

1 Introduction

On implémente dans ce TP l'algorithme de condensation en vision par ordinateur. Cet algorithme, construit à partir de l'algorithme SIR, nous permettra de suivre une zone dans une séquence d'images numériques couleur en se basant sur l'hypothèse que l'histogramme de couleur de la zone à suivre est constant le long de la séquence.

On représente matriciellement les images numériques qu'on manipule. Ainsi, chaque image est représentée par un tableau 2D dont chaque case est un pixel défini l'espace Ω , et tel que Ω est l'espace des valeurs de codage de la couleur (ou du niveau de gris).

Les axes X et Y de l'image sont orientés selon les sens de balayage effectué par l'écran, c'est à dire de gauche à droite et de haut en bas respectivement.

2 Principe de l'algorithme de suivi visuel

La zone à suivre est sélectionnée par l'utilisateur dans la première image de la séquence. On considère que cette zone est rectangle et caractérisée par la position en pixel de centre de rectangle $r = (x, y)$ et par un paramètre d'échelle s exprimé en %.

Au pas de temps k (la k ème image de la séquence), l'état du système à estimer est $X_k = (r_k, s_k)$. La sélection de la zone s'effectue à l'aide du fichier `selection_zone.m` : l'utilisateur définit la zone qu'il veut suivre à l'aide de 4 points afin de former un rectangle. Ce rectangle sera paramétré par les coordonnées du point haut/gauche, une largeur et une hauteur.

2.1 Équation d'état

On suppose que notre système évolue selon l'équation d'état :

$$X_k = X_{k-1} + W_k$$

où W_k est un bruit blanc gaussien centré en 0 et de matrice de covariance C de dimension 3×3 d'éléments diagonaux c_1, c_2 et c_3 . La diffusion de particules se fait selon la loi $p(x_k|x_{k-1})$, et on a :

$$\begin{aligned} p(x_k|x_{k-1}) &= \mathbb{P}(X_k = x_k | X_{k-1} = x_{k-1}) \\ &= \mathbb{P}(W_k = x_k - x_{k-1}) \end{aligned}$$

Avec : W_k admet la densité de probabilité : $\frac{1}{(2\pi)^{3/2} \sqrt{c_1 c_2 c_3}} e^{-\frac{1}{2}(\frac{x^2}{c_1} + \frac{y^2}{c_2} + \frac{s^2}{c_3})}$

2.2 Modèle de couleurs

La zone qu'on souhaite suivre et qu'on sélectionne sur la première image est caractérisée par un histogramme de couleur construit à partir des Nb couleurs les plus représentatifs de la zone. Cet histogramme constituera un histogramme de référence pour la suite, on le notera $q^* = \{q^*(n), n = 1..Nb\}$ où $q(n)$ représente le nombre normalisé de pixels de la zone initiale dont la couleur la plus proche est la

couleur n .

Afin d'effectuer le suivi de la zone initiale dans une séquence d'images, on part de l'hypothèse que l'histogramme de couleur est invariant dans le temps. Au pas de temps k (kème image), l'histogramme de couleur $q_k(x)$ d'un état hypothèse x sera comparé au modèle de couleur de référence q^* , et on définit la distance entre deux histogrammes par :

$$\begin{aligned} D(q^*, q_k(x)) &= (1 - \sum_{n=1}^{N_b} \sqrt{q^*(n)q_k(x, n)})^{1/2} \\ &= (\frac{1}{2}(2 - 2 \sum_{n=1}^{N_b} \sqrt{q^*(n)q_k(x, n)}))^{1/2} \\ &= (\frac{1}{2}(\sum_{n=1}^{N_b} q^*(n) + \sum_{n=1}^{N_b} q_k(x, n) - 2 \sum_{n=1}^{N_b} \sqrt{q^*(n)q_k(x, n)}))^{1/2} \end{aligned}$$

Car $\sum_{n=1}^{N_b} q^*(n) = 1$ et $\sum_{n=1}^{N_b} q_k(x, n) = 1$ Ainsi :

$$D(q^*, q_k(x)) = (\frac{1}{2} \sum_{n=1}^{N_b} (\sqrt{q^*(n)} - \sqrt{q_k(x, n)})^2)^{1/2}$$

Ce qui définit bien une distance entre l'histogramme de référence et l'histogramme de la zone centrée à x (distance de Helliger). Pour favoriser les états hypothèses dont l'histogramme de couleur associé est proche de l'histogramme de référence, on introduit la fonction de vraisemblance :

$$g_k(x) \propto \exp(-\lambda D^2(q^*, q_k(x)))$$

Le poids d'importance d'une particule ξ_k^i sera calculé par : $\frac{g_k(\xi_k^i)}{\sum_{j=1}^N g_k(\xi_k^j)}$. Ce poids sera grand si la distance

$D^2(q^*, q_k(\xi_k^i))$ est petite.

3 Structure de l'algorithme

On implémente l'algorithme de suivi d'une zone le long d'une séquence d'image dans le fichier matlab TP2 comme suit :

- Lecture de l'ensemble des images de la séquence.
- Initialisation des paramètres de l'algorithme (N , N_b , λ , $c1$, $c2$, $c3$).
- Affichage de la première image et sélection de la zone à suivre.
- Construction de l'histogramme de couleur de référence.
- Initialisation des particules (N particules ayant l'état X_0).

- Pour chaque étape de temps $k \in [1, \text{nbr d'images dans la séquence}]$, on effectue :
 - l'étape de propagation : $\xi_k^i = \hat{\xi}_{k-1}^i + W_k^i, i \in [1, N]$.
 - Calcul de l'histogramme de couleur associé à chaque état et les poids d'importance.
 - Normalisation des poids.
 - Ré-échantillonnage à l'aide de la fonction `fct_mult` utilisée dans le TP précédent pour sélectionner les particules $\hat{\xi}_k$ ayant les plus grands poids.

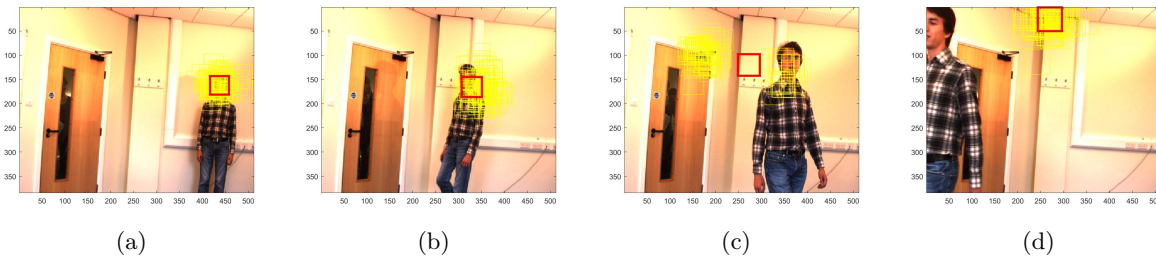
4 Tests sur l'algorithme

4.1 Zone de mouvement important

On teste l'algorithme pour les deux séquences d'images sur une zone de mouvement important, on choisit pour cela le visage de la personne qui conduit l'expérience.

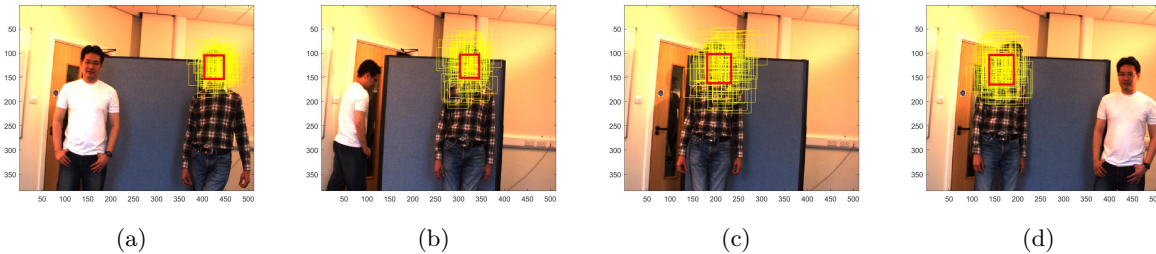
On obtient les résultats suivants sur la séquence 1:

Avec : $c1 = 300$; $c2 = 200$; $c3 = 10$;



On voit que notre approximation suit assez bien le visage de la personne jusqu'au grandissement de ce dernier, où le filtre perd le visage. Cela est sûrement dû à la faible valeur de variance du scale.

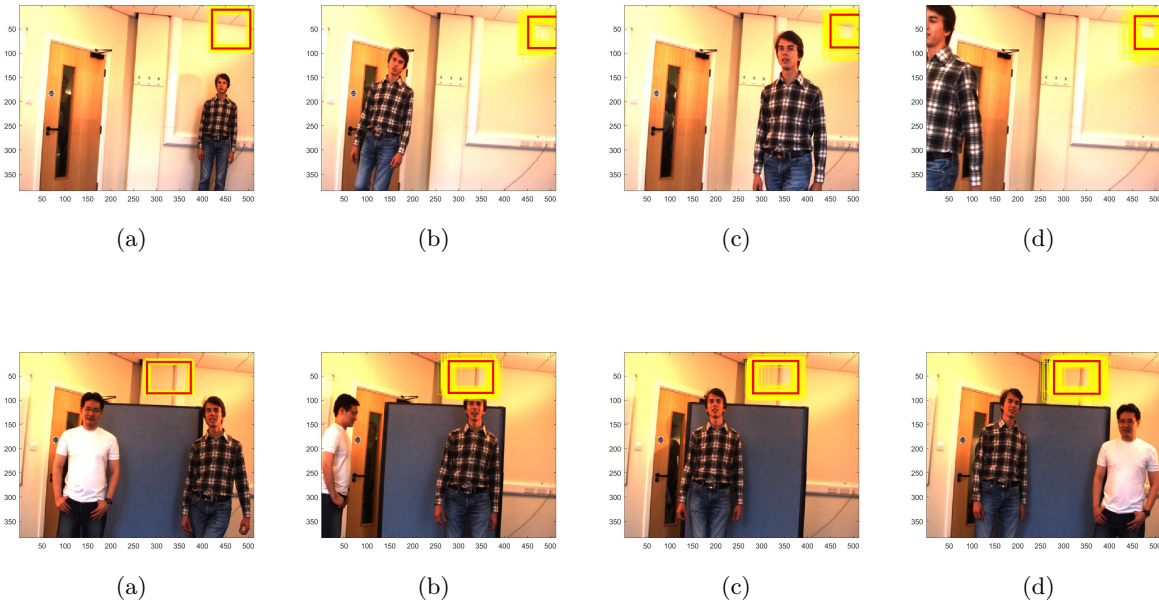
On obtient les résultats suivants sur la séquence 2:



4.2 Zone immobile

On teste également l'algorithme sur une zone immobile des deux photos, non occultée durant toute la séquence.

On utilise les valeurs suivantes de variance : $c1 = 10$; $c2 = 10$; $c3 = 10$;



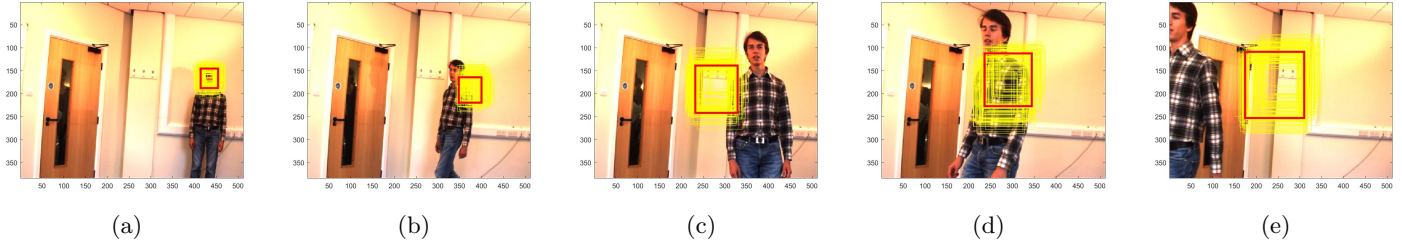
Que ce soit pour la séquence 1 ou 2, l'algorithme reste stable et le suivi est immédiat. Pour la séquence 1, la pondération des particules sort de l'image, cela pourrait être dû à la proximité de la zone choisie avec le coin de l'image.

4.3 Zone avec zoom

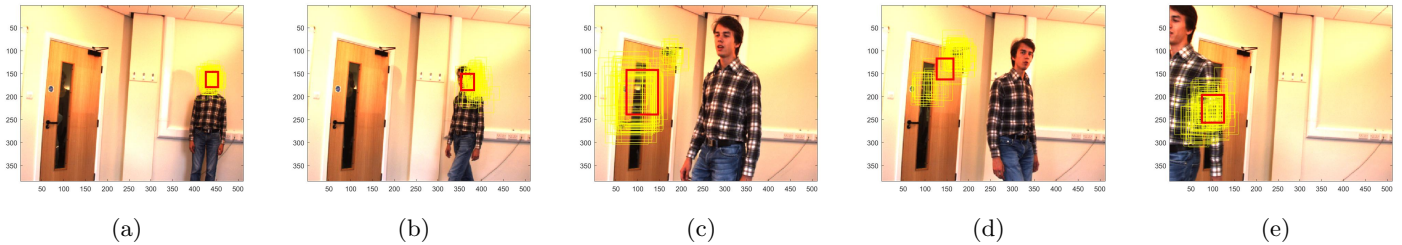
On travaille sur la séquence 1 où le visage de la personne subit un zoom arrière et avant dans son déplacement tout au long de la séquence.

On essaie pour différentes valeurs de variance.

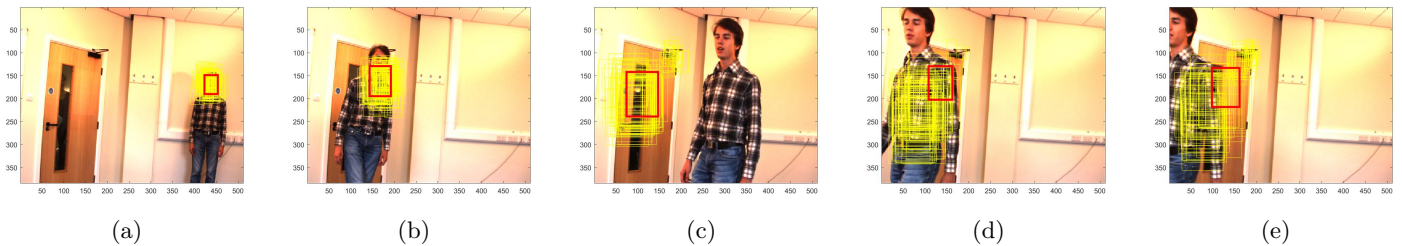
$c1 = 50; c2 = 50; c3 = 200;$



$c1 = 100; c2 = 100; c3 = 100;$

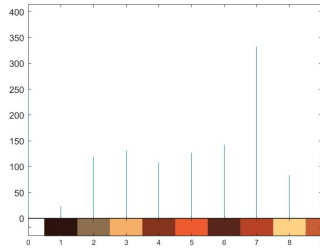


$c1 = 120; c2 = 120; c3 = 180;$

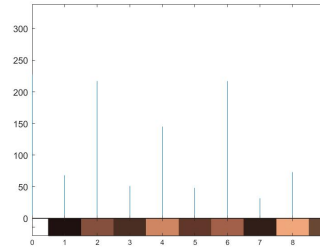


Dans nos différents essais, le modèle a des difficultés à faire le suivi du visage de la personne pendant tout le long. Selon les valeurs de variance, on va pouvoir plus ou moins récupérer le suivi puis le perdre. Cela peut s'expliquer par la similarité d'histogrammes de couleur pour plusieurs parties de l'image.

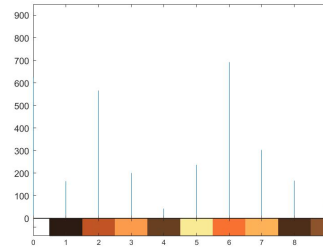
Ci-dessous nous avons tracé les histogrammes de couleur de certaines zones qui ont été indiquées par l'algorithme sur la séquence. Et on peut remarquer certaines similarités.



(a) Histogramme visage



(b) Histogramme chemise



(c) Histogramme porte

4.4 Zone occultée durant la séquence

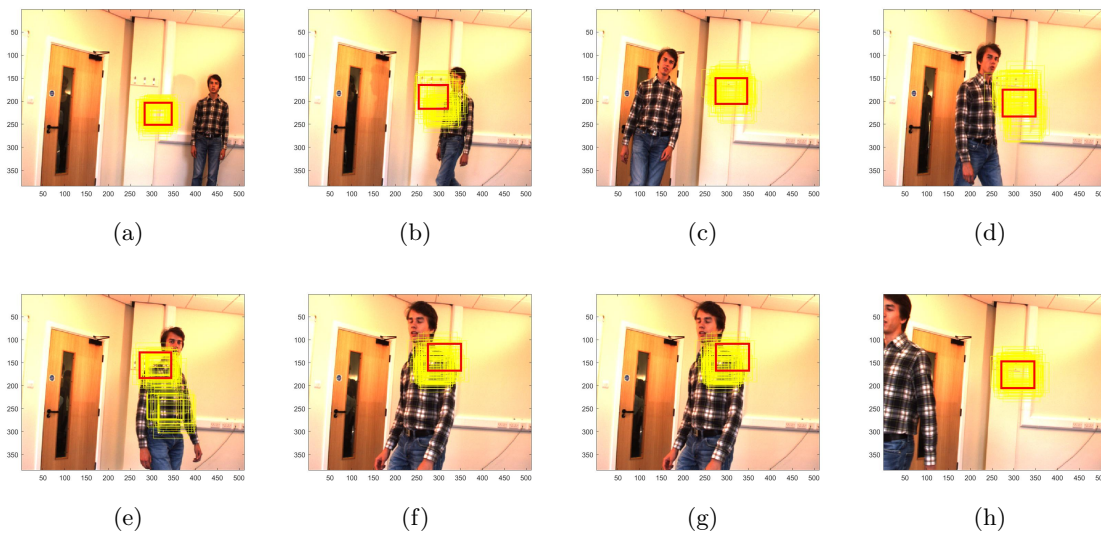
Dans cette partie, on teste le suivi d'une zone qui a occultée partiellement ou complètement au moins une fois durant le déroulement des images.

On choisit des valeurs de : $c1 = 50$; $c2 = 50$; $c3 = 10$; pour les variances.

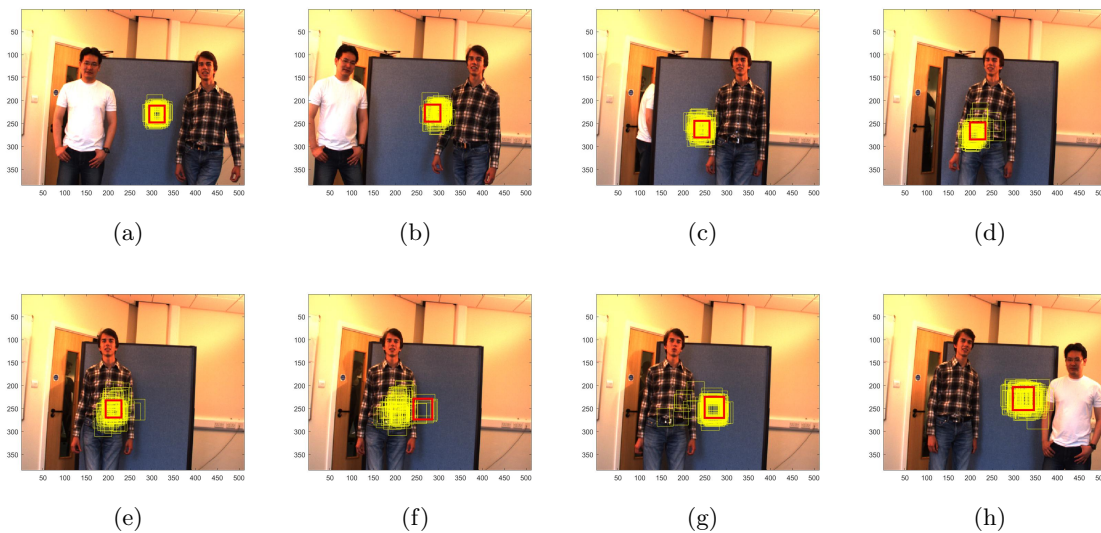
Pour les deux séquences, l'occultation gêne le suivi de la zone occultée pendant quelques instants. L'estimation se décale alors vers la zone qui se rapproche le plus de la zone sélectionnée d'un point de vue colorimétrie. L'estimation arrive néanmoins à retrouver la zone vers la fin de la séquence.

Il faut aussi prendre en compte le fait que les zones occultées choisis pour les deux séquences sont presque uniformes en couleur et représente la couleur dominante dans les images.

Séquence 1



Séquence 2



5 Faiblesses et améliorations de l'algorithme

Le choix des paramètres de l'algorithme permet d'améliorer ses performances et d'éviter certains problèmes. Par exemple :

- Le Nb (nombre de couleurs de l'histogramme) doit être suffisamment grand pour éviter qu'il y ait des zones différentes avec des histogrammes de couleurs semblables voire identiques.
- Les valeurs des éléments diagonaux de la matrice de covariance doivent nous permettre de balayer tout l'espace lors des étapes de diffusion afin de suivre les zones qui font des déplacements considérables, mais elles doivent en même temps prendre en considération le fait qu'on doit rester dans le cadre de l'image et ne pas dépasser ses dimensions.