

一种基于字典学习的压缩感知视频编解码模型

郭继昌 金卯亨嘉

(天津大学电子信息工程学院, 天津, 300072)

摘要: 无线多媒体传感器网络中针对视频信号处理的两个重要的问题是如何高效编解码和如何抵抗信道误码。结合压缩感知和字典学习理论, 提出了一种应用于无线传感器网络的基于字典学习的压缩感知视频编解码模型。模型整体采用压缩感知理论以降低编码端复杂度, 提高系统抗误码性。编码端应用差分编码和跳帧模式大大减少了信道传输数据量; 解码端采用字典学习算法增强图像的稀疏表示能力, 从而提高视频重构精度。本模型在实现高效编码的同时将计算复杂度从编码端转移到解码端, 从而满足编码端资源受限的应用场合。理论分析和仿真实验表明该模型可行并且有效。

关键词: 压缩感知; 字典学习; 视频编解码; 无线多媒体传感器网络

中图分类号: TN911.7 **文献标志码:** A

Dictionary Learning-Based Compressive Video Sensing Codec Model

Guo Jichang, Jin Maohengjia

(School of Electronic Information Engineering, Tianjin University, Tianjin, 300072, China)

Abstract: The problems of efficient codec and resiliency to channel errors are important in video processing of wireless multimedia sensor networks (WMSNs). Based on the compressed sensing (CS) and dictionary learning algorithm, a dictionary learning-based compressed video sensing codec model is proposed for the WMSNs. The model uses CS to reduce the complexity of encoder effectively and improve resiliency to channel errors. In the encoder, the application of difference structure and skip mode reduces the amount of data transmitted in the channel. And in the decoder, dictionary learning algorithm helps enhance images' sparse representation, thereby improve reconstructed video quality. The model switches the computational complexity from the encoder to the decoder and has high coding efficiency, so it can be applied to the resource-constrained embedded devices. The theory analysis and experiment results have verified the feasibility and efficiency of the model.

Key words: compressed sensing; dictionary learning; video codec; wireless multimedia sensor networks

引 言

近年来, 无线多媒体传感器网络(Wireless multimedia sensor networks, WMSNs)^[1] 作为一个研究热

点,被广泛应用于视频监控、环境监控以及目标跟踪等项目中。无线多媒体传感器网络是随着半导体技术的发展,在无线传感器网络(Wireless sensor network, WSN)中引入低功耗视频和音频传感器而形成的。因其使用的嵌入式设备和无线信道的局限性,此应用领域要求编码端具有低复杂度和低功耗以保证长期稳定工作,编码方式具有高抗信道误码性以保证接收端精确重构视频,而接收端可以进行大量的数据存储和复杂的解码计算。

然而,传统视频编码技术均采用编码端复杂、解码端简单的系统结构来去除时间和空间冗余,从而获得高压缩效率。这种结构使得整个系统对编码器的计算能力和内存容量的要求非常高,远高于解码器。同时,预测编码和变换编码使得编码视频帧之间具有很大的依赖性,若在信道传输过程中丢失关键帧或数据,会对重造成很大干扰。因此,传统的视频编码方式并不适用于无线多媒体传感器网络。

压缩感知(Compressed sensing, CS)^[2-4]是近年来信号处理领域提出的一种理论。该理论不需要复杂的数据编码计算,而是将其转移到解码端去完成,因此非常适用于编码端资源受限的无线多媒体传感器网络中。文献[5]提出一种基于运动检测的低复杂度视频编码算法。Cossalter 等人在运动估计中结合压缩感知提出了一种联合视频编解码模型^[6]。另一些学者则将压缩感知和其他编码方式相结合,提出了多种分布式视频压缩感知理论(Distributed compressed video sensing, DCVS)^[7-9]。然而其关键帧与非关键帧采用单独编码、联合解码的方式,因此编码端复杂度依旧较高,且系统的硬件实现有一定的难度。为了提高重构质量,一些模型^[8-9]采用已重构帧训练稀疏字典以提高压缩感知重构精度。

本文提出了基于字典学习的压缩感知视频编解码模型(Dictionary learning-based compressed video sensing, DL-CVS)。本模型将视频序列分为关键帧(I帧)和非关键帧(P帧),均采用压缩感知进行编解码以降低编码端计算复杂度,提高系统抗误码性能,为了减少信道传输数据量,对I帧和P帧的具体编码方式不同。与其他方法相比,DL-CVS在编码端加入了差分编码和解码模式判断,从而减少信道传输数据量;在解码端采用字典学习算法增强图像的稀疏表示能力,从而提高视频重构精度。

1 压缩感知和字典学习基本原理

1.1 压缩感知

考虑一个实值的有限长一维离散信号 $x \in \mathbf{R}^{N \times 2}$, 假设 $\{\psi_i, i = 1, \dots, N\}$ 是 \mathbf{R}^N 的一组向量基, 则 \mathbf{R}^N 空间的任何信号 x 可以被线性表示为

$$x = \sum_{i=1}^N \psi_i \alpha_i = \Psi \alpha \quad (1)$$

式中: $\Psi = [\psi_1, \psi_2, \dots, \psi_N]$ 是 $N \times N$ 的稀疏字典, α 是 x 在 Ψ 域的稀疏系数, 满足 $\alpha_i = (x, \psi_i)$ 。若 $\|\alpha\|_0 = K$ 且 $K \ll N$, 则称 α 是 K 项稀疏的, 或称 x 在 Ψ 域是 K 项稀疏的。

压缩感知理论指出, 当信号 x 具有稀疏性时, 可以线性投影得到远小于奈奎斯特采样定理所要求的测量值 $M = O(K \log(N/K))$, 其中 $K < M \ll N$, 并通过凸优化等算法高精度地重构出原始信号。其具体过程为: 设 y 是一个长度为 M 的测量矢量, 满足 $y = \Phi x$ 。其中, Φ 是一个 $M \times N$ 的测量矩阵, 则 y 可以用 α 表示为

$$y = \Phi \Psi \alpha = A \alpha \quad (2)$$

式中 $A = \Phi \Psi$ 。文献[3, 4]表明如果测量矩阵 Φ 设计恰当, 那么信号 x 可以通过求解最小范 l_0 数问题从测量值 y 中精确重构出来

$$\min \|\alpha\|_0 \quad \text{s. t.} \quad y = \Phi \Psi \alpha = A \alpha \quad (3)$$

目前常用的重构算法主要有松弛算法、贪婪追踪法和组合算法, 各有其优缺点。

1.2 稀疏字典与字典学习

信号 x 具有稀疏性是压缩感知的重要前提和理论基础,因此压缩感知理论的首要研究任务为稀疏字典 Ψ 的设计。合适的稀疏字典可保证信号的稀疏表示系数具有足够的稀疏性和衰减性,从而在压缩测量的同时保证信号的重构精度。

稀疏字典主要分为正交基字典、紧框架字典和过完备字典。其中过完备字典通过增加原子个数来增强信号逼近的灵活性,具有最强的稀疏表示能力。目前过完备字典主要有手工构造和字典学习两大构造方法。手工构造出的稀疏字典具有固定的结构,而字典学习方法是从数据中动态学习得到最优稀疏表示,使得字典中原子尺度和特性更接近于需要表示的图像信号。目前已涌现出众多的字典学习算法,其中广泛应用的有最优方向算法(Method of optimal directions, MOD)^[10]和 K-奇异值分解算法(K-singular value decomposition, K-SVD)^[11]。

已知一组样本 $x=[x_1, x_2, \dots, x_n]$,字典学习的目标就是要找到一个字典 D 和稀疏矩阵 Γ 使得表示误差最小

$$\arg \min_{D, \Gamma} \|X - D\Gamma\|_F^2 \quad \text{s. t.} \quad \|\gamma_i\|_0 \leq T \quad \forall i \quad (4)$$

式中 γ_i 是 Γ 的第 i 列。此优化问题是一个组合的且高度非凸的问题,因此最多只能找到局部最小值。MOD算法和 K-SVD算法均交替使用稀疏编码和字典更新来完成式(4)的计算。与 MOD相比, K-SVD具有更好的效果。

2 DL-CVS 编解码模型结构

压缩感知理论可以显著降低编码复杂度、减少编码端数据量和对传输带宽的需求等,适合大数据量、低复杂度编码要求的应用场合。基于以上特点,本文提出了如图 1 所示的视频编解码模型 DL-CVS。

在编码端,根据重构精度和实时性的要求,将视频序列划分为两类。一类为关键帧(I帧),另一类为非关键帧(P帧)。本模型中,每两帧组成一组,即图像组(Group of pictures, GOP)为 2,其中奇数帧为 I 帧,紧随其后的偶数帧为该组的 P 帧。对于 I 帧,选择适当的测量矩阵对其进行测量。对于 P 帧,先与对应 I 帧进行差分,判断解码模式,再使用相同的测量矩阵对差值进行测量。

和编码端相比,解码端的复杂度较高。首先根据编码端发送的判断信号确定解码模式,之后对该 GOP 的两帧进行相应的解码操作。若为跳帧模式,则只重构 I 帧, P 帧用 I 帧重构结果替代;若为普通模式, I 帧和差值分别进行重构,再通过图像重组得到 P 帧的重构结果;若为字典学习模式,则先按照普通模式的步骤进行重构,再根据 I 帧的重构结果进行字典更新。最后按照帧序和帧率输出重构视频流。

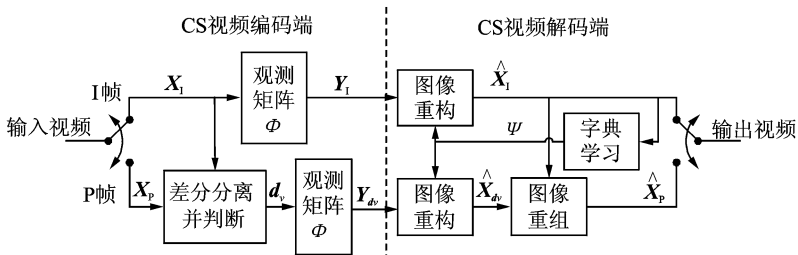


图 1 DL-CVS 系统

Fig. 1 Block diagram of DL-CVS

2.1 DL-CVS 编码端

编码端主要包括以下三个过程:差分编码、解码模式判断和压缩测量。视频序列相邻两帧具有较强的相似性,因此其差值图像较原始图像具有更大的稀疏性,能够使用更少的量化等级和采样比特率来表示^[12]。为降低信道传输数据量,本模型对每个 GOP 中的 P 帧与相应 I 帧进行差分运算,即

$$\mathbf{d}_v = \mathbf{X}_p - \mathbf{X}_I \quad (5)$$

为了更大程度地降低信道传输数据量和编解码复杂度,提高视频解码精度,在编码端引入解码模式判断,G 包括:跳帧、普通和字典学习 3 种模式。判断依据为 \mathbf{d}_v 的均方误差值 (Mean square error, MSE)。若 MSE 小于阈值下限 $thrL$,则判定这两帧非常相似,该 P 帧无需重构,可直接使用 I 帧重构结果作为其重构结果;若 MSE 大于阈值上限 $thrH$,表明这两帧差别很大,其拍摄场景或对象发生了较大改变,应更新字典以适应新的场景,因此解码端在完成此 GOP 内两帧的重构后更新字典;若 MSE 在阈值范围内,则直接对两帧分别进行重构。

在 DL-CVS 模型中,对每帧均采用分块压缩感知。即先将帧 \mathbf{X} 分为 N 个 $b \times b$ 的不重叠块 $\mathbf{b}_i, i=1, 2, \dots, N$,再将每个块 \mathbf{b}_i 看作一个长度为 $N_b = b \times b$ 的列向量,至此帧 \mathbf{X} 被转化为 $N_b \times N$ 的矩阵。对其进行差分分离和判断,并根据 $\mathbf{Y} = \Phi \mathbf{X}$ 对其进行测量,最后将测量值和模式判断信号传给解码端。

测量矩阵 Φ 采用随机块哈达玛矩阵 (Scrambled block Hadamard ensemble, SBHE)^[13]。SBHE 具有以下五个优点:接近最优的性能、和大多数现有的稀疏字典 Ψ 不相关、计算速度快、高效存储和方便硬件实现。因此,使用 SBHE 作为测量矩阵 Φ 可有效降低编码端所需内存,提高系统的编码效率。 Φ 的计算公式为

$$\Phi = \mathbf{Q}_M \mathbf{W} \mathbf{P}_N \quad (6)$$

$$\mathbf{W} \begin{bmatrix} \mathbf{W}_B & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \mathbf{W}_B \end{bmatrix} \quad (7)$$

式中: \mathbf{W} 为一个 $N \times N$ 块对角阵; \mathbf{W}_B 表示 $B \times B$ 的哈达玛矩阵; \mathbf{P}_N 是一个随机算子,对 \mathbf{W} 的 N 列重新排序; \mathbf{Q}_M 也是一个算子,随机地从 $\mathbf{W} \mathbf{P}_N$ 中挑 M 出行。这种块对角矩阵 \mathbf{W} 能快速并行计算,复杂度为 $O(N \log B)$,并且只需很小的存储单元。本模型中 \mathbf{W}_B 使用的是 32×32 的哈达玛矩阵。

2.2 DL-CVS 解码端

在解码端,I 帧和差值的压缩测量值首先通过重构算法解码,得到 I 帧的重构结果和差值的重构信号值。之后,因为 I 帧和差值使用相同的测量矩阵进行压缩测量、相同的稀疏字典和重构算法进行解码,因此将 I 帧的重构结果与差值的重构信号值相加即可得到 P 帧的重构图像,即

$$\hat{\mathbf{X}}_p = \hat{\mathbf{X}}_I + \hat{\mathbf{X}}_v \quad (8)$$

若编码端传来更新字典的信号,则在重构完这两帧图像后,使用 I 帧重构结果作为训练信号,根据字典学习算法更新字典。最后按照帧序和帧率整合成重构视频并输出。

在 DL-CVS 模型中,重构算法选择平滑 l_0 范数 (Smoothed norm l_0 , SL0) 算法^[14]。SL0 算法具有匹配度高、重构时间短、计算量低等优点,最适合视频编解码应用。

2.3 字典学习方法

如果图像稀疏字典的原子可以由其自身得到,那么使用该字典可实现图像的最稀疏表示。然而,解码端无法直接得到该图像原始信号来训练稀疏字典对其进行重构,因此,鉴于视频序列连续帧之间存在很大的相似性,可使用已重构的相邻帧作为训练集进行字典学习,更新稀疏字典,进而用其对目标帧进

行重构。

为满足视频重构的精度和实时性,应用于视频编解码模型的字典学习算法应具有简单、快速和稀疏表示能力强的特点。DL-CVS模型采用先从单帧内提取训练信号,然后用K-SVD算法进行字典更新的字典学习方法。具体过程为:首先,将1帧重构图像分为 $b \times b$ 的不重叠块,针对每个不重叠块,提取它本身及距离其最近的4个重叠块,将这些块看作长度为 $N_b = b \times b$ 的列向量,组成训练信号 \mathbf{x} ;其次,用K-SVD算法训练出字典 \mathbf{D} , $\mathbf{D} \in \mathbf{R}^{N \times P}$, $N \ll P$ 。

在本字典学习方法中,从单帧内提取训练信号可简化解码端结构,降低运算复杂度和重构时延。重叠块的提取可提供一定的运动补偿,对运动较剧烈视频序列的重构精度也有较大提高。K-SVD算法的稀疏编码阶段采用系数复用正交匹配追踪算法(Coefficient reuse orthogonal match pursuit, CoefROMP)^[15],CoefROMP使用前一次迭代得到的稀疏系数中最大的 $K/3$ (K 为稀疏度)个系数来进行初始化,之后再逐渐找到剩余的 $2K/3$ 个系数。因为有了 $K/3$ 个初始系数,该算法可以有效地提高字典学习效率,降低字典学习时间。

CoefROMP核心算法步骤为:

输入:字典 \mathbf{D} ,训练信号 \mathbf{x} ,前一次迭代稀疏系数 α_0 ,稀疏度 K ;

输出:稀疏系数 α ;

初始化参数:(1)索引集 $T_0 := \text{sort}(|\alpha_0|)$, $K/3$;(2)残差 $\mathbf{r}_0 := \mathbf{x} - \mathbf{D}_{T_0} \alpha_{T_0}$;(3) $\epsilon_0 := \|\mathbf{r}_0\|^2$ 。

循环执行步骤1~7:

步骤1:从残差的投影中取前 $K/3$ 个元素,即 $S_n := \text{sort}(|\mathbf{D}^T \mathbf{r}_{n-1}|, K/3)$;

步骤2:更新索引集, $\tilde{T}_n := (T_{n-1}, S_n)$;

步骤3:由最小二乘法得到稀疏系数, $\tilde{\alpha}_n := (\mathbf{D}_{\tilde{T}_n})^+ \mathbf{r}_n$;

步骤4:对 $\tilde{\alpha}_n$ 排序,取前 K 个更新索引集, $T_n := \text{sort}(|\tilde{\alpha}_n|, K)$;

步骤5:用最小二乘法更新稀疏系数, $\alpha_n := (\mathbf{D}_{T_n})^+ \mathbf{x}$;

步骤6:更新残差,即 $\mathbf{r}_n := \mathbf{x} - \mathbf{D}_{T_n} \alpha_{T_n}$ 和 $\epsilon_n := \|\mathbf{r}_n\|^2$;

步骤7:判断 $\epsilon_n > \epsilon_{n-1}$,若满足,则停止迭代,若不满足,则执行步骤1。

3 实验结果及分析

本文通过3个实验验证DL-CVS模型的有效性。第一个为字典学习方法的对比实验,字典学习是本模型的重点组成部分,将其与稀疏字典和其他模型中提及的字典学习算法作比较,从而说明本文方法可在有效降低视频重构时间的同时提高重构精度,非常适合应用于本视频编解码模型中。其次为本文模型整体性能测试,在验证其可行性的同时分析不同解码模式下图像组的重构质量及其特点。最后为本文模型与其他视频编解码模型的性能对比。所有实验均在Matlab R2012b环境下编写调试,并在CPU 3.10 GHz和4 GB内存的双核PC机上运行。

3.1 字典学习方法对比

为验证本模型字典学习方法的有效性,将其与已有字典学习方法以及正交基字典进行比较。对比文献[8]算法1,每帧对应的字典是由其前后两1帧重构结果及其生成的边信息中提取不重叠块作为训练信号,再通过K-SVD算法学习得到的。对比文献[9]算法2,每帧中各个图像块对应的字典由已重构的相邻两帧中以目标块为中心的 17×17 的正方形窗口内提取出来的图像块直接构成,不用进行字典训练,且每个图像块对应的字典各不相同。算法3为正交小波基字典(Discrete Wavelet Transform, DWT)。表1中归纳总结了本文和文献[8,9]提出的3种算法所得字典特征,并对其进行比较。可以看

出,相较于其他算法,本文方法一帧或多帧图像使用同一个字典,且字典学习的稀疏编码阶段只需更新 $2K/3$ 个系数,有效降低字典学习过程的计算复杂度和时间。

图 2 为采样率为 50% 的 foreman 第 15 帧的重构图像峰值信噪比(Peak signal to noise ratio, PSNR)比较。第 15 帧图像与前一帧变化较大,因此对其重构效果可以反映出算法对运动幅度较大的视频序列的适应性。其中图 2(a)和(b)分别为第 15 帧原始图像和使用本文字典学习方法根据第 13 帧(I 帧)重构结果更新的稀疏字典。图 2(c~f)分别为使用 DWT、文献[8]、文献[9]和本文方法所得稀疏字典对第 15 帧的重构结果和字典学习时间。可以看出,本文方法对应的图像重构精度最高,对运动幅度较大的视频序列有最好的适应性。另一方面,本文方法的字典学习时间较文献[8]缩短了将近 1/2,同时,DL-CVS 模型只在特定情况下才更新字典,较文献[8]的编解码模型更具有实时性和实用性。因此,本文字典学习方法具有重构精度高、速度快,字典更新高效等特点,更适用于 DL-CVS 模型。

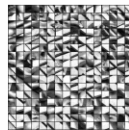
表 1 本文方法与文献[8,9]方法的字典特征比较

Table 1 Comparisons of the proposed scheme and schemes in Refs. [8,9]

特征	本文方法	文献[8]	文献[9]
字典组成	对前一个 I 帧重叠分块,将其作为训练信号进行字典学习	从前后两 I 帧和边信息提取不重叠块作为训练信号,再进行字典学习	将已重构相邻两帧中与目标块在空间上对应的块提取出来,直接构成字典原子
字典大小	256×256	256×256	空间对应的窗口大小相邻帧数(个原子)
每帧字典个数	1	1	每个分块对应一个字典,对于的 QCIF 格式视频流,一帧对应 99 个字典
字典学习类型	全局型	全局型	局部型
字典学习	是	是	否
复杂度	对个系数进行稀疏编码和字典更新的迭代	对 256 个系数进行稀疏编码和字典更新的迭代	对每一分块的 578 个系数进行稀疏编码



(a)原始图像
(a)Original image



(b)本文方法训练的字典
(b)Training dictionary of the proposed algorithm



(c) DWT
(c) DWT
(PSNR=30.39 dB)



(d)文献[8]算法
(d)Ref.[8] algorithm
(PSNR=33.47 dB, $t=14.36$ s)



(e)文献[9]算法
(e)Ref.[9] algorithm
(PSNR=25.31 dB, $t=0.33$ s)



(f)本文方法
(f)The proposed algorithm
(PSNR=35.29 dB, $t=7.95$ s)

图 2 Foreman 第 15 帧的字典和重构图像比较

Fig. 2 Dictionary and comparisons of recovery images made by different schemes of the 15 th frame of Foreman sequence

3.2 模型性能分析

为验证 DL-CVS 模型的有效性,本实验使用 QCIF 格式的 Forman 和 Salesman 测试序列(176×144)对 DL-CVS 模型进行仿真测试。其中,将每帧图像分为 16×16 的不重叠块。阈值下限 $thrL = 30$ dB,阈值上限 $thrH = 1\ 000$ dB,该值可根据视频运动剧烈程度进行调整。

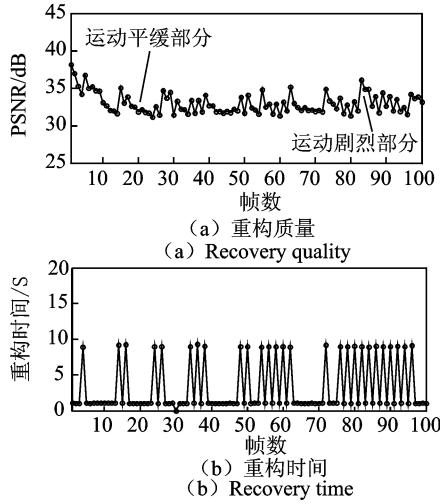


图3 Foreman 视频序列的重构质量和时间

Fig.3 Recovery quality and time of Foreman sequence

图3给出了 Foreman 序列在采样率为 50% 时的重构质量和重构时间。其中,若进行字典更新,则 P 帧重构时间为该帧重构时间与字典学习时间之和,单帧重构时间约为 1 s,单次字典更新时间均为 8 s。通过实验可以得出,序列包含的运动越平缓,其重构质量越一致,有时甚至不需要对 P 帧进行重构,同时因其字典更新次数较少,所需重构时间很短;运动越复杂越剧烈,其相邻帧重构质量差别越大,需更新字典,且因为字典更新而使得重构时间增长。若不更新字典,则连续帧的重构质量逐渐下降,而更新字典后的帧重构质量较前一 GOP 会有所上升。

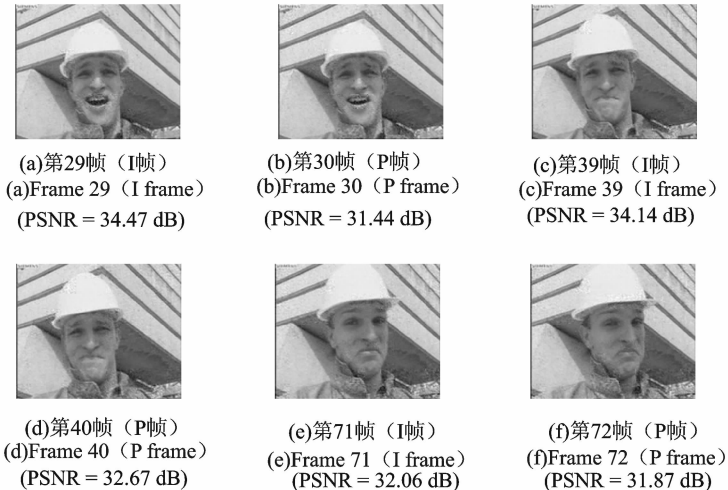


图4 3种模式对应 GOP 的重构图像

Fig.4 Recovery images of GOP of three modes

图 4 为 3 种模式对应 GOP 的重构图像。其中, (a) 和 (b) 为跳帧模式; (c) 和 (d) 为普通模式; (e) 和 (f) 为字典学习模式。可以看出, 跳帧模式中两帧均采用 I 帧重构结果, 但因其原图像略有不同, 因此 P 帧重构质量较 I 帧差些; 普通模式中两帧重构质量差别不很大, 造成差别的原因是因为差分值量化级较少, 从而传输数据量小, 重构质量略低; 字典学习模式中两帧的平均重构质量在 3 个模式中最低, 因为其环境场景变化, 字典不再合适, 间接说明应更新字典。

3.3 不同模型性能对比

表 2 为不同采样率和编解码结构下 Foreman 和 Salesman 视频序列的平均 PSNR 比较。实验结果说明, 相较于文献[7]的分布式压缩感知模型, 文献[8,9]基于字典学习的分布式压缩感知模型和基于 DWT 的编解码模型, 本文提出的模型在不同采样率下均有较好的重构精度; 同时本文的字典学习方法在不同视频序列中均更适合 DL-CVS 模型。

表 2 4 种算法的平均 PSNR 对比

Table 2 Comparison of PSNR of four schemes

MR/ %	Foreman					Salesman				
	本文	文献[7]	文献[8]	文献[9]	DWT	本文	文献[7]	文献[8]	文献[9]	DWT
50	33.19	29.44	32.37	25.56	30.18	33.91	31.26	33.81	26.22	30.57
40	31.75	27.21	30.84	24.01	26.27	32.45	29.37	32.04	25.01	27.67
30	30.03	26.30	29.76	22.53	21.52	31.48	27.64	31.21	23.93	22.90
20	28.28	20.75	26.22	20.99	9.61	30.69	23.59	30.40	22.82	16.81

4 结束语

压缩感知理论将复杂度从编码端转移到解码端, 字典学习算法可有效提高解码精度, 本模型结合二者的特点, 对视频序列进行编解码。理论分析可知, 该模型可以显著降低编码端的系统复杂度和功耗, 提高抗误码性; 实验结果表明, 该模型对于视频的解码重构质量从视觉效果和 PSNR 值两方面都较传统模型略优, 字典学习方法更是将字典更新时间缩短了 1/2。理论分析和实验结果表明, 该方法适合低复杂度、低功耗的应用场合。未来的研究工作将侧重于对模型结构和所采用字典学习算法的进一步完善和改进, 以期能够灵活方便地应用于无线多媒体传感器网络中。

参考文献:

- [1] Akyildiz I F, Melodia T, Chowdhury K R. A survey on wireless multimedia sensor networks [J]. *Computer Networks*, 2007, 51(4): 921-960.
- [2] Candès E J, Wakin M B. An introduction to compressive sampling [J]. *Signal Processing Magazine, IEEE*, 2008, 25(2): 21-30.
- [3] Donoho D L. Compressed sensing [J]. *Information Theory, IEEE Transactions on*, 2006, 52(4): 1289-1306.
- [4] 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 Guangming, Liu Danhua, Gao Dahua, et al. Advances in theory and application of compressed sensing [J]. *Chinese Journal of Electronics*, 2009, 37(5): 1070-1081.
- [5] 卓力, 刘博仑, 沈兰荪. 面向无线视频传感器网络的低复杂度视频编码算法[J]. *数据采集与处理*, 2009, 24(1): 119-124.
Zhuo Li, Liu Bolun, Shen Lansun. Low-complexity video coding for wireless video sensor networks [J]. *Journal of Data Acquisition and Processing*, 2009, 24(1): 119-124.
- [6] Cossalter M, Valenzise G, Tagliasacchi M, et al. Joint compressive video coding and analysis [J]. *Multimedia, IEEE Transactions on*, 2010, 12(3): 168-183.
- [7] 武晓嘉, 郭继昌, 姜丹, 等. 一种多视点立体视频分布式压缩感知编解码模型[J]. *天津大学学报*, 2012, 45(10): 895-902.
Wu Xiaojia, Guo Jichang, Jiang Dan, et al. A distributed compress sensing code model in multi-view stereo video [J]. *Journal of Tianjin University*, 2012, 45(10): 895-902.

- [8] Chen H W, Kang L W, Lu C S. Dynamic measurement rate allocation for distributed compressive video sensing [C]//Proceedings of Visual Communications and Image Processing. Huangshan, China; SPIE, 2010; 77440I-77440I-10.
- [9] Prades-Nebot J, Ma Y, Huang T. Distributed video coding using compressive sampling [C]// Proceedings of Picture Coding Symposium (PCS). Chicago, USA; IEEE, 2009; 1-4.
- [10] Engan K, Aase S O, Hakon Husoy J. Method of optimal directions for frame design [C]//Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Phoenix, USA; IEEE, 1999, 5; 2443-2446.
- [11] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation [J]. Signal Processing, IEEE Transactions on, 2006, 54(11): 4311-4322.
- [12] Pudlewski S, Prasanna A, Melodia T. Compressed-sensing-enabled video streaming for wireless multimedia sensor networks [J]. Mobile Computing, IEEE Transactions on, 2012, 11(6): 1060-1072.
- [13] Gan L, Do T, Tran T D. Fast compressive imaging using scrambled block Hadamard ensemble [C]//Proceedings of European Signal Processing Conference (EUSIPCO). Lausanne, Switzerland; EURASIP, 2008;1-5.
- [14] Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed l0 norm [J]. Signal Processing, IEEE Transactions on, 2009, 57(1): 289-301.
- [15] Smith L N, Elad M. Improving dictionary learning: Multiple dictionary updates and coefficient reuse [J]. Signal Processing Letters, IEEE, 2013, 20(1): 79-82.

作者简介:郭继昌(1966-)男,教授,博士生导师,研究方向:数字图像处理、滤波器理论及设计等,E-mail:jcguo@tju.edu.cn;金卯亭嘉(1990-),女,硕士研究生,研究方向:压缩感知、稀疏表示和字典学习算法。

