

Mesure de distance des vulnérables pour la conduite autonome.

1 Introduction

Le but de ce projet est de détecter et d'estimer la distance de piétons à partir d'un capteur stéréo positionné sur un véhicule. Le capteur utilisé est une caméra Zed (cf figures 1, 3). pour cela, il faudra mettre en commun des compétences en géométrie avec des compétences en *Machine Learning*. D'un point de vue technique, la librairie *Opencv* sera utilisée.



FIGURE 1 – Caméra ZED : caméra stéréo 1080p avec une *baseline* de 12cm



FIGURE 2 – Exemple d'image obtenue avec une caméra zed

2 Méthode

Le projet se divise en quatre principales phases :

1. Calibration de la tête stéréo (paramètres intrinsèques, distorsions, paramètres extrinsèques)
2. Détection des piétons dans les images stéréo rectifiée
3. Estimation de la distance entre les piétons et la caméra
4. Validation expérimentale (utilisation de la plate-forme PAVIN)

La librairie **OpenCV** comporte toutes les fonctions utiles à ce projet. Il existe des exemples liés à calibration, la rectification, la triangulation pour la stéréo, ainsi que la détection de piétons. Des méthodes de type *Deep Learning* peuvent aussi être utilisées en alternative à la méthode HOG-SVM présente dans **OpenCV** (YOLO, FasterRcnn, SSD) avec différents *frameworks* (tensorflow, pytorch,...).

Dans le cadre de ce projet, des acquisitions d'images pourront être réalisées, soit pour l'étalonnage, soit pour l'évaluation (Pavin)

2.1 Calibration

La phase de calibration consiste à acquérir des images d'une mire avec la caméra Zed, puis à rechercher les paramètres intrinsèques de chaque caméra ; et enfin estimer la pose entre les deux caméras. Il existe des exemples **OpenCV** qui font ce travail. L'utilisation d'un de ces exemples est possible. Une fois la calibration effectuée, il faut rectifier les images de la caméra afin d'aligner les points correspondants entre les deux images sur la même ligne (notion de disparité).

2.2 Détection de piétons

La méthode utilisée dans **OpenCV** est issue des travaux de Dalal et Triggs [1]. La détection utilise un descripteur de type HOG (Histogrammes d'Orientation des Gradients), suivi d'un classifieur de type *SVM* (*Support Vector Machine*). L'alternative à cette méthode est d'utiliser l'apprentissage profond, plus performant. Les étudiants auront le choix entre ces deux classes de méthodes.

2.3 Estimation de la distance entre les piétons et la caméra

Lorsque l'on dispose d'une base stéréo et d'images rectifiées, la distance d'un point à la caméra est liée à sa disparité. Elle se calcule facilement par la formule :

$$\frac{f \times baseline}{disp}$$

où f représente la focale en pixels, $baseline$ représente la distance entre les deux centres optiques et $disp$ représente la disparité.

Références

- [1] Navneet Dalal and Bill Triggs. Histograms of Oriented Gradients for Human Detection. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *International Conference on Computer Vision & Pattern Recognition (CVPR '05)*, volume 1, pages 886–893, San Diego, United States, June 2005. IEEE Computer Society.

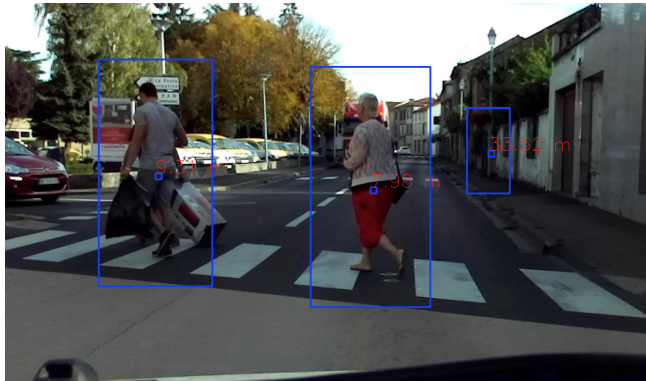


FIGURE 3 – Exemple de résultat final fourni par ma méthode. On note, sur cette image, une fausse détection