

# 彩色图象三维 DCT 变换压缩编码

朱艳秋 陈贺新 戴逸松

(吉林工业大学电子工程系, 长春 130022)

**摘要** 由于彩色图象是以三分量形式表示的,因此在对其进行压缩编码时不仅要考虑它在空间、结构等方面的冗余信息,还应充分利用各分量间的相关性。此文提出了一种基于彩色图象三维矩阵表示模型的三维 DCT 变换压缩编码方法。这种方法将彩色图象中各分量上的全部信息同时进行变换压缩,有效地去除了彩色图象中的各种冗余信息。仿真结果表明,该方法实用、简便,能在获得较好重建图象质量的同时具有较高的压缩比。

**关键词** 彩色图象压缩, 变换压缩编码

## 1 引言

长久以来,人们一直在探索降低图象数据量的方法。80年代初,研究者们突破了传统的 Shannon 理论的框架,开始重视对图象信息的适当分解和表征,并根据人眼的视觉机理决定信息的取舍和码字的分配,使图象压缩编码技术取得了突破性进展。进展的标志就是建立了关于图象编码的三个国际标准,即 JPEG、H. 261、MPEG 标准。这 3 种图象编码标准总结了人们数十年的研究成果,它们已被广泛应用到多种视频通讯系统中,并发挥了巨大的作用。然而,随着数字技术的进步,以及计算机、通信和大众传播技术的不断发展与结合,灰度图象已基本上被彩色图象取代。由于一幅彩色图象的原始图象数据是灰度图象的 3 倍,因此图象编码技术必须达到更高的压缩比,才能适应海量数据的彩色图象存贮和传输的要求。但是,目前图象压缩编码技术基本是针对二值图象或灰度图象,彩色图象的压缩编码基本上是将灰度图象的压缩编码方法或其改进方法分别独立地应用于彩色图象的各个分量(如 RGB、YUV 等)上。由于图象压缩编码的本质就是尽可能地去除图象信息中存在的各种冗余信息,以降低表示图象所需的数据量,而彩色图象中不仅具有灰度图象所具有的各种冗余特征,同时各个色分量之间

也存有大量的冗余信息<sup>[1]</sup>。因此,对彩色图象各分量独立编码的结果是未能把彩色图象的各种冗余信息统筹考虑,从而不能充分发挥彩色图象的内在压缩潜力,极大地限制了压缩比的进一步提高。

彩色图象在人的视觉中具有丰富多彩的颜色结构,它更为逼真地反映了现实世界。在一幅彩色图象的象素点阵中,每个象素点信息由三基色(RGB)各自的亮度构成。由于三分量同出于同一物理模型的事实,各分量之间存在着极强的相关性,在对彩色图象进行压缩编码时,通常先进行颜色空间转换,将采样量化后的 RGB 数据转换为 YUV(在 JPEG 标准<sup>[2]</sup>中使用)数据,然后各分量独立地进行压缩编码。一般地,亮度(Y)信号携带图象信息所占的比重最大,而色度(U、V)信号相对地要小得多,因此,这种 RGB 到 YUV 空间的转换实际上是能量集中的一种过程。但这种初级转换仅仅是统计意义上的一种简单的去除冗余的过程,并没有真正有效地去除各分量间的冗余信息,例如:当将色空间的三分量 YUV 分别以灰度图象单独显示时,它们仍表现为极大的视觉相似性<sup>[3]</sup>,这反映为色空间冗余。我们认为,真正的彩色图象压缩编码的关键应是将彩色图象中的全部信息用一个模型表达,在去除冗余的过程中统筹考虑彩色图象在空间结构上及各分量之间所存在的冗余信息,以期达到更好的压缩效果。本文针对彩色图象压缩编码中存在的上述问题,采用彩

色图象的三维矩阵模型<sup>[4]</sup>表示方法,用三维 DCT 结合人类视觉系统特性实现了彩色图象的压缩编码。实验结果表明,这种方法在具有较好的重建图象质量的同时获得了较高的压缩比。

## 2 三维矩阵模型及三维 DCT 变换

在图象压缩编码中,变换编码对图象统计特性的变化不是那么敏感,能有效地消除图象数据空间的相关性。同时,因为压缩编码是在变换域中进行的,而变换域内的一些分量往往反映了图象的一些基本特征。以二维 DCT 为例,根据变换系数的能量分布,可以将图象划分为不同的区域,其中变换后幅值较大的图象系数大多数集中于图象块的左上角,与其它系数相比,这些低频系数具有的能量大,包括了图象的大部分内容,变换图象中的其它系数代表了图象的边缘和方向,幅度很小,只具有原图象中很

小比例的能量,对图象质量影响甚微。人们可以根据人的视觉机理对变换域的系数分别对待,使得编码质量很高。通过人们长期的研究和实际应用,人们发现变换编码、特别是离散余弦变换压缩编码方法非常适合于电视类图象的压缩编码,具有压缩比高、误差影响小等明显优点。现在,DCT 已成为图象编码的核心技术之一,如 JPEG、H. 261、MPEG 中都采用的是 DCT 技术。

### 2.1 彩色图象的三维矩阵模型表示

如前所述,一幅尺寸为  $I \times J$  的彩色图象由 3 幅同样大小的灰度图象组成,我们可以用三维矩阵  $A_{I \times J \times 3} = [a_{ijk}]_{I \times J \times 3}$  表示,其定义如下。

**定义 1:** 由  $I \times J \times K$  个实数  $a_{ijk} (1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K)$  构成的  $I$  个横行、 $J$  个竖列、 $K$  个纵序的立方体形数据排列空间,称为  $I \times J \times K$  阶三维矩阵,记作:

$$A_{I \times J \times K} = [a_{ijk}]_{I \times J \times K} = \begin{bmatrix} a_{111} & a_{121} & \cdots & a_{1J1} \\ a_{211} & a_{221} & \cdots & a_{2J1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I11} & a_{I21} & \cdots & a_{IJ1} \end{bmatrix} \begin{bmatrix} a_{112} & a_{122} & \cdots & a_{1J2} \\ a_{212} & a_{222} & \cdots & a_{2J2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I12} & a_{I22} & \cdots & a_{IJ2} \end{bmatrix} \cdots \begin{bmatrix} a_{11k} & a_{12k} & \cdots & a_{1Jk} \\ a_{21k} & a_{22k} & \cdots & a_{2Jk} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1k} & a_{I2k} & \cdots & a_{IJk} \end{bmatrix}$$

**定义 2:** 在三维矩阵模型中,可以用一些横线、竖线和纵线划分成尺寸小的其它一些三维数据排列空间:

$$A_{U \times V \times W}(a, b, c)$$

$$(a = 1, \dots, E; b = 1, \dots, F; c = 1, \dots, G) \quad (2)$$

其中:  $E, U < I; F, V < J; G, W < K$ 。式(2)叫作式(1)的三维子矩阵。

特别地,三维矩阵  $A_{I \times J \times K}$  (或三维子阵)又可分割为二维的横向、竖向、纵向平面向量,分别表示为:

$$A_{(i=I)}^{J \times K} = [a_{ijk}]_{(i=I)}^{J \times K}, i_1 = 1, 2, \dots, I;$$

$$A_{(j=J)}^{I \times K} = [a_{ijk}]_{(j=J)}^{I \times K}, j_1 = 1, 2, \dots, J;$$

$$A_{(k=K)}^{I \times J} = [a_{ijk}]_{(k=K)}^{I \times J}, k_1 = 1, 2, \dots, K.$$

彩色图象的三维矩阵表示方法使彩色图象各分量信号统一在一个表示模型中,使图象的空间结构关系和各分量间的关系统一起来,为有效地去除彩色图象中的各种冗余信息打下了基础。

关于三维矩阵的详细讨论参见文献[4]。

### 2.2 三维 DCT 变换

类似于二维 DCT 变换,我们定义三维 DCT 变

换如下:

**定义 3:** 以三维矩阵  $X_{I \times J \times K} = [x(i, j, k)]_{I \times J \times K}$  表示的  $I \times J \times K$  个其值有限的三维实数信号序列的集合,  $i = 0, 1, \dots, I-1; j = 0, 1, \dots, J-1; k = 0, 1, \dots, K-1$ , 其三维 DCT 变换定义为:

$$F_{I \times J \times K}(l, m, n) = \sqrt{\frac{8}{I \times J \times K}} c(l) \cdot c(m) \cdot$$

$$c(n) \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} x(i, j, k) \cdot \cos \frac{2i+1}{2I} l \pi \cdot$$

$$\cos \frac{2j+1}{2J} m \pi \cdot \cos \frac{2k+1}{2K} n \pi =$$

$$[(O_1 \cdot X(i, j, k) \cdot O_2)_{III} \cdot O_3]_I$$

(3)

其中:

$$O_1 = [o_1(l, i, k)]_{I \times J \times K} =$$

$$\sqrt{\frac{2}{I}} [c(l) \cos \frac{2i+1}{2I} l \pi]_{I \times J \times K}$$

$$l(\text{横行}), i(\text{竖列}) = 0, 1, \dots, I-1;$$

$$K(\text{纵序}) = 0, 1, \dots, K-1$$

$$O_2 = [o_2(j, m, k)]_{J \times J \times K} =$$

$$\sqrt{\frac{2}{J}} [c(m) \cos \frac{2j+1}{2J} m\pi]_{J \times J \times k}$$

$j$ (横行),  $m$ (竖列) = 0, 1, ...,  $J - 1$ ;  
 $k$ (纵序) = 0, 1, ...,  $K - 1$

$$O_3 = [o_3(i, k, n)]_{I \times K \times K} =$$

$$\sqrt{\frac{2}{K}} [c(n) \cos \frac{2k+1}{2K} n\pi]_{I \times K \times K}$$

$i$ (横行) = 0, 1, ...,  $I - 1$ ;  $k$ (竖列),  
 $n$ (纵序) = 0, 1, ...,  $K - 1$

$$c(\alpha) = \begin{cases} \frac{1}{\sqrt{2}}, & \alpha = 0 \\ 1, & \alpha \neq 0 \end{cases}$$

$$X_{I \times J \times K}(i, j, k) = [O_1 \cdot (F_{1 \times J \times K}(l, m, n) \cdot O_3)_I \cdot O_2]_{III} =$$

$$\sqrt{\frac{8}{I \times J \times K}} \sum_{l=0}^{I-1} \sum_{m=0}^{J-1} \sum_{n=0}^{K-1} c(l)c(m)c(n)F_{1 \times J \times K}(l, m, n) \cdot \cos \frac{2i+1}{2I} l\pi \cos \frac{2j+1}{2J} m\pi \cos \frac{2k+1}{2K} n\pi \quad (4)$$

其中:

$$O_3 = [o_3(l, n, k)]_{I \times K \times K} =$$

$$\sqrt{\frac{2}{K}} [c(n) \cos \frac{2k+1}{2K} n\pi]_{I \times K \times K}$$

$l$ (横行) = 0, 1, ...,  $I - 1$ ;  $n$ (竖列),  
 $k$ (纵序) = 0, 1, ...,  $K - 1$

$$O_2 = [o_2(m, j, k)]_{J \times J \times K} =$$

$$\sqrt{\frac{2}{J}} [c(m) \cos \frac{2j+1}{2J} m\pi]_{J \times J \times k}$$

$m$ (横行),  $j$ (竖列) = 0, 1, ...,  $J - 1$ ;  
 $k$ (纵序) = 0, 1, ...,  $K - 1$

$$O_1 = [o_1(i, l, k)]_{I \times I \times K} =$$

$$\sqrt{\frac{2}{I}} [c(l) \cos \frac{2i+1}{2I} l\pi]_{I \times I \times K}$$

$i$ (横行),  $l$ (竖列) = 0, 1, ...,  $I - 1$ ;  
 $k$ (纵序) = 0, 1, ...,  $K - 1$

$$c(\alpha) = \begin{cases} \frac{1}{\sqrt{2}}, & \alpha = 0 \\ 1, & \alpha \neq 0 \end{cases}$$

式(3)和(4)就是三维 DCT 变换和反变换表达式。

### 3 彩色图象的三维 DCT 变换压缩编码

众所周知,二维 DCT 能够把二维灰度图象信息变换到频率域上并能集中到低频分量上,结合人类视觉特性,可以达到信息压缩的目的。三维 DCT

**定理 1:** 三维 DCT 变换是可分离的,即此三维 DCT 变换可分解成三重一维 DCT。

该定理很容易得到证明。从式(3)可见,先以  $i$  为变量,对图象三维矩阵的横行方向逐次进行一维 DCT 得到一个中间结果  $F'_{I \times J \times K}(l, j, k)$ ; 再对  $F'_{I \times J \times K}(l, j, k)$  以  $j$  为变量,在竖列方向上逐次进行第二个一维 DCT 得到另一个中间结果  $F''_{I \times J \times K}(l, m, k)$ ; 然后再对  $F''_{I \times J \times K}(l, m, k)$  以  $k$  为变量在纵序方向上进行第三个一维 DCT 得到最终的变换结果  $F_{I \times J \times K}(l, m, n)$ 。

根据定理 1,三维 DCT 的反变换可直接由一维的反变换 IDCT 推出:

变换把二维 DCT 处理灰度图象的方法延展到三维数据的彩色图象中来,因此我们提出如图 1 的压缩编码模型:

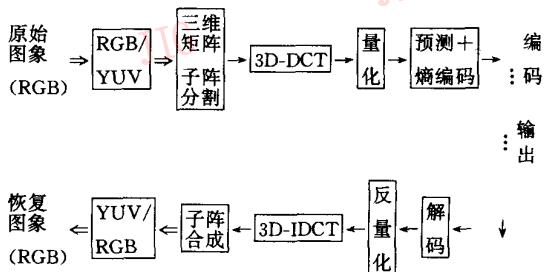


图 1 三维 DCT 变换压缩编码系统模型

Fig. 1 3D-DCT compression coding system model

首先,根据人类视觉特性可知,人眼对亮度信号和色度信号的敏感程度不同,对于色度信号不敏感。因此,普遍认为,以亮度(Y)、色度(U、V)信号方式编码是方便的,因为人类视觉系统的性能可以被充分利用以达到信息压缩的目的。

显然,转换(即 RGB 到 YUV)后的 YUV 信号无论是在空间结构上还是在各分量之间仍留有大量的冗余信息,因此我们采用 3D-DCT 变换来进一步去除冗余。从有效性和实用性的角度出发,我们选用  $8 \times 8 \times 3$  子阵块进行 3D-DCT 变换。具体变换过程如图 2 中 I  $\rightarrow$  II 所示。图中是针对某一幅图象进行变换的。为了统计 3D-DCT 变换系数方差,现利用 4 种图象进行变换,它们是猫,花(1),花(2),女孩等

256×256 彩色图象。每种图象 YUV 经变换后得到图 2 的  $Y_{DCT}$ 、 $U_{DCT}$  和  $V_{DCT}$  象素阵列, 分别计算 3 个纵向平面向量  $Y_{DCT}$ 、 $U_{DCT}$  和  $V_{DCT}$  的算术均值  $\bar{Y}_{DCT}$ 、 $\bar{U}_{DCT}$  和  $\bar{V}_{DCT}$ , 然后计算它们各元素对应的方差, 于是得到图 2 中 III 的方差数据阵列。对应 4 种彩色图象共得到四组方差数据阵列, 最后按各元素点计算 4 种彩色图象的变换系数方差的均值  $\bar{\sigma}Y$ 、 $\bar{\sigma}U$  和  $\bar{\sigma}V$

(即 4 个数值相加后除以 4), 得到图 3 所示的三维 DCT 归一化系数方差分布。在图 3 中, 为了形象地用 4 维空间表示系数的方差分布, 采用了 3 个曲面相比较的方式, 即曲面 1 对应  $\bar{\sigma}Y$ , 曲面 2 对应  $\bar{\sigma}U$ , 曲面 3 对应  $\bar{\sigma}V$ 。其中,  $(X, Y)$  是图象子块的空间位置,  $Z$  是归一化系数方差值。

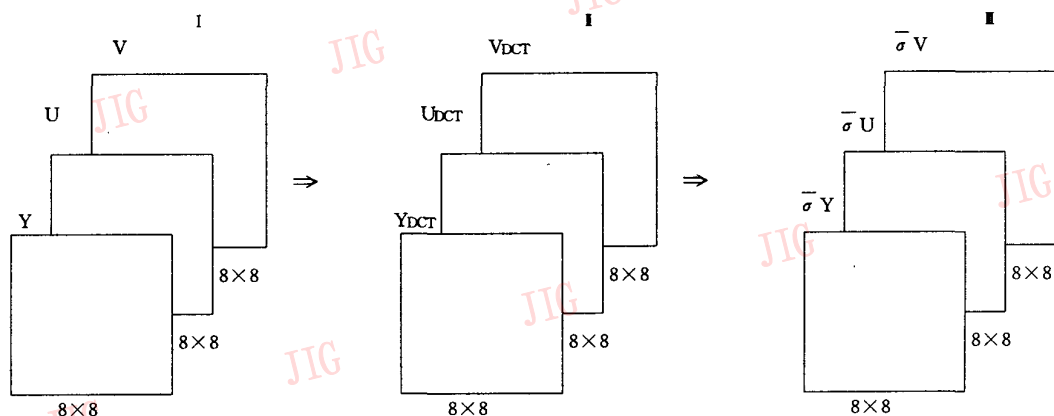


图 2 三维 DCT 变换过程和系数方差计算过程

Fig. 2 The process of 3D-DCT and variance calculation

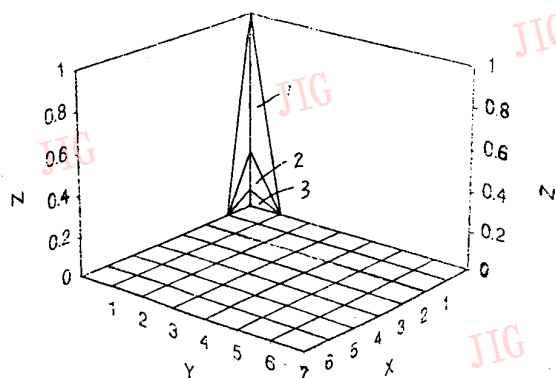


图 3 三维 DCT 归一化系数方差分布

Fig. 3 Normalized variance distribution of 3D-DCT transform coefficient

可见, 3D-DCT 变换的结果把彩色图象的能量大部分都集中于三维矩阵表示模型的左上角。考虑到 3D-DCT 系数的统计特性及人类视觉对不同 DCT 频率系数的不同敏感程度, 容易得出 192 个 DCT 系数的量化阶距应按如下原则选取: 对于频率较低的系数, 由于其方差较大, 即动态范围较大, 需要较小的量化阶距才能以很小的失真代价描述此系

统所包含的信息; 而对于频率较高的系数, 由于其方差较小, 只要用很少的比特数就可以描述, 故量化阶距要大一些。另外, 高频 DCT 系数的量化取大一些, 还可以把那些能量非常小的系数样值当作背景噪声抑制掉, 输出零值。

考虑到水平方向、垂直方向和纵序方向三维频率的变化, 我们将  $8 \times 8 \times 3$  个 DCT 系数按如下顺序排列成一个长度为 192 的 DCT 系数向量: 即对于  $F_{8 \times 8 \times 3}(i, j, k)$  系数矩阵, 以纵向平面向量  $F_{8 \times 8 \times k=0}$  为基准平面向量, 并先把该平面向量中的系数按 Z 形排列, 扫描时, 从  $F_{8 \times 8 \times 3}(0, 0, 0)$  系数开始, 先沿纵序方向依次排序, 即  $F_{8 \times 8 \times 3}(0, 0, 0) \rightarrow F_{8 \times 8 \times 3}(0, 0, 1) \rightarrow F_{8 \times 8 \times 3}(0, 0, 2)$ , 然后回到基准平面上按 Z 形排序的下一个系数点, 即  $F_{8 \times 8 \times 3}(0, 1, 0)$ , 再由此点开始作纵向排序, 如此循环排列下去直到最后一系数点  $F_{8 \times 8 \times 3}(8, 8, 3)$ 。

上述排列顺序实际上是将  $8 \times 8 \times 3$  个 DCT 系数按频率由低到高的顺序排列的。根据系数方差曲线可知, 系数方差主要分布在低频分量上, 高频分量上的系数很小, 大多数为零, 而多个零值连续出现非常有利于熵编码部分做进一步有效的压缩。

另外,由于每次变换都是作用于  $8 \times 8 \times 3$  子阵区域内,因此 3D-DCT 变换有效地消除了相邻象素点之间的相关性,但对于彩色图象的整体来说还需进一步去除块与块之间的相关性。因此,我们对数值较大的直流分量系数及低频分量系数  $F_{8 \times 8 \times 3}(0, 0, k), k=0, 1, 2$  分别作差分预测处理后再采用 Huffman 熵编码。

#### 4 实验与评价

实验选用的图象数据是从标准图象库中选取的 4 幅  $256 \times 256 \times 8 \times 3$  彩色图象。仿真实验在一台 486 微机上完成。

压缩比的计算为:

$$Cr = (256 \times 256 \times 24) / b_{p0}$$

其中,  $b_{p0}$  为编码后的总比特数。

图象质量的客观评价采用平均峰值信噪比

$$\overline{PSNR} = \frac{PSNR(R)}{3} + \frac{PSNR(G) + PSNR(B)}{3}$$

其中,  $PSNR = 10 \log \frac{255 \times 255}{MSE}$ ;

$PSNR(R), PSNR(G), PSNR(B)$  分别为重建图象 RGB 三分量的峰值信噪比。

为了评价 3D-DCT 变换压缩编码的性能,我们做了与 JPEG 基本彩色系统的对比实验(JPEG 实验采用常用软件“JPEG 基本系统”程序<sup>[5]</sup>)。众所周知,量化系数的不同选取,恢复后图象的质量会有很大区别。为了清晰、定量地进行 JPEG 方法和本文提出的方法之间的性能比较,在此量化系数选取较大(选为 40)。在两种方法的 PSNR 都较低而且基本相同的情况下,可以看出本文提出的方法  $Cr$  是 JPEG 方法  $Cr$  的 2.5 倍左右,如表 1 所示。

由图 4 可见“猫”图象在表 1 所示实验结果下的重建图象。从压缩效果上看,三维 DCT 变换压缩编码所得到的图象的“方块”效应比 JPEG 基本彩色系统编码下的要有所减弱,这是因

为,3D-DCT 把彩色图象的 3 个色分量上的能量更有效地集中在一个平面向量上,通过对其低频系数及高频系数量化的调整,可以使得“方块效应”得到控制。

表 1 三维 DCT 变换压缩编码方法与 JPEG 基本彩色系统对比实验结果

Table 1 Comparison of experiments result for 3D-DCT and JPEG

	JPEG 基本彩色		3D-DCT 变换压缩编码系统	
	$C_r$	PSNR	$C_r$	PSNR
猫	15.2	30.2	39.4	30.4
花(1)	12.4	27.3	30.1	27.2
花(2)	13.5	28.1	33.3	28.3
女孩	14.2	29.9	37.6	30.2



(a)



(b)



(c)

(a)原始图象 (b)3D-DCT 重建图象 (c)JPEG 重建图象

图 4 原始图象及 3D-DCT、JPEG 重建图象(从左至右分别为 R、G、B)

Fig. 4 Original image and 3D-DCT、JPEG reconstructed image

另外,对 DCT 系数的量化分层值,即量化阶的大小取决于它的变换域坐标或“频谱”位置,量化分层表应该是根据大量视觉心理实验导出的,本文由于时间和条件关系仅采用了均匀量化器对 3D-DCT 系数进行量化,因此,此方法有望进一步提高其压缩性能。

## 5 结 论

本文利用三维矩阵新概念,将彩色图象统一表示,进而实现统一的三维 DCT 变换。这种一次性三维 DCT 变换方法比传统的分帧处理具有在同样 PSNR 下压缩比更高、处理更方便的特点,为彩色图

象压缩编码开辟了一条新的实用途径。

## 参 考 文 献

- 1 Nasraddi N M. A multilayer address vector quantization technique. IEEE Trans on Circuits and System, 1990, 37(7): 912~921.
- 2 曾国良,丁宗豪. 静象数据压缩及其标准 JPEG(上)和(下). 电视技术, 1994, 4, 5.
- 3 Memon N D. Lossless compression of RGB color images. Optical Engineering, 1995, 34(4): 1711~1717.
- 4 朱艳秋,陈贺新,戴逸松. 彩色图象三维矩阵变换压缩编码. 电子学报.
- 5 黎洪松. 数字图象压缩编码技术及其 C 语言程序范例, 北京: 学苑出版社. 1994, 5: 180~230.



朱艳秋, 1987 和 1990 年分获西安交通大学电气工程学士和硕士学位, 1990 年任教于吉林工业大学电子系, 1994 年于该系攻读博士学位。主要研究领域包括彩色图象压缩编码, 变换理论及其在图象处理和计算机视觉中的应用。

# Compression Coding of Color Image via 3D-DCT Transform

Zhu Yanqiu, Chen Hexin, Dai Yisong

(Department of Electronic Engineering, Jilin Technology University, ChangChun 130022)

**Abstract** Due to the 3-components representation of a color image, it's essential that we should not only take into account the spatial and structural redundancies, but also fully exploit the spectral correlations present in color image data. In this paper, we present a 3D-DCT algorithm for the compression coding of color image, which is based on 3D matrix model. All the color image components are jointly transform compressed and the redundancies are effectively reduced. Simulation results demonstrate good quality reproduction and high compression ratio of this method, which is also simple and convenient.

**Keywords** Compression of color image, Transform-compressed coding