高技术通讯 2010 年 第 20 卷 第 2 期:117 ~ 121

doi:10.3772/j.issn.1002-0470.2010.02.002

基于字符串匹配技术的图像检索算法①

赵 珊②*** 郑清洁*

(*河南理工大学计算机科学与技术学院 焦作 454000)

(**南京邮电大学江苏省图像处理与图像通信重点实验室 南京 210003)

摘 要 为提高图像检索的效果,提出了一种基于字符串匹配技术的图像检索算法。该算法根据人眼的视觉特性及方块编码(BTC)的原理首先对图像进行分割,构造对表征图像内容有意义的图像特征。在此基础上,根据字符出现的概率对字符表征意义的重要性,把图像特征动态映射成字符串形式,然后采用字符串匹配技术进行图像检索。该算法不仅利用了图像中的边缘及纹理分布,而且将字符串匹配技术引入到图像检索中,在提高检索率的同时又加快了检索速度。实验结果表明,该算法具有较高的检索效率。

关键词 基于内容的图像检索(CBIR),方块编码(BTC),字符串匹配,视觉特性

0 引 言

图像检索即基于内容的图像检索(content-based image retrieval, CBIR),是指利用图像自身所包含的丰富的视觉信息(如颜色、纹理、形状及空间分布等)来进行图像的检索[1,2]。图像检索是从文本信息检索技术发展而来的,由于图像和文本两者本身的特征间存在的差异性,文本检索技术中许多成功的索引技术并不能应用到图像检索中,为此,许多专家进行了深入的研究[3-5]。有效地将文本检索技术引入图像检索领域的关键在于如何定义图像中的"关键子块"。文献[3]在进行检索时,仅仅利用文本检索模型来确定视觉特征的权值,从严格意义上来说,这并没有利用文本索引技术。文献[4]则是采用特征矢量量化和聚类的方法定义图像的关键块,然后采用文档分析技术来实现图像检索,但由于该方法在定义图像关键块时使用的矢量量化方法不能保证各关键子块之间的独立性,并且生成码书的算法既复杂又费时,同时在提取图像关键子块时忽略了人眼的视觉特性,因而影响了最后的检索效果。

基于此,本文提出了一种基于字符串匹配技术的图像检索算法。该算法结合人眼的视觉特性构造对表征图像内容有意义的图像特征,根据图像描述符的特点将这些特征映射为字符串形式,然后采用字符串匹配技术进行图像检索。此算法不仅有效地提取了图像的内容特征,而且在进行匹配时充分考虑了字符串的结构信息对检索结果的影响。实验表明该算法具有较高的检索效率。

1 特征提取

对人类视觉系统而言,图像中具有不同活跃度(类别)的子块对图像间相似度的匹配所做的贡献是不同的,那些在图像中数量相对较少的边缘块、细节块(高活跃度块)对决定图像间相似与否,或者说对图像内容的贡献往往大于那些占有大部分区域的平滑块。如果两幅图像大多数的平滑块都比较相近,而数量相对较少的边缘块相差较大,由于人眼对边缘及条状结构比较敏感,所以不会认为二者相似。为了有效地提取图像的特征,本文借鉴方块编码(block truncation code, BTC)的思想[6]来进行特征提取。由于人的大脑皮层的感觉细胞感受的视觉信息是有方向性的,基于这种视觉模型,我们把图像分割成 $m \times m$ 的小块,根据分块变化的形式,设计一组对人眼有意义的模块。即用一个低频分量和若干个方向的高频分量的图像块来表示图像的平坦区域和各个方向的边缘,从而来表示分块内的灰度变化,以匹配视觉的生理机制,寻求对图像准确而有效的描述。

① 国家自然科学基金(50804013),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)和河南省教育厅科学基础研究基金(2008B520015,2009B520013)资助项目。

② 女,1975 年生,博士,副教授;研究方向:基于内容的图像检索,模式识别;联系人,E-mail: zhaoshan_9228@163.com

(收稿日期:2008-09-09)

假设 I 是一幅大小为 $M \times N$ 的彩色图像。首先采用公式

$$I = 0.3R + 0.59G + 0.11B \tag{1}$$

将其转化为灰度图像，然后再将图像分割为 $m \times m$ 大小的互不重叠的子块，对于每个子块，根据公式

$$\mu = \frac{\sum_{\forall i,j} p(i,j)}{m \times m} \tag{2}$$

$$\sigma = \frac{\sum_{\forall i,j} \| p(i,j) - \mu \|}{m \times m} \tag{3}$$

计算块内像素的灰度均值 μ 和平均色差 σ，其中 $p(i,j)$ 表示图像中位于 (i,j) 处像素的灰度值。然后根据方块编码的思想，在每个子块中，将灰度值大于均值 μ 的像素点赋值为 1，反之赋为 0，这样就得到了一系列 $m \times m$ 的二进制块。这些二进制块在一定程度上体现了图像中像素的灰度分布，相似的分布会产生相同的二进制块，在算法中定义这些二进制块为图像特征。部分特征的提取过程如图 1 所示（这里 $m = 2$）[5]。图 2 给出了采用图像特征表示的原始图像，可以看出，这些图像特征在一定程度上能很好地体现原图像的内容信息。

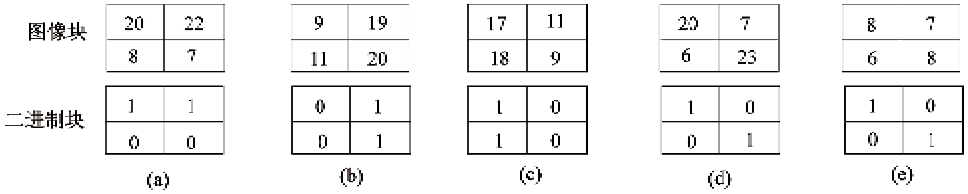


图 1 部分图像特征的提取

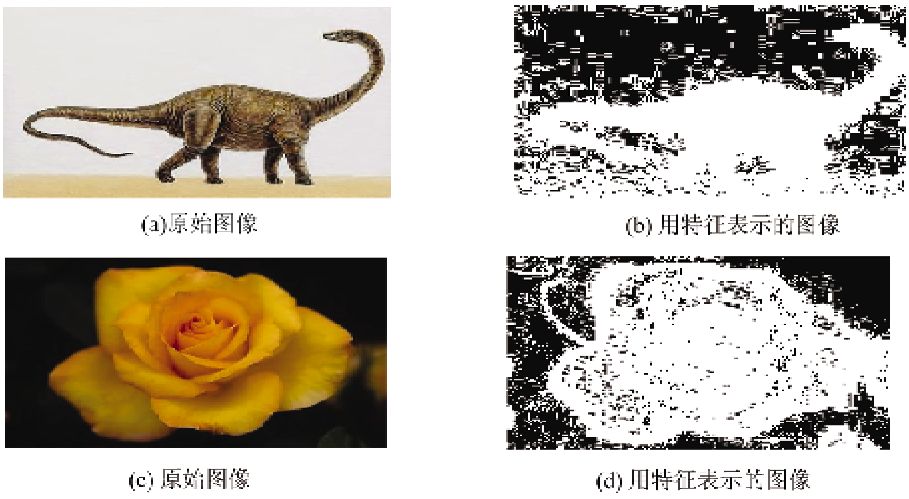


图 2 原始图像及相应的特征表示图像

在提取图像特征时，会出现如图 1 中(d)和(e)所示的情况，灰度相差很大的两个图像块提取的特征也可能会相同。因此，在算法中，设定了一个阈值 γ，当图像块的平均色差小于这个阈值时，就把这个块看作是均匀块，四个灰度像素全设为“0”。反之，就按上述的方法提取它的子特征。我们采用统计的方法，从图像库中随机选出 8 类图像，每类图像选出 3 幅，共组成 24 幅图像经过分割后选取某一种分布特征，统计其均值及方差的分布，在方差范围内经过 30 次实验，发现当这个阈值为整个图像平均色差的 0.0025 倍时可以取得满意的效果。在后面的实验中，我们取 $\gamma = 0.0025\sigma$，其中 σ 是图像的平均色差。

我们对每个特征块从左到右，从上到下进行扫描，以形成的二进制序列作为其索引值，然后采用一些常用的字符如 a、b、c 等来对应表示相应的子特征。算法中，我们用 a 来表示索引值为 0 的子特征块，即均匀块，b 来表示索引值为 1 的子特征块，依次类推。在此基础上，定义一个特征描述符

$$I = \{\{c_i, p_i\}, i = 1, \cdots, M\} \tag{4}$$

其中，M 是图像中总的子特征数目，c_i 是子特征相应的表示符，p_i 是子特征 c_i 在图像中所占的比例，且 $\sum p_i = 1$。针对每幅图像，将占比例最大的字符作为字符串中的第一个向量，其次是占比例第二的

子特征块向量,依次形成了一组字符串序列。这样我们就可以采用一组字符串来描述每幅图像的内容信息。实际上,从这个意义上来说,图像检索已经转化为基于字符串的匹配技术,计算图像间的相似度就转化为计算两个字符串的匹配程度。

上述算法在映射字符串时认为在图像中出现概率越大的关键特征越重要,然而实际上,如果关键特征在几乎所有的图像中都是反复出现的话,则其不具备将一幅图像与另一幅图像区分开来的能力。据此,我们在映射字符串序列时,以每个字符出现的概率以及是否频繁出现为依据,设定重要性权值

$$k_i = H_i/\log(p_i) \tag{5}$$

其中,p_i 的定义同式(4),H_i 为任意图像中索引值为 i 的关键子块的信息熵。这样,根据字符在表征图像内容的重要性来确定字符串序列,可以更好地体现图像特征。

2　相似性度量

将图像映射为字符串后,我们采用文本检索技术中的字符串匹配技术进行相似性度量。设 $S_1 = \{a_1, a_2, \cdots, a_m\}$,$S_2 = \{b_1, b_2, \cdots, b_n\}$ 表示两个字符串,Q_n 是特征数目相似程度,Q_s 表示特征相似程度,则根据相似学中测度相似性系统的相似度基本方法[7,8],两个字符串的相似度定义如下:

$$sim(S_1, S_2) = \alpha Q_n + \beta Q_s = \alpha \cdot \frac{K}{M + N + K} + \beta \cdot \sum_{i=1}^{K} \lambda_i q(s_i) \quad i \in [1, k] \tag{6}$$

其中 α, β 分别表示 Q_n 和 Q_s 对相似元素的整体相似度的互补性,且 $\alpha, \beta \in [0,1]$,$\alpha + \beta = 1$。K 表示字符串 S_1, S_2 间相似元素的数量,M, N 分别表示字符串 S_1, S_2 中的元素数量。λ_i 为反映相似元素对字符串相似度的影响程度的权值,$\lambda_i \in [0,1]$,且 $\sum_{i=1}^{K} \lambda_i = 1$,其定义如式

$$\lambda_i = \frac{i/\sum_{t=1}^{K} t + j/\sum_{l=1}^{K} l}{2} \tag{7}$$

所示。其中 i, j 分别表示相似元素 S_{ij} 为字符串 S_1 的第 i 个相似元,S_2 中的第 j 个相似元素。$q(s_i)$ 表示每一相似元素相似程度大小的相似元素值。

通常我们在只考虑字符串的字面特征时,也就是不考虑字符串中字符的位置信息对匹配结果的影响时,$q(s_i) = 1$,这时就会出现 $Sim(ABC, BCA) = 1$ 的匹配结果。但是,针对本文算法提取的特征来说,位置信息是很重要的一个因素。因此,在算法中,我们考虑到平移代价因素,引入字符串相似元素的平移距离来对 $q(s_i)$ 进行修正。定义

$$q(s_i) = \frac{1}{1 + |i - j|} \tag{8}$$

$|i - j|$ 表示字符串 S_1, S_2 左对齐或右对齐后,相似元素 s_{ij} 在字符串 S_1 的位置 i 与在字符串 S_2 的位置 j 的差的绝对值。

3　实验及讨论

为验证本文算法的有效性,我们在包含 3000 幅图像的通用图像库中进行了实验,图像库中包括了种类丰富的动物、建筑、自然景物、花卉、山脉等彩色图像。实验运行环境为 Windows XP 操作系统,系统配置为 PIV1.7GHz CPU,512MB 主存。采用"精确度(precision)"和"检索率(recall)[9]"作为相似检索的评价准则,进行了两组不同的对比实验。

首先,将本文算法同文献[4]和文献[5]的检索性能进行了对比实验。在图像库中随机选取 8 类图像,每类图像抽取 6 幅图像共组成 48 次查询,取这 48 次检索结果精确度和检索率的平均值作为算法的平均检索结果。实验中在利用式(6)进行相似性度量时,取 $\alpha = \beta = 0.5$,这样就赋予特征数目相似程度和特征相似程度同等重要性。图 3 给出了本文算法同文献[4]和文献[5]算法在"精确度"和"检索率"上的对比曲线。图 4 给出了其中的一次检索结果,其中每行最左边的一幅图像为示例图像,随后为检索结果图像队列,相似度从左向右按从大到小

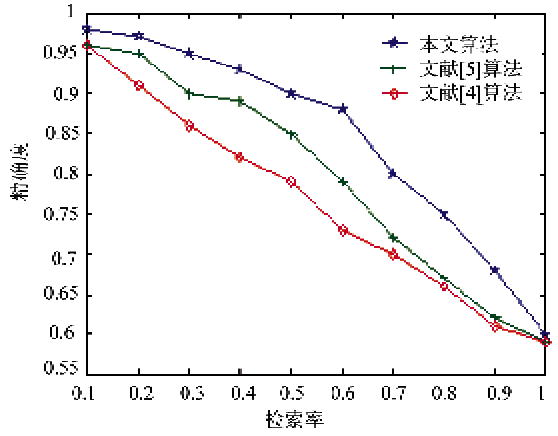
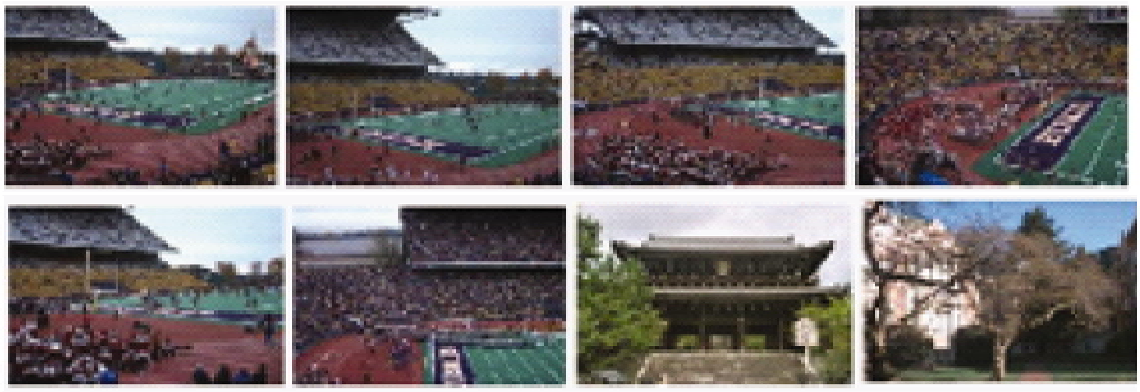


图 3　三种算法的检索率和精确度对比曲线

(a) 采用本文算法的实验结果

(b)采用文献[5]算法的实验结果



(c)采用文献[4]算法的实验结果

图 4 部分实验结果对比

顺序排列。由实验结果可以看出,本文算法的检索效率均优于其他两种算法,而且由于本文算法在提取图像特征时融合了人的视觉特性,检索结果更符合人类的视觉感受。同时,由于采用的字符串匹配方法中考虑了字符在字符串中的位置信息,因此和文献[5]采用直方图模型相比,检索效率要高。

为了检验本文算法的复杂度,在相同的软硬件环境下,从特征提取时间复杂度和图像检索时间复杂度两个方面将本文算法同其它两种方法进行了比较:(1) 特征提取时间复杂度:实验中在图像库中任意取 100 幅图像,利用本文算法提取特征所需的平均时间为 36.28s,采用文献[5]提取特征所需的平均时间为 48.12s。采用文献[4]算法提取特征所需的时间为 196.53s。(2)图像检索时间复杂度:实验采用从 1000 幅图像中检索出与例子图像相似的前 20 幅图像并显示,其中图像的索引特征已保存在不同的文件中。本文算法的平均检索时间为 2.36s,文献[5]的平均检索时间为 2.59s,文献[4]算法的平均检索时间为 3.17s。从实验结果可以看出,本文算法的特征提取时间复杂度和检索时间复杂度都小于其他