

基于 DCT 域的任意分辨率视频可伸缩编码

王 峥, 郑江滨

(西北工业大学 计算机学院, 陕西 西安 710129)

摘 要: 空间分辨率可伸缩编码能够满足网络视频对不同码率的要求以及不同用户显示终端的需求, 成为近几年研究热点. 由此提出一种基于 DCT 域准卷积的任意分辨率下采样方法, 该方法通过分析子块与宏块 DCT 系数之间的关系, 实现了由子块 DCT 系数预测宏块 DCT 系数, 通过低通滤波器截取预测的宏块低频 DCT 系数, 实现任意分辨率下采样. 实验证明: 该方法适合诸如 H. 264、AVS 等基于 DCT 变换的视频编码标准, 且提供灵活的空间可伸缩能力.

关键词: DCT 域; 卷积; 下采样; 多分辨率

中图分类号: TN919.8

文献标识码: A

文章编号: 1000-7180(2012)04-0058-04

Scalable Video Coding of Arbitrary Resolution in DCT Domain

WANG Zheng, ZHENG Jiang-bin

(School of Computer, Northwestern Polytechnique University, Xi'an 710129, China)

Abstract: Spatial scalable video coding technique becomes a hot problem in recent years because this kind of multi-resolution coding technique can satisfy with the requirement of online video transmitting in low bit-rate and various terminal devices displaying. In this paper, we propose a scalable video coding of arbitrary resolution in DCT domain based on convolution. Firstly, we estimate each macro-block's coefficients based on its sub-block coefficients referring to the relationship between the sub-block coefficients and the macro-block coefficients. Secondly, in order to reduce the original video resolution, a filter is designed to get the low frequency components of the estimated macro-block. Finally, these low frequency components can be saved or transmitted as one part of the current video coding stream, and a low computing complexity and an arbitrary on-demanded resolution video stream can be obtained, it is own to different size of low-pass filter and a scheme which can get macro-block DCT coefficients directly. Experimental results demonstrate that our scheme can provide flexible spatial scalable ability, and it can be used to process all kinds of the DCT-based video encoders including AVS and H. 264, etc.

Key words: DCT domain; convolution; down-sampling; multi-resolution

1 引言

随着多媒体和网络传输技术的发展, 在图像和视频的网络传输应用中, 由于用户需求不同、终端能力不同和网络传输条件的变化, 需要有不同分辨率的视频信号满足不同应用需求. 例如, 输入视频源为高分辨率, 而显示终端只能显示低分辨率, 在这种情况下, 就需要对输入的视频信号进行下采样, 得到终

端需要的分辨率后再进行编码. 近年来大量的图像下采样算法被提出. 文献[1]中 Dugad 和 Ahuja 提出了一种简单图像下采样方法, 主要利用了 DCT 变换能量的集中性. 文献[2]中 Mukherjee 和 Mitra 等人对 Dugad 方法进行的改进. 文献[1-2]中的下采样方法只能将分辨率降低到一半, 不具有灵活的可伸缩性. 文献[3-5]中实现了图像任意分辨率下采样, 其中文献[3]实现了 $2n:1$ 的下采样; 文献[4]的

采样方法适合静止图像,利用DCT域卷积,先进行 $m:1$ 上采样,再进行 $n:1$ 下采样, m 和 n 为互质的整数;文献[5]实现了长宽采样比不同的下采样,这种方法改变了长宽的比例,使下采样后的图像发生形变.针对文献[5]中的问题,文献[6-7]提出了一种新的改变图像分辨率的方法——细缝裁剪方法.该方法在一定范围内,能够实现图像分辨率变化且不失真,但超过了一定的限度,图像会发生形变,这种方法在理论方面很新颖,在实际应用中还需要进一步研究.文献[8]提出了一种快速的运动补偿算法,能够有效提高可伸缩编码的效率.

针对上述问题,本文提出一种适用于诸如H.264、AVS等视频编码标准的图像任意分辨率下采样方法.该方法通过分析子块与宏块DCT系数之间的关系,实现了由子块DCT系数预测宏块DCT系数,通过低通滤波器截取宏块低频DCT系数,实现任意分辨率下采样.实验结果表明:本文方法优于Dugad方法并且适合任何基于DCT变换的视频编码标准,能够实现任意分辨率下采样.

2 视频可伸缩编码的一般模型

视频可伸缩编码的一般过程包括:下采样、变换、量化、运动估计、运动补偿、熵编码等.在传统的视频编码过程中,应用比较多的下采样方法有Dugad方法和双线性插值bilinear方法.

2.1 Dugad 方法

Dugad方法实现了2:1下采样,该方法在编码过程中实现下采样的过程为:

- (1) 截取四个 8×8 块的低频DCT变换系数;
- (2) 分别对截取的四个低频部分进行反DCT变换;
- (3) 将得到的4个空域块重新组成一个 8×8 的子块;
- (4) 对新的子块进行DCT变换和编码.

由dugad方法的实现过程可以看出,该方法以舍去每个子块的高频部分为代价,形成新的采样块.

2.2 Bilinear 方法

Bilinear方法是空间域的采样方法,空间域插值方法容易计算.双线性插值算法放大后的图像质量较高,不会出现像素不连续的情况,然而下采样缩小图像的过程具有低通滤波器的性质,使高频分量受损,会使图像轮廓在一定程度上变模糊.

在使用Bilinear算法时,为了提高运算速率,可以使用加速手段:

- (1) 使用可分离方式,先后处理行和列
- (2) 构建缓冲区,以避免对某些行重复插值
- (3) 使用整型运算,避免浮点运算

3 DCT域准卷积下采样的视频可伸缩编码模型

本文在提出任意分辨率下采样方法的基础上,构造了一种视频可伸缩编码模型.该模型通过本文提出的采样方法,实现了任意分辨率视频可伸缩编码.图1中给出了编码器了虚线框中下采样和空域补0方法.其中下采样方法在第二部分已经给出,提出的空域补0方法能够使本文提出的下采样方法应用于基于DCT变换的编码标准.

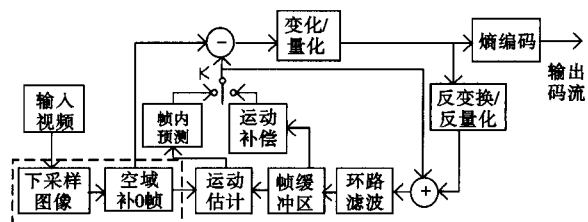


图1 视频可伸缩编码器模型

3.1 DCT域任意分辨率下采样方法

基于DCT域的准卷积下采样方法利用宏块与子块在DCT域之间的关系,由子块DCT变换系数预测宏块的DCT系数,最终在宏块中截取相应的系数实现任意分辨率可伸缩. Dugad方法截取每个子块的低频部分,经反变换后组成采样后的块,由于Dugad方法舍去了每个子块DCT系数的高频信息,所以块效应比较明显;而准卷积下采样方法能够保留更多高频信息.图2所示为任意分辨率准卷积下采样的过程.其中虚线框部分是公式推导过程,也是常规方法的操作过程,卷积方法的实现是由(1)子块DCT系数直接预测出(4)宏块DCT系数.

给定视频图像矩阵 f ,设 x 为 16×16 的宏块矩阵.式(1)是 16×16 的DCT变换矩阵,式(2)为 16×16 的DCT变换矩阵:

$$X = \text{DCT}(x) = Ax A^T \quad (1)$$

$$A_{16 \times 16} = \begin{bmatrix} A_{LL} & A_{LH} \\ A_{HL} & A_{HH} \end{bmatrix} \quad (2)$$

采样过程中使用低通滤波器,可以直接得到采样后的矩阵系数,研究过程为了计算方便,用式(3)代替低通滤波器:

$$L_{16 \times 16} = \frac{\sqrt{n}}{4} \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \quad (3)$$

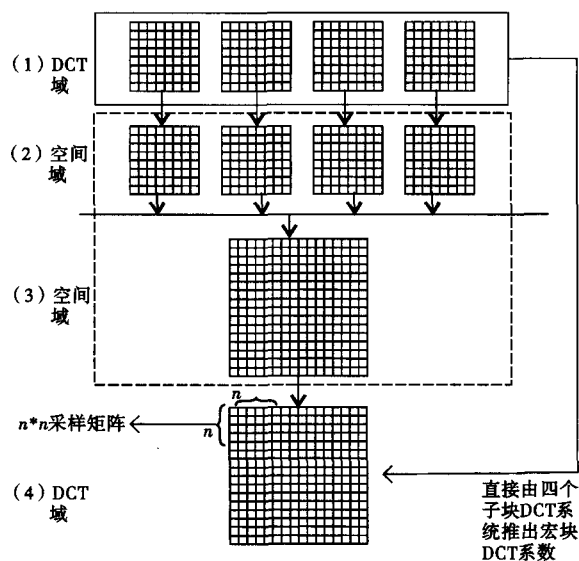


图2 任意分辨率卷积下采样

其中,当 $n \leq 8$ 时: $B_1 = \begin{bmatrix} I_{n \times n} & 0_{(8-n) \times n} \\ 0_{n \times (8-n)} & 0_{(8-n) \times (8-n)} \end{bmatrix}$, $B_2 = 0_{8 \times 8}$

当 $n > 8$ 时: $B_1 = I_{8 \times 8}$, $B_2 = \begin{bmatrix} I_{(n-8) \times (n-8)} & 0_{(16-n) \times (n-8)} \\ 0_{(n-8) \times (16-n)} & 0_{(16-n) \times (16-n)} \end{bmatrix}$ 下采样过程要用到下采样算子 T , 令 X_s 为下采样后的 DCT 系数矩阵, 式(4)为下采样计算方程, 其中 A_8 为 8×8 的 DCT 变换矩阵, 式(5)为下采样算子:

$$X_s = TXT^T \quad (4)$$

$$T_{16 \times 16} = L_{16 \times 16} A_{16 \times 16} \begin{bmatrix} A_8^T & 0 \\ 0 & A_8^T \end{bmatrix} \quad (5)$$

由式(3)~(5)可以计算出下采样方程(6)

$$X_s = \begin{bmatrix} X_{LL} & X_{LH} \\ X_{HL} & X_{HH} \end{bmatrix} = \frac{n}{16} T_{16} \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} T_{16}^T \quad (6)$$

如果对式(6)直接进行行列式计算, 相当于图2虚线框中常规方法, 计算过程有乘法运算 1 536 次和加法运算 3 800 次. 卷积方法是通过常规方法进行化简推导, 对式(6)进行化简得到式(7):

$$\begin{cases} X_{LL} = (M+N)E^T + (M-N)F^T \\ X_{LH} = (M+N)P^T + (M-N)Q^T \\ X_{HL} = (C+D)E^T + (C-D)F^T \\ X_{HH} = (C+D)P^T + (C-D)Q^T \end{cases} \quad (7)$$

其中

$$\begin{cases} E = \frac{\sqrt{n}}{8} B_1 (A_{(16 \times 16)LL} + A_{(16 \times 16)LH}) A_8^T \\ F = \frac{\sqrt{n}}{8} B_1 (A_{(16 \times 16)LL} - A_{(16 \times 16)LH}) A_8^T \\ P = \frac{\sqrt{n}}{8} B_2 (A_{(16 \times 16)HL} + A_{(16 \times 16)HH}) A_8^T \\ Q = \frac{\sqrt{n}}{8} B_2 (A_{(16 \times 16)HL} - A_{(16 \times 16)HH}) A_8^T \end{cases}$$

$$\begin{cases} M = E(X_1 + X_3) + F(X_1 - X_3) \\ N = E(X_2 + X_4) + F(X_2 - X_4) \\ C = P(X_1 + X_3) + Q(X_1 - X_3) \\ D = P(X_2 + X_4) + Q(X_2 - X_4) \end{cases} \quad (8)$$

式(7)是本文使用的采样方法, 其中计算 X_{LL} 需要乘法 864 次和加法 1 056 次, 并且 X_{LL} 、 X_{LH} 、 X_{HL} 、 X_{HH} 四个部分可以并行计算, 提高了计算效率.

3.2 空域补 0 去 0 方法

任意分辨率下采样得到的图像不一定能够按照 H. 264、AVS 等视频编码标准进行编码, 针对这个问题, 本文提出空域补 0 方法. 该方法首先将图像扩展到能够满足编码条件的最小分辨率(例如下采样后分辨率为 88×72 的图像, 要补 0 为 96×80); 然后对补 0 的图像按照编码标准进行编码. 在得到解码图像前, 需要去掉增加的尺寸部分.

4 实验结果

本文主要对 D1、CIF 和 QCIF 的视频源进行了实验, 实验过程采用本文提出的下采样方法和补 0 去 0 算法, 并对计算量、率失真度与可伸缩性进行了评价. 实验证明本文提出的方法在解码视频质量上优于 Dugad 和 Bilinear 方法, 同时提出的补 0 去 0 方法适合 H. 264、AVS 等主要的编码标准.

4.1 下采样算法计算量分析

本文通过统计采样过程中乘法和加法运算的次数来评价算法的计算量. 对于常规下采样方法, 要先求出宏块的像素值, 再计算 16×16 的 DCT 变换, 需要乘法运算 1 536 次和加法运算 3 800 次; 而本文的卷积方法, 当 $n \leq 8$ 时, 需要乘法 864 次和加法 1 056 次, 比常规方法的计算量节省了 30% 多. 在硬件实现时, 卷积方法求出子块 DCT 系数后, 推导宏块的 DCT 系数可以通过并行计算, 提高计算效率.

4.2 编码器率失真性

本文将提出的方法与 Dugad 和 Bilinear 方法进行了率失真性比较. 因 Dugad 下采样的系数为 $1/2$, 所以本文在进行率失真比较时, 首先对 foreman_CIF 分别用这三种方法进行 $1/2$ 的下采样, 然后再用相同的上采样方法恢复图像, 求出恢复后图像的信噪比 PSNR. 图 3 中给出了采样后恢复的图像以及信噪比值. 通过实验结果可以看出本文的算法优于 Dugad 和 Bilinear 方法, 且较 Dugad 方法提高了 0.2 左右的信噪比.

表 1 所示为不同下采样系数、不同分辨率视频源下三种方法率失真度的比较, 对于非 $1/2$ 采样, 没

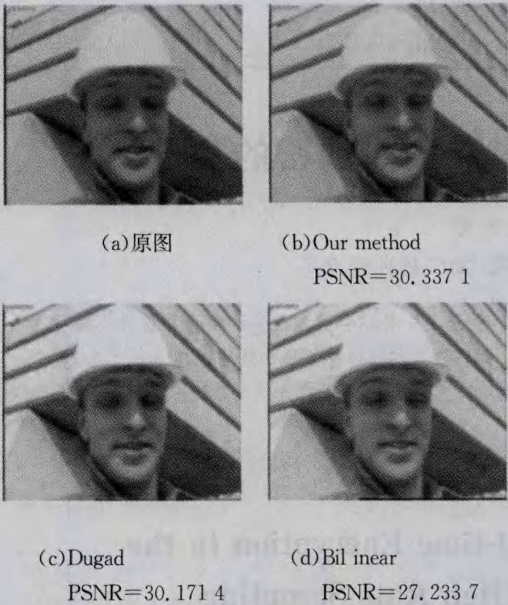


图 3 有下采样系数为 1/2 时 foreman_cif 的恢复图像

对 Dugad 方法进行. 由表 1 可以看出, 对于任意分辨率恢复出的图像, 本文方法优于其他两种方法.

4.3 编码器可伸缩性

图 4 中显示了对视频进行下采样解码的一个例子. 这里下采样系数分别为 4/16、8/16、11/16. 可见, 本文所提出的编码模型能够实现任意分辨率的下采样, 与传统编码器相比较, 具有灵活的可伸缩性.

5 结束语

本文提出了一种 DCT 域任意分辨率下采样方法, 并将该方法应用与编码模型中. 在编码过程中, 应用空域补 0 去 0 算法, 使本文提出的下采样方法能够应用于任何基于 DCT 变换的编码标准. 本文对提出的下采样方法与 Dugad 和 Bilinear 等方法进行了大量实验比较, 实验结果表明: 本文提出的方法在质量和灵活的可伸缩性方面都具有优势, 适合在低码率传输和移动终端业务中应用.

表 1 率失真度比较

视频源	下采样方法	PSNR		
		4/16(采样系数)	8/16(采样系数)	11/16(采样系数)
Walk_pal 720×576	Bilinear	23.955 6	27.233 7	31.373 7
	Dugad	无	31.713 7	无
	Our method	25.874 9	31.719 1	34.190 1
Foreman_cif 352×288	Bilinear	23.706 4	25.657 2	28.651 4
	Dugad	无	30.171 4	无
	Our method	24.996 6	30.337 1	31.757 7

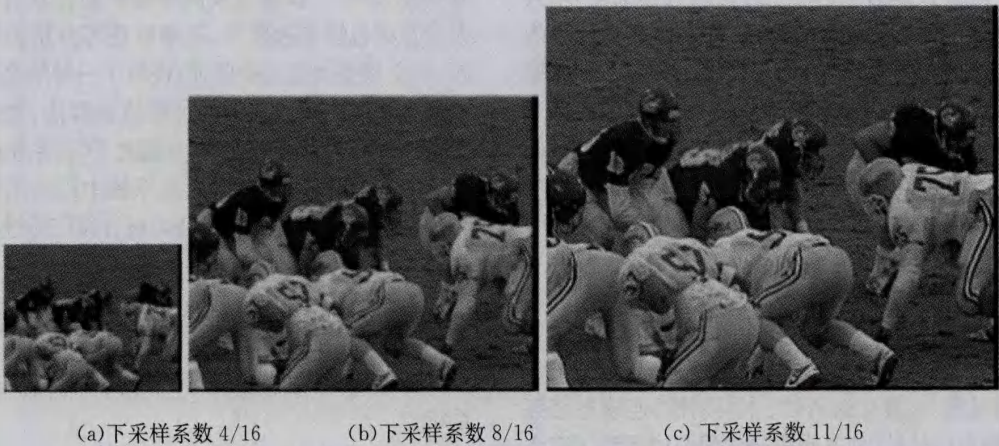
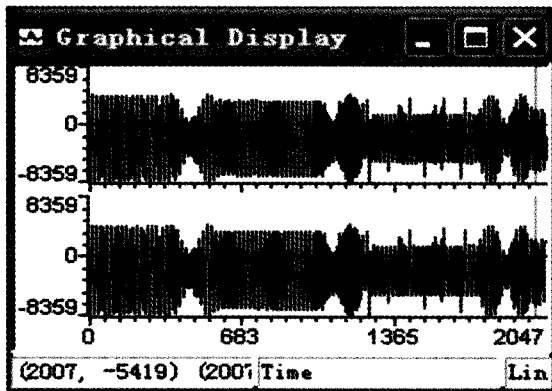
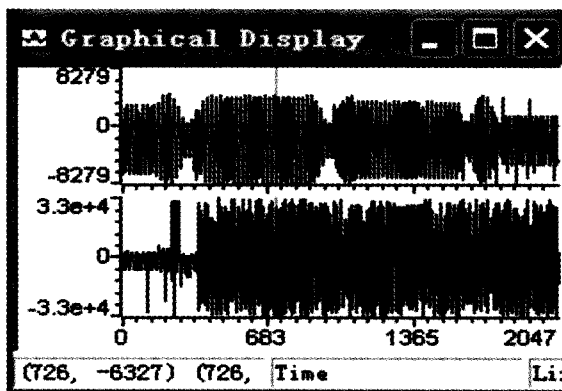


图 4 准卷积下采样方法对 football(352×288) 在不同分辨率下的恢复图像

(下转第 66 页)



(a)原始音频(上)和 c 取 30 的解密波形(下)



(b)原始音频(上)和 c 取 30.001 的解密图(下)

图 9 密钥敏感性测试

参考文献:

- [1] 刘运毅,覃团发,倪皖荪,等. 基于 G. 729A 和 AES 的分组语音保密通信[J]. 南京大学学报:自然科学版, 2005,41(4):1-10.
- [2] 潘桐. G. 726 语音压缩算法在低码率下的改进与实现[J]. 微计算机应用, 2009,30(4):76-78.
- [3] 游鹏,季晓勇,卢鹏. 对 G. 723.1 高码率编码语音中进行压缩域信息隐藏的性能分析[J]. 计算机科学, 2008,35(5):194-195.
- [4] 罗启彬,张健. 混沌密钥序列的 DSP 实现算法研究[J]. 微电子学与计算机, 2009,26(2):44-46.
- [5] 盛苏英,陆国平,吴新华. 混沌语音实时保密通信研究[J]. 微电子学与计算机, 2008,25(3):84-86.
- [6] Annie Lang. Information processing perspective defining audio/video redundancy from a limited-capacity[J]. Communication Research, 1995(22): 86
- [7] Zhou Xiaobing, Zhang Jinfu, Jiang Murong. Passivity-based adaptive chaos control of a new 3D chaotic system with uncertain parameter[C]// Information Technology and Applications, 2009. IFITA '09. International Forum on. 2009(3):57-60.
- [8] 高颖. 定点 DSP 中运算精度的提高[J]. 电子工程师, 2001,27(9):3-4.

(上接第 61 页)

参考文献:

- [1] Dugad R, Ahuja N. A fast scheme for image size change in the compressed domain[J]. IEEE Transactions on Circuits System Video Technology, 2001, 11(4): 461-474.
- [2] Mukherjee J, Mitra S K. Image resizing in the compressed domain using subband DCT[J]. IEEE Transactions on Circuits System Video Technology, 2002, 12(7): 620-627.
- [3] Shu H T, Chau L P. An efficient arbitrary downsizing algorithm for video transcoding[J]. IEEE Transactions on Circuits System Video Technology, 2004, 14(6): 887-891.
- [4] Park H W, Park Y S, Oh S K. L/M-image folding in block DCT domain using symmetric convolution [J]. IEEE Trans. Image Process, 2003, 12(9): 1016-1034.
- [5] Tan Eeleng, Gan Woonseng, Wong Mengtong. Fast image resizing in discrete cosine transform domain with relationship between DCT block and its sub-blocks[C] // 2010 International Conference, Audio Language and Image Processing (ICALIP). China: Shanghai, 2010: 1683-1687.
- [6] Avidan S, Shamir A. Seam carving for content-aware image resizing[J]. ACM Transactions on Graphics, 2007,26(3):10.
- [7] Han J W, Choi K S, Wang T S, et al. Wavelet based seam carving for content-aware image resizing[C]// Proc. of IEEE International Conference on Image Processing. Cairo, Egypt, 2009: 345-348.
- [8] 史海滨,朱光喜,喻莉,等. 一种基于预测频度的 AVS/H. 264 的快速 4x4 帧内预测模式选择算法[J]. 微电子学与计算机, 2007,24(8):8-10.

作者简介:

王 峰 女, (1984—), 硕士研究生. 研究方向为多媒体智能信息处理.

郑江滨 男, (1971—), 教授. 研究方向为语音图像处理与计算机视觉.