

省略 B 帧运动补偿的转码算法实现 MPEG2 到 MPEG4 视频流转码

解 蓉,王兴国,刘济林

(浙江大学信息与通信工程研究所,浙江杭州 310027)

摘 要: MPEG2 和 MPEG4 作为最流行的通用编码国际标准,将长期并存于不同的系统和网络服务中.为实现不同视频系统和网络间对以上两种压缩标准的兼容性,本文在详细分析它们压缩算法运算复杂度的基础上,提出三种压缩视频流转码算法.重点讨论了两种高性能转码算法的原理与实现过程,它们均是运算复杂性和视频质量之间的折衷方案:一是针对演播室和后处理环境的最高目标视频质量要求的运动参数重用算法;另一个是针对网络服务器和用户终端系统处理性能有限,具有最小延时要求的实时低复杂度转码算法.大量测试序列的仿真结果证明了算法的良好性能.

关键词: MPEG2; MPEG4; 转码; 压缩视频流; 视频信号处理

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2003) 04-0548-04

Omit Motion Compensation of B-pictures Algorithm for MPEG-2 to MPEG-4 Video Transcoding

XIE Rong, WANG Xing-guo, LIU Ji-lin

(Institute of Information and Communication Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract: Abstract: MPEG2 and MPEG4 are two most popular international video compression standards. They will coexist in different systems and networks for a long time. For compatibility of these two video standards, this paper put forward three transcoding algorithms on the basis of detailed analysis of computational complexity. In the paper, the principle and implementation of two efficient transcoding algorithms are emphasized. Both of them are compromise between the computational complexity and reconstructed video quality. One is motion information reusing algorithm, which meets the highest video quality requirement of studio. The other is real-time low complexity transcoding algorithm, which applies to lowest delay requirement of the network server and user terminal with limited computational performance. The simulation results of many test sequences testified the excellent performance of the algorithms.

Key words: MPEG2; MPEG4; transcoding; data compression; video signal processing

1 引言

MPEG2 和 MPEG4 是两种应用最广泛的压缩编码标准. MPEG2 是为数字视频广播应用制订的音视频压缩标准^[1]; MPEG4 标准支持更大范围的目标视频流码率,为个人移动通信、卫星网络系统和跨网络环境应用提供了解决方案^[2]. 然而要实现完全端到端交互的 MPEG4 视频系统还需要一段时间. 基于 MPEG2 和 MPEG4 这两种标准的视频系统,将会在很长时间内共存. 因此,对于未来全球网络的通用性而言,不同标准压缩视频流跨网络连接是必须考虑的关键问题,而在网络间设置视频转码器是有效的解决方案.

视频转码主要有两大类:一是把编码数据流从适合某一传输环境的压缩形式转换成适合于另一传输环境的压缩形式,如同一种标准码流的不同码率适配、时空采样率变化转码

等^[3]. 另一类是把压缩数据流从某一种标准转换成另一种标准. 它们一般用于具有相似压缩算法的标准间的转换,如均为基于帧间运动补偿、帧内 DCT 变换、量化的多种编码标准之间的转码(包括 MPEG, H. 261/263, DV, MPEG-1/2/4 等),本文研究的 MPEG2 到 MPEG4 视频转码即属于该类型转码器.

MPEG2 解码和 MPEG4 编码级联结构是最直观的转码方法,特点是算法运算代价巨大、多代降级明显. 分析该结构可发现运动估值、运动补偿、DCT 和 IDCT 是最复杂的过程^[4,5]. 因此本文转码器设计的主要目标是在完全级联式转码结构的基础上如何尽量避免运动估值和 DCT/IDCT 这两大运算瓶颈,同时保持良好的视频质量. 基于空域的运动参数重用转码结构避免了运动估值过程,可得到很高的视频质量,但仍然需要 IDCT 和 DCT 过程,运算较为复杂. 同时需要帧重排过程,时延大、缓存要求高. 在这种结构的基础上省略运动补

收稿日期:2001-09-13;修回日期:2002-03-02

偿和 DCT/IDCT,即得到基于 DCT 域的重量化结构,去除转码中所有运算瓶颈,使复杂性降至最低,延迟最短。但从下文对 I、P 和 B 帧转码公式的具体分析中可看出,该算法由于没经过运动补偿会导致误差积累,造成图像漂移,严重降低目标视频质量。它只适用于低端消费设备,如机顶盒上应用。为解决这些问题,本文提出了一种省略 B 帧运动补偿和 DCT/IDCT 的转码算法,该算法较好地同时解决了运动估值和 DCT/IDCT 这两大运算瓶颈,是运算复杂性和视频质量之间更好的折衷方案。

2 省略运动估值的级联转码器

转码过程中要考虑的几个关键问题是:视频质量、运算复杂度、处理延迟和硬件复杂性。由于运动估值占整个运算时间的 90%以上,是最费时的过程^[6]。因此在转码器设计时,首先应该从 MPEG2 码流的运动信息中提取 MPEG4 码流的运动参数,从而避免运动估值过程,有效地降低计算复杂性。由此得出一种无漂移的省略运动估值级联转码器,如图 1 所示。这种转码器中 MPEG4 编码器的运动估值过程被省略,比完全级联式转码器简化了一大步。

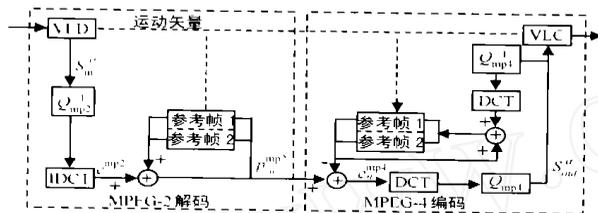


图 1 省略运动估值的级联式转码器框图

MPEG2 码流中有三种类型的图像: I 帧、P 帧和 B 帧。B 帧使用了图 1 中的全部过程, P 帧只使用了反馈回路中两个帧缓存的一个,而 I 帧不需要运动补偿反馈回路。为简化起见,图中省略 Z 扫描和 MPEG4 的 DC/AC 扫描过程, S_{in} 和 S_{out} ($= I, P, B$) 分别代表相应视频帧没进行变长编码和 Z 扫描的输入和输出码流。下面分别分析三种类型视频帧的转码。

2.1 I 帧

I 帧只利用自身信息进行编码,输出 I 帧可由式(1)描述:

$$S_{out}^I = Q_{mp4}[DCT(p_n^{mp2})] \quad (1)$$

其中 p_n^{mp2} 是 MPEG-2 解码重建的像素域视频帧:

$$p_n^{mp2} = IDCT[Q_{mp2}^{-1}(S_{in}^I)] \quad (2)$$

将式(2)代入式(1),考虑到 DCT 运算完全一致,而且具有正交性,可得 I 帧转码公式(3):

$$S_{out}^I = Q_{mp4}[DCT(IDCT[Q_{mp2}^{-1}(S_{in}^I)])] = Q_{mp4}[Q_{mp2}^{-1}(S_{in}^I)] \quad (3)$$

从式(3)看, I 帧转码时可以省略 DCT 和 IDCT 过程,直接在 DCT 域重新量化转码。

2.2 P 帧

P 帧是前向预测编码帧。如果输入的第 n 帧是 P 帧,则它需要重建第 $(n - M)$ 帧作参考帧,此时 M 是两个参考帧之间的距离。输出和输入的预测残差 e_n^{mp4} 和 e_n^{mp2} 与输出和输入码流 S_{out}^P 和 S_{in}^P 的关系分别如式(4)和式(5):

$$S_{out}^P = Q_{mp4}[DCT(e_n^{mp4})] \quad (4)$$

$$e_n^{mp2} = IDCT[Q_{mp2}^{-1}(S_{in}^P)] \quad (5)$$

式(4)中 e_n^{mp4} 可由式(6)得出:

$$e_n^{mp4} = p_n^{mp2} - M_P(p_n^{mp4-M}) \quad (6)$$

而此时 p_n^{mp2} 由式(7)得出:

$$p_n^{mp2} = e_n^{mp2} + M_P(p_n^{mp2-M}) \quad (7)$$

其中 p_n^{mp2-M} 和 p_n^{mp4-M} 分别是 MPEG2 解码和 MPEG-4 编码时重建的第 $n-M$ 帧参考帧, M_P 代表 P 帧的运动补偿运算。将式(5~7)代入式(4)可得:

$$\begin{aligned} S_{out}^P &= Q_{mp4}[DCT\{e_n^{mp2} + M_P(p_n^{mp2-M}) - M_P(p_n^{mp4-M})\}] \\ &= Q_{mp4}[DCT\{IDCT[Q_{mp2}^{-1}(S_{in}^P)] + M_P(p_n^{mp2-M}) - M_P(p_n^{mp4-M})\}] \\ &= Q_{mp4}\{Q_{mp2}^{-1}(S_{in}^P) + DCT[M_P(p_n^{mp2-M}) - M_P(p_n^{mp4-M})]\} \quad (8) \end{aligned}$$

式中 $M_P(p_n^{mp2-M}) - M_P(p_n^{mp4-M})$ 是输入输出运动补偿参考帧间的转码误差。式(8)意味着 P 帧的转码要进行运动补偿将参考帧的转码误差加到输入码流的预测残差上,否则会在一个 GOP 期间造成转码误差的传递与积累,导致图像漂移降质,直到出现下一个 I 帧刷新预测关系,停止漂移。视频序列的漂移是一种严重的多代降质类型,一般 GOP 越长,降质越明显。

2.3 B 帧

对 B 帧的转码只是 P 帧转码运算的一个扩展,唯一的区别是 B 帧的运动补偿包括前向和后向两个参考帧。因此 B 帧的转码公式可由式(8)修改得到:

$$\begin{aligned} S_{out}^B &= Q_{mp4}\{Q_{mp2}^{-1}(S_{in}^B) + DCT[M_B(p_f^{mp2}, p_b^{mp2}) \\ &\quad - M_B(p_f^{mp4}, p_b^{mp4})]\} \quad (9) \end{aligned}$$

其中 M_B 代表 B 帧的运动补偿运算, p_f^{mp2} 、 p_b^{mp2} 、 p_f^{mp4} 和 p_b^{mp4} 分别表示 MPEG2 解码和 MPEG-4 编码重建的前向和后向参考帧。与 P 帧类似, $M_B(p_f^{mp2}, p_b^{mp2}) - M_B(p_f^{mp4}, p_b^{mp4})$ 是 B 帧两个参考帧的转码误差,在进行 B 帧的转码时,也要进行运动补偿,来消除转码误差积累造成的图像漂移。

3 省略 B 帧运动补偿的转码算法

上述转码器避免了最费时的运动估值过程,使转码器大为简化,但没有解决 DCT/IDCT 这一运算瓶颈。转码过程中 DCT/IDCT 是为了进行空域运动补偿,以便消除误差积累造成的图像漂移。如上所述 I 帧无需运动补偿,如果省略图 1 中 P 帧和 B 帧的运动补偿回路,则 I、P 和 B 帧都按照 I 帧转码公式(3)在 DCT 域转码,可称之为 DCT 域重量化转码器。这种结构避免了运动估值、运动补偿、DCT/IDCT 等运算瓶颈,是最简单的转码器结构。然而根据上述分析,由于码流不经过运动补偿, P 帧和 B 帧由于误差积累会造成漂移。GOP 越长,漂移越明显。因而这种转码器适用于 QoS(Quality of Service)要求较低的实时环境。

由于 P 帧是参考帧,如果 P 帧不做运动补偿,在一个 GOP 内以该 P 帧为参考帧的后续 P 帧或 B 帧的误差会越来越大,使视频质量明显下降。而 B 帧不是参考帧,如果 B 帧不做运动补偿,不会造成后续视频帧的误差积累,对视频质量的影响不大。另外,由于 B 帧通常占整个码流的 2/3,数量较多。而且 B 帧运动补偿最复杂,需要有两个参考帧分别作前向和后向

运动补偿,因此对 B 帧省略运动补偿过程从而省略 DCT/IDCT 过程,对降低转码器运算复杂度非常有利,而且这样使得转码器只需一个缓冲器,也减少了整个系统的缓存开销。

由此我们设计出省略 B 帧运动补偿的转码器,该算法对 P 帧按照式(8)进行运动补偿转码,而 I 帧和 B 帧按照式(3)转码,不做运动补偿。由于 I 帧是参考帧,所以 I 帧要进行 MPEG-2 的 IDCT 和 MPEG-4 的反量化和 IDCT,重建视频帧,但可以省略 MPEG-4 的 DCT。而 B 帧由于省略了运动补偿,可省略所有 DCT/IDCT 过程,直接在 DCT 域转码,如图 2 所示。

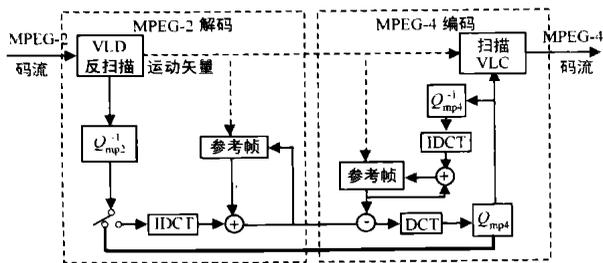


图 2 省略 B 帧运动补偿的转码器结构(粗实线仅对应于 B 帧)

4 仿真结果

我们使用了两个具有不同运动特性的 MPEG2 TM4 测试

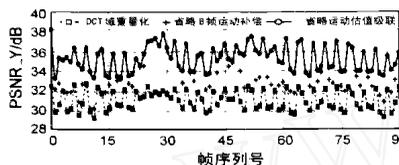


图 3(a) 三种算法对 Stefan 序列转码 Y 分量的 PSNR 曲线

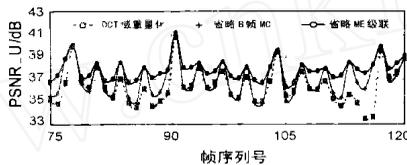


图 3(b) 三种算法对 Stefan 序列转码 U 分量的 PSNR 曲线

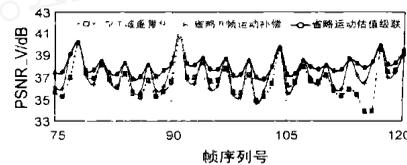


图 3(c) 三种算法对 Stefan 序列转码 V 分量的 PSNR 曲线

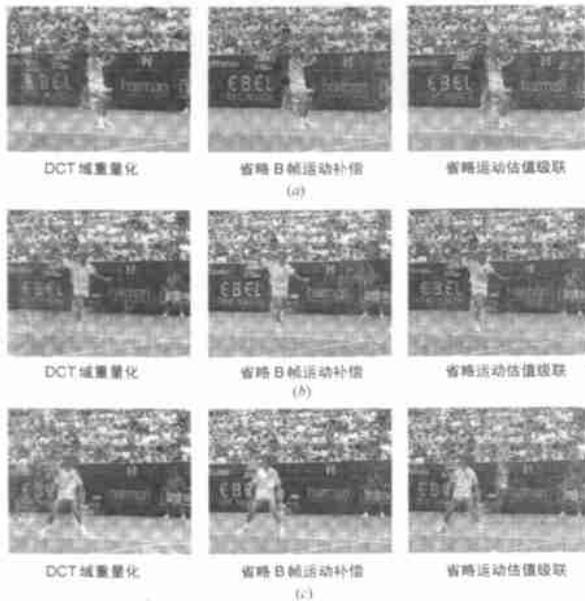


图 4 (a) 三种方法对 Stefan 序列 I 图像转码结构比较;
(b) 三种方法对 Stefan 序列 P 图像转码结构比较;
(c) 三种方法对 Stefan 序列 B 图像转码结构比较;

视频序列: Mobile 和 Stefan。前一个序列运动较为缓慢细节丰富,而后者运动较为剧烈。两序列均为 300 帧 CIF 352 x288 4:2:0 格式,比特率为 2Mbit/s。为便于分析比较省略 B 帧运动补偿转码器的性能,我们同时也实现了 DCT 域重量化转码器和省略运动估值级联式转码器。这三种算法中,省略运动估值方法所需运算最复杂,延迟最长。与它相比,DCT 域重量化方法可节省约 70% 的运算时间,省略 B 帧运动补偿方法可节省约 25% 的运算时间。

图 3 是三种算法对 Stefan 序列转码所得 Y、U、V 分量的 PSNR 曲线图。表 1 是对两序列转码的平均 PSNR 值。图 4 是目标码流中的 I、P 和 B 图像。从图 3 和表 1 可看出 DCT 域重量化方法视频质量比其他两种方法低 3dB 左右,从图 4(c) 可见该方法的 B 图像有较明显的块效应,此时 PSNR 值下降为不到 30dB。省略运动估值级联方法视频质量最好。而对 B 帧做运动补偿和不做运动补偿的差别非常小,后者比前者稍低 1dB,从图 4 中可见主观视觉不易察觉,这主要是由于 B 帧的误差不会积累。与省略运动估值级联式转码方法相比,省略 B 帧运动补偿方法节省了 B 帧的运动补偿、DCT/IDCT 过程,运算时间约有 1/4 的节省,而视频质量仅少许下降。与 DCT 域重量化转码方法相比,这种方法增加了 P 帧 DCT/IDCT 和运动补偿过程,转码延迟有所增加,但视频质量明显提高。因而该

方法对上述两种方法取长补短,较好地解决了 DCT/IDCT 计算瓶颈问题,是操作复杂性和视频质量一种更好的折衷。

表 1 三种方法对 Stefan 和 Mobile 序列转码的平均 PSNR 比较 (单位: dB)

平均 PSNR (dB)	Stefan 序列			MOBILE 序列		
	Y	U	V	Y	U	V
DCT 域重量化	30.8	36.8	36.9	31.2	32.8	31.9
省略 B 帧运动补偿	33.8	37.0	37.1	30.9	34.1	33.4
省略运动估值级联	35.0	38.1	38.2	31.2	34.5	34.1

5 结论

本文首先在完全级联式转码算法的基础上,提出了省略运动估值级联式 MPEG-2 到 MPEG-4 视频转码算法,该算法避免在运动估值这一运算瓶颈,且视频质量很好,但没有避免 DCT/IDCT 过程,速度较慢。DCT 域重量化转码器在此基础上又省略了解码和编码运动补偿环路和帧缓存,最大限度地节省了硬件和算法复杂性。然而由于没有运动补偿环路,视频质量在一个 GOP 中逐渐下降。为了更好的对视频质量和运算速度进行折衷,本文提出了省略 B 帧运动补偿的转码算法,这种算法通过省略占序列 2/3 的 B 图像的运动补偿和 DCT/ID-

CT 过程而降低了算法的复杂性,比省略运动估值级联式转码器节省了约 25% 的时间,而主观视频降质难以察觉.大量的测试序列仿真结果证明,省略运动估值的级联式转码器虽然运算节省有限,但能提供很高的目标视频质量,是非线性后处理中一种较好的选择,省略 B 帧运动补偿转码算法适合运算能力有限、目标视频质量要求较高的实时网络应用环境.

参考文献:

- [1] ISO/ IEC 13818 - 2 Amendment 2. Generic Coding of Moving Pictures and Associated Audio [S].
- [2] ISO/ IEC 14496 - 2. Information technology-Coding of audio-visual objects-Part 2:Visual [S].
- [3] Assuncao, P, et al. Post-processing of MPEG2 coded video for transmission at lower bit rates [J]. IEEE Int Conf Acoust, Speech, Signal Processing, 1996, 4:1998 - 2001.
- [4] Chang, S F, et al. A New Approach to Decoding and Compositing Motion-Compensated DCT-Based Images [J]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993, 5:421 - 424.
- [5] Pedro A, et al. A Frequency-Domain Video Transcoder for Dynamic Bit-Rate Reduction of MPEG2 Bit Streams [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8:953 - 967.
- [6] Bjork N, et al. Transcoder architectures for video coding [J]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1998, 5:2813 - 2816.

作者简介:



解 蓉 女,1975 年 6 月生于山东省济宁市,1996 年毕业于东北电力学院信息系,现为浙江大学通讯与信息专业博士生,主要研究领域是数字视频处理和传输技术.

王兴国 男,1972 年 11 月生,1992 年毕业于浙江大学信息与电子工程学系,2000 年在浙江大学信息与电子工程学系获博士学位,主要研究领域为视频压缩编码技术、多媒体技术与通信系统.

www.cnki.net