

三维物体的表面重建

李亚杰, 王锴, 王以桢, 麻文杰

摘要 —估计二维图像中像素点对应的空间坐标, 具有重要的理论和应用价值。本文参考 Bouguet 等人的三维重建方法, 使用台灯, 签字笔, 相机, 标定板相继实现了相机的标定, 光源位置确定。在光源下移动签字笔, 使得签字笔的阴影扫过桌面上的三维物体, 获得多帧图像。图像序列包含时域信息, 每帧图像具有不同的参考点(阴影边缘), 可以提取空间信息。利用时域与空域信息, 即可求取二维图像中像素点对应的空间坐标。整个装置构建容易, 操作简便, 精度较高。

关键词 —标定, 插值, 三维重建.

I. 简介

三维重建是利用计算机建立对客观世界真实物体描述的关键技术^[1], 它是对三维场景的恢复, 也可以看作场景图像描述的逆过程, 即由二维的图像还原出三维的立体信息。基于图像的三维重建是从若干张图片提取场景点的深度信息, 再根据这些信息重建出场景的三维模型。

目前, 三维重建的方法主要有单目立体视觉, 双目立体视觉, 多目立体视觉, 光度立体视觉, 结构光立体视觉, 基于运动的三维重建等^[2]。双目立体视觉方法主要是模拟人类双眼的立体成像原理, 左右两边各有一台摄像机在同一时间拍摄物体, 利用三角关系和视差原理, 获得两个摄像机视角下物体表面点的坐标信息, 从而构建物体的位置和形状。多目立体视觉解决了双目立体视觉两幅图像进行像素匹配容易出错的问题, 但硬件成本更昂贵, 系统更复杂。光度立体视觉方法中, 利用摄像机拍摄物体在不同方向光照条件下的同一处表面, 利用获得的这组图像来重建物体表面形状。结构光立体视觉方法使用标准的光栅条纹结构光投射到物体表面, 投影光条跟随物体表面形状的起伏而发生变化, 摄像机拍摄物体表面图像, 从被物体表面形状所调制的条纹模式中, 提取出物体的三维信息。

本文使用的方法基于单目视觉, 先后进行了相机标定, 光源位置确定, 再利用几何关系, 求取二维图像中的像素点对应的场景点的空间坐标。目前的工作, 扫描时只用了一个视角, 因而只对物体的部分表面进行了三维重建。若需要进行完整的重建, 则要变换视角进行扫描, 得到不同角度的点云数据, 再进行点云配准和数据融合。

II. 实现方法

整个处理过程包含相机标定, 点光源位置确定, 利用几何关系求解像素点对应场景点的空间坐标。

A. 相机标定

相机标定所做的工作是寻找对象在图像与现实世界相互转换的数学关系, 从而实现从图像中测量出实际空间数据的目的。通过相机标定, 我们可以获得相机的内部参数, 包含焦距 fx , fy , 图像坐标系原点在像素坐标系下的表示 (cx, cy) , 可利用这些参数将图像坐标系中的点转换为相机坐标系下的表示。外部矩阵包含旋转矩阵 R 和平移矩阵 T , 可利用这层关系将相机坐标系中的点转换为世界坐标系下的表示。目前使用最广泛的标定方法是棋盘格标定法, 利用角点的透视不变性, 将角点在图像中的坐标与实际空间中(标定板坐标系)坐标相互对应, 可以求得相机的内外参数矩阵。实验中采用 28 角点棋盘格, 处理 20 张图片, 保证标定得到的内部矩阵参数的准确度。启动相机, 开始捕捉图像。以 10 帧为单位, 每计数到 10 帧图像, 处理一次, 故 20 张图片需要采集 200 帧。调用 opencv 函数寻找棋盘角点, 计算角点数据, 并进行标定。最终可得到一组内参矩阵和 20 组不同的外参矩阵, 外参矩阵取棋盘平面与桌面平面相同时的参数。表格 1 列出了最终求解的相机参数。取第一张图的外部参数矩阵:

表 I 相机参数

参数	数值
fx	1641.7
fy	1399.5
cx	0
cy	0

$$R = \begin{bmatrix} -0.0390 & -0.0448 & -0.9982 \\ -0.9957 & 0.0857 & 0.0351 \\ 0.0840 & 0.9953 & -0.0480 \end{bmatrix} \quad (1)$$

$$T = \begin{bmatrix} 86.4894 \\ 682.1272 \\ 828.1897 \end{bmatrix} \quad (2)$$

B. 点光源位置确定

通过前面的相机标定，我们获得了相机的参数矩阵。给定世界坐标系 xoy 平面上的不共线的三点，使用公式 (3) 将它们变换成相机坐标系下的表述：

$$[X_c \ Y_c \ Z_c]^T = R * [X_w \ Y_w \ Z_w]^T + T \quad (3)$$

这三个点可组成三个向量。任取其中两个向量，求与它们正交的法向量，该法向量也与桌面平面垂直。此时可确定桌面平面的方程。固定光源，将签字笔立在桌面上，不断变换位置，拍摄影子图。实验中使用了四张图片，标出签字笔影子的顶端与底端，获取并保存这些点的图像坐标。一共 4 组，共 8 个点。

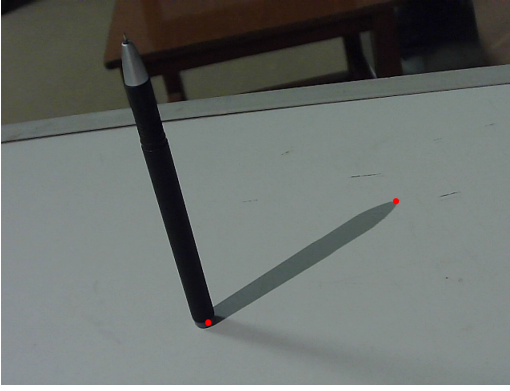


图 1 标记端点

标记的 8 个点的坐标均为像素坐标系下的坐标，通过公式 4 转化为相机坐标系下的描述：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} dx & 0 & 0 \\ 0 & dy & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} - \begin{bmatrix} cx \\ cy \\ 0 \end{bmatrix} \quad (4)$$

已知上述标记点在摄像机坐标系下的坐标，分别将它们与相机坐标系原点连接，得到对应的向量，并求解这些向量与桌面平面的交点。为方便理解，我们进行下面的分析：

$\triangle ABC$ 所在平面为桌面平面， S 为点光源， $M'N'$, $M'N'$ 为不同位置下的签字笔， MB , $M'C$ 分别为对应的笔的阴影。前面所求的交点即为 M, B, M', C 。由于桌面平面方程已知，故可求得 N, N' 在相机坐标系下的坐标。此时，有向量表达式：

$$\overrightarrow{N'C} = \overrightarrow{M'C} - \overrightarrow{M'N'} \quad (5)$$

$$\overrightarrow{NB} = \overrightarrow{MB} - \overrightarrow{MN} \quad (6)$$

$$\overrightarrow{BC} = \overrightarrow{OC} - \overrightarrow{OB} \quad (7)$$

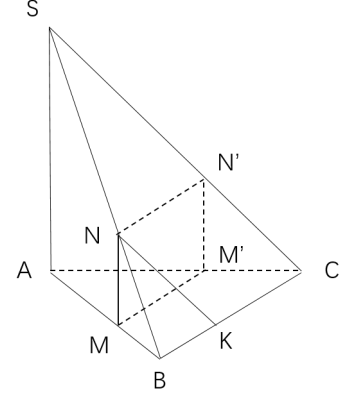


图 2 光源构成的几何关系

作 SA 平行于 $N'M'$ ，故 S, A, N', M' 四点共面。由于 $N'M'$ 垂直于底面 ABC ，且平行于 MN ，故 S, A, M, N 四点共面，从而可知 $\triangle SAB$ 相似于 $\triangle NMB$ ， $\triangle SAC$ 相似于 $\triangle N'M'C$ ，得到 $\triangle AMM'$ 相似于 $\triangle ABC$ 。因为四边形 $NN'M'M$ 为矩形，故 NN' 平行于 BC ，作 NK 平行于 $N'C$ ，可得如下关系：

$$\overrightarrow{BK} = \overrightarrow{N'C} - \overrightarrow{NB} \quad (8)$$

记 $d1$ 和 $d2$ 分别为向量 \overrightarrow{BK} 和 \overrightarrow{BC} 的模长。由公式 (9) 可确定光源位置。

$$\overrightarrow{OS} = \overrightarrow{OC} - \frac{d2}{d1} * \overrightarrow{N'C} \quad (9)$$

为了提高估计的准确性，采用多种方法求解点光源 S 的位置，利用组合学思想分别计算，如任取两张图片为一组，最后取平均结果。

C. 场景点空间坐标求解

此处所指的空间坐标是在相机坐标系下的描述，由于坐标系间的转换关系已知，故也能容易地转换为世界坐标系下的表达。启动相机，采集图像。从采集的第一帧中获取图像的大小，设置上下参考行的序号，最大处理帧数及对比度阈值。初始化参考点的列序号为 -1。对每一帧求上下参考点 (本文求的是左参考点)，即扫描上下参考行，当在某个像素点处的对比度超过设定的阈值时，计算 $Ishadow$ ，若该像素点的值小于 $Ishadow$ ，记录此时该点所处的列序号 [3]。该像素点的值比前面一列的像素点的值要小，其 $\Delta I < 0$, ΔI 和 $Ishadow$ 的定义式分别为：

$$\Delta I = I(x, y, t) - Ishadow \quad (10)$$

$$Ishadow = \frac{Imax + Imin}{2} \quad (11)$$

对于固定的一帧, t 为固定值, y 代表上参考行或下参考行, I 为 x 列对应的像素值。参考点前一列像素的 $\Delta I \geq 0$, 使用线性插值方法, 求解亚像素值。以此值作为最终的参考点 [3]。

从时域角度看, 阴影会先后扫过整个物体表面, 相机也在不断捕捉图像。假设点 (x, y) 是物体表面上任意一点的像素坐标, 与上面类似, 计算该点的对比度是否超越阈值, 符合条件则计算 I_{shadow} , 若该点像素值大于 I_{shadow} , 且前一帧图像相同位置点的像素值小于 I_{shadow} , 则在这两帧间进行线性插值, 获得亚帧数, 将此亚帧作为最终的参考时刻。为求得此时刻对应的参考点, 取两帧图像, 利用插值思想即可, 计算公式如下:

$$K = \frac{X_{top}(posttime) - X_{top}(pretime)}{posttime - pretime} \quad (12)$$

$$X_{top} = K * (time - pretime) + X_{top}(pretime) \quad (13)$$

其中, X_{top} 为参考帧对应的上参考点, $pretime$ 为前一帧对应的时刻, $pretime$ 取 $int(time)$, $time$ 代表当前亚帧, $posttime$ 表示下一帧对应的时刻, $X_{top}(posttime)$ 表示下一帧代表的参考点, $X_{top}(pretime)$ 表示前一帧代表的参考点, 同理, 下参考点的计算也是类似的。给定场景点对应的上下参考点计算出来后, 令上参考点为 A , 下参考点为 B , 点光源为 S , S, A, B 构成了一个平面, 示意图如下:

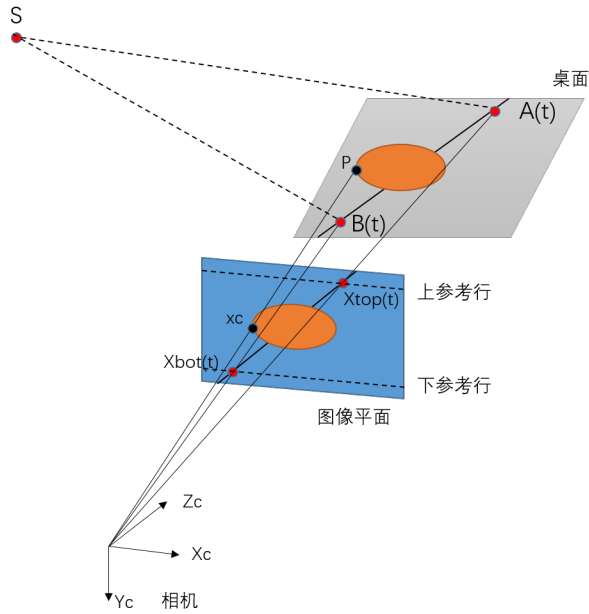


图 3 场景点求解

已知三点坐标, 可求解 S, A, B 所在平面的方程。对于给定场景点 (图中标出的 x_c), 有向量 $\overrightarrow{O_c X_c}$, 求解该向量

所在直线与平面 SAB 的交点 P 的坐标, 即可得到实际空间点在相机坐标系下的表示。类似地, 凡是可以扫描到的物体表面上的所有点, 都可以按此方法求解。将所有求出的点的三维坐标保存起来, 即为点云数据。

这里, 我们使用 python 的 mayavi 库绘制三维图, 效果如图:

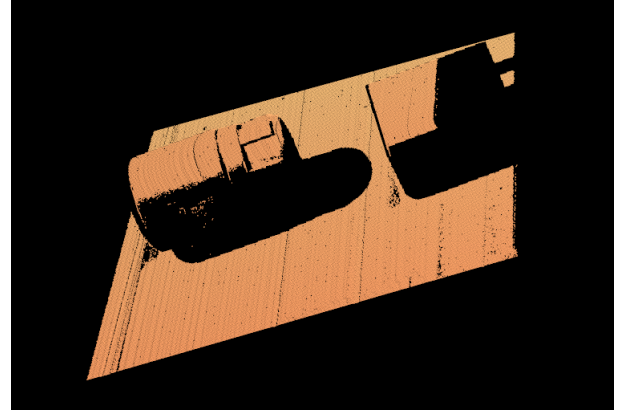


图 4 三维表面重建效果图

III. 实验结果与分析

我们搜集了相机的部分参数, 包括焦距, 畸变参数, 同时手工测量了光源高度, 被扫描物体的高度, 与程序估计的值进行比较, 并计算对应的误差。

表 II 实验结果

参数	估计值	误差
f	21.29mm	6.45%
LH	476.8mm	4.64%
OH	28.34mm	3.93%

从结果来看, 误差控制在 7% 以内, 具有较好的准确度。对于产生偏差的因素, 我们认为噪声是其主要来源, 由于实验光线暗淡, 相机捕捉的图像具有相当多的噪点, 明显降低了图像质量。另外, 我们未对获得的图像进行预处理, 在求解参考点时, 也未使用低通滤波器进行噪声滤除, 这些都直接影响了最后的估计精度。

目前, 由于只使用一个摄像机, 这种方法只能重建一个固定视角下的物体表面, 且对光照条件的要求较高, 亮度不足和遮挡的部分将无法进行重建。

IV. 未来工作

我们将以削减噪声为目标, 探索将重建图质量提升到更高水准的方法。为了获得完整物体表面的三维重建, 以不同角度扫描图像, 获取各视角的点云数据, 再进行点云配准与数据融合。

参考文献

- [1] 孙向军. 场景三维重建的若干关键技术研究 [D]. 南京: 南京理工大学, 2004.
- [2] 王克鸿, 杨嘉佳, 孙科. 基于视觉的焊接三维重建技术研究现状 [J]. 机械制造, 2013: 41(1): 1-5.
- [3] Jean-Yves Bouguett, Pietro Peronat. 3D photography on your desk[J]. Electrical Engineering, 2007.