A logo with a bee and text

Description automatically generated

Rapport TP3

Indexation du contenu pictural

*Rédigé par*

Mariana Chang leong — 2074331 Zakarya Khnissi — 1989641

Équipe 8

Technologies multimédias INF8770

Mehdi Miah

École Polytechnique de Montréal Département de GIGL

19 février 2023

Question 1

1. D’après nos mesures nous avons réussi à avoir un taux de bonnes réponses de 70% (704/1000) avec un écart temporel moyen de 2.08 sec.
2. La taille totale en octets des 100 vidéos MP4 sur le disque dur est 443 831 579.
3. L’ordre de grandeur de la taille en octet si on extrayait toutes les trames des 100 vidéos en JPEG est de 6 899 117 116.1.
4. L’ordre de grandeur en octets si on conservait toutes les trames en mémoire RAM est 104 772 096 000.
5. La taille la moins élevée de la question b) peut être expliquée par le fait que les codecs compressent mieux les données ce qui réduit la quantité de données car ça exploite la successivité des trames ce qui n’est pas le cas pour la compression JPEG qui compresse les images individuellement. La taille la plus élevée serait celle de la conservation de toutes les trames avec leurs résolution complète.

Nos résultats ont été génères dans notre script python en utilisant les méthodes disponibles dans la librairie cv2.

Question 2

1. Hypothèse 1: L’algorithme prend une image à chaque n\_images clefs, en divisant le nombre de trames par vidéo par n\_images clefs réduisant par conséquence le volume de données à comparer pendant la recherche. Cela fait un équilibre entre le temps de calcul et la précision de recherche contrairement aux autres méthodes de détermination de ce paramètre. Sélectionner toutes les trames de la vidéo mène à une précision élevée mais couteux en termes de performances et ressources, qui est le cas aussi si on décide de prendre n images par seconde et la vidéo est longue. On a jugé aussi la sélection dynamique des trames couteuse en termes de performance et qu’elle ne répondra pas aux contraintes de notre TP.
2. Hypothèse 2 : Pour avoir une mesure d’affinité précise, on préfère utiliser l’espace colorimétrique RGB avec une dimensionnalité 3D en découpant l’histogramme en 8 intervalles pour les trois canaux de couleur cela peut être couteux en terme de performance vu que notre ensemble de données est grand, on est arrivé a la solution des arbres KD qui permet de partitionner les données ce qui donne une recherche rapide pour un grand ensemble de données ( [K-Dimensional Tree in Data Structures (scholarhat.com)](https://www.scholarhat.com/tutorial/datastructures/k-dimentional-tree-in-data-structures) ) avec la métrique de distance euclidienne comme mesure d’affinité, cette dernière est une bonne métrique pour les histogrammes de couleurs car ça permet de bien détecter la différence entre les distributions de couleurs, on juge que ces mesures d’affinités nous permet de compenser pour le cout de calcul qui peut être élevée en utilisant l’espace colorimétrique RGB en 3D.

Question 3

1. Le pourcentage de vidéos correctement retrouvées (en %) : 70.4% (704/1000)
2. L’écart (en valeur absolue) moyen sur le minutage lorsque la vidéo est correctement retrouvée et présente dans la banque de données (en seconde) : 2.08 sec
3. Le taux de compression des données indexées, calculé avec 1− Tc To avec To la valeur obtenue à la question 1b) et Tc la taille en octet de la matrice de dimension N × D : 1.00
4. Moyenne temps d'indexation (par image): 0.2772 sec
5. Moyenne temps de recherche (par image): 0.0073 sec