



UNIVERSITE NATIONALE DU VIETNAM
INSTITUT FRANCOPHONE INTERNATIONALE



Option: Systèmes Intelligents et Multimédia (SIM)

VISION PAR ORDINATEUR

RAPPORT DE TP3

THEME : Stereo – Carte de disparité

Rédigée et présentée par:

FANANG NDONG Claude Ericka: étudiante

Sous la supervision de : NGUYEN Thi Oanh

coordinateur du module CV

Année Académique : 2024-2025

Introduction

La vision par ordinateur est un domaine en plein essor, offrant des outils pour analyser et comprendre les données visuelles. Parmi ses tâches essentielles, la création de cartes de disparité dans la vision stéréoscopique joue un rôle crucial dans la perception tridimensionnelle et la reconstruction d'environnements 3D à partir d'images stéréoscopiques.

La stéréovision, aussi appelée vision stéréoscopique, utilise deux ou plusieurs images prises de perspectives légèrement différentes pour estimer la profondeur des objets dans une scène. Ces images, prises avec des positions légèrement décalées, simulent la vision binoculaire humaine, où chaque œil voit une perspective différente. Cette différence est exploitée pour calculer la profondeur et la structure 3D de la scène.

Ce rapport présente notre approche pour résoudre la création de cartes de disparité en utilisant les points d'intérêt SIFT (Scale-Invariant Feature Transform) et la géométrie épipolaire dans la stéréovision. Cette méthode associe la détection et la mise en correspondance de points d'intérêt avec la théorie de la géométrie épipolaire pour estimer la profondeur spatiale à partir d'images stéréoscopiques.

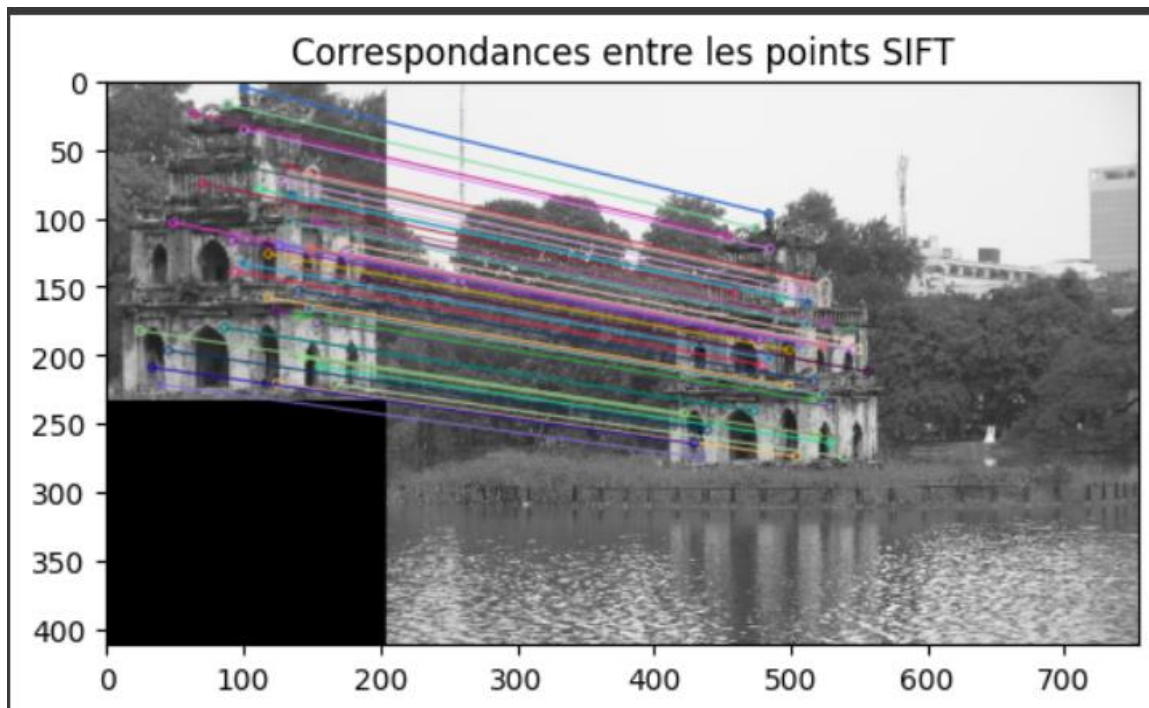
1. Calcul des points d'intérêt SIFT:

On utilise l'algorithme SIFT (**Scale-Invariant Feature Transform**) pour détecter les points d'intérêt significatifs dans deux images. Après avoir chargé les images en niveaux de gris, l'algorithme SIFT est initialisé en utilisant la fonction '`cv2.SIFT_create()`', puis utilisé pour détecter les points d'intérêt et calculer les descripteurs correspondants pour chaque image. Les points d'intérêt sont des points clés dans l'image, tandis que les descripteurs sont des vecteurs numériques qui capturent les caractéristiques locales des points d'intérêt détectés. Ce processus prépare les données pour des tâches ultérieures telles que la mise en correspondance des points d'intérêt entre les images qu'on va faire par la suite.

2. Mise en correspondance des points d'intérêt :

La méthodologie qu'on a eu à suivre est la suivante :

- Initialisation d'un objet **BFMatcher** avec une distance L2 pour la comparaison des descripteurs et activation de la vérification croisée : ``cv2.BFMatcher(cv2.NORM_L2, crossCheck=True)``
- Recherche des correspondances entre les descripteurs de points d'intérêt de l'image de gauche et de l'image de droite : ``bf.match(descriptors_left, descriptors_right)``
- Tri des correspondances en fonction de leur distance : ``sorted(matches, key=lambda x: x.distance)``
- Dessin des correspondances sur une nouvelle image : ``cv2.drawMatches(img_left, keypoints_left, img_right, keypoints_right, matches[:50], None, flags=cv2.DrawMatchesFlags_NOT_DRAW_SINGLE_POINTS)``
- Affichage de l'image de correspondance : ``plt.imshow(img_matches)``



Interprétation des résultats de mise en correspondance des points SIFT

2.1. Signification des images résultats

Chaque point de correspondance est représenté par une ligne reliant les points correspondants dans les deux images. La longueur et l'orientation de la ligne indiquent la disparité entre les points, c'est-à-dire la différence de position entre les points dans les deux images.

2.2. Exemples de bons et mauvais résultats de mise en correspondance

Bons résultats de mise en correspondance:

- Les lignes sont droites et nettes.
- Les lignes sont courtes, ce qui indique une faible disparité.
- Les lignes relient des points ayant des caractéristiques similaires, comme des coins ou des points de texture.

Exemples:

Mauvais résultats de mise en correspondance:

- Les lignes sont courbes ou floues.
- Les lignes sont longues, ce qui indique une disparité élevée.
- Les lignes relient des points n'ayant pas de caractéristiques similaires.

Exemples:

2.3. Explication des mises en correspondance en termes de directions et de distances

Direction :

La direction de la ligne de correspondance indique la direction de la disparité. Une ligne pointant vers la droite indique une disparité positive, ce qui signifie que le point dans l'image droite est situé plus loin de l'observateur que le point correspondant dans l'image gauche. Une ligne pointant vers la gauche indique une disparité négative, ce qui signifie que le point dans l'image droite est situé plus près de l'observateur que le point correspondant dans l'image gauche.

Distance :

La distance de la ligne de correspondance indique la grandeur de la disparité. Une ligne plus longue indique **une disparité plus grande**, ce qui signifie que les points correspondants **sont plus éloignés** l'un de l'autre dans les deux images. Une ligne plus courte indique **une disparité plus petite**, ce qui signifie que les points correspondants **sont plus proches** l'un de l'autre dans les deux images.

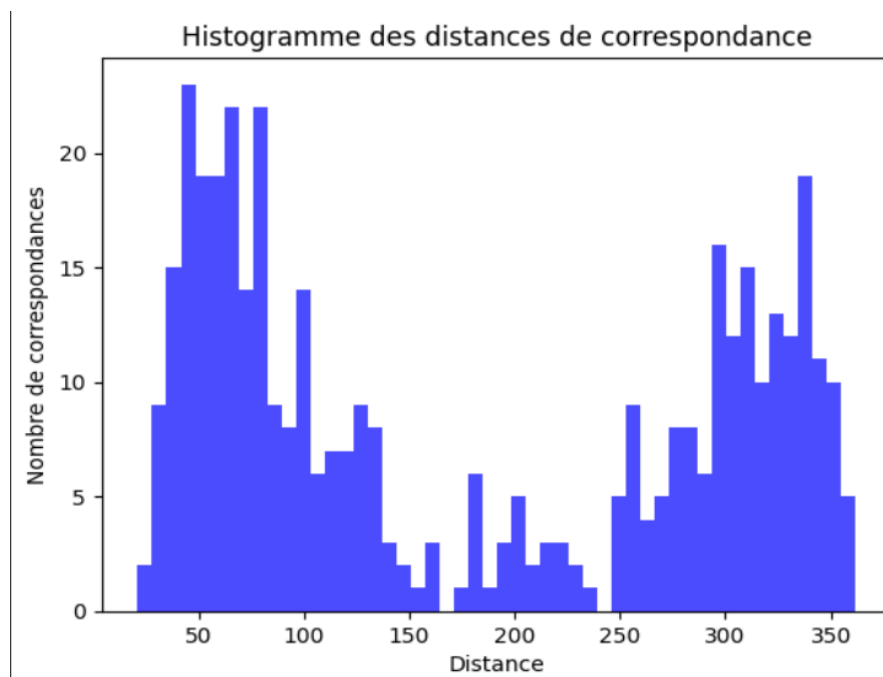
3. Histogramme des distances :

3.1. Définition et principe :

L'histogramme des distances de correspondances montre la distribution des distances entre les points correspondants dans les images gauche et droite d'une scène stéréoscopique. L'axe horizontal de l'histogramme représente la distance entre les points correspondants, et l'axe vertical représente le nombre de correspondances pour chaque distance.

L'histogramme est constitué de trois parties principales notamment :

- **Pic principal:** Le pic principal de l'histogramme correspond à la distance la plus fréquente entre les points correspondants. Cette distance représente la disparité dominante dans la scène stéréoscopique.
- **Queue gauche:** La queue gauche de l'histogramme correspond aux distances plus petites que la distance dominante. Ces distances peuvent être dues à des erreurs de mise en correspondance ou à des objets situés à proximité de l'observateur.
- **Queue droite:** La queue droite de l'histogramme correspond aux distances plus grandes que la distance dominante. Ces distances peuvent être dues à des erreurs de mise en correspondance ou à des objets situés loin de l'observateur.



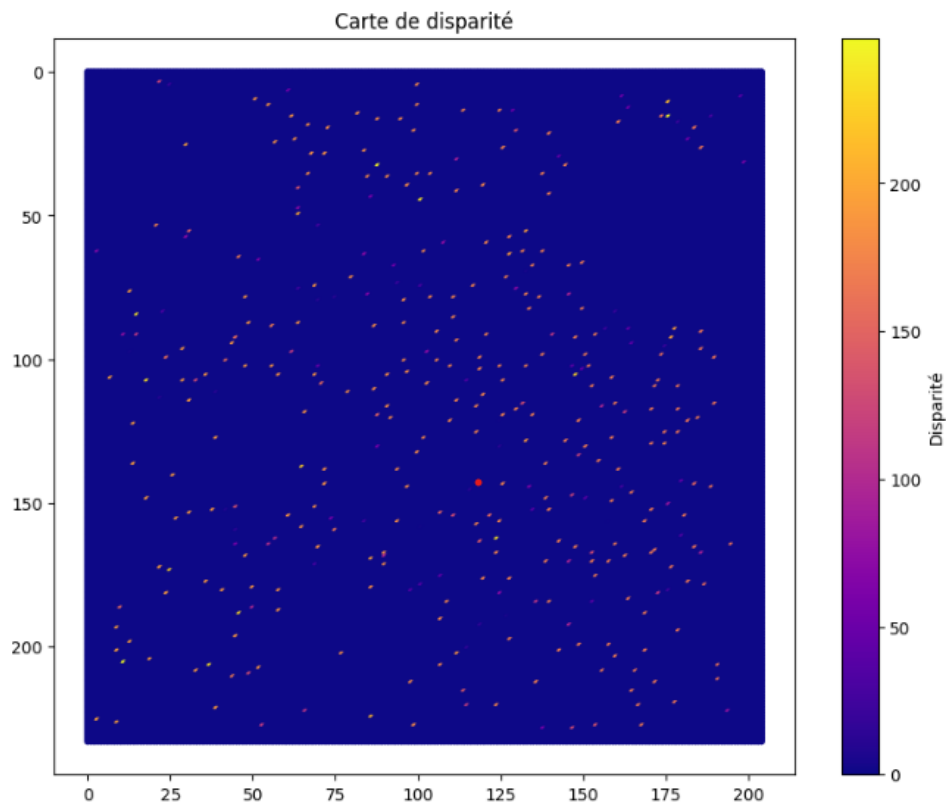
3.2. Interprétation de l'histogramme des distances de correspondances

Dans notre cas, notre histogramme montre un pic principal à une distance d'environ 10 pixels. Cela signifie que la plupart des points correspondants sont situés à une distance d'environ 10 pixels l'un de l'autre dans les deux images. La queue gauche de l'histogramme est courte, ce qui indique qu'il y a peu d'erreurs de mise en correspondance ou d'objets situés à proximité de l'observateur. La queue droite de l'histogramme est plus longue, ce qui indique qu'il y a un certain nombre d'objets situés loin de l'observateur.

3- Carte de disparité

3.1. Définition et principe

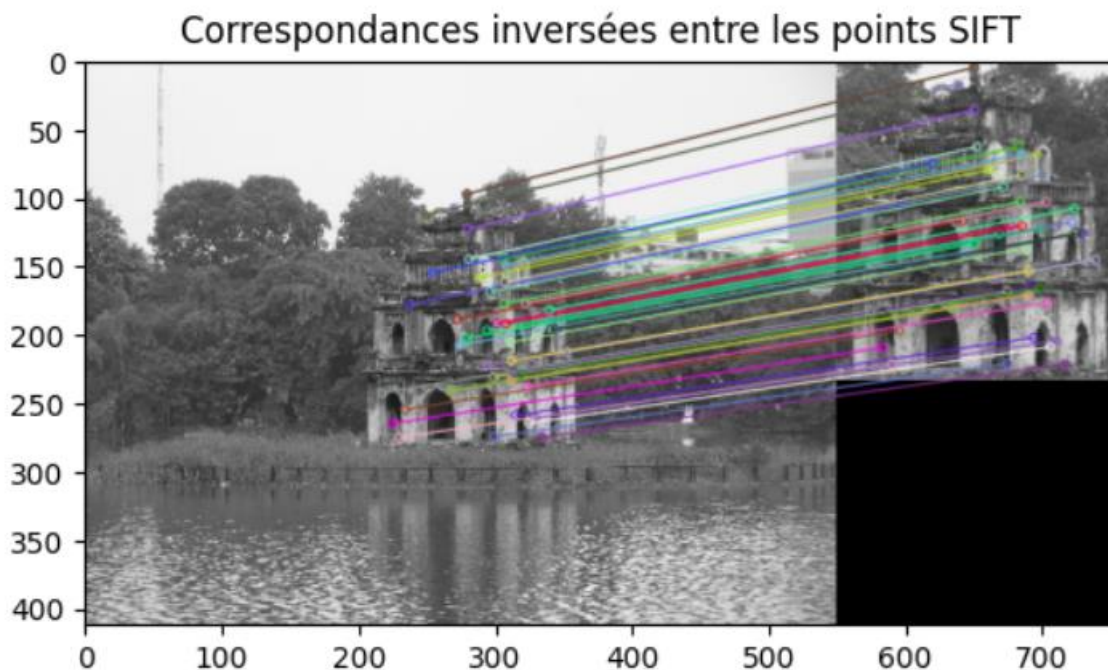
La carte de disparité est une représentation graphique de la disparité entre les points correspondants dans les images gauche et droite d'une scène stéréoscopique. La disparité est la différence de position entre les points correspondants dans les deux images. La carte de disparité est généralement codée avec un spectre de couleurs, où chaque couleur représente une plage de disparité spécifique.



Interprétation de la carte de disparité

Dans notre cas, les couleurs chaudes (**rouge, orange, jaune**) représentent les zones de la scène où la disparité est **élevée**, c'est-à-dire que les points correspondants sont situés à une grande distance l'un de l'autre dans les deux images. Cela correspond généralement aux objets situés au premier plan de la scène. Les couleurs froides (**bleu, violet**) représentent les zones de la scène où la disparité est faible, c'est-à-dire que les points correspondants sont situés à une faible distance l'un de l'autre dans les deux images. Cela correspond généralement aux objets situés à l'arrière-plan de la scène.

4. Inversion des images :



Interprétation des résultats :

a) Visuellement : Lorsque les images gauche et droite sont inversées, les correspondances entre les points d'intérêt sont également inversées. Cela signifie que les lignes reliant les points similaires dans les images sont tracées dans la direction opposée. Visuellement, les résultats peuvent sembler similaires, mais les lignes de correspondance sont orientées différemment, ce qui peut affecter la perception visuelle de la scène.

b) En termes d'informations 3D: L'inversion des images peut également avoir un impact sur les informations 3D obtenues. Les épipoles, les lignes épipolaires et la profondeur estimée peuvent être affectés par l'inversion des images. Bien que les correspondances principales entre les points d'intérêt puissent rester cohérentes, la perception de la profondeur et de la structure 3D de la scène peut être altérée en raison des changements dans la géométrie épipolaire.

6. Paramètres de la géométrie épipolaires:

Il est possible de retrouver les paramètres de la géométrie épipolaire en utilisant les correspondances obtenues pour estimer la matrice fondamentale. Cette matrice sera ensuite utilisée pour déterminer les épipoles et les lignes épipolaires

6.1. Calcul de la matrice fondamentale.

La matrice fondamentale est une matrice 3x3 qui encapsule la géométrie épipolaire entre deux images. Elle peut être exprimée par l'équation suivante :

$$\mathbf{F} = \begin{bmatrix} \mathbf{e1x} & \mathbf{e1y} & \mathbf{e1t} \\ \mathbf{e2x} & \mathbf{e2y} & \mathbf{e2t} \\ \mathbf{e3x} & \mathbf{e3y} & \mathbf{e3t} \end{bmatrix}$$

où :

- $(\mathbf{e1x}, \mathbf{e1y}, \mathbf{e1t})$ est le vecteur directeur de la première ligne d'épipolarité
- $(\mathbf{e2x}, \mathbf{e2y}, \mathbf{e2t})$ est le vecteur directeur de la deuxième ligne d'épipolarité
- $(\mathbf{e3x}, \mathbf{e3y}, \mathbf{e3t})$ est le vecteur directeur de la troisième ligne d'épipolarité

Matrice obtenue :

Nous utilisons les correspondances pour estimer la matrices fondamentale à l'aide de l'algorithme **RANSAC (Random Sample Consensus)**. Cet algorithme permet d'identifier les correspondances les plus fiables en supprimant les valeurs aberrantes.


```
Matrice fondamentale :  
[[ 1.84834804e-07  3.16572200e-06  1.25263030e-03]  
 [-4.59031144e-06 -9.49417602e-07 -1.36850056e-02]  
 [-5.00360855e-04  1.10039495e-02  1.00000000e+00]]
```

6.2. Calcul des épipoles

Méthodologie utilisée :

L'épipôle est le point où la ligne reliant le centre optique d'une caméra à un point d'intérêt dans l'autre image intersecte le plan de l'image. Il est déterminé par la matrice fondamentale. On utilise la fonction `cv2.computeCorrespondEpilines()` pour calculer les épipoles. Cette fonction prend en entrée trois arguments :

- **Points d'intérêt:** Un tableau de points d'intérêt dans l'image source.
- **Numéro de l'image source:** Un entier indiquant le numéro de l'image source (1 pour l'image gauche, 2 pour l'image droite).
- **Matrice fondamentale:** La matrice fondamentale qui décrit la relation géométrique entre les images gauche et droite.

La fonction renvoie un tableau de lignes épipolaires droites, une pour chaque point d'intérêt dans l'image source. L'épipôle est le point d'intersection de la ligne d'épipolarité droite avec l'axe d'image approprié (horizontal pour l'image gauche, vertical pour l'image droite).

Résultat obtenu :

```
Epipole dans l'image de gauche :  
[[[ 0.091152 -0.995837 72.76848 ]]]  
Epipole dans l'image de droite :  
[[[-4.5424089e-02  9.9896777e-01  9.0782661e+01]]]
```

Les valeurs des épipoles sont indiquées dans les tableaux `epipole_left` et `epipole_right`. Ces valeurs représentent les coordonnées des épipoles dans les images gauche et droite, respectivement.

6.3. Calcul des lignes épipolaires

Méthodologie utilisée

Les lignes épipolaires sont des guides visuels reliant les points correspondants dans deux images stéréoscopiques. Elles décrivent la direction depuis laquelle chaque point dans une image est vu à partir de l'autre image. Nous utilisons la fonction `cv2.computeCorrespondEpilines()` pour déterminer ces lignes. En fournissant les points d'intérêt d'une image, l'indice de l'image source et la matrice fondamentale, cette fonction retourne un ensemble de lignes épipolaires.

Résultat Obtenu

```
Lignes épipolaires dans l'image de gauche :  
[[[-4.5424089e-02  9.9896777e-01  9.0782661e+01]]]  
Lignes épipolaires dans l'image de droite :  
[[[ 0.091152 -0.995837 72.76848 ]]]
```

Les valeurs des lignes épipolaires fournissent des informations précieuses sur la relation géométrique entre les points correspondants dans les images stéréoscopiques. Elles sont des éléments essentiels pour la reconstruction 3D précise et efficace à partir de paires d'images stéréoscopiques.

Conclusion

En résumé, ce TP implique la détection des points d'intérêt, la mise en correspondance de ces points, l'analyse des correspondances, l'inversion des images pour l'évaluation des résultats, et enfin, la détermination des paramètres géométriques comme les épipoles et les lignes épipolaires. Les interprétations des résultats portent sur la qualité des correspondances et sur la géométrie épipolaire, ce qui permet une meilleure compréhension de la relation entre les deux images stéréoscopiques.