

RBFD: 一种鲁棒的图像局部二值特征描述子

耿利川^{1,2)}, 成 运³⁾, 苏松志^{1,4)}, 林贤明^{1,4)}, 李绍滋^{1,4)*}

¹⁾ (厦门大学信息科学与技术学院 厦门 361005)

²⁾ (许昌学院城乡规划与园林学院 许昌 461000)

³⁾ (湖南人文科技学院通信与控制工程系 娄底 417000)

⁴⁾ (福建省仿脑智能系统重点实验室(厦门大学) 厦门 361005)

(szlig@xmu.edu.cn)

摘 要: 针对传统浮点型特征描述子占用空间大、匹配速度慢的问题, 提出一种基于梯度统计信息比较的局部二值特征描述子. 通过对比特征点邻域梯度统计信息生成二值特征描述子, 再利用多邻域和多分块策略提高描述子判别力, 最后通过近似简化的 AdaBoost 算法实现描述子降维. 实验结果表明, 与已有描述子相比, 文中提出的描述子在实现快速生成的同时其鲁棒性更强.

关键词: 局部特征; 二值特征描述子; 图像匹配
中图法分类号: TP391

RBFD: a Robust Image Local Binary Feature Descriptor

Geng Lichuan^{1,2)}, Cheng Yun³⁾, Su Songzhi^{1,4)}, Lin Xianming^{1,4)}, and Li Shaozi^{1,4)*}

¹⁾ (School of Information Science and Technology, Xiamen University, Xiamen 361005)

²⁾ (School of Urban Planning & Landscaping, Xuchang University, Xuchang 461000)

³⁾ (Department of Communication and Control Engineering, Hunan University of Humanities, Science and Technology, Loudi 417000)

⁴⁾ (Fujian Key Laboratory of the Brain-like Intelligent Systems (Xiamen University), Xiamen 361005)

Abstract: The traditional floating feature descriptors are in high memory load and slow in matching. To best address these problems, this paper proposed a novel binary feature descriptor based on gradient statistic information comparison. Firstly, the image patch around the keypoint is divided into sub-regions, and our binary descriptor is constructed by comparing the gradient statistic information of these sub-regions. Then, a multi-gridding and multi-support region strategy is applied to boost the discrimination of our descriptor. Finally, a simplified AdaBoost algorithm is applied to realize the descriptor dimension reduction. The experimental results show that our descriptor is both efficient in construction and robust to compare with the state-of-the-art methods.

Key words: local feature; binary feature descriptor; image matching

收稿日期: 2014-03-17; 修回日期: 2014-07-24. 基金项目: 国家自然科学基金(61373076, 61202143); 厦门大学中央高校项目(2013121026, 2011121052); 厦门大学 985 平台建设费项目; 高等学校博士学科点专项科研基金(201101211120024); 福建省自然科学基金(2013J05100, 2010J01345, 2011J01367); 湖南省自然科学基金(12JJ2040); 湖南省教育厅科研项目(09A046); 厦门市科技重点项目(3502Z20123017); 深圳市战略性新兴产业发展专项基金(JCYJ20120614164600201). 耿利川(1982—), 男, 博士研究生, 主要研究方向为遥感图像处理、计算机视觉等; 成 运(1966—), 男, 博士, 教授, 主要研究方向为视频编码算法、数字图像处理等; 苏松志(1982—), 男, 博士, 助理教授, 主要研究方向为计算机视觉、模式识别; 林贤明(1980—), 男, 博士研究生, 主要研究方向为行为识别与理解、计算机视觉、模式识别; 李绍滋(1963—), 男, 博士, 教授, 博士生导师, 论文通讯作者, 主要研究方向为计算机视觉、图像处理、计算智能.

图像局部特征描述子是计算机视觉中的一项重要研究内容,它广泛应用于图像匹配、全景图拼接、图像检索、场景三维重建等多个方面^[1].该研究的目的是生成一种紧凑的图像内容描述,使其能够对各种图像变换具备不变性,如图像光照变化、视角变化、旋转、尺度变化等.尺度不变特征变换(scale invariant feature transform, SIFT)^[2]及加速鲁棒特征算法(speed up robust feature, SURF)^[4]由于对图像旋转、尺度缩放等变换具有较高的鲁棒性,已被应用于图像匹配、图像检索等.随着图像数据库越来越大,以及视觉算法在移动设备,如智能手机上的应用的普及,对局部特征描述子的标准也越来越高.要求描述子不仅能够实现对图像局部特征的鲁棒描述,而且对其占用的存储空间及匹配效率提出了更高的要求.SIFT及SURF算法使用的描述子均为浮点型向量,不仅占用存储空间大,而且匹配速度慢.近年来,研究人员提出通过比较图像灰度值来生成二值特征描述子的方法^[5-8].相比浮点型描述子,该类描述子具有占用存储空间少、匹配速度快的特点,因此成为了计算机视觉领域的研究热点.为进一步提高二值特征描述子的鲁棒性,本文提出了一种基于图像梯度信息对比的二值特征描述子(robust binary feature descriptor, Rbfd),并通过实验验证了其有效性.

1 相关工作

特征提取是特征描述子生成的前提,常用的特征提取算子有Moravec算子^[9-10]、Harris算子^[11]、SUSAN算子^[12]、SUSurE算子^[13]、DOG算子^[1]等.Harris算子是一种提出较早且至今仍比较流行的角点提取算法.Mikolajczyk等^[14]比较了几种流行的特征检测算子,指出没有任何一种算子适合于所有图像.之后提出的FAST算子^[15]由于具有提取速度快的特点,而逐渐在一些实时性要求比较高的场合得到应用,如同步定位与建图^[16].AGAST算子^[17]对FAST进行了改进,算法的速度得到进一步提升.

自Lowe提出具有旋转和尺度不变的SIFT算法以后,在图像处理领域引起了人们极大关注,并衍生出了一系列改进算法.针对SIFT算法描述子占用空间大、匹配时间长的问题,Ke等^[18]提出了基于主成分分析的改进SIFT算法.该算法利用主成分分析算法提取特征的局部邻域信息,并生成36

维的特征描述子,提高了描述子匹配速度,但算法性能有所下降,而且由于是在生成原始描述子基础上的降维,算法总体时间复杂度改进不大.Mikolajczyk等^[14]提出了另一种梯度位置与方向直方图算法(gradient location and orientation histogram, GLOH),它将SIFT中4×4棋盘格的分块方式转换为极坐标形式的同心圆分块方式,将图像块分为17个子块,并计算其中的梯度方向直方图(梯度方向分为16个),因此生成了272维的描述子,然后进行主成分分析变换将其降为128维.该描述子在保持跟SIFT描述子相同长度的基础上,提高了算法的鲁棒性,其缺点是算法复杂度比SIFT更高.SURF算法是一种比较成功的改进,该算法利用积分图像和Harr小波的概念生成64维的描述子,大大加快了算法的速度,其不足之处是旋转不变性较SIFT算法有所降低.

尽管上述算法在特征描述及匹配性能上获得了很好的效果,但随着图像数据量越来越大,对算法的性能及速度提出了更高的要求.近年来,通过比较特征点邻域图像块内像素的灰度值,生成二值特征描述子成为了计算机视觉领域的研究热点.Calonder等^[5]提出了一种二值鲁棒独立成分特征描述子(binary robust independent elementary features, BRIEF),其主要思路是在特征点周围按某种概率分布方式生成若干点对,将这些点对对应位置上像素的灰度值大小进行对比和二值化后,生成一个二进制串作为该特征点的特征描述子.该描述子的优点是描述子之间的距离可以利用汉明距离计算,匹配速度非常快,而且该描述子仅占用32B存储空间,为SIFT描述子所占512B的1/16,但其缺点是不具备尺度不变和旋转不变性.Rubeer等^[6]提出的旋转不变的BRIEF(oriented BRIEF, ORB)解决了BRIEF不具备旋转不变性的问题.该描述子利用图像不变矩计算特征点的主方向,同时针对旋转后由于主方向的变化,随机点对的相关性会比较大,从而降低描述子的判别性的问题;同时采取贪婪搜索法,寻找相关性较低的随机点对.Leutenegger等^[7]提出了一种旋转且尺度不变的二值特征(binary robust invariant scalable keypoints, BRISK),在生成多尺度图像的基础上提取AGAST特征点,使算法具备尺度不变性.上述二值特征描述子的共同特点是利用图像的灰度信息进行对比生成,通过图像块内部信息对比生成的描述子能够消除光照变化、模糊等引起的描述子

不变性下降问题, 但对视角变换等具有较大敏感性. 上述二值特征描述子的方式均为手工设定、无法保证生成的描述子性能最优, 为此, Trzcinski 等^[19]提出一种基于监督学习的描述子 (boosting binary keypoint descriptors, BinBoost), 将图像块中子块的梯度方向阈值化后作为弱分类器, 利用 Boost 思想对弱分类器进行加权后生成的强分类器作为最终的描述子. 但该方法对梯度方向通过计算阈值进行二值化的方式, 当待匹配图像与训练集使用图像存在较大差异时, 该描述子性能会在较大下降的可能, 实验部分也验证了该推论.

2 Rbfd 描述子生成方法

二值特征描述子在生成和匹配效率上具有很大的优势, 但简单通过像素灰度值对比生成描述子的判别力较差. 原因主要包含 2 个方面: 1) 简单对比不同点位的灰度值生成的描述子其区分力相对较低, 如图 1a 所示. 图中左右 2 个图像块内容并不相同, 通过对比由线段连接区域灰度生成二值描述后, 它们的描述子相同; 2) 当图像存在视角变化时, 描述子性能会严重下降. 原因在于存在视角变换的图像块之间, 如果以相同的测试点分布方式生成二值特征描述子, 2 个图像块测试点对应的像素灰度值不匹配, 通过对比对应位置上像素的灰度生成描述子会因为灰度值的不匹配造成算法性能下降严重, 如图 1b 所示. SIFT 算法将图像块分为 4×4 子块, 并统计子块中加权梯度方向累加直方图, 使算法具备了一定的仿射不变性. 然而该算法中梯度的计算容易受光照变化影响, 当图像在较大光照变化的情况下, 如果不进行均衡化处理, 算法性能会明显下降. BinBoost 算法利用计算梯度方向阈值生成二值特征描述子的方式同样存在对光照变化鲁棒性差的问题.

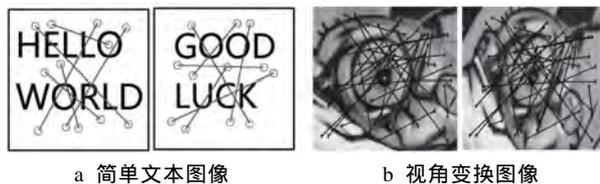


图 1 二值特征描述子性能分析示意图

本文描述子的提出主要受二值特征描述子匹配速度快内存占用少, 而图像梯度信息更具判别力启发, 其生成流程如图 2 所示. 主要包括以下步骤:

Step1. 将特征点邻域图像分块, 利用积分图像对各子块提取梯度统计信息.

Step2. 对各子块进行两两排列组合, 对组合后子块梯度统计信息进行两两对比, 生成二值特征描述子为进一步提高描述子判别力, 本文利用多分块和多邻域策略提取特征点在不同策略下的描述子.

Step3. 利用近似简化的 AdaBoost 算法抽取描述子中最具判别力的字节, 保证描述子的紧凑性.

实验结果表明, 本文描述子在具备二值特征描述子优点的基础上其鲁棒性更强.

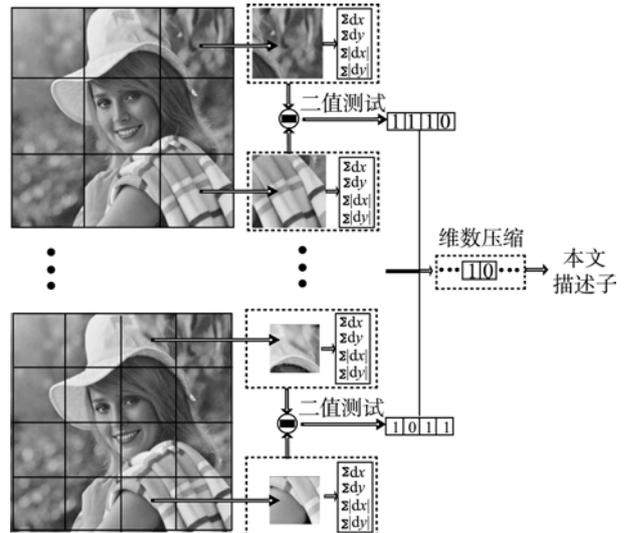


图 2 Rbfd 描述子生成流程图

本文利用 τ 测试生成二值特征描述子, τ 测试定义为

$$\tau(F(m), F(n)) := \begin{cases} 1, & \text{if } (F(m) > F(n)); \\ 0, & \text{otherwise} \end{cases}$$

其中 $F(\cdot)$ 为提取分块中梯度信息的函数, 该函数由 4 个子函数组成 $F(\cdot) = \{F_{dx}(\cdot), F_{dy}(\cdot), F_{|dx|}(\cdot), F_{|dy|}(\cdot)\}$, 分别用来提取分块中的水平梯度均值、竖直梯度均值、水平梯度绝对值均值和竖直梯度绝对值均值, 其中 m 和 n 表示各图像分块.

2.1 多分块与多邻域策略

用于提取描述子的分块方式对描述子的判别力具有直接影响. SIFT, SURF 以及 GLOH 算法描述子均采用单一的分块方式; SIFT 和 SURF 算法描述子将图像邻域划分为 4×4 等大的方形区域, GLOH 算法描述子将图像邻域按环形方式分割: 它们提取的特征点邻域图像块均为单一大小. 采用较大的分块方式, 如将图像块分割为 3×3 等大小的子块, 生成的描述子具有较强的稳定性, 但其判别性将小于更小的分块方式, 如 4×4 分块方式. 通过采取不同的分块方式以及使用不同大小的邻域图像, 能够使描述子包含更多的空间信息,

使描述子更具判别力和鲁棒性. 本文采取图 1 所示 3×3 和 4×4 2 种分块方式, 并且采用邻域图像大小分别为 24×24 , 36×36 , 48×48 像素. 矩形分块方式可以利用积分图像来实现梯度统计信息的快速计算.

2.2 基于近似简化 AdaBoost 的描述子特征选择
利用不同的分块策略和特征点局部邻域图像块大小的特征生成方式, 最终生成描述子的维数为 $(C_9^2 + C_{16}^2) \times 3 \times 4 = 1872$. 描述子维度过高将直接影响描述子匹配速度和存储空间的大小, 而且维数过高使得各维之间容易产生强相关性. 因此需要对高维的描述子进行降维处理, 从中选择最具判别力的维度, 同时消除描述子之间的相关性. 本文提出利用改进的 AdaBoost 算法对原始描述子进行降维.

AdaBoost^[20]由 Freund 等在 Boosting 算法的基础上提出, 它在机器学习领域受到了极大的关注. 该算法是一种迭代算法, 其核心思想是针对同一个训练集训练不同的分类器(弱分类器), 然后将这些弱分类器集合起来, 构成一个更强的最终分类器(强分类器). 算法通过改变数据分布来实现, 根据每次训练集中每个样本的分类是否正确, 以及上次的总体分类的准确率来确定每个样本的权值. 而且将修改过权值的新数据集送给下层分类器进行训练, 最后将每次训练得到的分类器融合起来作为最终的决策分类器.

AdaBoost 算法对每个弱分类器均赋予浮点型的权值来组成最终的强分类器. 本文描述子为二值型数据, 添加浮点型权值后将变为浮点型数据, 无法利用海明距离进行快速匹配; 因此, 本文对 AdaBoost 算法最后生成的弱分类器均添加相同的权值. 虽然这种改进只是对原始算法的一个近似, 但实验结果表明, 本文算法仍然能够获得较好的结果.

基于近似简化 AdaBoost 的特征选择算法流程如下:

Step1. 输入训练数据 $T_i = (x_i, y_i, l_i)$, 其中 (x_i, y_i) 为一对图像块, 如果图像块匹配则 $l_i = 1$; 否则 $l_i = -1$, $1 \leq i \leq N$, N 为训练集样本大小. 计算所有训练图像块的原始 M 维特征描述子.

Step2. 对所有训练集 T_i 赋予初始权重 $w_i = 1/N$.

Step3. for $t=1$ to K

$$\text{对权重进行归一化, } w_{t,i} = \frac{w_{t,i}}{\sum_{q=1}^N w_{t,q}};$$

利用描述子中的每一维特征, j , 对训练集中的图像对进行判断, $C_j(x_i, y_i)$. 计算累积误差, 选择具有最小加权累积误差的比特, $b_t = \arg \min \varepsilon_j$, 其中 $\varepsilon_j =$

$$\sum_{i=1}^{|T|} w_i [l_i \neq C_j(x_i, y_i)];$$

根据累积误差更新每个训练数据的权重 $w_{t+1,i} = w_{t,i} \beta^{1-e_i}$. 其中, 如果判断正确 $e_i = 0$, 否则 $e_i = 1$, $\beta = \frac{\varepsilon_j}{1-\varepsilon_j}$.

Step4. 输出 K 维二值特征描述子.

3 实验与分析

本文使用 Brown 数据集^[22]作为描述子降维训练集, 它包含 3 个子数据集, 每个数据集含有 40 万幅 64×64 像素大小含视角和亮度变化的图像块. 该数据集图像块从图像中的 DoG 特征点邻域抽取, 并且给出了匹配和非匹配图像块的标准结果. 为测试描述子对不同图像变化的鲁棒性, 本文使用 Mikolajczyk05 标准数据集^[14]进行测试; 它包含视角变化(graf, wall 图像)、模糊(bikes 图像)、光照变化(leuven 图像)、JPEG 压缩(ubc 图像)等变化条件下的场景图像. 每种场景图像包含 6 幅图像, 本文将每种场景图像中的第 1 幅图像与其余 5 幅图像进行匹配, 共生成 5 对图像, 分别记为 1/2~1/6. 由于描述子旋转不变性的强弱只是方向估计算法性能的体现, 无法体现描述子本身性能的优劣^[5], 因此实验没有选取存在尺度和旋转变换图像. 如同算法的尺度不变性只与生成的尺度空间的大小有关. 当需要匹配算法具备旋转不性时, 只需根据检测到的特征点的局部图像结构计算一个基准方向, 如利用图像不变矩计算特征点方向^[21]. 为保证实验比较的公平, 实验中用于对比的 SIFT, SURF, BRIEF, ORB, BRISK, BinBoost 也均采用不计算主方向的方式. 本文除 BinBoost 算法利用文献[19]提供的代码外, 其余实验均在基于 OpenCV2. 4. 4 的 VS2008 环境中实现, 计算机硬件配置为主频 3.01GHz 处理器, 3GB 内存.

3.1 Brown 数据集测试结果

本组实验分别利用其中的 Liberty 和 Notre Dame 子数据集作为训练集, 对原始描述子进行降维, 并对训练结果在不同的图像集上进行测试. 为对比描述子压缩对描述子性能的影响, 以及权衡描述子性能与维度之间的关系, 实验采用文献[22]

中的评价方式给出了原始维度描述子以及不同压缩维度描述子的 90% 错误率; 该错误率越小说明描述子性能越好. 同时为了验证本文描述子维度压缩算法的有效性, 图 3 还给出了利用随机抽取方式进行维度压缩后所得描述子 90% 错误率, 本文分别随机抽取了 256 维和 512 维. 实验分别利用 2 个子数据集得到的训练结果在另一个图像集中测试, 训练集和测试集大小均为 10 万对图像块, 匹配对和非匹配对各占 50%.

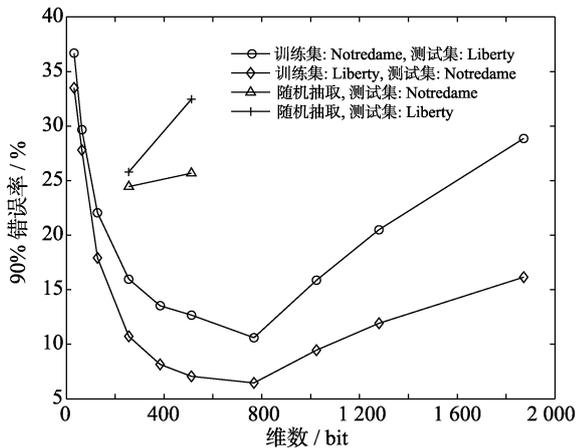


图 3 本文算法生成的不同压缩维度、随机抽取维度以及原始维度描述子的 90% 错误率对比

通过对比图 3 中本文算法生成的不同维度描述子的 90% 错误率可以看出, 较高的压缩率和较低的压缩率或不压缩均会产生较大的错误率, 而当描述子的维度为 500~800 时, 性能达到最佳. 考虑描述子存储空间占用以及匹配速度, 将 512 维作为本文描述子的最终维度. 通过与利用随机抽取进行描述子压缩的方法对比可以看出, 当压缩维度相同时, 本文算法较随机抽取进行维度压缩更加有效.

图 4 给出了本文描述子在 Brown 数据集上测试结果的 ROC 曲线, 并且与 SIFT, SURF, BRIEF, ORB, BRISK, BinBoost 描述子匹配结果进行了对比. 图 4 同时还给出了随机抽取 256 维 rand-Rbfd32 和 512 维描述子 rand-Rbfd64 的测试结果. 图 4 a 给出了以 Liberty 数据集为训练集对原始描述子进行降维, 分别生成的 256 bit 的 Rbfd32 和 512 bit 的 Rbfd64, 其性能在 Notre Dame 数据集上的测试结果. 通过对比可以发现, Rbfd32 性能与 SIFT

接近, 当错误率较低时甚至优于 SIFT, 只在错误率为 0.15~0.4 时略微低于 SIFT; Rbfd64 描述子性能在所有描述子中为最优. BRISK 描述子性能最差, ORB 和 BRIEF 描述子性能接近, SURF 描述子性能仅次于 SIFT; BinBoost32 描述子与本文 Rbfd64 描述子性能近似. 图 4b 的结果与图 4a 近似, 不同之处在于 Rbfd32 描述子性能在 Liberty 测试集上整体略优于 SIFT 描述子. 本文描述子性能好的原因在于其使用了梯度信息, 并且采用了多分块以及多邻域策略, 不同分块以及邻域策略可以使描述子包含更多的图像空间信息, 提升了描述子的判别力. 近似简化的 AdaBoost 算法对描述子性能同样具有增强作用. 表 1 给出了各描述子在正确率为 85% 时的错误率. 其中; 错误率越低, 说明描述子的性能越好. 由表 1 对比可以看出, 本文描述子的错误率最低; 相比表现最差的 BRISK 描述子, 本文算法错误率要低 30%; 与表现最好的 BinBoost32 算法在 Notre Dame 测试集上对比, Rbfd 描述子错误率比其低约 0.62%, 在 Liberty 测试集上低约 0.33%.

3.2 Mikolajczyk 数据集测试结果

本组实验利用 FAST 特征点提取算法^[6], 从每种场景第 1 幅图像中提取 1000 个特征点, 为得出待匹配图像中的实际对应特征点, 实验利用数据库给出的图像间的标准单应变换矩阵, 将第 1 幅图像中的特征点变换至第 2 幅图像. 这种方式在一定程度上会提高匹配的准确率, 但本文实验各描述子均采用该方式进行比较, 保证了对比的公平性. 实验采用文献[5]中的识别率作为评价标准, 其计算公式为 $R = \lambda / \kappa \times 100\%$; 其中, λ 为正确匹配数, κ 为总匹配数, 各描述子的总匹配数均为 1 000. 本文使用最近邻匹配策略匹配特征点, 利用 RANSAC 算法^[23]消除错配点并计算正确匹配数. 图 5 给出了各描述子在各场景图像中的识别率对比. 实验中本文描述子利用由 Liberty 训练集得到的降维方式进行降维, 经实验验证, 由 Notre Dame 训练获取的降维结果其性能与由 Liberty 测试集获取结果性能相近.

图 5a 与图 5b 分别给出了描述子在存在视角变化图像中的匹配性能, 图 5a 中的 graf 场景图像

表 1 各描述子匹配正确率为 85% 时的错误率对比

测试集	SIFT	SURF	BRIEF	ORB	BRISK	BinBoost	Rbfd64
Notre Dame	5.04	8.60	20.83	20.05	37.93	2.70	2.08
Liberty	8.67	14.08	26.33	25.71	42.25	5.54	5.21

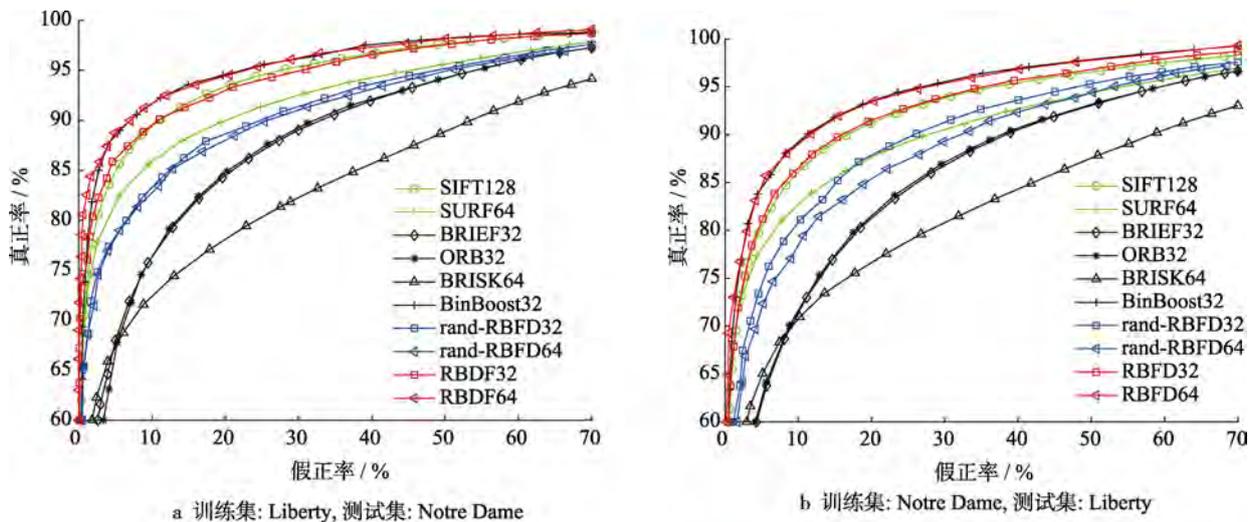


图 4 各描述子在不同测试集上的 ROC 曲线对比

不仅包含倾斜角度,而且含有一定的旋转角度; wall 场景图像只包含倾斜角. 由 graf 图像集中的识别率对比可以发现,本文描述子 RBFD32 与 RBFD64 性能均优于其他二值特征描述子, RBFD64 性能优于 SIFT, 只在 graf1/2 和 graf1/4 中较 SURF 略差,但在 graf1/3, graf1/5 和 graf1/6 中要优于 SURF 描述子. 在图 5b 的 wall 图像中,本文描述子性能均优于 SIFT 及 SURF. 与二值特征描述子相比,在视角变化不大的 wall1/2 和 1/3 中,本文描述子性能与其他二值特征描述子近似,但随着视角逐渐变大,本文描述子识别率均优于其他二值特征描述子. 在 wall 图像中 BRIEF, ORB 和 BRISK 性能要优于 SIFT 及 SURF,而在 graf 图像中相对较差;其原因在于 wall 图像只含有一定的倾斜角. 不同倾斜角图像在竖直方向上进行像素灰度对比,其结果仍然具有一定的判别性. 而 graf 图像由于视角变化复杂,不仅具有倾斜角,还有一定旋转角,所以利用灰度对比来生成描述子的算法,判别力正如论文前面分析的原因,性能会存在较大的下降. 当视角变化不大时,利用灰度对比生成二值特征描述子具有较大的优势. 图 5c 为图像存在模糊时的实验结果,本文描述子的识别率均优于 SIFT, SURF, BRIEF, ORB 以及 BRISK. SIFT 及 SURF 描述子与通过对比方式生成的二值特征描述子相比,性能较差,说明通过对比方式可以消除由图像模糊带来的描述子之间的差异. 由图 5d 可以发现,本文描述子对光照变化具有很强的鲁棒性,其识别率均为 100%;而 SURF 算法对光照变化较为敏感,性能在所有描述子中最差;其原因在于通过对比方式生成的描述子在一定程度上可以克服光照变化引起的描述子性能下降问题. 由

图 5 可以发现, SIFT 描述子也能够取得较好的匹配效果,其原因在于 SIFT 描述子对图像亮度进行了均衡化处理. 图 5e 为描述子在 JPEG 压缩图像上的匹配结果由对比结果可以得出,各描述子对 JPEG 压缩均具有很强的鲁棒性,识别率均达到 95% 以上,本文描述子识别率均达到 100%; SIFT 描述子相对较差, SURF 描述子最差.

为验证各描述子对噪声的鲁棒性,实验将 graf 图像中的第 1 幅添加了 10%~30% 的高斯噪声. 生成了 5 幅噪声图像. 对不含噪声的第 1 幅图像提取 1000 个 FAST 特征点,同时将提取的特征点定位到含噪声图像上,提取特征点对应位置的描述子,以对比各描述子的识别率. 图 5f 给出了本文算法与其他算法的结果对比. 通过对比可以发现, BinBoost 算法对高斯噪声比较敏感, SURF 算法其次,本文算法与 ORB, BRIEF 算法性能相近, SIFT 与 BRISK 算法较本文算法略有优势.

特征点提取过程中容易产生定位误差,为比较描述子对特征点定位误差的鲁棒性,本文同样将 graf 中的第 1 幅图像提取 1000 个 FAST 特征点后,加入 1~5 个像素的定位误差,然后分别提取描述子,将不存在定位误差和存在定位误差的描述子经行匹配,图 5g 给出了实验结果. 由图 5g 结果可以看出,本文描述子和 SIFT 描述子均对定位误差有很强的鲁棒性, BRISK 描述子其次, ORB 描述子性能最差;随着定位误差的加大, SURF, BRIEF, BinBoost 描述子性能也存在较大幅度的下降.

为对比各描述子区分匹配与非匹配特征点的性能,本文采用文献[14]方式给出了部分图像匹配结果的查全率-查错率曲线,结果如图 6 所示. 好的描述子应该能够在提高查全率的同时,降低查错

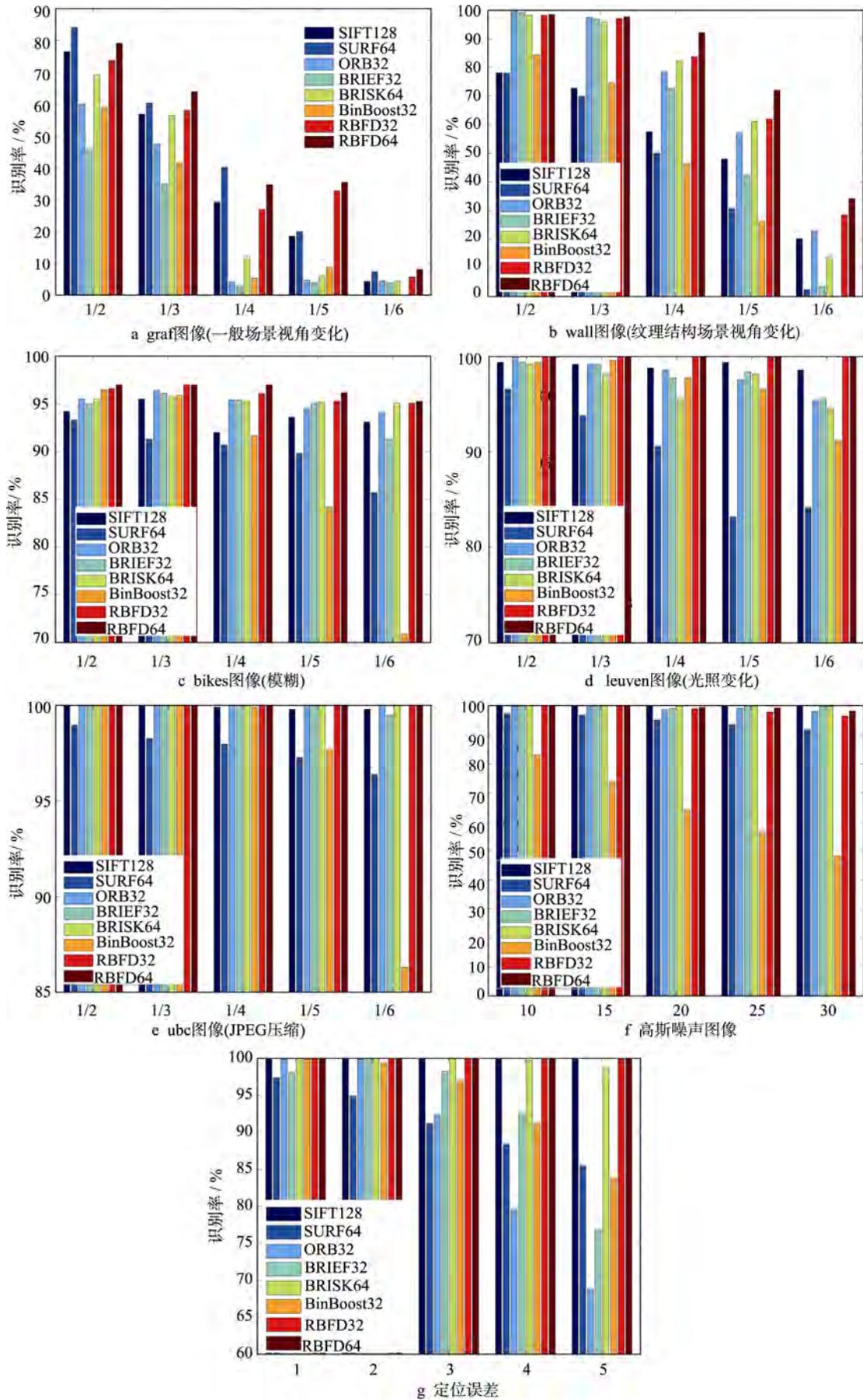


图 5 各描述子在 Mikolajczyk 数据集上的识别率对比

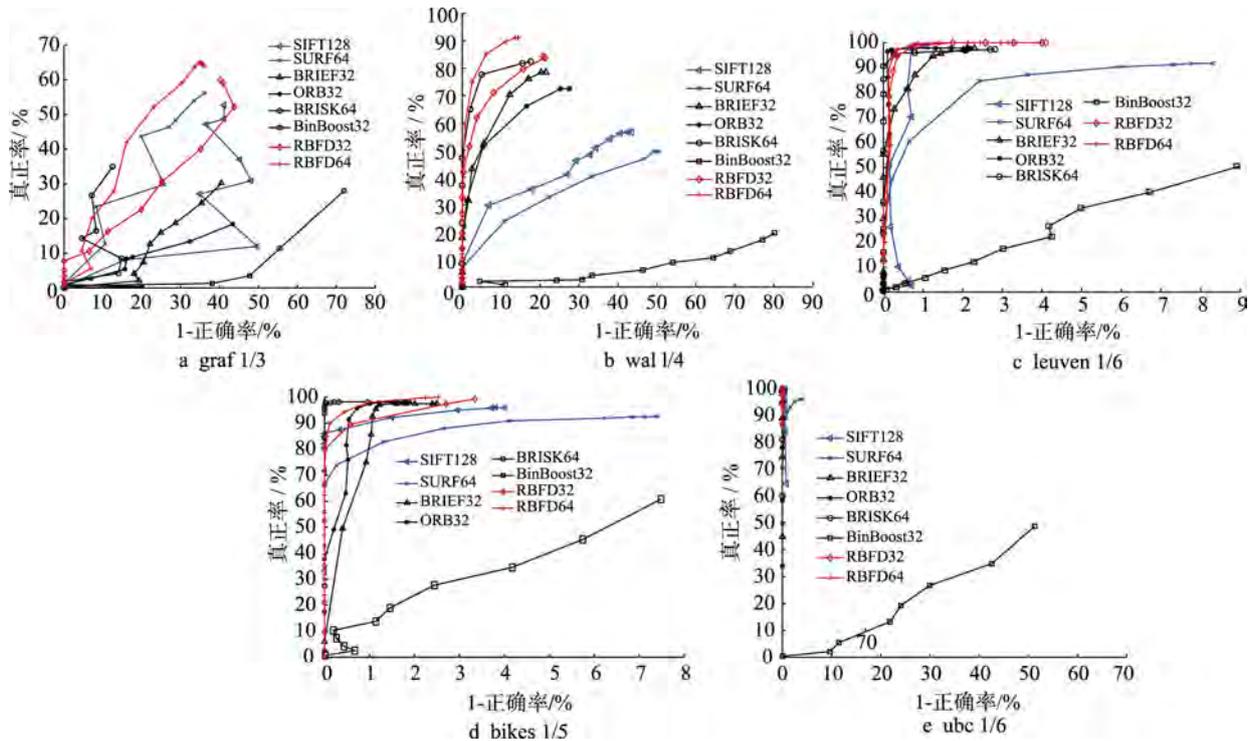


图 6 各描述子的真正率与 1-正确率曲线对比

率。通过比较可以发现，除 bikes 场景外，本文 Rbfd64 性能均优于或近似于其他描述子。特别是针对视角变化图像的匹配，本文描述子的匹配性能均有显著提升。本文描述子在 bikes 场景中性能稍差的原因在于 bikes 场景图像为模糊图像，由于图像的模糊会对图像梯度的计算产生较大的影响，因此以梯度信息作为描述子的算法性能普遍较差，这一点可以从 SIFT 和 SURF 的结果看出；但通过对比方式可以在一定程度上消除这种影响。Rbfd32 描述子相比 BRISK 性能略差，但整体优于 ORB, BRIEF, SIFT, SURF 和 BinBoost 描述子。BRISK 描述子在本实验中性能较好，部分原因在于 BRISK 长度为 64 Byte，相比 ORB 和 BRIEF 的 32 Byte 要长一倍，另一方面原因在于 BRISK 采取的灰度对比方式有助于提升描述子性能^[7]。由图 6e 的 JPEG 压缩图像曲线可以看出，本文描述子与 BRIEF, BRISK 对 JPEG 压缩均具有很强的鲁棒性，查错率在不同查全率的取值均为 0，说明在不同查全率情况下，其匹配的正确率均为 100%。SIFT 与 SURF 描述子对 JPEG 压缩具有一定的敏感性。BinBoost 算法在所有 Mikolajczyk 数据集上的 ROC 曲线均最差。

通过比较可以发现，SIFT 和 SURF 描述子在 Mikolajczyk 图像集上的结果要明显差于 Brown 数

据集上的结果。一个可能的原因是特征描述子的性能与特征点检测算法直接相关，不同的描述子适用于不同的特征点检测算法。Brown 数据集中的图像块中的特征点由 DoG 斑点检测算子提取，而本身 SIFT 算法使用的特征点检测子即为 DoG 算子。另外一个原因在于数据集中的图像块包含的图像变化类型不同。Brown 数据集中的图像主要包含视角以及光照的变化，而 Mikolajczyk 数据集还包含有模糊以及 JPEG 压缩图像。SIFT 与 SURF 描述子对视角变化具有较强的鲁棒性，因而性能在 Brown 数据集上的测试结果较好。BinBoost 在 Mikolajczyk 数据集上的测试结果说明，该描述子只是适用于原论文中的测试数据集，而不具有很好的通用性。本文描述子在以上 2 个数据集上均取得了较好的测试结果，说明本文描述子对上述各种图像变化均具有较强的鲁棒性。

3.3 描述子生成实时性

为了对比各描述子生成效率，本文对各描述子在 Mikolajczyk 数据集上的描述子生成时间进行了统计。表 2 给出了各算法单个描述子生成的平均时间。本文 Rbfd64 生成时间较 BRIEF 和 BRISK 描述子生成时间略有增加，与 ORB 描述子生成时间近似，但相比 BinBoost, SURF 和 SIFT 描述子生成时间，约有 1 个数量级的提高。

表2 描述子生成时间对比 ms

	SIFT	SURF	ORB	BRIEF	BRISK	BinBoost	Rbfd
生成时间	4.353	1.351	0.107	0.042	0.051	1.023	0.119

4 结 语

局部特征描述子是图像检索、三维重建、图像拼接以及目标识别与跟踪的研究重点与难点。二值特征描述子因其具有的快速匹配以及低存储消耗的优点而成为目前计算机时觉领域的研究热点。本文设计了一种新型的二值特征描述子 Rbfd; 并与当前流行的浮点型及二值特征描述子进行了对比。大量实验结果表明, 本文描述子在具备快速生成优点的基础上, 对图像的各种变化具有更强的鲁棒性。

参考文献(References):

- [1] Ge Juan, Cao Weiguo, Zhou Wei, *et al.* A local feature descriptor under color affine transformation [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2013, 25(1): 25-33(in Chinese)
(葛娟, 曹伟国, 周炜, 等. 一种颜色仿射变换下的局部特征描述子[J]. *计算机辅助设计与图形学学报*, 2013, 25(1): 25-33)
- [2] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [3] Lowe D G. Object recognition from local scale-invariant features[C]//*Proceedings of the International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 1999, 2: 1150-1157
- [4] Bay H, Tuytelaars T, van Gool L. SURF: speeded up robust features[M]//*Lecture Notes in Computer Science*. Heidelberg: Springer, 2006, 3951: 404-417
- [5] Calonder M, Lepetit V, Strecha C, *et al.* Brief: binary robust independent elementary features[M]// *Lecture Notes in Computer Science*. Heidelberg: Springer, 2010, 6314: 1-14
- [6] Rublee E, Rabaud V, Konolige K, *et al.* ORB: an efficient alternative to SIFT or SURF[C]//*Proceeding of IEEE International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 2011: 2564-2571
- [7] Leutenegger S, Chli M, Siegwart R Y. BRIS: binary robust invariant scalable keypoints[C]//*Proceedings of IEEE International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 2011: 2548-2555
- [8] Alahi A, Ortiz R, Vanderghaynst P. Freak: fast retina keypoint[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2012: 510-517
- [9] Morevec H P. Towards automatic visual obstacle avoidance[C]//*Proceedings of the 5th International Joint Conference on Artificial Intelligence*. San Francisco: Morgan Kaufmann Press, 1977: 584-584
- [10] Moravec H P. Visual mapping by a robot rover[C]//*Proceedings of the 6th International Joint Conference on Artificial Intelligence*. San Francisco: Morgan Kaufmann Press, 1979: 598-600
- [11] Harris C, Stephens M. A combined corner and edge detector[C]//*Proceedings of the 4th Alvey Vision Conference*. South Yorkshire: University of Sheffield Printing Office, 1988: 147-151
- [12] Smith S M, Brady J M. SUSAN—a new approach to low level image processing[J]. *International Journal of Computer Vision*, 1997, 23(1): 45-78
- [13] Ebrahimi M, Mayol-Cuevas W W. SUSurE: speeded up surround extrema feature detector and descriptor for realtime applications[C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2009: 9-14
- [14] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615-1630
- [15] Rosten E, Drummond T. Machine learning for high-speed corner detection[C]//*Proceedings of the 9th European Conference on Computer Vision*. Heidelberg: Springer, 2006, 1: 430-443
- [16] Davison A J, Reid I D, Molton N D, *et al.* MonoSLAM: real-time single camera SLAM[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(6): 1052-1067
- [17] Mair E, Hager G D, Burschka D, *et al.* Adaptive and generic corner detection based on the accelerated segment test[M]// *Lecture Notes in Computer Science*. Heidelberg: Springer, 2010, 8312: 183-196
- [18] Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors[C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2004, 2: 506-513
- [19] Trzcinski T, Christoudias M, Fua P, *et al.* Boosting binary keypoint descriptors[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2013: 2874-2881
- [20] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. *Journal of Computer and System Sciences*, 1997, 55(1): 119-139
- [21] Rosin P L. Measuring corner properties[J]. *Computer Vision and Image Understanding*, 1999, 73(2): 291-307
- [22] Brown M, Hua G, Winder S. Discriminative learning of local image descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(1): 43-57
- [23] 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