

基于二代小波变换的图像 目标编码方法*

张郑擎 莫玉龙

(上海大学通信与信息工程学院, 上海, 200072)

摘要 研究了二代小波用于任意形状图像目标变换编码的可行性, 文中采用一种需要对目标作适当的延拓后再作变换的方法, 另一种则在修改二代小波提升算法的基础上获得一种不需延拓的形状自适应变换方法, 实验结果表明, 这二种方法与 Katata 使用的一代小波的方法相比, 有更好的压缩性能。

关键词 MPEG-4; 图像目标变换编码; 小波提升算法。

CODING OF IMAGE OBJECTS BASED ON WAVELET TRANSFORM USING LIFTING SCHEME*

ZHANG Zheng-Qing MO Yu-Long

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

Abstract Wavelet lifting scheme was applied to the transform coding of arbitrarily shaped image objects. There are two approaches; one needs the extrapolation of the image object and the other is a shape-adapted algorithm by modifying the lifting scheme. The experiment result shows that the image compression performance of the two approaches is better than Katata's approach which uses the classical wavelet transform.

Key words MPEG-4, transform coding of arbitrarily shaped image objects, wavelet transforms using lifting scheme.

引言

MPEG-4 是支持“基于内容的交互性”为核心的多媒体通信视频音频编码的算法。在 MPEG-4 中, 包含场景中特定对象或视频内容的任意形状的图像区域称为视频对象平面 (Video Object Plane VOP), 连续属于同一对象的 VOP 序列构成视频对象 (Video Object, VO), 该算法的编码是基于视频对象的, 即对多个 VO 的 VOP 序列的形状、运动、纹理以及接收端为辨别不同 VO, 重建原图和序列所需的相关信息分别进行编码, 在每一个 VOP 中, 大多数情况下是任意形状的, 这意味着必须研究对任意形状图像目标的变换编码方法。

在该标准的制定过程中, 建立了描述 MPEG-4 核心视频编码算法平台的视频检验模型, 在检验模型中, 由于 VO 的形状不规则, 对包含 VOP 边界

轮廓的宏块, 需要使用图像填充技术, 即用一定的灰度填充轮廓宏块中位于边界以外的像素; 检验模型的纹理编码变换方法, 采用 8×8 DCT 与填充技术结合的方法, 或使用形状自适应 DCT 变换 (shape-adaptive DCT, SA-DCT)^[2~4], 但这类方法仍然具有与 MPEG-1 等标准相同的分块处理所带来的问题, H. Katata^[5] 等人使用一代小波变换取代分块 DCT, 小波变换是一种全局变换, 避免了分块 DCT 的不足, 又能较好地反映人眼视觉特性, 一代小波用于任意形状图像的变换编码, 必须先将原图填充成一幅边长为 2 的整数次幂的矩形图像。

文中研究二代小波用于任意形状图像变换编码的方法, 二代小波具有能对任意尺寸图像变换的能力, 计算快速, 不存在一代小波中不同延拓方式会对重构产生不同影响的问题, 文中提出了二种方案, 第一种需要对原图作适当的延拓和填充, 第二种修改

* 国家自然科学基金(编号: 69875009)资助项目

稿件收到日期 2000-12-14, 修改稿收到日期 2001-02-14

* The project supported by the National Natural Science Foundation of China (No. 69875009)

Received 2000-12-14; revised 2001-02-14

了目标边界附近系数的小波变换算法以达到形状自适应的目的.通过实验与 Katata^[1]的方法进行了对比,结果表明这二种方案具有更好的压缩性能.

1 二代小波的基本原理

1994年,Swelden^[6]提出了一种不依赖于傅里叶变换的新的的小波构造方法—提升方法,有人称之为第二代小波或整数小波变换.二代小波提升算法的基本原理如下:

对于原始信号 s_j ,经小波分解成为低频信号 s_{j-1} 与高频细节信号 d_{j-1} ,由提升方法构成的小波变换过程可以分为分裂、预测、更新三个步骤.

1.1 分裂

将信号 s_j 分裂成为两个互不相交的子集 s_{j-1} 和 d_{j-1} ,通常是将一个数列分偶数序列和奇数序列,即

$$\text{split}(s_j) = (\text{even}_{j-1}, \text{odd}_{j-1}) = (s_{j-1}, d_{j-1}). \quad (1)$$

1.2 预测

利用数据间的相关性,用 s_{j-1} 去预测 d_{j-1} ,即采用一个与数据集结构无关的预测算子 P ,使得 $d_{j-1} = P(s_{j-1})$,由于在重构 s_j 时,可以用 s_{j-1} 来预测失去的子集 d_{j-1} ,因此就可以用子数据集 s_{j-1} 代替原始的数据集 s_j .在实际的应用中,无法用预测算子去精确地预测子集 d_{j-1} , $P(s_{j-1})$ 只是尽可能地接近 d_{j-1} .若用子集 d_{j-1} 与预测值 $P(s_{j-1})$ 的差值去代替 d_{j-1} ,此差值反映了两者的逼近程度.如果预测是合理的,则差值数据集所包含的信息比原始子集 d_{j-1} 包含的信息要少得多.预测过程的表达式为:

$$d_{j-1} = \text{odd}_{j-1} - P(\text{even}_{j-1}) = d_{j-1} - P(s_{j-1}). \quad (2)$$

1.3 更新

经过分理解步骤产生的系数子集 s_{j-1} 的某些整体性质(如均值)并不和原始数据中的一致,因此需采用更新过程.其方法是通过算子 U 产生一个更好的子数据集 s_j ,使之保持原数据集 s_j 的一些特性, s_j 的定义为

$$s_j = \text{even}_{j-1} + U(d_{j-1}) = s_{j-1} + U(d_{j-1}). \quad (3)$$

对于更新后的数据集 s_{j-1} 进行相同的分裂、预测和更新的过程,即把 s_{j-1} 分解成 d_{j-2} 和 s_{j-2} ,经过 n 次分解后,原始数据 s_j 的小波表示为 $(s_{j-n}, d_{j-n}, d_{j-n-1}, \dots, d_{j-1})$,其中 s_{j-n} 代表了信号的低频部分,而 $(d_{j-n}, \dots, d_{j-1})$ 则是信号的高频部分.

重构数据时的提升公式与分解公式相同,区别仅在于计算次序及正负号,可表达为:

$$s_{j-1} = s_{j-2} - U(d_{j-1}), \quad (4)$$

$$d_{j-1} = d_{j-1} + P(s_{j-1}), \quad (5)$$

$$s_j = \text{Merge}(s_{j-1}, d_{j-1}). \quad (6)$$

式(6)中 Merge 指将 s_{j-1} 和 d_{j-1} 分别作为偶数序列和奇数序列拼接成 s_j .

二代小波提升算法的提出,引出了一代小波变换是否也可以转化为提升步骤解决的问题. Daubechies 和 Swelden^[7]证明任意具有有限长滤波器的高散小波变换都可以通过对它的多相矩阵进行因式分解化为有限的提升步骤来解决.而且一代小波变换转化为二代提升算法后运算速度有所提高, Daubechies 和 Seelden 又证明:当滤波器长度较长时,提升算法比直接用一代小波运算次数约减少一半.

与一代小波变换相比,二代提升算法除具有速度快的优点外,其它特点有:能实现任意图像尺寸的小波变换,而一代二进小波变换要求图像的长宽为 2 的幂数;它不依赖傅里叶变换实现小波构造,可直接在时域中完成小波构造;它对内存的需求量小,在当前位置即可进行小波变换,便于用 DSP 芯片实现;它不受边界延拓方式的影响等.

2 基于二代小波提升算法的任意形状图像目标变换编码方法

图像分割后的各个目标有些具有矩形的形状,但大多数情况下是任意形状的.本文利用二代小波算法的特点,采用了二种方法进行了二代小波用于任意形状图像目标变换编码的研究.第一种方法根据变换的需要对目标进行适当的延拓,第二种方法则在提升算法基础上获得一种较简便的形状自适应变换算法.

2.1 局部延拓方法

二代提升算法虽然对图像尺寸没有要求,但对任意形状图像目标进行二维变换时,会产生以下问题,如图 1 所示,当先对目标的每一行进行一维变换后,由于各行的起始位置并不一定在同一列上,因此变换后行与行之间低频系数或高频系数并不能对准,即同一列中低频系数和高频系数会交替出现,不能保证每一列只有低频系数或高频系数.如果再进行列方向上的一维变换,则需要移位,再作变换.这样做计算量大,又会破坏图像自身的空间相关性,而且会丧失提升算法可直接在原位置上进行的优点.

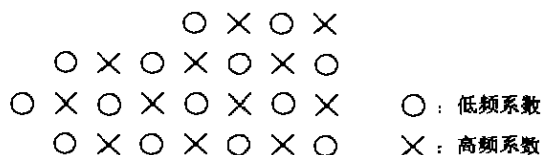


图1 行提升后低频(高频)系数不能对准的情况

Fig. 1 After lifting of lines, low(high) frequency coefficients are not in the same column



图2 延拓并填充后的视频对象
Fig. 2 The video object after extrapolating

通过对目标进行适当的延拓解决不能对准问题,方法如下:提升算法的主要优点是可以直接在原位置上进行,对一幅矩形的图像经过 n 次二维提升算法后,原图将变换为多个 $2^n \times 2^n$ 的正方形块组成的图像(在实际使用中,与一代小波相同,一般对图像作3~4次二维变换),而对应于每一个块,左上角第一个系数为变换后的低频系数,而其余 $2^n \times 2^n - 1$ 个系数分别对应于不同分辨率级下,不同方向上的高频系数,可以看出,若保证图像由若干个 $2^n \times 2^n$ 的正方形块组成,则不论其总的形状如何,都可保证在 n 次二维变换中行(或列)方向上低频系数(高频系数)的对准。

因此,对于一幅任意形状的图像目标,在对其作 n 次小波提升算法前,有必要将其延拓成每边边长是 2^n 的整数倍的一个多边形,例如为满足4次二维变换的需要,该多边形每边像素数都为16的整数倍(如图2所示),包含目标边缘的块中可能有位于原图之外的像素,这些像素的值可以用该块内有值像素的灰度均值来填充,对于延拓后的图像目标,虽然每一行/列的像素数并不相同,但提升算法不受信息尺寸影响,按正常的二代提升算法进行计算,信号重构公式与分解公式相同,区别仅在于计算次序及正负号。

2.2 形状自适应提升算法

在这种方法中,直接对目标作变换,而不加以延拓,同时又要避免直接对非方形图像作二维变换时会出现的问题,对于位于目标边缘附近的像素作提升算法时需要作一下修改,修改方法如下:

给定一幅任意形状的图像目标,可以确定一个完全包含该目标的一个各边边长为 2^n 的整数倍的多边形,与局部延拓方法相同,在这样的多边形基础上作 n 次二维变换,可保证变换中低频系数(高频系数)的对准,并且若以该多边形为基础作 n 次变换,可确定最终每个像素位置上将被哪个分辨率级下,

哪个方向上的高频系数或低频系数所取代,对于位于目标内的每一行(或列)像素则会存在四种情况:

- (1) 该行首像素为低频系数所取代,尾像素为高频系数所取代。
- (2) 该行首像素为高频系数所取代,尾像素为低频系数所取代。
- (3) 该行首像素为低频系数所取代,尾像素为低频系数所取代。
- (4) 该行首像素为高频系数所取代,尾像素为高频系数所取代。

上述四种情况中,第一种行(列)可以用正常的一维提升算法计算,并可完全重构,而后三种情况只有在任意形状图像目标中遇到,直接用一维提升算法处理,在边界处不能做到完全重构,必须进行填充。

本方案对小波变换边缘延拓作了修改,修改后的边缘延拓算法规定:在第一次二维小波提升中,以边缘所在局部范围的原图均值作为目标边缘外延拓的像素值;在以后各次二维小波提升中,当目标边缘外延拓的像素位置对应于变换后低频系数位置时,以该边缘所在局部范围的原图均值作为该像素的灰度值,当目标边缘外延拓的像素位置对应于变换后高频系数位置时,以0作为该像素的灰度值,反变换也使用同样的延拓方法,这一方法需要将原图目标边界处局部均值作为附加信息编码传输,而不需要对原图做填充,达到了形状自适应的目的,选用局部均值代替低频系数的原因是小波提升算法要求低频系数能够保持原数据的均值特性,而用0代替高频系数是基于小波提升算法希望预测步骤精确以使高频系数较小,实验结果也表明采用这种修改是可行的。

以Cohen-Daubechies-Feauveau(2,4)双正交滤

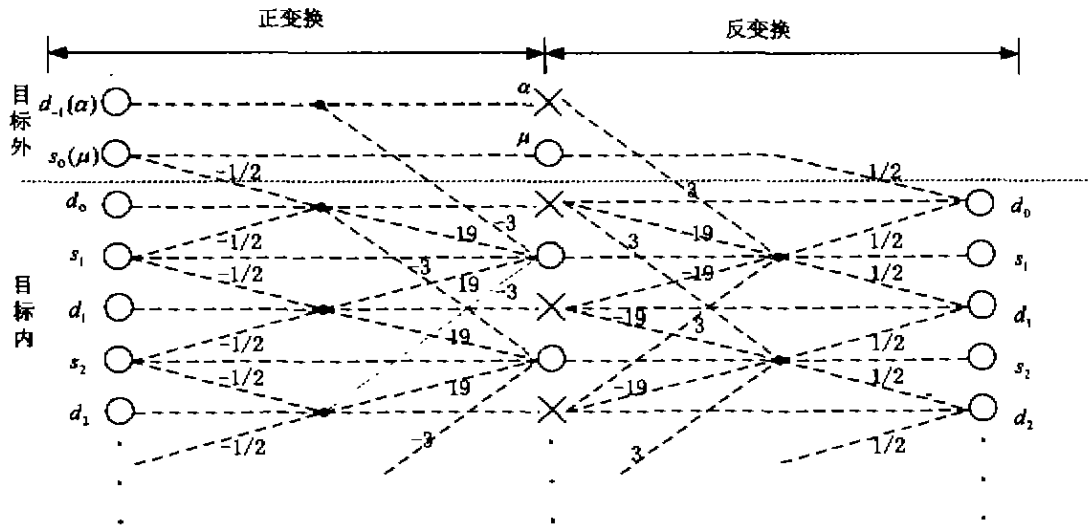


图 3 边缘附近的形状自适应提升算法

Fig. 3 The shape-adapted lifting algorithm near the edge

波器为例(见图 3),若边缘由 d_0 上部经过, d_0 变换后将由高频系数取代,采用形状自适应方案延拓后,在边缘外对应变换后低频系数位置处填充 μ ,在对应变换后高频系数位置处填充 α , μ 为 d_0 所在局部范围的原图均值, α 在第一次二维变换时为 d_0 所在局部范围的原图均值,在以后各次变换中为 0,这样相应 d_0, s_1, d_1 的计算式修改为:

$$d_0 \leftarrow d_0 - (s_0 + s_1)/2 = d_0 - (\mu + s_1)/2. \quad (7)$$

$$d_1 \leftarrow d_1 - (s_1 + s_2)/2. \quad (8)$$

$$s_1 \leftarrow s_1 - (3d_{-1} - 19d_0 - 19d_1 + 3d_2)/64$$

$$= s_1 - (3 \times \alpha - 19d_0 - 19d_1 + 3d_2)/64. \quad (9)$$

反变换时仍以同样方法延拓, d_0, s_1, d_1 的反提升计算式只需将式(7)、(8)、(9)中第一个减号改为加号,边缘处的像素可以得到完全重构。

表 1 实验结果

Table 1 Experimental Results

实验	编码率 (bpp)	PSNR(dB)	编码时间
1(轮廓向外扩张 3个系数)	0.4633	28.52	1'20"
1(轮廓向外扩张 5个系数)	0.474	30.62	1'20"
1(轮廓向外扩张 10个系数)	0.5057	32.09	1'32"
2	0.44	32.53	1'03"
3	0.4157	33.37	51"

3 实验结果与分析

实验中采用的任意形状的图像取自 QCIF 的视顿流(见图 4)。对于轮廓的编码采用 S. Lee^[8]等人的方法,码长为 698bit。小波系数的编码参照了 Lewis 和 Knowles^[9]的树形编码方法。小波基采用 Cohen-Daubechies-Feauveau (2,4) 双正交基。所有实验在 Pentium II 350 上完成,编程语言使用 Matlab5.3。

实验结果中,实验 1 为 Katata 的一代小波变换编码方法,保留的系数范围为目标向外扩张 3 个、5 个和 10 个系数;实验 3 为局部延拓方法;实验 2 为形状自适应提升算法,图 5、图 6 和图 7 为各实验对应编解码结果图。

由实验结果可见,在相近编码率条件下,采用 Katata 的一代小波方法 PSNR 较低,这是由于该方法需要由图像轮廓沿轮廓法线方向向外扩张 10 个像素左右才能保证解码恢复的图像质量,但这样做需要对更多的系数进行编码,压缩比较低。采用二代提升算法的第一种方法,除位于目标内的系数外,还需对目标外填充像素位置上的系数进行编码,由图 2 可以看出,该方法比 Katata 方法所需编码系数少,而且包含目标边缘的块用该块内有值像素的灰度均值来填充,可以保证在目标边缘处不产生过大的高频系数。采用二代提升算法的第二种方法不需要对目标进行延拓,除作为附加信息要多传边缘附近局部均值(在实验中局部均值包含边缘的各块中



图4 原图
Fig. 4 Original image

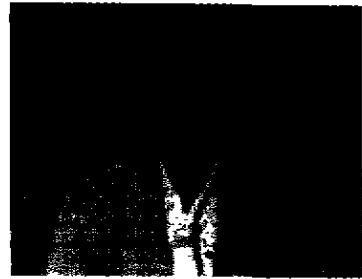


图6 实验2
Fig. 6 Recover image of experiment 2.



图5 实验1(轮廓向外扩张10个系数)
Fig. 5 Recover image of experiment 1.
(expanding 10 coefficients
along normal line of edge)



图7 实验3
Fig. 7 Recover image of experiment 3

目标内像素的均值(见图2),并沿顺时针的顺序排列后作DPCM编码,所需比特数不多)外,只需对目标内的变换系数进行编码,因此后二种方法压缩比更高,相近编码率时恢复效果更好。Katata方法在变换编码过程中需要将目标延拓成方形,变换后又要对目标轮廓进行扩张并将部分小波系数置0,运算量较大,而实验二只作局部延拓,实验三无延拓,因此基于提升算法的后二种方法编码速度较快。

二代提升算法除保留了对应一代小波的变换压缩性能外,二代小波可在空间域内对信号作计算,且对信号尺寸无要求,因此与一代小波相比,具有算法简单和易于DSP实现的优点,实验分析表明二代小波在任意形状图像目标变换编码中具有潜在的应用前景。

REFERENCES

- [1] Sikora T. The MPEG-24 video standard verification model. *IEEE Trans. on CSVT.*, 1997, 7(1): 19-31
- [2] Sikora T. Shape-adaptive DCT for generic coding of video. *IEEE Trans. on CSVT.*, 1995, 5(1): 59-62
- [3] Kauff P, Schuur K. Shape-adaptive DCT with block-based DC separation and Δ DC correction. *IEEE Trans. on CSVT.*, 1998, 8(3): 237-242
- [4] Kauff P. Functional coding of video using a shape-adaptive DCT algorithm and an object-based motion prediction toolbox. *IEEE Trans. on CSVT.*, 1997, 7(1): 181-196
- [5] Katata H, Aono N I T, Kusao H. Object wavelet transform for coding of arbitrarily shaped image segments. *IEEE Trans. on CSVT.*, 1997, 7(1): 234-237
- [6] Sweldens W. The lifting scheme: a custom-design construction for biorthogonal wavelets. Technical Report, University of South Carolina, 1994; 1-39
- [7] Daubechies I, Sweldens W. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 1994, 4(3): 247-269
- [8] Lee S H, Cho D, Cho Y, Son S H. Binary shape coding using baseline-based method. *IEEE Trans. on CSVT.*, 1999, 9(1): 44-53
- [9] Lewis A S. Image compression using the 2-D wavelet transform. *IEEE Trans. on IP.*, 1992, 1(2): 244-250