深度估计方法整理

**传统方式**

1. 图像模糊度建模
2. 大气散射（霾）散射模型推断深度（一般有天空被拍摄下来的图像，通过散射模型能够推断像素深度。这里给出的是图像亮度C和深度z之间计算的公式：C0是没有散射情况下的图像亮度值，S是天空的图像亮度值。）
3. 表面阴影变化提供深度信息（图像的阴影边界蕴涵了图像的轮廓信息，因此可以根据不同的光照条件下图像的阴影恢复物体表面的三维形状。Michael提出了一种由阴影恢复形状的优化算法。）
4. 利用图像亮度和其中物体表面的形状之间的关系（利用单幅图象中物体表面的明暗变化来恢复其表面各点的相对高度或表面法方向等参数值，由单幅图像灰度明暗变化恢复三维形状是在一定的约束条件下从平滑变化的灰度图像恢复出表面法向信息，即根据物体表面反射模型建立物体表面三维形状与采集的图像灰度之间关系的反射图方程，以及由先验知识所建立的对物体表面形状参数的约束条件，对这些关系求解可得到物体表面三维形状。传统SFS方法均进行了如下假设：( 1)光源为无限远处点光源；( 2)反射模型为朗伯体表面反射模型( Lambertian)； ( 3)成象几何关系为正交投影。）
5. 光度立体法（一幅图像像素点的灰度主要由如下因素决定：物体的形状、物体相对于光源和摄像机的位置、光源和摄像机的相对位置，以及物体的物理表面反射特性等。光度立体法固定摄像机和物体的位置，通过控制光源方向，在一系列不同光照条件下采集图像，然后由这几幅图像的反射图方程求解物体表面法向量，进而重构物体三维形状。）
6. 通过纹理的变形得到深度信息（必须加上几个纹理元素（textels）约束条件：homogeneity，isotropy，stationary纹理由纹理元组成，纹理元可以看作是图像区域中具有重复性和不变性的视觉基元，纹理元在不同的位置和方向反复出现。由纹理元的变化可以对物体表面法向量方向进行恢复。常用的纹理恢复形状方法有三类：利用纹理元尺寸变化、利用纹理元形状变化以及利用纹理元之间关系变化对物体表面梯度进行恢复。Gibson在1950年首先提出了由纹理或纹理梯度表面深度的变化，Kender提出了一种恢复由规则的平行线组成的栅格表面取向的方法。这种方法对物体表面的纹理信息要求严格,需要掌握成像投影中纹理元的畸变信息，只有在纹理特性确定的条件下才能应用。）



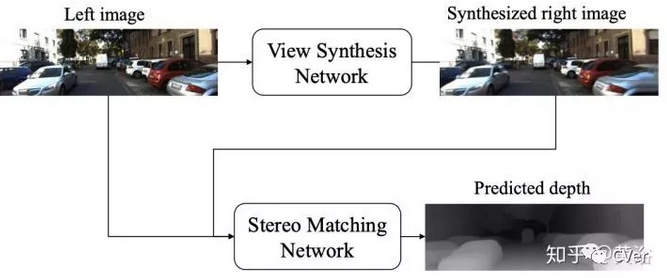
1. 遮挡提供深度信息（曲率，封闭的物体外轮廓，一般通过对图像灰度设门限可以得到，而它的曲率用来推导对应该门限的深度）
2. 从消失点得到深度信息（消失点即某平面的一组平行线在透视投影下会聚的点。那么，它相应的平面就能得到深度图，沿着平行线的平面，靠近消失点的赋予大的深度值）

**机器学习**

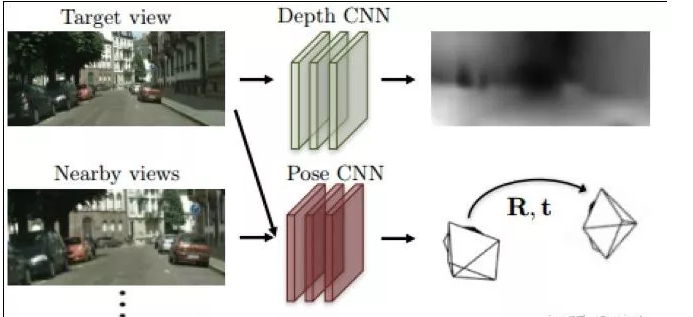
1. 基于MRF（各种约束关系放到模型的data term和connectivi term）
2. 把深度估计问题看做是搜素问题——>相似的图像有相似的深度图

**深度学习**

1. 神经网络（Depth Map Prediction from a Single Image using a Multi-Scale Deep Network 2015）
2. 神经网络+CRF（2017-CVPR-Multi-Scale Continuous CRFs as Sequential Deep Networks for Monocular Depth Estimation）
3. 双目立体视觉的空间约束被用作无监督学习单目的深度估计
4. 利用视差估计深度。使用对极几何约束，利用图像重建损失训练网络产生视差图
5. 使用视图合成作为监督（monoDepth2017使用的是DispNet网络架构，这是一个带跳跃连接(skip connections)和多尺度边(multi-scale side)预测的编码器-解码器网络，DispNet预测的是视差，而视差和深度互为倒数，因此用来求深度）
6. 神经网络结合表面法向图的联合训练：GeoNet（Geometric Neural Network）
7. 利用光场进行深度图估计（光场相机由于能够记录相机内部整个光场，可以实现重聚焦(模糊线索)和不同视角的变换(视差线索)，同时也可以利用这个特性进行深度估计）
8. 神经网络结合运动和边缘信息的联合训练
9. 神经网络结合view synthesis也是利用几何约束和成像特性的工作（有文章直接通过Deep3D模型做单目到双目的合成，这里的工作只是最终结果是深度估计而不是图像）



1. 自然地采用帧间运动为单目视频的深度估计提供帮助，实际上是双任务联合训练

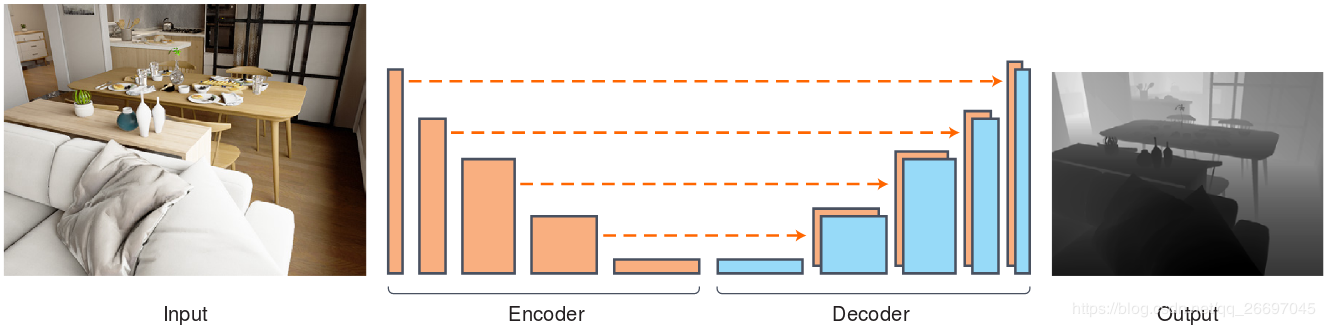


1. 延续了camera motion的结合，同时加上了其中物体各自运动的信息：可以认为是将optic flow分成了camera ego motion和object motion的工作，和加入语义分割有类似的思路
2. 在训练时把自动深度看成一个图像重建的问题。作者用的全卷积网络不需要深度数据，而是用合成深度作为一个中间产物。网络学习预测像素级的相关（2017-CVPR-Unsupervised Monocular Depth Estimation With Left-Right Consistency）
3. 有监督学习＋编解码网络进行无监督学习（2017-CVPR-Semi-Supervised Deep Learning for Monocular Depth Map Prediction第一个成分是　稀疏的ground-truth和预测深度值的差值；　第二成分是　左深度图＋右原图＝生成左图，然后和源左图的Identity的差值；　第三部分是　泛化损失）
4. 用点云对齐做三维空间约束vid2depth(2017,Mahjourian)（相机在两个位置看到同一些空间点(点云)，则可以利用特征匹配(传统方法)或深度学习(我们现在讨论的方法)估计出这两个相机的相对位姿和空间点到相机的距离(深度)，然后根据位姿和深度把在坐标系中的坐标变换的另一个坐标系中，变换得到的坐标和中本来的坐标应该是重合的，点云对齐就是调整位姿和深度使点云更好地重合，这样就优化了位姿和深度。）
5. 无监督学习

14.4、网络估计：深度，位移，表面法向量，图片之间的光流，匹配的confidence（2017-CVPR-Depth and Motion Network for Learning Monocular Stereo）

14.2无监督的深度估计一般利用基于对极几何的图像重建损失训练CNNs产生视差图。（对偶网络来自Godard，分别输出左右视差图。使用空间变换网络提出的双线性采样方法用视差图重建图像：用左视差图和右图像重建左图像；用右视差图和左图像重建右图像。）

1. 通过粗略的场景和光系统信息，用SGD选择性的优化metrics,用stereo限制规则化预测（2017-ICCV-Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction）
2. 利用语意关联和关注驱动损失函数进行单目深度估计（Look Deeper into Depth: Monocular Depth Estimation with Semantic Booster and Attention-Driven Loss 2018 1.像素数量在深度值和语义标签上都呈长尾分布；2.深度估计、语义分割相结合）
3. 用一个关联层对不同图像位置的关系建模（Monocular Depth Estimation with Ainity, Vertical Pooling, and Label Enhancement2018基于CNN的方法一般只提取诸如边界、纹理这样的绝对特征，而忽略邻接像素的深度关系(constraint)(相对特征)。针对这个问题，用一个关联层对不同图像位置的关系建模，位置相邻且外观相似的像素应该有相近的深度，外观不同则深度很可能相差很大。因此像素间的关系，称作关联。）
4. 迁移学习



1. 序数回归（把深度估计建模成回归问题然后用最小化均方误差训练网络把连续的深度估计离散化）