

# INF6804

## VISION PAR ORDINATEUR

### *H2019 – Travail Pratique #3*

#### *Détection et suivi multiobjet*

---

##### **Objectifs :**

- Permettre à l'étudiant de se familiariser avec les méthodes de détection d'objet et/ou de suivi à travers des séquences d'images

##### **Remise :**

- Tout votre code source (si on veut ré-exécuter, on devrait être capable!)
- Un rapport (**format .pdf**) de 8 à 15 pages (environ)
- Remise avant le 12 avril 2019, 17h00, sur Moodle – **aucun retard accepté**
- **Vous devez aussi soumettre votre rapport sur TurnItIn!**
  - Enregistrez-vous sur [www.turnitin.com](http://www.turnitin.com) en utilisant les informations sur Moodle!

##### **Références :**

- Voir les notes de cours sur Moodle

##### **Autres directives :**

- Les TPs s'effectuent seul, ou en équipe de deux – ne remettez par contre qu'une version du rapport/code!
- 

### **Présentation**

Dans ce TP, vous devez proposer une méthode capable d'effectuer la détection et le suivi de multiples objets d'intérêt à travers une séquence vidéo. Les objets en question sont identifiés par une boîte englobante au début de la séquence; votre but est de les relocaliser de façon similaire dans toutes les autres images de cette séquence.

Nous vous fournissons une séquence vidéo sur Moodle pour laquelle seule la première boîte englobante vous est fournie. Vos résultats seront évalués sur cette séquence après la remise du TP à l'aide du *groundtruth* que nous gardons secret, et le classement final des méthodes de toutes les équipes sera publié sur Moodle. Notez que votre rang dans ce classement n'influence aucunement la note finale dans ce TP — tout comme précédemment, la présentation de votre méthode, de ses forces/faiblesses (dans votre rapport), et de vos expériences sont les parties les plus importantes du travail (voir le barème à la fin).

L'évaluation finale pour le classement des équipes sera effectuée par le calcul de l'intersection sur l'union (IoU) des boîtes englobantes produites par chaque méthode en fonction du *groundtruth* secret de la séquence fournie sur Moodle. Afin de valider vous-même la performance de votre méthode lors de son développement, on s'attend à ce que vous évaluiez vos résultats sur les séquences d'une autre base de données.

Pour la méthode à développer elle-même, vous pouvez encore une fois implémenter votre propre solution, ou bien utiliser/modifier une solution déjà existante. Notez que deux approches s'offrent à vous pour ce TP: vous pouvez utiliser une solution basée sur le suivi par détection (*tracking by detection*, qui implique l'entraînement *a priori* d'un détecteur d'objet), ou bien utiliser une solution de suivi plus classique (i.e. basée sur la modélisation de l'objet d'intérêt en ligne à partir de la première image, et sur la recherche locale de celui-ci dans les images subséquentes). Le choix de la solution devra être justifié par les défis identifiés dans la séquence sur Moodle.

Dans votre rapport, en plus de la description (**en profondeur!**) de votre solution, nous vous demandons de:

1. Identifier les difficultés propres à la séquence vidéo fournie sur Moodle pour la détection et le suivi de l'objet d'intérêt. En d'autres mots, qu'est-ce qui fait en sorte que cette séquence est difficile à traiter?
2. Justifier le choix de la solution proposée en fonction des problèmes identifiés ci-dessus. Quels problèmes devrait-elle résoudre facilement? Lesquels resteront difficile à traiter, et pourquoi?
3. Décrire l'implémentation de la solution choisie. Si vous n'avez pas tout écrit le code vous-même, d'où provient-il? A-t-il demandé des modifications? Sinon, de quels articles (ou site web) vous êtes-vous inspirés pour l'écrire? Dans tous les cas, quels sont les paramètres principaux utilisés? Comment ont-ils été choisis?
4. Présenter les résultats d'évaluation tirés de votre validation de performance sur l'autre ensemble de données public utilisé. Utilisez un format approprié — on s'attend à voir des tableaux, des figures, ...
5. Discuter des résultats du quatrième point en fonction des défis identifiés au premier point. Quel défis semblent être résolus par votre solution? Lesquels ne le sont pas? Peut-on s'attendre à une performance similaire pour la séquence sur Moodle?

Lors des séances de laboratoire, n'hésitez pas à poser des questions aux chargés — ils peuvent vous aider avec tout problème technique si vous travaillez sur Windows/Linux, ou bien avec votre code si vous travaillez avec du C/C++ ou avec Matlab.

## **Barême**

### **Rapport :**

- Présentation de la solution (allez dans les détails! 1-2 pages) = 4 pts
- Identification des difficultés dans la séquence sur Moodle = 3 pts
- Justification de la méthode par rapport aux difficultés identifiées = 2 pts
- Description de l'implémentation utilisée = 2 pts
- Présentation des résultats de validation = 3 pts
- Discussion des résultats = 3 pts
- Lisibilité, propreté et complétude = 3 pts

Total sur 20 pts.

Vous serez pénalisés de 50% de la note totale si vous ne remettez pas votre code. De plus, si votre rapport n'est pas remis sur *TurnItIn*, celui-ci ne sera pas corrigé. Pour ce qui est de la longueur, on s'attend à quelque chose entre 8 et 15 pages, mais si c'est plus long, ce n'est pas trop grave (ne remettez tout de même pas 50 pages...). L'ordre de la présentation n'est pas non plus très important, tant que tous les éléments ci-haut sont présents. N'oubliez pas de citer vos références!

## **Références**

- Aide-mémoire (« Cheat sheet ») Matlab :  
<http://web.mit.edu/18.06/www/Spring09/matlab-cheatsheet.pdf>
- Guide complet Matlab :  
[http://www.mathworks.com/help/pdf\\_doc/matlab/getstart.pdf](http://www.mathworks.com/help/pdf_doc/matlab/getstart.pdf)
- C++ : Open Source Computer Vision Library (OpenCV)  
<http://opencv.org/>  
<http://docs.opencv.org/doc/tutorials/tutorials.html>