

一种改进的提取显著点的图像检索技术^①

赵 珊^{1,2} 王 水¹ (1.河南理工大学 计算机科学与技术学院 河南 焦作 454000)

(2. 南京邮电大学 江苏省图像处理与图像通信重点实验室 江苏 南京 210003)

摘 要: 在基于目标区域的图像检索中,显著点是一种重要的点特征。针对经典的显著点提取算法 SPARSE (Salient Points Auto-Reduction using Segmentation) 存在的复杂度高等问题,提出了一种改进算法,利用动态阈值分割算法中的类间方差和类内方差对图像进行分割,然后用三个颜色特征和三个纹理特征对分割出的显著点进行特征标注,最后用欧氏距离对显著点特征向量进行相似性度量。实验结果表明,改进后算法提取的显著点用于图像检索具有较好的检索效果。

关键词: 显著点 动态阈值分割算法 图像检索

Improved Image Retrieval of Extraction Salient Points

ZHAO Shan^{1,2}, WANG Shui¹

(1.School of computer science and technology, Henan Polytechnic University, Jiao zuo 454000, China; 2.Jiangsu Province Key Lab. on Image Processing and Image Communication, Nanjing University of Posts and Communication, Nanjing 210003, China)

Abstract: In ROI-based image retrieval, salient point is an important point feature. The classic extraction algorithm of salient points-SPARSE(Salient Points Auto-Reduction using Segmentation) is more complex. This paper presents an improved algorithm, using dynamic threshold segmentation algorithm in the between-class variance and within-class variance of the image segmentation. And then it uses three color features and three texture features of the segmentation to point out the significant features of labeling. Finally it makes similarity measurement with Euclidean distance of feature vectors of salient points. Experimental results show that the improved algorithm's extraction of salient points for image retrieval has better search results.

Keywords: salient points; dynamic threshold segmentation algorithm; image retrieval.

1 引言

在图像中,显著点是表征图像中具有显著变化细节的图像的局部特征。显著点仅代表它本身所在图像中位置的部分区域,以及非常接近的一些邻域,而这些局部区域结合在一起恰恰可以表征整幅图像的特征。因此,为了获取稳定的图像局部显著点特征,Q Tian 等^[1]和 E Loupias^[2]提出了基于小波变换的显著点特征的提取技术;J Wang 等^[4]提出组合了兴趣点和边缘特征的图像检索技术;C Wolf 等^[5]提出了基点检测算法;D G Lowe^[3]提出了图像尺度不变关于兴趣点和纹理特征的图像检索技术。这些显著点提取和检索方法的提出,

大大促进了基于感兴趣目标区域图像检索技术的发展。然而,由于显著点经常位于目标区域的边界上,所有以上这些显著点特征的提取方法,均不同程度地包含进了不同目标区域的像素,存在一定的误差;另外这些显著点经常在图像的不同纹理部分,或在图像特征显著变化的纹理的边上,因此多个显著点获取的是图像的不同部分,容易产生特征冗余,不利于图像的相似性比较;而且,显著点还可能聚集在高纹理区域和图像特征的显著变化带(如边缘),造成一个显著点的特征集非常庞大,需要包含所有可能感兴趣对象的显著点,复杂度过高。R Rahmani 等^[6]结合多实例

^① 基金项目:南京邮电大学江苏省图像处理与图像通信重点实验室开放基金(ZK208002);河南省教育厅自然科学基金基础研究基金(2008B520015, 2009B520013);河南理工大学博士基金(B2008-61,B2009-91)

收稿时间:2009-12-28;收到修改稿时间:2010-02-06

学习(Multiple Instance Learning, MIL)的优点提出了 SPARSE 算法,他们使用图像分割,形成一个模板,限制在每个部分的显著点数量,取得了较好的效果。但却存在着分割的区域不稳定、经常产生过分割、计算量较大、每个区域中的显著点数量不好确定、且对显著点的多样性有一定的限制作用等问题。针对上述问题,本文采用图像的动态阈值分割方法,形成一种遮掩,以限制每一分块内的相同特征显著点的数量;并能对不同特征的显著点进行更加明显的区分,以维持显著点的多样性。实验结果证明,我们改进后的显著点提取方法,不但计算速度得到了较大的提高,而且用于图像检索也具有较好的效率。

2 基于分割的显著点自动减少技术(Salient Points Auto-Reduction using Segmentation, SPARSE)

在图像中,颜色特征可以是各种颜色的比例分布,也可以是颜色的空间分布,目前大多数的检索技术采用颜色比例分布作为颜色基本特征。而纹理(texture)特征是图像表面的结构模式,例如图像中的木纹、水泥、谷物、沙子、玻璃等物体。同时纹理特征可以看作是在灰度和彩色上的规律图案,这些图案可以认为是许多基本纹理小单元的重复生成,这些小单元可称之为纹理元。而且每个纹理元均可以用它的方向向量来表征它本身的纹理特征。具体方法如下。

①用 3 个颜色特征(Y, C_b, C_r)和 3 个纹理特征(HL, LH 和 HH)描述图像中的每个像素。 Y 为亮度分量, C_b 为蓝色分量和参考值的差, C_r 为红色分量和参考值的差; $V_r^h(p(i, j))$ 表示纹理的水平(HL)分量, $V_r^v(p(i, j))$ 表示纹理的垂直(LH)分量, $V_r^d(p(i, j))$ 表示纹理的对角(HH)分量。求取颜色特征和纹理特征的公式如下:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

$$V_r^h(p(i, j)) = \left(\sum_{r=1}^{k/2} \left(\sqrt{\sum_{x=0}^{k/2-r} \sum_{y=0}^{k/2-r} C_{HL}(x, y)^2} / k \cdot 2^r \right) \right) / \sum_{r=1}^{k/2} \quad (2)$$

$$V_r^v(p(i, j)) = \left(\sum_{r=1}^{k/2} \left(\sqrt{\sum_{x=0}^{k/2-r} \sum_{y=r}^{k/2-r} C_{LH}(x, y)^2} / k \cdot 2^r \right) \right) / \sum_{r=1}^{k/2} \quad (3)$$

$$V_r^d(p(i, j)) = \left(\sum_{r=1}^{k/2} \left(\sqrt{\sum_{x=0}^{k/2-r} \sum_{y=0}^{k/2-r} C_{HH}(x, y)^2} / k \cdot 2^r \right) \right) / \sum_{r=1}^{k/2} \quad (4)$$

其中, $k * k$ 表示模板窗口的大小, $C_r(x, y)$ 表示纹理 S 元在位置 (x, y) 的第 r 级小波变换后的小波系数, $s \in \{HL, LH, HH\}$ $0 \leq x < k, 0 \leq y < k$, 且 $r \leq \log_2 k$ 。

②用一个基于聚类的分割算法,对具有显著特征的像素进行合并,以形成显著点。更具体一点,每个分割开始时是单个像素,为每个像素赋予用 3 个颜色特征和 3 个纹理特征,并且根据每个像素所拥有的这六个特征,在相邻像素间进行相似度比较,并把相似度高的归并到一组,重复执行,直到分割达到所期望的结果。其中像素间的相似性比较,用特征向量的欧氏(Euclidean)距离来进行度量。

③用一个模板(如 3×3 或 9×9 等)对图像中的显著点进行检测。具体就是:用模板从图像的左上角开始进行水平扫描直至图像的右下角结束为止,来检测图像中的显著点。模板在图像中任意显著点的响应由 3 个颜色特征(Y, C_b, C_r)和 3 个纹理特征(HL, LH 和 HH)给出。在这里,模板被当作掩模来用,当掩模的中心位于一个孤立的显著点时,掩模的响应必须最强,而在亮度不变的区域中响应为零。如图 1 为一个点检测模板以及其对图像进行显著点检测的结果。



(a)模板 (b)原图像 (c)显著点检测结果

图 1 点模板检测实例

④最后对每幅图像进行分区(一般分为 10—50 个区域),在每个区域中保持 k ($k=3$ 或其它值,要求最能表达这个区域的特征)个最有价值的显著点,并删除其它;如果一个区域中的显著点少于 K 个,则全部保留。这样每个区域中 K 个显著点的特征即代表了这个区域,又可以用它们进行下一步的相似性度量。

虽然 SPARSE 算法取得了较好效果。但也存在过分割、计算量较大、每个区域中的显著点数量不好确定等问题。因此,需要对 SPARSE 算法进行必要的改进。

3 改进的SPARSE算法(Improved Salient Points Auto-Reduction using Segmentation, ISPARSE)

显著点的性质决定了它可以捕获比图像分割所获得的较大的区域更精细的细节。如图像分割方法获取相同的细节的唯一途径是对图像进行过高、过多的分割,这样会不利于图像特征表示。因此,改进 SPARSE 算法中,首先利用动态阈值分割法对图像进行初始分割,以减少 SPARSE 中针对每个像素进行标注的计算量。然后利用类间方差和类内方差阈值法分离出每个区域中的显著点,并对显著点进行特征提取(用 3 个颜色特征和 3 个纹理特征)。接着在每个区域中根据显著点特征的相似性进行合并,以限制每个分割区域内的显著点数量;同时分裂具有不同特征的显著点,以维持显著点的多样性。从而达到对各个区域的重新定义,以利于特征提取和相似性比较。

3.1 用动态阈值进行初始分割

设 $f(x, y)$ 为 $N \times M$ 像素的图像 F 在图像中的坐标点 (x, y) 处的灰度值,灰度级为 L , 设 T 为该图像的一个阈值,则对目标区域和背景的有效分割准则为:目标区域部分: $\{f(x, y) \geq T\}$; 背景部分: $\{f(x, y) \leq T\}$ 。阈值 T 的选取步骤如下:

- ①为 T 选取一个初始估计值(一般取图像中最大亮度值和最小亮度值的中间值)。
- ②使用 T 分割图像,所产生的两组像素值 G_1 和 G_2 。
- ③计算 G_1 和 G_2 范围内的像素的平均亮度值和。
- ④计算一个新阈值: $T = (\mu_1 + \mu_2) / 2$ 。
- ⑤重复步骤②和步骤④,直到 T 的值比预先指定的参数 T_0 小为止。

其中参数 T_0 , 以选取图像颜色直方图峰谷的平均阈值为最佳。

首先针对一幅图像,计算它的像素灰度离散概率密度函数的归一化直方图,

$$p_r(r_q) = \frac{n_q}{n} \quad q=0,1,2,\dots,L-1 \quad (5)$$

其中, n 是图像中的像素总数, n_q 是灰度级为 r_q 的像素数目, L 是图像中所有可能的灰度级数。并找到最大化类间方差 σ_b^2 的阈值 k ,

$$\sigma_b^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \quad (6)$$

其中 $\omega_0 = \sum_{q=0}^{k-1} p_q(r_q)$, $\omega_1 = \sum_{q=k}^{L-1} p_q(r_q)$, $\mu_0 = \sum_{q=0}^{k-1} qp_q(r_q) / \omega_0$, $\mu_1 = \sum_{q=k}^{L-1} qp_q(r_q) / \omega_1$, $\mu_T = \sum_{q=0}^{L-1} qp_q(r_q)$ 。这时会产生阈值 $T=K$, 用阈值 T 对图像进行分割。具体分割示例如图 2(a), 这里 $T=128$ 。

然后,在用阈值 T 进行初始分割的基础上,对分割出的局部区域利用直方图的最大化类内方差的阈值进行进一步的分割。

$$\sigma_c^2 = \omega_2(\mu_2 - \mu_T)^2 + \omega_3(\mu_3 - \mu_T)^2 \quad (7)$$

其中, ω_2 , ω_3 , μ_2 , μ_3 分别是在类间方差基础上直方图阈值的进一步细化。从而可以有效地将目标区域和显著点分离出来。如图 2(b), 这里 $T=64$ 。



(a)类间方差分割结果 (b)类内方差分割结果

图 2 对图 1 中的原图像用动态阈值分割的结果

3.2 显著点特征提取

分割出显著点以后,就要对显著点进行必要的标注,以利于下一步对显著点的识别。算法中对显著点的标注,同样采用 3 个颜色特征 (Y, C_b, C_r) 和 3 个纹理特征 $(HL, LH$ 和 $HH)$ 。颜色特征的提取比较简单,把图像的原始颜色空间(RGB)转换到 YCbCr 颜色空间,即可得到 Y, C_b, C_r 三个颜色特征;而纹理特征的提取,需要我们利用 Haar 小波对显著点进行变换处理,可以得到显著点的 HL, LH 和 HH 三个方向的特征向量。如图 3, (a)表示经过处理的显著点图像,其中红色区域表示要进行纹理特征向量提取的部分, (b)表示 (a)图中红色显著点的放大图像, (c)图是显著点经过 Haar 小波变换后的结果。

对每个显著点特征表示后,按照欧氏距离在每个目标区域中进行显著点的合并,以改善目标区域的可识别性。最后在每个目标区域中,保持了 k 个最有价值的显著点,并删除其它;如果一个区域中的显著点少于 k 个,则全部保留。这样每个区域中的 k 个显著点的特征即代表了该区域的特征,可以用它进行下一步的图像间的相似性度量。

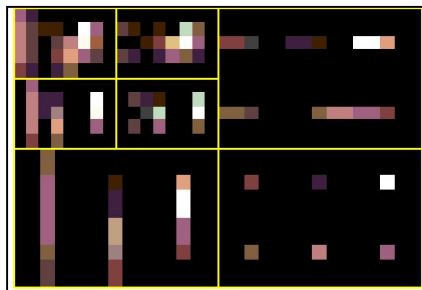
本文中采用的图像动态阈值分割法,借助于人机交互式系统,实现了显著点和目标区域的有效分离。不但提高了运算速度,而且对于目标区域的提取也是

一个较大的突破。

因为分割被当作一个减少显著点的滤波器来使用的，所以 CBIR 系统的性能不依靠分割范围的精确与否，而每个显著点的描述是基于以这个显著点为中心的、独立的分割区域。换言之，这个分割仅影响显著点的选取而不影响它们的表示。



(a) 一幅显著点图像 (b) 图(a)中红色区域的 50 倍放大图像



(c) 二维 Haar 小波对图(b)的变换结果

图 3 用 Haar 小波对显著点纹理特征的提取

4 相似性度量

针对示例图像和图像库中的图像，本文采用欧式距离的相似性度量准则。设 Q 为示例图像， I 为图像库中图像，针对显著点提取的特征向量为 $f_Q = \{u_1, u_2, \dots, u_n\}$ ， $f_I = \{v_1, v_2, \dots, v_n\}$ ，其中 $i = 1, 2, \dots, n$ 为图像中显著点的个数。则 Euclidean 距离为：

$$D_{Q,I}^E = \sqrt{(f_Q - f_I)^2} = \sqrt{\sum_{i=1}^n \sum_{j=1}^6 (u_{ij} - v_{ij})^2} \quad (8)$$

这里 u_{ij} 和 v_{ij} 分别是示例图像和图像库中图像的显著点特征向量的分量。 $D_{Q,I}^E$ 值越小，说明两个图像越相近，当为 0 时，两个图像完全一致。这种方法计算简单且受噪声的影响相对较小。

5 仿真实验结果和分析

为了验证本文算法的工作性能，采用 Coerl 图像库进行了多特征彩色图像检索实验，并与三种传统的显著点提取方法进行了对比。实验中，选用了

SIMPLicity 系统使用的测试集 (<http://wang.ist.Psu.edu/docs/related>) 作为图像库，该测试集是从 Corel 图像库中选取的，共包含有 10 类(包括 Bus、Flower、Horse 等)，其中，每类包含有 100 幅图像。另外，为评价图像检索算法的效果，本文采用“查准率(precision)”和“查全率(recall)”曲线作为相似检索效果的评价准则(如图 4)。

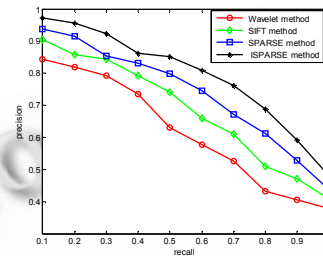


图 4 四种显著点提取方法的比较结果

为了说明算法的鲁棒性，首先从测试图像库的每类图像中随机抽取 30 幅图像(共计 300 幅)；然后对这 300 幅图像分别进行加噪处理(包括模糊、锐化、光照、滤波等)、旋转处理、平移和缩放处理；最后将上述 300 幅处理后的图像与原来的 1000 幅图像集中在一起，构成新的测试图像库进行检索。检索时，从上述新测试图像库中，每类随机抽取 5 幅作为示例图像，共构成 50 次查询，每次查询选取前 20 幅最相似的图像作为检索结果。对于每类图像而言，需计算其 5 次查询结果的查准率平均值、查全率平均值，并将其作为该类图像的平均查准率和平均查全率。

本文采用的是 Matlab 7.0 仿真系统。实验中，SPARSE 和小波方式用的是相同的显著点提取和表示的方式(3 个颜色和 3 个纹理空间)，它们之间的主要差别是：部分显著点在图像中的位置不同。而 SIFT^[3]特征向量有 128 个，它的描述围绕着一个显著点的局部梯度方向直方图，因为高维性使它变得非常难以计算；但 SIFT 的执行效果比小波的所有类别好 5.9%。尽管 SPARSE 只用 6 个方面的特征表示，但 SPARSE 还是优于 SIFT 3.2%。而 ISPARSE 算法在 10 个类别中的 9 个优于小波方式，其中统计学意义的有 6 个。

结果显示，ISPARSE 作为单独的技术，以改进图像特征的提取和表示，可以帮助提高检索性能(对所有类别的平均值有 4.8% 的改善)。如果和更好的分割算法一起使用时，可以进一步改善性能。

(下转第 172 页)

6 结论

本文提出了一种改进的 ISPARSE 算法,它利用了原始图像分割可以作为过滤器的原理,把动态阈值分割法中简洁快速的分割优点融入到 SPARSE 中,通过使用分割遮掩技术分散和减少一些显著点,但同时仍然保持着显著点的多样性,从而提高了检索性能。改进的 ISPARSE 的使用还可以提高运行性能,如通过减少每个特征集中的特征向量来降低了计算的复杂性。结果表明,这个算法使基于目标区域的图像检索技术得到了进一步改善。该算法也改善了基于分割的表示和基于显著点表示的方式的性能。

参考文献

- 1 Tian Q, Sebe N, Lew MS, Loupias E, Huang TS. Image retrieval using wavelet-based salient points. *Journal of Electronic Imaging, Special Issue on Storage and Retrieval of Digital Media*,78:138 – 156, 2001.
- 2 Loupias E. Salient points detection using wavelet transform, <http://telesun.Insaly-on.fr/loupias/points/demo.html>, 2009,3.
- 3 Lowe DG. Distinctive image features from scale-invariant keypoints. *Int. Journal of Computer Vision*, 37(2):151 – 172,2004.
- 4 Wang J, Zha H, Cipolla R. Combining interest points and edges for content-based image retrieval. *IEEE Int. Conf. on Image Processing*, 2005.947 – 963.
- 5 Wolf C, Jolion JM, Kropatsch W, Bischof H. Content based image retrieval using interest points and texture features. *IEEE Int. Conf. On Pattern Recognition*, 2000. 644 – 655.
- 6 Rahmani R, Goldman S, Zhang H, Krettek J, Fritts J. Localized content-based image retrieval. *ACM Workshop on Multimedia Image Retrieval*, 2005.227 – 236.