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 92.5% avec un écart temporel moyen de 1.53 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

* Hypothèse 1: Notre algorithme sélectionne n-images clefs à chaque n images dépendamment du nombre de trames dans chaque vidéo, on juge que ce choix est le meilleur choix parmi les autres vu que les vidéos en question ne sont pas trop longues ce qui va faire un équilibre entre la complexité temporelle et de mémoire ainsi que l’écart temporel moyen plus bas. La dimensionnalité 1D économisera plus de temps et de mémoire aussi, vu que chaque canal de couleur est traité séparément, couplée avec l’espace colorimétrique RGB. On juge aussi que la distance euclidienne est le meilleur choix afin d’avoir un équilibre entre le temps de calcul et la précision, contrairement à la distance Bhattacharya qui donne des résultats plus précis mais couteux en termes de performance.
* Hypothèse 2 : On juge que les paramètres choisis dans la première hypothèse sont les meilleurs choix, sauf pour la dimensionnalité ou on suppose que la capture des interactions entre les trois canaux de couleurs augmentera bien la précision de recherche sans affecter trop la performance temporelle et spatiale.

Question 3

Hypothèse 1

1. Le pourcentage de vidéos correctement retrouvées (en %) : 92.5% (925/1000)
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) : 1.53 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 : 0.999
4. Moyenne temps d'indexation : 107.4342 seconds, 1 min et 47 secondes
5. Moyenne temps de recherche (par image): 8.1512 ms/image

Hypothèse 2

1. Le pourcentage de vidéos correctement retrouvées (en %) : 87.7% (877/1000)
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) : 1.68 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 : 0.999
4. Moyenne temps d'indexation : 117.0786 seconds, 1 min et 57 secondes
5. Moyenne temps de recherche (par image): 7.8971 ms/image

Question 4

* Hypothèse 1 : n\_images clefs fixes , resnet-18, cpu , battacharayya distance (more accurate results, takes longer to compute)
* Hypothese 2: n\_images clefs, resnet-18, cpu, cosine distance (faster less accurate results)

Question 5

Hypothèse 1

1. Le pourcentage de vidéos correctement retrouvées (en %) : 95.1% (951/1000)
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) : 0.83 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 : 0.99
4. Moyenne temps d'indexation : 241 seconds, 4 min
5. Moyenne temps de recherche (par image): 9.15 ms/image

Hypothèse 2

1. Le pourcentage de vidéos correctement retrouvées (en %) : 94.7% (947/1000)
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) : 0.87 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 : 0.999
4. Moyenne temps d'indexation : 241 seconds, 4 min
5. Moyenne temps de recherche (par image): 6.5931 ms/image

Question 6