

文章编号:1005-0523(2013)01-0052-06

交通监控中基于压缩感知的WMSN视频编解码

罗 晖,王世昌,褚红亮,杨成武

(华东交通大学信息工程学院,江西 南昌,330013)

摘要:在交通监控中,利用WMSN节点获取路况视频,适用于复杂多变的地理环境,但视频图像数据量大,而WMSN节点资源受限。基于压缩感知技术和传统视频编码理论,提出了一种新的WMSN视频编解码方法。在编码端,依据视频图像的相关性对视频流进行分帧,再根据其特点分别进行压缩感知处理。在解码端,运用OMP算法重构关键帧和残差帧视频图像,进而恢复视频流。仿真实验表明,在保证视频图像重构质量的前提下,可以减少监控网中视频传输的数据量。

关键词:交通监控;WMSN;压缩感知;视频编解码

中图分类号:TN919.8

文献标志码:A

目前,在交通监控中,为实现对交通路况的实时视频监控,需要布置复杂的通信网络,以保证交通监控网络^[1]长期稳定、可靠地运行。而无线多媒体传感器网络(wireless multimedia sensor networks, WMSN)^[2]具备部署快速、组网灵活、感知信息丰富等优点。将WMSN用于交通监控中,不但可以实时获取路况场景的视频信息,而且能适应复杂多变的地理环境,其基本网络体系结构如图1所示。在监控区域内,视频传感器节点负责采集路况视频信息;簇头节点冗余部署在视频传感器节点周围,负责处理视频信息,并通过自组网以多跳中继方式将视频信息传输给汇聚节点;汇聚节点以无线通信方式将视频信息发送给交通远程监控中心。

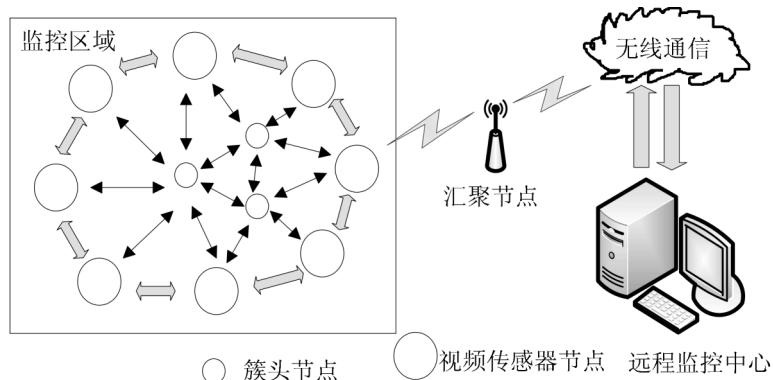


图1 交通视频监控WMSN网络体系结构

Fig.1 WMSN network architecture of traffic video surveillance

在上述交通视频监控WMSN网络中,视频信息数据量巨大,而WMSN传感器节点的能量、处理能力和

收稿日期:2012-11-25

基金项目:国家自然科学基金项目(61261040)

作者简介:罗晖(1969—),男,副教授,研究方向为无线通信,多媒体信息处理及压缩感知。

存储资源受限。因此,研究一种高效的视频压缩编码方法就成为了WMSN应用于交通视频监控的关键。

近年来,由Candes和Donoho等人提出的压缩感知(compressed sensing, CS)^[3]理论在信号采集和处理领域引起了国内外学者的广泛关注。CS理论指出,将一个稀疏或可压缩的高维信号投影到低维空间上,获取原始信号的观测值;同时,借助少量的观测值,可以通过一定的线性或非线性的优化算法重构出原始信号。在CS处理过程中,因为观测值的个数远远少于Nyquist采样数,所以,如果将CS理论引入到视频图像编码中^[4-6],就有望实现更有效的数据压缩和更准确的数据重构。这为视频编解码提供了一种新思路,也为WMSN应用于交通实时监控提供了可能。

2 压缩感知理论

假设一维有限长离散信号 $\mathbf{x} \in R^{N \times 1}$, 将其表示为 $N \times 1$ 的列向量, 所选取变换域空间的基函数为: $\Psi = [\Psi_1, \Psi_2, \dots, \Psi_N]$, 则 $\mathbf{x}$ 在 Ψ 域表示为

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

式中: α 是 $\mathbf{x}$ 在 Ψ 上的表示系数, 若它有 K 个分量不为零, 而 $N-K$ 个分量为零或者非常接近于零, 就称 $\mathbf{x}$ 在 Ψ 域是 K 稀疏的或是近似 K 稀疏, 也就是说, $\mathbf{x}$ 是可压缩的。

在稀疏条件下, 用一个 $M \times N$ 的观测矩阵 Φ 作用于信号 $\mathbf{x}$, 获取的观测向量为 $\mathbf{y}$, 则 $\mathbf{y} = \Phi \mathbf{x}$, 将式(1)代入, 得到压缩感知数学表达式^[7]

$$\mathbf{y} = \Phi \mathbf{x} = \Phi \Psi \alpha = \Theta \alpha \quad (2)$$

式中, Θ 是CS算子, 在数学上表现为一 $M \times N$ 矩阵, 由于 $M \ll N$, 因此, 式(2)是个病态方程。但当 Θ 满足有限等距性质(restricted isometry property, RIP)^[8]时, 等价于 Ψ 和 Φ 不相关, 通过数学优化方法能够获得精确解^[9]。因此, 重构原始信号等价于求解一个优化问题, 其优化目标表示如下

$$\min_{\alpha} \|\hat{\alpha}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \Phi \Psi \hat{\alpha} = \Theta \hat{\alpha} \quad (3)$$

式(3)是一个 l_0 范数的优化问题, 即NP-难问题, 在多项式时间内难以求解, 甚至无法验证其解的可靠性。理论分析表明, 在一定条件下, l_1 最小范数和 l_0 最小范数具有等价性, 可以得到相同的解。那么将式(3) l_0 最小范数转换为 l_1 最小范数

$$\min_{\alpha} \|\hat{\alpha}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \Phi \Psi \hat{\alpha} = \Theta \hat{\alpha} \quad (4)$$

式(4)是一个非线性优化问题, 可用贪婪追踪算法解此优化问题, 如正交匹配追踪(OMP)算法^[10], 也可以将非凸问题转化为凸问题寻找信号的逼近, 如基追踪(BP)算法^[11]、快速迭代收缩阈值(FIST)算法^[12]等。

3 基于CS的WMSN视频编解码

3.1 视频编码

在交通视频监控中, 监测区域场景一般是固定的, 连续视频帧图像的背景相同, 且视频图像存在帧内帧间相关性。基于此特点, 将视频图像组(group of picture, GOP)中的图像帧定义为关键帧和非关键帧再分别进行编码。

在编码过程中, 将WMSN节点采集的每个GOP的第一帧作为视频图像关键帧, 标记为 I 帧, 其余的视频图像帧称作非关键帧, 标记为 P_j 帧。同时, 因一个GOP中的视频图像背景相似, 帧间存在强相关性, 借助 I 帧的图像信息, 并利用帧间差值技术, 可以获得 P_j 帧的残差帧视频图像 ΔP_j 。

3.1.1 稀疏表示

对于交通视频图像, 因其内容中车辆等目标图像的像素存在邻域相关, 且细节丰富。而Symlets小波具有双正交、近似对称和紧支撑等特点, 其小波系数在保持稀疏性时, 能有效刻画图像细节信息。因此, 可

以选取Symlets小波对交通视频图像的 I 帧和残差帧视频图像 ΔP_j 进行稀疏表示。

在稀疏表示过程中,设 $x(i, j)$ 为一幅图像的像素点, i 、 j 表示像素点在图像中的位置。 $\varphi(i, j)$ 代表二维小波,则小波变换表达式为

$$WT_x(a, b_1, b_2) = \langle x(i, j), \varphi_{a, b_1, b_2}(i, j) \rangle = \frac{1}{a} \sum_{i=0}^M \sum_{j=0}^N f(i, j) \varphi\left(\frac{i-b_1}{a}, \frac{j-b_2}{a}\right) \quad (5)$$

式中: $\varphi_{a, b_1, b_2}(i, j)$ 表示 $\varphi(i, j)$ 的尺度伸缩与二维位移。

3.1.2 观测过程

在CS理论中,对信号进行重构,要求CS算子满足RIP特性,即观测矩阵与稀疏基函数相关性非常小。已有理论证明:随机高斯测量矩阵与绝大多数稀疏基不相关,且存储简单,因此,选用随机高斯测量矩阵作为观测矩阵对小波稀疏系数进行观测,可以最终获得重构视频图像所需要的采样信号。

1) I 帧观测过程。设高斯测量矩阵是 $\Phi \in R^{M \times N}$, 其元素相互独立并服从0均值、方差为 $1/\sqrt{M}$ 的高斯分布即

$$\Phi = \Phi_{i,j} \sim N(0, \frac{1}{\sqrt{M}}) \quad (6)$$

经过观测过程,获取的采样观测值为

$$y_{i,k} = \Phi \cdot WT_x = \sum_{j=1}^N \Phi(i, j) WT_x(j, k) \quad (7)$$

式中: $\Phi(i, j)$ 表示第 i 行、第 j 个元素; $WT_x(j, k)$ 表示第 k 列、第 j 行小波系数。

对二维图像的观测是按列进行的,观测矩阵的行数,决定了视频图像采样值的多少。设计合适的观测矩阵,可以获得最优的观测值以重构出高质量的视频图像。

2) P_j 帧观测过程。稀疏表示已获得了 P_j 帧的残差帧视频图像 ΔP_j , ΔP_j 与 P_j 相比较, ΔP_j 中的像素值更加稀疏,因此,经小波变换后,非零的小波系数非常少,为加快观测的实时性,由 I 帧观测矩阵构造出综合观测矩阵 Φ_p , 对GOP中所有 P_j 帧的 ΔP_j 进行综合观测, Φ_p 表示为

$$\Phi_p = \begin{bmatrix} \Phi_1 & 0 & 0 & 0 \\ 0 & \Phi_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \Phi_j \end{bmatrix} \quad (8)$$

式中: $\Phi_1 = \Phi_2 = \dots = \Phi_j = \Phi$, Φ_j 对应 P_j 帧的残差帧视频图像,残差帧视频图像 ΔP_j 按式(5)进行处理后获得相应的小波系数,进而得到采样观测值为

$$y_{\Delta P_j} = \Phi_p \cdot WT_{x_{\Delta P_j}} = \begin{bmatrix} \Phi_1 & 0 & 0 & 0 \\ 0 & \Phi_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \Phi_j \end{bmatrix} \begin{bmatrix} \langle (P_1 - I), \varphi_{i,j} \rangle \\ \langle (P_2 - I), \varphi_{i,j} \rangle \\ \vdots \\ \langle (P_j - I), \varphi_{i,j} \rangle \end{bmatrix} \quad (9)$$

由于非关键帧视频图像 P_j 经帧间差值技术处理和小波变换后,小波系数已经变得相当稀疏,且 Φ_p 中的矩阵元素 Φ_i ($i=1, \dots, j$) 相同,因此,设计的综合观测矩阵降低了编码复杂性。

3.1.3 视频图像编码具体过程

基于以上分析,交通视频图像的编码过程如下:在每个GOP中, I 帧经小波变换直接感知观测;对于 P_j 帧,先获取残差帧视频图像 ΔP_j ,再经小波变换后,利用综合观测矩阵获取观测值,观测值经量化^[13]和熵编码^[14-15]后完成编码。具体如图2所示。

3.2 视频解码

在传统视频图像解码中,视频解码就是实现编码的逆过程。然而基于CS的交通监控WMSN视频解码中,重构视频是一个求解数值优化问题的过程。

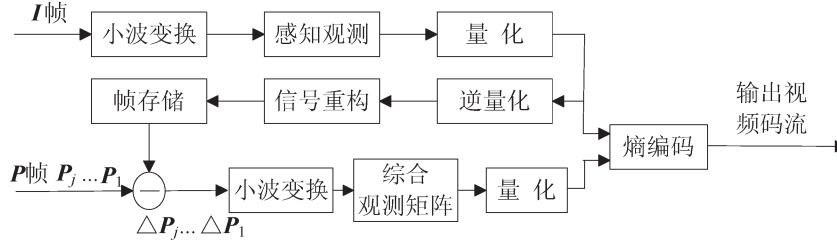


图2 基于CS的WMSN视频编码过程
Fig.2 WMSN video encoding based on CS

3.2.1 解码算法

在解码端,CS重构过程中的优化目标函数为

$$\min \|\hat{\mathbf{x}}(i, j)\|_1 \quad \text{s.t.} \quad \mathbf{y}_{i, j} = \Phi_{i, j} \varphi(i, j) \hat{\mathbf{x}}(i, j) \quad (10)$$

利用拉格朗日算子将式(10)转化为无约束的优化问题

$$\min \|\mathbf{y}_{i, j} - \Phi_{i, j} \varphi(i, j) \hat{\mathbf{x}}(i, j)\| + \lambda \|\hat{\mathbf{x}}(i, j)\| \quad (11)$$

利用OMP算法迭代求解式(11),重构I帧和残差帧视频图像 ΔP_j ,算法流程描述如下

输入: $\Theta = \Phi\Psi \in R^{M \times N}$:CS采样算子矩阵

$\mathbf{y} \in R^{M \times M}$:采样观测值

K :视频图像帧的稀疏度

输出: $\hat{\mathbf{x}} \in R^{N \times N}$:重构的视频图像帧或残差帧视频图像

中间变量: $\mathbf{r} \in R^{M \times M}$:每次迭代产生的残差

Λ_t : t 次迭代后,选出的所有算子向量索引 λ_t 的集合

初始化: $\mathbf{r}_0 = \mathbf{y}$, $\Lambda_0 = \varnothing$, $t = 1$

搜索最佳匹配算子向量,记录其索引: $\lambda_t \leftarrow \arg \min_{j=1, 2, \dots, M} |\langle \mathbf{r}_{t-1}, \Theta_j \rangle|$

$$\Lambda_t \leftarrow \Lambda_{t-1} \cup \lambda_t$$

计算新的 $\hat{\mathbf{x}}_t$ 和残差 $\mathbf{r}_t$: $\hat{\mathbf{x}}_t = \sum_{i=0}^{t-1} \Theta_i \mathbf{r}_i$, $\mathbf{r}_t = \mathbf{y} - \hat{\mathbf{x}}_t = \mathbf{y} - \sum_{i=0}^{t-1} \Theta_i \mathbf{r}_i$

循环迭代 K 次,获得原始视频图像的近似表达: $\hat{\mathbf{x}} \approx \sum_{i \in \Lambda_t} \Theta_i \mathbf{r}_i$

3.2.2 视频图像解码具体过程

在视频图像解码中,视频码流经过熵解码和逆量化后,获得了视频图像I帧和残差帧视频图像 ΔP_j 的观测值。然后,在图像压缩感知重构中,通过运行OMP算法迭代处理相应的观测值,重构出GOP中的I帧图像和残差帧视频图像 ΔP_j ,进而恢复出视频流。其解码具体过程如图3所示。

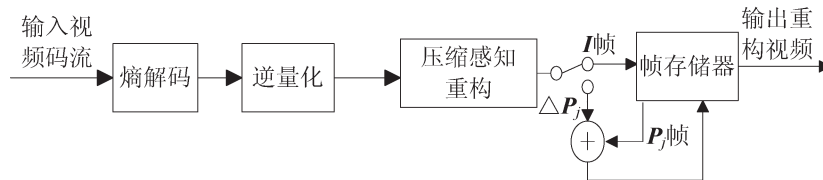


图3 基于CS的交通监控WMSN视频图像解码
Fig.3 WMSN video decoding of traffic surveillance based on CS

4 实验仿真与分析

4.1 I 帧视频图像编解码

在交通视频监控^[15]中, P_j 帧借助 I 帧信息完成重构, 因此 I 帧的重构是视频重构的关键。在实验中, 设置 GOP 中的原始 I 帧图像大小为 256×256 , 通过改变观测次数重构图像, 并对其进行比较, 具体仿真结果如图 4 所示。图 4(a) 是未经处理的原始视频 I 帧图像; 图 4(b) 是在观测次数为 200 时, 重构出的 I 帧视频图像, 该图中, 在车辆聚集处, 车辆之间的间隔不清, 路标牌上的字迹模糊, 这表明视频图像质量较差, 其 PSNR 值仅为 27.12 dB; 图 4(c) 是在观测次数为 300 时, 重构出的 I 帧视频图像, 与图 4(b) 相比, 其视觉效果更好, 能辨出车间距离, 字迹更清晰, 没有明显的模糊效应, 而且 PSNR 值达到 32.94 dB。

因此, 可以得出结论: 在视频图像重构过程中, 随着观测次数的增加, 重构出的视频图像的质量更高、PSNR 值也更大。

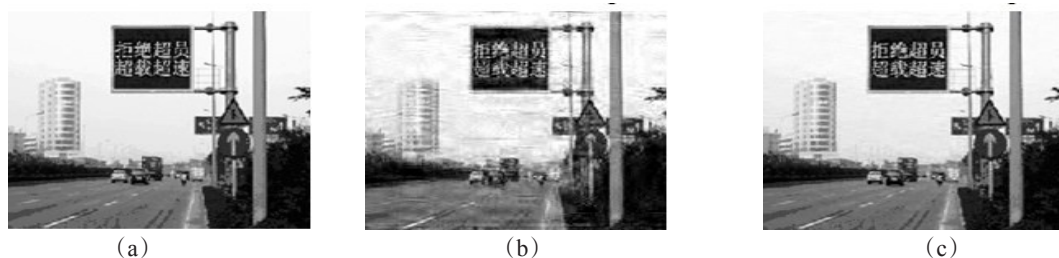


图 4 实验仿真结果图

Fig.4 The result pictures of experimental simulation

从视频图像质量的主观评价角度考虑, 图像质量越高越好, 那么需要相应地增大观测次数。但是, 如果观测次数太大, 就延长了重构时间, 因此, 需要依据交通视频监控标准设置合适的观测次数。

4.2 P_j 帧视频图像编解码

对于 P_j 帧视频图像处理过程, 采用与 I 帧相同的 CS 参数, 当观测次数增加到一定程度时, 重构出的 P_j 帧视频图像的 PSNR 值没有太大变化, 结果如图 5 所示。在图 5 中, 当观测次数大于 300 时, PSNR 值基本不变。这种现象是因为: 残差帧视频图像经 Symlets 小波变换后, 小波系数已经相当稀疏, 并且能重构出高质量的视频图像。因此, 增加观测次数, 重构出的 P_j 帧视频图像 PSNR 不会显著变化。

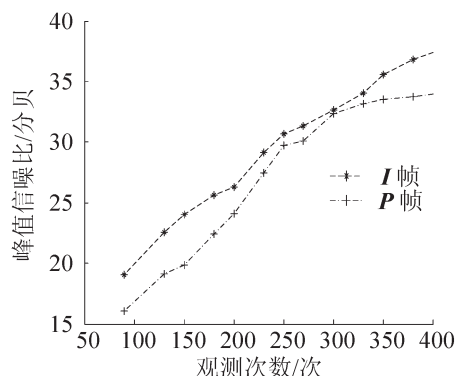


图 5 实验仿真结果图

Fig.5 The result pictures of experimental simulation

5 总结

在交通视频监控过程中, 利用监控视频图像的帧内、帧间相关性, 并采用压缩感知技术, 对视频图像中的关键帧与非关键帧进行编解码。经实验仿真分析得出: 该方法在恢复原始的视频图像过程中, 不但大大减少了传输数据量, 而且重构视频图像的 PSNR 值达到了 30 dB 以上。

参考文献:

- [1] KUMAR P. Framework for real-time behavior interpretation from traffic video [J]. IEEE Transaction on Intelligent Transportation Systems, 2005, 6(1): 43-53.

- [2] AKYILDIZ I F, MELODIA T, CHOWDHURY K R. Wireless multi-media sensor network: a survey [J]. IEEE Transactions on Wireless Communication, 2007, 14(6): 32-39.
- [3] DONOHO D L. Compressed sensing [J]. IEEE Transaction Information Theory, 2006, 52(4): 1289-1306.
- [4] WAKIN M B, LASKA J N, DUARTEM. F, et al. Compressive imaging for video representation and coding[C]//Processing Picture Coding Symposium, special session on Next Generation Video Representation, Beijing, 2006: 1-6.
- [5] ROMBERG J. Imaging via compressive sampling [J], IEEE Signal Processing Magazine, 2008, 25(2): 14-20.
- [6] DUARTE M, DAVENPORT M, TAKHAR D, et al. Single-pixel imaging via compressive sampling[J], IEEE Signal Processing Magazine, 2008, 25(2): 83-91.
- [7] 石光明, 刘丹华, 高大化, 刘哲, 林杰, 王良君. 压缩感知理论及其研究进展[J]. 电子学报, 2009, 37(5): 1070-1081.
- [8] TONG ZHONG. Sparse recovery with orthogonal matching pursuit under RIP [J]. IEEE Transaction, 2011, 57(9): 6215-6221.
- [9] CANDES E J, WAKIN M B. An introduction to compressive sampling [J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30.
- [10] TROPP J A, GIBERT A C. Signal recovery from random measurements via orthogonal matching Pursuit [J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655-4666.
- [11] CHEN S B, DONOHO D L, SAUNDERS M A, Atomic decomposition by basis pursuit [J]. Siam Journal on Scientific Computing, 1998, 20(1): 33-61.
- [12] A.BECK, M.TEBOULLE. A fast iterative shrinkage-thresholding algorithm for linear inverse problems [J]. Siam Journal on Imaging Sciences, 2009, 2(1): 183-202.
- [13] Y.BAIG, E.M.K.LAI, J.P.LEWIS. Quantization effects on compressed sensing video[C]//17th International Conference on Telecommunications. IEEE, 2010: 443-446.
- [14] THOMAS WIEGAND, GARY J SULLIVAN, AJAY LUTHRA, et al. Overview of the H.264 video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 1-19.
- [15] 郭厚焜, 吴峰, 黄萍. 基于压缩感知和字典学习的背景差分法[J]. 华东交通大学学报, 2012, 29(1): 43-47.

WMSN Video Encoding/Decoding Based on Compressed Sensing in Traffic Surveillance

Luo Hui, Wang Shichang, Chu Hongliang, Yang Chengwu

(School of Information Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: In traffic surveillance, WMSN is applied to capture the video of road state, which can adjust to complex and variable environments. But WMSN sensor's resource is limited and the number of video images yielded is tremendous yet. Based on compressed sensing and traditional video coding theory, this paper proposes a new method of WMSN video encoding/decoding. In encoding, the video stream is divided into key frames and non-key frames according to video image correlation, which are processed in line with their characteristics respectively. In decoding, video stream is further reconstructed after the key frames and residual video images are reconstructed through OMP algorithm. The simulation experiments demonstrate that the data volume of video transmission is largely reduced on the premise of keeping the reconstructed video image quality.

Key words: traffic surveillance; sensor network; compressed sensing; video coding and decoding