

硕 士 研 究 生 读 书 报 告



## 题目 “用镜子和玻璃表面重建场景”的学习和总结

作者姓名 姚宇辉

作者学号 21851090

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 二○一八年十二月

“Reconstructing Scenes with Mirror and Glass Surfaces” Reading

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: ×××

By

×××

Zhejiang University, P.R. China

2018

摘要

本文学习和研究了SIGGRAPH 2018上关于对镜子和玻璃表面重建3D真实场景的最新研究成果。先对作者相关工作背景进行了概述，其中包括前人的研究进展。之后我阐述了作者论文中的所提到的将镜面玻璃面重建和原始场景重建分离，并采用全新的在捕获阶段使用标签的方法估算镜面玻璃面的位置和平面参数并结合前人工作完成重建的过程。

**关键词**：扫描装置，反射面，镜面，玻璃面

Abstract

This article has studied the latest research paper in SIGGRAPH 2018 on the reconstruction of 3D real scenes on mirrors and glass surfaces. Firstly, the author's relevant work background is summarized, including the research progress of the predecessors. After that, I discuss the main content mentioned in the paper that the separation of the mirror glass surface reconstruction and the original scene reconstruction and the new method of using the label in the capture stage to estimate the position and plane parameters of the mirror glass surface, and then combining the work of the predecessors to complete the reconstruction process

**Keywords：**Scanning, Reflective Surface, Mirror, Glass

1引言

对于大多数当前的3D扫描技术而言，诸如玻璃和镜子的平面反射表面是非常难以重建的。如果仅仅是粗糙地对待它们，那么它们只能给出与镜外完全相同的重复的场景结构，从而完全忽略了各种光线阴影和视角的影响，完全破坏了重建。

最近使用类似Kinect传感器的主动扫描技术在重建室内场景的过程中应用非常成功[Kähler et al. 2015; Newcombe et al. 2011; Nießner et al. 2013]。其中在场景中一个经常出现的元素：镜子和玻璃等反射表面，在以前的工作中很少得到处理。虽然这听起来像是一个小小的遗漏，但实际上镜面对任何重建系统都是一个重要的问题。一面完美的镜子展示了对世界的完美反映，这与直接观察镜像世界无法区分。因此镜子基本上是不可见的。这样的话，镜像场景仍将使用标准视觉技术重建。从而得到图1中显示的来自ScanNet数据集的浴室示例[Dai et al. 2017] 在镜子后面重建的反射场景部分。



**图1 来自ScanNet数据集的示例扫描[Dai et al. 2017]镜中的家居**

这种镜像几何体与位于镜子后面的真实几何体重叠，并干扰其重建。因此有必要将镜子重建为场景中的标准元素，以避免整个场景的无效重建。类似地，玻璃表面通常不由传感器重建，但仍应包括在重建模型中。由于镜子和玻璃在普通室内环境中普遍存在，最近的场景重建方法如Matterport3D [Chang et al. 2017]要求用户在扫描中手动选择窗口和镜像。

**2 前人研究进展**

**2.1检测镜面和玻璃面**

在常规图像中，平面玻璃和镜子难以被检测到。当光线照射到玻璃表面时，大部分光线进入玻璃，被折射，然后停留在玻璃背面的不同位置。只有小部分的光（通常小于10％，由Fresnel方程确定）在前后的玻璃表面上反射，并且可以作为两个截然不同的微弱镜像被观察到，除非它们被玻璃后面的场景事物所覆盖。 [Shih et al. 2015]。当人们可以实际观察到来自场景的透射和反射部分的两个不同信号后，玻璃是部分透明且部分反射的表面的事实便提供了有价值的信息。这被广泛用于多信号的飞行时间传感器[Foster et al. 2013; Jiang et al. 2017; Koch et al. 2017b,a]。有几种被动成像方法也依赖于检测和解开这两个不同的图像分量。应用包括反射去除[Arvanitopoulos et al. 2017; Xue et al. 2015]和场景重建[Sinha et al. 2012; Wang et al. 2015; Wanner and Goldluecke 2013]。其他方法依赖于一些间接观察的结果，例如玻璃吸收一些光，尤其是光谱的红外部分[Klank et al. 2011]，或者光线的偏振状态的变化[Miyazaki et al. 2004]。

标准的镜子在玻璃的背面有一层薄反射层。该背面的反射通常比前反射强得多，导致众所周知的镜像效果即几乎没有任何可观察到的双重图像。用于科学用途的高质量镜子在前表面上也有反射层（所谓的第一面镜），从而产生更均匀的反射。鉴于我们无法观察到两个不同的场景，检测镜面非常具有挑战性。一种方法是切换模态，例如，使用超声波测量距离的方法，其从镜子表面反射并且可以在特定几何的成像下观察到[Yang and Wang 2008; Zhang et al. 2017]。

在光学领域中，有些弱信号可以使用。例如，沿着视线返回多个观察的飞行时间扫描系统可以返回在实际镜面处反射的弱信号。从镜框边界观察到的深度不连续性（也称为跳跃边缘）也可以推断出镜框的存在[Käshammer and Nüchter 2015; Yang and Wang 2011]。有源和纯无源光学方法可以检测捕获图像，深度图像或完整场景模型中的镜像对称性[Yang and Wang 2011]。

虽然这种方法原则上可以给出场景中镜子存在与否的可靠指示，但是如果仅在镜子中观察到场景的一部分并且从未直接看到，这种方法就无效了。最后，可以使用其独特的失真模式检测曲面镜面[DelPozo and Savarese 2007].。

**2.1重建镜面和玻璃面**

通常通过检测扩展目标的镜面反射来恢复镜面表面的几何形状，从而能够估计在给定像素处观察到的表面的法线方向。这种方法的关键挑战是目标需要覆盖所有相关的角度，这通常要求它完全包围物体[Balzer et al. 2014]。因此，大多数实际方法采用有限大小的的目标，参见[Liu et al. 2015; Tarini et al. 2005]，其可以在对象周围移动且记录在多个捕获帧中[Balzer et al. 2011]。在极限情况下，人们甚至可以使用聚集在许多帧上的单点光源的镜面反射[Chen et al. 2006]。或者，被动捕获方法利用平面中反映的完整环境[Sinha et al. 2012]或曲面和更复杂的表面[Godard et al. 2015]。对于近平面表面的特殊情况，[Ding and Yu. 2008]将反射解释为一般线性相机下观察到的正交投影，并通过观察目标棋盘确定其参数。Jacquet et al. [2013] 使用大窗格中的光线反射来分割3D立体图像来重建法线场。

**2.1场景中的镜面和玻璃面**

对于主动式扫描的应用，大多数系统能够扫描玻璃，因为折射可能会产生偏移。只要所有光线沿同一平面镜面反射，就可以扫描镜面。这就重建了位于镜中的真实场景几何的镜像。Fasano et al. [2003]使用镜子得到使用激光条纹图案扫描仪难以到达的区域。由于常规镜子提供低质量的扫描数据，他们使用了第一面镜。他们还提出了一个手持式镜子的方法，他们使用镜子上的标记检测其位置。

Foster et al. [2013] 依赖于使用基于飞行的多回程系统在玻璃（和镜子）表面上观察到的微弱漫反射。基于物理反射特性，他们只能在非常有限的角度范围内以这种方式观察玻璃。它们扩展了流行的占用网格映射，使其能够包含仅在一小部分帧中可观察到的表面。

Jiang et al. [2017]使用飞行时间（TOF）传感器的返回信号的强度，到表面的距离和入射角作为室内环境中基于神经网络的玻璃表面分类的特征。虽然足以用于机器人导航，但这种方法不能产生高精度的扫描数据。

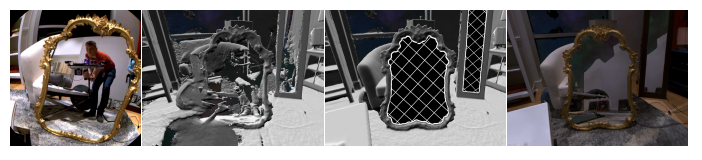
Käshammer and Nüchter [2015] 检测TOF数据中已知几何的镜像。他们使用深度相关的跳跃边界找到镜框。Yang and Wang [2011]引入了类似的度量标准来检测候选镜像位置。我们将跳跃边缘度量推广到还包括无框镜像。此外，我们使用有源标签的反射来确定镜像平面。因此，我们能够在场景中检测具有任意未知几何体的平面镜。

另一种工作方式是将超声波传感器与经典光学扫描相结合检测玻璃和镜面。 Yang and Wang [2008]将声纳传感器集成到激光扫描系统中检测玻璃。在激光扫描中的深度不连续处可以检测到镜子。通过反射几何体的基于ICP的指数来判断镜子的实际范围。

Zhang et al. [2017] 增加了带有超声波距离传感器的Kinect扫描设备。他们寻找光学和超声波深度的差异，并使用精细的基于MRF的推断来检测玻璃和镜子表面的范围。他们还可以通过将参数曲面拟合声学传感器的点集来重建曲面。超声深度传感器的一个关键的局限是需要以接近正交的角度观察表面，而我们的相机具有宽视场，并且我们的标签可从扫描装置的更大一组取向中看到。

**3 文章所做工作**

在文中，作者提出了一种自动检测和重建场景中镜子和玻璃表面的方法。大致流程是用文中提到的扫描装置在镜头上加上新的标签AprilTag捕获场景图像；重建没有镜子的场景，这在之前的研究工作中就已经做到；再重建第一步中检测到的镜面；最后对镜面中的场景进行渲染。如下图所示：



**图2 重建场景中镜面的大致流程**

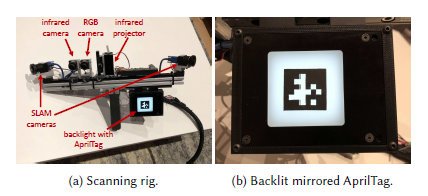
其中比较关键的是捕获场景检测镜面和重建镜面两个实现过程。

作者的主要想法是在捕获图像的装置上添加一个标签，该标签只有在摄像机面向镜子或玻璃表面时才能观察到。在作者的工作中，他们使用了AprilTag的镜像版本**[Olson 2011; Wang and Olson 2016]**。基于对该标签的观察结果不仅能够检测到反射表面，而且还能稳健地估计表面的平面参数。此外，他们还开发了一种基于多特征的方法来检测平面镜像表面的边界。

具体地来说，文中主要分成几个步骤来实现他们的想法。

为确保能够有效检测目标，他们使用AprilTag基准标记。经过一些基本的图像处理完成局部强度归一化后，候选位置被检测为包含暗区域的连续亮区域。接下来，提取暗区域的边缘和角落，并且对与标签对应的数字代码进行解码。最后细化边缘位置。这里作者制作了AprilTag的镜像版本，只有在镜面中时才能检测到该版本的AprilTag。AprilTag库的输出是每帧检测到的AprilTag的集合。每次检测都包含四个角和标签中心的有序图像位置，标签ID以及有关检测质量的其他信息。

这里文中也给出了AprilTag的大致解释。AprilTag是通过将3M 468MP压敏粘合剂应用于Thorlabs BKF12黑色铝箔而构建的。然后在水流上切割箔片夹在ABS塑料片之间，粘合剂面朝上。然后将层压图案安装在点光源上以确保均匀照明。作者也提到他们调整了背光的强度以匹配整体系统设置，确保捕获的标签的任何部分都不会过度曝光。图3显示了安装在背光上并连接到扫描装置上的镜像AprilTag。铝箔标签的宽度和高度为28.34mm。



**图3扫描装置的原型与附加的背光AprilTag**

检测到镜面之后，就可以对镜面的各个平面参数进行估算。文中假设一个典型的SLAM系统可以准确给出扫描装置每个帧的位置估计[Engel et al. 2018; Mur-Artal et al. 2015]。因此估算镜面参数的问题就转换成最小化镜头上的AprilTag的各个角落顶点重投影到摄像机视图时产生的投影误差。也就是说给定误差函数求解最优化问题即可得到各个平面参数。

大致地，这里定义了必要的未知参数，变换矩阵，镜面平面向量，投影生成模型以及误差项等数学式子，运用了基于自动束调整的校准方法，可准确检测镜面平面以及相对标签位置。

文中也提到了只需要在单个图像中看到AprilTag就可以使用场景几何信息和SLAM将观察信息传输到所有其他帧。在附加帧中观察镜子（可能是AprilTag）可以提高估计镜像平面的精度，并有助于后面准确地检测镜像边界。

文中也对多镜面参数的估算的情况进行了讨论。假设每个AprilTag观察属于它本身的镜子，然后我们就可以如上所述估算。另外，他们将2D标签检测的中心与估计的平面相交，以找到镜子表面上的观察点，AprilTag通过该观察点被反射。鉴于AprilTags的低误检率，这一点一定属于镜面。为了将观察结果分离为场景中每个镜像的集合，作者使用了非参数聚类算法，我们将其表示为DP平面，来自于DP-MEANS [Kulis and Jordan 2011] and DP-vMF-means [Straub et al. 2015]。除了使用不同的距离度量之外，这些算法执行相同的类似K-MEANS的交替优化：

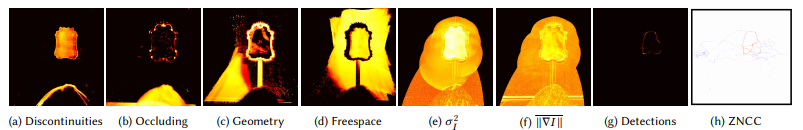
1.递增地将数据点分配给最近的现有点集，除非最近的点集超过某个阈值λ。在后一种情况下，从查询数据点实例化新的点集。

2.给定所有关联数据点重新计算点集的中心。他们采用现有算法，但将距离的度量改为对称的点到平面距离。

在上述算法中，如果距离在以米为单位度量的某个阈值λ内，则DP平面算法通过对称的点到平面的距离将数据点分配给聚类中心。 如果距离超过此阈值，则使用观察值初始化新簇。最后发现λ= 10 cm在他们所有的非共面反射镜实验中都能产生良好的镜面分割。对于共面镜，文中使用边界检测细化分割并重新计算新分割镜的镜像平面参数。

在边界检测中，文中提出了一种精确重建有框和无框镜和玻璃表面边界的自动方法。给定前面估计的镜像平面，将其用作边界提取的自然参数。将平面离散为具有方形单元格的网格，通常使用5mm的分辨率，将下面讨论的所有特征投影到该平面上，并使用基于变化的总分割来提取边界。

文中探索了八种不同的特征通道以改善镜像分割。除最右边的ZNCC通道外，所有通道都使用热度颜色方案以对数刻度显示，该通道从-1显示为蓝色，1显示为红色。如下图所示：



**图4 不同特征通道显示的镜像分割效果**

根据上述特征，作者使用g加权总变差分割[Unger et al. 2008a,b]做镜面边界提取。最后使用行进方块算法来提取子单元精确的分段线性镜像边界，作为分割图像中值为0.5的等值轮廓。行进方块算法[Lorensen and Cline 1987]相当于在2D网格上行进立方体。

另外，文中将结论迁移到玻璃面上，并作出相应的改动。

玻璃表面在多种方面与镜子不同。首先，玻璃表面的图像通常是透射和反射场景之间的混合。因此，反射的图像的亮度都会降低，并且可能会受到来自直接光路的纹理的破坏。因此，反射场景中的任何特征检测必须对相对低的对比度和信噪比具有鲁棒性。

其次，反射的场景反射在玻璃的正面和背面上，产生双重图像。这种效果取决于扫描台与玻璃表面的距离。根据作者的经验，如果偏移量太大，AprilTag库将不会检测标签，否则通常会重建两个标签位置中的一个。因此，在扫描时保持与玻璃板的最小距离就足够了。

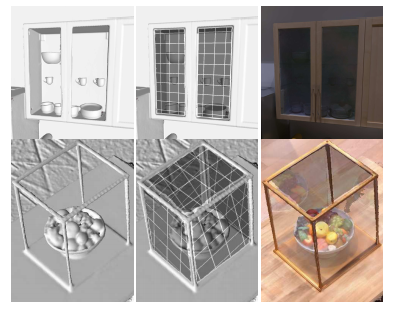
第三，需要区分玻璃和镜子表面。如果观察到检测到的AprilTag的投影区域内的几何图形既不在AprilTag的深度也不在反射平面的δ邻域内，就将表面分类为玻璃。这只适用于玻璃，而对于镜子，AprilTag可用作拦截器。换句话说，通过AprilTag的图像检测几何形状意味着正透过一层表面看到图像。另外，对于诸如相框之类的浅物体，这种区分会失败，导致玻璃表面被错误分类为镜子。另一种分类方法是检测反射的AprilTag的强度，玻璃的强度低于镜面。

除了这些变化之外，前面的方法可以直接重建平面以及框架玻璃表面的边界。

阐述了实现方法和理论之后，作者给出了一些实验结果并进行了对比，都达到了很好的效果如下图所示：



**图5 左图是真实场景，右图是重建效果**



**图6 玻璃柜的重建**

最后，文中提到了一些阶段性的不足之处。虽然他们的方法通常非常稳健，但仍然能在处理流程的各个阶段观察到偶然的故障情况。如果在任何输入框架中未检测到AprilTag，该方法将发生灾难性的失败。这通常是由于不良的成像条件造成的，例如由于扫描仪快速移动导致的模糊图像，标签的部分可见性，低对比度或高度弯曲的反射面。这可以通过更显著的目标来缓解，例如，标记平面标签边角的一组LED。

而无边框玻璃也呈现出一种具有挑战性的情况，其中光度测量结果太弱而不能约束边界，当估计平面的δ邻域内存在几何形状时，玻璃分类中会出现故障情况。可能的一种解决方法是校准镜子和玻璃上标签的反射强度，并使用该提示来区分两者，因为玻璃的反射比镜子的反射明显更暗。

**4 小结**

通过对该文章的阅读，我对三维场景中镜像重建的研究有了更加明确的认识。

其中对镜面和玻璃面的重建过程，文章采用了一种全新的思路，在重建的过程中把镜面玻璃面和原始场景分离重建，在捕获时利用摄像头上的标签的投影对场景中的镜面和玻璃面的位置和参数进行检测和还原，从而得到很好的效果。

作者同时也借鉴了许多前人的结论和算法来完成特征和边界提取等其他部分。而在最后作者提出的新的不足也给后面改进和研究的方向提供了新的道路。