

Poursuite de cible

Vision réalité virtuelle et augmentée

Emery Bourget-vecchio

16/11/2021

Le but de ce projet est de réaliser une poursuite de cible sur une séquence d'image.

Pour ce faire, je me suis servi de python et d'openCV pour lire et créer une vidéo à partir de la séquence d'image. Le code et les résultats vidéos sont disponibles en annexe.

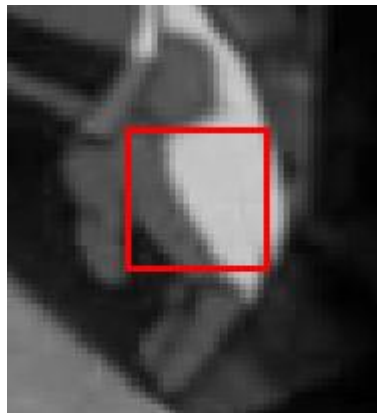
Méthode de mise en correspondance de motif

Le principe de la mise en correspondance est de mesurer la similarité entre deux motifs. Nous allons voir ici deux méthodes : distances statistiques et corrélation.

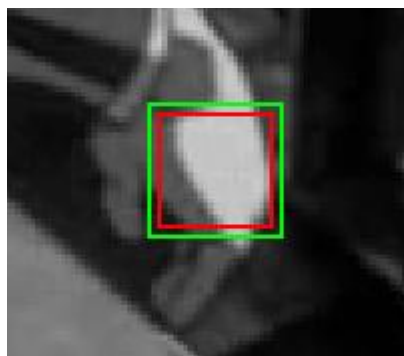
La distance statistique cherche à trouver le motif avec le moins de différence alors que la corrélation cherche à trouver le motif avec le plus de ressemblance.

Comme les deux algorithmes fonctionnent quasiment sur le même principe, je vais d'abord parler des étapes qui sont communes aux deux.

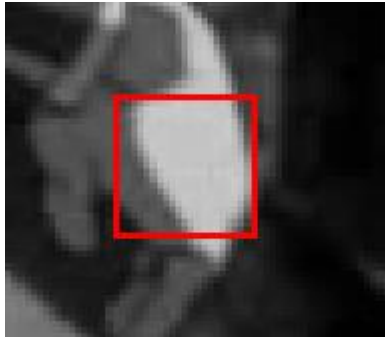
Dans un premier temps, j'ai tracé un rectangle rouge pour délimiter le motif.



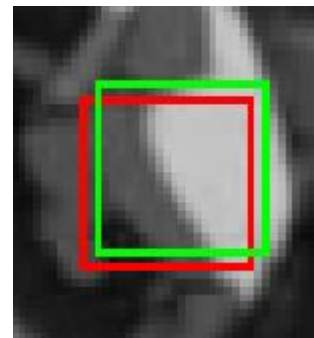
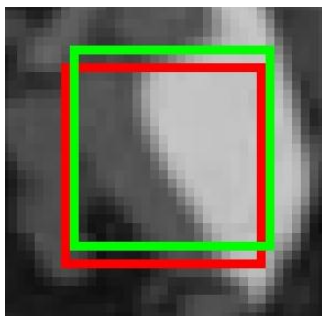
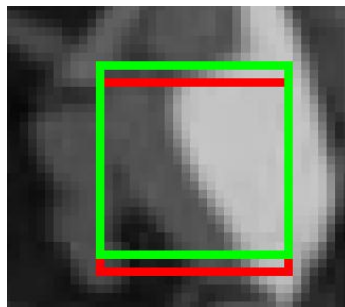
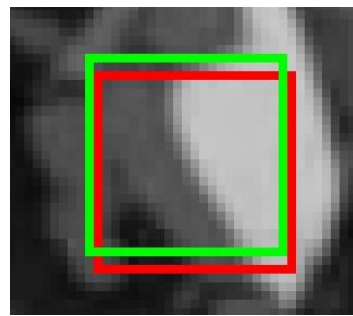
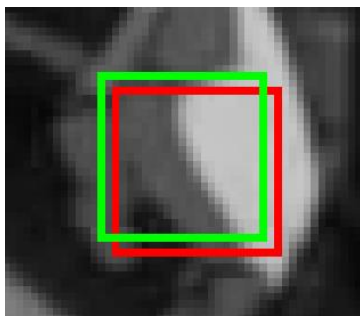
Puis j'ai créé une nouvelle zone qui sera une zone de recherche pour trouver le motif sur la prochaine image. Ici la zone de recherche est de 2.



Je récupère ensuite le motif de l'image suivante placée aux mêmes coordonnées que sur l'image 1.



Puis je cherche pour chaque cas possible dans la zone de recherche, la valeur qui minimise le SSD.
Voici un exemple de quelques cas qui va être testé.



1 Distance

Avec la distance statistique, il va être possible de mesurer la distance entre les deux motifs.

Ainsi, pour chacun des cas, le motif vert va être comparé avec le motif de base rouge de la première image avec la formule des écarts quadratiques normalisés centrée.

$$SSD = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} |I_1(x, y) - I_2(x, y)|^2$$

On va garder le motif vert qui minimise la distance avec le motif rouge de base. Cela va nous permettre de récupérer un décalage u et v qui correspond à la modification à effectuer sur le motif rouge de l'image courante pour obtenir le motif rouge de l'image suivante. On va donc appliquer la transformation $-u, -v$ sur l'image suivante pour que les motifs rouges de l'image n et $n+1$ focus la même zone.

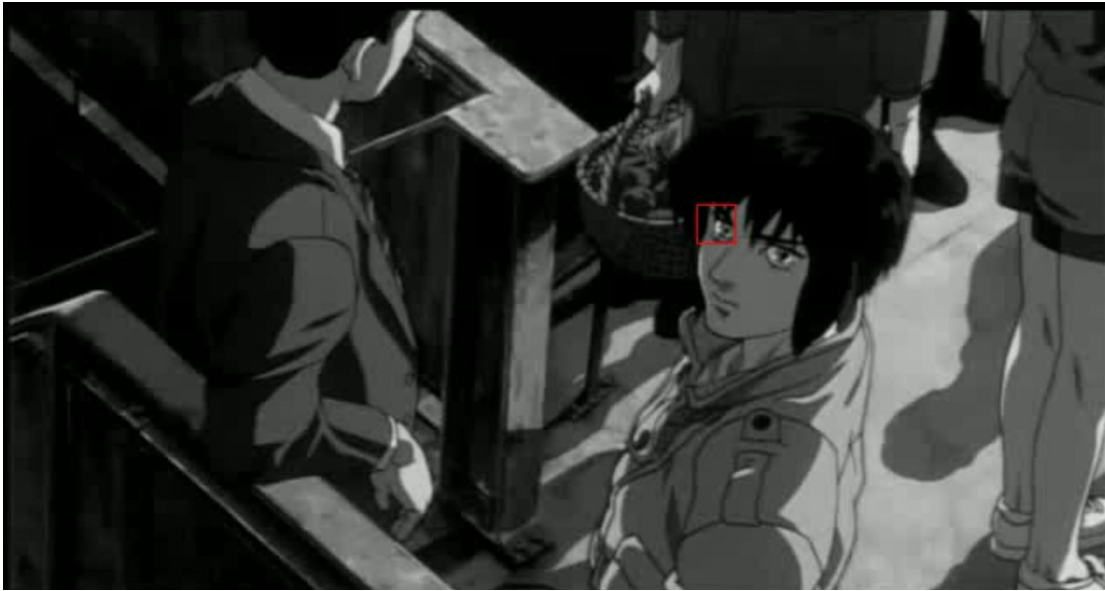
Ainsi on remarque dans la vidéo « Poursuite de cible main » le bon fonctionnement de l'algorithme. Dans l'exemple ci-dessous, on peut voir que le carré rouge se trouve à peu près à la même position entre la première image et la dernière.



La poursuite de cible fonctionne plutôt bien dans la vidéo « **Poursuite de cible nez** », le motif finit par dévier à la fin, ce qui est normal, car le contour du nez disparaît au fur et à mesure de la vidéo ce qui déplace le motif sur la bouche pour combler le manque de couleur sombre qu'apportaient les contours du nez.



Enfin, dans la vidéo « poursuite de cible œil », la poursuite de cible se passe plutôt mal, car il y a trop de variation de couleur lorsque le personnage tourne la tête. Le motif va ainsi décrocher.



On peut donc conclure sur cette partie que l'algorithme des distances fonctionne plutôt bien s'il n'y a pas trop de variation d'intensité lumineuse dans le motif. Dans le cas du motif se trouvant sur la main, il n'y a quasiment pas de variation donc la poursuite de cible se réalise très bien. Cependant lorsque la zone présente plus de variation lumineuse, la poursuite de cible est plus ou moins erronée comme dans le cas de la bouche ou de l'œil du personnage.

2 Corrélation

La corrélation consiste à rechercher, sur l'image courante, le sous-motif dont la corrélation est maximale avec le motif à poursuivre.

Ainsi, pour chacun des cas, le motif vert va être comparé avec le motif rouge de l'image suivante avec la formule de coefficient de corrélation de Pearson.

$$\rho = \frac{1}{N_x N_y} \frac{\sum_x \sum_y (I_1(x, y) - \mu_1)(I_2(x, y) - \mu_2)}{\sigma_1 \sigma_2},$$

On va garder le motif vert qui maxime la corrélation avec le motif rouge de l'image suivante. Cela va nous permettre de récupérer un décalage u et v qui correspond à la modification à effectuer sur le motif rouge de l'image courante pour obtenir le motif rouge de l'image suivante. Comme pour la distance, on va ainsi appliquer le décalage $-u, -v$ sur le motif suivant pour que le motif de base et le motif suivant focus la même position.

Pour comparer avec l'algorithme des distances statistique, nous allons prendre les mêmes motifs en exemple.

Dans le cas de la poursuite de la main, on remarque encore une fois que la poursuite fonctionne vraiment bien, encore mieux qu'avec la méthode des distances, car les deux focus presque qu'exactement la même partie de la main.



Pour le nez, on a également le même résultat qu'avec la méthode des distances statistique, c'est-à-dire un bon suivi même si à la fin le motif décroche un peu du modèle du motif original. Encore une fois pour les mêmes raisons, c'est-à-dire la disparition du nez qui entraîne un décalage du motif vers la bouche.



Enfin pour les yeux, nous avons également à peu près le même résultat qu'avec la distance statistique. Le motif décroche assez vite, car il y a beaucoup de variation de couleur.



3 Variation de l'espace de recherche

Dans le dossier « Résultat », on retrouve des exemples de résultat pour la corrélation et la distance. Dans chacun des deux, on retrouve plusieurs dossiers montrant la poursuite de cible avec différent niveau de zone de recherche (i.e 2,3 et 10). Dans les exemples ci-dessus, j'ai utilisé une zone de recherche de 2. On remarque qu'avec une zone de recherche de 3, la poursuite de cible est relativement équivalente à celle avec 2. Cependant lorsque l'on prend une zone de recherche de 10, dans aucun cas la poursuite de cible ne se passe bien, ce qui est normal, car il est plus probable qu'une zone qui n'est du tout la bonne soit considérée adéquate du fait que la valeur de la distance ou de la corrélation est respectivement faible ou élevée.

Voici un exemple avec une zone de recherche de 10 pour la corrélation.



4 Motif fixe et motif mobile

Dans l'algorithme, j'ai considéré que le motif de base est fixe, c'est-à-dire qu'au fur et à mesure des images, je compare mon motif de l'image n avec mon motif à l'image 1. Cela permet d'avoir une grande stabilité pour les séquences d'image sans trop de mouvement. Parfait pour juste du défilement d'image sans mouvement dans la scène.

Cependant on peut imaginer également faire varier le motif de base en l'assignant au motif de son image suivante. Ainsi cela va permettre de mieux gérer le mouvement de personnage, les variations de lumière, etc.

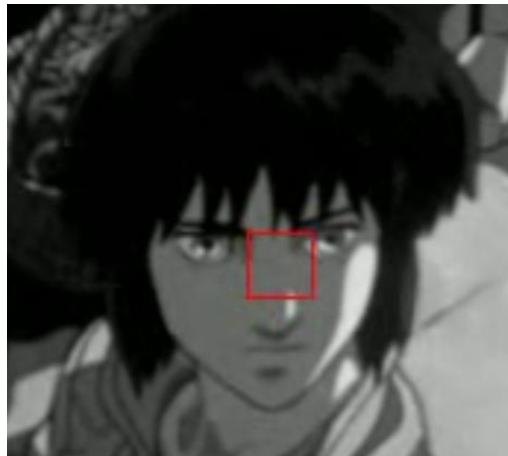
Pour illustrer mes propos, voici deux résultats de poursuite de cible avec pour le premier un motif fixe (déjà vu au-dessus) et pour le deuxième un motif mobile.



Image au frame 1



Dernier frame avec motif fixe

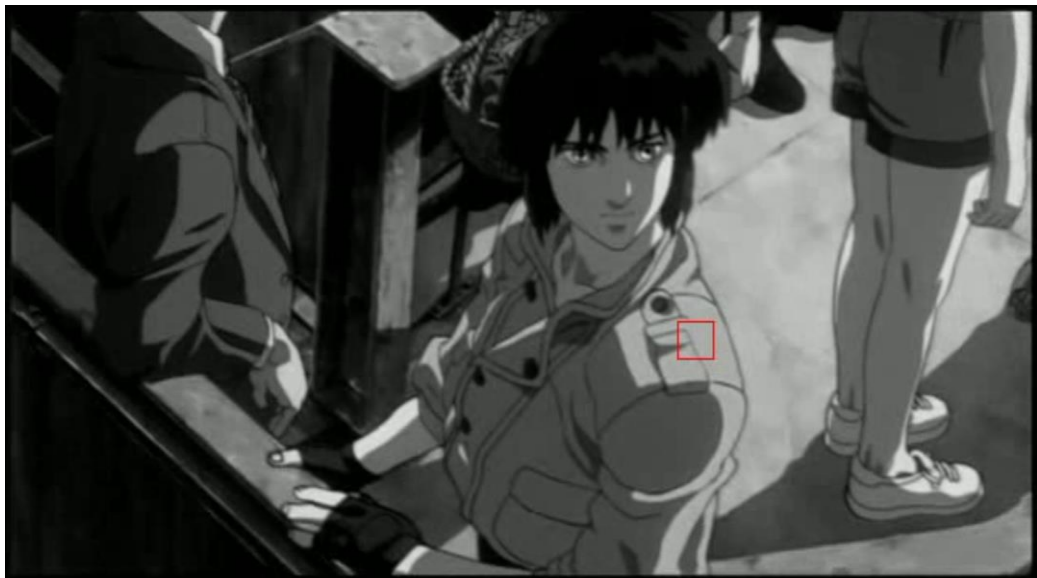


Dernier frame avec motif mobile

Comme dit au-dessus, avec le motif fixe on se retrouve à la dernière image au niveau de la bouche, ce qui est normal, car le nez du personnage disparaît et donc le motif cherche au plus proche pour récupérer les pixels sombre du nez manquant.

Avec le motif mobile, on voit qu'on reste parfaitement au niveau du centre du visage alors que la caméra défile et que la tête du personnage tourne. En effet, comme le motif de base s'actualise à chaque image, le motif a le temps de s'adapter et d'assimiler que des nuances de l'image disparaissent (ici le nez).

Cette technique a cependant des limites lorsque les variations sont trop importantes.



Ici la variation de lumières est trop importante, et on remarque qu'on finit par décrocher du motif original.