

# A Survey on Distributed Compressed Video Sensing\*

Chen Xie, Shengrong Gong<sup>#</sup>

School of Computer Science & Technology, Soochow University, Suzhou  
Email: xwadec@163.com, <sup>#</sup>shrgong@suda.edu.cn

Received: Oct. 29<sup>th</sup>, 2012; revised: Nov. 8<sup>th</sup>, 2012; accepted: Nov. 16<sup>th</sup>, 2012

**Abstract:** This electronic distributed video coding is a new paradigm for video compression. Compared to conventional video coding standards in which the video sequence is coded jointly and decoded jointly, distributed video coding system codes the video sequence separately for two or more sources that are independent identically distributed and decodes jointly with the statistical correlation between different sources, then the coder becomes as simple as possible, so as to solve the problems of the limited video terminal. On the other hand, an emerging signal acquisition technology (Compressive Sensing, CS) provides a new way for the signal sampling, signal compression reconstruction based on the sparsity of signal, random measurement matrix and nonlinear optimization algorithm. It broke through the limitations of traditional Nyquist sampling theorem, which has been applicable to directly capture compressed image data efficiently. Combination of distributed video coding and CS (Distributed Compressed Video Sensing, DCVS) results in more low-complexity and low-cost for video coding. This paper reviews the theory of distributed video coding, classic schemes involved, as well as theoretical knowledge of compressed sensing, and development status of distributed compressed video sensing at this stage. Finally we present some problems and the probably corresponding solutions, then discuss its possible applications in the future prospects.

**Keywords:** Distributed Video Coding; Compressive Sensing; Dictionary Learning; Sparse Representation; Signal Reconstruction

## 分布式压缩视频感知综述\*

解 晨, 龚声蓉<sup>#</sup>

苏州大学计算机科学与技术学院, 苏州  
Email: xwadec@163.com, <sup>#</sup>shrgong@suda.edu.cn

收稿日期: 2012年10月29日; 修回日期: 2012年11月8日; 录用日期: 2012年11月16日

**摘 要:** 本分布式视频编码(Distributed video coding, DVC)是一种全新的视频编解码技术。与传统的联合编码、联合解码不同, 分布式视频编码对两个或多个独立同分布的信源进行独立编码, 然后由单一解码器利用信源之间的相关性对所有编码的信源进行联合解码, 故分布式视频编码使得低复杂度编码成为可能, 从而很好地解决了视频编码终端受限的情况。另一方面, 压缩感知(Compressive Sensing, CS)为信号采样提供了新的方式, 它基于信号的稀疏性、测量矩阵的随机性和非线性优化算法对信号的压缩测量重构, 从而突破了传统 Nyquist 采样定理的限制。压缩感知理论应用于分布式视频编码, 能够使编码端更加低复杂化、低消耗、简单化。本文主要综述了分布式视频编码相关理论、当前的经典方案, 以及压缩感知理论和现阶段分布式压缩视频感知的发展现状、涉及到关键方法。最后分析了现存方案的一些问题和思路, 并讨论了其未来可能的发展方向。

**关键词:** 分布式视频编码; 压缩感知; 字典学习; 稀疏表示; 信号重构

\*基金项目: 基于 DVC 的无线网络视频传输错误隐藏技术研究(SYG201116); 苏州市应用基础研究计划资助。

<sup>#</sup>通讯作者。

## 1. 引言

近年来,伴随着无线通信技术的迅速发展以及无线网络带宽的不断提高,在移动环境下提供多媒体通信业务成为可能。许多新的视频应用,如移动可视电话、无线视频监视系统、无线 PC 摄像机、无线、低功耗的多媒体传感器网络等不断涌现,加之野外军事通信、灾后紧急救援等特定场合视频传输已经越来越迫切,例如,抗震救灾时将前线受灾面貌和路貌等视频实时传送到后方以供后方决策等,使得这些应用对视频终端有了新的要求:

需要对视频进行实时编码传输;编码端计算力、存储容量有限,难以执行复杂的传统编码;从客户端的无线化、便携式及电池容量等角度考虑,要求编码端视频压缩算法必须是低复杂度、低功耗的;由于采用无线接入与传输,要求视频信号传输具有强健壮性。

这些特点与传统的视频编码系统相比有明显的不同,传统的视频编解码技术在编码端普遍采用运动补偿预测,充分利用了视频序列帧间的时域相关性进行预测编码,成功地解决了视频的存储和传输,很好地应用于视频广播、视频点播等领域。但由于在编码端执行变换、量化、熵编码以及计算量较大的运动估计与运动补偿等操作,从而占用了大量的资源,导致其编码端复杂度通常是解码端的 5~10 倍。因而,编码器与解码器相比,要求具有更高的计算能力、存储容量等。此外,传统视频编码算法在编码端采用预测编码模式,网络传输导致的丢包或比特错误会造成编码端和解码端重构缓存的不匹配,从而出现误码扩散现象,容错能力较差。因此,传统视频编码标准主要适用于服务器端的视频信号处理,即视频内容经一次高效编码压缩后传输给广大低端用户下载和解码播放,如广播系统、流媒体点播系统等“一次编码、多次解码”类型的应用场合。

但是,移动可视电话、无线视频监视系统、无线 PC 摄像机等这些新的应用领域中,通常编码设备比较简单,而解码设备拥有较多的资源可以进行复杂的计算,这与传统视频编码标准适用的场景正好相反。传统的视频编码技术难以适应这类新的应用。为了更好地适用于这类新型无线网络的视频应用,要求视频编码系统必须满足以下几个方面:

- 1) 低时延;
- 2) 编码端复杂度较低;
- 3) 较高的编码效率,能够适应于带宽受限的视频传输;
- 4) 鲁棒视频传输。

针对上述情况,一种全新的视频编码框架——分布式视频编码,即 DVC(Distributed Video Coding)被提出并得到快速发展。

分布式视频编码是指对两个或多个独立同分布的信源进行独立编码,然后由单一解码器利用信源之间的相关性对所有编码的信源进行联合解码,它使得低复杂度编码成为可能,从而很好地解决了视频编码终端受限的情况。DVC 基于 Slepian 和 Wolf(1973)提出的无损分布式编码理论以及 Wyner 和 Ziv(1976)提出的使用译码端辅助信息的有损分布式编码理论。基于这两个理论,使得对两个统计相关的视频序列进行分布式编码(独立编码,联合译码)成为可能。其中编码端对源信号进行独立编码,不进行或只进行非常简单的运动估计;解码端利用视频序列的时域、空域等相关性进行联合解码,将复杂的运动预测、补偿技术从编码端移到了解码端,即将编码器运算复杂度转移到解码器。分布式视频编码可以取得与传统的预测编码机制(联合编解码)非常接近的编码效率。因此,分布式视频编码具有编码简单,解码较复杂,并且能够实现较为高效的压缩,抗误码特性好的特点<sup>[1]</sup>。

然而,在一些比较苛刻的资源约束下,视频编码要求完全的低复杂度、低消耗、简单化并且实时性更强,常规的分布式视频编码就不太适用,因为视频序列的获取还依赖于高消耗的机制,例如采样、压缩,不但让视频中的原始数据有所损失,而且效率也比较低。压缩感知(Compressive Sensing)的出现完美的解决了这样的问题。自然界的大部分信号是有冗余度的,或者说可压缩的,如果能够直接获取压缩后的数据并且能够完美地重建信号,那么意义将是非凡的。Candès 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<sup>[2]</sup>。Candès 和 Donoho 在相关研究基础上于 2006 年正式提出了压缩感知的概念<sup>[1,3]</sup>。它建立在矩阵分析、统计概率论、拓扑几何、优化与运筹学、泛函分析与时频分析等基础上,其核心思想是将压缩与采样合并进行:首先将采集的信号非自适应线性投影(测量

值),然后根据相应重构算法由测量值近乎完美重构原始信号<sup>[1]</sup>。压缩感知的优点在于信号的投影测量数据量远远小于传统采样方法所获的数据量,突破了香农采样定理的瓶颈。CS 理论避开了高速采样,意味着信号的采样与处理都可以以非常低的速率进行,这将显著降低数据存储和传输代价,以及信号处理时间和计算成本,给信号处理领域带来新的冲击。

DCVS(Distributed Compressive Video Sensing)结合了分布式视频编码中的分布式信源对立编码和联合解码的思想,建立的一种全新的编码体系,具有更低复杂度的编码端,以及重构图像质量高,鲁棒抗噪声能力强等特点,同时适用于实时、非实时的应用,比分布式视频编码有着更广泛的应用。本文主要阐述了分布式视频编码的基本理论以及常用的典型编码框架,然后引入压缩感知理论,并对分布式压缩视频感知国内外研究现状作了全面的分析,包括涉及的关键技术和压缩视频感知的应用等,最后对分布式压缩视频感知存在的问题提出了相应的思路,对其应用前景作了展望。

## 2. 分布式视频编码(DVC)基本思想

分布式视频编码是一种全新的视频压缩编码框架,它是 20 建立在 Slepian 和 Wolf 分布式无损编码理论基础上的<sup>[4]</sup>。Wyner 和 Ziv 进一步提出了 Wyner-Ziv 有失真信源编码理论<sup>[5]</sup>,为分布式信源编码奠定了理论基础。但是,由于缺乏具体的实现方法,DVC 的发展较为缓慢。直到近年来,无线移动视频和无线传感网络的广泛发展,DVC 重新成为研究热点。不同于传统的视频编码方法,分布式视频编码在译码端产生边信息来利用当前帧与边信息的相关性进行帧间译码,降低了编码端的复杂度。本节主要阐述分布式视频编码的理论基础及典型编码框架。

### 2.1. 分布式无损信源编码

分布式压缩是指编码两个独立的任意序列;每个具有分离的编码器,每一个编码器发送一个独立的码流到一个独立的解码器,解码器联合解码所有的码流并且计算统计相关,如图 1 所示。

分布式压缩是指编码两个独立的任意序列;每个具有分离的编码器,每一个编码器发送一个独立的码

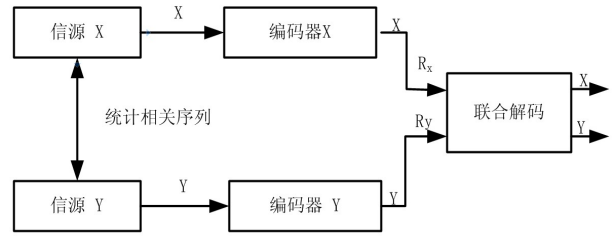


Figure 1. Distributed source coding  
图 1. 分布式信源编码

流到一个独立的解码器,解码器联合解码所有的码流并且计算统计相关,如图 1 所示。

假设两个统计相关独立同分布的无限长随机序列  $X$  和  $Y$ , 传统的熵编码和解码可以达到  $R_y \geq H(Y)$  和  $R_x \geq H(X)$  的码率,  $H(X)$  和  $H(Y)$  为  $X$  和  $Y$  的熵。然而联合解码具有更好的率失真性能(独立编码)。Slepian-Wolf<sup>[4]</sup>理论证明了互相独立的对两路信号  $X$  和  $Y$  编码,总码率  $R$  可以达到联合熵  $H(X, Y)$ , 我们建立码率区:

$$\begin{aligned} R_x &\geq H(X|Y), R_y \geq H(Y|X) \\ R_x + R_y &\geq H(X, Y) \end{aligned} \quad (1)$$

由此可以发现,  $R_x + R_y$  可以达到联合信息熵  $H(X, Y)$ 。

在图 2 中的  $N$  点,对  $X$  编码的码率  $R_x = H(X)$ ,而对  $Y$ ,进行压缩时所需要的码率仅为  $R_y = H(Y|X)$ 。同样在  $M$  点,对  $Y$  编码的码率为,  $R_y = H(Y)$  而对  $X$  进行压缩时所需要的码率仅为  $R_x = H(X|Y)$ 。这就是在解码端具有边信息的无损信源编码问题的理论限。

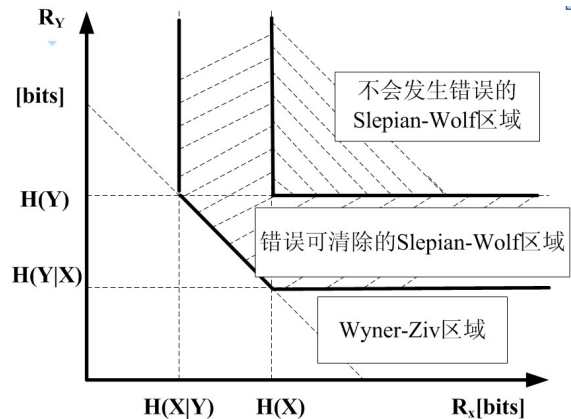


Figure 2. The reaching rate area of independent identical distribution source  $X$  and  $Y$   
图 2. 独立同分布信源  $X$  和  $Y$  的分布式编码可达速率区

### 2.2. 有损分布式编码的 Wyner-Ziv 理论

Wyner 与 Ziv 在此基础上提出了边信息辅助解码的分布式信源编码方案<sup>[5]</sup>, 无论边信息  $Y$  是仅仅作用在解码端, 还是同时作用在编码端与解码端, 对  $X$  序列均方率失真性能的影响是相同的。Slepian-Wolf 理论及 Wyner-Ziv 理论为分布式视频编码的发展奠定了理论基础, 其中无损分布式信源编码通常被称为“Slepian-Wolf 编码”, 在 Slepian-Wolf 编码前加上量化器而构成 Wyner-Ziv 编码器。

### 2.3. 当前分布式视频编码系统的经典方案

在分布式视频编码的各种方案中, 以 Bernd Girod 等提出的算法最具代表性, 像素域 Wyner-ziv 视频编码框如图 1 所示<sup>[6-8]</sup>, 编码端采用了两种编码方式: 一种是 Wyner-ziv 帧, 图像信号均匀量化后, 进行基于 turbo 码的 Slepian-Wolf 编码, 其目的是保持视频信号的基本轮廓(见图 3); 另一种关键帧编码采用的是传统视频编码中的帧内编码(如 H.263 的帧内编码)作为关键帧(Key Frame)。

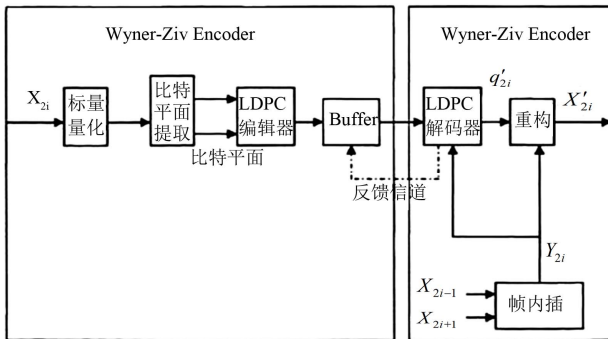


Figure 3. Distributed video coding based on pixel domain  
图 3. 基于像素域的分布式视频编码

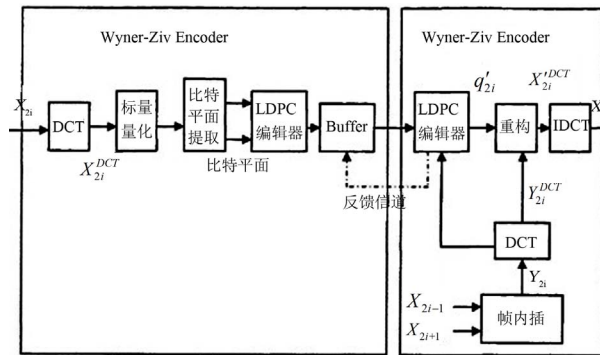


Figure 4. Distributed video coding based on transform domain<sup>[10]</sup>  
图 4. 基于变换域的分布式视频编码<sup>[10]</sup>

变换域的分布式视频编码<sup>[9]</sup>(见图 4)的主要过程跟像素域的编码过程基本一致, 还是关键帧  $K$  使用传统的帧内编解码方式对其处理; 两个关键帧间的 Wyner-Ziv 帧则采用帧内编码、帧间解码的技术对其进行处理。但在对 Wyner-Ziv 帧进行 DCT 变换, 然后将 DCT 量化系数按位平面分层送至 Slepian-Wolf 编码器进行熵编码。DCT 变换域分布式视频编码的系统结构比基于像素域的分布式视频编解码复杂, 但其率失真性能也要比像素域的方法高 2~2.5 dB<sup>[9]</sup>。

### 3. 压缩感知基本理论

Candès 和 Donoho 在相关研究基础上于 2006 年正式提出了压缩感知的概念<sup>[1-3]</sup>, 即假设  $x$  是一个  $N$  维数字信号, 如果它是  $K$ -稀疏的(即有  $K$  个非零元, 且  $K \ll N$ ), 已知某一个测量矩阵  $\Phi \in R^{M \times N}$  ( $M \ll N$ ) 以及某未知信号  $x$  在该矩阵下的线性测量值  $y \in R^M$

$$y = \Phi x \quad (2)$$

方程(2)也可以看作原信号  $x$  在  $\Phi$  下的线性投影, 现在考虑由  $y$  重构  $x$ 。很显然, 由于  $x$  的维数远远大于  $y$  的维数, 方程(2)有无穷多个解, 是一个欠定方程, 很难重构原始信号。然而, 由于原始信号  $x$  是  $K$  稀疏的, 并且  $y$  与  $\Phi$  满足一定条件, 理论证明, 信号  $x$  可以由测量值  $y$  通过求解最优  $\ell_0$  范数问题精确重构<sup>[2]</sup>:

$$\hat{x} = \arg \min \|x\|_0 \quad s.t. \quad \Phi x = y \quad (3)$$

式中,  $\|x\|_0$  为向量  $\ell_0$  范数, 表示向量  $x$  中非零元素的个数, Candès 等指出, 如果要精确重构  $K$  稀疏信号  $x$ , 测量次数  $M$ (即  $y$  的维数)必须满足:

$$M = O(K \ln(N)) \quad (4)$$

并且矩阵  $\Phi$  必须满足有限等距性条件<sup>[10,11]</sup>。

常见的自然信号在时域内几乎都是不稀疏的, 因此, 上述信号重构过程不能直接应用于自然信号的重构。信号稀疏表示理论指出, 自然信号可以通过某种变换  $\Psi$  进行稀疏表示, 即  $f = \Psi x$ ,  $x$  为信号  $f$  在  $\Psi$  变换域的稀疏表示。考虑测量公式  $y = \Phi f$ , 并且  $f$  是可以稀疏表示的,  $f = \Psi x$ , 且令  $A^{CS} = \Phi \Psi$  则有

$$y = \Phi f = \Phi \Psi x = A^{CS} x \quad (5)$$

通过以上的分析可以看到, 压缩感应的核心思想与传统信号采样方法对原始信号  $x$  先采样后压缩不

同，它由少量线性测量通过求解最优化问题得到  $x$ ，突破了 Nyquist 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。

#### 4. 分布式压缩视频感知(DCVS)

针对压缩感知的特点，即根据视频帧的稀疏性只需要少量的测量采样、量化、编码，在视频的解码端就可以利用某一特定算法进行近乎完美的重建。而分布式视频编码则强调低复杂度的编码端，以及传输的鲁棒性。所以，压缩感知和分布式视频编码的结合就显得很自然。其基本框架如图 5。

类似于传统的分布式视频编码，将视频帧进行等份分组，一般为 2 帧或者 4 帧一组，每帧组的首帧称为 Key 帧，作为参考帧，用作预测本帧组内其余的帧，或者产生边信息帧。Key 帧由 H.264 编码标准的帧内编码，帧内解码进行重构帧图像，其余帧称为 CS (Compressive Sensing)帧，CS 帧用压缩感知理论进行图像重构，编码端仅仅是负责视频帧图像信号的线性投影测量采样。得到的测量值经过量化熵编码传送到解码端，运动估计与运动补偿的边信息生成，及运算量较大的图像稀疏重建算法求解过程都转移到解码端。

##### 4.1. DCVS 研究现状

分布式压缩视频感知框架的首次提出是 2009 年，Thong T. Do 在基于分布式视频编码的传统框架下融入了压缩感知理论<sup>[10]</sup>，Key 帧利用传统的帧内编码，帧内解码进行重建画面，CS 帧分别进行基于块的测量和基于整帧的测量，保存基于整帧的测量值  $y_f$ ，利用基于块的测量值  $y_b$  稀疏重建得到近似的 CS 帧  $\hat{x}$ ，将整帧的测量值  $y_f$  与近似生成的 CS 帧的测量值  $\hat{y} = \Phi\hat{x}$  作差得到  $y_r = y_f - \hat{y}$ ，稀疏重建得到预测误差  $x_r$ ，然后与近似 CS 帧相加得  $x_{rec} = \hat{x} + x_r$ ， $x_{rec}$  是精度更高的重建近似帧。其框架如图 6。

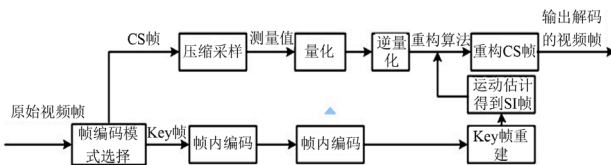


Figure 5. Basic DCVS codec framework  
图 5. DCVS 基本编解码框架

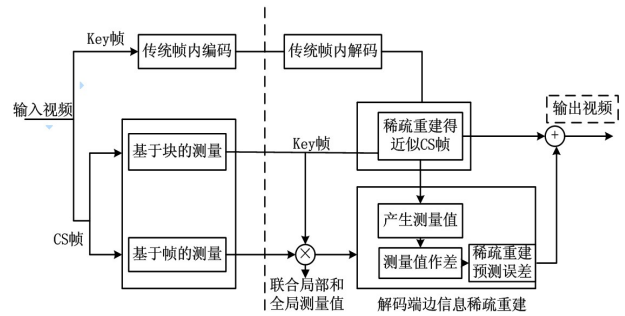


Figure 6. The proposed DCVS codec framework by Thong  
图 6. Thong 提出的 DCVS 编解码框架

同年，Josep Prades 提出的基于压缩感知的分布式视频编码<sup>[11]</sup>，总体思想与 Thong 一致，而处理关键帧和 CS 帧的细节有所不同，同样利用待解码图像块对应位置 Key 帧周边的图像块建立局部字典，并且加入了量化机制。在待解码块重建时加入了跳跃模式和独立模式，以及反馈信道机制有效的缩短了图像块的重构时间和提高了图像重建精度。但是由于采用传统的编码方式及反馈方式，对于编码端的负荷并没有过多降低。并且不适合离线视频编码。

Li-Wei Kang 也提出另一种分布式压缩视频感知框架<sup>[12]</sup>，他将 Key 帧和 CS 帧都采用 CS 技术进行测量采样和重构，根据联合稀疏模型理论分析，Key 帧采用较高的测量比率(指的是测量值数目,  $M$  与信号维数  $N$  的比值，即  $M/N$ )，保证画面的重构质量。由于与帧间有较高的相关性，CS 帧用较低的测量比率测量采样。且用改进的 GPSR(Gradient Projection for Sparse Reconstruction)算法与运动估计内插得到的边信息帧共同重建 CS 帧图像，能够取得很不错的画面重构效果，他们科研小组之后的研究中，Hung-Wei Chen 对<sup>[12]</sup>作了适当的改进<sup>[13]</sup>，加入了自适应的字典学习手段，同时对前后相邻关键帧及运动估计内插得到的边信息帧用 K-SVD 算法<sup>[14]</sup>学习并建立完备字典。用迭代阈值法 SpaRSA<sup>[15]</sup>进行稀疏重建。相比之前 DCVS 编码系统，画面重构效果有较大进步，其框架如图 7。

Huang-Wei Chen 优化了基于字典学习的 DCVS 框架<sup>[16]</sup>，提出了一种动态分配测量比率的优化算法，进一步提高了算法重构的效率，实现了性能的优化，但只是基于 CS 帧的动态测量，同时，由于加入了反馈机制，其适用场合具有很大的局限性。

北京邮电大学 Xun Wang, Cong Ma 分别在 DCVS

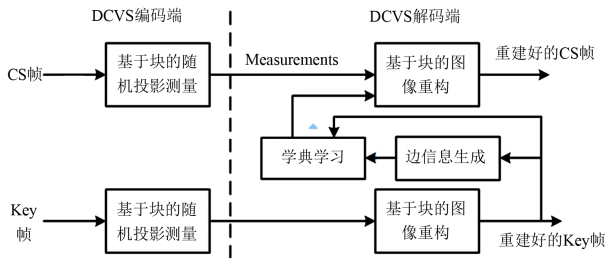


Figure 7. DCVS framework based on dictionary learning  
图 7. 基于字典学习的 DCVS 编码框架

框架中加入了稀疏滤波相关模型和部分已知支撑的平滑  $\ell_0$  范式重构算法<sup>[17,18]</sup>, 提高了算法的效率和画面的重建效果。台湾大学的 Hsiao-Yun Tseng 引入了改进的边信息生成算法, 对边信息和原始帧之间的建立拉普拉斯分布, 采用基于置信传播的压缩感知重构算法, 进一步对 DCVS 框架作了性能优化<sup>[19]</sup>。

#### 4.2. 联合稀疏模型

D. Baron 在 2009 年提出了分布式压缩感知理论<sup>[20]</sup>, 建立了联合稀疏模型。联合稀疏模型为充分利用两个时域相邻的两帧之间的相关性提供了思路。假设两个解码帧为  $x_t$  和  $x_{t+1}$ , 正常情况下,  $x_t$  和  $x_{t+1}$  含有共有的部分和各自私有的部分, 可以用以下式子表示:

$$\begin{aligned} x_t &= x_C + x_{t\_U} \\ x_{t+1} &= x_C + x_{t+1\_U} \end{aligned} \quad (6)$$

其中  $x_C$  为  $x_t$  和  $x_{t+1}$  的共有部分, 而  $x_{t\_U}$ ,  $x_{t+1\_U}$  为两者的私有部分, 如果将  $x_t$  作为  $x_{t+1}$  的一个参考帧, 对  $x_{t+1}$  进行编解码, 在传统的视频编码中, 编码端利用运动估计找出  $x_C$  并加以保存, 然后对  $x_{t+1}$  与  $x_C$  的差值进行编码传送, 因此, 对于帧  $x_t$  和  $x_{t+1}$  的压缩即可以用对  $x_t$  与  $x_{t+1\_U}$  的压缩来代替。假设  $\Psi$  为一个稀疏基矩阵, 能够将  $x_t$  和  $x_{t+1\_U}$  稀疏表示, 如下式:

$$\begin{aligned} x_t &= \psi \theta_t, \|\theta_t\|_0 = K_t, \\ x_{t+1\_U} &= \psi \theta_{t+1\_U}, \|\theta_{t+1\_U}\|_0 = K_{t+1\_U} \end{aligned} \quad (7)$$

其中  $\theta_t$  和  $\theta_{t+1\_U}$  是  $x_t$  和  $x_{t+1\_U}$  的稀疏表示, 通常情况下  $K_t \geq K_{t+1\_U}$ , 根据公式 (3) 我们可以得知  $M_t \geq M_{t+1\_U}$ , 其中  $M_t$  和  $M_{t+1\_U}$  是  $x_t$  和  $x_{t+1\_U}$  的测量值数目, 然而在 DCVS 编码系统中, 因为编码端低复杂度的约束条件下不可能在编码端获取到  $x_{t+1\_U}$ , 但

能确定的是,  $x_{t+1}$  的测量值数目要小于  $x_t$  的测量值数目。因此在 DCVS 编码系统中如果关键帧和非关键帧 (CS) 帧都采用压缩感知理论进行测量重构, 那么我们在编码端进行线性投影测量时, CS 帧的测量比率就可以小于 Key 帧的测量比率, 根据少量的测量值和已解码帧生成的边信息就可以重构精度不差于 Key 帧的 CS 帧画面。

#### 4.3. 图像帧信号的稀疏表示

信号的稀疏表示就是将信号投影到正交变换基时, 绝大部分变换系数的绝对值很小, 所得到的变换系数是稀疏或者近似稀疏的, 可以将其看作原始信号的一种简洁表达<sup>[21]</sup>, 这是压缩感知的先决条件, 即信号必须在某种变换下可以稀疏表示。

视频帧信号的稀疏表示可以看成图像信号的稀疏表示, 自然图像信号都是存在稀疏性的, 即可压缩性, 在通常情况下将二维图像信号向量转化为一维向量, 然后可以根据信号本身的特点灵活选取, 常用的有 DCT 基、FFT 基、DWT 基<sup>[22]</sup>、Curvelet 基<sup>[23]</sup>等。根据调和理论, 一个长度为  $N$  的一维离散时间信号  $f$ , 可以表示为一组标准正交基的线性组合

$$f = \sum_{i=1}^N x_i \psi_i \quad \text{or} \quad f = \Psi x \quad (8)$$

$\psi_i$  是列向量,  $\Psi = [\psi_1, \psi_2, \dots, \psi_N]$  为正交基矩阵, 展开系数中  $x = [x_1, x_2, \dots, x_N]^T$ , 系数向量  $x$  是  $K$ -稀疏的。稀疏表示不仅仅是局限于一维信号, 视频中单帧画面就可以利用稀疏基产生二维的稀疏表示系数, 还有文献<sup>[24]</sup>引入了 3D 小波作为变换基, 即将一组视频帧看成 3D 图像, 时域看成第三维, 帧间相关性得到充分挖掘。使用 3D 小波基的重建效果较好, 但是帧组体积太大, 计算量较为庞大。

实际上, 对于信号的稀疏表示不一定是在某个正交基上稀疏, 在冗余字典下的稀疏分解在最近几年变为一个研究热点, 这是一种全新的信号表示理论: 用超完备的冗余函数库取代基函数, 称之为冗余字典, 即字典中的元素被称为原子, 形如:

$$D = \{d_t \in \mathbb{R}^N \mid \|d_t\| = 1, 1 \leq t \leq L\} \quad (9)$$

$D$  中的每个原子为  $N$  维单位长度向量  $d_t$ ,  $L$  为字典原子个数, 如果字典  $D$  能够张成  $N$  维欧氏空间  $\mathbb{R}^N$ ,

则称  $D$  是完备的字典, 当原子个数  $L > N$  时, 则字典是冗余的, 如果同时还能张成  $N$  维欧氏空间  $R^N$ , 则称字典  $D$  是超完备的。

字典的选择应尽可能好地符合被逼近信号的结构, 其构成可以没有任何限制。从冗余字典中找到具有最佳线性组合的  $K$  项原子来表示一个信号, 能够代替基函数, 使得信号更加稀疏化。但因为字典是冗余的, 所以需要的测量数会有所增多。

在 DCVS 当前帧的冗余字典稀疏表示中, 视频帧的压缩采样一般是进行分块处理, 将图像划分成互不重叠的面积相等的正方形小块, 冗余字典的构造方式主要两种: 第一种是通过直接获取相邻的已解码帧或者运动估计与补偿得到边信息帧中位置相对应的周边块, 其中每一块相当于字典中一个原子, 调整选择适当的数目, 然后将每个原子作归一化处理, 即可组成一个简单冗余字典, 如图 8。

第二种冗余字典的构造方式是通过字典学习的方式来构造自适应的冗余字典, 自适应冗余字典设计的思路是通过字典学习算法获得更符合信号内容特征, 或者纹理信息的原子。自适应冗余字典可以随着不同的输入信号做出调整<sup>[14,25-28]</sup>, 常用的字典学习的算法有 MOD 算法<sup>[27]</sup>、RLS-DLA<sup>[25]</sup>、OLDL<sup>[28]</sup>、SLO-DLA<sup>[26]</sup>、K-SVD 算法<sup>[14]</sup>等, 这里可以最常见的 K-SVD 算法对相邻已解码关键帧, 及边信息帧中的所有块信息进行学习, 学习出适当数量的原子, 文献[13]实验表明通过学习得到完备字典比稀疏正交基和非学习建立的字典具有更好的稀疏表示性能, 并且在相同重建算法的情况下, 画面重建的质量效果也更加的好。

然而事实上, 自然复杂图像或者运动目标较多图像的向量表示的稀疏性往往并不理想, 而这些图像的在其本身梯度上较为稀疏, Candès 在 2006 年从自然图像的梯度出发提出了一种最小梯度全变分算法重建图像(又称 TV 算法)<sup>[29]</sup>, 如公式(10):

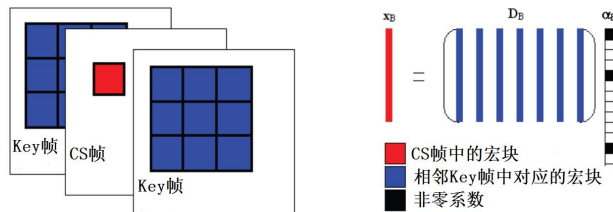


Figure 8. Sparse representation by dictionary building through adjacent blocks  
图 8. 直接获取邻块建立冗余字典的稀疏表示

$$\hat{x} = \arg \min \|x\|_{TV} \quad \text{s.t. } \Phi x = y \quad (10)$$

$$\text{其中 } \|x\|_{TV} = \sum_{i,j} \sqrt{(X_{i+1,j} - X_{i,j})^2 + (X_{i,j+1} - X_{i,j})^2},$$

看成图像梯度的近似  $\ell_1$  范式。<sup>[30]</sup>即用 TV 算法重建基于块的图像/视频信号, 结果表明, 最小梯度全变差算法的图像重建效果较好, 但计算复杂度非常高, 从计算量来比较, TV 算法并不适合实际应用的场合。

#### 4.4. 测量矩阵及测量策略

##### 4.4.1. 测量矩阵

测量矩阵在测量向量获取和信号重建的过程中起着关键性的作用, 它的合理选择对于信号重构所需的测量值数目、重构算法的算法和重构的质量都存在着内在的联系。如果测量矩阵列向量的非线性相关性越强, 矩阵元素稀疏度越高, 将更有利于信号的重构。Candès 和 Tao 给出并证明了测量矩阵中必须满足有限等距条件<sup>[31]</sup>。即对于任意  $K$  稀疏信号和常数  $\delta k \in (0, 1)$ , 如果

$$(1 - \delta_k) \|c\|_2^2 \leq \|\Phi_T c\|_2^2 \leq (1 + \delta_k) \|c\|_2^2 \quad \forall c \in R^{|T|} \quad (11)$$

成立, 其中  $T \in \{1, \dots, N\}$ ,  $|T| \leq K$ ,  $\Phi_T$  为  $\Phi$  中由索引  $T$  所指示的相关列构成的大小为  $K \times |T|$  的子矩阵, 则称矩阵  $\Phi$  满足有限等距性质 (Restricted Isometry Property, RIP)。

在一般的基于压缩感知的视频编码中, 测量矩阵通常采用随机高斯矩阵, 文献[32]证明了随机高斯矩阵能满足有限等距性质。我们可以选择一个大小为  $M \times N$  的随机高斯矩阵作为测量矩阵, 其中每一个值都满足  $N(0, 1/N)$  的独立正态分布, 高斯测量矩阵的优点在于它几乎与任意稀疏信号都不相关, 因而所需的测量次数最小。其他常见的能满足约束等距性的测量矩阵还包括一致球矩阵、二值随机矩阵、局部傅里叶矩阵、局部哈达玛矩阵以及托普利兹矩阵等<sup>[32]</sup>。然而, 二维图像信号转化为一维向量时, 往往维数过大, 采用的随机高斯矩阵或其他的一些固定生成的矩阵就变的异常巨大。当 DCVS 系统应用实际场合时, 就需要巨大存储空间用来存储测量矩阵, 就违背了编码资源有限的原则。为了解决这类问题, Thong 等人提出一种结构化随机矩阵<sup>[33]</sup>, 该矩阵具有几乎与所有其他正交矩阵(单位矩阵和极度稀疏矩阵除外)不相关的特点, 可以分解成定点、结构化分块对角矩阵与随

机置换向量与伯努利向量点积的形式，该矩阵可以看成是随机高斯、伯努利矩阵和局部傅里叶变换矩阵的混合模型，并保持了各自的特点，并且在存储器只需少量的内存空间保存相应的投影机制就可以完成大维数信号的随机投影测量。因此该矩阵被证明具有硬件友好、内存消耗低、以及计算速度快等优点。Thong在结构随机矩阵基础上提出了混沌哈达玛集合矩阵(Scrambled Hadamard Ensemble, SBHE)<sup>[34]</sup>，将结构化随机矩阵进一步优化。绝大多数的DCVS的编码系统中都采用SBHE矩阵。

#### 4.4.2. 测量策略

测量策略的实施直接关系到DCVS编码端当前帧的测量向量的获取，以及之后重构算法的重构精度等，对于视频画面的恢复有着极其重要的意义。确定好测量矩阵后，视频帧组中Key帧往往采用传统的帧内编解码的方式来保证Key帧画面的解码质量。而CS帧即采用压缩感知理论进行测量重构，通常将帧图像分成互不重叠的正方形小块，在考虑了画面效果和运算速度的均衡的前提下，一般取边长为16或32，文献[35]对图像进行分块处理的好处作了阐述：因为每个分块采用的测量矩阵尺寸很小能够有效地降低内存消耗，并且编码端无需等待整个图像被测量投影后才将测量向量发送出去，处理的效率更高。文献[36]证明了分块处理后相同图像重构效果下所需要的测量值要小于直接对整幅图像测量所需要的测量值。

由于相邻两帧相对应帧块之间存在着空域相关性，文献[37]利用了这种相关性，将帧差法的思想引入编码端，采用相同测量矩阵的前提下，即对每一个CS帧块的测量值与其参考帧Key帧的测量值的差值进行编码传送。因为<sup>[38]</sup>证明了当测量值数目 $M > \log N$ 时，可以用测量值间的相似性代替原信号向量间的相似性。

整幅帧画面内容中存在感兴趣区域和不感兴趣区域，这是由人眼的生物视觉特征来决定的，如果只是从像素域投影测量，就没有综合考虑到感兴趣区域带来的影响，文献[39]在压缩视频感知中结合视觉注意模型，对于感兴趣区域采用较高的测量比率进行独立的采样重构画面，而对于其他非感兴趣区域中就采用传统视频编码中的帧差法压缩采样，由于视频帧间的空域相关性，残差帧一般比较稀疏，所需要的测量

值就会相应降低，实验也证明了在相同测量比率的情况下，结合视觉注意模型的编码系统的画面重构视觉效果要好于一般的压缩视频感知编码系统。文献[40]则是在小波变换的频域下利用压缩感知理论进行重构图像，利用多尺度小波变换，对于不同分级的高频和低频区域采用分尺度加权方法确定相应测量比率，达到了一种均衡分配测量值的效果，是一种新颖的且效果较好的测量策略。

不同种类的视频帧的画面复杂度也有所不同，如何动态的对视频帧的测量值分配策略依然是现在的一个难题，文献[16]采用的动态测量的方法依据公式 $M = O(K \ln(N/k))$ ，即图像的稀疏度与测量值数目息息相关，是通过相邻已解码的关键帧块进行稀疏分解来估计当前帧的稀疏度，然后求得分解得到的稀疏系数的方差占整个画面总方差和的比重，然后按照这个比重反馈给编码端动态分配测量值数目，公式如下：

$$M_{(t+1),i} = \frac{V_{(t+1,i)}}{\sum_{j=1}^B V_{(t+1,j)}} \times (MR_{CS} \times N) \quad (12)$$

其中 $MR$ 为CS预先设定好该帧的平均测量比率， $N$ 为方块的大小， $v_{t+1,i}$ 为 $t+1$ 帧中第 $i$ 块的预测方差。文献<sup>[11-13,16-19]</sup>提出了利用类似于分布式视频编码的不断反馈修正的机制，主要通过重构近似画面的测量值与初始测量值差值的方差进行阈值处理，如果大于阈值就会通过反馈信道再传送更多的测量值，直到方差值小于阈值则停止迭代。文献<sup>[41]</sup>则是通过比较非关键帧和关键帧之间的测量值的差异性来对图像块分类，依据分类的不同采用不同的动态测量策略。文献<sup>[42]</sup>采用了边缘检测来估计视频帧的稀疏度，用边缘的重构错误率来估计整幅图像的重构精度。

#### 4.3. 帧的稀疏重构

假设图像原始信号为 $f$ ，在某种变换或者字典下转换为 $X$ 稀疏或可压缩的前提下，求解最小10范数问题(4)：它需要列出 $X$ 中所有非零项位置的 $C_N^K$ 种可能的线性组合才能得到最优解。因此，求解式(8)的数值计算极不稳定而且是NP问题。这和稀疏分解问题从数学意义上讲是同样的问题，于是稀疏分解已有算法可以应用到CS重构中。研究人员已经提出一系列的



求解次优解的算法,有常见的最小  $\ell_1$  范数法、匹配追踪系列算法、迭代阈值法、以及基于贝叶斯理论的重构算法等。

上述算法都是针对单一独立的信号重构,如果应用在分布式压缩视频感知编码系统时,基于分布式信源联合解码的特征,就需要在考虑如何使重构算法充分利用边信息解码,达到当前帧只要很少的测量值能够有不输于关键帧的解码质量的目的。在文献[10,11]是利用已解码关键帧局部生成字典代替边信息的生成,省略了计算复杂的运动估计和运动补偿的操作,将生成的字典应用在求解最小  $\ell_1$  范数算法中。<sup>[12]</sup>则是利用运动估计和运动补偿生成的边信息帧来当作梯度投影法(GPSR)<sup>[43]</sup>的初始参数。加快了算法的运行速度,提高画面恢复的精度。<sup>[13,16]</sup>则是利用 K-SVD 算法对已解码关键帧和边信息帧动态学习出字典,将该字典融于迭代阈值法中,能够得到较好的图像重构效果。文献[17,19]采用了基于置信传播(Belief Propagation)的压缩感知重建算法,也是将边信息帧作为算法的一个初始参数。文献[41]采用了更适合二维图像的最小变分法(Total Variation)算法,虽然画面重建质量比较出色但是运算复杂度太过巨大。

上述图像重构时比较突出的问题有:1) 解码的计算复杂度太高,比如求解最小  $\ell_1$  范数的 BP 算法,最小全变分算法等。2) 由于  $\ell_1$  范数无法区分稀疏系数尺度的位置,尽管整体上重构信号在欧式距离上逼近原信号,但存在低尺度能力搬移到了高尺度的现象。容易出现一些人工效应。针对这种现象,文献[40,44]分别是基于多尺度压缩重建算法,有效的利用了不同尺度下的小波系数,还有 La 和 N do 提出的树形匹配追踪算法<sup>[45]</sup>,该算法针对 BP(Basic Pursuit)<sup>[46]</sup>,MP(Match Pursuit)<sup>[47]</sup>,OMP(Orthogonal Match Pursuit)<sup>[48]</sup>等追踪算法没有考虑多尺度分解稀疏信号的在各子带的位置关系,将信号的树形结构加以利用,进一步提高了图像重构的精度与速度。暂时还没有在分布式压缩视频感知编码系统中应用此类重构算法,这将是一个值得研究的点。

#### 4.6. 帧的稀疏重构

随着视频业务的不断促使了视频技术的不断发展,传统的视频编码标准在网络传输上有了新的挑战,网络带宽的波动极可能出现丢包,因此编码技术

不但要考虑如何使视频流适合网络的波动性,还需要考虑怎样克服信道中出现的差错和丢失。

可分级视频编码及多描述视频编码即为此类问题提供了解决方案,它们的主要思想是视频帧编码流分为多种描述,解码端收到的描述越多则视频的重构质量越好,多描述视频编码里的多种描述之间的地位是平等的,而可分级视频编码在多种描述里则有基本层及增强层,基本层的描述则不能丢失,是视频恢复的关键所在,增强层仅仅用于辅助基本层解码,在网络中丢失也影响不大,由于压缩感知本身的欠采样特点,即使丢失少量测量值也能重构出令人接受的画面,使得压缩视频感知在网络环境恶劣情况下应用有了可能。例如文献[49]将压缩感知应用于多描述视频编码、将视频帧的 DCT 系数分为两个随机相等数目的描述,即使丢失了其中一个描述也能重构出可以接受的画面质量效果,还有文献[50,51]将压缩感知与可分级视频编码结合,利用压缩感知降维随机投影的特点,基本层和增强层由于测量值数目少,体积占用低,容易在波动的网络带宽下完好保存,保证了信号在解码端画面重构质量。文献[52]则是在分布式压缩视频感知编码系统,动态的调整的 CS 帧测量值数目的变化以适应网络带宽的需求。

可以看出如果将多描述视频编码或可分级视频编码应用在 DCVS 中,能够拓宽 DCVS 的应用场合,且适应复杂多变的网络环境,更为贴合实际,很具有研究价值。

### 5. 分布式压缩视频感知(DCVS)展望

CS 在分布式视频编码上的应用有着巨大的潜在优势,文献[53]已对三种编码模式(传统的 H.264/AVC 编码模式,典型的 DVC 模式,以及 DCVS 模式)做了各项性能上的对比,DCVS 不但在帧画面的重构效果上与传统分布式视频编码有的竞争,而且在编码端框架上也更加的简洁,不需要反馈信道,实时和离线场合都能适用。另一方面,即使网络环境比较差的情况下,由于压缩感知本身的欠采样特点,即使丢失少量测量值也能重构出令人接受的画面,因此具有高鲁棒性,低编码复杂等优势特点。分布式压缩视频感知的研究主体部分依赖于压缩感知理论不断推进,随着性能更卓越的测量矩阵、及速度更快,重构效果更优异的稀疏重建算法的不断涌现,再结合传统视频编码

中的一些高效视频帧间帧内处理技术,相信分布式压缩视频感知技术会越来越成熟。但是由于分布式压缩视频感知技术才刚刚起步,一些简单的仿真结果表明了其可行性,但是具体应用时仍然有许多尚待解决的技术难题。具体来讲,分布式压缩视频感知技术需要在以下几个方面进一步展开研究并给予了一定的解决思路:

#### 1) 无反馈信道动态测量问题

对于 Key 帧和 CS 帧的同时自适应的测量比率的问题,因为不同画面的视频帧,往往需要的测量数值是不一样的,因为有的画面比较简单则需要很少的测量值,而一些复杂的画面则需要的较多的测量值,如果只是用同一种测量比率才进行采样,对于简单的图片往往是计算浪费,对于复杂的图片则会重建效果差。解决思路主要是在编码端作适当预处理,估计当前帧的稀疏度,当然同时要保持编码段的低复杂化和解码效率的平衡。

#### 2) 量化问题

没有适合实际应用的量化方案,因为在视频编解码实际传输过程中,量化、编码是必不可少的步骤。我们对于这一问题的解决思路是可以从对各种类型的视频中进行稀疏采样,来发现采样值的分布规律,来确定合适的量化矩阵方案,依据量化矩阵的方案,可以指定不同的熵编码。

#### 3) 编解码端结构优化问题

如何有效的利用视频帧之间的时域空域相关性,因为往往在一些自然的图像和视频帧中,画面信号的稀疏度并不理想,但是由于帧间或者帧间测量值的相关性,我们就可以通过某种方式(最简单的方式即帧差法)降低图像的稀疏度,结合压缩感知的特点,将可以使测量值进一步减少,达到极大化视频压缩的目的,或者图像本身利用梯度稀疏的特性,找到速度更快、效率更高的最小梯度重建算法。

#### 4) 压缩感知理论研究推进问题<sup>[13]</sup>

实验已证明画面在稀疏基表示的系数越稀疏,则相同重构算法的重构的画面精度越高,如何找到能够使画面最大化稀疏基是一个值得研究热点,自适应字典学习一个很好的解决方案,然后学习建立字典的过程往往较为低效,怎么选用合适的字典学习方法,或者找到性能与质量的平衡是一个难点。还需要研究效果更好,速度更快的信号重建算法,以及更为高性能

的,硬件友好的测量矩阵,能够让画面能够更快速的重建,重构的质量效果更好,这主要依据压缩感知相关技术的不断推进。

## 参考文献 (References)

- [1] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306.
- [2] E. Candès, J. Romberg and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 2006, 52(2): 489-509.
- [3] E. Candès. *Compressive sampling*. Proceedings of International Congress of Mathematicians, Madrid: European Mathematical Society Publishing House, 2006: 1433-1452.
- [4] J. D. Slepian, J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 1973, 21(4): 471-480.
- [5] A. Wyner, J. Ziv. The rate distortion function for source coding with side information at the decoder. *IEEE Transactions on Information Theory*, 1976, 22(1): 1-10.
- [6] A. Aaron, R. Zhang and B. Girod. Wyner-Ziv coding of motion video. *Conference Record of the Asilomar Conference on Signals, Systems and Computers, Pacific Grove*, 2007: 240-244.
- [7] A. Aaron, E. Setton and B. Girod. Toward practical Wyner-Ziv coding of video. *International Conference on Image Processing, Barcelona*, 2003: 869-872.
- [8] A. Aaron, S. Rane and R. Zhang. Wyner-Ziv coding for video: Applications to compression and error resilience. *Proceedings of the IEEE Data Compression Conference, Snowbird*, 2003: 93-102.
- [9] A. Aaron, S. Rane and B. Girod. Wyner-Ziv video coding with hash based motion compensation at the receiver. *The International Conference on Image Processing, Singapore*, 2004, 2: 3097-3100.
- [10] T. T. Do, Y. Chen, D. T. Nguyen and N. Nguyen. Distributed compressed video sensing. *Proceedings of the IEEE International Conference on Image, Baltimore*, 2009: 1393-1396.
- [11] J. Prades-Nebot, Y. Ma and T. Huang. Distributed video coding using compressive sampling. *Proceedings of the Picture Coding Symposium, Chicago*, 2009: 1-4.
- [12] L. Kang, C. Lu. Distributed compressive video sensing. *IEEE International Conference on Acoustics, Speech, and Signal Processing, Taipei*, 2009: 1169-1172.
- [13] H. W. Chen, L. W. Kang and C. S. Lu. Dictionary learning-based distributed compressive video sensing. *Proceedings of the Picture Coding Symposium, Nagoya*, 2010: 210-213.
- [14] 14M. Aharon, M. Elad and A. M. Bruckstein. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representations. *IEEE Transactions on Image Processing*, 2006, 54(11): 4311-4322.
- [15] 15S. J. Wright, R. D. Nowak and M. A. T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 2009, 57(7): 2479-2493.
- [16] H.-W. Chen, L.-W. Kang and C.-S. Lu. Dynamic measurement rate allocation for distributed compressive video sensing. *Proceedings of the SPIE—The International Society for Optical Engineering, Bellingham*, 2010.
- [17] X. Wang, H. Fang, X. Zhu, B. Li and Y. Liu. Sparse filter correlation model based joint reconstruction in distributed compressive video sensing. *IEEE International Conference on Network Infrastructure and Digital Content, Beijing*, 2010: 483-487.
- [18] C. Ma, Y. Liu, L. Zhang and X. Q. Zhu. Distributed compressive video sensing based on smoothed  $\ell^0$  norm with partially known support. *IEEE International Conference on Multimedia and Expo*, 2011: 11-15.

- [19] H.-Y. Tseng, Y.-C. Shen. Distributed video coding with compressive measurements. MM'11 Proceedings of the 19th ACM International Conference on Multimedia. New York, 2011: 1273-1276.
- [20] D. Baron, M. B. Wakin and M. Duarte. Distributed compressed sensing. <http://www.dsp.rice.edu/~drorb/pdf/DCS112005.pdf>
- [21] B. A. Olshausen, D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 1996, 381(6583): 607-609.
- [22] S. Mallat. *A wavelet tour of signal processing*. San Diego: Academic Press, 1996.
- [23] E. Candès, D. Donoho. Curvelets: A surprisingly effective non-adaptive representation for objects with edges. Technical Report 1999-28, Department of Statistics, Stanford: Stanford University, 1999.
- [24] M. Wakin, J. Laska, M. Duarte and D. Baron. Compressive imaging for video representation and coding. Proceedings of Picture Coding Symposium, Beijing, 2006.
- [25] K. Skretting, K. Engan. Recursive least squares dictionary learning algorithm. *IEEE Transactions on Signal Processing*, 2010, 58(4): 2121-2130.
- [26] H. Zayyani, M. Babaie-Zadeh. Thresholded smoothed-L0 (SL0) dictionary learning for sparse representations. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, 2009: 1825-1828.
- [27] M. Elad, M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 2006, 15(12): 3736-3745.
- [28] J. Mairal, F. Bach and J. Ponce. Online dictionary learning for sparse coding. *ICML '09 Proceedings of the 26th Annual International Conference on Machine Learning*, New York, 2009: 689-696.
- [29] E. Candès, J. Romberg. Robust signal recovery from incomplete observation. *IEEE International Conference on Image Processing*, Atlanta, 2006: 1281-1284.
- [30] Y. F. Zhang, S. L. Mei. A novel image/video coding method based on compressed sensing theory. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, 2008: 1361-1364.
- [31] E. Candès, T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 2005, 51(12): 4203-4215.
- [32] E. Candès, T. Tao. Near optimal signal recovery from random-projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 2006, 52(12): 5406-5425.
- [33] T. T. Do, T. D. Trany and L. Gan. Fast compressive sampling with structurally random matrices. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Washington DC, 2008: 3369-3372.
- [34] L. Gan, T. T. Do and T. D. Trany. Fast compressive imaging using scrambled block hadamard ensemble. *European Signal Processing Conference*, 2008.
- [35] G. Lu. Block compressed sensing of natural images. *International Conference on Digital Signal Processing*, Cardiff, 2007: 403-406.
- [36] H. Lee, H. Oh and S. Lee. A new block compressive sensing to control the number of measurements. *IEEE International Conference on Image Processing*, Brussels, 2011: 2713-2716.
- [37] J. Zheng, E. L. Jacobs. Video compressive sensing using spatial domain sparsity. *Optical Engineering*, the International Society for Optical Engineering, 2009, 48(8): 087006.
- [38] Z. L. Wang, I. Lee. A study of video coding by reusing compressive sensing measurements. Proceedings of the 7th International Conference on Ubiquitous Intelligence & Computing and Automatic & Trusted Computing, Xi'an, 2010: 64-69.
- [39] J. Xu, J. W. Ma. Compressive video sensing based on user attention model. *The 28th Picture Coding Symposium*, Nagoya, 2010: 90-93.
- [40] J. E. Fowler, S. Mun. Multiscale block compressed sensing with smoothed projected landweber reconstruction. Proceedings of the 19th European Signal Processing Conference. Barcelona, 2011: 564-568.
- [41] Z. R. Liu, V. Zhao. Block-based adaptive compressed sensing for video. *IEEE of the 17th International Conference on Image Processing*, Hong Kong, 2010: 1649-1652.
- [42] A. Masomeh, A. Ali. Compressed video sensing using adaptive sampling rate. *The 5th International Symposium on Telecommunication*, Tehran, 2010: 710-714.
- [43] M. A. T. Figueiredo, R. D. Nowak and S. J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(4): 586-597.
- [44] S. Dekel. Adaptive compressed image sensing based on wavelet-trees. <http://dsp.rice.edu/files/cs/adaptiveCSimag.pdf>
- [45] C. La, M. N. Do. Signal reconstruction using sparse tree representation. Proceedings of the International Society for Optical Engineering, San Diego, 2005, 5914: 273-283.
- [46] S. S. Chen, D. L. Donoho and M. A. Saund. Atomic decomposition by basis pursuit. *SIAM Review*, 2001, 43(1): 129-159.
- [47] R. Neff, A. Zakhor. Very low rate video coding based on matching pursuits. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, 7(1): 158-171.
- [48] J. A. Tropp, A. C. Gilbert. Signal recovery from partial information by orthogonal matching pursuit. 2005. [www-personal.umich.edu/~jtropp/papers/TG05-Signal-recovery.pdf](http://www-personal.umich.edu/~jtropp/papers/TG05-Signal-recovery.pdf)
- [49] Y. F. Zhang, S. L. Mei. A multiple description image/video coding method by compressed sensing theory. *IEEE International Symposium on Circuits and Systems*, Seattle, 2008: 1830-1833.
- [50] S. Y. Xiang, L. Cai. Scalable video coding with compressive sensing for wireless videocast. *IEEE International Conference on Communications*, Kyoto, 2011: 1-5.
- [51] H. Jiang, C. B. Li. Scalable video coding using compressive sensing. *Bell Labs Technical Journal*, 2012, 16(4): 149-169.
- [52] M. Mashud, K. Mahata. A scalable distributed video coder using compressed sensing. *Annual IEEE on India Conference*, Gujarat, 2009: 1-4.
- [53] N. Imran, B.-C. Seet and A. C. M. Fong. A comparative analysis of video codecs for multihop wireless video sensor networks. *MultiMedia Systems*, 2012, 18(5): 373-389.