

# Questions/réponses

## Classe inversée

## Cours Vision

15 octobre 2025

### 1 Remarques générales

- **Si je n'ai pas répondu à votre question ou si la réponse n'est pas satisfaisante, vous avez le droit de la poser à nouveau.**
- Si certaines questions n'apparaissent pas c'est simplement parce qu'elles sont redondantes avec d'autres.
- Autant que possible, l'ordre d'apparition des questions/réponses correspond au déroulé du cours.

### 2 Questions/Réponses

1. *Pages 39, § 4.3.1 et 42, § 4.4.1* : le cours explique que la réponse du détecteur de Moravec est donnée par la valeur minimale de  $E_{(u,v)}(x, y)$  parmi les directions testées, et que Harris propose une amélioration pour répondre à quelques problèmes qui ont été identifiés. Cependant, pour SUSAN, il est mentionné que la réponse du détecteur SUSAN est basée sur la surface USAN centrée sur le pixel étudié, et que cette surface est comparée à un seuil géométrique  $T_g$ . Ma question porte sur le lien entre ces deux approches : Le détecteur de Harris utilise une approximation de la matrice des dérivées secondes pour évaluer comment chaque point est semblable à un coin, tandis que SUSAN utilise une surface circulaire et un seuil géométrique. Quelle est la différence entre ces deux méthodes lorsqu'une image contient des textures répétitives ou du bruit ? Laquelle serait théoriquement la plus robuste, et pourquoi ? Je crois comprendre que Harris utilise une approche analytique basée sur les dérivées, tandis que SUSAN utilise une approche géométrique basée sur la similarité des niveaux de gris. Mais je ne comprends pas clairement comment ces différences influencent la manière dont ces détecteurs sont robustes dans des conditions réelles (bruit, variations d'éclairage...). Est-ce qu'une de ces méthodes est généralement privilégiée selon le type d'application, ou faut-il simplement tester les deux pour choisir celle qui fonctionne le mieux ?

**Réponse :** Ces deux détecteurs sont très différents bien qu'ils considèrent tous les deux l'étude d'un voisinage. Dans le cas de Harris, on cherche le point le plus différent de ces voisins. Dans le cas de SUSAN, on cherche les centres des petites régions qui sont très différentes de leurs régions voisines. Dans le cas de Harris on utilise l'information de dérivée (et donc la notion de contours) alors que SUSAN n'exploite la photométrie. Alors, oui, je te conseille de toujours tester les deux pour identifier ou consolider celle que tu penses le plus adapté. Enfin, pour choisir il faut identifier si les points d'intérêt dont tu as besoin sont plutôt des points de type coins ou plutôt des petites régions très différentes de leur voisinage. Enfin, pour les invariances, aucun des deux n'en possèdent.

2. *Pages 60, équation 5.2* : Comment est défini le coût global  $E_{global}$  dans la mise en correspondance, et quel est le rôle du paramètre  $\lambda$  dans cette formulation ? Comment choisir sa valeur ?

**Réponse :** C'est la suite de ce chapitre qui détaille tous les éléments qui composent le coût global, donc, je ne suis pas sûre de comprendre la question. Le poids  $\lambda$  permet de pondérer l'importance des contraintes par rapport à l'importance du coût de correspondance. Comme indiqué page 62, la valeur de  $\lambda$  n'est pas souvent considérée. Ce que je voulais dire c'est que souvent  $\lambda = 0.5$  et chaque terme a le même poids.

3. *Page 60* : Quels types de contraintes peuvent être pris en compte dans le calcul du coût de correspondance ?

**Réponse :** Toutes celles qui sont décrites dans le chapitre dédié, le chapitre 7.

4. *Page 60* : Comment savoir quels groupes de pixels choisir ?

**Réponse :** J'ai supposé que la question porte sur le support  $\mathcal{S}$ . Cela va être choisi en fonction du type d'approches (locales ou globales) et des points considérés dans l'image (tous les points ou un sous ensemble).

5. *Pages 61-64* : Dans quelle mesure le choix entre une approche locale et une approche globale influence-t-il la robustesse du coût global de mise en correspondance en vue des situations extrêmes telles que les zones faiblement texturées, occultations ou variations photométriques ? Dans quels cas est-il préférable d'utiliser une approche locale de mise en correspondance plutôt qu'une approche globale ?

**Réponse :** Suivant la méthode choisie, la robustesse aux zones difficiles va être compliquée. Pour les zones faiblement texturées, une approche globale permettra de considérer plus de pixels et donc d'avoir un comportement plus global pour calculer l'appariement. En revanche, une approche locale type CENSUS sera également robuste à ce genre de difficulté. En ce qui concerne les occultations, les approches locales ont tendance à lisser les résultats sur les zones de discontinuités alors que les approches locales se comportent moins bien. C'est seulement en utilisant des mesures robustes que l'on peut s'approcher de la qualité des mesures de corrélation robuste.

6. *Page 65, § 5.4.3* : on fait l'hypothèse qu'une région de couleur homogène correspond à une surface de la scène pouvant être approximée par un modèle de disparité simple comme un plan. Quelle serait la limite de cette hypothèse dans les scènes réelles qui justifierait l'utilisation d'une sur-segmentation (mentionné dans le paragraphe 2 : Appariement en utilisant une segmentation en régions) plutôt qu'une segmentation classique ? Comment les cartes de régions et de disparités interagissent entre elles pour faire une meilleure estimation de la carte de disparités. Pourquoi on n'a pas besoin de faire une carte de région par image avant de passer à la disparité ?

**Réponse :** La limite c'est la taille de la région. Plus une région est étendue, plus l'hypothèse de planéité va être invalidée. Si on souhaite des régions de petites tailles alors l'approche par superpixels s'impose. Pour interagir, il y a différents niveaux : cela peut être de la fusion, des mécanismes de vote. Je reprends cette notion en cours. J'avoue ne pas avoir compris la seconde partie de la question.

7. *Pages 67* : Différentes mesures de similarité (MOR, SAD, GC, ISC, etc.) sont présentées. En dehors de l'appréciation visuelle des résultats, quelles métriques quantitatives permettent d'évaluer la précision ou de mesurer l'erreur d'une méthode de similarité, afin de comparer objectivement leurs performances.

**Réponse** : Nous pouvons utiliser tous les outils classiques d'évaluations : calculs du pourcentage de vrais positifs, vrais négatifs, faux positifs, faux négatifs. Mais cela signifie qu'il nous faut une vérité terrain.

8. *Page 70, § 6.6* : Pourquoi la direction du gradient, lorsqu'elle est de faible norme, est-elle une source d'erreurs pour la mise en correspondance, quelles sont les conséquences de cette fragilité sur le résultat final de l'appariement et quelle genre de données initiales peut donner une faible norme au gradient ?

**Réponse** : Un gradient de faible norme, cela veut dire que nous sommes en présence d'un point faiblement caractérisé et donc difficile à suivre pour la suite. Dès qu'une image présente des zones floues ou uniformes, la norme du gradient pourra en être affaiblie.

9. *Page 72* : En quoi les mesures robustes améliorent-elles la qualité des résultats d'appariement ? Pour les mesures robustes comme SMPD ou SMAD, le fait d'utiliser la médiane à la place de la moyenne rend la méthode moins sensible aux valeurs aberrantes et extrêmes. Mais du coup, est-ce que ça ne fait pas perdre un peu de précision dans les zones où il n'y a pas de bruit ou d'occultations ?

**Réponse** : Les mesures robustes sont, par définition, plus efficace dans les zones d'occultation ou de discontinuités. C'est normal car elles sont définies pour ça. En revanche comme on peut le voir dans les exemples de la figure 6.1, effectivement, les mesures robustes se comportent moins bien dans les zones simples d'une image.

10. *Pages 81* : Pourquoi dans la formule 7.8 la contrainte de rang ne prend en compte que les pixels avec des écarts en  $y$  et pas en  $x$  ?

**Réponse** : Je l'ai mentionné en cours mais pas dans le polycopié. Par soucis d'avoir des écritures simplifiées, je me place toujours dans le cas où les images sont rectifiées. Donc, les points correspondants sont sur les mêmes lignes de l'image. D'où le fait que seul  $y$  varie.

11. *Page 82, § 7.12* : Comment le seuil  $T_G$  (limite maximale du gradient de disparité) est-il traditionnellement choisi ou calculé dans la littérature et quelles sont les implications pratiques d'un  $T_G$  trop élevé ou trop faible sur la fiabilité et la densité de la carte de disparité obtenue ?

**Réponse** : Je n'ai pas trouvé de préconisations pour ce seuil et je pense qu'il faut faire un peu d'apprentissage pour le choisir.

12. *Page 85, § 8.1.2* : Comment la normalisation du vecteur descripteur et le seuillage des forts gradients permettent d'être invariant aux changements linéaires ou non d'illumination ? Comment le descripteur SIFT parvient-il à être invariant aux rotations et aux changements d'illumination, et quelle est l'importance de la normalisation du vecteur-descripteur ?

**Réponse** : Il n'y a pas d'invariance affine pour SIFT. Le fait de normaliser permet d'avoir des réponses comparables pour tous les points. SIFT n'est pas invariant affine. Le fait d'éliminer les forts

gradients permet de ne pas sélectionner de points sur les contours. En effet, ces points sont souvent les plus difficiles à apparier. Le détecteur est invariant à la rotation car on oriente les histogrammes descripteurs à partir de l'orientation principale. Je reprends cette notion en cours.

13. *Question d'ouverture* : On a remarqué qu'on sépare souvent la détection, la description (comme SIFT) et la mise en correspondance dans des étapes distinctes. Est-ce qu'il existe aujourd'hui des approches qui regroupent tout ça en un seul modèle ou une seule optimisation, pour éviter les pertes entre les étapes ?

**Réponse** : Pas à ma connaissance.

14. *Question d'ouverture* : Est-ce qu'il existe des contraintes supplémentaires liées à des applications différentes comme le tracking d'objet en vidéo ?

**Réponse** : Il y a la cohérence temporelle qui peut être ajoutée. On vérifie que l'évolution des déplacements sont cohérents les uns envers les autres.

15. *Question d'ouverture* : Est-ce qu'il y a des couples de contraintes connues qui fonctionnent bien ensemble ? Ou alors est-ce qu'il n'y a pas de cas connus, et donc selon le problème de mise en correspondance il faut prendre le temps de toutes les essayer ?

**Réponse** : Oui, par exemple la contrainte de symétrie + le seuillage de la réponse, que nous allons voir en TP. Mais, il est souvent conseillé de tester un ensemble de contraintes pour avoir une idée du comportement et de l'intérêt.