



Version française / [English version](#)

Propositions de post-doc (2 postes ouverts)

Détection « à la volée » d'Événements dans les Séquences vidéo par des méthodes structurales et Bayésiennes (projet ANR DESCRIBE) - [LIP6/UPMC](#), [MATIS/IGN](#) et [LTCI/Télécom ParisTech](#)

Dans le cadre de ce projet, nous nous intéressons à la détection « à la volée » d'événements remarquables dans les séquences vidéo. Nous appelons « événement remarquable » tout événement qui se distingue de son environnement de par son comportement spatial et/ou temporel. Dans notre cas, nous ne disposons pas d'apprentissage, mais cherchons à caractériser des « ruptures » par rapport au contexte, ce qui nous permet de fournir un cadre plus général. En effet, ce n'est plus seulement un « acteur » (toute entité se distinguant de son environnement par sa dynamique ou son apparence visuelle) et son comportement individuel qui nous intéressent, mais son comportement par rapport à son environnement. Chaque acteur sera observé dans une zone (dite « zone de surveillance »), qui sera détectée à l'aide de primitives génériques bas niveau locales ou semi-locales (points d'intérêt, edgels, régions, etc.) à spécifier, sur lesquelles seront appliquées des méthodes de regroupement pour former les zones de surveillance et leur caractérisation. Les liens des acteurs les uns avec les autres ou avec le contexte seront modélisés par des relations spatiales, temporelles et spatio-temporelles, définies entre les zones de surveillance ou encore entre ensembles de descripteurs, afin d'en donner une description compacte. Ces relations seront comparées, et leurs évolutions au cours du temps étudiées. La modélisation de relations entre ensembles épars de primitives visuelles (notamment de points d'intérêt), de relations spatio-temporelles et la définition de mesures de comparaison constituent une partie originale du projet, pour laquelle nous proposons de nous appuyer sur le formalisme de la morphologie mathématique et des ensembles flous. Un module de suivi avancé, correspondant à un filtre particulière modélisé par un réseau bayésien dynamique, permettra de suivre au cours du temps les zones de surveillance. L'intérêt est double : d'abord, le filtre optimal sera guidé dans son processus de prédiction et d'estimation entre deux instants par les relations spatiales calculées à l'instant antérieur, ce qui affinera la qualité de suivi ; ensuite, ces estimations seront exploitées de manière à améliorer la localisation des descripteurs et des régions d'observation dans la tranche de temps suivante. De plus, de manière à prendre en compte l'évolution des relations entre zones de surveillance ainsi que celle du nombre de zones, nous proposons de faire évoluer au cours du temps la structure du réseau bayésien dynamique (qualifié ainsi de « non stationnaire ») : toutes les caractéristiques de la scène observée seront ainsi intégrées « à la volée » dans le réseau bayésien, ce qui constitue une autre originalité du projet. La reconnaissance d'événements se fera soit de manière temporelle, soit de manière spatiale. Le réseau bayésien dynamique non stationnaire sera apte à détecter des ruptures en cas de changement de sa structure entre deux tranches de temps. De plus, la mesure de la qualité de l'estimation permettra de détecter un mauvais suivi à l'instant courant et sera donc un signe de rupture contextuelle puisque les relations spatiales n'auront pas permis de guider correctement le processus de prédiction. Enfin, la mesure de l'évolution des relations spatiales au cours du temps nous permettra elle aussi de qualifier des ruptures locales. L'établissement d'un certain nombre de scénarios types (personne abandonnant un objet, mouvement d'une personne contrastant avec celui de la foule environnante...) nous permettra de tester notre approche, qui sera validée ensuite sur des bases de séquences dédiées à la détection d'événements dont nous construirons une vérité terrain.

Nous nous intéresserons aux aspects suivants du projet : **définition de régions d'intérêt, points caractéristiques ou objets dans les images, modélisation et calcul de relations spatiales et spatio-temporelles, définition et calcul de mesures de comparaison adaptées à ces relations.**

Deux postes sont proposés :

1. Dans le premier, une attention particulière sera accordée à la modélisation de **relations entre ensembles de points épars**, qui est une partie originale du projet. Il s'agira donc de développer

des méthodes de description et modélisation des relations spatiales strictes et floues entre entités visuelles complexes représentées par des ensembles épars (primitives locales de bas niveau comme les points d'intérêt).

2. Dans le second, on s'intéressera à la **comparaison de relations**, en fonction de leur modélisation (par des nombres, des nombres flous, des intervalles, etc.), et la modélisation de **relations spatio-temporelles** entre objets (représentés sous la forme d'ensembles, d'ensembles flous, puis éventuellement de points épars).

Le travail sera effectué principalement avec Valérie Gouet-Brunet à l'IGN et Isabelle Bloch à Télécom ParisTech, tout en gardant une collaboration étroite avec les partenaires du LIP6 et un lien avec les autres parties du projet.

Organisation

- Projet financé par l'ANR
- Diplôme requis : doctorat
- Connaissances requises : traitement d'images, extraction de caractéristiques, segmentation, reconnaissance des formes, programmation en C/C++
- Date limite de candidature : jusqu'à ce que les postes soient pourvus
- Dossiers de candidature : ils doivent comporter
 - un CV détaillé avec liste de publications
 - une lettre de motivation
 - un résumé de la thèse
 - 2 ou 3 lettres de recommandations
 - les dates et la durée souhaitées
 - le sujet (1 ou 2) préféréet être envoyés à Valérie Gouet-Brunet (valerie.gouet@ign.fr) et à Isabelle Bloch (Isabelle.Bloch@telecom-paristech.fr) **sous la forme d'un seul fichier PDF.**
- Durée du post-doc : 12 à 18 mois
- Date de début envisagée : premier semestre 2013

English version / [Version française](#)

Post-doc positions (2 open positions)

Online event detection in video sequences using structural and Bayesian approaches (French ANR Project DESCRIBE) - [LIP6/UPMC](#), [MATIS/IGN](#) et [LTCI/Telecom ParisTech](#)

The goal of this project is to detect "online" noticeable events in video sequences. By "noticeable event" we mean any event that draws attention by its spatial and/or temporal behavior. In this project, we wish to develop a general-purpose detection framework. This rules out supervised learning-based algorithms which require specific training data. Rather, we propose to characterize noticeable events as break points w.r.t. their context, which is more suited for our goal. Actually, we are not only interested in studying only the actors of the scene (i.e., any entity whose dynamics or visual aspect draws attention) and their individual behaviors, but also in studying their behaviors w.r.t. their environment. Each actor will be observed in a so-called "observation" region that will be determined using local or semi-local low-level generic primitives (points of interest, edgels, areas, etc.) still to be defined, as well as clustering methods applied on these primitives. The relationships between different actors or between the actors and the context will be modeled by spatial, temporal and spatio-temporal relations. Those will be defined over observation regions or over sets of descriptors to result in compact descriptions. In addition, they will be compared and their evolution over time studied. Modeling relations between scattered sets of visual primitives (including points of interest) as well as spatio-temporal relations, and defining their comparison measures is one of the novel contributions of this project. For this purpose, we propose to rely on both mathematical morphology and fuzzy sets. In addition, observation regions will be accurately

tracked by an advanced tracking module based on a particle filter relying on a dynamic Bayesian network. The advantage of exploiting this module is twofold: first, the optimal filter will be guided in its prediction and estimation processes between consecutive time slices by the spatial relations computed before, hence resulting in better trackings; second, these estimations will be exploited to improve the determination of the descriptors and the observation regions in the next time slice. As observation regions evolve over time (their number may for instance vary) as well as their relations, we propose to make the structure of the dynamic Bayesian network evolve accordingly (in this case, such a network is called non-stationary). Thus, all the features of the studied scene can be integrated "online" into the dynamic Bayesian network. This is another originality of the project. Noticeable event detection will be performed either on a temporal or a spatial basis. For the former case, the non-stationary dynamic Bayesian network will allow us to detect break points due to its structure shifts over time. In addition, measuring the quality of the estimations provided by the particle filter will allow us to detect inaccurate trackings, which may be an indicator of contextual break points highlighted by a misguidance of the prediction process due to the spatial relations. Finally, measuring the evolution over time of the spatial relations will also provide evidence of local break points. Our approach will be tested against different kinds of scenarios (somebody leaving an object, somebody moving very differently from the surrounding crowd, etc.). It will then be validated on databases of sequences dedicated to event detection whose ground truth we will construct.

We will focus on the following aspects of the project: **definition of regions of interest, key-points or objects in images, modeling and computation of spatial and spatio-temporal relations, definition and computation of comparison measures adapted to these relations.**

Two positions are open, on the following topics:

1. In the first one, a particular attention will be paid to the modeling of **relations between sparse sets of points**, as an original part of the project. The aim is then to develop methods for description and modeling of strict and fuzzy spatial relationships between complex entities represented with sparse sets (low-level local features such as interest points).
2. The second topic concerns the **comparison of relations**, according to their modeling (using numbers, fuzzy numbers, intervals, etc.), and the modeling of **spatio-temporal relations** between objects (represented as sets, fuzzy sets, and eventually sparse sets).

This work will be performed mainly with Valerie Gouet-Brunet at IGN and Isabelle Bloch at Telecom ParisTech, in close collaboration with the partners at LIP6 and while keeping a link with the other parts of the project.

Organization

- Project funded by ANR
 - Required diploma: PhD thesis
 - Required skills: image processing, feature extraction, segmentation, pattern recognition, programming in C/C++
 - Deadline for applications: open until filled
 - Applications must include:
 - a detailed CV with a description of realized projects
 - a motivation letter
 - a summary of the thesis
 - 2-3 recommendation letters
 - preferred starting dates and duration
 - preferred topic (1 or 2)
- and must be sent to Valerie Gouet-Brunet (valerie.gouet@ign.fr) and Isabelle Bloch (Isabelle.Bloch@telecom-paristech.fr) **in one PDF file.**
- Duration: 12 to 18 months
 - Beginning of the post-doc: first semester 2013