

LTrSAC：局部三角拓扑结构约束下的误匹配消除

摘要：图像匹配是纯视觉应用不可或缺的关键环节，然而在图像匹配之后误匹配始终是人们多年来关注的问题。当只有少量的误匹配时，著名的 RANSAC(Random Sample Consensus)等经典方法能够轻松应对，但当场景变得复杂且误匹配增多时，误匹配消除就会成为一个主要问题。为此，本文提出一种基于局部三角拓扑相似性结合重采样的误匹配消除法 LTrSAC(Local Triangle Topology Similarity Sample Consensus)。该方法首先基于图像特征点构造 Delaunay 三角拓扑网络结构，并针对匹配的特征点，基于三角拓扑关系在其局部范围内，计算与其它相邻特征点的相似性，确定出相似节点的数量与不相似节点的数量，进而求取综合量化相似性得分；然后基于相似性得分最高的一部分匹配点来计算两幅图像之间的单应性矩阵，并通过投影变换误差检测实现对误匹配的去除。实验结果表明，与经典的 RANSAC 相比，准确率平均提高了 2%以上，综合性指标 F 值平均提高了 1.5%；与目前最好的 LGSC(Local Graph Structure Consensus)对比，准确率平均提高了 4%，F 值平均提高了 3%。

关键词：图像匹配，误匹配消除，RANSAC，三角拓扑，单应性矩阵

1 引言

图像匹配是通过在同一物体或场景的两幅不同图像之间识别并构建可靠的对应关系，广泛支撑了诸多计算机视觉的任务^[1-10]，是一个经典的但一直被人们追随和完善精度的难题。图像匹配包括基于灰度信息的匹配、基于特征的匹配、基于模板的匹配、基于关系的匹配等，其中基于特征的匹配应用更为广泛，但其中的误匹配消除也不可避免。

一般来说，基于特征的图像匹配可以分为特征点提取、特征描述子构造、特征点匹配和误匹配删除四个阶段。为了获得准确的特征点匹配，人们提出了众多基于局部特征描述符进行匹配的方法，从基于欧氏距离匹配的著名方法 SIFT(Scale Invariant Feature Transform)^[11]和 SURF(Speeded Up Robust Feature)^[12]，到基于汉明距离的快速方法 ORB(Oriented FAST and Rotated BRIEF)^[13]、BRISK(Binary Robust Invariant Scalable Keypoints)^[14]和 FREAK(Fast Retina Keypoint)^[15]。虽然基于局部特征描述符进行特征匹配的方法已是应用最广泛的方法，但误匹配却不可避免。一般来说，即使不考虑异常值，一个简单的 N 个特征点与另外 N 个特征点匹配，则会伴随着 $N!$ 种排列，当 N 越大，出现误匹配的数量越大。为了防止匹配中出现过多的误匹配，一些基于神经网络的方法开始出现，如 LIFT (Learned Invariant Feature Transform)^[16]、SuperPoint(Self-Supervised Interest Point Detection and Description)^[17]等，在一定程度上规避了误匹配，提高了准确率和精密度。一方面，由于深度学习模型往往依赖于大量的标注数据进行训练，且模型的性能很大程度上取决于训练数据的质量和多样性，因此深

度特征匹配器在通用性方面具有明显的局限性；另一方面，基于深度学习方式在特征匹配过程中，难以直观地理解模型是如何提取和匹配特征的，导致决策过程难以解释与理解，这种不透明性会影响到特征匹配的准确性和可靠性；此外，基于深度的特征匹配方法在训练过程中需要高性能的 GPU 与大量内存等资源，成本远高于基于局部描述符的方法。因此，在实际应用中，基于局部特征的图像匹配方法仍然是人们所关注的重要方法。

总体来看，基于局部描述符的匹配方法在实际应用中出现的问题可概括：首先，如果图像存在低质量、遮挡、重复结构等问题，仅使用局部描述符信息将不可避免地导致更加严重的误匹配。其次，两幅图像之间的转换模型多种多样，目前误匹配消除算法适应能力不足。为此，本文提出了一种基于局部三角形拓扑结构相似性重采样误匹配消除的方法。该方法首先利用特征点并结合其一定邻域内的点进行 Delaunay 三角拓扑结构构建，基于三角拓扑结构计算匹配点相似性得分并进行排序，进而选取相似性得分最高的匹配点来实现单应矩阵的求解，最后基于该矩阵在实现重投影的基础上，对比筛选并消除误匹配，并取得了比目前最佳方法 LGSC(Local Graph Structure Consensus)^[18]更优秀的表现。本文的特色和主要贡献概括为：

(1) 提出了一种基于邻域拓扑结构的误匹配程度量化判定方法。该方法通过引入 Delaunay 三角剖分来获取特征点与其邻域点的三角拓扑关系信息，并以此为依据计算特征节点的相似度得分，进而实现了对误匹配程度的判定。

(2) 设计了一种基于局部重采样的单应矩阵或基础矩阵求解方法。该方法通过计算特征点相似性得分并选取得分最高的若干点，进而进行局部重采样来求解单应性矩阵，利用单应矩阵进行全局误匹配的消除，实现局部估计全局，避免了基于几何的方法易受到局部范围内较多误匹配干扰的问题。

2 相关工作

尽管存在很多特征匹配的方法并得到了广泛的应用，但伴随误匹配消除一直以来是不可避免的特征匹配的构成部分。为了消除误匹配，已经出现了众多著名的方法，可分为三类：基于回归的、基于重采样的和基于几何的。

(1) 基于回归的误匹配去除方法

基于回归的误匹配去除方法的优点是过程相对较为简单，收敛速度也比较快。最常见的回归方法是最小二乘法^[19]，通过计算得到最小平方和误差，来找到一个最佳的参数并得到最

佳的模型，进而通过该模型尽可能多地找到正确的匹配点对。诸如 Li 等人^[20]提出了子区域最小二乘迭代拟合的误匹配消除方法，通过对模型进行迭代回归，直到所有匹配点的误差达到阈值。但是基于二乘法的回归方案需要所有正确匹配点基本满足一定的函数；同时在拟合模型之前，可能还需要对数据进行预处理，否则会影响模型检测匹配正确与否的准确性。此外，Maier 等人^[21]提出了一种基于统计光流的引导匹配方案，在精度和效率方面都取得了可喜的结果；Niu 等人^[22]提出的多项式回归，在去除彩色图像不匹配方面表现良好。

尽管基于回归的方法在不断被优化，消除误匹配的效率也提高了很多。但是此类方法的抗干扰能力弱，在少量不匹配的数据中可以获得一个比较好的结果，但在参杂较多的误匹配点对时，使得不匹配检测变得非常困难，这也导致了基于回归的误匹配检测方法应用的局限性。

(2) 基于几何与拓扑的误匹配去除方法

近年来，基于几何的错误匹配去除方法成为了特征点匹配与误匹配消除的热点方法^[18]，其核心是将匹配点对与其周围邻域点相结合，构建特征点邻域之间的几何或拓扑关系来去除误匹配。诸如 Aguilar 等人提出了一种基于图变换匹配的方法 GTM(Generative Topographic Mapping)^[23]，这是一种经典的基于几何的方法，通过特征点的最近邻特征点构造 KNN 无向图，利用无向图之间的长度以及角度进行误匹配的剔除。Izadi 等人提出了一种基于 GTM 的改进算法 WGTM(Weighted Graph Transformation Matching)^[24]，WGTM 在 GTM 算法的基础上对不同的节点和连接关系赋予不同的权重，权重是基于节点之间的相似性度量、节点的重要程度、连接关系的可靠性等因素进行计算的，WGTM 能够更好地适应大规模数据和复杂的几何变换。huang 等人利用梯度结合匹配描述子的角度余弦相似性进行误匹配的消除^[25]，提高了匹配的正确率，然而该方法在出现仿射变化的图片中会失效。Bian 等人提出了一种基于网格运动统计的方法 GMS(Grid-based Motion Statistics)^[26]，GMS 假设一个正确的匹配点在其周围存在更多正确的匹配点对并以此来剔除误匹配，大大提高了匹配准确率，然而 GMS 需要数量良好的初始匹配点对。Jia 等人提出了一种基于局部图结构一致性的图像特征点匹配方法 LGSC^[18]，该方法通过计算每个特征点局部区域内节点与邻居点的相似性得分、节点之间的边的相似性得分，来综合判断该特征点是否正确匹配，并将相似程度高的特征点保留下来，将相似程度低的点筛除掉，得到比较正确的匹配结果。当局部区域内大多数是误匹配时，该方法会退化，反而会将正确的匹配点对消除。

总体来看，虽然基于几何约束或拓扑结构的方法可以有效地检测错配，并具有较高的精度，但是当局部范围内存在较多的误匹配时，因为局部范围内的误匹配过多会影响正确匹配

的计算，因而此类方法会受到巨大的干扰导致不稳定。

(3) 基于重采样的误匹配去除方法

基于重采样的误匹配去除方法是去除误匹配邻域内应用最广最流行的方法，其是通过基本矩阵或单应性矩阵中的异常值来判定不匹配的。基于重采样的方法与回归方法相似，都是在寻找一个最优的模型，不同的是，基于重采样的方法是尽可能地在误差范围内囊括最多匹配数量。与基于回归的方法相比，大的失配对重采样方法的影响相对较小，因为在基于重采样方法中并非所有点都需要符合模型。

目前应用最广泛的图像误匹配剔除算法是 1987 年由 Fischler 等人提出的基于统计模型的随机采样一致性算法 RANSAC^[27]。RANSAC 通过不断地迭代计算最优基础矩阵或者单应性矩阵得到内点，也即正确的匹配点。RANSAC 拥有良好的鲁棒性，但是当特征点数量众多且密集时，仍然会存在较多的误匹配。此后，越来越多的改进 RANSAC 的方法被提出，包括基于采样的优化、基于损失函数的优化，以及基于模型估计的优化等改进的 RANSAC 方法。诸如 Torr 改进 RANSAC，提出了 MLESAC(Maximum Likelihood Estimator Sample Consensus)^[28]，其引入了统计学的思想，通过似然度来对随机抽样产生的模型假说进行评估，可将求解最大似然估计问题转化为求解代价函数最小值问题。Torr 之后又提出了一种新的采样技术 NAPSAC(N-Adjacent points sample consensus)^[29]，它逐渐从局部移动到全局。NAPSAC 算法对高维、内点比例低的数据集效果明显，但是容易产生退化，对于距离都很近的数据集效果差。Chum 等人提出了一种对 RANSAC 优化的渐进采样算法 PROSAC(Progressive Sample Consensus)^[30]，该方法是从不断增大的最佳对应点集合中进行采样的，在保证准确率的情况下减少了计算量，提高了运算速度。Barath 等人提出了一种无需设置内点阈值的算法 MAGSAC(Marginalizing Sample Consensus)^[31]，其设计了一种新的模型质量评估函数，由于该函数与内点阈值无关，因此不必求解内点集以评估模型，显著提高了鲁棒估计的精度。

基于重采样的方法从全局内估计一个符合所有匹配点的最佳模型，然而当误匹配过多时，可能会导致模型误差过大，也即求解的单应矩阵或者基础矩阵偏差过大，往往会导致将正确的匹配消除。

综上所述，基于几何的方法和基于重采样的方法效果，比基于回归的方法具有广泛的适应性。但是目前最好的基于重采样地方法诸如 RANSAC 和 PROSAC，都很难随机选择正确的点集。而基于几何的方法在局部范围内误匹配较多时，计算是否为正确匹配则更容易受到干扰。为此，基于优势互补的策略，本文提出了一种基于局部三角形拓扑结构的相似性重采样误匹配消除方法，即将基于几何的方法与基于重采样的方法实现融合，可提升特征点误匹

配去除方法的鲁棒性。

3 方法概述

本文方法的技术路线图如图 1 所示，主要包括三部分，即初始匹配点集获取、三角拓扑结构构建、相似性计算与重采样。

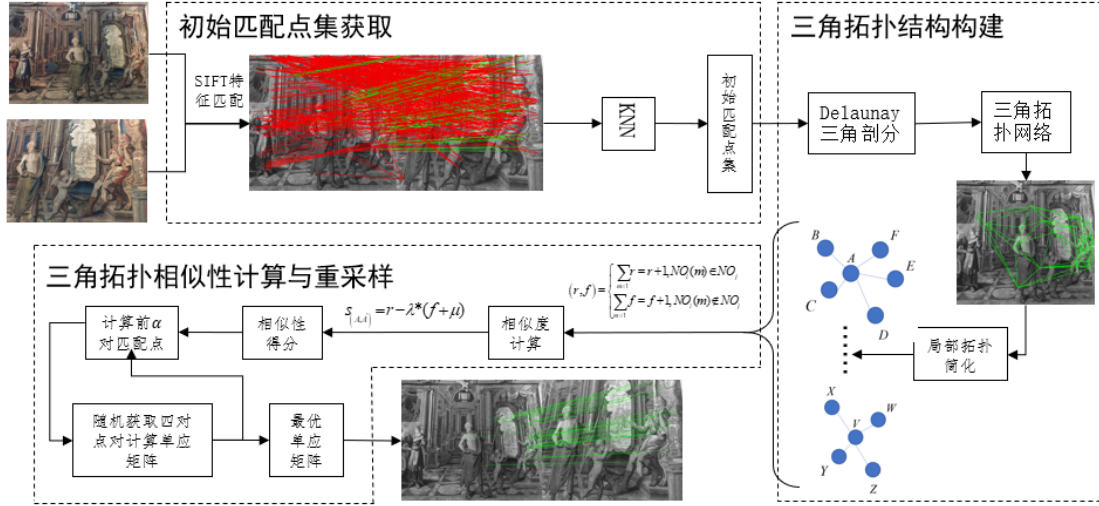


图 1 方法技术路线图

首先，通过 SIFT 获取匹配点集，利用最近邻方法消除掉一部分误匹配点对获得初始匹配点集。

其次，基于初始匹配点对构建三角拓扑网络结构。虽然构建三角拓扑网络的方法有很多，诸如 Delaunay 三角剖分、Power Crust 三角剖分、基于密度的三角剖分等。但是由于 Delaunay 三角剖分具有可自动过滤重复的点能力，更重要的是 Delaunay 三角剖分是基于最近邻原则的，它能够均匀地分布三角形，从而使三角形网格结构更加合理。因此本文选择 Delaunay 三角剖分来构造三角拓扑网络结构，并对其实施局部拓扑简化，以便提高计算效率。

最后，基于简化后的三角拓扑结构，计算每对匹配点的三角拓扑相似性。该过程以获取排名最高的一部分匹配点对，完成重采样并拟合出最佳模型，获得单应矩阵(或基础矩阵)，进而基于单应矩阵来实现误匹配的消除。

本文的方法由于是基于一个局部最佳匹配点对基础上，构建了单应性矩阵，而非从全局选择，因此使获取单应性矩阵的效率更高，同时也能确保特征点误匹配消除的准确性。

4 方法细述

4.1 初始匹配点对获取

获取大量特征匹配点对最常见的方法是 SIFT^[11], SURF^[12]以及 ORB^[13], 鉴于 SIFT 是一种具有尺度不变性的算法, 同时也对噪声、视角变化和光照变化具有良好的鲁棒性, 因此本文选择了 SIFT 获取初始的特征匹配点对。

SIFT 在多个尺度空间上获取特征点, 接着对特征点进行描述子的构建。具体来说, SIFT 将周围区域划分成多个子区域, 计算每个子区域内每个像素的梯度量值大小与方向, 并以此加入此子区域的直方图中, 构造了一个 128 维的描述子。基于描述符的特征点匹配在本质上来说是对描述子的比较, 找到最相似的两个特征点描述子。衡量两个描述子具有多大差异的标准是描述子之间的距离, 通过距离将描述子之间的差异转换成一个单一的数字, 这样用它使得相似性度量变得非常简单。对于 SIFT 来说, 使用的是欧式距离, 如式(1)所示。

$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

通过欧式距离并结合快速最近邻原则 Fast Library For Approximate Nearest Neighbors(FLANN)进行匹配^[32]。在匹配阶段时, 会根据描述子距离选择出两个最近的特征点, 如果满足次近邻与最近邻之比在一定的范围内(一般取 0.6~0.75), 那么则是一对正确的匹配, 否则被抛弃。通过最近邻当图像出现仿射变换、重复等问题时, 该最近邻方法则会受到一定的干扰, 此时会出现比较多的误匹配。为此, 本文接下来的步骤就是通过构建三角拓扑结构来获得相似性得并提高误匹配消除的误差的。

4.2 三角拓扑结构构建

在初始特征点匹配完成后, 能够获得大量的特征匹配点对, 需要对每个特征点构建其局部三角拓扑结构。三角结构在力学中是一种非常稳定的结构, 其受力均匀和不易损坏, 而我们在此处引入三角拓扑结构也是因为其具有稳定性。无论是经过普通的旋转平移变化, 还是较为复杂的仿射变化, 甚至是透视变换, 三角结构仍然是三角结构, 构成的三角拓扑仍然是保持不变。

对于图像中一个特征点, 构建三角拓扑结构需要遵循一定的规则。首先是距离最短规则,

三角结构需要通过最近的三点形成三角形；其次是不相交规则，各线段即三角形的边皆不相交；第三是不重复性规则，由于匹配点集中可能会出现一些重复的特征点，因此需要将重复的特征点去除；最后是全局不变性规则，即当新增、删除、移动某一个顶点时只会影响临近的三角形，不会影响整体结构的改变。我们使用 Delaunay 三角剖分对图像中的特征点构建三角形拓扑网络，Delaunay 三角剖分可以很好地满足以上规则。对图像构建三角拓扑网络结构展示如图 2 所示。

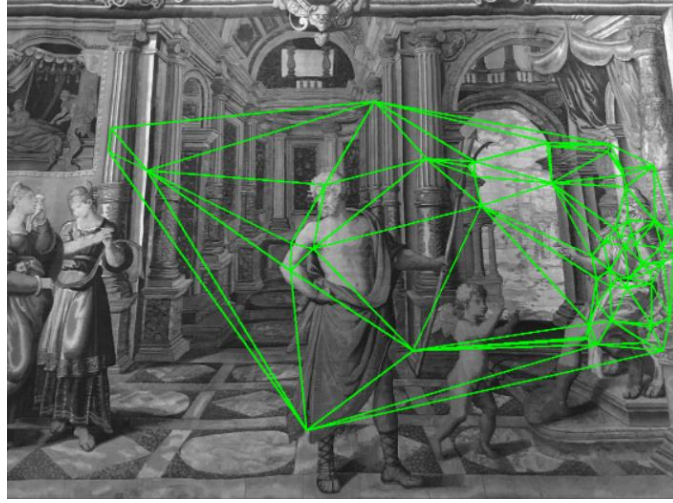


图 2 图像上基于特征点的三角拓扑网络结构

点集 $P\{p_1, p_2, \dots, p_n\}$ ， $Q\{q_1, q_2, \dots, q_n\}$ 分别表示两幅不同图像中获得的匹配点对中的 n 个匹配点，通过对两幅图像中的特征点进行三角拓扑结构的构建，可以得到三角拓扑集合 T 和 T' ， T 和 T' 的表达如式(2)所示。

$$\{T, T'\} = \left\{ \sum_{n=1}^{n=S} t_x [p_i, p_j, p_k], \sum_{n=1}^{n=S'} t_y [q_i, q_j, q_k] \right\} \quad (2)$$

其中， T, T' 分别表示点集 P 和点集 Q 构成的拓扑结构， S, S' 分别表示拓扑结构 T, T' 的数量， $t_x [p_i, p_j, p_k]$ 表示第 x 个拓扑结构，该拓扑结构由第 i, j, k 三个点 $\{p_i, p_j, p_k\}$ 构成， t_y 分别表示第 y 个拓扑结构，由第 i', j', k' 三个点 $\{q_i, q_j, q_k\}$ 构成。

不论经过仿射变化还是透视变换，正确的匹配点与其周围的特征点组成局部三角结构会比较稳定，虽然局部三角拓扑形状会发生变化，位置也可能变化，但是局部范围内仍然由对应的匹配点组成三角拓扑结构不变。如图 3 所示，蓝色节点表示特征点，蓝色连线表示节点与节点之间的边长，红色连线表示对应的匹配点对，对于一对正确的匹配点 A, A' ，节点与

节点的之间的边长长度以及方向会出现变化，然而拓扑结构不会变化。点 A 周围的拓扑结构由 B 、 C 、 D 、 E 、 F 构成，点 A' 周围的拓扑点为 B' 、 C' 、 D' 、 E' 、 F' ，点 A 与点 A' 周围的拓扑点是相对应的匹配点对，因此 A, A' 组成的三角拓扑结构是相似的。当一对误匹配点 A, A' 周围包围着较多的误匹配时，其对应的局部三角拓扑结构会比较紊乱，主要体现在周围的三角拓扑结构会由别的匹配点构成，如图 4 所示，点 A 周围的拓扑结构由 B 、 C 、 D 、 E 、 F 构成，点 A' 周围的拓扑点为 M' 、 N' 、 R' 、 L' 、 F' ，点 A 与点 A' 周围的拓扑点绝大多数或者全部不是对应的匹配点对，因此 A, A' 组成的三角拓扑结构是不相似的。

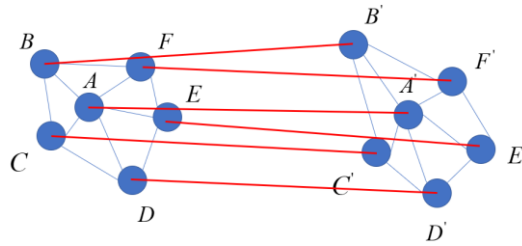


图 3 正确匹配对应的三角拓扑结构图

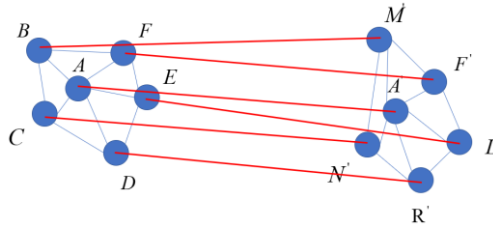


图 4 错误匹配对应的三角拓扑结构图

通过以上分析，可以得出一个结论，在图像对中正确的匹配点周围的局部三角拓扑结构是相似的。基于这个规律，我们判断一对匹配点正确与否，可通过判断其周围的三角拓扑结构是否相似来实现。一般来说，对于一对匹配点，衡量其局部领域内的三角拓扑结构是否相似需要衡量节点的相似性、节点与节点之间边的长度、节点与边的关系等。然而这样衡量过于复杂。为此，我们仅仅利用三角拓扑相似性，不仅可以确保精度，而且可大幅简化计算相似性的复杂度。

对于一对匹配点 A, A' ，只需分别计算点 A 与点 A' 周围有哪些点与其构成三角拓扑结构，因此我们可以将局部的三角拓扑结构进一步简化。如图 5 所示，对于点 A ，其周围的三角拓扑结构包括 ABC 、 ACD 、 ADE 、 AEF 、 AFB ，点 A' 周围的三角拓扑结构包括 $A'B'C'$ 、 $A'C'D'$ 、 $A'D'E'$ 、 $A'E'F'$ 、 $A'F'B'$ ，将其简化后，模型变为点 A 的邻居拓扑节点集

合为 $NO_A\{B,C,D,E,F\}$ ，点 A' 的邻居拓扑节点集合为 $NO_{A'}\{B',C',D',E',F'\}$ ，对于所有节点的拓扑节点集合 NO 为表示如式错误!未找到引用源。所示。

$$NO(P,Q) = \left(\sum_{i=1}^N NO_{P_i} \left(\sum_{m=1} p_m \right), \sum_{j=1}^N NO_{Q_j} \left(\sum_{n=1} q_n \right) \right) \quad (3)$$

其中， NO_{P_i} 表示节点集合 P 的第 i 个节点的邻居拓扑节点集合， NO_{Q_j} 表示节点集合 Q 的第 j 个节点的邻居拓扑节点集合， N 表示匹配点对的个数， p_m 表示邻居拓扑节点集合 NO_{P_i} 中的第 m 个节点， q_n 表示邻居拓扑节点集合 NO_{Q_j} 中的第 n 个节点。

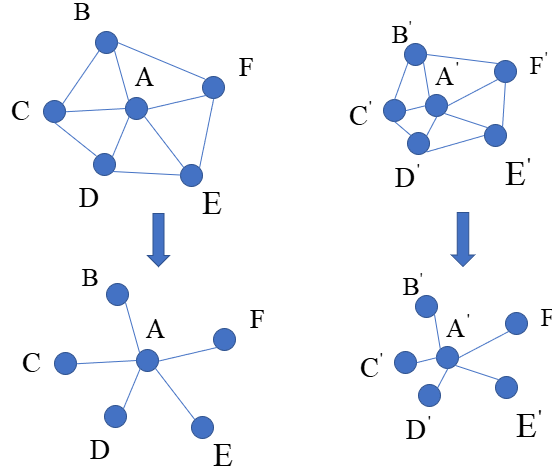


图 5 三角拓扑结构简化图

4.3 相似性计算与重采样

对于 NO_A ，其邻居拓扑节点为 $B、C、D、E、F$ ， $NO_{A'}$ 的邻居拓扑节点为 $B'、C'、D'、E'、F'$ ，其中 $B、C、D、E、F$ 与 $B'、C'、D'、E'、F'$ 共有五对匹配点，那么 NO_A 与 $NO_{A'}$ 的相似节点数量为 5，不相似节点数量为 0。换言之，对于任意一对匹配点 i,j 的两个邻居拓扑点集合 NO_i 与 NO_j ，如果 NO_i 中的一个元素与 NO_j 中的某个元素是一对匹配点，那么两者的拓扑节点相似节点数量 r 加 1，否则不相似节点数量 f 加 1。任意的匹配点对 i,j ，其相似节点数量 r 与不相似节点数量 f 的计算公式如式(4)所示。

$$(r, f) = \begin{cases} \sum_{m=1} r = r + 1, NO_i(m) \in NO_j \\ \sum_{m=1} f = f + 1, NO_i(m) \notin NO_j \end{cases} \quad (4)$$

其中， $NO_i(m)$ 表示节点 i 的邻居拓扑节点集合中的一个点。

仅仅使用相似节点数量评价三角拓扑相似性是片面的，对于一对拥有众多邻居拓扑点的匹配点对来说，其相似节点数量与不相似节点数量可能都很高，该匹配点对不一定是正确的匹配点对，而由于其相似节点数量高，可能会被选择进入后续的局部重采样环节中；对于一对邻居拓扑点较少的正确匹配点对来说，其相似节点数量可能会比错误匹配点对的相似节点数量少，这可能会使得错误匹配点对进入后续的局部重采样环节。以上情况可能会使得重采样得到的模型出现较大误差，导致最后结果中保留较多误匹配，因此我们需要综合相似节点数量与不相似节点数量进行考虑。对于一对匹配点对而言，不相似节点数量越多，该匹配点对为错误匹配的概率越高，因此针对不相似节点数量，需要进行一些惩罚。不相似节点数量越多，惩罚越重。在获得一对匹配点的拓扑邻居节点的相似节点数量与不相似节点数量后，则可以计算该匹配点对的相似性得分。匹配点对 A, A' 的相似性得分 $s_{(A, A')}$ 的计算公式如式(5)所示。

$$s_{(A, A')} = r - \lambda * (f + \mu) \quad (5)$$

其中， r 表示节点 A 的拓扑邻居节点与节点 A' 的拓扑邻居节点的相似节点数量， f 表示不相似节点数量， λ 表示相似节点数量与不相似节点数量的相减系数， μ 表示惩罚项，每当不相似节点数量增加 1，则惩罚项也相应增加 1。

最后就是相似性得分的重采样。计算得到每对匹配点的邻居拓扑结构的相似性得分后，可以得到相似性得分集合 S ，如式(5)所示。

$$S = enum\{s_{(i, j)}, i \in P, j \in Q\} \quad (6)$$

其中， i 表示节点集合 P 的一个节点， j 表示节点集合 Q 的一个节点， (i, j) 表示是一对匹配点对。

获取到所有匹配点对的相似性得分集合后，对得分进行排序，接着获取相似性得分最高的 e 对匹配点，目的在于通过这部分局部最优的匹配点对进行一个局部最优的单应性矩阵或者基本矩阵的求解，这与 RANSAC 从所有匹配点对中随机抽样进行最优单应矩阵或者基础矩阵的做法完全不同。对于获得的 e 对匹配点，每次随机从其中获取四组点对进行求解矩阵。

通过计算重投影误差，我们可以多次迭代得到两幅图像的最优单应性矩阵。单应性矩阵 H 可以表示如式(7)所示。

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \quad (7)$$

其中， h_{ij} 表示单应矩阵 H 中第 i 行第 j 列的元素。

对于一对匹配点对 $X(x, y)$ 和 $X'(x', y')$ ，通过投影变换可以建立关联的表达式，表达式如式(8)所示。

$$X = HX' \quad (8)$$

对于式(8)也可以表示为式(9)所示。

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \quad (9)$$

对于式(9)，展开后变为式(10)和式(11)所示。

$$x = \frac{h_{11}x' + h_{12}y' + h_{13}}{h_{31}x' + h_{32}y' + 1} \quad (10)$$

$$y = \frac{h_{21}x' + h_{22}y' + h_{23}}{h_{31}x' + h_{32}y' + 1} \quad (11)$$

最后，对于每一对匹配点进行投影变换，在误差范围内的点即为内点，内点即为我们认为的正确匹配。相反，重投影误差较大的匹配，即模型的异常值，是需要去除的误匹配。

5 实验结果与分析

我们在本节中进行了实验来评估本文方法与最经典的方法、以及最新最先进方法的对比。使用的数据集来自于公开数据集 Hpatches^[33]和 Mikolajczyk VGG^[34]，两种数据集包含明显的光照变化场景、巨大的仿射角度变化场景、巨大的旋转角度变化场景、模糊场景、尺度变化场景、以及前五者的综合场景，这样场景能够很好地模拟大多数自然场景，这样能更有效地验证本文方法的有效性。被对比的方法包括 RANSAC、PROSAC、GMS、MAGSAC、GTM、WGTM 和 LGSC，这些方法包括最优秀的基于几何的方法和基于重采样的方法。

实验在 Windows 10 操作系统的电脑上进行，配备英特尔酷睿 i5-1150K 处理器和 8gb

RAM。算法实现都采用 Python 编写，工具包括公开的计算机视觉库 Opencv-python 和高级科学计算库 scipy。

5.1 评价标准选择

由于公开数据集提供了对应图片的单应性矩阵，所以这里我们选择召回率(recall)、准确率(precision)和 F-score 三个公认的评价指标^[35]对实验结果进行评估。

召回率的定义是通过误匹配消除算法获得的正确匹配数占初始匹配点对中的所有正确结果的比例，准确率是通过误匹配消除算法获得的正确匹配数占获得的所有匹配数的比例，而 F-score 是召回率和准确率的调和值。召回率、准确率和 F-score 的计算分别如式(12)、式(13)和式(14)所示。

$$recall = \frac{a}{b} \quad (12)$$

$$precision = \frac{a}{c} \quad (13)$$

$$F - score = \frac{2 * a * b}{a + b} \quad (14)$$

其中， a 表示经过误匹配算法消除误匹配后保留的正确匹配数， b 表示经过误匹配算法消除误匹配之前的正确匹配数， c 经过误匹配算法消除误匹配后保留的所有匹配数。

5.2 参数变化实验

本小节将展示关键参数对特征提取和匹配效果的贡献，以及参数值变动时所带来的影响，从而方便我们找到效果最佳的参数值。在本文方法中，有一个重要参数，该参数的作用是挑选 e 对匹配点对，从 e 对匹配点对中获得最佳的单应性矩阵。为了能够获得这个参数的最佳值，使用不同的 e 值在公开数据集 Hpatches 与 Mikolajczyk VGG 进行实验，实验结果图如图 6 所示。

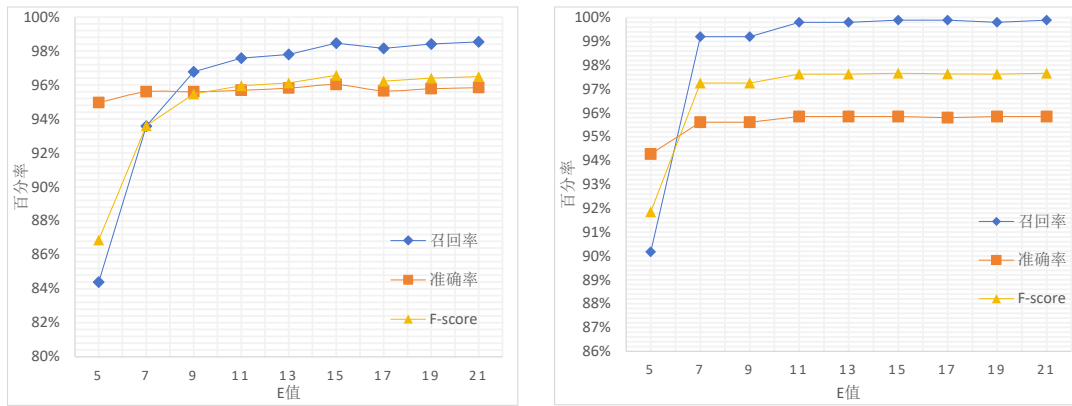


图 6 左图为数据集 Hpatches 中不同参数下的召回率、准确率和 F-score，右图为数据集 Mikolajczyk VGG 中不同参数的召回率、准确率和 F-score

从图 6 中可以清晰看出当 e 值范围在 11~15 时在数据集 Hpatches 和 Mikolajczyk VGG 中均能获得最高的召回率、准确率和 F-score，因此参数 e 的最佳范围为 11~15。

5.3 实验结果

本小结将展示本文方法在公开数据集 Hpatches 以及 Mikolajczyk VGG 上的实验结果，我们设置 e 值为 15，分别如图 7 与图 8 所示。

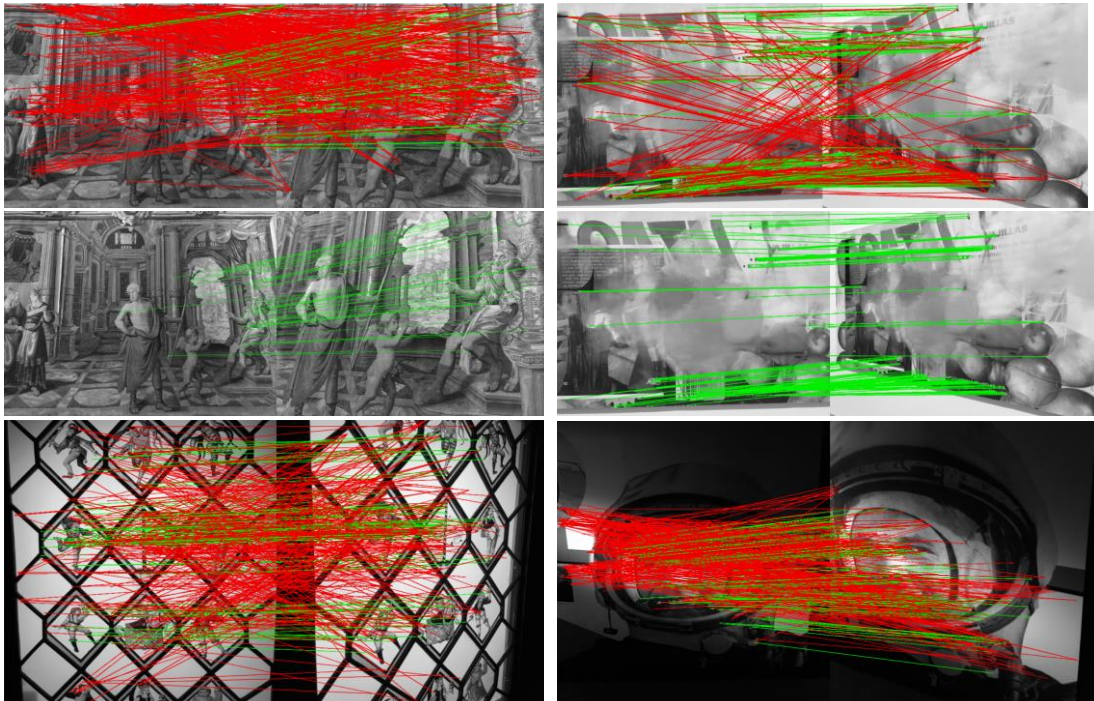




图 7 Hpatches 数据集部分图像匹配效果图，其中红色的匹配表示错误的匹配，绿色的匹配表示正确的匹配

从图 7 可以清晰看出，Hpatches 数据集中两幅图像存在较大的旋转，并且带有一定的尺度变化，因此在两幅图像中获得的初始匹配点对中包含大量的误匹配点，其比例高于正确匹配，但是通本文方法能够很好地将误匹配消除，获得较好的匹配结果。

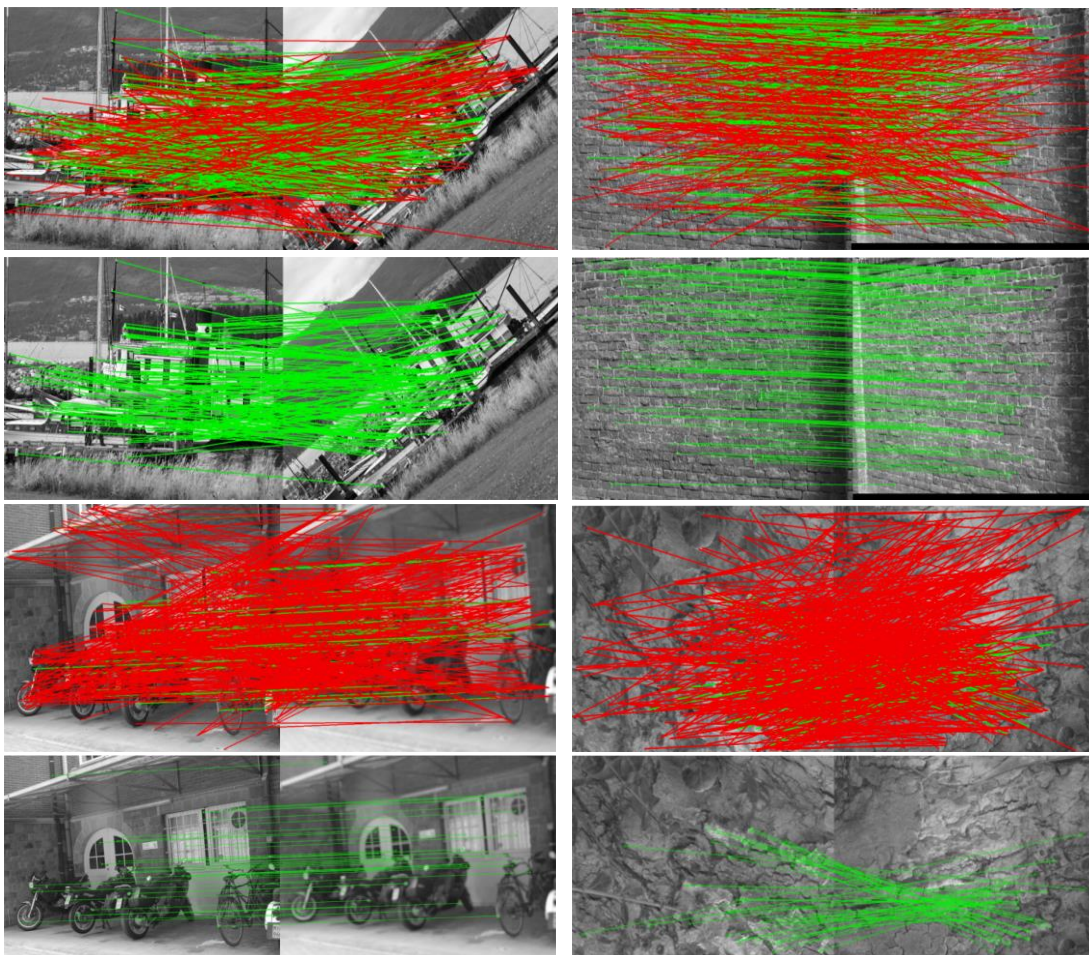


图 8 Mikolajczyk VGG 数据集部分图片匹配效果图，其中红色的匹配表示错误的匹配，绿色的匹配表示正确的匹配

由于 Mikolajczyk VGG 数据集中两幅图像之间存在很大的角度旋转，同时伴有一定的仿射变化，或者是很强的光照变化。正因这些问题的存在，Mikolajczyk VGG 数据集中获得的

初始匹配点队中存在大量的误匹配点，其比例高于正确匹配。如图 8 所示，本文方法也能够很好地过滤误匹配点，留下正确匹配点对。

综合在 Hpatches 与 Mikolajczyk VGG 两个数据集中的实验结果表现，可以看出本文方法能够在存在巨大仿射角度、巨大旋转角度以及明显的光照变化场景中，都发挥作用。

5.4 方法对比分析

本小节对常见的几种经典方法、以及目前最好的方法进行公开数据集 Hpatches 和 Mikolajczyk VGG 的实验对比分析，以便展现本文方法的优势。选择的方法包括 RANSAC, PROSAC, MAGSAC, GTM, WGTM, GMS, LGSC。同样的设置阈值 α 为 15。

(1) Hpatches 数据集

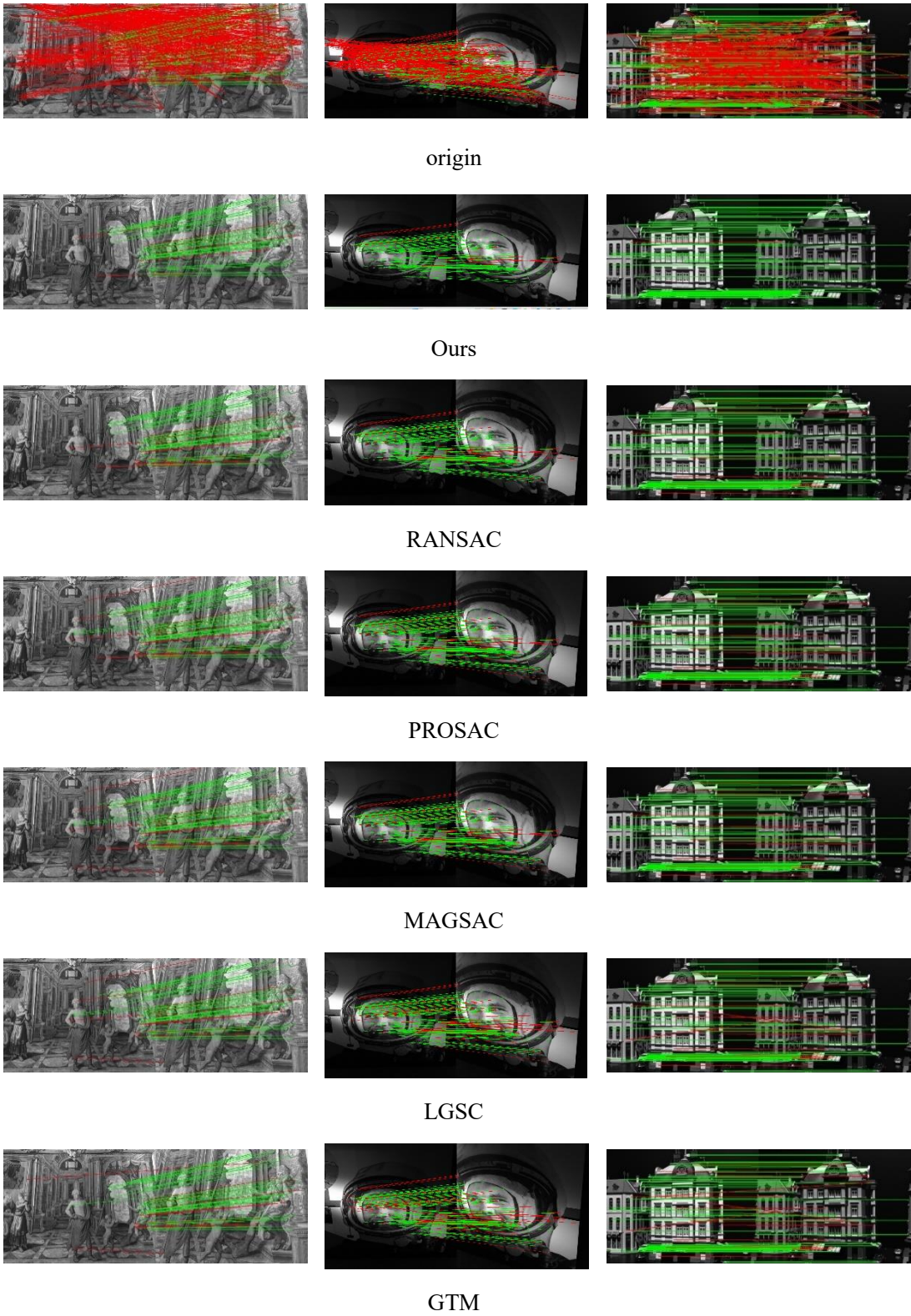
从 Hpatches 数据集中选择角度变化明显，光照变化较强的数据进行实验，这些数据内点占比很少，其参考性更强。实验结果如表 1 所示。

表 1 各种方法在 Hpatches 中的表现

	Recall	Precision	F-score
RANSAC	98.26%	93.48%	95.04%
PROSAC	<u>99.16%</u>	93.01%	95.05%
MAGSAC	<u>99.71%</u>	93.02%	95.50%
GTM	82.09%	87.70%	83.13%
WGTM	80.01%	92.11%	85.19%
GMS	36.57%	79.89%	46.77%
LGSC	96.71%	89.86%	92.32%
Ours	98.46%	96.02%	96.56%

从表 1 可以看出，本文方法的准确率比 RANSAC 高出 2.54%，比 PROSAC 与 MAGSAC 高 3%，比 GMS 高出 16.13%，相比最新算法 LGSC 高出 7%。在更为综合的 F-score 这一指标上，本文方法比 RANSAC, PROSAC, MAGSAC 都高出 1%以上，比 GMS 高出将近 50%，对比 LGSC 高出 4%以上。在召回率方面，本文方法仅次于 MAGSAC 和 PROSAC，比 GMS 高出 60%以上，比 LGSC 高出 2%。综合来说，本文方法相对于 RANSAC、PROSAC、GMS 等方法，在内点率较低的匹配点对中能够获得了更好的表现，在准确率和 F 值两种指标上获

得双高的表现。为了能够更直观地看出几种算法的效果对比，选择了从 Hpatches 中的实验效果展示如图 9 所示。



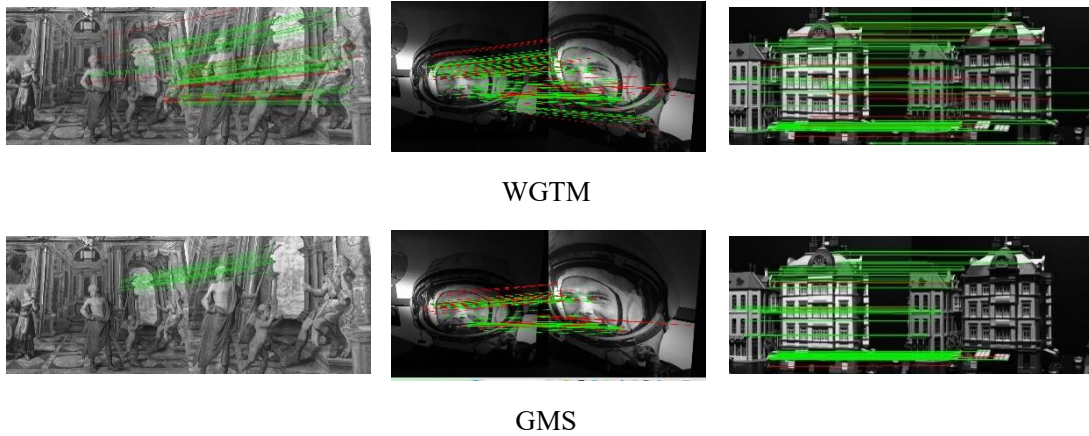


图 9 不同的方法在原匹配图基础上消除误匹配的结果图

(2) Mikolajczyk VGG 数据集

对 Mikolajczyk VGG 数据集进行实验得到的结果如表 2 所示。

表 2 各方法在 Mikolajczyk VGG 中的表现

	Recall	Precision	F-score
RANSAC	98.19%	94.11%	95.94%
PROSAC	<u>100.00%</u>	94.03%	96.77%
MAGSAC	<u>100.00%</u>	94.03%	96.77%
GTM	83.37%	91.83%	87.18%
WGTM	82.60%	94.00%	87.69%
GMS	38.69%	78.47%	49.72%
LGSC	98.28%	92.40%	95.12%
Ours	99.89%	95.85%	97.67%

从表 2 可以看出，在 Mikolajczyk VGG 数据集中，本文的方法仍然取得了最好的表现综合表现。本文方法的准确率比 RANSAC，PROSAC，MAGSAC 高出 1.7% 以上，比 GMS 高出 17% 以上，相比于 LGSC 高出 3.4%。在 F-score 这一指标上，本文方法仍然比 RANSAC，PROSAC，MAGSAC 高，比 GMS 高出将近 50%，对比 LGSC 高出 2% 以上。对于召回率，本文方法仅次于 MAGSAC 和 PROSAC，比 GMS 高出 60% 以上，比 LGSC 高出 1.6%。本文方法在 Mikolajczyk VGG 数据集中，准确率和 F 值两种指标仍然达到了最高。

6 结论

本文提出了一种更加鲁棒的特征点误匹配去除方法——LTrSAC，其将基于几何的方法与基于重采样的方法相融合，不仅仅避免了从全局内计算最佳的矩阵低效率，也能在存在大量误匹配的数据中获得更好的结果，特别是有效降低基于几何方法容易受局部范围内较多误匹配的干扰问题。

参考文献

- [1] Ma J, Jiang X, Fan A, et al. Image Matching from Handcrafted to Deep Features: A Survey [J]. International Journal of Computer Vision, 2020, 129(1): 23-79.
- [2] Snavely N, Seitz S M, Szeliski R. Photo tourism: exploring photo collections in 3D [M]. ACM siggraph 2006 papers. 2006: 835-46.
- [3] Fraundorfer F, Scaramuzza D. Visual odometry: Part ii: Matching, robustness, optimization, and applications [J]. IEEE Robotics & Automation Magazine, 2012, 19(2): 78-90.
- [4] Liu Y, Zhang D, Lu G, et al. A survey of content-based image retrieval with high-level semantics [J]. Pattern recognition, 2007, 40(1): 262-82.
- [5] Xie R, Xia M, Yao J, et al. Guided color consistency optimization for image mosaicking [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 135: 43-59.
- [6] Nebiker S, Lack N, Deuber M. Building change detection from historical aerial photographs using dense image matching and object-based image analysis [J]. Remote Sensing, 2014, 6(9): 8310-36.
- [7] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-63.
- [8] Goesele M, Snavely N, Curless B, et al. Multi-view stereo for community photo collections [C]. In: Proceedings of the 2007 IEEE 11th International Conference on Computer Vision, 2007. 1-8.
- [9] Wang Y, Hao W, Wang G, et al. A method of realistic leaves modeling based on point cloud [C]. In: Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, 2013. 123-130.
- [10] Wang Y, Zhao Y, Wang N, et al. A hole repairing method based on edge-preserving projection [C]. In: Proceedings of the E-Learning and Games: 12th International Conference, Springer, 2019. 115-122.
- [11] Lowe D. Distinctive image features from scale-invariant key points [J]. International Journal of Computer Vision, 2003, 20: 91-110.
- [12] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-59.
- [13] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C]. In: Proceedings of the IEEE International Conference on Computer Vision, 2011. 2564-2571.
- [14] Leutenegger S, Chli M, Siegwart R Y. BRISK: Binary robust invariant scalable keypoints [C]. In: proceedings of the 2011 International Conference on Computer Vision. IEEE, 2011. 2548-2555.
- [15] Alahi A, Ortiz R, Vandergheynst P. Freak: Fast retina keypoint [C]. In: proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012. 510-517.

- [16]Yi K M, Trulls E, Lepetit V, et al. Lift: Learned invariant feature transform [C]. In: Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Springer, 2016. 467-483.
- [17]Detone D, Malisiewicz T, Rabinovich A. Superpoint: Self-supervised interest point detection and description [C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018. 224-236.
- [18]Jiang X, Xia Y, Zhang X-P, et al. Robust image matching via local graph structure consensus [J]. Pattern Recognition, 2022, 126: 108588.
- [19]Zhao W, Tian Z, Yang L, et al. Image registration using a kernel partial least squares based mismatches removal method [J]. AEU-International Journal of Electronics and Communications, 2016, 70(4): 427-35.
- [20]Wang L, Niu Z, Wu C, et al. A robust multisource image automatic registration system based on the SIFT descriptor [J]. International Journal of Remote Sensing, 2012, 33(12): 3850-69.
- [21]Maier J, Humenberger M, Murschitz M, et al. Guided matching based on statistical optical flow for fast and robust correspondence analysis [C]. In: Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Springer, 2016. 101-117.
- [22]Niu H, Lu Q, Wang C. Color correction based on histogram matching and polynomial regression for image stitching [C]. In: Proceedings of the 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), IEEE, 2018. 257-261.
- [23]Aguilar W, Frauel Y, Escolano F, et al. A robust graph transformation matching for non-rigid registration [J]. Image and Vision Computing, 2009, 27(7): 897-910.
- [24]Izadi M, Saeedi P. Robust weighted graph transformation matching for rigid and nonrigid image registration [J]. IEEE Transactions on Image Processing, 2012, 21(10): 4369-82.
- [25]Liu L, Huang Z, Lan S, et al. Research on ORB mismatch elimination algorithm based on fusion of gradient and cosine similarity [C]. In: Proceedings of the 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT), IEEE, 2021. 519-523.
- [26]Bian J, Lin W-Y, Matsushita Y, et al. Gms: Grid-based motion statistics for fast, ultra-robust feature correspondence [C]. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. 4181-4190.
- [27]Fischler M A, Bolles R C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography [J]. Communications of the ACM, 1981, 24(6): 381-395.
- [28]Torr P, Zisserman A. MLESAC: A New Robust Estimator with Application to Estimating Image Geometry [J]. Computer Vision and Image Understanding, 2000, 78(1): 138-56.
- [29]Torr P H, Nasuto S J, Bishop J M. Napsac: High noise, high dimensional robust estimation-it's in the bag [C]. In: Proceedings of the British Machine Vision Conference (BMVC), 2002. 2: 3
- [30]Chum O, Matas J. Matching with PROSAC-progressive sample consensus [C]. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), IEEE, 2005. 1: 220-226.
- [31]Barath D, Matas J, Neskova J. MAGSAC: marginalizing sample consensus [C]. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 10197-10205.
- [32]Muja M, Lowe D G. Fast approximate nearest neighbors with automatic algorithm configuration [J]. VISAPP (1), 2009, 2(331-340): 2.
- [33]Balntas V, Lenc K, Vedaldi A, et al. HPatches: A benchmark and evaluation of handcrafted and learned local descriptors [C]. In: Proceedings of the IEEE Conference on Computer Vision and

Pattern Recognition, 2017. 5173-5182.

[34]Mikolajczyk K, Tuytelaars T, Schmid C, et al. A comparison of affine region detectors [J]. International Journal of Computer Vision, 2005, 65: 43-72.

[35]Mikolajczyk K, Schmid C. A performance evaluation of local descriptors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-30.