

文章编号:1000-582X(2007)04-0097-05

一种基于三维小波嵌入编码的上下文模型*

孙晓南

(重庆工程职业技术学院 计算机系,重庆 400037)

摘要:受 EZW、SPIHT、EBCOT 模型及 3D-EZW、3D-SPIHT、3D-ESCOT 模型的启发,综合利用小波系数不同子带间的自相似性、偏序性和同子带内相邻系数的相关性,提出了一种三维小波嵌入编码的上下文模型(3D-OWNER)。实验结果表明:利用这种模型获得的 PSNR 高于 3D-ESCOT 模型,同时也具有 3D-ESCOT 模型的优良性质。

关键词:上下文模型;可扩展编码;小波;嵌入编码;视频压缩

中图分类号:TP301.6

文献标志码:A

数字视频和图像压缩编码技术的研究,已历经半个世纪,在理论和工程上都取得了许多成果。进入 20 世纪 90 年代,在 ISO 和 ITU 等国际组织制定了几个通用的压缩编码标准,包括用于连续灰度和彩色静止图像的 JPEG^[1]、面向高品质数字视频和音频传输及存储的 MPEG-1、2^[2],以及适于低码率视频编码的 H. 26x^[3]。这些标准的算法主要由四类技术混合构成,即运动补偿、正交变换、量化和熵编码。20 世纪 90 年代中后期,在无线信道和 Internet 网上传输视频和图像成为人们研究的热点,于是,视频和图像编码的目标也就从传统的面向存储变为现在的面向传输。视频编码可分为两类:可伸缩视频编码与非伸缩性视频编码。新一代的视听对象编码的国际标准 MPEG-4^[4]中的精细可扩展性视频编码更提供了完全可扩展的增强层码流,它可以在任何地点截断,具有极强的网络带宽适应能力。后来又提出了渐进的精细可扩展性的编码方案,它保留了 FGS 的所有优良性能,并且把编码效率提高了将近 1 dB。

现有标准的伸缩性编码方案依赖于基本层的全部解码,如果带宽过低,导致基本层不能完全解码,则增强层不能解码。这将导致出现解码错误。为了满足不同网络、不同终端和不同需求条件下的应用,国际上提出了嵌入式编码解决方案^[5]。

嵌入式编码有两类:一类是分层的嵌入编码方案,

典型的代表有 EZW^[5]、SPIHT^[6]、ZTE^[7]等,主要是利用不同子带之间小波系数的偏序性和自相似性,将小波系数组织成树结构,形成一定的层次关系。另一类是基于上下文模型的编码方案,如 EBCOT 等^[8]。主要是利用子带内小波系数之间的相关性,组织高效的上下文,利用上下文来预测待编码的系数,编码的效率取决于上下文的选取。嵌入式编码的特性是码流可在任意处截断进行解码得到低码率的图像,而不需要码流最后解码的条件。因此,在每个码率下,嵌入的码流应具有那些能有效减少失真的码流应该放在压缩码流的前面这个特性。从这个原理出发,大多数位平面编码采用子带系数渐进从 MSB 到 LSB 编码,并且以 MSE 作为衡量失真的标准。

笔者受 EZW、SPIHT、EBCOT 模型及 3D-EZW、3D-SPIHT、3D-ESCOT 模型的启发,综合利用小波系数不同子带间的自相似性、偏序性和同子带内相邻系数的相关性,提出了一种新的三维小波嵌入编码上下文模型(3D-OWNER)。

1 模型概述

在这一节中,结合 EZW、SPIHT 提出的小波系数子带间的偏序性、自相似性及 EBCOT 提出的小波子带内相邻系数的相关性,综合利用这两种性质,修改 EBCOT 模型的上下文,加入待编码结点的父结点信息,

* 收稿日期:2007-01-17

基金项目:重庆市自然科学基金资助项目(8674)

作者简介:孙晓南(1972-),女,重庆工程职业技术学院讲师,研究方向为计算机硬件及应用。(Tel.):023-65210577;

E-mail:redapple4832@sina.com.

构造上下文信息,用来预测当前编码结点,提出了一种三维基于预测的子带编码模型(3D-OWNER).它同3D-ESCOT一样,具有良好的可伸缩性,支持率伸缩、分辨率伸缩和帧率伸缩.

1.1 三维小波变换

在进行三维的小波变换时,采用了三维的可分离的小波变换方案.即对视频序列图像的变换采用先时间域,再空间域变换的方法.考虑算法的效果,采用Daubechies 9/7小波作为标准的滤波器,因为它实现简单,具有较低的复杂度.在时间域的滤波中,为减少实现所需存储的帧数,采用32帧一组,进行三级三维小波变换.

1.2 小波系数编码

经过实验表明,小波系数统计近似满足Markov性:小波系数的显著性仅仅依赖于与其相邻的系数的值^[8],即待编码的小波系数的显著性与其直接相邻的系数的显著性情况有关.

JPEG2000中的EBCOT模型和3D-ESCOT模型充分利用这一性质,假定待编码的系数与其周围系数的显著性密切相关.根据小波变换的特点,3D-OWNER模型假设当前编码的结点与周围时间、空间、对角线方向上的10个结点密切相关.构造上下文.这点与3D-ESCOT相似.但3D-ESCOT没有利用EZW和SPIHT模型中提出的小波系数的偏序性和自相似性.

1.2.1 小波系数所构造的零树具有偏序性和自相似性

研究表明:不同子带之间具有自相似性,因此,如果从小波系数的同一空间方向来研究,有望得到比较好的偏序零树^[5-6].小波变换后,粗尺度子带的小波系数和细尺度子带的小波系数在空间上有一定的对应关系,即偏序性和相似性,大部分的能量集中在低频分量上,小波系数的大小从高层子带向低层子带递减.而且,不同子带之间具有自相似性,因此,如果从小波系数的同一空间方向来研究,可用粗尺度的小波系数来预测细尺度的小波系数.

1.2.2 小波系数编码

为了得到良好的编码效果,3D-OWNER充分利用小波系数的零树偏序性、自相似性和周围系数Markov性,构造良好的上下文.对每个变换子带的编码时,采用三次扫描,即显著性扩展扫描、细化扫描、清除扫描,并构造出相应的上下文,以便后续的算术编码(图1).

1) 周围系数的上下文模型

根据小波系数Markov性,构造的当前扫描系数上下文结点.假如当前结点是 (I, J, K) , I, J, K 分别表示该系数在时间、水平、垂直方向上的坐标.则时间域相邻的结点是 $(I, J, K-1), (I, J, K+1)$.水平的邻居是

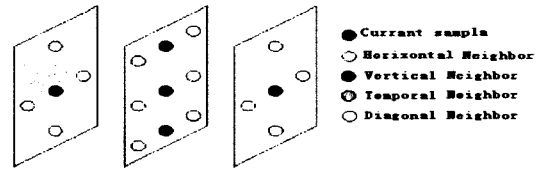


图1 当前编码系数的上下文模型

$(I-1, J, K), (I+1, J, K)$. 垂直的邻居是 $(I, J-1, K), (I, J+1, K)$. 对角线方向的邻居是 $(I, J-1, K-1), (I, J+1, K-1), (I, J-1, K+1), (I, J+1, K+1)$ 共4个对角线邻居.

2) 小波的父子结点的上下文模型的构造

根据小波系数的偏序性和相似性.利用当前结点的显著性与它父结点的显著性之间相关性,该模型构造了父子结点的上下文模型:对最粗尺度的小波系数,它的父系数不存在,设定它父系数的显著性为0,对其它尺度的系数,若结点的父系数是显著的,构造上下文时就用“1”,否则,就用“0”.根据上述,则小波系数树的上下文索引 P 的值在0到1之间(图2).

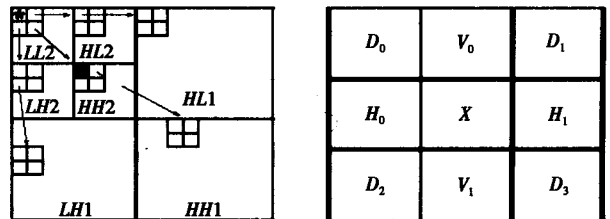


图2 小波系数树中的父子结点上下文模型

1.2.3 算法描述

算法:基于三维小波变换的系数编码;

输入:小波变换系数;

输出:编码后的比特流;

$c(i, j)$:量化后位置 (i, j) 上系数的绝对值;

K :变换后系数帧的总子带数;

T :当前的判决门限;

$S_n(i, j)$:显著性测试参数.若 $c(i, j) \geq T$ 则 $S_n(i, j) = 1$,否则 $c(i, j) \leq T, S_n(i, j) = 0, S_n(i, j) = 1$ 表示 (i, j) 位置上的系数是显著的, $S_n(i, j) = 0$ 表示 (i, j) 位置上的系数是不显著的.显著性判断过程中,门限 T 的值是不断递减的,因此每轮扫描都要对不显著的系数进行显著性判断,对那些已经判断是显著的系数,不再进行判断.对 K 个小波系数的子带,分别进行如下的编码过程:

步骤1 初始化

计算出“最大不为零的比特位”所在的比特平面 n 和当前子带的判决门限 T .

步骤2 扫描编码

对当前位平面中系数,按照先扫描列后扫描行的

顺序,进行一轮三次扫描,即显著性扩展扫描、幅度精细化扫描、清除扫描。每轮扫描中,三次扫描是互斥的,也就是说,每轮三次扫描中,只执行一次扫描。

显著性扩展扫描:若该点 $S_n(i, j) = 0$,且上下文满足下列条件之一时:1)该系数属于 HHH 子带时,至少有一个对角线邻居是显著的;2)该系数属于其它子带时,至少有一个垂直或水平或时间方向上的邻居是显著的。ENCODE(1, Index),然后对该系数的符号位进行编码 ENCODE(signbit, signIndex)。上述两个条件都不满足时 ENCODE(0, Index)。其中 signbit 是编码系数的符号位与符号位上下文模型中的 X 异或后的结果。系数为正时,符号位是 0,否则符号位是 1;signIndex 是该系数符号位的上下文索引。Index 为该系数的上下文索引。

幅度精细化扫描:若 $S_n(i, j) = 1$,进行精细扫描编码。当编码的系数绝对值大于等于门限 T 时,ENCODE(1, magIndex),小于 T 时 ENCODE(0, magIndex)。其中 magIndex 是编码系数当前的上下文索引。若该系数是第一次进行幅度精细时 magIndex = 24,若该系数已经精细扫描且该系数至少有 4 个显著的上下文邻居时, magIndex = 25,否则 magIndex = 26。

清除扫描:对没有被上述两次扫描中扫描的系数进行编码。如果系数大于当前域值,那么就有一个比特位被编码, ENCODE(1, Index),紧接着就是符号位编码 ENCODE(signbit, signIndex),其中 signbit 是编码系数的符号位与一个参数异或后的结果(见符号位上下文模型)。系数为正时,符号位是 0,否则符号位是 1;

如果系数小于当前域值,那么就编码 ENCODE(0, Index)。Index 和 signIndex 同显著性扩展扫描扫描中的 Index 和 signIndex。

步骤 3 量化步长更新 $T = T/2, n = n - 1$,并回到步骤 2。

1.2.4 上下文的构造

小波滤波后,低通子带包含低频信号,具有较多的能量。若某个子带是某个方向上的低通信号,也就意味着在这个方向上,这个子带包含较多的相关性,编码的算法应利用这种相关性,构造的上下文的索引与这种方向上的相关性密切相关,对不同的子带采用不同的上下文模型。变换后,对不同的子带单独编码,这是因为:

1)不同的子带单独编码,意味着这些子带可单独解码,因此分辨率和帧率伸缩很容易实现。

2)各个子带是滤波器对原始视频信号经过滤波后得到的,如上所述,各个子带内部系数在不同方向上的相关性是不同的,因此对各个类型的子带采取不同的上下文模型,以便于算法的高效实现。

下文由两部分组成,一部分是当前系数的零树上下文 P ,第二部分是借鉴 EBCOT、3D-ESCOT 的模型,由当前系数周围 10 个系数来创建上下文,根据 11 个相邻系数和零树父系数 P 中有多少个为显著,那些位置是显著的情况,将上下文分为 10 种。如果一个系数是显著的,那么在构造上下文时就用“1”,否则,就用“0”。上下文的索引同时还与系数所在的子带有关系。显著性的上下文如表 1 所示。

表 1 显著性扩展扫描及清除扫描的上下文

	$H + P$	3	2	1	1	0	0	0	0	0	0	0
	V	X	≥ 1	0	0	3	2	1	0	0	0	0
LLL 和 LLH	D	X	X	X	X	X	x	x	3	2	1	0
	T	X	X	≥ 1	0	0	0	≥ 1	0	0	0	0
	Context lable	0	0	1	2	3	4	5	6	7	8	9
	$H + P$	3	2	2	1	1	1	0	0	0	0	0
LHH	$V + T$	X	≥ 3	≥ 1	≥ 1	0	0	≥ 2	≥ 1	≥ 1	0	0
	D	X	X	≥ 2	X	≥ 1	0	x	≥ 2	x	≥ 1	0
	Context lable	0	1	1	2	3	4	5	6	7	8	9
	$D + P$	≥ 5	≥ 3	≥ 3	≥ 2	≥ 1	≥ 1	0	0	0	0	0
HHH	$H + V + T$	X	≥ 2	< 2	≥ 2	≥ 1	0	> 2	> 1	1	0	0
	Context lable	0	1	2	3	4	5	6	7	8	9	9

注: x 表示取任意值。

对于符号位的编码,采用另一种上下文。计算该上

下文可以分为两步,第一步计算垂直相邻显著性、水平

相邻及时间域相邻系数显著性的和;第二步将这些计算出来的值分为9种情况.在第一步计算中,每个方向的每一个系数的显著性都有3种状态:显著正、显著负、不显著(表2、表3).

精细化扫描是将10个直接相邻的系数显著性和父系数显著性求和,然后根据结果分为3种上下文.

2 实验结果与分析

为了比较3D-ESCOT和3D-OWNER模型中,选取了两种类型QCIF格式的视频序列图像:foreman、carphone,采用32帧为一组对两种模型进行了测试,对算法进行了仿真实验,每一组视频图像分别采取了时

间和空间域的三级Daubechies 9/7小波进行变换.实验结果如(表5、表6):

表2 垂直(或水平)相邻系数的显著性对符号上下文的影响

$V_0(H_0, T_0)$	$V_1(H_1, T_1)$	$V(H, T)$
显著正	显著正	1
显著负	显著正	0
不显著	显著正	1
显著正	显著负	0
显著负	显著负	-1
不显著	显著负	-1
显著正	不显著	1
显著负	不显著	-1
不显著	不显著	0
不显著	不显著	0

表3 符号的上下文

H	V	Context									
		-1	-1	-1	0	0	0	1	1	1	1
H = -1	T	-1	0	1	-1	0	1	-1	0	1	1
	X	0	0	0	0	0	0	0	0	0	0
	Context	10	11	12	13	14	15	16	17	18	19
	V	-1	-1	-1	0	0	0	1	1	1	1
H = 0	T	-1	0	1	-1	0	1	-1	0	1	1
	X	0	0	0	0	0	1	1	1	1	1
	Context	20	21	22	23	24	25	23	22	21	20
	V	-1	-1	-1	0	0	0	1	1	1	1
H = 1	T	-1	0	1	-1	0	1	-1	0	1	1
	X	1	1	1	1	1	1	1	1	1	1
	Context	18	17	16	15	14	13	12	11	10	10
	V	-1	-1	-1	0	0	0	1	1	1	1

表4 幅度精细化的上下文

$\sum H + \sum V + \sum D + \sum T + \sum P$	第一次精细化扫描	上下文
X	是	25
≥ 4	否	26
0	否	27

表5 carphone 测试视频的实验结果 PSNR (dB) 32 帧/s

Sequence	Carphone				
	分量	All	Y	U	V
35 kbps	3D-ESCOT	25.63	25.04	26.64	26.13
	3D-OWNER	26.01	24.93	26.80	26.54
250 kbps	3D-ESCOT	31.13	30.20	31.32	31.87
	3D-OWNER	31.34	30.40	31.76	32.15

表6 foreman 测试视频的实验结果 PSNR (dB) 32 帧/s

Sequence	Foreman				
	分量	All	Y	U	V
15 kbps	3D-ESCOT	25.79	24.04	26.93	27.13
	3D-OWNER	25.96	24.18	26.53	28.04
200 kbps	3D-ESCOT	31.46	30.15	32.20	32.42
	3D-OWNER	31.58	30.23	32.47	32.45

2种模型的曲线效果对比示意图:

从表5.6和曲线的对比分析(图3-6),可以看

出,3D-OWNER比3D-ESCOT好一些,有时较明显,有时差异很小.在变化缓慢的场景中,3D-OWNER模型与3D-ESCOT模型性能差异明显些,分析原因,主要是缓慢变化的场景中,小波系数的父结点对子孙系数有

200 kbps 曲线对比图

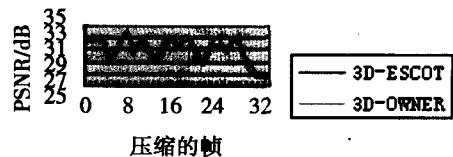


图3 200 kbps 码率下两种模型的曲线对比图 (foreman)

15 kbps 曲线对比图

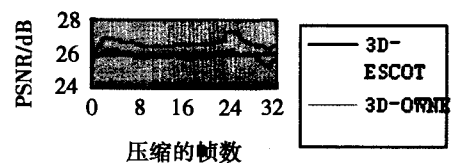


图4 15 kbps 码率下两种模型的曲线对比图 (foreman)

一定的预测作用,而在变化剧烈的场景中,3D-OWNER模型与3D-ESCOT的效果差异不明显,原因是剧烈变

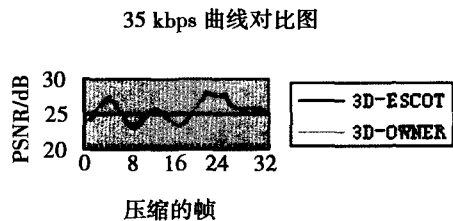


图5 35 kbps 码率下两种模型的
曲线对比图 (carphone)
250 kbps 下曲线对比图

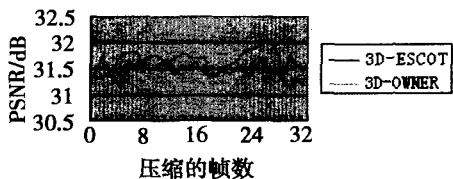


图6 250 kbps 码率下两种模型的
曲线对比图 (carphone)

化的场景中,父结点的系数对子系数的预测影响不如周围结点对待编码的结点的显著性影响大,加上父结点的预测,对模型性能的改善影响很小。

3 结论

笔者结合 EZW、SPIHT 提出的小波系数子带间的偏序性、自相似性及 EBCOT 提出的小波子带内相邻系数的相关性,综合利用这两种性质,修改 EBCOT 模型的上下文,加入待编码结点的父结点信息,构造上下文信息,用来预测当前编码结点,提出了一种三维基于预

测的子带编码模型 3D-OWNER. 利用这种模型获得的 PSNR 高于 3D-ESCOT 模型,同时它同 3D-ESCOT 一样,具有良好的可伸缩性,支持率伸缩、分辨率伸缩和帧率伸缩。

参考文献

- [1] WALLACE G K. The JPEG Still Picture Compression Standard[J]. *Comm ACM*, 1991, 34(4): 30-44.
- [2] ISO/IEC. Information Technology — Coding of Moving Pictures and Associated Audio for Digital Storage Media at u Pto 1.5 Mbps[S]. ISO/IEC 11172-2, 1992.
- [3] ITU Telecommunication Standardization Sector of ITU. Video Coding for Low Bitrate Communication[S]. ITU - T Recommendation H. 263 Version 2, Jan. 1998.
- [4] ISO/IEC, Information Technology — Coding of Audio - Visual Objects — Part2: Visual[S]. ISO/IEC 14496-2, Dec. 1999.
- [5] J M Shapiro. Embedded image coding using zerotree of wavelet coefficients [J]. *IEEE Trans Signal Processing*. 1993,41(12):3445-3462.
- [6] A. Said, W. A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1996, 16(3): 243-250.
- [7] S A Martucci, I Sodagar, T Chiang, et al. A zerotree wavelet video coder [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, 7(1): 109-118.
- [8] David Taubman. High Performance Scalable Image Compression with EBCOT [J]. *IEEE Trans On Image Processing*, 2000, 9(7): 1158-1170.

Context Model Based on 3D Wavelet Coefficients

SUN Xiao-lan

(Department of Computer, Chongqing Vocational Institute of Engineering, Chongqing 400071, China)

Abstract: Based on by typical coding schemes such as EZW, SPIHT, EBCOT, 3D-IEZW, 3D-SPIHT and 3D-ESCOT, a context model(3D-OWNER) of wavelet coefficients is proposed, which makes use of partial order, self-similarity between subbands of wavelet coefficients and the correlation of neighboring wavelet coefficients. Experimental results show that the model can achieve higher PSNR than 3D-ESCOT.

Key words: context model; scalable video coding; wavelet; embedded coding; video compression