

三维重建综述

三维重建方法大致分为两个部分 1、基于结构光的（如杨宇师兄做的）2、基于图片的。这里主要对基于图片的三维重建的发展做一下总结。

基于图片的三维重建方法：

基于图片的三维重建方法又分为双目立体视觉；单目立体视觉。

A 双目立体视觉：

这种方法使用两台摄像机从两个（通常是左右平行对齐的，也可以是上下竖直对齐的）视点观测同一物体，获取在物体不同视角下的感知图像，通过三角测量的方法将匹配点的视差信息转换为深度，一般的双目视觉方法都是利用对极几何将问题变换到欧式几何条件下，然后再使用三角测量的方法估计深度信息。这种方法可以大致分为图像获取、摄像机标定、特征提取与匹配、摄像机校正、立体匹配和三维建模六个步骤。王涛的毕业论文就是做的这方面的工作。双目立体视觉法的优点是方法成熟，能够稳定地获得较好的重建效果，实际应用情况优于其他基于视觉的三维重建方法，也逐渐出现在一部分商业化产品上；不足的是运算量仍然偏大，而且在基线距离较大的情况下重建效果明显降低。

代表文章：AKIMOIO T Automatic creation of 3D facial models 1993

CHEN C L Visual binocular vision systems to solid model reconstruction

2007

B 基于单目视觉的三维重建方法：

单目视觉方法是指使用一台摄像机进行三维重建的方法所使用的图像可以是单视点的单幅或多幅图像，也可以是多视点的多幅图像。前者主要通过图像的二维特征推导出深度信息，这些二维特征包括明暗度、纹理、焦点、轮廓等，因此也被统称为恢复形状法（shape from X）

1、明暗度（shape from shading SFS）

通过分析图像中的明暗度信息，运用反射光照模型，恢复出物体表面法向量信息进行三维重建。SFS 方法还要基于三个假设 a、反射模型为朗伯特模型，即从各个角度观察，同一点的明暗度都相同的；b、光源为无限远处点光源；c、成像关系为正交投影。

提出：Horn shape from shading: a method for obtaining the shape of a smooth opaque object from one view 1970（该篇文章被引用了 376 次）

发展：Vogel 2008 年提出了非朗伯特的 SFS 模型。

优势：可以从单幅图片中恢复出较精确的三维模型。

缺点：重建单纯依赖数学运算，由于对光照条件要求比较苛刻，需要精确知道光源的位置及方向等信息，使得明暗度法很难应用在室外场景等光线情况复杂的三维重建上。

2、光度立体视觉（photometric stereo）

该方法通过多个不共线的光源获得物体的多幅图像，再将不同图像的亮度方程联立，求解出物体表面法向量的方向，最终实现物体形状的恢复。

提出：Woodham 对 SFS 进行改进（1980 年）：photometric method for determining surface orientation from multiple images（该文章被引用了 891 次）

发展：Noakes：非线性与噪声减除 2003 年；

Horocitz：梯度场合控制点 2004 年；

Tang：可信度传递与马尔科夫随机场 2005 年；

Basri：光源条件未知情况下的三维重建 2007 年；

Sun：非朗伯特 2007 年；

Hernandez：彩色光线进行重建方法 2007 年；

Shi : 自标定的光度立体视觉法 2010 年。

3、纹理法 (shape from texture SFT)

通过分析图像中物体表面重复纹理单元的大小形状,恢复出物体法向深度等信息,得到物体的三维几何模型。

纹理法的基本理论为: 对于一个具有光滑表面并覆盖了重复纹理单元的物体,当被投影在二维图像上时,其上的纹理单元会发生变形,这种变形分为投影变形和透视收缩变形。投影变形使离图像平面越远的纹理单元看上去越小,透视收缩变形使与图像平面角度越大的纹理单元看上去越短。由于这两种变形都可以从图像中测量得到,因此就可以分析变形后的纹理单元,反向求取物体表面法向和深度信息,进行三维重建。

提出: WIKTIN recovering surface shape and orientation from texture (1987) (被引用 454 次)。

发展: Warren 2010 对 wiktin 方法进行改进使用了透视投影;

Liboy 2006 给出了在纹理单元结构发生改变的情况下的重建方法。

优点: 精度高, 对光照和噪声不敏感。

缺点: 只应用于具有规则纹理的物体。

4、轮廓法 (shape from silhouettes)

这种方法将物体所在的三维空间离散化为体素,再将图像分割为前景(物体)和背景,通过投影测试判定一个体素点的投影是在前景还是背景中 如果一个体素点在不同视角都投影到了前景内,那么就认为它是物体的一部分,将所有这样的体素点集合起来就可以重构出物体的三维形状。

提出: Martin Volumetric descriptions of objects from multiple view 1983 年(引用 152 次)

发展: Szeliski : 使用了八叉树的体素存储结构,通过由粗到精的判断模式,极大地提高了建模速度。1993

Tarini : Marching intersection 结构。2002

Snow : 使用背景剪除算子,结合全局优化,实现了直接从亮度图像建模的方法,避免了繁琐的图像前背景分割操作。2001

Kuhn : 对已知环境中的未知物体进行重建 2009

Landabso : 解决由于标定不准确等误差导致的体素在不同视角图像中不一致的问题 2008

优点: 效率高

缺点: 对输入信息的要求非常苛刻。

5、调焦法 (shape from focus)

通过分析摄像机焦距光圈与图像清晰度之间的关系恢复物体深度信息,进而得到三维模型的方法: 物体只有位于摄像机焦距处时才会投影出清晰的图像,因此,通过建立物体到投影中心的距离与图像清晰度之间的关系,就可以恢复出深度信息。

提出 : Rajagopalan optimal selection of camera parameters for recovery of depth from defocused images 马尔科夫随机场对物体形状和外观进行重建 1997 (引用 30 次)

发展 : Lou : 给出了自标定和非标定的离焦重建方法 2007

Pradeep : 提出了一种将聚焦法与离焦法结合的方法,综合利用了两种方法的优势 2007

Sahay : 提出了一种去模糊的重建方法,可以实现高分辨率的重建 2008

Manna : 提出一种使用低通滤波器的重建方法 2009

Hasinoff : 够对复杂物体进行重建 2009

Thelen : 提出了一种综合考虑焦距算子邻域尺寸和高度差值的方法, 获得了比较精细的建模效果 2009

优点: 调焦法可以使用少量图像计算物体表面的稠密深度信息, 且对光源条件要求也比较宽松, 重建效果比较精细。

缺点: 不足的是需要不断改变摄像机的焦距光圈等设置, 很难实现自动重建。

6、运动法 (structure from motion SFA)

通过多幅未标定图像中检测匹配特征点集, 使用数值方法恢复摄像机参数与三维信息的一种方法。运动法首先在图像中检测需匹配的特征点集, 以恢复摄像机之间的位置关系。

提出 : Harris : A combined corner and edge detector 1988 (被引用 7391 次)

发展 : Shi : 提出了一种效果更好的角度提取方法 1994

SIFT : 目前使用比较广泛的特征点提取和匹配方法 2004 年提出

PCA-SIFT : 2004

GIOH : 2005

SURF : 2008

优点: 运动法对图像的要求非常低, 可以采用视频图像序列进行三维重建。可以使用图像序列在重建过程中实现摄像机的自标定, 省去了预先对摄像机进行标定的步骤, 可以对大规模场景进行重建, 输入图像数量也可以达到百万级, 非常适合自然地形及城市景观等三维重建。

缺点: 运算量比较大, 同时由于重建效果依赖特征点的密集程度, 对特征点较少的弱纹理场景的重建效果比较一般。

7、基于特征统计学习法

这种方法建立在大型的目标数据库基础上, 如人脸数据库、场景数据库。首先通过学习的方法, 对数据库中的每个目标进行特征提取(包括亮度、深度、纹理、几何形状); 然后对重建目标的亮度、深度、纹理、几何形状等特征各自建立概率函数; 最后将重建目标与数据库中相似目标的相似程度表示为概率的大小, 取概率最大的目标深度为重建目标深度, 再结合纹理映射或插值方法进行三维重建。

目前已用的概率模型有马尔科夫模型、隐马尔科夫模型(HMMs, Hidden Markov Model)和 PHMMs 模型。

总结:

单幅图像三维重建的主要思想是通过单张数码影像提取目标的颜色、形状、共面性等二维、三维几何信息, 从而利用少量已知条件获取该目标的空间三维信息。

重建过程简单、速度快、只需拍摄一张角度合适的数码相片即可获得该目标的三维几何信息, 它投入少, 不需要多个摄像机或投影仪进行标定; 技术上只对一幅图像进行预处理, 无需多幅图像的匹配, 避开了多幅图像重建的难点。

而基于机器学习的方法又有如下优点:

只要数据库足够完备, 任何和数据库目标一致的对象都能进行三维重建, 并且重建质量和效率都很高, 很少需要人工交互。这种重建技术最大的困难是建立完备的数据库。此外, 如果能在深度预测上有更精确的估计, 重建效果会更精确。