基于优化采样模式的紧凑而快速的二进制描述子

惠国保 [1] 李东波 [1] 童一飞 [1]

摘 要 进一步挖掘图像补丁特征信息, 提出了一种鲁棒性更高的二进制描述子. 针对传统的二进制描述子对旋转和视角变化鲁棒性差的问题, 本文通过优化采样模式和分解图像补丁对其改进. 首先, 通过对最新提出的采样模式特点的分析、测试, 发现采样点密度和平滑重叠度对产生的描述子独特性有重要影响; 据此, 调整这两个影响因子, 设计出一种优化的采样模式. 其次, 利用像素点灰度值排序方法分解图像补丁, 产生多个对应不同灰度段的亚补丁. 最后, 将优化的采样模式映射到亚补丁上, 随机提取样本点进行灰度值比较测试. 所得到的二进制描述子不仅包含了补丁像素的灰度比较信息, 而且包含了灰度排序信息. 通过对比实验看到本文的二进制描述子对特征识别匹配效果最好. 本文的特征描述方法可应用于实时性要求高、内存紧凑的高质量目标识别.

关键词 关键点, 图像补丁, 二进制描述子, 采样模式, 灰度排序, 亚补丁

引用格式 惠国保, 李东波, 童一飞. 基于优化采样模式的紧凑而快速的二进制描述子. 自动化学报, 2014, **40**(2): 255−266

DOI 10.3724/SP.J.1004.2014.00255

A Binary Descriptor Based on Both Optimized Sampling Pattern and Image Sub-patches

HUI Guo-Bao[1] LI Dong-Bo[1] TONG Yi-Fei[1]

Abstract This paper proposes a more robust binary descriptor through further excavating feature information of image patch. Conventional binary descriptors such as binary robust independent elementary features (BRIEF) are not robust to rotation and viewpoint invariance, which is improved from two aspects in this paper. Firstly, an optimized sampling pattern is presented by tuning the density of sampling points and the overlapping size of receptive fields. Secondly, all pixels in the patch are classified according to their intensity order, so that the patch is decomposed into several sub-patches. Then, random tests on each sub-patch mapped with the optimized sampling pattern are repeatedly taken, and each test result is concatenated to form a distinct binary string of the sub-patch. The proposed descriptor encodes not only intensity-comparison information but also information about relative relationship of intensities. As a result, results based on experiments of performance evaluation have shown that the proposed binary descriptor outperforms the-state-of-the-art binary descriptors.

Key words Keypoint, image patch, binary descriptor, sampling pattern, intensity order, sub-patch

Citation Hui Guo-Bao, Li Dong-Bo, Tong Yi-Fei. A binary descriptor based on both optimized sampling pattern and image sub-patches. *Acta Automatica Sinica*, 2014, **40**(2): 255−266

许多计算机视觉应用涉及到图像配准技术, 如目标识别、3D 重构、场景拼接等. 有许多方法可以实现图像配准, 其中特征匹配是一个重要的方法. 特征匹配的性能与局部特征描述方式紧密相关. 以某个关键点 (兴趣点) 为中心的图像局部区域 (图像补丁), 根据其内部像素点的灰度信息, 对该区域特征进行描述, 就得到一个能表达该局部区域特征的描述子. 局部特征描述子是表达图像关键点周围局部特征的向量. 局部特征描述子能代表关键点进行特征匹配识别等图像处理任务.

一般认为, 一个好的特征描述子, 既能显著地不同于同幅图像其他特征描述子的同时, 又能在各种图像转换下保持其描述特征不变性, 即在匹配中具有较好的查全率[1]. 但是, 要将图像配准应用到资源有限的嵌入式设备, 只有好的查全率是不够的, 还要能适应计算能力有限、存储空间小的条件. 所以, 一个好的描述子, 在各种变换条件下具有较高的鲁棒性的同时, 还要求具有结构紧凑和计算简单的特点. 为此, 获得这样高性能的局部特征描述子是图像配准研究的一项重要任务.

对于同一关键点, 不同的描述方式产生的描述子性能是不同的. 为了获得高性能的描述子, 近些年

收稿日期 2013-05-28, 录用日期 2013-09-18
Manuscript received May 28, 2013; accepted September 18, 2013
国家自然科学基金 (61104171) 资助
Supported by National Natural Science Foundation of China (61104171)
本文责任编委 周杰
Recommended by Associate Editor ZHOU Jie
1. 南京理工大学机械工程学院 南京 210094
1. School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094

来涌现了大量的描述子构建方法. 根据描述子的表现形式, 主要分为两类: 浮点型描述子和二进制型描述子. 对于浮点型描述子, 两个最具代表性的方法是尺度不变特征转换 (Scale-invariant feature transform, SIFT)[2] 和快速鲁棒特征 (Speeded up robust features, SURF)[3], 它们各自提出了一套从特征检测到匹配的方法. 此后, 有许多学者试图改进它们以提高其鲁棒性, 如主成分分析尺度不变特征转换 (Principal component analysis-SIFT, PCA-SIFT)[4]、梯度位置方向直方图 (Gradient location and orientation histogram, GLOH)[1]、DAISY[5] 和多支持域直方图 (Multi support region of gradient histogram, MROGH)[6] 等. 浮点型描述子虽然具有良好的性能, 但是匹配速度慢, 占用内存多.

随着嵌入式手持设备对图像处理应用的需求日益增大, 浮点型描述子已不适用于低端处理器和存储资源有限的设备. 于是在这些应用场合采用二进制型描述子取代浮点型描述子成为必然. 二进制描述子的优点是可以通过计算描述子之间的汉明距离来进行匹配. 汉明距离只计算涉及到两个位之间的异或操作, 所以计算效率很高、速度很快. 近几年, 学者们也提出了许多方法构建二进制描述子, 可以归纳为以下三类:

1) 浮点型描述子二进制化. 有三种方式: a) 通过某个阈值, 将浮点型向量进行二值划分, 最终转化为二进制串, 如 Torralba 等[7] 用这种方式将 GIST[1] 描述子二进制化; b) 通过随机旋转矩阵量化浮点向量, 加快了匹配速度, 如 Jégou 等[8] 用这种方式将 SIFT 描述子二进制化; c) 采用非线性相邻单元分析, 虽然二进制化的效果很好, 但处理过程慢, 如 Salakhutdinov 等[9] 采用这种方法将 GIST 描述子二进制化.

2) 采样点灰度值比较法. 因为相对少量成对像素灰度值比较能有效地对图像补丁分类[10−11], 所以基于这个思想, 以关键点为中心的局部区域, 随机选择两个样本点进行灰度值比较测试, 串联每次比较结果, 即构成了二进制描述子. 二值鲁棒独立单元特征 (Binary robust independent elementary features, BRIEF)[12] 就是用这种方式构建二进制描述子. 因为在一个图像补丁里随机采样, 容易造成某些采样不能增加描述子的独特性, 即随机采样容易产生冗余信息, 于是通过采样模式来增加特征信息含量, 如二值鲁棒不变尺度关键点 (Binary robust invariant scalable keypoints, BRISK)[13]、ORB[14]、快速描述神经元类关键点 (Fast retina keypoint, FREAK)[15] 等. 在采样模式里, 随机提取样本点进行测试, 一定程度上增强了描述子的独特性.

[1] http://gistdescriptor.codeplex.com/

3) 通过机器学习训练出二进制描述子. 从机器学习的角度, 通过多个对应的图像补丁对和非对应的图像补丁对进行训练, 由几个简单的过滤器的线性组合计算得到几个独特投影的线性组合. 将要描述的图像补丁投影到这线性组合里, 通过一个阈值来二值划分投影结果, 再将这些投影结果串联起来, 就能得到一个二进制串, 即为二进制描述子. 如 D-Brief 等[16] 和 LDA-Hash 等[17] 都是采用这种方式构建二进制描述子. BINBOOST 等[18] 引入非线性组合过滤器到机器学习里, 以此增强投影结果的独特性.

上述三类二进制描述子构建方式大不相同, 得到的二进制描述子的大小和包含的信息也不同. 但是, 构建二进制描述子的目标是相同的, 即提高计算速度和节省存储空间. 在性能上, 这些二进制描述子的鲁棒性都要逊色于浮点型描述子.

本文通过改进采样点灰度值比较法, 提出一个鲁棒性更高的二进制描述子. 首先, 在分析 BRISK 和 FREAK 采样模式特点的基础上设计一种独特的采样模式. 接着, 对图像补丁内所有像素点灰度值进行分类, 据此将图像补丁分为多个亚补丁 (即包含某段灰度值的像素集合). 然后, 将亚补丁的像素灰度信息映射到一种采样模式上, 在采样模式上随机选择采样点进行类似 BRIEF 的比较测试, 每次测试结果依次串联起来, 这就构成了图像补丁的二进制描述子. 因为对图像补丁进行了分解处理, 所以描述子具有旋转和仿射不变性. 这样得到的描述子不仅具有很好的尺度不变性, 而且也增强了对光度变化的鲁棒性.

1 相关研究

近年来提出了许多浮点描述子, 最具代表性的浮点描述子是 SIFT[2]. SIFT 虽然对各种变化具有很强的鲁棒性, 但是描述子生成时间长且匹配速度慢. 后来, 有许多学者对其描述算法进行改进, 产生了许多 SIFT 的变体.

PCA-SIFT[4] 从减维角度筛选 SIFT 描述子分量, 留下较少但关键的分量, 使描述子从 128 维减少到 36 维. GLOH[1] 采用极坐标, 替换 SIFT 方法中的 4×4 采样窗口, 来获得梯度方向直方图, 使得产生的描述子比 SIFT 更独特性. DAISY[5] 多次用不同大小的高斯核, 卷积方向梯度直方图每个方向上梯度值, 通过改进直方图改进描述子.

除了 SIFT 及其变体, 也有许多学者通过不同方法得到浮点描述子. SURF[3] 利用 Hessian 矩阵行列式检测特征点, 再对特征点周围局部区域用 Haar 小波响应和得到描述子. MROGH[6] 将灰度排序

用于构建梯度方向直方图的浮点描述子. 局部灰度序列模式 (Local intensity order pattern, LIOP)[19] 采用查表法找到每组灰度序列对应的 24 位索引值累加得到描述子. 刘萍萍等[20] 通过强度差对比, 建立了一种适用于一定范围内尺度变化、平面旋转和线性照度变换下的快速局部特征描述算法. 唐永鹤等[21] 提出了一种基于拉普拉斯的局部特征描述算法. 葛娟等[22] 提出了一种颜色仿射变换下的局部特征描述子.

上述描述子均用浮点数表达, 对计算时间和内存使用都有要求, 所以学者们试图采用局部二值模式来解决这个瓶颈问题. 局部二值模式的特征提取方法在纹理分析和人脸识别应用中取得了显著成果[23]. 将其应用于特征描述, 目的是得到特征的二进制描述子, 也就是用一串二进制位来表征特征描述子. 较早采用二进制描述子的方法是 BRIEF[12].

BRIEF (又称 U-BRIEF) 方法是基于根据少量成对像素灰度值的比较结果能对图像补丁有效地分类的思想提出的, 在补丁内随机取两个像素点灰度值作比较测试, 根据多次测试的结果得到表征补丁特点的位串. 虽然 BRIEF 描述子简单、速度快, 但是对图像旋转和尺度变化很敏感, 所以后来有许多 BRIEF 改进版, 如 (O, S)-BRIEF 分别从旋转和尺度等方面改进. D-BRIEF[16] 是通过建立数据库模板达到旋转和尺度的不变性. ORB[14] 提出具有方向表征的二进制描述子, 通过训练补丁里的所有随机采样点对来获得更好的描述子.

BRISK[13] 是用 AGAST[24] 检测特征点, 通过采样模式获取特征点对, 再进行灰度值比较测试. 但是, 所选的成对采样点高度关联, 而且不具有独特性. FREAK[15] 构建类似视网膜神经的采样模式, 从训练数据库里学习最好的采样点对进行比较. Boosting 能够将预测精度仅比随机猜度略高的弱学习器增强为预测精度高的强学习器[25], BINBOOST[18] 利用 Boosting 这个特点, 试图通过弱学习确定描述子每个位的值, 改进二进制描述子的鲁棒性. Brown 等[26] 开发了从训练数据里学习得到局部图像描述子. Cai 等[27] 通过线性独特投影减少局部图像描述子的维数, 增加描述子的独特性. Strecha 等[17] 从样本描述子中学习描述子的不同点, 以更短的二进制串来表示描述子.

本文将图像补丁分解为亚补丁, 是受到 MROGH[6] 或 LIOP[19] 的灰度值排序思想的启发. 然后在亚补丁上采用采样模式提取样本点测试, 与 BRISK、FREAK 方式类似. 这样的二进制描述子, 因为包含了不同灰度段的采样信息, 而具有更强的描述性. 通过后面的实验可以发现这样的二进制描述子比 BRISK 和 FREAK 鲁棒性高, 而且比 MROGH 和 LIOP 速度快、占用内存少. 本文描述子标记为 CFRBD, 意思为紧凑、快速、鲁棒的二进制描述子.

2 设计采样模式

本文描述子是建立在优化的采样模式上的, 优化的结果是提高了采样模式性能. 性能好的采样模式, 是指在其上获得的成对采样点关联度低, 而且一系列点对比较结果方差高[14−15]. 在介绍本文采样模式之前, 首先分析 BRISK、FREAK 和 DAISY 的采样模式特点, 本文采样模式是在汲取它们优点的基础上提出的.

BRISK 采样模式是一种采样点在不同半径的同心圆上均匀分布的圆形模式. 随着同心圆半径增大, 圆上采样点个数增多. 每个采样点的灰度值是通过对图像相应区域高斯核平滑得到的, 整个采样模式内 60 个采样点几乎均匀分布. 每个同心圆上两个采样点的距离确定了平滑范围, 也就确定了高斯核的标准差大小, 并且每个采样点的高斯平滑范围不重叠.

FREAK 提出了类似于神经元采样网格模型的采样模式, 也是对称圆形模式. 该模式共有 43 个采样点, 每个同心圆上采样点的个数相同, 所以采样点在整个模式内非均匀分布, 离中心点近的采样点密度高, 远离中心密度呈指数下降. 距离中心点越远的采样点的高斯平滑范围呈指数增大, 所以平滑范围大量重叠, 产生大量的冗余信息. FREAK 指出高斯核大小若为指数变化, 采样模式性能会更好; 同时指出重叠部分也能提高采样模式的性能, 尽管重叠区里只有部分新信息被编码.

DAISY[5] 采样模式也是一种采样点对称分布的圆形模式, 共有 25 个采样点. 在同一个同心圆上, 采样点均匀分布且平滑范围相同, 越往外层, 则平滑范围成倍增加. 整个模式里, 平滑范围有大量重叠. DAISY 描述子是对 SIFT 描述子的改进, 使用该采样模式来替代 SIFT 的采样窗口, 避免了手动调整窗口的参数.

所以, 采样点密度和高斯平滑范围重叠程度是影响采样模式性能的两个因素. 首先分析这两个因素是如何影响采样模式性能, 然后据此衡量这两个因素以构建性能最优的采样模式.

2.1 采样点密度对采样模式的影响

采样点密度是反映采样模式里的采样点向中心集中的程度. 一般来说, 采样点向中心集中程度高的采样点密度就大; 否则, 采样点比较疏散或均匀分布于采样模式的, 采样点密度就小. 在确定了采样模式最小同心圆的情况下, 采样点集中度是指一层同心

圆上的两个采样点距离与其紧邻外层同心圆上的两个采样点距离比值的补数, 而且此后每紧邻两层的比值都一样. 图 1 中, 采样点在三个同心圆上分别均匀分布, 同一个同心圆上相邻的两个采样点的距离与紧邻外层圆上两个采样点的距离比值反映采样点密集程度, 采样点集中度表示为

$$\frac{1-L_1}{L_2}=\frac{1-L_2}{L_3} \tag{1}$$

为了探究采样模式性能与采样点密度之间的关系, 在采样点密度不同的采样模式上进行了描述独特性测试, 观察所得描述子相似度变化, 测试结果如图 2 所示. 描述子独特性越好, 相似度就越低. 两个描述子的相似度可由下式计算出:

$$\text{相似度}=\frac{\text{描述子维数}-\text{汉明距离}}{\text{描述子维数}}\times 100\,\% \tag{2}$$

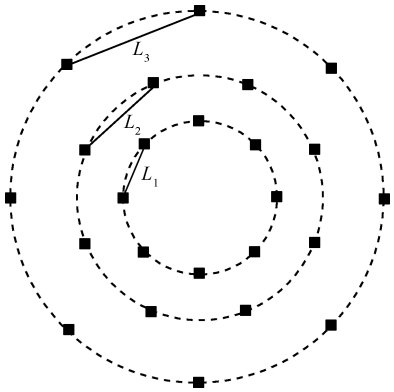


图 1 采样点密度示意图

Fig. 1 The diagram of sampling-point density

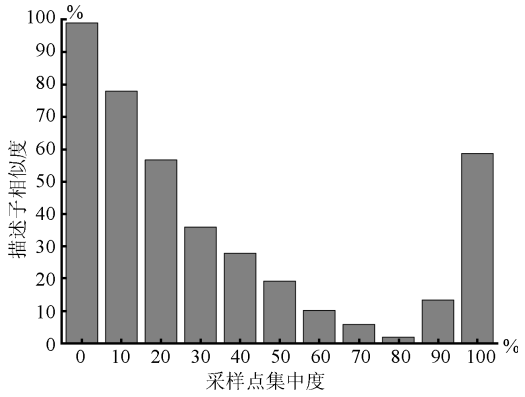


图 2 采样点密度对描述子独特性影响效果

Fig. 2 The relationship between sampling-point density and discriminative power of descriptor

图 2 中, 横坐标的 0 % 是指所有采样点在采样模式里呈均匀分布, 100 % 是指采样点集中程度非常高. 图 2 显示采样点密度对描述子的独特性是有影响的. 如果采样点集中度很小, 即采样模式里的采样点是内密外疏, 这种情况下对局部区域的信息捕获就很少, 致使描述子独特性很差. 如果采样点集中度很高, 采样点呈均匀分布, 就不能捕获有层次的信息, 也会降低描述子的独特性. 在圆形采样模式里, 影响采样点密度主要有两个变量: 一个是同心圆之间的距离; 另一个是同心圆上采样点的个数. 本文是根据上述分析，调整这两个变量得到的测试结果. 在采样点集中度小于 80 % 的描述子相似度变化较大于 80 % 的相似度变化缓慢, 所以采样点集中度要比 80 % 略低, 这里取采样点的中心集中度为 78 %.

2.2 平滑范围重叠度对采样模式的影响

影响采样模式特性的另一个因素是高斯平滑范围重叠. 一般来说, 平滑范围越大, 捕获的新信息越多, 就会增加对关键点独特性描述. 但是, 在一个多采样点的采样模式里, 采样点之间的平滑范围会重叠. 重叠区域产生冗余信息会压缩平滑区里的新信息, 从而会降低特征描述子的独特性.

在讨论平滑范围重叠对采样模式影响之前, 先定义两个名词. 信息冗余度, 指冗余信息与平滑范围内全部信息的比值; 平滑范围重叠度, 指采样模式里某一同心圆上所有采样点与相邻采样点平滑范围重叠面积总和, 与该圆上所有采样点平滑范围面积总和的比值, 所以该值能反映平滑范围和重叠部分之间的关系. 由上述定义可知, 平滑范围重叠度是对信息冗余度的更具体地阐述. 因为采样点在同心圆上均匀分布且平滑范围重叠度都相同, 所以一个采样点的平滑范围重叠度也就是整个同心圆的平滑范围重叠度.

下面介绍一下采样点重叠度的计算. 如图 3 所示采样点 P 平滑范围, 即中间的圆圈所示, 与紧邻四个圆圈都有交集. 计算出所有交集面积和, 再比上该平滑区域的面积, 结果就是该采样点的重叠度. 图中阴影部分面积为

$$S_1=\alpha_1 {r_1}^2+\alpha_2 {r_2}^2-r_1 d\sin(\alpha_1) \tag{3}$$

其中, r_1、r_2 分别两个采样点平滑范围的半径, d 是两个采样点的距离, $\alpha_1=\arccos((r_1{}^2+d^2-r_2{}^2)/(2r_1d))$, $\alpha_2=\arccos((r_2{}^2+d^2-r_1{}^2)/(2r_2d))$. 类似的, 可以计算出其他三个交集 S_2、S_3 和 S_4. 最后得到采样点 P 的平滑范围重叠度为

$$\frac{S_1+S_2+S_3+S_4}{\pi {r_1}^2}$$

为了探究平滑重叠度对采样模式性能影响, 这里做了个测试. 这里仍用描述子相似度来表达描述子的独特性, 测试结果如图 4 所示. 通过测试发现, 高斯平滑范围适度重叠可以捕获到重叠后的交叉信息, 增加局部特征的独特性, 这与 FREAK[15] 所提

的结论吻合. 但随着平滑范围不断增加, 重叠区也就不断增大, 每个平滑范围包含新的局部区域不断减少, 反而又使局部特征的独特性减少, 从图 4 中看到, 22 % 的重叠度能使描述子独特性达到最大化. 而且, 重叠次数不宜过多; 否则, 描述子独特性会迅速下降. 图 4 横坐标轴两端代表两个极端的采样模式: 左端代表没有重叠的 BRISK[13], 右端代表有大量重叠的 FREAK[15] 和 DAISY[5]. 调整局部区域重叠次数可以控制平滑范围重叠程度, 根据这一特点, 适度调整一个区域的重叠范围, 使之达到局部特征独特性上的最优化.

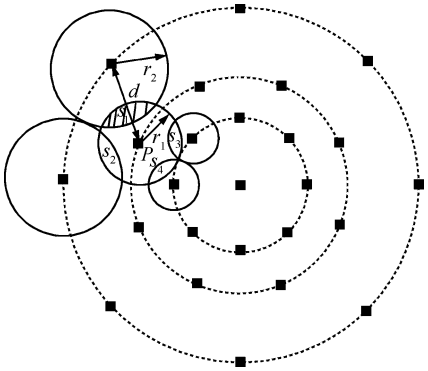


图 3　平滑范围重叠度示意图

Fig. 3　The diagram of overlapping degree

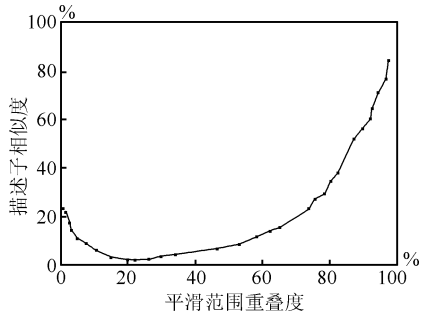


图 4　平滑范围重叠度对描述子独特性的影响

Fig. 4　The relationship between overlapping degree and discriminative power of descriptor

综上分析, 本文采样模式也是圆形采样模式, 如图 5 所示, 图 5 中每个圆圈的圆心为采样点. 同心圆之间距离由内向外逐渐增大, 每个同心圆有 8 个采样点, 每个采样点高斯平滑范围相同, 不同同心圆上采样点的高斯平滑范围大小不等, 平滑范围有适度的重叠, 重叠次数不大于 2.

本文采样模式与 FREAK 和 DAISY 的采样模式有很大不同. 首先, FREAK 采样模型每个同心圆有 6 个采样点, DAISY 有 8 个, 本文也是 8 个, 但采样点个数比 DAISY 要多. 这就使得本文采样模式采样点密度与众不同. 其次, FREAK 和 DAISY 采样模式的采样点高斯平滑范围高度重叠, 局部区域有三次平滑范围重叠, 而本文的采样模式高斯平滑范围重叠适度, 一个区域最多重叠两次. 最后, FREAK 采样模式是根据视网膜神经模型得到的, 而 DAISY 采样模式是为替代 SIFT[2] 窗口而设计的, 但是本文是从采样点密度和平滑重叠程度分析得到的, 兼具了三种采样模式的优点. 通过后面对比实验可以看到本文采样模式性能是其中最好的.

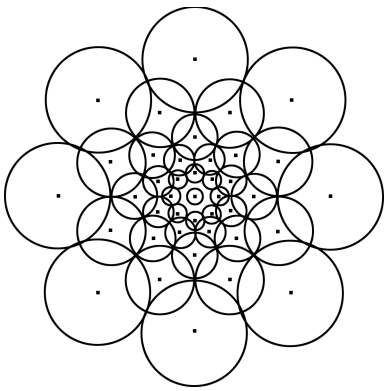


图 5　本文所提的采样模式

Fig. 5　The proposed sampling pattern

3　构建二进制描述子

传统的二进制描述子是直接在原图像补丁上随机选择两个像素点进行灰度值比较测试, 得到的二进制描述子里的每个位都是对等的, 这样只能根据同一个补丁里所有样本点灰度差值信息有限地表达补丁的特点. 所以这样的二进制描述子不具有旋转和视角不变性, 它只能表达基础的图像补丁特点.

本文利用灰度排序方法将原图像补丁分解成若干亚补丁, 在将本文的采样模式映射到亚补丁上, 最后在这样的亚补丁上做测试. 所得描述子因包含图像补丁灰度排序信息, 使描述子对旋转和视角具有不变性, 以此来提升二进制描述子的性能. 原图像补丁分解需要对补丁里所有像素点分类, 于是涉及到像素点灰度值的排序和范围的划分.

本文用到补丁像素点按灰度值分类排序方法, 与 MROGH[6] 或 LIOP[19] 类似. 根据几个灰度值范围将所有像素点按灰度值序列进行分类, 在补丁上属于某类的像素点都保留下来, 不属于该类的像素点全部移除, 这样就得到一个对应该灰度值范围的像素点群聚, 即一个亚补丁. 这就相当于将原图像补丁分解为若干层, 每一层对应一个亚补丁, 也就是对应一段灰度值范围, 属于该范围的像素点都留在该亚补丁上, 如图 6 所示. 图 6 上面部分是补丁对应的灰度直方图; 将灰度值划分范围后, 每个范围对应补丁的像素点作为一个分类, 如图 6 下面部分所示. 为了直观地显示, 将属于该类的像素点的灰度值都设为 256, 即白色, 非该类的像素点的灰度值都设

为 0, 即黑色.

在每个亚补丁上, 将灰度信息投影到本文采样模式上, 操作与 BRISK[13] 或 FREAK[15] 类似. 然后, 在每个映射后的采样模式上, 采用与 BRIEF[12] 一样的随机测试, 就会产生一串二进制位. 顺序连接所有亚补丁对应的位串, 就得到了原图像补丁的二进制描述子, 如图 7 所示. 图 7 中, 补丁所有像素点按灰度值分成 5 类, 即 5 个亚补丁, 依次串联这 5 个的亚补丁的二进制位串, 就得到了对关键点具有独特性描述的二进制描述子.

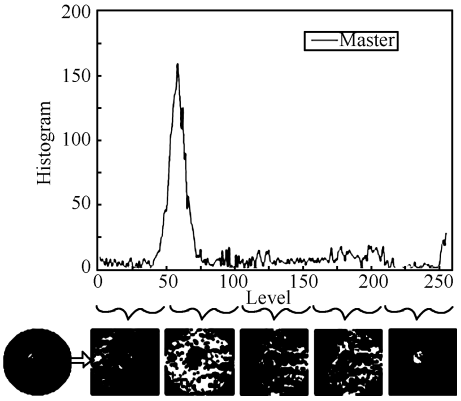


图 6 利用灰度值排序对补丁像素点分类

Fig. 6 Sorting pixels by their intensity order

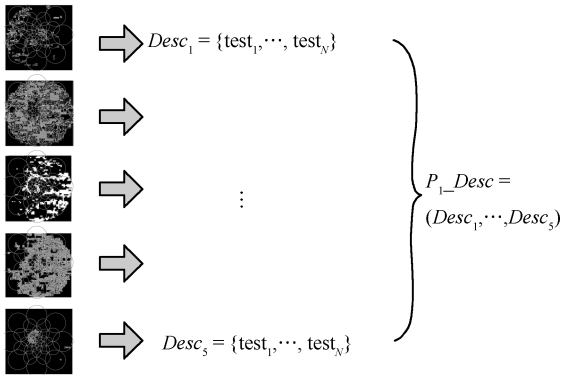


图 7 二进制描述子生成过程

Fig. 7 The process of building binary descriptor

3.1 二进制描述子生成

要将原图像的一个补丁 P 分解为多个亚补丁, 需将补丁像素点按灰度值大小进行分类. 像素点的灰度值范围是 $0 \sim 255$, 可以将这一范围等分成 M 个灰度段 I_i $(0 \le i < M)$, 于是每个灰度段的灰度值范围可以表示为: $(256/M)i \le I_i < (256/M)(i+1)$. 然后, 根据段的灰度值范围, 在补丁里保留相应灰度值范围段的像素点, 去除该范围段外的所有像素点, 这样就得到了该灰度段对应的像素点集合, 即一个亚补丁. 有多少个灰度段就有多少个亚补丁.

接着将本文采样模式映射到亚补丁上. 首先, 将亚补丁关键点的位置作为采样模式中心, 将采样模式最外层同心圆上的任意采样点与中心连线方向与关键点方向对齐. 然后, 按照采样模式采样点位置及其平滑范围的大小, 将亚补丁相应区域的灰度信息映射为采样点的值. 所以, 有多少亚补丁就有多少映射的采样模式 P_i $(0 \le i < M)$.

在采样模式映射的亚补丁上, 随机选择两个采样点进行灰度比较测试[12]:

$$T(P_i; x, y) := \begin{cases} 1, & I(P_i, x) < I(P_i, y) \\ 0, & \text{否则} \end{cases} \tag{4}$$

其中, x, y 是在采样模式 P_i 映射的亚补丁上随机选择的两个采样点, $I(P_i, x)$ 是采样模式 P_i 里采样点 x 的值.

将每次测试结果保存起来, 直到测试 N 次为止. 串联每次测试的结果, 就得到了一个维数为 N 的亚补丁描述子 $Desc_i$:

$$Desc_i := \sum_{1 \le j \le N} 2^{j-1} T(P_i; x_j, y_j) \tag{5}$$

将所有亚补丁描述子依次串联在一起, 就构成了维数为 $N \times M$ 的原图像补丁描述子 P_{Desc}:

$$P_{Desc} := (Desc_1, \cdots, Desc_M) \tag{6}$$

该描述子是对图像补丁 (即关键点) 的特征描述.

3.2 二进制描述子匹配

一个图像上的某个描述子 $CFRBD_1{}^a$, 要在第 2 幅经过转换的图像上找到对应的描述子, 这就需要匹配操作, 本文采用最近邻距离匹配. 计算 $CFRBD_1{}^a$ 与第 2 幅图像所有描述子的汉明距离, 并根据这些距离从小到大的顺序, 对应得到第 2 幅图像所有描述子的一个有序排列, 如 $\{CFRBD_2{}^a, CFRBD_2{}^b, CFRBD_2{}^c, \cdots\}$. 如果这两个汉明距离 $DIS(CFRBD_1{}^a, CFRBD_2{}^a)$ 与 $DIS(CFRBD_1{}^a, CFRBD_2{}^b)$ 满足如下关系, 则认为 $CFRBD_2{}^a$ 为第 2 幅图像上能与 $CFRBD_1{}^a$ 匹配的描述子.

$$DIS(CFRBD_1{}^a, CFRBD_2{}^a) < T \times DIS(CFRBD_1{}^a, CFRBD_2{}^b)$$

其中, T 是控制 $CFRBD_1{}^a$ 与 $CFRBD_2{}^a$ 是否匹配的阈值, 而且这个阈值是非常必要的. $CFRBD_1{}^a$ 在第 2 幅图像对应的描述子是唯一受到转换过程的影响, 即使是 $CFRBD_1{}^a$ 的对应描述子与 $CFRBD_1{}^a$ 的汉明距离也很少为零, 但这个距离比 $CFRBD_1{}^a$ 与其在第 2 幅图像非对应的描述子的距离要小得多. 所以, 要使 $CFRBD_2{}^a$ 为 $CFRBD_1{}^a$ 在第 2 幅图像上的匹配点, 只要二者的汉明距离至

少比 $CFRBD_2{}^b$ 与 $CFRBD_1{}^a$ 的距离小一个数量级即可, 就是由这个阈值控制的. 如果这个阈值太大的话, 会放弃许多原本正确的匹配.

4　实验与评估

4.1　数据和评估标准

本文采用牛津标准图像集评估特征描述子[2]. 这个图像集包括结构场景图像和纹理场景图像, 而且这些图像带有不同程度的几何和光度转换. 这 6 个转换包括: 视角变化、尺度变化及图像旋转、图像模糊、亮度变化和 JPEG 压缩.

由第 3.2 节可知, 如果两个匹配的描述子是真正的两个对应的描述子, 称这对匹配为正确匹配; 否则, 就不是两个真正对应的描述子, 称为错误匹配. 本文使用的文献 [1] 的评估标准称为查全率, 是基于两幅图像的正确匹配数量和错误匹配数量提出的, 又称为查准率 – 查错率. 查准率 (Recall) 指两幅图像正确匹配数量与真正对应描述子总数的比例, 查错率 (1-precision) 指错误匹配数量占所有已匹配数量的比例.

4.2　鲁棒性能评估

为了展示所提描述子的性能, 将所提描述子与 DAISY[5]、FREAK[15]、和 BRISK[13] 做对比实验. 作对比的描述子参数是根据各文献作者实验得出的, 能使其性能达到最好. 本文方法参数包括亚补丁个数和随机测试次数分别设为 5 和 20, 即描述子维数为 100, 实验证明这种设置描述子性能最好.

为了公平评估每个描述子的性能, 采用相同的关键点检测方法. 本实验采用两种不同的关键点检测方法, 即 Hessian 检测法[2] 和 Harris 检测法[3], 前者检测到的关键点为斑点, 后者检测到的关键点为角点, 而且两种方法检测到的关键点对仿射都不变. 将 4 个描述子分别对两种方法检测的关键点进行测试, 于是每一对图像有 8 组实验. 实验结果如图 8 所示.

由图 8 可知, 尽管不同关键点检测方法对描述子性能影响很大, 但相对不同特征描述方法来说, 检测方法不同对描述子性能的影响较小. 而且, 无论是 Hessian 检测子, 还是 Harris 检测子, 所提描述子的性能都要好于其他的描述子. 这是因为所提描述子包含了图像局部灰度信息、灰度对比信息、灰度排序信息. 灰度对比信息增加了描述子的独特性, 灰度排序信息使描述子对视角和选择变化具有鲁棒性. 总之, 本文描述子性能好, 要归功于好的采样模式设计、图像补丁进行分层处理、以及好的参数选择.

4.3　实时性评估

实时性是检验一个描述子算法性能的重要方面. 本次实验给出 4 种算法 (CFRBD、FREAK、BRISK 和 DAISY) 实时性对比测试结果. 对比的时间数据包括从对比项中的时间指关键点检测、描述过程的时间, 可以把这一处理过程时间再细致划分, 具体包括所有关键点检测时间、所有关键点描述时间、所有关键点处理的总时间和每个关键点处理的平均时间. 本文选择的关键点检测子为 AGAST[24]. 实验均在 Ubuntu-10.04 (32 位) 环境 (虚拟机, 512 MB 内存, Intel Core Duo-2.2 GHz) 中完成. 所有方法的测试选用的图像是 Graffiti 图片集里的第 4 幅图像, 图像大小都为 800 像素 × 640 像素, 均为 PPM 格式. 测试结果如表 1 所示. 由表 1 可知, 本文方法在时间上可与 FREAK 和 BRISK 相媲美, 比 DAISY 快很多倍, 因为 DAISY 为浮点描述子.

表 1　关键点检测描述时间对比结果 (Graffiti-4)

Table 1　The results of time comparisons of detecting and describing all key-points of Graffiti-4

	DAISY	BRISK	FREAK	CFRBD
关键点数量 (个)	1 373	1 050	1 036	1 053
检测所有关键点时间 (ms)	1 548.3	17.1	18.97	18.22
描述所有关键点时间 (ms)	1 486.1	22.06	28.12	18.43
所有关键点处理总时间 (ms)	3 034.4	39.16	47.09	36.65
平均每个关键点处理时间 (ns)	2 210.1	37.3	45.45	34.81

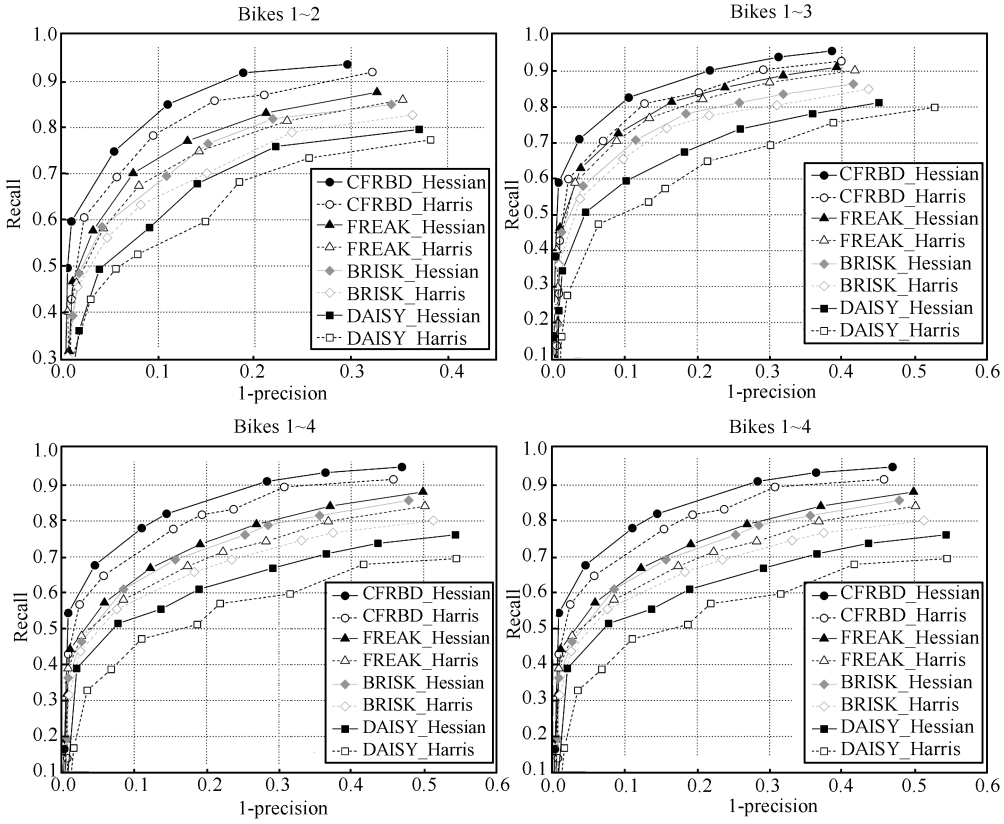
为了测试所提描述子匹配的实时性, 同样做了对比测试. 测试环境和对比对象同上. 对比的时间数据包括两个图像关键点匹配总时间和每对描述子平均匹配时间. 测试图像采用 Graffiti 的第 2 幅和第 4 幅图像. 测试结果如表 2 所示. 由表 2 可知, 所提描述子匹配的实时性是最好的. 总之, 所提方法从关键点描述到匹配都很快, 所以能满足实时识别目标的要求.

表 2　匹配实时性对比结果 (Graffiti-2 Vs. Graffiti-5)

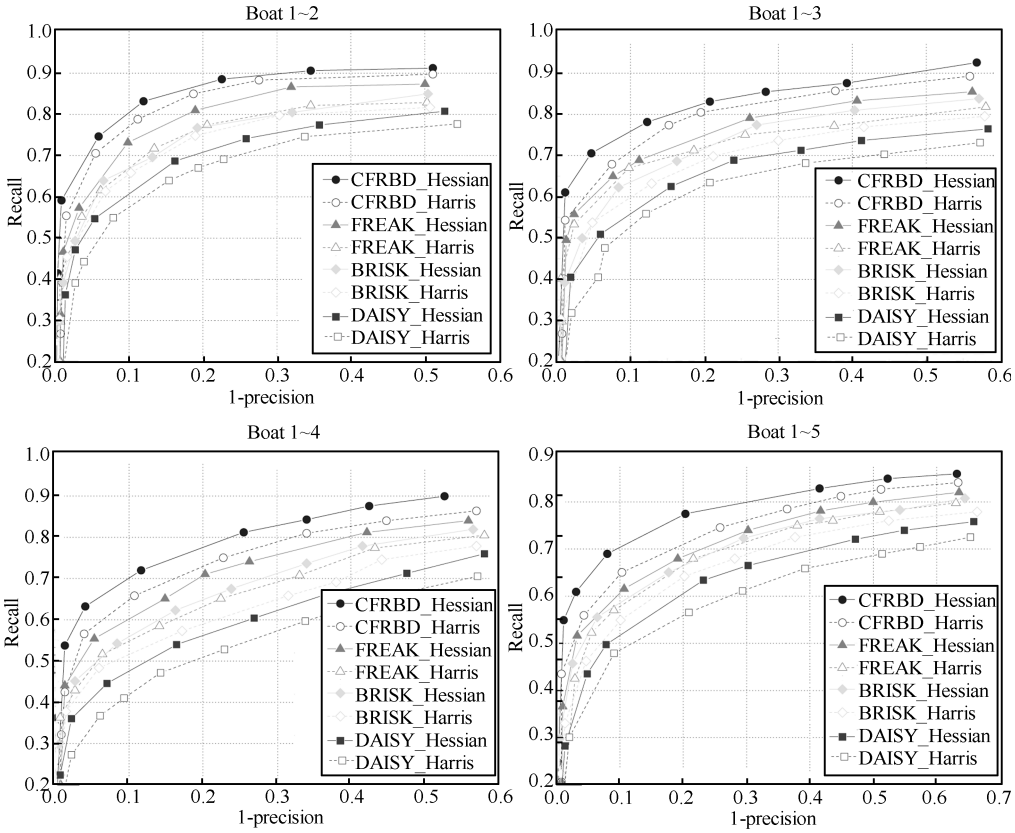
Table 2　The results of real-time performance comparisons of matching key-points in Graffiti-2 to that in Graffiti-5

	DAISY	BRISK	FREAK	CFRBD
Graffiti-2 图像关键点数量 (个)	1 468	1 064	1 017	1 063
Graffiti-5 图像关键点数量 (个)	1 984	1 259	1 160	1 302
两图像关键点匹配总时间 (ms)	246.16	32.53	34.68	30.41
每对描述子平均匹配时间 (ns)	68.44	22.23	24.52	20.12

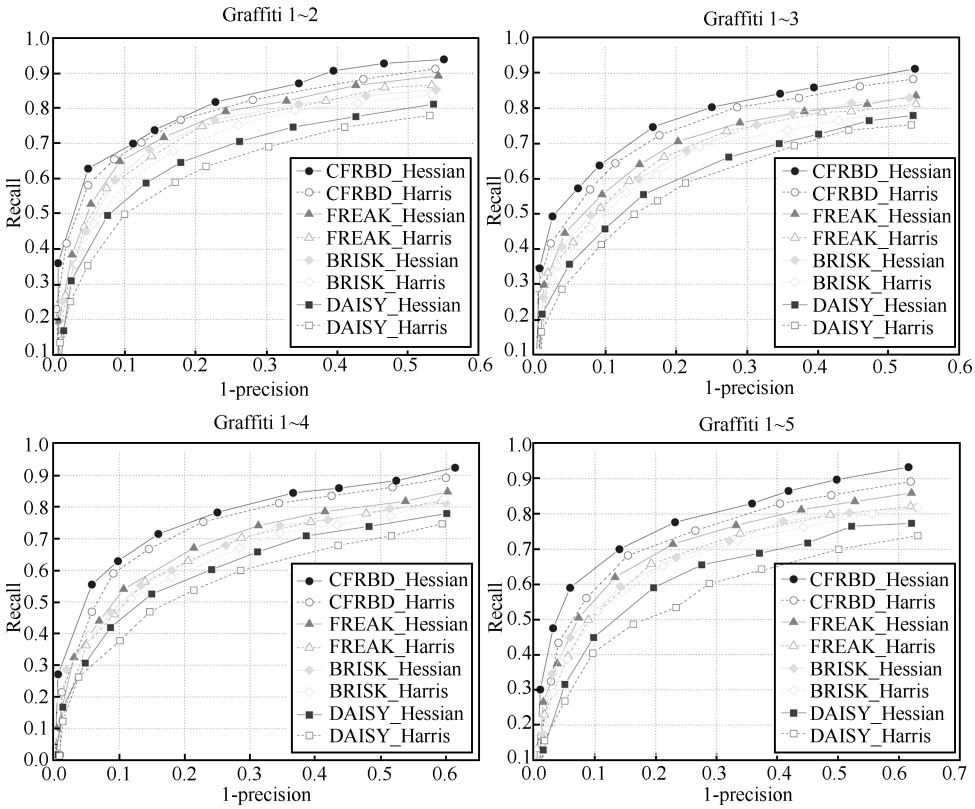
[2] http://www.robots.ox.ac.uk/vgg/research/affine/



(a) 图像模糊变化下的匹配结果
(a) Performance of blurred transform

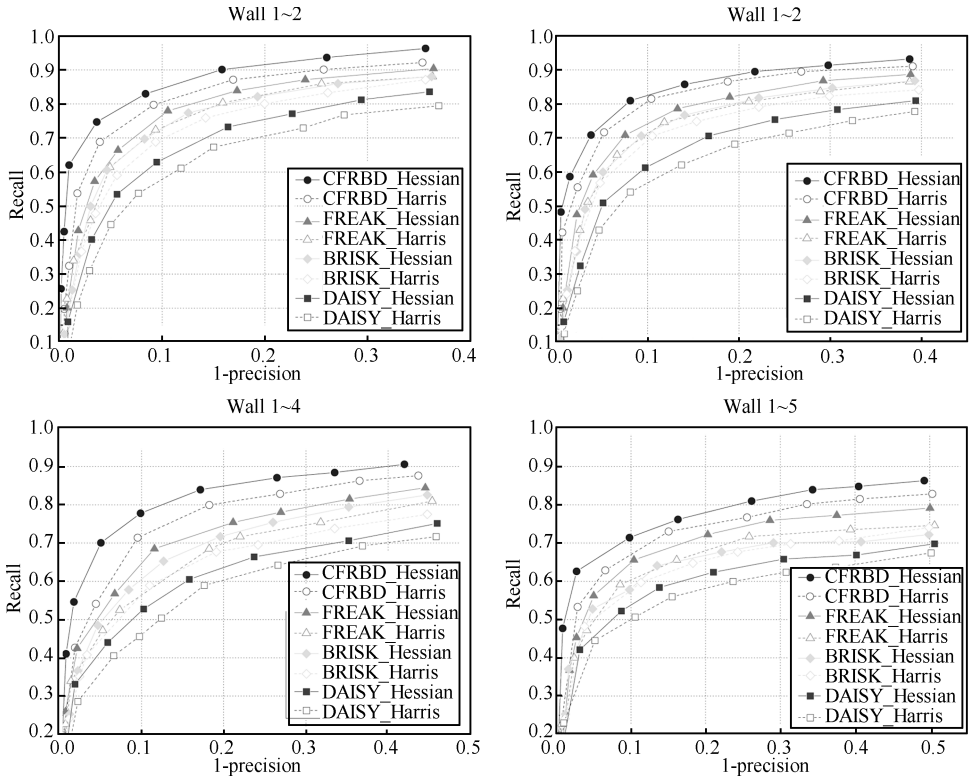


(b) 旋转和尺度变化下的匹配结果
(b) Performance in scale and rotation changes



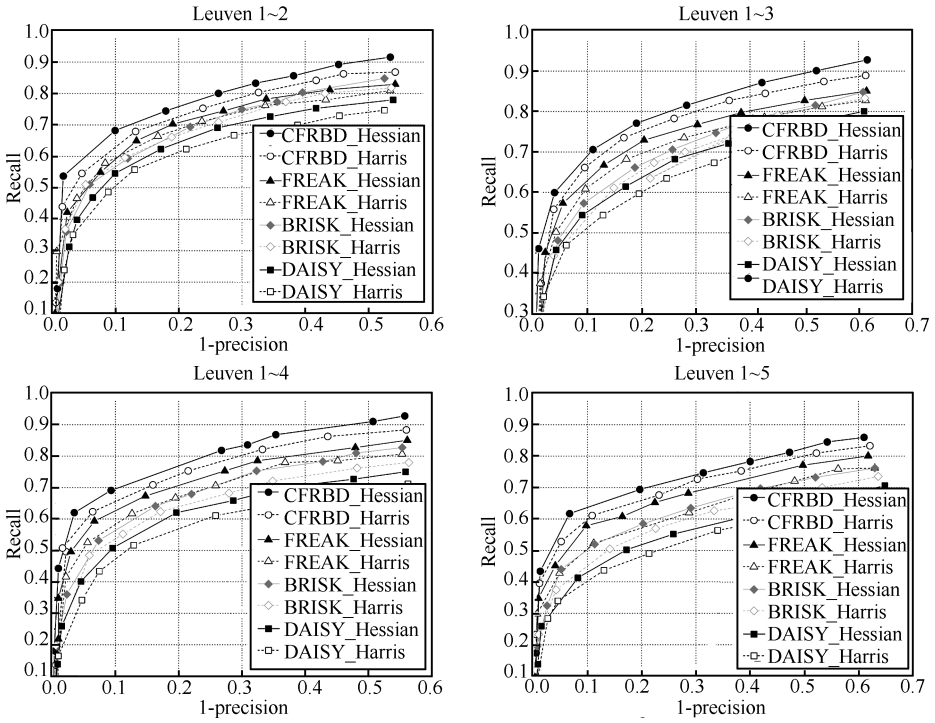
(c) 结构图像在视角变化下的匹配结果

(c) Viewpoint change of structured image



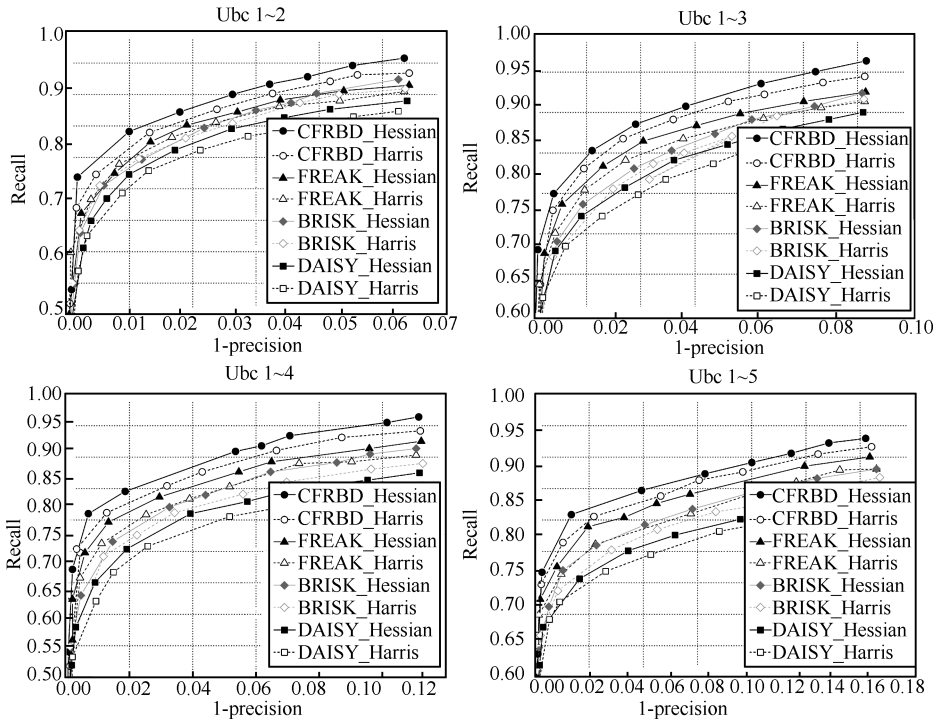
(d) 纹理图像在视角变化下的匹配结果

(d) Viewpoint change of textured image



(e) 亮度变化下点的匹配结果

(e) Performance of illumination changes



(f) JPEG 压缩变化下的匹配结果

(f) Performance of JPEG compression

图 8 各种图像转换下的匹配结果

Fig. 8 Matching effect under various image transformations

5 结论

本文进一步挖掘图像补丁的特征信息, 提出了一种具有更高鲁棒性的二进制描述子. 该二进制描述子是基于最优的采样模式、灰度排序技术和灰度比较测试技术构建起来的. 首先, 分析最新采样模式的特点, 设计了一种独特的、最优化的采样模式. 其次, 介绍了将图像补丁分解生成若干个亚补丁, 及二进制描述子生成和匹配. 最后, 通过对比实验客观地对所提描述子的性能进行评估. 通过实验发现, 在所有对比的描述子中, 所提的描述子在各种图像转换条件下性能是最好的. 而且所提的描述子具有计算简单、匹配速度快、占用内存少等优点. 下一步工作将此种方法引入到 RGB 图像, 构建 RGB 图像的二进制描述子.

References

1 Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1615−1630

2 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91−110

3 Bay H, Ess A, Tuytelaars T, van Gool L. SURF: speeded up robust features. *Computer Vision and Image Understanding*, 2008, **110**(3): 346−359

4 Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, D. C., USA: IEEE, 2004. II-506−II-513

5 Tola E, Lepetit V, Fua P. Daisy: an efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(5): 815−830

6 Fan B, Wu F C, Hu Z Y. Aggregating gradient distributions into intensity orders: a novel local image descriptor. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, Providence, USA: IEEE, 2011. 2377−2384

7 Torralba A, Fergus R, Weiss Y. Small codes and large image databases for recognition. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA: IEEE, 2008. 1−8

8 Jégou H, Douze M, Schmid C. Improving bag-of-features for large scale image search. *International Journal of Computer Vision*, 2010, **87**(3): 316−336

9 Salakhutdinov R, Hinton G. Learning a nonlinear embedding by preserving class neighbourhood structure. In: Proceedings of the 2007 International Conference on Artificial Intelligence and Statistics. Monte Carlo Resort, Las Vegas, Nevada, USA: IEEE, 2007. 412−419

10 Ozuysal M, Calonder M, Lepetit V, Fua P. Fast keypoint recognition using random ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(3): 448−461

11 Lepetit V, Fua P. Keypoint recognition using randomized trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(9): 1465−1479

12 Calonder M, Lepetit V, Ozuysal M, Trzcinski T, Strecha C, Fua P. BRIEF: computing a local binary descriptor very fast. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(7): 1281−1298

13 Leutenegger S, Chli M, Siegwart R. BRISK: binary robust invariant scalable keypoints. In: Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 2548−2555

14 Rublee E, Rabaud V, Konolidge K, Bradski G. ORB: an efficient alternative to SIFT or SURF. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 2564−2571

15 Alahi A, Ortiz R, Vandergheynst P. FREAK: fast retina keypoint. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, Providence, USA: IEEE, 2012. 2069−2076

16 Trzcinski T, Lepetit V. Efficient discriminative projections for compact binary descriptors. In: Proceedings of the 12th European Conference on Computer Vision. Firenze, Italy: IEEE, 2012. 228−242

17 Strecha C, Bronstein A, Bronstein M, Fua P. LDAHash: improved matching with smaller descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(1): 66−78

18 Trzcinski T, Christoudias M, Fua P, Lepetit V. Boosting binary keypoint descriptors. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Washington, D. C., USA: IEEE, 2013. 2874−2881

19 Wang Z H, Fan B, Wu F C. Local intensity order pattern for feature description. In: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 603−610

20 Liu Ping-Ping, Zhao Hong-Wei, Zang Xue-Bai, Dai Jin-Bo. A fast local feature description algorithm. *Acta Automatica Sinica*, 2010, **36**(1): 40−45

(刘萍萍, 赵宏伟, 臧雪柏, 戴金波. 一种快速局部特征描述算法. 自动化学报, 2010, **36**(1): 40−45)

21 Tang Yong-He, Lu Huan-Zhang, Hu Mou-Fa. Local feature description algorithm based on Laplacian. *Optics and Precision Engineering*, 2011, **19**(12): 2999−3006
(唐永鹤, 卢焕章, 胡谋法. 基于 Laplacian 的局部特征描述算法. 光学精密工程, 2011, **19**(12): 2999−3006)

22 Ge Juan, Cao Wei-Guo, Zhou Wei, Gong Ming, Liu Liu, Li Hua. A local feature descriptor under color affine transformation. *Journal of Computer-Aided Design & Computer Graphics*, 2013, **25**(1): 26−33
(葛娟, 曹伟国, 周炜, 公明, 刘浏, 李华. 一种颜色仿射变换下的局部特征描述子. 计算机辅助设计与图形学学报, 2013, **25**(1): 26−33)

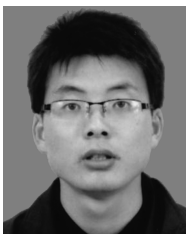
23 Song Ke-Chen, Yan Yun-Hui, Chen Wen-Hui, Zhang Xu. Research and perspective on local binary pattern. *Acta Automatica Sinica*, 2013, **39**(6): 730−744
(宋克臣, 颜云辉, 陈文辉, 张旭. 局部二值模式方法研究与展望. 自动化学报, 2013, **39**(6): 730−744)

24 Mair E, Hager G D, Burschka D, Suppa M, Hirzinger G. Adaptive and generic corner detection based on the accelerated segment test. In: Proceedings of the 11th European Conference on Computer Vision. Berlin, Heidelberg: Springer-Verlag, 2010. 183−196

25 Cao Ying, Miao Qi-Guang, Liu Jia-Chen, Gao Lin. Advance and prospects of AdaBoost algorithm. *Acta Automatica Sinica*, 2013, **39**(6): 745−758
(曹莹, 苗启广, 刘家辰, 高琳. AdaBoost 算法研究进展与展望. 自动化学报, 2013, **39**(6): 745−758)

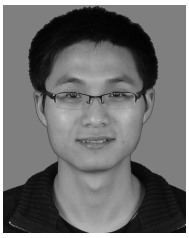
26 Brown M, Hua G, Winder S. Discriminative learning of local image descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(1): 43−57

27 Cai H P, Mikolajczyk K, Matas J. Learning linear discriminant projections for dimensionality reduction of image descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(2): 338−352



惠国保 南京理工大学机械工程学院博士研究生. 主要研究方向为图像配准和模式识别. 本文通信作者.
E-mail: huiguobao@163.com
(**HUI Guo-Bao** Ph. D. candidate at the School of Mechanical Engineering, Nanjing University of Science and Technology. His research interest covers image registration and pattern recognition. Corresponding author of this paper.)

李东波 南京理工大学机械工程学院教授. 主要研究方向为先进制造技术.
E-mail: db_calla@yahoo.com.cn
(**LI Dong-Bo** Professor at the School of Mechanical Engineering, Nanjing University of Science and Technology. His main research interest is advanced manufacture technology.)



童一飞 南京理工大学机械工程学院副研究员, 博士. 主要研究方向为网格制造技术. E-mail: tyf51129@yahoo.com.cn
(**TONG Yi-Fei** Ph.D., associate professor at the School of Mechanical Engineering, Nanjing University of Science and Technology. His main research interest is grid manufacture technology.)