

视频转码技术研究现状与最新进展

尚书林^{1,2} 杜清秀¹ 卢汉清² 唐小军¹

摘要 简要回顾了视频转码技术的研究发展状况,系统地归纳了各种转码方法,分析了相关的转码算法的特点.针对最新压缩标准 H.264 的出现,文中特别介绍总结了最新出现的转码方面的研究成果.最后,分析了转码研究未来的发展和可能存在的难点和问题.

关键词 转码, H.264, 码率, 误差扩散, 同质转码

中图分类号 TN919.81; TP37

Video Transcoding Architecture and Recent Advances

SHANG Shu-Lin^{1,2} DU Qing-Xiu¹ LU Han-Qing² TANG Xiao-Jun¹

Abstract The development and recent results of video transcoding are reviewed in this article. Then, the latest research is introduced on transcoding between H.264 and others video coding standards due to the emergency of the latest international video coding standard H.264. Finally, the problems faced by transcoding research are described.

Key words Transcoding, H.264, bit rate, drift error, homogeneous transcoding

1 引言

视频技术和视频压缩标准的不断发展以及视频节目在不同应用环境中的需要,促使了视频转码研究的出现和不断发展.目前,视频转码技术已被人们广泛研究.具体来说,什么是视频转码呢?顾名思义,视频转码就是从一种视频编码流到另一种编码流的转换^[1],目的是把输入的压缩流转换成符合接收端要求的压缩流.

视频压缩的标准有很多种,如 MPEG-1^[2]、MPEG-2^[3]、H.263^[4]、MPEG-4^[5]、H.264^[6] 等等.不同视频格式(或相同视频格式)之间转换的需要促使了视频转码研究的开展,而新标准的制定和应用又推动了新的转码技术的研究.首先出现的转码需求是码率变换转码.在数字电视广播领域和 IPTV 系统中的视频服务器端,为了使不同带宽的用户能够收看视频点播,视频服务器端需要利用转码器将高码率视频转换为低码率视频;然后,在低分辨率显示和处理器性能较弱的移动设备(如 3G 手机、移动 MP4)的应用中,需要考虑降低空间分辨率和

时间分辨率的转码;近年来新制定的 H.264 编码标准采用多项先进编码技术,进一步提高了压缩效率,在数字电视、IPTV、移动视频等领域都将得到广泛的应用.因此,将 MPEG-2 的视频转为 H.264 格式成为了转码研究的重要方向.

最基本的转码结构是“全解全编”结构(如下页图 1 所示),直接由解码器和编码器串联而成.先对输入流进行完全解码,然后再按照新的码率要求进行编码,当然还可以根据实际要求进行空间分辨率或帧率调整.然而,“全解全编”转码的计算复杂度很高,适合一些离线应用场合.实用的转码结构一般是在上述结构的基础上采取了许多优化措施,譬如,在解码环路中利用输入码流中如视频序列头信息、宏块编码模式信息、运动矢量信息、频域的压缩系数等压缩信息.在保持视频质量下降不大的条件下,尽可能降低转码计算复杂度.

视频转码器的类型多种多样,按照结构主要分为两大类:像素域转码(Pixel-domain transcoder, PDT)^[7,8]和压缩域转码(Compressed-domain transcoder, CDT)^[7,9,10],CDT 也可以称为变换域转码. PDT 是指对输入流进行完全解码得到像素域图像,在像素域进行转码; CDT 则不需要完全解码,而是在变换域中进行转码^[9]. PDT 有着实现灵活的特点,可以满足多种不同的转码需要,输出图像质量较高,但在转码时需要做 IDCT 和 DCT 计算,复杂度较高; CDT 则是直接处理压缩域系数,减少重新编码的计算量,但其实现框架不够灵活,转码质量下降较大.

本文内容安排如下:第 2 节主要介绍了视频转码技术的研究现状,包括码率转换、空间分辨率转

收稿日期 2007-1-4 收修改稿日期 2007-6-4
Received January 4, 2007; in revised form June 4, 2007
国家高技术研究发展计划(863 计划)(2006AA01Z130)和国家科技支撑计划(200613AH02A13)资助
Supported by National High Technology Development Program of China (863 Program)(2006AA01Z130), and National Key Technologies R&D Program of China (200613AH02A13)
1. 中国科学院自动化研究所高技术创新中心 北京 100080 2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100080
1. Hi-Tech Innovation Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100080 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080
DOI: 10.1360/aas-007-1233

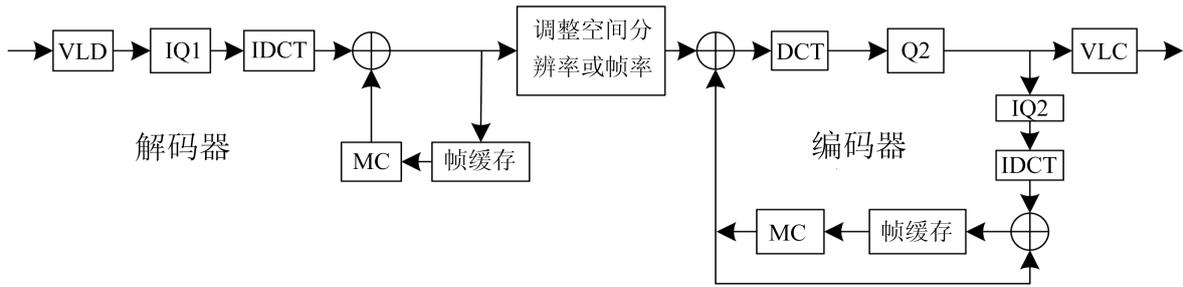


图 1 全解全编转码

Fig. 1 Cascade transcoder

码、时间分辨率转码、容错转码和语法转码的各种技术;第3节介绍了最新的转码研究进展,主要是关于新标准 H.264 的转码技术;第4节对转码做了简单的总结。

2 研究现状

视频转码主要有以下几种类型:码率转换 (Bit-rate reduction)、空间分辨率转换 (Spatial resolution reduction)、时间分辨率转换 (Temporal resolution reduction)、语法转换等^[8]。另外,随着网络技术和 3G 技术的发展,容错视频转码 (Error-resilience video transcoding) 获得了人们的广泛关注。

2.1 码率转换

码率转换主要应用于电视广播、因特网视频流传输、无线网络等传输领域。其目的是:有效地降低码率(节省带宽),使得网络、无线传输或数字设备存储应用中输出码率符合条件要求。降码率转码研究主要集中在两方面:降低复杂度和减少误差漂移 (Drift reduction)^[1]。

误差漂移问题可被理解为连续的预测帧图像的逐渐模糊^[1]。在视频编码过程中,使用参考帧对一个图像帧进行预测,仅所得残差被编码。只有当存储在解码器预测器中的重建参考帧与编码器预测器的重建参考帧一致时,解码器才能输出正常的编码图像。当编解码器预测器中的重建图像不相匹配时,由此导致的误差积累问题被称为误差漂移^[11]。误差漂移问题导致在一段时间内图像质量逐渐变坏,直到一个帧内编码帧的出现(帧内图像是独立的,不依赖于其他帧,不受误差漂移问题影响),图像质量重新变好。

作为最早的研究成果,文献 [7] 针对 MPEG-2 视频压缩流之间码率转换方面提出了四种方法。第一种是截断高频 DCT 分量(如图 2),通过丢弃部分高频 DCT 系数,使各块的比特数满足要求,其

缺点是丢弃高频系数会造成图像细节的损失,易在重建帧中产生块效应^[11]。文献 [12] 中有选择地丢弃 DCT 系数,提高了转码质量。第二种方法选取合适量化步长,对解码后频域系数进行再量化 (Requantization) 来降低码率(文献 [13] 也有相关研究),其转码框架大致如图 3(a) 所示。第三种方法利用提取的运动矢量和编码模式 (Coding decision modes) 对图像重新编码,避免重做运动估计和编码模式选择,而且重新运动补偿防止了误差积累。第四种方法在第三种方法的基础上重新选择最优的编码模式。

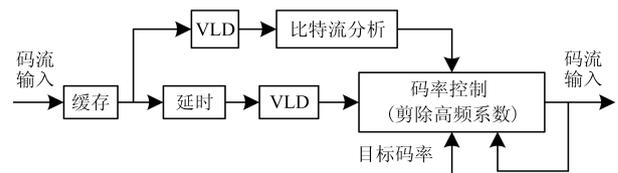


图 2 截断高频系数方法的结构图

Fig. 2 Architecture of cutting high frequencies

四种方法中,方法一、二属于开环系统,不需要进行 DCT/IDCT 且重建图像,这样在很大程度上降低了复杂度。然而,由于编解码器的预测器中的重建图像不相匹配导致了漂移误差的产生,在 GOP 内,图像误差逐渐积累。方法三、四属于闭环系统,虽然复杂度略高,但由于对解码图像进行了重建,反馈参考帧图像,重新计算了各宏块的残差,有效地去除了误差扩散,因此其转码视频质量较高。关于转码质量的比较,开环方法由于误差扩散所以质量较差,“全解全编”转码的质量最好,闭环方法的质量接近“全解全编”,但转码复杂度比较低。

除上述四种方法外,后面又有了许多发展。如图 3(a),就是根据数学推导对“全解全编”转码结构的一种简化^[14]。转码器从输入视频流中提取相应的编码信息,不需重新运动估计,而且 DCT、IDCT 只需

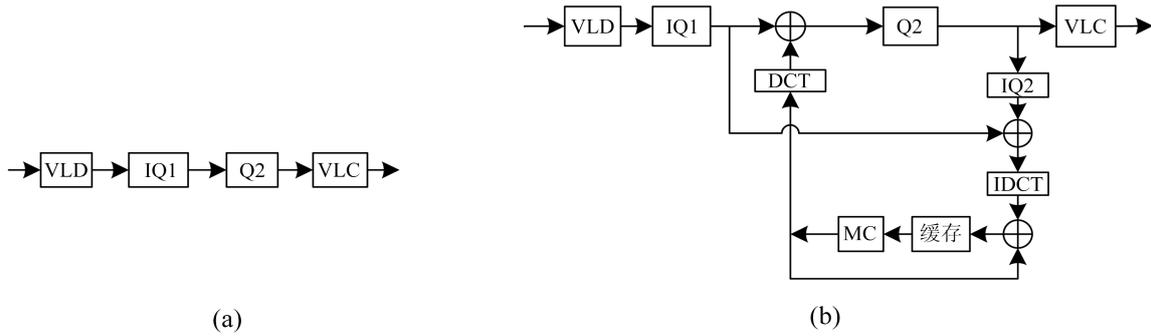


图 3 (a) 开环结构; (b) 对再量化数据进行漂移误差补偿的闭环结构
 Fig. 3 (a) Open loop; (b) Close loop compensating for the drift error due to re-quantization

进行一次, 大幅降低了复杂度, 同时相对于开环形式 (如图 3(a)), 取得不错的转码质量. 对于转码所需的运动信息, 相比较上面方法三、四中采用的简单的重用策略, 文献 [15, 16] 分析了调整量化步长导致运动矢量非最优, 采用运动矢量优化方法 (Motion vector refinement scheme, MVR) 提高运动估计精度. 为降低闭环形式的复杂度, 文献 [17] 分析了闭环结构中的量化误差积累补偿问题, 提出一种阈值自适应算法, 利用积累量化误差的大小决定是否误差补偿, 进一步降低了计算复杂度.

降低码率转码还可以在变换域着手^[9, 14, 18~21]. 变换域转码可以在图 3(b) 的基础上省略了 IDCT 和 DCT 计算, 进一步降低计算复杂度, 主要问题是研究在变换域上的运动估计, 进行漂移误差补偿. 但是需要注意的是, 在无误差扩散降帧率转码、空间分辨率转码和异质转码 (不同编码标准之间转码) 中, 变换域转码计算会十分繁琐.

还有从 CBR (Constant bit rate) 到 VBR (Variable bit rate) 的转码研究, CBR 编码中码率恒定而图像质量不恒定, VBR 中采用适当的码率分配使得码率不恒定而图像质量恒定. 文献 [22] 把 two-pass 码率控制算法^[23] 中 VBR 控制策略应用到转码中的码率控制方面, 给出了一个非常有效的方法.

2.2 空间分辨率转码

进行空间分辨率转码的动因是由于移动多媒体设备的显示分辨率较低, 人们需要寻找措施把现有的大量 MPEG-2 高质量视频降低空间分辨率以便能在这些设备上播放^[1]. 同时, 从高清电视到标清电视的转换也是原因之一. 与降低码率转码相比, 空间分辨率转码的“全解全编”结构中需要添加一个采样模块^[1]. 同样, 为提高转码视频质量和压缩效率, 并降低转码器复杂度, 人们进行了许多方面的研究. 在降低分辨率的研究中主要有以下几方面内容: 降低空间分辨率的下采样算法^[24, 25]、运动矢量的映射

算法和伸缩算法^[26, 27]、宏块编码模式决策^[28]、帧内刷新机制^[29, 30] 等等.

降低空间分辨率的下采样算法. 在降分辨率转码中, 需要缩小图像的空间尺寸. 文献 [25] 中提到三种方法: 1) 像素平均方法, 即用 2×2 像素的平均来表示一个像素; 2) 滤波加子采样方法, 其中采用了一七抽头滤波器对原图像滤波, 然后抽样^[25]; 3) 用截断原输入 DCT 系数的方法来获得低分辨率的图像, 此方法截断原来四个 8×8 块成四个 4×4 块, 分别对所得四个 4×4 块进行 IDCT 获得一个新的 8×8 块.

上面讨论的是当转码后图像大小为原图像的整数分之一时的情况. 当转码后图像尺寸不为原图像的整数分之一时, 文献 [24] 提出先对原图像进行补零插值, 然后用低通滤波器 $H(z)$ 去除插值带来的高频残差, 最后再进行下采样得到所需的图像, 流程图如图 4 所示, 这里 L 为上采样 (补零插值) 倍数, M 为下采样倍数.

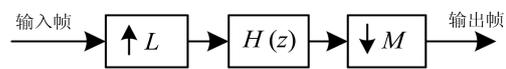


图 4 降分辨率流程图
 Fig. 4 Reducing resolution

运动矢量的映射算法和伸缩算法. 如图 5 所示, 空间分辨率降低后, 低分辨率图像的一个宏块会对应原来的多个宏块. 宏块运动矢量选择也存在多对一映射的问题, 采用一定的方法计算选择合适的运动矢量作为新宏块的运动矢量. 最后, 所得的运动矢量还需要除以分辨率压缩比, 使之与低分辨率图像相对应. 运动矢量映射算法的相关研究有很多^[10, 24, 27, 28, 31~33], 如文献 [27] 给出了三种方法: 平均值法、中值法和任意选取一个矢量作为当前宏块的运动矢量; 文献 [32] 利用图像活动度做加权平均来计算运动矢量; 文献 [24] 用各个宏块在降采样宏

块上对应的面积,对候选运动矢量做加权平均来计算运动矢量,这种方法的缺点是对脉冲噪声敏感^[28];为克服这个缺点,文献[10,28]提出使用加权中值滤波方法,来计算预测运动矢量;文献[33]则提出了一种 DC-max 的方法,DC-max 法利用原各宏块的 DC 系数作为活动度量,取最大 DC 系数宏块的运动矢量作为要求的运动矢量,此法计算复杂度低,所得 PSNR 值高于平均值法;文献[10]讨论了在输入图片格式为场同时存在跳帧的情况时,预测运动矢量的计算方法,最后,在上述方法所得运动矢量的附近进行 MVR,可得到更准确的运动矢量^[27].

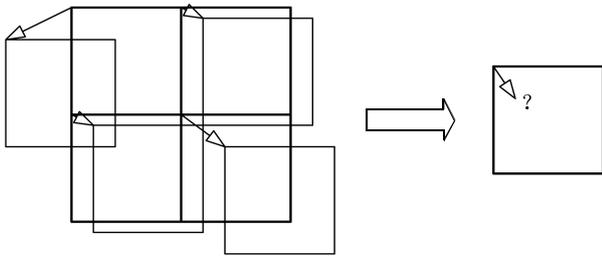


图 5 一个宏块对应原来的 4 个宏块 (2: 1 下采样)

Fig. 5 One MB corresponds to four MBs (2: 1 downscale)

宏块编码模式决策.在运动估计之后,编码器进行宏块编码模式选择.在转码中,通过利用输入流的模式决策信息来为宏块选择新的编码模式,节省了运动估计的部分计算量.降分辨率后的每个宏块对应着原来的多个宏块.这些宏块的类型并不一定一样.因此,需要根据候选宏块的类型为新的宏块选择合适的类型.文献[27]的处理方法是:

- 1) 如果四个宏块中至少有一个 INTRA 类型,则新宏块选择为 INTRA 类型;
- 2) 否则,如果一个宏块为 INTER 类型,则选择 INTER 类型;
- 3) 当所有宏块类型为 SKIP 类型时,选择 SKIP 类型;
- 4) 在编码器中重新估计宏块类型.

文献[28]则在宏块编码模式决策过程中先后进行了帧内-帧间决策和预测方向决策.对于帧内-帧间决策,采用了“少数服从多数”的原则,选取占多

数的宏块模式为当前宏块的模式.当选择帧间编码模式编码 B 帧时,面临着前向、后向、双向三种预测方向,同样采取上述原则来选择.

帧内刷新机制.降低空间分辨率开环转码结构有着复杂度低的优点,但其中的漂移误差现象会使输出视频图像质量变坏.文献[29,32]提出在开环结构中采用帧内刷新机制,即在 P 帧中提供一定比例的帧内宏块编码,由于误差漂移只能发生在帧间编码模式中,所以增加帧内宏块可以在一定程度上阻止误差的积累扩散.

2.3 时间分辨率转码

时间分辨率转码,即降低视频序列帧率的转码.降低帧率的同时可以降低码率,降低对解码设备处理能力的需求,适合应用于网络带宽较窄、用户端设备处理能力弱、显示设备分辨率低等场合.在转码过程中,降低帧率意味着需要提出合适的丢帧策略、必要时需要进行帧类型转换、需要利用丢弃帧的运动信息来重新合成运动矢量.因此,时间分辨率的转码的主要研究内容包括:丢帧策略^[34]、帧类型转换^[10,31]、场帧转换^[31]、运动矢量合成和优化算法^[10,15,16].

丢帧策略.考虑到 B 帧对 I、P 帧没有影响,文献[31]提出首先考虑丢弃 B 帧.帧率太低,往往会导致视频图像视觉上的不连续,为使转码后的视频序列在输出码率符合带宽要求的同时图像中的运动比较平滑,需要采用合适的跳帧策略.文献[34]根据运动矢量的累积幅度来决定是否跳帧.

帧类型的转换.文献[10]给出了 MPEG-2 到 MPEG-4 SP 转码时,由于 MPEG-4 SP 不支持 B 帧,需把原来的 B 帧转码成 P 帧的方法,即把原来存在的前向、后向、双向运动矢量按照依赖关系全部转换成前向运动矢量.图 6 是文献[10]中 MPEG-2MP@ML、30 fps、GOP(3,16)到 MPEG-4 SP 转码时,丢帧及更新参考帧和运动矢量的示意图.

场帧转换.转码中有时会遇到场编码方式转换为帧编码方式的情况,文献[31]针对 MPEG-2 到 H.263 转码,给出了“I 场~P 帧”和“P 场~P 帧”转换时运动矢量的调整算法.

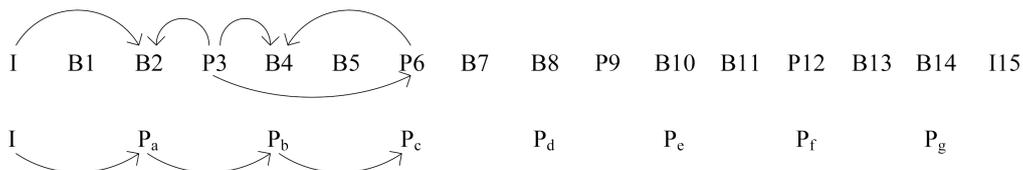


图 6 降低帧率转码时的运动矢量更新实例

Fig.6 Update motion vector after reducing temporal resolution

运动矢量合成算法 (Motion vector composition, MVC). 视频各帧之间由于运动矢量带来的依赖关系, 在丢帧之后产生中断. 利用被丢掉参考帧中的运动信息, 采用运动矢量合成算法, 把运动矢量进行累加以产生指向新的参考帧的运动矢量^[34, 25]. 具体算法有双线性插值方法^[34], 前向主矢量选择法 (Forward dominant vector selection, FDVS)^[15](如图 7), 伸缩矢量合成法 (Telescopic vector composition, TVC)^[25], 宏块活动度主运动矢量选择法 (Activity-dominant vector selection, ADVS)^[35] 等.

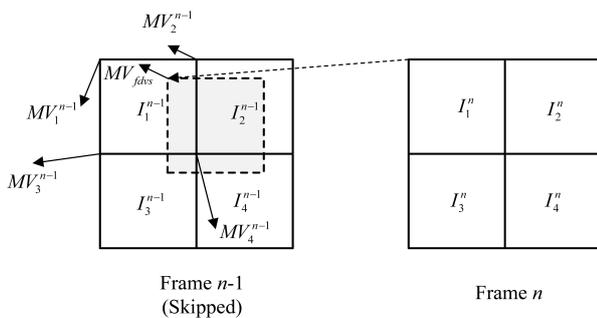


图 7 前向主矢量选择法

Fig. 7 Forward dominant vector selection

当前宏块在被丢掉参考帧中最多与四个彼此相邻的宏块相重叠. 双线性插值法对这四个宏块的运动矢量进行线性插值, 求得重叠区域在前面未被丢弃的参考帧中对应的运动矢量, 通过叠加求得当前宏块指向新参考帧的运动矢量. FDVS 方法则在这四个宏块中, 选取重叠面积最大的宏块为主宏块, 取主宏块的运动矢量为所需的中间运动矢量. 如图 7, 选择面积最大的宏块 I_2^{n-1} 为主宏块, 取宏块 I_2^{n-1} 的运动矢量 MV_2^{n-1} 为中间运动矢量 MV_{fdvs} . TVC 与 FDVS 十分相似, 但复杂度较小, 仅仅是把当前宏块在被丢掉帧对应宏块的运动矢量逐个累加在一起, 其效果只比 FDVS 稍微差一点, 甚至相同. ADVS 则是对 FDVS 的一种改进, 利用宏块活动度来选择主宏块和运动矢量. 在 MVC 之后, 利用得到的运动矢量作为参考运动矢量, 进行 MVR, 达到近似最优匹配, 从而提高运动估计的精度^[25].

2.4 容错转码

相对于有线信道, 无线信道有着信道窄、误码率高的特点. 因此, 在无线信道上传输视频, 需要加一些容错措施. 相对于简单使用在信道编码中的前向纠错 (Forward error correction)^[36, 37] 和自动重发请求 (Automatic error request) 等容错措施^[36], 文献 [38, 39] 则是从信源编码方面采取相应的容错

措施. 具体措施有: 加快空间域上的再同步, 减少每个 Slice 所包含的宏块个数以增加空间域容错; 增加每帧帧内宏块的比例以减少运动估计导致的错误扩散 (Error propagation); 帧内宏块自适应刷新机制 (Adaptive intra refresh); 发送反馈控制信号 (Feedback control signaling)^[40] 等技术.

3 视频转码研究的最新发展

近两年, 国际视频标准协会 ITU-T 制定了 H.264 标准, 编码效率比 MPEG-2 高出许多, 在相同画面质量下, 码率只有 MPEG-2 的 1/3~1/2. H.264 的制定带领着新一轮视频转码技术的发展, 下面首先介绍 H.264 的特点, 然后详细分析了 H.264 变换域转码技术, 最后分析总结了 H.264 模式选择的快速算法.

3.1 H.264 的特点

H.264 采用了许多新技术提高编码效率: MPEG-2 最多支持 2 个参考帧 (在 B 帧时), H.264 可以支持多达 16 个参考帧; MPEG-2 中只支持 DC 系数的 DPCM 预测, 而 H.264 中支持多种帧内宏块预测模式, 如 INTRA 4×4 支持九种预测模式, INTRA 16×16 支持四种预测模式; MPEG-2 只支持 16×16、16×8 (场模式) 块运动补偿模式, 而 H.264 支持七种块运动补偿模式 (16×16、16×8、8×16、8×8、8×4、4×8 和 4×4); MPEG-2 只支持 1/2 像素运动估计, 而 H.264 支持 1/4 像素运动估计. 正是因为 H.264 采用多参考帧、多编码模式以及 1/4 像素运动估计, 提高了运动估计的准确度, 所以才有很高的编码效率. 但是多种预测模式大幅度增加了模式选择的计算量, 因此, 在转码过程中, 需要研究模式选择的快速算法, 以提高计算速度. 另外, 相比较 MPEG-2、H.263 采用的 8×8 DCT 变换, H.264 采用的是 4×4 整数变换 (在增强的高清部分, 采用 8×8 整数变换); H.264 还支持容错编码, 而 MPEG-2 不支持.

上面提到的是应用最普遍、技术非常具有典型性的 MPEG-2 到 H.264 之间的转码. 实际上其他标准与 H.264 相比也有着相似的技术差距. 因此, 在传统标准到 H.264 的转码需要有效地研究利用转码前后两标准的相似之处来降低转码的复杂度, 研究各种视频转码架构和技术以在加快转码速度的同时平衡所得的视频质量.

大体说来, 需要进行如文献 [41, 42] 所述几部分内容的研究: 变换域编码; 帧内图像预测; 转码中的多模式选择问题和运动估计等.

下面介绍一些已有的技术研究成果.

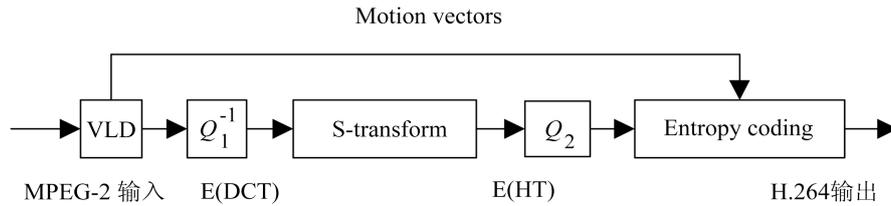


图 8 MPEG-2 到 H.264 转码的开环结构

Fig. 8 MPEG-2 to H.264 open-loop architecture

3.2 H.264 变换域转码

直接变换域转码可以省略 IDCT 和 HT 计算, 有利于降低计算复杂度, 因此有必要对其进行研究.

DCT-HT 变换. 变换域转码首先要解决 8×8 DCT 变换与 H.264 的 4×4 整数变换 (HT) 之间的转换. 在文献 [40, 43] 中都提出了 DCT-HT 变换算法, 基本思路是将 IDCT 和 HT 这两个变换结合成为一个矩阵算子, 并使用快速算法降低计算量.

变换域帧内转码. 文献 [44] 给出了一种完整变换域帧内转码算法. 该算法可以与快速帧内模式选择算法结合进一步减小计算量. 其关键技术有两点: 1) 在变换域计算预测残差; 2) 在变换域计算帧内转码失真函数^[45]. 计算残差过程中先计算帧内预测信号的 HT 变换, 然后用源图像的 HT 系数减去预测信号的 HT 系数, 以得到预测残差的 HT 系数, 由于预测信号的 HT 变换计算量很小^[46], 该算法可以节约计算量.

变换域帧间转码. 文献 [46] 对变换域帧间转码做了详细的分析研究, 如图 8 所示用算子将 DCT 系数转换为 HT 系数即可形成开环转码结构, 然而开环结构中存在两个主要的漂移误差: 亚像素插值滤波器引起的误差和量化不同产生的误差. MPEG-2 采用的是 2 拍插值滤波器, 而 H.264 采用的是 6 拍插值滤波器, 这样会导致运动估计出现可扩散偏差; MPEG-2 与 H.264 量化不同也会产生扩散偏差. 该文献中推导了变换域偏差补偿的公式, 使编码质量几乎达到了像素域的效果, 并降低了计算复杂度.

需要注意的是在变换域进行模式和 MVR 的计算量反而比像素域算法的计算量大, 因此文献 [46] 只能直接采用原编码的模式和运动矢量, 没有选择最优模式和运动矢量, 达不到 H.264 应有的压缩效率, 所以 H.264 的变换域转码只适用于简单的实时转码.

3.3 H.264 快速模式选择算法

H.264 提供了大量的帧内和帧间模式可供选择, 可以用率失真优化^[47] 的算法来选择最优的模式 (如 H.264 的 JM 参考软件中的模式选择算法), 但是遍

历式的率失真函数计算需要大量的计算量, 速度不能满足实际应用的要求, 因此有必要研究像素域转码中的快速模式选择算法以提高转码速度.

利用 MPEG-2 的模式和运动矢量信息. 在 DVD 等介质上的许多 MPEG-2 视频流是采用模式决策 (Mode decision, MD) 方法来决定帧的类型, 因此, 在做 MPEG-2 到 H.264 转码时可以重新利用这些帧类型信息, 根据 MPEG-2 编码的宏块帧内帧间类型来确定 H.264 转码中采用帧内还是帧间方式, 从而节省帧内/帧间判别的计算量. 在文献 [48] 中利用 MPEG-2 中的模式信息进行宏块帧内帧间的判定, 利用 MPEG-2 中的运动矢量作为帧间宏块的运动矢量, 使转码速度提高了很多, 而且 PSNR 下降不大. 当然帧内或帧间的模式选择还需要其他方法来确定, 而且 MPEG-2 的运动矢量也需要优化才能达到最优运动矢量.

利用 MPEG-2 宏块的 DCT 系数选择 H.264 帧内编码模式. MPEG-2 采用的帧内编码模式是简单的频域 DPCM, 而不同于 H.264 的空间域帧内编码模式. 因此在做 MPEG-2 到 H.264 帧内转码时很难直接利用原编码信息进行帧内模式选择. 文献 [49] 则利用 MPEG-2 帧内宏块的 DCT 系数选择 H.264 帧内编码模式, 取得较好效果. 其基本思想是: 如果宏块内的 4 个块的 DCT 系数差异较小, 则说明该宏块比较平滑、细节较少, 可以采用 INTRA 16×16 模式; 反之说明宏块不平滑、细节较多, 应当采用 INTRA 4×4 模式. 另外该文献中使用 SAD (绝对误差和) 来确定 INTRA 16×16 的 4 种模式选择, 取得比较好的效果. 然而文中进一步利用 DCT 系数确定 INTRA 4×4 的 9 种模式选择时, 效果并不理想, INTRA 4×4 的模式选择采用其他方法来确定会更好.

使用提前终止机制. 提前终止机制的思路是将各种模式出现概率从大到小排列, 逐个模式进行选择判断, 当某个模式的编码代价已经足够小时, 提前终止模式选择, 不再考虑其他模式. 文献 [48, 50] 中采用此机制降低模式选择的计算量. 其中文献 [50] 是利用 DIRECT 模式的编码代价作为依据进

行模式选择, 如果 DIRECT 模式的编码代价小, 则说明该宏块适合用大块的帧间模式, 否则需要考虑小块的帧间模式或者帧内模式, 其实验显示该算法在 PSNR 值降低不多的前提下, 大幅提高了转码速度. 文献 [48] 则是假设编码代价随宏块的大小而单调变化, 先尝试大块宏块模式, 如果较小宏块模式对应的编码代价大于较大模式的编码代价, 则选择较大宏块对应的编码模式, 提前终止模式选择. 另外, 相关实验表明 B 帧 DIRECT 模式 (或者 P 帧 SKIP 模式) 和 16×16 模式占了所有使用模式的很大部分, 如果能够提前判断此模式, 则其他模式就不用考虑了, 从而提高了转码速度^[50].

减少模式类别. 排除一些不常用的编码模式, 可以降低模式选择的复杂性. 文献 [50] 提出了: 在色度帧内预测时, 只考虑 DC 模式; 在 INTRA 16×16 预测中去掉 plane 模式. 此方法以损失很小视频质量的代价换得转码速度的提高.

3.4 H.264 转码快速运动估计技术

H.264 有多种帧间模式可供选择, 需要计算各种模式的运动矢量, 用全搜索方法选择运动矢量计算量显然是非常大的. 在 H.264 转码中, 需要对其中的快速运动估计算法进行研究. 一般转码快速估计算法包括三个步骤: 首先, 挑选一些已知的预测运动矢量 $PMV\{\overline{MV}_1, \overline{MV}_2, \dots, \overline{MV}_n\}$, 在其中选择一最优矢量作为搜索中心的运动矢量; 然后, 在得到的搜索中心周围小邻域内做 MVR, 以提高运动补偿的效果; 并使用提前终止准则来加快估计速度, 当预测精度达到某个阈值时就停止运动矢量搜索.

预测运动矢量是对最优运动矢量的初始估计, 在转码中可供选择的 PMV 有以下三类: 1) 输入视频流中的运动矢量, 该运动矢量是编码原视频流时得到的最佳运动矢量, 因此该矢量与最优运动矢量的相关性最大; 2) 空间邻域块 (包括左、上、右上、左上邻域块) 的运动矢量, 由于运动矢量的空间相关性, 所以相邻块的运动矢量一般比较接近; 3) 已计算的其他模式的运动矢量, 对于同一个宏块, 不同编码模式对应的运动矢量一般比较接近. 文献 [51] 在做 MPEG-4 到 H.264 转码时, 直接将在这三类运动矢量中选择的最优矢量作为 H.264 的最优运动矢量, 在不做 MVR (考虑到 MPEG-4 的运动矢量已经比较精确) 的情况下就能取得比较好的效果. 文献 [48] 在做 MPEG-2 到 H.264 转码时, 先在 MPEG-2 运动矢量和空间邻域运动矢量中选择最优矢量作为搜索中心, 并在搜索中心的小邻域内做 MVR, 将得到的运动矢量作为 H.264 的 16×16 模式运动矢量, 这样首先得到了大尺寸块模式的运动矢量, 可以作为小尺寸块模式的 PMV, 用以确定小块模式的运动

矢量. 编码中的快速估计算法 EPZS 在文献 [50] 中得到了新的应用, 并将输入视频流的运动矢量作为 EPZS 的预测运动矢量之一.

4 结束语

最新的视频编码标准 H.264 的颁布, 影响了新一轮的视频转码研究, 已有的 MPEG-2 素材到 H.264 的转码研究和 H.264 不同应用之间的转码研究正在如火如荼地进行. 同时, AVS^[52] 和 VC-1^[53] 的出现为转码研究添加了新的内容^[54]. 虽然, 在过去的时间内, 针对不同的标准、不同的应用场合, 出现了许多实用的转码技术. 但同时, 从上述文献分析中可以看出, 视频转码方面的研究并不是已经臻于至善了. 在已有的多种转码结构中, 每一种结构都有自己的局限性. 例如, 变换域上做运动矢量估计和模式选择, 计算量会非常大, 所以在降低空间分辨率转码、时间分辨率转码和 H.264 转码中, 变换域的处理很受限制. 许多研究工作的开展或者是基于前者或者是基于后者进行的. 当然, 并没有简单的“哪种方法更好”的结论, 不同的应用场合, 会有不同的合适选择.

视频转码方面的研究仍然有许多工作需要完成, 譬如, 视频传输中容错转码方面、H.264 转码的降分辨率、降帧率等方面的研究, 需要人们来细化、深入, 就像文献 [1] 中提到的那样: “展望视频转码的未来, 仍有许多方面的问题需要进一步的研究. 其中问题之一是怎样找到一种最优的转码策略”.

References

- 1 Anthony V, Charilaos C, Sun H F. Video transcoding architectures and techniques: an overview. *IEEE Transactions on Signal Processing Magazine*, 2003, **20**(2): 18~29
- 2 Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s, Part 2: Video, ISO/IEC 11172-2, 1991
- 3 Information Technology-Generic Coding of Moving Pictures and Associated Audio Information: Video, ISO/IEC 13818-2: 2000, 2000
- 4 Information Technology-Video Coding for Low Bit-rate Communications, ITU-T Recommendation H.263, 1998
- 5 Information Technology-Coding of Audio-visual Objects, Part 2: Visual, ISO/IEC 14496-2: 2001, 2001
- 6 Information Technology-Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264, 2005
- 7 Sun H F, Kwok W, Zdepski J W. Architectures for MPEG compressed bitstream scaling. *IEEE Transactions on Circuits and Systems for Video Technology*, 1996, **6**(2): 191~199
- 8 Chen Gong, Huang Xiang-Lin, Shen Lan-Sun. Video transcoding techniques. *Measurement and Control Technology*, 2003, **22**(5): 36~39

- (陈功, 黄祥林, 沈兰荪. 视频转码技术. 测控技术, 2003, **22**(5): 36~39)
- 9 Assuncao P A A, Ghanbari M. A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit streams. *IEEE Transactions on Circuits and Systems for Video Technology*, 1998, **8**(8): 953~967
 - 10 Xin J, Sun M T, Chun K. Motion reestimation for MPEG-2 to MPEG-4 simple profile transcoding. In: International Packet Video Workshop. Pittsburgh: 2002
 - 11 Xin J, Lin C W, Sun M T. Digital video transcoding. *Proceedings of the IEEE*, **93**(1): 84~97
 - 12 Eleftheriadis A, Anastassiou D. Constrained and general dynamic rate shaping of compressed digital video. In: Proceedings of International Conference on Image Processing. Washington D. C., USA: 1995. 3396
 - 13 Nakajima Y, Hori H, Kanoh T. Rate conversion of MPEG coded video by re-quantization process. In: Proceedings of International Conference on Image Processing. Washington D. C., USA: 1995. 3408
 - 14 Assuncao P A A, Ghanbari M. Post-processing of MPEG-2 coded video for transmission at lower bit rates. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. Atlanta, USA: 1996. 1998~2001
 - 15 Jeongnam Y, Sun M T, Lin C W. Motion vector refinement for high-performance transcoding. *IEEE Transactions on Multimedia*, 1999, **1**(1): 30~40
 - 16 Jeongnam Y, Sun M T, Lin C W. Motion estimation for high performance transcoding. *IEEE Transactions on Consumer Electronics*, 1998, **44**(3): 649~658
 - 17 Yuan L J, Wu F, Chen Q, Li S P, Gao W. The fast close-loop video transcoder with limited drifting error. In: Proceedings of the 2004 International Symposium on Circuits and Systems. Beijing, China: IEEE, 2004. 769~772
 - 18 Wang J S, Yu S L. Dynamic rate scaling of coded digital video for ivod applications. *IEEE Transactions on Consumer Electronics*, 1998, **44**(3): 743~749
 - 19 Chang S F, Messerschmidt D G. Manipulation and compositing of MC-DCT compressed video. *IEEE Journal on Selected Areas in Communications*, 1995, **13**(1): 1~11
 - 20 Sun H, Vetro A, Bao J, Poon T. A new approach for memory efficient ATV decoding. *IEEE Transactions on Consumer Electronics*, 1997, **43**(8): 517~525
 - 21 Safranek R J, Kalmanek C R J, Garq R. Methods for matching compressed video to ATM networks. In: Proceedings of IEEE International Conference on Image Processing. Washington D. C., USA: IEEE, 1995. 13~16
 - 22 Yuan L J, Sun H F, Gao W. MPEG transcoding for DVD recording. In: Proceedings of the 4th International Conference on Information Communication and Signal Processing and the 4th IEEE Pacific-Rim Conference on Multimedia. Singapore, Singapore: 2003. 545~548
 - 23 Westerink P H, Rajagopalan R, Gonzales C A. Two-pass MPEG2 variable-bit-rate encoding. *IBM Journal of Research and Development*, 1999, **43**(4): 471~488
 - 24 Shen G B, Zeng B, Zhang Y Q, Liou M L. Transcoder with arbitrarily resizing capability. In: Proceedings of IEEE International Symposium on Circuits and Systems. Sydney, Australia: IEEE, 2001. 25~28
 - 25 Shanableh T, Ghanbari M. Heterogeneous video transcoding to lower spatio-temporal resolutions and different encoding formats. *IEEE Transactions on Multimedia*, 2000, **2**(2): 101~110
 - 26 Xin J, Sun M T, Chun K, Choi B S. Motion re-estimation for HDTV to SDTV transcoding. In: Proceedings of IEEE International Symposium on Circuits and Systems. USA: IEEE, 2002. 715~718
 - 27 Bjork N, Christopoulos C. Transcoder architectures for video coding. *IEEE Transactions on Consumer Electronics*, 1998, **44**(1): 88~89
 - 28 Xin J, Sun M T, Choi B S, Chun K W. An HDTV-to-SDTV spatial transcoder. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, **12**(11): 998~1008
 - 29 Yin P, Vetro A, Liu B D, Sun H F. Drift compensation for reduced spatial resolution transcoding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, **12**(11): 1009~1020
 - 30 Vetro A, Yin P, Liu B D, Sun H F. Reduced spatio-temporal transcoding using an intra refresh technique. In: Proceedings of IEEE International Symposium Circuits and Systems. USA: IEEE, 2002. 723~726
 - 31 Wee S J, Apostolopoulos J G, Feamster N. Field-to-frame transcoding with spatial and temporal downsampling. In: Proceedings of IEEE International Conference on Image Processing. Kobe, Japan: IEEE, 1999. 271~275
 - 32 Shen B, Sethi I K, Vasudev B. Adaptive motion-vector resampling for compressed video downscaling. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, **9**(6): 929~936
 - 33 Chen M J, Chu M C, Lo S Y. Motion vector composition algorithm for spatial scalability uncompressed video. *IEEE Transactions on Consumer Electronics*, 2001, **47**(3): 319~325
 - 34 Hwang J N, Wu T D, Lin C W. Dynamic frame-skipping in video transcoding. In: Proceedings of the 2nd Workshop on Multimedia Signal Processing. Redondo Beach, USA: IEEE, 1998. 616~621
 - 35 Chen M J, Chu M C, Pan C W. Efficient motion-estimation algorithm for reduced frame-rate video transcoder. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, **12**(4): 269~275
 - 36 Ortega A, Khansari M. Rate control for video coding over variable bit rate channels with applications to wireless transmission. In: Proceedings of IEEE International Conference on Image Processing. Washington D. C., USA: IEEE, 1995. 3388
 - 37 Wenger S, Knorr G D, Ott J, Kossentini F. Error resilience support in H.263+. *IEEE Transactions on Circuits and Systems for Video Technology*, 1998, **8**(7): 867~877
 - 38 Reyes G de los, Reibman A R, Chang S F, Chuang J C I. Error-resilient transcoding for video over wireless channels. *IEEE Journal on Selected Areas in Communications*, 2000, **18**(6): 1063~1074
 - 39 Dogan S, Cellatoglu A, Uyguroglu M, Sadka A H, Kondoz A M. Error-resilient video transcoding for robust internet-work communications using GPRS. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, **12**(6): 453~464

- 40 Xin J, Anthony V, Sun H F. Converting DCT coefficients to H.264/AVC transform coefficients. *Lecture Notes in Computer Science*, 2004, **3332**: 939~946
- 41 Ahmad I, Wei X H, Sun Y, Zhang Y Q. Video transcoding: an overview of various techniques and research issues. *IEEE Transactions on Multimedia*, 2005, **7**(5): 793~804
- 42 Kalva H. Issues in H.264/MPEG-2 video transcoding. In: Proceedings of Consumer Communications and Networking Conference. Las Vegas, USA: IEEE, 2004. 657~659
- 43 Shen B. From 8-tap DCT to 4-tap integer-transform for MPEG to H.264/AVC transcoding. In: Proceedings of International Conference on Image Processing. Palo Alto, CA: 2004. 115~118
- 44 Su Y P, Xin J, Anthony V, Sun H F. Efficient MPEG-2 to H.264/AVC intra transcoding in transform domain. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS). USA: IEEE, 2005. 1234~1237
- 45 Xin J, Anthony V, Sun H F. Efficient macroblock coding-mode decision for H.264/AVC video coding. In: Proceedings of the 24th Picture Coding Symposium (PCS). San Francisco, USA: 2004. 53~58
- 46 Qian T J, Sun J, Li D, Yang X K, Wang J. Transform domain transcoding from MPEG-2 to H.264 with interpolation drift-error compensation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, **16**(4): 523~534
- 47 Wiegand T, Schwarz H, Joch A, Kossentini F, Sullivan G J. Rate-constrained coder control and comparison of video coding standards. *IEEE Transactions on Circuits and Systems for Video Technology*, 2003, **13**(7): 688~703
- 48 Zhou Z, Sun S J, Lei S, Sun M T. Motion information and coding mode reuse for MPEG-2 to H.264 transcoding. In: Proceedings of IEEE International Symposium on Circuits and Systems. Seattle, USA: IEEE, 2005. 1230~1233
- 49 Wang L, Wang Q, Liu Y, Lu W. A fast intra mode decision algorithm for MPEG-2 to H.264 video transcoding. In: Proceedings of IEEE the 10th International Symposium on Consumer Electronics. Tianjin, China: IEEE, 2006. 1~5
- 50 Lu X A, Tourapis A M, Yin P, Boyce J. Fast mode decision and motion estimation for H.264 with a focus on MPEG-2/H.264 transcoding. In: Proceedings of IEEE International Symposium on Circuits and Systems. New York, USA: IEEE, 2005. 1246~1249
- 51 Kim S E, Han J K, Kim J G. Efficient motion estimation algorithm for MPEG-4 to H.264 transcoder. In: Proceedings of International Conference on Image Processing. Seoul, South Korea: IEEE, 2003. 656~659
- 52 Information Technology—Advanced Audio and Video Coding Standard: Video, GB/T 20090. 2-2006 (信息技术先进音频编码第二部分: 视频, GB/T 20090. 2-2006)
- 53 Standard for Television: VC-1 Compressed Video Bitstream Format and Decoding Process, SMPTE 421M-2006

- 54 Shen G B, He Y W, Cao W Y, Li S P. MPEG-2 to WMV transcoder with adaptive error compensation and dynamic switches. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, **16**(12): 1460~1476



尚书林 中国科学院自动化研究所博士研究生. 2002 年获得北京师范大学信息学院电子系学士学位, 主要研究方向为图像与视频压缩技术. 本文通信作者. E-mail: shangsl@gmail.com

(**SHANG Shu-Lin** Ph. D. candidate at Institute of Automation, Chinese Academy of Sciences. He received his bachelor's degree from Beijing Normal University in 2002. His research interest covers image compression and video coding. Corresponding author of this paper.)



杜清秀 中国科学院自动化研究所副研究员. 主要研究方向是人机交互、虚拟现实、多媒体技术.

E-mail: qingxiu.du@ia.ac.cn

(**DU Qing-Xiu** Associate professor at Hi-Tech Innovation Center, Institute of Automation, Chinese Academy of Sciences. Her research interest covers human computer interaction, virtual reality, and multimedia technology.)



卢汉清 中国科学院自动化研究所研究员. 主要研究方向是图像视频分析、多媒体系统. E-mail: luhq@nlpr.ia.ac.cn

(**LU Han-Qing** Professor at National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. His research interest covers image and video analysis, and multimedia system.)



唐小军 中国科学院自动化研究所硕士研究生. 2001 年获得北方工业大学自动化系学士学位, 主要研究方向是图像和视频压缩技术.

E-mail: xjtang@hitic.ia.ac.cn

(**TANG Xiao-Jun** Master student at Institute of Automation, Chinese Academy of Sciences. He received his bachelor's degree from North China Technology University in 2001. His research interest covers video compression and video coding.)