



**POLYTECHNIQUE
MONTRÉAL**

**UNIVERSITÉ
D'INGÉNIERIE**

INF8770 – Technologies multimédias

Hiver 2023

Travail pratique 3

Groupe 2

2073519 – Charles-Antoine Vézina

2017113 - Yun Ji Liao

17 Avril 2023

Question 1 (/3)

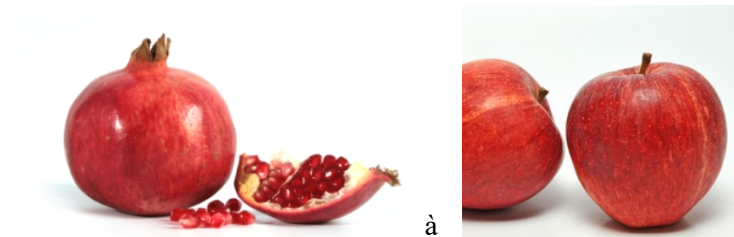
Formulez au moins TROIS hypothèses à tester pour caractériser la performance des deux méthodes. Justifiez vos différentes hypothèses par le fonctionnement théorique des méthodes.

Exemple : La méthode 1 sera plus précise que la méthode 2 si les deux plans...

Hypothèse 1 :

La prise de vues par la décomposition par l'histogramme est moins précise si les deux plans ont des couleurs similaires, mais réparties différemment dans l'image. En effet, l'histogramme collecte les couleurs RGB des images et les met l'ensemble dans les histogrammes afin de les comparer. Le problème est le fait qu'il serait possible d'avoir les deux plans qui ont le même histogramme de couleur, mais qui ne sont totalement pas pareils. Vu qu'un vidéo en noir et blanc a seulement deux couleurs à identifier, il serait plus possible d'arriver que deux scènes différentes aient un histogramme semblable.

Illustration:



Hypothèse 2 :

Puisqu'un problème de la comparaison des arêtes est la définition de seuil, deux plans qui ont des contrastes variés dans des régions affectent la précision de cette méthode. En effet, les pixels passent par un filtre déterminé. Puis, les calculs donnent les gradients en x et en y de chaque pixel, ce qui permet de définir les arêtes selon un seuil déterminé. Donc, il peut y avoir des arêtes non détectées, très pâles et des arêtes très foncées. Bref, si les contrastes sont très différents dans les zones différentes dans les deux plans, la méthode d'histogramme est plus précise.

Illustration:



Hypothèse 3 :

La prise de vue par histogramme sera plus précise pour les fondus lors d'un changement de couleur. Si la même image dans un vidéo fait un ou des changements de couleur, la méthode histogramme devient moins précise que la méthode des arêtes parce que cette dernière ne détectera seulement les arêtes, alors que la

décomposition par histogramme détectera que c'est une image différente à cause de la différence de l'apparition des couleurs différentes, ce qui affectera la précision des identifications des coupures et des fondus enchaînés de ce type de vidéo.

Illustration:



Hypothèse 4 :

La méthode par histogramme est plus précise que l'autre méthode s'il y a une fondue enchaînée dans les deux plans. En effet, une solution au problème des gradations et fondus chaînés est déjà existante dans la méthode par histogramme. Il suffit d'examiner les trames selon les critères avec deux seuils déterminés. Cependant, il est difficile de l'identifier par la méthode des comparaisons des arêtes parce qu'il y aura des régions confondues des deux plans et des régions identiques. La méthode s'appuie sur la détection des changements brusques d'intensité des pixels pour identifier les bords/contours.

Hypothèse 5 :

La méthode par histogramme sera plus rapide pour récupérer les coupures et les fondues enchaînées que la méthode par décomposition des arêtes. En effet, il est plus rapide de calculer l'histogramme de couleurs de chaque image que de calculer les arêtes entrantes et sortantes entre 2 images. La décomposition par histogramme va donc nécessiter moins de calcul et moins de mémoire pour obtenir ses résultats.

Question 2 (/3)

Décrivez en détail les expériences que vous allez réaliser pour vérifier les hypothèses formulées (vidéos ou extraits de vidéos utilisés, métrique d'évaluation de la performance, etc.).

Tout d'abord, nous voulons tester les vidéos fournies pour ce TP. Nous allons créer un algorithme pour chacune des deux méthodes. Ces algorithmes permettent d'afficher les résultats des vidéos, afin d'analyser les différences entre les méthodes quant à identifier les coupures et les fondus enchaînés. La visualisation permet aussi d'évaluer ces méthodes. Les résultats seront les différences entre les histogrammes de couleurs par la distance euclidienne pour la méthode d'histogramme et la valeur de rho pour la méthode par décomposition des arêtes.

Pour comparer les résultats, nous allons comparer la distance euclidienne obtenue par la méthode des histogrammes avec les valeurs de rho obtenue par la méthode des arêtes. Il nous sera donc plus facile de visualiser ces résultats de cette façon, comme cela, nous pourrions proposer des valeurs de t_s et t_b , ou des valeurs de ρ pour les coupures de rho, puisque ces valeurs sont souvent propres à chaque vidéo.

Question 3 (/3)

Décrivez les codes informatiques utilisés pour réaliser les expériences et donnez les sources. Décrivez, si applicable, comment vous avez adapté les codes informatiques pour réaliser les expériences décrites à la question 2.

Nous avons découpé le code en trois fichiers distincts. Le premier transforme la vidéo en une série d'images, le deuxième effectue la décomposition par histogramme et le troisième effectue la décomposition par arête. La majorité des fonctions utilisées proviennent de la librairie OpenCV. Voici les méthodes utilisées pour chaque fichier.

Pour la transformation de la vidéo, nous avons utilisé la méthode VideoCapture(path) afin d'ouvrir la vidéo dans l'environnement de développement. Ensuite, en faisant une boucle While infinie, nous avons ouvert chaque trame de la vidéo et l'avons enregistrée en tant qu'image avec la fonction imwrite.

Pour la décomposition en histogramme, nous avons fait la lecture de chaque image avec imread et avons fait la conversion en gris avec cvtColor. Ensuite, nous avons calculé l'histogramme de chaque trame avec calcHist, en mentionnant la méthode que nous voulions utiliser. Selon nos tests, le paramètre qui donnait les meilleurs résultats était la corrélation (HISTCMP_CORREL). Nous avons enfin normalisé ces histogrammes avec la fonction norm et le paramètre NORM_L2 pour avoir la distance euclidienne.

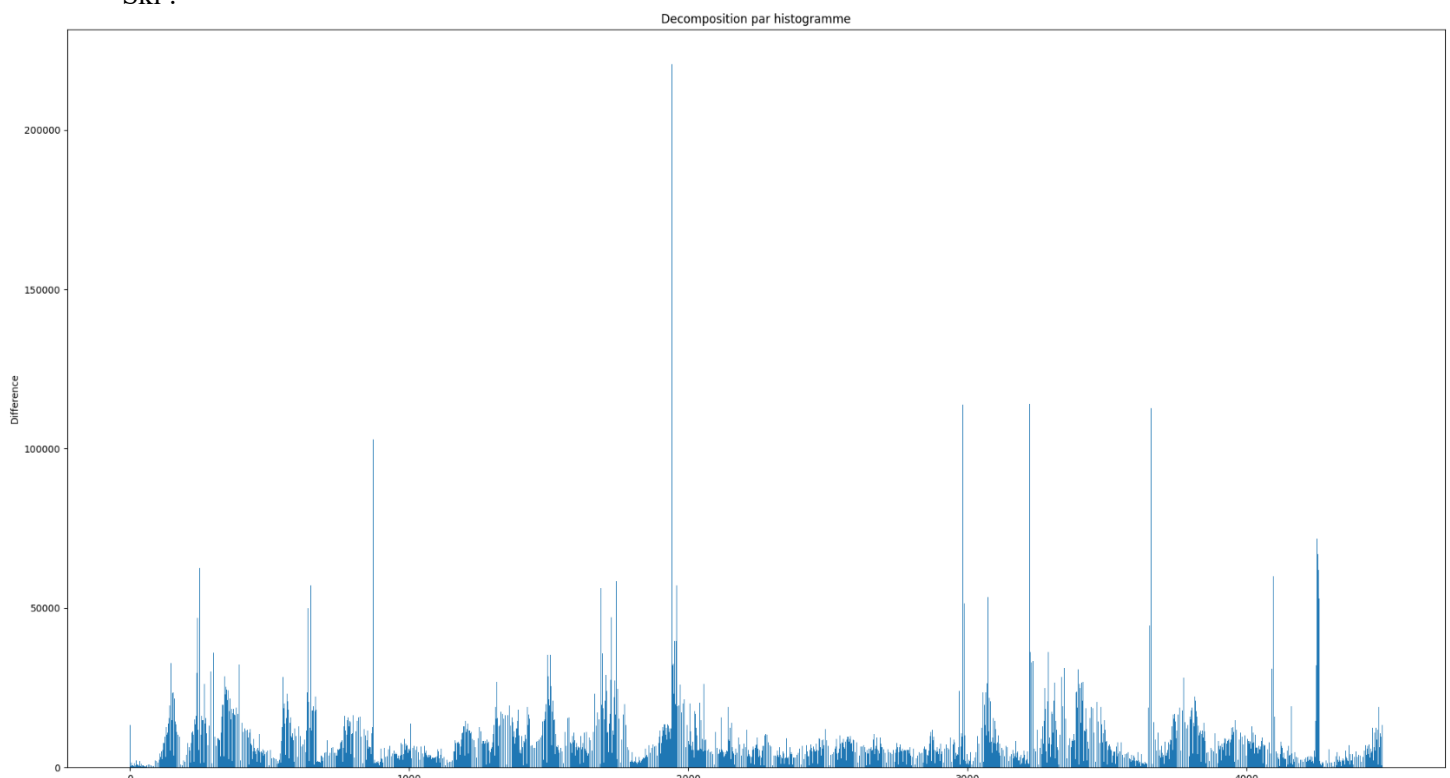
Pour la décomposition par arête, nous avons fait la lecture et la conversion des images de la façon précédente. Nous avons ensuite fait le floue et la détection des arêtes avec les méthodes GaussianBlur et Canny. Nous avons découvert que la méthode Canny était plus précise que la méthode Sobel. Enfin, nous avons fait une dilatation des arêtes afin d'obtenir des arêtes plus importantes.

Source : <https://www.geeksforgeeks.org/extract-images-from-video-in-python/>

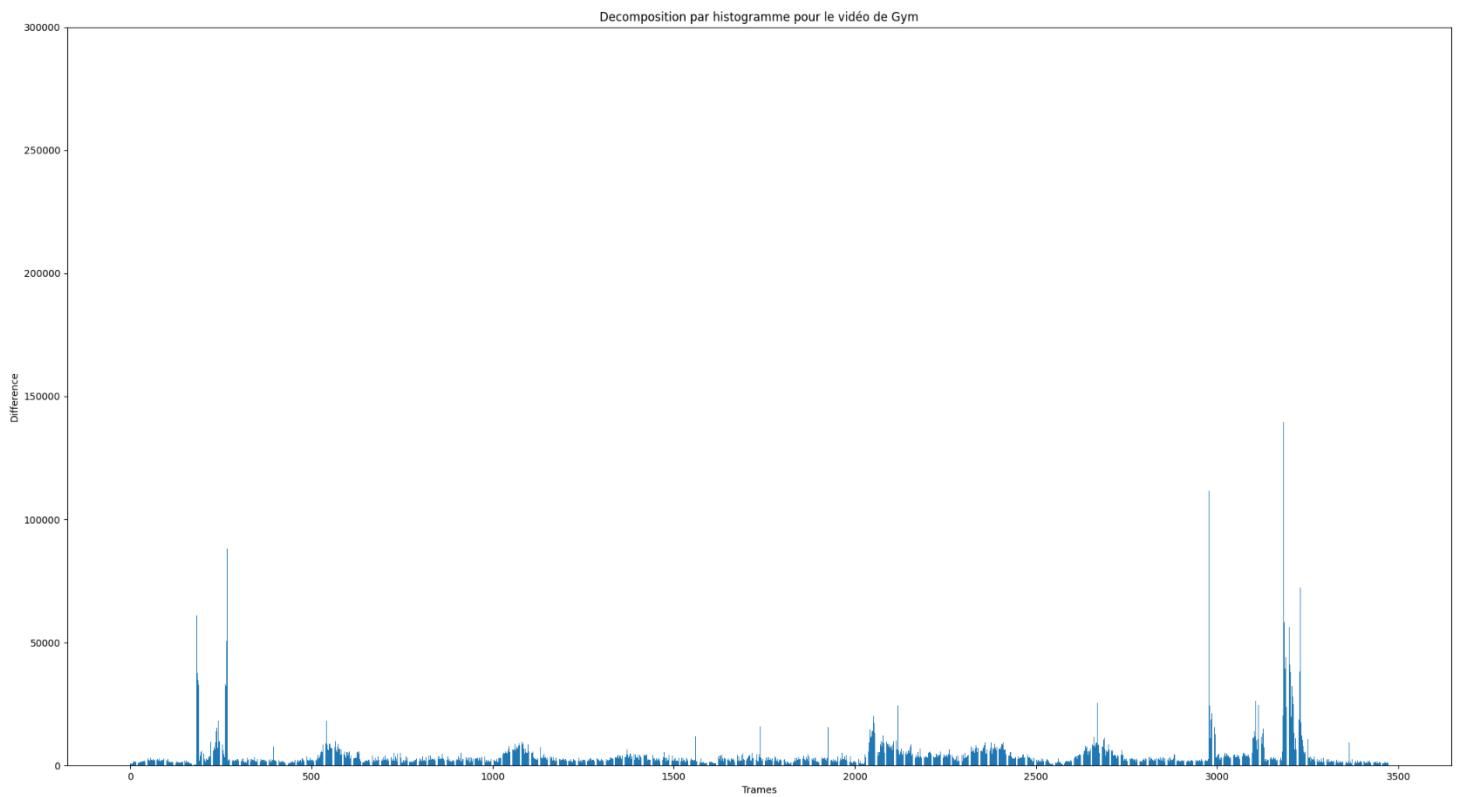
Question 4 (/3) Donnez les résultats obtenus pour les expériences décrites à la question 2 sous un format approprié (figures, tableaux...).

Décomposition par histogramme :

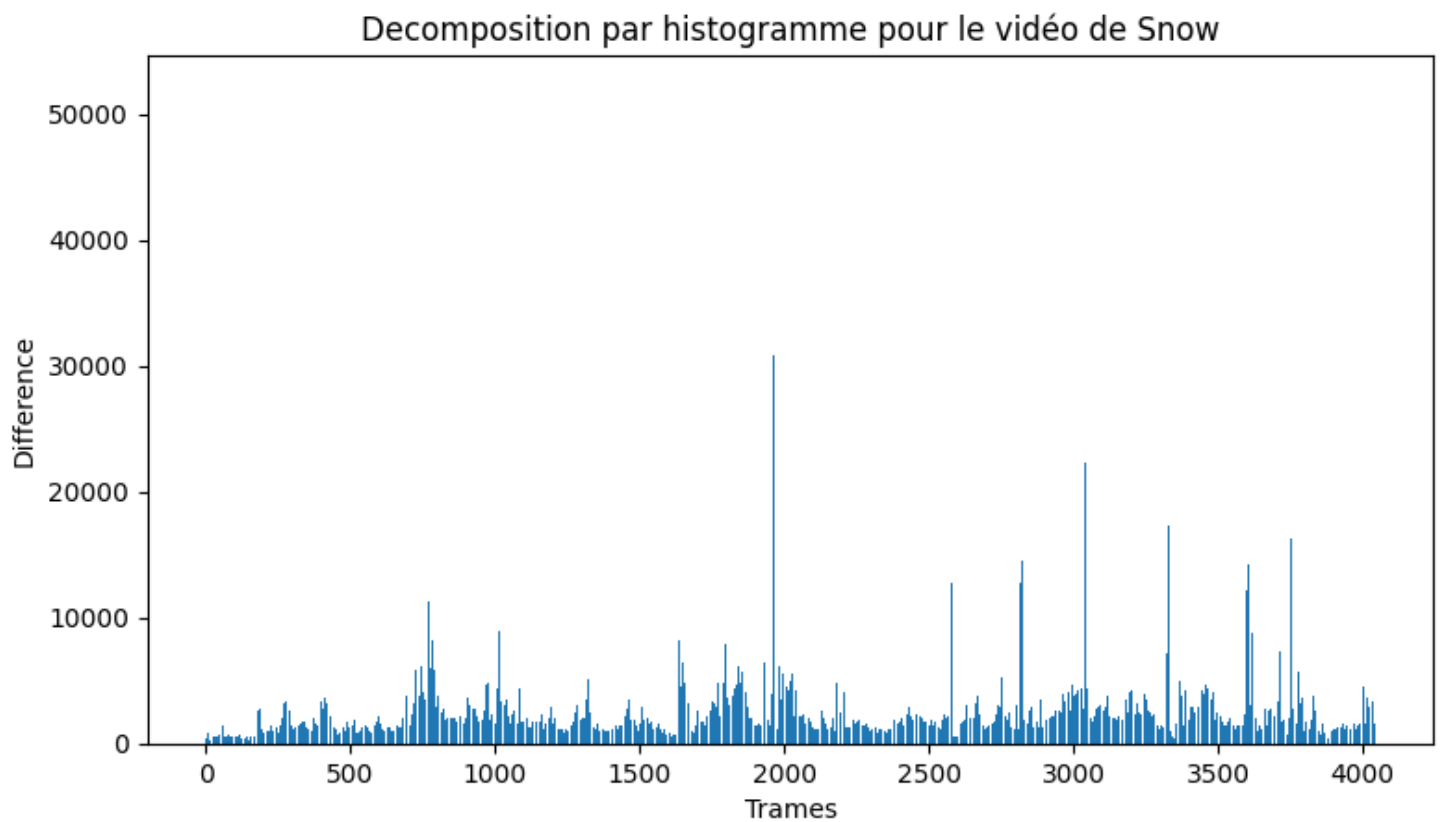
Ski :



Gym :

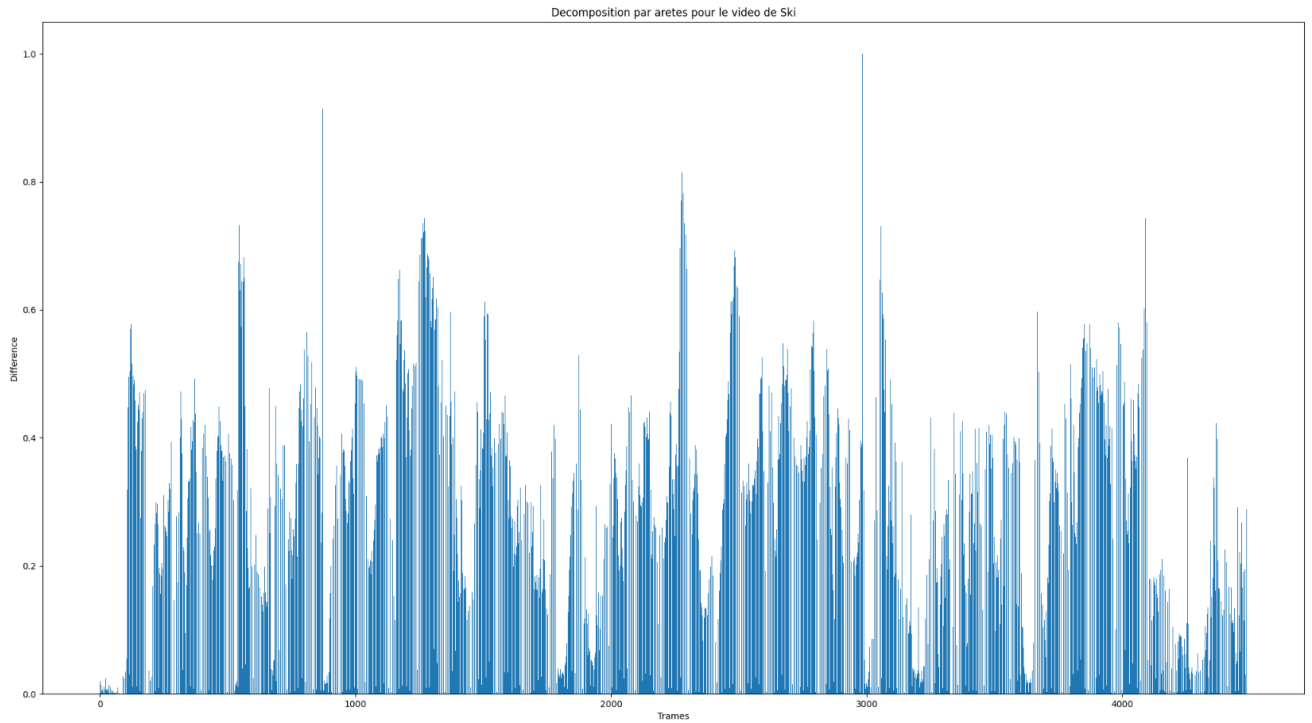


Snow :

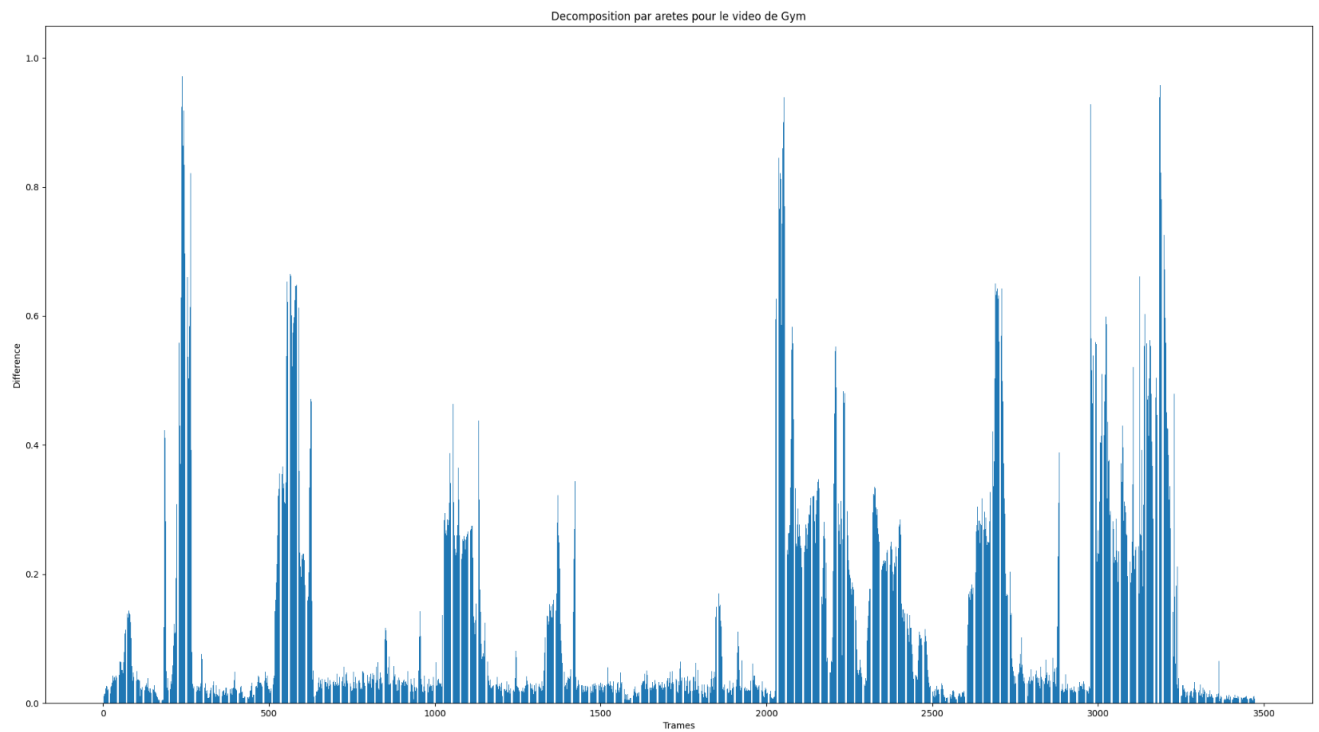


Décomposition par arêtes :

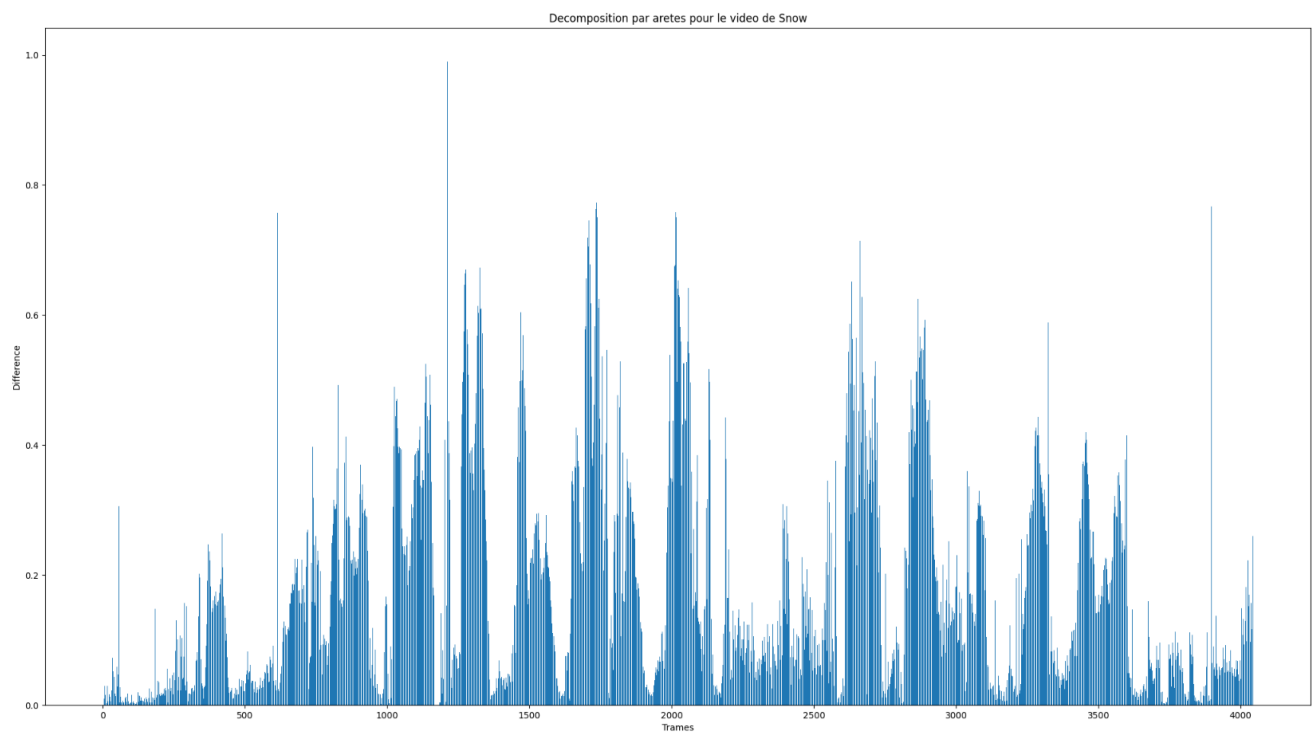
Ski :



Gym :



Snow :



Temps d'exécution :

Temps d'exécution	Ski	Gym	Snow
Histogramme	31.17	24.14	5.97
Arête	39.09	26.99	4.09

Question 5 (/8, 4 points par qualité)

Analysez les résultats obtenus et mettez-les en relation avec les hypothèses de la question 1. Est-ce que les hypothèses sont supportées par les résultats ? Quels sont les tests qui ne permettent pas de conclure ? Comment pourrait-on les améliorer ?

Il faut noter que l'algorithme par décomposition par arêtes donne parfois des résultats inattendus, et nous n'avons pas trouvé la source de ces comportements après avoir consulté les chargés et même le professeur. De plus, en changeant le taux de dilatation, nous pouvons arriver à des résultats complètement différents pour une même vidéo.

Pour la vidéo de Ski, si nous choisissons un tb d'environ 50 000, presque toutes les coupures sont décelées par les algorithmes. Par contre, les coupures des trames 1131 et 1171 ne sont pas découvertes. En effet, ces coupures ont été créées en changeant l'orientation de la caméra. Or, puisque les images sont très similaires, la décomposition par histogramme ne voit pas de différence entre les couleurs des images, et ne remarque

donc pas qu'il y a une coupure, contrairement à l'algorithme de décomposition par arêtes. Ceci confirme notre hypothèse 1. En se basant sur ce résultat, on peut imaginer le comportement contraire semblable à l'hypothèse 3, puisqu'on peut conclure que la décomposition par histogramme est sensible aux couleurs tandis que la décomposition par arête est sensible aux formes présentes, donc si on change la couleur des formes, la décomposition par arête ne remarquera pas de coupure alors que la décomposition par histogramme constate la différence. Ceci confirme donc l'hypothèse 3.

En ce qui concerne les fondues enchaînées, en fonction des t_b et t_s choisis, la technique par décomposition par histogramme est plus précise que la décomposition par arêtes, comme il est possible de voir dans les résultats. Effectivement, la décomposition par histogramme utilise deux seuils, t_b et t_s , et s'ils sont bien choisis, l'algorithme reconnaît mieux les fondus, alors que la décomposition par arête reconnaît quelques fondus, mais en ajoute d'autre aussi, car il est possible que les valeurs de ρ pour les arêtes entrantes et sortantes font en sorte que l'algorithme reconnaît une fondue alors qu'il n'y en a pas une. Ceci confirme l'hypothèse 4. Il faut noter par contre que l'algorithme par décomposition des histogrammes peut parfois couper une fondue en quelques morceaux, comme on peut bien voir dans la vidéo de ski. Les valeurs de différences sont bien plus grandes que t_s tout au long de la fondue, mais il arrive que certaines valeurs ne soient pas entre t_s et t_b , mais bien plus grande que t_b , ceci veut donc dire que l'algorithme peut par exemple reconnaître une fondue pendant 5 trames, ensuite 1 coupure, et ensuite une nouvelle fondue pendant 4 trames. Ainsi, au lieu d'avoir 1 fondue de 10 trames, il y a deux fondus différents et une coupure. Par contre, nous pensons que ceci est mieux que de ne pas représenter la fondue.

En raison des petits problèmes avec l'algorithme de décomposition par arête, il n'est pas possible de conclure sur l'hypothèse 2. En effet, il faudrait améliorer l'algorithme afin de pouvoir tester ce scénario.

Enfin, si on se fie à nos résultats, l'algorithme par décomposition par histogramme est généralement plus rapide que l'algorithme par décomposition par arêtes. En effet, il faut faire beaucoup d'opérations sur les images pour obtenir les arêtes (floue, détection des arêtes, dilatation des arêtes), et ensuite il faut comparer plusieurs images ensemble (image des arêtes, images des arêtes dilatées, prochaines images des arêtes, prochaine image des arêtes dilatées) afin d'obtenir une valeur pour ρ . Mais si nous nous attendions à un plus grand écart de temps entre les deux algorithmes, ceci confirme l'hypothèse 5.