

基于运动补偿的三维小波视频编码

林刚 刘泽民

(北京邮电大学电信工程学院 57#, 北京 100876)

摘要 提出了一种新的基于运动补偿的三维小波视频编码方案。通过对原始图象序列沿着运动轨迹进行时间维小波分解以及空间上的二维小波分解,得到不同的时间-空间三维频率子带。然后,将这些子带中的小波系数构成三维方向等级树结构,并采用改进的SPIHT零树编码算法进行压缩。实验表明,此方法不仅提高了视频编码效率,而且易于进行码率控制,以及实现时间、空间分辨率上的可伸缩编码。

关键词 视频编码,小波变换,运动补偿

1 引言

近些年来,随着网络与通信技术,尤其是ATM&B-ISDN的飞速发展,通信应用领域不断拓展,新的多媒体通信业务不断出现,如可视电话、电视会议、远程医疗及点播电视等。这些综合业务包括数据、话音和视频,其中视频业务尤为重要。正因为现实生活中的业务量需求不断增长,而网络带宽是有限的,所以这给多媒体数据压缩编码提出了更高的要求。信源编码,尤其是视频压缩已成为ATM&B-ISDN网中的一项关键技术。目前,世界各国的许多专家都在致力于此领域的研究。

小波分析,这门十年前才发展起来的新兴学科,在图象压缩方面显示了很好的前景。它克服了传统的FFT的一些弱点,能对图象进行时频局部化,对图象的平滑部分进行粗处理,对细节部分进行精细处理。同时,它也能将图象分解到多个尺度上,进行多分辨率分析。这样,就可根据各尺度上的子图象的特性不同,而进行不同的处理,用此压缩方法所重建的图象更符合人眼的主观特性,彻底消除了块效应,图象的编码效率也得到了很大提高,因此,小波编码备受众多研究者的青睐。对于活动图象(视频),要想进行更充分地压缩。一方面要进行帧内编码,消除每帧图象的空间冗余,另一方面还要进行帧间编码,即

解除各帧图象在时间轴方向的相关性,以便进一步压缩。C. I. Podilchuk等人提出了三维小波变换(3DWT)^[1],将图象序列看成三维信号,在水平、垂直及时间方向分别进行小波变换,将原始图象序列分解到不同的时频子带中。由于小波变换系数是相互独立的,这样就可消除三方向上的线性冗余,然后根据各时频子带的特性分别进行压缩。

实验表明,对于那些活动性不强的图象序列,如可视电话业务,三维小波变换具有很好的压缩效果。而对于那些纹理较复杂且活动性强的图象序列,此方法的压缩效率不很理想。这是由于此类图象序列的帧与帧之间的空间相关性较弱,沿时间轴方向作小波变换起不到多大作用。事实上,活动图象的帧与帧之间不仅存在基于象素的线性相关,还在宏观上存在着很大的运动相关性,即后一帧图象总是前帧图象经过平移、缩放、旋转等各种运动而得来的。为了充分利用图象序列的运动信息,消除冗余,必须采用运动补偿技术,以提高视频的压缩效率。

正是基于以上的考虑,本文提出了一种新的基于运动补偿的三维小波编码方法(3DMCWT)。首先对原始图象序列进行运动估值,获得运动矢量,然后沿着运动轨迹进行时间方向的小波变换,从而将视频信号分解到不同的时间子带,并对各子带中的变换图象帧进行水平和垂直方向上的二维小波变换,

以消除空间相关性。针对各个时间—空间频率子带, 本文提出了一种改进的三维零树编码方法(3D SPI-

HT)。总体算法框图如图 1 所示。以下分别对上述各部分进行介绍。

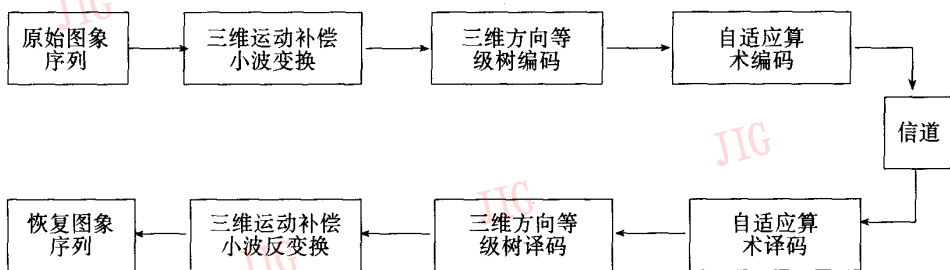


图 1 编码系统框图

2 三维小波变换

变换编码是通过将原始图象数据的能量集中, 用较少的比特数来表征原数据, 以实现压缩。目前, 离散余弦变换(DCT)已被许多国际压缩标准采用, 如 JPEG、MPEG^[2]、H. 261 及 H. 263。然而近些年来, 小波变换(DWT)被广泛引起重视, 使编码效率得到了很大改善, 这得益于小波变换的时频局部化、多分辨率等特性。本文就是采用三维小波变换方法来实现视频压缩。对原始图象序列进行运动估值, 沿运动轨迹进行时间维小波变换, 然后对分解后的各子带进行二维空间小波变换。

2.1 沿轨迹的时间维小波变换

对于不同的图象序列, 其运动的形式各不相同, 包括平移、缩放和局部物体运动等, 如图 2(a)所示。在进行运动估值时, 我们采用块匹配方法。假设原始图象序列中有 E、A、B 三个相邻帧(如图 2), 先将 B 图象划分为大小为 $I \times J$ 的不重叠块。对于 B 中的左上角坐标为 $(i \times I, j \times J)$ 的块, 求运动矢量 $[k, l]^{(i, j)}$ 的算法公式如下:

$$[k, l]^{(i, j)} = \arg \min_{[k, l] \in \Omega} \sum_{m=i \cdot I}^{(i+1) \cdot I-1} \sum_{n=j \cdot J}^{(j+1) \cdot J-1} d(B(m, n), A(m+k, n+l)) \quad (1)$$

其中, $d(\cdot, \cdot)$ 是某种误差准则, 如均方误差、最小绝对值和等。 Ω 为搜索范围, 可取为 $[2 \times I, 2 \times J]$, $B(m, n)$ 表示图象 B 中坐标为 (m, n) 的点的象素值, $A(m+k, n+l)$ 表示图象 A 中坐标为 $(m+k, n+l)$ 的点的象素值。

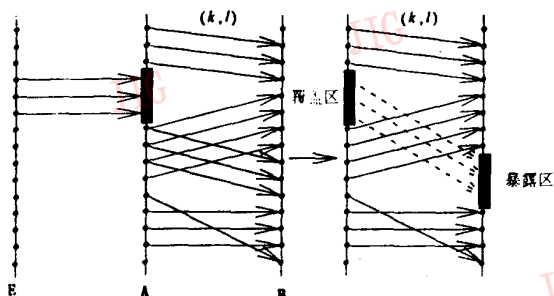
从图 2(a)可见, 图象 A 中存在着某些象素不与 B 中的象素相对应, 称之为“覆盖”区, 而在图象 B 中, 存在着多个象素同时对应于图象 A 中的某一个象素, 为了便于进行小波变换, 须将它变为一一对应

的形式, 因此, 图象 B 中也会存在一些与 A 中象素不连接的点, 称之为“暴露”区。如图 2(b)所示。

在连接区(除暴露区外), 可根据前述方法所求的运动矢量, 求出 B 帧的运动估值 B' :

$$B'(i, j) = A(i+k(i, j), j+l(i, j)) \quad (2)$$

其中, $(k(i, j), l(i, j))$ 表示图象 B 中点 (i, j) 的运动矢量。



(a) 图象 A、B 可能的运动矢量 (b) “覆盖区”和“暴露区”

图 2

欲对图象帧 A、B 沿运动轨迹方向作小波变换, 就可转化为对 B 和 B' 作变换。选取 Haar 函数作为小波函数, 其对应高通滤波器系数为 $[\sqrt{2}/2, -\sqrt{2}/2]$, 低通滤波器系数为 $[\sqrt{2}/2, \sqrt{2}/2]$ 。假设所求的低、高通小波系数分别记为 $L(i, j)$ 和 $H(i, j)$ 。则:

在连接区:

$$\begin{cases} L(i, j) = \frac{\sqrt{2}}{2} (B(i, j) + B'(i, j)) \\ H(i, j) = \frac{\sqrt{2}}{2} (B(i, j) - B'(i, j)) \end{cases} \quad (3)$$

实际上, $L(i, j)$ 就是图象 B 和 B' 的均值, $H(i, j)$ 对应 B 与 B' 的差值。

在暴露区,另行处理如下:

$$\begin{cases} L(i, j) = B(i, j) \\ H(i, j) = E(i_1, j_1) - A(i_1, j_1) \end{cases} \quad (4)$$

其中, (i_1, j_1) 点对应 A 帧的覆盖区中的点。之所以取高通为 E 与 A 在 (i_1, j_1) 点的差值, 而不是简单地取为 $A(i_1, j_1)$, 是因为我们认为 A 帧的覆盖区中的点是由前一帧 E 中相同位置的点而来, 事实上, 它们的确存在着相关性。这样, 就可减小高通部分的能量, 有利于进行压缩。由于 A 、 B 两帧的连接区的点是一一对应的, 且它们的象素总数相同, 所以 B 帧中的暴露区的任何一点总能从 A 中的覆盖区找到一点与之对应。对应的原则是: A 、 B 两帧都同时从左到右、从上到下的扫描顺序进行对应, 如图 2(b) 中的虚线所示。图 3 是 Mobile&Calender 序列的第 5、6 帧原图、帧间差以及沿运动轨迹的小波变换后的高通图。

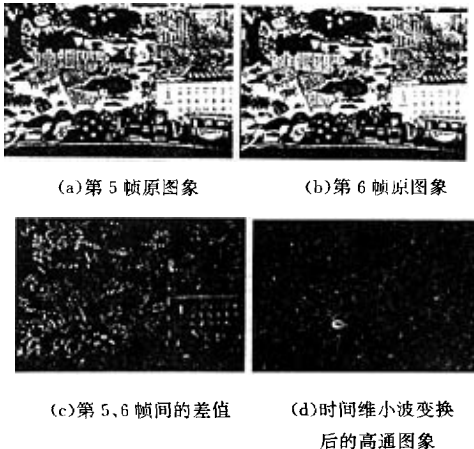


图 3

同理, 进行小波合成时, 对连接区和覆盖区中的点分别进行处理:

在连接区, 先根据 $L(i, j)$ 、 $H(i, j)$ 求得 B 及其估值 B' , 公式如下:

$$\begin{cases} B(i, j) = \frac{\sqrt{2}}{2} (L(i, j) + H(i, j)) \\ B'(i, j) = \frac{\sqrt{2}}{2} (L(i, j) - H(i, j)) \end{cases} \quad (5)$$

上式与 (3) 相对应, 然后根据运动矢量 $(k(i, j), l(i, j))$, 由 B' 重建出图象 A , 即

$$A(i+k(i, j), j+l(i, j)) = B'(i, j) \quad (6)$$

在覆盖区, 与式 (4) 相对应, 由重建出的 E 帧图象及 B' , 可恢复出图象 A :

$$A(i_1, j_1) = E(i_1, j_1) - B'(i, j) \quad (7)$$

A 中的点 (i_1, j_1) 与 B 中的点 (i, j) 相对应的原则与前述相同。

2.2 二维空间小波变换:

将原始图象序列的每 N 帧划分为一个图象组 (GOP), $N=2^n, n=0, 1, 2, \dots$ 。我们取 $N=8$, 则可进行 $\log_2 N=3$ 级时间维小波分解。对每一级, 先进行沿轨迹小波变换, 得到高、低两个频带。然后, 对变换后的各子带的每一帧进行水平、垂直方向二维小波变换。这样, 经 3 级分解后, 共得到 H、LH、LLH 和 LLL 4 个时间频带, 以及 22 个时间-空间三维子带, 编号如图 4 所示。时间维小波采用 Haar 函数, 空间方向采用 Daubechies26 小波函数^[3]。小波变换包括滤波和抽样两个过程, 所以每进行一级变换时, 沿时间轴图象帧数减半; 在二维空间上每帧图象尺寸减为四分之一, 参见图 4。

3 三维零树编码

一般地, 绝大多数图象的低频成分能量较高, 而小波变换能将图象数据分解到不同的尺度中, 构成从大尺度到小尺度的金字塔型结构。系数的幅值从大尺度到小尺度逐渐减小。而且, 同一空间方向 (水平、垂直和对角) 上的系数具有自相似性。正是利用小波变换的这一特性, Shapiro 提出了嵌入式零树小波算法 (EZW)^[4], 随后, Said 和 Pearlman 又对它作了改进, 提出了基于空间方向等级树编码 (SPIHT)^[5], 使该方法的编码效率进一步提高。

本文对 Said 和 Pearlman 的 SPIHT 编码算法作了改进, 并将它从二维推广到三维, 称之为“三维方向等级树 (3D SPIHT)”编码。以下只简单阐述所作的改进, 欲了解详细内容, 请参阅文献 [5]。

如前所述, 对原始图象序列进行三维小波变换后, 得到如图 4 所示的子带图。将这些子带 (除 a_1 外) 根据不同的空间方向 (水平、垂直和对角) 构成如下的等级树结构 (图 5):

图 5 中, a_2 表示图象帧 a 中的子带 2, 参考图 4。其它符号含义依此类推。可见, 不同级中的子带块的大小不同, 上一级子带中的点 (称为父节点) 对应着下一级同方向子带中的四点 (称为子节点)。分别对此 7 个等级树单独采用 SPIHT 算法^[5]进行编码, 合成为同一码流。实验表明, 原始图象经过小波变换

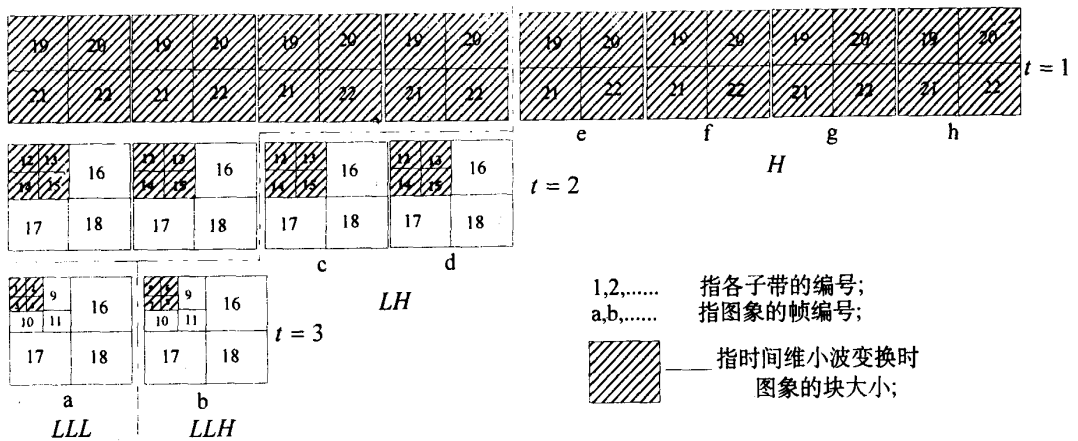


图 4 三级时间-空间小波变换后的各子带及其编号

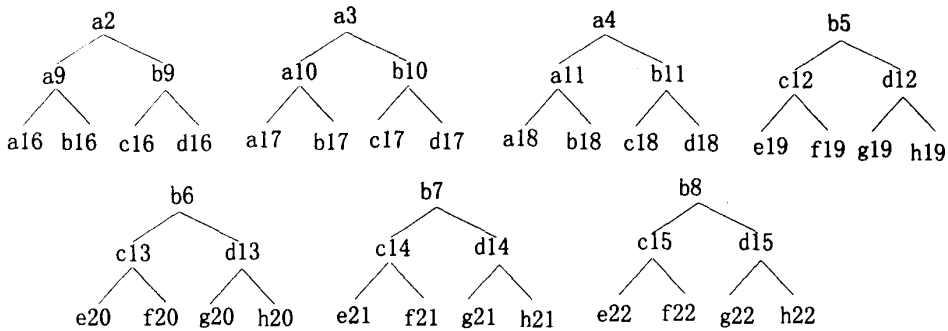


图 5 三维等级树结构图

后,各个空间方向上的小波系数的最大值不尽相同,因此,我们针对各个方向的三维等级树分别进行初始化、编码。而 Said&Pearlman 算法是将不同方向的等级树合并在一起进行初始化、编码,这无疑会增加不必要的码流。图 6 是本算法与 SPIHT 算法在二维时的仿真结果的比较,所用的原始图象为 Claire(SIF 360×288格式)标准图象。可见,我们的方法可大大提高编码效率

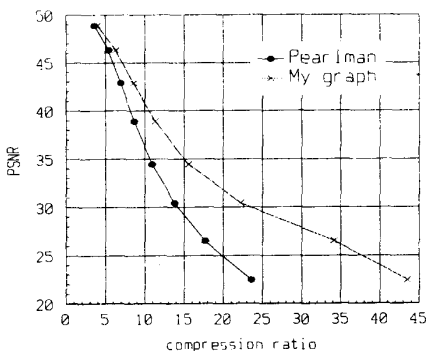


图 6 本算法与 Said&Pearlman 算法的信噪比-压缩比曲线

4 实验结果及结论

仿真实验采用 SIF 352×240 格式的 Mobile&Calendar 图象序列,此序列被认为最复杂、最不易压缩。为了更好的评价本文所提出的 3DMCWT+3D SPIHT 方法,我们将实验结果与三维小波变换而不进行运动补偿方法(3DWT)方法结果进行比较。在进行运动补偿时,采用的分块大小为 8×8,搜索范围是 16×16。图 7 是该序列的前 3 个 GOP(即第 1~24 帧)平均峰值信噪比(Average PSNR)-平均比特率(Average Bit Rate)曲线的比较。从图中可以看出,本文所提出的方法提高了编码效率。

在本方法中,运动补偿采用的是块匹配技术,它较适合于图象的平移运动,因此,若能采用其它更合适的运动模型,进行更精确的运动估值,例如面向对象技术,无疑将会进一步提高编码效率。另外,本方法所得到的码流是嵌入式的,易于进行码率控制。而且它将视频信号分解到多个时间-空间频率子带中,因此在时间、空间分辨率水平上进行可伸缩编码

是可行而有意义的。这些是我们今后要研究的内容。

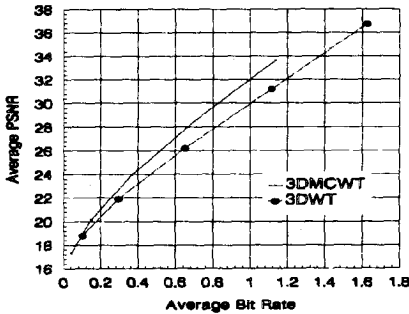


图7 本方法与3DWT方法的平均PSNR—Bit Rate曲线的比较

参考文献

- 1 Podilchuk Christine I, Jayant Nikil S, Nariman Farvardin. Three Dimensional Subband Coding of Video. *IEEE Trans. Image Processing*, 1995, 4(2):125~139.
- 2 LeGall D J. The MPEG video compression algorithm. *Signal Processing, Image Commun.*, 1992, (4):129~140.
- 3 Villasenor John D, Benjamin Belzer, Judy Liao. Wavelet Filter Evaluation for Image Compression. *IEEE Trans. Image Processing*, 1995, 4(8):1053~1059.

- 4 Shapiro J M. Embedded image coding using zerotrees of wavelets coefficients. *IEEE Trans. Signal Processing*, 1993, 41: 3445~3462.
- 5 Amir Said, Pearlman. William A. A New, Fast, and efficient Image Codec Based on Set Partitioning in Hierarchical Trees. *IEEE Trans. Circuits and Systems for Video Technology*, 1996, 6(3):243~249.
- 6 Chen Ying Wei, Pearlman William A. Three Dimensional Subband Coding of Video using the Zero-Tree Method. *SPIE*, 1994, 2727: 1302~1309.



林刚 1996年7月毕业于中国矿业大学北京研究生部,获工程力学专业硕士学位,现在北京邮电大学攻读电路与系统专业博士学位,主要研究方向是:视频编码,ATM网中的关键技术。



刘泽民 1951年毕业于上海交通大学。北京邮电大学教授、博士生导师,从事网络理论、数字信号处理神经网络、模糊系统的研究。目前,主要领导863项目、邮电部重点科技专题项目——智能型ATM的研究。

3D Wavelet Video Coding with Motion Compensation

Lin Gang, Liu Zemin

(Beijing University of Posts and Telecommunications, Radio Engineering, Beijing 100876)

Abstract In this paper, we present a new video coding scheme based on three dimensional wavelet transformation with motion compensation (3DMCWT). First, we apply 1-D temporal wavelet decomposition along the motion trajectory and 2-D spatial wavelet decomposition, and obtain spatio-temporal multiresolution representation of the original video signal. Then, we construct several 3-D orientation hierarchical trees for the wavelet coefficients in all subbands, and encode them with an extension of SPIHT algorithm^[7] introduced by Said and Pearlman (3D SPIHT). Our simulations demonstrate that this method is very efficient. What's more, it can provide both multirate and spatio-temporal multiresolution video scalability.

Keywords Video coding, Wavelet transform, Motion compensation