

基于小波变换的视频编码算法

蒋雯 勒中鑫

(郑州信息工程学院, 郑州 450002)

摘要 针对小波变换域特有的图象数据结构,对视频编码算法提出了若干改进策略。实验表明这些改进的编码方法有利于图象压缩比的提高和视觉效果的改善。

关键词 小波变换, 视频编码, 多分辨率运动估计, 自适应标量量化, 自适应矢量量化

1 引言

目前,在各种视频压缩编码中,帧内图象和运动补偿后的误差信号通常被分成图象子块进行变换编码。如 DCT 编码,当比特率较低时,常会出现人眼十分敏感的“方块效应”。帧间预测编码中所采用的运动补偿技术,往往存在运动估计(ME)问题,即各种运动估计算法普遍存在搜索运算量和检测精度之间的矛盾。针对这些问题,本文企图利用小波变换对图象的全局分解性能及图象经小波变换后所形成的特殊的锥形数据结构,在小波域进行量化编码,量化误差仅产生在某些频率分量上,并被淹没在重建图象的某些内容中,避免了过多地影响重建图象的主观质量。利用变换系数和原始图象在空间上的对应关系,引入的小波域多精度运动估计(MRME)算法,不仅可有效的减少搜索运算量,同时能提高运动矢量的编码效率。

2 小波变换图象的特征

2.1 小波变换域的图象数据结构

一幅图象经过三级小波分解后,其编号如图 1 所示。它表示一幅图象被分解为 3 层,共 10 个小波变换子图象。各层之间相互独立,分别对应于不同的频段和分辨率,每一层包括水平、垂直和对角线三个频道上的子图象。最低分辨率层还包含一个低分辨率版本的原始子图象,或称为最高层低频子图。因

此,这些子图便构成了一种锥形数据结构。

经理论分析,这种锥形数据结构具有以下特点:

(1) 与人眼视觉特性相匹配。因此,可根据 HVS 特性,对小波域各个子图象的量化误差进行合理的分配,在同样的平均比特数下,获得主观质量更好的重建图象,或者在相同的主观评价下,获取更高的压缩比。

(2) 所形成的多分辨率描述结构,没有扩展图象数据量,其传输可采用逐次浮现式。

(3) 基本上实现了图象中平稳成分和非平稳成分的分离。低频子图为平稳部分,高频子图为非平稳部分。因而图象边缘,纹理等非平稳部分,可有效地进行编码。

(4) 各级子图象与原始图象在空间域和频率域存在对应关系,同一物体在不同精度下,具有高度一致的运动特性。

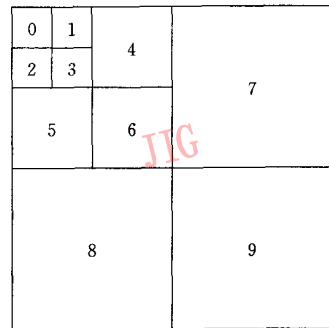


图 1 图象三级小波分解的频率划分
Fig. 1 Image frequency distribution after three-level wavelet transformation

2.2 小波变换图象的统计特性

为了对小波变换域图象数据结构有一个具体的了解,以便得到一些有益的启示,为后续量化编码算法的选择提供一定的理论依据。对“Miss American”标准图象的三级小波子图进行了均值和方差的统计。结果如表 1 所示。

表 1 小波变换图象“Miss American”的统计特性

Table 1 Statistical performance of wavelet transformation for “Miss American” image

分解层数	方向(子图标号)	均值	方差
第一层	水平(7)	0.029 4	2.676 5
	垂直(8)	0.013 1	1.305 0
	对角(9)	0.001 7	1.316 5
第二层	水平(4)	-0.019 2	3.414 3
	垂直(5)	0.055 7	4.437 0
	对角(6)	-0.000 6	0.914 8
第三层	水平(1)	-0.027 1	7.695 6
	垂直(2)	0.012 6	7.579 3
	对角(3)	-0.009 5	1.861 6
最高层低频子图(0)		64.000 6	774.420 7

由表 1 可以看出:

(1) 最高层低频子图的图象尺寸虽然仅为原图的 $1/64$,但却包含了原始图象的绝大部分信息,因此,小波变换具有潜在的数据压缩能力。

(2) 除了最高层低频子图外,其余小波变换子图象的均值非常接近于零。

(3) 各高频子图象的均值接近于零,其方差却有很大差别。但基本符合尺度越大、方差越大,水平与垂直方向的高频子图方差大于对角线方向高频子图方差的特点。子图象的方差越大则说明该子图象所包含原图象的信息越多,反之所包含的信息越少。因此,可根据方差的大小自适应的分配编码比特数;

正是由于小波变换域图象可形成这种特殊的锥形数据结构。所以,本文将双正交小波基的变换引入视频编码,以便对某些算法进行改进。

3 小波域中的视频编码技术

利用小波变换从帧内和帧间进行视频编码,主要目的是研究去除帧内图象和预测误差信号的空间相关性,去除相邻图象间的时间相关性。

3.1 帧内编码

在各种视频编码中,帧内编码普遍采用基于 8×8 分块的快速 DCT 变换编码。对于相关性接近 0.9 的图象,变换是准最佳的,加之采用与人眼视觉系统相匹配的量化表进行量化,所以是一种比较好的实用编码方法。但是,当比特率较低时,这种分块技术在画面上会出现“方块效应”,常常影响重建图象的视觉效果。

根据小波域的图象数据结构,小波变换编码的优点:首先小波变换是对图象进行全局分解,量化失真随机分布于整幅图象之中,人眼不易察觉这种失真,更不会出现方块效应;其次,它可将图象逐层分解成不同分辨率下的小波变换子图象,各子图象所含的频率信息不同,低频信息集中于尺度较大的子图象,高频信息集中于尺度较小的子图象。这种信息的高度分离,比 DCT 更显著,因此根据人眼对不同信息的敏感程度,对各子图象进行的自适应量化编码会更有效,在重建图象的主观视觉效果相同的情况下,会大大提高压缩比。另外,由于最高层低频子图象和各级高频子图象,具有不同的数据特点,可分别采用不同的方法进行有效的编码。

3.1.1 最高层低频子图的编码

最高层低频子图的系数代表原始图象中近似直流分量的成分,因此,虽然其抽取间隔映射到原始图象象素间的距离要变成 8 倍,仍需要进行去相关性的操作。但最高层低频子图所含信息量非常大,量化误差对重建图象会产生严重的影响。所以,采用无失真的方法对它进行编码,保留它的全部信息,去除残余的相关性。以“Miss American”标准序列图为例,10 幅三级小波变换的最高层低频子图的二维无失真 DPCM 编码,平均压缩比为 2.05 倍。而 JPEG 无失真编码,平均压缩比可达到 2.23 倍。但 DPCM 编码的运算量比 JPEG 的要小得多。

3.1.2 自适应标量量化

如果企图实现对各级高频子图象的最佳标量量化,应先估计各个子图象的概率分布密度函数,然后按照最小均方量化失真准则,设计最佳非均匀量化器。显然这一运算过程复杂,实现起来也比较困难。如果对所有子图象进行统一的均匀量化,但由于各个子图象的概率分布密度函数不相同,对应的灰度范围也不相同,因此,只有当量化电平数 N 值取得很大时,才可保证各子图象都获得较好的量化效果。显然,这样做将会急剧增加熵编码的复杂度和编码

比特数。

针对这种数据的特点,设计了一种自适应标量量化器。它以不同的量化阶距对各个子图象分别进行均匀量化。令各个子图象的量化阶距

$$z = \frac{S}{\log \sigma^2}$$

其中 σ 为子图象的方差。这样做实现了方差大的子图象即包含能量多的子图象采用较小的量化阶距,进行细量化;反之采用较大的量化阶距粗量化。最后对量化后的数据进行熵编码。其中子图象的方差小于门限 D_0 ,就不进行编码。即去除那些方差非常小、包含信息很少的子图象数据。以“Miss American”标准序列图作为实验样图,对自适应标量量化方法进行测试,结果见表 2。由表看出,只要选取的门限值 D_2 合适,重建图象的质量不仅不会受影响,还可较大地改善压缩性能。同时,由于量化时自适应分配量化误差,当压缩比高达 35 倍以上时,重建图象仍然具有较好的视觉效果。

表 2 自适应标量量化性能

Table 2 The property of Adaptive scalar quantization

编码参数	$D_0=0$ $s=1.3$	$D_0=1.5$ $s=1.3$	$D_0=2.0$ $s=1.3$	$D_0=1.5$ $s=3.0$
压缩比	13.26	15.48	15.68	35.63
PSNR(dB)	40.82	40.23	39.11	38.00

3.1.3 自适应矢量量化

上述自适应标量量化,尽管取得的压缩比和重建图象的视觉效果都比较好,但它没有充分利用和发挥小波域多级分解后所形成的与人眼视觉特性相匹配的多频带锥形数据结构的特点。相对于标量量化,应用矢量量化更有可能减少均方量化失真,提高重建图象的质量。

矢量量化是对矢量的集合量化,它只需传输或存储矢量地址。小波变换域的自适应矢量量化,就是对各个小波变换子图象分别进行矢量量化,并在对各子图象进行码书训练时,根据人眼视觉特性,自适应地选定失真准则、码书尺寸和码本矢量的大小。与各子图象使用统一码书相比,增加了码书训练的复杂度、需要训练多个码书,但能有效地减少码书尺寸,减少编码误差,加快码本搜索速度,提高编码效率。具体实现时,由于不同尺度下的数据量化误差对重建图象的质量会产生不同的影响。尺度 2^j 下的数据在重建图象时,要经过 j 次滤波。因此,同样大小

的误差,如发生在较大尺度下,则在重建图象中,其误差值和所影响的空间范围都会变得更大,人眼对大尺度的图象数据信息更为敏感。另外,小波变换子图象的方差越大,所包含原图象的信息就越多。所以对于尺度较大、方差较大的子图象,应选取相对较大的码书尺寸,并对它进行误差补偿。反之,选取较小的码书尺寸。这样一来,就可在保证重建图象主客观质量的前提下,尽可能减少平均码书的尺寸,缩短码书训练时间,提高编码效率。按照参数表 3 中数据对“Miss American”标准图进行实验测试,结果见表 4。其中矢量编码所用码书采用 SOFM 自组织特征映射神经网络方法训练得到。

表 3 自适应矢量量化编码所用参数表

Table 3 Parameter list for adaptive vector quantization

分级子图标号	码书尺寸	图象矢量块大小
1	16	2×2
2	16	2×2
3	16	2×2
4	64	4×4
5	64	4×4
6	32	4×4
7	64	8×8
8	32	8×8
9	32	8×8

表 4 自适应矢量量化性能

Table 4 The property of Adaptive vector quantization

编码算法	JPEG 压缩算法	自适应矢量量化
压缩比	24.21	34.14
PSNR(dB)	40.33	40.38

实验结果表明,选取适当的码书尺寸和码本矢量的大小,基于小波变换域的自适应矢量量化方法与 JPEG 算法相比,可获得较好的图象压缩效果。从图 2 可以看出,重建图象的主观视觉效果也非常好。因此可以说,基于小波变换域的自适应矢量量化方法是一种有效的帧内编码方法。在高压比下,重建图象的主客观评价令人满意,其运算量与使用统一码书的矢量量化方法相比大为减少。

3.2 帧间编码

序列图象压缩编码的基本问题是帧间编码。经过人们的研究,运动补偿已作为一种有效的时域预



原始图象



重建图象

图2 重建图象与原始图象的比较

Fig. 2 The comparison between original and reconstructed image

测手段用于许多视频信号的编码中。它的关键是目标运动参数的估计,即运动估计(ME)问题。目前一些运动估计算法普遍存在搜索运算量和检测精度之间的矛盾。在各种视频压缩编码中,运动补偿后的误差信号通常被分成许多图象子块进行变换编码,如DCT变换编码,在低比特率下,常会出现“方块效应”。小波域的运动补偿,正是针对这考虑而提出的。

如前所述,小波变换可将图象分解成不同分辨率下的小波变换子图象,且原始图象物体在不同层次对同一物体小波变换系数具有相似性。因此,将多分辨率运动估计(MRME)算法用于小波变换域的序列图象,可以获得运动矢量估计。这种算法称为小波变换运动估计(WTME)算法。利用它可有效的减少运动矢量的搜索复杂度。

3.2.1 小波变换运动估计算法

小波变换运动估计算法是一种变尺度的块匹配算法,其示意图见图3。由于小波变换由低层向高层过渡时,子图象的大小变成原来的 $(2 \times 2)^{-1}$ 。因此,从高层到低层,块匹配中的分块尺寸应变为原来的 (2×2) 倍。一帧图象经过 M 级小波变换后,可得到 $3M+1$ 个小波变换子图象,所有 $3M+1$ 个子图象表示同一物体在不同分辨率下的方向特性,因此,具有高度一致的运动特性。假定最高层子图象的分块尺寸是 $(P \times P)$,则第 i 层的分块尺寸选择为 $P2^{M-i} \times P2^{M-i}$,这样就保证了估计的是同一物体的运动,而和所在分辨率层无关。

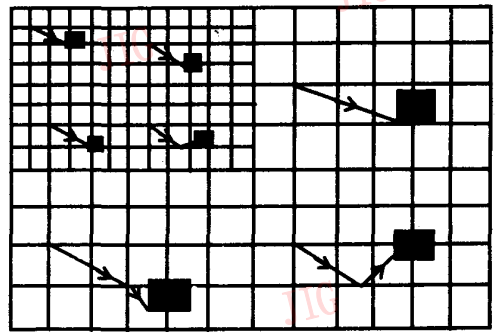


图3 块匹配算法示意图

Fig. 3 Diagram of algorithm based on block matching

令图象序列为 $\{I_{0,1}, I_{0,2}, \dots, I_{0,N}\}$ 经小波变换后的子图象序列为:

$$\{I_{M,n}; W_{k,n}^j\}, n = 1, 2, \dots, N;$$

$$j = 1, 2, 3; k = 1, 2, \dots, M$$

其中 n 表示图象序号, M 为小波变换的最大分解层数, k 表示第几层, $j=1, 2, 3$ 分别表示一层中的水平,垂直和对角线方向的子图象。对于三级小波变换,若对某层子图象的分块尺寸大小为 B 的位移矢量估计,则在搜索窗内进行全局搜索,使绝对值误差 MAD 为最小。以低频子图为例,有:

$$\min_{(r+v) \in SR} MAD =$$

$$\sum_{r \in B} |I_{M,k}(r) - I_{M,k-1}(r+v)|, k \geq 2$$

其中, $r=(x, y)$ 为空间坐标, $v=(v_x, v_y)$ 为位移矢量, SR 为搜索窗,当 $I_{M,k}(r)$ 代表最高层低频子图时, $I_{M,k-1}(r)$ 就是下一层低频子图。这样WTME算法步骤为:

(1) 首先对最高层四个子图象的位移矢量进行估计,即设定分块尺寸的大小,在给定的搜索窗内进

行全局搜索;

(2) 在得到最高层各方向子图的位移矢量后,将其尺度扩大(2×2)倍,作为下一层相应方向子图的位移矢量初始估计值,并在以这个初始估计值为中心的搜索窗内进行同样的搜索过程,可得到该层次各方向子图的位移矢量,该过程可重复进行到第一层各子图象。

在各分辨率层上,分块尺寸随层次的变小而递增,但搜索窗大小保持不变。除了最高层子图以外,其它各层高频子图仅需对修正矢量进行估值和编码。因此,可大大减少搜索运算量,提高编码效率。

具体实现时,由于人眼视觉系统对低频成分的误差比对高频成分的误差更为敏感,本文给出以下调整策略:

(1) 为了提高高层子图象运动估值的精度,可采用亚像素精度法修正位移矢量,对不在网格上的样本点,采用双线性内插来获得;

(2) 由于最高层子图象所含信息较多,可对其进行预测误差补偿,对预测误差信号可进行 DCT 变换编码,格型矢量量化,最简单可采用标量量化,然后进行熵编码。这样可用牺牲少量比特数,来改善重建图象质量;

(3) 最低层高频子图包含了图象的高频成分和绝大多数高频随机噪声,其帧间相关性非常小,帧间预测效果已不显著,而且其信号能量相当小,绝大部分系数为零,可不考虑对其进行帧间预测。因为,对这些子图象可实现有效编码。相反,如果对他们进行运动补偿,不仅会大幅度增加运算复杂度,而且将产生大量运动矢量信息。

3.2.2 实验结果

为了测试算法的运动估计效果,以连续的 10 幅“Miss American”标准序列图象作为实验样图,并对每幅图象进行三级小波变换。在运动估计过程中,最高层子图象的子块大小取 2×2 ,其它层可类推,搜索窗大小均为 5×5 。最高层低频子图进行二

级双线性内插,即给出 $1/4$ 样本精度时的位移矢量,最高层三个高频子图给出 $1/2$ 样本精度时的位移矢量,其它子图未进行位移矢量修正,给出一个样本精度时的位移矢量。

按照上述参数,每一帧图象都以前一帧图象作为参考图象进行运动估计(第一帧除外),得出 9 帧图象的运动矢量及运动估计值,进行统计平均。一幅图象运动矢量的平均编码比特数为 1776bits 时(原始图象的比特数为 101 376bits),运动补偿图象的平均 PSNR 为 33.62dB。

实验结果表明,小波变换运动估计算法利用一个原始图象物体在不同层次对同一物体小波变换系数的相似性,可有效的减少运动矢量的搜索区域,可获得较好的运动估计效果,是一种有效的运动估计方法。

4 结 论

本文通过对小波变换域图象数据结构的研究,成功地将自适应矢量量化和多分辨率运动估值算法用于视频编码,取得了较好的实验结果。研究表明:小波变换不仅适合于描述非平稳的视频信号,还是一种有希望的视频编码技术。

参 考 文 献

- 1 Mallat S. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intel., vol. PAMI-11, July 1989. 673~692.
- 2 Zhang Y Q, Sohail Z. Motion-compensated wavelet transform coding for color video compression. IEEE Trans CAS for Video Technology, 1992, CASVT-2(3).
- 3 John D Villasenor, Benjamin Belzer, Judy Liao. Wavelet Filter Evaluation for Image Compression. IEEE Trans IP, 1995, 4(8): 1053~1060.
- 4 Ramamurthi. A. Gersho. Classifical Vector quantization of image. IEEE Trans, on COM, 1986, COM-34, 1105~1115.



蒋雯,1994年毕业于郑州解放军信息工程学院信息科学系,获理工科学士学位;同年在该院就读图像识别和智能化信息处理专业硕士研究生。目前从事小波理论在视频图象压缩编码方面的研究及其应用。

(下转 739 页)



王小睿,1995年毕业于东南大学计算机科学与工程系,获计算机应用学士学位,现为
中国地质大学信息工程学院硕士研究生,主要研究方向为:遥感图象的自动处理与分析、
遥感数据的模式识别。已发表论文4篇。

The Automatic Registration Method of Remote Sensing Multi-Image

Wang Xiaorui, Wu Xincai

(China University of Geosciences, Wuhan 430074)

Abstract Automatic registration is an important technique of remote sensing multi-image processing. To overcome the drawbacks of existing approaches in application, this paper presents a method which is simple and effective, this method can achieve an accuracy up to 0.1 pixels.

Keywords Automatic registration, Multi-image, Least Squares Matching (LSM)

(上接 724 页)

Video Coding Algorithm Based on Wavelet Transform

Jiang Wen, Le Zhong Xin

(Information Engineering Institute, ZhengZhou 450002)

Abstracts Aimed at the particular image data structure of wavelet transform, video coding algorithm is further studied and several improved strategies are presented in this paper. Experiments show that the proposed coding methods are advantageous to raise compression ratio and improve visual effect.

Keywords Wavelet transform, Video coding, Multiresolution motion estimation, Adaptive scalar quantization, Adaptive vector quantization