

一种视频压缩感知中两级多假设重构及实现方法

欧伟枫^{①②} 杨春玲^{*①} 戴超^①

^①(华南理工大学电子与信息学院 广州 510640)

^②(华为技术有限公司 深圳 518129)

摘要: 视频压缩感知在采集端资源受限的视频采集应用场景有重要研究意义。重构算法是视频压缩感知的关键技术, 基于多假设预测的“预测-残差重构”框架具有良好的重构性能。但现有的多假设预测算法大多在观测域提出, 这种预测方法由于受到不重叠分块的限制, 造成了预测帧的块效应, 降低了重构质量。针对此问题, 该文将像素域多假设预测与观测域多假设预测相结合, 提出两级多假设重构思想(2sMHR), 并设计了基于图像组(Gw_2sMHR)和基于帧(Fw_2sMHR)的两种实现方法。仿真结果表明, 所提 2sMHR 重构算法能有效减小块效应, 相比于现有最好的多假设预测算法具有更低的时间复杂度和更高的视频重构质量。

关键词: 视频压缩感知; 重构; 预测; 多假设; 稀疏

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2017)07-1688-09

DOI: 10.11999/JEIT161142

A Two-stage Multi-hypothesis Reconstruction and Two Implementation Schemes for Compressed Video Sensing

OU Weifeng^{①②} YANG Chunling^① DAI Chao^①

^①(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China)

^②(Huawei Technologies CO., LTD., Shenzhen 518129, China)

Abstract: Compressed Video Sensing (CVS) has great significance to the scenarios with a resource-deprived video acquisition side. Reconstruction algorithm is the key technique in compressed video sensing. The Multi-Hypothesis (MH) prediction based “prediction-residual reconstruction” framework has good reconstruction performance. However, most of the existing multi-hypothesis prediction algorithms are proposed in measurement domain, which cause block artifacts in the predicted frames and decrease reconstruction accuracy due to the restriction of non-overlapping block partitioning. To address this issue, this paper proposes a two-stage Multi-Hypothesis Reconstruction (2sMHR) idea by incorporating the measurement-domain MH prediction with pixel-domain MH prediction. Two implementation schemes, GOP-wise (Gw) and Frame-wise (Fw) scheme, are designed for the 2sMHR. Simulation results show that the proposed 2sMHR algorithm can effectively reduce block artifacts and obtain higher video reconstruction accuracy while requiring lower computational complexity than the state-of-the-art CVS prediction methods.

Key words: Compressed Video Sensing (CVS); Reconstruction; Prediction; Multi-Hypothesis (MH); Sparsity

1 引言

传统视频信号采集建立在奈奎斯特采样定理之上, 先对视频信号进行高速采样, 然后通过复杂的传统视频编码压缩算法丢弃大量的冗余数据, 以实现高效存储与传输。该做法造成了采样资源的巨大

浪费, 加上过高的压缩编码复杂度, 使得传统视频编码方案在采集端资源受限的应用场景并不适用。压缩感知突破了奈奎斯特采样定理的限制, 实现了信号的采样与压缩同时进行, 节省了采样资源的同时降低了采集端信号处理复杂度, 适用于采集端资源受限的应用场景。将压缩感知应用于视频采集的视频压缩感知研究领域^[1-4]具有广阔的应用前景(如无线视频监控, 无线多媒体传感器网络), 近年来引起了广泛关注。

重构算法是视频压缩感知研究的关键。从理论上讲, 压缩感知能保证稀疏信号的完全恢复, 但对于真实的视频/图像信号, 由于其并不是真正稀疏

收稿日期: 2016-10-26; 改回日期: 2017-03-21; 网络出版: 2017-04-25

*通信作者: 杨春玲 eeclyang@scut.edu.cn

基金项目: 国家自然科学基金(61471173), 广东省自然科学基金(2016A030313455)

Foundation Items: The National Natural Science Foundation of China (61471173), The Natural Science Foundation of Guangdong Province (2016A030313455)

的，因此，充分挖掘视频信号的稀疏表示潜力是提高重构性能的关键技术。现有的重构方法大致可分为4类。第1类是将视频序列各帧视为独立的图像，分别进行独立帧内重构，这种方法由于没有利用视频帧间相关性，因此重构质量较低。第2类方法是利用3维稀疏变换^[5]对视频序列整体进行重构。这种方法计算复杂度非常高，而且由于忽略了视频帧中的物体运动，导致重构质量并不高。第3类方法利用视频时空相关性建立正则化模型^[6-9]，效果比较好的有：基于高阶FIR差分滤波器的TV正则化模型^[6]、高斯混合模型^[7]、联合TV最小化残差重建模型^[8]以及加权残差稀疏模型^[9]等。该类重构方法虽然有效利用了视频时空相关性进行建模，得到较高的视频重构质量，但模型的建立和求过程往往较为复杂，普适性并不高。

第4类预测-残差重构方法具有简单的算法思路和较高的重构性能，是目前视频压缩感知研究中重构性能最好的算法。其基本思想是，先利用帧间相关性对当前帧进行预测，然后用CS重构算法恢复出原始帧与预测帧之间的残差，最后把恢复的残差与预测帧相加得到最终重构帧。在预测-残差重构框架下，如何利用帧间相关性对视频帧进行高效预测(稀疏化)是视频重构的关键步骤。文献[10]借鉴了传统视频编码中的运动估计/运动补偿方法，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，以高计算复杂度为代价换取了非关键帧质量的提升。文献[11,12]提出了观测域运动估计方法，可以利用视频帧的观测值直接在观测域进行运动估计，但重构质量并不高。由于单假设运动补偿^[10-12]未能充分挖掘视频序列的时间相关性，又由于在重构端没有码率的限制，最近不少学者对多假设(MH)预测展开研究，以得到更高的预测精度。文献[13]最先提出了基于 ℓ_1 正则化的多假设预测算法，通过对多假设加权系数加以稀疏约束，用最少数量的假设块的线性加权来预测当前块，具有比单假设预测更高的预测精度。文献[14]认为对多假设加权系数加以稀疏约束并不合适，提出了基于Tikhonov正则化的多假设预测算法，可直接通过解析求得线性加权系数，算法简单而有效。文献[15]在Tikhonov正则化模型的基础上加入了当前块在DCT域稀疏的 ℓ_1 范数约束，并基于ADMM算法设计了该多假设预测模型的权值求解方法。文献[16]提出的基于弹性网模型的多假设预测算法结合了Lasso模型和Tikhonov模型的优点，可得到更加准确的加权系数。文献[17]对文献[16]的多假设预测模型作了进一步改进，在 ℓ_1 范数正则

化项中加入了权重调整，并提出了假设集优化技术，进一步提高了预测精度。

本文的研究工作建立在预测-残差重构框架的基础上。预测-残差重构框架下的多假设预测算法能充分利用多个假设块的信息，具有较高的视频重构质量。但是，现有的多假设预测算法^[13-17]均在观测域提出，在观测域进行多假设预测的最大缺点是受到不重叠分块的限制，从而造成了预测帧的块效应，降低重构质量。

为了克服观测域多假设预测的限制，以得到视频帧更加稀疏的表示，获得更高的视频重构质量，本文提出了两级多假设重构(2sMHR)思想，设计了两种实现方法，一种是基于帧的实现方法，另一种是基于GOP的实现方法，并给出了这两种实现方法的详细实现过程和理论分析。本文在第4节给出了所提思想和实现方法的仿真实验结果，并进行了全面的对比和分析，验证了其有效性。

2 视频压缩感知中的多假设预测

预测-残差重构框架的一般处理流程如图1所示。

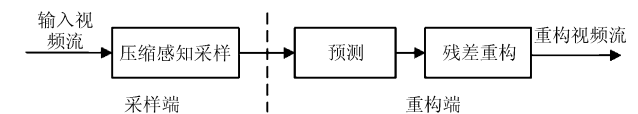


图1 预测-残差重构的一般流程

为了防止误差累加传播，通常以图像组(Group Of Pictures, GOP)为单位对视频序列进行处理。每个GOP的首帧为关键帧，其余帧为非关键帧。

在采样端，对各视频帧进行独立观测。分块压缩感知^[18](Block-based Compressed Sensing, BCS)对存储及计算能力要求较低，是目前使用最为广泛的采样方式。每个视频帧被分成若干个不重叠的 $B \times B$ 大小图像块，对每个子块采用相同的观测矩阵 Φ 进行独立观测：

$$\mathbf{y}_{t,i} = \Phi \mathbf{x}_{t,i} \quad (1)$$

式中， $\mathbf{x}_{t,i}$ 为 $B^2 \times 1$ 维列矢量，表示当前帧 \mathbf{x}_t 的第 i 个 $B \times B$ 不重叠图像块； $\mathbf{y}_{t,i}$ 为 $M \times 1$ 维列矢量，表示图像块 $\mathbf{x}_{t,i}$ 的观测值； Φ 为 $M \times B^2$ 维的高斯随机观测矩阵，定义 M/B^2 为采样率。通常，关键帧的采样率高于非关键帧采样率。

在重构端，对关键帧进行独立帧内重构；对于非关键帧，先利用视频帧间相关性进行预测，然后在预测帧基础上进行残差重构。在现有的预测算法中，基于Tikhonov正则化的多假设预测^[14]算法在预

测精度和计算复杂度上较其他算法更有优势,文中在该算法基础上,提出了改进算法。

多假设预测的核心思想是,用参考帧中的一组假设块的线性组合来估计当前块,即

$$\tilde{\mathbf{x}}_{t,i} = \mathbf{H}_{t,i} \hat{\mathbf{w}}_{t,i} \quad (2)$$

式中, $\tilde{\mathbf{x}}_{t,i}$ 表示 $\mathbf{x}_{t,i}$ 的预测块, $B^2 \times K$ 维矩阵 $\mathbf{H}_{t,i}$ 为该组假设块, $\mathbf{H}_{t,i}$ 的各列 $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K$ 为各假设块的矢量化形式, K 为假设块的数量; $K \times 1$ 维列矢量 $\hat{\mathbf{w}}_{t,i}$ 为最佳线性组合系数矢量。如何求得式(2)中的多假设线性加权系数 $\hat{\mathbf{w}}_{t,i}$ 是多假设预测的核心问题。

由于在重构端只有当前块的观测值,一种直接的做法是在观测域求解如式(3)所示最小二乘问题:

$$\hat{\mathbf{w}}_{t,i} = \arg \min_{\mathbf{w}} \|\mathbf{y}_{t,i} - \Phi \mathbf{H}_{t,i} \mathbf{w}\|_2^2 \quad (3)$$

式中,由于当前块的测量点数 M 通常小于假设块数量 K ,即方程个数小于未知数个数,此时式(3)是病态问题,无法解得 $\hat{\mathbf{w}}_{t,i}$ 。文献[14]提出了 Tikhonov 正则化的最小二乘模型:

$$\hat{\mathbf{w}}_{t,i} = \arg \min_{\mathbf{w}} \|\mathbf{y}_{t,i} - \Phi \mathbf{H}_{t,i} \mathbf{w}\|_2^2 + \lambda \|\Gamma \mathbf{w}\|_2^2 \quad (4)$$

式中, λ 为拉格朗日参数, Tikhonov 矩阵 Γ 为对角阵形式,其对角线元素由当前块观测值与各假设块观测值的欧氏距离所组成。

$$\Gamma = \text{diag}(\|\mathbf{y}_{t,i} - \Phi \mathbf{h}_1\|_2, \dots, \|\mathbf{y}_{t,i} - \Phi \mathbf{h}_K\|_2) \quad (5)$$

其中, \mathbf{h}_j 对应 $\mathbf{H}_{t,i}$ 的第 j 列, $j = 1, \dots, K$ 。式(4)具有如式(6)所示闭式解:

$$\hat{\mathbf{w}}_{t,i} = \left[(\Phi \mathbf{H}_{t,i})^T (\Phi \mathbf{H}_{t,i}) + \lambda \Gamma^T \Gamma \right]^{-1} (\Phi \mathbf{H}_{t,i})^T \mathbf{y}_{t,i} \quad (6)$$

得到当前帧的预测帧 $\tilde{\mathbf{x}}_t$ 后,在预测帧基础上进行残差重构^[10],以进一步补偿原始帧与预测帧之间的残差,得到当前帧更稀疏的表达。

3 两级多假设重构算法

现有的多假设预测方法^[13-17]均在观测域进行,由于受到采集端 BCS^[18]采样方式的限制,在重构端只能对固定大小的不重叠块进行多假设预测,导致在预测过程中不能灵活地分块,从而造成预测帧的块效应明显,降低了预测精度,使得预测残差稀疏度不高,限制了视频重构质量的提升。针对该问题,文中提出了两级多假设重构方案(2sMHR),该重构方案分为两个重构阶段,第1阶段的观测域多假设预测及第2阶段的像素域多假设预测,每次预测后,都要进行残差重构。2sMHR 重构方案总体流程图如图2所示。

3.1 像素域多假设预测

为了克服观测域多假设预测的限制,我们提出在第1阶段重构帧基础上进行第2阶段的像素域多

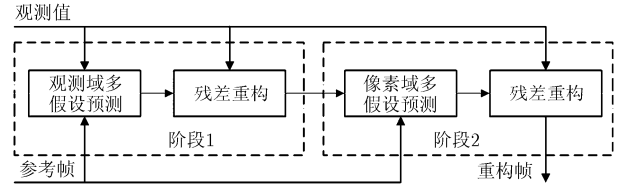


图2 2sMHR 总体流程图

假设预测。得到第1阶段的重构帧 $\tilde{\mathbf{x}}_t$ 以后,第2阶段的多假设预测以重叠分块的方式在像素域进行。图3给出了像素域多假设预测的流程框图,分步骤解释如下:

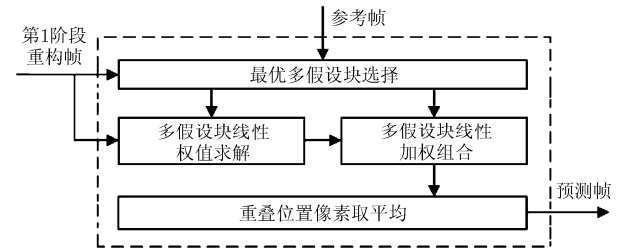


图3 像素域多假设预测框图

步骤1 对当前帧(第1阶段重构帧)进行重叠分块;

步骤2 在确定的搜索窗内对参考帧进行运动全搜索,选取出与当前块之间 MSE(或 SAD)最小的前若干个匹配块(文中试验中取8个匹配块)作为当前帧的最优多假设块;

步骤3 通过求解最小二乘问题得到多假设线性加权系数:

$$\hat{\mathbf{w}}_{t,i} = \arg \min_{\mathbf{w}} \|\hat{\mathbf{x}}_{t,i} - \mathbf{H}_{t,i} \mathbf{w}\|_2^2 = (\mathbf{H}_{t,i}^T \mathbf{H}_{t,i})^{-1} \mathbf{H}_{t,i}^T \hat{\mathbf{x}}_{t,i} \quad (7)$$

式中, $\hat{\mathbf{x}}_{t,i}$ 表示第1阶段重构帧 $\tilde{\mathbf{x}}_t$ 中的一个重叠块, $\mathbf{H}_{t,i}$ 表示该重叠块 $\hat{\mathbf{x}}_{t,i}$ 在参考帧中的假设块集;

步骤4 利用所求权值 $\hat{\mathbf{w}}_{t,i}$, 对多假设块进行线性加权求和,得到当前帧的预测块 $\tilde{\mathbf{x}}_{t,i}$:

$$\tilde{\mathbf{x}}_{t,i} = \mathbf{H}_{t,i} \hat{\mathbf{w}}_{t,i} \quad (8)$$

步骤5 得到当前帧各个块的预测块后,对重叠位置的像素值取平均,得到最终预测帧。

通过采取重叠块运动补偿和灵活的分块,所提像素域多假设预测有效消除了块效应和抑制重构噪声,从而提高了预测质量。更重要的是,在像素域中,可以借鉴许多传统视频编码中高效的运动补偿技术来改进预测精度。

需要指出的是,在观测域进行多假设预测和在像素域进行多假设预测各有优缺点。在观测域进行多假设预测,优点是可利用当前帧准确的观测值信息进行块匹配,缺点是重构端只能对各不重叠块进

行预测, 难免造成块效应。在像素域进行多假设预测, 优点是重构端可以进行灵活的重叠分块多假设预测, 从而有效抑制重构噪声, 减小块效应, 而且, 多假设线性权值可直接通过求解像素域误差最小化的普通最小二乘问题得到, 无需对加权矢量 w 加以正则化, 使得加权矢量的求解更为准确; 缺点是只能把第 1 阶段重构的图像块当做目标块做预测匹配, 其预测性能受到第 1 阶段重构质量的影响。像素域多假设预测中, 重叠分块很关键, 分块大小及重叠范围对预测性能有重要影响: 分块过小, 预测精度受第 1 阶段重构帧质量影响较大, 分块过大, 会导致视频帧运动区域的块匹配失真; 重叠范围太大, 容易过多平滑掉细节信息, 同时会大大增加分块数量, 导致算法复杂度剧增, 重叠范围太小, 则难以有效抑制块效应。研究中我们根据理论分析对大量视频序列做了仿真实验, 得出了像素域预测过程中重叠分块大小和重叠范围的经验值, 相关数据和分析在实验部分(4.2 节)。

由上述分析可知, 两级多假设重构方案通过先后进行第 1 阶段的观测域多假设预测和第 2 阶段的像素域多假设预测, 有效结合了观测域多假设预测和像素域多假设预测的优点。下面介绍 2sMHR 重构方案的两种实现方法。

3.2 基于 GOP 和基于帧的实现方法

针对图 2 的两级多假设重构方案, 文中提出了两种实现方法, 一种是基于图像组的实现方法 (Gw_2sMHR), 另一种是基于帧的实现方法 (Fw_2sMHR)。图 4 以 GOP 长度等于 4 为例给出了两种实现方法的示意图, “S1” 和 “S2” 分别表示阶段 1 重构和阶段 2 重构, “KEY” 表示关键帧, “CS” 和 “PXL” 分别表示观测域多假设预测和像素域多假设预测, 带箭头实线表示参考帧的参照方向, 带箭头虚线表示重构阶段 1 到重构阶段 2 的切换。

对于图 4 (a) 基于图像组的实现方法, 重构阶段 1 到重构阶段 2 的切换以图像组(GOP)为单位进行。具体步骤为: (1)对关键帧进行帧内重构; (2)对当前 GOP 内的所有非关键帧进行第 1 阶段的多假设预测重构; (3)以第 1 阶段的重构帧为目标匹配帧, 对当

前 GOP 内的非关键帧进行第 2 阶段的像素域多假设预测重构。在第 2 阶段的多假设预测中, 由于此时已对当前 GOP 内的各个非关键帧完成了第 1 阶段重构, 因此每个非关键帧将增加一个额外的后向参考帧, 以利用双向的时间信息, 得到更准确的预测帧。

对于图 4 (b) 基于帧的实现方法, 重构阶段 1 到重构阶段 2 的切换以帧为单位进行。具体步骤为: (1)对关键帧进行帧内重构; (2)对当前非关键帧进行第 1 阶段的多假设预测重构; (3)以第 2 阶段的重构帧为目标匹配帧, 对当前非关键帧进行第 2 阶段的像素域多假设预测重构, 至此, 当前非关键帧重构完成; (4)对下一非关键帧重复步骤(2)和(3), 直到当前 GOP 内的所有非关键帧重构完成。由于当前非关键帧在第 1 阶段多假设预测时利用到了前一帧更加精确的两阶段预测重构信息, 因此可以得到更精确的预测帧。

需要指出的是, 上述两种实现方法均各有优点。基于图像组的实现方法在重构阶段 2 可利用双向的参考帧信息; 而基于帧的实现方法在重构阶段 1 则具有更高质量的参考帧信息, 从而使其在重构阶段 2 具有更高质量的目标匹配帧。通常, 当视频信号的内容具有平移特点时, 基于图像组的实现方法会有更高的重构质量, 因为此时双向的参考帧能提供更多有用的时间信息。

4 仿真结果与分析

文中仿真实验分为 4 部分, 第 1 部分分析像素域预测阶段的分块大小及重叠范围对重构性能的影响; 第 2 部分验证所提 2sMHR 重构方案的有效性; 第 3 部分将所提 2sMHR 重构方案与现有最好的多假设预测算法进行比较; 第 4 部分分析算法的时间复杂度。2sMHR 重构算法的少量初步仿真结果我们已在刚结束的 ICIP 国际会议上展示^[19], 本文给出了更详细的算法描述, 更全面的仿真结果以及更详尽的结果分析。

4.1 仿真条件

为方便表示, 在后文中, 将阶段 1 观测域多假设预测方法记为 MRMH, 将基于图像组实现方法的 2sMHR 重构方案记为 Gw-2sMHR, 将基于帧实现方法的 2sMHR 重构方案记为 Fw-2sMHR。

实验中采用随机高斯观测矩阵, 分块大小为 16×16 。每个 GOP 的关键帧采样率为 0.7, 采用帧内多假设预测重构; 非关键帧采样率依次取 0.1 ~ 0.5, 采用文中提出的 2sMHR 算法进行重构。2sMHR 算法的多假设预测基于多参考帧^[20]实现, 最

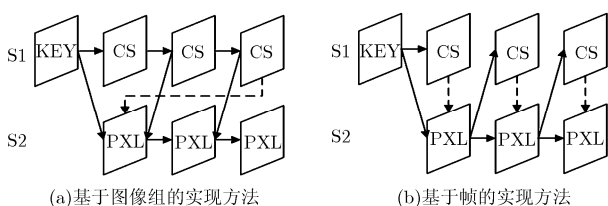


图 4 2sMHR 的两种实现方法

大参考帧数量设为 4(2 帧长期参考帧, 2 帧短期参考帧), 对于 Gw_2sMHR, 在预测阶段 2 的最大参考帧数量为 5, 这是因为 Gw_2sMHR 在重构阶段 2 有一个额外的最近邻后向参考帧, 以利用更多的时间信息。运动搜索范围为当前块在参考帧对应位置的周围±16 像素。第 1 阶段多假设预测最优假设块数量的选取, 根据采样率不同而不同, 采样率 0.1, 0.2, 0.3, 0.4, 0.5 分别对应的假设块数量为 81 块, 100 块, 121 块, 144 块, 169 块^[20], 分块大小为 16×16; 第 2 阶段像素域多假设预测的最优假设块数量固定为 8, 分块大小为 8×8, 相邻块之间有 4 像素重叠。采用 BCS-SPL-DDWT^[21]算法进行残差重构。

4.2 分块大小及重叠范围对重构性能的影响分析

由前文分析(3.1 节)可知, 在像素域多假设预测环节, 分块大小及重叠范围对预测性能有重要影响。本文通过大量的仿真实验综合考虑重构质量和重构时间到经验值, 部分实验结果见表 1 和表 2。

表 1, 表 2 给出了当 GOP 大小为 16, 关键帧采样率 0.7, 非关键帧采样率 0.1 时, 由 2sMHR 重构所得非关键帧的平均 PSNR 及平均重构时间, 试验环境为搭载 Inter Core i5 3.5 GHz 处理器, 内存 8 GB 的 PC 电脑, 操作系统为 Windows 7 SP1, 仿真软件为 MATLAB R2016a。其中, 重叠范围为 0 表示分块无重叠, 重叠范围为 x 表示相邻块之间有 x 行或 x 列像素重叠。由表 1, 表 2 可见, 当分块大小为 8×8, 相邻块重叠 4 行或 4 列像素时, Foreman 和 Soccer 的平均 PSNR 均高于其他组合, 且重构时间也不是太大, 因此综合考虑重构性能和算法复杂度, 本文接下来的实验中选取了该经验值。

4.3 2sMHR 重构方案的有效性验证

仿真中 GOP 长度为 16, 采用 4 个 qcif@15 Hz 的标准视频序列(每个序列取前 96 帧, 共 6 组 GOP) Foreman, Coastguard, Hall, Suzie 进行测试, 3 种多假设预测算法在不同采样率下的平均 PSNR 值示于图 5。

表 1 不同分块大小和不同重叠范围时非关键帧平均重构质量 PSNR(dB)

	分块大小\ 重叠范围	平均重构质量 PSNR(dB)				
		0	1	2	4	8
Foreman	4×4	32.21	32.42	32.50	-	-
	8×8	32.47	32.86	32.95	33.01	-
	16×16	32.20	32.81	32.74	32.85	32.91
Soccer	4×4	27.40	27.58	27.64	-	-
	8×8	27.69	28.14	28.23	28.35	-
	16×16	27.33	28.04	28.10	28.13	28.04

表 2 不同分块大小和不同重叠范围时非关键帧平均重构时间(s)

	分块大小\ 重叠范围	平均重构时间(s)				
		0	1	2	4	8
Foreman	4×4	2.57	3.20	4.93	-	-
	8×8	2.58	2.82	3.30	4.60	-
	16×16	3.01	4.16	3.58	4.22	4.96
Soccer	4×4	2.68	3.28	5.00	-	-
	8×8	2.67	2.82	3.22	4.66	-
	16×16	3.10	4.15	3.60	4.25	5.09

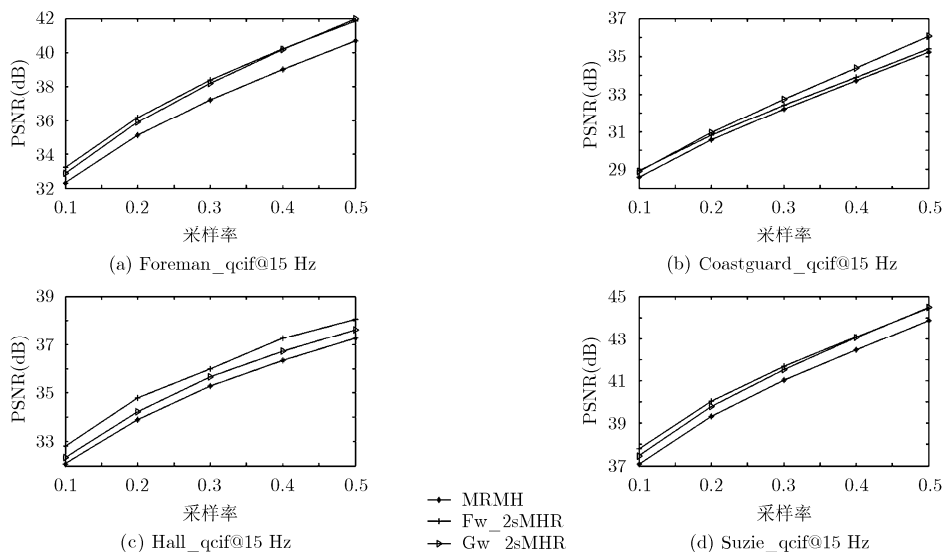


图 5 文中不同重构算法的率失真性能对比

由图 5 可见，在进行两阶段重构以后，Gw_2sMHR 和 Fw_2sMHR 均在 MRMH 基础上有较明显提升，表 3 给出了采样率为 0.1 时，Gw_2sMHR 和 Fw_2sMHR 较 MRMH 的 PSNR 提升情况。

2sMHR 重构质量提升的主要原因为，2sMHR 在第 2 阶段的像素域多假设预测中采取了重叠分块以及更小的分块，从而更充分地挖掘了视频帧间相关性，有效地减小块效应并抑制重构噪声。此外，从图 5 还可看出，对于 Coastguard 序列，Gw_2sMHR 较 Fw_2sMHR 具有更高的重构质量，而对于其它 3 个序列则反之。这是因为 Coastguard 序列的视频内容具有背景平移的特点，此时 Gw_2sMHR 的双向参考帧比 Fw_2sMHR 的单向参考帧提供了更多的有用信息，因此 Gw_2sMHR 对于

Coastguard 序列的重构质量较明显地高于 Fw_2sMHR。

为了更清楚地看到所提 2sMHR 算法对于减小块效应的效果，图 6，图 7，图 8 分别给出了 Foreman 第 48 帧，Soccer 第 52 帧以及 Suzie 第 53 帧进行局部放大后的重构视觉效果图。

由图 6 中 Foreman 的主观质量对比可见，在 MRMH 重构效果图中，Foreman 的嘴巴四周存在着较明显的块效应，而在 Gw_2sMHR 及 Fw_2sMHR 的重构效果图中，该区域的块效应得以有效减小，这是因为第 2 阶段的像素域多假设预测采用了重叠分块，从而平滑了块与块之间的边界，减小块效应；由图 7 中 Soccer 的主观质量对比可见，在 MRMH 的重构效果图中，足球运动员的小腿及手臂区域存在着较明显的块效应，并伴随有重构噪声(白色斑点)，而在 Gw_2sMHR 及 Fw_2sMHR 的重构效果图中，该区域的块效应以及重构噪声均得以有效抑制；图 8 中 Suzie 下巴左侧区域的块效应也得到了不同程度的减小，这同样归因于第 2 阶段的重叠分块多假设预测，可见，文中的 2sMHR 重构算法能有效减小块效应，抑制重构噪声。

表 3 Gw_2sMHR 和 Fw_2sMHR 较 MRMH 的 PSNR 提升情况(dB)

	Foreman	Coastguard	Hall	Suzie
Gw_2sMHR	+0.58	+0.32	+0.25	+0.38
Fw_2sMHR	+0.93	+0.35	+0.71	+0.70

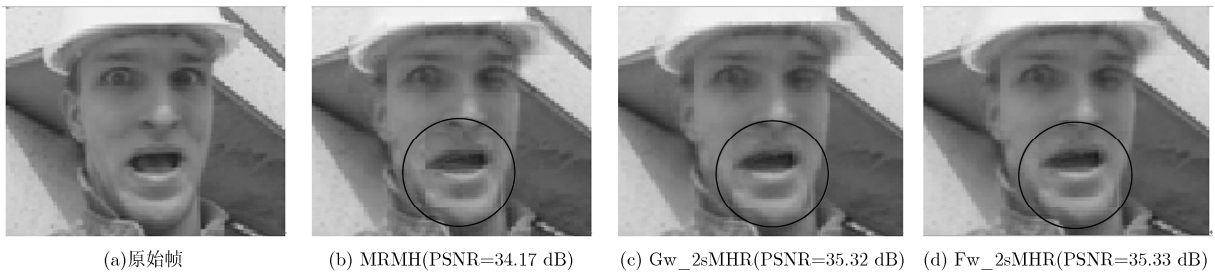


图 6 Foreman 第 48 帧主观质量对比

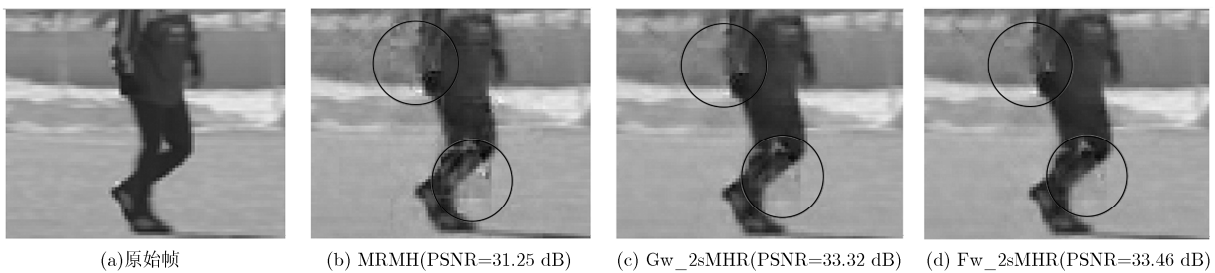


图 7 Soccer 第 52 帧主观质量对比

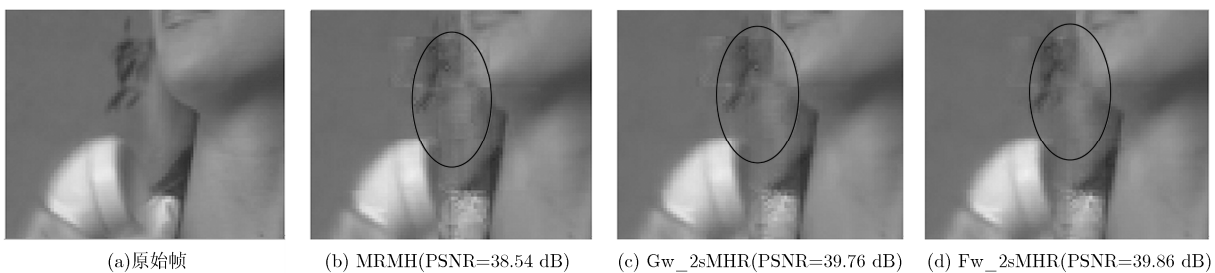


图 8 Suzie 第 53 帧主观质量对比

4.4 2sMHR 与最新文献结果的对比

本小节对比了本文所提算法和最新文献提出的视频压缩感知重构算法^[17], 仿真实验分为两部分, 第 1 部分将文中的 Gw_2sMHR 和 Fw_2sMHR 算法与基于加权弹性网的多假设预测算法 wElasticNet^[17]进行对比, 把文献[17]提出的加权弹性网模型也基于文中的多参考帧实现, 但未加入该文献的其他改进技术; 为了和文献[17]的整个重构方案作比较, 在第 2 部分实验中, 文中算法 Gw_2sMHR 和 Fw_2sMHR 的仿真条件完全按照该文献进行设置, 所得结果与该文献所给的仿真结果数据做比较。

(1)2sMHR 与 wElasticNet 对比: 仿真中 GOP 长度为 16, 在文中的多参考帧条件下, Gw_2sMHR, Fw_2sMHR 及 wElasticNet 在各采样率下的平均 PSNR 值示如图 9 所示。

由图 9 可见, Gw_2sMHR 和 Fw_2sMHR 对于各序列的率失真性能均较明显地高于 wElasticNet, 在采样率为 0.1 时, Gw_2sMHR 对于 Foreman, Coastguard, Hall 及 Suzie 序列的平均 PSNR 值比 wElasticNet 分别高了 0.87 dB, 0.53 dB, 0.19 dB 及 1.14 dB; Fw_2sMHR 对于 Foreman, Coastguard, Hall 及 Suzie 序列的平均 PSNR 值比 wElasticNet 分别高了 1.22 dB, 0.57 dB, 0.64 dB 及 1.46 dB。wElasticNet 算法本质上属于阶段 1 的观测域多假设预测算法, 不同之处在于其利用了加权弹性网正则化模型来求解线性权值, 相比于 Tikhonov 正则化模型有更高的权值求解精度, 但算法复杂度更高。2sMHR 算法之所以比基于加权弹性网模型的多假设预测算法有更高的重构质量, 同样归因于第 2 阶段的像素域重叠分块多假设预测, 消

除了观测域预测时产生的块效应, 提高了预测精度, 进而提高了预测残差的稀疏性, 最终提高了重构性能。

(2)2sMHR 与文献[17]的仿真结果对比: 在本部分仿真实验中, 仿真条件完全按照文献[17]给出的设置, 与文献[17]所给的仿真结果数据进行比较。仿真中 GOP 长度为 8, 采用 3 个 cif@30 Hz 的标准视频序列(每个序列取前 88 帧, 共 11 组 GOP) Foreman, Coastguard, Hall 进行测试, 各算法在不同采样率下的平均 PSNR 示于表 4。

由表 4 可见, 2sMHR 算法的 PSNR 值明显高于 Up-Se-AWEN-HHP^[17], 在采样率为 0.1 时, Gw_2sMHR 算法对于 Foreman, Coastguard, Hall 序列的平均 PSNR 值比 Up-Se-AWEN-HHP 算法分别高出了 0.79 dB, 0.83 dB, 1.96 dB; Fw_2sMHR 算法对于 Foreman, Coastguard, Hall 序列的平均 PSNR 值比 Up-Se-AWEN-HHP 算法分别高出了 1.14 dB, 0.80 dB, 1.75 dB。可见, 通过进行第 2 阶段的重叠分块多假设预测, 文中 Gw_2sMHR 和 Fw_2sMHR 算法较现有最好的多假设预测算法有更高的视频重构质量。

4.5 时间复杂度分析

本节给出了 2sMHR 算法的重构时间消耗及时间复杂度分析。实验中 GOP 长度为 16, 实验条件按照 4.1 节进行设置, 不同多假设预测算法对每个非关键帧(qcif:176×144)进行多假设预测平均所消耗的时间示于表 5。实验环境为搭载 Inter Core i5 1.3 GHz 处理器, 内存 4 GB 的 MacBook Air(2013 Mid, 操作系统为 OS X 10.9.4, 仿真软件为 MATLAB R2013a。

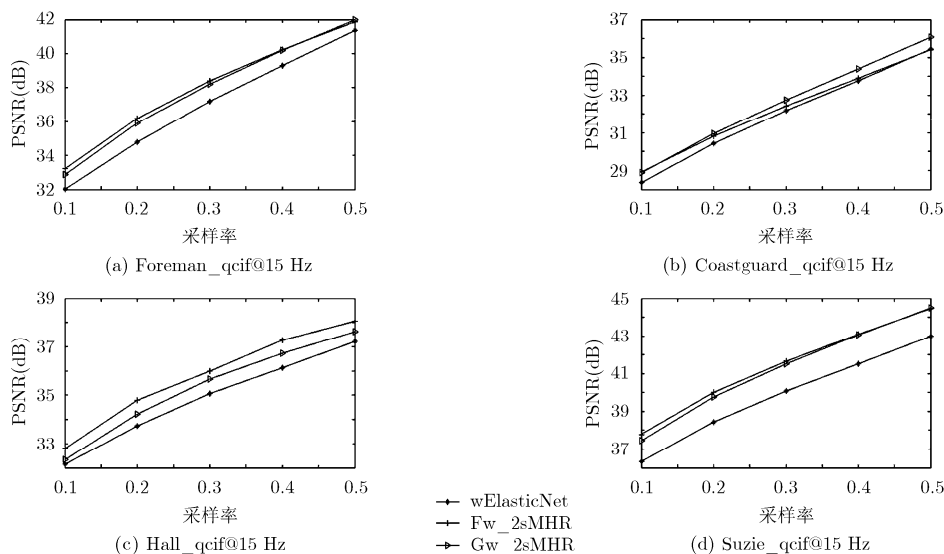


图 9 GOP=16 时, 2sMHR 与 wElasticNet 的率失真性能对比

表 4 GOP=8 时, 3 种算法的 PSNR 比较(dB)

视频序列	算法	采样率				
		0.1	0.2	0.3	0.4	0.5
Foreman	Up-Se-AWEN-HHP	36.17	38.15	39.65	40.86	42.08
	Fw_2sMHR	37.31	39.17	40.51	41.60	42.63
	Gw_2sMHR	36.96	38.79	40.17	41.29	42.32
Coastguard	Up-Se-AWEN-HHP	30.86	32.94	34.67	36.29	37.99
	Fw_2sMHR	31.66	33.44	34.93	36.35	37.79
	Gw_2sMHR	31.69	33.58	35.23	36.82	38.40
Hall	Up-Se-AWEN-HHP	35.15	36.50	37.51	38.37	39.11
	Fw_2sMHR	36.90	38.48	39.62	40.51	41.29
	Gw_2sMHR	37.11	38.54	39.60	40.43	41.16

表 5 各算法的时间复杂度(s)

算法	采样率				
	0.1	0.2	0.3	0.4	0.5
wElasticNet ^[17]	29.2	54.7	98.2	156.5	237.1
MRMH	2.9	3.1	3.4	3.8	3.9
Fw_2sMHR	7.2	7.5	7.8	7.9	8.4
Gw_2sMHR	8.2	8.5	8.8	9.1	9.3

由表 5 可见, 首先, 对比 Gw_2sMHR, Fw_2sMHR 及 MRMH 可知, 文中的两级多假设重构算法 Gw_2sMHR 和 Fw_2sMHR 的重构时间消耗明显地高于 MRMH 算法, 这是因为 Gw_2sMHR 和 Fw_2sMHR 在 MRMH 的基础上增加了第 2 阶段的重叠分块多假设预测; 其次, 对比 Gw_2sMHR 和 Fw_2sMHR 可知, Gw_2sMHR 的重构时间消耗略高于 Fw_2sMHR, 这是因为在第 2 阶段的多假设预测中, Gw_2sMHR 比 Fw_2sMHR 多了 1 个额外的后向参考帧。

值得注意的是, MRMH, Fw_2sMHR 及 Gw_2sMHR 算法的重构时间消耗均明显低于 wElasticNet, 且重构时间随采样率的增加变化不大, 而 wElasticNet 算法的重构时间消耗随采样率的增加而快速增加。这是因为, wElasticNet 通过迭代的方式求解线性权值, 一方面, 迭代次数随着采样率的增加而增加, 另一方面, 每次迭代中涉及大量的矩阵运算(如 cholesky 分解, 矩阵求逆, 矩阵相乘, 排序等), 随着采样率的增加, 每次迭代中参与矩阵运算的矩阵维度也随之增加, 再加上迭代次数的增加, 从而导致最终时间复杂度的显著增加; 而本文算法由于直接通过解析求得线性权值, 无需进行迭代, 涉及的矩阵运算较少(只需一次的矩阵求逆操作), 再加上采取了最优多假设块选择^[20], 大大减小

了参与线性权值求解的假设块数量, 从而时间复杂度较低。因此, 本文算法的时间复杂度低于 wElastic, 且受采样率的影响较小。

5 结束语

信号的稀疏性是压缩感知精确重构的前提, 多假设预测是挖掘视频信号稀疏表示潜力的有效手段。本文提出了两级多假设重构方案(2sMHR), 所提重构方案结合了观测域多假设预测和像素域多假设预测各自的优点, 有效解决了现有观测域多假设预测算法存在的块效应明显问题, 提高了视频压缩感知的重构性能。针对所提 2sMHR 重构方案, 文中设计了两种实现方法, 基于帧的实现方法(Fw_2sMHR)在重构阶段 1 具有更高质量的参考帧信息, 基于图像组的实现方法(Gw_2sMHR)在重构阶段 2 可利用双向的参考帧信息。仿真结果表明, 通过进行第 2 阶段的像素域重叠分块多假设预测, 所提 2sMHR 算法在低采样率(0.1)下相比于阶段 1 的观测域多假设预测算法有 0.25~0.93 dB 的 PSNR 提升, 与现有最好的多假设预测重构方案相比有 0.79~1.96 dB 的 PSNR 增益, 以及更低的重构复杂度。

鉴于两级多假设重构算法(2sMHR)具有较高的重构质量, 且基于帧的实现方法(Fw_2sMHR)和基于图像组的实现方法(Gw_2sMHR)各有优点, 能否将两种实现方法相结合, 通过进行多级的多假设预测重构来进一步提升视频重构质量, 是我们下一步的研究重点。

参考文献

- [1] LIU Y and PADOS D A. Compressed-sensed-domain L1-PCA video surveillance[J]. *IEEE Transactions on Multimedia*, 2016, 18(3): 351-363. doi: 10.1109/TMM.2016.2514848.
- [2] GUO J, SONG B, and DU X. Significance evaluation of video data over media cloud based on compressed sensing[J]. *IEEE Transactions on Multimedia*, 2016, 18(7): 1297-1304. doi: 10.1109/TMM.2016.2564100.
- [3] REHMAN A U, SHAH G A, and TAHIR M. Compressed sensing based adaptive video coding for resource constrained devices[C]. *IEEE International Wireless Communications and Mobile Computing Conference*, Paphos, Cyprus, 2016: 170-175.
- [4] WANG J, GUPTA M, and SANKARANARAYANAN A C. LiSens — A scalable architecture for video compressive sensing[C]. *IEEE International Conference on Computational Photography*, Houston, TX, 2015: 1-9.
- [5] LLULL P, LIAO X J, YUAN X, et al. Coded aperture compressive temporal imaging[J]. *Optics Express*, 2013, 21(9):

- 10526–10545. doi: 10.1364/OE.21.010526.
- [6] HOSSEINI M S and PLATANIOTIS K N. High-accuracy total variation with application to compressed video sensing [J]. *IEEE Transactions on Image Processing*, 2014, 23(9): 3869–3884. doi: 10.1109/TIP.2014.2332755.
- [7] YANG J B, YUAN X, LIAO X J, *et al.* Video compressive sensing using Gaussian mixture models[J]. *IEEE Transactions on Image Processing*, 2014, 23(11): 4863–4878. doi: 10.1109/TIP.2014.2344294.
- [8] 常佩, 覃团发, 唐振华. 基于联合总变分最小化的视频压缩感知重建算法[J]. *电子学报*, 2014, 42(12): 2415–2421. doi: 10.3969/j.issn.0372-2112.2014.12.012.
- CHANG K, QIN T F, and TANG Z H. Reconstruction algorithm for compressed sensing of video based on joint total variation minimization[J]. *Acta Electronica Sinica*, 2014, 42(12): 2415–2421. doi: 10.3969/j.issn.0372-2112.2014.12.012.
- [9] ZHAO C, MA S W, ZHANG J, *et al.* Video compressive sensing reconstruction via reweighted residual sparsity[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2016, to be published. doi: 10.1109/TCSVT.2016.2527181.
- [10] MUN S and FOWLER J E. Residual reconstruction for block-based compressed sensing of video[C]. *IEEE Data Compression Conference, Snowbird*, 2011: 183–192.
- [11] NARAYANAN S and MAKUR A. Compressive coded video compression using measurement domain motion estimation [C]. *IEEE International Conference on Electronics, Computing and Communication Technologies*, Bangalore, 2014: 1–6.
- [12] GUO J, SONG B, LIU H X, *et al.* Motion estimation in measurement domain for compressed video sensing[C]. *IEEE International Conference on Computer and Information Technology*, Xi'an, 2014: 441–445.
- [11] DO T T, CHEN Y, NGUYEN D T, *et al.* Distributed compressed video sensing[C]. *IEEE International Conference on Image Processing*, Cairo, 2009: 1393–1396.
- [12] TRAMEL E W and FOWLER J E. Video compressed sensing with multihypothesis[C]. *IEEE Data Compression Conference, Snowbird*, 2011: 193–202.
- [13] AZGHANI M, KARIMI M, and MARVASTI F. Multihypothesis compressed video sensing technique[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2016, 26(4): 627–635. doi: 10.1109/TCSVT.2015.2418586.
- [14] CHEN J, CHEN Y, QIN D, *et al.* An elastic net-based hybrid hypothesis method for compressed video sensing[J]. *Multimedia Tools & Applications*, 2013, 74(6): 2085–2108. doi: 10.1007/s11042-013-1743-y.
- [15] KUO Y H, WU K, and CHEN J. A scheme for distributed compressed video sensing based on hypothesis set optimization techniques[J]. *Multidimensional Systems and Signal Processing*, 2017, 28(1): 129–148. doi: 10.1007/s11045-015-0337-4.
- [16] GAN L. Block compressed sensing of natural images[C]. *IEEE International Conference on Digital Signal Processing*, Cardiff, 2007: 403–406.
- [17] OU W F, YANG C L, LI W H, *et al.* A two-stage multi-hypothesis reconstruction scheme in compressed video sensing[C]. *IEEE International Conference on Image Processing*, Phoenix, AZ, USA, 2016: 2494–2498.
- [18] 杨春玲, 欧伟枫. CVS 中基于多参考帧的最优多假设预测算法[J]. *华南理工大学学报(自然科学版)*, 2016, 44(1): 1–8. doi: 10.3969/j.issn.1000-565X.2016.01.001.
- YANG C L and OU W F. Multi-reference frames-based optimal multi-hypothesis prediction in compressed video sensing[J]. *Journal of South China University of Technology (Natural Science Edition)*, 2016, 44(1): 1–8. doi: 10.3969/j.issn.1000-565X.2016.01.001.
- [19] MUN S and FOWLER J E. Block compressed sensing of images using directional transforms[C]. *IEEE International Conference on Image Processing*, Cairo, 2009: 3021–3024.
- 欧伟枫: 男, 1990 年生, 硕士, 研究方向为视频压缩感知。
- 杨春玲: 女, 1970 年生, 教授, 博士生导师, 研究方向为图像/视频压缩。
- 戴超: 男, 1993 年生, 硕士生, 研究方向为视频压缩感知。