软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2015,26(5):1213−1236 [doi: 10.13328/j.cnki.jos.004822]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

3D 多尺度几何分析研究进展*

宋传鸣[1,2], 赵长伟[1], 刘 丹[1,3], 王相海[1,2]

[1](辽宁师范大学 计算机与信息技术学院,辽宁 大连 116081)

[2](计算机软件新技术国家重点实验室(南京大学),江苏 南京 210093)

[3](大连理工大学 信息与通信工程学院,辽宁 大连 116024)

通讯作者: 王相海, E-mail: xhwang@lnnu.edu.cn, http://mtl.lnnu.edu.cn/member/wangxianghai/homepage.html

摘 要: 3D 多尺度几何分析是图像、视频和几何模型等数字可视媒体处理的技术基础,其目的在于高效地表示这些媒体中存在的点、线、面奇异.为此,依据不同变换捕获奇异的能力演进及其非线性逼近效率的提高,从 2D 图像多尺度几何分析的研究进展切入,着重阐述视频 3D 多尺度几何分析的发展,并将其归纳为 3 类:由 2D 基函数直接扩展的 3D 多尺度几何分析、基于 3D 基函数的 3D 多尺度几何分析和基于时空非局部相关性的 3D 多尺度几何分析,深入探讨了各种典型变换方法的主要思想、非线性逼近效率、计算复杂度、优势和不足.同时,概要介绍了数字几何模型的 3D 多尺度几何分析研究进展.在此基础上,对 3D 多尺度几何分析的未来发展趋势进行了展望.

关键词: 多尺度几何分析;稀疏表示;综述;视频;图像;几何模型

中图法分类号: TP391

中文引用格式: 宋传鸣,赵长伟,刘丹,王相海.3D 多尺度几何分析研究进展.软件学报,2015,26(5):1213−1236. http://www.jos.org.cn/1000-9825/4822.htm

英文引用格式: Song CM, Zhao CW, Liu D, Wang XH. Advances in three-dimensional multiscale geometrical analysis. Ruan Jian Xue Bao/Journal of Software, 2015,26(5):1213−1236 (in Chinese). http://www.jos.org.cn/1000-9825/4822.htm

Advances in Three-Dimensional Multiscale Geometrical Analysis

SONG Chuan-Ming[1,2], ZHAO Chang-Wei[1], LIU Dan[1,3], WANG Xiang-Hai[1,2]

[1](School of Computer and Information Technology, Liaoning Normal University, Dalian 116081, China)

[2](State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210093, China)

[3](School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: Three-Dimensional (3D) multiscale geometrical analysis is the technological fundamental for the processing of digital visual media, such as images, videos, and geometrical models. Its objective is to efficiently represent the point singularity, curve singularity, as well as surface singularity presented in those visual media. This study first reviews the research advances in two-dimensional (2D) multiscale geometrical analysis. It then elaborates on the development of 3D multiscale geometrical analysis for video according to the capability evolution in capturing singularity and nonlinear approximation efficiency improvement of various transforms. State-of-the-Art 3D multiscale geometrical analysis is classified into three categories: the extended multiscale geometrical analysis from 2D basis functions, the multiscale geometrical analysis based on 3D basis function, and the multiscale geometrical analysis based on spatiotemporal non-local correlation. The basic ideas of typical transforms are thoroughly discussed subsequently, and so are their nonlinear approximation efficiency, computational complexity, advantages, and disadvantages. Meanwhile, this study also presents a general review

* 基金项目: 国家自然科学基金(61402214, 41271422); 教育部高等学校博士学科点专项科研基金(20132136110002); 辽宁省博士科研启动基金(20121076); 辽宁省教育厅科学研究一般项目(L2011192, L2013406); 大连市科学技术基金计划(2012J21DW008, 2013J21DW027); 计算机软件新技术国家重点实验室开放基金(KFKT2011B09, KFKT2010B11); 江苏省图像处理与图像通信重点实验室开放基金(LBEK2011001);

收稿时间: 2014-09-05; 修改时间: 2014-12-22; 定稿时间: 2015-01-21

© 中国科学院软件研究所 http://www.jos.org.cn

on the development of the 3D multiscale geometrical analysis for geometrical models. Based on the study above, the development trend of the 3D multiscale geometrical analysis is forecast in the near future.

Key words: multiscale geometrical analysis; sparse representation; survey; video; image; geometrical model

图像、视频和几何图形等可视媒体是人类接受信息的最主要途径,人类感知系统所获取的信息中大约有80%~85%是视觉信息[1].作为大数据的重要组成部分,可视媒体数据量逐年递增,人们对可视媒体产品的消费数量和消费期望远大于以往任何时期.以视频为例,在网络领域,以 Youtube.com 为代表的在线视频服务风靡全球;在广电领域,720p,1080i 和 1080p 这 3 种格式的高清、超高清电视已成为未来数字电视的必然走向.IMS Research 的研究结果显示,3D 电视也将成为媒体产业下一个里程碑式的内容终端.而随着移动终端的普及和微博、微信等社交网络的发展,Cisco 公司 2014 年 2 月发布的《视觉化网络指数报告》中预测:在 2013~2018 年间,移动视频数据量将增长 14 倍,2018 年视频数据量将占全球移动数据总量的 2/3.因此,无论是在研究领域还是在应用领域,可视媒体的处理,如图像和视频的获取、表示、去噪、分割、特征提取、融合、压缩、复原和超分辨率以及 3D 几何造型的编辑和检索等,在今后若干年都将是多媒体技术领域的一个重要课题.

可视媒体数据的紧致表示(compact representation)是上述诸多处理的基础,所谓紧致表示,是指用较少的数学描述来捕获图像、视频和图形中的重要信息[2].基于调和分析的稀疏表示是紧致表示的方法之一[3]:一方面,可视媒体处理的发展给调和分析提出了挑战;另一方面,调和分析理论的完善与进展则为可视媒体处理提供了更加完备、有效的工具.典型地,小波变换在表示 1D 信号的点奇异性方面具有最优的逼近性能,成为近 10 年来自然信号处理和分析的强有力工具,并应用在去噪、特征提取、压缩等众多领域,然而对于 2D 图像或者更高维的数据,小波变换的稀疏表示能力则比较有限.2D 可分离小波基函数的支撑区间为一个正方形,且各向同性,只能表示 2D 或高维信号在竖直、水平和对角线方向的直线奇异性.自然界中多数物体所具有的平滑边缘使得自然图像的奇异点往往不是独立分布的,而是聚集成具有某些几何特征的奇异曲线.在这种情况下,小波的各向同性的基函数不能有效地利用数据本身特有的几何特征来捕获沿着边缘方向的曲线奇异,无法为图像及高维数据提供最优的或者最稀疏的表示.这样,处于数学分析、计算机视觉、模式识别、统计分析等多个领域的很多专家在小波变换的理论基础上分别独立地发展着一种彼此极其相似的理论,称为多尺度几何分析(multiscale geometrical analysis,简称 MGA)[4,5],以检测、表示、处理某些高维数据,并为这些数据提供多尺度的、多方向的表示.现有的多尺度几何分析工具主要包括脊波(ridgelet)变换[6,7]、曲波 Curvelet 变换[8,9]、条带波(bandelets)变换[10,11]、小线(beamlets)变换[12]、Directionlets 变换[13]、轮廓波(contourlet)变换[14]、剪切波(shearlet)变换[15]、Surfacelets 变换[16],同时,Cortex 变换[17]、Steerable Pyramid 变换[18]、2-D 方向小波变换、Brushlets 变换[19]和复数小波变换[20]也是多尺度、多方向的变换.这些多尺度变换具有很好的图像各向异性特征,成为处理 2D 图像数据的更优算法,很好地解决了小波变换不能有效表示边缘、轮廓等线奇异的问题.文献[4,5]对图像多尺度几何分析方法的研究进展进行了综述.有关这些多尺度几何变换的详细理论分析可参见文献[2].

上述的多尺度几何分析工具都只适用于 2D 静态图像,若将其应用在 3D 数据的处理中,则需对这些变换向 3D 方向进行扩展.本文约定,将这类变换称为 3D 多尺度几何分析.3D 多尺度几何分析既可以用来处理视频数据,又可以应用到诸如核磁共振、计算机断层扫描等医学图像、高光谱遥感影像等专门图像以及几何造型、几何图组的处理中.考虑到视频与医学图像、遥感影像的数据分布存在相似性,本文主要以视频处理为例展开讨论,再简要论述用于 3D 数字几何模型的多尺度几何分析.

3D 多尺度几何分析是近年来多尺度几何分析领域的研究热点之一.本文以 3D 多尺度几何分析的研究进展为主线,将现有方法分为由 2D 基函数直接扩展的 3D 多尺度几何分析、基于 3D 基函数的 3D 多尺度几何分析和基于时空非局部相关性的 3D 多尺度几何分析这 3 类,同时,对各类算法的优势和存在的不足进行分析和讨论,最后,为 3D 多尺度几何分析的未来发展进行了展望.

© 中国科学院软件研究所 http://www.jos.org.cn

1 2D 多尺度几何分析概述

多尺度几何分析研究之初主要讨论 2D 数据的处理和分析,它是 3D 多尺度几何分析的基础,故而本节首先对典型的 2D 多尺度几何分析进行概述.

1D 小波在信号处理中的成功,促使研究人员将小波引入到 2D 数据的处理和分析中.典型的 2D 小波基是由两个 1D 小波基通过张量积扩展而成的.一次 2D 小波变换需要分别沿水平和竖直方向进行一次 1D 变换,这样的变换过程使得 2D 可分离小波基只沿着水平和竖直方向才有消失矩,即,它只能稀疏表示点奇异以及沿水平、竖直和对角线方向的直线奇异.而 2D 数据,尤其是数字图像中还含有大量的曲线奇异,理论研究结果表明:为了表示一条具有曲线奇异的边缘,2D 小波基产生的逼近误差的衰减速度仅为 $O(M^{-1})$[4,5].为了实现高效的逼近性能,Do 指出,一种真正的 2D 图像表示工具应具有 5 项特征[21]:多分辨率特性、时空局部性、严格下采样、多方向性和各向异性.然而,小波变换只具备前 3 个特性.为了实现真正的变换,众多领域的专家陆续提出了多种多尺度几何分析的理论与方法.

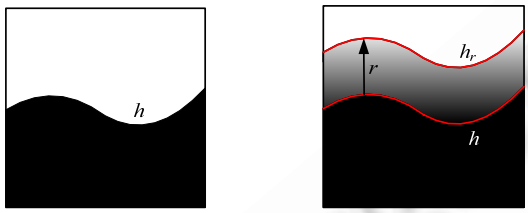
第 1 类开创性工作主要由 Candès 等人开拓,其基本思想是:考虑到小波能有效表示信号的点奇异,可先将线奇异分解为点奇异,再由小波变换进行处理.Candès 通过 Radon 变换将直线奇异映射为 Radon 域上的点奇异,进而提出了脊波变换[6].Donoho 等人则构造了正交脊波以及相应的脊波变换,可用来捕获分片光滑函数的直线奇异[22].借鉴微积分学中以短直线趋近曲线的思想,Candès 等人进一步把滤波和多尺度脊波变换相结合,提出了曲波变换,利用平滑分割将曲线奇异分解为直线奇异,再利用局部的脊波变换进行处理[8,9].由于曲波基函数不具备多分辨率特性,其离散化实现非常困难,于是出现了两个方面的典型工作:一方面,Candès 等人提出了快速的离散曲波变换方法[23];另一方面,受到信号处理领域出现的一系列方向滤波思想[24,25]的启发,Do 等人提出了一种金字塔型方向滤波器组(pyramidal directional filter band),亦称轮廓波变换,首先利用拉普拉斯金字塔变换捕获图像的点奇异,再用方向滤波器组进一步将分布在同方向上的点奇异合并表示.轮廓波变换继承了曲波变换的各向异性尺度关系(anisotropy scaling relation),并具有与后者相同的频带划分能力.文献[26]则实现了对轮廓波变换的加速.

第 2 类工作主要由 Velisavljevic 等人开展,其基本思想是:对 2D 图像数据或基函数进行平移、缩放和剪切等几何变换,从而使基函数能够沿着曲线方向捕获奇异点.Velisavljevic 等人首先选取两个线性无关的整数向量,其中,第 1 个整数向量的方向称为变换方向,第 2 个整数向量的方向称为队列方向.变换方向和队列方向可确定一个像素子集及其陪集,直观上相当于沿着某一特定方向对 2D 图像进行采样;其次,沿着变换方向和队列方向施以各向异性的斜小波变换.上述这种基于整数格理论的变换称为方向波变换[13],它可以稀疏表示位于任意两个有理斜率方向间的曲线奇异.类似地,Easley 等人则利用合成小波理论定义了尺度矩阵和剪切矩阵,并由此对小波基函数进行合成膨胀和仿射变换,所得基函数的集合构成了可有效表示 2D 图像数据的 Parseval 框架.该基函数称为剪切波,相应的变换就称为剪切波变换[15,27].2010 年,Lim 进一步给出了剪切波的离散实现[28].

第 3 类工作主要由 Donoho 等人提出.与前两类方法不同,该类变换的主要思想是,采用自适应的基函数来稀疏地表示图像.Donoho 将一族定义在不同位置、尺度和方向的二进楔形区域的函数作为过完备基函数字典,其中的每个基函数可简洁表示区域内的一条直边缘;再用基于率失真约束的四叉树分解把图像自适应地分割为不连续的子区域;最后,选择那些能够获得紧致表示的基函数来逼近每个子区域上的像素值.这种变换称为楔形波(wedgelet)变换[29].然而,该变换仅能有效地表示图像中的阶跃信号,却难以处理图像中大量包含的像素值渐变.为此,Lisowska 采用水平函数表示曲线奇异,引入模糊函数表示沿着奇异曲线法向的图像渐变(如图 1 所示),进而利用剪枝算法从字典中选取合适的基函数,提出了一种平滑波(smoothlets)变换[30].除了设计自适应逼近函数以外,Pennec 等人则采用沿着奇异曲线方向做弯曲小波变换的方法实现自适应分解,提出了一种条带波变换[10,11],其基本思路是:将图像进行二进迭代剖分,使得每个剖分区间中最多只包含一条轮廓线;对于包含轮廓线的每个区域,将其中所有像素沿着该轮廓的切线方向进行小波变换,进而达到稀疏表示曲线奇异的目的.Mallat 将条带波的思想进一步发展,提出了群波(grouplet)变换[31],它利用块匹配方法为每个像素计算出一个关联域(association field),即正则方向,再沿着该方向进行加权的提升方案哈尔(Haar)小波变换.与条带波相比,群

© 中国科学院软件研究所 http://www.jos.org.cn

波变换更加灵活,滤波过程不必在两个相邻像素间进行,滤波方向更加丰富,而且计算量低.需要特别指出:楔形波、条带波和群波的基函数需要根据目标图像的特性自适应地选取,并以辅助信息(亦称“边信息”)的形式存储起来.这意味着若将三者应用在视频编码中会产生额外的数据,在一定程度上影响编码效率.



(a) 水平函数示意图 (b) 模糊函数示意图

Fig.1 Basis function diagram of smoothlets transform[30]

图 1 平滑波变换基函数示意图[30]

上述各种多尺度几何分析的本质目的都是为了有效地捕获曲线奇异,然而其逼近效率却不尽相同.下面从非线性逼近性能的角度来简要比较各种变换.所谓的非线性逼近,是指对一个信号 f 进行某种多尺度变换后,选取绝对值最大的前 M 个分量所获得的重构信号 f_M.原始信号 f 与重构信号 f_M 的差 $\varepsilon_M=\|f-f_M\|^2$ 称为非线性逼近的重构误差,其中,ε_M 是 M 的函数.随着 M 的增大,ε_M 不断减小,本文称其衰减速度为非线性逼近效率.上文各种典型变换用于逼近二次连续可微(C^2)分片光滑函数的非线性逼近效率见表 1,从中可见,条带波、群波和平滑波的非线性逼近效率最高,标准小波变换和傅里叶变换的效率最低,其余变换介于两者之间.

Table 1 Nonlinear approximation efficiency comparison among different transforms

表 1 不同变换的非线性逼近效率比较

变换名称	适合捕获的奇异特征	逼近 C^2 曲线奇异的效率
傅里叶变换	连续周期信号	$O(M^{-1/2})$[15]
标准 2D 小波变换	点奇异	$O(M^{-1})$[15]
脊波变换	直线奇异	$O(M^{-3/2})$[4]
曲波变换	C^2 连续的曲线奇异	$O(M^{-2}\log_2^3 M)$[15]
轮廓波变换	C^2 连续的曲线奇异	$O(M^{-2}\log_2^3 M)$
方向波变换	C^2 连续的曲线奇异	$O(M^{-1.562})$[13]
剪切波变换	C^2 连续的曲线奇异	$O(M^{-2}\log_2^3 M)$[27]
楔形波变换	C^2 连续的曲线奇异	$O(M^{-2})$[13,29]
条带波变换	C^α 连续的曲线奇异($\alpha\geqslant 2$)	$O(M^{-\alpha})$[32]
平滑波变换	C^α 连续的曲线奇异($\alpha\geqslant 2$)	$O(M^{-3})$[30]
群波变换	C^α 连续的曲线奇异($\alpha\geqslant 2$)	$O(M^{-\alpha})$[31]

尽管目前的2D多尺度分析工具已经达到了较为理想的非线性逼近效率,但是其计算复杂度却较高,尤其是第三类的自适应变换.比如,平滑波变换一幅 $n\times n$ 像素的图像需要 $O(n^4)$次操作.故此,具有方向自适应特性的小波变换同样受到了研究者们的关注,出现了一系列以提升小波变换(lifting wavelet transform)[33]为基础框架的方向小波变换.Taubman[34]提出,在提升小波的预测阶段,可沿着 0°,±30°和±45°等方向中选取最优方向完成预测;Gerek 等人[35]设计了沿着 0°,45°和 135°进行滤波的方向自适应的类 Daubechies 5/3 小波.但是这两种变换的预测方向有限,仍无法满足表示图像中复杂纹理的需要.这样,文献[36,37]提出了一种能够自适应地调整滤波方向的小波变换方案,其预测过程是依照局部图像块中最高像素相关度的方向进行的,因而可以更好地适应图像的任意方向特征,变换本身的能量聚集性明显提高;考虑到文献[36,37]在多个候选方向间动态选取最优滤波方向仍需要较高的计算量,文献[38]提出利用方向滤波器进行预滤波,排除那些不可能的候选方向,从而缩小了后续变换的方向搜索范围,降低其计算复杂度.此外,文献[39]探讨了根据图像局部特性确定滤波器阶数的方法,文献[40]讨论了自适应滤波器的设计方法,文献[41]研究了具有方向特性的离散余弦变换.

© 中国科学院软件研究所 http://www.jos.org.cn

虽然方向自适应小波变换在计算量方面具有一定的优势,但其稀疏表示能力仍低于典型多尺度几何变换,而且就我们所掌握的文献来看,目前很少见到关于方向自适应小波变换的非线性逼近效率的严格证明,尚需进一步的理论研究.

2 由 2D 基函数直接扩展的 3D 多尺度几何分析

3D 多尺度几何分析是处理和分析海量视频数据的基础.与 2D 多尺度几何分析一样,3D 多尺度几何分析也经历了不同的发展阶段,并且出现了多种变换工具.本文将现有 3D 多尺度几何分析方法的研究进展分为由 2D 变换通过张量积扩展的 3D 多尺度几何分析、基于 3D 基函数的多尺度几何分析和基于时空非局部性的 3D 多尺度几何分析这 3 个阶段.其中,第 1 阶段的发展相对成熟,而后两个阶段的发展则刚刚起步.下文对各个阶段展开详细阐述.

2.1 不包含运动估计的扩展2D多尺度几何分析

不进行运动估计/补偿的扩展 2D 多尺度几何分析可以追溯到 1987 年,这是最早出现的一类 3D 数据分析方法,其基函数由 1D 小波基通过张量积构造而成[42],基本思想类似于由 1D 小波向 2D 小波的扩展,即,分别沿水平、竖直和时间轴方向做一次 1D 小波分解,从而建立 3D 空间的可分离正交小波基.

根据 2D 小波和时间维滤波的先后顺序,可将该类 3D 多尺度分析划分为(2D+t)结构和(t+2D)结构,前者先做 2D 变换再进行时间维滤波,后者则与之相反.典型的(2D+t)结构的 3D 多尺度变换如文献[43,44],其优势表现在易于产生具有不同空间分辨率的频带划分.但是,由于低空间分辨率下相邻视频帧间的时间相关性会降低,不利于有效表示沿时域方向的奇异,因而适用于运动量低、空间复杂度较高的视频序列.典型的(t+2D)结构的 3D 多尺度变换则包括文献[45,46],该结构有利于捕获时域奇异,但是由于现实的滤波器并非理想的带通滤波器,不同的时间子带在频域上往往存在混叠,会出现能量在低空间分辨率子带和高空间分辨率子带之间相互泄露的情况以及鬼影效应(ghosting artifact)(如图 2(a)所示),适用于运动相关性高、空间纹理简单的视频序列.因此,文献[47,48]认为应根据视频数据分布特点自适应地确定 2D 变换和时间维滤波的顺序,提出了(2D+t+2D)结构的 3D 多尺度变换,允许用户对时间维滤波和 2D 空域小波的变换顺序和分解层数进行选择,即,首先对原始视频进行 P 级空间域小波变换,得到(P+1)个具有不同空间分辨率的序列,然后对各个空间分辨率序列进行 T 级时间维滤波,最后再对最低空间分辨率的各时间域子带进行(S−P)级空间域小波变换.这样,可以得到经过 T 级时间域变换和 S 级空间域变换的多个具有不同时间、空间分辨率的子带.合理选择 P,T 和 S,能够有效地缓解能量泄漏和鬼影效应.

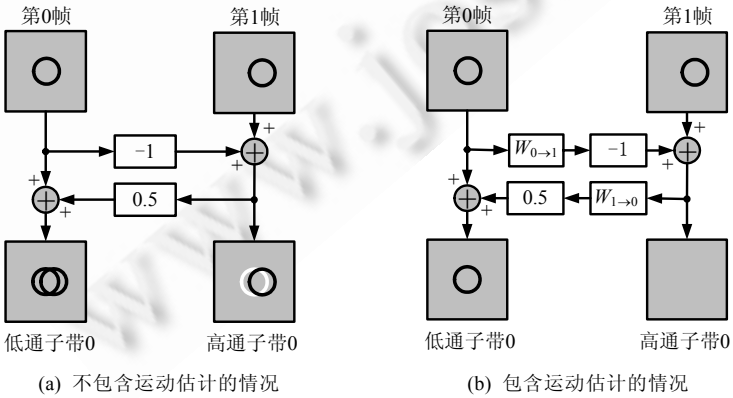


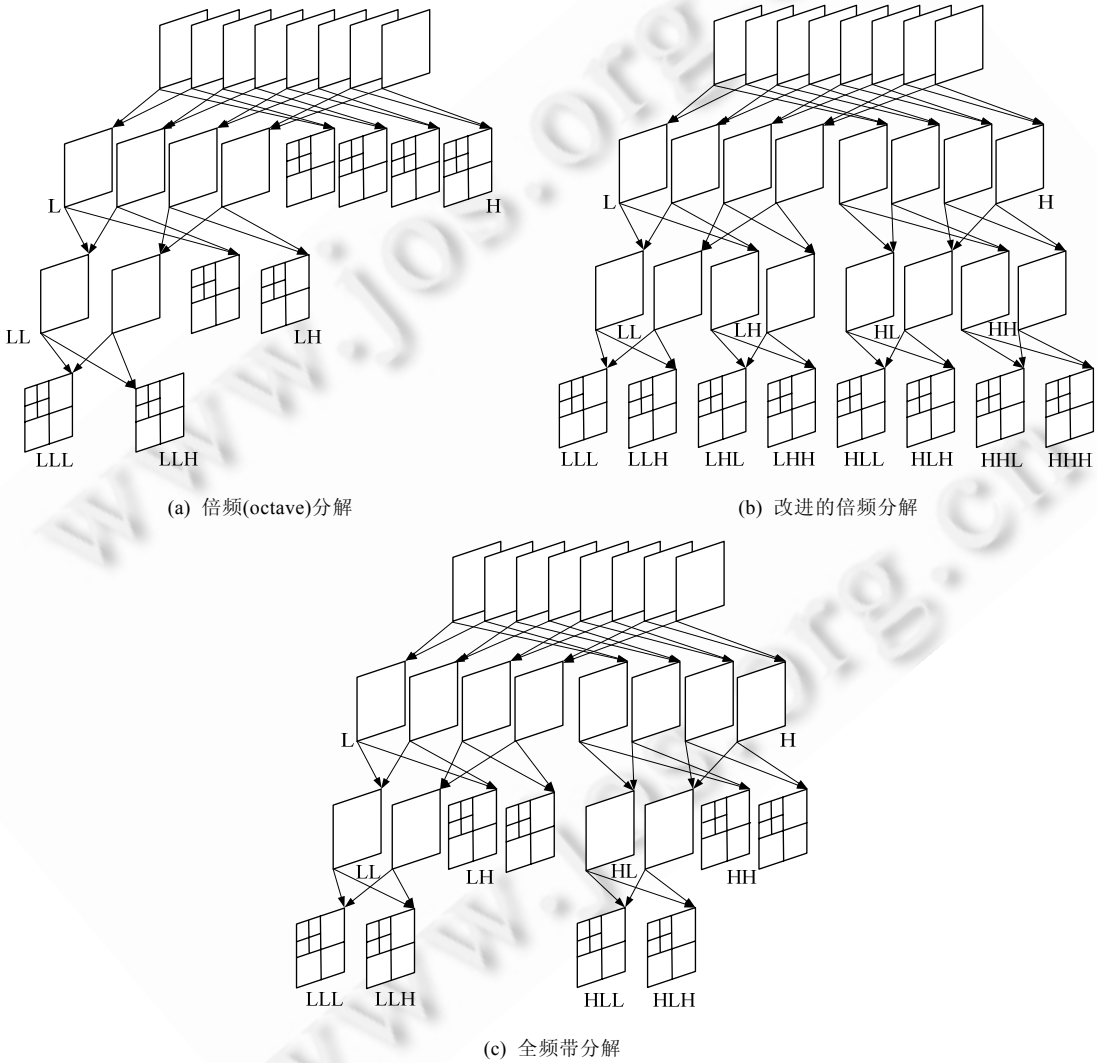
Fig.2 Diagram of ghosting artifacts[49]

图 2 鬼影效应示意图[49]

与 2D 变换相比,3D 多尺度几何分析的分解方式更为复杂多样,频带划分方式也更为灵活.根据时间维分解

© 中国科学院软件研究所 http://www.jos.org.cn

方式的不同,文献[50]探讨了倍频(octave)分解、改进的倍频分解和全频带(full-band)分解这 3 种频带分解方式.图 3 以(t+2D)结构的 3 级时间维滤波为例给出了频带划分情况,其中,倍频分解的计算量最低,但是仅能去除相邻帧间的时间冗余;改进的倍频分解的计算量最高,频带划分最细致,但是由于高频帧间的时间相关性随着分解级数的增加而快速降低,此时再在高频帧间进行时间滤波不仅不能实现稀疏表示,而且会产生幅值更大的高频系数;全频带分解只对时间维高频分量进行 2 级变换,对时间维低频分量则进行多级变换,能在稀疏表示能力和计算量之间实现很好的折中.后两者的频带划分方式类似于小波包分解,倘能依据视频序列的数据特点自适应地确定时间维滤波方式,则有利于实现最优的稀疏表示.需要指出的是,目前还缺乏关于最优分解方式的相关理论研究.



(a) 倍频(octave)分解

(b) 改进的倍频分解

(c) 全频带分解

Fig.3 Frequency band division diagram of three-level temporal filtering with (t+2D) structure

图 3 (t+2D)结构 3 级时间维滤波的频带划分示意图

若将视频视作一个定义在三维立方体上的像素集,那么上述三维小波基只能捕获该集合上所有的点奇异、沿水平、竖直和对角线(包括体对角线)方向的直线奇异以及沿立方体表面上各对角线所在平面的面奇异.时空体中的边缘、轮廓和纹理等信息所形成的几何特征往往表现为更加复杂的线奇异和面奇异,所以上述 3D 小波基有限的方向捕获能力导致其不能为视频数据提供最佳的稀疏表示,对于 C^2 连续的奇异曲面仅能达到

© 中国科学院软件研究所 http://www.jos.org.cn

$O(M^{-1/2})$的非线性逼近效率;同时,该类变换去除时间冗余的能力有限,对于中等以上运动复杂度、纹理复杂度也较高的视频序列,其高频子带中往往出现运动物体的边缘,稀疏表示能力明显下降[51].

2.2 包含运动估计的扩展2D多尺度几何分析

为了解决时间维滤波效率不高的问题,文献[52]提出一种可逆的帧弯曲(frame-warping)算子来缓解摄像机运动对去除时间维冗余的影响;文献[53]则提出在时间维滤波前利用 Mosaic 方法进行全局运动估计.但是,可逆算子和 Mosaic 技术均无法表示视频序列中普遍存在的局部运动.文献[50,54]进一步明确将运动估计/补偿引入到时间维滤波过程,其基本思路是,在时间维上沿着物体的运动轨迹进行 1D 小波分解.这种进行运动估计的 3D 小波变换又称为运动补偿的时域滤波(motion compensation temporal filtering,简称 MCTF).

依据 MCTF 过程所在域的不同,可将该类变换划分为基于空间域 MCTF(spatial domain MCTF,简称 SDMCTF)的 3D 多尺度分析和基于变换域 MCTF(in-band MCTF,简称 IBMCTF)的 3D 多尺度分析.本质上,前者相当于第 2.1 节的(t+2D)结构,后者则相当于(2D+t)结构.

对于第 1 类基于空间域 MCTF 的 3D 多尺度分析来讲,由于现实世界的视频序列中难以避免地包含复杂运动,导致当前帧和参考帧大约有 3~5%的像素不能通过运动向量一一对应;并且这一比例会随着分解级数的增加而显著提高,甚至高于 50%.于是,参考帧中有的像素可能会对应着当前帧的一个或多个像素,而有的像素则可能与当前帧的任何像素都不对应(如图 4(a)所示),前者称为连接像素(connected pixel),后者称为非连接像素(unconnected pixel).为此,文献[54]采用一种前向-后向同时块匹配(simultaneous forward-backward block matching)来降低非连接像素的比例,并提出将当前帧中非连接像素值直接插入到时间维低频帧,而将参考帧中非连接像素的后向块匹配误差的一半插入到时间维高频帧(如图 4(b)所示);文献[50]将文献[54]时间维滤波的精度拓展到了半像素精度.文献[55]则从降低运动估计运算量的角度采取了与文献[54]的频带划分和滤波方法截然相反的处理思路(如图 4(c)所示),并引入变块大小的运动估计技术(variable size block matching),提高了运动向量的精度以及时间维滤波的效率.然而,上述方法在滤波精度达到 1/2 像素时不能实现精确重构,即使不对变换系数进行任何有损处理,两级 MCTF 后的重构峰值信噪比也仅仅是 40dB~45dB.并且,由于哈尔小波基的消失矩小,不利于去除存在于物体运动路径上的多帧相关,其时间维滤波的效率不高.鉴于此,后续研究主要在 3 个方面进行改进:① 运动向量为分数像素精度下的 MCTF 精确重构;② 选用具有更高消失矩的小波基(如 Daubechies 5/3 和 9/7 小波基);③ 连接像素和非连接像素的处理.

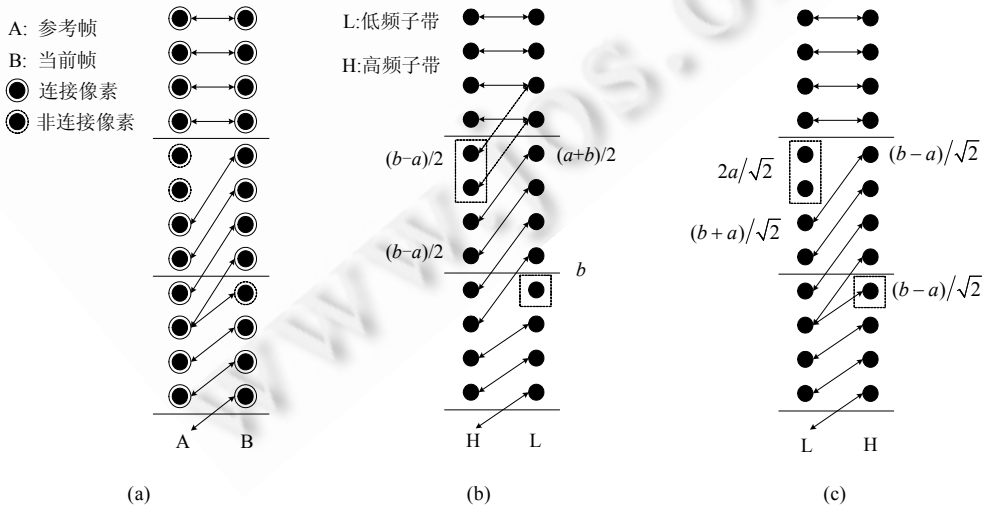


Fig.4 Connected pixels and unconnected pixels of motion compensation temporal filtering[54]

图 4 运动补偿的时间维滤波的连接像素与非连接像素[54]

© 中国科学院软件研究所 http://www.jos.org.cn

在第一方面,文献[56,57]通过将文献[54,55]中的哈尔小波变换改写为提升方案实现,修正了其滤波方式,达到了 1/2 像素滤波精度下的精确重构.文献[58,59]将文献[57]的精度又进一步提高至 1/4 像素和 1/8 像素.2003 年,Secker 的 3 项连续性工作[49,60,61]最终提出了基于提升方案和可逆运动映射算子的广义 MCTF 框架 LIMAT (lifting-based invertible motion adaptive transform).该框架具有很高的灵活性,不仅能够保证 MCTF 在分数像素下的精确重构,而且不依赖于所选的小波基,可逆运动映射算子又独立于所采用的运动模型(平移模型、仿射模型和变形网格模型等)和运动估计算法,这为后续算法的研究奠定了理论框架.在此基础上,文献[62]建立了不精确的运动向量对 LIMAT 的影响模型,其结论对于视频数据在时间维上的稀疏表示有一定的意义;文献[63]通过取消提升过程的更新(update)算子,提出了一种更为灵活的时间维滤波方法 UMCTF(unconstrained MCTF),既能根据视频内容的特点确定 MCTF 变换的级数,也可在不同的变换层次、不同帧间选择不同的小波基和分解结构.另外需要指出:虽然在上述工作中的 MCTF 能够达到很高的像素精度,但是高于 1/8 像素精度的 MCTF 的增益很小[64],实用性较差.

在第二方面,Luo 等人[65]在 2001 年首次将 5/3 小波引入 MCTF;同年,Secker 等人也在 LIMAT 方法中采用了 5/3 小波.文献[49,59,66]将基于 5/3 小波的 MCTF 应用到视频编码,并通过大量实验结果表明,该方法的稀疏表示能力明显优于基于哈尔小波的 MCTF.故此,后续的研究工作,如文献[49,56−63],大多数采用 5/3 小波作为 MCTF 的变换核.另外,有部分研究者也对基于 9/7 小波的 MCTF 进行了研究,如文献[46,67,68]等,但是由于其消失矩较大,以及相邻的多个视频帧之间的相关性会随着变换级数的增加而快速降低,反而不利于视频数据的稀疏表示,仅适用于纹理细节和物体运动较简单、变换级别较低的情况.

在第三方面,文献[46]在 2001 年首次提出了运动线索(motion thread)的概念,它是指由一串首尾相连的运动向量所链接起来的一系列像素(如图 5 所示).在此基础上,除了一一连接的普通像素以外,文献[46]将非连接像素作为一个运动线索的起始,而将与多连接像素(见图 5 中用五角星标出的像素)连接的 n 个像素中的(n−1)个作为其所在运动线索的终止.这样,非连接像素和多连接像素所在的运动线索可以与普通的运动线索同样处理.与文献[54,55]相比,文献[46]的时间维滤波更加统一、简洁,且每条运动线索的滤波可根据其特性选用哈尔、5/3 和 9/7 等不同小波基,灵活性更高,稀疏表示能力也明显提高.不过,非连接像素和多连接像素确实增加了运动线索的数量(尤其是在 MCTF 级数较高时),而运动线索两端的滤波效率并不高,造成总体滤波效率受到影响.

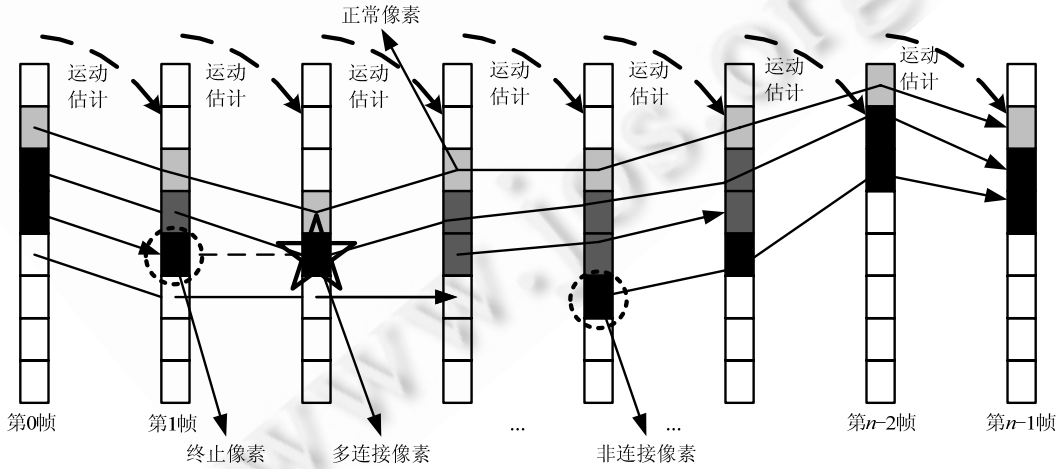


Fig.5 Example of motion thread[46]

图 5 运动线索示例[46]

文献[68]对这一不足进行了改进,其基本思路是:对于非自然终止的运动线索终点(也就是指向多连接像素的像素,如图 6 第 1 帧中被圈出的像素所示),用其运动向量所指向的前后两帧中的像素进行预测;对于非自然开始的运动线索起点(也就是非连接像素,如图 6 第 5 帧中被圈出的像素所示),则用其上方像素运动向量的反向向

© 中国科学院软件研究所 http://www.jos.org.cn

量所指向的像素及其自身向量所指向的后一帧像素共同进行预测.与此同时,文献[69]则提出通过自适应调整更新算子的参数来处理非连接像素和多连接像素.

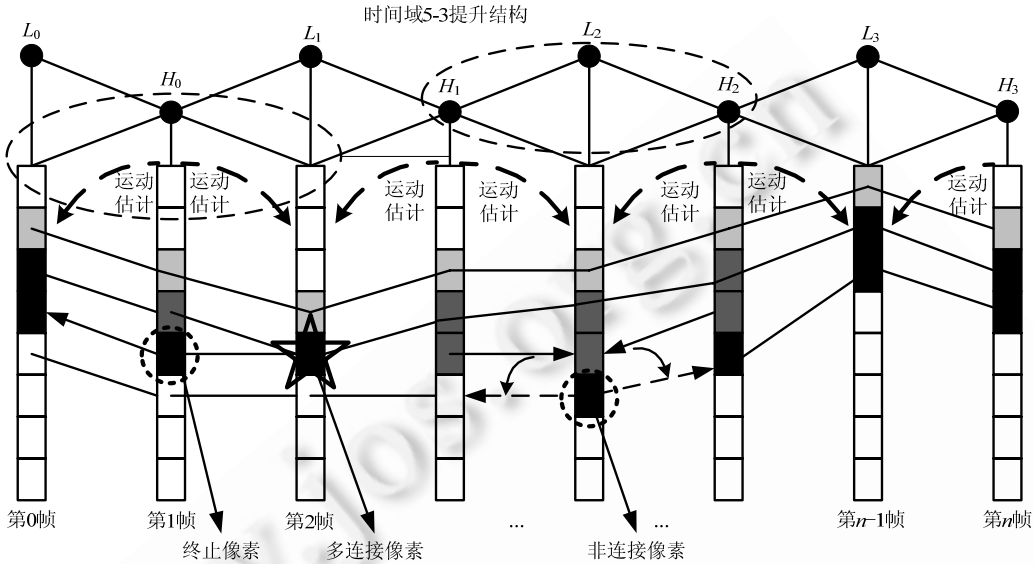


Fig.6 Example of improved motion thread filtering[68]

图 6 改进的运动线索滤波方法示例[68]

然而,无论是文献[46,68]还是文献[69],在 MCTF 过程中都需要利用边信息标识出非连接像素和多连接像素的位置,这难免要浪费一定的信息量;而且当运动向量为分数像素精度时,文献[49,56−69]等工作在很大的概率下会采用分数像素对奇数帧像素进行预测(如图 7 中用 C 预测 A),但在更新过程中 C 是不需要被更新的,只有其最邻近的整数像素才需被更新(如图 7 中用 E 更新 D).此时,就会发生预测线索与更新线索的偏离,这在理论上不能保证最有效地集中信号能量.尽管文献[70]试图通过预测算子将高频子带能量向低频子带传播、文献[71]通过建立预测算子与更新算子间的函数关系来最小化重构误差等方式缓解该问题,但是仍未彻底解决.

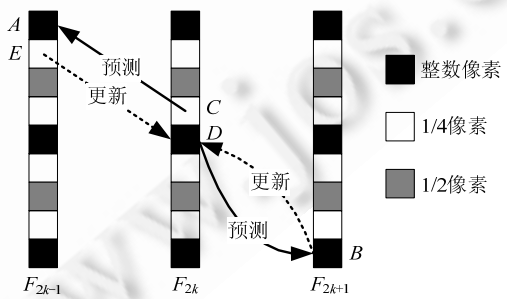


Fig.7 Deviation of predict thread from update thread

图 7 预测线索与更新线索的偏离

Xiong 等人[72,73]利用基于 Barbell 函数的小波变换巧妙地解决了上述问题,其主要思路是:在某个像素(如图 8(a)中 F_1 帧的像素 x_1)的预测阶段,采用由运动向量所指向像素的周围像素的线性组合作为其预测值;在像素(如图 8(b)中 F_2 帧的像素 x_2)的更新阶段,也做类似处理.所有的线性权值均由 Barbell 函数具体给出.一方面,因为该方法采用一个像素集合参与预测和更新,所以其计算得到的预测误差更小,高频系数幅值更低,能量更多地集中到低频子带;另一方面,该方法既无需对非连接像素和多连接像素进行特殊处理,又不存在预测线索与更新线索

© 中国科学院软件研究所 http://www.jos.org.cn

的偏离,只要保证具有预测与更新关系的像素对之间采用相同的权值即可.

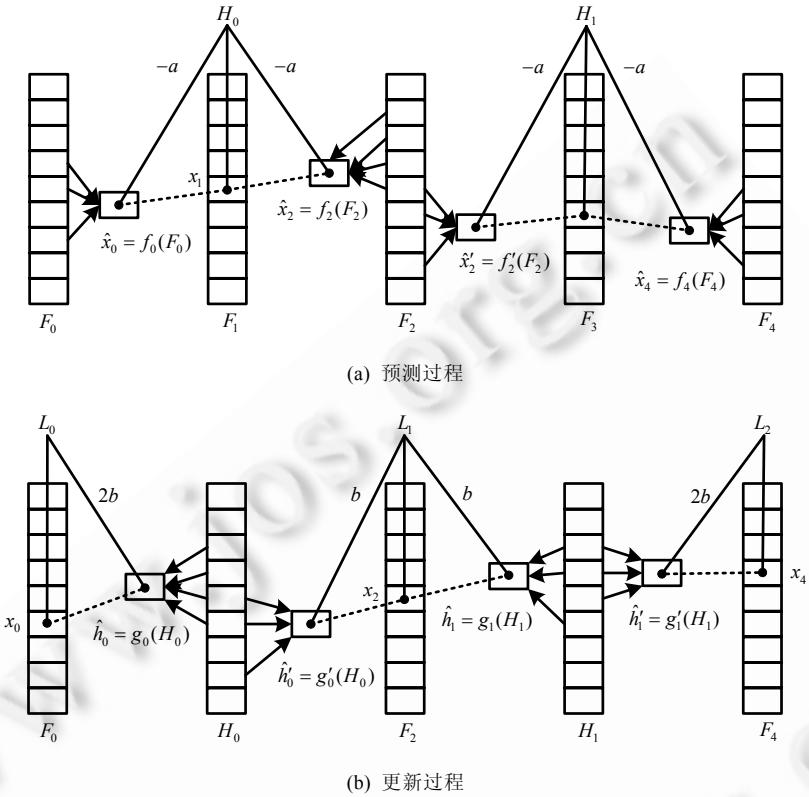


Fig.8 Procedure of MCTF based on Barbell lifting wavelet[72]

图 8 基于 Barbell 提升小波的 MCTF 过程[72]

需要注意的是:上述工作的初衷都是利用运动信息去除时间维相关性,而未从滤波器或正交基设计的角度来考虑,无法保证变换的正交性,尤其是在运动较为复杂的情况下.于是,Flierl 等人开展了一系列研究[74−77].其中,文献[74]提出一种正交运动补偿变换,它将 MCTF 分解为一系列两两正交的变换序列,称为增量变换(incremental transform),并利用尺度因子和去相关因子控制变换的正交性和能量集中能力,其性能优于以哈尔小波为变换核的 MCTF(对于运动复杂的序列尤为明显);文献[75]将文献[74]的时间维滤波精度扩展到 1/2 像素精度;文献[76]则将双向运动估计引入正交运动补偿变换,改进后的性能略优于以 5/3 小波为变换核的 MCTF;文献[77]进一步将由运动向量联系起来的像素集合建模为顶点加权图(vertex-weighted graph),以此为基础,提出一种子空间约束(subspace-constrained)变换,提高了文献[74]的能量集中能力.文献[77]的实验结果表明,其低频子带的能量仅比 K-L(karhunen-loeve)变换的低频子带高 1%左右.迄今为止,基于 Barbell 提升小波的 MCTF 和正交运动补偿变换是众多 MCTF 方法中变换效率最高、稀疏表示能力最强的两种.

对于第 2 类基于小波域 MCTF 的 3D 多尺度分析来讲,由于它在小波域内进行运动估计和时间域滤波,除了上文所讨论的问题以外,它面临的难题主要在于如何有效地解决小波的平移可变性所导致的 MCTF 性能下降.所谓平移可变性是指小波变换的下采样操作破坏了原始信号频率的线性相位关系,造成频率的混叠,进而使得信号在空间域中的平移与其小波系数在频率域中的相变不具有直接的线性关系(除了平移的像素数是下采样周期的整数倍以外)[78].为此,研究者们提出了大致两类方法来抑制混叠效应的发生,包括子带间接运动估计方法和克服平移可变性的运动估计方法.

子带间接运动估计方法的主要思路是,对参考帧的各子带进行必要的处理后再进行运动估计.常见的处理

© 中国科学院软件研究所 http://www.jos.org.cn

有移去小波变换中的下采样过程[79]、预测环补偿[80]、插值滤波[81]等.但是,这类方法仍然不能从根本上消除移变性的影响.

克服平移可变性的运动估计方法的主要思路是,构造出因下采样而丢失的相位信息.该类方法又主要有两个系列的研究工作:

- 首先,文献[82]在 2000 年提出了一种低频子带平移方法(low band shift),开启了第 1 系列的研究.具体做法是:将参考帧的低频子带分别沿着水平、竖直和对角线方向平移 1 个像素后,对当前帧和参考帧及其 3 个平移帧进行 1 阶小波变换,然后计算运动向量.对于多级小波分解,可以对前一次分解生成的低频子带重复上面的平移和分解操作.文献[83]在此基础上提出了基于 LBS 的运动估计算法,其准确度和主观质量明显高于空间域全搜索,但是其计算量比后者大约高 10.3%,空间复杂度是后者的 10 倍.文献[84]则利用层次策略提出了一种改进的 LBS 方法,运动估计仅在每层的低频分量进行,计算复杂度较之文献[83]降低了 70%,空间复杂度降低了 60%.通过建立多尺度变换系数的自相关模型,文献[85]证明了高频系数在运动估计中发挥的作用有限,甚至会误导宏块的搜索,从而在理论上证明了文献[84]的合理性和有效性.
- 与第 1 系列的子带平移思路不同,Li 通过研究发现,奇数相位的小波系数和偶数相位的小波系数可由唯一的相移矩阵(phase-shifting matrix)联系起来[86],据此能够恢复出任何相位的信息;进一步,Li 提出了基于过完备小波的可精确重构的 MCTF[87].与此同时,Andreopoulos 等人[88,89]也利用类似的预测矩阵计算出每一层低频子带的不同相位而无需反复采样、平移和分解,并提出了分层次的 LBS 以及相应的 MCTF 算法.限于篇幅,有关 MCTF 和小波域视频运动估计的具体技术细节详见文献[90,91].

除了上述基于 MCTF 的 3D 多尺度几何分析以外,还有少量相关的研究报道,如文献[92]阐述了基于三通道小波变换的 MCTF 的 3D 多尺度分析;文献[92−94]讨论了基于离散余弦变换的 MCTF 的 3D 多尺度分析;文献[95]探索了基于码激励线性预测(CELP)的 3D 多尺度分析;文献[96]受到 2D 多方向滤波器的启发,设计了一种用于时间域高频帧的非均匀方向滤波器,并提出了(t+2D)结构的具有方向多分辨率分析能力的 3D 多尺度分析.但由于这些工作相对孤立,本文也不再赘述,感兴趣的读者可参阅相应的文献.

与不包含运动估计的扩展 2D 多尺度几何分析相比,由于 MCTF 同时采用了运动补偿和时间维 1D 小波分解去除时间冗余信息,它不仅明显降低了时域高通子带的能量,也减少了时域低通子带中大幅值系数的个数和鬼影效应(如图 2(b)所示),进而显著提高了变换的能量集中能力和稀疏表示效率,可是视频序列往往具有复杂的 3D 相关性.文献[97]经过统计发现:在包含快速运动物体的视频序列中,如 Football,沿水平或竖直方向的数据相关性强于沿着时间方向的相关性;而在相对静止的视频序列中,如 Akiyo,数据相关性的强弱对比则相反.因此,无论是(2D+t)结构还是(t+2D)结构,均无法有效表示所有视频序列.另外,与方向小波变换类似,目前仍缺乏对这类多尺度几何分析的非线性逼近性能的理论探讨.

3 基于 3D 基函数的 3D 多尺度几何分析

由于视频具有复杂的 3D 相关性,本文将视频序列内部的面特征称为奇异曲面(如图 9 中播音员身体轮廓所组成的曲面穿过了视频序列).事实上,视频中存在的强相关性往往是由奇异曲面所形成的;并且这种强相关性在高清、超高清视频中表现得更加明显.若要实现更高的非线性逼近能力,就必须找到一种能够自适应地沿着奇异曲面方向进行滤波的变换.本文将满足这一条件的变换称为基于 3D 基函数的 3D 多尺度几何分析.

为了捕获 3D 面奇异性,现有文献大致报道了 3 类解决思路:

第 1 类研究工作的基本思路是频带分解,也就是将 3D 频率谱划分为众多楔形区域(如图 10 所示),每个区域代表某特定尺度下特定方向的频率,从而由该频谱提取出视频数据相应的方向信息.为实现这一目的,文献[98]对 2D 离散曲波变换进行扩展,在 3D 傅里叶(Fourier)域定义了半径窗函数和角度窗函数,二者分别控制尺度和方向的选择,它们可共同确定一个特定尺度、特定方向的楔形频带窗.通过将视频数据的 3D 傅里叶频谱与不同楔形窗函数相乘后做傅里叶逆变换,即可得到 3D 多尺度曲波系数.文献[99]则是对 2D 剪切波的扩展,它利用改

© 中国科学院软件研究所 http://www.jos.org.cn

进的拉普拉斯金字塔滤波完成尺度分解,再通过尺度矩阵和剪切矩阵张成在某个楔形区域具有支撑集的窗口函数,进而将该窗口函数与视频数据的傅里叶频谱相乘提取在某尺度下的 3D 剪切波方向系数.文献[100]已经从理论上证明 3D 剪切波对于 C^2 连续的奇异曲面可达到 $O(M^{-1}\log_2^2 M)$ 的最优非线性逼近阶.

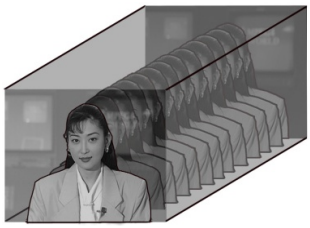


Fig.9 Singular face in spatiotemporal cube
图 9 时空体内的奇异曲面

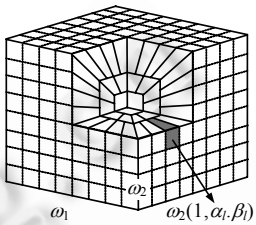


Fig.10 Diagram of 3D frequency spectrum partitioning[98]
图 10 3D 频谱划分示意图[98]

上述两种变换都是在傅里叶域完成的,其计算复杂度达到 $O(n^3\log_2 n)$以上[99];而且,二者均未考虑到视频数据在下采样过程中存在的频率混叠,使得变换的冗余率显著增大,3D 曲波和 3D 剪切波的冗余率分别约为 40 和 208.因此,文献[16]延伸了 2D 轮廓波的方向滤波思想,提出一种表面波(surfacelet)变换:首先,用沙漏型(hourglass)滤波器组分离出视频中沿 ω_1,ω_2 和 ω_3 轴向的信息;再用棋盘型滤波器和重采样矩阵建立方向滤波器组,对每个轴向信息进行连续两次 2D 方向滤波(如对 ω_1 轴向的信息先后沿 ω_1-ω_2 平面和 ω_1-ω_3 平面进行滤波),其相应的频带划分过程如图 11 所示.由于表面波变换完全在空间域进行,其计算量比 3D 剪切波降低了 87%.此外,方向滤波器的使用不仅保证了表面波变换与 3D 曲波变换有相同的频带分辨率,也有效地避免了频率混叠,使表面波仅有约 4.02%的冗余率[16].

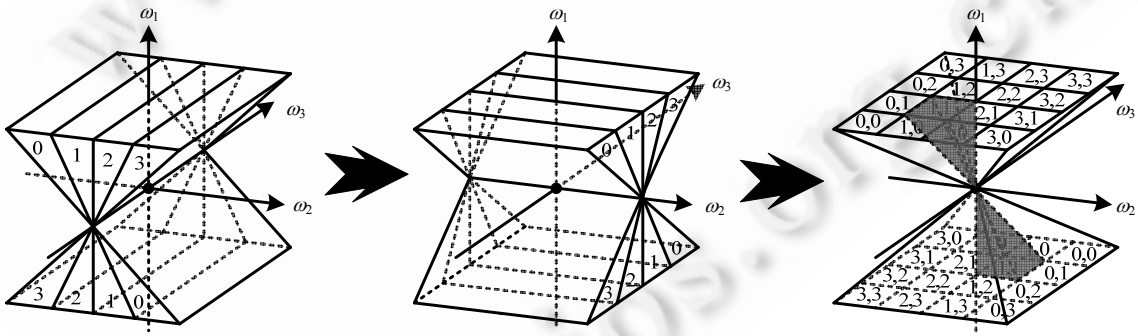


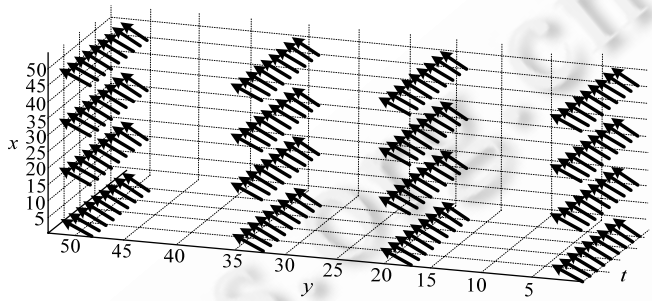
Fig.11 Frequency band division diagram of surfacelet by two sequential directional filtering[16]
图 11 Surfacelet 两次方向滤波对应的频带划分示意图[16]

第 2 类研究工作的基本思路是:选用本身具有方向性的基函数实现变换,如双树小波基(dual-tree wavelet)等.由于 1D 双树小波基可捕获沿±15°,±45°和±75°方向的线奇异,文献[101]利用张量积将其扩展为 3D 小波基,每个双树小波基由 4 个可分离的传统小波基组成.一级 3D 双树小波分解得到 28 个高频子带,每个子带可稀疏表示沿一个特定方向的相关信息.类似地,可操纵小波基(steerable wavelet)能够通过方向插值函数实现滤波方向的自适应控制.文献[102]在此基础上将多尺度分析和广义里斯(Riesz)变换相结合,提出对经小波分解后的子带进行里斯变换的 3D 可操纵小波变换框架及其快速实现方法.该变换继承了高阶里斯变换的方向可控性,较之 3D 双树小波变换更加灵活.

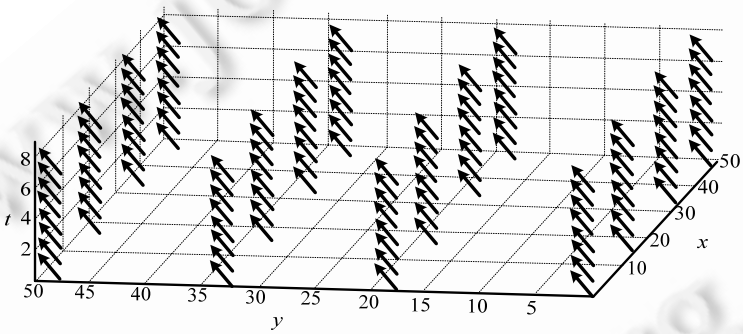
上述两类工作均是非自适应的变换,其基函数不依赖于具体的视频序列,但是它们仅可对分片光滑的 C^2 奇异曲面实现较好的逼近效率,尚不具备捕获时空体中复杂曲面奇异的能力.而第 3 类研究工作的基本思路则是

© 中国科学院软件研究所 http://www.jos.org.cn

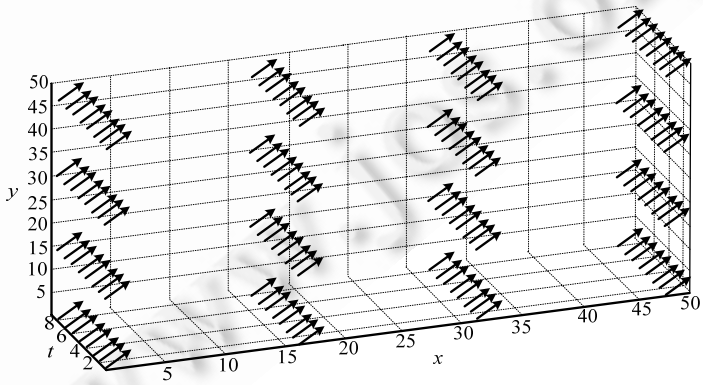
建立自适应的基函数来稀疏表示曲面奇异.文献[103]借助三维向量场将视频划分为一系列子视频体,并将这些子视频体按照其时空正则流(spatiotemporal regularity flow,简称 SPREF)方向(也就是梯度变化误差最小的方向)分成 3 类,即,*x*-*y* 平行流方向、*x*-*t* 平行流方向和 *y*-*t* 平行流方向(如图 12 所示).所谓 *x*-*y* 平行流方向是指向量场的分布只依赖于像素的空间坐标(*x*,*y*),其他两种情况可依此类推.再对每个子视频体沿着相应的正则流方向构造正交小波基,从而实现对曲面奇异的稀疏表示.



(a) *x*-*y* 平行流 50



(b) *y*-*t* 平行流



(c) *x*-*t* 平行流

Fig.12 Diagram of three kinds of SPREF[103]

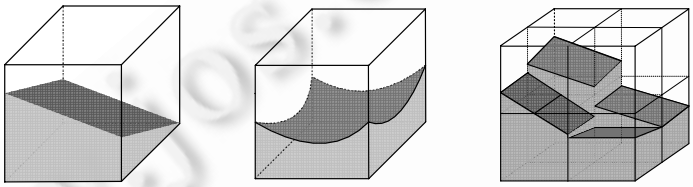
图 12 3 类时空正则流方向示意图[103]

实验结果显示,该方法的稀疏表示能力优于 LIMAT.但是,一方面,为了保证变换的正交性,上述方法要求向量场中的像素必须一一对应,不能出现无连接像素和多连接像素,并且只能采用平移运动估计模型;另一方面,3

© 中国科学院软件研究所 http://www.jos.org.cn

类平行流方向无法刻画所有的正则流和运动情况,仅是对物体复杂纹理和运动的近似.所以,这种方法仍不能精确表示视频中的曲面奇异.文献[104]则利用冗余字典描述视频的几何结构及其运动,并通过加权的匹配追踪(matching pursuit)来构造最优的稀疏逼近.不过,这种方法对运动模型的选取和运动向量场的精度较为敏感.

此外,Chandrasekaran 等人[105]指出:大部分 3D 多尺度分析方法均是为表示 C^2 连续的奇异曲面而设计的(如 3D 曲波等),而实际上,视频中往往存在更为复杂的曲面奇异.这样,文献[105]建立了一个定义在二进超立方体上的原子函数字典,并提出了一种自适应的高维多尺度几何分析——小面波(surflets)变换.对于 3D 情形,原子函数的形式如图 13(a)、图 13(b)所示,其中的 2D 奇异曲面可达到高阶光滑,曲面两侧是常值区域.可见,该原子函数字典可有效表示 2D 常值分片多项式曲面奇异(如图 13(c)所示).为进一步表示两侧存在平滑渐变区域的曲面奇异,文献[105]考虑到小波基适用于表示光滑函数的特点,提出了采用小波基和小面波基联合表示的方法,其基本思路是:通过对变换系数树进行剪枝来控制基函数的影响域,从而提高小波基函数在平滑区域的权值和小面波基函数在奇异曲面区域的权值.理论研究结果表明:小面波对具有 C^s 连续奇异曲面的 3D 分片常值函数的非线性逼近效率为 $O(M^{-s/2})$,而对具有 C^s 连续奇异曲面的 3D 分片 C^d 光滑函数的非线性逼近效率为 $O(M^{-\min\{s/2,2d/3\}})$.



(a) (b) (c)

Fig.13 Basis function diagram of surflets transform[105]

图 13 Surflets 变换基函数示意图[105]

4 基于时空非局部相关性的 3D 多尺度几何分析

作为时-频局部分析工具,上文的多尺度几何分析能够为存在局部数据依赖的图像和视频提供有效的稀疏表示.2005 年,Buades 等人[106]发现:图像内部尚普遍存在着另一类长久被忽视的数据相关,称为非局部相关性(non-local correlation),即,图像中的一个块往往与同一图像中位置不相邻的某块具有相似的边缘、纹理或轮廓等.文献[107]的研究也表明:由于像素沿奇异曲面的强相关,视频形成了四维空间$(x,y,t,I(x,y,t))$中的一个低维流形(其中,x 和 y 为空间维坐标,t 为时间维坐标,I 表示像素值).这表现为不仅一幅视频帧中存在非局部相关性[108],而且视频体内还存在 3D 非局部相关[109],即,视频体内的一个子视频体和与其不相邻的某个子视频体具有相似的奇异曲面(如图 14 所示).然而,无论是小波等有限支撑的基函数,还是傅里叶变换等无限支撑的基函数,均无法同时稀疏表达图像和视频数据的这种局部与非局部联合相关性.于是,近年来就出现了若干种联合非局部相关分析和多尺度分析的变换思想,本文将其称为基于时空非局部相关性的 3D 多尺度几何分析.

2007 年,文献[110]首次将非局部相关性引入到变换中,提出一种用于图像的 3D 稀疏变换,其核心思路是:将具有高度相似性的像素块分在一组,组成一个像素块序列(可视作一个视频序列),再对其进行可分离的 3D 小波变换.由于相似像素块往往具有类似的 2D 曲线(直线)奇异,经过第 3 维小波变换后,图像的能量进一步集中,其非线性逼近效率优于标准 2D 小波变换.但考虑到小波变换不能有效捕获曲线奇异,若能联合采用 1D 小波与具有更好方向奇异表示能力的 2D 多尺度变换,如 2D 轮廓波变换+1D 小波变换,则有望达到更高的非线性逼近阶.另外,该变换的计算复杂度偏高,处理一幅 $n\times n$ 像素的图像需要 $O(n^2W^2)$次操作(W 为搜索窗口尺寸).与此类似,文献[109]在 2011 年提出了一种信号依赖的图像变换(signal-dependent transform,简称 SDT),主要思想是:将图像分块后,对于每个块,利用模板匹配搜索到多个与其具有高度相似性的因果候选块集合;接下来计算候选块的协方差矩阵及其特征向量,进而采用这些特征向量对当前块施加 K-L 变换.虽然 SDT 方法在严格意义上并不是一种多尺度分析,并且计算复杂度过高,但它在图像变换中较早地融入了非局部相关分析的思想,值得借鉴.

© 中国科学院软件研究所 http://www.jos.org.cn

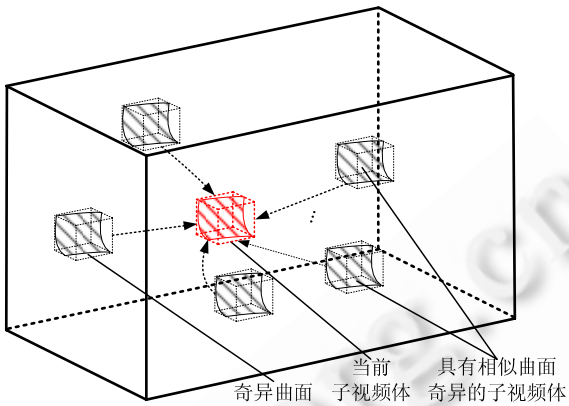


Fig.14 Diagram of non-local correlation in video cube

图 14 视频体中存在的非局部相关性示意图

在视频处理方面,文献[111−113]考虑到了子视频体之间的非局部相关性,提出一种 4D 非局部时空变换思想,它的主要过程是:首先,将相邻各帧中由运动向量联系起来的像素块(或像素面片)组织成一个子视频体;然后,利用非局部搜索找出内容相似的一系列子视频体,它们在逻辑上组成了一个 4D 的数据集合(每个子视频体本身就是 3D 数据,多个子视频体构成了第 4 维);最后,对每个子视频体集合进行可分离的 4D 小波或离散余弦分解.一方面,该变换不仅发掘了子视频体内部像素之间的时空局部相关,也兼顾了具有相似内容的子视频体之间的非局部时空相关,其逼近效率高于仅去除了时空局部相关的包含运动估计的扩展 2D 多尺度几何分析和基于 3D 基函数的多尺度几何分析;另一方面,由上面两节的论述可知,可分离 3D 小波基无法捕获每个子视频体的曲面奇异,并且在用非局部搜索组织相似子视频体时,若控制子视频体间内容相似度的阈值选取不当,则同族子视频体在进行第 4 维小波分解过程中就难免出现多处点奇异,影响非局部数据相关的去除能力.故而,可分离 4D 小波变换不是理想的基函数,并且关于多尺度变换基的选取和相似子视频体的组织策略尚待进一步优化.同时,该 4D 变换处理一个 $n\times n\times n$ 像素的子视频体需要 $O(n^3(W_1^2+W_2^3))$ 次操作(W_1 和 W_2 分别表示运动估计阶段和非局部搜索阶段的搜索窗口尺寸),而为了充分去除视频体的非局部相关,非局部搜索的搜索窗口要大于运动估计的搜索窗口,甚至还大于子视频体本身的尺寸(此时就需要多达 $O(n^6)$次的操作),这样的计算复杂度显然高于包含运动估计的扩展 2D 多尺度几何分析和基于 3D 基函数的多尺度几何分析.

需要指出的是,文献[109−113]的变换方法在严格意义上尚不能构成图像或视频数据的正交表示甚至完备表示,目前仅作为一个滤波环节应用在图像、视频的编码和去噪中.但是,其中蕴含的局部、非局部联合滤波的思想将是进一步改善多尺度几何分析效率的可行途径之一,有着重要的理论意义.此外,对上述变换的非线性逼近效率的分析仍然是一个难点,其原因在于:现有的非线性逼近分析模型均只考察某一多尺度变换在一个连续可微的局部曲线段(或曲面片)处的逼近阶,却未考虑曲线(或曲面)自身的非局部相关.例如,有一连续曲线(面)由 s 个相同的曲线段(或曲面片)组成,采用上述基于时空非局部相关性的多尺度几何分析对其进行稀疏表示,则与其他各类变换相比,其逼近效率将提高约 s 倍.因此,只有对连续曲线(面)的可微性、自相似程度等性质做出比以往更细致的假设之后,才能对该类多尺度几何分析的逼近效率进行刻画.

5 3D 数字几何模型的多尺度几何分析概述

3D 数字几何模型主要通过 3D 曲面来描述 3D 形体,已被广泛地应用于计算机辅助几何设计、计算机数学和数字娱乐等领域[114].利用多尺度几何分析研究曲线、曲面及由其组成之几何图组的频谱特征,有助于对其进行分类、检索、特征匹配和模式识别等处理.然而,与视频、核磁共振图像和高光谱遥感影像等性质不同,在 3D 几何造型中,几何模型往往是分片光滑的,且由多个子图组成,子图间又大都存在各类间断;另外,由于采样点分

© 中国科学院软件研究所 http://www.jos.org.cn

布不均匀,在曲线、曲面分析时,不同尺度下的尺度函数和小波函数之间并不是简单的伸缩和平移关系,多尺度分析中整平移空间的概念也产生了相应的变化,这使得传统的基于傅立叶变换的小波构造方法不再有效[115].在这种情况下,齐东旭等人[116]认为,并不是任何一个正交完备函数系都适用于几何信息的分析和综合.例如,傅里叶变换、勒让德(Legendre)多项式函数系等由于吉布斯(Gibbs)现象的存在而不能实现对曲面信息的有限且精确的表达;沃尔什(Walsh)变换和哈尔(Haar)变换又因其间断性太强,不适于对连续函数进行有限项逼近.因此,有关 3D 几何模型多尺度几何分析的研究在文献中并不如图像、视频的多尺度几何分析那么广泛,其要解决的关键问题亦与后者有所不同.考虑到 3D 曲面在实际应用中一般被参数化为三角网格或细分曲面,本文也仅关注在此二者基础上展开的多尺度几何分析.就我们掌握的资料来看,目前的研究成果主要可划分为如下两类.

第 1 类研究是 Sweldens 等人对传统小波变换所进行的一系列改进.针对网格点不均匀的情况,Chui 等人提出了多种小波构造方法,诸如区间小波(interval wavelet)[117]、用于非规则采样的样条小波[118]和加权小波(weighted wavelet)[119]等,统称为第二代小波.Sweldens 等人在此基础上将多尺度分析理论引入到第二代小波的构造中,提出了基于分裂、预测和提升这 3 个步骤的提升小波构造框架[33,115,120]以及适用于定义在任意曲面上的函数的提升小波变换[121,122].其基本思想是:利用网格细分实现采样点的分裂以及细尺度向粗尺度的转化;利用粗尺度网格点对细尺度网格点进行线性插值获得其对应的小波系数;再通过小波系数不断地对粗尺度系数进行提升与对偶提升,进而可从任意的双正交基开始,逐渐构造出满足实际所需的、具有理想消失矩和光滑性的小波基.该变换的主要优点在于:计算复杂度为 $O(n)$,变换速度快;可实现精确重构和原位运算,不需要额外的辅助空间.但是,文献[123]指出:提升小波变换并不是 3D 数字几何模型的正交表示,而且不能有效表示函数在三角网格边界或子图连接处的奇异.

第 2 类工作是由齐东旭等人开展的一系列关于 $L^2[0,1]$非连续正交完备函数系的研究,其核心思想是:在连续正交基函数系中引入间断,构造出分段连续的正交函数系,从而克服连续函数系下用有限个基函数的线性组合无法稀疏表示间断以及沃尔什变换和哈尔变换间断性太强的不足.文献[116]提出首先采用勒让德多项式的前$(k+1)$项作为尺度函数,再构造$(k+1)$个分段 k 次生成元函数,最后通过反复的压缩-复制过程由生成元函数建立非连续的 k 次正交系,称为 U-系统.理论研究结果表明:这种包含函数本身及其导数的各种层次间断性的正交函数系可以用有限项精确表示那些间断点仅出现在 $q/2^r$ 处(q 和 r 均为整数)的分段 k 次曲线,消除了吉布斯效应,但是不具有多分辨率特性.于是,文献[124]将 U-系统的构造过程进行了变形,从而提出了 V-系统.k 次 V-系统构造过程的前 2 步与 U-系统相同,而在压缩-复制过程中,则把每个生成元函数压缩 2^{n-2} 倍后分别复制到[0,1]上的 $1/2^{n-2}$ 等分区间.文献[125]证明:V-系统是一类有限区间上的特殊多小波,不仅与 U-系统具有相同的分段多项式稀疏表示能力,而且具备多尺度分析性质,其应用范围更为广泛.进一步来说,还可通过张量积的方法将上述 1D 形式的 U-系统和 V-系统向高维扩展,用于表达 3D 数字几何模型.然而,通过该形式构造的 U-系统和 V-系统是定义在矩形区域上的,它们在表示定义在三角域上的 3D 几何模型时就遇到了困难.为解决这一问题,文献[123,126]首先构造了定义在一个多层次自相似剖分三角域上的正交分片二元一次多项式;然后,将 3D 几何模型的所有三角面片一一映射到该三角域上的自剖分网格,使得每个三角面片与且仅与一个剖分子区域对应;再计算三角面片与剖分网格上对应的两个分片多项式的内积,即完成了几何模型的 V-系统变换.理论研究结果表明:该方法能够实现分片光滑的 3D 几何图组的整体正交分解,而其基函数的自相似剖分结构又为 3D 几何模型的多分辨率分析提供了可能,是一种有效的 3D 几何模型多尺度几何分析方法.关于 U-系统和 V-系统的详细了解请参阅文献[124],这里不再赘述.此外,文献[124]又研究了非均匀剖分下的各向异性 V-系统,并将其应用在 2D 图像的自适应稀疏表示中,但该方法仅能有效捕获水平和竖直方向的奇异,对复杂边缘的逼近效率有限,且计算过程较复杂,仍有待完善.

需要说明的是:虽然上述多尺度几何分析可实现 3D 数字几何模型的频谱特征提取,甚至有限项的精确重构,但目前尚未见到有关其非线性逼近效率的数学模型及证明.

© 中国科学院软件研究所 http://www.jos.org.cn

6 展 望

目前,3D 数据的多尺度几何分析方兴未艾,始终是多个学科的热点研究领域,其基础理论虽然日臻完善,但尚未完备,许多问题还处于不断探索中,尤其是基于3D基函数和基于非局部相关的多尺度几何分析的研究仍刚刚起步,甚至相关文献都不多见.鉴于此.我们认为,未来 3D 多尺度几何分析将有望在以下 4 个方面取得进展:

(1) 建立 3D 正交基函数以及有效的自适应多尺度几何分析框架

尽管 2D 多尺度几何分析经过近 30 年的发展已经相对完备,但 3D 多尺度分析却崭露头角:一方面,除文献[105]以外,现有研究几乎全部假设图像和视频数据具有 C^2 连续的曲线奇异或曲面奇异,并据此提出设计基函数的各向异性尺度关系原则,但 3D 数据分布实际远较此复杂;另一方面,物体运动的复杂性导致沿着时间维建立高效率的正交基变得非常棘手.在这种情况下,采用自适应变换机制是一种可选的解决思路,比如条带波变换、SPREF 方法和小面波变换均取得了优秀的逼近效率.故此,兼顾正交性、非线性逼近能力和计算复杂度的自适应 3D 多尺度几何分析,是需要深入研究和解决的问题.

(2) 建立评价基函数非线性逼近效率的理论和方法

由第 1 节可知,关于 2D 多尺度几何分析的非线性逼近效率的理论分析已经较为成熟,但是,除了极少数几个基于 3D 基函数的多尺度几何分析以外,对 3D 多尺度分析逼近能力的评价方法仍几乎空白.性能评价方法及其理论研究的发展,不仅有利于现有多尺度变换的推广应用,而且也能为 3D 多尺度分析指明一个改进方向,乃至促进相关研究的快速提升.

(3) 探索基于非局部相关性的多尺度几何分析方法

近几年来,图像和视频处理的多个领域,如增强、去噪、复原、超分辨率和纹理合成等,均已经发现并重视数据的非局部相关性,且有高效率的算法见诸报道,如文献[108−113,127−129]等,其结论表明,联合了非局部相关性的变换具有更好的稀疏表示能力.可是,多尺度分析的基函数是有限支撑的,傅里叶变换基函数又是无限支撑的,均无法很好地表示非局部数据相关.如何找到一种同时具有局部、非局部表示能力的基函数,或者是将现有基函数与非局部搜索相结合,建立一个多尺度变换框架,都是值得进一步探索的课题.

(4) 建模多尺度几何分析的系数分布规律

离散余弦变换虽不能为图像和视频提供最优表示,但其变换系数的幅值往往依 Zig-Zag 扫描顺序递减.相比之下,小波等多尺度变换的系数幅值分布却分散得多.实验发现:尽管多尺度变换系数大多具有压缩性、持续性和聚集性等特性,对变换系数分布模型的研究目前却仍不够深入.比如,文献[130]的最佳线性无偏估计模型不能描述尺度内系数的相关性,文献[131]的互信息模型不能解释零树的存在和分析系数分布规律,文献[85]的自相关模型能够建模多尺度变换系数的压缩性、持续性和聚集性等统计特性,文献[132]的带方向特征的轮廓波隐马尔科夫树模型得出了方向特征在父子子带间形成传递的结论,文献[133]的隐马尔科夫树模型则建模了非下采样轮廓波系数的持续性、聚集性和平移不变性,但文献[85,132,133]均未能刻画这些系数沿着时空奇异曲面方向上的统计特性;而且,这些模型只能实现系数分布规律的定性,却做不到像空间域模型[134]那样定量的刻画.同时,对于 3D 几何模型的多尺度变换系数分布规律的研究则更为缺乏,文献中鲜有报道.

实际上,对于大部分应用研究者来讲,对系数分布规律进行建模的意义绝不亚于找到一种最优变换的意义.因此,建模多尺度几何分析系数的分布规律从而实现其定性甚至定量刻画,是一个亟待研究的难题.

7 结 论

20 世纪 50 年代以来,3D 可视媒体处理成为多媒体技术领域最重要的研究内容之一.而大数据时代下,对 3D 可视媒体的处理和分析则愈加关键.现有处理方法不外乎空间域方法与变换域方法两类,且后者的研究和应用范围更为广泛.故此,3D 多尺度几何分析作为一类重要的变换工具,对其进行深入研究甚是必要.

本文从图像、视频和几何模型这 3 个方面阐述了 3D 多尺度几何分析的研究进展:首先,从 2D 图像多尺度几何分析切入,以 2D 基函数向 3D 基函数的演进为主线,将现有视频多尺度几何分析方法分为由 2D 基函数直

© 中国科学院软件研究所 http://www.jos.org.cn

接扩展的 3D 多尺度几何分析、基于 3D 基函数的 3D 多尺度几何分析和基于时空非局部相关性的 3D 多尺度几何分析这 3 类;其次,简要介绍了用于 3D 数字几何模型的多尺度几何分析研究进展.经过对各类分析工具发展脉络及其基本思想的详细梳理和比较,本文一方面发掘了不同变换算法的优势、不足和适用范围;另一方面,发现目前 3D 正交基函数及其非线性逼近性能分析的研究仍很不成熟,相关报道较少,而且 3D 数据的非局部相关性尚待挖掘、多尺度变换系数的分布规律亟需进一步明确等.据此,本文对 3D 多尺度几何分析的未来发展进行了展望.

References:

[1] Gao W, Zhao DB, Ma SW. Principles of Digital Video Coding Technology. Beijing: Science Press, 2010 (in Chinese).

[2] Jiao LC, Hou B, Wang S, Liu F. Image Multiscale Geometric Analysis: Theory and Applications—Beyond Wavelets. Xi'an: Xi Dian University Publish House, 2008 (in Chinese).

[3] Donoho DL, Vetterli M, De Vore RA, Daubechies I. Data compression and harmonic analysis. IEEE Trans. on Information Theory, 1998,44(6):2435−2476. [doi: 10.1109/18.720544]

[4] Jiao LC, Tan S. Development and prospect of image multiscale geometric analysis. Acta Electronica Sinica, 2003,31(Z1):1975−1981 (in Chinese with English abstract).

[5] Wang XH, Sun Q, Song CM, Liu D. Advances in image coding based on multiscale geometric analysis. Journal of Computer Research and Development, 2010,47(6):1132−1143 (in Chinese with English abstract).

[6] Candès EJ. Ridgelets: Theory and applications [Ph.D. Thesis]. Stanford: Stanford University, 1998.

[7] Candès EJ, Donoho DL. Ridgelets: A key to higher-dimensional intermittency? Philosophical Trans. of the Royal Society of London Series A, 1999,357(1760):2495−2509. [doi: 10.1098/rsta.1999.0444]

[8] Candès EJ, Donoho DL. Curvelets: A surprisingly effective nonadaptive representation for objects with edges. In: Cohen A, Rabut C, Schumaker LL, eds. Proc. of the Int'l Conf. on Curves and Surfaces. Nashville: Vanderbilt University Press, 1999. 105−120. [doi: 10.1.1.161.9294]

[9] Candès EJ, Donoho DL. Curvelets and curvilinear integrals. Journal of Approximation Theory, 2001,113(1):59−90. [doi: 10.1006/jath.2001.3624]

[10] Pennec EL, Mallat S. Bandelet image approximation and compression. SIAM Journal of Multiscale Modeling and Simulation, 2005,4(3):992−1039. [doi: 10.1137/040619454]

[11] Pennec EL, Mallat S. Sparse geometric image representation with bandelets. IEEE Trans. on Image Processing, 2005,14(4):423−438. [doi: 10.1109/TIP.2005.843753]

[12] Huo XM, Chen JH. JBEAM: Multiscale curve coding via beamlets. IEEE Trans. on Image Processing, 2005,14(11):1665−1677. [doi: 10.1109/TIP.2005.857273]

[13] Velisavljević V, Beferull-Lozano B, Vetterli M, Dragotti PL. Directionlets: Anisotropic multidirectional representation with separable filtering. IEEE Trans. on Image Processing, 2006,15(7):1916−1933. [doi: 10.1109/TIP.2006.877076]

[14] Do MN, Vetterli M. Contourlets: A new directional multiresolution image representation. In: Friedlander B, ed. Proc. of the Asilomar Conf. on Signals, Systems and Computers. Los Alamitos: IEEE Computer Society Press, 2002. 497−501. [doi: 10.1109/ACSSC.2002.1197232]

[15] Easley G, Labate D, Lim WQ. Sparse directional image representations using the discrete shearlet transform. Applied and Computational Harmonic Analysis, 2008,25(1):25−46. [doi: 10.1016/j.acha.2007.09.003]

[16] Lu YM, Do MN. Multidimensional directional filter banks and surfacelets. IEEE Trans. on Image Processing, 2007,16(4):918−931. [doi: 10.1109/TIP.2007.891785]

[17] Watson AB. The cortex transform: rapid computation of simulated neural images. Computer Vision, Graphics, and Image Processing, 1987,39(3):311−327. [doi: 10.1016/S0734-189X(87)80184-6]

[18] Simoncelli EP, Freeman WT, Adelson EH, Heeger DJ. Shiftable multiscale transforms. IEEE Trans. on Information Theory, 1992, 38(2):587−607. [doi: 10.1109/18.119725]

[19] Meyer FG, Coifman RR. Brushlets: A tool for directional image analysis and image compression. Applied and Computational Harmonic Analysis, 1997,4(2):147−187. [doi: 10.1006/acha.1997.0208]

[20] Kingsbury N. Complex wavelets for shift invariant analysis and filtering of signals. Applied and Computational Harmonic Analysis, 2001,10(3):234−253. [doi: 10.1006/acha.2000.0343]

© 中国科学院软件研究所 http://www.jos.org.cn

[21] Do MN. Directional multiresolution image representations [Ph.D. Thesis]. Swiss Federal Institute of Technology, 2001.

[22] Donoho DL. Orthonomal ridgelets and linear singularities. SIAM Journal on Mathematical Analysis, 2000,31(5):1062−1099. [doi: 10.1137/S0036141098344403]

[23] Candès EJ, Demanet L, Donoho D, Ying L. Fast discrete curvelet transforms. Multiscale Modeling & Simulation, 2006,5(3): 861−899. [doi: 10.1137/05064182X]

[24] Bamberger RH, Smith MJT. A filter bank for the directional decomposition of images: Theory and design. IEEE Trans. on Signal Processing, 1992,40(4):882−893. [doi: 10.1109/78.127960]

[25] Nguyen TT, Oraintara S. A directional decomposition: Theory, design, and implementation. In: Antoniou A, Trajković L, eds. Proc. of the IEEE Int'l Symp. on Circuits and Systems. Los Alamitos: IEEE Computer Society Press, 2004. 281−284. [doi: 10.1109/ISCAS.2004.1328738]

[26] Yan CM, Guo BL, Yi M. Fast algorithm for nonsubsampled Contourlet transform. Acta Automatica Sinica, 2014,40(4):757−762. [doi: 10.3724/SP.J.1004.2014.00757]

[27] Guo KH, Labate D. Optimally sparse multidimensional representation using shearlets. SIAM Journal on Mathematical Analysis, 2007,39(1):298−318. [doi: 10.1137/060649781]

[28] Lim WQ. The discrete shearlet transform: A new directional transform and compactly supported shearlet frames. IEEE Trans. on Image Processing, 2010,19(5):1166−1180. [doi: 10.1109/TIP.2010.2041410]

[29] Donoho DL. Wedgelets: Nearly minimax estimation of edges. The Annals of Statistics, 1999,27(3):859−897. [doi: 10.1214/aos/1018031261]

[30] Lisowska A. Smoothlets—Multiscale functions for adaptive representation of images. IEEE Trans. on Image Processing, 2011, 20(7):1777−1787. [doi: 10.1109/TIP.2011.2108662]

[31] Mallat S. Geometrical grouplets. Applied and Computational Harmonic Analysis, 2009,26(2):161−180. [doi: 10.1016/j.acha.2008.03.004]

[32] Mallat S, Peyré G. A review of Bandlet methods for geometrical image representation. Numerical Algorithms, 2007,44(3):205−234. [doi: 10.1007/s11075-007-9092-4]

[33] Sweldens W. The lifting scheme: A construction of second generation wavelets. SIAM Journal on Mathematical Analysis, 1998, 29(2):511−546. [doi: 10.1137/S0036141095289051]

[34] Taubman D. Adaptive, non-separable lifting transforms for image compression. In: Ozawa S, Sakai Y, eds. Proc. of the Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 1999. 772−776. [doi: 10.1109/ICIP.1999.817221]

[35] Gerek ÖN, Çetin AE. A 2-D orientation-adaptive prediction filter in lifting structures for image coding. IEEE Trans. on Image Processing, 2006,15(1):106−111. [doi: 10.1109/TIP.2005.859369]

[36] Chang CL, Girod B. Direction-Adaptive discrete wavelet transform for image compression. IEEE Trans. on Image Processing, 2007,16(5):1289−1302. [doi: 10.1109/TIP.2007.894242]

[37] Ding WP, Wu F, Wu XL, Li SP, Li HQ. Adaptive directional lifting-based wavelet transform for image coding. IEEE Trans. on Image Processing, 2007,16(2):416−427. [doi: 10.1109/TIP.2006.888341]

[38] Tanaka Y, Hasegawa M, Kato S, Ikehara M, Nguyen TQ. Adaptive directional wavelet transform based on directional prefiltering. IEEE Trans. on Image Processing, 2010,19(4):934−945. [doi: 10.1109/TIP.2009.2038820]

[39] Claypoole RL, Davis GM, Sweldens W, Baraniuk RG. Nonlinear wavelet transforms for image coding via lifting. IEEE Trans. on Image Processing, 2003,12(12):1449−1459. [doi: 10.1109/TIP.2003.817237]

[40] Boulgouris NV, Tzovaras D, Strintzis MG. Lossless image compression based on optimal prediction, adaptive lifting, and conditional arithmetic coding. IEEE Trans. on Image Processing, 2001,10(1):1−14. [doi: 10.1109/83.892438]

[41] Muramatsu S, Han DD, Kobayashi T, Kikuchi H. Directional lapped orthogonal transform: Theory and design. IEEE Trans. on Image Processing, 2012,21(5):2434−2448. [doi: 10.1109/TIP.2011.2182055]

[42] Muramatsu S, Han DD, Kobayashi T, Kikuchi H. Three-Dimensional subband coding of video. IEEE Trans. on Image Processing, 1995,4(2):125−139. [doi: 10.1109/83.342187]

[43] Kim BJ, Xiong Z, Pearlman WA. An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT). In: Storer JA, Cohn M, eds. Proc. of the Data Compression Conf. Los Alamitos: IEEE Computer Society Press, 1997. 251−260. [doi: 10.1109/DCC.1997.582048]

© 中国科学院软件研究所 http://www.jos.org.cn

[44] Kim BJ, Xiong ZX, Pearlman WA. Low bit-rate scalable video coding with 3D set partitioning in hierarchical trees (3D SPIHT). IEEE Trans. on Circuits and Systems for Video Technology, 2000,10(8):1374−1387. [doi: 10.1109/76.889025]

[45] Chen YW, Pearlman W. Three-Dimensional subband coding of video using the zero-tree method. In: Ansari R, Smith MJT, eds. Proc. of the SPIE 2727, Visual Communications and Image Processing. Bellingham: SPIE Press, 1996. 1302−1309. [doi: 10.1117/12.233203]

[46] Xu JZ, Xiong ZX, Li SP, Zhang YQ. Three-Dimensional embedded subband coding with optimized truncation (3-D ESCOT). Applied and Computational Harmonic Analysis, 2001,10(3):290−315. [doi: 10.1006/acha.2000.0345]

[47] Mehrseresht N, Taubman D. Spatial scalability and compression efficiency within a flexible motion compensated 3D-DWT. In: Kot A, Gray R, Ma KK, eds. Proc. of the Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 2004. 1325−1328. [doi: 10.1109/ICIP.2004.1419743]

[48] Mehrseresht N, Taubman D. A flexible structure for fully scalable motion-compensated 3-D DWT with emphasis on the impact of spatial scalability. IEEE Trans. on Image Processing, 2006,15(3):740−753. [doi: 10.1109/TIP.2005.860619]

[49] Secker A, Taubman D. Lifting-Based invertible motion adaptive transform (LIMAT) framework for highly scalable video compression. IEEE Trans. on Image Processing, 2003,12(12):1530−1542. [doi: 10.1109/TIP.2003.819433]

[50] Ohm JR. Three-Dimensional subband coding with motion compensation. IEEE Trans. on Image Processing, 1994,3(5):559−57l. [doi: 10.1109/83.334985]

[51] Hsiang ST. High scalable subband/wavelet image and video coding [Ph.D. Thesis]. New York: Rensselaer Polytechnic Institute, 2002.

[52] Taubman D, Zakhor A. Multirate 3-D subband coding of video. IEEE Trans. on Image Processing, 1994,3(5):572−588. [doi: 10.1109/83.334984]

[53] Wang A, Xiong ZX, Chou PA, Mehrotra S. Three-Dimensional wavelet coding of video with global motion compensation. In: Storer JA, Cohn M, eds. Proc. of the Data Compression Conf. Los Alamitos: IEEE Computer Society Press, 1999. 404−413. [doi: 10.1109/DCC.1999.755690]

[54] Ohm JR. Advanced packet video coding based on layered VQ and SBC techniques. IEEE Trans. on Circuits Systems and Video Technology, 1993,3(3):208−221. [doi: 10.1109/76.224231]

[55] Choi SJ, Woods JW. Motion-Compensated 3-D subband coding of video. IEEE Trans. on Image Processing, 1999,8(2):155−167. [doi: 10.1109/83.743851]

[56] Pesquet-Popescu B, Bottreau V. Three-Dimensional lifting schemes for motion compensated video compression. In: Mathews VJ, Swindlehurst AL, eds. Proc. of the Int'l Conf. on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2001. 1793−1796. [doi: 10.1109/ICASSP.2001.941289]

[57] Hsiang ST, Woods JW. Invertible three-dimensional analysis/synthesis system for video coding with half-pixel-accurate motion estimation. In: Aizawa K, Stevenson RL, Zhang YQ, eds. Proc. of the SPIE 3653, Visual Communications and Image Processing. Bellingham: SPIE Press, 1999. 537−546. [doi: 10.1117/12.334704]

[58] Woods JW, Chen PS. Improved MC-EZBC with quarter-pixel motion vectors. MPEG2002/M8366: Fairfax, MPEG: ISO/IEC JTC1/SC29/WG11, 2002. 1−16.

[59] Chen PS, Woods JW. Bidirectional MC-EZBC with lifting implementation. IEEE Trans. on Circuits and Systems for Video Technology, 2004,14(10):1183−1194. [doi: 10.1109/TCSVT.2004.833165]

[60] Secker A, Taubman D. Motion-Compensated highly scalable video compression using an adaptive 3D wavelet transform based on lifting. In: Pitas I, Venetsanopoulos AN, Pappas TN, eds. Proc. of the IEEE Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 2001. 1029−1032. [doi: 10.1109/ICIP.2001.958672]

[61] Secker A, Taubman D. Highly scalable video compression using a lifting-based 3D wavelet transform with deformable mesh motion compensation. In: Tekalp AM, Reibman A, Knox K, eds. Proc. of the IEEE Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 2002. 749−752. [doi: 10.1109/ICIP.2002.1039080]

[62] Secker A, Taubman D. Highly scalable video compression with scalable motion coding. IEEE Trans. on Image Processing, 2004, 13(8):1029−1041. [doi: 10.1109/TIP.2004.826089]

[63] Turaga DS, van der Schaar M, Andreopoulos Y, Munteanu A, Schelkens P. Unconstrained motion compensated temporal filtering (UMCTF) for efficient and flexible interframe wavelet video coding. Signal Processing: Image Communication, 2005,20(1): 1−19. [doi: 10.1016/j.image.2004.08.006]

© 中国科学院软件研究所 http://www.jos.org.cn

[64] Song CM, Wang XH, Zhang FY. Visually lossless accuracy of motion vector in overcomplete wavelet-based scalable video coding. Journal of Computers, 2009,4(9):821−828. [doi: 10.4304/jcp.4.9.821-828]

[65] Luo L, Li J, Li SP, Zhuang ZQ, Zhang YQ. Motion compensated lifting wavelet and its application in video coding. In: Ali MAT, Das DK, eds. Proc. of the IEEE Int'l Conf. on Multimedia and Expo. Los Alamitos: IEEE Computer Society Press, 2001. 365−368. [doi: 10.1109/ICME.2001.1237732]

[66] Flierl M, Girod B. Video coding with motion-compensated lifted wavelet transforms. Signal Processing: Image Communication, 2004,19(7):561−575. [doi: 10.1016/j.image.2004.05.002]

[67] Xu JZ, Xiong ZX, Li SP, Zhang YQ. Memory-Constrained 3D wavelet transform for video coding without boundary effects. IEEE Trans. on Circuits and Systems for Video Technology, 2002,12(9):812−818. [doi: 10.1109/TCSVT.2002.803231]

[68] Luo L, Wu F, Li SP, Xiong ZX, Zhuang ZQ. Advanced motion threading for 3D wavelet video coding. Signal Processing: Image Communication, 2004,19(7):601−616. [doi: 10.1016/j.image.2004.05.004]

[69] Pau G, Tillier C, Pesquet-Popescu B, Heijmans H. Motion compensation and scalability in lifting-based video coding. Signal Processing: Image Communication, 2004,19(7):577−600. [doi: 10.1016/j.image.2004.05.003]

[70] Feng B, Xu JZ, Wu F, Yang SQ, Li SP. Energy distributed update steps (EDU) in lifting based motion compensated video coding. In: Gray R, MA KK, eds. Proc. of the IEEE Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 2004. 2267−2270. [doi: 10.1109/ICIP.2004.1421550]

[71] Girod B, Han S. Optimum update for motion-compensated lifting. IEEE Signal Processing Letters, 2005,12(2):150−153. [doi: 10.1109/LSP.2004.840874]

[72] Xiong RQ, Wu F, Xu JZ, Li SP, Zhang YQ. Barbell lifting wavelet transform for highly scalable video coding. In: van der Schaar M, ed. Proc. of the Picture Coding Symp. Los Alamitos: IEEE Computer Society Press, 2004. 237−242.

[73] Xiong RQ, Xu JZ, Wu F, Li SP. Barbell-Lifting based 3-D wavelet coding scheme. IEEE Trans. on Circuits Systems and Video Technology, 2007,17(9):1256−1269. [doi: 10.1109/TCSVT.2007.905507]

[74] Flierl M, Girod B. A motion-compensated orthogonal transform with energy-concentration constraint. In: Liu ZC, Dumitras A, Zhang ZY, eds. Proc. of the IEEE Int'l Workshop on Multimedia Signal Processing. Los Alamitos: IEEE Computer Society Press, 2006. 391−394. [doi: 10.1109/MMSP.2006.285337]

[75] Flierl M, Girod B. Half-Pel accurate motion-compensated orthogonal video transforms. In: Storer JA, Marcellin MW, eds. Proc. of the Data Compression Conf. Los Alamitos: IEEE Computer Society Press, 2007. 13−22. [doi: 10.1109/DCC.2007.43]

[76] Flierl M, Girod B. A new bidirectionally motion-compensated orthogonal transform for video coding. In: Kuh A, Huang YF, eds. Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2007. 665−668. [doi: 10.1109/ICASSP. 2007.365995]

[77] Liu D, Flierl M. Motion-Adaptive transforms based on vertex-weighted graphs. In: Bilgin A, Marcellin MW, Serra-Sagrista J, Storer JA, eds. Proc. of the Data Compression Conf. Los Alamitos: IEEE Computer Society Press, 2013. 181−190. [doi: 10.1109/DCC.2013.23]

[78] Tsunashima K, Stampleman JB, Bove VM. A scalable motion-compensated subband image coder. IEEE Trans. on Communications, 1994,42(2/3/4):1894−1901. [doi: 10.1109/TCOMM.1994.582899]

[79] Nosratinia A, Orchard MT. Multi-Resolution backward video coding. In: Chellappa R, ed. Proc. of the IEEE Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 1995. 563−566. [doi: 10.1109/ICIP.1995.537541]

[80] Wang XH. Research on spatial scalable video coding based on wavelet. Computer Science, 2002,29(7):115−117 (in Chinese with English abstract).

[81] Yang XG, Ramchandran K. Scalable wavelet video coding using aliasing-reduced hierarchical motion compensation. IEEE Trans. on Image Processing, 2000,9(5):778−791. [doi: 10.1109/83.841519]

[82] Park HW, Kim HS. Motion estimation using low-band-shift method for wavelet-based moving-picture coding. IEEE Trans. on Image Processing, 2000,9(4):577−587. [doi: 10.1109/83.841935]

[83] Kim HS, Park HW. Wavelet-Based moving-picture coding using shift-invariant motion estimation in wavelet domain. Signal Processing: Image Communication, 2001,16(7):669−679. [doi: 10.1016/S0923-5965(00)00044-8]

[84] Song CM, Wang XH. A new scalable video motion estimation scheme in the wavelet domain. Chinese Journal of Computers, 2006,29(12):2112−2118 (in Chinese with English abstract).

[85] Song CM. Research on scalable video coding based on multiscale analysis [Ph.D. Thesis]. Nanjing: Nanjing University, 2010 (in Chinese with English abstract).

© 中国科学院软件研究所 http://www.jos.org.cn

[86] Li X. New results of phase shifting in the wavelet space. IEEE Signal Processing Letter, 2003,10(7):193−195. [doi: 10.1109/LSP.2003.811587]

[87] Li X. Scalable video compression via overcomplete motion compensated wavelet coding. Signal Processing: Image Communication, 2004,19(7):637−651. [doi: 10.1016/j.image.2004.05.006]

[88] Andreopoulos Y, Munteanu A, Barbarien J, Van der Schaar M, Cornelis J, Schelkens P. In-Band motion compensated temporal filtering. Signal Processing: Image Communication, 2004,19(7):653−673. [doi: 10.1016/j.image.2004.05.007]

[89] Andreopoulos Y, Munteanu A, Van der Auwera G, Cornelis JPH, Schelkens P. Complete-to-Overcomplete discrete wavelet transforms: Theory and applications. IEEE Trans. on Signal Processing, 2005,53(4):1398−1412. [doi: 10.1109/TSP.2005.843707]

[90] Song CM, Wang XH. Research progress in video motion estimation techniques in the wavelet domain. Chinese Journal of Computers, 2005,28(10):1716−1727 (in Chinese with English abstract).

[91] Wang XH, Song CM. Scalable Image and Video Coding. Beijing: Science Press, 2009 (in Chinese).

[92] Tillier C, Pesquet-Popescu B, van der Schaar M. 3-Band motion-compensated temporal structures for scalable video coding. IEEE Trans. on Image Processing, 2006,15(9):2545−2557. [doi: 10.1109/TIP.2006.877411]

[93] Atta R, Ghanbari M. Motion-Compensated temporal filtering based on the DCT. In: Ottersten B, Swindlehurst AL, eds. Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2008. 1033−1036. [doi: 10.1109/ICASSP.2008.4517789]

[94] Nicoulin A, Mattavelli M, Li W, Basso A, Popat AC, Kunt M. Image sequence coding using motion compensated subband decomposition. In: Sezan MI, Lagendijk RL, eds. Proc. of the Motion Analysis and Image Sequence Processing. New York: Springer-Verlag, 1993. 225−256. [doi: 10.1007/978-1-4615-3236-1_8]

[95] Foodeei M, Dubois E. Coding image sequence intensities along motion trajectories using EC-CELP quantization. In: Bovik AC, ed. Proc. of the Int'l Conf. on Image Processing. Los Alamitos: IEEE Computer Society Press, 1994. 720−724. [doi: 10.1109/ICIP.1994.413409]

[96] Xiong HK, Zhu LC, Ma NN, Zheng YF. Scalable video compression framework with adaptive orientational multiresolution transform and nonuniform directional filterbank design. IEEE Trans. on Circuits and Systems for Video Technology, 2011,21(8):1085−1099. [doi: 10.1109/TCSVT.2011.2133310]

[97] Liu AM, Lin WS, Paul M, Zhang F, Deng CW. Optimal compression plane for efficient video coding. IEEE Trans. on Image Processing, 2011,20(10):2788−2798. [doi: 10.1109/TIP.2011.2134858]

[98] Ying LX, Demanet L, Candès E. 3D discrete curvelet transform. In: Papadakis M, Laine AF, Unser MA, eds. Proc. of the SPIE 5914, Wavelets XI. Bellingham: SPIE Press, 2005. 1−11. [doi: 10.1117/12.616205]

[99] Negi PS, Labate D. 3-D discrete shearlet transform and video processing. IEEE Trans. on Image Processing, 2012,21(6):2944−2953. [doi: 10.1109/TIP.2012.2183883]

[100] Guo KH, Labate D. Optimally sparse 3D approximations using shearlet representations. Electronic Research Announcements in Mathematical Sciences, 2010,17(10):125−137. [doi: 10.3934/era.2010.17.125]

[101] Selesnick IW, Li KY. Video denoising using 2D and 3D dual-tree complex wavelet transforms. In: Unser MA, Aldroubi A, Laine AF, eds. Proc. of the SPIE 5207, Wavelets: Applications in Signal and Image Processing X. Bellingham: SPIE Press, 2003. 607−618. [doi: 10.1117/12.504896]

[102] Chenouard N, Unser M. 3D steerable wavelets in practice. IEEE Trans. on Image Processing, 2012,21(11):4522−4533. [doi: 10.1109/TIP.2012.2206044]

[103] Alatas O, Javed O, Shah M. Video compression using spatiotemporal regularity flow. IEEE Trans. on Image Processing, 2006, 15(12):3812−3823. [doi: 10.1109/TIP.2006.884930]

[104] Escoda ÒD, Monaci G, Figueras i Ventura RM, Vandergheynst P, Bierlaire M. Geometric video approximation using weighted matching pursuit. IEEE Trans. on Image Processing, 2009,18(8):1703−1716. [doi: 10.1109/TIP.2009.2021315]

[105] Chandrasekaran V, Wakin MB, Baron D, Baraniuk RG. Representation and compression of multidimensional piecewise functions using surflets. IEEE Trans. on Information Theory, 2009,55(1):374−400. [doi: 10.1109/TIT.2008.2008153]

[106] Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. In: Schmid C, Soatto S, Tomasi C, eds. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2005. 60−65. [doi: 10.1109/CVPR.2005.38]

[107] Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. Science, 2000,290(5500):2323−2326. [doi: 10.1126/science.290.5500.2323]

© 中国科学院软件研究所 http://www.jos.org.cn

[108] Cherigui S, Guillemot C, Thoreau D, Guillotel P, Pérez P. Correspondence map-aided neighbor embedding for image intra prediction. IEEE Trans. on Image Processing, 2013,22(3):1161−1174. [doi: 10.1109/TIP.2012.2227772]

[109] Lan CL, Xu JZ, Shi GM, Wu F. Exploiting non-local correlation via signal-dependent transform (SDT). IEEE Journal of Selected Topics in Signal Processing, 2011,5(7):1298−1308. [doi: 10.1109/JSTSP.2011.2165273]

[110] Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3D transform-domain collaborative filtering. IEEE Trans. on Image Processing, 2007,16(8):2080−2095. [doi: 10.1109/TIP.2007.901238]

[111] Maggioni M, Boracchi G, Foi A, Egiazarian K. Video denoising, deblocking, and enhancement through seperable 4-D nonlocal spatiotemporal transforms. IEEE Trans. on Image Processing, 2012,21(9):3952−3966. [doi: 10.1109/TIP.2012.2199324]

[112] Maggioni M, Katkovnik V, Egiazarian K, Foi A. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. IEEE Trans. on Image Processing, 2013,22(1):119−133. [doi: 10.1109/TIP.2012.2210725]

[113] Li W, Zhang J, Dai QH. Video denoising using shape-adaptive sparse representation over similar spatio-temporal patches. Signal Processing: Image Communication, 2011,26(4-5):250−265. [doi: 10.1016/j.image.2011.04.005]

[114] Hu SM, Yang YL, Lai YK. Research progress of digital geometry processing. Chinese Journal of Computers, 2009,32(8): 1451−1469 (in Chinese with English abstract). [doi: 10.3724/SP.J.1016.2009.01451]

[115] Sweldens W. The lifting scheme: A new philosophy in biorthogonal wavelet constructions. In: Laine AF, Unser MA, eds. Proc. of the SPIE 2569, Wavelet Applications in Signal and Image Processing III. Bellingham: SPIE Press, 1995. 68−79. [doi: 10.1117/12.217619]

[116] Qi DX, Tao CJ, Song RX, Ma H, Sun W, Cai ZC. Representation for a group of parametric curves based on the orthogonal complete U-system. Chinese Journal of Computers, 2006,29(5):778−784 (in Chinese with English abstract).

[117] Chui CK, Quak E. Wavelets on a bounded interval. Numerical Methods of Approximation Theory, 1992,9:53−75. [doi: 10.1007/978-3-0348-8619-2_4]

[118] Buhmann MD, Micchelli CA. Spline prewavelets for non-uniform knots. Numerische Mathematik, 1992,61(1):455−474. [doi: 10.1007/BF01385520]

[119] Sweldens W. The construction and application of wavelets in numerical analysis [Ph.D. Thesis]. Belgium: Katholieke Universiteit Leuven, 1995.

[120] Sweldens W. The lifting scheme: A custom-design construction of biorthogonal wavelets. Applied and Computational Harmonic Analysis, 1996,3(2):186−200. [doi: 10.1006/acha.1996.0015]

[121] Schröder P, Sweldens W. Spherical wavelets: Efficiently representing functions on the sphere. In: Cook R, ed. Proc. of the Computer Graphics Proceedings. New York: ACM Press, 1995. 161−172. [doi: 10.1145/218380.218439]

[122] Schröder P, Sweldens W. Spherical wavelets: Texture processing. In: Hanrahan P, Purgathofer W, eds. Proc. of the Eurographics Workshop on Rendering Techniques. New York: Springer-Verlag, 1995. 252−263. [doi: 10.1.1.48.8557]

[123] Li J, Song RX, Ye MJ, Liang YY, Qi DX. Orthogonal reconstruction of 3D model based on V-system over triangular domain. Chinese Journal of Computers, 2009,32(2):193−201 (in Chinese with English abstract). [doi: 10.3724/SP.J.1016.2009.00193]

[124] Qi DX, Song RX, Li J. Non-Continuous Orthogonal Function—U-Systems, V-System, Multi-Wavelet and Its Applications. Beijing: Science Press, 2011 (in Chinese).

[125] Huang C, Yang LH, Qi DX. A new class of multi-wavelet bases: V-System. Acta Mathematica Sinica, English Series, 2012,28(1): 105−120. [doi: 10.1007/s10114-012-9424-8]

[126] Song RX, Wang XC, Ou MF, Li J. The structure of V-system over triangulated domains. In: Chen FL, Jüttler B, eds. Proc. of the Advances in Geometric Modeling and Processing. LNCS 4975, Berlin: Springer-Verlag, 2008. 563−569. [doi: 10.1007/978-3-540-79246-8_48]

[127] Zhang J, Zhao DB, Zhao C, Xiong RQ, Ma SW, Gao W. Image compressive sensing recovery via collaborative sparsity. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2012,2(3):380−391. [doi: 10.1109/JETCAS.2012.2220391]

[128] Zhang J, Zhao DB, Jiang F, Gao W. Structural group sparse representation for image compressive sensing recovery. In: Bilgin A, Marcellin MW, Serra-Sagrista J, Storer JA, eds. Proc. of the Data Compression Conf. Los Alamitos: IEEE Computer Society Press, 2013. 331−340. [doi: 10.1109/DCC.2013.41]

[129] Liu XM, Zhai DM, Zhao DB, Gao W. Image super-resolution via hierarchical and collaborative sparse representation. In: Bilgin A, Marcellin MW, Serra-Sagrista J, Storer JA, eds. Proc. of the Data Compression Conf. Los Alamitos: IEEE Computer Society Press, 2013. 93−102. [doi: 10.1109/DCC.2013.17]

© 中国科学院软件研究所 http://www.jos.org.cn

[130] Shapiro JM. Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans. on Signal Processing, 1993,41(12): 3445−3462. [doi: 10.1109/78.258085]

[131] Po DDY, Do MN. Directional multiscale modeling of images using the contourlet transform. IEEE Trans. on Image Processing, 2006,15(6):1610−1620. [doi: 10.1109/TIP.2006.873450]

[132] Wang XH, Chen MY, Song CM, Xu MC, Fang LL. Contourlet HMT model with directional feature. Science China Information Sciences, 2013,43(5):626−643 (in Chinese with English abstract). [doi: 10.1360/112010-1288]

[133] Wang XH, Ni PG, Su X, Fang LL, Song CM. The nonsubsampled Contourlet HMT model. Science China Information Sciences, 2013,43(11):1431−1444 (in Chinese with English abstract). [doi: 10.1360/112012-584]

[134] Kamisli F. Intra prediction based on markov process modeling of images. IEEE Trans. on Image Processing, 2013,22(10):3916−3925. [doi: 10.1109/TIP.2013.2264679]

附中文参考文献:

[1] 高文,赵德斌,马思伟.数字视频编码技术原理.北京:科学出版社,2010.

[2] 焦李成,侯彪,王爽,刘芳.图像多尺度几何分析理论与应用——后小波分析理论与应用.西安:西安电子科技大学出版社,2008.

[4] 焦李成,谭山.图像的多尺度几何分析:回顾和展望.电子学报,2003,31(Z1):1975−1981.

[5] 王相海,孙强,宋传鸣,刘丹.基于多尺度几何分析的图像编码研究进展.计算机研究与发展,2010,47(6):1132−1143.

[80] 王相海.基于小波的视频空间可分级编码研究.计算机科学,2002,29(7):115−117.

[84] 宋传鸣,王相海.一种新的小波域视频可分级运动估计方案.计算机学报,2006,29(12):2112−2118.

[85] 宋传鸣.基于多尺度分析的视频可分级编码技术研究[博士学位论文].南京:南京大学,2010.

[90] 宋传鸣,王相海.小波域视频运动估计研究进展.计算机学报,2005,28(10):1716−1727.

[91] 王相海,宋传鸣.图像及视频可分级编码.北京:科学出版社,2009.

[114] 胡事民,杨永亮,来煜坤.数字几何处理研究进展.计算机学报,2009,32(8):1451−1469. [doi: 10.3724/SP.J.1016.2009.01451]

[116] 齐东旭,陶尘钧,宋瑞霞,马辉,孙伟,蔡占川.基于正交完备 U-系统的参数曲线图组表达.计算机学报,2006,29(5):778−784.

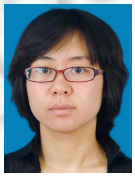
[123] 李坚,宋瑞霞,叶梦杰,梁延研,齐东旭.基于三角域上 V-系统的三维几何模型的正交重构.计算机学报,2009,32(2):193−201. [doi: 10.3724/SP.J.1016.2009.00193]

[124] 齐东旭,宋瑞霞,李坚.非连续正交函数——U-系统、V-系统、多小波及其应用.北京:科学出版社,2011.

[132] 王相海,陈明莹,宋传鸣,徐孟春,方玲玲.带方向特征的 Contourlet HMT 模型.中国科学(信息科学),2013,43(5):626−643. [doi: 10.1360/112010-1288]

[133] 王相海,倪培根,苏欣,方玲玲,宋传鸣.非下采样 Contourlet HMT 模型.中国科学(信息科学),2013,43(11):1431−1444. [doi: 10.1360/112012-584]

宋传鸣(1980－),男,辽宁沈阳人,博士,副教授,CCF 会员,主要研究领域为图像及视频编码,多尺度几何分析,多媒体水印技术.



刘丹(1981－),女,博士生,讲师,CCF 会员,主要研究领域为图像及视频编码,多尺度几何分析.

赵长伟(1989－),男,硕士生,主要研究领域为图像及视频编码,多尺度几何分析.

王相海(1964－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为图像及视频信息处理,遥感影像信息处理,计算机图形学.

© 中国科学院软件研究所 http://www.jos.org.cn