

无线多媒体传感器网络中基于多类传感器信息的分布式视频处理

王玉斐¹⁾ 王汝传^{1), 2)} 黄海平¹⁾ 孙力娟¹⁾

¹⁾(南京邮电大学计算机学院, 南京 210003) ²⁾(南京大学计算机软件新技术国家重点实验室, 南京 210093)

摘要 无线多媒体传感器网络中的多媒体传感器节点的功率、计算和存储能力及带宽资源受限, 迫切需要低复杂度的视频处理技术对大数据量的视频进行压缩传输。基于来自网络中监控同一目标场景的多种类型传感器的信息, 给出了一个新的无线多媒体传感器网络分布式视频处理系统框架; 提出了一种基于多类型传感数据融合和多视角的 GOP (group of picture) 划分方法; 在解码端, 考虑单一视角视频序列之间较强的时间相关性, 产生时间相关的边信息; 利用来自多个视频传感节点的视频序列间的空间多视角相关性, 产生多视角相关的边信息, 并提供两种边信息的融合和选择机制, 提高边信息的准确度和可靠性。最后仿真实验结果表明该方法的有效性和优越性。

关键词 无线多媒体传感器网络 分布式视频编码 多类传感器 图片组 (GOP) 信息融合 边信息

中图分类号: TP393.0 文献标志码: A 文章编号: 1006-8961(2010)01-0161-06

Multimodal Sensors Information Based Distributed Video Processing for Wireless Multimedia Sensor Networks

WANG Yu-fei¹⁾, WANG Ru-chuan^{1), 2)}, HUANG Hai-ping¹⁾, SUN Li-juan¹⁾

¹⁾(College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003)

²⁾(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract In wireless multimedia sensor networks (WMSN), sensor devices are constrained in terms of power, processing memory and bandwidth capability. Low-complexity video processing technique is highly desired to encode and transmit large numbers of video data. Based on the information from multimodal sensors which monitor the same scene, a new framework of distributed video processing system for WMSN is presented. A group of pictures (GOP) partition algorithm based on multimodal data and multiview is also put forward. In the decoder, considering the temporal correlation between adjacent frames from single view, temporal side information is generated. Multiview side information is acquired by using spatial correlation among the frames from adjacent video sensors. In order to improve the prediction accuracy and reliability, a mechanism to fusion and select two kinds of side information is provided. The validity and advantage of the proposed method are verified through simulations.

Keywords wireless multimedia sensor networks, distributed video coding, multimodal sensors, group of pictures (GOP), information fusion, side information

基金项目: 国家自然科学基金项目 (60573141, 60773041, 60973139); 国家高技术研究发展计划 (863) 基金项目 (2006AA01Z201, 2006AA01Z219); 南京市高科技项目 (2007 软资 106); 现代通信国家重点实验室基金项目 (9140C1105040805); 江苏省博士后基金项目 (0801019C); 六大高峰人才计划和江苏高校科技创新计划基金项目 (CX08B-085Z, CX08B-086Z)

收稿日期: 2009-02-23 改回日期: 2009-03-11

第一作者简介: 王玉斐 (1982—), 男。2006 年入南京邮电大学攻读信息网络专业博士学位。主要研究方向为无线多媒体传感器网络等。E-mail: wyf_1999@163.com

0 引言

无线多媒体传感器网络^[1] (WMSN) 是由无线互连的设备组成的网络, 能够无所不在地从监测环境中获取多媒体信息, 如视频和音频、静态图像及标量传感数据等。无线多媒体传感器网络视频编码技术的基本设计目标是高的压缩效率, 有效地限制带宽和能量消耗, 并具有低复杂度和低功率, 从而可以嵌入传感器设备并延长节点的生存时间。

为了取得较高的压缩效率, 使用帧内压缩减少帧内冗余, 帧间压缩 (预测编码或运动估计) 利用序列帧之间的相关性减少传送和存储的数据量, 达到较好的率失真性能; 但是由于运动估计具有较大的计算复杂度, 需要复杂的编码器、强大的处理算法和较高的能量消耗, 目前典型的视频压缩标准, 如 MPEG, H. 263^[2], H. 264^[3], 编码器的复杂度是解码器的 5 到 10 倍^[4], 并不适用于无线多媒体传感器网络中节点低成本、低能耗的要求, 必须研究更加简单但仍具有高压缩效率的编码技术。建立在 Slepian 和 Wolf 的无损编码码率界限^[5] 以及 Wyner 和 Ziv 的在解码端使用边信息的有损编码码率界限^[6] 的理论基础上, 分布式信源编码对相互之间不进行通信的多个相关传感器输出进行联合编码, 编码器简单, 低功率, Sink 节点的解码器具有较高的复杂度, 并承担大部分的处理和能量负担, 非常适合用于解决无线多媒体传感器网络中的多媒体信息编码, 尤其是视频编码。因此, 全新分布式视频编码^[7] 方法被用来解决无线多媒体传感器网络中的视频编码问题。

文献 [8-9] 提出了一种实用的像素域帧内编码和帧间解码相结合的 Wyner-Ziv 编码器, 相比而言, 其编码非常简单, 编码器仅需要两个反馈移位寄存器和一个插帧, PSNR (峰值信噪比) 比传统的帧内编码大约低 2~5 dB, 但该方案依赖于在线反馈, 不适合于视频被编码存储以备后续使用的场合, 同时在多跳网络中会引入额外的延迟。文献 [10] 提出了一种变换域 Wyner-Ziv 编码, 对每个 Wyner-Ziv 帧进行分块 DCT (离散余弦变换), 其率失真性能介于传统帧内变换编码和运动补偿变换编码之间。斯坦福大学的 Aaron 等人提出了基于 RCPT (rate compatible punctured turbo code) 分布式视频编码系统^[7], UC 伯克利分校的 Puri 和 Ramchandran 设计

了 PRISM (Power-efficient robust high-compression syndrome-based multimedia coding) 分布式视频编码系统^[7], 编码器基于当前帧和先前帧之间的帧能量差选择编码模式, 解码器的灵活性是受限的。这些分布式视频编码技术仅限于讨论编解码理论本身, 并没有考虑无线多媒体传感器网络中多种类型数据共存的特征, 没有充分利用网络中其他相关数据对分布式视频编码进行优化。

本文综合考虑 WMSN 中存在的多种类型传感器, 以及视频传感器的时间和空间关系, 研究适合于无线多媒体传感器网络的分布式视频编码方法。提出一种基于多传感器信息融合和多视角的 GOP (group of pictures) 划分机制; 在解码端, 综合利用视频序列之间的时间相关性和空间多视角相关性, 产生边信息。

1 基于多传感器信息融合的分布式视频处理机制

1.1 基于多传感器信息融合的分布式视频处理系统

本文基于多传感器信息融合的分布式视频处理系统, 如图 1 所示, GOP 的大小由基于多传感数据融合和多视角的 GOP 划分策略确定, 在编码端输入的原始视频被划分为固定大小的 GOP。每个 GOP 都由一个关键帧 (Key 帧) 开始, 余下的为 Wyner-Ziv 帧。

在编码端, 输入的视频帧被分成 Key 帧和 Wyner-Ziv 帧, Key 帧采用传统的帧内编码方法进行编码。这里, 使用 H. 264/AVC 帧内编码。对于 Wyner-Ziv 帧, 没有使用 DCT 将每个块由像素域变换到频域, 而是对每个像素值以步长 2^l 统一量化, 量化系数为 q 。虽然 DCT 能够很好地利用块的空间相关性, 但会使编码复杂度提高。然后进行 Turbo 编码。

在解码端, Key 帧采用 H. 264/AVC 帧内解码方法进行解码。对于 Wyner-Ziv 帧, 首先使用边信息产生与融合机制, 从时间、空间多视角两个角度出发, 产生边信息, 记为 \hat{S} 。Turbo 码解码器利用校验位和边信息进行解码, 如果 Turbo 码不能可靠地解码则向编码端的缓存请求更多的校验位, 直到解码错误率达到一定的阈值。接收端解码出量化系数 q' , 然后结合产生的边信息重构出 Wyner-Ziv 帧。

基于多传感器信息融合的分布式视频处理系统的改进之处在于: 1) GOP 大小的动态分配策略, 根

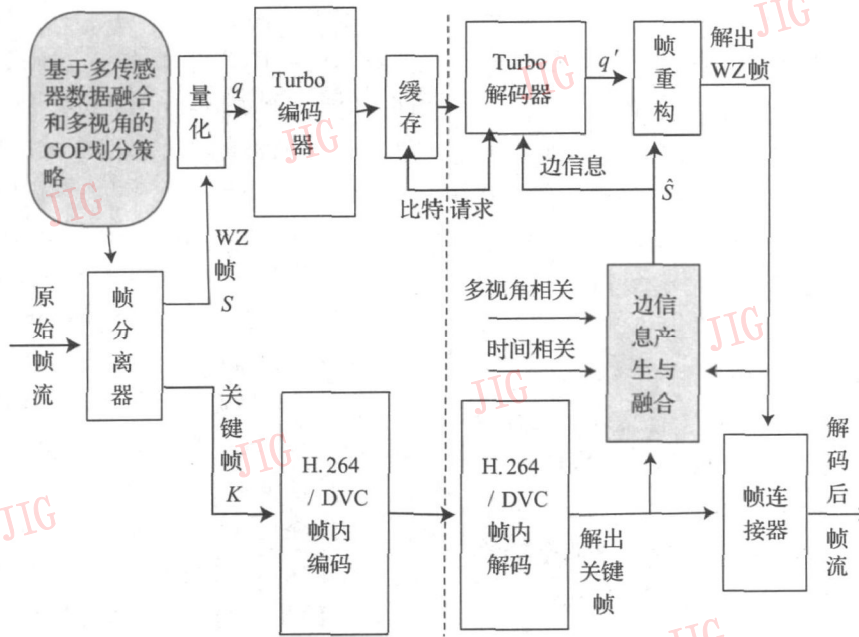


图 1 基于多传感器信息融合的分布式编解码

Fig 1 Distributed coding based on multimodal sensors information fusion

据目标监测场景的其他传感数据信息和邻近视频传感器节点之间的位置信息, 动态生成各视频传感器采集的视频流的分组大小, 调整 Key 帧和 Wyner-Ziv 帧的编码频率, 通过减少 Key 帧, 节省有限的节点资源; 2) 在解码端, 综合考虑不同 GOP 大小情况下, 时间相关边信息的产生, 空间多视角相关性边信息的产生, 以及两者的融合。

1.2 基于多类传感器数据融合和多视角的 GOP 划分

根据标量传感数据监测的目标场景信息及视频传感器节点的位置信息, 制定相关视频传感节点的视频序列的 GOP 划分机制, 每个视频传感器节点将捕捉到的视频序列划分为几个 GOP_s 。当 GOP 大小为 1 时, GOP 由 1 个 Key 帧组成, 当 GOP 为 2 时, GOP 由 1 个 Key 帧和 1 个非 Key 帧即 Wyner-Ziv 帧组成, 当 GOP 为 3 时, GOP 由 1 个 Key 帧和 2 个 Wyner-Ziv 帧组成, 即当 GOP 为 n 时, GOP 由 1 个 Key 帧和 $n-1$ 个 Wyner-Ziv 帧组成。

GOP 的大小由 GOP 划分机制确定, 根据场景内的物理信息, 如加速度、磁力、温度、湿度、光度及视频传感器的物理位置等动态确定监控同一目标场景的一组视频传感器节点编码时对视频序列划分时的 GOP_s 大小, 而视频传感器节点编码时根据分配给自己的 GOP_s 大小划分感知的视频序列, 由于 Key 帧采取的是传统的编解码方法, 将会有较大的复杂度和资源消耗, 因此, 可以根据物理场景变化程度、目

标物体是否出现及其移动速度及一组视频传感节点的物理位置和监控视角, 动态调整和分配 GOP 的大小, 在尽量不影响监控效果的情况下, 减少需要更多编码资源的 Key 帧数量, 降低资源消耗, 节省能量。下面以监控一个物理场景中移动目标的无线多媒体传感器网络为例, 描述基于多类传感数据融合和多视角的 GOP 划分算法。算法的输出是一个 $1 \times n$ 的矩阵, n 表示邻近视频传感器的数量, 矩阵中的每个元素对应于一个视频传感器节点的 GOP 大小。假设监控环境中部署有声波传感器、雷达波传感器和视频传感器, 声波传感器的数量为 4 个, 记为 SW_0, SW_1, SW_2, SW_3 ; 雷达波传感器的数量为 2 个, 记为 RW_0, RW_1 ; 在任一时刻保证有 3 个邻近的视频传感器节点 W_0, W_1, W_2 位置固定, 监控同一移动目标, 并且场景中只存在单个移动目标, 声波传感器和视频传感器之间可以相互通信, 交换信息。

GOP 划分算法:

- 1) 若场景内没有移动目标, 则返回 $[4 \ 4 \ 4]$;
- 2) 若场景内出现单个移动目标, 4 个声波传感器节点 SW_0, SW_1, SW_2, SW_3 , 以 SW_0 为基准, 计算出其他 3 个节点到目标的距离, 然后根据三角形算法计算出移动目标的位置;
- 3) SW_0 从接收 W_0, W_1, W_2 广播的位置信息 (以 SW_0 为基准), 判断出当前观测移动目标的最佳视角视频传感器节点, 并设置其 GOP 大小为 1 或 2 (根据需要的监控质量而定);

4) 雷达波传感器节点 RW_0, RW_1 , 测量移动目标反射波的相位差, 根据运动学原理确定移动目标的速度和运动方向, 然后, 根据视频传感器节点的位置信息, 设置处于目标运动方向上的非最佳视角传感器节点 GOP 大小为 2 设置处于目标运动反方向上的非最佳视角传感器节点 GOP 大小为 3 或 4 (根据目标的移动速度和监控质量需求确定);

5) 返回 GOP 大小分配矩阵。假设在无线多媒体传感器网络中, 有 n 个邻近的视频传感器节点 WN 监控同一目标场景, $n \geq 3$ 。对于每个视频传感器节点 $WN_i (i=1, 2, \dots, n)$, 视频序列根据分配的 GOP 大小 GOP_Size_i 被划分为若干 GOPs。 GOP_Size_i 由基于多类型传感数据融合和多视角的帧分离调度策略动态确定。当 $n=4, GOP_Size_1=2, GOP_Size_2=3, GOP_Size_3=2, GOP_Size_4=3$ 时, 多视角分布式视频编码的 GOP 结构如图 2 示。

		空间相关			
		WN_1	WN_2	WN_3	WN_4
时间相关	T	Key	Key	Key	Key
	T+1	Wyner-Ziv	Wyner-Ziv	Wyner-Ziv	Wyner-Ziv
	T+2	Key	Wyner-Ziv	Key	Wyner-Ziv
	T+3	Wyner-Ziv	Key	Wyner-Ziv	Key
	T+4	Key	Wyner-Ziv	Key	Wyner-Ziv
	T+5	Wyner-Ziv	Wyner-Ziv	Wyner-Ziv	Wyner-Ziv
...	

图 2 多视角分布式视频编码 GOP 结构图

Fig 2 GOP structure of multiview distributed video coding

1.3 边信息的产生与融合

1.3.1 基于时间相关性的边信息

参考文献 [11] 中的方法, 使用当前待解码帧的 previous 帧和 forward 帧产生基于时间相关性的边信息。计算从前一帧到后一帧的基于块的运动向量。然后, 在中间点插入运动向量来产生边信息, 如图 3 所示。

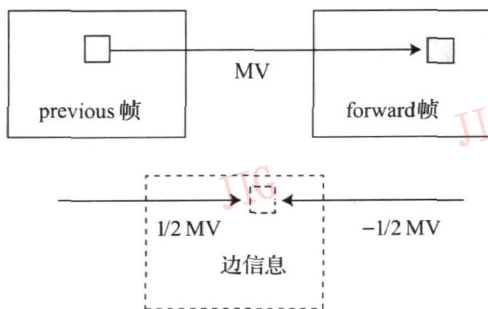


图 3 时间相关边信息的产生^[11]

Fig 3 The generation of time-related side information^[11]

由于编码时, 将帧分为 Key 帧和 Wyner-Ziv 帧, 同时不同的视频传感器节点所采用的 GOP 大小不同, 如图 2 所示, 因此, 这里基于时间相关性的边信息的产生机制有以下几种情况:

- 1) Key-Key-Side-Information (KCSI): previous 帧和 forward 帧取距离当前帧最近的 Key 帧;
- 2) WZ-Key-Side-Information (WKSI): previous 帧取当前帧的前一帧, 是 Wyner-Ziv 帧, forward 帧取距当前帧最近的 Key 帧;
- 3) KeyWZ-Side-Information (KWSI): previous 帧取距当前帧最近的 Key 帧, forward 帧取当前帧的最后一帧, 是 Wyner-Ziv 帧。

产生的边信息的准确性取决于当前帧和两个参考帧之间的时间相关程度, 以及两个参考帧本身的质量。但相对与 Key 帧而言, Wyner-Ziv 帧的质量较差。在当前帧紧邻的前一帧和最后一帧均为 Wyner-Ziv 帧的情况下, 为了保证边信息的准确性, 情况 2) 和情况 3) 取其中的一个参考帧为距离当前帧最近的一个 Key 帧。情况 1) 相对于情况 2) 和情况 3) 具有更高的参考帧质量, 但当 GOP 较大时, 两个 Key 帧之间的时间距离较大, 时间相关程度较小。3 种边信息产生的方法需要根据具体情况选择使用, 或进行融合后使用。

1.3.2 基于多视角的边信息

多个邻近视频传感器节点往往从不同角度和不同位置监控同一目标场景。以其中一个视频传感器节点为中心, 其他邻近的视频传感器节点的视频可以作为辅助信息。由于这种多视角相关性, 可以利用同一时刻来自邻近视频传感器节点的视频序列, 产生边信息, 这里假设无线多媒体传感器网络中所有节点时间同步。使用现有多视角视频编码机制中广泛使用的全局运动模型^[11]来表示来自邻近视频传感器节点的帧的不同。

通过计算与不同视角对应的邻近视频传感器节点之间的单应性矩阵产生基于多视角的边信息。假设在无线多媒体传感器网络中, 视频传感器节点的位置和视角固定, 则单应性矩阵只需要计算一次, 与运动估计相比, 具有较低的复杂度。对于监控同一目标场景的邻近传感器节点 WN_i 和 WN_j , 采用的是透视运动模型, WN_i 的帧中的每个像素点 (x_1, y_1) 通过下面的转换关系映射到 WN_j 的像素点 (x_2, y_2) :

$$x_2 = \frac{(a_2 + a_0x_1 + a_1y_1)}{(a_6x_1 + a_7y_1 + 1)} \quad (1)$$

$$y_2 = \frac{|a_5 + a_3x_1 + a_4y_1|}{|a_6x_1 + a_7y_1 + 1|} \quad (2)$$

式中, a_0, a_1, \dots, a_7 是 8 个运动参数。当 $a_6 = a_7 = 0$ 时, 模型为仿射几何转换; 当 $a_0 = a_4 = 1, a_1 = a_3 = a_6 = a_7 = 0$ 时, 是一个纯粹的转换; 当 $a_0 = a_4, a_1 = -a_3, a_6 = a_7 = 0$ 时, 模型是转换—缩放—旋转; 当 $a_0 = a_4, a_1 = a_3 = a_6 = a_7 = 0$ 时, 模型是转换—缩放。参数的计算采用梯度下降方法。

1.3.3 边信息的融合

在无线多媒体传感器网络中, 单个传感器节点的视频序列具有时间相关性, 而来自多个监控同一目标场景节点的视频序列之间则存在着较强的空间多视角相关性。根据本文所描述的基于多传感器信息融合的分布式视频处理系统, 在解码端, 一个 Wyner-Ziv 帧通常具有大于等于 4 个的参考帧: 2 个来自同一视频传感器的时间相关的帧, 即 previous 帧和 forward 帧; 其他的是来自邻近视频传感器的帧, 多视角相关的帧的个数取决于监控同一目标场景的视频传感节点的数量, 一般情况下为 2 或 3。

边信息的准确性很大程度影响到分布式视频编解码的性能和质量。同时利用基于时间相关性和基于多视角相关的边信息, 提出一种灵活的可扩展的边信息融合机制, 来提高边信息的可靠性和准确性, 如图 4 所示。

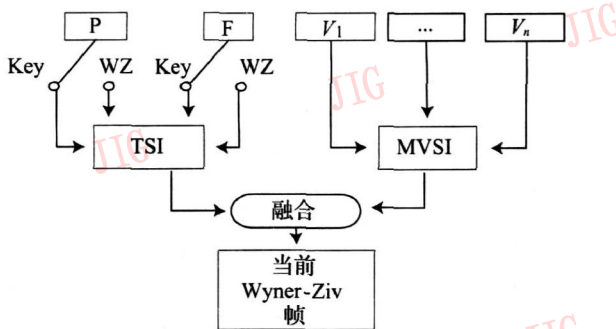


图 4 边信息的融合

Fig 4 Fusion of side information

如图 4 所示, P 和 F 代表时间上和当前 Wyner-Ziv 邻近的来自同一视频传感器节点的 previous 帧和 forward 帧; V_1, \dots, V_n 代表来自邻近视频传感器节点的不同视角的帧, 一般情况下, n 取 2 或 3 从多视角观测同一目标场景的角度出发, 若 $n > 3$ 则会产生较大的冗余, 增加资源消耗。TSI 代表由时间相关的 P 帧和 F 帧预测得到的边信息。MVSII 是由空间多视角相关的 V_1, \dots, V_n 预测得到的边信息。解码当前 Wyner-Ziv 帧时可以选择使用 TSI 和 MVSII 或者两者的融合。

使用 P 帧和 F 帧进行预测, 根据目标场景所产

生的视频序列的运动级别, 选择使用 Key 帧和 Wyner-Ziv 帧, 当运动级别较高时, 连续帧之间具有更高的时间相关性, 所以 P 帧和 F 帧均选择和当前帧紧邻的帧: GOP 大小为 2 的情况下, P 帧和 F 帧均为 Key 帧; GOP 大小等于 3 时, 一个为 Key 帧, 一个为 Wyner-Ziv 帧; GOP 大小大于等于 4 时, 存在与 Wyner-Ziv 帧紧邻的均为 Wyner-Ziv 帧的情况, 此时, 为了保证预测的准确性和可靠性, P 帧和 F 帧其中一个选择与当前帧具有最小时间间隔的 Key 帧, 另一个仍然选择紧邻的 Wyner-Ziv 帧。当运动级别较低时, 连续帧之间的时间相关性相对较小, P 帧和 F 帧均选择与当前帧具有最小时间间隔的 Key 帧, 以提高预测信息的可靠性。使用 V_1, \dots, V_n 帧, 基于全局运动模型计算并预测产生多视角相关的边信息 MVSII。显然, TSI 和 MVSII 具有不同的预测准确度和可靠性。因此, 应该根据监控目标场景的具体情况选择其中之一, 必要时, 使用数据融合算法将二者融合产生更准确和可靠的边信息。也可以设置阈值判断运动级别, 从而选择更准确的边信息, 或使用两种边信息的融合结果。

2 仿真结果

使用 breakdancers^[12] 测试序列进行仿真, 每秒 15 帧 (15 fps), 采用 QCIF (quarter common intermediate format) 分辨率 176×144 。视频传感器节点 W_1, W_2, W_3, W_4 分别对应 breakdancers 中的 camera0, camera1, camera2 和 camera3。图 5 为 GOP 大小对分布式视频编码率失真性能的影响, 解码时边信息只使用 TSI 参考帧选择与当前帧紧邻的 previous 帧和 forward 帧。

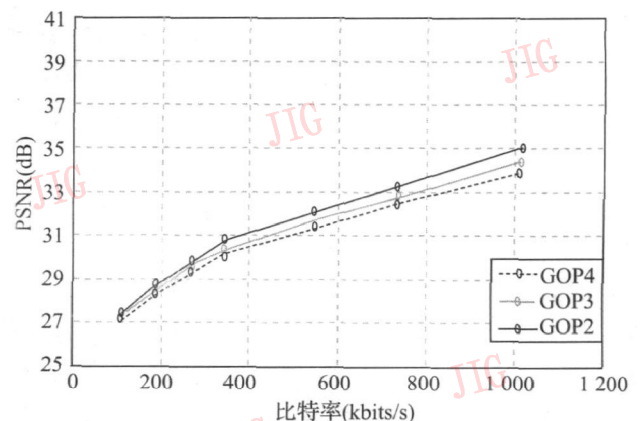


图 5 GOP 大小对率失真的影响

Fig 5 The influence to rate distortion by GOP size

从图 5 可以看出 GOP 较大时, 视频序列中连续的 Wyner-Ziv 帧就较多, Wyner-Ziv 帧相对 Key 帧具有更小的编码复杂度, 但在解码时提供的边信息的准确性和可靠性较差, 影响到编码性能。基于无线多媒体传感器网络中其他类型传感器提供的信息, 为能够从最佳角度监控目标的视频传感器分配较小的 GOP, 为邻近的视频传感器分配相对较大的 GOP, 则可以在保证监控质量的基础上, 在非最佳视角传感器节点上降低 Key 帧的编码频率, 节省能量。

图 6 为采用 TSI, MVSI 和融合边信息 (或逻辑) 时分分布式视频编码率失真性能的比较。所有的曲线均为 4 个 cameras 的平均值。从图 6 可以看出, TSI 和 MVSI 的使用, 提高了分布式视频编码的性能。

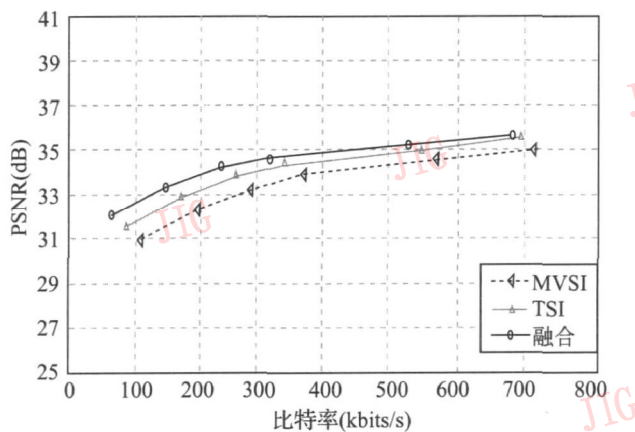


图 6 解码端使用边信息的率失真

Fig 6 The rate distortion of decoding with side information

3 结 论

考虑在一个多媒体传感器网络监控应用中, 可能存在着多种类型的传感器, 提出了一种适合于 WMSN 的分布式视频处理系统。在此框架内, 采用一种动态设置一组邻近视频传感器的 GOP 大小的机制, 根据监控目标的变化, 改变每个传感器编码时的 Key 帧和 Wyner-Ziv 帧的数量, 在保证监控效果的前提下, 从数量上减低了编码资源消耗。在不同 GOP 大小的情况下, 产生时间相关边信息和多视角空间相关边信息, 并根据需要对两者进行融合。

但是根据标量传感数据监测的目标场景的信息及视频传感器节点的位置信息, 制定相关视频传感器节点的视频序列的 GOP 划分机制依赖于其他类型传感器节点返回的监控场景信息的准确性, 以及多

类型传感器节点与视频编码传感器节点之间的协同。因此, 如何提高监控场景参数信息以及多类型传感器节点之间协作的准确性将是下一阶段重点解决的问题, 同时, 在解码端进行边信息融合时, 采取更加灵活的智能策略, 进一步提高性能。

参考文献 (References)

- [1] Akyildiz I F, M elodia T, Chowdhury K R. A survey on wireless multimedia sensor networks [J]. Computer Networks 2007, 51(4): 921-960
- [2] ITU-T. H. 263 Video Coding for Low Bit Rate Communication [EB/OL]. [2008-02-16] <http://www.itu-int/rec/T-REC-H.263/e>
- [3] ITU-T. H. 264 Advanced Video Coding for Generic Audiovisual services [EB/OL]. [2008-02-16]. <http://www.itu-int/rec/T-REC-H.264/e>.
- [4] Girod B, Aaron A M, Rane S et al. Distributed video coding [J]. Proceedings of the IEEE, Special Issue on Video Coding and Delivery, 2005, 93(1): 71-83
- [5] Slepian D, Wolf J K. Noiseless coding of correlated information sources [J]. IEEE Transactions on Information Theory, 1973, 19(4): 471-480
- [6] Wyner A D. On source coding with side information at the decoder [J]. IEEE Transactions on Information Theory, 1975, 21(3): 294-300
- [7] Puri R, Majumdar A, Ishwar P, et al. Distributed video coding in wireless sensor networks [J]. IEEE Signal Processing Magazine, 2006, 23(4): 94-106
- [8] Aaron A, Rane S, Zhang R, et al. Wyner-Ziv coding for video Applications to compression and error resilience [C] // Proceedings of IEEE Data Compression Conference. Snowbird Utah, USA: IEEE Computer Society Press, 2003, 93-102
- [9] Aaron A, Setton E, Girod B. Towards practical Wyner-Ziv coding of video [C] // Proceedings of IEEE 2003 International Conference on Image Processing, Barcelona, Spain, IEEE Press, 2003, 869-872
- [10] Aaron A, Rane S, Setton E, et al. Transform-domain Wyner-Ziv Codec for Video [EB/OL]. [2008-03-12]. <http://www.stanford.edu/~bgirod/pdfs/AaronVCIP04.pdf>
- [11] Quaret M, Dufaux F, Ebrahimi T. Fusion-based multiview distributed video coding [C] // Proceedings of ACM International Workshop on Video Surveillance and Sensor Networks, Santa Barbara, California, USA: ACM Press, 2006, 139-144
- [12] Microsoft. Breakdancing Sequence [EB/OL]. [2008-04-22]. <http://research.microsoft.com/en-us/um/people/sbkang/3dvideodownload/>.