|  |
| --- |
| Le 22 Février 2018  Yihong XU  Junshuai ZHU |
|  |

**Descripteurs locaux pour la recherche d’images similaires**

**TP Lot1**

Ingénieur Généraliste

F4B 516

Encadrant : John Puentes

1. problématique

Le but de ce projet est de classifier de manière automatique les objets similaires dans des images dont l’angle de visualisation change et la condition luminaire change. Le data entrant est un objet dans une image et le data sortant est des images des objets les plus similaires. Les algorithmes de détection des points caractéristique sont ORB, Brief, SIFT.

1. Introduction des algorithmes
   1. SIFT

Intro :

Premièrement, l'axe de coordonnées doit être tourné vers la même orientation du point caractéristique pour assurer l'invariance de rotation. L’orientation est calculée depuis l’histogramme d’orientation généré par la magnitude de gradient et l’orientation. Puis calculer le gradient des pixels adjacents dans la fenêtre de 16\*16 autour du point caractéristique. Le gradient de ces pixels peut être pondéré par une fenêtre gauss et les pixels les plus proches ont un poids plus important. Une sous-région est composée de 4\*4 pixels. Un histogramme d'orientation pour la sous-région peut être formé en accumulant l'orientation du pixel 4\*4 qu'il contient. Parce qu'il y a 4\*4 sous-régions et chaque sous-région a 8 dimensions, il donne un vecteur descripteur de 128 dimensions pour chaque point caractéristique.

Bon ：Rotation invariance ; Echelle invariance ; Luminance invariance ; Taux de correspondance haut

Con ：Lent

Image :

* 1. Brief

Intro : Choisir la région pour calculer le descripteur ; Calculer la Convolution Gaussienne dans cette région pour éliminer le bruit ; Choisir de manière aléatoire des pairs de point, ex. <x1, y1>, <x2, y2>…, autour du point caractéristique ; Calculer l’élément du descripteur en comparant la valeur de la paire de point choisie (ex, si x<y, retourne 1 ; Sinon, retourne 0) ; Générer le descripteur binaire en répétant plusieurs fois(ex.256) l’étape précédente.

Bon ：Rapid pour le calcul des descripteurs ; Rapid pour le pairage

Con ：Non rotation invariance ; Non échelle invariance ; Sensitif au bruit ; Taux de correspondance bas

Image :

* 1. ORB

Intro : rBRIEF utilise l'algorithme glouton pour filtrer les meilleures paires de points en choisissant la méthode : Construire un ensemble de 300k points caractéristiques. Pour chaque point caractéristique, considérez ses 31\*31 points adjacents comme son patch. Pour la capacité anti-bruit, utilisez les 5\*5 quartiers du point pour représenter la valeur de celui-ci. Donc, il y a (31-5+1) \* (31-5 +1) = 729 sous-fenêtres de taille 5\*5 et il y a M = (728+1) \* 728/2 = 26536 façons de choisir des paires de points. Dans ces 26535 façons de choisir, 256 meilleures méthodes de choix doivent être filtrées.

Pour chaque point clé, utilisez M façons de choisir des paires de points voisins et de générer une matrice Q égale à 300k \* M. Pour chaque colonne de Q, cela signifie que les descripteurs de ces 300k points clés sont générés en utilisant l'une des méthodes de choix de M.

Calculer la moyenne de chaque colonne dans Q ; Réorganiser le vecteur colonne de Q en utilisant la distance entre 0,5 et la moyenne de la colonne. Après que Q est réorganisé, il donne la matrice T. Mettez la première colonne de T dans R.

Calculez la corrélation entre la colonne suivante dans T et toutes les colonnes dans R. Si la corrélation est plus petite que le seuil correspondant, déplacez cette colonne vers R. Répétez l'étape précédant (le calcul de la corrélation), jusqu'à ce qu'il y ait 256 colonnes dans R. S'il n'y a pas assez de colonnes dans R, augmentez le seuil et réessayez.

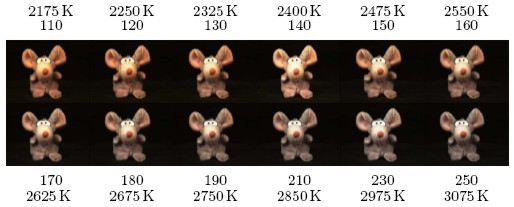
Bon ：Rapid, Rotation invariance, Echelle invariance(openCV), Moins sensitif au bruit

Con ：Taux de correspondance bas

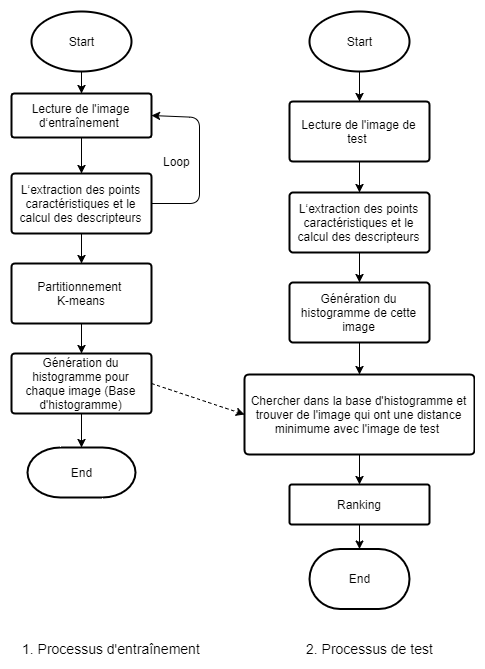
Image :

1. Introduction du fichier de données

Nous utilisons la base d’images ALOI (Amsterdam Library of Object Images). Nous avons choisi 2 groupe d’images :

* 1. Changement de couleur de la lumière
  2. Changement de l’angle de visualisation

Donc, comme indiqué dans ces figures précédentes, pour chaque class d’objet, il y a ensemble 15 images. Nous avons divisé ce groupe en deux parties. Une partie de 10 image comme la groupe d’entraînement, l’autre partie de 5 images comme la groupe de test.

1. Chaîne de traitement
2. L’analyse
   1. Performance des différents descripteurs
   2. Performance pour quantité des points caractéristiques différente
   3. Performance pour quantité de cluster différente
3. Conclusion