

几种基于内容的图像检索的方法

吴波, 王保保

(西安电子科技大学 计算机学院, 陕西 西安 710071)

摘要: 基于内容的图像检索(CBIR)技术依赖于对图像特征(例如颜色直方图、纹理、草图、形状等)的提取,相对于传统的基于文本的图像检索方式,这种方式提高了检索效率和检索的准确率。文中主要介绍了基于颜色和基于纹理特征这两种特征提取方法。这两种方法既能够反映全局特征,又能够兼顾所感兴趣区域的局部特征,是基于内容的图像检索的两种非常有效的方法。

关键词: 颜色; 纹理; 图像检索

中图分类号: TN911.73; TP391.3

文献标识码: A

文章编号: 1673-629X(2006)06-0191-02

Some Methods about Content - Based Image Retrieval

WU Bo, WANG Bao-bao

(School of Computer Science and Engineering, Xidian University, Xi'an 710071, China)

Abstract: Content - based image retrieval(CBIR) techniques rely on the features that are abstracted from the images. This method is more efficient and more accurate than the traditional method of text - based image retrieval. This paper introduces two methods to abstract the features of the image. The one is color - based image retrieval, the other is texture - based image retrieval. At the same time those two methods can reflect global feature and regions - of - interest feature. At the same time those are effective ways to solve the problems about content - based image retrieval.

Key words: color; texture; image retrieval

0 引言

图像检索领域的研究主要分为基于文本和基于内容两大类^[1]。

传统的图像检索基于文本方式,使用关键字或自由文本描述图像数据库中的每幅图像,采用文本匹配检索。该方法需要人工对每幅图像按其内容进行标注,然后将标注信息存到文本数据库中用于后来的检索。显然,随着图像的增多,人工标注非常困难,而且,每个人对图像内容的理解不同会造成标注的主观性过强,不利于用户检索。同时,图像的某些可视信息,例如:纹理,形状等,很难用文本准确描述。鉴于以上原因,基于内容的图像检索(Content - Based Image Retrieval, CBIR)逐渐被人们所重视。这种技术由机器自动提取包含图像内容的可视特征:颜色、纹理、形状、对象的位置和相互关系等,对数据库中图像和查询样本图像在特征空间进行相似匹配,检索出与样本相似的图像。此技术的核心是包含图像内容的可视特征,称为基于内容的图像检索技术。相对于传统的基于文本的检索方式,这种方式提高了检索效率和检索的准确率。

收稿日期:2005-10-11

作者简介:吴波(1980-),男,内蒙古人,硕士研究生,研究方向为网络数据库与智能计算;王保保,副教授,研究方向为网络数据库与智能计算。

1 系统总体结构

1.1 系统结构

基于内容的图像检索主要思路是:从图像中分析抽取底层视觉特征用于检索,特征主要有颜色(Color)、纹理(Texture)、草图(Sketch)、形状(Shape)等来衡量图像之间的相似度以实现基于内容的检索。在底层的视觉特征上,特征的提取主要采用计算机视觉和数字图像处理技术,而图像相似性的匹配主要利用模式识别技术。

系统结构如图1所示。

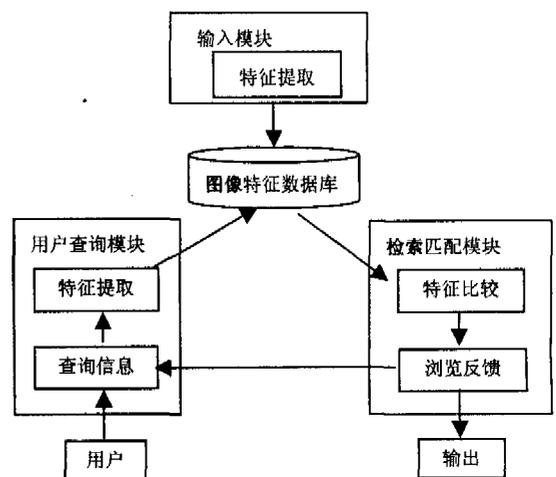


图1 系统结构图

1.2 性能评估

检索策略的性能评估的比较通用的两个准则是效率(efficiency)和有效性(effectivity),分别代表检索的速度和查找与查询图像相似图像的成功率。检索的有效性用查准率(Precision)、查全率(Recall)来评估,分别定义如下:

$$\text{查准率} = \frac{\text{检索出的图像中相关图像的数目}}{\text{检索出图像的数目}} \quad (1)$$

$$\text{查全率} = \frac{\text{检索出的图像中相关图像的数目}}{\text{所有相关的图像数目}} \quad (2)$$

查全率反映系统检索相关图像的能力,而查准率则反映系统拒绝无关图像的能力。

2 图像特征检索

2.1 基于颜色特征的检索

2.1.1 欧氏距离法

衡量两幅图片是否相似可以采用欧氏距离来描述:

$$D = (A, B) = \sqrt{\sum_{i=1}^N (A_i - B_i)^2} \quad (3)$$

(A, B 为两幅图片的颜色直方图, D 越小相似度就越大)

检索后,全图直方图的相似度的定量度量如下表示:

$$\text{Sim}(A, B) = \frac{1}{N} \sum_{i=1}^N (1 - \frac{|A_i - B_i|}{\text{Max}(A_i, B_i)}) \quad (4)$$

(N 为颜色级数, Sim 越靠近 1 两幅图片越相似)

如果对公式(4)加以改进,即对某些相对重要的颜色乘上一个权重,就可以做寻找某一前景或组合的查询。

全图的颜色直方图算法过于简单,因此带来很多问题,如:可能会有两幅根本不同的图像具有完全一样的颜色直方图,不反映颜色位置信息^[2],这样导致查准率和查全率都不高,因此笔者提出了一个改进,即将图像进行了分割,形成若干子块。这样就提供了一定程度的位置信息,而且可以对含用户感兴趣的子块加大权重,提高检索的查询智能性和查全率。相应的公式有子块 A_{ij} 与 B_{ij} 的相似性度量:

$$\text{Sim}_{ij}(A_{ij}, B_{ij}) = \frac{1}{P} \sum_{k=1}^P (1 - \frac{|A_k - B_k|}{\text{Max}(A_k, B_k)}) \quad (5)$$

(P 为所选颜色空间的样点数)

再引入子块权重 W_{ij} , 选取 L 个最大的 Sim 值作 $\text{Sim}_k(G_k, S_k)$, 就有:

$$\text{Sim}(A, B) = \sum_{k=1}^L W_k \text{Sim}_k(A_k, B_k) \quad (6)$$

(W_k 的选取应根据图像的特点决定,可以使图像中间或用户指定的区域权重,以反映图像的位置信息)

由以上公式算法可以推断,根据 Sim 的值就可以得出最终查询结果。

2.1.2 直方图的交集的方法

直方图的交集算法,根据颜色索引检索出查询图像的颜色直方图,将其与图像库中的每一图像的直方图取交集,根据交集的值来选出最佳匹配图像^[3]。

直方图交集的公式如下:

$$D(H_a, H_b) = \frac{\sum_{i,j,k} \min\{H_a(i, j, k), H_b(i, j, k)\}}{\sum_{i,j,k} H_b(i, j, k)} \quad (7)$$

其中 H_a 为查询图像直方图, H_b 为图像库中的任一图像直方图, $D(H_a, H_b)$ 为两图像的匹配值。 $D(H_a, H_b) \in (0, 1)$, 它的值越接近 1, 则两图像越相似。

直方图交集方法能对两幅图像进行详细的比较,然而对于许多合成的图像,如商标等,它们有大量的一致颜色,三维直方图只有几个域的值很高,而其它许多域的颜色信息变化有限。因而,对于这样的图像进行过细的比较是不必要的。又由于在扫描图像时容易产生一些噪声,所以这种过细的比较容易产生错误的结果。

2.2 纹理分析的统计方法

傅里叶频谱法。

该方法的依据是:图像的频域能量谱能在一定的程度上反映某些纹理的特征。由于计算纹理一定要选窗口,仅一个点是无纹理可言的^[4],所以纹理是二维的。设灰度变化函数为 $f(x, y)$, 因此就可以利用二维傅里叶变换 $F(u, v)$ 将图像变换到频域:

$$F(u, v) = \iint_{-\infty}^{\infty} e^{-2\pi j(xu + yv)} f(x, y) dx dy \quad (8)$$

并且,二维傅里叶变换的频谱能描述纹理的粗细程度。所谓傅里叶频谱定义为: $|F|^2 = FF^*$, 其中 F^* 为 F 之共轭复数。可以利用频谱的下列特性来做纹理分析:

(1) $|F|^2$ 值的分布反映了纹理的粗糙程度,在粗糙的纹理中, $|F|^2$ 值在当中有一个显著的高峰。若纹理较细腻,则 $|F|^2$ 值就比较分散,无显著高峰。因此可以选取一个环形区域,将 $|F|^2$ 用极坐标表示,取该区域上 $|F|^2$ 的平均值,计算 φ_r 。

$$\varphi_r = \int_0^{2\pi} F(r, \theta)^2 d\theta \quad (9)$$

r 为各种不同的值。若 r 小时, φ_r 很大; r 大时, φ_r 很小,则说明纹理是粗糙的。反之,若 r 变化对 φ_r 值影响不大,就说明纹理是比较细腻的。

(2) $|F|^2$ 值能反映纹理的方向性,若纹理中具有指向某一个方向 θ 的许多线条及边界条纹,则 $|F|^2$ 值在垂直于 θ 的方向上就有高值。因此,可以在某一个区域中计算 φ_θ 。

$$\varphi_\theta = \int_0^{\infty} |F(r, \theta)|^2 dr \quad (10)$$

其中 θ 为各种不同的数值。这是一个扇形区域,改变 θ 可知道 $|F|^2$ 有无方向性,从而检测出纹理是否有线条和边界条纹。从而根据这些纹理特征的匹配来进行图像的检索。

3 总结

除了上面介绍的通用 CBIR 系统所常用的技术手段以外,还有许多 CBIR 系统由于其图像库的内容的专用性

(下转第 195 页)

后一个值作为输出节点的期望值,这样可构成训练样本对 $(X_k, f_k) (k=1, 2, \dots, p)$, 设 $\theta = (\bar{f}, w_i, t_i, a_i, d_i)^T, (i=1, 2, \dots, N)$ 作为参数向量, 网络目标函数为 $E(\theta) = \frac{1}{2} \sum_{k=1}^p (\hat{X}(m+1) - f_k)^2$ 。用共轭梯度下降算法极小化上述目标函数得到相应的最优参数 θ , 迭代过程为: $\theta(k) = \theta(k-1) + \lambda_k \nabla E(\theta(k-1))$, 其中 $\lambda_k > 0$, 为学习步长, 而 $\nabla E(\theta) = [\frac{\partial E(\theta)}{\partial w_0}, \frac{\partial E(\theta)}{\partial w_k}, \frac{\partial E(\theta)}{\partial a_k}, \frac{\partial E(\theta)}{\partial d_k}]^T$ 。参数初始值选取是按随机梯度算法规定的方法, w_i 设置为 0, \bar{f} 取时间序列的平均值, 参数 a_i, d_i, t_i 选取采用计算样本重心方法, 即在 $[0, n]$ 上选择一个 p 点, 该点为 $[0, n]$ 区间的重力中心, 设 $t_1 = p, d_1 = \xi * n$, 一般取 $\xi = 0.5$, 这样点 p 把区间 $[0, n]$ 分成 2 部分, 在每一个子区间再重复上述过程, 直到 t_N 为止。

2 股市预测

从 <http://www.stock.com> 网站获取一些股票的历史交易数据进行实证分析, 选择的股票力求反映一定的代表性, 既有业绩稳定出色的绩优股如 600019(宝钢股份)、600177(雅戈尔), 也有流通量相对较小、股价起伏大、易被庄家操纵的高科技股如 600690(青岛海尔), 还一定程度上反映国家金融状况的金融类股如 000001(深发展)、600000(浦发银行)等, 用文中的小波神经网络方法, 以 2000 年 6 月 7 日到 2003 年 6 月 27 日的日收盘价为例进行预测。这里计算了预测均方差(MSE), 表 1 显示了文中方法与神经网络(ANN)、指数自回归(EAR)方法的比较, 利用文中方法估计出最小嵌入维数为 5, 根据实验取 12 个小波元, 小波函数选取 Meyer 小波, 这样小波神经网络的输入层数为 5, 隐层数为 12, 输出层数为 1。结果表明, 对一般的非线性股市时序, 训练次数、预测精度明显优于其它两种方法。

3 结论

给出了小波神经网络时间序列的预测模型, 它有效地

避免了普通神经网络预测模型中固有的缺陷。通过伸缩和平移因子构成标准正交基来生成小波级数, 使小波神经网络达到最佳的模型逼近效果, 从而达到最优拟合, 有效地提高了预测精度。将小波分析良好的局域化性质与神经网络的自适应学习和非线性映射能力相结合, 充分显示了小波神经网络极佳的函数逼近能力, 这为非线性时间序列流数据提供了新的思路和方法。

表 1 5 种股市时序的预测标准差的比较

股票代码	训练集(MSE)			测试集(MSE)		
	WNN	ANN	EAR	WNN	ANN	EAR
600000	0.0478	0.0727	0.0630	0.0075	0.0878	0.0249
600019	0.0259	0.0445	0.0388	0.0267	0.2165	0.0344
600177	0.1222	0.2538	0.2136	0.1266	0.5205	0.1379
000001	0.0371	0.0470	0.0491	0.0356	0.0787	0.0678
600690	0.0975	0.2767	0.2765	0.2763	0.4078	0.1987

参考文献:

- [1] 杨一文, 刘贵忠. 基于小波网络的非线性时间序列预测及其在股市中应用[J]. 模式识别与人工智能, 2001, 14(2): 213-247.
- [2] 刘志刚, 王晓茹. 小波变换、神经网络和小波网络的函数逼近能力分析比较[J]. 电力系统自动化, 2002, 26(20): 39-43.
- [3] 潘维民, 沈理. 基于神经网络的时间序列动态预测器的调整学习算法[J]. 电子学报, 1999, 27(11): 1-4.
- [4] Zhang Qinghua, Albert B. Wavelet Neural Networks[J]. IEEE Trans on Neural Networks, 2003, 3(6): 889-898.
- [5] Mallat S. A theory for multiresolution signal decomposition: the wavelet representation[J]. IEEE Trans on PAM I, 1989, 11(7): 674-693.
- [6] Hull J R. Neural Algorithm using Wavelets and Auto Regressive Inputs for System Identification[A]. IEEE International Conference on Networks[C]. [s. l.]: [s. n.], 1997. 724-727.

(上接第 192 页)

或其他原因, 往往对检索提出特殊要求, 此时就需要设计专门的检索方法来完成基于内容的检索。由于其约束信息不足, 在返回目标图像的同时往往会返回大量其他也满足此检索要求的图像^[5]。采用多个检索手段相结合的方法无疑可提供更多的约束而使得返回图像中目标图像的比率得到提高, 但检索手段间的融合是所要解决的问题。此外, 使系统在与操作者的检索交互中进行学习, 更好地理解检索的内容, 以及使检索性能更接近人类视觉的特性, 也是未来研究中所要解决的问题。

参考文献:

- [1] 陈纯. 计算机图像处理技术与算法[M]. 北京: 清华大学出版社, 2003.
- [2] Metal F. Query by Image Content: The QBIC System[J]. Computer, 1995(2): 23-32.
- [3] 容观澳. 计算机图像处理[M]. 北京: 清华大学出版社, 2000.
- [4] 章毓晋. 图像工程(上、下册)[M]. 北京: 清华大学出版社, 2000.
- [5] KSrihari R. Automatic Indexing and Content-Based Retrieval of Captioned Images[J]. Computer, 1995(1): 49-56.