

文章编号:1001-9014(2004)04-0265-06

视频帧组与其残差帧组交替的 3D-DWT-SPIHT 压缩编码方法研究

胡 佳, 丁文奇, 张立明, 胡 波
(复旦大学 电子工程系, 上海 200433)

摘要:提出一种改进的方法,它采用视频帧组与残差帧组交替来进行三维小波变换,减少视频帧组内的帧数、时间延迟和存储空间,同时在时间轴上采用两种小波基相结合的方法,提高了压缩效率和解码图像的质量;尤其是对背景复杂的视频序列有较好的效果.实验结果表明本文方法在相同条件下比现有方法的 PSNR 平均提高约 1dB,在内存减小近一半时仍有所提高.

关键词:三维小波变换(3D-DWT);三维等级树集合划分算法(3D-SPIHT);三维小波的等级树集合划分算法(3D-DWT-SPIHT);视频帧组(GOF);残差帧组

中图分类号:TN919.8 文献标识码:A

RESEARCH ON A NOVEL 3D-DWT-SPIHT CODING BY ALTERNATE GROUPS OF ORIGINAL AND RESIDUAL ERROR FRAME

HU Jia, DING Wen-Qi, ZHANG Li-Ming, HU Bo
(Department of Electronic Engineering, Fudan University, Shanghai 200433, China)

Abstract: An improved method that uses 3D-DWT-SPIHT in original video group and residual error group alternatively was proposed. In addition, the multi-wavelet bases were combined in time domain. The proposed method reduces coding latency time and releases memory requirement, and improves both compression efficiency and reconstructive image's quality, especially to the video sequences with complicated background. Simulation results show that, in the case of equal compression ratio and memory space, our method has higher quality (PSNR above 1dB) than available methods. It is still slightly better than others', when memory is reduced to half.

Key words: three dimension wavelet transform (3D-DWT); three dimension set partitioning in Hierarchical tree (3D-SPIHT); the set partitioning in Hierarchical tree for three dimension wavelet transform (3D-DWT-SPIHT); GOF (Group of Frame); group of residual error frame

引言

现有的视频编码标准,如 ITU 的 H. 261, H. 263 及 ISO/IEC 的 MPEG-1, 2, MPEG-4, 大都基于离散余弦变换(DCT). 基于小波变换的视频图像压缩近几年发展迅速. 小波系数不仅具有 DCT 变换中高频部分零较多的特点,同时相同方向子带的相似性进一步提高压缩效率,促使基于小波变换的后编码研究有了很大突破. 1996 年, A. Said 在 1993 年 Shapiro

提出嵌入式零树编码(EZW)^[1]的基础上,提出 SPIHT 的编码算法^[2],它不需多次扫描,采用位平面编码,考虑树内\树间的相似性,同时把有效系数位置的传输隐含在算法的执行过程之中,并允许峰值信噪比下降 0.3~0.6dB 时,不采用算术编码,使执行速度更快. 1997 年后发展了基于小波变换的视频编码,包括:基于运动估计和运动补偿的小波视频编码^[3],三维小波变换编码方法^[4,5,9]等. 三维小波变换编码对运动速度比较慢的视频信息,通过过去时间

收稿日期:2003-09-03,修回日期:2003-12-15

基金项目:“863”高技术计划基金资助项目(2002AA742051)

作者简介:胡佳(1980-),女,云南昆明人,复旦大学硕士研究生,主要从事图像处理研究.

Received date: 2003-09-03, revised date: 2003-12-15

冗余,提高了压缩率.典型的三维小波变换编码算法有:3D-EZW[4]算法、3D-SPIHT[5]算法和 3D-SL-CCA[6]算法.三者都是相应二维算法在三维上的扩展,因此继承了二维算法的特点.

文献[5]首次提出将三维小波变换结合三维等级树集合分割算法(本文中称为:3D-DWT-SPIHT)用于视频流,并与 MPEG2 比较.文献[7]尝试用 3D-DWT-SPHIT 进行视频编码并与其他的小波变换压缩方法做比较,文献[8]是通过构造新的时空方向性树降低时间延迟和减少存储容量,文献[9]采用改变存放小波系数位置的列表,更好地组织小波变换后的系数,视频压缩性能只比 H.263 高 0.4dB.本文通过对 3D-DWT-SPHIT 视频流压缩编码方法的分析,提出时域多小波基、视频帧组与残差帧组交替的方法对 3D-DWT-SPIHT 进行改进.对比文献[7]与文献[8],在同样的条件下,不仅重构图像有更高质量,而且降低了处理的时间延迟,减少了编解码器的存储容量.

1 基于 3D-DWT 和 3D-SPIHT 结合的视频编码

图 1 是一个 3D-DWT 和 3D-SPIHT 结合的编解码系统,系统的编码主要包括三部分:3D-DWT、SPIHT 编码和自适应算术编码(备选).图像序列按 T 帧为一组(GOF)输入 3D-DWT 的模块,进行三维小波变换.根据小波变换后系数在各子带内分布的特点,用 SPIHT 算法后编码.文献[2]已证明在 SPIHT 后用算术编码 PSNR 仅能提高 0.3~0.6dB,所以在很多系统中为了简单起见,可不采用算术编码,故在图中用虚线框表示.编解码的模块相同,只是内容上互为逆运算.如图 1 所示.

下面简单介绍三维小波变换(3D-DWT)、SPIHT

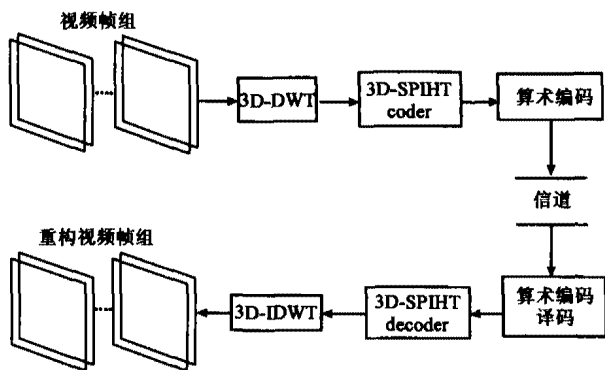


图 1 3D-DWT-SPIHT
Fig.1 3D-DWT-SPIHT system block diagram

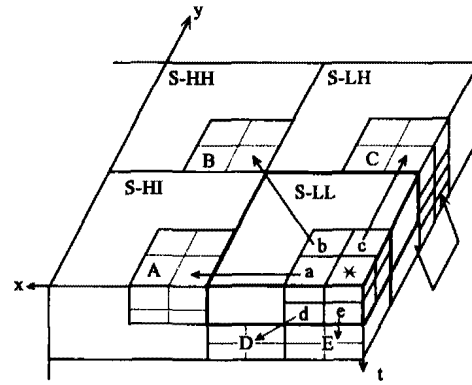


图 2 三维小波分解及父子节点关系示意图
Fig.2 3D-DWT and node relationship sketch map

编码模块:

1.1 三维小波变换(3D-DWT)

2D-DWT 是二维变换在时间轴上的扩展,目前可通过 2 种方式来实现.第一种为时间分离的三维小波变换,即由帧内二维和帧间一维变换组合(本文采用的方法);第二种是真三维小波变换,即直接构造三维小波基来实现.若我们采用 T 帧为一组(GOF),编码器首先沿时间轴进行 P 级一维小波变换,再对所得各时间子带图像帧进行空间二维小波分解.因小波变换的子带分割以 2 的幂次进行,级数 P 与 T 满足公式: $T \geq 2^P$.

1.2 三维等级树集合划分算法(3D-SPIHT)

3D-SPIHT 编码是 SPIHT 算法的扩展.它定义三维方向树结构以及相应节点间的父子关系,有效地组织了 3D-DWT 的系数.图 2 为一个三维小波分解图,为便于分析最高级只留下 2 个时间帧和空间的 4 个频带(S-LL,S-LH,S-HH,S-HL)在树根上的 8 个像素(*,a,b,c,d,e,f,g,)如图 2 的右下角示,除打星号的像素外,其它 7 个像素在各自同方向的子带上 8 个直接后代(图 2 中用箭头标出),分别在(A,B,C,D,E,F,G)的位置上,对下一级,每个直接后代又有 8 个孙辈…….小波分解的级数越高,树根内节点数越少,每棵树的子孙越多,小波系数能量越集中,算法的编码效率越高.

令 (i,j,k) 表示小波系数 $C(i,j,k)$ 的位置坐标, $O(i,j,k)$ 表示 $C(i,j,k)$ 的直接后代, $D(i,j,k)$ 表示它的所有后代, $L(i,j,k)$ 为除了直接后代外的子孙,有关系: $L(i,j,k) = D(i,j,k) - O(i,j,k)$.假设对 T 帧作 P 级的三级小波变换,得到最高级的低频成分为 S-LL,如图 2 所示.H 表示所有的树根节点集合,有直接后代的小波系数 $C(i,j,k)$ 的 8 个直接后代表

示为

$$O(i, j, k) = \begin{cases} (2i, 2j, 2k), (2i+1, 2j, 2k), (2i, 2j+1, 2k), (2i+1, 2j+1, 2k), \\ (2i, 2j, 2k+1), (2i+1, 2j, 2k+1), (2i, 2j+1, 2k+1), (2i+1, 2j+1, 2k+1). \end{cases}$$

$$\text{用函数: } S_n(q) = \begin{cases} 1 & \max_{(i,j,k) \in q} \|c(i,j,k)\| \geq 2^{n_0} \\ 0 & \text{其它} \end{cases} \quad (1)$$

表示小波变换系数 $C(i, j, k)$ 的重要性, q 表示一个集合, n_0 表示当前所取的系数幅值的阈值. 在 3D-SPIHT 编码过程中, 经 3D-DWT 变换后系数的信息保存在 3 个链表中: 不重要像素链表 LIP, 存放不重要小波系数像素点的坐标; 重要像素链表 LSP, 存放重要像素的坐标; 不重要集合链表 LIS, 存放不重要系数集合在最高小波分解级的像素坐标.

3D-SPIHT 编码的实现思路是: 对树根中的小波系数 $C(i, j, k)$ 用式(1)判别其重要性, 并把 S_n 输出, 在接收端得到树根各位置上的信息, 对重要的系数输出其符号, 并将其位置坐标移入 LSP 中, 有后代的树根节点(初始化 LIS 集时设为 A 类)对它的 $D(i, j, k)$ 进行式(1)的判别, 并把其 S_n 输出, 在接收端得到了其位置信息, 若 $D(i, j, k)$ 为重要集合, 则将此集合进行分割, 分为 $O(i, j, k)$ 和 $L(i, j, k)$, 判断 $O(i, j, k)$ 中各值对应的小波系数的重要性, 并把 S_n 输出, 在接收端得到了其位置, 如是重要系数则输出符号信息并将其位置坐标移入 LSP, 否则送入 LIS 的表尾. 若 $L(i, j, k)$ 为空集, 则从 LIS 中去掉 (i, j, k) , 否则把 (i, j, k) 送入 LIS 的表尾, 设为 B 类. 在对所有的系数重要性判别后, 把 LSP 中的在 n_0 位平面上所有重要系数的码字(除了当前处理的外), 根据所要求的编码精度形成码流输出, 最后再用 $n_0 - 1$ 代替 n_0 , 继续以上的操作.(具体算法参考文献[5]) 根据嵌入式编码的特点, 编解码对称.

从上可知, 在任何阈值 n_0 下, 一旦小波系数位置坐标被移入 LSP, 在以后的各个位平面上就可自动输出它在该位平面上的值, 不需重复扫描. 3D-SPIHT 利用位平面分层划分的方法, 间接实现了空间小波树的比特平面排序, 有效减少编码符号集的规模, 提高压缩比.

2 改进的 3D-DWT-SPIHT

虽然现有的 3D-DWT-SPIHT 算法的编码效率较高, 但因存储空间和实时性的限制, 一次三维编码图像组的帧数不能取的太多. 此外, 时间上图像组的尺度过大, 对有些变化较大、背景较复杂的视频序列, 时间轴上为零的像素值数目增加有限, 有一定能提高图像压缩的效率. 同时小波基的选择是否有潜力可挖. 针对上述问题, 在文献[8]的 3D-SPIHT 算法上, 我们提出采用多小波基方法和视频与残差帧组交替的方法对 3D-DWT-SPIHT 进行改进.

2.1 时间轴的多小波基方法

现有小波压缩方法多采用空间域和频率域光滑性好的 Daubechies 9/7 双正交小波基, 具体操作中为了减少边界效应, 在空间和时间上图像必须作对称延拓处理, 而 9/7 双正交小波基的滤波器阶数高, 对于时间域, 则需要对 T 帧的头尾进行相当 4 帧的延拓, 当分解级高时会发生困难, 如 $T=16$ 时, 做 4 级时间轴分解, 第 3 级后只剩 4 帧, 就无法进行延拓, 因此在文献[7~9]中在空间域用 Daubechies 9/7 小波, 时间域都采用 Harr 小波变换. 但因 Harr 小波基在时间域不光滑, 频率域无限, 使沿时间轴的小波分解不能达到最佳.

分析在时间域上的小波变换, 可发现高频主要表现在帧间运动的地方, 而大量的树内自相似性主要是在不运动的背景上, 因此采用 2 种不同的小波基并不影响其相似性, 同时可改善 Harr 小波基的不

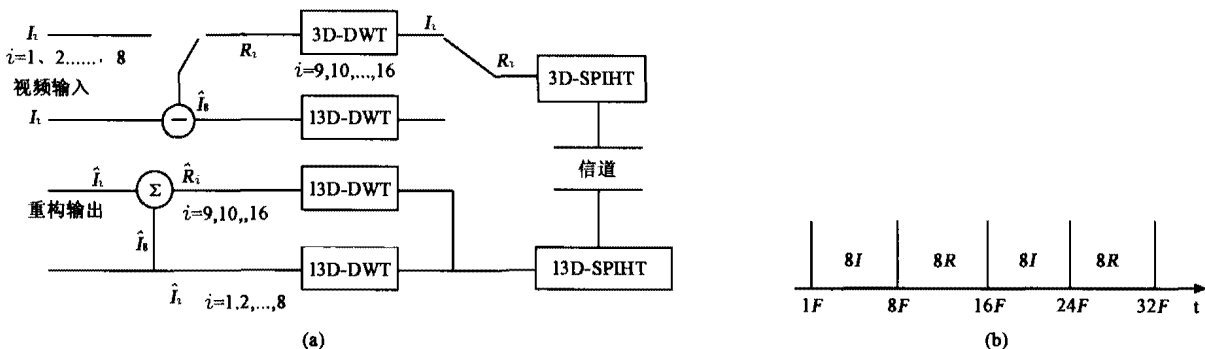


图3 (a) 视频和残差帧组法 3D-DWT-SPIHT 流程图 (b) 视频和残差帧组法中视频流的划分

Fig.3 (a) The flow chart of 3D-DWT-SPIHT with video and residual frame method (b) subsection of alternate frame

足. 同样取 $T = 16$, 用 2 次 Daubechies 9/7 小波基, 然后用 Harr 小波基; $T = 8$ 时可用一次 Daubechies 的 9/7 小波基. 实验表明这种方法在一定压缩比条件下能提高重建图像的 PSNR 值.

2.2 视频与残差帧组交替的方法

从 SPIHT 算法本身的特点来看, 小波分解的级数越高, 每个树根可组织的节点数越多, 树也就越大, 编码效率越高, 但因内存和延迟等原因, GOF 中帧数不能太多, 一般取 $T = 16$. 在文献[8]中, 13D-SPIHT 取 $T = 8$ 编码效果有所提高. 本文取 $T = 8$, 但用视频和残差帧交替的方法, 相当于将视频流中的 16 帧图像分为 2 个 8 帧进行处理, 没有降低处理效率. 设残差帧是当前帧和参考帧的差值帧, 主要反映两帧运动边界的信息. 系统如图 3 所示.

具体描述如下: 开始时在发送端首先作 8 帧原始图像(表示为 I_i)的 3D-DWT 接着进行 3D-SPIHT 压缩编码, 同时将前 8 帧恢复出来, 保留恢复的第 8 帧 \hat{I}_8 , 后面的每一帧与恢复的第 8 帧做差值, 可分别得到第 9 帧和第 8 帧的残差帧(表示为 R_i), 第 10 帧和第 8 帧的残差帧, 以此类推直到得到后 8 帧的残差帧. 可用公式描述为

$$R_i = I_i - \hat{I}_8. \quad (i = 9, \dots, 16) \quad (2)$$

然后, 对残差帧进行 3D-DWT-SPIHT 处理. 下一次再作 $8I + 8R$, 直到视频流结束为止(如图 3(b)所示). 实验结果表明, 在不影响压缩效果, 甚至还对图像的 PSNR 有所提高的条件下, 该方法有效的减小了时延和所需要的存储单元. 考虑视频流和残差帧的特点, 根据变化较慢的视频流, 其帧与帧之间具有一定地相似性, 用残差法可以更好地去除时间冗余, 提高压缩比; 而在一定压缩比下, 提高重建图像的 PSNR 值. 同时因 SPIHT 的码流是以位平面输出的, 很容易控制, 本文中视频帧组和残差帧组的码流比例分配有所不同, 在实验中我们分配 60% 给视频组 40%

给残差帧组, 也可自适应的调整码流的分配, 以取得更好的 PSNR.

3 实验结果

下面通过对标准序列大小为 $352 * 288$ 像素(CIF 格式) 每个像素值为 8 比特 MissA, Claire, Salesman 的灰度信号进行编解码的测试, 来证实我们提出的改进方法. 所有的测试均用视频图像的平均压缩比(CR, 单位: 倍)和解码图像与原图像的峰值信噪比(PSNR, 单位: dB) 2 个指标来比较的.

1) 单独采用时间轴 Daubichies 9/7 和 Haar 小波结合的时间轴多小波基方法

相同压缩比, $T = 16$, 对 1 ~ 48 帧的 MissA 视频图像作 3D-DWT-SPIHT 编解码测试, 比较结果见表 1. 可以看到单独采用时间轴 Daubichies 9/7 和 Haar 小波结合的方法比文献[5]中的方法提高 2dB ~ 4dB, 比改进的文献[7]和文献[8]的结果还提高了 0.2 ~ 0.3dB.

表 2 中第 4 列是单独采用时间轴多小波基方法与文献[7]的比较, 结果时域多小波基的引入, 对 PSNR 都有改善, 对 MissA 尤为明显, 提高了 0.03 ~ 0.22dB.

2) 采用视频与残差帧组交替的方法对视频图像的测试($T = 16$)

为了与文献[7]进行比较, 在相同的内存容量

表 1 单独采用时间轴 Daubichies 9/7 和 Haar 小波结合的方法与文献比较(PSNR)

Table 1 Multi-wavelet bases method comparison with documents

	CR = 100	CR = 200	CR = 400
文献[5]的方法	39.46	36.53	32.18
文献[8]的方法	40.03	38.46	36.32
时间轴多小波基的方法	40.30	38.78	36.55

表 2 视频与残差帧组交替方法与文献[7]比较(PSNR)

Table 2 Alternate video and residual frame method comparison with [7]

视频序列	CR	文献[7]的结果	时间轴多小波基的结果	视频与残差帧组交替
MissA	102.4884	39.9724	40.1963	39.9823
	219.4471	38.2224	38.4250	38.2830
	501.7708	35.7607	35.7912	35.7556
Claire	59.9943	41.8048	41.9986	42.7040
	113.2031	38.1878	38.2473	39.4567
	226.0669	34.9302	34.7271	36.1067
Salesman	24.3697	36.4545	36.5591	37.7339
	52.7924	32.9131	32.9090	33.9789
	136.6277	29.6187	29.5543	30.3150

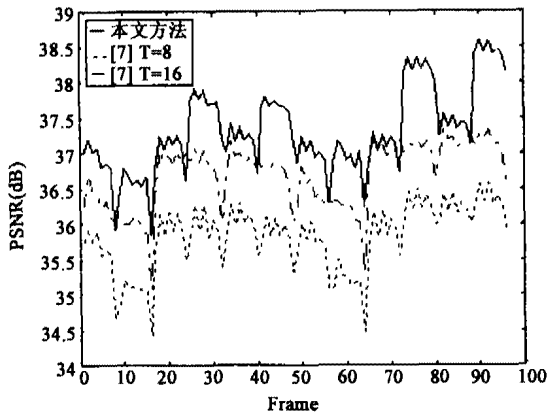


图4 Salesman(1~96 frame)每帧 PSNR

Fig.4 PSNR of salesman(1~96 frame)

表3 3D-DWT-SPIHT 在不同帧数时的结果(摘自文献[8])

Table 3 3D-DWT-SPIHT results of different frames

T (frame)	CR = 100	CR = 200	CR = 400
8	39.19	36.22	31.57
16	40.03	38.46	36.32

下,我们首先把提出的残差帧组的方法结合时间轴的多小波基的方法用 16 帧进行交替,对视频流(1~64Frame)作了测试,结果如表 2 所列。

本文方法与文献[7]结果比较:在不同类的图像和不同的压缩比下,除了 MissA 在 501 倍的情况 PSNR 低了 0.0051dB 外,其他情况都要好。

表 3 说明时间轴的帧数 T 越大,压缩效率就越高。在相同 CR 下,以 16 帧为一个处理单元比以 8 帧为一个处理单元所得的 PSNR 高约 1~4dB。这是由于对 16 帧可作 4 级小波变换,而 8 帧只能作 3 级小波变换,如前所述小波变换的级数越多,图像的能量越集中,有利于提高 3D-DWT-SPIHT 算法的编码效率,下面用我们提出的方法与直接用 $T=8$ 的 3D-SPIHT 及文献[7]中的 $T=16$ 进行比较。

3)采用视频与残差帧组交替方法对视频图像的测试($T=8$)

为了说明本文方法的优越性,我们采用背景比较复杂,一般方法压缩效果不太好的 Salesman 视频

图像序列作为测试。用 8 帧视频,8 帧残差交替的方法。码率分配为 62% 给视频帧组,38% 给残差帧组,对 Salesman 图像序列的 1~48 帧作了 3D-DWT-SPIHT 编解码,从表 4 中的测试结果可以看出:

1)交替帧组法所得结果的 PSNR 明显优于直接以 8 帧进行 3D-DWT-SPIHT 处理的结果,高 1dB 以上,二者的时延时间;

2)交替帧组法无论在存储容量,还是实时性方面都优于直接用 16 帧进行三维压缩,而且在相同压缩比下其重建图像的 PSNR 都有所提高。

综上所述:交替帧组法节省了内存,减小了延迟,所得结果优于 8 帧、16 帧的 3D-DWT-SPIHT 视频压缩编码。

图 4 给出了在 $T=8$, CR = 24.3697 时,文献[8]的 13D-SPIHT 方法在 8 帧和 16 帧为一组的情况下和视频与残差帧组交替方法对 Salesman 1~96 帧的编解码的结果。本文方法比原来 $T=8$ 的方法所得的信噪比高得多。在某些帧(如第 8、17、64 帧) PSNR 都下降。这是由于在该处两帧之间运动较为剧烈,但当一般的 3D-DWT-SPIHT 方法性能有较大下降时,本文方法性能下降相对较小。

图 5 给出 CR = 100 倍时, Salesman 序列用文献[7]的方法和本文的方法重建的第 1 和第 24 帧的图像。选这两帧是因为考虑到我们的方法对是将视频流分成视频帧组和残差帧组处理的,所以分别用视频帧重建图像(第 1 帧)和用残差帧重建的图像(第 9 帧)与所对应的文献[8]的重建帧进行对比。

从图 5 可以看出,用本文方法重建视频帧和残差帧的 PSNR 都比文献[8]方法重建的图像高,其中残差帧重建图像比视频帧提高更多,主观效果也有明显改善。

5 结论

本文采用视频帧组与残差帧组交替来进行三维小波变换,并在时间域上采用 2 种小波基结合的方法,提高了视频码流的压缩效率,在保证一定压缩比的条件下,有效的提高了重建视频图像的 PSNR 值。

表 4 用改进的方法的峰值信噪比与压缩比

Table 4 PSNR and CR of improved 3D-DWT-SPIHT method

方法	PSNR			内存量	时延
	CR = 24.3697	CR = 52.7924	CR = 136.6277		
交替帧组法	37.2896	33.2976	29.4588	9 帧	8 帧
直接用 8 帧	35.6896	31.6478	28.0256	8 帧	8 帧
文献[7](直接用 16 帧)	36.4545	32.9131	29.0659	16 帧	16 帧

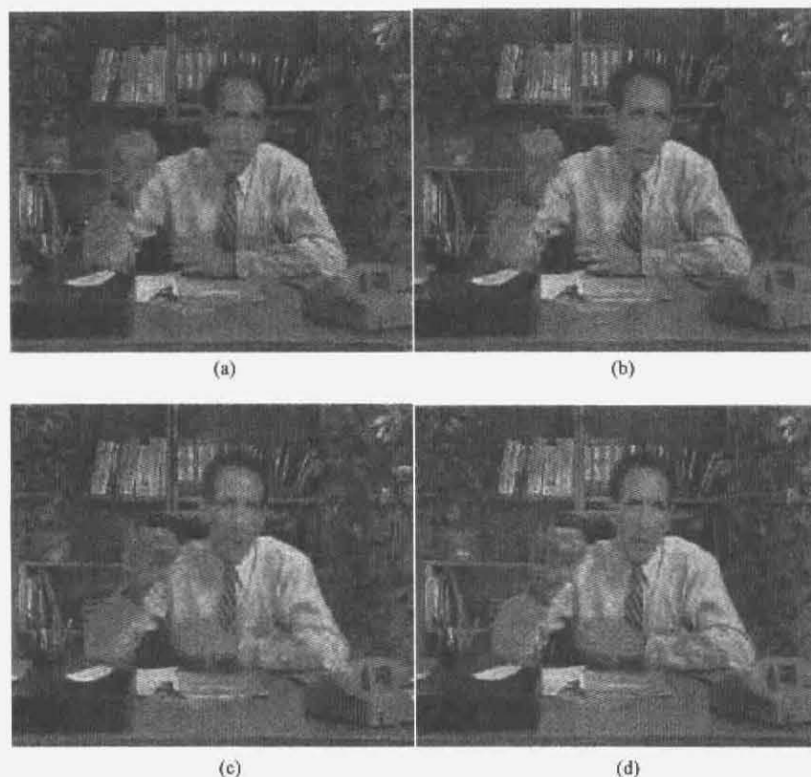


图5 用文献[8]的方法和本文方法重建的结果比较(a)[8]中方法第1帧重建图像 PSNR = 29.80dB (b)本文方法第1帧重建图像 PSNR = 30.47dB (c)[8]中方法第9帧重建图像 PSNR = 28.80dB (d)本文方法第9帧的重建图像 PSNR = 29.88dB

Fig. 5 Reconstruction results comparison (a) No. 1 reconstructive frame of [8] PSNR = 29.80dB (b) No. 1 reconstructive frame of our method PSNR = 30.47dB (c) No. 9 reconstructive frame of [8] PSNR = 28.80dB (d) No. 9 reconstructive frame of our method PSNR = 29.88dB

同时用本文提出的方法,可在时间轴的帧数减少一半的情况下,仍有相当于多帧时的压缩效率,尤其对于背景比较复杂的视频流,可优于多帧的结果,从而节省内存,减小帧间延迟。此外,因3D-3DWT-SPIHT编码器的算法具有码流嵌入式特点,能实现码率的精确控制,在确保视觉效果的同时,可任意截断码流。文中对视频帧组和残差帧组采用不同的固定码率分配,以提高编码效率,编码结果没进行熵编码,根据文献[2],如后编码采用游程编码和熵编码还可提高0.3~0.6dB。如采用自适应的方法对不同的情况改变视频帧组和残差帧组的码率分配,将可进一步提高压缩的效率,有关码率分配我们将在下一步工作中进行研究。

REFERENCES

- [1] Shapiro M. Embedded image coding using zerotrees of wavelet coefficients[J]. *IEEE Trans. Signal Processing*, 1993, **41**(12): 3445—3462.
- [2] Amir Said. A new fast and efficient image codec based on set partitioning in hierarchical trees[J]. *IEEE Transaction on Circuits and Systems for Video Technology*, 1996, **6**(6): 243—250.
- [3] Vass Jozsef, Chai Bing-Bing. Significance-linked connected component analysis for very low bit-rate wavelet video coding[J]. *IEEE Trans. Circuits Syst, CAS For Video Tech.*, 1999, **9**(4): 630—647.
- [4] Chang Y, Pearlman W A. Three-dimensional subband coding using the Zero-tree method[J]. *Proc. SPIE's 1996 Symp. on Visual Communications and Image Processing' 96 Mar, 1977*, **2727**: 1302—1312.
- [5] Kin BJ, Pearlman W A. An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT)[C]. *Proc Data Compression Conf, Utah: snowbird, 1997*: 251—260.
- [6] Vass J, Chai B B, Zhuang Xinhua. 3 D-SLCCA—a highly scalable very low bit rate software-only wavelet video codec[J]. *IEEE Workshop Multimedia Signal Processing*, 1998: 474—479.
- [7] YANG Chun-Ling, YU Ying-Lin. Research on embedded video compression based on 3D-wavelet transformation[J]. *Acta Electronica Sinica*(杨春玲,余英林.基于三维小波变换嵌入式视频压缩算法的研究. *电子学报*), 2001, **29**(10): 1381—1383
- [8] YANG Chun-Ling, YU Ying-Lin. An improved wavelet video coder using three-dimensional set partitioning in hierarchical trees[J]. *Acta Electronica Sinica*(杨春玲,余英林.改进的小波域三维等级树集合分割视频编码方法. *电子学报*), 2001, **22**(5): 86—92
- [9] ZHANG Zong-Ping, LIU Gui-Zhong, HOU Xing-Song. Improved 3D-wavelet video coder[J]. *Journal of Xi'an Jiaotong University*(张宗平,刘贵忠,侯兴松.一种改进的三维小波视频编码. *西安交通大学学报*), 2001, **35**(6): 595—599.