



单目、几何一致性、密集的视频深度估计

其他工作：

1、 传统的重建方法需要双摄像头或者多摄像头，但是由于立体匹配的时候，会受到噪声和视觉盲区的严重干扰，所以会产生许多空洞，[Szeliski 2010]

2、近几年出现了一些基于学习的对于单张图片的深度信息估计，但是对于视频的深度估计表现不太好，因为把视频当作独立的许多帧，会出现几何不一致性的现象。

3、近几年也出现了基于视频的深度信息估计技术，如[Patil et al. 2020; Wang et al. 2019b]的RNN的运用，或者[Patil et al. 2020; Wang et al. 2019b]的直接的多视场3D重建，但是他们有一个共同局限，就是智能针对静态场景。

相关工作：

1、监督化的单目深度估计。主要问题：需要大量的ground-truth，一些人通过人造场景来训练，另一些人通过传统的比较好的多目视觉方法去获取深度的gt，但是仍然是pseudo

2、自监督的单目深度估计。核心思想是通过可微的细节变形，去最小化重投影的光度学误差。但是光度学值一旦在整个场景短时间发生了剧烈变化，这些模型效果不太行了

3、多视场重建。其对于动态场景的估计不佳

此论文主要思想：

1、结合传统重建方法和近期的深度学习方法：用深度学习方法来对单目图像进行深度估计，但是存在几何不一致性。随后用在视频中提取的几何限制来微调网络，来实现几何一致性问题。

缺陷：

1、依赖于现有模型COLAMAP、FlowNet2，所以也会受到这个模型短板的限制。

2、速度很慢，由于需要根据特定的场景再进行训练才能得到深度图，实时基本上不太可能