

基于 3D 混合树和视觉特性的视频可分级编码算法

付明哲¹, 王相海^{1,2}, 宋传鸣¹

(1. 辽宁师范大学 计算机与信息技术学院, 辽宁 大连 116029; 2. 智能计算与信息处理教育部重点实验室, 湖南 湘潭 411105)

摘要: 分析了视频数据的 3D 小波系数分布特性, 提出了一种基于混合 3D 树型结构和 HVS 特性的视频可分级编码算法。首先, 依据小波低、高频系数的自相关性, 确定相应的树型结构来扫描和处理时间维上的低、高频系数, 明显减少了用于定位重要系数的同步信息; 其次, 依据人类视觉系统对各频率子带敏感程度的不同, 对各子带系数进行加权, 使得重构视频的重要系数得以排在码流前端, 从而在很大程度上提高了中低码率下视频的重构质量。对多种标准测试视频的仿真实验验证了本文算法的有效性, 与非对称树型结构编码方案和单一时空方向树结构方案相比, 该算法解码图像的 Y、U 和 V 3 个分量的均峰值信噪比分别高出 0.65dB、1.75dB、1.77dB 和 0.23dB、2.11dB、1.72dB。此外, 算法有效抑制了振铃效应, 并获得了更好的主观效果。

关键词: 视频编码; 混合树结构; 3D 小波变换; 零树; 人类视觉系统

中图分类号: TP391

文献标识码: A

文章编号: 1000-436X(2012)11-0100-08

Scalable video coding algorithm based on 3D hybrid tree and visual characteristics

FU Ming-zhe¹, WANG Xiang-hai^{1,2}, SONG Chuan-ming¹

(1. College of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China;

2. State Key Laboratory for Smart Computing and Information Processing, Xiangtan 411105, China)

Abstract: The distribution characteristic of three-dimensional wavelet coefficients of video data was analyzed, and a scalable video coding algorithm was subsequently addressed based on hybrid three-dimensional tree and human visual system (HVS) characteristics. First, the hybrid tree structure was adaptively determined according to the auto-correlation of low-pass and high-pass coefficients. It reduced obviously the number of synchronization bits locating significant wavelet coefficients when scanning and processing low-pass and high-pass coefficients in temporal dimension. Second, each wavelet coefficient was weighted in terms of HVS sensitivity to its corresponding subband. Significant coefficients thus tended to be coded with high priority and arranged at the front of bitstream, and the reconstructed video quality was improved at low and medium bitrates to a great extent. Experimental results in terms of peak signal-to-noise ratio (PSNR) verified the effectiveness of the proposed algorithm on several test videos with varying characteristics. 0.65dB, 1.75dB, and 1.77dB higher PSNR are gained than asymmetric 3-D orientation tree for Y, U, and V components, respectively. Moreover, 0.23dB, 2.11dB, and 1.72dB higher PSNR are reached than single temporal-spatial orientation tree separately for Y, U, and V components. Besides, better subjective quality is obtained through effectively attenuating ringing artifact.

Key words: video coding; hybrid tree structure; 3D wavelet transform; zerotree; human visual system

收稿日期: 2011-03-24; 修回日期: 2011-08-26

基金项目: 国家自然科学基金资助项目(41271422); 辽宁省自然科学基金资助项目(20102123); 辽宁省博士科研启动基金资助项目(20121076); 计算机软件新技术国家重点实验室开放基金资助项目(KFKT2011B09, KFKT2011B11); 南京邮电大学图像处理与图像通信江苏省重点实验室开放基金资助项目(LBEK2010003, LBEK2011001); 苏州大学江苏省计算机信息处理技术重点实验室开放课题基金资助项目(KJS1124); 智能计算与信息处理教育部重点实验室(湘潭大学)开放课题基金资助项目(2011ICIP06)

Foundation Items: The National Natural Science Foundation of China(41271422); The Natural Science Foundation of Liaoning Province(20102123); The Scientific Research Foundation for PhD of Liaoning Province(20121076); The Open Foundation of State Key Laboratory for Novel Software Technology of Nanjing University(KFKT2011B09, KFKT2010B11); The Jiangsu Key Laboratory's Open Foundation of Image Processing and Image Communication of Nanjing University of Posts & Telecommunications(LBEK2010003, LBEK2011001); The Jiangsu Key Laboratory's Open Foundation of Computer Information Processing Technology of Soochow University(KJS1124); The Open Project Program of Key Laboratory of Intelligent Computing & Information Processing of Ministry of Education, Xiangtan University(2011ICIP06)

1 引言

网络及移动计算技术的稳步发展使用户可通过多种接入设备实时访问网络视频资源。然而在视频传输过程中,对用户所提出的不同要求和终端性能的差异、网络异构特性所带来的各支路服务能力的差异以及网络拥塞和噪声等影响传输的因素等使得对视频信息的多质量服务需求急剧增加,这样对视频压缩编码技术的新需求——可分级编码便成为了一个热点问题^[1,2]。近年来,小波在图像编码方面的巨大成功使人们转去思考基于小波的视频可分级编码问题,出现了 2 类基于小波的视频可分级编码方法^[3],一类是基于运动补偿的 2D 小波视频编码方法,该类方法的难点是如何有效地实现小波域帧间运动估计,避免误差“漂移”问题的产生;另一类是基于 3D 小波视频编码方法^[4],该类方案以其无块效应和可分级特性的易于实现而受到关注^[5-8],然而由于在时间维的滤波一般并不与视频对象的运动方向相一致,这样通常会影响到时间域冗余的去除,编码效率不是很高。JVT 也曾在 2004 年成立了一个特别研究组来探索基于小波的更高效的编码^[9],并进一步支持除可分级编码外的其他特性^[10],如视频高分辨率存储与传输和极细粒度的信噪比可分级等。总之,基于小波的视频可分级编码技术还处于不断发展中,很多问题还有待进一步研究^[11]。

本文首先对视频 3D 小波系数的分布特性进行了分析,提出了一种混合的 3D 空间树结构,并结合 HVS 对各子带敏感程度的差异,给出了一个子带系数加权方案,进一步提出本文的视频编码架构,架构充分利用了视频 3D 小波变换后时间域低频帧与高频帧的系数分布特点,根据尺度间和尺度内小波系数幅值相关性,在时间域低频子带和高频子带分别采用相应的系数结构来组织重要系数的编码,从而使得幅值较大的系数排在码流前端,提高了编码效率,特别对较低码率情况。

2 相关工作分析

对视频小波系数进行可分级编码需要解决的一个重要问题是如何对重要系数进行定位。早期方案大都采用树型结构^[12,13]对时间域滤波后的系数进行组织,这种树型对时间域低、高频系数的组织具有相同的结构。但从文献^[11]的研究可以看出,

视频在空间和时间维所具有的相关性不尽相同,并且通常会存在较大的差异,时间维的相关性往往要大于空间维的相关性,这样时间维的高频系数幅值一般要小于空间维的高频系数。可见这种传统对称树型结构一般不利于大幅值重要系数的前置,将会对编码效率产生影响,特别是对中、低码率的编码。文献^[5]采用了一种近似对称的空间方向树,获得了高于对称树结构的编码效率,但它要求时、空方向的小波分解级数必须相同,这不利于发掘不同方向的系数相关性;文献^[14]改进了 3D-SPIHT 的空间方向树,将时间低频帧中空间低频的部分系数与前一粗尺度下的高频帧系数建立了关联,对时间维的低频系数和空间维的高频系数进行优先编码,但由于该方法没有考虑高频系数相关性较低这一特点,会在一定程度上影响时间高频帧的编码效率;文献^[15]以零树为基础建立了一种非对称树结构,将那些对父节点具有较高依赖程度的时空系数置于同一层,与对称的树结构相比,该结构有利于零树包含更多系数,在一定程度上改进了编码效率;文献^[6,16]提出了“虚拟树”的概念,即在构造树结构时假定最低频子带又进行了一次小波分解,从而增加分解层次来使一棵树包含更多系数,但仍然忽略了位于同一层零树上时空系数成为重要系数的概率不同的事实,其编码效率尚有提高空间;文献^[17]从时空系数相关性不一致出发,将时空方向树分解为时间方向树和空间方向树,并优先处理前者。只有当时间方向树的某个系数为重要系数时,才去扫描该系数所在的空间方向树,但该方案会推迟对空间方向树中孤独零的处理;文献^[18]在零树的基础上,将亮度分量作为色度分量的父亲,以增加树结构的长度,在低码率下提高了编码效率。文献^[19]在充分考虑几种基本树结构的基础上,将“零块”思想引入其中,以挖掘系数的子带内相关性,但其同一层系数成为重要系数的概率不同这一事实仍然存在。总之,现有算法对各子带大幅值系数的同等对待忽略了这些系数对重构视频质量的差异,不利于提高视频解码的主观质量。

3 3D 混合树结构的提出

3.1 小波变换系数的相关性分析

对基于树结构的编码方案,不同子带系数的相关性是设计树形结构的重要依据^[15]。

记 $p(x, y)$ 为视频帧中 (x, y) 位置的灰度值,则 2

像素的自相关函数 $r(m, n)$ 定义为

$$r(m, n) = E\{p(x, y)p(x + m, y + n)\} \quad (1)$$

其中, $m, n \in \mathcal{C}$, $E\{\cdot\}$ 为数学期望。电视画面的自相关系数一般在 0.95 至 0.98 之间^[20]。

对于 $l(l-1)$ 级的视频帧小波变换, 2 个变换系数间的自相关函数 r_w^l 为^[11]

$$r_w^l = \sum_{k=1}^K \sum_{s=1}^S \sum_{j=1}^J \sum_{t=1}^T h_v(k)h_v(s)h_v(j)h_v(t) \cdot r_w^{l-1}(2m + s - t, 2n + k - j) \quad (2)$$

其中, h_h 和 h_v 分别表示 2D 小波变换水平和垂直方向的滤波器, K 和 J 为 h_v 的长度, S 和 T 为 h_h 的长度, $r_w^0(m, n) = r(m, n)$ 。

对于低频子带(LL), 其自相关函数为

$$r_{LL}^l(m, n) = \sum_{k=1}^K \sum_{s=1}^S \sum_{j=1}^J \sum_{t=1}^T h_L(k)h_L(s)h_L(j)h_L(t) \cdot r_{aLL}^{l-1}(2m + s + t, 2n + k - j) \quad (3)$$

而由小波低通滤波器特性, $\sum_{k=1}^K \sum_{s=1}^S \sum_{j=1}^J \sum_{t=1}^T h_L(k)h_L(s)h_L(j)h_L(t)$ 通常为 4 (当对其进行归一化时, 其值为 1)。可见时间和空间低频系数间的相关性一般会较高。

对于高频子带 $B \in \{HL, LH, HH\}$, 自相关函数为

$$r_{HL, LH, HH}^l(m, n) = \sum_{k=1}^K \sum_{s=1}^S \sum_{j=1}^J \sum_{t=1}^T h_B(k)h_B(s)h_B(j)h_B(t) \cdot r_{HL, LH, HH}^{l-1}(2m + s - t, 2n + k - j) \quad (4)$$

依据低通和高通以及水平和垂直高通 2 对滤波器间的正交性, 有

$$\sum_{k=1}^K \sum_{s=1}^S \sum_{j=1}^J \sum_{t=1}^T h_L(k)h_H(s)h_L(j)h_H(t) = 0 \quad (5)$$

则 $r_{HL, LH, HH}^l(m, n) \approx 0$ 。特别对于严格服从宽平稳情况的信号有 $r_{HL, LH, HH}^l(m, n) = 0$ 。可见, 高频系数间具有去相关 (或近似去相关) 属性。

3.2 时、空方向树混合结构

3.2.1 传统树形结构分析

图像编码的 EZW^[12]和 SPIHT 算法^[13]的核心是采用如图 1 所给出的零树来组织图像的小波系数, 其中 EZW 零树基于尺度间系数幅值的衰减来关联父子带与孩子子带间的系数; SPIHT 零树则在此基础上增加了对相邻系数所具有相关属性的考虑。由

前一节知道, 时域低频与空域低频子带的相邻系数间的相关性较高, 如果直接采用 EZW 树结构来对系数进行组织, 这种系数间的相关特性将会被忽略掉, 同时会较 SPIHT 零树增加更多的同步信息; 此外, 对于相关性较低的高频系数 (尽管如此, 这些高频系数仍将会近似满足尺度间幅值的衰减特性, 如图 1(b)中的白、灰色像素), 若采用 SPIHT 树结构来对高频系数进行组织, 其父系数 (参见图 1(b)所示 LL 中的黑色像素) 及对应的子带中的系数可能不会构成零树, 这是由于此时可能会错过诸如 EZW 树结构中相应的节点, 进而“孤立点”的数量将大幅增加, 势必会影响最后的编码效率。对于通过 MCTF 所产生的帧差, 通常其高频系数的相关性会更低。总之, 单一的树形结构很难同时有效地兼顾时间域低、高频帧的系数分布情况, 难以提高整体编码效率。

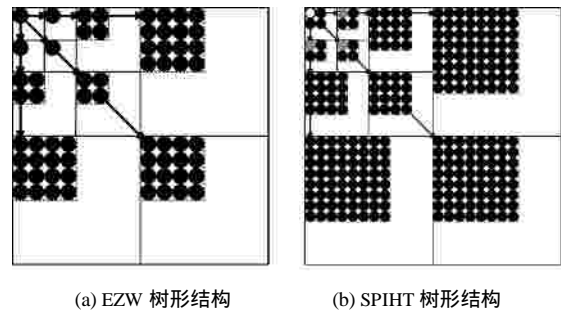


图 1 2D 情况下的 2 种树结构

3.2.2 混合树形结构

本文提出一种混合时空方向树结构, 图 2 为 4 帧情况的示意, 其中在时间域上进行了 2 级滤波, 空间域为 1 级, 箭头经父节点至子节点, 父子关系由不同尺度子带中的相同字母来表示, 而兄弟关系则由同尺度的相同字母来表示。

提出树形结构的内部构成: 1) 时间域低、高频帧分别采用 SPIHT 树形结构和 EZW 树形结构; 2) 由于经 MCTF 处理后的视频帧的能量主要集中在时间低频帧, 这样只将时间低频帧中空间 LL 子带与时间高频帧中子带构成父子关系, 这样时间低频与空间高频子带中的系数势必会得以优先编码; 3) 由于时间高频帧中的信息主要是由运动补偿的误差所引起的, 这样运动物体的边缘上将会聚集主要的能量, 可见时域高频与低频帧系数的分布是不同的, 时间高频帧的空间高频子带将会出现较大的幅值系数, 通常具有更高比例的能

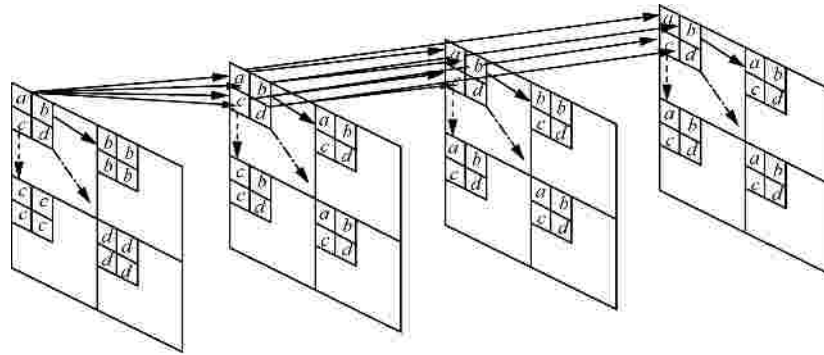


图 2 时、空混合方向树

量。这样，提出的混合树形结构将时域高频帧的空间 LL 系数与时域更高频帧对应的子带系数构建父子关系，即将空域高、低频系数同等看待，这将会使物体边缘的大幅值系数编码的优先级得以提高。

4 基于 HVS 特性的小波系数加权

HVS 对信号频率的敏感程度是不同的^[21]，一般而言，对于高、中、低 3 个频率分量，HVS 对中频会更加敏感一些。此外，对于方向不同的频率分量也会表现出一定不同的敏感程度，通常 HVS 对竖直和水平方向的频率会较对角线方向的频率更加敏感。而对于图像及视频，若使 HVS 敏感的频率系数优先编码，在相同的率失真下将会获得主观效果更好的重构图像。依据 HVS 所具有的这种对频率系数敏感性的差异，提出一种对小波子带的加权方案，目的是使 HVS 敏感的子带系数被优先编码。

因为小波多尺度特性与 HVS 初级视觉的空域频率多通道相吻合，因而可考虑通过空域频率揭示各子带的特性。文献[22]采用 CSF(contrast sensitivity function)来描述 HVS 的频率敏感程度

$$H(f) = 2.6 \times (0.192 + 0.114f) e^{-(0.114f)^{1.1}} \quad (6)$$

其中， f 为空间频率，其单位为周期/度，取值范围为 $[0, 32]$ ， $H(f)$ 为 HVS 对 f 的敏感程度。首先对 f 进行采样并通过式(6)获得一系列 CSF 数值，对这些数值进行小波分解，并计算各子带的平均值，对这些平均值进行线性拉伸，使其最小值为 1；其次，将所得到的 1D 子带的权重值通过图 3 的顺序进行 2D 子带映射，以此形成 2D 小波子带的视觉权重（如表 1 所示，其中，小波变换为 3 级）。

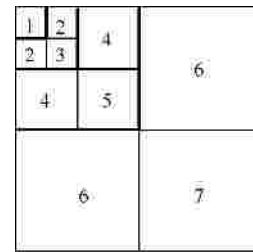


图 3 子带权重顺序

表 1 所示权重来自于自然图像，考虑到时域低频帧系数的分布与自然图像相类似，本文算法仅对时域低频帧进行了如此加权处理，而对于时域高频帧不适合进行相应的加权处理。

表 1 3 级 9/7 小波分解后得到的视觉权重

	LL	LH	HL	HH
粗尺度层 (3 层)	3.500	5.300	5.300	7.200
2 层	—	4.740	4.740	3.750
细尺度层 (1 层)	—	2.330	2.330	1.000

5 本文提出的视频可分级编码算法的实现

在混合时空方向树结构和基于 HVS 加权处理的基础上，给出如下基于 3D 小波的视频可分级编码算法。

符号约定如下。

T_i 为进行第 i 轮量化所采用的阈值，初值为 T_0 。

LIP、LSP 分别为不重要系数表和重要系数表。

LIS 为不重要像素的集合。

C 为标志位，用于标识当前所处理的图像分量，其取值分别为 Y、U 和 V。

$O(x, y, z)$ 表示节点 (x, y, z) 的直接子节点集合。

$D(x, y, z)$ 表示节点 (x, y, z) 的所有子系数集合。

$L(x, y, z)$ 表示节点 (x, y, z) 的除直接子节点外的所有子系数集合。

A、B 型树分别指在编码 LIS 的每个节点 (x, y, z) 时分别需要检查 $D(x, y, z)$ 、 $L(x, y, z)$ 是否有重要系数。

W 和 h 分别为空间 LL 的宽、高度。

F 为时域最低频帧的数量。

算法的主要过程如下。

Step1 对输入视频进行 GOP 划分 (每 16 帧为一个 GOP), 并且对每个 GOP 分别进行 L_t 级 MCTF 和 L_s 级小波变换 (其中, L_t 和 L_s 实验中分别选为 4 和 3)。

Step2 对 3D 小波子带按第 4 节的加权模型加权处理。

Step 3 进行初始化。

Step 3.1 对每个 GOP 的时空小波系数, 计算其初始量化阈值 T_0

$$T_0 = \left\lfloor \text{lb}(\max_{c_n \in C} |c_n|) \right\rfloor \quad (7)$$

其中, C 为加权处理的小波系数集合;

Step 3.2 赋 $i=0$, $LSP=f$; 将时、空方向同时为最低频子带的所有系数 (参见图 2 中第一帧左上角的 a, b, c, d 系数) 放入 LIP 和 LIS 中, 其中, Y 分量系数在前 (此时令 $C=Y$), U, V 分量系数在后 (此时分别令 $C=U$ 和 $C=V$)。同时记 LIS 中每个节点的类型为 A 型。

Step 4 重要系数搜索。

Step 4.1 根据分量类型, 对 LIP 中的节点 (x, y, z) , 将其对应的系数与相应的阈值 T_i 比较, 当系数的绝对值大于或等于 T_i 时, 输出 “1” 及符号位, 同时将系数移至 LSP 表; 否则输出 “0”。

Step 4.2 按 3.2 节时空混合方向树搜索重要系数并对 LIS 进行调整。对于父节点 (x, y, z) , 其直接子节点可划分为下列 4 种情况。

Case 1 若 (x, y, z) 是时间域及空间域低频子带中的节点 (如图 2 第一帧左上角的 a, b, c, d 所示), 其直接子节点 (下一层) 又被划分为 2 种情况。

1) 如果为左上角 2×2 块中的 a 系数, 其子节点集合为

$$O_a(x, y, z) = \{(x, y, z + f), (x + 1, y, z + f), (x, y + 1, z + f), (x + 1, y + 1, z + f)\}$$

2) 若为 2×2 小块中的 b, c, d 节点, 此时子节点的集合为

$$O_b(x, y, z) = \{(x + w - 1, y, z), (x + w, y, z), (x + w - 1, y + 1, z), (x + w, y + 1, z)\}$$

$$O_c(x, y, z) = \{(x, y + h - 1, z), (x + 1, y + h - 1, z), (x, y + h, z), (x + 1, y + h, z)\}$$

$$O_d(x, y, z) = \{(x + w - 1, y + h - 1, z), (x + w, y + h - 1, z), (x + w - 1, y + h, z), (x + w, y + h, z)\}$$

Case 2 若 (x, y, z) 为时域最高频的空间 LL 中的节点 (参见图 2 三、四帧左上角 a, b, c, d), 其直接子节点集合:

$$O(x, y, z) = \{(x + w, y, z), (x, y + h, z), (x + w, y + h, z)\}$$

Case 3 若 (x, y, z) 是时间域其他空域低频子带中的节点 (参见图 2 中第 2 帧的左上角 a, b, c, d 系数), 此时其直接子节点集合为

$$O(x, y, z) = \{(x + w, y, z), (x, y + h, z), (x + w, y + h, z), (x, y, 2z), (x, y, 2z + 1)\}$$

Case 4 若 (x, y, z) 为各帧中空域高频子带的节点 (参见图 2 中第一帧右上角 4 个 b , 以及第二帧中右上角 a, b, c, d 系数), 此时其直接的子节点集合为

$$O(x, y, z) = \{(2x, 2y, z), (2x + 1, 2y, z), (2x, 2y + 1, z), (2x + 1, 2y + 1, z)\}$$

Step 4.3 针对当前分量 C , 对 LIS 表中对应每个节点 (x, y, z) 的重要性进行检测。如果当前树是 A 型, 转至 Step 4.3.1; 否则转到 Step 4.3.2。

Step 4.3.1 检测 $D(x, y, z)$ 的重要性, 如果重要, 输出 “1”, 并通过阈值 T_i 判断该节点的直接子节点的重要性, 若重要, 输出 “1” 及符号位, 并将其移入 LSP; 如不重要则输出 “0”, 并将其移至 LIP 表。进一步, 如果 $L(x, y, z)$ 存在, 将在 LIS 表的末端移入该节点, 来作为 B 型树。若 $D(x, y, z)$ 不重要, 则输出 “0”。

Step 4.3.2 检测 $L(x, y, z)$ 的重要性, 如果重要则输出 “1”, 同时把该节点的直接子节点移至 LIS 表的末端作为 A 型树, 并把父节点自 LIS 中删除; 如果 $L(x, y, z)$ 不重要, 输出 “0”。

Step 5 重要系数幅值细化。

Step 5.1 针对当前分量 C , 考察 LSP 表中对应的重要系数, 如果它位于区间 $[T_i, 1.5T_i)$ 则输出 “0”, 否则输出 “1”。如果形成的码流达到了目标码率, 则结束算法; 否则转向 Step 5.2。

Step 5.2 $i = i + 1$, $T_i = T_{i-1} / 2$ 。如果 $T_i = 0$, 结束算法；否则转向 Step 4。

6 实验与分析

对 CIF 30Hz 格式的 6 个视频序列的 1~96 帧或 1~128 帧进行仿真实验，这 6 个视频序列分别为：Hall Monitor、Akiyo、Miss America、Mobile & Calendar、Foreman 和 Mother & Daughter。实验在 VidWav 平台上进行，每个 GOP 为 16 帧，选用 $t+2D$ 方式的 3D 小波变换，时间维进行了 $L_t = 4$ 级 $5/3$ 小波的 MCTF，空间采用了 $L_s = 3$ 级的 $9/7$ 提升小波的分解，对 Y、U、V 3 个分量进行顺序编码。实验结果与文献[15]的非对称树结构方案和文献[17]的时空方向树方案进行了比较，其中，文献[15]与本文算法的结果在 VidWav 平台上得到，文献[17]的结果源自原文。

6.1 主观解码质量

由于兼顾了 HVS 视觉特性，有利于解得高主观质量的视频图像。所提出算法与非对称树编码方案在 128kbit/s 和 512kbit/s 码率下编、解码 Miss

America 和 Mobile & Calendar 序列，并进行了比较。图 4 为 2 序列第 1 帧的解码结果。从解码图像可看出，在 128kbit/s 码率下，所提出算法的“振铃效应”要弱于文献[15]算法；在 512kbit/s 码率下，所提出算法获得了较高的主观解码质量，而此时文献[15]算法解码帧中振铃效应仍较为明显（参见图 5 中数字边缘及上方的日历）。可见本文小波子带系数加权方案将会在一定程度上改善重构视频的主观质量。

6.2 解码客观质量

本文对 Mother & Daughter、Hall Monitor、Akiyo、Miss America 和 Mobile & Calendar 的 1~96 帧在 128kbit/s、256kbit/s 和 512kbit/s 码率下解码的 PSNR 进行了统计（如表 2 所示），对于运动幅度低的 Miss America 和运动幅度高且细节丰富的 Mobile & Calendar，本文算法均获得了较高的 PSNR。提出算法 Y 分量解码的平均 PSNR 高于文献[15]0.65dB，而 U、V 解码分量较文献[15]平均高 1.75dB 和 1.77dB。

表 3 列出了 Foreman 和 Miss America 的 1~128 帧



图 4 本文算法与文献[15]中非对称树结构算法解码的主观质量比较

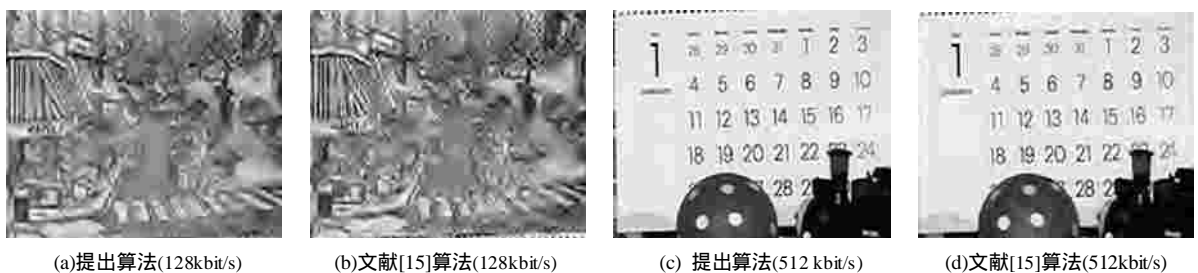


图 5 本文算法与文献[15]中非对称树结构算法解码的细节主观质量比较

表 2 不同码率下所提出算法与非对称树结构算法的重构结果比较

视频序列	码率/(kbit·s ⁻¹)	PSNR/dB					
		文献[15]方法			本文方法		
		Y	U	V	Y	U	V
Hall Monitor	128	30.920	36.460	38.820	31.500	38.020	40.130
	256	34.580	37.500	39.570	35.930	39.150	41.240
	512	37.350	38.710	40.800	38.530	39.880	42.020
Miss America	128	40.040	38.380	40.500	40.510	39.280	42.030
	256	40.590	38.750	41.070	42.000	39.750	43.130
	512	42.330	40.270	42.570	43.030	40.780	43.760
Mother & Daughter	128	35.400	40.680	41.330	35.740	42.600	42.610
	256	38.740	42.470	43.600	38.930	44.870	45.060
	512	41.540	44.560	45.550	41.880	46.430	47.090
Akiyo	128	36.730	39.410	40.610	37.480	41.890	43.970
	256	40.370	42.810	43.710	40.990	45.190	46.800
	512	43.510	45.860	46.880	44.170	47.690	48.500
Mobile & Calendar	128	20.450	26.120	24.730	20.670	28.040	26.280
	256	23.240	28.470	27.240	23.430	30.560	28.860
	512	25.790	30.440	29.590	26.530	33.060	31.620

表 3 不同码率下所提出算法与时空方向树结构算法的重构结果比较

视频序列	码率/(kbit·s ⁻¹)	PSNR/dB					
		文献[17]方法			本文方法		
		Y	U	V	Y	U	V
Foreman	500	34.840	37.630	39.420	34.960	40.980	41.630
	1 000	37.760	39.630	41.760	37.460	42.740	43.890
	1 500	39.400	40.500	42.620	38.880	43.690	45.040
Miss America	500	41.760	40.640	42.270	42.980	40.740	43.660
	1 000	43.230	40.690	43.460	43.710	41.580	44.270
	1 500	43.990	40.690	43.460	44.370	42.720	44.840

表 4 本文算法与基于非对称树结构算法以 GOP(16 帧)为单位的平均编、解码时间比较

视频序列	码率/(kbit·s ⁻¹)	文献[15]方法			
		文献[15]方法		本文方法	
		编码时间/ms	解码时间/ms	编码时间/ms	解码时间/ms
Miss-America	128	247.5	32.8	137.3	30.3
Mobile	512	370.7	126.7	236.7	113.5

分别在 500kbit/s、1 000kbit/s 和 1 500kbit/s 下本文算法与文献[17]算法重构 PSNR 的统计结果,对于 Y 分量,所提出算法获得的平均 PSNR 较文献[17]高 0.23dB,而解码 U、V 分量的平均 PSNR 分别高出文献[17]2.11dB 和 1.72dB。可见中低码率下本文算法的解码质量与[17]相比有一定改进,这源于所提出的混合时空方向树可以更加有效地对大幅值系数进行定位。

6.3 编、解码时间比较

表 4 给出了本文算法与基于非对称树结构算法

以 GOP(16 帧)为单位,前 96 帧视频序列的平均编解码时间比较,从表中可以看出,本文算法在编码方面所用时间要低于基于非对称树结构算法,解码时间与对比文献相差不大,而且从文献[17]可知,基于非对称树结构算法与基于时空方向树的编码算法在时间复杂度方面近似,因此本文算法的时间复杂度要低于基于非对称树结构与基于时空方向树结构的 2 种对比算法。分析原因主要在于本文算法提出的混合时空方向树结构更有利于定位重要

系数, 从而减少扫描次数, 降低了编码时间。

7 结束语

本文在分析 3D 小波变换时域低、高频帧系数分布特点的基础上, 提出一种混合时空方向树结构, 并依据 HVS 特性提出一种基于小波系数敏感函数的子带加权模型, 进一步提出了一种视频可分级编码算法。所提出的算法的优点体现在: 1) 根据低、高频系数自相关特性, 采用了相应的树形结构, 减少了用于重要系数定位的同步信息; 2) 根据 HVS 对不同频率分量的敏感差异, 对小波系数进行了加权改造, 提高了解码的主观视觉质量。

参考文献:

- [1] SCHWARZ H, MARPE D, WIEGAND T. Overview of the scalable video coding extension of the H.264/AVC standard[J]. *IEEE Trans on Circuits and Syst for Video Technol*, 2007, 17(9):1103-1120.
- [2] 王相海. 视频细粒度可分级编码研究进展[J]. *计算机科学*, 2006, 33(7):75-80.
WANG X H. Research progress on fine granularity scalability video coding[J]. *Computer Science*, 2006, 33(7):75-80.
- [3] 王相海, 宋传鸣. 图像及视频可分级编码[M]. 北京: 科学出版社, 2009.
WANG X H, SONG C M. *Image and Video Scalable Coding*[M]. Beijing: Science Press, 2009.
- [4] DAS A, HAZRA A, BANERJEE S. An efficient architecture for 3-D discrete wavelet transform[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(2):286-296.
- [5] KIM B, XIONG Z, PEARLMAN W A. Low bit rate, scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)[J]. *IEEE Trans on Circuits and Syst for Video Technol*, 2000, 10(12): 1374-1387.
- [6] 丁文奇, 胡佳, 张立明. 充分减小树间冗余的优化三维 VSPiHT 视频编码方法[J]. *计算机辅助设计与图形学学报*, 2005, 17(3):563-569.
DING W Q, HU J, ZHANG L M. Optimal 3D-SPIHT video coding method by reducing redundancy between trees[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2005, 17(3):563-569.
- [7] CHENG C C, PENG G J, HWANG W L. Subband weighting with pixel connectivity for 3-D wavelet coding[J]. *IEEE Trans on Image Process*, 2009, 18(1):52-62.
- [8] WANG H M, PAN X Z. Video compression coding based on the improved 3D-SPIHT[A]. 2010 International Conference on Computer Application and System Modeling (ICCAASM)[C]. Taiyuan, China 2010.108-111.
- [9] ISO/IEC JTC 1/SC 29/WG 11. Wavelet Video Coding – an Overview:w7824[R]. MPEG 2006.
- [10] ISO/IEC JTC 1/SC 29/WG 11. Status Report on Wavelet Video Coding Exploration:N7822[R]. MPEG 2006.
- [11] 宋传鸣. 基于多尺度分析的视频可分级编码技术研究[D]. 南京: 南京大学, 2010.
SONG C M. Research on Scalable Video Coding Based on Multiscale Analysis[D]. Nanjing: Nanjing University, 2010.
- [12] SHAPIRO J. Embedded image coding using zerotree of wavelet coefficients[J]. *IEEE Trans on Signal Process*, 1993, 41(12):3445-3462.
- [13] SAID A, PEARLMAN W A. A new, fast, and efficient image codec based on set partitioning in hierarchical trees[J]. *IEEE Trans on Circuits and Syst for Video Technol*, 1996, 6(3):243-250.
- [14] 张宗平, 刘贵忠, 侯兴松. 一种改进的三维小波视频编码[J]. *西安交通大学学报*, 2001, 35(6):595-599.
ZHANG Z P, LIU G Z, HOU X S. Improved 3D-wavelet video coder[J]. *Journal of Xi'an Jiaotong University*, 2001, 35(6):595-599.
- [15] HE C, DONG, ZHENG, GAO Z. Optimal 3-D coefficient tree structure for 3-D wavelet video coding[J]. *IEEE Trans on Circuits Syst and Video Technol*, 2003, 13(10):961-972.
- [16] KHAN E, GHANBARI M. Very low bit rate video coding using virtual SPIHT[J]. *IEE Electronic Letters*, 2001, 37(1):40-41.
- [17] ZHANG L, WANG D, VINCENT A. Decoupled 3-D zerotree structure for wavelet-based video coding[J]. *IEEE Trans on Broadcasting*, 2008, 54(3):430-436.
- [18] KASSIM A A, TAN E H, LEE W S. 3D color set partitioning in hierarchical trees[J]. *Circuits Syst Signal Process*, 2009, 28:41-53.
- [19] MOINUDDIN A A, KHAN E, GHANBARI M. The impact of tree structures on the performance of zerotree-based wavelet video codecs[J]. *Signal Processing: Image Communication*, 2010, 25(3):179-195.
- [20] 姚庆栋, 毕厚杰, 王兆华等. 图像编码基础[M]. 北京: 清华大学出版社, 2006.
YAO Q D, BI H J, WANG Z H, *et al.* *Fundamentals of Image Coding*[M]. Beijing: Tsinghua University Press, 2006.
- [21] 李玲, 王向阳. 基于视觉敏感特性的小波域图像编码算法研究[J]. *小型微型计算机系统*, 2010, 31(4):780-783.
LI L, WANG X Y. New image coding algorithm contrast sensitivity[J]. *Mini-MicroSystems*, 2010, 31(4):780-783.
- [22] BEEGANA P. Wavelet-based image compression using human visual system models[A]. *Electrical Engineering Department of Virginia Polytechnic Institute*[C]. Virginia, USA, 2001.

作者简介:



付明哲 (1985-), 男, 辽宁鞍山人, 辽宁师范大学硕士生, 主要研究方向为视频编码与通信。

王相海 (1964-), 男, 吉林汪清人, 博士, 辽宁师范大学教授、博士生导师, 主要研究方向为计算机图形学及多媒体信息处理。

宋传鸣 (1980-), 男, 辽宁沈阳人, 博士, 辽宁师范大学讲师, 主要研究方向为图像及视频编码。