

基于三维小波变换嵌入式视频压缩算法的研究

杨春玲, 余英林

(广州华南理工大学电子与通信工程系, 广东广州 510640)

摘 要: 本文把二维图像的等级树集分割编码方法(SPIHT)推广应用于三维的视频编码中,提出了一种用于三维视频序列图像的扩展等级树集分割编码方法.此方法所编码码流是嵌入式码流,在解码端可随意截取一段码流解码以达到所要求的码率和视频质量.另外此方法无需进行运动补偿,降低了算法的复杂度.通过理论分析和仿真实验得出,此方法有很好的编码效率,但对于帧内图像较复杂的序列,编码效率有所下降.

关键词: 视频压缩;小波变换;等级树集分割编码;扩展等级树集分割编码方法;嵌入式码流

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 0372-2112(2001)10-1381-03

Research on Embedded Video Compression Based on 3D-Wavelet Transformation

YANG Chun-ling, YU Ying-lin

(Dept. Of Electronics Engineering, South China University of Technology, Guangzhou, Guangdong 510640, China)

Abstract: The Set Partitioning in Hierarchical Trees (SPIHT) coding method is extended to 3D video coding in this paper. The extended SPIHT (ESPIHT) is proposed. This method creates an embedded compressed bit stream that can be truncated to achieve a desired bit rate or a desired video quality. Furthermore, motion estimation and motion compensation are needless in this method. By theoretical analysis and a lot of simulations, it can be found that ESPIHT is very effective for video sequence coding, but its compression effectiveness is somewhat reduced when the complicated video sequence is coded.

Key words: video compression; wavelet transformation; set partitioning in hierarchical trees (SPIHT); 3D-SPIHT; embedded compressed bit stream

1 引言

视频编码技术在多媒体、高清晰度电视、可视电话、会议电视等视频通信服务中起着至关重要的作用,因而具有广阔的应用前景和巨大的市场价值.

与静止图像的编码方法相比,视频图像序列编码有其自身的特点.使图像序列编码压缩比远远超过静止图像的编码压缩比成为可能.传统的视频编码方法是运动补偿加帧内DCT变换编码,这种结构已被现有的各种编码标准所采用,如H.261/H.263^[2,3]、MPEG^[4,5]等.嵌入式编码方法对图像进行编码时,图像的重要信息先编码而不重要信息后编码,所得码流是完全嵌入式的,在解码端任意截取一段码流进行解码都可以得到一幅完整的图像,且所截码流越长,解码图像的质量越高.这种编码方法对于网络通信及无线通信很有用处,现已引起了广大学者的关注^[6,7].Shapiro提出的嵌入式零树编码方法^[8]以及Amir Said和W.A. Pearlman提出的SPIHT编码方法为基于小波变换的嵌入式编码奠定了基础.本文把等级树集分割编码(SPIHT)方法推广应用到三维的视频编码中,提出

了扩展等级树集分割编码方法,此方法的编码码流是嵌入式的,且无需进行运动估计和运动补偿.

2 视频序列的三维等级树集分割编码方法

本文以十六帧图像为一组进行讨论三维小波变换和三维等级树集分割编码方法.

2.1 三维小波变换

本文用到的三维小波变换是基于文献[1]中所提出的空间三维小波变换结构,只是在本文的应用中进行每级小波变换时,先对时间轴方向进行变换,再对空间方向进行水平和垂直方向的变换,仿真实验证明,小波变换时对三个方向上变换的先后顺序对小波系数无影响.对 $L \times N \times M$ 的图像序列进行四级小波变换后的子带示于图1, $N \times M$ 表示视频序列每帧的像素点数, L 为一组序列的帧数.小波变换后仍是大小为 $L \times N \times M$ 的三维数组.经四级小波分解后得到29个子带,分别表示不同方向上不同分辨率的信息.在图1中,为了描述等级树结构的方便,把最低级子带放大重画在下端.图中方向向下的箭头表示小波分解的方向.

收稿日期:2000-05-23;修回日期:2001-02-21

基金项目:国家自然科学基金(No. 69772026);广东省自然科学基金(No. 970484);华南理工大学自然科学基金(No. 106-E50950)

2.2 三维空间等级树

本文所述的等级树集分割编码方法是文献[9]中提出的二维等级树集分割编码方法在三维空间的推广. 在小波域进行编码主要是利用小波变换系数在不同的分辨率下同一方向的子带中的像素具有相似性. 等级树集分割编码方法较好地利用了这一性质, 因此此方法是目前认为最好的一种小波域编码方法. 在编码过程中, 是以树为单位进行编码的. 如图 1 所示, 方向向下的箭头表示小波分解方向, 向上的箭头是等级树的结构, 最大的等级树是由最低频中某一像素开始到某一分解方向的最高频为止而形成的. 在编码过程中, 随着所选阈值的减小, 等级树不断进行分割, 从而使得等级树不断减小.

本文所用符号尽量和文献[9]中所用符号保持一致, 但要注意, 本文中符号的意义已发生了少许变化. 首先给出本文中将要用到的一些集合的表示方法: $O(i, j, k)$ 表示小波系数像素点 (i, j, k) 的子代; $D(i, j, k)$ 表示小波系数像素点 (i, j, k) 的所有后代; H 表示所有的树根节点; $L(i, j, k) = D(i, j, k) - O(i, j, k)$ 表示小波系数像素点 (i, j, k) 除子代外的所有后代.

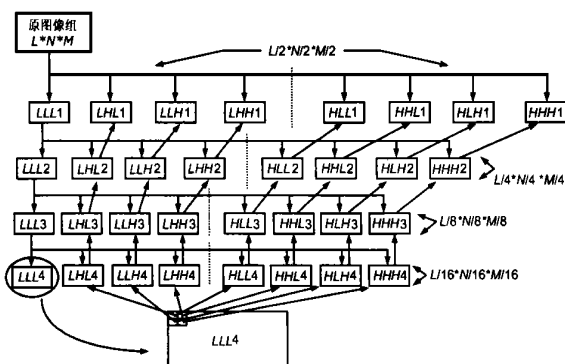


图 1 三维小波分解子带结构示意图

在三维等级树结构中, 最低频子带中的节点是树根. 除了最高频七个子带和一个最低频子带外, 每一个像素点都在同一方向上相应的高一级分辨率空间中有八个子代, 构成 $2 * 2 * 2$ 的立方块. 最高频子带中的像素无子代. 树根中各像素点的子代的形成方式不同, 分四种情况 (见图 1 所示).

2.3 编码方法

扩展等级树集分割编码方法在编码过程中把小波系数的重要信息放在三个表中, 分别为不重要集合表 (LIS), 其中的像素坐标 (i, j, k) 表示集合 $D(i, j, k)$ 或 $L(i, j, k)$, 分别称为 A 类集合或 B 类集合. 不重要像素表 (LIP), 其中是不重要小波系数像素的坐标; 重要像素表 (LSP), 其中是重要像素的坐标.

编码的开始是对三个表进行初始化, 重要像素表初始化为空集; 把最低频子带中的各像素坐标 $(i, j, 1)$ 写入不重要像素表和不重要集合表中, 其中 $i \leq M/2^4, j \leq N/2^4$ (因为本文是以四级三维小波分解为例来描述算法的, 若是 L 级三维小波分解, 则应为 $i \leq M/2^L, j \leq N/2^L$).

对三个表进行初始化后, 选定最大阈值, $n_{max} = \log_2(\max_{(i,j,k)}\{|c_{i,j,k}|\})$, 开始对不重要像素表和不重要集合表

进行扫描并编码, 用式 (1) 来计算扫描的像素或集合是否重要的.

$$S_n(T) = \begin{cases} 1, & \max_{(i,j,k)} |c_{i,j,k}| \geq 2^n \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

对于重要像素, 输出此像素的正负符号码流, 并把此像素的坐标从不重要像素表中移入重要像素表中; 对于重要集合, 按其分为 A 类集合或 B 类集合进行不同的处理, 处理方法和文献[9]相同, 这里不再详细介绍了, 有兴趣的读者可参考文献[9].

3 仿真实验

本文利用上节所述的三维小波变换系数的扩展等级树集分割编解码方法对 CIF 格式的标准序列 Miss American, Claire, Salesman 的亮度信号进行了编解码, 序列图像每帧的像素数为 $352 * 288$, 每像素用 8bit 表示. 在编解码过程中, 都是取 16 帧图像序列为一组进行小波变换和等级树集分割编码, 分别对 Miss American 和 Salesman 的 0~47 帧进行了测试, 对 Claire 序列的 0~31 帧图像进行了编解码测试, 解码图像的平均压缩比和峰值信噪比列于表 1 (注: 本文的实验结果是在扩展等级树集分割编码后没有进行自适应算术编码得到的). 峰值信噪比由下式计算:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (2)$$

其中 MSE 定义为:

$$MSE = \frac{1}{L * M * N} \sum_{k=1}^L \sum_{x=1}^M \sum_{y=1}^N [f_k(x, y) - \hat{f}_k(x, y)]^2 \quad (3)$$

其中: $N * M$ 为图像序列中每帧图像的像素点数, L 为进行测试的序列的帧数.

$$\text{压缩比} = \frac{N * M * L * 8}{\text{sumbit}} \quad (4)$$

sumbit 为达到某一精度时总的码流长度.

表 1 解码图像的峰值信噪比和平均压缩比

MissA		Claire		Salesman	
PSNR	CR	PSNR	CR	PSNR	CR
39.9724	102.4884	41.8048	59.9943	36.4545	24.3697
38.2224	219.4471	38.1878	113.2031	32.9131	52.7924
35.7607	501.7708	34.9302	226.0669	29.6187	136.6277

在达到一定压缩比时各帧图像的信噪比变化曲线示于图 2~4. Miss America 测试序列中信噪比最低的一幅图像的原图和恢复图像示于图 5.

本文所提算法和国内近年发表的利用小波变换进行视频编码的文献中所得结果比较, 发现本文的方法编码效率远远高于别的文献中的结果. 比较结果列于表 2. 文献[10]中, 作者提出的分级零树扫描的视频编码方法是在 EZW 方法的基础上发展而来的. 文献[11]中作者在第七章提出了一维小波 + 3D DCT 的混合编码方法. 在文献[12]中, 提出了改进的 SPIHT 方法, 也是在等级树集分割编码方法的基础上发展起来的一种视频编码方法, 此文献中作者只给出了在一定压缩比下的图像信噪比变化曲线, 故本文中和其结果比较也只能给出一个估计值.

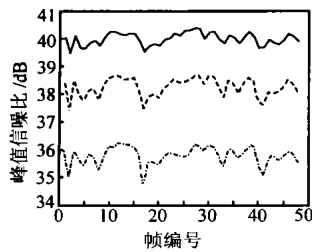


图 2 MissA 编码结果

实线平均压缩比为 102.4884
虚线平均压缩比为 219.4471
点划线平均压缩比为 501.7708

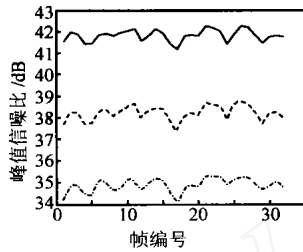


图 3 Claire 编码结果

实线平均压缩比 59.9943
虚线平均压缩比为 113.2031
点划线平均压缩比为 226.0669

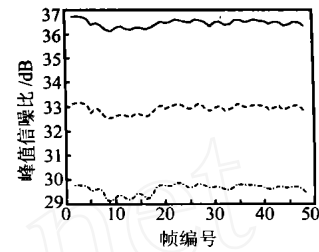


图 4 Salesman 编码结果

实线平均压缩比为 24.4545
虚线平均压缩比为 52.7924
点划线平均压缩比为 136.6277

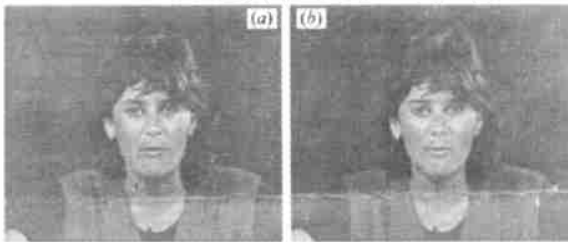


图 5 (a) MissA 测试序列中第 17 帧原始图像
(b) 压缩后的解码图像 (PSNR = 39.4661)

表 2 几种算法对 Miss American 进行编码的结果比较

算法	平均压缩比	信噪比
本文算法	219.4471	38.2224
文献[10]	100	37.56
文献[11]	160.39	38.27
文献[12]	112.6400	< 37.5

由本文对不同序列图像的仿真实验可知,对于每帧变化不太大的图像序列,用扩展的三维等级树集分割编码方法能达到很好的编码结果,但随着图像中边缘和细节信号的增加,此方法的编码效率在降低,这是因为边缘信息和细节信息多的图像中经小波变换后的高频分量的能量增多,因此在进行编码时,要达到一定的编码精度就需要较多 bit 数,因此基于小波变换的三维等级树集分割编码方法也不能很好地解决复杂图像的编码问题。

4 结论

本文把二维平面的等级树集分割编码方法推广应用于三维视频编码中,给出了十六帧为一组的图像序列进行四级小波变换后扩展的三维等级树集分割编码方法和步骤,仿真实验和理论分析发现,本文提出的扩展等级树集分割编码方法是一种较好的三维视频编码方法,和目前国内已有文献报导的基于小波变换的视频编码方法相比,有较好的编码效率。而且,本文所提方法是嵌入式编码方法,在编解码过程中,可根据需要随时停止编解码过程,得到不同精度和不同码率的恢复图像,这对现在发展起来的网络通信及多媒体通信有很重要的意义。另外利用本方法进行视频压缩无需运动补偿,使算法的复杂度降低。

通过本文的分析和实验还得出,对于较复杂的视频图像,扩展的等级树集分割编码方法的编码效率也不太高,如何更好地组织三维小波变换系数进一步提高复杂的视频图像编

码,将是本文作者下一步的研究目标。

参考文献:

- [1] 沈兰荪. 图像编码与异步传输 [M]. 北京:人民邮电出版社, 1998.
- [2] CCITT Recommendation H. 263. Video codec for audio visual services at P *64kbit/s [S]. 1990.
- [3] IUT-T Recommendation H. 263. Video coding for low bitrate communication [S]. Oct. 1995.
- [4] ISO/ IEC JCT1/ SC29 CD11172-2: MPEG1, international standard for coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s [S]. 1991.
- [5] ISO/ IEC JCT1/ SC29 CD13818-2: MPEG2, coding of moving pictures and associated audio [S]. 1993.
- [6] Ke Shen, Edward J Delp. Wavelet Based rate Scalable Video Compression [J]. IEEE Trans. On CAS for Video Technology, 1999, 9(1): 109 - 122.
- [7] Charles D Creusere. Fast embedded compression for video [J]. IEEE Trans. on IP, 1999, 8(12): 1811 - 1816.
- [8] J M Shapiro. Embedded image coding using zerotrees of wavelet coefficients [J]. IEEE Trans. On Signal Processing, 1993, 41(12): 3445 - 3462.
- [9] Anir Said, William A Pearlman. A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees [J]. IEEE Trans. On CAS For Video Technology, 1996, 6(3): 243 - 249.
- [10] 张旭东, 王德生, 彭应宁. 基于三维子波变换和分级零树扫描的视频编码算法研究 [J]. 电子学报, 1999, 27(7): 35 - 37.
- [11] 狄红卫. 低码率图像压缩研究 [D]. 华南理工大学, 1999, 4.
- [12] 林刚, 刘泽民. ATM 网中可变比特率小波视频编码和码率控制 [J]. 电子科学学刊, 2000, 22(1): 98 - 103.

作者简介:



余英林 (见本期第 1375 页)

杨春玲 女. 1970 年生于河南省长垣县. 现在华南理工大学电子与通信工程系工作. 已在国内外期刊及会议上发表多篇论文. 主要从事数据融合、图像处理方面的研究。

使得感知信息位于【0,1】区间内. 为了便于说明, 我们将其记作向量表示形式, 即 $K = \{ K_1, K_2, K_3 \}$, $k = \{ k_1, k_2, k_3, k_4, k_5 \}$.

5 试验结果

图像感知信息计算的初步实验采用图 2(a) 和图 3(a) 作为测试图像, 把它们经过低通滤波器的输出图像作为失真图像

像, 同时采用基于最小均方误差准则的最速下降算法进行图像恢复. 对于在一定迭代次数下得到的恢复图像, 我们采用 Sobel 算子进行边缘检测, 计算特征点, 特征线参数 x . 对恢复图像统计直方图的信息, 并用 FFT 工具统计频谱的变化, 计算参数 s . 对恢复图像做三级小波分解, 计算参数 m .

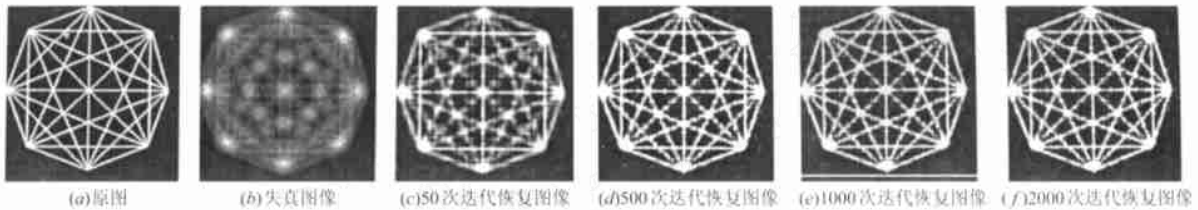


图 2

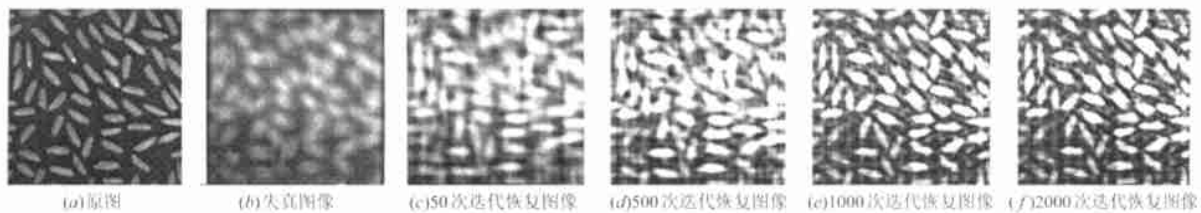


图 3

对于图 2, $K = \{ 1, 1, 1 \}$, $k = \{ 0.5, 0.5, 0.4, 0.6, 0.4 \}$, 图像的各项感知参数如表 1 所示.

表 1

	模糊图像	50 次迭代恢复图像	500 次迭代恢复图像	1000 次迭代恢复图像	2000 次迭代恢复图像
Y1	0.1074	0.1558	0.7910	0.9955	1
Y2	0.1329	0.3135	0.5279	0.9457	1
Y3	0.1333	0.3931	0.8695	1	1

表 2

	模糊图像	50 次迭代恢复图像	500 次迭代恢复图像	1000 次迭代恢复图像	2000 次迭代恢复图像
Y1	0.0263	0.0880	0.2629	0.6351	0.9457
Y2	0.1900	0.3270	0.5976	0.8760	0.9408
Y3	0.1951	0.4116	0.8524	0.9051	1

对于图 3, $K = \{ 1, 1, 1 \}$, $k = \{ 0, 1, 0.4, 0.6, 0.4 \}$, 图像的各项感知参数如表 2 所示. 试验结果表明在失真图像处理过程中, 图像的感知信息量是逐渐增加的, 感知信息量的计算方法也是有效可行的.

6 结论和展望

本文从观察者对图像的感知理解出发, 提出了从图像包含的感知信息的角度进行图像质量评价, 给出了图像感知信息的初步计算和度量方法, 试验结果初步说明了感知信息的有效性和可行性.

参考文献:

[1] Carlo colombo, Alberto Del Bimbo, Pietro Pala. Semantics in visual information retrieval [J]. IEEE journal of multimedia. 1999, 7(9) :38 - 53.

[2] Frank Nack, Adam T. lindsay. everything you wanted to know about

Mpeg-7 [J]. IEEE multimedia 1999, 7(9) :65 - 77.

[3] 魏征刚, 袁杰辉, 蔡元龙. 一种基于视觉感知的图像质量评价方法 [J]. 电子学报. 1999, 27(4) :79 - 84.

[4] 汪孔桥, 沈兰荪, 邢昕. 一种基于视觉兴趣性的图像质量评价方法 [J]. 中国图像图形学报. 2000, 4:300 - 303.

[5] Yinglin Yu, Chunling Yang, Jing Tian. Studies on information aspects of image processing system [A]. SCI '2000 国际学术会议论文集 [C]. 2000, Orlando USA.

[6] Anca L. Ralescu, James G. Shanahan. Perceptual organization for inferring object boundaries in an image [J]. Pattern recognition, 1999, 32: 1923 - 1933.

[7] 郭雷, 刘天明. 随机时分神经计算理论用于视觉曲线自激增强 [J]. 电子学报. 2000, 28(1) :35 - 38.

[8] Alberto Del Bimbo. Visual information retrieval [M]. Morgan kaufmann Inc. San Francisco california. 1999.

作者简介:

余英林 男. 1932 年 9 月生于广东省台山市. 1961 年获中国科学院副博士学位. 现为华南理工大学电子与通信工程系教授、博士生导师. 目前的主要研究方向包括图像与图形处理、神经网络、信号处理、模式识别、模糊技术等.



田 菁 男. 1979 年 11 月出生于山东省枣庄市. 1999 年获华南理工大学通信与信息系统专业学士学位. 现为华南理工大学信号与信息处理专业硕士研究生. 主要研究方向包括视频编码与图像处理、模式识别、图像分析等.