

引用格式:李文浩,贾同,吕朝辉,孙小钧,黄俊文,张松娜.大场景三维扫描与建模方法研究[J].中国传媒大学学报(自然科学版),2022,29(02):79-88.

文章编号:1673-4793(2022)02-0079-10

## 大场景三维扫描与建模方法研究

李文浩<sup>1</sup>,贾同<sup>1\*</sup>,吕朝辉<sup>2</sup>,孙小钧<sup>1</sup>,黄俊文<sup>1</sup>,张松娜<sup>1</sup>

(1.东北大学信息科学与工程学院,沈阳 110819;2.中国传媒大学信息与通信工程学院,北京 100024)

**摘要:**三维扫描与建模是一种能够获取目标场景三维信息的重要立体表示方法,在观演空间舞台效果展示、虚拟看房、文物保护等领域具有广泛应用。本文主要研究了一种三维扫描与建模方法,包括三维深度感知、点云配准与三维点云建模等内容。首先提出了全向环结构光深度感知方法;然后建立了混合点云配准模型,粗配准采用穷举搜索与特征匹配算法,精配准采用基于双向匹配与比率筛选的ICP(Iterative Closest Point, ICP)配准算法;最后通过Geomagic等软件对点云进行主动建模。实验结果表明,本文方法可以获得高精度、高鲁棒性的大场景三维扫描与建模结果。

**关键词:**深度感知;点云配准;三维建模;结构光

中图分类号:TP231 文献标识码:A

## Research on 3D scanning and modeling method of large scene

LI Wenhao<sup>1</sup>, JIA Tong<sup>1\*</sup>, LV Zhaohui<sup>2</sup>, SUN Xiaojun<sup>1</sup>, HUANG Junwen<sup>1</sup>, ZHANG Songna<sup>1</sup>

(1.Northeastern University, Shenyang 110819, China; 2. Communication University of China, Beijing 100024, China)

**Abstract:** 3D scanning and modeling is an important three-dimensional representation method that can obtain the three-dimensional information of the target scene. It is widely used in the fields of performance space stage effect display, virtual house viewing, cultural relics protection and so on. This paper mainly studies a 3D scanning and modeling method, including 3D depth perception, point cloud registration and 3D point cloud modeling. Firstly, an omni-directional ring structured light depth sensing method is proposed. Then a hybrid point cloud registration model is established. The coarse registration adopts exhaustive search and feature matching algorithm, and the fine registration adopts iterative closest point(ICP) registration algorithm based on two-way matching and ratio screening. Finally, the point cloud is actively modeled by Geomagic and other software. Experimental results show that this method can obtain high-precision and robust 3D point cloud scanning and modeling results of large scenes.

**Keywords:** depth perception; point cloud registration; 3D modeling; structured light

基金项目:观演空间静态感知技术与系统研发(2018YFB1404101);国家重点研发计划(20200197);

自适应环境结构光生成方法研究(62173083);国家自然科学基金(20210379)

作者简介(\*为通讯作者):李文浩(1993-),男,博士研究生,主要从事结构光深度感知。Email:1910274@stu.neu.edu.cn;贾同(1975-),男,博士,教授,主要从事计算机视觉、模式识别、机器学习等方面的研究。Email:jiatong@isc.neu.edu.cn

## 1 引言

随着计算机视觉的快速发展,三维视觉相比二维视觉具有“降维打击”优势,已在观演空间舞台效果展示<sup>[1]</sup>、虚拟看房<sup>[2]</sup>、文物保护<sup>[3]</sup>等多领域进行应用。根据成像方式分类,目前获取场景三维信息可分为主动视觉与被动视觉两类方法。主动视觉主要包括结构光<sup>[4][7]</sup>与TOF(Time of flight)<sup>[5]</sup>技术等。被动视觉主要包括双目与多目视觉技术<sup>[6]</sup>等。

近年来,结构光已成为三维视觉的主流技术。其主要原理为将编码光投射至物体表面,再使用相机接收该物体表面反射的结构光图案。由于接收图案必会因物体的立体形状而发生变形,故可以通过该图案在相机上的位置和形变程度来计算物体表面的空间信息。结构光与TOF、双目视觉<sup>[32]</sup>等技术比较而言,具有精度高、非接触等特点,在三维扫描、精准识别以及形状测量等方面表现出了优异效果。

根据时空编码方式,结构光可分为时间编码和空间编码。时间编码方式主要包括多步相移法和格雷码法等。空间编码主要包括点、线、面结构光等。其中,点和线结构光<sup>[8]</sup>可以直接通过相似三角形进行测距。面结构光<sup>[9]</sup>需要进行编解码后,才能进一步算出距离,主要有M阵列<sup>[9]</sup>、随机散斑编码<sup>[10]</sup>等。

获取场景三维信息后,即可转化为空间点云信息。由于现有技术不能单次获取大场景的全部点云信息,因此需要对多帧点云进行配准与重建<sup>[3]</sup>。点云配准分为粗配准和精配准两步,粗配准算法可分为两类,一种是基于穷举的配准算法,如RANSAC(Random Sample Consensus)配准算法<sup>[11]</sup>、四点一致集配准算法(4-Point Congruent Set, 4PCS)<sup>[12]</sup>、Super4PCS算法<sup>[13]</sup>等;另一种是基于特征匹配的配准算法,如基于点FPFH(Fast Point Feature Histograms)特征的SAC-IA(Sample Consensus Initial Alignment)<sup>[14]</sup>、以及基于线特征的ICL(Iterative Closest Line)<sup>[15]</sup>算法等。精配准主要采用PNP(Perspective N Point)<sup>[16]</sup>、ICP(Iterative Closest Point)<sup>[17]</sup>及其改进算法等。

ICP可分为阈值法和单向邻近匹配法。阈值法是通过设定的阈值对点云筛选的一种方法,原理相对简单。单向邻近匹配是一种比较常用的匹配算法,同时也是ICP算法在搜寻点对过程中的主要方法。

点云配准之后,需要对点云进行建模才能获得完整的大场景模型。三维点云建模通常有两种方式,一种是三维建模软件主动建模,即将点云导入软件中进

行建模,如Geomagic软件系统等;另一种是点云逆向建模,主要有泊松算法<sup>[18]</sup>及其改进算法、Delaunay Triangulation<sup>[19]</sup>、Slicing-based method<sup>[20]</sup>等算法。

本文对大场景扫描、点云配准和点云建模的相关技术进行了介绍,并采用改进算法实现大场景三维扫描与建模,获得了高精度、高鲁棒性的三维重建结果。

## 2 三维深度感知方法

如前所述,结构光是主动深度感知视觉的主要方法。本文提出一种基于全向环结构光<sup>[21]</sup>的深度感知方法,可归类为线结构光范畴<sup>[8][9][10]</sup>。

### 2.1 全向环结构光深度测量系统

如图1所示为全向环结构光深度测量模型<sup>[21]</sup>。图中显示了由全景相机和全向环结构光组成的深度测量单元,其中S为双曲面焦距到激光平面的垂心距离,下面简称为基线距离S。基线距离由机械测量确定,决定了最终的测量精度。

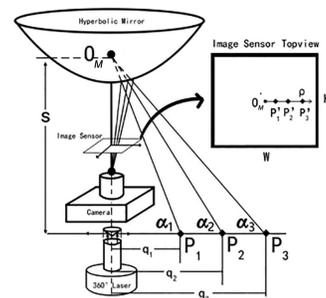


图1 全向环结构光深度测量模型

图中描述了空间点与图像平面的映射关系,给出了不同空间测量点在平面内的成像情况,以及测量深度与各参数之间的关系。图中 $O_M$ 为双曲面镜的焦点,如果能在相机平面内成像则所有入射光线都交于焦点 $O_M$ <sup>[21]</sup>。图中省略了相机的镜头,使用透视投影的方式来描述空间几何关系,ImageSensor表示相机内的感光元件,Camera所指位置表示感光元件所在位置,感光元件所成图像如ImageSensor Topview所示,其中H表示图像的行像素,W表示图像的列像素。

空间中的点 $P_1, P_2, P_3$ 在激光平面中的测量深度分别为 $q_1, q_2, q_3$ ,  $\alpha_1, \alpha_2, \alpha_3$ 为测量点 $P_1, P_2, P_3$ 到双曲面焦距与激光所在平面的夹角。

当测量物体挡住激光的通路时,物体表面反射激光光线到双曲面镜,通过相机镜头将图像聚焦在图像传感器上。如图中可以看到不同的测量位置 $P_1, P_2,$

$P_3$  在相平面上对应位置为  $P'_1, P'_2, P'_3$ 。其中  $O_M$  为双曲面镜的焦点,  $O'_M$  为过双曲面镜焦点垂直于传感器平面的交点, 后面称为中心投影。由图可知, 测量深度  $q$  可由  $S$  和  $\cos \alpha$  的乘积得到。

由此关系式可以得出测量深度  $q$ 、基线距离  $S$  和激光平面与被测点和双曲面镜焦距连线的夹角  $\alpha$  之间的关系, 由此关系可以明显看出, 测量距离与基线长度和光斑在传感器平面内坐标相关。

## 2.2 全向环结构光深度测量

$$\left\{ \begin{array}{l} \tan \theta = \frac{Y}{X} = \frac{y}{x} \\ \alpha = \tan^{-1} \theta \frac{(b^2 + c^2) \sin \gamma - 2bc}{(b^2 - c^2) \cos \gamma} \\ \gamma = \tan^{-1} \frac{f}{\sqrt{(x^2 + y^2)}} \end{array} \right. \quad (1)$$

式(1)中  $b$  为双曲面镜短轴,  $c$  为双曲面镜的长轴,  $f$  为双曲面镜的焦距。其中  $f$  为摄像机镜头的焦距, 通过求得  $\alpha$  可以求出三维空间中点  $P(X, Y, Z)$  在图像传感器中的映射点  $p(x, y)$ 。当使用的双曲面镜参数已知,  $\alpha$  只与相机平面内参  $\theta$  有关, 则只要标定出全景相机中心在相平面内的投影中心, 再通过激光点定位算法进行激光点定位, 则可以求出图2中的角度  $\alpha$ , 带入公式就可以求出测量深度:

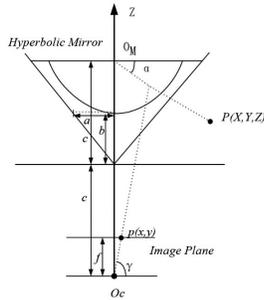


图2 深度测量几何模型俯视图

$$q = \frac{s(b^2 - c^2) \cos \gamma}{(b^2 + c^2) \sin \gamma - 2bc} = \frac{s(b^2 - c^2) \cos \gamma}{(b^2 + c^2) \frac{f}{\rho} - 2bc \sqrt{1 + \frac{f^2}{\rho^2}}} \quad (2)$$

式(2)中  $\rho = \sqrt{x^2 + y^2}$ ,  $S$  为基线距离。由前述可知,  $\tan \gamma = \frac{f}{\rho}$ , 因为传感器尺寸较小, 远小于焦距长度, 所以  $\angle \gamma$  接近 90 度且随着像素值的变化角度值变化很小, 所以可以认为  $\tan(\gamma)$  是一个远大于 1 的定值。将

公式(2)化简为下式(3):

$$q = \frac{s(b^2 - c^2) \cos \gamma}{(b^2 + c^2) \sin \gamma - 2bc} \approx \frac{A}{C} + \frac{AB}{(B - C\rho)C'} \quad (3)$$

$$A = s(b^2 - c^2), B = f(b^2 + c^2), C = 2bc \tan(\gamma)$$

式中  $A, B, C$  均为简化后的参数, 受误差影响较难测量参数, 后续采用模型拟合的方法确定参数。从模型图中可以看出, 结构光光斑在二维图像传感器上的投影呈径向分布, 光斑在相机平面的坐标和真实测量距离大致为  $\frac{1}{\rho}$  倒数关系。由公式(3)可以进一步推导出深度测量的敏感度如下(4)所示:

$$\frac{dq}{d\rho} = \frac{(b + c)^2 + 2bcq^2(b - c)^2}{fs(b^2 - c^2)} \quad (4)$$

从公式(4)可以看到敏感度  $\frac{dq}{d\rho}$  随着距离的增长,

误差呈几何增长。通过上式可以计算设定精度内的双曲面镜参数、镜头焦距和基线距离, 对系统设计有较大的指导意义。当基线距离增加, 其他参数均不变化的情况下, 相同测量位置的精度也随之增加。随着测量距离的增长, 其他参数均不变的情况下, 则每增加测量单位长度距离, 测量点附近的精度都会呈现几何增长。

## 2.3 点云三维坐标计算

根据图像坐标和空间坐标的对应关系, 可以计算出空间点的三维坐标信息。全向环感知系统<sup>[22]</sup>是一种以观察者为中心的单视点全景视觉感知系统, 因此在点云计算过程中采用笛卡尔坐标系<sup>[23]</sup>, 坐标系如图3所示。以单目的中心作为坐标系的原点, 以中心轴沿着双曲面镜方向作为  $Z$  轴正方向在空间中建立三维坐标系。其中,  $q$  为空间中点到摄像头的距离,  $\zeta$  为舵机旋转角度,  $\phi$  为全向环结构光中每一点相对于像素坐标系的角度。

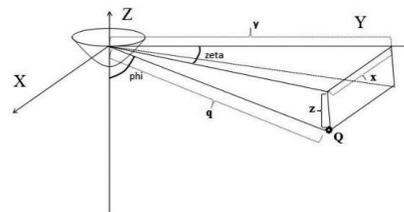


图3 点云三维测量原理图

根据像素坐标系到空间坐标系的转换,即可以得到空间坐标系中点的三维坐标公式如下式(5):

$$\begin{cases} x = q \times \sin(\phi) \times \cos(\zeta) \\ y = q \times \sin(\phi) \times \sin(\zeta) \\ z = q \times \cos(\phi) \end{cases} \quad (5)$$

### 3 点云配准方法

大场景空间大,且空间场景相对复杂。因此,本文首先联合多种点云滤波模型完成对初始点云数据的精简;其次提取点云关键点并进行多维度特征化描述;然后以特征向量为索引单位,在最邻近搜索的基础上,结合比率抑制和双向筛选完成对应点对搜索,以达到快速收敛;最后根据对应点构建点云变换模型,求出最优解,实现点云配准。

#### 3.1 点云预处理

大场景三维点云地图的点云数据量较大,点的数量一般在十万到百万级别,如果对所有的点进行点云配准操作势必产生过高的计算负荷。为了能够在保留原有云点特征的情况下有效减少点的数量,最终提高点云配准效率,本文采用联合多种滤波方法对点云进行精简操作。

##### 3.1.1 点云体素网格滤波

点云处理的第一步是对点云进行体素网格滤波<sup>[24]</sup>。每个体素网格都是一个三维空间立方体,边长为L。体素网格的重心 $P(\bar{x}, \bar{y}, \bar{z})$ 近似地表示该网格中所有N个空间点,边长L决定了体素网格滤波的程度。重心P的表达形式如下式(6)所示:

$$\begin{cases} \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \\ \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \\ \bar{z} = \frac{1}{N} \sum_{i=1}^N z_i \end{cases} \quad (6)$$

体素网格的重心P近似地表示了该网格中其他所有空间点,边长L决定了体素网格滤波的尺度。

##### 3.1.2 点云离群点剔除

点云分布较为密集的地方信息量就越大,相反点云分布稀疏地方信息量较少。因此离群点所表达的信息量较少,同时离群点容易影响当前点云呈现效果,而且还可能影响整个点云的配准过程。为了能够最大限度降低离群噪点产生的上述不良影响,本文采用统计滤波算法<sup>[25]</sup>对离群噪点进行

剔除。

统计滤波算法是一种对每个点的邻域都进行统计分析的算法。具体流程为:计算每个点到其最近的K个点的平均距离 $d_i$ ,其中 $i = 1, 2, 3 \dots N$ 。假设所有平均距离 $d_i$ 满足高斯分布,则均值 $\mu$ 代表了点云的平均密度,标准差 $\sigma$ 代表了点云的密度分布。将平均距离小于设定范围的点定义为离群点,从点云数据中去掉,保留绝大部分满足设定范围的点。

### 3.2 点云关键点提取

#### 3.2.1 ISS关键点提取

ISS(Intrinsic Shape Signatures)<sup>[26]</sup>关键点是一种通过点云局部空间几何信息加权确定的关键点。遍历点云中所有点,以每个点 $P_c$ 为中心,建立空间局部坐标系,以 $r$ 为半径,搜索 $r$ 邻域内的球形空间所有点,假设邻域中所有点的个数为 $n$ ,可以根据下式(7)确定每个点所对应的权值 $w_{ij}$ :

$$\begin{cases} w_{ij} = 1 / \|P_c - P_i\| \\ |P_c - P_i| < r \\ i \in P_c \{0, 1, 2 \dots n\} \end{cases} \quad (7)$$

根据式(7)计算得出的权值构建球形邻域内所有点的协方差矩阵 $cov(P_i)$ ,见式(8):

$$cov(P_i) = \frac{\sum w_{ij} (P_c - P_i)(P_c - P_i)^T}{\sum w_{ij}} \quad (8)$$

根据协方差矩阵计算特征值 $\lambda_i$ ,其中 $i \in \{0, 1, 2\}$ ,如下式(9)所示:

$$cov(P_i) \cdot v_i = \lambda_i \cdot v_i \quad (9)$$

对特征值 $\lambda_i$ 从小到大进行排序,假设得到的特征值序列为 $\lambda_0, \lambda_1, \lambda_2$ 。如果特征值满足 $\lambda_1/\lambda_2 \leq \varepsilon_1, \lambda_0/\lambda_1 \leq \varepsilon_2$ 约束关系则判定该中心点 $P_c$ 为ISS特征点。

#### 3.2.2 SIFT关键点提取

点云的3D SIFT(Scale-invariant Feature Transform)<sup>[27]</sup>特征点是从图像SIFT二维特征点到3D空间的拓展。SIFT是在不同的尺度空间中提取得到的特征点,SIFT特征点具有较强的抗光照、噪声干扰能力,同时也不会因点云仿射变换而发生变化。SIFT关键点提取可归纳为以下两个主要步骤。

##### 1. 尺度空间极值检测

要在尺度空间中进行极值检测首先需要构造高斯差分金字塔模型。三维空间中的尺度信息 $L(x, y, z, \sigma)$ 定义如下式(10):

$$\begin{cases} L(x,y,z,\sigma) = G(x,y,z,\sigma) * P(x,y,z) \\ G(x,y,z,\sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-\frac{(x^2+y^2+z^2)}{2\sigma^2}} \end{cases} \quad (10)$$

式(10)中,\*表示卷积, $\sigma$ 为尺度因子,一般情况下, $\sigma$ 取 $\sqrt[3]{2}$ 。 $G(x,y,z,\sigma)$ 表示具有尺度信息的高斯核函数, $P(x,y,z)$ 表示三维点坐标。

为了能够高效检测出三维空间中的特征点,采用高斯差分 DoG(Difference of Gaussian)算子<sup>[28]</sup>进行尺度空间极值点检测。

$$DoG(x,y,z,k^i\sigma) = L(x,y,z,k^{i+1}\sigma) - L(x,y,z,k^i\sigma) \quad (11)$$

式(11)中, $k$ 为乘法因子,利用 $k$ 可以得到不同尺度。

每个空间点都要在其邻域内进行筛选比较从而得到极值,包括极大值和极小值。二维图像中的中间点需要和邻域内的8个点进行比较,三维空间中的中间点则需要和周围26个点进行比较。如下图4所示:

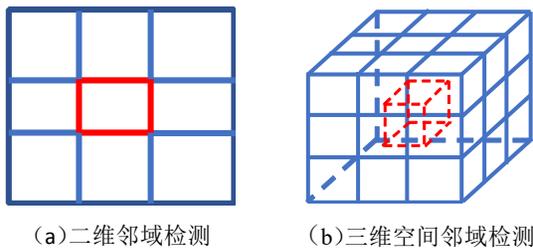


图4 多维空间邻域检测

## 2. 极值点方向确定

极值点方向角 $\theta$ 确定如下式(12)所示:

$$\theta(x,y,z) = \tan^{-1}\left(\frac{(y_i - y_c)}{(x_i - x_c)}\right) \quad (12)$$

式(12)中, $(x_c, y_c)$ 为三维空间中心点P坐标。

### 3.2.3 关键点对比与分析

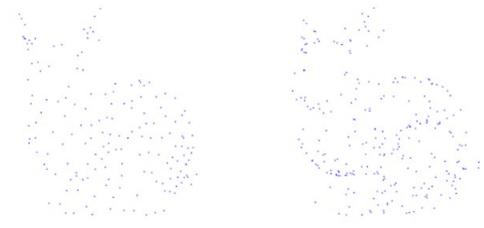
前两小节分别介绍了最常用的两种3D特征点检测算法,本小节对这两种3D特征点进行对比实验。



图5 斯坦福点云模型

图5为斯坦福大学利用三维扫描仪构建的兔子点云模型,该模型包含14806个点。本文对该模型进行

了点云的两种关键点提取实验,实验效果图如下图6所示。



(a)ISS特征点提取效果 (b)SIFT特征点提取效果

图6 两种特征点效果提取图

表1 两种关键点数据分析

特征点种类	特征点个数	提取时间	平均每点耗时
ISS	162	8.539 ms	0.05 ms
SIFT	296	16.77 ms	0.06 ms

从外观上看,ISS关键点在所有区域的分布更为均匀,而SIFT关键点在平坦区域特征点分布相对较少,在轮廓区域分布更为充分。在表达点云空间结构信息的过程中,轮廓信息尤为重要,因此SIFT关键点能够更为良好地表达出点云的空间结构信息。

从统计数据上比较,在同样阈值限定的情况下,SIFT关键点的数量更多,这也正是SIFT关键点耗时较长的原因。但是大规模点云配准问题的核心在于配准结果的准确性,而不是配准过程的实时性。因此本文可以在牺牲时间的基础上来提升结果的准确性。

综上所述,本文选取3D SIFT关键点作为后续过程中的特征点。

### 3.3 点云特征描述

点云特征描述符是对点云特征点的信息表达,一种好的描述符能够充分表达特征点自身以及特征点周围的空间信息,同时具备较强的鲁棒性。点云特征描述符最终以特征向量的形式进行表达。

一种合格的点云描述符需要满足以下四点准则:一是点云经过任意三维变换,特征向量不发生变化;二是点云经过不同的密度采样,特征向量保持不变;三是当点云存在部分噪点时,特征向量基本保持不变。四是特征向量本身能够全面化表述特征点的多维信息。本文依据上述准则对点云的特征描述符进行了分析与选取。

#### 3.3.1 点云固有特征信息

##### (1)颜色信息

点云中的每个点除了包含对应的XYZ空间位置信息之外,往往包含了彩色信息RGB三种分量。由于相机彩色摄像头精度较高,在同一时间和同一场景下,物体颜色信息不会发生较大变化,同一个物体在不同角度采集得到的点云中表现出相同且稳定的RGB分量。因此每个点的RGB分量能够为后续点云匹配提供相应的约束。

### (2) 曲率信息

曲率描述了点及其局部点的弯曲程度,形状不同的点云曲率信息相差较大,因此曲率是点云中比较重要的几何特征。曲率的计算依赖于该点及其周围点的坐标,但是并不会因为点云旋转平移而发生变化,因此曲率信息是点云的固有内在特征。可以采用直接计算的方式求取任意特征点 $P(X,Y,Z)$ 的曲率,具体的计算过程为:

1) 以任意一个特征点 $P(X,Y,Z)$ 为中心,采用kd-tree<sup>[29]</sup>的方式选取周围最邻近 $n$ 个点,计算 $n$ 个点的均值得到点 $\bar{P}(\bar{X},\bar{Y},\bar{Z})$ 。

2) 计算出协方差矩阵 $C$ 。

3) 根据特征向量与特征值计算公式(8)计算得出协方差矩阵 $C$ 的三个特征值 $\lambda_1, \lambda_2, \lambda_3$ 。式中, $\lambda_j$ 表示第 $j$ 个特征值, $v_j$ 表示第 $j$ 个特征值对应的特征向量。

4) 取 $\lambda_1, \lambda_2, \lambda_3$ 中的最小特征值 $\lambda_{min} = \min(\lambda_1, \min(\lambda_2, \lambda_3))$ 。

5) 根据式(12)计算特征点 $P(X,Y,Z)$ 对应的曲率 $\rho_p$ ,见式(13)。

$$\rho_p = \frac{\lambda_{min}}{\lambda_1 + \lambda_2 + \lambda_3} \quad (13)$$

### 3.3.2 PFH特征描述

颜色信息和曲率信息只能描述当前点的特征信息,当点云相对复杂时,上述特征信息不够全面,对于相似性较高区域或者对称区域容易出现较大的误差。为了更好地描述该点和周围点更深层次的空间位置关系,从而得到置信度更高的匹配点对,本文采用PFH特征描述符。

PFH<sup>[30]</sup>特征描述实质上是曲率特征更加进一步的表达。它描述了中心点和周围点之间的法线方向差异,同时建立统计直方图来表述中心点和周围点之间的几何特征。

(1) 为了计算两点之间的法线差异,需要选取两点连线与法向量夹角较小的一点构建 $uvw$ 坐标系。计算公式为式(14):

$$p_s = \begin{cases} p_i, n_i \cdot (p_s - p_i) \leq n_j \cdot (p_i - p_j) \\ p_j, n_i \cdot (p_j - p_i) > n_j \cdot (p_i - p_j) \end{cases} \quad (14)$$

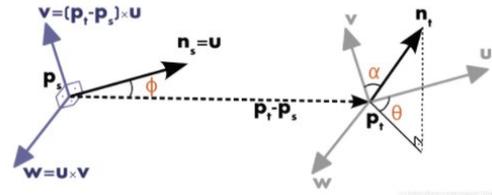


图7 局部空间坐标系

(2) 两点的信息包括各自的位置坐标 $(x,y,z)$ 和法线方向信息 $(n_x, n_y, n_z)$ 。

如图7所示,两点之间的相互关系可以用一组角度 $\alpha\phi\theta$ 以及两点之间距离 $d$ 共四个元素进行表示。如式(15)所示:

$$\begin{aligned} \alpha &= v \cdot n_i \\ \phi &= u \cdot \frac{(p_i - p_s)}{d} \\ \theta &= \arctan(w \cdot n_i, u \cdot n_i) \end{aligned} \quad (15)$$

(3) 构建中心点P的统计分布直方图。

首先计算查询点 $P_q$ 近邻内对应的所有四个元素,如图8所示,表示的是一个查询点 $P_q$ 的PFH计算的影响区域, $P_q$ 用红色标注并放在圆球的中间位置,半径为 $r$ , $P_q$ 的所有 $k$ 邻元素(即与点 $P_q$ 的距离小于半径 $r$ 的所有点)全部互相连接在一个网络中。最终的PFH描述子通过计算邻域内所有两点之间关系而得到直方图,计算复杂度为 $O(k)$ 。为了创建最终的直方图,将所有四元素组以统计的方式放入一个直方图中,这个过程首先把每个特征值范围划分为 $b$ 个子区间,并统计落在每个子区间的点数量,前三个元素均是角度,都和法向量有关,可以将三个元素标准化并放到同一个区间内。

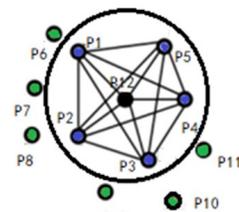


图8 空间结构示意图

本文将点云颜色、曲率以及PFH特征相结合,具体包含两个分量。第一个分量表达了点云中点的自身信息,包括颜色信息以及曲率信息,是一个4维度分

量。第二个分量表达了点云关键点的PFH特征,是一个125维度分量。

### 3.4 点云精确配准方法分析及改进

#### 3.4.1 ICP算法中所用到的匹配方法与不足

ICP<sup>[17]</sup>算法用到的匹配方法可以细分为两种,分别为阈值法和单向邻近匹配法。

##### (1) 阈值法

阈值法是通过设定的阈值进行筛选的一种方法。主要思想为:假设得到了两个待配准点云的特征描述子集合 $D_p$ 和 $D_q$ ,对于源点云的特征描述子集合 $D_p$ 中的任意一个特征描述子 $p^i$ ,从目标点云的特征描述子集合 $D_q$ 中通过阈值判断找到满足阈值的描述子 $q^j$ 。由于阈值的限定,满足阈值的描述子数量可能是一个也可能是一个集合。

阈值的设定和实验结果密切相关,对于不同密度、不同大小、原始空间位置相差较大的点云,事先设定的阈值需要动态变化才能满足实验要求,否则会出现错误的匹配。同时,由于点云特征点数量较为繁多,满足阈值要求的点可能较多,巨大的信息干扰使得ICP算法的迭代次数不断增多。因此阈值法通常只适用于点云空间结构较为简单、点云匹配步骤较少的情况。

##### (2) 单向邻近匹配

单向邻近匹配是ICP算法在搜寻点对过程中的主要方法。单向邻近匹配的过程为:假设得到了两个待配准点云的特征描述子集合 $D_p$ 和 $D_q$ 。遍历源点云的特征描述子集合 $D_p$ 中的每一个元素 $p^i$ ,在目标点云的特征描述子集合 $D_q$ 中找到与 $p^i$ 欧式距离最近的描述子 $q^j$ ,得到一对匹配的描述子 $(p^i, q^j)$ 。与阈值法不同的是, $p^i$ 具有唯一性。

单向邻近匹配算法在寻找对应点过程中是一种单向的搜索,只能满足从源点云到目标点云之间的搜索。假设利用单向邻近匹配寻找得到了一组对应点对 $(p^i, q^j)$ , $p^i$ 对应于 $q^j$ 点,但是并无法证明 $q^j$ 对应于 $p^i$ 点,很可能 $q^j$ 与其他点相对应。因此从原理上分析可知,单向的搜索过程容易出现更多的潜在错误点对。

#### 3.4.2 基于双向匹配与比率筛选的ICP配准

不论是利用阈值法进行特征点对的筛选,或者利用单向邻近搜索策略进行特征点对筛选,以上策略均会出现大量潜在的误匹配,因此会造成搜索过程的效率低下。造成效率低下的原因主要有三:

1) 阈值的设定较为单一,匹配结果无法令人

满意。

2) 单向搜索无法保证匹配点之间的双向对应。

3) 特征维度较高,只利用最邻近欧氏距离进行比较,相似的结果可能包含有其它大量错误的匹配,数据干扰较大。

##### (1) 双向筛选

ICP<sup>[17]</sup>方法中单向搜索的基本原理如图9所示。由于单向的搜索方式无法保证特征点对之间的相互对应,为了能够最大限度地提升对应点对正确性,本文将原有的单向搜索策略改为双向搜索策略,原理示意图如图10所示:

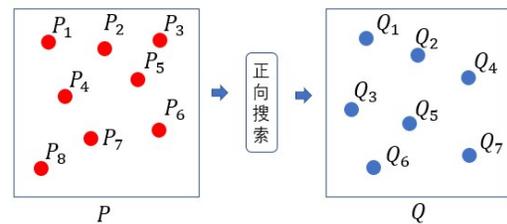


图9 单向搜索示意图

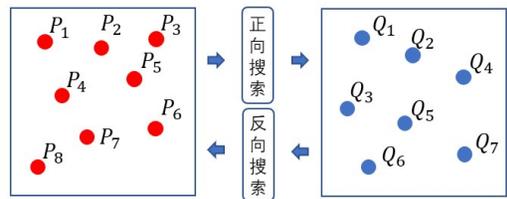


图10 双向搜索示意图

图10中,左侧区域红色标记点和右侧区域蓝色标记点分别为源点云 $P$ 和目标点云 $Q$ 中提取得到的3D关键点。在遍历源点云 $P$ 的所有3D关键点过程中,红色标记点 $P_3$ 在目标点云 $Q$ 中寻找得到了距离最近的点 $Q_5$ 。在ICP的单向搜索过程中,正向匹配点对 $(P_3, Q_5)$ 被认定为最佳匹配点对。但是在本文双向筛选过程中,在得到正向匹配点对 $(P_3, Q_5)$ 后,需要对点 $Q_5$ 反向从源点云 $P$ 中继续寻找最佳匹配点对。如果在反向搜索过程中, $P_3$ 是源点云 $P$ 中距离 $Q_5$ 点欧氏距离最近的点,则反向匹配点对 $(Q_5, P_3)$ 被判定为最佳反向点对。由于正向和反向的搜索结果相吻合,因此 $(Q_5, P_3)$ 为最佳匹配点对。反之,重新遍历源点云 $P$ 中的下一个关键点。

##### (2) 比率抑制

由于特征描述维度较高,当特征点存在疑似点时,如果直接利用阈值法,很容易得出错误的结果。如图11所示:

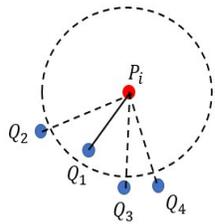


图11 存在干扰时的特征匹配

图11中的 $P_i$ 为源点云中的待匹配特征点, $Q_1, Q_2, Q_3, Q_4$ 皆为目标点云中距离 $P_i$ 点最近的四个特征点。根据阈值法可以计算得出, $P_i$ 的匹配点为 $Q_1$ ,而其他点 $Q_2, Q_3, Q_4$ 因到 $P_i$ 点距离大于 $Q_1$ 到 $P_i$ 的距离,皆被舍弃。但是阈值法计算结果不具有可靠性,由于特征点对应的特征描述具有高维度性质,欧氏距离最近的点不一定是最佳匹配点。而在干扰点中可能存在最佳的对应点。阈值法的处理过程很容易将潜在的目标点当作干扰点去除。

为了能够使得计算结果具有可靠性,本文提出比率抑制的方法。对于源点云中的待匹配点 $P_i$ ,在寻找对应点过程中,计算最邻近点 $Q_j$ 和次邻近点 $Q_s$ 分别到原始特征点之间的距离比,如式(16)所示:

$$d_i = \frac{d_{(P,Q_j)}}{d_{(P,Q_s)}} \quad (16)$$

设定阈值 $\varepsilon$ ,Lowe<sup>[10]</sup>在文章中已经给出了阈值 $\varepsilon$ 的经验阈值为0.8。比较距离比 $d_i$ 与阈值 $\varepsilon$ 之间的大小关系。如果 $d_i < \varepsilon$ 则认为次临近点的干扰性较弱,最临近点的可靠性较高。否则舍弃待匹配特征点 $P_i$ ,

继续对源点云其它特征点进行查找。比率抑制的示意图如下图12所示。

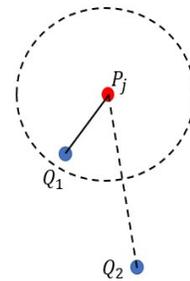


图12 基于比率抑制的提取效果图

图12中, $P_j$ 和 $Q_1$ 为经过比率抑制得到的匹配点对,而 $Q_2$ 因为距离远远小于 $Q_1$ ,认为 $Q_2$ 在特征匹配过程中的干扰性较弱,舍弃掉 $Q_2$ 点。

#### 4 三维点云建模方法

三维点云建模方法主要是以软件主动建模为主,通过CloudCompare<sup>[30]</sup>和Geomagic软件<sup>[31]</sup>,对大场景点云进行建模。如图13(a)所示为实验室的三维点云图,以实验室为对象,对实验室进行三维建模。

如图13所示,将实验室点云文件加载到CloudCompare内。首先通过软件中的裁剪功能,手动将原始点云中的离群点和噪点去除掉,去除后的点云图如图13(b)所示;然后将实验室分成若干部分,分类时可按照物品种类进行分类,比如墙面、地板、桌椅等,通过手动裁剪,将点云进行裁剪,被裁剪后的点云文件如图13(c)所示,并将裁剪后的点云保存成pts点云文

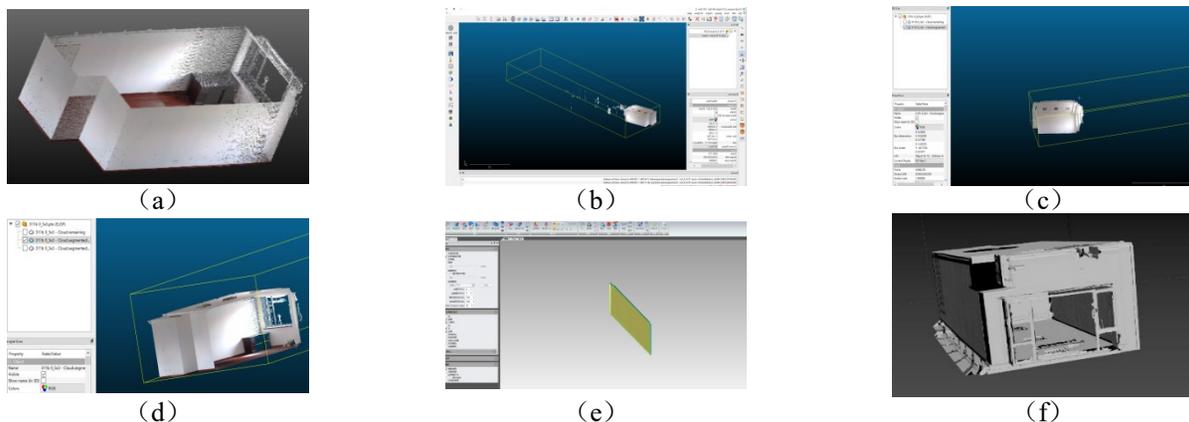


图13 三维点云建模方法。(a)通过三维扫描与点云配准生成实验室原始点云;(b)将原始点云加载到CloudCompare;(c)去除离群点和噪点;(d)对实验室进行分类并裁剪;(e)加载到Geomagic中进行封装;(f)通过3DMax对封装后的三角面片进行拼接,生成完整的实验室模型。

件;将裁剪后的点云加载到Geomagic软件中,对点云进行去噪、修补漏点和优化点云后,进行封装并转为obj三角面片格式,如图13(d)所示;最后将封装后的模型导入到3DMax中,并进行拼接,拼接后将其保存为max模型文件。如图13(e)所示为建模后的实验室三维模型。

通过CloudCompare、Geomagic和3DMax软件,可以快速的对原始点云进行建模。同时Geomagic软件中本身自带点云去噪、漏洞补全优化功能;3Dmax软件也可以对模型进行渲染,进一步对模型进行优化,从而达到更好的效果。

## 5 结论

本文面向大场景三维扫描与建模关键问题,首先采用全向环结构光深度感知方法获取场景的三维点云信息;然后采用基于双向匹配与比率筛选的ICP配准算法对扫描获得的点云进行配准,从而获得完整的大场景点云;最后采用CloudCompare和Geomagic三维软件对点云进行主动建模,获得高精度、高鲁棒性的大场景三维点云扫描与建模结果。

### 参考文献(References):

- [1] Shen S. Accurate multiple view 3d reconstruction using patch-based stereo for large-scale scenes[J]. IEEE Transactions on Image Processing, 2013, 22(5): 1901-1914.
- [2] Dai A, Nießner M, Zollhöfer M, et al. Bundlefusion: real-time globally consistent 3d reconstruction using on-the-fly surface reintegration[J]. ACM Transactions on Graphics (ToG), 2017, 36(4): 1.
- [3] Song L, Li X, Yang Y, et al. Structured-light based 3D reconstruction system for cultural relic packaging[J]. Sensors, 2018, 18(9): 2981.
- [4] Geng J. Structured-light 3D surface imaging: a tutorial[J]. Advances in Optics and Photonics, 2011, 3(2): 128-160.
- [5] Foix S, Alenya G, Torras C. Lock-in time-of-flight (ToF) cameras: a survey[J]. IEEE Sensors Journal, 2011, 11(9): 1917-1926.
- [6] Kanade T, Okutomi M. A stereo matching algorithm with an adaptive window: theory and experiment[J]. IEEE transactions on pattern analysis and machine intelligence, 1994, 16(9): 920-932.
- [7] Salvi J, Pages J, Batlle J. Pattern codification strategies in structured light systems[J]. Pattern recognition, 2004, 37(4): 827-849.
- [8] 李文浩,贾同,薛宇,等.一种基于线激光的目标深度测量方法[C]//中国金属学会.第十一届中国钢铁年会论文集-S16(分析检测).北京:冶金工业出版社,2017.
- [9] 贾同,吴成东,陈东岳,等.一种基于全向结构光的深度测量方法[J].自动化学报,2015,41(9):10.
- [10] Fanello S R, Rhemann C, Tankovich V, et al. Hyperdepth: learning depth from structured light without matching[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 5441-5450.
- [11] Fontanelli D, Ricciato L, Soatto S. A fast ransac-based registration algorithm for accurate localization in unknown environments using lidar measurements[C]. IEEE International Conference on Automation Science and Engineering, 2007: 597-602.
- [12] Ge X. Automatic markerless registration of point clouds with semantic-keypoint-based 4-points congruent sets[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130: 344-357.
- [13] Mellado N, Aiger D, Mitra N J. Super 4pcs fast global pointcloud registration via smart indexing[J]. Computer graphics forum, 2014, 33(5): 205-215.
- [14] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration[C]. IEEE International Conference on Robotics and Automation, 2009: 3212-3217.
- [15] Alshawa M. ICL: Iterative closest line A novel point cloud registration algorithm based on linear features[J]. Ekscentar, 2007 (10): 53-59.
- [16] Lepetit V, Moreno-Noguer F, Fua P. Epnp: an accurate o(n) solution to the pnp problem[J]. International Journal of Computer Vision, 2009, 81(2): 155-166.
- [17] Chetverikov D, Svirko D, Stepanov D, et al. The trimmed iterative closest point algorithm[C]. International Conference on Pattern Recognition, IEEE, 2002, 3: 545-548.
- [18] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction[C]. Proceedings of the fourth Eurographics Symposium on Geometry Processing, 2006: 61-70.
- [19] Lee D T, Schachter B J. Two algorithms for constructing a Delaunay triangulation[J]. International Journal of Computer & Information Sciences, 1980, 9(3): 219-242.
- [20] Wang Y, Wang L, Hao W, et al. A novel slicing-based regularization method for raw point clouds in visible IoT[J]. IEEE Access, 2018, 6: 18299-18309.
- [21] Jia T, Chen D Y, Bai L Q, et al. Depth perception based on omnidirectional ring structured light[J]. Optics & Laser Technology, 2018, 104: 123-132.
- [22] Ukida H, Yamato N, Tanimoto Y, et al. Omni-directional 3d

- measurement by hyperbolic mirror cameras and pattern projection [C]. IEEE Instrumentation and Measurement Technology Conference, 2008: 365-370.
- [23] Russell C T. Geophysical coordinate transformations [J]. Cosmic Electrodynamics, 1971, 2(2): 184-196.
- [24] Nießner M, Zollhöfer M, Izadi S, et al. Real-time 3D reconstruction at scale using voxel hashing [J]. ACM Transactions on Graphics (ToG), 2013, 32(6): 1-11.
- [25] Balta H, Velagic J, Bosschaerts W, et al. Fast statistical outlier removal based method for large 3D point clouds of outdoor environments [J]. IFAC-PapersOnLine, 2018, 51(22): 348-353.
- [26] Nießner M, Zollhöfer M, Izadi S, et al. Real-time 3D reconstruction at scale using voxel hashing [J]. ACM Transactions on Graphics (ToG), 2013, 32(6): 1-11.
- [27] Zhong Y. Intrinsic shape signatures: A shape descriptor for 3d object recognition [C]. IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009: 689-696.
- [28] Wang S, Li W, Wang Y, et al. An improved difference of Gaussian filter in face recognition [J]. J Multim, 2012, 7(6): 429-433.
- [29] Zhou K, Hou Q, Wang R, et al. Real-time kd-tree construction on graphics hardware [J]. ACM Transactions on Graphics (TOG), 2008, 27(5): 1-11.
- [30] Girardeau-Montaut D. CloudCompare [J]. France: EDF R&D Telecom ParisTech, 2016.
- [31] Liu L, Yang H. 3D reverse engineering design on seed tube based on Geomagic Design software [J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(11): 40-45.
- [32] Chi C, Wang Q, Hao T, et al. Feature-level collaboration: Joint unsupervised learning of optical flow, stereo depth and camera motion [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 2463-2473.

编辑:王谦,王雨田