

L'ETUDE DE L'APORTEE DE LA THEORIE DES GRAPHE A LA CLASSIFICATION NON SUPERVISEE ET LA CLASSIFICATION SUPERVISEE POUR L'ANALYSE DE LA BAIE DU MONT SAINT-MICHEL A PARTIR DES IMAGES SATELLITAIRES

BOURAOUI S.

Université de Strasbourg, STRASBOURG, FRANCE

RÉSUMÉ

Dans ce papier, une nouvelle démarche pour la cartographie basée sur la notion d'objets et de relations entre ces objets est proposée afin de tirer profit des deux méthodes de classification utilisées. D'une part, les objets obtenus après une classification supervisée sont représentés par un graphe d'adjacence modèle. D'autre part, les objets obtenus après une classification non supervisée sont représentés par un graphe d'adjacence de donnée, et le but est de faire la mise en correspondance entre ces deux graphes afin d'améliorer les résultats de la classification non supervisée en les associant à ceux de la classification supervisée. Cette étude a concerné le littoral de la baie de Mont saint-michel, les données utilisées sont des images satellitaires multi-temporelles acquises lors du projet ECOSGIL (Extraction des Connaissances Spatiales pour la Gestion Intégrée du Littoral).

MOTS CLÉ

Classification, théorie des graphes, relations spatiales, cartographie d'objets littoraux.

1. INTRODUCTION

La baie du Mont saint-michel est un espace littoral soumis à de grandes pressions environnementaux et anthropiques. Les contraintes naturelles et les activités humaines présentent des pressions croissantes auxquelles le littoral doit faire face, nécessitent des solutions rapides pour la gestion de ce territoire.

Dans ce souci, l'imagerie occupe une place de plus en plus importante, l'étude du milieu littoral à partir d'images satellitaires permet d'aborder des domaines variés allant de la simple cartographie de l'occupation du sol à l'étude de la zone d'estran et de la dynamique hydro-sédimentaire.

La baie du Mont Saint-Michel a été retenue comme site d'étude d'une part pour son intérêt scientifique qui lui est porté depuis plusieurs décennies. Ce territoire a aussi une portée internationale avec des enjeux environnementaux et patrimoniaux très importants. D'autre part, la baie est l'une des plus complexes et des plus dynamiques zones au monde. Les enjeux multiples que représente la gestion de la baie nécessitent une connaissance approfondie de cet espace et de ses évolutions actuelles et à venir.

La baie est un milieu très dynamique qui évolue rapidement, les images satellitaires sont très bien adaptées pour un suivi régulièrement mis à jour et elles permettent une étude multiscalaire, en complément des données traditionnelles utilisées par les gestionnaires. De plus la précision des images est de plus en plus fine, se rapprochant aujourd'hui de celles des photographies aériennes. Enfin les images satellitaires permettent de nombreux traitements supplémentaires de part leur étendue spectrale avec l'infrarouge. Cet intérêt d'utiliser des images satellitaires pour l'étude du milieu littoral à déjà fait l'objet de plusieurs travaux, Gourmelon et al. (2004), Le Berre et al.(2003), Marieu et al. (2004), Simon et al. (2004).

A l'arrivé des images à THR ou les techniques dites traditionnelles de traitement d'image qui se basent sur les pixels ou les régions ne sont plus applicables, la communauté est actuellement intéressée par les techniques de traitement d'objets et de relations entre ces objets. L'intérêt dans une telle approche est d'extraire le maximum d'information et d'avoir un traitement intelligent de l'information de la même manière que dans le raisonnement humain.

Ce papier est organisé comme suite : dans la section 2, on présente la description du système de classification. Section 3, présente la classification supervisée. Dans la section 4, on présente la méthode de classification non supervisée utilisée. Dans la section 5, on présente la construction des graphes d'adjacence et sa mise en correspondance. Enfin, dans la section 6, on donne la conclusion et quelques perspectives.

2. LA DESCRIPTION DU SYSTÈME DE CLASSIFICATION

La description du système est donnée par la figure 1. Deux différentes méthodes de classification s'exécutent en parallèle pour cartographier l'image, une supervisée, et l'autre non supervisée. On va détailler chacune de ces méthodes dans ce qui suit ainsi que les différentes étapes qui constituent cette architecture.

La classification supervisée s'effectue à partir des échantillons, la sélection et le choix des échantillons est une étape primordiale dans le processus de classification supervisée, ainsi que la qualité dépend de cette phase d'échantillonnage. C'est pour quoi il faut bien choisir les échantillons qui couvrent toute les classes d'objets. Ce n'est pas très claire pour un non expert du domaine de distinguer entre les différents objets surtout on milieu littorale ou les objets sont très semblables. Donc pour faire un bon échantillonnage, ça nécessite l'expert du domaine (connaissance du domaine géographique), tandis que la deuxième méthode de classification non supervisée n'exige aucune connaissance du domaine.

C'est deux méthodes de classification sont indépendantes l'une de l'autre et s'exécutent en parallèle. A partir des résultats de chacune de ces deux méthodes, on va construire un graphe d'adjacence. Dans le cas de la classification supervisée, le système crée un graphe modèle ou chaque nœud représente une étiquette qui correspond à un objet classifié par cette méthode et chaque arc représente la relation d'adjacence (graphe d'adjacence). Pour la classification non supervisée, le système crée un graphe de donnée ou chaque nœud représente une forme trouvée et chaque arc représente la relation d'adjacence entre ces formes.

On a utilisé qu'une seule relation entre objets (A-coté-de) qui peut exprimer implicitement d'autres relations comme la relation (entouré-par) qui peut être exprimée par (tout les 4 ou 8 A-coté-de d'un objet sont le même objet).

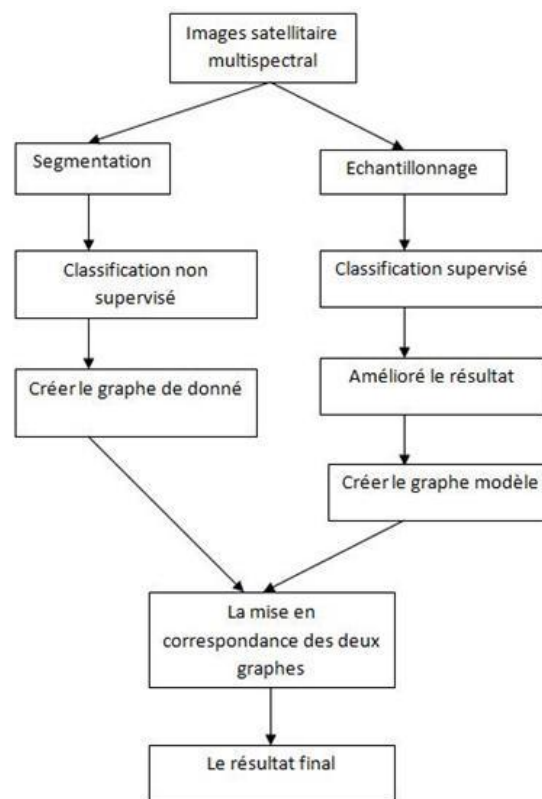


Figure 1 : L'architecture détaillée du système.

3. LA CLASSIFICATION SUPERVISÉE

Cette méthode de classification part d'un ensemble d'échantillons qui représente les différentes classes d'objets qu'on souhaite trouvé dans cette image. On a appliqué l'algorithme standard de classification par Maximum de vraisemblance, L'avantage majeur de cette méthode de classification est sa rapidité, mais elle n'est pas très efficace surtout dans notre cas d'étude ou les repenses spectrales des objets s'interfères. C'est pour quoi si on utilise seulement la repense spectrale on ne peut pas distinguée entre les objets « mer », « cours d'eau, étendue d'eau » et « chenal et mer peu profonde ». (Figure 2)

Pour cela, et après cette classification on applique une phase d'amélioration des résultats qui repose sur la notion de voisinage directe en appliquant des règles qui tiennent en comptes des voisins (ici les 4 ou 8 voisins) d'un pixel étiquette. Dont voici un exemple d'une des règles utilisées :

- Format d'une règle : *si (étiquette(i,j) == «Mer ») et (étiquette (i-1,j) == surfaces agricoles avec faible couvert végétal, sol nu) ou (étiquette (i+1,j) == « surfaces agricoles avec faible couvert végétal, sol*

nu) ou (étiquette (i,j-1) == « surfaces agricoles avec faible couvert végétal, sol nu ») ou (étiquette (i,j+1) == « surfaces agricoles avec faible couvert végétal, sol nu ») alors étiquette (i,j) == « canal d'eau ».

Les classes « eau » sont facilement identifiables. Il est possible d'isoler cette catégorie d'objets en effectuant un simple seuillage des valeurs numériques, dans une seule bande spectrale (seuillage monospectral). Les surfaces en eau sont observables au moyen d'un capteur sensible au proche infrarouge : tous les pixels dont la valeur numérique est inférieure à une valeur seuil peuvent être affectés à la classe « eau », car l'eau absorbe le rayonnement. Mais c'est difficile de distinguer entre ces sous classe (Mer, Canal et mer peu profonde, cours d'eau, étendue d'eau).

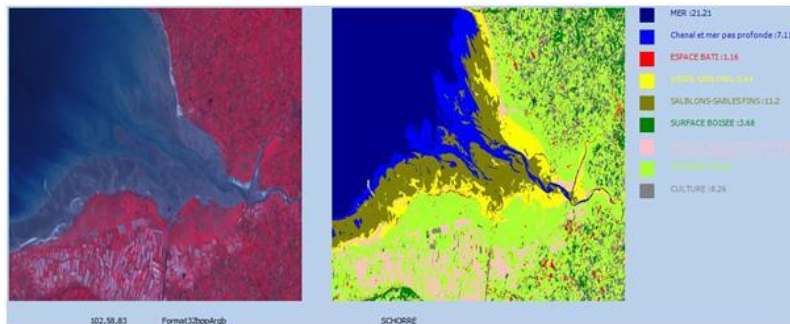


Figure 2 : Les résultats de la classification supervisée pour neuf classes.

On générale le format des règles de classification pour l'étape d'amélioration des résultats tiennent en compte de l'aspect spatiale des objets. En d'autres termes, la première elle raffine les résultats de la première classification en introduisant la contrainte de voisinage directe.

- Si (objet classer chenal) et (non (A-cote-de (objet, mer))) alors objet devient cours d'eau, et étendue d'eau.

Il reste aussi comment extraire le schorre, l'information de voisinage directe ne suffit pas, on a utilisé la notion spatiale entre deux objets pour trouver et limité le schorre sur une carte de gradient sur le néo-canal NDVI (Normalized Difference Vegetation Index) :

$$NDVI = (IR - R) / (IR + R) \quad (1)$$

Il varie entre -1 et 1. Cet indice est très efficace pour la détection de la végétation active. Ainsi qu'un autre indice et calculé afin de mettre en évidence les surfaces minérales, sols nu, c'est l'indice de Brillance (IB) donné par la formule suivante :

$$IB = (PIR^2 + R^2)^{1/2} \quad (2)$$

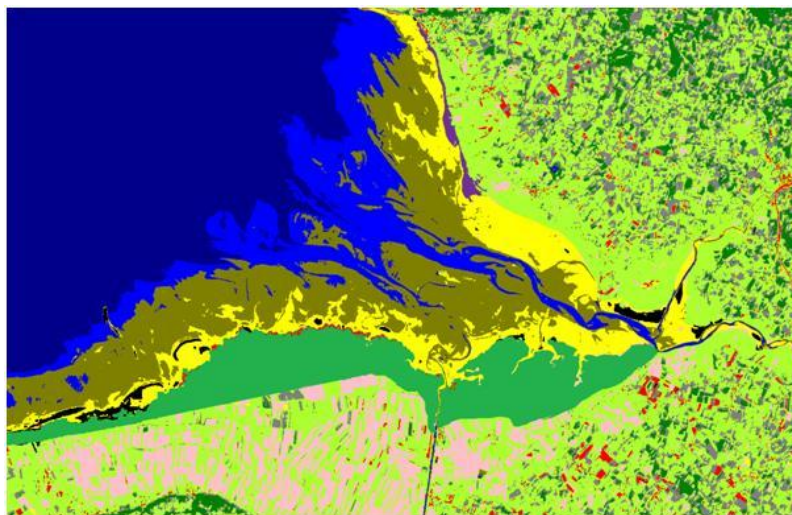


Figure 3 : Le résultat final de la classification supervisée.

4. LA CLASSIFICATION NON SUPERVISÉE

Ce type de classifications est guidé par la phase de segmentation, on a utilisé un algorithme de segmentation par croissance des régions avec une quelques changements pour mieux adapter l'algorithme a nos besoins.

Le principe est de regrouper les pixels ayant le même niveau de gris dans un seul groupe (cluster), le problème majeur ici, c'est qu'un simple changement dans la valeur de niveaux de gris d'un pixel va lui

affecter à un nouveau groupe. Pour cela, on a utilisé un paramètre $k = 5$; qui permet de regrouper les pixels ayant le niveau de gris (+ / - k) par rapport à celui du pixel de départ, ainsi qu'on recalcule le niveau de gris à chaque itération par la formule :

$$\text{NewNg} = (\text{OldNg} * 2 + \frac{1}{4} (\text{Ngg} + \text{Ngd} + \text{Ngh} + \text{Ngb}))/ 3 \quad (3)$$

Avec : NewNg : le nouveau niveau de gris du pixel.

OldNg : l'ancien niveau de gris du pixel.

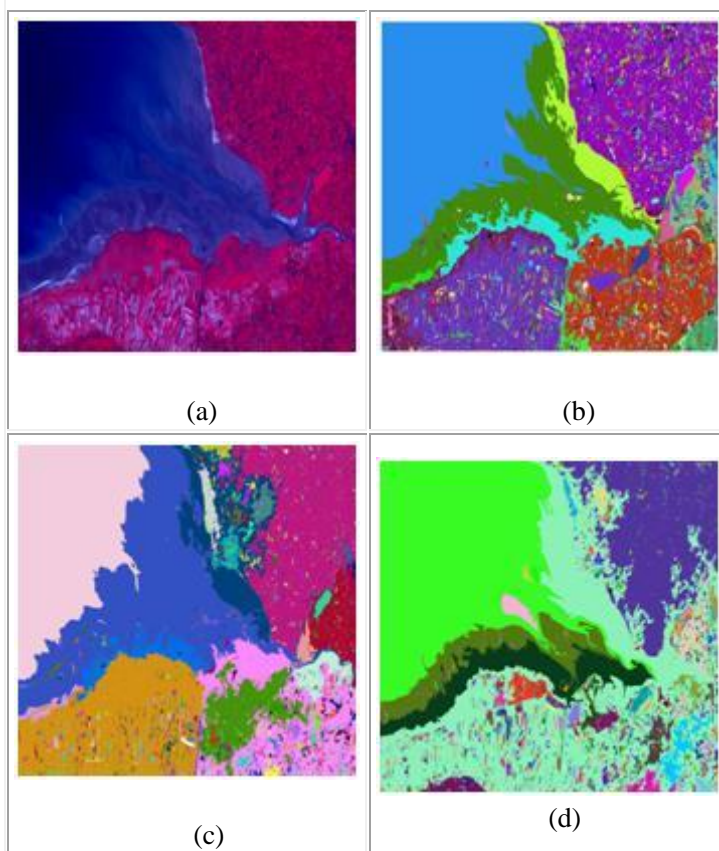
Ngg : le niveau de gris du voisin de gauche.

Ngd : le niveau de gris du voisin de droite.

Ngh : le niveau de gris du voisin en haut.

Ngb : le niveau de gris du voisin en bas.

Tableau 1 : (a) l'image originale, (b) segmentation avec $k = 5$, (c) segmentation avec $k = 10$, (d) segmentation a partir d'indice NDVI avec $k = 5$



Dans un deuxième temps, on appliqué cette algorithmme de segmentation sur une carte de gradient pour cette image, ensuite pour chaque région trouver, on calcule un indice de consistance entre ces régions voisins afin de les regrouper en une seule région.

L'indice de consistance entre deux régions est donnée par :

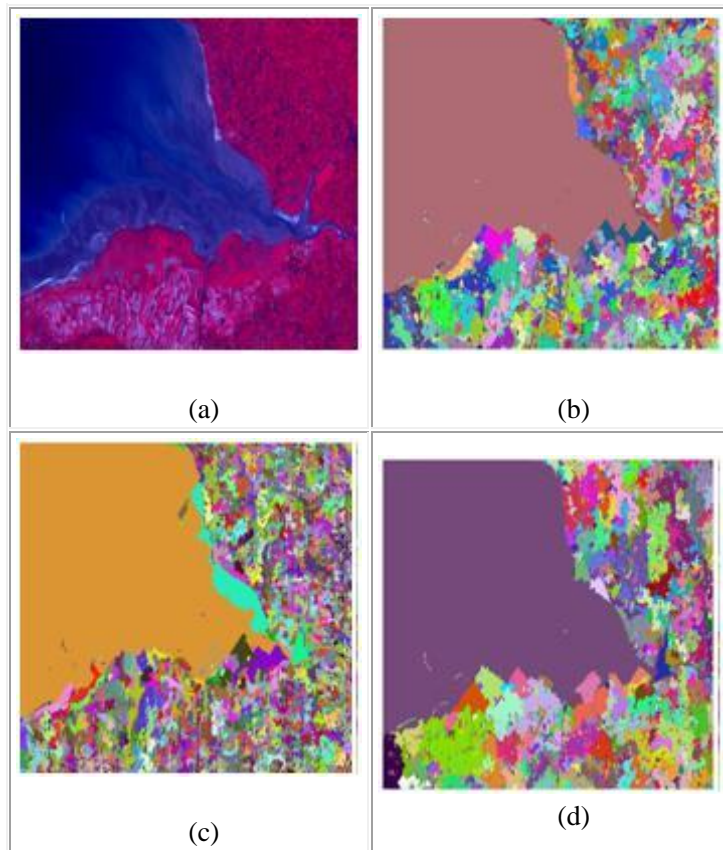
$$\text{IC} = 1/(i*j) * \sum_{-i} \sum_{-j} (|\text{PiR1} - \text{PjR2}|) \quad (4)$$

Avec : IC : l'indice de consistance.

PiR1 : pixel i de la région 1.

PjR2 : pixel j de la région 2.

Tableau 2 : (a) l'image originale, (b) segmentation avec $k = 5$, $\text{IC} = 15$, (c) segmentation avec $k = 10$, $\text{IC} = 20$, (d) segmentation a partir d'indice NDVI avec $k = 5$, $\text{IC} = 20$



5. LA CONSTRUCTION DES DEUX GRAPHS ET LA MISE EN CORRESPONDANCE

Cette étape est divisée en trois parties, dans la première partie, on construit un graphe modèle à partir de la classification supervisée, les nœuds représentent les étiquettes et la moyenne de niveaux de gris et la taille de chaque classe (en nombre de pixels) pour chaque objet, et les arcs représentent la relation d'adjacence (voisinage directe).

Ainsi qu'on construit un graphe de donnée à partir d'une segmentation non supervisée avec ($k = 10$, $IC = 25$). Les nœuds représentent la moyenne de niveaux de gris pour chaque région ainsi que sa taille. En générale, la taille de graphe de donnée est plus grande que celle du graphe modèle.

Dans la deuxième partie, on regroupe les nœuds du graphe de donnée ayant le niveau de gris le plus près de celui du nœud dans le graphe modèle, ainsi les régions se regroupent au tour des nœuds du graphe modèle et à chaque regroupement on recalcule le niveau de gris moyenne des régions par la formule (3). L'information spatiale utilisée ici est la relation A-coté-de, elle est bijective.

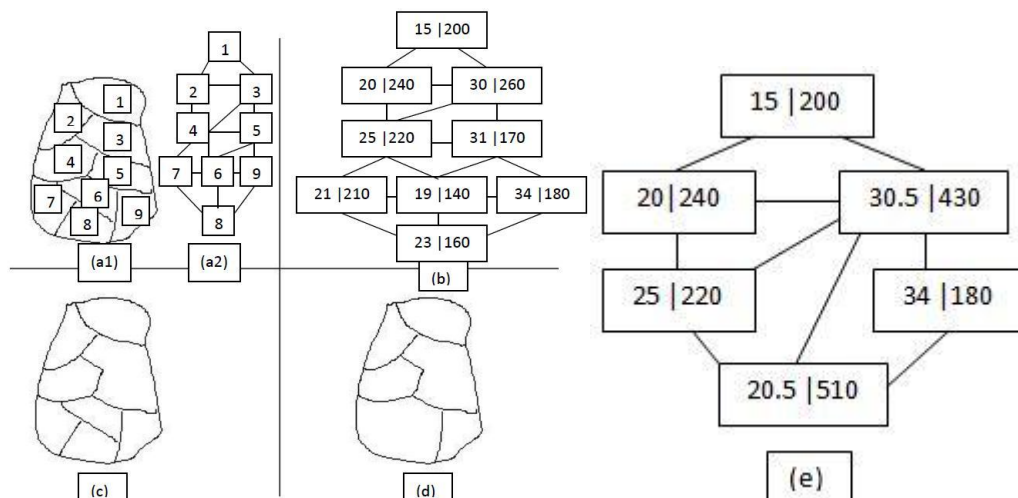


Figure 4 : (a1) le résultat de la segmentation non supervisée, (a2) son graphe d'adjacence, (b) la composition des nœuds la valeur à gauche le niveau de gris et à droit la taille de la région, (c) le résultat de la première itération, (d) le résultat finale avec Nb-nœud = 6, (e) le graphe résultat finale.

Par fois, une ambiguïté dans la décision est apparue lorsqu'on se retrouve dans une situation où il y a plus qu'un seul nœud voisin vérifiant la condition de regroupement (son niveau de gris et le proche du nœud). Dans ce cas, la décision est pour la région la plus grande (en nombre de pixels). En d'autres termes, ce nœud critique va être ajouté au nœud qui représente la plus grande région.

À la fin de cette étape, on se retrouve avec deux graphes qui ont le même nombre de nœuds, et l'étiquetage du graphe de donnée s'effectue comme suit :

Pour chaque nœud dans le graphe modèle faire

- trouve le nœud dans le graphe de donnée qui minimise la distance, cette distance est calculée par la formule : $Distance = ((NGM - NGD) * (TailleM - TailleD))^{1/2}$ (5)
- if (taille nœud donnée) est inférieure à celle (du nœud modèle) alors :
 - ▶ nœud donnée := nœud modèle.
 - ▶ Supprimer tous les pixels qui ont été pris par ce nœud dans le graphe modèle.
- étiqueter ce nœud du graphe de donnée avec la même étiquette dans le graphe modèle.
- Return le graphe de donnée.

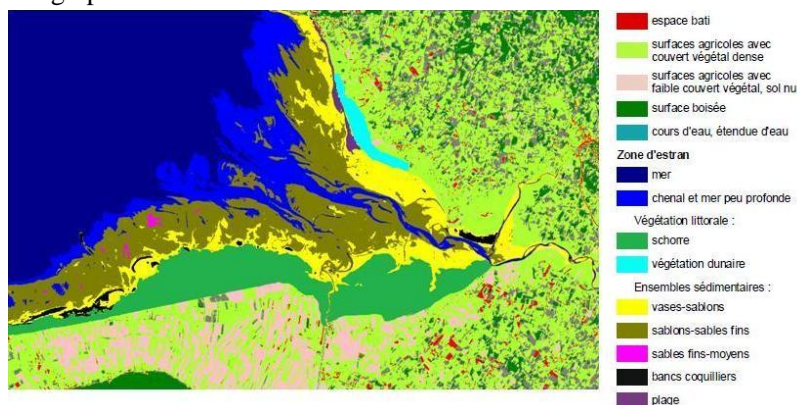


Figure 5 : Le résultat de l'application de la mise en correspondance des deux graphes et l'étiquetage.

Tableau 3 : Connaissance sous forme d'un tableau de voisinages construits à partir des résultats présentés dans la figure 5.

Objets	A comme voisin
MER	Chenal et mer peut profonde
Chenal et mer peut profonde	MER Sablons-sables fin Sables fins-moyens Vases-sablons bancs coquilliers Schorre
Sablons-sables fin	Chenal et mer peut profonde Sables fins-moyens Vases-sablons bancs coquilliers
Sables fins-moyens	Chenal et mer peut profonde Sablons-sables fin Vases-sablons bancs coquilliers
Vases-sablons	Chenal et mer peut profonde Sablons-sables fin Sables fins-moyens bancs coquilliers Plage Dune Schorre Surface Boisée Sol nu Agricoles dense bâti
Bancs coquilliers	Chenal et mer peut profonde Sablons-sables fin Sables fins-moyens Vases-sablons Schorre
Plage	Vases-sablons Dune bâti
Dune	Vases-sablons Plage Schorre Sol nu Agricoles dense bâti
Schorre	Chenal et mer peut profonde Vases-sablons bancs coquilliers Dune Sol nu Agricoles dense bâti
Cours d'eau	Surface Boisée Sol nu Agricoles dense
Surface Boisée	Vases-sablons cours d'eau Sol nu Agricoles dense bâti
Sol nu	Vases-sablons Dune Schorre cours d'eau Surface Boisée Agricoles dense bâti
Agricoles dense	Vases-sablons Dune Schorre cours d'eau Surface Boisée Sol nu bâti
Bâti	Vases-sablons Plage Dune Schorre Surface Boisée Sol nu Agricoles dense

6. CONCLUSION ET PERSPECTIVES

Dans ce travail, on a présenté deux méthodes différentes de segmentations l'une supervisée et l'autre non supervisée pour la cartographie du milieu littoral. La représentation sous forme de graphe d'adjacence permet de mieux guider la classification non supervisée et apporte la notion spatiale dans le processus de classification. La mise en correspondance est faite à partir du graphe modèle qui a été construit à partir de la classification supervisée et du graphe de donnée qui a été construit à partir de la classification non supervisée pour étiqueter cette dernière. Le processus est testé sur quelques images, il reste à tester sur plusieurs images avec différentes résolutions, ensuite étudier et suivre le changement de la baie à partir des images multi-temporelles.

7. RÉFÉRENCES

Gourmelon F., Robin M., SIG et littoral. Paris, Hermes, 2005.

Gourmelon F., Sauve L., Bioret F., Potentialités de l'imagerie satellitaire SPOT 5 pour la cartographie de la végétation terrestre, application à l'île d'Ouessant. CNES-IFEN, 2004.

Le Berre I., Hénaff A., Wenzel F., cartographie synthétique de l'environnement littoral du Finistère : exploitation de SPOT5 pour la cartographie de l'estran, du trait de côte et de l'occupation du littoral. Géormer, 2003.

Marieu V., Dupuis H., Bretel P., Evaluation du potentiel de l'imagerie SPOT 5 dans le secteur du bassin d'Arcachon : Cartographie du milieu lagunaire et du littoral océanique. Université Bordeaux, 2004.

Simon B., Apport des données du satellite SPOT 5 aux études environnementales de la zone du delta du Rhône et de l'étang de Berre. 2004.