

一种高效的运动补偿三维小波视频编码方案

常铮 卓力 沈兰荪

(北京工业大学信号与信息处理研究室 北京 100022)

摘要 该文提出了一种高效的运动补偿三维小波视频编码方案。该方案采用运动补偿时域分析技术,以有效去除视频序列中存在的时间冗余。然后,基于视频序列的运动特征,自适应确定帧组结构,在提高编码效率的同时降低内存需求和运算复杂度。最后,根据小波图像系数特性,采用基于四叉树分裂的小波图像编码方法对三维时空子带进行编码,以获得更高的压缩效率。实验结果表明,与其它运动补偿三维小波视频编码方法相比,该文提出的编码方案能够获得更好的性能。

关键词 视频编码, 运动补偿时域分析, 三维小波编码, 四叉树分裂小波编码, 自适应帧组结构

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2006)02-0237-05

An Efficient Motion-Compensated Three Dimension Wavelet Video Coding System

Chang Zheng Zhuo Li Shen Lan-sun

(Signal and Info Processing Lab., Beijing Polytechnic University, Beijing 100022, China)

Abstract This paper describes an efficient Motion-Compensated (MC) three dimension wavelet video coding system. In this system, MC temporal analysis is used to remove temporal redundancy. Then an adaptive group of pictures structure according to characteristics of motion can increase the coding efficiency and reduce computational complexity and memory requirement at the same time. Finally, a quadtree-based wavelet image coding method is proposed, which can adapt the property of wavelet image coefficient. The experimental results show that, the coding system is better than other motion-compensated three dimension wavelet video coding methods.

Key words Video coding, MC temporal analysis, Three-dimension wavelet coding, Quadtree-based wavelet coding, Adaptive group of pictures (GOP) structure

1 引言

随着小波理论不断发展,基于小波的图像压缩技术取得了巨大成功,极大地推动了小波视频编码技术的研究,小波视频编码方法已经成为视频编码领域的研究热点^[1-3]。

视频图像区别于静止图像的一个重要特征是不仅一帧图像内部存在空间相关性,相邻图像帧之间还存在时间相关性。为了有效去除视频的时间相关性,Podilchuk等提出了将帧间一维小波变换和帧内二维小波变换相结合的三维小波变换^[4],将序列中各个图像所对应的同一位置的各个点构成一个一维序列,采用一维小波变换去除时间冗余。Kim和

Pearlman等则提出了三维SPIHT算法对三维小波系数进行编码^[5]。对于活动性不强的图像序列,该方法具有良好的压缩效果,但对于纹理复杂且活动性强的图像序列,由于其帧间相关性较弱,沿时间轴方向作一维小波变换无法有效去除时间冗余,因此压缩效果并不理想。

随后,Choi等^[1]提出了一种具有代表性的运动补偿三维子带编码方法(Motion Compensated Three Dimension Subband Coding, MC_3DSBC),编码质量优于MPEG-2国际标准。但是这种方法采用固定大小的帧组(Group of Pictures, GOP)结构,无法对运动剧烈程度不同的各种序列都实现最优

2004-07-29 收到,2004-12-07 改回

国家自然科学基金(90304001, 60472036),北京市自然科学基金(4032008, 4052007)和北京市教委(KM200410005022, KZ200310005004)资助课题

编码,采用三维子带有限状态标量量化对三维时/空域子带进行编码时也没有充分考虑小波图像的特点。

本文提出了一种高效的运动补偿三维小波视频编码方案。该方案采用运动补偿时域分析技术,以有效去除视频序列中存在的时间冗余。然后,基于视频序列的运动特征,自适应确定 GOP 的大小,在提高编码效率的同时降低内存需求和运算量。最后,根据小波图像系数特性,采用基于四叉树分裂的小波图像编码方法对三维时/空子带进行编码,以获得更高的压缩效率。实验结果表明,与文献[1]中提出的运动补偿三维小波视频编码方法相比,本文提出的方案能够获得更好的编码性能。

2 运动补偿三维小波视频编码方案

2.1 运动补偿三维小波视频编码系统框图

本文提出的运动补偿三维小波视频编码系统框图如图 1 所示。可以看出,该系统包括两个关键技术:运动补偿时域分析技术和三维时/空子带编码技术。输入的视频信号首先经过运动补偿时域分析去除时间冗余,利用空域小波变换去除空间冗余,得到三维时/空子带。然后,三维子带经过基于四叉树分裂的小波图像编码方法,与运动矢量一同形成最后的输出码流。

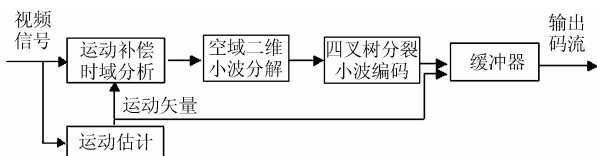


图1 运动补偿三维小波视频编码系统结构框图

与 MPEG 类似,本文采用 GOP 编码结构,一个 GOP 中可以包括 2, 4, 8 或者 16 个编码帧。合理选择 GOP 的大小不仅能够使运动补偿时域分析更有效地去除相邻帧之间的时间冗余,从而实现更高效的压缩,也会对编码系统的内存需求和运算复杂度产生重要影响。因此,如何根据视频序列的具体特性合理选择 GOP 大小也是该编码系统的关键技术之一。

下面将分别对编码系统所涉及的 3 项关键技术作详细论述。

2.2 运动补偿时域分析

视频序列的相邻帧之间存在很大的相关性,如何去除时间冗余信息就成为提高视频编码性能的关键。与传统的运动估计后求预测残差不同,运动补偿时域分析(Motion Compensated Temporal Analysis, MCTA)在进行运动估计后,

沿着运动轨迹的方向进行时间域的一维小波分解,运动轨迹可以由任何块匹配方法来确定。图 2 所示的是运动补偿时域分析示意图。

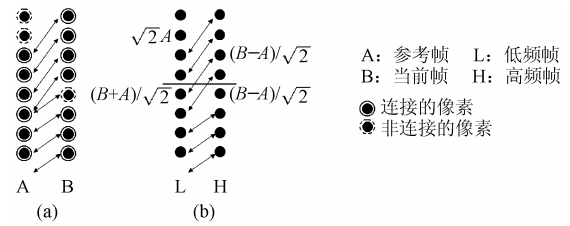


图2 运动补偿时域分析示意图

除了两帧图像没有运动或是纯粹平移运动的情况以外,在参考帧中找到的匹配块必然存在着与邻块相互重叠的现象,即参考帧中一个像素对应当前帧中两个或多个像素的现象,见图 2(a)。如果沿着所有的运动轨迹都进行滤波,就会导致参考帧中的一部分像素进行不止一次的滤波,而还有一部分始终没有进行滤波。因此,当存在重叠现象时,我们只能选择其中一条轨迹进行滤波,这就意味着参考帧和当前帧中都只有一个像素的子集可以进行沿运动轨迹的时域滤波。这部分像素被称为“连接的像素”,采用 Haar 小波沿运动轨迹进行小波分解:

$$L[m-d_m, n-d_n] = \frac{1}{\sqrt{2}}(B[m, n] + A[m-d_m, n-d_n]) \quad (1)$$

$$H[m, n] = \frac{1}{\sqrt{2}}(B[m, n] - A[m-d_m, n-d_n]) \quad (2)$$

而另一部分无法沿运动轨迹进行滤波的像素,被称为“非连接的像素”,按下面公式进行分解:

$$L[m, n] = \sqrt{2}A[m, n] \quad (3)$$

$$H[m, n] = \frac{1}{\sqrt{2}}(B[m, n] - A[m-d_m, n-d_n]) \quad (4)$$

其中 $L[m, n]$ 和 $H[m, n]$ 分别表示时域低频帧和时域高频帧, $A[m, n]$ 和 $B[m, n]$ 分别表示参考帧和当前帧, (d_m, d_n) 是当前帧中 (m, n) 点像素的运动矢量。经过运动补偿时域分析后,低频帧包含的是两幅图像的平均信息,而高频帧则可以看作是两幅图像运动估计补偿后的残差信号,包含了较少的能量。因此,运动补偿时域分析可以有效地去除时间冗余,提高编码效率。

2.3 基于视频序列运动特征的自适应 GOP 结构

运动补偿时域分析可以有效去除连续两帧之间的冗余信息,但也是编码系统中运算复杂度最高的部分。对于运动平缓的视频序列来说,在进行连续两次的运动补偿时域分析

之后，相邻的两个时域低频帧之间仍然存在着很强的相关性，需要再进行运动补偿时域分析以进一步去除冗余信息。而对于运动剧烈的视频序列来说，相邻时域低频帧之间存在的相关性则较弱，继续进行运动补偿时域分析，对去除时间冗余意义不大，而且需要更多的内存开销和运算量。因此，如何根据视频序列的运动特征，合理选择 GOP 的大小，将对提高编码性能，降低系统内存需求和运算复杂度有着重要的影响。

为此，本文提出了一种基于视频序列运动特征，自适应选择 GOP 大小的方法。可供选择的 GOP 大小共有 4 种，分别包含 2 帧，4 帧，8 帧和 16 帧等 4 种情况，记作 $GOP=2$ ， $GOP=4$ ， $GOP=8$ 和 $GOP=16$ ，如图 3 所示。

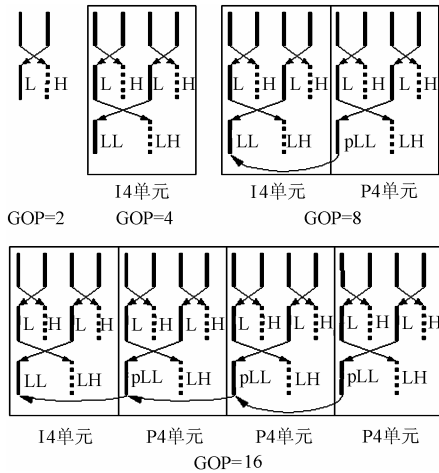


图 3 4 种 GOP 编码结构示意图

$GOP=2$ 时，仅进行一次运动补偿时域分析，得到时域低频帧 L 和时域高频帧 H。 $GOP=4$ 时，先进行两次运动补偿时域分析，然后对得到的两个时域低频帧 L 再进行一次运动补偿时域分析，得到第二级的时域低频帧 LL 和时域高频帧 LH，并称这样的连续 4 帧构成一个 I4 单元。 $GOP=8$ 时，先进行两组与 $GOP=4$ 相同的两级运动补偿时域分析，第一组连续 4 帧同样构成一个 I4 单元，然后以第一个 LL 为参考，对第二个 LL 进行运动估计补偿，得到一个预测残差帧 pLL，并与后续的 3 个时域高频帧构成一个 P4 单元。 $GOP=16$ 时，先进行 4 组与 $GOP=4$ 相同的两级运动补偿时域分析，第一组连续 4 帧构成 I4 单元，然后对后续的 3 个 LL 帧，均以前一个 LL 为参考，进行运动估计补偿，得到 3 个预测残差帧 pLL，并相应构成 3 个 P4 单元。

表 1 所示的是采用不同的 GOP 所需要的运动补偿时域分析的次数。从前面的分析和表 1 可以看出，对于运动平缓

表 1 不同 GOP 所需的运动补偿时域分析次数比较

GOP	每个 GOP	连续 16 帧
2 帧	1 次	8 次
4 帧	3 次	12 次
8 帧	7 次	14 次
16 帧	15 次	15 次

的序列来说，虽然采用大的 GOP 结构会增加运动补偿时域分析的次数，内存和运算复杂度需求量大，但由于能够充分去除时间冗余，因此会提高编码性能。而对于运动剧烈的序列来说，由于相邻时域低频帧之间的相关性较弱，则应该采取较小的 GOP，这样不仅不会影响编码性能，还可以减少运动补偿时域分析的次数，从而大大降低内存需求和运算量。

运动矢量是表征视频序列运动剧烈程度的一种有效方法，本文采用下式来确定视频序列的运动剧烈程度：

$$MA = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [mv_x^2(i, j) + mv_y^2(i, j)] \quad (5)$$

其中 $mv_x(i, j)$ 和 $mv_y(i, j)$ 分别为像素点 (i, j) 运动矢量的 x, y 分量， M, N 分别为一帧图像的宽度和高度。MA 的值小，表明视频序列的运动平缓，相邻时域低频帧之间的相关性强，在这种情况下应选择大的 GOP。MA 值较大时，说明视频序列运动剧烈，则应选择小的 GOP。

考虑到视频序列在连续的几帧或十几帧内，其运动特征变化不大，因此，当前 GOP 得到的 MA 值，可以作为后续 GOP 选择的依据。本文通过大量实验，选定 3 个阈值 $TH1=26$ ， $TH2=20$ 和 $TH3=14$ ，用以衡量视频序列的运动剧烈程度，作为选择 GOP 大小的依据，如表 2 所示。

表 2 阈值法选择 GOP

运动特征	GOP
$TH1 < MA$	2 帧
$TH2 < MA \leq TH1$	4 帧
$TH3 < MA \leq TH2$	8 帧
$MA \leq TH3$	16 帧

2.4 基于四叉树分裂的小波图像编码方法

小波视频编码技术的发展要依赖于小波静止图像编码技术的发展，小波静止图像编码技术是否充分利用了小波图像的特点，直接影响着视频编码的性能。本文考虑到时域高频帧为非自然图像的特点，根据小波系数在频率和空间上的能量聚类特性，提出一种基于四叉树分裂的小波图像编码方法(Wavelet Quadtree-based Coding, WQC)^[6]。

二叉树分裂是一种简单但非常有效的小波系数组织方式。每个集 S 包括一个方块区域中的所有像素, 可以用 $\{c_{i,j}\}$ 表示, (i,j) 表示该小波系数在图像中的位置。对于一个 n 值, 集 S 是否重要, 可以如下定义: 如果 $\max_{(i,j) \in S} \{|c_{i,j}|\} \geq 2^n$, 则集 S 是重要的; 否则, 集 S 是不重要的。这样, 集 S 是否重要可以表示为 n 值的函数:

$$F(n, S) = \begin{cases} 1, & 2^n \leq \max_{(i,j) \in S} \{|c_{i,j}|\} < 2^{n+1} \\ 0, & \text{其它} \end{cases} \quad (6)$$

对于不重要的集, 用一个比特 0 就可以表示该集中所有系数的状态; 对于重要的集, 则将这一方块区域分裂成 4 个尺寸相同的子块得到 4 个子集, 然后对这 4 个子集进行重要性检测, 对重要的子集仍然采用相同的方法继续分裂, 直到确定了当前阈值下所有的重要系数。一个二叉树分裂的过程如图 4 所示。

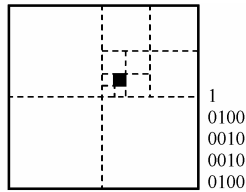


图4 二叉树分裂示意图

与 SPIHT 算法类似, WQC 采用两个列表: 不重要集列表(List of Insignificant Sets, LIS)和重要像素列表(List of Significant Pixels, LSP)。前者包括的是当前阈值下不重要的集, 后者为当前阈值下被检测为重要的小波系数, 两个列表在编码过程中动态更新。WQC 的编码过程可以用伪代码描述如下, 其中 n 表示比特平面, k 表示子带:

初始化过程:

```
each subband is initialized as  $S_k$ ;
Add each  $S_k$  to LIS $_k$  and set LSP $_k = \emptyset$ ;
Output  $b_n = \log_2 \{ \max_k [ \max_{(i,j) \in S_k} (|c_{i,j}|) ] \}$ .
```

编码过程:

```
For  $n = b_n, b_n-1, b_n-2, \dots$  逐比特平面进行编码
  For  $k = 1, 2, \dots, N$  子带进行编码
    Sorting Pass(分类过程)
      For each  $S \in \text{LIS}_k$ 
        —Process_Set( $S$ );
```

Refine Pass (细化过程)

```
For each  $c_{i,j} \in \text{LSP}_k$ , except those included in the
last sorting pass
  —output the  $n$ -th MSB of  $|c_{i,j}|$ ;
  —if  $c_{i,j} \geq 2^n$ , subtract  $2^n$  from  $c_{i,j}$ ;
```

其中

Process_Set(S)

{

Output $F(n, S)$;

If $F(n,S)=1$

—if S is a pixel, output sign of S and add S to LSP $_k$;

—else Code_Set(S);

}

Code_Set(S)

{

Partition S into four equal subsets $O(S)$;

For each $O(S)$

—output $F(n, O(S))$;

—if $F(n,O(S))=1$

* if $O(S)$ is a pixel, output sign of $O(S)$ and add $O(S)$ to LSP $_k$;

* else Code_Set($O(S)$);

—else

* add $O(S)$ to LIS $_k$;

}

在本文中, WQC 采用的熵编码是基于上下文的算术编码方法, 分类和细化两个过程输出的比特要经过算术编码来进一步提高压缩效率。

WQC 本质上是一种比特平面的编码方法, 码流具有嵌入的特点。在每一个比特平面下, 编码又是逐子带进行的。这样, 只要在编码过程中适当添加一些头信息, 就可以避免误码扩散, 提高码流的抗误码性能, 有利于码流在易发生误码的环境下传输。

3 实验结果

为了验证本文方法(简称 A_MC_3DWQC)的有效性, 我们将本文方法与文献[1]中提出的方法(简称 MC_3DSBC)进行了比较。表 3 给出了对各种测试序列分别采用两种编码方法得到的实验结果, 测试序列均为 96 帧 CIF 格式, 码率为 1.2Mbps, 帧率为 30 帧/秒。

表3 MC_3DSBC 和 A_MC_3DWQC 实验结果比较(dB)

测试序列	MC_3DSBC	A_MC_3DWQC	增加质量
Akiyo	40.84	46.21	+5.37
Mother_Daughter	40.24	42.84	+2.60
Football	31.65	33.65	+2.00
Foreman	36.48	37.17	+0.69
Mobile	30.54	30.51	-0.03
Bus	29.07	29.04	-0.03

图5和图6所示的是分别采用两种方法 Mother_Daughter 和 Foreman 两个测试序列各帧的重建质量对比结果,图7所示的是分别采用两种方法得到的 Akiyo 序列第11帧的重建图像。

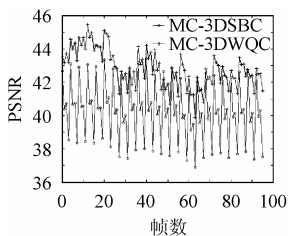


图5 Mother_Daughter 测试序列两种编码方法实验结果比较

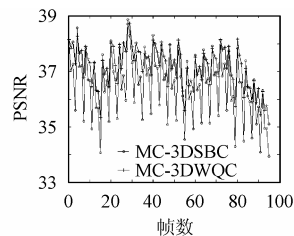


图6 Foreman 测试序列两种编码方法实验结果比较



(a) MC-3DSBC方法



(b) A_MC_3DWQC方法

图7 分别采用两种编码方法得到的 Akiyo 测试序列第11帧重建图像

从表3,图5,图6以及图7可以看出,在相同编码条件下,对于运动平缓或者运动剧烈的测试序列,本文提出的运动补偿三维小波视频编码方案均可以获得更好的性能,最大时可以提高5dB左右。而对于运动剧烈程度中等的测试序列来说,编码性能基本相当。

表4所示的是 Football 序列前96帧两种方法所采用的 GOP 大小的比较结果。由于 Football 测试序列的运动剧烈程度较高,本文方法采用了 GOP=2 的结构42次, GOP=4 和 GOP=8 的结构各1次。而 MC_3DSBC 则采用了固定的 GOP=16 的结构,共有6次。因此,本文算法需要进行运动补偿时域分析共 $1 \times 42 + 3 \times 1 + 7 \times 1 = 52$ 次,而 MC_3DSBC 需要进行运动补偿时域分析 $15 \times 6 = 90$ 次。这说明,对于运动剧烈的视频序列来说,本文方法不仅可以提高编码性能,还能够大大降低内存需求和运算复杂度。

表4 96帧 football 测试序列采用的 GOP 个数比较

GOP	MC_3DSBC	A_MC_3DWQC
2	0个	42个
4	0个	1个
8	0个	1个
16	6个	0个

4 结束语

本文提出了一种高效的运动补偿三维小波视频编码方案。实验结果表明,与文献[1]中具有代表性的运动补偿三维小波视频编码方法相比,在相同的测试条件下,对于运动平缓的序列,本文提出的运动补偿三维小波视频编码方案可以获得更好的编码性能,而对于运动剧烈的测试序列,在提高编码性能的同时还可以大大降低内存需求和运算复杂度。

参考文献

- [1] Choi Seung-Jong, Woods J W. Motion-compensated 3-D subband coding of video. *IEEE Trans. on Image Processing*, 1999, 8(2): 155-167.
- [2] 沈兰荪,卓力,田栋,汪孔桥. 视频编码与低速率传输. 北京: 电子工业出版社, 2001: 69-82.
- [3] 张宗平,刘贵忠. 基于小波的视频图像压缩研究进展. 电子学报, 2002, 30(6): 883-885.
- [4] Podilchuk C I, Jayant N S, Farvardin N. Three dimensional subband coding of video. *IEEE Trans. on Image Processing*, 1995, 4(2): 125-139.
- [5] Kim Beong-Jo, Pearlman W A. An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT). *Proceedings DCC'97 (Data Compression Conference)*, Snowbird, Utah, 1997: 251-260.
- [6] 卓力. 无线环境下图像/视频流编码与传输技术研究. [博士论文], 北京: 北京工业大学, 2004.

常 铮: 男, 1980年生, 硕士生, 研究方向为视频编码与传输。

卓 力: 女, 1971年生, 博士, 副教授, 近年来发表论文多篇, 研究方向为视频编码、无线IP视频传输等。

沈兰荪: 男, 1938年生, 教授, 博士生导师, 研究领域为图像编码、处理、传输与应用, 以及光谱信号的检测与处理等。