

仿视网膜采样的二进制描述子

袁庆升^{1,3,4}, 靳国庆², 张冬明^{1,4}, 包秀国^{1,4}

(1. 国家计算机网络应急技术处理协调中心, 北京 100029; 2. 中国科学院计算技术研究所, 北京 100190;
3. 中国科学院大学网络空间安全学院, 北京 100049; 4. 中国科学院信息工程研究所, 北京 100193)

摘要: 现有二进制描述子生成采用随机点对或均匀采样方式, 顽健性弱、计算复杂。针对这一问题, 提出了一种模仿人眼视网膜特性的采样模式 (RBS), 首先通过设计采样密度、多尺度平滑、视野重叠等采样方法来模仿视网膜神经节细胞层 (ganglion cell layer), 也称为视神经层, 将光信号转换为视信息的方式, 再通过对典型数据学习来选择特征点对, 最后使用区块均值代替单像素点计算点对比较值, 生成顽健的紧致二进制描述子。在 Mikolajczyk 提出的数据集上进行了实验, 实验结果表明, 128 bit 的 RBS-128 相对于 512 bit 的 FREAK 和 BRISK 正确率分别提升 16.4% 和 5.3%。

关键词: 二进制描述子; 仿视网膜采样; 神经节细胞; 比较点对

中图分类号: TP37

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2019021

Retina-imitation sampling based binary descriptor

YUAN Qingsheng^{1,3,4}, JIN Guoqing², ZHANG Dongming^{1,4}, BAO Xiuguo^{1,4}

1. National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing 100029, China

2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

3. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

4. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100193, China

Abstract: The existing binary descriptors, generated from random or uniform point pairs sampling, suffer from low robustness and high computation. A novel sampling method, named RBS (retina-imitation based sampling), was proposed, which combines different densities sampling, multi-scale smoothing and reception field overlapping to imitate the converting from light signal to vision of ganglion cells of human retina cells, and further selects most discriminative comparison pairs based on learning on training data. Finally, compact binary descriptor was generated based on comparisons between the neighbor mean instead of single sampled point. The experimental results show the RBS-128 with 128 bit outperforms FREAK and BRISK with 512 bit about 16.4% and 5.3% in precision on the dataset provided by Mikolajczyk.

Key words: binary descriptor, retina-imitation based sampling, ganglion cell, comparison pairs

1 引言

随着互联网图像数量迅速增长, 对海量图像进行内容分析与检测的需求越来越大, 基于内容的图像检索技术成为研究热点。这些技术通常使用局部

特征作为图像内容的描述, 这是由于局部特征具有良好的区分性, 对图像的多种变换, 如遮挡、模糊、噪声、剪切等, 具有较高的稳定性。局部特征的提取主要分为 2 个步骤: 1) 特征点的提取, 提取的信息包括特征点位置、特征点主方向、特征点尺度信

收稿日期: 2018-03-13; 修回日期: 2018-12-01

通信作者: 张冬明, dmzhang@ict.ac.cn

基金项目: 国家自然科学基金资助项目 (No.61672495, No.61273247); 国家重点研发专项基金资助项目 (No.2016YFB0801203, No.2016YFB0801200)

Foundation Items: The National Natural Science Foundation of China (No.61672495, No.61273247), The National Key Research and Development Program of China (No.2016YFB0801203, No.2016YFB0801200)

息等；2) 局部特征点描述子的计算，即依据特征点周围的像素信息，计算其对旋转、光照、尺度等变化顽健的特征，提取到的特征通常以向量表示，能有效描述特征点。

常见的局部特征点描述子主要分为 2 种：实数描述子和二进制描述子。实数描述子提取复杂，在内存和磁盘上以浮点数的形式存储，对存储要求高且不能满足实时性需求。二进制特征描述子使用二进制位进行编码，计算简单且占用的存储空间小，适合实时性要求高的应用场景。常用的二进制描述子主要有 BRIEF (binary robust independent elementary features)^[1]、ORB (oriented FAST and rotated BRIEF)^[2]、BRISK (binary robust invariant scalable keypoint)^[3]、FREAK (fast retina keypoint)^[4]等。

BRIEF 通过随机选择点对而得到二进制描述子，计算速度快^[1]，但是其对旋转和尺度变化敏感，在不同数据集下检索性能差异大。BRISK 抛弃了随机采样模式，提出类似 Daisy^[5]的采样方式，以特征点为圆心，在半径依次增大的同心圆上均匀采样候选点，同心圆上候选点采样密度从中心开始均匀下降，这在一定程度上增强了对噪声的顽健性。FREAK 则模仿视网膜特性在 BRISK 技术上进一步引入了多尺度平滑^[4]。李兵等对 BRIEF 进行改进，不同于 ORB 特征使用区域质心计算主方向，该方法通过在特征点邻域内均匀采样 8 对对称子区域，然后累加各子区域以其像素均值加权的梯度方向作为主方向，从而实现了描述子对旋转的不变性^[6]。朱英宏等通过提取特征点邻域内 33 个像素点的二进制纹理特征，得到一个 132 bit 的二进制描述子^[7]。卢鸿波等在特征点邻域内选取 R 个同心圆，每个圆上采样 N 个像素点，然后使用同心圆层内编码和层间编码的方式得到一个 $2RN$ bit 的二进制描述子^[8]。惠国保等研究了不同采样模式和不同高斯平滑半径对二进制描述子的影响，提出了一种内密外疏的采样模式^[9]。

此外，近年来研究者们还提出了很多基于有监督学习的二进制特征生成方法。RFD (receptive fields selection) 特征^[10]对多个子区域的实值特征进行量化构建二进制特征，子区域的选择和量化阈值通过学习得到。BinBOOST^[11]特征则是通过学习多个弱分类器组合成一个强分类器，用分类器来判断二进制位取值为 0 或 1，采用的弱分类器形式是量化函数。RFD 特征和 BinBOOST 特征在本质上都是通过有监督学习的方式对实值特征进行量化来得

到二进制描述子。随着深度学习卷积神经网络 (CNN, convolution neural network)^[12]在图像分类上快速发展，也出现了许多基于深度学习提取二进制描述子的算法^[13-15]，这类方法不同于上述局部的二进制描述子，获得的是图像的全局二进制描述，缺少描述图像底层细节的有用信息，且计算复杂度高，提取描述子时需要占用大量内存，效率低。

本文综合考虑采样点选择和描述子生成，基于视网膜的视神经细胞分布和感知特性，研究二进制描述子生成方法。特别地，针对特征点对顽健性低的问题，提出了模仿视网膜特性的采样模式。对比同样模仿视网膜特性的 FREAK^[4]描述子，本文方法二进制描述子的生成过程不同。本文所提出的采样模式混合使用多尺度平滑、视野重叠等模仿视网膜特性，且通过数据实验对关键参数进行优化，同时还提出了基于学习的方法，在典型数据集上对最终采样点对进行选择。所提方法具有计算快速、占用内存少、顽健性高的优点，且匹配正确率有明显提高。

2 基于 RBS 的二进制描述子

图像局部二进制特征提取包括特征点检测和描述子提取。特征点检测往往需要在多个尺度空间上进行极值点检测，以应对图像的尺度变化，保证特征点的可重复性。描述子提取是在特征点检测的基础上进行的，通常依据特征点周围的像素分布按照预设计算模型生成一串特征数值，而二进制描述子一般通过比较像素点之间亮度的大小来确定描述子取值，如式(1)所示。

$$D = \sum_{0 \leq i < l} 2^i T(P_i) \quad (1)$$

其中， l 为比较点对的数目，它决定二进制描述子的最终长度；点对 P_i 由 2 个像素点 P_i^1 和 P_i^2 定义； $T(P_i)$ 表示点对 i 的比较结果，定义如式(2)所示。

$$T(P_i) = \begin{cases} 1, & I(P_i^1) \leq I(P_i^2) \\ 0, & I(P_i^1) > I(P_i^2) \end{cases} \quad (2)$$

其中， $I(\cdot)$ 表示对应点的灰度值。

由式(1)可以看出，二进制描述子完全取决于点对的选择，相应地，描述子的区分能力和顽健性完全依赖特征点周围的点对采样模式。选择好的采样模式，一方面可以减少描述子长度，即降低 l 的取值，进而降低计算量和内存开销；另一方面由此产生的

二进制描述子对各种常见的变化具有很好的顽健性，即不随图片尺度、方向及视角等变化而改变。本文根据视网膜的细胞分布特性，设计了相应的候选点选取方式和点对采样模式，从而获得高性能的 RBS (retina-imitation based sampling) 特征。下面详细介绍候选点选取方式、点对采样模式以及 RBS 特征的生成方式。

2.1 候选点选取

候选点选择方面，ORB 和 BRIEF 采用随机采样，BRISK 采用均匀采样，FREAK 使用类视网膜采样，均取得了较大的性能提升。视网膜由 3 层神经元组成。第一层神经元是视细胞层，包括锥细胞和柱细胞，负责感光。第二层是双节细胞，数量为十至数百个，负责联络作用，视细胞通过双节细胞与一个神经节细胞相联系。第三层是神经节细胞层，将光信号转化为视觉信号，是实际的视觉神经，所以也称为视神经层。神经科学的最近研究表明视觉神经细胞的分布并非均匀分布，在视网膜中心视觉神经细胞密度并不是很高，而在中心周围密度急剧上升，再往外密度则快速下降，如图 1 所示^[16]。视网膜中心位置叫做视网膜小凹，其中只有很少的细胞，但在这个凹槽周围分布有大量的细胞，细胞数量与其到中心小凹的距离成反比。此外，视神经层还具有类似高斯差分机制的感知特性，抽取信息的神经节细胞位于视网膜外围，而形成高层视觉信息的细胞位于视网膜中心。

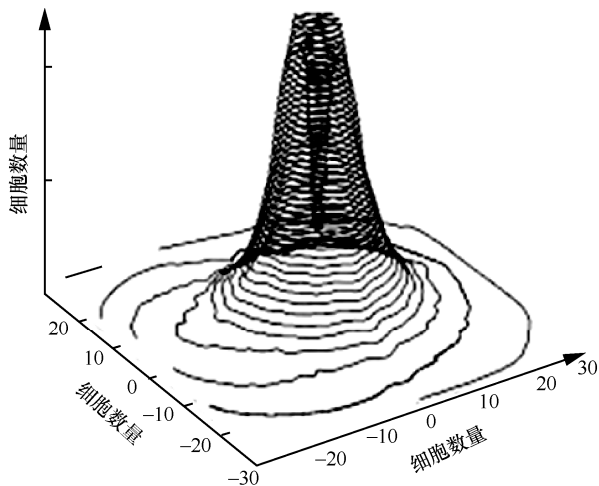


图 1 人眼视网膜视神经细胞分布

本文模仿视网膜视神经层细胞的分布特性，设计了一种新型的候选点采样模式。模拟视神经层的细胞具有“低-高-低”的密度分布特性，通过实验

选择，确定对特征点采用 5 层采样。具体地，以特征点为候选点中心，在中心外围第一层候选点的密度比较低，在第二层候选点的密度最高，第三层~第五层候选点的密度呈指数下降。如图 2 所示，中间点为特征点，从内向外每层采样点数分别为 4、24、12、8、4。本文将这种采样模式称为 RBS。FREAK 采样点设置为 6:6:6:6:6:6，共 7 层 42 个采样点；RBS 采样点设置为 4:24:12:8:4，共 5 层 52 个采样点。通过比较可知，RBS 的特点如下：一方面采样点数增加了 10 个，增长率为 23.8%；另一方面突出特征点周围像素点对描述子的作用，分布更贴合视网膜的细胞“低-高-低”的密度分布。

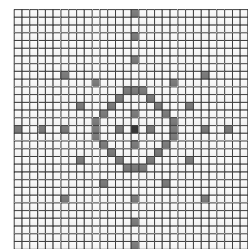


图 2 候选点分布

由于二进制描述子是基于像素点之间的亮度比较，所以对噪声较为敏感，为了降低噪声的影响，对候选点进行高斯模糊已成为通行做法。对于不同候选点，BRIEF 和 ORB 使用相同的高斯模糊半径；而 BRISK 和 FREAK 使用不同的模糊半径，更符合人眼视觉特性。本文模仿视神经的感知特性，采用不同大小的高斯模糊半径，中心模糊半径最小，外围候选点模糊半径逐渐增大。通过匹配正确率实验对比，本文设定初始模糊半径为 2，层间相差为 2。此外，本文还设计特殊的模糊范围来改善模糊效果，同一层的模糊范围保证一定的重叠，第二层和第三层的重叠面积最大，各层的重叠面积满足“少-多-少”的规律。经实验，选定比例为 1:4:3:2:1，进一步提高平滑效果，多级高斯模糊如图 3 所示。

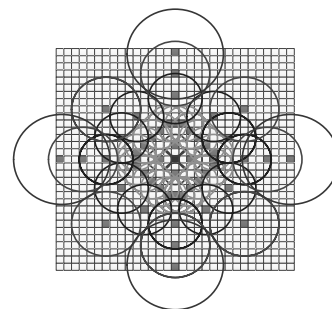


图 3 多级高斯模糊示意

2.2 点对采样模式

接下来研究如何从潜在的大量候选点对中选出最佳的比较点对。假设有 n 个候选点，两两组合一共有 C_n^2 个候选点对。为了提高特征顽健性，同时降低特征维数，必须进一步筛选以去除噪声影响，同时消除强相关的点对。为此，本文提出从训练集中学习筛选比较点对，从而确定最后的采样模式，具体方法如下。

1) 在训练集上提取 r 个特征点，每个特征点计算长度为 $l(l=C_n^2)$ 的描述子。这样组成一个 r 行 C_n^2 列的二进制矩阵。在本文的实验中取 $n=53$ ，则这个矩阵一共有 1 378 列。

2) 在上述矩阵中，每一行表示一个特征点描述子，每一列表示一个候选点对的所有计算结果，点对之间可能存在较强的相关性，导致某些列可能全为 0 或全为 1。而如果某个候选点对的区分性高，则该列的按位均值应接近 0.5。所以本文选取均值最接近 0.5 的那一列记为 c ，把其所对应的候选点对加入最终的比较点对集合。

3) 对剩下的所有列按照其均值与 0.5 的接近程度排序，计算各列与列 c 的余弦距离，选择余弦距离大的列，把对应点对加入最终点对集合。重复此步骤，直到选出 M 组点对。

通过上述方法选出的 M 组点对区分性较强、相关性低。可以看出，本文方法和 FREAK 中方法类似，均采用构建特征矩阵、降维方式，来获取点对中对提升描述子区分力作用最显著的点对，本质上是一种统计学习方法。但 RBS 得益于本文设计的采样模式，可以消除更多的冗余点对，构建更短的二进制描述子，并构建了 32、64、128、160 bit 的描述子以适应不同应用场景，分别记为 RBS-32、RBS-64、RBS-128、RBS-160。图 4 中列出了当 M 设置为 32 时选择的部分点对。结果显示点对选择并没有像 FREAK 那样呈现明显的对称规律，这一方面是由于 32 组点对数量偏少，另一方面也表明 FREAK 的点对中可能存在较大冗余。

2.3 二进制描述子

为了使生成的二进制描述子具有旋转不变性，在提取特征点的二进制描述子之前，首先将特征点邻域旋转至特征点主方向上，这里使用特征点邻域的亮度质心来计算特征点主方向。特征点邻域的亮度质心 C 的具体计算如式(3)~式(6)所示，其中 $I(x,y)$ 表示点 (x,y) 上像素亮度值。

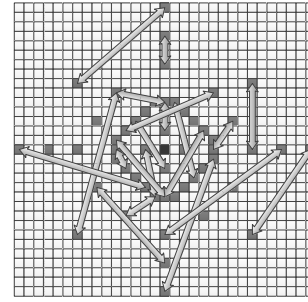


图 4 比较点对示意

$$M_{00} = \sum_{x=-r}^r \sum_{y=-r}^r I(x,y) \quad (3)$$

$$M_{10} = \sum_{x=-r}^r \sum_{y=-r}^r xI(x,y) \quad (4)$$

$$M_{01} = \sum_{x=-r}^r \sum_{y=-r}^r yI(x,y) \quad (5)$$

$$C = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right) \quad (6)$$

特征点中心 O 与质心 C 的连线方向 \overrightarrow{OC} 即是特征点的主方向。

传统的描述子计算按照式(1)和式(2)进行计算，由于只是像素点灰度的比较，所以噪声对其影响比较大，为了提高顽健性，本文使用像素点及其周围 8 个邻域像素的信息作为比较信息量，即

$$T(P_i) = \begin{cases} 1, & AI(P_i^1) \leq AI(P_i^2) \\ 0, & AI(P_i^1) > AI(P_i^2) \end{cases} \quad (7)$$

其中， $AI(\cdot)$ 表示 9 个像素点（8 个邻域像素点以及自身）灰度的均值。按照 2.2 节中介绍的点对采样模式进行采样后，按照式(7)进行描述子计算，获取该像素点的二进制描述子，通过该方法可以进一步提高描述子对噪声的顽健性。

对比同样基于视网膜特性的 FREAK，本文方法取得两方面改进：1) 通过设计更符合视网膜特性的采集模式和典型数据的学习进行点对选择，大幅降低了点对数目，提高了二进制描述子的紧致性；2) 对模糊半径、模糊范围等关键参数进行优化，设计了更符合视网膜特性的参数模型，在生成二进制描述子时，使用区块均值代替单个像素点，顽健性更高。

3 实验结果与分析

3.1 数据集与测试环境

为了测试描述子 RBS 的顽健性，本文使用专用

描述子测试集进行评测，该测试集由 Mikolajczyk 等^[17]提出。这个数据集由 8 张图片组成，每张图片有 5 张变体，覆盖了视角变化 (wall 和 graffiti)、压缩退化 (ubc)、光照变化 (leuven)、旋转和焦距变化 (boat 和 bark)、图像模糊 (tree 和 bike) 这些常见的图像变化，第一张变体变化最小，第五张变化最大。另外，这个数据集还提供了变体与原图之间的单应变化矩阵，方便计算匹配正确率。

进行测试的硬件平台为 Dell 台式机，其指标如下：CPU 为 Intel(R) Core(TM) i7-4770 @ 3.40 GHz，内存为 8 GB，操作系统为 Ubuntu 14.04。

实验中，提取二进制描述子前需要进行特征点检测。Rostenand 等通过制定特征点规则并进行机器学习的方式得到了 FAST (feature from accelerated segment test) 特征检测子^[18]，使用极大值抑止的方法筛选特征点。针对检测速度慢的问题，Mair 等^[19]提出了 AGAST (adaptive and generic corner detection based on the accelerated segment test) 检测子，BRISK 进一步提出了一个多尺度的 AGAST 检测子，他们以 fast 得分为衡量指标在不同的尺度空间中寻找极大值，从而找到顽健的特征点。本文综合考虑稳定性和检测速度因素，使用 AGAST 检测子进行测试。

3.2 测试指标

本文使用 2 个指标来衡量不同方法的性能，即描述子平均计算时间和描述子正确匹配率。描述子平均计算时间是多次计算时长的平均。描述子正确匹配率按以下步骤计算。

- 1) 分别计算原图和变体的特征点和描述子。
- 2) 判断原图和变体的描述子是否匹配，得到总匹配数 N 。
- 3) 判断每个匹配是否正确，得到正确匹配数 N_c 。
- 4) 计算匹配率 $r = \frac{N_c}{N}$ 。

其中，使用文献[17]中介绍的最近邻比率匹配方法来判断描述子是否匹配。假设变体中的描述子 D_B 和 D_C 分别为与原图中的描述子 D_A 海明距离最小和次小的描述子，如果 $\frac{\|D_B - D_A\|}{\|D_C - D_A\|} < t$ ，则认为 2 个描

述子匹配，这里 t 取值为 $\frac{1}{2}$ 。另外，本文使用数据集提供的单应变化矩阵来寻找正确的匹配点。

3.3 实验结果

首先，测试了 RBS 描述子的距离分布，即匹配

描述子之间的距离和不匹配描述子之间的距离分布。对数据集内每张图片及其变体图片使用 AGAST 检测子提取 1 000 个特征点，并计算对应的 RBS-128 描述子，进而计算特征点之间的海明距离，图 5 列出了 wall 图与其他 3 个变体的特征点对的距离分布情况，其中，海明距离单位为 bit，频率是指对应距离点所占比例。可以看出，匹配点的方差和均值都在不断增大。

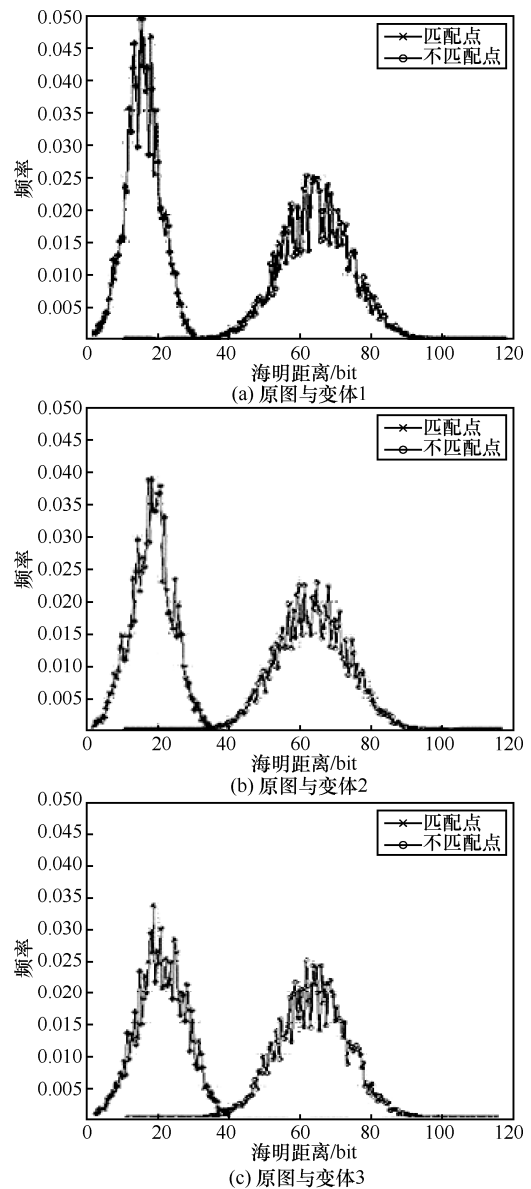


图 5 wall 图片与其不同变体特征点距离分布情况

如图 5 可知，不匹配点之间的距离大致呈正态分布，均值约为 64 bit。匹配点的距离分布也呈正态分布，对于不同的变体图片，匹配点的距离均值不同，第一张图变化和距离均值最小，最后一张图变化和距离均值最大。随着变化程度加剧，匹配点之间的海明

距离变化范围逐渐增大，这符合实际情况。

接下来，对不同描述子匹配正确率进行比较，对比对象包括 SIFT (scale-invariant feature transform)^[20]、ORB、BRISK、BRIEF、FREAK 这 4 种描述子，其中 SIFT 是最流行的实数描述子，其余 3 个是主流的二进制描述子。程序采用了 OPENCV 中的开源代码或者作者提供的代码，均使用默认参数。对数据集中 wall、graffiti、ubc、boat、tree、leuven

的匹配正确率对比结果如图 6 所示，具体数据如表 1 所示。

通过分析结果可以发现，SIFT 平均正确率最高，尤其在有大视角变化和焦距变化的图片 (graffiti 和 boat) 下。4 种二进制描述子中，随机采样模式 (BRIEF 和 ORB) 性能低于特定模式采样 (BRISK、FREAK 和 RBS)。RBS-128 分别超过 BRIEF 和 ORB 约 8.6%和 4.1%，RBS-64 则分别超过

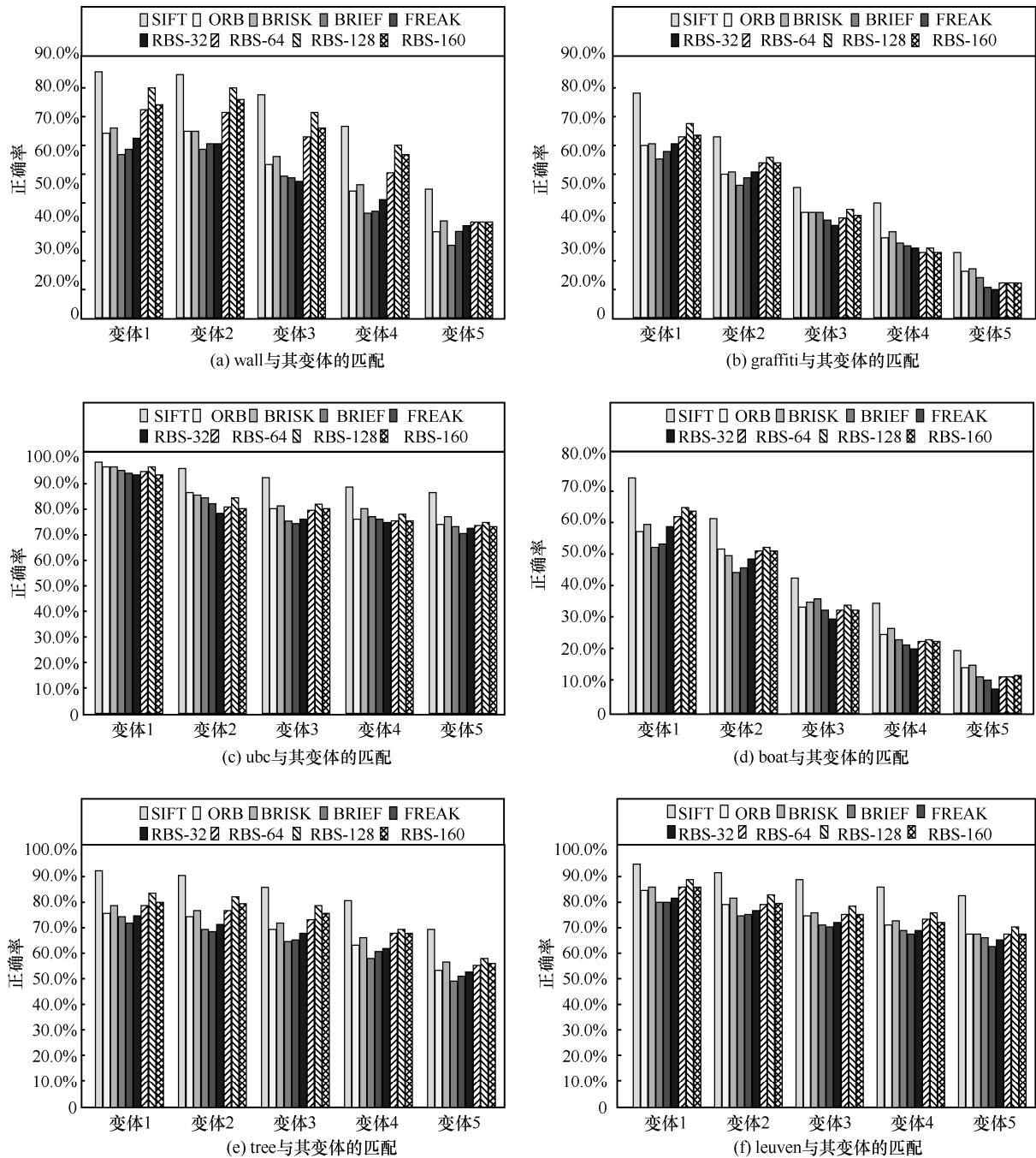


图 6 不同描述子正确率对比

表 1 不同描述子的正确率对比

图像	变换	SIFT	ORB	BRISK	BRIEF	FREAK	RBS-32	RBS-64	RBS-128	RBS-160
wall	1-2	85.3%	64.1%	65.9%	56.9%	58.4%	62.3%	72.5%	79.7%	74.7%
	1-3	84.5%	63.1%	64.9%	58.6%	58.6%	60.4%	71.3%	79.7%	72.7%
	1-4	77.4%	53.2%	55.9%	49.3%	48.5%	47.5%	63.1%	71.4%	64.9%
	1-5	66.3%	44.0%	46.7%	36.7%	37.6%	41.1%	50.6%	59.7%	53.1%
	1-6	45.1%	30.1%	34.0%	23.0%	25.5%	30.1%	32.0%	33.2%	32.4%
	平均	71.7%	50.9%	53.5%	44.9%	45.7%	48.3%	57.9%	64.7%	59.5%
graffiti	1-2	77.2%	59.5%	60.5%	54.9%	57.6%	60.3%	62.4%	66.9%	62.9%
	1-3	62.6%	49.3%	50.4%	45.6%	48.8%	50.7%	53.1%	55.2%	54.0%
	1-4	45.3%	36.5%	36.5%	36.5%	33.9%	32.0%	34.7%	37.6%	35.4%
	1-5	39.5%	27.5%	29.6%	25.6%	24.8%	24.0%	22.9%	24.0%	22.6%
	1-6	22.2%	15.7%	16.8%	13.9%	10.4%	9.3%	12.3%	12.0%	11.8%
	平均	49.4%	37.7%	38.8%	35.3%	35.1%	35.3%	37.1%	39.1%	37.3%
ubc	1-2	94.9%	93.6%	93.3%	91.9%	90.9%	90.5%	91.6%	93.2%	91.4%
	1-3	92.7%	83.8%	82.9%	81.7%	79.1%	75.9%	78.3%	81.7%	77.6%
	1-4	89.2%	77.7%	78.6%	72.8%	71.9%	73.6%	76.8%	79.1%	76.6%
	1-5	85.8%	73.6%	77.4%	74.8%	73.6%	72.5%	73.0%	75.4%	73.7%
	1-6	83.9%	71.9%	74.5%	70.7%	68.1%	69.9%	71.0%	72.2%	70.7%
	平均	89.3%	80.1%	81.3%	78.4%	76.7%	76.5%	78.2%	80.3%	78.0%
boat	1-2	74.6%	57.6%	59.7%	52.0%	53.1%	58.7%	61.6%	65.1%	63.9%
	1-3	61.4%	51.5%	49.6%	44.0%	45.6%	48.5%	50.7%	52.3%	51.1%
	1-4	42.4%	33.1%	34.9%	36.0%	32.0%	29.3%	32.3%	33.9%	32.4%
	1-5	34.5%	24.8%	26.7%	23.2%	21.3%	20.3%	22.4%	22.9%	22.6%
	1-6	19.3%	14.1%	14.7%	11.2%	10.1%	7.5%	11.5%	11.5%	11.8%
	平均	46.4%	36.2%	37.1%	33.3%	32.4%	32.9%	35.7%	37.1%	36.4%
tree	1-2	91.5%	74.9%	77.9%	73.3%	71.2%	74.1%	77.6%	82.4%	76.6%
	1-3	90.1%	73.3%	76.0%	68.5%	67.7%	70.4%	75.7%	81.3%	76.6%
	1-4	85.1%	68.8%	71.2%	64.0%	64.8%	66.9%	72.3%	77.6%	74.7%
	1-5	79.5%	62.4%	64.8%	57.6%	59.7%	61.3%	66.7%	68.8%	66.8%
	1-6	69.0%	53.1%	56.0%	48.8%	50.4%	52.3%	54.9%	57.1%	54.0%
	平均	83.0%	66.5%	69.2%	62.5%	62.8%	65.0%	69.4%	73.4%	69.8%
leuven	1-2	93.3%	82.7%	84.6%	78.7%	78.2%	79.8%	84.3%	87.0%	83.5%
	1-3	90.4%	77.7%	79.8%	73.1%	73.9%	75.0%	77.4%	81.4%	76.6%
	1-4	87.2%	73.1%	74.2%	69.7%	68.9%	70.7%	73.9%	76.9%	73.7%
	1-5	84.6%	69.9%	71.5%	67.6%	66.2%	67.8%	71.8%	74.5%	72.7%
	1-6	81.2%	66.2%	66.0%	64.6%	61.4%	64.1%	66.0%	68.6%	64.9%
	平均	87.4%	73.9%	75.2%	70.7%	69.7%	71.5%	74.7%	77.7%	74.3%
平均	—	71.2%	57.6%	59.2%	54.2%	53.7%	54.9%	58.8%	62.1%	59.2%

约 4.8%和 1.3%，RBS-32 优于 BRIEF 约 0.7%，但比 ORB 低约 2.7%。RBS 优于 512 bit 的 BRISK，与 FREAK 性能接近，RBS-32、RBS-64 和 RBS-128 性能随着长度增加匹配正确率不断上升，RBS-128 性能最高，其平均正确率超过 FREAK 约 8.4%，超过 BRISK 约 1.9%，正确率提升比例分别达到 16.4%和 5.3%。二进制描述子 BRISK 性能高于 FREAK，分析可能是由于 FREAK 并没有很好地模拟视网膜分布，而被使用均匀采样的 BRISK 超出。实验还显示，160 bit 时匹配正确率相对 128 bit 已经开始下降，这表明随着点对选择方法能够有效地筛选出稳定点对，随着排序靠后点对数量的增加，描述子开始受到噪声点的负面影响。

最后，使用 leuven 原图对不同描述子的计算速度进行对比，结果如表 2 所示。

表 2 描述子提取算法速度对比

比较内容	描述子总个数	总时间/ms	平均时间/ μ s
SIFT	2 415	3145	1302
ORB	1 516	76	50
BRISK	1 752	79	45
BRIEF	1 345	65	48
FREAK	1 428	51	36
RBS-32	1 456	15	10
RBS-64	1 456	30	21
RBS-128	1 456	60	41

由表 2 可以看出，RBS 描述子在速度上有很大的优势，比 SIFT 快了近 2 个数量级，与其他二进制描述子（ORB、BRISK、BRIEF）相比，也有一定优势，RBS-32 和 RBS-64 均快于其他二进制描述子，而 RBS-128 仅慢于对比描述子中最快的 FREAK。

4 结束语

本文模仿人眼视网膜细胞分布特性，实现了一种仿视网膜视神经分布和感知特性的采样学习方法，以生成更顽健、更紧致的二进制描述子。实验表明，描述子对光照、旋转、退化、模糊具有较高的顽健性，RBS-128 匹配正确率超过 512bitFREAK 和 BRISK 二进制描述子，计算性能比 SIFT 描述子提高了近 2 个数量级，相比其他二进制描述子也有一定的优势。但是，由于本

文的方法仍旧是像素点对之间的比较，对于大角度的视角变化顽健性较低，如图 6 中 boat 和 graffiti，需要进一步研究如何提高二进制描述子对视角变化的顽健性。

参考文献:

- [1] CALONDER M, LEPETIT V, STRECHA C, et al. Brief: binary robust independent elementary features[C]//European Conference on Computer Vision. 2010: 778-792.
- [2] RUBLEE E, RABAUDE V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]//International Conference on Computer Vision. 2011: 2564-2571.
- [3] LEUTENEGGER S, CHLI M, SIEGWART R Y. BRISK: binary robust invariant scalable keypoints[C]//International Conference on Computer Vision. 2011: 2548-2555.
- [4] ALAHI A, ORTIZ R, VANDERGHEYNST P. Freak: fast retina keypoint[C]//Computer Vision and Pattern Recognition. 2012: 510-517.
- [5] TOLA E, LEPETIT V, FUA P. Daisy: an efficient dense descriptor applied to wide-baseline stereo[J]. Pattern Analysis and Machine Intelligence, 2010, 32(5): 815-830.
- [6] 李兵, 刘磊波, 魏志强, 等. 一种具有强实时性、强顽健性的图像匹配算法[J]. 软件学报, 2014, 25(7):1583-1592.
LI B, LIU L, WEI Z Q. A strong Robust real-time image matching algorithm[J]. Journal of Software, 2014, 25(7): 1583-1592.
- [7] 朱宏宏, 李俊山, 郭莉莎, 等. 基于 LBP 的尺度不变特征的描述和匹配算法[J]. 计算机辅助设计与图形学学报, 2011, 23(10): 1758-1763.
ZHU Y H, LI J S, GUO L S, et al. A local invariant feature description and matching algorithm based on LBP [J]. Journal of Computer-Aided Design & Computer Graphics, 2011, 23(10):1758-1763.
- [8] 卢鸿波, 孙愿, 张志敏. 采用旋转匹配的二进制局部描述子[J]. 中国图象图形学报, 2013, 18(10):1315-1321.
LU H B, SUN Y, ZHANG Z M. Binary local descriptor based on rotative matching [J]. Journal of Image and Graphics, 2013, 18(10): 1315-1321.
- [9] 惠国保, 李东波, 童一飞. 挖掘图像补丁特征信息增强二进制描述子独特性[J]. 计算机辅助设计与图形学学报, 2014, 26(9): 1419-1429.
HUI G B, LI D B, TONG Y F. A discriminative binary descriptor built on further mining marginal information [J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(9): 1419-1429.
- [10] FAN B, KONG Q, TRZCINSKI T, et al. Receptive fields selection for binary feature description[J]. IEEE Transactions on Image Processing, 2014, 23(6):2583-2595.
- [11] TRZCINSKI T, CHRISTOUDIAS M, FUA P, et al. Boosting binary keypoint descriptors[C]// Computer Vision and Pattern Recognition. 2013: 2874-2881.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classifi-

cation with deep convolutional neural networks[C]//Neural Information Processing Systems, 2012:1097-1105.

- [13] XIA R, PAN Y, LAI H, et al. Supervised hashing for image retrieval via image representation learning[C]//The 28th AAAI Conference on Artificial Intelligence.2014: 2156-2162.
- [14] LIN K, YANG H F, HSIAO J H, et al. Deep learning of binary hash codes for fast image retrieval[C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015:27-35.
- [15] ZHAO F, HUANG Y, WANG L, et al. Deep semantic ranking based hashing for multi-label image retrieval[C]// Computer Vision and Pattern Recognition. 2015:1556-1564.
- [16] GARWAY-HEATH D F, CAPRIOLI J, FITZKE F W, et al. Scaling the hill of vision: the physiological relationship between light sensitivity and ganglion cell numbers[J]. Investigative Ophthalmology and Visual Science, 2000, 41(7): 1774-1782.
- [17] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors[J]. Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- [18] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[C]//European Conference on Computer Vision. 2006: 430-443.
- [19] MAIR E, HAGER G D, BURSCHKA D, et al. Adaptive and generic corner detection based on the accelerated segment test[C]//European Conference on Computer Vision.2010: 183-196.
- [20] LOWE D G. Object recognition from local scale-invariant features[C]//The Seventh IEEE International Conference on IEEE. 1999: 1150-1157.

[作者简介]



袁庆升(1980-)，男，山东济南人，中国科学院信息工程研究所博士生，国家计算机网络应急技术处理协调中心副高级工程师，主要研究方向为多媒体大数据处理、网络与信息安全。



靳国庆(1988-)，男，山东单县人，博士，中国科学院计算技术研究所助理研究员，主要研究方向为多媒体内容检索、模式识别等。



张冬明(1977-)，男，江苏盐城人，博士，国家计算机网络应急技术处理协调中心研究员、硕士生导师，主要研究方向为多媒体内容检索、模式识别、视频编码等。

包秀国(1963-)，男，江苏如皋人，博士，国家计算机网络应急技术处理协调中心教授级高级工程师、博士生导师，主要研究方向为网络与信息安全。