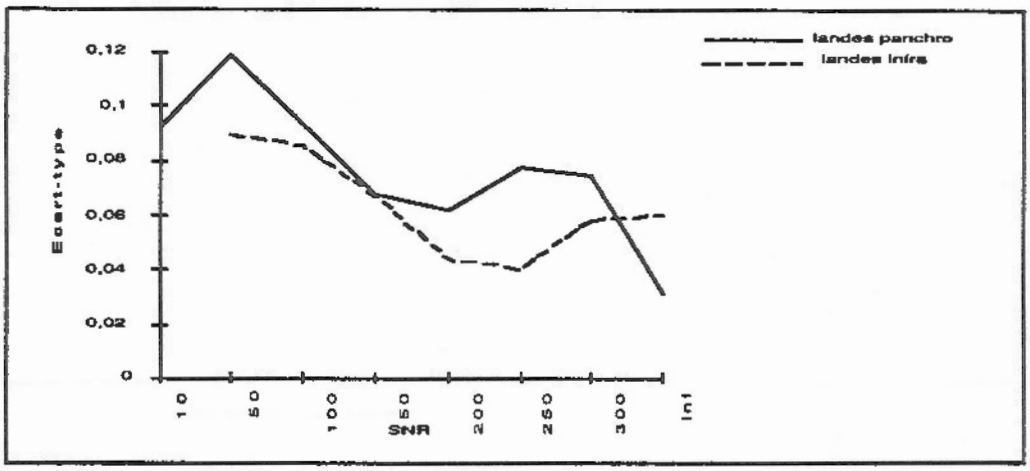

Fig. 4 : Influence de la bande spectrale pour "Landes"

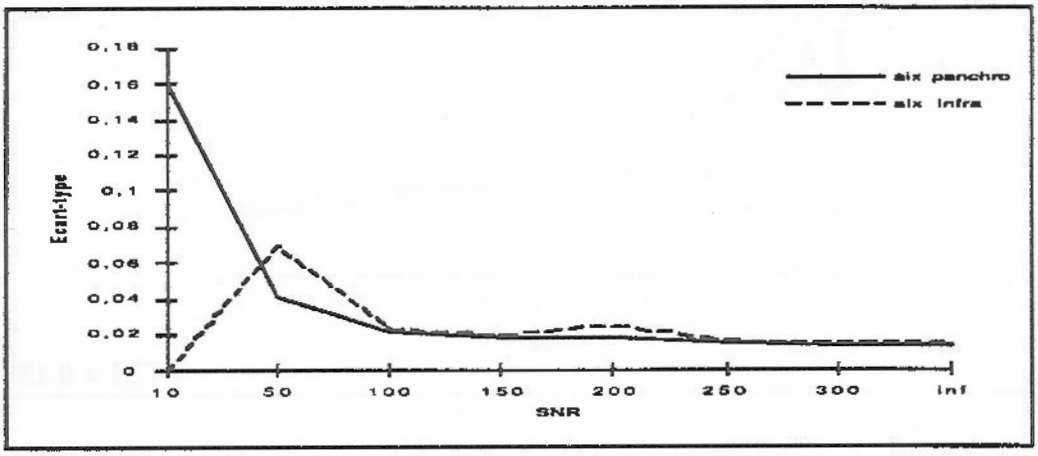

Fig. 5 : influence de la bande spectrale pour "Aix"