

彩色视频的四维矩阵离散余弦变换编码

赵 岩

(吉林大学通信工程学院通信工程系, 长春 130012)

陈贺新

(吉林大学通信工程学院测控与通信仪器系, 长春 130025)

摘 要 为了在高信噪比条件下获得对彩色视频的高倍压缩, 提出了 4 维矩阵及 4 维矩阵离散余弦变换理论, 并将该理论应用于彩色视频编码。基于 4 维矩阵离散余弦变换的彩色视频编码方法将彩色视频的多个帧放在一个统一的数学模型(即, 4 维矩阵)中考虑, 利用 4 维矩阵离散余弦变换去除其间的各种相关性, 并通过矢量量化对变换系数进行压缩编码。其方法可以同时全面利用彩色视频图象相邻像素间、彩色空间 Y、U、V 3 分量间, 以及视频图象相邻帧之间的相关性。实验结果证明, 对可视电话和视频会议等应用中的彩色视频序列图象, 利用该方法可在高信噪比条件下获得较高压缩比的编码效果。

关键词 计算机图象处理(520·6040) 彩色视频 4 维矩阵 离散余弦变换 视频编码

中图法分类号: TN919.8 **文献标识码:** A **文章编号:** 1006-8961(2003)06-0620-05

Color Video Coding via Four-dimensional Matrix DCT

ZHAO Yan

(Department of Communication Engineering, College of Communication Engineering, Jilin University, Changchun 130012)

CHEN He-xin

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

Abstract Four-dimensional (4D) matrix and four-dimensional matrix DCT (4D-MDCT) theory are proposed and applied in color video coding, in order to obtain high compression for color video under high signal-to-noise ratio. The four-dimensional matrix (4D Matrix) theory is a new and useful theory for color video processing. The color video coding method based on 4D-MDCT put several frames of color video into a unified mathematical model (i. e. , 4D matrix), and various correlations in color video are eliminated utilizing 4D-MDCT. The four dimensions of 4D-MDCT used in the proposed algorithm are image width (row component), image height (column component), three channels (Y, U and V) in a color image (vertical component), and adjacent three frames of the video sequence (depth component) respectively. Similar to the method of 2D-DCT, the whole three frames of a color video sequence are also divided into "4D-blocks", that is 4D sub-matrices, and the order of each sub-matrix is $3 \times 3 \times 3$. The transform coefficients are coded by vector quantization further. The method can fully exploit correlations between adjacent pixels of an image, YUV frames of color space, and adjacent frames of video at the same time. Experimental results has proved that for the color video in applications such as videophone and video conferencing, the proposed method can achieve higher compression ratio under high signal-to-noise ratio.

Keywords Computer image processing, Color video, 4D matrix, DCT, Video coding

0 引 言

随着网络带宽的迅速增加和计算机能力的不断

增强, Internet 上的视频流即将经历从其作为一种新鲜事物到成为主流媒体通信的转变。视频流技术作为目前 Internet 上所有视频应用中最成功的应用, 已经渗透到每个人的日常生活和工作中。除了阅

基金项目: 国家自然科学基金项目(60172046); 国家自然科学基金重大国际合作项目; 吉林省科技发展计划项目(20020323-1)

收稿日期: 2002-08-05; 改回日期: 2002-11-18

读电子邮件和浏览网页外,越来越多的人将成为网络上视频内容浏览的用户.在 Internet 上发展视频流业务的主要瓶颈即为传输视频所需的带宽,而 Internet 的本质特征限制了对其带宽的有效利用.无疑,Internet 上的视频流面临着许多技术和商业挑战,它需要全新的视频编码器、协议(标准)、播放器,以及系统.

目前,国际上建议并在相关产品中广泛使用的视频压缩编码标准,如 H. 261、H. 263、MPEG-1(运动图象专家组)、MPEG-2 等^[1-4],以及 20 世纪 90 年代发展起来的一些基于分形理论、小波分析理论等的新编码方法,大多是针对灰度图象的.而对于彩色图象和视频的压缩编码,一般都是先将 R、G、B 数据转换成 Y、U、V 数据,然后将灰度图象的压缩编码方法或其改进方法独立地应用于各个分量上.而实际上,R、G、B 各颜色分量间的相关性很强,转换后生成的亮度及色度各个分量间的相关性也很强.正如人们所知,彩色图象的 3 个分量是对同一个物理模型的统一反映,具有相同的纹理、边缘和灰度变化梯度,每一分量都几乎可以反映除颜色外的全部信息,各分量之间具有很强的视觉相似性.人类的视觉特性表明,彩色图象各个分量之间的依赖关系是非线性的.可是,如果分别对 Y、U、V 各分量进行压缩编码,就不能充分利用彩色图象中各分量间潜

在的相关性,也就是不能去除这部分视觉冗余,这就限制了压缩比和峰值信噪比等性能的进一步提高.事实也证明这些方法很难达到高信噪比条件下的高压缩比要求.

文献[5]运用三维矩阵变换方法将彩色图象的 Y、U、V 3 分量统一在一个模型中进行编码,获得了优于 JPEG(联合图象专家组)标准方法的信噪比和压缩比性能.本文则采用全新的数学模型——4 维矩阵来表示彩色视频,将彩色视频表示在一个统一的数学模型中,并利用新的变换方法——4 维矩阵离散余弦变换,如同魔方游戏运算规则一样把对彩色视频多个帧的处理统一起来,转化为对空间 4 维变换的数学处理,全面去除彩色视频中各灰度图象的像素之间,彩色空间的 YUV 3 分量之间,以及视频的连续帧之间的相关性,从而实现高信噪比条件下的高倍压缩.

1 4 维矩阵及其离散余弦变换

1.1 4 维矩阵基本定义

由 $I \times J \times K \times L$ 个实数 a_{ijkl} ($1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K, 1 \leq l \leq L$) 构成的 I 个横行, J 个竖列, K 个纵序和 L 个深度的超立方体形数据排列,称为 $I \times J \times K \times L$ 阶 4 维矩阵,记为

$$A_{I \times J \times K \times L} = [a_{ijkl}]_{I \times J \times K \times L} = \begin{bmatrix} \begin{bmatrix} a_{1111} & a_{1211} & \dots & a_{1j11} \\ a_{2111} & a_{2211} & \dots & a_{2j11} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i111} & a_{i211} & \dots & a_{ij11} \end{bmatrix} & \dots & \begin{bmatrix} a_{11k1} & a_{12k1} & \dots & a_{1jk1} \\ a_{21k1} & a_{22k1} & \dots & a_{2jk1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1k1} & a_{i2k1} & \dots & a_{ijk1} \end{bmatrix} \\ \vdots & \ddots & \vdots & \vdots \\ \begin{bmatrix} a_{11l1} & a_{12l1} & \dots & a_{1jl1} \\ a_{21l1} & a_{22l1} & \dots & a_{2jl1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1l1} & a_{i2l1} & \dots & a_{ijl1} \end{bmatrix} & \dots & \begin{bmatrix} a_{11k1} & a_{12k1} & \dots & a_{1jk1} \\ a_{21k1} & a_{22k1} & \dots & a_{2jk1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1k1} & a_{i2k1} & \dots & a_{ijk1} \end{bmatrix} \end{bmatrix}$$

其中, a_{ijkl} 称为 4 维矩阵 $A_{I \times J \times K \times L}$ 的元素.

1.2 4 维子矩阵定义

对任何一个 4 维矩阵 $A_{I \times J \times K \times L}$ 都可以用一些横线,竖线,纵线和深线划分成小尺寸的 4 维子矩阵,这样的小 4 维子矩阵称为 4 维矩阵 $A_{I \times J \times K \times L}$ 的 4 维子矩阵 $A_{U \times V \times W \times S}$,其中, $U < I, V < J, W < K, S < L$.

1.3 4 维矩阵的运算准则

(1) 相等

指两个同阶的 4 维矩阵 $A_{I \times J \times K \times L} = [a_{ijkl}]_{I \times J \times K \times L}$ 和 $B_{I \times J \times K \times L} = [b_{ijkl}]_{I \times J \times K \times L}$,有

$$a_{ijkl} = b_{ijkl} \quad (1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K, 1 \leq l \leq L)$$

(2) 加法

指两个同阶的 4 维矩阵 $A_{I \times J \times K \times L} = [a_{ijkl}]_{I \times J \times K \times L}$ 和 $B_{I \times J \times K \times L} = [b_{ijkl}]_{I \times J \times K \times L}$ 相加:
 $A_{I \times J \times K \times L} + B_{I \times J \times K \times L} = [a_{ijkl} + b_{ijkl}]_{I \times J \times K \times L} \quad (1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K, 1 \leq l \leq L)$

(3) 乘法

指两个 4 维矩阵 $A_{I_1 \times J_1 \times K_1 \times L_1} = [a_{ijkl}]_{I_1 \times J_1 \times K_1 \times L_1}$ 和 $B_{I_2 \times J_2 \times K_2 \times L_2} = [b_{ijkl}]_{I_2 \times J_2 \times K_2 \times L_2}$ 的 6 类乘法:

① 行列乘 (I 类乘法)

$$(A_{I_1 \times J_1 \times K_1 \times L_1} \cdot B_{I_2 \times J_2 \times K_2 \times L_2})_I = [c_{ijkl}]_{I_3 \times J_3 \times K_3 \times L_3}$$

其中, $K_3 = K_2 = K_1, I_3 = I_2 = L_1, J_1 = I_2, I_3 = I_1, J_3 = J_2$, 并且 $c_{ijkl} = \sum_{e=1}^I a_{iek} \cdot b_{ejl} (i = 1, 2, \dots, I_3, j = 1, 2, \dots, J_3, k = 1, 2, \dots, K_3, l = 1, 2, \dots, L_3)$.

② 行纵乘 (II 类乘法)

$$(A_{I_1 \times J_1 \times K_1 \times L_1} \cdot B_{I_2 \times J_2 \times K_2 \times L_2})_I = [c_{ijkl}]_{I_3 \times J_3 \times K_3 \times L_3}$$

其中, $J_3 = J_2 = J_1, L_3 = L_2 = L_1, K_1 = I_2, I_3 = I_1, K_3 = K_2$, 并且 $c_{ijkl} = \sum_{e=1}^I a_{ije} \cdot b_{ekl} (i = 1, 2, \dots, I_3, j = 1, 2, \dots, J_3, k = 1, 2, \dots, K_3, l = 1, 2, \dots, L_3)$.

③ 行高乘 (III 类乘法)

$$(A_{I_1 \times J_1 \times K_1 \times L_1} \cdot B_{I_2 \times J_2 \times K_2 \times L_2})_{II} = [c_{ijkl}]_{I_3 \times J_3 \times K_3 \times L_3}$$

其中, $J_3 = J_2 = J_1, K_3 = K_2 = K_1, L_1 = I_2, I_3 = I_1, L_3 = I_2$, 并且 $c_{ijkl} = \sum_{e=1}^I a_{ike} \cdot b_{ejl} (i = 1, 2, \dots, I_3, j = 1, 2, \dots, J_3, k = 1, 2, \dots, K_3, l = 1, 2, \dots, L_3)$.

④ 列纵乘 (IV 类乘法)

$$(A_{I_1 \times J_1 \times K_1 \times L_1} \cdot B_{I_2 \times J_2 \times K_2 \times L_2})_{II} = [c_{ijkl}]_{I_3 \times J_3 \times K_3 \times L_3}$$

其中, $I_3 = I_2 = I_1, L_3 = L_2 = L_1, K_1 = J_2, J_3 = J_1, K_3 = K_2$, 并且 $c_{ijkl} = \sum_{e=1}^I a_{ije} \cdot b_{ekl} (i = 1, 2, \dots, I_3, j = 1, 2, \dots, J_3, k = 1, 2, \dots, K_3, l = 1, 2, \dots, L_3)$.

⑤ 列高乘 (V 类乘法)

$$(A_{I_1 \times J_1 \times K_1 \times L_1} \cdot B_{I_2 \times J_2 \times K_2 \times L_2})_V = [c_{ijkl}]_{I_3 \times J_3 \times K_3 \times L_3}$$

其中, $I_3 = I_2 = I_1, K_3 = K_2 = K_1, L_1 = J_2, J_3 = J_1, L_3 = L_2$, 并且 $c_{ijkl} = \sum_{e=1}^I a_{jke} \cdot b_{iel} (i = 1, 2, \dots, I_3, j = 1, 2, \dots, J_3, k = 1, 2, \dots, K_3, l = 1, 2, \dots, L_3)$.

⑥ 纵高乘 (VI 类乘法)

$$(A_{I_1 \times J_1 \times K_1 \times L_1} \cdot B_{I_2 \times J_2 \times K_2 \times L_2})_V = [c_{ijkl}]_{I_3 \times J_3 \times K_3 \times L_3}$$

其中, $I_3 = I_2 = I_1, J_3 = J_2 = J_1, L_1 = K_2, K_3 = K_1, L_3 = L_2$, 并且 $c_{ijkl} = \sum_{e=1}^I a_{jke} \cdot b_{iel} (i = 1, 2, \dots, I_3, j = 1, 2, \dots, J_3, k = 1, 2, \dots, K_3, l = 1, 2, \dots, L_3)$.

(4) 转置

指 4 维矩阵 $A_{I \times J \times K \times L}$ 的 6 类转置

① 行列转置 (I 类转置)

$$(A_{I \times J \times K \times L})^{T_I} = [a_{ijkl}]_{I \times J \times K \times L}$$

② 行纵转置 (II 类转置)

$$(A_{I \times J \times K \times L})^{T_{II}} = [a_{ijkl}]_{K \times J \times I \times L}$$

③ 行深转置 (III 类转置)

$$(A_{I \times J \times K \times L})^{T_{III}} = [a_{ijkl}]_{L \times J \times K \times I}$$

④ 列纵转置 (IV 类转置)

$$(A_{I \times J \times K \times L})^{T_{IV}} = [a_{ijkl}]_{I \times K \times J \times L}$$

⑤ 列深转置 (V 类转置)

$$(A_{I \times J \times K \times L})^{T_V} = [a_{ijkl}]_{I \times L \times K \times J}$$

⑥ 纵深转置 (VI 类转置)

$$(A_{I \times J \times K \times L})^{T_{VI}} = [a_{ijkl}]_{I \times J \times L \times K}$$

1.4 4 维矩阵离散余弦变换及其逆变换

$$\text{设 } [c_{UVWS}]_L \begin{cases} \sqrt{\frac{2}{N}} \cos \frac{i\pi(2j+1)}{2N} & i \neq 0 \\ \frac{1}{\sqrt{N}} & i = 0 \end{cases}$$

$L=1$ 时, $i=u, j=v$

$L=2$ 时, $i=u, j=w$

$L=3$ 时, $i=u, j=s$

$L=4$ 时, $i=v, j=w$

$L=5$ 时, $i=v, j=s$

$L=6$ 时, $i=w, j=s$

4 维矩阵离散余弦变换定义为

$$B = (C_6(C_1AC_1^T)_I C_6^T)_V = (C_1(C_6AC_6^T)_V C_1^T)_I = (C_5(C_2AC_2^T)_I C_5^T)_V = (C_2(C_5AC_5^T)_V C_2^T)_I = (C_3(C_4AC_4^T)_V C_3^T)_I = (C_4(C_3AC_3^T)_I C_4^T)_V$$

4 维矩阵离散余弦逆变换定义为:

$$A = (C_6^T)_V (C_1^T)_I BC_1 = (C_1)_V (C_6^T)_V BC_6 = (C_5^T)_V (C_2^T)_I BC_2 = (C_2)_V (C_5^T)_V BC_5 = (C_3^T)_I (C_4^T)_V BC_4 = (C_4)_I (C_3^T)_I BC_3$$

2 彩色视频的 4 维矩阵离散余弦变换编码

2.1 彩色视频的 4 维矩阵表示

由 L 幅大小为 $I \times J$ 的彩色图象构成的视频序列, 每幅由 3 帧同样大小的灰度图象组成, 可以由如下的 4 维矩阵来表示: $A_{I \times J \times 3 \times L} = [a_{ijkl}]_{I \times J \times 3 \times L}$, 其中, $(1 \leq i \leq I, 1 \leq j \leq J)$ 表示彩色图象中像素的空间位置, $1 \leq k \leq 3$ 代表彩色空间的 3 个不同分量, $1 \leq l \leq L$ 表示视频序列的时间分量. 因此, 彩色视频的 4 个分量可以分别由 4 维矩阵的 4 个平面向量表示, 彩色视频中的任一像素点可用 4 维矩阵中相应位置的点来表示. 彩色视频的 4 维矩阵表示使其各像素间的位置关系以及相应的各分量间的关系构筑在一个统一的数学模型中.

2.2 编码系统模型

彩色视频的 4 维矩阵离散余弦变换编码系统原理框图如图 1 所示。

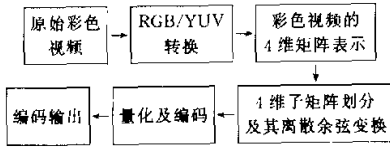


图 1 4 维矩阵离散余弦变换编码系统原理框图

(1) 4 维子矩阵划分及其离散余弦变换

将彩色视频的 $I \times J \times 3 \times L$ 阶 4 维矩阵分割成 $3 \times 3 \times 3 \times 3$ 的一系列 4 维子矩阵, 对此 4 维子矩阵 $A_{3 \times 3 \times 3 \times 3} = [a_{ijkl}]_{3 \times 3 \times 3 \times 3}$ 进行 4 维矩阵离散余弦变换, 得到变换系数 4 维子矩阵。

(2) 量化及编码

彩色视频经上述 4 维矩阵变换后, 还要通过量化器将变换系数进行压缩编码。矢量量化总是优于标量量化, 是一种非常有效的信号压缩方法。因此, 对经上述变换后的各 4 维子矩阵 (即, 4 维变换系数子矩阵) 利用矢量量化进行编码, 每个 $3 \times 3 \times 3 \times 3$ 的 4 维变换系数子矩阵可形成一个 81 维矢量, 通过

矢量量化将变换系数转换为压缩码流。

3 实验结果

分别对图象大小为 180×144 (QCIF 格式) 的原始彩色视频序列“Miss America”和“Suzie”进行压缩编码, 并将编码结果同 MPEG-2 标准中的 2D-DCT/运动补偿方法进行了比较。解码恢复图象质量的客观评价采用峰值信噪比, 即

$$PSNR = 10 \log \frac{255^2}{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y) - F(x,y)]^2} \quad (\text{dB})$$

其中, M 是图象的宽度, N 是图象的高度, $f(x,y)$ 是原图象的像素值, $F(x,y)$ 是解码恢复图象的像素值。

在本方法中所用的矢量量化码书是采用 LBG 算法生成的^[6], 码书的大小分别为: 64、128 和 256。每一个码字的维数为 81 (即: $3 \times 3 \times 3 \times 3$)。

表 1 列出了 4 维矩阵离散余弦变换编码与 2D-DCT/运动补偿算法的压缩结果。

压缩比为 108 时, Miss America 和 Suzie 第 12 帧 Y 分量的原图象和恢复图象如图 2 所示。

表 1 4 维矩阵离散余弦变换编码与 2D-DCT/运动补偿算法的压缩结果

单位: dB

视频序列	本文方法						2D-DCT/运动补偿算法	
	压缩比	帧	PSNR _V	PSNR _U	PSNR _V	平均 PSNR	压缩比	平均 PSNR
Miss America	108	12	33.777	37.836	35.057	35.557	78	31.836
		13	33.831	37.996	35.071	35.633		
		14	33.813	37.782	34.985	35.527		
	92.6	12	35.436	38.556	37.084	37.025		
		13	35.552	38.703	37.200	37.152		
		14	35.488	38.338	37.015	36.947		
	81	12	38.995	39.840	38.742	39.192		
		13	38.126	40.035	38.861	39.007		
		14	37.964	39.608	38.620	38.731		
Suzie	108	12	30.717	38.712	37.691	35.707	48	30.103
		13	31.139	38.651	37.754	35.848		
		14	30.412	38.579	37.738	35.576		
	92.6	12	31.859	39.569	38.445	36.624		
		13	32.398	39.546	38.542	36.829		
		14	31.737	39.448	38.503	36.563		
	81	12	33.338	40.364	39.596	37.766		
		13	33.904	40.403	39.705	38.004		
		14	33.219	40.310	39.744	37.758		



(a) Miss America 原图象 (b) Miss America 恢复图象 (c) Suzie 原图象 (d) Suzie 恢复图象

图 2 Miss America 和 Suzie 第 12 帧 Y 分量的原图象和恢复图象

4 结 论

4 维矩阵及 4 维矩阵离散余弦变换理论分别是对二维矩阵论及传统二维离散余弦变换理论的发展. 4 维矩阵离散余弦变换应用于彩色视频的压缩编码, 可以将彩色视频内在的全部相关性统一起来考虑, 从而解决传统 MPEG 等方法中没有考虑彩色空间中 Y、U、V 3 分量之间相关性的问题. 实验中对可视电话和会议电视等类应用中的典型彩色视频图象用文中方法进行了压缩编码, 并同 MPEG-2 标准中的 2D-DCT/运动补偿方法进行了压缩效果比较. 结果表明, 本文方法在恢复图象的信噪比和压缩比上, 都优于 MPEG-2 标准中的 2D-DCT/运动补偿方法. 该方法的提出, 从基础理论上为未来视频压缩技术的发展提供了一个新的思路.

本文方法的探讨尚属初步, 如何使基于 4 维矩阵离散余弦变换理论的视频压缩方法对复杂和运动激烈的视频图象在满足质量较好的恢复图象效果下, 达到更高的压缩比, 将是进一步研究的方向.

参 考 文 献

- 1 CCITT Recommendation H. 261. Video codec for audio visual services at $p \times 64$ kbits/s[S]. 1990.
- 2 ITU-T Draft Recommendation H. 263 Version 2. Video coding for low bitrate communication[S]. 1998.

- 3 ISO/IEC 11172-2. Coding of moving pictures and associated audio for digital storage media at up or about 1.5 Mbit/s (MPEG1 video)[S]. 1993.
- 4 ISO/IEC 13818-2. Generic coding of moving pictures and associated audio information (MPEG2 video)[S]. 1994.
- 5 朱艳秋, 陈贺新, 戴逸松. 彩色图象压缩矩阵变换编码[J]. 电子学报, 1997, 25(7): 16~21.
- 6 Linde Y, Buzo A, Gray R M. An algorithm for vector quantizer design[J]. IEEE Transactions on communications, 1980, 28(1): 84~89.



赵 岩 1971 年生, 1999 年获吉林工业大学硕士学位, 2000 年攻读博士学位. 主要从事多维矩阵理论、视频编码和图象通信研究.



陈贺新 1949 年生, 吉林大学通信工程学院教授, 博士生导师, 1989 年毕业于吉林工业大学电子工程系并获工学博士学位, 主要从事数字图象处理、多维矩阵理论、多维数字信号处理、计算机视觉等研究.