



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信息采集是近年来计算机视觉领域的热点问题，基于结构光的三角测量方法是3D信息采集的重要途径和有效方法。本文建设了可用的具有高可靠性的3D重建系统，并从时间序列、one shot两个角度讨论了结构光重建的途径，通过不同的结构光颜色编码方式分别予以实现，取得了很好的重建效果。

关键词： 3D重建; 结构光; 编码; 三角测量

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 one-shot color structured light system, the color of stripe patterns are usually distorted with respect to color crosstalk, ambient light and albedo of the scanned objects, leading to mismatch in the correspondence of color stripes between the projected and captured images. In this paper, an adaptive color calibration as well as a Discrete Trend Transform algorithm are presented to achieve high-resolution 3D reconstructions. The adaptive color calibration, according to the relative albedo in RGB channels, can improve the accuracy of labeling stripe by alleviating the effect of albedo and ambient light while decoding the color. Furthermore, the Discrete Trend Transform in M channel makes it effective to detect weak stripes due to the uneven surface or reflectance characteristics of the scanned objects. With this approach, the presented system is suitable for scanning the moving objects and generating high-resolution 3D reconstructions without the need of dark laboratory environments.

Keywords: three-dimensional image acquisition; adaptive color calibration; discrete trend transform; structured light.

目 录

| | |
|----------------------------------|-----------|
| 目录 | iii |
| 第一章 绪论 | 1 |
| 1.1 引言 | 1 |
| 1.2 3D重建的技术综述 | 1 |
| 1.2.1 被动式非接触重建技术 | 1 |
| 1.2.2 主动式非接触重建技术 | 1 |
| 1.3 结构光技术综述 | 1 |
| 1.3.1 结构光技术的原理 | 1 |
| 1.3.2 结构光技术的发展 | 1 |
| 1.4 本文章节组织形式 | 1 |
| 第二章 投影仪-摄像机系统校正 | 2 |
| 2.1 摄像机模型及定标 | 2 |
| 2.1.1 坐标系的定义 | 2 |
| 2.1.2 世界坐标系到相机坐标系的转换 | 3 |
| 2.1.3 相机坐标系到成像平面参考系的转换 | 4 |
| 2.1.4 摄像机畸变 | 5 |
| 2.1.5 摄像机定标 | 6 |
| 2.2 投影仪定标 | 7 |
| 2.3 三维信息的计算 | 9 |
| 第三章 结构光为颜色条纹序列的3D重建 | 10 |
| 3.1 系统 | 10 |
| 3.1.1 架构和流程 | 10 |
| 3.1.2 特征投影 | 10 |
| 3.2 crosstalk校正 | 11 |
| 3.3 相对反射率校正 | 12 |

| | | |
|-------------------|-------------------------------|-----------|
| 3.3.1 | 相对反射率对观测的影响 | 12 |
| 3.3.2 | 相对反射率的计算 | 13 |
| 3.3.3 | 反射率校正的结果 | 14 |
| 3.4 | 特征识别 | 15 |
| 3.4.1 | M 通道定义 | 15 |
| 3.4.2 | 离散趋势变换 | 16 |
| 3.4.3 | 峰值定位 | 19 |
| 3.4.4 | 颜色提取 | 20 |
| 3.4.5 | 建立检测特征及投射特征间的对应关系 | 20 |
| 3.5 | 重建结果 | 22 |
| 第四章 | 结构光为格雷码投影光栅的3D重建 | 24 |
| 4.1 | 特征投影 | 24 |
| 4.1.1 | 格雷码光栅 | 24 |
| 4.1.2 | 位移码图 | 25 |
| 4.2 | 编码图像的处理 | 27 |
| 4.2.1 | 有效区域的选择 | 27 |
| 4.2.2 | 解码格雷码光栅 | 28 |
| 4.2.3 | 解码位移图码光栅 | 29 |
| 4.3 | 重建结果 | 29 |
| 第五章 | 结束语 | 32 |
| 5.1 | 总结 | 32 |
| 5.2 | 创新点 | 32 |
| 5.3 | 展望 | 32 |
| 参考文献 | 33 | |
| 致谢 | 35 | |

第一章 绪论

1.1 引言

1.2 3D重建的技术综述

1.2.1 被动式非接触重建技术

1.2.2 主动式非接触重建技术

1.3 结构光技术综述

1.3.1 结构光技术的原理

1.3.2 结构光技术的发展

1.4 本文章节组织形式

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，容易获取并安装。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。

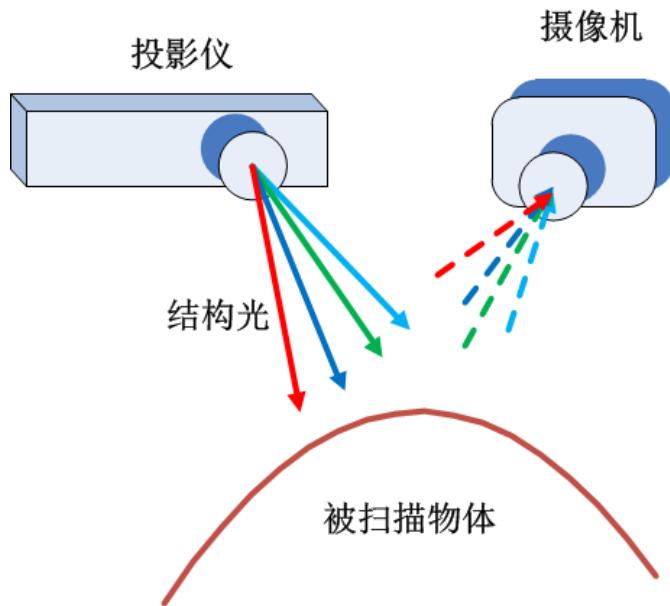


图 2.1: 系统结构图

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

讲世界坐标系的点 $Q_w = [X_w \ Y_w \ Z_w]^T$ 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轴平行并且反向一致
2. 旋转操作 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[6]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：

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

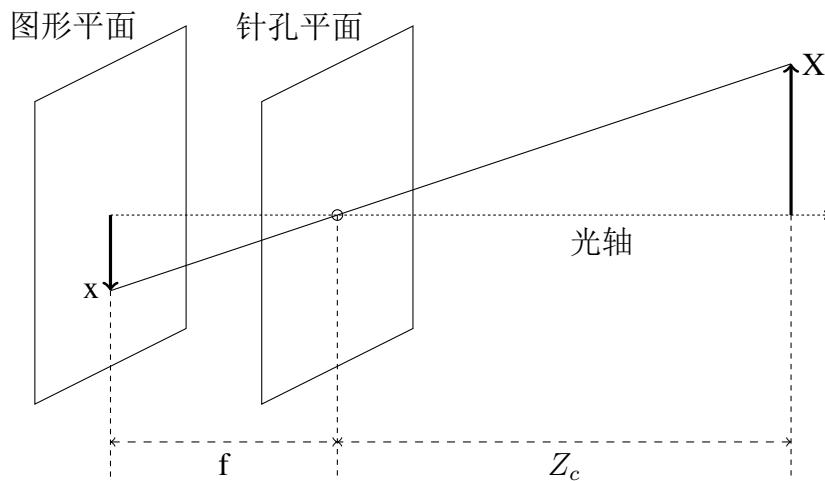


图 2.2: 小孔成像模型

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[1]。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：

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

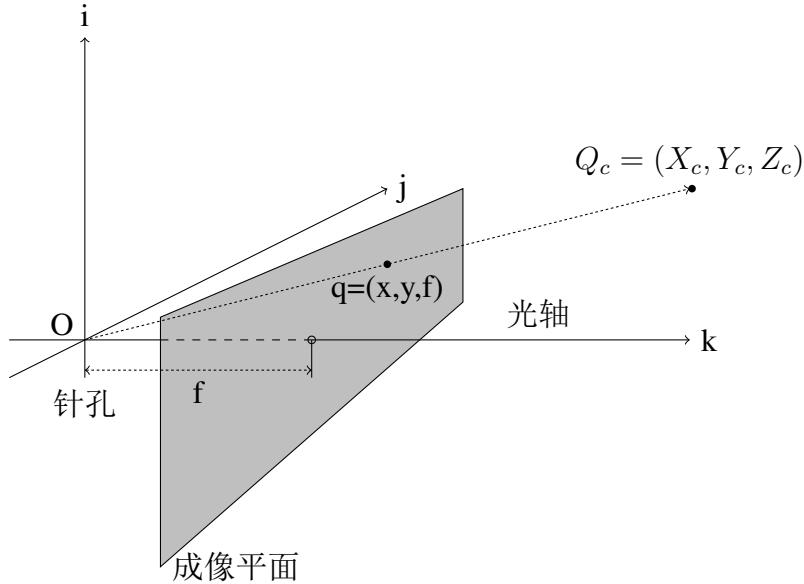


图 2.3: 针孔模型的等价变换

其中, $q = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $Q_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$, $M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ 。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[7]。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[5]。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, 随着感光元件的边缘移动, 畸变程度越显著, 参见图 2.4。在实际中, 这样的畸变比较小, 可以利用光学中心($r = 0$)位置周围的泰勒级数展开来描述, 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。

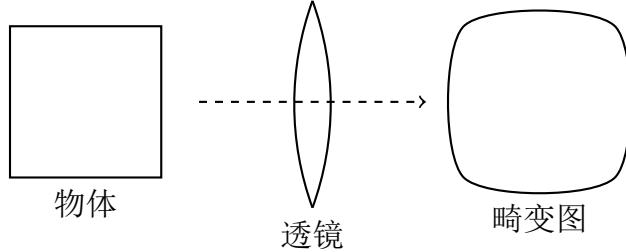


图 2.4: 筒形畸变

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[2]：

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

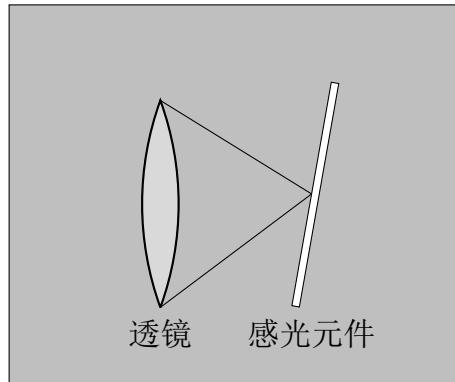


图 2.5: 摄像机结构，感光元件与透镜不平行

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。

本文的定标板如图 2.6(a)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

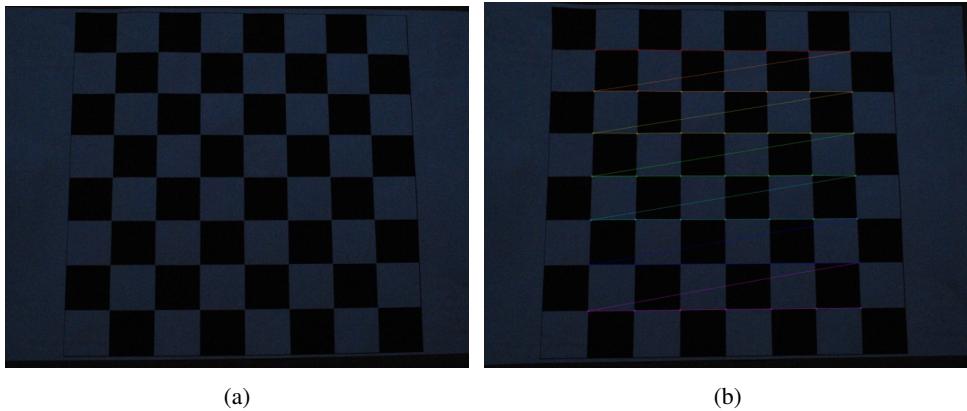


图 2.6: (a),棋盘格作为定标板(b), 定标板上找到的角点

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 2.6(b)中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的cvCalibrateCamera2函数实现的。

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 );
```

cvCalibrateCamera2函数的输出包括定标点的世界坐标和图像坐标，输出则包含内参数矩阵 M 、畸变系数(k_1, k_2, p_1, p_2)、外参数中的旋转矩阵 R 和平移矩阵 T 等。函数cvInitUndistortMap和函数cvRemap可用于对径向畸变和切向畸变的校正。内参数矩阵和外参数矩阵描述了世界坐标系和成像平面参考系之间的转换关系。

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。

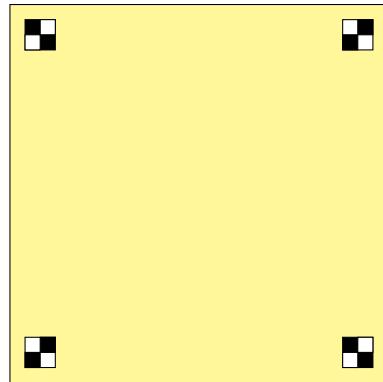


图 2.7: 投影仪校正的定标板

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：

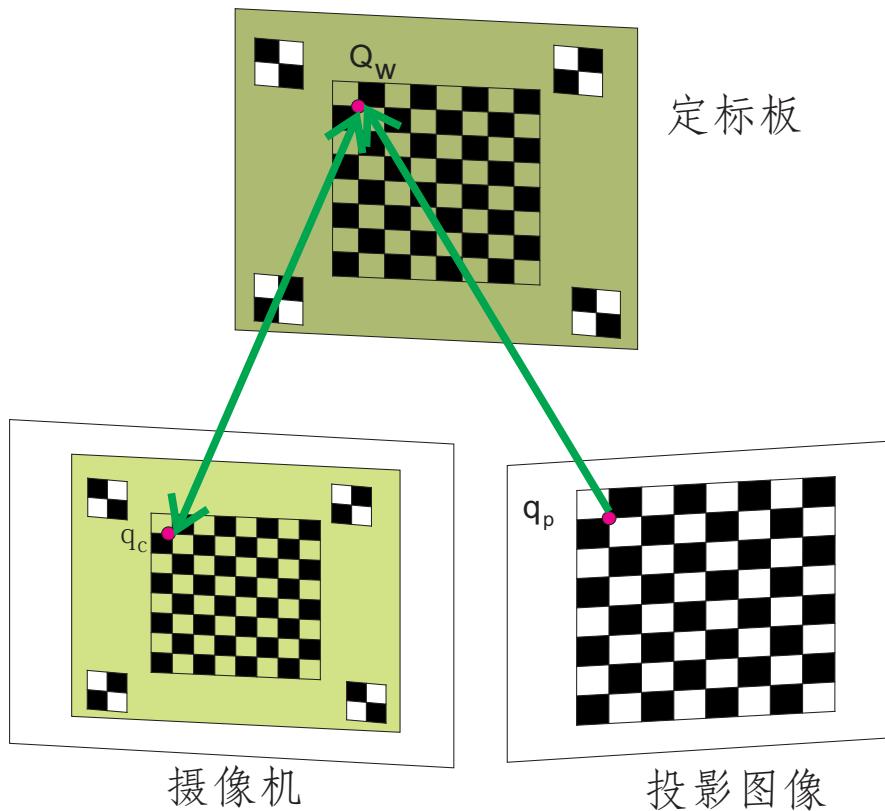


图 2.8: 投影仪投射定标图像后的定标板

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。

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

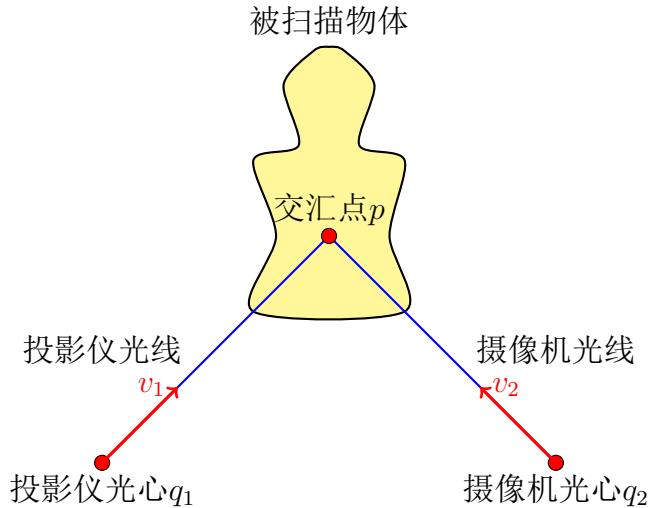


图 2.9: 三角测量的简化模型

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：

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

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：

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

第三章 结构光为颜色条纹序列的3D重建

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[13]:

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° |

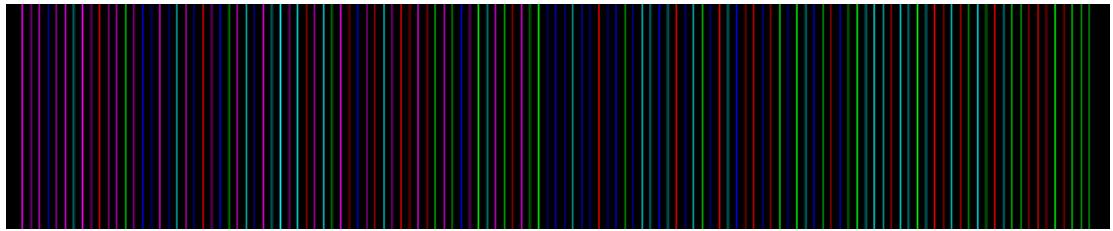


图 3.1: 投影背景

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
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；反射率校正以使个颜色通道的反射率出在相同水平

认为颜色系统的模型为公式 3.1[3]：

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

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)[10]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

\tilde{a} 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。通过流动矩阵的特性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， $H_i^r (1 \leq i \leq n)$ 直方图的灰度值， $W_i^r (1 \leq i \leq n)$ 是对应灰度的值。假设 r、g 和 b 分别代表红、绿和蓝通道，为了匹配 H^r ， H^g 被转换为 $H^{g'}$ ，并且其转换过程可表示为流矩阵的形式： $f = f_{ij} (1 \leq i \leq n, 1 \leq j \leq n)$ 。 f 可通过最小化公式 3.6 的目标函数求得。

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

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。

3.3.3 反射率校正的结果

反射率校正的目的是通过使得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, 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。

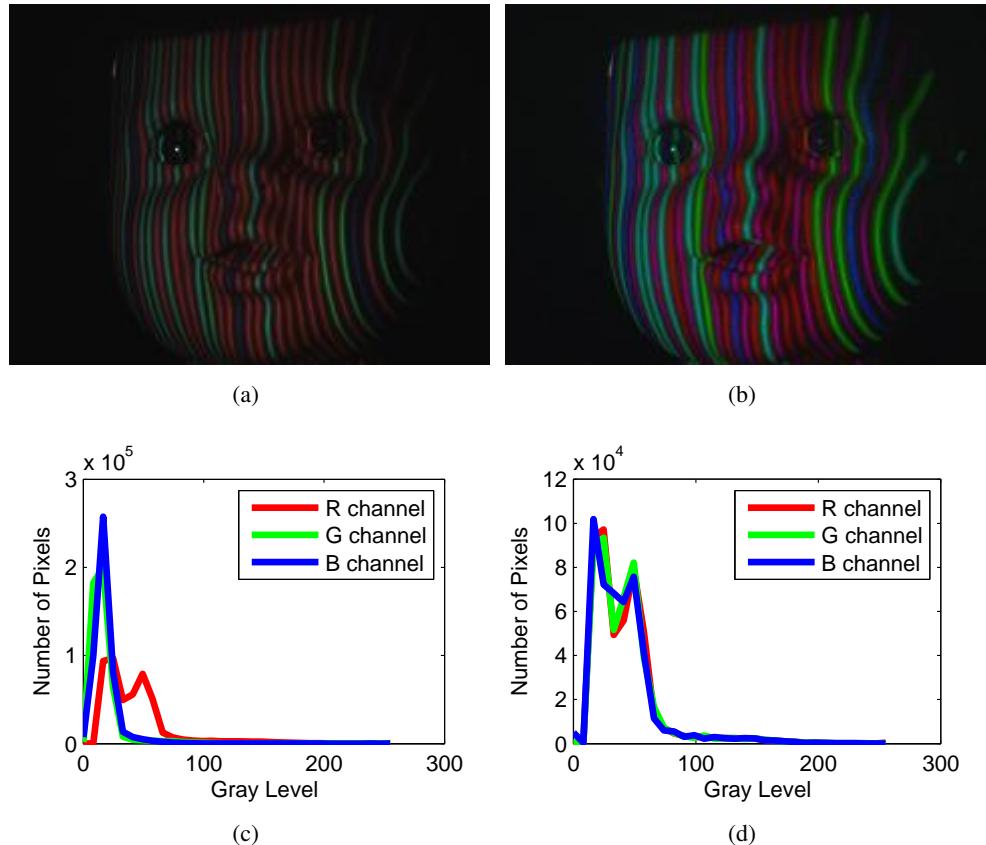


图 3.2: 反射率校正的结果. (a) 加照了结构光的脸部模型, (b) 反射率校正后的图像, (c), (a)的直方图, (d), (b)的直方图.

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%。

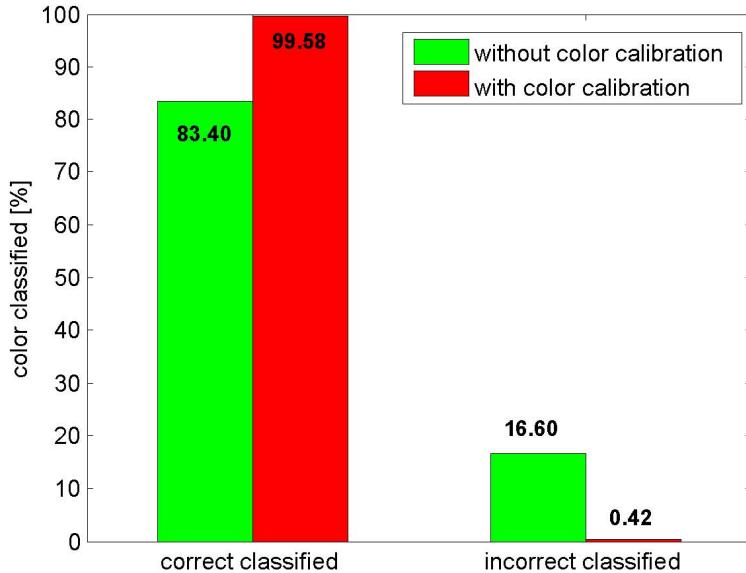


图 3.3: 颜色校正前后的颜色识别准确性比较.

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。为了能够提高检测颜色线条的召回率和准确率，本文定义了两个概念： M 通道和离散趋势变换。

M 通道将R、G、B三个通道融合为一个通道，避免了处理的复杂性，结合投射颜色序列的特殊性，最大限度地保留了原始线条的特性，同时具有一定的去背景光干扰的能力。离散趋势变换将线条的变化特性转换为易检测的形式，提高了线条检测的成功率。

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[12]。

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

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:

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

注意公式 3.14 中存在如下关系:

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

这是因为, 所投射的各颜色中, 至少存在一个为0 的通道。具体可参看表 3.1。

可见, M 通道只和两个因素相关:

- 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 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:

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

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

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，在物体表面起伏坡度较大的地方，由于投射线条被拉扯，常常导致拍摄到的线条起伏较为平坦，不管是亮度还是 M 通道都缺少较高的峰值。DTT能够天然地处理这种情况。这是因为，DTT将线条的先递增、后递减趋势直接变化为最大值最小值模式，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 3.6 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。

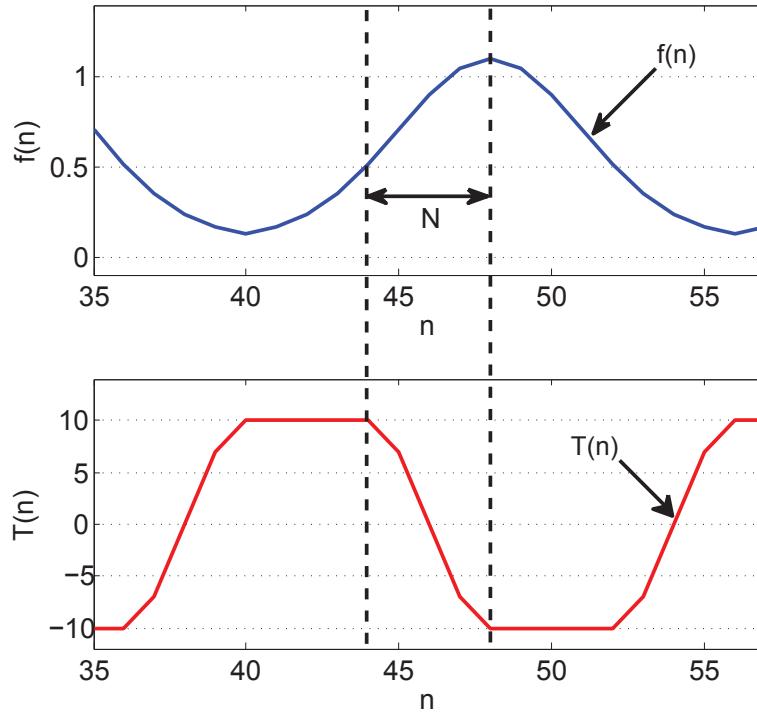


图 3.4: DTT. $T(n)$ 代表了 $f(n)$ 的局部趋势. $T(n)$ 的最大值意味着 $f(n)$ 位于上升沿, 而 $T(n)$ 达到最小值意味着 $f(n)$ 位于下降沿, 且 $T(n)$ 最大值的结束点与 $f(n)$ 的峰值点的距离恰为窗口大小 N .

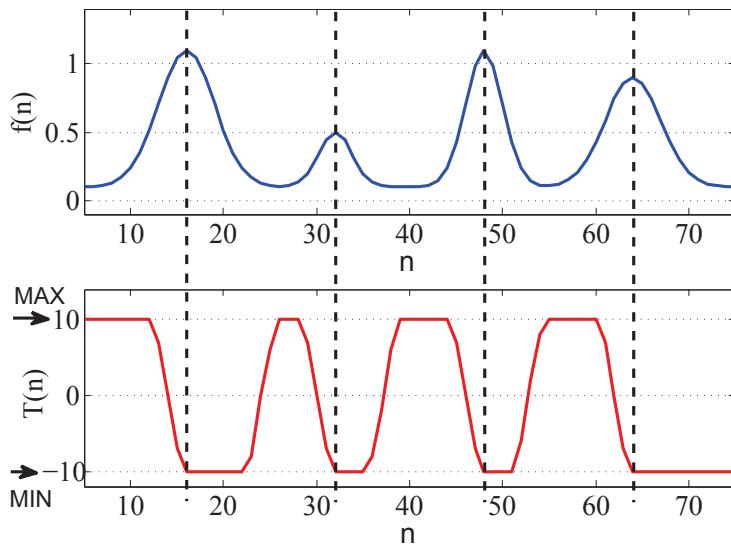


图 3.5: $f(n)$ 和 $T(n)$: 如图, $f(n)$ 有四个符合高斯分布的峰值, 各有不同的宽度和高度, 在对 $f(n)$ 实施DTT 变换得到的 $T(n)$ 中, 四个峰值点则对应相同的最大值到最小值的变化模式, 寻找和定位这样的变化模式可以用于回溯寻找峰值位置.

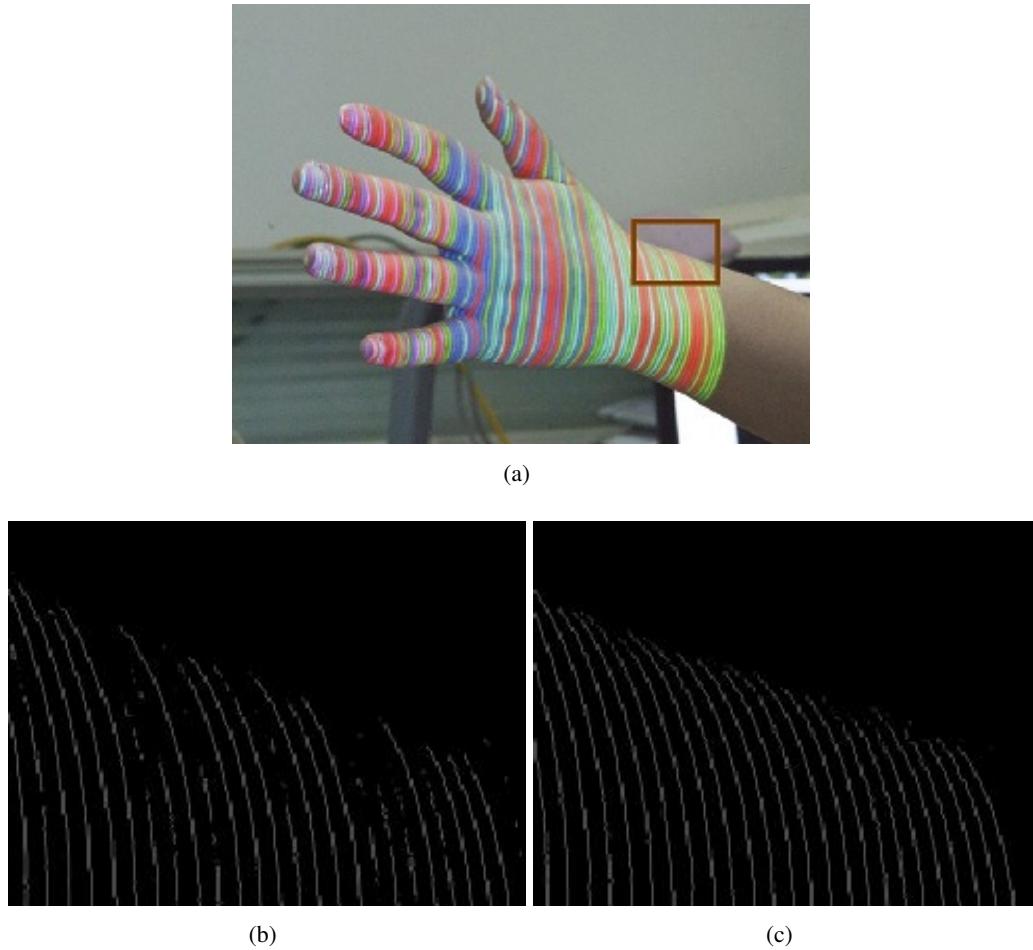


图 3.6: 线条检测的结果对比. (a) 原图像, (b) 自适应阈值二值化方法检测的线条, (c) 使用DTT检测到的线条.

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，现列举如下。

- **最大值方法 Jordi[9]** 使用线条中具有最大值得像素位置作为峰值位置，为了便于对比结果，将此方法简记为MAX。
- **中心点方法 Thomas等[8]** 则使用线条的中心位置作为峰值位置，记吃方法为MID。
- **加权平均法 Zhiqiang等[11]** 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作为峰值位置，记此方法为WA。

- 概率法 Fechteler等[4] 则结合亮度、线条现状等因素，构建抛物线模型并计算峰值位置，记此方法为PA。

本文确定峰值的方法不同于以上方法，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，本文的方法记为MMWA，可表述为公式 3.20。

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

在实际应用中， α 选取的值为0.6。图 3.7 对比了各方法得到的峰值位置对比。

3.4.4 颜色提取

为了进一步解码，还需要对探测到的线条进行颜色识别。从公式 3.4可知，即使在crosstalk和相对反射率校正后得到的图像中，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，本文中的颜色识别是在hue空间中计算色度来进行的。

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：

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

由公式 3.22计算得到的色度与各干扰因素均无关，有效地提高了颜色识别的准确性。

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的。

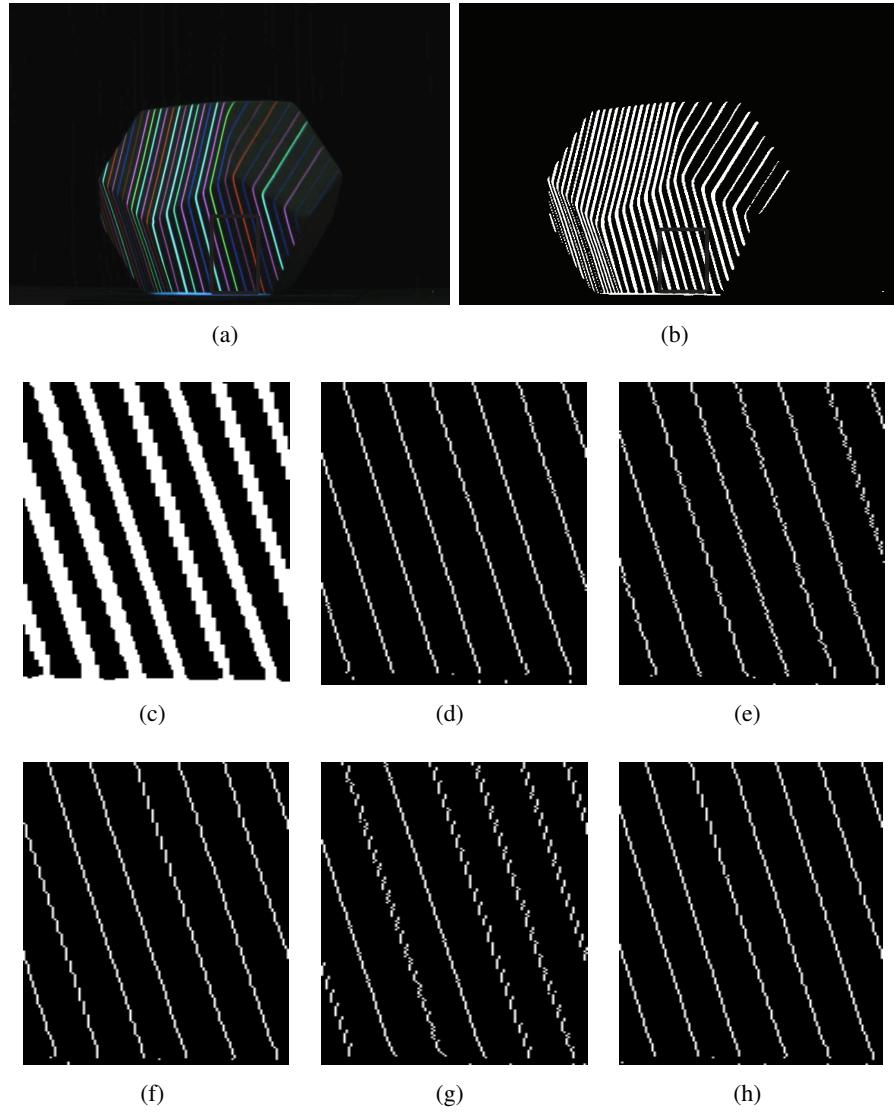


图 3.7: 以不同方法定位的峰值. (a)多面体. (b)线条区域. (c)放大的线条区域. (d)MMWA. (e)PA. (f)MID. (g)MAX. (h)WA

假设 $A_i, 1 \in [1, N]$ 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:

- 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:

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

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

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，实际使用的 $cost_{ij}$ 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：

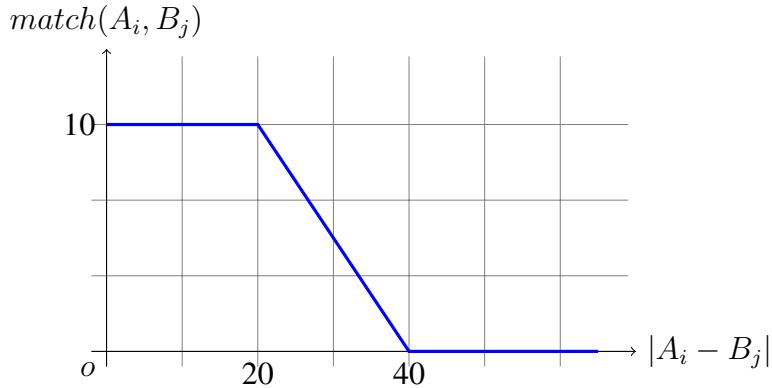


图 3.8: $match(A_i, B_j)$ 的图形表示

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 3.9 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 3.9(b)(d) 分别有 85687 和 75371 个三维点。

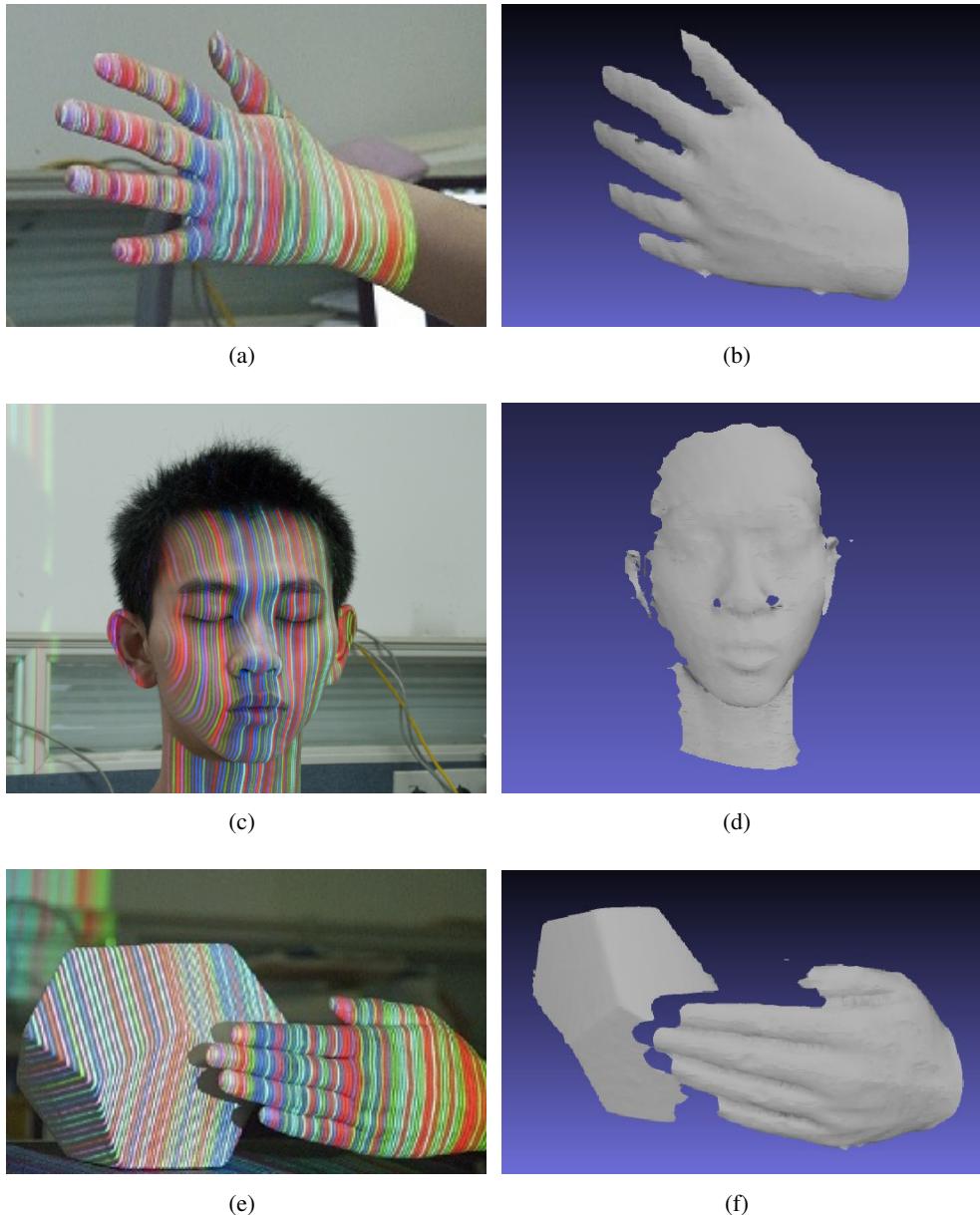


图 3.9: 重建结果: (a) 手, (c) 脸, (e) 手和多面体. (b) 手的重建表面, (d) 脸的重建表面(f) 手与多面体的重建表面.

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，形成形成 2^n 个序列，每个序列的长度为 $n - 1$ 。最后在前 2^{n-1} 个序列



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 |

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，10幅格雷码光栅的分辨率为1024，而7幅格雷码光栅的分辨率为128，额外投射宽度为8的位移码可以填补损失的分辨率。宽度为8的位移

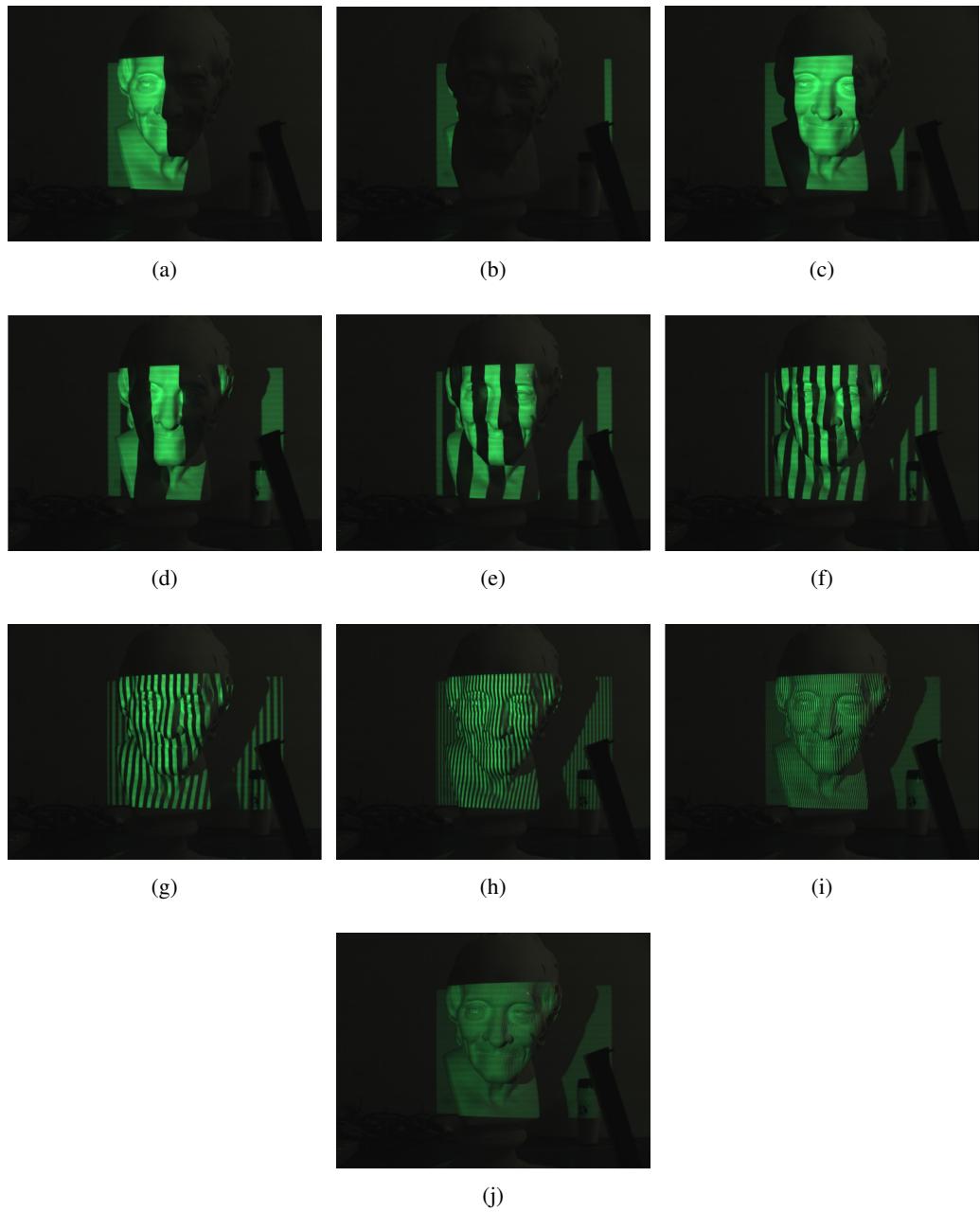


图 4.2: 投射了格雷码光栅的物体

码如图 4.3所示。

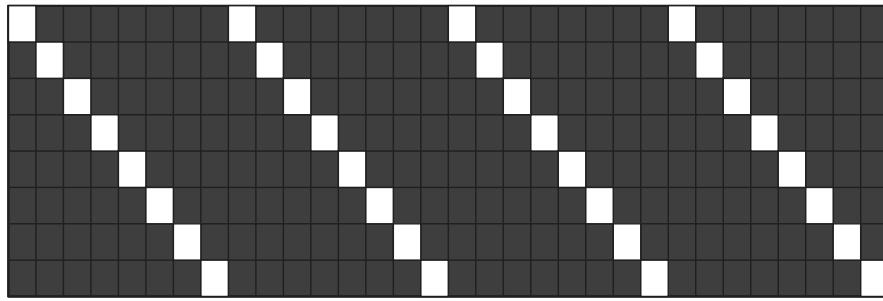


图 4.3: 位移码图

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。

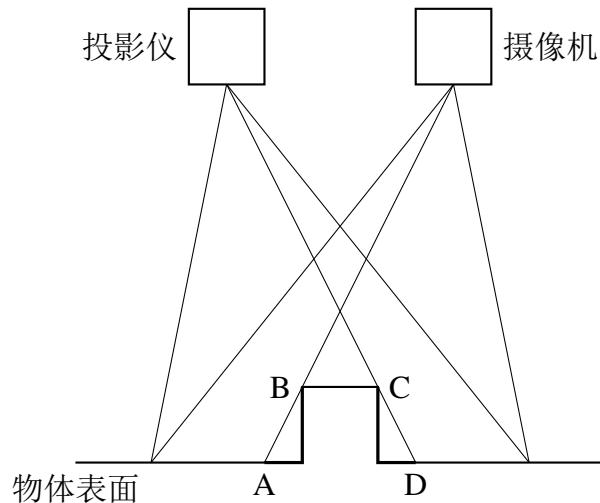


图 4.4: 测量死角的示意图

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。

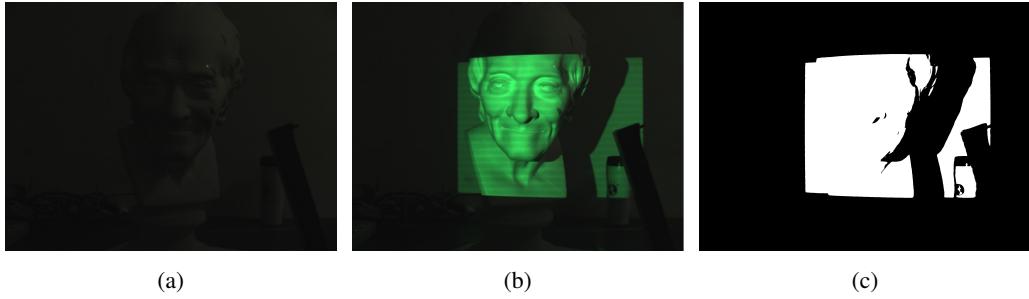


图 4.5: 有效区域提取

4.2.2 解码格雷码光栅

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 $I_{background}$ 。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：

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

3. 二值化 $I_{channel}$

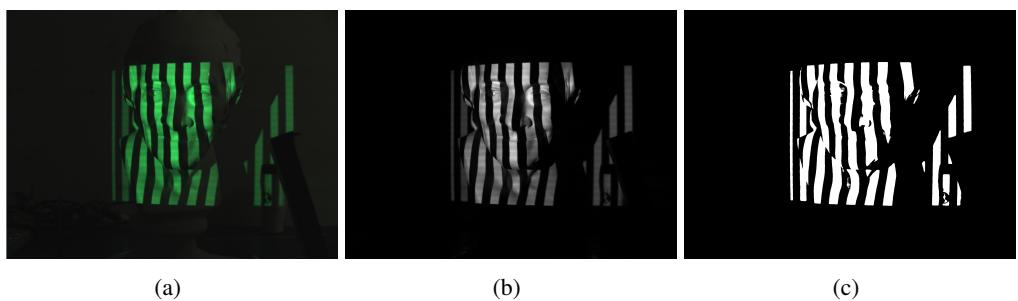


图 4.6: (a),一幅加照了格雷码光栅的物体(b),合并为单通道(c),二值化的结果

图 4.6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)从RGB转换为灰度图后得到的图像，(c)是二值化(b)得到的处理结果。

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，系统中采用的是局部阈值法。

4.2.3 解码位移图码光栅

解码位移图码图像同样需要 $I_{background}$ ，处理步骤为：

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

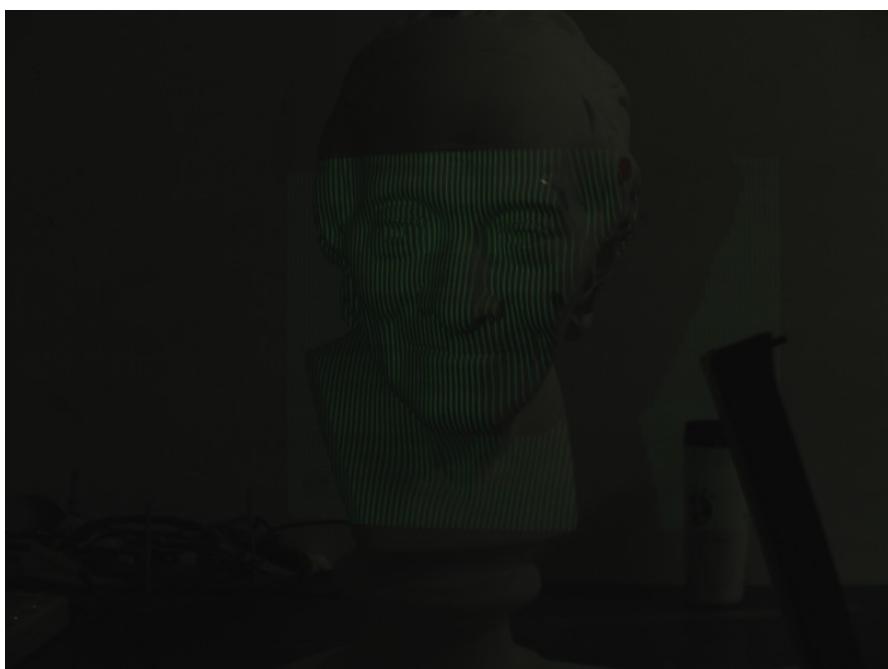
假设投射光的颜色为(R, G, B)，那么：

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

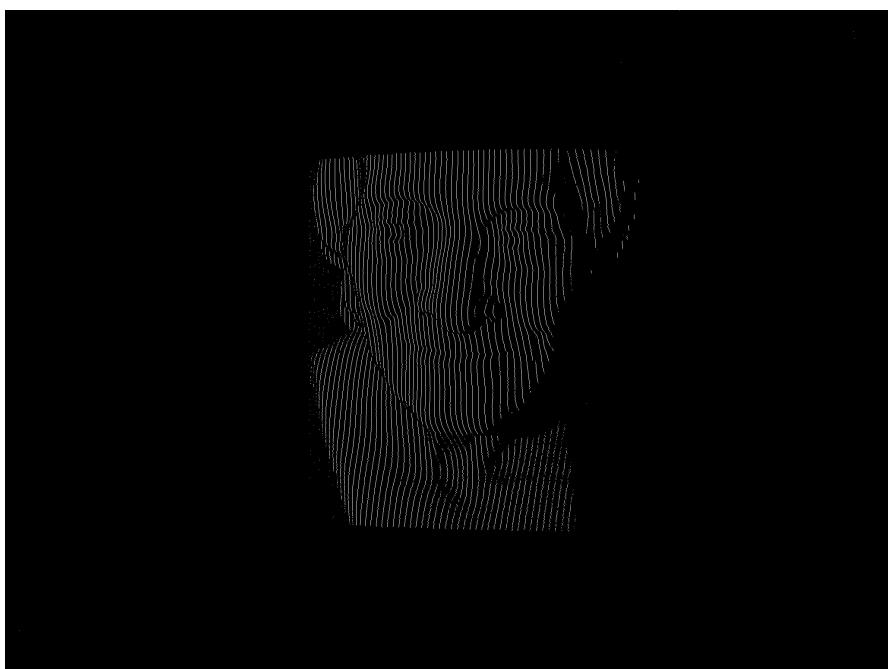
3. 用类似颜色条纹序列中的峰值定位方法定位 $I_{channel}$ 中的峰值

4.3 重建结果

重建的结果如图：



(a)



(b)

图 4.7: 位移码图像的处理(a),拍摄到的源图像(b),检测出的线条

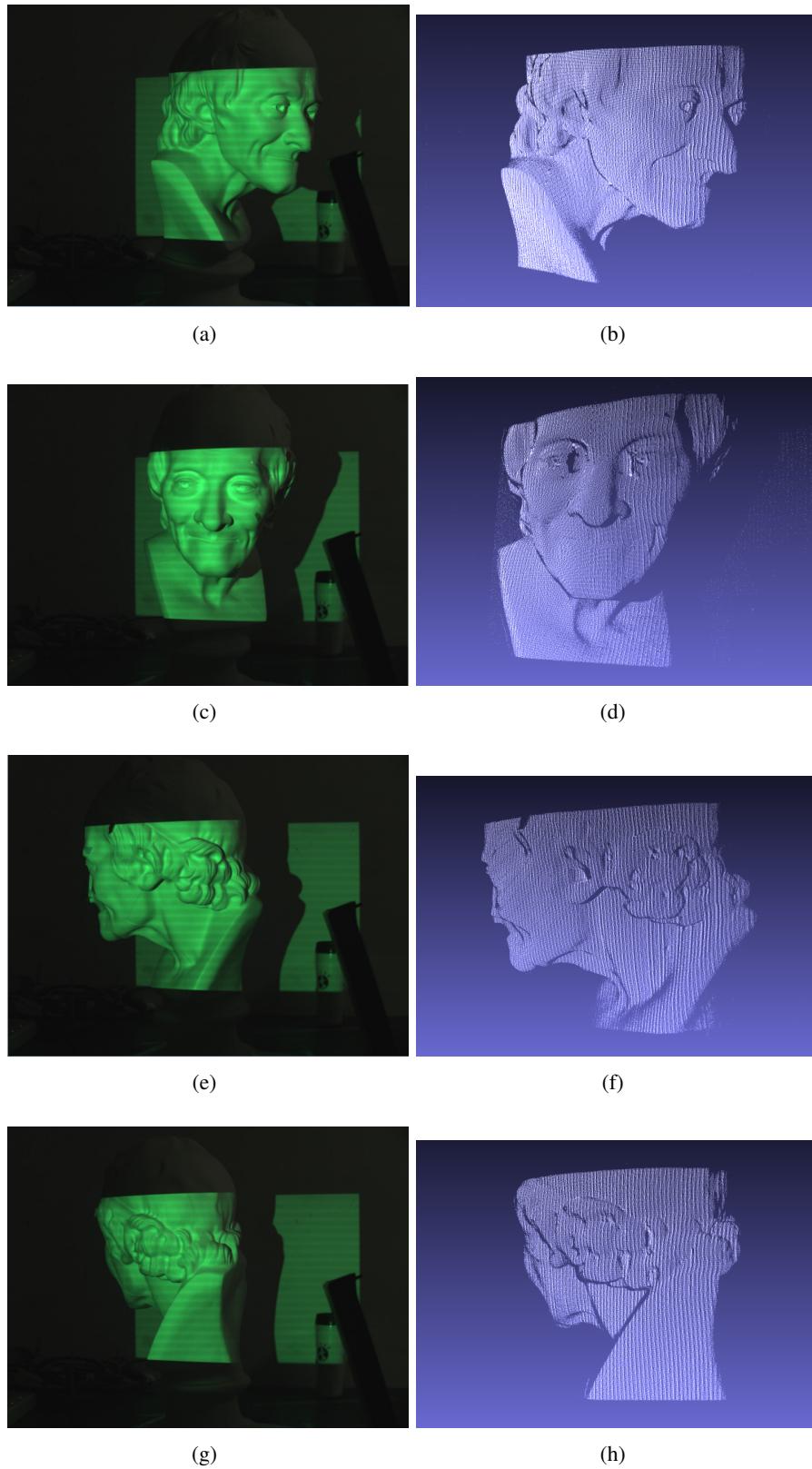


图 4.8: 重建结果

第五章 结束语

5.1 总结

5.2 创新点

5.3 展望

参考文献

- [1] Bradski, G., Kaehler, A., 2008. Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media, Incorporated.
- [2] Brown, D.C., 1966. Decentering distortion of lenses , 444–462.
- [3] 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.
- [4] Fechteler, P., Eisert, P., 2009. Adaptive colour classification for structured light systems. *Computer Vision, IET* 3, 49 –59.
- [5] Fryer, J.G., Brown, D.C., 1986. Lens distortion for close-range photogrammetry. *Photogrammetric engineering and remote sensing* 52, 51–58.
- [6] Hartley, R., Zisserman, A., 2000. Multiple view geometry in computer vision. volume 2. Cambridge Univ Press.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.

- [12] 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.
- [13] Zhou, Y., Zhao, D., Yu, Y., Yuan, J., Du, S., 2012. Adaptive color calibration based one-shot structured light system. Sensors 12, 10947–10963.

致 谢

感谢都思丹老师、周余老师、于耀老师的指导和帮助。