

SEGMENTATION DES LIGNES ET DES COLONNES D'UNE MATRICE : APPLICATION AUX SÉQUENCES DE NAVIGATION VISUELLE

Vincent Brault¹ & Adeline Leclercq-Samson¹ & Jean-Charles Quinton¹

¹ *Univ. Grenoble Alpes, LJK, F-38000 Grenoble, France*
CNRS, LJK, F-38000 Grenoble, France

Résumé. Dans la recherche sur les véhicules autonomes, nous sommes souvent amenés à synthétiser de longues séquences vidéos de l'environnement (Korrapti et al. 2013) en séparant les lieux qualitativement distincts (e.g. ligne droite, intersection...) (Birem et al. 2014). En résumant la séquence par une matrice de similarité entre les images, l'objectif revient alors à estimer les frontières entre des blocs relativement homogènes.

Pour répondre à cette question, il existe des algorithmes développés pour l'analyse des données Hi-C issue de la biologie (Dixon et al. 2012) dont la problématique est similaire. En particulier, Brault et al. (2016a) proposent une segmentation rapide en supposant que les observations sont constantes par bloc et Brault et al. (2016b) étudient une segmentation fondée sur des statistiques de rang que nous proposons d'utiliser.

Dans cet exposé, nous proposons de comparer les deux procédures dans le cas particulier de la segmentation de séquences vidéo obtenues dans le cadre de l'étude sur les véhicules autonomes. Après avoir rappelé le principe de chaque méthode, nous montrerons les segmentations obtenues.

Mots-clés. Segmentation, Statistique non paramétrique, LASSO, Algorithme de programmation dynamique.

Abstract. In the research about the autonomous vehicles, we need to synthesize large videos footage of the environment (Korrapti et al. 2013) by separating the distinct places (e.g. straight line, intersection...) (Birem et al. 2014). By summarizing the footage by a matrix of similarity between the images, the goal is to estimate the boundaries of homogeneous blocks.

This problem is similar to the analyse of the biologic Hi-C data (Dixon et al. 2012) and some algorithms exist. Notably, Brault et al. (2016a) have developed a quickly segmentation if the matrix is block wise constant and Brault et al. (2016b) have studied a segmentation based on a rank statistic.

In this talk, we compare the two procedures for the segmentation of some videos footage obtained in the framework of the autonomous vehicles. After detailing each algorithm, we will present the estimated segmentations.

Keywords. Break-points, non parametric statistic, LASSO, Dynamic programming algorithm.

1 Introduction

Dans la recherche sur les véhicules autonomes, de longues séquences vidéos des environnements sont capturés afin de permettre aux véhicules de se localiser et de naviguer. De larges environnements peuvent alors être synthétisés par des cartes connectant des lieux qualitativement distincts (e.g. ligne droite, intersection...). Par diverses méthodes de description des images (e.g. points d'intérêt et caractéristiques locales), on définit une signature visuelle de chaque image et on peut ainsi calculer une similarité entre images observées à différents moments (Birem et al. 2014). Une fois regroupées dans une matrice symétrique (voir la partie gauche de la figure 1), ces valeurs de similarité forment des blocs relativement homogènes.

Notre but est de segmenter les lignes et les colonnes de cette matrice afin de faire ressortir ces blocs et de pouvoir étudier les zones en interactions les unes avec les autres. Pour cela, nous étudions la vidéo faite par Korrapati et al. (2013) dont nous avons représenté les coordonnées GPS des images sur la partie droite de la figure 1.

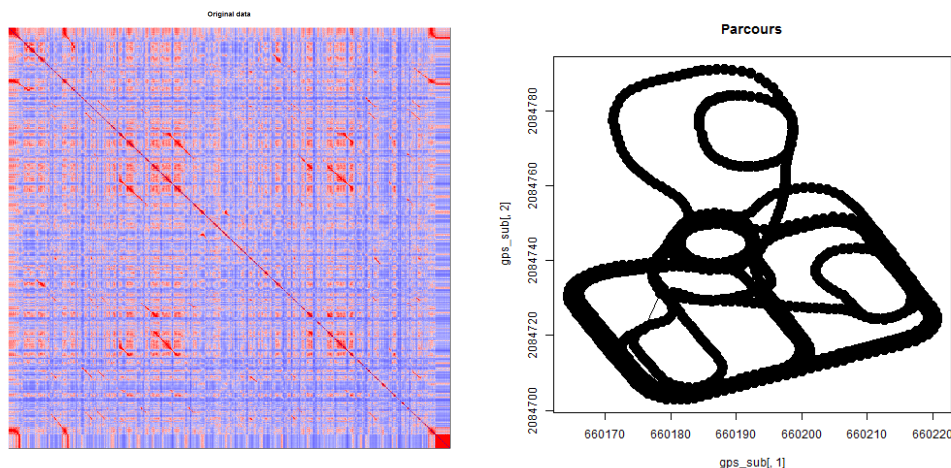


FIGURE 1 – À gauche : représentation de la matrice étudiée ; une couleur rouge indique que les images se ressemblent fortement tandis qu’une couleur bleue correspond à une faible ressemblance. À droite : coordonnées GPS du véhicule correspondantes à chaque image de la séquence vidéo étudiée.

2 Présentation des modèles

Dans cette partie, nous explicitons les procédures utilisées. Pour cela, nous commençons par présenter les hypothèses communes aux deux modèles puis les particularités de

chaque procédure.

2.1 Notations générales

Les données sont représentées sous la forme d'une matrice symétrique $\mathbf{Y} = (Y_{ij})_{1 \leq i, j \leq n}$ de taille $n \times n$ où chaque case Y_{ij} représente la similarité entre les images i et j . Nous supposons qu'il existe K^* ruptures en lignes (et par conséquent en colonnes), notées $1 \leq t_1^* < \dots < t_{K^*}^* \leq n-1$ représentant les frontières des $(K^* + 1)^2$ blocs recherchés (voir la figure 2). Pour tout couple $(k, \ell) \in \{1, \dots, K^* + 1\}^2$, nous notons $\mathbf{R}_{k, \ell}^*$ le bloc défini par :

$$\mathbf{R}_{k, \ell}^* = \{(i, j) \in \{1, \dots, n\}^2 \mid t_{k-1}^* + 1 \leq i \leq t_k^* \text{ et } t_{\ell-1}^* + 1 \leq j \leq t_\ell^*\}$$

avec la convention que $t_0^* = 0$ et $t_{K^*+1}^* = n$.

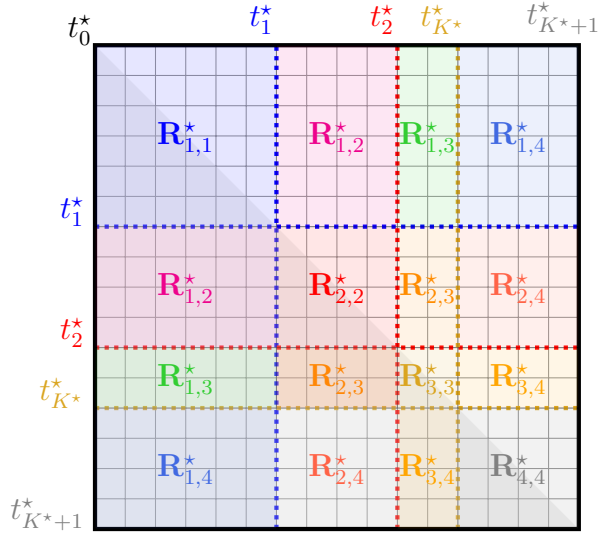


FIGURE 2 – Représentation schématique de la matrice.

2.2 Méthode paramétrique blockseg

La première méthode proposée par Brault et al. (2016a) suppose que la matrice $\mathbf{Y}$ peut se décomposer en somme de deux matrices $\mathbf{Y} = \mathbf{U} + \mathbf{E}$ où $\mathbf{U} = (U_{i,j})$ est une matrice constante par blocs définie par

$$U_{i,j} = \mu_{k, \ell}^* \quad \text{si } (i, j) \in \mathbf{R}_{k, \ell}^*.$$

Les coefficients $E_{i,j}$ de la matrice $\mathbf{E} = (E_{i,j})_{1 \leq i, j \leq n}$ sont supposés indépendants, de même loi et de moyenne nulle. Ainsi, les coefficients $Y_{i,j}$ sont supposés être des variables indépendantes avec des moyennes constantes par bloc.

En effectuant deux transformations, les auteurs montrent que cette modélisation est équivalente au système linéaire suivant :

$$\mathcal{Y} = \mathcal{X}\mathcal{B} + \mathcal{E}$$

où $\mathcal{Y}$ et $\mathcal{E}$ sont les vectorisations respectives des matrices $\mathbf{Y}$ et $\mathbf{E}$, $\mathcal{X} = \mathbf{T} \otimes \mathbf{T}$ est la matrice issue du produit de Kronecker de deux matrices $\mathbf{T}$ triangulaires inférieures avec uniquement des 1 et $\mathcal{B}$ est un vecteur de $\mathbb{R}^{n^2}$ comportant au plus $(K^* + 1)^2$ valeurs non nulles.

Ils montrent alors que la recherche des ruptures est équivalente à la recherche des valeurs non-nulles du vecteur $\mathcal{B}$ et proposent d'estimer les ruptures à l'aide d'une procédure LASSO associée à une adaptation de la méthode *stability selection* de Meinshausen et Bühlmann (2010).

Notons que, dans le cas de matrices symétriques, l'hypothèse d'indépendance n'est pas vérifiée. Nous avons donc choisi de suivre la procédure proposée dans leur article et disponible dans le package `blockseg`¹.

2.3 Méthode non paramétrique MuChPoint

Pour l'autre procédure, Brault et al. (2016b) se placent dans un cadre non-paramétrique. Les auteurs supposent que les variables Y_{ij} situées au-dessus de la diagonale (donc vérifiant $j \geq i$) sont indépendantes et qu'à chaque bloc $\mathbf{R}_{k,\ell}^*$ est associé une loi $\mathbb{P}_{k,\ell}$. Notons que, comme la matrice $\mathbf{Y}$ est symétrique, les lois $\mathbb{P}_{k,\ell}$ et $\mathbb{P}_{\ell,k}$ sont identiques.

Pour estimer les ruptures, les auteurs étendent au cas multivarié la statistique de rang de Wilcoxon dans le prolongement de l'article de Lung-Yut-Fong et al. (2012). Pour cela, ils définissent pour chaque ligne $i \in \{1, \dots, n\}$ la statistique

$$U_{n,i}(t_1, \dots, t_K) = \frac{4}{n^2} \sum_{k=0}^K (t_{k+1} - t_k) \left(\bar{R}_k^{(i)} - \frac{n+1}{2} \right)^2$$

où $\bar{R}_k^{(i)} = \frac{1}{t_{k+1} - t_k} \sum_{j=t_k+1}^{t_{k+1}} R_j^{(i)}$ est la moyenne des rangs $R_j^{(i)}$ au sein de la ligne i des cases $Y_{i,t_k+1}, \dots, Y_{i,t_{k+1}}$. À l'aide de ces statistiques, ils définissent

$$S_n(t_1, \dots, t_K) = \sum_{i=1}^n U_{n,i}(t_1, \dots, t_K) = \frac{4}{n^2} \sum_{k=0}^K (t_{k+1} - t_k) \sum_{i=1}^n \left(\bar{R}_k^{(i)} - \frac{n+1}{2} \right)^2$$

et estiment les ruptures en minimisant cette statistique :

$$(\hat{t}_1, \dots, \hat{t}_K) \in \underset{1 \leq t_1 < \dots < t_K \leq n-1}{\operatorname{argmin}} S_n(t_1, \dots, t_K).$$

Pour faire cette estimation, les auteurs utilisent l'algorithme de programmation dynamique mais ne fournissent pas de méthodes d'estimation du nombre de ruptures.

1. Disponible à l'adresse <https://cran.r-project.org/web/packages/blockseg/index.html>

3 Résultats sur les données de navigation visuelle

Dans cette partie, nous présentons quelques résultats obtenus sur la matrice réduite de dissimilarité obtenue à partir de la séquence vidéos d'un véhicule autonome (figure 1). Étant donné un nombre de ruptures, nous avons comparé les distances de Hausdorff entre les partitions renvoyées par chacune des procédures. Sur la figure 3, nous observons que cette valeur diminue avec le nombre de ruptures mais qu'elle ne semble pas tendre vers 0.

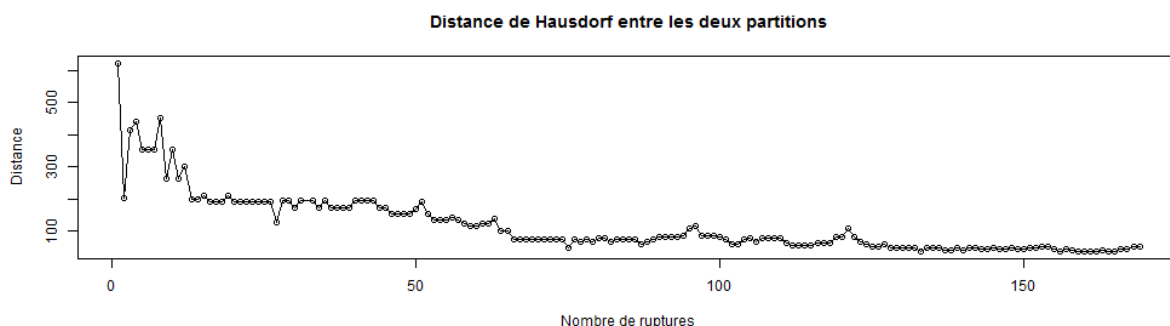


FIGURE 3 – Évolution de la distance de Hausdorff entre les partitions obtenues par les deux méthodes en fonction du nombre de ruptures conservé.

Si nous observons de plus près les partitions obtenues pour 50 ruptures (voir la figure 4), nous observons que la méthode **blockseg** met des ruptures très rapprochées plus souvent que la méthode **MuChPoint**. Il serait intéressant de vérifier si ces ruptures correspondent à des virages. Nous constatons que dans les deux cas, le premier et le dernier bloc de la diagonale ont été assez bien séparés (même si la procédure **MuChPoint** a mieux subdivisé la dernière partie). Ces blocs correspondent aux moments où le véhicule était à l'arrêt sur deux zones spatialement très proches.

4 Conclusion

Dans cet exposé, nous comparerons les partitions obtenues en fonction du nombre de ruptures souhaitées et évaluerons la qualité de ces partitions en fonction des coordonnées GPS associées aux images à partir de séquences plus complètes.



FIGURE 4 – Représentation de la matrice initiale (à gauche) et des matrices résumées par la moyenne pour chaque bloc dans le cas de 50 ruptures : la méthode **blockseg** au milieu et la méthode **MuChPoint** à droite. Pour chaque matrice résumée, un bloc est représenté par la valeur moyenne.

Bibliographie

- [1] M. Birem, J.-C. Quinton, F. Berry et Y. Mezouar(2014). Sail-map : Loop-closure detection using saliency-based features. Dans *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4543–4548. IEEE, 2014.
- [2] Brault, V., Chiquet, J. et Lévy-Leduc, C. (2016). Fast Detection of Block Boundaries in Block-Wise Constant Matrices. In *Machine Learning and Data Mining in Pattern Recognition*, pages 3–10, 2010. Springer International Publishing.
- [3] Brault, V., Ouadah, S., Sansonnet, L. et Lévy-Leduc, C. (2016). Nonparametric homogeneity tests and multiple change-point estimation for analyzing large Hi-C data matrices. *arXiv preprint arXiv :1605.03751*.
- [4] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu et B. Ren (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398) : 376–380.
- [5] H. Korrapati, J. Courbon, S. Alizon et F. Marmonton (2013). " the institut pascal data sets" : un jeu de données en extérieur, multicateurs et datées avec réalité terrain, données d'étalonnage et outils logiciels. Dans *Orasis, Congrès des jeunes chercheurs en vision par ordinateur*.
- [6] A. Lung-Yut-Fong, C. Lévy-Leduc, and O. Cappé (2012). Distributed detection/localization of change-points in high-dimensional network traffic data. *Statistics and Computing*, 22(2) :485-496.
- [7] Meinshausen, N. et Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 72(4), 417-473.