

# Bibliographie et Étude de Cas (BEC)

## Première Partie: Résumé

KAFANDO Rodrique

**Abstract**—Ce document fait l'objet du résumé d'un article consacré à l'étude de la première partie du module bibliographie et étude de cas (BEC). Dans cette première partie, il est question de faire ressortir les idées maîtresses évoquées dans l'article afin de se rassurer que les approches évoquées par l'auteur ont été bien comprises par l'étudiant afin de pouvoir aborder la seconde partie qui est l'analyse et l'étude comparative. Notre article traite un sujet relatif à la fouille de données, et au traitement d'images. L'intitulé de notre article est "A graph-based approach to detect spatiotemporal dynamics in satellite image time series [1]".

**Index Terms**—Satellite image time series, Monitoring, OBIA, Data mining, Graph-based techniques, Land-cover.

### 1 INTRODUCTION

L'exploration spatio-temporelle des territoires est de nos jours un phénomène qui reste de plus en plus difficile du fait que les composantes qui le composent changent et évoluent dans le temps de même que dans l'espace. De nos jours, ce problème semble être résolu grâce aux satellites de surveillances terrestres qui ne cessent de tourner afin de fournir la structure de la terre sur des périodes bien définies. Cependant, la maîtrise des informations contenues dans ces images restent méconnues ou non maîtrisées due au manque de bonnes méthodes d'analyses sur ces dernières. Suivre l'évolution morphologique et/ou temporelle d'un environnement reste donc une tâche qui fait l'objet de plusieurs recherches. Dans le cas présent, il est question de suivre l'évolution des surfaces agricoles qui ne cessent de varier au fil du temps à partir des images satellitaires récoltées sur des séries temporelles ou SITS. Pour plus de compréhension, nous aborderons le travail réalisé à partir des points des grands points comme la problématique, la motivation, l'approche ou la méthodologie utilisée, la phase d'expérimentation et les résultats, une analyse critique sur le sujet et une conclusion pour terminer.

### 2 PROBLÉMATIQUE

Les satellites fournissent des images sur des séries temporelles (SITS) qui sont considérées comme une banque d'informations à partir de laquelle il devient possible de mettre en évidence les évolutions spatio-temporelle et même d'établir les relations dynamiques qui pourraient en dégager entre ces différentes périodes. Jusqu'à présent, les méthodes ou les approches traitant les mêmes types de données sont uniquement basées sur la classification et ne peuvent fournir des informations sur leurs évolutions et leurs comportements dans le temps. C'est ainsi, que l'objectif de ce travail était donc de proposer une nouvelle approche permettant de détecter et d'extraire de façon automatique des informations spatio-temporelles à partir des séries chronologiques d'images satellitaires. L'analyse de ces différentes images contribuera donc à construire une image générale qui

représentera toute une zone d'étude tout en laissant voir son comportement dans le temps.

### 3 CONTEXTE

Le travail effectué dans cet article se repose essentiellement sur deux domaines à savoir le traitement d'images et la fouille de données. En effet, les images obtenues sont soumises à des processus de traitement avant d'être appliquées aux processus de recherches.

### 4 MOTIVATION

Comme mentionné ci-haut, le traitement des données spatio-temporelles obtenues à partir des séries chronologiques d'images satellitaires (SITS) se limitait toujours la technique de classification et ne peuvent fournir des informations sur leurs évolutions et leurs comportements dans le temps. Pourtant il est bien nécessaire de lever cette insuffisance qui renferme sans doute une grande partie des informations. L'étude sur l'évolution des sols agricoles est un exemple palpable qui permettra de constater plus facilement par exemple les migrations de cultures et les surfaces occupées dans le temps.

### 5 APPROCHE PROPOSÉE

#### 5.1 MÉTHODOLOGIE

En rappel, l'objectif de cet travail est de trouver un moyen qui permettra de suivre l'évolution spatio-temporelle décrivant la façon dont évolue une entité dans le temps.

L'idée générale de la technique proposée consiste à sélectionner l'ensemble des entités que l'on souhaite suivre tout au long de la période d'observation. Cet ensemble forme alors ce que l'on appelle les Bounding Boxes (BBs). Plusieurs sous-ensembles BBs peuvent appartenir à des horodatages différents. Dans le but de pouvoir suivre l'évolution de chaque BBs, un graphe d'évolution est généré par la suite en tenant compte de tous les objets présents dans ce dernier. La composition du graphe est la suivante:

- Chaque nœud représente un objet dans le BBs;
- Les nœuds sont reliés par des arrêtes.

Deux nœuds sont reliés lorsque les deux objets représentatifs se chevauchent. Le poids de l'arrête représente le degré de chevauchement entre ces deux objets. Plus le chevauchement est fort, plus le poids de l'arrête est fort et vice-versa. L'analyse se focalise sur deux niveaux. Le premier niveau consiste à l'analyse des trajectoires temporelles d'un événement spatio-temporel particulier, tandis que le second niveau permet de faire ressortir une image plus générale résumant la dynamique temporelle détectée sur l'ensemble du milieu d'étude. Le principe et le mode d'élaboration est explicité dans le point suivant.

## 5.2 PRINCIPE DE FONCTIONNEMENT

Comme précédemment évoqué, l'une des phases la plus importante est la sélection des Boundings Boxes ou BBS. La première étape consiste donc à la sélection des BBs que l'on désire surveiller de façon cohérente. L'une des particularité de ce sous-ensemble est qu'il doit couvrir au maximum l'ensemble du site d'étude. Ce choix réside de l'hypothèse à partir de laquelle chaque BB possède une extension maximale au point de vue spatiale à une certaine période. L'exemple concret est celui d'un lac dont la quantité d'eau varie en fonction du temps, ce qui veut dire qu'à partir d'une même zone, il est possible d'obtenir des objets différents sur des horodatages différents. Pour ce faire, il s'agit tout d'abord de sélectionner les BBs respectant l'hypothèse de l'emprunt maximal de tel sorte à ce que les BBs se chevauchent le moins possible sur l'ensemble du site. Ce sous-ensemble est appelé "candidate BBs". Cependant, un problème se pose car étant donné que les horodatages sont du même site, il est fort probable que des objets se superposent fortement lors de la sélection de "candidate BbsB". Un algorithme a donc été mise en place qui permet de sélectionner un ensemble d'objets pour minimiser autant que possible le degré de superposition.

Après le principe de sélection des BBs vient la phase de la construction du graphe. Dans cette étape, chaque BBs étant un ensemble d'objet peut être organisés et stockés sous forme d'un graphe d'évolution. La construction du graphe se fait en reliant les objets de l'horodatage  $i$  à ceux de l'horodatage  $i+1$ . Chaque objet correspond à un nœud du graphe et le poids de l'arrête représente le degré de chevauchement qui existe entre ces objets. Par contre, il est important de noter que deux objets connectés proviennent nécessairement de deux horodatages différents. Le nombre de sommet sera donc égale au nombre d'images obtenues au cours de la série temporelle. Ce type de graphe est appel graphe orienté temporelle.

Un graphe d'évolution est associé à un et un seul BB, mais peut représenter plusieurs couvertures spatiales. Nous distinguons principalement deux types de couvertures à savoir le graphe de la couverture boîte limite (BBCov) et le graphe de la couverture entière (WholeCov). Le BBCov est la représentation d'un BB dans une ancienne BB ou encore dans une carte par exemple tandis que le WholeCov représente l'étendue spatiale totale de tous les Bbs dans le graphe le long des horodatages. Afin de pouvoir représenter ou de faire la part entre les couvertures qui reviennent

de façon régulière et celles qui ne le sont pas, il a été nécessaire de subdiviser le WholeCov en deux parties. Il se compose d'une couverture éphémère (EphemCov) et d'un graphe de couverture principale (CoreCov). Le EphemCov est l'ensemble des zones qui ont été couvertes une seule fois durant le temps de surveillance (les horodatages) et le CoreCov regroupe l'ensemble des zones qui ont été couvertes au moins deux fois sur les mêmes horodatages. Lorsque le pourcentage de EphemCov est élevé, cela indique une forte instabilité des objets contenus dans le graphe. Cette instabilité peut être expliquée par des évolutions transitoires dans la zone d'étude ou encore par une mauvaise segmentation. La phase de la segmentation joue donc un rôle très important pour l'extraction du graphe d'évolution. Cependant comment faire pour savoir si l'on a une bonne segmentation ou pas? A priori, il est évident que le CoreCov englobe le BBCov ainsi qu'une zone tampon autour de celle-ci. L'ensemble du graphe fourni une bonne représentation lorsque l'écart entre ces deux couvertures est négligeable. En ce qui concerne la mesure des évolutions spatio-temporelle, il est essentiel d'analyser et de comprendre l'information qui pourrait se cacher derrière les graphes d'évolutions. C'est ainsi qu'intervient l'analyse sur les deux niveaux à savoir le niveau du graphe et aussi au niveau du site d'étude en question.

L'analyse du graphe est axée sur la façon dont les objets sont connectés et aussi de comment leurs attributs évoluent le long des horodatages. Dans le dernier cas, l'accent est mis sur l'ensemble du site d'étude mais on se focalise plus sur la répartition des entités spatiales les plus stables et les plus dynamiques. Étant donné un graphe  $G$ ,  $G(i)$  l'ensemble des objets couverts par  $G$  à l'horodatage  $i$  et avec  $W(j,k)$  le poids du lien entre l'objet  $O(j)$  et  $O(k)$ . La formule ci-dessous permet de calculer la variation (Var) entre deux horodatages consécutifs.

$$Var(G_i, G_{i+1}) = \sum_{o_j \in G_i} \frac{size(o_j)}{size(G_i)} \cdot \frac{\sum_{o_k \in G_{i+1}} w_{j,k} \cdot dist(o_j, o_k)}{\sum_k w_{j,k}}$$

Fig. 1. Formule de la variation (Var)

Dans cette formule nous distinguons principalement deux grandes parties. La première est proportionnelle à l'importance de l'objet  $O(j)$  sur l'ensemble des objet des à l'horodatage  $i$ . La deuxième partie évalue l'évolution d'un objet entre deux horodatages consécutifs. L'ensemble de la variation ou variation globale du graphe est obtenue en cumulant la contribution de chaque paire d'horodatages comme illustré dans la formule ci-dessous:

Le GlobalVar est un indicateur variant de 0 à l'infini qui permet d'estimer l'évolution de la surface d'étude au cours de la période couverte par la série chronologique. Lorsqu'elle est faible, cela implique un comportement temporel est stable tandis qu'une valeur élevée traduit une importante évolution ou variation temporelle du milieu.

Pour une analyse au niveau du site d'étude, les différentes valeurs du GlobalVar obtenue pour chaque graphe d'évolution peut être utilisé pour faire ressortir une carte GlobalVar. Dans ce cas, l'une des quelconques couvertures

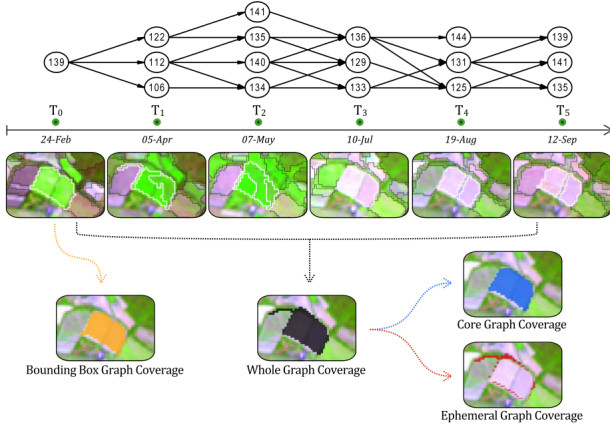


Fig. 2. Exemple d'un graphe d'évolution

$$GlobalVar(G) = \sum_{i=1}^{n-1} Var(G_i, G_{i+1})$$

Fig. 3. Formule de la variation globale (GlobeVar)

peut être utilisée pour faire ressortir la carte (ex: CoreCov). La carte GlobaVar ainsi obtenue résume la distribution des différents phénomènes présents sur le site et fournit des informations relatives à l'intensité des évolutions au cours du temps.

De façon générale, le modèle comporte trois paramètres que sont  $a$ ,  $b1$  et  $b2$ . Le paramètre  $a$  permet de limiter la superposition entre les BBs tandis que  $b1$  et  $b2$  permettent d'éviter la sélection d'objets non représentatifs. Ces trois paramètres sont déterminés en fixant un seuil (il est exprimé en pourcentage de la zone entière) qui définit la couverture minimale acceptée.

## 6 EXPÉRIMENTATION

Dans le but de pouvoir valider le modèle proposé, des expériences ont été menées principalement sur deux sites différentes que sont Libron Valley (A) et Lower Aude Valley (B) sur un suivi total de six (06) horodatages allant de février à septembre 2009. Ces images sont issues du satellite Landsat.

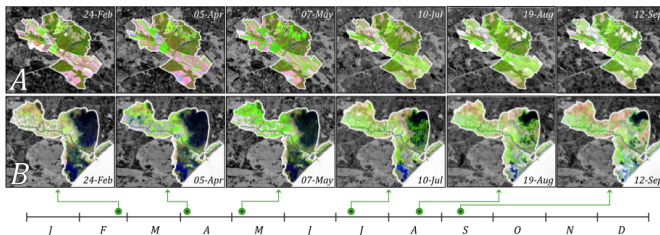


Fig. 4. Séries temporelles des sites A et B

|                 | Lower Aude Valley |      |       | Libron Valley |      |       |
|-----------------|-------------------|------|-------|---------------|------|-------|
|                 | Min               | Mean | Max   | Min           | Mean | Max   |
| Number of nodes | 7                 | 15.2 | 38    | 6             | 13.0 | 26    |
| Number of edges | 7                 | 24.5 | 77    | 5             | 18.9 | 53    |
| Number of paths | 2                 | 79.7 | 1050  | 1             | 46.7 | 480   |
| BBcov (ha)      | 1.6               | 16.2 | 125.0 | 3.2           | 13.8 | 67.1  |
| WholCov (ha)    | 5.3               | 46.2 | 175.7 | 6.1           | 34.5 | 107.9 |
| CoreCov (ha)    | 2.1               | 29.9 | 142.5 | 3.4           | 22.6 | 94.6  |
| CoreCov (%)     | 13.9              | 65.1 | 90.9  | 14.4          | 66.3 | 93.3  |
| EphemCov (ha)   | 0.8               | 16.4 | 134.0 | 1.2           | 11.8 | 65.1  |
| EphemCov (%)    | 9.1               | 34.9 | 86.1  | 6.7           | 33.7 | 85.6  |

Fig. 5. Tableau récapitulatif des observations

## 7 RÉSULTATS OBTENUS

Le tableau ci-dessus nous montre les résultats obtenus à l'issue de l'étude des sites. Nous pouvons par exemple remarquer que le site de Aude Valley, le graphe généré se composait d'un nombre de nœuds allant de 7 à 38 et aussi d'une couverture spatiale WholCov allant de 5,3 hectares à 175,7 hectares. La figure suivante nous montre un exemple d'évolution d'un des graphes extrait parmi l'ensemble de graphes obtenus sur les deux sites. A partir de cette fig-

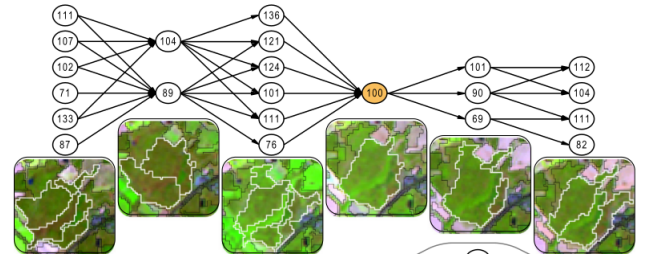


Fig. 6. Exemple d'un graphe obtenu

ure, nous pouvons bien constater que le BB correspond à l'image obtenue à la période T3 ou au quatrième horodatage qui correspond à un début de l'été où la zone est plus homogène, tandis que les situations les plus hétérogènes sont observées dans le premier et le troisième horodatage. Il a été démontré également qu'il est possible, à partir de ce modèle de mettre en évidence la pertinence des zones à partir de leurs indices spectrales. Cette technique a été utilisée dans le cas d'une étude sur zones agricoles. Les indices mis en évidence étaient l'indice de végétation (NDVI), l'indice de l'eau (NDWI), l'indice de la sécheresse (VSDI). L'étude a permis de mettre en évidence et de faire ressortir différents types de cultures pratiquées dans une zone et de suivre leurs évolutions.

## 8 CONCLUSION

cette étape de notre étude s'est essentiellement basée sur une analyse non approfondie sur le sujet mais avait pour but de cerner l'idée générale des auteurs de cet article. La prochaine étape de notre étude sera consacrée sur une analyse critique et comparative afin de lever les forces et les faiblesses de cette études par rapport à d'autres approches et de proposer des perspectives si nécessaire.

## REFERENCES

- [1] F. Guttler, D. Ienco, J. Nin, M. Teisseire, and P. Poncelet, "A graph-based approach to detect spatiotemporal dynamics in satellite image time series," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, no. Supplement C, pp. 92 – 107, 2017.