

综合视觉注意模型的显著性局部特征提取算法研究

杨族桥¹ 陈跃鹏² 张青¹

(黄冈师范学院数学与计算机科学学院 黄冈 438000)¹ (武汉理工大学自动化学院 武汉 430063)²

摘要 图像检索过程中往往会提取大量的局部特征,这将加大图像检索的计算量和复杂度,影响其应用。针对这一问题,提出了一种应用综合视觉注意模型的显著性分析提取局部特征的方法:在图像尺度空间中提取关键点,利用模糊增长技术查找原始图像的显著性区域,计算其综合视觉显著性权值并分类,提取 SIFT 描述因子,保留最突出的局部特征以提高检索性能。相比于传统的局部特征提取算法,本方法在图像检索精度和检索速度方面都具有明显优势。

关键词 综合视觉显著性,局部特征,局部特征选择,基于内容的图像检索

中图分类号 TP391 文献标识码 A

Salient Local Feature Extraction Algorithm Based on Integrated Visual Attention Model

YANG Zu-qiao¹ CEHN Yue-peng² ZHANG Qing¹

(College of Mathematics and Computer Science, Huanggang Normal University, Huanggang 438000, China)¹

(Automation College, Wuhan University of Technology, Wuhan 430063, China)²

Abstract Local features are widely used for content-based image retrieval recently. During image retrieval, a lot of local features are extracted, which increases the amount of calculation and complexity of image retrieval, and as a result, affecting the practical applications. With an eye towards this problem, a novel method based on integrated visual attention model was proposed to extract salient local features. Using this method, first, the key points in an image scale-space are extracted, and the salient area of the original image is found using fuzzy growth technology, then the integrated visual saliency is calculated and classified, and SIFT factors are extracted and ranked according to their integrated visual saliency, and at last, only the most distinctive features are kept to enhance the retrieval performance. The experimental results demonstrate that compared to traditional local feature extraction algorithms, this salient local feature extraction algorithm based on integrated visual attention model provides significant benefits both in retrieval accuracy and speed.

Keywords Integrated visual saliency, Local features, Local feature selection, Content-based image retrieval

1 引言

人类具有极强的视觉感知能力,能快速从复杂背景图像中检测并识别目标,且通过少量实例学习能辨识并记忆一个新的物体类别。但目前的计算机仍不能在图像中准确可靠地辨识日常物体和场景。视觉注意的显著性计算可以为选择少数的几个视觉显著区域和减少冗余提供快速的计算机制^[1]。以视觉注意为代表的图像显著性区域检索技术成为提高海量数据筛选实时性和分析准确性的重要技术途径之一。

尺度不变特征变换(SIFT)已证明是最具鲁棒性的局部特征提取算法,在空间域和频率域都可以很好地定位,也能减小遮挡、噪音等影响。SIFT 通过探测尺度空间极值、定位关键点、任务定向、建立关键点描述因子等主要步骤生成图像特征集,并进行尺度变换^[2]。但该方法会产生大量特征,并密集地分布于整个图像的尺度和位置空间。一幅 500 * 500 像素的典型图片会产生约 2000 个稳定的特征。特征的数量对目

标识别尤为重要,要具备探测杂乱背景下的小目标的能力,要求每个目标至少有 3 个特征被正确匹配才是可靠的识别。PCA-SIFT^[3]将主成分分析(PCA)方法应用于初始特征提取,并将每个斑块特征维数从 128 减少到 36 来提高匹配的速度。然而,计算量依旧太大,也难以应用到实时图像检索系统中。因此,本文集中研究这个问题,并提出一种基于视觉注意分析来选择显著性 SIFT 特征的新方法:在提取图像所有 SIFT 特征点的基础上,计算这些特征点的综合视觉显著性值,综合考虑图像局部和全局分布信息的同时对特征点进行排序,仅保留最独特的特征点,以此作为图像检索特征。这样就可以减少计算量,提高检索速度。

2 相关研究

人们对图像预处理阶段就能很容易地判断图像中的显著性区域,同时解释复杂的场景信息^[4,5]。注意是认知与感知的有效连接,人类视觉系统能力快速且可靠。这种智能行

到稿日期:2012-10-28 返修日期:2013-04-11 本文受湖北省科技厅自然科学基金项目(2012FFC036,2011CDC028),湖北省教育厅重点项目(D20102901,D20122701)资助。

杨族桥(1972-),男,博士,副教授,主要研究方向为空间数据挖掘与数字图像处理,E-mail:jsjyzq@hgnu.edu.cn;陈跃鹏(1971-),男,博士,教授,主要研究方向为信息自动化处理;张青(1965-),女,博士,教授,主要研究方向为人工智能、空间信息处理等。

为的数值模拟计算对计算机视觉来说非常有用,因为我们可以通过显著性区域来优先分配图像分析与合成所需的计算资源。视觉显著性源于视觉的独特性、不可预测性、稀有性以及奇异性,并通过颜色、梯度、边缘和边界等图像属性来表示,显著性图像区域的提取在计算机视觉许多实际应用中广泛使用,包括目标识别、图像自适应压缩、内容认知的图像大小调整以及图像检索^[6,7]。由 Itti 等^[2]所定义的注意理论提出,视觉注意机制分为两个阶段:快速的、下意识的、自下而上的、数据驱动的显著性提取和慢速的、任务依赖的、自上而下的、目标驱动的显著性提取。由于后者具有良好的可扩展性,故这里集中研究自下而上数据驱动的视觉显著性检测方法。

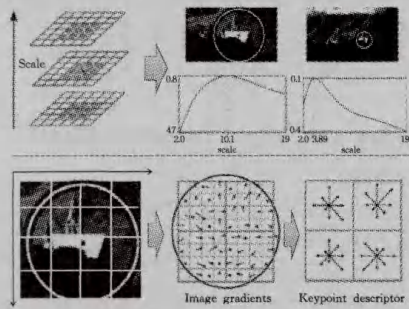


图1 原始 SIFT 特征的检测与描述^[3]

在许多应用中,视觉显著性分析方法提出了一种可供选择的方法来理解图像语义,如自适应内容分发和基于局部的图像检索。人们研究了大量的基于局部对比度的可计算注意模型,其利用低级处理方法来计算图像区域同周围的对比度,使用亮度、颜色和边缘等图像特征属性^[7]。通过跨多尺度图像特征的中心-周围的差异性计算,以特征点周围像素的梯度大小和方向信息来表示特征描述子(见图1),进而提出一种基于显著性注意模型的场景分析方法^[10]。研究表明,基于局部对比度的方法具有鲁棒性、可重复性、局部几何不变性等性质。在 Itti 的工作中,视觉输入首先被分解成一组特征图(42个低级特征图分别从不同的颜色、亮度和方向通道通过8空间尺度来提取);然后,使用正则化算子及“显著性图”通过自下而上的方式生成特征图的线性组合。图2中显著性图与输入图像在拓扑结构上是对应的,表示其邻域的每个像素的局部“显著性”。与显著性图的最大亮度像素对应的是原始图像的最显著位置。所以利用显著局部特征检索目标,同时还能解决目标的定位问题,估计目标在图像中出现的大致位置。此外,区别图像不同区域的显著性,可以非均匀地分配计算资源,提高视觉算法的效率。Ming-ming Cheng 等提出提取高分辨率的全局显著性图像的分析方法,其采用基于直方图对比度的方法检测显著性,并用色彩空间的平滑操作来控制量化的缺陷^[9]。

虽然基于局部对比度的显著性分析方法易于理解和计算,但这种方法对噪声数据非常敏感,而且在边缘附近往往会产生较高的显著性值,而不是均匀地突出显著性目标。Hou 等提出一种分析输入图像基于频率方法的日志-频谱方法^[8],但该方法尽在固定尺度上进行,忽略了图像尺度的影响。

这种视觉显著性分析系统中的基本原理是减少对频繁发生目标的响应,同时又对偏离标准的目标保持敏感。此性质

也称为 $1/f$ 法则,如图3所示。它表明自然图像集合的平均傅立叶频谱的振幅 $A(f)$ 服从某种分布:

$$E\{A(f)\} \propto 1/f \quad (1)$$

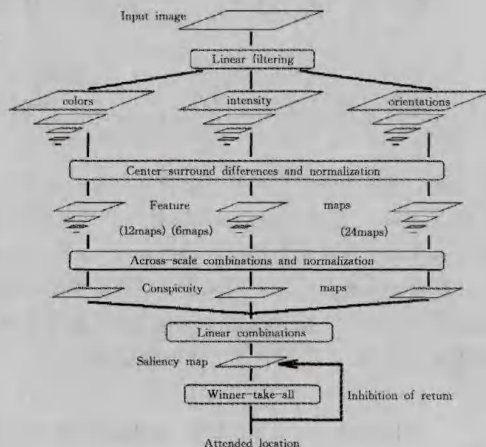


图2 基于显著性分析的局部对比度的体系结构

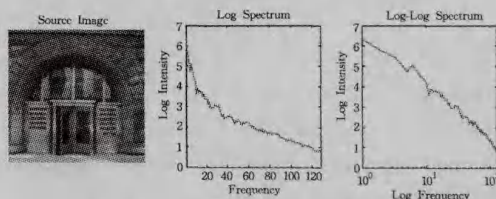


图3 日志-频谱方法示例

综上所述,基于局部对比度或基于全局对比度的方法仅适用于特定的图像内容。此外,这些方法往往忽视了图像各部分之间的空间关系,而这些空间关系在显著性检测的可靠性和一致性评价中发挥着重要作用。所以本文提出一种将局部对比度和全局对比度优势综合互补的新方法,它可有效地计算每个局部目标的显著性值,并提取最显著特征。

3 基于综合视觉显著性的局部特征选择

SIFT 算法的基本思路:先构建图像的尺度空间表示,再在尺度空间中搜索图像的极值点,然后由极值点构建特征点的描述因子^[2]。为了有效地选择显著的和独特的局部特征,本文给出一种结合基于局部对比度和全局对比度优点的视觉显著性分析方法。本方法的主要步骤如下:

- (1) 构建图像的尺度空间;
- (2) 查找图像尺度空间的极值点,并提取图像的关键点;
- (3) 利用模糊增长技术查找原始图像中的所有显著性区域;
- (4) 计算显著性区域的权值并排序;
- (5) 提取选取的关键点的 SIFT 描述符。

本方法为过滤局部特征而重点研究原 SIFT 算法的第3阶段。对每个图像,在 SIFT 关键点提取之后,利用基于局部对比度和基于全局对比度的视觉显著性分析技术^[1]同时分别生成显著性图 A 和 B。最后的显著性图是与输入图像相应的拓扑结构,并且每个像素的显著性程度是 A 和 B 的总和。规范化后像素局部“显著性值”会调整在 $[0 \sim 255]$ 范围之内,与这个显著性图中最大亮度像素相对应的是原始图像中最显

著的位置。然后,利用模糊增长技术查找原始图像中所有显著性区域。考虑到计算量的大小以及计算的复杂度,规定每个图像的显著性区域数目不超过3。图4给出了一些示例。

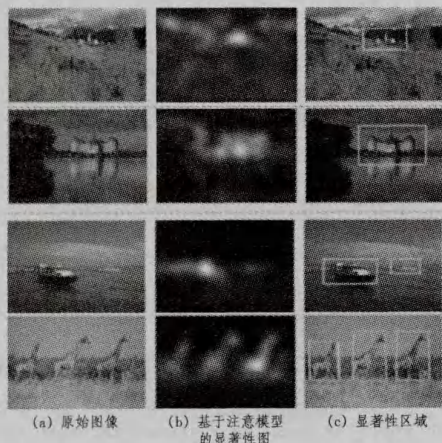


图4 基于注意模型和模糊增长的显著性区域检测

通常情况下,一个SR由原始图像中的像素点集表示,显著性图中的显著性区域(SR)可为任意图形。在本文中,为简单起见,使用矩形来表示,其定义为: $\{Center_x, Center_y, Width, Height\}$ 。其中 $(Center_x, Center_y)$ 表示的矩形的中心位置, $(Width, Height)$ 表示的此矩形SR的宽度和高度值。假定任何矩形都不会相互重叠,在信息损失不大的条件下将会明显提升计算速度。

根据SR的定义,每个SIFT的关键点附加了显著性权值 KP_{weight} ,并通过如下公式计算得到:

$$KP_{weight} = KPR_{dis} * SR_{weight} \quad (2)$$

$$KPR_{dis} = 1 - \frac{2 * \sqrt{(x - Center_x)^2 + (y - Center_y)^2}}{\sqrt{Width^2 + Height^2}} \quad (3)$$

$$SR_{weight} = R_{area} * R_{pos} \quad (4)$$

KPR_{dis} 表明此关键点和包含它的区域的中心之间的距离成反比例。 (x, y) 表示关键点的位置。

如果关键点在其相应区域的中心,则 KPR_{dis} 小于等于1;若关键点在边缘上,或在检测到的显著性区域之外,则 KPR_{dis} 为0。 SR_{weight} 表示显著性区域的显著性权值。移位检测到的区域其重要性通常由该区域面积权值 R_{area} 和位置权值 R_{pos} 反映出来。若显著性区域面积太小,无法提供任何有用的信息,则可以忽略不计。因此,只有面积大于总图像5%的区域才被划分出来,并且最多保留3个作为SR区域。假设图像包含 n 个SR区域(n 介于0和3之间),且 $area_i$ 表示第 i 个SR的面积。当前SR的位置权值是通过式(5)来计算的:

$$R_{area} = \frac{area_{current}}{\sum_{i=1}^n area_i} \quad (5)$$

因为人们往往更注重图形中心附近的区域,图像中心归一化高斯模板用来给位置权值 R_{pos} 赋值,如图5所示。如上所述,显著性权值 KP_{weight} 的每个局部特征(本文中侧重SIFT关键点)就可以计算出来。我们对所有关键点按照 KP_{weight} 值排序,并仅保留前 N 个关键点来提取SIFT描述符。

在实际应用中, N 的值是由保持检索精度和检索速度的适当平衡来确定的。通过这种方式,应用注意模型就可以完成SIFT关键点的选择。

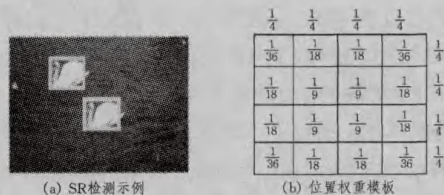


图5 位置权重的高斯模板

4 实验分析

为了验证本方法的有效性,我们利用大数据集对图像检索性能进行了评估。测试数据从ALOI图像数据集和加州理工图库集中下载。根据不同的对象,数据集共分为10类,每个类有10多个相近图像。在所有图像数据集已提取约40000标准SIFT关键点。

4.1 评价矩阵

为匹配图像,我们采用一个类似于文献[11]提出的标准:从图像中获得正确的匹配数。给定一个查询图像 i ,数据库的每个图像 j 都可通过计算来获取其匹配度 $m_{i,j}$ 。图像对 i 和 j 中所有关键点都要检查。如果特征间的欧氏距离低于给定的阈值,就产生一个匹配。 $m_{i,j}$ 是两个图像间的匹配数,且所有图像都将按照匹配数排序。

我们采用平均检索精度来描述本数据集中图像检索的排序序列。前 n 幅图像的检索精度利用式(6)来计算。这里 q 是查询图像, p_i 表示每个图像的排序结果,且 n 为10。

$$AP_q = \frac{1}{n} \left(\sum_{i=1}^n \left(\frac{1}{i} \sum_{j=1}^i \psi(p_j, q) \right) \right) \quad (6)$$

$$\psi(p_j, q) = \begin{cases} 1, & p_j \text{ 匹配 } q \\ 0, & p_j \text{ 不匹配 } q \end{cases}$$

4.2 结果与分析

实验对基于注意模型的SIFT关键点的过滤算法(AF-SIFT)与标准SIFT及PCA-SIFT的结果进行了比较。标准SIFT描述符是通过生成平滑的方向直方图来描述关键点周围的斑块,采用 $4 * 4$ 列的直方图、8方向束来计算,因此标准SIFT的维度是128。实验结果如表1所列。

表1 实验结果
(独立运行10次以上的平均时间)

	每一幅图像表示时间(s)	数据集匹配时间(s)	
SIFT	1.41	SIFT	184.9
PCA-SIFT	1.55	PCA-SIFT	67.4
AF-SIFT1	1.69	AF-SIFT1	57.6
AF-SIFT2	1.71	AF-SIFT2	22.3

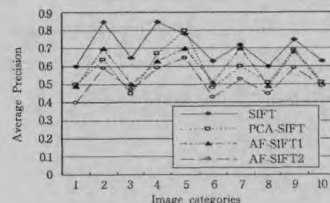


图6 不同算法的比较

每个关键点的 PCA-SIFT 描述符维度是 36。标准方式 AF-SIFT1 使用 128 维描述符,而 AF-SIFT2 使用 8 方向束和 2×2 数组,其维度是 32。图 6 比较这 4 种算法的检索准确性。我们可以看到 PCA SIFT 和 AF-SIFT1 有相近的准确性,而 AF-SIFT2 比 SIFT 算法快很多。说明本文提供了一种比标准 SIFT 算法更有效的算法且其非常适合图像检索。

结束语 本文研究了一种基于综合视觉注意分析的局部特征选择的新方法,即在一幅图像中提取局部特征,按其综合显著性权值划分等级,并且仅保留最独特的局部特征。使用这种方法,在线检索速度将明显增加且精度损失很小。本文 AF-SIFT 算法对 SIFT 算法进行了改进,在图像检索的正确率和查全率方面具有很好的性能,且简单高效,适合大规模图像检索。

在以后的工作中,还要进一步研究融合多种局部特征的表达,综合多种局部特征检测算子和描述因子,建立多通道的视觉物体表示机制,以发挥不同局部特征的互补优势;还要进一步考虑图像的语义概念信息,避免造成底层局部特征和高层语义概念的脱节;同时还要进一步研究局部特征的同现性规律并建立更加有效的视觉显著性表达模型,以便进行更高效的目标检测、可靠的遥感图像分类、鲁棒的图像景观分析,提高图像检索性能。

(上接第 251 页)

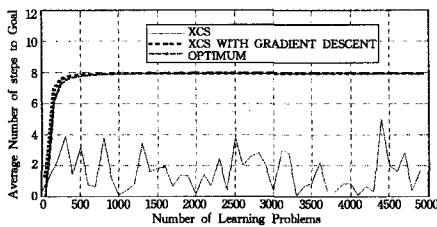


图 7 “6”字迷宫中机器人强化学习($N=2500$)

结束语 本文提出了用 XCSG 来解决多机器人的强化学习问题。XCS 建立低维的状态-动作对逼近函数 Q 和状态-动作对到环境回报值的映射关系。梯度下降技术利用在线知识建立稳定的逼近函数 Q ,使状态-动作对到环境回报值的映射 Q 表格一直保持在低维状态。算法的分析和仿真表明,XCS 和梯度下降法相结合用于多机器人的强化学习是高效、稳定收敛的。

参考文献

- [1] 邵杰,杨静宇.基于多 LCS 和人工势场法的机器人行为控制[J].计算机学报,2011,38(1):264-267
- [2] 朱美强,程玉虎,李明,等.一类基于谱方法的强化学习混合迁移算法[J].自动化学报,2012,38(11):1765-1776
- [3] Wiering M. Multi-agent reinforcement learning for traffic light control[C]//Proc. 17th Int. Conf. Mach. Learn. (ICML-00). Stanford Univ. Stanford, CA, 2009:1151-1158

- [1] Ma Y F, Zhang H J. Contrast-based Image Attention Analysis by Using Fuzzy Growing[C]//ACM MM. 2003:374-381
- [2] Itti L, et al. Visual Attention and Target Detection in Cluttered Natural Scenes[J]. Optical Engineering, 2001, 40(9):1784-1793
- [3] Ke Y, et al. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors[C]//CVPR. 2004:506-513
- [4] Lowe D. Distinctive Image Features from Scale-Invariant Keypoints[J]. Int'l J. Computer Vision, 2004, 2(60):91-110
- [5] Tuytelaars T. Matching Widely Separated Views Based on Affine Invariant Regions[J]. Int'l J. Computer Vision, 2004, 1(59):61-85
- [6] Mikolajczyk K, Schmid C. Performance Evaluation of Local Descriptors[J]. IEEE Trans. of PAMI, 2005, 27(10):1615-1630
- [7] Hakim A. CSIFT: A SIFT Descriptor with Color Invariant Characteristics[C]//CVPR. 2006:1978-1983
- [8] Hou X. Saliency detection: A spectral residual approach[C]//CVPR. 2007:1-8
- [9] Cheng M M, et al. Global Contrast based Salient Region Detection [C]//CVPR. 2011:456-463
- [10] 陈硕,吴成东,陈东岳.基于视觉显著性特征的快速场景配准方法[J].中国图象图形学报,2011(7):1241-1247
- [11] 程仁贵,刘书焱.基于边缘检测的影像多线自动测量算法[J].重庆理工大学学报:自然科学版,2013,27(2):89-92

- [4] Dixon P W, Corne D W, Oates M J. A preliminary investigation of modified XCS as a generic data mining tool[C]//Lanzi P L, Stolzmann W, Wilson S W, eds. LNAI, Advances in Learning Classifier Systems. vol. 2321, Berlin, Germany: Springer-Verlag, 2002:133-150
- [5] 欧世峰,高颖,赵晓晖.基于随机梯度的变向量因子自适应白化算法[J].自动化学报,2012,38(8):1370-1374
- [6] Butz M V, Goldberg D E, Lanzi P L. Gradient descent methods in learning classifier systems: Improving XCS performance in multistep problems[J]. IEEE Trans. Evol. Comput., 2005, 9(5):452-473
- [7] Bernad E, o-Mansilla, Garrell J. Accuracy-based Learning Classifier Systems: Models, analysis and applications to classification tasks[J]. Evolutionary Computation, 2003, 11(3):209-238
- [8] Hung K-T, Liu J-S, Chang Y-Z. Smooth path planning for a mobile robot by evolutionary multiobjective optimization [C]//IEEE Int. Symposium on Computational Intelligence in Robotics and Automation. Jacksonville, Florida, June 2007
- [9] Butz M V, Lanzi P L, Wilson S W. Function approximation with XCS: Hyperellipsoidal conditions, recursive least squares, and compaction[J]. IEEE Trans. Evol. Comput., 2008, 12(3):355-376
- [10] Bagnall A J, Cawley G C. Learning classifier systems for data mining: A comparison of XCS with other classifiers for the Forest Cover dataset[C]//Proc. IEEE/INNS Int. Joint Conf. Artificial Neural Netw. vol. 3, Portland, OR, 2003:1802-1807