

# XYZ 实时视频编码技术

朱秀昌

(南京邮电学院信息工程系, 南京 210003)

**摘要** 论述了3D-DCT 实时视频压缩编码的原理及其特点。3D-DCT 是一种对称的编解码方法, 和当前国际标准(如 H. 261、MPEG 等)中的混合编码方法相比, 它具有较低的复杂度, 较高的压缩比及较好的重建图像质量。因此, 3D-DCT 技术可望在视频编码中取得良好的应用前景。

**关键词** 图像通信 视频压缩 XYZ 技术 三维 DCT

## 1 重提 XYZ 编码

XYZ 编码实际上是三维离散余弦变换(3D-DCT)的一种形象化的代名词。在图像压缩编码中, 离散余弦变换(DCT)是一种仅次于 K-L 变换的正交变换, 目前广泛应用于各种视频压缩标准, 如 H. 261、MPEG、JPEG 等。既然 2D-DCT 如此, 人们自然会想到 3D-DCT。事实确是如此, 早在 20 年前, 就有人提出将 3D-DCT 应用于活动图像的压缩编码, 但是受当时技术条件的限制, 如算法的复杂性和对大容量的存储体的需求, 限制了硬件或软件的实时实现。经过 20 年的技术发展, 高速计算机、大容量存储体的出现, ASIC 的性能的不断提高等, 这一切都为 3D-DCT 的实现提供了技术保证。因此, 很有必要重提 XYZ 三维压缩编码技术, 并对它加以改进, 使之更适应于当前技术水平, 适应于实时应用。

如果说实时实现的时机已经成熟是我们重提 3D-DCT 的一个重要原因的话, 那么, 另一个更为重要的理由就是 XYZ 技术优良的编码性能。和目前以 2D-DCT 及运动估计(MC)为核心的视频压缩编码方法相比较, 在相同的压缩比较条件下, XYZ 方法所得到的重建图象的质量要略优于 H. 261 或 MPEG (如压缩比为 34 左右时, XYZ 的 NRMSE 为 0.079, 而 MPEG 为 0.125); 在相同的重建图象质量的条件下, XYZ 的方法的压缩比要高于这些标准所规定的混合编码方法(如当 NRMSE 约为 0.08 时, XYZ 的

压缩比 34.5, 而 MPEG 只有 15.6)。而且, XYZ 方法的编解码实现的复杂性要低于标准所规定的方法(见表 1)<sup>[1]</sup>。

表1 各种压缩方法的大致比较

算 法	平 均 压缩比	相对编码复杂度 (参考 MIPS 数)	相对解码复杂度 (参考 MIPS 数)
XYZ	75:1	2(240)	2(240)
H. 261/H. 263	50:1	5(970)	1(200)
MPEG	30:1	10(1 120)	1(120)
M-JPEG	10:1	1	1

MPEG 技术是基于 I 帧、P 帧和 B 帧的算法, 它是一种非对称性算法, 编码器要比解码器复杂得多。而 XYZ 技术对所有的视频帧只需要一种算法, 编码器和解码器的复杂程度一样。这样, 非常有利于 VLSI 及编解码设备的实现。

## 2 XYZ 编码的基本原理

### (1) 3D-DCT

在二维 DCT 基础上, 加入时间维信息形成若干个  $N \times N \times N$  的图像“立方体”, 其三维离散余弦变换可用下式表示:

$$F(U,V,W) =$$

$$C(U)C(V)C(W) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{t=0}^{N-1} f(x,y,t) \times \cos[(2x+1)U\pi/2N] \times \cos[(2y+1)V\pi/2N] \cos[(2t+1)W\pi/2N]$$

式中  $f(x,y,t)$  表示图像时空值,  $F(U,V,W)$  为 DCT 变换域的系数。其中

$$C(i) = \begin{cases} \sqrt{2}, & i=0 \\ 1, & i=1,2,\dots,N-1. \end{cases}$$

### (2) XYZ 编码系统的构成

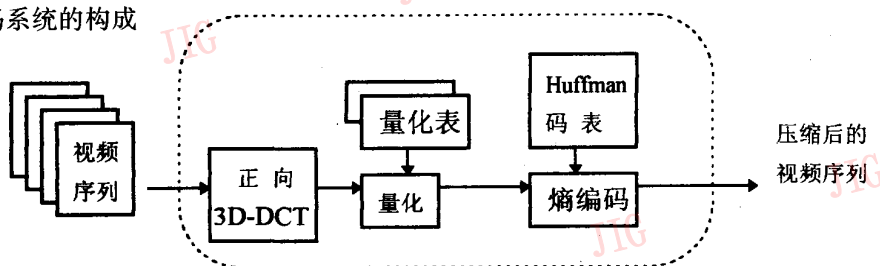


图1 XYZ 编码器

### (3) 快速3D-DCT

实时实现 XYZ 的编码, 最关键的是寻找快速 3D-DCT 算法。目前 3D-DCT 的快速算法, 大部分都是基于一维或二维 DCT 的算法。除此之外, 我们还可以利用目前计算机的 32 位、甚至 6 位数据宽度, 利用适合图像处理的指令集。例如, 在 32 位数据线上可以将 4 个 8 位的数据“四合一”地进行运算, 一次可以处理 4 个数据。如何建立专门针对 3D-DCT 特点的快速算法, 还有很多问题值得深入研究。

## 3 量化和变长编码

3D-DCT 除去了图像数据之间的大部分的相关性, 将原来分散分布的时空数据变为三维频率域系数, 主要集中分布在低频部分。但是, 即使是经过 3D-DCT, 也没有对图像信息有所压缩, 还必须通过量化和变长编码这两个步骤才能对数据进行有效的压缩。

#### (1) 量化

目前, 对 DCT 系数的量化方法主要有两类, 一类为均匀量化, 即对某一变换块中各个系数的值(频率分量)以同一种量化步长进行量化(如 H. 261、H. 263)。另一类是考虑到人眼的视觉特性, 对低频分量用较小的量化步长, 而对高频分量则用较大的量化步长。显然, 在相同的主观重建图像质量的条件下, 后者可以取得要更高的压缩比(如 JPEG、MPEG 等)。在 XYZ 技术中也可采用这种量化方

XYZ 视频实时编码系统的构成框图如图 1 所示。在编码过程中, 输入的数字视频序列经过存取变换以后, 形成一个个顺序的  $N \times N \times N$  的“立方体”像素值单位, 编码过程就是对这一个个立方体进行 3D-DCT 运算, 对所得的 DCT 系数进行量化。再把量化后的系数值按一定的规则划分成不同的游长事件进行 Huffman 编码。和编码过程完全相对称, 解码过程是解编码处理的逆运算, 顺序颠倒过来进行。

法。例如, 一维量化步长由下式给出:

$$1D-Q[U] = (0, 1, 2, 3, 6, 11, 20, 25),$$

其中,  $U=0, 1, \dots, 7$ 。

则三维量化步长为:

$$3D-Q[U, V, W] = 5 + 1D-Q[U] + 1D-Q[V] + 1D-Q[W],$$

其中,  $U, V, W$  的范围都是 0 到 7。

按此公式所形成的量化步长分布如图 2 所示。这里只是一种可能的方法, 更好的 3D-DCT 的量化还有待于进一步的研究和实验。

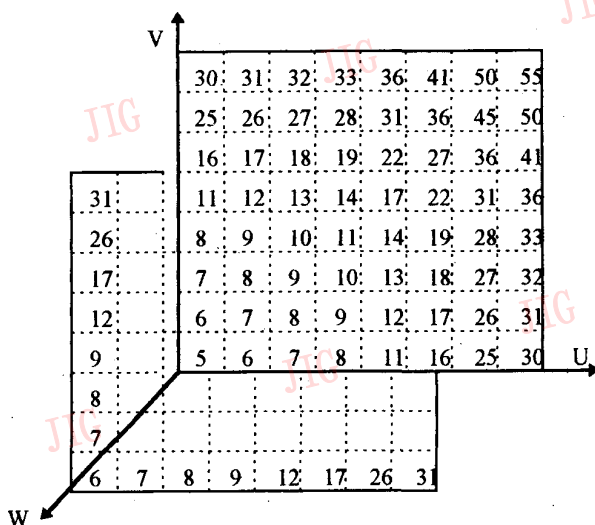


图2 三维量化步长分布

#### (2) 变长编码

在目前的编码标准中, 将二维块(8×8)的 DCT

系数进行 Zigzag 扫描,形成二维(游长,系数)事件,再对这些事件进行哈夫曼编码。在 XYZ 编码的情况下,由于加入了第三维,一种较为简单的办法就是将 2D-DCT 的 Zigzag 扫描方法应用到三维,将三维的情况看成若干个二维平面块迭合,这种方法的好处就是比较简单,可以直接套用现有标准的变长编码的码表。显然这不是最好的方法,寻找符合三维 DCT 系数分布特点的变长编码,进一步提高编码效率是 3D-DCT 的另一个有待解决的问题。

#### 4 实时应用前景

XYZ 技术的应用前景主要是在活动图像的压缩编码领域。

一种应用是软件编解码。由于 XYZ 的对称编解码形式以及较低的编码复杂度,工作站或微机的计算能力飞速提高,再加上适于图像处理的 MMX 功能,使得采用 XYZ 技术的软件编码成为可能。如在一个计算机高速局域网上,如果每台计算机都配备有 XYZ 编解码软件及数字摄像机,那么在这个网上就可以进行视频图像的交互。另一种应用是硬件编解码。由于 XYZ 编解码的关键是 3D-DCT,因此,只要设计一种专用集成电路(ASIC)或数字信号处理器(DSP)来完成这一工作,就可使得硬件编解码设备或类似的计算机插卡成本有所降低,设计和调试工作得到简化,同时也增加了可靠性。当然,XYZ 技术还可以用于 CATV 的压缩视频传输,激光视盘的节目制作,视频图像信息库,点播视(VOD)等领域。

对于 XYZ 技术,除需进一步改进快速算法和提高量化效率以外,还必需克服 XYZ 技术实时应用的一个最大的弱点:需要较大的储存容量。减少储存容量的方法之一就是寻找一种实时的迭代方法来进 3D-DCT,能够实现逐帧运算。

总之,XYZ 视频压缩技术对比于现有的视频压缩标准(如 H261、MPEG 等),它提供了较高的压缩比和较优良的图像质量,XYZ 视频技术以它的对称性和简洁性,将导致廉价的 VLSI 的实现,导致较为简单的视频压缩编码解码器的产生和应用。

#### 参考文献

- 1 Raymond Westwater, Borko Furht. Real-Time Video Compression Techniques and Algorithm. U. S. A. Kluwer Academic Publishers. July 1997.
- 2 Natarajan T, Ahmed N. On Interframe Transform Coding. IEEE Transaction on Communications, 1977, 25: 1323~1329.
- 3 Westwater R, Furht B. The XYZ Algorithm for Real-Time Compression of Full-Motion Video. Journal of Real-Time Imaging, 1996, 2(1):19~34.
- 4 Allen J D. An Approach to Fast Transform Coding in Software. Signal Processing : Image Communication, 1996, 1~11



朱秀昌 南京邮电学院信息工程系教授。1987年在南京邮电学院获通信与电子系统专业硕士学位。长期从事图像通信方面的科研和教学工作。曾先后获邮电部科技进步一等奖和国家科技进步三等奖。1993年获国家政府特殊津贴。

## The XYZ Technique for Real-Time Video Compression

Zhu Xiuchang

(Department of Information Engineering, Nanjing Institute of Posts and Telecommunications, Nanjing 210003)

**Abstract** This paper describes the theory and characteristics of 3D-DCT technique for real-time video compression. 3D-DCT technique is a symmetric method used in video coding and decoding. As compared with the hybrid coding in existent international standards (e. g. H. 261, MPEG, etc. ), it has lower complexity, higher compression rate and better reconstructed image quality. Therefore the 3D-DCT technique would have a very good spectacle in the area of real-time coding.

**Keywords** Picture communication, Video compression, XYZ technique, 3D-DCT