

ENSTA Paris 3e année - Cours CSC_5R013

TP4 : Analyse vidéo et Tracking

Janvier 2025

L'objectif de ce TP est de comprendre les enjeux et difficultés du suivi (tracking) d'objets dans les séquences vidéo, et d'expérimenter et programmer les solutions fondées sur les algorithmes de Mean Shift et de Transformées de Hough Généralisées.

Pour cela, nous utiliserons la bibliothèque de traitement d'images OpenCV sous Python (code testé avec python3 et OpenCV 4.1.0). Le code nécessaire pour le TP et des vidéos de test sont téléchargeables aux liens suivants :

- https://perso.ensta-paris.fr/~manzaner/Cours/CSC_5R013/Tracking_MeanShift.py
- https://perso.ensta-paris.fr/~manzaner/Cours/CSC_5R013/Test-Videos.zip

Vous devez rédiger - par binôme - un rapport que vous téléchargerez sur le moodle <https://moodle.ip-paris.fr/> Merci d'utiliser le format `Nom1+Nom2_Rapport_Tracking.pdf` pour nommer votre rapport.

Le but du suivi est de fournir, à chaque trame de la vidéo, la position d'un objet d'intérêt, ici sous la forme de coordonnées d'un rectangle englobant (RoI), dont les valeurs sont initialisées à la main sur l'image initiale de la vidéo (on définit le rectangle à la souris avec le bouton de gauche, puis on valide avec la touche 'q').

1 Mean Shift

Q1 EXPÉRIMENTER le suivi réalisé par le code de base `Tracking_MeanShift.py` fourni qui utilise l'algorithme de Mean Shift, avec la densité marginale f_H sur la composante H de teinte. Rappeler le principe de l'algorithme Mean Shift, et illustrer par vos expériences ses avantages et ses limitations.

Q2 ANALYSER plus finement le résultat en affichant la séquence des composantes couleur, ainsi que des poids définis par la rétro-projection R_H de l'histogramme f_H de teinte : $R_H(x, y) = f_H(H(x, y))$. Proposer et programmer des améliorations, en modifiant la densité calculée ou en combinant différentes cartes de poids.

2 Transformée de Hough

Q3 CALCULER à chaque trame, l'orientation locale, i.e. l'argument du gradient des pixels de l'image, ainsi que le module du gradient. Définir un seuil sur le module du gradient pour masquer les pixels dont l'orientation n'est pas significative. Afficher ainsi la séquence des orientations où les pixels masqués apparaissent en rouge. L'objectif de cette question est de définir l'index de la Transformée de Hough (l'orientation), ainsi que l'ensemble des pixels votants, i.e. ceux dont l'orientation est significative. Un exemple de résultat attendu est illustré Figure 1.

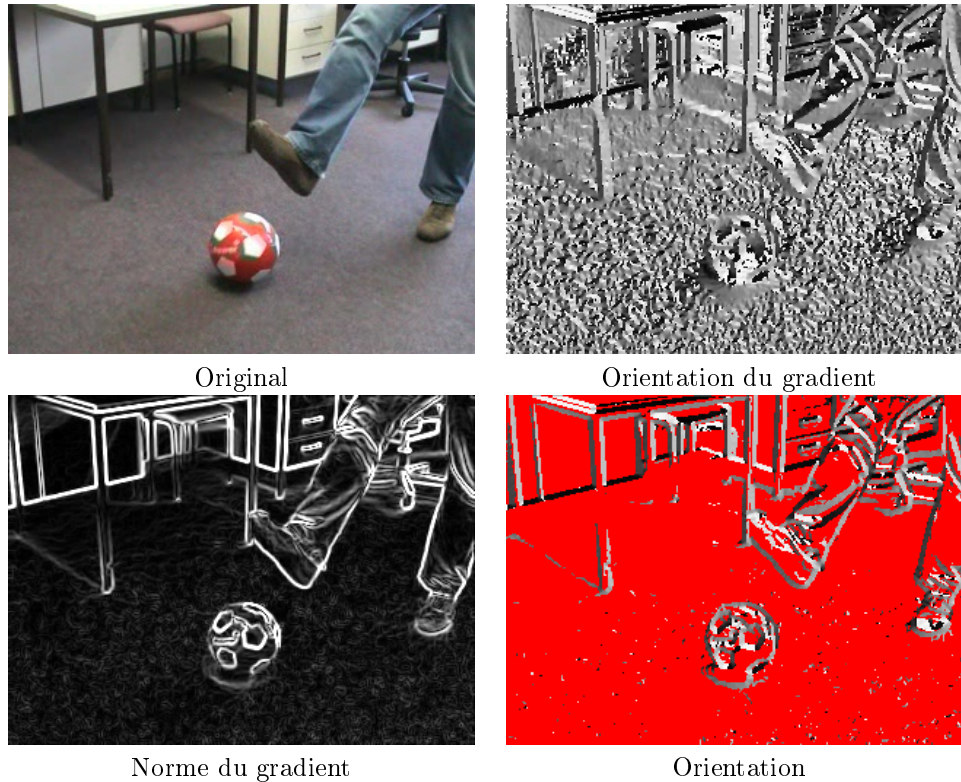


FIGURE 1 – Calcul de l'index de vote (orientation du gradient), avec sélection des pixels votants (norme du gradient).

Q4 CONSTRUIRE un modèle de l'objet défini initialement sous la forme d'un modèle implicite indexé sur l'orientation (R-Table). Puis calculer la transformée de Hough associée sur toutes les images de la séquence. Calculer le suivi correspondant à la valeur maximale de la transformée de Hough à chaque image. Commenter et critiquer le résultat obtenu. Illustrer vos réponses en montrant des exemples de transformées de Hough et des détections correspondantes (voir Figure 2).

Q5 PROPOSER (1) une stratégie simple de prédiction permettant d'améliorer le tracking en tirant parti de la régularité du déplacement, et (2) une stratégie de mise à jour du modèle pour gérer les changements d'aspect de l'objet ou les occultations.

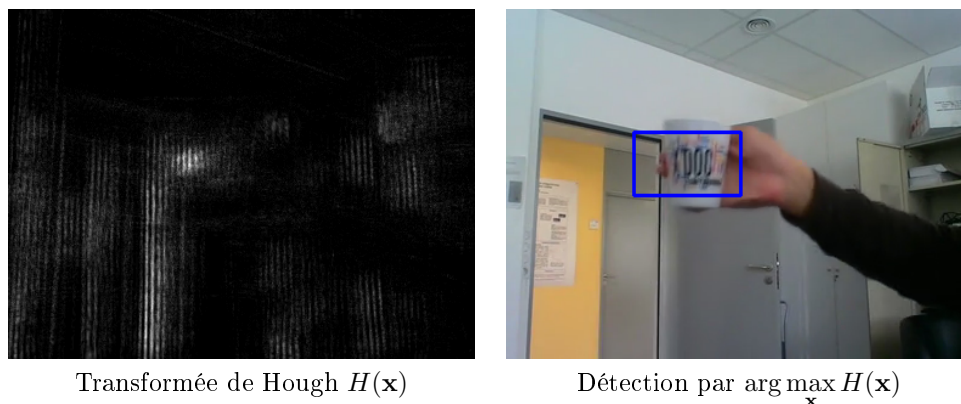


FIGURE 2 – Tracking par transformée de Hough.

3 Synthèse

Q6 EXPLIQUER comment améliorer les deux méthodes précédentes (i.e basées histogramme ou transformée de Hough) en utilisant des caractéristiques apprises par un réseau profond (Deep Features). Choisir un réseau existant, une couche et des canaux dans cette couche et utiliser ces caractéristiques dans le Mean Shift ou la transformée de Hough, au choix. Comment fonder le choix de la meilleure couche ? Et des meilleurs canaux ?