

基于非局部自相似序列集的一类视频图像盲去噪算法

邢远秀^{1,2}, 李军贤^{1,2}, 王文波^{1,2}, 王松³

(1. 武汉科技大学理学院, 湖北武汉 430081; 2. 冶金工业过程系统科学湖北省重点实验室, 湖北武汉 430081; 3. 南卡罗莱纳大学工程与计算学院, 美国南卡罗莱纳州哥伦比亚市 29208)

摘 要: 为提高具有帧间位移平移特性的视频图像的信噪比和去噪时效, 本文提出了基于非局部自相似序列集的视频图像盲去噪算法. 选取与待去噪视频图像前后相邻的若干图像帧, 在每一图像帧中寻找具有典型特征的图像块群, 并通过在前一帧图像中查找和该图像块群具有最小差异度的块群来确定帧间的精确位移; 将待去噪视频图像划分成若干图像块, 根据帧间位移快速构建每个图像块的自相似序列集; 随后将每个自相似序列集中的二维图像块整合成三维矩阵后进行三维变换, 并对变换系数进行自适应阈值处理; 再将三维逆变换后的图像块融合生成去噪图像. 实验结果表明, 在噪声方差未知的情况下, 本文算法所得去噪视频图像具有较好的信噪比和视觉效果, 并且有较高的运行效率.

关键词: 视频图像去噪; 块匹配; 自相似序列集; 三维变换; 自适应阈值

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2021)08-1498-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20191207

Blind Video Image Denoising Based on Nonlocal Self-Similarity Series Sets

XING Yuan-xiu^{1,2}, LI Jun-xian^{1,2}, WANG Wen-bo^{1,2}, WANG Song³

(1. College of Science, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China;

2. Hubei Province Key Laboratory of Systems Science in Metallurgical Process, Wuhan, Hubei 430081, China;

3. College of Engineering and Computing, University of South Carolina, Columbia, SC 29208, USA)

Abstract: A blind video image denoising algorithm based on non-local self-similar series sets is proposed to improve the peak signal-to-noise ratio and denoising efficiency of video images with displacement characteristics. Image block-groups with typical features are detected in each frame, and accurate inter-frame displacement between the image and previous frame is computed based on image block-matching. The noise image is divided into several image blocks, and the self-similar series set of each image blocks is constructed quickly according to the inter-frame displacement. A 3D transform is applied to the self-similar series set, followed by an adaptive-threshold of the transform coefficients to attenuate the noise. The 3D estimate after inverse 3D transformation are aggregated to obtain the final image. Experimental results show that the proposed algorithm has significant advantages in PSNR and visual effect in unknown noise variance, meanwhile it also has higher efficiency.

Key words: video image denoising; block-matching; self-similarity series set; 3D transform; adaptive threshold

1 引言

视频图像在采集、传输及存储过程中, 机械、光照等因素使得视频图像中往往包含较多的噪声, 影响人们对视频图像的分析与理解^[1]. 如何在保持图像真实信息不被破坏的同时有效降低图像中的噪声是图像处

理领域研究的主要目标之一.

目前图像去噪方法主要有基于深度学习的去噪方法、基于变换域的去噪方法、基于稀疏字典的去噪方法和基于非局部自相似性的去噪方法等. 随着大数据技术和计算机硬件的发展, 基于深度学习的图像去噪方

法引起了人们的关注. 该类方法^[2,3]主要利用神经网络构建去噪网络模型, 通过对噪声图像进行特征提取、非线性映射和图像重构实现图像去噪. 该类算法在利用训练集对去噪网络模型进行训练时, 训练集需要大量的噪声图像和对应的干净图像作为样本对, 在难以获得干净图像或者缺少训练集的情况下, 该类方法的去噪性能会受到较大的影响. 基于变换域的去噪方法^[4,5]将时域空间信号转换到变换域空间, 在变换域中进行滤波后再进行反变换重构图像信号, 该类方法过于依赖变换模型和阈值, 并且边缘可能存在一定的伪吉布斯现象或模糊现象. 基于稀疏字典的去噪方法^[6,7]构建与图像信息结构相适应的字典, 以尽可能少的原子组合来表示向量化后的图像, 该类方法需要构造过完备的字典, 字典的优劣影响去噪的性能.

基于非局部自相似性先验的图像去噪方法因为在去噪的同时能较好地保留图像细节, 可获得较好的性能, 因此涌现了大量的相关算法. 其中在基于非局部自相似均值 (Non-Local Means, NLM) 方法^[8]中, 当前像素的估计值由图像中与它具有相似邻域结构的像素加权平均求取, 由于需要对每个像素点计算它与搜索窗口内若干个像素间的相似度, 导致算法复杂度较高. 为了降低运算时间, 文献^[9]利用积分图快速进行相似度计算, 文献^[10]利用统计策略查找图像块的相似邻域, 有效提高了算法的时间效率. 为了进一步提高去噪效果, 一些算法^[11-15]将非局部自相似先验和图像的低秩特性相结合, 对非局部自相似块矩阵进行奇异值分解, 并对奇异值进行收缩, 将非局部自相似块滤波转化成低秩矩阵的近似问题实现图像去噪, 较好地挖掘了图像自身的信息和结构化稀疏特征. 还有一些算法将非局部自相似先验和结构稀疏表示相结合进行图像去噪, 其中经典的 BM3D 方法^[16]融合了时域和频域降噪方法, 构建三维相似图像块矩阵并进行 3D 变换, 在频域空间对 3D 变换系数进行阈值处理, 去噪的同时更好地保留了图像的细节, 是目前公认的最好的去噪方法之一. 随后大量改进的 BM3D 算法^[17-19]通过改进相似块的搜索过程或三维变换过程等提高算法的执行效率和去噪性能. 最近, 一些算法将机器学习方法应用到基于非局部自相似性去噪算法中, 比如文献^[20]提出的 SNSS 算法将噪声图像非局部自相似块的先验和外部干净图像的非局部自相似块的先验结合, 文献^[21]提出的 GSRC-NLP 算法利用图像的非局部自相似性先验和自监督学习方法对每个自相似块矩阵的稀疏性进行估计, 均获得了不错的去噪效果.

对于视频图像, 除了使用上述方法进行单帧图像去噪之外, 近年来基于视频序列的视频图像去噪方法

引起了研究者的关注^[22], 其中基于非局部自相似性先验的视频图像去噪算法在视频信息组成的三维空域时域上滤除噪声, 其中空域部分利用帧间冗余在相邻视频帧间搜索目标的相似区域, 时域部分则对目标进行运动预测. 经典的算法, 如文献^[23]使用光流法对视频进行运动估计, 得到目标块在前后帧的大概位置, 然后在相邻帧中进行相似块搜索, 并利用 NLM 方法进行视频图像去噪. 文献^[24]提出的 SALT 算法将基于低秩的图像去噪方法扩展到视频图像序列, 采用自适应稀疏性和低秩性结合的方法, 在相邻图像帧上进行非局部相似块搜索并进行低秩近似, 从而实现视频图像去噪. 文献^[25,26]提出的 VBM3D 算法基于结构稀疏特性, 利用块匹配方法在相邻帧中搜索相似图像块, 在变换域中通过阈值收缩变换系数来去除噪声. 文献^[27]提出的 VBM4D 算法通过追踪目标块在下一帧的位置, 估计目标块的运动轨迹, 将时域和空域上的相关性通过四维数组的方式直接联系起来, 去噪性能有较大的提升, 但该算法的计算量较大. 文献^[28]提出的 E-RF3D 算法基于运动补偿思想, 沿着运动轨迹搜索自相似块, 采用自适应三维阈值数组对自相似块的三维变换系数进行收缩实现图像去噪. 上述方法针对于特定应用的视频图像, 能够取得较高的峰值信噪比, 且视觉效果较好.

对于在生产生活中出现的大量由视频采集装置移动所拍摄的视频图像, 如提升矿山井筒设施监控视频图像和手机运动所拍摄的场景视频图像等, 图像帧间存在位移平移等特性. 针对此类视频图像, 本文提出了基于非局部自相似序列集的视频图像盲去噪方法, 通过块匹配方法快速精确地定位视频帧间位移, 在此基础上直接构建待去噪图像的自相似序列集, 然后对自相似序列集对应的三维矩阵进行自适应协同滤波, 最后聚合生成去噪后的图像. 和目前大多数基于非局部自相似的图像去噪方法相比较, 本文算法的主要优势有: 其一, 本文算法利用具有典型特征的图像块群快速确定帧间的精确位移, 在此基础上直接确定每个待去噪图像块的自相似序列集, 而不需要对每个待去噪图像块进行相似块的搜索, 提高了算法的运行效率; 其二, 自相似序列集中的图像块是在相邻帧中获取的, 同一场景下同一目标的图像块理论上具有较高的相似度, 提升了算法的去噪性能; 其三, 通过平均帧间差异度估计滤波系数, 从而实现了视频图像的盲去噪. 实验结果表明, 在图像噪声方差未知情况下, 本文算法在保持视频图像边缘及纹理特征不被破坏的同时, 大幅提高了去噪的时间效率.

2 视频图像盲去噪算法

2.1 自相似序列集构建

工程中所采集的视频图像包含的噪声常常是高斯噪声,在 t 时刻,定义被噪声污染的视频图像帧 u^t 在坐标 x 处的像素值为 $u(x; t) = y(x; t) + \eta(x; t)$, $x \in X, X \rightarrow R$, 其中, t 表示 t 时刻的图像帧,又表示图像帧所在的帧序号, $y(x; t)$ 为图像真实像素值, $\eta(x; t)$ 为高斯噪声,噪声服从 $\eta \sim N(0, \sigma^2)$ 正态分布.

第 t 时刻的视频图像帧 u^t 中包含 $N_1 \times N_1$ 个像素的参考图像块 u_{xR}^t , 和第 k 时刻的视频图像帧 u^k 中的图像块 u_x^k 之间的差异度 $\text{dis}(u_{xR}^t, u_x^k)$ 定义为

$$\text{dis}(u_{xR}^t, u_x^k) = \frac{\|T_{2D}(u_{xR}^t) - T_{2D}(u_x^k)\|_2^2}{(N_1)^2} \quad (1)$$

其中,下标 xR 和 x 分别为参考图像块 u_{xR}^t 和图像块 u_x^k 的左上角像素在相应图像中的二维坐标; $\|\cdot\|_2^2$ 为二维欧氏距离; T_{2D} 为二维离散余弦变换. 考虑噪声的影响,给定差异度阈值 T_m , 满足:

$$\text{dis}(u_{xR}^t, u_x^k) < T_m \quad (2)$$

则 u_{xR}^t 和 u_x^k 被判别为相似图像块.

将视频图像帧 u^t 以步长 N_{step} 划分为 $N_1 \times N_1$ 大小的图像块作为参考块,在以 u^t 为中心的前后连续 $2l+1$ 帧视频图像 $u^{t-l}, \dots, u^t, \dots, u^{t+l}$ 上,依次查找每个参考块的相似图像块并构建相应的自相似序列集. 本文假设视频是由采集装置自上向下运动采集,当给定图像帧间的位移 Δ 时,可快速确定视频图像帧 u^t 中的每个参考图像块 u_{xR}^t 在相邻的 $2l+1$ 图像帧中对应的相似图像块. 每个参考图像块 u_{xR}^t 的自相似序列集 S_{xR}^t 为

$$\begin{cases} x = xR - i \times \Delta \\ S_{xR}^t = \{(x, t+i) : \text{dis}(u_{xR}^t, u_x^{t+i}) < T_m\} \\ i = -l, -l+1, \dots, 0, \dots, l-1, l \end{cases} \quad (3)$$

S_{xR}^t 保存的是 $2l+1$ 个视频图像帧中和参考图像块 u_{xR}^t 相似的图像块的左上角像素坐标和所在的图像帧序号, $|S_{xR}^t|$ 表示 u_{xR}^t 的自相似序列集中相似块的个数,取值范围为 $1 \leq |S_{xR}^t| \leq 2l+1$.

2.2 帧间位移精确定位

图像块的自相似序列集主要由视频图像帧间的位移快速确定,帧间位移定位的准确度对后继去噪性能有着直接的影响. 本文利用图像中具有典型特征的图像块群来精确定位帧间位移.

帧间位移精确定位阶段的图像块大小可能与自相似序列集构建阶段的图像块大小不同,设该阶段的图像块大小为 $N_2 \times N_2$. 以确定视频图像帧 u^{t-1} 和 u^t 的帧间位移定位为例,为了防止光照等因素影响,对视频图

像帧 u^t 中待判别的图像块 u_x^t 进行归一化,生成 \bar{u}_x^t , 图像块的均值 μ_x^t 和均方差 s_x^t 为

$$\bar{u}_x^t(k) = \frac{u_x^t(k)}{\max_{\Omega_x} u_x^t} \quad (4)$$

$$\mu_x^t = \frac{1}{N_2^2} \sum_{k \in \Omega_x} \bar{u}_x^t(k) \quad (5)$$

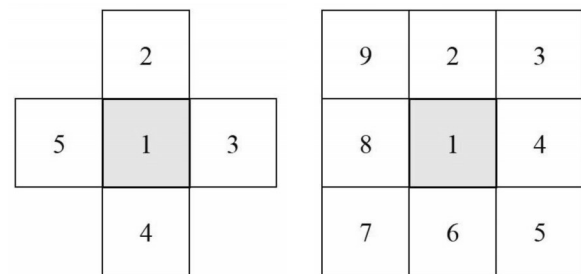
$$s_x^t = \frac{1}{N_2^2} \sum_{k \in \Omega_x} (\bar{u}_x^t(k) - \mu_x^t)^2 \quad (6)$$

其中, Ω_x 为图像块 u_x^t 中所有像素的坐标集合. 给定阈值 T_s , 待判别图像块 u_x^t 满足 $s_x^t > T_s$ 时被判别为典型特征图像块. 在图像帧 u^t 中,根据初始给定的帧间位移估计 Δ_{ini} , 在最大可能和前一图像帧 u^{t-1} 重合的区域内,从左到右、从上到下顺序搜索 n 个彼此不重叠的典型特征图像块,将其联合起来共同确定帧间位移.

设搜索到的第 i 个典型特征图像块 $u_{x_i}^t$ 的左上角像素坐标为 x_i , 根据帧间位移估计 Δ_{ini} , 映射到前一视频图像帧 u^{t-1} 中的坐标为 $x_i + \Delta_{\text{ini}}$. 在视频帧 u^{t-1} 中,以坐标 $x_i + \Delta_{\text{ini}}$ 为中心,在宽度为 $2\Delta_{\text{ini}}$ 的正方形邻域 $\Omega_{\Delta_{\text{ini}}}$ 内,搜索与 n 个典型特征图像块 $u_{x_i}^t$ 具有最小差异度的图像块,确定帧间的精确位移 $\hat{\Delta}^t$ 为

$$\hat{\Delta}^t = \arg \min_{\Delta, x_j + \Delta \in \Omega_{\Delta_{\text{ini}}}} \left(\sum_{i=1}^n \text{dis}(u_{x_i}^t, u_{x_i + \Delta}^{t-1}) \right) \quad (7)$$

为防止噪声干扰,以每一个典型特征块为中心,将与其直接相邻且大小相同的邻域图像块一起作为一个特征块群,来替代单一的典型特征块. 图1显示了两种不同类型的块群结构.

(a) $m=5$ (b) $m=9$ 图1 块群结构示例(m 为块群中块的个数)

序号为1的图像块为典型特征图像块,采用差异度加权均值,式(7)重写为

$$\hat{\Delta}^t = \arg \min_{\Delta, x_j + \Delta \in \Omega_{\Delta_{\text{ini}}}} \left(\sum_{i=1}^n \sum_{j=1}^m \text{dis}(u_{x_j}^t, u_{x_j + \Delta}^{t-1}) \times w_{ij} \right) \quad (8)$$

$$w_{ij} = \frac{\exp(-d_{ij})}{\sum_{j=1}^m \exp(-d_{ij})} \quad (9)$$

其中, x_{ij} 为第 i 个特征块群中序号为 j 的图像块的左上角像素坐标; w_{ij} 为第 i 个特征块群中序号为 j 的图像块的距离权重, 对于每个块群 i , 其典型特征图像块的块距离 $d_{i1} = 0$, 直接相邻的上下左右四个图像块距离 $d_{ij} = 1$, 四个角上的块距离 $d_{ij} = \sqrt{2}$. 利用典型特征块群和图像帧间位移, 求取视频图像帧 u^t 和 u^{t-1} 间的差异度为

$$\text{dist}(t, t-1) = \left(\sum_{i=1}^n \sum_{j=1}^m \text{dis}(u_{x_{ij}}^t, u_{x_{ij}+\hat{\Delta}^t}^{t-1}) \right) \times w_{ij} \quad (10)$$

当 $2l+1$ 帧的帧间精确位移 $\hat{\Delta}$ 被确定后, 利用式(3)可快速地构建自相似序列集 S_{xR}^t , t 时刻相邻的 $2l$ 个视频图像帧的平均帧间差异度定义为

$$\text{dist}^t = \frac{1}{2l} \sum_{k=t-l+1}^{t+l} \text{dist}(k, k-1) \quad (11)$$

在没有噪声的情况下 dist^t 的值为 0, 在存在噪声的情况下, dist^t 的大小和噪声的强度相关.

2.3 自相似序列集盲去噪

参考块 u_{xR}^t 的自相似序列集 S_{xR}^t 对应的每个图像块内差异主要由边缘和纹理等特征引起, 而不同图像块间同位置的像素差异主要由噪声引起. 如果 $|S_{xR}^t| < 2$, 则采用基于硬阈值的离散余弦变换, 反之则采用 3D 自适应滤波变换. 将 S_{xR}^t 中对应的图像块组合成三维矩阵 $U_{S_{xR}^t}^t$, 通过 3D 变换及逆变换后的三维矩阵 $\hat{Y}_{S_{xR}^t}^t$ 为

$$\hat{Y}_{S_{xR}^t}^t = T_{3D}^{-1} \left(\Upsilon \left(T_{3D} \left(U_{S_{xR}^t}^t \right) \right) \right) \quad (12)$$

其中, T_{3D} 为 3D 变换算子, 本文中采用 3D 傅里叶变换; T_{3D}^{-1} 为 3D 逆变换算子; Υ 为基于自适应阈值 λ_{3D} 的变换算子, 将 3D 变换后的系数小于阈值 λ_{3D} 的系数置为 0, 大于阈值 λ_{3D} 的系数保持不变. 实验获得多幅不同噪声方差的视频图像帧在较好去噪效果时的阈值 λ_{3D} 和对应的平均帧间差异度 dist^t . 对 λ_{3D} 和 dist^t 进行拟合, 得到自适应阈值 λ_{3D} 为

$$\lambda_{3D} = \min \left(\begin{array}{l} 0.1824 \times (\text{dist}^t)^2 + 0.7285 \times \text{dist}^t \\ +0.0394, 6.386 \end{array} \right) \quad (13)$$

三维矩阵 $\hat{Y}_{S_{xR}^t}^t$ 由 $|S_{xR}^t|$ 个经过 3D 滤波的二维图像块组成, 每个滤波后的二维图像块表示为 $\hat{Y}_{sn}^{t,xR}$, 下标 sn 表示二维图像块所在的帧序号, 同时将每个图像块 $\hat{Y}_{sn}^{t,xR}$ 中所有元素的坐标映射为和参考图像块 u_{xR}^t 对应的像素的坐标相同. 每个图像块 $\hat{Y}_{sn}^{t,xR}$ 的权重定义为

$$w_{xR}^t(sn) = \exp \left(- \left| \text{dis}(u_{xR}^t, u_k^{sn}) \times (t - sn) \right| \right) \quad (14)$$

其中, 图像块 $u_k^{sn}(k, sn) \in S_{xR}^t$ 为 $\hat{Y}_{sn}^{t,xR}$ 对应的滤波前的图像块. u_k^{sn} 和参考图像块 u_{xR}^t 的差异度越小, 且所对应的

两图像帧越邻近, 块 $\hat{Y}_{sn}^{t,xR}$ 对去噪视频图像的最终估计值贡献也越大.

将这些滤波后的图像块进行融合, 每个像素点 x 去噪后的像素灰度估计值 $\hat{y}^t(x)$ 定义为

$$\hat{y}^t(x) = \frac{\sum_{xR \in X} \sum_{sn \in S_{xR}^{t(2)}} w_{xR}^t(sn) \hat{Y}_{sn}^{t,xR}(x)}{\sum_{xR \in X} \sum_{sn \in S_{xR}^{t(2)}} w_{xR}^t(sn)}, \quad x \in X \quad (15)$$

其中, $S_{xR}^{t(2)}$ 表示保存图像块所在的帧序号.

3 实验结果与分析

3.1 评价指标

(1) 峰值信噪比 (Peak Signal-Noise Ratio, PSNR). 是目前评价图像去噪效果最常见的指标之一, PSNR 值越大表明去噪效果越好, 单位为 dB.

(2) 处理时间. 算法核心部分运算时间, 单位为 s.

(3) 视觉效果. 将去噪图像的整体结构是否完整、图像的边缘纹理细节是否得以保持等作为主观评价标准.

为了验证本文算法的有效性, 将本文算法与目前表现性能良好的 NL-Means 算法^[8]、FNL-Means 算法^[9]、WNNM 算法^[12]、BM3D 算法^[16]、SNSS 算法^[20]、GSRC-NLP 算法^[21]、SALT 算法^[24] 和 E-RF3D 算法^[28] 在峰值信噪比、处理时间和视觉效果上进行对比实验. 实验平台为 2.5GHz 双核 CPU 和 8GB 内存 PC, 编程语言为 Matlab.

3.2 实验数据

本文选取 2 张大小为 512×512 的具有丰富边缘和纹理特征的干净人物图像, 如图 2(a) 和图 2(b) 所示, 对每张图像进行裁剪, 生成包含 15 幅大小为 520×360 的具有 10 个像素的垂直位移平移的合成视频, 对每一个视频的 15 幅图像分别加载方差为 $\sigma = 5, 10, 15, \dots, 35, 40$ 的高斯噪声. 另外通过罐笼上采集的提升井筒设施视频^[29] 来验证算法的性能, 视频包含 4775 帧视频图像, 图 2(c) 显示了其中一帧视频图像帧, 在裁剪去除了时间水印等信息后, 视频图像帧的大小为 320×190 .

3.3 参数选择

本文算法需要设置 8 个参数, 在帧间位移精确定位阶段, 图像块的大小 N_2 为 16, 典型特征图像块群个数 n 为 2, 每个块群内块个数 m 为 5, 典型特征块阈值 T_s 为 0.06; 在视频图像帧去噪阶段, l 为 3, 即在连续 7 帧视频图像帧上构建自相似序列集, 步长 N_s 为 3, 待去噪图像块的大小 N_1 为 8; 在相似图像块判别时, 在图像的噪声方差小于 30 时差异度阈值 $T_m = 0.15$, 大于等于 30 时 $T_m = 0.20$.



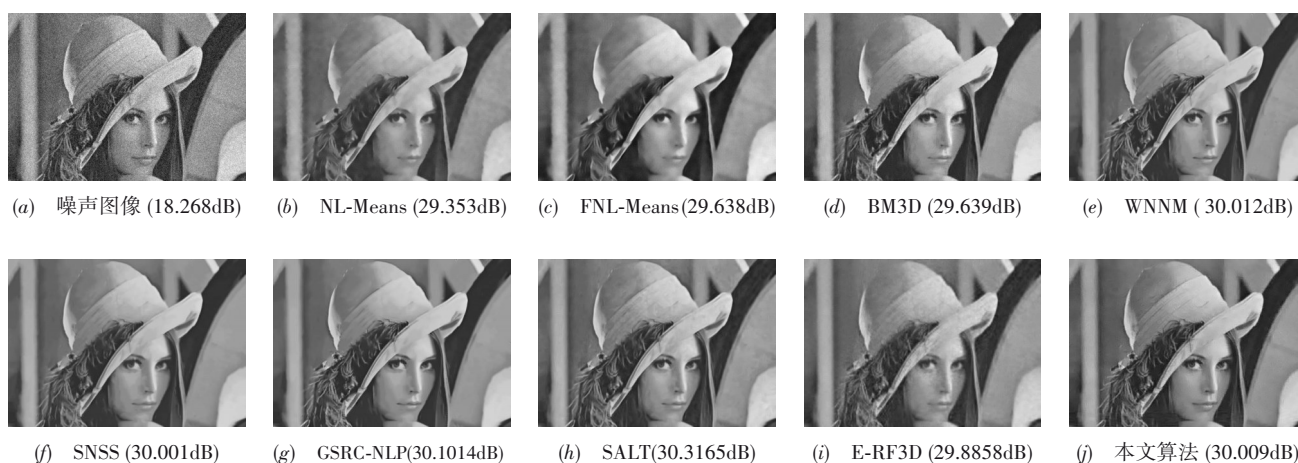
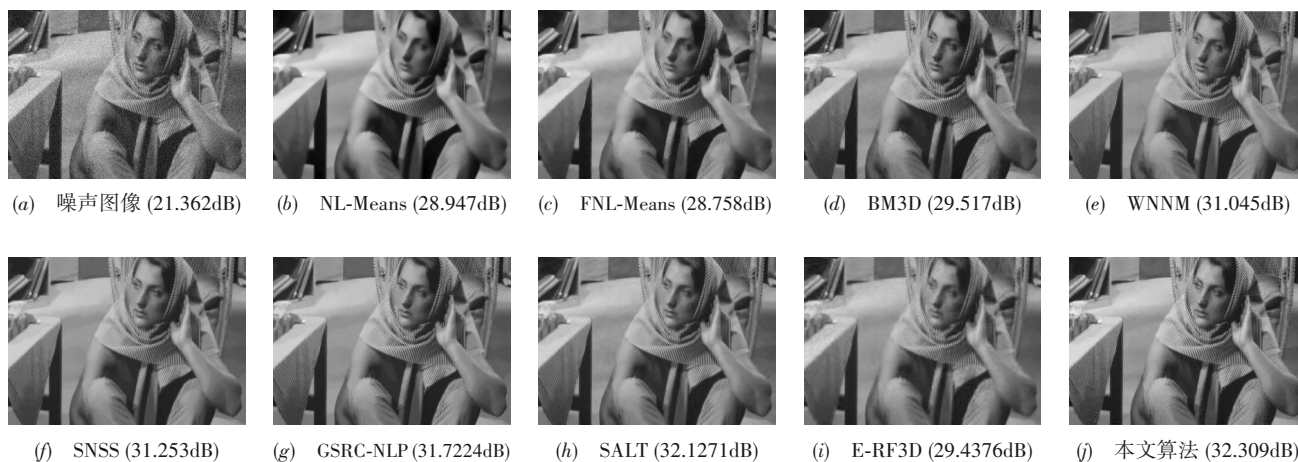
图2 测试图像

3.4 实验结果

3.4.1 合成视频图像上的去噪结果

为了验证算法的有效性,对叠加了不同噪声的

个合成人物视频的视频图像进行了去噪处理,9种算法的部分去噪结果如图3和图4所示,不同噪声方差下多帧视频图像帧上去噪所得的平均PSNR如图5所示.

图3 9种算法在高斯噪声方差 $\sigma=30$ 的人物图像1上的去噪效果图4 9种算法在高斯噪声方差 $\sigma=20$ 的人物图像2上的去噪效果

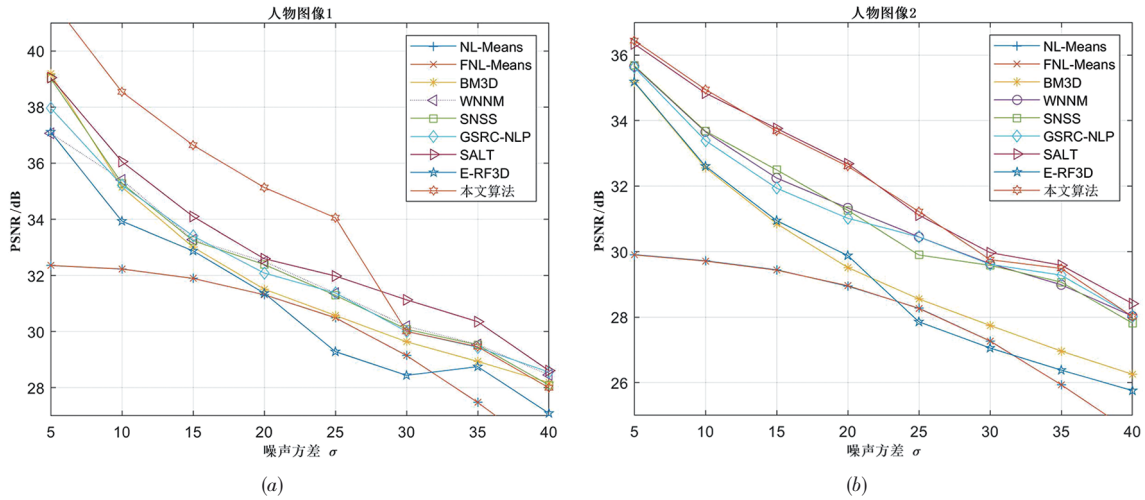


图5 9种算法在不同高斯方差下去噪后的平均信噪比

从图3~5中可以看出,随着噪声方差的增加,9种算法所得去噪图像的PSNR均呈现下降的趋势.在同一噪声方差下,NL-Means算法和FNL-Means算法所得去噪图像的PSNR几乎相同,当图像噪声方差较小时,PSNR值较高,同时去噪图像的视觉效果较好,但是随着噪声方差的增加,边缘和纹理细节信息减少,并出现块效应和模糊现象. BM3D算法的去噪效果比NL-Means算法的去噪效果好,可以较好地去除图像中大部分的噪声,并且在噪声方差较高的情况下,边缘和纹理细节仍保持较好,同时去噪后的图像视觉效果也好. E-RF3D算法在帧间进行快速块匹配时误差稍大,导致所得去噪图像边缘出现了轻微的伪影. 相比上述算法, WNNM算法、SNSS算法和GSRC-NLP算法在去除噪声的同时能够较好地保留图像的细节信息,使得去噪后图像的视觉效果较好. SALT算法和本文算法所得去噪图像的PSNR值比较高,噪声去除得较为干净,同时边缘纹理细节保持较好,主要原因是SALT算法在相邻视频帧上获得的相似块的质量比较好,同时

采用了稀疏表示和低秩表示. 而本文算法进行了精确的位移定位,使得自相似序列集的图像块间相似度相对较高,同时利用帧间的差异度估计噪声的强度,对3D变换系数进行自适应阈值处理,弥补了BM3D算法需要给定噪声方差的不足.

表1给出了9种算法在人物图像上多次运行的平均处理时间,可以看出,SNSS算法、WNNM算法和GSRC-NLP算法进行多次迭代运算,导致算法运行时间较长. SALT算法两次在相邻多帧视频图像上进行相似块的查找,同样需要较长的运行时间. BM3D算法和NL-Means算法需要在大窗口进行相似块的搜索,处理时间也较长. FNL-Means算法利用积分图像进行处理,使得算法的运行时间仅需要2s左右. E-RF3D算法通过估计帧间位移,在较小范围内进行相似块查询,大幅提高了算法的执行效率,算法所需平均运行时间仅需4s左右. 本文算法在确定了帧间位移后,直接构建自相似序列集,从而节约了大量寻找相似块的时间,算法整体运行效率也有了大幅提高.

表1 9种不同算法在人物图像上所需平均处理时间(单位:s)

	NL-Means	FNL-Means	BM3D	WNNM	SNSS	GSRC-NLP	SALT	E-RF3D	本文算法
人物图像1	70.068	2.347	124.503	327.614	977.470	212.411	334.612	4.011	9.1684
人物图像2	67.757	2.228	121.449	345.235	946.561	203.144	301.087	4.039	9.0511

3.4.2 真实视频图像上的去噪结果

图6和表2给出了提升矿山井筒视频图像帧的部分去噪结果. 由于复杂环境下获得的井筒设施视频图像无对应的干净图像,并且存在采集装置抖动及运行速度不均匀等情况,因此无法明确给出图像帧间的实际位移,也无法计算去噪后图像的信噪比. 由于井筒视频图像的噪声方差未知,根据经验设定 $\sigma=10$,从视觉效果可以看出, BM3D算法、WNNM算法、GSRC-NLP算法、SNSS算法、

SALT算法和本文算法都能较好地保持对比度较低的设施的纹理细节,在进行噪声方差估计时,噪声方差估计过小则不能较好地去除噪声,噪声方差估计过大图像细节又会出现模糊现象,而本文算法在去噪时无须考虑图像的噪声方差也可以获得较好的去噪结果. 同时由表2可知,SNSS算法去噪所需时间最长,本文算法、FNL-Means算法和E-RF3D算法所需时间相对较少. 总体来说,本文提出的算法在处理帧间存在位移差的视频图像时表现较好.

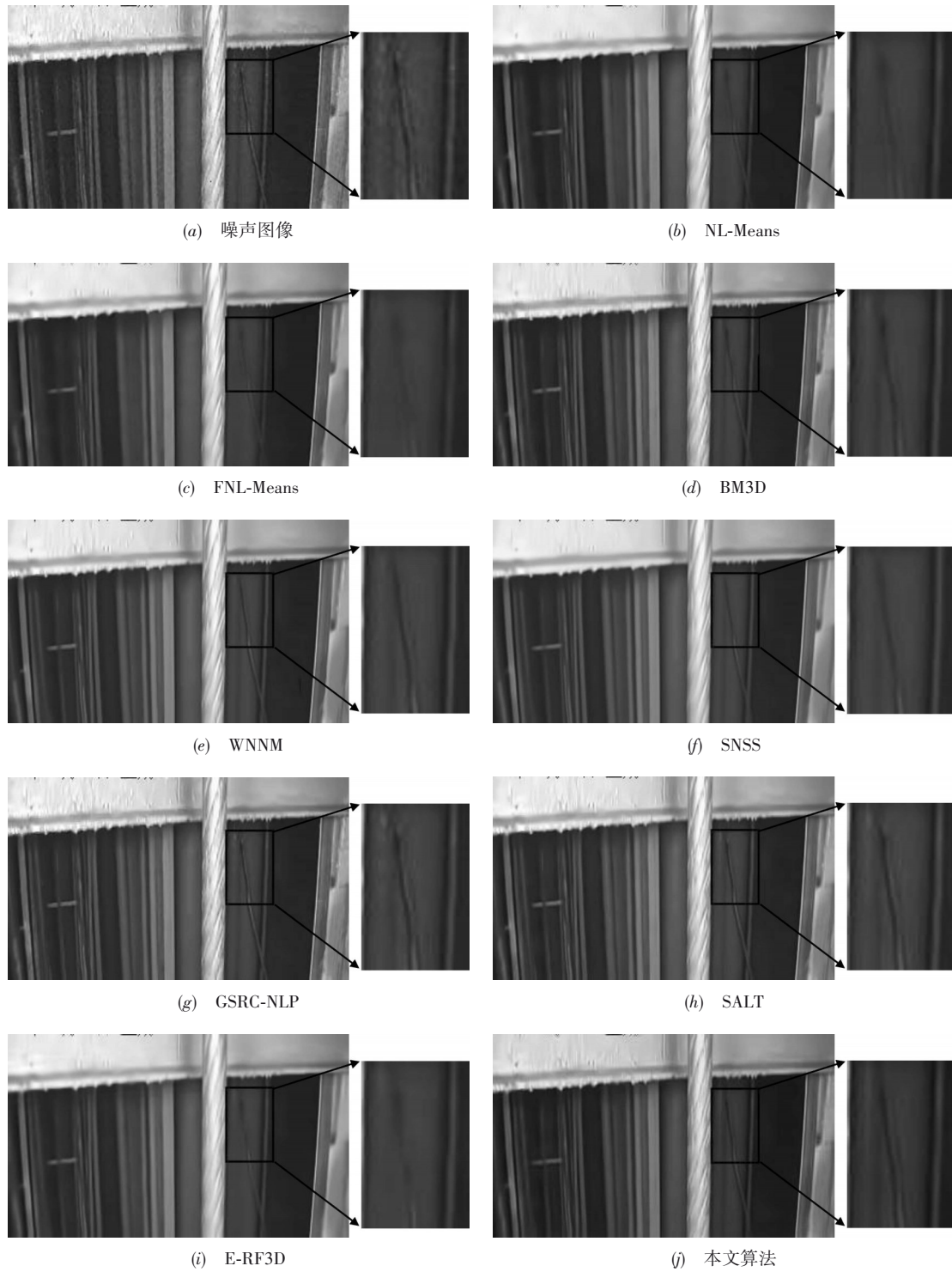


图6 9种算法在井筒视频图像帧上的去噪效果

表2 9种算法在井筒视频图像帧上所需平均处理时间(单位:s)

	NL-Means	FNL-Means	BM3D	WNNM	SNSS	GSRC-NLP	SALT	E-RF3D	本文算法
时间	20.065	0.397	38.775	88.217	187.384	54.221	76.862	1.5756	2.7837

4 结论

本文针对帧间具有位移平移特性的视频图像,利用帧间信息冗余,在确定了前后相邻的若干视频图像帧的帧间精确位移后,快速构建待去噪视频图像帧的自相似序列集,并根据帧间差异度估算 3D 变换时的自适应阈值,实现了视频图像帧的盲去噪. 本文算法对于有位移平移特性的视频图像去噪具有较好的信噪比和较好的视觉效果,同时具有较高的运行效率,能有效应用于如矿山竖井井筒视频监测和场景拼接等领域的视频图像快速去噪. 但是当噪声方差大于 40 时,位移定位精度受到较大影响,导致算法去噪性能不理想,还需要进一步研究和改进.

参考文献

- [1] Dodge S, Karam L. Understanding how image quality affects deep neural networks[A]. 2016 Eighth International Conference on Quality of Multimedia Experience[C]. Lisbon, Portugal: IEEE, 2016. 1 – 6.
- [2] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[A]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C]. Las Vegas, USA: IEEE, 2016. 770 – 778.
- [3] Chen J W, Chen J, Chao H Y, et al. Image blind denoising with generative adversarial network based noise modeling [A]. 2018 IEEE Conference on Computer Vision and Pattern Recognition[C]. Salt Lake City, USA: IEEE, 2018. 3155 – 3164.
- [4] Foi A, Katkovnik V, Egiazarian K. Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images[J]. IEEE Transactions on Image Processing, 2007, 16(5): 1395 – 1411.
- [5] Yang J H, Feng R, Deng W. A new algorithm of image denoising based on stationary wavelet multi-scale adaptive threshold[A]. Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology[C]. Harbin, China: IEEE, 2011. 4550 – 4553.
- [6] Yang J, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861 – 2873.
- [7] Wen B, Ravishanker S, Bresler Y. Structured overcomplete sparsifying transform learning with convergence guarantees and applications[J]. International Journal of Computer Vision, 2015, 114(2/3): 137 – 167.
- [8] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising[A]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) [C]. San Diego, USA: IEEE, 2005. 60 – 65.
- [9] Froment J. Parameter-free fast pixelwise non-local means denoising[J]. Image Processing on Line, 2014, 4: 300 – 326.
- [10] Frosio I, Kautz J. Statistical nearest neighbors for image denoising[J]. IEEE Transactions on Image Processing, 2019, 28(2): 723 – 738.
- [11] Wang S L, Zhang L, Liang Y. Nonlocal spectral prior model for low-level vision[A]. Computer Vision - ACCV [M]. Berlin, Heidelberg, GER: Springer, 2013. 231 – 244.
- [12] Gu S H, Xie Q, Meng D Y, et al. Weighted nuclear norm minimization and its applications to low level vision[J]. International Journal of Computer Vision, 2017, 121(2): 183 – 208.
- [13] 宋云, 李雪玉, 沈燕飞, 等. 基于非局部相似块低秩的压缩感知图像重建算法[J]. 电子学报, 2017, 45(3): 695 – 703.
Song Y, Li X Y, Shen Y F, et al. Compressed sensing image reconstruction based on low rank of non-local similar patches[J]. Acta Electronica Sinica, 2017, 45(3): 695 – 703. (in Chinese)
- [14] Zha Z Y, Yuan X, Wen B H, et al. From rank estimation to rank approximation: Rank residual constraint for image restoration[J]. IEEE Transactions on Image Processing, 2020, 29: 3254 – 3269.
- [15] Wen B H, Li Y J, Bresler Y. Image recovery via transform learning and low-rank modeling: The power of complementary regularizers[J]. IEEE Transactions on Image Processing, 2020, 29: 5310 – 5323.
- [16] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. IEEE Transactions on Image Processing, 2007, 16(8): 2080 – 2095.
- [17] Djurović I. BM3D filter in salt-and-pepper noise removal [J]. EURASIP Journal on Image and Video Processing, 2016, 10(1): 1 – 11.
- [18] Li Y J, Zhang J W, Wang M N. Improved BM3D denoising method[J]. IET Image Processing, 2017, 11(12): 1197 – 1204.
- [19] Yang J Y, Zhang X, Yue H J, et al. IBM3D: Integer BM3D for efficient image denoising[J]. Circuits, Systems, and Signal Processing, 2019, 38(2): 750 – 763.
- [20] Zha Z Y, Yuan X, Zhou J T, et al. Image restoration via simultaneous nonlocal self-similarity priors[J]. IEEE

- Transactions on Image Processing, 2020, 29: 8561 – 8576.
- [21] Zha Z Y, Yuan X, Wen B H, et al. Group sparsity residual constraint with non-local priors for image restoration [J]. IEEE Transactions on Image Processing, 2020, 29: 8960 – 8975.
- [22] Buades A, Lisani J L. Enhancement of noisy and compressed videos by optical flow and non-local denoising [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(7): 1960 – 1974.
- [23] Liu C, Freeman W T. A high-quality video denoising algorithm based on reliable motion estimation[A]. European Conference on Computer Vision[C]. Berlin, Germany: Springer, 2010. 706 – 719.
- [24] Wen B H, Li Y J, Pfister L, et al. Joint adaptive sparsity and low-rankness on the fly: An online tensor reconstruction scheme for video denoising[A]. 2017 IEEE International Conference on Computer Vision[C]. Venice, Italy: IEEE, 2017. 241 – 250.
- [25] Dabov K, Foi A, Egiazarian K. Video denoising by sparse 3D transform-domain collaborative filtering[A]. 2007 15th European Signal Processing Conference[C]. Poznań, Poland: IEEE, 2010. 145 – 149.
- [26] 肖进胜, 姜红, 彭红, 等. 一种改进的3维块匹配视频去噪算法[J]. 四川大学学报(工程科学版), 2014, 46(4): 81 – 86.
- Xiao J S, Jiang H, Peng H, et al. An improved video denoising algorithm based on 3D block matching[J]. Journal of Sichuan University (Engineering Science Edition), 2014, 46(4): 81 – 86. (in Chinese)
- [27] Maggioni M, Boracchi G, Foi A, et al. Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms[J]. IEEE Transactions on Image Processing, 2012, 21(9): 3952 – 3966.
- [28] Maggioni M, Sánchez-Monge E, Foi A. Joint removal of random and fixed-pattern noise through spatiotemporal video filtering[J]. IEEE Transactions on Image Processing, 2014, 23(10): 4282 – 4296.
- [29] 邢远秀. 矿山井筒视觉监测与故障智能识别系统设计[J]. 金属矿山, 2013, (9): 138 – 141.
- Xing Y X. Design of the vision monitoring and fault intelligent recognition system for security of mineshaft[J]. Metal Mine, 2013, (9): 138 – 141. (in Chinese)

作者简介



邢远秀(通信作者) 女, 1980年8月出生
于河南省南阳市. 博士, 武汉科技大学理学院副
教授, 硕士生导师. 主要研究方向为图像处理、
机器视觉等.

E-mail: yuanxiu@126.com



李军贤 女, 1994年1月出生于河南省安
阳市. 现为武汉科技大学硕士研究生, 主要研究
方向为图像处理、计算机视觉.

E-mail: 1156325080@qq.com