

DOI:10.11918/j.issn.0367-6234.201902009

基于动态多模式匹配的视频压缩感知三步重构

周 健, 刘 浩

(东华大学 信息科学与技术学院, 上海 201620)

摘 要: 相较于奈奎斯特-香农定理所要求的采样数据量, 压缩感知理论表明采用较少的测量值就可以实现高维信号的重构, 因此压缩感知在视频信号传感中具有很大的潜力. 现有的视频压缩感知重构算法是利用多假设预测来得到残差模型, 大量文献采用基于最小均方误差的方法挑选多假设匹配块, 由此对视频信号进行重构, 然而没有考虑最大化重构视频的整体结构相似性, 在图像重构的整体质量效果上存在较大改进空间, 并且挑选匹配块的模式没有采用自适应的选择机制, 挑选匹配块的方式较为单一. 通过增加一定的复杂度, 本文提出了一种基于动态多模式匹配的视频压缩感知三步重构算法, 该算法主要包括三大步骤: 第一步, 对压缩感知的每一视频帧进行独立的重构; 第二步, 从参考帧中动态地挑选匹配块进行非关键帧的重构; 第三步, 基于整体结构相似性对重构帧进行最终挑选, 完成多帧重构. 实验结果表明, 所提算法能够在重建端有效提高多假设过程的预测精度, 与当前最优的视频重构算法相比, 进一步提升了重构质量.

关键词: 视频重构; 压缩感知; 多假设预测; 整体结构相似性

中图分类号: TN919

文献标志码: A

文章编号: 0367-6234(2019)11-0167-07

Three-phase video compressive sensing reconstruction via dynamic multi-pattern matching

ZHOU Jian, LIU Hao

(College of Information Science & Technology, Donghua University, Shanghai 201620, China)

Abstract: Compressive sensing theory indicates that high-dimensional signal reconstruction can be obtained from far fewer measurements than those required by the Nyquist-Shannon sampling theorem, and compressive sensing has a great potential in video signal sensing. The existing reconstruction algorithms utilize the multi-hypothesis prediction to derive the residual model. A large number of studies adopt the method based on the least mean square error to select multiple hypothesis matching patches and reconstruct the video, while the maximization of the structural similarity is not considered in these algorithms, and there is much room for improvement in the overall quality effect of image reconstruction. Therefore, a three-phase video compressive sensing reconstruction algorithm is proposed in this paper on the basis of dynamic multi-pattern matching, in which the first phase reconstructs each frame independently, the second phase dynamically selects the hypothesis patches from the reference frames and reconstructs the frames, and the final reconstruction result is obtained in the third phase with the best structural similarity. Experimental results demonstrate that compared with the state-of-the-art algorithm, the proposed algorithm could obtain better prediction accuracy and reconstruction quality for video compressive sensing.

Keywords: video reconstruction; compressive sensing; multi-hypothesis prediction; structural similarity

传统的视频通信系统根据奈奎斯特-香农采样定理对每帧采集一组大量的信号样本, 采样率要求不小于视频信号带宽的两倍, 随后进行高复杂度的压缩编码, 以消除像素之间存在的冗余. 压缩感知 (CS) 技术的兴起提供了一种新型的信号采集与重构方法, 根据 CS 理论, 系统只需通过将信号投射到随机基上以获得少量的测量值, 在某一变换域中具有稀疏或近稀疏表示的信号就可以通过这些测量值

得到高可靠性的重构^[1-3].

在初期的压缩感知文献中, 测量端对视频帧采用直接观测, 但由于视频帧的信息量较大, 在进行重构时对计算资源有着非常高的要求. 为减少计算复杂度, 文献[4]提出了分块压缩感知 (BCS) 框架及其重构算法, 即先对视频帧进行分块测量, 之后再利用投影 Landweber 迭代重构, 并在迭代过程中执行维纳滤波以消除重构中的块效应. BCS 框架稳定了视频压缩感知的重构质量, 大大降低了重构计算的复杂度, 被认为是图像/视频压缩感知的典型框架, 在图像/视频压缩感知研究中广泛使用. 基于 BCS 框架, 文献[5]提出了在离散双树小波变换

收稿日期: 2019-02-03

基金项目: 上海市自然科学基金 (18ZR1400300)

作者简介: 周 健 (1996—), 男, 硕士研究生

通信作者: 刘 浩 (1977—), liuhao@dh. edu. cn

(DDWT)域进行图像重构的算法,由于这类重构算法没有借鉴传统图像编码中对块间相关性的充分利用,其重构效果并不理想。

对于视频压缩感知,研究人员已经提出在重建端通过块匹配进行运动估计和补偿^[6-7],即将相邻帧作为参考帧,从参考帧中选取匹配块作为当前块的一个预测。文献[8]提出通过稀疏化当前帧与其运动补偿帧之间的残差进行重构^[9-10],这是因为残差通常比原始信号具有更大的稀疏性。多假设预测是目前视频压缩感知重构的主流技术^[11]。在变换域多假设预测的研究中,文献[12]对文献[5]进行了改进,采用多假设(MH)预测的线性组合来估计变换域中的运动补偿帧,在重构质量与复杂度两方面取得了较好的性能折中。对于多个匹配块的权值分配问题,文献[13]提出通过增加 Tikhonov 正则化项来求解最优权值;文献[14]则提出在 l_2 正则化项中加入权重调整函数;文献[15]通过联合 lasso 模型中的 l_1 范数正则化与回归模型中的 l_2 正则化,求解最优权值。

在匹配块的选取方面,文献[16]提出多参考帧假设集合的优化算法,增加匹配块数目以扩大假设选择范围;文献[17]则提出基于菱形快速搜索的双匹配算法,即利用菱形快速搜索方式确定当前解码块的前景/背景的运动矢量,获得两个最佳搜索窗,从中搜索匹配块。这些挑选匹配块的方法均从最小化每个匹配块与原始块均方误差的角度进行考虑,没有对挑选的匹配块进行整体结构质量的衡量,同时这些挑选匹配块的方法存在复杂度过高的问题。

文献[18]在文献[12]的基础上进一步利用了视频序列的时空相关性,提出了一种基于判别加权残差稀疏性(RSS)的视频重构算法,该 RSS 算法首先对每帧图像进行高质量的独立恢复,随后执行帧间的运动估计,对非关键帧进行预测。RSS 算法建模了信号间残差的稀疏性而不是信号内的稀疏性,对不同残差系数的稀疏性进行判别加权,并对权重进行迭代更新。RSS 算法在现有视频重构算法中具有相对突出的重构性能,本文试图在增加一定复杂度的情形下进一步提升其重构质量。此外,现有重构算法均只是试图去最小化块间的均方误差,没有考虑视频整体结构相似性(SSIM)的最大化。在某些情况下,均方误差的最小化并不能得到满意的重构效果。因此,本文提出基于 SSIM 和动态多模式匹配的視頻压缩感知三步重构算法,以期进一步提升视频压缩感知的重构质量。

1 视频压缩感知

1.1 多假设预测

压缩感知理论认为,如果一个信号在某变换域中足够稀疏,那么该信号可以从少量的测量值中恢复。若 $\mathbf{x} \in \mathcal{R}^N$ 为 N 维信号, $\boldsymbol{\varphi} \in \mathcal{R}^{M \times N}$ 代表一个测量矩阵,则 BCS 框架的测量过程为

$$\mathbf{y} = \boldsymbol{\varphi} \mathbf{x}. \quad (1)$$

式中: $\mathbf{y} \in \mathcal{R}^M$ 为得到的测量值。测量值的个数 M 通常比原始信号中的数据个数 N 小得多,因此该重构恢复问题是欠定的。采样率定义为 M/N 。一个视频序列表示为 $\{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^1, \dots\}$, 其中 $\mathbf{x}^1$ 为即将进行重构的当前帧信号, $\hat{\mathbf{x}}$ 表示相应的重构帧。

$\hat{\mathbf{x}}^p$ 为重构帧 $\hat{\mathbf{x}}$ 通过前一次迭代后获得的重构结果, p 为前一次重构。在式(1)的基础上,为有效消除潜在的块效应,常将 $\hat{\mathbf{x}}^p$ 拆分成 D 个重叠的块: $\hat{\mathbf{x}}_k^p \in \mathcal{R}^{S^2}$, $1 \leq k \leq D$, 其中 S^2 是重叠块的大小。上述的拆分过程为

$$\hat{\mathbf{x}}_k^p = \mathbf{E}_k \hat{\mathbf{x}}^p$$

式中 $\mathbf{E}_k \in \mathcal{R}^{S^2 \times N}$ 为矩阵算子,以交叉重叠块的方式从 $\hat{\mathbf{x}}^p$ 中获取重构块 $\hat{\mathbf{x}}_k^p$, 对于拆分获得的每个块,从指定的参考帧中挑选与该块 $\hat{\mathbf{x}}_k^p$ 具有高度相似性的匹配块。一般地,块与块之间的像素相似性可用均方误差(MSE)衡量,具体的表达式为

$$M_{se}(\hat{\mathbf{x}}_k^p, \hat{\mathbf{x}}_{ref,i}^p) = \frac{1}{S^2} \|\hat{\mathbf{x}}_k^p - \hat{\mathbf{x}}_{ref,i}^p\|_2^2. \quad (2)$$

式中: $\hat{\mathbf{x}}_{ref,i}^p$ 为在某个参考帧中的一个 S^2 大小的块,该块在以块 $\hat{\mathbf{x}}_k^p$ 为中心的 $L \times L$ 搜索窗内。基于式(2)挑选 C 个匹配块,也就是均方误差较小的 C 个匹配块,将它们定义为 $\mathbf{v}_{k_i}$, $1 \leq i \leq C$ 。上述的 C 个块,将视为重构块 $\hat{\mathbf{x}}_k^p$ 的多假设匹配块。

1.2 加权残差稀疏模型

当前块 $\mathbf{x}_k$ 的残差 $R(\mathbf{x}_k)$ 是块原始值与多个匹配块线性组合之间的差值,该差值表示为

$$R(\mathbf{x}_k) = \mathbf{x}_k - \sum_{1 \leq i \leq C} \alpha_{k_i} \mathbf{v}_{k_i}.$$

式中: α_{k_i} 为某匹配块的权重,将影响块匹配的预测精度。一般而言,匹配块如果和当前块相似性更高,那么该匹配块将赋予较大的权重;匹配块与当前块的相似性不高,赋予该匹配块一个较小的权重。因此,权重应与匹配块和当前块的均方误差值成反比关系。可以选取指数函数来计算不同匹配块的残差权重,由于原始块 $\mathbf{x}_k$ 无法获取,我们利用前一次的重

构结果 $\hat{\mathbf{x}}_k^p$ 进行如下计算

$$\alpha_{k_i} = \frac{\exp(-\text{MSE}(\hat{\mathbf{x}}_k^p, \mathbf{v}_{k_i})/h)}{\sum_{1 \leq j \leq C} \exp(-\text{MSE}(\hat{\mathbf{x}}_k^p, \mathbf{v}_{k_j})/h)}$$

式中: h 为一常数, $\text{MSE}(\cdot)$ 为计算均方误差的公式.变换域内信号残差的稀疏性已经被广泛运用.因此,可采用 DDWT 对残差进行稀疏化.残差 $R(\mathbf{x}_k)$ 的 DDWT 变换系数表示为

$$R_\Psi(\mathbf{x}_k) = \Psi^T R(\mathbf{x}_k) = \tilde{\mathbf{x}}_k - \sum_{1 \leq i \leq c} \alpha_{k_i} \tilde{\mathbf{v}}_{k_i}$$

式中: Ψ 为 DDWT 基, Ψ^T 为其转置,并且 $\tilde{\mathbf{x}}_k = \Psi^T \mathbf{x}_k$, $\tilde{\mathbf{v}}_{k_i} = \Psi^T \mathbf{v}_{k_i}$.不同系数为 0 的概率存在区别,所以可以对残差稀疏度进行加权.加权残差稀疏模型表示为.

$$F(\mathbf{x}_k) = \mathbf{W}_k R_\Psi(\mathbf{x}_k)$$

式中: $\mathbf{W}_k$ 为一个对角阵,对角线上元素为 $w_{k,1}, \dots, w_{k,s^2}$,其余位置为 0. $w_{k,1}, \dots, w_{k,s^2}$ 是残差系数的权重,每个值反映了对应系数为 0 的概率.利用对于 $F(\mathbf{x}_k)$ 的稀疏化,并得到一个加权的 l_1 范数问题

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \sum_{1 \leq k \leq D} ||F(\mathbf{E}_k \mathbf{x})||_1 \text{ s. t. } \mathbf{y} = \Phi \mathbf{x}.$$

式中: $\mathbf{x}_k = \mathbf{E}_k \mathbf{x}$.通过求解上式可获得当前帧的重构 $\hat{\mathbf{x}}$.当前重构获得的帧 $\hat{\mathbf{x}}$ 相较于前一次迭代获取的重构帧 $\hat{\mathbf{x}}^p$ 具有显著的质量改进.所提算法将不断迭代进行多假设预测的块匹配,构建加权残差稀疏模型进行重构,获得最终重构结果.

2 三步重构算法

为控制计算复杂度和内存需求,测量端逐帧地采用基于 BCS 框架的随机测量进行视频信号的采集^[5,19].视频序列由若干图像组构成,每个图像组由一个关键帧和后面的一些非关键帧组成.关键帧使用相对较高的采样率进行测量,而非关键帧使用较低的采样率进行测量.接下来,着重介绍本文提出的基于 SSIM 和动态多模式匹配的视频压缩感知三步重构算法,所提算法的整体流程如图 1 所示,主要包括三大步骤:首先对于序列中的每帧进行独立的重构;其次非关键帧利用第一步中重构的关键帧进行多假设预测的块匹配,在挑选匹配块总个数时进行动态分配;最后通过整体结构相似性的最大化,获取最终的重构结果.

步骤 1:视频序列中每帧均通过帧内模型独立恢复.这一步采用基于多假设预测的残差重构算法,匹配块从每帧自身进行挑选.一般而言,采样率越高的帧重构质量越好,采样率越低的帧重构质量越差.

受传统视频压缩编解码的启发,所提算法采用具有较高采样率的帧作为参考帧.

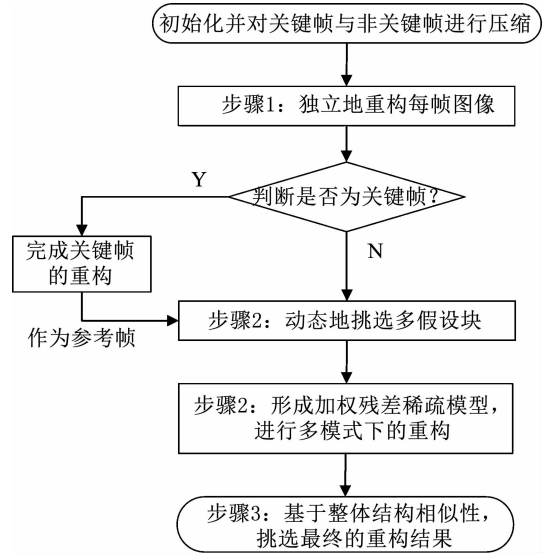


图 1 视频压缩感知三步重构算法的流程图

Fig. 1 Flowchart of three-phase video compressive sensing reconstruction

步骤 2:重构非关键帧.非关键帧可从相邻的两参考帧获得匹配块,由于参考帧在步骤 1 后得到了较好的重构结果,不需要进一步处理.如前所述,对于每个非关键帧,应该考虑从前面参考帧和后面参考帧选取匹配块的个数.所提算法采用动态多模式挑选匹配块.基于每一非关键帧与前后两帧的距离决定从两参考帧挑选匹配块的个数,而不是固定每一帧从某参考帧中获取块个数.也就是,如果当前的非关键帧距离前向参考帧较近,从前向参考帧中将挑选更多的匹配块而从后向参考帧中挑选的匹配块将相应减少.

为降低计算复杂度,基于帧距离挑选匹配块总数量的机制,可简化为:假定从两个参考帧中挑选的总匹配块个数为 M ,不失一般性, M 为偶数,可设置三组模式进行动态挑选:1、从前后两参考帧各挑选 $M/2$ 块;2、从距离较近的参考帧挑选 $M/2 + 1$ 块,从距离较远的参考帧挑选 $M/2 - 1$ 块;3、从距离较近的参考帧挑选 $M/2 + 2$ 块,从距离较远的参考帧挑选 $M/2 - 2$ 块.在获得上述匹配块之后,对于 1 2 3 每种模式,计算每块与对应的匹配块线性组合之间残差.然后,根据残差系数为零的概率对残差系数进行加权,得到一个加权的 l_1 最小化问题,根据 Split Bregman 迭代算法解决该问题,从而完成一帧的重构.例如,当 M 设置为 10 时,从前后两个参考帧中基于帧距离挑选匹配块的个数分配如表 1 所示.

表 1 基于帧距离的匹配块选择机制

Tab. 1 Matching patch selection mechanism according to frame distance

参考帧	分配 I	分配 II	分配 III	分配 IV	分配 V
前向	5	3	7	4	6
后向	5	7	3	6	4

文献[18]的 RRS 算法挑选匹配块的模式较为固定,所提算法根据视频帧之间的时域相关性,采用了多种模式动态地挑选匹配块,两者的对比如图 2 所示。

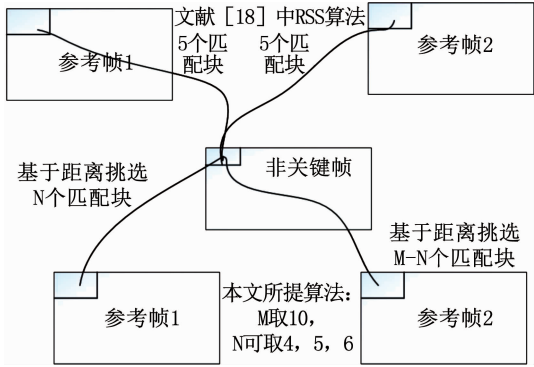


图 2 RRS 算法与本文算法在挑选匹配块的差别示例
Fig. 2 Difference between the RRS algorithm and the proposed algorithm in selecting matching patches

步骤 3:挑选最终的重构结果. 由于在某些情况下最小化均方误差并不会得到满意的效果,因此本步骤基于整体结构相似性从模式 1~3 中挑选最终的重构结果. 为降低复杂度,设置一阈值 $T(0 < T < 1)$,可依据重构质量需求自动调节所提算法的复杂度. 视频序列中的每个非关键帧按照模式 1~3 的顺序依次执行残差重构,并在每种模式重构完成后计算其与参考帧的 SSIM,若某种模式下对应的 SSIM 大于 T ,则重构完成,该模式对应的重构结果为最终结果;否则,接着执行其余的模式;若三种模式下得到的 SSIM 均小于阈值 T ,则从三种模式中挑选较大的 SSIM 对应的重构结果为最终重构结果。

从大量实验可以分析得出,阈值 T 的合理设置决定了最终的重构质量以及算法的复杂度:若阈值过小,则最终的重构质量效果不好;相反若阈值过大,重构时间较长,算法复杂度较高,甚至在耗费很长的计算时间之后还是达不到高阈值情况下所期望的重构效果. 一般情况下,阈值 T 在 0.9 左右时所提算法在重构效果和复杂度方面均能取得较好的性能,因此本文在实验结果部分设置阈值 $T=0.9$ 进行所提算法的性能评估. 在步骤 3 中挑选最终重构结果的详细算法流程图如图 3 所示。

相比于现有固定挑选匹配块的模式,所提算法挑选匹配块的模式更加灵活,实现了多模式的动态

挑选. 同时,在基于整体结构相似性的考虑下,所提算法可以从总体上衡量每个视频帧的质量,进一步消除了块效应,在增加一定复杂度的情况下有效地提升了视频压缩感知的重构质量。

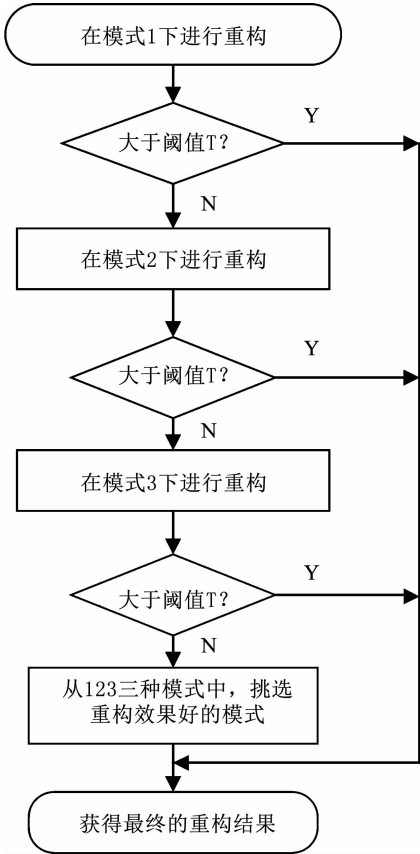


图 3 步骤 3 挑选最终重构结果流程
Fig. 3 Flow chart of selecting final results under phase 3

3 实验结果

提出算法的实验仿真与分析,以验证所提算法的性能. 以下的实验过程均在装有 MATLAB R2018a 软件的 Lenovo Y430p 计算机上进行,具体硬件配置如下:CPU 型号为 Intel 酷睿 i7 4710MQ、CPU 主频为 2.5 GHz、内存容量 4 GB、显存容量 2 GB,该计算机的操作系统是 64 位的 Windows 8.1。

所提算法需要进行一些参数设置:搜索窗大小 $L \times L$ 、从前后两参考帧挑选匹配块的总个数 C 、重叠块的大小 S^2 、每个匹配块权重的参数常量 h 、阈值 T . 基于算法复杂度的考虑,实验中的参数值选取如下: $S^2 = 8 \times 8$ 、 $C = 10$ 、 $h = 8$ 、 $T = 0.9$. 由于所提算法的目标是从整体质量上改进视频重构的效果,本文选取 SSIM 作为算法性能的主要评价指标,同时辅以峰值信噪比 (PSNR) 进行指标分析。

结合式(2)与峰值信噪比 PSNR 的定义,可得

$$P_{\text{PSNR}} = 20 \log_{10} \left(\frac{\text{MAX}}{\sqrt{\text{MSE}(\hat{x}_k^p, x_{\text{ref},i})}} \right)$$

式中:MAX 为常数,如果每个像素用 8 位表示,那么其值为 255.从式中可看出 PSNR 与 MSE 成反比,而 MSE 表征了重构图像与原始图像的最小均方误差,其值越小表明重构效果越好.基于以上分析可得出 PSNR 值越大,均方误差意义上的重构质量越高.

为方便叙述,规定如下:文献[12]中的算法为多假设(MH)算法,文献[18]中的算法为 RRS 算法,本文所提算法(Proposed)主要与这两种代表性的算法作对比.

在进行匹配块的选取过程中,可发现当按照多种分配从前后两个参考帧中挑选匹配块时,不同分配会在 SSIM 与 PSNR 上产生区别,基于帧距离的匹配块选择机制的实验示例如图 4 所示.图中的“前”“后”表示从前后两个参考帧挑选的匹配块的个数,例如“前 4 后 6”表明当前重构块从前向参考帧中挑选 4 个匹配块、而从后向参考帧中挑选 6 个匹配块.从图中可看出,图 4(a)中的两曲线相差较大,图 4(b)中的两曲线基本重合,这表明两种挑选匹配块的配置导致重构 PSNR 基本一致,但在 SSIM 上却有较大的差别.一般而言,SSIM 值越大,重构视频在整体质量会越好.因此,所提算法可基于帧距离的匹配块选择机制,利用多模式重构动态地挑选匹配块.

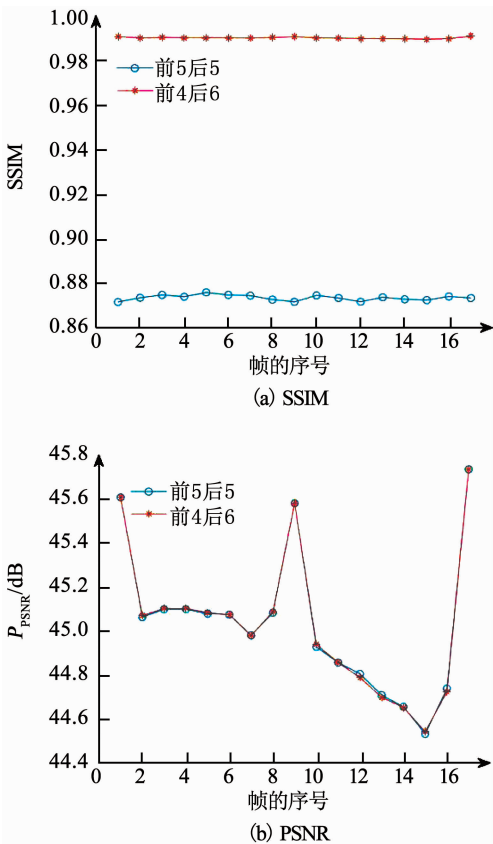


图 4 不同挑选匹配块配置的 SSIM 与 PSNR

Fig. 4 SSIM and PSNR results of different matching patch combinations

由于本文算法从 MSE 与 SSIM 两方面全面考虑了视频重构的质量,通过增加一定的复杂度进行了更为精细的预测,从理论上可以预期所提算法在 PSNR 与 SSIM 上均会有较好的表现;同时,在进行大量 CIF 格式视频序列实验的基础上,本文挑选了 3 个具有代表性的 Foreman、Akiyo 与 Container 视频序列在多种采样率下依次执行 MH 算法、RRS 算法与本文所提算法,得出的实验结论基本一致.

在实验中,每种序列选取连续的 17 帧,其中第 1、9、17 选用较高的采样率并作为参考帧,其余的各帧为非关键帧.使用三种算法对应的重构结果如表 2、表 3 所示.其中,表 2 展示重构结果的 PSNR 平均值;表 3 展示重构结果的 SSIM 平均值.对于两表中的联合采样率,逗号前为参考帧采样率,逗号后的为非关键帧的采样率.从表 2 可以看出,在 (0.4, 0.3)、(0.6, 0.4)、(0.7, 0.2)、(0.5, 0.4) 4 种联合采样率下,RRS 算法在 PSNR 方面相对 MH 算法有所改进,在 3 个序列各采样率下提高大概 1 ~ 2dB,而本文所提算法在 PSNR 方面基本与 RRS 算法基本持平,甚至在某些采样率下有较小的改善.由表 3 可以看出,在整体结构相似性上,MH 与 RRS 两种算法的性能基本一致,本文算法对于上述两种算法在多种采样率下具有较好的改进.所提算法在 3 个视频序列上的 SSIM 平均值相较于 RRS 算法提高大约 10%.

表 2 三种算法在不同视频序列上的 PSNR/dB				
Tab.2 PSNR of three algorithms on different sequences/dB				
联合采样率	视频序列	MH	RRS	Proposed
(0.4,0.3)	Foreman	38.32	38.75	39.01
	Akiyo	40.42	43.03	43.04
	Container	32.34	33.90	33.90
	Average	37.02	38.56	38.65
(0.6,0.4)	Foreman	40.66	41.57	41.59
	Akiyo	44.34	46.02	46.03
	Container	35.41	38.01	38.05
	Average	40.14	41.87	41.89
(0.7,0.2)	Foreman	37.34	39.01	39.34
	Akiyo	43.96	44.50	44.61
	Container	30.59	31.74	31.73
	Average	37.30	38.41	38.56
(0.5,0.4)	Foreman	40.33	40.95	40.96
	Akiyo	41.66	45.03	45.03
	Container	34.67	36.06	36.06
	Average	38.89	40.68	40.68

表 3 三种算法在不同视频序列上的 SSIM

联合采样率	视频序列	MH	RRS	Proposed
(0.4,0.3)	Foreman	0.94	0.79	0.96
	Akiyo	0.79	0.88	0.91
	Container	0.75	0.79	0.90
	Average	0.82	0.82	0.92
(0.6,0.4)	Foreman	0.90	0.74	0.97
	Akiyo	0.84	0.99	0.99
	Container	0.81	0.87	0.97
	Average	0.85	0.86	0.98
(0.7,0.2)	Foreman	0.94	0.96	0.96
	Akiyo	0.76	0.89	0.99
	Container	0.83	0.73	0.84
	Average	0.84	0.86	0.93
(0.5,0.4)	Foreman	0.93	0.76	0.97
	Akiyo	0.83	0.87	0.99
	Container	0.66	0.84	0.86
	Average	0.80	0.82	0.94

本文算法与 MH 算法、RRS 算法从原理上进行简单的比较,MH 算法利用多假设的方法进行重构,

只进行了一次重构;RRS 重构分为两个阶段,第一阶段基于 MH 算法进行初步的重构,在第二阶段进一步进行重构.上述两种方法均采用分块的方法进行重构,并通过 MSE 度量重构质量,因此在 PSNR 质量上表现较好,但在 SSIM 度量上欠缺考虑.本文算法从 MSE 与 SSIM 两方面进行度量,因此在整体重构质量上有更好的表现.

为更加直观地比较所提算法的性能,依次挑选 Akiyo 序列在 (0.5,0.4) 联合采样率下、Foreman 序列在 (0.5,0.2) 联合采样率下、Container 在 (0.4,0.3) 联合采样率下分别执行三种重构算法.图 5 分别给出了三种重构算法的 SSIM 与 PSNR 对比图.由各图中可看出,所提算法在 SSIM 指标上远高于 MH 算法和 RRS 算法,也就是在视频重构的整体质量上,所提算法具有较好的效果;在 PSNR 指标方面,本文算法与 RRS 算法基本一致,都远高于 MH 算法.对于所选取的 Foreman 序列,在 SSIM 与 PSNR 指标方面也有同样的实验结果.综上所述,本文算法在保持较高 PSNR 质量的同时,对于 SSIM 质量具有较为显著的改善.

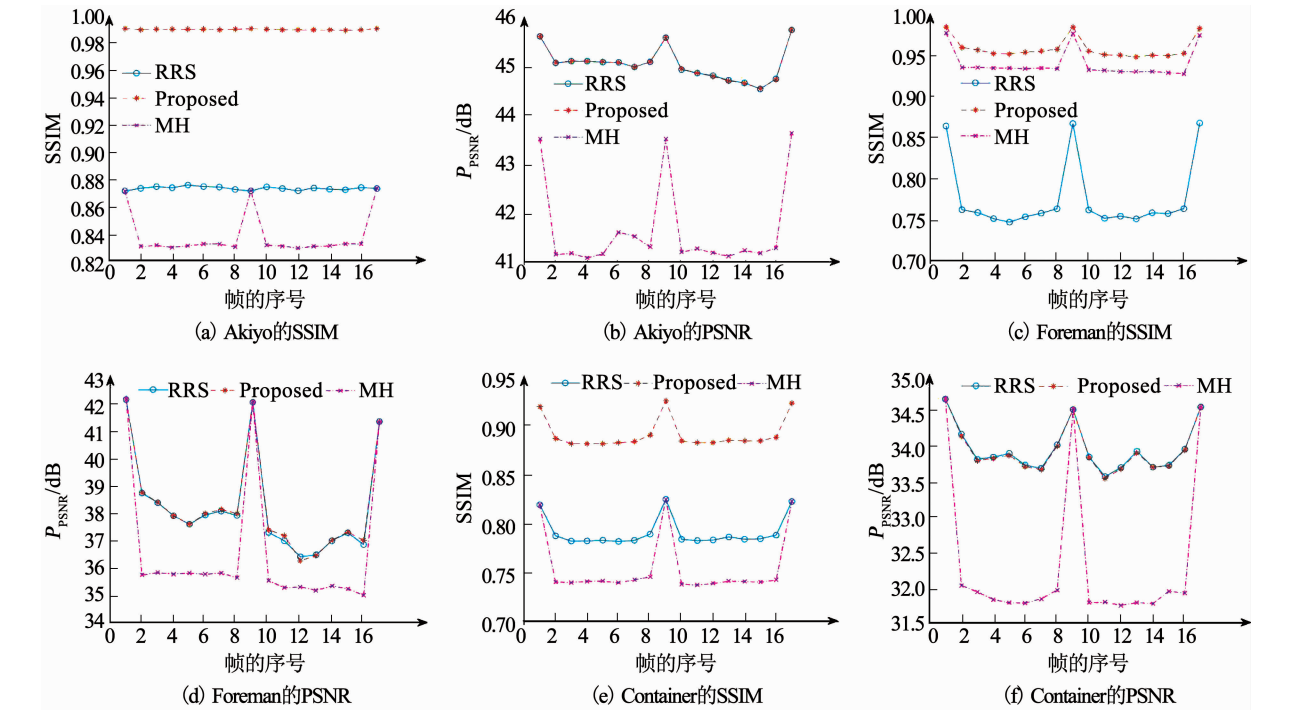


图 5 3 个序列在不同算法下重构结果的 SSIM 与 PSNR 对比

Fig. 5 SSIM and PSNR comparison of reconstruction results under different algorithms on three sequences

基于以上分析,可以得出本文所提算法相较于 MH 算法,无论从 PSNR 还是 SSIM 两方面均取得较大性能的提升,图像的重构质量得到明显的提升;与 RRS 算法相比较,虽然在 PSNR 上改进的效果不是特别明显,但是在 SSIM 方面却有较大的改善,可以

得出从整体重构质量方面,本文算法相较于 RRS 也有一定程度上的改善.同非压缩感知系统相比,压缩感知系统的相对优势是测量端简单、重建端复杂,可将测量端的复杂度向重建端转移,适合于无线多媒体传感应用.一般而言,重建端具有充足的计算资

源,其核心难题是如何提升重构质量. 本文算法通过适当增加重建端的复杂度,在现有方法的基础上进一步提升了视频信号的重构质量.

提出算法的复杂度主要由挑选匹配块的过程所决定. 所提算法在极端情况下才达到 RRS 算法的 3 倍复杂度,因此重构每帧的算法复杂度大概为 $O(N)$. 基于前文所述实验平台的具体软硬件条件以及实验参数的设置,对于本文所使用的 CIF 格式的视频序列,所提算法重构每帧图像平均大约花费 5~6 分钟;相较于 RRS 与 MH 算法的 3~4 分钟,所提算法在增加一定算法复杂度的情况下,重构效果得到了改进. 由于硬件并行计算的发展,上述复杂的计算问题将得到一定程度的解决.

4 结 论

现有的压缩感知重构算法基于均方误差的基础上,进行重构,存在一定的弊端. 本文提出了一种基于动态多模式挑选匹配块的视频压缩感知重构算法,在迭代重构中有效融合了均方误差与结构相似准则,通过多模式视频重构获取最终的重构结果. 所提算法充分利用了视频帧的时空相关性,提高了重构视频的整体质量,在 PSNR 与 SSIM 性能指标方面相比于当前代表性的算法均有所改善,无论在重构图像的细节和整体质量上都有提升. 但本文所提算法复杂度有一定提升,对计算硬件要求较高.

参考文献

- [1] Donoho D L. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52 (4): 1289. DOI: 10.1109/TIT.2006.871582
- [2] ZHAO Yingxin, LIU Zhiyang, WANG Yuanyuan, et al. Sparse coding algorithm with negentropy and weighted l_1 -norm for signal reconstruction[J]. Entropy, 2017, 19(11): 599. DOI: 10.3390/e19110599
- [3] Prades-Nebot J, Ma Y, Huang T. Distributed video coding using compressive sampling [C]//Picture Coding Symposium. USA: IEEE, 2009: 1. DOI: 10.1109/PCS.2009.5167431
- [4] LU Gan. Block compressed sensing of natural images [C]//15th International Conference on Digital Signal Processing. Welsh: IEEE, 2007: 403. DOI: 10.1109/ICDSP.2007.4288604
- [5] MUN S, FOWLER J E. Block compressed sensing of images using directional transforms[C]//IEEE International Conference on Image Processing (ICIP). Egypt: IEEE, 2009: 3021. DOI: 10.1109/ICIP.2009.5414429
- [6] Sankaranarayanan A C, Studer C, Baraniuk R G. CS-MUVI: Video compressive sensing for spatial-multiplexing cameras [C]//IEEE International Conference on Computational Photography (ICCP). USA: IEEE, 2012: 3021. DOI: 10.1109/ICCPH.2012.6215212
- [7] Park J Y, Wakin M B. A multiscale framework for compressive sensing of video [C]//IEEE Picture Coding Symposium. USA: IEEE, 2009: 1. DOI: 10.1109/PCS.2009.5167440
- [8] Mun S, Fowler J E. Residual reconstruction for block-based compressed sensing of video[C]//IEEE Data Compression Conference. USA: IEEE, 2011: 183. DOI: 10.1109/DCC.2011.25
- [9] Vaswani N. LS-CS-residual (LS-CS): Compressive sensing on least squares residual [J]. IEEE Transactions on Signal Processing, 2010, 58(8): 4108. DOI: 10.1109/TSP.2010.2048105
- [10] Jung H, Sung K, Nayak K S, et al. A general compressed sensing framework for high resolution dynamic MRI[J]. Magnetic resonance in medicine, 2010, 61(1): 103. DOI: 10.1002/mrm.21757
- [11] YANG Jianbo, YUAN Xin, LIAO Xuejun, et al. Video compressive sensing using Gaussian mixture models [J]. IEEE Transactions on Image Processing, 2014, 23 (11): 4863. DOI: 10.1109/TIP.2014.2344294
- [12] TRAMEL E W, FOWLER J E. Video compressed sensing with multihypothesis [C]//IEEE Data Compression Conference. USA: IEEE, 2011: 193. DOI: 10.1109/DCC.2011.26
- [13] Chen C, Tramel E W, Fowler J E. Compressed-sensing recovery of images and video using multihypothesis predictions [C]//45th Asilomar Conference on Signals, Systems and Computers (ASIOMAR). Pacific: IEEE, 2012: 1193. DOI: 10.1109/ACSSC.2011.6190204
- [14] KUO Yonghong, WU Kai, CHEN Jian. A scheme for distributed compressed video sensing based on hypothesis set optimization techniques [J]. Multidimensional Systems & Signal Processing, 2017, 28(1): 1. DOI: 10.1007/s11045-015-0337-4
- [15] CHEN Jian, CHEN Yunzheng, DONG Qin, et al. An elastic net-based hybrid hypothesis method for compressed video sensing [J]. Multimedia Tools and Applications, 2015, 74 (6): 2085. DOI: 10.1007/s11042-013-1743-y
- [16] 阔永红, 王需泉, 陈健. 基于多参考帧假设优化的压缩感知重构算法[J]. 通信学报, 2017, 38(12): 1. DOI: 10.11959/j.issn.1000-436x.2017297
- KUO Yonghong, WANG Ruquan, CHEN Jian. Reconstruction algorithm based on multi-reference frames hypothesis optimization for compressive sensing[J]. Journal on Communications, 2017, 38 (12): 1. DOI: 10.11959/j.issn.1000-436x.2017297
- [17] 杨春玲, 戴超. 视频压缩感知中基于菱形快速搜索的双匹配区域预测[J]. 华南理工大学学报(自然科学版), 2018, 46(3): 55. DOI: 10.3969/j.issn.1000-565X.2018.03.008
- YANG Chunling, DAI Chao. A prediction scheme based on fast diamond search and two match regions in compressed video sensing [J]. Journal of South China University of Technology (Natural Science Edition), 2018, 46(3): 55. DOI: 10.3969/j.issn.1000-565X.2018.03.008
- [18] ZHAO Chen, Ma Siwei, Zhang Jian, et al. Video compressive sensing reconstruction via reweighted residual sparsity [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(6): 1182. DOI: 10.1109/TCSVT.2016.2527181
- [19] Fowler J E. Block-based compressed sensing of images and video [J]. Foundation Trends Signal Processing, 2010, 4 (4): 297. DOI: 10.1561/20000000033

(编辑 苗秀芝)