

**NOM et prénom :**

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

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

8 décembre 2023

### **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. Calculer la matrice de projection perspective qui permet d'établir la correspondance entre les points 3D de la scène et les points 2D dans l'image, c'est effectuer le calibrage de la caméra : **vrai** ou **faux** ?
2. La rectification épipolaire est indispensable pour effectuer l'étape de mise en correspondance : **vrai** ou **faux** ?
3. Pour réaliser la reconstruction 3D d'une scène avec un capteur stéréoscopique, l'étape de mise en correspondance de points 2D entre images est indispensable : **vrai** ou **faux** ?
4. Le principe du détecteur de Harris est de s'appuyer sur la notion de courbure de la surface dans l'image : **vrai** ou **faux** ?
5. Nous imaginons que nous avons réalisé une détection de points d'intérêt dans une première image. Si nous voulons maximiser la probabilité de trouver les correspondants corrects de ces points d'intérêt dans une deuxième image, il est préférable de réaliser une recherche exhaustive, plutôt qu'une recherche parmi des points d'intérêt dans cette seconde image : **vrai** ou **faux** ?
6. Les méthodes de mise en correspondance par corrélation robuste permettent de prendre en compte les difficultés liées aux occultations : **vrai** ou **faux** ?
7. La contrainte d'unicité garantit la contrainte de symétrie **vrai** ou **faux** ?
8. Pour réaliser la mosaïque d'images en travaux pratiques, nous avons eu besoin de calibrer la caméra : **vrai** ou **faux** ?







## 2 Réalité augmentée

1. La société AR7 a développé une application pour tablettes mobiles destinée à la maintenance des machines industrielles (voir Figure 3). L'utilisateur ·rice peut suivre les instructions pour intervenir sur la machine en pointant la caméra de la tablette vers la machine : les différentes pièces de la machine sont reconnues et mises en évidence sur l'écran et des animations montrent les opérations à effectuer (indicateurs rouges et verts en Figure 3).



FIGURE 3 – L'application est capable de reconnaître la machine et ses pièces (ici en rouge et vert) et de montrer les opérations à effectuer à travers des animations graphiques.

- (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.

2. Considérons l'homographie  $H_{3 \times 3}$  qui relie des points entre deux images de telle manière qu'un point  $\mathbf{a}$  de la première image est relié au point  $\mathbf{a}'$  de la seconde image par la relation  $\mathbf{a}' \sim H \mathbf{a}$  (où  $\sim$  signifie à un facteur d'échelle près et  $\mathbf{a}_{3 \times 1}$  est le vecteur des coordonnées homogènes du point).

— Prouver que, sous la transformation  $H$ , toute ligne  $\mathbf{l}$  de la première image est reliée à une autre ligne  $\mathbf{l}'$  de la seconde image.

Pour la preuve, on peut exploiter la propriété du produit vectoriel pour une transformation matricielle  $M_{3 \times 3}$  :

$$(M\mathbf{a}) \times (M\mathbf{b}) = (\det M) M^{-T}(\mathbf{a} \times \mathbf{b})$$

— Supposons de disposer d'un ensemble de lignes correspondantes entre les deux images,  $\{\mathbf{l}_i \leftrightarrow \mathbf{l}'_i\}$ . Pouvons-nous estimer l'homographie en utilisant ces correspondances et, si c'est le cas, combien de correspondances sont nécessaires ?

3. Revenons au problème de l'estimation de la pose de la caméra à partir de l'image d'un damier vu lors des travaux pratiques. La Figure 4 montre une configuration similaire, avec le damier placé sur le plan  $\pi$  et avec un repère  $XYZ$  du damier choisi comme indiqué.

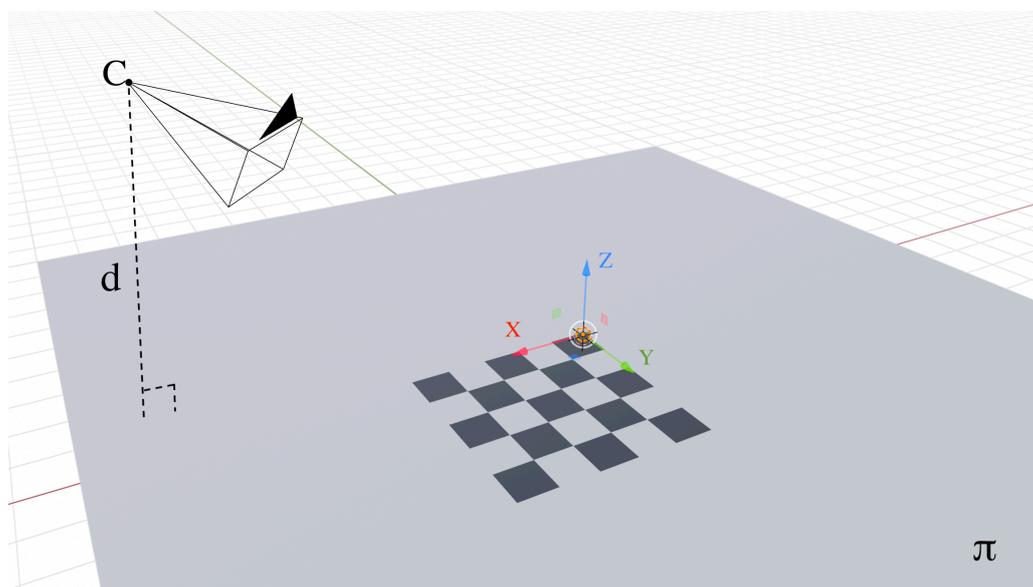


FIGURE 4 – Le damier est placé sur le plan  $\pi$ .

- Expliquez comment vous avez pu estimer l'homographie une fois que les coins sont détectés sur l'image.
  
- Supposons maintenant que l'homographie ait été estimée, correctement décomposée et que la rotation  $\mathbf{R}$  et la translation  $\mathbf{t}$  soient connues. Considérons le centre de la caméra  $\mathbf{C}$ .
  - Quelle est sa position dans le repère  $XYZ$  du damier ? (indice : dans le repère de la caméra  $\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$ ).

— Quelle est la distance  $d$  du centre de la caméra au plan  $\pi$  ?

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.