

# 基于交叠正交变换的视频压缩编码分析与实现

肖平平 余松煜 宋利 熊红凯

(上海交通大学 图像通信与信息处理研究所, 上海 200030)

**摘要** 混合视频编码标准由于采用基于块的运动补偿预测和变换量化,从而使解码图像在块边界产生干扰视觉的不连续性,尤其在低码率下,易形成块效应。由于滤波法、凸集投影法、交叠正交变换等可作为信号处理中去除块效应的技术方案,为此在对视频编码的系统结构进行分析的基础上,实现了基于交叠正交变换时间域预滤波和后置滤波的去块效应视频混合编码方案,并讨论了实现这一方案的框架结构和交叠模式选择等关键技术问题,最后给出了实验结果,对比实验结果显示,该方案在提供相当的编码性能基础上,能有效地改善视频图像的块效应现象。

**关键词** 块效应 交叠正交变换 离散余弦变换 时域预滤波和后置滤波

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2005)11-1345-05

## Analysis and Implementation of Video Coding with Lapped Orthogonal Transform

XIAO Ping-ping, YU Song-yu, SONG Li, XIONG Hong-kai

(Institute of Image Communication and Information Processing, Shanghai Jiaotong University, Shanghai 200030)

**Abstract** The visual disturbing discontinuities at the block boundaries, which are called block artifacts, might arise in the reconstructed pictures, especially in low bit-rate applications, because the block-based motion compensation prediction, transform and quantization are adopted in the hybrid video coding standards. A deblocking filter, projections onto convex sets, and lapped orthogonal transform behave as the candidate technologies for reducing the block effects in signal processing. This paper analyzes and integrates the systematic architecture of video coding, achieves a deblocking hybrid video coding framework on the basis of a lapped orthogonal transform with time-domain pre- and post-filtering, and discusses the key technology issues including overlap mode choice within the framework. The experimental outcomes are shown to improve the subjective quality of video pictures with a comparative rate-distortion performance, compared to the traditional framework.

**Keywords** block artifact, lapped orthogonal transform, discrete cosine transform, time-domain pre- and post filtering

## 1 引言

目前多数混合视频压缩编码标准,如 MPEG-2、H. 263、JPEG-2000、MPEG-4、H. 264/AVC 都基本采用基于块离散余弦变换(discrete cosine transform, DCT)的编码框架,但在低码率情况下,块效应则成为比较突出的问题。为了克服分块 DCT 所带来的块效应问题,包括滤波法、凸集投影法(projections

onto convex sets, POCS)、交叠变换(lapped transform, LT)等在内的用于消除块效应的技术被提出。其中滤波法作为传统消除块效应的方法,可分为环路滤波和后置滤波,其不仅计算简单,不增加码流,而且对边缘平滑有一定的效果,但由于低通滤波器的低通特性,易损失块边界的高频信息,因此常造成块边界模糊<sup>[1]</sup>;POCS 作为信号重建中的重要理论方法,它是将解码后的块往一系列平滑约束集上进行连续投影变换,其得到的重构图像能有效消除块效应,优

基金项目:国家自然科学基金项目(60502033);上海市自然科学基金项目(04ZR14082)

收稿日期:2005-08-20;改回日期:2005-09-07

第一作者简介:肖平平(1978~),女。1999年毕业于解放军信息工程学院电子工程系,获监测工程专业学士学位,现在上海交通大学图像通信与信息处理研究所攻读硕士学位。研究方向为图像通信与多媒体技术、图像压缩编码等。E-mail: pingxiao@sjtu.edu.cn

点是在消除块效应的同时,能在一定程度上保持图像的细节,但由于它是一个迭代算法,算法复杂度较高,而且由于受平滑限制条件制约很难保证不损伤原有的频率分量<sup>[2]</sup>。

LT 是一种可以有效消除块效应的变换方法<sup>[3]</sup>,其在以往的应用中多用于信号处理、静止图像压缩编码领域<sup>[4]</sup>。本文通过分析由时间域预滤波和后置滤波实现的交叠变换 (time-domain lapped transform, TDLT)<sup>[4]</sup>,结合视频编码的系统结构,实现了基于 TDLT 的去块效应视频混合编码方案,并讨论了实现这一方案的框架结构和交叠模式选择等关键技术问题,通过对比实验,在提供相当的编码性能基础上,有效地改善视频图像的块效应现象。

## 2 混合视频压缩编码技术及块效应的产生

运动补偿预测、变换量化、熵编码技术是构成混合视频编码框架的主要环节。由于受视频序列中物体变化速度的限制,致使相邻帧之间存在很高的相关性,即时间冗余,而运动补偿预测的目的就是为了消除视频序列图像间的时间相关性;同时由于视频图像内相邻像素之间存在空间相关性,因此变换的目的在于集中图像内的能量分布,以有效减少空间冗余度,而量化则是实现信号有损压缩的精简表示;熵编码的目的是为了消除输出量化符号序列的统计冗余度。正是为了提供实时视频编解码器的可实现性和运算复杂度,并获得匹配于图像帧的时域模型,因此基于块的运动补偿预测 (16 × 16、16 × 8、8 × 16、8 × 8、8 × 4、4 × 8、4 × 4) 以及变换编码 (8 × 8、4 × 4) 就成为实现变换的基本单元。图 1 实线所示为传统基于 DCT 的视频编码框架。

必须明确的是,基于块的混合视频编码框架存在以下两个主要问题:一是由于各块进行独立编码,没有考虑块与块之间相关性,因而降低了编码效率;二是块效应现象。之所以后者作为当前混合视频压缩编码的主要限制之一,其主要来源于以下两个环节:(1)由于利用帧内及帧间残差进行分块 DCT 时,纹理系数的非均匀分布会使相邻块间变换系数的重要性权重产生差异,尤其是在低码率下量化,会在相邻块边界造成干扰视觉的不连续性,形成“虚假边缘”;(2)由于按不同自适应块尺度进行适配的运动补偿块,可以通过将不同参考帧内不同位置块进

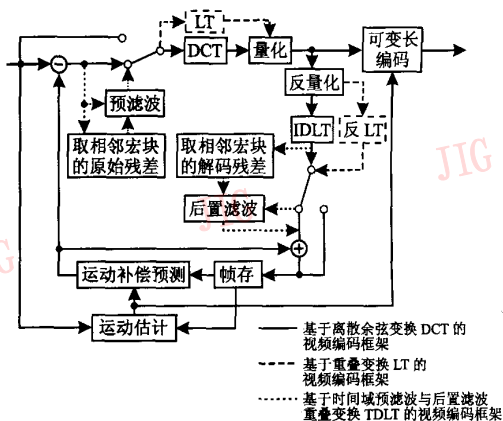


图 1 基于 3 种变换模式的混合视频编码框架  
Fig. 1 The hybrid video coding framework with three transform modes

行加权内插来获取,而且由于不存在对任何当前编码块边界、纹理的完美配准,因此参考帧的边界非连续性会因被引入当前编码块内部而产生新的“虚假边缘”。图 2 所示为量化系数较大或码率较低时,“Foreman”序列经混合视频编码框架 H. 263 编码后的一帧解码重建图像,从图中可以看到块效应明显,且主观视觉效果较差。



图 2 低码率下混合编码框架压缩后的解码图像  
Fig. 2 The reconstructed picture in hybrid video coding

## 3 基于传统交叠变换的去块效应视频编码框架

### 3.1 LT

LT 最早由 Malver 提出,图 3 为 LT 示意图。其与普通的块变换具有以下区别:普通的块变换是将输入信号分块后单独进行变换,且每块的变换矩阵

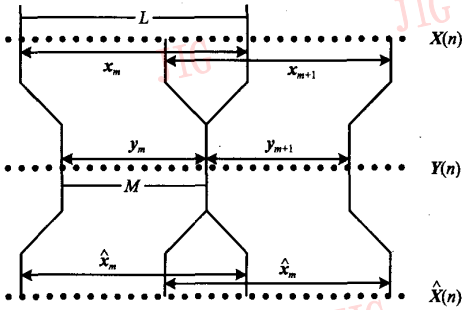


图 3 LT 示意图

Fig. 3 The lapped transform diagram

为方阵;LT 中,当信号分块之后,由于它是通过将各块信号与相邻块的部分信号做变换来得到该块信号的变换系数,因此其变换的输入值和输出值数目不等。由于块交叠变换的输入信号包括了相邻块中与该块相邻的边缘信息,且充分利用了块与块之间的相关性,因此可以有效地消除块效应,且变换本身不会增加码流。

基于 LT 的可分离性,可先考虑 1 维变换:设输入信号  $X$  分割成大小为  $M$  的块,并假设第  $m$  块为  $x_m$ ,其进行 LT 时,输入取样值数目为  $L(L > M)$ ,即包括当前  $x_m$  块分别与相邻  $x_{m-1}$ 、 $x_{m+1}$  块进行交叠的  $L-M$  个样值,一般选取  $L = KM$ , $K$  称为交叠因子;交叠变换矩阵  $H$  的维数为  $M \times L$ ,变换输出矢量  $y_m$  长度为  $M$ ,表示为: $y_m = Hx_m$ ,其中, $H$  中的  $M$  个行向量表征基函数的转置。在解码端重构时,反变换矩阵  $F$  的维数为  $L \times M$ ,重建信号表示为  $\hat{x}_m = FHx_m$ 。

为便于表述,设分别用  $K$  个  $M \times M$  子矩阵来表示  $H$  和  $F$ ,即: $H = [h_0, h_1, \dots, h_{K-1}]$  和  $F = [f_0, f_1, \dots, f_{K-1}]^T$ ,设  $l$  代表交叠点数,则只要  $H$  和  $F$  满足下式,就可在解码端完全重建信号。

$$\sum_{i=0}^{K-1-l} F_i^T H_{i+1} = \sum_{i=0}^{K-1-l} F_{i+1}^T H_i = \delta(l) I_M$$

### 3.2 基于传统交叠变换的视频编码框架

LT 可以通过 DCT 快速算法来实现,其传统的方法是先对信号进行 DCT 变换,再进行交叠,然后通过对 DCT 系数的后处理,即 DCT 域的交叠来实现的,如果将这种传统 LT 用于视频压缩编码中来消除块效应,则编码框架如图 1 中所示。

传统 LT 是介于 DCT 和量化之间的,如在 2 维信号中,若先按行进行 1 维交叠变换,再按列进行 1 维交叠变换,则其过程为:按行 1 维 DCT→交叠→按列 1 维 DCT→交叠,由此可见这种方法破坏了原有分块 2

维 DCT 框架的完整性,比较适合于一帧图像进行整体处理,而不便于对图像进行分块处理,由于现有的视频压缩标准中多数采用分块 2 维 DCT 的框架,因此这种通过 DCT 域交叠实现的传统 LT 不易于与现有视频压缩编码框架兼容,另外由于传统 LT 不能灵活地选择交叠率,即交叠点数,因而在实际应用中不便于在图像质量、编码效率和运算复杂度之间进行均衡。Trac 提出的 TDLT<sup>[4]</sup>,不仅能很好地保持原有框架的完整性,还易于与现有视频压缩标准兼容,而且可以选择交叠点数,以便均衡各种性能。

## 4 基于 TDLT 的去块效应视频混合编码方案

### 4.1 TDLT

TDLT 是与传统 LT 等效的一种交叠变换方法,它是通过对块边界进行时间域预滤波和后置滤波来实现的,其多相矩阵为

$$D_z = C_{M,II} \hat{\Lambda}_z P$$

其中,  $P = \frac{1}{2} \begin{bmatrix} I & J \\ J & -I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} I & J \\ J & -I \end{bmatrix}$ ,  $\hat{\Lambda}_z =$

$$\begin{bmatrix} I & 0 \\ 0 & zI \end{bmatrix} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = \begin{bmatrix} 0 & I \\ zI & 0 \end{bmatrix}$$
,  $C_{M,II}$  是  $M \times M$  大小的第 II 类 DCT 矩阵, $J$  为反对角矩阵, $\hat{\Lambda}_z P$  表示对相邻块边界的时间域进行预处理反变换的多相矩阵,即  $D_z$  的逆矩阵为  $R_z = D_z^{-1} = P^{-1} \hat{\Lambda}_z^{-1} C_{M,II}^T$ ,其中  $P^{-1} \hat{\Lambda}_z^{-1}$  表示对相邻块边界的时间域进行后处理,由此可得对块边界进行处理的过程(如图 4 所示)。

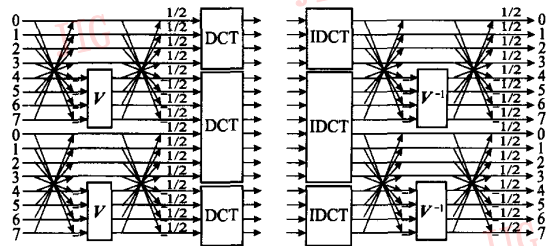


图 4 TDLT 边界处理示意图

Fig. 4 The diagram of TDLT

在 TDLT 中,由于预滤波和后置滤波只在时间域处理信号,因此可完全在现有基于块的 2 维 DCT 框架之外进行, $P$  作为对块边界信号进行处理的预滤波器,用于消除块与块之间的相关性,而经预处理后的样值作为 DCT 的输入进行一般的分块 DCT 编码;而

在重建阶段,  $P^{-1}$  则作为后置滤波器, 以交叠方式重建信号, 用于消除块效应, 可见由于 TDLT 在去除块效应的同时, 能够很好地保持分块 2 维 DCT 框架的完整性, 而且多数现有视频压缩编码中所使用的都是这一框架, 因此下面以 H. 263 编解码器为平台, 重点讨论基于 TDLT 的去块效应视频编码框架。

#### 4.2 基于 TDLT 的视频编码框架

H. 263 将视频图像分为 I 帧、P 帧和 B 帧, 而对每帧的处理又分为帧内编码模式和帧间预测编码模式, 由于它们均以宏块为单元进行, 所以本文在使用 TDLT 时也以宏块为单元, 由于 I 帧和 P、B 帧的宏块编码模式不同, 因此对 I 帧和 P、B 帧编码的框架也有所不同。

由于 I 帧所有宏块均采用帧内编码模式, 且在用 TDLT 对每个宏块进行变换时, 需要先取出与其相邻的宏块进行预滤波, 再将所得宏块进行 DCT、量化和熵编码, 其 DCT 等后续过程与 H. 263 原有框架相同, 由此得到的基于 TDLT 的 I 帧编码框架如图 1 中的点划线系统所示; B 帧所有宏块都采用帧间预测编码模式, 因此变换编码只对残差信号进行; P 帧中宏块不是使用统一的编码模式, 其大部分宏块为帧间预测编码模式, 少量宏块采用帧内编码模式, 因此, 在变换编码过程中, 要先对 P、B 帧宏块的编码模式进行判断, 若为帧内模式, 则不进行交叠, 而采用传统的 DCT, 若为帧间模式, 则需先取相邻宏块残差进行预滤波, 再将所得宏块进行 DCT。在取相邻宏块残差时, 对相邻宏块编码模式也要进行判断, 若为帧内模式, 则由于其不进行运动估计, 无残差信息, 应将该宏块残差置 0。

#### 4.3 交叠点数的选择

预滤波器  $P$  的作用相当于对输入信号进行平滑处理, 目的是使 DCT 的输入信号尽可能均匀, 以增加总的能量集中性, 由于预滤波器将高频分量集中在相邻块边缘, 因此平滑后其高频分量并不是被消除了, 只是在时间上被轻微偏移, 由于每块内的样值被平滑增强了 DCT 能量集中的效果, 因此 DCT 输入信号相邻块边缘之间的不连续性不会影响编码效率。由于相邻块边缘各样值的相关性是随着样值之间距离的增加而减小的, 即距离越近, 相关性越大, 距离越远, 相关性越小, 因此在 TDLT 中, 如果交叠点数过少, 则可能因为没有充分利用块边缘之间的相关性而减小编码增益, 对块效应的改善也不能达到最佳效果; 如果交叠点数过多, 也可能会因为对相关性过小的样值进行交叠而引入人为噪声, 并且由

于对图像过度平滑因而也会造成块边缘模糊, 此外随着交叠点数增加, 运算复杂度也会相应增加, 可以通过适当选择  $V$  的阶数, 即块与块之间交叠的点数 (从相邻块所借点数) 来控制图像质量和运算复杂度, 以有利于进行均衡。

对于大小为  $8 \times 8$  的块来说, 交叠点数可以为 1、2、3、4, 其相应的变换依次记为  $8 \times 10$ TDLT,  $8 \times 12$ TDLT,  $8 \times 14$ TDLT,  $8 \times 16$ TDLT。图 5 为量化系数相同的条件下, 用不同交叠模式的 TDLT 对“Foreman”序列进行变换的解码图像第 2 帧。由图 5 可见, 与其他 3 种模式相比, 选择 2 点交叠深度的  $8 \times 12$ TDLT 能得到更好的主观图像质量, 因为它既能充分利用相邻块的相关性, 以改善块效应现象, 又不致因借用相邻块中相关性过小的相邻样值引入人为噪声而造成图像模糊和因过度平滑而造成边缘模糊, 另外, 运算复杂度也不会太大。因此对于大小为  $8 \times 8$  的块来说, 在相同的量化系数下, 选择交叠点数为 2 就能得到较好的图像质量, 且运算复杂度也较低。

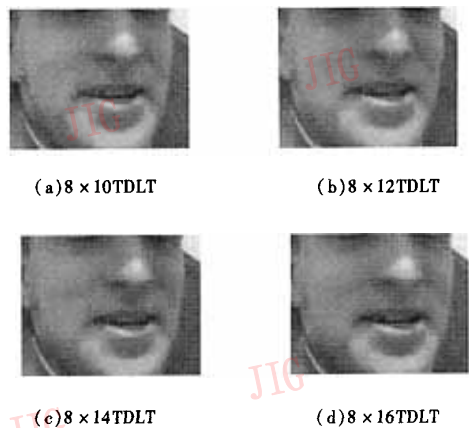


图 5 Foreman 序列不同 TDLT 交叠模式下的解码图像 (第 2 帧)  
Fig. 5 The 2nd reconstructed frame quality with different TDLT lapped modes in Foreman

## 5 实验结果

为验证本文算法效果, 利用 CIF 格式的测试图像序列 Foreman、football 和 Stefan 进行了压缩实验, 每个序列长度为 30 帧, 编码采用第 1 帧为 I 帧, 其余帧都为 P、B 帧, 在 I 帧量化系数相同的前提下, 选择使比特率相等或相近的 P、B 帧量化系数, 交叠点数分别选择为 1 点 ( $8 \times 10$ TDLT) 和 2 点 ( $8 \times 12$ TDLT)。

(1) 解码图像峰值信噪比 (peak signal-to-noise ratio, PSNR) 的比较

图 6(a)、图 6(b) 和图 6(c) 分别为“Foreman”、“football”和“Stefan”序列在两种 TDLT 模式下,与传

统 DCT 的亮度分量峰值信噪比比较图,由图中显示,低码率下“Foreman”序列经 TDLT 变换后的峰值信噪比与 DCT 相比有所提高,football 和 Stefan 序列经 TDLT 变换后的峰值信噪比与 DCT 基本接近。

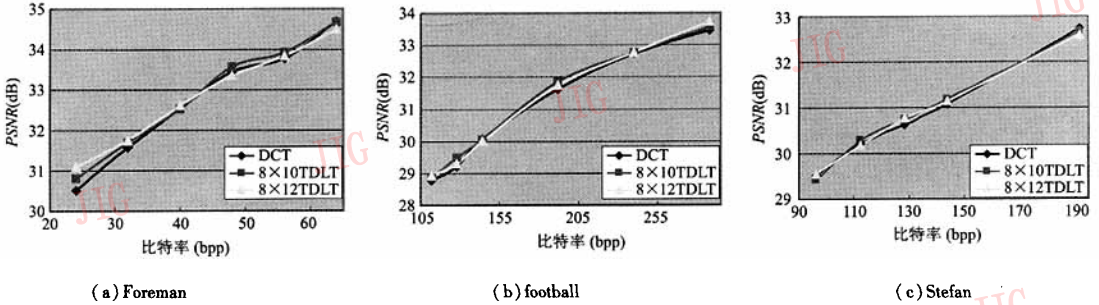


图 6 3 种变换模式下 PSNR 比较

Fig. 6 The PSNR comparison of three transform modes

(2) 主观效果的比较

图像编码最终效果要由人们主观来判断,反复比较 3 种变换模式解码后图像序列的主观效果可得出结论,与 DCT 相比,经 8x10TDLT 和 8x12TDLT 变换解码后的图像序列,其主观效果都有较明显的提高,且块效应现象都得到较大改善,而其中 8x12TDLT 变换的改善效果更为明显。下面给出图像序列的第 2 帧亮度分量的局部放大图像,图 7 和图 8 分别为“Foreman”、“football”序列的 3 种变换解码后的第 2 帧局部图像。

经对上述解码图像序列的峰值信噪比和主观效果的比较可以显示:基于 TDLT 的混合视频编码方案可使解码后的视频序列图像质量明显提高,块效应现象明显改善。

### 6 结论

本文结合传统块变换的视频编码结构,分析和实现了 TDLT 的混合视频编码方案,并通过选择交叠模式均衡各种性能,在附加计算量不大的情况下,有效地抑制基于块变换编码图像序列中的块效应,不仅使视频序列的峰值信噪比得到一定的提高而且具备与传统 DCT 相当的率失真性能,具有很好的应用价值。

### 参考文献 (References)

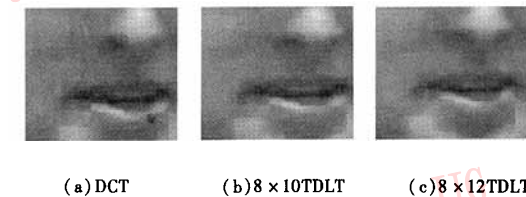


图 7 Foreman 序列的 3 种变换模式下的解码图像(第 2 帧)  
Fig. 7 The 2nd reconstructed frame quality with three transform modes in Foreman



图 8 Football 序列的 3 种变换模式下的解码图像(第 2 帧)  
Fig. 8 The 2nd reconstructed frame quality with three transform modes in football

- 1 Peter L, Anthony J, Jani L, et al. Adaptive deblocking filter[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 614 ~ 619.
- 2 Yang Yong-yi, Nikolas P G. Removal of compression artifacts using projections onto convex sets and line process modeling [J]. IEEE Transactions on Image Processing, 1997, 6(10): 1345 ~ 1357.
- 3 Henrique S M, David H S. The LOT: transform coding without blocking effects[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989, 37(4): 553 ~ 559.
- 4 Tran Tran D, Liang Jie, Tu Cheng-jie. Lapped transform via time-domain pre- and post-filtering [J]. IEEE Transactions on Signal Processing, 2003, 51(6): 1557 ~ 1571.