

# 基于光学图像的 三维重建 理论与技术

孙玉娟 著



清华大学出版社



基于光学图像的  
三维重建  
理论与技术

孙玉娟 著

清华大学出版社  
北京

## 内 容 简 介

本书详细论述了基于光学图像进行三维重建的基本理论和方法,从摄像机模型、相机校准、常用的光学模型等基础理论出发,介绍了基于明暗形状恢复、光度立体、结构光三维测量、双目立体视觉等技术的基本原理和实现方法。从每种技术的理论基础、系统组成、基本约束、重建精度、应用场合等方面分析了其理论依据和具体算法,为基于光学图像的三维重建的实际应用提供了基本的理论参考。

全书共分8章:第1章为概述部分;从第2章到第4章分别论述了本书所涉及的数学基础、基本的摄像机校准方法和光照模型;从第5章到第8章着重介绍了目前基于光学图像进行三维重建的主要方法。在介绍三维重建理论的同时,书中提供了大量应用实例,可对理论部分进行直观的解释和验证。

本书可供三维重建、图像处理、计算机视觉等领域的工程技术人员参考,也可作为高等学校电子信息类、计算机类专业高年级本科生或研究生的学习教材。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

基于光学图像的三维重建理论与技术/孙玉娟著. —北京:清华大学出版社,2018  
ISBN 978-7-302-49000-5

I. ①基… II. ①孙… III. ①光学—影像图—图像数字化处理 IV. ①O43-39

中国版本图书馆CIP数据核字(2017)第294403号

责任编辑:白立军 李 晔

封面设计:傅瑞学

责任校对:梁 毅

责任印制:王静怡

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市金元印装有限公司

经 销:全国新华书店

开 本:185mm×230mm 印 张:11.75

版 次:2017年11月第1版 字 数:204千字

印 数:1~800 印 次:2017年11月第1次印刷

定 价:59.00元

产品编号:077227-01

# 前 言

基于光学图像的三维重建是计算机视觉领域重要的研究方向之一，经过了几十年的不断探索，研究人员提出了很多经典有效的算法，如：明暗恢复法、结构光三角法、多目立体视觉、近景工业测量、光度立体技术、双目立体技术等。基于拍摄到的光学图像对目标物体进行三维重建，可将物体或者场景的三维形状准确地恢复出来，去除由于环境变化或者视角偏差引起的对物体外观的理解错误，对于煤炭、钻井、勘探、考古等领域具有重要的应用前景。目前，虽有部分图像处理或者机器视觉方面的书籍涉及其中的某种技术，但并没有系统全面地介绍基于光学图像的三维重建技术的参考书。本书基于光学图像详细介绍三维重建领域的基本理论与典型算法，期望为从事该领域相关工作的人员提供一些有效的参考。

本书的内容，除第1章概述外，分为两部分：第一部分（第2章至第4章）涉及三维重建需要的数学基础，包括平面几何、射影几何、射影变换的基本概念和基本定理，摄像机的三维射影几何与校准等，可为后续基于图像的三维重建技术的理解提供理论参考。第二部分（第5章至第8章）介绍目前主流的基于图像的三维重建技术，每种技术独占一章，从基本的光学模型理论入手，对基于图像的三维重建进行比较深入的阐述和分析，对每一种技术采用“理论简介、算法分析、实验仿真”的方式进行论述，便于读者从理论到实践对现有的基于光学图像的三维重建技术进行理解和掌握。

在编写过程中，我们参考和借鉴了许多专家学者的研究成果，在书后均一一列出，同时也参考和选用了许多博客和组织机构网站中的内容和图片，由于数量太多未能全部列出，在此向这些成果的所有者和组织机构表示诚挚的谢意。

本书的研究工作得到国家自然科学基金(61602229、61771231)、省自然科学基金(ZR2016FM13、ZR2017MF010)和鲁东大学博士基金(LY2015032)的资助，为本书的顺利出版提供了资金支持。

感谢中国海洋大学董军宇教授、青岛理工大学吴则举副教授、鲁东大学邹海林教授给予作者的热心帮助。也感谢和作者朝夕相处的许多同事、朋友和学生的鼎力相助,多年来他们对本书内容提出的建议和意见使作者受益匪浅,在此向我的同事、朋友和学生表示由衷的感谢。

全书的策划和编写工作由孙玉娟负责。鲁东大学邹海林教授、中国海洋大学董军宇教授对书稿进行了认真审阅,并提出了许多宝贵的建议。殷波老师,国圳宇、孙晓婷、周子淇等同学在文字录入、书稿校对等方面付出了辛勤的劳动,在此向他们一并表示感谢。

特别感谢清华大学出版社,感谢责任编辑及其他参与此书编辑工作的各位老师为本书顺利出版而付出的辛勤劳动。

在计算机视觉领域,基于图像的三维重建发展非常迅速,本书虽力求全面并紧跟其发展趋势,但由于作者水平和时间有限,书中难免出现疏漏之处,恳请作者不吝批评指正。

孙玉娟

2017年7月于烟台

# 目 录

第 1 章 概述 .....	1
1.1 人眼的空间知觉 .....	1
1.2 三维重建技术的分类 .....	3
1.3 光学三维重建技术 .....	6
1.3.1 被动光学重建技术 .....	7
1.3.2 主动光学重建技术 .....	8
1.4 本章小结 .....	12
第 2 章 射影几何与变换基础 .....	13
2.1 2D 射影几何 .....	13
2.1.1 点与直线 .....	13
2.1.2 理想点与无穷远点 .....	15
2.2 射影变换 .....	17
2.3 射影变换的层次 .....	20
2.3.1 射影变换的层次简介 .....	20
2.3.2 射影变换的分解 .....	24
2.4 从图像恢复仿射和度量性质 .....	24
2.4.1 无穷远线 .....	25
2.4.2 由图像恢复仿射性质 .....	25
2.4.3 虚圆点及其对偶 .....	27
2.5 不动点和直线 .....	28
2.6 3D 射影几何和变换 .....	29

2.6.1	点和射影变换 .....	30
2.6.2	平面表示和变换 .....	30
2.6.3	直线的表示和变换 .....	32
2.7	3D 射影变换的层次 .....	36
2.8	无穷远平面 .....	38
2.9	本章小结 .....	40
<b>第 3 章</b>	<b>摄像机标定与校准 .....</b>	<b>41</b>
3.1	相机射影几何 .....	41
3.1.1	相机射影几何简介 .....	41
3.1.2	二维射影几何 .....	42
3.1.3	三维射影几何 .....	42
3.2	相机成像理论 .....	43
3.2.1	空间坐标系之间的关系 .....	43
3.2.2	摄像机的理论成像模型 .....	48
3.2.3	实际成像模型 .....	49
3.3	相机标定理论知识 .....	50
3.3.1	张正友标定理论 .....	51
3.3.2	相机参数的初始估计 .....	51
3.3.3	拍摄图像的径向畸变 .....	53
3.3.4	优化函数迭代求精 .....	54
3.3.5	实验步骤及注意的问题 .....	54
3.4	本章小结 .....	58
<b>第 4 章</b>	<b>光照模型 .....</b>	<b>59</b>
4.1	朗伯模型 .....	59
4.2	双向反射分布模型 .....	60

4.3	Phong 模型	62
4.4	Cook-Torrance 光照模型	64
4.5	球谐光照模型	65
4.6	各种光照模型的对比	69
4.7	本章小结	74
<b>第 5 章</b>	<b>利用明暗信息恢复物体的三维形状</b>	<b>76</b>
5.1	根据明暗信息恢复形状概述	77
5.1.1	恢复形状的常用方法	77
5.1.2	SFS 方法需要解决的主要问题	79
5.1.3	根据明暗信息恢复形状的应用	80
5.2	根据明暗信息估计图像的光源参数	81
5.2.1	光源方向的估计	83
5.2.2	方位角 $\tau$ 的估计	85
5.2.3	倾角 $\delta$ 的估计	87
5.3	根据明暗信息恢复形状算法	89
5.3.1	最小化方法	89
5.3.2	演化方法	92
5.3.3	局部分析方法	93
5.3.4	线性化方法	96
5.4	根据明暗信息重建三维形状	98
5.5	本章小结	100
<b>第 6 章</b>	<b>基于单目视觉的三维重建</b>	<b>101</b>
6.1	经典的 PMS 技术	101
6.1.1	朗伯模型	102
6.1.2	基于光度立体技术估计目标物体表面高度信息	103



6.2	未校准的 PMS 技术	107
6.3	基于参照物的三维重建技术	112
6.4	基于学习的三维重建技术	117
6.5	本章小结	120
<b>第 7 章</b>	<b>基于结构光的三维重建技术</b>	<b>122</b>
7.1	结构光的成像系统	124
7.2	系统标定	127
7.3	基于结构光的三维数据测量	129
7.3.1	结构光激光扫描法测量原理	130
7.3.2	相位测量轮廓法测量原理	132
7.4	投影图像编码	136
7.4.1	常用投影图像编码策略	137
7.4.2	基于格雷码+相移的编码策略	138
7.5	图像特征提取方法	140
7.5.1	选取有效区域的方法	140
7.5.2	边界提取方法	141
7.5.3	三维点云重建	142
7.6	本章小结	144
<b>第 8 章</b>	<b>基于双目视觉的三维重建</b>	<b>145</b>
8.1	双目立体视觉的原理	145
8.2	双目立体视觉的系统组成	147
8.3	双目立体视觉的摄像机标定技术	148
8.3.1	张正友标定法简述	149
8.3.2	双目立体视觉摄像机标定过程	151
8.4	特征提取与立体匹配	152

8.4.1	立体匹配的约束条件 .....	152
8.4.2	特征提取与立体匹配 .....	154
8.5	双目立体视觉的三维重建 .....	157
8.5.1	基本的三维坐标求解方法 .....	158
8.5.2	基于最小二乘法的三维坐标求解方法 .....	159
8.5.3	基于视差测距法的三维坐标求解方法 .....	161
8.5.4	三维重建的结果 .....	162
8.6	本章小结 .....	166
参考文献 .....		167

# 第1章 概述

计算机视觉是一门研究如何使机器“看”的科学,更进一步地说,就是指用摄像机和电脑代替人眼对目标进行识别、跟踪和测量,并用计算机进一步做图形或图像处理,使之成为更适合人眼观察或传送给仪器检测的图像。本章首先对人眼的空间知觉做简要介绍,然后详细论述基于图像的三维重建技术的主流方法,及各方法的基本原理、特点和适用场合。

## 1.1 人眼的空间知觉

人类通过视觉、触觉、听觉、嗅觉等感觉器官从外界环境获取信息,其中视觉是人类认识世界和传递信息的主要途径,70%~80%的信息是通过人眼视觉获取的。当获取到这些视觉信息后,就会将有效的信息传输给人脑进行处理,进而形成有效的视觉信息,辅助人脑进行定位、识别、判断等功能。

眼睛是人类视觉重要的组成部分,是实现光学过程的物理基础。动物的眼睛在生物学上是非常复杂的,但从物体成像的角度考虑可以将人的眼睛比拟为照相机。人眼的视网膜是一个曲面(如果从成像的角度看仅相当于二维空间的一个平面),人能够从从一个2D的视网膜中感知到一个3D的视觉空间,这种能力就是所谓的空间视觉<sup>[1]</sup>。

图1.1为眼球横切面图,外界物体反射来的光线,经过眼角膜,由瞳孔进入眼球内部,再经过晶状体和玻璃体的折射作用,在视网膜上能形成清晰的物像,物像刺激了视网膜上的感光细胞,这些感光细胞产生的神经冲动,沿着视神经传到大脑皮层的视觉中枢,就形成视觉。

摄像机和眼睛的成像原理非常相似<sup>[2]</sup>,如图1.2的摄像机简化模型所示,摄像机拍摄照片时,光线从物体表面反射后通过透镜(镜头)到达摄像机,在像平面(底片)上成像。我

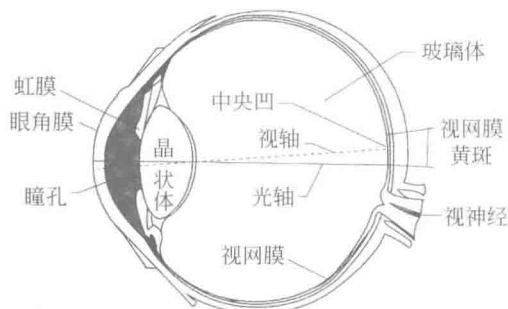


图 1.1 眼球横切面图

们可以将眼球的瞳孔抽象为相机的光圈,眼球的晶状体抽象为相机的镜头,眼球的视网膜相当于相机的底片,眼球的玻璃体相当于相机的暗箱。现在的一些照相机或者摄影机就是利用了小孔成像的原理——镜头是小孔(大多数安装凸透镜以保证光线成像距离),景物通过小孔进入暗室,物像被一些特殊的化学物质(如显影剂等)留在胶片上(数码相机、摄影机等则是把物像通过一些感光元件存储在存储卡内)。

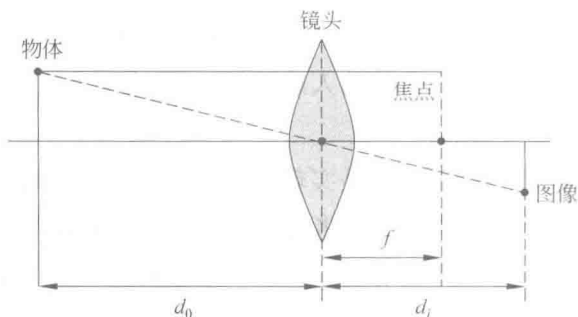


图 1.2 摄像机简化模型(小孔成像模型)

但是小孔成像模型仅为数学意义上的摄像机,即摄像机的基本模型,实际的相机或者摄像机一般都采用透镜作为镜头,因为如果只是一个小孔,CCD 采集的光线强度很弱,图像成像不够清晰,透镜可尽可能多地采集光线,让胶片上的图像更清晰。但是也带来了一些问题,如图像畸变<sup>[3]</sup>,因此对于拍摄高精度的图像或者测量精度要求较高时,还要做相机标定,以测出畸变参数。

立体感和人类的视觉系统是分不开的,人眼平时看到的任何东西都是立体的,那是因为人的两只眼睛之间有一段小小的距离,人眼看到的是两幅有细微差别的画面,这种细微差别能让大脑换算出一个个空间坐标,我们就能够区分出物体的远近,也就具有了立体感。这种立体感是灵长类的专长,只有灵长类和少数捕食动物(眼睛必须在头部前方)拥有,像牛、羊、马是不会有立体视觉的,这些动物的眼睛已经进化为观看更广角视野的广角模式,便于及早发现危险并逃生。所以独眼的人实际上丧失了立体空间感,因为单眼对距离感知困难。当然也不代表着独眼的人一点也感知不到立体,例如,通过在空间中的移动,人脑会通过运动的相对位差来补充一定的立体感;或者通过物体表面上光线的明暗也可推算出一定的立体感;以及景物远近的移动速度,大脑均可推断出一定的立体信息。

## 1.2 三维重建技术的分类

摄像机在对周围景物拍照时,将立体的空间信息投影到了摄像机的二维像平面上,获取的图像损失了空间中的深度信息。通过记录的二维图像重建空间或者物体的三维数据已经成为近几年计算机视觉领域研究的关键问题,获取物体的三维数据也已经成为很多学科对物体特性进行研究和分析的重要手段<sup>[4]</sup>。如:在医学领域,三维重建技术已经成为医学诊断、骨骼建模、器官病变检测等方面的重要手段;在工业领域,三维重建技术是逆向工程、产品质量三维检测、零部件高精度测量等研究领域的关键技术;在文物收藏和历史研究领域,三维重建技术能够提供二维图像无法展示的高精度三维数据,以辅助古董、文物、历史遗迹的三维展示与修复工作;在数字媒体方面三维重建技术可以用于动漫人物造型、三维场景虚拟;在水下或者海洋探测方面,三维重建技术可实现海底地形、地貌的三维观测,水下机器人近距离三维导航,以及海洋工程检测等。

当前,三维重建技术可以分为三大类<sup>[5,6]</sup>:接触型三维重建技术、反射型三维重建技术、透射型三维重建技术,如图1.3所示。

接触型三维测量技术最早起源于手工直接测量,手工直接测量精度低,只能测量表面结构简单、规则且利于直接测量的物体,测量结构复杂的物体时工作量较大,有可能根本无法完成。伴随着工业技术的蓬勃发展,市场上逐渐出现了半自动或者全自动三维坐标

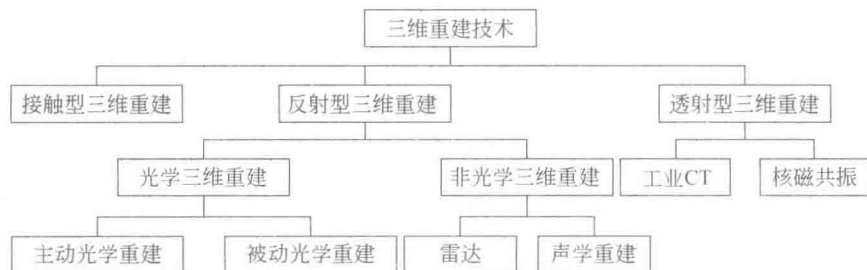


图 1.3 三维重建技术分类

测量机,如图 1.4 中海克斯康 MultiGage 关节臂三维测量机,是一个由空间六轴工业机器人臂和高精度测头组成。根据测得的三维点云数据,利用图形学相关技术进行插值,得到目标对象的三维数据。接触型三维重建测得的精度很高,但测算效率较低,测绘范围大的三维坐标测量机价格较昂贵。

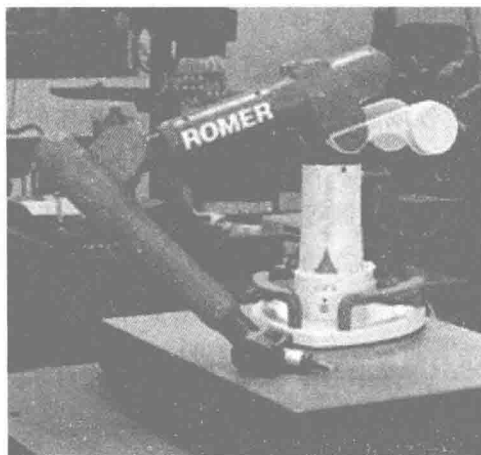


图 1.4 海克斯康关节臂三维测量机 MultiGage

反射型三维测量分为基于光学图像的三维重建和非光学图像的三维重建。非光学三维重建主要包括水下声学探测和空间雷达探测。水下声学探测通过采集反射声波的强度,组成声学探测图像,用于描述探测区地形。低频声波信号传输波长较长,绕射能力强,

受水介质散射影响小,适合工作在较差水质条件下,但其水平分辨率和垂直分辨率低。图 1.5 和图 1.6 分别显示了单波束声纳和多波束声纳的对某海域海底地形的扫描数据。图 1.7 显示了水下轮胎的声纳扫描图,仅能看出轮胎基本形状,无法观察到细节三维形状。

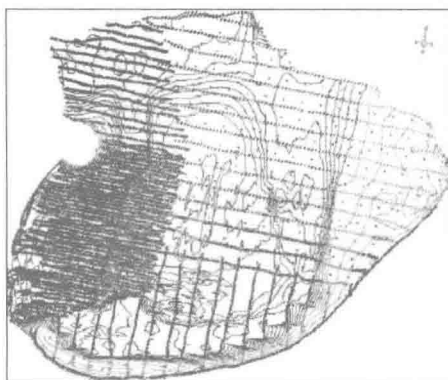


图 1.5 Souther Gulf 单波束声纳扫描数据

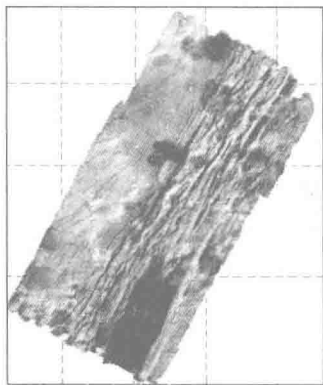


图 1.6 多波束声纳扫描数据(来自 Marine Electronics Ltd)

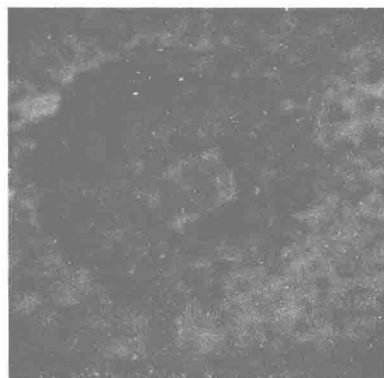


图 1.7 三维实时声纳扫描水下轮胎结果  
(来自 ECHO SCOPE)

透射型三维重建技术利用投射到物体内部的射线能量衰减程度,判断物体的纵切形状,并对纵切结果进行重构得到物体的三维信息,主要包括工业 CT 和医学中的核磁共振

技术。前者指应用于工业中的核成像技术,广泛用于汽车、材料、煤炭、石油等领域;后者是继 CT 后医学影像的又一重大进步,该技术可直接作出横断面、矢断面、冠状面和各种斜面的体层图像,不会产生 CT 检测中的伪影。

随着计算机、图像处理和机器视觉技术的高速发展,基于光学成像的视觉测量逐渐成为三维重建领域研究的热点。光学三维重建技术以光学图像为基础,通过图像处理和机器视觉技术从二维图像中提取有用信息,利用空间几何方法获取被测目标物体的三维信息。光学三维重建不仅有较高的水平分辨率,而且有较高的垂直分辨率,重建精度可达微米级。

根据原始数据采集过程中的光源控制程度,光学三维重建技术<sup>[7]</sup>分为被动光学重建技术和主动光学重建技术。如图 1.8 所示,通常主动光学重建的精度要高于被动光学重建,我们会在以后的章节中对相关技术进行详细的介绍和讨论。



图 1.8 基于光学的三维重建技术分类

### 1.3 光学三维重建技术

随着电子技术和电子产品的不断发展与更新,图像获取技术也飞速发展,高分辨率摄像机和照相机大量涌现,基于光学的三维重建技术在工业和农业等领域得到了广泛的应用与推广。根据图像或者数据采集过程中光源的控制程度,可分为被动光学重建技术和



主动光学重建技术。

### 1.3.1 被动光学重建技术

被动光学重建技术在进行三维重建的过程中,不需要直接控制任何光源,而是完全依赖于环境照明光拍摄图像,即可实现目标物体的三维重建,如立体视觉技术、基于运动的结构重建等。

#### 1. 立体视觉

立体视觉分为双目立体视觉和多目立体视觉技术。立体视觉包括两个过程:融合两个(或多个)摄像机观察到的特征和重建这些特征点的三维信息。后一个过程相对简单,这是由于场景中点的像会出现在经过成像点和相应摄像机光心的直线交点处(如图 1.9 所示)。

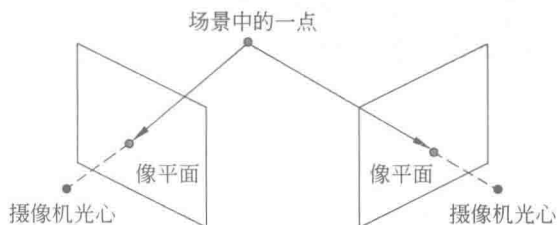


图 1.9 双目立体视觉原理图

因此,在任意时刻,当单个图像特征点被确定后,可以很容易地利用立体视觉定位空间中该点的深度信息。然而一幅图像中往往包含几十万个像素,同时又有几万个图像特征,例如边缘、梯度、纹理特征等。因此,必须采用某些方法来建立空间中同一场景点在不同摄像机中成像的对应点的匹配关系。双目或者多目立体视觉的关键就是要解决立体匹配问题,这一问题已经成为立体视觉研究中的一个具有开放性、挑战性的研究课题。近几年,虽然已经出现基于双目或者多目摄像机进行三维估算的商业化产品,但是立体视觉中的这一难点问题仍然制约着该技术的推广应用,尤其是对于比较平坦或者具有周期性纹理特征的目标物体,基于立体视觉的三维信息估算仍然面临巨大的挑战。