

# Estimation de flux d'objets fluorescents dans les séquences de vidéo-microscopie

Jérôme BOULANGER<sup>1</sup>, Thierry PÉCOT<sup>2</sup>, Charles KERVRANN<sup>3</sup>, Patrick BOUTHEMY<sup>3</sup>, Jean SALAMERO<sup>1</sup>

<sup>1</sup>UMR144 Institut Curie/CNRS, 12 rue Lhomond, 75005 Paris

<sup>2</sup>The Ohio State University, Columbus, Ohio 43210, USA

<sup>3</sup>Inria, Centre Rennes Bretagne Atlantique, Campus de Beaulieu, 35042 Rennes Cedex

jerome.boulanger@curie.fr, thierry.pecot2@osumc.edu, charles.kervrann@inria.fr,  
patrick.bouthemy@inria.fr, jean.salamero@inria.fr

**Résumé** – L'analyse automatique du contenu dynamique de séquences de vidéo-microscopie est un enjeu important pour la compréhension de certains mécanismes moléculaires impliqués dans les fonctions cellulaires. Nous proposons dans cet article une approche originale pour appréhender les mouvements de vésicules dans les séquences de vidéo-microscopie. Nous réduisons en effet le contenu dynamique de ces séquences à des flux entre régions d'intérêt. Nous étudions dans ce cadre les performances de différentes méthodes pour l'estimation de ces flux.

**Abstract** – Automatic analysis of dynamic content of video-microscopy image sequences is crucial for understanding molecular mechanisms involved in cell functions. In this paper, we propose an original approach for the analysis of the motions of vesicles in these sequences. We handle this issue by considering flows of objects within predefined regions. We investigate several methods to solve this problem and compare their performances in different situations.

## 1 Introduction

Les mécanismes moléculaires régissant la coordination du transport de protéines membranaires au sein des cellules sont encore méconnus. La microscopie de fluorescence permet de localiser au moyen de sondes la distribution de protéines d'intérêt au cours du temps. L'analyse du comportement dynamique de ces protéines permet de mieux comprendre les mécanismes sous-jacents. Une des méthodes les plus utilisées pour extraire ce type d'information consiste à réaliser un suivi individuel des objets tels que les vésicules de transport qui apparaissent alors dans les séquences d'images comme des particules brillantes de quelques pixels en déplacement rapide [6]. En pratique, les fréquences d'échantillonnage temporel sont limitées par les phénomènes de photo-toxicité, la densité des objets élevée, leur grande similarité et leurs vitesse de déplacement sont autant d'obstacles à la mise en œuvre d'algorithmes de suivi.

Dans cet article, nous décrivons une approche originale permettant l'analyse des flux d'objets entre régions connexes dans des séquences d'images de vidéo-microscopie à partir de la variation du nombre d'objets dans ces régions [4]. Les flux entre régions permettent de comprendre ainsi les relations entre les différents domaines de la cellule. Il s'agit d'un problème générique qui se retrouve dans l'étude de différentes voies et fonction (endocytose recyclages, motilité, etc). Dans le cadre de cette approche, les apparitions et les disparitions d'objets peuvent être prises en compte en utilisant une région virtuelle

supplémentaire. Ainsi en plus des flux observés entre régions dans le domaine de l'image, le nombre des événements d'apparitions et de disparitions telles qu'ils peuvent être observés dans des modalités d'imagerie où seule une fraction seule de la cellule est imagée (par exemple en imagerie par ondes évanescentes). Une autre application possible de cette approche est la comptabilisation du nombre d'objets entrant dans une région d'intérêt sur une période donnée. Cela permet ainsi de quantifier le nombre d'interactions entre la zone considérée et les objets nécessaires au déclenchement de certains processus.

Hors quelques cas triviaux, l'estimation de ces flux est mal posée et nous proposons deux approches de régularisation pour l'estimation de ces flux. Nous évaluons ces deux approches sur des données simulées et nous décrivons les résultats obtenus sur des données réelles de vidéo-microscopie. L'approche ainsi proposée peut constituer une étape préliminaire dans l'estimation des flux entre régions non-connexes (tomographie de réseaux) telle que proposée dans [5].

## 2 Méthode

### 2.1 Détection des vésicules en microscopie de fluorescence

Dans un premier temps, afin de localiser les vésicules dans la séquence d'images, nous avons recours à un détecteur de "blob" tel que le filtre des différences de Gaussienne. La dis-

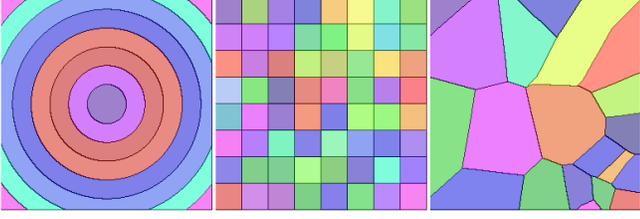


FIGURE 1 – Trois types de partition du domaine de l’image.

tribution de la réponse du filtre est modélisée par une loi de Laplace, ce qui nous permet d’établir à partir d’une probabilité de fausse alarme  $\alpha$ , un seuil pour la détection des objets d’intérêt.

## 2.2 Partition de l’image

Nous considérons le partitionnement du domaine  $\Omega$  de l’image en régions et nous associons à cette partition un graphe orienté  $G(E, V)$  comprenant  $|E|$  arêtes correspondant aux bords des régions et  $|V|$  sommets correspondant aux régions elles-mêmes. En plus des régions visibles dans l’image, nous définissons une région virtuelle (notée avec l’indice  $i^*$ ) connectée à toutes les régions et permettant de prendre en compte les apparitions et disparitions d’objets dans  $\Omega$ . Le graphe est orienté car nous considérons par la suite que les flux entre les régions seront positifs.

Différentes stratégies de partition sont possibles (voir Figure 1). La première consiste simplement à diviser le domaine de l’image en une grille régulière de taille  $n \times m$ . Dans ce cas de figure, le nombre de sommets du graphe est  $|V| = nm$  et le nombre d’arêtes est  $|E| = 4nm - 2n - 2m$ . Une tessellation de Poisson peut être aussi considérée pour obtenir une partition de l’image. Enfin, l’information provenant d’un marquage complémentaire peut être exploitée pour définir ces régions d’intérêt à partir d’une segmentation de l’image correspondante.

## 2.3 Estimation des flux

### 2.3.1 Méthode par appariements

Une méthode simple pour estimer les flux sur les arêtes du graphe consiste à chercher d’une image à l’autre une correspondance entre les objets. Nous considérons ici une approche très naïve basée sur un appariement avec le plus proche voisin à l’instant  $t + 1$ . Dans cette version, il n’est pas possible de prendre en compte l’apparition et la disparition d’objets. Enfin, la complexité de cette méthode dépend principalement du nombre total d’objets et peu du nombre de régions.

### 2.3.2 Méthode basée sur l’analyse des variations du nombre d’objets dans chaque région

Nous nous intéressons à une autre classe d’approche ne nécessitant pas de mise en correspondance. En effet, une fois les régions définies, il est possible de compter le nombre  $Y(t) =$

$\{y_i(t), i \in [1, |V|]\}$  d’objets détectés situés à l’intérieur de ces régions d’indice  $i$  à chaque instant  $t$ . L’équation de conservation (loi de Kirchoff) suivante :

$$\frac{\partial Y}{\partial t} - MX = 0 \quad (1)$$

permet de relier le flux  $X(t) = \{x_j(t), i \in [1, |E|]\}$  de particules sur les arêtes d’indice  $j$  avec la variation temporelle du nombre d’objets dans chaque région. Dans ce contexte, l’opérateur  $M$  de divergence est donné par une matrice de taille  $|V| \times |E|$  telle que pour chaque région d’indice  $i$  (y compris la région virtuelle d’indice  $i^*$ ) et chaque arête d’indice  $j$  :

$$m_{ij} = \begin{cases} 1 & \text{si l’arête } j \text{ aboutit au sommet } i, \\ -1 & \text{si l’arête } j \text{ part du sommet } i, \\ 0 & \text{sinon.} \end{cases} \quad (2)$$

Nous pouvons noter que comme le nombre d’arêtes est plus grand que le nombre de sommets, cette matrice a plus de colonnes que de lignes ( $|E| > |V|$ ) en général.

L’estimation des flux  $X(t)$  à partir des mesures  $Z = \partial Y(t)/\partial t$  nécessite l’inversion de l’opérateur  $M$  ou encore la minimisation de l’erreur  $\|Z - MX\|^2$ . Si le nombre de particules situés dans la région invisible  $i^*$  n’est pas connue, il est possible de déduire ses variations à partir du nombre total d’objets présents dans l’ensemble des régions :

$$z_{i^*}(t) = - \left( \sum_{i \in V \setminus i^*} y_i(t+1) - \sum_{i \in V \setminus i^*} y_i(t) \right) \quad (3)$$

Étant donné que le graphe est orienté, les flux sur les arêtes sont toujours positifs et nous pouvons utiliser cette contrainte pour réduire l’espace des solutions. Ainsi, un algorithme de type moindres carrés avec contrainte de positivité permettrait d’estimer le flux le long des arêtes du graphe si le nombre d’arêtes est plus faible que le nombre de sommets. La matrice  $M$  étant mal conditionnée dans la plupart des configurations (nous avons vu que  $|V| < |E|$  dans l’exemple d’une grille régulière), nous proposons de régulariser dans un premier temps le problème en utilisant une méthode de coupure spectrale et nous estimons le flux le long des arêtes avec l’algorithme des moindres carrés avec contrainte de positivité NNLS de Lawson et Hanson [3].

La probabilité que la trajectoire d’une particule intersecte l’interface entre deux régions et donc traverse l’arête du graphe est relativement faible. Ainsi la solution recherchée à chaque instant est un vecteur parcimonieux. Nous proposons donc en seconde approche à minimiser une énergie du type :

$$J(X, Z) = \frac{1}{2} \|Z - MX\|^2 + \kappa |X| + \frac{1}{2} d_C^2(X) \quad (4)$$

avec  $d_C$  représente la fonction distance à l’espace  $C$  des solutions positives et  $\kappa$  un réel positif. Pour estimer des flux  $X$  sur les arêtes du graphe, nous proposons d’utiliser un algorithme de type PPXA [1] qui fait appel aux opérateurs proximaux associés aux différents termes de l’énergie proposée (voir

algorithme 1) afin de prendre en compte les différentes contraintes proposées. En particulier, l'opérateur associé à la fonction distance au carré est donné par [1] :  $\frac{1}{2}(x + P_C(x))$ , avec  $P_C(x) = x$  si  $x > 0$  et 0 sinon.

**Data:**  $M, Z, \varepsilon \in [0, 1], \gamma > 0, \eta > 0$

**Result:**  $X_N$

$U_{1,1} = 0, U_{2,1} = 0$  **for**  $n \in [1, N]$  **do**

$$P_{1,n} = (I + \gamma M^T M)^{-1}(U_{1,n} + \gamma M^T Y)$$

$$P_{2,n} = \frac{1}{2}(U_{2,n} + P_C(U_{2,n}))$$

$$P_{3,n} = \max(U_{3,n} - \frac{1}{\kappa}, 0) \text{sign}(U_{3,n})$$

$$P_n = P_{1,n} + P_{2,n} + P_{3,n}$$

$$\varepsilon \leq \lambda_n \leq 2 - \varepsilon$$

**for**  $i \in [1, 3]$  **do**

$$U_{i,n+1} = U_{i,n} + \lambda_n (2P_n - X_n - P_{i,n})$$

**end**

$$X_{n+1} = X_n + \lambda_n (P_n - X_n)$$

**end**

**Algorithm 1:** Algorithme de minimisation de type PPXA utilisé pour estimer le flux sur les arêtes du graphe. On reconnaîtra l'intervention des opérateurs proximaux associées aux différentes contraintes.

## 3 Résultats expérimentaux

### 3.1 Données synthétiques

Dans un premier temps, nous évaluons les performances des méthodes proposées à partir de données synthétiques. Afin de découpler l'évaluation de la méthode de détection et celle de la méthode d'estimation de flux, nous considérons que les détections sont idéales en utilisant les positions des particules correspondant à des trajectoires simulées et nous cherchons à mesurer à l'influences des modèles de mouvements ainsi qu'à la densité de particules dans l'image.

Pour évaluer la qualité de l'estimation, nous comparons les passages sur les arêtes du graphes obtenues par les différentes méthodes d'estimations de flux et la vérité terrain. Les flux vrais doivent être calculés à partir des trajectoires des objets et cette vérité terrain est obtenue avec une méthode similaire à la celle par appariement, mis à part que dans ce cas, la mise en correspondance est exacte. Nous calculons ensuite pour chaque simulation les probabilité de détection (sensibilité)  $P_d$  comme le rapport entre le nombre de bonnes détections et le nombre d'évènements simulés ainsi que la probabilité de fausses alarmes (spécificité)  $P_{fa}$  comme le rapport entre le nombre de fausses détections et le nombre d'évènements détectés.

Nous avons simulé différents scénarios correspondant à différentes classes de mouvement pour un nombre donné  $K$  de particules dans le domaine de l'image. Le premier scénario correspond à une marche aléatoire, le second à un mouvement dirigé (en cercle) et le dernier à une marche aléatoire avec des apparitions (avec une probabilité fixée à 0.05) et des disparitions (avec une probabilité identique). Nous étudions l'influ-

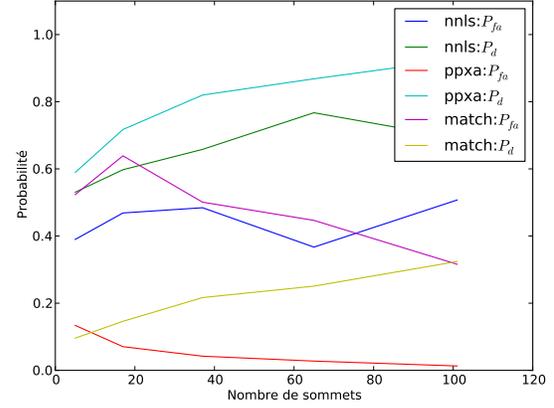


FIGURE 3 – Évaluation de l'influence du nombre de régions sur l'estimation dans le cas d'un mouvement correspondant à 30 marches aléatoires avec apparition et disparition d'objets.

ence du nombre d'objets  $K$  en le faisant varier dans l'intervalle  $[10, 100]$  et pour chaque nombre d'objets, nous faisons la moyenne de 10 réalisations pour chacune des méthodes (ce qui correspond à  $3 \times 45 \times 10 = 1350$  simulations par scénario). La géométrie retenue pour cette série d'expériences est une partition régulière en  $2 \times 2 = 4$  blocs. Pour ces simulations le temps de calcul de la méthode par appariement croît avec  $K$  de  $10ms$  à  $1s$ , la méthode exploitant l'algorithme NNLS prend  $2s$  et celle s'appuyant sur l'algorithme PPXA de  $10ms$  à  $200ms$  en fonction du nombre d'itérations nécessaires pour atteindre un plateau pour l'énergie  $J$ . Dans ces deux derniers cas, les temps de calculs ne dépendent pas explicitement de  $K$ .

Nous remarquons sur la figure 2, que la méthode reposant sur l'analyse des plus proches voisins (match) est plus performante dans le cas des marches aléatoires (Fig. 2.a). L'approche par étude des variations temporelles du nombre d'objets devient plus compétitive dans les cas où les mouvements sont dirigés (Fig. 2.b) ou que des évènements d'apparition et de disparition se produisent (Fig. 2.c). Pour cette approche, la méthode utilisant l'algorithme PPXA semble plus performante que celle utilisant l'algorithme NNLS.

Nous avons aussi évalué l'influence du nombre de régions sur ces algorithmes dans le cadre du dernier scénario avec  $K = 30$  trajectoires. La figure 4 montre que dans ce cas précis, l'augmentation du nombre de régions améliore les résultats. Nous pouvons interpréter ce résultat comme une réduction du nombre d'ambiguïtés.

### 3.2 Données réelles

Nous avons appliqué la méthode basé sur l'algorithme PPXA à l'analyse du trafic membranaire de la protéine Rab 11 impliquée dans la voie d'endocytose recyclage (voir Figure 4)[2]. La protéine Rab11-mCherry fluorescente est associée à certaines vésicules qui quittent l'endosome pércentriolaire (au centre de l'image de gauche et situé en profondeur dans la cellule)

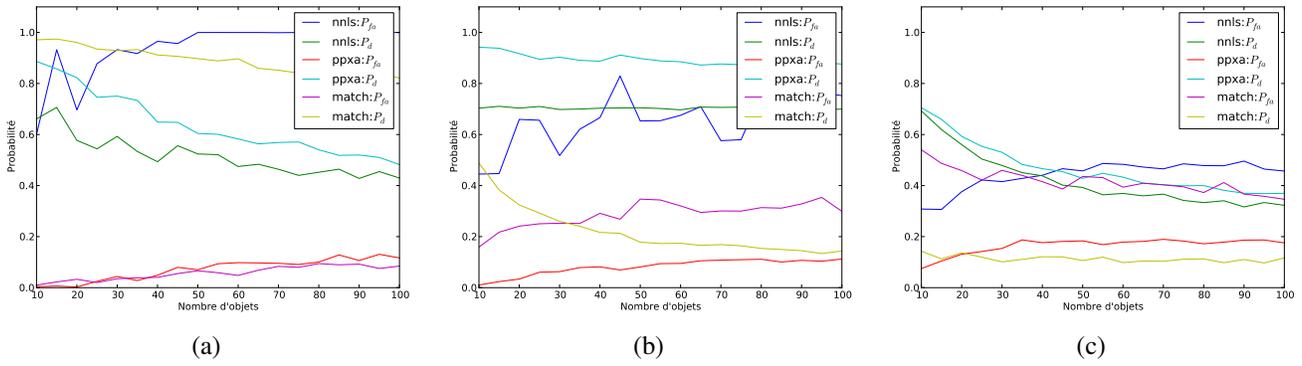


FIGURE 2 – Evaluation des performances des trois méthodes (NNLS, PPXA et appariement) sur 3 types de scénarios avec une partition régulières à 4 régions. (a) marches aléatoires, (b) mouvements circulaires, (c) marches aléatoires avec apparition et disparition d’objets.

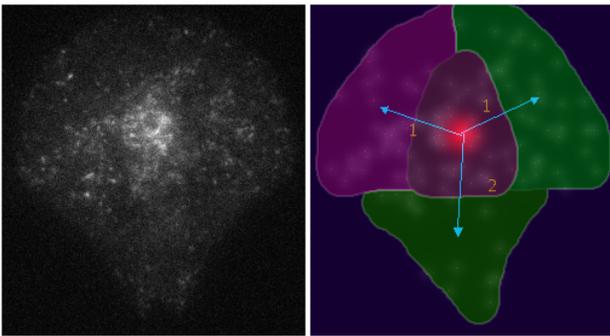


FIGURE 4 – Analyse du trafic de Rab11 avec la méthode basée sur l’algorithme PPXA. À droite les données originales sont représentées. À gauche l’estimation des flux entre les régions est représentée sous forme de flèches. Les petits points rouges indiquent la localisation des particules détectées tandis que le signe rouge plus intense indique la disparition d’objets dans la région.

et qui vont rejoindre la membrane plasmique (en particulier à l’interface avec la boîte de culture). Les vésicules associées à Rab11 ont la particularité d’exhiber une dynamique très rapide ce qui rend leur suivi particulièrement délicat. Une partie de ces vésicules vont venir fusionner avec la membrane et ainsi disparaître de l’image. Les images ont été acquises avec un microscope à onde évanescente (TIRF), ainsi seule une partie de la cellule est imagée.

L’utilisation d’une région virtuelle permet de prendre en compte un nombre variable d’objets mais aussi de pouvoir localiser approximativement leur apparition et leur disparition dans la séquence d’images.

## 4 Conclusion

Nous avons proposé différentes méthodes pour estimer les flux d’objets entre les régions obtenus par une partition de l’image. Nous avons conduit un certain nombre de simulation numériques pour étudier l’influence de différents paramètres du problèmes.

Nous avons observé que la méthode basée sur l’analyse des variations du nombre d’objets dans chaque région permet de mieux prendre en compte les évènements d’apparitions et de disparitions et se comporte mieux dans le cas de mouvement dirigés. Ces deux cas de figures sont fréquents en vidéo-microscopie ce qui rend cette approche intéressante dans ce contexte.

## Références

- [1] P. L. Combettes and J.-C. Pesquet. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, chapter 10. Proximal splitting methods in signal processing, pages 185–212. Springer, New York, 2011.
- [2] A. Gidon, S. Bardin, J. Boulanger, F. Waharte, B. Cinquin, L. Heliot, H. de la Salle, D. Hanau, C. Kervrann, B. Goud, and J. Salamero. Dynamics of Rab11a-MyoVb-FIP2 complex frames the late recycling steps of langerin from the ERC to the plasma membrane. *Traffic*, 13 :815–833, 2012.
- [3] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*. Society for Industrial and Applied Mathematics, 1974.
- [4] T. Pécot. *Modélisation et estimation du trafic intracellulaire par tomographie de réseaux et microscopie de fluorescence*. PhD thesis, Université de Rennes I, Rennes, March 2010.
- [5] Th Pécot, Ch Kervrann, and P. Bouthemy. Minimal paths and probabilistic models for origin-destination traffic estimation in live cell imaging. In *IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’08)*, pages 843–846, may 2008.
- [6] I. Smal, E. Meijering, K. Draegestein, N. Galjart, I. Grigoriev, A. Akhmanova, M. E. van Royen, A. B. Houtsmuller, and N. Niessen. Multiple object tracking in molecular bioimaging by Rao-Blackwellized marginal particle filtering. *Medical Image Analysis*, 12(6) :764–777, 2008.