

# ENSTA-Paris 3e année - Cours ROB317

## TP2 : Découpage et indexation de vidéos

Novembre 2020

L'objectif de ce TP est de se familiariser avec le traitement de vidéos sur OpenCV, en particulier l'estimation du champ de vitesses apparentes (flot optique), et de réfléchir à la problématique de l'indexation automatique de vidéos, dont le but est de sectionner la vidéos en plans, en attachant à chaque plan une description textuelle, ainsi qu'une image représentative du plan. L'indexation de vidéos comprend donc classiquement les tâches suivantes :

1. Détecter les changements de plans dans la vidéo.
2. Trouver pour chaque plan l'image la plus représentative.
3. Associer à chaque plan une description textuelle.

Dans ce TP on restreindra la description textuelle à la nature du mouvement de la caméra/optique et/ou des objets présents, en identifiant la nature du plan parmi les types (non exhaustifs) suivants :

- Plan fixe
- Panoramique horizontal (pan) ou vertical (tilt)
- Rotation / axe optique (roll)
- Travelling horizontal, vertical, avant, arrière...
- Zoom avant ou arrière

On pourra ajouter au type de plan une description du mouvement d'objets mobiles le cas échéant.

Vous devez fournir un rapport, au format pdf, en justifiant autant que possible vos réponses avec des expériences personnelles.

## 1 Préliminaires informatiques

Une base de code est fournie, qui fonctionne avec une installation standard d'OpenCV (sans opencv-contrib, version testée 4.1.0) et Python3. L'archive contenant cette base peut être téléchargée au lien suivant :

[http://perso.ensta-paris.fr/~manzaner/Cours/ROB317/TP2\\_Videos\\_Codes.zip](http://perso.ensta-paris.fr/~manzaner/Cours/ROB317/TP2_Videos_Codes.zip)

Tous les codes fournis sont des scripts Python3 éditables et exécutables en lançant la commande, par exemple :

```
$ python3 Dense-Optical-Flow.py
```

Des vidéos de test pour le découpage, ainsi que des vidéos d'expérimentation / réglage peuvent être téléchargées au lien suivant :

[http://perso.ensta-paris.fr/~manzaner/Cours/ROB317/TP2\\_Videos\\_Exemples.zip](http://perso.ensta-paris.fr/~manzaner/Cours/ROB317/TP2_Videos_Exemples.zip)

Merci à Juliette Drupt et Caroline Pascal (ENSTA 2019) qui ont produit ou compilé les vidéos d'expérimentation.

OpenCV est une bibliothèque très utilisée en analyse d'images et on pourra consulter avec profit de nombreux guides et tutoriels en ligne (attention toutefois aux versions de Python et d'OpenCV

fournies dans les exemples). On se limitera ici à mentionner les sites officiels pour le guide de référence des fonctions :

<https://docs.opencv.org/4.1.0/>

ainsi que le tutoriel :

<https://opencv-python-tutroals.readthedocs.io/en/latest/>

## 2 Histogramme de couleurs

Q1 CALCULER et afficher pour chaque image de la vidéo, sous la forme d'une image, l'histogramme 2d correspondant à la probabilité jointe des composantes chromatiques  $(u, v)$  du codage  $Yuv$  des images couleur. Observer les variations de l'histogramme au cours d'une séquence et proposer des premières mesures pour la détection des changements de plan. Que faire dans le cas de vidéos monochromes ?

## 3 Flot optique et histogramme de vitesses

Q2 ANALYSER le comportement et les arguments de la fonction de calcul du flot optique dense dans le script *Dense-Optical-Flow.py*, qui utilise la méthode de Farnebäck [1]. Expliquer le principe de cet algorithme, en particulier expliquer comment il peut calculer un flot dense, et comment il peut estimer de grandes vitesses.

Q3 CALCULER et afficher pour chaque image de la vidéo, sous la forme d'une image, l'histogramme 2d correspondant à la probabilité jointe des composantes  $(V_x, V_y)$  du flot optique. Observer les variations de l'histogramme et expliquer comment il peut être exploité pour identifier le type de plan.

## 4 Découpage et Indexation

Q4 PROPOSER une technique de découpage en plans à partir des mesures développées précédemment.

Q5 PROPOSER une mesure permettant d'extraire une image-clef représentative d'un plan.

Q6 PROPOSER une technique d'identification de plan (parmi les catégories données en introduction) permettant une description textuelle simple du plan.

Pour chaque réponse dans cette partie, on tentera de montrer quelques résultats en les critiquant et en pointant les principales difficultés rencontrées en fonction de la nature des plans, de la qualité de la vidéo ou des transitions.

## Références

- [1] Gunnar Farnebäck "Two-frame motion estimation based on polynomial expansion". 13th Scandinavian conf. on Image analysis (SCIA'03). Springer-Verlag, 2003, p.363-370