

MAT 4501
INFÉRENCE BAYÉSIENNE DANS DES MODÈLES
MARKOVIENS



Filtrage particulaire



Eloi CAMPAGNE



Saad AMMARI

Table des matières

Table des figures

1	Le modèle de Kitagawa	1
2	Suivi de visage sur une séquence vidéo	5
2.1	Suivi 2D	5
2.1.1	Influence du paramètre λ	6
2.1.2	Influence des variances	8
2.2	Suivi 3D	11

Table des figures

1.1	Trajectoires réelle, observée et estimée ($T = 50, N = 50, Q = 1, R = 1$)	2
1.2	Évolution de l'erreur moyenne en fonction du nombre de particules . . .	3
1.3	Évolution du temps d'exécution du programme en fonction du nombre de particules	3
1.4	Courbe proportionnelle à t et à $\frac{EQM}{t}$ (ici : $y = 50t + \frac{EQM}{t}$)	4
2.1	Histogramme de couleur pour $N_b = 10$	6
2.2	Histogramme de couleur pour $N_b = 50$	6
2.3	Nombre effectif de particules pour $\lambda = 5$ (séquence 3)	7
2.4	Nombre effectif de particules pour $\lambda = 20$ (séquence 3)	7
2.5	Suivi de la personne sur la séquence 1 ($N = 25, C_1 = C_2 = 300,$ $N_b = 10, \lambda = 5$)	8
2.6	Suivi de la personne sur la séquence 1 ($N = 25, C_1 = 2000, C_2 = 300,$ $N_b = 10, \lambda = 5$)	9
2.7	Suivi de la personne en blanc sur la séquence 3 ($N = 25, C_1 = C_2 =$ $300, N_b = 10, \lambda = 5$)	9
2.8	Suivi de la personne en blanc sur la séquence 3 ($N = 25, C_1 =$ $2000, C_2 = 300, N_b = 20, \lambda = 5$)	10
2.9	Algorithme de reconnaissance faciale s'appuyant sur du <i>deep learning</i>	11
2.10	Suivi du surfeur ($N = 25, C_1 = C_2 = 300, C_3 = 0.02, N_b = 10, \lambda = 5$)	11
2.11	Suivi du surfeur ($N = 25, C_1 = C_2 = 300, C_3 = 0.02, N_b = 10, \lambda = 30$)	12
2.12	Suivi du surfeur ($N = 25, C_1 = C_2 = 300, N_b = 10, \lambda = 30$)	12

Le modèle de Kitagawa

L'**objectif** de ce TP est d'implémenter le filtrage particulaire afin dans un premier temps d'estimer des trajectoires cachées à partir d'observations bruitées, puis dans un second temps de faire du suivi de visage sur une vidéo.

Le modèle de Kitagawa repose sur les équations suivantes :

$$\begin{cases} X_n = 0.5X_{n-1} + 25\frac{X_{n-1}}{1+X_{n-1}^2} + 8\cos(1.2n) + U_n \\ Y_n = \frac{X_n^2}{20} + V_n \end{cases} \quad (1.1)$$

où $U_n \sim \mathcal{N}(0, Q)$ et $V_n \sim \mathcal{N}(0, R)$.

L'équation d'observation ne nous permet pas d'obtenir directement l'état caché, en effet la présence du carré nous donne deux solutions possibles pour X_n : il nous faut une connaissance supplémentaire afin d'estimer l'état caché.

Du système (1.1) et de la connaissance sur les bruits gaussiens, nous pouvons en déduire la loi de transition de X_n .

$$f_{n|n-1}(x_n|x_{n-1}) = 0.5x_{n-1} + 25\frac{x_{n-1}}{1+x_{n-1}^2} + 8\cos(1.2n) + \frac{1}{\sqrt{2\pi Q}} \exp\left\{-\frac{x_{n-1}^2}{2Q}\right\} \quad (1.2)$$

De même, on a la vraisemblance :

$$g_n(y_n|x_n) = \frac{x_n^2}{20} + \frac{1}{\sqrt{2\pi R}} \exp\left\{-\frac{x_n^2}{2R}\right\} \quad (1.3)$$

Avec (1.1), (1.2) et (1.3), on peut estimer l'état caché X_n . La FIGURE 1.1 compare les trajectoires observées et estimées par l'intermédiaire du filtrage particulaire.

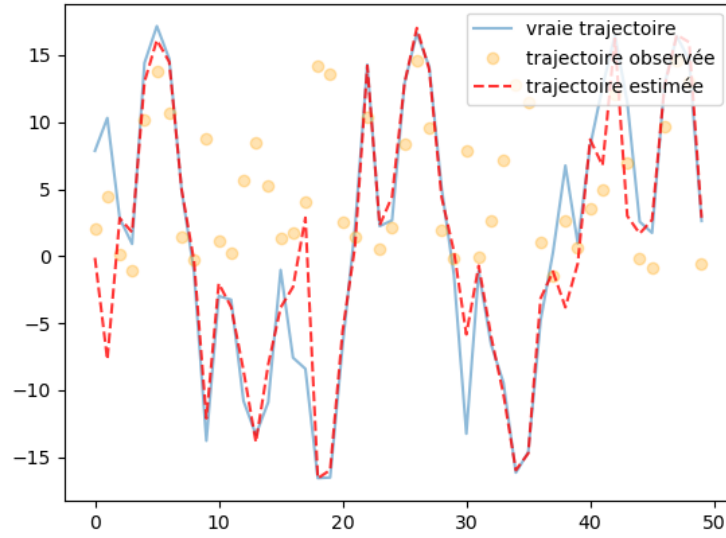


FIGURE 1.1 – Trajectoires réelle, observée et estimée ($T = 50$, $N = 50$, $Q = 1$, $R = 1$)

On peut mesurer l'erreur quadratique moyenne entre la véritable trajectoire et la trajectoire estimée par filtrage particulaire. On conserve les paramètres $T = 50$, $N = 50$, $Q = 1$, $R = 1$.

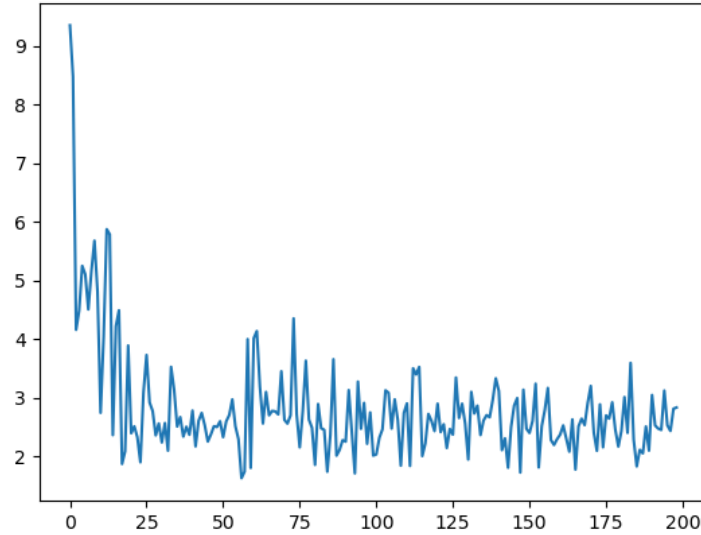


FIGURE 1.2 – Évolution de l’erreur moyenne en fonction du nombre de particules

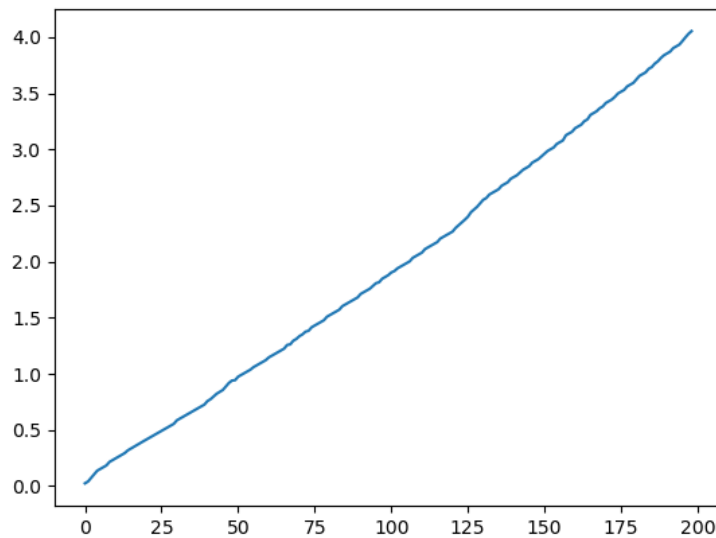


FIGURE 1.3 – Évolution du temps d’exécution du programme en fonction du nombre de particules

On remarque que l’erreur moyenne diminue rapidement en fonction du nombre de particules, cependant on remarque aussi qu’à partir d’un certain seuil (ici environ 25), peu importe le nombre de particules mises en jeu, l’erreur sera sensiblement

la même. Ceci doit tout de même être mis en relation avec le temps d'exécution du programme qui est linéairement proportionnel au nombre de particules : la complexité est en $O(N)$ (avec N le nombre de particules). Il est donc judicieux d'utiliser le filtrage particulaire avec un nombre de particules juste suffisant pour que l'erreur moyenne soit minimale. On peut par exemple se référer au graphique suivant pour déterminer le nombre de particules nécessaires. En effet, prendre en compte uniquement le rapport de l'EQM sur le temps ne rend pas bien compte de la contrainte temporelle, d'où l'ajout d'un terme en t . De cette manière, en 0 on a comme équivalent $\frac{1}{t}$ et en $+\infty$ on a comme équivalent t .

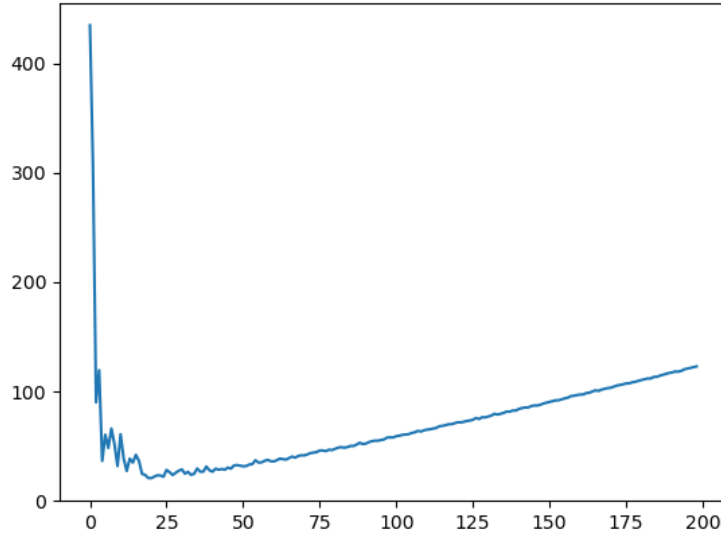


FIGURE 1.4 – Courbe proportionnelle à t et à $\frac{EQM}{t}$ (ici : $y = 50t + \frac{EQM}{t}$)

D'après cette figure, 25 particules semblent être suffisantes pour le programme.

Suivi de visage sur une séquence vidéo

2.1 Suivi 2D

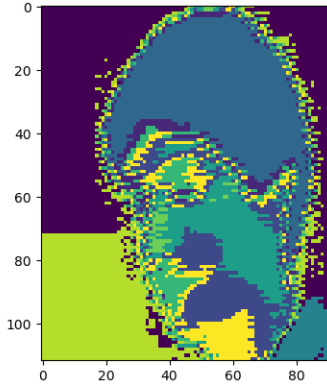
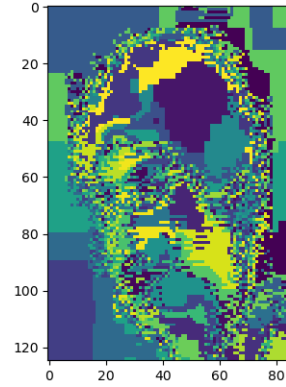
On s'intéresse dans cette partie au suivi d'un visage sur une séquence vidéo. Pour ce faire, on se donne comme équation de transition d'un état caché $n - 1$ à un état caché n :

$$X_n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} X_{n-1} + U_n \quad (2.1)$$

où $U_n \sim \mathcal{N} \left(0_{2 \times 2}, \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \right)$. On a de plus la vraisemblance qui nous est donnée par :

$$g_n(y_n | x_n) \propto \exp \left\{ -\lambda D^2(q, q'(x_n)) \right\} \quad (2.2)$$

où $D(q, q') = \left(1 - \sum_{i=1}^{N_b} \sqrt{q(n)q'(n)} \right)^{\frac{1}{2}}$. q et q' étant des histogrammes de couleur. Ici, N_b représente le nombre de couleurs représentatives. Les figures 2.1 et 2.2 représentent deux histogrammes de la même image mais avec des nombres de couleurs représentatives différents.

FIGURE 2.1 – Histogramme de couleur pour $N_b = 10$ FIGURE 2.2 – Histogramme de couleur pour $N_b = 50$

On observe logiquement une plus grande précision pour un N_b plus grand. Cependant, le temps de calcul est directement impacté par le nombre de couleurs représentatives, ce qui explique en partie pourquoi nous n'utilisons pas $N_b = 255$.

Le modèle d'état utilisé ici est pertinent dans le sens où l'on ne possède pas d'information *a priori* sur le mouvement de la cible. Ici, on fait l'hypothèse la plus simple possible qui est de supposer que la position de la cible à l'instant $n + 1$ est sa position à l'instant n , plus une certaine variation qui suit une loi gaussienne. C'est un modèle qui peut présenter de bons résultats mais qui peut s'avérer très limité dans certains cas particuliers, notamment lorsqu'il y a des mouvements brusques. Aussi, les paramètres tels que le nombre de particules, le nombre de couleurs représentatives, les variances ou encore le paramètre λ de la loi exponentielle, doivent très souvent être adaptés au cas d'étude considéré. Étudions leur influence.

2.1.1 Influence du paramètre λ

Les figures 2.3 et 2.4 nous montrent l'influence du paramètre sur le nombre effectif de particules. N_{eff} nous donne un aperçu de la diversité des échantillons de particules. En effet, la vraisemblance est proportionnelle à une exponentielle décroissante en

λ . Or, $N_{eff} = \frac{1}{\sum_{i=1}^N (w_n^i)^2} \propto g_n(y_n | x_n)$, ie plus λ est important, plus g_n diminue, plus N_{eff} est faible, donc moins il y a de diversité des échantillons. C'est ce que l'on observe par l'expérience.

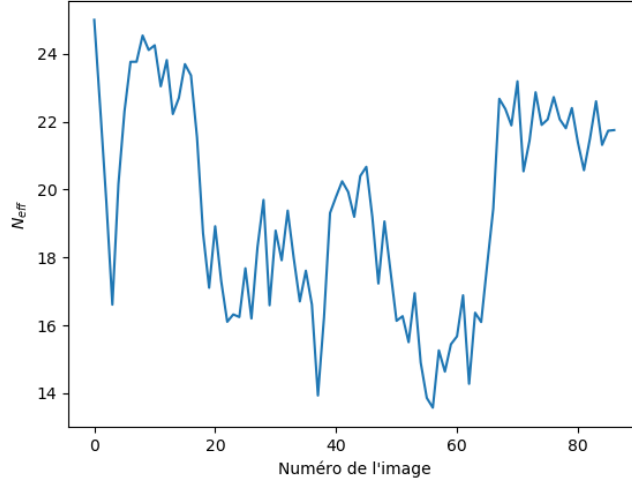


FIGURE 2.3 – Nombre effectif de particules pour $\lambda = 5$ (séquence 3)

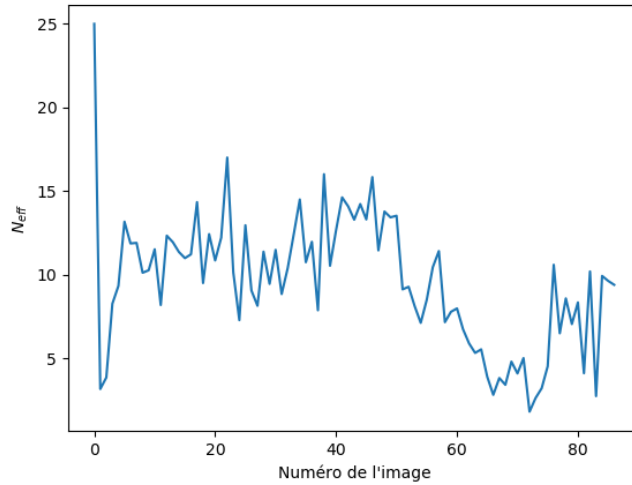


FIGURE 2.4 – Nombre effectif de particules pour $\lambda = 20$ (séquence 3)

Dans certains cas, il peut être intéressant d'avoir un peu moins de diversification dans les particules afin de se "concentrer" sur la cible et non sur l'environnement

autour (comme pour la séquence du surfeur). En effet, plus λ est faible, plus N_{eff} est important et se rapproche de N et donc plus la distribution des poids se rapproche d'une distribution uniforme, ce qui n'est pas ce que l'on recherche.

2.1.2 Influence des variances

C_1 et C_2 représentent respectivement la variance en *abscisse* et la variance en *ordonnée*. Les figures 2.5 à 2.8 illustrent l'impact des variances sur le suivi de visage. On remarque que pour des mouvements horizontaux, il convient d'augmenter C_1 et que pour des mouvements verticaux, il convient d'augmenter C_2 . En effet, en augmentant la variance pour un axe, on "élargit" la fonction de vraisemblance, c'est-à-dire que les probabilités pour que les particules s'éloignent de la position moyenne (sur un axe) augmentent. Ainsi, si la cible accélère brusquement, dans un cas de variance faible, il y aura moins de particules autour de la position moyenne et il y aura une faible probabilité qu'une particule se trouve au même endroit que le visage. De fait, il sera moins probable d'avoir un histogramme de couleur "proche" de l'histogramme de référence (puisqu'il sera compliqué d'avoir un rectangle sur la cible) et donc le suivi ne sera pas bon. Pour pallier au problème d'accélération brusque, il est donc nécessaire d'augmenter la variance.

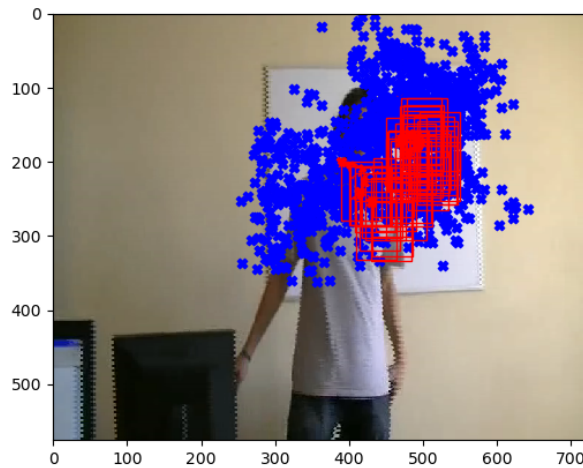


FIGURE 2.5 – Suivi de la personne sur la séquence 1 ($N = 25$, $C_1 = C_2 = 300$, $N_b = 10$, $\lambda = 5$)

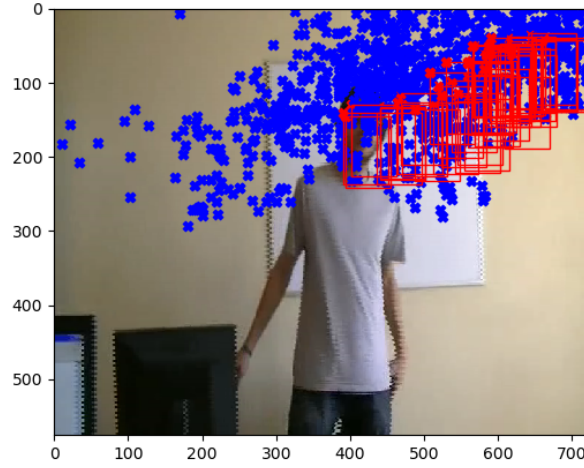


FIGURE 2.6 – Suivi de la personne sur la séquence 1 ($N = 25$, $C_1 = 2000$, $C_2 = 300$, $N_b = 10$, $\lambda = 5$)

Entre les figures 2.5 et 2.6, on a augmenté C_1 . On remarque que les rectangles rendent mieux compte du déplacement de la personne puisque celle-ci se déplace horizontalement sur la séquence 1.

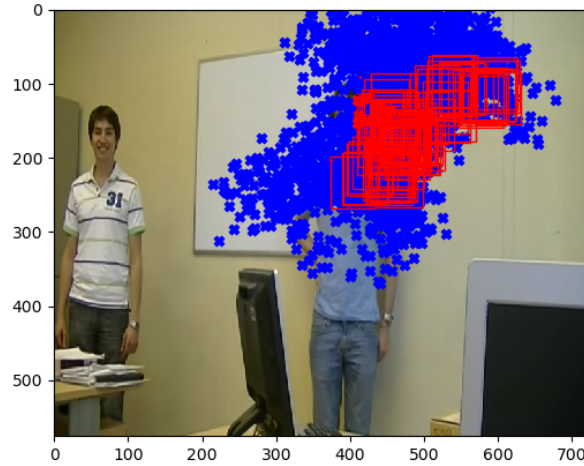


FIGURE 2.7 – Suivi de la personne en blanc sur la séquence 3 ($N = 25$, $C_1 = C_2 = 300$, $N_b = 10$, $\lambda = 5$)

La figure 2.7 nous montre le problème du croisement de cibles. En effet, la personne en blanc se déplace de droite à gauche en passant devant la personne en bleu. Sur

cette figure, les premiers rectangles (en haut à droite) suivent correctement la personne en blanc puis lorsqu'il y a croisement, les rectangles se mettent à suivre la personne en bleu. Ceci s'explique par le fait que lorsque la personne en blanc accélère horizontalement vers la gauche de l'image, la variance n'est pas assez grande pour pouvoir assurer un bon suivi. Comme les visages des personnes présentent des histogrammes de couleur similaires et que les particules tirées sont majoritairement autour de la personne en bleu, l'algorithme "confond" les deux et se met donc à suivre la personne en bleu. Pour pallier au problème du croisement de cibles, il faudrait augmenter la variance selon l'axe considéré, ainsi que le nombre de couleurs représentatives.

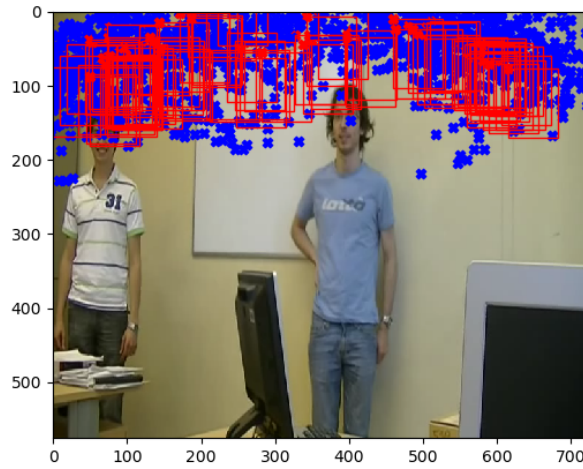


FIGURE 2.8 – Suivi de la personne en blanc sur la séquence 3 ($N = 25$, $C_1 = 2000$, $C_2 = 300$, $N_b = 20$, $\lambda = 5$)

Sur la figure 2.8, on a augmenté C_1 ainsi que N_b pour bien différencier les deux visages lors du croisement. Les résultats du suivi du visage de la personne en blanc sont encourageants.

L'algorithme que nous utilisons dans ce TP diffère des algorithmes de reconnaissance implémentés dans les appareils numériques, qui eux utilisent le *deep learning* : certains points du visage (nez, bouche, yeux...) sont analysés puis sont extraites dif-

férentes caractéristiques comme des distances ou des tailles, qui vont finalement être comparées à des bases de données.

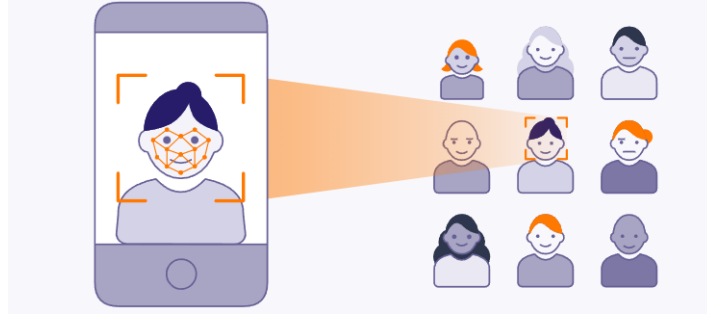


FIGURE 2.9 – Algorithme de reconnaissance faciale s'appuyant sur du *deep learning*

2.2 Suivi 3D

Dans cette partie, on s'intéresse aux effets de zoom sur le suivi de la cible. Pour cela on intègre un troisième paramètre au vecteur d'état X_n : le paramètre d'échelle, de variance C_3 . Il est à noter que pour un C_3 trop faible, le paramètre d'échelle est parfois trop faible conduisant l'algorithme à utiliser des rectangles trop petits, ce qui conduit à l'arrêt du programme.

On souhaite étudier l'influence de la diversification des particules. Les figures 2.10 et 2.11 présentent un suivi du surfeur pour respectivement $\lambda = 5$ et $\lambda = 30$.

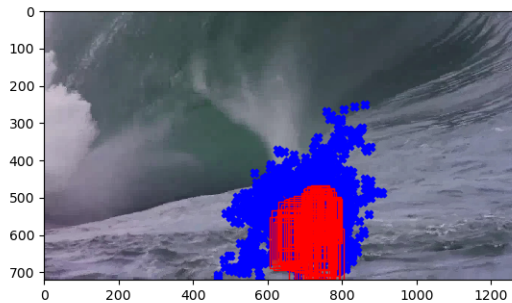


FIGURE 2.10 – Suivi du surfeur ($N = 25$, $C_1 = C_2 = 300$, $C_3 = 0.02$, $N_b = 10$, $\lambda = 5$)

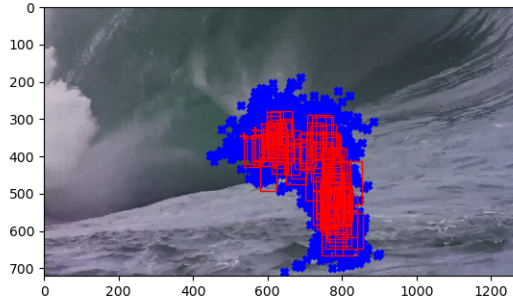


FIGURE 2.11 – Suivi du surfeur ($N = 25$, $C_1 = C_2 = 300$, $C_3 = 0.02$, $N_b = 10$, $\lambda = 30$)

Sur cette séquence, le surfeur va de droite à gauche, puis passe derrière un rouleau. Cette séquence est filmée de loin ce qui explique la présence d'un zoom, que l'on peut voir à travers la taille des rectangles qui diffère selon que la cible est grossie ou non. On observe que pour $\lambda = 5$ le suivi n'est pas fidèle à la réalité, contrairement à $\lambda = 30$. Avec les paramètres choisis, le programme arrive relativement bien à suivre le surfeur, même lorsque celui-ci passe derrière la vague. On peut aussi regarder si l'ajout du paramètre d'échelle est vraiment utile pour le suivi de cible en trois dimensions. La figure 2.12 montre le suivi du surfeur réalisé sans l'ajout du paramètre d'échelle au modèle : le paramètre améliore bel et bien le suivi.

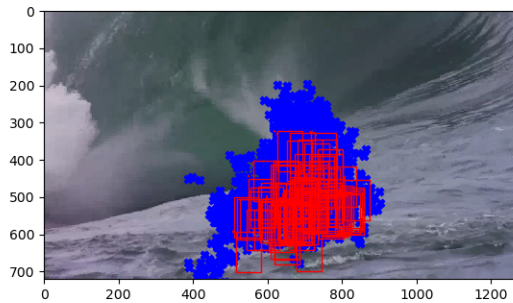


FIGURE 2.12 – Suivi du surfeur ($N = 25$, $C_1 = C_2 = 300$, $N_b = 10$, $\lambda = 30$)