

INF8770

Technologies multimédias

A2020 - Travail pratique #3

Indexation de contenu pictural

Remise du travail :

- Au plus tard, le 7 décembre 2020, 8h30 sur Moodle - **aucun retard accepté**

Documents à remettre :

- Votre code source ainsi qu'un rapport en **PDF** dans une archive (.zip/.7z/...).
- Vous devez remettre votre rapport sur www.turnitin.com. Les instructions sont sur Moodle.

Autres directives :

- Il vous est interdit d'utiliser du code écrit par d'autres équipes.
 - Il vous est permis d'utiliser du code trouvé sur internet et des bibliothèques externes, mais vous devez nous donner toutes les références.
 - Le code lui-même n'est pas évalué, mais il doit être remis au complet.
-

Préambule :

Ce travail pratique inclut une évaluation des qualités 2, 3 et 5 du Bureau canadien d'agrément des programmes de génie (BCAPG).

- **Qualité 2 Analyse de problèmes** : capacité d'utiliser les connaissances et les principes appropriés pour identifier, formuler, analyser et résoudre des problèmes d'ingénierie complexes et en arriver à des conclusions étayées.
- **Qualité 5 Utilisation d'outils d'ingénierie** : capacité de créer et de sélectionner des techniques, des ressources et des outils d'ingénierie modernes et de les appliquer, de les adapter et de les étendre à un éventail d'activités simples ou complexes, tout en comprenant les contraintes connexes.

Travail à réaliser : Dans ce TP, vous devez implémenter un algorithme de reconnaissance de vidéos : à partir d'une image et d'une banque de données, vous devez retrouver la vidéo dans laquelle cette image est extraite. L'algorithme implémenté devra répondre à certains critères de performances (temps d'exécution, volume de stockage, capacité à trouver la bonne vidéo et le bon moment)

Objectif de l'algorithme d'indexation

Étant donnée une image, vous devez retrouver la vidéo dans laquelle cette image est extraite. Si vous trouvez une vidéo candidate, vous devez également retourner le moment où cette trame apparaît. Si au contraire, aucune vidéo ne semble être candidate, aucun moment n'est retourné. La figure 1 illustre l'approche générale de l'algorithme.

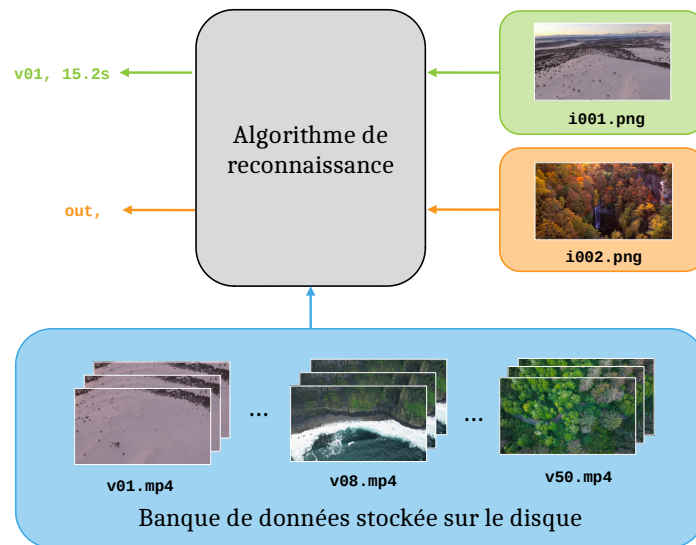


FIGURE 1 : Illustration de l'algorithme de reconnaissance de vidéos

Description des données

La banque de vidéos contient 50 vidéos en HD (1280×720), sous le format MP4, en une seule prise de vue, libres de droit, extraites du site www.pexels.com (ces vidéos sont disponibles sur Moodle). En plus de ces vidéos, vous trouverez 200 images qui seront les requêtes pour la recherche ainsi qu'un fichier `solution.csv` qui indique pour chaque image le nom de la vidéo dans laquelle elle est extraite et son minutage en secondes (dans le cas où l'image provient d'une vidéo absente de la banque de données, ces deux valeurs sont respectivement à "out" et à vide). Les images sous le format PNG ont été extraites sans perte des vidéos et celles sous le format JPEG ont subi une perte.

Mesure des performances

Les performances des algorithmes seront mesurées sur quatre critères :

1. le temps d'exécution moyen pour une image (en seconde) ;
2. le pourcentage de vidéos correctement reconnues (en %) ;
3. l'écart (en valeur absolue) moyen sur le minutage lorsque la vidéo est correctement reconnue et présente dans la banque de données (en seconde) ;
4. la mémoire utilisée pour stocker sur le disque les informations sur les vidéos (en Mo)

Question 1 (/10)

Alice, Bob et Carol sont trois étudiants n'ayant pas suivi le cours INF8770. Ils proposent cependant trois algorithmes permettant de retrouver une vidéo à partir d'une image.

Alice propose de stocker les vidéos sous le format MP4 telles quelles sans traitement particulier. Puis, étant donnée une image, elle propose de parcourir toutes les vidéos, trame par trame, et de calculer une distance euclidienne entre les pixels de l'image d'intérêt et la trame de la vidéo. Dès que la distance est nulle, elle retourne la vidéo en question ainsi que le minutage. Si la distance nulle n'est jamais atteinte, elle renvoie un message indiquant que la vidéo n'a pas été trouvée.

Bob propose de stocker les vidéos sous la forme d'images (et d'un fichier d'indexation optionnel) car il pense que cela accélère le traitement compte tenu qu'il s'agit de comparer des images avec d'autres images. Pour

cela, il hésite entre les deux formats PNG et JPEG. Pour la suite, il reprend les idées d'Alice : il propose de parcourir toutes les trames extraites et de calculer une distance euclidienne avec l'image d'intérêt. Dès que la distance est nulle, il retourne la vidéo associée à la trame ainsi que le minutage. Si la distance nulle n'est jamais atteinte, il renvoie un message indiquant que la vidéo n'a pas été trouvée.

Carol propose de sélectionner une trame à chaque seconde de chaque vidéo et de la décrire par un histogramme de couleurs (un fichier d'indexation additionnel peut être utilisé). Puis, étant donnée une image, Carol calcule son histogramme de couleurs et retourne la vidéo et le minutage de l'histogramme de couleurs ayant la plus faible distance euclidienne.

D'après vos connaissances en technologies multimédias, discutez de la pertinence de chaque algorithme lors de l'application sur les images en PNG et en JPEG. Quelles recommandations feriez-vous à Bob quant au format de l'image dans son algorithme ?

- **Qualité évaluée : 2.1 Identifier et formuler un problème**

Critère d'évaluation : Analyse des algorithmes présentés par rapport aux quatre mesures de performance.

Justification du format d'image approprié pour l'algorithme de Bob.

Question 2 (/10)

Implémentez l'un des trois algorithmes proposés et appliquez-le sur les images en PNG et en JPEG. Est-ce que les résultats en termes de performance sont conformes à vos remarques pour la question 1 ?

NB1 : si vous choisissez l'algorithme de Bob, prenez le format d'images qui vous semble le plus pertinent.

NB2 : si vous choisissez l'algorithme de Carol, vous êtes libres quant aux choix des paramètres de l'histogramme des couleurs.

- **Qualité évaluée : 5.3 Créer ou adapter un outil**

Critère d'évaluation : Capacité à créer un code informatique fonctionnel.

- **Qualité évaluée : 5.4 Intégrer des outils**

Critère d'évaluation : Capacité à bien intégrer les fonctions de bibliothèques de traitement d'images/vidéos dans son propre code.

Question 3 (/30)

Implémentez votre propre algorithme de reconnaissance de vidéos. Discutez de vos choix d'implémentation (sélection d'images clés, format des données sauvegardées, algorithme de recherche, caractérisation d'une image, etc.)

Comparez vos mesures de performances avec la méthode implémentée à la question 2 vis-à-vis des quatre mesures de performances sur les images en PNG et en JPEG. Synthétisez vos résultats dans un tableau ou un graphique.

- **Qualité évaluée : 2.2 Explorer des approches de résolution et planifier la démarche**

Critère d'évaluation : Justification des choix de votre algorithme au regard des limites rencontrées par ceux décrits à la question 1.

- **Qualité évaluée : 2.4 Produire des résultats**

Critère d'évaluation : Pertinence des résultats et du protocole de test pour démontrer la qualité de la solution proposée.

- **Qualité évaluée : 2.5 Valider ses résultats et recommander**

Critère d'évaluation : Qualité et exhaustivité de l'analyse des résultats.

- **Qualité évaluée : 2.6 Analyser l'incertitude, la sensibilité et les limites des approches**

Critère d'évaluation : Qualité des explications sur les limites de la méthode.