



# Analyse de la structure des haies dans les vergers pour la définition de paysages mieux adaptés contre les bioagresseurs

S. Da Silva<sup>\*,\*\*</sup>, C. Lavigne<sup>\*\*</sup>, Florence Le Ber<sup>\*,\*\*\*</sup>

<sup>\*</sup>LORIA UMR 7503, 54500 Vandœuvre-lès-Nancy

<sup>\*\*</sup>INRA, UR 1115 Plantes et Systèmes de culture Horticoles, 84000 Avignon

<sup>\*\*\*</sup>LHyGeS UMR 7517 - ENGEEES, 67000 Strasbourg

Conférence SAGEO - Juillet 2011

**Résumé.** Afin de modéliser la structure des haies de vergers, nous utilisons des méthodes de fouille de données. Ainsi, les chaînes de Markov couplées avec les chemins de Hilbert-Peano et les processus ponctuels servent à déterminer les paramètres propres à la disposition des haies dans le paysage. Ceci permet, par la suite, de simuler des paysages valides afin de créer des scénarios paysagers luttant contre les ravageurs.

**Abstract.** We use data mining methods in order to modelize the structure of hedges in orchards. Markov chains combined with Hilbert-Peano curves, or point processes are able to reveal parameters describing the hedge structures. The final aim is to simulate realistic landscapes to evaluate scenarios of pest control

**Mots clés.** Modélisation de paysage, Fouille de données spatiales, Chaînes de Markov, Processus Ponctuel, Chemin de Hilbert-Peano

**Keywords.** Landscape modelling, Spatial Data mining, Markov chains, Point process, Hilbert-Peano curves

## 1 Introduction

Que cela soit pour l'aménagement du territoire, le suivi de l'implantation des arbres dans les forêts ou bien, la gestion des cultures, la modélisation et la simulation du paysage et de ses composantes sont un sujet de préoccupation des instances publiques. En particulier, la prise en compte récente de la nécessité de préserver l'environnement oblige les acteurs du secteur agricole à chercher des solutions permettant de réduire l'utilisation des pesticides et autres produits phytosanitaires. Dans ce souci de création d'une agriculture responsable, une solution envisagée serait de rendre les cultures et les paysages intrinsèquement défavorables aux organismes néfastes. Il s'agit là de production intégrée des cultures, et c'est dans ce cadre que se déroule le travail décrit ci-dessous.

L'objectif finalisé de ces recherches consiste à créer des scénarios paysagers pour contrer les ravageurs des cultures, en particuliers le carpocapses des pommes et des poires. Pour cela, il est nécessaire de pouvoir s'appuyer sur des simulations de paysages reprenant certaines caractéristiques des paysages réels. La première partie de ce travail mènera justement à la création de modèles permettant la simulation des haies dans les paysages agricoles. À celui-ci sera couplé un simulateur de parcellaire inspiré de celui déjà présenté dans [Le Ber *et al.*, 2009]. Nous obtiendrons ainsi un outil capable de simuler des paysages agricoles ayant des caractéristiques pouvant être contrôlées. La suite du travail sera d'utiliser

les données existantes sur le carpacape pour confronter sa dynamique aux descripteurs des paysages simulés.

## 2 Méthodes et données

Nous ferons appel pour ce travail à deux catégories de méthodes. Les méthodes de Markov sont utilisées pour la résolution de problèmes venant de différents horizons. Ainsi, elles sont particulièrement bien implantées dans le secteur de la reconnaissance de la parole ou de l'écriture [Rigoll *et al.*, 1996], elles apparaissent dans les outils de représentation de molécule en 3D [Alexandrov *et al.*, 2004], et même, dans la modélisation d'un comportement piétonnier [Makris *et al.*, 2002]. D'une manière plus parlante pour nous, elles sont également introduites de la création de certaines cartes économiques liées au tissu urbain [Huang *et al.*, 2008], dans la reconnaissance de forme sur des images satellitaires ou de simples photos, bruitées ou non [Benmiloud *et al.*, 1995], [Pieczynski, 2003], [Benboudjema *et al.*, 2007] ou dans l'étude de successions de cultures agricoles [Le Ber *et al.*, 2006], [Lazrak *et al.*, 2010], [Mari *et al.*, 2006].

En ce qui concerne les processus ponctuels, il existe deux catégories intéressantes pour nous, tout d'abord, les processus ponctuels marqués présents par exemple dans la détection de forme ou l'extraction d'objets dans une image [Chatelain *et al.*, 2009, Lafarge *et al.*, 2010, Stoica, 2001]. Il existe aussi les processus ponctuels de Markov toujours dans le domaine de la télédétection [Van Lieshout, 2000] ou dans celui de la gestion des forêts [Gaetan *et al.*, 2008].

Pour construire les modèles, nous possédons un jeu de données fourni par l'Unité PSH de l'INRA d'Avignon, sous forme d'un fichier *shape*. Pour sa création, un opérateur humain a, grâce à un logiciel SIG, pointé les haies, une à une, sur une photo aérienne représentant une zone de vergers de fruits à pépins. Cette zone s'étend au nord des Bouches du Rhône, couvre environ  $70 \text{ km}^2$ , et abrite 11501 haies. Une fois le fichier obtenu, nous extrayons les points d'attaches de chaque haie et, après son conditionnement, nous déterminons le barycentre, la longueur et l'angle, par rapport à l'axe Est/Ouest, de chacune. Ceci nous permettra d'effectuer dans un premier temps des calculs de statistique simples. Nous pouvons, par exemple, effectuer une segmentation des haies avec d'une part, les haies dites "brise vent" étant parallèles par groupe et d'autre part, les haies "de bordure".

## 3 Traitement envisagé

Une fois ces premiers résultats obtenus, nous nous attacherons à explorer les données avec les différents points de vue cités précédemment avec la chronologie suivante. Dans un premier temps, nous utiliserons les méthodes de Markov, plus précisément, les chaînes de Markov couplées à l'utilisation des chemins de Hilbert-Peano [Benmiloud *et al.*, 1995]. Puis, nous envisagerons l'utilisation des processus ponctuels, marqués d'une part et de Markov de l'autre. Il existe une partie commune à toutes les méthodes évoquées : nous réduisons chaque haie à son milieu, auquel nous attachons un vecteur contenant des informations (descripteurs géométriques, biologiques...).

Pour les chaînes de Markov utilisant les chemins de Hilbert Peano, nous découpons la carte des données en fenêtre de taille fixe (on pourra la faire varier ensuite pour déterminer l'influence de ce paramètre dans les résultats). Nous nous retrouvons alors avec un plan contenant des vecteurs associés à des points. Ensuite, nous définissons un chemin de Hilbert-Peano parcourant ce plan (la granularité et le sens de parcours pourront être également envisagés comme "paramètres"). Chaque barycentre sera alors ordonné suivant ce chemin. Au terme de cette étape, nous aurons un ensemble de suites de barycentres qui devra être traité par un algorithme permettant d'extraire les paramètres d'une chaîne de Markov. Ceux-ci seront ensuite utilisés pour la simulation qui reprendra un chemin identique à la modélisation. Le problème peut venir de la difficulté à simuler une zone d'une autre taille

que celle modélisée. On peut également envisager dans cette partie de ne pas prendre des fenêtres de taille fixe, mais des zones ayant une réalité agronomique et utiliser avec cela des modèles hiérarchiques en plus des chaînes de Markov. Les processus ponctuels appréhenderont l'espace en entier mais utiliseront des règles de voisinage différentes selon qu'ils sont marqués ou de Markov.

## References

- [Alexandrov *et al.*, 2004] Alexandrov V., Gerstein M., « Using 3D Hidden Markov Models that explicitly represent spatial coordinates to model and compare protein structures », *BMC Bioinformatics*, vol. 5, n 1, p. 2, 2004.
- [Benboudjema *et al.*, 2007] Benboudjema D., Pieczynski W., « Unsupervised Statistical Segmentation of Nonstationary Images Using Triplet Markov Fields », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, n 8, p. 1367 -1378, aug., 2007.
- [Benmiloud *et al.*, 1995] Benmiloud B., Pieczynski W., « Estimation des paramètres dans les chaînes de Markov cachées et segmentation d'images », *Traitement du signal*, vol. 12, n 5, p. 433 - 454, 1995.
- [Chatelain *et al.*, 2009] Chatelain F., Descombes X., Zerubia J., « Estimation des paramètres de processus ponctuels marqués dans le cadre de l'extraction d'objets en imagerie de télédétection », *GRETSI*, Dijon France, 2009.
- [Gaetan *et al.*, 2008] Gaetan C., Guyon X., *Modélisation et statistique spatiales*, Springer, 2008.
- [Huang *et al.*, 2008] Huang R., Kennedy C., « Uncovering Hidden Spatial Patterns by Hidden Markov Model », in T. Cova, H. Miller, K. Beard, A. Frank, M. Goodchild (eds), *Geographic Information Science*, vol. 5266 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, p. 70-89, 2008.
- [Lafarge *et al.*, 2010] Lafarge F., Gimel'farb G., Descombes X., « Geometric Feature Extraction by a Multimarked Point Process », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, n 9, p. 1597 -1609, sept., 2010.
- [Lazrak *et al.*, 2010] Lazrak E., Mari J.-F., Benoît M., « Landscape regularity modelling for environmental challenges in agriculture », *Landscape Ecology*, vol. 25, p. 169-183, 2010. 10.1007/s10980-009-9399-8.
- [Le Ber *et al.*, 2006] Le Ber F., Benoît M., Schott C., Mari J.-F., Mignolet C., « Studying crop sequences with CarrotAge, a HMM-based data mining software », *Ecological Modelling*, vol. 191, n 1, p. 170 - 185, 2006. Selected Papers from the Fourth International Workshop on Environmental Applications of Machine Learning, September 27 - October 1, 2004, Bled, Slovenia.
- [Le Ber *et al.*, 2009] Le Ber F., Lavigne C., Adamczyk K., Angevin F., Colbach N., Mari J.-F., Monod H., « Neutral modelling of agricultural landscapes by tessellation methods - Application for gene flow simulation », *Ecological Modelling*, vol. 220, n 24, p. 3536 - 3545, 2009.
- [Makris *et al.*, 2002] Makris D., Ellis T., « Spatial and probabilistic modelling of pedestrian behaviour », 2002.
- [Mari *et al.*, 2006] Mari J.-F., Le Ber F., « Temporal and spatial data mining with second-order hidden markov models », *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 10, p. 406-414, 2006. 10.1007/s00500-005-0501-0.

- [Pieczynski, 2003] Pieczynski W., « Modèles de Markov en traitements d'images », *Traitement du Signal*, vol. 20, n 3, p. 255 -278, 2003.
- [Rigoll *et al.*, 1996] Rigoll G., Kosmala A., Rattland J., Neukirchen C., « A comparison between continuous and discrete density hidden Markov models for cursive handwriting recognition », *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, vol. 2, p. 205 -209 vol.2, August, 1996.
- [Stoica, 2001] Stoica R., Processus ponctuels pour l'extraction de réseaux linéïques dans les images satellitaires et aériennes, PhD thesis, Université de Nice-Sophia Antipolis, Nice, France, 2001.
- [Van Lieshout, 2000] Van Lieshout M., *Markov point processes and their applications*, Imperial College Press, 2000.