

基于小波的视频图像压缩研究进展

张宗平, 刘贵忠

(西安交通大学信息与通信工程系, 陕西西安 710049)

摘 要: 近十几年来, 随着小波理论研究的不断深入, 小波图像压缩的应用研究日臻成熟, 在此背景下, 基于小波的视频图像压缩的理论和应用研究在业界已引起广泛关注. 本文在对近十几年来小波图像压缩研究的回顾和总结基础上, 着重介绍了小波技术在视频图像压缩应用研究的最新进展, 探讨了目前小波视频压缩研究的若干热点问题及其今后的发展方向.

关键词: 小波; 视频编码; 基于小波的视频图像压缩; JPEG2000

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2002) 06-0883-07

Advances in Wavelet-Based Video Compression

ZHANG Zong-ping, LIU Gui-zhong

(Department of Information & Communication Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China)

Abstract: The application research of Wavelet-based image compression is becoming mature as the wavelet theory research proceeds in last decades. Under this background, the wavelet-based video compression research has drawn much attention in this field. Based on summarizing the wavelet-based image compression techniques for the last decades, this paper describes the advances of wavelet-based video compression in details, and then discusses several hot problems as well as their development in future.

Key words: wavelet; video coding; wavelet-based video compression; JPEG2000

1 引言

近十几年来数据、语音和多媒体业务的需求在不断增加, 同时人们对这些业务的服务质量的期望也越来越高. 为了有效防止信号在传输和存储过程引入噪声和导致波形畸变, 获得更好的品质, 模拟信号(如语音、音频、图像和视频信号)一般都先进行采样和数字化, 然后再进行存储、传输和接受重建. 但这些数字化信号的数据量极大, 尽管海量存储技术、处理器的速度以及数字通信系统的性能的迅猛发展, 但对数据存储的能力和数据传输带宽的需求仍然超出了现有技术的能力所及. 为了使通信成为可能并尽可能地降低通信费用, 信号的压缩是必需的. 近几年在已经或将要出现的如视频会议、可视电话、高清晰度电视、交互式电视和远程医疗等诸多商业应用的推动下, 图像和视频压缩编码受到了极大的重视. 随着 Internet 网、电信网、广播电视网的逐步融合, 图像和视频压缩编码也成为图像通信的中心问题之一.

Fourier 变换是信号处理领域中研究最早应用最广的分析工具. 图像和视频压缩中最常使用的工具是一种简化了的 Fourier 变换——DCT 变换^[1,2]. 基于分块 DCT 变换的压缩编码技术是已有图像和视频压缩标准的核心技术^[3~7], 这主要来自两个方面的原因: (1) DCT 具有良好的去相关性和能量压缩特性; (2) DCT 变换存在快速实现算法. 但随着应用和研究的不断深入, 分块 DCT 变换编码的缺点逐步暴露出来, 尤其

在低比特率环境下, 压缩图像不可避免地出现了方块效应和飞蚊噪声. 这是因为一般情况下图像信号是高度非平稳的, 很难用 Gauss 过程来刻画, 并且图像中的一些突变结构例如边缘信息远比图像平稳性重要, 用余弦基作图像信号的非线性逼近其结果不是最优的^[8].

信号压缩就是要用尽可能少的比特数表示一个信号, 因此要提高信号的压缩效率首先要给出信号的稀疏表达式. 数学界和工程界在共同研究数据表示技术的过程中所发展起来的小波分析技术和多分辨分析理论^[8~12], 摒弃了传统 Fourier 分析所必须的前提假设——平稳性, 成为分析非平稳信号的有力工具. 它的出现导致了我们从新的视角去研究信号压缩、噪声滤波等信号处理问题: 一方面, 由于小波基的紧支性和小波分解的多尺度结构, 非线性小波逼近实质上等价于一个自适应的网格逼近, 网格的分辨率在信号奇异点的邻域内被适当加细了; 另一方面, 由于小波基的无条件基特性, 使它成为一大类信号的非线性逼近的最优基, 许多信号在小波基的表示下, 都可以获得稀疏的表示式. 因此自从八十年代小波技术诞生以来, 伴随着小波理论研究的不断深入, 在小波理论的指导下基于小波的图像应用研究取得许多成果, 这些成果正逐步标准化, 汇集成拟订中的工业标准——JPEG2000^[13]. 在此背景下, 小波视频压缩业已成为当前视频压缩的研究热点, 可以预见基于小波的视频压缩技术将是未来视频压缩标准的一

收稿日期: 2000-10-21; 修回日期: 2001-04-16

基金项目: 国家自然科学基金(No. 69872030); 国家教育部高校骨干教师资助计划(No. 2000G17); 国家教育部高校博士点基金(No. 2000069828)

项核心技术。

2 小波图像压缩研究进展

早在 1986 年, Woods^[14]等人就构造了第一个真正意义上的小波图像编码器, 但那时在图像领域人们却把小波编码称为子带编码。因为在当时人们还没有完全搞清楚子带编码的工作机理, 更不清楚它与数学领域的小波基有着千丝万缕的联系, 直到多分辨分析理论^[9, 12]提出, 信号处理领域的专家们才从理论上统一了众多子带编码器的工作原理。子带分解实质上是小波分解, 基于子带的图像编码实质上是基于小波的图像非线性逼近, 因此后来的“子带编码”和“小波编码”两个词就不再加以区别。伴随着数学领域的小波理论的发展和非线性逼近论研究的深入, 小波图像压缩取得了极大的进展。一个极具表现力的实证是 Daubechies^[10]所发现的小波双正交基立即被转化为图像分解所使用的具有线性相位多通道完全重建滤波组。自从小波理论引入图像处理领域, 图像处理工程师们就开始逐渐从非线性逼近论和泛函分析的角度出发, 分析、设计和改进图像压缩应用算法。

2.1 基于小波的静止图像压缩

图像的小波变换具有内在的多尺度结构, 因此基于小波的图像压缩无需对图像进行分块, 可以消除基于块变换而所导致的方块效应。小波基的紧支集特性和小波变换的多尺度展开结构使得小波变换同时又是一种时频分解, 每个小波系数都只反映图像的一个特定的空间范围和特定的频率范围内的信息。一般说来, 一幅典型自然图像主要是由齐次区域和纹理区域以及相当少的视觉敏感的边界组成的。齐次区域的跳变很少, 主要是由低频成分组成; 纹理区域则具有一定的跳变, 因此这部分区域既含有部分高频成分又包含部分低频成分; 而边界区域则具有极大的跳变, 这部分区域则集中了图像绝大部分的高频能量。因此, 小波变换将图像的绝大部分能量压缩到低频子带中, 只有少数的能量分布在个高频子带。结果图像小波变换系数具有下面几个统计特性:

- 空-频局部化;
- 能量压缩特性;
- 子带内小波系数的聚类特性;
- 子带间小波系数的相似性;
- 小波系数幅度从低频子带到高频子带的衰减特性。

而且所有小波系数在作非线性逼近时都具有(大致)相同的失真度, 这些特性已被证实对图像压缩是至关重要的。图 1(b) 是基于张量积小波对原始 Lena 图像进行 Mallat 塔式小波分解结果, 读者由此可窥一斑。

2.1.1 小波图像压缩新技术 近几年小波图像压缩取得了很大的成功, 这些算法都是成功地使用了各自不同的具有创新性的策略更好地组织和表示了塔式结构的小波系数, 从而更好地和更完整地利用了小波系数所具有的统计特性。由于各种算法所使用的具体方法不同, 因此无法统一给出小波压缩算法结构模式, 但总结起来, 这些算法^[15~18]主要利用了下面几种技术:

(1) 空间小波树^[15, 16]

图 1(d) 演示了图像小波分解后的空间小波树结构。由于

小波系数的空-频局部化特性可知, 图 1(d) 中的每棵空间小波树实质上定义了图 1(c) 所示的一个原始图像块。因此, 每棵空间小波树与分块 DCT 变换的方块地位相当, 但它们又有着本质的区别:

(a) 每棵空间小波树随小波分解级数的改变而发生变化, 因此它所定义的空间图像块的尺寸是可调的;

(b) 由于小波变换是处理整幅图像, 每相邻两个小波系数是来自原图像两个相互重叠区域, 因此这些空间小波树所代表的图像块是相互重叠的。

(c) 每棵空间小波树小波系数与分块 DCT 变换系数相类似, 但空间小波树上的小波系数却有着显著统计特性。

空间小波树的三个不同分支分别表示了图像中三个不同方位信息: 水平分支表示了图像中的水平信息; 垂直分支表示了图像中的垂直信息; 对角分支则代表了图像中的对角信息。

空间小波树上小波系数存在一定相似性, 这实质上隐含着系数高阶相关性。

空间小波树上的节点分散在不同子带, 可以有效地支持视觉带权量化。

空间小波树上系数幅值按频带的衰减规律, 可以有效地支持渐进式编码和分层量化。

空间小波树的多尺度特性, 可以支持多分辨图像编码。

因此利用空间树的特性可以有效地进行感兴趣区域(ROI) 编码、多分辨编码和嵌入渐进式编码, 这些都是人们期待已久的编码方式。

(2) 比特面编码技术^[15, 16, 18]

传统的编码器如 JPEG 和 MPEG-1, 2 一般先量化, 然后对每个量化系数进行熵编码。嵌入式编码器则不同, 嵌入式编码采用的是逐次量化、渐进编码的方法。它是先把待编码信源排列在比特面上, 优先编码所有信源的高比特位。对于变换的小波系数, 进行比特面编码必须设法编码三类信息, 即位置信息、系数符号信息和系数幅度信息。若规定比特面编码在塔式结构的小波系数的执行顺序, 则位置信息可以隐藏在预定的编码执行路径中, 因此基于空间树的小波系数无需显式编码位置信息, 但预置的隐式编码路径会对编码器的性能产生决定性的影响^[15~17]。每个系数只有一个符号, 并且系数符号只有在系数幅值非零的情况下才有意义, 因此符号编码一般在其系数幅值编码结束后有条件地进行。

(3) 基于上下文的自适应算术编码技术^[13, 17, 18]

同 JPEG 和 MPEG-1, 2 一样, 小波图像压缩编码的最后一步是无失真的熵编码, 所不同的是小波图像编码器一般都使用算术编码器^[19, 20]而不是 Huffman 编码器。算术编码器较 Huffman 编码器优越之处在于它可以耦合一定的模型去推断待编码符号的概率, 这个特点对小波图像编码器就显得特别重要。这是因为小波系数如此多的统计特性使得我们有可能设计出较好的统计推断模型, 更好地估计编码符号的概率, 进一步降低信源熵。

一般说来, 固定模型的算术编码器只适合平稳信源的编码, 而自适应算术编码器每编码一个符号就会更新相应的条

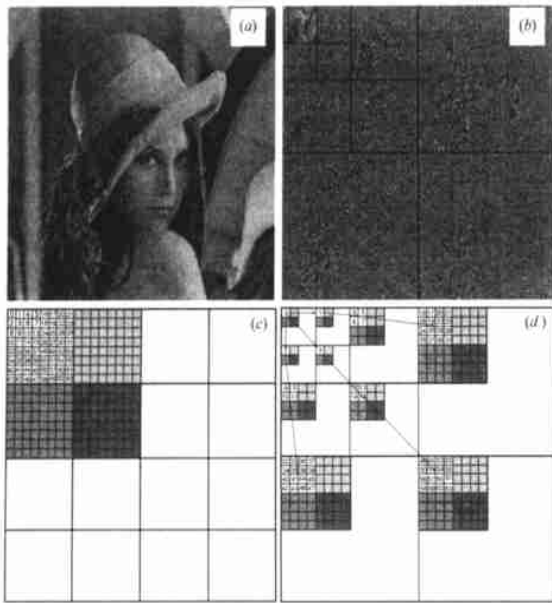


图1 Lena 图像的小波分解及空间结构。(a) 原始图像;(b) 小波分解结果;(c) 对应的图(d) 原始图像块;(d) 对应图(b) 空间小波树

件概率估计,因此它特别适合非平稳信源的编码。而对于非平稳信源如图像所产生的数据流,条件概率分布或者说局部概率分布会随所处的区域位置不同会有所变化,使用局部自适应模型所收集到的局部概率估计则可以很好的跟踪局部统计变化,比全局自适应模型所收集到的概率估计更稳健。因此对于图像这类高度非平稳信号,多数小波图像编码算法都使用基于上下文的局部自适应模型的算术熵编码。模型上下文通常是由当前编码像素周围节点状态或和其父节点的状态链来确定。一般说来,由状态链所定义模型数目极大,而这些模型的统计推断通常是高度聚类的,因此实际编码一般要将这些模型进行归并成为数很少的几个典型模型,以有效地提高算法执行效率,同时又不致损失其压缩效率。

2.1.2 几个典型的小波图像编码器 随着小波图像压缩应用研究的深入,各种类型的小波图像编码器相继提出,其中下面三种小波图像编码器因其卓越的性能常常被用着比较标准。

(1) EZW 编码器^[15]

1993 年,Shapiro 根据小波系数的统计分布规律,提出了一个假设:如果空间树的粗尺度的小波系数对 T 门限是不重要的,则在这棵小波树的细尺度的小波系数也极可能是不重要的。基于此 Shapiro 引入一个小波“零树”,通过定义 POS、NEG、IZ 和 ZIR 四种符号进行空间小波树递归编码,有效地剔除了对高频系数的编码,极大地提高了小波系数的编码效率。这个算法采用的是渐进式量化和嵌入式编码模式,算法复杂度低。EZW 算法的出现打破了信息处理领域长期笃信的准则:高效的压缩编码器必须通过高复杂度的算法才能获得,因此 EZW 编码器在数据压缩史上具有里程碑意义。

(2) SPIHT 编码器^[16]

由 Said 和 Pearlman 提出的分层小波树集合分割算法

(SPIHT) 则利用空间树分层分割的方法,间接实现空间小波树的比特面排序,有效地减小了比特面上编码符号集的规模。虽然 SPIHT 算法实现与 EZW 算法有较大差别,但其核心思想与 EZW 如出一辙。同 EZW 相比较,它的主要成功之处还在于构造了两种不同类型的空间零树,从而可以更好的利用小波系数的幅值衰减规律。同 EZW 编码器一样, SPIHT 编码器的算法复杂度低,产生的也是嵌入式比特流,但编码器的性能较 EZW 有很大的提高。

(3) EBCOT 编码器^[18]

可扩展图像压缩是指一次压缩编码流可作多种不同类型的解码。最常见的是 SNR 可扩展和分辨率可扩展, EZW 和 SPIHT 的嵌入式编码实现了 SNR 可扩展,分辨率可扩展也是目前人们所期望的一种压缩码流品质,尤其在大图像的远程浏览和图像数据库检索等应用领域。为了能够支持这种类型的解码特性, Taubman 使用了优化截断点的嵌入块编码方法(EBCOT)对图像进行编码。在算法实现上 EBCOT 编码是先将每个子带分成一个个相对独立的码块,然后使用优化的分层截断算法对这些码块进行编码,产生压缩码流,结果图像的压缩码流不仅具有 SNR 可扩展而且具有分辨率可扩展,还可以支持图像的随机存储。比较而言, EBCOT 算法的复杂度较 EZW 和 SPIHT 有所提高,由于是采用优化截断方式, EBCOT 的压缩性能比 SPIHT 略有提高。

2.2 图像编码新标准——JPEG2000^[13]

自从十年前 JPEG 制定以来,图像压缩编码取得了很大的进展,小波编码技术或者说子带编码技术日趋成熟。同时随着应用需求不断提升,原有的 JPEG 标准的不足逐渐暴露出来:如方块效应(尤其在低比特率下),没有足够的可扩展功能(如支持空间可扩展和品质可扩展),缺少基于对象和区域的表示,也不具有容错性能。鉴于此,1996 年 JPEG 专家组着手制定图像压缩新标准——JPEG2000。期望 JPEG2000 系统能够解决具有不同特征(自然场景,医学,遥感,文本以及复合文档等等)的各种类型静止图像(包括二值图,灰度图,彩色或多光谱图像)压缩问题,同时允许使用各种不同的模式(如客户机/服务器,实时传输,图像库查询,有限缓冲和带宽资源等等)。同基线 JPEG 相比, JPEG2000 至少要提高 30% 的压缩效率,并且要支持更多的基线 JPEG 所没有提供的高级功能选项,如同时支持无损压缩和有损压缩,空间和品质可扩展编码,容错编码, ROI 编码,顺序组合、拼接以及码流的随机存取和处理等。

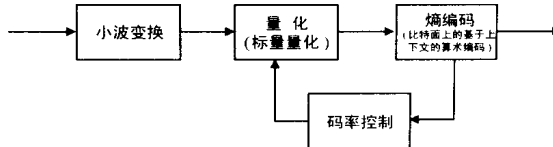


图2 JPEG2000 的编码结构框图

图2 是 JPEG2000 的编码结构框图,输入图像首先进行小波变换,然后对小波变换系数量化并进行无损熵编码。它可以支持不同的小波和小波包分解模式,其中小波变换滤波器有两种选项,一种是支持无损压缩编码的整数定点小波滤波器(B57),另一种是支持有损压缩编码的浮点小波滤波器

(B97). 量化器既可以是统一的死区标量量化器也可以是模型量化器. 熵编码选用的是比特面上基于上下文的自适应算术编码. 每个比特面有三个编码通道, 即重要性演化通道 (Significance Propagation Pass), 幅值细化通道 (Magnitude Refinement Pass) 和清除通道 (Cleanup Pass). 重要性演化通道是针对那些邻域内存在重要系数但其本身却是不重要系数的节点, 幅值细化通道则是进一步细化那些重要系数的节点, 清除通道则是处理前两个通道所没有处理的剩余所有节点. 这三个编码通道是由四个编码操作即重要性编码 (Significance Coding)、符号编码 (Sign Coding)、幅值细化编码 (Magnitude Refinement Coding) 和清除编码 (Cleanup Coding) 来协调完成. JPEG2000 采取的是基于码块的分层编码方法, 基本操作与 EBCOT 算法类似, 最后采取分组打包方法产生编码流以支持容错.

静止图像编码新标准——JPEG2000 可以看成是近十年来小波理论和应用研究在工程实践中的一个具体结晶, JPEG2000 新标准成功地集成了近年来小波图像压缩的研究成果, 基本完成了 JPEG2000 所预期的目标.

3 小波视频图像压缩研究进展

由于小波压缩技术在消除方块效应和飞蚊噪声的特殊性能和良好的率失真品质, 自从小波引入图像压缩领域以来, 基于小波的视频压缩研究一直在不断进行. 在这部分首先对近几年小波视频图像压缩研究现状做一个总结, 然后再探讨一下小波视频图像压缩的几个热点研究问题.

3.1 小波视频图像压缩算法

如果给图像编码方案提供一个消除时间冗余的机制, 一个普通小波图像编码器就可以扩展成视频图像编码器. 根据时间冗余的利用不同, 目前出现的基于小波变换的视频压缩算法主要可以分为以下三种不同类型.

3.1.1 算法类型之一: 基于空域运动补偿的小波视频编码 (MC-DWT) MC-DWT 视频编码方案是采取空域运动估计和运动补偿技术消除时间冗余, 然后使用整域小波变换消除空间冗余. 与传统的基于 DCT 的混合编码的主要差别在于对补偿帧差的处理, 它是用整域小波变换代替分块 DCT 变换. 如果使用的是非重叠的块补偿技术, 这种简单的取代并不会带来较大的改进, 其原因是显而易见的, 基于非重叠块运动补偿将在补偿帧差内产生块状的虚象, 这对于整域的小波变换来说, 显然不利于信号能量的压缩. 比较而言, 重叠块运动补偿技术 (OBMC) 不仅可以有效地消除方块效应, 而且能够有效地降低补偿帧差的能量提高运动补偿效率, 极大改善视频编码器的性能. 因此近来基于空域运动补偿的混合小波视频编码广泛采用 OBMC 技术.

1996 年, David Sarnoff 研究中心的研究人员在 Shapiro 的 EZW 算法基础上, 提出了一个基于小波域空间分块的自适应系数单次量化编码和优化算术熵编码的“零树熵编码” (ZTE)^[21] 视频图像压缩算法. 虽然 ZTE 失去了嵌入特性, 但它的压缩效率比 EZW 有了很大的提高, 尤其是低码率条件下. ZTE 视频编码器在基于对率失真 LR 分析, 提出了一个速率控制方案: 当前帧的编码比特数为前一帧编码比特数和平均目标比特数的平均值, 然后根据所分配的比特数作帧内自适应

量化和编码. 这个编码器的性能与 H. 263 的编码性能相当, 但它的最大优点在于可以支持基于对象编码和空间可扩展编/解码, 实现 MPEG-4 的部分功能. 后来 D. Marpe 和 L. Cycon^[22] 在进一步优化信源分离、改进零树映射和使用局部自适应熵编码等予编码处理基础上提出了一个 PACC 小波视频压缩编码算法. 他们针对低码率图像压缩条件, 对零树的失效性作了进一步分析和处理, 结果 PACC 视频编码器的性能比 ZTE 视频编码器的性能有了进一步的提高. 从本质上看, PACC 编码器的帧内编码结构是对 EZW 和 SPIHT 综合优化的结果.

Vass^[23] 等人则不同, 他们将 B. B. Chai 等人提出的 SLCC 图像压缩编码器推广到传统的混合视频编码方案中, SLCC 算法是近年来在小波域利用形态算子进行图像压缩的最成功的范例之一. 其成功之处在于利用了形态算子来描述数字图像小波域内高频子带能量的边缘聚集特性, 它不仅利用子带内重要系数的聚类特性而且利用了子带间的重要系数聚类相似特性, SLCC 算法采用了与零树编码器类似的方法对相邻频带间的重要系数进行父子连接以此提高编码效率. SLCC 算法与 EZW 和 SPIHT 算法的最大区别在于, 它利用的是重要系数的父子相关性而不是 EZW 和 SPIHT 所利用的不重要系数的父子相关性. 比较而言 SLCC 算法可以更好地编码具有丰富纹理信息的图像, 结果 Vass 等人提出的 SLCC 视频编码器的压缩性能比 ZTE 视频编码器的性能有了显著的提高.

这类编码器的差别仅在于运动补偿残差的小波变换编码. 近来随着小波图像压缩算法研究的不断深入, 这类编码器的在基低比特的压缩性能, 无论主观品质还是客观品质大都超出 H. 263 编码器性能.

3.1.2 算法类型之二: 基于变换域运动补偿的小波视频编码 (DWT-MC) DWT-MC 视频编码方案是先使用整域小波变换小波消除空间冗余, 再通过小波域的运动估计和运动补偿技术消除时间冗余, 最后对变换域的补偿残差进行小波压缩编码. 由于小波分解内在的多分辨结构, 在小波域内很容易执行多分辨运动估计和运动补偿 (MRMC), 同传统的时域运动估计相比可以节省大量的搜索匹配时间.

但不幸的是, 由于小波变换不具有空间平移不变性, 因此在小波变换域内由于无法精确界定相邻帧间的平移运动而导致的大量误配预测, 大大地削弱了运动补偿效果, 因此最初由 Y.-Q. Zhang^[24] 提出的基于小波域的运动补偿视频编码器只适用于中高比特率的视频图像压缩, 如 HDTV 等应用领域, 低码率条件下的压缩性能极差. 最近 H.-W. Park^[25] 提出了一种改进的小波域运动估计和运动补偿技术, 即所谓基于低频子带平移方法的运动估计 (Motion Estimation Using Low-Band-Shift Method, ie. LBS-ME). LBS-ME 方法的小波域估计的搜索对象集实质上是一个未执行下抽样的二进小波变换集, 小波域的运动估计等价于空域的全域搜索. 由于小波基的有限支撑性, LBS-ME 等价于空域的重叠块估计, 而且是随小波基的不同而自适应的, 它的估计效果达到了最佳. 实验结果表明 LBS-ME 方法比空域的运动估计和小波域的直接运动估计都要好得多. 但这种改进方法也存在一个严重的缺陷, 那就是它需要较多缓冲帧存, 而且匹配搜索的时间较长.

最近 Xuguang Yang 和 K Ramchandran^[26] 提出一种基于正反双向联合运动估计的小波域运动补偿方案. 通过对低分辨率图像进行内插上抽样和反混叠滤波得到精细尺度的逼近像, 再用这个逼近像作反向运动估计得到相应子块的估计运动矢量, 以此估计运动矢量去预测当前帧的精细尺度逼近像, 然后再进行小波分解得到相应分辨率的高频子带的预测系数. 这种反向估计无需编码运动矢量, 可以提高编码效率, 但它也存在一个严重的缺点——容错性能差, 容易受噪声干扰, 尤其在低比特条件下, 其运动估计性能就会下降. 为了改善其性能, 他们又使用正向估计方法去估计相应子块的运动并记录相应运动矢量. 小波域内的逼近图像双向估计策略有效的避免源于小波变换导致高频子带估计失真问题, 运动补偿效果得到很大的提高, 获得了比 H. 263 更好的压缩效果.

目前看来要提高这类编码器压缩性能, 应设法适当减小小波域的搜索匹配集合, 优化小波域的运动估计策略, 提高小波域运动估计速度.

3.1.3 算法类型之三: 含运动补偿的三维小波视频编码 (MC-3DWC) 基于三维小波变换的视频编码是二维空间小波图像压缩向三维视频空间的简单推广. 值得注意的是: 视频序列在时间方向上存在着很强的相关性, 但这些相关性明显不同于空间相关性. 这表现在: 当前帧一般是由先前帧的部分区域变化的结果, 引起这些变化有许多可能因素, 如摄像机的移动, 镜头聚焦的变化、场景中物体的运动以及照明条件的变化等等. 如果不考虑这些因素, 简单地使用三维小波变换进行视频图像压缩, 编码器的性能将大打折扣. Taubman 和 Zakor^[27] 提出的多速率三维子带 (小波) 编码器 (3DSBC) 关注更多的是编码器的可扩展性, 在时间分析上仅考虑到相机的整体运动. Ohm^[28] 提出的运动补偿三维子带编码器 (MC-3DSBC) 的时间方向上的变换采取的是沿运动轨迹进行分析/合成的滤波方案, 其运动轨迹是由传统的块匹配方法确定的. 这个编码方案较前一种在时间分析方面有了很大的改进. 后来 Choi 和 Woods^[29] 等人对这个时间分析系统作了进一步完善, 调整了相应的三维变换结构并优化了相应码率分配. 改进后的 MC-3DSBC 编码器性能有了很大的改善.

其次对视频三维小波 (包) 的空间数据结构利用对视频图像压缩也有着决定性的影响. J - Y Tham^[30] 等人通过改进三维小波 (包) 分解, 基于一种所谓三重零树结构 (Tri-Zerotree, i.e. TRF-ZIR) 提出的高度可扩展三维小波视频编码器, 这也是一种 MC-3DWC 编码器. 它将 Shaprio 的 EZW 编码算法改进推广在 TRF-ZIR 结构树中, 通过组合分层量化 (渐进式) 编码技术和嵌入分辨率块编码技术, 将编码比特流划分成不同的子段, 以支持不同空间分辨率、不同帧率和不同品质的多速率的解码输出. 视频编码性能客观评价同 H. 263 相当, 但主观品质优于 H. 263. TRF-ZIR 视频编码器的时间分析滤波忽略了对非关联像素的细致处理, 并且他们的时间分析/合成系统不是完全重建的, 因此 TRF-ZIR 视频编码器不适合高品质的视频压缩. Rensselaer 的图像处理研究中心的 Kim 和 Pearlman^[31, 32] 等人最近基于三维空间八叉树将二维 SPIHT 算法推广到三维空间, 提出了一个三维小波视频编码方案, 实验结果表明中高码率

条件下, 视频压缩结果图像要优于 MPEG2, 但低码率条件下的压缩结果略逊于 H. 263. 我们的实验结果表明^[50]: Pearlman 对三维空间数据结构的利用还不够有效, 八叉树特别不利于时间低频子带的空间高频系数的编码, 通过优化定义空间树的定义, 我们得到了比 H. 263 更优的结果.

由于三维小波变换的视频编码器在支持容错和可扩展码流输出方面的显著特性, 它一直是国外的研究热点之一, 这些算法已初见成效, 高码率压缩条件下性能优于 MPEG2, 但在低码率压缩条件下, 与 H. 263 相比效果不是非常明显.

3.2 小波视频压缩的发展状况

总起来看, 目前小波视频压缩技术远不及小波图像压缩技术成熟. 小波视频图像压缩研究之所以滞后于小波图像压缩研究不仅仅因为小波视频图像的进展有赖于小波图像的研究和信号维数的提高, 还由于视频压缩的应用背景与图像压缩的应用背景有很大的不同. 由于视频应用实时性的原因, 视频压缩算法的要对压缩算法的复杂度和压缩效率作综合考虑. 实际上, 推动小波视频压缩研究不断深入主要有两个技术因素, 一个是小波图像压缩技术研究, 另一个是运动补偿技术研究. 另外不断激增商业需求也在不断刺激小波视频压缩理论和应用研究的不断深入. 从目前的业界研究水平来看, 下面几个研究热点在未来一段时间仍将继续.

3.2.1 研究热点之一: 小波图像压缩新技术的研究 小波图像压缩技术虽然已经取得了很多成果, 但目前尚无一种方法可以利用小波系数的所有统计特性, 目前这些成功的算法都只是侧重对一部分统计特性的利用. 另外部分研究已证明小波压缩同其他技术如分形压缩技术^[33] 和矢量量化技术^[34, 35] 的结合也会产生更高效的压缩性能. 数学形态学近来在视频和图像压缩中已显示出了相当的成功^[17, 36], 但这些算法还有待进一步深入研究. 当然小波理论自身的发展也是推动小波图像/视频压缩的重要因素, 第二代小波如提升小波、内插小波、区间小波、M 带小波和非线性小波等的研究在改善图像压缩品质和提高执行速度方面都产生了积极效果^[37~39], 形状自适应小波变换^[40] 在图像和视频压缩中已引起人们的重视, 因此小波图像压缩新技术的研究, 目前和未来很长一段时间仍将是图像和视频压缩应用研究的热点.

3.2.2 研究热点之二: 运动补偿技术的研究 基于混合方案的小波视频编码器的空域运动补偿技术目前只是对传统混合视频编码器的简单继承, 缺乏针对性, 对小波变换域的运动补偿技术还有待于进一步研究, 在算法复杂度和匹配精度方面作出适当折衷. 三维小波的运动补偿目前的研究不够深入, 目前的应用也仅限于传统的匹配方法的简单应用, 如果能够有效地提高对运动矢量的编码, 有望提高这类编码器在低码率的压缩性能.

作为视频压缩的核心技术之一的运动补偿技术, 近几年也取得了很大的进展^[41~46], 其中可变块运动补偿技术较传统的固定块运动补偿技术在率失真性能方面有较大提高, 值得注意的是目前这些研究大都只是针对传统的混合编码器提出的, 它们在小波视频压缩编码器中的研究所见不多. 因此应用这项技术有望进一步提高小波视频编码器的性能. 普遍认为

运动补偿技术是一种非线性技术,运动矢量的编码无法耦合到运动补偿残差编码统一框架体系中,因此大多数小波视频编码器对运动矢量和补偿残差分开进行。在码率约束条件下这两者之间不是没有关系的,目前已有部分文献开始着手解决这类问题^[29]。

3.2.3 研究热点之三:利用人的视觉生理特性的进行小波视频图像压缩的应用研究^[47] 人对视频图像的主观评价也明显不同于对图像的主观评价,由于每帧视频在视场中暂留时间极短,人脑接受的是帧更新的动态信息,因此考虑人脑的视频图像认知结构的生理特征也可能提高视频压缩性能。虽然静止图像压缩考虑了人眼的视觉频率响应特性,但对视频图像这些方法可能不完全适用,因此考虑人对视频图像的视觉心理感知特性有可能进一步提高目前的小波视频图像压缩的性能。

3.2.4 研究热点之四:容错信源编码技术研究^[13,48,49] 信源编码数据流在信道传输时,由于噪声、干扰等一些原因可能导致视频图像严重受损。虽然传统的信道编码技术如前向纠错(Forward Error Correction)可以降低误码率,但残留的比特误码仍然会对图像的品质产生严重的影响,研究结果表明容错信源编码技术有助于这个问题的有效解决,因此为了改善图像的恢复品质,最近制定的一些标准如 JPEG2000 和 MPEG4 都广泛使用了容错信源编码技术。采用容错信源编码技术必然要牺牲部分压缩效率,因此高效的容错信源编码技术研究将是未来一段时间的研究重点。

3.2.5 研究热点之五:小波视频压缩应用研究^[21~32] 今后一段时间小波视频压缩应用研究将沿着三个不同方向进行:(1)面向甚低比特率的应用研究;(2)面向中高比特率的应用研究;(3)支持高度可扩展应用研究。

由于视频处理所要求的实时性缘故,视频压缩编码研究一般都是根据特定的应用环境在算法的复杂度和算法的压缩性能作适当的折衷处理。目前在面向低比特率的应用研究中,MC-DWT 小波视频压缩方案取得了比较满意的效果,这些算法无论是主观评价还是客观评价大都超出了现有的 H.263 标准,下一步的研究应该集中在进一步扩展其处理功能,如增加帧率可扩展功能和空间可分级功能等以满足实际应用的需要。在面向中高比特率的应用研究中,MC-DWT 和 DWT-MC 方案都显示出很好的压缩效果,但目前存在的主要问题在于这些算法的扩展功能不强,尚不能满足实际应用的需要。目前最大的困难是第三类视频压缩编码的应用研究。MC-DWT 和 DWT-MC 两种方案由于它们的递归循环编码结构,不易实现支持高度可扩展功能。MC-3DWC 是最理想的备选方案,但目前这类视频编码器的编码效率普遍偏低,尤其是在低码率条件下。广泛的商业应用,如视频数据浏览,远程医疗,远程教育,远程购物,多级多点视频会议以及交互式多媒体应用等急需支持高度可扩展视频压缩编码的出现,因此这类视频编码算法有着广泛的市场前景。

图像/视频的小波分解的多分辨结构使得小波编码器更易实现可扩展的编/解码,完成 MPEG4 视频对象压缩编码的目标要求。因此基于小波的压缩编码方案是 MPEG4 视频对

象编码一个的优选方案,可以预见今后一个阶段基于小波的 MPEG4 视频对象压缩的研究仍将是理论研究的一个热点。

4 总结

小波基的时频局部化特性和小波变换的多尺度展开结构使得基于小波的图像/视频压缩编码不仅具有良好的率失真特性,而且更容易实现压缩码流的全伸缩特性。因此,随着网络通信的融合,基于小波的图像/视频压缩编码在未来将会显示出技术优越性。本文主要对近十几年的小波视频压缩研究状况作一简要介绍,回顾了近十年来小波图像压缩技术的发展及其在最新静止图像压缩标准 JPEG2000 中的应用,重点阐述了近几年小波视频图像压缩研究进展,并探讨了目前和今后一个阶段小波视频图像压缩研究的几个热点问题。

参考文献:

- [1] 沈兰荪. 图像压缩与异步传输 [M]. 北京:人民邮电出版社, 1998.
- [2] A M Takalp. Digital Video Processing [M]. Now York: Prentice-Hall Inc, 1996.
- [3] ITU-T Recommendation H.261 [S].
- [4] ITU-T Recommendation H.262 ISO/IEC 13818-2:1996 [S].
- [5] ITU-T Recommendation H.263 [S].
- [6] ITU-T Recommendation H.324 [S].
- [7] T Sikara. MPEG digital video-coding standards: delivering picture-perfect compression for storage, transmission, and multimedia application [J]. IEEE Signal Processing Magazine, 1997, 14(5): 82 - 100.
- [8] S Mallat. A Wavelet Tour of Signal Processing [M]. Boston: Academic Press, 1998.
- [9] S Mallat. A theory for multiresolution decomposition: the wavelet representation [J]. IEEE Trans Pattern Anal Mach Intell, 1989, 11(7): 674 - 693.
- [10] I Daubechies. Ten Lectures on Wavelets [M]. SIAM, Philadelphia, PA, 1992.
- [11] D Donoho, I Johnstone. Ideal spatial adaptation via wavelet shrinkage [J]. Biometrika, 1994, 81: 425 - 455.
- [12] S Mallat. Multiresolution approximation and wavelet orthonormal bases of $L^2(\mathbb{R})$ [J]. Trans Amer Math Soc, 1989, 315: 69 - 87.
- [13] JPEG2000 Part I Final Committee Draft Version 1.0 [S].
- [14] J W Woods, S D O'Neil. Subband coding of images [J]. IEEE Trans Acoust Speech, and Signal Proc, 1986, 34(5): 1278 - 1288.
- [15] J M Shaprio. Embedded image coding using zerotree of wavelet coefficients [J]. IEEE Trans on Signal Processing, 1993, 41(12): 3445 - 3462.
- [16] A Said, W A Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees [J]. IEEE Trans on Circuits and Systems for Video Tech. 1996, 6(3): 243 - 250.
- [17] B-B Chai, J Vass, X Zhuang. Significance-linked connected component analysis for wavelet image coding [J]. IEEE Trans on Image Processing, 1999, 8(6): 774 - 784.
- [18] H Taubman. High performance scalable image compression with EBCOT [J]. IEEE Trans on Image Processing, 2000, 9(7): 1158 - 1170.
- [19] H I Witten, R M Neal, J G Cleary. Arithmetic coding for data compression [J]. Communications of the ACM, 1987, 30(6): 520 - 540.

- [20] P G Howard, J S Vitter. Arithmetic coding for data compression [J]. Proceedings of IEEE, 1994, 82(6): 857 - 865.
- [21] S A Martucci, I Sodagar, T Chiang, Y-Q Zhang. A zerotree wavelet video coder [J]. IEEE Trans on Circuits and Systems for Video Tech, 1997, 7(1): 109 - 118.
- [22] D Marpe, H L Cycon. Very low bit-rate video coding using wavelet-based techniques [J]. IEEE Trans on Circuits and Systems for Video Tech, 1999, 9(1): 85 - 94.
- [23] J Vass, B-B Chai, K Palaniappan, X Zhuang. Significance-linked connected component analysis for low bit-rate wavelet video coding [J]. IEEE Trans on Circuits and Systems for Video Tech, 1999, 9(4): 630 - 647.
- [24] Y-Q Zhang, S Zafar. Motion-compensated wavelet transform coding for color video compression [J]. IEEE Trans on Circuits and Systems for Video Tech, 1992, 2(3): 285 - 296.
- [25] H-W Park, H-S Kim. Motion estimation using low-band-shift method for wavelet-based moving-picture coding [J]. IEEE Trans on Image Processing, 2000, 9(4): 577 - 587.
- [26] X Yang, K Ramchandram. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation [J]. IEEE Trans on Image Processing, 2000, 9(5): 778 - 791.
- [27] D Taubman, A Zakhor. Multirate 3-D subband coding of video [J]. IEEE Trans on Image Processing, 1994, 3(5): 572 - 588.
- [28] J-R Ohm. Three-dimensional subband coding with motion compensation [J]. IEEE Trans on Image Processing, 1994, 3(5): 559 - 571.
- [29] S-J Choi, J W Woods. Motion-compensated 3-D subband coding of video [J]. IEEE Trans on Image Processing, 1999, 8(2): 155 - 167.
- [30] J Y Tham, S Ranganath, A A Kassim. Highly scalable wavelet-based video codec for very low bit rate environment [J]. IEEE Journal on Selected Areas in Communication, 1998, 16(1): 12 - 27.
- [31] W A Pearlman, B-J Kim, Z Xiong. Embedded video subband coding with 3D SPIHT [M]. P N Topiwala. Wavelet Image and Video Compression, Holland: Kluwer Academic Publishers, 1998.
- [32] B-J Kim, Z Xiong, W A Pearlman. Very low bit-rate embedded video coding with 3d set partitioning in hierarchical trees (3D SPIHT) [J]. IEEE Trans Circuits and Systems for Video Tech, 2000, 10(8): 1374 - 1387.
- [33] J Li, C-C J Kuo. Image Compression with a hybrid wavelet-fractal coder [J]. IEEE Trans on Image Processing, 1999, 8(6): 868 - 874.
- [34] H Jafarkhani. Wavelet coding of images: adaptation, scalability and transmission over wireless channel [D]. USA: CSHCN University of Maryland, 1997.
- [35] S P Voukelatos, J J Soraghan. Very low bit-rate color video coding using adaptive subband vector quantization with dynamic bit allocation [J]. IEEE Trans Circuits and Systems for Video Tech, 1997, 7(2): 424 - 428.
- [36] S D Servetto, K Rannan. Image coding based on a morphological representation of wavelet data [J]. IEEE Trans on Image Processing, 1999, 8(9): 1161 - 1174.
- [37] W Sweldens. The lifting scheme: a custom-design construction of biorthogonal wavelets [J]. Journal of Appl and Comput Harmonic Analysis, 1996, 3(2): 186 - 200.
- [38] I Daubechies, W Sweldens. Factoring Wavelet and Subband Transforms into Lifting Steps [M]. Preprint, 1996.
- [39] M Cottronei, D Lazzaro, L B Montefusco, L Puccio. Image compression through embdedd multiwavelet transform coding [J]. IEEE Trans on Image Processing, 2000, 9(2): 184 - 189.
- [40] K Cinkler. Very low bit-rate wavelet video coding [J]. IEEE Journal on Selected Areas in Communication, 1998, 16(1): 4 - 11.
- [41] J Zhang, M O Ahmad, M N S Swamy. Quadtree structure region-wise motion compensation for video compression [J]. IEEE Trans Circuits and Systems for Video Tech, 1999, 9(5): 808 - 822.
- [42] J Lu, M L Liu. A simple and efficient search algorithm for block-matching motion estimation [J]. IEEE Trans Circuits and Systems for Video Tech, 1997, 7(2): 429 - 433.
- [43] C Stiller, J Konrad. Estimating motion in image sequences: a tutorial on modeling and computaion of 2D motion [J]. IEEE Signal Processing Magazine, 1999, 16(4): 70 - 91.
- [44] F Dufaux, J Konrad. Efficient, robust and fast global motion estimation for video coding [J]. IEEE Trans on Image Processing, 2000, 9(3): 497 - 501.
- [45] X Q Gao, C J Duanmu, C R Zou. A multilevel successive algorithm for block matching motion estimation [J]. IEEE Trans on Image Processing, 2000, 9(3): 501 - 504.
- [46] S Zhu, K-K Ma. A new diamond search algorithm for fast block-matching motion estimation [J]. IEEE Trans on Image Processing, 2000, 9(2): 287 - 290.
- [47] J Gu. 3D wavelet-based video codec with human perceptual model [D]. USA: CSHCN University of Maryland, 1999.
- [48] I Moccagatta, S S J Liang, H Chen. Error-resilient coding in JPEG2000 and MPEG4 [J]. IEEE Journal on Selected Areas in Communication, 2000, 18(6): 899 - 914.
- [49] Y Wang, S Wenger, J Wen, A K Katsaggelos. Error resilient video coding techniques: Real-time video communications over unreliable networks [J]. IEEE Signal Processing Magazine, 2000, 17(4): 61 - 82.
- [50] 张宗平, 刘贵忠, 等. 一种改进的三维小波视频编码器 [J]. 西安交通大学学报, 2001, 35(6).

作者简介:



张宗平 男, 1971 年生于河南固始, 1995 年 7 月于河南师范大学获理学学士学位, 1998 年 7 月于中国科学院西安光学精密机械研究所获理学硕士学位, 现为西安交通大学电信学院在读博士生, 主要研究兴趣包括小波分析、时频分析、非平稳信号处理、视频图像压缩和电磁场数值计算等。



刘贵忠 男, 1962 年生, 1989 年 6 月获荷兰 Eindhoven 大学博士学位, 现为西安交通大学电子与信息工程学院教授, 博士生导师, 主要从事非平稳信号的分析与处理、音视频数据压缩、模式识别、反演等方面的理论、方法与应用研究。