

基于整体区域相似的交互式图象检索模型

陈添丁 刘济林

(浙江大学信息与电子工程系, 杭州 310027)

任志国

(杭州商学院信电学院, 杭州 310035)

摘要 高效的基于内容的图象检索在许多领域得到了广泛的应用, 基于内容的图象检索研究领域已经建立了一些系统, 但在实际使用中, 这些系统均有如下欠缺: (1) 这些系统均期望以相同的方法来处理各种不同类型的图象检索; (2) 这些系统在设计时, 均缺乏从使用者的需求出发. 实际上, 由于不同的检索方式是针对不同类型的图象, 为此, 提出了一个基于整体区域相似匹配的图象交互式检索系统, 该系统是一个基于小波变换的特征提取和图象整体区域相似的、语义分类和互动方法的图象检索系统. 与其他检索方法相比较, 此方法允许自适应查找和互动, 因此可缩小查找范围, 以提高检索效率. 实验结果表明, 该系统比其他一些系统精确和高效.

关键词 计算机信息管理系统(520·6070) 图象检索 整体区域相似 小波特征提取 互动方法

中图分类号: TP391.41 TP391.3 **文献标识码**: A **文章编号**: 1006-8961(2003)07-0810-07

Image Interactive Retrieval Model Based on Integrated Region Similarity

CHEN Tian-ding, LIU Ji-lin

(Department of Information and Electronics engineering, Zhejiang University, Hangzhou 310027)

REN Zhi-guo

(College of Information and Electronics engineering, Hangzhou University of Commerce, Hangzhou 310035)

Abstract Efficient content-based image retrieval has developed tremendously in many application areas. Content-based image retrieval research areas had established many systems, but these systems have deficiency for actually usage: (1) these systems are expected to process different kinds of image retrieval by same method and (2) these systems are designed without considering user requirement. In fact, different kinds of image retrieval need different retrieval mode. So it presents image interactive retrieval based-on integrated region similarity match. This system is an image retrieval system with interaction and semantic classification of based-on wavelet feature extraction and image integrated region similarity method. Compared with other retrieval methods of based-on image content, the method enhances retrieval efficiency by permitting adaptive search and interaction and narrowing search area. Experiment result shows that the system is more precise and efficient than other retrieval systems.

Keywords Computer information management system, Image retrieval, Integrated region similarity, Wavelet feature extraction, Interaction method

0 前言

高效的基于内容的图象检索在许多领域得到了广泛的应用, 例如, 生物医学、军事、商业、教育和 Web 图象分类和查找, 因此如何在图象资料中进行检索已成为一个亟需解决的问题.

目前在基于图象内容的检索领域中, 已有许多研究成果, 并且也建立了一些商用或研究用的系统,

如: QBIC^[1]、Virage、Photobook^[2]、VisualSEEK^[3]、Netra^[4]、PicToSeek^[5]、MARS^[6]等, 但这些系统在实际使用中, 却不如预期的理想, 其原因为:

(1) 这些系统均期望以相同的方法来处理各种不同类型图象的检索;

(2) 这些系统在设计时, 均缺乏考虑使用者的需求.

在上述研究中, 几乎都倾向让系统具有处理各种不同类型图象的能力, 这本身固然是很好的研究课

题,但实际上不同类型的图象需要不同的检索方式,例如:一张风景照片与一张动物照片,由于使用者对于它们有不同的需求,因此也需要以不同的方式进行检索.基于内容的图象检索(Content based image retrieval, CBIR)是依据自动获取的图象特征(色彩,纹理,形状等)来从一个图象数据库检索到语义相关的图象的一系列技术总称.

在本文的研究中,以使用者需求为出发点和依据,采用整体区域相似匹配策略来对图象的二维空间进行划分,提出了一种以各区块的特征作为图象相似的表示方法,即对图象的语义表达起关键作用的重要区块给予更多的权重,并提出了整体区域相似匹配的主要思想;同时,对整体区域相似匹配算法融入互动式技术,从而有效地改善了基于颜色和位置特征的查询效果.

1 整体区域相似

对一般题材的大型图象数据库,通过计算机理解图象是很困难的,而且明确表述一个查询和适当评估查询结果也是很困难的,这是 CBIR 面临的一个很大的挑战.已有的一般题材的 CBIR 系统,按照提取信号的方法大致可以分成直方图、色彩分布和基于区域的查询 3 类,也有的系统采用一个权值总和匹配度量、或和个别算法结合起来的融合检索方案^[1,7,8].

对大多数的图象检索系统而言,查询是通过一幅例子图象来进行匹配的,这是全局查询,因为相似度是基于图象的所有特性.与之形成对比的是局部查询检索系统,其检索的是基于图象的一个特殊区域.

1.1 小波变换的多分辨率色彩分布图象索引

本文提出一个能较好表示图象语义的基于 Daubechies 小波变换的色彩分布索引方法,即目标结构和局部色彩变量二者均由 Daubechies 小波系数表示.对大型图象数据库,特征矢量采用多级小波变换来获得,以提高查询速度.在数据库里,每一幅图象里的 R、G、B 三色成分,其每一色均应用

Daubechies 小波来进行快速小波变换(Fast wavelet transform, FWT),一些小波变换系数和它们的标准差都作为特征矢量储存起来.本文所处理的是二维彩色图象,对特定色彩像素,其时域是局部空间,而频域是在一个像素周围的色彩变量.基于此,本文的小波变换采用一个能高效表示图象每一个局部空间区域色彩变量的基函数,即 Daubechies 小波基, Daubechies 的正交基^[9]定义为:

$$\varphi_{r,j,k}(x) = 2^{j/2}\varphi_r(2^jx - k) \quad j, x \in \mathbf{Z} \quad (1)$$

式中, j 为尺度索引, k 为位置索引.

这里,函数 $\varphi_r(x)$ 有个性质,即 $\langle \varphi_r(x-k) | k \in \mathbf{Z} \rangle$ 是 $L^2(\mathbf{R})$ 空间的一个正交系列.

那么在尺度 2^{-j} , 函数 $f \in L^2(\mathbf{R})$, 势函数 f_j 可定义为

$$f_j(x) = \sum_k \langle f, \varphi_{r,j,k} \rangle \varphi_{r,j,k}(x) \quad (2)$$

函数细节或波动定义为

$$d_j(x) = f_{j+1}(x) - f_j(x) \quad (3)$$

对给定的尺度,在分析这些细节时,可定义一个正交基 $\psi_r(x)$,它跟上面所提到的 $\varphi_r(x)$ 有相似的性质.即 $\int_{-\infty}^{+\infty} \psi_r(x)dx = \dots = \int_{-\infty}^{+\infty} x^r \psi_r(x)dx = 0$. 由于小波变换能提供好的时间和频率定位,所以小波这个工具适合于数据压缩,特别是具有大量局部变量的函数.比如,基函数对尺度索引 j 和位置索引 k 都有相当的柔性,甚至可以应用小波变换的多次递归把图象进一步分解.图 1 是 Daubechies-8 父小波 $\varphi(x)$ 、母小波 $\psi(x)$ 的波形,图 2 是图象的 Daubechies-8 小波变换的多尺度结构.

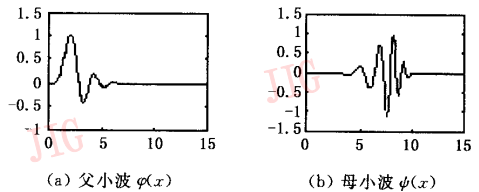


图 1 Daubechies-8 小波波形

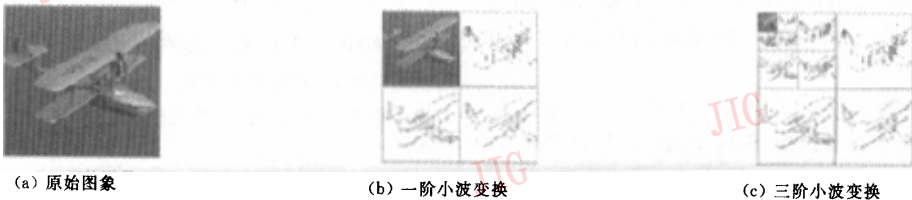


图 2 飞机图象小波变换的多尺度结构

1.2 图象整体区域分割

图象区域分割是进行相似匹配的基础,大家知道,图象分割是一个极为困难的过程,且一直是计算机视觉的一个难题,比如,一个图象分割算法可以把一幅小狗图象分割成小狗和背景两个区域,而相同的算法也可以把另一幅小狗图象分割为小狗的身体、前腿、后腿、眼睛、草地背景和天空 6 个区域,本文的整体区域是指图象按其颜色分布,自适应地分割成符合原始图象特征的区块,以避免传统的平均分割或仅依靠几种主色进行分割,而不能很好地反映原始图象基本特征的弊端.并且本文方法对出现频率高或对图象的识别有重要影响的颜色区块,即所谓的重要区域给予特别关注.因此,整体区域相似度量有如下主要优点:

(1) 与基于个别区域检索对比,整体区域相似采用全面“软相似”的方法来进行匹配,这不仅可以减少不准确分割的不利影响,而且对大小不一的图象数据库更加适合,以前的一些系统缺乏的正是这个重要特性;

(2) 对一些对象用语义分类来帮助理解,通常有的目标会在一些特殊区域出现,其中很典型的特征,如花朵有绿叶的陪衬,小船通常在有水的地方.

本小节描述的是基于 k 均值算法^[10]的使用色彩和空间变量特征的图象分割,为分割图象,系统先把图象分成 4×4 的像素块,并且对每一块提取一个特征矢量,然后用 k 均值算法把特征矢量聚成几个类,每一类对应图象的一个分割区域.由于块很小,并且边缘短小,因此对检索的影响是很小的,本文选择块方式的分割,而不是像素方式的分割,是为了减少计算开销.

假设数列 $\{x_i; i=1, 2, \dots, k\}$, k 均值算法的目的是把数列分成 k 组具有最小均值的 $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_k$, 其失真函数式为

$$D(k) = \sum_{i=1}^k \min_{1 \leq j \leq k} (x_i - \hat{x}_j)^2 \quad (4)$$

k 均值算法没有指定选择多少聚类,本文通过试探性地逐渐增加 k 值来确定聚类数目,当满足以下准则时,则采用停止其增长的方法来选择聚类 k 的数目,从 $k=2$ 时开始,如果满足以下条件之一就停止 k 的增长:

(1) 失真函数 $D(k)$ 低于一个阈值. 一个低 $D(k)$ 值表示较纯的聚类过程,但这个阈值不是很关键的,因为整体区域相似的方法就是对 k 值不敏感;

(2) 失真函数的差式 $D(k) - D(k-1)$ 低于一个

阈值(与 $D(3) - D(2)$ 的平均差式相比较). 一个低的 $D(k) - D(k-1)$ 值表示聚类的过程收敛,阈值大小决定图象分割的全部时间,它虽对速度是关键的,但不能体现最后图象分割的质量,阈值可以按照实验的运行时间来调整;

(3) k 的值超过上限. 本文允许图象分割数最大为 16 块,即假设一个图象少于 16 个典型的目标. 图 3 就是原始图象及分割的结果.

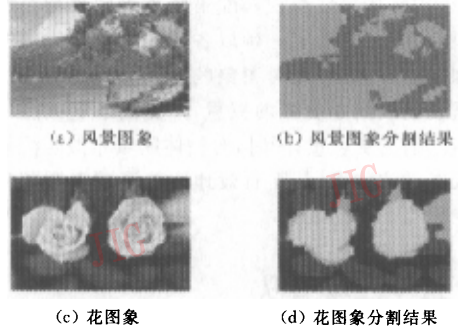


图 3 两种图象及对应的区域分割示例

1.3 整体区域相似的匹配原则

使系统能捕捉图象语义的另一个策略是定义一个鲁棒的基于区域的相似措施,即整体区域相似匹配. 它能结合所有区域分割的特性,以便于一幅图象的有关信息都能充分地使用. 系统通过定义一个全面的图象到图象的相似措施来为用户提供一个简单的查询接口,如果需要,系统还能增加一个函数来允许用户查询基于一个明确区域或一些区域的图象,再融合互动式的查询,效果将能得到更好改善.

设图象 1 和图象 2 分别由区域集合 $R_1 = \{r_1, r_2, \dots, r_n\}$, $R_2 = \{r'_1, r'_2, \dots, r'_m\}$ 组成. 那么区域 r_i, r_j 的相似度量 $s(r_i, r_j) = e^{-d(r_i, r_j)/\sigma}$, 其中, $d(r_i, r_j)$ 是两区域特征矢量的欧氏距离, σ 是两区域集合分割数抽样标准差(一般取 $0.2 \sim 0.4$). 图象 I 与区域 r 的相似度量 $s(r, I) = \max_{k \in \{1, \dots, n\}} (s(r, r_k))$, 图象集合 $I_i = \{I_1, I_2, \dots, I_n\}$ 与区域 r 的相似度量 $s(r, I_i) = \sum_{k=1}^n (s(r, I_k))$, 两图象之间区域与区域的相似匹配如图 4 所示,显然整体区域相似应该是对区域之间的多对象的考虑,而不是简单的一对一,如此才能提高相似度量的准确性和鲁棒性. 由此可得两图象 $I_n = \{r_1, r_2, \dots, r_n\}$, $I_m = \{r'_1, r'_2, \dots, r'_m\}$ 的相似度量 $s(I_n, I_m) = \sum_{i,j} \beta_{i,j} s(r_i, r'_j)$, 其中, $\beta_{i,j}$ 是信度

系数,表示两图象的区域之间相似的程度.这样图象之间的整体区域相似匹配如图 4 所示.

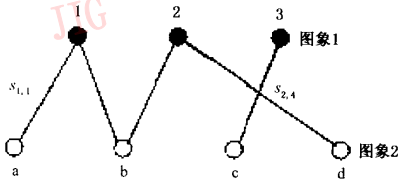


图 4 图象整体区域相似匹配

2 互动式查询模式

在最近的基于图象内容的检索研究中,越来越多的研究者指出:要让系统有好的查询功能,人的介入是很重要的^[1].本文在整体区域相似匹配的基础上,融合了相似查询与特征查询的互动模式,使检索结果更为可靠有效.

这个查询模式解决了图象检索的一个难点,即不同使用者之间的认知差异问题.所谓相似查询是指使用者选取一张图象作为例子来进行查询的方法;而特征查询就是使用者使用一些特征来描述一张图象的方法.由于简单、直觉、易于使用是使用者对于系统的基本要求,因此从使用者需求的角度入手,建立一个合适的查询模式是很必要的.本文在整体区域相似匹配查询和语义分类方法的基础上,结合互动式的查询模式,让使用者可以在互动的过程中,不断依据图象的颜色、纹理(花纹)、形状等特征,或是依据图象间的相似性来进行查询.并且在查询结果中,使用者还可以通过选取其中任一张图象来查询与那张图象相似的其他图象;系统也将一些可以辅助查询的相关信息在查询结果中同时提供给使用者,让使用者可以据此再进行查询.

在整体区域相似匹配基础上,融合了相似查询的方法,其相似度度量是模糊特征对比度量(Fuzzy feature contrast measure,即 FFC) P 和图象相交直方图(Histogram Intersection,即 HI) R 结合起来的查询方法,简单称之 FFC+HI.

$$S(k, h) = \underbrace{w_c \times R(H_{L_c(k)}, H_{L_c(h)})}_{\text{利用Histogram(自动)}} + \underbrace{w_s \times P_1(k, h)}_{\text{利用“感觉分层”}} + \underbrace{w_s \times P_s(k, h)}_{\text{(\lambda E)}}$$

其中, P_1 、 P_s 分别为纹理、形状对比度.

若希望在整体区域相似度量中,颜色(c)、纹理

(t)、形状(s)所占的比重相当,则颜色权重

$$w_c \gg 1, w_t \approx \frac{1}{E\left[\sum_{f \in F_t} \mu(L_k^{(f)})\right]}, w_s \approx \frac{1}{E\left[\sum_{f \in F_s} \mu(L_k^{(f)})\right]}$$

在特征查询中,特征的界定是很重要的,界定的原则一般是:(1)明显度,即特征是否容易被使用者观察到;(2)鉴别性,即是否有助于鉴别图象资料的不同;(3)涵盖性,即是否可用来描述大多数图象资料.

本文把图象的色彩、纹理、形状的描述用“感觉分层”来分级描述,即将图象对于特征的符合度分为 5 级,分别为:第 0 级,非常符合;第 1 级,符合;第 2 级,略为符合;第 3 级,不符合;第 4 级,反例.以 $L^{(f)}(i)$ 表示图象 i 对特征 f 的感觉分层: $L^{(f)}(i) \in \{0, 1, 2, 3, 4\}$,在感觉分层的权值度量中,一幅图可以一个颜色直方图 H 来表示:

$$H = \{(C_j, P_j) | C_j \in C, 0 \leq P_j \leq 1, \sum_{1 \leq j \leq N} P_j = 1, \text{且 } 1 \leq j \leq N\}$$

应该指出的是,相似查询与特征查询的互动模式是有区别的,其中特征查询使用的场合是使用者知道如何用什么样的特征来描述图象,由于它的性质是属于粗略查询,因而与文字查询中用的“关键字”查询有些类似.相似查询使用的场合是使用者不知要如何用特征来描述图象,由于它的性质是属于详细查询,因而跟整体区域相似查询的方式更为接近.

3 实验与评价

基于上述思想,在 Linux 操作系统上实现了一个图象数据库检索系统.利用该系统在 P II-450、RAM128M 的 PC 机上对一个具有 1 000 幅、大小为的彩色图象数据库进行处理,大约需要 23min,一次检索的反馈时间达毫秒级.数据库共有 1 000 幅图象(ID 号分别从 0~999),依语义分成 10 类,有人物,风景,建筑,汽车,动物,花朵等,系统依据整体区域相似匹配和互动查询来进行设计,并且还提供一些文本信息,实验结果是满意的.

实验 1 是对整体区域相似(Integrated region similarity, IRS)互动的融合算法与模糊特征对比(FFC)方法、相交直方图(HI)方法进行比较,即进行平均查全率和检索精度的对比实验,从图 5 可以看出,FFC 和 HI 方法的检索效果相差不多,而整体区域相似互动的融合方法在查全率较小(查找范围较小)的情形下,优势更为明显.

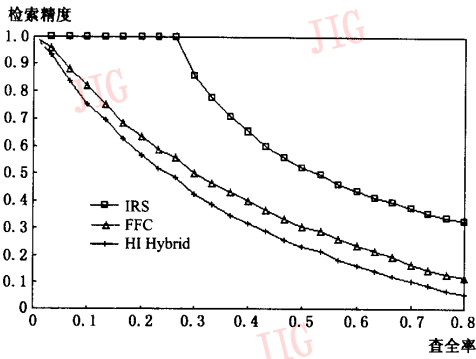
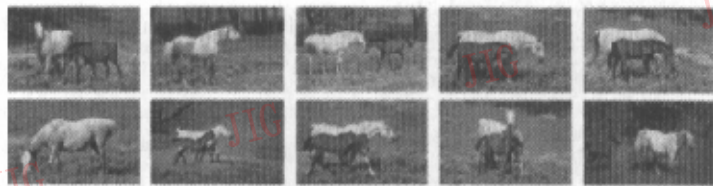


图 5 整体区域相似互动融合方法与其他方法检索效果的比较

实验 2 则是给出几个典型的结果,用以验证算法的有效性,一次检索返回 10 幅图象,第 1 幅图象为例子图象(查询图象),而整体区域相似距离是指与例子图象的相似程度的度量.其中马图背景有的复杂,有的简单,背景和前景明朗,实验结果如图 6 与表 1 所示;而花朵的背景和前景区分度不高,轮廓复杂,实验结果如图 7 与表 2 所示;选择的飞机图象则是大小不一,这样可以验证图象大小悬殊的情形下的匹配查询效果,结果如图 8 与表 3 所示.由图 6~图 8 可见,图象检索结果准确率都很高,较复杂的图象偶尔有一两张(返回 10 张)串题了(如同为草地和树木背景的小狗图象)但经一次互动反馈后,都能获得满意结果,限于篇幅,文中选择其中 3 类情形.

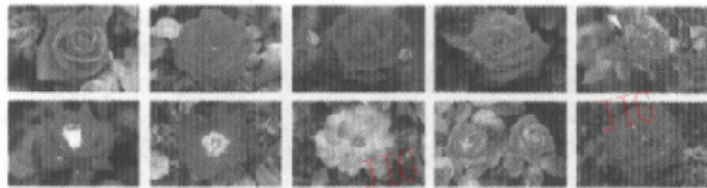


(a) 马图的检索结果

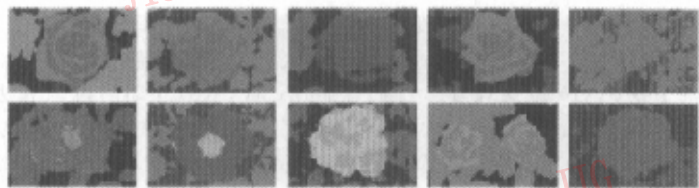


(b) 用于查询的对应特征矢量

图 6 马图象检索结果



(a) 花朵图象的检索结果

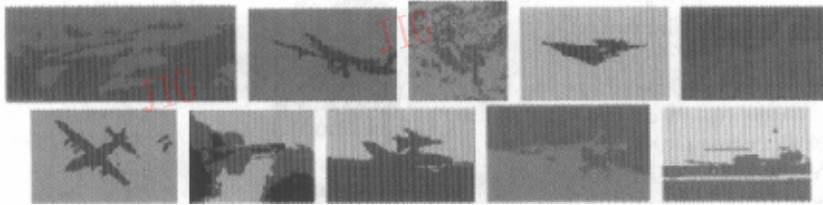


(b) 用于查询的对应特征矢量

图 7 花朵图象检索结果



(a) 飞机图象的检索结果



(b) 用于查询的对应特征矢量

图 8 飞机图象检索结果

表 1 图 6 系统反馈的检索结果注释文本

图象 ID 号	729	725	705	734	783	747	749	739	745	713
图象分割区块数	5	4	4	4	6	4	4	4	4	3
整体相似距离	0.00	8.50	10.10	10.89	10.94	11.71	12.21	12.25	13.58	14.34
CPU time: 0.03 seconds / Database size: 1000 images										

表 2 图 7 系统反馈的检索结果注释文本

图象 ID 号	624	634	616	618	669	650	620	619	672	657
图象分割区块数	6	4	4	4	6	4	4	4	4	3
整体相似距离	0.00	13.01	17.02	18.48	19.61	19.94	20.18	20.30	21.81	22.60
CPU time: 0.02 seconds / Database size: 1000 images										

表 3 图 8 系统反馈的检索结果注释文本

图象 ID 号	38	42	29	45	55	46	91	82	99	66
图象分割区块数	4	4	4	5	4	5	4	4	3	3
整体相似距离	0.00	6.87	10.3	10.85	11.00	12.05	12.23	12.69	12.86	13.02
CPU time: 0.03 seconds / Database size: 1000 images										

4 结论及今后的工作

从实验结果看,对于外观检索来说,整体区域相似匹配是很有效的,即对于一般题材的图象,都能快速有效地提高总体检索效果,并且结合互动查询模式,能满足使用者对于图象检索的需求,还能使不同的使用者可以利用对图象特征及相似的不同认知来查询.通过整体区域相似匹配方法与一般常用的方法在查询效率上的优劣对比的实验结果显示,整体区域相似匹配方法查询效率是优于其他方法的,检索效果也得到了显著的改善.由此可见,融合互动技术的整体区域相似匹配方法对于提高基于内容的图

象检索系统的查询效果是有效的.虽然该研究是针对一般题材的图象,但其中的方法及技术都可以用于各式各样的图象检索.

今后可在原型系统的基础上对系统的功能进行进一步的扩展,主要是通过融合相关互动技术和改进算法来更适应图象意思(语义),而不仅是外观的检索,以便在查询过程中更好地实现系统的自适应和自学习能力,从而使系统应用在大规模图象数据库时,还能有好的查询表现.

参考文献

1 Filickner M, Sawhney H, Niblack W *et al*, Query by image and video content: The QBIC system[J]. IEEE Transactions on

- Computer, 1995,28(9):67~74.
- 2 Pentland A, Picard R W, Sclaroff S. Photobook: Content-based manipulation of image databases [J]. International Journal of Computer Vision, 1996,18(3): 322~330.
 - 3 Smith J R, Chang S F. VisualSEEK: A fully automated content-based image query system [A]. In: Proceedings of Association for Computing Machinery Multimedia'96 [C], New York, 1996: 142~149.
 - 4 Ma W Y, Manjunath B S. Netra: A toolbox for navigating large image databases [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Berkeley, CA, USA, 1997: 231~238.
 - 5 Gevers T, Smeulders A W M. PicToSeek: Combining color and shape invariant features for image retrieval [J] IEEE Transactions on Image Processing, 2000,9(1):102~111.
 - 6 Ortega M, Rui Y, Chakrabarti K *et al.*, Supporting Similarity Queries in MARS [A]. In: Proceedings of Association for Computing Machinery Multimedia'97 [C], Orlando, FL, USA, 1997:269~277.
 - 7 Gupta A, Jain R. Visual Information Retrieval [J]. Communication of Association for Computing Machinery, 1997, 40(5):70~79.
 - 8 Sheikholeslami G, Chang W, Zhang A. Semantic clustering and querying on heterogeneous features for visual data [A]. In: Proceedings of Association for Computing Machinery Multimedia [C], New York, 1998:3~12.
 - 9 Mukherjea S, Hirata K, Hara Y. AMORE: A World Wide Web image retrieval engine [J]. Proceedings of World Wide Web, 1999,2(3):115~132.
 - 10 Hartigan J A, Wong M A. Algorithm AS136: A k-means Clustering Algorithm [J]. Applied Statistics, 1979. 28: 100~108.
 - 11 Rui Y, Huang T S, Chang S F. Image Retrieval: Current Techniques, Promising Directions, and Open Issues [J]. Journal of Visual Communication and Image Representation, 1999,10:53~61.



陈添丁 1968年生,2003年获浙江大学博士学位。主要研究领域是图象处理、图象(视频)检索和 MPEG-7。



刘济林 1947年生,博士生导师。主要研究领域是图象处理、机器人视觉与导航、ITS、MPEG 系列及视频检索。

任志国 1963年生,硕士生导师,教授,杭州商学院信电学院院长。主要研究领域是多媒体通信与网络安全。