

基于三维小波变换的视频图像的压缩算法研究

马 宣, 马海武, 周 杰

(西安建筑科技大学 信控学院, 陕西 西安 710055)

摘 要:文中提出了一种时域加强的基于三维小波变换的视频图像序列的编码。同一般的三维小波变换编码相比,该算法根据人类视觉系统的特性对视频图像对不同频率的数据进行粗细不同的量化,特别是以均方误差(MSE)为准则的阈值量化方法对时间高频域的量化,解决了当图像运动变化较大时所产生的大数据量的问题。该算法无运动估计和补偿环节,降低了算法的复杂度。仿真实验表明该算法的有效性和可行性。

关键词:小波;三维小波变换;视频图像;量化

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2006)10-0089-02

Research of Compressing Algorithm for Video Coding Based on 3D - Wavelet Transformation

MA Xuan, MA Hai-wu, ZHOU Jie

(Department of Communication and Automation, Xi'an University of
Architecture and Technology, Xi'an 710055, China)

Abstract: Presents an algorithm for video coding based on 3D - wavelet transformation with temporal enhancement. Compared to normal 3D - wavelet transformation, the algorithm quantized various data according to human visual system. Especially, the quantizing method using MSE(mean square error) as criterion to quantize temporal high - frequency, solving the matter that the large number of data is created when motion is acuity. Furthermore, the approach will not be concerned with motion estimation and motion compensation. The simulation experiment shows its validity and feasibility.

Key words: wavelet; 3D - wavelet transformation; video image; quantization

0 引 言

随着通信技术的极速发展,视频编码技术也越来越显示出其重要性。视频编码技术在多媒体、高清晰电视(HDTV)、可视电话、视频点播(VOD)、电视会议等视频服务中起着至关重要的作用。传统的视频编码方法是运动补偿加帧内DCT变换编码。该方法已被各种编码标准所采用,如H.261/H.263, MPEG1, MPEG2及近几年才制定的标准MPEG4等。但是运动估计和补偿的算法复杂,且基于分块DCT编码方法在较高压缩比时解码图像的块效应明显,使其在低码率压缩的性能不令人满意。而小波变换在时域、频域均具有良好的局域化特性,可以更好地捕捉到视频图像的非平稳信息,获得更高的压缩比,并且它的多分辨率特性更接近HVS(human visual system,人类视觉系统)的视觉特性。更重要是它能更好地提供可分级编码的能力,有效地防止误码传输问题,减低复杂度。

收稿日期:2005-12-14

作者简介:马 宣(1978-),男(回族),山西长治人,硕士研究生,研究方向为无线电应用技术、视音频信号处理、多媒体数据处理、数字图像处理等;马海武,副教授,主要从事国防科研和通信专业教学工作。

文中提出了一种时域加强的基于三维小波变换的视频图像序列的编码。同一般的三维小波变换编码相比,该算法根据人类视觉系统的特性对视频图像对不同频率的数据进行粗细不同的量化,特别是以均方误差(MSE)为准则的阈值量化方法对时间高频域的量化,很好地解决了当图像运动变化较大时所产生的大数据量的问题。该算法无运动估计和补偿环节,降低了算法的复杂度。

1 基于三维小波变换的视频编码系统

如图1所示,基于三维小波变换的视频编码系统主要是由三维小波、量化和熵编码三部分组成的。如果以 L 帧为一组,设图像的大小为 $M \times N$ 个像素,其中 M, N 分别代表列数和行数,则该组图像共有 $M \times N \times L$ 个像素。首先,对输入的 L 帧图像进行时间上的一维小波变换(t -DWT),直到固定的级数 m ,然后再对每一帧视频图像进行空间域的二维小波变换(s -DWT),然后用改进的阈值量化矩阵对所得的小波系数进行量化,最后在比特分配阶段对量化后的系数进行熵编码,解码过程是其逆过程。对于该编码系统中的三维小波基的选取有以下几点考虑:空间小波变换技术已经很成熟(如文献[1,2]中所采用的

分解方法),它常用的是双正交 B97 滤波器组,如果采用 Lifting 提升算法可以加快算法的计算速度。由于小波变换对信号长度的要求以及传输延迟时间和计算存储量的需要,经折衷考虑以固定数量 L 帧组成的时域分解组进行时域小波变换。

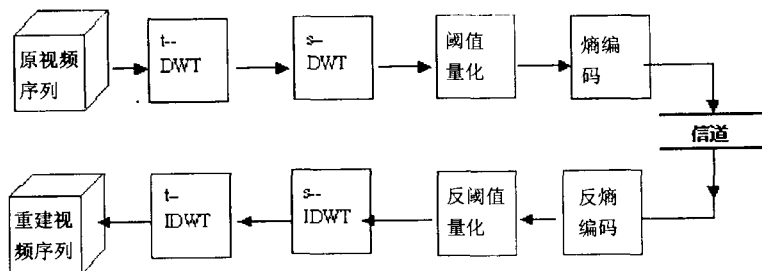


图 1 基于三维小波变换的视频编码系统

2 阈值量化法

视频序列经小波分解得到一个低分辨率下的图像逼近 LLL4 和其他代表不同分辨率、不同方向上的视频序列细节信息的高频三维小波变换子带 (LHHm, LHLm, LLHm, HLHm, HHm, HLLm, 其中 m 表示分解级数,即 $m = 1, 2, 3, 4$)。低频子带 LLLm 代表了视频序列上下文信息,它包含了图像大部分的能量,所以应该重点保护。从时间轴上看,时间轴低频子带的能量要高于高频子带,但是当小波分解中出现同一场景中的物体快速运动时,变换后的时间轴高频子带内的系数会急剧加大,从而影响了算法的整体压缩效果。阈值量化法就是针对此项不足,采用的一种新的量化方法。如图 2 所示。

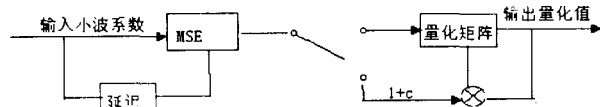


图 2 阈值量化法的方框图

时间轴的高频子带主要包括 HLHm, HHm, HLLm, 它们分别代表时间高频域内水平方向、垂直方向和对角方向上不同频率的子带,包含了视频图像不同方向上的运动信息。文中所用的阈值量化方法就是针对该 4 个系列的子带来进行量化的。该方法中阈值量化矩阵是利用人类的视觉效应来得到,它利用了文献[3]中所提到的取得阈值量化因子的方法:对于小波变换的图像压缩技术,首先对小波系数作处理使之达到视觉无失真,其次影响人眼感知的视觉对象的频率、对比度和光强等因素,通常用检测门限来表示某个对象的可视性。Watson 等人作的关于二维小波变换典型的视觉效应的试验研究,已成为在各多分辨率子带确定检测门限值的基础。阈值量化因子由下面的公式取得:

$$Q(u, \varphi) < \frac{2a}{V(u, \varphi)} \cdot 10^{k(\lg(3u_{\max}/2^{(m+1)}g_{\varphi}u_0))^2}$$

其中 g_{φ} 为可视性空间曲线的最小值转化为整数方向 g_{LH} 和 g_{HL} 的函数, u_0 为曲线在 g_{φ} 方向取最小值时的名义中心频率, a, k 为曲线拟合系数, m 为小波分解的级数。文献

[4] 给出了图像压缩中最佳小波基 9/7 小波分解图像的阈值量化因子矩阵。而文中空间变换小波同样采用双正交 B97 滤波器组,因而可以利用该矩阵作为参考和启发。文中采取的一维时域变换的小波基是 Haar 小波,该小波在高频域的变换可以看作变换数据的差值运算,结合三维小波系数分布的特点,对上述量化矩阵进行适当的调整,可得到三维小波变换的亮度阈值量化矩阵 Q 。

阈值量化法的主要过程见图 2,设阈值为 T (T 的选取须大量的试验结果来折中选择),以小波分解系数的均方误差 (MSE) 作为判断准则,对系数进行量化。若量化矩阵为 Q ,则可按照下式进行量化。

$$Q = \begin{cases} Q & \text{MSE} \leq T \\ Q(1+c) & \text{MSE} > T \end{cases}$$

其中 c 为非零一个任意数。

由上面的公式对小波系数进行量化,并对量化后的系数进行游程编码 (RLC) 和霍夫曼编码。

采用阈值量化方法的三维小波视频编码算法根据三维小波分解系数的特点,参考文献[5]的方法,文中也将小波系数分为几部分来考虑:

(1) 对于 LLLm 子带占据的信息能量区域比较集中,要重点保护,文中采用文献[2]的方法采用 DPCM 编码,因该区域系数的绝对值很大,所以现将系数进行排序,可以用不带符号的 PCM 编码误差信号。

(2) 对于 HLHm, HHm, HLLm 子带,采用上述的阈值量化法进行编码。

(3) 对于其他时间轴低频子带,因其积聚了图像信息的大部分能量,因而不能引入太大的失真,文中采取三维嵌入式零树小波编码 (EZW),可以取得良好的压缩效果。

3 仿真实验

文中利用上述的视频编码方法对 CIF 格式的标准视频序列 Claire, Salesman 的亮度信号进行编解码,取 $L = 16$ 帧为一组进行三维小波变换,分别对时间和空间上进行 4 级的小波分解。由此视频图像序列被分解为 169 个不同频率的子带。解码的平均压缩比和峰值信噪比 (PSNR) 列于表 1。

表 1 仿真结果

	Claire		Salesman	
峰值信噪比 (PSNR) dB	42.77	41.68	38.93	37.51
平均压缩比	41.29	60.02	40.82	59.804

由以上的仿真试验数据可以看出,新的小波变换方法对于平稳图像 Claire,可以保持三维小波变换编码的特点,在高的压缩比之下重建图像仍然能够得到较高的图像质量。对于像 Salesman 这样图像背景比较复杂,并且图像

(下转第 93 页)

推论 1 设 $S = (U, C \cup D)$ 是一个信息决策系统, C 是条件属性集, $D = \{e\}$ 是决策属性集, 则二进制可辨矩阵中, 在 U 中除去所有全为 0 的行所对应的对象对的并集后, 余下的元素组成 $POS_C(D)$ 。

推论 2 在二进制可辨矩阵中, 若存在每一行都不全为 0, 则 $POS_C(D)$ 为 U 的全集。

4 粒度计算的应用

给定信息系统 $S = (U, R)$, 其中属性集 $C = \{a, b, c, d\}$, 如表 1 所示。表 2 为表 1 的二进制可辨矩阵。

表 1 $S = (U, R)$

U	a	b	c	d
1	1	0	1	1
2	1	0	1	0
3	0	0	1	1
4	0	0	1	0
5	1	1	1	1

表 2 表 1 的二进制可辨矩阵

	a	b	c	d		a	b	c	d
1,2	0	0	0	1	2,4	1	0	0	0
1,3	1	0	0	0	2,5	0	1	0	1
1,4	1	0	0	1	3,4	0	0	0	1
1,5	0	1	0	0	3,5	1	1	0	0
2,3	1	0	0	1	4,5	1	1	0	1

由定义 3 和 7 知: $DIS(R) = 2 \times 10/5^2 = 4/5$, $GRD(R) = 5/25 = 1/5$

由定义 8 知: $Sig_{R-\{a\}} = 1 - 5/(5 + 2 \times 2) = 4/9$, $Sig_{R-\{b\}} = 1 - 5/(5 + 2 \times 1) = 2/7$, $Sig_{R-\{c\}} = 1 - 5/(5 + 2 \times 0) = 0$, $Sig_{R-\{d\}} = 1 - 5/(5 + 2 \times 2) = 4/9$

若表 1 是信息决策系统 $S = (U, R)$, $R = C \cup D$, 其中属性集 $C = \{a, b, c\}$, $D = \{d\}$, 则其二进制可辨矩阵如表 3 所示。

由定理 3 知, 表 3 的第一、第五行全为 0, 则该信息决策系统是不一致的; 由定理 4 知, $NEG_{C(D)} = \{1, 2\} \cup \{3,$

$4\} = \{1, 2, 3, 4\}$; 由定理 4 的推论 1 知: $POS_C(D) = U - \{1, 2, 3, 4\} = \{5\}$, $kc(D) = 1/5$, $NEG_{C-a}(D) = \{1, 2\} \cup \{1, 4\} \cup \{2, 3\} \cup \{3, 4\} = \{1, 2, 3, 4\}$, $POS_{C-a}(D) = U - \{1, 2, 3, 4\} = \{5\}$, $NEG_{C-b}(D) = \{1, 2\} \cup \{2, 5\} \cup \{3, 4\} = \{1, 2, 3, 4, 5\}$, $POS_{C-b}(D) = U - \{1, 2, 3, 4, 5\} = \emptyset$, $NEG_{C-c}(D) = \{1, 2\} \cup \{3, 4\} = \{1, 2, 3, 4\}$, $POS_{C-c}(D) = U - \{1, 2, 3, 4\} = \{5\}$ 。

表 3 决策表的二进制可辨矩阵

	a	b	c
1,2	0	0	0
1,4	1	0	0
2,3	1	0	0
2,5	0	1	0
3,4	0	0	0
4,5	1	1	0

5 结束语

文中是利用二进制可辨矩阵讨论知识的粒度的计算及其应用, 得到了二进制可辨矩阵若干定理及推论, 提出计算知识粒度和属性重要度的新方法, 并给出这些方法的应用。利用文中得出的二进制可辨矩阵若干定理及推论, 用于属性及属性值约简, 将另文给出。

参考文献:

- [1] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356.
- [2] Skowron A. Rough Sets in KDD[R]. Special Invited Speaking, WCC2000 in Beijing, 2000. 36-39.
- [3] 苗夺谦, 范世栋. 知识的粒度计算及其应用[J]. 系统工程理论与实践, 2002, 22(1): 48-56.
- [4] 郑书富. 分辨矩阵与知识粒度的应用[J]. 聊城大学学报, 2004, 10: 16-18.
- [5] 支天云, 苗夺谦. 二进制可辨矩阵的变换及高效属性约简算法[J]. 计算机科学, 2002, 29(2): 140-142.

(上接第 90 页)

场景中的物体运动也较大的图像, 该方法仍能在较大的压缩比下得到重建质量良好的视频图像, 这说明文中所提出的方法是有所改进的, 特别是对于运动变化较大的图像。

4 结束语

文中采用了三维小波变换对视频图像序列进行压缩编码, 它较好地改善了当图像场景中的物体进行快速运动时, 使得时间域的小波系数突然变大而使得压缩效率变低的不足。该算法可以快速、高效地压缩图像, 并且无运动补偿和运动估计, 降低了算法的复杂度。今后对该算法在量化矩阵的构造和阈值的选取上还可以进一步地进行研究和完善。

参考文献:

- [1] 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[J]. 电子学报, 2001(10): 1381-1383.
- [2] 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.
- [3] Watson A B, Yang G U, Solomon J A, et al. Visibility of wavelet quantization noise[J]. IEEE Tran Image Process, 1997, 6: 1164-1175.
- [4] Ferguson K L, Allinson N M. Psychophysically derived quantisation model for efficient DWT image coding[J]. IEE Proc - Vis Image signal Process, 2002, 149(1): 51-56.
- [5] 李春华, 白云飞. 一种基于三维小波变换的图像编码算法[J]. 仪器仪表学报, 2002, 23(增刊): 199-200.