

doi:10.3969/j.issn.1001-893x.2013.03.009

基于残差重构的分布式视频压缩感知^{*}

常侃^{**},覃团发,唐振华

(广西大学 计算机与电子信息学院,南宁 530004)

摘要:为了改进分布式视频压缩感知方案的性能,提出了一种基于残差重构的分布式视频压缩感知方案。该方案在编码端逐帧独立进行测量,在解码端依靠视频信号的时域相关性提升重构信号质量。首先,对关键帧独立进行重构;其次,利用已重构关键帧做运动估计/运动补偿以生成非关键帧的边信息;接下来,对边信息采用与编码端相同的测量矩阵进行测量并计算测量残差值;最后,采用全变分最小化重构残差信号值并将其与边信息相加生成最终的重构图像。实验结果表明,在相同采样率下,与已有的分布式视频压缩感知方案相比,提出的方案可获得2.8 dB以上的峰值信噪比增益。

关键词:压缩感知;分布式视频压缩感知;残差重构;全变分最小化;边信息

中图分类号:TN919.8 **文献标志码:**A **文章编号:**1001-893X(2013)03-0274-05

Residual Reconstruction Based Distributed Compressed Video Sensing

CHANG Kan, QIN Tuan-fa, TANG Zhen-hua

(School of Computer and Electronic Information, Guangxi University, Nanning 530004, China)

Abstract: To improve the performance of Distributed Compressed Video Sensing(DCVS), a residual reconstruction based DCVS framework is proposed. The proposed framework samples each video frame independently at the encoder. However, it recovers frames at the decoder by exploiting inter frame correlation. Firstly, the key frame of a Group of Pictures(GOP) is independently reconstructed. Secondly, Side Information(SI) is generated by performing bi-directional Motion Estimation(ME) and Motion Compensation(MC) through the reconstructed key frames. Afterwards, the generated SI frame is sampled by the same matrix as the one at the encoder, and the measurement of SI is used to calculate the residual of measurement. Finally, total variation minimization is applied to reconstruct the residual signal, and the output frame is formed by adding SI to the residual signal. Experimental results show that compared with the existing DCVS method, the proposed one can get more than 2.8 dB Peak Signal to Noise Ratio(PSNR) increment at the same sampling rate.

Key words: compressed sensing(CS);distributed compressed video sensing(DCVS);residual reconstruction;total variation minimization; side information(SI)

1 引言

分布式视频编码(Distributed Video Coding, DVC)^[1]是一种特殊的视频编码框架。这类框架在

编码端对各帧进行独立编码,在解码端则对连续的视频帧进行联合解码以获取更高质量的解码结果。与传统混合视频编码框架不同,DVC把计算量较大的运动估计(Motion Estimation, ME)/运动补偿(Mo-

^{*} 收稿日期:2012-08-31;修回日期:2012-11-09 Received date:2012-08-31;Revised date:2012-11-09

基金项目:国家自然科学基金资助项目(61261023);广西自然科学基金资助项目(2011GXNSFD018024);广西教育厅科研项目资助(201203YB001)

Foundation Item: The National Natural Science Foundation of China (No. 61261023); The Natural Science Foundation of Guangxi (2011GXNSFD018024);The Foundation of Education Commission of Guangxi(201203YB001)

^{**} 通讯作者:pandack0619@163.com **Corresponding author:** pandack0619@163.com

tion Compensation, MC)操作转移到解码端,具有“轻编码、重解码”的特点,特别适合于电力和计算能力有限的应用场合,例如无线多媒体传感器、无线视频监控、基于移动设备的视频会议等。

压缩感知(Compressed Sensing, CS)^[2]是近年在信号处理领域得到广泛重视的新兴理论体系。该理论指出,若某个待采样信号是稀疏的(或者在某个变换域是稀疏的),那么我们从其有限的(远低于原始信号维数)不相关测量值中准确恢复出原始信号的可能性极高。该理论突破了奈奎斯特采样定理的限制,使得信号获取的复杂度大大降低。CS 的测量过程运算量较低,而重构过程则需要迭代求解最优化问题,复杂度相对较高。因此,CS 理论结构特别适合各类分布式应用。

为了结合 CS 与 DVC 的优点,已经有一些学者将 CS 理论应用于 DVC 框架,设计分布式视频压缩感知(Distributed Compressed Video Sensing, DCVS)框架。其中,Kang 等人在解码端采用 GPSR(Gradient Projection for Sparse Reconstruction)方法对关键帧进行独立重构,采用非关键帧与边信息(Side Information, SI)间的相关性改进非关键帧的 GPSR 重构过程^[3]。但是,Kang 的方法没能很好地利用时域相关性,因此信号的重构质量不高。Do 等人采用已重构关键帧中的空域相邻块对待重构非关键帧中的块进行稀疏表示,以提高 SI 的准确性^[4]。在 Do 的方法中,SI 的生成需要块信号的测量值,而一般来说,帧级别测量比块级别测量的性能更好;另外,SI 的生成需要逐块求解 l_1 最小化问题,复杂度很高。文献[4-5]都在重构端对非关键帧进行预测,并对预测值进行测量,最后对测量残差值进行重构。在预测准确的前提下,残差信号比原始信号在变换域下更为稀疏,因此残差信号的重构误差在很大概率上要小于原始信号的重构误差。但是,文献[5]首先对非关键帧进行独立重构,其次以已重构关键帧为参考进行 ME/MC,然后再次对非关键帧做残差重构。为了保证重构质量,上述过程还需迭代 n 次(一般取 5 次),可见解码复杂度很高,并不实用。

为了提高 DCVS 方案的重构信号质量,本文提出了一种基于残差重构的 DCVS 方案。该方案利用相邻关键帧迭代进行 1/4 精度的 ME/MC 操作以保证 SI 的准确性;对 SI 进行测量,并对测量残差值进行全变分最小化(Total Variation Minimization, TVmin)重构。与 Kang 的方案相比^[3],本文方案没有过多增加解码端负担,却较大幅度地提升了非关键帧的重

构质量。

2 压缩感知理论

CS 理论^[2]的出现,突破了香农采样定理的瓶颈,降低了对传感器分辨率的要求,使得超分辨率信号获取成为可能。具体地,假设 $\mathbf{x}$ 是 n 维原始数据, Φ 是 $m \times n$ 维测量矩阵($m \ll n$), $\mathbf{y}$ 是 m 维测量数据。投影的过程可以采用公式描述如下:

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

若原始信号 $\mathbf{x}$ 足够稀疏,则可以将信号重构过程转换为一个最小 l_0 范数的最优化问题:

$$\hat{\mathbf{x}} = \arg \min \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{y} = \Phi \mathbf{x} \quad (2)$$

若 $\mathbf{x}$ 本身不稀疏,但是在某种基 Ψ 下稀疏,可以通过求解以下 l_0 范数的最小化问题来重构基 Ψ 下的原始信号:

$$\hat{\mathbf{s}} = \arg \min \|\mathbf{s}\|_0 \text{ s.t. } \mathbf{y} = \Theta \mathbf{s} \quad (3)$$

其中, $\Theta = \Phi \Psi$, $\mathbf{x} = \Psi \mathbf{s}$ 。

在 CS 理论中,传感器并不是直接获取信号 $\mathbf{x}$ 本身,而是获取其在测量基 Φ 上的投影值。对于测量矩阵 Φ 而言,希望其与稀疏矩阵 Ψ 不相干,这样所需的测量数为

$$m > c \cdot k \cdot \ln n \quad (4)$$

其中, k 是 $\mathbf{x}$ 在 Ψ 下的非零值个数, c 是常数。

比较普遍的方法是选取 Φ 为随机矩阵,因为随机矩阵与任何稀疏基都能以极大的概率不相干。但是,采用此类矩阵需要较高的内存开销和计算量,所以并不实用。Gan 等人提出了结构化的随机矩阵(Structurally random matrix, SRM)^[6],在该方法中,首先对输入信号进行随机置乱,其次进行简单、高效的变换(例如 DCT 或哈达玛变换),最后做随机下采样操作。测量值的产生仅需要少量的计算和内存开销,并不需要存储庞大的测量矩阵。鉴于该方法的良好性能,在本方案中编码端采用 SRM 方法进行测量。

由于求解 l_0 范数的最小化问题是 NP 难,目前已有许多替代解法,其中最著名的方法是基追踪(Basic Pursuit, BP)算法,该方法将最小化 l_0 范数问题转化为求解最小化 l_1 范数问题,即

$$\hat{\mathbf{x}} = \arg \min \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{y} = \Phi \mathbf{x} \quad (5)$$

求解 BP 问题的复杂度较高,当信号的维度较高时(例如图像与视频信号),信号重构过程所需时间过长,使得 CS 的实用性降低。为提高信号的重构速度,出现了基于梯度下降的方法(如 GPSR^[7]),并出现了许多贪婪算法,如正交匹配追踪(Orthogonal Matching Pursuit, OMP)^[8]等。

对于图像/视频信号而言,可以采用 TVmin 来取代 l_1 最小化,从而更好地反映图像/视频信号的特征。TVmin 由 Candes 等人引入 CS 中作为图像/视频信号的重构方法^[2]。经典的求解 TV 最小化的方法是采用内点法迭代求解 log-barrier 问题,但是需要较高的复杂度,并不实用。Li 将 TV 最小化问题转化为 Augmented Lagrangian 问题的形式,并将该问题拆成两个子问题交替求解,大大简化了解题过程^[9],该方法被命名为 TVAL3。本文采用 TVAL3 算法进行信号重构。

3 基于残差重构的 DCVS 方案

3.1 整体结构

在编码端,将视频序列划分为若干图像组(Group of Pictures, GOP),每个 GOP 包含一个关键帧和若干非关键帧,对关键帧和非关键帧均独立采用 SRM^[6]进行 CS 测量,关键帧的采样率需大于非关键帧的采样率。

解码端的重构总体流程如图 1 所示。首先重构每个 GOP 中的关键帧——对关键帧的测量值独立采用 TVAL3 算法^[9]求解 TVmin 问题;其次,采用相邻已重构关键帧经过 1/4 精度的迭代 ME/MC 操作生成非关键帧 SI 值 $\tilde{x}$;接下来,利用非关键帧的测量值及 SI 逐帧进行残差重构。

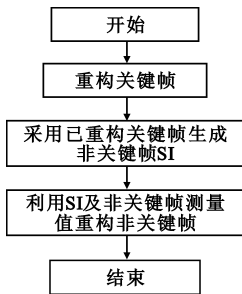


图 1 解码端图像重构总流程

Fig.1 Overall algorithm for decoder side image reconstruction

非关键帧的残差重构过程如图 2 所示,具体包括如下步骤:

- 步骤 1:对 SI 值 $\tilde{x}$ 进行 SRM 测量,得到 SI 的测量值 $\tilde{y}$;
- 步骤 2:计算测量残差值 $y_r = y - \tilde{y}$;
- 步骤 3:对 y_r 采用 TVAL3 算法^[9]求解 TVmin 问题,得到重构结果 $\hat{x}_r$;
- 步骤 4:将重构残差 $\hat{x}_r$ 与非关键帧预测值 $\tilde{x}$ 相加,得到非关键帧重构值 $\hat{x}$ 。



图 2 非关键帧残差重构流程

Fig.2 Residual reconstruction algorithm for non-key frames

3.2 残差重构

残差重构是本文提出的 DCVS 结构的核心,下面详细对该方法的流程和性能进行分析。

假设原始信号为 x ,采用测量矩阵 Φ 进行测量,得到测量值 y 。在本文算法中,并不直接采用测量值 y 进行重构,而是采用残差重构的方法。具体地,假设在重构端有待重构信号的预测值 $\tilde{x}$,对预测值 $\tilde{x}$ 进行测量,得到

$$\tilde{y} = \Phi \tilde{x} \tag{6}$$

其次,求实际测量值与预测测量残差值:

$$y_r = y - \tilde{y} = \Phi(x - \tilde{x}) = \Phi x_r \tag{7}$$

从公式(7)可知, y_r 实际上就是原始帧与预测值的残差 x_r 做随机投影的结果。假设 $\hat{x}_r$ 为从 y_r 中恢复出来的残差信号,则可以通过下式获取原信号 x 的重构值:

$$\hat{x} = \hat{x}_r + \tilde{x} \tag{8}$$

根据上述流程,原始信号 x 的重构误差为

$$\|x - \hat{x}\|_2 = \|(\tilde{x} + x_r) - (\tilde{x} + \hat{x}_r)\|_2 = \|x_r - \hat{x}_r\|_2 \tag{9}$$

根据式(9)可知,在残差重构算法中,原始信号 x 的重构误差由预测残差 x_r 直接决定。另一方面,若 $\tilde{x}$ 与 x 足够接近,则 x_r 定会比原信号 x 更为稀疏。因此,采用 CS 重构算法从 y_r 中恢复 x_r 会比从 y 中恢复 x 的误差更小。由此可以推断出,残差重构算法可以比直接重构算法获得更好的重构质量。

3.3 边信息生成

根据 3.2 节的分析,预测值 $\tilde{x}$ 的准确性是残差重构算法性能的决定因素。在本文提出的 DCVS 结构中,预测值 $\tilde{x}$ 就是非关键帧的 SI,SI 越准确,原始图像与其残差越小,因此残差重构的效果也越好。

为了保证 SI 准确性,在 Do 的方法中^[4],需要逐块求解 l_1 最小化问题以获取当前预测块的稀疏表示;而在 Mun 的方法中^[5],则需要先直接重构非关

键帧,接着多次迭代进行“ME/MC-残差重构”操作。两种 SI 生成方法的复杂度都过高,并不实用。

类似于 DVC^[1],可以通过在关键帧之间进行 ME,并对运动矢量调整后进行 MC 得到 SI。在本文方案中,首先对关键帧进行 6-tap FIR 滤波器插值得到 1/2 像素精度图像,继而进行双线性插值得到 1/4 像素精度图像;之后在前后两个关键帧之间双向迭代进行 1/4 像素精度的 ME/MC 操作。图 3 给出了 GOP 大小为 4 时的双向迭代 ME/MC 结构。

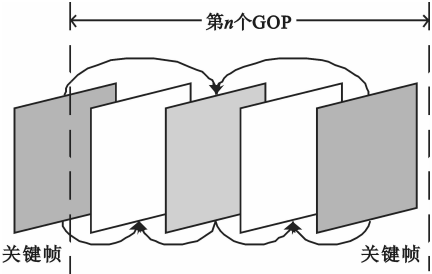


图 3 双向迭代 ME/MC 结构
Fig.3 Bi-direction iterative ME/MC structure

4 实验结果及讨论

为了测试算法性能,在 MATLAB 平台上仿真了本文方案,并与 Kang 的方法^[3]进行比较。GOP 大小设定为 4,关键帧的采样率固定为 0.7,对于非关键帧,分别采用 0.1、0.2、0.3、0.4 和 0.5 5 种采样率。用于信号测量的 SRM 选用哈达玛变换;用于信号重构的 TVL3 算法中的 μ 选为 2^{12} , β 选为 2^6 ,外循环门限设为 10^{-6} ,内循环门限设为 10^{-3} ,最大迭代次数设为 150。采用的测试序列为 CIF 格式的“foreman”与“football”,帧率为 30 frame/s。“foreman”序列含有较慢的运动,相对较平缓;相比之下,“football”序列细节信息丰富,包含快速运动。

表 1 列出了两种方法在 5 个采样率下的非关键帧重构图像的平均峰值信噪比(Peak Signal to Noise Ratio, PSNR),图 4 和图 5 则更直观地显示出了非关键帧重构图像 PSNR 值随采样率变化情况。

表 1 非关键帧重构质量比较				
Table 1 Reconstruction quality comparison for non-key frames				
采样率	foreman		football	
	文献[3]	本文方法	文献[3]	本文方法
0.1	28.30	31.17	21.86	24.95
0.2	28.96	34.16	22.77	27.25
0.3	29.97	36.48	24.04	29.93
0.4	31.21	38.57	25.56	32.26
0.5	33.29	40.52	26.49	34.88

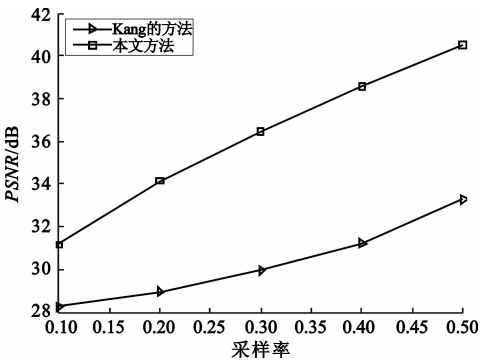


图 4 “foreman”序列 PSNR 比较
Fig.4 PSNR comparison for sequence “foreman”

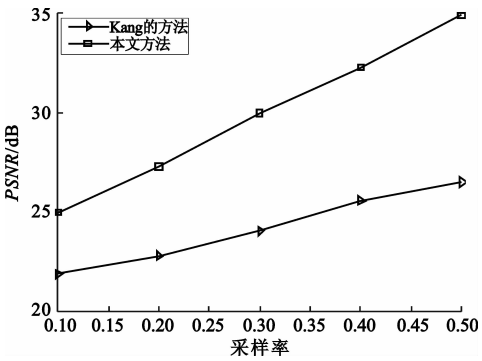


图 5 “football”序列 PSNR 比较
Fig.5 PSNR comparison for sequence “football”

从表 1 和图 4、图 5 中可以总结出,首先,在相同的采样率下,本文方法比 Kang 的方法有较高的 PSNR 提升;随着采样率的升高,本文方法 PSNR 提升值逐渐增大。例如,对于“foreman”序列,在 0.1 的采样率下,本文方法可以获得 2.87 dB 的 PSNR 增益;当采样率提升到 0.5 时,该值提升到 7.23 dB。其次,当序列中含有快速运动和丰富细节信息时,本文方法的非关键帧重构图像质量下降,主要原因有二:第一,丰富细节信息导致关键帧重构质量下降;第二,快速运动导致生成的 SI 准确性降低。

图 6 和图 7 分别展示了“foreman”序列与“football”序列的 SI 与残差值(取绝对值显示)。由于“foreman”序列运动缓慢,因此 SI 较为准确,图 6 中的残差值很小。在此情况下,采用 TVL3 算法求解 TVmin 问题,整帧图像的离散梯度值较稀疏,因此可断定重构效果较好,该分析结论也与表 1 的实验结果吻合。相比之下,“football”序列包含快速运动,生成的 SI 失真较大,图 7(b)中的残差图像梯度值不够稀疏,因此求解 TVmin 问题的效果明显差于“foreman”序列。

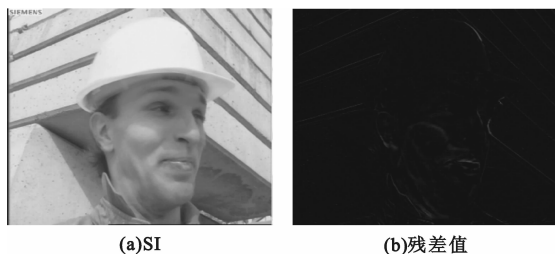


图 6 “foreman”序列第 6 帧 SI 与残差值
Fig.6 SI and residual value for the 6th frame of sequence “foreman”

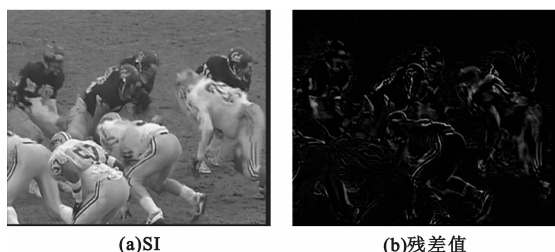


图 7 “football”序列第 6 帧 SI 与残差值
Fig.7 SI and residual value for the 6th frame of sequence “football”

5 结束语

本文将残差重构方式引入 DCVS 方案,并且综合应用了双向迭代的 SI 生成方法,充分利用了视频信号的时域相关性。依据实验结果可知,在 SI 准确的前提下,残差重构方案可以较大幅度地提升重构信号质量,从而证明了残差重构方案的有效性。需要注意的是,本文方案未考虑测量值的量化和熵编码问题。虽然量化和熵编码可以有效降低输出数据量,但是量化对非线性重构结果的影响还需要进一步地研究和分析。因此,下一步的工作重点是将量化和熵编码合理、有效地引入到提出的 DCVS 方案中,以进一步推进 DCVS 方案的实用化。

参考文献:

- [1] Guillemot C, Pereira F, Torres L, et al. Distributed monoview and multiview video coding: basics, problems and recent advances [J]. IEEE Signal Processing Magazine, 2007, 24(5):67–76.
- [2] Candes E, 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.
- [3] KANG Li-wei, LU Chun-shien. Distributed compressive video sensing [C]// Proceedings of 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing. Taipei, Taiwan: IEEE, 2009: 1169–1172.
- [4] Do T, Chen Yi, Nguyen D, et al. Distributed compressed video sensing [C]// Proceedings of 2009 IEEE International

Conference on Image Processing. Cairo, Egypt: IEEE, 2009: 1393–1396.

- [5] Mun S, Fowler J. Residual reconstruction for block-based compressed sensing of video [C]// Proceedings of 2011 Data Compression Conference. Snowbird, Utah, USA: IEEE, 2011:183–192.
- [6] Do T, Can Lu, Nguyen N, et al. Fast and Efficient Compressive Sensing Using Structurally Random Matrices [J]. IEEE Transactions on Signal Processing, 2012, 60(1):139–154.
- [7] Figueiredo M, Nowak R, Wrigth S. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems[J]. IEEE Journal of Selected Topics in Signal Processing, 2007, 1(4):586–597.
- [8] Tropp J, Gilbert A. Signal recovery from random measurements via orthogonal matching pursuit[J]. IEEE Transactions on Information Theory, 2007, 53(12):4655–4666.
- [9] LI Cheng-bo. An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing[D]. Houston: Rice University, 2009.

作者简介:



常侃(1983—),男,广西南宁人,2010年于北京邮电大学获博士学位,现为广西大学计算机与电子信息学院副教授,主要研究方向为压缩感知、视频编码与传输;

CHANG Kan was born in Nanning, Guangxi Zhuang Autonomous Region, in 1983. He received the Ph. D. degree from Beijing University of Posts and Telecommunication in 2010, and is now an associate professor of School of Computer and Electronic Information, Guangxi University. His research interests include compressed sensing, video coding and transmission.

Email: pandack0619@163.com

覃团发(1966—),男,广西宾阳人,1997年于南京大学获博士学位,现为广西大学计算机与电子信息学院副院长、教授、中国电子学会高级会员、中国通信学会高级会员,主要研究方向为无线多媒体通信、网络编码、视频编码和图像检索;

QIN Tuan-fa was born in Binyang, Guangxi Zhuang Autonomous Region, in 1966. He received the Ph.D. degree from Nanjing University in 1997. He is now a professor and vice Dean of School of Computer and Electronic Information, Guangxi University. He is also the senior member of China Institute of Electronics and China Communications Institute. His research interests include wireless multimedia communications, network coding, video encoding and image retrieval.

Email: tfqin@gxu.edu.cn

唐振华(1979—),男,广西玉林人,2009年于华中科技大学获博士学位,现为广西大学计算机与电子信息学院副教授,主要研究方向为无线多媒体通信。

TANG Zhen-hua was born in Yulin, Guangxi Zhuang Autonomous Region, in 1979. He received the Ph. D. degree from Huazhong University of Science and Technology in 2009, and is now an associate professor of School of Computer and Electronic Information, Guangxi University. His research concerns wireless multimedia communications.

Email: tangedward@126.com