

基于独立关键子块和三角树的 快速图象检索新方法

赵雪雁 庄越挺 吴飞 刘骏伟

(浙江大学计算机学院, 浙大 微软视觉感知联合实验室, 杭州 310027)

摘要 由于图象存储数据量非常大, 因此提取图象特征和检索极为耗时. 为了提高图象检索效率, 将文本检索中的有效检索方法(基于关键字频率与关键字逆文档频率乘积的索引模型)结合三角树索引机制应用到基于内容的图象检索, 提出了一种基于独立关键子块和三角树的快速图象检索新方法. 该方法首先用独立分量分析将样本图象子块中的直方图特征映射到色彩概念空间来得到类似于文本中关键字的独立关键子块; 然后再用训练好的模糊支持向量机去识别每幅图象中所包含的独立关键子块, 由于独立分量分析能够使特征彼此保持高阶独立性, 因此该方法与主成分分析方法对比, 具有较高检索效率; 最后, 再通过构造三角树来为图象数据库建立分层索引结构, 以加快检索速度.

关键词 计算机信息管理系统(520·6070) 独立分量分析 模糊支持向量机 独立关键子块 主成分分析 三角树

中图分类号: TP391.3 **文献标识码:** A **文章编号:** 1006-8961(2003)11-1327-07

Fast Image Retrieval Method based on Independent Keyblock and Triangle Tree

ZHAO Xue-yan, ZHUANG Yue-ting, WU Fei, LIU Jun-wei

(Department of Computer Science and Engineering, Zhejiang University,
Microsoft Visual Perception Laboratory of Zhejiang University, Hangzhou 310027)

Abstract Because image database is very huge, the feature extraction and retrieval process are usually time consuming. In order to effectively use existing text information retrieval methods in content based image retrieval, especially the index mechanism of the product of $tf * idf$ by term frequency (tf) and inverse document frequency (idf) for each text document, this paper cooperates $tf * idf$ model with triangle tree to improve the retrieval performance. First, after pixel-based histogram features of sub-block in certain image class are mapped to color concept space through independent component analysis (ICA), we would obtain all of independent keyblock of such image class; then well-trained fuzzy support vector machine is used to recognize all of independent keyblocks contained by each image. Similar to text retrieval, in which the whole text document is indexed by keywords, the recognized independent keyblock is used to index each image in database. Because independent component features are naturally high order independent with each other, compared to principle component analysis (PCA) method, this algorithm achieves higher performance. At last, triangle tree is used to hierarchically index image database and thereof speed up retrieval.

Keyword Independent component analysis, Fuzzy support vector machine, Independent keyblock, Principle component analysis, Triangle Tree

0 前言

图象检索随着多媒体技术和计算机网络的飞速

发展已成为一个非常活跃的研究领域. 由于早期基于文本的图象检索(text-based image retrieval)存在着人工标注费时、费力, 而且主观理解不统一等问题, 因此 20 世纪 90 年代初期, 基于内容的图象检索

基金项目: 国家自然科学基金资助项目(60272031); 教育部博士点科研基金项目(20010333049);

国家“十五”重大科技攻关项目(2001BA101A07-03); 浙江省科技计划项目重点科研项目(2003C21010)

收稿日期: 2002-05-20; 改回日期: 2003-06-13

技术开始兴起^[1]。这种图象检索系统是将自动提取得到的每幅图象视觉特征作为索引,然后根据用户提交的一幅或多幅例子图象来查找在视觉特征上相似的图象。如今已经开发出一些原型系统,如 QBIC^[2], VisualSeek^[3], PhotoBook^[4]和 Virage^[5]等,但是由于图象的存储数据量大,从而使得提取图象特征和检索很耗时。本文结合文本关键字检索中的有效方法和三角树来加快图象检索的速度。

虽然图象检索是从文本检索发展而来的,但文本检索中许多成功技术仍不能有效应用到图象检索中。如采用关键字频率(term frequency)与关键字字逆文档频率(inverse document frequency)积的关键字索引匹配模型(简称 tf×idf 模型),该模型是首先通过对文本进行全文分析^[6]来提取关键字,再采用关键字频率 F_{term} 和关键字逆文档频率 F_{id} 的积 $F_{term} \times F_{id}$ 来为文档建立索引的检索方法^[7]。由于该模型在文本检索中的成功应用,tf×idf 模型方法已被应用到图象检索领域:如文献[8]首先使用特征向量量化和聚类来得到图象中的关键块,并将图象看成文档,然后使用文档分析技术来实现图象检索;文献[9]使用 tf×idf 模型来为视觉特征建立权值,以便进行相关反馈检索。

将文本检索技术引入图象检索领域,难点在于如何定义图象中的“关键子块”,并使子块之间保持彼此独立性。但由于文献[8]使用向量量化方法不能保证关键子块之间的独立性,而文献[9]也没有定义关键子块,因此文献[10]将主成分分析(Principle Component Analysis, PCA)技术应用在人脸识别中,即通过构造“特征人脸(Eigenface)”来将图象从像素空间转换到人脸空间,从而实现了较高的识别正确率,并将其成功应用到其他模式识别领域。本文首先采用 PCA 技术以及在此基础上能保持更高阶独立的独立分量分析(Independent Component Analysis, ICA)方法来生成某类图象的独立关键子块空间,并对两者做出比较;然后用学习机将独立关键子块识别出来,再使用类似 tf×idf 模型的文本检索方法进行图象检索。虽然利用关键子块检索可以使提取图象特征变得容易简单,但是如果检索任一幅图象都要计算其与数据库中每幅图象的相似度,则将产生不必要的冗余;为此,最后通过构造三角树来加快检索速度和降低复杂度。

在为每个独立关键子块构造学习分类机时,为保证识别的准确率要考虑 3 点因素:(1)如何表征独立关键子块,由于提取的特征(如颜色、纹理和形状

等)向量之间要尽可能相互独立,因此太多的相关性信息会造成识别分类性能稳定性差^[11];(2)使用的样本数目要适中,因为过多将导致过学习问题,若太少,则又收集信息不足;(3)学习机要有良好的学习推广能力,即能够对测试样本和实际未知数据都达到良好的分类目的。

1 支持向量机原理

支持向量机起源于统计学习理论,它是研究如何通过构造学习机来实现模式分类问题^[12]。支持向量机使用结构风险最小化(Structural Risk Minimization, SRM 准则)原理来构造决策超平面,以使每一类数据之间的分类间隔(Margin)最大。SRM 准则认为:学习机对未知数据分类所产生的实际风险是由如下两部分组成的,且满足如下关系

$$R \leq R_{emp} + \sqrt{\frac{h(\log(2n/h)+1) - \log(\eta/4)}{n}}$$

其中, R 是实际风险,右边叫做风险边界, R_{emp} 称为

经验风险, $\sqrt{\frac{h(\log(2n/h)+1) - \log(\eta/4)}{n}}$ 叫做“VC 置信值”, n 是训练样本个数, h 是学习机的 VC 维,它反映了学习机的复杂程度, η 是置信水平,其思想就是在样本数目适宜的前提下,选取比较好的 VC 维 h , 使经验风险 R_{emp} 和置信值达到一个折中,最终使实际风险 R 变小,以便在样本数目适宜的前提下,取得实际最好分类效果,其在语音识别、视频字幕提取和人脸识别等模式分类领域已取得了成功^[13,14]。在支持向量机中,存在线性可分和线性不可分两种情况,但在现实世界中,由于很多问题是非线性可分的,而根据 Cover 定理,一个复杂的模式识别分类问题,在高维空间比低维空间更容易线性可分^[15],因此要首先通过核函数把训练样本中的低维数据映射到高维特征空间,然后才能在高维空间构造一个最佳分类平面。由于构造的核函数满足 Mercer 条件,所以在训练中只需考虑核函数 K , 而不必明确知道低维向高维的映射函数 Φ ^[16,17]。

2 图象独立关键子块空间形成

独立分量分析起源于盲源分离问题(Blind Source Separation, BSS),它与主成分分析和奇异值分解(Singular Value Decomposition, SVD)均属

于线性变换,但是后者只能按能量大小对数据进行分解,以消除数据之间的二阶相关性,而 ICA 则能够消除输入数据的高阶相关性^[18]。在图象、视频和声音识别检索等应用中,特征之间虽存在相关性,但是由于重要特性一般隐藏在高阶统计特性中,因此使用 ICA 方法能够减少特征维数,并且能保持高阶相互独立^[19]。下面介绍如何使用 ICA 方法来为每类图象生成独立关键子块空间:假设将某类图象中的每幅图象分割成 $N=n_1 \times n_2$ 子块(表示每个子块为 n_1 行 n_2 列),总共有 L 个图象子块,并且将每一个 $n_1 \times n_2$ 的图象子块均表示为一个行向量,其维数为 N 。若 L 个子块样本组成 $L \times N$ 的样本矩阵 x (其中必须保证 $L \geq N$) ,则 x 的 ICA 重构式为 $y=Ax(A$ 为分解方阵)。

生成独立关键子块空间的具体步骤如下:(1)计算 x 的特征向量和特征根,首先将特征值按从大到小的次序进行排序,然后取与其前 m 个特征根(含有 x 最可能多的能量)对应的特征向量 $p_i(N \times 1$ 列向量, $i=1, \dots, m$) 组成矩阵 $P_m(N \times m)$; (2)用 P_m 的转置矩阵 P_m^T 代替重构式中的 x 来得到 $y=AP^T \rightarrow P^T=A^{-1}y$, 其中 y 的每一行代表一个独立关键图象子块,即构成此类图象的独立关键子块空间, $A_{m \times m}$ 是在训练中得到的; (3)将属于此类图象的所有子块投影到独立关键子块空间,并用独立关键子块系数表示每个子块样本,即 $R_m=xP_m, R_m$ 是 $L \times m$ 大小矩阵,其每一行表示相应子块样本对于 m 个独立关键子块的投影系数坐标;同时使用最小平方误差求 x 的逼近值 x_{mse} , 使 $x_{mse}=R_mP_m^T=xP_mP_m^T=xP_mA^{-1}y$, 由此可以知道, xP_mA^{-1} 中第 i 行是第 i 个子块样本相对于 y 中独立关键子块的 m 个线性组合系数,也就是第 i 个图象子块样本的 ICA 特征。因此,对于任意测试图象子块 $I_{1 \times N}$, 其相对于这类图象的独立关键子块 ICA 特征就是 $c=IP_mA^{-1}$ 。本文使用颜色直方图来构造独立关键子块空间。

3 构造三角树

虽然,利用关键子块检索可以使提取图象特征变得容易简单,但是如果检索任一幅图象都要计算与数据库中的每幅图象的相似度,则产生不必要的冗余。如果在数据库中只保存计算出来的距离,那么就不用频繁地计算相似度了。这里,引入了三角树的概念和方法^[20],其基本思想就是,如果两个图象相

似,则与相同的参照物之间的距离也是相近的。

定义 1 假设 B 和 C 是两幅图象, δ 是个 $[0, 1]$ 之间的实数。对于参照物 $K_i(1 \leq i \leq m)$, 如果 $|S(K_i, B) - S(K_i, C)| < \delta$ (S 为两图象相似度), 则说明 B 与 C 相似。其中称 $\{K_i\}$ 为参照系, 称其中的图象为参照图象。

根据定义 1, 对于任意的两幅图象, 如果它们与所有参照图象之间的距离差异不大于 δ , 则认为它们是相似的, 这样只需保存它们与所有参照图象之间的距离即可。这种检索方法的复杂度为 $O(Nm)$, 其中, m 是参照系中参照图象的数目, N 为数据库中图象的数目。

这里, 可用一个三角树来实现快速检索机制。图 1 是三角树的基本结构。

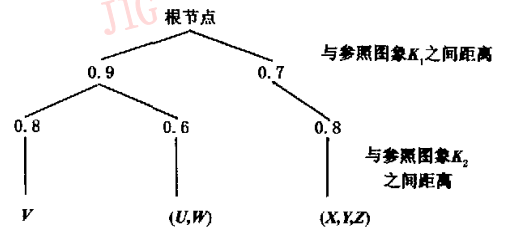


图 1 三角树的基本结构

其中, 每一层保存数据库中图象与参照系中图象的距离, 例如: 图象 V 与 K_1 之间的距离是 0.9, 图象 U, W 与 K_2 之间的距离是 0.6。检索的时候, 如果用户提交图象 Q , 用于寻找与 Q 的距离不超过 δ 的图象, 则首先, 计算 Q 与所有参照图象 $\{K_i\}$ 的距离 $D(K_i, Q)(1 \leq i \leq m)$, 并且按照深度优先的顺序遍历三角树。如果在 l 层上节点 P 处的值为 c , 且有 $|c - D(K_i, Q)| > \delta$, 那么就可以剪除节点 P 以下保存的所有图象。这样, 就可以大大缩减计算相似度的频繁度, 从试验结果来看, 这样可大大加快检索的速度。

4 基于独立关键子块和三角树的快速图象检索

4.1 提取直方图的 ICA 特征与独立关键子块

提取直方图的 ICA 特征时, 首先将样本中的每幅图象分割为 $N \times N$ 个子块, 并从 RGB 每种色素中分别提取 n 个 Bin 色彩直方图, 每个子块表示为 $1 \times 3n$ 行向量。若 M 表示图象子块总数, 那么样本图象集合就可用 $M \times 3n$ 矩阵 O 表示; 然后, 按照第 3 节的方法得

到 O 的独立关键子块空间和每个图象子块的 ICA 特征: (1) 计算 O 的特征向量和特征根, 得到由 $3n$ 个与特征值对应的特征向量 p_i ($3n \times 1$ 列向量, $i=1, \dots, 3n$) 组成的 $3n \times 3n$ 矩阵 P ; (2) 用 P 的转置矩阵 P^T 代替重构式 $y = W \times O$ 中的 O , 即得到 $y = WP^T \geq P^T = W^{-1}y$, 其中 y 的每一行是一个基于直方图的独立关键子块, 这样所有行就构成独立关键子块空间, 且每行相互独立, 也称为“直方图独立基概念空间”; (3) 若将 M 个图象子块中的每个子块均投影到“直方图独立基概念空间”, 则对于任意 M 个子块中的任意子块 $I_{1 \times 3n}$, 其 ICA 特征就是 $c = IPW^{-1}$, 再由所有子块的 ICA 直方图特征组成矩阵 $R_{M \times 3n}$.

此外, 需要考虑如下两点: (1) 选择合适的子块大小: 子块的多少决定了矩阵 O 的行数. 若子块太大, 则失去作为子块的意义; 若太小, 则矩阵 O 的行数太多. 通过实验对几种子块大小的比较分析说明了这一点; (2) 直方图特征维数 n 的大小要适中: 若太大, 则丢失许多色彩信息; 太小, 则增加计算的复杂度. 实验中, n 取为 10.

4.2 训练识别独立关键子块的支持向量机

通过 ICA 算法, 可得到 $3n$ 个独立关键子块, 总共要构造 $3n$ 个支持向量机. 现以独立关键子块 p_i 为例, 分输入层、隐含层和输出层来说明构造识别 p_i 的支持向量机 (用 S_i 表示) 的过程:

(1) 输入层

其功能是输入样本数据以及对数据的标注. 对数据进行标注时, 首先, 计算其欧拉距离; 然后将与 p_i 最接近的图象子块, 作为 S_i 的正样本, 标注为 1; 将与 p_i 距离最远的图象子块, 作为负样本, 标注为 -1, 而且样本数目的选取取决于子块总数和要训练的 SVM 个数. 一般来说, 每个 SVM 的训练样本数目都是 $[M/3n]$, 且正样本和负样本各占一半.

(2) 隐含层

该层有如下两个功能, 一是用非线性映射 Φ 把输入的 ICA 特征从原始低维空间映射到高维特征空间, 以达到高维可分目的; 二是计算高维特征和支持向量的内积. 这两步是通过核函数 Γ 来实现的, 核函数满足: $\Gamma(x, y) = (\Phi(x) \cdot \Phi(y))$. 其中, Γ 是核函数, Φ 是高维非线性映射, “ \cdot ” 是内积. SVM 中研究最多的核函数主要有多项式、径向基函数 (RBF)、多层 Sigmoidal 神经网络 3 类. 这里使用的是 RBF 核函数, 形式为: $K(x, y) = \exp\left\{-\frac{\|x-y\|^2}{\sigma^2}\right\}$.

(3) 输出层

用于输出识别结果, 并用支持向量机 S_i 识别图象中的独立关键子块 p_i (也就是关键子块所代表的特征向量), 若是独立关键子块, 则输出为 1, 否则标注为 -1.

同样, 对于其他的独立关键子块也可先分别训练一个 SVM, 表示为 S_i ($i=1, \dots, 3n$). 然后分别识别一幅图象中哪些子块属于独立关键子块 p_i , 也就是文本文件中的关键字.

4.3 FuzzySVM 构造

传统支持向量机的输出只是将识别物体标注为 1 或 -1, 即表示属于与不属于关系. 但是在独立关键子块识别中, 可能一个子块被两个或两个以上的 SVM 均标注为 1. 这样, 就无法判断这个图象子块到底属于哪个独立关键子块. 如果 SVM 的输出为概率, 那么就可以将某个图象子块归类为对应最大概率的 S_i , 为此, 本文引入模糊支持向量机 (Fuzzy Support Vector Machine, FuzzySVM) 的概念来将几个支持向量机的距离输出转换为概率输出, 并且由最大概率隶属度来判断某个图象子块到底属于哪个支持向量机. 在此可先用一个 Sigmoid 函数把每个 S_i 的输出, 映射到后验概率上^[21]

$$P(S_i | X) = \frac{1}{1 + \exp(A_i d_{out}^{(x)} + B_i)}$$

其中, A_i 和 B_i 是每个 S_i ($i=1, \dots, 3n$) 的 Sigmoid 参数, $d_{out}^{(x)}$ 是任意图象子块 X 相对于 S_i 的距离输出; 再用隶属函数: $\mu_{\text{ndkey}}(X) = \max_{S_i} P(S_i | X)$ 来判断某个图象子块 X 到底属于哪个独立关键子块. 同时设置一个阈值 ϵ , 如果 $\mu_{\text{ndkey}}(X) \leq \epsilon$, 则该图象子块不属于任何独立关键子块.

4.4 图象相似度度量

使用文本检索方法中的 $tf \times idf$ 模型来检索图象. 若 $X_{\text{ndkey}}(k)$ ($k=1, \dots, 3n$) 代表数据库中图象所包含的所有独立关键子块, $\omega_{i,k}$ 为图象 I_i 中某个独立关键子块 G_{key} 对应的权重 (如果图象中不包含该关键子块, 则此关键子块在此图象中的权重为 0), 则图象 I_i 可以用其包含的独立关键子块的权重向量来表示: $I_i = [\omega_{i,1}, \dots, \omega_{i,k}, \dots, \omega_{i,n}]$.

计算每个关键子块的权重有两个方面的因素: (1) 独立关键子块 G_{key} 在 I_i 中的权重 $\omega_{i,k}$ 随着出现频率的增高而变大, 可用独立关键子块频率 F_{ndkey} 来表示这种变化关系, 但如果某个独立关键子块在几乎

所有的图象中都是反复地出现,则不具备将一个图象与另一个图象区分开来的能力;(2)引入逆图象频率因子 u 来体现这种关系,其值与包含该独立关键子块的图象数目呈反比,计算方法如下:首先计算逆图象频率因子 $u = \log_2 \frac{N_{image}}{f_{key}} + 1$,其中, f_{key} 是包含某个独立关键子块 G_{key} 的图象数目, N_{image} 是数据库中图象数目;最后,将 F_{mdkey} 和 u 的乘积作为该独立关键子块在图象 I_i 中的权重 $\omega_{i,k}$,类似于文本检索中广泛采用的 $tf \times idf$ 检索模型。

若将用户提交的例子图象 Q 用于查询与之相似的图象,那么,应首先通过 SVM 识别 Q 中所有独立关键子块.若 Q 被表示为独立关键子块的权重向量: $Q = [\omega_{q,1}, \dots, \omega_{q,k}, \dots, \omega_{q,n}]$,则 Q 与数据库中所有图象之间的相似度 S 可通过如下余弦距离定义:

$$S(I_i, Q) = \frac{I_i \cdot Q}{\|I_i\| \|Q\|}, \text{其中 } \|\cdot\| \text{ 表示向量的模.}$$

4.5 采用三角树加快检索速度

为更进一步地提高检索速度,可通过构造三角树来仅仅保存如上计算出来的相似度^[21],而且参照系是自动生成的.当新的图象 R 插入数据库的时候,首先与参照系中原有的图象进行比较,如果 $\min(D(K_i, R)) > \sigma$,则将 R 作为一个新的参照图象插入参照系, σ 是一个事先设定的阈值.由于参照系图象的数目控制了三角树的深度,因此,三角树中的节点可以用如下的四元组表示

$$T_{node} = (\text{Nicode}, \text{Prefix}, \text{Layer}, \text{Value})$$

其中, Nicode 为节点编号; Prefix 为父节点编号; Layer 为层编号,对应每个参照图象; Value 为参照图象与数据库中图象之间的距离。

这样检索时,就可以首先按照深度优先的顺序遍历三角树,然后只要遍历符合相似度限制的节点即可。

5 实验结果和数据分析

为了验证本文方法的有效性,使用 Matlab6.0 在 Windows2000 下对本文的算法进行了仿真实验.实验用的各种类型图象共 1000 幅,从中分别选取 A、B 和 C 3 种类型的样本图象(表 1)作为样本。

表 1 几组不同的样本构造

样本类型	样本个数	样本组成
A	10 幅图象	5 幅亮色图象,5 幅暗色图象
B	20 幅图象	10 幅亮色图象,10 幅暗色图象
C	50 幅图象	25 幅亮色图象,25 幅暗色图象

衡量图象检索性能好坏使用的主要标准是查准率 (precision) 和查全率 (recall) 两个指标.其中查准率的含义是在一次查询过程中,系统返回的查询结果中相关正确图象的数目占所有返回图象数目的比例;而查全率则指系统返回的查询结果中的相关正确图象的数目占图象库中所有相关图象数目(包括返回的和没有返回的)的比例。

表 2、表 3 和表 4 分别列出了 A、B 和 C 3 类样本图象基于不同的子块大小,分别按照 ICA 方法得到的独立关键子块和按照 PCA 方法得到的特征子块,再通过训练 FuzzySVM 进行识别,最后得到的检索相似图象的检索正确率对比.从表 2 到表 4 可以看出:在样本图象数目为 10 的情况下,ICA 方法得到的独立关键子块与 PCA 方法得到的特征子块的检索效果相差不多,查准率和查全率都很低。

表 2 A 类样本检索正确率

提取特征类型		子块大小 (pixels)		
		33×33	66×66	99×99
PCA	查准率 (%)	65.6	62.8	60
	查全率 (%)	66	64	63
ICA	查准率 (%)	65.6	64.31	61.2
	查全率 (%)	65.6	66	64.4

表 3 B 类样本检索正确率

提取特征类型		子块大小 (pixels)		
		33×33	66×66	99×99
PCA	查准率 (%)	80.6	76.33	73
	查全率 (%)	86.1	80.2	74.96
ICA	查准率 (%)	85.47	83.3	79.4
	查全率 (%)	90.6	86	80.51

表 4 C 类样本检索正确率

提取特征类型		子块大小 (pixels)		
		33×33	66×66	99×99
PCA	查准率 (%)	83.55	81.625	79.9
	查全率 (%)	84.6	81.63	78.66
ICA	查准率 (%)	96.189	90.3	85
	查全率 (%)	97.4	93.975	84.35

这说明样本数量太少,由于提取出的色彩特征非常单纯,特征之间的相关性也不大,所以采用在高阶上保持独立的 ICA 独立关键子块就不必要;但随着样本数目的增加,查全率和查准率也有了明显的提高,在样本图象数目为 50 的时候,通过 ICA 得到独立关键子块的方法达到了 96% 以上的查准率和查全率,相同条件下,也取得了比 PCA 特征子块更

好的检索效果. 因为样本数目的增加, 使提取出的色彩特征越来越丰富, 其间的相关性也逐渐增大, 且 SVM 可通过核函数把低维特征向高维空间映射, 而 ICA 独立关键子块恰恰在高阶上保持独立, 所以 ICA 比 PCA 取得了更好的检索效果.

在 SVM 的生成过程中, 用 50 个图象作为样本便取得了令人满意的检索结果, 这说明小样本情况下, SVM 的识别效果就很好, 具有很强的学习推广能力. 另外, 统计结果还表明, 在样本类型一定的情况下, 小子块一般要比大子块的查全率和查准率都高一些, 但小子块的高查准率和高查全率都是以复杂度的提高为代价的. 当子块大小为 33×33 的时候, 已经取得了非常好的检索结果, 而且复杂度也不高, 因此没有必要划分更小的子块.

表 5 的数据表明, 在引入了三角树检索结构以后, 由于检索时, 计算相似度距离的次数减少, 大量缩减了检索的时间, 因此比原来没有使用三角树检索结构的系统提高 3 倍的效率.

表 5 原始系统与引入三角树检索机制的对比

系统原型		A	B	C
原始系统 ⁽¹⁾	时间(s)	0.71	3.48	14.0
	返回匹配结果数	100	400	2000
新系统 ⁽²⁾	时间(s)	0.32	1.54	6.12
	返回匹配结果数	100	400	2000

(1) 没有用三角树检索的原始系统; (2) 引入三角树检索

图 2 是实验检索结果, 其中左上角图象为用户提交的例子图象, 其余 5 幅图象是检索出来的与该幅图象色彩相似的结果集.



图 2 实验检索结果

6 总结

如今在图象检索中, 利用图象自身的视觉特征进行检索已经成为主流, 但是文本检索的很多方法

在图象检索中仍难以应用. 本文提出的首先通过 ICA 分析来得到独立关键子块空间, 然后使用 SVM 学习机来识别每幅图象的“独立关键子块”方法, 就可以用文本检索技术来实现对图象的检索. 实验表明, 这种方法可以取得很好的检索结果. 今后的工作集中在如下两个方面: (1) 在构造独立关键子块空间时, 本文使用的是直方图色彩特征, 而其他的纹理等信息同样有用, 因此将其他视觉特征结合起来是今后要考虑的一个问题; (2) 本文只考虑了每幅图象所包含的独立关键子块, 而没有考虑子块之间的空间约束, 因此如何实现基于子块空间约束机制的图象检索是要考虑的又一问题.

参考文献

- 1 姚玉荣, 章毓晋. 利用小波和矩进行基于形状的图象检索[J]. 中国图象图形学报, 2000, 5A(3): 206~210.
- 2 Flickner M, Sawhney H, Niblack W *et al.* Query by image and video Content: The QBIC System[J]. IEEE Computer, 1995, 28(9): 23~32.
- 3 Smith J, Chang Shih-Fu. VisualSeek: A fully automated content-based image query system [A]. In: Proceedings of Association for Computing Machinery Multimedia 96 [C], Boston, MA, USA, 1996, 87~98.
- 4 Pentland A, Picard R, Sclaroff S. Photobook: Tools for content-based manipulation of image databases [A]. In: Proceedings of the SPIE Conference on Storage and Retrieval of Image and Video Databases II[C], San Jose, CA, USA, 1994, 34~47.
- 5 Bach J R, Fuller C, Gupta A. The virage image search engine: An open framework for image management[A]. In: Proceedings of SPIE, Storage and Retrieval for Still Image and Video Databases IV[C], San Jose, CA, USA, 1996, 76~87.
- 6 Hearst M. TextTiling: Segmenting text into multi-paragraph subtopic passages[J]. Computational Linguistics, 1997, 23(1): 33~64.
- 7 Salton G, Buckley C. Term weighting approaches in automatic text retrieval [J]. Information Processing and Management, 1997, 24(5): 513~523.
- 8 Zhu Lei, Zhang Aidong, Rao Aibing *et al.* Keyblock: An approach for content-based image retrieval[A]. In: Proceedings of ACM Multimedia 2000[C], Los Angeles, CA, USA, 2000.
- 9 Rui Y, Huang T S, Mehrotra S. Content-based image retrieval with relevance feedback in MARS [A]. In: Proceedings IEEE International Conference on Image Processing [C], Santa Barbara, California, USA, 1997.
- 10 Turk M, Pentland A. Face recognition using eigenfaces [A]. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition [C], Hawaii, USA, 1992, 586~591.

- 11 Proriot J. Selection of variables for neural network analysis[J]. Journal of Nuclear Instrument and Methods, 1996, 361:581~585.
- 12 Vapnik V. The nature of statistical learning theory[M]. New York: Springer Verlag, 1995.
- 13 Burges C. A tutorial on support vector machines for pattern recognition[J]. Knowledge Discovery and Data Mining, 1998, 2(2):121~167.
- 14 庄越挺,刘骏伟,吴飞等. 基于支持向量机的视频字幕自动定位与提取[J]. 计算机辅助设计与图形学学报, 2002, 14(8):750~753.
- 15 Cover T M. Geometrical and statistical properties of systems and linear inequalities with applications in pattern recognition[J]. IEEE Transactions on Electronic Computers, 1965, 14:326~334.
- 16 Cortes C, Vapnik V. Support vector networks[J]. Machine Learning, 1995, 20(3):273~297.
- 17 Courant R, Hilbert D. Methods of Mathematical Physics[M]. New York, Interscience Publishers, Inc., 1953.
- 18 Hyvarinen A. Survey on independent component analysis[J]. Neural Computing Surveys, 1999, 2:94~128.
- 19 Bartlett M S, Lades H M, Sejnowski T J. Independent component representations for face recognition [A]. In: Proceedings of SPIE Conference on Human Vision and Electronic Imaging II [C], San Jose, CA, USA, 1998, 528~539.
- 20 Berman A P, Shapiro L G. Efficient content-based retrieval; experimental results [EB/OL]. <http://www.cs.washington.edu/research/imagetdatabase/reportfin.htm>, 2000-11-4
- 21 Platt J C. Probabilistic outputs for support vector machines for pattern recognition[A]. In: Smola A, Bartlett P, Scholkopf B Editor: Advances in Large Margin Classifiers[M], Boston, MA, USA: Kluwer Academic Publishers, 1999.



赵雪雁 1978年生,2003年获浙江大学计算机应用硕士学位。主要研究方向为图象检索、音频检索及数字化图书馆。



庄越挺 1965年生,1998年获浙江大学计算机应用博士学位,教授,博士生导师。主要研究领域为多媒体数据库、人工智能、基于内容的图象/视频信息检索、及视频动画等。



吴飞 1973年生,2002年获浙江大学计算机应用博士学位,讲师。主要研究方向为多媒体分析、计算机视觉和遥感处理。



刘骏伟 1978年生,2003年获浙江大学计算机应用硕士学位。主要研究方向为多媒体处理与数字化图书馆。

爱普生 GIS/CAD 行业解决方案 2003 年全国巡展

为推动我国各行业信息化、数字化建设的发展,促进数字城市建设的全面启动,爱普生公司联合了南方数码科技有限公司、微软(中国)有限公司、方正数码有限公司共同组织了爱普生 GIS/CAD 行业解决方案 2003 年全国巡展,从 8 月 26 日从昆明起程,历时一个月,走访了昆明、广州、长沙、南京、济南、太原、西安、北京、沈阳十个城市。这次巡展的每站都有知名专家做有关数字化、信息化建设的报告,并运用最好的数码工具和技术,使用户全面掌握从软件使用、设计到输出的全套 CAD/GIS 解决方案。爱普生公司协同其合作伙伴向用户分别介绍了其大幅面打印机 Epson Stylus 7600/9600/10600 在 GIS 行业的应用方案及使用伙伴的最新产品与解决方案报告。

作为技术领先的爱普生公司拥有多种型号的大幅面输出设备,针对不同的用户需要,提供不同的解决方案,并在 CAD/GIS 行业中拥有大量成功的用户。此次巡展展出的产品和解决方案主要包括:Epson Stylus PRO 7600/9600/10600 和大幅面打印机,以及与扫描仪厂商 Spotimage 及 Ikonos 卫星影像厂商等联合推出的解决方案。

本次 CAD/GIS 巡展活动让更多的用户享受到大幅面数字化流程的最新产品和解决方案。此次巡展将对我国的数字城市/数字省建设产生极大的推动作用。