

Plan de stage pour l'expérimentation

Objectif du stage :

Les objectifs du projet sont à terme de pouvoir estimer la profondeur d'un objet sur une image et de pouvoir calibrer la caméra utilisée dans le cadre d'images professionnelles et non professionnelles.

En calibrant une caméra on peut par la suite déterminer de nombreux paramètres autres que la position, telle que la vitesse d'un objet sur une vidéo ou sa taille.

La calibration de caméra consiste à déterminer la matrice de passage ainsi que les coordonnées X, Y et Z

La première étape consiste à déterminer la profondeur d'un objet par rapport à la caméra

Pour extraire la coordonnées Z d'un objet sur une image (sa profondeur) on va :

- Créer notre propre base de données avec des mesures de références (distance des objets par rapport à la caméra)
- Calculer et annoter les bounding box de notre propre dataset (YOLO ou avec un logiciel)
- Télécharger le dataset COCO et y extraire les informations sur les bounding box (hauteur, largeur et diagonale)
- Entraîner le réseau de neurones avec le dataset COCO
- Vérifier avec notre propre dataset la précision de notre estimateur de profondeur

Afin de pouvoir déterminer les coordonnées X et Y manquante et ainsi achever la calibration de la caméra, il nous faut un keypoint regressor. Cependant pour mettre en œuvre un réseau de neurones capable de retrouver il nous faut un dataset comprenant les coordonnées 3D des différents objets, et si elles ne sont pas disponibles il nous faudra alors la matrice de passage pour pouvoir avoir un dataset complet et ainsi pouvoir entraîner notre réseau de neurone.

En vue de déterminer la matrice de passage, on va mettre en place une méthode qui détermine la distorsion et la distance focale.

Pour cela il a fallu utiliser le dataset SUN360 (image omnidirectionnelle) et le modèle mathématique de la projection stéréographique.

Avec le modèle mathématique de la projection stéréographique et le dataset SUN360 on va créer des images avec des distorsions et des distances focales diverses.

Une fois le CNN entraîné avec le dataset personnalisé, l'objectif est d'obtenir la distance focale et la distorsion de l'image en sortie et donc la matrice de passage P qui nous permettra de déterminer les coordonnées.

Avec cette matrice de passage on va ensuite déterminer les coordonnées X,Y et Z sur un dataset (KITTI et nuScenes) afin d'entraîner notre keypoint regressor à déterminer les coordonnées X,Y et Z.

Une fois les coordonnées 3D acquises, la calibration de la caméra peut se faire.

*