
Examen de Bases de données multimédia

5 février 2014

durée : 2h00

Documents autorisés : transparents du cours, notes de cours.

Calculatrice non autorisée.

Bonus = points en plus hors barème.

EXERCICE 1. Questions à réponses courtes (10 pts)

Répondez de manière *concise*. Sauf si cela est explicitement demandé, vous n'êtes pas obligé de faire des phrases ni de justifier vos réponses.

[Question 1.] QCM : Choisissez une réponse aux propositions suivantes.

Attention : les mauvaises réponses sont pénalisées.

1. Instagram est une application et un service de partage de photos et de vidéos disponible sur plates-formes mobiles. Apple l'a désignée comme "Application de l'année" en 2011. Instagram a été racheté par Facebook pour 1 milliard de dollars. [Source : wikipedia]. On peut en conclure que sur Instagram, il y a (a) 16×10^3 images, (b) 16×10^6 images, (c) 16×10^9 images, (d) 16×10^{12} images.
2. Pour trouver les images exactement identiques dans un ensemble de fichiers, il vaut mieux (a) comparer les fichiers octet par octet comme des chaînes de caractères (b) calculer un histogramme de couleurs dans l'espace de couleurs Lab, (c) calculer un bag-of-words sur des descripteurs SIFT.
3. La photo suivante est commercialisée sur une banque d'images professionnelle. Le photographe l'a annotée avec le texte (a) "Charlotte et Médor pendant les vacances de Pâques", (b) "footing femme chien pré", (c) "mon programme minceur pour l'été".



4. Quand on mélange deux encres de couleur (par exemple le jaune et le cyan), la lumière réfléchie (le vert) est (a) l'union (b) l'intersection des lumières réfléchies par les encres individuelles.

[Question 2.] Comment la matrice d'auto-corrélation du détecteur de Harris caractérise-t-elle un contour droit ? Comment avoir une indication de la direction de ce bord ?

[Question 3.] Quels sont les deux cas de prise de vue d'une scène rigide qui garantissent qu'on peut mettre en correspondance 2 vues avec une homographie ?

[Question 4.] Un plan d'un film cadre les roues à rayons d'une locomotive à vapeur qui démarre et accélère. Au début du plan, les rayons tournent dans un sens consistant avec le déplacement du train, mais après un certain temps, les roues semblent immobiles, puis paraissent tourner dans l'autre sens !

(a) À quel phénomène d'échantillonnage est-ce dû ?

Le film est tourné à 24 images par seconde. La roue a 12 rayons et fait 50 cm de diamètre.

(b) Quelle vitesse le train a-t-il atteint quand la roue paraît immobile ?

[Question 5.] La recherche de plus proches voisins par kd-tree n'est pas efficace en grande dimension car elle n'élimine que peu de cellules candidates à visiter. Pourquoi ?

[Question 6.] Pour un système d'indexation audio, analyse des fichiers son en les divisant en tranches de 1 s et en calculant un descripteur VLAD à partir de descripteurs MFCC par tranche. Les descripteurs VLAD sont ensuite normalisés et comparés avec des produits scalaires pour produire des similarités.

Le fichier son B est monté à partir de morceaux du fichier son A, en respectant l'ordre des morceaux.

(a) quel méthode peut-on utiliser pour retrouver les morceaux correpondants ?

La table suivante présente les similarités entre les tranches de la piste son A (par ligne) et celles de B (par colonne), les intervalles de temps en secondes sont indiqués.

(b) Appliquez la méthode choisie pour trouver les morceaux correspondants.

B \ A	[0, 1]	[1, 2]	[2, 3]	[3, 4]	[4, 5]	[5, 6]	[6, 7]
[0, 1]	13	4	4	3	4	3	5
[1, 2]	4	19	2	4	2	4	4
[2, 3]	2	2	4	10	4	4	5
[3, 4]	3	4	3	3	13	3	5
[4, 5]	5	4	3	5	2	4	14

EXERCICE 2. Quantificateur produit (product quantizer)

On cherche à calculer les distances euclidiennes de $x \in \mathbb{R}^d$ à toutes les colonnes de la matrice $Y \in \mathbb{R}^{d \times n}$.

[Question 7.] Combien d'opérations flottantes cela nécessite-t-il ?

Maintenant, chaque colonne y de Y est représentée à l'aide d'un quantificateur produit. Le vecteur y est divisé en sous-vecteurs de taille d/q . Chaque sous-vecteur est représenté par un code de 0 à 255.

[Question 8.] Quelle est la taille de la matrice de codes en octets ? Quel est le facteur de compression, en supposant que X est représenté avec des flottants en 32 bits ?

[Question 9.] Quel est le coût de calcul (nombre d'opérations flottantes) des distances approchées en utilisant une table de distances précalculée ?

[Question 10.] Peut-on appliquer cet algorithme pour d'autres types de distances :

- L1 ?
- L ∞ ?
- χ^2 (Chi-2) ?
- Mahalanobis ?

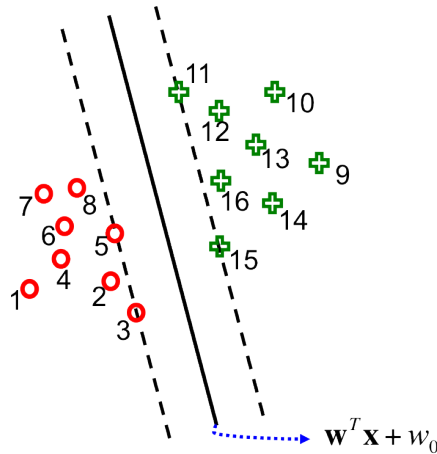
Décrivez les adaptations de l'algorithme.

On souhaite maintenant calculer une approximation de distances d'une colonne à une autre de X (il y en a donc $n \times n$).

[Question 11.] Décrivez un algorithme qui permet ce calcul, en utilisant une table précalculée.

EXERCICE 3. Machine à vecteur de support (Support Vector Machine)

Considérez un plan (espace à deux dimensions) dans lequel sont répartis deux groupes de points. On peut trouver un séparateur linéaire, obtenu par combinaison linéaire du vecteur \mathbf{x} avec un vecteur de poids $\mathbf{w} : \mathbf{w}^T \mathbf{x} + w_0$.

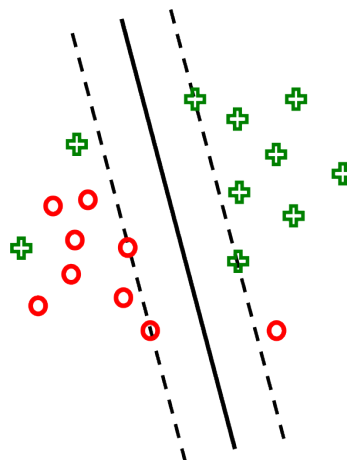


[Question 12.] Écrire la largeur de la marge pour cet exemple. Expliquer comment la calculer.

[Question 13.] Identifier les vecteurs supports dans cet exemple.

[Question 14.] Il existe une infinité d'hyperplans séparateurs. Écrire le problème d'optimisation pour trouver le meilleur séparateur en utilisant la marge. Détailler avec les contraintes.

[Question 15.] **Marge souple** : La technique de marge souple tolère les mauvais classements et permet de trouver une séparatrice linéaire pour quelques cas non séparables (ci-dessous par exemple). Écrire les modifications nécessaires dans le problème d'optimisation.



EXERCICE 4. Reconnaissance d'actions

On décide d'implémenter un système de reconnaissance d'actions, spécifiquement, l'action *boire un café*, sur une base de données de films d'Hollywood.



[Question 16.] Quel descripteur local voulez-vous utiliser ? Justifier la réponse.

[Question 17.] Quelles sont les étapes de calcul de votre descripteur de choix ? Sont-ils calculés sur une vidéo ou une image ?

[Question 18.] Comment agréger les descripteurs pour produire un descripteur de la vidéo complète ?

On suppose maintenant qu'on a un descripteur de taille fixe pour la vidéo complète.

[Question 19.] Comment faire l'apprentissage de ce système ? Vous pouvez utiliser la machine à vecteur de support par exemple. Expliquer les étapes, les données en entrée et en sortie.