

Généralisation d'images satellites classées pour leur intégration à un SIG

Vincent SIMONNEAUX

Orstom, 72 route d'Aulnay, 93143 Bondy cedex, France.

Résumé

Nous présentons ici deux méthodes de généralisation d'une image classée, pouvant notamment servir de préalable à sa vectorisation pour son utilisation dans un SIG vectoriel. Cela permet de réduire le volume des données à vectoriser et d'homogénéiser les précisions des différentes sources d'informations contenues dans le SIG (image de résolution 30 m pour une cartographie à 1/200 000). La généralisation présentée ici consiste à faire disparaître les petites plages et à simplifier les contours restants. L'exemple traité concerne la constitution d'une base de données sur les sols au Togo, destinée notamment à l'évaluation des ressources en terre et de leur dégradation (projet R2T, « Ressources en Terres du Togo »).

Les algorithmes proposés sont (a) l'agrégation de petites plages dispersées en plages « synthétiques » plus grandes et (b) la suppression pure et simple des petites plages. Ils sont basés sur des opérateurs de morphologie mathématique.

La première approche concerne les classes représentées par des plages de surface faible mais que leur juxtaposition sur de grandes étendues rend significatives (ex : cultures dispersées dans la savane). On agrège la classe initiale en fonction de sa densité (ex : cultures seules, cultures et savane seule).

La deuxième approche consiste à supprimer les plages en dessous d'une certaine épaisseur, puis à boucher les lacunes obtenues par la classe la plus proche.

L'implémentation de ces deux algorithmes est fortement automatisée grâce aux possibilités de macroprogrammation des logiciels PLANETE et OSIRIS, développés à l'Orstom.

Ces deux méthodes ont donc répondu à quelques-uns des problèmes typiquement posés par l'intégration télédétection-SIG, en permettant sans intervention manuelle fastidieuse d'incorporer des résultats de télédétection à une base de données vectorielle constituée.

Mots-clés : généralisation, image satellite, SIG, morphologie mathématique, classification.

Abstract

Generalization of classified satellite images for integration into a GIS.

Presented here are two methods for the generalization of a classified image as a preliminary step to converting the image into the vector format required to enter it into a vector-based GIS. The operation reduce the volume of data to be converted and results in a more uniform accuracy of the data sets (30 m pixels is too accurate for a 1:200,000 database). The generalization concerned here consist of removing the smallest patches from the image and smoothing the outlines of the remaining patches. The example taken concerns a database on the soils of Togo, designed for land resource and degradation assessment (Project R2T, *Land resources of Togo*).

The algorithms shown are for (a) the aggregation of small, scattered patches into larger, "synthetic" ones and (b) the elimination of small patches. Both are based on mathematical morphology operators.

The first algorithms is used for classes consisting of small patches that are nonetheless significant because scattered over a large area — e.g. : scattered patches of cropland in savanna. The initial class is aggregated according to density (ex : crops only, crops present, or savanna only).

The second algorithms is used to eliminate patches below a certain thickness, filling in the resulting gaps with the nearest other class.

Implementation of the algorithms was handled virtually automatically, thanks to the macro-programmation capacities of the PLANETE and OSIRIS software, packages developed by Orstom.

These two procedures resolve some of the classic problems that arise when using both GIS and remotely sensed data, providing a way of incorporating raster-based data into a existing GIS vector database, without tedious manual processing.

Keywords : generalization, satellite image, GIS, mathematical morphology, classification.

Introduction

L'imagerie satellitaire constitue une source privilégiée d'information spatialisée sur l'environnement physique et elle est donc destinée à s'intégrer de plus en plus aux Systèmes d'Informations Géographiques (SIG). Cette association bénéficie aux deux parties. Ainsi, le traitement de l'image peut s'appuyer sur des informations exogènes archivées dans le SIG, qui fournissent un contexte à l'analyse de l'image (ex : probabilités *a priori* pour une classification). Plus simplement, l'habillage d'une image satellite par un fond topographique est souvent indispensable pour une bonne lisibilité de celle-ci. En retour, l'information issue du traitement de l'image ira naturellement enrichir la base de données gérée par le SIG.

Toutefois, la combinaison d'informations spatialisées de nature différente pose de nombreux problèmes méthodologiques, plus ou moins spécifiques des thématiques étudiées, et qui concernent le choix des données à utiliser et la manière de les combiner (ex : les équations empiriques permettant d'évaluer les pertes en sols par érosion). On se heurte également très souvent à des problèmes techniques, qui cachent eux-mêmes souvent des problèmes méthodologiques. C'est typiquement le cas pour les transferts de format, notamment le passage de l'image *raster* au plan vectoriel et vice versa (ESTES et *al.*).

Nous présentons ici quelques aspects techniques d'une intégration télédétection-SIG, en prenant comme exemple la constitution d'une base de données sur les sols du Togo destinée notamment à l'évaluation des ressources en terre et de leur dégradation (Projet R2T « Ressources en terres du Togo », BRABANT et *al.*). Une classification de l'occupation du sol est obtenue sur l'ensemble du pays à partir d'une couverture Landsat TM (4 scènes complètes plus 1/4 de scène). Ce thème est essentiel pour évaluer ensuite la dégradation de terres. Nous montrons ici comment cette classification est « mise en forme » pour pouvoir être ultérieurement transformée en plan vectoriel géré par un SIG (« vectorisation », voir figure 1).

1. Vectorisation d'une image classée

1.1. Origine de l'image classée

Les grandes classes d'occupations du sol (essentiellement savane, cultures et forêt) sont obtenues par classification automatique des canaux Landsat TM utilisables (c'est à dire non bruités ou endommagés), plus un indice de végétation (NDVI). Le choix d'une classification automatique vient de ce que les classes reconnues sont souvent très hétérogènes et correspondent plus à des continuums variables qu'à des catégories bien homogènes. Elles sont de plus difficiles à délimiter sur les images, et on aurait donc des difficultés à définir des parcelles d'apprentissage représentatives.

Les nuages¹ obtenus par agrégation automatique sont interprétés par l'utilisateur, qui détermine en fonction de sa connaissance du terrain si ils sont homogènes ou hétérogènes vis à vis des classes recherchées. Les nuages hétérogènes sont redécoupés jusqu'à obtenir des sous-nuages homogènes. Lorsque le redécoupage dans l'espace spectral ne donne pas de résultats satisfaisants, et si il s'avère que la confusion entre les classes est localement solvable, on réalisera la stratification géographique adéquate de l'image. Bien souvent, il est judicieux de réaliser cette stratification dès le départ si on voit nettement apparaître sur l'image des types d'organisation bien différenciés et spatialement bien localisés. Notons d'ailleurs que les régions apparaissant sur l'image ne correspondent pas toujours à la

¹ On appelle nuage un ensemble de pixels de l'image placés dans l'espace spectral. Le nuage initial de l'image totale est ainsi découpé en nuages (ou « sous-nuages ») que l'on peut représenter dans l'espace géographique.

géologie ou à la pédologie, mais plutôt à des combinaisons entre géologie/pédologie, climatologie et histoire agricole de la zone.

Les classifications ainsi obtenues sont alors simplifiées pour pouvoir être ensuite vectorisées et incorporées à la base de données vectorielle gérée par le logiciel ARC/INFO.

1.2. Pourquoi vectoriser ?

On suppose ici que l'image doit être vectorisée pour pouvoir être utilisée au mieux dans le SIG. Cela n'est pas *a priori* une obligation conceptuelle ; mais de fait, les systèmes actuels réalisent encore assez mal la gestion réellement conjointe des images et des plans vectoriels. D'autre part, pour des informations spatialisées synthétiques comme les classifications, le stockage vectoriel est nettement moins volumineux, et d'autant moins que l'on consent une « simplification » de cette image.

On peut néanmoins vouloir conserver dans un SIG une image « brute », par exemple sous forme d'une composition colorée, à des fins d'illustration, en considérant qu'elle véhicule plus information que ne pourrait en extraire n'importe quel traitement. Ces spatiochartes restent alors sous forme d'image, à condition que le SIG puisse en assurer une gestion minimale, et il n'est plus question alors de la simplifier.

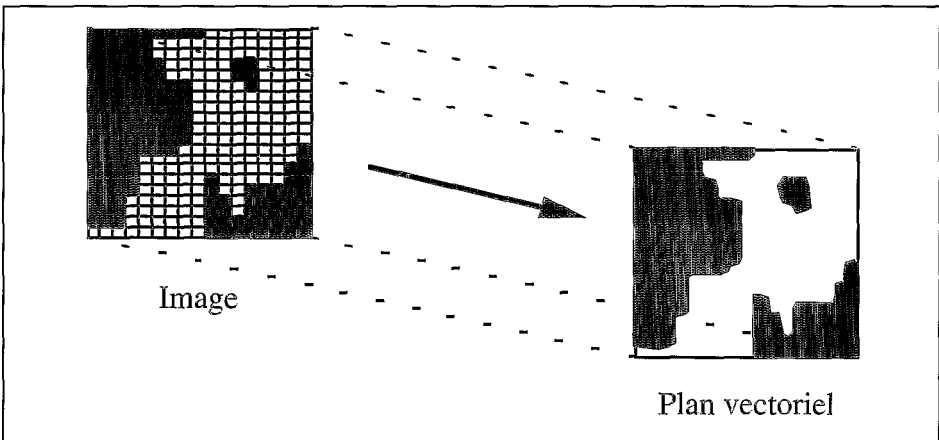


Figure 1. Exemple de vectorisation des contours d'une classification.

1.3. Pourquoi simplifier ?

La vectorisation doit le plus souvent être précédée d'une « simplification » préalable de l'image classée. Cette opération, bien connue en géomatique sous le nom de généralisation, désigne l'ensemble des opérations visant à modifier les informations contenues dans une base de données géographiques (dans le sens d'une réduction de ces informations) afin de

les adapter à une représentation à une échelle donnée (généralisation cartographique) ou bien de dériver de la base initiale un modèle spécifique de données adapté à une application thématique donnée (généralisation statistique ou conceptuelle). Des articles de synthèse permettent d'avoir une vue d'ensemble de cette discipline (MULLER, BRASSEL et *al.*, RIGAUD).

Dans notre exemple, la généralisation va consister à faire disparaître les petites plages de l'image et à simplifier les contours restants (simplification ou lissage, *stricto sensu*).

Le premier intérêt, purement technique, est de diminuer le nombre d'objets manipulés et donc le volume de ces données, qui pourrait sans cela dépasser rapidement les capacités de traitement du SIG (LUNETTA). Ceci est d'autant plus vrai que certains algorithmes de vectorisation ont tendance à créer des contours complexes qui suivent plus ou moins les bords des pixels de l'image, et contiennent donc un grand nombre de segments. Or la généralisation possible de ces contours vectoriels n'est souvent effectuée que dans un deuxième temps (SYSAME).

Une deuxième raison pour généraliser l'image est d'homogénéiser les précisions des différentes sources d'information contenues dans le SIG. Il est inutile de stocker une information très précise, donc plus lourde à manipuler, si on n'en a pas l'utilisation ou bien si les autres cartes de cette base ont une précision nettement inférieure. Dans ce cas, le croisement de ces cartes de précision différente (*overlay*), créerait des contours de précision spatiale variable, ce qui n'est pas souhaitable.

Enfin, il faut voir que cette généralisation n'est pas toujours une perte réelle d'information car la résolution importante d'une image est bien souvent une illusion quand on prend en compte l'incertitude sur le classement de chaque pixel. La fiabilité du classement ne devient correcte qu'à condition de généraliser les résultats sur un ensemble de pixels, ce qui a pour effet de compenser les erreurs. Ainsi, le thématicien ne valide généralement le résultat d'une classification que plus ou moins globalement, il néglige le « bruit » associé au pixel. Il est donc logique de se débarrasser de ce bruit par la généralisation de l'image.

Dans l'exemple étudié ici, la précision spatiale de l'information apportée par l'image (résolution 30 m) est excessive par rapport à la précision de données gérées par le SIG (saisie de cartes à 1/200 000). Il est donc bien justifié de simplifier l'image avant de la vectoriser, d'une part pour rester cohérent avec la précision du SIG, et d'autre part pour ne pas dépasser les capacités de traitement du SIG lors de la vectorisation.

1.4. Méthode

La généralisation que l'on désire effectuer ici comporte *a priori* deux aspects : la suppression des petites plages et le lissage du contour des plages conservées.

Un premier problème est de définir ce qu'est une « petite » plage, soit par sa surface totale, soit par son épaisseur² (ce qu'il en reste après x érosions³ morphologiques).

La « loi du quart » bien connue des cartographes suggère qu'une plage n'ait pas une surface inférieure à $\frac{1}{4}$ cm² (2.5 mm sur 2.5 mm). Par ailleurs, il est admis qu'une plage quelconque ne doit pas avoir une épaisseur inférieure à 2 mm. Ici, ces règles ne sont pas directement utilisables car il n'y a pas *a priori* d'échelle fixée de restitution. On cherche à conserver la précision maximale compte tenu des contraintes sus-évoquées de volume de données et de précision intrinsèque de la classification (ou du SIG à alimenter).

Un autre problème vient de ce que toutes les « petites » plages ne sont pas nécessairement à supprimer car certaines peuvent avoir une importance particulière (économique, thématique, culturelle) qui justifie de les conserver. C'est la cas ici des forêts qui ont une importance écologique particulière. Cette remarque sur l'importance variable des plages est faite également par ROBBEZ-MASSON qui évoque à ce sujet la notion de « contraste sémantique » de la plage avec son environnement.

Pratiquement, nous envisagerons deux manières de supprimer des petites plages, par leur **élimination** (*stricto sensu*) pure et simple ou par leur **agrégation** (*stricto sensu*) en plages « significatives ». Les deux approches envisagées permettront simultanément de lisser les contours de toutes les autres plages (Fig. 2).

1.4.1. Agrégation de classes

Certaines classes sont représentées par des plages de faible surface mais que leur juxtaposition sur de grandes étendues rend significatives. C'est typiquement le cas ici pour les cultures qui sont présentes en quantité variable dans la savane, réalisant un mitage de celle-ci. On observe ainsi un passage progressif entre deux pôles, la savane pure et les cultures pures. Or la densité des cultures est un paramètre important pour l'évaluation de la dégradation des terres.

Ce type d'agencement spatial étant difficile à gérer dans des systèmes vectoriels, on aimerait pouvoir synthétiser l'information en agrégeant les pixels en néo-classes en fonction des densités respectives des deux classes initiales (exemple : cultures denses (pas de jachères) / cultures peu dense (jachères) / savane non cultivée). Or cette densité de cultures est un paramètre important pour l'évaluation de la dégradation des terres dans la mesure où elle est inversement proportionnelle à l'importance de la jachère. Nous proposons dans la suite une méthode permettant de réaliser cette agrégation.

2 Ce terme est à prendre dans un sens purement géométrique et n'a pas de rapport avec son équivalent pédologique.

3 Opération de base en morphologie mathématique. L'érosion d'une plage revient à « l'amincir » en lui retirant les pixels de bordure. Ceci s'obtient pratiquement par l'application d'un filtre « minimum » sur une image binaire (0-1). L'opération inverse est la dilatation (filtrage « maximum »)

1.4.2. Élimination de petites plages

L'importance des petites plages est fonction de la thématique étudiée. Ainsi, les plages allongées, de surface pas nécessairement faible, mais peu épaisses, peuvent être dues à la présence de pixels mixtes dans l'image, auquel cas on souhaitera les éliminer. Il peut s'agir également d'une classe intergrade réellement présente sur le terrain que le thématicien peut ou non décider de négliger. Enfin, il peut s'agir d'une classe thématiquement bien différente de son voisinage et que l'on considère significative malgré sa taille et que l'on souhaite à tout prix conserver (forêt galerie par exemple). Nous proposons également une méthode d'élimination des petites plages (pour les classes choisies par le thématicien).

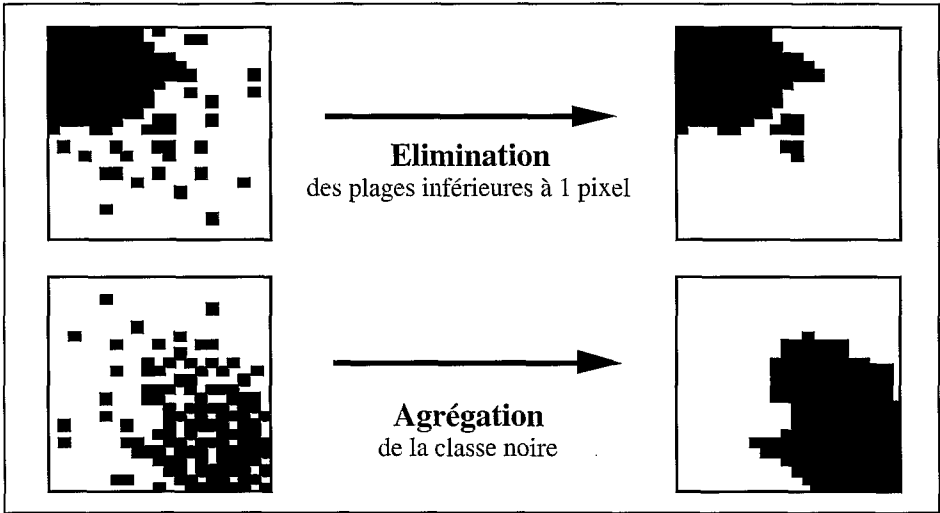


Fig.2. Exemple d'opérations d'élimination et d'agrégation.

1.5. Après la vectorisation...

Malgré les simplification que l'on aura pu appliquer à l'image, des problèmes vont encore se poser lors de la vectorisation de celle-ci. En effet, les contours des plages de l'image sont classiquement réalisés par l'enchaînement des côtés des pixels séparant deux classes. Cette procédure de vectorisation a l'avantage d'être la plus conservatrice qui soit, mais elle n'est évidemment pas économique. En effet, malgré la simplification préalable de l'image, elle va créer un très grand nombre de vecteurs « inutiles » (une ligne droite oblique ce traduira par un « escalier »). Il est donc souhaitable ensuite, en fonction de la précision souhaitée, de simplifier ces contours par des procédures de routine disponibles dans la plupart des SIG vectoriels (ex : algorithme de la corde mobile).

2. Agrégation des classes dispersées (effet « sel et poivre »)

Pour réaliser ces synthèses spatiales, plusieurs types de traitement sont utilisables, comme l'analyse de la texture, la morphologie mathématique, ou l'analyse des densités.

L'approche texturale consiste à reclasser la première classification (obtenue en général à partir de critères uniquement spectraux) en se basant cette fois sur les caractéristiques du voisinage de chaque pixel (ROGALA). On définit des compositions de référence caractérisant les nouvelles classes (ex : Forêt = 100 % forêt, Savane = 100 % savane, Cultures peu denses = 80 % savane et 20 % cultures, etc.), et on reclassifie chaque pixel en le rattachant à la composition la plus proche. Cette méthode séduisante n'a pas été étudiée ici en détail mais on notera simplement qu'elle peut provoquer la formation de liserés artificiels entre deux plages homogènes pour peu que celles-ci entrent dans la composition d'une nouvelle classe mixte. Certes, il est toujours possible de supprimer ces liserés par la procédure d'élimination décrite plus bas. L'avantage de cette méthode est de prendre simultanément en compte plusieurs classes.

Dans les cas plus simples où on s'intéresse à l'agrégation d'une seule classe (ici les cultures), il est possible d'utiliser d'autres méthodes. On pourra par exemple calculer des images de densité des cultures par un filtrage approprié (problème du choix de la taille du filtre), puis segmenter et simplifier l'image continue ainsi obtenue en fonction des intervalles de densité que l'on souhaite voir apparaître sur la carte finale.

Une deuxième possibilité, qui est celle retenue ici, est l'utilisation d'opérateurs de morphologie mathématique (MULLER, COSTER et *al.*). On part d'une image binaire I représentant les cultures par la valeur 1, les autres classes étant mises à 0 (Fig. 3).

2.1. Grandes lignes du traitement effectué

On retiendra essentiellement de cette description la démarche générale suivie, étant entendu que les paramètres de réglage sont dépendants du problème traité et notamment de la résolution de l'image initiale et du document final. L'idée de base est de réaliser une agrégation des cultures en fonction de la densité de leur répartition.

On commence donc par réaliser une fermeture⁴ morphologique sur l'image I afin de boucher les lacunes qui existent entre les plages de cultures. Globalement, un taux de fermeture croissant agrège des zones de cultures de moins en moins denses (Fig. 4).

L'opération suivante qui suit logiquement est une ouverture⁵ morphologique consistant à éliminer les plages trop petites et les portions de plages trop étroites par rapport au degré de généralisation induit par la fermeture précédente. Contrairement à la fermeture, un taux d'ouverture croissant élimine des zones de plus en plus denses (Fig. 5).

4 La fermeture est la succession de n dilations puis de n érosions inverses, ce qui a pour effet de boucher les petits trous et les concavités étroites d'une plage.

5 C'est l'inverse de la fermeture (érosions puis dilations). Elle a pour effet de supprimer les petites plages et les convexités étroites (les « presqu'îles »).

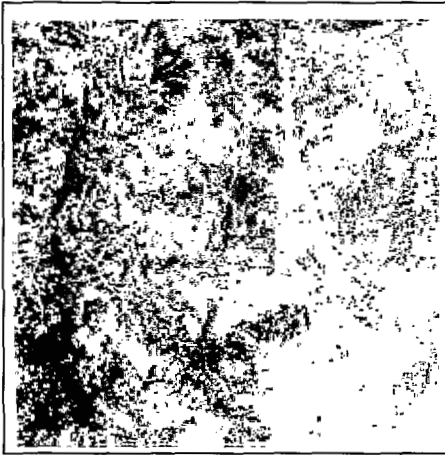


Figure 3. Classe initiale : I.

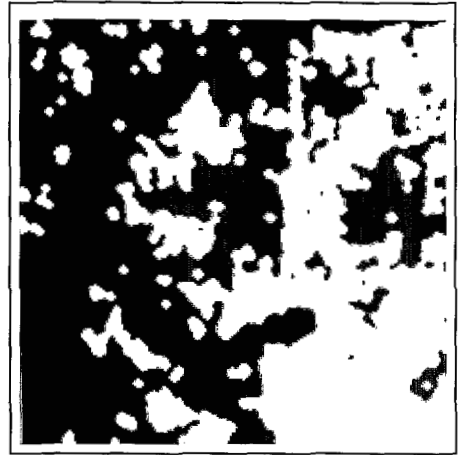


Figure 4. Fermeture : I+2F.

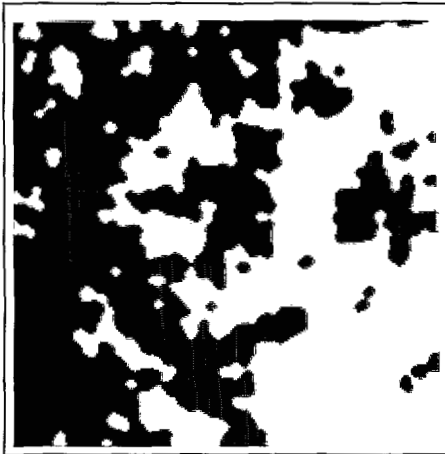


Figure 5. Fermeture + Ouverture :
I+2F+O.

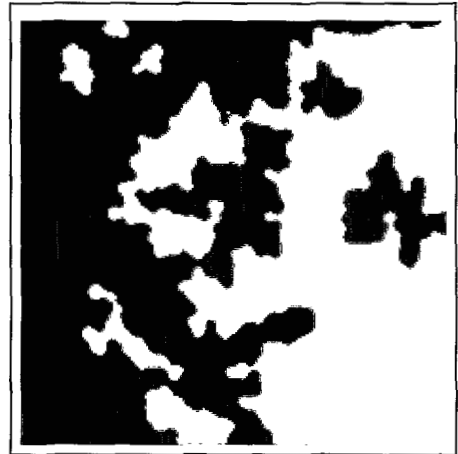


Figure 6. Fermeture + Ouverture + Nettoyage
: I+2F+2O+N(4E).

Au terme de ces deux opérations, on obtient une image binaire simplifiée sur laquelle apparaissent des petites plages pseudo-circulaires ou des petites lacunes pseudo-circulaires dans les plages plus grandes. Ces petits éléments ont un contour excessivement simplifié et déformé par rapport à leur surface, du fait que le traitement précédent est indépendant de la surface des plages qu'il traite. Or intuitivement, le lecteur d'une carte aura tendance à attendre plus de précision dans la localisation du contour d'une petite plage que dans celui d'une plus grande plage. En termes plus objectifs, cela revient à juger non pas la précision absolue du positionnement du contour, mais le pourcentage de surface correctement attribué de chaque plage. On voit bien alors que celle-ci sera très mauvaise pour les petites plages sus-évoquées. Dans le souci d'homogénéiser la carte, on va donc éliminer ces petits éléments

par une opération de nettoyage morphologique (reconstruction de l'image à partir d'une érodée⁶ de celle-ci) (Fig. 6). Remarquons que ce nettoyage n'aura d'effet que s'il est d'un niveau supérieur à la fermeture ou à l'ouverture.

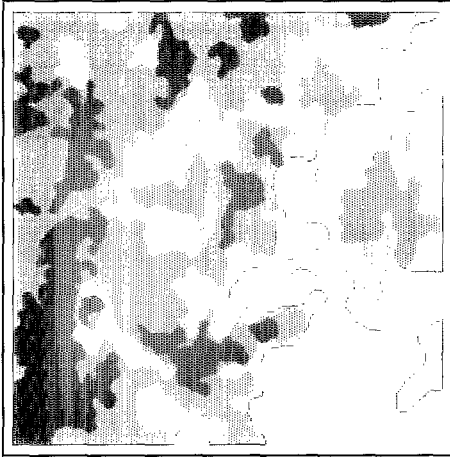


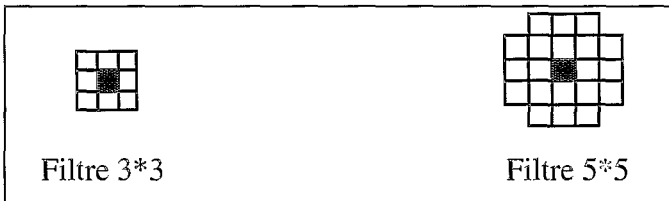
Figure 7. Combinaison des 3 niveaux d'agrégation :

$$I+N(1e)+I_f+IO+N(3E)$$

$$I+2F+2O+N(4E)$$

$$I+4F+3O+N(6E)$$

A ce schéma général s'ajoute une série de filtrages majoritaires avant le nettoyage final, afin d'améliorer l'aspect des contours parfois irréguliers du fait de la forme de l'élément structurant (carré 3*3 ou carré 5*5 sans les coins, Fig. 8).



Filtre 3*3

Filtre 5*5

Figure 8. Eléments structurants utilisés pour les opérations de morphologie mathématique.

Par ailleurs, pour délimiter les zones de cultures les plus denses, on effectue en plus de ce schéma un nettoyage préalable de l'image de départ en prenant comme marqueur une érodée de celle-ci. Ce nettoyage permet d'éviter d'agréger des zones où les pixels de culture sont proches les uns des autres sans qu'aucun noyau compact n'apparaisse vraiment.

2.2. Réalisation pratique

Les algorithmes mis au point ont été développés à partir des logiciels de traitement d'images PLANETE et OSIRIS de l'Orstom. Une des particularités intéressantes de ceux-ci est de pouvoir enchaîner les procédures élémentaires de traitement à partir d'un fichier de commandes. Cela permet d'automatiser en grande partie les procédures qui autrement

⁶ Ce terme désigne l'image résultant de l'érosion d'une image initiale (morphologie mathématique).

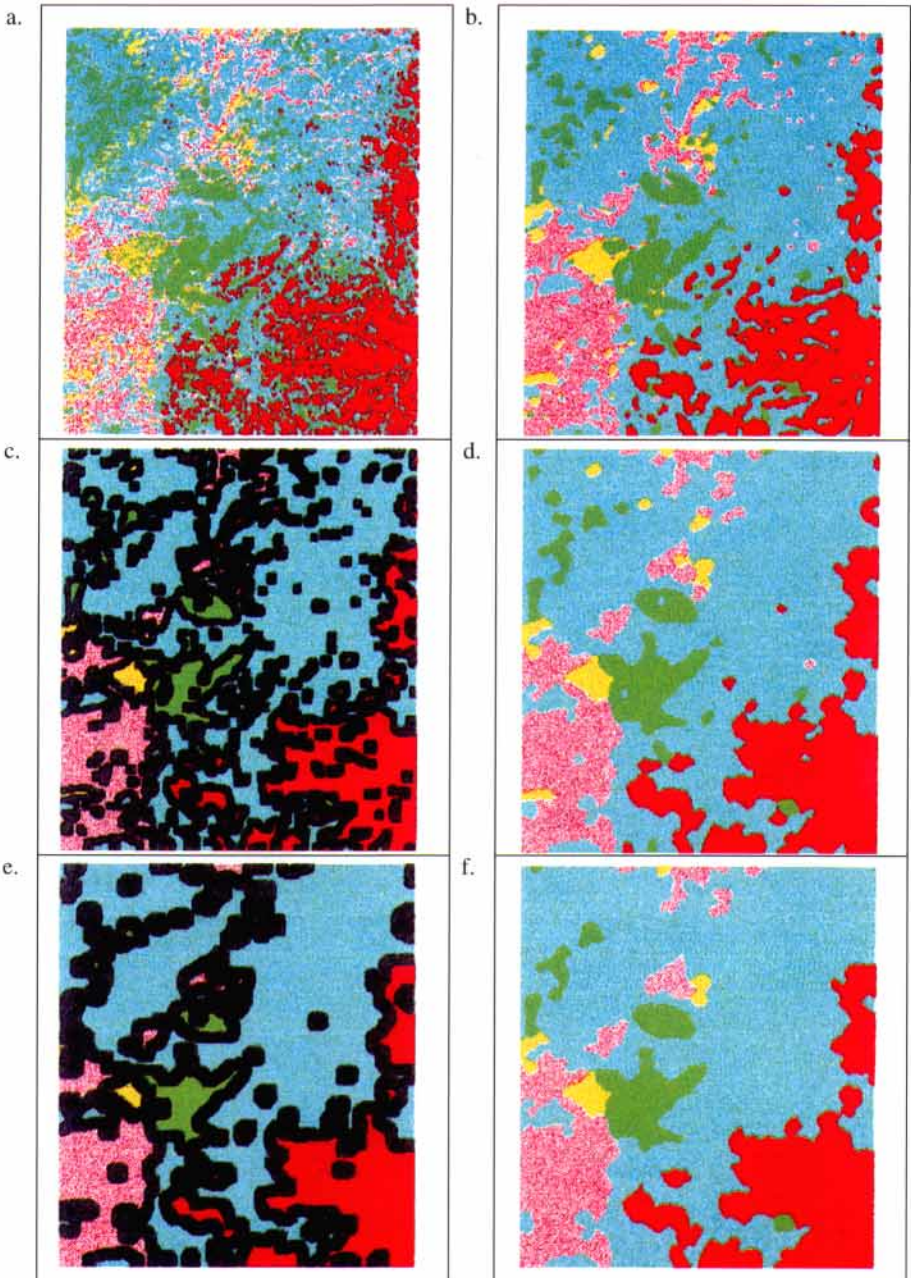


Planche 6. a. Classification initiale ; b. Lissage majoritaire (4M) ; c. Erosion à partir des contours (2E) ; d. Bouchage des lacunes par dilatation en connexité 4 ; e. Seconde érosion à partir des contours (4E) ; f. Image résultat : reconstruction des plages puis bouchage des trous par dilatation en connexité 4.

seraient très fastidieuses à mettre en oeuvre, du fait de leur complexité et des temps de calcul nécessaires.

Toute la difficulté dans l'utilisation de ce schéma de traitement est de déterminer les paramètres permettant d'obtenir un zonage cohérent. Cela signifie qu'il doit segmenter l'espace cultivé sans créer de « classes de densité » trop ou pas assez étendues, et que d'autre part ces classes doivent correspondre plus ou moins à ce que l'utilisateur final peut interpréter ou veut connaître.

Les traitements sont illustrés par les figures 3 à 7 et décrits en utilisant la codification suivante :

- X_m → x lissages majoritaires 5*5,
- x_e → x érosions 3*3,
- x_E → x érosions 5*5 tronquées,
- x_d → x dilatations 3*3,
- x_D → x dilatations 5*5 tronquées,
- x_o → x ouvertures 3*3 de niveau x (= x_ex_d),
- x_O → x ouvertures 5*5 tronquées de niveau x (= x_Ex_D),
- x_f → x fermetures 3*3 de niveau x (= x_dx_e),
- x_F → x fermetures 5*5 tronquées de niveau x (= x_Dx_E),
- N(x_e) → nettoyage par reconstruction utilisant comme marqueur l'érodée x_e de l'image.

Nous avons ainsi obtenu une partition en trois classes de densité emboîtées (cultures denses, moyennement denses, peu denses) (Fig. 7). Les paramètres correspondant à chacune de ces classes sont les suivants (Classification de l'image Sokode) :

	Cultures denses	Cultures moyennement denses	Cultures peu denses	Savane
Densité moyenne	75%	29%	10%	<1%
Densité max.	100%	33%	13%	5%
Densité min.	63%	28%	13%	0%
Surface (en pixels)	850 000	1 370 000	1 330 000	4 800 000
Réglage	N(1e)1f1ON(3E)	2F2ON(4E)	4F3ON(6E)	
Densité moyenne à la limite entre les deux classes	48%	20%	9%	

Etant donné l'incertitude sur le classement des cultures (environ 80 % de pixels correctement classés), il est préférable de ne pas chercher à délimiter les zones dans lesquelles la densité des cultures est très faible (quelques %). Dans ces cas en effet, les erreurs de classement sont peu compensées par le nombre de pixels et on risque donc de produire des contours sans signification. Il faut donc considérer que toute zone classée « savane » comporte potentiellement quelques rares cultures.

On remarque que la simplification du contour est plus importante pour les zones peu denses. Ce fait est lié au traitement effectué, mais reste en accord avec la constatation faite plus haut que la précision du contour est d'autant moins importante que la plage est étendue. Par ailleurs, plus les cultures sont espacées, plus l'incertitude sur leur classement augmente, ce qui justifie également de ne pas accorder trop d'importance à la précision de ce contour.

Afin de caractériser le mieux possible les différentes classes de densité ainsi obtenues, on peut fournir simplement la densité moyenne des cultures dans celles-ci. Toutefois, étant donné le gradient de densité qui existe dans ces plages, ce renseignement est insuffisant pour bien les décrire, on aimerait donc proposer un intervalle de densité. Il faut alors choisir la taille du voisinage sur lequel calculer cette densité, ce qui pose des problèmes à la limite des plages. Nous effectuons donc simplement l'évaluation de la densité sur la frange de pixels située en bordure de chaque classe de densité. Une fois l'algorithme défini, la difficulté pour l'utilisateur va être de trouver les paramètres du processus permettant d'obtenir les classes d'agrégation à la densité souhaitée, et au taux de généralisation souhaité.

3. Elimination des petites plages et lissage d'une classification

Certains algorithmes de classement basés sur la relaxation permettent de réduire partiellement la dispersion dans l'image classée finale, mais celle-ci est encore trop complexe pour être incorporée telle quelle dans un SIG (Fig. 9, planche 6).

Une solution classiquement adoptée pour simplifier les classifications est de lui appliquer une série de lissages majoritaires. L'inconvénient de cette méthode est de déformer parfois excessivement les contours des plages mais surtout, même après un lissage ultime (lissages successifs jusqu'à invariance de l'image), il reste des petites plages qui ne peuvent être éliminées.

Nous proposons ici une autre méthode, qui consiste à supprimer les plages en dessous d'une certaine épaisseur, définie par érosion morphologique. Ce choix revient à prendre en compte la lisibilité de la carte finale (déterminée par l'épaisseur des objets qui la constituent), alors que la prise en compte de la surface des plages relèverait de préoccupations plus « économiques ». Rappelons que si certaines classes spatialement peu étendues nécessitent d'être conservées en l'état (forêts-galeries), il suffit de les extraire avant d'effectuer le traitement puis de les réincorporer ensuite. Le traitement comporte essentiellement deux étapes : le lissage des contours des plages puis l'élimination des petites plages subsistantes.

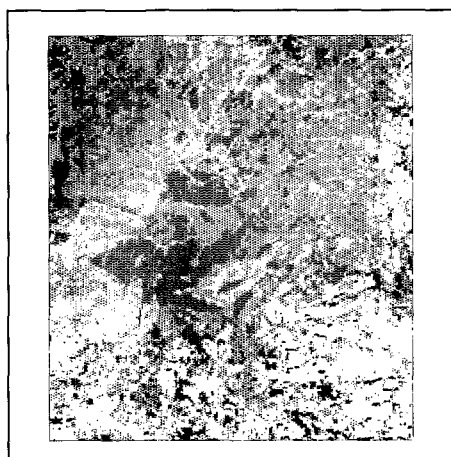


Figure 9. Classification initiale.



Figure 10. Lissage majoritaire (4M).

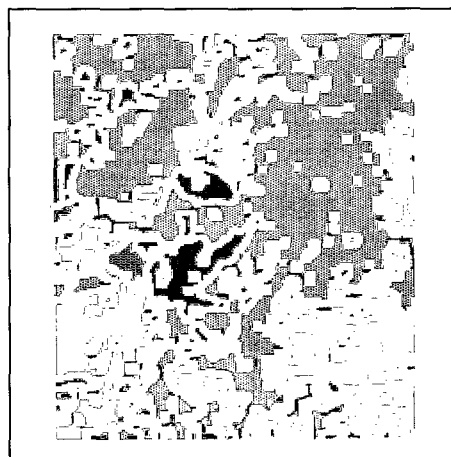


Figure 11. Erosion à partir des contours (2E)

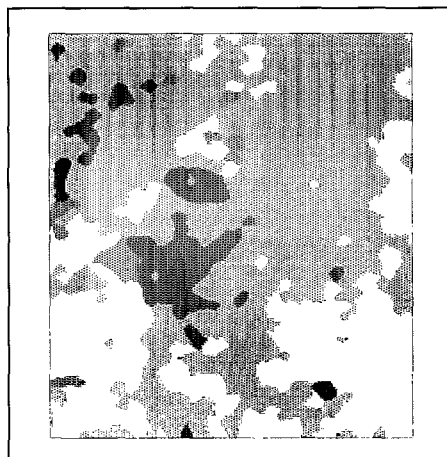


Figure 12. Bouchage des lacunes par dilatation en connexité 4.

3.1. Lissage des contours

Une classification brute est souvent constituée par endroit de la juxtaposition d'un grand nombre de très petites plages. Dans ces zones morcelées, l'élimination appliquée directement peut créer des vastes lacunes dont le remplissage donnera de mauvais résultats. Pour éviter cela, on applique sur la classification brute une série de filtres majoritaires (deux ou trois suffisent) qui provoquent l'apparition de plages de dimensions suffisantes pour résister aux érosions à venir (Fig. 10). Ainsi, indépendamment du degré d'érosion choisi ensuite, on a moins de chances de voir apparaître des lacunes trop vastes dont le comblement serait trop

imprécis. Le deuxième intérêt de ces filtrages majoritaires est de lisser en même temps les contours des grandes plages.



Figure 13. Seconde érosion à partir des contours (4E)

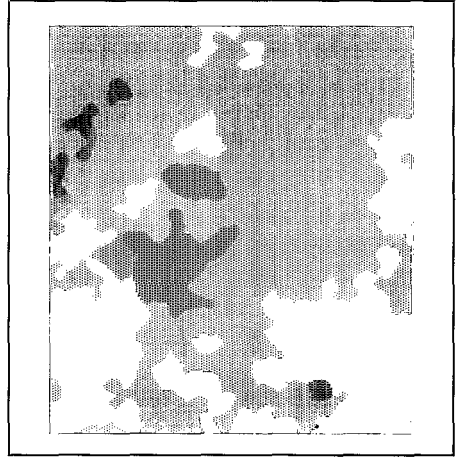


Figure 14. Image résultat : reconstruction des plages puis bouchage des trous par dilatation en connexité 4.

On code alors à 0 sur la classification tous les pixels de « bordure », qui sont en contact avec une autre classe que la leur en connexité 8. Ceci s'obtient par un simple filtrage. On réalise ensuite sur cette image une série de E érosions à partir de cette bordure, E étant proportionnel au degré de simplification souhaité. On obtient une image sur laquelle les parties restantes des plages sont séparées par des corridors de valeur 0 (Fig. 11).

Les corridors sont bouchés par des dilations successives des plages restantes (Fig. 12). On utilise pour cela un élément structurant comprenant le pixel central et ces quatre voisins en connexité 4, et en ne remplissant que les zones nulles. Ces précautions permettent d'éviter que les classes empiètent sur leurs voisins de valeur inférieure. On obtient ainsi un remplissage des vides équivalent à une affectation au plus proche voisin. Cette phase de bouchage donne des résultats peu différents d'une affectation « au plus proche voisin » (ex : GONG).

3.2 Elimination des petites plages

A ce stade du traitement, les contours des plages restantes ont été lissés de manière satisfaisante. Toutefois, comme on l'a remarqué pour le traitement précédent (§ 2.1), il reste des petites plages dont le degré de simplification du contour est excessif par rapport à leur surface. Elles sont éliminées par érosion comme à l'étape précédente (Fig. 13), mais le bouchage des corridors se fait ici par reconstruction des plages initiales et non par dilatation. Cela permet de conserver la forme des plages qui avait été « arrêtée » à l'étape précédente. Il résulte de cette reconstruction la même image qu'à la fin de l'étape 1, à la différence près

que les petites plages sont remplacées par des trous codés 0. Ceux-ci sont remplis comme précédemment par dilatation (Fig. 14).

Les généralisations obtenues sont très significatives. Par exemple, à partir d'une image classée comportant 282 000 polygones, un lissage majoritaire 5*5 en laisse 20 500, alors que le traitement cité en exemple (Fig. 9 à 14) abaisse ce nombre à 632. Outre la meilleure cohérence de l'information, cela facilite beaucoup sa manipulation par le SIG.

Cette méthode a toutefois l'inconvénient de nécessiter des calculs importants, de durée imprévisible, notamment à cause des reconstructions morphologiques qui peuvent être longues si on ne dispose que de simples algorithmes itératifs (ex : 8 heures sur un SUN LX pour une image 3000*3000 et une reconstruction à partir de 5 érosions avec un élément structurant 5*5 (Fig. 8)).

Par ailleurs, comme pour l'agrégation, la difficulté réside pour l'utilisateur dans le choix des paramètres donnant le niveau de généralisation souhaité.

Conclusion et perspectives

Nous venons de présenter deux aspects de la généralisation d'images classées permettant de faciliter leur vectorisation ainsi que leur lisibilité. Ces méthodes sont basées sur l'utilisation de la morphologie mathématique et présentent l'avantage de ne pas nécessiter de développement informatique important, puisqu'elles utilisent des primitives classiques de traitement d'images qu'il suffit de pouvoir enchaîner.

L'importance de la généralisation d'images ne peut aller que croissant étant donné le développement que connaissent les bases de données géographiques dont l'imagerie satellite constitue une source d'alimentation privilégiée. Toutefois, même en restant dans le domaine de l'image, la généralisation constitue avec la surimposition d'information topographique un moyen essentiel d'améliorer la lisibilité des documents finaux.

Les traitements mis au point nous ont permis de transférer de manière satisfaisante les résultats d'une classification de l'occupation du sol, réalisée à partir d'images satellites Landsat TM, vers une base de données vectorielle préexistante. Cette base de données constitue le noyau d'un projet consacré à l'étude de l'état de dégradation des terres du Togo, sujet faisant l'objet d'une autre communication à ce Symposium (BRABANT *et al.*).

Un des prolongements possibles de cette étude serait de chercher à quantifier la qualité de l'image généralisée résultante (notamment la justesse géographique), afin de déterminer les paramètres optimaux de généralisation. Il serait également intéressant de pouvoir tenir compte de certaines priorités lors du remplacement des petites plages par les classes voisines (imposer par exemple qu'une petite plage de « cultures » située à la limite entre « forêt dense » et « savane » soit remplacée exclusivement par de la « savane »).

Bibliographie

- BRABANT P., SIMONNEAUX V., EGUE K. (1995). « Etat actuel de la dégradation des terres et de l'environnement au Togo déterminé à l'aide de la télédétection », Symposium « Surveillance des sols dans l'environnement par télédétection et SIG », Ouagadougou 6-10 Février 1995, 12 p.
- BRASSEL K.E., WEIBEL R. (1988). « A review and conceptual framework of automated map généralisation », *International Journal of Geographical Information Systems*, Vol. 2, N°3, p 229-244.
- COSTER M., CHERMANT J.L. (1989). *Précis d'analyse d'images*, Presses du CNRS, pp 65-88.
- EHLERS M. (1992). Remote sensing and GIS : Image-Integrated GIS, in *Geographic Information System (GIS) and Mapping - Practices and Standards*. ASTM STP 1126, A.I.JOHNSON, C.B.PETTERSSON and J.L.FULTON (eds), American Society for Testing and Materials, Philadelphia, pp.53-67.
- ESTES J.E., STAR J.L. (1993). « Remote sensing and GIS integration : Toward a prioritized reasearch agenda ». 25th int. Symposium, Remote Sensing and Environmental Change, GRAZ, Austria, 4-8 April.
- GIRARD M.C., MOUGENOT B., RANAIVOSON A. (1990). « Présentation d'un modèle d'organisation et d'analyse de la structure des informations spatialisées : OASIS », Deuxièmes journées de télédétection, Caractérisation et suivi des milieux terrestres en régions arides et tropicales, Bondy 4-6 décembre, Orstom éditions.
- GONG P. (1994). « Reducing boundary effects in a kernel based classifier », *International Journal of Remote Sensing*, Vol.15, n°5, pp.1131-1139.
- LUNETTA R.S. (1991). « Remote Sensing and Geographic Information System Data Integration : Error Sources and Research Issues », *Photogrammetric Engineering and Remote Sensing* Vol.57, n°6, pp.677-687.
- MOMAL F. (1990). « Structuration et exploitation numérique des informations géographiques. Application à la cartographie pédologique ». *Sols*, n°18, Institut National Agronomique Paris Grignon, 78850 Thiverval Grignon, France.
- MULLER J.C. (1991(?)). « Generalization of spatial database », In: MAGUIRE D., GOODCHILD M.F., RHIND D. (eds), *Geographical Information Systems - Principles and Applications*. Longman, London, pp. 457-475, Vol 1.
- RIGAUX P. (1994). « La représentation multiple dans les Systèmes d'informations géographiques », *Revue internationale de géomatique*, Vol. 4, N°2, p 137-164.
- ROBBEZ-MASSON J.M. (1994). *Reconnaissance et délimitation de motifs d'organisation spatiale - Application à la cartographie des pédopaysages*. Thèse de Doctorat, Ecole Nationale Supérieure Agronomique de Montpellier, 2 place Viala, 34060 Montpellier, France. 160 p.
- ROGALA J.P. (1982). *Approche numérique de l'espace agricole. Analyse de l'hétérogénéité et de la struicure des unités cartographiques*. Thèse de docteur ingénieur, Institut National Agronomique Paris-Grignon, 78850 Thiverval Grignon, France.
- SYSAME-CNES (1991). « Télédétection et SIG pour la Gestion des Périmètres Irrigués de l'Office du Niger », Rapport Final, 67p.