

Analyse du mouvement dans les vidéos et Suivi d'objets

Fabien Baldacci

Enseirb-Matmeca, IPB

Introduction

Vidéo

- ▶ Séquence d'images
- ▶ Chaque image pourrait être traitée indépendamment

Pourquoi l'analyse de vidéo ?

- ▶ Cohérence temporelle
- ▶ Temps de calcul
- ▶ Mouvement : information supplémentaire

Introduction

Que peut on faire ?

- ▶ Calculer le mouvement
- ▶ Déetecter les objets en mouvement
- ▶ Segmentation de mouvement
- ▶ Suivre les objets

Plan

Détection des objets en mouvement

Détection de mouvement

Extraction des objets à partir de la carte de mouvement

Flot optique

Segmentation de mouvement

Detect-before-track

Suivi basé sur les distributions

Segmentation dynamique

Détection des objets en mouvement

1. Détection de mouvement (quels pixels ont changé)
2. Extraction des objets à partir de la carte de mouvement

Détection de mouvement

Objectif :

identifier dans chaque image quels pixels ont bougé
i.e. quels pixels appartiennent à des objets mobiles



Détection de mouvement

Objectif :

identifier dans chaque image quels pixels ont bougé
i.e. quels pixels appartiennent à des objets mobiles



Détection des objets en mouvement

Détection de mouvement

Identifier dans chaque image quels pixels ont bougé

$$J(x, y, t) = |I(x, y, t) - I(x, y, t - 1)|$$

ou

$$J(x, y, t) = |I(x, y, t) - I_{ref}(x, y, t)|$$

Si $J(x, y, t) \neq 0$ alors le pixel a bougé

Sinon pas de mouvement.

Détection de mouvement

Identifier dans chaque image quels pixels ont bougé

$$J(x, y, t) = |I(x, y, t) - I_{ref}(x, y, t)|$$



Détection des objets en mouvement

Détection de mouvement

Identifier dans chaque image quels pixels ont bougé

$$J(x, y, t) = |I(x, y, t) - I(x, y, t - 1)|$$



Détection des objets en mouvement

Détection de mouvement

Identifier dans chaque image quels pixels ont bougé

$$J(x, y, t) = |I(x, y, t) - I(x, y, t - 1)|$$

ou

$$J(x, y, t) = |I(x, y, t) - I_{ref}(x, y, t)|$$

Si $J(x, y, t) \neq 0$ alors le pixel a bougé

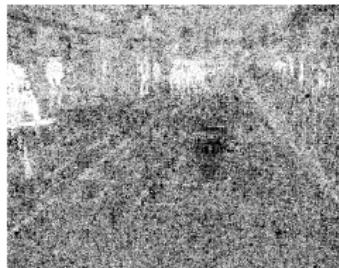
Sinon pas de mouvement.

En pratique, utilisation d'un seuil (Si $J(x, y, t) < \tau$).

Détection de mouvement

Identifier dans chaque image quels pixels ont bougé

$$J(x, y, t) = |I(x, y, t) - I_{ref}(x, y, t)|$$



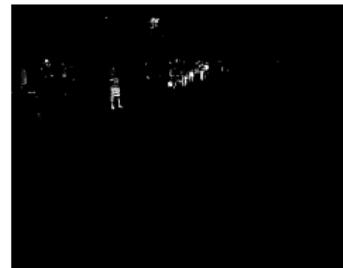
$$\tau = 0$$



$$\tau = 20$$



$$\tau = 50$$



$$\tau = 100$$

Détection de mouvement

Limitation : difficulté du choix du seuil, de la fréquence

$$D(n) = |I(x, y, t + n) - I(x, y, t)|$$



I(t)



D(-1)



D(-3)



D(-5)



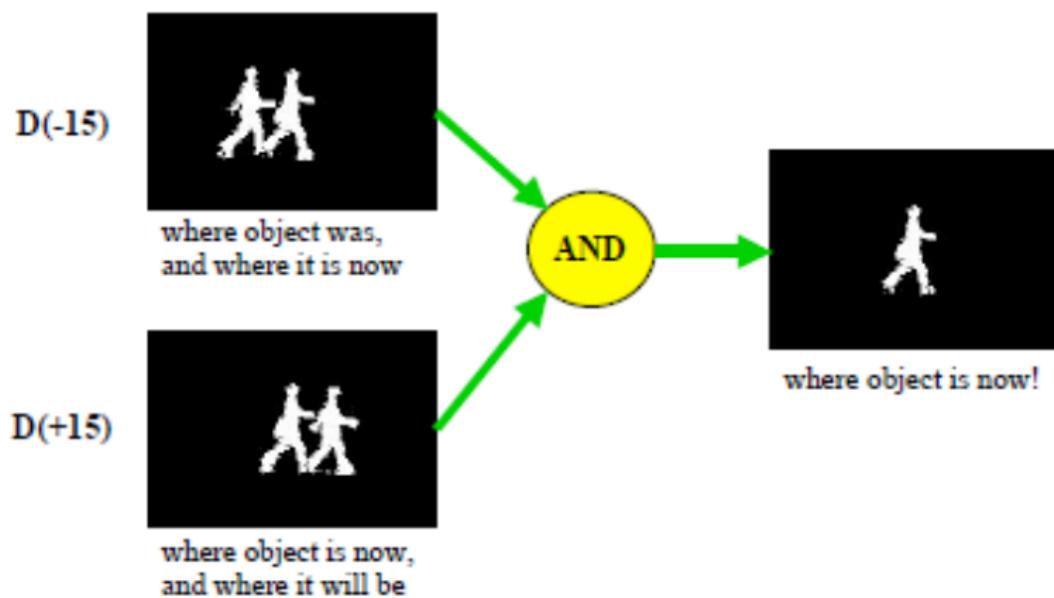
D(-9)



D(-15)

Détection de mouvement

Utilisation de trois images :



Détection de mouvement

Contraintes (sans seuil)

- ▶ Caméra fixe
- ▶ Illumination constante
- ▶ Environnement contrôlé
- ▶ Capteur parfait

Contraintes (avec seuil)

- ▶ Caméra bougeant lentement
- ▶ Faible variation d'éclairement
- ▶ Bruits de capteur

Environnement non contrôlé → nécessité d'estimer le mouvement ou d'apprendre le fond

Soustraction de fond

Différence d'images limitée à des séquences simples

Estimation statistique en chaque pixel

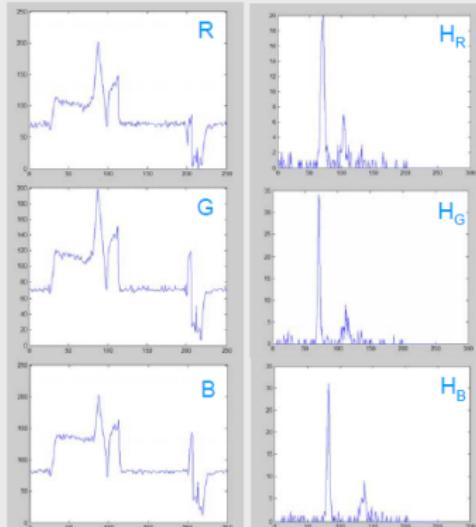
- ▶ Moyenne arithmétique des n dernières images (coûteux)
- ▶ Moyenne récursive ($M(x, y, t) = \alpha I(x, y, t) + (1 - \alpha)M(x, y, t - 1)$)
- ▶ Modèle gaussien ou GMM (variance et moyenne récursive)
C. Stauffer and E. Grimson, "Adaptive Background Mixture Models for Real-time Tracking," CVPR, 1999.

Soustraction de fond

- Example:



pixel
location



sequence of
pixel values

histograms

1

Détection des objets en mouvement

Extraction des objets à partir de la carte de mouvement

1. Extraction des composantes connexes, suppression des pixels parasites
2. Etiquetage des composantes connexes

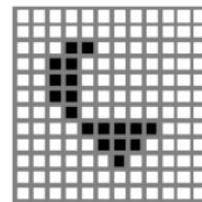
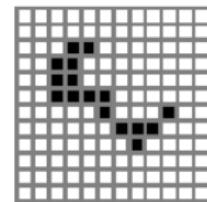
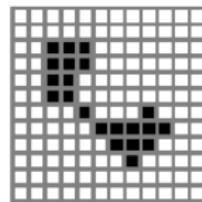
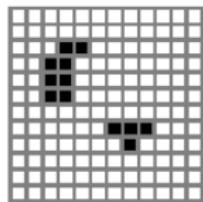
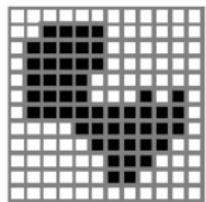
Extraction des objets à partir de la carte de mouvement

Extraction des composantes connexes

- ▶ Opérateurs morphologiques (érosion, dilatation)
- ▶ Erosion : supprime les pixels isolés
- ▶ Dilatation : renforce les groupes denses de pixels
- ▶ Ouverture morphologique : érosion puis dilatation
- ▶ Fermeture morphologique : dilatation puis érosion

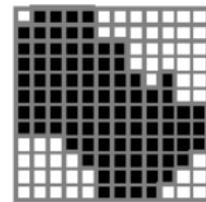
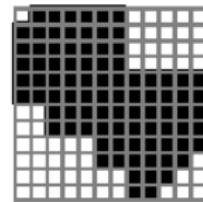
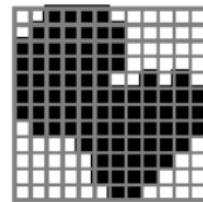
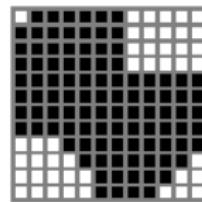
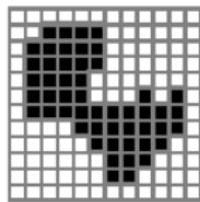
Opérateurs morphologiques

Erosion



Opérateurs morphologiques

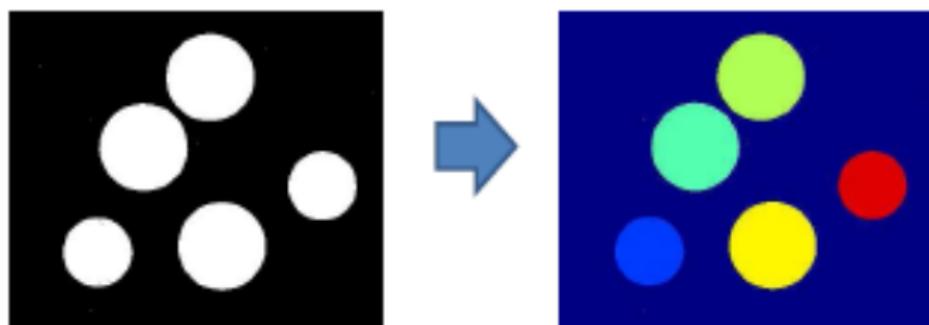
Dilatation



Détection des objets en mouvement

Extraction des objets à partir de la carte de mouvement

1. Extraction des composantes connexes, suppression des pixels parasites
2. Etiquetage des composantes connexes



Plan

Détection des objets en mouvement

Flot optique

Segmentation de mouvement

Detect-before-track

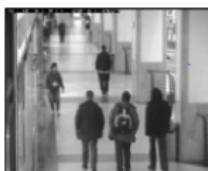
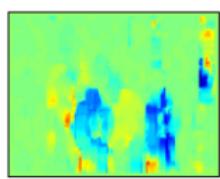
Suivi basé sur les distributions

Segmentation dynamique

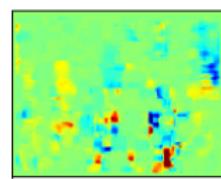
Flot optique

- ▶ Flot optique = Champ de déplacement
- ▶ Hypothèse de conservation de l'intensité :

$$I(x, y, t) = I(x + \delta_x, y + \delta_y, t - 1)$$

 I_1  I_2  v_x

déplacement horizontal

 v_y

déplacement vertical



(v_x, v_y) superposé à I_1
Champ de déplacement

Flot optique

- ▶ Block matching
- ▶ Equation du flot optique (ECMA : équation de contrainte du mouvement apparent)

$$\begin{aligned} I(x, y, t) &= I(x + \delta_x, y + \delta_y, t + \delta t) \\ &\simeq I(x, y, t) + \frac{\delta I}{\delta x} \cdot \delta x + \frac{\delta I}{\delta y} \cdot \delta y + \frac{\delta I}{\delta t} \cdot \delta t \quad (\text{Taylor}) \end{aligned}$$

$$\frac{\delta I}{\delta x} \cdot \frac{\delta x}{\delta t} + \frac{\delta I}{\delta y} \cdot \frac{\delta y}{\delta t} + \frac{\delta I}{\delta t} = 0$$

$$\nabla I \cdot \mathbf{v} + \frac{\delta I}{\delta t} = 0$$

Flot optique

$$\nabla I \cdot v + \frac{\delta I}{\delta t} = 0$$

Lucas & Kanade

- ▶ Champ de mouvement constant sur des petites régions

$$v = \arg \min_v \sum_{p \in \Omega} [\nabla I(m, t) \cdot v + I_t(m, t)]^2$$

- ▶ Résolution par moindres carrés

Plan

Détection des objets en mouvement

Flot optique

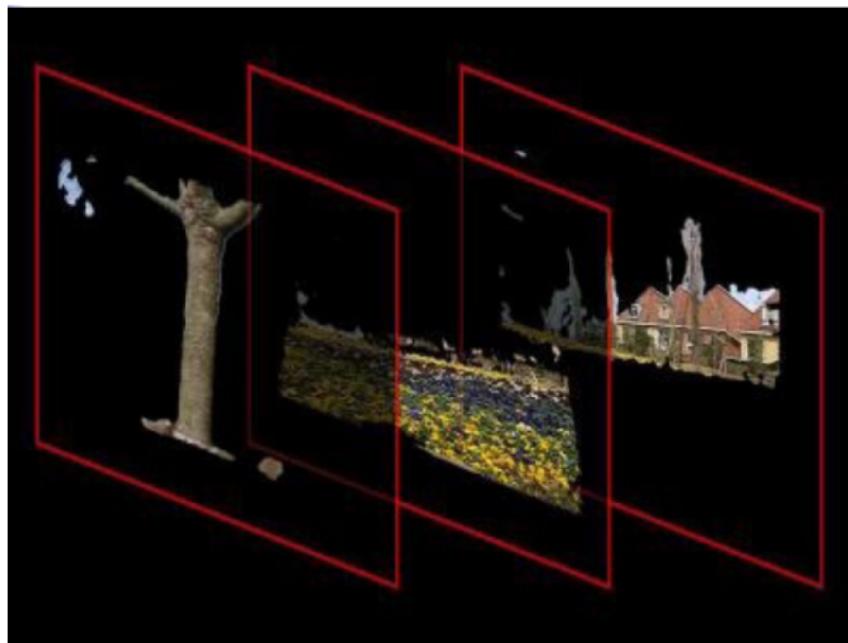
Segmentation de mouvement

Detect-before-track

Suivi basé sur les distributions

Segmentation dynamique

Segmentation de mouvement

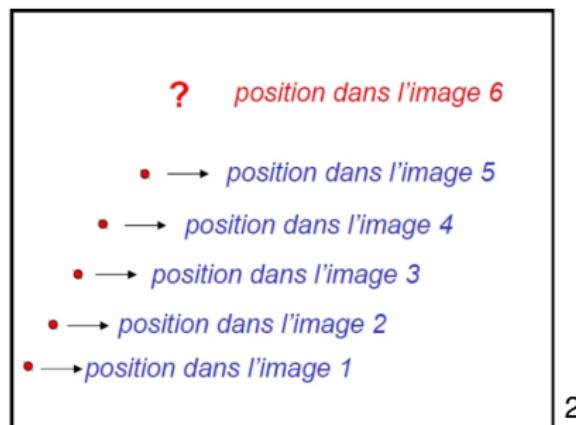


Suivi d'objets dans les vidéos

Suivi temporel, de points, d'objets segmentés ou de boîtes englobants des objets.

Suivi temporel

À partir des positions de l'objet dans les images précédentes, trouver la position de l'objet dans la nouvelle image



Suivi d'objets dans les vidéos

Suivi temporel, de points, d'objets segmentés ou de boîtes englobants des objets.

Difficultés

- ▶ Zoom
- ▶ Occultations
- ▶ Changement d'apparence
- ▶ Mouvement de la caméra
- ▶ Objets non rigides

Introduction

3 catégories :

1. “Detect-before-track”
2. Suivi basé sur les distributions
3. Segmentation dynamique

Plan

Détection des objets en mouvement

Flot optique

Segmentation de mouvement

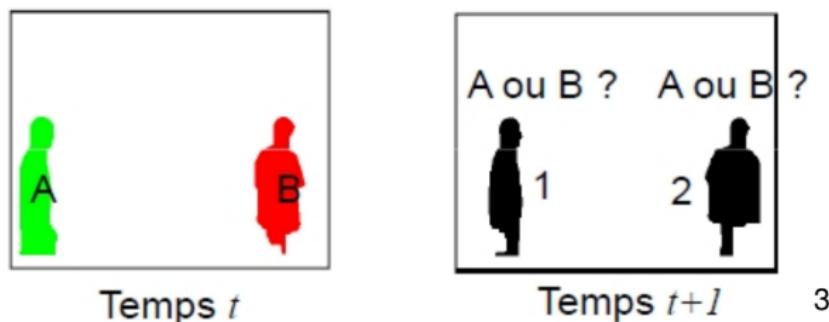
Detect-before-track

Suivi basé sur les distributions

Segmentation dynamique

Detect-before-track

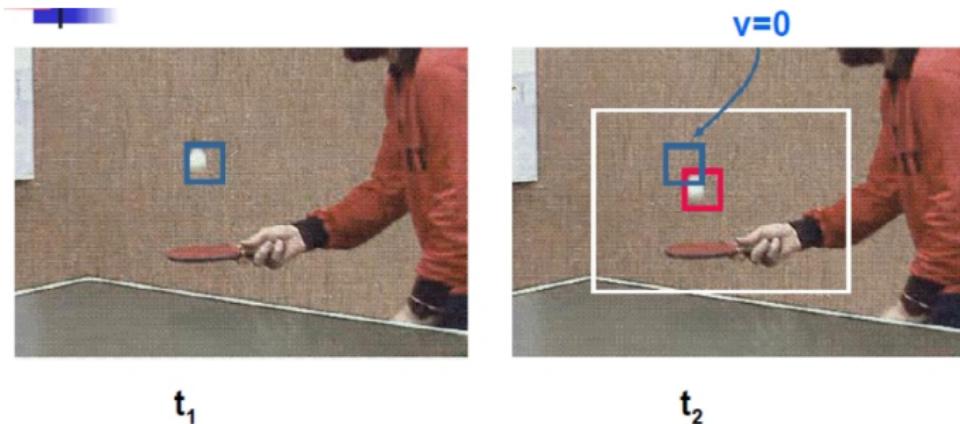
Mise en correspondance de l'objet suivi à $t - 1$ avec les objets détectés à l'instant t



Critères d'appariement

- ▶ Visuels : taille, couleur, forme ...
- ▶ Spatiaux : distance entre centres ...

Block Matching

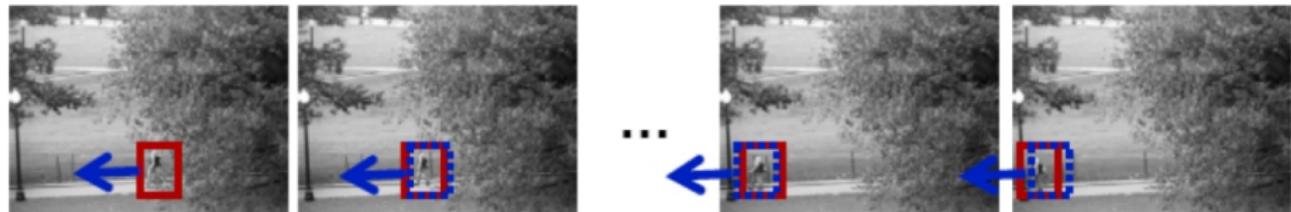


Choisir le bloc le plus représentatif parmi tous les blocs possibles à l'intérieur d'une fenêtre

4

Utilisation du mouvement

Mouvement permet de prédire la nouvelle position de l'objet



5

Prédiction

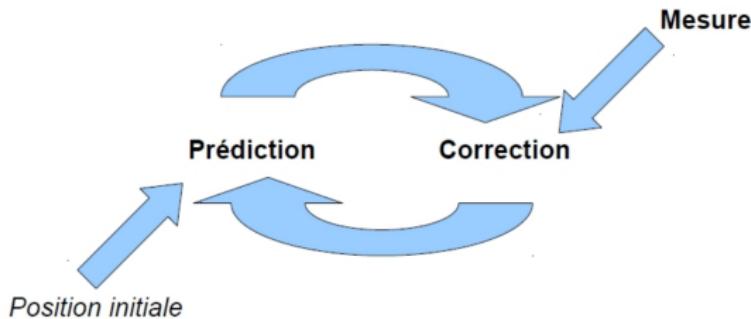
- ▶ Flot optique
- ▶ Vitesse constante, accélération constante ...

Appariement

- ▶ Recherche dans un voisinage autour de la prédiction
- ▶ Méthodes de filtrage

5. Source : Elise Arnaud

Méthodes de filtrage



6

2 étapes

1. Prédiction
2. Correction à l'aide d'une mesure

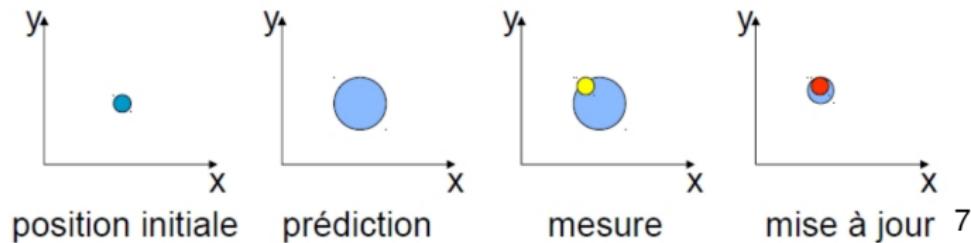
6. Source : Alain Boucher - IFI

Méthodes de filtrage

2 étapes

1. Prédiction
2. Correction à l'aide d'un mesure

Erreurs de prédiction et de mesures



Avantages/Inconvénients

- + Peut gérer l'entrée et la sortie d'objets dans la scène
- + Suivi robuste si les détections sont de bonne qualité
- + Utilisation du mouvement : gestion des occultations
- Dépend beaucoup des détections

Plan

Détection des objets en mouvement

Flot optique

Segmentation de mouvement

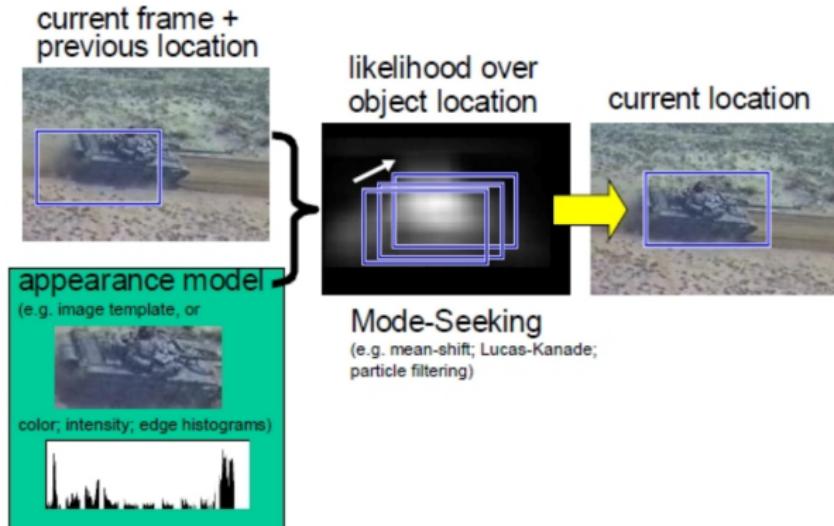
Detect-before-track

Suivi basé sur les distributions

Segmentation dynamique

Suivi basé sur les distributions

Meilleure position correspond à la position pour laquelle la distribution des pixels est la plus proche de celle de référence



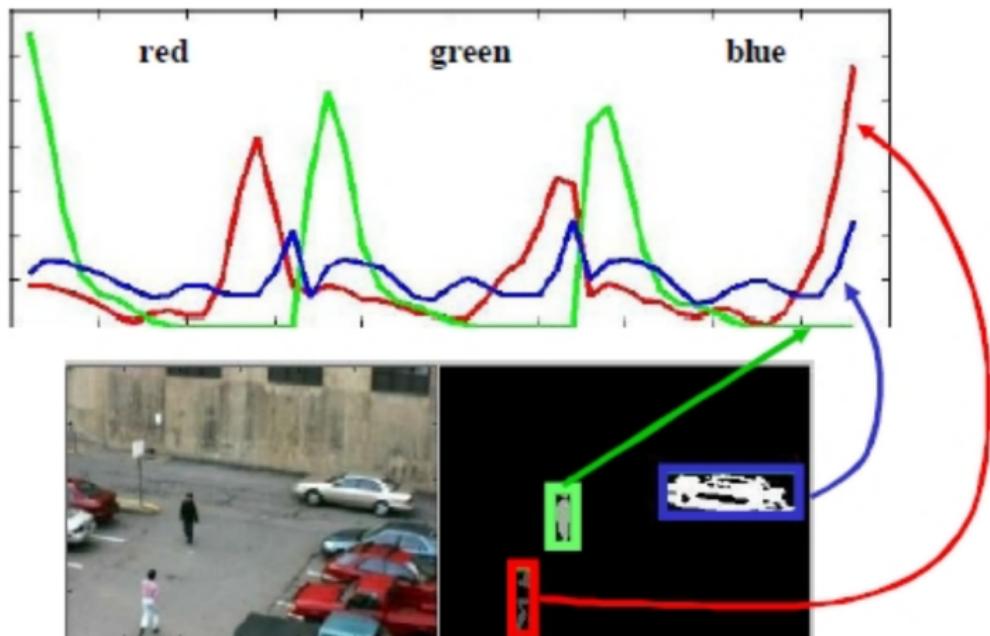
8

(poids élevé pour les pixels correspondant au modèle de couleur)

8. Robert Collins

Suivi basé sur les distributions

En général distribution de couleur



9

Avantages/Inconvénients

- + Suivi rapide et robuste dans des vidéos couleurs
- Ne gère pas l'entrée et la sortie de nouveaux objets
- Sensible aux changements d'illumination
- Boîtes englobantes seulement

Plan

Détection des objets en mouvement

Flot optique

Segmentation de mouvement

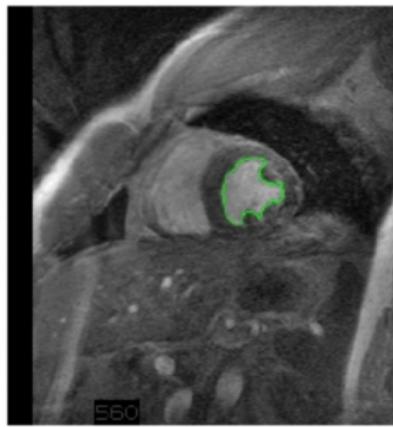
Detect-before-track

Suivi basé sur les distributions

Segmentation dynamique

Segmentation dynamique

Faire évoluer le contour obtenu à l'instant précédent vers les contours de l'image courante.



10

Utilisation des contours actifs

10. Kristen Grauman