



**POLYTECHNIQUE
MONTRÉAL**

**UNIVERSITÉ
D'INGÉNIERIE**

Département de génie informatique et de génie logiciel

**INF8770
Technologies multimédias**

Travail pratique 3

Hiver 2020
Section 01

Soukaina Moussaoui - 1728325
Carine Thouin - 1837383

Le 15 avril 2020

Question 1

La vidéo présentée est un vieux documentaire sur des satellites. Nous pouvons remarquer que ce montage présente quelques fondus et des coupures. Nous pouvons aussi constater qu'il y a une faible variété de mouvements et de couleurs, qui tournent souvent autour des tons de blanc, de noir et de marron. À la seconde 5, nous pouvons remarquer une coupure, car l'objet central est différent de celui de la prise de vue précédente. Néanmoins, les couleurs des deux prises de vues sont très similaires et uniformes. Ceci rendrait alors la détection de cette coupure à l'aide de l'analyse par histogrammes moins efficace. En outre, au temps 4min15s, nous pouvons remarquer un léger fondu avec le même objet central (la Lune). Seulement l'ombre sur la Lune change de position entre une prise de vue et une autre, ce qui altérerait les arêtes et les gradients. Néanmoins, l'analyse par histogrammes de couleurs serait encore induite en erreur.

Voici d'autres exemples où l'analyse par histogrammes de couleurs ne serait pas fiable. Au temps 2min48 et au temps 2min51, il existe une coupure. Étant donné que le contenu des deux prises de vues est très similaire, car il s'agit d'étoiles, les couleurs ne varient pas. Nous restons sur des tons de blanc et de noir même si quelques étoiles sont légèrement plus intenses en luminosité d'une prise de vue à une autre. En outre, entre les temps 3min58 et 4min02 nous pouvons compter dix fondus. Néanmoins, dans toutes les prises de vues, nous avons une image de cratères de couleur grise seulement. Certes, la forme des cratères change d'une prise de vue à une autre, ce qui ferait varier les gradients et arêtes, mais les couleurs restent uniformes. L'analyse par histogrammes de couleurs serait alors encore induite en erreur. Par ailleurs, entre les secondes 24 et 29, nous pouvons remarquer des zooms brusques sur la Lune, ce qui pourrait causer des variations dans l'analyse de gradients et d'arêtes et nous faire penser qu'il s'agirait de coupures (ce qui n'est pas le cas). Un autre cas où l'analyse des gradients et arêtes serait induite en erreur est au temps 6min49 où la caméra suit le satellite qui se dirige de manière très rapide vers Saturne, entre quasiment en impact et reprend son envol. Au moment du soi-disant impact, la caméra montre pendant une moitié de seconde un fond complètement obscur et suit le satellite encore pendant son envol. Ici, avec l'analyse par gradients et arêtes, nous pourrions identifier une coupure étant donnée la grande vitesse à laquelle la caméra suit le satellite et d'un coup change de perspective après le soi-disant impact.

De manière générale, la vidéo présente des coupures très évidentes étant donné que d'une prise de vue à une autre, nous avons un paysage très différent : ce ne sont pas les mêmes objets ni la même distribution de couleurs. Néanmoins, comme nous l'avons mentionné précédemment, il y a quelques cas de coupures et de fondus qui ne pourront pas être facilement détectables à l'aide de l'analyse par histogrammes étant donnée l'uniformité des couleurs d'une prise de vue à une autre. Il y aura aussi des risques de faux positifs à l'aide de l'analyse par gradients et arêtes quand la caméra bouge de manière rapide et change rapidement de perspective. Ces cas sont très moindres dans la vidéo présentée mais représentent quand même un défi.

En bref, en ayant analysé ces possibles difficultés, nous pouvons affirmer qu'il faudra utiliser plus d'une technique d'analyse de détection de coupures et fondus afin d'atténuer les risques d'erreurs.

Question 2

Après avoir analysé précédemment les possibles difficultés de détection de coupures et fondus, nous avons fait le choix d'utiliser deux méthodes de décomposition de prises de vues. Étant donné que la vidéo présentée contient des séquences pouvant représenter une difficulté de détection de coupures à l'aide de comparaison de couleurs ou de contours, nous avons opté pour une solution qui combine l'analyse des des gradients et des arêtes et l'analyse par histogrammes de couleurs.

La décomposition en prises de vues par histogrammes de couleurs consiste à extraire les valeurs RGB pour chaque pixel d'une trame. Nous vérifierons la présence d'effets de caméra en faisant la différence d'histogrammes entre une trame et celle qui la précède selon des seuils. Ces seuils seront fixés selon des critères que nous établirons. Étant donné que la majorité des prises de vues entre les coupures présentent une grande différence de couleurs et d'intensité, l'histogramme de couleurs sera efficace. Néanmoins, il existe certains cas comme ceux cités à la question 1 dans lesquels certains fondus et coupures pourraient ne pas être détectés si la différence de variations de couleurs est moindre. Nous allons alors tenir compte de cela au moment d'établir nos seuils.

Afin de palier à ces limitations, nous avons décidé d'ajouter une technique totalement indépendante des couleurs : les arêtes. Dans notre cas, nous appliquerons le filtre Sobel à chaque trame afin d'en extraire les arêtes entrantes et sortantes et ainsi, analyser leurs mouvements et trouver les coupures et fondus. Néanmoins, comme nous l'avons mentionné à la question antérieure, certaines scènes présentent des mouvements rapides, ce qui pourrait nous donner quelques faux positifs. Finalement, en ayant en main les vérités de terrain et avec quelques essais-erreurs, nous allons pouvoir déterminer les seuils pour les deux méthodes d'analyse.

L'algorithme suivant en Python fait appel à l'une des méthodes décrites ci-haut selon l'option en paramètre. De façon générale, nous avons créé deux fonctions indépendantes : l'une, nommée *methodHistogram* détecte les coupures et fondus à l'aide des histogrammes et l'autre, nommée *methodEdges*, détecte les coupures et fondus à l'aide des gradients et des arêtes. Les détails concernant ces fonctions seront décrits à la question 3.

Main.py

```
import cv2
import histograms
import edges
FILE_NAME = "anni009.mpg"

def process_video(file: str, option):
    frameNb = 0
    cap = cv2.VideoCapture(file)
    print("Methode " + str(option))
    print("Name " + file)
    while(True):
        frameNb += 1
        ret, frame = cap.read()
        if ret:
            if option == 1:
                histograms.methodHistogram(frame, frameNb)
            if option == 2:
                edges.methodEdges(frame, frameNb)

            cv2.imshow('frame', frame)
            if cv2.waitKey(1) & 0xFF == ord('q'):
                break
        else:
            break
    cap.release()
    cv2.destroyAllWindows()

process_video(FILE_NAME, 1)
```

Question 3

Comme nous l'avons mentionné à la question 2, nous avons utilisé le langage Python pour faire notre algorithme. En effet, ce langage nous offre un large éventail de bibliothèques pertinentes pour ce genre d'analyse. La première bibliothèque utilisée dans notre cas est OpenCV. Cette dernière est spécialisée pour le traitement d'images. Dans notre cas, nous l'avons employée pour lire les vidéos, découper les trames, créer les histogrammes de couleurs, comparer les histogrammes, etc. La deuxième bibliothèque utilisée est NumPy. De son vrai nom Numerical Python, cette bibliothèque est spécialisée dans le calcul scientifique et fournit des fonctions pour la manipulations d'objets à n dimensions et en algèbre linéaire. Nous avons implémenté nos deux algorithmes à l'aide des notes de cours et des informations trouvées sur le web.

Pour le premier algorithme, nous avons créé les histogrammes des images de la vidéo pour les trois canaux de couleur avec la fonction *calcHist* de OpenCV [1]. Nous comparons les histogrammes entre deux images afin d'obtenir une corrélation [2] et de pouvoir trouver les coupures et les fondus selon certains paramètres. La distance euclidienne donnait des résultats extrêmement élevés qui rendaient plus difficile notre capacité d'analyse.

Pour le deuxième algorithme, nous avons obtenu les arêtes et les arêtes dilatées en mélangeant la méthode du professeur [3] et nos recherches. L'image est d'abord convertie en gris [4] et nous y ajoutons des lignes et des colonnes additionnelles pour les calculs aux frontières. Par la suite, nous calculons nos filtres de Sobel avec la fonction *Sobel* de OpenCV [5]. Nous calculons ainsi la force des gradients et nous obtenons les arêtes. Les arêtes sont ensuite dilatées [6], afin de pouvoir calculer par la suite les arêtes entrantes et sortantes. Les formules utilisées pour calculer ces dernières viennent des notes de cours et nous utilisons NumPy afin d'exécuter les calculs. Par la suite, nous utilisons le maximum entre les deux, afin de comparer avec certains critères pour déterminer si nous avons un effet de caméra et si c'est une coupure ou un fondu.

Question 4

Les tableaux suivants présentent les différentes coupures et fondus de chacune des méthodologies implémentées. Dans le cas des coupures avec la méthode des arêtes et de fondus avec la méthode des histogrammes, étant donnés les nombreux cas de faux positifs, nous avons seulement inclus une partie dans le tableau dédié. La totalité des coupures et fondus se retrouvent dans le fichier *anni009edges.txt* et dans le fichier *anni009histo.txt*.

Tableau I : Coupures de la vérité de terrain

146	1921	2268	2440	2474	2492	2552	2604	2658	2712
2763	2819	2886	4055	4108	4227	4327	4715	5996	6134
6519	6684	6746	7062	7122	7140	7158	7177	7197	7218
7238	7255	8217	8290	8322	9160	9241	9365		

Tableau II : Coupures avec les histogrammes

400	427	432	434	439	440	465	559	1545	1658
1756	2270	2476	2494	2554	2606	2660	2714	2765	2821
3389	4110	5531	5866	6521	7160	7179	8219	8292	8324
9162	9243	9367							

Tableau III : Coupures avec les arêtes

3	152	336	342	345	645	650	654	657	668
673	681	695	705	711	722	724	731	736	740
745	752	757	770	785	800	807	815	820	830
839	850	880	906	911	915	920	924	935	940
945	950	960	964	966	975	977	980	985	991
1000	1005	1010	1015	1021	1025	1027	1031	1037	1040
1045	1047	1051	1163	1171	1173	1236	1243		

Tableau IV : Fondus avec les histogrammes

322 - 323	349 - 350	354 - 355	361 - 362	376 - 377	385 - 386
389 - 390	425 - 426	446 - 447	448 - 449	456 - 458	467 - 468
469 - 470	472 - 473	483 - 484	485 - 486	543 - 544	575 - 576
589 - 590	591 - 592	603 - 604	609 - 610	623 - 625	633 - 634
640 - 641	643 - 656	740 - 749	807 - 808	945 - 946	1224 - 1225
1231 - 1232	1257 - 1258	1501 - 1503	1520 - 1521	1538 - 1539	1548 - 1549
1552 - 1553	1573 - 1574	1576 - 1577	1590 - 1592	1608 - 1609	1613 - 1615
1619 - 1620	1621 - 1622	1650 - 1651	1663 - 1666	1677 - 1678	1680 - 1681
1692 - 1693	1716 - 1718	1733 - 1734	1753 - 1754	1771 - 1772	[...]

Tableau V : Fondus avec les arêtes

683 - 685	699 - 701	725 - 727	759 - 761	764 - 766	779 - 781
789 - 791	809 - 811	834 - 836	859 - 861	864 - 867	869 - 871
884 - 887	889 - 892	954 - 956	970 - 972	994 - 996	1054 - 1056
1062 - 1064	1254 - 1256	1308 - 1310	1497 - 1499	1777 - 1779	1797 - 1799
1802 - 1804	1807 - 1809	1847 - 1849	1852 - 1854	1874 - 1876	2052 - 2054
2072 - 2074	2102 - 2105	2271 - 2273	2443 - 2457	2495 - 2497	2528 - 2530
2568 - 2570	2673 - 2677	2703 - 2705	2748 - 2750	2758 - 2760	2804 - 2806
3027 - 3029	3117 - 3119	3142 - 3144	3203 - 3205	3225 - 3227	3468 - 3470
3478 - 3480	3518 - 3522	3617 - 3619	3734 - 3736	3779 - 3781	3794 - 3796
3831 - 3833	4129 - 4131	4144 - 4146	4154 - 4157	4169 - 4172	4174 - 4176
4224 - 4226	4339 - 4342	4369 - 4371	4389 - 4391	4404 - 4406	4409 - 4412
4429 - 4431	4434 - 4437	4457 - 4460	6010 - 6012	6140 - 6142	6160 - 6162
6165 - 6167	6185 - 6188	6341 - 6343	6526 - 6528	6546 - 6549	[...]

Question 5

Si nous nous référons aux résultats, nous pouvons remarquer qu'une grande partie des effets de caméra ont pu être détectés. Néanmoins, nous pouvons constater qu'il y a un grand nombre de faux positifs. Autrement dit, malgré le fait que nous ayons tenté de choisir des seuils qui se rapprochaient le plus de la vérité de terrain, notre algorithme a souvent identifié des effets n'ayant pas lieu. Nous pouvons aussi remarquer quelques décalages avec la vérité de terrain, soient d'environ deux trames dans le cas des coupures avec les histogrammes. Étant donné que les vérités de terrain ne sont pas exactes, nous pouvons accepter ces erreurs d'incertitude.

Si nous comparons nos résultats avec nos hypothèses à la question 1, nous avons prévu que certaines coupures auraient pu ne pas être détectées par la méthode par histogrammes. Si nous regardons les résultats, une petite partie des coupures ont été ignorées par notre algorithme dû aux faibles variations de couleurs entre les prises de vues des coupures. Comme dans le cas de la trame 146 affichant le satellite, nous pouvons constater que les images entre les deux coupures présentent bel et bien les mêmes variations chromatiques et de luminance aussi. Ce n'est alors pas surprenant que l'algorithme par histogrammes ne l'ait pas détectée. Néanmoins, nous sommes surpris que cet algorithme ait pu détecter plus de 70% des coupures compte tenu des faibles variations de couleurs entre quelques coupures. Il est vrai qu'il existe une légère erreur d'incertitude avec la vérité de terrain, mais encore une fois, cette dernière est approximative aussi.

Par ailleurs, toujours en référence à nos hypothèses à la question 1, nous avons prévu d'avoir quelques faux positifs avec la méthodologie par convolution dû à quelques cas de mouvements brusques de la caméra. Néanmoins, les résultats montrent une quantité importante de ces fausses détections, même beaucoup plus de ce que nous prévoyions. En effet, puisque les mouvements de cette vidéo à vue d'oeil ne sont pas très rapides, nous ne prévoyons pas autant de fausses détections de coupures. Néanmoins, en analysant les résultats par rapport aux images de la caméra, nous nous sommes rendues compte que notre algorithme était vraiment sensible aux mouvements. Comme dans le cas des trames 1000 à 1050, où un satellite tourne autour de lui-même, notre algorithme a détecté cela comme des coupures. Cela est dû au fait

que le satellite tourne à une certaine vitesse ce qui génère de fortes variations dans le calcul d'arêtes et de gradients.

En outre, dans le cas des fondus, même si les deux méthodologies ont détecté la plupart des fondus de la vérité de terrain, celles-ci ont donné aussi une grande quantité de faux positifs, soient d'environ de plus de 50 % par rapport à la vérité de terrain. Quand nous comparons les résultats avec les séquences de la vidéo, nous nous rendons compte que certains mouvements dans une même prise de vue font varier de manière progressive les couleurs d'où la détection de nombreux faux fondus par la méthodologie par histogrammes. Quant à la détection de fondus par convolution, celle-ci a donné moins de faux positifs que l'algorithme par histogrammes. Cette différence est due au fait que le changement progressif des couleurs ne modifie pas énormément les contours des objets ce qui ne provoque pas autant de détection de fondus par la méthodologie d'arêtes. Néanmoins, dû à la qualité des images, certains contours sont légèrement flous durant la séquence vidéo, comme dans le cas du faux fondu entre les trames 889 et 892. Dans cette séquence, la Lune est présentée en forme de croissant et un compteur à côté. Entre ces deux trames, les contours deviennent relativement flous avant de passer au zoom. Notre algorithme devient alors sensible à ce type d'effet. En outre, entre les trames 650 et 690, où il y a plusieurs fondus avec des images de cratères, notre algorithme détecte des coupures au lieu de fondus. Cela est dû à la vitesse de changement des images et des zooms rapides sur ces dernières.

En bref, malgré le fait que nous ayons été pessimistes avec la méthode par histogrammes au départ, celle-ci semble le plus se rapprocher de la vérité de terrain surtout dans le cas des coupures. Néanmoins, dans le cas des fondus, celle-ci donne un grand nombre de faux positifs. Cette méthode présente alors une grande sensibilité quant à la détection de fondus. Quant à la méthode par convolution du filtre Sobel, celle-ci donne un grand nombre de fausses détections dans le cas des coupures et des fondus. Aussi, nous avons pu constater que la qualité d'images de la vidéo était souvent la source de ces faux positifs, notamment dans le cas des fondus. Ceci dit, pour cette vidéo, la méthodologie par histogrammes des couleurs semble le plus se rapprocher de la vérité de terrain.

Question 6

Voici un résumé des données obtenues sur les trois vidéos avec la vérité de terrain et les deux méthodes. Pour la vidéo *anni009*, la vérité de terrain indique qu'il y a 38 coupures et 65 fondus. La méthode des histogrammes trouve 33 coupures et 180 fondus. Quant à la méthode des arêtes, celle-ci trouve 704 coupures et 130 fondus. Pour la vidéo *anni005*, la vérité de terrain indique qu'il y a 38 coupures et 27 fondus. La méthode des histogrammes trouve 29 coupures et 135 fondus. La méthode par convolution trouve 614 coupures et 176 fondus. Finalement, pour la vidéo *NAD57*, la vérité de terrain indique qu'il y a 44 coupures et 23 fondus. La méthode des histogrammes trouve 28 coupures et 40 fondus. Avec la méthode des arêtes, nous trouvons 411 coupures et 131 fondus.

En comparant les résultats de la méthode des histogrammes avec les vérités de terrain respectives, les vidéos *anni005* et *NAD57* ont moins de faux positifs de coupures que la vidéo *anni009*, malgré le fait que nous ayons adapté les paramètres pour cette dernière. Cette méthode est limitée, car les taux de couleurs varient selon les vidéos. Les seuils pour détecter les effets efficacement ne sont pas identiques pour toutes les vidéos. Dans notre cas, les seuils semblent être bons, puisque la méthode reste efficace pour détecter les coupures, mais ce ne serait pas le cas pour toutes les vidéos.

Comme mentionné dans la question précédente, notre méthode par convolution n'est pas efficace pour la détection des coupures et des fondus. Il est donc difficile de comparer la performance de notre solution sur les trois vidéos. Nous pouvons hypothétiser que la méthode aurait probablement encore une quantité de faux positifs, car elle dépend des circonstances de la vidéo et les paramètres ne sont pas facilement ajustables. Ainsi, cette méthode n'est pas prédictible et donc nous ne pouvons pas déduire que la quantité de faux positifs sera petite ou grande.

Références

1. https://docs.opencv.org/master/d1/db7/tutorial_py_histogram_begins.html
2. https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram_comparison/histogram_comparison.html
3. <https://github.com/gabilodeau/INF8770/blob/master/Gradients%20et%20extraction%20d'aretes%20sur%20une%20image.ipynb>
4. <https://stackoverflow.com/questions/37203970/opencv-grayscale-mode-vs-gray-color-conversion>
5. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_gradients/py_gradients.html
6. <https://www.geeksforgeeks.org/erosion-dilation-images-using-opencv-python/>