基于 Surfacelet 变换的 3D Context 模型视频去噪新方法

肖弘智¹,闫敬文¹,屈小波²

(1. 汕头大学电子工程系,广东汕头 515063; 2. 厦门大学通信工程系,福建厦门 361005)

摘 要: 本文将 3D Context 模型应用于 Surfacelet 变换域,提出一种新的视频去噪方法. Surfacelet 变换(ST) 是一种 新的 3D 变换,具有多方向分解、各向异性和低冗余度等性质. 根据视频信号 ST 域内系数和噪声分布的特征,将 2D Context 模型拓展到 3D,按照能量分布将 ST 系数分成多个子块,每个子块有独立的能量和阈值估计. 实验结果表明,本 文算法噪声抑制效果明显优于分层 2D 去噪声方法和其它现有的 3D 方法,去噪视频的 PSNR 值提高了约 2dB. 从视觉 效果来看,本文算法在去除噪声的同时,能很好的保留视频图像细节,运动物体非常平滑,有效解决传统算法中存在的 拖影、闪烁等问题,尤其适合于包含剧烈运动和丰富纹理图像的视频.

关键词: 视频去噪; Surfacelet 变换; 三维 Context 模型; 方向滤波器组 中图分类号: TN911.73 **文献标识码**: A **文章编号**: 0372-2112 (2008) 07-1460-05

A Novel Video Denoising Method with 3D Context Model Based on Surfacelet Transform

XIAO Hong-zhi¹, YAN Jing-wen¹, QU Xiao-bo²

(1. Department of Electronic Engineering, Shantou University, Shantou, Guangdong 515063, China;

2. Department of Communication Engineering, Xiamen University, Xiamen, Fujian 361005, China)

Abstract: We propose a novel video denoising method with 3D Context Model in Surfacelet Transform Domain (3DCMST) in this paper. In order to take advantage of the characteristic of the coefficients, the Context model was extended from 2D to 3D. The ST coefficients were divided into several parts according to their energy distribution by 3D Context model and each part had independent energy estimate and threshold. Experimental results show that the proposed method achieves better denoising performance than other 3D or hierarchical 2D denoising methods, and remarkably improves the PSNR of video about 2dB. In terms of visual quality, the proposed method can effectively preserve the video detail, and the trajectory of motion object is very smooth, which is especially adequate to process the video frames with acute movement and plenty of texture.

Key words: video denoising; Surfacelet transform; 3D Context model; directional filter bank

1 引言

传统的视频去噪方法基本上可以按空间域、时域、 频域和小波域来进行划分^[1].简单的空域滤波不考虑时 域信息,而帧之间同一位置的噪声存在随机性,因此容 易导致滤波后相邻帧之间明显的闪烁现象.简单的时域 滤波会由于运动的存在导致匹配失败或误差,出现噪声 残留或"鬼影"现象.频域滤波利用低通滤波器来去除集 中在高频段的噪声,不可避免会丢失部分细节信息,使 视频变得比较模糊.传统的小波域去噪通常是各帧去 噪,没有考虑各帧之间的运动相关性,会发生运动物体 拖尾的现象.

新的视频去噪算法可分为两类. 第一类算法是在传统算法的基础上增加运动检测和运动估计,利用各帧之

间的运动相关性,能够较好地解决运动图像拖尾和运动 边缘模糊的问题^[1,2].但是当视频图像中存在运动轨迹 模糊或者光照剧烈变化等现象时,这类算法的效果比较 差,鲁棒性不足^[3].第二类算法是把视频信号看成一个 特殊的三维信号,即两个空间维和一个时间维,采用三 维变换把它当成一个整体来处理.文献[4,5]对这类算 法进行了研究.这类算法对视频中的运动物体进行整体 处理,有效解决了前述算法中运动物体拖尾、闪烁以及 算法鲁棒性不足等问题.

绝大多数的三维变换都是可分一维变换的组合,不 是真正意义上的三维变换,这类变换得到的子带会混淆 三维信号的方向信息,不适合处理视频.文献[5]中采用 的三维双树复数小波把信号分解为6个方向,有较好的 方向选择性,但是当视频运动非常剧烈、图像纹理非常

基金项目:国家自然科学基金(No.60472081);航空科学基金(No.05F07001)

收稿日期:2007-08-14;修回日期:2008-01-02

丰富时,其方向数显然不足.Lexing Ying 和 Candes 等人 提出的 Curvelet 变换的三维形式^[6],即 3D-Curvelet,具有 良好的方向选择性,但是其过大的冗余度(约 40 倍)使 得算法效率非常低.

ST(Sufacelet Transfrom) 是 Yue Lu 和 M. N. Do 在 2007 年提出的一种新的三维变换^[7],具有多方向分解、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,非常适合于视频处理.本文提出一种 3DCMST (3D Context Model in ST domain)视频去噪算法,把视频信 号作为特殊的三维信号,对信号的空间和时间信息整体 处理.在阈值选取中采用 3D Context 模型,按照能量相 近与否把 ST 系数矩阵分成多个部分,每个部分都有独 立的能量估计和相应的阈值.实验表明,该方法能达到 更好的视觉效果和更高的 PSNR 值,尤其适用于包含剧 烈运动物体和丰富纹理图像的视频.

2 Surfacelet 变换

2.1 三维方向滤波器组

Bamberger 和 Smith 在 1992 年首先提出方向滤波器 组(DFB, Directional Filter Banks)的概念^[8].2002 年, M. N. Do 和 M. Vetterli 提出一种" 真"二维的图像表示方法,即 Contourlet 变换^[9]. Contourlet 变换首先利用拉普拉斯金字 塔对图像进行多尺度分解,再通过二维方向滤波器组来 实现完整的变换.

DFB 和 Contourlet 变换都只能处理二维信号,为了 处理三维甚至多维信号,Yue Lu 和 M. N. Do 在 2007 年 提出一种新的多维方向滤波器组设计方法——NDFB (N-dimensional Directional Filter Banks).本文只讨论三维 情况下的 NDFB,即 3D-DFB.图1(b)为 3D-DFB 的频率分 割示意图.



2.2 多尺度分解

因为 3D-DFB 只能处理信号的高频部分,在信号通 过 3D-DFB 之前,需要先进行多尺度分解.在 Contourlet 变换中,图像的多尺度分解由拉普拉斯金字塔(LP, Laplacian Pyramid)实现.而在 ST中,采用一种新的塔式 结构来实现,如图 2 所示.图中 $D_i()(i=0,1)$ 表示高 通滤波器, $L_i()(i=0,1)$ 表示低通滤波器,S()为反 混频滤波器,可以避免上抽样操作带来的混频现象.多 尺度分解和 3D-DFB 结合就构成完整的 ST 变换,其结构 图如图 2 所示.



2.3 ST的性质

ST 具有多方向分解、高效率的树结构滤波器组、完 全重建和低冗余度等性质,它能够用不同尺度、不同频 率的方向子块准确地捕获三维信号中的面状奇异,并且 变换后的 ST 系数能量非常集中.图 3 为对一个三维图 形进行 ST 分解后,选取不同的方向子块进行重建,重建 后的图形如图 3(b) ~ (d) 所示.从图中可以看出,ST 可 以选取不同的方向,并进行完全重建.



3 3DCMST 视频去噪算法

3.1 ST 系数分析

在阈值去噪算法中,阈值的选取最为关键.比如小 波域图像去噪,常根据小波分解后各层系数的分布情况 而选择不同的阈值.这种自适应阈值去噪算法在去除噪 声的同时还能较好地保留图像细节,因而得到了广泛的 应用.在讨论如何确定 3DCMST 去噪算法的阈值之前, 先对 ST 系数进行分析.

图 4 是对 Mobile 视频进行 ST 分解后,最精细层和 最粗糙层的某一方向子块的系数分布,图中 Y 表示系 数值, N 表示系数个数. 从图中可以看出,视频经 ST分 解后的系数变得较为稀疏,能量集中在有限的变换域系 数上,其余大部分变换域系数的幅值接近于零.

1462



图 5 是高斯白噪声进行 ST 分解后,任意选取两个 方向子块的系数分布.可以看出,高斯白噪声经 ST 分解 后能量均匀地分布在所有的变换域系数上,并且实验分 析可知其功率谱密度函数近似服从均匀分布,因此高斯 白噪声经 ST 分解后仍可看作高斯白噪声.



在实际应用中,还要考虑到算法的运算效率.表1 为针对不同大小数据,ST运行时间.运行环境为 CPU P4 2.8GHz,内存1.5GB,编译环境为 Visual C++6.0.

数据大小	层数 ×方向数						
$N \times N \times N$	3 ×1	3 × 4	3 ×16	3 × 64			
128 ×128 ×128	0.73	0.89	1.00	1.11			
192 ×192 ×192	2.88	3.78	4.03	4.33			
256 ×256 ×256	8.58	8.82	8.87	9.71			

表1 ST运行时间(秒)

3.2 3DCMST去噪算法的阈值确定

假设噪声模型为:

y(i,j,k) = f(i,j,k) + n(i,j,k) (1) 其中 f(i,j,k)为理想的不含噪声的视频经 ST 分解后的系 数, y(i,j,k)是观察到的 ST 系数. n(i,j,k)是服从独立同 分布的高斯噪声,即 $n(i,j,k) \sim N(0, ^2)$,均值为零,为 噪声等级. 若 f(i,j,k)和 n(i,j,k)不相关,从y(i,j,k)中 恢复 f(i,j,k)就是一个经典的统计估计问题.

通过分析视频经 ST 分解后的所有方向子块,发现 其系数分布都近似服从近似高斯分布或者近似拉普拉 斯分布^[10],即 GCD (General Gaussian Distribution)分布. Grace Chang 等人在文献[11]中专门对这种 GCD 模型的 阈值做了探讨.定义 "为噪声标准差, "为信号标准 差,则 GCD 模型的阈值表达式为:

$$T(\)={}^{2}_{n}\sqrt{} \tag{2}$$

其中 为 GOD 模型中的可变参数,它与信号标准差 " 的关系为:

$$x = 1/\sqrt{3}$$

所以式(1)可以重写为:

$$T() = r \times \frac{2}{n} / x$$
 (4)

其中 r 是可调节的参数,经多种测试发现 r=√2时能取 得最好的去噪效果.

从式(4)可知,如何估计噪声标准差 "和信号标准 差 "就成为阈值选取的关键.文献[12]中,给出了静止 图像中噪声标准差 "的估计式:

"=[^]"= median(| y(i,j,k)|/0.6745 (5)
其中 median 为求均值运算.在 ST 中,经过测试,发现式
(5) 的噪声估计也是可信的.在下一节将对信号标准差
"的估计进行讨论.

3.3 3D Context 模型建立

对含噪视频进行 ST 分解后的系数 y(i,j,k) 是信 号和噪声的混合矩阵,这个矩阵的方差为信号和噪声的 能量和,定义该矩阵方差为 var(y),则信号标准差 _x 可 用下式估计:

$$x = \sqrt{\operatorname{var}(y) - \frac{2}{n}} \tag{6}$$

由于在同一方向子块中 ST 系数变化范围很大,而式 (6) 通过求整个矩阵方差减去噪声能量来估计信号的标 准差,显然没有考虑到每个 ST 系数的情况,它只是一种 平均的方法.针对这个问题,本文提出一种 3D Context 模 型.Context 模型最早由文献[10]提出,主要针对静止图像 处理,本文算法把传统的Context 模型从 2D 拓展到 3D,能 够充分利用视频信号的三维时空信息,把能量相近的 ST 系数放在一起,求出这一部分系数的方差作为这一部分 系数信号和噪声的能量和.将 ST 系数矩阵分成多个部 分,每一个部分都有各自的能量估计和相应的阈值.

定义每一个 ST 系数的 context 值为 Z(i, j, k):

$$Z(i,j,k) = \frac{1}{N_{(i,j,k)}} | y(i,j,k) |$$
(7)

*S*为 ST 分解后方向子块中系数 *y*(*i*,*j*,*k*)的邻域,一般 取3 x3 x3 的方块, *N*为 *S*内的系数个数.图 6(*b*)为 *N* 取3 x3 x3 时系数 *y*(*i*,*j*,*k*)的 3D Context 模型图.实际 应用中,邻域 *S*需要根据视频内容作相应调整,当视频 内容较丰富时,邻域范围应取较小值.

作为一个例子,下面考察用 3D Context 模型是否能 衡量变换系数能量接近情况.如图 7,对含噪视频进行 ST 分解后,任选其中一个方向子块,发现系数越小它们 所对应的 context 值的分布就越密,而系数越大则 context 值的分布就越稀疏.由图 7 可见 Z(*i*,*j*,*k*) 值越小时,分



布的越密; Z(i, j, k) 值越大,则分布越稀疏.因此当 context 值小于 80 时,context 值每隔 5 分一级,如图 7 窗 口 a;当 context 值大于 80 时则增大分级间隔,如图 7 中 窗口 b.



图 7 ST 方向子块系数与其对应 context 值的分布关系图

3DCMST视频去噪算法具体步骤如下:

对含噪视频进行 ST 分解; 用 3D Context 模型将 ST 分解得到的各方向子块分 成多个部分,分别计算每一部分的方差 var(g)和噪声标准

差・

表 2 各种算法去噪结果(dB)

利用式(6) 计算每一部分的信号方差 x 及其相 对应的阈值 $T(x) = \sqrt{2} \frac{2}{n} / x;$

利用步骤 3 中得出的阈值分别对 ST 分解后各方 向子块进行去噪;

对去噪后的 ST 系数进行重构.

4 实验结果和分析

4.1 实验环境

本文采用 SIF 格式的标准视频序列进行实验,包括 小幅度运动的 Mobile、大幅度运动的 Football 和图像纹 理丰富的 Suise.首先对视频序列添加不同大小方差的 零均值高斯白噪声,然后采用基于小波去噪算法(WT)、 基于双树复数小波去噪算法(DTWT)、基于三维 Curvelet 去噪算法(CT)、基于硬阈值的 ST 去噪算法(ST)^[3]以及 本文的基于三维 Context 模型的 ST 去噪算法(3DCMST) 分别对含噪视频进行去噪.实验中,各种算法的变换均 只分解三层.为了考察分解层数对本文算法的影响,ST 分别采用三层和四层的分解,对应 3D-DFB 的方向数分 别为 64、16、4 和 64、64、16、4,只对最精细两层的方向子 块进行去噪处理.

4.2 实验结果分析和讨论

表2给出了五种算法去噪后 PSNR 值比较,图8为 Football 视频去噪效果图.从实验结果可以得到以下结论:

	Mobile			Football			Suise				
噪声标准差。	30	40	50	30	40	50	30	40	50		
Noisy	18.59	16.09	14.15	18.58	16.09	14.14	18.58	16.09	14.15		
WT	23.12	22.75	21.28	21.62	20.08	18.64	26.31	24.82	23.34		
СТ	23.54	23.19	22.86	23.16	22.11	21.14	26.79	25.31	24.76		
DTWT	24.56	23.41	22.58	23.67	22.14	20.90	27.83	26.35	24.78		
ST	25.86	24.23	23.15	24.98	23.58	21.98	28.69	26.97	25.56		
3DCMST(3)	26.48	25.08	23.89	26.31	24.89	23.51	29.78	28.16	26.88		
3DCMST(4)	27.54	26.68	25.49	27.87	26.43	25.03	31.41	29.72	28.44		

(1)本文算法在 PSNR 值上要明显优于其它四种算法,比如 Mobile 视频,本文算法比其它算法的 PSNR 提高了约2~3dB.基于硬阈值的 ST 算法的去噪效果已经优于传统算法,但无论是 PSNR 还是视觉效果,其效果都不如本文算法.

(2)本文算法在处理包含剧烈运动和丰富纹理图像 的视频时,效果更加明显.三个视频中,Mobile 属于小幅 度运动视频,视频中的所有物体作整体平移运动,各帧 之间差异较小;而 Football 属于大幅度运动视频,其中包 含的都是杂乱无章的大面积的运动;Suise 视频的运动幅 度较小,但是包含头发、皮肤等纹理丰富的图像.从表1 可以看出,本文算法在处理 Mobile 视频时,PSNR 值比 DTWT算法提高了约1.5~2dB,而在处理 Football 和 Suise 视频时,PSNR 值比 DTWT算法提高了约2~2.7dB.

(3) 从图 8 的去噪效果图可以看出,本文算法能够 更好地恢复图像细节信息,视觉效果明显优于其他算 法.限于客观条件,本文中只能给出视频中某一帧的效 果图,实际上,当播放完整的视频序列时,本文算法的 视觉效果会更加的突出,不存在其他算法中常见的拖 影、闪烁等现象.



1464



(a) Original Video







图 8 Football视频去噪效果图

(c) WT PSNR=21.62dB



(d) CT PSNR= 23.16dB

(g) 3DCMST(3) PSNR=26.31dB

(4) 当增加分解层数时,本文算法的性能会得到较 大提高,从表1可知,当 ST 的分解层数由3 层增加为4 层时,PSNR值提高了约1~1.5dB.而其它三种算法,尤 其是 WT和DTWT.在增加分解层数时.去噪效果只是略 有提升.实际上,因为WT和DTWT的分解方向数较少, 在分解层数达到一定程度时,再继续增加分解层数性 能提高非常小.

本文在国内首先提出基于 ST 的视频去噪算法 ,针 对 ST 系数特点,在阈值选取中采用 3D Context 模型.实 验结果表明,本文算法显著提高了去噪视频的 PSNR 值,能够很好的保留视频图像细节,运动物体非常平 滑,不存在传统算法中的拖影、闪烁等现象.因 ST 具备 的多方向性、各向异性和低冗余度等性质,使得其在视 频增强、运动目标提取、三维医学影像处理和三维数据 压缩等方面也有很好的应用前景.这将是我们下一步 研究工作的重点.

参考文献:

[1] 陈勇兵,郭晓云,沈未名,一种鲁棒性的基于运动估计的 自适应时空域视频去噪算法 [J]. 计算机应用, 2006, 26 (8) :1882 - 1887.

Chen Yong bing, Guo Xiao yun, Shen Wei-Ming. Robust adaptive spatio-temporal video denoising algorithm based on motion estimation [J]. Computer Applications, 2006, 26(8):1882 -1887. (in Chinese)

[2] F A Mujica, J P Leduc, R Murenzi, et al. A new motion parameter estimation algorithm based on the continuous wavelet transform [J]. IEEE Transactions on Image Processing, 2000, 9(4) : 873 - 888.

- [3] Yue Lu, M N Do. Video processing using the 3-Dimensional Surfacelet transform [A]. Invited Paper, Fortieth Annual Asilomar Conference on Signals, Systems and Computers [C]. Pacific Grove, USA, 2006. 883 - 887.
- [4] E Chang, A Zakhor. Subband video coding based on velocity filters [A]. IEEE International Symposium on Circuits and Systems [C]. San Diego, USA, 1992. 2288 - 2291.
- [5] I W Selesnick, K Y Li. Video denoising using 2D and 3D dualtree complex wavelet transforms [A]. Proc. of SPIE conference on Wavelet Applications in Signal and Image Processing X [C]. San Diego, USA, 2003. 607 - 618.
- [6] Lexing Ying, Laurent Demanet, Emmanuel Candes. 3D discrete Curvelet transform [R]. Applied and Computational Mathematics. California Institute of Technology, USA :2005

作者简介:



肖弘智 男,1984年出生于广东梅州,汕头 大学电子工程系硕士研究生.主要研究方向:图 像/视频处理、计算机视觉.

(下转第1440页)

E-mail:s_hzxiao @stu.edu.cn

男,1964年出生于吉 闫敬文(通信作者) 林磐石,博士,教授,博士生导师,汕头大学电子 工程系教授,中国图像图形学会理事.目前主要 研究方向:小波理论及其应用、超小波分析及应 用、图像处理、遥感图像处理等. E-mail : jwyan @stu. edu. cn

学习、帧转发和地址老化,所以 CPU 占用率都比较高. 使用基于 IVL 的非对称 VLAN 时,CPU 占用率基本保持 在 58 %以上;使用基于 CSVL 的非对称 VLAN 时,CPU 占 用率基本保持在 47 %以上.总体来说,使用基于 IVL 的 非对称 VLAN 时的 CPU 占用率要比使用基于 CSVL 的 非对称 VLAN 时高 12 %左右,而且变动相对要剧烈些.

从 FDB 中 MAC 地址表项的数量和 CPU 占用率可 以看出,在同等硬件条件下,运行基于 CSVL 的对称 VLAN 的交换机能胜任的网络规模至少是运行基于 IVL 的非对称 VLAN 交换机的两倍.基于 CSVL 的对称 VLAN 相较基于 IVL 的非对称 VLAN,前者有更好的扩展性和 适应能力.

5 总结

1440

本文首先介绍了非对称 VLAN 要解决的问题,然后 分析了基于 SVL 的非对称 VLAN 和基于 IVL 的非对称 VLAN 的不足.为此,本文提出了一种适用于类似非对 称 VLAN 应用的 CSVL 地址学习及转发模型.并以此为 基础,提出了基于 CSVL 的非对称 VLAN.通过实验验证 了基于 CSVL 的非对称 VLAN 相较基于 IVL 的非对称 VLAN,前者在不增加额外 MAC 地址表项的情况下,实 现了非对称 VLAN 的功能,有效降低了 CPU 的占用时 间,并具有更好的可扩展性和适应能力.

参考文献:

- Marina Smith. Virtual LANs : A Guide to Construction, Operation and Utilization [M]. McGraw-Hill Companies, 1997.
- [2] Gilbert Held. Virtual Lans: Construction, Implementation, and Management[M]. John Wiley & Sons, 1997.
- [3] IEEE 802. 1Q-2003, IEEE Standards for Local and metropolitan area networks : Virtual Bridged Local Area Networks [S].
- [4] L Sun, H Zhang, Z Wang, Z Wu. Asymmetric VLAN manag

(上接第 1464 页)

- [7] Yue Lu, M N Do. Multidimensional directional filterbanks and surfacelets [J]. IEEE Transactions on Image Processing, 2007, 16(4):918 - 931.
- [8] R H Bamberger, M J T Smith. A filter bank for the directional decomposition of images : Theory and design [J]. IEEE Transactions on Signal Processing, 1992, 40(4):882 - 893.
- [9] M N Do, M Vetterli. The contourlet transform : an efficient directional multiresolution image representation [J]. IEEE Transactions on Image Processing, 2005, 14(12):2091 - 2106.

ement protocol for distributed architecture [A]. Fifth International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 02) [C]. IEEE Press, 2002. 143 - 147.

- [5] R Perlman. Interconnections : Bridges , Routers , Switches , and Internetworking Protocols [M]. Amsterdam : Addison-Wesley , 2000.
- [6] Galileo Inc. GT-48360 Switched Gigabit Ethernet Controller with On-chip Memory and Layer 3 QoS[M/CD],2001.
- [7] Marvell Inc. GT-48370A/ GT48371A GalNet 2 + Octal Fast Ethernet Switch Controller with Integrated Transceiver and Memory, Revision A [M/ CD] ,2002.
- [8] VLAN : Virtual Local Area Network and IEEE 802. 1Q[OL]. http://www.javvin.com/protocolVLAN.html.
- [9] Broadcom Corporation. StrataXGS BCM5690 THEORY OF OPERATIONS [M/ CD] ,2002.

作者简介:



王显雷 男,1974年11月出生于内蒙古赤 峰.1997年毕业于华中科技大学计算机科学与 工程系,其后在NEC-软件所有限公司从事软件 开发工作,2001年进入中国科学院研究生院.现 为硕博连读生,从事计算机网络和多媒体通信有 关研究.

E-mail :xianlei01 @ios. cn



刘琼女,1959年3月出生于云南昆明. 博士,研究员,博士生导师.主要研究领域为多媒体通信,网络测量. Email:liug@iscas.ac.cn

- [10] S G Chang, B Yu, M Vetterli. Spatially adaptive Wavelet thresholding with context modeling for image denoising [J].
 IEEE Transactions on Image Processing, 2000, 9(9):1522 -1531.
- [11] S G Chang, B Yu, M Vetterli. Image denoising via lossy compression and wavelet thresholding [J]. IEEE Transactions on Image Processing, 2000,9(1):1532 - 1546.
- [12] David L, Donoho. De-noising by thresholding [J]. IEEE Transactions on Information Theory, 1995, 41 (3):613 - 627.