

图像可分级编码研究进展

王相海

(辽宁师范大学计算机与信息技术学院, 大连 116029) (中国科学院研究生院, 信息安全国家重点实验室, 北京 100039)

摘 要 图像渐进传输、多质量服务以及图像数据库浏览等多分辨率环境下的多媒体应用导致了图像可分级编码思想的产生。本文首先对图像可分级编码的概念进行了解释, 然后对 JPEG 和 JPEG-2000 中的图像可分级情况进行了讨论, 并以实现图像可分级编码的工具和方法为线索, 对基于小波、DCT 和匹配追踪 (matching pursuit) 的图像可分级编码的发展情况进行了分析和讨论, 同时对 3 类可分级编码技术进行了比较, 并对其发展趋向和在视频可分级编码中的推广情况进行了讨论。最后对图像编码技术的发展方向进行了展望。

关键词 图像 可分级编码 DCT 小波 匹配追踪

中图分类号: TP391 **文献标识码**: A **文章编号**: 1006-8961(2006)08-1051-11

Research Progress on Scalable Image Coding

WANG Xiang-hai

(College of Computer and Information Technology, Liaoning Normal University, Dalian 116029)

(State Key Laboratory of Information Security, Graduate School of Chinese Academy of Sciences, Beijing 100039)

Abstract The multimedia applications in a multiresolution environment, such as progressive transmission of image, image database browsing, etc., arise scalable image coding algorithm. This paper first explains the concepts of all kinds of scalable image coding, and gives an analyse for the scalable image coding algorithms in JPEG and JPEG-2000. Secondly, an overview of major methods already used or under investigation for scalable image coding based on Wavelet, DCT and Matching Pursuit with the scalable coding tools and methods is introduced, and the decoder performance comparison among the three kinds of scalable coding is given. In addition, the trend of the three kinds of scalable coding and their extension in video scalable coding is discussed. Finally, the future development of image coding is also analyzed and conjectured in the paper.

Keywords image, scalable coding, DCT, wavelet, matching pursuit

1 引 言

随着网络和多媒体技术的发展, 人们对图像编码的要求越来越高, 不仅要求编码技术具有较好的压缩效果, 而且要求它能适应网上渐进传输图像的需求, 比如在许多实际应用中, 由于用户要求不同、终端能力不同、异构网络的不同支路所能提供的 QoS 不同或网络传输条件的变化 (比如噪声、拥塞等) 等原因, 需要提供不同质量的图像信号, 解决此类问题的最好方法是用单个编码器产生分层次的压

缩码流, 对不同层次的码流解码可以获得不同的图像质量, 例如, 低档次终端只对码流的一部分进行解码, 获得低分辨率的图像, 而处理能力高的终端对整个码流进行解码获得高分辨率的图像, 这种编码机制通常被称为可分级编码。可分级编码分为视频可分级和静态图像可分级, 本文重点讨论静态图像可分级。与传统的图像编码方法相比, 静态图像的可分级编码机制除了能满足可分级需求之外, 在相同的码率下还具有更加优秀的编码效率。

实际上, 静态图像可分级编码的思想在 20 世纪 90 年代的 JPEG 图像编码标准中便有所体现, 随着

基金项目: 国家自然科学基金项目 (60372071); 辽宁省自然科学基金项目 (20032105); 辽宁省高等学校优秀人才支持计划 (RC-04-11)

收稿日期: 2005-05-16; **改回日期**: 2005-10-09

第一作者简介: 王相海 (1965 ~), 男, 教授, 博士。主要研究方向为 CG、CAGD、图像编码和多媒体信息处理。E-mail: xhwang@lnnu.

图像编码技术的发展,人们对图像的可分级编码技术进行了进一步的研究,比如在 MPEG-4 视频标准校验模型中,对静态图像纹理的编码便给出了空间分辨率可分级、数率可分级以及二者的混合可分级的编码模式^[1];此外在 JPEG-2000 中,还提出了基于 ROI(regions of interest)的可分级编码思想^[2]。

2 图像可分级编码的基本概念

一个图像编码的比特流具有可分级性 (scalability)是指利用该比特流的一部分(或子集)可以产生对该图像一个有用表示的能力,即解码器能够根据分辨率的情况对码流的一部分进行解码。图像的可分级编码通常包括以下情况^[3,4]:空间分辨率可分级编码、数率可分级(或称 SNR 或质量可分级)编码、空间和数率混合可分级编码以及 ROI 可分级编码等。

数率可分级指对一个压缩的码流,用户可根据带宽或系统能力的具体情况从码流的开始端来截取相应的一段码流进行解码。在对码流的截取过程中,如果在一定范围内可任意选择数率,则此时的数率可分级被称为连续数率可分级(图 1)。空间分辨率可分级指以不同空间分辨率对图像进行解码(图 2)。



图 1 图像数率可分级示意图
Fig. 1 Sketch map of image rate scalability

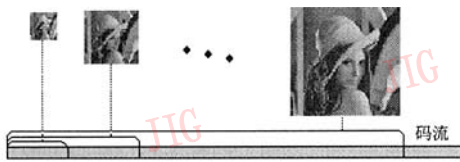


图 2 图像空间可分级示意图
Fig. 2 Sketch map of image spatial scalability

除了单独的数率可分级和空间可分级编码方案外,文献[1]中还提出了静态纹理图像的数率可分级与空间可分级的混合编码方案,即码流在总体上实现基于分层的空间可分级编码情况下,对每一层次的空间可分级图像码流,又可实现其数率可分级解码(图 3),这在一定程度上提高了可分级的灵活性。

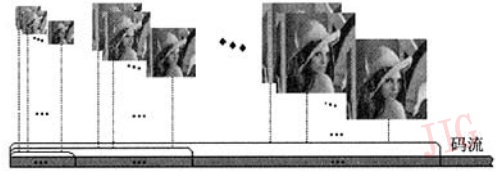


图 3 图像空间与数率混合可分级示意图
Fig. 3 The sketch map of image spatial/rate scalability

ROI 可分级编码技术是指在图像编码码流中,允许图像的某一部分(即感兴趣区域)以更好的质量被解码(图 4)。



图 4 以左上角矩形区域为“感兴趣区域”的 ROI 可分级解码图像

Fig. 4 Scalable decoding images which ROI is rectangle region

3 JPEG 和 JPEG-2000 中的可分级编码

3.1 JPEG 中的图像可分级编码分析

尽管在 JPEG 图像编码标准^[5]中还没有明确提出图像可分级编码的概念,但其提供的“累进编码”方式和“层次编码”方式则分别对应了前面所介绍的数率可分级和空间可分级编码方式。

(1) JPEG 中的数率可分级编码

JPEG 中的累进编码其目的是为了解码过程中能在屏幕上逐步浮现重建的图像,改善人机交互效果。它的原理如图 5 所示。

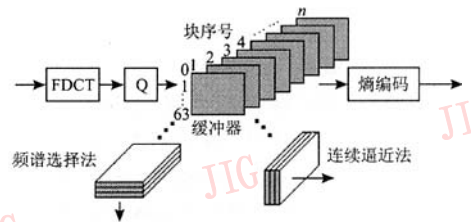


图 5 JPEG 的累进编码模式
Fig. 5 Progressive coding mode in JPEG

JPEG 编码器中有一个缓冲器,在累进编码模式中,缓冲器的容量需要能存储图像中的全部 n 个子

块数据,等所有子块数据都到达缓冲器之后,把它们组织成一个数据长方体,然后可采用两种方式实现累进编码,一种是频谱选择法,即将数据长方形按水平方向切片,自上而下(从低频数据到高频数据)将数据送给熵编码器进行编码;另一种是连续逼近法,即将数据长方体按垂直方向切片,自左向右(从高位数据到低位数据)将数据送给熵编码器去编码。

(2) JPEG 中的空间分辨率可分级编码

JPEG 层次编码模式的目的是为了适应各种不同分辨率的显示器展现图像的需要,用户在解码时只要从网络上接收能满足分辨率要求的数据即可,而不必要接收编码图像的全部数据。其编码器原理如图 6 所示。

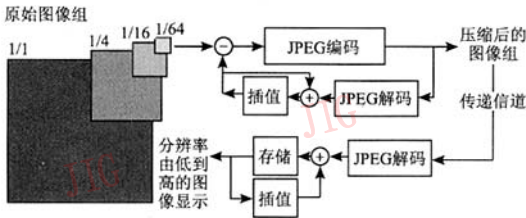


图 6 JPEG 的层次编码模式

Fig. 6 Hierarchical coding mode in JPEG

JPEG 的层次编码模式首先把原始图像的分辨率逐级降低,得到一组分辨率由低到高的图像,然后把最低的图像进行 JPEG 编码,同时对该图像进行解码,并利用插值方法提高其分辨率,作为高一级的分辨率原始图像的预测图像;求出预测图像与相应原始图像的差值图像,并对差值图像进行 JPEG 编码。

3.2 JPEG-2000 中的 ROI 可分级编码分析

JPEG-2000 编码采用的是以小波编码为基础的分块处理方式,JPEG-2000 中提出了一种 Maxshift 的 ROI 可分级编码方法^[2],这在一定程度上体现了 JPEG-2000 所具有的基于内容的编码能力。

Maxshift 的基本思想是将与重构 ROI 无关的变换系数(不妨称为背景系数)适当按比例缩小,使得与 ROI 相关连的系数(不妨称为 ROI 系数)在进行编码时排在背景系数的前面,这样在解码端,在一定的解码率下 ROI 系数能得到充分的解码。其编码过程如下:

(1) 确定 ROI 面罩(即对最后重构 ROI 有贡献的所有小波分解系数的集合)。首先利用 ROI 的形

状信息(比如二值 α 平面)确定属于 ROI 的像素;根据所采用的小波变换的逆变换确定重构当前层 ROI 系数所需要的上一层系数,此过程重复进行,直到找到所有的 ROI 系数。图 7 给出了 $W9 \times 7$ 滤波器逆变换的示意图,为了重构 ROI 中的 $X(2n)$ 和 $X(2n+1)$,需要其上一层低频子带中系数 $L(n-1)$ 、 $L(n)$ 、 $L(n+1)$ 、 $L(n+2)$ 和高频子带中的 $H(n-2)$ 、 $H(n-1)$ 、 $H(n)$ 、 $H(n+1)$ 、 $H(n+2)$ 。

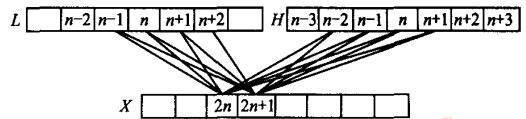


图 7 9/7 小波逆变换示意图

Fig. 7 Sketch map of inverse 7/9 wavelet transform

(2) 确定背景系数的缩小因子。首先求出所有方块中背景系数的最大值,然后根据此最大值求出背景系数的缩小因子,使得背景系数被缩小后其最大值要比 ROI 系数中非零最小值还要小。

(3) 根据 ROI 面罩和背景系数的缩小因子对背景系数进行缩小操作(图 8)。

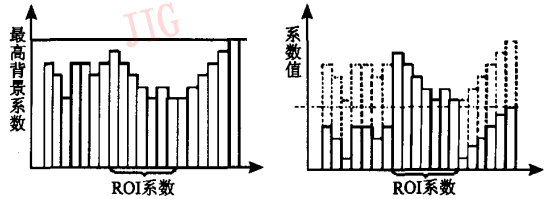


图 8 缩小背景系数示意图

Fig. 8 Sketch map of reducing the background coefficients

(4) 将背景系数缩放因子写入码流相应的标志位中。

Maxshift 方法的解码过程变得非常简单,首先从码流相应的标志位中获取缩小因子并利用此缩小因子确定出一个阈值,该阈值介于 ROI 系数与背景系数之间;然后将码流中相应的每一个系数与阈值相比较,如果比阈值小,则按缩小因子对其进行等比例放大。

Maxshift 编码方法具有许多好的性质,具体表现如下:它支持任意形状的 ROI 编码,并且具有较小的复杂度;无需向解码端发送 ROI 的形状信息和 ROI 面罩信息,并且在解码端也无需对 ROI 进行特殊考虑;此外,该方法还支持对 ROI 的质量和分辨率的可分级解码等特性。

4 基于小波的图像可分级编码

借助小波变换可以把图像分解成许多具有不同空间分辨率、频率特性和方向特征的子带信号,实现低频长时特征和高频短时特征的同时处理,图像小波分解的这些特性,为图像的可分级编码的实现提供了很好的基础。以下分别对基于小波图像编码的空间可分级、数率可分级编码以及空间和数率的混合可分级编码的发展和典型算法进行分析和讨论。

4.1 空间可分级编码

基于小波的图像分解具有的将图像分解成不同空间分辨率层次结构这一特性使得基于小波的图像空间可分级编码的实现变得有效和简单,为了实现空间可分级编码,只需在某种量化策略下按图像的分解层次对各子带进行编码;在解码时按照指定的空间分辨率的要求,对相应层次子带中的系数进行解码,随着对空间分辨率要求的提高,只需将更高层次的子带系数参与到解码数据中即可。

4.2 数率可分级编码

小波变换具有许多良好的空间-频率局部化特性,这些特性主要表现在以下 3 个方面:频率压缩特性,即从频率分解上看,原始图像的能量大部分聚集在低频子带;空间压缩特性,即从系数在空间的分布看,原始图像的能量大部分集中在图像的边缘轮廓等对应的位置;系数的分布相似特性,即同一方向上各级高频子带系数幅值分布大体一致。人们除了利用图像小波变换的频率特性进行编码外,还利用其空间压缩特性和系数分布的相似性提出了一些卓有成效的具有数率可分级特性的编码方案。下面对一些典型的算法进行分析。

Shapiro 在文献[6]的基础上,提出了“零树”数据结构的全新方法来表征小波系数分布的空间自相似性,进而提出了一个嵌入式零树小波编码模型 EZW(embedded zerotree wavelet)^[7]。所谓零树是指:如果一个小波系数在一个粗的尺度上关于给定的阈值 T 是不重要的,之后在较细尺度上在同样的空间位置中的所有小波系数也关于阈值 T 是不重要的,则称这些小波系数形成了一个零树,这时在粗尺度上的那个小波系数称为树根,较细尺度上相应位置上的小波系数称为孩子。整个 EZW 的编码过程是在阈值的意义下用上一级的小波系数预测下一级的小波系数,每次阈值的选取按递减方式来实现

逐级量化,这样使最后所产生的比特流可以按对应的幅值大小分层次地排列,从而实现对图像的逐次逼近和对每个小波系数进行编码的过程。

由 Said 和 Pearlman 提出的 SPIHT(set partitioning in hierarchical trees)算法^[8]是对 EZW 算法的改进。EZW 算法使用了零树结构来揭示小波分解图像不同子带间不重要系数分布的自相似性,取得了较好的编码效果。然而,该方法还存在着一些不足,比如在进行零树编码时,对不能被包括进零树的一些不重要系数必须个别编码为孤立零码,这一部分编码往往会具有较高的代价;另外由于孤立零码的过多可能导致高频子带上的一些边缘重要系数在低码率要求下未能充分解码。SPIHT 算法采用了与 EZW 算法相同的假设,但与 EZW 算法相比该算法注意到对“零值”除了采用零树进行描述外,对未包含进零树的其他零值在很大程度上亦可以采用与零树相类似的其他符号予以描述。由于该算法采用了更精细的编码模式,所以取得了较 EZW 算法更好的编码效果,在 PSNR 上平均要高出 EZW 算法 0.3 ~ 0.6 dB。

SPIHT 算法包括两个主要的核心技术,即集合分割分类算法和空间方向树及其层次结构上的集合操作,后者与 EZW 中的零树结构相类似。而集合分割分类算法是通过减小阈值,将小波系数按阈值从最重要到最不重要进行层次分类,进而构建其重要图。对每一个阈值,编码过程一般包括两个过程,即分类过程和提炼过程。整个 SPIHT 算法有 4 个实现步骤,即初始化、分类过程、提炼过程和量化尺度的修改。

Servetto 等人提出了一种基于形态学方法的图像编码算法 MRWD(morphological representation of wavelet data)算法^[9],该方法基于小波图像分解后各子带中重要系数的“聚类”特性,利用数学形态学中的膨胀算子直接对各子带的重要系数进行检测、提取和编码,该方法弥补了零树编码的不足,降低了运算的复杂度,取得了很好的编码效果。MRWD 算法的基本思想是在阈值意义下对小波分解图像各子带按重要性顺序进行如下两方面的操作:第一,对每个子带,由重要系数的聚类特性可知,重要系数更有可能聚集在一起,然而,在每个重要系数的聚类簇中,一些零值往往会嵌在其中,这样很难采用类似于“链式”编码的方式对聚类簇的形状进行编码,为此,算法利用形态学中的膨胀操作进行区域增长,同时对聚类簇中的重要系数进行测试和提取;第二,对

同一方向各子带,由于重要系数的分布是相似的,即如果高层子带中的父亲系数是重要的,那么其低层子带上的孩子系数也极有可能是重要的,这样对低层次上的重要系数簇便可能利用高层的重要系数簇来预测确定低层次上膨胀操作的“种子”系数,从而提高了编码效率。

MRWD 算法以形态学中的膨胀操作作为工具,在没有增加额外开销的情况下即可以确定各子带中任意形状的重要系数簇,并以一种隐含的逻辑顺序对其进行编码,取得了优于 EZW 算法,而与 SPIHT 算法相近的编码效果,特别对低码率编码情况。

EBCOT (embedded block coding with optimized truncation) 算法^[10]是一种具有对嵌入式码流进行优化截取机制的嵌入式分块编码方法,该方法以其具有灵活的可分级性、可在一定程度上对码流进行随机访问以及较低的复杂度等优异性质,而被 JPEG-2000 标准的 VM3 (verification model version 3) 所采纳,该算法非常适用于对较大压缩图像的远程浏览。

与前面的可分级编码方案不同,EBCOT 采用了如下两种机制实现对图像的可分级性编码:第一,对于每一个图像小波分解子带,EBCOT 将其进一步分解为一个称为“码块”的样本块(这些码块的大小通常被选为 32×32 或 64×64),对每一个这样的码块,EBCOT 将产生一个具有高可分级性的嵌入式码流,使得对该码流的大部分截取点均位于所对应的率失真曲线的“凸壳”上,为了达到这样的效果,算法采用了分数位平面 (fractional bit-plane) 编码方法^[11,12];第二,对于每一个码块的码流,EBCOT 按照一种率失真优化原则,将其分为若干个“质量层”,对于给定的数率,EBCOT 可以从最低质量层开始,依次提取位于该层上的各码块码流,形成对原始图像的率失真优化表示。

在表 1 中给出了本节所讨论的几种编码算法的比较结果,采用的是 Daubechies 的 9-7 双正交样条

滤波器^[13]和标准的 512×512 Lena 8bpp 图像,图像被分解了 5 层。

4.3 空间和数率混合可分级编码

MPEG4 中对静态图像纹理给出了一种基于零树小波的空间与数率混合可分级的编码方案^[14],整个方案首先对纹理图像进行小波分解 (MPEG-4 提供了一种省缺的 9/3 奇次对称双正交小波),对于低频子带,方案进行了单独处理,即首先采用一种“均匀中点提升”量化器对小波系数进行量化,然后采用 DPCM 编码模式和自适应算术编码器对量化后的数据进行编码;对其他高频子带,方案则采用了基于零树的多尺度编码方法,具体过程如图 9 所示。

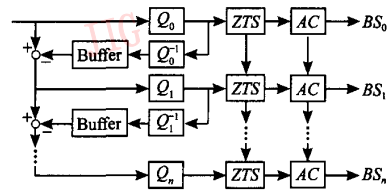


图 9 MPEG-4 中静态图像纹理多尺度零树小波编码流程图
Fig. 9 Flow chart for static texture image multiscale zerotree coding in MPEG-4

此外,在文献 [15] 中实现了基于 EZW 的图像空间和数率可分级编码两种方案,其中方案一是在 EZW 算法的基础上,对一定范围内任意选取的数率获得相应的嵌入式码流,然后对该码流进行空间层次重排,即将指定数率的整幅图像的码流重新分配到相应的空间子带中,这样对于指定的空间分辨率,利用相应空间子带系数便可以实现可分级解码;方案二是在空间可分级确定情况下,实现连续数率可分级编码,该方案采用了一种基于空间分层的零树小波编码机制,即随着空间分辨率的递增,对前一空间分辨率状态下的嵌入式码流进行调整和扩充,以形成新的空间分辨率下的嵌入式码流,在此基础上实现该空间分辨率情况下的数率可分级解码。

5 基于 DCT 的图像可分级编码

近年来,随着 Shapir 嵌入式零树编码器的提出和 JPEG-2000 中采用了小波作为其编码工具,人们对图像编码技术的研究已有从基于 DCT 转向基于 DWT 的发展趋势。然而,尽管小波变换与 DCT 相比能够表现出更加灵活的空间-频率分辨率的特性,

表 1 几种算法对 Lena 图像编码的 PSNR 性能

Tab.1 PSNR performance effects of the algorithms on Lena image

bpp	PSNR			
	EZW	SPIHT	MRWD	EBCOT
1.00	39.55	40.41	40.33	40.57
0.50	36.28	37.22	37.18	37.41
0.25	33.17	34.12	34.12	34.29

但与 DCT 一样,小波变换也只是线性变换集合的一元,EZW 等算法的高效特性,部分是小波变换的功绩,但更重要的是其所采用的基于“零树”的量化机制,在 DCT 系数上也可以采用嵌入式零树量化器^[16]。进一步,Xiong 等人在文献[16]中提出了一个基于 DCT 的嵌入式图像编码器。该方法首先将图像进行 8×8 DCT 分块变换,每块 DCT 系数的逻辑关系被看成为与 EZW 中小波系数相同的深度为 3 的系数“树”,算法中采用了与 SPIHT 相同的嵌入式零树量化器。算法不仅可以形成具有可分级特性的嵌入式码流,而且其编码效率要高于 JPEG 和 EZW。

Monro 等人在文献[17]中也提出了一种 DCT 系数的零树编码算法,与文献[16]中算法不同的是,该方案在对分块 DCT 系数进行零树量化之前,首先通过对分块相同空间方向和频率位置上的 DCT 系数进行系数重排,形成一个整体的“DCT 层次子带结构”,然后再对该结构实施类似 EZW 的零树量化。该方案同样取得了高于 JPEG 和 EZW 的编码效果。此外,该文还对分块的大小进行了试验和讨论,指出该方案对压缩低于 0.4bpp 的情况,采用 16×16 的分块可取得最佳效果,而对高于 0.4bpp 的情况,则采用 32×32 可以实现最佳。

Wei 等人提出一种基于整数 DCT^[18]的嵌入式图像编码方案^[19]。该方案首先进行 8×8 DCT 分块变换,并将每一块中的 DCT 系数看成是一个频率系数,然后按相近的频率系数分成一类,将每一个 DCT 块系数分成 10 个不同的频率子带,同时对每一块中相同频率的系数根据块的顺序进行重排,形成类似小波分解的频率子带。在此基础上,文中对形成的子带结构分别进行了基于 EBCOT 和 SPIHT 的零树量化方案,从而形成嵌入式的可分级码流,试验结果表明,所提出的整数 DCT 编码方案对无损编码要优于浮点型 DCT 编码,而对有损压缩,二者比较接近。

此外,Li 等人在文献[20]中提出了一种基于平面的分层 DCT 嵌入式图像编码器。与 JPEG 相同,该方法也首先对图像进行分块 DCT 编码,并对每块的系数进行“Z”字形顺序重排,与 JPEG 不同的是在对重排的 63 个 AC 系数进行编码时,该方法采用了按照各系数的比特位由“重要”到“不重要”的顺序(文中称为“按层次”)进行编码,从而形成了嵌入式的编码码流。实验结果表明,该方案的编码效果明显高于 JPEG 情况,特别在低码率情况效果更加明显。

6 基于 MP (matching pursuit 匹配追踪) 的图像可分级编码

在信号处理中,正交基由于能用很少的向量来有效地逼近某一类信号而受到重视,在傅里叶分析与小波基中,这种逼近对一致正则的信号是非常精确的,但依赖信号来选取基向量会得到更好的逼近。MP 技术首先由 Mallat 和 Zhang 在文献[21]中给出。该方法通过迭代的方式从一个称为“冗余字典”(redundant dictionary)的函数集中选取与信号结构最佳匹配的波形函数,实现信号在这些波形函数上的次最优线性展开。这里,冗余字典是特殊设计的用于捕捉信号特征的函数集合。MP 通常不考虑正交性的限制条件,它是通过从时频原子的冗余字典中一个一个的挑选函数,每一步都使得信号的逼近更为优化。这一技术已被应用于视频和图像编码中^[22-24],并取得了非常好的压缩效果。

MP 的特性可以概括为两方面^[24]:一方面是算法本身所固有的特性,即不管采用哪一种函数进行信号分解,MP 算法都具有可逆性、能量守恒性等;另一方面是依赖冗余字典的特性,比如在冗余字典的选择中通常采用具有平移不变性、旋转不变性、膨胀不变性的函数,这样 MP 算法也就具有这样的特性。MP 的这种对冗余字典的依赖特性和其自身对信号提取的渐进逼近特性为实现图像编码的可分级性打下了基础。近年来,基于 MP 的图像可分级编码技术受到人们的关注,出现了一些较为有效的编码技术,特别对低码率下的可分级编码。

Bergeaud 和 Mallat 在文献[23]中首先将 MP 应用于图像编码中,通过伸缩、平移和调整高斯窗,构造了时间和频率平移不变的 Gabor 字典。Gabor 原子具有比小波包更好的时频局部性,且在一个更加精细的时频网格上进行平移,从而保证了可精确定位图像的边缘和纹理区域的局部分解,逐步精细的提取图像的主要特征。Rabiee 等人在文献[25]中使用卷积样条^[26]所构成的有限字典,利用 MP 方法采用递归方式逐步精细地逼近图像,直到渴望的峰值信噪比或误差计算被满足。从而得到了一个自适应的多分辨率的图像可分级编码算法,该算法的计算复杂度和低码率下的峰值信噪比和解码图像的主观质量要优于基于 DCT 的 JPEG 编码。

在 MP 的实施过程中,对最佳函数的搜索意味

着将要进行大量的数积运算,这样对尺寸较大的图像,MP 通常具有很高的计算代价。为了提高 MP 的编码速度,文献[26]提出了一种 MRMP (multiresolution MP) 方案,该方案首先对原图像进行几次二下取样,并将 MP 算法应用于最小的图像,当在最低分辨率层上获得了所渴望的系数个数后,利用所获得的系数进行下一层图像的重构,同时将该重构图像与下一层分辨率图像进行差运算,得到相应的差图像,并将 MP 作用于该差图像。由于多分辨率减少了字典中的原子数量,所以 MRMP 通常较单分辨率下的 MP 有更好的效果;此外,由于 MRMP 与 HVS 具有很多相似的特性,从而使得 MRMP 在低 PSNR 下的重构图像较 DCT 和小波具有更好的视觉效果。

Frossard 等人在文献[27]中提出了一个基于 MP 的高灵活性的图像可分级编码器。为了有效地捕捉到图像的边缘信息,编码器的冗余字典采用对高斯函数进行平移、旋转和各向异性加细操作得到,同时字典也保持了对各向同性放缩操作的不变性,该特性使图像空间大小的调整操作变得非常简单,解码器能够通过修改单位范数原子的参数串来重构放缩后的图像,从而使得编码器具空间可分级特性。该方案所提供的空间可分级对于任意空间分辨率都是有效的,这要比基于子带的空间可分级更加灵活。此外,由 MP 所产生的嵌入式码流对于任意的数率限制,能够使接受端获得最重要的图像重构信息,从而保证了解码器具有高效的数率可分级特性。

Frossard 在该文献中还提到了可利用文献[28]所提出的一种对 MP 系数进行基于率失真优化的后量化方案来改进该编码器的效率。此外,文献[29]还提出了一种基于 MP 的彩色图像可分级编码方案,编码器采用了与文献[28]相似的基函数同时利用了彩色信道间的冗余来减少码率,取得了很好的空间和数率可分级编码效果^[29]。

7 3 类图像可分级编码技术的比较

7.1 基于 DCT 的图像可分级编码与基于小波的图像可分级编码的比较

在表 2 中结合文献[8]、[10]、[17]、[30]给出了一些基于 DCT 和小波的典型可分级算法对 Lena 和 Barbara 图像进行编码的比较结果。从比较结果可以看出,尽管 DWT 较 DCT 具有更好的空间分辨率、频率特性和方向特性,但仅从编码效果上看,基于 DWT 的零树小波编码方案之所以有较好的效果,除采用了 DWT 对图像进行分解外,很重要的一点是采用了“零树”量化模式,如果对于图像的 DCT 系数采用相应的优化量化模式,同样可以取得较好的编码小波,甚至可以取得超过 EZW 的编码效果,而 DCT 一般又较 DWT 有更小的计算复杂度。当然,DWT 好的性质为实现具有其他特性的编码机制提供了可能,比如可分级性、鲁棒性、对码流的直接操作性等等。这也是 JPEG-2000 采用 DWT 作为其核心技术的一个重要原因。

表 2 基于 DCT 和 DWT 典型图像可分级编码算法的 PSNR

Tab.2 PSNR performance effects of the typical image scalable coding algorithms based DCT and DWT

Rate(bpp)	Optimized JPEG ^[17]		EZ_DCT ^[31]		EZW ^[7]		SPIHT ^[8]	
	Lena	Barbara	Lena	Barbara	Lena	Barbara	Lena	Barbara
0.25	32.30	26.70	32.25	26.83	33.17	26.77	34.11	27.58
0.50	35.90	30.60	36.00	30.82	36.28	30.53	37.21	31.39
0.75	38.10	33.60	38.06	33.70	—	—	39.04	34.25
1.00	39.60	35.90	39.62	36.10	39.55	35.14	40.40	36.41

7.2 基于 MP 的图像可分级编码与基于小波的图像可分级编码的比较

为了具有一般性,本文将文献[27]中所提出的基于 MP 的图像可分级编码器(简记为[27]编码方案)和通过文献[28]所提出的基于率失真优化的后量化方案改进的可分级编码器(简记为[27]+[28]

编码方案)与以分块小波编码为基础的 JPEG2000 的编码效率进行比较,图 10 给出了 3 类编码方案的率失真特性比较,其中图像采用了 128 × 128 的 Lena 图像,分别通过 3 种编码器并以 1.7bpp 形成嵌入式码流^[27]。

从图中的曲线可以看出,在低码率情况下,基于

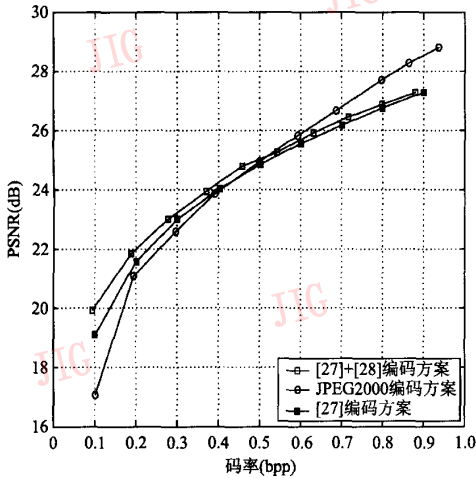


图 10 MP 与 JPEG2000 编码的率失真比较

Fig. 10 Rate distortion comparison between MP image coding and JPEG2000 image coding

MP 所截取的码流总是优于以分块小波变换为基础的 JPEG2000 所截取的码流。但随着码率的增加, MP 的编码效率便不再与小波编码具有可比性, 一般并不能实现无损编码。此外, 该方案编码时一般需要较大的计算量, 但解码器具有简单和速度快的特性。

7.3 3 类图像可分级编码技术的发展趋向

尽管目前基于 DCT 的图像可分级编码的算法还不是很多, 特别是在空间可分级和空间-速率混合可分级方面。但由于 DCT 编码的简洁性和硬件的易实现性, 对该领域继续深入的研究具有重要的实际意义。

近年来, 除了基于 DCT 图像可分级编码技术的发展之外, 基于 DCT 的视频可分级编码技术也得到了蓬勃发展, 最为典型的要数基于 DCT 的细粒度可分级 (FGS) 视频编码技术的发展。视频可分级编码机制在 MPEG-2 和早期的 MPEG-4 版本中采用的是基于 DCT 的分层编码模式, 即视频流被编码为一个基层和一个增强层, 在解码过程中首先解码基层码流, 然后根据解码分辨率的情况, 通过解码增强层来提高解码精度。上述分层可分级编码技术的一个共同特性是要想使得解码精度得以提高, 增强层的码流必须完全解码, 否则不会增强解码的质量。MPEG 组织注意到了网络传输对视频编码的新的要求, 开始征集精细的可分级视频编码方案, 该方案要求将视频编码成一个可以单独解码的基层码流和一

个可以在任何地点截断的增强层码流, 即实现解码的“连续可分级”, 其中基层码流适应最低的网络带宽, 而增强层码流用来适应网络带宽变化的动态范围。世界各地的学者提出了许多解决方案, 最终 Li 的 FGS 方案^[31] 被选中, 其基本思想是将一个视频序列编码为一个基层码流和一个增强层码流, 基层采用非可分级编码方式进行编码, 其码率作为解码端可分级解码的下界; 增强层则采用了基于 DCT 系数的位平面编码机制对原始图像和重构图像的帧差图像进行编码。该方案的一个主要特点是预测帧中的基层和增强层都是以重构的参考帧中的基层作为预测信息进行预测的, 这样 FGS 编码方案对于偶尔的数据丢失或增强层中的数据错误通常具有很好的错误恢复能力, 此外由于利用基层信息来预测增强层, 从而使得在码流传输过程中增强层信息的丢失或被侵蚀不会像传统的 SNR 可分级编码方案那样引起误差漂移现象。然而, 同样是由于预测总是在基于最低质量的基层上进行, 这样对增强层的预测一般不是很准确, 形成的 DCT 参差较大, 编码效率不是很高, 有时甚至比传统的 SNR 可分级方案的编码质量还要糟糕^[32]。近年来, 人们在 FGS 编码结构的基础上对细粒度的视频可分级编码技术进行了较深入的研究, 提出了许多有效的改进算法^[33-36]。

随着 JPEG-2000 标准的颁布, 基于小波的图像编码技术越来越受到关注, 未来除了小波自身的理论方法不断完善外, 它与其他工具的相互结合也将成为今后的一种发展趋势。事实上, 多年来与图像编码技术相关的许多学科和技术的发展, 诸如图像处理、数学形态学、计算机视觉和神经网络等等, 已为图像编码准备了许多有用的工具, 多种工具的有机结合有可能会提高编码性能的进一步提高。

目前基于小波的视频编码技术还不像图像编码技术那样成熟, 有很多问题还有待于进一步研究。目前的编码技术主要有两类, 一类是基于 3D 小波变换的视频编码方法, 另一类是基于运动补偿的 2D 小波视频编码方法。前者是将视频信息看成一个 3 维整体, 将基于小波图像编码的思想直接扩展到 3 维空间的视频信息中。该方法以其无块效应和比较容易实现可分级编码而受到关注^[37], 然而它也明显存在着一些不足, 比如它在时间方向的滤波并不是沿着视频对象运动的方向, 这通常会影响到时域冗余量的减少, 即编码效率不是很高; 此外在算法的具体操作中要利用多帧信息, 一般需要较多的帧存储

器,并且会带来较大的延迟。这样人们更看重基于运动补偿的 2D 小波视频编码方案^[39]。在基于运动补偿的 2D 小波可分级编码方法中,通常需要解决的问题是如何有效地实现帧间的运动估计和如何对剩余帧差图像进行有效编码。

运动估计在小波可分级编码方案中是重点也是难点问题,其主要原因表现在视频序列的连续两帧经小波分解后,相应的高频子带间具有很少的“相似”之处,一般而言相互之间并不能通过简单的运动估计/补偿得到。这是因为基于二下取样的小波图像分解关于平移操作是可变的,即二下取样操作与平移操作是不可换的^[38]。为此人们进行了积极的研究,Wang 等人^[39]给出了一种分层预测的方案,该方法的运动估计是在小波变换域上进行的,利用小波分解每一层中的 LL 子带的运动估计矢量作为该层其他子带的运动估计矢量进行估计,取得了一定的效果;Cheng 等人提出了一种基于下层低频子带的分层运动估计方法^[40],该方法对最低频子带进行运动估计,然后从最低频开始对各高频子带进行“间接”估计,即首先利用估计层的下一层低频子带进行运动估计获得相应的预测 LL 子带,然后对该 LL 子带进行一级小波分解,从而获得相应高频子带信息;在此基础上,文献[41]对其系统框架进一步调整,提出了基于下层 LL 子带的分层运动估计方法,并将小波域运动估计方法应用于空间可分级编码的具体框架和方法。此外,为了适应视频的可分级编码,Delp 等人提出了一种 SAMCoW 方案^[42],该方案为了保持数率可分级解码的视频质量,在解码端增加了一个反馈环路,使得编、解码端对帧差的解码被锁定在同一数率上,保证了编解码端与解码端使用相同的参考帧,从而避免了较大的误差“漂移”现象。

基于 MP 的图像可分级编码是近年来新兴起的一项编码技术,由于其作用于整幅图像,所以没有块效应,并且其图像质量的损失也只是在图像的细节方面。目前来看,该方案在低码率下的编码存在着明显的优势,进一步在算法的复杂性和高码率下的编码效率还有待于进一步改进。

人们在对基于 MP 的图像编码算法研究的同时,也积极展开了基于 MP 的视频编码的研究。在文献[43]中,Neff 等人提出了一种基于迭代 MP 算法^[44]的运动帧差编码方案,该方案使用一个贪心匹配技术将帧差图像分解为原子,并利用完备 Gabor

库中的基函数对其进行编码,实验结果表明,该方案在可视性和 PSNR 两方面均优于基于 DCT 的视频编码器^[45];进一步,他们又通过应用最佳自适应模数量化改进了算法的效率^[46]。文献[47]采用一种启发式的搜索过程来提高基本算法的效率,同时给出了一个 SNR 可分级编码器。Rahmoune 等人提出了一种新的基于 3 维 MP 的高灵活性的可分级视频编码方案(MP3D)^[48],其编码效果优于 MPEG-FGS^[31],特别在低码率情况;此外由 MP3D 所产生的码流在自适应性方面更加灵活。

8 图像编码技术发展展望

尽管压缩效果一直是任何编码方案所追求的目标,但它不是唯一的目标,特别是随着 Internet 和多媒体技术的进一步发展,静态图像编码的一些更高的要求会被提出来,我们认为未来图像编码技术在以下几个方面将会有更好的发展:

(1) 对图像编码信息的直接操作和访问。对编码信息的可操作性是未来图像编码技术所追求的一个重要目标,它不仅可以节省存储开销,减少运算的复杂度,而且在操作的过程中可以降低噪音的引入。这一目标在 JPEG2000 中已初露端倪,相信这将是未来的一个研究热点。

(2) 图像的可分级编码技术将会得到进一步发展和完善。可分级编码作为适应于多分辨率环境下多媒体应用的重要方案将随着编码工具的不断发展和完善而更加灵活和有效,特别是基于对象的图像可分级编码技术将会得到进一步发展。

(3) 适应于有噪信道、特别是无线通信网络的静态图像可分级编码技术将会得以更好的发展,这也是未来的一个重要的研究领域。

(4) 适应基于内容检索需求的图像编码技术。基于内容的检索是多媒体和 Internet 技术发展过程中所提出的一项重要的实用技术,诸如图像数据库的检索、数字化图书馆、Web 网页上的图像检索等应用领域都有这方面的需求。如何使图像编码技术有效地适应基于内容检索的需求是亟待解决的问题。

参考文献 (References)

- 1 ISO/IEC document JTC1/SC29/WG11 N3312: MPEG-4 Video Verification Model Version 16.0[S], 2000.

- 2 ISO/IEC 15444-1: Information Technology-JPEG 2000 Image Coding System: Core Coding System[S], 2001.
- 3 Wang X H. Research on Scalable Image and Video Coding Based on Wavelet[D], Nanjing: Nanjing University, 2001. [王相海. 基于小波的图像和视频可分级编码研究[D]. 南京: 南京大学, 2001.]
- 4 Wang X H. Research on scalable image coding based on wavelet[J]. *Mini-Micro Systems*, 2003, 24(9): 1677 ~ 1681. [王相海. 基于小波的图像可分级编码研究[J]. 小型微型计算机系统, 2003, 24(9): 1677 ~ 1681.]
- 5 Pennebaker W B, Mitchell J L. JPEG Still Image Data Compression Standard[M]. New York: Van Nostrand Reinhold, 1992.
- 6 Lewis A S, Knowles G. Image compression using the 2-D wavelet transform[J]. *IEEE Transactions on Image Processing*, 1992, 1(4): 244 ~ 250.
- 7 Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients[J]. *IEEE Transactions on Signal Processing*, 1993, 41(12): 3445 ~ 3462.
- 8 Said A, Pearlman W. A new, fast and efficient image codec based on set partitioning in hierarchical trees[J]. *IEEE Transactions Circuits and System for Video Technology*, 1996, 6(3): 243 ~ 249.
- 9 Servetto S D, Ramchandran K, Orchard M T. Image coding based on a morphological representation of wavelet data[J]. *IEEE Transactions on Image Processing*, 1999, 8(9): 1161 ~ 1174.
- 10 Taubman D. High performance scalable image compression with EBCOT[J]. *IEEE Transactions on Image Processing*, 2000, 9(7): 1158 ~ 1170.
- 11 Ordentlich E, Weinberger M, Seroussi G. A low-complexity modeling approach for embedded coding of wavelet coefficients[A]. In: *Proceedings of IEEE Data Compression Conference [C]*, Snowbird, UT, 1998: 408 ~ 417.
- 12 Li J, Lei S. Rate-distortion optimized embedding[A]. In: *Proceedings of Picture Coding Symposium [C]*, Berlin, Germany, 1997: 201 ~ 206.
- 13 Antonini M, Barland M, Mathieu P, *et al.* Image coding using wavelet transform[J]. *IEEE Transactions on Image Processing*, 1992, 1(2): 205 ~ 220.
- 14 ISO/IEC JTC1/SC29/WG11, Overview of the MPEG-4 Standard, N3536[S], 2000.
- 15 Wang X H. Researches on hybrid spatial and rate scalable image coding based on zerotree wavelet[J]. *Computer Science*, 2002, 29(6): 123 ~ 125. [王相海. 基于零树小波的图像空间与数率混合可分级编码研究[J]. 计算机科学, 2002, 29(6): 123 ~ 125.]
- 16 Xiong Z, Guleryuz O, Orchard M T. A DCT-based embedded image coder[J]. *IEEE Signal Processing Letters*, 1996, 3(11): 289 ~ 290.
- 17 Monro D M, Dickson G J. Zerotree coding of DCT coefficients[A]. In: *Proceedings of 1997 International Conference on Image Processing (ICIP'97) [C]*, Washington, DC, USA, 1997: 625 ~ 628.
- 18 Yan Y S, Shi Q Y. Reversible integer DCT and lossless image coding[J]. *Software Journal*, 2000, 11(5): 620 ~ 627.
- 19 Wei C, Hao P W, Shi Q Y. Integer DCT-based image coding[A]. In: *Picture Coding Symposium 2001 [C]*, Seoul, Korea, 2001: 175 ~ 178.
- 20 Li J, Li J, Kuo C J. Layered DCT still image compression[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, 7(2): 440 ~ 443.
- 21 Malla S G, Zhang Z F. Matching pursuits with time-frequency dictionaries[J]. *IEEE Transactions on Signal Processing*, 1993, 41(12): 3397 ~ 3415.
- 22 Al-Shaykh O K, Miloslavsky E, Nomura T, *et al.* Video compression using matching pursuits[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, 9(2): 123 ~ 143.
- 23 Bergeaud F, Mallat S. Matching pursuit of images[A]. In: *Proceedings of IEEE International Conference Image Processing [C]*, Washington DC, USA, 1995, 1(10): 53 ~ 56.
- 24 Figueras i Ventura R M. Image Coding with Matching Pursuits[D]. Lausanne, Switzerland: Swiss Federal Institute of Technology, 2001.
- 25 Rabiee H R, Kashyap R L, Safavian S R. Adaptive multiresolution image coding with matching and basis pursuits[A]. In: *Proceedings of IEEE 3rd International Conference on Image Processing, ICIP'96 [C]*, Lausanne, Switzerland, 1996: 16 ~ 19.
- 26 Figueras i Ventura R M, Frossard P, Vanderghyest P. Evolutionary multiresolution matching pursuit and its relations with the human visual system[A]. In: *Proceedings of the European Signal Processing Conference [C]*, Toulouse, France, 2002, 2: 395 ~ 398.
- 27 Frossard P, Vanderghyest P, Figueras i Ventura R M. High flexibility scalable image coding[A]. In: *Proceedings of Visual Communication and Image Processing [C]*, Lugano, Switzerland, 2003: 127 ~ 134.
- 28 Frossard P, Vanderghyest P, Figueras i Ventura R M, *et al.* A posteriori quantization of progressive matching pursuit streams[J]. *IEEE Transactions on Signal Processing*, 2004, 52(2): 525 ~ 535.
- 29 Figueras i Ventura R M, Vanderghyest P, Frossard P, *et al.* Scalable color image coding with matching pursuit[A]. In: *Proceedings of ICASSP 2004 [C]*, Montreal, Canada, 2004, 3: 53 ~ 56.
- 30 Cruz D S, Ebrahimi S. An analytical study of JPEG 2000 functionalities[A]. In: *Proceedings of IEEE International Conference on Image Processing [C]*, Vancouver, BC, Canada, 2000, 2: 49 ~ 52.
- 31 Li W. Overview of fine granularity scalability in MPEG-4 video standard[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(3): 301 ~ 317.
- 32 ISO/IEC JTC1/SC29/WG11, MPEG99/m5122, Results on Fine Granularity Scalability[S]. 1999.
- 33 Wu F, Li S P, Zhang Y Q. A framework for efficient progressive fine granularity scalable video coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(3): 332 ~ 344.
- 34 Sun X Y, Wu F, Li S P, *et al.* Macroblock-based progressive fine granularity scalable video coding[A]. In: *Proceedings of IEEE International Conference on Multimedia and Expo (ICME) [C]*, Tokyo, 2001: 461 ~ 464.

- 35 Sun Xiao-yan, Gao Wen, Wu Feng, *et al.* Macroblock-based progressive fine granularity scalable video coding [J]. *Journal of Software*, 2002, **13**(11): 2134 ~ 2141. [孙晓艳, 高文, 吴枫等. 基于宏块的渐进-精细可伸缩的视频编码[J]. *软件学报*, 2002, **13**(11): 2134 ~ 2141.]
- 36 Sun Xiao-yan, Gao Wen, Wu Feng, *et al.* Macroblock-based progressive fine granularity scalable(PFGS) video coding with flexible temporal-SNR scalabilities[J]. *Chinese Journal of Computers*, 2003, **26**(3): 345 ~ 352. [孙晓艳, 高文, 吴枫等. 基于宏块的具有时域和 SNR 精细可伸缩的视频编码[J]. *计算机学报*, 2003, **26**(3): 345 ~ 352.]
- 37 Yang X G, Ramchandran K. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation [J]. *IEEE Transactions on Image Processing*, 2000, **9**(5): 778 ~ 791.
- 38 Yang X G, Ramchandran K. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation [J]. *IEEE Transactions on Image Processing*, 2000, **9**(5): 778 ~ 791.
- 39 Wang Q, Ghanbari M. Scalable coding of very high resolution video using the virtual zerotree [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, **7**(5): 719 ~ 727.
- 40 Cheng P, Li J, Kuo C. Multiscale video compression using wavelet transform and motion compensation [A]. In: *Proceedings of International Conference on Image Processing[C]*, Washington DC, USA, 1995: 606 ~ 609.
- 41 Wang Xiang-hai. Researches on spatial scalable video coding based on wavelet[J]. *Computer Science*, 2002, **29**(7): 115 ~ 117. [王相海. 基于小波的视频空间可分级编码研究[J]. *计算机科学*, 2002, **29**(7): 115 ~ 117.]
- 42 Shen K, Delp E. Wavelet based rate scalable video compression[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, **9**(1): 109 ~ 122.
- 43 Neff R, Zakhor A. Matching pursuit video coding at very low bit rates [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, **7**(2): 158 ~ 171.
- 44 Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries [J]. *IEEE Transactions on Signal Processing*, 1993, **41**(12): 3397 ~ 3415.
- 45 Neff R, Nomura T, Zakhor A. Decoder complexity and performance comparison of matching pursuit and dct-based MPEG-4 Video codecs [A]. In: *Proceedings of International Conference on Image Processing[C]*, Chicago, Illinois, USA, 1998, **1**: 783 ~ 787.
- 46 Neff R, Zakhor A. Modulus quantization for matching pursuit video coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2000, **10**(6): 895 ~ 912.
- 47 Al-Shaykh O, Miloslavsky E. Video compression using matching pursuits[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1998, **8**(9): 123 ~ 143.
- 48 Rahmoune A, Vanderheynt P, Frossard P. MP3D: highly scalable video coding scheme based on matching pursuit[A]. In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing[C]*, Montreal, Canada, 2004, **3**: 133 ~ 136.