

2015 湖南省研究生数学建模竞赛参赛承诺书

我们仔细阅读了湖南省研究生数学建模竞赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们授权湖南省研究生数学建模竞赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号是（从组委会提供的试题中选择一项填写）：A 题

我们的参赛报名号为（如果组委会设置报名号的话）：201518001005

所属学校（请填写完整的全名）：国防科技大学

参赛队员（打印并签名）：

1. 米杨

2. 李硕

3. 赵本东

指导教师或指导教师组负责人(打印并签名)：王丹

日期：2015 年 9 月 15 日

评阅编号（由组委会评阅前进行编号）：

2015 湖南省研究生数学建模竞赛

编号专用页

评阅编号（由组委会评阅前进行编号）：

评阅记录（可供评阅时使用）：

[illegible]

湖南省首届研究生数学建模竞赛

题 目 平行双目非接触式 3D 扫描系统

摘 要:

本文利用平行双目原理,用两个性能完全相同、平行放置的数码相机代替拍照式三维扫描仪,以自然光模拟3D扫描仪中的白光光栅,通过Matlab编写图像处理方法和重构算法对目标物体实现三维重构。该方法操作简单,设备常见、便携,可以在绝大多数场合下对目标物体进行拍照扫描,对拍摄背景没有限定要求,可有效提取杂乱环境中的三维信息,实现对小尺寸物体的三维重建。

本方案的基本原理及工作方式:在光照均匀的情况下,两个平行放置、相距规定间距的相机对同一物体进行成像,以获取不同视角下的图像。成像得到的左右两幅图像进行立体匹配找出对应点,通过计算得到对应点之间的视差,进而从视差中确定其深度距离信息。其关键问题是确定出物体点在两幅图像中的视差,然后根据给定的投影中心,有效焦距,光轴方向和摄像机几何位置,由三角几何原理恢复出物体的三维结构信息。然后通过仿射变换,还原物体纹理信息。

根据上述原理,系统的实现步骤及方法如下:

固定两相机位置,保证其光轴平行且间距不变,安装系统后进行校准,并对相机内参数进行标定。测量和算法计算过程中所用到的实际参数均通过标定得到,标定后两相机的相对位置不再发生改变。

标定过程是已知相机和目标物体距离,拍摄左右图片进行匹配后计算机内参数。本文将系统设备与目标物体放置于同一水平面,两相机光轴平行,间距取 10cm 进行标定。分别在相机距离目标物体 30cm、40cm、60cm、80cm 处捕获图像并进行匹配,根据匹配到的特征点分别计算得到 α_x, α_y 值,对其取平均,再根据图像像素得到 u_0, v_0 值,即可完成相机内参数的标定。

系统标定后,即可对目标物体进行拍照扫描。在用相机得到左右图像后,首先采用加权平均法对图像进行灰度化预处理并降噪,得到降噪后的灰度图像后,用灰度直方图均衡化来改善对比度,并利用拉普拉斯微分法对图像进行锐化,增强边缘信息。

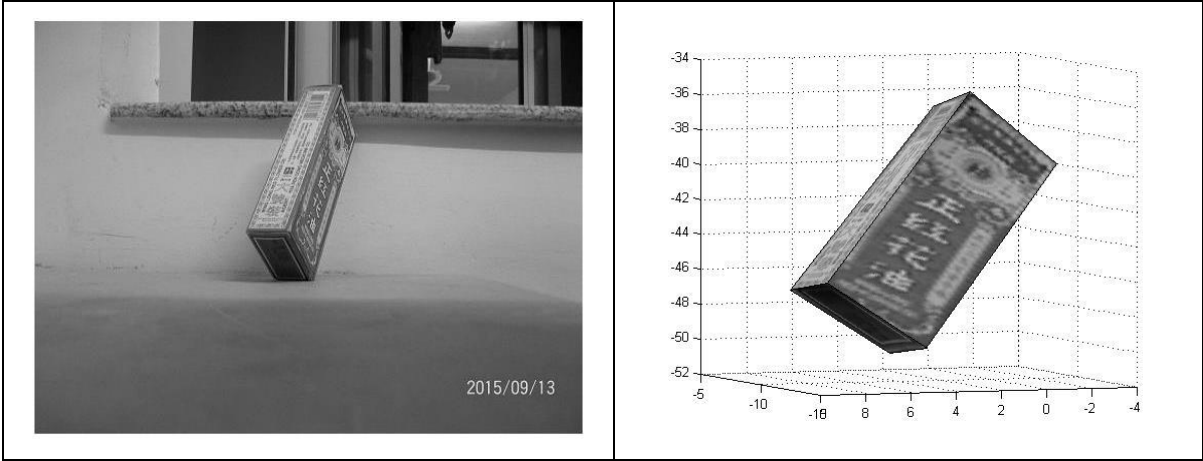
对处理后的图像提取特征点时,我们采用了一种改进的特征点提取算法,即先通过点与点之间的相似性度量初步选定候选角点集,再在候选角点集中运用 Harris 算法提取特征点。该方法可以大大减少乘法运算次数,降低算法复杂度。

在进行图像匹配时,本文采取极线约束和互应性约束,即只在纵坐标相同的像素点进行匹配,并且分别以左右图像为基准,遍历另一幅图像,实现双向搜索。由于摄像机的安装误差等因素的影响,对应匹配点可能不会准确地出现在图像平面中纵坐标相等的极线上,因此本文对第一幅图像某个特征点,在另一幅图像对应纵坐标正负五个像素的带状区域进行搜索匹配,极大地提高了匹配的正确率。为验证匹配算法的准确度,我们选取了一副标准图像及一副自己拍摄图像进行匹配,并用表格的形式给出了所有匹配点的坐标值。匹配结果显示,标准图像的正确率达到100%,自己拍摄图像的匹配正确率大于80%,这是由于光照因素及拍摄误差造成的。下图是标准图像匹配结果。



根据图像的匹配结果，可以得到左右图像中的像素点信息，利用三维重建算法和坐标变换原理可得到物体的三维点坐标，在 Matlab 中将其三维结构还原，最后通过仿射变换及纹理映射还原其外观信息。

本文采用柯达 M320 数码相机，以正红花油外盒作为目标物体构建算法验证实例。数码相机和目标物体处于同一水平桌面，两相机平行放置，间距保持与标定时不变，假设光照均匀，同时拍摄，得到左右图像。经过上述图像处理、特征点匹配、三维点重构及纹理贴图，最终得到的三维结果和原图的对比如下图所示：



我们计算了经过三维重建后的坐标信息及实物测量信息，得到在本次三维重建过程中物体结构误差如下表所示：

	长(cm)	宽(cm)	高(cm)
实物测量	13.6	5.0	2.8
重构信息	13.64	5.14	2.73
误差	0.29%	2.8%	2.5%

通过表中数据计算得到平均误差为 1.86%。从该实例可以看出，本文所提模型可以很好的实现三维重建，是一个简易可行的 3D 扫描系统方案。

关键词：平行双目；三维重建；参数标定；特征点提取；图像匹配

1. 问题重述

3D 扫描仪用于采集现实物体的形状、外观等数据，在 3D 打印、工业设计、动画制作等方面有广泛的用途。3D 扫描仪有多种类型，其扫描方式、测距原理、重构算法各不相同，分辨率、速度、造价、便携性等方面差异较大。最常见的是非接触主动式扫描仪，其基本原理是先将某种光投射到物体表面，再根据采集到的反射信息重构物体表面三维信息。

商用专业 3D 扫描仪通常需要专用硬件设备，价格较高，因此需要提出利用普通投影仪和家用数码相机、摄像机等常见设备构建一个简单的 3D 扫描系统的方案。请详细描述该系统的设备构成，工作方式和重构算法，并讨论系统的分辨率、速度、扫描区域的尺寸限制等技术指标。最后，最好构造合适的实例来说明此 3D 扫描系统的实际应用效果。

2. 问题分析

3D 扫描仪可用来侦测并分析现实世界中物体或环境的几何构造与外观数据。本题要求我们用常用设备构造一个简单的 3D 扫描系统方案，实现对物体的三维重建。

扫描仪工作原理可分为主动和被动两种。主动式主要指激光扫描技术，需主动发射光源并收集反射信息；被动式主要指拍照式扫描技术，通过双目视觉原理得到立体景深信息。

问题背景中提供的思路即为主动扫描。在主动扫描过程中，需向待扫描物理发射激光，逐点扫描目标物理，通过接收到的信息进行三维重建。在常见设备中，需利用投影仪代替激光光源，相机代替激光接收端采集反射信息。然而常见投影仪无法给出点光束，只能得到目标物体的面信息而不是精确点信息，这会增加系统复杂度，增大算法误差。另外投影仪体积大，需要连接电源，不方便携带，需要单纯背景捕获信息，因此对实验场地要求高，不能进行实时测量。

我们从便携、设备易获取的角度出发，将问题解决的范围限定在小尺寸物体扫描，问题不涉及建筑物、房间等大体积物体，因此根据工作原理的特点，选取被动拍照式扫描方式。本文利用平行双目原理[1][2]，用两个性能完全相同、平行放置的数码相机代替拍照式三维扫描仪，以自然光模拟白光光栅，通过Matlab编写图像处理方法和重构算法对目标物体实现三维重构。该方法操作简单，设备常见、便携，可以在绝大多数场合下对目标物体进行拍照扫描，对拍摄背景没有限定要求，可有效提取杂乱环境中的三维信息，实现对小尺寸物体的三维重建。

本方案的基本原理及工作方式为：在光照均匀的情况下，两个平行放置、相距规定间距的相机对同一物体进行成像，以获取不同视角下的图像。成像得到的左右两幅图像进行立体匹配找出对应点，通过计算得到对应点之间的视差，进而从视差中确定其深度距离信息。其关键问题是确定出物体点在两幅图像中的视差，然后根据给定的投影中心，有效焦距，光轴方向和摄像机几何位置，由三角几何原理恢复出物体的三维结构信息。最后通过仿射变换，得到物体的纹理信息。

平行双目法求得重构三维模型的流程如图 2-1 所示。

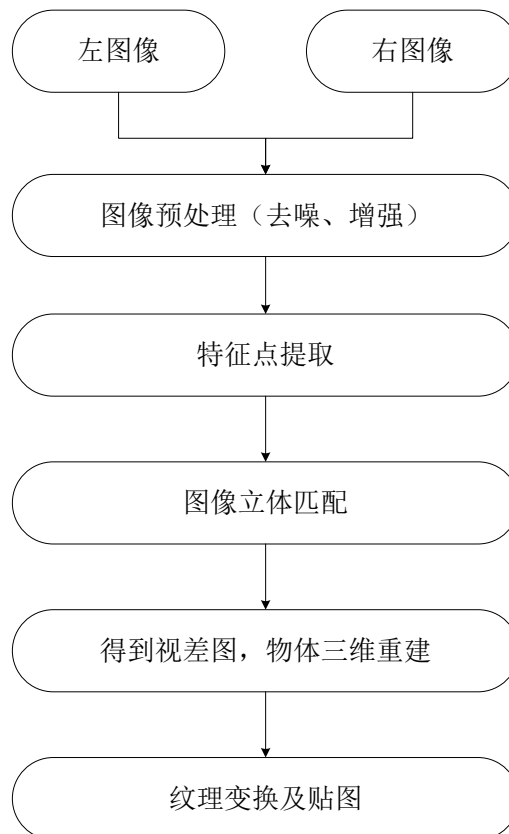


图 2-1 平行双目 3D 扫描仪流程图

3. 模型假设

- (1) 假设两个摄像头同时捕捉目标，无时间差
- (2) 假设两个摄像头性能参数完全一致
- (3) 假设摄像头间距不大，保证满足图像立体匹配约束条件
- (4) 假设摄像头主光轴严格平行
- (5) 假设光照均匀
- (6) 假设目标物体有明显的边缘或角点特征
- (7) 假设标定过程中所有测量无误差

4. 符号说明

符号说明见表 1。

表 1 符号说明表

符号	名称	备注
$f(x+i, y+j)$	原图像像素灰度值	
$g(x, y)$	变换后图像像素灰度值	原图像像素灰度的加权和
$w(i, j)$	权值	
$g(r_0)$	中心点 r_0 的图像区域像素	邻域内与中心点 r_0 相似的

	的相似度	像素点个数
$C(r_0)$	r_0 的 N 邻域	
$D(r, r_0)$	邻域内某点 r 与中心点 r_0 是否相似的判断函数	$D(r, r_0)=1$ 或 0 , 对应相似或不相似
$I(r_0)$	中心点像素 r_0 的灰度值	
$I(r)$	中心点 N 邻域内某像素点 r 的灰度值	
t	点与点相似的判定阈值	$ I(r) - I(r_0) \leq t$
$R(p, q)$	特征点 p, q 的相似度	$-1 < R(p, q) < 1$
$\bar{u}_l, \bar{u}_r$	左(右)图像匹配窗口内的灰度平均值	
$\sigma(I_l), \sigma(I_r)$	左(右)图像匹配窗口内灰度的标准差	
(X_w, Y_w, Z_w)	世界坐标系	
(X_c, Y_c, Z_c)	摄像机坐标系	
(x, y)	像平面坐标系	
(u, v)	图像坐标系	
(u_0, v_0)	图像坐标系中的主点坐标	
d	视差	
b	基线	
f	摄像机焦距	
N_u	图像平面 u 方向单位距离上的像素点数	
N_v	图像平面 v 方向单位距离上的像素点数	
D	点 P 的双目视差	

5. 问题的建模与求解

5.1 系统构成

平行双目视觉模型几何关系如图5-1所示：

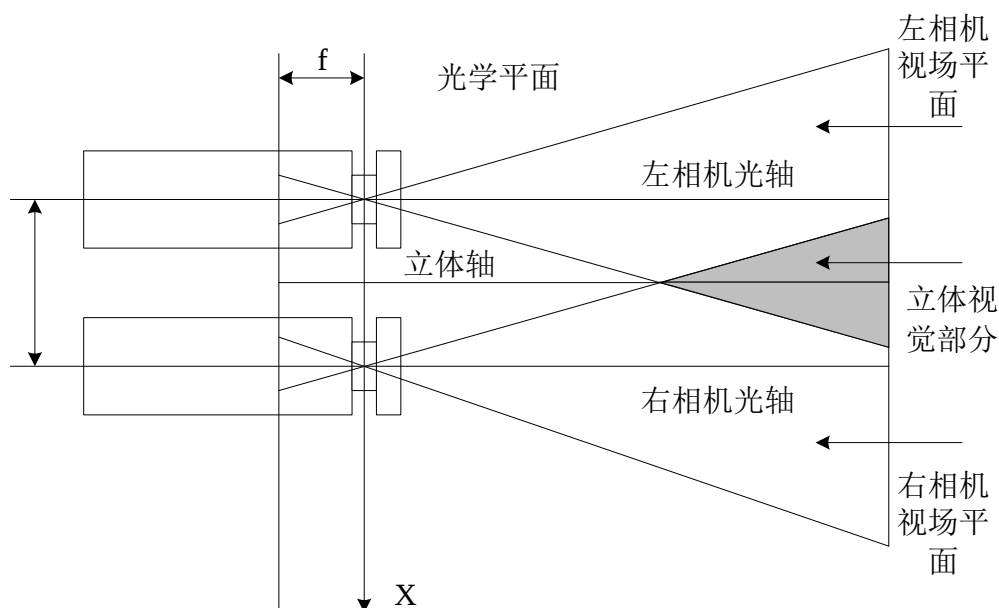


图 5-1 平行双目视觉模型

模型由两个完全相同的相机构成，所成图像位于同一平面，两个相机的光轴相互平行，且 X 轴重合，相机之间在 X 轴方向上的间距称为基线距离 b ，图中立体视觉部分即为可恢复三维信息的部分。

本文构造的系统中，需要两个性能参数及外观尺寸完全相同的相机，对相机品牌、型号没有特殊要求。我们出于设备易获取，系统方便安装固定，精度较高的角度考虑，选择同一型号的两台卡片式数码相机作为测试设备。为保证两台相机的光轴严格平行，相机间隔保持不变，采用固定装置固定卡片机的前后两端，将两台照相机固定在卡尺上，间距为 10cm。这个距离参数是我们模仿人眼双目距离而设定的，可以根据实际需求 and 限制因素进行更改，但要保证在标定后该参数不再改变，否则会对系统精度产生很大影响。

在系统安装的过程中，需对相机间距、相机主光轴是否平行进行校准。对于相机间距，我们采用多次测量取平均值的方法；对于相机主光轴是否平行，我们采取在阴暗环境下，打开两台闪光灯并使光斑打在深色平板上的方式进行校准，根据光斑在平板上的大小、形状、相对位置等参数判断相机主光轴是否平行。在测量过程中，如果间距发生改变，则需重新进行标定，否则会对重构算法带来很大误差；如果两相机无法保持平行，角度差过大时会使算法失效。重构算法中使用的实际参数，均已标定结果为准。

5.2 参数标定

要对障碍物进行三维坐标计算，首先需要对摄像机的内参数 $u_0, v_0, \alpha_x, \alpha_y$ 进行标定。

本文使用柯达 M320 数码相机拍摄得到 3472×2604 像素的彩色照片，为了减少运算量，提高系统算法效率，首先使用图像压缩软件，将图片压缩至 533×400 像素。通过对已知距离和已知尺寸的口香糖盒子拍照，然后分析计算得出所需摄像机参数。具体分析如下：

(1) u_0, v_0 的标定：由于所处理的图片均被压缩至 533×400 像素，故很容易确定

$u_0 = 267, v_0 = 200$ 。

(2) 参数 b 可人为设定，本文每次对同一目标物体拍摄左右两幅图像时均使摄像机平移10cm，即 $b = 10\text{cm}$ 。

(3) α_x 的标定：对距离摄像机固定距离的口香糖盒子拍摄左右两幅图像，通过图像匹配找出对应点，算出视差，便可根据公式 $z_w = \frac{b\alpha_x}{u_r - u_l}$ 得出 α_x 的值。本文对距离为80cm，60cm，40cm，30cm的盒子分别算出各自对应的 α_x 值，然后取其平均。

根据上述分析，将目标物体与摄像头放置在同一水平桌面，标定过程如下：
口香糖盒距离摄像机为80cm时左右两幅图像匹配结果如图5-2所示：

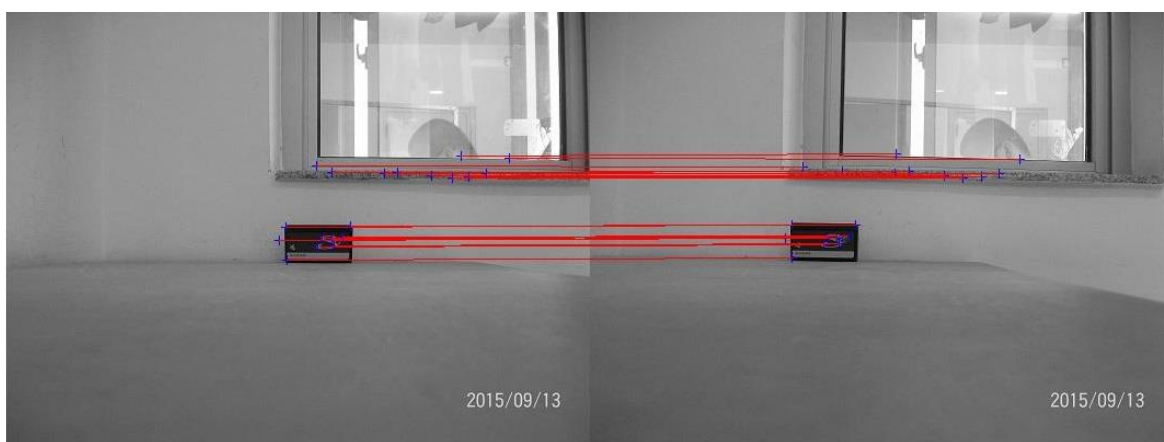


图5-2 距离为80cm时左右两幅图像匹配结果图

由于已知距离80cm为口香糖盒与摄像机的距离，故应找出口香糖盒上某匹配点对，通过计算其视差而根据公式得出 α_x 值。

取盒子上某匹配点对如图5-3所示：



图5-3 距离为80cm时盒子上某点匹配图

图中两点图像坐标分别为 $(u_l, v_l) = (185, 196)$ ， $(u_r, v_r) = (256, 198)$ ，视差

$u_r - u_l = 71$ 。带入公式 $z_w = \frac{b\alpha_x}{u_r - u_l}$ 得 $\alpha_x = 568$ 。

同理分析距离为60cm, 40cm, 30cm的情况, 如图5-4、5-5和5-6所示:

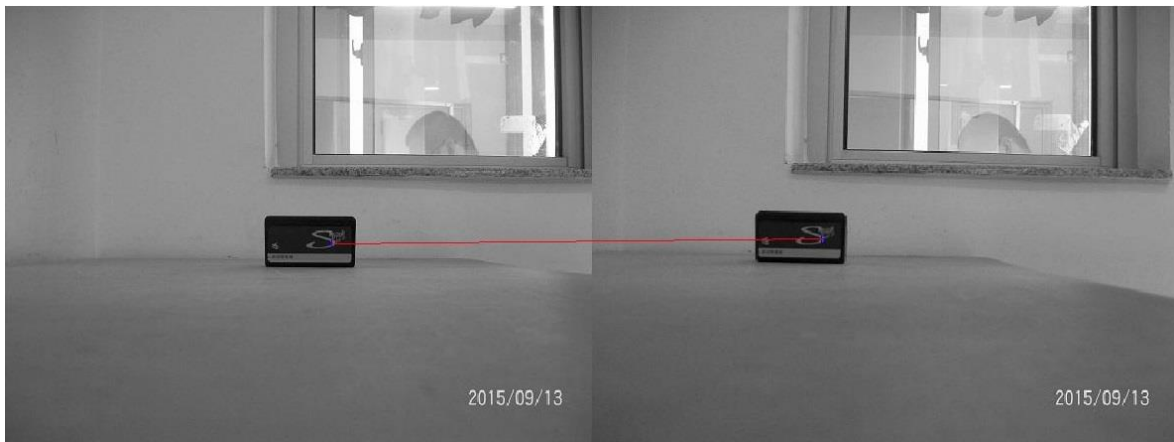


图5-4 距离为60cm时盒子上某点匹配图



图5-5 距离为40cm时盒子上某点匹配图

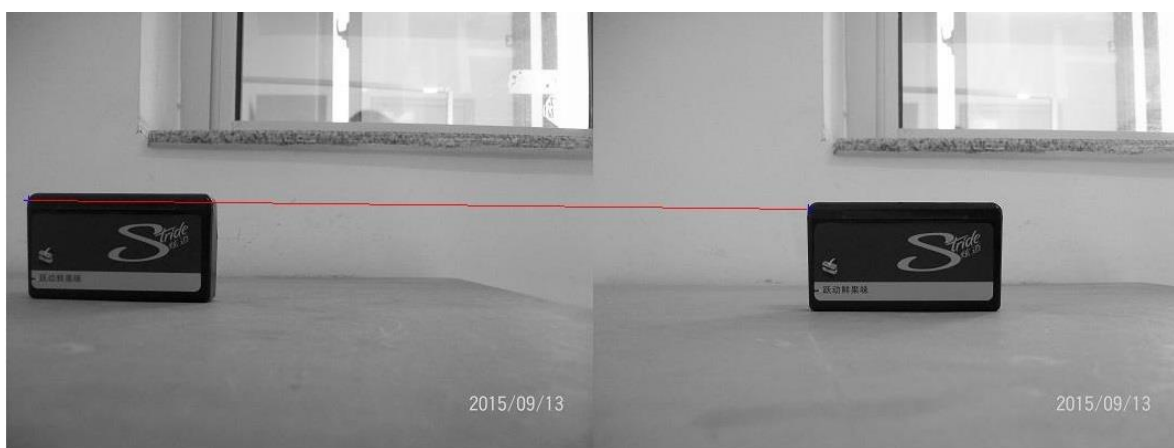


图5-6 距离为30cm时盒子上某点匹配图

距离为60cm时, 匹配点对图像坐标为 $(u_l, v_l) = (267, 210)$, $(u_r, v_r) = (169, 212)$, 视差 $u_r - u_l = 98$, 计算得出 $\alpha_x = 558$; 距离为40cm时, 匹配点对图像坐标为

$(u_l, v_l) = (319, 220)$ ， $(u_r, v_r) = (181, 220)$ ，视差 $u_r - u_l = 138$ ，计算得出 $\alpha_x = 552$ ；距离为30cm时，匹配点对图像坐标为 $(u_l, v_l) = (46, 227)$ ， $(u_r, v_r) = (243, 225)$ ，视差 $u_r - u_l = 197$ 。计算得出 $\alpha_x = 562$ ；取其平均得 $\alpha_x = 560$ 。

(4) α_y 的标定：同 α_x 的标定类似，对已知高度的口香糖盒拍摄左右两幅图像，完成匹配后根据公式 $y_w = \frac{b\alpha_x(v_l - v_0)}{\alpha_y(u_r - u_l)}$ 得出 α_y 的值。本文采取的方式是将相机立着拍摄左右图像对，那么 α_y 的计算过程则与 α_x 的计算过程完全一样，计算得 $\alpha_y = 740$

5.3 重构算法

用相机进行图像采集即进行成像变换[3]，需进行不同坐标系统之间的变换，模型涉及到的坐标系有以下四种：

(1) 世界坐标系

世界坐标系 (X_w, Y_w, Z_w) 也称为全局坐标系，它是由用户来定义的三维空间坐标系，用来真实地反应客观景物世界。

(2) 摄像机坐标系

摄像机坐标系 (X_c, Y_c, Z_c) 一般以摄像机光心为原点， Z_c 轴与摄像机光学轴重合，与摄像机成像平面垂直，取摄影方向为正向。

(3) 像平面坐标系

像平面坐标系 (x, y) 是在摄像机内所形成的平面坐标系统，一般以光轴与图像平面的交点为原点， x 、 y 轴分别平行于 X_c 、 Y_c 轴。

上述三种坐标系如图5-7所示， $P(X_c, Y_c, Z_c)$ 是摄像机坐标系下的物点坐标， p 点是像平面坐标系下的像点坐标。

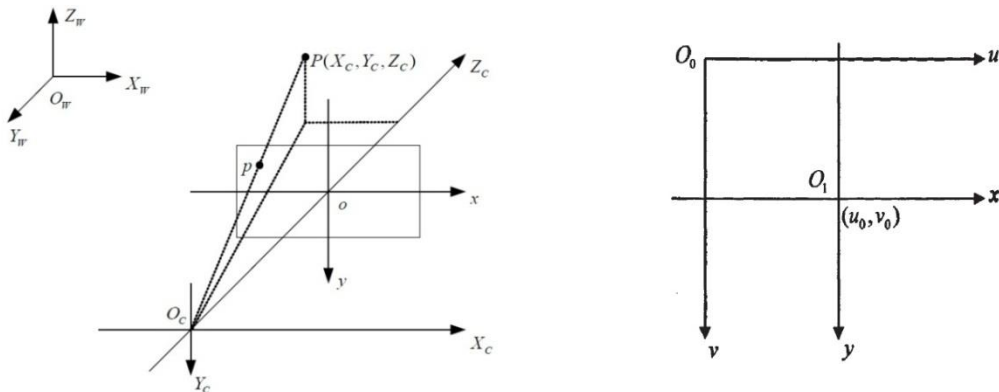


图5-7 世界坐标系，摄像机坐标系，
像平面坐标系

图5-8 计算机图像坐标系
与成像平面坐标系

(4) 计算机图像坐标系

在计算机中，图像是以 $M \times N$ 的二维数组的形式存储的，计算机图像坐标系 (u, v) 是在计算机内部数字图像所使用的坐标系统，以图像左上角为原点，以像素为单位， u 、 v 分别表示该像素在数字图像中的列数与行数，如图5-8所示。

本文采用的摄像机模型为针孔模型，是一种理想成像模型，如图5-9所示。

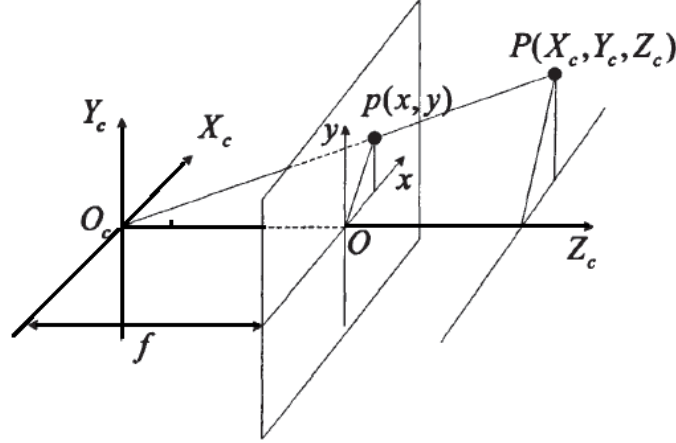


图5-9 针孔摄像机模型

在针孔摄像机模型中，三维空间物体上任一点P投影到二维图像平面，主要经过下列坐标变换得到[4]，各步骤具体分析如下：

(1) 从世界坐标系到摄像机坐标系的变换

假设景物为刚体（物体在变换过程中只有位置和角度发生变化，形状不变），变换可表示为式(5.1)：

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = R \begin{bmatrix} X_W \\ Y_W \\ Z_W \end{bmatrix} + T \quad (5.1)$$

其中， R 是 3×3 的旋转矩阵， T 是 3×1 的平移矢量，

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad T = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

(2) 从摄像机坐标系到像平面坐标系的变换

由三角几何比例关系，得式(5.2)：

$$\frac{f}{Z_C} = \frac{x}{X_C} = \frac{y}{Y_C} \quad (5.2)$$

即：

$$x = f \frac{X_C}{Z_C} \quad y = f \frac{Y_C}{Z_C} \quad (5.3)$$

(3)从像平面坐标系到计算机图像坐标系间的变换如式(5.4):

$$u = N_u x + u_0 \quad v = N_v y + v_0 \quad (5.4)$$

其中, (u_0, v_0) 是计算机图像坐标系的中心坐标, 即标定得到的参数值, N_u 和 N_v 是图像平面 u 和 v 方向单位距离上的像素点数。

以左摄像机坐标系作为世界坐标系, 设空间点 $P(x_w, y_w, z_w)$ 在 C_l 、 C_r 摄像机坐标系下坐标为 $P_{cl}(x_{cl}, y_{cl}, z_{cl})$, $P_{cr}(x_{cr}, y_{cr}, z_{cr})$, 成像平面上的投影点分别为 $P_l(x_l, y_l)$ 、 $P_r(x_r, y_r)$, 由三角几何关系易知:

$$\frac{x_l}{x_{cl}} = \frac{y_l}{y_{cl}} = \frac{f}{z_{cl}}, \quad \frac{x_r}{x_{cr}} = \frac{y_r}{y_{cr}} = \frac{f}{z_{cr}} \quad (5.5)$$

左摄像机坐标系与世界坐标系重合, x 、 y 、 z 坐标均相同, 则点在两坐标系下的关系如公式(5.6)所示:

$$\begin{cases} x_w = x_{cl} = x_{cr} - b \\ y_w = y_{cl} = y_{cr} \\ z_w = z_{cl} = z_{cr} \end{cases} \quad (5.6)$$

则:

$$x_{cr} - x_{cl} = b$$

两个像点在成像平面坐标系中的 y 坐标相同, x 坐标不同, 分别设为 x_l 、 x_r , 若 $P_l(x_l, y_l)$ 、 $P_r(x_r, y_r)$ 已知, 则:

$$x_{cl} = \frac{z_{cl} x_l}{f} = \frac{z_w x_l}{f} \quad x_{cr} = \frac{z_{cr} x_r}{f} = \frac{z_w x_r}{f} \quad (5.7)$$

$$z_w = \frac{bf}{x_r - x_l} \quad (5.8)$$

设 $x_r - x_l = d$, 则:

$$z = \frac{bf}{d} \quad (5.9)$$

其中, d 实际上就是两个像点相对于图像坐标系中 $P_l(x_l, y_l)$ 与 $P_r(x_r, y_r)$ 坐标之差, 即视差; b 是左右两个摄像机光心之间的距离, 即基线; f 是摄像机焦距。

于是 $P(x, y)$ 的世界坐标可表示为式(5.10):

$$\begin{cases} x_w = \frac{bx_l}{x_r - x_l} \\ y_w = \frac{by_l}{x_r - x_l} \\ z_w = \frac{bf}{x_r - x_l} \end{cases} \quad (5.10)$$

如果由计算机像素坐标表示，则可表示为式(5.11):

$$\begin{cases} x_w = \frac{b(u_l - u_0)}{u_r - u_l} \\ y_w = \frac{b\alpha_x(v_l - v_0)}{\alpha_y(u_r - u_l)} \\ z_w = \frac{b\alpha_x}{u_r - u_l} \end{cases} \quad (5.11)$$

式中: $\alpha_x = f \cdot N_u, \alpha_y = f \cdot N_v$ 。 N_u 和 N_v 是图像平面 u 和 v 方向单位距离上的像素点数;

(u_0, v_0) 是 O_l 在计算机图像坐标系中的主点坐标。

设 $D = u_r - u_l$ ，记为点 P 的双目视差(由像素坐标表示，在编程中计算方便)，由此可见，只要得到两个摄像机的内参数 $u_0, v_0, \alpha_x, \alpha_y$ 和两个摄像机之间的距离 b ，就可以重构该点的三维坐标。相机内参数已在 5.2 节标定得到，摄像头间距离定位 10cm，故可根据重构算法得到目标物体三维坐标点，还原物体结构信息。

5.4 仿射变换

上述重构算法可以还原出目标物体的三维结构信息，想要得到外观信息，则需截取图像中的面，通过仿射变换[5]及纹理映射完成。

仿射变换模型不但考虑两幅图像之间的平移、旋转和缩放外，而且考虑畸变中的倾斜、纵横比变化等复杂几何变化，适用于从同一视角、不同位置获取的同一场景的两幅图像间的匹配。仿射变换模型可表示如下：

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} t_u \\ t_v \end{bmatrix} \quad (5.12)$$

式中 a, b, c, d 和 $[t_u, t_v]$ 为模型参数。通过待定系数法确定仿射变换的模型参数，进而将二维图像中的纹理映射到三维空间中，完成物体的外观重建。

5.5 模型求解过程

在问题分析中，我们给出了本文的实现流程图，可以看出，实现 5.2、5.3 节中的参数标定及重构算法需要进行图像处理及匹配，下面给出具体过程。

5.5.1 图像处理

1) 图像预处理

首先对摄像机获取彩色图像灰度化，采用加权平均法将 R 、 G 、 B 分量值等量化，即令：

$$R = G = B = w_R R + w_G G + w_B B \quad (5.13)$$

通过实验证明，当其加权系数 $w_R = 0.3$ 、 $w_G = 0.59$ 、 $w_B = 0.11$ 时，能得到最合理的灰度图像。

2) 图像去噪

- (1) 将模板在图像中滑动，并将模板中心与图像中某个像素位置重合；
- (2) 将模板上系数与模板下对应像素相乘；
- (3) 将所有乘积相加；
- (4) 将和(模板的相应输出)赋给图像中对应模板中心位置的像素。即：

$$g(x, y) = \sum_i \sum_j w(i, j) f(x+i, y+j) \quad (5.14)$$

其中， $g(x, y)$ 为变换后的图像像素， $f(x+i, y+j)$ 为原图像像素， (x, y) 为窗口中心元素， i, j 的取值范围决定窗口的大小， $w(i, j)$ 为加权值，为防止加权算法

造成图像变亮或变暗，需 $\sum_i \sum_j w(i, j) = 1$ 。

3) 图像均衡

本文用灰度直方图均衡化来改善对比度，该方法的基本思想是把原始图像的直方图变换成均匀分布的形式，这样就增加了像素灰度值的动态范围，从而达到增强图像整体对比度的效果。由于需进行处理的图像为数字图像，所以灰度直方图是一个离散函数。具体步骤如下：

- (1) 在 0 到 255 的灰度范围内，遍历整幅图像，得到第 k 灰度级出现的次数 n_k 。
- (2) 用频数近似代替概率值，如式(5.15)：

$$p_r(r_k) = n_k / n \quad (5.15)$$

式中 $r_k = \frac{f(x, y)}{255} \in [0, 1]$ ， $k \in [0, 255]$ ， n 为图像中的总像素。

- (3) 算变换后的灰度值，如式(5.16)：

$$g(x, y) = 255 \times s_k = 255 \times \sum_{j=0}^k p_r(r_j) \quad (5.16)$$

经过以上处理，图像的亮度分布均匀化，立体图对间的亮度差异得到平衡，有利于后面需要进行的图像立体匹配。

4) 图像锐化

图像锐化可以增强图像的边缘信息。本文采用微分法中的拉普拉斯锐化法对图像进行锐化，具体过程如下：

用 $f(x, y)$ 来表示数字图像 (x, y) 这一像素的灰度值，那么它的二阶差分可以表示为式(5.17)

$$\begin{cases} \frac{\partial^2 f(x, y)}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y) \\ \frac{\partial^2 f(x, y)}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y) \end{cases} \quad (5.17)$$

由拉普拉斯算子的定义可知

$$\begin{aligned} \nabla^2 f(x, y) &= \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \\ &= f(x-1, y) + f(x, y-1) + f(x+1, y) + f(x, y+1) - 4f(x, y) \end{aligned} \quad (5.18)$$

对于因扩散引起的图像模糊，可以用下式进行锐化：

$$g(x, y) = f(x, y) - k \nabla^2 f(x, y) \quad (5.19)$$

上面公式中， k 是一个与扩散有关的系数，本文选定 $k = 0.5$ 。

本文所采用的图像处理算法均用Matlab编程实现。为测试算法性能及后续特征点提取及匹配算法，我们上网查阅到一组平行双目法拍摄的标准图像，对其进行灰度化、降噪及锐化处理，效果如图5-10到图5-12所示。

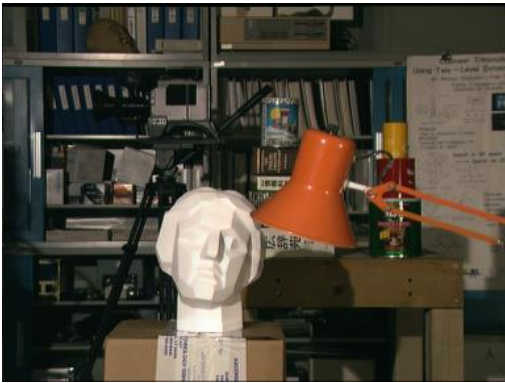


图 5-10 摄像机获取的彩色原图

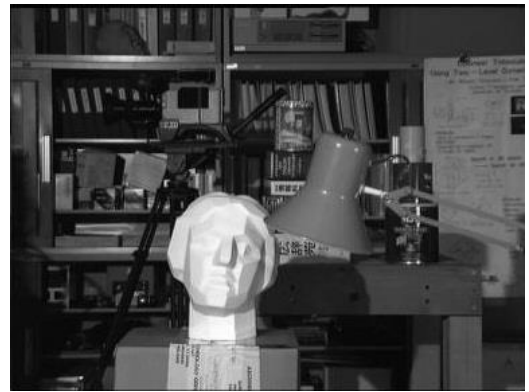


图 5-11 彩色图像灰度化得到的灰度图



图 5-11 图像去噪后的灰度图

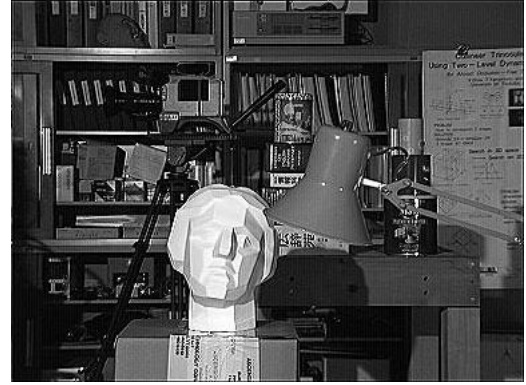


图 5-12 图像增强后的灰度图

5.5.2 特征点提取

本文采用的特征点提取法,是在Harris算法基础上的一个改进,以达到快速准确提取角点信息的目的,其主要思路是通过图像区域像素的相似度找到一个候选角点集,再在候选角点集中采用Harris算法提取角点。

像素的相似度是指N邻域内所有像素点与中心像素点相似的个数。点与点的相似是通过两点的灰度值之差来衡量的,如果邻域内某点灰度值与中心点的灰度值之差的绝对值在一个阈值 t 范围内,那么就认为这个点与中心点相似。具体可表述为公式(5.20):

$$g(r_0) = \sum_{r \in C(r_0)} D(r, r_0) \quad (5.20)$$

式中, $g(r_0)$ 为中心点 r_0 的图像区域像素的相似度, $C(r_0)$ 为 r_0 的N邻域, $D(r, r_0)$ 为

邻域内某点 r 与中心点 r_0 是否相似的判断函数, 它的计算方法为:

$$D(r, r_0) = \begin{cases} 1 & \text{if } |I(r) - I(r_0)| \leq t \\ 0 & \text{if } |I(r) - I(r_0)| > t \end{cases} \quad (5.21)$$

其中 $I(r_0)$ 为中心点像素灰度值; $I(r)$ 为中心点N邻域内某像素点灰度值; t 为点与点相似的判定阈值。

通过图像区域像素的相似度确定角点候选集的分析如下:

本文选择 3×3 的检测窗口, 对于中心像素点, 在下面的讨论中只考虑其8邻域内像素点的相似度。计算该邻域内的像素点与中心像素点的灰度值之差的绝对值(记为 $\Delta = |I(r) - I(r_0)|$), 如果该值小于等于设定的阈值(记为 t), 则认为该像素点与目标像素点相似。

从定义中可以看出: $0 \leq g(r_0) \leq 8$ 。现在讨论 $g(r_0)$ 值的含义, 初步确定该中心点是否可能为角点。

(1) $g(r_0) = 8$, 表示当前中心像素点的8邻域范围内都是与之相似的像素点, 所

以该像素点邻域范围内的梯度不会很大，该点肯定不是角点。

(2) $g(r_0) = 0$, 表示当前中心像素点的8邻域范围内没有与之相似的像素点，所以该像素点为孤立像素点或者是噪声点，该点肯定不是角点。

(3) $g(r_0) = 1$ 和 $g(r_0) = 7$ 时可能的角点只能是8邻域中和它不相似的点，故中心像素点不应该作为角点的候选点。

(4) $2 \leq g(r_0) \leq 6$, 情况比较复杂，我们无法从直观上确认该中心像素点准确的性质。故本文将这些点列入候选角点之中。

通过上述步骤，我们得到图像候选角点集，为后续的角度提取剔除了大量的非角点像素点，而不必计算每个像素点的响应函数值，大大减少了乘法的计算次数，使得计算效率得到很大程度上的提高。

在得到候选角点集后，我们采用Harris算法进行特征值提取[6]，其过程如下：

由于图像中的角点与自相关函数的曲率有关，所以求得自相关函数：

$$E(x, y) = \sum_{u,v} w_{u,v} |I_{x+u, y+v} - I_{u,v}| \quad (5.22)$$

式中， $E(x, y)$ 是两个图像窗口偏移 (x, y) 后图像灰度的平均变化， w 为图像窗口， I 为图像灰度。

在角点处时， $E(x, y)$ 随着窗口的偏移将有显著变化。 $E(x, y)$ 在像素点 (u, v) 处可近似表示为公式(5.23)：

$$E(x, y) = \begin{bmatrix} x & y \end{bmatrix} M \begin{bmatrix} x \\ y \end{bmatrix} \quad (5.23)$$

式中，矩阵 M 为：

$$M(x, y) = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 \otimes h(x, y) & \left(\frac{\partial I}{\partial x} \bullet \frac{\partial I}{\partial y}\right) \otimes h(x, y) \\ \left(\frac{\partial I}{\partial x} \bullet \frac{\partial I}{\partial y}\right) \otimes h(x, y) & \left(\frac{\partial I}{\partial y}\right)^2 \otimes h(x, y) \end{bmatrix} \quad (5.24)$$

其中， $h(x, y)$ 为一阶高斯平滑滤波函数。可以看出矩阵 M 具有如下形式：

$$M(x, y) = \begin{bmatrix} A(x, y) & C(x, y) \\ C(x, y) & B(x, y) \end{bmatrix} \quad (5.25)$$

像素点 (u, v) 处的 $E(x, y)$ 的极值曲率可由矩阵 M 的特征值近似表示。有以下三种情况：

- (1)两个特征值都很大，可以认为该点是一个角特征点；
- (2)两个特征值都很小，可以认为该点是一个面上的点；
- (3)两个特征值一个大，一个小，可以认为该点是一个边上的点。

矩阵 M 行列式的值正比于两个正交方向上的极值曲率的乘积，故定义Harris角点响应函数(CRF)为：

$$CRF(x, y) = \det[M(x, y)] - k \bullet \text{trace}^2[M(x, y)] \quad (5.26)$$

式中， $\det[M(x, y)] = AB - C^2$ ， $\text{trace}[M(x, y)] = A + B$ ，当 $CRF(x, y)$ 超过某一阈值时，就可认为 (x, y) 为角点。Harris建议 k 取值为0.04以获得比较好的结果。本文依次从以每个像素为中心的 3×3 的窗口中提取CRF值，如果中心点像素的CRF值大于给定的门限值，则该点就被判作特征点。

5.5.3 图像立体匹配

经过角点提取后，我们得到两幅图片里的匹配点集 P 和 Q 。设左图像某角点的图像坐标为 $p = (x_l, y_l)$ ，右图像某角点图像坐标为 $q = (x_r, y_r)$ ，取匹配窗口大小为 $(2n+1) \times (2m+1)$ 。则定义角点 p, q 的特征相似度为公式(5.27)：

$$R(p, q) = \frac{\sum_{i=-n}^n \sum_{j=-m}^m [I_l(x_l + i, y_l + j) - u_l] \times [I_r(x_r + i, y_r + j) - u_r]}{(2n+1)(2m+1) \sqrt{\sigma^2(I_l) \times \sigma^2(I_r)}} \quad (5.27)$$

式中， u_l 为左图像匹配窗口内的灰度平均值， u_r 为右图像匹配窗口内的灰度平均值， $\sigma(I_l)$ 为左图像匹配窗口内灰度的标准差， $\sigma(I_r)$ 为右图像匹配窗口内灰度的标准差，即：

$$u_l = \frac{\sum_{i=-n}^n \sum_{j=-m}^m I_l(x_l + i, y_l + j)}{(2n+1)(2m+1)} \quad (5.28)$$

$$u_r = \frac{\sum_{i=-n}^n \sum_{j=-m}^m I_r(x_r + i, y_r + j)}{(2n+1)(2m+1)} \quad (5.29)$$

$$\sigma(I_l) = \sqrt{\frac{\sum_{i=-n}^n \sum_{j=-m}^m I_l^2(x_l + i, y_l + j)}{(2n+1)(2m+1)} - u_l^2} \quad (5.30)$$

$$\sigma(I_r) = \sqrt{\frac{\sum_{i=-n}^n \sum_{j=-m}^m I_r^2(x_r + i, y_r + j)}{(2n+1)(2m+1)} - u_r^2} \quad (5.31)$$

由公式可知特征点相似度 $R(p, q)$ 取值在-1到1之间，如果角点 p, q 越相似，则

$R(p,q)$ 取值越大。

本文根据上述定义的相似性度量，采取极线约束和互应性约束[7][8]，在基于角点的正确提取上完成左右两幅图像的匹配。

极线约束即只在纵坐标相同的像素点进行匹配，由于摄像机的安装误差等因素的影响，对应匹配点可能不会准确地出现在图像平面中纵坐标相等的极线上，因此本文对第一幅图像某个特征点，在另一幅图像对应纵坐标正负五个像素的带状区域进行搜索匹配。

互应性约束指的是进行角点匹配时，首先对于左图像中的任一角点，寻找在右图像中和它相似度最大的角点，然后对于右图像中的任一角点，寻找在左图像中和它相似度最大的角点，当双向搜寻到的最大相似度的角点彼此对应且相似度大于某一阈值时，我们就认为这是一对正确的匹配点。反之，如果双向搜寻的角点不满足彼此对应，则认为该匹配不可靠，应该被排除。

为验证本文匹配算法的准确性，我们将经图像处理过后的标准图像进行特征点匹配，共找出48对匹配点，各点图像坐标如表2所示，图5-13显示出各匹配点对的连线。

表2 标准图像对特征点匹配结果

左图 纵坐标	左图 横坐标	右图 纵坐标	右图 横坐标	左图 纵坐标	左图 横坐标	右图 纵坐标	右图 横坐标
100	29	100	24	221	135	222	125
132	29	132	24	124	136	124	131
128	34	128	29	142	141	141	134
165	51	165	46	77	147	77	141
109	52	109	47	177	167	177	157
278	54	279	49	48	163	48	158
65	74	65	69	34	168	34	164
143	77	143	72	47	169	47	164
180	83	180	78	34	174	34	169
114	85	114	81	161	190	162	176
183	94	183	88	134	195	134	188
71	95	71	90	141	196	141	188
141	98	141	93	193	202	193	192
221	99	221	94	76	209	76	202
102	101	102	95	133	219	133	205
37	104	37	99	138	265	138	252
152	104	151	99	142	277	142	264
184	105	184	99	193	297	193	289
113	116	113	110	164	304	164	296
254	121	255	110	152	305	152	297
152	117	151	111	157	305	159	297
161	125	162	114	193	305	193	297
77	126	77	120	161	318	162	310
253	132	253	122	96	374	96	369



图5-13 标准图像对的本文匹配算法仿真结果

从结果中可以看出，此次匹配的正确率达到 100%。同时我们也对图 5-5 中用于标定的盒子的匹配结果进行了分析，共得到 30 对匹配点，其坐标如表 3 所示。由于光照等误差影响，匹配结果无法达到完全正确，但可达到 80%以上正确率。

表3 拍摄图像对特征点匹配结果

左图像 纵坐标	左图像 横坐标	右图像 纵坐标	右图像 横坐标	左图像 纵坐标	左图像 横坐标	右图像 纵坐标	右图像 横坐标
115	86	114	75	137	163	139	237
127	120	126	90	98	165	100	225
114	121	113	90	111	166	110	136
110	135	108	104	26	214	25	203
132	135	130	104	119	223	118	213
90	137	89	106	179	225	179	205
161	137	163	106	29	226	29	216
15	140	15	117	118	233	118	222
94	141	93	111	180	235	180	225
25	143	24	132	20	243	19	233
95	150	94	118	31	245	30	235
127	154	126	123	152	245	153	233
91	157	90	126	119	246	119	235
115	158	114	127	179	246	179	246
149	162	148	127	14	250	13	239

6. 模型的实例验证

我们用 20ml 永龙牌正红花油外盒作为验证算法的目标物体，将数码相机和红花油放置在水平桌面，两个相机间距为 10cm，用平行双目法拍摄得到的原始图像如图 6-1 所示。



左图像



右图像

图 6-1 原始双目图像

将相机图像导入电脑，压缩到指定像素后，利用 Matlab 编程进行图像处理，得到的灰度图如图 6-2。



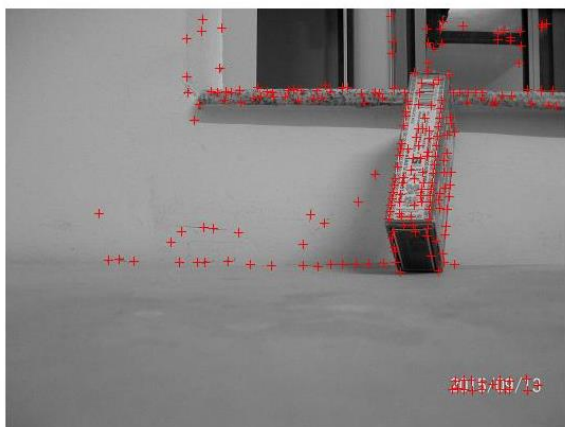
左图像



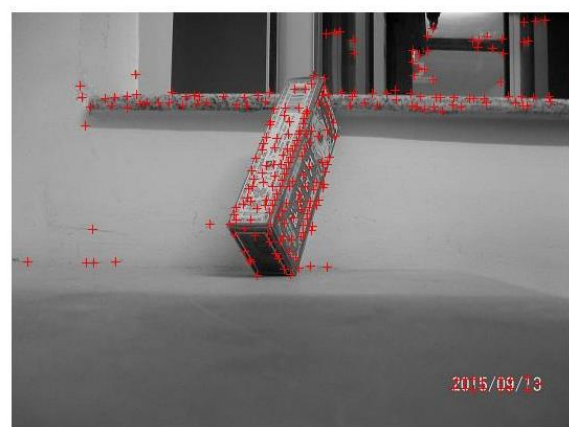
右图像

图 6-2 灰度化后图像

用 5.4.2 中的特征点提取算法对左右图像分别进行特征点提取，得到的特征点如图 6-3 所示。



左图像



右图像

图 6-3 特征点提取

进行图像匹配：

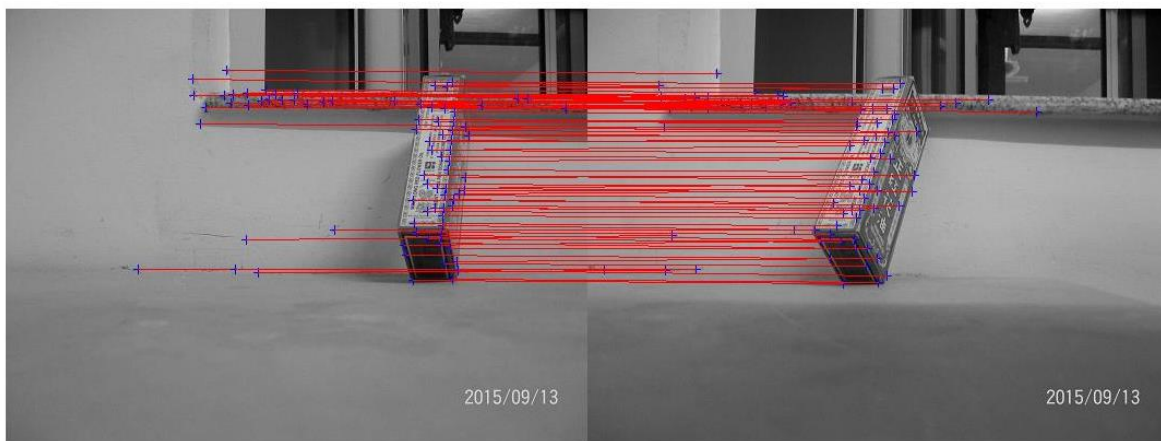


图 6-4 图像匹配

上图中匹配得到了 67 对特征点，由于目标物体是照片中间的盒子，所以其他部分的匹配特征点为额外信息。为减小计算量，将匹配进行优化，只保留目标物体上的匹配特征点，共 46 对，如图 6-5 所示。

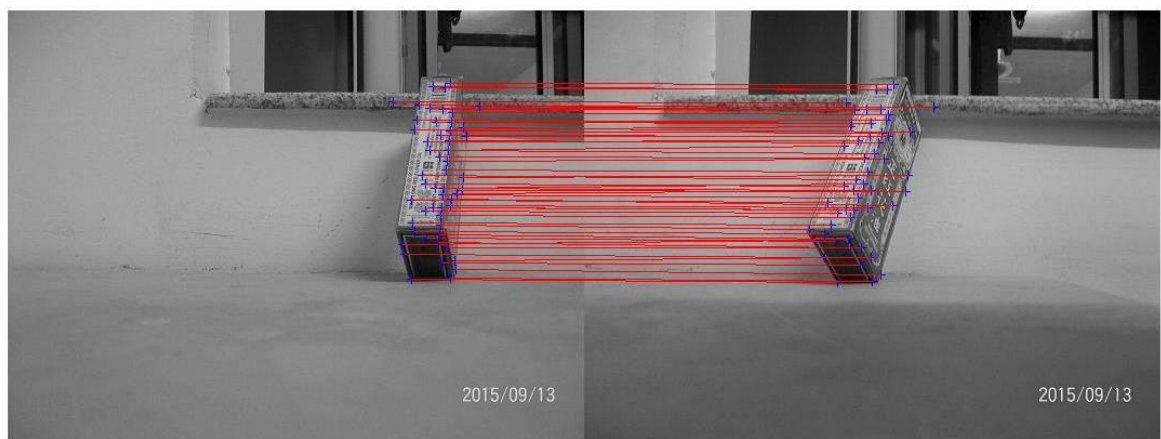


图 6-5 图像匹配优化

将匹配得到的特征点按照重构公式求得其三维坐标，就可以在 Matlab 中看到其点云信息，恢复三维结构。

此处以顶点为例，简述三维重建过程。由于实验中只利用了一组照片，因此只能得到 7 个顶点，第 8 个顶点被遮挡住。顶点匹配如图 6-6 所示。



图 6-6 顶点匹配

如图 6-6 所示，对匹配的顶点进行编号，依此得到左右两幅图中 7 个点的像素信

息，代入 5.3 节的公式，计算其三维世界坐标，具体数值见表 4。

表 4 目标物体顶点坐标信息

顶点	左图像	右图像	三维世界坐标
1	(384,63)	(264,63)	(-9.83,8.61,-46.67)
2	(421,63)	(299,63)	(-12.7,8.47,-45.9)
3	(427,108)	(314,108)	(-14.25,4.92,-49.56)
4	(358,205)	(206,205)	(-6,0.25,-36.84)
5	(406,205)	(249,205)	(-8.92,0.25,-35.69)
6	(375,252)	(237,252)	(-7.9,-2.84,-40.58)
7	(414,252)	(273,252)	(-10.5,-2.78,-39.72)

依照上表数据在三维空间中恢复得到的物体结构信息如图 6-7 所示，进行旋转可看到重建出的完整坐标信息。

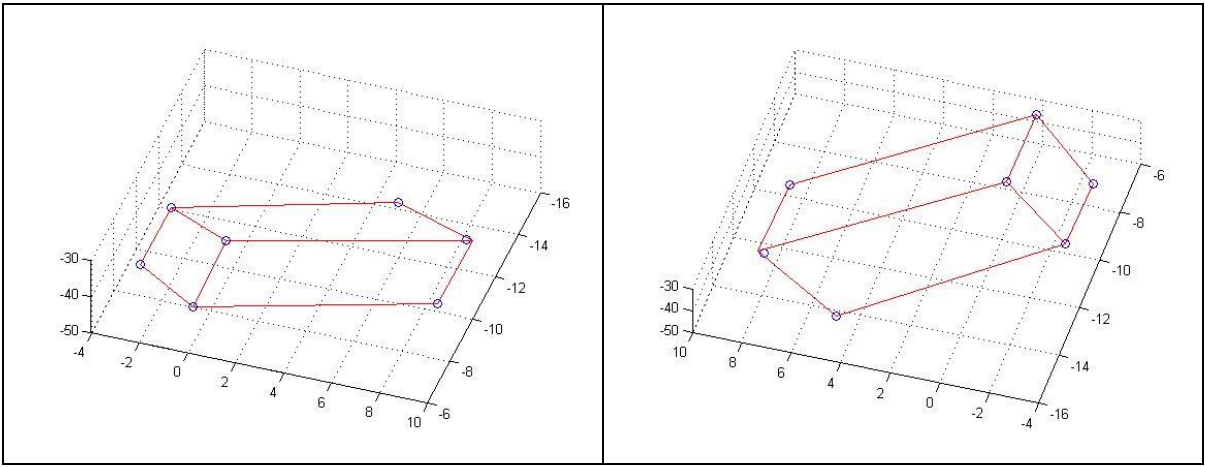


图 6-7 重建得到的三维结构信息

至此已完成了三维结构的重构，最后一步就是对目标物体的外观信息进行三维重建，即仿射变换和纹理映射。

纹理映射需将左右图像的灰度图中的面进行截取，通过仿射变换得到每个面的矩形变换图，将得到的矩形图按坐标匹配到三维结构图中，即可得到目标物体最终的三维重建图，结果如图 6-8、6-9 所示。



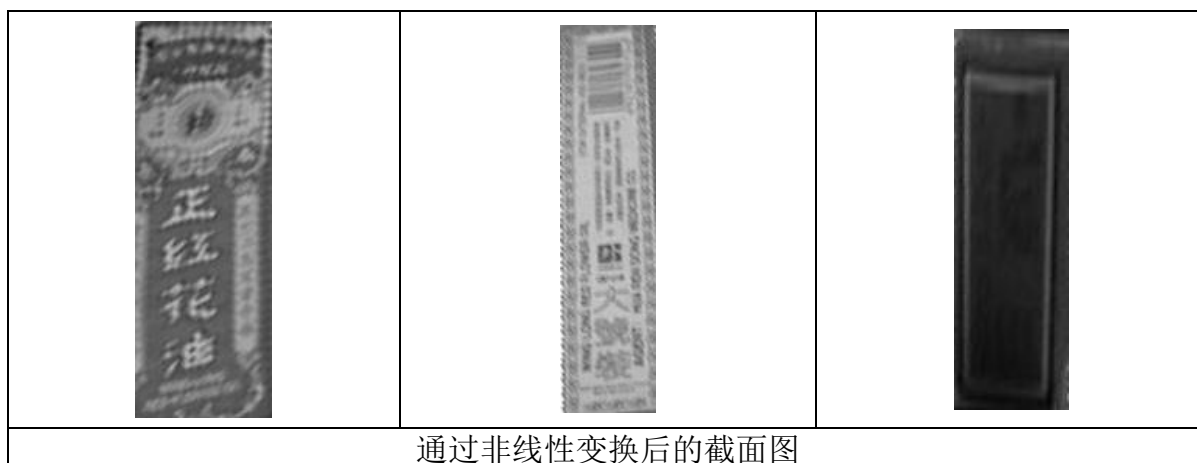


图 6-8 目标物体面变换

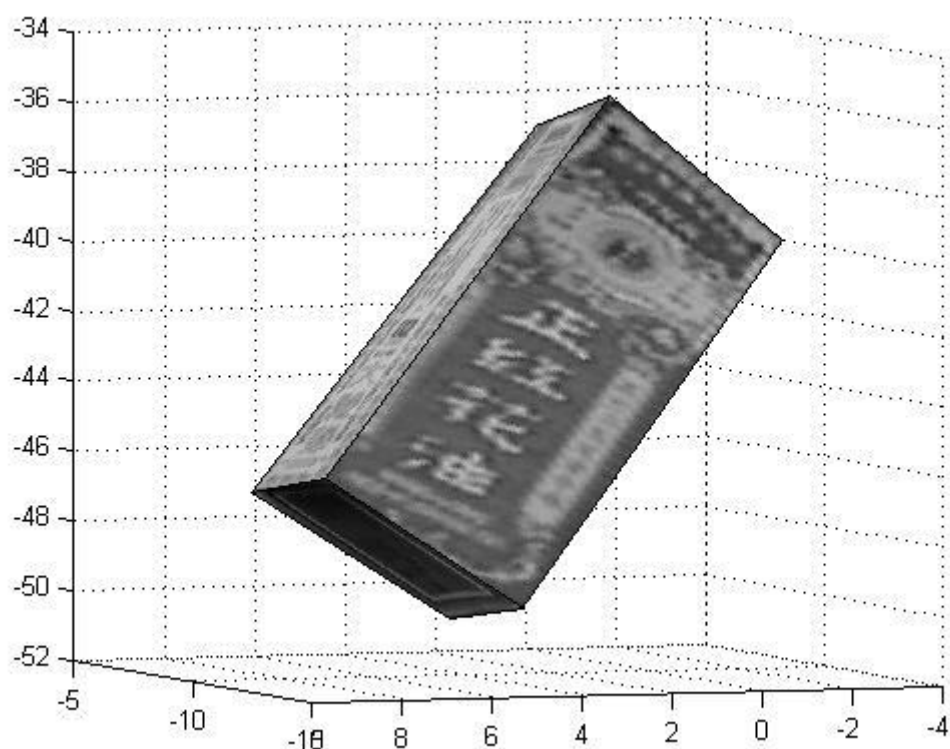


图 6-9 三维重建结果

本次三维重建的误差分析如下：

通过重建算法得到的三维点信息，我们可以确定目标物体的长宽高（可根据表 4 数据计算得到），用尺子可测得实物的长宽高信息，具体数值及误差见表 5。

表 5 误差分析

	长(cm)	宽(cm)	高(cm)
实物测量	13.6	5.0	2.8
重构信息	13.64	5.14	2.73
误差	0.29%	2.8%	2.5%

通过上表计算得到平均误差为 1.86%。从重建结果及误差可以看出，本文所提模型可以很好的还原目标物体的三维信息，实现三维重建，是一个简易可行的 3D 扫描

系统方案。

7. 技术指标分析

1) 设备型号及参数

本文中采用柯达 M320 数码相机，传感器类型为 CCD，最大像素数 930 万，最高分辨率为 3472×2604 。

测距工具为米尺，最小精确单位为毫米(mm)，计算与测量时四舍五入到厘米(cm)。

算法运行仿真平台为 Matlab 2013，处理器为 Intel 酷睿 i5 450M。

2) 算法复杂度

本文提出了一种快速确定候选角点集的方法，因此在特征点提取时可以大大减少算法复杂度。

Harris 算法计算了图像中的每一个像素点的 CRF 值。该算法采用 9×9 的矩形高斯窗口(共 81 个像素)，方向导数采用 Prewitt 算子，则在计算每个像素的自相关矩阵 M 过程中乘法次数为： $(9 \times 9 + 1) \times 3 = 246$ 次，加法次数为： $(9 \times 9 - 1) \times 3 = 240$ 次，因此，计算点(x,y)的 CRF 就要花进行乘法 $246 + 2 = 248$ 次，加法 $240 + 1 = 241$ 次。对于分辨率为 $X \times Y$ 的图像，不考虑边界的影响，要进行乘法 $248 \times X \times Y$ 次，加法 $241 \times X \times Y$ 次。由于一次乘法运算所耗时间要远大于一次加法所耗时间，所以 Harris 算法的时间复杂度较高，计算速度较慢。

本文的改进算法主要是在计算 CRF 值之前对像素点做了一个初始选择，得到候选角点集。这一部分主要是加法运算，由于选用 3×3 检测窗口，运算量为 $(3 \times 3 - 1) \times X \times Y$ 次。由于不涉及乘法运算，为后续的角度提取剔除了大量的非角点像素点，而不必计算每个像素点的 CRF 值，大大减少了乘法的计算次数，计算效率在分辨率高的图像中得到显著提高。

3) 系统的精确度

设备精度：系统最大可处理分辨率为相机最高分辨率，本文中所采用的相机最高分辨率为 3472×2604 ，若更改相机类型，分辨率相应改变。

匹配精度：忽略拍摄时光线分布不均匀造成的左右图像灰度值误差，本文应用的匹配算法正确率可达到 100%。

三维重建精度：系统中特征点提取的数量是根据图像复杂度和灰度相似度的改变而自适应的，特征点匹配对数越多，三维重构时恢复的精度越高。本文构造的实例中，目标物体匹配点数为 46 对，重构后的尺寸与真实物体的误差仅为 1.86%。

4) 扫描尺寸限制

限制可扫描目标尺寸的因素主要为硬件因素。

平行双目模型所采用设备仅为两个性能相同的照相机，忽略测量误差，相机参数是唯一需考虑的硬件因素。实验中所采用的相机为智能数码相机，拍摄一般物体时可实现自动对焦，其对焦范围为 600mm-无穷远；需要拍摄近距离物体时，可通过切换模式实现 100mm-700mm 范围内对焦。故理论上可在 100mm-无穷远处实现对焦，得到物体的像素信息。

相机中另一个需要考虑的因素为分辨率，实验用相机最高分辨率为 3472×2604 ，当图像中包含的信息过多时，由于分辨率的限制，相机会压缩一部分信息，即造成像素的丢失。

综上，当目标物体尺寸极小时，若在 10cm 内拍摄则无法实现聚焦，在 10cm 外拍摄则无法得到有效特征点；当目标物体尺寸过大（如建筑物）时，为了拍摄物体的全部信息，由于像素的限制，导致局部信息被压缩，无法准确得到匹配信息，这两种

情况均无法实现重构。考虑到现实生活中肉眼可见的物体均可在 10cm 处捕获其特征，故近似认为该方法对物体尺寸的最小范围无限制。

扫描尺寸的限制根据硬件设施而变化，总体而言，本方法适用于小尺寸目标的三维重建。

8. 模型的评价与改进方向

8.1 模型的评价

平行双目模型是一种模拟人眼原理进行目标三维重建的方法，该方法操作简单，设备常见易得，成本低，且携带方便，可以随时在户外进行目标拍摄，实现三维重建，对周围环境和拍摄背景均无要求。本文所采用的特征提取算法具有运算复杂度小，速度快的特点；图像匹配算法的准确度高，容错率高，可以很好地对特征点进行匹配。从实例可以看出，通过模型可以直接得到目标物体的三维点信息，进而重构三维结构信息及外观信息。

然而，该方法是一种被动式非接触的三维重建方式，不适用于大尺寸物体的测量与重建。另外，模型的精度受相机像素限制。其主要不足之处在于，由于需要得到物体的照片的灰度信息以提取特征点，因此不适用于测量透明物体，反光性较强的物体，这也是被动三维重建相较于主动激光法的不足之处。

8.2 模型的改进方向

由于时间原因，本文仅构造了对规则物体进行三维重建的实例来验证模型。该实例利用相对少量特征点即可得到完整信息，简化了重建的复杂度。为使该方法有更广的适用范围，下一步工作需探究对于复杂物体的重建，细化特征点的提取，优化线面构建方法。

另外由于在实验中，两个相机的平行放置是人为标齐的，并没有专业工具，容易造成细微误差，在以后的实验中应规范拍摄及测量步骤，避免人为因素造成的误差。

参考文献

- [1] Kenneth. R. Castleman 著，朱志刚等译，数字图像处理[M]，北京：电子工业出版社，2009。
- [2] 高文，陈熙霖，计算机视觉算法与系统原理[M]，北京：清华大学出版社，1999：76-78。
- [3] 金德闻，Terrain Identification for prosthetic Knees Based on Electromyographic Signal Features[J]，清华大学学报自然科学英文版，2006(11)：74-79，2006。
- [4] 吴若鸿，基于特征匹配的双目立体视觉技术研究[D]，武汉科技大学，2010。
- [5] 张虎，机器视觉中二维图像的三维重建[D]，北方工业大学，2006。
- [6] Harris C, Stephens M, A combined corner and edge detector[J]. Proceedings of the Fourth Alvey Vision Conference, Manchester, 1988。
- [7] Y. Dufournaud, C. Schmid and R. Horaud, Matching images with different resolutions. In: Proceedings of the Conference on Computer Vision and Pattern Recognition[C], Hilton Head Island, South Carolina, USA, 2000。
- [8] 高文娟，双目视觉中的立体匹配技术研究[D]，陕西科技大学，2009。