

DCT 域视频转码技术综述

杜耀刚^{1,2}, 蔡安妮¹

(1. 北京邮电大学电信工程学院, 北京 100876; 2. 北京电子科技学院基础部, 北京 100070)

摘 要: 随着数字视频技术的广泛应用, 实现各种视频编码格式之间转码的要求越来越迫切。视频转码是数字媒体网络网关的关键技术。本文从 DCT 域视频转码的体系结构、关键技术及最优转码策略三个方面, 介绍了现有的各种算法, 分析和总结了各自的特点, 提出了 DCT 域视频下采样帧内刷新体系结构和最优视频转码策略解的概念, 并指出了下一步的研究方向。

关键词: DCT 域视频转码; 率失真; 最优转码策略; 帧内刷新

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2005) 09-1644-07

Overview of the Video Transcoding Techniques in DCT Domain

DU Yaogang^{1,2}, CAI Anni¹

(1. Telecommunication Engineering School, Beijing University of Posts & Telecommunications, Beijing 100876, China;

2. Foundation Department Beijing Institute of Electronic Science and Technology, Beijing 100070, China)

Abstract: With widely applications of digital video, the requirement of video transcoding is very exigent. Video transcoding is defined as the conversion of one coded video format to another and becomes a key technique of digital multimedia network gateway. In this paper, a survey on video transcoding in DCT-domain, including several typical architectures, key techniques and an optimal transcoding strategy, and various existing algorithms to analyze their advantages and disadvantages respectively are presented. An intra-block refreshing architecture for reduced spatial resolution transcoding in DCT domain and a solution of an optimal transcoding strategy are introduced. Several research aspects in the future are proposed.

Key words: video transcoding in DCT domain; rate distortion; an optimal transcoding strategy; intra refresh

1 引言

视频压缩是最具挑战的编码问题, 它有大量的商业应用, 基本目标是为了在给定的码率下达到最优的解码质量。视频压缩方法依据所用信源模型的不同可以分为两类: 一类是基于波形的视频压缩, 比如像 MPEG 和 H. 26x 视频压缩技术都采用了基于块的混合编码方法; 它综合了预测编码、运动补偿、离散余弦变换(DCT)编码和熵编码等技术; 另一类则是基于内容的视频压缩。最好的图像压缩算法是用余弦基或小波基的变换编码, 这种基的效率来源于它们能用很少几个非零系数构造精确非线性图像逼近的能力。DCT 是一种次最优的特征向量(K-L)变换, 具有能量集中特性、去空间相关特性、熵不变特性以及类似于小波多分辨率图像分解之特性^[1], 且多数 DCT 的计算较小波变换简单, 例如, 时域信号的卷积对应于 DCT 域的直接乘积^[2], 而小波变换则无此计算之简单性。因此, 本文重点讨论 DCT 域视频转码问题。

当不同用户试图通过不同的通信链路访问相同的视频, 或流媒体在异构网络中传输时, 常常期望该视频具有可分级

特性或对不可分级的视频能进行视频转码, 而前者的复杂度较高同时功能也很有限。随着数字视频技术在我国国民经济和社会信息化建设中的广泛应用(如视频点播、视频会议、数字图书馆、远程教育、电子商务及家庭娱乐等), 随着互联互通的电信网、计算机网和电视网等网络资源的无缝融合, 实现各种视频编码格式之间的转码要求越来越迫切。

将某种编码格式的压缩码流转换为另一种编码格式的压缩码流称为视频转码(Transcoding)。非实时视频转码就是在转码代理服务器中保存一个质量足够好的压缩视频流, 当需要降低输出码率时, 服务器只需进行部分的解码和编码就可以实现, 而实时视频转码则需要根据信道条件和用户要求提供最好的转码质量。视频转码主要包括再量化减少比特率、下采样减少空间分辨率及丢弃 B 帧减少帧率^[3]。视频转码分为同质转换和异质转换两类, 前者是在同一编码标准中进行转换, 如将 MPEG 2 变为 MPEG 1, 而后者是在不同视频编码标准中进行转换, 如将 MPEG 2 转为 MPEG 4。典型转码结构也有两种^[4]: 一是基于像素域的级联转码, 它由三部分组成: 译码、下采样和再编码, 三部分相互独立, 转码很灵活, 但计算复杂性

很高;二是基于 DCT 域的级联转码,即直接在 DCT 域上进行各种率的转换,不需要费时的运动估计、DCT 及逆 DCT 运算,计算复杂性较低,且能较好地保持图像质量.因此,DCT 域的级联转码普遍受到欢迎.

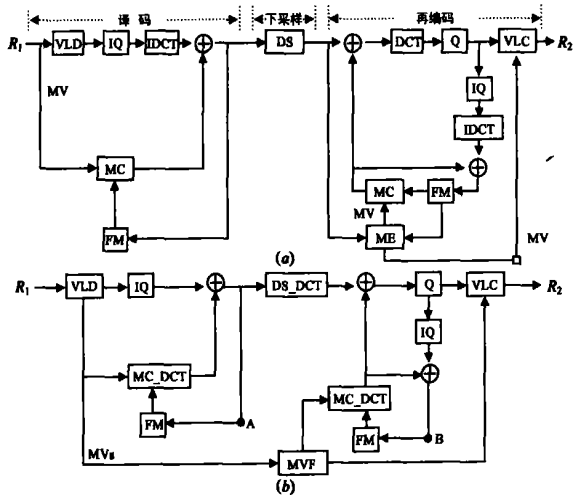
DCT 域视频转码的关键技术包括运动向量的再使用、运动补偿技术、下采样方法、宏块类型确定、码率控制、视频水印转码技术以及最优视频转码策略等.根据用户对转码的质量、有效性、安全性和可靠性的不同需求,这些关键技术(转码参数)既可以独立使用启发式方法(充分利用输入码流中的运动向量及宏块类型等信息,快速确定输出码流相应参数的方法)来确定,也可以联合使用率失真(Rate Distortion, R-D)方法来优化.与启发式方法相比,R-D 方法有较高的复杂度和较好的转码质量.

2 视频转码体系结构

基于块的混合视频编码器有效地联合了运动补偿和变换编码.首先,用基于块的运动估计(ME)由前面已编码的参考帧对块进行预测.运动向量(MV)确定当前块和最佳匹配块之间的位移.然后,依据 ME 精度的不同,用 DCT 对块本身(帧内模式)或预测误差块(帧间模式)进行变换,量化 DCT 系数.最后,将量化的 DCT 系数、MV、编码模式、量化参数及其它信息用可变长度编码转化为比特流 R_1 .视频转码的任务就是如何在尽可能地保持图像质量的同时快速地将源比特流 R_1 转换为目标比特流 R_2 ,其典型的体系结构有如下两种:

2.1 基于像素域的级联转码(CPDT)体系结构

视频转码最简单最灵活的实现方式就是先将输入的压缩流完全解码、用滤波器技术进行下采样^[4],然后再按照输出格式的要求直接压缩成另一种格式的视频流(如图 1(a)所示).CPDT 的计算复杂性最高,即使再使用 MV 和编码模式,也很



VLD: 可变长度译码 IDCT: 逆离散余弦变换 DCT: 离散余弦变换 Q: 量化 IQ: 逆量化 MC: 运动补偿 ME: 运动估计 MV: 运动向量 FM: 帧存储器 DS: 下采样 VLC: 可变长度编码 A: 译码重构帧 DS_DCT: DCT 域下采样 MVF: 运动向量滤波器 MC_DCT: DCT 域运动补偿 B: 转码重构帧

图 1 视频转码体系结构: (a) CPDT; (b) CDDT

难满足实时应用的需求.

2.2 基于 DCT 域的级联转码(CDDT)体系结构

由于可以再使用输入视频流中的 MV、能避免费时 DCT 和逆 DCT 的计算以及有 DCT 域运动补偿(MC-DCT)^[5]和 DCT 域下采样(DS-DCT)的快速算法^[6],导致了该体系结构有较低的复杂度.CDDT 能保持较好的图像质量,但其灵活性受到限制^[7](如图 1(b)所示).CDDT 最基本的思想就是将像素域的处理过程用等价的 DCT 域公式来代替.

3 DCT 域视频转码关键技术

3.1 运动向量的再使用

运动估计(ME)是基于块的匹配技术,主要目的是去除或减少时域的冗余性,使编码预测误差的比特率最小化,ME 算法是基于恒定亮度假设,且每个块的运动被假定是恒定的. ME 占整个编码时间的 60% 以上^[8],所以视频转码中,运动向量再使用技术(MVR)受到欢迎.例如在 2:1 下采样视频转码中,MVR 的启发式方法主要有三种:一是自适应运动向量下采样(AMVR)方法,即下采样 MV 是相邻四个输入宏块 MV 的加权和^[9];二是主运动向量(MMVM)方法,即从相邻四个输入宏块 MV 中,按照某种准则挑选出其中一个做为下采样 MV (称为主 MV),典型的准则有:最小欧氏距离准则^[10]和低频 DCT 系数用可测量化矩阵归一化准则^[11];三是主 MV 精细方法(MVRM)^[12],即在主 MV 和其余三个输入 MV 的简单平均所构成的搜索窗内,求最小均方差值(MSE)所对应的 MV.就重构图像的峰值信噪比(PNSR)、转码复杂性和编码所需的比特数而言,最大最小距离准则方法较好.至于任意比率下采样视频转码的 MVR 技术,采用 MV 的加权和^[13]为宜,它特别适用于输入宏块与编码宏块的边界不对齐的情况.

传统的 ME 方法只注意使运动补偿预测误差码率 R_{DFD} 最小化,而 R-D 最佳化还应考虑产生的 MV 以预测方式编码所需的码率 R_{MV} ^[14]. R_{MV} 随运动向量精度 Δ 的减少而单调增加,但 R_{DFD} 随运动向量精度 Δ 的减少而单调减少,最佳运动向量精度 $\Delta^* = \arg \min_{\Delta} (R_{MV} + R_{DFD}) = \frac{1}{2}$ ^[15],这就是在 MPEG 转码中经常使用半像素精度的原因.

3.2 DCT 域运动补偿技术(MC-DCT)

MV 确定了当前块与最佳匹配块之间的位移.在像素域中,用于运动补偿的预测块 $\hat{x}$ 很容易由当前块 x 及相应的 MV 获得.但在 DCT 域中预测块与参考帧 DCT 块的边界通常是不对齐的,所以不太容易获得预测块 $\hat{x}$,需要用式(1)来计算,它约占整个转码时间的 30%.

设当前帧块 X 的预测误差是 E ,而经 MC_DCT 得到的预测块是 $\hat{X}$,则译码时,有 $X = \hat{X} + E$;编码时,有 $E = X - \hat{X}$.

$$\text{其中: } \hat{X} = \sum_{i=1}^4 P_i Z_i Q_i \quad (1)$$

$$P_i = C_8 p_i C_8^t, Q_i = C_8 q_i C_8^t, \quad i = 1, 2, 3, 4 \quad (2)$$

$$p_1 = p_2 = u_h = \begin{bmatrix} 0 & I_h \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & I_7 \\ 0 & 0 \end{bmatrix}^{8-h} \quad (3)$$

$$q_1 = q_3 = l_w = \begin{bmatrix} 0 & 0 \\ I_w & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ I_7 & 0 \end{bmatrix}^{8-w} \quad (4)$$

$$p_3 = p_4 = l_8 - h, q_2 = q_4 = u_8 - w \quad (5)$$

C_8 代表 8×8 的 DCT 变换矩阵, 而 I_7 表 7×7 的单位矩阵 (6)

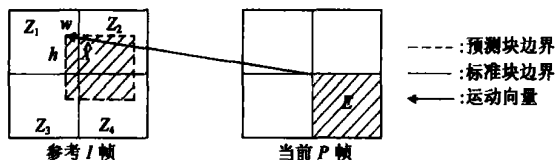


图 2 P 帧中的帧宏块运动补偿示意图

w, h 分别是预测块 X 与 Z_1 相交的列数和行数 ($0 \leq w, h \leq 7$), Z_i 为参考帧中的块 (如图 2 所示). 左乘 p 矩阵和右乘 q 矩阵分别相当于矩阵的行、列变换, 式 (1) 表示将预测块 $\hat{x}$ 与 z_i 相交的矩形块分别移动到 z_1 的左上角、 z_2 的右上角、 z_3 的左下角及 z_4 下角上去, 而将 z_i 其余部分变成零, 目的是将预测块 $\hat{x}$ 分解为 z_i 的代数和 (见图 3):

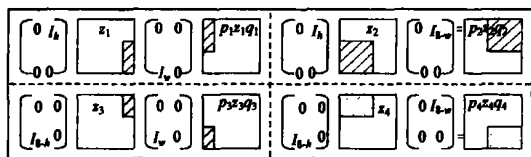


图 3 利用矩阵乘法抽取并平移子块示意图

MC-DCT 快速算法^[5]的基本思想就是利用 DCT 变换的稀疏特性、线性叠加性以及提前将矩阵 P_i 和 Q_i ($i = 1, 2, 3, 4$) 计算出来, 它分为两类: 基于 DCT 变换的快速算法和基于视频内容的快速算法. 基于 DCT 变换的快速算法就是将 DCT 变换矩阵分解为稀疏矩阵或提升矩阵 (即对角线元素全是 1、非对角线元素中仅有一个为非零的矩阵) 之积^[6], 以减少乘法和加法的数目, 与传统方法相比, 可节省 37% 到 50% 的运算量^[5]. 在此基础上, 再利用宏块之间的共享信息可进一步提高 DCT 域逆运动补偿的运算速度^[17].

现将式 (1) 重新改写为:

$$\hat{X} = P_1(Z_1Q_1 + Z_2Q_2) + P_3(Z_3Q_1 + Z_4Q_2) \quad (7)$$

就得到了基于视频内容 MC 的快速算法^[18], 其基本思想是利用视频内容 DCT 系数的能量集中特性来减少矩阵的计算量, 比如计算 $Z_1Q_1 + Z_2Q_2$ 时, 若 Z_1 第 1 行非零 DCT 系数的最大下标是 2 (假设 DC 系数的下标是 (0, 0)), 而 Z_2 第 1 行非零 DCT 系数最大下标是 1, 则计算 $Z_1Q_1 + Z_2Q_2$ 第 1 行时, 仅需要计算矩阵的前 3 个元素, 其余元素均可置为零, 从而大大减少了矩阵乘法和加法的数目.

3.3 DCT 域下采样方法

视频转码主要包括再量化减少比特率、下采样减少空间分辨率及丢弃 B 帧减少帧率. 依据应用的不同, 各种率的减少, 既可以用 R-D 方法联合优化, 也可以单独处理.

由于 DCT 的 AC 系数通常假定服从 Laplacian 分布而非均匀分布, 在再量化之前, DCT 系数的均匀重构可能不是最优的^[19]. 如果均匀重构被合适的量化区间的质心所代替, 则再量化错误就可以被减少^[20].

时域分辨率的上采样或下采样是通过在码流中添加 B 帧或丢弃 B 帧来实现的. B 帧使用与它在时间上最近邻的前

后两个 I 帧或 P 帧来预测, 而自己并不作为任何其它帧的参考图像, 因此, 在传输中丢弃 B 帧^[21]并不影响其它帧的质量, 而仅仅会降低帧率.

在许多多媒体应用中, 人们常常希望同一视频内容能有不同空间分辨率的压缩版本, 例如, 希望相同的图像可以在不同带宽信道上, 以不同的空间分辨率进行传输; 又如浏览视频数据库时, 希望可以先检索低分辨率图像, 如果感兴趣, 再浏览高分辨率的图像等等. 而从高分辨率的图像获得低分辨率的图像则需要进行下采样. 对已压缩视频进行图像下采样的方法可分为两类: 空间域技术和压缩域技术. 空间域技术将已压缩的视频数据先解压缩, 再下采样, 然后重新进行压缩, 生成低分辨率的码流. 这种方式的主要缺点, 一是计算复杂性高; 二是在解压缩时需要额外的存储; 三是会损失一些编码精度. 而压缩域上的下采样技术, 可以有效地减少计算复杂性, 同时可较好地保持原始图像的质量, 因此越来越受到重视. 典型的 DCT 域 2 1 下采样方法主要有三个 (基本思想是将四个块变为一个块): 一是 Mehav 提出的抗混叠滤波器^[22] (又称为双线性滤波器), 即将每一个 2×2 的子块用一个像素代替, 其强度是原始 4 个像素的平均值; 二是 Dugad 等人提出的从每一个 8×8 块上截取 4×4 低频率子块的方法^[23], 即首先对截取的子块进行 4×4 的逆 DCT 变换, 四个这样的子块便构成像素域上一个 8×8 块, 然后进行 8×8 的 DCT 变换即可; 三是我们提出的能较好保持高频能量的准卷积下采样技术^[6], 其基本思想简要说明如下:

假设将组成 DCT 域上任意一个宏块 $X = \begin{bmatrix} X_1 & X_3 \\ X_2 & X_4 \end{bmatrix}$ 的每一个块 X_i , $i = 1, 2, 3, 4$ 进行 8×8 的逆 DCT 变换, 得到像素域上的对应宏块 $x = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$; 再对宏块 x 进行 16×16 的 DCT 变换, 然后截取其 8×8 的低频矩阵, 并将该低频矩阵的各系数均除以 2, 作为 DCT 域上的 2 1 下采样结果 X_{LL} (见图 4).

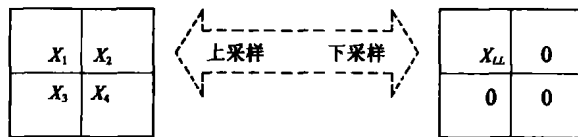


图 4 基于宏块的下采样方案示意图

实际上, 任意 2 1 下采样可统一写成:

$$X_{LL} = AXA^T \quad (8)$$

相应的上采样可统一写为:

$$X = BX_{LL}B^T \quad (9)$$

其中 A 是 8×16 的下采样算子, 而 B 是 16×8 的上采样算子.

值得一提的是, 上述准卷积下采样思想可推广到 DCT 域任意下采样上去^[24].

3.4 下采样宏块类型的确定

在下采样视频转码中, 转码失真就是转码重构帧与译码重构帧之间的均方误差 (MSE), 即图 1(b) 中 A、B 两点之间的 MSE. 引起转码失真的主要原因是: 再量化、运动向量的截断、下采样以及传输错误等. 帧间编码能使转码失真沿运动补偿

后者通常是建立在率失真的基础上,依据缓存器的充满程度、图像空间复杂性及信道可利用带宽来调整帧率及选择最佳宏块模式和量化参数 $(Q)^{[39,40]}$,在尽可能保证图像质量稳定的条件下,使转码码流的速率适应恒定信道的要求。

码率控制算法分为两类:一类是 R NZ 线性模型^[27],即量化 DCT 系数所需比特率 R 与量化后的非零 DCT 系数个数的百分比 ρ 存在着线性关系 $(R = \theta \cdot (1 - \rho) + R_H)$;另一类是联系 Q 与编码模式的 R Q 非线性模型 $(R = S \cdot \left[\frac{X_1}{Q} + \frac{X_2}{Q^2} \right])$,这里 S 是转码复杂性, X_1 和 X_2 是模型参数,其中假设 DCT 系数近似服从拉普拉斯分布^[41]。

建立图像 DCT 系数的准确数学模型,直接影响利用 R D 优化技术进行码率控制的效率。Lam 认为图像 DCT 系数近似服从拉普拉斯分布^[42];如果图像方差为一个定值,则由中心极限定理,图像 DCT 系数近似服从高斯分布;但是对于大多数图像来说,图像块之间的方差是变化的,已经证明,压缩视频数据(例如已量化和游程编码的 DCT 系数, MV 差值,头信息等)可由所谓的广义高斯分布(GGD)很好地建模^[43]:

$$R(D) = \sigma^2 \log_2 \left(\frac{b}{D} \right) \quad (10)$$

$$R(Q) = k + m \log_2(Q) \quad (11)$$

其中 D , σ^2 和 Q 分别是该宏块的均方失真、方差和量化参数,而 a , b , k , m 均是模型参数。

如果能有效地利用输入视频和输出视频之间的相关性,视频转码的码率控制效果会更好。譬如当分辨率和比特率发生改变时,图像编码的输入复杂性和输出复杂性不同,但其比值近似为常数且对场景不敏感^[10],利用此特性就可进行转码的目标比特分配;又如在 R NZ 线性模型中,转码前后的也具有一定的线性关系^[44]。当然,转码前后 Q 之间的最优表达式一直是我们所期望的,目前仍然没有理想的结果。

以上所述的视频转码关键技术中,可以独立进行研究,也可以部分联合研究,譬如使用 Lagrange 乘子法得到了最优量化参数的解析表达式^[45],再例如在量化技术与 DCT 方法的融合^[46]以及在下采样技术与运动补偿技术的融合^[47]方案中找最优的运动补偿方案^[48]等等。

4 最佳转换策略

视频转码最终目标是在转码质量、有效性、可靠性及安全性等方面取得最佳平衡。优化问题的本质是根据视频转码的特点及 R D 理论来联合优化比特分配问题和缓存器控制问题,即联合信源及信道确定使 Lagrange 费用函数 $D + \lambda_1 \left(\frac{R}{R_{\text{channel}}} - T \right) + \lambda_2 C + \lambda_3 S$ 达到最小的转码参数就是我们要寻找的最优视频转码策略解,关键问题是确定 Lagrange 乘子 λ 和费用函数的阈值。其中, R 是信源和信道编码的总比特数; R_{channel} 是信道的传输率; T 是帧传输延迟阈值; C 是转码的计算复杂性; S 是转码的安全性。在误码环境中图像的转码失真由两部分组成:信源转码失真(主要由再量化引起)与信道误码失真(主要由信道错误引起)。假定 $f(n, i)$ 为当前 n 帧

像素点 i 上的原始值,转码端重建此值为 $\tilde{f}(n, i)$,经过信道传输后在解码端的重建值为 $\hat{f}(n, i)$,假设宏块的信源转码失真与信道误码失真统计独立,并采用 MSE 作为失真衡量标准,则转码宏块的总失真期望函数为:

$$D(n) = E \left\{ [f(n, i) - \tilde{f}(n, i)]^2 \right\} = D_q(n) + D_c(n) \quad (12)$$

其中, $D_q(n) = E \left\{ [f(n, i) - \tilde{f}(n, i)]^2 \right\}$ 为信源转码失真, $D_c(n) = E \left\{ [\tilde{f}(n, i) - \hat{f}(n, i)]^2 \right\}$ 为信道误码失真。实际上 $D_q(n)$ 和 $D_c(n)$ 都是信道编码率和帧内刷新率的函数^[26,49,50]。

R D 优化技术主要有两种: Lagrange 乘子法和动态程序(DP)方法,两者最主要的区别有两点:一是, Lagrange 乘子法的最优解只能在 R D 曲线的凸包络线上取得,而 DP 方法没有此限制;二是,就计算复杂性而言, DP 方法随转码宏块数目的增多而呈指数增长,但 Lagrange 乘子法仅线性增加^[51]。

5 结论与展望

视频转码是数字媒体网络网关的关键技术和应用热点。本文从 DCT 域视频转码技术的体系结构、关键技术及最佳视频转换策略三个方面,回顾了常用的方法和目前的研究进展。提出了 DCT 域视频下采样帧内刷新体系结构和最优视频转码策略解的概念。

在压缩域上如何根据信道条件、终端类型、视频特征及用户对转码的质量、有效性和安全性的不同需求,确定是独立使用启发式方法(充分利用输入码流中的运动向量、宏块类型及量化参数等信息,快速确定输出码流相应参数的方法)、还是联合使用 R D 方法来优化:使转码的视频质量最好、转码的计算复杂性最低,同时使嵌入的视频水印和抗误码信息的容量最大、鲁棒性最强,仍然面临极大的挑战。显然,更精确、更鲁棒的视频转码方法依赖于信源所采用的编码方案、信道错误传输模型、网络的拥塞状况、接收端采用的错误隐藏方案、视频数据的保密方式以及信源和信道的 R D 模型等。

另外,目前 MPEG-2 到 H.264 的快速转码尚无较好的解决办法。MPEG-2 主要应用于大约 4Mbps 速率的 TV 广播以及 15Mbps 的高质量视频方面,而 H.264^[52] 广泛应用于移动通信及 IP 信道传输等领域中。目前国际上对于 MPEG-2 到 H.264 的转码研究仅就“宏块模式选择、运动向量再使用、变换核转换及 DCT 域帧内预测”等方面开始做了一些简单讨论。由于 MPEG-2 与 H.264 在编码方式上存在较大差别,因此在快速转码和最优转码方面有许多新的问题需要解决。

参考文献:

- [1] Debin Z, Wen G, Chan Y K. Morphological representation of DCT coefficients for image compression[J]. IEEE Trans on CSVT, 2002, 1(12): 1287-1293.
- [2] Martucci S A. Symmetric convolution and the Discrete Sine and Cosine Transforms[J]. IEEE Transactions on Signal Processing, 1994, 42(5): 1038-1051.
- [3] Sugano M, Nakajima Y, Yanagihara H, et al. An efficient transcoding from MPEG-2 to MPEG-1[A]. IEEE Image Processing 2001 Proceedings

- 2001 International Conference[C]. Vancouver, Canada: IEEE, 2001, 1: 417–420.
- [4] Zhu W, Yang K, Beacken M. CIF to QCIF video bitstream down conversion in the DCT domain[J]. Bell Labs Technical Journal, 1998, 3(3): 21–29.
 - [5] Mehav N, Bhaskaran V. Fast algorithms for DCT-Domain image downsampling and for inverse motion compensation[J]. IEEE Trans on CSVT, 1997, 7(3): 468–476.
 - [6] 杜耀刚, 蔡安妮, 孙景鳌. 一种高效的 DCT 域图像下采样方法[J]. 中国图像图形学报, 2005, 10(4): 511–515. (Chinese Source) DU Yao Gang, CAI An Ni, SUN Jing Ao. An Efficient Method for DCT-Domain Image Downsampling[J]. Journal of Image and Graphics, 2005, 10(4): 511–515. (Chinese Source)
 - [7] Song M Z, Cai A, Sun J A. Motion estimation in DCT domain[A]. IEEE Communication Technology Proceedings, 1996. ICCT'96., 1996 International Conference[C]. Washington, DC: IEEE, 1996, 2: 670–674.
 - [8] Pirch P, Demassieux N, Gehrke W. VLSI architectures for video compression A survey[J]. Proceeding of the IEEE, 1995, 83(2): 220–246.
 - [9] Shen B, Ishwar I K, Bhaskaran V. Adaptive motion vector resampling for compressed video downscaling[J]. IEEE Trans on CSVT, 1999, 9(6): 929–936.
 - [10] Xin J, Sun M T, Choi B S, et al. An HDTV to SDTV spatial transcoder[J]. IEEE Trans on CSVT, 2002, 12(11): 998–1008.
 - [11] Yusuf A A, Mushed M, Dooley L S. An HVS based motion vector composition algorithm for spatial resolution transcoding[A]. IEEE Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference[C]. Orlando, FL: IEEE, 2004, 2: 682–688.
 - [12] Seo K D, Kim J K. Motion vector refinement for video downsampling in the DCT domain[J]. IEEE Signal Processing Letters, 2003, 9(11): 356–359.
 - [13] Zhang Y Q, Chau L P, Tan Y P. Arbitrary downsizing video transcoding using fast motion vector reestimation[J]. IEEE Signal Processing Letters, 2002, 9(11): 352–355.
 - [14] Chen M C, Wilson A N. Rate distortion optimal motion estimation algorithms for motion compensated transform coding[J]. IEEE Trans on CSVT, 1998, 8(2): 147–158.
 - [15] Jordi R C, Neuhoff D L. Optimizing motion vector accuracy in block based video coding[J]. IEEE Trans on CSVT, 1999, 11(4): 497–511.
 - [16] Zeng Y, Cheng L, Bi G, et al. Integer DCTs and Fast Algorithms[J]. IEEE Transactions on Signal Processing, 2001, 49(11): 2774–2782.
 - [17] Song J, Yeo B L. A Fast Algorithm for DCT-Domain inverse motion compensation based on shared information in a macroblock[J]. IEEE Trans on CSVT, 2000, 10(5): 767–775.
 - [18] Lius, Bovik A C. Local bandwidth constrained fast inverse motion compensation for DCT-domain video transcoding[J]. IEEE Trans on CSVT, 2002, 12(5): 468–476.
 - [19] Werner O. Requantization for transcoding of MPEG-2 Intraframes[J]. IEEE Trans on Image Processing, 1999, 8: 179–191.
 - [20] Werner O. Requantization for transcoding of MPEG-2 Intraframes[J]. IEEE Trans on Image Processing, 1999, 8(2): 179–191.
 - [21] Funk T, Chan Y L, Siu W C. Low-complexity and high quality frame skipping transcoder for continuous presence multipoint video conferencing[J]. IEEE Trans on Multimedia, 2004, 6(1): 31–46.
 - [22] Mehav N, Bhaskaran V. Fast Algorithms for DCT-Domain Image Down Sampling and for Inverse Motion Compensation[J]. IEEE Trans on CSVT, 1997, 7(3): 468–476.
 - [23] Dugad R, Ahuja N. A Fast Scheme for Image Size Change in the Compressed Domain[J]. IEEE Trans on CSVT, 2001, 11: 461–474.
 - [24] Shu H, Chau L P. An efficient arbitrary downsizing algorithm for video transcoding[J]. IEEE Trans. on CSVT, 2004, 14(6): 887–891.
 - [25] Yin P, Vetro A, Liu B, et al. Drift compensation for reduced spatial resolution transcoding[J]. IEEE Trans on CSVT, 2002, 12(11): 1009–1020.
 - [26] Stuhlmuler K, Farber N, Link M, et al. Analysis of video transmission over lossy channels[J]. IEEE on Selected Areas in Communications, 2000, 18(6): 1012–1032.
 - [27] Zhihai H, Mitra S K. A linear source model and a unified rate control algorithm for DCT video Coding[J]. IEEE Trans on CSVT, 2002, 11(6): 970–981.
 - [28] Wiegand T, Lightstone M, Mukherjee D, et al. Rate distortion optimized mode selection for very low bit rate coding and the emerging H. 263 standard[J]. IEEE Trans on CSVT, 1996, 6(2): 182–190.
 - [29] Kim H, Altunbask Y. Low-complexity Rate distortion optimal macroblock mode selection for MPEG like video coders[A]. IEEE Multi media and Expo, 2003. ICME'03[C]. Proceedings, Atlanta, GA, 2003, 3: 513–516.
 - [30] Reyes G D, Reibman A R, Chang S F, et al. Error-resilient transcoding for video over wireless channels[J]. IEEE J Select Areas Commun, 2000, 18: 1063–1073.
 - [31] Wu D, et al. An end-to-end approach for optimal mode selection in Internet video communication: Theory and application[J]. IEEE Journal on Selected Areas in Communications, 2000, 18(6): 977–995.
 - [32] Wang Y, Zhu Q F. Error control and concealment for video communication: a review[J]. Proceedings of the IEEE, 1998, 86(5): 974–997.
 - [33] Nikolaidis N, Pitas I. Robust image watermarking in the spatial domain[J]. IEEE Transactions on Signal Processing, 1998, 66(3): 385–403.
 - [34] Cox I J, Kilian J, Leighton T, et al. Secure spread spectrum watermarking for multimedia[J]. IEEE Trans on Image Processing, 1997, 6(12): 1673–1687.
 - [35] Watson A B. Visually optimized DCT quantization matrices for individual images[A]. Data Compression Conference, DCC'93[C]. Los Angeles: DCC'93, 1993. 178–187.
 - [36] Wade Trappe, et al. Key management and distribution for secure multimedia multicast[J]. IEEE Trans on Multimedia, 2003, 5(4): 544–557.
 - [37] Wee S J, Apostolopoulos J G. Secure scalable streaming enabling transcoding without decryption[A]. in Proc IEEE Int Conf Image Processing[C]. Thessaloniki, Greece: PIEEE, 2001, 1: 437–440.
 - [38] Pao I M, Sun M T. Encoding stored video for streaming applications[J]. IEEE Trans on CSVT, 2001, 11(6): 199–209.
 - [39] Sun H, Kwok W, Chien M, et al. MPEG Coding performance improvement by jointly optimizing coding mode decisions and rate control[J]. IEEE Trans on CSVT, 1997, 7(3): 449–458.
 - [40] Grecos C, Jiang J. On-Line Improvements of the Rate Distortion Per-

- formance in MPEG 2 Rate Control[J]. IEEE Trans. on CSVT, 2003, 13 (6): 519– 528.
- [41] Chiang T, Zhang Y Q. A new rate control scheme using quadratic rate distortion model[J]. IEEE Trans Circuits Syst Video Technol. 1997, 7 (1): 246– 250.
- [42] Lam E Y, Goodman J W. A mathematical analysis of the DCT coefficients distribution for images[J]. IEEE Transactions on Image Processing, 2000, 9(10): 1661– 1666.
- [43] Ortega A, Ramchandran K. Rate distortion method for image and video compression [J]. IEEE Signal Processing Magazine, 1998, 23– 50.
- [44] Chen L, Cheng L J, Yu S Y. Accurate rate control method in transcoding[J]. IEEE, Electronics Letters, 2004, 40(1): 16– 18.
- [45] Jordi R C, Lei S. Rate control in DCT video Coding for low delay communications[J]. IEEE Trans Circuits Syst Video Technol. 1999, 9(1): 172– 185.
- [46] Docef A, Kossentini F, Phi K N, et al. The quantized DCT and Its application to DCT-based video coding[J]. IEEE Trans on Image Processing, 2002, 11: 177– 187.
- [47] Lee Y R, Lin C W, Chen Y W. Computation reduction in cascaded DCT-domain video downscaling transcoding[A]. IEEE Circuits and Systems, 2003. ISCAS' 03. Proceedings of the 2003 International Symposium[C]. Lausanne, Switzerland: IEEE, 2003, 2: II- 860 II- 863.
- [48] Vetro A, Sun H. Frequency domain down conversion of HDTV using an optimal motion compensation scheme[A]. Proceedings International Conference on[C]. PIC, 1997, 1. 763– 766.
- [49] Vetro A, Christopoulos C, Sun H. Video transcoding architectures and techniques: an overview[J]. IEEE SignalProcessing Magazine, 2003, 20 (2): 18– 29.
- [50] He Z H, Cai J F, Chen C W. Joint source channel Rate distortion analysis for adaptive mode selection and rate control in wireless video coding[J]. IEEE Trans Circuits Syst. Video Technol, 2002, 12(6): 511– 523.
- [51] He Z H, Mitra S K. A linear source model and a unified rate control algorithm for DCT video Coding[J]. IEEE Trans Circuits Syst Video Technol, 2002, 11(6): 970– 981.
- [52] T Wiegand, G J Sullivan, G Bjontegaard, A Luthra. Overview of the H. 264/AVC Video Coding Standard[J]. IEEE Trans CSVT, 2003, 13 (7): 560– 576.

作者简介:



杜耀刚 男, 1964 年 3 月生于山西万荣, 北京电子科技学院基础部副教授, 现为北京邮电大学通信与信息系统专业在职博士研究生, 研究方向为多媒体通信、视频转码和视频水印等。
E-mail: duyao gang@best.i.edu.cn.



蔡安妮 女, 1943 年生于浙江德清, 教授, 博士生导师, 1989 年在美国加利福尼亚大学圣巴巴拉分校电机与计算机工程系统获博士学位, 研究领域为多媒体通信、流媒体技术、图像处理与图像识别等, 发表论文 60 余篇, 出版专著 6 部。