

基于均匀区域分割的图像检索研究

肖 潇¹ 王宪保¹ 庄德文³ 王守觉^{1,2}

(浙江工业大学信息学院 杭州 310014)¹ (中国科学院半导体研究所 北京 100083)²

(浙江工业大学应用物理系 杭州 310014)³

摘要 为减少图像分割准确性对基于内容的图像检索效率的影响,提出了一种基于均匀区域分割的图像检索算法。首先对图像进行均匀区域划分,提取其区域直方图颜色特征和 Gabor 小波纹理特征,再利用与所提取的图像特征相适应的相似性度量实现有效检索。实验结果表明,与 SIMPLIcity 系统相比,该算法平均检索性能提高了 3.6%,具有良好的平均查找率。

关键词 区域分割,基于内容的图像检索,Gabor 滤波,颜色,纹理

中图法分类号 TP391 文献标识码 A

Image Retrieval Based on Uniform Region Segmentation

XIAO Xiao¹ WANG Xian-bao¹ ZHUANG De-wen³ WANG Shou-jue^{1,2}

(Information College, Zhejiang University of Technology, Hangzhou 310014, China)¹

(Institute of Semiconductors, Chinese Academy of Science, Beijing 100083, China)²

(Department of Applied Physics, Zhejiang University of Technology, Hangzhou 310014, China)³

Abstract In order to reduce the effect of the accurate image segmentation on the content-based image retrieval, a new image retrieval algorithm based on uniform region segmentation was proposed. At first, the image was segmented using uniform region, and then the histogram color feature and Gabor wavelet texture feature were extracted. Finally, a reasonable similarity measure was designed to improve the retrieval efficiency. The experiments results show that the algorithm in this paper is more accurate than SIMPLIcity system. The average retrieval performance is increased by 3.6%. Better average precision ratio is obtained.

Keywords Region segmentation, Content-based image retrieval, Gabor filter, Color, Texture

1 引言

基于内容的图像检索,从根本上来说,可定义为根据图像的视觉内容对图像档案进行组织的一门技术。随着计算机计算能力的迅猛增长,以及数字化图像获取设备的普及,如何有效地检索大型数字图像库,是具有很强挑战性的问题^[1]。

基于内容的图像检索的两个关键问题可表示为:(1)如何从数学上描述一幅图像;(2)如何基于从图像中所抽取的抽象描述来评价一对图像间的相似性。用于图像检索图像数学描述,称为签名。近年来,基于区域视觉签名的构造受到越来越多的关注^[2,3]。为得到基于区域的签名,关键一步是对图像的分割,可靠的图像分割也是得到形状描述的关键。

图像分割技术也是视觉理解的关键,一直受到研究者的重视。最简单同时也是广泛采用的有基于 k -平均聚类的方法,但其显然的缺陷是不够精确。近来所提出的基于图论的规范化切割图像分割方法受到普遍的关注^[4-6]。毫无疑问,图像分割越准确,就越能将目标对象从其环境中分离,以便做更

准确的对象匹配。但由于基于内容图像检索中所面对的图像类型众多,有些图像中并没有一个清晰的目标对象,因此完全依赖图像准确分割的图像检索系统并不实用。所以,减少对图像准确分割的依赖性,更是目前人们努力寻求的图像检索系统所追求的目标^[7]。

2 均匀区域视觉特征提取

图像的视觉特征主要包括色彩、纹理和形状。对视觉特征的自动提取和表示是基于内容图像检索的基础,也是早期图像研究的主要问题^[8]。

2.1 颜色空间的选择

由于准确分割图像困难,又要结合区域匹配的优点,因此我们将图像分割为均匀的图像块,每块分别计算其视觉特征。由于单个视觉特征并不能很好地反映图像特点,因此往往由多个特征综合来构造。常用的 RGB 颜色空间与人的感知特性有差异,所以我们采用与人的感知特性较为接近的 HSV 颜色空间。H, S, V 分别对应于彩色信号的色度 h 、饱和

到稿日期:2011-07-08 返修日期:2011-11-30 本文受国家自然科学基金(60871093),浙江工业大学科研基金(2011XZ004)资助。

肖 潇(1980—),女,博士生,主要研究方向为计算机视觉、模式识别、神经网络,E-mail: xiaoxiao800412@163.com;王宪保 男,博士,讲师,主要研究方向为模式识别、神经网络、图像处理及其在智能监控中的应用等;庄德文 男,博士,讲师,主要研究方向为模式识别、人工神经网络、计算机视觉;王守觉 男,中国科学院院士,主要研究方向为模式识别。

度 s 和量度 v 。

2.2 颜色特征的提取

颜色直方图是常用的视觉特征之一,具有计算简单、高效的特点。考虑计算效率和检索性能的平衡,对图像做 4×4 子块均匀分割(见图 1),每块计算其颜色直方图。在做颜色直方图时,区间(bin)的大小影响着计算效率及检索准确性。根据实验分析,对于 HSV 颜色空间的 (hsv) ,采用 $(32\ 16\ 16)$ 区间能较好地平衡计算效率及检索准确性。为减少对区间中心位置的敏感性,对所得到的直方图做平滑处理,即对邻近区间的直方图值计算平均,得到 $(31\ 15\ 15)$ 即共 61 维特征向量,每个子块的颜色特征向量记作 $fcolor_i = \{c_{1i}, c_{2i}, c_{3i}, \dots, c_{61i}\} (i = 1, 2, \dots, 16)$ 。

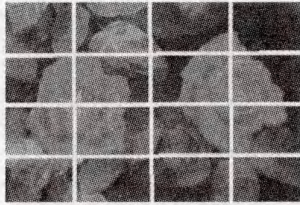


图 1 4×4 均匀区域分割

2.3 基于 Gabor 小波的纹理特征提取

纹理特征是一种不依赖于颜色或量度的、反映图像中颗粒度和表面的重复性模式,它包含了物体表面结构的重要信息及周围环境的联系。典型的有 Tamura 纹理特征^[9]。

小波变换是一种常见的纹理分析工具。Gabor 小波是由 Gabor 函数通过尺度伸缩和旋转生成的一组复函数系,Gabor 小波变换具有优良的空间局部性和方向选择性。2D Gabor 核函数定义为:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2} [e^{i k_{u,v}^T z} - e^{-\sigma^2 / 2}] \quad (1)$$

式中, $k_{u,v} = k_v e^{i\varphi_u}$, $z = (x, y)$, $k_v = k_{\max} / f^v$, $\varphi_u = \pi u / 8$, u 和 v 分别代表滤波器的方向和尺度, k_{\max} 是最大频率, f 为频率域中的采样步长。通常,取 5 个尺度 $v \in \{0, \dots, 4\}$ 和 8 个方向 $u \in \{0, \dots, 7\}$ 的 Gabor 小波。图 2 分别是 Gabor 核在 5 个尺度上的幅值和 5 尺度、8 方向的 Gabor 核的实部。

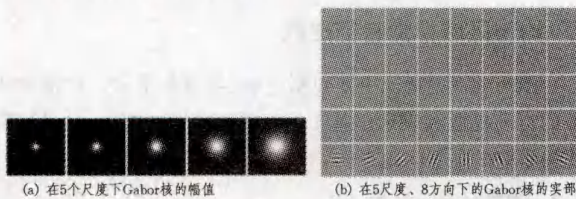


图 2

图像 $I(z) = I(x, y)$ 的 Gabor 特征表达是图像与一系列由式(1)定义的 Gabor 核进行卷积:

$$O_{u,v}(z) = I(z) * \psi_{u,v}(z) \quad (2)$$

式中, $*$ 表示卷积操作, $O_{u,v}(z)$ 是图像和方向 u 、尺度 v 的 Gabor 核进行卷积的结果。当使用 5 个尺度、8 个方向的 Gabor 小波时,可以得到 40 个卷积结果 $O_{u,v}(z)$ 。因此,总的维数增加了 40 倍,往往先对 $O_{u,v}(z)$ 进行 ρ 下采样降维处理,并进行零均值、单位方差的归一化处理,得到 $O_{u,v}^{\rho}(z)$,最后的特征表达是将不同方向、不同尺度的归一化图像 $O_{u,v}^{\rho}(z)$ 按式(3)进

行连接,得到最终的特征表达 χ^{ρ} :

$$\chi^{\rho} = (O_{0,0}^{\rho}, O_{0,1}^{\rho}, \dots, O_{4,7}^{\rho})^T \quad (3)$$

我们采用能较好反映局部与整体纹理特性的 Gabor 滤波^[10],利用预计算滤波模板的方法,提高计算效率。Gabor 滤波采用 5 个尺度、8 个方向的滤波模板,仅对 HSV 颜色空间中亮度做滤波,对所得到的每个滤波子图,取其平均 E 和方差 σ 作为最后的纹理特征,记为 $ftexture_i = \{E_{1i}, E_{2i}, \dots, E_{40i}, \sigma_{1i}, \sigma_{2i}, \dots, \sigma_{40i}\} (i = 1, 2, \dots, 16)$ 。合并颜色特征得到块特征,最后每个子块的特征为 141 维 $(61 + 40 + 40)$,记为 $F_i = \{fcolor_i, ftexture_i\} (i = 1, 2, \dots, 16)$ 。每张图像有 16 个子块,得到的特征是 16×141 维,每张图像的特征向量为 $F = \{F_1, F_2, \dots, F_{16}\}$ 。

2.4 特征向量加权

根据大量的图像检索实验表明,图像中心的视觉关注度最高。视觉的关注度越高,在相似性度量计算中的地位越高,所以我们将得到的图像的特征向量以加权的形势表示出来,

$$\text{经过实验得到的权重矩阵 } W = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & 2 \\ 2 & 4 & 4 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix} \text{ 其效果是最理}$$

想的。

所以最终得到的特征向量为:

$$F' = \begin{bmatrix} F_1 & 2F_2 & 2F_3 & F_4 \\ 2F_5 & 4F_6 & 4F_7 & 2F_8 \\ 2F_9 & 4F_{10} & 4F_{11} & 2F_{12} \\ F_{13} & 2F_{14} & 2F_{15} & F_{16} \end{bmatrix}$$

3 相似性度量

如何利用所得到的图像签名做准确的图像检索,是基于内容图像检索中研究的第二个关键问题。由于存在语义差距,即所提取的底层视觉特征与高层语义内容之间的一致性,因而基于底层提取的是视觉签名,给图像相似性度量设计提出了很大的挑战。根据视觉签名的形式不同,相应的相似性度量,即距离的数学形势也不同,有欧氏距离以及流行嵌入的测地线距离^[11]。对于基于区域签名所得到的矢量集间的距离,有矢量距离加权和形式^[1]、Hausdorff 距离^[12]、集成区域匹配距离^[13]等。对于区域矢量综合所得到的图像签名,通过高斯混合模型的拟合,可有效地得到密度估计,并利用 Kullback-Leibler 散度,即相对熵,来度量两个分布之间的差异^[14]。



图 3 按最相似块顺序排列均匀区域匹配

我们使用的相似性度量类似于“最相似优先原则”^[1,15]：

$$d(I_1, I_2) = \sum_{i,j} \text{mind}(F_{1i}, F_{2j}) \quad (4)$$

式中, F_{1i} 、 F_{2j} 分别为图像 I_1 和 I_2 的子块图像特征, $d(F_{1i}, F_{2j})$ 为其欧氏距离。依次找出最相似图像块, 直到图像块匹配完毕, 示例见图 3。

4 实验结果分析

4.1 实验数据库

基于内容的图像检索并没有一个标准的测试数据库, 我们采用常用的 COREL 数据库中的一个子集^[1], 它有 10 类图像, 每类 100 张图像, 图像尺寸为 256×384 或 384×256 , 具体类别名称见表 1。检索示例见图 4。

表 1 COREL 子集图像类别

ID	Category Name
1	Africa people and village
2	Beach
3	Buildings
4	Buses
5	Dinosaurs
6	Elephants
7	Flowers
8	Horses
9	Mountains and glaciers
10	Food

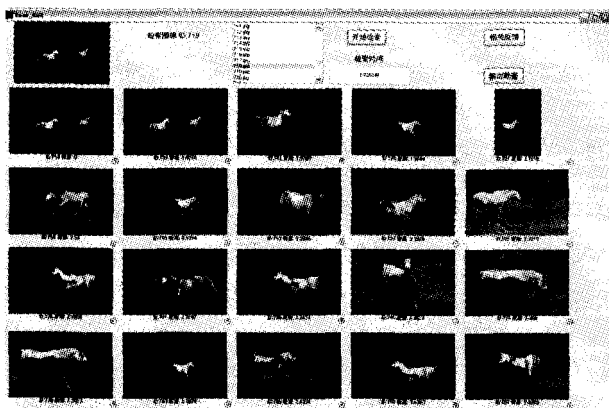


图 4 均匀区域方法图像检索示例图

4.2 实验结果比较

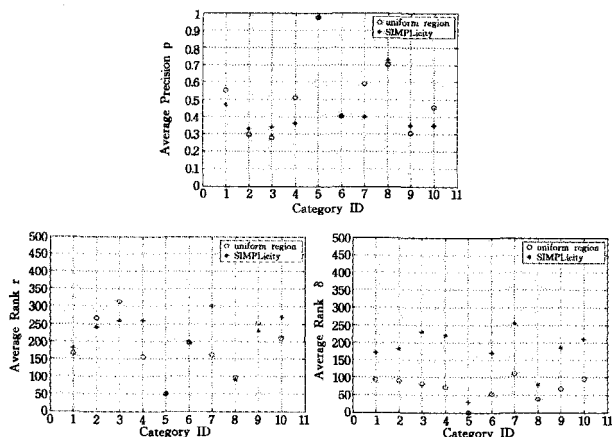


图 5 均匀区域方法图像检索性能比较

基于内容图像检索的评估标准仍是有待研究的问题, 因为与用户的意愿有关, 一个实用的图像检索系统离不开人的

主观评价。为了能较为客观地反映图像检索系统的性能, 我们采用文献[1]对 COREL 子集的评估方法, 并与之加以比较。即用平均准确率 p 、所匹配图像平均排序 r 及所匹配图像排序的标准方差 σ , 具体性能指标见图 5。SIMPLcity 检索的平均准确率为 0.471, 均匀区域检索平均准确率为 0.507。

由图 5 可知, 对于平均准确率, 类别 1、4、7 和 10 都有不同程度的提高, 最高约有 20% 的提高; 而类别 2、3、8 和 9 虽相比 SIMPLcity 略有降低, 但最低仅相差约 6%; 类别 5 和 6 与 SIMPLcity 性能相似, 总体平均检索准确率为 0.507, 比 SIMPLcity 要高 3.6%。相应的平均排序性能有类似的结果, 而标准方差均略低于 SIMPLcity。

结束语 本文提出了基于均匀区域视觉特征图像检索方法, 该方法对所提取的直方图区域色彩特征和 Gabor 小波纹理特征采用相应的相似性度量, 其图像检索性能优于 SIMPLcity 系统。算法避免了图像准确分割的困难。但同时看到, 对一些类别, 其检索性能略低。另外, 尽管采用了预计算 Gabor 滤波模板的方法来提高滤波计算效率, 但其仍是特征提取中的计算瓶颈。如何更好地提取视觉特征、提高检索效率, 以及如何利用反馈技术提高检索系统的性能, 使其能够应用于大规模数据库的图像检索^[16-19], 仍是我们以后工作中需要进一步解决的问题。

参考文献

- [1] Wang J Z, Li J, Wiederhold G. SIMPLcity: semantics-sensitive integrated matching for picture libraries[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2001, 23(9): 947-963
- [2] Chen Y, Wang J Z. A region-based fuzzy feature matching approach to content-based image retrieval[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2002, 24(9): 1252-1267
- [3] Chen Y, Wang J Z. Image categorization by learning and reasoning with regions[J]. Journal of Machine Learning Research, 2004, 5: 913-939
- [4] Shi J, Malik J. Normalized cuts and image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905
- [5] Yu S X, Shi J. Segmentation given partial grouping constraints[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2004, 26(2): 173-183
- [6] Carballido-Gamio J, Belongie S, Majumdar S. Normalized cuts in 3-D for spinal MRI segmentation[J]. IEEE Trans. Medical Imag., 2004, 23(1): 36-44
- [7] Carson C, Belongie S, Greenspan H, et al. Blobworld: Image segmentation using expectation-maximization and its application to image querying[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2002, 24(8): 1026-1038
- [8] Smeulders A W, Worring M, Santini S, et al. Content-based image retrieval at the end of the years[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2000, 22(12): 1349-1380
- [9] Tamura H, Mori S, Yamawaki T. Texture features corresponding to visual perception[J]. IEEE Transaction on Systems, Man, and Cybernetics-Part B: Cybernetics, 1978, 8(6): 460-473
- [10] Jain A, Farrokhnia F. Unsupervised texture segmentation using gabor filters[C]// Proceedings of the International Conference on Systems, Man and Cybernetics, 1990

(下转第 292 页)

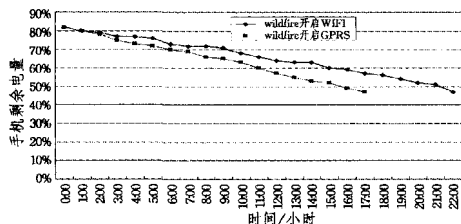


图9 wildfire 开启 WIFI 和开启 GPRS 功耗对比

由图 8、图 9 计算得出,在相同流量模式下,Lephone 在 GPRS 模式下电池使用时间相对 WIFI 模式减少 37.5%,wildfire 在 GPRS 模式下电池使用时间相对 WIFI 模式减少 22.7%。

综合以上各图,可以得出如下结论:

(1)小流量模式下,用户在使用 WIFI 策略或 GPRS 策略接入网络模式比未接入网络模式下电池使用时间减少约 50%。这说明不管采用 WIFI 策略还是 GPRS 策略接入网络,电池电量下降均很快。

(2)大流量模式下,使用 WIFI 策略接入网络模式与未接入网络模式相比,电池使用时间减少约 85%。这说明在大流量模式下,电池电量下降得更快。

(3)对比相同流量模式下分别使用 WIFI 策略和 GPRS 策略连接网络,同等条件下 WIFI 模式比 GPRS 模式省电。

结束语 随着计算机技术的迅速发展和移动互联网的普及,智能机移动网络模块(GPRS/WIFI)的功耗在系统总功耗中占据了越来越重要的位置。本文通过实验分析了在使用 WIFI 和 GPRS 两种策略接入网络时手机系统功耗的变化情况,得出了移动网络模块的功耗在系统总功耗中占据了十分重要的位置这一结论。因此,分析手机移动网络模块的功耗具有十分重要的意义,能为将来嵌入式移动设备的低功耗研究提供一定的指导。

参 考 文 献

[1] Carroll A, Heiser G. An analysis of power consumption in a smartphone[C]// Proceedings of the 2010 USENIX Annual Technical Conference, Boston, MA, USA, June 2010; 1-14
 [2] Bhowmik A K, Brennan R J. System-level Display Power Reduc-

tion Technologies for Portable Computing and Communications Devices[C]// Proc. IEEE Int. Conf. Portable Information Devices. Orlando, Florida, USA, 2007
 [3] Mallik A, Lin Bin, Memik G, et al. User-driven Frequency Scaling[J]. IEEE Computer Architecture Letters, 2006, 5(2): 16
 [4] Vallina-Rodriguez N, Hui P, Crowcroft J, et al. Exhausting battery statistics: understanding the energy demands on mobile handsets[C]// Proc. ACM Mobiheld. New Delhi, India, 2010; 9-14
 [5] Paul K, Kundu T K. Android on Mobile Devices: An Energy Perspective[C]// Proc. CIT 2010 IEEE 10th International Conference. Bradford, 2010; 2421-2426
 [6] 卜爱国. 嵌入式系统动态低功耗设计策略的研究[D]. 南京: 东南大学, 2006
 [7] Benini L, Bogliolo A, Micheli G D. A Survey of Design Techniques for System-Level Dynamic Power Management [J]. IEEE Transactions on VLSI Systems, 2000, 8 (3): 299-316
 [8] Kim W, Gupta M, Wei G, et al. System level analysis of fast, per-core DVFS using on-chip switching regulators[C]// IEEE 14th International Symposium on High Performance Computer Architecture(HPCA 2008). 2008; 123-134
 [9] Seokwoo L, Shidhartha D, Toan P. Reducing pipeline energy demands with local DVS and dynamic retiming low power electronics and design[C]// Proc of ISLPED. IEEE Press, 2004; 319-324
 [10] Yan L, Luo J, Jha N. Joint dynamic voltage scaling and adaptive body biasing for heterogeneous distributed real-time embedded systems [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2005, 24(7): 1030-1041
 [11] Android developers[EB/OL]. <http://developer.android.com>
 [12] Zhang Tao, Gurung M S, et al. Reducing Energy Consumption on Mobile Devices with WiFi Interfaces[C]// Proceedings of Global Telecommunications Conference (GLOBECOM' 05). IEEE, 2005; 561-565
 [13] Rehfuess U, Ivanov K. Estimating the Gains of Adaptive Antenna Systems for GPRS and EDGE Data Services in GSM Networks [C]// Proc. Vehicular Technology Conference. IEEE VTS-Fall VTS, 2000; 3026-3032

(上接第 257 页)

[11] He X, Ma W-Y, Zhang H-J. Learning an image manifold for retrieval[C]// Proceedings of the ACM International Conference on Multimedia. 2004
 [12] Ko B, Byun H. Integrated region-based image retrieval using region's spatial relationships[C]// Proceedings of the IEEE International Conference on Pattern Recognition(ICPR). 2002
 [13] Li J, Wang J Z, Wiederhold G. IRM: International region matching for image retrieval[C]// Proceedings of the ACM International Conference on Multimedia. 2000
 [14] Do M N, Vetterli M. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance [J]. IEEE Trans. Image Process, 2002, 11(2): 146-158

[15] 王守觉. 基于仿生形象思维方法的图像检索算法[J]. 电子学报, 2010, 38(5): 993-997
 [16] 杨芳宇. 一种基于边缘综合特征的彩色图像检索算法[J]. 计算机科学, 2010, 37(2): 256-260
 [17] Philbin J, Chum O, Isard M, et al. Object retrieval with large vocabularies and fast spatial matching[C]// Proc. CVPR. 2007
 [18] Fergus R, Weiss Y, Torralba A. Semi-supervised learning in gigantic image collections[C]// Advances in neural information processing systems. 2009
 [19] Schroff F, Criminisi A, Zisserman A. Harvesting image databases from the Web[C]// Proceedings of the 11th International Conference on Computer Vision. Riode Janeiro, Brazil, 2007