

DOI:10.3969/j.issn.1673-4785.201211017

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20131012.1814.005.html>

基于压缩感知的监控视频重构

李昂¹, 马强², 岑翼刚¹, 赵瑞珍¹, 岑丽辉³

(1. 北京交通大学 信息科学研究所, 北京 100044; 2. 宜昌供电公司 电能计量中心, 湖北 宜昌 443000; 3. 中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘要:传统对视频先采样再压缩的方法极大地浪费了硬件资源, 针对这一问题, 基于压缩感知理论提出了一种对监控视频采样及重构的方法. 该方法先获得帧间差值, 再将差值投影到小波稀疏域后进行压缩采样. 选取每一分组中的中间帧作为关键帧, 对关键帧不进行处理, 完全保留所有采样点. 恢复时利用关键帧和差值可以得到初步重构的视频序列, 最后通过运动估计和运动补偿得到优化. 实验结果表明, 与仅使用压缩感知对差分帧进行重构的方法相比, 该方法对监控视频重构帧序列图像的平均峰值信噪比有较大的提升, 且受采样点数影响较小, 具有很好的鲁棒性.

关键词:压缩感知; 监控视频; 关键帧; 帧间差值; 运动补偿

中图分类号: TP18; TN911.72 **文献标志码:** A **文章编号:** 1673-4785(2013)06-0512-05

中文引用格式: 李昂, 马强, 岑翼刚, 等. 基于压缩感知的监控视频重构[J]. 智能系统学报, 2013, 8(6): 512-516.

英文引用格式: LI Ang, MA Qiang, CEN Yigang, et al. Reconstruction of surveillance video based on compression sensing[J].

CAAI Transactions on Intelligent Systems, 2013, 8(6): 512-516.

Reconstruction of surveillance video based on compressed sensing

LI Ang¹, MA Qiang², CEN Yigang¹, ZHAO Ruizhen¹, CEN Lihui³

(1. Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China; 2. Electric Energy Measurement Center, Yichang Power Supply Company, Yichang 443000, China; 3. Institute of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract: The traditional first-compression-then-sampling method used for video greatly wastes hardware resources. In order to solve this problem, a sampling and reconstruction method used for the surveillance video is proposed based on compressed sensing. With this method, first, the interframe difference is calculated and projected to the sparse domain of the wavelet, then the wavelet coefficients are sensed by the sensing matrix. The middle frame in every group is taken as the key frame, which is not processed, and all sampling points are completely preserved. In the stage of recovery, the preliminary reconstruction of the video sequence can be obtained by using the key frames and interframe difference. Finally, the optimization is realized by motion estimation and motion compensation. The experimental results show that, compared with the reconstruction for difference frames by only using compressed sensing, this method can greatly lift the mean peak signal-noise ratio of the sequence image on the reconstruction frames of a surveillance video. In addition, the influence caused by the quantity of the sampling points is small, and the robustness is excellent.

Keywords: compressed sensing; surveillance video; key frame; interframe difference; motion compensation

随着社会的进步和通信技术的不断发展, 人与

人之间的视频传输变得越来越频繁, 这就对通信过程中的数据存储和传输设备提出了更高的要求. 传统的奈奎斯特采样定理指出, 信号的采样频率必须大于或等于信号带宽的 2 倍, 才能准确地恢复和重构出原信号. 显然, 对于视频信号的传输, 奈奎斯特采样定理具有很大的局限性.

收稿日期: 2012-11-12. 网络出版日期: 2013-10-12.

基金项目: 国家自然科学基金资助项目(61272028, 61104078, 61073079); 中央高校基本科研业务费资助项目(2013JBZ003); 上海交通大学系统控制与信息处理教育部重点实验室开放基金资助项目(SCIP2011009); 高等学校博士学科点专项科研基金资助项目(20110162120045, 20120009110008).

通信作者: 岑翼刚. E-mail: ygcen@bjtu.edu.cn.

近年来, 由 Donoho、Candes^[1] 及 Tao 等^[2] 提出

了一种新的信息获取理论称作压缩感知 (compressed sensing 或 compressive sampling, CS). 这一理论的着眼点放在了信号的稀疏性, 即信号在合适的稀疏基下可以变换到相应的稀疏域, 从而呈现出信号的稀疏性. 另外, 对于如何将信号在稀疏域进行压缩感知处理, 很多算法被提出, 例如利用峰值变换^[3]、利用单层小波变换^[4]等.

在处理监控视频信号时, 还要充分考虑到视频中帧与帧之间的相关性. 对相邻 2 帧做差, 得到的差值图像本身就具有一定的稀疏性. 另外, 在帧图像重构的过程中, 可以通过参考帧图像对当前帧图像进行运动估计和运动补偿, 从而进一步提高视频重构质量.

1 压缩感知理论

考虑一个可压缩实值信号 f ($f \in \mathbf{R}^{N \times 1}$), 信号长度为 N , 则该信号可以由一组正交基展开 $\Psi = [\psi_1 \ \psi_2 \ \dots \ \psi_N]$, 如式(1):

$$f = \sum_{i=1}^N \psi_i \alpha_i \text{ 或 } f = \Psi \alpha. \quad (1)$$

式(1)为正交变换的反变换, 正变换为 $\alpha = \Psi^H f$. 其中, $\Psi \Psi^H = \Psi^H \Psi = I$, $\Psi \in \mathbf{C}^{N \times N}$, I 为单位阵, α 是信号 f 在 Ψ 下的稀疏表示. 若式(1)中的 α 只有 K 个非零元素 ($N \gg K$), 其余的元素为 0 或者接近于 0, 则认为该信号是稀疏的.

信号可稀疏表示是压缩感知的先验条件. 在已知信号是可压缩的前提下, 压缩感知过程可分为如下 2 步^[5]:

1) 设计一个大小为 $M \times N$ ($M \ll N$) 的测量矩阵 Φ 对稀疏信号进行观测, 得到 M 维的测量向量.

测量值 y 是一个 M 维向量, 这样使测量对象从 N 维降为 M 维. 观测过程是非自适应的, 即测量矩阵的选择不依赖于信号 f . 测量矩阵的设计要求信号从 f 转换为 y 的过程不会破坏原始信号的信息, 否则, 重构是不可能的^[6]. 由于信号 f 是可稀疏表示的, 测量过程可以表示为

$$y = \Phi f = \Phi \Psi \alpha = \Theta \alpha. \quad (2)$$

式中: Θ 是一个 $M \times N$ 矩阵. 若 Φ 满足有限等距性质 (restricted isometry property, RIP), 即对于任意 K -稀疏信号 f 和常数 $\delta_k \in (0, 1)$, 矩阵 Φ 满足:

$$1 - \delta_k \leq \frac{\|\Phi f\|_2^2}{\|f\|_2^2} \leq 1 + \delta_k,$$

则 K 个系数能够从 M 个测量值中准确地被重构. RIP 性质的等价条件是测量矩阵 Φ 和稀疏基 Ψ 不相关.

2) 由 M 维的测量向量重构信号.

当矩阵 Φ 满足 RIP 准则时, 重构 K -稀疏系数 α 最直接的方法是通过在 l_0 范数下求解式(2)的最优化问题:

$$\min_{\alpha} \|\alpha\|_{l_0}, \text{ s.t. } y = \Phi \Psi \alpha, \quad (3)$$

从而得到稀疏系数 α 的估计. 由于式(3)求解时, 方程个数远小于未知数的个数, 即 l_0 范数下的求解是 NP-Hard 问题. Donoho 等^[7]指出, 当 Φ 和 Ψ 不相关时, 求解一个更加简单的 l_1 范数优化问题会产生相同的解. 即 l_0 极小化问题可转化为 l_1 范数下的最小化问题:

$$\min_{\alpha} \|\alpha\|_{l_1}, \text{ s.t. } y = \Phi \Psi \alpha. \quad (4)$$

通过求解式(4), 可以得到稀疏域重建的结果. 求解 l_1 范数下极小化问题常用算法有 3 类: 贪婪追踪法、凸松弛法和组合算法, 例如匹配追踪法 (matching pursuit, MP)^[8]、正交匹配追踪法 (orthogonal matching pursuit, OMP)^[9]、梯度投影法 (gradient projection, GP)^[10]、链式追踪法^[11]等.

2 基于压缩感知的监控视频恢复

监控视频的特点是背景基本不变, 视频图像数据有着极大的相关性, 也就是有大量的冗余信息. 对于监控视频, 背景不是重要信息, 压缩感知主要是针对监控视频中的运动物体进行处理的.

对于 YUV 视频数据来说, 本文研究的是对其 Y 分量所构成的部分进行处理. 首先, 将视频序列进行分组, 每组包含相同的帧数. 为了达到更好的恢复效果, 在关键帧的选取上, 选取一组帧图像 (奇数个) 的中间帧作为关键帧. 帧与帧之间差值的获取比较简单, 即后一帧减去前一帧的结果作为差分的结果. 对差值图像运用压缩感知技术进行采样, 而关键帧不进行压缩感知处理, 完全保留所有采样点. 另外也可先对帧图像先感知, 再做差分, 此时则应该对 2 帧图像采用相同的测量矩阵, 即应有式(5)成立:

$$\Phi(f_1 - f_2) = \Phi f_1 - \Phi f_2. \quad (5)$$

式中: f_1 和 f_2 是相邻的 2 帧. 由式(5)可知, 对 2 帧图像的差值采样等效于分别对 2 帧图像先采样再进行差分^[12].

由于监控视频中每一帧图像的背景总是基本不变的, 当场景中的物体运动比较缓慢或者运动物体尺寸比较小时, 帧与帧的差值在空域就具备稀疏的性质. 但当场景中物体运动比较迅速或者运动的物体尺寸比较大时, 空域的稀疏性将急剧下降. 因此, 为了使差值都具有稀疏的性质, 本论文对差值进行进一步处理, 即利用二维小波变换, 将差值图像投影到比较稀疏的域.

以视频序列的第 1 组为例进行说明,该组由 5 帧图像组成,第 3 帧为关键帧.视频的采样过程如图 1 所示.

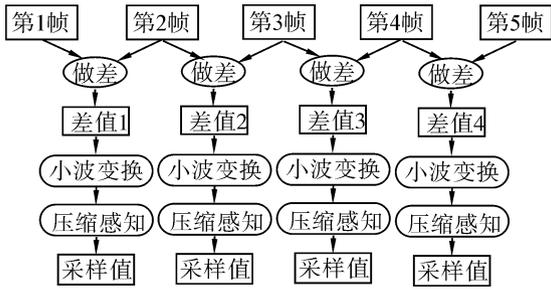


图 1 视频采样过程

Fig.1 The sampling process of surveillance video

获得采样值以后,接下来就根据采样值重构原来的视频序列.首先,根据最小 l_1 范数优化方法重构出差值图像,本文采用正交匹配追踪算.然后,用关键帧加上或减去差值图像获得与其相邻的 2 帧图像的重构,之后,这 2 帧再做参考帧,和差值图像做加法或减法得到下一帧或上一帧的重构图像.视频的恢复过程如图 2 所示.

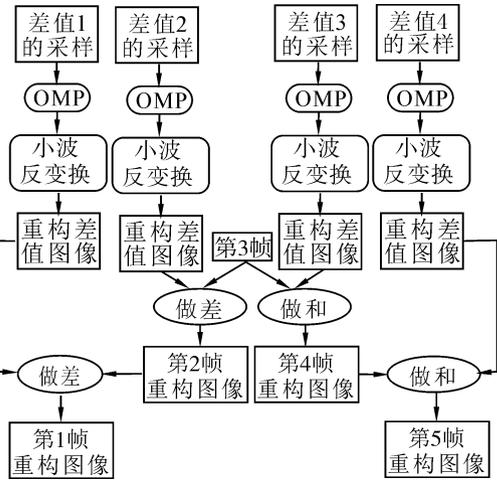


图 2 视频恢复过程

Fig.2 The reconstruction of surveillance video

在恢复视频序列的过程中,应该充分利用视频序列中帧与帧之间的相关性和连续性,把运动估计和运动补偿技术添加到视频序列的恢复过程中来.运动估计可以用来估计物体的位移,从而得到运动矢量.运动补偿可以根据得到的运动矢量,对参考帧中由于运动而产生的位移进行调整,从而得到尽可能接近于当前帧的预测帧.

根据运动估计和运动补偿的特点和作用,本文将应用到视频序列中非关键帧的重构过程中,采用的是全搜索算法.该算法分为 2 步:1)在搜索窗口内,按从左至右、从上到下的顺序逐个像素进行搜

索,分别计算各个像素的 MAD 值,直到遍历搜索窗口内所有的点.2)在所有记录下来的 MAD 值中寻找 MAD 值最小的点,即最小块匹配误差点,该点对应的位置即为所求的运动矢量.这里, MAD 是指平均绝对误差,按照式(6)进行计算:

$$MAD(i, j) =$$

$$\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N |f_k(m, n) - f_{k-1}(m + i, n + j)|. \quad (6)$$

式中: (i, j) 是位移矢量, f_k 和 f_{k-1} 分别为当前帧和参考帧的灰度值, $M \times N$ 为宏块的大小.

基于全搜索算法,运动估计和运动补偿的算法原理介绍如下:先确定当前帧中的待匹配块,设其左上角的坐标为 (p, q) ,然后在参考帧中对应的搜索窗口内找到与待匹配块 MAD 值最小的块,即预测块,设其左上角坐标为 (u, v) ,那么所求的运动矢量即为 $(u - p, v - q)$,这就是运动估计的过程.在运动补偿过程中,参考帧中的预测块根据运动矢量平移到当前帧中待匹配块的位置.如此反复,当参考帧中所有预测块都按照相应的运动矢量平移到了相应的待匹配块的位置时,就完成了对当前帧的运动补偿.

首先将关键帧作为参考帧,关键帧的前一帧的重构图像或后一帧的重构图像作为当前帧,进行运动估计得到运动矢量,再根据运动矢量进行运动补偿,得到前一帧或后一帧的优化重构图像.再将得到的优化图像作为参考帧,其前一帧的重构图像或后一帧的重构图像作为当前帧,重复前面的运动估计和运动补偿的过程,从而得到对应帧的优化重构图像.以此类推,完成整个视频序列的每一组中的非关键帧的运动估计和运动补偿,从而恢复出整个视频序列.整个过程如图 3 所示.

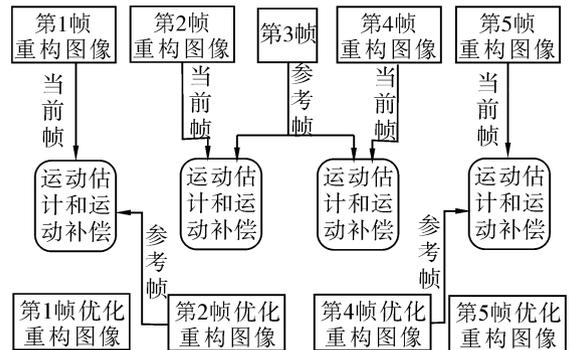


图 3 视频序列的运动补偿

Fig.3 The motion compensation of surveillance video

对于一个实际的处理系统来讲,上述内容分别

对应编码端和解码端的处理过程.编码端:1)将视频序列进行分组,每组包含相同的帧数;2)在关键帧的选取上,选取一组帧图像(奇数个)的中间帧作为关键帧;3)用传统的方法对关键帧进行熵编码;4)每组最后一帧减去前一帧获得差值图,并将差值图进行小波变换;5)在小波域对差值进行压缩感知,获得差值的采样值,并进行熵编码.解码端:1)通过熵解码获得关键帧和差值的采样值;2)将差值的采样值利用 OMP 算法进行重构,并进行小波反变换,获得重构的差值图像;3)由关键帧和差值图像获得该组非关键帧的重构图像;4)通过运动估计和运动补偿对该组非关键帧进行优化.

3 实验仿真

实验是建立在 MATLAB 运行环境下实现的.实验中处理的视频对象为监控视频 hall(QCIF 格式),共有 300 帧图像,每帧图像的大小为 144×176 ,播放时间为 10 s.采取服从 $N(0,1)$ 正态分布的随机矩阵作为测量矩阵,用 sym8 小波函数对差值图像进行小波变换,分解层数为 4 层.将视频序列分成 60 组,每组 5 帧,并且中间帧作为关键帧.

每幅差值图像的采样点数为 44×176 时,在未进行运动估计和运动补偿的情况下,重构序列的平均峰值信噪比为 40.263 4 dB.重构视频序列如图 4 所示.

当对图 4 的结果进行运动估计和运动补偿之后,重构序列的平均峰值信噪比可以提高到 42.178 4 dB.得到的重构视频序列,如图 5 所示.



图 4 未进行运动补偿的重构视频序列

Fig.4 The reconstructed surveillance video without motion compensation



图 5 已进行运动补偿的重构视频序列

Fig.5 The reconstructed surveillance video with motion compensation

由图 4~5 可以得出,在未进行运动估计和运动补偿的情况下,视频序列虽然可以重建出来,但背景和运动的人物还是存在比较明显的模糊现象.经过运动补偿以后,这种模糊现象明显降低,反映到数据上就是重构视频序列平均的峰值信噪比增加.也就是说,运动补偿可以减少较小的采样率给重构视频序列带来的不利影响,可以实现对重构视频序列的优化.

2 种情况下,重构图像与原来图像的平均峰值信噪比(PSNR)随采样点数的变化情况如图 6 所示.

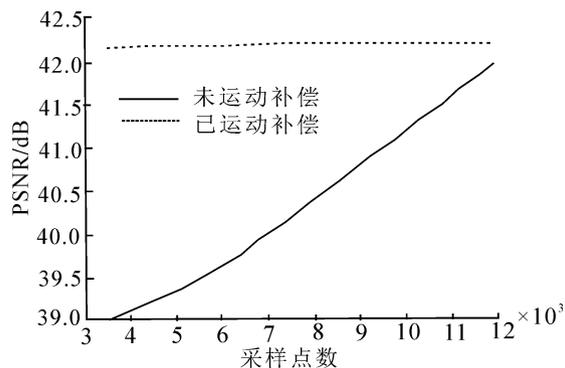


图 6 平均峰值信噪比随采样点数的变化情况

Fig.6 The change of the average PSNR along with the number of sampling

由图 6 可以看出,随着采样点数的增加,2 种情况的平均峰值信噪比都是逐渐增加的,但是经过运动补偿的视频序列的重构质量相对较高,且平均峰值信噪比受采样点数影响较小,具有很好的鲁棒性.

为了验证在不同场景下本文算法的有效性,对背景变化的 foreman(QCIF 格式)的视频序列进行实验,仿真结果如图 7 所示.说明本文算法仍可获得比较好的重构效果,具有一定的普适性.

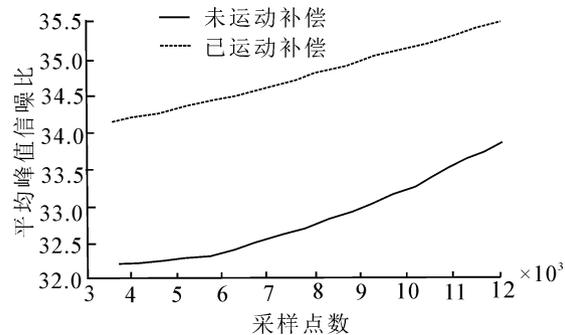


图 7 平均峰值信噪比随采样点数的变化情况

Fig.7 The change of the average PSNR along with the number of sampling

4 结束语

在压缩感知理论的基础上,提出了一种基于帧

间差值和运动补偿的视频压缩感知重构的方法,该方法充分利用了帧与帧之间的连续性和相关性,并进行了实验,获得了较好的实验效果.在实验仿真时,对视频序列的分组比较固定,采样点数比较统一,并且是对整帧图像进行处理.若是采用根据运动剧烈程度自适应分组和自适应采样,并且将帧图像分成子块进行处理,效果会进一步得到优化.

参考文献:

- [1] DONOHO D. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [2] CANDES E. Compressive sampling[C]//Proceedings of the International Congress of Mathematicians. Madrid, Spain, 2006: 1433-1452.
- [3] 岑翼刚,岑丽辉. 基于峰值变换的信号系数表示及重建[J]. 电子与信息学报, 2011, 33(2): 326-331.
CEN Yigang, CEN Lihui. Sparse representation and reconstruction of signals based on the peak transform[J]. Journal of Electronics and Information Technology, 2011, 33(2): 326-331.
- [4] 岑翼刚,陈晓方,岑丽辉,等. 基于单层小波变换的压缩感知图像处理[J]. 通信学报, 2010, 31(8A): 52-55.
CEN Yigang, CHEN Xiaofang, CEN Lihui, et al. Compressed sensing based on the single layer wavelet transform for image processing [J]. Journal on Communications, 2010, 31(8A): 52-55.
- [5] 石光明,刘丹华,高大化. 压缩感知理论及其研究进展[J]. 电子学报, 2009 (5): 1070-1078.
SHI Guangming, LIU Danhua, GAO Dahua. Advances in theory and application of compressed sensing [J]. Acta Electronica Sinica, 2009 (5): 1070-1078.
- [6] BARANIUK R. A lecture on compressive sensing[J]. IEEE Signal Processing Magazine, 2007, 24(4): 118-121.
- [7] DONOHO D L, ELAD M, TEMLYAKOV V N. Stable recovery of sparse overcomplete representations in the presence of noise[J]. IEEE Transactions on Information Theory, 2006, 52(1): 6-18.
- [8] TROPP J A. Greed is good: algorithmic results for sparse approximation[J]. IEEE Transactions on Information Theory, 2004, 50(10): 2231-2242.
- [9] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655-4666.
- [10] FIGUEIREDO M A T, NOWAK R D, WRIGHT S J. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problem[J]. Journal of Selected Topics in Signal Processing: Special Issue on Convex Optimization Methods for Signal Processing, 2007, 1(4): 586-598.
- [11] GILBERT A C, STRAUSS M J, TROPP J A, et al. Algorithmic linear dimension reduction in the l_1 norm for sparse vectors[C]//Proceedings of the 44th Annual Allerton Conference on Communication, Control and Computing. Monticello, USA, 2006: 1-27.
- [12] 翟冲. 基于压缩感知的视频编码技术研究[D].大连:大连理工大学, 2010: 18-20.
Zhai Chong. The research of video coding based on compressive sensing [D]. Dalian: Dalian University of Technology, 2010: 18-20.

作者简介:



李昂,男,1987年生,硕士研究生,主要研究方向为压缩感知算法研究、基于压缩感知理论的图像处理以及视频处理.



岑翼刚,男,1978年生,副教授,博士,主要研究方向为压缩感知、小波分析、信号处理、图像处理,发表学术论文约 40 篇.