

视频压缩感知多假设局部增强重构算法

汤瑞东¹ 杨春玲¹ 禩韵怡¹

摘要 在基于多假设预测的视频压缩感知重构中,不同图像块对应的假设集匹配程度差异较大,因此重构难度差异明显.本文提出多假设局部增强重构算法(Local enhancement reconstruction algorithm based on multi-hypothesis prediction, MH-LE),利用帧间相关性对图像块进行分类后针对运动图像块提出像素域双路匹配策略,通过强化图像块基本特征来提高相似块匹配效果,获取更高质量的假设集;同时将结构相似度评价标准引入假设块权值分配过程,提高预测精度.仿真结果表明,所提算法的重构质量明显优于其他多假设预测重构算法.和基于组稀疏的重构算法相比,所提算法具有更快的重构速度,在大部分的采样率条件下具有更高的重构质量.

关键词 视频压缩感知,多假设预测,相似块匹配,增强重构

引用格式 汤瑞东,杨春玲,禩韵怡.视频压缩感知多假设局部增强重构算法.自动化学报,2022,48(8):1984–1993

DOI 10.16383/j.aas.c190408

Local Enhancement Reconstruction Algorithm Based on Multi-hypothesis Prediction in Compressed Video Sensing

TANG Rui-Dong¹ YANG Chun-Ling¹ XUAN Yun-Yi¹

Abstract In multi-hypothesis prediction-based compressed video sensing reconstruction algorithms, the matching degrees of the hypothesis set corresponding to different image blocks are quite different, so the reconstruction difficulty of different blocks is obviously different. In this paper, a local enhancement reconstruction algorithm based on multi-hypothesis (MH-LE) is proposed. Image blocks are classified into two categories and a pixel domain dual channel matching strategy is proposed for moving image blocks, where the basic features of the image blocks are enhanced to improve the matching effectivity of similar blocks and obtain a higher quality hypothesis set. Besides, the structural similarity evaluation criteria are introduced into the matching block weight assignment process to improve prediction accuracy. The simulation results show that the reconstruction quality of the proposed algorithm is superior to other multi-hypothesis prediction-based reconstruction algorithms. Compared with the group sparsity-based reconstruction algorithms, the proposed algorithm possesses faster reconstruction speed and higher reconstruction quality at most sampling rates.

Key words Compressed video sensing (CVS), multi-hypothesis prediction, similar block matching, enhancement reconstruction

Citation Tang Rui-Dong, Yang Chun-Ling, Xuan Yun-Yi. Local enhancement reconstruction algorithm based on multi-hypothesis prediction in compressed video sensing. *Acta Automatica Sinica*, 2022, 48(8): 1984–1993

压缩感知 (Compressed sensing, CS)^[1] 理论突破了 Nyquist 采样理论的瓶颈,指出对于稀疏信号或可压缩信号,可以用欠采样得来的数据高概率恢复出原信号,采样率下限不再受限于信号带宽.视频压缩感知 (Compressed video sensing, CVS) 是

基于 CS 理论的视频信号编解码方案,在编码端直接通过线性投影合并采样与压缩过程,将计算压力转移至重构端,十分适用于无线多媒体传感网络、无线视频监控等采样端资源受限的应用场景.

重构算法是压缩感知理论研究中的核心任务.在图像压缩感知重构中,文献 [2] 提出的分块压缩感知 (Block-based compressed sensing, BCS) 成为大多数重构算法采用的处理方式,整个编解码过程以独立图像块作为基本处理单元,显著降低了存储传输压力.平滑 Landweber 投影算法 (Smoothed projected landweber, SPL)^[3] 通过维纳滤波减轻分块处理造成的块效应,因其重构的高效性成为图像压缩感知中的经典重构算法,并广泛应用于 CVS

收稿日期 2019-05-27 录用日期 2019-10-28

Manuscript received May 27, 2019; accepted October 28, 2019

广东省自然科学基金重点项目 (2017A030311028), 广东省自然科学基金 (2016A030313455) 资助

Supported by Key Program of Natural Science Foundation of Guangdong Province (2017A030311028) and Natural Science Foundation of Guangdong Province (2016A030313455)

本文责任编辑 黄庆明

Recommended by Associate Editor HUANG Qing-Ming

1. 华南理工大学电子与信息学院 广州 510641

1. School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641

重构. 组稀疏重构算法 (Group-based sparse representation, GSR)^[4] 提出组稀疏概念, 以相似块组进行奇异值分解后的奇异值具备稀疏性作为假设前提, 对每个相似块组建立自适应稀疏表示字典. 近年来, 研究人员尝试利用深度学习框架处理图像压缩感知问题, 并提出了几种信号重构的实现框架^[5-6]. 基于深度学习框架的重构算法具有极高的重构速度, 但灵活性较差, 在不同采样率下均需训练出对应模型.

视频信号在图像信号的基础上增加了时间维度, 利用好帧间相关性是提高 CVS 重构质量的关键. 文献 [7] 将多假设预测 (Multi-hypothesis prediction, MH) 运用到 CVS 重构中, 并通过 SPL 算法对残差进行重构, 提出了多假设预测重构算法 (Multi-hypothesis prediction BCS-SPL, MH-BCS-SPL). 在此基础上, 文献 [8-9] 对假设集的构造方式提出了优化方案, 有效提升了多假设预测精度. 文献 [10] 通过解弹性网回归问题求解各匹配块权重, 文献 [11] 在其基础上对 l_2 范数正则化项引入了观测域距离权重, 进一步提高了预测精度. 文献 [12] 提出一种基于多参考帧的两阶段多假设预测算法 (Two-stage multi-hypothesis reconstruction, 2sMHR), 在观测域多假设得到的重构帧基础上对视频帧进行重叠分块后再进行一次像素域多假设预测, 有效减轻了重构帧的块效应, 明显提高了视频重构质量. 为了提升图像组 (Group of picture, GOP) 中间帧的重构质量, PBCR-DCVS 算法 (Position-based cross reconstruction distributed CVS)^[13] 在对关键帧进行二次迭代重构的基础上提出位置交叉重构策略, 改进非关键帧的重构顺序并根据残差大小自适应扩大搜索窗. 基于 MH 的 CVS 重构算法重构速度较快, 具有较高的实用性. 与此同时, 受 GSR 算法在图像压缩感知重构的出色表现所启发, 研究人员基于 GSR 算法提出了一批新的视频压缩感知重构算法. 文献 [14] 提出加权残差稀疏算法 (Reweighted residual sparsity, RRS), 根据信号残差的稀疏性建立求解模型, 在分裂 Bregman 迭代求解的过程中不断更新信号残差 DCT (Discrete cosine transform) 系数各分量的权重, 实现重构帧质量的不断提升. 文献 [15] 提出基于结构相似度 (Structural similarity, SSIM) 的帧间 (Inter frame, InterF) 组稀疏表示重构算法 (SSIM-InterF-GSR), 将结构相似度 (SSIM) 作为相似块选取时的匹配准则, 提高了相似块组的组稀疏特性, 并引入阶梯递减方案调整组内相似块个数, 提高了重构精度. 基于 GSR 的 CVS 重构算法重构质量较高, 但因

其繁琐的迭代过程导致算法复杂度较高.

多假设预测因其对视频帧间相关性的高效利用成为了 CVS 重构算法的关键技术之一, 但在重构中不同运动特征的图像块找到高质量匹配块的难易程度明显不同, 由于假设集与当前待重构图像块的匹配程度对图像块重构质量影响较大, 低质量的假设集合会极大增加图像块的重构难度. 针对此问题, 本文提出了一种局部增强的多假设预测方案 (Local enhancement reconstruction algorithm based on multi-hypothesis prediction, MH-LE), 对初始重构质量不好的图像块进行增强重构. 主要创新点有: 1) 根据图像块的运动情况对图像块进行分类, 采用不同的多假设预测方法进行预测重构; 2) 提出像素域双路匹配策略, 同时利用双边滤波前后的图像进行相似块匹配, 提高假设集质量; 3) 将结构相似度 (SSIM) 评价标准引入假设块权重分配过程, 提高预测精度.

1 CVS 多假设预测算法

在采集端, 传统视频编码需要先以高于信号最高频率的两倍的速率采集视频信号, 再通过一系列如运动估计、运动补偿、信号变换等高复杂度的压缩算法来丢弃冗余数据. 而 CVS 直接通过观测矩阵 Φ 对视频帧图像块 \mathbf{x} 进行随机采样, 得到维数远小于原信号的观测信号 \mathbf{y} , 同时实现信号的采样与压缩, 观测模型为

$$\mathbf{y} = \Phi \mathbf{x} \quad (1)$$

在分布式 CVS 信号采集中, 通常以 GOP 为单位对 GOP 内每个视频帧单独进行采样. GOP 首帧为关键帧, 其他帧为非关键帧, 关键帧的采样率一般比非关键帧高, 在重构阶段能以更高的重构质量为非关键帧重构提供高质量边信息. 同时, 考虑到实际应用过程中的采样资源和传输压力等因素, 在对各个视频帧进行观测的过程中采用非重叠分块的 BCS 独立观测方法, 将每个图像块的观测值 \mathbf{y} 分别传递到解码端再进行联合解码重构. 在 CVS 实现中, 目前主要采用空间多路复用方式 (Spatial multiplexing, SM). SM 方式可通过多个单像素相机 (Single pixel camera, SPC) 同时进行逐帧采样, 将多个 SPC 得到的观测结果拼接后即为该帧的目标观测值^[16-17].

在重构端, 对当前帧中的某个图像块进行多假设预测时, 首先在参考帧中寻找当前块的相似块组成假设集, 然后利用假设块的线性组合预测当前块, 通过最小化观测域欧氏距离得到最优线性加权系数矢量

$$\hat{w} = \arg \min_w \|y - \Phi H w\|_2^2 \quad (2)$$

其中, H 为当前块 x 对应的假设集矩阵, H 中的每一列为一个假设块的列矢量, $H\hat{w}$ 即为当前块的预测结果.

基于 Tikhonov 正则化的多假设预测加入了 l_2 范数约束项, 具有简单的闭式解, 同时能够在一定限度内处理假设块高度相关的情况^[7]. 求解目标为

$$\hat{w} = \arg \min_w \|y - \Phi H w\|_2^2 + \lambda \|\Gamma w\|_2^2 \quad (3)$$

其中, Γ 为对角矩阵, 对角线上元素为对应的假设块与当前块的观测值之间的欧氏距离. 可直接求得式 (3) 的闭式解

$$\hat{w} = [(\Phi H)^T(\Phi H) + \lambda \Gamma^T \Gamma]^{-1}(\Phi H)^T y \quad (4)$$

基于弹性网的多假设预测进一步引入 l_1 范数约束项, 通过强调 w 的稀疏性对假设块进行自适应选取, 在此基础上文献 [11] 引入距离惩罚矩阵 Γ 对 w 进行修正, 提出自适应加权弹性网 (Adaptive weighted elastic net, AWEN)

$$\hat{w} = (1 + \lambda_2) \arg \min_w \|y - \Phi H w\|_2^2 + \lambda_1 \|\Gamma w\|_1 + \lambda_2 \|\Gamma w\|_2^2 \quad (5)$$

式 (5) 可通过 LARS-EN 算法^[18] 进行迭代求解, 在假设块相关性较差的情况下该方法可得到更好的加权系数, 但算法复杂度较高.

2 多假设局部增强重构算法 (MH-LE)

多假设预测算法利用参考帧作为边信息对当前帧进行预测, 假设集中假设块和当前块的匹配程度是决定预测质量的关键因素之一, 能否找到合适的匹配块在一定程度上反映了准确重构当前块的难度. 根据与相邻帧的相关程度, 可将图像块分为稳定块和运动块两个状态类别. 一方面, 不同状态的图像块在相邻帧中找到高度匹配块的难易程度不同. 对于基本没有出现位移的稳定块, 直接在相邻帧以当前位置为中心建立搜索窗进行搜索即可高概

率地找到高度匹配块; 而重构运动块时需利用运动估计等方法来确定搜索窗位置, 且无法保证一定能找到有效匹配块. 另一方面, 假设集匹配程度不同导致不同状态的图像块适用的预测方法不同. 当假设块与当前块高度相关时, 基于弹性网的多假设预测算法会出现不稳定现象, 算法时间复杂度急剧增长, 预测精度反而不如计算量很小的基于 Tikhonov 正则化的预测方法 (Multi-hypothesis Tikhonov, MH_Tikhonov)^[11, 13]. 由于稳定块较容易找到高度相关的匹配块, 采用 MH_Tikhonov 方法即可高质量重构, 引入过多的正则化条件反而会因过度强调泛化能力而造成欠拟合. 因此, 本文提出多假设局部增强重构算法 (MH-LE), 对不同类别图像块采用不同的重构策略, 通过进一步挖掘像素域信息对运动块进行二次重构, 提升视频整体重构质量.

2.1 MH-LE 算法框架

本文所提算法采用 coarse-to-fine 策略, 先进行粗尺度的整体重构, 再针对部分高概率低质量的图像块进行精细化重构. 重构算法实现的整体框架如图 1 所示.

关键帧采用重构性能较优的 GSR 算法进行独立重构, 为非关键帧的重构提供高质量的参考信息. 考虑到 MH_Tikhonov 预测算法的高效性, 本文将其作为非关键帧的初始重构算法. 在此基础上, 通过计算当前帧与相邻帧的帧差情况, 根据图像块与相邻帧的相关程度将图像块分为稳定图像块和运动图像块, 对运动图像块进行增强重构. 在增强重构中, 采用提出的像素域双路匹配策略来构造假设集, 然后对当前运动块进行 AWEN 多假设预测, 在预测过程中同时结合观测距离和结构相似度距离调整各假设块权值. 最后, 对整幅图像重叠分块后进行像素域多假设预测^[12] 以减轻块效应. 图 1 中虚线框内容为本文所提算法的核心部分, 符号 \oplus 为取并集操作, 即完整图像为运动块与稳定块的集合.

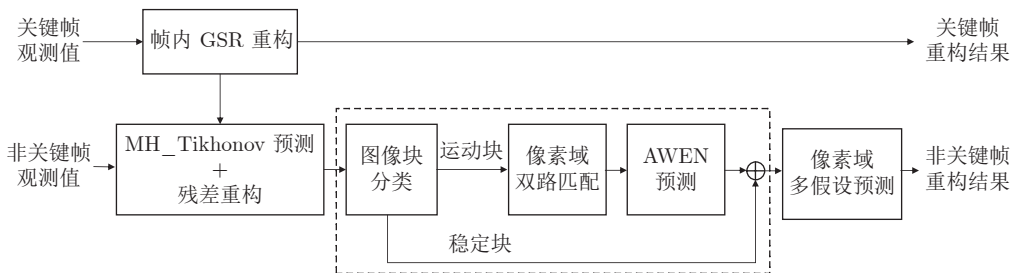


图 1 MH-LE 算法框架

Fig. 1 The framework of the proposed MH-LE algorithm

2.2 图像块状态分类

在 MH_Tikhonov 初始重构的基础上, 本文通过计算相邻帧帧差的方式来简单确定图像块状态类别, 将计算力集中到运动图像块的增强重构上. 判断图像块状态类别的步骤分为两步: 1) 计算帧差并进行二值化处理, 得到变化显著点分布情况; 2) 统计各图像块内变化显著点比例, 若高于给定阈值则判别为运动块, 反之为稳定块. 以 hall 序列中的一个视频帧为例, 该过程中间结果图如图 2 所示.

由于非关键帧的重构有先后顺序, 帧差的计算方式应根据帧重构顺序来进行设计. 本文算法采用文献 [19] 提出的参考帧选择方案, 从 GOP 两端向中间帧靠拢的重构顺序来对非关键帧进行重构. 因此, 在 GOP 的前半部分和后半部分, 帧差计算方式为当前帧与相邻已重构帧的差值, 而在 GOP 的中间帧因为同时有前后两个相邻帧作为参考帧, 需要同时计算当前帧与两相邻参考帧的差值. 然后, 对帧差图像进行阈值大小为 τ_1 的二值化处理, 得到变化显著点分布图, GOP 中间帧最终的变化显著点分布图为两幅二值化帧差图像的交集. 最后, 通过各图像块内变化显著点的比例即可判断该图像块是否为运动图像块, 当固定大小的图像块中变化显著点比例超过设定的数量比例阈值 τ_2 时, 则判定该图像块为运动块, 否则为稳定块.

2.3 像素域双路匹配策略

初始重构时运动图像块找到合适匹配块的难度远大于稳定图像块, 造成运动图像块的初始重构质

量较差, 再次进行匹配时准确度较低. 本文提出像素域双路匹配策略提高假设集质量, 首先引入双边滤波^[20] 滤除部分噪声, 将滤波结果作为新的标准块参与到匹配过程, 匹配过程中更关注图像块的主要特征; 其次提出细致化筛选方式构建高质量假设集, 参与后期的 AWEN 多假设预测. 像素域双路匹配策略流程见图 3.

在像素域双路匹配策略中, 运动块的假设块匹配工作分为两路, 一路直接在初始重构结果上进行匹配; 另一路在匹配前先对当前图像块和参考帧执行滤波操作, 利用滤波后的信息进行匹配. 在这两路匹配流程分别建立起各自的假设集后将假设集进行合并, 得到最终的假设集, 参与到最后进行的 AWEN 预测. 当前初始重构帧和前后关键帧、已重构相邻非关键帧共同组成了完整的参考帧集合, 作为第一条匹配路线的参考帧集合. 但由于关键帧重构质量已经较高, 滤波操作反而会抹去原有的真实信息, 因此第二条匹配路线不将关键帧作为参考帧.

接下来具体介绍图 3 中执行双边滤波操作的匹配路线.

步骤 1. 对当前运动块和参考帧进行双边滤波

当前帧中运动图像块初始重构质量通常较差. 由于重构时只以在观测域上逼近真实观测值作为优化目标, 而没有引入足够的图像先验信息, 重构结果往往存在明显的块效应, 且图像块内像素值的连续性较差. 这样的图像块如果不经处理就直接在像素域寻找匹配块, 就很容易找到那些同样充满噪点及块效应, 且像素间灰度值不连贯的图像块.

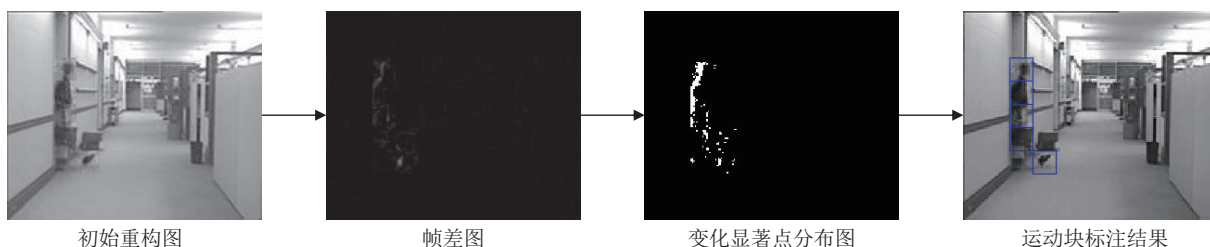


图 2 运动块判断过程中间结果图

Fig.2 The intermediate result of motion block judgment process

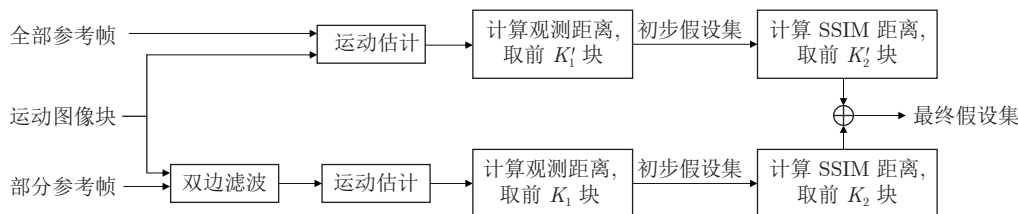


图 3 像素域双路匹配策略

Fig.3 Dual channel matching strategy in pixel domain

因此,为了利用像素间相关性且在不至于过度平滑的情况下缓解情况严重的噪点问题,采用经典的双边滤波对当前图像和参考帧进行处理.双边滤波同时考虑了像素点间的欧氏距离和灰度差异,将周围像素灰度值的加权平均作为当前像素灰度值.周边像素点对应权值的计算式为

$$w(i, j; m, n) = \exp \left(-\frac{(i-m)^2 + (j-n)^2}{2\sigma_d^2} - \frac{\|I(i, j) - I(m, n)\|_2^2}{2\sigma_r^2} \right) \quad (6)$$

其中, (i, j) 为中心像素坐标, (m, n) 为当前周边像素的坐标, $I(i, j)$, $I(m, n)$ 分别为二者对应的灰度值, σ_d 和 σ_r 为平滑参数.

双边滤波有效地利用了像素相关性,使得滤波结果更符合自然图像常见的内部平滑特征,同时减少了干扰信息,有效地突出了图像块的主要特性.在图像质量较差的情况下,匹配时将图像块的主要初级特征而非细节等较难准确重构的特征作为关注点,能够提高匹配准确度.

步骤 2. 快速运动估计寻找候选假设块

由于运动图像块找到匹配块的难度较大,因此通过快速运动估计方法来扩大搜索范围.该步骤采用已经发展成熟的四步快速搜索算法在滤波后的参考帧中找到最佳匹配点,然后把搜索窗内的所有图像块作为候选假设块.

步骤 3. 筛选假设块

步骤 2 中得到的候选假设块数量庞大,且质量参差不齐,若全部参与加权计算,不仅会极大增加运算量,还会降低预测精度.因此,需要对候选假设块进行筛选,只保留匹配程度较高的假设块.筛选过程在观测域和像素域共同进行,由于观测过程为多对一的映射过程,即使重构图像块的观测结果接近于真实观测值,也无法确保重构图像块接近真实图像块,因此需要像素域信息作为辅助信息来进一步判断匹配程度.筛选假设块时先在观测域上判断匹配程度,再在像素域上匹配,遵循这样的匹配顺序主要有 3 个原因. 1) 观测值相近是图像块内容相近的前提,观测结果相差太远的图像块基本不可能匹配; 2) 当前块的真实观测值是目前已有的唯一的真实信息,而像素域信息为重构信息,匹配可靠性严重依赖于重构结果的质量,依此找出来的假设块可能出现与真实观测值相差较大的情况; 3) SSIM 需对图像块进行亮度、对比度以及结构上的比较,计算复杂度较高,不适合处理大量数据.因此,筛选假设块时先在观测域上进行粗筛选,再在像素域上

进行细筛选.

筛选过程的具体操作为:在当前已滤波的处理假设集中选出前 K_1 个观测结果和当前块真实观测值最接近的假设块,然后在这些假设块中选出和已滤波的当前块 SSIM 值最大的前 K_2 个假设块作为当前重构路线的假设集.

另一条重构路线采用类似的假设集构造方法,其假设集与当前块均为未经滤波的原始结果.最终,AWEN 预测方法中使用的假设集为这两条路线构造的假设集之和 \mathbf{H}_N .

2.4 基于观测距离和 SSIM 距离的权值调整方法

得到全新的假设集后,利用 AWEN 算法对当前块进行预测重构,即

$$\hat{\mathbf{w}} = (1 + \lambda_2) \arg \min_{\mathbf{w}} \|\mathbf{y} - \Phi \mathbf{H}_N \mathbf{w}\|_2^2 + \lambda_1 \|\Gamma_N \mathbf{w}\|_1 + \lambda_2 \|\Gamma_N \mathbf{w}\|_2^2 \quad (7)$$

大部分的多假设预测算法采用距离惩罚方式来设计 Γ_N ,即将各假设块与当前块的观测域距离作为权值调整矩阵 Γ_N 对角线上的元素.然而,由于观测值是原始信号经低维映射后所得的结果,观测值与原始信号为一对多的关系,观测结果相近的信号在映射前并不一定相近.因此,对于那些假设集质量较差的图像块,在对其假设块进行权值分配的过程中,如果只利用到观测域信息,可能会出现噪声块被赋予大权重的意外情况.由于在预测前已经有了初始重构结果,可一并结合像素域信息来设计权值调整矩阵,减轻度量失真程度.

同时考虑图像块的观测结果与图像块本身的结构特性,本文提出基于观测距离和 SSIM 距离的权值调整矩阵设计方法,即

$$\Gamma_N = \text{diag}\{\|\mathbf{y} - \Phi \mathbf{h}_1\| \times d_1, \|\mathbf{y} - \Phi \mathbf{h}_2\| \times d_2, \dots, \|\mathbf{y} - \Phi \mathbf{h}_K\| \times d_K\} \quad (8)$$

其中, \mathbf{h}_i 表示假设块总个数为 K 的假设集中的第 i 个假设块, d_i 表示 \mathbf{h}_i 与当前块的结构相似度距离,定义为

$$d_i = 1 - SSIM^2(\mathbf{h}_i, F(\mathbf{x}_{\text{cur}}; \mathbf{h}_i)) \quad (9)$$

其中, $SSIM(\cdot)$ 为结构相似度度量函数,由于 $SSIM$ 值变化幅度较小,对 $SSIM$ 结果引入平方操作以拉大差距. \mathbf{x}_{cur} 为当前块, $F(\cdot)$ 为条件选择函数,其定义为

$$F(\mathbf{x}_{\text{cur}}; \mathbf{h}_i) = \begin{cases} BF(\mathbf{x}_{\text{cur}}), & \text{若 } \mathbf{h}_i \text{ 已过滤} \\ \mathbf{x}_{\text{cur}}, & \text{否则} \end{cases} \quad (10)$$

其中, $BF(\cdot)$ 为双边滤波操作.在计算结构相似度时,条件选择函数 F 的作用是确保参与计算的当前

块的滤波状态与假设块保持一致. F 的输出取决于 h_i , 若 h_i 为经过滤波的假设块, 则函数输出滤波后的当前块; 反之, 则仍输出当前块.

滤波操作的目的是在像素域上发掘图像块间的潜在相似性, 导致在强化图像内部主要特征的同时难免会对原图内容造成一定程度的更改, 该情况下如果和滤波前图像块进行比较, 则无论是观测误差还是像素域误差均会明显增加, 滤波后的图像块被认为和当前图像块匹配程度较低, 导致被赋予更大权重的块基本只能是那些未经滤波的假设块. 因此, 在进行权值调整时, 参与计算的当前块的状态与假设块的状态应保持一致, 才能保证匹配计算的公平性, 使更多的优质假设块能够参与到多假设重构中.

3 仿真结果及分析

为验证本文所提 MH-LE 算法的有效性, 在本节将对 MH-LE 进行仿真实验并与最近几年提出的性能较好的视频压缩感知重构算法进行比较分析. 参与比较的算法包括基于多假设预测的两种重构算法 2sMHR^[12] 和 PBCR-DCVS^[13], 以及基于组稀疏的两种重构算法 RRS^[14] 和 SSIM-InterF-GSR^[15]. 本节实验使用的视频序列有: CIF 格式 (352 × 288) 的 foreman, coastguard, soccer, mother-daughter, QCIF 格式 (176 × 144) 的 foreman, hall, coastguard, suzie, salesman, soccer. 观测时, 视频帧分块大小为 16 × 16, 观测矩阵为高斯随机矩阵.

MH-LE 算法中各参数设置如下: 1) 图像块状态分类的判断中, 帧差二值化阈值 τ_1 的大小决定被判定为变化显著点的难易程度, 数量比例阈值 τ_2 的大小决定被判定为变化块的图像块个数. 结合视频监控中判断运动物体时常见的帧差阈值设置情况进行大量实验发现, τ_1 取值在 [5, 20] 的区间内即可满足大部分序列的运动情况判断需求. 而在进行数量比例阈值的选取实验时发现, 当显著点比例约在 20% ~ 40% 时, 对图像块状态的判断结果基本合理. 另外, 考虑到变化区域重构复杂度较高, 变化区域图像块个数不应过多, 即 τ_2 不应太小. 因此, 最终选取 $\tau_1 = 5$, $\tau_2 = 35\%$. 2) 运动估计中, 四步快速搜索算法搜索步长设置为 8, 8, 8, 2. 在双边滤波平滑参数的设置中, σ_d 控制空间距离权重, σ_r 控制图像灰度变化权重, σ_d 和 σ_r 的大小比例影响着对去噪和保边这两方面的重视程度, σ_d 和 σ_r 的比值越大, 去噪强度越大, 但为避免过度平滑, 需对该比值的取值上限进行限制. 考虑到所提算法引入双边滤波的目的是对图像降噪以强化图像基本特征, 因此对 σ_d 设置更大的数值. 经过对多个序列的多次实验,

选取整体最优值 $\sigma_d = 1$, $\sigma_r = 0.1$. 式 (7) 的计算由 spasm 软件包^[21] 中的 elasticnet 函数模块实现, 该模块需要设置 λ_2 和 $stop$ 参数. 设置 $\lambda_2 = 0.1$, $stop = -1000 \times \text{采样率}$, $stop$ 参数直接决定了 λ_1 的具体值.

3.1 与 PBCR-DCVS 结果对比及分析

本小节将 MH-LE 算法与目前文献中基于多假设预测的重构算法中重构峰值信噪比 (Peak signal for noise ratio, PSNR) 值最高的 PBCR-DCVS^[13] 进行性能对比, 由于 PBCR-DCVS 算法未开源, 为保证对比结果的公平性, PBCR-DCVS 实验结果均取自原文献, 本文实验条件与文献一致, 即实验序列为 CIF 格式的 foreman, coastguard, soccer, mother-daughter, 关键帧采样率为 0.7, GOP 大小为 8, 取序列前 88 帧. 各非关键帧采样率下 MH-LE 与 PBCR-DCVS 重构 PSNR 对比如图 4 所示.

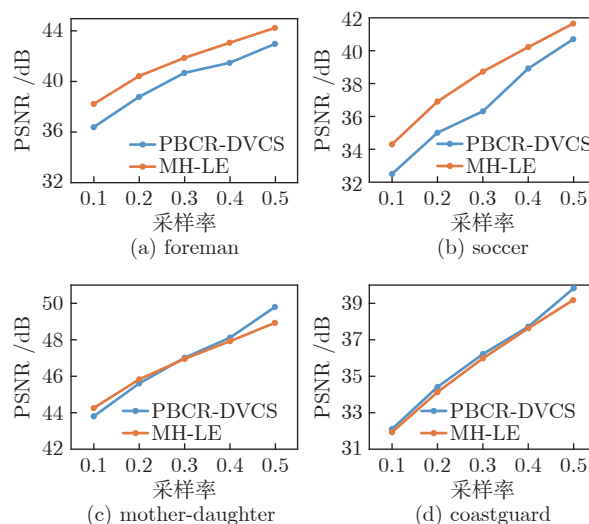


图 4 MH-LE 与 PBCR-DCVS 重构 PSNR 对比

Fig. 4 Reconstruction PSNR performance comparison between MH-LE and PBCR-DCVS

从图 4 可见, 对于 foreman 和 soccer 序列, 本文所提 MH-LE 算法相比于 PBCR-DCVS 算法有明显的性能提升, 在各采样率下均有 1 ~ 2 dB 的提升. Foreman 和 soccer 序列中运动块的初始重构质量相对较差, MH-LE 算法对图像块进行划分后, 利用像素域信息为运动块构造更高质量的假设集, 并在假设块权值分配过程中结合 SSIM 评价标准, 有效提升了运动块的重构质量. 对于 mother-daughter 序列, MH-LE 算法与 PBCR-DCVS 算法分别在低采样率和高采样率下显示出其各自的优越性. 这是因为 mother-daughter 序列内容极其稳定, 在采样率达到一定大小时即可获得较高的初始重构质量. 低采样率时图像块初始重构质量一般, MH-LE

算法中的滤波操作能带来图像块质量的改善;而在高采样率时细节内容成为相似块匹配程度的重要依据,在该情况下滤波操作会造成图像细节的丢失. Coastguard 序列包含大量如水面波纹、植被等细节较多的内容,非关键帧初始重构质量较差. PBCR-DCVS 更关注对关键帧信息的利用,通过对关键帧进行二次重构来提供更多的高质量假设块;而 MH-LE 算法中对非关键帧的滤波操作容易造成细节内容的丢失,因此对 coastguard 序列中运动块重构质量的提升幅度相对有限.

MH-LE 算法中预处理操作带来的额外开销主要来自于运动块的相似块匹配过程,运动块的参考帧集合增加了原参考帧的滤波结果,虽然增加了匹配次数,但也提高了假设集质量. MH-LE 和 PBCR-DCVS 均采用了基于 elasticnet 的预测算法,该算法需通过迭代进行求解,上述两个算法的时间复杂度主要来源于预测算法的求解过程. 由于 MH-LE 算法只对运动图像块采用基于 elasticnet 的预测算法,因此大大减少了需要进行迭代计算的图像块个数,和 PBCR-DCVS 相比具有更低的算法复杂度.

3.2 算法重构性能对比与复杂度分析

3.2.1 重构性能对比分析

本节将 MH-LE 算法与 RRS^[14]、2s-MHR^[12]、

SSIM-InterF-GSR^[15] 进行性能对比. 实验序列为 QCIF 格式的 foreman, hall, coastguard, suzie, salesman, soccer, 关键帧采样率为 0.7, GOP 大小为 16, 取序列的前 96 帧. 各非关键帧采样率下 MH-LE 与各算法重构 PSNR 对比如表 1 所示.

从表 1 仿真结果可见, MH-LE 算法在大部分采样率下和其他重构算法相比重构性能均有较明显的提升. 分别和 RRS、2sMHR、SSIM-InterF-GSR 比较所有采样率下 PSNR 的平均值, MH-LE 算法在 foreman 序列上分别提升了 3.49, 2.49, 1.21 dB, 在 hall 序列上分别提升了 4.73, 3.03, 0.81 dB, 在 coastguard 序列上分别提升了 2.89, 1.78, 1.16 dB, 在 suzie 序列上分别提升了 3.04, 0.51, 1.16 dB, 在 salesman 序列上分别提升了 2.92, 1.97, 2.16 dB. 在此类物体运动处于正常运动速度的视频序列下, MH-LE 算法能够较好地地区分出稳定块和运动块,稳定块用经过高度筛选的假设块直接进行多假设预测即可得到高质量的预测结果,避免过度重构;而运动块大概率上能够在相邻的视频帧内找到相关内容, MH-LE 算法加强对相邻帧像素域信息的利用程度,首先引入像素域双路匹配策略来丰富假设集资源,其次在预测过程中结合 SSIM 准则来提高预测精度,这两个措施大幅度改善了重构质量. 在重构 suzie 序列时,基于多假设预测的 2sMHR 和 MH-

表 1 各采样率下算法重构 PSNR 对比

Table 1 Reconstruction PSNR performance comparison of various algorithms at different sampling rates

采样率	重构算法	视频序列					
		foreman	hall	coastguard	suzie	salesman	soccer
0.1	RRS	31.10	26.73	25.96	34.39	31.18	29.31
	2sMHR	33.25	32.88	28.91	37.82	35.32	29.32
	SSIM-InterF-GSR	34.63	33.87	29.09	36.68	34.39	29.51
	MH-LE	35.60	35.57	30.67	38.18	37.09	30.32
0.2	RRS	35.78	33.45	30.32	37.94	37.15	32.29
	2sMHR	36.17	34.76	30.82	40.06	36.56	32.42
	SSIM-InterF-GSR	37.71	37.34	31.70	39.55	36.52	34.60
	MH-LE	38.79	37.82	32.61	40.61	38.59	33.75
0.3	RRS	37.91	38.56	32.57	39.60	38.07	34.05
	2sMHR	38.38	36.21	32.44	41.63	37.38	34.81
	SSIM-InterF-GSR	39.57	39.29	33.25	41.36	37.80	37.17
	MH-LE	40.87	39.54	34.26	42.28	39.50	36.16
平均值	RRS	34.93	32.91	29.61	37.31	35.47	31.88
	2sMHR	35.93	34.61	30.72	39.84	36.42	32.18
	SSIM-InterF-GSR	37.21	36.83	31.34	39.19	36.23	33.76
	MH-LE	38.42	37.64	32.50	40.35	38.39	33.41

LE 算法的重构性能均高于 SSIM-InterF-GSR 算法. 由于 *suzie* 序列运动极其缓慢, 当前块能够直接从参考帧中找到高度相关的匹配块, 多假设预测算法显示出明显的优越性. 对于 *soccer* 序列, MH-RE 算法在 0.1 采样率下保持了算法的性能优越性, 但在 0.2 采样率以上时出现不及 SSIM-InterF-GSR 算法的情况. 上述情况的出现和序列运动剧烈程度有关. 基于多假设预测的算法高度依赖假设块的匹配程度, 在处理快速运动序列时, 基于多假设预测的算法能利用到的帧间相关信息不足, 在参考帧中难以找到优质匹配块, 因此预测结果相对较差; 而 SSIM-InterF-GSR 直接利用帧内自相似性, 对相似块组多次进行低秩约束重构, 通过不断迭代提高重构质量.

对于 *foreman* 序列第 8 帧在 0.1 采样率下的重构结果, 各算法的重构视觉效果如图 5 所示, 并给出了每个重构结果的 PSNR 值. 在采样率较低的情况下, 对原图像的重构难度较大, RRS 重构结果中出现了明显的噪点, 而 2sMHR, SSIM-InterF-GSR 和 MH-LE 算法基本重构出了图像的主要信息. 2sMHR 算法重构, 在人物的耳朵、衣领、下巴处仍出现较明显的块效应; 在 SSIM-InterF-GSR 算法的重构图像中虽然没有出现块效应, 整体视觉效果十分柔和, 但却出现过度平滑现象, 下巴褶皱、耳窝、屋檐下阴

影等细节内容被大量抹除. 相比之下, MH-LE 算法既减轻了块效应, 又避免了过度平滑, 更忠实于原图像.

3.2.2 算法复杂度对比分析

本节将所提 MH-LE 算法与上述提供算法代码的 RRS^[14], 2s-MHR^[12], SSIM-InterF-GSR^[15] 算法进行时间复杂度比较, 并根据比较结果进行分析. 本节所有实验均在搭载 Intel Core i5 3.30 GHz 处理器的 PC 端完成, 仿真软件为 Matlab R2016a, 操作系统为 Windows 7. 实验中采用 QCIF 格式的视频序列: *suzie*, *hall*, *foreman*, *soccer*, 以上序列的运动剧烈程度依次递增. GOP 长度为 16, 关键帧采样率为 0.7. 实验结果如表 2 所示, 表中数据为各采样率下视频序列每个非关键帧的平均重构时间.

由表 2 可知, 算法运行速度从快到慢依次为: 2sMHR, MH-LE, RRS, SSIM-InterF-GSR. 2sMHR 和 MH-LE 算法的时间复杂度明显低于 SSIM-InterF-GSR 和 RRS 算法, 以 *hall* 序列为例, SSIM-InterF-GSR 算法的运行时间约为 2sMHR 算法的 20 倍, 约为 MH-LE 算法的 10 倍. 多假设预测算法直接利用假设集对当前帧进行预测, 涉及到的处理过程较简单, 所需时间较少. 其中 2sMHR 算法在预测时因为没有涉及到迭代过程, 所以重构速度最快. 而基于组稀疏的重构算法则因为涉及到矩阵分解问



图 5 *foreman* 第 8 帧重构视觉效果图

Fig. 5 Visual quality comparison of various algorithms recovery on the 8th frame of *foreman*

表 2 不同算法下每帧平均所需重构时间 (s)
Table 2 Running time comparison with various algorithms for reconstructing a video frame at different sampling rates (s)

采样率	重构算法	视频序列			
		suzie	hall	foreman	soccer
0.1	RRS	103.8	110.3	115.2	104.3
	SSIM-InterF-GSR	168.9	177.5	167.2	184.2
	2sMHR	6.9	7.1	7.2	7.3
	MH-LE	13.5	15.2	23.3	60.7
0.2	RRS	98.5	125.0	106.4	99.6
	SSIM-InterF-GSR	174.3	188.2	166.2	182.5
	2sMHR	7.1	7.3	7.4	7.7
	MH-LE	14.8	15.5	25.1	67.2
0.3	RRS	99.0	123.7	103.5	100.7
	SSIM-InterF-GSR	171.4	173.4	169.3	176.4
	2sMHR	7.3	7.4	7.4	7.6
	MH-LE	15.1	14.6	24.8	86.2

题, 每次迭代所需时间较长; 同时由于观测域约束优化过程与像素域优化过程是分开进行的, 因此需要一定的迭代次数来保证算法已达到收敛状态. 由表中数据可知, MH-LE 算法重构 suzie, hall, foreman, soccer 序列所需时间依次递增. MH-LE 算法的运行时间与序列内容的运动状态密切相关, 取决于视频帧内容中运动块的比例. 重构运动块时采用了 AWEN 预测方法, 该算法求解时需要迭代操作, 所需预测时间稍长; 另外, 对运动块进行重构前增加了双边滤波操作, 且对假设集进行了扩充, 增加了每次预测所需时间. 所以, 运动越快的序列中运动块的数量越多, 所需重构时间也越长.

4 结束语

在多假设预测重构中, 假设集与当前待重构图像块的匹配程度对图像块重构质量影响较大, 低质量的假设集合极大地增加了图像块的重构难度. 针对此问题, 本文提出了一种对部分图像块进行增强重构的多假设预测方案 (MH-LE), 通过挖掘像素域上的有用信息来进一步提高图像块的重构质量. 该方案通过帧间像素变化情况将图像块划分为稳定块和运动块, 结合图像块自身结构特点对图像块采取不同的重构策略. 所提方案针对运动块提出了像素域双路匹配策略, 建立更高质量的假设集; 并将结构相似度评价标准引入假设块权值分配过程, 进一步提高预测精度. 仿真结果表明, 所提方案和其他

基于多假设预测的重构算法相比具有明显的质量提升; 和基于组稀疏的重构算法相比, 在具有更低的算法复杂度这一优势的同时仍在大部分采样率下具有更高的重构质量.

References

- 1 Donoho D L. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, **52**(4): 1289–1306
- 2 Gan L. Block compressed sensing of natural images. In: Proceedings of the 2017 International Conference on Digital Signal Processing. Cardiff, UK: IEEE, 2007. 403–406
- 3 Mun S K, Fowler J E. Block compressed sensing of images using directional transforms. In: Proceedings of the 2019 International Conference on Image Processing. Cairo, Egypt: IEEE, 2009. 3021–3024
- 4 Zhang J, Zhao D B, Gao W. Group-based sparse representation for image restoration. *IEEE Transactions on Image Processing*, 2014, **23**(8): 3336–3351
- 5 Zhang J, Ghanem B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. In: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 1828–1837
- 6 Shi W Z, Jiang F, Zhang S P, Zhao D B. Deep networks for compressed image sensing. In: Proceedings of the 2017 IEEE International Conference on Multimedia and Expo. Hong Kong, China: IEEE, 2017. 877–882
- 7 Tramel E W, Fowler J E. Video compressed sensing with multi-hypothesis. In: Proceedings of the 2011 Data Compression Conference. Snowbird, USA: IEEE, 2011. 193–202
- 8 Azghani M, Karimi M, Marvasti F. Multi-hypothesis compressed video sensing technique. *IEEE Transactions on Circuits & Systems for Video Technology*, 2016, **26**(4): 627–635
- 9 Chen J, Wang N, Xue F, Gao Y N. Distributed compressed video sensing based on the optimization of hypothesis set update technique. *Multimedia Tools and Applications*, 2016, **74**(14): 1–20
- 10 Chen J, Chen Y Z, Qin D, Kuo Y H. An elastic net-based hybrid hypothesis method for compressed video sensing. *Multimedia Tools and Applications*, 2015, **74**(6): 2085–2108
- 11 Kuo Y H, Wu K, Chen J. A scheme for distributed compressed video sensing based on hypothesis set optimization techniques. *Multidimensional Systems and Signal Processing*, 2017, **28**(1): 129–148
- 12 Ou Wei-Feng, Yang Chun-Ling, Dai Chao. A two-stage multi-hypothesis reconstruction and two implementation schemes for compressed video sensing. *Journal of Electronics and Information Technology*, 2017, **39**(7): 1688–1696
(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. *电子与信息学报*, 2017, **39**(7): 1688–1696)
- 13 Zheng S, Chen J, Kuo Y H. An improved distributed compressed video sensing scheme in reconstruction algorithm. *Multimedia Tools and Applications*, 2018, **77**(7): 8711–8728
- 14 Zhao C, Ma S W, Zhang J, Xiong R Q, Guo W. Video compressive sensing reconstruction via reweighted residual sparsity. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, **27**(6): 1182–1195
- 15 He Zhi-Jie, Yang Chun-Ling, Tang Rui-Dong. Research on structural similarity based inter-frame group sparse representation for compressed video sensing. *Journal of Electronics*, 2018, **46**(3): 544–533
(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. *电子学报*, 2018, **46**(3): 544–533)
- 16 Baraniuk R G, Goldstein T, Sankaranarayanan A C, Studer C, Veeraraghavan A, Wakin M. Compressive video sensing: algorithms, architectures, and applications. *IEEE Signal Processing*

Magazine, 2017, **34**(1): 52–66

- 17 Xu K, Ren F B. CSVideoNet: A real-time end-to-end learning framework for high-frame-rate video compressive sensing. In: Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe, USA: IEEE, 2018. 1680–1688
- 18 Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, 2005, **67**(2): 301–320
- 19 Yang Chun-Ling, Ou Wei-Feng. Multi-reference frames-based optimal multi-hypothesis prediction algorithm for compressed video sensing. *Journal of South China University of Technology (Natural Science Edition)*, 2016, **44**(1): 1–8
(杨春玲, 欧伟枫. CVS中基于多参考帧的最优多假设预测算法. 华南理工大学学报(自然科学版), 2016, **44**(1): 1–8)
- 20 Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of the 6th International Conference on Computer Vision. Bombay, India: IEEE, 1998. 839–846
- 21 Sjöstrand K, Ersbøll B. SpaSM. A MATLAB toolbox for sparse statistical modeling [Online], available: <http://www2.imm.dtu.dk/projects/spasm/>, October 30, 2019



汤瑞东 华南理工大学电子与信息学院硕士研究生. 主要研究方向为视频压缩感知.

E-mail: eerdtang@scut.edu.cn

(TANG Rui-Dong Master student at the School of Electronic and Information Engineering, South China

University of Technology. His main research interest is video compressed sensing.)



杨春玲 华南理工大学电信学院教授. 主要研究方向为图像/视频压缩编码, 图像质量评价. 本文通信作者.

E-mail: eeclyang@scut.edu.cn

(YANG Chun-Ling Professor at the School of Electronic and Information Engineering, South China University of Technology. Her research interest covers image/video compression coding and image quality evaluation. Corresponding author of this paper.)



禩韵怡 华南理工大学电子与信息学院硕士研究生. 主要研究方向为视频压缩感知.

E-mail: eeyunyixuan2014@scut.edu.cn

(XUAN Yun-Yi Master student at the School of Electronic and Information Engineering, South China University of Technology. Her main research interest is video compressed sensing.)