

一种高效的 DCT 域图像下采样方法

杜耀刚^{1),2)} 蔡安妮¹⁾ 孙景鳌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 在许多多媒体应用中,人们往往希望压缩的视频数据可以在 DCT 域上进行降低空间分辨率的处理。常见的 Dugad 方法是从每一个 8×8 的 DCT 块中截取 4×4 的低频分量,但该方法容易产生错误漂移及块效应,其主要原因是没有很好地保持图像的高频分量。为此提出了一种称之为准卷积下采样方法的 DCT 域图像下采样方法:从 8×8 块 DCT 系数推算出 16×16 宏块的 DCT 系数,然后截取 8×8 的低频分量。并通过分析 DCT 系数矩阵的稀疏特性给出了快速算法。该下采样方法适用于任何基于 DCT 压缩的视频转码方案。实验结果表明,该方法重构图像的峰值信噪比比 Dugad 方法平均高 0.3dB,并且在图像细节处,视觉效果较好。

关键词 DCT 域 准卷积下采样 快速算法 峰值信噪比

中图分类号: TN919.8 文献标识码: A 文章编号: 1006-8961(2005)04-0511-05

An Efficient Method for DCT-domain Image Downsampling

DU Yao-gang^{1),2)}, CAI An-ni¹⁾, SUN Jing-ao¹⁾

¹⁾(Telecommunication Engineering School, Beijing University of Posts & Telecommunications, Beijing 100876)

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

Abstract Spatial resolution reduction of compressed video in the DCT domain are highly desirable in many multimedia applications. A popular approach presented by Dugad is to extract the 4×4 low frequency components from each 8×8 DCT block, but this method may cause error drift and artifacts due to over-loss of high frequency components. In this paper, we propose a new method of DCT-domain image downsampling, in which 16×16 DCT coefficients of a macroblock are deduced from the coefficients of the four blocks that constitute the macroblock, and the 8×8 low frequency components are then preserved. We also present a fast algorithm to implement this method with exploitation of the sparseness of DCT matrices. The proposed method is applicable to any DCT-based compressed video transcoding. Experimental results confirm that our scheme performs about 0.3dB better than Dugad's method in terms of PSNR of reconstructed image, and better subjective quality can be observed on details of the image.

Keywords DCT domain, quasi-convoluted down-sampling, fast algorithm, peak signal to noise ratio

1 引言

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。而从高分辨率图像获得低分辨率图像则需要

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。空间域技术是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。

空间域下采样的方法有:Merhav 提出的抗混叠

收稿日期:2003-11-20; 改回日期:2004-09-13

第一作者简介:杜耀刚(1964 ~),男,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、数字水印和信息安全。E-mail:duyaogang@besti.edu.cn

虑滤波器^[1],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;Hu 假定相邻 4 个像素的强度相等,任抽取一个,得到了其 DCT 域上的关系式^[2]。压缩域的图像下采样一般在 DCT 域上进行,DCT 域 2 比 1 的下采样方法主要有: Dugad 等人提出的从每一个 8×8 块上截取 4×4 的低频率系数的方法^[3],这主要是利用了 DCT 变换的能量集中特性^[4];Mukherjee 等人对 Dugad 方法的改进^[5];Yim 等人提出的一种下采样算子^[6],该方法实质上是空间域技术和压缩域技术的结合等等。但這些技术存在以下两个问题:一是简单地舍去 DCT 的高频分量,往往会在转换后的视频码流中引入新的误差,而这种误差在一个 GOP 内是逐步传播的,即带来错误漂移;二是截去 DCT 的高频分量后,逆 DCT 变换可能使量化噪声集中到某一个小块内,使相邻像素的误差显著不同,而产生图像失真。

针对上述问题,提出了 DCT 域上一种新的下采样方案:从每一个 16×16 的宏块中截取 8×8 的低频分量,并将相应的各低频系数均除以 2 作为 DCT 域上的下采样结果,这种方法尽可能地保持了图像高频分量。给出了任意 1 个宏块与组成该宏块的 4 个子块 DCT 系数之间的关系式,并利用下采样 DCT 系数矩阵的稀疏特性,给出了快速计算方法。实验结果表明:经上采样后,重构图像的峰值信噪比(PNSR)优于 Dugad 方法。

2 DCT 域上下采样的一般框架

图 1 所示的下采样过程可用如下的下采样算子 T 来描述:

$$T_{16}: X = \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \rightarrow X_1 \quad (1)$$

其中, $X_i = T_{16} X T_{16}^T$, T 的下标表示变换矩阵的维数。下采样算子构成方式的不同则形成了不同的下采样算法。

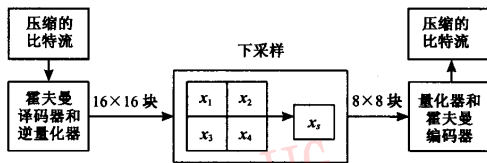


图 1 DCT 域上 2:1 下采样的一般框图

Fig. 1 DCT domain based down-sampling by factor of two

3 准卷积下采样方案

假设将组成 DCT 域上任意一个宏块 $X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}$ 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 $x = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$;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 2)。

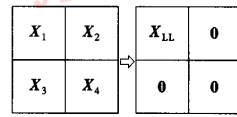


图 2 基于宏块的下采样方案示意图

Fig. 2 Macroblock-based down-conversion from four blocks to one block

给定图像数据 $x = \{x(m, n)\}_{m, n=0}^{N-1}$, 则它的 $N \times N$ 的 DCT 变换为

$$X(k, l) = \frac{2}{N} C(k) C(l) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left[x(m, n) \times \cos \frac{(2m+1)k\pi}{2N} \cos \frac{(2n+1)l\pi}{2N} \right] \quad (2)$$

$$k, l = 0, 1, \dots, N-1$$

其中, $C(y) = \begin{cases} \frac{1}{\sqrt{2}} & y=0 \\ 1 & \text{其他} \end{cases}$

定义 $N \times N$ 的 DCT 变换矩阵 $C_N = \left\{ \sqrt{\frac{2}{N}} C(k) \cos \frac{(2i+1)k\pi}{2N} \right\}_{k,i=0}^{N-1}$, 它满足 $C_N^{-1} = C_N^T$, 而式(2)可写为矩阵表达式:

$$X = \text{DCT}(x) = C_N x C_N^T \quad (3)$$

将 16×16 的 DCT 变换矩阵 C_{16} 表示为

$$C_{16} = \begin{bmatrix} C_{16LL} & C_{16LH} \\ C_{16HL} & C_{16HH} \end{bmatrix} \quad (4)$$

用 W_{16} 表示 16×16 的低通滤波器: $W_{16} = \frac{1}{\sqrt{2}} I_2 \otimes$

$\begin{bmatrix} I_8 & 0 \\ 0 & 0 \end{bmatrix}$, 其中, \otimes 表示矩阵 Kronecker 积, 则图 2 所示方案对应的下采样算子为

示方案对应的下采样算子为

$$T_{16} = W_{16} C_{16} \begin{bmatrix} C_8^T & \mathbf{0} \\ \mathbf{0} & C_8^T \end{bmatrix} \quad (5)$$

其中, C_8 代表 8×8 的 DCT 变换矩阵。根据式(1) 得到:

$$\begin{aligned} \begin{bmatrix} X_{LL} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} &= T_{16} X T_{16}^T \\ &= \frac{1}{2} \begin{bmatrix} I_8 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} C_{16LL} & C_{16LH} \\ C_{16HL} & C_{16HH} \end{bmatrix} \times \\ &\quad \begin{bmatrix} C_8^T & \mathbf{0} \\ \mathbf{0} & C_8^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \begin{bmatrix} C_8 & \mathbf{0} \\ \mathbf{0} & C_8 \end{bmatrix} \times \\ &\quad \begin{bmatrix} C_{16LL}^T & C_{16LH}^T \\ C_{16LH}^T & C_{16HH}^T \end{bmatrix} \begin{bmatrix} I_8 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (6) \end{aligned}$$

化简整理得

$$\begin{aligned} X_{LL} &= \frac{1}{2} (C_{16LL} C_8^T X_1 C_8 + C_{16LH} C_8^T X_3 C_8) C_{16LL}^T + \\ &\quad \frac{1}{2} (C_{16LL} C_8^T X_2 C_8 + C_{16LH} C_8^T X_4 C_8) C_{16LH}^T \quad (7) \end{aligned}$$

令

$$P = \frac{1}{\sqrt{2}} C_{16LL} C_8^T, Q = \frac{1}{\sqrt{2}} C_{16LH} C_8^T \quad (8)$$

则

$$\begin{aligned} X_{LL} &= P X_1 P^T + Q X_3 P^T + P X_2 Q^T + Q X_4 Q^T \\ &= \frac{1}{2} (P+Q)(X_1+X_3)P^T + \frac{1}{2} (P-Q)(X_1-X_3)P^T + \\ &\quad \frac{1}{2} (P+Q)(X_2+X_4)Q^T + \frac{1}{2} (P-Q)(X_2-X_4)Q^T \\ &= \frac{1}{4} (P+Q)(X_1+X_3+X_2+X_4)(P+Q)^T + \\ &\quad \frac{1}{4} (P+Q)(X_1+X_3-X_2-X_4)(P-Q)^T + \\ &\quad \frac{1}{4} (P-Q)(X_1-X_3+X_2-X_4)(P+Q)^T + \\ &\quad \frac{1}{4} (P-Q)(X_1-X_3-X_2+X_4)(P-Q)^T \quad (9) \end{aligned}$$

令

$$E = (P+Q)/2, F = (P-Q)/2 \quad (10)$$

$$\begin{aligned} X_{+++} &= X_1 + X_2 + X_3 + X_4 \\ X_{--} &= X_1 - X_2 + X_3 - X_4 \\ X_{+-} &= X_1 + X_2 - X_3 - X_4 \\ X_{-+} &= X_1 - X_2 - X_3 + X_4 \end{aligned} \quad (11)$$

则

$$X_{LL} = E X_{+++} E^T + E X_{--} F^T + F X_{+-} E^T + F X_{-+} F^T$$

$$\begin{aligned} &= (E(E X_{+++})^T)^T + (F(E X_{--})^T)^T + \\ &\quad (E(F X_{+-})^T)^T + (F(F X_{-+})^T)^T \quad (12) \end{aligned}$$

由式(8)及式(10),得

$$P = E + F, Q = E - F \quad (13)$$

代入式(7),得

$$\begin{aligned} X_{LL} &= ((E+F)X_1 + (E-F)X_3)(E+F)^T + \\ &\quad ((E+F)X_2 + (E-F)X_4)(E-F)^T \\ &= (E(X_1+X_3) + F(X_1-X_3))(E+F)^T + \\ &\quad (E(X_2+X_4) + F(X_2-X_4))(E-F)^T \\ &= (A+B)E^T + (A-B)F^T \quad (14) \end{aligned}$$

其中

$$\begin{aligned} A &= E(X_1+X_3) + F(X_1-X_3) \\ B &= E(X_2+X_4) + F(X_2-X_4) \end{aligned} \quad (15)$$

式(14)给出了下采样后的 DCT 系数矩阵与下采样前拼成宏块的 4 个 8×8 DCT 系数矩阵之间的关系。离散傅里叶变换的卷积定理告诉我们:频域上的乘积等效于像素域上卷积。从 16×16 的 DCT 系数中截取其 8×8 的低频分量,相当于 DCT 域上系数矩阵与窗口矩阵的乘积,式(14)表示的 X_{LL} 与 $X_i, i=1,2,3,4$ 之间的复杂关系可以看成是像素域卷积后再变换的结果,所以将上述方法称为 DCT 域上的准卷积下采样方案。

相对于 Dugad 等人在 4 个 8×8 DCT 系数矩阵上分别截取 4×4 的低率分量,再拼凑出下采样后系数矩阵的作法而言,准卷积方案能给出更精确的结果。以下面一个极端的情况为例:假设 $X_i = \begin{pmatrix} \mathbf{0} & 80I_4 \\ 80I_4 & 80I_4 \end{pmatrix} (i=1,2,3,4)$, Dugad 方法^[3]的下采样结果是 $X_{LL} = \mathbf{0}_{8 \times 8}$ (能量为数值 0),而准卷积下采样的结果为

$$X_{LL} = \begin{pmatrix} 0 & 2 & 0 & -7 & 0 & 31 & -20 & 54 \\ 2 & -2 & 0 & 4 & 0 & -13 & -18 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 0 & -2 & -1 & 5 & 6 & 2 \\ 0 & 2 & 0 & -3 & 0 & 6 & 0 & -12 \\ 6 & 0 & 0 & 1 & 1 & -2 & -4 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -12 & 0 & 0 & -1 & -2 & 2 & 3 & 0 \end{pmatrix}$$

其中, X_{LL} 的能量为 5344。进一步,用视频 flwr 做实验,比较用准卷积和 Dugad 两种方法获得的下采样块 X_{LL} 的能量,前者大于后者的宏块约占 82%。上述结果说明了准卷积下采样方案能更好地保持高频分量。

4 算法复杂性分析

假设用式子 $mM + aA$ 表示计算中需要 m 个乘法 (M) 和 a 个加法 (A)。下面对用行列法^[7]、基方法^[8]和准卷积方法计算式(6)的计算量给予比较。

方法 1 先求出每个宏块的像素值,再给出该宏块 16×16 的 DCT 变换,直接按式(6)计算。用快速行-列方法^[7],首先计算出 4 个 8×8 的 IDCT (X_i) ($i=1,2,3,4$),需要 $12 \times 8 \times 2 \times 4M + 29 \times 8 \times 2 \times 4A = 768M + 1856A$;再计算出 1 个 16×16 的 DCT (x),需要 $32 \times 16 \times 1.5M + 81 \times 16 \times 1.5A = 768M + 1944A$;共需 $1536M + 3800A$ 。

方法 2 用基方法^[8],即利用 16×16 DCT 变换的两组正交基之间关系式,计算式(6)。

$$X = \begin{bmatrix} A_{0,0} & A_{0,1} \\ A_{1,0} & A_{1,1} \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \begin{bmatrix} A_{0,0}^T & A_{1,0}^T \\ A_{0,1}^T & A_{1,1}^T \end{bmatrix} \quad (16)$$

基方法共需要 $8 \times 4 \times 8 \times 6M + (7 \times 4 \times 8 \times 6A + 64 \times 3A) = 1536M + 1536A$ 。

基方法实质上等价于提前计算:

$$\begin{bmatrix} C_{16LL} & C_{16LH} \\ C_{16HL} & C_{16HH} \end{bmatrix} \begin{bmatrix} C_8^T & \mathbf{0} \\ \mathbf{0} & C_8^T \end{bmatrix} = \begin{bmatrix} A_{0,0} & A_{0,1} \\ A_{1,0} & A_{1,1} \end{bmatrix}$$

不难证明:

$$\begin{cases} A_{0,0}(i,j) = (-1)^{i+j} A_{0,1}(i,j) \\ A_{0,0} = \sqrt{2}P \\ A_{0,1} = \sqrt{2}Q \\ A_{1,1}(i,j) = (-1)^{i+1} A_{1,1}(i,j), i,j=0,1,\dots,7 \end{cases} \quad (17)$$

如果将矩阵的稀疏性定义为矩阵零元素的个数占总元素个数的比例,则矩阵的稀疏性越高,进行矩阵运算所需要的乘法和加法次数就越少。矩阵 P , Q 的稀疏性为 43.8%,但是若令 $E = (P + Q)/2$, $F = (P - Q)/2$,则 E, F 稀疏性就成了 71.9% (见表 1 和表 2)。这也是我们提出准卷积方法的原因。

表 1 矩阵 $E = (e_{ij})_{i,j=0}^7$ 的稀疏特性

Tab.1 The sparseness of $E = (e_{ij})_{i,j=0}^7$

j	i							
	0	1	2	3	4	5	6	7
0	0.5000	0	0	0	0	0	0	0
1	0	0.2112	0	0.0170	0	0.0050	0	0.0013
2	0	0	0	0	0	0	0	0
3	0	0.3851	0	-0.0672	0	-0.0166	0	-0.0040
4	0	0	0.5000	0	0	0	0	0
5	0	-0.1569	0	0.2834	0	0.0349	0	0.0078
6	0	0	0	0	0	0	0	0
7	0	0.1067	0	0.3515	0	-0.0763	0	-0.0138

表 2 矩阵 $F = (f_{ij})_{i,j=0}^7$ 的稀疏特性

Tab.2 The sparseness of $F = (f_{ij})_{i,j=0}^7$

j	i							
	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0.4509	0	-0.0414	0	-0.0088	0	-0.0028	0
2	0	0.5000	0	0	0	0	0	0
3	-0.1522	0	0.2696	0	0.0309	0	0.0090	0
4	0	0	0	0	0	0	0	0
5	0.0938	0	0.3592	0	-0.0750	0	-0.0179	0
6	0	0	0	0.5000	0	0	0	0
7	-0.0697	0	-0.1431	0	0.2874	0	0.0336	0

方法 3 准卷积算法,即利用式(14)计算式(6)。

计算式(14)和式(15)只需要 $((18M + 12A) \times 8 \times 2 + 64A \times 2 + 32A) \times 3 = 864M + 1056A$ 。它比方法一的计算量节省了 35%。

3 种方法的运算量如表 3 所示。

表 3 进行 DCT 域下采样时,每个像素需要的平均乘法和加法数目

Tab.3 The number of multiplications and additions per pixel for downsampling in the DCT domain

方法	乘法	加法
方法 1	1536/256 = 6	3800/256 = 14.84375
方法 2	1536/256 = 6	1764/256 = 6.5
准卷积法	864/256 = 3.375	1056/256 = 4.125

5 实验结果

对于给定的下采样方法,通常采用峰值信噪比 (最大像素值的平方与 MSE 之比)来评价该下采样图像的质量,其中, MSE 为经过上采样恢复的图像 \hat{x} 与原始图像 x 之间的均方误差。

本文方法的上采样过程为

$$\begin{bmatrix} \hat{x}_1 & \hat{x}_2 \\ \hat{x}_3 & \hat{x}_4 \end{bmatrix} = \begin{bmatrix} C_{LL}^T & C_{HL}^T \\ C_{LH}^T & C_{HH}^T \end{bmatrix} \begin{bmatrix} 2X_{LL} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \times \begin{bmatrix} C_{LL} & C_{LH} \\ C_{HL} & C_{HH} \end{bmatrix} \quad (18)$$

在仿真实验中,分别用截取 4 个 4×4 的低频率系数方法^[3]及其改进的方法^[5]、每 4 个相邻像素中任取一个的方法^[2]及本文方法,对 4 个原始视频图像进行 2:1 的下采样,其实验结果如表 4 和图 3 所示。实验的硬件执行环境如表 5 所示。

从表 4 可以看出,本文方法与文献[3]比较接近,但由于本文方法比较好地保持了图像的高频分

表4 用3种方法对4个原始图像进行2:1的下采样性能比较

Tab.4 PSNR comparison of downsampling by factor of two on four different videos using three methods

视频	分辨率	帧速率 (fps)	比特率 (bps)	帧数	PSNR(dB),时间(ms)			
					本方法	文献[3]	文献[2]	文献[5]
susi	704 × 480	29	4M	12	39.6,237	39.3,190	34.3,387	39.4,191
mobl	704 × 480	29	4M	12	25.3,246	24.9,206	20.6,527	25.0,208
tens	704 × 576	25	6M	12	25.2,270	25.1,230	20.8,701	25.1,233
flwr	704 × 576	25	4M	12	27.9,301	27.3,213	22.1,624	27.5,215

注:时间为处理时间。

表5 进行下采样运算的软硬件执行环境

Tab.5 Executing enviroment of software and hardware for downsampling

硬件		软件		视频名
内存	硬盘	操作系统	编程环境	
128MB SDRAM	ATA-66 7200 RPM	Windows 2000 Professional	Visual C++ 6.0	susi, mobil tens, flwr



(a) 文献[3]方法

(b) 本文方法

图3 视频 flwr. mpg 的下采样结果

Fig.3 flwr. mpg downsampled result

量,因此比文献[3]的PSNR平均高0.3dB。这两种方法均比文献[2]明显要好一些。另外,仔细查看图3的树枝不难看出,图3(a)较图3(b)的视觉效果要好一些。

6 结论

压缩域上的视频转码技术越来越受到重视。基于DCT域的帧图下采样方法,适用于研究DCT域上MPEG-2转为MPEG-1,而本文的下采样方案及快速算法则适用于任何基于DCT压缩的视频转码方案,并且可将准卷积下采样思想推广到DCT域任意下采样上。

参考文献 (References)

- 1 Merhav N, Bhaskaran V. Fast algorithms for DCT-domain image down sampling and for inverse motion compensation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(3):468~476.
- 2 Hu Q, Panchanathan S. Image/video spatial scalability in compressed domain[J]. IEEE Transactions on Industrial Electronics, 1998, 45(1):23~31.
- 3 Dugad R, Ahuja N. A fast scheme for image size change in the compressed domain[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(4):461~474.
- 4 Jung S H, Mitra S K, Mukherjee D. Subband DCT: definition, analysis, and applications [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(3):273~286.
- 5 Mukherjee J, Mitra S K. An image resizing in the compressed domain using subband DCT[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(7):620~627.
- 6 Yim C, Isnardi M. An efficient method for DCT-domain image resizing with mixed field/frame-mode macroblocks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(5):696~700.
- 7 Lee B G. A new algorithm to compute the discrete cosine transform [J]. IEEE Transactions on Acoustics, Speech, Signal Processing, 1984, 32(6):1243~1245.
- 8 Jiang J, Feng G. The spatial relationship of DCT coefficients between a block and its sub-blocks [J]. IEEE Transactions on Signal Processing, 2002, 50(5):1160~1169.