



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2013
08
VOL.18

ISSN1006-8961
CN11-3758/TB



北魏



隋



初唐

数字文化遗产研究 P968

综述

尺度不变特征变换算子综述

刘立, 詹茵茵, 罗扬, 刘朝晖, 彭复员 885

图像处理和编码

改进的H.264帧间模式选择算法

任克强, 张旭光, 罗会兰 893

均质雾天下摄像机动态标定的车速估计

宋洪军, 陈阳舟, 郜园园 901

基于人脸五官特征的空域差错掩盖算法

张江鑫, 谢晋, 邝万坤 913

采用圆谱-傅里叶矩的图像区域复制粘贴篡改检测

秦娟, 李峰, 向凌云, 殷蓓蓓 919

图像分析和识别

基于视频感知哈希的视频篡改检测与多粒度定位

朱映映, 文振焜, 杜以华, 邓良太 924

视觉注意模型的道路监控视频关键帧提取

刘云鹏, 张三元, 王仁芳, 张引 933

格拉斯曼流形上的半监督判别分析

姜伟, 陆瑶, 杨炳儒 944

精神疲劳实时监测中多面部特征时序分类模型

陈云华, 张灵, 丁伍洋, 严明玉 953

图像分类中的概率乘积核函数

杨赛, 赵春霞 961

轮廓整体结构约束的壁画图像相似性度量

唐大伟, 鲁东明, 杨冰, 许端清 968

应用色彩空间聚类方法实现道路建模

向宸薇, 王拓, 于舰 976

图像理解和计算机视觉

杂乱背景和摄像机移动下的时空兴趣点检测

刘长红, 陈勇, 王明文 982

形状和颜色混合不变量在图像检索中的应用

公明, 曹伟国, 李华 990

计算机图形学

结合统计分类和边缘特征的最优视图提取

潘翔, 陈敖, 章国栋 1004

医学图像处理

局部熵驱动下的脑MR图像分割与偏移场恢复耦合模型

张建伟, 杨红, 陈允杰, 方林, 詹天明 1011

改进的模糊聚类实质肺结节3维分割

李翠芳, 聂生东, 王远军, 孙希文, 郑斌 1019

遥感图像处理

高光谱海岸带区域分割的活动轮廓模型

王相海, 金弋博 1031

改进的基于非局部均值的极化SAR相干斑抑制

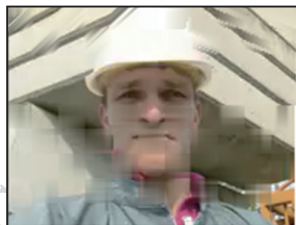
赵忠民, 赵拥军, 牛朝阳 1038

小波变换和稀疏表示相结合的遥感图像融合

刘婷, 程建 1045

《中国图象图形学报》论文摘要写作要求

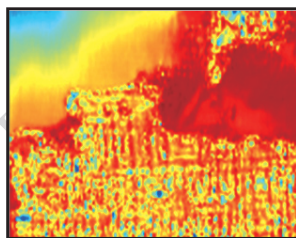
1054



基于人脸五官特征的空域差错掩盖算法 (第913页)



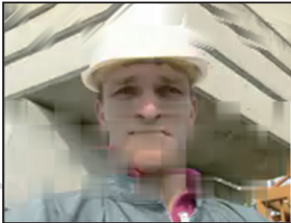
应用色彩空间聚类方法实现道路建模 (第976页)



改进的基于非局部均值的极化SAR相干斑抑制 (第1038页)

CONTENTS

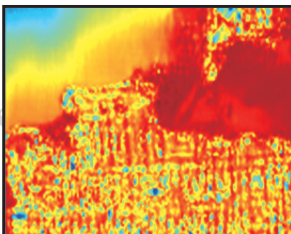
JOURNAL OF IMAGE AND GRAPHICS



Spatial error concealment algorithm based on the facial features (P913)



Road modeling using color spatial clustering (P976)



Improved polarimetric SAR speckle filter based on non-local means (P1038)

Review

- Summarization of the scale invariant feature transform
Liu Li , Zhan Yinyin, Luo Yang, Liu Chaohui, Peng Fuyuan 885

Image Processing and Coding

- Improved inter frame mode selection algorithm for H.264
Ren Keqiang, Zhang Xuguang, Luo Huilan 893
- Vehicle velocity estimation based on camera calibration under homogenous fog
Song Hongjun, Chen Yangzhou, Gao Yuanyuan 901
- Spatial error concealment algorithm based on the facial features
Zhang Jiangxin, Xie Jin, Kuang Wankun 913
- Detection of image region copy-move forgery using radial harmonic Fourier moments
Qin Juan, Li Feng, Xiang Lingyun, Yin Changming 919

Image Analysis and Recognition

- Video forgery detection and multi-granularity location based on video perceptual hashing
Zhu Yingying, Wen Zhenkun, Du Yihua, Deng Liangtai 924
- Key frame extraction based on the visual attention model for lane surveillance video
Liu Yunpeng, Zhang Sanyuan, Wang Renfang, Zhang Yin 933
- Semi-supervised discriminant analysis on Grassmannian manifold
Jiang Wei, Lu Yao, Yang Bingru 944
- Time-series classification model based on multiple facial feature for real-time mental fatigue monitoring
Chen Yunhua, Zhang Ling, Ding Wuyang, Yan Mingyu 953
- Probability product kernels in image classification
Yang Sai, Zhao Chunxia 961
- Similarity metrics between mural images with constraints of the overall structure of contours
Tang Dawei, Lu Dongming, Yang Bing, Xu Duanqing 968
- Road modeling using color spatial clustering
Xiang Chenwei, Wang Tuo, Yu Jian 976

Image Understanding and Computer Vision

- Spatio-temporal interest point detection in cluttered backgrounds with camera movements
Liu Changhong, Chen Yong, Wang Mingwen 982
- Application of combined shape and color invariants to image retrieval
Gong Ming, Cao Weiguo, Li Hua 990

Computer Graphics

- Best view selection based on statistical classification and edge features
Pan Xiang, Chen Ao, Zhang Guodong 1004

Medical Image Processing

- Brain MR image segmentation and bias correction model based on local entropy
Zhang Jianwei, Yang Hong, Chen Yunjie, Fang Ling, Zhan Tianming 1011
- Segmentation of sub-solid pulmonary nodules based on improved fuzzy C-means clustering
Li Cuifang, Nie Shengdong, Wang Yuanjun, Sun Xiwen, Zheng Bin 1019

Remote Sensing Image Processing

- The active contour model for segmentation of coastal hyperspectral remote sensing image
Wang Xianghai, Yin Yibo 1031
- Improved polarimetric SAR speckle filter based on non-local means
Zhao Zhongmin, Zhao Yongjun, Niu Chaoyang 1038
- Remote sensing image fusion with wavelet transform and sparse representation
Liu Ting, Cheng Jian 1045

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2013)08-0885-08

论文引用格式: 刘立, 詹茵茵, 罗扬, 刘朝晖, 彭复员. 尺度不变特征变换算子综述[J]. 中国图象图形学报, 2013, 18(8): 885-892. [DOI: 10.11843/jig.20130801]

尺度不变特征变换算子综述

刘立^{1,2}, 詹茵茵¹, 罗扬¹, 刘朝晖¹, 彭复员²

1. 南华大学计算机科学与技术学院, 衡阳 421001; 2. 华中科技大学电信系, 武汉 430076

摘要: 随着计算机软件与硬件技术的发展, 计算机视觉算法逐渐成为图像处理领域的研究热点。其中 SIFT (scale invariant feature transform) 算法是目前机器视觉领域应用最成功的算法之一。由于在尺度不变、旋转不变、光照不变等方面的独特优势, SIFT 被广大视觉领域的研究者借鉴与学习。但是 SIFT 算法本身也存在一些问题, 如仿射性能不太理想, 计算复杂度过高等, 因此针对它的多种改进算法不断出现。本文对 SIFT 的发展历史、SIFT 算法的演变以及它不同领域的典型应用给出了一个比较全面的综述, 比较了各类算法的优缺点。最后给出了该算法未来可能的发展方向, 为视觉研究者提供参考。

关键词: 尺度不变; SIFT; 计算机视觉; 图像匹配

Summarization of the scale invariant feature transform

Liu Li^{1,2}, Zhan Yinyin¹, Luo Yang¹, Liu Chaohui¹, Peng Fuyuan²

1. Department of Computer Science and Technology, NanHua University, Hengyang 421001, China;

2. Department of Telecommunication, HuaZhong Science and Technology University, Wuhan 430076, China

Abstract: With the development of software and hardware technique, computer vision has become a hot research fields in image processing. Scale invariant feature transform (SIFT) is one of the most successful vision algorithm nowadays and it is widely studied by the computer vision community because of its unique features. SIFT is scale invariant, rotation invariant and illumination invariant. However, it also has some problems such as it is only part affine has a rather the high computation complexity. Many extended or modified algorithms of the SIFT are developed unceasingly. In this paper, we summarize the history, the evolved processing, and the application of the SIFT and compares those algorithm effects. At last, the paper discusses the feature direction and provides reference for computer vision researchers.

Key words: scale invariant; scale invariant feature transform (SIFT); computer vision; image match

0 引言

自从 Marr 的视觉理论产生以来, 人们对利用机器视觉来解决实际生活问题产生了极大的兴趣。Marr 理论的一个重要观点认为, 图像要素组成的基元与物理实在之间有重要联系, 而这些要素有: 边

缘、棒、斑点和端点, 它们可以通过对图像中灰度变化的检测和定位得到, 同时这些要素就构成了图像的视觉特征。国内外的许多学者对特征点的检测投入了大量的精力进行研究, 其中以 Moravec^[1] 与 Harris 等人^[2] 的成果最为显著。Moravec 首次提出了兴趣点的概念, 但是该方法提取的兴趣点不具备旋转不变性。Harris 改进了 Moravec 的算法, 用微分

收稿日期: 2012-11-13; 修回日期: 2013-01-10

基金项目: 国家自然科学基金项目 (60475024, 60872076); 航天技术创新基金项目 (2006AA09Z203); 湖南省科技计划 (2010GK3012)

第一作者简介: 刘立 (1971—), 男, 副教授, 湖南省青年骨干教师, 2008 年于华中科技大学获信号与信息处理专业博士学位, 主要研究方向为数字图像处理、计算机视觉。E-mail: LL710915@yahoo.com.cn

算子代替亮度块构造了具有结构信息的 2×2 的 Harris 矩阵,并通过特征值来判断该点是否是角点。由于 Harris 提出的算法具有良好的旋转不变性和光照不变性而被视觉领域广泛采用与借鉴,并成为了经典的特征检测算法。但是 Harris 算法的最大缺点是不具备尺度不变性,一些学者想在保持 Harris 算法的优点的前提下增加尺度不变性,如 Mikolajczyk 和 Schmid 提出的 Harris-Laplacian^[3] 以及 Harris-affine^[4] 检测算子就是在 Harris 算子的基础上增加了尺度不变性以及仿射不变的特性。与他们不同,加拿大著名学者 Lowe 在 Linderberg 关于尺度空间理论的基础上^[5],于 1999 年提出了一种全新的特征点提取方法 SIFT (scale invariant feature transform)^[6],并在 2004 年进一步完善^[7]。该算法与人类视觉机制息息相关^[8]。与 Harris 一样,该算法也是局部特征研究领域的里程碑式的工作^[9]。截止 2012 年 10 月,对比 Harris、Canny 以及 SIFT 3 种经典的特征提取算法的文献引用数据如表 1 所示。

表 1 文献引用数据比较

Table 1 The comparison of Citation data

算法	发表年份	Google 学术/篇	平均引用量 (篇/年)
Harris	1988	6 946	289.4
Canny	1986	14 067	541.1
SIFT	2004	15 292	1 911.5

从表 1 可以看出,SIFT 算法的引用已经超过了 Harris 与 Canny 算法,成为视觉研究领域的热点。

1 SIFT 算法简介

SIFT 算法包含 4 个主要步骤:

1) 尺度空间极值检测

为了提高效率,Lowe 在算法中用 DOG (difference of gaussian) 算子近似代替了 LOG (laplacian of gaussian) 算子来建立图像的差分尺度空间,然后在差分图像以及尺度 3 维空间中搜索极值点作为潜在的尺度与旋转不变特征点。

2) 关键点的定位

为了提高定位精度,算法采用了一些辅助手段,如泰勒公式、插值等数学工具来提高关键点在图像中的位置精度以及空间的尺度精度。

3) 指定方向参数

利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数。这样每个特征点就拥有 4 个重要参数,在图像中的横坐标与纵坐标、尺度参数以及方向参数,保证旋转不变性以及尺度不变性。

4) 生成 SIFT 特征向量

统计以特征点为中心的局部区域梯度,形成 128 维梯度特征向量,最后对特征向量归一化,去除光照影响,保证光照不变性。通过以上步骤产生的特征点具有尺度不变、旋转不变以及光照不变等性能,如图 1 所示。

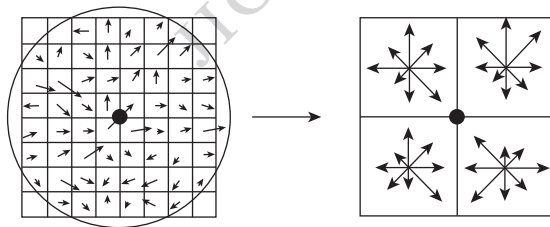


图 1 SIFT 算法生成特征向量

Fig. 1 The feature vectors of the SIFT algorithm

2 SIFT 算法族

SIFT 算法是 Lowe 在尺度空间理论基础上通过详尽实验验证的算法,经受了不同类型图像的考验,具有极高的综合性能^[10]。虽然 SIFT 是一种能力很强的算子,但面对各种应用场景,仍然需要进一步的扩展和增强,主要表现在以下 4 个地方:

1) 算法的复杂度比较高,每个特征点均采用 128 维向量来表示,对于特征比较多的图像来说,所需时间代价比较大;

2) 算法对图像特征的提取还不够充分,匹配数量不够多,匹配效果还不够理想;

3) 算法的仿射不变性不够好;

4) 对于具有相似背景环境的图像,误匹配较多。

针对这 4 方面的问题,国内外的许多学者做了大量的改进或扩展,取得了比较好的效果。

2.1 通过降维提高 SIFT 算法效率

SIFT 算法中,每个特征点形成 128 维特征向量,这无论是在匹配或者向量的形成等过程中都将消耗大量的时间,因此可以考虑降维的方法来提高效率,其中 Ke 与 Sukthankar 于 2004 年提出一种 PCA-SIFT 算法,他认为可以将 PCA (principle component analysis) 方法与 SIFT 算子结合起来减少时间复

杂度^[11]。具体做法是:不同于 SIFT 算法中采用的 $4 \times 4 \times 8$ 描述子,而是在 41×41 的图像块上计算 $39 \times 39 \times 2$ 个梯度导数,获得 3 042 维向量后通过 PCA 方法降至 36 维,从而提高运算速度。算法虽然通过降维提高了匹配速度,但是由于采用了 PCA 方法以及前期形成的 3 042 维向量,因此在向量形成过程中所消耗的时间比原算法更多,比较数据如表 2 所示。

表 2 PCA-SIFT 算法与 SIFT 算法时间性能实验数据

Table 2 The time performance of PCA-SIFT and SIFT

	时间性能/s
特征点检测与定位	2.63
SIFT 特征向量	1.59
PCA-SIFT 特征向量	1.64
SIFT 匹配	2.20
PCA-SIFT 匹配	0.58

可以看出,总时间从 6.42 s 降低到 4.85 s,速度提高有限,而算子的稳定性以及尺度不变性存在一定程度的下降,而且由于算法中采用的投影矩阵需要一系列有代表性的图像,因此算法不具有普适性。

出于同样的目的,Bay 提出了 SURF (speeded up robust features) 算法对 SIFT 进行了 3 个方面的改进^[12],首先图像没有采用 2:1 的下采样方式建立高斯金字塔,每幅图像都是采用同样的分辨率。其次,采用不同大小的方框滤波器同原始图像进行卷积运算,以提高计算速度。为了进一步减少计算时间,还采用了基于积分图形的快速海森矩阵的方法来检测特征点。最后分别形成 36 维或 128 维向量满足不同的要求,实验结果表明,与 SIFT 相比,SURF 在保持其他性能不明显下降的情况下,速度提高了 3 倍。但是文献[13]实验数据显示,SURF 在尺度不变以及旋转不变等性能上表现并不理想。

国内学者在减少运算时间上也做了许多有益的工作,比较有代表的是刘立等人研究的 SSIFT (simplified SIFT) 算法^[14],算法采取的关键步骤包括:1) 在多尺度空间特征点形成之后,以特征点为中心采用圆形窗体来确定需要统计的邻域范围,这样就可以保证完全的旋转不变性;2) 在该区域里统计 12 个梯度方向,形成 12 维特征向量,大大缩短了计算时间;3) 在该向量里查找最大的梯度方向元素,并将该元素循环左移至向量头部,再归一化该向量,以保证光照不变性,如图 2 所示。算法不但因为降维

而使得算法的效率提高,而且因为将原 SIFT 算法中的两次梯度统计合并为一次而进一步提升了算法的速度。实验结果表明,SSIFT 算法所消耗的时间降低 1/10 以上,同时旋转不变性也强于原算法。该算法在多幅图像匹配过程中初步过滤非目标图像以及在仿射与尺度变化不大(如全景拼图)的情况下有较大实用价值^[15]。

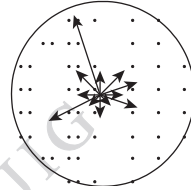


图 2 SSIFT 算法生成特征向量

Fig. 2 The feature vectors of the SSIFT algorithm

2.2 通过扩展提升算法效果

虽然 SIFT 的性能综合性能比较高,但是许多学者依然在匹配的稳定性以及匹配的数量上提出了更多的要求,他们认为图像的特征并没有在 SIFT 算法中得到充分的展示,如 SIFT 算法只是利用了灰度图像进行匹配,没有充分利用彩色特征,为了提高算法的效果,Abdel-Hakim 与 Farag 基于颜色不变量理论,利用图像 RGB (red、green、blue) 色彩信息构建 CSIFT (colored SIFT) 描述符^[16],再进行特征匹配,求解出图像变换矩阵参数,实现了基于彩色图像的配准。图像配准实验表明:CSIFT 算法能够充分利用图像的色彩信息,CSIFT 描述符不仅继承了 SIFT 算法良好的特征不变性,对图像平移、旋转、尺度变化、光照变化也具有更高的独特性和稳健性,图像配准误差矩阵说明 CSIFT 算法可以提高图像配准精度。目前该算法被国内外学者广泛应用在场景匹配^[17]、彩色图像配准^[18]等领域。也有一些学者将彩色图像进行分解,分别求得不同分量的 SIFT 特征,再融合特征进行匹配。如 Bosch^[19] 在 HSV (hue、saturation、value) 颜色空间的 3 个通道中分别计算 128 维 SIFT 算子,最后生成 128×3 维的向量。而国内的赵坤^[20]、高健^[21]等人采用同样的方法在 RGB 颜色空间中分别计算 SIFT 向量。Burghouts^[22] 通过实验表明彩色 SIFT 算法具有更好的性能,而 SANDE^[23] 则总结对比了不同的彩色 SIFT 算子性能。

Mikolajczyk 等人提出了一种 SIFT 变体的描述子 GLOH (gradient location and orientation histogram)^[10],使用对数极坐标分级结构替代 Lowe 使用

的4象限。即以特征点为中心,在不同的圆环带上将整个区域分解为大小相同的扇形区间并统计梯度(图3),形成272维的梯度向量,最后借鉴PCA-SIFT算法,通过训练最终用PCA方法生成128维向量。实验结果表明,GLOH与SIFT相比,大多数情况下在尺度不变、光照不变以及选择不变等方面略有优势,但时间复杂度增加。

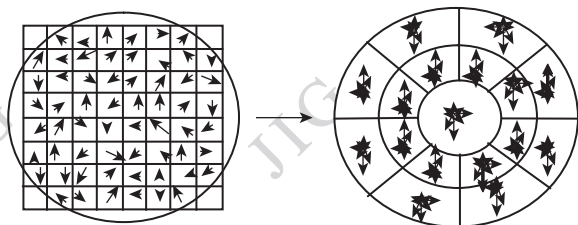


图3 GLOH算法生成特征向量

Fig. 3 The feature vectors of the GLOH algorithm

Scovanner等人则在原算法的基础上增加时间维度,在连续多帧图像中计算SIFT特征向量,形成3DSIFT(3-dimension SIFT)算子,并通过行为识别对比实验,表明3DSIFT算子的精度大大优于原SIFT算法^[24]。另外针对SIFT算子过于稀疏,不利于在物体识别中的应用,提出了Dense SIFT算法并在场景分类中取得了很好的效果^[25]。

受SIFT算法的启发,Dalal以及Triggs将直方图统计方法进一步扩展,提出了HOG(histograms of oriented gradients)方法^[26],该算法目前被广泛应用在行人检测等方面^[27-31]。使用该方法,图像被分割为较小的连通单元(cell),然后计算并统计该单元中各个像素点的梯度或边缘方向直方图形成该单元的特征向量。把各个单元组合成大的、空间上连通的区间(block)并对梯度强度进行归一化,以保证光照不变性。针对该区间的形状又分为R-HOG(rectangular histogram of gradients)和C-HOG(circular histogram of gradients)算法,如图4所示。

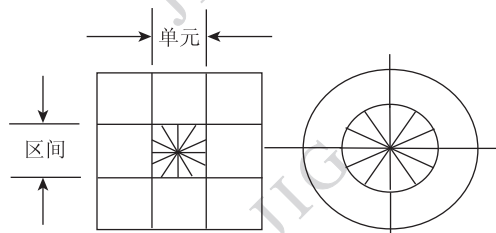


图4 R-HOG和C-HOG算法的区间形状

Fig. 4 The region shape of R-HOG and C-HOG

2.3 提高仿射不变性

SIFT算法不是完全的仿射不变性,PCA-SIFT算法在这个方面也只是稍微有所提升,提高算法的仿射不变性也是一个难点。Morel等人提出了一种仿射不变性极高的SIFT算法ASIFT(affine SIFT)^[32],该算法不但在尺度上进行采样以保证尺度不变性,而且还模拟了导致仿射变化的两个角度,即相机的经度与纬度,在这两个角度上也进行全方位的采样,如图5所示。这样就考虑到了图像的各种仿射情况,再针对这些采样生成的图像进行一一匹配,以保证仿射不变性。实验数据表明,该算法仿射不变性大大优于原算法。但由于增加了两个参数,算法的复杂度增加了,计算时间也成倍增加。

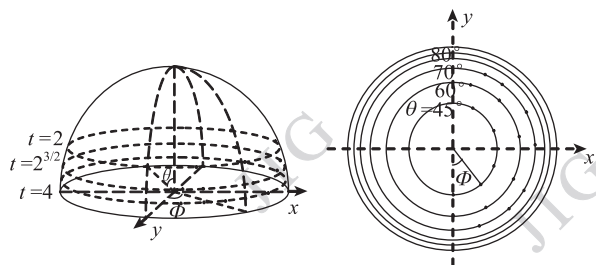


图5 ASIFT算法中对经度与纬度的采样

Fig. 5 The longitude and dimension of sampling in ASIFT algorithm

2.4 增加全局信息,提高相似环境适应性

SIFT算法是基于局部区域的梯度统计算子,如果存在大量相似局部区域,就会使得不同位置特征点的特征向量会有很大程度相似,从而导致大量的误匹配,如纹理背景。Mortensen在SIFT特征中加入全局纹理特征,即SIFT+GC(SIFT+global context)^[33],对于每一个被检测到的特征点,建立一个由两部分组成的向量,一部分是代表局部特性的SIFT描述子,另一部分是用来区分相似局部特征的全局纹理向量,即

$$F = \begin{bmatrix} \omega L \\ (1 - \omega) G \end{bmatrix} \quad (1)$$

式中, L 是长度为128的SIFT向量, G 是全局向量, ω 是相对权重因子。在基于局部信息的SIFT向量中加入基于全局纹理信息的全局向量,降低了由于局部信息相似而造成的误匹配的概率。纪华等人^[34]将该算法中k-d树搜索算法改为BBF(best bin first)算法,取得了更好的效果。

到目前为止,主要的SIFT家族图谱如图6所示。

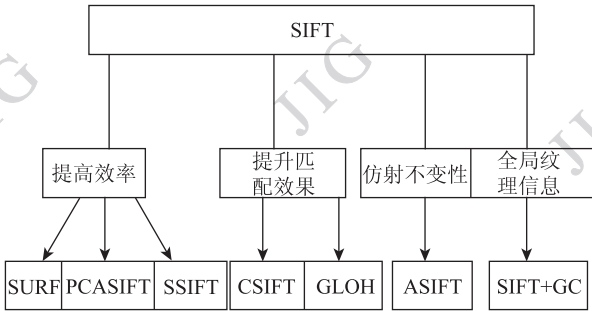


图 6 SIFT 主要家族谱 (截止 2012 年)

Fig. 6 SIFT main family (until 2012)

以上各种算法在研究中均与 SIFT 进行了对比实验,根据文献提供的评估结果可以在尺度不变性、旋转不变性、仿射不变性以及时间复杂度等方面对 SIFT 算法族给出一个定性的比较,以 + 号表示算法在某性能具有一定的优势,以 ++ 号表示算法在某性能表现最有优势,结果如表 3 所示。

表 3 SIFT 算法族的性能比较

Table 3 The performances of SIFT family

算法	尺度不变性	旋转不变性	仿射不变性	时间性能
SIFT	++	+	+	
SURF	+		+	++
PCA-SIFT	+			+
GLOH	+	+	+	
CSIFT	+	+	+	
ASIFT	+	+	++	
SSIFT		++		++
SIFT + GC	+	+		

虽然针对 SIFT 的改进与研究还在继续,但是到目前为止,这些算法的优势只是体现在某个或者某些方面,从整体综合性能来讲,还是 SIFT 最具有实用价值。

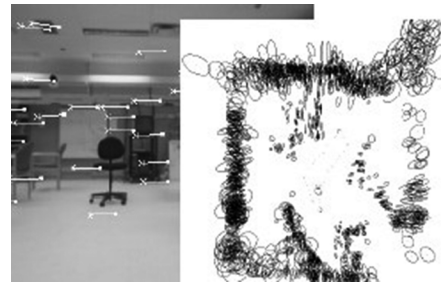
3 SIFT 算法在视觉领域的典型应用

SIFT 从一开始产生就受到广泛关注,在其为基础的各种应用也应运而生,以 SIFT 为基础, Lowe 的研究中心将该算法应用在 4 个主要的研究方向上,分别是: 移动机器人定位识别与地图生成、增强现实、全景拼图、FLANN(fast library for approximat nearest neighbors)。

3.1 移动机器人定位识别与地图生成

移动机器人的一个基本功能就是建立一个地图

环境以便随时知道自己的位置,使用 SIFT 算法可以在没有先验知识的前提下建立详细的 3 维地图并具备更新地图的能力。具体做法是:采用双相机或多相机针对同一场景在不同的角度拍摄两幅或多幅图像,通过 SIFT 算法进行匹配,获得基本矩阵与本质矩阵,采用立体视觉方法得到每个特征点的 3 维信息实现目标定位,最后这些 3 维坐标信息被存入数据库中作为地图标记(图 7)。Stephen 带领的团队在这个方面做了大量的工作并取得了很好的成绩^[35-44]。



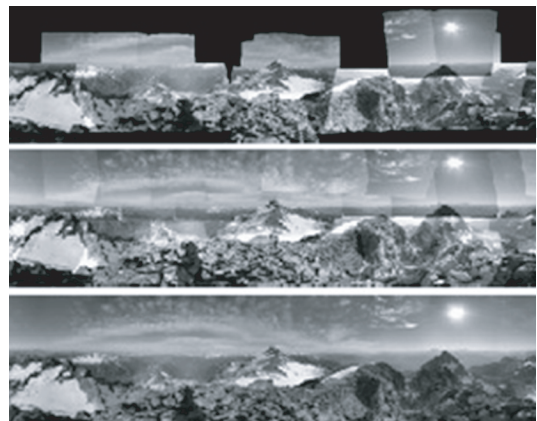
(来自<http://www.cs.ubc.ca/~lowe/research.html>)

图 7 基于 SIFT 的移动机器人定位与地图生成

Fig. 7 Mobile robot localization and map building based on SIFT

3.2 全景拼图

通过对多幅图像进行拼接处理,一个普通的数码相机也可以创建广角高分辨率的全景图像。然而,以往的制作全景图像的方法需要复杂的用户接口、固定焦距和曝光装置以及按一个特定的线性序列排列的图像。采用 SIFT 算法可以不受这些条件的限制,不同视角拍得的图像一一匹配,然后通过匹配的特征点将图像拼接在一起,最后在图像拼接位置采用平滑处理获得最后的全景图(图 8)。Brown



(来自<http://www.cs.ubc.ca/~lowe/research.html>)

图 8 基于 SIFT 的全景拼图

Fig. 8 Panoramas based on SIFT

团队从 2002 年开始进行了该方面的研究^[45-47]。

3.3 增强现实

增强现实可以将人工生成的虚拟目标与现实世界中的自然目标集成在一起,实现现实世界中不可能存在的效果。这在电影业、旅游业、汽车头盔显示以及可视化手术等方面有着重要的地位。采用 SIFT 算法可以为这些虚拟目标建立 3 维模型,该模型可以在序列图像中被识别出来并实现精确定位和跟踪^[48-49]。图 9 中茶壶是被虚拟出来的物品,其他的目标是现实世界中的物品,可以看出两个世界的物品被完美地结合起来了。



(来自 <http://www.cs.ubc.ca/~lowe/research.html>)

图 9 基于 SIFT 的增强现实

Fig. 9 Augmented reality based on SIFT

3.4 FLANN

在目标识别算法中最耗时的部分是在大规模的数据集中寻找匹配的高维向量对。而线性搜索是目前最快的方法。采用近似算法虽然有可能使得性能上有些损失,但是带来的比线性搜索快千倍以上的效率。以 Muja 为代表的团队采用了快速近似算法 FLANN 解决了最近领域的搜索问题,为提高 SIFT 特征向量集的匹配速度作出了重大贡献^[50-52],同时也进一步提高了 SIFT 算法的实用性。

3.5 其他研究

SIFT 算法其他许多领域也有广泛的应用,如生物特征识别,包括人脸^[53-60]、虹膜识别^[61-62]等,以及 3 维重建^[63-64]、车牌识别^[65-66]等。随着硬件的高速发展,SIFT 算法的应用也会越来越广泛。

4 结 语

对 SIFT 算法的基本原理,SIFT 算法的扩展与简

化以及它的应用作了详细的论述与说明。SIFT 算法虽然在许多领域取得成功,但其本身还具有许多难以解决的问题。未来的主要发展方向可能集中在两个方面:1)与其他算法结合(如各种聚类分析算法)进一步完善自身算法,使之具备极强的环境适应能力(如人类视觉的学习与自适应能力),可以结合人工智能或神经网络等算法,由于增加了更多的环节,因此对硬件也提出了更高的要求,可以考虑在 FPGA (field programmable gate array) 等硬件中实现算法;2)研究更加快速的简便算法,在实时性要求比较高的场合对快速检索图像数据有非常重要的意义。随着这些研究的日益成熟,SIFT 算法将在更多的应用领域如反恐安全中的面部特征识别,指纹识别,运动目标跟踪,太空探索中的机器人目标定位,深海资源探索中的地图生成以及地形 3 维重建等方面发挥强大的作用。

参考文献 (References)

- [1] Moravec H. Rover visual obstacle avoidance [C]//Proceedings of International Joint Conference on Artificial Intelligence. Vancouver, Canada: University Of British Columbia, 1981:785-790.
- [2] Harris C, Stephens M. A combined corner and edge detector [C]//Proceedings of the 4th Alvey Vision Conference. Manchester, UK:IEEE,1988:147-151.
- [3] Mikolajczyk K, Schmid C. Indexing based on scale invariant interest points [C]//Proceedings of the 8th International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001: 525-531.
- [4] Mikolajczyk K, Schmid C. An affine invariant interest point detector [C]// Proceedings of the 8th International Conference on Computer Vision. Vancouver, Canada: IEEE, 2002:128-142.
- [5] Lindeberg T. Scale-space theory: a basic tool for analyzing structures at different scales [J]. Journal of applied statistics, 1994, 21:224-270.
- [6] Lowe, D G. Object recognition from local scale-invariant features [C]//Proceedings of International Conference on Computer Vision. Corfu, Greece: IEEE, 2009:1150-1157.
- [7] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [8] Lowe D G. Towards a computational model for object recognition in IT cortex [C]//Proceedings of the 1st IEEE International Workshop on Biologically Motivated Computer Vision. Seoul, Korea: IEEE, 2000: 20-31.
- [9] Wang Y M, Wang G J. Image Local Invariant Features and Descriptors [M]. Beijing: National Defense Industry, 2010. [王永明, 王贵锦. 图像局部不变性特征与描述 [M]. 北京:国防工业出版社, 2010.]
- [10] Mikolajczyk K, Schmid C. A performance evaluation of local de-

- scriptors[C]//Proceedings of International Conference on Computer Vision and Pattern Recognition. Madison, USA: IEEE, 2003:17-122.
- [11] Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors [C]//Proceedings of CVPR. Washington DC, USA: IEEE, 2004: 506-513.
- [12] Bay H, Tuytelaars T, Gool L. SURF: speeded up robust features [C]//Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006:404-417.
- [13] Luo J, Oubong G. A comparison of SIFT, PCASIFT and SURF [J]. International Journal of Image Processing, 2009, 3(4): 143-152.
- [14] 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. [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.]
- [15] Sui L S, Wang Y, Wang L. An image mosaic algorithm based on disparity gradient [J]. Chinese journal of Stereology and Image analysis, 2009, (14):156-161. [隋连升, 王钰, 王磊. 基于视差梯度的图像拼接算法研究[J]. 中国体视学与图像分析, 2009, (14):156-161.]
- [16] Abdel-Hakim A E, Farag A A. CSIFT: a SIFT descriptor with color invariant characteristics[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE, 2006, 2:1978-1983.
- [17] Deng L Q, Chen D W, Yuan Z M, et al. Cartoon scene matching system using color and CSIFT feature [C]// Proceedings of the 3rd International Congress on Image and Signal Processing. Québec. Canada: Springer, 2010, 5:2519-2524.
- [18] Zhang R J, Zhang J Q, Yang C, et al. Study on color image registration technique Based on CSIFT [J]. Acta Photonica Sinica, 2008, 28(11):2097-2103. [张锐娟, 张建奇, 杨翠, 等. 基于 Csift 的彩色图像配准技术研究[J]. 光学学报, 2008, 28(11):2097-2103.]
- [19] Bosch A, Zisserman A, Muñoz X. Scene classification using a hybrid generative /discriminative approach [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(4): 712-727.
- [20] Zhao K. Research on feature extraction and matching for underwater unstructured environment [D]. WuHan: Huazhong Science and Technology University, 2007. [赵坤. 水下非结构化环境中的特征提取与匹配算法研究[D]. 武汉: 华中科技大学, 2007.]
- [21] Gao J, Huang X H, Peng G, et al. Color-based SIFT feature point detection and matching [J]. Computer and Engineering and Applications, 2007, 43(34):10-14. [高健, 黄心汉, 彭刚, 等. 基于彩色的 SIFT 特征点提取与匹配[J]. 计算机工程与应用, 2007, 43(34):10-14.]
- [22] Burghouts G J, Geusebroek J M. Performance evaluation of local color invariants [J]. Computer Vision and Image Understanding, 2009, 133:48-62.
- [23] Sande K E A, Gevers T, Snoek C G M. Evaluation of color descriptors for object and scene recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage. Alaska, USA: IEEE, 2008:1-8.
- [24] Scovanner P, Ali S, Shah M. A 3-dimensional SIFT descriptor and its application to action recognition [C]//Proceedings of the 15th International Conference on Multimedia. Augsburg, Germany: ACM, 2007: 357-360.
- [25] Bosch A, Andrew Z, Xavier M. Scene classification via Plsa [C]//Proceedings of ECCV 2006. Graz Austria: Springer, 2006, 3954: 517-530.
- [26] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//Proceedings of CVPR 2005. San diego, USA: IEEE, 2005:886-893.
- [27] Zhu Q, Avidan S, Yeh M C, et al. Fast human detection using a cascade of histograms of oriented gradients [C]//Proceedings of CVPR 2006. New York, USA: IEEE, 2006:1492-1498.
- [28] Grabner H, Bischof H. On-line boosting and vision [C]//Proceedings of CVPR 2006. New York, USA: IEEE, 2006:260-267.
- [29] Suard F, Rakotomamonjy A, Bensrhair A, et al. Pedestrian detection using infrared images and histograms of oriented gradients [C]//The IEEE Intelligent Vehicles Symposium. Tokyo, Japan: IEEE, 2006:206-212.
- [30] Laptev I. Improvements of object detection using boosted histograms [C]//Proceedings of Machine Vision Conference. Edinburgh, UK: ACM, 2006: 949-958.
- [31] Enzweiler M, Gavrilu D M. Monocular pedestrian detection: survey and experiments [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(12): 2179-2195.
- [32] Morel J M, Yu G. ASIFT: a new framework for fully affine invariant image comparison [J]. SIAM Journal on Imaging Sciences, 2009, (2):438-469.
- [33] Mortensen E N, Deng H L, Shapiro L. A SIFT descriptor with global context [J]. Computer Vision and Pattern Recognition, 2005, (1):20-25.
- [34] Ji H, Wu Y H, Sun H H. SIFT feature matching algorithm with global information [J]. Optics and Precision Engineering, 2009, 17(2):339-444. [纪华, 吴元昊, 孙宏海. 结合全局信息的 SIFT 特征匹配算法[J]. 光学精密工程, 2009, 17(2):339-444.]
- [35] Stephen S, David G L, Jim L. Local and global localization for mobile robots using visual landmarks [C]//Proceedings of IEEE International Conference on Intelligent Robots and Systems. Maui, Hawaii: IEEE, 2001:414-420.
- [36] Stephen S, David G L, Jim L. Vision-based mobile robot localization and mapping using scale-invariant features [C]//Proceedings of IEEE International Conference on Robotics and Automation. Seoul, Korea. :IEEE, 2001:2051-2058.
- [37] Stephen S, David G L, Jim L. Vision-based mapping with backward correction [C]//Proceedings of International Conference on Intelligent Robots and Systems. Lausanne, Switzerland: IEEE, 2002:153-158.
- [38] Stephen S, David G L, Jim L. Global localization using distinctive visual features [C]//Proceedings of International Conference

- on Intelligent Robots and Systems. Lausanne, Switzerland; IEEE, 2002: 226-231.
- [39] Stephen S, David G L, Jim L. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks [J]. International Journal of Robotics Research, 2002, 21 (8): 735-758.
- [40] Stephen S, David G L, James J L. Vision-based global localization and mapping for mobile robots [J]. IEEE Transactions on Robotics, 2005, 21 (3): 364-375.
- [41] Mario E M, Paolo P, Enrico D B, et al. SIFTing through features with ViPR: application of visual pattern recognition to robotics and automation [J]. IEEE Robotics and Automation Magazine, 2006, 26 (3): 72-77.
- [42] Per-Erik F, David M, Kevin L, et al. Informed visual search: Combining attention and object recognition [C]//Proceedings of IEEE International Conference on Robotics and Automation. Pasadena, CA: IEEE, 2008: 935-942.
- [43] Meger D, Forssén P E, Lai K, et al. Curious george: an attentive semantic robot [J]. Robotics and Autonomous Systems Journal, 2008, 56 (6): 503-511.
- [44] David Meger. Curious george: an integrated visual search platform [C]//Proceedings of Canadian Conference on Computer and Robot Vision. Ottawa, Ontario: IEEE, 2010: 107-114.
- [45] Matthew B, David G L. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007: 59-73.
- [46] Matthew B, David G L. Recognizing panoramas [C]//Proceedings of International Conference on Computer Vision. Nice, France: IEEE, 2003: 1218-1225.
- [47] Matthew B, David G L. Invariant features from interest point groups [C]//Proceedings of British Machine Vision Conference. Cardiff, Wales: IEEE, 2002: 656-665.
- [48] Iryna G, David G L. What and where: 3D object recognition with accurate pose [J]. Toward Category-Level Object Recognition, 2006, 30 (12): 67-82.
- [49] Iryna G, David G L. Scene modeling, recognition and tracking with invariant image features [C]//International Symposium on Mixed and Augmented Reality. Arlington, USA: IEEE, 2004: 110-119.
- [50] Marius M, David G L. Fast approximate nearest neighbors with automatic algorithm configuration [C]//Proceedings of International Conference on Computer Vision Theory and Applications. Lisbon, Portugal: Springer, 2009: 1-10.
- [51] Marius M, Radu B R, Gary B, et al. REIN-a fast, robust, scalable recognition infrastructure [C]//Proceedings of International Conference on Robotics and Automation. Shanghai, China: IEEE, 2011: 2939-2946.
- [52] Marius M, David G L. Fast matching of binary features [C]//Proceedings of the Conference on Computer and Robot Vision. Toronto, Canada: IEEE, 2012: 404-410.
- [53] Sancho M, David G L. Local naïve Bayes nearest neighbor for image classification [C]//Proceedings of the Conference on Computer Vision and Pattern Recognition. Rhode Island, USA: IEEE, 2012: 3650-3656.
- [54] Luo J, Ma Y, Takikawa E, et al. Person-specific SIFT features for face recognition [C]//Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing. Acoustics: IEEE, 2007: 593-596.
- [55] Geng C, Jiang X. Face recognition using sift features [C]//Proceedings of 16th IEEE International Conference on Image Processing. Nanyang Technol. Univ. Singapore: IEEE, 2009: 7-10.
- [56] Krizaj J, Struc V, Pavesic N. Adaptation of SIFT features for face recognition under varying illumination [C]//Proceedings of the 33rd International Convention. Opatija, Croatia: IEEE, 2010: 691-694.
- [57] Majumdar A. Discriminative SIFT features for face recognition [C]//Proceedings of Canadian Conference on Digital Object Identifier. Canadian: IEEE, 2009: 27-30.
- [58] Kisku D R, Rattani A, Grosso E, et al. Face identification by SIFT-based complete graph topology [C]//Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies. Alghero, Italy: IEEE, 2007: 63-68.
- [59] Zhou Z M, Yu S Y, Zhang R, et al. An algorithm for face recognition based on SIFT descriptor [J]. Journal of image and Graphics, 2008, 13 (10): 1882-1885. [周志铭, 余松煌, 张瑞, 等. 一种基于 SIFT 算子的人脸识别方法 [J]. 中国图象图形学报, 2008, 13 (10): 1882-1885.]
- [60] Xu G Y, Wang Y. Single-sample face recognition using SIFT features [J]. Journal of Information Engineering University, 2008, 9 (2): 164-167. [许广毅, 王杨. 基于 SIFT 特征的单样本人脸识别研究 [J]. 信息工程大学学报, 2008, 9 (2): 164-167.]
- [61] Liu T G, Li X Y, Wng Y X, et al. Recognition algorithm for multi-biometric fusion based on SVM [J]. Nanotechnology and Precision Engineering, 2011, (1): 44-47. [刘铁根, 李秀艳, 王云新, 等. 基于 SVM 的多生物特征融合识别算法 [J]. 纳米技术与精密工程, 2011, (1): 44-47.]
- [62] Belcher C, Du Y. Region-based SIFT approach to iris recognition [J]. Optics and Lasers in Engineering, 2009, 47 (1): 139-147.
- [63] Ke J P, Xin C, Dong X Z, et al. 3D reconstruction based on SIFT and Harris feature points [C]//Proceedings of Paper presented at the Robotics and Biomimetics. Changsha, China: IEEE, 2009: 960-964.
- [64] Hu J F. Research on the 3D reconstruction method based on scale invariant feature transform feature matching algorithm [J]. Process Automation Instrumentation, 2011, (1): 30-32. [胡影峰. 基于 SIFT 特征匹配算子的三维重建方法研究 [J]. 自动化仪表, 2011, (1): 30-32.]
- [65] Psyllos A P, Anagnostopoulos C N, Kayafas E. Vehicle logo recognition using a SIFT-based enhanced matching scheme [J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 11 (2): 322-328.
- [66] Yang T, Zhang S L. Approach of vehicle plate extraction based on HSV color space and SIFT feature [J]. Application Research of Computers, 2011, 28 (10): 3937-3939. [杨涛, 张森林. 一种基于 HSV 颜色空间和 SIFT 特征的车牌提取算法 [J]. 计算机应用研究, 2011, 28 (10): 3937-3939.]