

# TP2 SOD 333 - Suivi visuel par histogramme de couleur

Luc de Montella - Joshua Wolff

Octobre 2021

## Table des matières

Contexte général	1
Question 1	1
Question 2	2
Question 3	2
Question 4	3
Question 5	5
Question 6	6
Question 7	6

## Contexte général

Ce TP propose d'implémenter un filtre particulaire permettant de suivre un élément dans une suite d'images. L'hypothèse de base de notre méthode repose sur la stationnarité de l'histogramme de couleur de l'élément que nous voulons poursuivre.

Nous nous intéressons dans un premier temps à la distance permettant de quantifier l'écart entre deux histogrammes, puis aux éléments théoriques nécessaires à la mise en place du filtre particulaire (probabilité de transition / poids d'importance des particules). La suite de ce rapport se concentre sur la présentation des résultats dans différentes situations (changement de l'élément à suivre / variation des paramètres du filtre). On propose dans un dernier temps quelques pistes d'amélioration de notre modèle.

## Question 1

L'histogramme de référence est noté  $q^* = \{q^*(n), n = 1, \dots, Nb\}$ , où  $q^*(n)$  désigne le nombre normalisé de pixels de la zone initiale dont la couleur la plus proche est la couleur  $n$ . On a donc  $\sum_{n=1}^{Nb} q^*(n) = 1$ .

À un temps  $k$  fixé, on note  $q_k(x)$  l'histogramme de couleur d'un état hypothèse  $x$ . Comme notre objectif est de suivre une zone de l'image sous l'hypothèse que son histogramme de couleur est invariant, nous cherchons à comparer  $q_k(x)$  avec notre référence  $q^*$ .

On introduit la quantité  $D(q^*, q_k(x))$  définie comme suis :

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

On note  $E$  l'espace des suites réelles composées de  $Nb$  termes positifs et de somme unité. Soient  $p, q \in E$ , montrons que  $D$  définit bien une distance sur  $E$  :

$$\begin{aligned}
D(p, q) &= \sqrt{1 - \sum_{n=1}^{Nb} \sqrt{p_n q_n}} \\
&= \sqrt{\frac{1}{2} \left( 2 - 2 \sum_{n=1}^{Nb} \sqrt{p_n q_n} \right)} \\
&= \sqrt{\frac{1}{2} \left( \sum_{n=1}^{Nb} p_n + \sum_{n=1}^{Nb} q_n - 2 \sum_{n=1}^{Nb} \sqrt{p_n q_n} \right)} \\
&= \sqrt{\frac{1}{2} \sum_{n=1}^{Nb} (p_n + q_n - 2\sqrt{p_n q_n})} \\
&= \sqrt{\frac{1}{2} \sum_{n=1}^{Nb} (\sqrt{p_n} - \sqrt{q_n})^2}
\end{aligned}$$

On note alors  $\sqrt{p}$  le vecteur de  $\mathbb{R}^{Nb}$  dont le  $n$ -ième coefficient est la racine carrée du  $n$ -ième terme de la suite  $p$  (licite car tous les termes de  $p$  sont positifs ou nuls). On note par ailleurs  $\|\cdot\|$  la norme euclidienne sur  $\mathbb{R}^{Nb}$  :

$$= \frac{1}{\sqrt{2}} \|\sqrt{p} - \sqrt{q}\|$$

Comme  $(x, y) \rightarrow \|x - y\|$  définit une distance sur  $\mathbb{R}^{Nb}$ ,  $D$  vérifie bien les hypothèses de symétrie, séparation ainsi que l'inégalité triangulaire. C'est donc bien une distance sur  $E$ .

## Question 2

On souhaite désormais exprimer la densité de probabilité de transition  $p(x_k | x_{k-1})$  :

$$p(x_k | x_{k-1}) = \mathbb{P}(X_k \in dx_k | X_{k-1} = x_{k-1})$$

Compte tenu de l'équation dynamique :

$$\begin{aligned}
&= \mathbb{P}(X_{k-1} + W_k \in dx_k | X_{k-1} = x_{k-1}) \\
&= q^W(x_k - x_{k-1}) dx_k
\end{aligned}$$

Avec  $q^W$  la densité d'une loi gaussienne multivariée  $\mathcal{N}(0_{\mathbb{R}^3}, \Sigma)$  avec  $\Sigma = \begin{pmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{pmatrix}$ .

On s'intéresse maintenant au calcul du poids d'importance d'une particule  $i$  au temps  $k$ . On introduit pour cela la fonction suivante :

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

On en déduit l'expression du poids d'importance  $w_k^i$  de la particule  $i$  à l'instant  $k$  (notée  $x_i^k$ ) :

$$w_k^i = \frac{g_k(x_i^k)}{\sum_{j=1}^N g_k(x_j^k)}$$

## Question 3

Cette question correspond aux codes en Matlab fournis en complément de ce rapport.

## Question 4

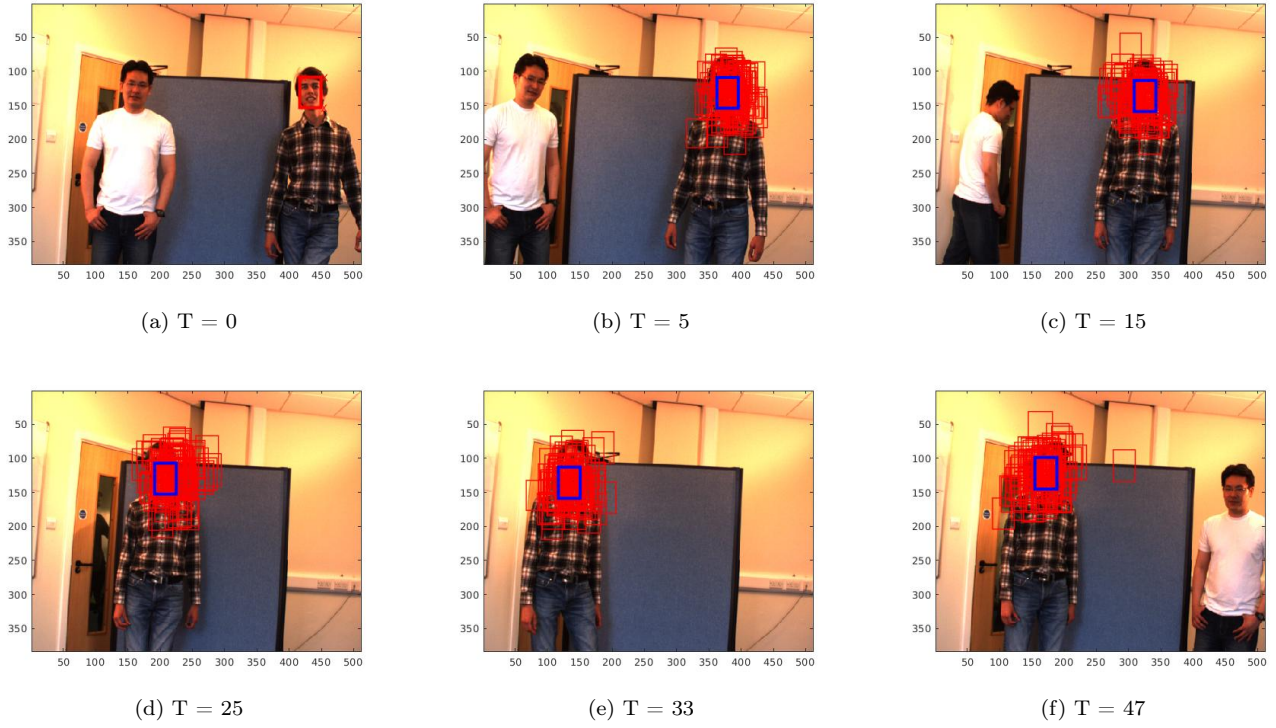


FIGURE 1 – Suivi d'un visage (paramètres :  $c_1 = c_2 = 300$ ,  $c_3 = 100$ ,  $N = 100$ ,  $Nb = 10$ ,  $\lambda = 20$ )

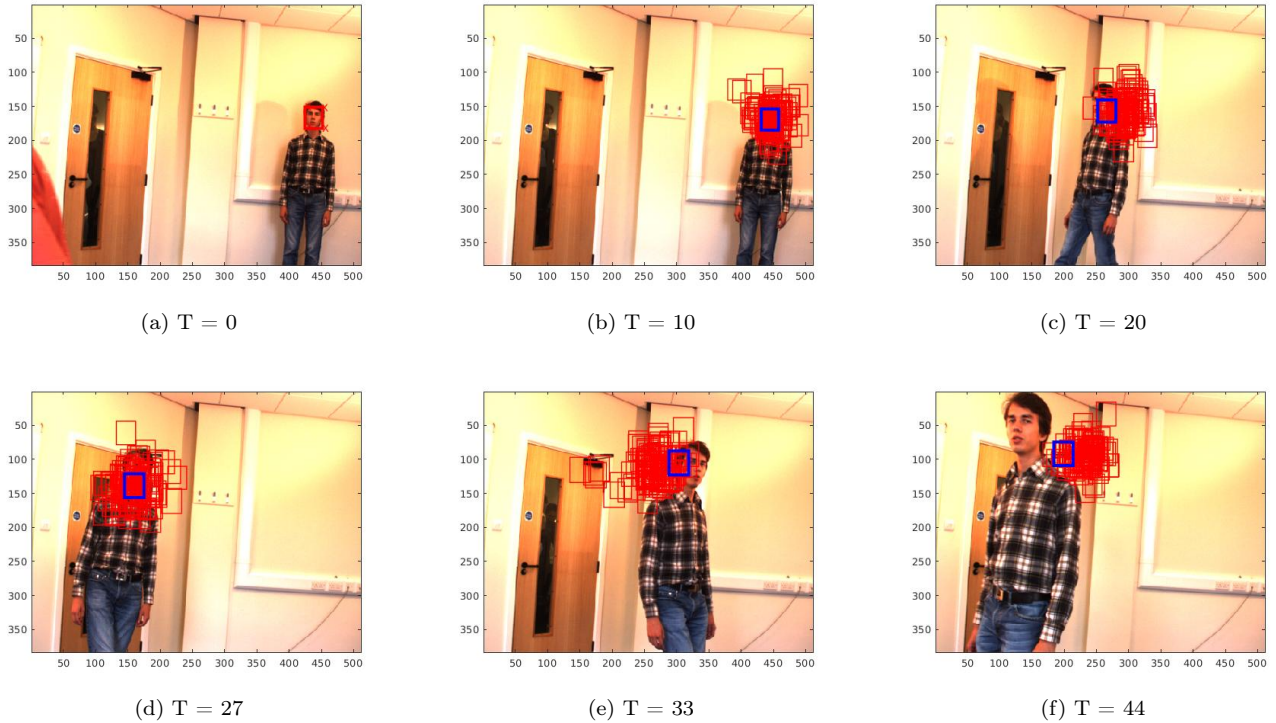
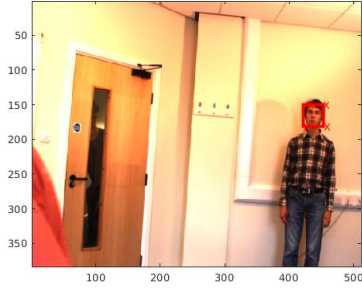
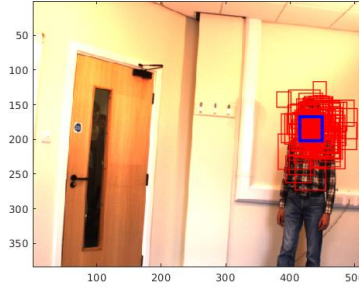


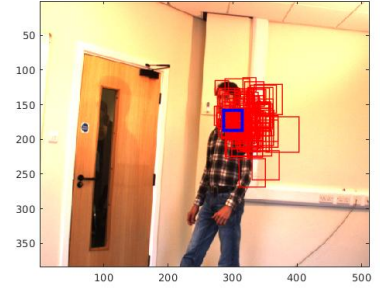
FIGURE 2 – Suivi d'un visage (paramètres :  $c_1 = c_2 = 300$ ,  $c_3 = 2$ ,  $N = 100$ ,  $Nb = 10$ ,  $\lambda = 20$ )



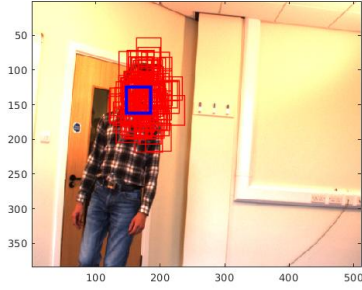
(a)  $T = 0$



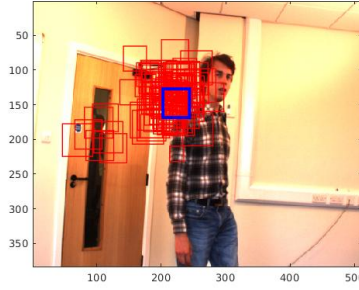
(b)  $T = 12$



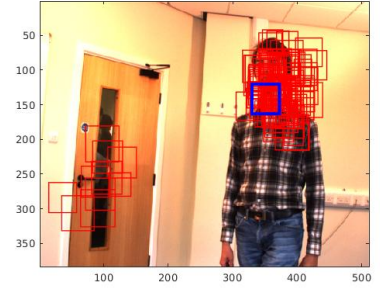
(c)  $T = 19$



(d)  $T = 25$

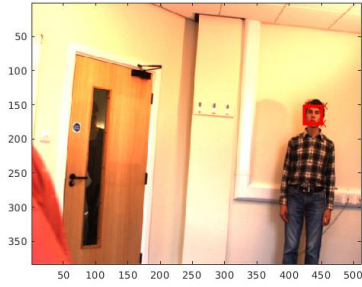


(e)  $T = 32$



(f)  $T = 39$

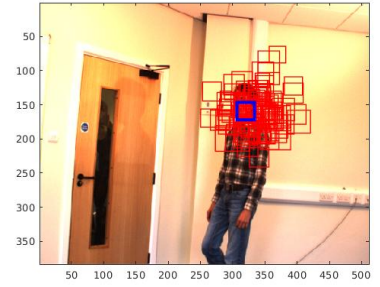
FIGURE 3 – Suivi d'un visage (paramètres :  $c_1 = c_2 = 300$ ,  $c_3 = 100$ ,  $N = 100$ ,  $Nb = 10$ ,  $\lambda = 20$ )



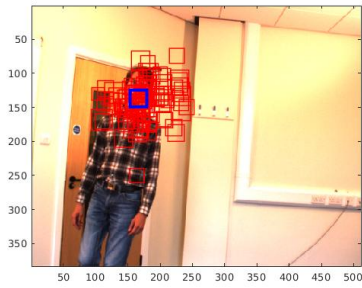
(a)  $T = 0$



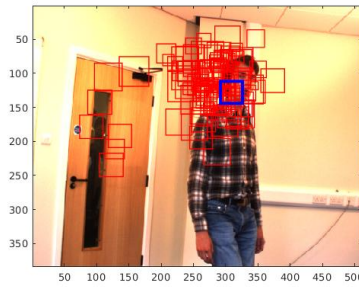
(b)  $T = 11$



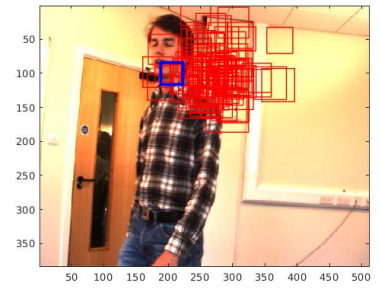
(c)  $T = 18$



(d)  $T = 25$



(e)  $T = 33$



(f)  $T = 43$

FIGURE 4 – Suivi d'un visage (paramètres :  $c_1 = 1000$ ,  $c_2 = 500$ ,  $c_3 = 100$ ,  $N = 100$ ,  $Nb = 10$ ,  $\lambda = 20$ )



La figure 1 correspond au suivi d'un visage dans un cas peu dynamique (déplacement lent et pas de zoom). On constate que le visage initialement sélectionné est correctement suivi.

Dans le cas de la figure 2 le visage sélectionné est désormais plus mobile. Il y a à la fois un rapprochement et des déplacements latéraux. Le suivi est initialement de bonne qualité mais se dégrade lorsque le visage se rapproche. Pour remédier à ce problème nous avons d'abord tenté d'augmenter la variance associée au paramètre d'échelle  $c_3$  (c.f. figure 3). On constate que le suivi de la zone sélectionnée s'améliore, mais qu'en augmentant aussi les variances associées au paramètres de position  $c_1$  et  $c_2$  on obtient des résultats encore plus probants. Ces résultats sont présentés dans la figure 4.

## Question 5

On cherche désormais à étudier l'effet d'une occultation de la zone sélectionnée sur la qualité du suivi par notre filtre particulière. On sélectionne pour cela une zone fixe qui va être cachée par le déplacement d'une personne. La figure 5 montre les performances du filtre dans cette situation. On observe que l'algorithme va se déplacer vers des zones similaires (il est dans un premier temps "poussé" par la personne). L'algorithme indique le panneau bleu autant que possible puis "passe" par le pantalon bleu de la personne lorsque le tableau n'est plus accessible. Il retourne finalement au centre du panneau bleu, qui est la zone initialement sélectionnée.

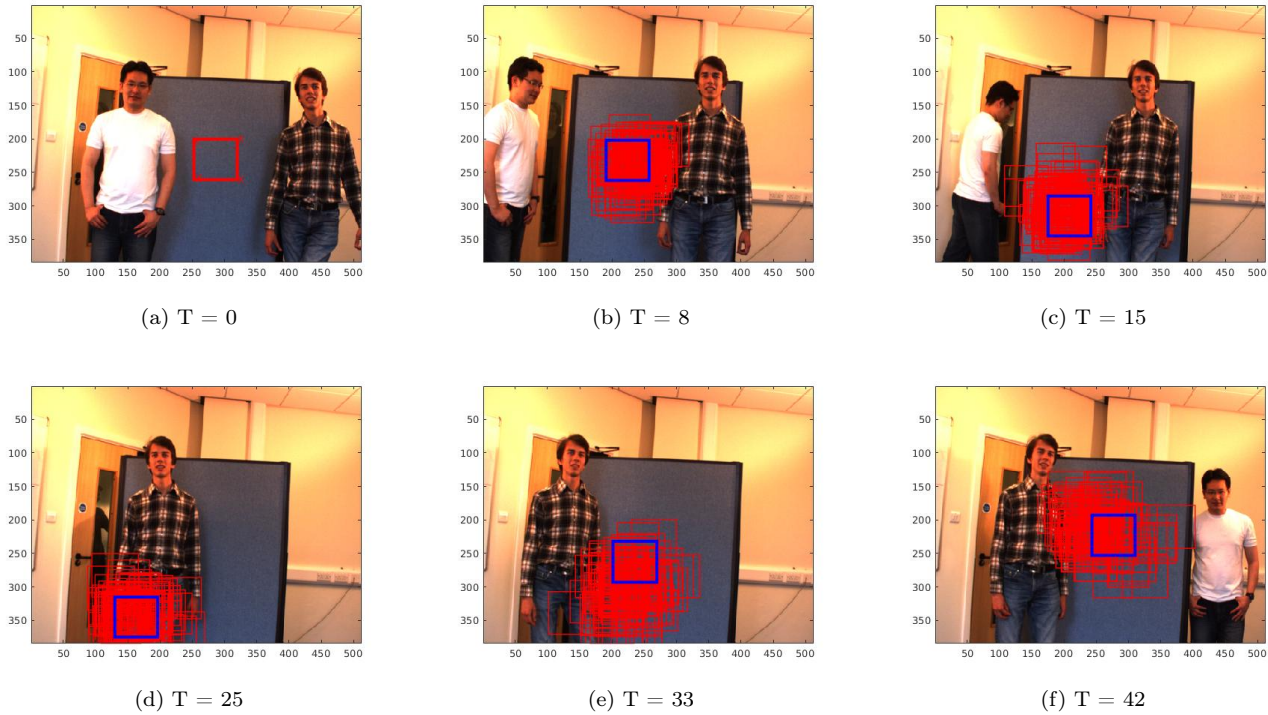


FIGURE 5 – Suivi d'une zone occultée (paramètres :  $c_1 = c_2 = 300$ ,  $c_3 = 2$ ,  $N = 100$ ,  $Nb = 10$ ,  $\lambda = 20$ )

On suit cette fois le visage d'un personnage qui disparaît derrière le tableau puis réapparaît (c.f. figure 6). Dans un premier temps le visage est correctement suivi. Lorsque le visage disparaît derrière le tableau, la zone de suivi reste dans la zone par laquelle le visage initial a disparu. Lorsque le visage réapparaît, notre algorithme n'est pas capable de le récupérer (aucune particule ne se trouve alors dans la zone d'intérêt). Nous sommes alors "coincé" sur le visage de la seconde personne qui est désormais à la place initialement occupée par le personnage au t-shirt blanc (dont nous avons sélectionné le visage).

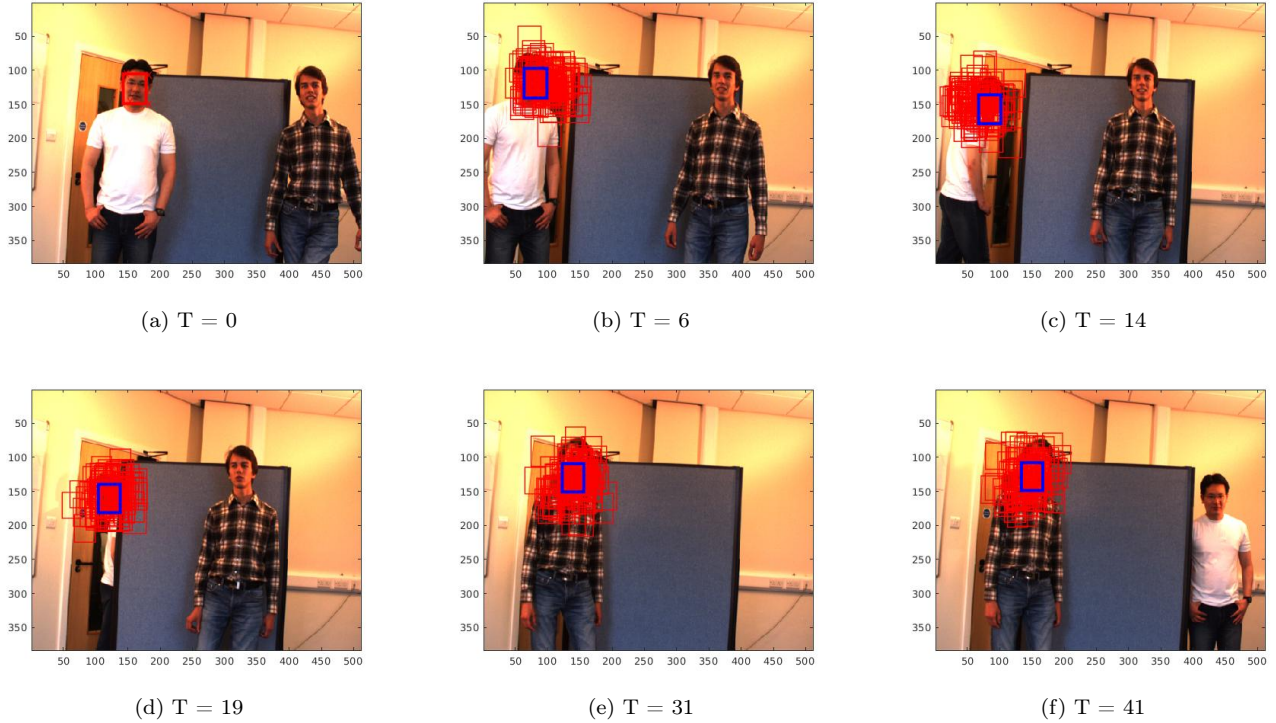


FIGURE 6 – Suivi d'un visage (paramètres :  $c_1 = c_2 = 300$ ,  $c_3 = 100$ ,  $N = 100$ ,  $Nb = 10$ ,  $\lambda = 20$ )

## Question 6

On peut observer que les résultats présentés en figure 5 illustre la performance du modèle dans le cadre du suivi d'une zone avec de faibles nuances de couleurs / peu discriminante. En effet, même si le reste du panneau semble présenter des caractéristiques colorimétriques semblables à celles du centre du tableau, le filtre particulière parvient à retourner sur la partie du tableau initialement sélectionnée (le centre) après que la personne l'ait occultée.

Remarquons cependant que lorsque l'on sélectionne une zone avec trop peu de nuances de couleurs (par exemple un pan des murs blancs), la fonction de calcul des histogrammes de couleurs retourne des histogrammes contenant moins de  $Nb$  valeurs. Ceci cause des erreurs empêchant le bon déroulement de notre algorithme. Nous n'avons pas réussi à résoudre ce problème d'implémentation.

## Question 7

Le principale faiblesse du filtre réside dans son incapacité à sortir de certains *minima locaux*. Ce genre de problème est notamment observé en figure 6. En effet, après la réapparition du visage initialement sélectionné le filtre est incapable de le récupérer. Le visage de la seconde personne est une sorte de minimal local dont le nuage de particule ne peut pas s'extraire pour retrouver le premier visage qui est désormais assez éloigné. Pour résoudre ce type de problème, on pourrait envisager d'augmenter ponctuellement les variances  $c_1$  et  $c_2$  pour tenter de re-capturer la zone initiale si la distance moyenne à l'histogramme de référence est au dessus d'un certain seuil.

Plus généralement, on pourrait chercher à avoir des variances  $c_1$ ,  $c_2$  et  $c_3$  adaptatives. Il faudrait pour cela être capable de calculer une grandeur qui décrit la "quantité de mouvement" entre deux images successives. Ainsi les variances associées au filtre devraient être des fonctions croissantes de la "quantité de mouvement" entre deux images.

Enfin, nous avons pu observer que dans certaines situations le nuage peut se scinder en plusieurs parties (c.f. dernière image de la figure 3). Il serait pertinent de "ramener" les particules égarées dans le nuage de particule ayant la distance à l'histogramme de référence la plus faible.