



南京大學

研究生毕业论文 (申请硕士学位)

论文题目 基于结构光的3D重建系统

作者姓名 赵东威

学科、专业方向 电路与系统

指导教师 都思丹 教授

研究方向 计算机视觉

2013年5月

学 号 : MG1023041

论文答辩日期 : 2013 年 5 月 20 日

指导教师 : (签字)

3D Reconstruction System Based on Structured Light

by

Dongwei Zhao

Directed by

Professor Sidan Du

School of Electronic Engineering
Nanjing University

May 2013

*Submitted in partial fulfillment of the requirements
for the degree of Master in Circuit System*

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目： 基于结构光的3D重建系统

电路与系统 专业 2010 级硕士生姓名： 赵东威
指导教师（姓名、职称）： 都思丹 教授

摘要

3D信息采集是计算机视觉领域的热点问题，在医疗、文物保护以及机器人导航领域有着广泛应用，近年来成为了到了越来越多的研究焦点。从被关注起，研究人员提出了多种方法用于3D信息采集，包括立体成像、激光雷达、3D重建等方法，其中以3D重建不需要专用设备、要求图像处理过程简单、重建结果的分辨率高等特点而受到重视。

本文实现了基于投影仪-摄像机三角测量系统的3D重建。首先通过摄像机定标和投影仪定标获得摄像机和投影仪之间的几何关系，建立三角测量的硬件环境和几何基础。其次利用两种结构光进行3D重建，一种结构光是格雷码加相移编码方式。这种结构光编码方式属于时间编码，需投射多幅结构光并记录多幅图像，故适用于静止物体的重建，其优点是能够得到高分辨率的重建结果。另一种结构光是基于De Bruijn序列编码的颜色条纹，这种编码方法属于空间编码，采用这种结构光的重建只需拍摄一幅图像即可实现，具有重建时间短、重建效果好的优点。由于只需要记录一幅图像，此方法适用于动态物体的重建，而由此引起的缺点是，相比于时间编码的结构光，此方法要求的编码解码处理较为严格，理想的重建结果也需要借由高超的特征编码和特征检测技术达到。针对空间编码结构光的关键技术，本文以De Bruijn序列编码的颜色条纹为结构光，在其编解码技术方法的多个环节进行了多项创新，包括投射颜色选择、颜色校正、线条检测、峰值定位和颜色识别等，提出了有效的方法并提升了特征检测的准确率，取得了很好的重建效果，同时扩展了重建系统的适用范围。而对时间编码的结构光，本文采用格雷码编码的结构光加相移编码方式进行重建，并提出了对格雷码光栅边缘进行准确定位的方法，同样提升了重建的效果。

关键词： 3D重建；三角测量；结构光；空间编码；时间编码

南京大学研究生毕业论文英文摘要首页用纸

THESIS: 3D Reconstruction System Based on Structured Light

SPECIALIZATION: Circuit System

POSTGRADUATE: Dongwei Zhao

MENTOR: Professor Sidan Du

Abstract

In the last decade, 3D shape acquisition has received more and more considerable research interests. Since the moment when the concept of 3D reconstruction was proposed, researchers have developed some groups of techniques to achieve 3D shape acquisition, some important methods include structured light, stereo vision and TOF(Time Of Flight). Among those methods, one of the main advantages of structured light based techniques is that it requires less and easier image processing, meanwhile, it can derive high accuracy in the 3D reconstruction result. Based on the math of triangle principle, structured light based techniques take three steps to fulfill final result. The first step is to project some designed coding patterns onto the surface of the scanned objects. After that, a camera (or more) is controlled to capture the scene and send the pictures to computer for further processing. In the second step, algorithms detect the coding patterns in the captured image and establish the correspondence with projected patterns. In the final step, 3D coordinates of the object surface are calculated from the correspondence established in the second step within the calibrated triangle model.

In the review of the development history of structured light techniques, the first appeared techniques are the time-multiplexing techniques, which are also called multi-shots techniques because multi patterns are projected onto the surface and multi images are taken. Actually there are both advantages and disadvantages using this method. Advantages include high accuracy and resolution and the disadvantage is that the scanning speed is slow. After that, one-shot structured light system is developed to achieve high resolution within a short time. In one-shot structured light system, there are three key processes which would effect final reconstruction result, which includes the choose of projected patterns, detection of pattern in captured image and the algorithms applied to establishment of correspondence between captured and projected patterns. Recently,

more and more efforts have been made in order to generate higher quality result, which makes this method widely used in industry design and cultural heritage protection.

In this paper, we have build a structure light system and this system uses two methods: De Bruijn code based colored stripe and Gray code as structure light, which belongs to one-shot structured light technique and time-multiplexing technique. To achieve a better reconstruction result using colored stripes, we propose an adaptive color calibration method, and a Discrete Trend Transform (DTT) algorithm to detect patterns. And we proposed a method to detect the edge of Gray code structured light edge. In this system, dark laboratory environments is not needed to derive high-resolution 3D point cloud.

Keywords: 3D Reconstruction; structured light; De Bruijn sequence; Gray code.

目 录

目录	iv
第一章 绪论	1
1.1 引言	1
1.2 3D 重建的技术综述	2
1.3 结构光技术综述	3
1.4 本文的主要工作和研究内容	4
第二章 投影仪-摄像机系统校正	5
2.1 摄像机模型及定标	5
2.1.1 坐标系的定义	5
2.1.2 世界坐标系到相机坐标系的转换	6
2.1.3 相机坐标系到成像平面参考系的转换	7
2.1.4 摄像机畸变	8
2.1.5 摄像机定标	10
2.2 投影仪定标	11
2.3 三维信息的计算	13
2.4 本章总结	14
第三章 结构光为颜色条纹序列的3D重建	15
3.1 系统	15
3.1.1 架构和流程	15
3.1.2 特征投影	16
3.2 crosstalk校正	16
3.3 相对反射率校正	17
3.3.1 相对反射率对观测的影响	17
3.3.2 相对反射率的计算	18
3.3.3 反射率校正的结果	19

3.4 特征识别	20
3.4.1 M 通道定义	21
3.4.2 离散趋势变换	23
3.4.3 峰值定位	24
3.4.4 颜色提取	26
3.4.5 建立检测特征及投射特征间的对应关系	28
3.5 重建结果	29
3.6 本章总结	30
第四章 结构光为格雷码投影光栅的 3D 重建	32
4.1 特征投影	32
4.1.1 格雷码光栅	32
4.1.2 位移码图	33
4.2 编码图像的处理	35
4.2.1 有效区域的选择	35
4.2.2 解码位移图码光栅	36
4.2.3 解码格雷码光栅	36
4.3 重建结果	39
4.4 本章总结	40
第五章 总结和展望	42
5.1 研究总结	42
5.1.1 摄像机-投影仪系统定标	42
5.1.2 结构光选择和设计	42
5.1.3 特征识别	42
5.2 展望	43
5.2.1 3D点云的处理	43
5.2.2 实现实时重建	43
参考文献	44
致谢	47

第一章 绪论

1.1 引言

人类生活在一个由三维的空间和时间组成的四维立体空间中，我们的眼睛不仅能感知周围事物的色彩光泽，还能够通过双目立体视觉的方式感知立体空间。在科技不发达的古代和近代，由于缺少描述和记录三维世界的方法，人们只能通过画画、照相、电影等行为以二维的方式记录周围的世界。随着计算机的发明和极快的普及，互联网和计算能力的快速增长深刻地影响和改变着人们的生产生活方式，其中尤其以3D技术的影响最为突出。3D技术使人们有能力对现实三维世界的认识和表达回归到原始的直观、立体的境界。在现代化的21世纪，3D的概念已然充斥着虚拟的世界，也在深刻影响着现实的世界。从3D游戏、动画、3D电影到计算机辅助制造、3D打印、医学成像、自动驾驶汽车等，处处体现3D技术的应用。如果现在的世界不能称为3D世界，那么可以预想的是，不久的未来，3D技术和设备的应用将会更加普及和大众化，届时3D信息的采集、处理等过程也将成为普通大众习以为常的日常行为。

3D技术包含一系列的过程，包括采集、处理、识别等，而这其中，3D信息采集技术是其他技术的基础，其质量、可靠性均影响着处理、识别等过程，可以说，有意义的3D信息才可能经由进一步处理和转化进而提供有意义的数据和信息。

在多种3D信息采集技术中，3D重建是重要手段，其应用包括文物保护、医疗、机器人导航、模型生成、计算机视觉等。当然在一些设计领域，我们可以通过设计软件设计物体虚拟的3D信息（如计算机辅助设计CAD），其优点是制作精细，简便可控，但同时具有需要事先设计模型各组成部分比例，需要熟练操作工人，开发周期长的缺点。而3D重建则基于对已有物体三维信息的采集和处理，并在此基础上建立模型，在时间成本和真实感上具有很大优势，例如人脸模型就广泛用于电影角色创建、计算机游戏、外科手术、视频会议和虚拟现实等应用领域，另外，对于一些特殊需求则必须通过3D重建实现，如扫描历史文物3D信息等。

另一方面，我们应当注意到，3D打印有越来越普及的趋势，使用3D打印机可方便地打印现有的3D模型。目前3D打印设备已经实现了产品化，其价格随着时间的推移也越来越接近普通大众的接受范围，可以预计的是，3D打印的风行也必将激发更多的3D信息的需求，同时也推动3D重建技术的进一步普

及和发展。而就 3D 重建本身来看，其实现不需要太过复杂或精密的设备，只需要简单的家用设备即可完成，甚至一种名为通过聚焦的重建方法(shape from focus) 只需要一台可变焦摄像机即可。对设备要求低的特点为 3D 重建的大范围应用提供了很好的可实现性。本文通过使用一台家用摄像机和一台家用投影仪，使用结构光的方法完成了 3D 重建系统[17]，系统采用两种结构光分别实现重建，一种是空间编码的颜色序列结构光，另一种是时间编码的格雷码光栅。

1.2 3D 重建的技术综述

从 3D 重建的概念被提出到目前为止，有多种技术被提出用于重建，主要可分为接触测量和非接触测量两大类。

A 接触测量法

此方法中，利用传感触头与被测物体的表面相接触并直接测量本点的深度信息。直接测量法的原理相当简单，且具有分辨率高，不受物体表面颜色、透明度等非形状因素干扰的优点。但其缺点是，传感触头与物体的直接接触则会一定程度地导致物体的轻微形变从而影响测量的精度，同时，传感触头的物理接触则不能排除对物品损坏的可能，这也使得本方法不适用于珍贵物品或极易形变物体的重建。

B 非接触测量法

顾名思义，非接触测量法在测量中不产生与被测量物体直接的物理接触（但可能会有光或声的接触），只通过算法对被测量物体反射的声或光信息进行计算以识别物体的表面形状特性。没有物理接触可以保证对被测物体不会产生伤害，这个因素保证了非接触测量法的可使用范围要比接触测量法要更为广泛。非接触测量法是 3D 测量的重要部分，其中主要的方法包含如下：

1 立体成像

人类的眼睛是天然的立体成像系统，两只眼睛在水平方向有一段距离，这段差距使得我们在观察三维物体左右眼产生视差，大脑在接收到双眼图像时检测视差并通过计算将视差转化为深度，从而感知到物体的距离、产生立体的感受。

立体成像[12]采用与人眼一样的原理，它利用两台或多台摄像机对物体从不同角度进行拍照并交由计算机处理并建立对应关系，再结合三角测量的原理就可以将探测得到的位置偏差转换为物体表面的三维坐标信息，进而实现从二维空间到三维空间的转换。此方法的优点包括设备简单，只需两台或多台家用摄像机即可，成本很低，且在测量中不与物体产生直接接触，但其缺点是要求被测物体有足够的纹理以便能够进行识别，故对于表面光滑、缺少颜色变化的

物体其重建效果并不太好。

2 飞行时间法

飞行时间法(TOF) 又被成为雷达测距，其过程是向被测物体发出经调制的激光信号，经反射后由探头接收到，利用相移特性计算被测物体表面到测量仪器的距离。这种方法的优点是不受物体表面颜色或环境光的干扰，但其要求精密的信号处理过程，对时间精度要求很高，需要使用专用、精密的调制、探测和解码设备。

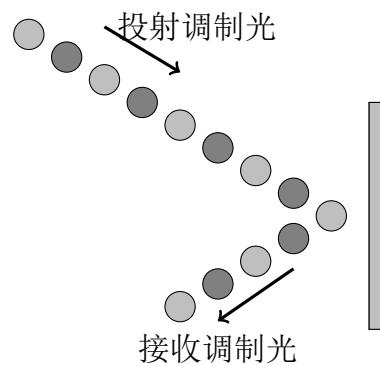


图 1.1: 飞行时间法

3 结构光法

结构光法[1]同样利用三角测量的原理，通过投影仪投射结构化（时间上或空间上）的特征图形，并在摄像机拍摄的图像中检测相应特征并与所投射图像中的特征建立相应关系，随后利用三角模型计算物体表面的距离信息。该方法原理简单，成本低廉，可用家用摄像机和投影仪实现。而结构光有多种选择，包括点、线和栅格图像等。结构光的编码方式是基于结构光的 3D 重建系统的基础。

1.3 结构光技术综述

基于结构光的 3D 重建系统是获取物体表面 3D 特性的重要方法，而其中，结构光的选择对精度、鲁棒性等因素有着深刻的影响。最先被提出的结构光是点结构光[19]，后来发展出了线结构光，很大地提升了处理速度。到目前为止，发展出的结构光包括：基于二进制码、格雷码的时间编码结构光，基于 M-arrays、De Bruijn 序列等空间编码的结构光，Salvi 等在[16] 中对各种形式的结构光进行了比较。

1.4 本文的主要工作和研究内容

本文实现了一个基于结构光的 3D 重建系统，并实现了 Gray 码为结构光和使用 De Bruijn 序列颜色条纹为结构光的两种结构光方法分别实现。本文提出了一系列的创新点用于提高系统对物体表面颜色和环境光的抗干扰能力，有效地提高了系统对环境光等干扰因素的免疫能力，拓宽了系统的适用范围。本文的章节组成如下。

第一章：对几种常用的 3D 重建方法进行了分类和介绍，并介绍了结构光的原理和发展。

第二章：本章重点分析和实现了摄像机和投影仪的畸变校正和定标，为精确实现 3D 测量提供准确的三角模型。

第三章：介绍了采用以 De Bruijn 序列编码的颜色条纹为结构光的 3D 重建系统，对其中的重点步骤如颜色校正、颜色线条查找、线条颜色识别等步骤进行了创新并予以重点说明。本章提出的创新点帮助系统有能力实现高分辨率的 3D 重建系统，并具有对环境光等干扰因素不敏感的特性。

第四章：本章介绍了采用格雷码光栅加相移码的 3D 重建系统，对格雷码光栅条纹的边界提取和确定提出了新的方法。

第五章：对本文所使用的方法、创新点进行总结，并讨论了进一步的可改进方面和方法。

第二章 投影仪-摄像机系统校正

投影仪-摄像机系统常用作三维重建系统，示意图见图 2.1，其优点是设备价格低廉，容易获取和安装。在系统中，为了实现测量，需要事先对摄像机和投影仪进行定标。

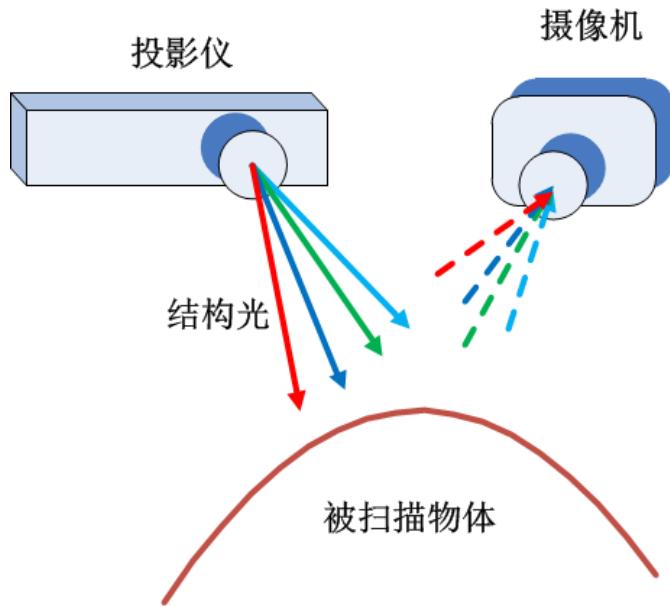


图 2.1: 系统结构图

2.1 摄像机模型及定标

2.1.1 坐标系的定义

对于相机模型的描述一般需要三个坐标系：世界坐标系、相机坐标系和成像平面参考系。其中世界坐标系是在真实的环境中选取的坐标系，可以任意设置。世界坐标系选定后，物体和摄像机就拥有世界坐标系的坐标。在接下来对坐标的描述和转换中，本文用 $Q_w = [X_w \ Y_w \ Z_w]^T$ 来表示世界坐标系中的一点。

摄像机是一个独立的成像系统，其成像的模型则需要用相机坐标系来描述。在相机坐标系中，坐标原点是摄像机的光心， Z 轴垂直于焦平面向外， X 轴和 Y 轴平行于焦平面。对于相机坐标系中的一点，本文用 $Q_c = [X_c \ Y_c \ Z_c]^T$ 予以描述。

成像平面参考系建立于摄像机的成像平面上，其坐标原点是相机坐标系 Z 轴与成像平面的交点，这一点也被称为摄像机的主点。理想情况下，主点应当位于成像平面的中心位置，而实际上，由于制作工艺等问题，成像平面的中心常偏离于主点，本文用 (c_x, c_y) 来描述偏差程度，用 $q = [x \ y]^T$ 描述成像平面参考系的点。

2.1.2 世界坐标系到相机坐标系的转换

将世界坐标系的点 $Q_w = [X_w \ Y_w \ Z_w]^T$ 转换为相机坐标系 $Q_c = [X_c \ Y_c \ Z_c]^T$ 的过程可以分为以下两步：

1. **旋转操作** 旋转世界坐标系，使得三个轴分别与相机参考系的 X 、 Y 、 Z 轴平行并且方向一致
2. **平移操作** 平移世界坐标系，使得世界坐标系的原点与相机坐标系的原点相重合

定义旋转矩阵 R 来描述旋转操作，定义平移矩阵 T 来描述平移操作，世界坐标系的坐标与相机坐标系的坐标则有如下关系：

$$Q_c = R \cdot Q_w + T \quad (2.1)$$

其中， R 是大小为 3×3 的正交矩阵：

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (2.2)$$

T 为 3×1 大小的矩阵：

$$T = [t_1 \ t_2 \ t_3]^T \quad (2.3)$$

将从世界坐标系转换到相机坐标系的过程表示为齐次形式：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2.4)$$

矩阵 R 和 T 构成了摄像机的外参数，表面世界坐标系和相机坐标系的相对位置关系。

2.1.3 相机坐标系到成像平面参考系的转换

在考虑从相机坐标系转换为成像平面参考系之前，这里先介绍摄像机的模型。在计算机视觉中，针孔模型[9]被广泛用于相机模型的描述。模型假设物体表面的反射光满足光的直线传播条件，并经过小孔投射在图形平面上，形成物体的倒像，其示意图见图 2.2，图中 X 为物体， x 为物体的倒像，像距和物距分别为 f 和 Z_c ，那么显然有如下关系成立：

$$-x = \frac{f}{Z_c} X \quad (2.5)$$

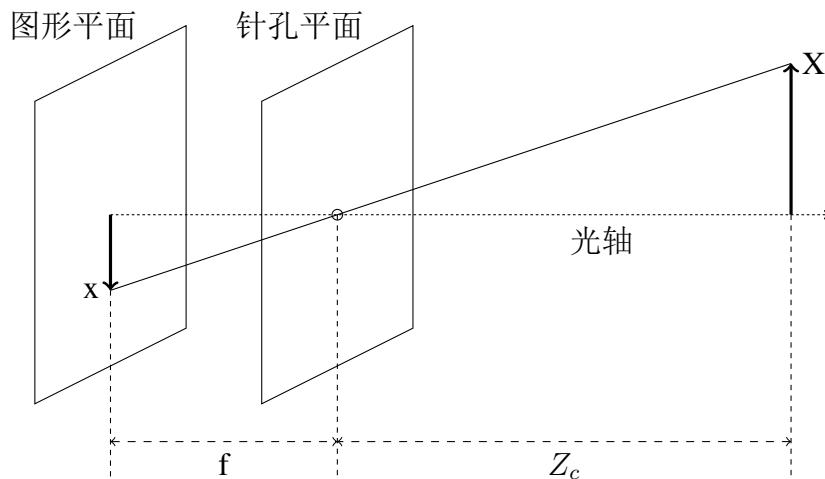


图 2.2: 小孔成像模型

为了简化针孔模型，将成像平面转移到针孔之前，可将针孔模型整理为一种等价形式，如图 2.3[3]。图中，针孔相当于投影中心，物体 Q_c 经过投影中心的光线投射到成像平面上，形成 q 。

图 2.3 中的针孔即为投影中心，但在实际的相机中，芯片的中心通常不位于光轴上。中心的偏差可以引进参数 c_x 和 c_y 来描述，同样引入两个不同的焦距： f_x 和 f_y 来描述两个方向上的焦距，于是将物体 $Q_c(X_c, Y_c, Z_c)$ 和像体 $q(x, y, f_x, f_y)$ 的数学关系描述为：

$$\begin{aligned} x &= X \frac{f_x}{Z_c} + c_x \\ y &= Y \frac{f_y}{Z_c} + c_y \end{aligned} \quad (2.6)$$

将公式 2.6 使用其次坐标重新表示为：

$$q = M Q_c \quad (2.7)$$

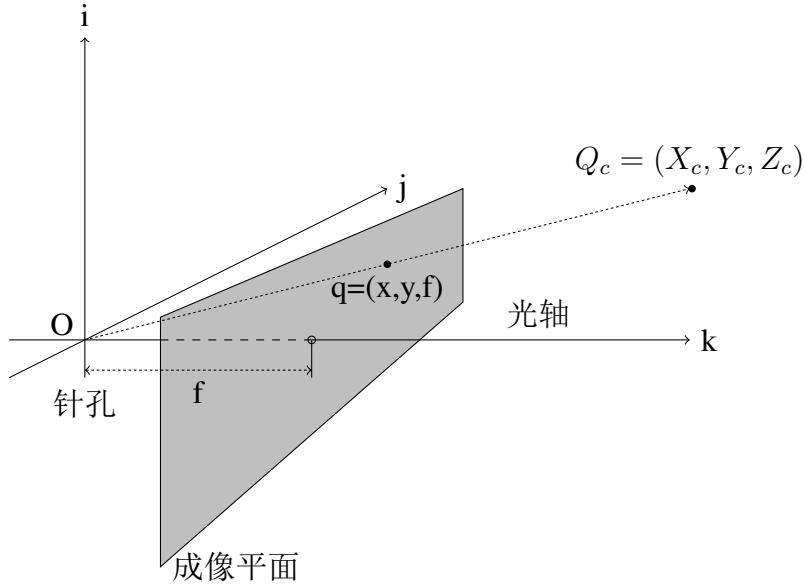


图 2.3: 针孔模型的等价变换

$$\text{其中, } q = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad Q_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}, \quad M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}.$$

公式 2.7 表明了从相机坐标系到成像平面参考系的转换过程，矩阵 M 包含相机内部参数的描述，故也被称作摄像机的内矩阵[10]。对摄像机定标的过程就是求得相机内参数与外参数的过程。

2.1.4 摄像机畸变

虽然摄像机可简化为针孔模型，然而在实际上，真实的针孔无法在短时间内为曝光收集足够的光线，因此在摄像机中使用透镜而非小孔。透镜的使用使得快速曝光成为可能，而其缺点则是引入了透镜的畸变。透镜的畸变主要分为两种：径向畸变和切向畸变[8]。径向畸变是由透镜形状引起，而切向畸变则由摄相机的组装过程引起。

径向畸变分为筒形畸变和枕形畸变，在光学中心的畸变程度为 0，越往感光元件的边缘移动，畸变程度越显著，参见图 2.4。通常，与整幅图片相比，这样的畸变程度比较小，可利用光学中心($r = 0$)位置周围的泰勒级数展开来近似描述，通常只取前两项或三项就足够了。假设 $(x_{corrected}, y_{corrected})$ 表示由点

(x, y) 校正后的位置，那么有如下公式：

$$\begin{aligned} x_{corrected} &= x(1 + k_1 r^2 + k_2 r^4) \\ y_{corrected} &= y(1 + k_1 r^2 + k_2 r^4) \end{aligned} \quad (2.8)$$

注意上式中没有出现 r 的奇次方项，这是由畸变程度 Δr 是 r 级数展开这个关系决定的。

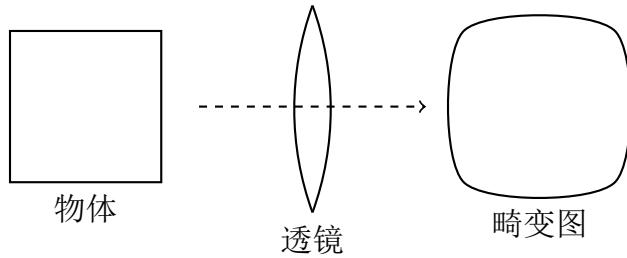


图 2.4: 筒形畸变

另一个常用的畸变是切向畸变，这是由于透镜不完全平行于图像平面的时候所产生的，示意图见图 2.5。引入两个额外的参数 p_1 、 p_2 来描述，得到如下的关系[4]：

$$\begin{aligned} x_{corrected} &= x + [2p_1y + p_2(r^2 + 2x^2)] \\ y_{corrected} &= y + [p_1(r^2 + 2y^2) + 2p_2x] \end{aligned} \quad (2.9)$$

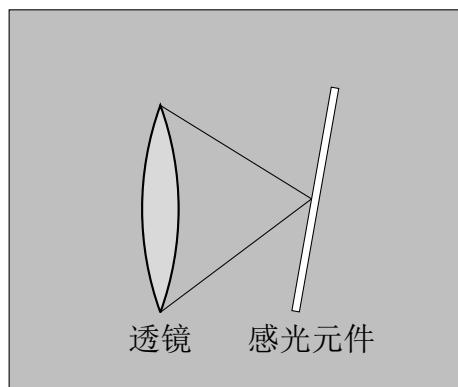


图 2.5: 摄像机结构，感光元件与透镜不平行

2.1.5 摄像机定标

摄像机定标的目标是求得摄像机的内参数、外参数和畸变的校正参数，分为内部参数(包括 $f_x, f_y, c_x, c_y, k1, k2, p1, p2$)和外部参数(包括 R, T)。本系统使用棋盘格作为定标物，调用 openCV 摄像机校正函数实现摄像机定标。

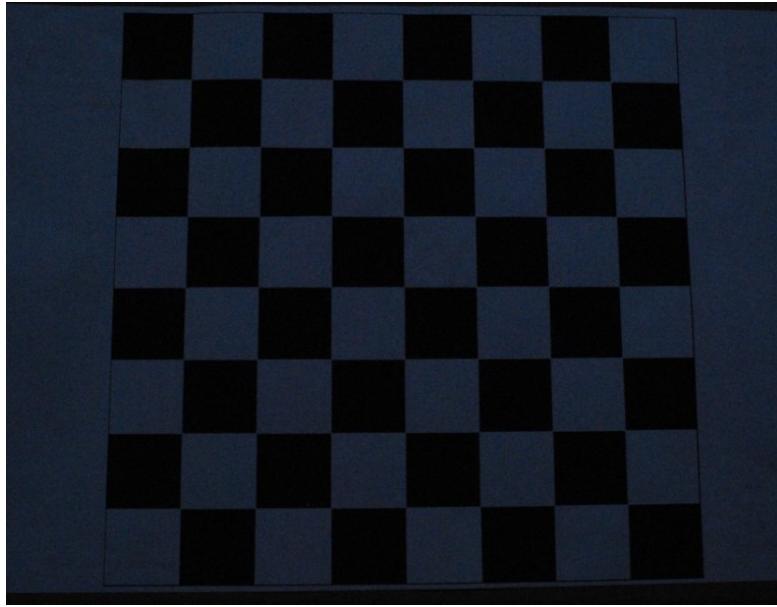


图 2.6: 棋盘格作为定标板

本文的定标板如图 2.6 所示，一幅图像中共有 7×7 个角点，即每两个深色方块的接触点。每个方块的边长为 3cm，将左上角的第一个角点设定为世界坐标系的原点， X 轴和 Y 轴平行于定标板， Z 轴垂直于定标板向外。根据角点的相对关系，很容易确定定标板上每个角点的世界坐标系坐标。图 2.7 是校正摄像机所使用的 9 幅定标图像，图中彩色点是从摄像机拍摄到图像中检测到的角点，其在图像中的位置可以确定成像平面坐标系的坐标，由此建立世界坐标系和成像平面坐标系的多个点对，并以此计算相机的内外参数。本文中这些参数的计算是调用 openCV 库中的 cvCalibrateCamera2 函数实现。

openCV 计算畸变参数的函数为：

```
void cvCalibrateCamera2( const CvMat* object_points, const CvMat* image_points,
                        const CvMat* point_counts, CvSize image_size,
                        CvMat* intrinsic_matrix, CvMat* distortion_coeffs,
                        CvMat* rotation_vectors=NULL,
                        CvMat* translation_vectors=NULL,
                        int flags=0 );
```

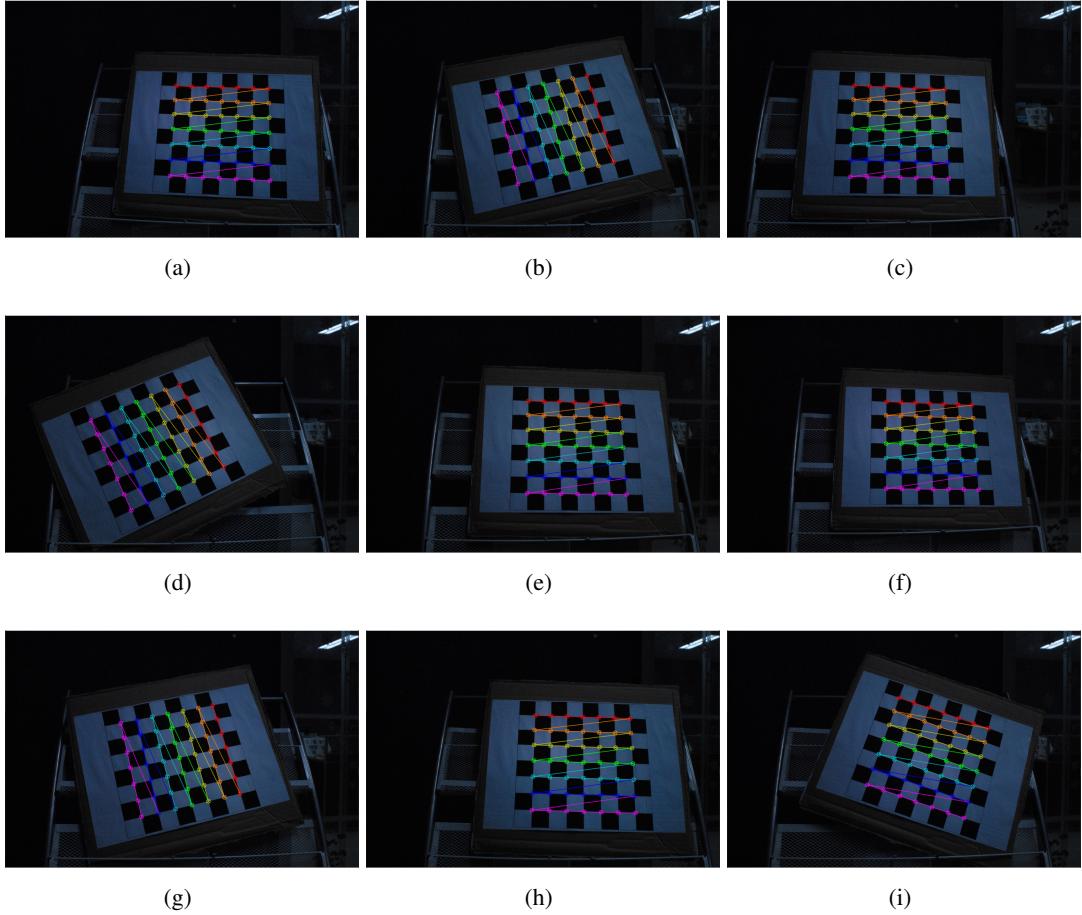


图 2.7: 棋盘格中找到的角点

`cvCalibrateCamera2` 函数依据张正友提出的校正算法[20]，以角点在世界坐标系和成像平面参考系中的坐标对为输入，输出则包含内参数矩阵 M 、畸变系数(k_1, k_2, p_1, p_2) 以及外参数的旋转矩阵 R 和平移矩阵 T 。内参数矩阵和外参数矩阵描述了世界坐标系和摄像机成像平面参考系之间的转换关系，而畸变系数可用于对摄像机拍摄图像进行的畸变校正。在 openCV 库中，畸变校正可通过调用函数 `cvUndistort2`、函数 `cvInitUndistortMap` 和函数 `cvRemap` 实现。

2.2 投影仪定标

摄像机定标后要进行投影仪定标。因为投影仪模型可以认为是相机模型的逆过程，即将 2D 图像映射为 3D 光线的过程，所以其定标过程与相机定标方法类似，同样需要在世界坐标系和投影图像平面参考系之间建立多个点对。投影图像平面参考系中点的坐标已知，而世界坐标系中点的坐标则需要借助已定标

的照相机完成。

定标板四周的四个角点经过摄像机检测以确定定标板的位置，而后将投影图像投射至定标板：

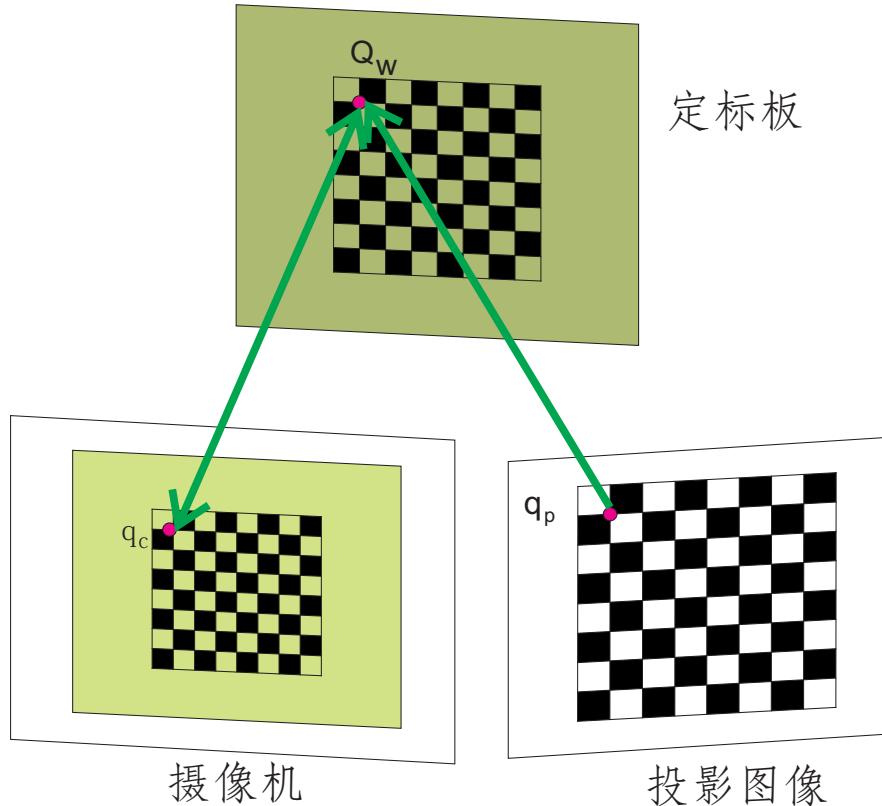


图 2.8: 投影仪定标中3D点与2D点的对应关系

在图 2.8 中，摄像机拍摄到投射了定标图像的定标板，并检测到一个角点 q_c ，根据已经定标的相机坐标系和成像平面参考系可推出世界坐标系中三维点 Q_w 的坐标，而在设计投影图像时可知投影图像中角点 q_p 在投影图像平面参考系的坐标，由此建立 Q_w 和 q_p 的坐标对应关系，并据此完成投影仪的定标。

总的来看，将摄像机定标和投影仪定标可表述为以下流程：

1. 架设摄像机、投影仪，并将其焦距调整合适，将定标板置于摄像机和投影仪视野下。
2. 摄像机拍摄定标板图像(记为 I_c)并检测角点，并匹配和记录角点在图像中的坐标、在世界坐标系的坐标，如定标板上角点未被全部检测出则重新拍摄图像，直至全部角点被检测出。

3. 定标板不动，在其表面铺设白纸，投影仪投射投影仪定标图像，见图 2.9。
4. 摄像机拍摄图像(记为 I_p)提取投影图像的角点，记录角点在所拍摄图像中的位置，如角点未被全部检测出则重新拍摄 I_p 。
5. 适度改变定标板位置和角度。
6. 重复步骤 2-5 共 9 次，调用 cvCalibrateCamera2 函数处理步骤2 得到的角点坐标对，计算摄像机内外参数，完成摄像机定标
7. 根据摄像机参数计算投影图像角点的三维信息，建立投影图像角点在世界坐标系下和投影图像下的坐标对，同样调用 cvCalibrateCamera2 函数计算投影仪内外参数，完成投影仪定标。

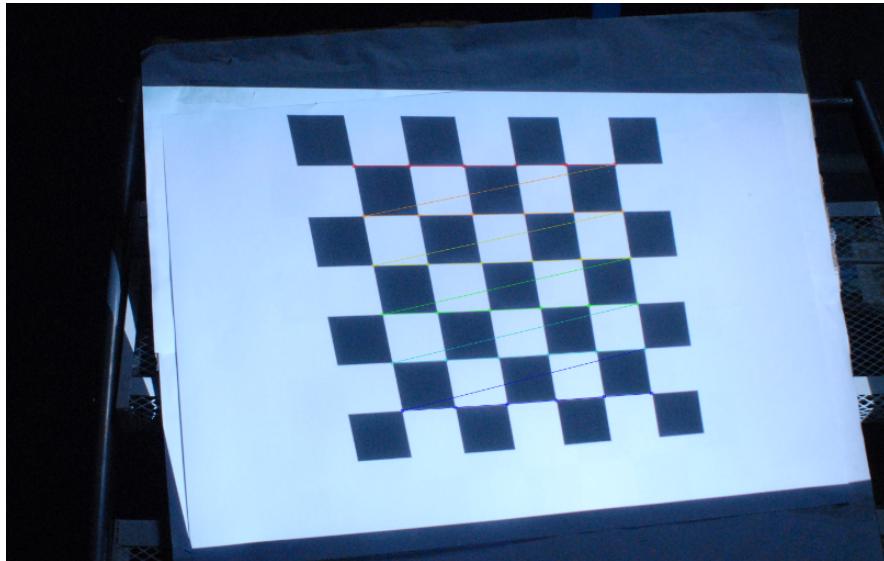


图 2.9: 投影仪投射定标图像后的定标板及找到的角点

2.3 三维信息的计算

基于投影仪-摄像机的三维重建系统可简化为如下模型： 经过定标，投影仪的光心 q_1 、摄像机光心 q_2 在世界坐标系下的坐标均为已知。摄像机光线与投影仪光线相交于物体表面上的 p 点，方向向量分别为 v_1 和 v_2 。容易推断得到如下关系：

$$\begin{aligned} p &= q_1 + \lambda_1 \cdot v_1 \\ p &= q_2 + \lambda_2 \cdot v_2 \end{aligned} \tag{2.10}$$

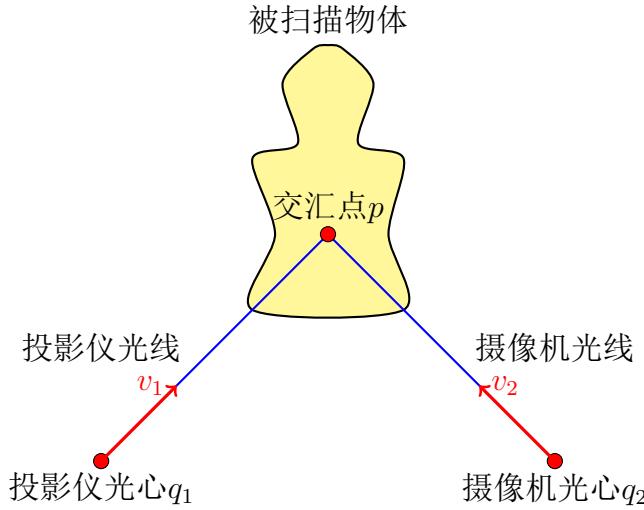


图 2.10: 三角测量的简化模型

根据以上的两式可以计算出 λ_1, λ_2 , 进而计算出 p 在世界坐标系的坐标。

在实际的情况下, 摄像机光线与投影仪光线并不一定会相交, 这种情况下上述两式无法解出。解决办法是, 求出 λ_1, λ_2 , 使得如下的目标函数最小:

$$\|(q_2 + \lambda_2 v_2) - (q_1 + \lambda_1 v_1)\|^2 \quad (2.11)$$

求解可得:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \|v_1\|^2 & -v_1^t v_2 \\ -v_2^t v_1 & \|v_2\|^2 \end{bmatrix}^{-1} \begin{bmatrix} v_1^t (q_2 - q_1) \\ v_2^t (q_1 - q_2) \end{bmatrix} \quad (2.12)$$

2.4 本章总结

本章首先介绍了三角测量的数学原理, 指出实现三角模型的精确计算需要预先知晓的摄像机与投影仪的相对位置关系, 其次介绍了摄像机的成像模型和基于定标板的定标方法, 并实现了投影仪的定标。投影系统的定标是实现精确 3D 重建的数学基础。

第三章 结构光为颜色条纹序列的3D重建

本章介绍以 De Burijn 序列为编码颜色条纹作为结构光的 3D 重建方法，能够实现在环境光下对物体的 3D 扫描。本文在此系统中提出了较多的创新点，包括提出对被扫描物体表面的反射率校正方法、提出局部趋势变换的概念并用于线条检测、峰值定位方法以及改进动态规划的 cost 函数并用以建立投射特征和被检测特征间的对应关系建立等。以上的创新方法使得系统具备一定对环境光和被扫描物体表面颜色的抗干扰能力，很好地扩展了系统的适用范围。同时所采用的高分辨率投影仪也使 3D 重建具有很高的分辨率，例如重建后一幅人脸的点云包含了多达 75371 个三维点。

3.1 系统

3.1.1 架构和流程

所使用的设备是：

- 投影仪: Acer H7350D (1920*1080 分辨率)
- 摄像机: Nikon D200

整个 3D 重建的流程表述如下[22]:

1. 投射颜色序列，并在环境光下抓取图像
2. 摄像机及投影仪畸变校正
3. 颜色校正，包括 crosstalk 校正和反射率校正
4. 分离线条并进行峰值定位
5. 颜色识别并找出识别出的特征与所投射特征的对应关系
6. 计算 3D 点云信息，并进行后续处理

3.1.2 特征投影

为了方便地建立投影特征和检测特征之间的对应关系，本文根据 De Bruijn 序列来设定线条颜色。一个拥有5个元素、3阶的 De Bruijn 序列为：4 4 4 3 4 4 2 4 4 1 4 4 0 4 3 3 4 3 2 4 3 1 4 3 0 4 2 3 4 2 2 4 2 1 4 2 0 4 1 3 4 1 2 4 1 1 4 1 0 4 0 3 4 0 2 4 0 1 4 0 0 3 3 3 2 3 3 1 3 3 0 3 2 2 3 2 1 3 2 0 3 1 2 3 1 1 3 1 0 3 0 2 3 0 1 3 0 0 2 2 2 1 2 2 0 2 1 1 2 1 0 2 0 1 2 0 0 1 1 1 0 1 0 0 0。投射的颜色选择见表 3.1，注意选取的颜色中排除了白色和黄色，排出白色是因为线条检测和颜色识别的特殊需要，而排除黄色的原因是，黄色与人的肤色相近，在对人体扫描时容易引进干扰。结合 De Bruijn 序列和各元素颜色可以生成投影背景，见图 3.1。

表 3.1: 投影颜色选择

编码	颜色(R, G, B)	色度
0	(255, 0, 0)	0°
1	(0, 255, 0)	60°
2	(0, 0, 255)	120°
3	(0, 255, 255)	180°
4	(255, 0, 255)	240°

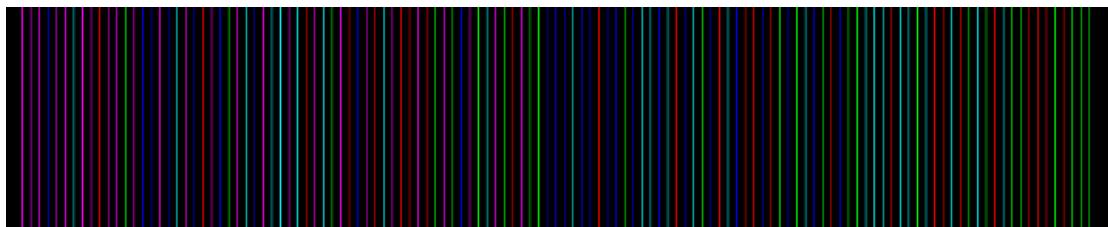


图 3.1: 投影背景

3.2 crosstalk校正

总体上讲，影响系统稳定性的因素包括以下几个：

1. 投影仪及摄像机存在的 crosstalk 现象
2. 环境光

3. 通常，物体表面对不同颜色通道具有不同的反射率，形成颜色的扭曲
4. 拍摄到的图像中，颜色线条的宽度和强度会受到物体表面的崎岖特性的影响

为了提高系统的鲁棒性，需要对以上的干扰因素采取措施。其中的一个方法是进行颜色校正。颜色校正包括两步：crosstalk 校正消除颜色通道间的 crosstalk 现象；反射率校正可以通过校正个颜色通道实现反射率处在相同水平。

认为颜色系统的模型为公式 3.1[5]：

$$\underbrace{\begin{bmatrix} c^r \\ c^g \\ c^b \end{bmatrix}}_{I_c} = \underbrace{\begin{bmatrix} x_{rr} & x_{rg} & x_{rb} \\ x_{gr} & x_{gg} & x_{gb} \\ x_{br} & x_{bg} & x_{bb} \end{bmatrix}}_X \underbrace{\begin{bmatrix} a^r & 0 & 0 \\ 0 & a^g & 0 \\ 0 & 0 & a^b \end{bmatrix}}_A \left(\underbrace{\begin{bmatrix} p^r \\ p^g \\ p^b \end{bmatrix}}_{I_p} + \underbrace{\begin{bmatrix} o^r \\ o^g \\ o^b \end{bmatrix}}_O \right) \quad (3.1)$$

这里的 X 矩阵为 crosstalk 矩阵，可以在暗室环境中通过投影 RGB 三幅纯色图至纯白色物体并计算拍摄到图片的颜色误差得到。矩阵 A 为各颜色通道的反射率。

3.3 相对反射率校正

3.3.1 相对反射率对观测的影响

公式 3.1 中的矩阵 A 为反射率矩阵，经过简单转换，可以得到：

$$A = \begin{bmatrix} a^r & 0 & 0 \\ 0 & \alpha a^r & 0 \\ 0 & 0 & \beta a^r \end{bmatrix} = a^r \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} \quad (3.2)$$

公式 3.2 中的 α 、 β 分别为 G 通道、B 通道相对 R 通道的相对反射率， a^r 代表 R 通道的反射率。为了进行颜色校正，对公式 3.1 进行转换，并将公式 3.2 代入得到：

$$\begin{aligned} I_p + O &= \tilde{A}^{-1} X^{-1} I_c \\ &= a^{-r} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{\alpha} & 0 \\ 0 & 0 & \tilde{\beta} \end{bmatrix}^{-1} X^{-1} I_c \end{aligned} \quad (3.3)$$

公式 3.3 中的 \tilde{A} 由 $\tilde{\alpha}$ 、 $\tilde{\beta}$ 组成，前者是 A 的近似估计，而后者分别是 G、B 通道相对于 R 通道的相对反射率估计。下文介绍 $\tilde{\alpha}$ 、 $\tilde{\beta}$ 的计算方法。

令 I_a 代表经过 crosstalk 校正和相对反射率校正后的颜色值，可以得到：

$$\begin{aligned} I_a &= a^r(I_p + O) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{\alpha} & 0 \\ 0 & 0 & \tilde{\beta} \end{bmatrix}^{-1} X^{-1} I_c \end{aligned} \quad (3.4)$$

观察公式 3.4 得到，经过 crosstalk 校正和相对反射率校正的颜色 I_a 由 3 部分组成：

- 投影颜色 I_p
- 环境光 O
- R 通道反射率 a^r

I_a 中虽然仍保留有环境光和 R 通道的反射率这两个干扰因素，但已经排除了 crosstalk 现象和不同反射率的影响。 I_a 也是作为特征检测的直接输入。在特种检测中，更多的策略被采用以消除环境光和 a^r 对线条检测和颜色识别的影响。而为了计算 I_a ，不仅需要进行 crosstalk 校正，还要进行相对的反射率校正。实际系统中，EMD(Earth Mover Distance)[15] 方法被用来分析颜色直方图并估算相对反射率。

3.3.2 相对反射率的计算

物体表面上 G 通道相对于 R 通道的相对反射率 $\tilde{\alpha}$ 的计算方法是，首先建立 R 通道和 G 通道各自的直方图，进而在直方图之间建立流动矩阵，通过分析流动矩阵的值可以实现相对反射率的估计。

考虑 H^r 代表 R 通道的直方图，如下：

$$H^r = \{(H_1^r, W_1^r), (H_2^r, W_2^r), \dots, (H_n^r, W_n^r)\} \quad (3.5)$$

其中， $H_i^r (1 \leq i \leq n)$ 是直方图的灰度值， $W_i^r (1 \leq i \leq n)$ 是对应灰度的像素个数。假设 r、g 和 b 分别代表红、绿和蓝通道，首先 H^g 被转换为 H^g' ， H^g' 作为 H^r 的最好匹配，并且其转换过程可表示为流矩阵的形式：

$\mathbf{f} = f_{ij}$ ($1 \leq i \leq n, 1 \leq j \leq n$)；齐次是计算流矩阵 f 以描述从 H^g 到 $H^{g'}$ 的转换过程。 f 可通过最小化以下的目标函数求得，目标函数表示为公式 3.6。

$$cost(H^g, H^r) = \sum_{i=1}^n \sum_{j=1}^n |j - i| f_{ij} \quad (3.6)$$

这里的 f_{ij} 代表从 H_i^g 到 H_j^r 的流， $|j - i|$ 代表流动距离。上述目标函数的约束为公式 3.7 到公式 3.10。

$$f_{ij} \geq 0 \text{ for } 1 \leq i \leq n, 1 \leq j \leq n \quad (3.7)$$

$$\sum_{i=1}^n f_{ij} \leq W_j^r, 1 \leq j \leq n \quad (3.8)$$

$$\sum_{j=1}^n f_{ij} \leq W_i^g, 1 \leq i \leq n \quad (3.9)$$

计算流矩阵的目标函数可以表示为：

$$\sum_{i=1}^n \sum_{j=1}^n f_{ij} = \min\left(\sum_{i=1}^n W_i^g, \sum_{j=1}^n W_j^r\right) \quad (3.10)$$

流矩阵 \mathbf{f} 中， f_{ij} 代表从 i 灰度级搬移到 j 灰度级的像素个数。于是相对反射率估算公式为：

$$\tilde{\alpha}(i) = \frac{(\sum_{j=1}^n f_{ij})i}{\sum_{i=1}^n \sum_{j=1}^n (f_{ij}j)}, 1 \leq i \leq n, 1 \leq j \leq n \quad (3.11)$$

B 通道的相对反射率可通过相似的过程求得，这里不再叙述。

3.3.3 反射率校正的结果

反射率校正的目的是通过校正使得 RGB 通道的反射率处于相同水平。图 3.2 中，(a)(b) 分别是反射率前后的图片，(c)(d) 是对应的各通道直方图。可以清晰看到，(a) 中颜色不鲜艳、不明朗，蓝色、绿色光均没有红色光强烈，对应的直方图中，R 通道具有比 G, B 通道更高的亮度。而经过反射率校正后，(d) 中的 RGB 通道的亮度分布基本相似。

由于校正了颜色的畸变，反射率校正同样提高了颜色的识别精度。图 3.3 对比了反射率校正前后的颜色识别准确率，经过反射率校正，颜色识别的准确率提高了 16.18%。

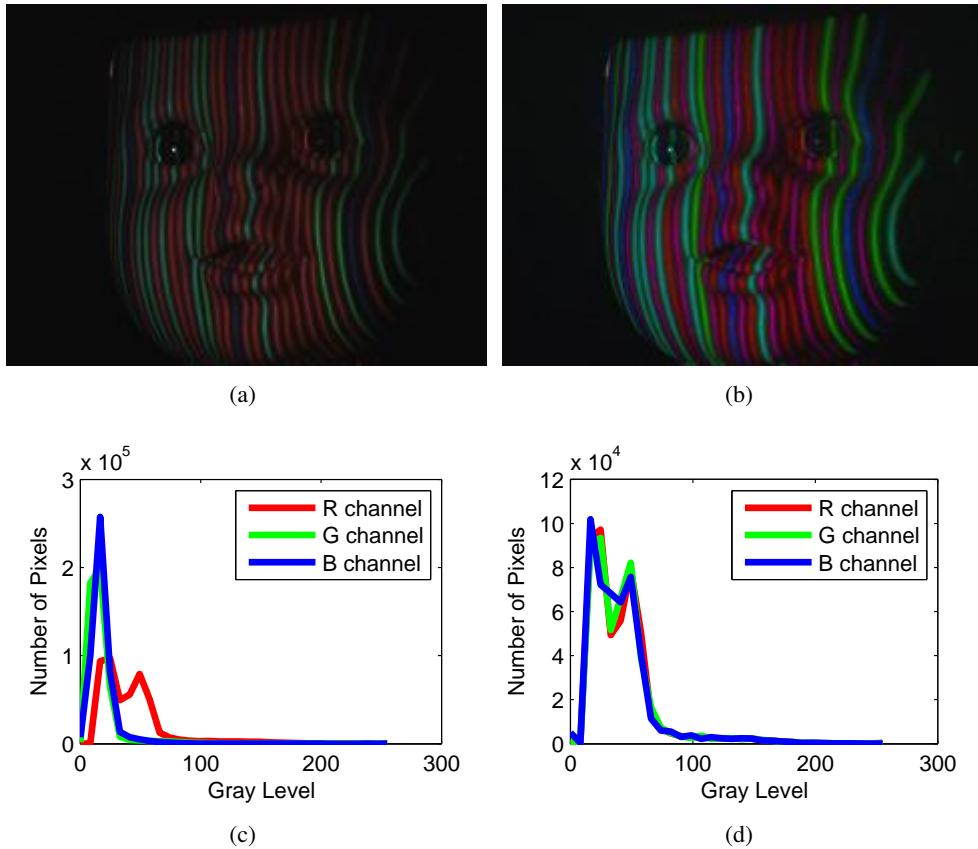


图 3.2: 反射率校正的结果. (a) 加照了结构光的脸部模型, (b) 反射率校正后的图像, (c), (a)的直方图, (d), (b)的直方图.

3.4 特征识别

crosstalk 校正和反射率校正的目的是校正所拍摄图片的颜色畸变，以便更准确地检测到所投影特征：线条位置及其颜色。为了能够提高检测颜色线条的召回率和准确率，本文引入了两个概念： M 通道和离散趋势变换。 M 通道是指将 R、G、B 三个通道融合为一个通道，避免了分别处理 R、G、B 三个通道的复杂性，同时结合所投射颜色序列的特殊性，最大限度地保留了原始线条的特征，另外， M 通道的定义也使其具有一定去背景光干扰的能力。离散趋势变换将 M 通道中线条的变化特性转换为更易检测的形式，很大程度上提高了线条检测成功率。

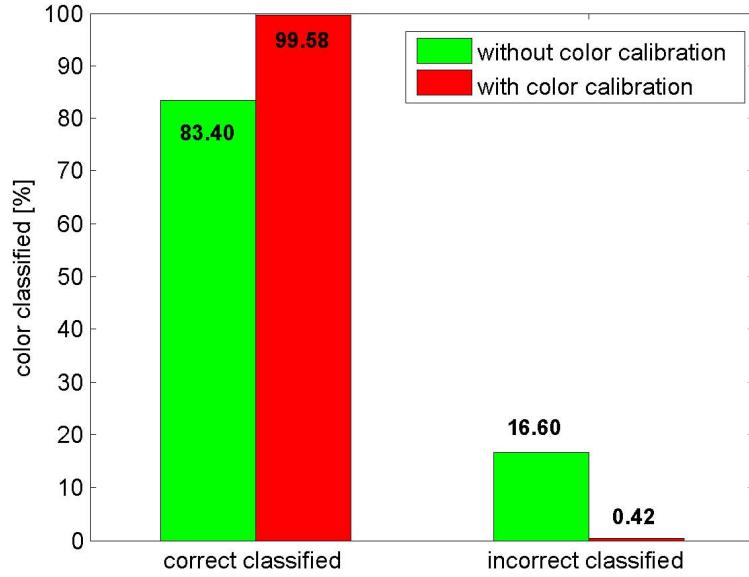


图 3.3: 颜色校正前后的颜色识别准确性比较.

3.4.1 M 通道定义

M 通道被定义为当前像素下，R、G、B 三个通道的最大值与最小值之差，如公式 3.12[21]。

$$M_{ij} = \max(C_{ij}^r, C_{ij}^g, C_{ij}^b) - \min(C_{ij}^r, C_{ij}^g, C_{ij}^b) \quad (3.12)$$

这里的 C_{ij} 是公式 3.4 中 I_a 图像中的像素值，即颜色校正后的像素值。接下来的处理如中心线查找和峰值定位都是基于 M 通道进行的。 M 通道的定义形式决定了其具有排除环境光干扰的能力，其证明见公式 3.13 和公式 3.14。

根据公式 3.4， C_{ij} 具有如下形式：

$$\begin{bmatrix} C_{ij}^r \\ C_{ij}^g \\ C_{ij}^b \end{bmatrix} = a^r \begin{bmatrix} p_{ij}^r + o_{ij}^r \\ p_{ij}^g + o_{ij}^g \\ p_{ij}^b + o_{ij}^b \end{bmatrix} \quad (3.13)$$

将公式 3.13 代入公式 3.12，并认为环境光主要有白色光组成，即 $o_{ij}^r \approx o_{ij}^g \approx o_{ij}^b$ ，可以得到：

$$\begin{aligned}
 M_{ij} &= a^r \max \begin{bmatrix} p_{ij}^r + o_{ij}^r \\ p_{ij}^g + o_{ij}^g \\ p_{ij}^b + o_{ij}^b \end{bmatrix} - a^r \min \begin{bmatrix} p_{ij}^r + o_{ij}^r \\ p_{ij}^g + o_{ij}^g \\ p_{ij}^b + o_{ij}^b \end{bmatrix} \\
 &= a^r \max \begin{bmatrix} p_{ij}^r \\ p_{ij}^g \\ p_{ij}^b \end{bmatrix} - a^r \min \begin{bmatrix} p_{ij}^r \\ p_{ij}^g \\ p_{ij}^b \end{bmatrix} \\
 &= a^r \max \begin{bmatrix} p_{ij}^r \\ p_{ij}^g \\ p_{ij}^b \end{bmatrix} \tag{3.14}
 \end{aligned}$$

注意公式 3.14 中存在如下关系：

$$\min([p_{ij}^r \ p_{ij}^g \ p_{ij}^b]^T) = 0 \tag{3.15}$$

这是因为，所投射的各颜色的 R、G、B 三个通道之中，至少存在一个为 0 的通道。具体可参看表 3.1。从以上的推演可见，M 通道只和两个因素相关：

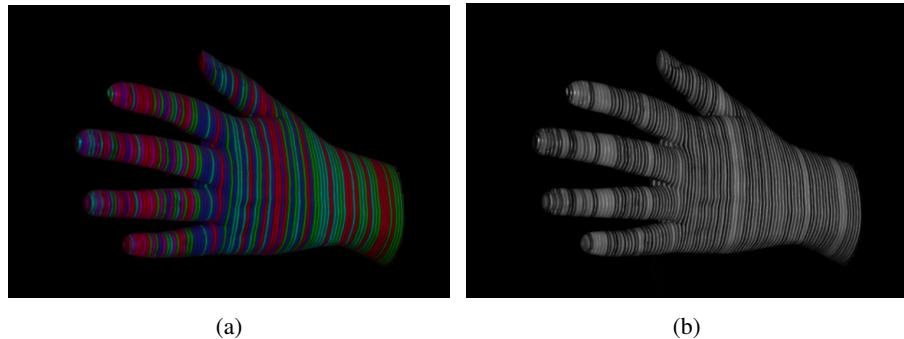


图 3.4: (a) 加照了光栅的手 (b) 将(a)转换后得到的M通道图

- 投影颜色 I_p

- R 通道反射率 a^r

M 通道与环境光 O 无关。

3.4.2 离散趋势变换

在 M 通道的基础上, 为了进一步排除 \mathbf{R} 通道反射率对中心线查找的影响, 本文提出离散趋势变换(DTT)的概念。离散趋势变换被定义为如下形式[13]:

$$T_{ij} = \sum_{k=h+1}^{j+N} \sum_{h=j}^{j+N-1} sign(M_{ik} - M_{ih}) \quad (3.16)$$

$$sign(a) = \begin{cases} 1 & a > 0 \\ 0 & a = 0 \\ -1 & a < 0 \end{cases} \quad (3.17)$$

其中的 N 为窗口大小, 实际应用中窗口大小可设定线条平均宽度的一半。考虑窗口大小远小于图像尺寸的实际因素, 则以下条件近似成立: $a_{ik}^r \approx a_{ih}^r$ 。此近似条件可以理解为, 在物体表面上, 相近的点具有接近的反射率。将这个近似条件代入公式 3.16, 可以得到如下推演:

$$\begin{aligned} & sign(M_{ik} - M_{ih}) \\ &= sign(a_{ik}^r \max \begin{bmatrix} p_{ik}^r \\ p_{ik}^g \\ p_{ik}^b \end{bmatrix} - a_{ih}^r \max \begin{bmatrix} p_{ih}^r \\ p_{ih}^g \\ p_{ih}^b \end{bmatrix}) \\ &= sign(\max(\mathbf{p}_{ik}) - \max(\mathbf{p}_{ih})) \end{aligned} \quad (3.18)$$

最终, M 通道的离散趋势变换被表示为:

$$T_{ij} = \sum_{k=h+1}^{j+N} \sum_{h=j}^{j+N-1} sign(\max(\mathbf{p}_{ik}) - \max(\mathbf{p}_{ih})) \quad (3.19)$$

公式(3.19)表达的 DDT 有如下性质:

- 在 M 通道的上升沿, T 达到最大值 $\frac{N(N+1)}{2}$.
- 在 M 通道的下降沿, T 到达最小值 $-\frac{N(N+1)}{2}$.
- 在非单调区间, T 的值位于最大值和最小值之间.

DTT 的另外一个重要作用是能够很好地检测出微弱线条。在实际采样的照片中, 在物体边缘或表面起伏坡度较大的地方, 由于投射线条被斜向拉扯, 线条变宽, 导致由投射线条所引起的亮度起伏变化较小, 使用传统的以亮度作为

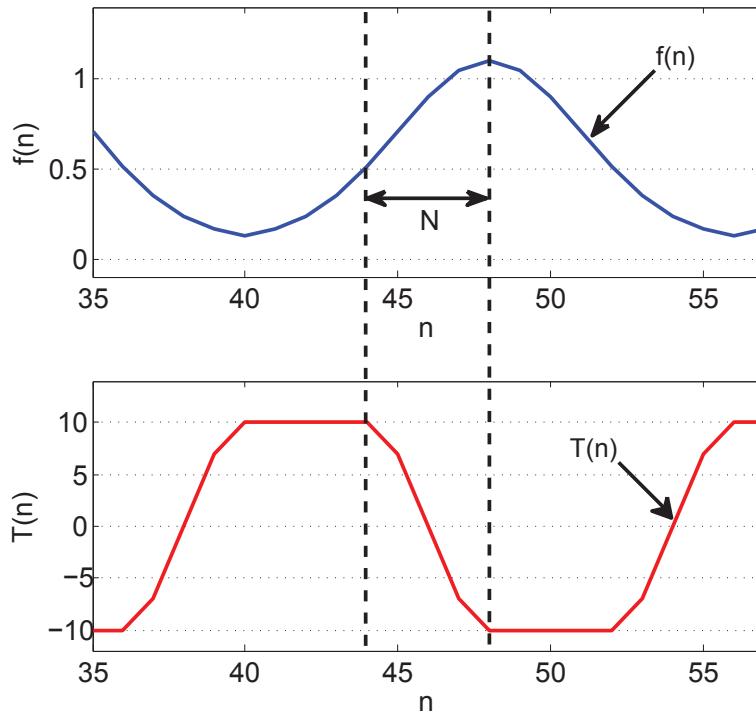


图 3.5: DTT. $T(n)$ 代表了 $f(n)$ 的局部趋势. $T(n)$ 的最大值意味着 $f(n)$ 位于上升沿, 而 $T(n)$ 达到最小值意味着 $f(n)$ 位于下降沿, 且 $T(n)$ 最大值的结束点与 $f(n)$ 的峰值点的距离恰为窗口大小 N .

判断标准方法很难在背景中检测到这些微弱线条。然而可想而知, 物体表面的 3D 信息在起伏较大区域最为集中, 所以如何将线条检测准确极为重要, 直接关系到 3D 重建的效果。DTT 之所以能够很好地处理这种情况是因为, DTT 仅关注亮度的变化趋势而不关注亮度的绝对值, 即通过变换, DTT 将线条的先递增、后递减趋势转换为最大值-最小值模式, 而最大值-最小值模式只与线条的变化趋势有关, 与线条的绝对亮度无关。图 3.7 对比了使用自适应阈值二值化方法和 DTT 方法分离的线条结果, 显然, 在坡度较大的物体边缘区域, DTT 方法检测除了绝大部分线条, 而自适应阈值的二值化方法的检测结果则存在不少漏检及断点。

3.4.3 峰值定位

在使用 DTT 的方法定位线条后, 还需要确定线条的峰值位置以用于三维点云的计算。峰值定位的方法有多种[7], 一些是精确到像素, 另一些是精确到亚像素, 现列举主要的方法如下。

- **最大值方法 Jordi[14]** 使用线条中具有最大值得像素位置作为峰值位置,

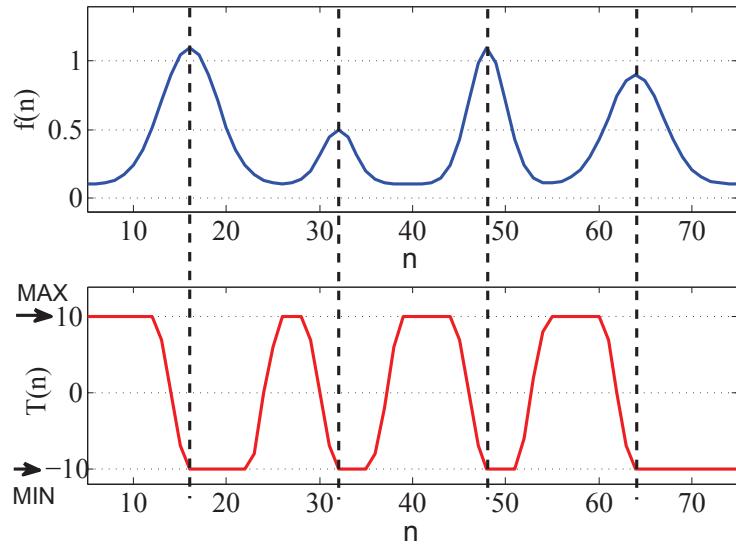


图 3.6: $f(n)$ 和 $T(n)$: 如图, $f(n)$ 有四个符合高斯分布的峰值, 各有不同的宽度和高度, 在对 $f(n)$ 实施 DTT 变换得到的 $T(n)$ 中, 四个峰值点则对应相同的最大值到最小值的变化模式, 寻找和定位这样的变化模式可以用于回溯寻找峰值位置.

为了便于对比结果, 将此方法简记为 **MAX**, 本方法只精确到像素。

- **中心点方法** Thomas等[11] 则使用线条的中心位置作为峰值位置, 记此方法为 **MID**, 本方法只精确到像素。
- **加权平均法** Zhiqiang等[18] 使用的峰值位置计算方法为, 在整个线条区域之中, 以各像素点的亮度为权重, 对各像素的位置求取加权平均值并作为峰值位置, 记此方法为 **WA**, 本方法将峰值位置计算到了亚像素。
- **概率法** Fechteler等[6] 则结合亮度、线条形状等因素, 构建抛物线模型并计算峰值位置, 记此方法为 **PA**, 本方法将峰值位置计算到了亚像素。

本文计算峰值的方法不同于以上, 其主要的步骤是, 在 M 通道和使用 DTT 方法确定的线条区域的基础上, 寻找线条区域中 M 通道值高于峰值一定比例的像素点, 并对这些点加权计算均值以作为峰值位置, 本文的方法记为 **MMWA**, 可表述为公式 3.20。本方法与以上方法的区别是, 首先公式 3.20 并不以亮度直接作为权重, 而是使用此像素点的 M 通道值, 另外, 此方法排除了线条外围更容易受噪声影响的点。

$$I_{estimated} = \frac{\sum I_i \cdot M_i}{\sum I_i}, \text{ for } M_i \geq \alpha M_{max} \quad (3.20)$$

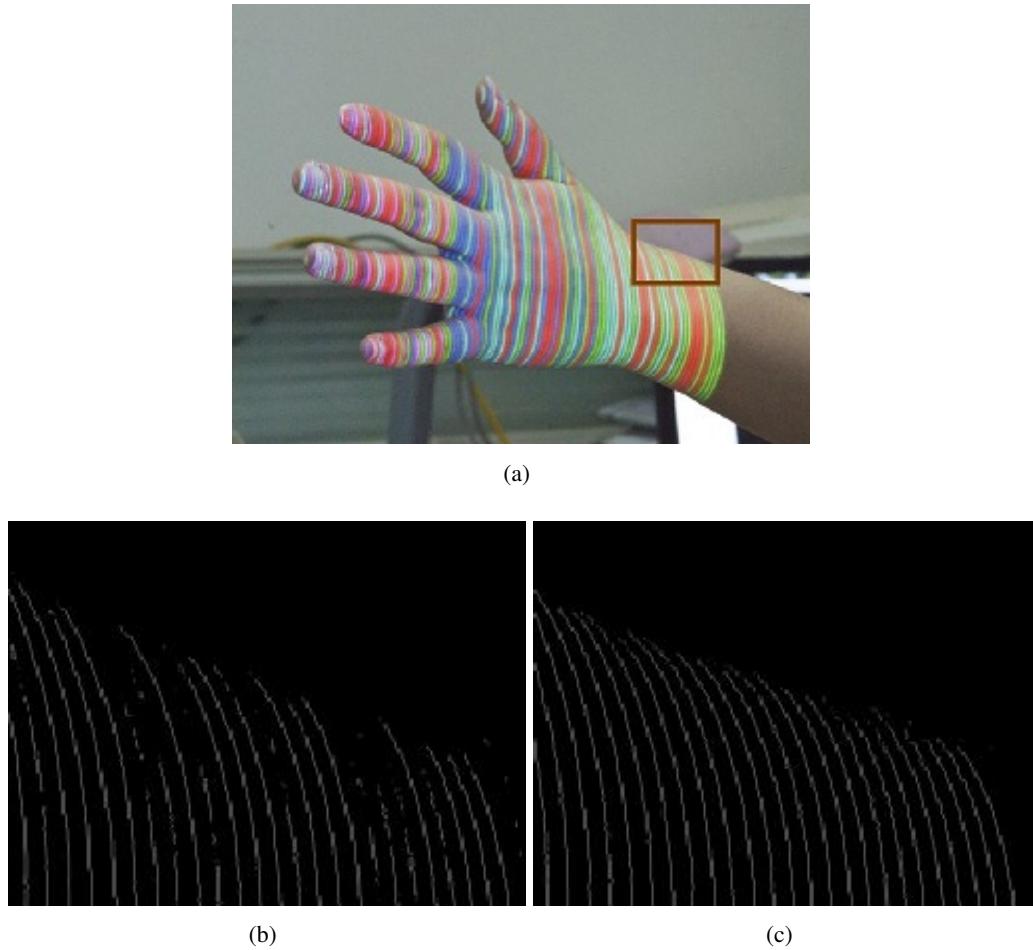


图 3.7: 线条检测的结果对比. (a) 原图像, (b) 自适应阈值二值化方法检测的线条, (c) 使用DTT检测到的线条.

在实际系统中, α 选取的值为 0.6。图 3.8 对比了各方法得到的峰值位置对比。可以看到最大值方法确定的峰值位置起伏最大, 本文方法检测的峰值则平顺得多。

3.4.4 颜色提取

线条颜色识别是特征识别的另一个重要部分, 颜色识别的准确率直接影响特征识别的准确率。从公式 3.4 可知, 原始图像即使经过 crosstalk 和相对反射率校正, 仍然保留有环境光和 R 通道反射率的影响。以简单的 RGB 颜色空间对颜色识别难以做到对上述干扰因素的免疫性。为了解决这个问题, 本文将已检测到线条位置的 R、G、B 值转化为色度, 并在 hsv 空间分别与 5 个所透射颜色的色度相比较, 若与某个投射颜色的色度相差最小且差值在阈值之内则判定线条颜色为此投射颜色。

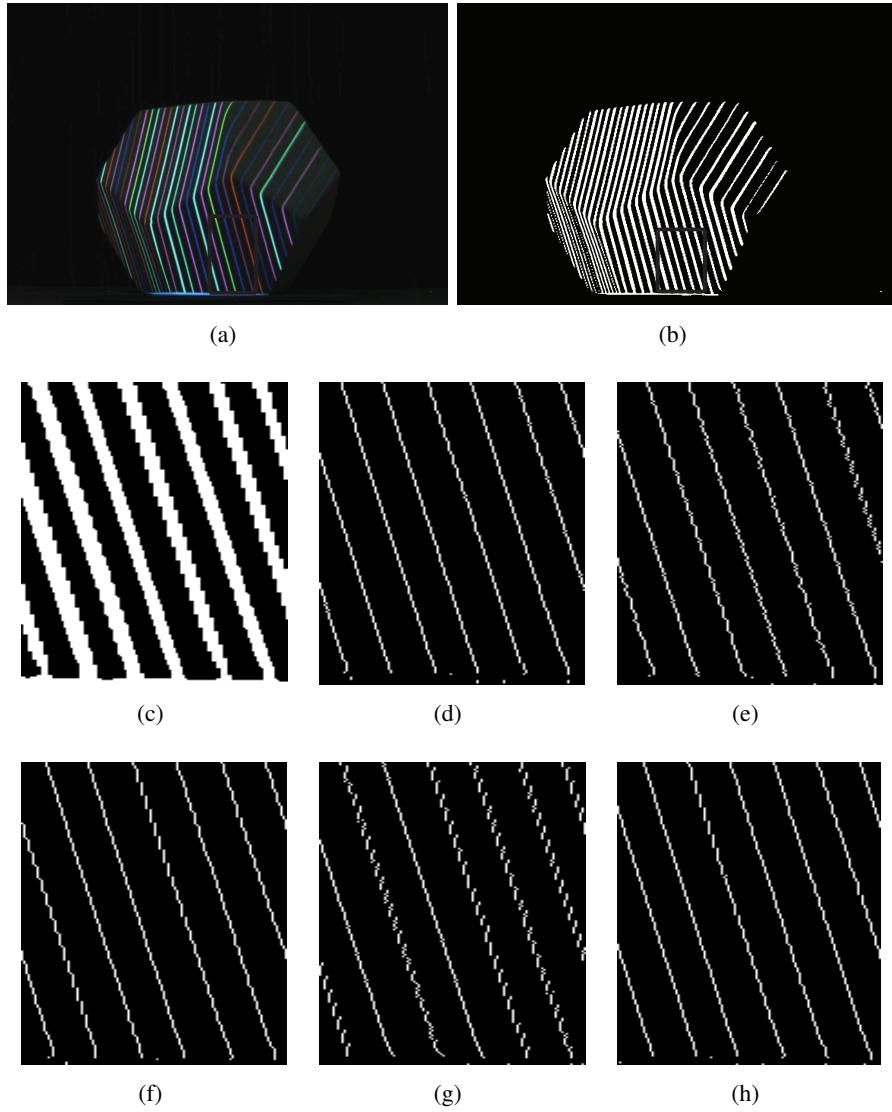


图 3.8: 以不同方法定位的峰值. (a)多面体. (b)线条区域. (c)放大的线条区域. (d)MMWA. (e)PA. (f)MID. (g)MAX. (h)WA

假设某像素点具有 (R,G,B) 三个通道值, 且满足 $R \geq G \geq B$, 那么这个点的色度表达为:

$$h = 60^\circ \cdot \frac{G - B}{R - B} \quad (3.21)$$

将公式 3.14 代入公式 3.21，那么得到：

$$\begin{aligned}
 h &= \frac{C^g - C^b}{C^r - C^b} \cdot \frac{\pi}{3} \\
 &= \frac{a^r(p^g + o^g) - a^r(p^b + o^b)}{a^r(p^r + o^r) - a^r(p^b + o^b)} \cdot \frac{\pi}{3} \\
 &= \frac{p^g - p^b}{p^r - p^b} \cdot \frac{\pi}{3}
 \end{aligned} \tag{3.22}$$

由公式 3.22 计算得到的色度与各干扰因素均无关，有效提高了颜色识别的准确性，图 3.4(a) 中食指部分的颜色识别效果图见图 3.9。

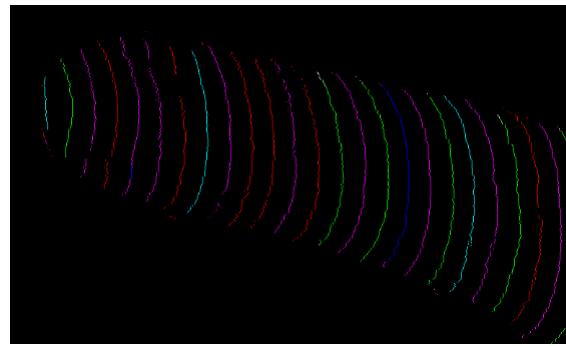


图 3.9：食指部分的线条颜色识别结果

3.4.5 建立检测特征及投射特征间的对应关系

完成线条检测、峰值定位和颜色识别等特征识别的步骤后，还要建立检测出特征及所投射特征间的对应关系以代入三角模型并计算三维点云。实际中这一步是采用动态规划匹配色度的方法实现，邹润等[23]则提出了概率模型以求解概率模型。

假设 $A_i, i \in [1, N]$ 代表投射线条序列的色度值， $B_j, j \in [1, M]$ 代表检测出线条序列的色度值，分两步进行 A_i, B_j 的匹配：

1. 建立 $C_{ij}(1 \in [1, N], j \in [1, M])$ 矩阵和 $P_{ij}(1 \in [1, N], j \in [1, M])$ 路径矩阵，建立方法为：

$$C_{ij} = \max(C_{i(j-1)}, C_{(i-1)j}, C_{(i-1)(j-1)} + cost_{ij}) \tag{3.23}$$

$$P_{ij} = \begin{cases} \nwarrow & C_{ij} = C_{(i-1)(j-1)} + cost_{ij} \\ \uparrow & C_{ij} = C_{(i-1)j} \\ \leftarrow & C_{ij} = C_{i(j-1)} \end{cases} \tag{3.24}$$

2. 从元素 P_{MN} 开始，沿着箭头跟踪 P 矩阵，其路径中的“ \nwarrow ”对应 A_i, B_j 的一个匹配，“ \uparrow ”表示 B_j 中缺失元素，即漏检线条，“ \leftarrow ”表示 B_j 中多出元素，即多检出了线条。

公式 3.23 和公式 3.24 中的 $cost_{ij}$ 是实现正确匹配的关键，为了鼓励连续的匹配，本文使用连续三根线条的匹配程度作为 cost 值， $cost_{ij}$ 可表示为：

$$cost_{ij} = match(A_{i-1}, B_{j-1}) + match(A_i, B_j) + match(A_{i+1}, B_{j+1}) \quad (3.25)$$

$$match(A_i, B_j) = \begin{cases} 10 & |A_i - B_j| \leq 20 \\ 0 & |A_i - B_j| \geq 40 \\ 10 \times \frac{40 - |A_i - B_j|}{20} & \text{else} \end{cases} \quad (3.26)$$

将 $match(A_i, B_j)$ 表示为图形为：

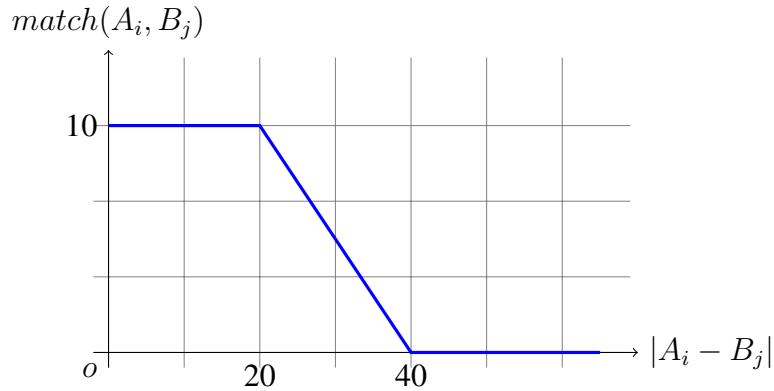


图 3.10: $match(A_i, B_j)$ 的图形表示

3.5 重建结果

本系统的输出结果是 3D 点云，由于采用了颜色校正、DTT 等改进方法，系统能够在非暗室的环境中正常工作，具有良好的对环境光抗干扰性能。图 3.11 展示了多面体、手部和脸部的图像采集环境和重建结果，即使存在环境光，系统也能够利用算法针对性地排除干扰，从而正确地检测特征并计算出三维点云。系统的另一个特点是分辨率高，图 3.11(b)(d) 分别有 85687 和 75371 个三维点。

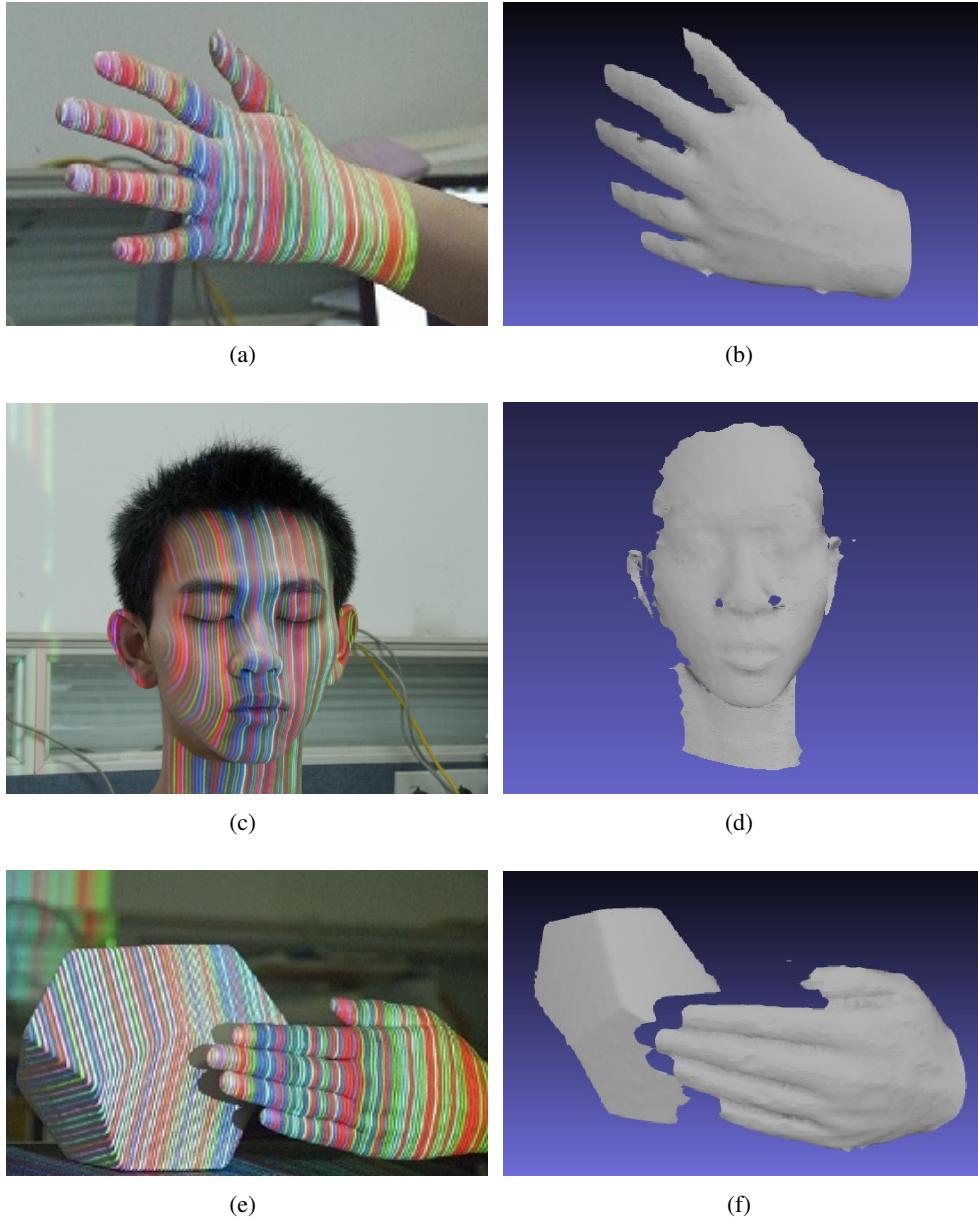


图 3.11: 重建结果: (a) 手, (c) 脸, (e) 手和多面体. (b) 手的重建表面, (d) 脸的重建表面(f) 手与多面体的重建表面.

3.6 本章总结

本章实现了一种基于 De Bruijn 序列编码颜色序列作为结构光的 3D 重建方法, 介绍了一般的流程并在关键处理步骤中, 如颜色校正、线条识别、峰值定位和颜色识别等方面进行了创新, 提出了 DTT 等概念用于特征识别, 使得系统在线条检测、定位和颜色识别等方面能够取得精确的结果, 同时也克服了结构

光方法易受环境光干扰的缺点，使重建系统即使在环境光下很好的工作，不需要暗室环境，很大程度上提升了重建系统的使用范围。

第四章 结构光为格雷码投影光栅的 3D 重建

格雷码投影光栅是时间编码方式的 3D 重建，相比于空间编码，时间编码方式具有分辨率高、测量误差小的特点。结构光为格雷码投影光栅的 3D 重建同样采用三角模型，系统架构与之前所述类似，此处不再重复。总体的处理步骤是：

1. 投影光栅并采集图像
2. 图像处理并提取有效区域
3. 解码并建立对应关系
4. 计算点的三维坐标

4.1 特征投影

通常 CCD 摄像机镜头上有增透膜，增透膜的作用是减少光在进入镜头时的反射，增加进入镜头的光强。因为人眼对偏绿的颜色敏感，大多数镜头都选择增强绿光的入射，因为采用厚度为绿光波长一半的增透膜。这样就造成由 CCD 摄像机获取图像时颜色的不均匀性，所以系统使用的结构光采用绿色。

经验证，绿色比红色光等其他光在相同条件下得到的点云噪点更少且正确点更多。

4.1.1 格雷码光栅

格雷码在任意两个相邻的数之间转换时，只有一个数位发生变化。这个特性使得其码值最多有一位被误判，且任意位被误判引起的误差只有一位。同时，由于格雷码较之二进制码器明暗光间隔要高出一倍，可有效减少明暗光线间的相互干扰。

以三幅二进制格雷码为例，如图 4.1 所示：

格雷码与十进制、二进制码的转换表如表 4.1。

格雷码的生成可采用递归方式实现。假如想要生成 n 位格雷码，则可先将 $n - 1$ 位格雷码的 2^{n-1} 个编码列出，而后再以从后向前的顺序，重新列出这 $n - 1$ 位格雷码，形成 2^n 个序列，每个序列的长度为 $n - 1$ 。最后在前



图 4.1: 格雷码投影光栅

表 4.1: 格雷码与实际值、二进制码的对应关系

十进制	格雷码	二进制码
0	000	000
1	001	001
2	011	010
3	010	011
4	110	100
5	111	101
6	101	110
7	100	111

2^{n-1} 个序列的最前补 0，在后 2^{n-1} 个序列的最前补 1，这样就完成了 n 位格雷码的生成。分别设定格雷码中 0 和 1 代表的颜色，并从高位向低位依次递进，形成 n 幅 2^n 位的格雷码编码的光栅图案并投射至被扫描物体表面。拍摄到加照格雷码光栅的物体在图 4.2 中体现，从图中我们看到，后三幅的格雷码光栅亮度转换频繁，尤其是最后一幅，亮区和暗区只占 1 个像素，这导致在实际处理中很难准确区分亮区和暗区。

4.1.2 位移码图

在实际的系统中，除了格雷码光栅还使用了位移码光栅作为补充。这是因为，当格雷码转换频率较高时，明暗光线间隔很近，其相互间干扰已不可避免和忽略，导致在识别光亮区和黑暗区的处理中引入误差。为了解决这个问题，实际的系统只使用明暗间隔较大的若干幅格雷码光栅，由此引起的分辨率缺失采用位移码标注。如在本文中采用的格雷码共 10 幅，理论上最高分辨率为 1024，但在变化频率最高的 3 幅图像中，光栅的明暗区域交替频繁、互相影响，

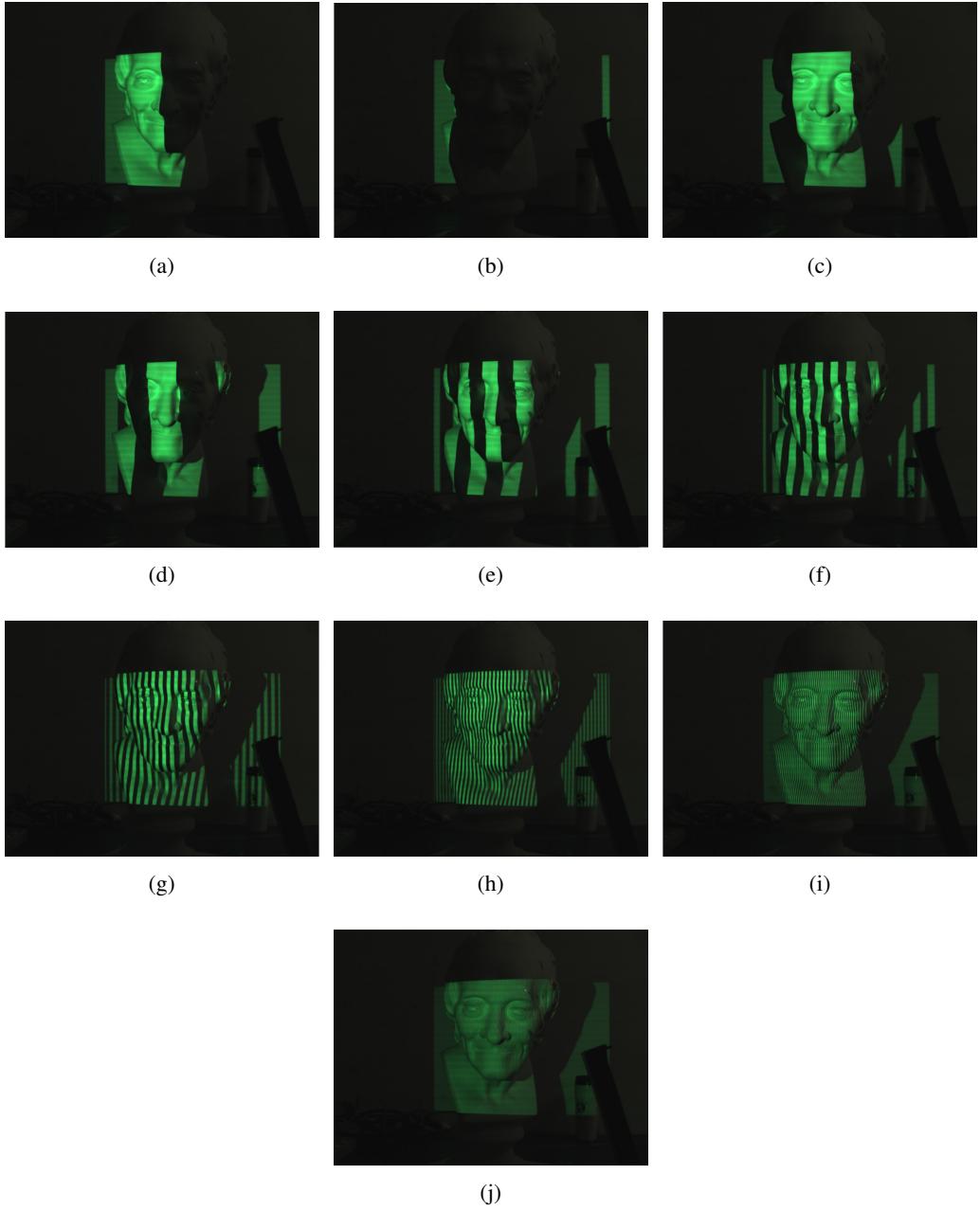


图 4.2: 投射了格雷码光栅的物体

导致很难准确定位边界。为了解决这个问题，实际的系统中只投射前 7 幅变化较小的格雷码光栅。与 10 幅格雷码光栅相比，7 幅格雷码光栅的最小变化单元为 8 像素，分辨率下降为原来的八分之一，为 128。为了达到相同分辨率，此时在格雷码光栅每个最小的变化单元之内分时 8 次，每次投射 1 条位移码，8 幅位移码使得分辨率升为 1024。宽度为 8 的位移码如图 4.3 所示。

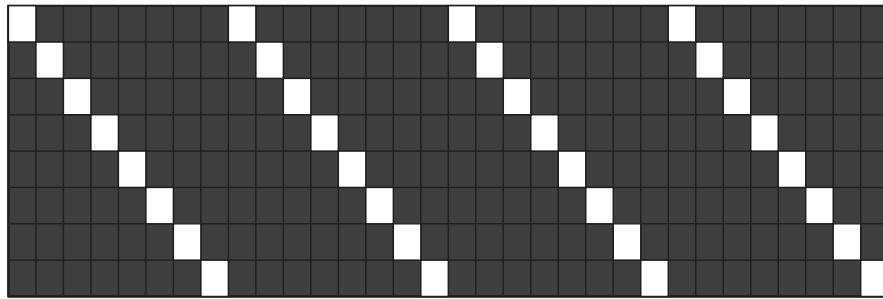


图 4.3: 位移码图

4.2 编码图像的处理

4.2.1 有效区域的选择

实际上，以结构光为基础的 3 维重建可能会存在测量死角，即结构光无法照射到或照相机无法捕捉的区域，其示意图如图 4.4。

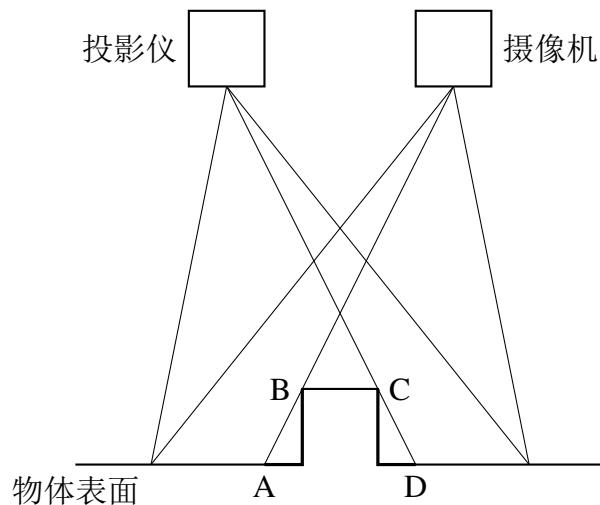


图 4.4: 测量死角的示意图

在图 4.4 中，物体表面的 AB 和 CD 部分为测量死角，其中 AB 部分是因为摄像机无法捕捉到，CD 部分表面则是因为投影仪无法照射到。

死角的存在会在解码图像引入干扰，应事先予以标记并在解码和计算点云时排除此部分。方法在投射格雷码光栅之前投射全亮的投射光，并通过二值化的方法检测出有效区域。图 4.5 中，(a)是在无任何投影下的拍摄图像，(b)是投射全屏光时拍摄的图像，(c)是提取的有效区域。

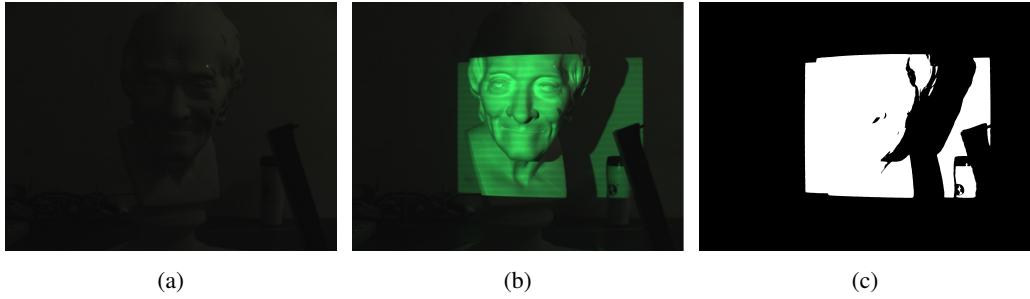


图 4.5: 有效区域提取

4.2.2 解码位移图码光栅

在解析格雷码光栅之前，本文先解码位移图码。在加照格雷码光栅之前，需要让摄像机记录在无任何光照条件下的环境图像，记为 $I_{background}$ 。提取位移图码的处理步骤可表示为：

1. 计算 $I_{sub} = I_{move} - I_{background}$
2. 按照原投射光中各通道的比例将 I_{sub} 中的 R, G, B 通道合并为单一通道，记为 $I_{channel}$

假设投射光的颜色为(R, G, B)，那么：

$$I_{channel} = \frac{R}{R+G+B} \cdot I_{sub}^R + \frac{G}{R+G+B} \cdot I_{sub}^G + \frac{B}{R+G+B} \cdot I_{sub}^B \quad (4.1)$$

3. 用类似颜色条纹序列中的峰值定位方法定位 $I_{channel}$ 中的峰值

位移图码的解析对点云的精确程度有很大影响，本处采用与颜色条纹序列中的峰值定位方法一样的算法可达到亚像素的精度。另一方面，接下来的格雷码光栅的解析则需要用到位移图码信息。

4.2.3 解码格雷码光栅

在 t 时刻拍摄到的加照格雷码光栅的图像记为 I_{gray} ，那么处理方式步骤可列为：

1. 计算 $I_{sub} = I_{gray} - I_{background}$
2. 合并 I_{sub} 的R, G, B通道为单一通道，记为 $I_{channel}$

假设投射光的颜色为(R, G, B)，那么：

$$I_{channel} = \frac{R}{R+G+B} \cdot I_{sub}^R + \frac{G}{R+G+B} \cdot I_{sub}^G + \frac{B}{R+G+B} \cdot I_{sub}^B \quad (4.2)$$

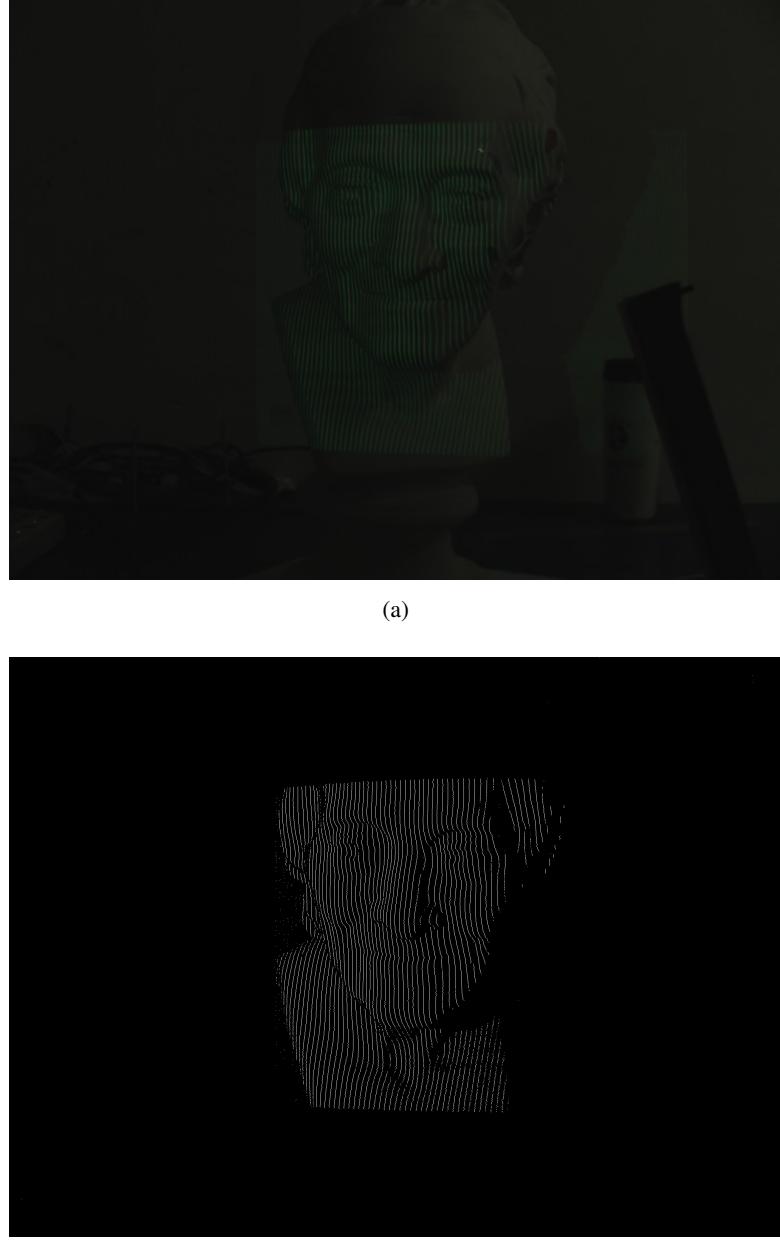


图 4.6: 位移码图像的处理(a),拍摄到的源图像(b),检测出的线条

图 4.7展示了编码图像的处理过程, (a)是摄像机拍摄到的图像, (b)是将(a)减去背景图像后从 RGB 转换为灰度图后得到的图像。

实际上, 二值化图像是准确解码的关键。将一幅图像二值化的方法有多种, 大致可分为全局阈值法和局部阈值法, 而阈值的选择至关重要。全局阈值法是指, 选取单一的阈值, 适用于整幅图像并二值化, 典型的算法包括直方图变换法, OTSU 法(大律法)。全局阈值法虽然简单但却不能表现图像的细节。而局部

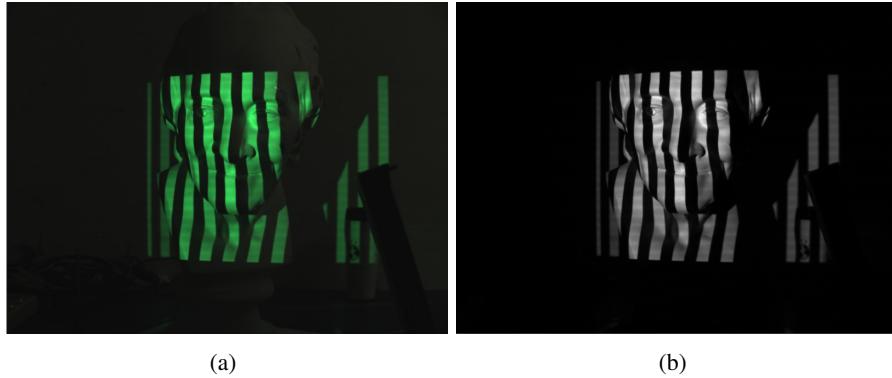


图 4.7: (a),一幅加照了格雷码光栅的物体(b),合并为单通道

阈值法是指，将图像按照某种规则划分为若干部分，在各部分之内选取阈值并适用于此部分以二值化。相比之下，局部阈值法虽然没有从根本上解决这个缺陷，但较全局阈值法能将细节表现地更好。

另一方面，在本系统中，稍加推理便可知道，格雷码光栅的边缘应当位于第 8 条位移码和相邻格雷码光栅变化单元的第 1 条位移码之间，也就是说，检测得到的位移码位置对格雷码光栅的边缘位置添加了约束和指导，光栅边缘的确定需要考虑光栅亮度的变化，也需要考虑位移码的约束条件。本文采用概率的方式综合考虑这两个因素并最终确定边缘位置，具体可表述为：

1. 使用局部阈值法二值化 $I_{channel}$ ，并初步确定边缘的大致位置，见图 4.8(b)。假设被确定的边缘位置为 t ，那么按照高斯衰减，与 t 相邻的位置 i 可能为边缘的概率可表示为：

$$p_t(i) = e^{\left(\frac{i-t}{c_1}\right)^2} \quad (4.3)$$

此处的 c_1 选择为 1。

2. 确定 i 附近第 8 条位移码和相邻格雷码光栅变化单元的第 1 条位移码的位置，见图 4.8(a)，分别记为 l 和 h ， l 和 h 为边缘位置提供了限制因素，在此约束条件下，位置 i 可能为边缘的概率可表示为：

$$p_c(i) = e^{\left(\frac{2i-l-h}{2c_2}\right)^2} \quad (4.4)$$

此处的 c_2 设定为 $\frac{1}{h-l}$ ，即 l 和 h 相距越远 $p_c(i)$ 越平缓，反之则越陡峭。

3. 结合上述两个因素，位置 i 最终可能为真正边界的概率为：

$$p(i) = p_t(i) \cdot p_c(i) = e^{\left(\frac{i-t}{c_1}\right)^2} \cdot e^{\left(\frac{2i-l-h}{2c_2}\right)^2} \quad (4.5)$$

最终，取 $p(i)$ 最大值处的 i 作为格雷码光栅的最终边缘位置。

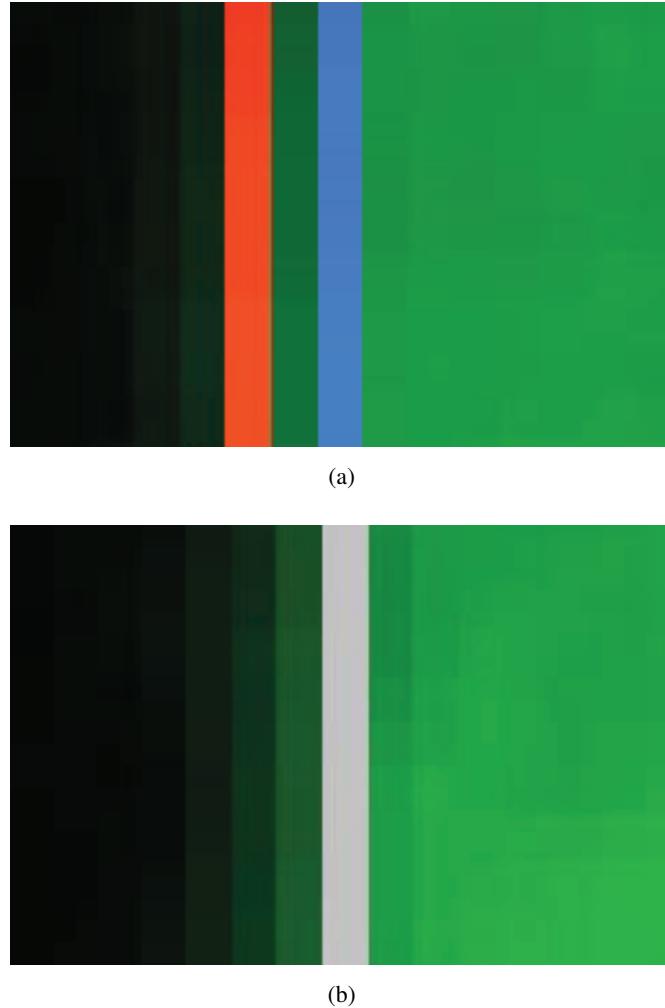


图 4.8: (a)两条相邻的边界位移码线条(b)阈值分割得到的边缘线

图 4.8(a)是两条相邻的边界位移码线条，理想情况下格雷码光栅的边缘应位于线条之间，按照式 4.4 将其转化为图 4.9(a)。图 4.8(b)是二值化得到的边界位置，按照式 4.3 将其转化为图 4.9(b)，最终按照式 4.5 合并得到图 4.9(c)并取其最大位置为格雷码光栅的边界，效果见图 4.10。

4.3 重建结果

本文从多个角度对模型进行了表面重建，得到了不同角度的 3D 点云数据，重建的结果如图。

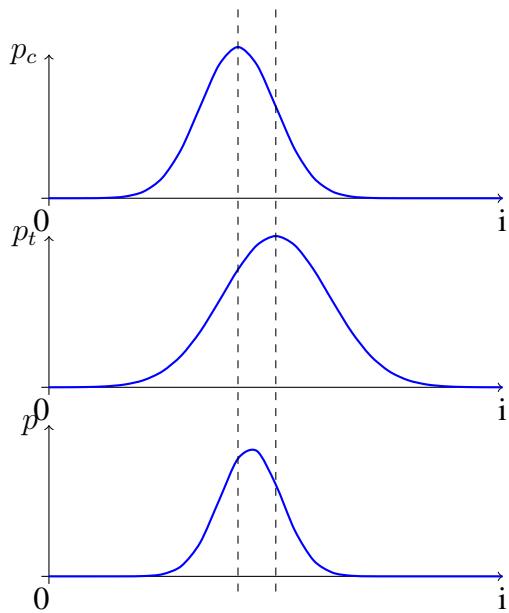


图 4.9: (a), p_c (b), p_t (c), p

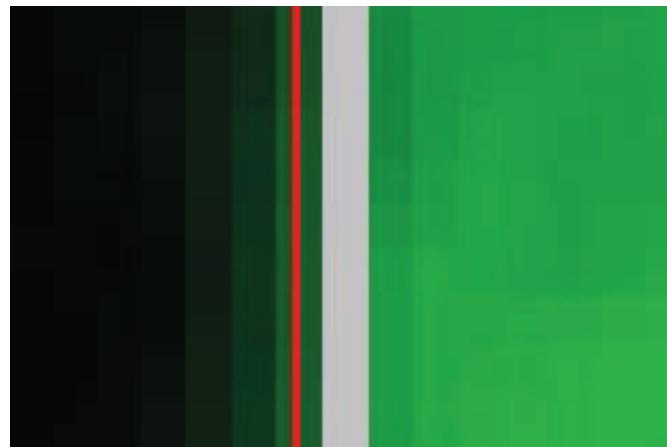


图 4.10: 最终确定的格雷码光栅边界

4.4 本章总结

本章介绍了基于格雷码光栅的 3D 重建方法。系统结合格雷码光栅和相移线条，并在格雷码光栅的边缘定位等方法进行了创新，提出了基于概率的边缘识别方法，取得了理想的重建效果。

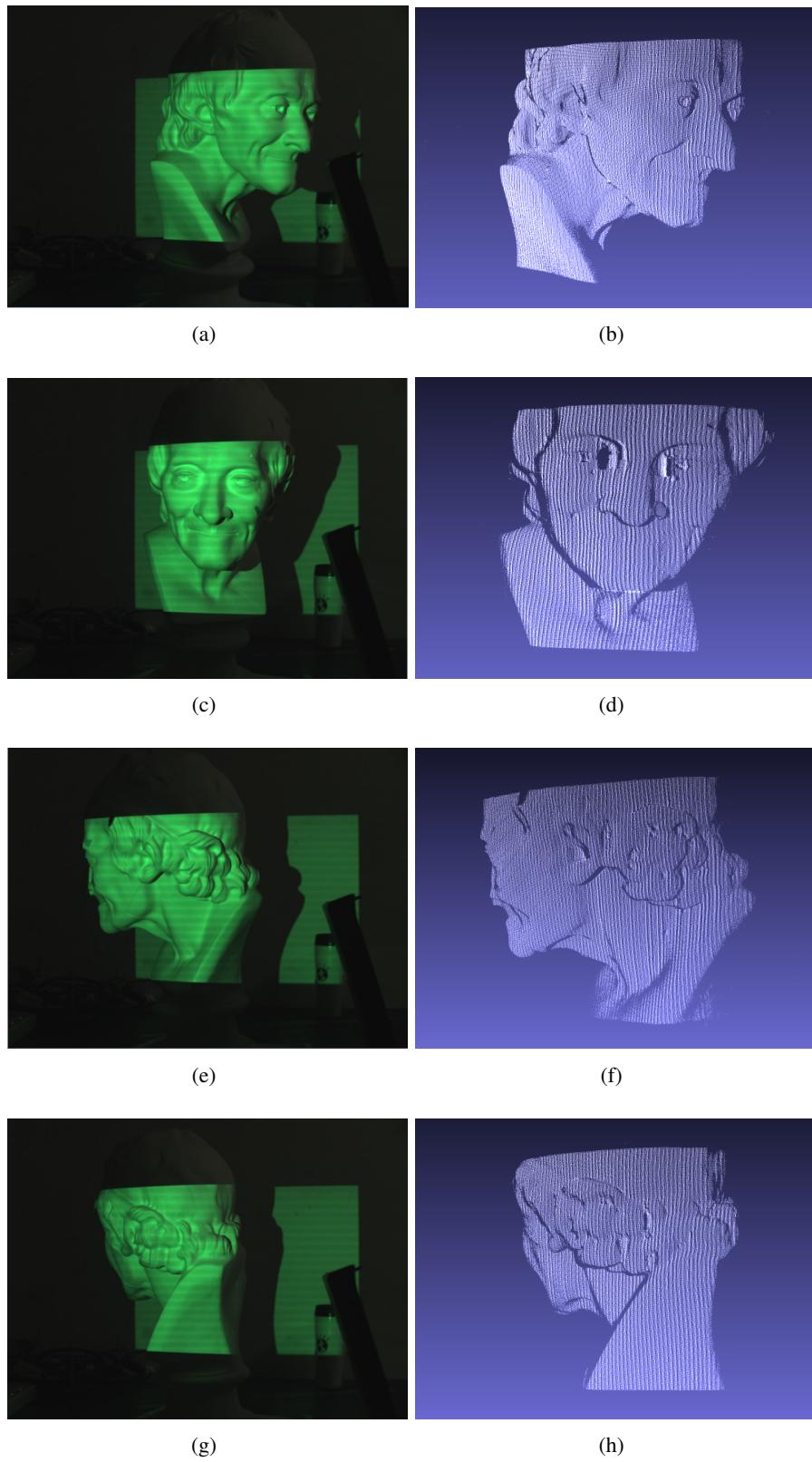


图 4.11: 重建结果，左侧图为场景图，右侧为重建的点云

第五章 总结和展望

本文首先对 3D 重建的原理和结构光技术进行了研究分析和总结，在综合考虑重建所需时间和重建精度等因素后选择基于格雷码光栅的重建方法和基于 De Bruijn 序列编码条纹的结构光方法实现了 3D 重建前者属于时间编码而后者属于空间编码。在实际使用中可综合考虑，灵活选择使用方法。

5.1 研究总结

本文主要的研究内容包括：摄像机-投影仪系统定标、结构光的选择、结构光特征识别等方法，提出了 M 通道、离散趋势变换和格雷码光栅边缘识别等概念和方法，具体总结为如下：

5.1.1 摄像机-投影仪系统定标

对摄像机-投影仪系统进行定标是实现精确 3D 重建的前提。本文分析了摄像机的成像模型和畸变原因，推导了坐标系转换关系，并使用棋盘格作为定标板而确立了摄像机定标过程。本文将投影仪视作逆成像系统，并利用摄像机定标参数实现了投影仪的定标。

5.1.2 结构光选择和设计

本文选择两类结构光实现 3D 重建：时间编码的格雷码光栅和空间编码的 De Bruijn 序列编码颜色条纹。在 De Bruijn 序列编码条纹的颜色选择中，本文选择在色度空间中分布较为离散的 5 种颜色作为编码颜色以便线条检测和颜色识别。在格雷码光栅中又结合相移条纹的方法实现了高分辨率的 3D 重建。

5.1.3 特征识别

在特征识别方面本文作了较多创新。对颜色条纹结构光方法，本文提出了颜色校正方法、颜色线条检测方法和颜色识别等方法。在颜色校正中，本文利用 EMD 实现了反射率校正，避免了物体表面颜色对颜色线条的影响。在线条检测中，本文提出了 M 通道以融合 R、G、B 通道，提出了离散趋势变换(DTT)实现了线条的识别。在对应点建立之中，本文使用动态规划方法，并设计了特殊的 cost 函数，实现了被检测出特征与投射特征之间的对应关系建立。最后，本文将线条的颜色识别放在色度空间之中进行，避免了环境光的干扰。

在基于格雷码光栅的系统中，本文提出了基于自适应阈值和概率的方法确定格雷码光栅的边缘，实现了很好的重建结果。

5.2 展望

本文利用两种方法，实现了具有适用范围广、鲁棒性高的 3D 重建系统。在后续的研究中，在以下的方面还可以继续深入探究。

5.2.1 3D点云的处理

本系统的直接输出是物体表面的 3D 点云，在后续的研究之中，还可以通过 Delaunay 三角化的方法将点云转化为表面。更进一步地，可以将物体表面的实际纹理覆盖于包围之上，使得重建结果更加鲜明和立体化。

另一方面，本系统一次 3D 重建输出的结果是从某一个角度观测到的物体表面，要想得到 360° 的全景点云则需要从不同角度重建并拼接。目前可以使用多台设备同时从不同角度对物体进行重建，也可以利用单设备和云台转动物体并多次重建并实现拼接。前者速度快但所需成本较高，后者成本低但需进行多次重建。点云拼接可利用 Besl 和 Mckay[2]提出的 ICP 算法实现。

5.2.2 实现实时重建

本文所使用的空间编码的结构光只需要拍摄一幅图像即可实现 3D 重建，可考虑使用并行处理的方式加速处理比较耗时的环节，如离散趋势变换、动态规划、点云计算，并最终实现实时的 3D 重建。实时的 3D 重建对移动物体具有很大意义。

参考文献

- [1] Besl, P.J., 1988. Active, optical range imaging sensors. *Machine vision and applications* 1, 127–152.
- [2] Besl, P.J., McKay, N.D., 1992. Method for registration of 3-d shapes, in: Robotics-DL tentative, International Society for Optics and Photonics. pp. 586–606.
- [3] Bradski, G., Kaehler, A., 2008. *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Incorporated.
- [4] Brown, D.C., 1966. Decentering distortion of lenses , 444–462.
- [5] Caspi, D., Kiryati, N., Shamir, J., 1998. Range imaging with adaptive color structured light. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 20, 470–480.
- [6] Fechteler, P., Eisert, P., 2009. Adaptive colour classification for structured light systems. *Computer Vision, IET* 3, 49 –59.
- [7] Fisher, R., Naidu, D., 1996. A comparison of algorithms for subpixel peak detection, in: *Image Technology*. Springer, pp. 385–404.
- [8] Fryer, J.G., Brown, D.C., 1986. Lens distortion for close-range photogrammetry. *Photogrammetric engineering and remote sensing* 52, 51–58.
- [9] Hartley, R., Zisserman, A., 2000. *Multiple view geometry in computer vision*. volume 2. Cambridge Univ Press.
- [10] Heikkila, J., Silven, O., 1997. A four-step camera calibration procedure with implicit image correction, in: *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, IEEE. pp. 1106–1112.
- [11] Koninckx, T., Geys, I., Jaeggli, T., Van Gool, L., 2004. A graph cut based adaptive structured light approach for real-time range acquisition, in: *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pp. 413 – 421.

- [12] Mishra, D.K., Chandwani, M., 1993. Ccd camera based automatic dimension measurement and reproduction of 3-d objects, in: TENCON '93. Proceedings. Computer, Communication, Control and Power Engineering.1993 IEEE Region 10 Conference on, pp. 530–533 vol.4.
- [13] Onoz, B., Bayazit, M., 2003. The power of statistical tests for trend detection. *Turkish Journal of Engineering and Environmental Sciences* 27, 247–251.
- [14] Page, J., Salvi, J., Collewet, C., Forest, J., 2005. Optimised de bruijn patterns for one-shot shape acquisition. *Image and Vision Computing* 23, 707 – 720.
- [15] Rubner, Y., Tomasi, C., Guibas, L.J., 2000. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision* 40, 99–121.
- [16] Salvi, J., Pages, J., Batlle, J., 2004. Pattern codification strategies in structured light systems. *Pattern Recognition* 37, 827–849.
- [17] Sansoni, G., Trebeschi, M., Docchio, F., 2009. State-of-the-art and applications of 3d imaging sensors in industry, cultural heritage, medicine, and criminal investigation. *Sensors* 9, 568–601.
- [18] Wei, Z., Jiang, S., Ji, X., 2010. 3-d reconstruction based on structured light of composite colorful stripes, in: Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on, pp. V6–123 –V6–125.
- [19] Will, P.M., Pennington, K.S., 1972. Grid coding: A preprocessing technique for robot and machine vision. *Artificial Intelligence* 2, 319–329.
- [20] Zhang, Z., 2000. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22, 1330–1334.
- [21] Zhao, D., Zhou, Y., Yu, Y., Du, S., 2011. A novel peak detection method of structured light stripes for 3d reconstruction, in: Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2011 International Conference on, IEEE. pp. 43–46.
- [22] Zhou, Y., Zhao, D., Yu, Y., Yuan, J., Du, S., 2012. Adaptive color calibration based one-shot structured light system. *Sensors* 12, 10947–10963.

- [23] Zou, R., Zhou, Y., Yu, Y., Du, S., 2012. A novel locally adaptive dynamic programming approach for color structured light system, in: Advances in Visual Computing. Springer, pp. 372–381.

致 谢

本文的研究内容是在我的导师都思丹老师、周余老师、于耀老师的悉心指导和帮助下完成的。在整个研究生学习期间，我的导师为我提供了卓越、轻松友好的学习研究环境，并在思想、研究等方面给予了很多的关怀和指引，使得我对生活和人生的认识和理解得到极大提升。导师严谨的治学态度和颇具前瞻性的眼光对我产生了深刻影响，是我在今后工作生活中的榜样，谨向导师致以崇高敬意！

周余老师、于耀老师很关心我的学习和科研，每次与周余老师、于耀老师的讨论都及其富有创意并都能够获益很深。记忆中最深的是周余老师、于耀老师曾带领我去上海参观 Intel 嵌入式设计大赛，在大赛上看到了全国各地大学生创作的作品，极大地开阔了眼界和见识。周余老师儒雅富有思想，于耀老师博学刻苦，两位老师对我帮助很大，感谢你们！

感谢盛玉宝老师、王育昕老师的关心帮助，感谢王自强老师、彭成磊老师等我的授业恩师！

还要感谢任结、宋宇、王研的相伴，感谢刑天威、陆桂亮、邹润等师弟师妹，正是你们使得实验室充满朝气和欢乐！感谢王大千、黄锋、杨瑜、许斌锋等师兄，你们同样对我帮助很大！