

# 提高频率域可分级视频信源编码效率的研究

郭丹鹏 邱俊海 张春田 刘峰

(天津大学 天津 300072)

**摘要** 在MPEG-2视频编码标准的框架基础上,本文引入了一种提高频率域可分级视频编码效率的方法。通过在 $8 \times 8$ DCT块中合理地分割系数及尽量减少在低分辨率层中所必需传送的块结束码(EOB)的数量,使为获得频率域可分级性所付出的代价明显减小。计算机模拟实验证实了采用该方法可提高编码效率。

**关键词** 视频编码,可分级编码,MPEG-2

## 1 引言

随着视频通信技术的迅速发展,可分级视频编码的研究日益受到重视。由于可分级视频编码可以实现同一视频信源的多分辨率接收,所以在数字HDTV/SDTV(高清晰度电视/标准电视)兼容传输、ATM网中的视频通信、多点会议电视、多媒体终端的窗口显示等许多领域都得到应用。为此,MPEG-2已在标准制定中包含进了可分级视频编码方式<sup>[1]</sup>。

视频编码的可分级性是指从全分辨率的视频压缩编码码流中舍弃其中的一部分,对余下的部分解码,仍能恢复出一完整的视频信号进行显示,只不过分辨率下降到一个较低的水平而已。以HDTV数字广播为例,如果视频信源编码器具有可分级性,就可以将HDTV广播设计成一多层系统,即由演播室产生的HDTV数字信号经信源解码后,能提供给不同码率的传输信道,可被不同价格/性能等级的接收机接收,所接收图象的分辨率,取决于信道可提供的码率和接收机解码器的复杂程度。高级的HDTV接收机可以接收并解码全部HDTV码流,在其大尺寸的屏幕上显示全分辨率的HDTV图象;而价格便宜的SDTV接收机,通过其简化的解码器只从HDTV码

流中解出与其屏幕尺寸相应的SDTV数据,显示普通电视分辨率的图象。此外,电视台也可从HDTV码流中单独取出SDTV部分,在一个低码率的信道中播出,由此实现HDTV与SDTV的同播<sup>[2,3]</sup>。

可分级视频编码可以通过空间域分级、频率域分级、信噪比分级以及和时间域分级结合的组合方案进行。由于离散余弦变换(DCT)在视频压缩编码中表现出的优越性能,它已被采用到JPEG和MPEG等视频编码国际标准中。因为DCT相当于图象信号的频域表示,所以,在DCT域通过将DCT系数块划分成子块的方法划分子带,可以方便地实现频率域的可分级视频编码<sup>[4]</sup>。

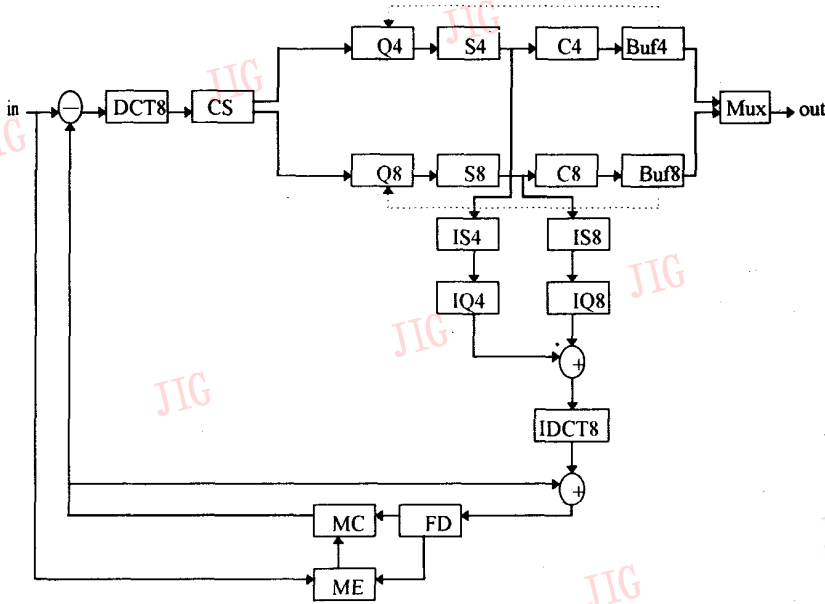
从信息论角度分析,为获得信源编码的可分级性,与不分级的单层编码相比,可分级的编码效率必然有所下降,即为了在每一分辨率层都获得与不分级编码相当的图象质量,必须多付出一定的码率开销作为代价<sup>[3]</sup>。本文的工作旨在提高频率域可分级编码的效率,减少由于实现分级所花费的码率开销。

## 2 频域可分级视频编码的基本方案

图1是一个由标准的MPEG-2视频编码器发展而成的两层的频域可分级活动图象视频编码基本方案的原理框图,频域分级是通过分割DCT块的

• 国家自然科学基金资助的课题

收稿日期:1996.06.10;收到修改稿日期:1996.07.31



图例: DCT: 离散余弦变换    IDCT: 反离散余弦变换  
 Q: 量化    IQ: 反量化  
 S: 扫描    IS: 反扫描  
 C 变字长编码    CS: 系数选择  
 MC: 运动补偿    ME: 运动估值  
 FD: 帧延迟    Buf: 缓存器

图 1 频域可分级视频编码原理框图

Fig. 1 Principle block diagram of a frequency scalable video encoder

系数实现的。

对于标准的单层 PMEG-2 编码器而言, 经过  $8 \times 8$ DCT (DCT8) 后的变换系数再经量化、zigzag 路径扫描、变字长编码和输出缓存器的码率控制后形成恒定码率输出码流。同时, 量化后的 DCT 系数经反量化、反离散余弦变换 (IDCT8) 与运动补偿 (MC) 预测值相加, 在编码端构成本地的解码复原图象, 做为后续图象时域预测的参考图象。

为了实现频率域分级, 图 1 中在 DCT8 后, 通过系数分割 (CS), 形成了低频、高频两个子带, 然后对两个子带分别进行量化、编码。将全分辨率图象  $8 \times 8$ DCT 系数块中左上角  $4 \times 4$  个系数单独选出形成子块 (图 2 中左上角虚线框内的 16 个系数), 其相应于低 (分辨率) 层的图象信息, 划为低频子带; 而把其余 48 个 DCT 系数划为高频子带, 所含为高 (分辨率) 层的图象增强信息。

低频和高频两个子带的 DCT 系数各自经过量化 (Q4, Q8)、 $4 \times 4$  或  $8 \times 8$  的 zigzag 路径扫描和变字长编码 (C4, C8), 而后分别经过输出缓存器 (Buf4, Buf8), 开成低层和高层两个恒定码率的码

流。两层码流的码率根据系统的总码率按一定比例分配。两层的码流按照 MPEG-2 可分级编码的句法规定, 写入规范的头部信息, 最后复合构成可分级的两层的视频编码码流。

同时, 高低两层量化后的 DCT 系数分别经过反 zigzag 扫描 (IS4, IS8), 反量化 (IQ4, IQ8) 后相加, 重新合成一全分辨率的  $8 \times 8$ DCT 块 (但含有量化误差), 再经 IDCT8, 反变换到空间域。其余的过程与前述的单层 MPEG-2 编码相同。

众所周知, 任何可分级编码方案, 由于可分级性的要求, 必须在总的编码码流中增加一部分头部信息, 做为区分不同层次数据的识别标志, 以便实现可分级解码, 这是不可避免的。但是, 上述的频域可分级编码基本方案还由于以下两方面的原因使编码效率下降:

(1) 块结束码 (EOB) 增加

由于每个  $8 \times 8$ DCT 块需要做  $4 \times 4$  (低层) 和  $8 \times 8$  (高层) 两次编码, 而每次编码都需要一个 EOB 码, 因此, EOB 码增加了一倍, 由其引起的码字增加相当可观。

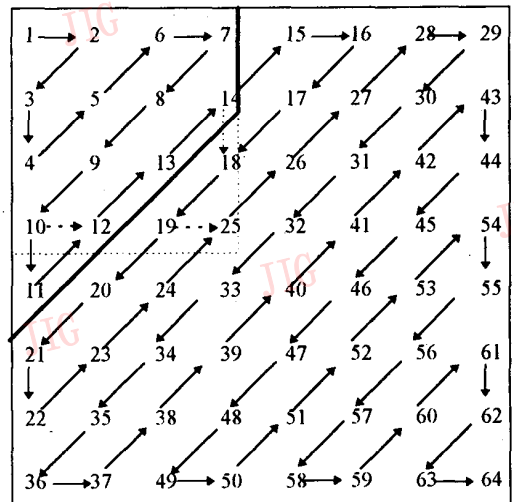


图 2 DCT 系数分割

Fig. 2 Segmentation of DCT coefficients

(2)系数分割造成连“0”串中断

在将二维 DCT 数据转成一维数组形式时,采用 zigzag 扫描路径的原因是根据统计结果,DCT 系数的方差有按照 zigzag 路径递减的趋势。因此,按 zigzag 扫描可以有较多机会获得较长的连“0”系数串,配合由“0”系数游程和在此游程之后的非“0”系数数值组成的二维事件(run,level)的 Huffman 编码,有利于码率压缩。

但采用基本方案的系数分割后,却使这种处理的效果受到影响。例如,在图 2 中,当系数  $a_{13}, a_{14}, a_{15}, a_{16}, a_{17}$  都为 0,而  $a_{18}$  不为 0 时,对于原  $8 \times 8$  块,由二维 Huffman 编码,可将  $a_{13} \sim a_{18}$  作为一个事件(5,  $a_{18}$ )编码。但在分出  $4 \times 4$  子块后,由于该子块内扫描路径是  $4 \times 4$  zigzag,沿此路径系数的排序为  $a_{13}, a_{14}, a_{18}$ ,因此造成了连“0”串的中断,使一个编码事件分成了两个事件(2,  $a_{18}$ )和(3,  $a_{20}$ )分别在低两层中编码,从而影响了编码效率。

针对上述问题,本文提出了如下提高编码效率的改进措施。

### 3 提高频域可分级视频编码效率的方法

本方法的特点是改进了系数分割方式,其划分是严格按  $8 \times 8$  DCT 系数块中 zigzag 扫描顺序进行的<sup>[4]</sup>。将图 2 中的 zigzag 路径上的系数  $a_1 \sim a_{14}$  划出做为低频子带,其余系数  $a_{15} \sim a_{64}$  做为高频子带。然后分别进行量化和(run,level)二维 Huffman 编码。该方法由于采取了如下的编码技巧,减少了低层的 EOB 数量,并且避免了“0”串的中断,提高了系统的编码效率。具体编码方法如下:

#### 3.1 低层编码及 EOB 处理

(1)判断  $a_{14}$  是否为 0。若  $a_{14} \neq 0$ ,则低层编码结束,不写 EOB;

若  $a_{14} = 0$ ,则将低层中最末的连“0”系数与高层中的第一个事件(H-event-1)共同组成一个新事件(N-event)。

(2)判断 N-event 在 MPEG-2 给出的变字长码表中有无对应码字。若有对应码字,

则将 N-event 编码到低层中,并且低层不写 EOB。由于 H-event-1 已被编码到低层中,所以低层应向高层反馈这一信息,用 1bit 标志符 skip=1 通知高层。高层编码时,应跳过 H-event-1 再开始编码;若无对应码字,则 H-event-1 不在低层编码,并在低层块编码结束时写 EOB。同时低层向高层反馈信息:skip=0,通知高层编码应从 H-event-1 开始。

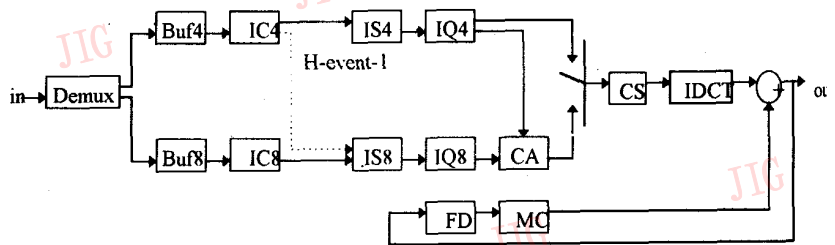
#### 3.2 高层编码

(1)若 skip=1,跳过 H-event-1,从 H-event-2 开始二维 Huffman 编码;若 skip=0,从 H-event-1 开始编码。

(2)按照常规,高层每块编码完毕,要写 EOB。与编码端相应的频域可分级解码原理框图如图 3 所示。

从总码流中分出的低层码流用作低分辨率图象显示。由于此层的 DCT 系数是  $a_1 \sim a_{14}$ ,其中  $a_{11}$  在  $4 \times 4$  块外,而  $a_{18}, a_{19}, a_{25}$  未在低层传送,故在做 ID-CT4 时,舍弃  $a_{11}$ ,并设  $a_{18} = a_{19} = a_{25} = 0$ 。由此会引起低层图象损失一些斜向高频,但统计结果表明,对于一般的自然图象来说,这种斜向高频(尤其经过视觉加权量化后)数值较小。

改进方案的其它部分与图 1 中所示的基本方案相同。运动估值在全分辨率图象上进行,运动矢量则由低层传出。由于低层的分辨率在水平和垂直两个方向上皆为全分辨率的 1/2,所以低层解码复原所用的运动矢量应当除 2。



图例: Demux:解多路复用 Buf:缓存器  
IC:反变字长编码 IS:反扫描  
IQ:反量化 CA:系数复合

H-event-1:将低层中编码的属于高层块的第一个事件送回原高层中

图 3 频域可分级视频解码原理框图

Fig. 3 Principle block diagram of the frequency scalable video decoder

## 4 计算机模拟实验

在 Sun Sparc 10 工作站上,利用两个 CCIR 601 活动图象序列: Flower Garden (50 帧) 和 Mobile (40 帧) 进行了本文第二、第三节给出的基本方案改进前后的性能对比实验。两个序列的图象格式为  $576 \times 720$ , 场频为 50Hz, 2:1 隔行扫描, 编码采用的色分量格式为 4:2:0。

在所进行的模拟实验中,每个图象组(GOP)由 12 帧图象组成,为 IBBPBBPBBPBB 形式,其中 I 帧为帧内编码帧,P 帧为前向预测帧,B 帧为双向预测帧。采用了 MPEG-2 TM5 建议中给出的量化加权矩阵和自适应量化方法以及关于可分级编码的缓冲存储器控制策略<sup>[5,6]</sup>。运动估值中运动矢量的搜索范围为  $-16 \sim +15.5$  像素,精度为  $1/2$  像素。

实验对改进前后的编码效率采用两种方法进行比较。

(1) 在规定的同一码率下,比较改进前后的解码重建图象的信噪比 PSNR,作为客观图象质量的比较。本实验规定系统编码总码率为 5.375Mb/s, 低层和高层按 6:4 分配码率,分别为 3.225Mb/s 和 2.150Mb/s。表 1 列出了两种情况下得到的低分辨率层图象和全分辨率层图象的解码信噪比,并且按 Y, U, V 三个分量分别给出。

表 1 解码重建图象的信噪比 PSNR (dB)

Table 1 PSNR of decoded pictures in dB

图象序列	可分级编码	低分辨率层图象			全分辨率层图象		
		Y	U	V	Y	U	V
Flower	改进前	28.0	31.6	33.2	29.7	31.8	33.4
	改进后	28.0	31.8	33.4	30.1	32.2	33.7
Garden	改进前	27.4	32.1	33.7	28.4	32.2	33.7
	改进后	27.4	32.3	33.9	28.8	32.5	34.1

表 1 表明,采用改进措施后,对于全分辨率的解码图象,可以获得 0.4dB 左右的质量改善;对于低分辨率层的解码图象,亮度分量(Y)的信噪比改进前后持平,而对于色度分量(U, V),改进后仍获得 0.2dB 的改善。这一结果所反映的是由于改进后的方案使可分级所花费的附加码率减小,因而使真正用于图象信息本身编码的码率得到增加,从而提高了图象质量。

应当注意,在采用了新的系数分割方式后,由于  $4 \times 4$  IDCT 少用了  $a_{18}, a_{19}, a_{25}$  三个系数,低层的斜向高频成分受到一定损失,但尽管如此,低分辨率层图

象的 Y 分量信噪比仍达到原有的水平,并且 U 和 V 分量还有所提高,这是由于低层 EOB 的减少所换取的改善。

(2) 比较改进前后为获得同样的图象质量,平均编码每帧图象各自所需的比特数。为此中断输出缓冲器的反馈控制,将对 DCT 系数的自适应量化变为固定量化。实验中对两种方案都选定量化等级为 TM5 中 31 级量化的第 11 级,即  $mquant = 11$ ,通过这种手段来保证进行两种方案比较时得到相同的输出图象质量(信噪比在 30dB 左右)。显然,由于失去了输出缓存的闭环控制,系统的输出码率不再恒定,输出码率随编码方案效率的提高而下降。

表 2 列出了采用改进措施前后进行上述对比实验,对于低分辨率层图象和全分辨率图象的 I, P, B 三种类型帧,编码每帧图象所用的平均编码比特数。表 2 的结果表明,采用改进措施后,无论对于低分辨率层图象还是对于全分辨率图象,编码效率都有所改善。为获得同样的图象质量,改进后约可节省 10% 左右的比特数,改善程度取决于被编码图象序列的复杂程度(包括运动),复杂程度较低的图象序列,改善更多。

表 2 固定量化等级  $mquant = 11$  时,平均编码每帧图象所用比特数(单位:  $10^5$  bits)

Table 2 Average coding bits ( $\times 10^5$ ) per frame by fixed quantization level  $mquant = 11$

图象序列	可分级编码	低分辨率层图象			全分辨率层图象		
		Y	U	V	Y	U	V
Flower Garden	改进前	4.35	1.42	0.88	6.90	2.54	1.49
	改进后	3.19	1.32	0.79	5.81	2.03	1.44
Mobile	改进前	5.22	1.75	0.81	8.64	3.20	1.33
	改进后	4.90	1.68	0.71	8.08	3.03	1.18

## 5 分析与讨论

本文获得的编码效率改善主要来源于低分辨率层 EOB 的处理。因为按照 MPEG-2 规定<sup>[1]</sup>,每一  $4 \times 4$  DCT 块,在任何情况下(包括最后为一连“0”串的情况)都要加一 EOB。而本文的方法,实际只在系数  $a_{14} = 0$  且由其与高层组成的事件 H-event-1 在 MPEG-2 变字长码表中无对应码字时,才在低层传一个 EOB,而这种情况是十分少见的。

本文的实验结果表明,所采取的改进措施,不仅对于帧内编码的 I 帧有效,而且对于帧间编码的 P 帧和 B 帧同样有效。

分析和实验表明<sup>[7]</sup>,由于P帧和B帧的编码对象——运动补偿帧差图象(残差图象)的统计特性决定,DCT加zigzag扫描的编码性能不再象对自然图象编码时那么优越。但同时也发现,尽管残差图象的DCT系数依空间频率递减分布趋势已不明显,但由于经过良好的运动补偿的这种残差信号能量很小,特别是加上所采取的依主观视觉特性的加权处理,对中、高频DCT系数作了很强的衰减,因此,在DCT系数块的中、高频区出现大量“0”和连“0”串的机会仍很高。这一事实或许是MPEG-2以及最近公布的极低码率视频编码标准H.263<sup>[8]</sup>仍在采用DCT+zigzag扫描的一个原因。(当然,寻找较DCT更好的编码方法来处理残差图象已成为当前的研究热点之一。)这一事实也是解释我们的改进措施对于I,P,B帧均能有效的根据所在。

## 6 结束语

可分级视频信源编码在视频通信领域有广泛的应用。本文讨论了影响频率域可分级视频编码效率的主要因素。在MPEG-2视频编码框架基础上所提出的改进可分级编码效率的措施通过合理划分子

带,减少EOB数量,明显减少了为获得可分级性所付出的码率开销。

本文采用的基本编码算法取自MPEG-2 TM5,对TM5算法进行的各种改进,原则上都可以被吸收到上述可分级编码系统中来,从而使该系统的编码效率进一步提高。

## 参考文献

- 1 ISO/IEC 13818-2, Generic coding of moving pictures and associated audio, 1993, (Nov).
- 2 D. Anastassion. Scalability for HDTV. Int. Workshop on HDTV'92, Kawasaki, 1992, (Nov).
- 3 张春田. 可分级视频信源编码及其在数字HDTV传输中的应用. 电视技术, 1994, (5).
- 4 M. R. Ciranlar and A. Puri. Scalable video coding in frequency domain. Proc. of SPIE Visual Communication and Image Processing, Boston, 1992. (Nov)
- 5 ISO/IEC 1/SC29/WG11. Coding representation of picture and audio information. Test Model 5, Draft Rev. 2, 1993(Apr).
- 6 郭丹鹏. 提高频率域可分级视频信源编码效率的研究. 天津大学硕士论文, 1996, 3.
- 7 张春田, 张劲松. 运动补偿视频编码中DCT编码效率的研究. 电子学报, 1996, (1).
- 8 Draft ITU-T Recommendation H. 263. Video coding for narrow telecommunication channels at <math>64\text{kbit/s}</math>. 1995, (June).



郭丹鹏, 1984年毕业于锦州师范学院物理系, 获理学学士学位, 毕业后在辽宁省丹东市电子职工大学任教。1996年毕业于天津大学研究生院信号与信息处理专业, 获工学硕士学位。现在丹东市人民政府人事局工作。

# Research on Efficiency Improvement of Scalable Video Source Coding in Frequency Domain

Guo Danpeng, Qiu Junhai, Zhang Chuntian, Liu Feng

(Tianjing University, Tianjin, 300072)

**Abstract** An efficient frequency scalable video coding method based on the framework of MPEG-2 video coding standard is introduced in this paper. The penalty due to frequency scalability is reduced considerably by reasonably partitioning coefficient in the  $8 \times 8$  DCT block and minimizing the number of the EOB to be transmitted in the low resolution layer. The efficiency improvement by using this coding method has been confirmed by computer simulations.

**Keywords** Video coding, Scalable coding, MPEG-2