利用分形变换的帧间图象运动补偿预测编码*

谢卫国　　沈兰荪
（北京工业大学信号与信息处理研究室，100022）

摘　要　首先把分形编码中的变换方法引入传统的运动补偿预测算法中；并进一步提出一种图象序列正交树自适应分块分形压缩编码新方法，用于三维图象块的编码，得到了令人满意的结果.

关键词　帧间图象编码，分形变换，运动补偿，预测编码
分类号　TN919.31

近年来，以分形理论为基础的图象压缩方法，由于其潜在的高压缩性能，引起人们极大兴趣. M.F.Barnsley首先把分形技术用于图象压缩中[1]，随后A.E.Jacguin[2]，D.M.Monro[3]等又发展了一些灰度图象的自动压缩技术. 以迭代函数系统（IFC）为基础的块图象压缩算法[2,3]对静止图象的压缩，已取得较好的结果.

本文首先把分形编码中的变换引入传统的运动补偿预测方法中，提出了采用分形变换的帧间图象运动补偿预测编码方法，克服了传统方法的一些局限性.

本文还进一步把几帧图象作为一个整体进行编码，提出一种图象序列正交树自适应分块分形压缩编码方法，即将分形编码方法与运动补偿预测编码有机结合，用于三维图象块的编码，得到了令人满意的结果.

1　帧间图象运动补偿预测编码

对于前向运动补偿预测编码，编码图象为$f(x, y)$，参考图象为$g(x, y)$，把编码图象$f(x, y)$分割成一系列大小为$B\times B$的图象块. 并假设一个图象块中各物体具有相同的运动. 这样对一个图象块F_k，其顶点坐标为(x_k, y_k)，在参考图象$g(x, y)$中寻找一个$B\times B$的图象块G_k，它的顶点坐标为$(x_k+\Delta x_k, y_k+\Delta y_k)$，使得$G_k$与$F_k$的均方误差$e_k$最小. G_k可以在图象$g(x, y)$上顶点为$(x_k+\Delta x, y_k+\Delta y)$，$(|\Delta x|, |\Delta y|\leqslant \Delta\max)$的所有$B\times B$的图象块中进行遍历搜索，但这种搜索方法效率太低. 现有许多快速搜索算法,本文采用的是文献[4]中介绍的方法，简述如下:

首先，对任意整数$m>0$，定义:

$M(m)=\{(0, 0), (m, 0), (0, m), (-m, 0), (0, -m)\}$

收稿日期: 1996—03—11
* 北京市自然科学基金资助项目

$N(m) = \{(i, j);\ -m \leqslant i, j \leqslant m\}$.

对 Δx_k、Δy_k 的对数递归搜索过程分以下几个步骤:

步骤 1(初始化):

设定 $\Delta\max$ ($|\Delta x_k|$, $|\Delta y_k| \leqslant \Delta\max$)

$n' = \lfloor \log_2 \Delta\max \rfloor$ (对任何大于 0 的实数 x, $\lfloor x \rfloor$ 为整数, 且 $x-1 < \lfloor x \rfloor \leqslant x$)

$n = \max\{2, 2^{n'-1}\}$

并置 $q = p = 0$.

步骤 2: 在 $M(n)$ 中寻找 (i, j), 使 $\Delta x = i + q$, $\Delta y = j + p$ 处的图象块与 F_k 的均方误差最小, 且满足 $|\Delta x|$, $|\Delta y| < \Delta\max$. 如果 $i=j=0$, 转到步骤 3, 否则转到步骤 4.

步骤 3: $n = n / 2$, 如果 $n > 1$, 转到步骤 2, 否则转到步骤 5.

步骤 4: $q = q + i$, $p = j + p$, 转到步骤 2.

步骤 5: 在 $N(1)$ 中寻找 (i, j) 使 $\Delta x = i + q$, $\Delta y = j + p$ 处的图象块与 F_k 的均方误差 e 最小, 且满足 $|\Delta x|$, $|\Delta y| \leqslant \Delta\max$, 则

$\Delta x_k = i + q$, $\Delta y_k = j + p$, $e_k = e$.

运动补偿预测编码是一种去除帧间图象中存在的时间冗余信息的有效方法. 它在图象中的物体只有平移运动时, 编码效果较好; 而当物体有旋转、变形等非平移运动或因光照等原因而有灰度变化时, 编码效果就比较差, 且“块效应”比较严重. 为了克服这些弱点, 本文引入静止图象分形编码中的变换编码, 即在运动补偿预测编码失败的地方增加分形变换编码, 也即编码的均方误差大于误差门限时, 使用分形变换编码. 通过调整误差门限, 可以在编码的质量与压缩比这两个指标间做出权衡.

2 分形变换用于帧间运动补偿预测编码

对于编码图象 $f(x, y)$ 中给定的一个 $B \times B$ 的图象块 F_k, 在参考图象 $g(x, y)$ 中寻找一个 $B \times B$ 的图块 G_k, 对 G_k 中的每个象素 $g(x, y)$, 作双线性变换 $g'(x, y) = a_1^k x + a_2^k y + a_3^k g(x, y) + a_4^k$, 得到图象块 G_k', 使 G_k' 与 F_k 的均方误差最小. 各个系数 a_1^k、a_2^k、a_3^k、a_4^k 的求法参见文献[3]. 与静止图象分形编码不同的是, 由于这里已不存在解码时的收敛问题, 也就是说没有了变换的收缩性问题, 因而寻找的图象块 G_k 的尺寸不必大于 B, a_1^k、a_2^k、a_3^k、a_4^k 的取值也没有限制.

整个编码过程如下:

(1) 把图象分割成大小为 $B \times B$ 的图象块.

(2) 对每个 $B \times B$ 的图象块 F_k, 进行运动补偿预测编码得到 Δx_k、Δy_k、e_k.

如果 e_k 小于误差门限, 则 Δx_k、Δy_k 为图象块的编码; 如果 e_k 大于误差门限, 则进行分形变换编码, 此时 G_k 的搜索范围缩小到顶点为 $(x_k + \Delta x, y_k + \Delta y)$ ($\Delta x = \Delta x_k$, $\Delta x_k \pm 1$; $\Delta y = \Delta y_k$, $\Delta y_k \pm 1$) 等 9 个图象块.

解码时不必象静止图象分形编码的解码过程那样反复进行迭代, 而只要在参考图象 $g(x, y)$ 上进行预测及变换, 即可得到编码图象 $f(x, y)$ 的解码图象 $f'(x, y)$.

我们在 PC 机上用本文提出的分形变换用于帧间运动补偿预测编码方法 (FTMC) 对 256 × 256, 8bits/pixel 的序列图象“Cronkite”进行了编码, 得到了令人满意的主观与客观结

果，下表1给出了对两幅帧间图象的编码结果，并与传统的运动补偿预测编码(MC)进行了比较．从表中可以看出，FTMC编码方法的性能要比传统的运动补偿预测编码(MC)方法有较大提高，特别是在编码图象与参考图象的峰值信噪比比较小时．从主观效果看，传统方法中存在的“块效应”得到了相当程度的抑制．

表1 MC与FTMC编码结果比较

参考图象与编码	MC				FTMC			
图象PSNR(dB)	B	Δmax	压缩比	PSNR(dB)	B	Δmax	压缩比	PSNR(dB)
25.43	12	7	157	33.48	16	7	147	34.40
23.69	8	7	59	33.35	16	7	78	34.46

3 图象序列正交树自适应分块分形压缩编码

图象序列中，相连的几帧往往变化不大，只是存在一定程度的运动和变形．相连的几帧图象，它们的某些部分之间有可能存在共同的特性——自相似性．显然，如果几帧图象的某些部分可用一个共同的分形变换来表示，那么就可以得到很高的压缩比．本文为此提出了一种图象序列正交树自适应分块分形压缩编码算法．

对于待编码的图象序列$f(x, y, t)$中顶点坐标为(x_k, y_k, t_k)，大小为$M\times M\times N$(M为图象块的平面尺寸，N为时间轴尺寸，即图象的帧数)的图象块R_k，可以寻找一个图象块D_k^*，它的顶点在(x_k', y_k', t_k')，大小为$2M\times 2M\times N$，首先把它缩小为$M\times M\times N$的图象块D_k'，然后对图象块中的每个象素的灰度值$f'(x, y, t)$做变换$g(x, y, t)=a_1^k x+a_2^k y+a_3^k f'(x, y, t)+a_4^k$，得到图象块$D_k^2$，它即是$R_k$的分形近似，而$a_1^k$、$a_2^k$、$a_3^k$、$a_4^k$及$D_k$的位置信息即为$R_k$的分形编码．这些参数选取的标准是使$D_k^2$与$R_k$之间的均方误差最小．具体的求取方法可参考文献[3]，只是这时公式中的积分应由二重积分$\iint_{x,y}\cdot$ 扩展到三重积分$\iiint_{x,y,t}\cdot$，亦即离散求和由$\sum\limits_x\sum\limits_y\cdot$扩展到$\sum\limits_x\sum\limits_y\sum\limits_t\cdot$．考虑到解码时的收敛问题，对$a_3^k$的取值范围应加以限制，实验表明，取$|a_3^k|<1.35$即可保证解码图象收敛．

在分形编码的过程中也引入预测，利用时间冗余信息来达到数据压缩的目的．即当$t_k>0$时，图象块D_k在顶点为$(x, y, t_k, -1)$的图象块中确定．这一改进提高了系统性能，并使解码时收敛速度加快．

在序列的各帧图象之间具有很大的相关性，即存在时间冗余信息，运动补偿预测编码是去除这种冗余信息的有效方法，它直接从参考图象预测出当前帧图象．编码过程是这样的，当$t_k>0$时在待编码的图象序列中，寻找一个图象块D_k，它的顶点在$(x_k+\Delta x_k, y_k+\Delta y_k, t_k-1)$，大小为$M\times M\times N$，使它与$R_k$的均方误差最小，那么$\Delta x_k$，$\Delta y_k$就是$R_k$的编码．现有许多快速运动补偿预测编码算法，这里采用文献[4]中的快速运动补偿对数递归搜索算法．

对于一个$M\times M\times N$的图象块，存在两个分割方向．它可以在空间域被分割为4个

$M/2 \times M/2 \times N$的图象块，亦可在时间域被分割成2个$M \times M \times N/2$的图象块. 一个图象块是否被分割由它编码的匀方误差决定，而采用哪个方向的分割由每个方向分割后各子块编码的平均均方误差决定. 下面我们介绍一个$M \times M \times N$的图象块R_k分割的具体过程.

1) 把R_k分割为4个$M/2 \times M/2 \times N$的图象子块R_{kj}，$j=1, 2, 3, 4$，对每个子块R_{kj}分别进行分形编码和运动补偿预测编码，求出编码系数及编码误差，根据编码的误差决定选用哪一种编码，然后对4个子块的编码误差求平均，得出空间域分割的平均误差e_1. 为提高压缩比，对图象块先进行运动补偿预测编码，如误差小于门限值，则不再进行分形编码，而直接选用运动补偿预测编码.

2) 把R_k分割为2个$M \times M \times N/2$的图象子块R_{kj}，$j=1, 2$，对每个子块R_{kj}进行编码，求出编码系数及编码误差，然后求平均得出时间域分割的平均误差e_2.

3) 如果$e_1 < e_2$则采用空间域分割，否则采用时间域分割，然后对分割后的每个子块R_{kj}(采用空间域分割时$1 \leqslant j \leqslant 4$，而采用时间域分割时$1 \leqslant j \leqslant 2$)进行处理. 如果它的编码误差已经低于误差门限，或它的尺寸已达到最小，即$M_0 \times M_0 \times N_0$，就不再进行分割，记录下编码系数；否则对该子块进一步分割.

整个分块过程用正交树表示，正交树中的每个节点对应于一个图象块，其中叶节点表示不再分割的图象块，而中间节点表示需要继续分割的图象块. 每个节点只需用两位b_0，b_1表示，具体含义如下：

$$b_0 = \begin{cases} 0 & \text{本节点为叶节点} \\ 1 & \text{本节点为中间节点} \end{cases}$$

当$b_0 = 0$时，

$$b_1 = \begin{cases} 0 & \text{采用分形编码} \\ 1 & \text{采用预测编码} \end{cases}$$

当$b_0 = 1$时，

$$b_1 = \begin{cases} 0 & \text{时间域分割} \\ 1 & \text{空间域分割} \end{cases}$$

根据人眼视觉对大灰度变化区域的编码误差不如小变化区域的编码误差敏感这一特性，误差门限的取值与待编码图象块的灰度变化大小有关，灰度变化大的图象块所用的误差门限比较大，反之，小灰度变化的图象块所用的误差门限就比较小，这样可保证解码图象的主观质量.

解码时，从任意的图象序列$f_0(x, y, t)$开始，把分形变换及运动补偿预测重复作用于其上，得到一系列图象序列$f_0(x, y, t)$，$f_1(x, y, t)$，…，它们收敛于图象序列$f'(x, y, t)$，这可以由以下两点来保证：

(1) 迭代函数系统(IFS)理论；

(2) 图象序列中第二帧以后的图象才有可能进行运动补偿预测编码.

图象序列$f'(x, y, t)$就是原序列$f(x, y, t)$的解码结果.

我们在PC机上对288×352, 8bits/pixei的图象序列“Miss America”(40-47帧)进行编码. 由于受内存限制，把序列分成4个144×176的子序列分别进行编码，为了消除由此可的块间痕迹，对每一子序列编码时空间域尺寸都略大于144×176，即4个子序列有部分重

叠. 表 2 列出了模拟结果.

表2 图象序列正交树自适应分块分形压缩编码

帧数	压缩比	平均PSNR/dB
4	80	36.65
4	54	38.28
8	98	35.64
8	65	37.28

4 结束语

本文提出的图象序列分形编码方案, 综合了分形编码与运动补偿预测编码的特点, 去除了图象序列的空间及时间冗余信息, 实现了高度压缩的目的. 由于编码过程采用正交树自适应分块方法, 因而编码参数 (如压缩比、解码图象信噪比等)易于控制, 且存储数据的冗余开支小.

参 考 文 献

1 Barnsley M F, Sloan A D, A Better Way to Compress Image, Byte. 1988. 215 ~ 223

2 Jacquin A E. Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations. IEEE Trans on Image Proccssing, 1992, 1 (1): 18 ~ 30

3 Monro D M, Dudbridge F. Fractal Approximation of Image Blocks. IEEE ICASSP. 485 ~ 488

4 Jain J R, Jain A K. Displacement Measurement and Its Application in Interframe Image Coding, IEEE Trans on Comm, VOl. com-29. 1799-1981, 12

Fractal Transformed Motion Compensation Predictive Coding of Interframe Image

Xie Weiguo Shen Lansun

(Signal and Information Processing Laboratory, Beijing Polytechnic University, 100022)

Abstract A new image sequence compressing coding based on fractal theory is introduced, of which the sequence is adaptively partitioned using quad-tree data structure, which overcomes the shortcomings resulted from the conventional motion compensation predictive technique. The fractal coding introduced is to remove spatial redundancy; the motion compensation predictive coding exploited is to remove temporal redundancy.

Keywords interframe image coding, fractal transform, motion compensation, prediction coding