

## 基于变采样率的多假设预测分块视频压缩感知

练秋生\* 田天 陈书贞 郭伟  
(燕山大学信息科学与工程学院 秦皇岛 066004)

**摘要:** 现有的分块视频压缩感知通常对所有图像块均采用相同的测量矩阵进行测量, 这种平均分配采样率的测量方式忽略了视频中不同区域的结构复杂度和变化程度不同的事实。针对这一问题, 该文根据视频帧间相关性的分布特点提出了一种自适应分配采样率的变采样率压缩感知方法。将图像块按照帧间相关性的大小分类并分配不同的采样率, 重构过程采用变采样率多假设预测算法以充分利用帧间相关性。实验结果表明该文算法能够在低采样率下重构出高质量的视频图像, 而且这种变采样率测量的方式有利于提高运动剧烈区域的重构质量。

**关键词:** 压缩感知; 变采样率; 自适应采样; 帧间相关性; 多假设预测

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2013)01-0203-06

DOI: 10.3724/SP.J.1146.2012.00590

## Block Compressed Sensing of Video Based on Variable Sampling Rates and Multihypothesis Predictions

Lian Qiu-sheng Tian Tian Chen Shu-zhen Guo Wei

(Institute of Information Science and Technology, Yanshan University, Qinhuangdao 066004, China)

**Abstract:** For most of those existing block-based compressed sensing of video, the same measurement matrix is usually utilized for all blocks, which underestimates the fact that the structural complexity and the movement varies from different regions. To address this issue, a novel block-based adaptive compressed sensing algorithm with variable sampling rate is proposed according to the distribution characteristics of the correlations between neighboring frames. It classifies blocks into different types depending on their inter-frame correlation, and adjusts the sampling rate accordingly. Multihypothesis predicting algorithm is used to reconstruct the videos to make full use of the inter-frame correlation. The experiment shows that the proposed algorithm reduces the number of sampled measurements while still improving the quality of the reconstructed frames. Also, with the variable sampling rate method, a higher reconstruction quality can be achieved for the regions containing relatively fast movement.

**Key words:** Compressed Sensing (CS); Variable sampling rates; Adaptive sampling; Inter-frame correlation; Multihypothesis prediction

### 1 引言

压缩感知理论 (Compressed Sensing, CS) 由 Donoho, Candès 和华裔科学家 Tao 等人在 2004 年提出, Candès 等人<sup>[1]</sup>在 2006 年从数学上证明了可以从部分傅里叶变换系数精确重构原始信号, 为压缩感知奠定了理论基础。由于高分辨率的数字图像和视频频带较宽, 基于香农定理进行先采样后压缩的传统数据处理方式造成了采样资源的严重浪费<sup>[2]</sup>。压缩感知理论提出将传统的数据获取与数据压缩合二为一, 在信号获取的过程中寻找最少的系数来表述

信号, 并在需要时利用稀疏性或低秩性<sup>[3,4]</sup>等先验知识通过重构算法从已知的少量测量值中以高概率重构出原信号<sup>[5,6]</sup>。

视频压缩采样过程可以通过连续使用单像素照相机实现<sup>[7]</sup>, 这种整帧采样的方式会导致存储空间和算法复杂度的问题。为此文献[8]提出了一种块采样模式, 对单帧图像分块并对所有块均用相同的测量矩阵单独采样。与直接对整帧图像采样重构的方法相比, 针对块采样的单帧重构算法在速度和质量上都显示了明显的优越性<sup>[9,10]</sup>。然而, 这些针对视频单帧处理的方法忽略了视频连续帧之间所具有的强相关性。基于运动估计和运动补偿的视频压缩感知重构算法不仅利用了视频帧间相关性, 而且把传统视频压缩中编码端的复杂度转移到了解码端, 降低了

2012-05-16收到, 2012-10-22改回

国家自然科学基金(61071200, 60772079)和河北省自然科学基金(F2010001294)资助课题

\*通信作者: 练秋生 lianqs@ysu.edu.cn

视频采集设备的功耗和成本<sup>[11,12]</sup>。视频压缩感知的运动估计过程是在接收端用已知的测量值粗略重构出带有明显噪声的近似帧,再对其进行运动估计得到预测。相比之下,Tramel等人<sup>[13,14]</sup>提出的在测量域直接预测的多假设帧预测方法(MultiHypothesis Frame Prediction, MHFP)能够明显提高预测帧的准确性。但是由于忽略了视频场景中不同区域的运动特点,这种方法对帧间相关性的利用率仍然有待提高。文献[15]虽然意识到了这一问题,却由于自适应采样时需要周期性地加入重构过程而大大增加了采样端的复杂度,违背了压缩感知理论的初衷。

考虑到视频序列不同区域的帧间相关性不同,本文提出一种简单有效的基于变采样率的分块视频压缩感知算法。首先利用相同的测量矩阵得到目标帧与其参考帧对应块的测量向量,根据其差值能量将块分为3类:近似不变块、缓慢变化块和快速变化块;再对不同类别的块分别选用不同的采样率处理。这种测量方法能够根据不同块具有的不同场景复杂度和变化强度自适应调整自身的测量点数(即块采样率),从而达到合理分配采样率的目的。

在接收端,本文通过变采样率的多假设预测得到预测帧,并用基于冗余残差字典的迭代加权最小二乘法的 $\varepsilon$ 正则化变型( $\varepsilon$ -regularized Iteratively Reweighted Least Squares,  $\varepsilon$ -regularized IRLS)<sup>[16]</sup>重构预测残差,最终得到高质量的重构帧。本文提出的基于变采样率测量的视频压缩感知结构具体过程如图1所示。

## 2 视频序列的变采样率测量

### 2.1 参考帧的固定高采样率测量

对于 $n \times n$ 的单帧图像,若将其作为一个 $n^2 \times 1$ 的向量直接采样,则需要 $m \times n^2$ 的测量矩阵以获得 $m$ 个测量值,显然如此规模的测量矩阵的存储和计算都将给视频压缩感知系统带来很大困难。本文采用文献[8]提出的块采样模式,将参考帧不重叠分块得到 $K$ 个 $B \times B$ 大小的块,并对各个图像块用相同的测量矩阵 $\Phi$ 单独采样。对于第 $i$  ( $i = 1, 2, \dots, K$ )个图像块 $\mathbf{x}_{t-1}^i$ ,其测量向量可以表示为

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

式中 $\Phi \in \mathbf{R}^{M \times N}$  ( $M = \lfloor (mN)/n^2 \rfloor$ ,  $N = B^2$ ),  $\mathbf{y}_{t-1}^i \in \mathbf{R}^{M \times 1}$ 。则整个参考帧的采样过程可以表示为 $\mathbf{Y}_{t-1} = \Phi \mathbf{X}_{t-1}$ ,其中 $\mathbf{X}_{t-1} \in \mathbf{R}^{N \times K}$ 和 $\mathbf{Y}_{t-1} \in \mathbf{R}^{M \times K}$ 的列向量分别为各图像块的像素向量和测量向量。

为了更好地满足视频压缩感知的高质量重构要求,本文采用K-SVD(K-Singular Value Decomposition)方法<sup>[17]</sup>训练的4倍冗余字典使视频信号在字典下的稀疏表示系数更为稀疏:

$$\mathbf{x}_{t-1}^i = \Psi \alpha_{t-1}^i \quad (2)$$

式中 $\Psi$ 为KSVD字典, $\alpha_{t-1}^i$ 为稀疏表示系数。另外,测量矩阵 $\Phi$ 的设计也是压缩感知理论中的重要问题,本文采用的是由KSVD字典 $\Psi$ 通过优化准则得到的最佳测量。设计最佳测量矩阵的关键在于使全息基 $\Theta = \Phi \Psi$ 中原子的互相关性最小,即构造一个测量矩阵 $\Phi$ 使其满足优化准则<sup>[18]</sup>:  $\Theta^T \Theta \approx I$ 。

### 2.2 非参考帧的变采样率测量

视频中不同的运动目标可能以不同方式运动,因此不同区域的帧间相关性不同。对于非参考帧的测量,文中按照不同区域相对参考帧的不同变化程度合理分配采样率以获得高效采样,即变化程度较小的区域分配较低的采样率,变化程度较大的区域分配较高的采样率,这样保证了在较低的总采样率下仍能高质量地重构出快速变化的区域。当场景发生突变时,使用本文提出的变采样率采样时,几乎所有块都将以最高采样率采样,可以保证突变场景的高质量重构。

**2.2.1 分类判决准则** 将非参考帧按照与参考帧相同的方式不重叠分块,则第 $i$ 块 $\mathbf{x}_t^i$ 与参考帧中对应位置的 $\mathbf{x}_{t-1}^i$ 之间的差别反映了该位置上的帧间相关性,即两对应块的残差能量大小 $E(\mathbf{x}_{t-1}^i, \mathbf{x}_t^i) = \|\mathbf{x}_{t-1}^i - \mathbf{x}_t^i\|_2^2$ 可以作为对块分类的判决准则。然而,由于不同视频的亮度和对比度相差较大,对应的残差能量会有很大差异,以此作为判决准则时其判决阈值的选择对视频的依赖性将会非常强,从而影响算法通用性。故本文采用残差与参考块的能量比值作为判决准则。

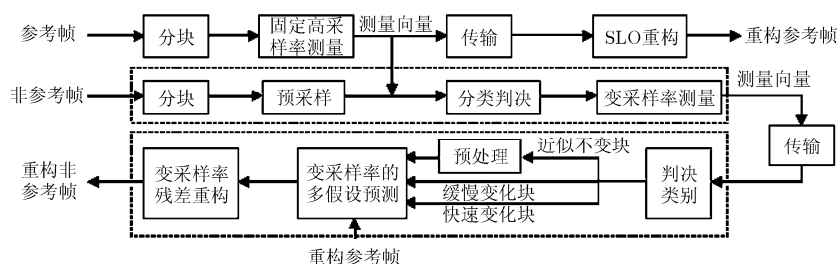


图1 基于变采样率的视频压缩感知结构

$$e(\mathbf{x}_{t-1}^i, \mathbf{x}_t^i) = \frac{E(\mathbf{x}_{t-1}^i, \mathbf{x}_t^i)}{E(\mathbf{x}_{t-1}^i)} \quad (3)$$

需要注意的是，压缩感知系统中视频帧的像素值是未知的，仅已知其测量值。因此本文采用测量域的残差能量与参考块能量的比值作为判决准则。

$$e(\mathbf{y}_{t-1}^i, \mathbf{y}_t^i) = \frac{E(\mathbf{y}_{t-1}^i, \mathbf{y}_t^i)}{E(\mathbf{y}_{t-1}^i)} \quad (4)$$

对 Foreman, News, Susie, Football, Caltrain, Garden, Tempete 和 Tennis 共 8 组标准视频序列的 12968 个图像块使用阈值  $T_1$  和  $T_2$  进行分类 ( $T_1 < T_2$ )。根据式(3)计算第  $i$  块的像素域判决函数值，如果  $e(\mathbf{x}_{t-1}^i, \mathbf{x}_t^i) < T_1$ ，则将该块判决为近似不变块。将所有近似不变块由式(4)计算其测量域判决函数值，如果也满足  $e(\mathbf{y}_{t-1}^i, \mathbf{y}_t^i) < T_1$ ，则表明利用测量域判决函数能够正确分类。对满足  $T_1 \leq e \leq T_2$  的缓慢变化块和满足  $e > T_2$  的快速变化块的测量域判决函数能否实现正确分类的测试方法同理可得。统计结果显示，对由像素域判决函数确定的近似不变块根据测量域分类判决时分类正确的概率为 93.35%，缓慢变化块和快速变化块分类正确的概率分别为 89.68% 和 97.02%。实验表明，测量域的残差能量比能够真实体现像素域的变化情况，可以实现正确分类。

**2.2.2 变采样率测量** 设近似不变块、缓慢变化块和快速变化块的采样率分别为  $S_1$ 、 $S_2$  和  $S_3$ ，其中  $S_3$  与参考帧的块采样率相同，且  $S_1 < S_2 < S_3$ ，对应的块采样点数分别为  $M_1 = \lfloor S_1 \cdot B^2 \rfloor$ ， $M_2 = \lfloor S_2 \cdot B^2 \rfloor$  和  $M$ 。

首先对非参考帧进行预采样。由参考帧测量矩阵  $\Phi$  的前  $M_2$  行构成预采样矩阵  $\Phi_2$ ，对非参考帧图像块  $\mathbf{x}_t^i$  进行投影得到测量值  $\mathbf{y}_{t,M_2}^i = \Phi_2 \mathbf{x}_t^i$ 。然后将预采样值  $\mathbf{y}_{t,M_2}^i$  与参考帧测量值的前  $M_2$  行  $\mathbf{y}_{t-1,M_2}^i$  作比较并根据设定的阈值  $T_1$  和  $T_2$  对其进行分类判决。由式(4)计算  $\mathbf{x}_t^i$  的判决函数值  $e$ ，如果  $e < T_1$ ，则将其判决为近似不变块，只保留测量向量中前  $M_1$  个测量值得到  $\mathbf{y}_{t,M_1}^i$ ，即减小该块采样率到  $S_1$ ；如果  $T_1 \leq e \leq T_2$ ，则判决该块为缓慢变化块，保持  $M_2$  个测量值不变，采样率仍为  $S_2$ ；如果  $e > T_2$ ，则判决为快速变化块，增大其采样率到  $S_3$ ，即由式(5)和式(6)增加其测量向量长度，得到新的测量向量  $\mathbf{y}_{t,M}^i$ ：

$$\Delta \mathbf{y}_t^i = \Phi_\Delta \mathbf{x}_t^i \quad (5)$$

$$\mathbf{y}_{t,M}^i = \left[ (\mathbf{y}_{t,M_2}^i)^T, (\Delta \mathbf{y}_t^i)^T \right]^T \quad (6)$$

其中  $\Phi_\Delta$  是由测量矩阵  $\Phi$  的  $M_2 + 1 \sim M$  行构成的附加测量矩阵。

为保证用式(4)作为分类判决函数时能够实现准确分类，需要预采样得到的测量  $\mathbf{y}_{t,M_2}^i$  能够正确反映当前块的真实变化情况，即测量向量的信息量应

该足够大，则预采样过程就不能使用太低的采样率，而预采样率太高又会影响到采样速度。因此需要适度选择预采样率，实现既能降低错误分类的风险又能避免采样速度受到影响。2.2.1 小节中的正确分类统计结果是在预采样率为 20% 时得到的，正确分类概率较高，且采样速度也适中。

由于本文的变采样率测量过程是自适应的，我们事先并不知道非参考帧的真实采样率，因此在实际传输过程中，每一块都需要多传输一个数据用以记录该块的类别。但是这些附加数据在实际传输过程中增加的额外负担并不明显，如果每个测量值用 1 个字节表示，而使用二进制数 01, 10 和 11 分别代表第 1、2、3 类，则每块只需用 2 bit 传输其类别信息。非参考帧的实际采样率为

$$S = \frac{K_1 \times M_1 + K_2 \times M_2 + K_3 \times M + K/4}{n^2} \quad (7)$$

式中  $n^2$  表示非参考帧总像素数， $K$  表示分块的总个数， $K_1$  表示判决为第 1 类图像块的个数， $K_2$  为第 2 类图像块的个数， $K_3$  为第 3 类图像块的个数，并且满足  $K = K_1 + K_2 + K_3$ 。

### 3 基于变采样率测量值的重构

#### 3.1 参考帧的光滑 $\ell_0$ 范数重构

对于固定高采样率测量的参考帧，直接应用单帧压缩感知的快速重构算法就可以得到高质量的重构效果。由于测量数  $M$  小于信号长度  $N$ ，从  $\mathbf{y}_{t-1}^i = \Phi \mathbf{x}_{t-1}^i$  中解  $\mathbf{x}_{t-1}^i$  就成了具有不唯一解的欠定方程组的求解问题。利用视频信号在字典  $\Psi$  下有稀疏表示的先验知识，可以将求解欠定方程组的问题转化为最小  $\ell_0$ -范数问题：

$$\hat{\alpha}_{t-1}^i = \arg \min \|\hat{\alpha}_{t-1}^i\|_0, \text{ s.t. } \mathbf{y}_{t-1}^i = \Phi \Psi \alpha_{t-1}^i \quad (8)$$

式中  $\|\alpha_{t-1}^i\|_0$  表示稀疏系数中非零元素的个数，即稀疏度。再由  $\hat{\mathbf{x}}_{t-1}^i = \Psi \hat{\alpha}_{t-1}^i$  得到重构信号  $\hat{\mathbf{x}}_{t-1}^i$ 。

然而，最小  $\ell_0$  范数的求解是一个 NP-hard 问题，可以将其转化为最小  $\ell_1$  范数或光滑  $\ell_0$  范数 (Smoothed  $\ell_0$ ，文中将其定义为 SL0 范数) 的优化求解<sup>[19]</sup>。本文选用求解最小 SL0 范数的方法来快速精确重构参考帧：

$$\hat{\alpha}_{t-1}^i = \arg \min \|\alpha_{t-1}^i\|_{\text{SL0}}, \text{ s.t. } \mathbf{y}_{t-1}^i = \Phi \Psi \alpha_{t-1}^i \quad (9)$$

式中  $\alpha_{t-1}^i = [\alpha_1, \dots, \alpha_L]^T$ ， $\ell_0$  范数的光滑近似 SL0 范数可以表示为： $\|\alpha_{t-1}^i\|_{\text{SL0}} = L - \sum_{l=1}^L \exp(-\alpha_l^2 / 2\sigma^2)$ 。

其中， $\sigma$  值用来平衡近似函数 SL0 范数的精确度和光滑程度， $\sigma$  越小与  $\ell_0$  范数越相似。

本文通过 <http://ee.sharif.ir/~SLzero> 的 SL0 软件包优化求解式(9)，再由  $\hat{\mathbf{x}}_{t-1}^i = \Psi \hat{\alpha}_{t-1}^i$  得到高质量重构参考帧。实验选用的  $\sigma$  序列为  $\sigma_k = \mu \sigma_{k-1}$  ( $k$

$\geq 2$ ),  $\sigma$  初始值  $\sigma_1 = 2 \max_i |\alpha_i|$ , 下降因子  $\mu = 0.25$ 。

### 3.2 非参考帧的变采样率多假设重构

通过上述过程能够在固定的高采样率下获得高分辨率的重构参考帧。将其作为非参考帧的变采样率多假设预测重构的参考, 有效提高了非参考帧的重构质量。

**3.2.1 近似不变块的预处理** 在接收端, 首先判断接收到的非参考帧各块的所属类别。如果是近似不变块  $\mathbf{y}_{t,M_1}^i$ , 则需要先通过预处理过程。用参考帧对应块的测量向量  $\mathbf{y}_{t-1}^i$  中最后  $M - M_1$  个值将  $\mathbf{y}_{t,M_1}^i$  补长到  $M$  :

$$\mathbf{y}_{t,M}^i = [(\mathbf{y}_{t,M_1}^i)^\top, \mathbf{y}_{t-1}^i(M_1 + 1), \dots, \mathbf{y}_{t-1}^i(M)]^\top \quad (10)$$

其中  $\mathbf{y}_{t-1}^i(j)$  表示  $\mathbf{y}_{t-1}^i$  中的第  $j$  个元素。

由于近似不变块相对参考帧中对应块的运动非常小, 包含的新信息最少, 已有的  $M_1$  个测量值足以代表两帧之间的微小差别。同时, 把参考块的部分测量值近似看做当前近似不变块的测量值使用, 相当于增加近似不变块的采样率, 有效提高了非参考帧中低采样率测量的近似不变块的重构质量。

**3.2.2 变采样率的多假设预测** 为充分利用视频所特有的帧间相关性, 本文选用在测量域预测的多假设预测方法<sup>[13]</sup>。

$$\tilde{\mathbf{X}}_t = \arg \min_{\tilde{\mathbf{X}}_t} \|\mathbf{Y}_t - \Phi \tilde{\mathbf{X}}_t\|_2^2 \quad (11)$$

由于被预测信号的测量值  $\mathbf{Y}_t$  是已知的, 这种直接在测量域预测的过程确保了预测的准确性。注意到缓慢变化块  $\mathbf{y}_{t,M_2}^i$  的采样率与其它两类不同(测量向量长度均为  $M$ ), 因此对整个非参考帧的多假设预测过程可推广为变采样率的多假设预测。

给定块  $\mathbf{x}_t^i$ , 所有位于重构参考帧中对应位置搜索窗内的图像块均被看做用来预测  $\mathbf{x}_t^i$  的候选块, 这些候选块的最佳线性组合便构成了  $\mathbf{x}_t^i$  的多假设预测, 即  $\hat{\mathbf{x}}_t^i = \mathbf{H}_t^i \hat{\mathbf{w}}_t^i$ 。  $\mathbf{H}_t^i$  表示由所有候选块组成的矩阵, 其列向量  $\mathbf{h}_c$  ( $c = 1, \dots, C$ ) 为不同候选块的向量化表示, 则对  $\mathbf{x}_t^i$  的多假设预测过程也就是求解候选块的最佳线性组合系数  $\hat{\mathbf{w}}_t^i$  的过程:

$$\hat{\mathbf{w}}_t^i = \arg \min_{\mathbf{w}} \|\mathbf{y}_{t,M}^i - \Phi \mathbf{H}_t^i \mathbf{w}\|_2^2 \quad (12)$$

用 Tikhonov 正则化的方法求解式(12), 为其加入惩罚项, 得到

$$\hat{\mathbf{w}}_t^i = \arg \min_{\mathbf{w}} \|\mathbf{y}_{t,M}^i - \Phi \mathbf{H}_t^i \mathbf{w}\|_2^2 + \lambda \|\mathbf{\Gamma} \mathbf{w}\|_2^2 \quad (13)$$

$\lambda$  为拉格朗日因子, 本文选择  $\lambda = 0.0625$ 。Tikhonov 矩阵  $\mathbf{\Gamma}$  选择对角元素为  $\|\mathbf{y}_{t,M}^i - \Phi \mathbf{h}_c\|_2$  的对角矩阵, 通过这种结构,  $\mathbf{\Gamma}$  矩阵对系数向量  $\hat{\mathbf{w}}_t^i$  实施了奖惩策略: 与目标块越相似的候选块其加权系数就越大, 反之其加权系数就越小。由 Tikhonov 标准解公式可

以直接得到最佳线性组合系数:

$$\hat{\mathbf{w}}_t^i = ((\Phi \mathbf{H}_t^i)^\top (\Phi \mathbf{H}_t^i) + \lambda \mathbf{\Gamma}^\top \mathbf{\Gamma})^{-1} (\Phi \mathbf{H}_t^i)^\top \mathbf{y}_{t,M}^i \quad (14)$$

如果  $\mathbf{x}_t^i$  是缓慢变化块, 由于其测量向量  $\mathbf{y}_{t,M_2}^i$  长  $M_2$ , 用相同的过程进行多假设预测时则需要改用与之相匹配的测量矩阵  $\Phi_2$ 。

**3.2.3 变采样率的残差重构** 多假设预测得到的预测帧  $\tilde{\mathbf{X}}_t$  与原信号  $\mathbf{X}_t$  仍然存在一定误差  $\mathbf{R} = \mathbf{X}_t - \tilde{\mathbf{X}}_t$ 。由测量过程的线性特性可以得出残差  $\mathbf{R}$  的测量值等于预测帧  $\tilde{\mathbf{X}}_t$  与原信号在测量域的残差:

$$\mathbf{Q} = \Phi \mathbf{R} = \Phi (\mathbf{X}_t - \tilde{\mathbf{X}}_t) = \mathbf{Y}_t - \Phi \tilde{\mathbf{X}}_t \quad (15)$$

其中  $\mathbf{Q} \in \mathbf{R}^{M \times K}$ ,  $\mathbf{Y}_t$  表示接收端预处理之后得到的测量, 仅缓慢变化块的测量向量  $\mathbf{y}_t^i$  长为  $M_2$ , 故  $i \in I_2$  时预测残差向量  $\mathbf{q}^i$  也只保留前  $M_2$  个元素  $\mathbf{q}^i = [q_1, \dots, q_{M_2}]^\top$ 。

文中通过求解最小  $\ell_1$  范数的优化方法重构残差向量  $\mathbf{r}^i$  ( $i = 1, 2, \dots, K$ ), 得到残差  $\hat{\mathbf{R}} = [\hat{\mathbf{r}}^1, \hat{\mathbf{r}}^2, \dots, \hat{\mathbf{r}}^K]$ 。

$$\hat{\mathbf{r}}^i = \arg \min \|\mathbf{D}^\top \mathbf{r}^i\|_1, \text{ s.t. } \mathbf{q}^i = \Phi \mathbf{r}^i \quad (16)$$

则重构非参考帧为  $\tilde{\mathbf{X}}_t = \tilde{\mathbf{X}}_t + \hat{\mathbf{R}}$ 。采用  $\varepsilon$ -正则化的 IRLS 算法<sup>[16]</sup> 优化求解式(16), 将目标函数中的  $\ell_1$  范数替换为加权  $\ell_2$  范数:

$$\hat{\boldsymbol{\gamma}}^i = \arg \min \sum_{l=1}^L v_l \gamma_l^2, \text{ s.t. } \mathbf{q}^i = \Phi \mathbf{D} \boldsymbol{\gamma}^i \quad (17)$$

式中  $\mathbf{D}$  为对标准视频测试序列的残差数据用 K-SVD 方法训练的残差通用字典,  $\boldsymbol{\gamma}^i = [\gamma_1, \dots, \gamma_L]^\top$  为残差向量  $\mathbf{r}^i$  在残差字典  $\mathbf{D}$  下的稀疏表示系数, 即  $\hat{\mathbf{r}}^i = \mathbf{D} \boldsymbol{\gamma}^i$ 。在对式(17)的迭代求解过程中, 每一次迭代的加权系数  $v_l$  均由前一次迭代的结果计算得出

$$v_l = ((\gamma_l^{(k-1)})^2 + \varepsilon)^{-1/2} \quad (18)$$

则本次迭代的最优稀疏系数为

$$\boldsymbol{\gamma}^{(k)} = \mathbf{Q}_k (\Phi \mathbf{D})^\top (\Phi \mathbf{D} \mathbf{Q}_k (\Phi \mathbf{D})^\top)^{-1} \mathbf{q}^i \quad (19)$$

式中  $\mathbf{Q}_k$  为由本次迭代的加权系数的倒数  $1/v_l$  ( $l = 1, 2, \dots, L$ ) 构成的对角矩阵。文献[16]将正则化因子  $\varepsilon$  初始值设为 1, 并随着迭代的不断进行逐渐减小  $\varepsilon$  值, 直至稀疏系数收敛于最优解则停止迭代。

## 4 实验结果

本文用 MHFP 软件包中的 4 组标准视频序列测试文中提出的算法性能: Foreman, News, Susie, Football。该软件包可以直接在文献[13]作者的网站 <http://www.ece.msstate.edu/~ewt16/publications> 下载。考虑测试序列前两帧, 并将第 1 帧作为参考帧。实验中所有算法的块大小均为  $8 \times 8$ , 参考帧采样率均固定为 50%。本文算法非参考帧的采样率会根据测试序列本身的结构复杂程度和帧间运动的程度自适应调节大小, 因此事先是未知的, 但是可以通过人为改变参数(各类别块采样率或分类阈值)控

制其采样率的大概范围，以适应应用环境的需求(是对重构质量要求较高还是对传输数据量有较大限制)。本文通过对多组视频序列进行实验选择了一组通用性较强的参数，既保证了采样率较低，又保证了重构质量较高，并且在没有特殊要求的情况下可以作为固定参数使用。设定 3 种不同类别的块采样率分别为： $S_1 = 5\%$ ， $S_2 = 20\%$ ， $S_3 = 50\%$ ，分类判决的阈值分别为： $T_1 = 0.003$ ， $T_2 = 0.15$ 。

将 MHFP 中的搜索窗尺寸设定为  $\pm 4$  个像素。首先，实验比较了在其余条件完全相同的情况下(选用与 MHFP 中相同的随机投影矩阵以及参考帧和残差重构算法 BCS-SPL-DWT)使用变采样率测量方法 VS-MHFP1 与 MHFP 算法的重构结果，如表 1 所示。分析表中数据可以看出，在完全相同的背景条件下，仅仅依靠变采样率的测量方式就能使测试序列 Foreman, News, Susie 和 Football 的重构 PSNR 分别提高 0.69 dB, 0.43 dB, 1.21 dB, 0.62 dB。

表 1 MHFP 与 VS-MHFP1 视频帧重构质量比较(PSNR(dB))

序列	采样率	MHFP	VS-MHFP1
Foreman	0.11	31.65	<b>32.34</b>
News	0.13	29.50	<b>29.93</b>
Susie	0.09	34.07	<b>35.28</b>
Football	0.25	26.77	<b>27.39</b>

另外，表 2 给出了本文提出的视频压缩感知重构算法 VS-MHFP2 与 MHFP 及 FS-MHFP 的性能比较。其中 FS-MHFP 表示固定采样率的多假设预测重构方法(除采样方式外，FS-MHFP 的其它条件与 VS-MHFP2 相同)。从表中可以看出，VS-MHFP2 与 MHFP 和 FS-MHFP 相比，其重构质量均有明显提高，Foreman, News, Susie, Football 分别提高 2.67 dB, 5.55 dB, 2.52 dB, 2.95 dB 和 1.29 dB, 2.70 dB, 1.76 dB, 0.64 dB。观察比较图 2 可知本文算法 VS-MHFP2 能够有效重构视频中的主要运动区域。此外从图 3 显示的 Foreman 局部放大图中也可以看出对于运动程度较大的嘴唇周围部分，FS-MHFP 的块效应明显，而 VS-MHFP2 能够有效消除块效应，提高重构质量。

对于较简单的测试序列如 Susie 序列，本文方法可以自适应降低总采样率，并且能够保证很好的重构质量。而对于较复杂的测试序列如 Football 序列，本文方法则能够自适应增大采样率，确保一个可以接受的重构质量。换言之，本文提出的变采样率测量方式不需要人为判断视频复杂度来选择采样

表 2 MHFP, FS-MHFP 和 VS-MHFP2 视频帧重构质量比较(PSNR(dB))

序列	采样率	MHFP	FS-MHFP	VS-MHFP2
Foreman	0.12	31.98	33.36	<b>34.65</b>
News	0.14	29.63	32.48	<b>35.18</b>
Susie	0.09	34.07	34.83	<b>36.59</b>
Football	0.26	26.99	29.30	<b>29.94</b>

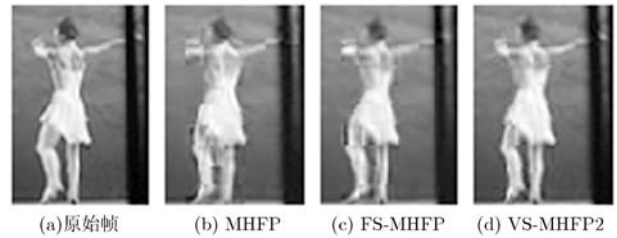


图 2 News 序列第 2 帧局部重构结果(采样率 14%)

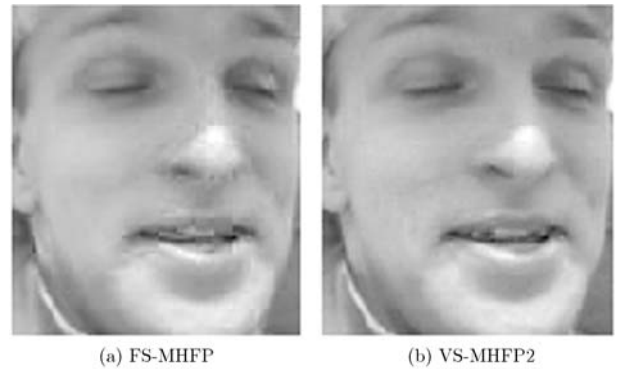


图 3 采样率为 12%时，Foreman 图像的重构结果(局部放大)

率，它可以自适应地调整总采样率大小，使本系统应用于不同场景时最终重构质量都不至于太低。

为了比较表 1 与表 2 中 4 种算法的复杂度，本文在相同的实验环境(CPU 配置: AMD Athlon(tm) II X2 255, 主频: 3.11 GHz, 内存: 1.75 GB, 运行环境: MATLAB R2010a)下记录了 4 种算法对表中 4 个视频序列进行处理的平均运行时间，MHFP 为 7.2001 s, VS-MHFP1 为 7.0366 s, FS-MHFP 为 4.4673 s, VS-MHFP2 为 18.6643 s。其中，VS-MHFP1 的运行时间与 MHFP 基本相同，表明本文所提出的变采样率思想基本没有给算法增加复杂度。而 VS-MHFP2 与 VS-MHFP1 相比重构速度较慢，算法复杂度较高，其主要原因在于本文选用的残差重构算法 IRLS 是一种迭代优化求解过程，重构速度比对图像进行整体小波变换的 BCS-SPL-DWT 算法慢。

## 5 结论

本文提出一种充分利用帧间相关性的变采样率

视频压缩感知算法,对结构复杂、运动较大的图像块以较高的采样率进行测量。在重构端采用变采样率多假设预测方法来达到充分利用帧间相关性的目的。实验结果表明,本文提出的变采样率测量方式可以根据不同视频场景的结构和运动自适应调整采样策略以合理分配采样率,在总采样率一定的条件下使整体效果达到最优,有效提高了采样效率。

### 参 考 文 献

- [1] Candes E J, Romberg J, and 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] 石光明, 刘丹华, 高大化, 等. 压缩感知理论及其研究进展[J]. *电子学报*, 2009, 37(5): 1070–1081.  
Shi Guang-ming, Liu Dan-hua, Gao Da-hua, *et al.* Advances in theory and application of compressed sensing[J]. *Acta Electronica Sinica*, 2009, 37(5): 1070–1081.
  - [3] Waters A E, Sankaranarayanan A C, and Baraniuk R G. SpaRCS: recovering low-rank and sparse matrices from compressive measurements[C]. Proceedings of Neural Information Processing Systems (NIPS), Granada, Spain, 2011.
  - [4] Lee K and Bresler Y. ADMiRA: atomic decomposition for minimum rank approximation[J]. *IEEE Transactions on Information Theory*, 2010, 56(9): 4402–4416.
  - [5] Donoho D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289–1306.
  - [6] Candes E J and Wakin M B. An introduction to compressive sampling[J]. *IEEE Signal Processing Magazine*, 2008, 25(2): 21–30.
  - [7] Wakin M B, Laska J N, Duarte M F, *et al.* Compressive imaging for video representation and coding[C]. Proceedings of the Picture Coding Symposium, Beijing, China, 2006.
  - [8] Gan L. Block compressed sensing of natural images[C]. Proceedings of the 15th International Conference on Digital Signal Processing, Cardiff, UK, 2007: 403–406.
  - [9] Mun S and Fowler J E. Block compressed sensing of images using directional transforms[C]. IEEE International Conference on Image Processing, Cairo, Egypt, 2009: 3021–3024.
  - [10] Fowler J E, Mun S, and Tramel E W. Multiscale block compressed sensing with smoothed projected landweber reconstruction[C]. Proceedings of the European Signal Processing Conference, Barcelona, Spain, 2011: 564–568.
  - [11] Mun S and Fowler J E. Residual reconstruction for block-based compressed sensing of video[C]. Proceedings of the IEEE Data Compression Conference, Snowbird, UT, 2011: 183–192.
  - [12] Jung H and Ye J C. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: what we can learn from video compression techniques[J]. *Imaging Systems and Technology*, 2010, 20(2): 81–98.
  - [13] Tramel E W and Fowler J E. Video compressed sensing with multihypothesis[C]. Proceedings of the IEEE Data Compression Conference, Snowbird, UT, 2011: 193–202.
  - [14] Chen C, Tramel E W, and Fowler J E. Compressed-sensing recovery of images and video using multihypothesis predictions[C]. Proceedings of the 45th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, 2011: 1193–1198.
  - [15] Liu Z, Zhao H V, and Elezzabi A Y. Block-based adaptive compressed sensing for video[C]. Proceedings of 2010 IEEE 17th International Conference on Image Processing, Hong Kong, China, 2010: 1649–1652.
  - [16] Chartrand R and Yin W. Iteratively reweighted algorithms for compressive sensing[C]. Proceedings of the 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, USA, 2008: 3869–3872.
  - [17] Aharon M, Elad M, and Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311–4322.
  - [18] Duarte-Carvajalino J and Sapiro G. Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization[J]. *IEEE Transactions on Image Processing*, 2009, 18(7): 1395–1408.
  - [19] Mohimani G H, Babaie-Zadeh M, and Jutten C. Fast sparse representation based on smoothed  $L^0$  norm[C]. Proceedings of the 7th International Conference on Independent Component Analysis and Signal Separation, London, UK, 2007: 389–396.
- 练秋生: 男, 1969年生, 教授, 博士生导师, 研究方向为压缩感知、稀疏表示、多尺度几何分析、图像处理。  
田天: 女, 1988年生, 硕士生, 研究方向为视频压缩感知。  
陈书贞: 女, 1968年生, 副教授, 研究方向为图像处理。  
郭伟: 男, 1972年生, 博士生, 研究方向为稀疏表示、雷达成像。