

Institut de la Francophonie pour l'informatique

Vision par ordinateur

TP 2 : Reconnaissance d'objets avec le descripteur SIFT

Date de remise : 23h 13/03/2024

Remise : Moodle (smartlearning)

Détecteur de points d'intérêt SIFT et descripteur SIFT

Pour ce tp, vous allez utiliser les points d'intérêt calculés par la méthode SIFT (voir cours sur les points d'intérêt). Bien que cette méthode soit a priori difficile à comprendre d'un point de vue théorique, son utilisation est très simple grâce à la démo disponible sur la page de David Lowe (<http://www.cs.ubc.ca/~lowe/keypoints/>).

Pour commencer, allez sur cette page, télécharger la démo et essayez-là avec les exemples donnés. Avec le code (en C), vous comprendrez comment calculer les points d'intérêt et comment faire la correspondance (*matching*) entre deux points. Tout y est. Cependant, vous n'avez pas accès à tout le code (pour raison de brevet). Mais les exemples donnés vous permettent de calculer les points d'intérêts dans un fichier, puis vous devrez lire ce fichier texte dans votre programme. Ne cherchez pas à tout inclure dans un seul programme, mais procédez en deux étapes comme dans la démo. Vous pouvez appeler le calcul de SIFT par des appel `system` dans votre code, ou en précalculant toutes les valeurs dans des fichiers texte par exemple.

Actuellement, SIFT est aussi implémenté dans OpenCV, donc vous pouvez l'utiliser à la place du fichier exécutable du David Lowe. Donc, c'est à votre choix.

N'hésitez pas à télécharger et à lire les articles de recherche disponibles sur la page de David Lowe. Outre les explications théoriques, on y trouve des explications pratiques et simples à comprendre sur l'utilisation de SIFT pour plusieurs applications.

Mise en correspondance entre deux images avec SIFT

Calcul des descripteurs SIFT

Vous allez extraire les descripteurs SIFT de chaque image. Chaque descripteur SIFT consiste en un vecteur de 128 valeurs. Pour une seule image, il est possible d'extraire des dizaines ou même des centaines de points SIFT.

Mise en correspondance (*matching*) de points d'intérêts

Etudiez la méthode de mise en correspondance présentée dans la démo SIFT (et aussi présentée en classe)). Il s'agit d'une méthode possible, mais pas la seule.

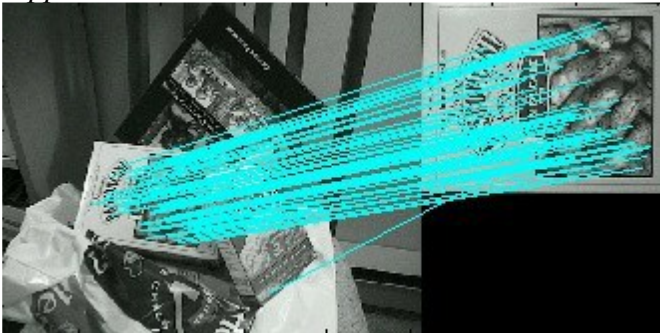
La mise en correspondance de deux images se fait en calculant les correspondances entre tous les descripteurs SIFT de ces deux images. La distance entre deux vecteurs SIFT peut se faire simplement en prenant la distance Euclidienne entre les deux vecteurs. On considère la distance la plus courte (petite) entre deux vecteurs comme la meilleure. C'est-à-dire, chaque descripteur f_1 (associé à un point d'intérêt) dans la 1ère image est apparié avec un descripteur f_2 (associé à un autre point d'intérêt dans la 2ème image) qui ont la distance minimale avec f_1 en comparant avec d'autres descripteurs

dans la 2eme image.

Afin de réduire le nombre de fausses correspondances, une métrique possible est de calculer le ratio entre la distance la plus courte et la deuxième plus courte distance. Si ce ratio est inférieur à un seuil (à déterminer), alors la correspondance peut être considérée comme robuste, sinon elle est rejetée :

$$ratio = \frac{d_{plusproche}}{d_{deuxiemeplusproche}} < seuil$$

Remarque : Comme indiqué auparavant, il s'agit d'un exemple de mesure possible (pas forcément la meilleure) pour déterminer les bonnes et les mauvaises correspondances. Il existe d'autres façons de déterminer cela et vous êtes libre d'utiliser une autre méthode, tant que vous l'expliquez dans votre rapport.



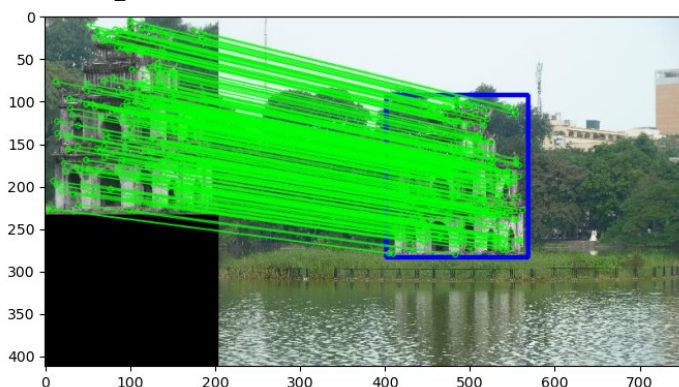
Localization d'objets avec SIFT

Remarque : La méthode décrite ci-dessous est simple. Elle n'est ni complète ni parfaite. Il vous restera plusieurs points à résoudre pour bien réussir ce tp. Je vous encourage fortement à étudier les différentes solutions et idées existantes sur Internet. Il existe une documentation abondante disponible pour ce tp, donc profitez-en. Dans votre rapport, expliquez-bien les idées que vous avez implémentées, en citant les références extérieures au besoin.

Supposer que vous avez l'image 1 qui contient un objet (objet de requête), et l'image 2 (image de recherche) qui contient peut-être plusieurs différents objets. Vous devez localiser la position de cet objet dans l'image 2 (comme illustrer dans l'image suivante).

Pour le faire, il faut :

- Mettre en correspondance des descripteurs locaux entre 2 images pour trouver des paires des points d'intérêt associés.
- Chercher la transformation homographie (H) entre l'ensemble des points d'intérêt de l'image 1 avec l'ensemble des points d'intérêt associés dans l'image 2.
(https://docs.opencv.org/3.4/d1/de0/tutorial_py_feature_homography.html)
- Calculer les positions transformées dans l'image 2 des 4 coins de l'objet de requête dans l'image 1.



Dans l'image 2 (l'image de recherche), peut-être il n'existe pas l'objet de requête. Dans ce cas, vous devez préciser l'absence de cet objet dans l'image de recherche. Une solution simple est d'calculer le score de correspondances réussies entre 2 images. Plus ce nombre de correspondances réussies, plus les deux images ont des chances de représenter la même classe d'objets. Et puis, vous pouvez mettre un seuillage sur ce score pour décider la présence ou l'absence de l'objet. Si ce score est inférieur à un seuil (*seuil_score*), l'objet n'est pas apparu dans l'image de recherche. Sinon, localiser-le dans cette image. Le score peut être calculé comme suit :

$$score = \frac{\# \text{ correspondances réussies}}{\# \text{ descripteurs de l'objet de requete}} < \textit{seuil_score}$$

Rapport à remettre

Dans votre rapport, expliquez bien les idées que vous avez implémentées et qui ne sont pas expliquées dans cet énoncé. Si vous avez d'autres idées pour améliorer les résultats, discutez-les. Donnez les références si nécessaire.

Montrez quelques résultats de mise en correspondances, bon et mauvais. Analysez et expliquez pourquoi vous obtenez ces résultats.

Vous avez des différents paramètres (comme des seuils), essayez de tester avec des valeurs différentes et choisissez des valeurs convenables. Analyser l'influence de ces valeurs sur vos résultats.

Images à tester pour ce TP

Vous pouvez analyser et tester avec d'autres images. Mais n'oubliez pas de faire des tests et des évaluations sur les images fournies dans ce TP.

- Objets de requêtes : fichier attaché
- Image de recherche : fichier attaché

Bon travail !