

图像检索 内容 关键词 信息检索系统 ②

计算机科学 1999 Vol. 26 No. 8

6-12

## 基于内容的图像检索的方法研究

On the Methods of Content Based Image Retrieval

李瑜 李磊

(中山大学软件所 广州 510275)

G354.4

**Abstract** Content Based Image Retrieval is the most active research area. This paper summarizes the important result reached by now in this area, and introduces some famous systems. At his childhood, CBIR has a lot of problems remained unsolved. Pointing out the default in the existing methods, this paper put over our ideas and suggested some future research directions.

**Keywords** Content based image query, High dimensional index, Feature extraction Similarity, Matching, Neural network

### 一、前言

近几年来,随着大容量存储器和数字化图像设备,诸如扫描仪、数字相机等技术的发展和普及应用,图像数据呈现级数的增长,由于无论是 WWW 还是企业单位部门都拥有大量的图像,因此从图像中提取有用的信息的前提是图像的有效检索,这也是数据转化为信息的关键一步。

图像检索的传统方法是基于文本的,使用关键字注释是最常用的方法,对图像的检索变成了对关键字的查找。这种方法可以利用传统的 DBMS 实现,简单易行,而且能够从用户角度表达图像内容的高层语义。但对于目前容量以 GB 或 TB 来计算的图像集合,要求对每一条记录进行注释是不可行的,同时,图像内容对于注释者的依赖性也局限了这种方法的有效性。除了关键字,图像的文件类型、大小、日期等附加信息也可以作为检索的辅助手段,但它们都不能反映图像的内容。

为了克服以上方法的局限性,90 年代初出现了基于内容的图像检索系统(CBIR—Content-based Image Retrieval)。CBIR 系统除了利用以上方法之外,主要是把图像的可视特征,例如颜色、纹理结构、轮廓、位置关系,作为图像的内容表进行匹配、查找。利用已有的算法,特征抽取和匹配完全可以由机器自动完成,这就克服了手工注释的低效和二义性。事实上,可视特征是对图像内容的压缩、抽象,与人类的观察十分相似,因此 CBIR 查询的结果能够满足

用户的一定要求。CBIR 在提出以后,得到了飞速的发展,在 WWW、图像数据库、医学等应用上得到研究,由于 CBIR 方法存在着不少问题,目前虽然已经有一些研究系统和商业系统,但功能和效率都有限,离真正的实用还有一定的距离。本文综述了 CBIR 的目前主要成果和未来重要研究方向,希望能够对 CBIR 的发展起到一定的促进作用。

### 二、特征抽取

特征是图像的压缩表示。特征抽取是 CBIR 系统的最基础部分,在很大程度上决定了 CBIR 系统的成败。目前,大量的 CBIR 的研究集中在特征抽取上,力求抽出与人类观察最相近的特征。

在不同的应用中,可以抽取不同的图像特征,例如对手写英文的识别系统可以利用表示笔画绝对曲率的特征量、表示笔画方向的特征量等。作为一个通用的图像检索系统,只能考虑通用的图像特征,例如颜色、纹理、轮廓等。本文将从通用的角度来介绍特征抽取方法。

#### 1. 颜色

颜色是图像最重要的特征。其中,色直方图(Color Histogram)是使用最广的特征。色直方图代表了三个颜色频道的强度的联合概率。色直方图是一个一维向量,向量的长度由用户定,例如,RGB 空间中,若划分为  $4 \times 4 \times 4$ ,则向量长度为 64。向量的值为图像中颜色属于该区段的像素点的个数。色直方图描述了图像的全局颜色分布,易于计算,对图像

中对对象的旋转以及观察位置的变化不敏感。但色直方图完全不考虑空间位置信息,因此,大的图像集合中相同色直方图的图很多,该方法只能应用在图像集合较小的情况,或者结合其他特征一起使用。

在文[1]中, Greg Pass 等提出了色相关向量(Color Coherence Vectors, CCV)方法。他们在色直方图基础上,将各颜色区段的像素点总数分为相关和不相关两类像素点的总数。相关的像素点是一大群相同颜色、相互临近的像素点。CCV 方法易于计算,而且包含了一定的空间信息。作者的实验证明,CCV 比色直方图的正确率提高 30%。

在文[2]中, Jing Huang 等提出了计算色相关图(Color Correlogram)的算法,色相关图是一张由颜色对组成的表,表的第  $k$  个入口是在颜色  $i$  的像素点的距离为  $k$  的范围内找到颜色  $j$  的像素点的概率。色相关图包含的空间位置信息比前两种方法要多,但对图像的旋转和位置变换相对敏感。

在文[3]中, John R. Smith and Shih-Fu Chang 提出了色集(Color Sets)的方法。他们先将 RGB 空间转化为 HSV 颜色空间,然后将 HSV 空间量化为一维空间,使用阈值过滤,大于阈值的为 1,小于阈值的为 0,该方法保留了最显著的颜色,而且这种二进制表示方便以后的索引构造。色集方法的空间信息最多,因此支持的查询种类也最为丰富。

在文[4]中, W. Y. Ma 等人使用 GLA 算法先将图像数据库中的颜色聚类,在预定义的误差范围内,将三维颜色空间量化为最小数目的一维,然后使用各颜色所占的比例来反映分割区域的颜色分布。这种方法在最低平均方差下聚类出最少的颜色数目,因此生成的特征向量是相对小的。

## 2. 纹理结构

纹理至今还没有一个标准定义,通常意义上的纹理是指图像像素灰度级或颜色的某种规律性的变化,而且这种变化是空间统计相关的。纹理反映了图像本身的属性,不同的物体具有明显不同的空间特征。

纹理的分析方法可分为统计方法和结构方法,统计方法主要有空间域和频率域方法。空间域方法中,最简单的直方图统计特征法由于只能反映一维的灰度变化,早已不用。在文[5], Haralick 和 Shanmugam 提出了共生(co-occurrence)矩阵法,共生矩阵定义为从灰度为  $i$  的像素点开始离开某个固定位置  $t=(DX, DY)$  的像素点的灰度为  $j$  的概率。从灰度共生矩阵中可以分析图像的纹理。共生矩阵法算

法简单,但共生矩阵通常较大,而且含有不少冗余信息。

在文[6]中, Tamura 等从视觉心理学角度提出了新的分析方法。Tamura 定义了粗糙、方向、规整等六种特征,这些特征都是相互独立而且具有可视性。与共生矩阵相比, Tamura 特性几乎没有冗余信息,效率相当高,因此 Tamura 方法在 CBIR 中应用较广。

基于频率域的方法主要是傅立叶变换和小波分析。图像在傅立叶变换后,其能量谱在一定程度上反映了图像的粗糙和方向性。90 年代以后,小波分析在纹理分析方面发展很快,出现了很多新的算法,在文[7]中, W. Y. Ma 分析了多种小波方法,结论是 GABOR 方法表现出最好的特征

## 3. 轮廓、空间位置关系和分割

图像的轮廓和空间位置关系都是在图像分割之后进行。图像分割的目的是对应用所需的物体进行抽取和描述,因此没有唯一的标准和方法。分割的原理是利用物体区域内部的某种同一性和均匀性进行划分。图像分割的经典算法有门限法、区域增长法等。目前较新的方法有空间聚类法,它主要利用待分割物体的统计特性,使用贝叶斯决策理论和聚类方法分割。

在对物体进行分割以后,可以得到物体的边界点的集合,把这些边界点进行傅立叶变换,可以生成一个复系数的集合。这些系数在频率域描述了物体的轮廓,低频系数代表了一般的轮廓属性,高频部分代表了轮廓的细节。目前还出现了有限元法、图灵函数、小波描述等。

一般的 CBIR 系统对空间位置关系不作太多的研究,因为大多数情况下,人们会把一幅猫在主人的左脚和另一幅猫在右脚的图当作一样。但在一些特殊的应用中,位置关系是很重要的。较多的分析方法是基于颜色和纹理的分析。主要成果有 Rickman 等的颜色元组直方图, Smith 的色反向投影, Quad-tree(二叉树), 和 Huang 等的色相关图。

小结 图像特征抽取方法主要在计算机视觉和图像模式识别两个领域进行研究。在 CBIR 系统出现以后,特征抽取的方法得到重新的审视,以上研究成果中,颜色特征和空间位置关系是来源于 CBIR 系统。而纹理和轮廓的方法主要是计算机视觉的成果,较多的小波方法是在 CBIR 系统中得到增强。总之, CBIR 系统根据自己的需要,采用不同的特性,因此,查询的效率和效果也很不同。

### 三、图像匹配

#### 1. 距离度量方法

图像特性抽取之后,很直观的方法是直接使用特征向量的距离来衡量两幅图的相似性。然而,不同距离度量方法对不同的特征抽取模型具有不同的性能。在文[8]中,W. Y. Ma 和 B. S. Manjunath 的实验显示欧几里德距离和加权欧几里德距离使不同的特征集具有较大的性能差异。例如,Mahalanobis 距离对马尔可夫模型最适合,而 Gabor 特征对不加权的距离公式最好。以下介绍目前在 CBIR 系统中较为常用的距离公式。

##### • Minkowsky 距离

$$D(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^r \right)^{\frac{1}{r}}$$

这是若干种距离公式的通式表示,以下的 Manhattan 距离和欧几里德距离都是它的特例。

##### • Manhattan 距离

$$D(x, y) = \sum_{i=1}^n |x_i - y_i|$$

这是 Minkowsky 距离当  $r=1$  时的特例。

• 欧几里德距离。两个特征向量的欧几里德距离定义为:

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

欧几里德距离是最简单的距离公式,也是在 CBIR 中应用较广的距离公式。然而欧几里德公式完全不考虑向量各维之间的关系,而且各维必须是同等重要的,这就大大影响其使用范围和有效性。欧几里德距离可以单独用于色直方图和纹理特性。

##### • 加权欧几里德距离

$$D(x, y) = \sqrt{\sum_{i=1}^n w_i (x_i - y_i)^2}$$

加权欧几里德距离考虑了不同维之间的不同重要性。由于 CBIR 系统通常抽取了大量的特征,不同特征的重要性都是不同的,因此加权欧几里德距离在 CBIR 系统中应用很广,可单独用于色直方图、纹理、形状或它们的联合向量。

##### • Mahalanobis 距离

$$D(x, y) = \sum_{i=0}^n \sum_{j=0}^n (x[i] - y[i]) \cdot a_{i,j} (x[j] - y[j])$$

A 矩阵是相应的协方差矩阵。马氏距离注意到样品的统计特性,排除了样品间的相关性影响,是较为常用的距离。

##### • 二次距离

$$D(x, y) = \sum_{i=0}^n \sum_{j=0}^n (x[i] - y[i]) \cdot a_{i,j} (x[j] - y[j])$$

该距离类似于马氏距离,它是 QBIC 系统为计算色直方图的相似性而提出的,其中  $A_{i,j}$  是特征  $i$  和特征  $j$  之间的相似因子,由系统在计算前预定义。该距离由于考虑了不同特征之间的关联,是比较准确的度量,同时也带来大的计算量。

##### • 直方图交(Histogram Intersection)

$$D(x, y) = \frac{\sum_{i=0}^n \min(x[i], y[i])}{\min(\sum_{i=0}^n x[i], \sum_{i=0}^n y[i])}$$

该距离只能用于直方图,当纹理用直方图表示时,也适合使用该公式。另外,Chamfer 匹配是轮廓匹配中常用的距离度量。

使用向量距离作为相似性度量是最简单和直观的方法,但缺乏严格的证明。事实上,距离公式来源于四条计量公理的限制:

•  $D(A, A) = D(B, B)$  自相似常数公理。

•  $D(A, B) \geq D(A, A)$  最小公理。

•  $D(A, B) = D(B, A)$  对称性公理。

•  $D(A, B) + D(B, C) \geq D(A, C)$  三角不等公理。

然而,这四条公理并不是真正的公理。实验证明,这四条公理都有反例<sup>[9]</sup>。例如,在认知实验中,人们容易把某特征不那么显著的物体认为象特征显著的物体,却不认为显著的物体象不显著的物体,这就违反了对称性公理。因此,使用由它们衍生的距离公式作为相似性度量是有缺陷的。

#### 2. 神经网络学习相似性

在文[10]中,W. Y. Ma 和 B. S. Manjunath 意识到欧几里德距离在 CBIR 中的缺陷,正如它在模式识别中的缺陷一样,几何相近的模式很可能属于不同类别,而几何相近的向量可能是由于噪音引起,它们是不相似的。于是他们提出了使用神经网络学习相似性。

W. Y. Ma 和 B. S. Manjunath 先使用无监督学习的自组织特征映射网(Kohonen Map)将图像分类,然后再用监督学习的向量量化(Learning Vector Quantization)调整决策边界。经过学习,同类的图像在视觉上比不同类的图像更加相似。

#### 3. 人类的视觉相似模型

70 年前,心理学家已经开始研究相似的概念。上述四条计量公理的提出正是心理学的某些结果形

式化。但 Tyersky 发现了这四条公理并不真正成立, 于是, 他提出了特征对照模型 FCM (Feature Contrast Model)<sup>[1-3]</sup>。Tyersky 将对象看成特性的集合, 而不是计量空间的一点, 通过集合运算得到相似性, FCM:

$$S(a, b) = f(A \cap B) - \alpha \cdot f(A - B) - \beta \cdot f(B - A)$$

其中,  $S(a, b)$  是  $a$  和  $b$  两幅图的相似函数,  $A$  与  $B$  分别是  $a, b$  的特征集,  $f$  是某个非负函数。这个模型可以合理地解释计量公理所碰到的所有问题。

由于 FCM 的特征值用 0 (不存在某特性) 和 1 (存在) 表示, 因此在实际应用中很不方便。于是 Simone Santini 和 Ramesh Jain 提出了模糊特征对照 FFC 模型 (Fuzzy Feature Contrast)<sup>[3]</sup>。他们引入模糊运算, 于是两幅图的相似性计算如下:

$$S(a, b) = \sum_{i=1}^p \min\{\mu_i(a), \mu_i(b)\} - \alpha \cdot \sum_{i=1}^p \max\{\mu_i(a) - \mu_i(b), 0\} - \beta \cdot \sum_{i=1}^p \max\{\mu_i(b) - \mu_i(a), 0\}$$

在 FFC 中, 图像的特征向量的各维值是模糊值, 例如, 纹理的亮度 (对应谓词: 该纹理是亮的), 纹理的垂直分布 (对应谓词: 该纹理呈现垂直方向排列)。作者证明了只要各个特征取值范围有限, FCM 具有 FCC 的所有特性。作者的大量实验表明 FCM 比传统的距离度量更为接近人类对相似性衡量的心理上的特点。

**小结** 从以上研究我们可以得出以下结论: ①关于图像相似性的度量不只一个模型, 但没有一个是最好的, 主要原因是相似性具有特征依赖的特点。不同的特征, 具有不同的度量。特征空间本身的特点应该决定了其度量模型, 因此, 对特征空间本身的研究是图像相似性必须重视的。②计量公理已经不能称之为公理, 但由此还不能否定距离模型的有效性, 只能对它们的有效性进行进一步的考察。③对图像相似性的研究是 CBIR 必须解决的课题。

## 四、特征索引

目前, 大量的研究集中在特征抽取和选用上, 而很少关于图像查询的索引模式的探讨。事实上, 没有一个有效的索引模式, 所有的查询方法都变得无效, 用户不会忍受一个很慢的查询系统, 因为查询系统的目标就是为用户迅速找到所需的图像。当数据库

容量大时, 该问题更为突出, 而且, 随着特征维数的增加, 所有的索引方法都面临着陷于顺序查找的困境。

### 1. 传统的高维索引方法

由于使用多维的特征向量表示, 图像索引可以看成是高维索引的一种。传统的数据库技术中有三种比较流行的高维索引方法: R-树、线性四叉树、栅格文件。它们最初提出是在 GIS 和 CAD 上, 作为矩形、点等几何对象的索引。基本思想都是使用几何分区的方法将多维向量空间划分超平面, 按照几何覆盖的关系组织为树, 查询时只须搜索几个相关的分区。

四叉树和栅格文件是早期的多维索引方法。一个二维的四叉树的非叶节点包含 4 个子节点, 叶节点指向一个区域或点。四叉树使用递归的方法将整个二维空间不断划分。一个  $n$  维的四叉树的非叶节点有 2 的  $n$  次方个子节点。因此, 当维数增加时, 会出现节点数目爆炸的情况。栅格文件将数据空间分区为格块 (grid block), 桶是包含记录的存储单元, 一个桶可以被几个格块共享。另有一个栅格字典, 包含着两个  $n$  维的数组维护格块和记录空间以及桶之间的对应。栅格文件同样面临着栅格字典随维数增大而增长爆炸的情形。

R-树及其变体是最为有效的多维索引方法, 具有以下优点: ①是一种类似于 B-树的动态树结构, 能够维持平衡, 所有的叶节点都出现在同一层。②达到 50% 的节点利用率。R-树是利用多维空间在几何意义上的覆盖、重叠等概念来构造一个树。每个节点都用一个多维矩形表示, 如三维:  $(X_{min}, X_{max}, Y_{min}, Y_{max}, Z_{min}, Z_{max})$ 。父节点矩形包含所有子节点的矩形。插入一个新的向量有可能导致父节点的分裂, 删除一个记录可能导致父节点的合并, 搜索是按照覆盖关系由上往下直到找到重叠的矩形。

实验证明, 当向量的维数在 2~20 时, R-树的效率很好。当维数超过 20 时, 以上方法的效率都迅速降低, 几乎和顺序查找差不多。另外, 由于 R-树的构造是基于几何意义上的覆盖关系, R-树只限于欧几里德距离作为相似性度量的查询。

### 2. 多维空间的转换

由于图像的特征向量的维数很大, 通常会超过 50 或达到 100, 为了能够有效地利用以上的索引方法, 必须将图像特征的  $n$  维空间转换为 20 以内的  $m$  维。较常用的方法有 Karhunen-Loeve (K-L) 变换, 离散傅立叶变换 (DFT) 和高散余弦变换 (DCT)。可以

证明<sup>[12]</sup>,以上算法都能保证基于范围的查询不会错失正确的记录。但维数的降低不可避免会带来信息的丢失,因此查询结果中会有较多的错误记录。

### 3. 神经网络和簇聚在索引上的应用

针对以上索引方法的局限性,文[13]中使用自组织特征映射网(SOM)来构造图像索引。Zhang 和 Zhong 先用 SOM 将所有图像分类,然后不断合并距离最小的拓扑相邻两节点,直到所有几何相邻的类间距离大于预定的阈值或类的数目大于预定的阈值,最后,以 SOM 图为底,使用投影方法构造出一个有层次的树。搜索时从顶层开始,一直搜到最低层,那里维护着一个搜索列表,里面是属于该类的所有记录的指针。与全局搜索比较,SOM 索引方法大大减少了搜索的次数。这种索引方法还具有以下好处:①通过无监督的学习将特征空间分类;②相似的种类拓扑相近,适合最近邻居搜索;③适合于具有大量的类别和小量的模式的情况,这种情况在图像数据库中很典型。

在文[14]中,Moses Charikar 等提出了通过使用递增式聚类方法,可以将高维的数据分类,并能使用非欧几里德距离作为相似性度量。

## 五、用户反馈

由于 CBIR 系统使用的特征是很低层的,而用户的需求是很高层的,两者之间具有一定的鸿沟。例如,用户提供的查询范例是身穿黄色队服的巴西球星罗纳尔多,系统找到的是—棵叶子变黄的秋天的树木。因此,CBIR 系统必须具有从用户的反馈中学习的能力,使系统的查询结果逼近用户的要求。

CBIR 系统有效的前提是用户的高层需求可以映射在图像的低层属性上。因此,不同的特征反映了用户不同的需求。例如,植物学家对图的纹理有兴趣,而画家关注图的颜色。因此,学习用户的反馈就是要调整不同特征的权值。

在文[15]中,Rui Yong 等将用户反馈分为 5 个等级:高度相关 3 分,相关 1 分,没意见 0 分,不相关 -1 分,很不相关 -3 分。K 个查询结果分别有一个等级分,然后单独按照各个特征来查出 K 个结果,有属于前面 K 个结果的就把结果的等级分加在该特征的权值上。最后把权值归一化。实验表明,该方法能快速地找到与用户目标相关的图。

Huang Jing 等提出两种学习方法。第一种也是学习权值,Huang Jing 把结果分为相关和不相关,对查询例子和相关结果之间距离贡献越大的特征越不

重要,公式如下:

$$\omega_{i,j} = \omega_{i,j} \cdot (1 + \delta - \delta_{i,j})$$

$$\omega_{i,j} = \omega_{i,j} \cdot (1 - \delta + \delta_{i,j})$$

其中第一条公式是处理相关的例子,第二条是处理不相关的例子。

第二种方法是学习查询向量,很自然的想法是将查询向量向相关例子靠近,向不相关的例子偏离。由于不相关的例子在特征空间中是分散的,因此将查询向量向不相关例子偏离很难实现。下面只考虑相关例子,公式如下:

$$A = (1 - \beta) \cdot \alpha(Q) + \beta \cdot \frac{1}{|F|} \sum_{i=1}^{|F|} \alpha(F_i)$$

其中 Q 为查询向量,  $F_i$  为反馈为相关的图像的特征向量,  $\beta$  用于控制反馈的速度。将 A 作为新的查询向量重新查找,该方法类似于 WWW 中的搜索引擎。

## 六、已有的 CBIR 系统

从 90 年代中期开始,CBIR 逐渐成为活跃的研究领域。到目前为止,已有不少的研究和商业系统。这些系统的框架和查询的处理过程很相似,只是采用的特征和搜索的方法各具特色,性能也互有差异,然而,没有一个是最好的,因为这些系统都还有许多共同的弱点。下面介绍一些比较有代表性的 CBIR 系统。

1. QBIC(Query By Image Content) 是由 IBM 公司推出的最早的 CBIR 系统。QBIC 使用  $K=64$  的色直方图。纹理是改进的 Tamura 模型,包含粗糙度、对比度、方向。轮廓特征有圆形、奇异度等。QBIC 使用 K-L 变换来降低特征维数,再用  $R^*$ -树构造索引树。QBIC 系统同时支持字符和内容检索。

2. VISUALSEEK 是哥伦比亚大学开发的。它使用颜色集和由小波分析提取的纹理特征。他们发明了二进树(Binary Tree)算法构造索引树。VISUALSEEK 支持基于空间位置关系的查询。

3. VIRAGE 由 Virage 公司研制。Virage 的特征集包含颜色、颜色布局、纹理和结构。用户可以调整各个特征的权值。

4. PHOTOBOK 是由麻省理工学院的媒体实验室开发。它包含轮廓、纹理、脸部特征三本书(BOOK),用户可以分别查询这三本书。为了把人的能动性加进系统,PHOTOBOK 提出了一种“SOCIETY OF MODEL”的方法,该方法根据用户的反馈动态采用不同的模型。

**5. IMAGEFINDER** 是 Attrasoft 公司开发的基于神经网络的图像搜索系统。它需要用户提供一组以备查询的图像集合对波尔兹曼机进行训练,然后把要查的图像训练,最后系统给图像集合按相似度打分。

**6. MARS (Multimedia Analysis and Retrieval System)** 是由伊利诺斯大学开发。MARS 系统从计算机视觉、DBMS、信息查询三个领域进行研究。MARS 系统提出一个相关反馈体系结构,该结构从用户的反馈对系统的参数进行了调整,以求满足用户的需求。

## 七、未来研究方向

从以上讨论可以看出, CBIR 还处于初级阶段, CBIR 的各个步骤都有不少算法,但这些算法都不能令人真正满意。因此, CBIR 中很多领域需要研究,主要如下:

**1. 图像理解** 目前的特征抽取都是计算机视觉已有的方法或其延伸,抽取出来的特征是很底层的,而人类观察图像后能对图像进行语义上的抽象,把低层特征直接作为高层语义的代表。由于以下原因会出现问题:1)低层特征的不完整性。除了颜色、纹理、轮廓、空间位置外,图像是否还有其他特征未发现?即使是已知的纹理,还没有一个标准的定义。2)低层特征与高级语义不一一对应。一块草地,可能和从空中拍摄的森林的顶部的颜色、纹理和形状都相似。因此,图像理解的研究是 CBIR 不能忽视的。

**2. 有效的高维索引** 四叉树、栅格文件和 R-树族都只限于 20 维以内的索引,神经网络 SOM 的方法具有静态的特征;图像数据必须在聚类前已知,而且类别数是固定的,不能动态增长。这对于大量的图像的插入和删除会引起问题。根据 ART 网络动态增长的特点,可以考虑使用 ART 作为聚类的工具,构造索引。由于有效的索引是 CBIR 真正实用的关键,对它的研究迫在眉睫。

**3. 标准测试集和性能标准** 目前, CBIR 还没有一个统一的测试图像集合,各类系统使用的集合不相同,必然会影响对该系统的性能的正确评价。作为标准测试集,最重要的是考虑样本的随机性,各个领域、各种色彩、各种大小的图像出现的概率要相同。

性能标准使用最多的是精度和回想率。精度是查询结果中,相关的数目占总数的比例。回想率则是

查询结果的相关数目除以整个图像集合中的相关数目,由于没有对于“相关”的标准衡量,目前都是以人的目测为准,这就带来了很大的主观性。

**4. 人类的视觉相似模型** CBIR 系统大多简单地使用欧几里德距离作为图像的相似性的度量。与模式识别一样,其缺陷显而易见,相似的图在特征空间中不一定几何相近,这已是共识,如何衡量两幅图的相似,从心理学角度对人类的视觉模型进行研究是一个很好的方向。Tyersky 的 FCC 模型已经显示了一定的效用。

**5. 真正的图像数据库** CBIR 的主要应用是 WWW 和图像数据库。目前很多的 CBIR 系统还只是一个简单的图像文件管理和查询系统,并不支持真正的图像数据共享。与传统的数据库相比,图像数据库具有以下特点:

- 数据库模式常变,而字段体不变。新的字段代表着新的特征抽取方法,它会随技术的发展和应用的变化而不断出现,而只要图像不变,它的旧的字段值就不会变化。

- 在模式定义时,特征的选用可以有 DBA 根据应用需要指定,允许特征抽取方法的定制。

- 把图像作为一个复合字段,而不是作为一个记录,图像的各个特征相对独立,但又必须同时出现。图像数据库应该维持这种关系的完整性。

**结束语** CBIR 是计算机视觉、DBMS、神经网络、模式识别和机器理解等领域的技术的综合应用。这些领域都在发展当中,因此,本文不可能覆盖所有的成果,只能对 CBIR 一些比较常用的方法进行探讨。由以上内容可以推断, CBIR 还处于幼年时期,但它的前景是诱人的。

## 参考文献

- 1 Pass G, et al. Comparing images using color coherence vectors. In: Proc. ACM Conf. On Multimedia, 1996
- 2 Huang J, et al. Image indexing using color correlograms. In: Proc. of IEEE Conf. On Computer Vision and Pattern Recognition, 1997
- 3 Smith J R, Chang S F. Local color and texture extraction and spatial query. In: Proc IEEE Int. Conf. on Image Proc. 1996
- 4 Ma W Y, et al. Tools for texture/color based search of images. In: SPIE Int. Conf. 3106, Human Vision and Electronic Imaging, 1997
- 5 Haralick R M, Shanmugam K. Texture Features for Image Classification. IEEE Trans. On Sys, Man, and

- Cyb.1973,SMC-3(6)
- 6 Tamura H, et al. Texture features corresponding to visual perception. *IEEE Trans. on Sys, Man, and Cyb.* 1978,SMC-8(6)
  - 7 Ma W Y, Manjunath B S. A comparison of wavelet transform features for texture image annotation. In *Proc. IEEE Int. Conf on Image Proc.* 1994
  - 8 Ma W Y, Manjunath B S. Texture Features for browse and retrieve of image; [Technical Report 95-06] University of California at Santa Barbara, 1995
  - 9 Santini S, Jain R. Similarity Match. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1996
  - 10 Manjunath B S, Ma W Y. Image Indexing Using a Texture Dictionary. In: *Proc of SPIE Conf. on Image Storage and Archiving Systems*, volume 2606
  - 11 Tversky A. Features of similarity. *Psychological review*, 1997, 84(4): 327~352
  - 12 Niu Youping, et al. A study of image indexing techniques for multimedia database systems: [Technical Report TR95-19]. 1995
  - 13 Zhang H, Zhong D. A Scheme for Visual Feature Based Image Indexing. *Storage & Retrieval for Image and Video Database*. SPIE, 1993
  - 14 Charickar M, et al. Incremental clustering and dynamic information retrieval. In: *Proc. of the 29th Annual ACM Symposium on Theory of Computing*, 1997. 626~635
  - 15 Rui Y, et al. Relevance FeedBack. A Power Tool for Interactive Content-Based Image Retrieval. *IEEE Trans. on Circuits and Systems for Video Technology*, Special Issue on Interactive Multimedia Systems for the Internet, 1998

(上接第5页)

为标签栈技术是针对特定路由的,它也不能用于路由总结节点,否则将破坏层次结构,将交换带来的益处抵消。

• MPLS 虽然支持 QoS,但是在现实中目前只有 IP 这种“尽力传输”业务, QoS 实现的效果不可知。

• MPLS 虽然定义了回路预防/检测,但何时要求回路预防/检测,是否在所有 LSR 上都实行,仍有待进一步研究。

**结束语** MPLS 是当今路由技术的发展新方向, MPLS 将路由选择与数据转发分离,既吸收了第二层交换的优点(快速高效、提供 QoS 服务),又保留了传统 IP 寻径的灵活性,是解决因特网所面临的一系列问题(伸缩性,向综合业务网的演变等)的一种优秀的解决方案,随着标准化的完成,是 IP 与 ATM 结合技术的发展方向,必将极大地推动当今的因特网的发展。

## 参 考 文 献

- 1 Rosen E, et al. Multiprotocol Label Switching Architecture. Internet draft, July 1998
- 2 Thomas B, et al. LDP Specification. Internet draft, Aug 1998
- 3 Davie B, et al. Use of Label Switching With ATM. Internet draft, July 1998
- 4 Ayandeh S, Fan Y. MPLS Routing Dynamics. Internet draft, Mar 1998
- 5 Awduche D, et al. Extensions to RSVP for Traffic Engineering. Internet draft, Aug 1998
- 6 Rosen E, et al. MPLS Label Stack Encoding. Internet draft, July 1998
- 7 Gray E. Issues With MPLS Architecture. Internet draft, Mar 1998
- 8 Fenner W. Internet Group Management Protocol (version 2). RFC 2236, Nov 1997
- 9 Davie B, et al. Use of Label Switching with RSVP. Internet draft, Nov 1997
- 10 张宏科,等. ATM 网络互连原理与工程. 清华大学出版社, 1996