

非局部低秩正则化视频压缩感知重构

田金鹏^{1,2†}, 杨洁¹, 刘通¹, 闵天¹

(1. 上海大学通信与信息工程学院, 上海 200444; 2. 上海大学特种光纤与光接入网省部共建重点实验室, 上海 200072)

摘要: 视频压缩感知在采样资源受限的视频采集领域具有重要研究意义, 重构算法是视频压缩感知系统的关键技术. 为了更好地从压缩采样数据中重构视频信号, 本文提出一种基于全变分与非局部低秩正则化的视频重构算法, 为视频重构提供一种新的思路. 该算法包括两个步骤: 第一步考虑视频帧内局部光滑特性和帧间相关性, 应用全变分模型作为先验约束得到初步恢复的视频帧. 第二步考虑视频帧内和帧间的非局部自相似性, 应用改进的非局部低秩正则化算法对其进一步重构, 该步骤对初步恢复的图像帧分块, 在本帧和关键帧中寻找相似块, 构建低秩矩阵进行低秩正则化重构. 仿真结果表明, 提出的算法能够精确重构视频信号, 相比主流的视频压缩感知重构算法具有更高的重构质量.

关键词: 视频压缩感知; 非局部自相似性; 块匹配; 低秩正则化; 全变分; 先验约束

中图分类号: TN911.73

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.0310

Compressive Video Sensing Reconstruction via Nonlocal Low-Rank Regularization

TIAN Jin-peng^{1,2†}, YANG Jie¹, LIU Tong¹, MIN Tian¹

(1. School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China; 2. Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai 200072, China)

Abstract: Compressive video sensing (CVS) has great research significance in the video acquisition system with limited sampling resources. In this paper, we proposed a reconstruction algorithm based on total variation (TV) and nonlocal low-rank regularization (NLR-CS) to better reconstruct video signal from compressive sampled data. This algorithm consists two steps: The first step considers local correlation between and within video frames, applies TV as the prior constraint to obtain the initial recovered frame; In the second step, the improved NLR-CS algorithm is utilized to further reconstruct video frame considering the nonlocal self-similarity (NLSS). This step first blocks the initial recovered frame, finds similar blocks in the current frame and the key frames to construct low-rank matrix, then a low-ranking regularization reconstruction is performed. Experimental results show that the proposed algorithm can reconstruct video signals well, obtains higher video reconstruction accuracy than other CVS reconstruction algorithms.

Keywords: compressive video sensing; nonlocal self-similarity; block matching; low-rank regularization; total variation; priori constraint

0 引言

压缩感知^[1-2](Compressive Sensing, CS) 是一种重要的信号采集方式, 在信号稀疏前提下, 能够以远低于 Nyquist 采样定理的采样率对信号进行压缩采样, 应用重构算法可以从采样数据中重构原始信号. 该技术实现了信号采样与压缩同时进行, 提高采样速率的同时减少了传输数据量, 被广泛应用于医学诊断^[3]、激光雷达^[4]等成像领域, 在含有大量冗余数据的视频采样系统中也具有较大应用前景^[5-6]. 传统的视频采样方案通常首先对视频信号进行高速采

样, 然后应用复杂的压缩编码算法丢弃大量冗余数据, 以便于视频的存储与传输. 其中压缩编码步骤需要消耗大量计算资源, 不适用于计算能力有限的低成本视频设备. 而视频压缩感知技术^[7](Compressive Video Sensing, CVS) 可以提供一种有效的视频采样方案. CVS 系统的信号采集端对视频信号进行低速压缩感知采样, 信号解码端根据视频信号特性设计重构算法从采样数据中恢复视频信号. 该技术有效地将计算复杂度从视频采集端转移到了视频解码端, 近年来得到广泛研究.

收稿日期: 2020-03-19; 修回日期: 2020-07-22.

基金项目: 国家自然科学基金项目 (61871261), 上海市科委重点项目 (19DZ1205802)

†通讯作者. E-mail: adaline@163.com.

CVS 系统信号采集端的压缩采样容易实现, 一般应用高斯随机测量矩阵对信号进行随机线性采样, 系统主要难点在于解码端重构算法的设计. CS 理论要求信号在变换域中具有稀疏表示, 这是 CS 重构的先验条件, 所以视频信号的稀疏先验信息对视频信号的重构具有重要作用. 对于具有显著结构化特征和高度数据冗余性的视频信号来说, 如何充分挖掘其先验信息设计合理的约束条件是视频信号精确重构的关键. 文献 [8] 基于稀疏表示, 联合所有的图像帧构建三维稀疏变换模型 3D-DWT 对视频信号进行总体恢复, 此类算法计算复杂度过高, 且重构效果不够理想. 文献 [9] 将传统视频编码中的运动估计和运动补偿应用到 CVS 中, 结合运动估计和补偿产生的残差模型, 交替重建视频序列的帧和相应的运动场. 文献 [10] 提出了一种基于多假设预测的“预测-残差”重构模型, 其核心思想是从一个或多个已经重构的参考帧中选取相似块组, 通过求解最优化问题得到预测块, 然后对预测的残差进行重构得到当前帧, 取得了较好的重构性能. 目前主流的 CVS 重构算法大都基于多假设预测算法, 如文献 [11] 提出的两级多假设预测算法 2sMHR 分别在观测域和像素域进行多假设预测, 由于在像素域进行了重叠分块, 有效避免了块效应. 文献 [12] 提出的加权残差稀疏算法 RRS 通过交替帧间多假设预测和残差加权稀疏建模, 迭代地提高重建质量. 上述多假设预测算法虽然在不同程度上提高了视频重构质量, 对视频帧间的相关性进行了充分的挖掘, 但忽略了本帧图像所具有的特性. 文献 [13] 提出的基于结构相似的帧间组稀疏表示重构算法 SSIM-InterF-GSR 同时考虑了帧内图像的组稀疏性和帧间相关性, 得到了更高的重构精度.

视频信号的本质是连续的图像序列, 既具有图像空间上的冗余, 又具有序列时间方向上的冗余. 近年来基于图像局部光滑特性^[14]、非局部相关性^[15,16]等先验的图像 CS 重构算法取得了很好的重构效果. 先验图像约束压缩感知算法^[17,18](Prior Image Constrained Compressed Sensing, PICCS) 考虑了视频帧内和帧间的局部光滑特性, 采用全变分^[14](Total Variation, TV) 模型作为先验约束, 算法复杂度低重构速度较快, 被用于医学动态 MRI 和 CT 图像重构. 文献 [15] 进一步挖掘了图像的非局部自相似性^[19](Nonlocal Self Similarity, NLSS), 提出了一种非局部低秩正则化压缩感知重构算法 (Nonlocal Low-Rank Compressive Sensing, NLR-CS), 针对样本块寻

找相似块组成低秩矩阵, 把最小低秩作为约束构建图像重构的目标函数, 获得了目前为止图像 CS 的最优重构性能. 基于以上分析, 本文充分考虑视频帧内和相邻帧间的局部光滑特性和非局部自相似性, 提出了一种基于时空全变差和非局部低秩正则化的视频压缩感知重构算法. 主要创新点包括: (1) 考虑图像的局部光滑特性和相邻帧间的高度相关性, 采用 PICCS 算法对图像帧进行初步恢复, 将基于图像梯度的 TV 约束模型用于本帧图像和相邻帧间差的差值, 得到的梯度图像具有较强的稀疏性; (2) 考虑视频帧内和帧间的 NLSS 特性, 对 NLR-CS 算法进行改进, 对初步恢复的图像帧进行分块后, 在本帧和相邻的两个关键帧中寻找相似块构建低秩矩阵, 对初步恢复的图像帧进一步低秩正则化重构. 由于关键帧的重构质量较高, 充分利用关键帧的信息能够更好地重构本帧图像.

1 视频压缩感知相关理论

1.1 压缩感知理论

CS 理论表明, 如果信号在某个变换域具有稀疏表示, 则可以通过一个测量矩阵将该信号线性映射为低维的测量向量, 该测量向量保存了原始信号足够的有用信息, 利用重构算法可以从中恢复原始信号. 对于 N 维稀疏信号 x , 用测量矩阵 $\Phi \in R^{M \times N}$, $M \ll N$ 进行压缩采样, 可以表示为:

$$y = \Phi x. \quad (1)$$

其中采样数据 y 的维度为 M , 数据量远小于 N 维信号 x , 观测矩阵 Φ 实现了从高维信号 x 到低维测量向量 y 的线性映射, 从而实现了信号的低采样率采样. CS 重构通过求解欠定方程 (1) 从测量向量 y 中恢复信号 x , 常用的优化模型为:

$$x = \arg \min_x \frac{1}{2} \|y - \Phi x\|_2^2 + \lambda \Psi(x). \quad (2)$$

其中 $\Psi(x)$ 是正则项先验约束, 用来保证图像在某个变换域中的稀疏性, λ 是保真项和正则项之间的调和参数.

1.2 TV 模型

自然图像普遍具有局部光滑特性, 通过相邻像素的减法运算得到的梯度图像具有较强的稀疏性. TV 模型根据图像在梯度域稀疏提出, 表示为:

$$TV(x) = \sum_i \|D_i^h x + D_i^v x\|_p = \sum_i \|D_i x\|_p. \quad (3)$$

其中 x 为图像, i 为像素点坐标, $D_i^h x$ 和 $D_i^v x$ 分别表示像素点 i 的水平方向梯度和垂直方向梯度. p 取 1 时对应 l_1 范数, 求取各向异性全变差 (anisotropic

TV), 即水平梯度与垂直梯度的绝对值之和; p 取 2 时对应 l_2 范数, 求取各向同性全变差 (isotropic TV). PICCS 算法同时考虑了本帧图像空间上的局部相关性和相邻帧间的局部相关性, 将 TV 模型作为本帧和相邻帧差的稀疏约束。

1.3 NLR-CS 算法

NLR-CS 算法^[16] 考虑了图像的 NLSS 特性^[9], 把图像的重构问题转换为求图像中各相似块矩阵的低秩问题, 是目前自然图像 CS 重构效果最好的算法. NLR-CS 算法包括 3 个步骤: 基于 NLSS 性质的块匹配、对相似块矩阵进行低秩优化、应用低秩矩阵进行图像恢复. 针对初步重构的图像, NLR-CS 算法首先应用块匹配对每个样本块 (尺寸为 $n \times n$) 在 k 邻域内寻找 $m - 1$ 个相似块, 并将这 m 个相似块按列展开聚合为 $n^2 \times m$ 的矩阵 $X_i = [x_{i0}, x_{i1}, \dots, x_{i(m-1)}]$, 其中 i 为像素点坐标, x_{i0} 为样本块本身, $x_{i1} \sim x_{i(m-1)}$ 为相似块. X_i 的列向量具有高度相似性, 因此 X_i 具有低秩属性. 考虑噪声, 令 $X_i = L_i + W_i$, 其中 L_i 表示低秩矩阵, W_i 表示噪声. L_i 可以通过优化式 (4) 求解:

$$L_i = \arg \min_{L_i} \text{rank}(L_i). \text{ s.t. } \|X_i - L_i\|_F^2 \leq \sigma_\omega^2. \quad (4)$$

式中 $\|\cdot\|_F^2$ 是 Frobenious 范数, σ_ω^2 是加性高斯噪声的方差. rank 最小化是 NP-hard 问题, 不能直接求解. 为了得到近似解, Dong 等人利用 log det 函数替代 rank 函数, 并转化为拉格朗日形式:

$$L_i = \arg \min_{L_i} \|X_i - L_i\|_F^2 + \lambda l(L_i, \varepsilon). \quad (5)$$

式中 $l(L_i, \varepsilon)$ 是替代 rank 函数的 log det 函数, 用来近似矩阵 L_i 的奇异值对数之和, ε 是一个小的常数值. 虽然 $l(L_i, \varepsilon)$ 是非凸的, 但可以通过局部最优化转化为加权核范数问题, 应用加权奇异值阈值 (Singular Value Thresholding, SVT) 方法^[20] 求解, 低秩矩阵 L_i 的第 $k+1$ 次迭代表示为:

$$L_i^{(k+1)} = U(\tilde{\Sigma} - \lambda \text{diag}(\omega^{(k)}))V^T. \quad (6)$$

其中 $U\tilde{\Sigma}V^T$ 表示 X_i 的奇异值分解, $\omega_j^{(k)} = 1/(\sigma_j^{(k)} + \varepsilon)$ 中 $\sigma_j^{(k)}$ 表示 $L_i^{(k)}$ 的第 j 个奇异值. 利用相似图像块矩阵的低秩特性作为先验约束, CS 重构的目标函数可以表示为:

$$(\hat{x}, \hat{L}_i) = \arg \min_{x, L_i} \|y - \Phi x\|_2^2 + \eta \sum_i \left\{ \|\tilde{R}_i x - L_i\|_F^2 + \lambda l(L_i, \varepsilon) \right\}. \quad (7)$$

其中, $\tilde{R}_i x$ 表示 x_i 的相似块组构成的低秩矩阵, 上式可以拆分为式 (5) 低秩优化问题和对低秩后的

矩阵进行图像恢复问题, 对于图像恢复问题, 可以使用交替方向乘子法 (ADMM)^[21] 求解.

1.4 VCS 系统框架

为了降低采样端的计算成本, 视频压缩感知通常采用“独立编码, 联合解码”模型, 采样端对每一帧图像进行独立的采样编码, 解码端利用帧间关系进行联合解码. 为了保证较少采样数据量的同时更好地还原每一帧图像, 通常将视频序列分成图像组 (Group of picture, GOP), 每组中包含一帧关键帧 (key 帧) 和随后的几帧非关键帧 (CS 帧)^[9]. 其中关键帧使用高采样率采样, 而 CS 帧使用低采样率采样. 为了有效存储测量矩阵, 通常需要对整幅图像进行分块, 然后使用同一个测量矩阵对每一块进行采样. VCS 系统可以表示为下图:

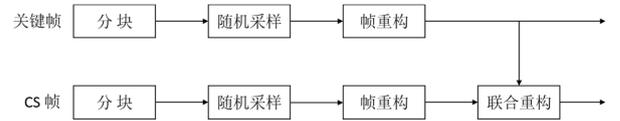


图1 VCS系统

2 非局部低秩正则化 CVS 重构算法

视频解码端获取视频采样数据后, 按照视频序列顺序对每一组数据帧进行重构. 关键帧的高质量重构是后序 CS 帧重构的保证, 关键帧的采样率相对 CS 帧较高, 采样数据中含有充足的图像信息, 为保证关键帧的重构质量, 采用重构效果可靠的 NLR-CS 算法. 而重构 CS 帧时, 首先充分利用帧内和帧间的局部相关性, 应用 PICCS 算法结合 TV 约束模型得到初步重构帧. 然后利用帧内和帧间的非局部自相似性, 应用改进的 NLR-CS 模型对初步重构帧进一步重构, 得到最终重构的 CS 帧.

2.1 PICCS 初步重构算法

根据视频信号相邻帧间的高度相关性, 本文利用 PICCS 算法对 CS 帧进行初步重构, 将前一帧图像作为先验, 解决下式凸优化问题:

$$\begin{aligned} \min_x & (1 - \alpha) \|\Psi_1 x\| + \alpha \|\Psi_2(x - pre)\|. \\ \text{s.t. } & y = \Phi x. \end{aligned} \quad (8)$$

式中, pre 表示前一帧图像, α 表示比例系数, Ψ_1 和 Ψ_2 表示稀疏变换矩阵. 由于视频在时间和空间方向都存在梯度稀疏特性, 因此 Ψ_1 和 Ψ_2 选择梯度稀疏变换矩阵, 则式中的第 1 项表示图像 x 的空间全变差, 第 2 项表示视频时间上梯度 $(x - pre)$ 的全变差. 结合 TV 模型, 式 (8) 可以表示为:

$$\begin{aligned} \min_x \sum_i (1 - \alpha) \|D_i x\|_p + \alpha \|D_i(x - pre)\|_p. \\ \text{s.t. } y = \Phi x. \end{aligned} \quad (9)$$

可以使用 ADMM 提高上式的求解速度, 引入变量 $w_i = D_i x, v_i = D_i(x - pre)$, 写成增广拉格朗日形式并拆分为关于 x, w, v 的 3 个子问题, 如下列式:

$$\begin{aligned} w_i = \arg \min_{w_i} (1 - \alpha) \|w_i\|_p - \sigma^T (D_i x - w_i) + \\ \frac{\beta}{2} \|D_i x - w_i\|_2^2; \end{aligned} \quad (10)$$

$$\begin{aligned} v_i = \arg \min_{v_i} \alpha \|v_i\|_p - \eta^T (D_i(x - pre) - v_i) + \\ \frac{\beta}{2} \|D_i(x - pre) - v_i\|_2^2; \end{aligned} \quad (11)$$

$$\begin{aligned} x = \arg \min_x \sum_i \{-\eta^T (D_i(x - pre) - v_i) - \\ \sigma^T (D_i x - w_i) + \frac{\beta}{2} \|D_i x - w_i\|_2^2 + \\ \frac{\beta}{2} \|D_i(x - pre) - v_i\|_2^2\} - \\ \delta^T (\Phi x - y) + \frac{\mu}{2} \|\Phi x - y\|_2^2. \end{aligned} \quad (12)$$

其中 $\sigma^T, \eta^T, \delta^T$ 为拉格朗日乘子, β, μ 为惩罚系数. 求解 w_i 和 v_i 时, 可以转化为类软阈值形式, 当取各向异性 TV 时, 有式 (13)、(14):

$$\begin{aligned} w_i = \max\{|D_i x - \frac{\sigma_i}{\beta_i}| - \frac{1 - \alpha}{\beta_i}, 0\} \\ \text{sgn}(D_i x - \frac{\sigma_i}{\beta_i}); \end{aligned} \quad (13)$$

$$\begin{aligned} v_i = \max\{|D_i(x - pre) - \frac{\eta_i}{\beta_i}| - \frac{\alpha}{\beta_i}, 0\} \\ \text{sgn}(D_i(x - pre) - \frac{\eta_i}{\beta_i}). \end{aligned} \quad (14)$$

更新 w_i 和 v_i 后, 将其代入到 x 子问题中, x 子问题是关于 x 的二次型函数, 求解时可以采用梯度下降 (Steepest Descent, SD) 法迭代:

$$x^{k+1} = x^k - \tau d^{k+1}. \quad (15)$$

上标 k 表示梯度下降的迭代次数, τ 表示迭代步长, d 表示梯度, 对于大规模问题, 多次迭代得到准确的 x 十分耗时, 而拉格朗日函数也是在循环迭代 w, v 和 x , 因此每次执行梯度下降时, 通过选取合适的 τ , 只执行一次迭代即可. 选取 τ 时, 可以采用 BB-like 步骤^[22], 即:

$$\tau_j = \frac{s_j^T s_j}{s_j^T u_j}. \quad (16)$$

其中 $s_i = x_j - x_{j-1}, u_j = d_j(x_j) - d_j(x_{j-1})$, 下标 j 表示 ADMM 循环迭代次数. 为验证步长选择合理, 可以使用非线性搜索算法 (NLSA), 即判断是否满

足 Armijo 条件, 如果不满足, 则缩放当前的 τ 值. 每次迭代中, 依次更新完 w, v 和 x 后, 更新拉格朗日乘子, 然后进入下次迭代. 当满足终止条件后, 停止迭代. PICCS 算法步骤如下:

算法 1 PICCS 算法.

- step 1. 输入: Φ, y , 前一帧 pre ;
- step 2. 初始化: $x_1 = \Phi^T y, \delta = 0, \eta = 0, \sigma = 0$;
- step 3. while $j < J, \|x_j - x_{j-1}\|_F / \|x_j\|_F > M$:
 - (a) 通过式 (13), (14) 计算 w_j, v_j ;
 - (b) 通过式 (16) 计算步长 τ_j , 判断调整 τ_j ;
 - (c) 通过式 (15) 计算更新 x_j ;
 - (d) 更新拉格朗日乘子 δ, η, σ ;
- end while;
- step 4. 输出: 初步重构图像 x .

2.2 改进的 NLR-CS 算法

NLR-CS 算法用于图像 CS 重构, 考虑图像的 NLSS 特性, 对初始重构的图像进行块匹配得到相似块集合, 对其进行低秩优化, 然后应用低秩矩阵进行图像恢复. 本文改进的 NLR-CS 算法用于视频序列的重构, 考虑帧内图像 NLSS 特性的同时, 充分利用相邻帧间的 NLSS 特性以提高重构性能. 对于 PICCS 算法初步重构的视频帧, 分别在本帧和前后两个相邻关键帧中寻找相似块构建低秩矩阵, 然后进行低秩优化和图像恢复. 由于关键帧的重构质量较高, 可以从关键帧中匹配到质量更高的相似块, 构建的低秩矩阵能够得到更好的低秩优化, CS 帧的重构性能可以得到明显提升. 算法可以表示为:

$$\begin{aligned} (\hat{x}, \hat{L}_i) = \arg \min_{x, L_i} \|y - \Phi x\|_2^2 + \\ \rho \sum_i \{ \|\left[\tilde{R}_{i1} x, \tilde{R}_{i2} Key \right] - [L_{i1}, L_{i2}] \|_F^2 + \\ \lambda \mathcal{L}([L_{i1}, L_{i2}], \varepsilon) \}. \end{aligned} \quad (17)$$

式中, i 是本帧像素坐标, $[\tilde{R}_{i1} x, \tilde{R}_{i2} Key]$ 表示本帧中的 m_1 个相似块和关键帧中的 m_2 个相似块构成的含噪低秩矩阵, $[L_{i1}, L_{i2}]$ 则表示去噪优化后的低秩矩阵, ρ 表示惩罚系数. 式 (17) 可以拆分为求解图像 x 和求解低秩矩阵 L , 求解 L 可以使用式 (6) 所示的奇异值阈值方法, 而对于 x 问题, 应用 ADMM, 引入变量 z 替代 $\tilde{R}_{i1} x$ 中的 x , 则有:

$$\begin{aligned} \hat{x} = \arg \min_{x, z} \|y - \Phi x\|_2^2 + \beta \|x - z\|_2^2 + \frac{\mu}{2\beta} \|z\|_2^2 + \\ \rho \sum_i \|\tilde{R}_{i1} z - L_{i1}\|_F^2. \end{aligned} \quad (18)$$

其中 β 为权重系数, μ 为拉格朗日乘子, 上式可以拆为 z 子问题和 x 子问题. z 子问题为关于 z 的二

次型函数, 因此有式 (19), 式中 $\tilde{R}_{i1}^T \tilde{R}_{i1}$ 项是对角矩阵, 对应像素点坐标, 于是图像对应位置相除, 容易求解. x 子问题同样是关于 x 的二次型函数, 求导有式 (20), 可以使用梯度下降法或共轭梯度法求解.

$$z = \arg \min_z \beta \|x - z + \frac{\mu}{2\beta}\|^2 + \rho \sum_i \|\tilde{R}_{i1} z - L_{i1}\|_F^2.$$

$$z = \{\rho \sum_i \tilde{R}_{i1}^T \tilde{R}_{i1} + \beta I\}^{-1} \{\beta x + \frac{\mu}{2} + \rho \sum_i \tilde{R}_{i1} L_{i1}\}.$$
(19)

$$x = \arg \min_x \|y - \Phi x\|_2^2 + \beta \|x - z + \frac{\mu}{2\beta}\|^2.$$

$$(\Phi^T \Phi + \beta I) x = (\Phi^T y + \beta z - \frac{\mu}{2}).$$
(20)

交替迭代 z, x 后, 更新拉格朗日乘子 μ , 达到终止条件后, 停止迭代并输出最终重构图像. 为了减少计算量, 考虑到相似块位置可能不变, 可以多次迭代 L, z, x 后, 再寻找一次相似块位置. 因此改进的 NLR-CS 算法步骤如下:

算法 2 改进的 NLR-CS 算法.

step 1. 输入: Φ, y , 相邻关键帧 key ;

step 2. 初始化: x_1 为 PICCS 初步重构帧, $\mu = 0$;

step 3. while $j < J, \|x_j - x_{j-1}\|_F / \|x_j\|_F > M$:

(a) 对 x_j 在本帧和 key 帧中进行块匹配;

(b) while $k < K$:

通过式 (6) 计算低秩矩阵 L_j^k ;

通过式 (19) 计算更新 z_j^k ;

通过共轭梯度法求解式 (20) 更新 x_j^k ;

end while;

(c) 更新 x_j 和拉格朗日乘子 μ

end while;

step 4: 输出: 最终重构图像 x .

3 实验与分析

为验证本文算法的重构性能, 将其与近年来提出的性能较好的视频压缩感知重构算法进行仿真对比, 包括 Video-MC^[9]、2sMHR^[11]、RRS-CS^[12] 和 SSIM-InterF-GSR^[14]. 图 1 所示 VCS 系统的参数设置如下: 视频帧的图像分块大小为 32×32 , 对图像块进行采样的测量矩阵为高斯随机矩阵, 视频序列中每个图像组包含 8 帧图像, 其中第 1 帧为关键帧, 采样率设为 0.7, 后 7 帧为 CS 帧, 采样率设为 0.2.

本文算法的相关参数设置如下: 第 1 步利用 PICCS 模型初步重构时, 比例系数 α 设为 0.8, 惩罚系数 μ, β 分别设为 256 和 32, 循环迭代终止条件设置为 $\|x_j - x_{j-1}\|_F / \|x_j\|_F < 10^{-6}$, 同时设置最大迭代次数 J 为 300; 第 2 步利用 NLR-CS 模型进一步

重构时, 为了避免块效应, 将本帧图像进行重叠分块, 块大小为 6×6 , 重叠 1 个像素, 之后针对每个图像块寻找相似块, 本帧中寻找的相似块数 m_1 为 16, 前后两帧关键帧中寻找的相似块数 m_2 为 20. 式 (17) 中的惩罚系数 ρ 设为 0.01, 循环迭代终止条件设置为 $\|x_j - x_{j-1}\|_F / \|x_j\|_F < 10^{-3}$, 循环迭代的最大迭代次数 K 和 J 均为 20.

3.1 重构性能对比分析

为比较重构性能, 考虑视频中每一帧图像, 应用图像的峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 和主观感受对比重构质量. 实验对比应用的标准视频序列为 CIF 格式 (352×288) 的 Forman、Akiyo、Mother-daughter、city. 图 2 为本文算法和 Video-MC、2sMHR、RRS-CS、SSIM-InterF-GSR 在重构 4 组标准视频序列的前两个 GOP 时每一帧图像的 PSNR 值折线图 (其中第 1 帧和第 9 帧是关键帧).

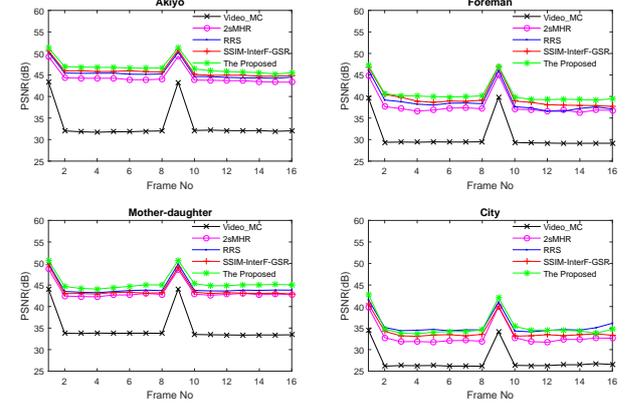


图 2 重构视频序列性能对比

对比可见, 对于 Forman、Akiyo、Mother-daughter 3 个视频序列, 本文算法与其他算法相比, 重构的每一帧图像都能得到最高的 PSNR 值, 而 RRS-CS 和 SSIM-InterF-GSR 的重构性能要略低于本文算法, 2sMHR 的重构性能略低于前三种算法, Video-MC 的重构性能则远落后于其他算法. 对于 Mother-daughter 视频序列, 本文算法和 SSIM-InterF-GSR、RRS-CS、2sMHR 重构视频帧的 PSNR 值非常接近, 这是因为 Mother-daughter 视频序列内容极其稳定, 各算法都考虑了视频帧间的相关性, 能够获得较高的重构质量. City 视频序列是对一片城市楼房的俯瞰, 视频的每一帧图像内容复杂高频分量较多, 可以看到各算法对 City 视频帧序列的重构质量与其他视频序列相比大幅降低且具有较强的不稳定性. 其中 RRS 算法在 CS 帧上的重构质量总体上要略优于本文算法, 由于 RRS 根据帧间残差的稀疏性建模, 在分裂 Bregman 迭代的过程中能够更好地对 City 视频序



图3 foreman 第6帧重构效果对比

表1 重构视频序列的PSNR均值(dB)

视频序列	Video-MC	2sMHR	RRS-CS	SSIM-InterF-GSR	the Proposed
Akiyo	33.39	44.54	45.56	46.04	46.87
Forman	30.61	37.96	38.86	39.74	40.69
Mother-daughter	34.92	43.67	44.43	43.90	45.55
City	27.34	33.16	35.51	34.27	35.37
Average	31.57	39.78	41.09	40.99	42.15

列进行重构。

基于多假设预测的 RRS-CS 和 2sMHR 算法高度依赖于参考帧上的假设块匹配程度, 在处理快速运动的视频序列时, 找到优质的匹配块非常困难, 所以预测的结果相对较差. 而本文算法和 SSIM-InterF-GSR 直接考虑了本帧图像的 NLSS 特性, 在参考帧与本帧图像相差较大时, 也能保证初始的重构质量. 本文算法的重构性能要优于 SSIM-InterF-GSR, 主要有以下两个原因: (1) 对于采样率较高的关键帧, 本文算法应用基于 NLSS 特性的 NLR-CS 算法进行重构, 能够得到非常高的重构质量. 而 SSIM-InterF-GSR 算法应用基于组稀疏特性的 GSR 算法进行重构, 文献 [16] 证明了 NLR-CS 算法的重构性能高于 GSR 算法.

(2) 对 CS 帧进行重构时, SSIM-InterF-GSR 算法仅仅考虑了前面的一个关键帧, 这使得一个 GOP 中靠后的视频序列不能很好地利用关键帧信息; 而本文算法分别在本帧和前后两个关键帧上寻找相似块, 由于关键帧的重构质量较高, 构建的低秩矩阵能够得到较好的低秩优化, CS 帧的重构性能提升明显.

图3是本文算法与 Video-MC、2sMHR、RRS-CS、SSIM-InterF-GSR 对 Forman 视频序列第6帧图像的

重构视觉效果直观对比, 并且给出了每个重构结果的 PSNR 值. 其中 Video-MC 的重构效果一般, 在人物与墙壁之间的边界产生了明显的振铃现象, 且在快速运动的地方比如人物嘴部非常模糊. 2sMHR 算法在像素域进行多假设预测时进行了重叠分块, 有效地避免了块效应, 但重构图像的墙壁上以及面部出现了非常明显的噪点. RRS 和 SSIM-InterF-GSR 的重构结果非常接近, 与 2sMHR 的重构结果相比, 重构性能有了明显的提升, 但重构图像的视觉效果非常柔和, 图中的墙壁整体上比较平滑, 丢失了部分高频细节. 本文算的重构结果与 RRS 和 SSIM-InterF-GSR 相比, PSNR 约提升 1.5dB, 图像的高频细节部分保留较好, 视觉上更加逼近原图.

表1对比了本文算法与其他重构算法对四组视频序列的重构 PSNR 均值, 本文算法与其他重构算法相比重构性能得到明显的提升. 本文算法与 Video-MC、2sMHR、RRS-CS、SSIM-InterF-GSR 相比, 在 Akiyo 序列上分别提升了 13.48、2.33、1.31、0.83dB; 在 Forman 序列上分别提升了 10.08、2.73、1.83、0.95dB; 在 Mother-daughter 序列上分别提升了 10.63、1.88、1.12、1.65dB. 在 City 序列上, 本文算法与 Video-MC、

2sMHR、SSIM-InterF-GSR 相比, 分别提升了 8.03、2.21、1.1dB.

3.2 算法复杂度分析

本文算法应用 NLR-CS 算法对关键帧进行重构, 应用 PICCS 算法对 CS 帧进行初步重构, 应用改进的 NLR-CS 算法对初步重构帧进一步重构得到最终的视频帧. 其中 PICCS 算法的复杂度较低, 把基于图像梯度的 TV 模型应用于本帧图像和本帧与前一帧的差值得到优化式 (9). 应用 ADMM 提高式 (9) 的求解速度, 将其拆分为关于本帧图像 x 、本帧图像 TV 模型 ω 和帧间差的 TV 模型 v 三个子问题进行迭代求解. 应用类软阈值算法求解 ω 和 v , 具有较快的收敛速度. x 子问题是关于 x 的二次函数, 采用梯度下降法求解时, 在每次 ADMM 迭代中仅迭代一次.

对于 NLR-CS 算法, 首先利用图像的 NLSS 特性在本帧和关键帧中找到指定图像块的相似块矩阵, 然后以矩阵秩最小化为稀疏先验对图像进行恢复. 算法耗时主要集中在式 (6) 基于加权 SVT 算法的秩最小化函数中的 SVD 函数. 具体来说, 对于 PICCS 算法初步重构的 QCIF 格式的视频帧, 大小为 176×144 , 对其进行重叠分块, 块大小为 6×6 , 重叠一个像素, 约得到 1000 个图像块, 对于每个图像块, 在本帧和关键帧中共匹配 36 个图像块, 得到 36×36 大小的低秩矩阵. 在 ADMM 的每次迭代中, 都需要对 1000 个互不相关的 36×36 大小的矩阵求 SVD 函数, 该过程约占整个算法时间复杂度的 50%.

为了对比各算法的重构速度, 对 QCIF 格式的视频序列 Forman、Akiyo、Mother-daughter、city 进行重构, GOP 长度设置为 8, 关键帧采样率为 0.7, CS 帧采样率为 0.2. 实验的运行环境是 Matlab R2015b, 电脑配置如下: CPU 为 Inter Core i5-6500, 主频为 3.2GHz, 运行内存为 8GB, 系统版本为 Win 7. 表 2 为各算法对一个 CS 帧的平均重构时间对比.

表 2 不同算法下每帧平均所需重构时间 (s)

重构算法	Mother-daughter	Akiyo	Forman	City
Video-MC	14.3	15.1	15.6	15.8
2sMHR	7.3	7.5	7.7	8.0
RRS-CS	99.4	110.3	106.6	113.8
SSIM-InterF-GSR	178.3	182.5	182.0	180.7
The Proposed	165.5	167.0	166.5	168.4

由表 2 可知, 算法重构速度由快到慢依次为 2sMHR、Video-MC、RRS-CS、本文算法和 SSIM-InterF-GSR. 2sMHR 由于在多假设预测时没有迭代过程, 直接利用假设集对当前帧进行预测, 处理过程

比较简单. Video-MC 先利用观测值对当前帧进行初始重构, 然后迭代地进行 $1/4$ 像素精度的像素域运动估计和残差重构, 处理过程也相对简单. 2sMHR 和 Video-MC 的重构速度与其他算法相比具有明显优势. RRS 算法涉及到交替的帧间残差加权稀疏建模和多假设预测, 模型的建立和求解过程较为复杂, 耗时约为 110s. SSIM-InterF-GSR 算法由于也涉及到矩阵分解问题, 每次迭代所需时间较长, 重构一帧图像耗时与本文算法相当, 约为 3min. 本文算法的耗时约为 166s, 通过对本文算法复杂度的分析可知, 矩阵低秩化处理中的 SVD 函数约占总耗时的一半, 但值得注意的是, 每次迭代中对上千个低秩矩阵进行处理时, 不同的低秩矩阵处理之间没有相关性和依赖性, 在 CPU 串行处理算法中, 必须等一个矩阵处理完成之后才能进行下一个矩阵的处理, 不可避免地导致重构时间过长. 可以考虑基于 GPU 对图像块的低秩优化部分进行并行处理, 从而提高重构速度.

4 结语

本文提出的基于全变分和非局部低秩正则化视频压缩感知重构算法在典型的 CVS 系统中实现了高性能重构. 对于采样率较高的关键帧, 主要考虑图像本身具有的 NLSS 特性, 应用性能可靠的 NLR-CS 重构算法进行重构, 保证了关键帧的高质量恢复; 对于采样率相对较低的 CS 帧, 充分考虑了视频帧内和帧间的局部光滑特性和非局部自相似性, 应用 PICCS 算法和改进的 NLR-CS 算法相结合进行重构. 实验结果表明, 相对于目前主流的 CVS 重构算法 Video-MC、2sMHR、RRS-CS 和 SSIM-InterF-GSR, 本文算法对视频帧的重构质量得到明显提升. 本文算法计算复杂度虽相对较高, 但通过分析发现本文算法在迭代过程中涉及到大量不相关的矩阵低秩优化问题, 可以应用 GPU 对其并行化处理大幅提高重构速度.

参考文献 (References)

- [1] Candes E J, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.
 - [2] Donoho D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
 - [3] Elahi S, Kaleem M, Omer H. Compressively sampled MR image reconstruction using generalized thresholding iterative algorithm[J]. 2018, 286(1): 91-98.
 - [4] 丁泽刚, 刘旻昆, 王岩, et al. 基于压缩感知的地基 MIMO SAR 近场层析成像研究 [J]. 信号处理, 2019(5): 729-740.
- (Ding Z G, Liu H K, Wang Y, et al. Near-Field

- Ground-Based MIMO SAR Tomography via Compressive Sensing[J]. *Journal of Signal Processing*, 2019(5): 729-740.)
- [5] Guo J, Song B, Du X. Significance Evaluation of Video Data Over Media Cloud Based on Compressed Sensing[J]. *IEEE Transactions on Multimedia*, 2016, 18(7): 1297-1304.
- [6] Kareth Leon-Lopez, Henry Arguello. Temporal Colored Coded Aperture Design in Compressive Spectral Video Sensing[J]. *IEEE Transactions on Image Processing*, 2018, 28(1): 253-264.
- [7] Shi J V, Sankaranarayanan A C, Studer C E, et al. Methods and systems for video compressive sensing for dynamic imaging[P]. US: 9552658B2, 2017-1-24.
- [8] Cossalter M, Valenzise G, Tagliasacchi M, et al. Joint compressive video coding and analysis[J]. *IEEE Transactions on Multimedia*, 2010, 12(3): 168-183
- [9] SUNGKWANG M, FOWLER J E. Residual reconstruction for block-based compressed sensing of video[C]. *Data Compression Conference(DCC)*. Snowbird, USA: IEEE, 2011: 183-192.
- [10] Chen C, Tramel E W, Fowler J E. Compressed-sensing recovery of images and video using multihypothesis predictions[C] *Signals, Systems and Computers*. Asilomar: IEEE, 2012: 1193-1198.
- [11] 欧伟枫, 杨春玲, 戴超. 一种视频压缩感知中两级多假设重构及实现方法 [J]. *电子与信息学报*, 2017, 39(7): 1688-1696.
(Ou W F, Yang C L, Dai C. A Two-stage Multi-hypothesis Reconstruction and Two Implementation Schemes for Compressed Video Sensing[J]. *Journal of electronics & information technology*, 2017, 39(7): 1688-1696.)
- [12] Zhao C, Ma S, Zhang J, et al. Video Compressive Sensing Reconstruction via Reweighted Residual Sparsity[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2017, 27(6): 1182-1195.
- [13] 和志杰, 杨春玲, 汤瑞东. 视频压缩感知中基于结构相似的帧间组稀疏表示重构算法研究 [J]. *电子学报*, 2018, 46(3): 544-553.
(He Z J, Yang C L, Tang R D. Research on Structural Similarity Based Inter-Frame Group Sparse Representation for Compressed Video Sensing[J]. *Acta Electronica Sinica*, 2018, 46(3): 544-553.)
- [14] Wang Y, Yang J, Yin W, et al. A new alternating minimization algorithm for image segmentation[J]. *Siam Journal on Imaging sciences*, 2008, 1(3): 248-272.
- [15] Dong W, Shi G, Li X. Nonlocal Image Restoration With Bilateral Variance Estimation: A Low-Rank Approach[J]. *IEEE Transactions on Image Processing*, 2013, 22(2): 700-711.
- [16] Dong W, Shi G, Li X, et al. Compressive sensing via nonlocal low-rank regularization.[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2014, 23(8): 3618-3632.
- [17] Chavarrias C, Abascal J F P J, Montesinos P, et al. Exploitation of temporal redundancy in compressed sensing reconstruction of fMRI studies with a prior based algorithm (PICCS)[J]. *Medical Physics*, 2015, 42(7): 3814-3821.
- [18] Abascal J F P J, Abella M, Sisniega A, et al. Investigation of Different Sparsity Transforms for the PICCS Algorithm in Small-Animal Respiratory Gated CT[J]. *Plos One*, 2015, 10(4): e0120140.
- [19] Xu J, Zhang L, Zuo W, et al. Patch group based nonlocal self-similarity prior learning for image denoising[C]. *Proceedings of the IEEE international conference on computer vision*. Chile: IEEE, 2015: 244-252.
- [20] Gu S, Zhang L, Zuo W, et al. Weighted Nuclear Norm Minimization with Application to Image Denoising[C]. *Computer Vision and Pattern Recognition*. Columbus, USA: IEEE, 2014:2862-2869.
- [21] Boyd S, Parikh N, Chu E, et al. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers[J]. *Foundations & Trends in Machine Learning*, 2010, 3(1): 1-122.
- [22] Li C. An Efficient Algorithm For Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing[D]. USA: Department of Computational and Applied Mathematics, Rice University, 2009: 4-6.

作者简介

田金鹏(1974—), 男, 硕士生导师, 博士, 从事压缩感知、信号处理等研究, E-mail: adaline@163.com.

杨洁(1995—), 女, 硕士研究生, 硕士, 从事视频、图像压缩感知的研究, E-mail: jieyang5252@163.com.

刘通(1995—), 男, 硕士研究生, 硕士, 从事压缩感知、通信信号处理的研究, E-mail: 535093824@qq.com.

闵天(1994—), 男, 硕士研究生, 硕士, 从事视频压缩感知的研究, E-mail: mtian168@163.com.