

UE Vision, Réalité Augmentée et Applications

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

9 décembre 2022

1 Vision (10 points)

Cette partie du cours est évaluée via des questions de cours et des questions d'ouverture en lien avec les thématiques abordées en cours.

1.1 Questions de cours (5 points)

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

1. Lorsque nous cherchons à apparier deux images, nous savons que les points correspondants se trouvent sur la même ligne de l'image : **vrai** ou **faux** ?
2. La rectification épipolaire permet de simplifier la recherche de correspondants entre deux images : **vrai** ou **faux** ?
3. Le mouvement 2D apparent correspond au mouvement que nous percevons d'une image à une autre lorsque nous considérons la projection de la scène dans chacune des images : **vrai** ou **faux** ?
4. Pour effectuer une détection de points d'intérêt, il est indispensable de suivre ces trois étapes : calcul d'une carte de réponse, suppression des non-maxima locaux, sélection des N points de plus forte réponse : **vrai** ou **faux** ?
5. Le principe du détecteur de Harris est de supposer qu'un point d'intérêt possède un voisinage faiblement corrélé aux voisinages de ses voisins : **vrai** ou **faux** ?
6. L'opérateur SIFT est intéressant car il permet une détection multi-échelle et il fournit également un descripteur : **vrai** ou **faux** ?
7. La mise en correspondance entre deux points homologues, proches d'une occultation, est difficile car l'occultation provoque des distributions photométriques différentes entre leurs deux voisinages : **vrai** ou **faux** ?
8. La mise en correspondance globale consiste à minimiser un coût de correspondance qui considère toutes les correspondances de points : **vrai** ou **faux** ?
9. Les méthodes de mise en correspondance globale ne peuvent pas utiliser de mesures de corrélation : **vrai** ou **faux** ?
10. La contrainte de symétrie est toujours valide quelle que soit le type de scène : **vrai** ou **faux** ?

1.2 Exercice sur la détection et le suivi de points (5 points)

Pour cet exercice, nous vous proposons de travailler sur la séquence de 5 images présentées dans la figure 1. Comme en TP, nous souhaitons construire une mosaïque d'images.



FIGURE 1 – Séquence d'images.

1. Quelles sont les difficultés que nous pouvons rencontrer pour détecter, puis pour apparier des points dans cette image ? Illustrez vos propos en indiquant clairement des zones précises de l'image.
2. D'après vous, est-ce que ces acquisitions respectent les conditions nécessaires pour réaliser une mosaïque en utilisant les mêmes outils que ceux étudiés en TP ?
3. Si vous souhaitez réaliser une détection puis un suivi de points d'intérêt, quelles approches de détection, puis de mise en correspondance utiliseriez vous ? Justifiez chaque choix en vous appuyant sur des exemples précis dans l'image.

2 Réalité augmentée

1. À l'occasion de la Coupe du Monde de football 2022 la FIFA a lancé une application pour dispositifs mobiles, FIFA+ Stadium Experience, qui permet de visualiser en direct les statistiques sur le match et la composition des équipes. Quand l'utilisateur pointe la caméra de son dispositif vers le terrain de jeu les différentes informations sont visualisées en surimpression sur le flux vidéo (voir Figure 2).

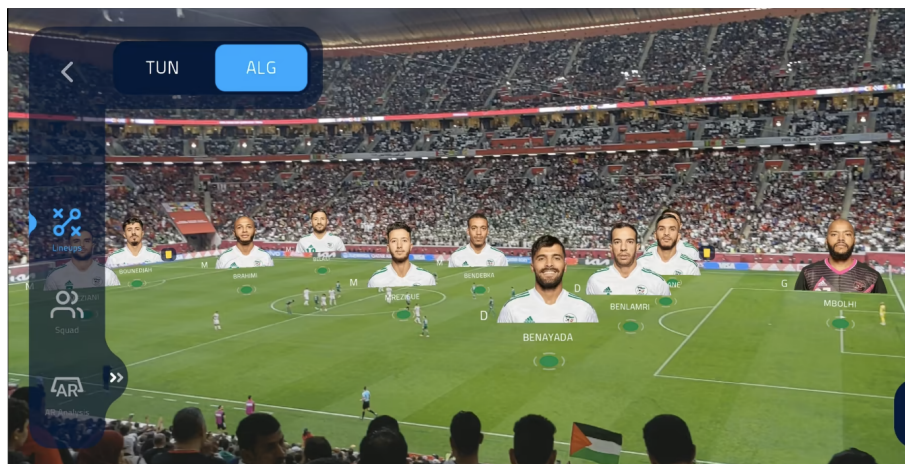
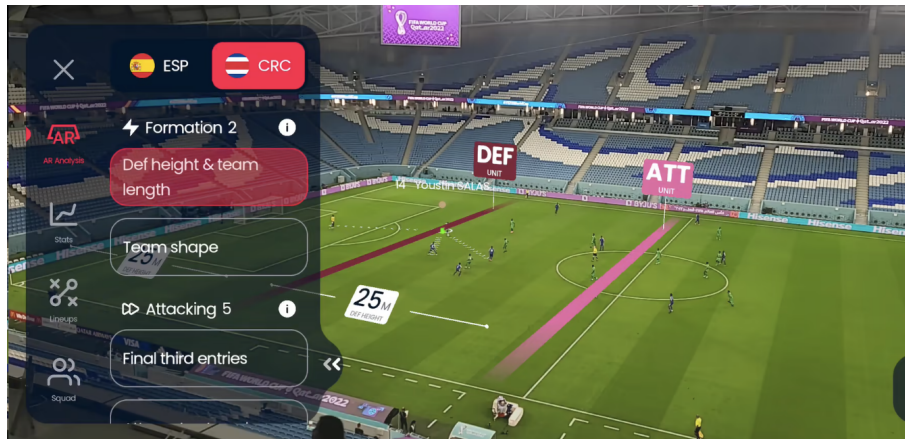


FIGURE 2 – L'application FIFA+ Stadium Experience permet de visualiser les statistiques du match comme si elles étaient dessinées sur le terrain de jeu.

- (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. On souhaite développer une application qui permet d'incruster une image donnée dans un flux vidéo. Pour cela on veut utiliser 4 marqueurs pour définir la zone où l'image doit être affichée. La [Figure 3](#) montre un exemple de cette application : les marqueurs sont placés sur une feuille rigide et l'image est incrustée dans la région délimitée par les marqueurs.



FIGURE 3 – Dans l'application on souhaite incruster (image au centre) une image donnée (à droite) dans le flux vidéo de la caméra à l'intérieur de la zone délimitée par les 4 marqueurs (à gauche).

- (a) Expliquez pourquoi généralement on préfère avoir des marqueurs carrés pour les applications de réalité augmentée.

- (b) On suppose qu'on a à disposition une bibliothèque logicielle qui détecte les marqueurs et qui nous fournit, pour chaque image, les coordonnées 2D **du centre** de chaque marqueur, toujours dans le même ordre (Figure 4). Expliquer pourquoi dans ce cas 4 marqueurs sont nécessaires.



FIGURE 4 – La bibliothèque logicielle dont on dispose, détecte les marqueurs (encadrés en vert) et donne en sortie la liste des 4 centres, toujours dans le même ordre (p.e. $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$, $\mathbf{q}_4$)

- (c) Détailler un algorithme qui permet de réaliser cette application.

3. 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.