

DOI: 10.13382/j.jemi.2016.02.001

局部二进制特征描述算法综述^{*}

白 丰¹ 张明路¹ 张小俊^{1,2} 孙凌宇¹

(1. 河北工业大学机械工程学院 天津 300130;

2. 哈尔滨工业大学机器人技术与系统国家重点实验室 哈尔滨 150080)

摘 要: 为引起更多机器人视觉领域的研究者对局部二进制特征描述算法进行探索和讨论,并推动局部二进制特征描述算法在图像匹配、目标识别和机器人导航等领域的应用,系统的综述局部二进制特征描述算法的发展过程以及研究进展。首先简要概述浮点型的特征描述算法,并指出研究局部二进制描述算法的必要性;然后详细的介绍基于手动设置的经典二进制描述算法及其改进方案,深入分析基于优化学习的二进制特征描述算法的优势和缺陷;最后对局部二进制特征描述算法存在的问题进行讨论,强调将当前深度学习理论与局部二进制特征描述原理相结合的重要性,同时随着深度学习理论的引入,局部二进制特征描述算法必然会更加适合机器人系统开展反恐防爆、核电安全、太空探索等特殊环境下的图像匹配和目标识别任务。

关键词: 局部特征; 二进制描述; 深度学习; 目标识别

中图分类号: TP301.6 **文献标识码:** A **国家标准学科分类代码:** 510.8040

Summarization of local binary feature description algorithm

Bai Feng¹ Zhang Minglu¹ Zhang Xiaojun^{1,2} Sun Lingyu¹

(1. School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China;

2. State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150080, China)

Abstract: To cause more researchers in the field of robot vision to explore and discuss the local binary character description algorithm and promote the application research in image matching, object recognition and robot navigation, the development process and research progress of local binary feature description algorithm is summarized systematically. First, the float point type description algorithm is briefly summarized, and the necessity of research on the local binary feature description algorithm is pointed out. Then, the binary feature descriptor algorithm based on manually and relevant improved scheme is introduced in detail, and the advantages and defects of the binary feature descriptor algorithm based on optimizing study are analyzed in-depth. Finally, the existing problem of local binary feature description algorithm is discussed, and the essentiality of combining the current mainstream deep learning theory with local binary feature description theory is emphasized. At the same time, along with the introduction of deep learning theory, the local binary feature description algorithm is certainly to be more suitable for robot system developing image matching and target recognition tasks in the special environment such as counter-terrorism explosion-proof, nuclear safety and space exploration.

Keywords: local feature; binary description; deep learning; target recognition

收稿日期: 2015-10 Received Date: 2015-10

^{*} 基金项目: 国家 863 计划(2015AA043101)、机器人技术与系统国家重点实验室开放研究(SKLR2013ZD04)资助项目

1 引言

机器人系统提取的视频画面中普遍存在各种视角、尺度、旋转甚至光照变换的干扰,严重影响图像匹配、目标识别和机器人导航等任务的顺利进行。有效的表示视频或者图像的局部邻域显著特征并对各种变换保持不变性具有挑战意义。相关调查显示,局部特征描述算法对于以上变换干扰以及相似背景信息的影响具有良好鲁棒性,成为当前的研究重点和必然趋势。

在局部特征描述算法的研究过程中,lowe等首先提出浮点型的尺度不变特征变换算法 SIFT^[1],该算法利用高斯差分模板构建图像组层,并在DOG空间中提取极值点;然后剔除边缘和低对比度点的伪点,通过梯度方向直方图构建128维的描述向量。虽然SIFT算法具有优越的区分性能,但是计算复杂度很高。学者们随后提出相应的浮点型改进方案,比如Ke等提出的PCA-SIFT算法^[2]在 41×41 的图像块上计算 $39 \times 39 \times 2$ 个梯度导数,将获得的3042维向量通过主成分分析的方式降低至36维,算法能够有效压缩特征描述向量的维数,减少匹配时间;但是构建描述向量所需的时间增加,总体的执行时间只略低于SIFT算法;Bay等提出的快速鲁棒特征SURF^[3]采用框状滤波器快速创建图像层,并利用积分图像和海森矩阵检测特征点,通过Haar小波构建浮点型描述向量,实时性能大约是SIFT算法的2~3倍;刘立等提出的简化算法SSIFT^[4]在创建的圆形框架中构建12维特征描述向量,并利用排序的方法合并两次梯度运算,总体的执行时间仅为SIFT算法的1/5。相比SIFT算法,以上浮点型的描述方案的实时性能有所提升。但是研究表明,这些改进方案的区分性能较SIFT算法均有明显下降。

另外,SIFT算法还包括一些改进方案,比如Mikolajczyk等提出的GLOH算法^[5],Abdel Hakim等提出的CSIFT算法^[6],Morel等提出的ASIFT算法^[7],蔡国榕等提出的PSIFT算法^[8]和唐永鹤等提出的基于拉普拉斯的局部特征描述算法^[9]等,虽然这些算法具有优越的区分性能,但是,计算复杂度甚至要高于SIFT算法。随着视频图像本身包含的信息以指数级的速度增加,浮点型的描述方案更加不适合当前高速高精度的工程应用需

求。因此,许多学者提出“利用二进制描述向量代替浮点型描述向量,通过汉明距离代替欧氏距离快速完成特征点匹配和目标识别”的手动设置描述算法。

本文首先在传统浮点型特征描述算法的基础上,指出研究局部二进制描述算法的必要性;随后,详细的介绍基于手动设置的经典二进制描述算法及其改进方案,深入分析基于优化学习的二进制特征描述算法的优势和缺陷;最后对局部二进制特征描述算法存在的问题进行讨论,强调将当前主流的深度学习理论与局部二进制特征描述算法相结合的重要性。同时,随着深度学习理论的引入,局部二进制特征描述算法也必然会更加适合机器人系统开展反恐防爆、核电安全、太空探索等特殊环境下的图像匹配和目标识别任务。

2 基于手动设置的二进制描述算法

在基于手动设置的二进制描述算法的研究中,学者们首先想到通过二值划分的方式将浮点型描述向量转化为二进制型^[10-11]。但是,这类算法需要先获得浮点型描述向量,再进行二值转化,因此计算复杂度有增无减。学者们进一步提出直接在原始图像的特征点邻域中构建二进制描述向量以获得优越的实时性能;其主要思想是:在以特征点为中心的局部区域里,根据特征点邻域的采样模式(或随机模式)进行采样点间的强度值比较,将所有的比较结果依次串联构成二进制描述向量。

在基于手动设置的描述方案中,最具代表性的算法当属Calonder等提出的二进制鲁棒独立基础特征BRIEF^[12],该算法通过比较随机选取的高斯分布点对构建二进制描述向量,相比浮点型的描述算法具有明显的实时性能优势,具体的随机采样模型如图1所示;

Rublee等提出的ORB算法^[13]利用加速分割检测特征FAST^[14]提取稳定的特征点,然后采用具有方向特性的改进BRIEF描述子获取二进制描述向量并结合汉明距离实现特征点的快速匹配。实验表明ORB算法的平均计算时间仅是SIFT算法的1/30,同时,由于能够在具有较高方差和期望值的方向范围内提取特征点对,并筛选掉候选特征点

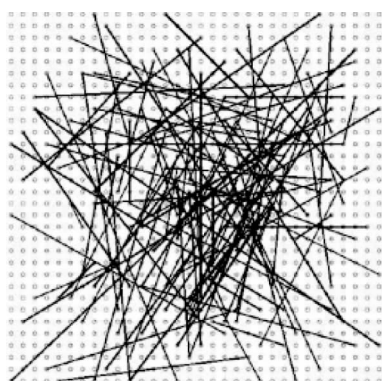


图1 随机采样点对分布图

Fig. 1 Distribution of random sampling points

对中相关性能偏高的集合,相比 BRIEF 描述算法具有更加优越的区分性能和鲁棒性能优势。该算法的具体采样点对如图 2~3 所示;

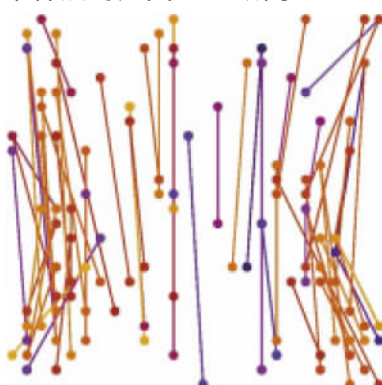


图2 按方向提取特征点对

Fig. 2 Extracting the pair of feature point according to the direction

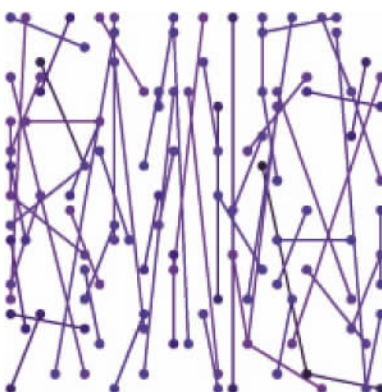


图3 按相关程度提取特征点对

Fig. 3 Extracting the pair of feature point according to the relevancy

Leutenegger 等进一步提出的二进制鲁棒尺度不变特征算法 BRISK^[15-16] 利用 AGAST 快速提取稳定的特征点,然后在一系列同心圆上等间距选取采样点并求取采样点间的距离集合,利用短距离的集合快速构造描述子,利用长距离的集合计算描述向量的主方向,该算法同样利用汉明距离实现特征点间的快速匹配,计算的复杂度与 ORB 描述算法相当。具体的 AGAST 特征点检测示意图如图 4 所示,获得的特征点均具有亚像素级的精度,图 5 表示 BRISK 二进制描述向量的采样模型。

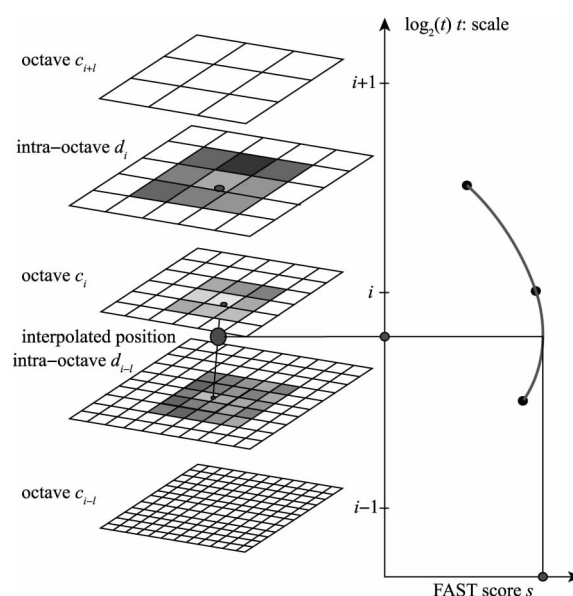


图4 AGAST 特征点检测算法

Fig. 4 AGAST feature point detection algorithm

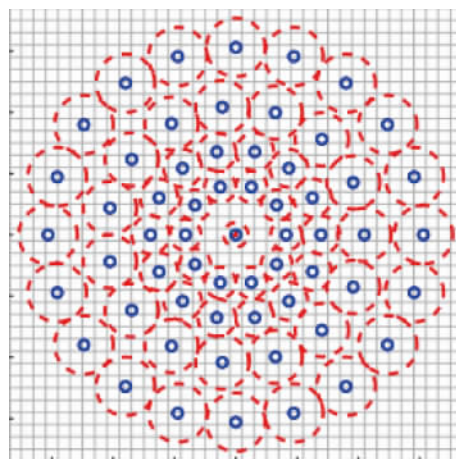


图5 BRISK 描述向量的采样模型

Fig. 5 Sampling model of BRISK description vector

等提出的快速视网膜关键点描述子 FREAK^[17] 来源于人类视网膜的分布特性^[18], 该算法以特征点为中心构建七层同心圆环(层数不固定), 所有圆环上均包含等间距分布的6个采样点, 采样点的接受域间存在着重合部分, 通过高斯平滑滤波除噪声后, 将采样点对的强度比较结果依次链接形成二进制串, 最后利用最大方差和最小标准差的迭代运算筛选出512维的二进制描述向量。这里的512维二进制描述向量可分为4个区域, 从第1区到第4区的采样点对位置信息代表着中心特征点由粗到精的描述方式。相比 BRISK 描述算法, FREAK 描述子的执行时间略有优势, 具体的 FREAK 采样模型如图6所示。

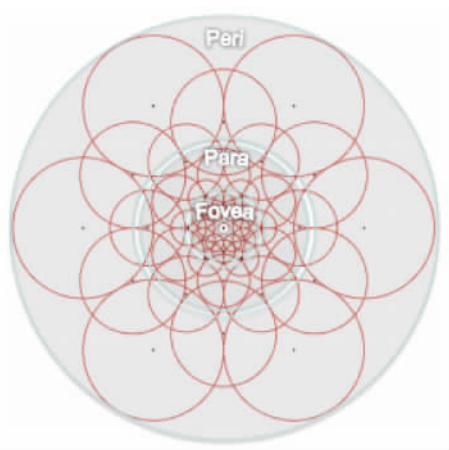


图6 FREAK 描述子的采样模型

Fig.6 Sampling model of FREAK descriptor

另外, 具有代表性的传统二进制描述算法还包括 ZHANG 等提出的 USB 算法^[19], LI 等提出的 BRINT 算法^[20], Kota 等提出的 BRIGHT 算法^[21] 和 Sajib 等提出的 ALOHA 算法^[22] 等。虽然以上二进制描述算法具有实时性能的优势, 但是由于区分性能和鲁棒性能明显低于浮点型的描述方案, 很难应用到实际的工程领域。

3 手动设置二进制描述算法的改进

传统二进制特征描述算法虽然具有优越的实时性能, 但是区分性能普遍低于浮点型的描述方案, 导致特征点间错误匹配、目标无效识别等现象严重。相关学者针对性的提出有效改进方案, 具有代表性的算法当属 Xianwei 等提出的

OSRI^[23], 该算法将特征点的局部邻域扩展到多层范围, 并对每层邻域按像素强度值和梯度方向顺序划分为多层子区域; 按圆形区域中的像素强度值或者梯度方向值将划分范围进行多次旋转(移位), 构建多组层的子区域集合; 最后通过子区域中多种不变特征的比较结果形成二进制描述向量。OSRI 算法在部分图像匹配和目标识别应用中能够接近甚至达到部分浮点型描述算法的区分性能, 并保有传统二进制描述算法良好的实时性能。具体按像素强度值划分图像并旋转划分区间得到的子区域集合如图7所示, 图8是 OSRI 算法的总体流程图。

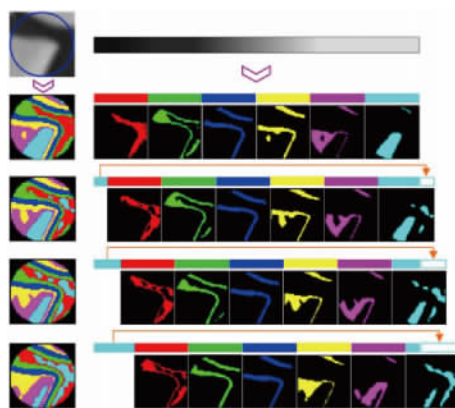


图7 按像素强度值划分图像

Fig.7 Dividing the image according to the pixel intensity value

与此同时, 时等提出的基于强区分性描述子的视觉识别方法^[24] 在快速视网膜关键点描述子 FREAK 的基础上, 通过深入分析人类视网膜各区域的功能结构, 建立特征点局部邻域的多组层子区域集合, 然后在新采样模型中增加周边采样点对比较结果的份额。该算法能够提升传统二进制特征描述算法的区分性能, 并兼有优良的实时性能, 具体的新采样模型如图9所示。

高洪波等提出的基于分层处理的特征点匹配算法^[25] 通过模拟 FREAK 算法和 BRIEF 算法的采样点对选取机制, 有效结合固定点抽样和随机抽样两种模式的优势, 设计出合理的采样点对辨识模型。在保证与传统二进制特征描述算法具有相当实时性能的前提下, 成功提升特征点间的正确匹配率。具体的粗精两种抽样模型如图10~11所示。

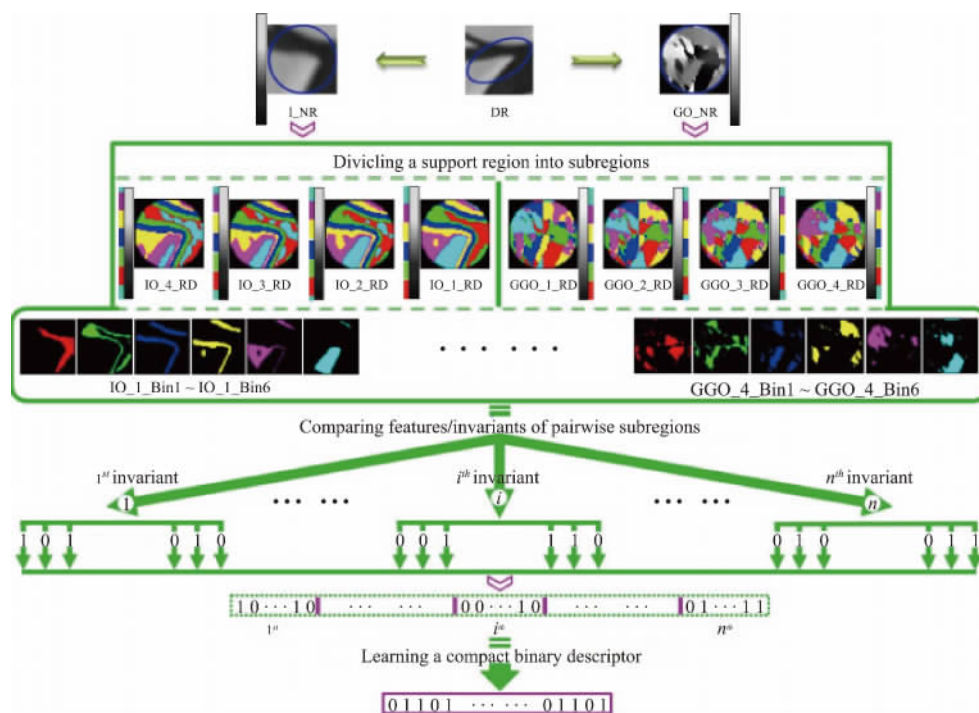


图 8 OSRI 算法流程图

Fig. 8 Flow chart of OSRI algorithm

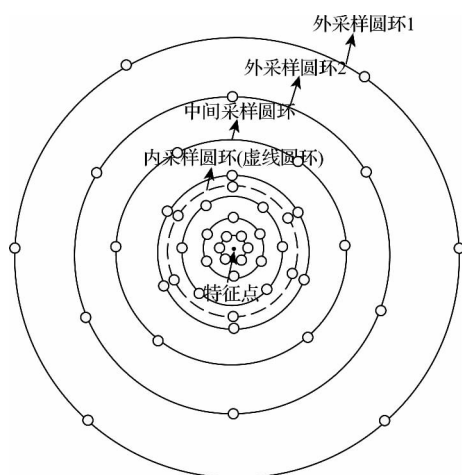


图 9 改进 FREAK 算法的新采样模型

Fig. 9 New sampling model of improved FREAK algorithm

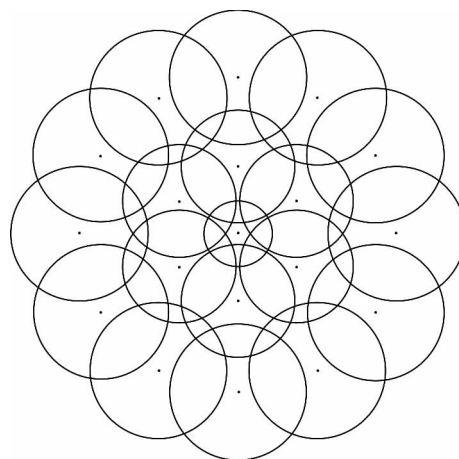


图 10 基于分层处理的粗采样模型

Fig. 10 Coarse sampling model based on hierarchical processing

卢鸿波等提出的 RMBD 算法^[26]在以特征点为中心的同心圆周上等间隔设置采样点,按照旋转不变模式选取采样点对进行灰度值比较,将比较结果依次连接成二进制串作为描述向量。匹配时比较描述向量在所有旋转变换模式中的汉明距离,取最小距离值作为结果,实现不依赖于主方向的旋转不变

性。另外,描述方案配合多尺度的特征检测方法,将采样模式进行缩放,并对图像进行亚像素插值,实现尺度不变性。该方案在一定程度上克服了各种变换,特别是旋转变换的干扰,相比传统二进制描述算法能够匹配到更多的正确特征点,具有更高的鲁棒性能。RMBD 算法具体的旋转不变采样模型如图 12 所示。

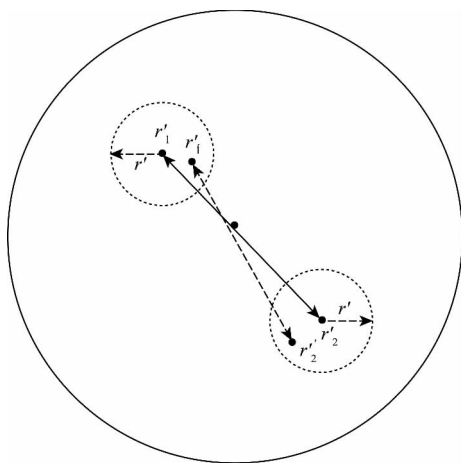


图 11 基于分层处理的细采样模型

Fig. 11 Fine sampling model based on hierarchical processing

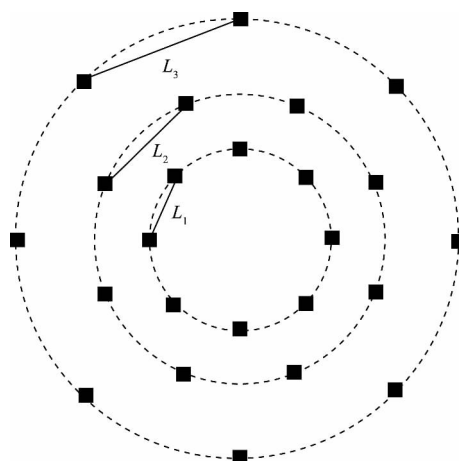


图 13 采样点密度示意图

Fig. 13 Schematic diagram of sampling point density

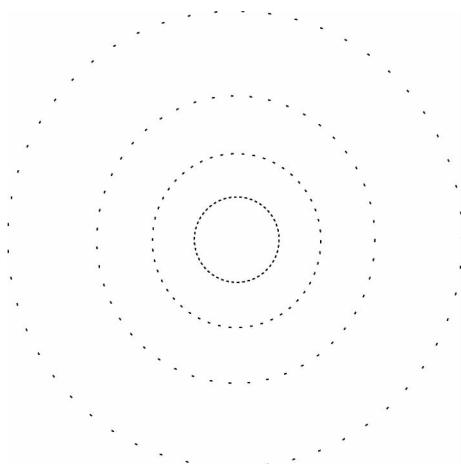


图 12 RMBD 算法的采样模型

Fig. 12 Sampling model of RMBD algorithm

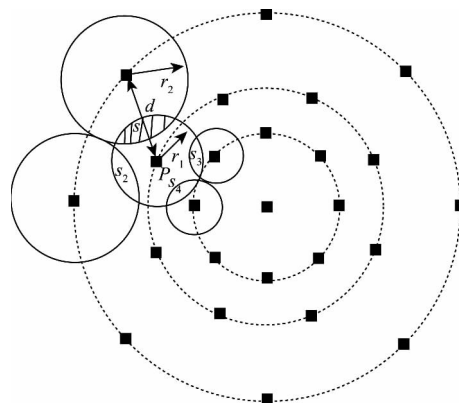


图 14 平滑范围重叠度示意图

Fig. 14 Schematic diagram of smooth range overlap degree

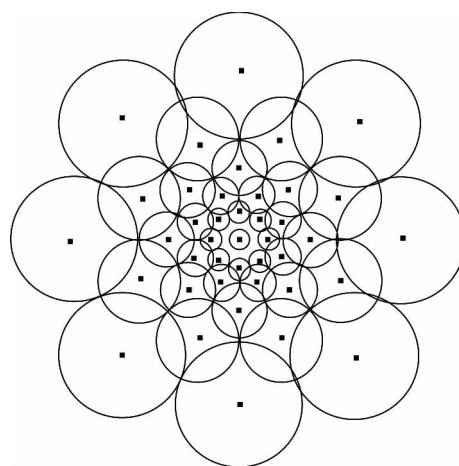


图 15 CFRBD 算法的采样模型

Fig. 15 The sampling model of CFRBD algorithm

惠国保等提出的基于优化采样模式的紧凑二进制描述算法 CFRBD^[27]通过对最新提出的采样模型特点进行分析和测试,发现采样点密度和平滑重叠度对描述向量的独特性具有重要影响,据此设计出一种优化的采样模型;其次利用像素点灰度值排序方法分解图像局部领域,产生多个对应不同灰度段的子图像集合;最后将优化的采样模式映射到子图像中,随机提取样本点进行灰度值的比较。相关实验表明,CFRBD 二进制描述算法能够应用于实时性能要求高、内存紧凑的目标识别任务。具体的采样点密度和平滑范围重叠度示意图如 13 ~ 14 所示,图 15 是 CFRBD 算法最终的采样模型。

另外,改进的二进制特征描述方案还包括庄东晔等提出的基于随机亮度差量化的二进制描述算法^[28]以及惠国保等提出的基于像素聚类的二进制描述算法^[29]等。以上算法均能够提升传统二进制描述方案的区分性能,并保有二进制描述算法原本优越的实时性能优势。但是有学者同时指出,经典的二进制描述算法及其改进方案均基于手动设置,采样模型的构建过程通常需要依靠丰富的经验和复杂的验证,限制着二进制描述算法的推广应用。

4 基于优化学习的二进制描述算法

为克服基于手动设置的经典二进制描述算法及其改进方案存在的缺陷,近些年基于机器视觉的优化学习算法(数据驱动算法)被越来越多的学者应用于构建二进制特征描述向量,比如相似敏感哈希算法^[30]、位置敏感哈希算法^[31]等。

在最初的研究中,无监督式的优化学习算法首先被应用于构建二进制描述向量。比如主成分分析和非负矩阵分解法^[32]。其中,影响最为广泛的方案当属 Salakhutdinov 等提出的语义哈希算法^[33],该算法首先将高维的数据信息转换为二进制的形式,并根据相似的数据必然具有相似的二进制编码的原理,计算描述向量间的最近邻匹配结果,高质量的二进制描述向量可以通过熵的最大化准则生成。具体的语义哈希算法示意图如 16 所示。

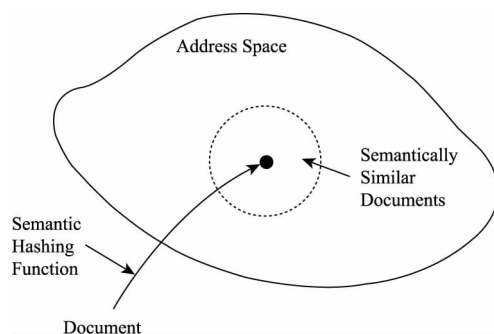


图 16 语义哈希算法的示意

Fig. 16 Schematic diagram of the semantic Hash algorithm

同时,Weiss 等提出的谱哈希算法 SH^[34]通过主成分分析的方法对高维数据进行降维,然后直接编码降维的结果,获取二进制描述向量。SH 算法

具有很高的效率,并且经过降维和编码处理后的二进制描述向量仍然能够保持原始高维数据间的相似程度。但是,SH 算法的所有高维特征均会参与编码过程,因此该算法面临如何处理冗余高维信息的问题。

Kulis 等提出的 Binary Reconstruction Embeddings(BRE) 算法^[35]通过学习哈希函数,显式最小化原始输入的欧氏距离和映射的汉明距离间的误差,实现浮点型描述向量匹配到二进制描述向量匹配的过度。通过多尺度坐标下降法(scalable coordinate-descent algorithm) 从不同位置优化哈希函数。BRE 算法能够达到较高的准确率,但是优化学习的过程较为复杂,需要较长的训练时间,限制着该算法在大规模数据集中的应用。

早期的无监督式学习算法大多以保留原始局部特征为目标,然而在实际应用中往往更倾向优化相关的结果。基于此,标签信息被应用到优化算法中,产生很多半监督式的学习方案。

比如 Wang 等^[36]提出利用特征对进行优化学习,使相似的特征对在二进制编码过程中保持最大程度的相似性;反之,不相似的特征对编码出最大程度的不相关性。同时保证二进制描述向量在未标记的数据上能够得到最大的熵。另外, Mu 等^[37]提出将原始数据中的一些特征对进行手动标注,分成语义相似对和语义不相似对两类;通过放松限制条件,将整体的优化问题转变成若干子问题,利用二次规划进行求解,得到用于构建描述向量的哈希函数。这些半监督式的优化学习算法与早期无监督式的优化学习算法相比,能够在很大程度上提高二进制描述向量的区分性能。

虽然无监督式和半监督式的优化学习算法具有良好的区分性能,但是算法中预先设置的相似性度量方式在很多情况下受到原始输入图像的精确程度影响,导致无监督式和半监督式的优化学习算法难以达到最优。因此,在无监督式和半监督式的优化学习算法之后,相关学者进一步提出充分利用数据标签信息的监督式优化学习算法。

监督式优化学习算法的核心思想是:通过标记过的训练样本学习有效的相似性度量方式,以获得最初期望的近似关系。将监督式的优化学习算法应用于构建二进制描述向量的方案有许多,最具代表性的算法当属 Trzcinski 等提出的 D-BRIEF^[38],该算

法在训练集中通过优化学习获得一系列相互正交的投影向量和最优阈值,然后将图像的局部领域信息投影到更有区分性能的特征空间,阈值化所有的投影结果依次建立二进制描述向量。投影均由一系列非线性滤波器的响应值通过线性组合构成,响应值可利用框状滤波器和高斯滤波器的卷积运算或者矩形滤波器的积分运算快速计算得到。具体优化后的投影结果如图17所示,其中 W^0 表示最终的投影结果, $|S|_0$ 表示投影向量的非零元素数量。

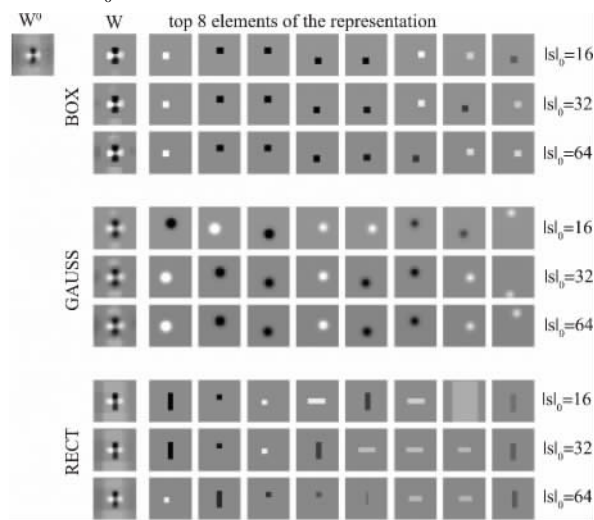


图17 优化后的投影向量

Fig. 17 The projection vector after optimizing

Liu等提出的 Kernel-Based Supervised Hashing (KSH) 算法^[39]利用目标函数操作编码内积以代替汉明距离实现特征匹配;根据编码内积的离散属性通过穷举优化算法(greedy optimization algorithm)使哈希函数能够按顺序依次训练;另外KSH的优势还包括能够通过基于核函数的哈希算法处理线性不可分离的数据以及最小化相似图像对的二进制描述向量距离值的同时最大化相异图像对的二进制描述向量匹配结果。具体利用编码内积代替汉明距离进行优化学习的示意图如图18所示。

Strecha等提出的 LDAHash 方案^[40]针对基于手动设置的二进制描述算法和浮点型的描述算法只能近似模拟图像邻域各种变换的缺陷,提出通过学习映射关系以最小化描述向量的类内方差和最大化描述向量的类间方差,并利用一维搜索的方式获得最佳阈值,构建具有完全仿射不变性能的二进制描述向量。LDAHash 算法相比基于手动设置的

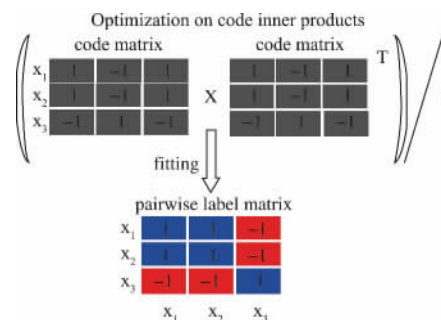


图18 编码内积的优化结果

Fig. 18 Optimal result of code inner product

二进制描述算法和浮点型的描述算法具有更加优良的区分性能,具体的映射方法包括线性判别分析和协方差的差分方式(Difference of Covariances)两种,阈值的选取通过最小化假正率和假负率获得。

另外,监督式的优化学习方案还包括毛晓蛟等提出的基于子空间学习的图像语义哈希算法^[41]、欧阳端飞等提出的结构化稀疏谱哈希算法^[42]等。以上利用监督学习方法构建的二进制描述向量通常是在原始图像的灰度空间或者特征空间中进行线性变换,然后利用预先设置的核函数获取非线性的特征,具有比无监督式和半监督式的优化学习算法更加优良的区分性能。但是,现有的监督式学习方法仍然存在明显的缺陷,比如通过学习哈希函数获取紧凑二进制描述向量的过程中,投影矩阵往往需要较强的判别性,能够同时区分多类特征,否则很难保证匹配和识别的准确率。但是,理论上很难学习到具有如此强判别性能的投影矩阵,若增加二进制描述向量的长度,则需要更多的运算时间和存储空间。

5 优化学习算法的最新研究进展

针对早期无监督式、半监督式和监督式的优化学习算法在构建二进制描述向量的过程中存在的缺陷,最新的研究指出,当前阶段要求基于优化学习的二进制描述算法能够学习弱分类器的最优权重并优化弱分类器的联合配置方式,以保证描述向量具有优越的区分性能和鲁棒性能。具体的算法原理概述如下:

Brown等^[43]分别模拟人类视觉系统的单细胞和复合细胞处理过程,对输入图像进行高斯平滑并利用梯度的角度量化和方向可调滤波器进行非线性变换(G-blocks和T-blocks模块),空

间联合响应值以最小化分类误差或通过线性判别分析学习最优参数 (S-blocks 或 E-blocks 模块), 最后对描述向量做归一化处理以消除外部光照变化以及自身尺寸等因素的影响 (N-blocks 模块)。所提算法虽然比早期的优化学习方案具有明显的区分性能优势, 但是由于没有利用 ROC 曲线进行详细的算法性能分析和参数评估, 获取的二进制描述向量很难达到最优。具体的模块组合形式如图 19 所示, 图 20 表示空间联合的两种典型模板。

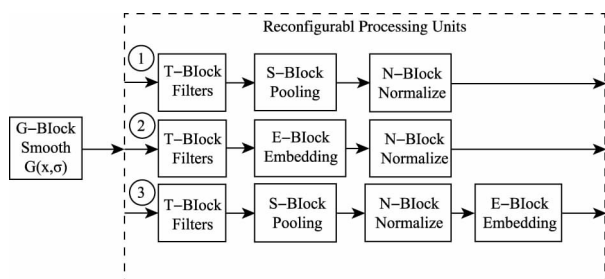


图 19 模块的组合方式

Fig. 19 The assemble way of modules

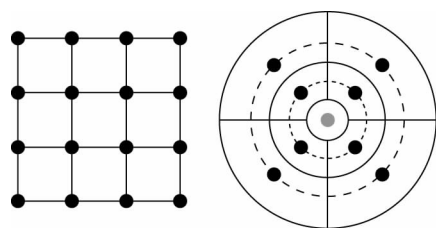


图 20 空间联合的模板

Fig. 20 The template of space combination

Simonyan 等^[44]基于凸优化理论学习稀疏接受域的有效联合方式, 并通过标准核的正交化实现降维; 最后利用随机优化算法处理损耗函数, 获取的二进制描述向量具有优良的区分性能。但是, 为使优化学习的过程可控, 算法在联合配置过程中加入了约束条件。包括限制采样区域为圆形或者方形, 必须关于邻域的中心或者关于主次对角线方向对称; 另外在具体的计算过程中, 关于中心对称的各接受域会被同时选择或者舍弃, 这实际上都会对优化学习的精确度产生影响。联合的配置方式如图 21 所示。

从理论上分析, 采用常规的优化学习算法实现形状和权重的联合配置非常困难, 因为潜在

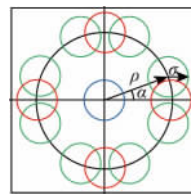


图 21 基于凸优化理论的联合配置模板

Fig. 21 Joint configuration template based on convex optimization theory

的位置和可能的联合配置方式有很多, 筛选的过程复杂繁琐。对于这类问题的解决, 最合理的方案当属 boosting 的相关算法。虽然采用穷举搜索的方式, 但是 boosting 算法能够将输入信息进行非线性的变换, 并自动参数化描述向量每维的形状和权重, 构建出具有高精度的紧凑二进制描述向量。

最先利用 boosting 学习图像相似度并构建二进制描述向量的算法是 boosted similarity sensitive coding (boosted SSC)^[45], 但是 boosted SSC 算法只能进行线性的投影, 并且获取的二进制描述向量维度偏高, 不适合实际的应用要求。另外, Low-dimensional Boosted Gradient Map (L-BGM) 算法^[46]利用非线性的变换将原始图像映射到高维的特征空间, 度量方式通过模拟弱分类器间的相似程度获得。但是由于 boosting-trick 采用显式映射, 同时学习到的嵌入方法被假设成为离散形式, 构建的二进制描述向量难以达到最优。

在所有利用 boosting 构建的描述向量中, 最具代表性的算法当属 Adaboost^[47]。在成熟的深度学习理论提出之前, Adaboost 与支持向量机 (SVM) 被认为是两种效果最好的分类方法, Adaboost 算法的核心思想是训练不同的弱分类器, 并把这些弱分类器组合, 构造一个最终的强分类器。算法本身是通过改变数据分布实现的, 根据每次训练集中的每个样本的分类是否正确, 以及上次总体分类的准确率来确定每个样本的权值。将修改后的权值发送给下层的分类器进行训练, 然后融合每次训练得到的分类器, 作为最终的决策分类器。

具有影响力的 boosting 方案还包括 Trzcinski 等提出的 BinBoost 算法^[48-49], 该算法利用新颖的优化学习模式获取高区分性能的二进制描述向量, 并且保持尺度、视角、旋转以及光照变换的不变性。

按照顺序学习的方法训练描述向量的最优形状和权重,并通过弱分类器的线性组合计算描述向量的所有维度。相比其他优化学习算法最大的优势在于能够充分利用 boosting 算法进行非线性特征变换的同时,有效的参数化描述子的形状和权重;另外在每个哈希函数中不会对弱分类器的联合配置方式强加任何限制,直接作用于原始图像并不需要任何的中间表示。具体 BinBoost 算法的二进制描述向量构建过程如图 22 所示,整个过程是通过迭代优化的方式实现的,图 23 表示基于强度和梯度信息的两组弱分类器。

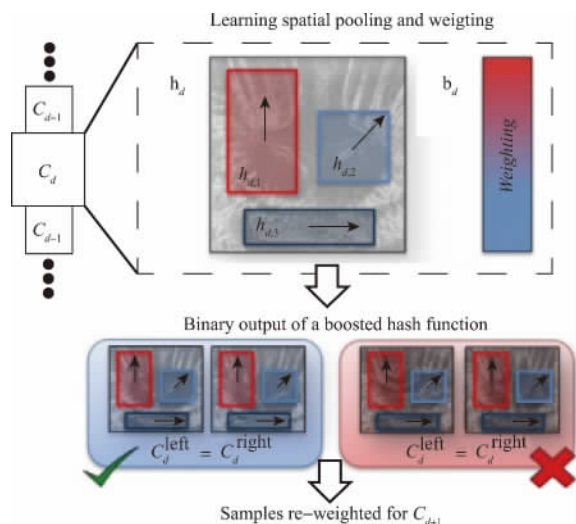


图 22 BinBoost 描述向量的构建过程

Fig. 22 The construction process of BinBoost description vector

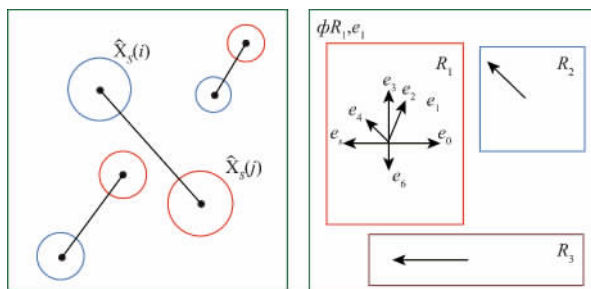


图 23 基于强度和梯度信息的弱分类器

Fig. 23 Weak classifier based on intensity and gradient information

值得注意的是, Bin 等受到 L-BGM 算法的启发, 提出 RFD_C 和 RFD_R 特征描述算法^[50]。该算法通过二值化接受域的响应值构建描述向量。接受

域从候选集中提取, 利用真正率和真负率选取其中区分性较强的接受域和对应的最优阈值, 然后通过相关系数公式筛选其中相关性能较低的接受域用以构建二进制描述向量。但是, 由于训练图像的数量以及候选接受域的数目仍然非常多, 造成实时性能并无明显优势; 同时 RFD 算法的区分性能与其他基于优化学习的二进制描述算法相比, 仍然还有进一步提升的空间, 具体的算法流程图如 24 所示。

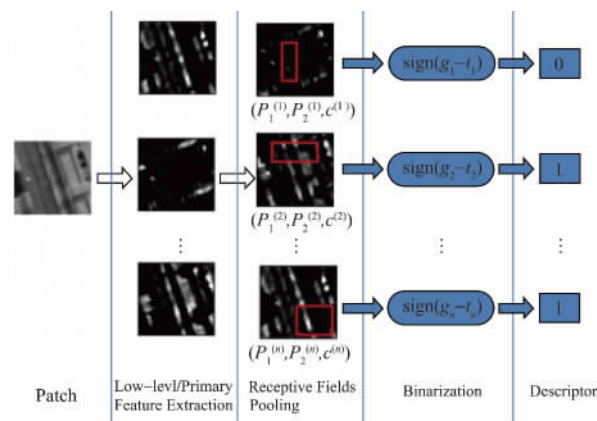


图 24 RFD 算法流程

Fig. 24 Flow chart of RFD algorithm

另外, 成功利用 boosting 构建局部二进制特征描述向量的算法还包括惠国保等提出的基于图像局部区域相似度学习架构的匹配算法^[51]; 耿利川等提出的基于 AdaBoost 分类方法的鲁棒图像局部二值特征描述算法^[52]等。相关研究表明, 基于 boosting 的优化学习算法构建的二进制描述向量在当前具有相对最优的区分性能。

6 局部二进制特征描述算法存在的问题和发展的趋势

无论是基于手动设置的二进制描述算法还是基于优化学习的二进制描述算法, 逐渐被学者们认可的关键在于具有实时性能优势的同时, 能够保持相对鲁棒的区分性能。但是相关的理论研究表明, 当前的局部二进制特征描述算法研究同样存在明显的缺陷, 具体概括如下:

在理论研究过程中, 部分学者在构建基于手动设置的二进制描述向量的过程中, 以牺牲算法整体的区分性能为代价, 盲目的追求速度方面的优势。与此同时, 在研究基于优化学习的二进制描述向量

的过程中,又以无限制牺牲算法的实时性能为代价,通过繁琐的迭代优化步骤构建二进制描述向量。由于学者们只追求实时性能或者区分性能的提升,从仿真实验里发现,这些局部二进制特征描述算法普遍不具有普遍适用性。

与此同时,具体的工程应用领域相比理论研究会有更高的精度和速度约束,例如在机器人系统的目标识别任务中,通常会明确要求所提识别算法必须能够克服各种尺度、视角和旋转变换干扰以及光照因素的影响,对于狭小空间环境下的各种遮挡现象具有良好的鲁棒性,并且保证快速有效的框选有待识别目标。显然,只具有实时性能或者区分性能的局部二进制特征描述算法根本无法满足当前高速高精度的工程应用要求。

针对局部二进制特征描述算法存在的问题,未来的发展方向必然倾向于提出“兼顾基于手动设置的二进制特征描述子优越的实时性能和基于优化学习的二进制特征描述子鲁棒的区分性能”的方案,并且能够得到理论研究的认证以及适用于具体的工程应用场合。近些年,深度学习以其优异的区分性能和鲁棒性能成为学者们的研究热点,根据学习单元的不同,深度学习理论主要包括基于限制玻尔兹曼机的深度置信度网络^[53]和基于自编码器的深度网络^[54]两类,自从Hinton等人提出训练限制玻尔兹曼机到目前为止,产生了很多优秀的深度学习算法,比如去噪自编码器^[55]、三元因子玻尔兹曼机^[56]、卷积限制玻尔兹曼机^[57]和神经自回归分布估计器^[58]等。将深度学习理论的区分性能优势与二进制描述向量自身的实时性能优势相结合,必然会成为未来的研究重点和发展趋势。

7 结 论

为引起更多机器人视觉领域的研究者对局部二进制特征描述算法进行探索和讨论,并推动局部二进制特征描述算法在图像匹配、目标识别和机器人导航等应用领域的研究,系统的概述了局部二进制特征描述算法的发展历史、演变过程以及研究进展。并对局部二进制特征描述算法在理论研究和工程应用中存在的局限性进行讨论,强调将当前主流的深度学习理论与局部二进制特征描述算法相结合的重要性。可以预见,随着深度学习理论的引

入,局部二进制特征描述算法必然会更加适合机器人系统开展反恐防爆、核电安全、太空探索等特殊环境下的图像匹配和目标识别任务。

参考文献

- [1] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [2] KE Y, SUKTHANKAR R. PCA-SIFT: a more distinctive representation for local image descriptors [C]. *Proceedings of the International Conference on Computer Vision and Pattern Recognition*. Washington DC, USA: IEEE, 2004: 506-513.
- [3] BAY H, ESS A, TUYTELAARS T, et al. SURF: speeded up robust features [J]. *Computer Vision and Image Understanding*, 2008, 110(3): 346-359.
- [4] 刘立,彭复员,赵坤,等. 采用简化 SIFT 算法实现快速图像匹配 [J]. *红外与激光工程*, 2008, 37(1): 181-184.
LIU L, PENG F Y, ZHAO K, et al. Simplified SIFT algorithm for fast image matching [J]. *Infrared and Laser Engineering*, 2008, 37(1): 181-184.
- [5] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors [C]. *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Madison, USA: IEEE, 2003: 17-22.
- [6] ABDEL-HAKIM A E, FARAG A A. CSIFT: a SIFT descriptor with color invariant characteristics [C]. *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Washington DC, USA: IEEE, 2006: 1978-1983.
- [7] MOREL J M, YU G. ASIFT: a new framework for fully affine invariant image comparison [J]. *SIAM Journal on Imaging Sciences*, 2009, 2(2): 438-469.
- [8] 蔡国榕,李绍滋,吴云东,等. 一种透视不变的图像匹配算法 [J]. *自动化学报*, 2013, 39(7): 1053-1061.
CAI G R, LI S H Z, WU Y D, et al. A perspective invariant image matching algorithm [J]. *Acta Automatica Sinica*, 2013, 39(7): 1053-1061.
- [9] 唐永鹤,卢焕章,胡谋法. 基于 Laplacian 的局部特征描述算法 [J]. *光学精密工程*, 2011, 19(12): 2999-3006.
TANG Y H, LU H Z, HU M F. Local feature description algorithm based on Laplacian [J]. *Optics and Precision Engineering*, 2011, 19(12): 2999-3006.
- [10] CHANDRASEKHAR V, TAKACS G, CHEN D, et al.

- CHoG: compressed histogram of gradients a low bit-rate feature descriptor [C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, Miami, USA: IEEE, 2009: 2504-2511.
- [11] JEGOU H, DOUZE M, SCHMID C, et al. Aggregating local descriptors into a compact image representation [C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, San Francisco, USA: IEEE, 2010: 3304-3311.
- [12] CALONDER M, LEPETIT V, OZUYSAL M, et al. Brief: computing a local binary descriptor very fast [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1281-1298.
- [13] RUBLEE E, RABAUDE V, KONOLIDGE K, et al. ORB: an efficient alternative to SIFT or SURF [C]. Proceedings of the International Conference on Computer Vision, Barcelona, Spain: IEEE, 2011: 2564-2571.
- [14] ENSOR A, HALL S. ColourFAST: GPU-based feature point detection and tracking on mobile device [C]. Proceedings of the International Conference of Image and Vision Computing New Zealand, Wellington, New Zealand: IEEE, 2013: 124-129.
- [15] LEUTENEGGER S, CHLI M, SIEGWART R. BRISK: binary robust invariant scalable keypoints [C]. Proceedings of the International Conference on Computer Vision, Barcelona, Spain: IEEE, 2011: 2548-2555.
- [16] MARI E, HAGER G, BURSCHKA G, et al. Adaptive and generic corner detection based on the accelerated segment test [C]. Proceedings of the European Conference on Computer Vision, Los Alamitos, USA: IEEE, 2010: 183-196.
- [17] ALAHI A, ORTIZ R, VANDERGHEYNST P. FREAK: fast retina keypoint [C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, Los Alamitos, USA: IEEE, 2012: 510-517.
- [18] FIELD G, GAUTHIER J, SHER A, et al. Functional connectivity in the retina at the resolution of photoreceptors [J]. Nature, 2010, 467(7316): 673-677.
- [19] ZHANG S, TIAN Q, HUANG Q M, et al. USB: ultra short binary descriptor for fast visual matching and retrieval [J]. IEEE Transactions on Image Processing, 2014, 23(8): 3671-3683.
- [20] LI L, YANG B, PAUL F, et al. BRINT: binary rotation invariant and noise tolerant texture classification [J]. IEEE Transactions on Image Processing, 2014, 23(7): 3071-3084.
- [21] KOTA I, RYOTA M, TOSHIYUKI N. BRIGHT: a scalable and compact binary descriptor for low latency and high accuracy object identification [C]. Proceedings of the International Conference on Image Processing, Melbourne, Australia: IEEE, 2013: 2915-2919.
- [22] SAJIB S, VINCENT D. ALOHA: an efficient binary descriptor based on Haar features [C]. Proceedings of the International Conference on Image Processing, Orlando, USA: IEEE, 2012: 2345-2348.
- [23] Xianwei X, LU T, Jianjiang F, et al. OSRI: A rotationally invariant binary descriptor [J]. IEEE Transactions on Image Processing, 2014, 23(7): 2983-2995.
- [24] 白丰, 张明路, 张小俊. 卫星装配中基于强区分性描述子的识别方法 [J]. 仪器仪表学报, 2014, 35(12): 2653-2662.
- BAI F, ZHANG M L, ZHANG X J. Identification method based on strong distinguishing descriptor in satellite assembly [J]. Chinese Journal of Scientific Instrument, 2014, 35(12): 2653-2662.
- [25] 高洪波, 王洪玉, 刘晓凯. 一种基于分层学习的关键点匹配算法 [J]. 电子与信息学报, 2013, 35(11): 2751-2757.
- GAO H B, WANG H Y, LIU X K. A keypoint matching method based on hierarchical learning [J]. Journal of Electronics & Information Technology, 2013, 35(11): 2751-2757.
- [26] 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 ZH M. Binary local descriptor based on rotative matching [J]. Journal of Image and Graphics, 2013, 18(10): 1315-1321.
- [27] 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 [J]. 自动化学报, 2014, 40(2): 255-266.
- HUI G B, LI D B, TONG Y F. A binary descriptor based on both optimized sampling pattern and image sub-patches [J]. Acta Automatica Sinica, 2014, 40(2): 255-266.
- [28] 庄东晔, 张冬明, 张勇东, 等. 随机亮度差量化的二进制特征描述中国图象图形学报 [J]. 中国图象图形学报, 2014, 19(4): 630-636.
- ZHANG D Y, ZHANG D M, ZHANG Y D, et al. Novel binary feature from intensity difference quantization [J]. Journal of Image and Graphics, 2014, 19(4): 630-636.
- [29] 惠国保, 李东波. 像素聚类改进二进制描述子鲁棒

- 性[J]. 中国图象图形学报, 2014, 19(4): 569-576.
- HUI G B, LI D B. Robust descriptor structured by pixel clustering[J]. Journal of Image and Graphics, 2014, 19(4): 569-576.
- [30] GIONIS A, INDIK P, MOTWANI R. Similarity search in high dimensions via hashing[C]. Proceeding of the International Conference on Very Large Databases, Edinburgh, Scotland: Chile, 2004: 518-529.
- [31] ANDONI A, INDYK P. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions[C]. Proceedings of the International Conference on Foundations of computer Science, Washington DC, USA: Crea Press, 2006: 459-468.
- [32] GUILLAMET D, VITRIA J. Non-negative matrix factorization for face recognition[C]. Proceedings of the Catalanian Conference on AI: Topics in Artificial Intelligence, London, Britain: Springer, 2002: 336-344.
- [33] SALAKHUTAINOV R, HINTON G. Semantic hashing[J]. International Journal of Approximate Reasoning, 2009, 50(7): 345-353.
- [34] WEISS Y, FERGUS R, TORRALBA A. Multidimensional spectral hashing[C]. Proceedings of the European Conference on Computer Vision, Firenze, Italy: IEEE, 2012: 340-353.
- [35] KULIS B, DARRELL T. Learning to Hash with binary reconstructive embeddings[C]. Proceedings of the Conference on advances in Neural Information Processing Systems, Vancouver, Canada: Curran Associates Inc. 2009: 1042-1050.
- [36] WANG J, KUMAR S, CHANG S F. Semi-supervised hashing for large scale search[J]. Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(12): 2393-2406.
- [37] MU Y D, SHEN J L, YAN S C. Weakly supervised hashing in kernel space[C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, Washington, USA: IEEE, 2010: 3344-3351.
- [38] TRZCINSKI T, LEPETIT V. Efficient discriminative projections for compact binary descriptors[C]. Proceedings of the European Conference on Computer Vision, Firenze, Italy: Springer Verlag, 2012: 228-242.
- [39] LIU W, WANG J, JI R, et al. Supervised Hashing with Kernels[C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, Rhode Island, USA: IEEE, 2012: 2074-2081.
- [40] STRECHA C, BRONSTEIN A, BRONSTEIN M, et al. LDAHash: improved matching with smaller descriptors[J]. Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(1): 66-78.
- [41] 毛晓蛟, 杨育彬. 一种基于子空间学习的图像语义哈希索引方法[J]. 软件学报, 2014, 25(8): 1781-1793.
- MAO X J, YANG Y B. Semantic hashing with image subspace learning[J]. Journal of Software, 2014, 25(8): 1781-1793.
- [42] 欧阳遹飞, 张寅, 张啸, 等. 结构化稀疏谱哈希索引[J]. 计算机辅助设计与图形学学报, 2012, 24(1): 60-67.
- OUYANG C F, ZHANG Y, ZHANG X, et al. Structured sparse spectral hashing[J]. Journal of Computer-Aided Design & Computer Graphics, 2012, 24(1): 60-67.
- [43] 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.
- [44] SIMONYAN K, VEDALDI A, ZISSERMAN A. Learning local feature descriptors using convex optimisation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(8): 1573-1585.
- [45] TORRALBA A, FERGUS R, WEISS Y. Small codes and large databases for recognition[C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, Alaska, USA: IEEE, 2008: 4587633.
- [46] TRZCINSKI T, CHRISTOUDIAS M, LEPETIT V, et al. Learning image descriptors with the boosting-trick[C]. Proceedings of the Conference on advances in Neural Information Processing Systems, Lake Tahoe, USA: Neural information processing system foundation, 2012: 269-277.
- [47] 刘冲, 张均东, 曾鸿, 等. 基于支持向量机的无穷维 AdaBoost 算法及其应用[J]. 仪器仪表学报, 2010, 31(4): 764-769.
- LIU C, ZHANG J D, ZENG H, et al. Infinite dimension AdaBoost algorithm based on support vector machine and its application[J]. Chinese Journal of Scientific Instrument, 2010, 31(4): 764-769.
- [48] TRZCINSKI T, CHRISTOUDIAS M, FUA P, et al. Boosting binary keypoint descriptors[C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, Portland, USA: IEEE, 2013: 2874-2881.

- [49] TRZCINSKI T, CHRISTOUDIAS M, LEPETIT V. learning image descriptors with boosting[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 597-610.
- [50] BIN F, QING Q K, TRZCINSKI T, et al. Receptive field selection for binary feature description[J]. IEEE Transactions on Image Processing, 2015, 23(6): 2583-2595.
- [51] 惠国保, 童一飞, 李东波. 基于改进的图像局部区域相似度学习架构的图像特征匹配技术研究[J]. 计算机学报, 2015, 38(6): 1148-1161.
HUI G B, TONG Y F, LI D B. Image feature matching based on improved patch similarity learning framework[J]. Chinese Journal of Computers, 2015, 38(6): 1148-1161.
- [52] 耿利川, 成运, 苏松志. RBFD: 一种鲁棒的图像局部二值特征描述子[J]. 计算机辅助设计与图形学学报, 2015, 27(5): 815-823.
GENG L CH, CHENG ZH, SU S Z, et al. RBFD: a robust image local binary feature descriptor[J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27(5): 815-823.
- [53] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [54] HINTON G E, ZEMEL R S. Autoencoders minimum description length and Helmholtz free energy[C]. Proceedings of the Advances in Neural Information Processing Systems, Burlington, USA: Morgan Kaufmann, 1994: 3-10.
- [55] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders[C]. Proceedings of the International Conference on Machine Learning, New York, USA: ACM, 2008: 1096-1103.
- [56] RANZATO M A, HINTON G E. Modeling pixel means and covariances using factorized third-order boltzmann machines[C]. Proceedings of the International Conference on Computer Vision and Pattern Recognition, San Francisco, USA: IEEE, 2010: 2551-2558.
- [57] LEE H, GROSSE R, RANGANATH R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]. Proceedings of the International Conference on Machine Learning, New York, USA: ACM, 2009: 609-616.
- [58] LAROCHELLE H, MURRAY I. The neural autoregressive distribution estimator[C]. Proceedings of the International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, USA: Microtome Publishing, 2011: 29-37.

作者简介

白丰(通讯作者) 2012 年于大连交通大学获得双专业学士学位, 现为河北工业大学硕博连读博士研究生。主要研究方向为机器人视觉与图像处理技术。

E-mail: bfhebut@163.com

Bai Feng (Corresponding author) received his B. Sc. from Dalian Jiaotong University in 2012. Now he is a Ph. D. candidate in Hebei University of Technology. His main research direction is robot vision and image processing technology.

张明路, 1985 年于河北工业大学获得学士学位, 1988 年于河北工业大学获得硕士学位, 1997 年于天津大学获得博士学位, 现为河北工业大学教授, 博士生导师, 研究生学院党委书记, 教育部创新团队发展计划“特殊环境下服役机器人关键技术”团队负责人, 主要研究方向为特种机器人机构与应用技术研究、多智能体技术。

E-mail: zhangml@hebut.edu.cn

Zhang Minglu received his B. Sc. and M. Sc. both from Hebei University of Technology in 1985 and 1988, and Ph. D. from Tianjin University in 1997, respectively. Now he is a professor, Ph. D. supervisor in Hebei University of Technology. His main research interests is special robot mechanism and application technology, and multi-agent technology.