

基于内容统计的彩色视频 4D-MDCT 编码

杜相文 陈贺新

(吉林大学通信工程学院测控技术与仪器系, 长春 130025)

摘要 为了提高彩色视频序列的压缩比, 根据四维矩阵及四维矩阵离散余弦变换(4D-MDCT)理论, 提出了一种基于内容统计的编码方法, 即先对彩色视频序列所构成的四维矩阵进行子阵划分和四维矩阵离散余弦变换, 然后对变换系数进行四维矩阵量化。为了能够充分利用彩色视频序列的帧间相关性和帧内相关性, 又通过对量化后的四维子阵系数进行统计来找出其中完全相同的四维子矩阵。对于系数完全相同的子矩阵, 只对第1个子矩阵进行编码, 而不对其他的子矩阵进行编码, 而且在对第1个四维子矩阵进行编码时, 采用了传统的游程编码和熵编码。实验结果表明, 在保持高信噪比的情况下, 此方法可以大幅度地提高压缩比。

关键词 彩色视频 四维矩阵 矩阵量化 内容统计

中图分类号: TN919.8 **文献标识码**: A **文章编号**: 1006-8961(2004)05-0616-04

Color Video 4D-MDCT Coding based on Content Statistics

DU Xiang-wen, CHEN He-xin

(Department of Survey and control Technology and Equipment, College of Communication Engineering, Jilin University, Changchun 130025)

Abstract This paper presents a new coding method based on content statistics according to four-dimensional matrix and four-dimensional matrix DCT (4D-MDCT) theories in order to increase the compression ratio of color video sequence. First, four-dimensional matrix, which is constructed by color video sequence, is divided into submatrices and 4D-MDCT is performed to the submatrices. Then four-dimensional matrix quantization is applied to the transformed coefficients. In order to use the correlation between the inter color frames and intra color frames, statistics on quantized four-dimensional submatrix coefficients is performed to search four-dimensional submatrices which have identical quantized coefficients. Only the first one of these identical quantized four-dimensional submatrices is coded. When the first quantized four-dimension submatrix is coded, conventional runlength coding and entropy coding are performed. Experimental results proved that the compression ratio can be increased greatly and the high PSNR is preserved. Because conventional runlength coding and entropy coding are simple, and easy to be realized by using hardware, the method in this paper has a prosperous future.

Keywords image processing, color video, four-dimensional matrix, matrix quantization, content statistics

1 引言

随着计算机技术和网络技术的发展, 各种视频业务, 如低码率视频会议、视频电话、远程低码率多媒体协同工作和 VOD(视频点播)等也正在蓬勃发展。由于在人们所获得的信息中有大约 70% 的信息来自视觉信息, 因此网络资源中的图片信息和视频信息受到了越来越多的关注, 但由于有限的网络带宽极大限制了图片信息和视频信息在网络上的自由

传输, 所以, 如何进一步压缩视频信息来突破网络带宽这一瓶颈已成为全球关注的焦点。

目前, 广播电视网络, Internet 网络以及电信网络和对应终端设备上处理视频数字信号的压缩编码方法采用的标准主要有 MPEG^[1]、H. 261^[2] 和 H. 263^[3] 等。这些标准首先将 R、G、B 3 帧图像转换成 Y、Cr、Cb 数据, 然后将灰度图像的压缩编码方法分别独立地应用于 Y、Cr、Cb 各分量上, 但这些方法都是利用图像的灰度等级进行压缩的, 其压缩比都不是很高, 且信噪比也比较低。正如人们所知, 由于

基金项目: 国家自然科学基金资助国际合作与交流项目(60211130507); 国家自然科学基金项目(60172046); 吉林科技发展计划项目(20020323-1)

收稿日期: 2003-07-24; **改回日期**: 2003-12-19

彩色图像的 R、G、B 3 帧是对同一个物理模型的统一反映,其不仅具有相同的纹理、边缘和灰度变化梯度,且每一帧都几乎可以反映除颜色外的全部信息,因此具有很强的视觉相似性。人类的视觉特征也表明,彩色图像各个分量之间的依赖关系是非线性的。可见,如果分别对视频序列中的 Y、Cr 和 Cb 各分量进行压缩编码,将不能充分利用彩色图像中的各分量间潜在的相关性,这就限制了压缩比和峰值信噪比等性能的进一步提高。

为了能够充分利用多光谱图像在光谱间的相关性,文献[4]提出并应用了彩色图像的三维矩阵模型,并得到了很好的压缩效果;在此基础上,文献[5]提出的彩色视频序列四维矩阵模型也得到了很高的压缩比和信噪比。本文以文献[5]的四维矩阵理论为基础,采用了矩阵量化和内容统计的方法,又进一步提高了彩色视频序列的压缩比。

2 四维矩阵理论及其离散余弦变换

2.1 四维矩阵的基本定义

假定 I_1, I_2, I_3, I_4 是整数, $I_j \in \mathbf{I}, (j=1, 2, 3, 4)$, 且 $a_{i_1, i_2, i_3, i_4} (1 \leq i_1 \leq I_1, 1 \leq i_2 \leq I_2, 1 \leq i_3 \leq I_3, 1 \leq i_4 \leq I_4)$ 是一个实数,即 $a_{i_1, i_2, i_3, i_4} \in \mathbf{R}$, 那么四维矩阵定义成一个以 a_{i_1, i_2, i_3, i_4} 为元素的四维数的排列,即

$$A = [a_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4} \quad (1)$$

2.2 四维矩阵点乘法(四维矩阵与二维矩阵的乘法)

(1) 第 1 类点乘法

给定任意四维矩阵 $A = [a_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4}$ 和二维矩阵 $B = [b_{p, q}]_{P \times Q}$, 其第 1 类点乘法表示为

$$(A \cdot B)_1 = C = [c_{k_1, k_2, k_3, k_4}]_{K_1 \times K_2 \times K_3 \times K_4} \quad (2)$$

其中, $c_{k_1, k_2, k_3, k_4} = \sum_{e=1}^{I_4} a_{e, i_2, i_3, i_4} \cdot b_{e, q}, I_1 = P, k_1 = q, K_1 = Q, k_n = i_n, K_n = I_n (n=2, 3, 4)$ 。这里 p, q, e 是下标变量, P 和 Q 是常量。

(2) 第 x 类点乘法

一般可以给出如下各类点乘法的通式:

给定任意四维矩阵 $A = [a_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4}$ 和二维矩阵 $B = [b_{p, q}]_{P \times Q}$, 其第 x 类点乘法表示为

$$(A \cdot B)_x = C = [c_{k_1, k_2, k_3, k_4}]_{K_1 \times K_2 \times K_3 \times K_4} \quad (3)$$

其中, $c_{k_1, \dots, k_x, \dots, k_4} = \sum_{e=1}^{I_x} a_{i_1, \dots, e, \dots, i_4} \cdot b_{e, q}, I_x = P, k_x = q, K_x = Q, k_n = i_n, K_n = I_n (1 \leq n \leq 4, n \neq x)$ 。

2.3 四维矩阵理论的阵乘法

给定任意两个四维矩阵 $A = [a_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4}$ 和 $B = [b_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4}$, 则 A 与 B 的阵乘法表示为

$$C = [c_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4} = A \times B \quad (4)$$

其中, $c_{i_1, i_2, i_3, i_4} = a_{i_1, i_2, i_3, i_4} \times b_{i_1, i_2, i_3, i_4}$ 。

同样,可以给出如式(5)所示的矩阵除法表达式

$$A = [a_{i_1, i_2, i_3, i_4}]_{I_1 \times I_2 \times I_3 \times I_4} = \frac{C}{B} \quad (5)$$

其中, $a_{i_1, i_2, i_3, i_4} = c_{i_1, i_2, i_3, i_4} / b_{i_1, i_2, i_3, i_4}, \forall b_{i_1, i_2, i_3, i_4} \neq 0$ 。

关于四维矩阵的其他运算准则与性质请参阅参考文献[5]。

2.4 四维矩阵离散余弦变换及其逆变换

为了保持与 MPEG 压缩标准一致以及计算的方便、快速,在对四维矩阵进行变换时,本文采用如下 DCT 二维矩阵作为变换核:

$$B_M = [b_{u, v}]_{M \times M} = \begin{cases} \sqrt{\frac{2}{M}} \cos \frac{u\pi(2v+1)}{2M}, & u \neq 0 \\ \frac{1}{\sqrt{M}}, & u = 0 \end{cases} \quad (6)$$

其中, M 为变换核矩阵的阶数, u 和 v 是变换核矩阵元素的下标。

四维矩阵离散余弦变换(4D-MDCT)的正向变换为

$$C_{\text{Sub}} = (((A_{\text{Sub}} B_8)_1)_2 B_3)_3 B_4 \quad (7)$$

四维矩阵离散余弦变换的逆向变换为:

$$A_{\text{Sub}} = (((C_{\text{Sub}} B_8^{-1})_1)_2 B_3^{-1})_3 B_4^{-1} = (((C_{\text{Sub}} B_8^T)_1)_2 B_3^T)_3 B_4^T \quad (8)$$

其中, A_{Sub} 表示对四维矩阵 A 进行分割后的子矩阵, C_{Sub} 表示 A_{Sub} 经过四维矩阵离散余弦变换后所形成的新的子矩阵,此处,下标 Sub 表明该矩阵是分割后的子矩阵。

3 基于内容统计的彩色视频 4D-MDCT 编码

3.1 四维矩阵模型的建立

一段彩色视频序列可以用四维矩阵来描述。若分别用 x, y 分量来表示彩色视频序列中图像的高度分量和宽度分量,用 z 分量表示图像的光谱分量(即图像的 R、G、B 三基色),用 t 分量来表示彩色视频序列在时间上的分量,则一段彩色视频序列的四维矩阵模型为

$$A = [a_{x, y, z, t}]_{I_1 \times I_2 \times I_3 \times I_4} \quad (9)$$

其中, $1 \leq x \leq I_1, 1 \leq y \leq I_2, 1 \leq z \leq I_3, 1 \leq t \leq I_4, I_1$ 是视频序列中图像的高度, I_2 是图像的宽度, I_3 是彩色图像的光谱数(这里 $I_3=3$), I_4 是一段彩色视频序列的帧数。这种四维矩阵模型所描述的一段彩色视频序列如图 1 所示。

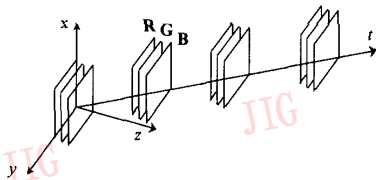


图 1 四维矩阵模型表示的彩色视频序列
(x 和 y 为空间分量, z 为光谱分量, t 为时间分量)

3.2 编码系统

基于内容统计的彩色视频 4D-MDCT 编码原理框图如图 2 所示。



图 2 基于内容统计的彩色视频 4D-MDCT 编码原理框图

3.2.1 子阵划分

将彩色视频序列所构成的四维矩阵进行子阵划分, 即可得到互不重叠的 $8 \times 8 \times 3 \times 8$ 的四维子矩阵。

3.2.2 4D-MDCT

对四维子矩阵进行如式(7)所示的 4D-MDCT。四维子矩阵经过变换后, 四维矩阵模型在各分量方向上的相关性得以减弱, 这样四维矩阵模型中的能量分布更加集中。四维子矩阵的 4 个下标构成了一个一维矢量, 记作

$$\xi = (x, y, z, t)^T \quad (10)$$

则 ξ 到 $(0, 0, 0, 0)^T$ 矢量的距离为

$$\|\xi\| = \sqrt{(\xi, \xi)} \quad (11)$$

变换后的低频分量分布在 $\|\xi\|$ 较小的元素上, 高频分量分布在 $\|\xi\|$ 较大的元素上, 而其能量主要集中在低频分量上。

3.2.3 四维矩阵量化

在量化之前, 首先需要构造四维量化矩阵。由于人眼对低频分量的变化敏感, 对高频分量的变化不敏感, 所以在构造四维量化矩阵时, 应赋予 $\|\xi\|$ 值较小的量化矩阵元素以较小的量化值, 而赋予 $\|\xi\|$ 值较大的量化矩阵元素以较大的量化值。若将四维量化矩阵记为 T , 将量化后的四维子矩阵记为 D , 则量化公式为

$$D = \frac{C_{\text{Sub}}}{T} \quad (12)$$

3.2.4 内容统计

为了充分利用彩色视频序列帧间和帧内的相关性, 可对量化后的四维子矩阵 D 进行内容统计, 统计步骤如下:

(1) 统计当前四维子矩阵 D 和与其处于同一空间位置并且在时间维上处于前一位的四维子矩阵的系数, 然后看两四维子阵是否完全相同; 若完全相同, 则标记该四维子矩阵;

(2) 统计当前四维子矩阵 D 和与其处于同一时间位置并且在空间上处于 D 之前的所有四维子矩阵的系数, 然后看是否存在与其系数完全相同的四维子矩阵, 若存在, 则标志该四维子矩阵, 并记录第 1 个雷同四维子矩阵的位置。

3.2.5 编码

对于四维子矩阵中的 DC 分量(即与下标矢量 $(1, 1, 1, 1)^T$ 对应的元素)先进行 DPCM 编码, 然后再进行熵编码; 而对于四维子矩阵中的扫描排序后的 AC 分量则需先进行游程编码; 再进行熵编码; 而对于已标记过的四维子矩阵则不进行编码, 只需将标记信息和位置信息写入输出码流即可。

4 实验结果

为验证本文方法的效果, 对彩色视频序列 Miss America 和 Salesman 的第 1~48 帧进行了编解码实验。

在 Miss America 的压缩比为 187.024 0 倍, Salesman 的压缩比 87.294 0 倍时, 两彩色视频序列中 Y 帧的峰值信噪比如图 3 所示。

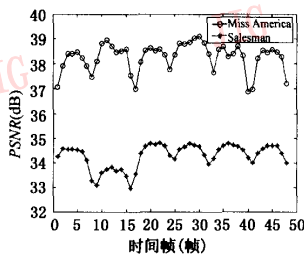


图 3 Miss America 与 Salesman 解码后各帧的峰值信噪比

由于 Salesman 序列的纹理、色彩等信息比较复杂, 所以信噪比要略低一些。

表 1 给出了 Miss America 和 Salesman 在量化矩阵不同的情况下的压缩比与平均峰值信噪比, 并与基于运动补偿的二维 DCT(MC/2D-DCT)方法进

表 1 基于内容统计的彩色视频 4D-MDCT 编码与 MC/2D-DCT 编码的实验结果比较

		Miss America			Salesman		
本文方法	压缩比(CR)(倍)	187.0240	227.3212	355.7185	87.2940	117.665	146.7982
	$PSNR_Y$ (dB)	38.2774	37.5543	35.9610	34.2977	33.0474	32.0980
	$PSNR_U$ (dB)	40.7462	40.3861	39.5922	44.0339	43.2400	42.6534
	$PSNR_V$ (dB)	40.9224	40.1309	38.3788	44.4770	43.7048	43.0546
MC/2D-DCT	压缩比(CR)(倍)		138.2693			110.3806	
	$PSNR_Y$ (dB)		35.2667			28.0533	
	$PSNR_U$ (dB)		37.6067			36.4800	
	$PSNR_V$ (dB)		35.1267			37.1133	

行了对比。从表中数据可以看出,与 MC/2D-DCT 方法相比,本文方法所得到的压缩比和信噪比都有大幅度的提高,其中, Miss America 在信噪比有所提高的情况下,压缩比提高了 1.57 倍。图 4 给出了

Miss America 和 Salesman 的原图像和解码后的恢复图像。从恢复图像中可以看出,在高压缩比的情况下,图像质量依然很好。Miss America 图像只是在压缩比达到 355.7185 倍时,才有轻微的方块效应。



图 4 Miss America 和 Salesman 第 28 帧原图像与恢复图像

5 结 论

本文根据四维矩阵理论,应用 4D-MDCT、四维矩阵量化和内容统计对彩色视频进行了编码,在保证图像质量的前提下,获得了很高的压缩比。由于 4D-MDCT 充分考虑了彩色视频序列在空间、光谱和时间上的各种相关性,因此将其作为一个整体来进行描述和变换,可有效地去除各种相关性;而内容统计的方法则进一步利用了帧内和帧间的相关性。由于本方法编码简单、快速,易于硬件实现,因此具有很好的应用前景。

参 考 文 献

- ISO/IEC. Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbit/s-part 2: Video[S]. IS 11172-2, ISO/IEC, 1993.
- ITU-T. Video codec for audio visual services at $p \times 64$ kbits/s

- [S]. Recommendation H. 261, ITU-T, 1990.
- ITU-T. Video Coding for low bitrate communication[S]. Draft Recommendation H. 263 Version 2, ITU-T, 1998.
- 桑爱军, 陈贺新. 三维矩阵彩色图象 WDCT 压缩编码[J]. 电子学报, 2002, 30(4): 594~597.
- 赵岩, 陈贺新. 彩色视频的三维矩阵离散余弦变换编码[J]. 中国图象图形学报, 2003, 8A(6): 620~624.



杜相文 1975 生, 2001 年获吉林大学硕士学位, 同年攻读吉林大学通信与信息系统专业博士学位。主要研究方向为计算机视觉、图像编码、多维矩阵理论等。



陈贺新 1949 生, 工学博士, 吉林大学通信工程学院教授, 博士生导师。主要研究兴趣是图像处理、多维数字信号处理、人工神经网络、计算机视觉等。