Marc Pollefeys
ETH Zurich

short abstract

Semantic 3D reconstruction

Obtaining 3D geometric information from images is one of the big challenges of computer vision. It is critical for applications such as robotics, autonomous vehicle navigation and augmented reality.

I will first briefly talk about some our recent work on vision-based autonomous micro-aerial vehicles, driverless cars and 3D on mobile devices.

While purely geometric models of the world can be sufficient for some applications, there are also many applications that need additional semantic information.

Next, I will focus on 3D reconstruction approaches which combine geometric and appearance cues to obtain semantic 3D reconstructions. Specifically, the approach I will present is formulated as a multi-label volumetric segmentation, i.e. each voxel gets assigned a label corresponding to one of the semantic classes considered, including free-space. We propose a formulation representing raw geometric and appearance data as unary or high-order (pixel-ray) energy terms on voxels, with class-pair-specific learned anisotropic smoothness terms to regularize the results. We will see how by solving both reconstruction and segmentation/recognition jointly the quality of the results for both subtasks can be improved and we can make significant progress towards 3D scene understanding.