

一种空间域 Wyner-Ziv 视频编码系统的性能改进算法

干宗良, 齐丽娜, 朱秀昌

(南京邮电大学江苏省图像处理与图像通信重点实验室, 江苏南京 210003)

摘 要: 分布式视频编码是建立在 Slepian-Wolf 和 Wyner-Ziv 信息编码理论基础上的全新视频编码框架, 具有编码复杂度低, 编码效率较高, 抗误码性能好的特点. 本文首先简单介绍了一种典型的分布式视频编码实现方案——空间域 Wyner-Ziv 视频编码, 随后提出一种空间域 Wyner-Ziv 视频编码系统的性能改进算法, 该算法在不增加编码复杂度的基础上, 在解码端利用双向运动估计预测获取更高质量的边信息, 同时采用基于 Huber-Markov 随机场约束的联合迭代解码算法重建图像. 实验结果表明, 在相同的输出码流情况下, 本文改进算法在解码端重建图像的峰值信噪比与空间域 Wyner-Ziv 视频编码算法相比平均提高 2dB, 并且主观效果有所改善.

关键词: 分布式视频编码; 空间域 Wyner-Ziv 视频编码; 边信息; 运动估计; Huber-Markov 随机场

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112(2007)10-2014-05

A Novel Performance Improved Method for Pixel Domain Wyner-Ziv Video Coding System

GAN Zong-liang, QI Li-na, ZHU Xiu-chang

(Image Processing and Image Communication Key Laboratory, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China)

Abstract: Distributed video coding (DVC) is a new video compression paradigm based on two key information theories: Slepian-Wolf and Wyner-Ziv theorems, and it has the advantage of very lightweight encoder, high compression efficiency and inbuilt robustness. In this paper, pixel domain Wyner-Ziv video algorithm, which is a typical DVC scheme, is analyzed at first. Then a performance improved method for pixel domain Wyner-Ziv video coding system is presented. The proposed method, without sacrificing the encoder complexity, can calculate high quality side information using bi-directional motion estimation and reconstruct decoding frames using side information by joint iteration function based on Huber-Markov random field. Simulation results show that a better visual quality and 2dB average gain of Power Signal-to-Noise Ratio (PSNR) can be achieved with the proposed method at the decoder.

Key words: distributed video coding; pixel domain Wyner-Ziv video coding; side information; motion estimation; Huber-Markov random field

1 引言

近年来, 一些新的视频应用需要低复杂度的编码器, 而在解码端可以具有较高复杂度的解码器, 如计算能力、内存容量、耗电量都受限的无线视频终端: 无线视频监控、无线 PC 相机、移动视频电话等. 引起较多关注的无线传感网络 (wireless sensor network) 中的视频传感节点 (VSN: video sensor node) 有两个基本的要求: (1) 要求无线视频编码器功耗低、复杂度低, 这对延长 VSN 的使用寿命至关重要; (2) 由于传输速率的限制, 要求编码器具有较高的压缩效率^[1]. 传统的视频编码技术 (如 H. 26x, MPEG) 不再适用于上述应用场合, 必须寻找新的编码压缩方法. 针对这些情况, 一种全新的视频编码框

架——分布式视频编码 (DVC: Distributed Video Coding) 开始受到关注, 其特点是编码简单, 解码较复杂, 并且能够实现较为高效的压缩, 重建图像质量能够接近传统的编码方式, 抗误码特性好.

分布式视频编码的思想最早分别由 Slepian 和 Wolf^[2], Wyner 和 Ziv^[3] 分别在信息编码理论上提出, 并且加以论证. Slepian-Wolf 定理^[2] 论述了相关信源进行无损编码需要满足的条件, 指出相关信源“独立编码, 联合解码”的压缩效率能够和相关信源“联合编码, 联合解码”的压缩效率相同. Wyner-Ziv 定理^[3] 主要讨论的是在有损编码的情况下, 相关信源编码的率失真函数描述, 通常 Wyner-Ziv 编码可以等效成信源经过量化后再进行 Slepian-Wolf 编码. 在传统视频通信中, 编码端采用时域

相邻帧间的联合编码(运动补偿), 解码端相邻帧的联合解码实现高效压缩。由 Wyner-Ziv 定理可知, 如果把视频序列的时域相邻帧看成相关信源, 通过相邻帧各自独立编码, 而在解码端利用相邻帧的统计相关性进行联合解码也能够实现高效的视频压缩, 并且在理论上这种分布式编码的效率能够与传统视频编码相当。更为详细的分布式视频编码的理论和相关应用的分析可以参阅文献[1, 4]。

分布式视频编码理论基础早在上个世纪七十年代就已确定, 然而其具体实现方案从 2002 年才开始有学者做相关问题的研究, 并逐渐引起关注, 目前已经成为视频编码的前沿课题之一。目前主要的分布式视频编码方案有: Bernd Girod 等提出的“Wyner-Ziv MPEG intra encoder and inter decode”框架^[4-7]; Ramchandran 等提出的 PRISM(Power efficient Robust high compression Syndrome base Multimedia) 框架^[8, 9]; Zixiang Xiong 提出的“Layered Wyner-Ziv video coding”^[10, 11]; 基于小波编码的分布式视频编码方案^[12, 13]。这些分布式视频编码方案与传统视频编码具有明显的区别: 基于运动估计的混合编码框架视频编码(如 MPEG、H. 26x)在编码端联合编码, 解码端联合解码; 基于帧内编码的视频编码(如 Motion JPEG)每帧基本上都是各自独立地进行编码和解码。虽然已实现的分布式视频编码的压缩效果优于帧内编码^[4, 8], 但是与传统的帧间编码相比还有不少差距, 还需要做大量更为深入的研究。在分布式视频编码中, 边信息是一个非常重要的概念。广义上, 边信息是指编码器(发送端)或解码器(接收端)能够得到的关于信源(信道)的额外信息。在分布式视频编码中, 边信息是解码器中所能得到的关于当前解码帧的额外信息, 一般可以通过对已解码的时域相邻帧进行运动估计, 时域内插, 得到的对当前解码帧的信息估计做为边信息, 因此分布式视频编码能够把计算量复杂的运动估计从编码端搬到解码端。空间域 Wyner-Ziv 视频编码框架^[4, 5]是一种最具有代表性的分布式视频编码算法, 其中边信息的获取和联合解码重建算法是影响系统率失真性能的主要因素。本文首先简要分析了空间域 Wyner-Ziv 视频编码系统, 然后提出了一种空间域 Wyner-Ziv 编码系统的性能改进算法, 该算法不改变编码复杂度, 在解码端利用双向运动矢量场估算的方法获得更准确的边信息, 并且采用基于 Huber-Markov 随机场约束的联合迭代解码算法, 可以获得更高质量的重建解码图像。仿真实验表明, 在相同输出码率情况下, 运用本文的改进算法在解码端获得的重建图像质量与空间域 Wyner-Ziv 编码算法相比, 峰值信噪比(PSNR)平均提高 2dB, 同时重建图像的主观视觉效果也有所改善。

2 空间域的 Wyner-Ziv 视频编码系统^[4-6]

空间域 Wyner-ziv 视频编码系统如图 1 所示, 采用了两种帧内编码: 一种是均匀量化后, 进行基于码率自适应删截 turbo 码(RCPT: Rate-compatible punctured turbo)的 Slepian-Wolf 编码, 目的是保持视频信号的波形, 并且 RCPT 通过解码器的反馈信息, 建立删截机制, 能够实现高效的压缩; 另一种采用的是传统视频编码中的帧内编码(如 H. 263 的帧内编码)作为关键帧(Key Frame)。解码端则利用已解码重建帧通过运动估计求取边信息, 进行联合解码。

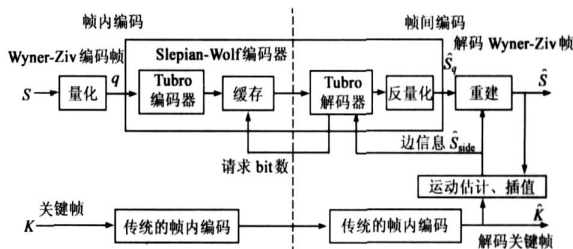


图 1 空间域 Wyner-Ziv 视频编码系统框图

在 Wyner-Ziv 视频编码系统框架中, 解码是最复杂的部分, 简单介绍如下:

(1) 解码后的关键帧 K 通过运动估计, 插值, 得到边信息 S_{side} ;

(2) 假定原图像 S 与边信息 S_{side} 的误差 $e = S - S_{side}$ 满足拉普拉斯(Laplacian)分布, 即 $P(e) = (1/2b) \cdot \exp(-|e|/b)$;

(3) 利用最大后验估计(MAP: maximum a posteriori) $\text{Max}\{P(S|S_{side}, S_q)\}$ 估计最佳的重构值 S , 重构公式如式(1)所示, 其中 $Q(x)$ 表示 2^M 级的量化;

$$S(x, y) = \begin{cases} S_q(x, y), & Q(S_{side}(x, y)) < Q(S_q(x, y)) \\ S_{side}(x, y), & Q(S_{side}(x, y)) = Q(S_q(x, y)) \\ S_q(x, y) + 2^M, & Q(S_{side}(x, y)) > Q(S_q(x, y)) \end{cases} \quad (1)$$

在相同输出码率的情况下, 解码端的空间域 Wyner-Ziv 视频编码重建图像的 PSNR 与传统的帧内编码(如 H. 263)相比平均高 1.5dB 以上, 同时编码的计算复杂度较低, 所需的缓存少, 抗误码性好。

3 空间域 Wyner-Ziv 视频编码系统的改进算法

空间域 Wyner-Ziv 视频编码的解码算法是提高系统性能的关键, 主要体现在两个方面: (1) 获取边信息 S_{side} 的精确度; (2) 如何在 S_{side} 和 S_q 发生可能发生错误的情况下, 尽可能正确估算 S 。本文主要从这两个方面考虑, 对空间域 Wyner-Ziv 视频编码系统加以改进。

3.1 边信息 S_k 获取的改进

运动估计是获得边信息 S_{side} 的主要手段, 较为精确的 S_{side} 的获取取决于两方面: (1) 关键帧 K 和参考帧 S_q 的重构质量; (2) 如何利用运动估计手段获得具有较高

可信度的边信息 S_{side} . 本文的改进措施如下:

(1) 关键帧 K 采用 H. 263 的帧内编码. 传统视频编码通常使用 DCT 和量化技术, 会产生块效应. 本文在解码端进行去块效应滤波, 具体采用 H. 263+^[14] 附录 J 中去块效应滤波算法, 这样在一定程度上可以提高关键帧的编码效率, 同时没有在编码端增加额外的计算量;

(2) Wyner Ziv 帧 S 的编码过程是经过 2^M 级均匀量化后, 再采用 Turbo 编码. 大量试验数据表明视频序列的灰度值在 $[0, 35]$ 和 $[235, 255]$ 两个区间内的像素点占所有像素点个数的 2% 以内. 因此本文采用非均匀的 LloydMax 量化, 在像素点灰度值分布较密集的区域采用较小的量化间隔, 在分布较稀疏的区域采用较大的量化间隔, 以使量化噪声功率最小, 在一定程度上可以提高 Wyner Ziv 帧 S 编码的率失真特性.

(3) 边信息 S_{side} 通过运动估计来估算, 首先估计运动矢量场

$$\langle d_x(x, y), d_y(x, y) \rangle$$

$$= \arg \min_{\substack{dx, dy \\ x, y \in \Omega}} \left[\sum_{x, y \in \Omega} |S(x, y) - S_k(x + d_x(x, y), y + d_y(x, y))|^2 \right] \quad (2)$$

式中, Ω 为单帧图像的像素点集合. 此时 S_{side} 的最佳值为

$$S_{\text{side}} = S_k(x + d_x(x, y), y + d_y(x, y)) \quad (3)$$

实际在解码端 $S(x, y)$ 和 $S_k(x, y)$ 均不可得, 分别用 $S_q(x, y)$ 和 $S_k(x, y)$ 来近似, 运动矢量场 $\langle d_x(x, y), d_y(x, y) \rangle$ 的计算可以看成

$$\langle d_x(x, y), d_y(x, y) \rangle$$

$$= \arg \min_{\substack{dx, dy \\ x, y \in \Omega}} \left[\sum_{x, y \in \Omega} |S_q(x, y) - S_k(x + d_x(x, y), y + d_y(x, y))|^2 \right] \quad (4)$$

可得到 S_{side} 的估算值为

$$S_{\text{side}} = S_k(x + d_x(x, y), y + d_y(x, y)) \quad (5)$$

考虑到在视频编码中, B 帧利用双向预测可比前向预测的 P 帧获得更多帧间冗余信息,

本文采用类似的方法, 如图 2 所示. 两个关键帧 $K(t-n)$ 和 $K(t)$ 分别对

$S_q(t-i)$ 做 4×4 块

匹配运动估计, 得到两个边信息 $S_{\text{side}}(t-n)$ 和 $S_{\text{side}}(t)$. 考虑到视频序列的相关性随间隔时间变长而减少, 最终的边信息 $S_{\text{side}}(t-i)$ 为

$$S_{\text{side}}(t-i) = \frac{n-i}{n} S_{\text{side}}(t) + \frac{i}{n} S_{\text{side}}(t-n) \quad (6)$$

这样估算出 $S_{\text{side}}(t-i)$ 将更准确, 并且在误码情况

下, 比如一个关键帧或者关键帧的一部分数据丢失时, 可以用另一个关键帧没有丢失的边信息来弥补. 所以, 本文的 S_{side} 算法具有更好的抗误码纠错机制.

3.2 基于 Huber Markov 随机场约束的 S 迭代解码

空间域 Wyner Ziv 视频编码中 S 解码的计算如式

(1). 注意到在 $Q(S_{\text{side}}(x, y)) \neq Q(S_q(x, y))$ 时, $S(x, y)$ 只取决于 $S_q(x, y)$, 有时可能引起较大的估计误差. 本文引入基于 Huber Markov 随机场 (HMRf) 的先验概率模型加以约束, 这种模型被广泛用于图像和视频处理领域, 并取得良好效果, 实验表明这个先验概率模型对图像做平滑约束的同时, 具有较好的边界保留特性^[15-17], Huber 函数如下

$$p(s) = \frac{1}{G} \exp \left\{ - \lambda \sum_{s \in S} \rho_T(d_s^t s) \right\} \quad (7)$$

式中 G 是一个大于零的常数, λ 是一个标准化参数, 阈值函数 $\rho_T(n)$ 定义如下

$$\rho_T(n) = \begin{cases} n^2 & |n| \leq T \\ T^2 + 2T(|n| - T) & |n| > T \end{cases} \quad (8)$$

本文取 $T=10$. 将这个先验概率考虑到 S 的重构中, 可以得到加平滑约束 S 的计算公式如下

$$S(x, y) = \arg \min_s \left[\sum_{s \in \Phi} \rho_T(d_s s) \right] \quad (9)$$

$$S(x, y) = \begin{cases} S \in [S_q(x, y) - q_m/2, S_q(x, y) + q_m/2], \\ Q(S_{\text{side}}(x, y)) < Q(S_q(x, y)); \\ S \in [S_{\text{side}}(x, y) - \beta, S_{\text{side}}(x, y) + \beta], \\ Q(S_{\text{side}}(x, y)) = Q(S_q(x, y)); \\ S \in [S_q(x, y) - q_m/2, S_q(x, y) + q_m/2], \\ Q(S_{\text{side}}(x, y)) > Q(S_q(x, y)); \end{cases} \quad (10)$$

其中式 (10) 是对式 (9) 中 $S(x, y)$ 取值范围的限定, 而 q_m 为量化间隔, 取 $\beta=2$, Φ 表示 $S(x, y)$ 点所有阈值的集合, 本文取与 $S(x, y)$ 相邻的 8 个点, 如图 3 所示, 从 4 个方向描述像素点的像素值变化, 具体阈值的计算由式 (11)~(15) 依次给出.

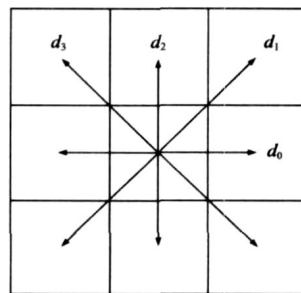


图 3 像素点 $S(x, y)$ 的阈值区域

$$\sum_{s \in \Phi} \rho_T(d_s s) = \sum_{m=1}^3 \rho_T(d_{x, y, m} S) \quad (11)$$

$$d_{x, y, 0} S = S(x+1, y) + S(x-1, y) - 2S(x, y) \quad (12)$$

$$d_{x, y, 1} S = \frac{1}{2} [S(x+1, y+1) + S(x-1, y+1) - 2S(x, y)] \quad (13)$$

$$d_{x, y, 2} S = S(x, y-1) + S(x, y+1) - 2S(x, y) \quad (14)$$

$$d_{x,y,3}S = \frac{1}{2}[S(x+1, y+1) + S(x-1, y-1) - 2S(x, y)] \quad (15)$$

Huber-Markov 随机场的先验概率模型可以从各个方向对图像点的像素值加以约束, 并且当 $|d_s| > T$ 时, 能够较好地保留图像边界。

通常根据式(9)、(10) 求解 S 是一个非常复杂的过程, 可以通过梯度递减的迭代算法来实现, 梯度定义为

$$g(S) = \nabla \left(\sum_{s \in \Omega} \rho_T(d_s S) \right) = \sum_{s \in \Omega} d_s \rho'_T(d_s S) \quad (16)$$

在解码端, 重建图像 S 的迭代计算如式(17)所示, 式中 $\lambda^{(n)}$ 是第 n 次迭代时的步长, 用来控制收敛的速度和结果的精度。当 $S^{(n+1)} = S^{(n)}$ 时, 停止迭代, 得到 S 的估计值。

$$S^{(n+1)} = S^{(n)} - \lambda^{(n)} g(S^{(n)}) \quad (17)$$

4 仿真实验结果与分析

在本文的仿真实验中, Wyner Ziv 帧 S 的 Turbo 编码器码率为 $4/5$, 生成矩阵为 $[1, (1 + D + D^3 + D^4)/(1 + D^3 + D^4)]$ 。通过对量化级数 $2^M \in \{2, 4, 8, 16\}$ 的调整, 产生不同的输出码率, 从而获得不同的压缩率。关键帧 K 采用 H. 263 的帧内编码, 并且和 Wyner Ziv 帧 S 的编码顺序为“K-S-S”, 即每 2 个关键帧 K 之间有 3 个 Wyner Ziv 帧 S 。

在仿真实验中, 本文提出的空间域 Wyner-Ziv 视频编码系统的改进算法分别与文献[5]的空间域 Wyner-Ziv 视频编码算法、H. 263 的帧内编码(I-I-I)、H. 263 的帧间编码(P-P-P)的进行性能比较, 实验采用 Foreman 和 Salesman 两个标准序列, 其图像格式为 QCIF (176×144), 编码 100 帧, H. 263 编码器采用 TMN8^[18]。解码端得到的率失真曲线如图 4、图 5 所示, 在相同码率情况下, 文献[5]算法重建图像的 PSNR 比 H. 263 帧内编码平均高 1.5dB 以上, 并且编码复杂度小于 H. 263 帧内编码; 而本文改进算法与文献[5]算法相比, 获得的重建图像 PSNR 平均提高 2dB。

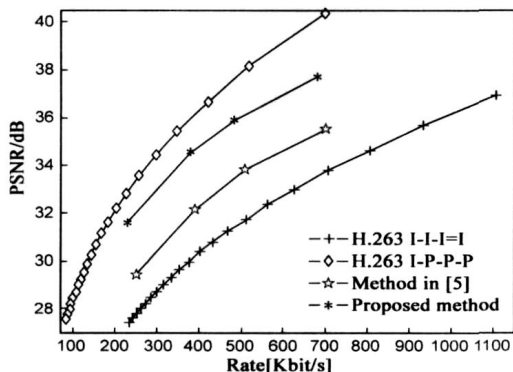


图 4 Salesman 序列 (25 帧/秒)

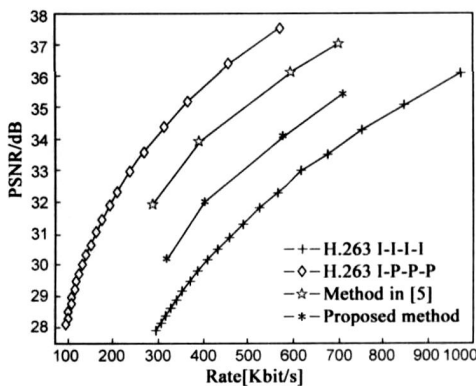


图 5 Foreman 序列 (25 帧/秒)

图 6 是 Foreman 序列第 10 帧的主观质量比较, 可以看出与文献[5]相比, 本文算法得到的解码图像的主观效果有所提高。总体来说, 本文算法相当有效, 并且没有增加编码器的额外开销, 而是在解码端充分挖掘统计信息, 从而提高了整个编码系统的率失真性能。在算法的运行时间方面, 与文献[5]算法相比, 本文算法增加了一定程度的解码复杂度, 其中有三个主要原因: (1) 解码端需要针对关键帧做环路滤波; (2) 通过双向运动估计和时域内插计算边信息; (3) 基于 HMRF 模型约束的联合解码采用了迭代算法, 而文献[5]算法的联合解码求解是一种非迭代算法, 虽然本文算法一般在迭代 5 次之内能够收敛, 但是迭代算法要比非迭代算法耗费更多的时间。实验统计, 本文算法的解码时间大约是文献[5]解码时间的 3~6 倍。



图 6 Foreman 序列 (176×144) 主观质量比较 (第 10 帧)

5 结束语

分布式视频编码的基本思想与传统视频编码有着明显的区别, 是一种全新的视频编码框架, 理论基础是 Slepian-Wolf、Wyner-Ziv 编码理论。由于分布式视频编码的特点是编码简单、解码复杂, 压缩效率高, 抗误码性能好, 易于分级和多描述处理, 因而非常适合于耗电低、存储容量小、运算速度有限的无线视频终端, 具有较大的应用前景。空间域 Wyner-Ziv 视频编码是一种典型的分布式视频编码实现方案, 本文从边信息获取和联合重构解码两方面对空间域 Wyner-Ziv 视频系统加以改进。实验表明, 采用本文改进算法不增加编码复杂度, 在相同码率情况下, 解码端重建图像的 PSNR 比文献[5]提出的空间域 Wyner-Ziv 视频编码算法平均提高 2dB, 并且具有更好的主观质量。

参考文献:

- [1] Z Xiong, D Angelos, C Samuel. Distributed source coding for sensor networks[J]. IEEE Signal Processing Magazine, 2004, 21(5): 80–94.
- [2] J Slepian, J Wolf. Noiseless coding of correlated information sources[J]. IEEE Transactions on Information Theory, 1974, 19(4): 471–480.
- [3] A Wyner, J Ziv. The rate distortion function for source coding with side information at the decoder[J]. IEEE Transactions on Information Theory, 1976, 22(1): 1–10.
- [4] B Girod, A Aaron, S Rane, D Rebollo Monedero. Distributed video coding[J]. Proceedings of the IEEE, 2005, 93(1): 71–83.
- [5] A Aaron, S Rane, R Zhang, B Girod. Wyner Ziv coding for video: applications to compression and error resilience[A]. Proc IEEE Data Compression Conf[C]. Snowbird, Utah, 2003. 93–102.
- [6] A Aaron, E Setton, B Girod. Toward practical Wyner Ziv coding of video[A]. IEEE Int Conf Image Processing[C]. Barcelona, Spain, 2003. 1–4.
- [7] A Aaron, S Rane, E Setton, B Girod. Transform domain Wyner Ziv codec for video[A]. Proc SPIE Visual Communications and Image Processing[C]. San Jose, CA, 2004. 520–528.
- [8] R Puri, K Ramchandran. PRISM: A Video Coding Paradigm Based on Motion Compensated Prediction at the Decoder[OL]. <http://www.eecs.berkeley.edu/~rpuri/researchlinks/papers/purirvc2003.pdf.gz>.
- [9] R Puri, K Ramchandran. PRISM: A ‘reversed’ multimedia coding paradigm[A]. IEEE Int Conf Image Processing[C]. Barcelona, Spain, 2003. 617–620.
- [10] Z Xiong. Layered Wyner Ziv video coding[OL]. <http://tlo.tamu.edu/tlo/documents/INDR/2101.pdf>.
- [11] Q Xu, Z Xiong. Layered Wyner Ziv video coding[A]. Proc. VCIP 04[C]. San Jose, CA, 2004. 83–91.
- [12] J Fowler, M Tagliasacchi, B Pesquet Popescu. Wavelet based distributed source coding of video[OL]. <http://www.ece.msstate.edu/~fowler/Publications/Papers/FTP2005.pdf>.
- [13] S Yaman, G AlRegib. A low-complexity video encoder with decoder motion estimator[A]. Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP 04)[C]. 2004. 157–160.
- [14] Video Coding for Low Bitrate Communication[S]. ITU-T Recommendation H. 263 Version 2, 1998.
- [15] T O'Rourke, R Stevenson. Improved image decompression for reduced transform coding artifacts[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1995, 5(6): 490–499.
- [16] R Schultz, R Stevenson. Extraction of high resolution frames from video sequences[J]. IEEE Transactions on Image Processing, 1996, 5(6): 996–1011.
- [17] M Robertson, R Stevenson. DCT quantization noise in compressed images[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(1): 27–38.
- [18] Video Codec Test Model. ITU-T/SG15[R]. TMN8, Portland, OR, June 1997.

作者简介:



干宗良 男, 1979年9月生于新疆维吾尔自治区石河子市, 现为南京邮电大学通信与信息工程学院博士研究生. 主要研究方向为分布式视频编码、视频通信.

E-mail: ganzongliang@163.com



齐丽娜 女, 1979年3月生于河北省衡水市, 现为南京邮电大学通信与信息工程学院博士研究生. 主要研究方向为超宽带无线通信系统.

E-mail: aigief@163.com



朱秀昌 男, 1947年8月生于江苏省丹徒市, 南京邮电大学通信与信息工程学院副院长, 教授, 博士生导师, 中国通信学会、电子学会高级会员, 江苏省图像处理与图像通信重点实验室主任. 主要研究方向为图像通信、视频处理.

E-mail: zhuxc@njupt.edu.cn