

基于 Surfacelet 变换和 SPIHT 算法的视频压缩

王海燕 殷 俊 潘显萌
(公安部第三研究所 上海 201204)

摘 要 提出一种基于 Surfacelet 变换并结合 SPIHT 算法的视频压缩编码方法。这种新方法把视频信号作为特殊的三维信号,对信号的空间和时间维进行整体处理。Surfacelet 变换具有多方向分解、各向异性、高效率的树结构滤波器组、可完全重建和低冗余度等性质。SPIHT 算法具有分辨率质量上的可伸缩性、渐进传输性等特性。利用 Surfacelet 变换的分解系数在各层间相关以及图像能量集中的特性,结合 SPIHT 算法完成视频数据的压缩编码。这种新的视频压缩编码方法能弥补三维小波变换的缺陷,达到更高的 PSNR 值和更好的视觉效果,尤其适用于纹理复杂度较高、运动幅度较小的视频。

关键词 Surfacelet 变换,多尺度分解,多维方向滤波器组,SPIHT 算法,视频压缩
中图分类号 TP37 文献标识码 A

Video Compression Based on Surfacelet Transform and SPIHT Algorithm

WANG Hai-yan YIN Jun PAN Xian-meng
(The Third Research Institute of Ministry of Public Security, Shanghai 201204, China)

Abstract A new video compression coding method based on Surfacelet transform and SPIHT algorithm was proposed. The new method treats the video signal in a special 3D signal with space and time axes in a whole. Surfacelet transform achieves properties of multi-directional decomposition, anisotropy, efficient tree-structured implementation, refinable angular resolution, low redundancy and so on. SPIHT algorithm is scalable and progressive in resolution. The video data compression coding is completed by using the properties of correlation and energy concentration between layers of decomposing factors in Surfacelet transform and SPIHT algorithm. The new video compression coding method can overcome the shortcoming of the method based on 3D wavelet transform, achieving higher PSNR and better vision effect, especially for video with complex texture and small motion.

Keywords Surfacelet transform, Multiscale decomposition, NDFB, SPIHT algorithm, Video compression

1 概述

随着计算机技术和半导体材料的发展,高分辨率的视频数据应用于多个工程领域,如多媒体视频流、医学成像、遥感测绘、计算机视觉等。为了解决这些数字化视频数据大容量存储和实时传输的问题,得到令人满意的视频画面质量以及处理和提取有效的特征信息,除了要提高计算机本身性能和通信信道带宽外,更重要的是对视频体数据进行有效压缩。

小波变换具有良好的空间和频率局部化特性,在信号处理中适用于非平稳信号的分析^[1]。以多分辨率分析理论为基础,基于小波变换的图像及视频压缩应用研究取得了许多成果,部分算法已融入 JPEG2000 等压缩标准中^[2]。但由于其方向选择性差,不能有效处理线状奇异性,该变换的视频压缩效果受到一定限制。多尺度几何发展的目的和动力就是克服小波变换的缺点,致力于在高维情况下改进一种新的高维函数的最优或“最稀疏”的函数表示方法,从而衍生出了多种超小波分析方法。

1998 年, Candès 和 Donoho 提出 Ridgelet 变换并其具有

线奇异性的多变量函数的最优基,但对图像中曲线边缘特征的表现和重构无法达到理想效果。2002 年, Do 和 Vetterli 提出 Contourlet 变换,它能更准确地捕获图像边缘系数,能量更加集中,但是没有基于临界采样的滤波器组,仍有划痕现象^[3]。2007 年, Yue Lu 和 M NDo 提出基于 Surfacelet 变换的三维变换,其具有的多方向分解、各向异性、高效率的树结构滤波器组、完全重建和低冗余度等性质。

本文提出一种基于 Surfacelet 变换,结合 SPIHT 算法的视频压缩编码方法,把视频信号作为特殊的三维信号(即两个空间维和一个时间维),对信号的空间和时间信息进行整体处理。

2 Surfacelet 变换

Surfacelet 变换对三维甚至更高维信号实现多尺度及多方向分解,捕获三维信号中的面状奇异,其主要包括多尺度分解和多维方向滤波器组(N-dimensional Directional Filter Banks, NDFB)。

2.1 多尺度分解

在 Surfacelet 变换中,多尺度分解采用多尺度的塔式结构来实现,如图 1 所示。

王海燕(1988—),女,硕士,助理研究员,主要研究方向为网络安全,E-mail:abwhy@163.com;殷俊(1990—),男,硕士,研究实习员,主要研究方向为网络侦查技术;潘显萌(1973—),男,硕士,助理研究员,主要研究方向为网络与通信技术。

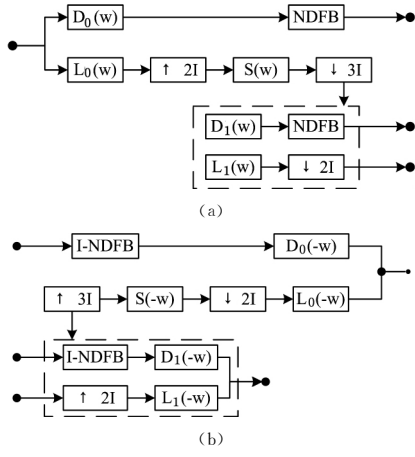


图1 Surfacelet 变换架构

在 Surfacelet 变换架构里, $D_i(\omega)$ ($i=0, 1$) 表示高通滤波器, $L_i(\omega)$ ($i=0, 1$) 表示低通滤波器, 用来实现多尺度分解得到对应信号的高频和低频成分。 $S_i(\omega)$ ($i=0, 1$) 表示抗混叠滤波器, 用来抑制上采样操作所造成的频域混叠问题。对图 1 中的虚线封闭框进行 N 次迭代操作, 能达到更多级尺度分解的效果。对各维第一级 $L_i(\omega)$ 进行 2 倍上采样后进行 3 倍下采样, 即非整数因子 1.5 倍下采样的操作, 虽然轻微增加了部分系统冗余, 但是大幅度减少了频域混叠现象^[4]。频域混叠主要来自离散信号频域周期 2π , 集中在 $[-\pi, \pi)^N$ 范围内。直观地说, 多尺度塔式结构只保留 NDFB 中间 $S_i(\omega)$ 部分的响应, 达到消除频域混叠的目的, 使得 Surfacelet 变换在时域和频域都有更好的细分效果。

图 1 实现的分解部分: 指定 $L_i(\omega)$ ($i=0, 1$) 为一维低通滤波器 $L_i^D(\omega) = d_i \cdot \prod_{n=1}^N L_i^{1D}(\omega_n)$, 其中 $d_1 = 6^{N/2}$, $d_2 = 2^{N/2}$, 通带频率为 $\omega_{p,i}$, 阻带频率为 $\omega_{s,i}$ 。

$$L_i^D(\omega) = \begin{cases} 1, & |\omega| \leq \omega_{p,i} \\ \frac{1}{2} + \frac{1}{2} \cos \left(\frac{|\omega| - \omega_{p,i}}{\omega_{s,i} - \omega_{p,i}} \pi \right), & \omega_{p,i} < |\omega| < \omega_{s,i} \\ 0, & \omega_{s,i} < |\omega| < \pi \end{cases}$$

指定 $S(\omega)$ 为一维抗混叠滤波器, 其中通带频率为 $\omega_{p,A}$, 阻带频率为 $\omega_{s,A}$ 。

$$S(\omega) = \begin{cases} 1, & |\omega| \leq \omega_{p,A} \\ \frac{1}{2} + \frac{1}{2} \cos \left(\frac{|\omega| - \omega_{p,A}}{\omega_{s,A} - \omega_{p,A}} \pi \right), & \omega_{p,A} < |\omega| < \omega_{s,A} \\ 0, & \omega_{s,A} < |\omega| < \pi \end{cases}$$

根据频带分解以达到抗频域混叠的目的, 选定 $L_i(\omega)$ ($i=0, 1$) 的通带频率 $\omega_{p,0} = \frac{\pi}{3}$, $\omega_{p,1} = \frac{\pi}{4}$, 以及阻带频率 $\omega_{s,0} = \frac{2\pi}{3}$, $\omega_{s,1} = \frac{\pi}{2}$; 选定 $S(\omega)$ 的通带频率 $\omega_{p,A} = \frac{\pi}{3}$, 以及阻带频率 $\omega_{s,A} = \frac{2\pi}{3}$ 。

在设计该频域滤波器时, 严格控制截止频率响应, 通过高频响应, 确保完全去除由上采样和下采样操作带来的频域混叠问题。对应的完美的多尺度金字塔状结构合成部分可以简化为:

$$\frac{|L_i(\omega)|}{d_i^2} + |D_i(\omega)|^2 = 1, i=0, 1$$

$L_i(\omega)$ 参数确定后即可推导出对应的 $D_i(\omega)$, 进行信号完整重建。

2.2 多维方向滤波器组 NDFB

基于视频的三维体数据处理, 选取三维情况下的 NDFB 结构进行分析, 即 3D-DFB。3D-DFB 结构如图 2 所示。

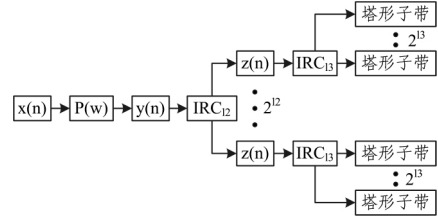


图2 3D-DFB 结构图

$x[n]$ 表示输入三维信号, 其 3 个维度分别为 n_1, n_2, n_3 , $P(w)$ 表示三维沙漏滤波器。 $y[n]$, $z[n]$ 表示输出三维信号, 其 3 个维度分别为 n_1, n_2, n_3 。 IRC_{12} 表示二维方向滤波器组。

$x[n]$ 首先通过 $P(w)$, 该滤波器由三通滤波器组组成。输出 $y[n]$ 通过 IRC_{12} , 基于 n_1, n_2 平面通过 2D-DFB 分解三维视频信号获得楔形子带。输出 $z[n]$ 再通过另一个 IRC_{13} , 基于 n_1, n_3 平面分解三维视频信号同样获得楔形子带。由以上两个楔形子带重合形成一个沿 n_1 轴方向的塔形子带, 构造了一个 3D-DFB。以同样的方式获得沿 n_2 轴和 n_3 轴方向的塔形子带, 最后共得到 2^{12+13} 的所有塔形子带。将不同尺度不同方向的分解子带作为 Surfacelet 变换后的系数矩阵, 也即视频序列经变换后的结果。Surfacelet 变换通过 3D-DFB 频率分解, 能够捕捉多维信号中的曲面奇异性, 实现对边缘信息的更全面提取^[5]。

2.3 Surfacelet 变换特性

Surfacelet 变换后的各个层系数之间存在着一定的关系, 与小波变换相类似, 它是一种很好的时频分析工具。同时由于 NDFB 对任意维信号进行方向分解, 迭代滤波器组可得到高效的树状结构, N 维的 NDFB 只有 N 倍的冗余度^[6], 使得 Surfacelet 变换能利用不同尺度、不同频率的方向子块准确地捕获三维信号中的面状奇异, 变换后的系数能量集中, 并能进行完全重建。

3 SPIHT 算法

基于三维的等级树集合分割编码 (Setpartitioning in Hierarchical Trees, SPIHT) 算法是通过 3 个链表把所有像素点按三维空间方向树的结构组织起来进行编码^[7]。重要像素链表 (LSP)、不重要像素链表 (LIP) 和不重要集合链表 (LIS) 存放像素坐标, 在编码过程中动态更新。SPIHT 算法在空间、时间方向上形成三维空间方向树, 即八叉树结构。算法过程如下。

计算编码时的最大阈值, 初始化链表 LIP、LIS、LSP。令 LSP 为空, 将低频根节点坐标加入 LIP, 将具有子孙的根节点加入 LIS, 然后从根节点开始扫描。排序扫描 LIP 中的节点, 检查变换系数的重要性。若为重要系数则进入细化过程, 输出细化有效位, 移入 LSP; 若为非重要系数则不输出有效位。排序扫描 LIS 中的节点, 检查 Surfacelet 系数树像素集合的重要性。以 A 类型、B 类型进行分类处理^[8]。细化扫描 LSP 中的节点, 除最后一次分类过程中所得到的重要系数外, 输出其他重要系数在对应平面上的码流。对排序扫描和细化扫描进行迭代操作、量化更新, 直到达到预定的压缩比或者失真目标为止。

SPIHT 算法计算复杂度较低,位速率容易控制^[9]。该算法编译过程简单,不用进行运动补偿和估计操作,自动完成嵌入式编码码流的分配,利用 Surfacelet 变换的分解系数在各层间相关和图像能量集中的特性,SPIHT 算法将根结点、子结点根据同一方向划为同一集合,利用集合分割排序策略来计算阈值、初始化有序表、排序扫描、细化扫描、迭代更新等。接收嵌入位流在任意点中断时,都可解压重建视频图像,提供分辨率、时域和质量上的伸缩性、渐进传输性^[10],同时可获得较高的峰值信噪比(Peak Signal to Noise Ratio,PSNR)。

4 视频压缩编码实验

4.1 基本框架

本文提出基于 Surfacelet 变换并结合 SPIHT 算法的视频压缩编码。视频信号帧内图像的相邻像素存在相关性,相邻帧之间在时间维方向也存在相关性^[11]。把视频信号作为特殊的三维信号,利用 Surfacelet 变换对三维体数据的空间和时间信息进行整体处理。图 3 是基于 Surfacelet 变换和 SPIHT 算法的视频压缩编码的基本框架。

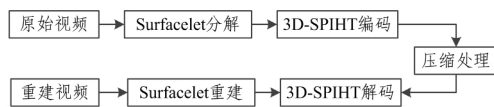


图 3 视频压缩基本框架图

以 QCIF 格式的标准视频序列进行实验,一组输入视频序列,帧数越多,压缩编码的效果越好。但是过大的矩阵数据输入会导致编码运行时间成倍增加,尤其在 SPIHT 编码部分,逐次循环扫描会耗费大量时间。综合 Surfacelet 变换编码效果和运行时间最优化,把 176×144 大小 QCIF 格式的 YUV 视频序列以 32 帧为一组进行划分。

把每组内的视频段作为一个三维体数据进行 n 级 Surfacelet 变换分解,即直接对视频组进行空间和时间轴的分解。依次迭代处理后,变换域中所有系数在时空方向上构成一个金字塔型结构。经过分解得到不同子带的 Surfacelet 系数。不同子带具有不同的特性^[12]。低频子带系数幅值和能量较大,高频子带系数幅值和能量较小,而且各方向系数差距大。利用 Surfacelet 变换后能量集中的特性,将分解系数通过对多方向的系数矩阵合并,转化成类似小波变换得到的树形结构^[13]。采用 SPIHT 编码,通过舍弃高频子带中幅值较小的不重要系数,对幅值较大的系数分层量化编码,从而消除冗余,实现压缩。

对视频序列 Y, U, V 3 个分量分别统计 PSNR,可知 Y, U, V 分量的 PSNR 主要受到亮度分量(PSNR Y)的影响。从生物视觉角度来讲,人眼视觉系统对图像亮度的变化更敏感^[14],因此后面的实验中采用 Y 分量的 PSNR 大小衡量视频序列的 PSNR 值。实验结果比较选定视频序列一组的中间几帧作为分析参考帧,以减少其他因素干扰。

4.2 与 3D-DWT 压缩方法比较

Surfacelet 变换作为改进版的三维小波变换(3D-DWT),与 3D-DWT 冗余度相近,都对三维信号具有方向多尺度分解特性^[15]。作为对比,取同一视频序列 Suzie,在结合相同 SPIHT 算法编码的情况下,对 Surfacelet 变换和 3D-DWT 不同压缩比下的 PSNR 大小进行比对。

4.3 不同视频压缩效果比较

取 3 种不同的 QCIF 视频序列,包括 Suzie, Mobile 和 Foreman,利用 Surfacelet 变换结合 SPIHT 算法在同一压缩比下对 PSNR 大小进行比对。

5 实验结果分析

5.1 与 3D-DWT 压缩方法比较

表 1 列出视频序列 Suzie 基于 Surfacelet 变换和 3D-DWT 进行视频压缩编码处理,在不同压缩比情况下的 PSNR 数值。

表 1 视频序列 Suzie 的 PSNR 比对

变换\压缩比	5	15	30	50
Surfacelet 变换	36.5195	33.1934	30.4896	25.9812
3D-DWT	30.65	25.43	12.78	9.75

在 PSNR 上,Surfacelet 变换比 3D-DWT 方法能达到更高的数值,视频序列 Suzie 基于 Surfacelet 变换处理的 PSNR 值比基于 3D-DWT 处理后的数值高出 5dB 以上,并且压缩比越高,PSNR 的差距越明显。

在视觉效果上,Surfacelet 变换明显优于 3D-DWT。压缩 30 倍时,Surfacelet 变换仍能较好地恢复图像信息,较好地保留低频,恢复高频,而 3D-DWT 重建图已经完全模糊,还有拖影分块现象,效果非常差,如图 4 所示。



(a) 未压缩 (b) 基于 Surfacelet 变换 (c) 基于 3D-DWT

图 4 Suzie 重建图(压缩比为 30)

Surfacelet 变换相比 3D-DWT 能分出更多的高频子带,大部分方向系数的幅值接近于零,利用 SPIHT 编码在相同压缩比下能保留更多相对重要的高频子带大系数的优势,更好地重建图像信息。

5.2 不同视频压缩效果比较

表 2 为 Suzie, Mobile, Foreman 3 种视频序列在压缩比为 30 的情况下中间几帧的 PSNR 数值比较。

表 2 3 种不同视频序列的 PSNR 比对

视频序列\帧数	5	6	7	8
Suzie	30.4896	30.6829	30.7219	30.6794
Mobile	20.0147	20.0361	20.0058	19.9594
Foreman	24.1616	24.4974	24.6856	24.6770

在 PSNR 上,由表 2 可见 Suzie 为 30 左右,Foreman 为 24 左右,Mobile 只有 20 左右。本文算法在处理 3 种不同视频序列时 PSNR 值差异较大,Suzie 最高,Foreman 最低。



(a) Suzie 原图 (b) Mobile 原图 (c) Foreman 原图

图 5 3 个视频序列未压缩原图

(下转第 267 页)

参考文献

- [1] Goldberg R P. Survey of Virtual Machine Research [J]. Computer, 1974, 7(6):34-45
- [2] Govindan S, Nath A R, Das A, et al. Xen and co.: communication-aware CPU scheduling for consolidated xen-based hosting-platforms[C]//VEE. 2007:126-136
- [3] Ongaro D, Cox A L, Rixner S, et al. Scheduling I/O in virtual machine monitors[C]//Virtual Execution Environments. 2008
- [4] 王凯,侯紫峰. 自适应调整虚拟机权重参数的调度方法[J]. 计算机研究与发展, 2011, 48(11):2094-2102
- [5] 暴锡文. 一种云计算环境下基于 Xen 的虚拟机调度机制[J]. 计算机测量与控制, 2014, 22(10):3381-3384

- [6] 张莹,李华. 一种虚拟机负载均衡调度算法[J]. 河南科技, 2015 (12):30-33
- [7] 陈锐忠,齐德昱,林伟伟,等. 一种面向非对称多核处理器的虚拟机集成调度算法[J]. 计算机学报, 2014, 37(7):1466-1477
- [8] 陈锐忠. 非对称多核处理器的若干调度问题研究[D]. 广州:华南理工大学, 2013
- [9] 常建忠. VCPU 组调度技术的研究与实现[D]. 长沙:国防科学技术大学, 2010
- [10] 王凯,侯紫峰. Xen 虚拟 CPU 空闲调度算法[J]. 计算机研究与发展, 2013, 50(11):2429-2435
- [11] 姚文斌,郑兴杰. 一种改进的 SEDF 调度算法[J]. 小型微型计算机系统, 2010, 31(3):446-450

(上接第 239 页)



(a) Suize 重建图 (b) Mobile 重建图 (c) Foreman 重建图

图 6 3 个视频序列重建图(压缩比为 30)

在视觉效果上,观察图 5 和图 6 可知, Suize, Mobile, Foreman 3 个视频重建图整体视觉良好,在纹理细节部分有点模糊,如 Suize 的头发、Mobile 的日历数字、Foreman 的眼睛部分。

3 个视频序列各有特点: Suize 纹理比较丰富,包含头发、皮肤等细节,运动幅度较小; Mobile 纹理特别复杂,画面中有多个运动物体,运动轨迹平缓; Foreman 纹理丰富度一般,画面中人物运动较多且有场景切换。根据 3 种视频序列的纹理复杂度和运动幅度特性分析可知, Surfacelet 变换和 SPIHT 算法编码相结合适用于纹理复杂度较高、运动不剧烈的视频。实际播放完整的视频序列时,视觉效果更加突出。

结束语 本文把视频信号作为特殊的三维信号,提出基于 Surfacelet 变换并结合 SPIHT 算法编码的视频压缩编码方法。SPIHT 算法自动完成嵌入式编码码流的分配,利用 Surfacelet 变换的分解系数在各层间相关和图像能量集成的特性,能获得较好的视频压缩效果。

实验证明, Surfacelet 变换结合 SPIHT 算法的视频压缩编码方法比 3D-DWT 结合 SPIHT 算法的方法实现了更高的 PSNR 和更好的视觉效果,且较适用于纹理复杂度较高、运动幅度较小的视频。

对于空间复杂度较高的视频序列,其图像的中频和高频 Surfacelet 系数成分中不重要系数的节点形成的零树较少,导致 SPIHT 算法对高频系数的压缩性能不如对低频系数的压缩性能, SPIHT 算法编码效率会受到影响。因此可以采取 Surfacelet 变换结合多种算法的方法,对高频系数用自适应编码,对低频系数用 SPIHT 算法编码,以改进视频压缩效果。在实际应用中,还要考虑到算法的运算效率。视频三维编码需要较多的帧存储器处理,会带来较大的延迟,这还有待提高。

参考文献

- [1] Chaudhury K N, Unser M. Construction of Hilbert transform pairs of wavelet bases and optimal time-frequency localization

- [J]. IEEE Transactions on Signal Processing, 2009, 57(9): 3411-3425
- [2] Unser M, Blu T. Mathematical properties of the JPEG2000 wavelet filters [J]. IEEE Transactions on Image Processing, 2003, 12(9):1080-1090
- [3] 冯鹏. 高分辨图像处理用抗混叠 Contourlet 变换的若干关键问题研究[D]. 重庆:重庆大学, 2007
- [4] Zhang J N, Gilling C J, Kreger K S. System and method for filtering frequency encoded imaging signals; US, US6633162 [P]. 2003
- [5] 江山,尹忠科,陈帆. 基于 Surfacelet 稀疏重构的视频修复[J]. 数据采集与处理, 2012, 27(4):444-449
- [6] Lu Y M, Do M N. Multidimensional directional filter banks and surfacelets. [J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2007, 16(4):918-931
- [7] Wang H, Pan X. Video compression coding based on the improved 3D-SPIHT[C]//2010 International Conference on Computer Application and System Modeling (ICCASM). IEEE, 2010: V14-108-V14-111
- [8] Tang G, Gu G. LIS-No-Classification Wavelet Image Coding Algorithm Based on Lifting Scheme[C]//International Workshop on Intelligent Systems and Applications. IEEE, 2009:1-4
- [9] Fradj B B, Zaid A O. Scalable video coding using motion-compensated temporal filtering and intra-band wavelet based compression[C]//European Workshop on Visual Information Processing. 2011:50-55
- [10] Lian S. Secure service convergence based on scalable media coding[J]. Telecommunication Systems, 2010, 45(1):21-35
- [11] Jeon B M, Park S W, Park J H. Method and apparatus for encoding/decoding video signal using block prediction information; US, US8228984[P]. 2012
- [12] An Z Y, Liu P Q, Jiang H L. Dynamic Textures Retrieval Using the Integrated Wavelet-Based Surfacelet Transform[J]. Applied Mechanics & Materials, 2012, 263-266:227-230
- [13] 袁琴,吴宣够,熊焰. 小波树结构在贝叶斯压缩感知图像重构中的应用研究[J]. 计算机科学, 2014, 41(3):314-318
- [14] 汤敏,陈秀梅,陈峰. 基于 Contourlet 变换和 SPIHT 算法的彩色医学图像压缩[J]. 计算机科学, 2014, 41(1):303-306
- [15] Kaveh H, Moin M S, Razzazi F. A novel steganography approach for 3D polygonal meshes using Surfacelet Transform[C]//2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP). IEEE, 2013:304-309