

基于 H. 264标准的混合编码算法分析

余胜生 张 剑 周敬利

(华中科技大学计算机科学与技术学院 国家外存储系统专业实验室 武汉430074)

摘 要 H. 264标准是 ITUT/VCEG 和 ISO/MPEG 委员会联合制定的视频通信新编码标准,它在继承现有混合编码视频标准分块变换、运动估计/补偿、量化和熵编码等成熟技术的基础上,还采用了多种新技术,保证了编码的有效性。本文介绍了最新 H. 264标准中的差分编码、块匹配运动预测法、归一化准则、可变尺寸块运动补偿、多假设运动补偿。最后,对运动补偿中不同的块大小编码方法进行了实验分析,实验结果表明:对于通常视频内容,选择 8×8 及更大的分块模式、 $1/4$ 像素运动补偿精度和混合编码,能够获得最经济的压缩效果。

关键词 视频编码标准,运动补偿,差分编码,归一化准则

The Algorithm Analysis of the Hybrid Encoding Based on H. 264 Standard

YU Sheng-Sheng ZHANG Jian ZHOU Jing-Li

(Huazhong University of Science and Technology, Wuhan 430074)

Abstract H. 264 is a new generation video coding standard being established up by ITUT VCEG and ISO/IEC MPEG for video communication. On the basis of mature techniques such as transform on block, motion esteem action/com sensation, quantization and entropy coding, H. 264 employ some new techniques, which ensure the coding efficiency. First We introduce the H. 264 coding standard, differential encoding, Cake marry movement Predicted method, normalization norm, Variable dimensions, motion compensation, Much assumption motion compensation. In the end, experiments are conducted on different block size coding methods. The experimental results proved that the selection of block partition modes that are 8×8 and larger, $1/4$ -pixel resolution motion estimation and universal variable length coding can achieve the most economical compression results for the common video contents.

Keywords Video coding standard, Motion compensation, Differential encoding, Normalization norm

1 引言

计算机的发展为视频编码技术的研究提供了强有力的工具,大规模、超大规模集成电路的出现,则为编码技术的实现提供了基础。混合编码是近年来提出的一类新的编码技术,它将波形编码和参量编码组合起来,克服了原有波形编码和参量编码的弱点,结合各自的长处,力图保持在 $4\text{-}16\text{ kbit/s}$ 低速率上,能够得到高质量的合成视频编码。混合编码是一种运动补偿、分块变换、差分编码相结合的一种视频编码方法,其多脉冲激励线性预测编码(MPLPC),规划脉冲激励线性预测编码(KPELPC),码本激励线性预测编码(CELP)等都是属于混合编码技术。目前,H. 264标准也是采用 DPCM 加变换编码的混合编码模式的结构和算法,加强了对各种信道的适应能力,有利于对误码和丢包的处理。该算法框架中,运动补偿、变换和差分编码相互影响,不同的运动补偿方法、不同的量化系数、不同的差分编码方法构成不同的编码模式,使 H. 264应用目标范围较宽,可以满足不同速率、不同解析度以及不同传输(存储)场合的需求。而每种编码模式针对不同的内容,又体现不同编码的率失真性能,因此,编码器的控制成为混合编码算法的关键问题,即必须明确“哪一部分视频内容应当用哪种编码方法,以及采用什么样的编码参数”^[1]。如图1所示,给出了典型混合编码算法的框图,其中,需要编码器解决的问题有:按照何种方式将图像划分成基本的编码单位;每个基本的编

码单位采用何种预测方式来进行差分编码;如果利用视频内容时间的相关性来进行预测,采用什么样的运动补偿方式。

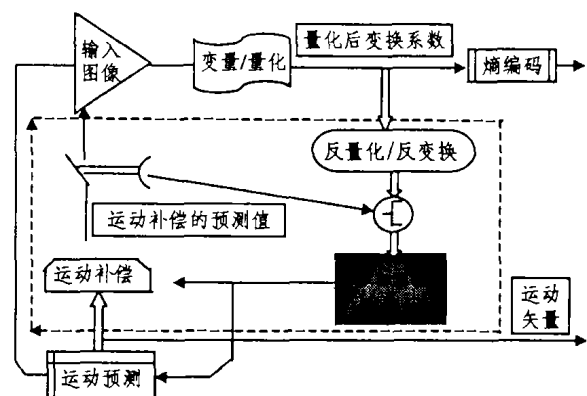


图1 混合编码拓扑图

2 混合编码算法

2.1 基于内容的视频对象划分

在早期的基于混合编码框架的视频编码标准 H. 261中,大部分视频编码算法采用规则矩形对图像进行划分,运动补偿的基本单位块为 16×16 ,图像正好被划分成一系列 16×16 的矩形区域宏块,编码器以这些矩形区域为单位对图像进行

编码。而在 H. 264 标准中, 图像进行划分单位不再局限于 16×16 矩形区域宏块, 编码后的每个宏块具备自己的编码模式、量化信息、运动信息和残差变换系数。为此, 引入了基于内容的思想和视频对象概念, 允许以内容为依据对视频图像进行任意形状的划分^[2]。但是, 视频对象划分涉及对视频内容的分析和理解, 这与人工智能、图像理解、模式识别和神经网络等学科有密切联系。目前, 人工智能的发展还不够完善, 计算机还不具有观察、识别、理解图像的能力, 同时, 要实现正确的图像划分, 需要在更高层次上对视频内容进行理解。尽管 MPEG-4 或 H. 264 框架已经制定, 但至今仍没有根本解决视频对象划分问题, 基于语义的划分则更是难题。鉴于视频对象定义的主观性, 多数划分技术仅停留在半自动的层面上, 只有一些简单的划分方法, 如运动物体和背景的分割等可以实现全自动化。

2.2 差分编码

充分利用基于内容视频对象活动图像的时间相关性, 进行帧间差分编码是视频压缩的主要技术, 也是视频编码和静止图像编码的根本区别。通常对于视频序列, 除去运动引起的一些细微差别, 相邻的图像基本是相同的, 因此, 仅对发生变化的部分进行编码能够大大地提高压缩效率。在早期的混合编码方案中, 每个宏块通常选择: SKIP、INTRA、INTER 和 INTER4V 四种编码模式, 作为基本的编解码单元。新近制定的标 H. 264, 由于采用了多种帧内预测技术、双向帧间预测技术以及更成熟的可变尺寸块运动补偿技术, 编码模式已增加到二十种。

运动补偿技术是 INTER 模式的基础, 如果直接对相邻两帧的差分数据进行编码, 对于发生运动的区域, 其编码效果一团糟。而视频序列各帧图像的差异多数是由运动产生的, 因此, 在差分编码之前, 首先需要得到待编码图像各点相对于前一帧图像的运动信息, 然后在此运动信息的基础上找到其预测值的位置, 这个过程称为运动补偿。实验证明, 运动补偿可以有效地减小帧间差分数据的方差, 使其近似于一个常数。由率失真的理论可知, 信源方差数据的降低可以有效降低码率。

定义1 如果假设第 n 帧图像在 $S=(x, y)$ 点的值为 $Y_n(s)$, $\hat{Y}_R(S^*)$ 为 n 帧图像在参考图像中 $S^*=(x^*, y^*)$ 点的亮度值, $V(s)$ 为当前图像中 $S=(x, y)$ 点相对于参考图像的运动矢量。则差分数据 $D_n(s)$, 为:

$$D_n(X, M) = Y_n(s) - \hat{Y}_R(S^*) \quad (1)$$

$$S^* = S + V(S) \quad (2)$$

式中, $\hat{Y}_R(S^*)$ 通常为前一帧或前几帧图像的本地解码重建图像。

2.3 块匹配运动预测法

我们把计算待编码图像各点的运动矢量的过程叫做运动预测, 也可以称为位移估计或位移测量。如: 块匹配法、梯度法、相位法等^[3]。其中, 块匹配法是视频编码中最常用的运动预测方法, 它基于图像中的运动均为方块的平移, 即同一方块内的各点具有相同的运动矢量。利用块匹配法进行运动预测的就是将图像中, 某个块内的所有像素点和参考图像中相同位置、相同大小块内的所有像素点进行比较, 如果没有发生运动的话, 它们的值将保持极高的相关性, 否则, 就在给定的区域内进行搜索, 寻找一个最佳匹配块。在实际中如果方块足够小, 多数复杂运动, 如旋转、缩放、形变等都是可以由基于方块的平移来近似的。

2.4 归一化误差最小准则

归一化互误差函数(NCCF, Normalized Cross Correlation Function)准则中误差最大准则和最小准则是常用的最佳匹配准则, 由于最大准则计算复杂度较高, 在块尺寸比较小的情况下精确度很差, 因此视频编码中的运动预测一般采用最小准则。

定义2 假设当前图像 $S=(x, y)$ 处, 一个大小为 $M \times N$ 的块, (i, j) 为对应的整数运动矢量, 在参考帧位移为 R 附近作双线性插值, 得到 $(2 \times P + 1) \times (2 \times P + 1)$ 的数组 $\hat{Y}_R(x+m+i, y+n+j)$, 则其相对于参考图像各位置的绝对误差和函数为:

$$S_{SAD}(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |Y_n(x+m, y+n) - \hat{Y}_R(x+m+i, y+n+j)| \quad (3)$$

这里, $-1 \leq i, j \leq 1$, 则最小准则为:

$$S_{SAD}(i_0, j_0) = \min\{S_{SAD}(i, j), -1 \leq i, j \leq 1\} \quad (4)$$

其中, 绝对误差和最小的位置就是最佳匹配位置。

如果规定 $-P < i < P$, $-P < j < P$, 求该运动矢量的过程, 就是在参考图像 $(2 \times P + 1) \times (2 \times P + 1)$ 区域内搜索的过程。为了寻找最佳匹配位置, 需要计算 $(2 \times P + 1) \times (2 \times P + 1)$ 区域内所有位置对应的 SAD 值, 但这样的计算量是很大的, 为此, 人们提出了许多简化的快速算法, 如: 搜索窗内均匀搜索的二维对数法和三步法、偏向中心搜索的三步法和四步法、基于块的梯度下降法(BBGDS)、交叉法(CS)和菱形法(DS)、线性正方形搜索算法、APDZS 算法、MVFAST 算法、HEXBS 算法以及基于遗传算法的运动搜索方法^[1]。这些算法对初始搜索位置的预测、搜索中止门限确定和搜索区形状的选择等问题进行了深入的研究, 并通过设计不同的搜索模板和搜索策略, 来减小计算复杂度。

2.5 可变尺寸块运动补偿

匹配块尺寸问题、多假设问题、长时期预测问题和复杂运动模型问题, 也是改进运动补偿技术性能的关键性问题。早期的运动补偿技术都是基于固定尺寸块, 文[4]从运动补偿的效果、描述运动矢量所需的数据量和算法复杂度三方面进行折衷考虑, 16×16 或 32×32 是最佳的运动补偿块尺寸。如前所述, 基于块的运动补偿是建立在把所有运动归结为以块为单位的平移这样一个假设上的, 因而块的尺寸越小, 对运动的描述就越精确, 但是不可能通过进一步减小块的尺寸来提高运动补偿的性能, 这样又增加了描述运动矢量所需的数据量。如果采用可变尺寸的块来进行运动补偿, 对符合上述假设的运动部分采用较大尺寸的块, 对不符合上述假设的运动部分(如旋转、缩放、形变等)采用较小尺寸块, 势必会在整体上改善运动补偿的效果^[4]。基于该思想, 文[4]中给出了一种可变尺寸块的运动补偿方案。其可变尺寸块的运动补偿算法, 在 H. 264/AVC 视频编码标准中, 已经得到了应用并取得了较好的效果。

2.6 多假设运动补偿

为了保证在随机变量进行预测过程中, 尽可能满足均方误差最小准则。新一代视频编码都采用了分数像素运动补偿、重叠块运动补偿(OBMC)以及环路滤波器多假设的思想, 这也是当前人们一直努力研究的一个课题。

定义3 假设概率分布为 $h_p(s)$ ($p = 1, 2, \dots, P$), s 点的预测值存在 P 中, $Y_n(s)$ 为待编码图像中 $s=(x, y)$ 点的值, $v_p(s)$ 是点 s 相对于参考图像的一个运动矢量, 而 $\hat{Y}_R(S + V_p(S))$ 为 s 点在参考图像中一个经过运动补偿的预测值, 则

差分编码数学期望值为:

$$D_n(s) = Y_n(s) - \sum_{p=1}^P h_n(s) \cdot \hat{Y}_R(s + v_p(s)) \quad (5)$$

式(4)是由式(1)和式(2)演化而来的,它表示对于s点有p个可能的运动矢量。

分数像素运动补偿技术要求以原参考图像插值后的图像作为新的参考图像,然后进行运动补偿预测,从而运动矢量的精度达到分数像素精度。改善分数像素运动补偿性能主要方法:一是提高分数像素精度,二是选择合适的插值滤波器。文[4]中对 Since 函数滤波器、双线性滤波器和维纳滤波器的插值性能进行了详细的分析和比较,同时指出超过一个极限精度后,继续提高精度来改善分数像素运动补偿的性能其意义不大。在 H. 264 标准制订过程中,虽然曾经企图采用 1/8 像素精度的运动补偿,但是最终还是定为 1/4 像素。不过 H. 264 采用了 6 阶 FIR 滤波器的内插获得 1/2 像素位置值,当 1/2 像素值获得后,1/4 像素值可通过线性内插获得。理论上,如果将运动补偿的精度增加一倍(例如从整像素精度提高到 1/2 像素精度),可有 0.5bit/Sample 的编码增益,但是实际验证发现,在运动矢量精度超过 1/8 像素后,系统基本上就没有明显增益了,因此,在 H. 264 中采用 1/4 像素精度的运动矢量模式,而不是采用 1/8 像素的精度,分像素运动矢量补偿与整像素运动矢量补偿相比,前者显著改善了编码复杂性方面的花费,1/4 像素的准确度大大超过 1/2 像素准确度产生的内插分像素样本。

OBMC 是 H. 264 标准中的附件技术。如果某宏块采用 8×8 的块进行运动补偿预测时,预测模式允许该宏块中 4 个 8×

8 亮度块,各对应一个运动矢量;两个色度块的运动矢量,取这 4 个亮度块运动矢量的平均值,每个块对应的 5 个运动矢量,为该块本身和上下左右四个相邻块的运动矢量。使用重叠的块运动补偿,各运动矢量的权值,因点的位置不同而不同,8×8 亮度块的每个像素的补偿值,由 3 个预测值加权平均得到,这样提高了预测精度,减少了块效应,提高了编码增益。

环路滤波器技术是为了尽量消除块变换造成的块效应现象。一个边缘两边的像素点是否需要滤波是由两边像素亮度之差的绝对值决定的,判定的门限值与量化系数有关。当绝对值较大时,表明存在方块效应需要滤波,但是,如果绝对值太大时,超出了当前量化系数所能引起的误差,说明这个差值是图像内容变化引起的,是不需要滤波的。对于需要滤波的边缘部分,还要根据两边块的预测方式和运动矢量等数据来决定滤波的强度。块边缘信息的不同采用不同的滤波权重,因而可以在有效消除块效应的同时又不会影响图像锐度,经过滤波的图像方块效应明显减小,同时图像中的细节部分基本没有变化,避免了滤波引起的图像模糊。此滤波器通常是基于 4×4 的块边缘,在这个块中每个边缘的两个像素可以用 3 抽头的滤波器来进行修正,使用块效应去除滤波器使图像的主观质量有相当大的提高。

3 JM-9.0 编码器的测试分析

使用 H. 264 JM-9.0 编码器,选择几个不同的数字视频序列,如图 2 所示^[5]。

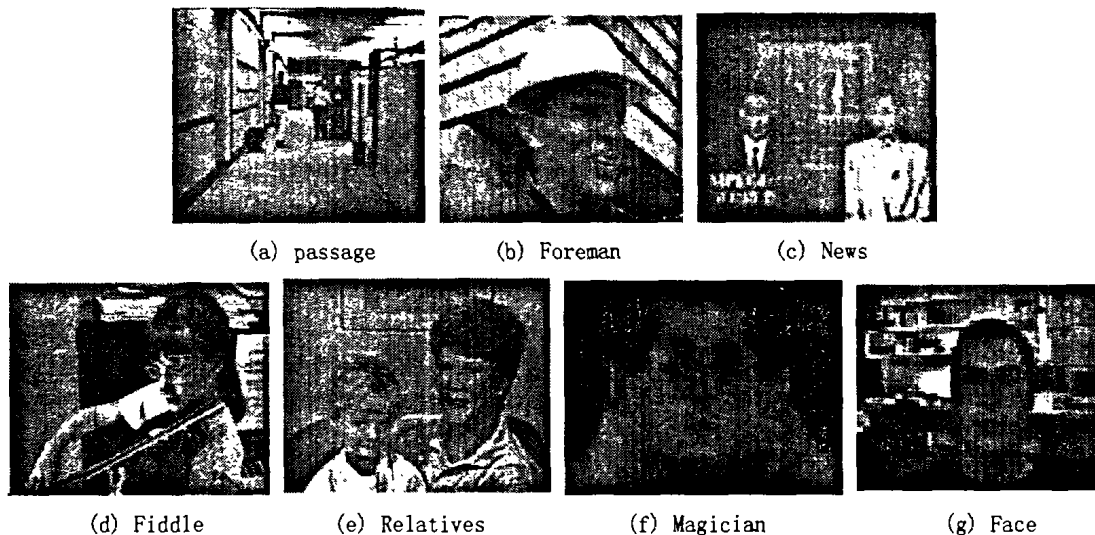


图 2 I 帧数字视频序列图

表 2 不同块大小模式编码

1/4 像素 UVLC				16×16	16×16, 18×8	16×16, 18×8, 8×16	16×16, 18×8, 8×16, 16, 8×8	16×16, 18×8, 8×16, 16, 8×8, 8×4, 4×8, 4×4	16×16, 18×8, 8×8, 16, 8×8, 8×4, 4×8, 4×4					
格式	帧频 (fps)	视频序列	压缩比	比特节省 (%)	比特节省 (%)	PSNR (Y)dB	比特节省 (%)	PSNR (Y)dB	比特节省 (%)	PSNR (Y)dB	比特节省 (%)	PSNR (Y)dB	比特节省 (%)	PSNR (Y)dB
QCIF	10	passage	200.992	36.004	9.241	0.14	11.172	0.111	11.605	0.153	12.540	0.216	12.487	0.214
		Foreman	65.011	35.501	7.761	0.298	10.872	0.254	11.731	0.332	12.885	0.422	12.600	0.457
		News	94.160	36.461	10.482	0.199	10.762	0.185	12.550	0.238	15.580	0.381	16.160	0.390
CIF	15	Fiddle	90.015	35.723	10.850	0.140	10.533	0.102	12.728	0.171	15.021	0.275	15.295	0.290
		Relatives	70.500	32.990	12.423	0.190	12.251	0.170	14.300	0.233	17.021	0.341	17.651	0.350
	30	Magician	30.002	31.682	6.404	0.214	7.470	0.290	8.211	0.252	8.658	0.375	8.580	0.386
		Face	51.040	33.167	3.899	0.231	5.335	0.250	5.500	0.310	5.390	0.390	4.493	4.415

(下转第 127 页)

商协议的最大缺点是协商各方不能根据目前情况调整自己的协商策略和协商提议,更好地达到协商共识。所以目前协商的趋势是如何制定合理的协商协议实现多轮协商过程,同时也可以更好地应用于现实环境。

3) 个性化服务协商:自治协商的优势就在于提供了可以根据协商方的不同个性,制定不同的协商策略,因此在协商过程中如何获取个性心理因素,如何设计不同心理特征下的协商机制也是目前研究的趋势。

4) 协商中交互语义:自治协商仍然是一个双向交流过程,Agent在协商过程中模拟一些现实应用的行为模式更有助于自治协商的发展。语义网和网格计算等研究热点同样可以在协商中起到不小的作用。

5) 协商中的交互方式:这个问题看起来似乎比较简单,但是目前这方面的研究比较缺乏。自治协商对于现实生活中交互方法的模拟和发展仍然有着巨大的研究空间,特别是在双边协商的模拟情境中。

6) 协商中的信任度问题:协商的交互,信息交换等过程都要依赖于信息度,同时这个因素也直接影响信息的可信度。因此,如何让自己的Agent的信任度提高,同时如何辨别协商对手的信任度,也是近两年的研究热点,并且将具有广阔的应用前景。

以上列出的仅仅是双边多问题自治协商的一些有挑战性问题,不仅如此,自治协商的总体发展趋势就是:个性化、智能化、动态化、高效化。

参考文献

- 1 Raiffa H. The Art and Science of Negotiation. Harvard University Press, Cambridge, MA, 1982
- 2 Chen M K. Agenda in Multi-Issue Bargaining: [when to Sweat the Small Stuff Technical Report]. Harvard Department of Economics, Cambridge, Nov. 2002
- 3 In Y, Serrano R. Agenda Restrictions in Multi-Issue Bargaining. Journal of Economic Behavior and Organization, 2004, 53: 385~399
- 4 Weinberger C J. Selective Acceptance and Inefficiency in a Two-Issue Complete Information Bargaining Game. Games and Economic Behavior, 2000, 31: 262~293
- 5 Fatima S S, Wooldridge M J, Jennings N R. An Agenda-Based Framework for Multi-Issue Negotiation. Artificial Intelligence Journal, 2004, 152: 1~45
- 6 Fatima S S, Wooldridge M, Jennings N R. Optimal agendas for multi-issue negotiation. In: Proc 2nd Int. Joint Conf. on Autonomous Agents and Multi-Agent Systems, Melbourne, Australia, 2003. 129~136
- 7 Fatima S S, Wooldridge M J, Jennings N R. Multi-issue negotiation under time constraints. In: Proc. 1st Intl. Joint Conf. on Autonomous Agents and Multi-Agent Systems, Bologna, Italy, 2002. 143~150
- 8 Sycara K. Negotiation Planning: an AI Approach. European Journal of Operational Research, North Holland, 1990, 46: 216~234
- 9 Sycara K. Utility Theory in Conflict Resolution. Annals of Operations Research, 1988, 12: 65~84
- 10 Sycara K. Problem Restructuring in Negotiation. Management Science, 1991, 37(10)
- 11 Sycara K. Persuasive Argumentation in Negotiation. Theory and Decision, 1990, 28(3): 203~242
- 12 Luo Xudong, Jennings N R, Shadbolt N, et al. A fuzzy constraint based model for bilateral, multi-issue negotiations in semi-competitive environments. Artificial Intelligence Journal, 2003, 148(1-2): 53~102
- 13 Faratin P, Sierra C, Jennings N R. Using similarity criteria to Make Negotiation Trade-Offs. In: the Proc. of the 4th Intl. Conf. on Multi-Agent Systems, Boston: USA, 2000. 119~126
- 14 Klein M, Faratin P, Sayama H, Bar-Yam Y. Negotiation Complex Contracts. MIT Sloan Working Paper No. 4196-01, 2001
- 15 Teasuro G. Efficient Search Techniques for Multi-Attribute Bilateral Negotiation Strategies. In: Proc. of the 3rd Intl. Symposium on Electronic Commerce, IEEE Computer Society, 2002. 30~36
- 16 Li C, Teasuro G. A Strategic Decision Model for Multi-attribute Bilateral Negotiation with Alternating Offers. In: ACM Conf. on Electronic Commerce 2003. 208~209
- 17 Keeney R L, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press, 1976
- 18 Satty T L. The Analytic Hierarchical Process. McGraw-Hill International, New York, USA, 1980
- 19 Linden G, Hanks S, Lesh N. Interactive Assessment of User Preference Models: The Automated Travel Assistant. In: Proc. of User Modeling '97, 1997
- 20 Shearin S, Lieberman H. Intelligent Profiling by Example. In: Proc. of the Conf. of Intelligence User Interfaces. ACM Press, 2001
- 21 Stolze M, Nart F. Well-Integrated Needs-Oriented Recommender Components Regarded as Helpful. CHI Extended Abstracts, 2004
- 22 Ha V, Haddawy P. Toward Case-Based Preference Elicitation: Similarity Measures on Preference Structures. In: Proc. of the Uncertainty in Artificial Intelligence Conf. 1998. 79~88
- 23 Ehrgott M. Multicriteria Optimization. Berlin: Springer-Verlag, 2000
- 24 Inderst R. Multi-issue Bargaining with Endogenous Agenda. Games and Economic Behavior, 2000, 30: 64~82
- 25 Raith M G. Fair-Negotiation Procedures. Mathematical Social Sciences, 2000, 39(3): 303~322
- 26 Brams S J, Taylor A D. Fair Division from Cake-Cutting to Dispute Resolution. Cambridge, MA: Cambridge University Press, 1996

(上接第111页)

对 H. 264 的运动补偿中不同的块大小,在不同视频内容情况下进行编码性能测试,如表2所示,给出了各数字视频序列实验基本参数。实验结果表明,一个 16×16 的宏块可以分成 16×8 、 8×16 、 8×8 、 8×4 、 4×8 、 4×4 七种模式,使用全部七种分块模式可以节省平均约12.5%的比特,而用 8×8 及更大的分块模式(前四种)可以节省平均约11%的比特,用 4×4 块而带来的PSNR提高并不明显,且在一些序列中还导致了比特增加,这是因为虽然小块可以提高运动补偿的精度,尤其是在具有精细运动细节的区域,但是也会增加编、解码器的复杂度和传输运动矢量。

参考文献

- 1 Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC). In: 7th Meeting: Pattaya, Thailand, March, 2003. 7~14
- 2 Saponara S, Blanch C, Denolf K, Bormans J. The JVT advanced video coding standard: Complexity and performance analysis on a tool-by-tool basis. 2003. <http://www.polytech.univ-nantes.fr/pv2003/papers/pv/papers/cr1008.pdf>
- 3 Halbach T, Wien M. Concepts and performance of next-generation video compression standardization. 2002. <http://www.norsig.no/norsig2002/Proceedings/papers/cr1121.pdf>
- 4 Wiegand T. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG. Geneva, Switzerland, January 29-February 1, 2002
- 5 Zhang axiao-feng. New Generation Video Coding Standard H. 264 and its Performance Analysis. Journal of Image and Graphics, 2003, 8(A)(8)