



Université  
de Lille  
1 SCIENCES  
ET TECHNOLOGIES

Université de Lille  
**MASTER 2 INFORMATIQUE**  
Premier semestre



# Vision artificielle

## TP n°02

Mise en correspondance stéréoscopique

**BARCHID Sami**

# Introduction

---

La matière vue dans ce TP traite de la mise en correspondance stéréoscopique. Le but de ce TP est d'appréhender la mise en correspondance stéréoscopique, une technique qui sert pour l'une des formes de reconstruction 3D vues en cours, à savoir la stéréovision.

La stéréovision est une technique de reconstruction 3D qui reconstitue la troisième dimension de la scène à partir de deux images prises à deux points de vue différents (en vision passive). Elle consiste en deux étapes principales :

1. Recherche des points homologues dans les deux images par **mise en correspondance stéréoscopique**
2. Reconstruction 3D de la scène à partir des propriétés des points homologues trouvés grâce à l'étape précédente.

La première étape de mise en correspondance sera le coeur du travail à réaliser pour ce TP. Pour ce faire, j'implémenterai la méthode de mise en correspondance tout en présentant, dans ce rapport, les étapes réalisées.

Ce rapport comprend 5 parties :

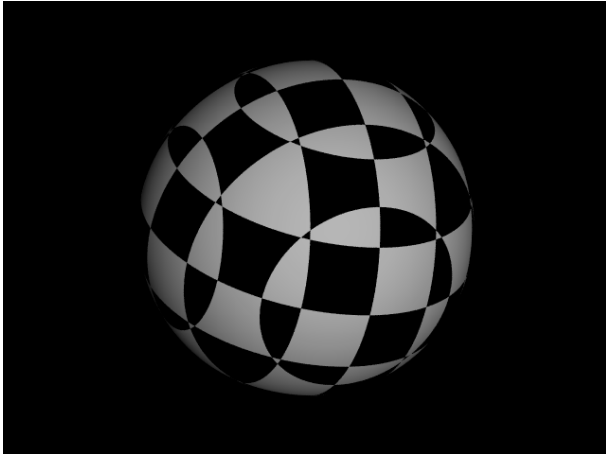
- **Expérience réalisée** : description de l'expérience effectuée pendant ce TP.
- **Calcul de la matrice fondamentale** : explication du calcul pour obtenir la matrice fondamentale du stéréoscope nécessaire pour la suite.
- **Extraction des coins** : description de la méthode utilisée pour récupérer des points caractéristiques de l'image (ici, des coins).
- **Calcul des distances** : description de l'utilité de calculer les distances entre chaque point caractéristique trouvé auparavant et explication de la méthode utilisée pour obtenir ces distances pour tous les points.

- **Mise en correspondance** : présentation de l'algorithme conçu pour obtenir la correspondance entre les points caractéristiques des deux images.

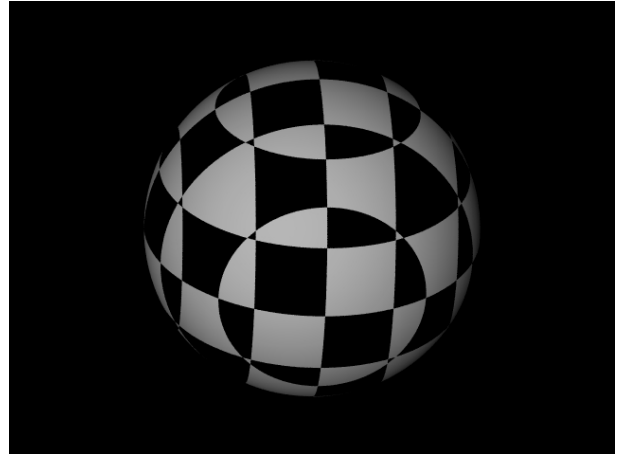
# Expérience réalisée

---

L'expérience est constituée de deux images d'une même scène prise au moyen d'un stéréoscope. Les deux images, nommées respectivement « left » et « right », sont les suivantes. Comme leurs noms l'indiquent, ces images représentent les images prises selon le capteur gauche et le capteur droit.



**Left** (point de vue gauche)



**Right** (point de vue droite)

Le but de l'expérience est de réaliser la mise en correspondance des points d'intérêts de **left** et de **right** pour pouvoir reconstruire la troisième dimension de la scène (cette reconstruction 3D n'est pas le but de ce tp, néanmoins).

Pour effectuer cette mise en correspondance, plusieurs opérations doivent être faites :

- Calculer la matrice fondamentale du stéréoscope
- Extraire les points d'intérêt de **left** et de **right**
- Calculer la distance entre chaque point d'intérêt de **left** et **right**
- Définir les points d'intérêts homologues grâce aux distances calculées auparavant.

# Calcul de la matrice fondamentale

---

Dans cette partie du TP, je vais expliciter le calcul nécessaire pour trouver la matrice fondamentale.

## Déterminer la matrice fondamentale

Nous disposons des paramètres intrinsèques et extrinsèques des caméras gauches et droites. Ces paramètres sont notés :

- Matrice des paramètres intrinsèques pour la caméra gauche :  $\mathbf{A}_1$
- Matrice des paramètres extrinsèques pour la caméra gauche :  $(\mathbf{R}_1 | \mathbf{t}_1)$
- Matrice des paramètres intrinsèques pour la caméra droite:  $\mathbf{A}_2$
- Matrice des paramètres extrinsèques pour la caméra droite :  $(\mathbf{R}_2 | \mathbf{t}_2)$

Nous pouvons utiliser ces paramètres des caméras pour trouver la matrice fondamentale.

On peut également estimer la projection d'un point  $\mathbf{M}$  de la scène sur les deux plans projectifs  $\pi_1$  (plan projectif de l'image **left**) et  $\pi_2$  (plan projectif de l'image **right**) :

$$\mathbf{m}_1 \equiv \mathbf{P}_1 \mathbf{M} = \mathbf{A}_1 (\mathcal{R}_1 | \mathbf{t}_1) \mathbf{M}$$

$$\mathbf{m}_2 \equiv \mathbf{P}_2 \mathbf{M} = \mathbf{A}_2 (\mathcal{R}_2 | \mathbf{t}_2) \mathbf{M}$$

où

- $\mathbf{m}_1$  est la projection de  $\mathbf{M}$  sur  $\pi_1$

- $\mathbf{m}_2$  est la projection de  $\mathbf{M}$  sur  $\pi_2$
- $\mathbf{P}_1$  est la matrice de projection d'un point  $\mathbf{M}$  sur  $\pi_1$ . On peut donc conclure que  $\mathbf{P}_1 = \mathbf{A}_1(R_1 | \mathbf{t}_1)$ .
- $\mathbf{P}_2$  est la matrice de projection d'un point  $\mathbf{M}$  sur  $\pi_2$ . On peut donc conclure que  $\mathbf{P}_2 = \mathbf{A}_2(R_2 | \mathbf{t}_2)$ .

D'autre part, pour chacune des caméras du stéréoscope, nous pouvons définir le centre de projection  $\mathbf{O}_i$  comme étant la **dernière colonne de l'inverse de la matrice des paramètres extrinsèques de la caméra i**. Grâce à ceci, nous connaissons  $\mathbf{O}_1$  et  $\mathbf{O}_2$ .

Avec ces données, nous pouvons alors opérer le calcul de la matrice fondamentale à partir des matrices de projection  $\mathbf{P}_1$  et  $\mathbf{P}_2$ . Le calcul est le suivant :

$$\mathbf{F} = (\mathbf{P}_2 \mathbf{O}_1)^\times H_\pi = (\mathbf{P}_2 \mathbf{O}_1)^\times \mathbf{P}_2 \mathbf{P}_1^{+\zeta}$$

où

- $\mathbf{F}$  est la matrice fondamentale
- $\mathbf{P}_2 \mathbf{O}_1$  est la projection de  $\mathbf{O}_1$  sur le plan projectif  $\pi_2$
- $H_\pi = \mathbf{P}_2 \mathbf{P}_1^{+\zeta}$  est l'homographie du plan  $\pi_M$  qui transforme  $\mathbf{m}_1$  en  $\mathbf{m}_2$ .
  - Note :  $\mathbf{P}^{+\zeta}$  est la pseudo inverse d'une matrice  $\mathbf{P}$ .

On remarque, dans la formule de la matrice fondamentale, une expression  $(\mathbf{P}_2 \mathbf{O}_1)^\times$  qui correspond à une matrice de produit vectoriel. Nous pouvons calculer cette matrice de produit vectoriel grâce à la formule suivante :

$$\mathbf{p}^{\times} = \begin{pmatrix} 0 & -p_z & p_y \\ p_z & 0 & -p_x \\ -p_y & p_x & 0 \end{pmatrix}$$

Avec la matrice fondamentale connue, il est possible de trouver l'équation de la droite épipolaire  $\mathbf{d}_2$  de l'image right (  $\pi_2$  ) associée à un point de l'image gauche  $\mathbf{m}_1$  et l'équation de la droite épipolaire  $\mathbf{d}_1$  de l'image left (  $\pi_1$  ) associée à un point de l'image droite  $\mathbf{m}_2$  :

$$\mathbf{d}_2 = \mathbf{F}\mathbf{m}_1$$

$$\mathbf{d}_1 = \mathbf{F}^T\mathbf{m}_2$$

## Résumé de l'algorithme

On peut alors expliciter les étapes successives à implémenter dans le code pour retrouver la matrice fondamentale.

### Entrée

- $A_1$
- $A_2$
- $(\mathbf{R}_1|\mathbf{t}_1)$
- $(\mathbf{R}_2|\mathbf{t}_2)$

## Sortie

$F$  , la matrice fondamentale

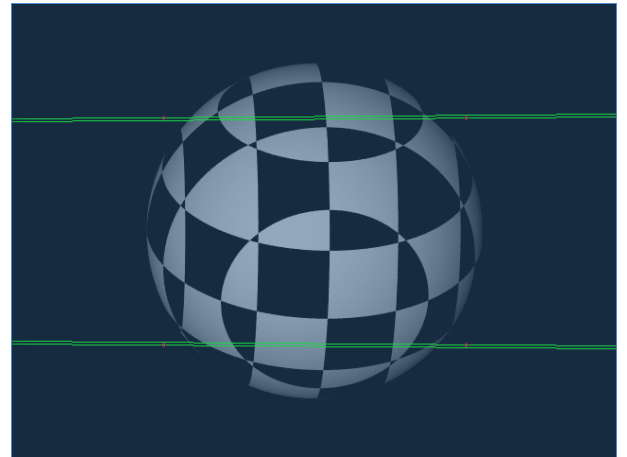
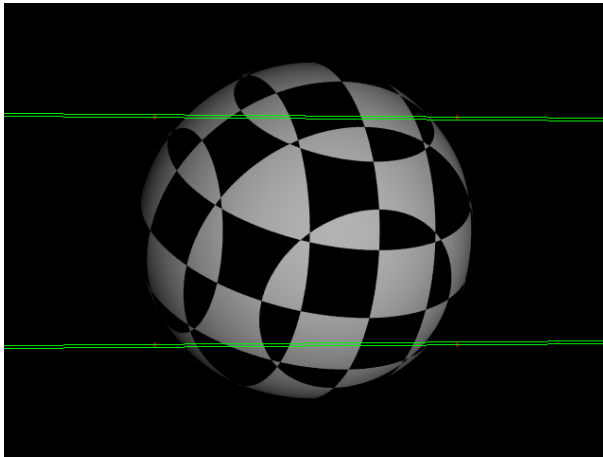
## Algorithme

- Calculer  $P_2 = A_2(R_2|t_2)$
- Calculer  $P_1 = A_1(R_1|t_1)$
- Calculer  $P_1^{+i}$  , la pseudo-inverse de  $P_1$
- Calculer  $H_\pi = P_2 P_1^{+i}$
- Calculer  $O_1$  qui est la dernière colonne de l'inverse de la matrice des paramètres extrinsèques  $(R_1|t_1)$  .
- Calculer  $(P_2 O_1)^x$  avec la formule de la matrice de produit vectoriel mentionnée précédemment.
- Calculer  $F = (P_2 O_1)^x H_\pi$  avec toutes les données que l'on a calculées avant.

## Résultats dans l'expérience

Il est alors possible de vérifier l'application de cette matrice fondamentale en traçant des droites épipolaires de points homologues quelconques entre les deux images de la scène. (**Note importante:** les points quelconques que je décris ont été fournis à l'initialisation du TP pour nous permettre de tester notre code).





**Droites épipolaires des points quelconques pour l'image left et right**

On remarque que les droites sont alignées, montrant que les droites épipolaires pour des points homologues sont alignées.

# Extraction des coins

---

Cette partie du TP présente brièvement la prochaine étape dans le procédé de mise en correspondance stéréoscopique, à savoir l'extraction de points d'intérêts dans les deux images **left** et **right**.

Les points d'intérêts à extraire qui ont été choisis pour ce TP sont des **coins**, c'est-à-dire des points pour lesquels le voisinage présente une grande variation d'intensité dans toutes les directions.

La méthode pour extraire ces coins est la méthode de **Shi et Tomasi**. Néanmoins, puisque la priorité dans ce rapport est de présenter la méthode de mise en correspondance stéréoscopique, seulement un résumé de la méthode de **Shi et Tomasi** sera énoncé.

## Méthode de Shi et Tomasi

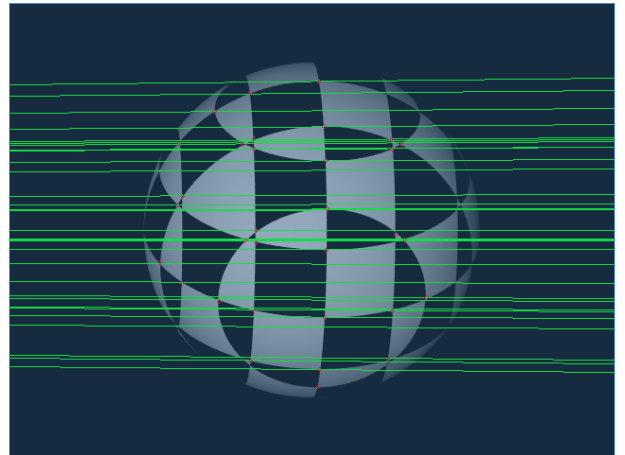
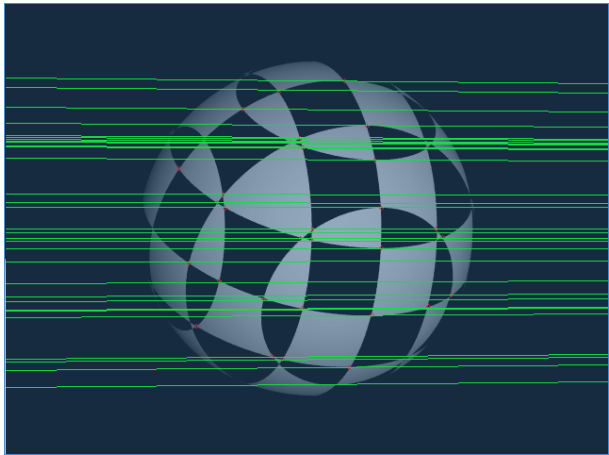
La méthode de Shi et Tomasi est une amélioration de la méthode de détection de coins de Harris.

Cette méthode de Harris consiste à déterminer mathématiquement la taille de la fenêtre de voisinage d'un pixel qui produit la plus grosse variation d'intensité quand on la bouge dans n'importe quelle direction. Avec chaque fenêtre trouvée, un score  $R$  est calculé pour chaque point qui est candidat à être élu en tant que point. Puis, on retient les points candidats dont le score dépasse un certain seuil fixé. Ces points candidats qui passent le seuil sont donc les coins.

La méthode de Shi et Tomasi est en réalité quasiment la même chose que la méthode de détection de coin de Harris, à la chose près que le calcul du score  $R$  est différent et donne de meilleurs résultats.

## Résultats dans notre expérience

En appliquant la méthode de Shi et Tomasi sur nos images **right** et **left**, nous obtenons les coins suivant :



**Coins trouvés pour left et right (avec droites épipolaires)**

Dans ces images, nous pouvons voir les coins trouvés entourés en rouge sur chacune des images. Les droites épipolaires pour chaque point de chaque image ont aussi été tracées en vert.

On remarque bien que, pour chaque coin trouvé, on observe une grande variation d'intensité dans le voisinage du coin à chaque déplacement d'une fenêtre dans n'importe quelle direction. La méthode permet donc bien de retrouver des coins de manière efficace.

# Calcul des distances

---

Cette partie du TP explicite l'étape suivante de la mise en correspondance stéréoscopique, à savoir le calcul de la distance entre chaque paire de points d'intérêt (un point d'intérêt de l'image **left** et un point d'intérêt de l'image **right**).

## Calcul de la distance entre deux points d'intérêts

La distance entre un point d'intérêt de **left** et un point d'intérêt de **right** est la somme des deux distances euclidiennes suivante :

- La distance euclidienne entre un point  $\mathbf{m}_1$  de l'image **left** et la droite épipolaire  $\mathbf{d}_1$  de l'image **left** associée à un point  $\mathbf{m}_2$  de l'image **right**.
- La distance euclidienne entre un point  $\mathbf{m}_2$  de l'image **right** et la droite épipolaire  $\mathbf{d}_2$  de l'image **right** associée à un point  $\mathbf{m}_1$  de l'image **left**.

Si on veut reformuler plus simplement cela, on peut dire que :

- Nous avons un point d'intérêt  $\mathbf{m}_1$  de l'image **left** et un point d'intérêt  $\mathbf{m}_2$  de l'image **right**
- On peut obtenir les droites épipolaires  $\mathbf{d}_1$  et  $\mathbf{d}_2$  grâce au calcul suivant (déjà expliqué précédemment)

$$\mathbf{d}_2 = \mathbf{F}\mathbf{m}_1$$

$$\mathbf{d}_1 = \mathbf{F}^T\mathbf{m}_2$$

- On calcule la distance euclidienne entre  $\mathbf{m}_1$  et  $\mathbf{d}_1$

- On calcule la distance euclidienne entre  $m_2$  et  $d_2$
- On fait la somme des deux distances

Il faut faire ces étapes pour chaque paire de points d'intérêt possible.

La distance entre un point A et une droite (d) se calcule de la manière suivante :

$$d(A, (d)) = \frac{|ax_A + by_A + c|}{\sqrt{a^2 + b^2}}$$

Avec  $(d)$  qui a pour équation  $ax + by + c = 0$  et le point A a pour coordonnées  $(x_a, y_a)$ .

Dans le cadre de nos calculs, nous possédons déjà les paramètres  $a, b, c$  des droites  $d_1$  et  $d_2$ . Même chose pour les coordonnées  $(x_A, y_A)$  de  $m_1$  et  $m_2$ . On peut donc calculer la somme des distances.

Toutes les distances entre chaque paire de point d'intérêt servent à créer une matrice des distances composée de la manière suivante :

- Un élément  $d_{ij}$  de la matrice des distances correspond à la distance entre le point d'intérêt  $i$  de l'image **left** et le point d'intérêt  $j$  de l'image **right**.
- La matrice de distance est donc une matrice  $n_{\text{left}} \times n_{\text{right}}$  où
  - $n_{\text{left}}$  est le nombre de points d'intérêts pour l'image **left**
  - $n_{\text{right}}$  est le nombre de points d'intérêts pour l'image **right**

Cette matrice des distances sera utilisée en entrée pour la prochaine étape de la mise en correspondance.

# Mise en correspondance

---

Cette étape est la dernière de la mise en correspondance stéréoscopique et sert à trouver pour chaque point d'intérêt de l'image **left**, son point d'intérêt homologue de l'image **right**.

Pour ce faire, un algorithme a été utilisé qui permet d'estimer, avec l'aide de la matrice des distances, si deux points sont suffisamment similaires pour pouvoir être associés.

## Description de l'algorithme

### Entrée

- $M_d$  , la matrice des distances définie dans la partie précédente
- $d_{max}$  Un seuil maximal de distance. Deux points ne peuvent être associés si la distance entre eux est plus grande que ce seuil maximal.

### Sortie

- $Homologues_{left}$  , la liste des points homologues de l'image **left**.
- $Homologues_{right}$  , la liste des points homologues de l'image **right**.

Ces deux listes contiennent les indices des points homologues pour leur image respective tel que l'on peut retrouver la  $i^{ème}$  paire de points homologues du stéréoscope grâce à

$Homologues_{left}[i]$  et  $Homologues_{right}[i]$  .

## Réalisation

```
TANT QUE [ il existe une distance  $< d_{max}$  dans  $M_d$  ] :  
|    $d_{ij}$  = la distance minimale de  $M_d$   
|   // Rappel :  $d_{ij}$  contient la distance entre le point  
|   //  $i$  de left et le point  $j$  de right  
|   SUPPRIMER  $i^{ème}$  ligne de  $M_d$   
|   SUPPRIMER  $j^{ème}$  colonne de  $M_d$   
|   AJOUTER indice  $i$  à  $Homologues_{left}$   
|   AJOUTER indice  $j$  à  $Homologues_{right}$   
FIN TANT QUE
```

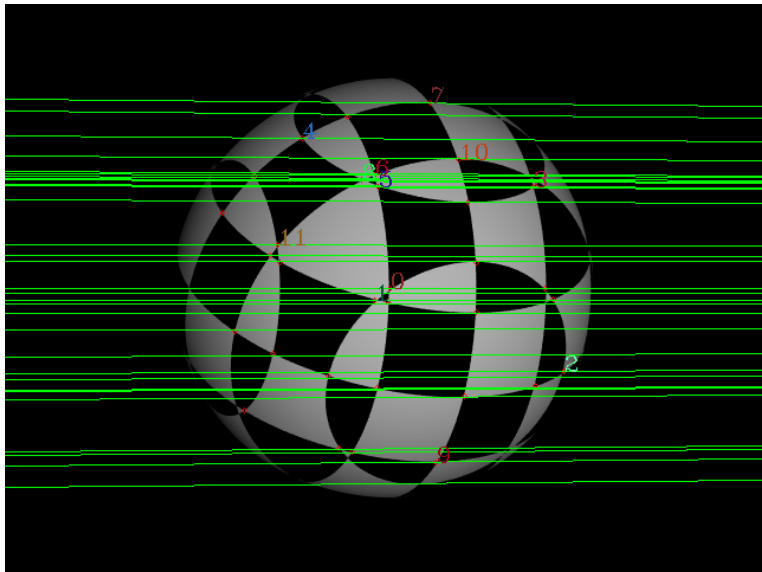
## Analyse des résultats

Pour analyser la mise en correspondance, le but sera de faire varier le seuil de distance maximal pour comprendre les résultats obtenus.

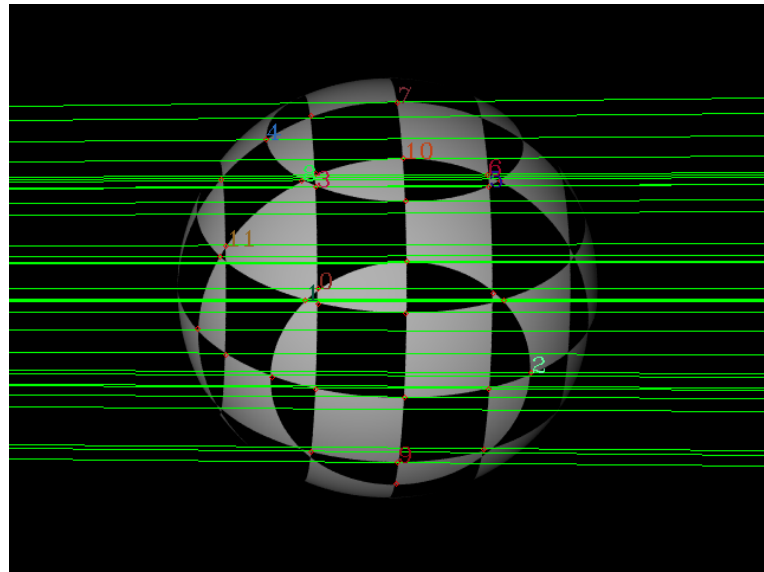
Ces résultats seront mesurés, pour chaque seuil testé, de la manière suivante :

- Le nombre de paires de points homologues
- Le nombre de points occultés (= sans homologue) sur l'image **left**
- Le nombre de points occultés (= sans homologue) sur l'image **right**
- Un schéma montrant, dans l'image, les points homologues obtenus. Chaque paire de point homologue se verra attribué un numéro et une couleur au hasard, permettant de retrouver cette paire dans les deux images.

*Seuil maximal = 1*



*left*



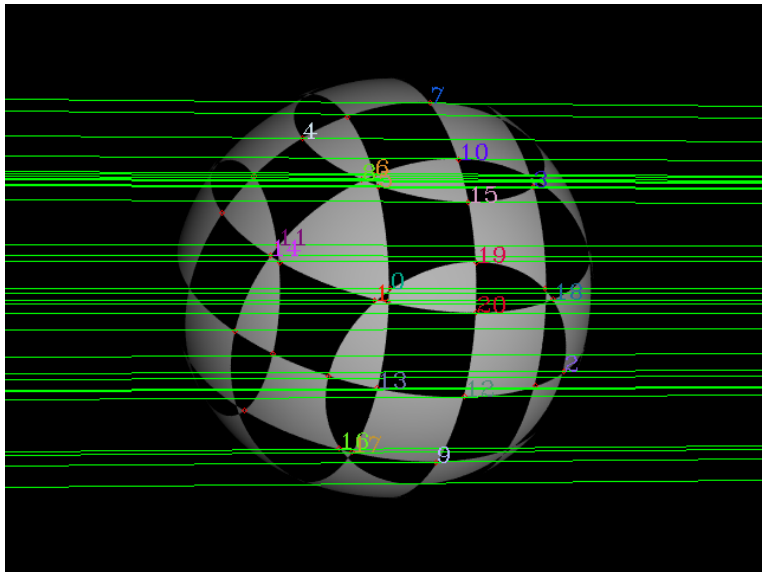
*right*

|                                    |    |
|------------------------------------|----|
| <b>Paires de points homologues</b> | 12 |
| <b>Points occultés</b>             | 20 |

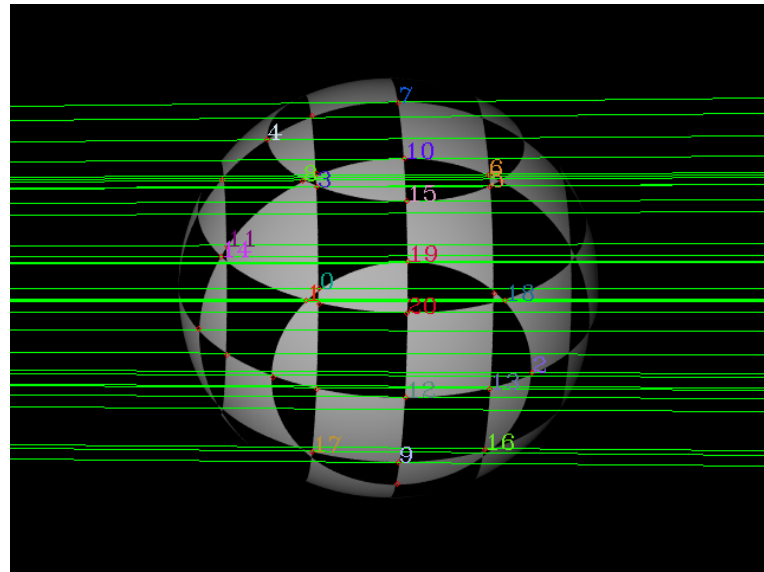
- On remarque une grande quantité de points occultés pour les deux images (et donc peu de paires de points homologues trouvées).
- Les paires de points sont en majeure partie exactes. Je n'ai relevé qu'une seule paire de points homologues faussée : la paire n°6 sur l'image.



*Seuil maximal = 2*



left

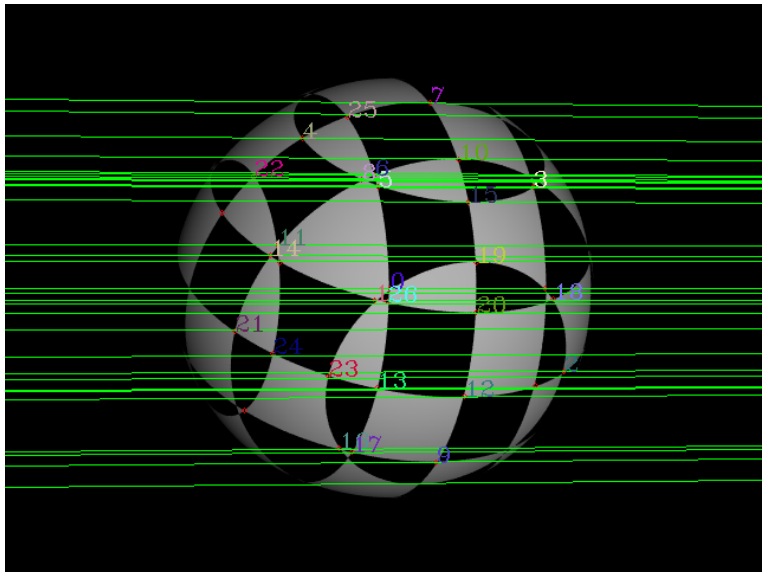


right

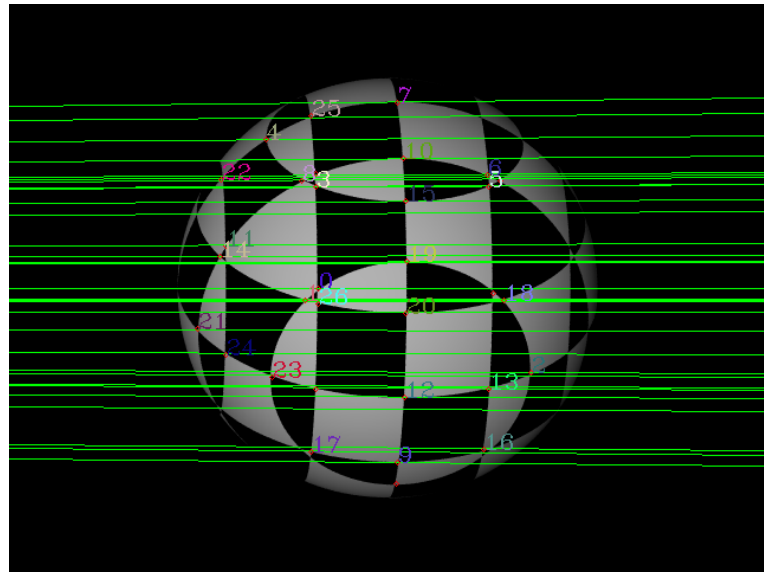
|                                    |    |
|------------------------------------|----|
| <b>Paires de points homologues</b> | 21 |
| <b>Points occultés</b>             | 11 |

- Il y a déjà plus un plus grand nombre de points homologues trouvés (et donc moins de points occultés).
- Un peu plus de paires de points homologues faussées (ex : les paires 6, 5 et 3). Cependant, la grande majorité des paires trouvées est correcte.

*Seuil maximal = 5*



left

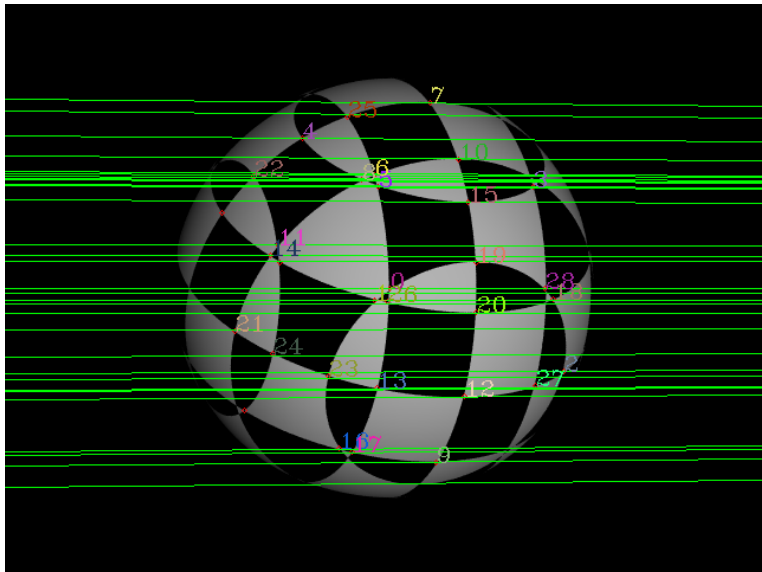


right

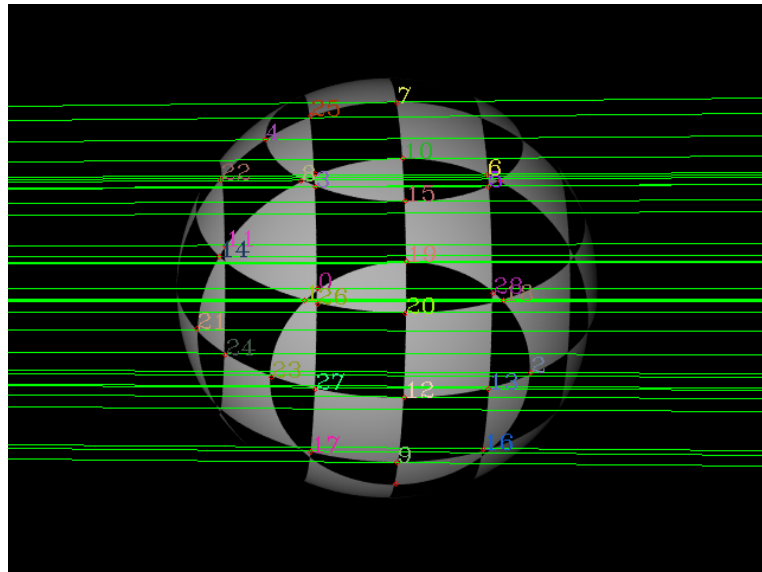
|                                    |    |
|------------------------------------|----|
| <b>Paires de points homologues</b> | 27 |
| <b>Points occultés</b>             | 5  |

- Quasiment tous les points d'intérêts ont été admis dans des paires de points homologues.
- Il y a un très grand problème de fiabilité : beaucoup de paires de points homologues sont fausses.

*Seuil maximal = 50*



**left**



**right**

|                                    |    |
|------------------------------------|----|
| <b>Paires de points homologues</b> | 29 |
| <b>Points occultés</b>             | 3  |

- On remarque que, par rapport au seuil max de 5, très peu de choses changent, ce qui montre que le seuil (pourtant déjà jugé trop grand à cause du peu de fiabilité observée) ne suffit pas à modifier le résultat obtenu auparavant.

## *Interprétation*

On remarque plusieurs choses face aux résultats obtenus :

- Un seuil maximal réduit occulte un grand nombre de points feature précédemment calculés mais permet d'obtenir des points homologues fortement similaires (vu que la distance entre ces deux points est faible)
- Un seuil maximal élevé permet de retrouver beaucoup de paires de points homologues. Cependant, ces paires de points peuvent s'avérer fausses étant donné que la distance entre deux points associés peut être élevée (et donc correspondre à deux points qui ne sont pas homologues en réalité).
- Certains points d'intérêts sur **left** ou sur **right** sont simplement seuls parce que la méthode de Shi & Tomasi n'a pas trouvé le coin sur l'autre image.

On peut donc résumer cette variation du seuil de la distance par :

- Réduire le seuil maximal = augmenter la précision de la méthode mais réduire le nombre de paires de points trouvées.
- Augmenter le seuil maximal = réduire la précision (voire risquer d'admettre des erreurs) mais avoir un nombre de paires de points plus important.

# Conclusion

---

Pour conclure, nous avons pu, dans ce TP, comprendre et tester la mise en correspondance stéréoscopique, qui est une des grandes étapes de la stéréovision.

De plus, nous avons pu rapidement étudier une méthode pour extraire des points d'intérêt d'une image sous la forme de coins grâce à la méthode de Shi & Tomasi. Il est important de mentionner, tout de même, que l'on aurait pu prendre une autre méthode de sélections de points d'intérêts (comme par exemple, une méthode pour trouver des points de jonction en T).

Nous avons également vu que faire varier un seuil de distance maximal pouvait avoir des conséquences sur la précision/fiabilité des paires de points homologues trouvées ainsi que sur leurs nombres.