

Utilisation des Points d'Intérêts Couleurs pour le Suivi d'Objets

Pierre Gabriel, Jean-Bernard Hayet, Justus Piater et Jacques Verly
Université de Liège, Département d'Électricité, Électronique et Informatique
Campus du Sart Tilman, Bât. B28, 4000 Liège, Belgique
{P.Gabriel,Jean-Bernard.Hayet,Justus.Piater,Jacques.Verly}@ulg.ac.be

Résumé

Ce papier présente une nouvelle approche pour le suivi d'objets dans des situations complexes comme celles des gens dans une foule ou des joueurs sur un terrain de football. Chaque objet présent dans l'image est représenté par plusieurs points d'intérêts (PI). Ces PI sont obtenus en utilisant la version couleur du détecteur de Harris. Chaque PI est caractérisé par l'apparence locale (local jet couleur du premier ordre) de l'objet autour de ce point et par sa position relative dans l'objet. Nous suivons les objets en faisant des correspondances entre des PI d'une image à l'autre, basées sur la distance de Mahalanobis. Nos résultats montrent que l'approche est robuste aux occultations et que le système de suivi est globalement performant.

Mots Clef

Suivi d'objets, points d'intérêts, apparence locale, local jet couleur, modèle géométrique, occultation.

Abstract

This paper presents a new approach for tracking objects in complex situations such as people in a crowd or players on a soccer field. Each object in the image is represented by several interest points (IP). These IPs are obtained using the color version of the Harris IP detector. Each IP is characterized by the local appearance (chromatic first-order local jet) of the object around the point and by geometric parameters. We track objects by matching IPs from image to image, based on the Mahalanobis distance. The approach is robust to occlusion. Performance is illustrated with some examples.

Keywords

Object tracking, interest points, local appearance, color local jet, geometric model, occlusion.

1 Introduction

Il y a eu une activité de recherche considérable dans le domaine du suivi d'objets dans les séquences vidéos depuis une vingtaine d'années. Cet intérêt a été motivé par les nombreuses applications comme la surveillance, la vidéo conférence, l'interaction homme-machine et les émissions télévisées.

Nous nous intéressons au suivi de plusieurs objets dans une séquence vidéo couleur contenant des scènes complexes où les objets peuvent avoir des tailles différentes. Ces objets peuvent être rigides (par exemple une voiture) ou non-rigides (par exemple une personne) et s'occulter les uns les autres. La plupart des systèmes de suivi actuels ne marchent pas bien avec ces conditions difficiles. En fait, la majorité d'entre eux est basée soit sur l'extraction de régions en utilisant un modèle du fond [17, 6, 21], soit sur l'extraction de contours avec un filtre à particules [8, 13].

Dans notre approche, un objet est caractérisé par un ensemble de points d'intérêts (PI) obtenus avec la version couleur du détecteur de Harris [16]. Chaque PI est caractérisé par son apparence locale (un vecteur de caractéristiques locales). L'utilisation d'un ensemble de PI nous permet de suivre un objet à travers une occultation partielle tant qu'un ou plusieurs points restent visibles. Pour augmenter la robustesse, nous exploitons aussi des relations géométriques existant entre les PI.

La combinaison de PI avec des descripteurs locaux a été utilisée avec succès pour l'appariement de points en reconstruction stéréo [18], indexation d'images [14] et reconstruction d'objets [19]. Cependant, à notre connaissance, ceci est la première tentative pour le suivi d'objets en utilisant des PI couleurs caractérisés par un descripteur local et un modèle géométrique. Récemment, Gouet et Lameyre [5] ont présenté un tracker qui utilise des PI et des *snakes* mais avec des images en niveau de gris, sans aucun modèle géométrique et pour un objet unique dans la scène.

Ce papier est organisé comme suit. Dans la section 2, nous décrivons la version couleur du détecteur de PI de Harris. Dans la section 3, nous présentons notre nouveau système de suivi basique qui combine un modèle d'apparence et un modèle géométrique pour chaque objet. Dans

la section 4, nous généralisons ce système basique pour le suivi d'objets au travers d'occultations partielles. Finalement, dans la section 5, nous concluons et proposons quelques orientations pour des recherches futures.

2 Le détecteur de Harris couleur

Nous pouvons définir un PI comme étant un point dans l'image où des changements significatifs se produisent. Des exemples de PI sont les coins, les jonctions, les points noirs sur fond blanc ou tout autre point marqué par un changement important de la texture comme illustré sur la figure 1 pour une image en niveau de gris.



FIG. 1 – Les points rouges indiquent les PI détectés par le détecteur de Harris traditionnel.

Depuis une vingtaine d'années, plusieurs détecteurs de PI ont été développés. Schmid et Mohr [20] ont comparé les performances de plusieurs d'entre eux.

Le détecteur de PI le plus populaire est le détecteur de Harris [7]. Ce détecteur étant seulement applicable aux images en niveau de gris, Montesinos *et al.* [16] l'ont généralisé pour les images couleur. Les PI produits par ce détecteur sont définis comme étant les extrema locaux positifs de l'image intermédiaire en niveau de gris

$$R(x, y) = \det(M(x, y)) - k \cdot \text{trace}(M(x, y))^2, \quad (1)$$

où k vaut typiquement 0.04 et $M(x, y)$ est la matrice 2×2

$$M(x, y) = \begin{pmatrix} M_{11}(x, y) & M_{12}(x, y) \\ M_{12}(x, y) & M_{22}(x, y) \end{pmatrix}$$

construite en chaque point de l'image en fonction des trois images intermédiaires en niveau de gris $M_{11}(x, y)$, $M_{12}(x, y)$ et $M_{22}(x, y)$ définies par

$$\begin{aligned} M_{11}(x, y) &= G_{\sigma_i} \otimes (r_x^2 + g_x^2 + b_x^2) \\ M_{12}(x, y) &= G_{\sigma_i} \otimes (r_x r_y + g_x g_y + b_x b_y) \\ M_{22}(x, y) &= G_{\sigma_i} \otimes (r_y^2 + g_y^2 + b_y^2), \end{aligned}$$

où G_{σ_i} est une gaussienne 2D isotropique de variance σ_i^2 , \otimes dénote l'opération de convolution 2D et c_x et c_y représentent les dérivées gaussiennes du premier ordre des canaux c de l'image originale $I(x, y)$ avec $c \in \{r, g, b\}$. Ces dérivées sont implémentées en utilisant une gaussienne 1D de variance σ_d^2 .

Selon les comparaisons faites par Gouet et Boujemaa [4], le détecteur ci-dessus apparaît comme étant le plus sûr parmi les détecteurs de PI couleur connus, et ce par rapport aux changements d'illumination, au bruit, aux rotations et aux changements de point de vue. La figure 2 donne quelques exemples de PI détectés dans différentes conditions.

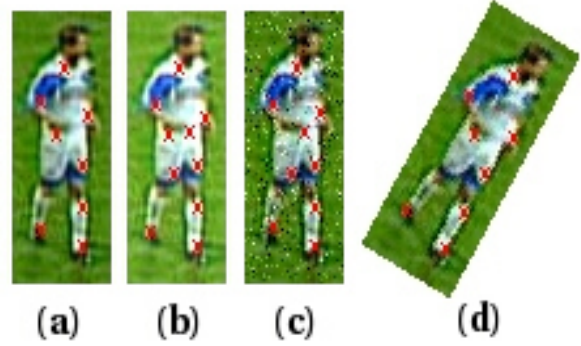


FIG. 2 – Sortie du détecteur de Harris couleur, avec $\sigma_d = 2$ pixels, dans quatre conditions différentes : (a) image originale, (b) changement d'illumination, (c) bruit et (d) rotation de l'image. Les PI détectés dans les quatre cas sont pratiquement les mêmes.

3 Le système de suivi

Le but du système proposé est de suivre des objets d'une image à l'autre dans une séquence vidéo couleur. Dans la configuration actuelle du système, pour la première image ($n = 0$) de la séquence, nous supposons que les objets à suivre sont bien séparés et que nous définissons manuellement une région d'intérêt (ROI) autour de ceux-ci. Ensuite, le système fonctionne de façon autonome en commençant par la seconde image ($n = 1$). A chaque image n , nous appliquons le détecteur de Harris couleur à chaque ROI. Pour l'image $n = 0$, nous ignorons les PI détectés en dehors de l'objet. Actuellement, chaque ROI garde sa taille initiale d'une image à l'autre. Pour suivre les objets, nous recherchons les correspondances entre les PI de l'image actuelle et ceux de l'image précédente, grâce à la combinaison d'un modèle d'apparence et d'un modèle géométrique.

3.1 Le modèle d'apparence

Le modèle d'apparence d'un objet consiste en un ensemble de PI qui se trouvent dans la ROI correspondante. Chaque PI est caractérisé par un vecteur contenant certains attributs locaux.

Les descripteurs les plus utilisés sont :

- **Le local jet** qui est un vecteur de coefficients caractérisant la surface locale en niveau de gris par un développement de Taylor [12]. Il est défini en un point (x, y) jusqu'à l'ordre n par

$$J^n(x, y, \sigma) = \left\{ \frac{\partial^{p+q}}{\partial x^p \partial y^q} I(x, y) \right\},$$

où $p = 0, \dots, n$, $q = 0, \dots, n$ et σ est l'écart type de la gaussienne utilisée pour le calcul des dérivées.

- **Les invariants différentiels** qui combinent les éléments du local jet pour avoir l'invariance en rotation de l'image [19].
- **Le SIFT** (*Scale Invariant Feature Transform*) qui représente le voisinage comme un vecteur de caractéristiques invariant à la rotation et à la mise à l'échelle et qui est robuste aux petits changements de translation, d'illumination et de transformations affines [14].

Vu que nous travaillons sur des séquences vidéos, nous n'avons pas de changements significatifs d'une image à l'autre. Donc, nous n'avons pas besoin des invariants différentiels ou du SIFT. Le local jet semble alors particulièrement adéquat à notre application. En effet, Gouet et Lameyre [5] ont montré que dans le cas de séquences vidéos l'appariement de points avec le local jet donnait de meilleurs résultats qu'avec les invariants différentiels. Comme nous travaillons avec des images couleurs, nous utilisons le local jet pour chaque canal. De plus les dérivées d'ordre supérieur étant sensibles au bruit de l'image, nous utilisons uniquement le local jet du premier ordre. Il se compose donc des trois canaux r , g et b plus les dérivées gaussiennes r_x , g_x , b_x , r_y , g_y et b_y .

En plus du local jet couleur, nous ajoutons la *cornerness* R locale donnée par l'équation 1. Nos tests ont indiqué que l'ajout de R aide le suivi des PI d'une image à l'autre. Donc nous caractérisons chaque PI avec 10 attributs locaux.

3.2 Le modèle géométrique

L'inconvénient d'utiliser le modèle d'apparence seul est qu'un PI peut être associé à un autre qui lui ressemble même si ces points sont loin l'un de l'autre. Par exemple, pour un joueur de football, les PI qui se trouvent à la jonction peau-chaussette sur les jambes gauche et droite ont une apparence très similaire et peuvent donc être confondus.

C'est pourquoi, pour suivre un objet de façon robuste, nous ajoutons un modèle géométrique en plus de son modèle d'apparence. Ce modèle géométrique décrit les coordon-

nées (x_c, y_c) de chaque PI par rapport au centre de l'objet (voir Fig. 3), tel que déterminé par une méthode expliquée plus bas.

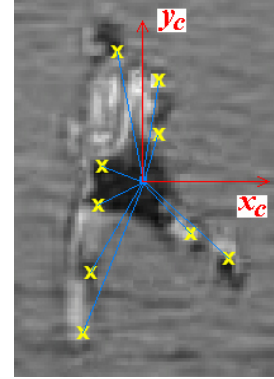


FIG. 3 – Un objet est caractérisé par un ensemble de PI. Chaque PI est, quant à lui, caractérisé par un vecteur d'attributs de son apparence locale et par sa position relative au centre de l'objet.

3.3 Combinaison des modèles d'apparence et géométrique

Dans la suite, nous notons les PI dans une ROI donnée dans l'image n par $j \in \{1, \dots, P(n)\}$ où $P(n)$ est le nombre de PI. L'élément de base dans notre approche pour le suivi est le vecteur (combinant les caractéristiques d'apparence et géométriques) défini pour chaque j dans l'image n par

$$V_j(n) = (r, g, b, r_x, g_x, b_x, r_y, g_y, b_y, R, x_c, y_c), \quad (2)$$

où tous les composants ont été introduits plus haut. Un objet O dans l'image n est alors modélisé par

$$V^O(n) = \{V_j(n) \mid j = 1, \dots, P(n)\}. \quad (3)$$

3.4 L'algorithme de suivi

Considérons une ROI particulière (entourant un objet d'intérêt) qui est suivie d'une image à l'autre, disons de $n - 1$ à n . Le problème de base est de trouver, pour chaque PI $i \in \{1, \dots, P(n - 1)\}$, le meilleur PI j correspondant avec $j \in \{1, \dots, P(n)\}$. Pour trouver j , nous comparons tous les $V_j(n)$ à chaque $V_i(n - 1)$ comme illustré à la figure 4.

Un choix naturel pour l'appariement est la distance de Mahalanobis entre $V_i(n - 1)$ et $V_j(n)$,

$$d_M(i, j) = [(V_{ij} - \mu)^T C^{-1} (V_{ij} - \mu)]^{1/2}, \quad (4)$$

où $V_{ij} = V_i(n-1) - V_j(n)$ et (μ, C) sont la moyenne et la matrice de covariance de V_{ij} mesurés pour les PI correctement appariés sur une séquence d'entraînement.

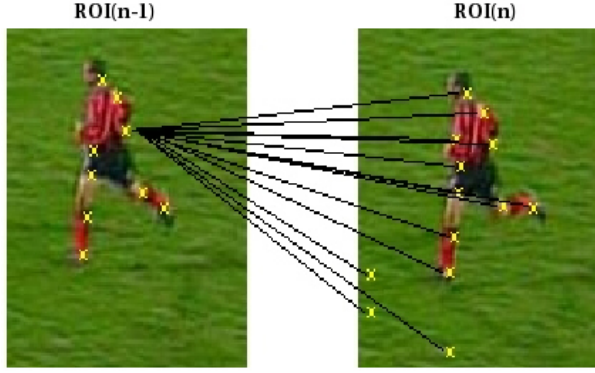


FIG. 4 – Pour suivre les PI (appartenant à un objet donné) de l'image $n-1$ à l'image n , nous considérons chaque PI à l'image $n-1$ et nous essayons de lui trouver un correspondant parmi tous les PI à l'image n .

Un PI j est déclaré comme étant le correspondant de i si j est le k qui correspond au minimum de $d_M(i, k)$, pour $k \in \{1, \dots, j, \dots, P(n)\}$, et si ce minimum est inférieur à un seuil θ . Évidemment, certains PI i et j risquent de ne pas avoir de correspondant.

Nous supposons que l'image de la scène ne change pas beaucoup de l'image $n-1$ à l'image n , même si la caméra est mobile. Dans notre système actuel, le centre d'une ROI à l'image n est égale au centre de l'objet correspondant à l'image $n-1$. Dans les versions futures, nous considérerons des techniques plus sophistiquées, par exemple basées sur le filtre de Kalman.

Pour chaque ROI de l'image n , en plus des PI qui ont été appariés avec succès, nous incluons dans $V^O(n)$ tout autre PI qui se trouve à l'intérieur du rectangle entourant ces PI correctement appariés. Cette étape permet de tenir compte des nouveaux PI qui apparaissent sur l'objet.

L'algorithme 1 décrit notre approche pour le suivi d'un objet donné d'une image à l'autre.

3.5 L'algorithme de centrage

Pour calculer les coordonnées (x_c, y_c) utilisées dans $V_j(n)$ pour l'équation 2, nous avons besoin de la position du centre de l'objet (CO) dans la ROI de l'image n . Quand nous appliquons le détecteur de Harris couleur à une ROI donnée, nous trouvons fort probablement des PI qui appartiennent au fond comme c'est le cas pour la ROI(n) à la figure 4. Dans ce cas, le centre de gravité des PI ne correspond pas au CO.

Algorithm 1 SUIVI (pour chaque objet)

- 1: Initialisation : (a) Définir manuellement la ROI de l'objet O dans l'image $n = 0$. (b) Appliquer le détecteur de Harris couleur à la ROI, sélectionner manuellement les PI appartenant à O et donc rejeter les PI appartenant au fond. (c) Calculer le centre de O comme étant égal au centre de gravité de ses PI et déterminer $V^O(0)$.
 - 2: **for** image $n = 1$ à N **do**
 - 3: Mettre $V^O(n) = \emptyset$.
 - 4: Mettre le centre de la ROI(n) à la valeur du centre de l'objet à l'image $n-1$.
 - 5: Appliquer le détecteur de Harris couleur à la ROI(n) et remplir les vecteurs $V_j(n)$ avec les caractéristiques d'apparence.
 - 6: Calculer la nouvelle position du centre de l'objet en utilisant l'algorithme 2 ci-dessous.
 - 7: Ajouter les caractéristiques géométriques aux vecteurs $V_j(n)$.
 - 8: **for** $i = 1$ à $P(n-1)$ **do**
 - 9: **for** $j = 1$ à $P(n)$ **do**
 - 10: Calculer $d_M(i, j)$ en utilisant l'équation 4.
 - 11: **end for**
 - 12: **if** $\min(d_M(i, j)) < \theta$ **then**
 - 13: Apparier j à i et ajouter $V_j(n)$ à $V^O(n)$.
 - 14: **end if**
 - 15: **end for**
 - 16: Inclure dans $V^O(n)$ tout autre PI qui se trouve à l'intérieur du rectangle entourant les j appariés correctement.
 - 17: **end for**
-

Si nous supposons que les PI ne bougent pas beaucoup par rapport au CO de l'image $n-1$ à l'image n , nous pouvons déterminer le CO présumé dans l'image n selon le processus suivant. Nous considérons les B (typiquement 5) meilleurs appariements $j_b \rightarrow i_b$ avec $b = 1, \dots, B$. Ces derniers étant calculés sur base des 10 premiers éléments des vecteurs $V_i(n-1)$ et $V_j(n)$. Pour chaque appariement b , nous déterminons une estimation du CO comme suit. Nous calculons le déplacement $(\Delta x_b, \Delta y_b)$ du point j_b par rapport au point i_b . L'estimation du CO correspondant à l'appariement b est alors obtenu en ajoutant ce déplacement au CO dans l'image $n-1$,

$$x_{CO}(n) = x_{CO}(n-1) + \Delta x_b \quad (5)$$

$$y_{CO}(n) = y_{CO}(n-1) + \Delta y_b$$

Nous calculons alors une estimation $(\bar{x}_{CO}, \bar{y}_{CO})$ plus précise, mais toujours temporaire, du CO pour l'image n en moyennant les $(x_{CO,b}, y_{CO,b})$ avec $b = 1, \dots, B$. Ceci est illustré par le point \bar{C} sur la figure 5. Ensuite nous éliminons les estimations grossières $(x_{CO,b}, y_{CO,b})$ qui sont loin de $(\bar{x}_{CO}, \bar{y}_{CO})$. Par exemple, dans le cas de la figure 5, le point

C_3 est éliminé. Finalement, nous recalculons la moyenne avec les points restant comme illustré par le point C à la figure 5. L'algorithme 2 résume la procédure complète pour calculer le nouveau centre de l'objet de façon robuste.

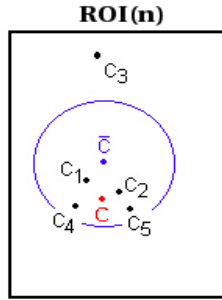


FIG. 5 – Illustration du processus d'estimation du centre de l'objet (CO) à l'image n en utilisant l'algorithme 2. C_3 est un exemple d'estimation intermédiaire à supprimer à l'étape 6 de l'algorithme 2. Le CO final est C . C sera utilisé comme centre de la ROI à l'image $n + 1$.

Algorithm 2 CENTRAGE (pour chaque objet)

- 1: Déterminer les B meilleurs appariements $j_b \rightarrow i_b$ en utilisant uniquement les caractéristiques d'apparence.
 - 2: Calculer les centres C_b correspondant à ces appariements en utilisant l'équation 5 avec $b = 1, \dots, B$.
 - 3: Calculer la valeur moyenne \bar{C} de ces centres.
 - 4: **for** $b = 1$ à B **do**
 - 5: **if** C_b est loin de \bar{C} **then**
 - 6: Supprimer l'appariement correspondant $j_b \rightarrow i_b$.
 - 7: **end if**
 - 8: **end for**
 - 9: Calculer la valeur moyenne C des centres restant.
-

Ce processus est très utile quand un objet n'est pas au milieu de la ROI correspondante ou lorsqu'il est occulté. Les figures 6 et 7 illustrent le fonctionnement du système proposé. Les objets sont correctement suivis même quand une ROI n'est pas centrée sur l'objet (comme dans l'image 16 de la figure 6) ou quand l'objet est occulté (comme dans les images 28, 36 et 46 de la figure 7).

4 Occultation mutuelle

Dans la section 3, nous nous sommes concentrés sur le suivi d'objets isolés. Suivre des objets dans une foule conduit nécessairement au problème de l'occultation mutuelle des objets cible.

Nous avons déjà étudié ce problème et proposé un cadre théorique simple qui décrit l'occultation en termes génériques [3]. Ce cadre nous a permis de classifier et de com-

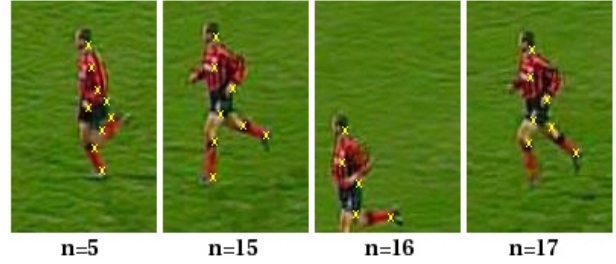


FIG. 6 – Exemple montrant la robustesse de l'algorithme 1 quand la ROI n'est pas centrée sur l'objet à suivre. (La ROI est l'image complète.)

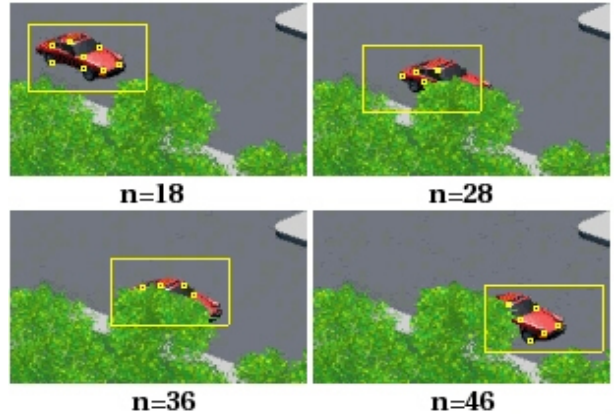


FIG. 7 – Exemple montrant la robustesse de l'algorithme 1 quand l'objet à suivre est partiellement occulté. (La ROI est définie par le rectangle.)

parer 23 systèmes de suivis qui traitent du problème d'occultation. Il existe essentiellement trois types de méthodes :

Méthodes *merge-split* : Il s'agit de la solution standard où les objets sont fusionnés (*merge*) dès qu'ils se touchent, en un groupe formant une seule entité. Cette entité est ensuite divisée (*split*) en plusieurs objets (correspondant aux objets initiaux) dès que ceux-ci se séparent [6, 15, 17].

Méthodes *straight-through* : Les objets sont suivis individuellement sans aucune fusion. Ces techniques sont souvent basées sur le calcul de la profondeur relative de l'objet [2, 11].

Méthodes utilisant plusieurs caméras : Le champ de vision des caméras classiques étant fort réduit, un intérêt croissant s'est développé pour l'utilisation de plusieurs caméras pour augmenter la scène d'intérêt. Cependant, l'utilisation de caméras multiples peut aussi aider dans le suivi d'objets qui sont occultés dans une ou plusieurs vues. Certains systèmes utilisent des caméras statiques [10, 9] alors que d'autres utilisent des caméras pilotées [1, 22].

Nous avons décidé dans notre travail d'utiliser une méthode *straight-through* pour traiter le problème de l'occlusion. En d'autres termes, nous avons choisi de continuer à différencier les objets quand ils sont partiellement occultés.

4.1 Les points disputés

Pour traiter les occultations, il est utile de définir un certain prédicat qui indique si une occultation est présente ou non. Ici, nous disons qu'une occultation a lieu entre deux ou plusieurs objets dès que leurs ROI s'intersectent.

Lorsque nous détectons une occultation, nous appliquons d'abord l'algorithme 1 à chaque ROI. Ensuite, nous calculons le rectangle entourant les PI détectés dans chaque ROI. Puis, nous collectons les PI qui sont à l'intersection de ces rectangles comme illustré à la figure 8 dans le cas de deux ROI. Nous définissons ces points comme étant les *points disputés*. Finalement, nous assignons chaque point disputé à un des objets impliqués dans l'occlusion, ceci en se basant sur le vecteur de caractéristiques complet.

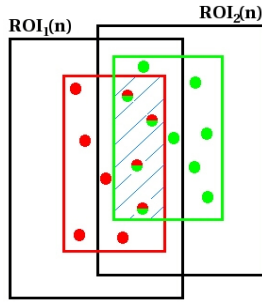


FIG. 8 – Le petit rectangle dans $ROI_1(n)$ (respectivement dans $ROI_2(n)$) entoure les PI détectés pour $ROI_1(n)$ (respectivement $ROI_2(n)$) en utilisant l'algorithme 1. Les points à l'intersection de ces rectangles sont définis comme étant les points disputés.

Un point disputé k est attribué à un objet O s'il a le plus petit minimum pour la distance $d_M^O(i, k)$ où $i = 1, \dots, P(n-1)$ et $P(n-1)$ est le nombre de PI qui décrivent l'objet O dans l'image $n-1$. Notons que nous utilisons tous les PI correspondant à O dans l'image précédente.

4.2 Retrouver un objet totalement occulté

Soit $\eta^O(n)$ le nombre d'appariements $j \rightarrow i$ trouvés pour l'objet O dans l'image n . Nous disons que O est totalement occulté si $\eta^O(n)$ est plus petit que 3. Pour retrouver O dans l'image suivante, nous utilisons un modèle d'apparence de référence pour O .

Quand un objet est totalement occulté, nous continuons à positionner sa ROI comme expliqué précédemment. En-

suite, nous appliquons le détecteur de Harris couleur à la ROI mais nous caractérisons chaque PI avec uniquement les 10 premiers éléments de l'équation 2. Puis, nous comparons cette description de l'objet avec un modèle de référence $V_r^O(n)$.

Le modèle de référence $V_r^O(n)$ correspond à l'apparence moyenne de l'objet O dans l'image n . Il est mis à jour d'une image à l'autre tant que l'objet est correctement détecté (c'est-à-dire avec $\eta^O(n) > 3$). La mise à jour utilise une moyenne pondérée de tous les vecteurs des PI, symboliquement écrite comme

$$V_r^O(n) = \alpha V_r^O(n-1) + (1-\alpha) \text{moyenne}(V^O(n)), \quad (6)$$

où $\alpha \in [0, 1]$ et les $V_j(n)$ de $V^O(n)$ contiennent uniquement les caractéristiques d'apparence.

L'algorithme 3 décrit notre approche pour le suivi à travers les occultations. Les figures 9 et 10 montrent le résultat l'algorithme 3 appliqué à deux situations impliquant une occultation. Dans les deux cas, les deux personnes près du centre de l'image sont correctement suivis à travers l'occlusion.

Algorithm 3 SUIVI A TRAVERS LES OCCULTATIONS

```

1: for chaque objet  $O$  do
2:   if  $\eta^O(n-1) < 3$  then
3:     Marquer  $O$  comme étant totalement occulté.
4:     Appliquer le détecteur de Harris couleur à la ROI.
5:     Rechercher les PI qui ressemblent à  $V_r^O(n-1)$ .
6:   end if
7: end for

8: Appliquer l'algorithme 1 à chaque ROI.

9: if deux ou plusieurs ROIs s'intersectent then
10:  Déterminer les  $K$  points disputés.
11:  for  $k = 1$  à  $K$  do
12:    for chaque objet  $O$  dans l'occlusion do
13:      Rechercher le  $\min_i(d_M^O(i, k))$ ,
        où  $i = 1, \dots, \eta^O(n-1)$ .
14:    end for
15:    Assigner  $k$  à l'objet  $O$  qui a le plus petit minimum.
16:  end for
17: end if

18: for chaque objet  $O$  do
19:  Calculer  $\eta^O(n)$ .
20:  if  $\eta^O(n) > 3$  then
21:    Mettre à jour  $V_r^O(n)$ .
22:  end if
23: end for

```

5 Conclusions et perspectives

Nous avons proposé une nouvelle approche pour le suivi d'objet utilisant des points d'intérêts (PI) couleurs. Dans cette approche, un objet est défini par un ensemble de PI détectés avec la version couleur du détecteur de Harris. Chaque point est d'abord caractérisé par son apparence locale c'est-à-dire le local jet couleur d'ordre 1 plus la *cornerness*. De plus, chaque point est caractérisé par sa position relative au centre estimé de l'objet. L'utilisation d'un ensemble de PI nous permet de suivre un objet à travers une occultation partielle aussi longtemps que quelques points restent visibles. Le système implémenté a été testé avec succès sur différentes scènes montrant ainsi sa généralité.

Dans la version actuelle du système, nous supposons que les objets à suivre sont bien séparés dans la première image de la séquence, nous plaçons manuellement les ROIs qui les entourent et nous supprimons manuellement les PI appartenant au fond. Dans les versions futures, nous allons, parmi d'autres choses, (a) automatiser complètement le processus d'initialisation, (b) traiter les objets qui entrent ou sortent de la scène et (c) faire un modèle géométrique indépendant de l'orientation.

Références

- [1] R.T. Collins, A.J. Lipton, and T. Kanade. A system for video surveillance and monitoring. *International Topical Meeting on Robotics and Remote Systems*, pages 497–501, 1999.
- [2] A. Elgammal and L. Davis. Probabilistic framework for segmenting people under occlusion. *International Conference on Computer Vision*, pages 145–152, 2001.
- [3] P. Gabriel, J. Verly, J. Piater, and A. Genon. The state of the art in multiple object tracking under occlusion in video sequences. *Advanced Concepts for Intelligent Vision Systems*, pages 166–173, 2003.
- [4] V. Gouet and N. Boujemaa. About optimal use of color points of interest for content-based image retrieval. *Internal Report, INRIA Rocquencourt*, 2002.
- [5] V. Gouet and B. Lameyre. SAP : A robust approach to track objects in video streams with snakes and points. *British Machine Vision Conference*, pages 737–746, 2004.
- [6] I. Haritaoglu, D. Harwood, and L. S. Davis. W4 - a real time system for detection and tracking people. *European Conference on Computer Vision*, pages 877–892, 1998.
- [7] J. Harris and M. Stephens. A combined corner and edge detector. *4th ALVEY Vision Conference*, pages 147–151, 1988.
- [8] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, pages 5–28, 1998.
- [9] V. Kettner and R. Zabih. Bayesian multicamera surveillance. *Computer Vision and Pattern Recognition*, pages 253–259, 1999.
- [10] S. Khan, O. Javed, Z. Rasheed, and M. Shah. Human tracking in multiple cameras. *International Conference on Computer Vision*, pages 331–336, 2001.
- [11] S. Khan and M. Shah. Tracking people in presence of occlusion. *Asian Conference on Computer Vision*, 2000.
- [12] J.J. Koenderink and A.J. Van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 1987.
- [13] P. Li, T. Zhang, and A.E.C. Pece. Visual contour tracking based on particle filters. *Image and Vision Computing*, pages 111–123, 2003.
- [14] D. Lowe. Object recognition from local scale-invariant features. *International Conference on Computer Vision*, pages 1150–1157, 1999.
- [15] S. McKenna, S. Jabri, Z. Duric, and H. Wechsler. Tracking groups of people. *Computer Vision and Image Understanding*, pages 42–56, 2000.
- [16] P. Montesinos, V. Gouet, and R. Deriche. Differential invariants for color images. *International Conference on Pattern Recognition*, pages 838–840, 1998.
- [17] J.H. Piater and J.L. Crowley. Multi-modal tracking of interacting targets using gaussian approximations. *International Workshop on Performance Evaluation of Tracking and Surveillance*, 2001.
- [18] F. Schaffalitzky and A. Zisserman. Multi-view matching for unordered image sets, or 'How do I organize my holiday snaps?'. *European Conference on Computer Vision*, pages 414–431, 2002.
- [19] C. Schmid and R. Mohr. Local greyvalue invariants for image retrieval. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, pages 530–535, 1997.
- [20] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, pages 151–172, 2000.
- [21] C. Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. *Computer Vision and Pattern Recognition*, pages 246–252, 1999.
- [22] S. Stillman and T. Tanawongsuwan. Tracking multiple people with multiple cameras. *International Conference on Audio- and Video-based Biometric Person Authentication*, 1999.

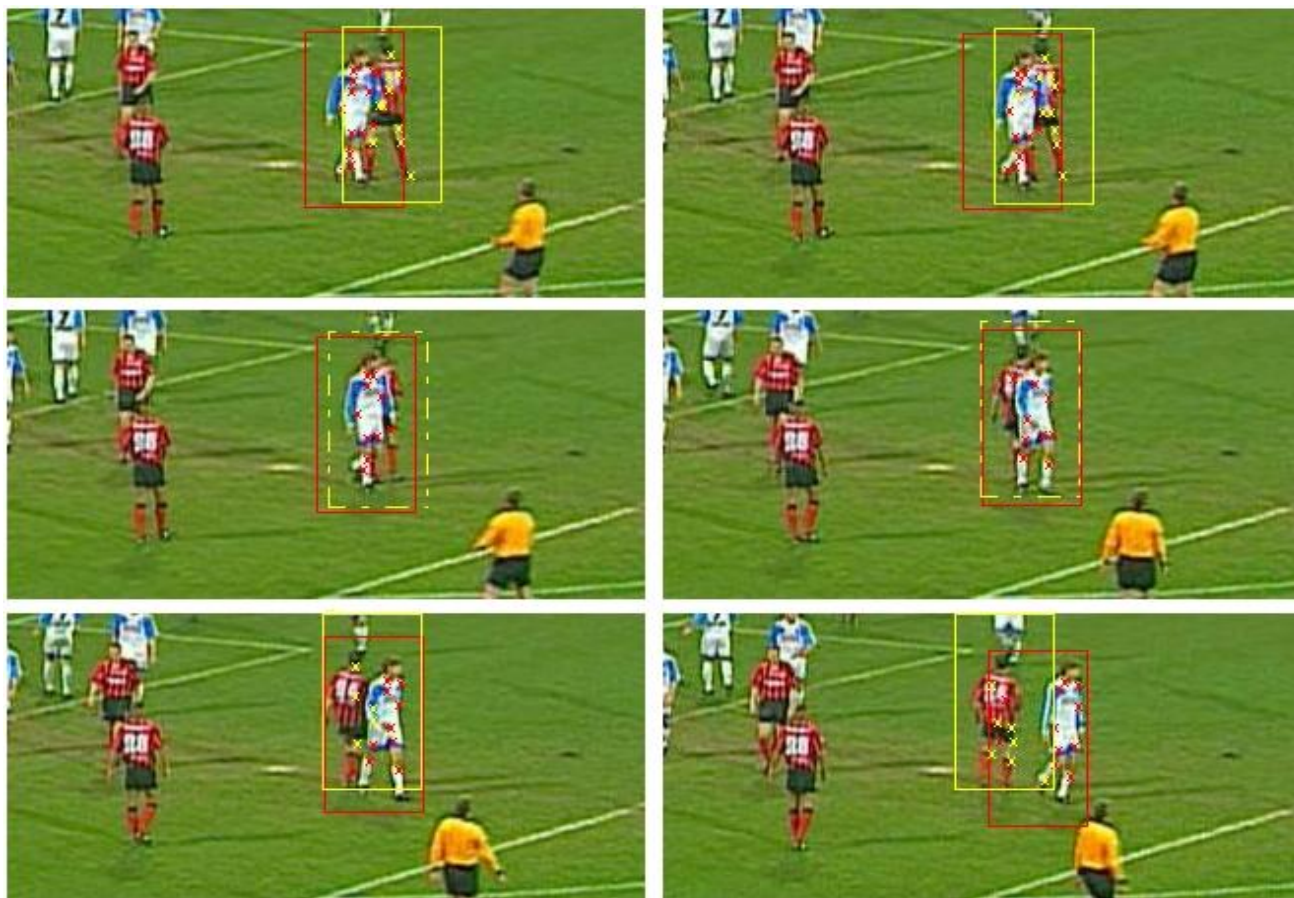


FIG. 9 – Premier exemple décrivant le résultat du suivi de deux objets à travers une occultation en utilisant l’algorithme 3. La ligne en pointillés sur les bords d’une ROI indique que l’objet O correspondant est totalement occulté, ce qui signifie que O est décrit par moins de 3 PI. Les deux joueurs près du centre de l’image sont correctement suivis à travers l’occlusion.

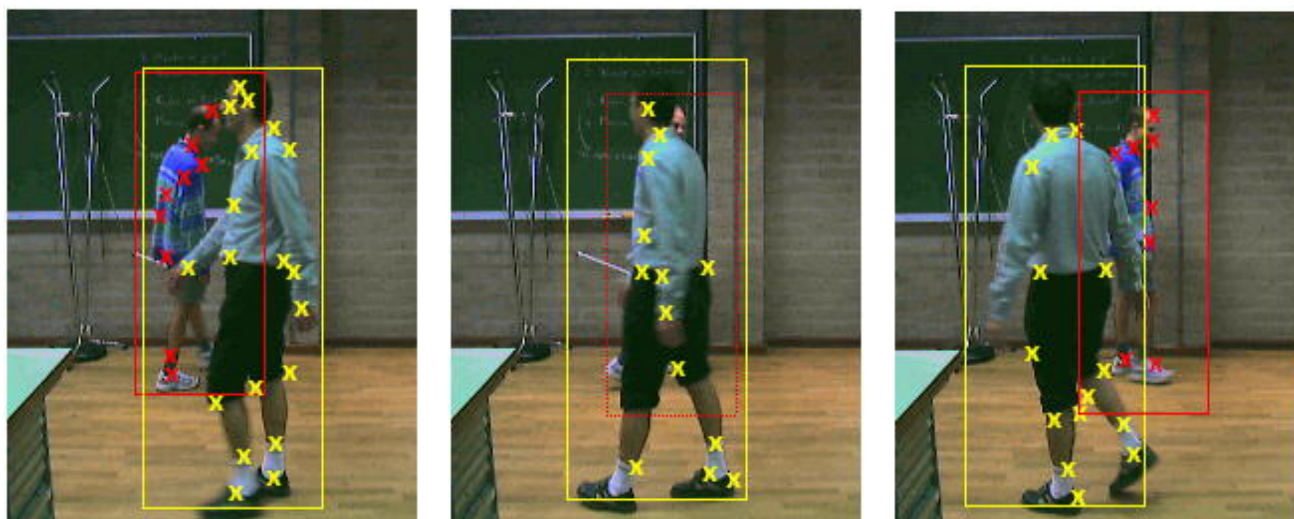


FIG. 10 – Deuxième exemple décrivant le résultat du suivi de deux objets à travers une occultation en utilisant l’algorithme 3. Les conventions et conclusions sont identiques à celle de la figure 9.