Learning Attention Maps for Self-supervised Depth and Camera Motion Estimation in Monocular Videos.

Julio Mendoza
j153646@dac.unicamp.br

Helio Pedrini
helio@ic.unicamp.br

Abstract

In this work, we propose an approach for unsupervised depth and camera motion estimation. We focus on two aspects of the view synthesis approach. We learn attention maps that are used to increase the error contribution of the meaningful regions on the error computation between the synthesized and the real views. Moreover, we propose to approximate the uncertainty of pixels of being visible by mapping the depth consistency to probabilities using a Gaussian function. The attention maps are learned simultaneously with the depth network in an end-to-end manner.

We evaluated our method on the depth estimation and odometry tasks of the KITTI benchmark. The inclusion of the attention and visibility maps shows an improvement in depth and camera motion estimation tasks when compared with a competitive baseline method.