

基于方向提升小波变换的多描述视频编码

陈锦春, 蔡灿辉

(华侨大学信息科学与工程学院, 泉州 362021)

摘要: 提出一种基于方向提升小波变换的多描述可分级视频编码方法, 该方法能增强视频编码的容错性和网络自适应性。对视频序列中的每一帧图像按梅花下采样产生 2 个子序列, 对每个子序列分别进行三维小波可分级视频编码形成 2 个描述。为解决传统二维小波变换方法对梅花下采样图像编码效率低的问题, 引入方向提升小波变换对图像进行空域分解, 提出 5 模式方向提升小波变换以进一步提高编码效率。实验结果证明了该方法的有效性。

关键词: 多描述可分级编码; 梅花下采样; 三维小波视频编码; 平面预测模式; 图像插值

Multiple Description Video Coding Based on Directional Lifting Wavelet Transform

CHEN Jin-chun, CAI Can-hui

(College of Information Science and Engineering, Huaqiao University, Quanzhou 362021)

【Abstract】 A multiple description scalable video coding based on directional lifting wavelet transform is proposed, it can enhance the error-resilient and network-adaptive ability of video coding. It divides each frame into two subframes to form two subsequences by using quincunx down-sampling, each subsequence is coded by a scalable 3D wavelet video coder to form one description. To improve coding efficiencies of the subframes, directional lifting wavelet transform is used to perform spatial decomposition. A five mode directional lifting is introduced to improve coding efficiency. Experimental results prove the validity of the proposed algorithm.

【Key words】 Multiple Description Scalable Coding(MDSC); quincunx down-sampling; 3D wavelet video coding; plane prediction mode; image interpolation

1 概述

随着网络与多媒体技术的不断发展, 视频编码与传输成为通信领域的研究热点, 但互联网和移动信道都是不可靠信道, 而且用户终端可用的带宽不尽相同, 如何让服务器具有提供不同层次视频质量的能力, 以及如何利用不可靠信道传输高质量的视频信号成为当前视频通信需要解决的难题。多描述可分级视频编码的出现, 为解决上述 2 个难题提供了途径。

多描述编码(Multiple Description Coding, MDC)^[1-3]是将图像分成几个描述, 各个描述都携带着原始数据的基本信息, 描述间存在互补关系。编码端分别对各个描述进行独立编码, 解码端则依据接收情况选择边缘解码或中央解码。当仅接收到 1 个描述时接收端仍可恢复出质量可以接受的图像。随着接收的描述数增多, 图像的重建质量也得到提高, 当接收到所有描述时, 图像获得最高的重建质量。如果对各描述进行可分级编码, 就形成多描述可分级编码(Multiple Description Scalable Coding, MDSC)。MDSC 使传输的信号不但具有良好的容错性能, 而且对网络带宽的变化有很强的自适应能力。文献[2]提出一种基于运动补偿时域滤波的多描述可分级编码。具体做法是, 按帧序的奇偶性, 将运动补偿时域滤波产生的高频帧划分为 2 个描述, 同时将低频帧与运动矢量作为冗余信息复制到各个描述。这样当部分高频帧丢失时(如只接收到一个描述时), 解码端可以依靠低频帧与运动矢量近似恢复出相应的帧。文献[3]则采用空域分解的方法, 通过按行(列)

下采样将图像划分成 2 个描述。在解码端, 利用像素的空间相关性, 通过差错掩盖重建原始分辨率大小的图像。文献[4]提出一种对原始图像进行梅花下采样形成 2 个描述的算法。由于梅花下采样产生的 2 个子图是高度相关的, 空域差错掩盖的效果好, 但采用传统的小波变换对采样子图编码效率低。为此, 该文引入了方向提升小波, 在一定程度解决了上述问题。遗憾的是文献[4]的算法只是一种静态图像编码算法。

本文在文献[4]的基础上, 提出基于方向提升小波变换的多描述可分级视频编码。首先对视频序列中每幅图像按梅花下采样形成 2 个子序列(即 2 个描述), 并利用方向提升小波解决各采样子图内相邻行之间的错位问题。采用可分级三维小波视频编码器^[5]分别对每个子序列进行编码, 产生码率可分级的多描述编码。为了获得更好的预测与更新效果, 本文在传统方向提升小波基础上增加一个平面预测模式, 进一步提高了方向提升小波的编码性能。

2 多描述可分级视频编码系统

2.1 框架概述

本文提出的多描述可分级视频编码系统框架见图 1。

基金项目: 国家自然科学基金资助项目(60772164); 福建省自然科学基金资助项目(A0710009)

作者简介: 陈锦春(1982 -), 男, 硕士研究生, 主研方向: 图像处理, 视频编码; 蔡灿辉, 教授

收稿日期: 2009-03-21 **E-mail:** donalds2008@163.com

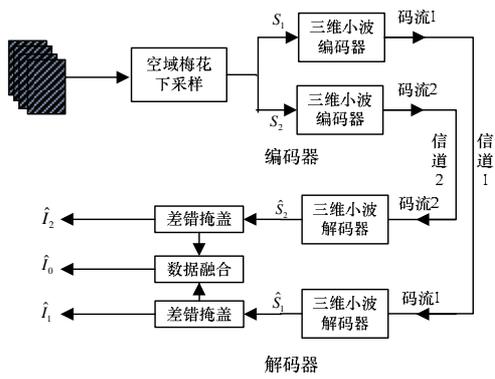


图1 MDSC 编解码器框图

令 I 表示原始序列的图像数据，经空域梅花下采样获得 2 个非空子集 S_1, S_2 ，则有：

$$I = S_1 \cup S_2, \quad S_1 \cap S_2 = \emptyset \quad (1)$$

S_1 和 S_2 构成原始序列的 2 个描述。这 2 个描述分别通过不同的三维小波编码器编码后，经不同的信道传送到接收端。在接收端，每个描述经过独立的三维小波解码和差错掩盖后，形成各自的边缘解码重建图像 \hat{I}_1 和 \hat{I}_2 。如果只接收到一个描述，则输出相应的重建图像，否则，对 \hat{I}_1 和 \hat{I}_2 进行融合，建立中央解码重构图像 \hat{I}_0 。

2.2 三维小波编码器

三维小波编码器由运动补偿时域滤波，空域二维小波变换和三维时空子带编码等部分组成，如图 2 所示。

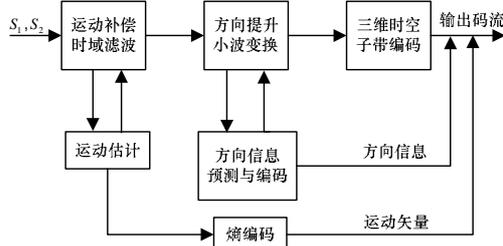


图2 三维小波编码器框图

运动补偿时域滤波和空域二维方向提升小波变换用于去除视频序列的时空冗余，并提供对视频数据在时空域的多分辨率分析能力。三维时空子带编码采用 MC-EZBC^[5]，产生的码流具有完全嵌入的特点，因此，可以实现对每个描述进行可分级编码，增强各个描述的网络自适应能力。

在空域二维小波变换部分，方向提升小波变换通过对图像进行空域插值预处理，使滤波总是沿着最优方向进行，能够比传统提升小波变换更有效地去除图像的空间冗余。

3 方向提升小波变换

3.1 方向提升小波变换的引入

传统的图像二维提升小波变换是在水平和垂直方向进行 2 次独立的一维提升小波变换。提升小波变换主要由分裂，预测，更新 3 个步骤组成。令 $x[m, n]$ 代表位于图像中第 m 行，第 n 列的像素值，不失一般性，假设对图像先进行一次垂直小波提升变换，则其分裂，预测，更新过程如下：

(1)分裂：将 $x[m, n]$ 分裂成偶像素集 $x_e[m, n]$ 和奇像素集 $x_o[m, n]$ ，其中：

$$x_e[m, n] = x[2m, n], x_o[m, n] = x[2m+1, n] \quad (2)$$

(2)预测：将与奇像素 $x_o[m, n]$ 相邻的偶像素集进行线性

组合，用于预测该奇像素：

$$P(x_o)[m, n] = \sum_i p_i x_e[m+i, n] \quad (3)$$

其中， p_i 为相应的高频滤波系数。用预测误差值 $d[m, n]$ 代替奇像素的值，获得高频系数：

$$d[m, n] = x_o[m, n] - P(x_o)[m, n] \quad (4)$$

(3)更新：将与偶像素 $x_e[m, n]$ 相邻的高频系数进行线性组合，用于更新该偶像素：

$$U(d)[m, n] = \sum_j u_j d[m+j, n] \quad (5)$$

其中， u_j 为相应的低频滤波系数。用更新值 $c[m, n]$ 代替偶像素的值，获得低频系数：

$$c[m, n] = x_e[m, n] + U(d)[m, n] \quad (6)$$

原始视频经过运动补偿时域滤波后，在时间域上被分解为一系列的时域低频帧和高频帧^[5]。这些帧仍保持一定的纹理结构。由于现有的小波变换没有直接针对梅花下采样图像的处理模型，必须将偶数行的像素左移一位，如图 3 所示。

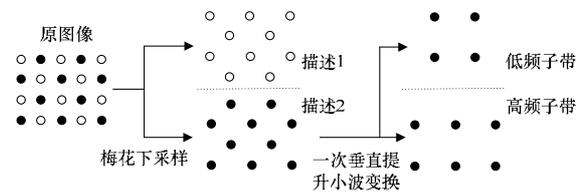


图3 图像经过梅花下采样和一次垂直小波变换后的像素位置

再根据式(2)~式(6)进行变换。该移位操作破坏了图像的纹理结构，将产生大量的高频成份，不利于压缩编码。因此，引入方向提升小波变换。为叙述方便，以经过奇偶行对齐像素间的空间位置关系作为参考，在垂直方向上按式(2)划分像素的奇偶性，如图 4 所示。

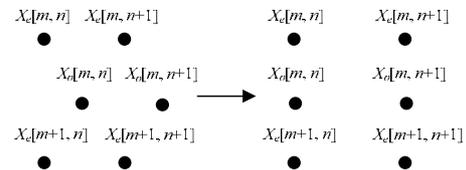


图4 错位像素的奇偶行对齐

在每一次预测或更新处理前，方向提升小波需要对由上一次更新或预测得到的系数进行插值。以第 1 次预测阶段为例，如图 5 所示，其中，黑色圆圈代表采样图像中的像素，灰色圆圈代表插值像素。

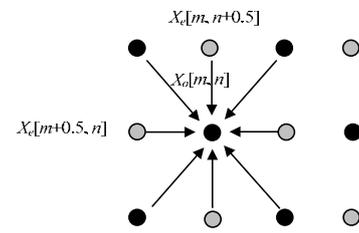


图5 预测方向选择

采样图像经过插值之后，便得到一个与原图像相同分辨率大小的图像。由图可见，对奇像素 $x_o[m, n]$ 在前向预测时共产生 4 个候选方向(0~3)，分别对应偶像素为： $x_e[m, n+1], x_e[m, n+0.5], x_e[m, n]$ 和 $x_e[m+0.5, n]$ 。这 4 个偶像素形成一个候选集合，在其中选择一个能获得最小预测误差 $d[m, n]$ 的像素用于预测 $x_o[m, n]$ ，也就确定了最优方向。

由于解码端需要借助方向信息才能进行相应的逆变换，而编码端传送方向信息需要增加比特开销。更新阶段，对偶像素 $x_e[m, n]$ 采用与预测方向相同的方向进行更新。这样处理一方面可以节省用于编码方向信息的比特开销，另一方面对于多数图像而言，最优的更新方向一般与最优的预测方向是一致的。此外，为了更好地权衡编码效率与方向信息的开销，本文采用二叉树结构进行图像的分块和方向信息的组织^[4]。分块大小为 $16 \times 16 \sim 4 \times 4$ 。

综上所述，类似式(3)、式(5)，方向提升小波变换中的预测，更新可表示为

$$P_{dir}(x_e)[m, n] = \begin{cases} \sum_i p_i x_e[m+i, n + \frac{1+(-1)^{i-1}(dir-1)}{2}] & dir = 0, 1, 2 \\ \sum_i p_i x_e[m+0.5, n+i] & dir = 3 \end{cases} \quad (7)$$

$$U_{dir}(d)[m, n] = \begin{cases} \sum_j u_j d[m+j, n + \frac{1+(-1)^j(dir-1)}{2}] & dir = 0, 1, 2 \\ \sum_j u_j d[m+0.5, n+j] & dir = 3 \end{cases} \quad (8)$$

其中， dir 代表预测与更新的方向。为叙述方便，本文以 P_{dir} 、 U_{dir} 分别代表式(7)和式(8)中的 $P_{dir}(x_e)[m, n]$ 和 $U_{dir}(d)[m, n]$ 。

值得注意的是，通常图像需要经过多级二维小波变换，而采样子图在进行一次方向垂直提升小波变换后，低频(高频)子带内的像素之间已不存在错位关系，如图 3 所示，此后的每一级垂直(水平)提升小波变换均采用传统一维小波变换的做法，也就是不插值，直接对高频(低频)子带内的像素进行垂直(水平)方向的一维小波变换。

3.2 基于欧氏距离的平面预测模式

必须指出，上述的二维图像插值是按行、按列进行的，所得到的插值点只利用到了像素的行相关性或列相关性。为了充分利用图像的平面相关性，本文在上述 4 个方向预测模式的基础上，增加一种新的预测模式——基于欧氏距离的平面预测模式。其基本思路是，对上述 4 个方向的预测结果，分别依据其最近的预测点(偶像素)与被预测点(奇像素)的空间距离分配权值，构造平面(无方向)预测模式，作为一种最优方向的候选。平面预测和更新计算公式如下：

$$P_4 = w_0 P_0 + w_1 P_1 + w_2 P_2 + w_3 P_3, \sum_{dir=0}^3 w_{dir} = 1 \quad (9)$$

$$U_4 = w_0 U_0 + w_1 U_1 + w_2 U_2 + w_3 U_3, \sum_{dir=0}^3 w_{dir} = 1 \quad (10)$$

其中， w_{dir} 为对应各方向偶像素的权值，和预测点与被预测点的空间距离成反比。距离越近的像素点，相关性越大，分配权值越大。设 d_{dir} 为对应于方向 dir 的预测点与被预测点之间的欧氏距离，获得归一化常数因子：

$$K = \sum_{dir=0}^3 \frac{1}{d_{dir}} \quad (11)$$

则加权值 w_{dir} 为

$$w_{dir} = \frac{1}{K} \times \frac{1}{d_{dir}}, \sum_{dir=0}^3 w_{dir} = 1 \quad (12)$$

实验结果表明，与传统的方向提升小波算法相比，增加平面预测的方向提升小波对边缘解码和中央解码的重建质量有不同程度的提高，如表 1 所示。考虑到平面预测通常只对较均匀的区域起作用，且增加该模式需要增加额外的比特开销，因此本文只在选择最大块(16×16)的方向时加入平面预测模式。为了叙述方便，本文把增加平面预测后的提升小波称为“5 模式方向提升小波”。

表 1 Salesman 序列的边缘解码和中央解码质量比较

中心码率 (Kb·s ⁻¹)	传统的方向提升小波		5 模式方向提升小波	
	边缘解码 (均值)/dB	中央解码 (均值)/dB	边缘解码 (均值)/dB	中央解码 (均值)/dB
128	28.27	28.81	+0.04	+0.11
256	30.92	32.51	+0.03	+0.05
512	32.74	35.95	+0.02	+0.03

4 边缘解码与中央解码算法

4.1 边缘解码与差错掩盖

当接收端收到一个描述的码流时(不失一般性，设为 \hat{S}_1)，则将码流送到相应的三维小波解码器重建该描述。解码端可以利用两采样图像间的强相关性，对 \hat{S}_1 进行空域差错掩盖得到边缘解码重建图像 \hat{I}_1 ：将重建子图 \hat{S}_1 的像素按梅花下采样前的位置进行放置，缺失像素可通过近邻的像素进行线性插值得到。

4.2 中央解码与数据融合

当解码端收到 2 个描述时，对解码得到的重建子图 \hat{S}_1 和 \hat{S}_2 进行与边缘解码相同的插值。这样，对应一帧原始图像可以得到 2 个重建图像版本 \hat{I}_1 和 \hat{I}_2 ，对应 2 个描述相同空间位置的像素，任一描述中包含的插值像素值对另一描述中经直接逆变换得到的像素值仍有一定补偿作用。因此，可对两者分配不同权重，重构出中央解码图像 \hat{I}_0 。一般地，取直接经逆变换得到的像素值的权重高于插值像素的权重。在本文实验中前者取 0.75，后者为 0.25。

5 实验结果与分析

为测试本文提出的视频多描述方案的性能，本文用 CIF 格式(352×288)的典型测试序列：Foreman 和 Salesman 进行测试。为方便与文献[3]的比较，对实验条件设定如下：(1)选取序列前 80 帧用于测试；(2)运动估计采用半像素精度；(3)空域分解采用双正交 9/7 滤波器；(4)采用 MC-EZBC 作为三维时空子带编码器。

表 2 分别给出采用传统提升小波变换和方向提升小波变换的视频多描述编码，在不同码率下得到的边缘解码和中央解码结果。表中的数据为重建图像的 Y 分量平均 PSNR 值。

表 2 采用不同提升框架得到的边缘解码和中央解码的结果比较

测试序列	中心码率 (Kb·s ⁻¹)	传统提升小波变换			方向提升小波变换		
		边缘解 码 I/dB	边缘解 码 II/dB	中央解 码/dB	边缘解 码 I/dB	边缘解 码 II/dB	中央解 码/dB
Foreman	128	26.79	26.74	27.10	+0.35	+0.37	+0.49
	256	29.99	29.96	30.92	+0.31	+0.31	+0.45
	512	31.66	31.64	33.24	+0.21	+0.17	+0.32
Salesman	128	27.89	27.96	28.35	+0.37	+0.39	+0.57
	256	30.64	30.68	32.02	+0.31	+0.28	+0.53
	512	32.62	32.66	35.61	+0.14	+0.10	+0.37

图 6 分别给出本文算法与文献[3]平均边缘解码(仅接收到一个描述)和中央解码(同时收到 2 个描述)的效果比较。图 6(b)给出采用三维小波编码器，对测试序列不进行 2 个描述划分而是直接编码(即单描述)情况下，在解码端获得的重建效果。测试序列为 Salesman。

从图 6 可以明显看出，在相同速率和相近中心解码输出条件下，本文所提出的多描述视频编码方法产生的边缘解码输出优于文献[3]提出的 2 个算法。

(下转第 175 页)