

基于偏好统计数据表征的鲁棒几何模型拟合方法

郭翰林¹⁾ 肖国宝²⁾ 严 严¹⁾ 林舒源¹⁾ SUTER David³⁾ 王菡子¹⁾

¹⁾ (厦门大学信息学院福建省智慧城市感知与计算重点实验室 福建 厦门 361005)

²⁾ (闽江学院计算机与控制工程学院福建省信息处理与智能控制实验室 福州 350108)

³⁾ (埃迪斯科文大学科学学院 珀斯 6027 澳大利亚)

摘 要 鲁棒几何模型拟合是计算机视觉的一个基础性研究问题,广泛应用于各类计算机视觉任务,如单应性矩阵或基础矩阵估计、图像匹配、医学图像分析等。它的主要任务是:在包含噪声点和离群点的数据集中估计模型实例的参数和个数。针对该任务,本文提出一种基于新型数据表征(称之为偏好统计数据表征)的模型拟合方法。该新型数据表征算法将残差值进行排序然后映射到不同的区间以构建残差直方图数据表征,来描述数据分布的特征。该算法充分利用传统模型拟合方法中偏好分析和一致性统计分析的优点,更加有效地对数据分布特征进行描述,从而有效地提高数据表征的准确性和鲁棒性。为了进一步有效地利用该数据表征中的统计信息(内点和离群点显示出显著的信息熵值差异),本文利用直方图中不同区间段所映射的残差值的出现频次,以分析直方图的特性。并且采用一种简单的自适应熵阈值算法,来区分内点与离群点以进行离群点检测。最后,为了能够更好地处理分布在交叉模型实例附近的数据点,本文引入一种基于相似矩阵学习的图聚类技术,提出一个有效的模型实例估计算法。该算法先是用聚类技术以实现数据的分割,进而估计模型实例的参数。同时,该模型实例估计算法结合拉普拉斯矩阵特征值的分析以及最小子集数目的约束,使其能够自适应地估计模型实例的个数。在合成数据集和真实数据集上的实验结果表明,本文提出的偏好统计数据表征算法能够有效提高模型拟合方法的准确性和鲁棒性。同时,与当前一些流行的模型拟合方法相比,本文基于偏好统计数据表征的鲁棒几何模型拟合方法取得了更好的拟合精度,并且在速度方面要比大部分拟合方法更加高效。

关键词 鲁棒模型拟合;多结构数据;偏好统计数据表征;离群点检测;模型参数估计
中图法分类号 TP391 DOI号 10.11897/SP.J.1016.2020.01199

A Preference-Statistic-Based Data Representation for Robust Geometric Model Fitting

GUO Han-Lin¹⁾ XIAO Guo-Bao²⁾ YAN Yan¹⁾ LIN Shu-Yuan¹⁾ SUTER David³⁾ WANG Han-Zi¹⁾

¹⁾ (Fujian Key Laboratory of Sensing and Computing for Smart City, School of Informatics, Xiamen University, Xiamen, Fujian 361005)

²⁾ (Fujian Key Laboratory of Information Processing and Intelligent Control, School of Computer and Control Engineering, Minjiang University, Fuzhou 350108)

³⁾ (School of Science, Edith Cowan University, Perth, 6027 Australia)

Abstract Robust geometric model fitting is a fundamental research problem in computer vision. It has been widely used in a variety of applications such as homography/fundamental matrix estimation, motion segmentation, image matching and medical image analysis. Given that data usually contain noise and outliers caused by sensing or preprocessing errors, the main task of robust geometric model fitting is to estimate the number and the parameters of model instances in data. Although a number of robust geometric model fitting methods have been proposed during

收稿日期:2019-06-19;在线发布日期:2019-12-03. 本课题得到国家自然科学基金联合基金(U1605252)、国家自然科学基金(61702431, 61571379, 61872307)资助。郭翰林,博士研究生,主要研究方向为计算机视觉、鲁棒模型拟合、健康医疗大数据。E-mail: hanlinguo@stu.xmu.edu.cn。肖国宝,博士,教授,主要研究领域为计算机视觉、模式识别。严 严,博士,副教授,主要研究方向为计算机视觉、机器学习。林舒源,博士研究生,主要研究方向为计算机视觉、鲁棒模型拟合。SUTER David,教授,博士生导师,主要研究领域为计算机视觉、机器学习。王菡子(通信作者),教授,博士生导师,中国计算机学会(CCF)会员,主要研究领域为计算机视觉、模式识别。E-mail: hanzi.wang@xmu.edu.cn。

the past few decades, it is still a challenging task since a scene typically contains multiple geometric structures, especially, when the observed data are largely contaminated with noise and outliers. Traditional robust model fitting methods can be roughly classified into consensus statistic based methods and preference analysis based methods. The consensus statistic based methods aim to search the maximum consensus sets that yield the underlying model instances in data, and then segment the data points into inliers and outliers by a certain threshold. Preference analysis based methods describe the relationship of data points based on the preference information which is usually used to define a similarity measurement, and then the data points are grouped into inliers and outliers based on a similarity-based clustering algorithm. However, the preference analysis potentially keeps redundant information (small differences in the residuals are probably unimportant), whilst the consensus statistic throws away too much potentially useful information. These methods might lead to poor fitting results. In this paper, we aim to use a compromise statistic that keeps sufficient and compact information for model fitting and segmentation purposes. We present an effective robust geometric fitting method based on a novel data representation (called preference-statistic-based data representation) algorithm to deal with multiple-structure data contaminated with noise and outliers. Specifically, the proposed data representation algorithm analyzes the residuals (of each data point with respect to model hypotheses) in a histogram for the data representation, taking advantage of both preference analysis and consensus statistic. Thus, the accuracy and robustness for the proposed data representation are boosted. To utilize the statistical information embedded in the proposed data representation, a simple entropy threshold algorithm is used for the adaptive outlier detection, based on using the frequency count of bins of histogram. In addition, we present an effective model selection algorithm (which is able to effectively deal with data points near the intersection of model instances), based on similarity matrix learning for graph clustering. Specifically, we analyze the eigenvalues of the Laplacian matrix and use a constraint of the number of a minimal subset to automatically estimate the number of model instances. We evaluate the performance of the proposed method on both synthetic datasets and real image pairs. We firstly compare the proposed data representation algorithm (preference-statistic-based data representation) with the consensus statistic based data representation algorithm and the preference analysis based data representation algorithm, in the framework of T-Linkage method. Experimental results demonstrate that the proposed data representation algorithm is effective and helpful to achieve better fitting results for different fitting tasks. Then, we compare the proposed fitting method with several state-of-the-art fitting methods to show the promising fitting results of the proposed method.

Keywords robust model fitting; multiple-structure data; preference-statistic-based data representation; outlier detection; model parameter estimation

1 引 言

作为人工智能的一个重要分支,计算机视觉的一项重要任务是利用计算机从图像和视频中获取人类所需要的信息.鲁棒几何模型拟合是计算机视觉的一个基础性研究问题.它的主要任务是:在包含噪声点和离群点的数据集中估计模型实例的参数和个数,以获取有用的信息.比如,在一张图片里分割出

圆形,以描述几何目标的位置、形状和大小;在两个视图中拟合基础矩阵以估计物体的运动关系;在视频的图像序列里构建线性子空间模型以分割运动目标等.因此,往往需要对图像和视频数据进行分析,发现蕴含在数据中的有用信息.在多数情况下,蕴含在这些数据中的有用信息可以通过参数模型来表示.这些参数模型可以通过模型拟合方法从输入的数据中估计得到,例如,直线模型、基础矩阵模型等.由此可见,模型拟合是计算机视觉领域中一个重要

的基础性研究问题。

在过去的几十年里,大量的模型拟合方法(例如,文献[1-3])已经被提出来,并被广泛地应用在单应性矩阵或基础矩阵估计^[1-2]、运动分割^[3-4]、医学图像分析^[5-6]等各类计算机视觉任务中。然而,实际数据可能包含噪声、测量误差、错误提取的特征点等问题;以及这些数据通常包含多个模型实例,而不同模型实例之间往往会产生相互干扰等问题。其中,一个模型实例的内点(inlier),对于其它模型实例来说是伪离群点(pseudo-outliers)。上述问题增加了模型拟合的难度。因此,实现高效且鲁棒地在数据中拟合模型依然是一个十分具有挑战性的难题。

鲁棒模型拟合的一个经典方法是 RANSAC (RANDOM SAMPLE CONSENSUS)^[7]。该方法通过指定一个内点噪声尺度来统计分析包含最多内点数的一个模型实例。由于 RANSAC 方法计算简单且易于实现,因而被广泛地应用在计算机视觉和相关的领域。然而该方法只能处理单一结构的数据。这无法满足实际应用的需求。为处理多结构数据,有许多优秀的拟合方法针对 RANSAC 进行改进,比如,AKSWH^[8],gpbM^[9],RansaCov^[10]。这些方法都是属于基于一致性统计分析的拟合方法。具体地说,基于一致性统计分析的拟合方法通过分析模型假设的最大化一致集,以选择有意义的模型假设作为模型实例。然后通过一个内点噪声尺度将数据点分割为内点和离群点。然而,这一类方法的拟合性能对于内点噪声尺度十分敏感,无法很好地应用于实际问题。

近年来,许多研究者从不同的角度来解决模型拟合这一问题。不同于采用基于一致性统计分析的方法(例如 RANSAC^[7]),另一个类别的拟合方法采用基于偏好分析的方法(例如 RHA^[11],J-Linkage^[12],T-Linkage^[13],RPA^[14]和 KF^[15])来解决模型拟合问题。其中,偏好分析中用到的偏好信息是一个数据点到一系列模型假设之间的残差值的分布信息。具体地说,基于偏好分析的方法通常采用偏好信息来描述数据点之间的关系,以构造数据的相似矩阵。然后采用基于相似矩阵的聚类算法将数据点分割为内点和离群点。此类方法在内点里估计模型实例的参数,有利于缓解基于一致性统计分析的方法对内点噪声尺度的敏感性问题。基于偏好分析的方法,在描述数据点分布特征时发挥了很大的优势。例如,J-Linkage^[12]提出基于二值偏好的方式来描述数据点分布的特征以表征数据,其中二值偏好是通过一系列候选模型的一致性统计分析获取的。为了获取更丰富的偏好

信息,T-Linkage^[13]和 RPA^[14]将二值偏好扩展为连续偏好来表征数据。并且这两种方法分别利用了一个连续的指数核函数和一个连续的鲁棒柯西函数来进行偏好分析。然而,基于连续偏好分析的拟合方法,在进行数据表征时保留了大量的残差信息(包括微小变化的残差值),使得特征表示存在冗余;同时也忽略了残差的一致性信息(一致性信息在决定数据点属于哪个模型实例时具有重要的作用)。因此,单独使用二值偏好分析(可由一致性统计分析获取)或者连续偏好分析进行数据表征时存在一定的局限性,可能导致模型拟合方法的拟合性能变差。

针对以上问题,本文提出了一种基于新型数据表征的模型拟合方法,以有效融合二值偏好分析(本质为采用一致性统计分析的思想)和连续偏好分析的优点,即同时考虑到一些重要的残差值以及微小变化的残差值。所提出的数据表征算法将残差值(指数据点与生成的模型假设之间残差的绝对值)映射到直方图的不同区间段来表征数据。基于此,为了有效利用数据表征中的统计信息(内点和离群点显示出显著的信息熵值差异),本文采用一种简单的自适应熵阈值算法,来区分内点和离群点以进行离群点检测。最后,本文引入基于相似矩阵学习的图聚类技术^[16]来构造更精确的相似矩阵以实现数据的分割,进而估计模型实例;并且本文提出基于拉普拉斯矩阵特征值的分析和利用最小子集数目的约束,自适应地估计模型实例的个数。

综上所述,本文主要有以下三个贡献点:

(1) 提出一种新型的数据表征算法。该算法将残差值进行排序然后映射到不同的区间以构建残差直方图数据表征。该数据表征算法融合了偏好分析和一致性统计分析的优点,以获取更加丰富且有效的残差信息,进而描述更准确的数据分布特征,提高模型拟合的精度。

(2) 提出一种基于熵阈值的离群点检测策略。该策略充分利用所提出的数据表征中内点和离群点之间显著的信息熵值差异,有效地去除离群点。

(3) 提出一种基于相似矩阵学习的图聚类技术的模型实例估计算法。该算法能够有效处理分布在交叉模型实例附近的数据点。更重要的是,该算法根据特征值的分析和最小子集数目的约束自适应地估计模型实例的个数。

本文第 2 节回顾一些相关的模型拟合方法;第 3 节详细描述本文所提出的偏好统计数据表征算法,以及在此基础上的离群点检测策略和模型实例

估计算法;第4节给出本文所提出的方法在不同数据集上的实验结果以及同其它方法的对比讨论;第5节总结本文工作以及给出了对未来工作的展望。

2 相关工作

传统的鲁棒几何模型拟合方法主要可以分为两大类:基于一致性统计分析的拟合方法和基于偏好分析的拟合方法。此外,基于图聚类的技术在模型拟合中也得到了广泛地应用。

一致性统计分析。一个模型假设的一致集是指该模型假设在指定的内点噪声尺度范围内所包含的数据点集合。一致性统计分析的思想最初起源于RANSAC方法^[7],并被广泛地应用于RANSAC的变种方法(详细叙述请参看综述文献^[17])。文献^[18]对所有的数据点依次执行RANSAC来处理多模型实例问题。然而,RANSAC执行过程中不精确的模型实例检测结果,会严重影响最终的模型拟合效果。AKSWH^[8]通过分析模型假设的一致集,对有意义的模型假设进行聚类,然后在聚类结果的每个类里面选出最有代表性的模型假设作为所拟合的模型实例。RansaCov^[10]采用贪婪式搜索的方法对多个具有最大化一致集的模型假设进行聚类,来搜索对应的模型实例所包含的内点。另外,一致性统计分析也是基于能量优化的几何模型拟合方法的基础(例如,PEARL^[19])。虽然一致性统计分析的方法能够根据内点噪声尺度直接二分内点和离群点,有效地决定数据点所属的模型实例。但是这类方法忽略了一些可能重要的残差信息,容易导致最终的拟合性能下降。

偏好分析。一个数据点的偏好集是指将该数据点归为内点的模型假设集。偏好分析的思想起源于RHA^[11]。该方法提出分析每个数据点对应于所有生成的模型假设之间的偏好信息,并且通过模式搜索估计模型实例。J-Linage^[12]提出基于二值偏好的方式来表征数据,并用杰卡德距离测量两个数据点之间的距离,然后采用层次聚类算法将数据分割为内点和离群点。为了获取更丰富的偏好信息,T-Linkage^[13]和RPA^[14]将二值偏好扩展为连续偏好来表征数据。KF^[15]提出按照数据点到生成的模型假设之间的残差的升序排列来表征数据,然后采用谱聚类分割数据。为了最大化内点之间共有的偏好信息,QP-MF^[20]通过偏好信息来排序对应的模型假设以表征数据,然后采用二次规划的方法进行模型拟合。RCMSA^[21]通过获取的偏好信息构建一个权

重邻接图,同时把模型拟合建模为图割问题,并且在模拟退火框架下有效地解决了该问题。MSHF^[22]将模型拟合建模为超图构建问题以及结合所获取的偏好信息进行模式搜索。尽管基于偏好分析的方法相比基于一致性统计分析的方法考虑了更多的残差信息,但是那些微小变化的残差值可能是不重要且冗余的,甚至影响数据分布特征描述的准确性。这在一定程度上也会损失模型拟合方法最终的拟合精度。

上述分析可以看出,基于一致性统计分析方法(可用于二值偏好分析)和基于偏好分析方法均存在各自的优点和缺点。在融合各自方法的优势下,本文提出一种新的分析方法,即偏好统计分析方法。本文方法先是对不同的残差进行排序以保留残差信息的偏好性,然后将排序后的残差值映射到直方图的多个区间以保持残差信息的一致性。本文的方法充分利用偏好分析和一致性统计分析的优点,从而能够更好地对数据分布特征进行描述,获得紧凑和具有鲁棒性的数据表征。

此外,近些年来基于图聚类的技术,来实现鲁棒模型拟合的研究工作^[3,14-15]成为研究热点。这些研究工作通常是通过构造数据点的相似矩阵以描述数据点之间的关系来进行图聚类。其中图顶点之间的边的权重值对应于构造的数据点的相似矩阵。例如,文献^[3,15]通过对数据点对应的残差值排序来构造相似矩阵,以进行聚类并估计模型实例。文献^[14]通过鲁棒PCA算法取代相似矩阵特征分解,然后通过对称的NMF算法进行聚类,以此取得更好的拟合性能。然而,这些方法在计算相似矩阵和图聚类时是两个独立的步骤。因此,不精确的相似矩阵将会导致后续图聚类的性能变差。最近,文献^[16,23]提出相似矩阵学习并同时聚类的方法。其中,学习的相似性矩阵能够获取更精确的相似值以进行聚类,有利于处理分布在交叉模型实例附近的数据点。

3 算法描述

本文提出一种新型的数据表征算法,即将残差值映射到直方图的不同区间来进行数据表征。该数据表征算法能够有效地获取残差的一致性信息和偏好信息。在此基础上,利用数据表征中的统计信息有效地对内点和离群点进行区分。最后,在保留下来的内点中,基于相似矩阵学习的图聚类技术,对数据点进行聚类以自适应地估计模型实例的个数和参数。

3.1 一种新型的数据表征算法

J-Linkage^[12]和T-Linkage^[13]利用数据点到一

系列模型假设之间的残差值组成的向量来表征数据。精确的数据表征有利于刻画数据点的本质特征，描述数据点之间的相似关系，进而有效地区分内点和离群点。然而，基于二值偏好分析的数据表征算法（如 J-Linkage^[12] 通过一系列候选模型的一致性统计分析获取数据点的二值偏好）忽略了许多可能重要的残差信息；而基于连续偏好分析的数据表征算法（如 T-Linkage^[13]）保留了大量冗余的残差信息，而且微小变化的残差值实际上可能是不重要的，甚至影响数据分布特征描述的准确性。因此，需要引入更为紧凑的数据表征算法，以得到更精确的数据分布特征描述。本文方法旨在综合利用偏好分析和一致性统计分析的优点，来保留充分且有效的残差信息以进行模型拟合。

给定输入 N 个数据点 $x = \{x_i\}_{i=1, \dots, N}$ ，并从输入的数据点中采样 M 个最小子集生成模型假设 $h = \{h_j\}_{j=1, \dots, M}$ ，其中一个最小子集是指估计一个模型所需要的最小数据点集，比如生成一条直线至少需要 2 个点，估计一个基础矩阵需要 7 个点或者 8 个点。针对每个数据点 x_i ，本文计算该点到第 j 个模型假设 h_j 的残差绝对值 $r(x_i, h_j)$ 。

本文所提出的数据表征算法，其核心思想描述如下：J-Linkage^[12] 提出的二值偏好数据表征的本质是采用一致性统计分析的思想。该方法统计所有的残差值中小于固定阈值的残差（其中阈值是由指定的或者估计的内点噪声尺度决定）。这种做法可看作一种非常简单的残差直方图数据表征，它将所有小于阈值的残差值映射到一个直方图区间并将大于阈值的残差值映射到另一个直方图区间中。本文提出的数据表征算法是将所有小于固定阈值的残差值映射到残差直方图的 K 个区间（ K 是一个相对小的整数且大于 1）以进行残差直方图数据表征。这是一种更加有效的数据表征算法。它比 J-Linkage^[12] 基于二值偏好分析的方法描述了更丰富的残差信息。从某种意义上说，本文所提出的表征算法是介于 J-Linkage^[12] 的二值偏好分析和 T-Linkage^[13] 的连续偏好分析之间的表征算法。如果将小于阈值的残差进行排序并且将 K 增长到极限值的情况下（即 K 为残差小于阈值时所对应的模型假设的数目），即每个直方图区间最多包含一个残差值。那么此时本文的表征方法就是 T-Linkage^[13] 中的连续偏好分析的方法。因此，本文将所提出的表征算法称为偏好统计数据表征。并且通过该数据表征算法，可以对一系列离散分布的整数集（对应于本文所构造的残差直方图区间所映射的残差值）进行统计分析。

更具体的说，针对每个数据点 x_i ，对该点关于每个模型假设 h_j 的残差绝对值 $r(x_i, h_j)$ 进行排序，并且将这些残差值映射到直方图的 K 个区间中。在此，为了统计这些残差值，本文定义一个偏好统计矩阵 $\Theta \in \mathbb{R}^{N \times M}$ ，其中 $\Theta(i, j)$ 是关于第 i 个点和第 j 个模型假设对应的残差值在直方图区间的映射值。

3.2 离群点检测策略

结合上文的研究分析，在随机采样最小子集生成模型假设的过程中，一部分模型假设会聚集在真实的模型实例附近^[8]。为寻找出聚集在真实模型实例附近的这些有意义的模型假设，本文引入非参数核密度估计算法^[24]，以计算每个模型假设的权重值。越高的权重值对应于越有意义的模型假设。其中，第 j 个模型假设 h_j 的权重，计算公式如下^[8]：

$$\omega_j = \frac{1}{N} \sum_{i=1}^N \frac{\mathbf{KN}(r(x_i, h_j)/\hat{h}_j)}{\bar{s}_j \hat{h}_j} \quad (1)$$

其中， \bar{s}_j 代表第 j 个模型假设 h_j 所估计的内点噪声尺度（本文通过 IKOSE^[8] 求得）； $r(x_i, h_j)$ 代表第 i 个数据点 x_i 对应于第 j 个模型假设 h_j 的残差绝对值； $\mathbf{KN}(\cdot)$ 代表经典的 Epanechnikov 核函数，其定义如下：

$$\mathbf{KN}(\psi) = \begin{cases} 0.75(1 - \|\psi\|^2), & \|\psi\| \leq 1 \\ 0, & \text{其它} \end{cases} \quad (2)$$

\hat{h}_j 表示第 j 个模型假设的带宽，计算公式如下^[24]：

$$\hat{h}_j = \left[\frac{243 \int_{-1}^1 \mathbf{KN}(\psi)^2 d\psi}{35N \int_{-1}^1 \psi^2 \mathbf{KN}(\psi) d\psi} \right]^{0.2} \bar{s}_j \quad (3)$$

根据式(1)，如果一个模型假设包含更多的内点且对应较小的残差值，则该模型假设将被赋予更高的权重值；反之，则被赋予较小的权重值。

接下来，本文采用文献[25]提出的熵阈值算法以获得一个自适应的阈值来区分这些模型假设的权重值 ω_j 。高权重值（对应于大于阈值的权重值）的模型假设即为获得的 \hat{M} 个有意义的模型假设 $\hat{h} = \{\hat{h}_j\}_{j=1, \dots, \hat{M}}$ 。从而得到一个简化的偏好统计矩阵 $\tilde{\Theta} \in \mathbb{R}^{N \times \hat{M}}$ 。该矩阵只包含了 \hat{M} 个有意义的模型假设 \hat{h} 。相比采用计算数据点之间的距离进行离群点检测的策略（例如 KF^[15]，T-linkage^[13]），本文提出基于偏好统计矩阵 $\tilde{\Theta} \in \mathbb{R}^{N \times \hat{M}}$ 的离群点检测策略（ $\tilde{\Theta}(i, :)$ 表示矩阵的第 i 行，对应于数据点 x_i 的一个表示向量）。

受文献[26]工作的启发，该工作结合信息论的优点分析每个类别属性的重要性，从而进行每个类别间的相似性度量。本文提出分析矩阵 $\tilde{\Theta}$ 中每一行

的信息内容,然后根据每行信息内容的变化来区分出内点和离群点.因为矩阵 $\tilde{\Theta}$ 中每一行 $\tilde{\Theta}(i,:)$ 的信息内容和该行的数值(对应于直方图区间所映射的残差值)变化情况密切相关.

为更清晰地描述本文所提出的离群点检测策略,本文给出了一个偏好统计矩阵 $\tilde{\Theta}$ (该矩阵基于所提出的数据表征算法所构造)的可视化例子,如图1所示.图1(a)为两张不同视角的图像对,其中显示的数据点包括两个模型实例的数据点(分别用三角形点和五角星点表示不同结构的数据点,并用直线连接对应的内点匹配对)和大量离群点(用交叉点表示).图1(b)为可视化的偏好统计矩阵.矩阵的行(代表数据点)和矩阵的列(代表模型假设)通过聚类进行了排序.图1(b)中颜色条上的不同数值对应于残差值在直方图的不同区间的映射值.从图1(b)可以看出,对于一个内点而言,无论它是属于哪一个模型实例,在相应的偏好统计矩阵行里面它都能显示出显著的数值变化(对应于直方图的不同区间所映射的残差值).这是因为它属于几个有意义的模型假设.相反,离群点在相应的行里面通常显示出微小的数值变化.因此,如果矩阵 $\tilde{\Theta}$ 的行对应于内点,那么该行将产生一个高的信息熵,反之亦然.因此,偏好统计矩阵每一行 $\tilde{\Theta}(i,:)$ 的信息熵能够以该行数值(对应于直方图区间所映射的残差值)变化情况的概率分布进行测量,其对应的信息熵计算公式如下:

$$I_{\tilde{\Theta}(i,:)} = - \sum_{t=1}^r p(a_{it}) \log p(a_{it}) \quad (4)$$

其中, $r(r \leq K)$ 是偏好矩阵 $\tilde{\Theta}$ 每行中所出现的数值(对应于残差直方图所映射的残差值)的数目; a_{it} 表示在 $\tilde{\Theta}$ 的第*i*行中的第*t*个数值出现的次数;以及 $p(a_{it})$ 是偏好矩阵 $\tilde{\Theta}$ 的第*i*行中第*t*个数值在 $\tilde{\Theta}$ 所有数值中出现的概率.通过归一化数值 a_{it} 来计算对应的概率值 $p(a_{it})$,计算公式如下:

$$p(a_{it}) = \frac{a_{it}}{\sum_{t=1}^r a_{it}} \quad (5)$$

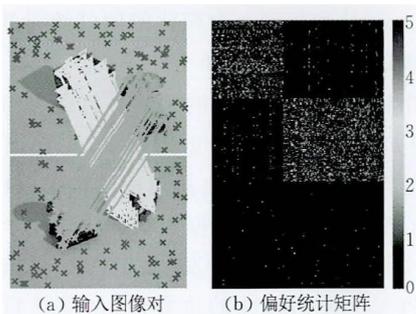


图1 在基础矩阵估计中,一个偏好统计矩阵可视化例子

根据式(4),可在偏好统计矩阵的每一行 $\tilde{\Theta}(i,:)$ 中获取不同的信息熵.因此,内点和离群点在偏好统计矩阵 $\tilde{\Theta}$ 中所对应的信息熵将显示出显著的区别.

因此,为了区分出内点和离群点,可以通过一个阈值来区分这些不同的信息熵 $I_{\tilde{\Theta}(i,:)}$,进而有效地检测离群点.信息熵值高的矩阵行对应于内点,信息熵值低的矩阵行对应于离群点.在此利用文献[25]提出的熵阈值算法自适应地决定该阈值.那么此时获得了只有 \hat{N} 个内点的数据 $\hat{x} = \{\hat{x}_i\}_{i=1, \dots, \hat{N}}$,以及一个新的偏好统计矩阵 $\hat{\Theta} \in \mathbb{R}^{\hat{N} \times \hat{M}}$,该矩阵只包含了 \hat{N} 个内点和 \hat{M} 个有意义的模型假设.

3.3 模型实例估计算法

通过离群点检测策略得到只有内点的数据后,下一个问题就是如何分配这些内点到不同的模型实例里.一般情况下,可以采用传统的聚类方法(例如, K -means^[27]和Ncut^[28])来聚类数据点到不同的模型实例.然而,这些聚类方法难于自适应地估计聚类的个数而且不能有效地处理分布在交叉模型实例附近的数据点.这是因为传统的聚类方法的设计目的是对数据进行划分,因此传统聚类方法无法有效地处理分布在交叉模型实例附近的数据点^[10].

为解决上述问题,本文提出一种基于相似矩阵学习的图聚类^[16]技术的模型实例估计算法.该算法能够学习相似矩阵并同时执行聚类来估计模型实例.由于来自相同模型实例的内点到对应的模型假设会有类似的残差分布特征^[11],那么来自相同模型实例的数据点间距离越小,其成为邻居的概率就会更大.基于以上分析,本文利用余弦距离作为数据点之间的距离度量.因此,根据偏好统计矩阵 $\hat{\Theta} \in \mathbb{R}^{\hat{N} \times \hat{M}}$ 中对应的两个行向量 $\hat{\Theta}(i,:)$ 和 $\hat{\Theta}(i',:)$,两个数据点 \hat{x}_i 和 $\hat{x}_{i'}$ 之间的余弦距离 $C(i,i')$ 的计算公式如下:

$$C(i,i') = \frac{1}{2} \left(1 - \frac{\hat{\Theta}(i,:) \cdot \hat{\Theta}(i',:)}{\|\hat{\Theta}(i,:) \times \|\hat{\Theta}(i',:)\|} \right) \quad (6)$$

本文定义数据点 \hat{x}_i 能够成为其它数据点邻居的概率为 $s_{ii'}$,其中 $s_{ii'}$ 可看作相似矩阵 $S \in \mathbb{R}^{\hat{N} \times \hat{N}}$ 的一个元素.两个数据点之间的距离越小,则成为邻居的概率越大.因此,概率值 $s_{ii'}$ 与两个数据点 \hat{x}_i 和 $\hat{x}_{i'}$ 的距离值成反比.那么,得到这些数据点之间成为邻居的概率值 $s_{ii'}$,即学习相似矩阵 S 的模型,可以借鉴文献[16]中的方法解决,其目标函数如下:

$$\min_S \sum_{i,i'=1}^{\hat{N}} C(i,i') s_{ii'} + \alpha s_{ii'}^2 \quad (7)$$

$$\text{s. t. } \forall_i, s_i^T \mathbf{1} = 1, 0 \leq s_{ii'} \leq 1, \text{rank}(L_S) = \hat{N} - c$$

式中 α 是正则化参数, $L_S = D - (S^T + S)/2$ 为相似矩

阵 S 对应的拉普拉斯矩阵, 其中 $D \in \mathbb{R}^{\hat{N} \times \hat{N}}$ 定义为一个对角矩阵, 对角矩阵 D 的第 i 行元素是 $\sum_{i'} (s_{ii'} + s_{i'i})/2$. 拉普拉斯矩阵的秩的约束 $\text{rank}(L_S) = \hat{N} - c$ 确保在学习过程中相似矩阵 S 包含 c 个不同的连接成份 (connected components), 其中学习得到的相似矩阵的一个连接成份对应于一个类. 为学习得到每个数据点合适的邻居安排, 本文利用文献[16]的方法以一种近似的优化方式优化目标函数(7). 经过优化得到的相似矩阵 S 是一个更精确的相似矩阵, 它具有 c 个不同的连接成份. 因为在相似矩阵 S 的学习过程中, 每个数据点获得对应的聚类标签, 当完成目标函数(7)的优化时, 即达到了聚类目的. 因此, 基于学习的相似矩阵 S 得到 c 个不同的连接成份, 以获取数据点的聚类标签, 即将内点分配到 c 个不同的模型实例.

基于相似矩阵学习的图聚类技术来估计模型实例是一种非常有效的算法, 它可以根据数据点之间的距离值自适应地安排数据点的邻居点以进行聚类, 有效地处理分布在交叉模型实例附近的数据点. 但是上述模型实例估计算法无法自适应地估计模型实例的个数, 这在一定程度上影响了该算法的实用性. 因此, 本文提出基于拉普拉斯矩阵特征值的分析和最小子集数目的约束, 以自适应地估计模型实例的个数.

在理想情况下, 数据中包含的聚类个数等价于拉普拉斯矩阵对应的特征值为 0 的个数^[28]. 然而, 由于数据中通常存在噪声点和计算精度的限制, 很难得到特征值为 0 的精确值. 那么寻找一个合适的大于 0 的阈值, 对于自适应地估计聚类的个数具有重要意义. 具体来说, 本文通过计算拉普拉斯矩阵 L_S 的特征值小于阈值 μ 的个数, 来初始化相似矩阵 S 的连接成分的个数 c .

在实际的应用中, 当聚类的尺寸 (一个类里面数据点的数目) 小于最小子集数目时, 那么此时的类里会出现无法拟合模型的情况. 同时, 初始化聚类的个数通常大于真实聚类的个数. 因此, 本文提出利用最小子集数目的约束进一步地估计模型实例的个数, 即聚类的最小尺寸 z 必须大于或者等于最小子集数目 p . 此时, 本文先是减少初始化聚类的个数 c , 然后重新运行该算法, 直到避免聚类的最小尺寸 z 小于最小子集数目 p 的情况. 尽管此时该算法可以自适应地估计模型实例的个数, 但是所估计到的模型可能对应于多个冗余的模型实例. 因此, 本文通过合并

两个具有相似残差分布的独立模型实例^[20], 进一步自适应地估计模型实例的个数以提升聚类性能. 最后根据聚类得到的结果估计每个类对应的模型实例参数.

本文所提出的模型实例估计算法能够有效地处理分布在交叉模型实例附近的数据点. 更重要的是, 该算法能够自适应地估计模型实例的个数. 具体算法描述见算法 1.

算法 1. 模型实例估计算法.

输入: 矩阵 $\hat{O} \in \mathbb{R}^{\hat{N} \times \hat{M}}$, 参数 α , 参数 μ , 最小子集数 p

输出: 数据点对应的标签和模型实例的参数

1. 根据式(6)计算两个数据点之间的余弦距离.
2. 利用文献[16]的方法初始化目标函数(7)以得到初始化的相似矩阵 S , 并计算 S 对应的拉普拉斯矩阵 L_S .
3. 计算拉普拉斯矩阵 L_S 的特征值小于参数 μ 的个数, 作为聚类的初始个数 c .
4. REPEAT
5. 优化目标函数(7)以得到学习的相似矩阵 S .
6. 计算 c 个类里面最小尺寸类的数据点的数目 z .
7. IF $z < p$ AND $c > 1$, THEN
8. $c = c - 1$
9. END IF
10. UNTIL $z \geq p$
11. 获得数据点对应的标签 (一个类对应为一个模型实例).

3.4 基于偏好统计数据表征的模型拟合方法

本文提出基于偏好统计数据表征的鲁棒几何模型拟合方法 (A Preference-Statistic-Based Data Representation for Robust Geometric Model Fitting), 简称 PSMF. 本文方法主要分为三个步骤: 首先是基于偏好统计分析进行数据表征, 其次是在偏好统计数据表征的基础上进行离群点检测, 最后是在去除离群点后剩下的数据点中聚类进而估计模型实例. 其总体算法流程见算法 2.

算法 2. 基于偏好统计数据表征的鲁棒几何模型拟合算法 (PSMF).

输入: 数据点 x , 参数 α , 参数 μ , 参数 K

输出: 数据点对应的标签和模型实例的参数

1. 生成一系列的模型假设以及通过 IKOSE^[8] 估计对应的内点噪声尺度.
2. 通过本文提出的数据表征算法构造偏好统计矩阵 Θ (如 3.1 小节所述).
3. 根据式(1)获得修剪的偏好统计矩阵 $\hat{\Theta}$.
4. 根据式(4)计算偏好统计矩阵 $\hat{\Theta}$ 每行的信息熵 $I_{\hat{\Theta}(i, \cdot)}$ (对应于每个数据点).
5. 自适应二分信息熵值 $I_{\hat{\Theta}(i, \cdot)}$ 并将内点从离群点中分

离以得到新的偏好统计矩阵 $\hat{\Theta}$.

6. 根据算法 1 聚类数据点并估计每个类对应的模型实例.
7. 根据文献[20]分析模型实例对应的残差分布特征的方法,融合冗余的模型实例并重新估计模型实例.

针对 PSMF 算法的时间复杂度,本文主要关注模型拟合中的模型选取问题,即没有考虑抽样最小子集和生成模型假设的时间.在 PSMF 中,计算两个数据点之间的余弦距离来得到初始化相似矩阵 S ,即算法 1 的步骤 1 需要最大的时间开销,其时间复杂度大约为 $O(\hat{N}^2)$,其中 \hat{N} 为去除离群点后剩余的数据点的数目.然而,PSMF 的其它步骤所需要的时间开销均比较少.因此,PSMF 算法的整体时间复杂度约为 $O(\hat{N}^2)$.

4 实验与分析

为验证本文方法的有效性,并对其性能进行分析.本节首先将本文提出的偏好统计数据表征算法,与现有的二值偏好数据表征算法和连续偏好数据表征算法进行对比,测试其有效性和鲁棒性.接着进一步比较本文提出的基于偏好统计数据表征的拟合方法(PSMF)与当前几种流行的模型拟合方法的拟合性能.实验中用到的数据集分别是 J-Linkage^[12]中使用的合成数据集和公开的 AdelaideRMF^[29]真实数据集.文献[1, 10, 13, 22]的实验测试,也是基于 AdelaideRMF 数据集.为使得对比更为公平,实验中所有的拟合方法都采用相同的采样方法(此处采用有效采样方法 Proximity Sampling^[12, 30]),生成相同数量的模型假设集.在每个数据集上,所有模型拟合方法中的参数(例如,内点噪声尺度以及目标优化函数涉及的参数等)分别调试出最好的性能以进行对比实验.本文的评估实验所使用的机器配置为 Intel Core i7-7700 3.6 GHz CPU 和 16 GB 内存.

本文首先将 PSMF 的(偏好统计)数据表征算法,与 T-Linkage 的(连续偏好)数据表征算法和 J-Linkage 的(二值偏好)数据表征算法进行对比.因为 T-Linkage 和 J-Linkage 方法主要的不同点是两者采用不同的数据表征算法.因此,本文统一在 T-Linkage 方法的框架下比较不同数据表征算法的性能.本文将 T-Linkage 的数据表征替代为 J-Linkage 的(二值偏好)数据表征(该拟合方法取名为 J-Linkage),以及替代为 PSMF 的(偏好统计)数据表征(该拟合方法取名为 P-Linkage),并同时比较这 3 种方法的拟合性能.然后,将本文拟合方法 PSMF

与当前几种流行的拟合方法进行对比,包括 KF^[15], AKSWH^[8], T-Linkage^[13], RansaCov^[10] 和 MSHF^[22].这是因为:KF 是与本文方法密切相关的拟合方法(该方法也是采用先区分内点和离群点,然后聚类内点以估计模型实例的策略).AKSWH 和 RansaCov 是基于一致性统计分析的拟合方法. T-Linkage 和 MSHF 是基于连续偏好分析的拟合方法.本文采用同文献[1, 9]一样的方法以计算实验拟合误差 FE (Fitting Error),来评价拟合方法的性能.被错误标签的数据点总数记为 N_{error} ,数据点总数为 N ,其 FE 可以定义为

$$FE = \frac{N_{error}}{N} \quad (8)$$

其中, FE 的值越小说明拟合效果越好.

4.1 参数分析与设置

本文方法涉及到的主要参数有: α , μ 和 K 值.参数 α 的选择已经在文献[16]给出了有效的分析和实验.针对参数 μ ,本文经验地设置为一个固定值 0.06,以取得较好的实验效果.参数 K 是本文所提出的偏好统计数据表征算法中残差直方图区间的组数.因此,本节主要分析参数 K 在本文所提出的数据表征算法中,对拟合性能的影响.

本文固定参数 α, μ 和变化 K 值,然后在 4 种不同的拟合任务(即直线拟合(Star5),圆形拟合(Circle5),单应性矩阵估计(Johnsone),基础矩阵估计(Breadtoycar))进行实验.针对每个任务,重复实验 50 次取拟合误差平均值.所采用的数据集来源于 J-Linkage^[12]中使用的合成数据集和公开的 AdelaideRMF^[29]真实数据集.

图 2 展示了对比实验的拟合误差平均值.从图 2 可以看出,针对直线拟合(Star5),在 K 取不同值情况下,它的拟合误差保持相对的稳定.针对圆形拟合(Circle5),当 $1 \leq K \leq 6$ 时,它的拟合误差逐渐变低.当 $K=8$ 时,它的拟合误差保持相对稳定.针对单应性矩阵估计(Johnsone),当从 $K=5$ 开始时,它的拟合误差变得相对稳定.针对基础矩阵估计(Breadtoycar),当 $K=1$ 或者 $K=2$ 时,过于粗糙的残差直方图区间划分降低了数据表征的准确性,使得数据在聚类过程中来自不同模型实例的数据被错误地合并,导致最终的拟合误差较高.但是当 $K > 3$ 时,较为紧凑的残差直方图区间划分,提高数据表征的准确性使得它的拟合误差变低并保持稳定.一般来讲,在这 4 种不同的拟合任务实验中,当 K 值从 1 增加到 6 时,所构造的残差直方图能够获得越来越

充分有效的残差信息以进行数据表征,使得它们的拟合误差逐渐降低.当 $K \geq 6$ 时,所构造的残差直方图能够保留丰富且有效的残差信息以进行数据表征,使得最终的拟合误差保持相对稳定.因此,本文选择 $K=6$ 进行以下的实验.

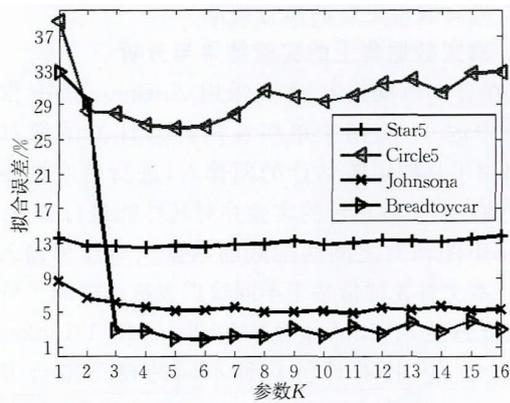


图 2 针对 4 种不同的拟合任务,本文方法 PSMF 在 K 取不同值情况时所获得的拟合误差平均值对比

4.2 合成数据集上的实验结果与分析

在合成数据集上,本文采用 J-Linkage^[12] 中使用的合成数据集,即直线拟合数据集 (Star5 和 Star11),圆形拟合数据集 (Circle5),进行模型拟合实验并对其性能进行分析.本文首先评价基于不同数据表征算法的 3 种模型拟合方法 (P-Linkage, J_1 -Linkage 和 T-Linkage) 的拟合性能.接着评价 6 种不同的模型拟合方法 (T-Linkage, KF, AKSWH, RansaCov, MSHF, PSMF) 的拟合性能.重复各个方法的每个实验 50 次,并在表 1 中给出了这些方法的拟合结果和 CPU 耗时 (s) 的平均值.其中,拟合结果中的“ \pm ”符号前面的数字代表拟合误差的平均值,“ \pm ”符号后面的数字代表拟合误差的标准方差;同时,该时间的计算不包括最小子集抽样和模型假设生成的时间.图 3 展示了 8 种拟合方法的对比结果.在图 3 中交叉点表示离群点,其它不同灰度的圆点分别代表不同模型实例的内点.

表 1 在合成数据集上 8 种拟合方法的拟合结果 (%) 和时间消耗 (s) (最优结果用加粗字体显示)

数据集	P-Linkage	J_1 -Linkage	T-Linkage	KF	AKSWH	RansaCov	MSHF	PSMF	
Star5	拟合误差	13.10 \pm 0.73	14.67 \pm 1.04	14.49 \pm 1.30	15.04 \pm 1.21	14.33 \pm 0.47	10.71\pm0.73	13.77 \pm 2.35	12.58 \pm 1.04
	耗时	77.34	77.37	87.27	11.18	0.67	12.20	0.77	0.61
Star11	拟合误差	28.33 \pm 3.52	33.59 \pm 2.25	33.93 \pm 5.35	46.69 \pm 4.78	31.80 \pm 2.26	25.78 \pm 0.86	26.46 \pm 2.49	25.31\pm2.56
	耗时	173.45	171.15	164.70	25.69	0.86	18.32	0.96	1.24
Circle5	拟合误差	26.98 \pm 4.72	31.32 \pm 5.57	27.29 \pm 4.74	54.92 \pm 3.35	20.36\pm3.77	26.11 \pm 5.95	25.49 \pm 5.92	23.10 \pm 3.01
	耗时	31.34	29.48	31.79	9.66	0.64	36.29	0.57	0.70
平均	拟合误差	22.81 \pm 2.99	26.53 \pm 3.83	25.24 \pm 3.80	38.88 \pm 3.11	22.16 \pm 2.88	20.87 \pm 2.52	21.91 \pm 3.59	19.74\pm2.21
	耗时	94.04	92.67	94.59	15.51	0.72	22.27	0.77	0.88

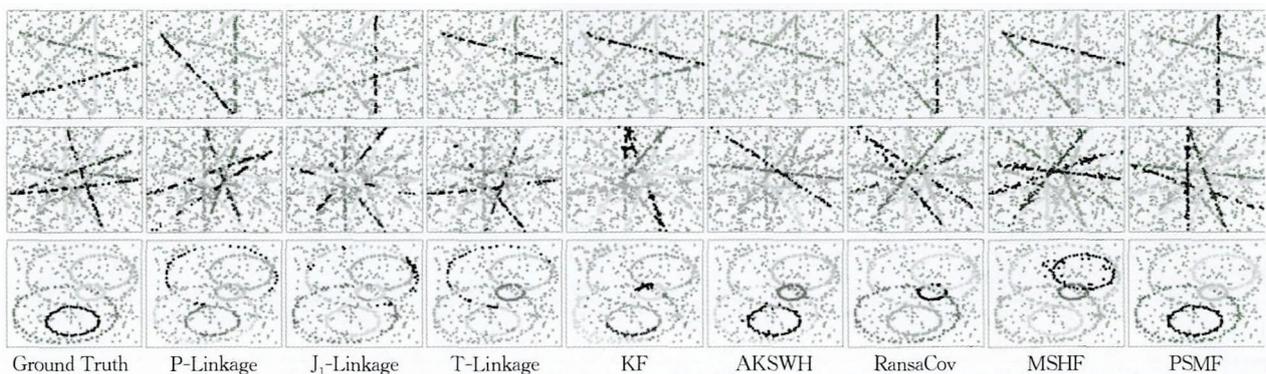


图 3 在合成数据集上 8 种拟合方法的拟合结果对比 (Ground Truth 代表真实的模型实例标签,其它为所拟合的模型实例标签)

对比不同数据表征算法对拟合性能的影响,从表 1 和图 3 中,可以看出:对于 Star5 数据集,基于不同数据表征的拟合方法 T-Linkage, J_1 -Linkage, P-Linkage 都获得了较低的拟合误差平均值和较好的拟合效果.特别地, P-Linkage 取得了最低的拟合误差平均值.对于 Star11 的数据集 T-Linkage 和 J_1 -Linkage 取得了相似的结果,但是它们的拟合结

果仍然比 P-Linkage 差.对于 Circle5 数据集,观察到 P-Linkage 取得比 T-Linkage 好的拟合性能.这是因为 T-Linkage 基于连续偏好的数据表征忽略了残差的一致性信息,不准确的数据表征使得一些内点被错误地去除,导致较差的拟合结果. P-Linkage 利用偏好统计数据表征的优点,融合了偏好信息和一致性信息,包含更丰富的残差信息,以提高拟合的性能.

对比 6 种不同拟合方法的拟合性能,从表 1 和图 3 中,可以看出:对于 Star5,本文拟合方法 PSMF 取得了第二好的拟合性能和最低的时间消耗,同时 RansaCov 取得了最好的拟合结果.然而,RansaCov 需要手动设置模型实例的个数且算法的运行非常耗时. T-Linkage, KF, AKSWH 和 MSHF 都正确拟合了这 5 个模型实例. 对于 Star11,基于参数空间的方法(AKSWH 和 MSHF)运行速度比基于数据点聚类的方法 PSMF 快. 但是 PSMF 取得了最低的拟合误差平均值,因为偏好统计数据表征保留了丰富且有效的残差信息,进而能够有效地进行离群点检测和模型实例估计,以提升拟合的精度. 对于 Circle5, AKSWH 取得了最好的拟合结果. 这是因为 AKSWH 是基于参数空间的拟合方法,该类方法有效地减缓了对数据分布的敏感性. PSMF 取得了次好的拟合结果. PSMF 的模型实例估计算法能够通过相似矩阵学习得到数据点合适的邻居安排,使得它同样也能有效地处理分布在交叉模型实例附近的数据点. MSHF 同样取得较好的拟合结果,这是因为 MSHF 能够保留大部分有代表性的超边(对应于不同的模型实例). T-Linkage 取得较低的拟合误差平均值,但是它无法有效地处理分布在交叉模型实例附近的数据点. KF 取得了高的拟合误差平均值,这是因为许多对应于不同模型实例的内点被错误地去除或者

合并. 而 PSMF 不仅能够很好地检测出离群点,还能够准确地估计出模型实例的个数. 实验的结果表明,本文方法 PSMF 能够有效地处理分布在交叉模型实例附近的数据点(而传统的基于数据点聚类的模型拟合方法无法有效地处理这类问题,例如 KF),从而取得更好的拟合性能.

4.3 真实数据集上的实验结果与分析

在真实数据集上,本文采用 AdelaideRMF 数据集^[29]中的 19 组用于单应性矩阵估计的图像对和 19 组用于基础矩阵估计的图像对,进行单应性矩阵估计和基础矩阵估计的实验并对其性能进行分析. 其中,每组图像对之间匹配好的数据点对作为输入的数据. 本文首先评价基于不同数据表征算法的 3 种模型拟合方法(P-Linkage, J₁-Linkage 和 T-Linkage)的拟合性能. 接着评价 6 种不同的模型拟合方法(T-Linkage, KF, AKSWH, RansaCov, MSHF, PSMF)的拟合性能. 同时重复各个方法的每个实验 50 次,并分别在表 2 和表 3 中给出了这些方法的拟合结果,以及 CPU 耗时(s)的平均值. 同时,在图 4 和图 5 展示了本文方法 PSMF 所获得的部分拟合结果. 在图 4 中十字形点表示离群点,其它不同灰度的圆点分别代表不同模型实例的内点. 在图 5 中交叉点表示离群点,其它不同形状的点分别代表不同模型实例的内点.

表 2 在 AdelaideRMF 数据集上 8 种拟合方法的单应性矩阵拟合结果(%)和耗时(s)(最优结果用加粗字体显示)

数据集		P-Linkage	J ₁ -Linkage	T-Linkage	KF	AKSWH	RansaCov	MSHF	PSMF
Bonython	拟合误差	22.39±17.34	28.49±22.64	28.33±11.54	33.89±8.76	3.49±3.38	2.68±0.23	2.98±2.39	0.10±0.21
	耗时	6.83	6.82	6.68	1.97	1.04	55.56	0.82	0.81
Physics	拟合误差	22.87±4.85	26.70±2.14	26.19±4.36	14.06±1.44	3.11±6.49	8.96±1.68	0.85±1.29	0.66±0.46
	耗时	2.24	1.75	2.18	0.84	1.32	42.73	0.86	0.72
Unionhouse	拟合误差	21.85±22.13	24.58±19.31	26.58±21.56	26.12±4.03	2.11±0.00	2.74±0.13	0.63±0.56	0.21±0.15
	耗时	20.18	20.16	23.91	2.11	0.82	82.74	0.70	0.72
Elderhalla	拟合误差	5.03±3.14	10.47±3.73	12.73±2.96	24.55±4.54	2.33±0.07	1.40±0.00	0.84±0.20	1.36±0.34
	耗时	9.63	7.87	9.66	1.01	0.92	89.77	0.80	0.76
Hartley	拟合误差	6.34±4.14	9.44±2.06	7.56±1.84	15.46±0.18	12.63±0.68	10.22±2.53	8.72±4.94	3.25±5.29
	耗时	22.79	25.69	26.33	2.46	1.08	61.71	0.81	0.80
Library	拟合误差	8.00±2.81	16.56±7.69	12.11±2.12	27.82±0.56	21.95±0.90	5.21±1.07	4.37±2.40	3.44±0.91
	耗时	8.11	7.92	9.46	1.62	0.93	47.53	0.75	0.75
Sene	拟合误差	0.50±0.20	3.36±0.69	2.78±0.86	10.24±7.22	2.28±5.81	1.48±0.19	2.00±1.47	0.60±0.21
	耗时	12.74	10.90	12.93	1.35	1.46	95.59	0.93	0.87
Nese	拟合误差	0.61±0.30	3.62±1.63	1.97±0.60	5.60±3.11	6.34±8.65	7.24±0.25	0.42±0.43	1.50±0.25
	耗时	16.29	11.38	14.13	2.86	2.91	104.93	1.13	1.03
Ladysymon	拟合误差	10.48±0.68	11.14±1.00	10.34±0.83	24.90±0.00	5.36±6.12	10.57±0.75	4.05±1.64	8.61±0.30
	耗时	13.25	10.09	10.92	1.25	1.54	85.88	1.10	0.89
Oldclassicswing	拟合误差	3.03±1.28	4.14±0.62	3.32±0.56	0.95±1.78	1.45±0.14	8.84±0.70	1.93±1.14	6.83±5.87
	耗时	31.15	35.60	35.94	4.47	2.13	211.20	1.26	1.52
Neem	拟合误差	8.54±4.72	14.44±5.94	11.54±4.44	31.19±1.12	6.22±0.52	13.58±1.89	3.36±1.65	3.07±0.35
	耗时	12.97	10.56	12.66	2.61	1.11	90.82	1.15	1.13
Napiera	拟合误差	17.98±3.63	20.40±4.06	20.94±4.25	31.42±0.83	27.78±0.63	17.30±2.92	28.87±9.71	16.32±2.43
	耗时	16.39	16.25	19.45	2.66	0.81	71.41	0.83	0.90

(续 表)

数据集		P-Linkage	J ₁ -Linkage	T-Linkage	KF	AKSWH	RansaCov	MSHF	PSMF
Barrsmith	拟合误差	22.66±6.12	29.17±5.88	25.52±6.63	25.00±2.33	28.05±8.43	23.86±0.86	34.94±18.02	11.41±4.10
	耗时	10.61	10.45	14.34	1.51	0.93	79.31	0.88	0.90
Elderhallb	拟合误差	12.96±3.67	14.75±3.17	13.19±3.28	34.41±0.17	11.84±0.55	14.90±1.82	16.43±6.15	7.45±3.80
	耗时	16.16	15.04	16.19	1.36	1.18	81.97	0.94	1.03
Napierb	拟合误差	20.31±4.44	23.94±2.68	23.04±3.99	24.13±4.73	34.25±0.26	25.14±4.46	16.45±3.13	11.51±3.76
	耗时	16.62	11.37	16.80	1.92	1.25	97.15	1.21	0.97
Johnsona	拟合误差	5.28±4.84	8.26±4.93	7.88±1.48	43.98±4.28	8.45±7.78	9.17±2.85	6.54±5.29	7.35±2.18
	耗时	28.58	28.85	26.73	14.49	2.50	169.83	1.27	1.17
Unihouse	拟合误差	12.38±0.17	13.66±2.52	12.65±0.37	26.09±8.61	12.38±0.09	8.95±1.80	15.53±1.15	12.03±0.34
	耗时	1154.89	1156.50	1145.79	115.17	4.32	1095.30	3.06	5.96
Bonhall	拟合误差	37.98±3.31	42.89±10.14	31.81±2.28	29.80±4.59	35.54±12.50	26.61±4.96	30.82±10.49	24.98±3.54
	耗时	297.26	289.02	397.49	48.05	4.51	640.58	2.36	3.14
Johnsonb	拟合误差	21.09±4.59	24.33±6.74	19.70±6.76	58.94±7.40	28.63±5.00	29.66±6.41	34.52±14.43	38.00±11.69
	耗时	90.96	95.73	92.81	26.65	2.69	407.03	1.36	1.71
平均	拟合误差	13.70±4.86	17.39±5.66	15.69±4.25	25.71±3.44	13.38±3.58	12.03±1.87	11.28±4.55	8.35±2.43
	耗时	94.09	93.30	99.71	12.33	1.76	190.05	1.17	1.36

表 3 在 AdelaideRMF 数据集上 8 种拟合方法的基础矩阵拟合结果 (%) 和耗时 (s) (最优结果用加粗字体显示)

数据集		P-Linkage	J ₁ -Linkage	T-Linkage	KF	AKSWH	RansaCov	MSHF	PSMF
Biscuit	拟合误差	14.61±20.38	24.30±23.50	27.03±24.88	0.61±0.00	1.03±0.26	1.76±0.50	2.21±0.78	2.00±0.61
	耗时	46.64	45.87	44.55	5.32	3.21	384.12	2.97	3.37
Book	拟合误差	11.38±14.49	11.71±14.92	10.28±13.56	5.88±0.00	4.06±0.68	2.78±0.70	2.62±1.48	0.32±0.52
	耗时	16.32	15.32	16.07	1.83	2.70	334.35	2.45	2.88
Cube	拟合误差	37.42±29.16	33.18±29.07	33.25±32.66	8.28±0.00	3.08±0.47	3.90±0.89	3.61±1.28	0.73±0.33
	耗时	34.62	36.01	34.49	3.10	3.25	123.14	3.10	3.26
Game	拟合误差	42.75±32.27	44.75±1.70	44.64±31.31	31.97±8.28	2.40±0.97	2.92±0.56	6.35±2.30	0.73±0.00
	耗时	20.32	20.00	19.81	5.66	3.08	226.85	2.96	3.07
Cubechips	拟合误差	4.01±2.84	5.35±3.43	4.12±2.91	5.00±2.75	5.74±1.16	9.04±2.01	3.87±1.67	2.71±1.73
	耗时	31.89	30.58	31.66	2.82	2.99	155.16	2.77	3.14
Cubetoy	拟合误差	4.42±3.71	3.25±1.88	6.15±5.03	19.04±3.40	11.37±10.60	4.66±1.35	4.06±2.26	1.93±0.53
	耗时	25.10	24.17	25.03	5.68	2.68	368.11	2.50	2.90
Breadcube	拟合误差	1.90±1.16	1.86±0.59	1.98±1.65	13.18±4.35	4.92±1.00	10.02±5.16	2.73±1.71	7.11±3.18
	耗时	26.68	25.50	25.97	3.17	2.89	257.42	2.63	2.99
Gamebiscuit	拟合误差	10.31±22.34	24.18±32.65	13.29±22.95	19.27±7.43	20.67±12.13	7.44±1.12	4.24±1.84	1.19±0.56
	耗时	44.38	45.10	43.92	11.04	3.75	427.90	4.86	3.28
Breadtoy	拟合误差	3.08±2.93	7.63±14.32	4.44±2.46	15.87±6.64	21.91±0.64	17.40±7.76	24.83±3.55	22.01±0.62
	耗时	47.45	45.40	44.75	4.55	3.11	322.13	2.86	3.05
Breadtoycar	拟合误差	6.51±6.73	9.16±7.95	6.96±5.25	31.49±3.15	10.84±1.99	21.08±7.76	9.82±1.73	1.75±0.53
	耗时	11.31	10.88	10.33	2.58	2.73	230.93	2.31	2.84
Biscuitbook	拟合误差	1.11±0.53	1.85±0.76	1.13±0.54	11.53±11.06	4.52±2.11	4.91±1.13	4.90±2.14	1.41±0.19
	耗时	45.22	46.90	44.66	6.05	3.33	167.38	2.81	3.39
Biscuitbookbox	拟合误差	3.67±4.64	7.30±7.05	4.40±4.53	17.54±5.93	14.21±6.03	17.59±5.61	8.92±5.25	2.32±1.69
	耗时	29.26	29.59	28.84	5.15	4.12	187.00	3.41	3.03
Breadcubechips	拟合误差	6.96±7.25	8.78±8.79	7.74±6.57	35.32±7.42	18.22±1.53	22.30±7.24	11.91±4.58	6.78±4.27
	耗时	21.47	21.16	21.44	3.66	2.71	139.44	2.86	3.05
Cubebreadtoy-chips	拟合误差	18.07±8.43	18.83±7.27	17.71±6.58	33.73±5.23	22.36±4.63	24.74±4.21	16.45±4.37	3.67±4.58
	耗时	42.99	43.63	47.12	21.77	2.77	179.96	3.41	3.37
Breadcartoy-chips	拟合误差	13.76±7.48	21.60±7.39	19.65±8.07	28.72±6.53	34.26±5.50	30.46±3.12	21.10±3.97	13.21±0.56
	耗时	23.00	21.66	21.68	9.03	2.76	296.40	2.77	3.05
Carchipscube	拟合误差	12.97±2.94	14.06±3.05	13.39±3.56	23.03±13.25	35.15±2.20	27.82±8.41	19.64±6.68	27.76±11.88
	耗时	10.82	10.26	10.36	2.19	2.39	95.76	2.43	2.71
Toycubecar	拟合误差	9.75±3.35	13.65±1.62	11.29±10.39	24.70±13.13	32.35±6.16	34.90±11.22	25.40±11.83	21.90±6.74
	耗时	17.74	20.16	16.22	3.53	2.69	145.78	3.14	2.78
Boardgame	拟合误差	14.25±0.86	14.49±0.86	14.37±0.94	28.07±9.34	18.03±1.61	20.72±2.10	21.15±1.16	17.20±0.70
	耗时	32.98	32.35	31.24	4.26	3.11	172.03	2.85	3.13
Dinobooks	拟合误差	20.04±1.44	27.14±1.25	22.06±2.62	29.57±5.93	29.64±6.31	27.31±2.50	23.00±1.29	24.19±1.42
	耗时	58.24	68.76	54.79	20.05	2.96	234.05	3.50	3.40
平均	拟合误差	12.47±9.10	15.42±8.84	13.89±9.81	20.15±5.99	15.51±3.47	15.35±3.85	11.41±3.15	8.36±2.16
	耗时	30.86	31.23	30.15	6.39	3.01	234.10	2.98	3.09

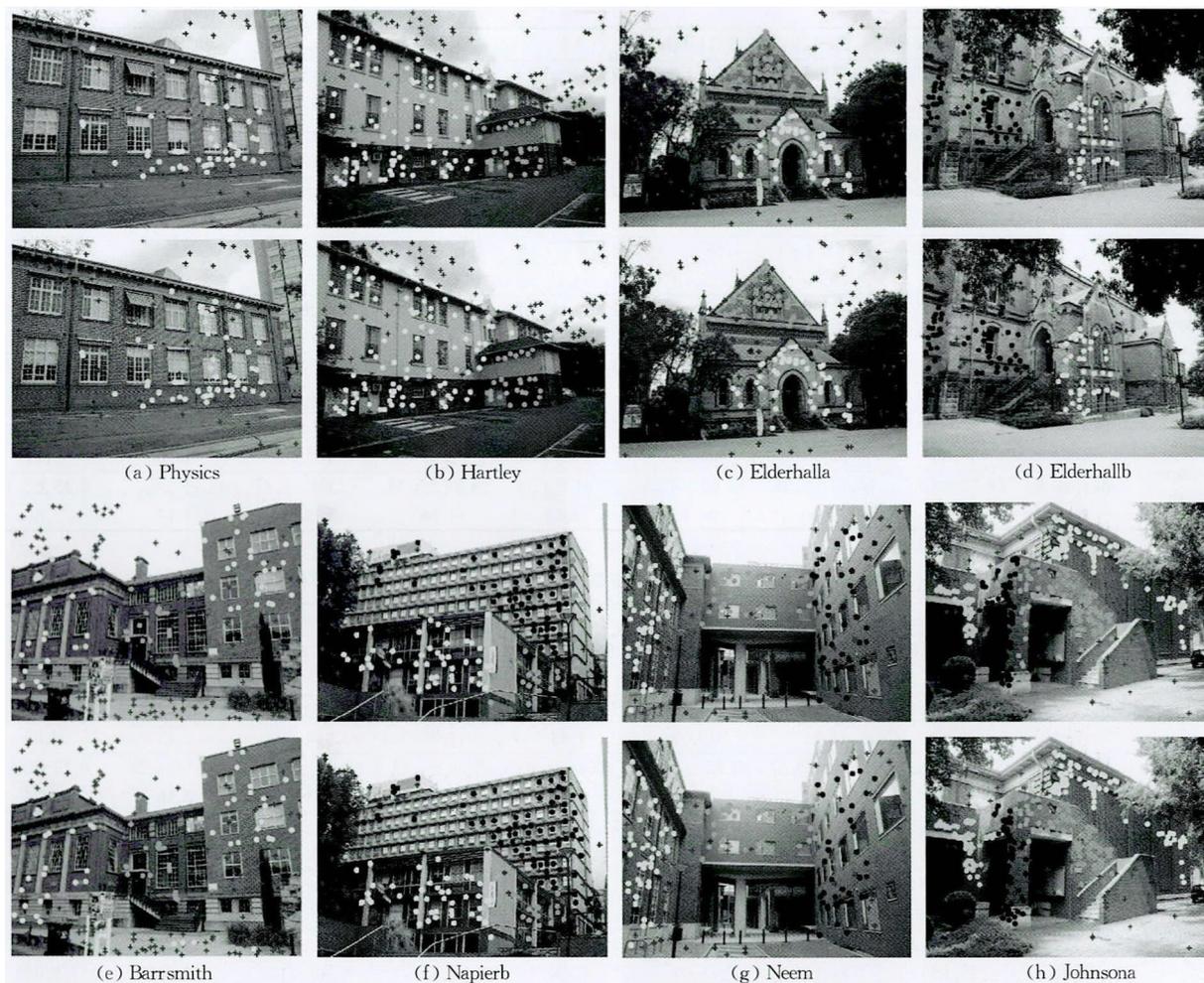


图 4 在 AdelaideRMF 数据上的单应性矩阵估计示例(第 1 行和第 3 行中的图片分别代表真实的模型实例标签,第 2 行和第 4 行中的图片分别代表本文拟合方法 PSMF 所拟合的模型实例标签.这里只显示两视图中的一张图片)

单应性矩阵估计. 本文采用直接线性变换(Direct Linear Transformation(DLT))^[31]来估计单应性矩阵,用桑普森距离(Sampson distance)^[31]计算其对应的残差.对比 3 种不同数据表征算法对拟合性能的影响,从表 2 可以看出:对于大部分的数据, J_1 -Linkage 取得了较高的拟合误差平均值,P-Linkage 取得了比 T-Linkage 低的拟合误差平均值.P-Linkage 取得了更好的拟合结果,这是因为本文提出的数据表征算法同时考虑了残差的偏好信息和一致性信息,进而有效地提高了数据点之间距离测量的准确性,以取得更好的拟合性能.

对比 6 种不同的模型拟合方法的拟合性能,从表 2 和图 4 中,可以看出:本文拟合方法 PSMF,在 19 组数据中的 11 组数据上取得了最低的拟合误差平均值.尽管,PSMF 的时间消耗稍微高于 MSHF,但是 PSMF 明显比 MSHF 提高了拟合精度.这是因为 PSMF 直接对数据点进行聚类.然而,基于模

型假设聚类的方法 MSHF 的拟合结果对内点噪声尺度十分敏感,其中内点噪声尺度往往难以被准确地估计. AKSWH 的时间消耗比 PSMF 大,同时它也取得比较高的拟合误差平均值.虽然 T-Linkage 和 RansaCov 也取得相对低的拟合误差平均值,但是这两种方法对数据中包含有较多离群点的数据分布比较敏感. KF 取得了最高的拟合误差平均值.注意到 PSMF 同 KF 一样,也是采用先去除离群点,然后聚类点进而估计模型实例的策略.尽管 KF 在单应性矩阵估计任务中能够检测到大部分的内点,但它经常错误地估计模型实例.这也验证了本文方法 PSMF 中的模型实例估计算法的有效性和鲁棒性.因此,本文拟合方法 PSMF 与其它拟合方法相比兼具拟合性能和运行时间的优势.

基础矩阵估计. 本文采用直接线性变换(DLT)^[31]来估计基础矩阵,用桑普森距离(Sampson distance)^[31]计算其对应的残差.为对比 3 种不同数据表征算法

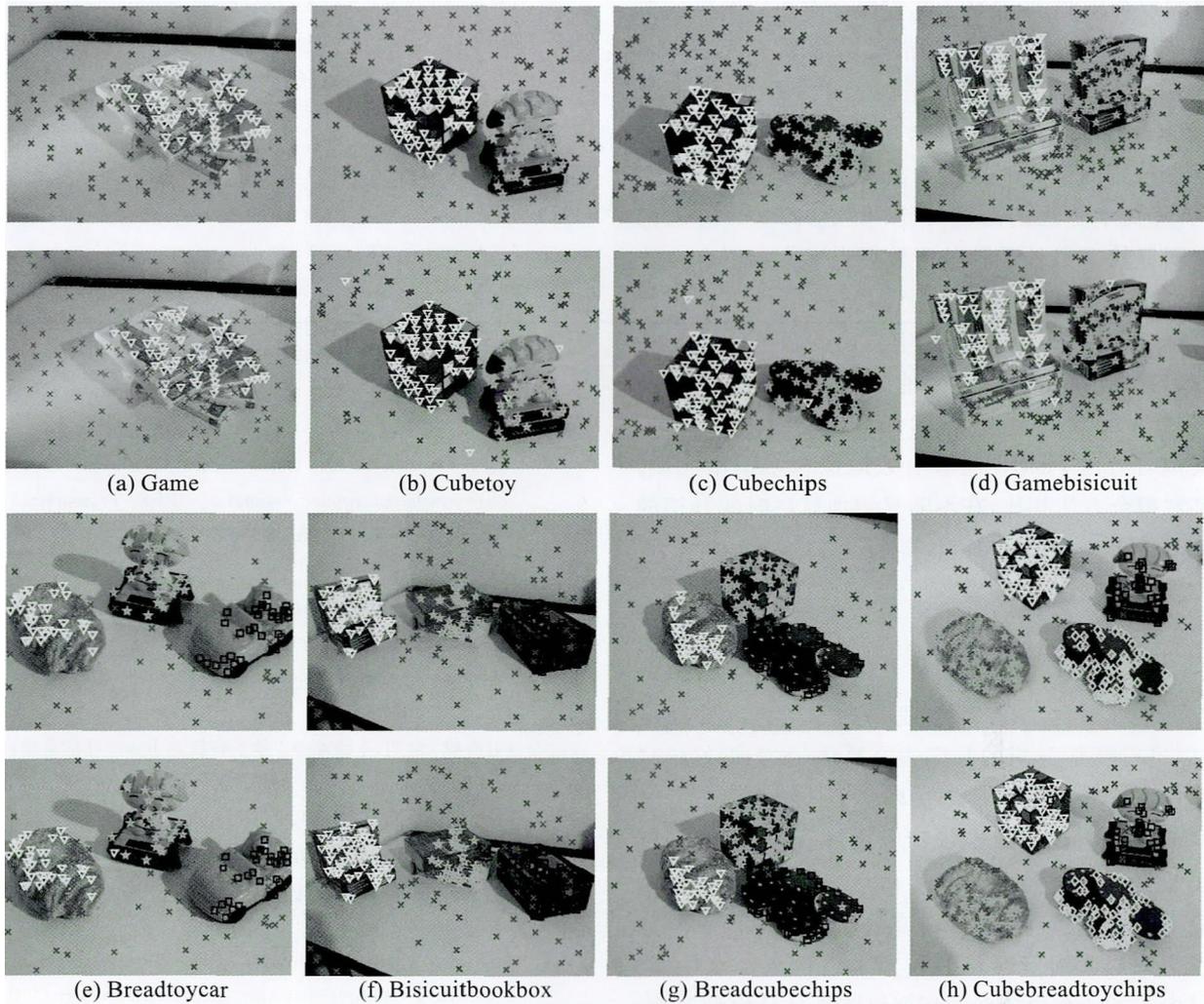


图 5 在 AdelaideRMF 数据上的基础矩阵估计示例(第 1 行和第 3 行中的图片分别代表真实的模型实例标签,第 2 行和第 4 行中的图片分别代表本文拟合方法 PSMF 所拟合的模型实例标签.这里只显示两视图中的张图片)

对拟合性能的影响,从表 3 中,可以看出 J_1 -Linkage 取得了比 T-Linkage 和 P-Linkage 差的拟合结果. P-Linkage 取得了比 T-Linkage 好的拟合结果.相比 J_1 -Linkage(将所有残差值映射到直方图的两个区间)和 T-Linkage(将所有小于阈值的每个残差值映射到直方图对应的一个区间),在 P-Linkage(将所有的残差值映射到几个区间)中,所提出的数据表征算法保留了丰富且有效的残差信息,进而提高模型拟合的性能.图 6 展示了 3 种基于不同数据表征算法的模型拟合方法(P-Linkage, J_1 -Linkage 和 T-Linkage)在不同的拟合任务中所取得的拟合误差平均值.从图 6 中可以看出,在不同的拟合任务中,基于偏好统计数据表征的拟合方法 P-Linkage 取得了最低的拟合误差平均值,这也验证了本文所提出的数据表征算法的有效性和鲁棒性.

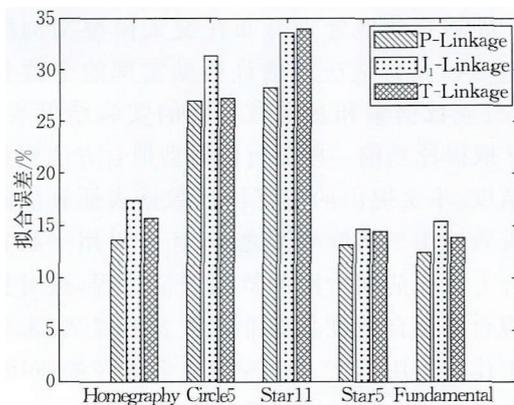


图 6 基于 3 种不同数据表征的拟合方法在不同拟合任务中的拟合结果

对比 6 种不同的模型拟合方法的拟合性能,从表 3 和图 5 中,可以看出:本文拟合方法 PSMF 能够有效分割出大部分的模型实例,并且在 19 组数据中的 11 组数据取得最低的拟合误差平均值. PSMF

取得的拟合精度,明显好于其它拟合方法所取得的拟合精度. 尽管 PSMF 的时间消耗稍微多于 AKSWH 和 MSHF 的时间消耗. 这是因为 PSMF 是基于数据点聚类的拟合方法,然而,AKSWH 和 MSHF 是基于参数空间聚类的拟合方法来搜索代表性的模型实例,降低了算法的时间开销. 但是,PSMF 的运行速度明显快于另外几种基于数据点聚类的拟合方法(KF, T-Linkage 和 RansaCov),同时取得了最高的拟合精度. 这是因为 PSMF 中精确的数据表征和鲁棒的离群点检测策略,有效地提高了 PSMF 的拟合性能. 实验结果表明,在包含大量离群点的数据中,本文拟合方法 PSMF 能够有效地拟合多模型实例. 与其它拟合方法相比,在不明显增加算法时间开销的情况下,本文方法 PSMF 取得了更好的拟合结果.

5 结 论

本文提出了一种基于残差直方图区间数值变化,以进行特征分析的数据表征算法(称为偏好统计数据表征). 该数据表征算法融合了传统模型拟合方法中偏好分析和一致性统计分析的优点,有效地保留了数据表征所需要的残差信息. 在此偏好统计数据表征的基础上,本文利用该数据表征中内点和离群点之间显著的信息熵值差异来进行离群点检测,同时构造了一个更精确的相似矩阵来估计模型实例. 本文将所提出的新的模型拟合方法称之为基于偏好统计数据表征的鲁棒几何模型拟合方法(PSMF). PSMF 能够有效地处理分布在交叉模型实例附近的数据点以及自适应地估计模型实例的个数和参数. 在合成数据集和真实数据集的实验结果表明,PSMF 取得比当前一些流行的模型拟合方法更好的拟合精度. 本文提出的偏好统计数据表征算法除了可以有效地用于离群检测之外,还可以用于其它模型拟合方法在前期阶段的数据表征过程,这对提高模型拟合的拟合精度具有非常重要的意义. 未来的研究工作将运用本文方法探索更多潜在的应用,比如在医学图像分析和增强现实的应用.

参 考 文 献

[1] Amayo P, Piniés P, Paz L M, Newman P. Geometric multi-model fitting with a convex relaxation algorithm//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018; 8138-8146

- [2] Pham T T, Chin T-J, Yu J, Suter D. Simultaneous sampling and multi-structure fitting with adaptive reversible jump MCMC //Proceedings of the Conference on Neural Information Processing Systems. Granada, Spain, 2011; 540-548
- [3] Chin T-J, Suter D, Wang H. Multi-structure model selection via kernel optimization//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA, 2010; 3586-3593
- [4] Wang Yu-Xia, Zhao Qing-Jie, Zhao Liu-Jun. Robust object tracking based on FREAK and P3CA. Chinese Journal of Computers, 2015, 38(6): 1188-1201(in Chinese)
(王宇霞, 赵清杰, 赵留军. 基于 FREAK 和 P3CA 的鲁棒目标跟踪. 计算机学报, 2015, 38(6): 1188-1201)
- [5] Wang H, Mirota D, Hager G D. A generalized kernel consensus-based robust estimator. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(1): 178-184
- [6] Zhao Hai-Feng, Lu Ming, Bu Ling-Bin, et al. Medical image registration based on feature points and Rényi mutual information. Chinese Journal of Computers, 2015, 38(6): 1212-1221(in Chinese)
(赵海峰, 陆明, 卜令斌等. 基于特征点 Rényi 互信息的医学图像配准. 计算机学报, 2015, 38(6): 1212-1221)
- [7] Fischler M A, Bolles R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 1981, 24(6): 381-395
- [8] Wang H, Chin T-J, Suter D. Simultaneously fitting and segmenting multiple-structure data with outliers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(6): 1177-1192
- [9] Mittal S, Anand S, Meer P. Generalized projection-based M-Estimator. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(12): 2351-2364
- [10] Magri L, Fusiello A. Multiple model fitting as a set coverage problem//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016; 3318-3326
- [11] Zhang W, Kösecká J. Nonparametric estimation of multiple structures with outliers//Proceedings of the European Conference on Computer Vision. Graz, Austria, 2006; 60-74
- [12] Toldo R, Fusiello A. Robust multiple structures estimation with J-Linkage//Proceedings of the European Conference on Computer Vision. Marseille, France, 2008; 537-547
- [13] Magri L, Fusiello A. T-Linkage: A continuous relaxation of J-Linkage for multi-model fitting//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014; 3954-3961
- [14] Magri L, Fusiello A. Multiple structure recovery via robust preference analysis. Image and Vision Computing, 2017, 67: 1-15

- [15] Chin T-J, Wang H, Suter D. Robust fitting of multiple structures: The statistical learning approach//Proceedings of the IEEE Conference on International Conference on Computer Vision. Kyoto, Japan, 2009: 413-420
- [16] Nie F, Wang X, Huang H. Clustering and projected clustering with adaptive neighbors//Proceedings of the Conference on Knowledge Discovery and Data Mining. New York, USA, 2014: 977-986
- [17] Choi S, Kim T, Yu W. Performance evaluation of RANSAC family//Proceedings of the British Machine Vision Conference. London, UK, 2009: 458-467
- [18] Torr P H. Geometric motion segmentation and model selection. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 1998, 356(1740): 1321-1340
- [19] Isack H, Boykov Y. Energy-based geometric multi-model fitting. International Journal of Computer Vision, 2012, 97(2): 123-147
- [20] Yu J, Chin T-J, Suter D. A global optimization approach to robust multi-model fitting//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA, 2011: 2041-2048
- [21] Pham T T, Chin T-J, Yu J, Suter D. The random cluster model for robust geometric fitting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(8): 1658-1671
- [22] Wang H, Xiao G, Yan Y, Suter D. Searching for representative modes on hypergraphs for robust geometric model fitting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(3): 697-711
- [23] Nie F, Wang X, Jordan M I, Huang H. The constrained Laplacian rank algorithm for graph-based clustering//Proceedings of the AAAI Conference on Artificial Intelligence. Phoenix, USA, 2016: 1969-1976
- [24] Wand M P, Jones M C. Kernel smoothing. Chapman and Hall, 1995
- [25] Ferraz L, Felip R, Martínez B, Binefa X. A density-based data reduction algorithm for robust estimators//Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis. Girona, Spain, 2007: 355-362
- [26] Cheung Y-M, Jia H. Categorical-and-numerical attribute data clustering based on a unified similarity metric without knowing cluster number. Pattern Recognition, 2013, 46(8): 2228-2238
- [27] MacQueen J. Some methods for classification and analysis of multivariate observations//Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, 1967, 1(14): 281-297
- [28] Shi J, Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905
- [29] Wong H S, Chin T-J, Yu J, Suter D. Dynamic and hierarchical multi-structure geometric model fitting//Proceedings of the IEEE Conference on International Conference on Computer Vision. Barcelona, Spain, 2011: 1044-1051
- [30] Kanazawa Y, Kawakami H. Detection of planar regions with uncalibrated stereo using distributions of feature points//Proceedings of the British Machine Vision Conference. London, UK, 2004: 247-256
- [31] Hartle R, Zisserman A. Multiple View Geometry in Computer Vision. UK: Cambridge University Press, 2004



GUO Han-Lin, Ph. D. candidate.

His research interests include computer vision, robust model fitting, and data science in health and medicine, etc.

XIAO Guo-Bao, Ph. D. , professor. His research interests include computer vision, and pattern recognition, etc.

Background

Robust geometric model fitting is a fundamental research problem in computer vision with wide spread applications, such as motion segmentation, homography/fundamental matrix estimation, image matching and medical image analysis. The

YAN Yan, Ph. D. , associate professor. His research interests include computer vision, and machine learning, etc.

LIN Shu-Yuan, Ph. D. candidate. His research interests include computer vision, and robust model fitting, etc.

SUTER David, professor, Ph. D. supervisor. His research interests include computer vision, and machine learning, etc.

WANG Han-Zi, professor, Ph. D. supervisor. His research interests include computer vision, and pattern recognition, etc.

main task of model fitting is to robustly estimate the number and the parameters of model instances in data. It is a challenging task since a scene typically contains multiple geometric structures and the observed data are usually contaminated

with noise and outliers. One of the most classical fitting methods is RANSAC that works effectively for fitting single-structure data. However it is difficult to handle multiple-structure data, even when the percentage of outliers is low.

In this work, we propose a robust model fitting method to deal with the multiple-structure data contaminated with noise and outliers. The key idea is to represent residuals below a threshold using K bins of a histogram for data representation. The method initially filters insignificant hypotheses using non-parametric density estimation, and then removes outliers based on the entropy of their residuals (with respect to each hypothesis) which are categorized into K -bins. Using the cosine distance of the binned residual representation, we construct a similarity matrix which is used to cluster the data points. Then, we learn the similarity matrix and assign labels to the data points. The number of model instances is estimated by re-applying the graph clustering algorithm for a

decreasing number of clusters, until the minimum members of each cluster satisfy a chosen threshold (i. e. , the members of minimal sampled subset). The proposed method is compared with several state-of-the-art methods on synthetic data and real image pairs to show the effectiveness of the proposed method. The main insight of this work is the regularization afforded by the binned representation of residuals that provides a good balance between the consensus based methods and preference based methods.

We have been working on robust geometric model fitting for many years and have been published in the international journals and conferences including the IEEE T-PAMI, IJCV, IEEE T-Cyber, PR, IEEE T-ITS, ICCV, ECCV, AAAI, etc.

This work is supported by the National Natural Science Foundation of China under Grant Nos. U1605252, 61702431, 61571379 and 61872307.