

基于小波的视频空间可分级编码研究^{*}

Research on Spatial Scalable Video Coding Based on Wavelet

王相海

(辽宁师范大学计算机与信息技术学院 大连116029)

(南京大学计算机软件新技术国家重点实验室 南京210093)

Abstract New applications in a multiresolution environment arise the need of video scalable coding. The video scalable coding based on wavelet transformation has received much attention in recent years. Based on a thorough examination of current research works on motion estimation and prediction, this paper gives a kind of video spatial scalable coding scheme based on a kind of indirect approach, which uses the motion prediction on lowpass image at the next finer resolution to instead of predicting on the highpass subbands at a given resolution, to avoid the motion tracking errors in high frequency bands. Simulation results show it is effective.

Keywords Wavelet transform, Motion estimation, Spatial scalable coding

1. 引言

在许多实际应用中,由于用户要求不同、终端能力不同、异构网络的不同支路所能提供的 QoS 不同、或网络传输条件的变化(比如噪声、拥塞等)等原因,需要提供不同质量的视频信号,比如视频会议、视频点播、多媒体数据库浏览等等,解决此类问题的最好方法是用单个编码器产生分层次的压缩码流,对不同层次的码流解码可以获得不同的图像质量,这种编码机制通常被称为可分级编码(Scalable Coding)。90年代,国际标准化协会(ISO)等领导制定了 MPEG 多媒体编码标准,其中的 MPEG-2和 MPEG-4中便体现了视频的可分级编码思想,即支持空间可分级编码、时间可分级编码和数率(或信噪比)可分级编码^[1~3]。然而,目前所采用的这种可分级编码方法还存在着一定的局限性,具体表现在两个方面:一是标准中对视频信息进行分层表示的“层数”是有限的,通常不能实现解码器端更细粒度的解码;二是视频标准中大都采用二维 DCT 方法实现帧内和预测帧差的编码,而固定大小分块的 DCT 通常不能很好地适应空间可分级编码的需求。此外,分块 DCT 编码所产生的“块效应”往往会使数率可分级编码不能取得较好的视觉效果,特别对低码率情况。所有这些促使人们去探索更加有效的视频可分级编码方法。

近年来,人们对基于小波变换的视频编码技术展开了积

极的研究,出现了一些较为有效的视频编码方法,主要有运动补偿的 2D 小波视频编码方法^[4,5]和 3D 小波视频编码方法^[6,7]。试验表明,单纯的 3D 小波变换在视频编码中效率不是很高,并且编码延迟较大,这样人们更看重基于运动补偿的 2D 小波视频编码方案,对该技术的研究已成为当前的热点问题。

本文首先介绍了视频可分级编码的基本概念,并对基于小波的视频运动估计方法进行了分析,在此基础上提出了一种基于下层 LL 子带分层运动估计的空间可分级编码方案,该方案有效地克服了一般高频子带间的运动估计误差问题,在一定程度上提高了解码端空间可分级解码图像的质量。实验结果验证了方法的有效性。

2. 视频可分级编码的基本概念

一个视频编码的比特流具有可分级性(Scalability)是指利用该比特流的一部分(或子集)可以产生对该视频的一个有用表示的能力,即解码器能够根据分辨率的情况对码流的一部分进行解码。视频的可分级编码通常有三种情况,即空间分辨率可分级编码、时间分辨率可分级编码和数率可分级(或称 SNR 可分级)编码。空间分辨率可分级是指以不同空间分辨率解码图像的能力(参见图1);时间分辨率可分级是指视频帧率的更新可以被调整(参见图2);而数率可分级则是指对一个

^{*} 南京大学计算机软件新技术国家重点实验室资助项目。王相海 博士,副教授,主要研究领域为计算机图形学、图像及多媒体信息处理。

一步运算都起到应有的作用。

选取了适当的结构元素后,我们再对图1的输入图像进行形态学变换,其开-闭运算结果如图6所示。

显而易见,只要结构元素选择合理,开-闭变换的结果不但可以达到而且要优于开-闭-闭-开变换的效果,无论直线段还是曲线段,其正、负脉冲噪声都基本被滤除干净,且图像变形很小,只是右下角处稍有些偏离,形态学变换的优良特性被充分发挥了出来。

结论 通过分析比较,我们得出结论,当结构元素真正实现与输入图像的良好匹配时,形态学会成为图像处理的有力

工具。我们在认真分析输入图像几何形状的基础上,根据本文提出的基本原则选取结构元素,取得了良好的图像处理效果。

参 考 文 献

- 1 章毓晋. 图像处理和分析. 清华大学出版社, 1999. 3
- 2 Michielsen K, De Raedt H. Integral-geometry morphological image analysis. Physics Reports, 2001, 347: 461~538
- 3 容太平, 夏玉洁. 形态滤波算法在油井测量数据处理中的应用. 华中理工大学学报, 2000, 28(5)
- 4 刘志敏, 杨杰, 施鹏飞. 数学形态学的细化算法. 上海交通大学学报, 1998, 32(9)

压缩的码流,用户可根据带宽或系统能力的具体情况在一定范围内按任意比率从码流的开始端来截取相应的一段码流进行解码(参见图3)。

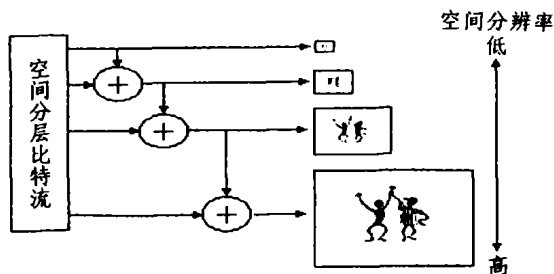


图1 空间可分级

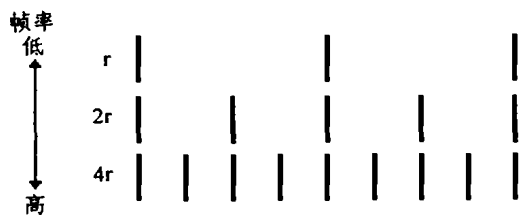


图2 时间可分级



图3 512×512 Lena 图像数率可分级

3. 基于小波的视频运动估计分析

视频运动估计作为去除帧间冗余的有效方法,直接影响到视频的编码效率和解码质量。在基于小波的视频编码方法中,视频序列的连续两帧经小波分解后,相应的高频子带间具有很少的“相似”之处,一般而言相互之间并不能通过运动估计得到。这是因为基于二下取样的小波图像分解关于平移操作是可变的,即二下取样操作与平移操作是不可换的。我们通过一个例子来说明这一问题。图4中a和b为包含简单周期信号的两帧,它们有一个像素的水平位移,图a1和b1分别是图a和b的多分辨率小波分解,我们通过对HL频带的二层分解可以看出,这两帧的HL频带相互间不能被预测,这是由于对两帧所进行的下取样操作丢失了一些像素的光亮度。文[9]中也从频率域的角度讨论了相类似的问题。

为了克服这些不利因素使得运动估计更加有效地结合到编码方案中,人们进行了积极的研究,J. Magarey等人提出了基于复值小波图像分解的运动估计方法^[10],该方法能保证以实数级精度获得相应的运动矢量,有效地克服了二下取样操作与平移操作的不可换问题,然而,复值小波变换的使用使得小波分解系数增加了4倍,这样很难使最后的编码有好的效果。为此人们提出了一些“折衷”方案,Q. Wang等人给出了一种分层预测的方案^[11],该方法的运动估计是在小波变换域上进行的,利用小波分解每一层中的LL子带的运动估计矢量作为该层其它子带的运动估计矢量进行估计,然而,该方法同

样存在着对各层高频子带的估计误差;P. Cheng等人提出了一种基于下层低频子带的分层运动估计方法^[12],取得了较好的效果,在我们所提出的视频空间可分级编码方案中,运动估计即采用了该方法的基本思想。

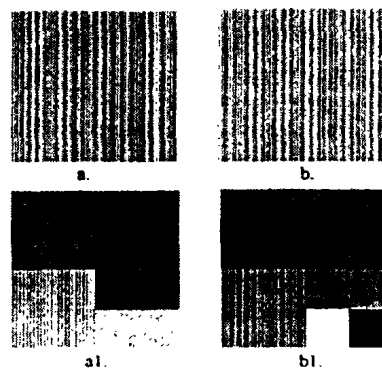


图4 带有一个像素位移的两个测试图像(a和b)以及相应的小波分解(a1和b1)

4. 一种基于小波的视频空间可分级编码方案

图5给出了所提出的空间可分级编码的一般结构。

4.1 奇次对称双正交小波滤波器

设奇次对称双正交小波滤波器中低通分解滤波器系数和高通分解滤波器系数分别为 $g(i)$ 和 $h(j)$;低通合成滤波器系数和高通合成滤波器系数分别为 $e(j)$ 和 $f(i)$,其中 $i=0, \dots, L_g-1, j=0, \dots, L_h-1, L_g$ 和 L_h 均为奇数。则有如下关系式成立:

$$\begin{cases} g(i) = g(L_g - 1 - i), i = 0, \dots, (L_g - 1)/2; \\ h(j) = h(L_h - 1 - j), j = 0, \dots, (L_h - 1)/2; \\ e(i) = (-1)^{i+1} h(i), i = 0, \dots, (L_h - 1); \\ f(j) = (-1)^j g(j), j = 0, \dots, (L_g - 1). \end{cases}$$

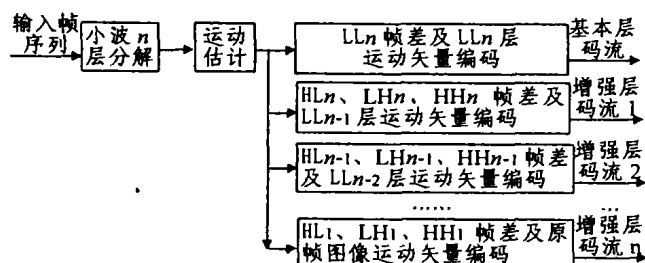


图5 一种空间可分级编码方案结构图

分解滤波过程如下:

$$T(i) = \sum_{k=0}^{L_g-1} x(i+k-(L_g-1)/2) g(L_g-1-k),$$

$$S(i) = \sum_{k=0}^{L_h-1} x(i+k-(L_h-1)/2) h(L_h-1-k).$$

其中 $T(i)$ 和 $S(i)$ 分别是二次抽样之前低通和高通滤波器的输出结果,对其进行二抽样即可得到分解的小波系数。二抽样既可在奇数位置进行也可在偶数位置进行,考虑到在算法中对各帧的边缘要进行对称延拓,我们采用了低通系数采样和高通系数采样交替进行的方式,即低通和高通小波系数分别为 $C(i) = T(2i-s)$, $D(i) = S(2i+1-s)$,其中当低通采样位置为偶数时 s 取0,否则取1。

为了进行小波系数合成,首先对系数进行如下二插值:P

$(2i-s)=C(i), P(2i+1-s)=0; Q(2i+1-s)=D(i), Q(2i+s)=0$, 其中 $P(k)$ 和 $Q(k)$ 分别是二插值低通和高通系数。则合成滤波器的重构信号为:

$$r(i) = \sum_{j=0}^{L_h-1} p(i-(L_h-1)/2+j) e(L_h-1-j) + \sum_{j=0}^{L_g-1} Q(i-(L_g-1)/2+j) f(L_g-1-j).$$

4.2 基于下层 LL 子带的分层运动估计方法

在文[11]的基础上,我们给出了基于下层 LL 子带分层运动估计的具体过程。整个过程首先对最低频子带进行运动估计,然后从最低频开始对各高频子带进行“间接”估计,即首先利用估计层的下一层低频子带进行运动估计获得相应的预测 LL 子带,然后对该 LL 子带进行一级小波分解,从而获得相应高频子带信息,具体参见图6。

在该运动估计过程中,对最低频子带的运动估计可按 MPEG 中基于分块的运动估计方法进行,搜索过程可采用“三步搜索”方法,块匹配原则可采用 MSE(最小均方误差函数)等等;此外,利用第 $i+1$ 层 LL 子带的运动矢量对第 i 层 LL 子带运动矢量的“提炼”可以采用以第 $i+1$ 层 LL 子带的运动矢量作为估计初始值,利用第 i 层 LL 子带的信息对该运动矢量进行精确调整。由于有一个较理想的初始值,因此可以使用相对较小尺寸的窗口进行搜索,这在一定程度上可以提高运行速度。

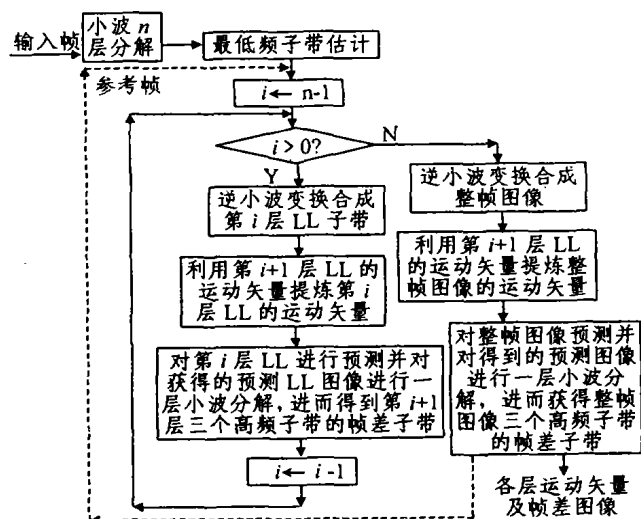


图6 基于下层 LL 子带的分层运动估计流程图

4.3 运动矢量及帧差的编码

在所提出的编码方案中,对各运动矢量的编码一般直接采用无损编码,比如算术编码;而对各帧差图像的编码可以采用一般的基于小波子带量化的编码方法。

实验结果与讨论 为了说明所提出方法的有效性,我们对一些视频序列进行了可分级编码实验,试验中所使用的滤波器为 Daubchies 的双正交 9/7 样条小波^[13],对各 LL 子带的运动估计采用了基于宏块(16×16)的交迭块(Overlapped Block)运动估计方法,所使用的“权值”参见文[14],运动矢量被限制在整数精度。

图7给出了对 Akiyo 序列的第249帧进行空间可分级解码的情况,图像被分解四层,其中图 a 是第249帧原图,图 b~e 是对该帧按空间分辨率为 1、1/2、1/4 和 1/8 解码的情况,其中对帧差采用了文[15]中的量化方案。全分辨率解码图像的

PSNR 为 35.18dB。

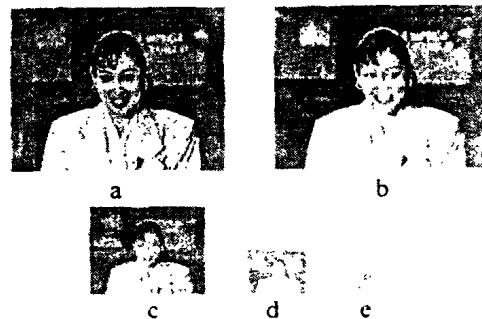


图7 a 图是 Akiyo CIF 序列的第249帧原图;b、c、d、e 分别是对该帧按空间分辨率 1、1/2、1/4、1/8 的解码图像

从试验结果可以看出,本文所提出的基于小波的视频空间可分级编码方案是有效的,它能够在一定程度上提高解码图像的质量,取得了较好的效果。然而,该方法也存在着一些不足,比如,在运动估计中由于需要重构相应帧图像的 LL 子带,这样从时间上要多占用一些;此外,由于采用了基于下层 LL 子带的运动估计,相应地增加了一些需要进行运动估计的块数,所有这些均有待于去寻找相应的最优“折衷”方案予以解决。

参考文献

- 1 ISO/IEC 13818-2, Generic coding of moving pictures and associated audio information. MPEG (Moving Pictures Expert Group), International Organization for Standardisation, 1994 (MPEG2 Video)
- 2 ISO/IEC JTC1/SC29/WG11. Overview of the MPEG-4 standard. N3536, Beijing, 2000, 7
- 3 王相海, 张福炎. MPEG-4 自然视频编码工具和算法. 计算机科学, 2001, 28(1): 1~7
- 4 Martucci S A, Sodagar I, Chiang T, Zhang Y O. A zerotree wavelet video coder. IEEE Trans. Circuits and Syst. Video Technol., 1997, 7(1): 109~118
- 5 Yang X-G, Ramchandran K. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation. IEEE Transactions on Image Processing, 2000, 9(5): 778~791
- 6 Taubman D, Zakhor A. Multirate 3-Subband coding of video. IEEE Trans. Image Processing, 1994, 3(5): 172~186
- 7 Khalil H, Atiya A F, Shaheen S. Three-dimensional video compression using subband/wavelet transform with lower buffering requirements. IEEE Trans. Image Processing, 1999, 8(6): 762~773
- 8 Sikora T. MPEG digital video-coding standards. IEEE Signal Processing Magazine, 1997, 14(5): 82~110
- 9 Yang X-G, Ramchandran K. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation. IEEE Transactions on Image Processing, 2000, 9(5): 778~791
- 10 Magarey J, Kingsbury N. Motion estimation using complex wavelets. SPIE, 2825: 674~685
- 11 Wang Q, Ghanbari M. Scalable coding of very high resolution video using the virtual zerotree. IEEE Trans. Circuits and systems for video technology, 1997, 7(5): 719~727
- 12 Cheng P, Li J, Kuo C-CJ. Multiscale video compression using wavelet transform and motion compensation. In: Intl. Conf. on Image Process., 1995. 606~609
- 13 Antonini M, Barland M, Mathieu P, Daubechies I. Image coding using wavelet transform. IEEE transactions on image processing, 1992, 1(2): 205~220
- 14 Shen K, Delp E. Wavelet based rate scalable video compression. IEEE Trans. Circuits and systems for video technology, 1999, 9(1): 109~122
- 15 Desarte P, et al. Signal - Adapted Multiresolution Transform for Image Coding. IEEE Trans. On Information and Theory, 1992, 38(2): 897~904