谨以此论文献给我的恩师、父母及朋友

—— 杨攀

深度哈希特征及其应用研究

学位论文答辩日期：\_\_\_\_\_\_\_\_\_\_\_\_\_\_

指导教师签字：\_\_\_\_\_\_\_\_\_\_\_\_\_\_

答辩委员会成员签字：\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**独 创 声 明**

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含未获得 （注：如没有其他需要特别声明的，本栏可空）或其他教育机构的学位或证书使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名： 签字日期： 年 月 日

---------------------------------------------------------------------

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权学校可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同时授权中国科学技术信息研究所将本学位论文收录到《中国学位论文全文数据库》，并通过网络向社会公众提供信息服务。（保密的学位论文在解密后适用本授权书）

学位论文作者签名： 导师签字：

签字日期： 年 月 日 签字日期： 年 月 日

深度哈希特征及其应用研究

# 摘 要

通过物体的二维图像估计其三维模型是计算机视觉领域的基础问题，在工业测量、人机交互、医学、考古、计算机动画等多个领域有着广泛的应用。

现有的三维重建理论和方法一般假设成像环境是在空气中，相关技术也较为成熟。然而，水下环境相对复杂，水中的悬浮物对光线产生吸收与散射，导致成像信息的衰减，因此无法直接使用现有的三维重建技术。

本文使用深度传感器获取物体的深度信息，在此基础上利用多光谱光度立体技术求解物体表面法向，将深度和法向融合完成物体的三维重建，并对已有的水下成像模型进行简化，使得该方法能够适用于水下物体的三维重建。

本文首先实现了基于深度传感器的多光谱光度立体三维重建技术。使用红、绿、蓝三种颜色的光源同时照射物体，在物体表面反射符合朗伯模型的假设下，根据颜色信息对物体表面划分不同的区域，利用深度数据从单幅彩色图像分离出三个通道的子图像计算光照矩阵，得到物体的表面法向，再与深度数据融合获得更高精度的三维重建结果。

由于水对不同频率的光的吸收和散射程度不同，本文建立了适用于多光谱光源条件下的水下成像模型，将上述技术在水下应用环境进行了拓展。通过在不同浑浊度的水下环境进行实验，结果表明本文提出的方法能够以更加简单的操作、廉价的设备和较低的计算量，对水下物体进行三维重建。

关键词：三维重建；多光谱光度立体；深度信息；朗伯模型

Fusing Depth and Multispectral Photometric Stereo for 3D Reconstruction and Its Applications in Underwater Environment

# Abstract

Reconstruction of the 3D geometry of objects from their 2D images is the basic problem in computer vision, which can be applied in the fields such as industrial measurement, human-computer interaction, medicine, archaeology, computer animation, et al.

Current theories and methods in 3D reconstruction were developed under the assumption of the air environment. However, the underwater environment is much more complicated. The absorption and scattering effect caused by the suspended particles attenuate the image signal, where the reconstruction algorithms failed under the water.

This thesis uses depth sensor to obtain the objects’ depth data, then uses multispectral photometric stereo to calculate the objects’ surface normal. Those two part are fused for 3D reconstruction. At the same time, this thesis simplify the under water image formation model, to make it apply to our method.

This thesis implements a multispectral photometric stereo 3D reconstruction technology based on depth sensor. The object was illuminated with three light sources of the color: red, green, and blue, simultaneously. Assume that an object's surface reflection complies with lambert model, partitioning the surface area based on color, then the color images are used to compute the normal vector of the object surface.

For the different effects of absorption and scattering on different spectrum light, this paper improve the underwater imaging model, and introducing the method explained in the previous paragraph to the underwater environment. Experiments on different turbidity of water environment show that the method of this thesis can reconstruct under water object in a more simple operation, cheap equipment and low amount of calculation.

Keywords： 3D Reconstruction; Multispectral Photometric Stereo; Depth Information; Lambert reflection model

目 录

[摘 要 I](#_Toc420244353)

[Abstract II](#_Toc420244354)

[1 绪论 1](#_Toc420244355)

[1.1 选题背景及意义 1](#_Toc420244356)

[1.2 国内外研究现状 2](#_Toc420244357)

[1.2.1 基于深度摄像机的三维重建 2](#_Toc420244358)

[1.2.2 水下三维重建技术 5](#_Toc420244359)

[1.3 论文的创新性 8](#_Toc420244360)

[1.4 论文的主要内容及结构安排 9](#_Toc420244361)

[2 光学三维重建技术 11](#_Toc420244362)

[2.1 被动光学重建技术 11](#_Toc420244363)

[2.1.1 立体视觉技术 12](#_Toc420244364)

[2.1.2 运动恢复法 14](#_Toc420244365)

[2.1.3 明暗恢复法 14](#_Toc420244366)

[2.1.4 纹理与表面朝向法 15](#_Toc420244367)

[2.1.5 焦距法 15](#_Toc420244368)

[2.2 主动光学重建技术 16](#_Toc420244369)

[2.2.1 光度立体技术 16](#_Toc420244370)

[2.2.2 飞行时间 18](#_Toc420244371)

[2.2.3 结构光测量法 19](#_Toc420244372)

[2.3 小结 21](#_Toc420244373)

[3 融合深度信息的多光谱光度立体三维重建 22](#_Toc420244374)

[3.1 深度信息的获取 22](#_Toc420244375)

[3.1.1 Kinect深度测距原理 22](#_Toc420244376)

[3.1.2 Kinect误差分析 23](#_Toc420244377)

[3.2 朗伯模型与多光谱光度立体 23](#_Toc420244378)

[3.2.1 朗伯模型 23](#_Toc420244379)

[3.2.2 多光谱光度立体 24](#_Toc420244380)

[3.2.3 多光谱光度立体的求解 26](#_Toc420244381)

[3.2.4 多光谱光度立体三维重建 28](#_Toc420244382)

[3.3 深度信息与光度立体的融合 29](#_Toc420244383)

[3.3.1 使用位置信息改善法向 29](#_Toc420244384)

[3.3.2 使用法向信息改善位置 30](#_Toc420244385)

[3.4 实验结果 32](#_Toc420244386)

[3.5 小结 36](#_Toc420244387)

[4 水下三维重建 37](#_Toc420244388)

[4.1 水下光学成像模型 37](#_Toc420244389)

[4.1.1 水下单散射模型 37](#_Toc420244390)

[4.1.2 水下光学成像模型 38](#_Toc420244391)

[4.2 水下多光谱光度立体 40](#_Toc420244392)

[4.2.1 水下多光谱光度立体模型 40](#_Toc420244393)

[4.2.2 水下多光谱光度立体求解 40](#_Toc420244394)

[4.2.3 水下多光谱光度立体三维重建 42](#_Toc420244395)

[4.3 深度信息与光度立体的融合 42](#_Toc420244396)

[4.4 实验结果 43](#_Toc420244397)

[4.5 小结 47](#_Toc420244398)

[5 总结与展望 48](#_Toc420244399)

[参考文献 50](#_Toc420244400)

[致谢 54](#_Toc420244401)

[个人简历 55](#_Toc420244402)

# 1 绪论

## 1.1 选题背景及意义

地球表面约71%的面积被海洋覆盖，与陆地相比，人们对海洋的了解十分有限。近年来，随着科学技术的发展及人类活动范围的扩大，人们对海洋环境的探索需求日益提高。另一方面，随着传感器技术、定位技术及计算能力的飞速发展，海底探索已经发展到了新的境界。

三维重建是指将真实场景中的三维物体在计算机中建立数学模型的过程，是计算机视觉领域的一个热门研究方向。相比于二维图像，三维模型能够提供物体的深度数据，从而能够更加全面的展示物体特性，因而在计算机动画、人机交互、现代医学等多个领域都有着广泛的应用。在三维重建领域中，空气中的重建技术及应用程序已经发展的较为完善，而在水下的应用中，由于水对光线存在吸收和散射作用，使得相机捕获的水下图像的质量有所下降，无形中增大了水下三维重建的难度。也正是由于水下环境的特殊性，使得水下三维重建成为当今世界上许多科研机构和研究人员一直以来的研究热点。

水下三维重建及其可视化应用程序能够提供水下环境和地质特征信息，展示水下地质结构在自然或人工干预状态下的发展及变化情况。因此，在定期管道检查、长途通信电缆维护、水下工程设备检测、海上平台建设、海洋环境研究、海洋生物研究、水产养殖和考古学等领域都有广泛的应用。

当前广泛使用的水下三维重建技术大都是使用各种方法的融合算法。常见的是使用光度立体技术、双目视觉或运动恢复结构技术与其他算法的结合。在这些技术中，通常需要对场景进行单幅或多幅图像序列的采集，或在同一视角下拍摄物体在不同光照方向下的图像，然后通过图像之间的特征点匹配或明暗变化特征进行三维重建。这不仅对待重建场景有局限，而且在数据处理过程中需要进行大量计算。而在非光学水下传感领域，声学相机只能提供水下环境的粗糙的三维模型，在考古学等需要精确数据的领域是无法使用的。

深度相机的出现打破了传统三维重建技术对彩色图像的过分依赖，由于其自身具有实时重建的算法，而且体积较小，操作简单，为三维重建领域提供了一个新的发展方向。2012年，Ken Mankoff博士仅带着背包、笔记本电脑、电池组和Kinect进入了一个位于挪威斯瓦尔巴特群岛Rieperbreen冰川的地下洞穴，使用Kinect对洞穴内部进行了扫描，并且绘制出洞穴内部的三维不规则表面。体现出Kinect相对于传统三维重建设备，具有体积小、对场景的三维数据采集快速且精确的优势，一时间迅速成为三维重建研究领域的重要的工具。

使用深度相机直接获取的场景深度数据，虽然会存在高频噪声，但是其低频数据是相对准确的；而光度立体技术能够恢复物体精细的低频细节，在高频区域却存在较大误差；多光谱光度立体技术将传统光度立体的时分复用改为频分复用，从而为动态场景的重建提供了可能。将深度数据与多光谱光度立体技术输出的结果在合适的频段做结合便能够获得比单独使用任何一种方法更精确的三维重建结果。与传统的方法相比，使用深度相机结合多光谱光度立体的方法所需设备简单（只需要在Kinect的基础上添加三个不同颜色的光源），数据采集方法便捷，重建过程中计算量也较小。目前，在多光谱光度立体技术的研究中，国外已经有学者取得了一些研究成果，但也只是局限于空气环境中，国内更是鲜有研究。

## 1.2 国内外研究现状

### 1.2.1 基于深度摄像机的三维重建

可靠并且快速的对场景进行三维重建一直以来都是计算机视觉领域的一个热门问题，目前已经有很多方法来提供解决方案，但是它们都或多或少的存在一定的问题。传统的立体视觉方法通过多幅图像间的纹理匹配来完成对场景深度信息的计算，但是这项工作非常耗时，并且如果拍摄到的图像纹理特征不明显或者待重建物体有明显的闭塞结构时，便很难在多视角图像序列中找到图像间的一致性，从而使得这种方法失效。尽管可以通过增加图像的数目或提高相机分辨率来改善这一状况，但是方法固有的问题始终存在。

深度摄像机能够非常快速的获得场景的深度信息，利用其实时的重建算法，能够迅速计算场景的三维结构，并且不依赖物体表面的纹理信息。但是这一方法也由于设备的特性而存在固有噪声，尤其是在深度不连续的地方。因此国内外学者提出了多种方法来改善深度相机的输出结果。

将较高分辨率的彩色图像与深度相机获得的数据进行结合是其中一种比较流行的做法。在同一场景中，深度的不连续通常伴随着彩色图像的颜色或者亮度的变化，这样就允许场景中的信息从彩色图像对应到深度图像中，从而使边界区域恢复的更加精确，同时降低平滑表面的噪声。Huhle 等人在文献[1]中提出了结合高分辨率彩色相机和飞行时间相机来提供高分辨率三维模型的系统。该方法通过使用马尔科夫随机场框架，将相机提供的高分辨率彩色图像用于提高飞行时间相机输出的深度信息的分辨率，进而提高三维模型的重建质量。

Reulke在文献[2]对飞行时间相机的成像原理进行了介绍，并测试了其性能，他也将飞行时间相机与高分辨率彩色相机组合：在将两个相机系统进行校准之后，使用分辨率为的彩色图像对飞行时间相机输出的分辨率为的深度图的重采样做引导，进而获得更高分辨率的深度图像。这种方法能获得的三维模型相对单独使用飞行时间相机的结果在整体质量方面有了很大的提高，但是在最终模型的精确度方面还有待进一步考量。

Kuhnert 和Stommel在文献[3]中提出了融合立体视觉和飞行时间相机成像结果的方法。通过将飞行时间相机的深度测量结果与立体视觉方法的输出做映射，飞行时间相机与高分辨率相机对之间的对应关系就可以找到，这样，飞行时间相机的数据就能与彩色图像中的信息对联接。同时在这个框架中开发了一个用于初始化和限制层次立体匹配的算法。通过这种数据融合方法能够获得比只用飞行时间相机更高分辨率且更密集的深度数据结果。但是论文中独立的使用深度相机和立体视觉相机计算场景的深度（对立体视觉中可信度低的像素，即没有明显匹配窗口的位置，使用深度相机的数据进行填充），这使得重建的结果存在很大的噪声。

Hahne 和 Alexa在文献[4]中提出的方法与[3]中的相似，都是将飞行时间相机与两个标清摄像头组成的立体装置结合在一起。在对场景进行三维重建时，使用低分辨率的飞行时间相机初始化场景，然后利用彩色摄像头捕获的图像来完善高分辨率深度图像中的不连续的深度信息。这一设置的重要特征是能够记录视频序列，然后计算深度信息，这对动态场景三维重建提供了保障。

Kim 等人在文献[5]中将多个飞行时间相机与多个彩色摄像机结合，使用多视角飞行时间相机对场景进行测量，获得粗糙但是完整的三维模型，然后利用多视角彩色摄像机对获取的初始模型进行优化，最终获得高质量的稠密三维模型。但是从重建的结果来看（如图1-1），这种方法最终被深度相机和立体视觉方法存在的高频噪声所局限。

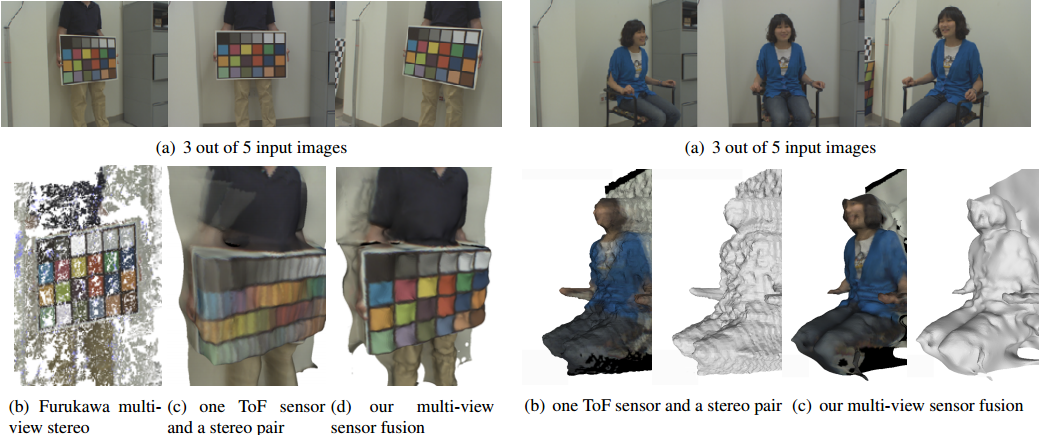


图1-1 文献[5]中展示的实验结果

文献[6]认为飞行时间相机与被动立体视觉的误差特性是互相补充的：飞行时间相机能够在被动立体视觉无法使用的情况下（例如在白色的墙面）获得深度数据；而在立体视觉擅长的纹理特征明显的区域，深度传感器存在很大的噪声。文章中对两种方法的概率密度函数进行了计算，然后进行融合，并通过实验证明，获取的重建结果比单独使用任何一种方法都要好（实验结果如图1-2），而且具有较好的鲁棒性。但是在复杂的场景，如：有镜面反射产生的高光或大范围的纹理变化，重建的精度就会大大降低。

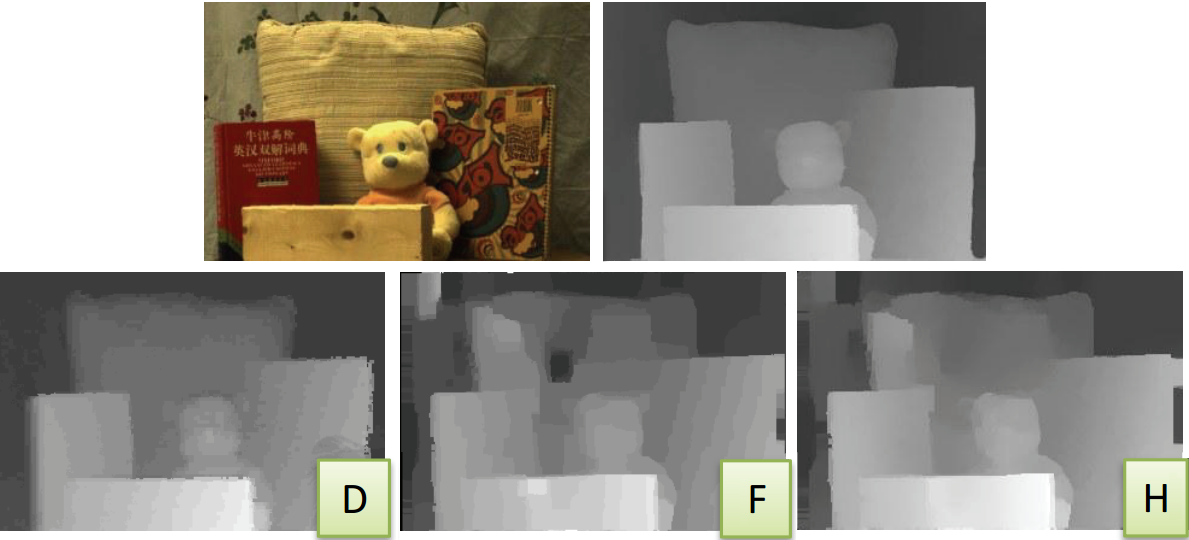


图1-2 文献[6]的实验结果：其中第一行分别为场景的彩色图像和真实的深度图；第二行中，D为深度传感器的输出结果，F为立体相机的输出结果，H为两种方法融合的结果。

从展示的实验结果中可以看出，论文[5][6]均受到了深度相机和立体视觉方法中存在的高频噪声的影响。

在论文[7]中提出了一个有效的结合深度相机捕获的深度数据和在不同照明条件下的彩色图像的方法，即将深度传感器与光度立体技术结合，从而显著提高三维重建的质量。质量的提高是通过对不连续和阴影区域进行明确的建模而实现的。论文中提出的融合方法能够在最小强度的光照变化甚至是环境光的条件下工作，而且计算开销较小。但是在阴影区域或局部深度分辨率较低的情况下，该方法不能正常工作。

文献[8]提出了一个基于阴影的形状细化算法，该算法使用Kinect输出的具有噪声且不完整的深度图来帮助明暗恢复法处理具有歧义的地方。文献[9]中也提出了多种方法来结合由光约束条件获得的物体表面法向和深度相机的输出数据。上述两种方法都是使用RGB图像与深度信息结合。其中论文[8]中使用的是Kinect获得的彩色图像，而论文[9]使用的是高分辨率彩色相机捕获场景信息。

Haque 在文献[10]中创新性得提出使用Kinect自带的红外相机捕获深度信息后，关闭红外发射器，使用红外摄像机捕获物体本身的红外辐射，也就是说，仅使用一个摄像头完成对物体的三维重建。

### 1.2.2 水下三维重建技术

在水下三维重建领域，常用的是声学三维重建光学三维重建两类技术。

#### 1.2.2.1水下声学三维重建

声学探测传感系统的水平分辨率和垂直分辨率均较低，也没有光学成像系统的较为真实的视觉效果，但是它能够在较大的范围内进行工作，即使在较差水质的状况中，仍能提供较为理想的三维信息。目前声学探测传感器在水下定位、水下导航、水下地形探测等方面有着广泛的应用。

自20世纪20年代以来，单波束声呐系统一直是应用最广泛的水下声呐系统之一。单波束声呐系统通常安装在船体下方，用于测量船体下方的竖直水深（如图1-3[11](a)）。单波束声呐系统获取的水下三位数据点较为稀疏，数据范围也较小，但是获取数据的成本低，方便与其他设备集成，数据的转化和处理也相对简单，因此在水下声学系统中有着广泛的应用。

单波束声呐系统利用声波的反射原理，每次只能探测一个波束覆盖的范围内的数据，探测范围较窄，数据的获取效率也低。20世纪70年代，出现了多波束声学探测系统（如图1-3[11](b)），它由发射换能器阵列和接收阵列组成，每次发射多个独立的声学波束组成的扇形，并通过接收的回波信号及波束角度和时间信息来计算深度。多波束声呐系统探测范围大，精度高，但硬件设备昂贵。

(a)单波束声呐工作示意图 (b)多波束声呐工作示意图

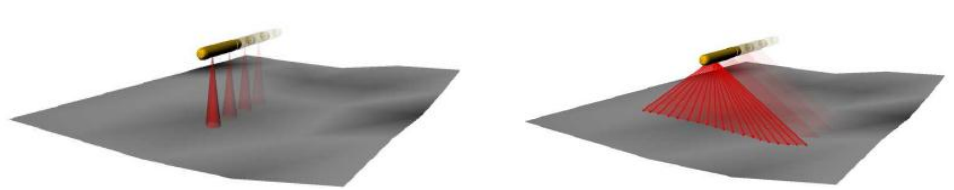
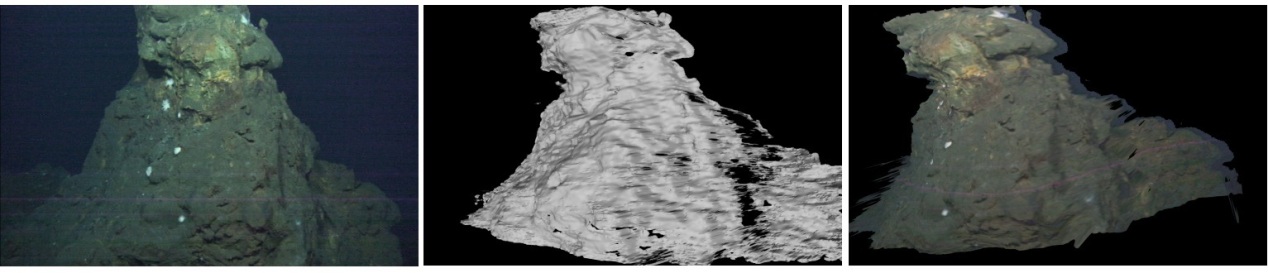


图1-3[11]两种声呐的工作示意图

#### 1.2.2.2水下光学三维重建

相比于空气中采集到的图像信息，水下的图像信息通常会产生较大的衰减。造成这一现象的原因有很多，如：物体表面接收到的光照不均匀、水对光线的散射和吸收以及水中悬浮微粒的存在等。这使得空气中的三维重建方法不能直接应用于水下光学三维重建。

文献[12]提出了一个对水下图像或者视频进三维重建的方法。在该系统中，除了一个被罩住的相机，无需其他设备。该论文中提出的算法本质上是传统的运动获取结构方法，为了适应水下的成像环境，对图像背景进行了滤波，从而减少悬浮颗粒对图像的影响，这样就能够鲁棒地获得相机的位置和稀疏的三维点云。根据相机追踪的计算，能够计算出物体表面的三维模型。重建结果如图1-4所示，从结果中可以看出，该方法的重建结果存在误差较大的地区。



(a)

(b)

(c)

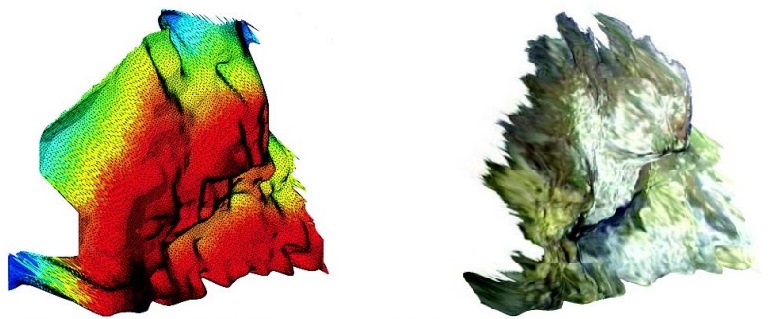
图1-4[12]其中(a)为输入的其中一幅彩色图像，(b)为恢复的未贴图的三维模型，(c)为贴图之后的模型在某一视角下的效果。

文献[13]中利用双目视觉技术与惯性传感器结合，成功地在水下和陆地重建物体完整的三维结构。但是该方法所需的设备复杂，需要融合位置传感器和相机的立体视觉信息，而且重建结果在很大程度上受水下位置传感器的误差影响。

Törnblom[14]在2010年提出了一个基于结构光的水下三维扫描方法。在该方法中，通过投影仪发射的结构化的光线照亮被扫描的物体表面，并使用照相机捕获物体表面反射的光线。通过投射一系列特定的线性图案，数字投影仪的列像素可以确定扫描表面的三维结构。但是从重建的结果来看，该水下结构光扫描方法无法覆盖所有的目标像素点，而且水下目标表面法向与结构光方向之间的角度会影响重建的精度。

随着水下三维重建技术的不断发展，加拿大2G Robotics公司[15]推出了基于结构光的多款水下三维扫描设备。其中ULS-500水下扫描仪对水下物体的纵向扫描精度在3mm以内，远优于声学探测设备，但是其扫描结果在高频具有较大噪声。

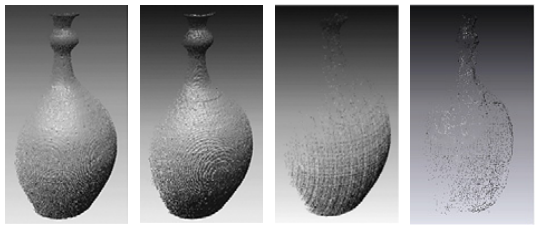
Brandou等人在文献[16]中提出了一个新的定量测量的图像采集方法，对小尺度场景的三维重建效果进行提高。该方法通过立体视觉系统获得图像序列，并保证每隔一定时间在预定的轨迹上获得物体在不同镜头下的图像。完整的三维重建方法产生稠密的有纹理贴图的三维模型，重建效果如图1-5。



(a) 重建后获得的三角网格图 (b) 纹理贴图后的效果

图1-5文献[16]的重建结果

Bruno在文献[17]中第一次提出使用主动光学技术来研究在水下环境中物体的三维重建。他在实验中测试了在空气中应用的基于投射结构光图案的立体视觉数据获取方法。并在水箱中使用不同浑浊程度的水进行测试，但是从文中给出的结果来看（如图1-6），该方法在水下环境的三维重建结果是非常稀疏的。



|  |  |  |  |
| --- | --- | --- | --- |
| (a) | (b) | (c) | (d) |

图1-6 文献[17]中展示的重建结果：(a)为空气中的重建结果、(b)为清水中的重建结果、(c)为浑水中的重建结果、(d)为非常浑浊的水中的重建结果

卡耐基梅隆大学的Narasimhan在文献[18]中提出，结构光在水下的应用中，其散射现象不能被忽略，因此在文章中综合分析了条纹光扫描技术和光度立体技术在散射介质中的应用方法，并推导出了散射介质中物体的物理成像模型，得出了散射介质中光度立体需要的光源数量及条纹光的使用条件。通过实验证明，当散射现象严重的时候，传统的三维重建方法会失效，但是论文中提出的方法能够精确的恢复场景（深度、法向、反照率）信息。

## 1.3 论文的创新性

深度相机体积较小、价格低廉并且能够实时获取场景的三维信息，成为三维重建领域的热门研究工具。但是其直接获取的数据往往存在较大的噪声。本文中使用深度相机输出的数据与多光谱光度立体技术结合，对空气中和水下的物体进行简易、快速的三维重建。本文的主要创新在于：

将深度相机与多光谱光度立体技术结合，对水下物体进行三维重建，考虑到水对光线存在的吸收和散射作用会使相机拍摄的水下图像质量大大下降，本文针对水对不同频率的光的吸收和散射作用的不同，建立了适用于多谱光源条件的水下成像模型，并将融合深度信息的多光谱光度立体技术应用到水下，通过实验证明了本文提出的方法能够以更加简单的操作、廉价的设备和较低的计算量，对水下的物体进行三维重建。

## 1.4 论文的主要内容及结构安排

本文使用深度相机结合多光谱光度立体技术对物体进行三维重建：

1）将多光谱光度立体技术应用到物体表面法向求解过程中：使用红、绿、蓝三种颜色的光源同时照亮物体，通过彩色图像通道分离，分别获得各光源照亮物体的图像，这样就能能够使用单幅彩色图像获得传统光度立体三幅图像的信息，进而可以使用单幅彩色图像对物体的表面法向进行求解。

2）将深度相机获取的场景粗糙深度数据与多光谱光度立体技术的计算结果进行融合：由于深度相机的数据在高频部分存在较大的噪声，但是低频区域是相对准确的，而光度立体技术在高频部分较精确，在低频位置却存在误差，因此，本文将两种数据分频段结合，并控制两部分的影响比例，从而获得优于单独使用任何一种方法的三维重建结果。

3）将结合深度信息的多光谱光度立体技术应用到水下物体的三维重建：由于水对光线存在吸收和散射作用，使得相机拍摄的水下图像质量大大下降，本文通过修改水下光学成像模型，使其适用于本文使用的多光谱光源条件，并通过实验证明了本文提出的方法的可靠性。

本文在结构上共分为五章：

第一章绪论：说明了本文研究内容的背景以及其研究意义，并对目前使用深度相机进行三维重建的方法及各类水下三维重建方法的国内外研究现状进行了综合介绍，同时阐述了本文的创新点，并在最后对文章的主要内容及结构安排进行了说明。

第二章光学三维重建技术：从主动光学重建与被动光学重建两个方面，介绍了目前常用的三维重建技术，并对其设备需求、计算特点及重建结果等特点进行了分析，进而体现出本文中选取深度相机结合多光谱光度立体技术进行水下三维重建的优势。

第三章融合深度信息和多光谱光度立体的三维重建。首先介绍了Kinect的数据获取原理和方法，然后将多光谱光度立体技术在空气中的原理进行了讲解，并将深度数据与多光谱光度立体技术获得的法向信息进行融合，展示了在空气中应用该方法的实验结果，从而为将该技术应用到水下提供了理论基础。

第四章水下三维重建：将第三章中在空气中获得良好实验结果的方法移植到水下环境，根据水下环境对不同波长光线存在的不同的吸收和散射特性，将水下光学成像模型进行了修改，并针对计算过程中存在的问题，对模型进行了简化，同时展示了在实验环境下对水下物体的三维重建结果。

第五章总结与展望。对本文中提及的工作进行了总结，并针对目前未解决的问题和可以提升结果的方面提出了今后的研究方向。

# 2 光学三维重建技术

随着电子技术的不断发展和图像获取技术的提高，高分辨率相机不断涌现，计算机技术、图像处理技术与机器视觉技术也得到了高速发展，使得光学三维重建技术在工业应用领域得到广泛的应用与推广，基于光学的三维成像技术也得到较大程度的提高。

光学三维重建技术一直以来都是计算机视觉领域的研究热点，它是以光学图像为基础，使用机器视觉技术、图像处理技术从二维图像中提取信息，利用空间几何方法获取三维空间信息。利用光学信息进行三维重建，不仅能够得到较高的水平分辨率，而且能获得较高的垂直分辨率和较高的光学特性。

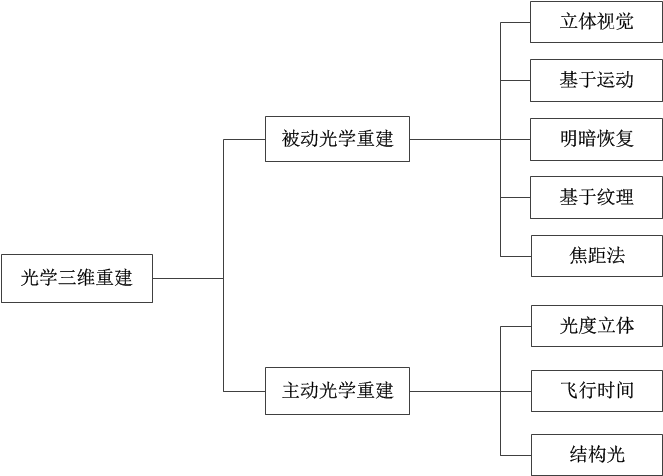


图2-1 光学三维重建技术分类

根据在采集数据过程中根据对光源的控制程度不同，光学三维重建技术可以分为被动光学重建技术和主动光学重建技术。

## 2.1 被动光学重建技术

所谓被动光学重建技术，就是在重建过程中，不需要对任何光源进行直接控制，而是完全依赖于环境光实现三维重建。

### 2.1.1 立体视觉技术

立体视觉技术的主要任务是同时采集多幅不同视点的图像，利用视差理论恢复场景的深度，并利用三角测量方法和相机的内外参数，重建世界坐标系下的三维坐标。立体视觉技术又可以分为双目立体视觉技术和多目立体视觉技术。

双目立体视觉技术模拟了人类的视觉过程其原理。图2-2中模拟的是两台相机观察同一点P的情形。

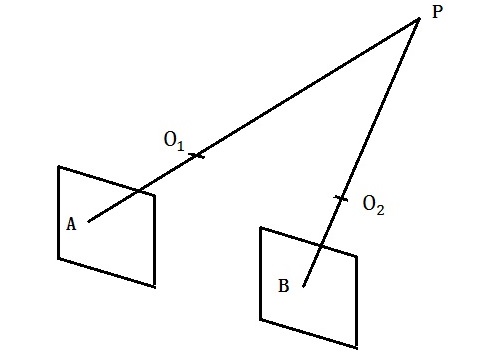


图2-2 双目视觉成像

点P在相机1中的成像点为A，它是由点P发出的光线经相机1的透镜中心与图像所在的平面的交点。反之，如果图像中已知点A的位置和相机透镜中心，便可以唯一确定一条射线，这样，所有可以在A点成像的物体，一定在这条射线上。此时，如果我们知道点P在相机2中的成像点B，那么根据点B与相机2的透镜中心可以确定另外一条射线。与的交点便是点P的位置。如果两台相机的几何位置已知，且相机是线性的，当同时知道一个物体在两台相机中的位置时，利用三角原理就可以将其空间位置计算出来。

Barnard在文献[19]中通过对广义的立体问题解决方案的研究，发现它在方法和目标都遵循一个相同的模式，这个模式可以划分为以下六个模块：

1）图像采集。立体图像可以通过多种方式获得：可以同时或者在任意时间间隔拍摄，或者通过观察位置和角度的转换来获得。图像采集过程中，可以产生影响的相关参数有：场景范围、时间差异、高光和阴影、光谱分布、相机相对位置等，通常是根据实际应用情况来做具体的取舍。

2）相机建模。相机的建模也可以称为相机校准或相机标定，主要目的是建立成像模型，并且确定相机的内参和外参，以便后续计算过程中能够正确建立物点与像点之间的对应关系。

3）特征采集。由于特征不明显的区域不能被可信的匹配，因此，在立体视觉的计算过程中，特征选择工作至关重要。一般来说，大尺度的特征数目较少，并且含有比较丰富的图像信息，因此容易得到较快的匹配，但是这种提取方法定位精度不高，且对特征的描述也相对复杂。小尺度特征的数目通常较多，能够提供较高精度的定位，对特征的描述比较简单，但是能够提供的信息量却大大减少。因此，在进行特征采集的过程中，需要采用较好的匹配策略和较强的约束准则。

4）图像匹配。图像匹配是场景分析的重要部分，它将采集的图像特征进行计算，从而建立图像特征之间的对应关系。这样，就能够将空间中的点与不同图像中的成像点对应起来，并通过这种对应关系得到视差图像。图像匹配主要有特征匹配和灰度相关两种方法。

5）距离判定。一旦计算出视差图，场景的深度（高度）信息便可以确定，这样就能够对场景进行三维重建。

6）后期处理。经过上述步骤获取的三维信息会因相机标定、特征检测等过程中的精确度问题而产生误差或深度信息不完整现象，这样就需要对计算得到的深度数据行进一步的处理。常用的处理方法是对深度数据进行插值、各部分误差的校正以及精度的改善等。

由于传统双目视觉技术在立体匹配的准确性和稳定性中存在一些问题，引入了三目视觉系统[20]和多目视觉技术[21]，它们的技术的关键都是解决立体匹配的问题。这一问题已然成为立体视觉领域中一个具有挑战性的课题。

由介绍可知，使用立体视觉方法对物体进行三维重建时需要采集多幅图像，还需对相机的相对位置进行确定和相机校准等操作；在特征点的采集和匹配过程中需要进行大量的计算，而重建的结果也会因特征点的选择和计算情况受到影响。此外，在对具有周期性纹理或者纹理信息不明显的物体进行三维重建时，立体视觉的重建效果便会大大下降。立体视觉方法得到的重建结果往往是非常稀疏的，在对计算结果进行插值、误差校正以及精度改善的过程中，可能会引入新的误差。这与本文中以简单、快速且精确的方法对物体进行三维重建的目标有出入，因此，并未选择立体视觉方法。

### 2.1.2 运动恢复法

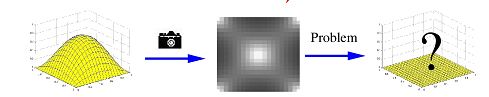
基于运动的光学三维重建技术[22]是指从二维图像序列中恢复三维空间结构的过程。在重建过程中，摄像机按照固定或非固定的路径对目标场景进行拍摄，从拍摄到的视频序列中进行空间对应点的匹配，从而建立摄像机运动和目标场景的空间几何关系，并从目标场景与相机间的相对运动约束关系中恢复场景的三维信息。

基于运动的光学三维重建过程具有对待重建场景光照环境要求低，对摄像机校准精度要求不高，操作灵活等方面的优点，尤其适用于对刚体目标场景进行三维重建。然而对于非刚体的物体，重建效果会受到影响；同时也存在图像序列中关键点的匹配和跟踪难度较大，重建模型复杂，需要较大的计算量等方面的缺点。

### 2.1.3 明暗恢复法

当物体被光源照射时，由于物体的表面法向与光线方向的夹角不同，会使物体表面的亮度产生差异，这种亮度在空间中的变化就是我们常说的物体的阴影。在物体成像后，物体的阴影便以图像中的阴影的形式再次表现出来。具体来说，当三维物体投影到二维平面时，会形成亮度的变化，这些变化通常以灰度的形式来表示，并被将其称作阴影。阴影的产生主要取决于四个因素，即：物体在成像工具可见表面的几何形状；物体的表面反射特性；光源的入射方向和强度；以及观察者的位置和相对物体的距离。

明暗恢复法描述的问题可以由图2-3展示，即利用相机拍摄的单幅灰度图像，根据图像的明暗变化与目标场景的表面梯度对应关系，对物体表面进行三维重建。



表面

图像

由图像恢复表面

图2-3 明暗恢复法

传统的明暗恢复法假设物体符合朗伯反射模型（漫反射模型），利用图像的亮度变化和光源的方向来确定物体的表面法向，进而通过积分计算出物体的表面高度。但是，明暗恢复法的解是不唯一的[23]，仅当物体表面反射特性或形状特性已知时，根据这些特性的约束，建立正则化模型，才可以解决这一问题。

### 2.1.4 纹理与表面朝向法

利用物体表面的纹理特征来帮助确定物体表面的方向信息，进而恢复物体表面形状的方法就是纹理与表面朝向法。纹理是由纹理基元组成的，它是图像中带有不变性和重复性的视觉基元。其中重复性指的是纹理基元在图像不同的方向和位置中重复出现，并且这中重复是在一定的分辨率下；而不变性指同一纹理基元内的像素在灰度或形状等方面，具有一些相同的特性。

利用物体的表面纹理确定其朝向信息需要满足一定的条件：在图像的获取过程中发生了透视投影，物体的纹理结构发生投影变换，这种变换因纹理所在表面的法向不同而有所差异，这样图像中就会包含物体表面的方向信息。值得注意的是，这里所说的不是物体的纹理带有三维信息，而是在成像过程中纹理产生的变化会包含三维信息。纹理的变化可以大致分为三类，对应三种常用的恢复方法：纹理基元尺寸变化（图2-4(a)）、纹理基元形状变化（图2-4(b)）和纹理基元之间的关系变化（图2-4(c)）。

x

o

y

|  |  |  |
| --- | --- | --- |
| (a) | (b) | (c) |

图2-4 纹理的三类变化

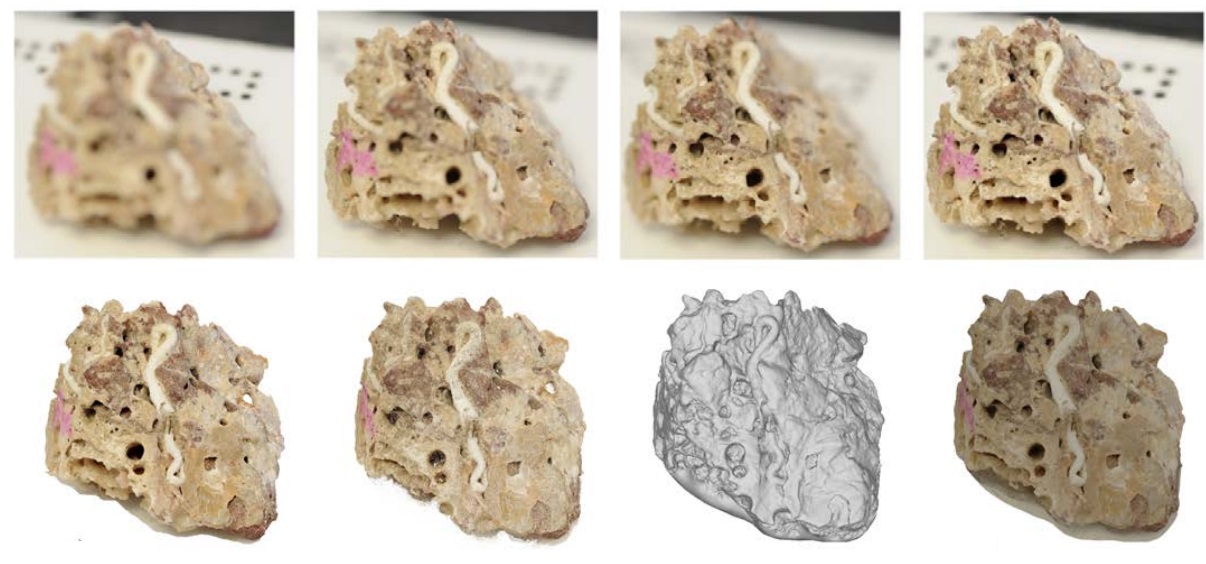
纹理与表面朝向法对待重建物体的表面纹理特性有较高的要求，对于没有纹理或者纹理特性复杂的物体，便不能得到理想的重建结果。然而，水下环境通常都是被泥沙覆盖的表面或者不规则的石块，使用纹理与表面朝向法显然不能获得理想的效果。

### 2.1.5 焦距法

焦距法三维重建是指利用相机的焦距信息来进行三维重建，又可以分为聚焦法和离焦法。

聚焦法是将被观测的点放置到相机的聚焦点上，然后根据相机的参数和透镜成像原理的公式，对空间被测点相对于相机的深度进行计算。在这种方法中，如果相机偏离聚焦位置，就会引入深度计算的误差。

离焦法与聚焦法不同，不需要通过待测空间点的聚焦位置来计算该点的深度，而是对偏离相机焦距平面的点进行标定，然后根据空间点偏离聚焦平面的模型信息，对被测空间点的深度信息进行计算。其中，对相机离焦模型的标定是利用离焦法实现三维重建的难点所在。



(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

图2-5 离焦法三维重建结果[24] 其中 (a)~(d)分别对应远焦距图像、中焦距图像、金焦距图像和最终图像；(e)~(h)为处理后的输入图像、三维点云、重建的网格以及贴图后的三维模型。

焦距法对物体三维重建的过程中，需要物体与相机之间相对位置不变，同时拍摄多幅不同焦距下的图像，这对拍摄设备和被拍摄的物体都有较大的局限，且操作繁琐不适合在水下环境中使用。

## 2.2 主动光学重建技术

主动光学重建技术与被动光学重建技术的主要区别在于，主动光学重建技术在数据采集过程中需要引入专门的光源进行照明，进而利用物体对光的反射或者散射进行三维重建。

### 2.2.1 光度立体技术

由2.1.3节中的描述可知，明暗恢复法利用单幅图像进行目标物体的三维重建是一个欠约束的问题，因为单幅图像会丢失很多三维信息。既然利用单幅图像恢复三维信息是一个欠约束的问题，那么在研究过程中，利用同一摄像机拍摄相同物体在不同光照条件下的多幅图像，便可以增加约束条件，光度立体技术也就应运而生。

光度立体技术最早由Woodham[25]在1980年提出，该技术利用相机从同一方向拍摄待重建物体在不同光照方向下的照片（拍摄过程中要求相机与物体的相对位置不变，只改变光照方向）获取物体的形状信息，进而对目标物体进行三维重建。自此，光度立体技术成为了计算机视觉领域的热门研究问题，大量学者针对光度立体技术开展了研究。

光度立体算法大致可以分为基于光照模型的方法和基于参照物的方法两类。

1）基于光照模型的方法：通过获取同一物体在相同视角但在不同光照方向下的多幅图像，根据朗伯反射模型，可以获得像素值与入射光线向量、物体表面法向量关系的方程组，进而计算出物体的表面梯度。



图2-6 [26]中的部分输入图像和重建结果

2）基于参照物的方法。在场景中放置一个与待重建物体材料相同且形状已知的参照物，通过物体与参照物的特征向量匹配，获得与待重建物体上的点相匹配的参照物上的点，由于参照物的形状已知，便可以对其表面法向进行求解，并直接将参照物上的法向复制为待重建物体的法向，这样就将目标物体的法向计算问题转换为物体与参照物的点的搜索与匹配问题。在进行搜索时，需要对待重建物体与参照物上的点的特征向量进行匹配度计算，并将其作为最佳匹配的标准。一旦找到最佳匹配，便可以利用匹配点进行法向值的计算。



图2-7[27]中使用的参照物及重建结果

在使用过程中，光度立体技术与立体视觉技术相比，不需要过多的中间校准步骤和后处理过程，并且能够获得较精准的重建结果。而且具有所需设备简单，拍摄过程简易，数据处理方式简明，适用范围广的特点，因此，本文中选取光度立体技术作为三维信息的获取方式之一。

### 2.2.2 飞行时间

飞行时间（Time-of-Flight）就是利用已知的光速来计算距离，也就是通过计算光信号在发射端与目标物体之间的飞行时间来确定距离。伴随着半导体工艺的快速发展，飞行时间相机（Time-of-Flight camera）在2000年左右开始在民用领域使用，其测量范围可以由几厘米到几公里不等[28]。市场上常见的飞行时间相机的分辨率是或，与常见的2D摄像机相比其分辨率是较低的。但是，飞行时间相机获取三维信息的速度是非常快的，通常能够提供高达每秒160幅图像的速度。



SwissRanger 4000 [29] PMD CamCube[30] FOTONIC-B70[31]

图2-8 几种飞行时间相机

目前已经有几种不同技术类型的飞行时间相机，通常由以下几个部分组成：

1）照明装置：用于照亮整个场景，通常使用的是红外线，使得发出的光线在物体表面并不明显。相位探测器成像系统的光源必须被调制到高达100MHz的速度。对于直接成像的飞行时间照相机，使用的是每帧显示单个脉冲（例如30HZ）。

2）光学设备：由一个镜头捕获被物体反射的光线，并且将环境信息在图像传感器中成像（使用的是焦平面阵列）。一个光学带通滤波器，仅允许与照亮环境的光的波长相同的光线通过，这有助于抑制无关的光线，从而降低图像噪声。

3）图像传感器：这是飞行时间相机的核心部件。用于计算图像中每个像素点光线离开照明设备到达物体并且返回焦平面矩阵所用的时间。

4）电子线路驱动：照明设备和图像传感器必须使用高速的信号进行控制和同步。为了获得较高的分辨率，这些信号必须非常精确。例如：如果照明信号与成像传感器之前的信号平移了10毫秒，距离就会有1.5毫秒的变化。作为比较：目前的CPU频率高达3GHz，对应大约300ps的时钟时间，相应的分辨率误差是45毫米。

5）运算与界面：在此阶段，距离在相机中直接被计算出来，为了获得更好的性能，也会使用一些校准数据，然后相机会通过界面输出距离图像。

飞行时间相机的优势：

1）装置简易：与立体视觉或者三角测量系统不同，飞行时间相机的整个系统非常简洁，照明装置就放置在镜头旁边，不需要进行手动的调节。并且在使用过程中不需要对相机进行校准，或对相机进行机械的移动。

2）计算方法高效：飞行时间相机对距离的计算是一个直接从飞行时间相机输出信号提取信息的过程。因此，这项技术只需要较小的处理能力，而且不会因物体表面的纹理或图案的存在对计算结果产生影响。

3）速度快：飞行时间相机能够在单个脉冲中计算整个场景的距离。每秒钟可以达到160帧，因此非常适用于在实时处理的情况中使用。

相对于传统的三维重建方法，飞行时间相机从设备、操作、计算几个方面，都有较大程度的提高，但是还是存在一些不足的地方，例如：背景光的影响无法完全消除，多个飞行时间相机同时运行时会产生互相干扰，以及光线在场景中多次反射对结果造成影响等。但是这些不足也正在因硬件设备的发展以及通过时间复用等方式而不断得到弥补[32]。

### 2.2.3 结构光测量法

结构光测量法是目前应用较为广泛、技术也相对成熟的一种三维测量方法，理论基础是光学三角法原理[33][34]，利用产生结构光的光源与相机之间的三角几何信息，对物体的三维信息进行测量。根据结构光形态的不同，可以分为点式结构光，线式结构光和光栅式结构光。在实际应用中，会根据测量对象的不同而选用不同的结构光。

利用结构光对物体进行三维重建，需要对相机和结构光的光线进行校准；然后对相机的每一个像素点在空间三维坐标系内的反向投影射线的解析表达式和结构光光线的空间解析表达式进行求解；最后，在像素坐标系内对结构光的特征点进行求解，并根据结构光在空间三维坐标系中的解析表达式和反向投影射线对目标物体的三维空间结构进行求解。

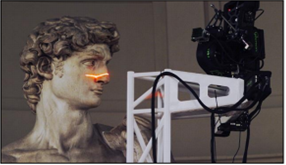


图2-9 利用结构光对文物进行扫描[35]

传统的结构光测量方法主要是使用结构光对物体进行扫描，这样得到的重建结构精度较高，但是需要使用结构光对整个目标物体进行遍历，因此大大降低了重建的效率。为了提高重建效率，出现了对结构光进行二进制编码的重建方法[36]（图2-10），和基于彩色编码的结构光三维重建方法[37]（图2-11）。

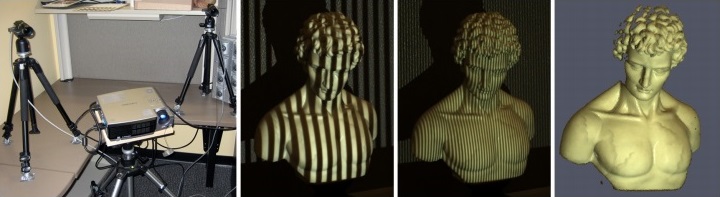


图2-10 二进制结构光编码三维重建[36]



图2-11 彩色结构光编码三维重建[37]

本文中深度相机也是使用的结构光原理，文章将在第3.1节中做详细的介绍。

## 2.3 小结

在本章中，我们介绍了目前常见的光学三维重建方法，通过对本章内容的分析可以知道，被动光学重建技术在对场景进行三维重建的过程中，虽然不需要使用特定的光源，并且对成像设备的要求也并不高，但是针对没有显著特征的图像（例如没有明显边缘、角点或纹理信息），图像匹配的准确度很难保证；并且在重建的过程中，需要进行大量的计算。而主动光学重建技术通过引入特定的光源，大大提高了成像的精确度和速度，同时具有较高的可靠性，从而使得快速的高精度的三维重建成为了可能。同时由于硬件设备的快速发展，对场景的实时三维重建也成为了可能。

综合分析后，我们可以发现，深度相机设备可以快速获得场景的粗糙的深度信息，这些信息虽然在细节方面存在较大误差，甚至是无法提供细节信息，但是它们提供的数据在低频区域是非常准确的，并且由于其自身具有实时处理的算法，能够实现对场景的实时三维重建。而光度立体技术在使用过程中，可以通过频分复用的方法，解决需要同时获取多幅图像的需求，从而在视频化处理的应用中具有了使用前景。同时，通过光度立体方法获取的三维信息在高频区域具有较准确的信息，而对于低频的数据却不够准确。这就使我们想到，可以利用光度立体技术与深度摄像机进行结合，利用深度信息与位置信息结合的算法[38]，获取更好的重建结果。

在下一章我们将针对多光谱的光度立体技术进行介绍，并详细说明如何将光度立体技术与深度相机的输出数据进行融合，从而获得更高分辨率的三维模型。并对该方法的重建结果进行了展示。

# 3 融合深度信息的多光谱光度立体三维重建

## 3.1 深度信息的获取

由第二章中的介绍可以知道，深度相机能够对场景的深度数据进行实时获取。考虑到飞行时间相机的价格昂贵，本文中选用价格较便宜、操作方便，并且具有二次开发功能的Kinect作为初始深度的获取工具。

### 3.1.1 Kinect深度测距原理

Kinect的光学成像系统中包含一个红外激光发射器、一个彩色摄像头和一个红外摄像头（其对应位置如图3-1）。其中，由红外激光发射器和红外摄像头结合来产生深度图像。



图3-1 Kinect

Kinect在对场景进行深度测距时，使用的是结构光技术，更准确的说，是一种光编码（Light Coding）技术，这是一种通过光源照明对空间进行编码的技术。与传统的结构光测量方法不同，Kinect使用的是一种具有三维纵深的“立体编码”，而不是周期性变化的二维图像编码。具体来说，Kinect 使用红外激光发射器对外发射单束红外激光，这束激光在到达物体之前，会经过发射器前的光栅，从而产生衍射效应，进而形成激光散斑。这种散斑具有高度的随机性，而且图案会随着距离的变化而产生变化。这也就是说，在同一空间的任何两个地方的散斑图案都是不相同的。通过这种方式，整个空间都被这种斑点做了标记。在进行深度测距之前，不同距离下的散斑模式需要被记录下来。为此，需要对光源进行标定，具体做法如下：在场景中，每隔一段距离取一个参考平面，并将该平面的散斑图案做记录。在进行深度测量时，只要通过红外摄像头捕获到物体表面的散斑图案，就能够根据图案的变化得知物体的位置信息。



(a)彩色图像 (b)散斑图像 (c)深度图像

图3-2 Kinect 获取的三种图像

### 3.1.2 Kinect误差分析

使用Kinect捕获深度数据，误差主要来源于传感器、测量条件及物体表面特性三个方面。其中传感器误差属于系统误差，主要影响单个数据点的精度。测量条件对Kinect的输出结果影响较大，如：过强的环境光会使散斑图案的对比度降低，从而影响到捕获数据的准确性；场景中若存在多个Kinect，则会产生相互干扰。由于Kinect的激光红外发射器和红外摄像头不在同一视点上，场景中物体可能对红外摄像头可见但对激光红外发射器不可见，这就会使红外图像中某些区域没有散斑图案，从而无法产生有效的深度值，这就导致了空洞的产生。此外，一些黑色物体会吸收大部分的光，以致红外相机接收的光线数量过少；一些光滑的表面反射绝大部分光，以致红外图像曝光过量，这两种情况也会导致空洞的产生。

## 3.2 朗伯模型与多光谱光度立体

### 3.2.1 朗伯模型

理论上来讲，如果在目标场景中，物体之间的多次反射可以忽略，那么物体表面反射到相机中的能量就仅与相机的位置、入射光线的强度、目标物体的表面反照率及目标物体的表面法向有关。用来表示这些参数之间关系的函数就是物体的表面反射模型。

物体的表面反射模型能够反应入射光线与反射光线之间的关系，其中镜面反射模型和朗伯模型是目前常见的光学反射模型。朗伯模型是由Johann Heinrich Lambert[39]于1760年提出的，它是一种理想的漫反射模型。在朗伯反射模型中，假设光源位于无穷远处，且入射光线被物体表面在各个方向上均匀反射；而镜面反射只在单一方向对光线进行反射。如图3-3所示：镜面反射的结果是仅能从一个方向观察到反射光线，而朗伯模型能够在各个角度观察到物体的反射光线，且亮度相同。

n

s

s

n

r





(a)漫反射

(b)镜面反射

图3-3 两种反射模型

对于朗伯反射模型，相机捕获的物体表面的亮度可以由朗伯反射公式给出：

(3-1)

其中，表示图像中点的亮度（灰度）值，是物体的表面反照率，是物体在点处的表面法向量，即

其中和是物体表面高度关于和方向的偏导数，即，，其中，就是物体的表面高度函数。是场景中的单一方向的光源。

### 3.2.2 多光谱光度立体

在传统的光度立体技术中，假设待重建物体符合朗伯反射模型，并且没有环境光照，在同一视角拍摄物体分别被三个非共面的光源照射的图像，其中三个光源均具有相同的频谱，并可以被认为是远点光源。

我们用 来表示第 幅图像中点处像素的亮度（其中）。假设在第 幅图像中，光源的方向为，光源光谱分布为，物体表面对不同波长的入射光线吸收情况用反射函数表示。此外，我们用 表示相机传感器对不同波长的响应， 表示物体局部的表面法向。这样，物体表面像素的亮度 就可以用下面的公式来表示：

|  |  |
| --- | --- |
|  | (3-2) |

其中，积分的值就是我们通常说的表面反照率，这样，公式 (3-2) 可以用一个简单的点积来表示：

(3-3)

如果我们将 写成 表示图像中某一点接收的三个光源的光照方向，并且 表示该点在三个通道的像素亮度，这样，上面的公式就只有一个解：

|  |  |
| --- | --- |
|  | (3-4) |

一旦我们求出了法向，物体的表面就能够由法向域的积分恢复得到。

使用多光谱光度立体技术[40]来捕捉物体的形状，就是将传统光度立体中的时分复用改为频分复用：使用红、绿、蓝三种颜色的光源同时照亮待重建的物体（如图3-4），并且假设物体按照朗伯反射规律对这些带颜色的光进行反射，且各光谱之间没有频率的混合。同时，红、绿、蓝三种颜色灯的反射量是沿物体表面法向的线性方程，并且满足非共面、频谱独立、可以看作远点光源三个条件。这样，在采集到的单幅RGB图像中，便可以通过通道的分离获得三幅具有不同光照方向的图像。由于彩色光度立体只依赖于单幅图像，不需要在每一帧之间使用复杂的光照变换，因此可以应用于视频处理。在多光谱光度立体技术中，每一帧图像中获得的信息都相当于是传统光度立体三个相机获得的信息。

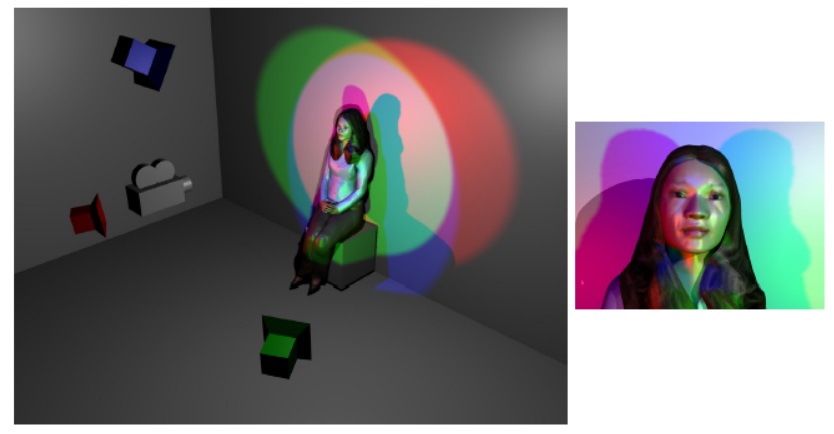


图3-4 多光谱光度立体三维重建实验环境模拟[41]

在多光谱光度立体中，像素在第 *i* 个通道中的亮度可以表示为：

|  |  |
| --- | --- |
|  | (3-5) |

现在，相机敏感性 和光源光谱分布 对每个相机和每个光源都是各不相同的。为了能够确定RGB值与法向的唯一映射，我们假设物体是单色的表面。这样我们令 ，其中 是单色的待重建表面的反照率，是待重建表面材料的颜色特征。令

|  |  |
| --- | --- |
|  | (3-6) |

表示矩阵中第行，第 列的元素。这样，在每个像素点，三个传感器的相应向量可以用下面的公式表示：

(3-7)

这样为了计算 只要求出矩阵即可。

### 3.2.3 多光谱光度立体的求解

对于矩阵的计算的方法基于论文[42]，与该论文不同的是，我们的初始粗略几何由深度摄像机直接获得。

与传统光度立体算法相似，多光谱光度立体的核心步骤是基于公式(3-7)，为了使用这个等式来执行光度立体算法，给定的输入是获得的图像亮度 和光照矩阵，未知的只有法向。为了完成校准的目的，需要输入成对的彩色图片的亮度和表面法向，来对未知的光照矩阵进行求解。

如果我们给定三个点 ，，，它们的表面反照率未知，但是都是相等的；它们的法向是，，（非共面）；相应的捕获的RGB图像亮度是，，。我们可以用下面的公式唯一确定：

(3-8)

对于多幅图像，相同的三个点能够通过相同的比例因子在每幅图像中提供光照矩阵。现在的问题是，我们怎样获得这样的三个点。答案是：如果在初步获得的粗略的几何模型中包含了足够多的准确点或者有效值，这样在所有的点中以三个点为一组进行随机反复取样，将有很高的可能性在取的所有点中，有一组数据的三个点全部为有效值。

同时我们知道，无效数据不会对任何特定的光照模型产生一致的影响，然而有效值对模型的影响是一致的。这样的一个发现使得论文[42]使用了随机抽样一致性算法（RANSAC）来分离有效值和无效值，并且用于计算光照矩阵。这个算法可以用下面的步骤概括描述：

1）在待重建物体表面随机抽取三个点，用它们的RGB亮度和法向为 估计一个值；

2）待重建物体表面的每一个点 都将用来验证这个假设预测的值与Kinect获得的图片亮度的差是否在给定的阈值范围内。

其中 中包含了量化误差，图像的噪声等。

3）重复步骤1和2一定的次数，使得光照矩阵的假设始终保持最大数目的支持。

值得注意的是，尽管我们在计算的过程中使用的是最简单光照模型，也就是一个3\*3的矩阵，这个算法也可以很简单的扩展到一阶球谐光照模型，也就是3\*4的矩阵来对三个点光源和环境光进行建模。使用随机抽样一致算法的方法也是一样的，只是现在需要选择4个点而不是三个来建立光照模型。然而，在所有的实验中，环境光是微不足道的，因此这种扩展是没有必要的[43]。

在上面的计算中，我们假设待重建的物体颜色一致，且具有相同的表面反照率，因此使用随机抽样一致算法得到对的多个估计之后，可以通过最小二乘法获得最优解，进而对场景的法向进行求解。然而，在实验中，待重建的物体往往不只具备一种颜色或材质，这样，就必须将上面的算法向多色场景进行扩展。

我们令[44]，使用前面的方法，随机抽取三对数据对对进行求解，利用随机抽样一致算法，找到求得的中有最多像素点支持的那一个，然后考虑删除满足下面条件的点：

|  |  |
| --- | --- |
|  | (3-9) |

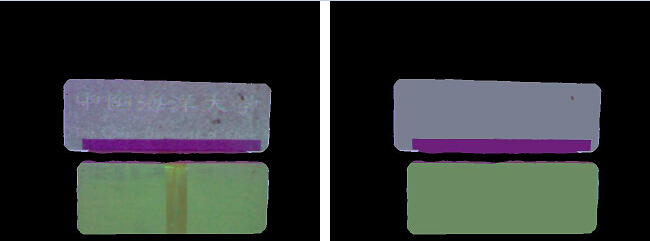
这些满足条件的点已经对进行了很好的建模。在这种情况下，起作用的只有颜色因素，而表面反照率不再有影响。所有具有相同颜色的点被移除，与反照率无关。当所有满足该条件的点被移除之后，继续对场景中第二显性的颜色寻找校准矩阵，采用与前面相同的计算方法，直到图像中所有的点被移除，或者满足停止循环的条件。

### 3.2.4 多光谱光度立体三维重建

#### 3.2.4.1图像分割

在将多光谱光度立体技术应用到多彩色物体的重建应用中，由于计算出的值不止一个，因此，为了给值找到最佳的计算区域，需要对图像进行分割。由于Kinect获取的图像在局部会存在颜色不均匀的现象，为了保证分割效果，我们首先对图像使用mean-shift算法进行初步的聚类。

Mean Shift算法（均值偏移算法），由Fukunaga[45]在1975年提出。它是一种无参估计方法，它不需要提供聚类数目，通过概率梯度上升方向来寻找数据分布的峰值。它是一个迭代过程：首先计算出当前点的偏移均值，并将该点移动到该偏移均值，然后以此点为新的起点继续移动，直到满足结束条件。



(a)输入图像 (b)聚类后的图像

图3-5 mean-shift算法效果

如图3-5，经过mean-shift算法聚类之后，图像中相似颜色区域会由同一种颜色表示。图像聚类之后，我们使用k-means算法根据求出的的个数，对图像进行分割。

K-means（K均值)算法在1967年由James MacQueen[46]提出，是一个应用最广泛的聚类算法，它是一种硬聚类算法。在聚类过程中，首先随机选择K个对象，每个对象代表一个初始的平均值或中心。根据剩余的每个对象到中心的距离，把他们分给距离最小的中心，然后重新计算每个区域的平均值，重复上述计算过程，直到聚类的准则函数收敛。

我们将M的个数赋值给K，便可以将图像分为K个部分（或小于K，当出现空类时，我们将该类舍弃），并对每个部分进行编号。

#### 3.2.4.2法向计算

在对法向进行计算之前，首先对深度相机输出的深度信息进行适当尺度的平滑操作，以降低高频噪声的影响。然后使用平滑之后的深度数据计算法向，同时对彩色图像进行与深度图像相同尺度的平滑，得到平滑之后的图像亮度值。我们将分割之后的图像的每一部分均与进行下面的计算：

|  |  |
| --- | --- |
|  | (3-10) |

记录在每个颜色区域产生的误差均值，完成全部运算后，在该颜色区域拥有最小误差的将用于计算该区域的法向。这样我们就使用多光谱光度立体方法获得了物体的法向数据。

## 3.3 深度信息与光度立体的融合

由第二章中对三维重建技术的分析可以知道，使用深度相机直接获取的三维数据存在较多的高频噪声，但是在低频区域数据是相对准确的；然而使用光度立体技术对物体进行三维重建时，高频数据是精确的，在低频区域却存在较大的误差。因此我们需要找到一种方法，区别对待两种方式产生的高频和低频信息，并对他们运用最佳的方式进行融合，使最终的结果比单独使用任何一种数据均有更好的效果。

### 3.3.1 使用位置信息改善法向

Rushmeier 和 Bernardini[47]在1999年提出了一个利用测量位置的数据来消除法向误差的方法，这种方法是特别针对光度立体环境的。Nehab等人在[38]中对这种方法进行了改进，使它适用于更广范围的输入数据。

正如我们之前描述的，使用光度立体方法计算的法向存在低频误差。另一方面，从深度信息中计算得来的法向的误差是高频的。通过结合适当的频率，我们能够获得更高质量的法向估计。

我们使用和分别表示通过深度相机获得的法向和计算得到的法向。从理论上将，我们希望使用中的低频替换中的低频部分。为了达到这个目的，我们对数据进行相同尺度的平滑，这个平滑需要足够大，以保证能够消除中的高频噪声以及中的高频细节。平滑之后的结果和相当于物体的低频部分。

然后我们计算旋转域用来描述将的每一个法向移动到的对应位置中进行的旋转。最后，我们计算修正后的法向域，通过把旋转域应用到平滑之后的法向域（从位置信息中获得的）。旋转域获得的高频细节，但是没有低频信息。细节信息前后一致的从一个法向域转移到另外一个，不管相应平滑域的角距离。（由于法向域只在大幅平滑之后使用，它可以通过任何方式从位置信息中获得）

### 3.3.2 使用法向信息改善位置

我们由深度相机获得的是一幅距离图像。物体表面的每一个像素点的坐标都是在相机参考系中的。因此，在透视投影中，物体表面点的坐标可以写成深度的函数。也就是说，给定一个像素的坐标，相应的表面点的位置只有一个自由度，：

|  |  |
| --- | --- |
|  | (3-11) |

其中和 是以像素为单位表示的相机焦距。我们要解决的问题是找到一个深度函数，符合我们对位置和法向在每一点的估计。为了完成这个目标，我们选择一个深度函数，使它满足拥有的位置误差和法向误差的和最小。

位置误差被定义为最佳位置与测量位置距离的平方：

|  |  |
| --- | --- |
|  | (3-12) |

其中是第个最优距离，是相应的测量距离。为了评估位置的误差，我们把深度值通过中心投影进行转换：

(3-13)

其中：

|  |  |
| --- | --- |
|  | (3-14) |

法向的误差可以用几种不同的方式定义，包括最优平面的对应法向与校正后的法向域的角度误差和；或者校准与未校准对之间的距离平方和。然而，多数公式会产生非线性的优化问题，这与我们的数据集的尺寸是不符合的。我们的解决方案是考虑优化表面的切线。修正的法向和最优平面的切线应该是垂直的。给定像素的表面的切线 和 可以写作一个深度值偏导数的线性方程：

|  |  |
| --- | --- |
|  | (3-15) |
|  | (3-16) |

定义：

(3-17)

其中， 对应的是点校准后的法向。 是最优表面的切线到校准后的法向投影的平方和。当所有的切线与校准的法向都垂直时，它的值是最小的。为了估计和，我们计算深度函数的偏导数，替代公式(3-15)和(3-16)。我们可以用3\*3的邻域来近似偏导数，下面的卷积核，假设所有的邻域都是可以得到的：

|  |  |
| --- | --- |
| 3-18.png | (3-18) |

有时候，在边缘和深度不连续的地方，有些邻域可能不存在。我们可以通过分析测量值，并且使用最佳离散导数，或者简单的单边导数来检测这些情况。在没有邻域的情况下，这个点可以通过最小值的判断来删除。

这样，最优表面可以用下面的公式给出：

(3-19)

其中，参数用来控制位置和法向信息在优化中起到多大作用。当是时，算法只考虑法向，测量的位置只在边界条件下有帮助，很像明暗恢复法。当是是，算法仅返回原来的位置。对于中间值，我们的方法找到最优法向和位置的权值。

每个像素点生成最多3个方程：一个是位置误差，一个是在x方向和y方向的法向误差。在取平方之前，误差项的方程都是深度值的线性方程。因此，整个最小化的过程可以用公式表示为一个大的过约束线性系统，可以通过最小二乘法来解决。

(3-20)

在这里，表示这样一个矩阵，当与一个未知向量Z相乘，每个点产生两行向量，来评估法向约束和。注意和在Z中是线性的，是独立于Z的。此外，这个方法可以被扩展到加权最小二乘法。

每个像素3个方程，一个的扫描可能产生成千上万个方程。幸运的是，这些矩阵是非常稀疏的。事实上，非零项的数量是像素数量的线性关系，因为每一行最多有7个非零项（一个系数是像素的深度，和至多6个用于相邻像素寻找偏导数）。Paige和Saunders[48]实现了他们提出的共轭梯度方法，来解决稀疏最小二乘问题可以很容易地处理这种类型的系统。处理之后的的深度图像在位置和法向上都是很精确的。

## 3.4 实验结果

我们的实验装置由三盏分散的红、绿、蓝三种颜色的LED灯，以及一个Kinect组成。待重建的物体被三盏灯同时照亮，使用Kinect捕获深度与彩色信息。由于我们基于彩色图像的三个通道的信息对物体进行三维重建，且假设物体符合的是朗伯模型，因此采集图像时需要在暗室环境。

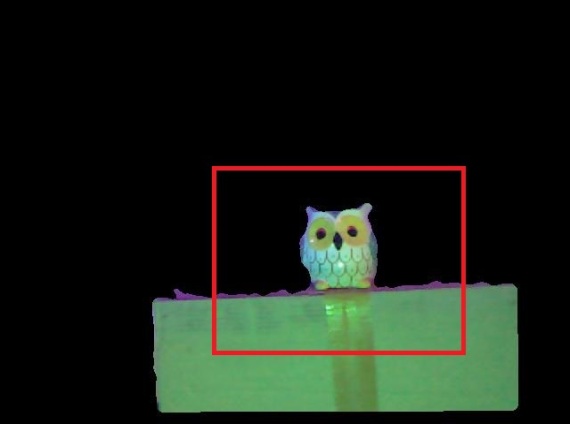
由于Kinect的深度摄像头与彩色摄像头之间存在距离，因此获得的两幅图像是存在误差的，所以在对物体进行三维重建时，首先要对Kinect获取的深度图像和彩色图像进行校准。在这里，Kinect提供的官方SDK和OPenNI中都提供了相应的函数。



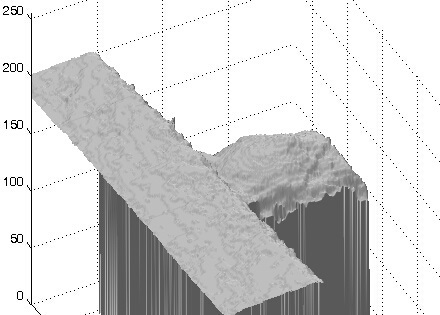
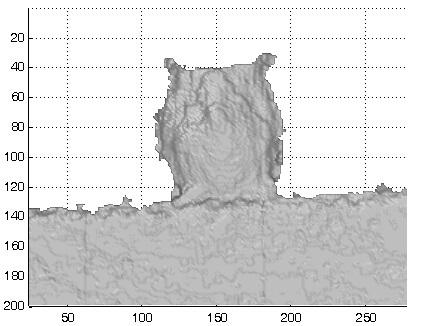
(a)场景彩色图像 (b)场景深度图像 (c)校准后的深度图像

图3-6 Kinect获取的原始图像及校准后的图像

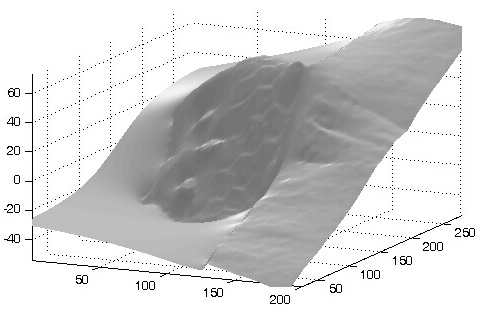
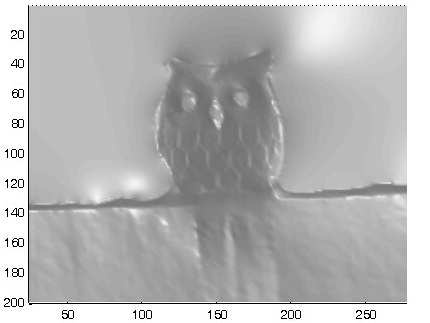
校准之后，使用本章中提到的方法对原始图像进行处理。下图展示了分别对三个不同的物体进行三维重建的结果。其中图3-7展示的是一个猫头鹰形状的花盆，图3-8展示的是一块刻有“中国海洋大学”字样的石板（刻痕深度在5mm左右），图3-9展示的是一块刻有花纹的石板（刻痕深度在1cm以下）。



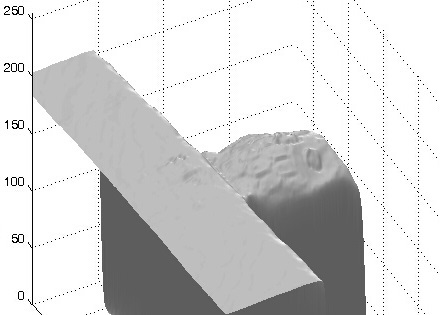
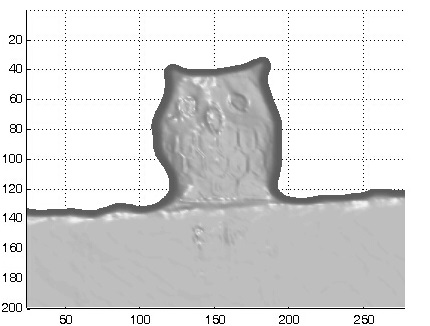
(a) Kinect采集的原始图像



(b)Kinect 获取的原始深度

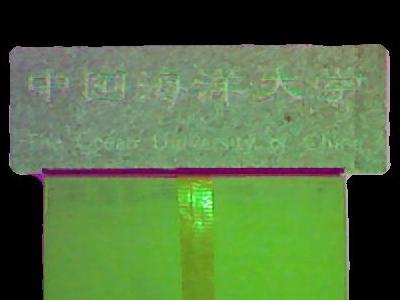
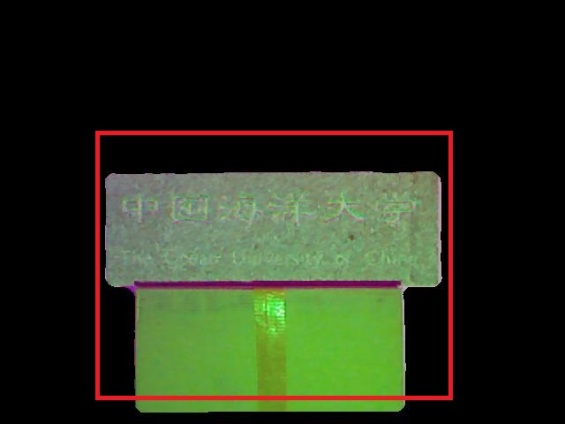


(c)光度立体的计算结果

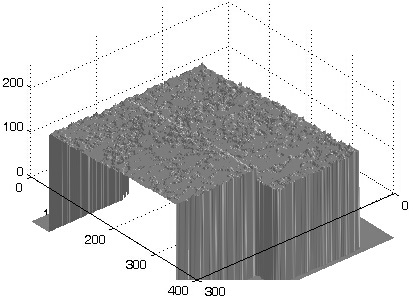
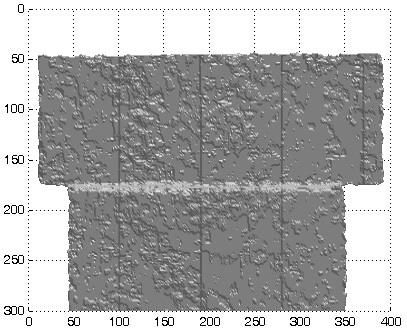


(d)本文方法重建的结果

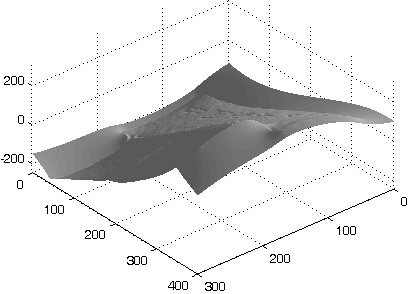
图3-7 对猫头鹰花盆的重建结果



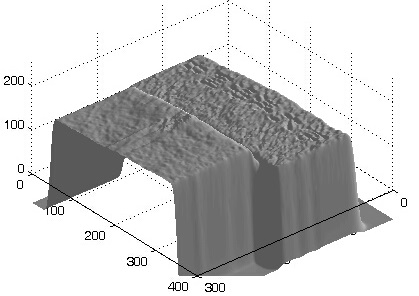
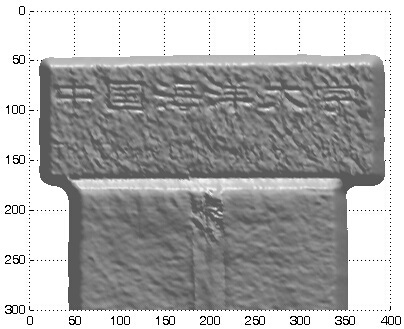
(a) Kinect采集的原始图像



(b)Kinect 获取的原始深度

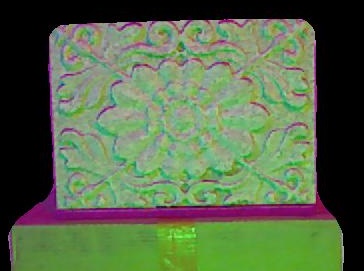
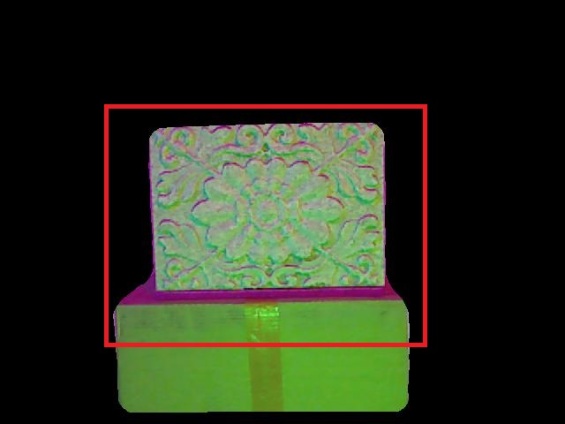


(c)光度立体的计算结果

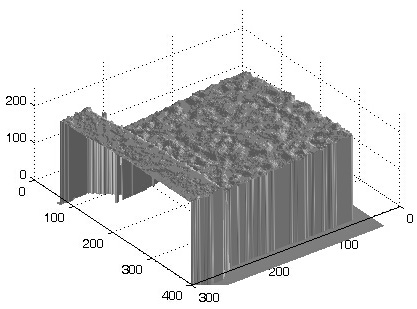
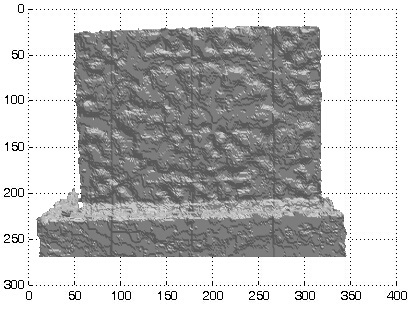


(d)本文方法重建的结果

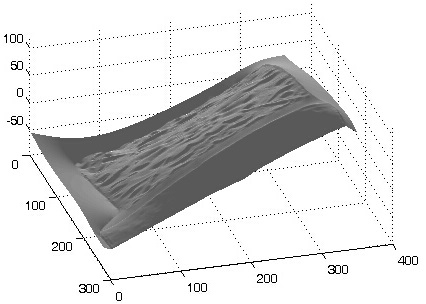
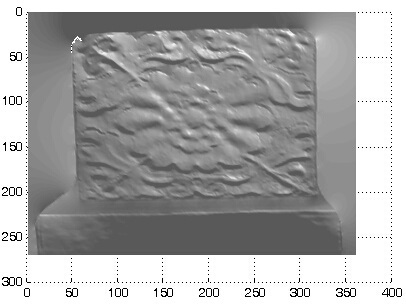
图3-8 对“中国海洋大学”字样的石板的重建的结果



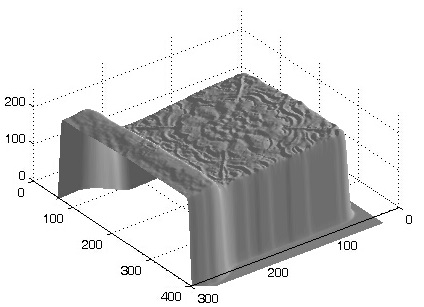
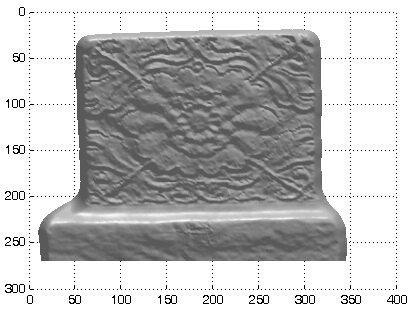
(a) Kinect采集的原始图像



(b)Kinect 获取的原始深度



(c)光度立体的计算结果



(d)本文方法重建的结果

图3-9 对刻有花纹的石板的重建结果

实验过程中，为了便于拍摄，我们将物体放在一个纸箱上。为了更好得展示实验效果，我们根据Kinect获取的深度信息去除了背景点，并仅展示了红色框线标出的区域的重建结果。

从展示的实验结果中可以看出，Kinect捕获的原始深度信息中存在较大的噪声，对于细节区域（如石板上的字和花盆微小起伏）不能准确获取甚至是获取不到。多光谱光度立体方法能够获取到物体在细节的高度变化，但是在对物体的整体形状判断却存在较大的误差。但是本章中的方法能够很好的结合两种方法的优势，从而在最终的重建结果中无论是低频还是高频都是较准确的。

## 3.5 小结

本章按照融合深度信息的多光谱光度立体三维重建方法的步骤，对各部分的背景知识进行了介绍：首先说明了Kinect的深度测距原理，并将产生误差的情况进行了分析，并在实验中尽量避免上述状况的产生。然后对多光谱光度立体的技术进行了详细的介绍，并对深度数据和位置数据的结合方法进行了讲解，并在最后展示了我们在真实场景中拍摄的图像的重建结果。从实验结果可以看出，融合深度信息的多光谱光度立体技术对原始深度数据有了很大的优化。

在下一章中，我们将把多光谱光度立体技术应用到水下物体的三维重建，并针对水下环境的特殊性，对水下光学成像模型针对我们的实验仪器进行修改，使得其能在多光谱光源的条件下获得更好的结果，并在最后展示了实验结果。

# 4 水下图像深度估计

## 4.1 水下光学成像模型

深度估计在场景理解应用中扮演重要角色，反过来，场景内物体的几何关系为我们表示物体提供了更好的表达。根据单幅图像图像估计深度在计算机视觉领域中有诸多重要应用，深度估计已经成为多个计算机视觉问题的瓶颈问题，例如：语义标注、姿态估计、三维建模等。尽管像微软Kinect等已经将深度获取设备变为一种普及的产品，但目前视觉领域主要使用的数据仍然是图像；同时，户外场景的应用中仍主要使用LiDAR或激光雷达等传感器，因为户外的光线环境更复杂，例如室外过强的光线会让获取的深度信息夹杂很多噪声。水下图像深度估计主要用于水下三维建模，不同于空气中的深度估计，水下图像面临更多挑战，例如设备防水，光线等。相对于用直接用深度传感器获取深度信息，从图像中估计深度在水下应用中更为方便。

对人来说，从单幅图像中推断隐含的3D结构容易，因为人的视觉系统可以利用先验知识，纹理，双目视差等信息。对计算机来说，由于没有足够多的线索可以利用，深度估计仍然是一个十分有挑战的任务。之前的工作主要将几何假设加入到模型中，通过引入更多的约束来求解问题，比如盒式模型。然而这类方法的适用场景有限，只能应用于特定场景结构的问题中。

### 4.1.1 水下图像深度估计模型

与之前的方法不同，我们将深度估计问题转化为典型的回归问题，不依赖于任何几何先验知识或额外信息。

深度估计的目标是推断图像中每个像素点的深度信息，与先前方法类似，我们假设一副图像是由一系列同质的小区域（超像素）组成，并且以超像素作为节点构建图模型。每个超像素由区域中心点近似。表示彩色图像，表示n个超像素点的深度值。与传统CRF类似，我们用条件概率来对像素深度值的概率分布进行建模：

(4-1)

为能量函数，为

### 4.1.2 水下光学成像模型

我们假设光照条件和拍摄环境符合图4-1：方向为的远光源照亮水下物体表面的点P（其表面法向为，表面反照率为）。这样，相机接收的能量可以分为照射到物体表面并反射回相机的能量（实线）和介质散射到相机中的能量（虚线）。其中：

(4-3)

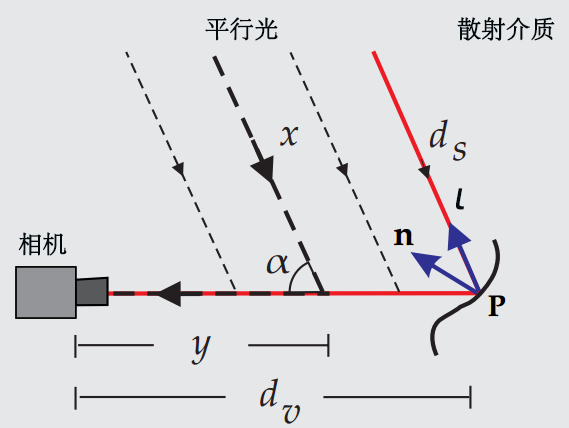


图4-1 水下光学成像模型

经过介质单一散射的光能量可以由视角方向的亮度积分获得（由公式(4-2)）：

|  |  |
| --- | --- |
|  | (4-4) |

其中、、和都是与无关的。此外，我们假设光源均匀的照亮视距，即：假设是一个等于常量，也是独立于的。这就使我们能将公式（4-4）写为：

(4-5)

相机接收到的能量就可以写为和的和：

(4-6)

在我们的实验中，将Kinect放置在水箱上方，并保持相机与水面平行，这种条件下，光线在空气和水中的折射是必须被考虑到的。图片4-2展示了实验的原理图：水下某物体表面的一点P的表面法向为，表面反照率为；入射光线折射之后的方向为；入射光线到达水下物体表面之前的传输距离为，水下物体表面反射出的光线到达相机传感器之前的传播距离为，它们之间的夹角为。由图片可知，和之间的关系可以用三角法来表示：

(4-7)

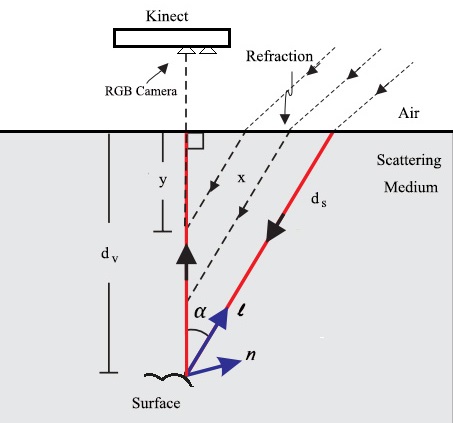


图4-2 实验原理图

值得注意的是，在这种条件下，散射到相机中的环境光在到达光轴之前在介质中的传播距离是不一样的（通过比较图4-1和图4-2中虚线的长度可以知道）。因此，公式(4-5)中将和假设为的常数在此条件下便不再成立。所以，我们利用公式(4-7)来修改：

|  |  |
| --- | --- |
|  |  |
|  | (4-8) |
|  |  |

这样，相机捕获的能量为：

(4-9)

## 4.2 水下多光谱光度立体

### 4.2.1 水下多光谱光度立体模型

由3.2.2节中对空气中多光谱光度立体的描述，我们可以将水下光学成像模型写成多光谱光度立体下的形式。

我们令表示彩色图像中某一点获取的能量， 表示该点在三个通道的光照方向，为第个光源的散射系数，为第个光源与相机方向的夹角，为第个光源折射之后的光照方向，,

则：

|  |  |
| --- | --- |
|  | (4-10) |
|  |

### 4.2.2 水下多光谱光度立体求解

文献[18]中使用测光球来对光源的参数进行求解，通过球上最亮点的位置可以获得折射后的光照方向和光源发出的能量。在我们的实验过程中发现，由于Kinect相机的分辨率太低，无法从测光球中获取任何信息（如图4-3）。因此，为了对公式进行求解，我们需要对成像模型进行简化。

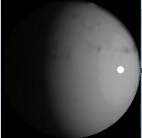


图4-3，左图为文献[18]中的测光球图像，右图为Kinect采集到的图像

对于图像中的某一通道，由于光源的方向和频谱固定，所以散射系数、光照强度以及光照角度都是固定的，在均匀的介质环境中，我们可以将相位函数近似为 [18]。这样，对于环境光的散射能量只与相机到物体表面点的距离有关。由于设备本身存在的限制，我们的方法在水下三维重建中，主要应用与表面起伏不是特别剧烈的环境或物体，因此的变化并不大，这样对于散射能量的变化影响也不大。由此分析，我们可以将散射能量当作一个常数项来处理。

对于物体表面反射到相机的能量，我们将介质对光线的衰减合并到光源频谱分别和相机对光照敏感性的混合矩阵中，即令这样，公式(4-9)可以简化为我们熟悉的第三章中的形式：

(4-11)

通过对公式(4-11)分析可知，由于项被简化为常数，但是对于每个通道的值不同，这样，利用Kinect获取的场景初始深度值求出的法向，我们需要选取四组对来对公式进行求解。

我们从表面选取四个点 ，，，，它们的表面反照率未知，但是都是相等的。它们的法向是，，，（非共面），相应的捕获的RGB图像亮度是，，，。则：

|  |  |
| --- | --- |
|  | (4-12) |

由于在Kinect获得的粗略的几何模型中包含了足够多的准确点或者有效值，这样在所有的点中以四个点为一组进行随机反复取样，将有很高的可能性在取的所有点中，有一组数据的四个点全部为有效值。同时根据第三章中的分析可以知道，无效数据不会对任何特定的光照模型产生一致的影响，然而有效值对模型的影响是一致的。这样可以使用随机抽样一致性算法来分离有效值和无效值，并且用于计算未知参数。对水下环境的光照模型，可以通过下面的步骤进行求解：

1）在图像中随机抽取四个点，用它们的RGB亮度和法向为估计一个值；

2）将待重建物体表面的每一个点 都使用公式（4-13）来验证这个假设预测的值与Kinect获得的图片亮度值的差是否在给定的阈值范围内：

|  |  |
| --- | --- |
|  | (4-13) |

其中 中包含了量化误差，图像的噪声等。

3）重复步骤1）和2）一定的次数，直到得到的预测值始终保持最大数目的支持。

4）当得到一组之后，用式（4-14）来删除表面已满足该式关系的点。

|  |  |
| --- | --- |
|  | (4-14) |

然后继续从步骤1）开始执行，对图像中剩余的点的寻找矩阵，直到图像中所有的点被移除，或者满足停止循环的条件。

### 4.2.3 水下多光谱光度立体三维重建

对于水下多光谱光度立体场景的重建方法与空气中的方法类似：首先为了保证图像分割效果，对图像使用mean-shift算法进行初步的聚类，然后根据计算得到的的个数，使用K-means算法对图像进行分割，并对分割得到的各个区域进行编号。

得到图像的分割结果之后，就可以对物体的法向进行求解了。为了给每个区域选取最适合的，首先对深度相机输出的深度信息进行适当尺度的平滑操作，以降低高频噪声的影响。然后使用平滑之后的深度数据计算法向，同时对彩色图像进行与深度图像相同尺度的平滑，得到平滑之后的图像亮度值。我们将分割之后的图像的每一部分都与进行下面的计算：

|  |  |
| --- | --- |
|  | (4-15) |

记录在每个颜色区域产生的误差均值，完成全部运算后，在该颜色区域拥有最小误差的将用于计算该区域的法向。这样我们就使用多光谱光度立体方法获得了物体的法向数据。

## 4.3 深度信息与光度立体的融合

在水下环境中使用的深度信息与多光谱光度立体的融合方法完全相同：使用位置信息改善法向，并用法向信息来修正位置数据。

需要注意的地方是，由3.1.1小节中的介绍可以知道，Kinect在获取深度信息时使用的是红外散斑，而红外线在水中的衰减比可见光要剧烈的多，因此在浑水中采集到的深度数据中很容易出现Kinect采集不到深度的点，即会有空洞产生。

现在已经有很多算法在对Kinect的深度数据填洞做研究，但是在图4-6的情况下（有一块石头的深度数据完全没有被采集到），无论是基于彩色图像的填洞算法，还是使用深度数据本身的方法，都无法对这种整片缺失的情况进行修复。考虑到使用填洞算法可能会增加时间成本，本文中仅使用形态学变换对空洞进行填充——从直径为5的圆形模板开始，对深度图像进行膨胀，然后用相同的模板进行腐蚀，这样，可以对深度图像中的小尺寸空洞填充，然后逐步增大模板的尺寸，直至所有的空洞都被填充。为了使深度数据更加准确，Kinect获取的深度数据不做改变，仅将空洞区域的值改为填洞后的深度。

实验证明，在空洞范围不大或很稀疏时，较差的填洞方法不会对结果产生很大的影响（如图4-5），但当有大范围的空洞时（如图4-6），简单的填洞会在低频引入较大误差，从而影响整体的重建效果。

## 4.4 实验结果

我们使用一个尺寸为的玻璃水箱作为实验环境，将Kinect水平放置于水箱上部，并将三个彩色光源放置在Kinect周围，使其从不同方向照射水下的物体。与第三章相同，我们同时打开三盏灯，对水下物体的彩色数据和深度信息进行拍摄，并根据本章提出的光学成像模型对数据进行处理。

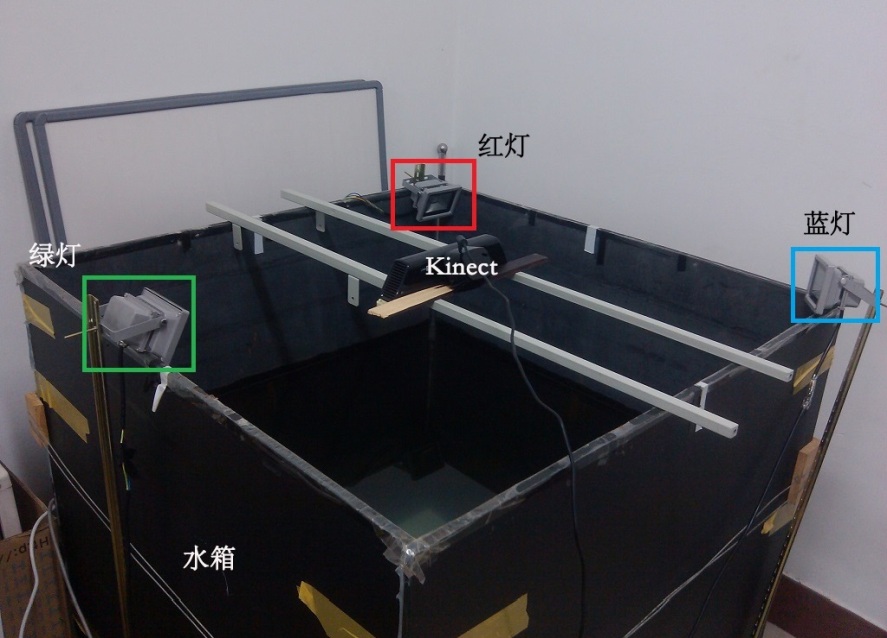
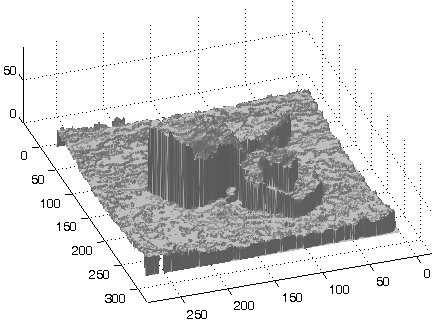
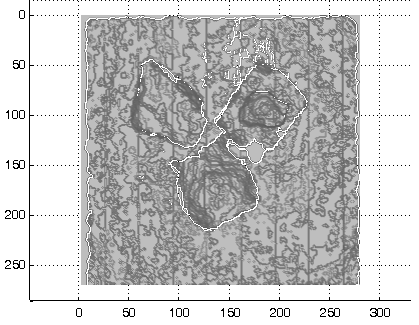


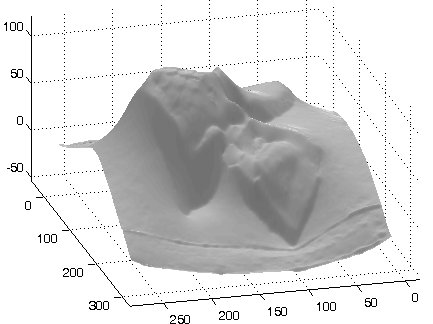
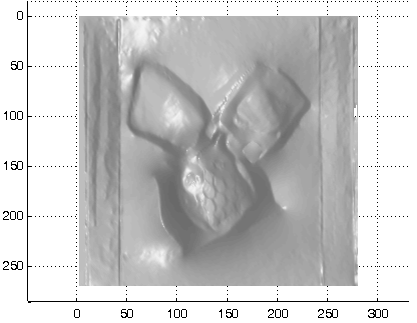
图4-4 实验环境

图4-5、4-6和4-7中展示了相同外部条件下对不同浑浊度的水中的物体进行三维重建的效果。其中，图4-5展示了我们在清水中的实验结果。之后，我们在水中加入了牛奶，并搅拌均匀，使水变得浑浊，然后又在两种不同的浑浊条件下，对水下的物体进行了三维重建（如图4-6、4-7）。

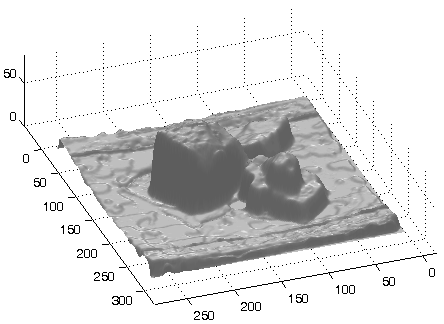
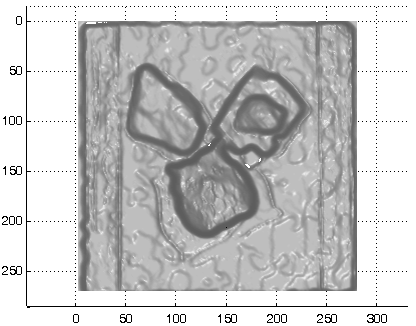
(a) Kinect采集的原始图像和局部放大的图像



(b)Kinect 获取的原始深度

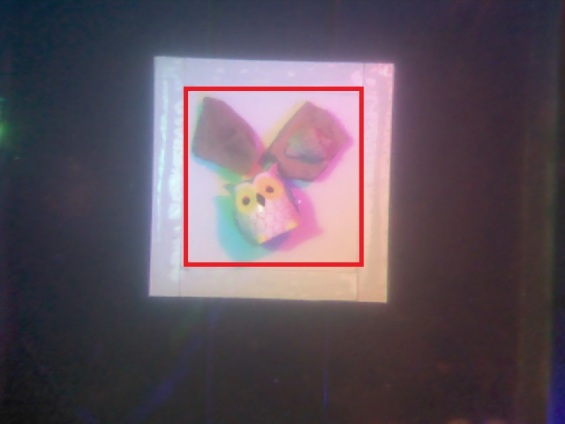


(c)光度立体的计算结果

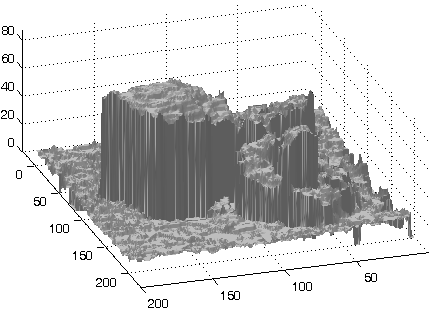
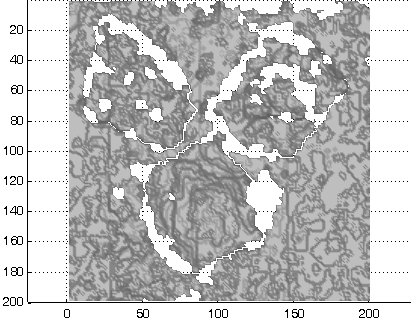


(d)本文方法重建的结果

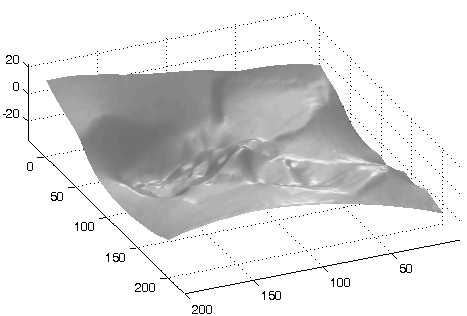
图4-5 清水中的实验结果



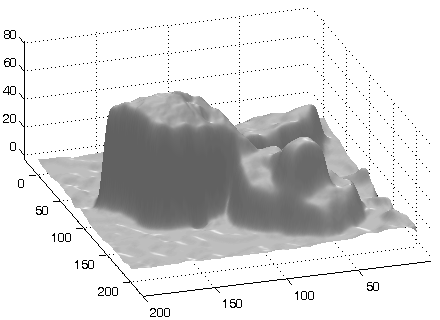
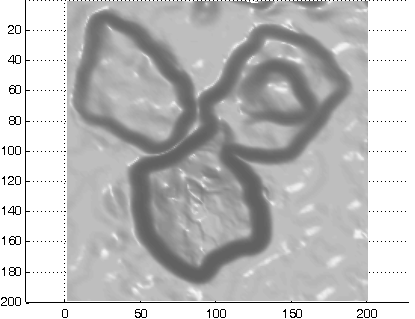
(a) Kinect采集的原始图像和局部放大的图像



(b)Kinect 获取的原始深度

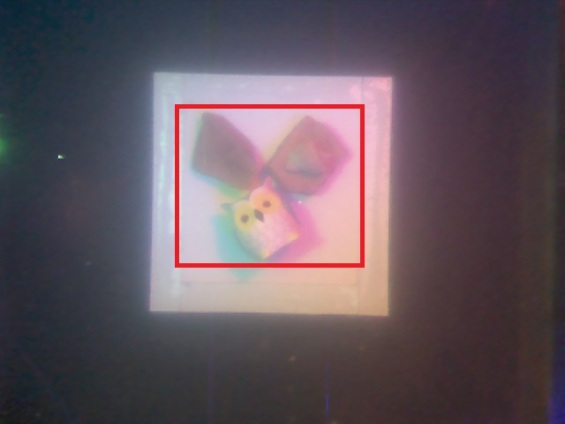


(c)光度立体的计算结果

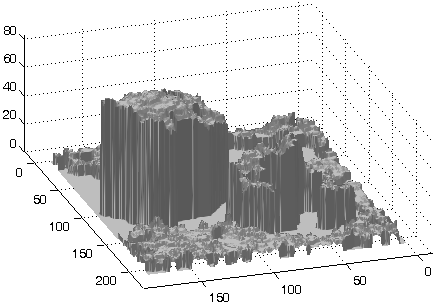
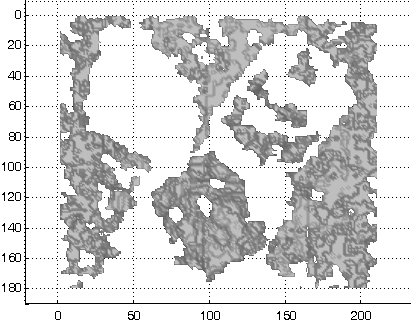


(d)本文方法重建的结果

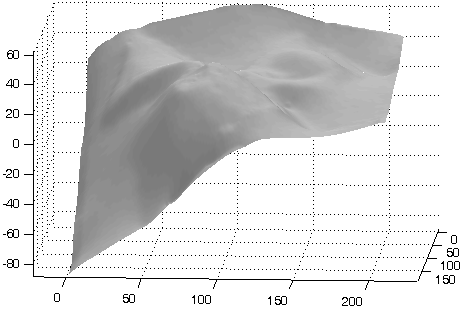
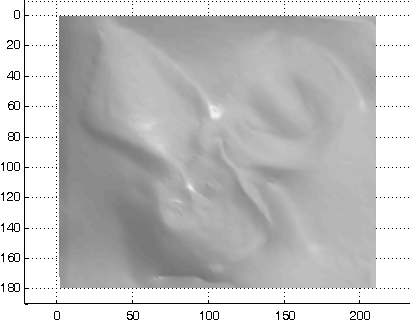
图4-6 对浑浊程度低的水中物体的重建结果



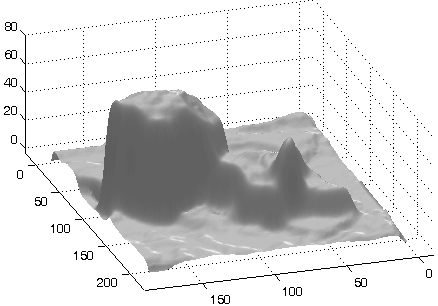
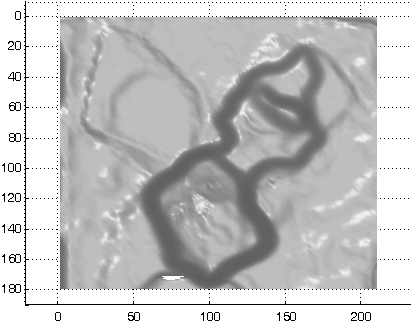
(a) Kinect采集的原始图像和局部放大的图像



(b)Kinect 获取的原始深度



(c)光度立体的计算结果



(d)本文方法重建的结果

图4-7对浑浊程度较高的水中物体的重建结果

从展示的三组重建结果中可以看出，在清水中，Kinect能够获取所有物体的深度信息，但是存在很大的噪声；使用光度立体技术能够获取物体的细节信息，但是在整体轮廓的计算中却存在较大误差；而本文中使用的方法能够在低频和高频区域均有较好的表现。在浑水条件下，Kinect能够获取的深度信息随着浑浊程度的增加而急剧减少，在图4-7(b)中可以看到，以经有大片区域无法获取深度数据，而本文的方法仍能够在全局获得计算结果，只是在原始深度大片缺失的区域会产生低频误差，而在高频区域能够保持准确。

## 4.5 小结

本章中，将第三章提到的基于深度的多光谱光度立体技术应用到了水下物体的三维重建，并建立了适用于多光谱光源条件的水下成像模型。从实验效果来看，本文提出的方法在Kinect的工作范围内，能够得到令人满意的重建效果。

# 5 总结与展望

随着计算机计算能力和硬件设备的飞速发展，三维重建技术作为计算机视觉领域的一个重要分支，它的发展越来越受到人们的重视。三维重建技术在计算机动画、人机交互、现代医学等多个领域有着广泛的应用，近年来，随着人们对海洋的探索，水下三维重建的研究也日益增多。

本文首先介绍了当前一些常见的光学三维重建方法，在对这些方法进行了详细的分析和比对之后，发现光度立体技术能够在较简易的实验环境下对物体进行三维重建，同时其计算结果在高频的细节区域具有较高的精确度；而深度相机的出现让实时获取场景的深度数据成为了可能。于是我们选用Kinect与多光谱光度立体技术结合，在设备简易、操作简单、计算量较小的情况下对物体进行三维重建。本文中对多光谱光度立体三维重建涉及的基本知识进行了详细的介绍，包括Kinect的测距原理、朗伯模型的基本概念、多光谱光度立体的求解方法以及深度信息与多光谱光度立体技术结合的方法等，并在空气中实现了融合深度信息的多光谱光度立体三维重建技术，在第三章中展示了实验结果。在对空气中的多光谱光度立体技术进行了全面的了解之后，我们将该技术引入水下，并针对水对光能量的吸收与散射特性，将水下成像模型进行了修改，并通过实验证明，本文提出的方法能够对水下的物体进行简单、快速的三维重建。

使用多光谱光度立体技术对物体进行三维重建的过程中，不需变换光照条件，也无需对光源的方向进行特殊设置，而且数据采集过程简单，设备简易，为水下手持操作提供了可能，因此具有较大的应用前景。

本文还存在一些不足和可以进行进一步的研究的地方，主要可以总结为以下几点：

1）本文假设物体都符合朗伯反射模型，这种模型过于理想，尽管对有镜面反射的物体进行重建实验时也能获得不错的结果，但是前提假设中存在的固有问题会在计算过程中引入误差，从而对重建的精度造成影响。

2）本文对水下吸收和散射影响的考虑中，假设光线在水中仅进行了单次散射，这也是对水下环境的一种理想化估计，对水下获得的图像信息的恢复会产生一定的影响；

3）Kinect获得的图像中存在深度获取不到的点，也就是深度图像中出现的空洞现象，本文使用了最简单的填洞算法，这会对空洞处的低频信息引入较大误差，可以在本文的研究基础上，尝试使用对Kinect输出的深度数据的填洞算法，从而获得更精确的结果；

4）由于红外光线在水中的衰减是以指数形式存在的，Kinect在水下的工作距离受到了较大的影响，但本文提出的方法具有较大的可移植性，可以在本文的研究基础上，使用其他方法（如立体视觉）获取物体的初始深度信息，进而使用本文提出的方法对水下物体进行三维重建，从而使本文的方法能够在更大的深度范围内对物体进行三维重建。

# 参考文献

1. Huhle B, Fleck S, Schilling A. Integrating 3D time-of-flight camera data and high resolution images for 3DTV applications[C]//3DTV Conference, 2007. IEEE, 2007: 1-4.
2. Reulke R. Combination of distance data with high resolution images[C]//ISPRS Commission V Symposium Image Engineering and Vision Metrology. 2006, 2.
3. Kuhnert K D, Stommel M. Fusion of stereo-camera and pmd-camera data for real-time suited precise 3d environment reconstruction[C]//Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on. IEEE, 2006: 4780-4785.
4. Hahne U, Alexa M. Combining time-of-flight depth and stereo images without accurate extrinsic calibration[J]. International Journal of Intelligent Systems Technologies and Applications, 2008, 5(3): 325-333.
5. Kim Y M, Theobalt C, Diebel J, et al. Multi-view image and tof sensor fusion for dense 3d reconstruction[C]//Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on. IEEE, 2009: 1542-1549.
6. Zhu J, Wang L, Yang R, et al. Fusion of time-of-flight depth and stereo for high accuracy depth maps[C]//Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008: 1-8.
7. Zhang Q, Ye M, Yang R, et al. Edge-preserving photometric stereo via depth fusion[C]//Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012: 2472-2479.
8. Yu L F, Yeung S K, Tai Y W, et al. Shading-based shape refinement of RGB-D images[C]// Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013: 1415-1422.
9. Han Y, Lee J Y, Kweon I S. High quality shape from a single rgb-d image under uncalibrated natural illumination[C]//Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE, 2013: 1617-1624.
10. Haque S, Chatterjee A, Govindu V M. High Quality Photometric Reconstruction Using a Depth Camera[C]//Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014: 2283-2290.
11. NH Vilarnau. Integration of optical and acoustic sensor data for 3D underwater scene reconstruction.2009.
12. Sedlazeck A, Koser K, Koch R. 3d reconstruction based on underwater video from rov kiel 6000 considering underwater imaging conditions[C]//OCEANS 2009-EUROPE. IEEE, 2009: 1-10.
13. Hogue A, German A, Jenkin M. Underwater environment reconstruction using stereo and inertial data[C]//Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on. IEEE, 2007: 2372-2377
14. Törnblom N. Underwater 3D Surface Scanning using Structured Light[J]. 2010.
15. http://www.2grobotics.com/
16. Brandou V, Allais A G, Perrier M, et al. 3D reconstruction of natural underwater scenes using the stereovision system iris[C]//OCEANS 2007-Europe. Ieee, 2007: 1-6.
17. Bruno F, Bianco G, Muzzupappa M, et al. Experimentation of structured light and stereo vision for underwater 3D reconstruction[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2011, 66(4): 508-518.
18. Narasimhan S G, Nayar S K, Sun B, et al. Structured light in scattering media[C]//Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. IEEE, 2005, 1: 420-427.
19. Barnard S T, Fischler M A. Computational stereo[J]. ACM Computing Surveys (CSUR), 1982, 14(4): 553-572.
20. Vaillant R, Faugeras O D. Using extremal boundaries for 3-D object modeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 157-173.
21. Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge university press, 2003.
22. Tomasi C, Kanade T. Shape and motion from image streams under orthography: a factorization method[J]. International Journal of Computer Vision, 1992, 9(2): 137-154.
23. Prados E, Faugeras O, Rouy E. Shape from shading and viscosity solutions[M]//Computer Vision—ECCV 2002. Springer Berlin Heidelberg, 2002: 790-804.
24. Gallo A, Bruno F, Muzzupappa M, et al. Multi-view 3D reconstruction of small stone samples deteriorated by Marine organisms[C]//Virtual Systems and Multimedia (VSMM), 2012 18th International Conference on. IEEE, 2012: 181-187.
25. Woodham R J. Photometric method for determining surface orientation from multiple images[J]. Optical engineering, 1980, 19(1): 191139-191139-.
26. Basri R, Jacobs D, Kemelmacher I. Photometric stereo with general, unknown lighting[J]. International Journal of Computer Vision, 2007, 72(3): 239-257.
27. Hertzmann A, Seitz S M. Example-based photometric stereo: Shape reconstruction with general, varying brdfs[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2005, 27(8): 1254-1264.

http://ssco.gsfc.nasa.gov/workshop\_2010/day3/Roger\_Stettner/Stettner\_ASC\_Workshop\_Presentation.pdf

http://commons.wikimedia.org/wiki/File:TOF\_Kamera.jpg

http://commons.wikimedia.org/wiki/File:PMDCamCube.jpg

http://commons.wikimedia.org/wiki/File:FOTONIC-B70.jpg

Aue J, Langer D, Muller-Bessler B, et al. Efficient segmentation of 3d lidar point clouds handling partial occlusion[C]//Intelligent Vehicles Symposium (IV), 2011 IEEE. IEEE, 2011: 423-428.

胡曙光, 陈静, 华艳秋等. 基于激光三角法的传感器设计[J]. 测控技术, 2006, 25(6):7-8. DOI:10.3969/j.issn.1000-8829.2006.06.002.

杨再华, 李玉和, 李庆祥等. 一种基于光学三角法的形貌测量系统[J]. 光学技术, 2005, 31:622-623. DOI:10.3321/j.issn:1002-1582.2005.04.012.

Levoy M, Pulli K, Curless B, et al. The digital Michelangelo project: 3D scanning of large statues[C]//Proceedings of the 27th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co., 2000: 131-144.

Lanman D, Taubin G. Build your own 3D scanner: 3D photography for beginners[C]//ACM SIGGRAPH 2009 Courses. ACM, 2009: 8.

Zhang L, Curless B, Seitz S M. Rapid shape acquisition using color structured light and multi-pass dynamic programming[C]//3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on. IEEE, 2002: 24-36.

Nehab D, Rusinkiewicz S, Davis J, et al. Efficiently combining positions and normals for precise 3D geometry[J]. ACM transactions on graphics (TOG), 2005, 24(3): 536-543.

Lambert J H. Photometria[M]. 1760.

A. Petrov. Light, color and shape. Cognitive Processes and their Simulation (in Russian), 1987: 350–358.

Vogiatzis G, Hernández C. Self-calibrated, multi-spectral photometric stereo for 3D face capture[J]. International journal of computer vision, 2012, 97(1): 91-103.

Hernández C, Vogiatzis G, Cipolla R. Multiview photometric stereo[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2008, 30(3): 548-554.

Hernández C, Vogiatzis G. Self-calibrating a realtime monocular 3d facial capture system[C]//Proceedings international symposium on 3D data processing, visualization and transmission (3DPVT). 2010.

Anderson R, Stenger B, Cipolla R. Augmenting Depth Camera Output Using Photometric Stereo[C]//MVA. 2011: 369-372.

Fukunaga K, Hostetler L. The estimation of the gradient of a density function, with applications in pattern recognition[J]. Information Theory, IEEE Transactions on, 1975, 21(1): 32-40.

MacQueen J. Some methods for classification and analysis of multivariate observations[C]// Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967, 1(14): 281-297.

Rushmeier H, Bernardini F. Computing consistent normals and colors from photometric data[C]//3D Digital Imaging and Modeling, International Conference on. IEEE Computer Society, 1999: 0099-0099.

Paige C C, Saunders M A. LSQR: An algorithm for sparse linear equations and sparse least squares[J]. ACM Transactions on Mathematical Software (TOMS), 1982, 8(1): 43-71.

Chandrasekhar S. Radiative transfer[M]. Courier Corporation, 2013.

# 致谢

时光飞逝，转眼间两年的硕士研究生生活即将结束，在此谨向在攻读研究生期间对我的学习、生活给予帮助和关心的所有老师和同学表示深深的谢意。

在此，我要衷心的感谢我的导师董军宇教授，董老师在我的研究生求学期间，在生活、学习以及工作等方面都给予了很大的帮助、支持和鼓励。还要特别感谢亓琳老师，从毕业论文选题开始，我在背景知识学习、代码编写及论文修改各个阶段都遇到了很多问题，是他的悉心指导和严格要求，让我顺利完成了毕业设计。两位老师都是我两年求学过程中的良师益友。

同时也要感谢实验室已毕业的同学前期所做的工作对本文研究的启发，是他们留下的资料让我以最快的速度了解了研究相关的背景知识；感谢吴则举博士制作了实验器材，让我的理论研究能够在实验中验证结果。

感谢实验室各位伙伴，是他们的热情帮助，让我能够顺利的进行科研，他们在生活上对我的关心和爱护，让我的研究生求学期间充满了欢乐。

最后要感谢的是我的父母，感谢他们多年来对我的养育和培养。

# 个人简历

1991年5月6日出生于山东省青岛胶南市。

2009年9月考入中国海洋大学信息科学与工程学院计算机科学与技术专业，2013年6月本科毕业并获得工学学士学位。

2013年9月保送入中国海洋大学信息科学与工程学院计算机技术专业攻读硕士学位至今。