

一种彩色视频图像序列的三维小波变换编码方法¹

伍 鹏 张妙兰 李象霖

(中国科技大学研究生院电子学部 北京 100039)

摘 要 该文给出了一种基于三维小波变换的彩色视频图像序列的编码方法。先将彩色视频图像序列的 RGB(红绿蓝) 色彩表示转换成 YCbCr(亮度色度色度) 表示, 再将表示帧序的 t 坐标代换成 z 坐标, 可将其看成是三维空间中的“体”。将其分成子块后, 仿照二维图像小波变换的方法, 对每个亮度色度分量做三维小波变换, 变换后的各亮度色度分量的能量主要集中于低频波段, 这些波段对该视频图像序列的视觉效果影响最大。将不同分量的各个不同波段按不同的精度量化并进行熵编码, 可以达到去除帧内和帧间冗余, 压缩数据的目的。试验表明, 使用本方法可以达到较好的压缩效果。此方法较直观, 速度也比较快。

关键词 小波, 三维小波变换, 彩色视频图像编码

中图分类号 TN911.73, O177.6

1 引 言

近年来, 数字图像处理技术在工业, 影视和科研等领域的应用非常广泛。其中数字视频图像序列编码是当前国内外特别热门的研究方向。这是因为效率更高的编码方法不仅能极大节省存储空间, 也使得在带宽较窄的网络上更快更廉价地传输视频信号成为可能。图像压缩国际标准, 如静止图像编码的 JPEG, 视频图像序列编码的 H.261, H.263, MPEG-1, MPEG-2 和 MPEG-4^[4] 和正在研究、讨论中的 MPEG-7 等都有非常广阔的应用前景。

小波变换由于其特有的时频局部性等特性, 使它比传统的 DCT 变换在图像处理等领域可以取得更好的效果, 尤其是在数字图像和视频图像编码领域, 基于小波变换的编码技术可取得比基于 DCT 变换的编码方法更高的信噪比(压缩比相同时) 或压缩比(信噪比相同时)。在 Mallat 等学者提出快速小波变换理论和 AT & T 公司 Bell 实验室的 Wim Sweldens 等学者提出 lifting 技术后, 小波变换的计算速度可和快速 DCT 变换相媲美甚至更快。快速小波变换算法的逐渐成熟也使小波变换的应用研究快速发展, 基于小波变换的数字图像编码已经成为当前国内外研究的热点。

当前国内外将小波变换理论应用于视频图像序列编码的研究多集中于帧内编码使用小波变换, 帧间编码使用基于运动矢量的运动估计和运动补偿的编码方法。这种方法也是当前各国际标准, 如 H.261, H.263, MPEG-1, MPEG-2 和 MPEG-4^[1] 等所广泛采用的。本文所描述的方法对帧内和帧间编码统一采用三维小波变换, 从而省略了基于运动矢量的运动估计和运动补偿的计算。对于彩色视频信号的不同分量, 采用不同的压缩比编码, 可取得更高的编码效率和更好的重建图像质量。

2 算法描述

2.1 小波变换概述

Chui^[2], Sweldens^[3], Davis^[4] 等学者在他们的文章中详细讨论了小波, 主要是一维小波的理论和应用。Chui 和 Sweldens 给出了小波理论的严格数学定义和相关结果的证明, Davis 主要从工程的角度描述了小波在图像处理等领域的应用。根据 Chui, Sweldens 和 Davis 等给出的结果, 离散的一维小波变换(DWT) 及其逆变换的表达式如下:

¹ 1999-08-31 收到, 1999-12-19 定稿

$$\phi(x) = \sum_k h_k \phi(2x - k) \tag{1}$$

$$\tilde{\phi}(x) = \sum_k \tilde{h}_k \tilde{\phi}(2x - k) \tag{2}$$

$$\psi(x) = \sum_k g_k \phi(2x - k) \tag{3}$$

$$\tilde{\psi}(x) = \sum_k \tilde{g}_k \tilde{\phi}(2x - k) \tag{4}$$

其中 $h_k, \tilde{h}_k, g_k, \tilde{g}_k$ 是小波基系数, $\psi(x)$ 是小波函数, $\phi(x)$ 是尺度函数.

一维小波变换非常适合于处理声音, 心电图等一维信号, 而对于图像和视频信号, 需要使用高维小波变换作为工具. 对于高维信号的每一维, 分别采用一维小波变换进行计算, 可以得到高维可分小波变换. 高维不可分小波变换是一种非线性的小波变换. 本文只讨论高维可分小波变换在视频图像序列编码中的应用.

将一维小波变换分别应用在图像等二维信号的行和列上可构成可分二维小波变换. 行变换将图像分为高频和低频两个部分, 通常, 左边的是低频部分, 用 L 标记, 右边的是高频部分, 用 H 标记. 列变换将这两个部分再分别分为两个部分, 上面的是低频部, 下面的是高频部分. 这四个部分可用 LL, LH, HL 和 HH 标记, 其规则是前面符号表示行变换的波段, 后面符号表示列变换的波段. 将小波变换后的低频部分 (大小是原图的 1/4) 看成一幅新图像, 可对它再次做小波变换. 重复这个过程, 直到图像的低频部分的数据数目不够进行小波变换为止. 为标记不同的变换次数, 可在 LL, LH, HL, HH 后加一个数字表示, 如 HH1 表示经一次变换的行列高频段, HL3 表示经三次变换的行高频列低频段. 小波变换后的图像, 能量 (图像的信息) 大都集中于变换后的低频部分, 也就是 LL 波段. 小波变换的高频部分包含图像中的高频信息, 如物体的边缘等. 经 3 次二维小波变换的图像波段分布如图 1 所示.

三维小波变换本质上与二维小波变换相同, 不同的是除分别在行 (x 坐标方向) 和列 (y 坐标方向) 上做一维小波变换外, 还要在 z 坐标方向上做一维小波变换. 经一次三维小波变换的三维信号被分成 8 个波段, 仿照二维小波变换的波段标记, 按照 x, y, z 坐标的顺

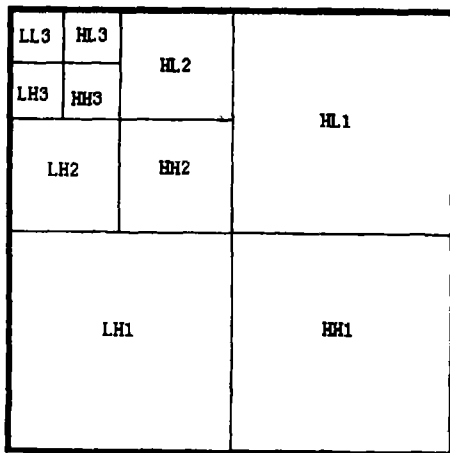


图 1 二维小波变换

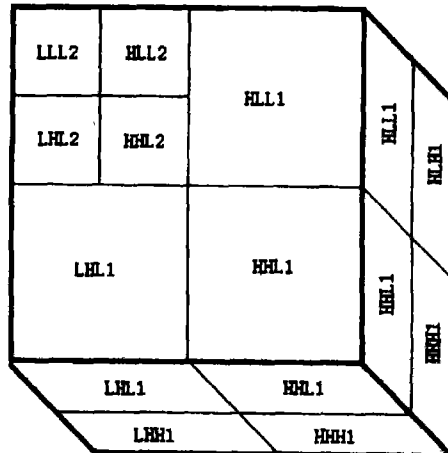


图 2 三维小波变换

序, 可用 LLL1, HLL2, HHH3 等记号来标记经不同次数小波变换后的不同波段。经 2 次三维小波变换后的波段分布如图 2 所示。

2.2 基于三维小波变换的彩色视频图像序列编码算法

2.2.1 彩色视频图像序列的三维小波变换 利用彩色视频信号的统计特性和人眼的视觉感知特性, 彩色视频图像序列的编码可用小波变换的方法来解决。彩色数字图像, 尤其是使用摄像机拍摄的彩色数字化自然图像, 常有比较大面积的变化平缓的区域, 如大面积的相似色块。这说明彩色数字图像的相邻像素间有很强的相关性, 称为帧内冗余。二维图像编码的目的就是要去除这些冗余。彩色视频信号的帧内统计特性和二维图像相同, 因而可采用二维图像的编码方法去除; 且通常相邻帧之间的变化也不大, 这是相邻帧间的相关性, 称为帧间冗余, 也可采用编码的方法去除^[5]。目前, 帧间冗余多采用基于运动矢量的运动估计和运动补偿的方法去除, 视频图像序列编码的国际标准普遍采用该方法。此外, 人眼对于高频信号的敏感度远低于对低频信号的敏感度, 对色度和饱和度的敏感度低于对亮度的敏感度, 因此在有损压缩中, 常常适当地舍弃部分高频信号和色度, 饱和度信息以提高编码效率^[6]。

用摄像机拍摄得到的彩色数字化视频图像序列一般用 RGB 三个色彩频道表示, 它的每个像素都由 RGB 三个分量组成, 一种彩色视频图像序列的编码方法是直接将灰度视频信号的编码方法分别应用在彩色视频信号的三个色彩频道上, 此方法可以直接从灰度视频信号的编码方法扩展而来, 因为不需做 RGB 色彩表示到 YCbCr 亮度色度表示的转换, 所以编码速度较快, 也比较直观, 但此方法没有充分利用人眼的视觉感知特性, 因而编码效率比较低。如果将 RGB 三个色彩频道转换为一个亮度频道 Y 和两个色度频道 Cb 与 Cr, 因人眼对亮度信号敏感, 对色度信号不敏感, 故可以对亮度频道做较精细的编码, 对两个色度频道做较粗糙的编码, 此编码方法可以得到比直接对 RGB 三个频道编码更高的压缩比或更高的信噪比。将 RGB 信号转换为 YCbCr 信号的公式为

$$Y = 0.299000 \times R + 0.587000 \times G + 0.114000 \times B \quad (5)$$

$$Cb = -0.168736 \times R - 0.331264 \times G + 0.500000 \times B \quad (6)$$

$$Cr = 0.500000 \times R - 0.418688 \times G - 0.081312 \times B \quad (7)$$

解码时, 需要将 YCbCr 表示的视频图像序列转换回 RGB 表示, 以便于视频图像序列的显示与输出。将 YCbCr 信号转换为 RGB 信号的公式为

$$R = Y + 1.402000 \times Cr \quad (8)$$

$$G = Y - 0.344136 \times Cb - 0.714136 \times Cr \quad (9)$$

$$B = Y + 1.772000 \times Cb \quad (10)$$

在将彩色视频信号的 RGB 表示转换成 YCbCr 表示后, 就可以分别对视频信号的每个频道做小波变换了。对于一个变化不太快的视频图像序列 (这一假设符合大多数摄像机拍摄到的彩色视频图像), 将视频信号中表示帧序的 t 坐标代换成 z 坐标后, 可把视频信号看成是三维空间中的“体”。仿照二维图像小波变换的方法, 将这个三维空间中的“体”沿 x , y , z 方向分别做一维小波变换, 亦即, 分别在行, 列和不同帧的相同位置的对应像素上做小波变换, 可构造三维可分小波变换。视频信号的三维小波变换及其波段分布如图 2 所示。对视频信号做一次三维小波变换可将原视频信号分成 8 个波段, 每次对变换所得的低频波段 (LLL 波段) 做一次变换又可得到 8 个新波段。对 LLL 波段重复做小波变换, 直到在 x , y , z 中任何一个方向上的系数个数不够为止。这样, 对视频信号做 N 次三维小波变换可得 $7 \times N + 1$ 个波段。经不同变换次数的波段内的系数数目是不同的, 第 $i + 1$ 次变换后, 每波段的系数数目是第 i 次变换后每波段系数数目的 $1/8$ 。

由于视频信号可能含有很多帧,如一段电视节目信号可能有数十万到数百万帧,若将这些帧存储到内存中编码数据量太大,因而是困难的,不经济的。另外,这样做还可能造成的编码延迟也过大。因此我们将视频信号分成许多块来编码。如果每块为 $h \times v \times f$ 个像素,这里, h 为水平像素数, v 为垂直像素数, f 为帧数,那么每次只须取 f 帧,缓冲区大小应为 $M \times N \times f$,这里 M 和 N 是一帧图像的宽和高, f 是帧数。试验表明,使用不同大小的块效果不同。使用较小的块,如 $8 \times 8 \times 8$,由于像素间相关性很强,可获得较高的信噪比,但用于量化和比特分配的时间较长,编码速度较慢;使用较大的块,如 $32 \times 32 \times 16$ 或更大,由于像素的相关性变弱,信噪比降低,但量化和比特分配的速度较快,因而编码速度较快。 $16 \times 16 \times 8$ 的块是一个较好的选择。因此我们取水平 16 像素,垂直 16 像素,8 帧。

对 Y, Cb, Cr 等亮度色度频道可采用不同大小的块。因为人眼对亮度信号敏感,所以可为亮度频道 Y 选取比较小的块,以有利于提高重建图像的主观质量;因为人眼对色度信号比较不敏感,所以可为色度频道 Cb 和 Cr 选取比较大的块,以有利于提高编码效率。通常可将色度块的大小选为亮度块的 2 倍或 4 倍。若色度块宽度为亮度块宽度的 2 倍,块的高度与帧数都相同,此时色度块大小为亮度块大小的 2 倍;若色度块宽度高度均为亮度块宽度高度的 2 倍,块的帧数相同,此时色度块大小为亮度块大小的 4 倍。可以对亮度块和色度块任意选取,比如,除 2, 4 倍这两种倍数外,还可以将色度块大小选取为亮度块大小的 3 倍或其它倍数,只要在编码前给定初始块大小即可。比 4 倍更高的倍数在实践中证明效果并不好,可能会对重建图像的质量有较大的不利影响。

小波基的选择^[7]可显著地影响编码器的性能, Haar 小波较适于处理有很多边缘和高频信息的视频信号,另外一些小波则较适于处理比较平滑的视频图像。理论上,对于不同的视频信号应自适应地选用最优小波基来做变换,以达到最好的处理效果。但实际上,根据图像和人眼的特性自适应地选取最优小波基是一个相当难的问题,现在还没有很好解决。为了简化编/解码器的构造,我们选取 Antonini 9-7 小波。9-7 小波在图像处理中很常用,主要是因为计算较简单,对高低频段的总体性能也比较好,是一种综合性能较优的小波。

彩色视频信号的各亮度色度频道经三维小波变换后,能量主要集中于其低频波段,这些低频波段对该信号的视觉效果影响最大。将这些低频波段和高频波段按不同的精度量化并进行熵编码,可以达到去除帧内和帧间冗余,压缩数据的目的。这样,直接利用三维小波变换的方法就能找出视频信号的帧内和帧间冗余,并同时分离了高低频信号,再根据人眼的视觉特性将其中不太重要的,对视觉效果影响不大的部分舍弃,就达到了编码的目的。

2.1.2 量化, 比特分配策略和熵编码 小波变换后的系数通常是含有小数部分的浮点数,不能直接编码成比特流的形式,必须先把这些系数量化为整型数。常用的量化方案有标量量化和矢量量化。矢量量化的效率较高,但计算复杂。为了降低编/解码器的实现难度并加快执行速度,我们采用了比较简单的标量量化^[8]方案。小波变换系数按不同波段量化精度不同,精度的确定由比特分配策略决定。不同波段的系数量化精度不同,但同一波段内的不同系数量化精度相同。量化阶均匀。

我们采用的比特分配策略是:将每一波段按 K 种不同精度量化,并同时计算出每种不同量化精度下的失真,得到码率/失真比;根据码率/失真比,在给定的压缩比限制下,计算每一波段对整帧图像的码率和失真贡献,循环调整每一波段的量化精度,直到找到每一波段的最优量化精度。在本文中, K 的取值为 10。在量化和比特分配阶段也可对不同的频道做不同的处理,如对亮度频道 Y 量化应选小的量化阶,以有利于提高重建图像的主观质量;对色度频道 Cb 和 Cr 量化选较大的量化阶,以有利于提高编码效率。实验中,我们并不直接选定各频道的量化阶,而是通过为不同频道选择不同的比特率来确定各频道的量化阶大小,如我们选择亮度频道 Y 和两个色度频道 Cb, Cr 的比特率比为 2:1:1, 3:1:1 或 4:1:1, 这样,两个色度频道的比特率是相同的,而亮度频道分配的比特数是色度频道的 2, 3 和 4 倍。通过这样的比特分配策略,我们可以达到亮度信号精细量化,色度信号粗糙量化的目的,既提高了编码效率又不会使重建图像质量损失太大。应当注意,通过不同频道比特分配的方法与通过改变亮度块和色度块大小之比的方法都可以提高色度块的编码效率,实际上

总的压缩比是它们的乘积, 也就是, 若亮度色度块比特分配比是 2:1:1, 色度亮度块大小之比是 2:1:1, 那么色度块的压缩比是亮度块压缩比的 4 倍。我们在实验中选择亮度和两个色度的块大小相同而比特分配比不同的办法来提高色度频道的编码效率, 这是因为各块大小相同的编 / 解码器实现上比较方便, 编码时改变比特分配比只要改动一个参数即可。

量化精度确定后就可对每一波段的系数按求得的最优量化精度量化。

为了获得最大的编码效率, 量化后的系数还需要做熵编码。熵编码一般使用游程编码 (RLE) 加 Huffman 编码的方法, 如 MPEG 等国际标所采用的, 也可以采用算术编码。本文的编 / 解码器采用了算术编码方案。

3 试验结果

实现上述算法的编码器的流程如图 3 所示, 解码器的流程如图 4 所示。在这个编 / 解码器中实现了彩色视频信号数据的读写, RGB 分量到 YCbCr 分量的转换, 三维小波变换, 量化, 比特分配, 熵编码等功能。

我们采用国际上常用的视频信号——Miss America 作为试验使用的视频信号, 它被广泛用于测试 H.261, H.263 和 MPEG 等国际标。Miss America 彩色视频图像序列的画幅为 360×288 , 帧数为 32。试验结果见表 1, 原帧图像和重建的帧图像见图 5。

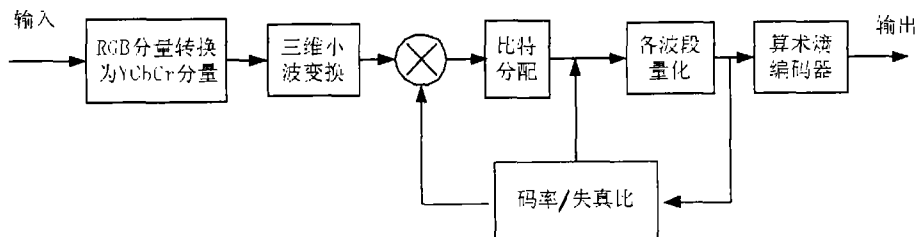


图 3 编码器的流程图

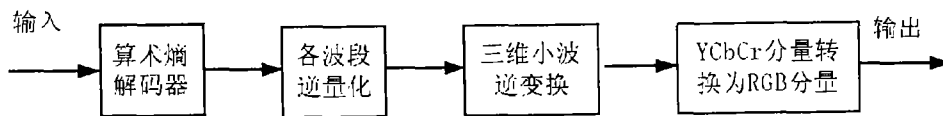


图 4 解码器的流程图

表 1 Miss America 视频图像序列的压缩比与信噪比

图号	亮度块压缩比	色度块压缩比	平均压缩比	均方误差 (MSE)	信噪比 (PSNR) dB
图 5(b)	50	100	75	17.7538	35.6379
图 5(c)	150	300	225	78.8920	29.1605
图 5(d)	150	450	270	198.6850	25.1492

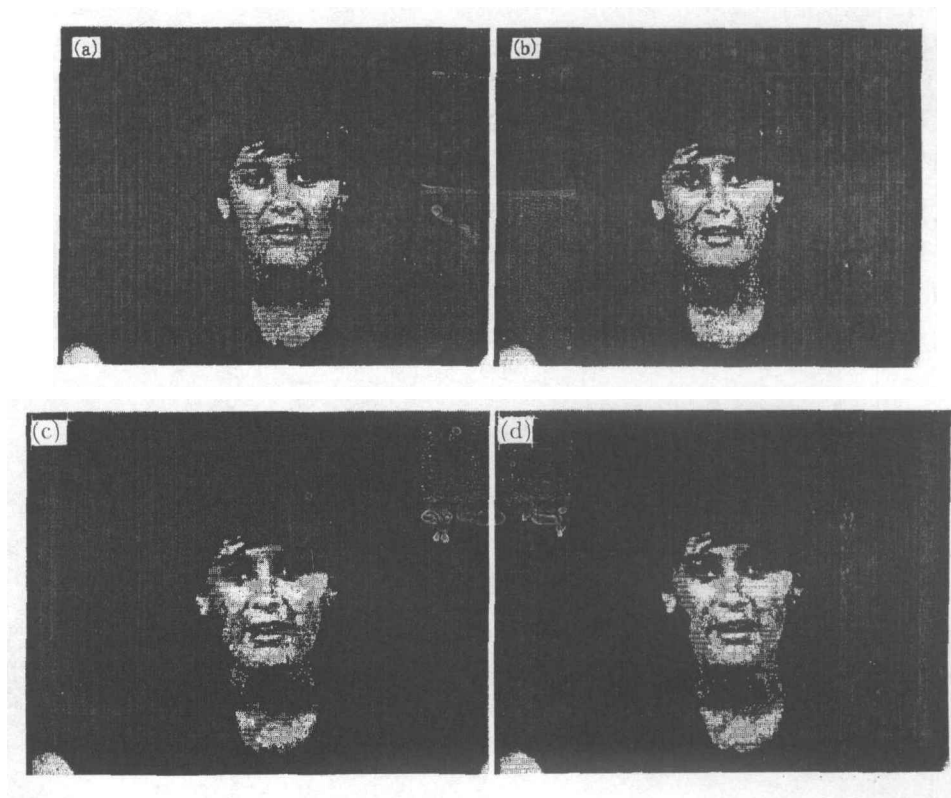


图 5 Miss America 的原视频帧与压缩后重建的视频帧
 (a) Miss America 的原视频帧
 (b) 平均压缩比为 75 时的重建视频帧
 (c) 平均压缩比为 225 时的重建视频帧
 (d) 平均压缩比为 270 时的重建视频帧

4 结 论

试验表明, 本编码算法可以达到较好的压缩效果: 如图 5 所示, 当平均压缩比分别为 75, 90 和 100 时, 重建的视频图像序列帧几乎与原帧一样, 很难用肉眼分辨出来; 当平均压缩比分别为 150 和 180 时, 重建的帧稍有模糊, 平均压缩比为 180 时有轻微的偏色 (重建帧色彩与原帧色彩不符) 现象; 当平均压缩比分别为 225 和 270 时, 重建的帧有较易分辨的模糊, 分块现象和偏色现象。我们的分析表明, 模糊和分块现象是由高压压缩比时舍弃的高频信息过多造成的; 偏色是由高压压缩比时舍弃的色度信息较多造成的。总的来讲, 重建的视频图像帧的质量还是比较好的, 编码效率也令人满意。

参 考 文 献

- [1] R. Koenen, Overview of the MPEG-4 standard, International Organization for Standardization, ISO/IEC JTC1/SC29/WG11, Coding of moving pictures and audio, ISO/IEC JTC1/SC29/WG11 N2725, March 1999/Seoul, South Korea
- [2] C. K. Chui, An Introduction to Wavelets, Academic Press, Inc., USA., 1992, 66-229.
- [3] W. Sweldens, The construction and application of wavelets in numerical analysis: Ph.D. Thesis, Belgium: Department of Computer Science, Katholieke Universiteit Leuven, May 25, 1994, 11-191.

- [4] G. M. Davis, A. Nosratinia, Wavelet-based image coding: An overview, *Applied and Computational Control, Signals and Circuits*, 1998, 1(1), 11-55.
- [5] M. L. Hilton, B. D. Jawerth, A. Sengupta, Compressing still and moving images with wavelets, *Multimedia Systems*, 1994, 2(3), 1-19.
- [6] 伍鹏, 张妙兰, 李象霖, 基于分块三维小波变换的视频图像序列编码方法的研究, *电子科学学刊*, 1999, 已接收, 即将发表.
- [7] J. D. Villasenor, B. Belzer B, J. Liao, Wavelet filter evaluation for image compression, *IEEE Trans on Image Processing*, 1995, Preprint
- [8] A. C. Popat, Scalar quantization with arithmetic coding, M.S. Thesis. Boston: Department of Electrical Engineering, Massachusetts Institute of Technology, 1990, 18-82.

AN ALGORITHM FOR CODING COLOR VIDEO SIGNAL USING 3D-WAVELET TRANSFORMATION

Wu Peng Zhang Miaolan Li Xianglin

(Dept. of Electron., Graduate School, Univ. of Sci. and Tech. of China, Beijing 100039, China)

Abstract This paper presents an algorithm for coding color video signal using 3D-wavelet transformation. When the 3 color channels of RGB of the video signal are converted to color channels of YCbCr, and the frame order t of the signal is replaced by order z , this signal can be considered as 3 blocks in a 3D-space. Each block presents a color channel. After splitting each block into smaller sub-blocks, imitate the method of 2D-wavelet transformation for image, the sub-blocks can be transformed with 3D-wavelet. Most of energy of each color channel is in the transformed low-frequency bands. These bands affect the visual quality of the video signal most. After quantizing different bands with different precision and entropy coding each band, the intra- and inter-frame redundancy of the color video signal can be eliminated and the data can be compressed. Our simulation experiments show that this algorithm can obtain very good results.

Key words Wavelets, 3D-wavelet transformation, Color video signal coding

- 伍 鹏: 男, 1972 年生, 在职研究生, 主要研究方向为: 预测控制的理论与应用, 小波变换及其应用, 图像与视频信号编码, 音频信号编码, 数字信号处理, 数字图像处理等
- 张妙兰: 女, 1940 年生, 教授, 主要从事数字信号处理的教学与科研, 感兴趣的科研领域是信号处理, 数字图像处理和计算机视觉.
- 李象霖: 男, 1938 年生, 教授, 主要从事数字图像处理, 三维运动分析的教学与科研, 感兴趣的科研领域是图像处理, 图像序列分析与计算机视觉.