

基于自适应分组与压缩感知的分布式视频编码方法

黄涛,熊继平,宣利峰,赵健

(浙江师范大学 数理与信息工程学院,浙江 金华 321004)

摘要: 针对快速场景变换环境中基于压缩感知的分布式视频编码所存在的问题,提出了一种基于自适应动态分组与压缩感知的分布式视频编码方法。对采集的视频帧在基本分组的基础上插入一个自适应分组分离器,通过设定门限阈值实现能够根据视频场景变换的动态调整分组。实验表明,与基本分组方法重构效果相比,采用自适应动态分组方法的重构效果有了较大改进,其峰值信噪比也有大大改善和提高,有效地解决了视频帧中场景变化较快造成的关键信息边缘纹理损失的问题。

关键词: 多媒体传感器网络;分布式视频编码;压缩感知;图像组

中图分类号: TP30116

文献标识码: A

文章编号: 1674-7720(2012)15-0067-03

Method of distributed video coding based on adaptive groups and compressive sensing

Huang Tao, Xiong Jiping, Xuan Lifeng, Zhao Jian

(Department of Information Science and Engineering, Zhejiang Normal University, Jinhua 321004, China)

Abstract: In this paper, we proposed an adaptive grouping and compressive sensing distributed videocoding method to the problem of the current distributed video encoding under fast changing scene. On the basis of the basic grouping method, adaptive grouping separator sets the threshold to achieve the purpose of changing groups dynamically. Experiments show that the adaptive method has been greatly improved compared to the basic grouping method, and the PSNR has significantly improved and enhanced, it can effectively solve the problem of the video frame edge of key information in the rapid scene.

Key words: multimedia sensor networks; distributed video coding; compressive sensing; group of picture

分布式视频编码^[1-2]DVC (Distributed Video Coding) 是多媒体视频传感器网络^[3-4]MVSN (Multimedia Video Sensor Networks) 中视频信号处理常用的一种技术。由于受到带宽和能耗等条件的限制,为了使编码端(多媒体视频传感器节点)对视频信息的编码复杂度降低,传统的基于奈奎斯特采样定理的视频编码标准并不适用于多媒体传感器网络的低成本、低能耗的要求。近年来信号与信息领域的研究热点压缩感知^[5-6]CS (Compressive Sensing) 技术的出现,由于其以全新的采样技术进行采样,实现了采样和压缩同时进行,大大减少了采样率,从而降低了编码端的运算复杂度。Guillemot C^[7]等人提出一种分布式视频编码方法来解决视频信号的编码问题,利用分布式视频编码大大降低了编码端的运算复杂度,实现了编码端的简单、低功率的要求。但这种方法还是采用的传统的奈奎斯特采样,其采集的海量数据使得性能

并不高,如果应用到 MVSN 中将使得编码复杂度极高,实用性差。周燕^[8]等人提出把压缩感知 CS 应用到 MVSN 的视频信息处理中,减少视频信号稀疏分解过程的计算量和存储量,但遗憾的是未采用 DVC 技术,所以重构帧的峰值信噪比 (PSNR) 不高,效果不佳。Pereira F^[9]等人提出一种基于 CS 的分布式视频编码框架,这种方法具有两大优点:(1) 在信号的稀疏处理上,利用每一帧对首帧的差值,得到稀疏视频信号,再利用 CS 低压缩采用率进行采样处理,最后恢复的视频信号效果得到改善;(2) 使 CS 帧在低采样率的情况下也可以良好地恢复重构。然而其缺点也十分明显,一旦视频信号中场景变化较快,其重构视频帧的效果将大打折扣。汪滢^[10]提出了一种基于基本图像组 GOP (Group of Picture) 和压缩感知的分布式视频编码,即把所有采集帧进行 GOP 分组,然后再对每一个 GOP 组分别进行处理,从而在一定程

技术与方法 Technique and Method

度上改善了重构效果,具有一定的场景变换适用性,然而不足之处是分组显得单调固定,不具备灵活变换性。针对以上问题,本文在研究总结前人所做工作的基础上,提出并实现了一种新颖的自适应动态图像组分组的CS分布式视频编码方法,这种方法首先继承了基本分组法的优点,同时避免了基本分组法造成的组数固定、不灵活的缺点,通过设定门限阈值 T 来实现自适应动态调整图像组分组的目的。实验仿真表明:(1)重构后的峰值信噪比(PSNR)得到大大改善和提高,重构效果非常良好;(2)具有场景变化的适用性。

1 基于自适应动态图像分组的CS分布式视频处理机制

本文首先介绍 Pereira F^[9]等人提出的基于差值处理的CS分布式视频编码方法(差值法),以及汪濛^[10]的基本图像组和压缩感知的分布式视频编码方法(基本分组法),最后提出一种自适应分组的CS分布式视频编码方法(动态分组法),并与上面两种方法对比证明所提出的动态分组法的优势。

1.1 差值法与基本分组法

图1所示为基于差值法处理的CS分布式视频编码框架,在这种框架下,首先对第1帧 x_1 标定为关键帧(KEY帧),即基本层。第2帧到最后一帧(x_1, \dots, x_j)标定为WZ帧,然后再对采集的视频帧从第2帧开始,每帧都对第1帧进行作差值处理,其产生的差值($\Delta x_1, \dots, \Delta x_j$)也叫做增强层,同时采用CS技术对数据进行压缩编码处理,在恢复端对首帧基本层采用传统的CS重构技术,对增强层采用联合稀疏解码(JSM-1)模型^[11]进行重构。最后再进行简单叠加运算即可得到重构视频信号。

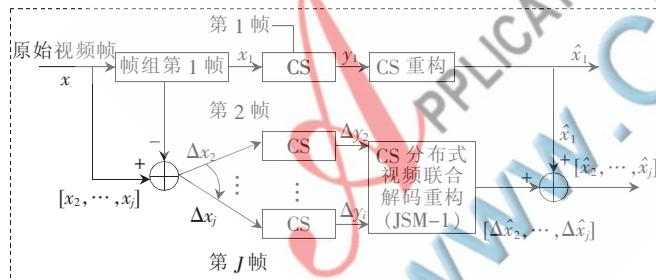


图1 差值法

图2所示为基本图像组分组的原理及框架结构,其基本流程是:首先对原始视频信号进行固定GOP分组,然后再对每一个GOP组分别进行处理,对每一GOP组的帧图像的处理方法采用上面介绍的差值法进行处理,最后再对每一个GOP分组的帧信息进行合并叠加得到总体的重构视频信号。可以看出,用这种基本分组方法得到的每一个GOP,其所含有的视频帧数相对固定,即每一分组的视频帧数相等,因此还是没有实现真正意义上的根据视频场景变换的动态调整GOP分组的目的。

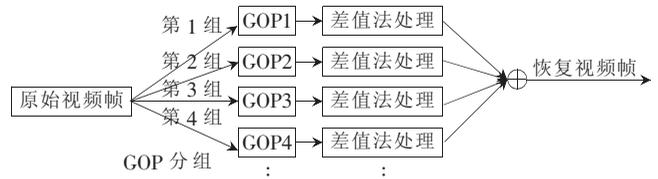


图2 基本分组法

1.2 动态分组法

针对第1.1节基本分组法所造成的GOP组数固定、不灵活的缺点,提出了一种自适应动态图像分组法,简称为动态分组法。图3所示为所提出的基于自适应分组和压缩感知的分布式视频编码框架图,与基本分组方法相比,在基本分组的基础上插入一个自适应分组分离器,目的是为了真正意义上实现能够根据视频场景变换的动态调整GOP分组的目的。



图3 动态分组法

其动态分组分离器的工作原理是在基本分组的基础上,通过设定动态门限阈值 T ,具体设定方法是对差值矩阵中的所有列向量均值求和, $T = \sum |x_t^{(i)} - x_{key}^{(i)}|$,以每一组为基本处理单元,不断更新阈值 T 来改变分组组数。首先设定门限阈值 $T=0$,只要门限阈值 T 有更新则判断视频场景内发生变换,之后从新标定每组的首帧作为关键帧(KEY帧),剩下的为非关键帧(WZ帧),直到下一个场景变换时,再重新定义KEY帧和WZ帧,不断反复循环该步骤直到标定将结束,最后对不同的帧进行后期不同的压缩感知编解码处理。在同一组内,进行CS处理的过程和定义KEY帧和WZ帧的过程是同时进行的。具体算法流程描述如算法1所示。

算法1 自适应分组算法

- (1) Input : $[i=1, x_{key}^{(i)} \leftarrow x_t^{(i)}, GOP \leftarrow 5, T \leftarrow 0$
 $NF \leftarrow \text{Number of Frame}, \Psi \leftarrow \text{sparse basis},$
 $\Phi \leftarrow \text{measurement metric}, \Theta \leftarrow \Phi\Psi]$
- (2) while $i \leq NF$ do
- (3) $\Delta x^{(i)} \leftarrow |x_t^{(i)} - x_{key}^{(i)}|, a \leftarrow \sum \Delta x_t^{(i)},$
 $y = \Phi \cdot \Delta x^{(i)} = \Phi\Psi \cdot \Delta s = \Theta \cdot \Delta s$
- (4) if $i \% GOP == 0$ or $a > T$ do
- (5) $T \leftarrow a, x_{key}^{(i)} \leftarrow x_t^{(i)}, i \leftarrow i+1$
- (6) end if
- (7) solve $\Delta \hat{s} = \min_s \|\Delta s\|_l \text{ s.t.}$

《微型机与应用》2012年第31卷第15期

$$y = \Phi \Psi \cdot \Delta s$$

$$(8) x_i^{(i)} \leftarrow \Delta \hat{s} + x_{key}^{(i)}$$

(9) end while

这样所得的每一组 GOP 组数由实际场景变换来决定,从而实现自适应动态调整图像分组的目的,分组后的后期处理方法和基本分组法相同,在此不再赘述。

2 仿真结果及分析

实验仿真基于 2.79 GHz CPU, 512 MB 内存的个人计算机环境下的 Matlab 7.0 实验平台。实验的测试视频信号为测试序列 foreman_cif.yuv^[12], 每帧尺寸大小 176×144, 采用序列的前 50 帧来进行实验, 分别利用文中介绍的三种不同的方法进行仿真测试。测试中在编码端利用 CS 技术进行压缩采样, 其测量率对 KEY 帧采样率取高采样率 0.7, 对 WZ 帧取采样率 0.2 进行处理, 设 GOP=5, 解码端重构方法用的是梯度投影算法 (GPSR), 如图 4 所示为 foreman 序列重构后各帧峰值信噪比。

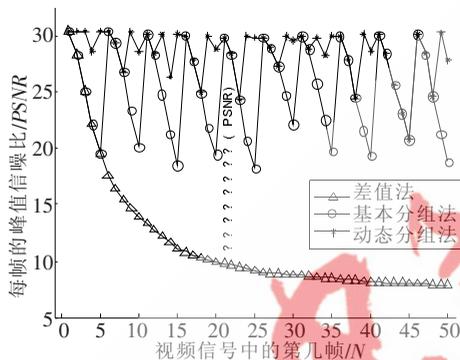


图4 重构后各帧峰值信噪比对比图

对每一帧视频信号, 首先从横向看, 使用差值法随着视频帧数增大, PSNR 值逐渐下降, 效果比较差。比较基本分组法和动态分组法, (1) 首先从分组上进行比较, 由于设置的 GOP=5, 所以基本分组每一个 GOP 就比较固定, 自始至终都以 5 个一组划分开来, 从而在分组上单调呆板, 不具有灵活性, 而所提出的动态分组法在分组上就显得动态灵活了很多; (2) 再从重构效果 (PSNR) 上进行比较, 虽然少数几个点的 PSNR 相同外, 动态分组法整体效果的 PSNR 值要明显高于基本方法的整体效果的 PSNR 值, 尤其在 WZ 帧上, 动态分组的主要目的就是改善 WZ 帧的重构效果。

差值方法模糊化了图像的重要关键信息, 所以重构出来的效果就很差, foreman 序列是人物表情变化 (即场景变化) 比较大序列, 所以差值法和基本分组方法整体重构效果不理想, 因此证明了所提出的动态分组方法具有良好的场景变化适应性。

从主观评价上来看, 由于第 30 帧图像恰好为 WZ 帧, 比较用上面叙述的三种不同的方法对第 30 帧进行重构, 得到的峰值信噪比 PSNR 值, 如表 1 所示, 可以看出所提出动态分组方法的优势, 其 PSNR 远远大于前两

种。再比较图像的纹理视觉效果, 如图 5 所示。

表 1 第 30 帧图像的 PSNR 比较

处理方法	差值法	基本分组法	动态分组法
foreman	8.8319	22.112	29.588



(a) 原始帧 (b) 差值法重构



(c) 基本分组法重构 (d) 动态分组法重构

图5 foreman 序列第 30 帧重构后的效果

从图 5 中可以看出, (d) 帧图像的图像纹理均比 (b)、(c) 好。首先与 (b) 比较, 可以发现 (b) 图像中效果特别差, 里面的关键人物几乎不见了, 这样就失去了图像的重要信息了; 再与 (c) 比较, 发现 (c) 图像中人物边缘出现白斑, 由于人物是不断运动的, 具有动态特征, 而基本分组并不具备这种动态特性, 所以最后重构的效果会比较差; 从 (d) 中可以看出采用本文所提出的基于自适应分组方法所重构的视频帧 (d) 图像的效果良好, 人物也清晰可见, 边缘纹理也比较理想。同样测试 hall 序列等也有类似的效果。

利用本文提出的基于自适应分组与压缩感知的分布式视频编码方法, 重构后的视频帧整体效果良好, 峰值信噪比大大改善, 能够适用于视频场景变化剧烈的场合。门限阈值的选择是本文自适应算法的关键, 本文中采用的是简单但行之有效的求列向量均值的方法设置门限阈值, 下一步的研究方向可以针对具体的应用场景在自适应算法复杂性和重构效率之间权衡, 选择其他智能算法来优化阈值, 进而达到更完善的分组目的。

参考文献

- [1] GIROD B, AARON A, RANE S. Distributed video coding [J]. Proceedings of the IEEE Special Issue on Advances in Video Coding and Delivery, 2005, 93(1): 71-83.
- [2] 王尊亮, 李学俊. 分布式视频编码技术研究进展 [J]. 计算机工程与设计, 2010, 31(3): 31-34.
- [3] 鲁琴. 无线多媒体传感器网络多节点协同图像采集与压缩技术研究 [D]. 北京: 国防科学技术大学, 2009.
- [4] <http://www.eefocus.com/article/07-12/4155121112071731.html>.
- [5] GUILLEMOT C. Distributed monoview and multiview video coding: basics, problems and recent advances [J]. IEEE Signal Processing Magazine, 2007, 24(5): 67-76.
- [6] DONOHO D. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.

[7] CANDÉS E, WAKIN M. A sensing paradigm that goes against the common knowledge in data acquisition[J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30.

[8] 周燕,王东,钟勇.压缩传感在无线视频监控中的应用研究[J].计算机应用研究, 2010,27(6): 73-75.

[9] PEREIRA F. Selecting the most promising application scenarios [J]. Signal Processing: Image Communication, 2008, 23(9):339-352.

[10] 汪滢.压缩感知在视频编码中的应用研究[D].北京:北京邮电大学,2010.

[11] 蔡述庭,王钦若.视频传感网中鲁棒多视角分布式视频

编码[J].计算机工程,2009,35(24):219-221.

[12] <http://trace.eas.asu.edu/yuv/>

(收稿日期:2012-02-27)

作者简介:

黄涛,男,1986年生,在读研究生,主要研究方向:图像处理,压缩感知。

熊继平,男,1982年生,副教授,主要研究方向:视频传感器网络、压缩感知。

宣利峰,男,1987年生,在读研究生,主要研究方向:分布式视频编码。

