



**UNIVERSITE NATIONALE DU VIETNAM**  
**INSTITUT FRANCOPHONE INTERNATIONALE**



**Option: Systèmes Intelligents et Multimédia (SIM)**

## **VISION PAR ORDINATEUR**

### **RAPPORT DE TP1**

**THEME : Détecteur de points d'intérêt SIFT et  
descripteur SIFT**

Rédigée et présentée par:  
FANANG NDONG Claude Ericka: étudiante

Sous la supervision de : NGUYEN Thi Oanh  
coordinateur du module CV

**Année Académique : 2024-2025**

## 1. Introduction

La vision par ordinateur a connu des avancées significatives avec l'émergence de techniques de détection et de description de points d'intérêt, parmi lesquelles le détecteur et descripteur SIFT. Ce TP explore la mise en correspondance entre deux images en utilisant le détecteur SIFT pour extraire les points d'intérêt et générer des descripteurs. Nous abordons également la localisation d'un objet spécifique dans une image de recherche à l'aide de la transformation homographique.

## 2. Détection des points d'intérêt et Extraction des descripteurs SIFT :

❖ La première étape de notre travail a été la détection des points d'intérêt grâce au détecteur SIFT. Cette méthode nous permet de repérer des zones clés dans les images, qui sont des endroits susceptibles de contenir des informations discriminantes. Le détecteur SIFT utilise une échelle invariante pour garantir que les points d'intérêt restent détectés même en présence de variations d'échelle. Dans le cas de notre projet, on a procédé comme suit :

- On crée un objet SIFT à l'aide de `cv2.SIFT_create()`.
- Ensuite, on utilise la méthode `detectAndCompute` pour détecter les points d'intérêt (`kp_objet` et `kp_recherche`) et calculer les descripteurs SIFT correspondants (`des_objet` et `des_recherche`) pour les images de l'objet et de recherche, respectivement.

Les points d'intérêt (`kp_objet` et `kp_recherche`) représentent les positions dans les images où le détecteur SIFT a identifié des caractéristiques significatives. Les descripteurs SIFT (`des_objet` et `des_recherche`) sont des vecteurs qui décrivent ces points d'intérêt, capturant les caractéristiques locales des régions détectées.

Le résultat de cette étape est essentiel pour la mise en correspondance ultérieure des points d'intérêt entre les deux images. Ces points d'intérêt et leurs descripteurs associés sont utilisés pour évaluer la similarité entre les caractéristiques des deux images.

- ❖ La deuxième étape de notre travail est l'extraction des descripteurs SIFT. Pour chaque point d'intérêt détecté ('kp\_objet' et 'kp\_recherche'), le descripteur SIFT est extrait. Dans notre cas, es descripteurs sont des vecteurs de **128 valeurs** qui encapsulent les caractéristiques locales de chaque région d'intérêt. L'approche SIFT assure une robustesse aux transformations géométriques et aux altérations d'éclairage, offrant ainsi une représentation fiable des caractéristiques de l'image.

### 3. La reconnaissance

La reconnaissance est le résultat de deux parties de notre programme : la mise en correspondance (matching) et le calcul de nombre de correspondances entre les images.

#### 3.1. Mise en correspondance des points d'intérêt

La mise en correspondance des points d'intérêt entre deux images a été réalisée en comparant les descripteurs SIFT extraits. La distance euclidienne entre les vecteurs de descripteurs a été calculée, et un ratio de distance (variant selon l'image) a été introduit pour filtrer les correspondances les plus pertinentes. Cette approche vise à réduire le nombre de fausses correspondances et à améliorer la robustesse du système. Si ce ratio est inférieur à un seuil, alors la correspondance peut être considérée comme robuste, sinon elle est rejetée :

$$ratio = (d_{plusproche} / d_{deuxiemeplusproche} < seuil$$

L'interprétation du résultat se fait en évaluant la qualité des correspondances obtenues. Un nombre élevé de bonnes correspondances suggère une forte probabilité que les deux images représentent la même classe d'objets. Ces correspondances robustes sont ensuite utilisées pour localiser l'objet dans l'image de recherche. Si le nombre de correspondances est insuffisant, le code indique qu'il n'y a pas assez de correspondances pour une localisation robuste, ce qui peut signifier que l'objet recherché n'est pas présent dans l'image de recherche ou qu'il est partiellement occulté.

### 3.2. Score de correspondance et Seuillage :

Une étape importante de notre approche a été l'introduction d'un score de correspondance. Ce score, calculé en comparant le nombre de correspondances réussies avec le nombre total de descripteurs de l'objet de requête, a servi de métrique pour décider de la présence ou de l'absence de l'objet dans l'image de recherche. Un seuillage a été appliqué sur ce score pour prendre cette décision.

$$\text{score} = (\# \text{ correspondances réussies} / \# \text{ descripteurs de l'objet de requete}) < \text{seuil\_score}$$

Si ce score est élevé, cela suggère une bonne mise en correspondance entre les deux images, augmentant ainsi la probabilité de localiser correctement l'objet dans l'image de recherche. En revanche, un score plus faible peut indiquer une mise en correspondance moins fiable, ce qui peut affecter la capacité à localiser l'objet avec précision. Les seuils et paramètres doivent être ajustés en conséquence pour optimiser les résultats en fonction des images spécifiques utilisées.

## 4. Localisation de l'objet avec la transformation homographique

Pour localiser l'objet de requête dans l'image de recherche, nous avons cherché la transformation homographique (H) entre les ensembles de points d'intérêt des deux images.

Pour l'interprétation du résultat, voici comment nous avons procédé :

- **Nombre de Correspondances :**

- Si le nombre de correspondances entre les descripteurs SIFT de l'objet de requête et de l'image de recherche est suffisant (supérieur à `MIN\_MATCH\_COUNT`), cela indique que des points d'intérêt similaires ont été trouvés dans les deux images.

- **Calcul de l'Homographie :**

- À l'aide de l'algorithme RANSAC, une transformation homographique ('M') est calculée. Cette transformation permet de comprendre comment les points d'intérêt dans l'objet de requête sont positionnés par rapport à l'image de recherche.

- **Projection des Coins de l'Objet dans l'Image de Recherche :**

- Les coins de l'objet de requête sont projetés dans l'image de recherche en utilisant l'homographie calculée. Cela signifie que nous "transposons" les coins de l'objet de requête pour les placer dans le contexte de l'image de recherche.

- **Dessin d'un Polygone :**

- Si le nombre de correspondances est suffisant, un polygone est dessiné autour des coins projetés de l'objet de requête dans l'image de recherche. Ce polygone délimite la position estimée de l'objet dans l'image de recherche.

- **Message en Cas d'Insuffisance de Correspondances :**

- Si le nombre de correspondances est inférieur à `'MIN_MATCH_COUNT'`, un message est affiché indiquant que l'objet n'a pas été correctement localisé. Cela peut être dû à une absence d'objet similaire dans l'image de recherche ou à une détection insuffisante de points d'intérêt.

## **6. Paramètres et Expérimentations**

Nous avons identifié divers paramètres cruciaux, tels que le seuil de ratio, le nombre maximal de correspondance, le seuil\_score, etc., et avons réalisé des expérimentations pour évaluer l'impact de ces paramètres sur les résultats. Des ajustements ont été effectués pour obtenir des valeurs optimales et assurer la fiabilité du système.

## **7. Quelques Résultats obtenus**

Dans cette étape nous allons détailler nos résultats obtenus. Nous avons travaillé sur Google colab, pour avoir une vitesse de traitement beaucoup plus grande. Notre notebook est disponible sur ce lien .

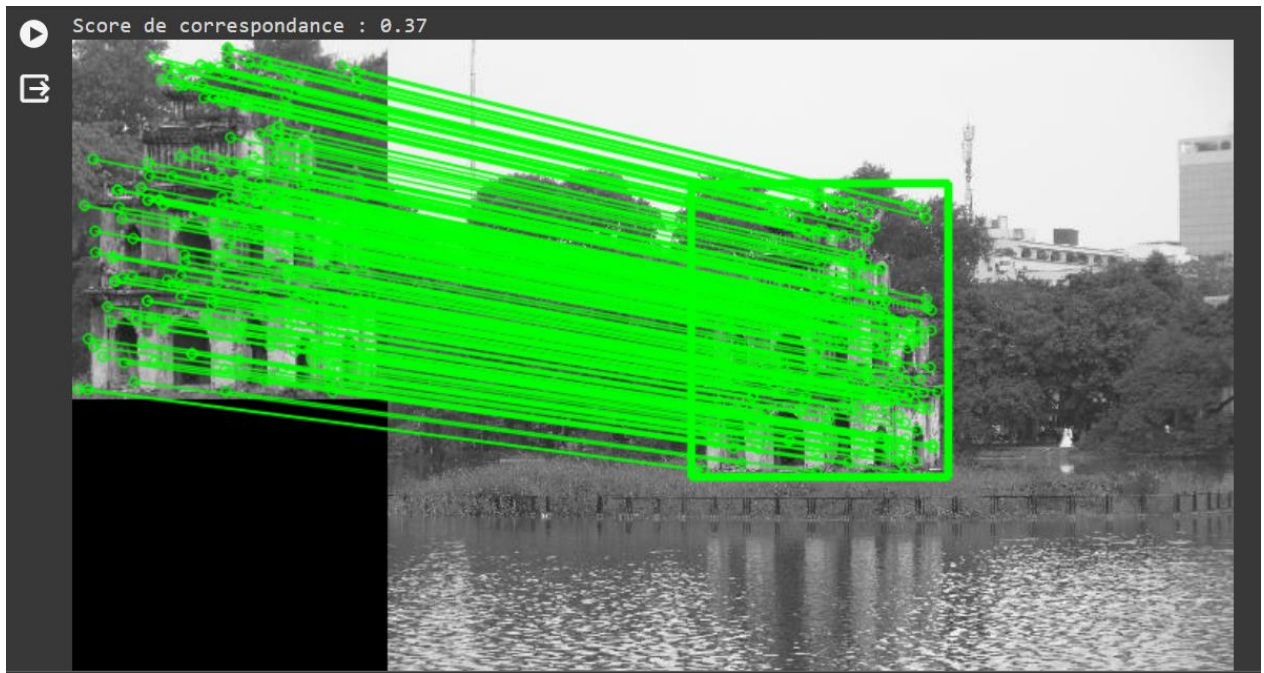
<https://colab.research.google.com/drive/1imDUgFknVpvOPpVgG5IXqThRk9H78Alo#scrollTo=LDPLUujJem4>

### **7.1- Résultat1 : Le temple**

Seuil de ratio : 5

Nombre maximal de correspondance : 10

Conslusion : Score de correspondance : 0.37 . On a une bonne correspondance

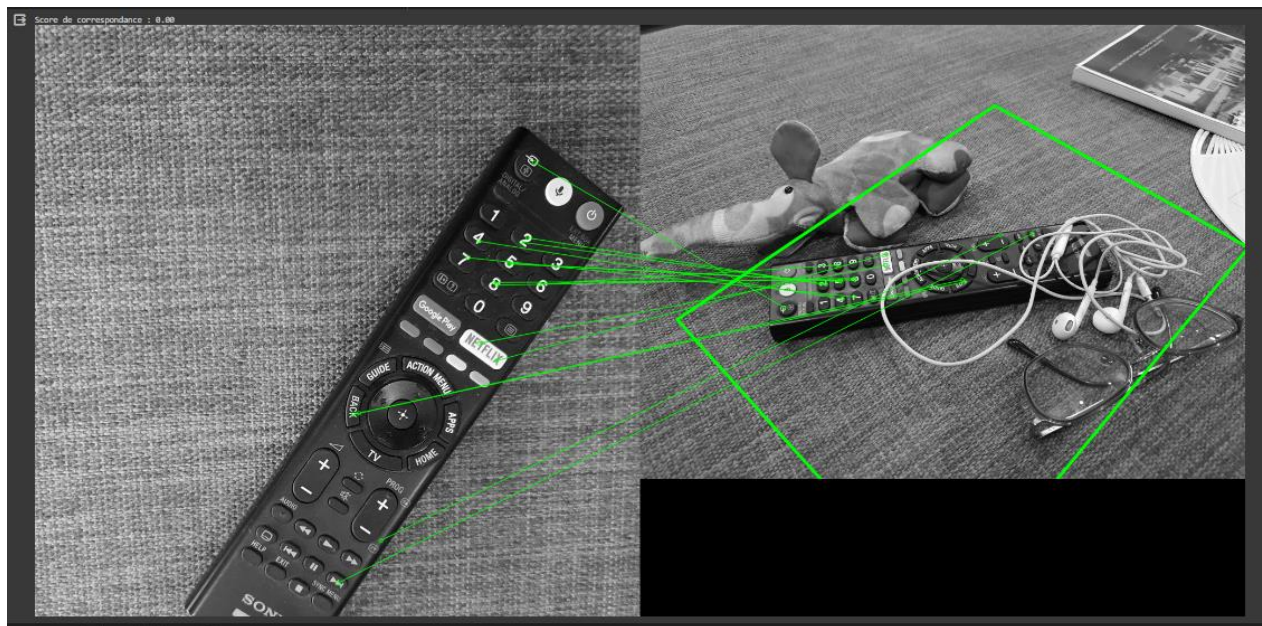


## 7.2. La télécommande

Seuil de ratio : 0.75

Nombre maximal de correspondance : 10

Conslusion : Score de correspondance : 0.5 . On a une bonne correspondance

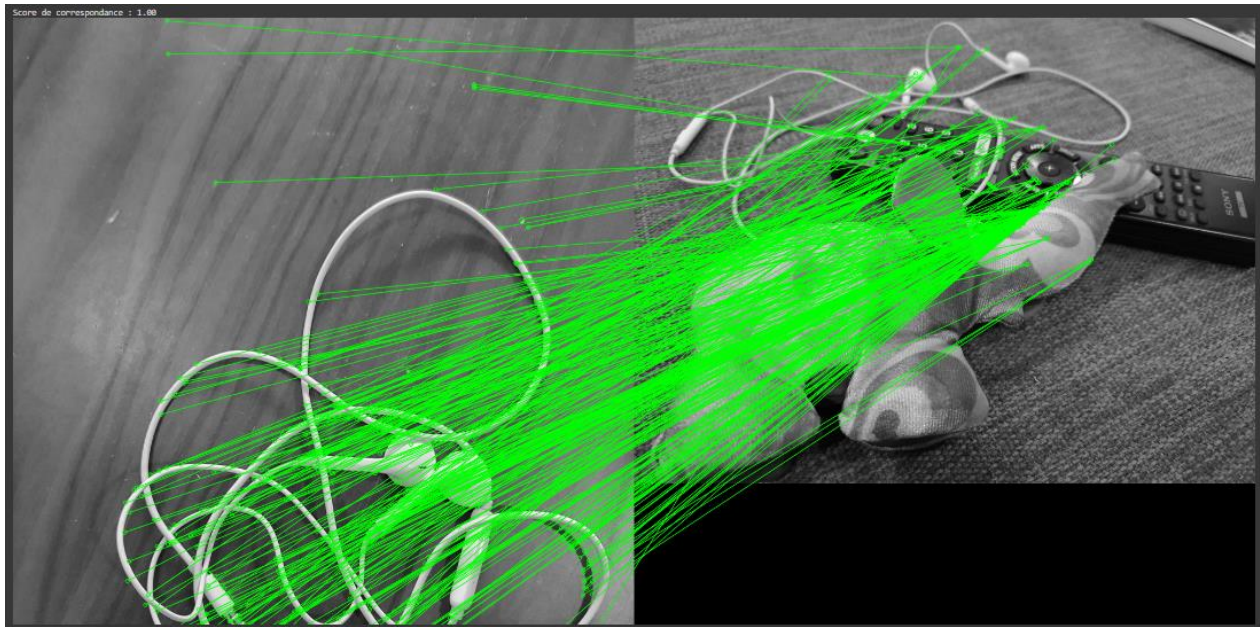


### 7.3. Les écouteurs

Seuil de ratio : 1.5

Nombre maximal de correspondance : 10

Conslusion : Score de correspondance :1. On a une mauvaise correspondance étant donné le seuil de ratio qui est très faible, cela provoque la présence de nombreux points de correspondances également très faibles



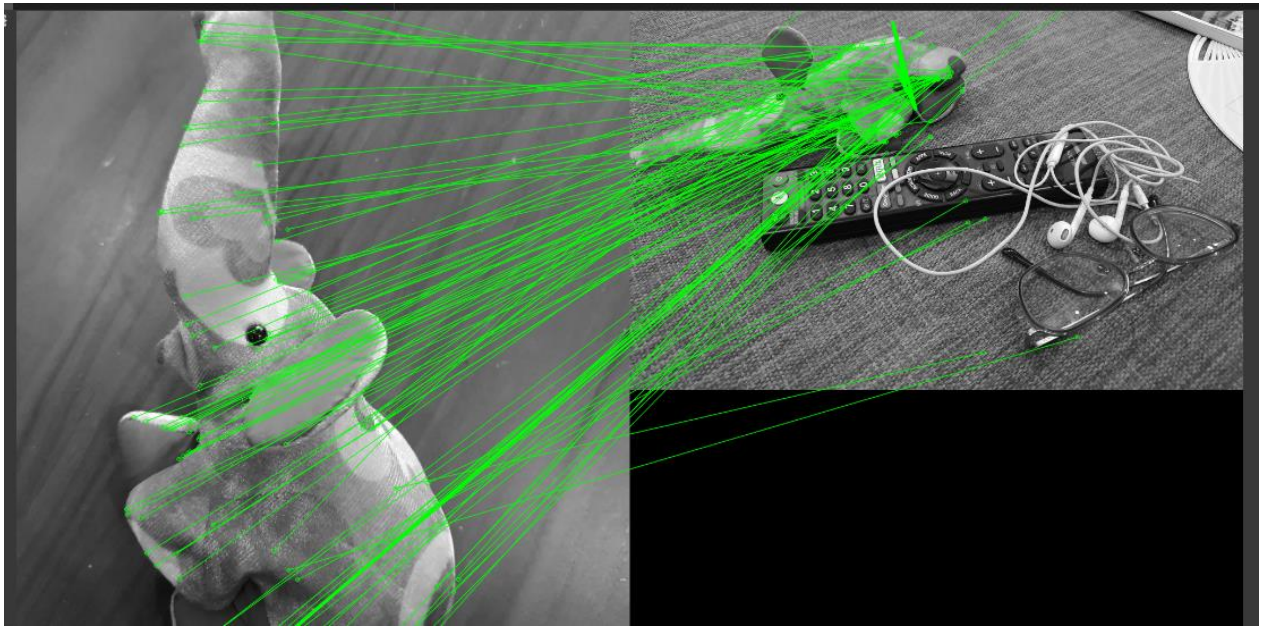
### 7.4. L'éléphant

Seuil de ratio : 10

Nombre maximal de correspondance : 10

Conslusion : Score de correspondance : 1 . On a une correspondance moyenne car il y a la présence de points d'interet faiblesdes





## 8. Conclusion

En conclusion, notre approche a démontré une utilisation efficace des techniques SIFT pour la détection et la mise en correspondance de points d'intérêt entre deux images. La méthode de transformation homographique a été un outil précieux pour localiser l'objet de requête dans l'image de recherche.