

## TP4 – Suivi d'un objet dans une vidéo par *mean-shift*

Les trois premiers TP de segmentation vous ont permis de découvrir trois méthodes de segmentation par classification : classification *supervisée* dans le TP1 (champs de Markov), classification *semi-supervisée* dans le TP2 (algorithme EM) et classification *non supervisée* dans le TP3 (méthode *mean-shift*). Nous vous proposons maintenant d'utiliser la méthode *mean-shift* pour effectuer le suivi d'un objet (*tracking*) dans une séquence vidéo. En suivant le lien ci-après, vous pourrez vous faire une idée du genre de résultats que l'on peut espérer atteindre :

<https://www.irit.fr/~Alain.Crouzil/jaffre/RECHERCHE/dea.html>

En lisant l'article suivant, vous pouvez également comprendre comment cela fonctionne :

[https://www.irit.fr/~Alain.Crouzil/jaffre/PUBLICATIONS/orasis2003\\_jaffre\\_crouzil.pdf](https://www.irit.fr/~Alain.Crouzil/jaffre/PUBLICATIONS/orasis2003_jaffre_crouzil.pdf)

Les deux vidéos qui vous sont fournies ont été découpées en séquences d'images, afin de faciliter leur lecture en Matlab (vous avez certainement entendu parler des problèmes de « codecs »). Lancez le script `lecture_video.m`. Les objets en mouvement se distinguent plus ou moins bien du fond, de par leur couleur. Le *suivi* d'un de ces objets nécessite de commencer par le détecter dans la première image de la séquence.

### Initialisation

La première image permet d'effectuer l'initialisation du suivi. Nous vous suggérons de procéder comme suit :

- Transformez l'image RVB en image LUV (cf. TP3 de segmentation).
- Demandez à l'utilisateur de prélever un échantillon à la souris (cf. TP1 de segmentation), de manière à « apprendre » sa couleur, qui peut par exemple être modélisée par une loi normale dans  $\mathbb{R}^3$ .
- Calculez la vraisemblance de chaque pixel, relativement à la loi apprise. Si possible, évitez les boucles `for` qui ralentissent énormément les calculs.
- Sélectionnez les pixels pour lesquels la vraisemblance est supérieure à un seuil  $S$  convenablement choisi.
- Les pixels sélectionnés forment un nuage de points de  $\mathbb{R}^2$  auquel la méthode *mean-shift* peut être appliquée. En choisissant comme initialisation le centre de l'échantillon, le mode vers lequel la méthode converge peut être considéré comme le point de départ de la trajectoire de l'objet suivi.

### Suivi

La suite du traitement est très simple : à chaque nouvelle image, la position de l'objet suivi est obtenue par *mean-shift*, en initialisant la recherche à la position trouvée à l'image précédente. Si le déplacement de l'objet dans l'image est continu, ce qui est généralement le cas, le déplacement entre deux images successives doit être petit. Par conséquent, un faible nombre d'itérations de la méthode *mean-shift* devrait suffire pour chaque image.

### Travail demandé

Pour ce TP, aucun script à trous ne vous est fourni. En suivant les conseils ci-dessus, vous devez parvenir à suivre certains des objets mobiles de la séquence **Pietons** (piétons ou véhicules), pour peu que la sélection à la souris d'un rectangle, sur la première image de la séquence, soit suffisamment précise.

Vous pourrez ensuite tester votre script sur la séquence **Football** : vous constaterez que, dans la plupart des cas, le suivi « décroche ». Vous pourrez tenter de corriger ce défaut en modélisant la zone sélectionnée, non plus par une loi normale dans  $\mathbb{R}^3$ , mais par un mélange de deux lois normales dans  $\mathbb{R}^3$  (cf. TP2 de segmentation).