澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Mutual Voting for Ranking 3D Correspondences

翁霄羽
wengxiaoyu2-009@163.com

一、背景介绍

3D点云特征匹配

1.特征匹配：需要在源点云和目标点云的点云序列中建立正确的点对点对应关系

2.是点云配准、三维物体识别等问题的基础

3.一般通过两个步骤进行匹配：生成初始对应关系、内点选择

一、背景介绍

基于几何方法的内点选择:

（1）label-based：通过分类器给点云对应关系分配二进制标签，找到形成一簇的正确的对应关系

缺点：在区分离群点和内点方面的判别能力有限

（2）score-based：通过计算输入对应项的置信度分数，根据评分结果对对应项进行排序（评分函数通常根据特征相似度或几何约束定义）

二、Mutual-voting（MV）

Mutual-voting（MV）：一种基于投票的内点选择方法

方法流程:

（1）图的构造：初始对应集被建模为兼容性图，其中每个节点表示单个对应关系，两个节点之间的每条边表示一对几何上兼容的对应关系。能够准确地呈现无序对应关系之间的亲和关系

（2）节点聚类系数计算：在复杂网络中，聚类系数用于衡量图节点对周围环境的联系程度，描述网络的密集程度或稀疏程度，这个概念被引入到MV方法中，它能够初步去除一部分异常值，为接下来的相互投票提供更好的基础

（3）相互投票：通过相互投票，每个对应关系都得到一个分数，选民和候选人双向投票，实现一个令人信服的得分结果。由于高质量的投票集合是产生可靠结果的基础，我们强制"候选人→选民"投票，以减少不可靠的选民，另外还有"选民→候选人"投票，形成了一个相互投票的方案

（4）对应关系排序：所有输入的对应关系根据投票分数降序排序，得分最高的作为MV选择的内点

二、 Mutual-voting（MV）

图的构造

The voting process will be performed in a graph. Compared with the Euclidean space, the graph space can better render the affinity relationship among correspondences. The initial correspondence set is modeled as a compatibility graph, where nodes represent correspondences and edges connect geometrically compatible nodes.

In particular, let $\mathbf{p}_i^s$ and $\mathbf{p}_i^t$ denote the points in the source point cloud $\mathbf{P}^s$ and target point cloud $\mathbf{P}^t$, respectively. Then, the rigidity between the two correspondences $\mathbf{c}_i$ and $\mathbf{c}_j$ can be quantitatively measured as:

$$S_{dist}(\mathbf{c}_i, \mathbf{c}_j) = \left| \left\| \mathbf{p}_i^s - \mathbf{p}_j^s \right\| - \left\| \mathbf{p}_i^t - \mathbf{p}_j^t \right\| \right|. \tag{1}$$

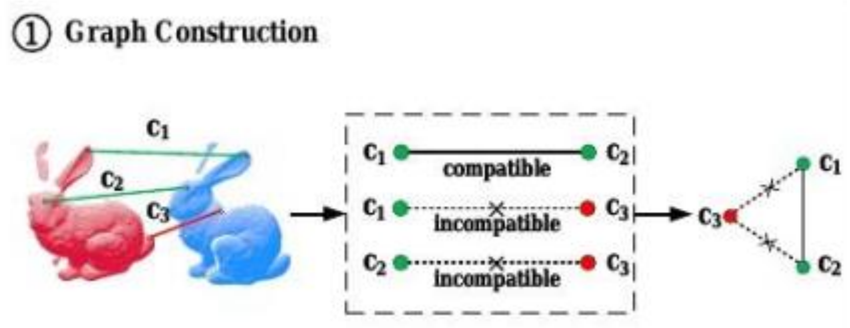
The compatibility score between $\mathbf{c}_i$ and $\mathbf{c}_j$ is given as:

$$S_{cmp}(\mathbf{c}_i, \mathbf{c}_j) = \exp(-\frac{S_{dist}(\mathbf{c}_i, \mathbf{c}_j)^2}{2d_{cmp}^2}), \tag{2}$$

参数

where d_{cmp} is a distance parameter and $S_{cmp} \in [0, 1]$. Ideally, $S_{cmp}(\mathbf{c}_i, \mathbf{c}_j) = 1$ if $\mathbf{c}_i$ and $\mathbf{c}_j$ are inliers.

Subsequently, as shown in Fig.1(a), given a set of initial correspondences $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_n\}$, we model them as a graph $G = (\mathbf{V}, \mathbf{E})$. Here, $\mathbf{V} = \{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_n\}$ and $\mathbf{E} = \{\mathbf{e}_{12}, \mathbf{e}_{13}, ..., \mathbf{e}_{ij}\}$ with $\mathbf{e}_{ij} = (\mathbf{c}_i, \mathbf{c}_j)$. Notably, if $S_{cmp}(\mathbf{c}_i, \mathbf{c}_j)$ is greater than a threshold t_{cmp}, $\mathbf{c}_i$ and $\mathbf{c}_j$ form an edge and $S_{cmp}(\mathbf{c}_i, \mathbf{c}_j)$ is the weight of the edge. To this end, a graph is generated for $\mathbf{C}$.



(a) Graph construction rule

$S_{cmp}(c_i, c_j) = 1 \Longrightarrow S_{dist}(c_i, c_j) = 0$

目标点云中i到j点的距离和源点云中i到j点的距离相同，所以Ci和Cj是内点

二、 Mutual-voting（MV）

节点聚类系数计算

3.2.1 Basic Concepts of the Clustering Coefficient

Let the degree of node $\mathbf{c}_i$ be d_i, then there are at most $d_i(d_i-1)/2$ edges among d_i neighbor nodes. Let w_i denote the number of edges that actually exist among the neighboring nodes of node c_i. In a weighted network, w_i denotes the sum of the weights of these edges. The clustering coefficient α_i for $\mathbf{c}_i$, as illustrated in Fig. 1(b), can be defined as:

$$\alpha_i = \frac{w_i}{(d_i*(d_i-1))/2}. \tag{3}$$

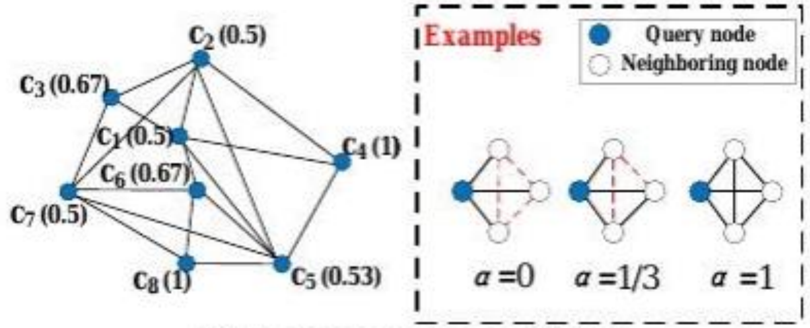
i:节点

W_i:通过i邻节点的关系数.

d_i: i的一阶邻节点

The nodal clustering coefficients reflect the significance of the nodes in the network and the degree of local aggregation of the network. In addition, average clustering coefficient $\overline{\alpha}$ and overall clustering coefficient $\alpha_{overall}$ can be used to express the degree of aggregation of the whole network, as defined in the following:

② Nodal Clustering Coefficient Calculation



(b) Nodal clustering coefficient

平均聚类系数 $$\overline{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\alpha_i, \tag{4}$$

全局聚类系数 $$\alpha_{overall} = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n}(d_i*(d_i-1))/2}. \tag{5}$$

二、 Mutual-voting（MV）

节点聚类系数计算

3.2.2 Application of Nodal Clustering Coefficient in MV

i) Remove outliers preliminarily. Inliers are consistent, and therefore are supposed to be compatible with each other. As such, inliers are more likely to form cliques in a graph. It is interesting to note that nodes with greater nodal clustering coefficients are more likely to be in cliques. Therefore, we set an adaptive threshold t_α to eliminate nodes with low nodal clustering coefficients, which is defined as:

$$t_\alpha = \min(\alpha_{overall}, \overline{\alpha}, otsu_\alpha), \tag{6}$$

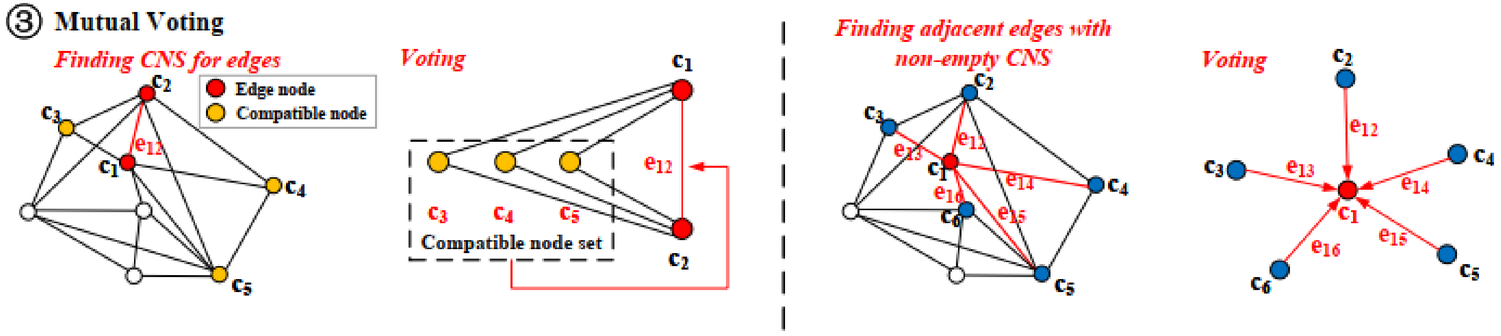
最大类间方差法确定阈值

where $otsu_\alpha$ represents the OTSU [40] threshold based on the nodal clustering coefficients of all nodes. The leverage of nodal clustering coefficient has two merits. First, a portion of outliers can be removed, which would alleviate the negative effects of outliers in the following mutual voting process. Second, less nodes will be involved in the voting process, therefore speeding up the selection process.

ii) Be the weight in the voting process. The nodal clustering coefficient is an informative cue for nodes. It will participate the "node→edge" voting process (Sect. 3.3)

1.剔除离群点
2.成为投票过程中的权重

二、 Mutual-voting（MV）



(c) Node→Edge voting
(d) Edge→Node voting

相互投票

（1）节点⟶边投票

由每条边的兼容节点集 $C_{cmp}(e_{ij})$ 进行投票，分数记为 $S(e_{ij})$

$$S(\mathbf{e}_{ij}) = \sum_{\mathbf{c}_k \in \mathbf{C}_{cmp}(\mathbf{e}_{ij})} \frac{\alpha_i + \alpha_j + \alpha_k}{3} [S_{cmp}(\mathbf{c}_i, \mathbf{c}_j) + S_{cmp}(\mathbf{c}_i, \mathbf{c}_k) + S_{cmp}(\mathbf{c}_j, \mathbf{c}_k)].$$

（2）边⟶节点投票

节点通过其兼容节点集不为0的相邻边进行投票，节点 C_i 的投票分数定义为

$$S(\mathbf{c}_i) = \sum_{\mathbf{C}_{cmp}(\mathbf{e}_{ij}) \neq \emptyset} S(\mathbf{e}_{ij})$$

二、 Mutual-voting（MV）

根据投票分数 $S(c_i)$，对初始对应集合进行降序排序，得到MV的最终输出

MV 输出的确定是灵活的。一种是采用 OTSU阈值化策略。另一种是选择 top-K 的作为内点，其中 K 可以根据特定的应用场景进行调整

三、实验评估

本篇文章介绍了三种类型的实验来评估MV方法:

（1）特征匹配

（2）点云配准

（3）3D 物体识别实验。

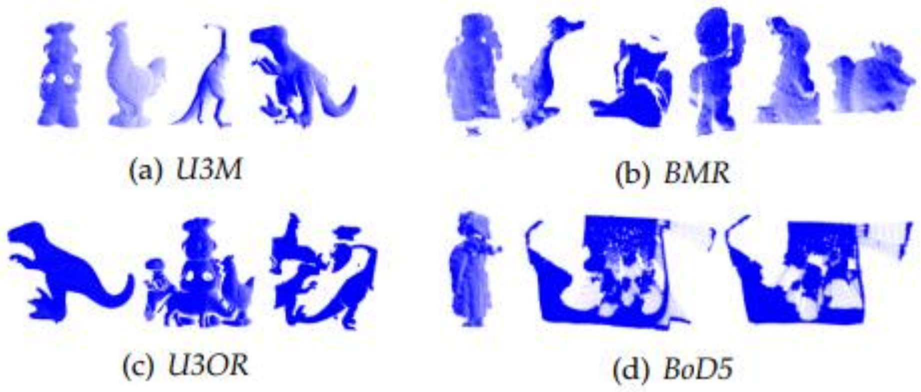
三、实验-特征匹配

特征匹配实验使用的数据集：

TABLE 1
Properties of feature matching experimental datasets.

Dataset	Data modality	Nuisances	Application scenario	# Matching pairs	Avg. inlier ratio
U3M [41]	LiDAR	Limited overlap, self-occlusion	Object Registration	496	14.80%
BMR [42]	Kinect	Limited overlap, self-occlusion, real noise	Object Registration	485	5.63%
U3OR [43], [44]	LiDAR	Clutter, occlusion	Object Recognition	188	8.09%
BoD5 [42]	Kinect	Clutter, occlusion, real noise, holes	Object Recogniiton	43	15.75%

数据集的可视化样本图：



(a) *U3M* (b) *BMR*
(c) *U3OR* (d) *BoD5*

三、实验-特征匹配

评价指标：对于一致的 $c=(p^s, p^t)$，当满足条件 $\left\|\mathbf{R}_{gt}\mathbf{p}^s + \mathbf{t}_{gt} - \mathbf{p}^t\right\| < d_{inlier}$ 时，则判定匹配成功

R_{gt}：基准真值旋转矩阵，改变向量方向，不改变向量大小的矩阵

t_{gt}：平移向量

d_{inlier}：距离阈值，本次实验中设置为5pr

pr：一个称为点云分辨率的距离单位，它是点云中一个点到最近邻居的最近距离的平均值

K的内点召回率：

$$recall_K = \frac{\#inliers\ in\ \mathbf{C}_K}{\#inliers\ in\ \mathbf{C}_{initial}}. \qquad (12)$$

C_k：top-k中的对应集合

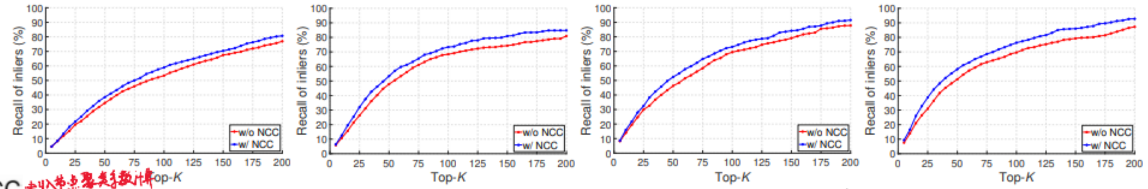
$C_{initial}$：初始的对应集合

三、实验-特征匹配

比较方法：MV 与 9 种现有的最先进的特征匹配方法进行了比较。这些方法是相似评分(SS)， 最近邻相似比 (NNSR)，光谱技术 (ST)，搜索内线(SI)，博弈论匹配(GTM）和一致性投票(CV)， PointDSC， TEASER++， SC2-PCR，其中 SI和 CV 也是基于投票的方法

实验使用点云库(PCL)实现

三、实验-特征匹配



w/o NCC 未引入节点聚类系数计算
w/ NCC 引入节点聚类系数计算步骤

(a) *Chef*　(b) *Chicken*　(c) *Parasaurolophus* 副栉龙　(d) *T-rex* 霸王龙

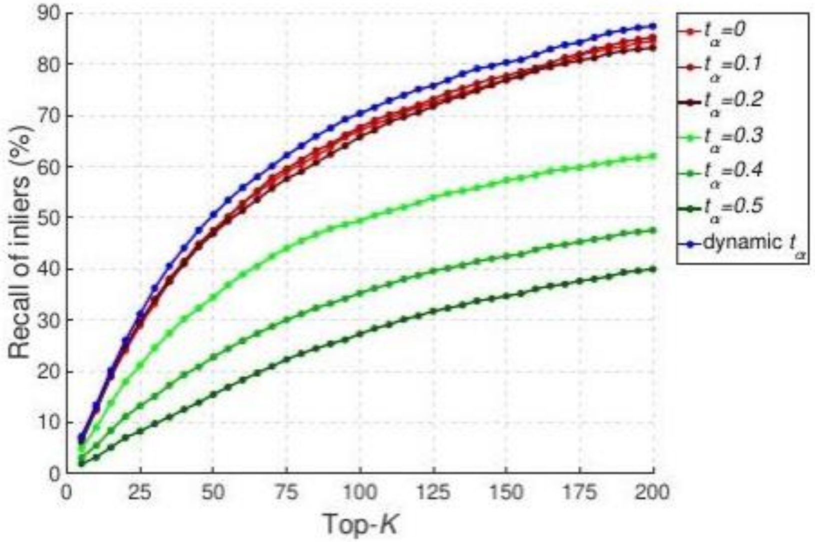
实验方法分析：

（1）兼容性约束：如图所示，在只有距离约束的情况下可以实现更稳定的召回率

TABLE 2
Recall performance (%) of applying different constraints to MV on the U3M dataset.

Std. of Gaussian noise (pr)	1.0	1.5	2.0	2.5	3.0
Distance+angular	71.39	68.32	60.17	51.51	48.00
Distance-only	**74.43**	**72.92**	**65.45**	**64.21**	**62.95**

（2）设置自适应的聚类系数阈值：设置自适应阈值能更好的剔除离群点，且相较于固定的聚类系数阈值，召回率明显提高，特征匹配的结果更好



（3）消融实验：为了验证使 用节点聚 类系数去除部分离群点的必要性，进行了消融实验，比较了有节点聚类系数计算步骤和没有节点聚类系数计算步骤时 MV 的召回率 （数据集：U3M）

三、实验-特征匹配

实验结果

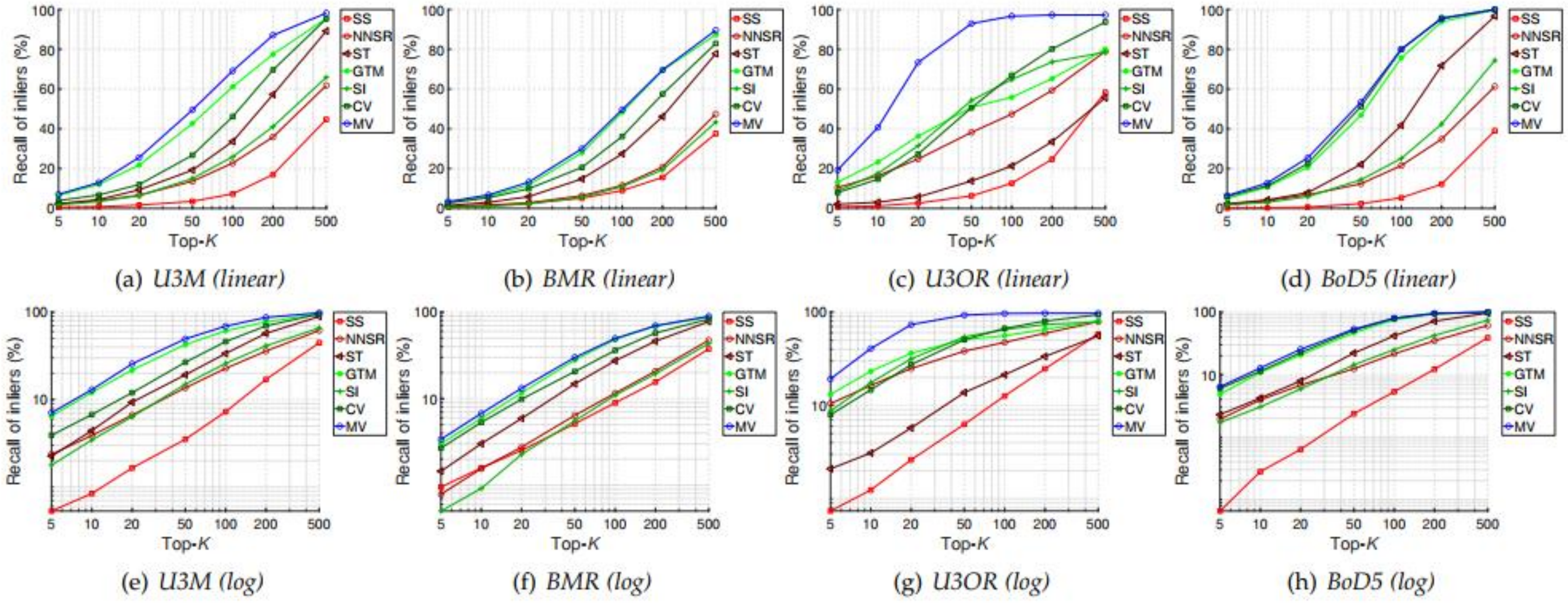


Fig. 5. Feature matching performance of tested methods on four feature matching datasets.

结果分析:
测试的基于分数的方法在实验数据集上的特征匹配结果如图所示。对于每个数据集，分别给出了具有线性和对数召回轴图的两幅图，它们可以更清晰地反映出 K值较小和较大时的性能差距。
在所有的实验数据集上， MV 的表现都优于其他两种，这表明MV可以推广到不同的应用场景和数据模态。在 U3OR 数据集上， MV 比其他方法有更明显的差距。从表1（见前页）中可以看出U3OR 的平均初始比率相对较低(8.09%)，从而验证了 MV 对低初始比率的稳健性。
下图可视化了几种测试方法的特征匹配结果。

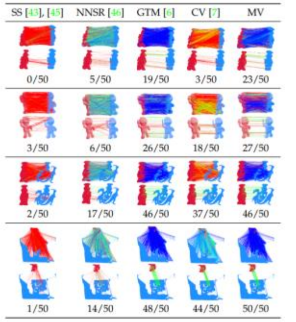


Fig. 8. Visual feature matching results of several tested methods. Each method has two visual results, where the top one is the scoring result rendered by pseudo-color (red→blue: high → low confidence scores) and the bottom one is the selection result (x/K indicates x inliers in the selected K correspondences). Green and red lines refer to correct and incorrect correspondences, respectively.)

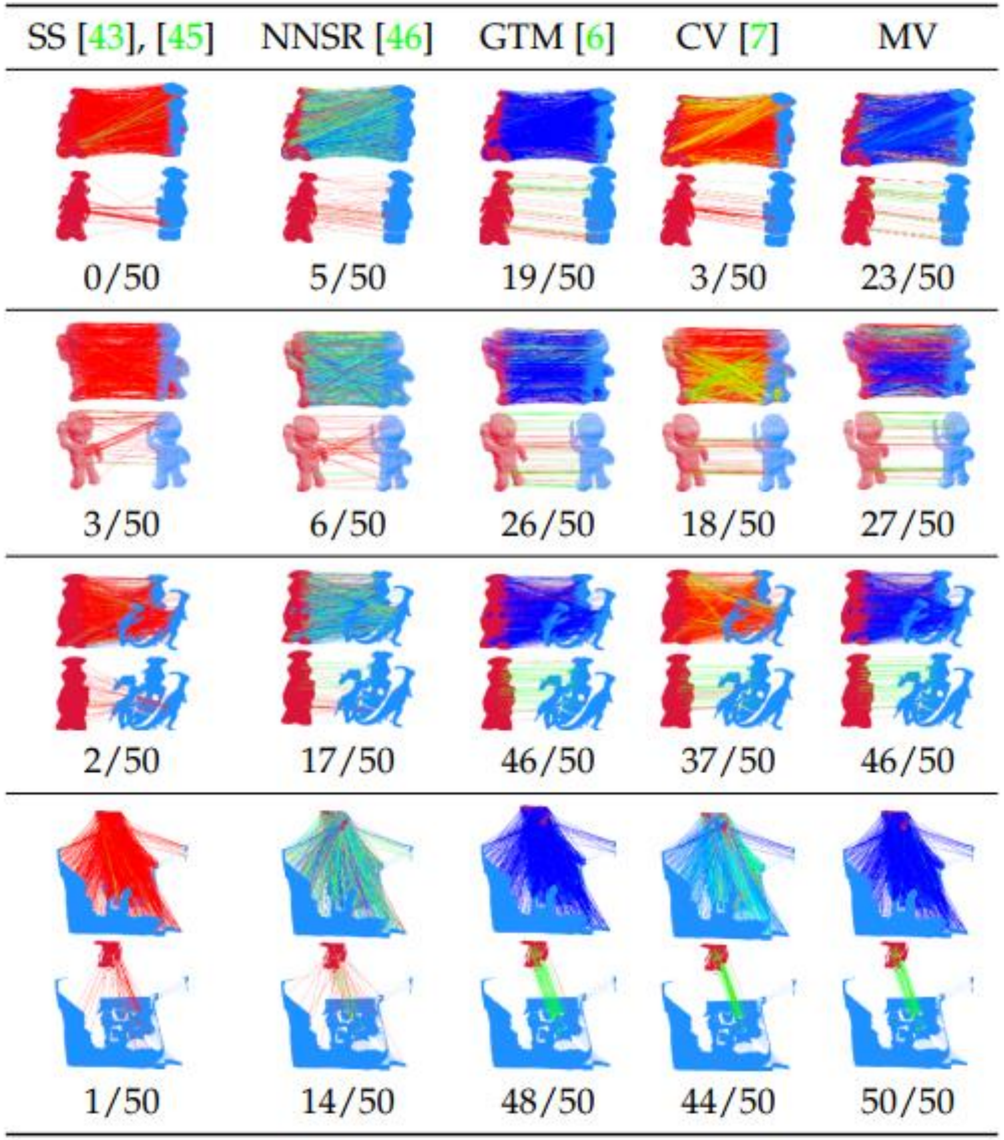


Fig. 8. Visual feature matching results of several tested methods. Each method has two visual results, where the top one is the scoring result rendered by pseudo-color (red→blue: high → low confidence scores) and the bottom one is the selection result (x/K indicates x inliers in the selected K correspondences). Green and red lines refer to correct and incorrect correspondences, respectively.)

三、实验-特征匹配

实验结果

除了召回率结果，实验还进一步得出了其他指标的结果，如下表所示：

TABLE 3
Feature matching performance (%) on U3M.

	Precision	Recall	F-score
SS [43], [45]	4.81	63.39	8.94
NNSR [46]	9.67	41.63	15.69
ST [3]	4.31	**100.00**	8.26
GTM [6]	35.63	46.22	40.24
SI [47]	14.28	43.81	21.54
CV [7]	13.31	94.92	23.35
PointDSC [38]	9.09	0.23	0.45
SC^2-PCR [9]	**49.81**	20.74	29.29
TEASER++ [34]	1.03	0.56	0.73
MV	31.00	78.26	**44.41**

TABLE 4
Feature matching performance (%) on BMR.

	Precision	Recall	F-score
SS [43], [45]	2.68	54.41	5.11
NNSR [46]	4.41	29.11	7.66
ST [3]	2.89	**100.00**	5.62
GTM [6]	**23.07**	26.07	24.48
SI [47]	4.57	14.00	6.89
CV [7]	8.04	88.29	14.74
PointDSC [38]	5.43	86.03	10.21
SC^2-PCR [9]	3.11	**100.00**	6.03
TEASER++ [34]	19.13	42.53	26.39
MV	18.19	62.8	**28.21**

TABLE 5
Feature matching performance (%) on U3OR.

	Precision	Recall	F-score
SS [43], [45]	2.40	80.53	4.66
NNSR [46]	6.42	63.45	11.66
ST [3]	1.69	**99.49**	3.32
GTM [6]	43.55	43.28	43.41
SI [47]	24.27	60.11	34.58
CV [7]	7.34	90.63	13.58
PointDSC [38]	3.84	0.27	0.50
SC^2-PCR [9]	62.34	45.14	52.36
TEASER++ [34]	0.19	1.30	0.34
MV	**64.62**	90.39	**75.36**

TABLE 6
Feature matching performance (%) on BoD5.

	Precision	Recall	F-score
SS [43], [45]	6.29	52.03	11.22
NNSR [46]	12.84	39.61	19.39
ST [3]	11.12	**100.00**	20.01
GTM [6]	**65.23**	50.18	56.72
SI [47]	27.27	41.80	33.01
CV [7]	37.96	99.98	55.03
PointDSC [38]	28.59	96.12	44.07
SC^2-PCR [9]	6.59	**100.00**	12.37
TEASER++ [34]	52.99	82.33	64.48
MV	55.12	94.39	**69.60**

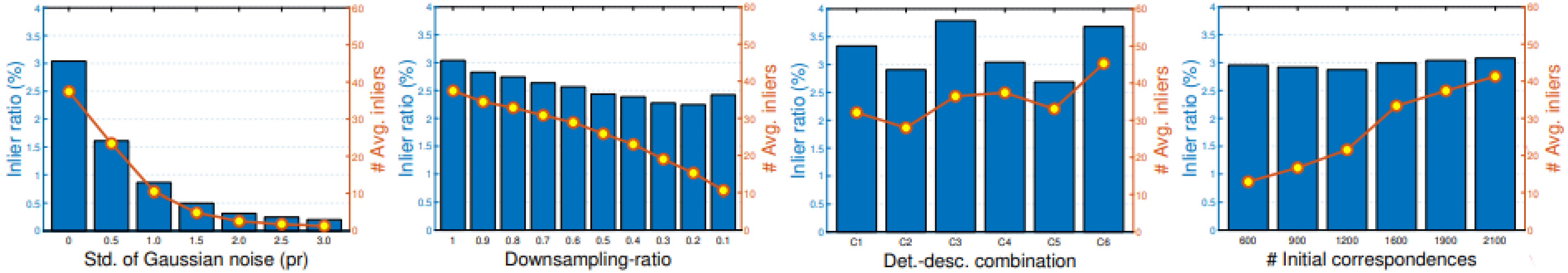
结果分析

MV 在所有数据集上的 F-score 表现都是最好的。当应用于对象数据集时， PointDSC 的性能非常差。在跨数据集实验中，所有比较方法的性能都有明显的波动。相比之下，MV 在这种情况下取得了稳定和突出的性能。

三、实验-特征匹配

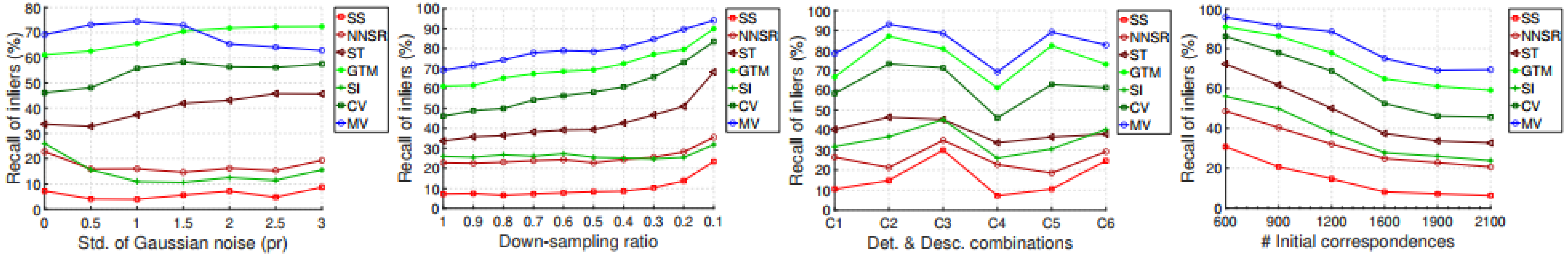
鲁棒性结果

实验分析了在存在高斯噪声、点密度变化、不同detector-descriptor 组合以及不同初始特征匹配数量的情况下，所有比较方法在 U3M 数据集上的鲁棒性。这些干扰对输入的影响在统计上如图 6 所示。将 K 设为 100，结果如图 7 所示。



(a) *Gaussian noise* 高斯噪声 (b) *Down-sampling* 降低采样率 (c) *Different det.-desc.* 不同的C集. (d) *Diff. input scales* 不同的输入规模.

Fig. 6. Information in terms of the inlier ratio and the number of inliers of the input correspondence sets under different experimental configurations on the U3M dataset.



(a) *Varying noise* (b) *Density variation* (c) *Varying det.-desc.* (d) *Varying input scales*

Fig. 7. Robust performance of tested methods with respect to different nuisances on U3M.

三、实验-特征匹配

MV方法的时间效率

为了比较测试特征匹配方法的时间效率，这里分析了两种情况:具有不同量级和不同内点比率的输入。对于前者，初始对应数设置为 600、 900、 1200、 1600 和 1900;对于后者，输入幅值固定为 1200，点云对分为 6 组，如图 8 所示是几种测试方法的视觉特征匹配结果。每种方法都有两个视觉结果，上面的是伪颜色渲染的评分结果(红色→ 蓝色:高置信度分数→低置信度分数)，下面的是选择结果(*x*/K 表示选择的 K 个对应中的 *x*内点)。绿色和红色的线分别表示正确和错误的对应关系。)其内点比范围为 0% - 30%，步长为 5%。实验在 U3M 数据集上进行，所测试方法匹配单点云对的平均时间成本如表 7 和表 8 所示。从表 7 中可以观察到，随着输入对应幅度的增加，每种方法所花费的时间成本都有增加的趋势。特别地，当输入对应数小于 1600 时， MV 排名靠前。从表 8 中可以发现，当内点比小于 5% 时， MV 是效率最高的。随着内点比的进一步提高， MV 的时间成本有所提高但仍然保持高效，因为当内点比在 10% - 30%之间时，它能够以 0.1秒左右的时间完成内点选择。总的来说， MV 是一种高效的 3D 特征匹配方法， MV 的召回性能明显优于其他对比方法。

TABLE 7
Varying the number of input correspondences (ms).

# Initial correspondences	600	900	1200	1600	1900
SS [43], [45]	54.40	86.39	135.00	**278.64**	**319.18**
NNSR [46]	61.73	97.90	154.56	318.31	366.26
ST [3]	1946.24	3560.74	5983.80	15007.80	17611.20
GTM [6]	**36.63**	**65.65**	**112.99**	288.72	336.87
SI [47]	95.52	150.32	241.58	519.92	600.39
CV [7]	102.76	177.85	297.38	709.77	844.76
MV	39.75	73.52	133.09	399.95	510.14

TABLE 8
Varying the inlier ratio of input correspondences (ms).

Inlier ratio	0%~5%	5%~10%	10%~15%	15%~20%	20%~25%	25%~30%
SS [43], [45]	53.79	54.76	**51.23**	34.89	**41.00**	31.00
NNSR [46]	60.94	63.54	59.07	43.89	50.50	38.00
ST [3]	2981.73	3099.22	3060.30	2017.33	2548.50	1580.00
GTM [6]	50.93	**52.89**	52.77	**34.44**	43.33	**26.67**
SI [47]	107.18	112.05	106.07	80.78	93.33	72.33
CV [7]	145.16	151.95	147.43	97.44	130.33	80.00
MV	**35.86**	57.84	103.30	107.11	207.33	115.33

三、实验-点云配准

点云配准实验使用的数据集：物体数据集 U3M，场景室内数据集3DMatch、3DLoMatch和场景室外数据集 KITTI， 3DLoMatch是3DMatch的子集

均方根误差

Evaluation metric. We follow [57] that employs the root mean square error (RMSE) metric to evaluate the 3D point cloud registration performance on the object-scale dataset, e.g., U3M. Given the estimated rotation matrix $\mathbf{R}_{est}$ and translation vector $\mathbf{t}_{est}$, the point-wise error ε_{p} between two

Rest: 旋转矩阵
test: 平移向量

truly corresponding points $\mathbf{p}^s$ and $\mathbf{p}^t$ is defined as:

逐点误差

$$\varepsilon_{\mathrm{p}}(\mathbf{p}^s, \mathbf{p}^t) = ||\mathbf{R}_{est}\mathbf{p}^s + \mathbf{t}_{est} - \mathbf{p}^t||. \tag{13}$$

Then, the definition of RMSE is:

$$\mathrm{RMSE} = \sqrt{\sum_{\mathbf{p}^s, \mathbf{p}^t \in \mathbf{C}_{gt}} \frac{\varepsilon_{\mathrm{p}}^2(\mathbf{p}^s, \mathbf{p}^t)}{|\mathbf{C}_{gt}|}}, \tag{14}$$

Cgt: 目标点云集和源点云集对应点的基准真值

where $\mathbf{C}_{gt}$ represents the ground-truth set of corresponding points between two point clouds. When the RMSE of a registration is smaller than a threshold t_{rmse}, we judge it as success.

阈值

RE: 旋转误差
TE: 平移误差

Also, we employ the rotation error (RE) and translation error (TE) to evaluate the registration results on scene-scale dataset. Given the estimated rotation matrix $\mathbf{R}_{est}$ and ground-truth rotation matrix $\mathbf{R}_{gt}$, estimated translation vector $\mathbf{t}_{est}$ and ground-truth translation vector $\mathbf{t}_{gt}$, RE and TE can be defined as:

Rgt: 基准真值旋转矩阵

$$\mathrm{RE}(\mathbf{R}_{est}) = \arccos \frac{\mathrm{Tr}(\mathbf{R}_{est}{}^{\top}\mathbf{R}_{gt}) - 1}{2}, \tag{15}$$

$$\mathrm{TE}(\mathbf{t}_{est}) = ||\mathbf{t}_{est} - \mathbf{t}_{gt}||_2. \tag{16}$$

By referring to the settings in [35], the registration is considered successful when the RE $\leq 15°$, TE ≤ 30 cm on 3DMatch & 3DLoMatch datasets, and RE $\leq 5°$, TE ≤ 60 cm on KITTI dataset. For a dataset, we define its registration accuracy as the ratio of success cases to the total number of point cloud pairs to be registered.

$$配准精确度 = \frac{成功配准对数}{待配准总对数}$$

$RMSE < t_{rmse}$ 判定配准成功

三、实验-点云配准

实验细节：我们使用基于RANSAC的方法，从一组对应关系中预测一个配准位置来进行配准。更具体地说，是使用 MV进行点集中对应关系的选择，并将MV 的输出作为 RANSAC 估计器的输入。默认情况下，我们使用 5000次RANSAC 迭代来执行配准。注意，还可以考虑更好的估计器，它们与MV 互补，因为 MV 的输出是刚性变换估计器的输入

方法介绍

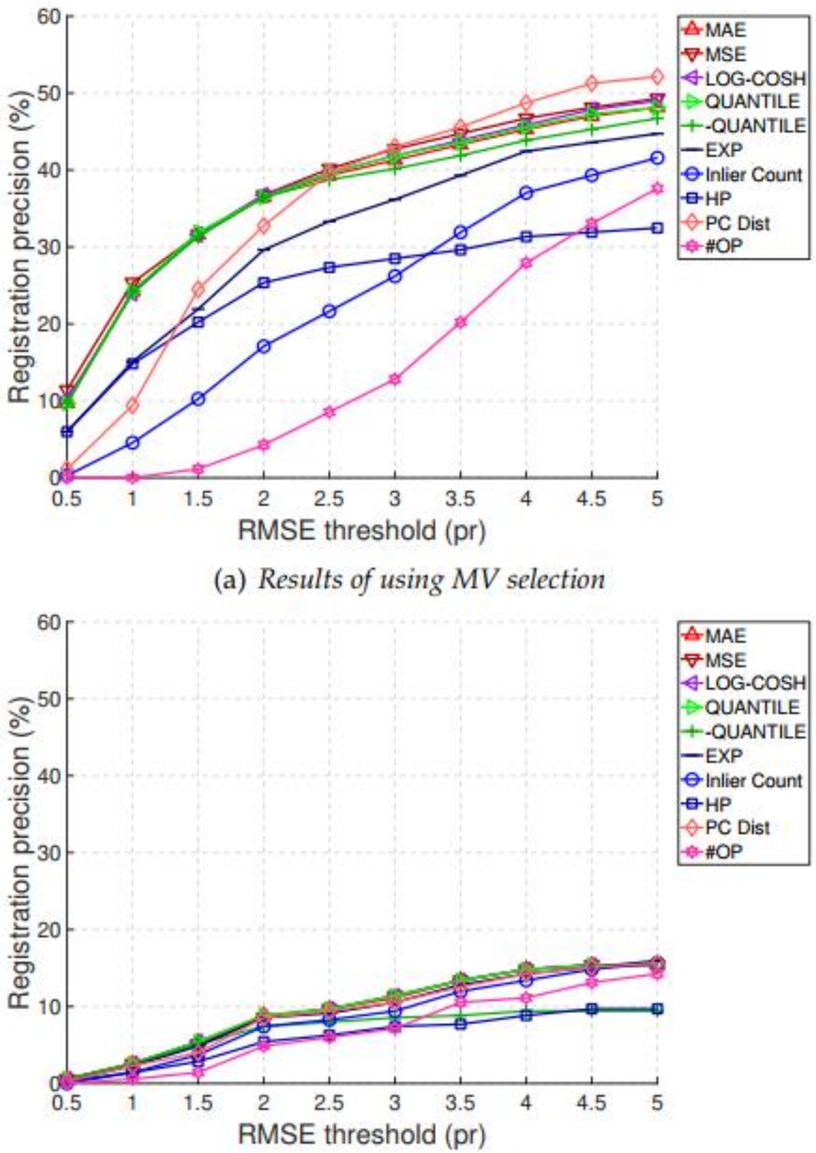
RANSAC（随机抽样一致性方法）：该方法可以通过迭代的方式在一组包含离群值的数据集中找到最优参数模型，在图像对齐和拼接中得到广泛应用

SC^2-PCR：引入一种二阶空间相容性度量来衡量对应关系之间的亲密度，利用全局谱技术对对应关系进行排序，然后得分最高的对应关系被作为种子来计算三维点云配准的最终刚性变换

三、实验-点云配准

实验结果（在U3M数据集上）

结果分析：所有RANSAC 的假设评估指标都实现了显著的性能提升。



(a) *Results of using MV selection*

(b) *Results without using MV selection*

Fig. 9. Registration performance of RANSAC-based methods with and without MV-selected correspondences when using different hypothesis evaluation metrics on U3M.

MAE、MSE等：RANSAC的10个假设评价指标

三、实验-点云配准

实验结果

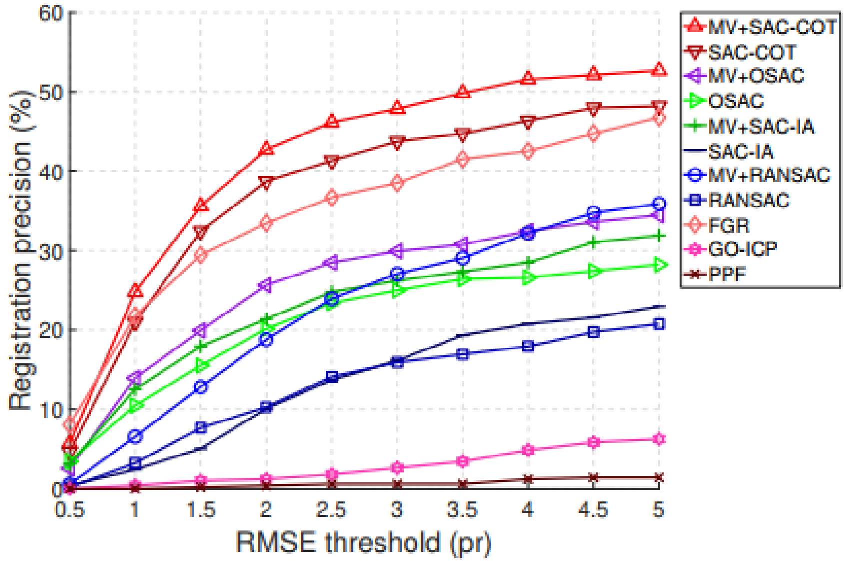


Fig. 10. Registration performance of tested point cloud registration methods on U3M.

结果分析：在测试的SAC-COT、MV+SAC-COT 、 OSAC、MV+OSAC 、SAC-IA、 MV+SAC-IA 、 RANSAC 、MV+RANSAC、 FGR、 GO-ICP和 PPF方法中（后三种与RANSAC无关），MV+SAC-COT 的性能最好。值得注意的是，MV 显著改善了所有被测试的RANSAC-fashion 估计器， 比如SAC-COT, OSAC,SAC-IA, and RANSAC

澳大

三、实验-点云配准

实验结果（3DMatch 和 3DLoMatch 数据集）

TABLE 10
Registration results (%) on 3DMatch dataset under 1k correspondences setting. The symbol '-' denotes unavailable benchmark record, **bold** and underlining indicate the best and the second best results, respectively.

Descriptor	Method	Kitchen	Home1	Home2	Hotel1	Hotel2	Hotel3	Study Room	MIT Lab	Avg.
FPFH	*i) Traditional*									
	SM [3]	-	-	-	-	-	-	-	-	55.88
	FGR [58]	44.07	51.28	37.98	46.9	38.46	48.15	26.71	46.75	41.16
	RANSAC-1K [2]	60.87	71.15	53.37	67.26	58.65	70.37	44.52	51.95	58.60
	RANSAC-10K [2]	77.08	**82.69**	67.31	82.30	73.08	81.48	61.99	66.23	73.75
	GORE [33]	78.06	80.12	66.34	86.54	75.96	83.30	66.09	59.74	74.79
	TEASER++ [34]	-	-	-	-	-	-	-	-	76.27
	LTGV [8]	80.63	**82.69**	67.79	**86.73**	75.00	**85.19**	67.47	**67.53**	76.83
	ii) Deep learned									
	3DRegNet [36]	-	-	-	-	-	-	-	-	26.31
	DGR [35]	69.17	74.36	57.69	68.58	65.38	74.07	56.16	55.84	65.19
	PointDSC [38]	76.88	81.41	65.87	83.63	70.19	79.63	61.99	62.34	73.14
	MV	**82.21**	**82.69**	**71.15**	**86.73**	**78.85**	83.33	**69.18**	**67.53**	**78.25**
FCGF	*i) Traditional*									
	SM [3]	-	-	-	-	-	-	-	-	86.57
	RANSAC-1K [2]	71.15	72.44	58.17	80.97	71.15	72.22	63.36	63.64	69.25
	RANSAC-10K [2]	74.90	73.72	59.62	81.86	70.19	70.37	62.33	66.23	70.67
	TEASER++ [34]	-	-	-	-	-	-	-	-	86.07
	LTGV [8]	95.65	91.67	**76.44**	94.69	89.42	81.48	82.53	**75.32**	88.48
	ii) Deep learned									
	3DRegNet [36]	-	-	-	-	-	-	-	-	77.76
	DGR [35]	93.68	91.03	75.00	95.13	89.42	**85.19**	81.85	67.53	87.31
	PointDSC [38]	94.86	91.03	75.48	92.48	87.5	81.48	83.22	71.43	87.55
	MV	**96.64**	**93.59**	75.96	**95.58**	**93.27**	83.33	**84.93**	74.03	**89.71**

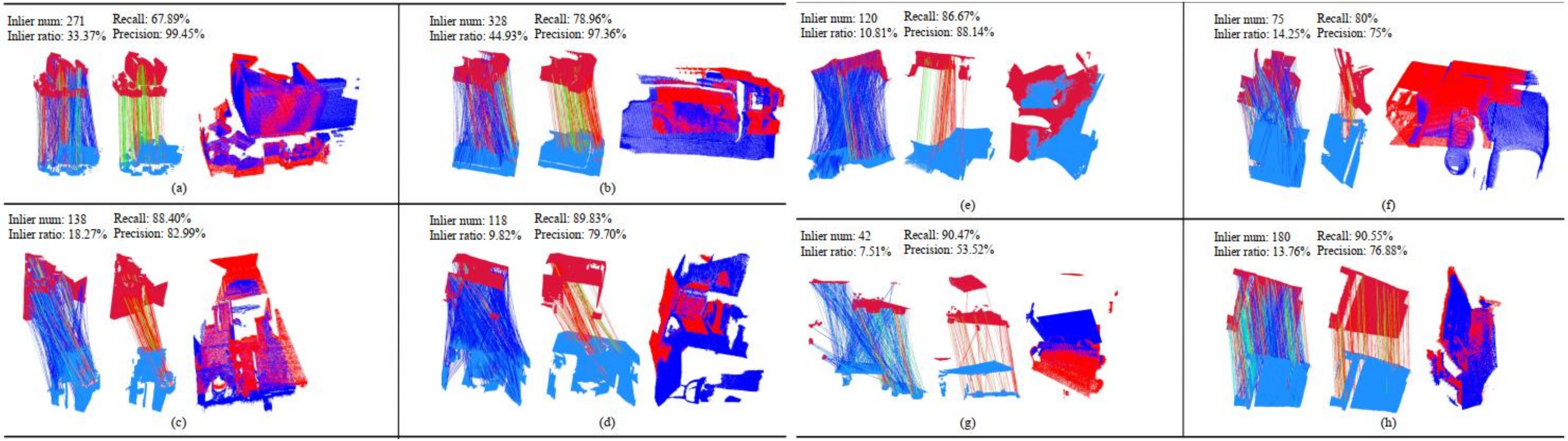
实验结果分析
1)无论使用哪种描述符， MV 在 3DMatch 和 3DLoMatch 数据集上都优于所有比较方法，表明其对室内场景点云的配准能力较强;
2)即使与深度学习的方法相比，MV 仍然在没有任何数据训练的情况下取得了更好的性能。图 11 给出了 MV 在 3DMatch 数据集上进行特征匹配和配准结果的一些可视化示例

三、实验-点云配准

实验结果（3DMatch 和 3DLoMatch 数据集）

TABLE 11
Registration results (%) on 3DLoMatch dataset under 1k correspondences setting.

Descriptor	Method	Kitchen	Home1	Home2	Hotel1	Hotel2	Hotel3	Study Room	MIT Lab	Avg.
FPFH	*i) Traditional*									
	FGR [58]	0.38	2.77	4.78	1.83	1.27	4.08	0.42	1.39	1.74
	RANSAC-1K [2]	10.29	9.34	16.96	17.43	12.03	32.65	2.92	4.17	11.40
	RANSAC-10K [2]	23.24	15.92	27.83	25.69	18.35	38.78	6.25	11.11	20.16
	GORE [33]	29.90	21.45	29.56	44.59	26.58	**40.82**	11.25	11.11	26.50
	TEASER++ [34]	-	-	-	-	-	-	-	-	28.9
	LTGV [8]	33.10	22.30	36.90	45.00	30.70	34.10	12.30	**21.70**	30.38
	ii) Deep learned									
	DGR [35]	19.05	14.53	20.00	37.61	24.05	34.69	8.33	13.89	19.93
	PointDSC [38]	24.90	16.00	27.00	27.30	14.60	29.30	5.10	15.90	19.99
	MV	**34.20**	**24.50**	**37.80**	**47.80**	**32.10**	36.60	**13.60**	20.30	**32.17**
FCGF	*i) Traditional*									
	RANSAC-1K [2]	18.67	9.69	20.00	19.27	19.62	22.45	12.08	16.67	16.68
	RANSAC-10K [2]	18.48	10.03	23.91	17.89	19.62	20.41	7.92	13.89	16.28
	TEASER++ [34]	-	-	-	-	-	-	-	-	42.11
	LTGV [8]	55.10	39.70	50.50	57.90	**39.40**	**36.60**	40.70	36.20	48.29
	ii) Deep learned									
	DGR [35]	37.90	20.07	35.22	31.19	28.48	28.57	17.50	23.61	29.42
	PointDSC [38]	51.40	34.00	**52.30**	57.40	38.00	**36.60**	33.50	**40.60**	45.14
	MV	**56.40**	**40.80**	50.00	**59.30**	38.00	**36.60**	**41.10**	39.10	**48.91**


Inlier num: 271 Recall: 67.89%
Inlier ratio: 33.37% Precision: 99.45%
(a)
Inlier num: 328 Recall: 78.96%
Inlier ratio: 44.93% Precision: 97.36%
(b)
Inlier num: 138 Recall: 88.40%
Inlier ratio: 18.27% Precision: 82.99%
(c)
Inlier num: 118 Recall: 89.83%
Inlier ratio: 9.82% Precision: 79.70%
(d)
Inlier num: 120 Recall: 86.67%
Inlier ratio: 10.81% Precision: 88.14%
(e)
Inlier num: 75 Recall: 80%
Inlier ratio: 14.25% Precision: 75%
(f)
Inlier num: 42 Recall: 90.47%
Inlier ratio: 7.51% Precision: 53.52%
(g)
Inlier num: 180 Recall: 90.55%
Inlier ratio: 13.76% Precision: 76.88%
(h)

三、实验-点云配准

实验结果（3DMatch 和 3DLoMatch 数据集）

TABLE 12
Registration results on 3DMatch dataset under 5k correspondences setting.

	FPFH			FCGF		
	RR (%)	RE (°)	TE (cm)	RR (%)	RE (°)	TE (cm)
i) Traditional						
SM [3]	55.88	2.94	8.15	86.57	2.29	7.07
FGR [58]	40.91	4.96	10.25	78.93	2.90	8.41
RANSAC-1M [2]	64.20	4.05	11.35	88.42	3.05	9.42
RANSAC-4M [2]	66.10	3.95	11.03	91.44	2.69	8.38
GC-RANSAC [10]	67.65	2.33	6.87	92.05	2.33	7.11
TEASER++ [34]	75.48	2.48	7.31	85.77	2.73	8.66
CG-SAC [62]	78.00	2.40	6.89	87.52	2.42	7.66
SC^2-PCR [9]	**83.73**	**2.18**	6.70	93.16	**2.09**	**6.51**
ii) Deep learned						
3DRegNet [36]	26.31	3.75	9.60	77.76	2.74	8.13
DGR [35]	32.84	2.45	7.53	88.85	2.28	7.02
DHVR	67.10	2.78	7.84	91.93	2.25	7.08
PointDSC [38]	72.95	**2.18**	**6.45**	91.87	2.10	6.54
MV	82.62	3.13	9.04	**93.47**	3.30	9.46

TABLE 13
Registration results on 3DLoMatch dataset under 5k correspondences setting.

	FPFH			FCGF		
	RR (%)	RE (°)	TE (cm)	RR (%)	RE (°)	TE (cm)
i) Traditional						
RANSAC-1M [2]	0.67	10.27	15.06	9.77	7.01	14.87
RANSAC-4M [2]	0.45	10.39	20.03	10.44	6.91	15.14
TEASER++ [34]	35.15	4.38	10.96	46.76	4.12	12.89
SC^2-PCR [9]	**38.57**	4.03	10.31	58.73	**3.80**	**10.44**
ii) Deep learned						
DGR [35]	19.88	5.07	13.53	43.80	4.17	10.82
PointDSC [38]	20.38	4.04	10.25	56.20	3.87	10.48
MV	36.16	5.07	12.73	**59.18**	4.99	12.92

实验结果分析：
当使用 FCGF 描述符时， MV 在3DMatch 和 3DLoMatch 数据集上都取得了最好的性能，表明 MV 非常灵活。当使用 FPFH 描述符时， MV 是第二好的描述符，略次于 SC2-PCR

三、实验-点云配准

实验结果（KITTI 数据集）

TABLE 14
Registration results on KITTI dataset.

	FPFH			FCGF		
	RR (%)	RE (°)	TE (cm)	RR (%)	RE (°)	TE (cm)
i) Traditional						
FGR [58]	5.23	0.86	43.84	89.54	0.46	25.72
TEASER++ [34]	91.17	1.03	17.98	94.96	0.38	**13.69**
RANSAC [2]	74.41	1.55	30.20	80.36	0.73	26.79
CG-SAC [62]	74.23	0.73	14.02	83.24	0.56	22.96
SC^2-PCR [9]	**99.28**	<u>0.39</u>	<u>8.68</u>	<u>97.84</u>	**0.33**	20.58
ii) Deep learned						
DGR [35]	77.12	1.64	33.10	96.90	<u>0.34</u>	21.70
PointDSC [38]	<u>98.92</u>	**0.38**	**8.35**	<u>97.84</u>	**0.33**	<u>20.32</u>
MV	<u>98.92</u>	0.56	10.82	**98.20**	0.35	20.41

实验结果分析:

如表所示，在配准结果方面，MV 分别在 FPFH和 FCGF 描述符设置下表现出最佳和第二好的结果(比最佳方法落后 0.36%)。另外，户外点云具有明显的稀疏性和非均匀分布性。在物体、室内场景和室外场景上的配准实验一致验证了 MV 选择的对应关系可以有效提升不同应用场景下的点云配准性能

三、实验-3D物体识别

实验设置:

数据集采用Queen 和U3OR

Evaluation metrics. We follow [30] and compute the ε_{res} and r_{ov} metrics. Assume that the transformation estimation matrix $\mathbf{M}_i = (\mathbf{R}_i, \mathbf{t}_i)$ is obtained in the i-th iteration of RANSAC, and the model point cloud $\mathbf{P}^s$ is transformed to obtain the point cloud $\mathbf{P}^{trans}$. Let $\mathbf{p}_i^t$ be the point of

$M_i=(R_i,t_i)$: 变换估计矩阵

从RANSAC的第i次迭代中获得.

$\mathbf{P}^s$, $\mathbf{p}_i^{trans}$ be the point of $\mathbf{P}^{trans}$ and the nearest neighbor to the point in the point cloud $\mathbf{P}^t$. If the distance $d(\mathbf{p}_i^t, \mathbf{p}_i^{trans}) = ||\mathbf{p}_i^t - \mathbf{p}_i^{trans}||$ is less than the threshold d_{rec}, which is set to 2 pr, $\mathbf{p}_i^{trans}$ will be classified into the point set $\mathbf{P}_{overlap}^{trans}$. The point cloud residual error ε_{res} and the point cloud overlap rate r_{ov} are defined as:

残差

$$\varepsilon_{res} = \frac{\sum\limits_{\mathbf{p}_i^{trans} \in \mathbf{P}_{overlap}^{trans}} d(\mathbf{p}_i^{trans}, \mathbf{p}_i^t)}{|\mathbf{P}_{overlap}^{trans}|}, \tag{17}$$

点云重叠率

$$r_{ov} = \frac{|\mathbf{P}_{overlap}^{trans}|}{|\mathbf{P}^t|}. \tag{18}$$

When the residual is less than the threshold d_{res} and the overlap rate is greater than the threshold $t_{overlap}$, the transformation estimation matrix $\mathbf{M}_i$ is considered to satisfy the condition and the iteration is stopped. Using this transformation matrix the model $\mathbf{P}^s$ can be identified from the scene $\mathbf{P}^t$. Following [68], we let d_{res} be 0.75 pr and $t_{overlap}$ be 0.04, or let d_{res} be 1.5 pr and $t_{overlap}$ be 0.2.

三、实验-3D物体识别

实验比较结果

对于 3D 物体识别，基于描述符的方法是主要解决方案，在不同的 RANSAC 迭代下，我们测试了有和没有 MV 选择对应关系作为输入后的几个代表性描述符的识别性能，Queen 和 U3OR 数据集识别结果分别如表 17 和表18 所示

TABLE 17
Object recognition results on Queen dataset (%).

Method	Angle	Bigbird	Gnome	Kid	Zoe	Average
EM [64]	-	-	-	-	-	81.90
VD-LSD(SQ) [65]	89.70	**100.00**	70.50	84.60	71.80	83.80
(500 iterations)						
Spin image [66]	2.63	51.16	18.42	36.59	9.52	24.14
RoPS [30]	34.21	34.88	44.74	17.07	11.90	28.57
RCS [67]	63.16	76.74	86.84	82.93	57.14	74.38
VOID [68]	84.21	90.70	92.11	95.12	61.94	86.21
Spin image+MV	21.05	67.44	47.37	78.05	47.62	53.69 (29.55↑)
RoPS+MV	60.53	72.09	68.42	58.54	73.81	67.49 (38.92↑)
RCS+MV	92.11	88.37	**94.74**	95.12	92.86	94.58 (20.20↑)
VOID+MV	**94.74**	83.72	**94.74**	**100.00**	**95.24**	**97.04** (10.83↑)
(1000 iterations)						
Spin image [66]	2.63	48.84	18.42	43.90	19.05	27.09
RoPS [30]	39.47	39.53	50.00	24.39	11.90	33.00
RCS [67]	68.42	79.07	86.84	87.80	69.05	79.31
VOID [68]	84.21	83.72	**94.74**	90.24	71.43	88.18
Spin image+MV	23.68	67.44	50.00	85.37	42.86	56.65 (29.56↑)
RoPS+MV	63.16	74.42	68.42	63.41	69.05	68.47 (35.47↑)
RCS+MV	**92.11**	86.05	**94.74**	95.12	**95.24**	95.07 (15.76↑)
VOID+MV	**92.11**	74.42	**94.74**	**100.00**	**95.24**	**97.04** (8.86↑)
(2000 iterations)						
Spin image [66]	5.26	51.16	23.68	46.34	26.19	31.03
RoPS [30]	50.00	51.16	68.42	41.46	26.19	47.52
RCS [67]	71.05	83.72	86.84	95.12	69.05	82.76
VOID [68]	**94.74**	95.24	**100.00**	97.56	78.57	93.07
Spin image+MV	23.68	65.12	50.00	80.49	47.62	56.16 (25.13↑)
RoPS+MV	63.16	72.09	71.05	70.73	73.81	70.94 (23.42↑)
RCS+MV	92.11	79.07	94.74	**100.00**	92.86	95.57 (12.81↑)
VOID+MV	**94.74**	95.24	**100.00**	**100.00**	**97.62**	**97.52** (4.45↑)

TABLE 18
Object recognition results on U3OR dataset (%).

Method	T-rex	Chef	Chicken	parasaurolophus	Average
RoPS [30]	-	-	-	-	98.90
TriLCI [69]	97.78	**100.00**	**100.00**	62.22	98.90
(500 iterations)					
Spin image [66]	35.56	92.00	62.50	40.00	58.51
RCS [67]	88.89	96.00	97.92	84.44	92.02
SHOT [42]	35.56	82.00	52.08	42.22	53.72
FPFH [55]	46.67	**100.00**	68.75	46.67	66.49
VOID [68]	97.78	**100.00**	**100.00**	95.56	98.40
Spin image+MV	68.89	**100.00**	75.00	57.78	76.06 (17.55↑)
RCS+MV	93.33	**100.00**	95.83	93.33	95.74 (3.72↑)
SHOT+MV	40.00	98.00	66.67	44.44	63.30 (9.58↑)
FPFH+MV	64.44	**100.00**	72.92	60.00	75.00 (8.51↑)
VOID+MV	**100.00**	**100.00**	**100.00**	**97.78**	**99.47** (1.07↑)
(1000 iterations)					
Spin image [66]	46.67	98.00	64.58	46.67	64.89
RCS [67]	93.33	**100.00**	95.83	91.11	95.21
SHOT [42]	44.44	84.00	62.50	46.67	60.11
FPFH [55]	44.44	**100.00**	66.67	51.11	66.49
VOID [68]	**97.78**	**100.00**	**100.00**	**97.78**	**98.94**
Spin image+MV	68.89	**100.00**	72.92	57.78	75.53 (10.41↑)
RCS+MV	95.56	**100.00**	97.92	93.33	96.81 (1.60↑)
SHOT+MV	44.44	**100.00**	70.83	44.44	65.96 (5.85↑)
FPFH+MV	60.00	**100.00**	72.92	60.00	73.94 (7.45↑)
VOID+MV	**97.78**	**100.00**	**100.00**	**97.78**	**98.94**
(2000 iterations)					
Spin image [66]	51.11	**100.00**	58.33	97.78	68.62
RCS [67]	**100.00**	**100.00**	93.75	97.78	97.87
SHOT [42]	90.00	48.89	37.78	58.33	59.57
FPFH [55]	**100.00**	53.33	48.89	68.75	68.62
VOID [68]	**100.00**	**100.00**	**100.00**	97.78	**99.47**
Spin image+MV	71.11	**100.00**	77.08	57.78	77.13 (8.51↑)
RCS+MV	**100.00**	95.56	97.78	**100.00**	98.40 (0.53↑)
SHOT+MV	**100.00**	51.11	48.89	77.08	70.21 (10.64↑)
FPFH+MV	**100.00**	55.56	62.22	70.83	72.87 (4.25↑)
VOID+MV	**100.00**	97.78	**100.00**	**100.00**	**99.47**

三、实验-3D物体识别

结果分析:

MV 显著提高了所有测试描述子下所有数据集上的 3D 物体识别性能。RANSAC 迭代次数越少，这种现象就越明显。 在 Queen 数据集上， MV 对 RoPS 的改进最为显著，在 500 次 RANSAC 迭代下提高了 38.92%;对于500 次迭代下表现最好的 VOID 描述子，也有 10.83%的提升。在 U3OR 数据集上， MV 在 500 次迭代的 spin 图像上提高了 17.55%，在 2000 次迭代的 SHOT 图像上提高了 10.64%。实验结果表明， MV 具有良好的泛化能力，能够适应不同的描述子。

文章结论

本文提出了一种新颖的相互投票的 3D 对应关系排序方法。它通过细化投票人和选民，可靠地为每个对应关系分配一个投票方式。在各种具有不同挑战和模态的数据集上进行的特征匹配、3D 点云配准和 3D 物体识别实验验证了两个结论:1)在不同挑战设置下， MV 对重离群点具有鲁棒性;2) MV 可以显著提升现有方式下的 3D 点云配准和 3D 物体识别性能

Thank You!

Avenida da Universidade, Taipa, Macau, China
Tel : (853) 8822 8833 Fax : (853) 8822 8822
Email : info@um.edu.mo Website : www.um.edu.mo

澳大