Chapter 1

Understanding the scope of 3D vision

At initiative state, 3D vision is understood as Principles of mapping from images to 3D geometry, typically focused in computer vision society which tries to interpret 3D information from images given by cameras. Traditionally, it is common to see discussion in 3D vision in topics of stereo-vision, monocular-vision, structure from motion, etc.

In fact, it is good to know that cameras, like human eyes, capture much rich information about the world, and well-known to be the best sensors in terms of gathering tremendous amount of data, but very hardly to extract useful information out of such huge and complex data. Some might concern that a human can easily understand surrounding environment with our vision system or eyes, why not a camera system doing so. Well, with trillions of neurons in human brain, we can do so many abstract learning (imagination, feeling, prediction, etc.) which are almost impossible for a normal Hardware or Computer to achieve. Note that at our current technology, some Super-Computers can establish billions of neurons (Deep Learning), and thus can do a few of such abstract learning, but still we, at a very rough speaking, are comparing trillions versus billions. On the other hand, other sensors such as LIDAR, RADAR, Ultrasonic, etc. can provide very reliable information which are much easier to be interpreted and applied in real autonomous applications, including obstacle avoidance, object tracking, terrain classification, etc. Nevertheless, those sensors just obtain some pieces of environmental information, like 3D structure, object location, and so on, and thus leading to an ambiguity to further understanding the environment.

A combination of 2D and 3D sensing becomes very common recently in Robotics,

which starts from the idea of having an active sensor (LIDAR, PMD, RADAR, Ultrasonic, etc.) to build up 3D structure of environment, and then mapping back to images to render objects' texture [18][2]. Such techniques have been applied in many fields, including computer vision, computer graphics, and robotics. Such 2D-3D mapping and 2D-3D calibration become a key-topic in 3D vision later on. Whereby, regardless of sensors (LIDAR, RADAR, SONAR, Stere-Camera, Monocular-Camer SFM, etc.) used for 3D structural reconstruction, a 2D-3D mapping is needed to understand both 3D structures and textures of objects. This makes 3D vision become not just a pure camera-based system, but a complex system which at least contains a camera.

Overall, the scope of 3D vision include Geometric Model, Core-Computer Vision, and 2D-3D mapping. Note that there is a significant amount of knowledge in Machine Learning within computer vision section.

Recently, autonomous car is a very HOT topic in Tech-trend, and thus 3D vision has got a great attention of scientists and researchers. Because it is the key component in Autonomous Vehicle's architect, and the most important block in Environmental Sensing [1][4][10][19]. Therefore, this is a chance to learn and review what has been done so far within 3D vision.