

基于 JPEG2000 的图象局部可分级检索

钟敏生 马争鸣

(中山大学电子系信息处理实验室, 广州 510275)

摘要 提出了一种基于 JPEG2000 的图象局部可分级检索方法. 首先, 对图象采用 JPEG2000 的基本算法进行压缩编码; 然后, 把压缩码流按零树结构编排, 每个零树结构对应着图象的某个网格区域. 从零树结构所包含的 DWT 系数可以分级重建这个网格区域的内容. 检索时, 先从图象的最低分辨率子图象(即 LL 子图象)的预览开始, 一旦发现目标, 则锁定目标; 然后在空间分辨率和信噪比两个方面逐级浮现目标所在网格的内容. 由于该方法仅仅使用部分压缩码流数据来恢复图象的局部内容, 从而节省了大量的计算资源; 最后给出了基于 JPEG2000 的图象局部可分级检索的实验实例.

关键词 小波图象编码 JPEG2000 图象检索

中图法分类号: TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2001)12-1191-07

JPEG2000-Based Scalable Lookup of Image's Local Regions

ZHONG Min-sheng, MA Zheng-ming

(Lab. of Information Processing, Dept. of Electronics and Communication Engineering, Zhongshan Univ., Guangzhou 510275)

Abstract A method of JPEG2000-based scalable lookup of image's local regions is presented in this paper. At first, we compress the image with the baseline algorithm of JPEG2000. Then we arrange the compressed data in light of zerotree structures, each zerotree is corresponding to a local region (mesh) of original image. When we want to look up a target on the image, we first browse the lowest resolution subimage (LL subimage) and locate mesh which the target belongs to. Then we reconstruct the mesh level by level in both spatial resolution and Signal-to-Noise (SNR). In this way we use only part of the compressed data to reconstruct the local region of image and show the target. It can save much computation resources. In the end of this paper, the experimental results of JPEG2000-based scalable lookup of image local regions are given.

Keywords Wavelet image coding, JPEG2000, Image lookup

0 引言

目前, 图象大都以 JPEG 格式存储在计算机里. 如果要把图象从计算机调出, 查找某个局部目标(景物), 例如在一张偌大的卫星照片中查找一个导弹发射架, 则不论这个局部目标所占的区域与整幅图象的尺寸相比多么微小, 都必须首先对图象的全部压缩码流进行解码, 重建整幅图象, 然后才能关注要查找的局部目标. 这种牵一发而动全身的劳顿源于所采用的 JPEG 算法. 当然, 人们可以采用其他的图象

压缩编码算法, 但是, 如果采用的算法不能与国际标准接轨, 推广应用的范围也就很有限了.

JPEG2000 是一个基于 DWT(离散小波变换)的、新的图象压缩编码国际标准, 已于 2001 年正式颁布. 现在, JPEG2000 标准可以从 JPEG 站点下载^[1]. 本文根据 DWT 的特点, 提出一种“基于 JPEG2000 的图象局部可分级检索”方法. 这种方法, 不但与即将推广的图象压缩编码国际标准 JPEG2000 接轨, 而且, 更为重要的是, 可以仅仅使用部分相关的压缩码流数据, 在空间分辨率和信噪比两个方面分级重现图象的局部目标. 这样做, 不但

免除了牵一发而动全身的劳顿,而且一旦发现找错目标或者目标的分辨率已达到应用的要求,可以立即终止解码过程。

1 图象的可分级编解码

近年来,由于图象传输、显示和检索等方面的需要,图象的可分级编解码颇受青睐^[2,3]。许多与图象编解码有关的国际标准(如 JPEG2000、MPEG2 和 MPEG4 等)都有图象或图象序列可分级编解码的语法和语义方面的规定。所谓图象的可分级编解码,是指对图象压缩编码后的数据进行分类,即根据其在图象重建中的重要性赋予不同的级别。这样当图象应用系统的服务质量(QoS)下降时,即可优先保证对级别较高的数据的服务质量。

图 1 是图象空间可分级(Spatial Scalability)编解码的示意图。

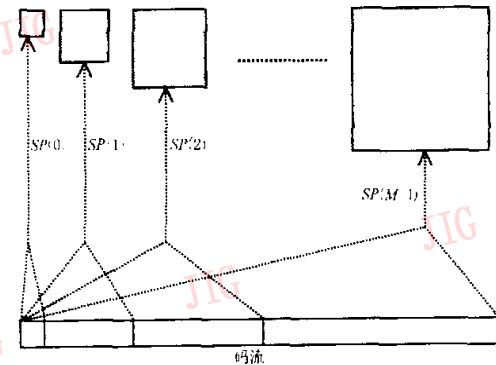


图 1 空间可分级(M层)

图 1 中,码流被分为 M 个部分,级别按从左到右的次序,由高到低。对第 1 部分数据解码,可得到



原图象的一幅空间分辨率较低的子图象;对第 2 部分数据解码,加上对第 1 部分数据解码的结果,可得到一幅空间分辨率比上一幅高一个级别的子图象。依此类推,解码的部分越多,得到的子图象空间分辨率也越高,但每一次得到的子图象的信噪比与原图象是一致的。

图 2 是图象信噪比可分级(SNR Scalability)编解码的示意图。

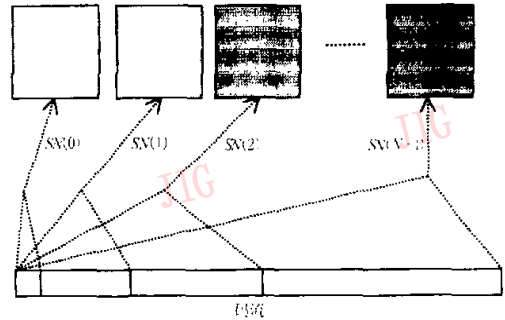


图 2 信噪比可分级(N层)

在图 2 中,码流也同样被划分为多个部分,对每个部分的数据依次解码,就可以得到一系列与原图象具有相同空间分辨率,但信噪比不断改善的图象。

图象的可分级编解码要受到图象编码算法的制约。如果采用小波图象编码算法,则可以进行空间和质量同时可分级的编解码,从而产生 Zoom-In 的视觉效果。图 3 是利用小波图象编码算法对 Lena 图象进行空间和质量同时可分级编解码的例子。从图 3 可见,图象不但愈来愈大(空间分辨率愈来愈高),而且愈来愈清晰(信噪比愈来愈高)。



图 3 基于 DWT 的图象空间和质量同时可分级编解码

2 基于 JPEG2000 的图象局部可分级检索

2.1 基本原理

2.1.1 图象的预览

目前, 图象预览的主要方法是先对图象的压缩码流进行完全解码, 然后对解码得到的图象进行多次亚抽样, 从而得到一个空间分辨率较低、但信噪比基本保持不变的图象缩影。这里, 图象都是采用基于 DWT 的 JPEG2000 进行压缩。利用图象 DWT 系数的特点, 可以仅从图象的部分压缩码流数据直接得到图象的预览, 既省工, 又省时。

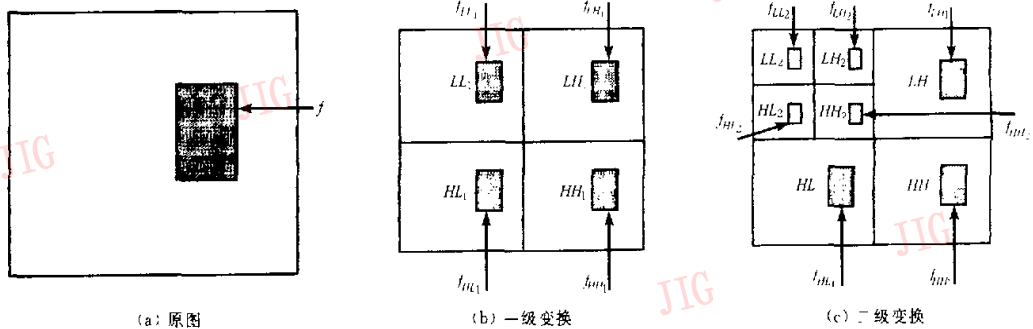


图 4 图象的小波变换

2.1.2 图象的局部重建

在图象预览时, 一旦发现目标, 就应提高目标所在局部区域的空间分辨率和信噪比。目前, 要重建图象的局部区域, 必须先对图象压缩数据完全解码, 然后再取所需的局部区域。

事实上, 由于原图象的景物与它的 LL_1 子图象的景物是完全一一对应的, 因此, 如果只对原图象的某个局部区域感兴趣 (如图 4(a) 中的 f 区域), 就只需考虑 LL_1 子图象中与之相对应区域的 DWT 系数 (如图 4(b) 中的 f_{LL_1} 区域)。如果想要提高 f_{LL_1} 区域的信噪比和空间分辨率, 那么按照小波变换的理论, 可以进一步考虑 LH_1 、 HL_1 和 HH_1 子图象中相应区域的 DWT 系数 (如图 4(b) 中的 f_{LH_1} 、 f_{HL_1} 和 f_{HH_1} 区域)。利用这些 DWT 系数, 通过 DWT 反变换, 可以重现图象 f 区域的内容。这就是图象 DWT 系数的局部空频属性, 也就是说, DWT 系数只是反映原图象中与 DWT 系数的空间位置相对应的局部区域的频率成分。根据图象 DWT 系数的局部空频属性, 可以由部分 DWT 系数重建图象的局部区域。

如图 4 所示, 图象经过 DWT 变换后, 分解为 4 幅子图象: LL_1 、 LH_1 、 HL_1 和 HH_1 。其中, LL_1 子图象是原图象经过 (横向+纵向) 低通滤波和亚抽样之后得到的子图象。由于低通滤波导致信噪比降低, 而亚抽样则导致空间分辨率降低, 因此, LL_1 子图象可以视为原图象的一个信噪比和空间分辨率都比原图象低一个级别的缩影 (预览)。在 JPEG2000 中, LL_1 子图象作为图象 DWT 系数的一部分, 经过量化和熵编码后, 放置在图象压缩码流的前端。这意味着对图象压缩码流的前端数据进行熵解码和反量化后即可得到图象的预览。这是其他图象编码算法 (如基于 DCT 的 JPEG 算法) 难以做到的事情。

2.1.3 塔台 (零树) 结构与和图象的局部可分级重建

如图 4(c) 所示, 可以对 LL_1 子图象再做小波变换, 将 LL_1 子图象进一步分解为 4 幅子图象: LL_2 、 LH_2 、 HL_2 和 HH_2 。同样道理, LL_2 子图象可以看成是 LL_1 子图象的一个在信噪比和空间分辨率都低 1 个层次的缩影, 而 LL_1 子图象又是原图象的一个在信噪比和空间分辨率都低一个层次的缩影。因此, LL_2 子图象可以看成是原图象的一个在信噪比和空间分辨率都低 2 个层次的缩影。如果只是对 LL_1 子图象中的局部区域 f_{LL_1} 的内容感兴趣, 就只需考虑 LL_2 、 LH_2 、 HL_2 和 HH_2 子图象中相应区域的 DWT 系数 (如图 4(c) 中 f_{LL_2} 、 f_{LH_2} 、 f_{HL_2} 和 f_{HH_2} 区域)。利用这些 DWT 系数可以完全重建 LL_1 子图象中 f_{LL_1} 区域的 DWT 系数。如果再加上 LH_1 、 HL_1 和 HH_1 子图象中 f_{LH_1} 、 f_{HL_1} 和 f_{HH_1} 区域的 DWT 系数, 则可以完全重建原图象中局部区域 f 的内容。

f_{LL_2} 、 f_{LH_2} 、 f_{HL_2} 、 f_{HH_2} 、 f_{LH_1} 、 f_{HL_1} 和 f_{HH_1} 区域的 DWT 系数构成了一个如图 5 所示的塔台结构。在

这个塔台结构之内的 DWT 系数都与原图象的局部区域 f 所包含的内容有关,而在这个塔台结构之外的 DWT 系数则与区域 f 所包含的内容无关.在使用 JPEG2000 进行图象编码时,通常要做多次小波

变换,每次小波变换都会使塔台向上延伸一层.塔台的每一层都对应着区域 f 所包含的内容在信噪比和空间分辨率某个层次的重现.

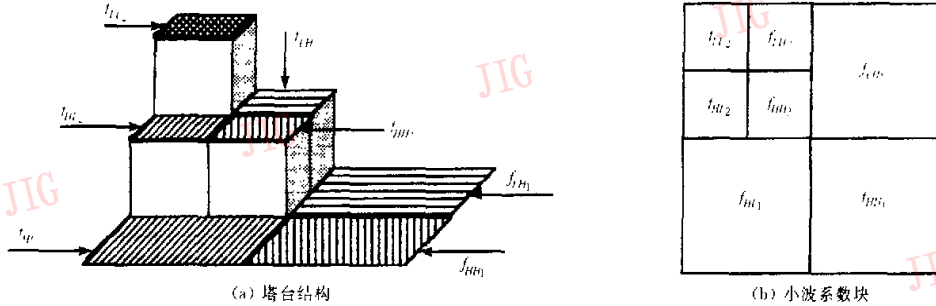


图 5 图象小波系数的塔台结构

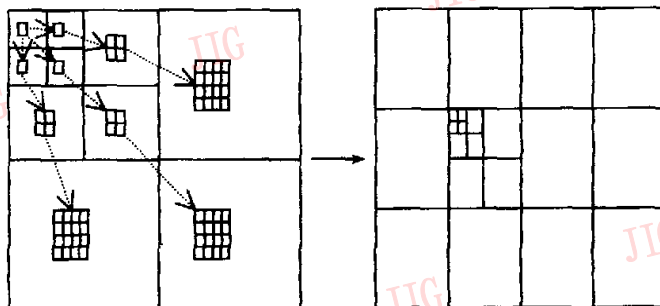
需要指出的是,这里所说的 DWT 系数的塔台结构,可以看成是零树结构的推广. DWT 系数的零树结构最早由 Shapiro 提出^[4],现已成为 JPEG2000 的一个重要组成部分.但是,塔台结构与零树结构的出处不同,作用也不同.零树结构源于量化后零值 DWT 系数的聚类,而塔台结构则是根据 DWT 系数的局域空频特性应运而生的.零树结构的作用在于:只要知道零树根的位置,就可以推出零树结构所包含的所有零值 DWT 系数的位置.零树结构的作用与 JPEG 对于量化后零值 DCT 系数的游程编码相似,都是用最少的符号来标记量化后零值变换系数的位置,以便提高图象编码的压缩比.塔台结构的作用在于收集与图象某个局部相关的 DWT 系数,利用这些 DWT 系数,通过多次小波反变换,可以分级重现图象这个局部的内容.为了避免引入过多概念,以后仍把塔台结构称为零树结构.

2.2 图象 DWT 系数的排列

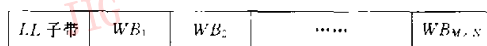
图象 DWT 系数的排列是图象局部可分级检索的重要内容.一个好的排列,能起到事半功倍的作用.首先采用固定网格模式确定图象局部的空间位置,然后按照固定网格的顺序来排列图象的 DWT 系数.

所谓固定网格模式就是把图象区域划分为 $M \times N$ 个网格,并按照从左到右、从上到下的顺序编号.图象小波变换后的各个子带图象也相应地划分为 $M \times N$ 个网格,处于相同空间位置的网格所包含的 DWT 系数组成一个 DWT 系数块(如图 6(a)所示).每个 DWT 系数块都与某个图象网格相对应,并且由这个块所包含的 DWT 系数,通过局部小波反变换,可以完全重建这个图象网格的内容.把图象网格的编号作为 DWT 系数块的编号,各个 DWT 系数块将按照编号顺序依次写入码流.

图 6(b)表示码流数据的排列方式.将 LL_1 子带



(a) 小波系数块的编号



(b) 码流结构

图 6 图象 DWT 系数的排列

图象的 DWT 系数作为原图象的预览放在码流的最前端,随后是各个 DWT 系数块按其编号依次排列.每个 DWT 系数块都剔除了属于 LL_1 子带图象的那部分 DWT 系数.另外,每个 DWT 系数块都加入一个表示编号的字符串,以便查找.在图象局部可分级检索时,先把位于码流最前端的 LL_1 子带图象拿出来作为图象的预览.如果在图象预览时发现目标,则根据目标所在的图象网格的编码,取出相应的 DWT 系数块,再加上 LL_1 子带图象相应区域的 DWT 系数,通过逐级小波逆变换,实现分级检索的目的.

2.3 实验结果

图 7(a)是 512×512 的 Lena 原图象,图 7(b)是对 Lena 原图象进行二次小波变换后 DWT 系数的分布图(因尺寸太大,2 个图象都缩小为原来尺寸的四分之一来显示).从图 7(b)可以看到,左上角处于高亮状态的 LL_2 子带图象基本保留了 Lena 头像的大概轮廓,可以作为 Lena 原图象的一个尺寸较小(空间分辨率较低)、细节模糊(信噪比较差)的预览.

图象的检索从 LL_2 子带图象(图 8(a))的预览开始.

如果只对 Lena 的眼睛(图 8(a)方框部分)感兴趣,那么只需拿出 LL_2 子带图象中这部分的 DWT 系数,再加上相邻高频子带图象(LH_2 、 HL_2 和 HH_2)中与方框相对应部分的 DWT 系数,进行小波反变换,就可得到图 8(b)所示的子图象.从视觉效果来看,这个子图象要比 LL_2 子带图象中方框部分大了一倍(空间分辨率提高了),而且清晰了,已经能够看出眼睛和眉毛(信噪比提高了).如果再加上更高分辨率的子带图象(LH_1 、 HL_1 、 HH_1)相应部分的数据,与图 8(b)的数据一起进行小波逆变换,可以得到图 8(c)所示的图象.现在,很多细节部分,例如瞳孔 and 睫毛,都能清楚显现,与原图象(图 7(a))相应部分的内容并无二致.

图 9 是用图象的部分 DWT 系数分级重建图象局部的另一例子.图 9(a)是 1024×1024 的原图象(被缩小八分之一显示),图 9(b)是原图象经过 3 次小波变换后的 LL_3 子带图象,并作为原图象的预览.图 9(c)是对图象中碑帖部分的分级重建.可以看到,“石林”二字愈来愈大(空间分辨率愈来愈高),也愈来愈清楚(信噪比愈来愈大).



(a) Lena 原图



(b) 二次小波变换系数图

图 7 Lena 图象的小波变换



(a)

局部窄带 级小波逆变换



局部窄带 级小波逆变换



(c)

(b)

图 8 Lena 眼睛部分的分级重建

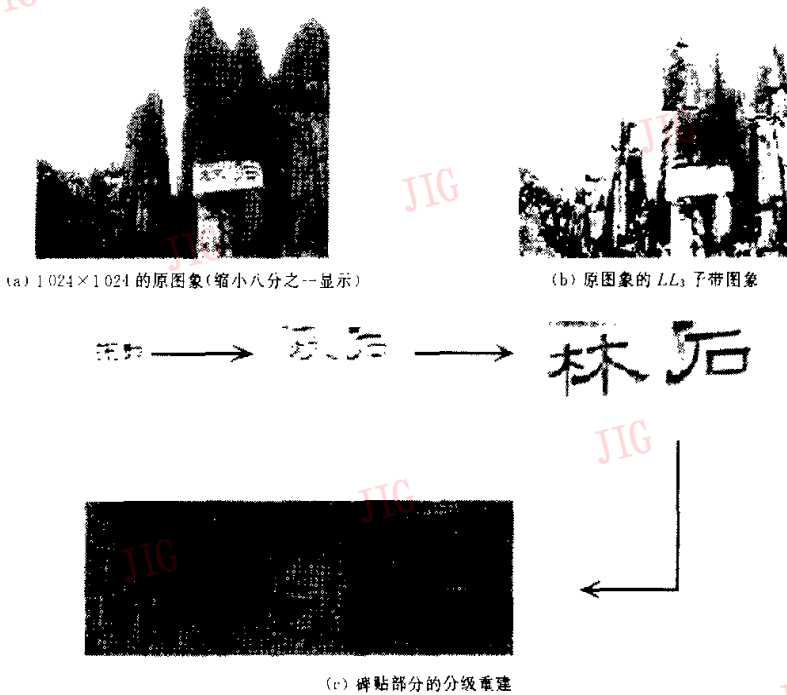


图9 基于 JPEG2000 的图象局部可分级重建的例子

3 结束语

(1) JPEG 自 1992 年颁布以来得到了广泛的应用. 例如, 许多视频编码国际标准中的帧内编码都是采用 JPEG 方法. 但是, 斗转星移, JPEG2000 作为比 JPEG 更为优秀的、新的图象压缩编码国际标准, 正在许多应用场合逐步取代 JPEG, 并且还开辟了新的应用领域. 可以预言, 随着 JPEG2000 的颁布和推广, 有关 JPEG2000 应用研究的热潮必将风生水起.

(2) JPEG 与 JPEG2000 同属变换编码, 但是, JPEG 所用的变换是 DCT (离散余弦变换), 而 JPEG2000 所用的变换是 DWT (离散小波变换). DCT 是一种从空间域到频率域的变换, 图象的 DCT 系数只反映图象的频率成分, 与图象内容的空间位置无关. DWT 是一种从空间域到空频域的变换, 图象的 DWT 系数不但反映图象的频率成分, 而且还与图象内容的空间位置有关. 因此, 人们可以仅从图象的部分 DWT 系数恢复图象的局部内容, 这也就是基于 JPEG2000 的图象局部可分级检索中“局部”的含义.

(3) 基于 DCT 的 JPEG 只能提供图象的质量

可分级编解码, 即所谓“基于 DCT 的累进编解码模式”. 基于 DWT 的 JPEG2000, 利用 DWT 的频域属性, 能够提供图象的质量可分级编解码; 利用 DWT 的空域属性, 又能够提供图象的空间可分级编解码. 基于 JPEG2000 的图象局部可分级检索充分利用 DWT 的空频二重属性, 对图象的局部区域提供空间和质量同时可分级的编解码, 从而产生 Zoom-In 的视觉效果.

(4) 图象可分级编解码的关键是压缩码流的布局. JPEG2000 规定, 图象的 DWT 系数可以按照零树结构排列(纵向排列), 也可以按照频带/分辨率排列(横向排列). 基于 JPEG2000 的图象局部可分级检索先把图象的 DWT 系数按照零树结构(即小波系数块)排列, 零树结构中的 DWT 系数再按照频带排列. 这样, 每个零树结构对应着图象的一个局部区域, 零树结构中处于不同频带的 DWT 系数则对应着这个局部区域不同级别的重现.

(5) 基于 JPEG2000 的图象局部可分级检索中, 图象区域被划分为一个个网格, 每个网格与图象小波系数的一个零树结构相对应. 在检索时, 一旦发现目标, 就要确定目标所在的网格. 如果目标占据多

个网格,为了避免边界效应,不是逐个网格来重现目标,而是把先用这些网格组建一个更大的网格(如果这些网格不能形成一个矩形网格的话,还必须把邻近的网格也一并考虑),然后把原来分散在各个零树结构中小波系数重新排列,组成一棵更大的零树,通过小波反变换,重现目标。

参 考 文 献

- 1 ISO/IEC JTC1/SC29/WG1, Document N1646R, JPEG2000 Image Coding System[S], Mar., 2000. <http://www.jpeg.org/FCD15444-1.htm>
- 2 Sodagar Iraj, Lee Hung-Ju, Hatrack Paul *et al.* Scalable wavelet coding for synthetic/natural hybrid images [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(2):244~254.
- 3 Chai Bing-Bing, Jozsef Vass, Xinhua Zhuang. Significance-linked connected component analysis for wavelet image coding [J]. IEEE Transactions on Image Processing, 1999, 8(6): 774~784.

- 4 Shapiro Jerome M. Embedded image coding using zerotrees of wavelet coefficients [J]. IEEE Trans on Signal Processing, 1993, 41(12):3145~3459.



钟敏生 1976 年生,中山大学通信与信息系统专业硕士研究生,研究方向为多媒体数据处理与传输。



马争鸣 1957 年生,1989 年获清华大学模式识别与智能控制专业博士学位,现为中山大学电子系副教授,主要学术兴趣为小波分析、分形几何和人工神经元网络。

第三届“Web 时代信息管理(WAIM'2002)”国际会议征文通知

第三届 Web 时代信息管理国际会议(WAIM'02)将于 2002 年 8 月 11 日至 13 日在北京召开。该会议由中国计算机学会数据库专业委员会和 ACM SIGMOD 主办,中国人民大学和清华大学协办。会议论文集由德国 Springer 出版社作为《Lecture Notes in Computer Science》发表,优秀论文将推荐到 JCST 杂志上发表。会议得到国家自然科学基金委的支持。WAIM 国际会议旨在促进中国数据库界与国际上的交流,向世界介绍中国最新的研究成果,并使国内的研究水平保持与世界同步。会议将邀请国际知名学者做辅导讲座和特邀报告。

会议主题:

讨论和交流国际范围内在“Web 时代的信息管理”这一研究领域的最新科研成果,主要包括:

高级数据库应用	数据安全	互操作性和异构系统	空间和时态数据库
代理技术及信息管理	数据库语言及语义学	位置相关的数据管理	文本数据库及信息检索
生物及基因信息系统	数据库与电子商务	元数据的使用和管理	事务管理
协同信息系统	数据库及数据仓库设计	多媒体信息系统	数据库结构与内容可视化
数据和信息质量	新兴 Web 技术	并行和分布式数据库系统	Web 和数据库技术
数据移植及集成	金融及工程数据库	性能及基准	工作流
数据挖掘及知识发现	基于英特网和 Web 的数据管理	查询处理与优化	XML 和半结构化数据
数据仓库	多维数据库及联机分析处理	存贮管理和存取方法	

投稿要求:

稿件应反映有关 Web 时代信息管理的最新理论、技术、产品或应用成果,并未在其他会议或杂志上公开发表过的论文。本次会议特别征集系统演示(Demo)投稿,详见会议网址。

稿件一律用英文,采用 11pt 以上的字体,双行(double space),不超过 20 页,要求电子邮件投稿,请用 PDF 或 PS 文件。

重要时间:

论文摘要截止日期:2002 年 2 月 22 日;论文全文截止日期:2002 年 3 月 1 日;系统演示截止日期:2002 年 3 月 1 日;论文录用通知日期:2002 年 4 月 15 日;论文付印稿日期:2002 年 5 月 20 日

有关信息:

有关大会程序:中国人民大学信息学院 孟小峰 邮编:100872
Tel/Fax: 010-62519453 Email: xfmeng@public.bta.net.cn;
大会网址:<http://www.cs.ucsb.edu/~waim02>; <http://www.cs.ust.hk/waim/>