

基于 ORB-BEBLID 的图像快速配准方法研究

赵柏山¹ 郭晋宇¹ 王庆祝²

ZHAO Baishan GUO Jinyu WANG Qingzhu

摘要

针对现有 ORB 算法配准误差较大, 配准率较低且运算时间长以及 RANSAC 算法随机性强且不稳定的问题, 提出一种基于改进 ORB 与改进 MAGSAC++ 结合的快速图像配准方法。首先构建多尺度金字塔特征点, 然后引入 BEBLID 描述符进行特征描述。通过 GMS 算法对所获得特征点进行粗匹配, 最后采用 MAGSAC++ 算法对匹配结果进一步筛选, 从而实现大视差图像快速配准。实验结果表明, 相较于传统 ORB+RANSAC 算法和 SIFT+RANSAC 算法, 特征点数量明显增加, 且匹配准确率分别提升了 24.45% 和 34.3%, 在图像尺度旋转, 光照等多种变换场景中均能保持良好的鲁棒性。

关键词

特征匹配; ORB 算法; MAGSAC++ 算法; 图像配准

doi: 10.3969/j.issn.1672-9528.2025.06.009

0 引言

随着计算机技术的发展, 计算机视觉技术得到了飞速发展, 图像配准技术是近几年发展迅速的图像处理技术^[1], 目前已经广泛应用在医学影像、全景地图、遥感技术、智能驾驶等领域。在图像配准研究领域, 目前的核心问题是如何平衡图像配准精度和计算效率。主要有两种图像配准方式: 基于图像区域的配准方法和基于图像特征的配准方法。基于

图像特征的配准方法相对于其他非特征的方法有较强的鲁棒性。

基于图像特征的配准方法主要有 SIFT (scale-invariant feature transform) 算法、SURF (speeded up robust feature) 算法以及 ORB (oriented fast and rotated brief) 算法等。SIFT 算法采用梯度直方图进行描述, 精度较高但计算量较大, 耗时长。ORB 算法结合并改进了 FAST 特征点检测和 BRIEF 特征描述, 在处理速度上得到显著的提升, 匹配准确率有一定下降。文献 [1] 简化了图像金字塔, 采用梯度方向改进了特征点主方向提取模式。文献 [2] 用阈值判断特征点, 并使用自适应范围的非极大值抑制算法提高特征点分布均匀性,

1. 沈阳工业大学信息科学与工程学院 辽宁沈阳 110870
2. 中国人民解放军空军 93057 部队 吉林四平 136400

[3] DACHSBACHER C, STAMMINGER M. Reflective shadow maps[C]//I3D'05: Proceedings of the 2005 symposium on Interactive 3D graphics and games. New York: ACM, 2005: 203-231.

[4] LAINE S, SARANSAARI H, KONTKANEN J, et al. Incremental instant radiosity for real-time indirect illumination[C]//EGSR'07: Proceedings of the 18th Eurographics conference on Rendering Techniques. New York: ACM, 2007: 277-286.

[5] SIMON F, HANIKA J, DACHSBACHER C. Rich - VPLs for improving the versatility of many-light methods[J]. Computer graphics forum. 2015, 34(2): 575-584.

[6] GEORGIEV I, SLUSALLEK P. Simple and robust iterative importance sampling of virtual point lights[EB/OL].[2024-05-12].<https://www.ilijan.com/publications/ImportanceVPL>.

[7] RITSCHER T, EISEMANN E, HA I, et al. Making imperfect

shadow maps view-adaptive: high-quality global illumination in large dynamic scenes[J].Computer graphics forum, 2011, 30(8): 2258-2269.

[8] BABAHENINI D, GRUSON A, CHAOUKI-BABAHENINI M, et al. Efficient inverse transform methods for VPL selection in global illumination[J]. Multimedia tools and applications, 2018, 77: 13571-13595.

[9] 陈胜, 陈纯毅, 邢琦玮, 等. 基于可视区域间接光照聚类的虚拟点光源采样 [J]. 系统仿真学报, 2020, 32(6): 1085-1093.

【作者简介】

吴玉莹 (1999—), 女, 江西抚州人, 硕士研究生, 研究方向: 计算机图形学。

(收稿日期: 2025-02-15 修回日期: 2025-06-05)

并没有提高计算效率。文献 [3] 利用匹配点邻域灰度分布相似性，将特征点显著性与余弦相似度结合，保证配准精度的同时减少了运算量，但是配准点数量较少。为了平衡大视差图像配准精度和计算效率，本文提出一种基于改进多尺度 ORB-BEBLID 特征的大视差图像快速配准方法。首先构造八层金字塔结构并对图像进行下采样获取特征点，然后采用 BEBLID 描述符对特征点二值化描述，利用 GMS 算法对特征点粗匹配，结合改进 MAGSAC++ 算法对特征点精匹配，有效提高图像匹配的准确率的同时减少了运算时间。

1 ORB 算法

ORB 算法是一种快速的特征点提取算法。通过查找图像关键点，并为每个关键点计算相应的特征向量来实现特征点提取，可以分为三个步骤。

1.1 FAST 角点检测

取图像的任一点 P 并以 P 为圆心，3 为半径的圆周上 16 个像素点，确定阈值 h 和 P 点灰度值 I_p ，如果圆圈上有 8 个以上的像素点灰度值小于 I_p-h 或大于 I_p+h ，则将 P 点选为关键点。

1.2 特征点方向

FAST 角点不具有方向性，ORB 算法选择一种简单有效的方法：强度质心法确定关键点方向。定义一个小的图像块的矩为：

$$m_{pq} = \sum_{x,y \in R} x^p y^q I(x,y) \quad p,q \in \{0,1\} \quad (1)$$

则图像块的质心定义为：

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2)$$

连接图像块的几何中心 O 与质心 C ，得到一个方向向量 \overrightarrow{OC} ，特征点主方向为特征点与质心的夹角，于是定义特征点的方向为：

$$\theta = \arctan(m_{01} / m_{10}) \quad (3)$$

1.3 特征点描述

ORB 算法采用 Streer-BRIEF^[4] 描述子，首先利用高斯核对图像进行平滑处理，然后对于给定关键点 x ，在以关键点 x 为中心的高斯分布中抽取一个点 y ，通过比较 x 和 y 的灰度值为关键点构建二进制描述符，随机选取 n 对特征点则变成长度为 n 的二进制描述符为：

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x, y) \quad (4)$$

最后通过两张图片描述子的汉明距离匹配特征点。

2 改进算法

2.1 改进的特征检测算

构建多级图像金字塔：创建一个包含 8 个层级的图像金字塔，以原始图像为基础，经过平滑处理后作为金字塔的最顶层。随后通过对每一层图像应用高斯模糊并进行降采样，

生成金字塔的下一层图像为：

$$I_1(x,y) = I(x,y) * G(x,y, \sqrt{\sigma_0^2 - \sigma_n^2}) \quad (5)$$

高斯卷积核为：

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (6)$$

对当前图像层进行卷积滤波和下采样得到图像金字塔的下一层图像为：

$$I_{n+1}(x,y) = I_n(2x,2y) * G(x,y,\sigma) \quad (7)$$

式中： I_{n+1} 、 I_n 分别为图像金字塔的第 $n+1$ 层和第 n 层图像。

对每一层进行 FAST 特征点检测，根据其 Harris 响应值排序，并进行非极大值抑制，保留响应值较大的点。同时若一个点在不同尺度被检测到，将其保留。对于保留下来的点，通过双线性插值法计算得到其精确位置。

2.2 特征点描述

BEBLID^[5] 描述符使用积分图像来有效地计算一对图像正方形区域的平均灰度值之间的差，结合 BoostedSCC 算法选择一组特征，并结合起来产生一个强描述。

假设一个图像块的训练集为 $\{(x_i, y_i, l_i)\}_{i=1}^M$ ，其中 $x_p, y_i \in X$ ， $l_i \in \{-1, 1\}$ 。其中， $l_i = 1$ 表示两个图像块具有相同的图像特征， $l_i = -1$ 表示不同的图像特征。在训练过程中需要将损失函数降到最低，损失函数定义为：

$$L_{\text{BEBLID}} = \sum_{i=1}^N \exp(-\gamma l_i \sum_{k=1}^K h_k(x) h_k(y)) \quad (8)$$

式中： γ 表示学习率参数； $h_k(z) = h_k(z; f, T)$ 表示第 K 个 WL，依赖于特征提取函数 $f: \rightarrow R$ 和阈值 T 。通过含 T 的 $f(x)$ 来定义 Weak Learner 为：

$$h(x; f, T) = \begin{cases} +1 \cdots f(x) \leq T \\ -1 \cdots f(x) > T \end{cases} \quad (9)$$

特征提取函数定义为：

$$f(x; p_1, p_2, s) = \frac{1}{s^2} \left(\sum_{q \in R(p_1, s)} I(q) - \sum_{r \in R(p_2, s)} I(r) \right) \quad (10)$$

式中： $I(q)$ 和 $I(r)$ 分别表示两个图像像素的灰度值； $R(p, s)$ 以像素 p 为中心且大小为 s 的矩形区域。因此， $f(x)$ 用来计算两个图像矩形框像素点的平均灰度值之间的差值。

通过不断迭代优化 L_{BEBLID} ，最终得到 K 个最佳的 WL 作为采样模板，如图 1 所示。

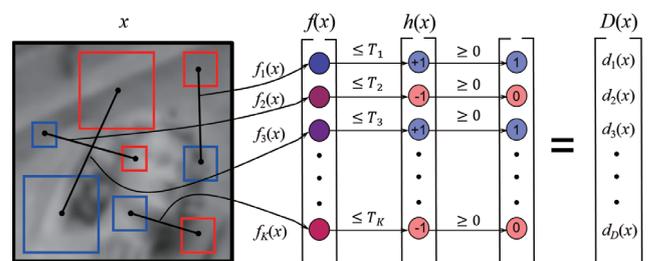


图 1 描述符提取流程

为描述图像块，特征提取函数 $f(x)$ 计算两个对应框中的像素平均灰度值之差。然后，对 $f(x)$ 应用阈值 T ，得到 $h(x)$ 和二值描述符 $D(x) = h(x) \geq 0$ 。最终，为了输出为二值 0 和 1，将 -1 强制转化为 0，+1 强制转化为 1，最终生成高效的二值局部图像描述符 $D(x)$ 。

2.3 GMS 粗匹配

在通过 BEBLID 描述符得到特征点的二值描述符后，接下来需要对特征向量进行匹配，以便更准确地确定图像集合之间的变换关系和具体参数。然而目前特征匹配方法面临的主要挑战在于同时保持较高鲁棒性和较好的实时性。而 GMS^[6] 特征匹配算法可以很好地平衡二者，该算法的主要思想是正确匹配的点会表现出一致的运动平滑性，而错误匹配的点则展现出的是随机的运动特性。因此，可以对每个待匹配的特征点对的邻域内的匹配对数量进行统计以筛选出匹配质量较高的内点。

2.4 MAGSAC++ 精匹配

经典 RANSAC 算法很大程度上依赖于内点和离群点的阈值参数，同时计算效率低下，迭代次数多，如果外点数量多于内点，RANSAC 可能会失败，找不到最合适的模型。针对 RANSAC 方法的内点与外点阈值选取不当导致模型拟合容易出错失败的问题，MAGSAC++^[7] 算法提出了一种新的阈值一致性算法，无需手动设定内点阈值，只需设置阈值的最大可能值。同时提出了一种新的模型质量评估函数和一种新的迭代终止条件，提高了算法的稳定性和鲁棒性。

算法依赖两个假设。首先，假设噪声水平 σ 是一个具有密度函数 $f(\sigma)$ 的随机变量，没有任何先验信息， σ 服从均匀分布。然后，对于给定的 σ ，由服从 n 自由度的卡方分布的 σ 的密度函数决定内点的误差。内点误差定义为：

$$g(r|\sigma) = 2C(p)\sigma^{-p}\exp(-\frac{r^2}{2\sigma^2})r^{p-1} \quad (r < \tau) \quad (11)$$

$$g(r|\sigma) = 0 \quad (r \geq \tau(\sigma)) \quad (12)$$

式中：常数 $C(n) = (2^{2n}\Gamma(\frac{n}{2}))^{-1}$ ； n 为欧几里得空间维度。

此外，MAGSAC++ 在最小二乘法基础上提出了一种迭代重新加权最小二乘法 (IRLS)，其中第 $(i+1)$ 的模型计算参数为：

$$\theta_{i+1} = \arg \min_{\theta} \sum_{p \in P} w(D(\theta_i, p))D^2(\theta, p) \quad (13)$$

其中 P 点权重为：

$$w(D(\theta_i, p)) = \int g(p|\theta_i, \sigma)f(\sigma)d\sigma \quad (14)$$

该算法同时提出一种新的采样方式，渐进式 NAPSAC 采样，该采样方式先在局部采样，随后转向全局采样。假设最

初局部采样点都是内点，如果假设成立，则继续局部采样并逐渐变成 RANSAC 随机采样。

由于局部所有的内点样本模型不太准确，且 GMS 匹配后仍然存在误匹配点，因此对采样方式进行改进。

改进后步骤为：

- (1) 计算出匹配点的汉明距离，按照从大到小排列。
- (2) 抽取前 95% 的点作为新的内点集合。
- (3) 随机从新的内点集合选取一个初始点 P 。
- (4) 找到以 P 为中心， r 为半径的 P 点邻域的点集 S 。
- (5) 若点集 S 中点的数量小于最小采样中值，重复 (1)。
- (6) 点 P 和点集 S 形成最小样本。

3 实验结果对比分析

3.1 实验环境

为验证本文改进算法的实时性和鲁棒性，设置两组图像对传统 ORB、传统 SIFT 和本文算法进行实验测试，其中实验二图像为牛津数据集标准图像。实验在个人计算机上进行，配置为处理器 Intel(R) Core(TM) i7-10875H CPU @ 2.30 GHz，16 GB 内存，64 位 Windows10 操作系统。所用算法开发为 OpenCV 平台，编程环境为 PyCharm，编程语言为 Python。

3.2 实验结果对比分析

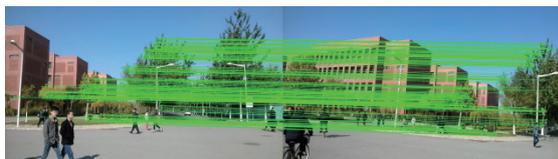
如表 1 所示，在实验 1 中（图 2），ORB 结合 RANSAC 算法的正确匹配点数量为 272 对，准确率为 67.1%；SIFT 结合 RANSAC 算法的特征点分布均匀但数量较少，正确匹配点数量为 355 对，准确率为 61.1%；本文算法的特征点分布均匀且数量显著增加，正确匹配点数量达到 585 对，准确率高达 89.7%。在实验 2 中（图 3），面对有明显旋转和缩放的图像，ORB 结合 RANSAC 算法正确匹配点数量为 187 对，准确率为 63.8%；SIFT 结合 RANSAC 算法的特征点分布均匀但数量较少，正确匹配点数量仅 94 对，准确率为 54.1%；本文算法的特征点数量多、分布均匀，正确匹配点数量为 438 对，准确率高达 94.1%。在运算时间方面，ORB 算法速度最快，其次为本文算法，SIFT 算法运算时间最长。

表 1 实验对比结果

	算法	匹配点数 / 对	准确率 / %	运行时间 / s
实验 1	ORB	272	67.1	0.386
	SIFT	355	61.1	0.678
	本文算法	585	89.7	0.515
实验 2	ORB	187	63.8	0.346
	SIFT	94	54.1	0.501
	本文算法	438	94.1	0.481



(a) ORB+RANSAC 筛选正确匹配点

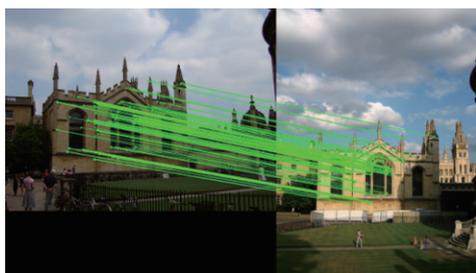


(b) SIFT+RANSAC 筛选正确匹配点



(c) 本文算法特征匹配

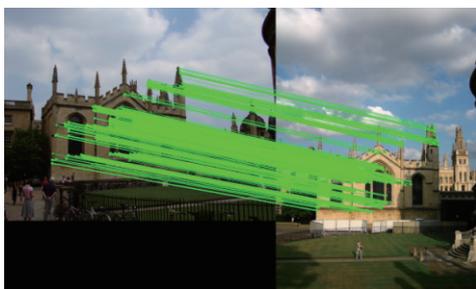
图 2 实验 1



(a) ORB 算法 +RANSAC 特征匹配



(b) SIFT 算法 +RANSAC 特征匹配



(c) 本文算法特征匹配

图 3 实验 2

综合比较, 本文算法特征点提取分布均匀, 数量较多, 正确率高, 同时运算时间适中。

4 结论

为满足图像配准高鲁棒性和实时性需求, 针对传统 ORB 算法精度不高, RANSAC 算法计算效率低下, 容易出现模型拟合失败的问题, 提出一种改进算法, 相对于原 ORB 算法本文算法构造多层金字塔进行特征点提取, 采用 BEBLID 描述符对特征点二值化描述, 同时利用 GMS 算法进行粗匹配并采用改进后 MAGSAC++ 算法对图像特征点进行快速提取和配准。经过实验测试, 在有明显旋转和缩放的图像匹配中, 本文算法较准确度较高, 且符合实时性要求, 有较高的性能, 可用于后续图像实时拼接中。

参考文献:

- [1] 樊彦国, 柴江龙, 许明明, 等. 基于 ORB 与 RANSAC 融合改进的图像配准 [J]. 光学精密工程, 2019, 27(3): 702-717.
- [2] 王险峰, 丁子琳, 邱祖泽, 等. 基于改进 ORB 算法的特征检测与匹配研究 [J]. 计算机与数字工程, 2024, 52(6): 1854-18589.
- [3] 尚明珠, 王克朝. 基于改进多尺度 ORB 的图像配准算法 [J]. 电光与控制, 2024, 31(10): 42-46.
- [4] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF: binary robust independent elementary features[C]//Computer Vision – ECCV 2010. Berlin: Springer, 2010: 778-792.
- [5] SUAREZ I, SFEIR G, BUENAPOSADA J M, et al. BEBLID: boosted efficient binary local image descriptor[J]. Pattern recognition letters, 2020, 133: 366-372.
- [6] BIAN J W, LIN W Y, MATSUSHITA Y, et al. GMS: grid-based motion statistics for fast, Ultra-robust feature correspondence[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 2828-2837.
- [7] BARÁTH D, NOSKOVA J, IVASHECHKIN M, et al. MAGSAC++, a fast, reliable and accurate robust estimator[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1301-1309.

【作者简介】

赵柏山 (1980—), 男, 辽宁铁岭人, 博士, 副教授, 研究方向: 信号处理与保密通信。

郭晋宇 (1998—), 男, 山西朔州人, 硕士研究生, 研究方向: 信号处理。

王庆祝 (1991—), 男, 黑龙江五大连池人, 硕士, 工程师, 研究方向: 信号处理与保密通信。

(收稿日期: 2024-12-19 修回日期: 2025-06-04)