

**NOM et prénom :**

## **UE Vision, Réalité Augmentée et Applications**

**Examen - Documents autorisés : photocopie de cours distribué en vision - Durée 1h15**

20 décembre 2024

### **1 Vision (9 points)**

Cette matière est évaluée via des questions sur les parties théoriques et pratiques du cours puis sur des questions d'ouverture en lien avec les thématiques abordées en cours.

**Rappel :** La participation à la classe inversée de ce cours compte pour 1 points (ce qui permet d'avoir une note sur 10 points pour la partie vision).

#### **1.1 Questions liées aux cours et aux travaux pratiques (4 points)**

Pour cette série de questions, nous indiquons **en gras des solutions possibles** : entourez la réponse juste.

1. Pour calibrer la caméra, on peut au choix estimer les paramètres extrinsèques ou estimer les paramètres intrinsèques : **vrai** ou **faux** ?
2. Pour la détection de points d'intérêt, l'élément qui permet de caractériser une méthode est surtout lié au calcul de la carte de réponse : **vrai** ou **faux** ?
3. Le principe de base du détecteur de Harris est de mesurer la corrélation du point étudié à l'ensemble de ces points voisins : **vrai** ou **faux** ?
4. Comme il permet une analyse multi-échelle et multi-espace, le détecteur SIFT peut détecter plusieurs fois le même point d'intérêt : **vrai** ou **faux** ?
5. Le coût global de mise en correspondance contient toujours un coût de similarité et un coût des contraintes : **vrai** ou **faux** ?
6. La mise en correspondance locale permet une recherche de correspondant qui est localement exhaustive : **vrai** ou **faux** ?
7. Le problème des occultations lors de l'appariement est pris en compte lorsqu'on utilise une mesure de corrélation croisée, comme vue en TP : **vrai** ou **faux** ?
8. L'algorithme vu en TP et permettant d'estimer l'homographie afin de construire la mosaïque d'images est robuste aux erreurs d'appariement : **vrai** ou **faux** ?





## 2 Réalité augmentée

1. Pour cette série de questions entourez la réponse correcte et motivez brièvement votre réponse.

(a) Une fois qu'une caméra est calibrée, si l'on capture une nouvelle image, il est possible de déterminer directement le point 3D dans l'espace associé à un pixel de cette image : **vrai** ou **faux** ?

(b) La décomposition de l'homographie permet de récupérer l'échelle métrique de la scène : **vrai** ou **faux** ?

(c) La matrice fondamentale est une matrice  $3 \times 3$  toujours singulière : **vrai** ou **faux** ?

(d) La matrice fondamentale peut être utilisée pour estimer la pose relative entre deux caméras : **vrai** ou **faux** ?

(e) Les algorithmes de SLAM (“Simultaneous Localization And Mapping”) ne peuvent être utilisés que dans des environnements parfaitement connus et structurés : **vrai** ou **faux** ?

2. La société Pool7 a développé un système d'aide à la pratique du billard (voir [Figure 2](#)). Celui-ci se compose d'un projecteur suspendu au-dessus de la table, qui permet de visualiser la trajectoire de la bille par rapport à l'angle de la queue grâce à un jeu de lumières. Le système est capable de détecter les balles, de repérer la queue et de positionner la table dans l'espace afin de gérer les rebonds. Cet assistant de billard fonctionne en temps réel et indique en permanence la trajectoire que va adopter la boule de billard lorsqu'elle est frappée.



FIGURE 2 – L'application reconnaît la bille et l'angle de la queue, puis calcule la trajectoire de la bille et la projette sous forme de lumière sur la table.

- (a) Est-ce qu'on peut qualifier cette application comme une application de réalité augmentée au sens de la définition d'Azuma? **Argumenter la réponse** (si nécessaire, faire des hypothèses sur les fonctionnalités de l'application).

- (b) Quel(s) type(s) de « tracker » (inertiel, optique, *etc.*) est (sont) le(s) plus adapté(s) pour réaliser cette application? Justifier la réponse.

3. On dispose d'une caméra embarquée dans un satellite utilisée comme "Star Tracker" pour estimer l'**orientation** du satellite en détectant les étoiles dans le ciel. Les étoiles peuvent être modélisées comme des points à l'infini, et on suppose que les mouvements effectués par la caméra sont négligeables par rapport à la distance des étoiles (*i.e.*  $\mathbf{t} = 0$ ).

On voudrait être capable d'estimer la rotation de la caméra entre 2 images en utilisant l'image des étoiles. Pour cela on dispose d'un algorithme de traitement d'image ("Star Detection", SD) qui détecte les étoiles sur l'image et renvoie leurs coordonnées 2D en pixels dans le plan image de la caméra. Le détecteur SD donne uniquement les positions 2D des étoiles dans l'image, sans les reconnaître (*e.g.* Proxima Centauri, *etc.*), et on ne dispose donc d'aucune autre information que ces points 2D.

L'objectif de cet exercice est d'estimer la matrice de rotation  $\mathbf{R}$  qui relie les deux poses de la caméra à partir des correspondances entre les points d'une première image et ceux d'une seconde image.

On considère que la caméra est calibrée, avec une matrice intrinsèque  $\mathbf{K}_{3 \times 3}$  connue.

- (a) Une étape essentielle pour résoudre ce problème est d'établir la correspondance entre les points détectés dans les deux images. Parmi les approches possibles, quelle méthode entre "Tracking by Detection" et "Detect and Track" vous semble la plus adaptée dans ce cas? **Justifier la réponse**

- (b) Écrivons la projection d'une étoile sur le plan image. Soit  $\hat{\mathbf{Q}}_{4 \times 1} = \begin{bmatrix} \mathbf{Q} \\ 1 \end{bmatrix}$  le point 3D **inconnu** correspondant à une étoile dans l'espace, avec  $\mathbf{Q}_{3 \times 1}$  le vecteur des coordonnées 3D cartésiennes. Ce point est projeté sur le plan de la première image dans le point image 2D  $\mathbf{q}_1$  selon l'équation :

$$\mathbf{q}_1 \sim \mathbf{K} \begin{bmatrix} \mathbf{I} & | & \mathbf{0} \end{bmatrix} \hat{\mathbf{Q}} = \mathbf{K} \mathbf{Q}. \quad (1)$$

Écrire l'équation de projection de  $\mathbf{Q}$  dans le point  $\mathbf{q}_2$  sur le plan de la deuxième image après une rotation  $\mathbf{R}$  de la caméra.

- (c) En utilisant cet équation et à l'aide de l'eq. (1) trouver la relation entre les deux points  $\mathbf{q}_1$  et  $\mathbf{q}_2$ .

- (d) Que pouvez-vous dire de cette relation? (*Hint : pensez aux dimensions des matrices et vecteurs.*) Dans quel contexte avez-vous déjà rencontré cette relation?

(e) Cette relation est-elle valable pour tous les points détectés dans les deux images par le détecteur SD? Pourquoi?

(f) Combien de correspondances  $\mathbf{q}_1^i \leftrightarrow \mathbf{q}_2^i, \forall i \in [1, N]$  sont nécessaires pour estimer les éléments de cette relation? *I.e.* quelle valeur pour  $N$ ?

(g) Sachant que la caméra est calibrée (matrice  $\mathbf{K}$  connue), expliquez comment estimer la rotation  $\mathbf{R}$  à partir des correspondances  $\mathbf{q}_1^i \leftrightarrow \mathbf{q}_2^i$ .

4. Lors des séances de TP vous avez réalisé deux versions d'un "camera tracker". À partir de cette expérience, illustrez brièvement, **pour chaque version**,
- le type d'algorithme de tracking réalisé ;
  - les algorithmes utilisés pour la détection et le suivi des points d'intérêt ;
  - les algorithmes utilisés pour l'estimation de la pose de la caméra ;
  - en général, les points forts et les points faibles de chaque approche.