文章编号:1674-7070(2010)02-0124-05

一种特征可控的图像检索方法

陈文兵[1]　王伟[1]　王霞[1]

摘要

提出了一种新的图像检索方法,它利用图像颜色特征和空间特征的加权平均作为图像的相似性度量.与目前常用的图像检索方法相比,该算法在查精度及查全度两个指标上均有明显的提高,且对结构相似但颜色完全不同的图像也能够有效地进行检索.同时,由于特征加权因子的可调节性使得所设计的系统可根据用户检索所反馈的信息进行调节及控制.

关键词

图像检索;聚类;相似度

中图分类号 TN911.73

文献标志码 A

0　引言

Introduction

基于内容的图像检索在类似于多媒体数据库系统的应用领域起重要作用.相关的研究主要利用图像的颜色、纹理、形状及布局等特征定义图像的相似性,在这些视觉特征中,颜色是最具代表性的特征.颜色直方图[1]、颜色矩[2]及颜色相关图[3]可以描述颜色特征.Fukunaga 等[4]提出了 Mean Shift 密度函数估计方法,它是一种非参数估计方法.近年来,Mean Shift 方法的应用很广泛,常用来对图像进行平滑、聚类及分割.利用适当改进的该方法对图像进行聚类,可以得到带有颜色和像素数的聚类集,进一步分析这些聚类,可以发现,这些聚类不仅包含了图像的颜色特征和位置特征,还包含了空间大小特征.对查询图像和目标图像,本文通过改进的 Mean Shift 分割算法分割聚类并对聚类集按从大到小进行排序,从排序后的 2 幅图像的聚类集中选取前若干个主聚类并比较对应聚类的颜色偏差、空间尺度偏差,基于这 2 个偏差给出一种基于颜色及空间尺度的加权平均相似度定义,以此作为查询图像与目标图像匹配的度量.同文献[1-3]相比,本文提出的图像检索方法不仅包含了颜色信息而且含有空间尺度及位置信息.实验分析与对比结果表明:本文算法图像检索的查精度及查全度明显提高;对结构相似但颜色完全不同的图像能够有效地进行检索;加权因子的可调节性使得系统可根据用户对特征的要求进行选择,使得检索系统具有高度的灵活性.

1　图像预处理

Image preprocessing

1.1　图像缩略及颜色空间变换

在图像检索之前先对图像进行预处理是非常必要的,本文对图像的预处理主要考虑 2 个因素,一是提高算法的性能,二是颜色空间的表示方法与人类视觉感知的一致性.为了提高性能,采用了对图像进行缩略的方法;同时,为了更好地反映人类视觉对颜色感知的一致性,需要将图像从 RGB 颜色空间变换到 CIELUV 颜色空间.

由于 Mean Shift 聚类过程需要迭代,计算量相对较大,但在图像检索中比较的是图像的颜色及结构特征,细节并不重要,故在利用

收稿日期 2009-11-09

资助项目 江苏省教育厅自然科学基金(06KJD110119,08KJB520004);南京信息工程大学科研基金(20070123)

作者简介

陈文兵,男,硕士,副教授,主要研究领域为图像处理、计算数学、计算机应用与软件.chenwb@nuist.edu.cn

1　南京信息工程大学　数理学院,南京,210044

Mean Shift 过程聚类之前先利用缩略方法[5-6]对图像进行预处理. 图 1 显示了与目标图片对比的 4 幅图片;表 1 是缩略前后相似度的对比. 由表 1 中的数据可以清楚地观察到颜色与像素比例偏差很小,对图像检索的影响可以忽略不计,这表明利用缩略算法预先对图像进行预处理没有改变图像的主要特征,因此利用缩略算法对图像进行预处理是可行并有效的.



图 1 目标图像及查询图像

Fig. 1 Target image and query images

表 1 缩略前和缩略后的目标图像与测试图像的颜色及像素比较

Table 1 color and pixel comparison of target/test image thumbnails and source images

图片	前色彩偏差	前像素比例偏差	后色彩偏差	后像素比例偏差
P1	0.881 0	0.914 6	0.879 5	0.908 4
P2	0.869 4	0.902 4	0.875 4	0.902 1
P3	0.754 5	0.885 4	0.745 2	0.862 1
P4	0.845 5	0.845 6	0.840 2	0.836 5

1.2 颜色空间变换

国际照明委员会推出 2 种色彩模式,即 CIELUV 空间和 CIELAB 空间,两者都具有感知上的均匀性,与人类对颜色的感知非常接近,即视觉上近似的 2 种颜色在 2 个颜色空间上位置相邻;它们还具有欧式距离不变性,即视觉上差异较小的 2 种颜色在 2 个颜色空间上的欧式距离较小,而视觉上差异较大的 2 种颜色对应的欧式距离也较大. 本文采用 CIELUV 颜色空间,其中,L 代表亮度信息,U 和 V 代表色度信息,2 点间的欧氏距离 $d = \sqrt{\Delta L^2 + \Delta U^2 + \Delta V^2}$ 表示了 2 点间的色度差异,d 越大表示 2 点间色彩差别越大. 而 CIEXYZ、RGB 和 HSV 颜色空间均不是感知均匀颜色空间.

为了准确地反映颜色的差异,需将缩略后的图像从 RGB 颜色空间转换到 CIELUV 颜色空间. 关于 RGB 颜色空间到 CIELUV 颜色空间的变换算法参见文献[7].

1.3 改进的 Mean Shift 聚类过程

Mean Shift 算法是一种有效的统计迭代算法,使用一个无参密度估计函数,得到特征向量 $\boldsymbol{x}$ 的概率密度函数的模.

给定 d 维空间 R^d 中的 n 个样本点 $\{\boldsymbol{x}_i, 1 \leqslant i \leqslant n\}$,$K(\boldsymbol{x})$ 为核函数,h 为核半径带宽,则 $\boldsymbol{x}$ 的核密度估计定义

$$f(\boldsymbol{x}) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left(\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right). \tag{1}$$

最优的核函数是采用具有最小平方差的核函数 Epanechnikov:

$$K(\boldsymbol{x}) = \begin{cases} (2C_d)^{-\frac{1}{2}}(d+2)(1+\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}), & \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} < 1; \\ 0. \end{cases} \tag{2}$$

其中,C_d 是 d 维空间中单位球的体积.

另一个是均匀核,称作高斯核函数:

$$K(\boldsymbol{x}) = (2\pi)^{-\frac{d}{2n}}\exp\left(\frac{1}{2}\left\|\boldsymbol{x}\right\|^2\right), \tag{3}$$

$$\nabla f(\boldsymbol{x}) = \frac{2c_k}{nh^{j+2}}\left[\sum_{i=1}^{n} g\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right)\right] \cdot \left[\frac{\sum_{i=1}^{n} \boldsymbol{x}_i g\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right)} - \boldsymbol{x}\right].$$

由文献[8]得到核密度估计的梯度,Mean Shift 向量为

$$M_h(\boldsymbol{x}) = \frac{\sum_{i=1}^{n} \boldsymbol{x}_i g\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right)} - \boldsymbol{x}. \tag{5}$$

由 Mean Shift 向量得到 Mean Shift 迭代公式:

$$\boldsymbol{y}_{j+1} = \frac{\sum_{i=1}^{n} \boldsymbol{x}_i\, g\left(\left\|\frac{\boldsymbol{y}_j - \boldsymbol{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right)}, \quad j = 1, 2, \cdots. \tag{6}$$

由文献[9]可知,当式(6)中 $j \to \infty$ 时,向量序列 $\{\boldsymbol{y}_j\}$ 收敛.

在空间范围域,利用 CIELUV 的欧式距离来度量颜色的差异,通过 Mean Shift 聚类算法得到图像中每个聚类的平均颜色值以及像素数,聚类过程中仅需要输入 (h_s, h_r),其中 h_s 和 h_r 分别代表空间核窗宽和范围核窗宽[10-11].

设 $\{\boldsymbol{x}_j\}_{j=1,\cdots,n}$ 为 5 维原始图像点,$\{\boldsymbol{z}_j\}_{j=1,\cdots,n}$ 为收敛点,$\{L_j\}_{j=1,\cdots,n}$ 为标签集. Mean Shift 聚类算法为

1) 运用 Mean Shift 算法计算 $\boldsymbol{x}_j$,记收敛后的值

为 z_j,其中 $j=1,\cdots,n$;

2) 在联合域中,通过比较所有的 z_j 之间的距离,选择更接近给定的距离阈值确定聚类 $\{C_i(p_i,n_i)\}_{i=1,\cdots,n}$ 收敛点,并且记录这样的点的个数.

其中:p_i 表示第 i 个聚类 C_i 的颜色平均值;n_i 则表示它的像素个数.

2 相似度定义
Similarity measure definition

给定一幅图像 I,首先利用 Mean Shift 聚类算法对图像 I 进行聚类,图像 I 经过聚类后,可以表示成一个聚类的集合 $\{C_i(p_i,n_i)\}_{i=1,\cdots,n}$,这里,$C_i(p_i,n_i)$ 是第 i 个聚类,n 是总聚类个数. 对第 i 个聚类,p_i 是第 i 个聚类的颜色平均值,n_i 是第 i 个聚类的像素个数,则对应于图像 I 的聚类集合 $\{C_i(p_i,n_i)\}_{i=1,\cdots,n}$,可以得到相应的聚类颜色平均值集合 $\{p_i\}_{i=1}^{n}$ 及聚类像素个数集合 $\{n_i\}_{i=1}^{n}$,显然聚类颜色平均值集合刻画了对应聚类构成的均匀区域的颜色特征,而聚类像素个数刻画了对应聚类构成的均匀区域的空间(大小)特征,进而,对一幅图像主要特征进行刻画仅需抽取若干个主控聚类即可.

针对 2 幅图像的查询,给定目标图像 I_1 及查询图像 I_2,$\{C_{1i}(p_{1i},n_{1i})\}_{i=1,\cdots,n}$ 及 $\{C_{2i}(p_{2i},n_{2i})\}_{i=1,\cdots,n}$ 分别表示图像 I_1 的聚类集合和图像 I_2 的聚类集合. 对图像 I_1 及图像 I_2 的聚类集合 $\{C_{1i}(p_{1i},n_{1i})\}_{i=1,\cdots,n}$ 和 $\{C_{2i}(p_{2i},n_{2i})\}_{i=1,\cdots,n}$,均按聚类的像素个数按从大到小排序,并对聚类颜色平均值集合的索引次序依据聚类像素个数集合的次序进行相应的调整,在 2 个图像聚类集中均取 N 个像素个数较大的聚类构成子集,称之为主聚类集,不妨设为 $\{C_{1i}(p_{1i},n_{1i})\}_{i=1,\cdots,N}$ 及 $\{C_{2i}(p_{2i},n_{2i})\}_{i=1,\cdots N}$.

定义 1 (聚类像素数绝对差比) 图像 I_1 及 I_2 的平均颜色值绝对差(简称为 ADRC)为

$$R_{\mathrm{ADRC}}=\frac{\sum_{i=1}^{N}\left|p_{1i}-p_{2i}\right|}{p},\tag{7}$$

这里

$$p=\max\left\{\sum_{i=1}^{N}p_{1i},\sum_{i=1}^{N}p_{2i}\right\}.$$

定义 2 (平均像素绝对差比)图像 I_1 及 I_2 的平均像素绝对差(简称为 ADRP)为

$$R_{\mathrm{ADRP}}=\frac{\sum_{i=1}^{N}\left|n_{1i}-n_{2i}\right|}{M},\tag{8}$$

这里

$$M=\max\left\{\sum_{i=1}^{N}n_{1i},\sum_{i=1}^{N}n_{2i}\right\}.$$

在式(7)、(8)中,N 表示主聚类的个数,$N\ll n$,可以由用户指定,n 表示一幅图像经 Mean Shift 聚类过程聚类后得到的聚类数.

定义 3 (图像相似度)图像 I_1 及 I_2 的相似度(简记为 S)为

$$S(I_1,I_2)=w_{\mathrm{ADRC}}(1-R_{\mathrm{ADRC}}(I_1,I_2))+w_{\mathrm{ADRP}}(1-R_{\mathrm{ADRP}}(I_1,I_2)).\tag{9}$$

式(9)中,w_{ADRC} 是图像颜色相似的权值,w_{ADRP} 是图像聚类空间结构相似性的权值,$w_{\mathrm{ADRC}}+w_{\mathrm{ADRP}}=1$. 显然,上述算法不仅强调了图像的颜色特征,而且强调了图像的空间结构特征.

3 图像检索过程
Procedure of image retrieval

本文算法描述如下:

1) 利用缩略方法将查询图像 I_1 缩略成图像 I_a;

2) 由 RGB 颜色空间转换到 CIELUV 颜色空间,得到图像 I_{al};

3) 应用 Mean Shift 聚类过程对进行聚类,得到图像 I_{al} 的聚类集合 $\{C_{1i}(p_{1i},n_{1i})\}_{i=1,\cdots,n}$;

4) 对聚类集合 $\{C_{1i}(p_{1i},n_{1i})\}_{i=1,\cdots,n}$ 按聚类像素个数按从大到小排序,选取前 N 个主聚类 $\{C_{1i}(p_{1i},n_{1i})\}_{i=1,\cdots,N}$;

5) 根据式(6),(7),(8)输入 w_{ADRC},w_{ADRP} 和阈值 S,计算图像 I_1 及 I_2 之间的相似度 $S(I_1,I_2)$,若 $S(I_1,I_2)$ 大于给定的阈值 S,则图像 I_1 及 I_2 相似;

6) 重复 1) ~5),用户可以根据自己需求调整颜色特征和结构特征的加权因子,即 w_{ADRC},w_{ADRP},其中图像 I_2 是存储在数据库中的图像.

4 实验结果
Experiment result

本文实现了一个原型系统,实验所采用测试图像数据库取自于网站 http://photo.itbulo.com/,总计 2 200 幅,100 个分类,每个分类含图像数为 14 到 30 幅不等,每次检索系统返回 14 幅图片.

实验随机地从 100 个分类中选取了 5 幅图片作为检索图像,选出的图片是儿童、飞机、地球、芯片及花等 5 个类别. 采用惯用查全率和查准率对算法进行评价. 这 2 个指标被定义如下:

$$查全率(\%)=\frac{检索到的相关图像数}{数据库中相关图像总数},\quad(10)$$

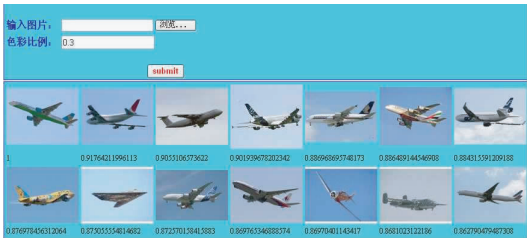
$$查准率(\%)=\frac{检索到的相关图像数}{检索返回的图像总数}.\quad(11)$$

实验目标一:对同一个语义图像类,选取同一个查询图像但取不同的加权因子,测试加权因子对检索结果的影响. 加权因子取值为 0.2、0.3、0.5、0.8,实验结果如图 2.

实验目标二:与基于主色直方图的图像检索方法对比,主要比较基于 RGB 颜色空间的直方图方法及基于 HSV 颜色空间的直方图方法.



a. w_{ADRC}=0.2



b. w_{ADRC}=0.3



c. w_{ADRC}=0.5



d. w_{ADRC}=0.8

图 2 对同一目标图像取不同加权因子的检索结果

Fig. 2 Retrieval results in different weighted factors for the same target images

图 2 显示检索图像为飞机时得到的检索结果,其中第 1 幅图像为检索图像. 图 3 为基于主色直方图的检索结果,表 2 记录了本文方法加权因子 $w_{ADRC}=0.8$ 时与另外 2 种方法进行同样的检索得到的评价数据. 图 4 为取不同的权值 w_{ADRC}得到的查全率和查准率的对比曲线. 从图 2 中可以观察到在 $w_{ADRC}=0.2$, 0.3,0.5 时,有 1 架颜色完全不同于查询图像的飞机(黄色),这表明虽然颜色完全不同,但 2 幅图像结构具有相似性,在权值 w_{ADRC}趋于 1 时检索更加强调的是颜色特征. 但在 $w_{ADRC}=0.5$ 附近时的查全率和查准率却最高(图 4),这表明平均考虑颜色特征和空间特征检索效果最好. 表 3 中的数据表明本文方法在整体性能上明显优于其它 2 种方法. 此外,由于加权因子 w_{ADRC}的可调节性,使得检索系统的灵活性得以提高.



a. RGB直方图



b. HSV直方图

图 3 基于主色直方图法的检索结果

Fig. 3 Retrieval results in histogram method based on main color

表 2 不同检索方法的查全率和查准率

Table 2 Precision and recall of different retrieval method

样本	本文算法(w_{ADRC}=0.8)		RGB 直方图		HSV 直方图	
	查全率	查准率	查全率	查准率	查全率	查准率
儿童	0.650	0.929	0.550	0.786	0.400	0.570
飞机	0.560	1.000	0.120	0.214	0.440	0.786
地球	0.700	1.000	0.600	0.857	0.650	0.929
芯片	0.923	0.857	0.692	0.643	0.846	0.786
花	0.450	0.643	0.250	0.357	0.450	0.643
平均	0.657	0.886	0.442	0.571	0.557	0.743

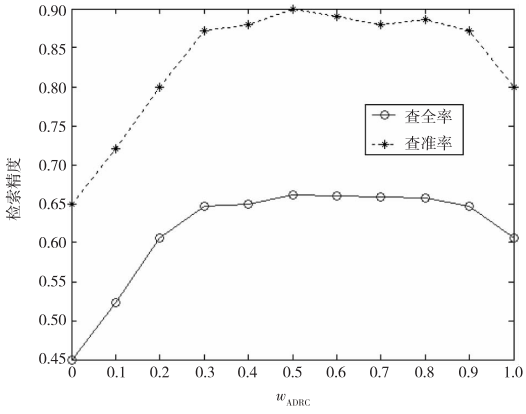


图 4　不同 w_{ADRC} 的查全率与查准率对比曲线

Fig.4　Comparison curve of precision and recall under different w_{ADRC}

5　结论
Conclusion

颜色特征由于其具有图像尺寸和旋转无关性从而被研究人员广泛使用,但单一的颜色特征在实际的图像检索应用中存在明显的不足之处. 本文方法不但利用了颜色特征,而且考虑了结构特征. 实验表明该方法在整体性能上优于目前被广泛使用的基于颜色直方图的方法,并能够有效地检索 2 幅结构相似而颜色完全不同的图像.

参考文献
References

[1] SONG Youngjun, Park W, Kim D, et al. Content-based image retrieval using new color histogram[C]//Intelligent Signal Processing and Communication Systems, Proceeding of International Symposium, 2004: 600-611

[2] Shih J L, Chen L H. Color image retrieval based on primitives of color mements[J]. Vision, Image and Signal Processing, IEEE Processings, 2002, 149(6): 370-376

[3] Rautiainen M, Doermann D. Temporal color correlograms for video retrieval[J]. 16th International Conference on Pattern Recognition, 2002, 1(2): 267-270

[4] Fukunaga K, Hostetler L D. The fstimation of the gradient of a density function, with application in pattern recognition[J]. IEEE Trans, On Information Theory, 1975, 21(1): 32-40

[5] Don L. A review of some image pixel interpolation algorithms[EB/OL]. [2009-08-29]. http://www.tinaja.com/glib/pixintpl.pdf

[6] Cohen H A. Retrieval and browsing of images using image thumbnails[J]. Journal of Visual Communication and Image Representation, 1997, 8(2): 226-234

[7] Wikipedia. CIELUV color space[EB/OL]. [2009-08-29]. http://en.wikipedia.org/wiki/CIELUV_color_space

[8] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis[J]. IEEE Transactions on PAMI, 2002, 24(5): 603-619

[9] LI Xiangru, HU Zhanyi, WU Fuchao. A note on the convergence of the mean shift[J]. Pattern Recognition, 2007, 40(6): 1756-1762

[10] 彭宁嵩,杨杰,刘志,等. Mean-Shift 跟踪算法中核函数窗宽的自动选取[J]. 软件学报, 2005, 16(9): 1542-1550
PENG Ningsong, YANG Jie, LIU Zhi, et al. Automatic selection of kernel-bandwidth for mean-shift object tracking[J]. Journal of Software, 2005, 16(9): 1542-1550

[11] Comaniciu D. An algorithm for data-driven bandwidth selection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(2): 281-288

An image retrieval method based on controllable feature

CHEN Wenbing[1]　WANG Wei[1]　WANG Xia[1]

1　College of Math and Physics, Nanjing University of Information Science & Technology, Nanjing　210044

Abstract　A new image retrieval method is proposed in this paper, which uses the weighted average of color and spatial features as similarity measure between two images. Experimental results show that both precision and recall are significantly improved, compared with most common image retrieval methods; and images with similar structures and completely different colors can be retrieved efficiently. With ajustable weighted factors for different features, this method enables its users to regulate and control retrieval process according to feedback information.

Key words　image retrieval; clustering; similarity measure