

## 基于二进制深度图像描述符的点云配准方法

江政远, 蔡 驭, 杨俊成, 范 丹  
(云南省水利水电勘测设计研究院, 云南 昆明 650021)

**摘要:** 基于局部特征的三维点云配准是计算机视觉和机器人技术领域的核心问题, 现有的大多数三维局部特征描述符均是浮点型的。首先, 本文提出了一种二进制局部特征描述符, 即二进制深度图像描述符(BDIF), 用于描述三维局部特征; 其次, 提出了一种基于 BDIF 的配准算法, 用于大场景的点云配准。BDIF 根据局部曲面到投影面的距离将局部结构编码为位串。具体来说, BDIF 描述符是在关键点附近建立一个局部坐标系, 以实现旋转不变性, 然后在三个正交投影面上对空间信息进行编码。之后, 根据阈值法完成二值化, 并利用最大类间方差确定分割阈值。基于 BDIF 开发了一种高效的点云配准算法, 该算法采用自适应尺度的 Welsch 估计空间变化参数, 能有效处理大场景采集到的点云数据。最后分别在 Retrieval 和 WHU-TLS 数据集上进行了广泛的实验, 实验结果证明了本文所提的 BDIF 和基于 BDIF 的点云配准算法的有效性和总体优越性。

**关键词:** 局部特征描述符; 二进制; 点云配准; Welsch

**中图分类号:** TP242; TN958.98 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2024.10.010

## Point cloud registration method based on binary depth image descriptors

JIANG Zheng-yuan, CAI Yu, YANG Jun-cheng, FAN Dan

(Yunnan Institute of Water & Hydropower Engineering Investigation, Design and Research, Kunming 650021, China)

**Abstract:** 3D point cloud registration based on local features is a core problem in the field of computer vision and robotics, and most of the existing 3D local feature descriptors are of floating-point type. In this paper, a binary local feature descriptor, the Binary Depth Image Descriptor (BDIF), is proposed for describing 3D local features, and a registration algorithm based on the BDIF is also put forward for point cloud registration of large scenes. The BDIF encodes the local structure as a bit string based on the distance of the local surface to the projection plane. Specifically, the BDIF descriptor establishes a local reference frame near the keypoints to achieve rotational invariance, and then encodes the spatial information on three orthogonal projection surfaces. After that, binarization is completed based on the thresholding method and the segmentation threshold is determined using the maximum inter-class variance. An efficient point cloud registration algorithm is developed based on BDIF, which employs the adaptive scale Welsch to estimate the spatial variation parameters, and can effectively deal with the point cloud data collected from large scenes. Finally, extensive experiments are conducted on Retrieval and WHU-TLS datasets, respectively, and the experimental results demonstrate the effectiveness and overall superiority of the BDIF and BDIF-based point cloud registration algorithm proposed in this paper.

**Keywords:** local reference frame descriptors; binary; point cloud registration; Welsch

**基金项目:** 云南省旱限(警)水位确定试点与重点地区抗旱应急预案研究项目(No. 2021YFC3000205-06)资助。

**作者简介:** 江政远, 男, 硕士, 主要研究方向为遥感技术应用、水利工程测绘。E-mail: 897196238@qq.com

**通讯作者:** 杨俊成, 男, 主要研究方向为水利工程测量、激光雷达点云研究。E-mail: 744943993@qq.com

**收稿日期:** 2023-12-05

## 1 引言

点云配准是三维计算机视觉和机器人技术领域的关键问题,其相关应用较多,如三维对象识别、三维建模和智能交通系统等<sup>[1]</sup>。然而,由于视野和视点遮挡有限,激光扫描仪捕获的点云只能覆盖场景的一部分。为了解决这个问题,需要从不同的角度进行多次三维扫描,以覆盖整个场景,形成每个相邻的点云对都有重叠区域的扫描序列。点云配准是一种将这些单独的扫描合并成一个全景图的技术,以一种寻求最佳的刚性转换的方式将单个坐标系,统一到一个共同的参考系统中<sup>[2]</sup>。由于人工配准对大尺度的点云数据而言是相对耗时和不可行的,因此相关领域的研究重点主要是点云的自动配准。然而,尽管取得了重大进展,但仍有一些问题需解决。如对噪声敏感、效率较低和高昂的存储成本。

点云配准主要包括两个步骤:粗配准和精配准<sup>[3]</sup>。精配准通常采用最近点迭代法(ICP)<sup>[4]</sup>算法及其变体<sup>[5-6]</sup>。虽然ICP已经得到广泛的应用,但ICP的局限性在于可能收敛到一个局部最小值,并高度依赖于模型参数的初始猜测。如果初始参数不够好,ICP可能会失效。因此,针对原始点云需要进行粗配准以完成点云的初始对齐。目前基于局部特征描述符的点云粗配准在点云配准中占主导地位,因此描述符应对遮挡、噪声和分辨率等具有鲁棒性。基于局部特征描述符的点云配准关键在于局部特征描述符的建立。在过去的二十多年中,研究人员已经提出了多种描述符,如自旋转图像(SI)<sup>[7]</sup>、快速特征直方图(FPFH)<sup>[8]</sup>、旋转统计特征(RoPS)<sup>[9]</sup>、正交投影的点云局部特征描述(TOLDI)<sup>[10]</sup>、旋转轮廓特征(RCS)<sup>[11]</sup>等。然而,应用在实际采集到数据时,这些方法都存在一定的缺陷,比如SI和FPFH的描述性有限,FPFH对噪声异常敏感,RoPS的计算相对耗时,TOLDI的维度较高。这些描述符在高质量的点云数据上具有很好的表现,但点云数据质量较差时难以表现出令人满意的效果。

此外,上述所有描述符均是浮点型的,因此这些描述符需要较高的内存需求和特征匹配效率较差。然而,二进制描述符可以有效的处理这些问题,它是以“1”和“0”的形式来表达局部特征,通过XOR操作可以实现特征匹配的快速操作。这个问题在机器人技术和移动电话中尤为重要,因为许多应用程序

都需要在线和实时处理。对于二维图像存在许多二进制描述符,例如,BRIEF<sup>[12]</sup>和ORB<sup>[13]</sup>,而二进制的三维局部特征描述符则相对较少。第一次尝试获得三维局部二进制表示的是B-SHOT<sup>[14]</sup>,它是SHOT<sup>[15]</sup>描述符的扩展。与SHOT相比,在特征描述符匹配方面要快6倍。不幸的是,由于B-SHOT是SHOT的直接量化,它丢失了一定的鉴别信息,同时继承了SHOT的缺点;即,对网格分辨率的变化很敏感。Yang等<sup>[11]</sup>提出二进制的旋转轮廓特征(B-RCS),但该方法直接通过比较的方式进行二值化,导致信息损失较多。Quan等<sup>[1]</sup>使用简单的体素化方式提出二进制描述符(LoVS),虽然计算简单,但描述符的描述性较差。

基于上述分析,本文提出一种新的三维局部特征描述符,称为二进制深度图像描述符(BDIF),并基于BDIF提出了一种点云配准方法。BDIF的关键思想是它在三个正交视角上统计局部曲面的深度,然后将三个投影面的特征进行均值统计进而减少特征维度,并采用双边滤波对“空洞”进行填充,最后采用阈值法进行二值化,最大类间方差被用来确定分割阈值。基于BDIF的配准方法沿用基于局部特征的方法的标准模块,即关键点检测、特征提取、特征匹配和变换估计。它与标准模块的不同之处在于,特征提取步骤使用BDIF进行,变换使用自适应尺度的Welsch进行鲁棒估计。

## 2 二进制深度图像的构建

点云是离散的数据结构,通常只包含空间位置属性,缺乏与二维图像类似能直接比较的信息(比如,亮度)。因此难以通过恰当的选择方式对点云直接生成二进制描述符,导致针对点云二进制的描述相对较少。

### 2.1 局部坐标系的建立

局部坐标系(LRF)的建立是BDIF的关键步骤,LRF的鲁棒性和稳定性会对描述符的描述性产生显著的影响。LRF通常基于协方差分析或基于关键点的法向量。描述符BRoPH通过对局部曲面内的点进行加权协方差分析构建LRF,用关键点代替质心,并根据距离为不同点赋予不同的权重。估计协方差矩阵最小特征值对应的特征向量 $n$ 作为 $z$ 轴, $x$ 轴和 $y$ 轴是将局部曲面内的点通过加权投影进行估计。

和之前的工作类似,我们同样采用协方差矩阵分析的方式来构建 LRF。然而,利用局部曲面内所有点进行协方差分析,分辨率不同的数据容易导致 LRF 的稳定性较差。这是由于不同采用模式下局部曲面的内的点云间距不同,进而导致相同关键点所处的空间属性不同。因此,构建的 LRF 具体细节如下:

以关键点  $p$  为球心,支撑半径  $r$  内的点作为  $p$  局部曲面内的点。为了避免点云分辨率对 LRF 稳定性的影响,仅采用局部曲面内的部分点作为构建 LRF 的必要信息。具体而言,通过一个缩放因子  $\alpha$  来空间支撑半径的大小,则构建 LRF 的支撑半径  $r_M = \alpha r (\alpha < 1)$ 。  $\alpha$  的值根据<sup>[16]</sup>的建议,本文设置为  $\alpha = \frac{1}{3}$ 。减小支撑半径不仅可以降低点云分辨率对 LRF 的影响,还可以提高 LRF 的计算效率。因此关键点  $p$  的协方差矩阵为:

$$M = \frac{1}{N-1} \sum_{d_{pp_i} < r_M} w_i (p_i - p) (p_i - p)^T. \quad (1)$$

式中,  $N$  是支撑半径  $r_M$  内的点;  $d_{pp_i}$  是局部曲面内点  $p_i$  到  $p$  的欧式距离。  $w_i$  是根据  $p_i$  到  $p$  的欧式距离为协方差矩阵赋予的一个距离权重,其计算式为:

$$w_i = \frac{r_M - \|p_i - p\|}{r_M} \quad (2)$$

根据式(1)构建的协方差矩阵  $M$ , LRF 的  $z$  轴定义为式(3),同时为了消除向量的二义性,将向量与局部曲面内的点所构建的向量进行判断。

$$n_z = \begin{cases} -n_z, & \text{if } n_z \cdot \sum_{i=1}^N \vec{p_i p} < 0 \\ n_z, & \text{otherwise} \end{cases} \quad (3)$$

式中,  $n_z$  是  $M$  最小特征值对应的特征向量;  $\vec{p_i p}$  是关键点  $p$  与其所在平面内点  $p_i$  构建的向量。

之后,将局部曲面内点投影到  $xy$  轴所在平面得到投影点  $p_j$ , 则  $x$  轴根据  $p$  与所构建的向量进行计算,如下式所示:

$$n_x = \sum_{i=1}^N \vec{p_i p} \quad (4)$$

最后  $y$  轴根据上述计算的  $n_z$  和  $n_x$  定义为:

$$n_y = n_x \times n_z \quad (5)$$

式中,“ $\times$ ”表示向量的叉乘。

## 2.2 深度图像描述符

一旦为局部曲面构造了 LRF,我们就需要对局部

曲面中包含的空间和几何信息进行编码。首先对局部曲面进行相对于 LRF 的变换,以实现刚性变换的不变性。旋转后的点为  $Q = \{q_0, q_1, q_2, \dots, q_N\}$ 。然后,需要设计一种有力的描述符以实现关键点特有空间信息表达。

为了有效编码局部形状的几何信息,本文采用局部深度特征来描述局部形状,原因有两个。首先,局部深度特征,有时也称为有符号投影距离,保留了形状的大部分基本信息。相比之下,其他方法的局限性是从局部形状的 3D 到 2D/1D 表示的信息丢失。其次,局部深度特征的计算效率相当高。如图 1 所示,  $q_j$  为旋转后局部曲面内的点,  $q'_j$  为  $q_j$  在  $xy$  轴所在平面的投影,  $d_j$  为  $q_j$  与  $q'_j$  之间的欧式距离。然而,由于杂波和遮挡,单个视图并不总是足以充分利用局部表面中所包含的几何信息。因此,本文在 LRF 中定义了几个虚拟视图平面,如  $xy$ 、 $xz$  和  $yz$  轴所在平面,以便对局部曲面进行一个全面的编码。

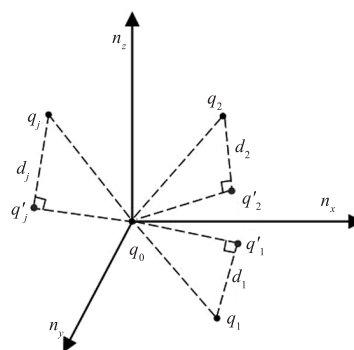


图 1 局部深度

Fig. 1 Localized depth

确定三个投影面后,分别根据下式计算局部曲面内的点分别在这三个投影面的局部深度:

$$\begin{cases} d_j^{xy} = \frac{r - q'_j \cdot n_z}{r} \\ d_j^{xz} = \frac{r - q'_j \cdot n_y}{r} \\ d_j^{yz} = \frac{r - q'_j \cdot n_x}{r} \end{cases} \quad (6)$$

式中,  $q'_j \cdot n_z$ 、 $q'_j \cdot n_y$ 、 $q'_j \cdot n_x$  分别代表  $q_j$  在  $xy$ 、 $xz$  和  $yz$  轴所在平面的深度。

三维局部特征描述符需要一个固定维度,才能使描述符之间具有可比性。本文使用一个二维阵列在每个投影面形成一个  $w \times w$  的图像  $I$ , 图像的像素值定义为相应投影点根据式(6)计算的最大值。三个投影面对应着三个特征图像  $I_{xy}$ 、 $I_{xz}$  和  $I_{yz}$ 。为了

有效地整合这些子特征且减少特征维度,三个子特征对应位置的像素均值构成浮点型的描述符:

$$I = \frac{1}{3}(I_{xy} + I_{xz} + I_{yz}) \quad (7)$$

由于点云的分布具有随机性,在投影的过程中可能造成部分像素内没有投影点,进而导致“空洞”像素。这种由点云分布不均匀产生的“空洞”会对局部特征描述的辨别性产生较大的影响。为了解决这个问题,本文采用双边滤波填充这些“空洞”像素<sup>[17]</sup>。双边滤波是一种经典的非线性空间滤波方法,它同时考虑了空间域和值域对于滤波产生的影响,因而能达到保持边缘,降噪平滑的效果,是一种良好的边缘保持滤波器。双边滤波的计算公式为:

$$I_f(u, v) = \frac{1}{W_p(u, v)} \sum_{(u, v) \in \Omega} I(u + i, v + j) F_s(i, j) F_r(I(x, y), I(u + i, v + j)) \quad (8)$$

式中,  $I_f(u, v)$  是滤波后的像素值;  $I(u, v)$  是原始图像中的像素值;  $W_p(u, v)$  是归一化的权重,用于归一化滤波结果;  $\Omega$  表示滤波器的邻域窗口,通常是一个固定大小的窗口;  $F_s(i, j)$  是空间域核函数,用于计算像素之间的空间距离权重;  $F_r(I(x, y), I(u + i, v + j))$  是像素值域核函数,用于计算像素值之间的相似性权重。

### 2.3 深度图像描述符二值化

经过2.2节计算后得到了浮点型的描述符。为了降低描述符的存储需求,本文采用阈值法将浮点型的描述符转换为二进制类型。阈值法是一种简单且高效的二值化方法。由于  $I_f$  由一系列的浮点型数值构成,依次转换每一个数值就能将其二值化。 $f(i)$  表示  $I_f$  的第  $i$  个特征,根据下式可以计算到一个位串  $f_b(i)$  :

$$f_b(i) = \begin{cases} 1, & \text{if } f(i) < \tau \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

式中,  $\tau$  是一个分类阈值,其可以通过对数据集进行训练或多次实验获得。图像阈值转换法中的最大类间方差(OTSU)已经被验证能为图像分割确定合适的阈值<sup>[18-19]</sup>。本文将其用于确定  $\tau$  的取值从而为数据的转换提供合理的依据。该方法有两个优点:首先,避免多次调参从而提高算法的运行效率;其次,特征描述符需要尽可能表达局部曲面包含的空间信息,所以二值化的过程中应使类间方差最大。

基 OTSU 方法,能最大限度使类间方差最大的分割,意味着错分概率最小。OTSU 计算阈值的具体细节如下:

(1) 假设此时选定的阈值为  $t$ , 则  $I_f$  小于  $t$  的像素点个数为  $n_0$ , 大于  $t$  的像素点个数为  $n_1$ 。那么小于阈值  $t$  的像素在整幅图的占比为:  $r_0 = \frac{n_0}{w \times w}$ , 大于阈值  $t$  的像素在整幅图的占比为:  $r_1 = \frac{n_1}{w \times w}$ 。

(2) 阈值为  $t$  的平均灰度值为:

$$I_a = n_0 u_0 + n_1 u_1 \quad (10)$$

式中, 其中  $u_0$  和  $u_1$  分别是指两个类别的平均灰度值(灰度值总和除以像素点个数)。

(3) 此时可求得类间方差为:

$$\sigma^2 = r_0 r_1 (u_0 - u_1)^2 \quad (11)$$

当取阈值  $t$  使得  $\sigma^2$  值最大时, 为最佳阈值  $\tau$ 。

经过式(9)计算后, 就得出多个“0”和“1”组成的字符串, 也就是最终的二进制深度图像描述符(BDIF)。

### 3 基于 BDIF 的点云配准

在本节中, 我们将 BDIF 应用于点云配准。一个有效的三维点云配准算法应该是自动和高效的。同时, 为了解决实际应用中点云配准存在耗时高、对重叠度、噪声、点分辨率鲁棒性弱托问题。我们提出一种基于 BDFI 的点云配准算法来解决这些问题。该算法由关键点提取、BDFI 估计、特征匹配和变换估计等模块组成。

#### 3.1 关键点提取

固定尺度和自适应尺度的关键点检测方法均有相关研究人员进行了广泛的研究。固定尺度的关键点检测方法可能在局部曲面变化较小的地方检测的关键点较少, 导致关键点无法反应物体的真实特征<sup>[20]</sup>。因此, 本文采用 Main 等<sup>[21]</sup>提出的自适应关键点检测算法。Main 等<sup>[21]</sup>利用多个尺度检测关键点, 并通过最大化局部表面的前两个主轴之间的比例来确定固有尺度。因此, 针对目标点云和源点云分别检测到关键点  $P_s$  和  $P_t$ 。

#### 3.2 BDFI 估计和特征匹配

当完成关键点检测后, 根据第2节内容分别计算  $P_s$  和  $P_t$  对应的 BDFI 特征  $F_s = \{f_1, f_2, \dots, f_{N_s}\}$  和  $F_t = \{f_1, f_2, \dots, f_{N_t}\}$ 。

计算了  $P_s$  和  $P_t$  对应的 BDFI 特征  $F_s$  和  $F_t$ , 我们需要建立  $P_s$  和  $P_t$  之间的点与点的对应关系。对于特征  $f_i^s$ , 根据式(12)在  $F_t$  中搜索与它最相似的特性。由于 BDFI 是二进制的形式, 因此我们采用汉明距离衡量两个描述符之间的相似性<sup>[22]</sup>, 汉明距离在现代 CPU 上计算的非常快。

$$f_i^t = \min_{j=1,2,\dots,N_t} (h(f_i^s - f_j^t)) \quad (12)$$

式中,  $h(f_i^s - f_j^t)$  表示  $f_i^s$  和  $f_j^t$  之间的汉明距离。

为  $F_s$  中所有的特征都找到最相似的特征后, 得到一个特征对应集  $\{(f_1^s, f_1^t), (f_2^s, f_2^t) \dots (f_{N_s}^s, f_{N_s}^t)\}$ 。 $(f_i^s, f_i^t)$  对应的点  $(p_i^s, p_i^t)$  则称为一个点对  $c_i$ 。根据这个规则, 最终在  $P_s$  和  $P_t$  之间生成一个初始点对应集  $C = \{c_1, c_2, \dots, c_{N_s}\}$ 。

### 3.3 变换估计

当为  $P_s$  和  $P_t$  生成对应集  $C$ , 需要从对应集  $C$  估计出正确的变换参数。然而, 由于  $P_s$  和  $P_t$  通常只存在部分重叠,  $P_s$  中的每个点无法都在  $P_t$  找到对应关系, 所以  $C$  中可能包含大量的错误匹配。为了剔除对应集中的错误匹配, 首先根据公式(12)比较特征匹配的相似性, 按降序的方式对  $C$  中的所有匹配进行排序, 然后从排序后的对应集中选取前  $k$  个对应集  $C' = \{c_1, c_2, \dots, c_k\}$ , 即相似性高的特征对更可能成为正确的匹配。由于点云的对称结构特征或受到噪声的影响,  $C'$  中仍然可能存在不正确的匹配。

为了保证配准结果的准确性, 通常是最小化  $C'$  中对应点之间的距离, 则目标函数为:

$$\min_R \sum_{(p_i^s, p_i^t) \in C'} \rho(\|p_i^t - (p_i^s R + t)\|) \quad (13)$$

式中,  $\rho(\cdot)$  是代价函数。

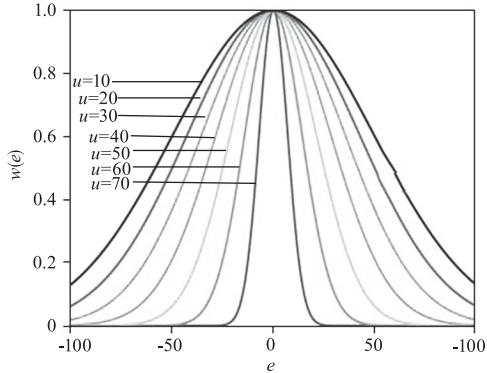


图2 从粗到细的 IRLS

Fig. 2 IRLS from coarse to fine

如果  $\rho(\cdot)$  是一个 M-估计问题, 比如 Welsch 估计, 则式(13)可以很容易地用迭代加权最小二乘

(IRLS)方法求解。然而, IRLS 可以得到局部最小的解决方案。在我们的配准方法中, 使用一个按比例 Welsch 函数作为代价函数, 其权重函数为:

$$w_i = \exp\left(-\left(\frac{e_i}{u}\right)\right) \quad (14)$$

式中,  $u$  是一个尺度参数;  $e_i$  是匹配对  $(p_i^s, p_i^t)$  的残差。然后, 引入了一个粗到细的 IRLS 进行优化, 它是由基于分级非凸性的方法。与传统的 IRLS 不同, 我们通过参数  $u$  和迭代来改变权重函数, 这在很大程度上避免了求解器陷入局部极小值。从图2中可以看出, 较大的  $u$  使函数曲线更宽, 更多的对应关系可以参与优化。相比之下, 函数曲线随着  $u$  的减小而变窄, 估计也变得更加精确。在自适应尺度 Welsch 估计中, 我们首先将一个大值  $u_0$  设为  $u$ ; 然后, 在每次迭代中,  $u_0$  除以一个步长  $\varepsilon$ , 直到 IRLS 优化收敛。

### 4 BDIF 分析

为检验本文算法的有效性, 本文分别设计了两组实验对描述符的描述性和配准精度进行了分析。数据测试在处理器为 12 th Gen Intel(R) Core(TM) i7-12700F 2.10 GHz、运行内存为 32.0 GB 的 Windows 10 64 位操作系统, Visual Studio 2022 Visual C++ x64 控制台应用程序上进行编程。

为了验证 BDIF 对于噪声和点云不同分辨率的鲁棒性, 我们在 Retrieval 数据集<sup>[21]</sup>上进行了一系列的实验。Retrieval 数据库所有的模型均来自 Stanford Repository。Retrieval 数据库包含了 18 个场景和 6 个模型, 其中场景是模型经过随机刚体变换和添加多个尺度的高斯噪声生成的。为了验证描述符对点云分辨率的鲁棒性, 将场景降采样为原来点云分辨了的 1/2 和 1/4。

实验结果的评价指标采用精度召回曲线 (PRC)<sup>[11]</sup>。根据召回率 (recall) 和 1-精度 (precision) 可绘制一条靠近坐标左上角的一条曲线。PRC 的面积 AUC 是评价描述符稳健性另一个常用的指标。在本文中, 我们利用 AUC 来衡量描述符的准确性。

#### 4.1 BDIF 的关键参数

BDIF 的关键参数主要包括: 支撑半径  $r$  和网格宽度  $w$ 。支撑半径  $r$  决定了局部曲面的大小, 较大的支撑半径虽然能包含较多的点, 但对算法



的运行效率产生严重的影响。 $w$  决定了描述的维度,虽然较小的网格数能降低描述的维度,但同时也导致描述符的描述性较差。为了分析出适当的参数,我们分别测试了不同参数时描述符的描述性。实验结果图3所示(图中 pr 表示点云的分辨率)。

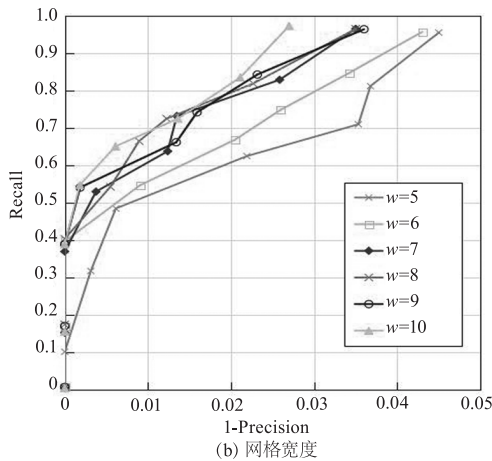
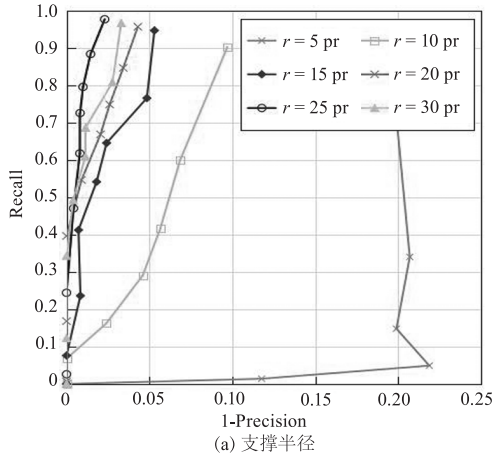


图3 BDIF 在不同参数下的表现

Fig. 3 Performance of BDIF with different parameters

图3(a)表明,当支撑半径  $r$  过小时,如  $r = 5$  pr 时,描述符基本没有鉴别力。随着  $r$  逐渐增大时,描述符的描述性也在随之增大。但  $r$  过大时,描述符的描述性也并没有随之改善,反而出现了一定的下降,如  $r = 30$  pr。虽然  $r = 25$  pr,描述符取得最佳的表现性能,但为了平衡描述符的计算效率,实际应用我们将支撑半径设置为  $20$  pr。

根据图3(b)可知,随着网格宽度的逐渐增大,描述符的描述性有了明显的提高。这是因为较大的网格能够表现出更多的局部特征。网格宽度进一步增大时,描述符的描述性却出现了下降,如  $w = 9$ ,这是因为过宽的网格并不能对点云分辨率和

噪声具有较好的鲁棒性。但  $w = 10$  时,描述符的性能又有了一定的提高,因此,实验中我们将网格宽度设置为  $10$ 。

4.2 BDIF 的描述性

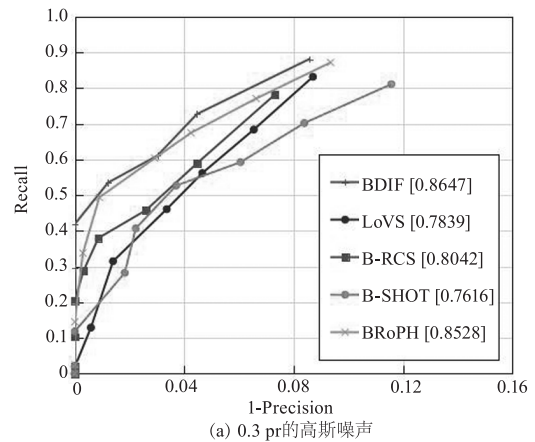
与先进描述符的对比分析也是我们实验内容的重点之一。本文将 BDFI 与现有的四种先进二进制描述符进行了对比分析,比如 B-SHOT<sup>[14]</sup>、B-RCS<sup>[11]</sup>、BRoPH<sup>[23]</sup> 和 LoVS<sup>[1]</sup>。所选描述符的关键参数表1所示。

表1 不同描述符的参数

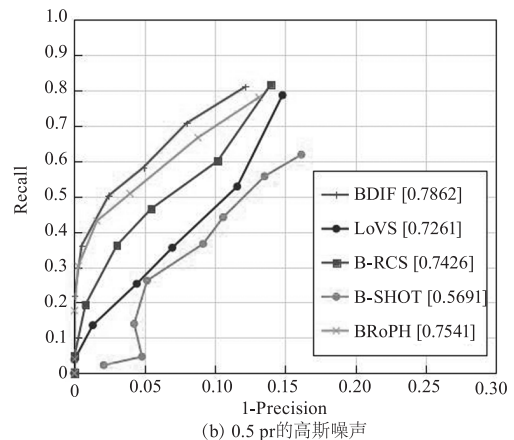
Tab. 1 Parameters of different descriptors

描述子	支持半径/pr	维度	长度	字节
BDIFI	20	10 × 10	100	100 × 1
LoVS	20	9 × 9 × 9	729	729 × 1
B-RCS	20	6 × 12 × 4	288	288 × 1
B-SHOT	20	8 × 2 × 2 × 11	352	352 × 1
BRoPH	20	2 × 9 × 24	432	432 × 1

根据表1中设置的参数,在 Retrieval 数据集上分别检验了所选描述符对噪声和点云分辨率的鲁棒性。实验结果如图4所示。



(a) 0.3 pr 的高斯噪声



(b) 0.5 pr 的高斯噪声

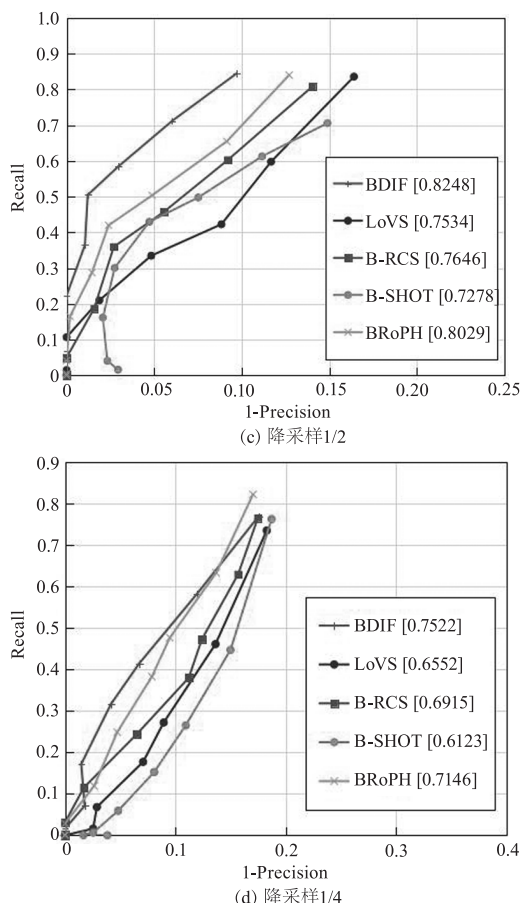


图4 不同描述符的描述性

Fig. 4 Descriptiveness of different descriptors

从表1和图4可以看出,与其他几种二进制描述符相比,BDIF的描述性和维度均具有一定的优势。具体来说,基于BDIF描述符具有一下特点:

(1) 构建LRF时,仅采用了支撑半径内的部分点,并且根据局部曲面内的点至关键点的距离为不同的点赋予了相应的权重,这样就能使远距离的点所占的权重较小,提高了LRF的稳定性。进而为BDIF良好的描述性提供了强有力的基础。

(2) 构建局部特征描述符的过程中,在三个正交视角上多方面解析局部曲面包含的空间信息;为了填补投影时产生的“空洞”,采用了同时考虑空域信息和值域信息的双边滤波对无投影点的网格进行赋值;为了降低描述符的维度采用均值的方式将三个正交视角获取的到特征进行合并。

(3) 二值化的过程中,本文采用阈值的方法将均值图像转为二进制。OTSU方法被用来确定最佳阈值,这样能在很大程度上确保“0”和“1”所对应的像素值方差最大。

## 5 配准精度分析

本节对所提出的基于BDIF的配准算法进行了

评价。我们在WHU-TLS数据集<sup>[24]</sup>中的Park、Mountain和Campus数据上测试了我们的算法。与最先进的基于局部特征描述配准算法进行了比较,并对实验结果进行了定性和定量的分析。

### 5.1 定性评价

图5为三个场景中点云的配准结果。从图中可以直观感知到初始输入点云的质量非常差:存在大量对称结构、点的不规则性、自遮挡和确实区域明显等,这些问题为点云配准带来巨大的挑战。然而,在两个待配准点云之间还是构建了多个匹配对。主要原因在于:一方面,所提出BDIF具有强大的描述性;另一方面,BDIF对点云分辨率具有鲁棒性。另外,图6中给出了15个相邻测站的配准结果,以显示我们配准算法的有效性。

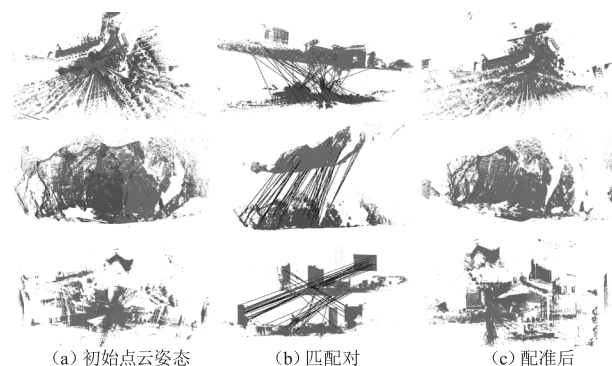


图5 所提算法在三种数据上的配准

Fig. 5 Registration of the proposed algorithm on three types of data

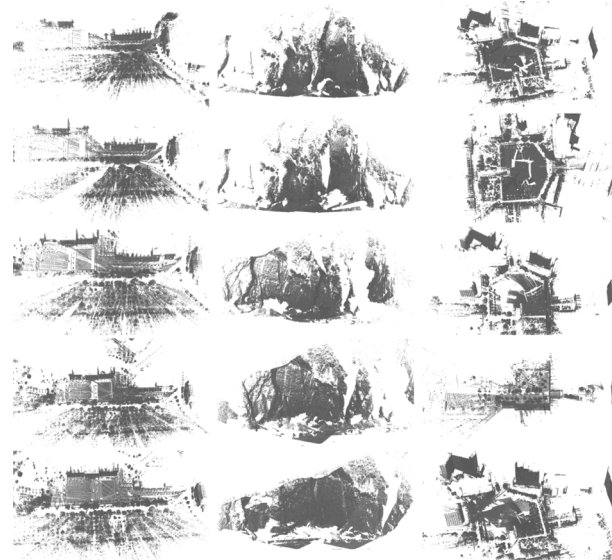


图6 基于BDIF的15个配准结果

Fig. 6 BDIF-based 15 registration result

图5和6的配准结果表明,本文所提的配准算法能成功将点云的坐标系进行统一。为了清晰地展示

相邻测站点云,本文使用不同的颜色对点云进行渲染,所以质量较高的点云配准应使颜色均匀分布。因此,在图中也可以看到所提方法均实现良好的配准。

整个配准过程中不需要人为干预或对初始姿态进行处理,因此所提算法是全自动的。只需将两个待配准点云作为数据输入即可。

## 5.2 定量评价

除了可视化的定性评价外,实验中还根据配准精度和时间开销对配准算法进行了定量评价。

### 5.2.1 配准精度

本文展示了配准后以数值描述的配准精度,并

与其他配准算法进行了比较。为了保证实验的公平性,实验中使用了与比较方法相同的配准流程。特别地,配准过程中仅替换了特征描述符,进而采用 RANSAC 解算配准参数。

根据表 2 可知,本文提出的基于 BDIF 的配准算法在所测数据集上均正确实现了点云高精度配准。基于 RANSAC 方法的点云配准也实现了点云配准,只是其整体精度相对本文所提方法较低。为了更加直观地分析 RANSAC 和 Welsch 两种方法解算配准参数的精度,表 3 统计了不同数据使用相同特征描述符时的配准精度均值。

表 2 不同描述符和配准方法的配准精度

Tab. 2 Registration accuracy for different descriptors and registration methods

配准方法	数据集	BDIF	LoVS	B-RCS	B-SHOT	BRoPH
RANSAC	Park	0.1526	0.1636	0.1765	0.2101	0.1824
	Mountain	0.2018	0.2206	0.2167	0.2347	0.2156
	Campus	0.1845	0.2087	0.2103	0.2237	0.1896
Welsch	Park	0.1354	0.1468	0.1619	0.1567	0.1391
	Mountain	0.1189	0.1411	0.1526	0.1804	0.1311
	Campus	0.1366	0.1737	0.1355	0.1799	0.1395

表 3 RANSAC 和 Welsch 解算的配准精度

Tab. 3 Registration accuracy of RANSAC and Welsch solvers

	BDIF	LoVS	B-RCS	B-SHOT	BRoPH
RANSAC	0.1796	0.1976	0.2012	0.2228	0.1959
Welsch	0.1303	0.1539	0.1500	0.1723	0.1365
提高/%	27.45	22.16	25.48	22.67	30.32

RANSAC 和 Welsch 解算空间变换参数时采用的关键点是完全相同的。根据表 3 可知,采用自适应尺度的 Welsch 估计配准参数更加准确。自适应尺度的 Welsch 估计配准参数时,首先为代价函数赋予一个较大的权重,在迭代的过程中利用控制参数剔除残差较大的观测值,计算时观测值是动态变化的。避免了尺度相同的代价函数,在某次迭代中使残差较大值获得了较小的权重,但无法避免在下次迭代中获得了较大的权重,容易出现迭代震荡的现象。

RANSAC 通过最小集仅能估计到近似近解,最优解需要所有观测值的参与。RANSAC 解算空间变化参数具有随机性,难以保证局部最优解。

自适应尺度的 Welsch 相比 RANSAC 估计的空间变化参数,均有较大的提高。另一个重要的观察

结果是通过比较配准精度,我们发现 B-SHOT 的配准精度最差,这是因为浮点型的 SHOT 对点云分辨率和噪声较为敏感,因此导致了其二进制描述符的鲁棒性也相对较差。BRoPH 所采用的多视角机制以及密度分布和深度信息的整合相辅相成,增强了其描述性,而中心对称局部二进制模式(CS-LBP)的多尺度扩展则为生成二进制字符串提供了一种高效、紧凑的方法。因此,BRoPH 的配准精度次之。B-RCS 虽然也采用视角编码局部曲面的空间信息,但其采用简单的对比方式进行二值化,致使信息损失严重。LoVS 简单地使用点空间位置将点云表示的局部形状结构编码为位串,由于局部曲面使用立方体进行划分,立方体的大小会对 LoVS 的描述性产生严重的影响,对质量较差的点云鲁棒性较弱。



因此,基于 LoVS 描述的配准精度也较差。BDIF 采用三个正交视角全面解析局部曲面的空间信息,且利用双边滤波对“空洞”进行赋值,此种方式能很好保留局部曲面的几何特征。因此基于 BDIF 的配准精度最高。

### 5.2.2 时间

表4统计了基于 BDIF 配准过程中的时间消耗。表中  $S_s$ 、 $T_s$ 、TK、TF、TM、TC 和 TS 分别代表源点云中的点数、目标点云中的点数、关键点提取的时间、BDIF 的计算时间、特征匹配和筛选的时间、空间变化参数的计算时间和总的时间消耗。

总的来说,所提的算法能在几秒钟内成功的完成

大尺度的点云配准,点云数据量高达3百万到1千万。所提算法的大部分时间都消耗在特征匹配的过程中。虽然 BDIF 是一个具有强鲁棒性的二进制描述符,且这种类型的描述符匹配速度较快,但实验中采用的数据均是大场景采集到的点云。具体来说, Park、Mountain 和 Campus 场景的点云含有大量的特征点,进而导致特征匹配和匹配对筛选耗费了大量的时间。从4表中可以发现,计算空间变换参数所花费的时间在整个配准过程中所占的比重最小且能达到毫秒级,这得益于所提配准算法采用自适应尺度的 Welsch 进行鲁棒估计,该方法在迭代过程中通过步长控制权重能快速剔除残差较大的观测值。

表4 配准时间消耗

Tab. 4 Registration time consumption

	$S_s$ /百万	$T_s$ /百万	TK/s	TF/s	TM/s	TC/s	TS/s
Park	5.65	5.06	1.485	0.682	2.922	0.004	5.093
Mountain	3.40	2.70	1.435	0.639	2.962	0.005	5.041
Campus	12.44	11.68	3.308	1.34	5.487	0.005	10.14

## 6 结论

本文提出了一种新的二进制局部特征描述符 BDIF,以及一种基于 BDIF 的点云配准算法。BDIF 的基本思想是在三个正交视角上根据局部曲面的投影高度对关键点的空间信息进行编码。与以往的大多数方法相比,BDIF 采用双边滤波对“空洞”进行填充,保持局部曲面的边缘特征。BDIF 的另外一个特殊点在于它是以二进制的形式进行表示的,这使得它能降低对存储内存的要求和提高匹配速率。本文根据 BDIF 提出了一种点云配准算法,它能快速实现大场景的点云正确配准,这对于数据城市的建设具有重大现实意义。

我们进行了多个实验来评估我们的方法。实验中评估了 BDIF 对噪声、点云分辨率、杂波、遮挡和不规则点云的鲁棒性。与现有的二进制描述符相比,BDIF 均表现出较优的性能。基于 BDIF 的配准算法在 WHU-TLS 数据集上进行了测试。定性和定量结果表明,所提配准算有得益于 BDIF 强有力的描述性和自适应尺度的 Welsch 从而完成了高质量的点云配准。

### 参考文献:

[1] Quan S, Ma J, Hu F, et al. Local voxelized structure for

3D binary feature representation and robust registration of point clouds from low-cost sensors [J]. Information Sciences, 2018, 444: 153 - 171.

- [2] Li J Y. A practical outlier removal method for correspondence-based point cloud registration clouds from low-cost sensors [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2022, 44(18): 3926 - 3939.
- [3] Wang Mingjun, Yi Fang, Li Le, et al. Local neighborhood feature point extraction and matching for point cloud alignment [J]. Infrared and Laser Engineering, 2022, 51(5): 431 - 440. (in Chinese)  
王明军, 易芳, 李乐, 等. 自适应局部邻域特征点提取和匹配的点云配准 [J]. 红外与激光工程, 2022, 51(5): 431 - 440.
- [4] Besl P J, McKay H D. A method for registration of 3-D shapes [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1992(2): 239 - 256.
- [5] Lv C, Lin W, Zhao B. KSS-ICP: point cloud registration based on Kendall shape space [J]. IEEE Transactions on Image Processing, 2023, 32: 1681 - 1693.
- [6] Shi X, Liu T, Han X. Improved iterative closest point (ICP) 3D point cloud registration algorithm based on point cloud filtering and adaptive fireworks for coarse registration [J]. International Journal of Remote Sensing, 2020, 41(8): 3197 - 3220.

- [7] Johnson A E, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(5): 433 – 449.
- [8] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration [C]//: *IEEE International Conference on Robotics and Automation*. 2009.
- [9] Guo Y L, Sohel F B M, Liu M, et al. Rotational projection statistics for 3D local surface description and object recognition [J]. *International Journal of Computer Vision*, 2013, 105: 63 – 86.
- [10] Yang J Q, Zhang Q, Xiao Y, et al. Toldi: an effective and robust approach for 3D local shape description [J]. *Pattern Recognition*, 2017, 65: 175 – 187.
- [11] Yang J Q, Zhang Q, Xian K, et al. Rotational contour signatures for both real-valued and binary feature representations of 3D local shape [J]. *Computer Vision and Image Understanding*, 2017, 160: 133 – 147.
- [12] Calonder M, Lepetit V, Ozuysal M, et al. Brief: computing a local binary descriptor very fast [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1281 – 1298.
- [13] Rublee E, Rabaud V, Konolige K, et al. Orb: an efficient alternative to sift or surf [C]//: *IEEE International Conference on Computer Vision (ICCV)*, 2011.
- [14] Prakhya S M, Liu B B, Lin W S, et al. B-shot: a binary feature descriptor for fast and efficient keypoint matching on 3D point clouds [J]. *Autonomous Robots*, 2017, 41(7): 1501 – 1520.
- [15] Tombari F, Salti S, Di Stefano L. Shot: unique signatures of histograms for surface and texture description [J]. *Computer Vision and Image Understanding*, 2014, 125: 251 – 264.
- [16] Dong Z, Yang B S, Liu Y, et al. A novel binary shape context for 3D local surface description [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, 130: 431 – 452.
- [17] Ai Jiaqiu, Wang Feifa, Yang Xingming, et al. SAR image speckle noise suppression algorithm based on background homogeneity and bilateral filtering [J]. *Journal of Remote Sensing*, 2021, 25(5): 1071 – 1084. (in Chinese)
- 艾加秋, 王非凡, 杨兴明, 等. 基于背景匀质性双边滤波的 SAR 图像斑点噪声抑制算法 [J]. *遥感学报*, 2021, 25(5): 1071 – 1084.
- [18] He Zhiyong, Sun Lining, Huang Weiguo, et al. Thresholding segmentation algorithm based on Otsu criterion and line intercept histogram [J]. *Optics and Precision Engineering*, 2012, 20(10): 2315 – 2323. (in Chinese)
- 何志勇, 孙立宁, 黄伟国, 等. 基于 Otsu 准则和直线截距直方图的阈值分割 [J]. *光学精密工程*, 2012, 20(10): 2315 – 2323.
- [19] Zhou Yunyan, Yang Kuntao, Huang Ying. Improved Otsu thresholding based on minimum inner cluster variance [J]. *Journal of Huazhong University of Science and Technology: Natural Science Edition*, 2007, (2): 101 – 103. (in Chinese)
- 周云燕, 杨坤涛, 黄鹰. 基于最小类内离散度的改进 Otsu 分割方法的研究 [J]. *华中科技大学学报: 自然科学版*, 2007, (2): 101 – 103.
- [20] Guo Y L, Bennamoun M, Sohel F, et al. 3D object recognition in cluttered scenes with local surface features: a survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(11): 2270 – 2287.
- [21] Mian A, M B, Owens R. On the repeatability and quality of keypoints for local feature-based 3D object retrieval from cluttered scenes [J]. *International Journal of Computer Vision*, 2010, 89(2): 348 – 361.
- [22] Qiu Liya, Chen Weilin, Li Fanming, et al. Fast Hash\_LBP moving target detection algorithm based on Hamming distance constraint in complex background [J]. *Acta Photonica Sinica*, 2022, 51(9): 910003. (in Chinese)
- 裘莉娅, 陈玮琳, 李范鸣, 等. 复杂背景下基于汉明距离约束的 Hash\_LBP 运动目标快速检测算法 [J]. *光子学报*, 2022, 51(9): 910003.
- [23] Zou Y W X Z T, Delearde R, Kurtz C, et al. Broph: an efficient and compact binary descriptor for 3D point clouds [J]. *Pattern Recognition*, 2018, 76: 522 – 536.
- [24] Dong Z, Liang F X, Yang B S, et al. Registration of large-scale terrestrial laser scanner point clouds: a review and benchmark [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 163: 327 – 342.