

分布式视频编码中基于多概率混合分布的相关噪声建模方法

唐振华, 梁祥严, 覃团发, 常 侃

(广西大学计算机与电子信息学院, 广西南宁 530004)

摘 要: 相关噪声建模(Correlation Noise Modeling, CNM)的准确度是影响分布式视频编码(Distributed Video Coding, DVC)系统性能的关键因素之一. 针对经过 8×8 离散余弦变换(Discrete Cosine Transform, DCT)后的相关噪声子带, 本文提出了全新的基于多概率混合分布的 CNM 方法. 该方法可根据子带的信息熵与子带在柯西、拉普拉斯和高斯三种概率分布下信息熵的相似度, 自适应选择合适的概率分布对子带进行建模, 提高了 CNM 的准确度. 实验结果表明, 采用本文提出的离线和在线 CNM 方法与现有典型的 CNM 方法相比, DVC 系统的率失真(Rate-Distortion, R-D)性能均获得显著的提高.

关键词: 分布式视频编码; Wyner-Ziv 视频编码; 相关噪声建模; 子带概率分布; 信息熵

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2015)02-0365-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.02.024

Correlation Noise Modeling Algorithm Based on Multiple Probability Distributions for Distributed Video Coding

TANG Zhen-hua, LIANG Xiang-yan, QIN Tuan-fa, CHANG Kan

(School of Computer and Electronics Information, Guangxi University, Nanning, Guangxi 530004, China)

Abstract: Accurate correlation noise modeling (CNM) is one of the key factors affecting the performance of distributed video coding (DVC). This paper presents a novel CNM algorithm based on multiple probability distributions for DVC, which mainly focuses on the subbands of correlation noise after 8×8 discrete cosine transform (DCT). The proposed CNM method can select the best suitable probability distribution to model the subbands of correlation noise adaptively by comparing the entropy of the subbands with the ones of the candidate probability distributions including Cauchy, Laplace and Gaussian. Experimental results show that compared with the existing typical CNM approaches, the proposed CNM method can improve the rate-distortion (R-D) performance of the DVC system significantly both in offline and online manner.

Key words: distributed video coding; Wyner-Ziv video coding; correlation noise modeling; subband probability distribution; information entropy

1 引言

分布式视频编码(Distributed Video Coding, DVC)^[1~5]与传统视频编码(如 MPEG-x 和 H.26x 等)相比,将计算的复杂度从编码端转移到解码端,且在理论上可达到与传统视频编码相当的压缩性能^[6,7],为功耗、计算能力及储存容量受限的无线视频终端提供了可行的视频编码方案.

Wyner-Ziv(WZ)视频编码是一种典型的 DVC 编码方案^[1~5].解码端为了有效地利用编码端发送的校验比特信息来纠正边信息(Side Information, SI)的估计偏差,往

往需借助原始的 WZ 帧与 SI 帧之间的误差信息,即相关噪声.由于解码端通常无法获得原始的 WZ 帧,寻求 WZ 帧与 SI 帧之差所服从的概率分布的过程就是相关噪声建模(Correlation Noise Modeling, CNM).CNM 的准确度越高,译码时所需要的校验位也越少,有利于进一步降低系统的码率.因而,提高 CNM 的准确度已成为提升 DVC 系统率失真(Ratio-Distortion, R-D)性能的关键因素之一.

目前,研究 DVC 系统中 CNM 方法的工作可分为三类.第一类是基于拉普拉斯概率分布^[2,8,9]及其改进方案^[10,11]的建模方法,其主要特点是通过各种不同的算法或方案来提高拉普拉斯分布参数的估计精度.典型工

作有: Brites 等人^[9]针对基于像素域和离散余弦变换(Discrete Cosine Transform, DCT)域的 DVC 系统,提出了视频帧、图像块、像素、DCT 子带等多个不同层次的 CNM 方法.然而,由于视频内容多样性及其时变的特点,仅采用拉普拉斯概率分布对相关噪声子带进行建模会存在较大偏差.第二类是单独采用非拉普拉斯分布来建模相关噪声的方法,例如采用指数分布来对相关噪声进行建模等^[12],但这类方法仅对部分视频内容能够获得较好的系统性能.第三类是基于多种概率混合分布的建模方法,该类方法以拉普拉斯概率分布为基础,结合其他概率分布来提高相关噪声建模的准确度.典型工作有:杨春玲等人^[13]针对相关噪声 DCT 子带具有尖锐的峰值特性和较长的拖尾情况,采用拉普拉斯-柯西混合分布来建模相关噪声改善 DVC 的性能.但该模型仅适用于子带数目较少的场合(如子带数为 16),当子带数增大时(如子带数为 64),子带尾部会变短,模型将带来较大偏差,最终导致 DVC 系统性能变差.

本文针对经过 8×8 DCT 变换后的相关噪声子带,提出了全新的基于多概率混合分布的 CNM 方法.该方法可根据相关噪声子带实际的信息熵与子带在拉普拉斯、柯西和高斯三种概率分布下的信息熵的相似度,自适应选择准确的概率分布对各个子带进行建模,提高了 CNM 的准确度,从而提升了 DVC 系统的 R-D 性能.

2 相关噪声子带的概率分布

2.1 相关噪声子带的定义

设 $WZ(x, y)$ 和 $SI(x, y)$ 分别表示原始 WZ 帧和 SI 帧,则相关噪声可表示为:

$$R(x, y) = WZ(x, y) - SI(x, y) \quad (1)$$

其中, (x, y) 表示视频图像中像素点的坐标.如果对相关噪声 R 进行 8×8 的 DCT 变换,则对应的变换帧可表示为:

$$D(u, v) = \text{DCT}[R(x, y)] \quad (2)$$

其中, (u, v) 表示变换帧中系数的坐标.此时,提取变换帧中各子带对应的系数,即可获得相关噪声的 64 个子带,即 $B_m, m \in \{1, 2, \dots, 64\}$.

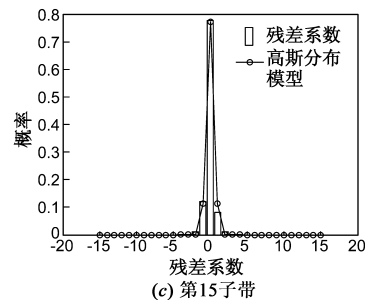
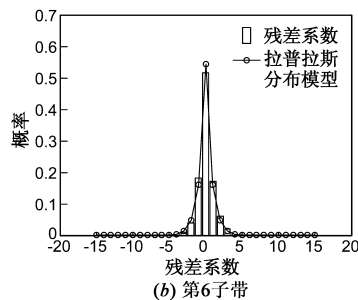
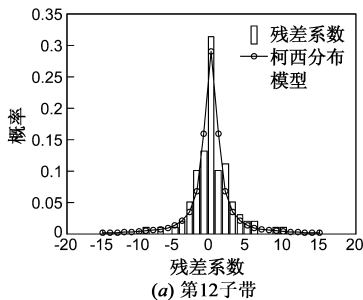


图2 Stefan序列中不同相关噪声子带的概率分布拟合

2.2 相关噪声子带的概率分布特性

本文对若干典型视频序列进行实验测试,提取它们的相关噪声子带进行分析.图 1 给出了 Foreman、Carphone、Coastguard 和 Stefan 四个视频序列中不同相关噪声子带的分布.由图 1 可看出,不同视频序列的相关噪声子带的概率分布特性存在一定的差异,这主要是由于各视频图像中内容细节的不同所造成.经分析后发现,相关噪声子带的分布可分为三种类型:拖尾较长、拖尾中等和拖尾较短.如果仅用拉普拉斯分布来对上述三种类型的子带进行建模,将难以获得高准确度的 CNM.

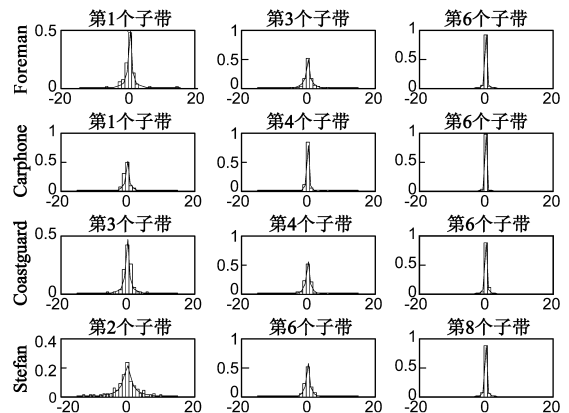


图1 不同视频序列中相关噪声子带的分布

从概率分布来说,尾部衰减较慢,拖尾很长的子带,可采用柯西分布来对这种类型的相关噪声子带进行建模^[13];对于尾部变化速度一般,拖尾中等的子带,可采用拉普拉斯分布来进行描述;而对于尾部衰减较快,拖尾很短的子带,则较适合于利用高斯分布对其构造模型.本文利用上述三种概率分布对 Stefan 序列的第 2 帧中不同的子带进行了分布曲线拟合,结果如图 2 所示.由图 2 可看出,上述的三种概率分布能较好地描述了不同子带的分布特性.

3 基于多概率混合分布的相关噪声建模方法

在 DVC 系统中,离线的 CNM 方法不仅是在线方法的重要组成部分,也将为在线方法的优化提供性能上限参考.

3.1 离线的相关噪声建模方法

设 χ_k 和 K_m 分别为当前相关噪声子带 B_m 的系数及系数的总数,其中 $k = 1, 2, \dots, K_m$, 则离线的 CNM 方法的具体步骤如下:

步骤 1 统计子带系数的概率分布,并计算当前子带的信息熵。

令 $\rho_m(\chi_k)$ 表示子带系数 χ_k 的出现概率, $N_m(\chi_k)$ 表示 χ_k 在 B_m 中出现的次数,则

$$\rho_m(\chi_k) = N_m(\chi_k) / K_m \quad (3)$$

根据信息熵^[14]的定义,当前子带的信息熵 H_m 可计算为:

$$H_m = - \sum_{\chi_k \in B_m} \rho_m(\chi_k) \log_2(\rho_m(\chi_k)) \quad (4)$$

步骤 2 利用子带系数分别估算柯西、拉普拉斯和高斯三种概率分布的参数。

为了获得当前子带的最佳概率分布,需得到子带在上述三种分布下的信息熵,而估算概率分布参数则是获取信息熵的首要步骤。设 C_m 、 L_m 和 G_m 分别表示当前子带在柯西、拉普拉斯和高斯三种分布下的参数。本文采用最佳熵法^[15]估算拉普拉斯和柯西两种分布的参数,而采用最小概率分布法来估算高斯分布的参数,具体计算步骤详见 3.2 节。

步骤 3 获取子带系数在柯西、拉普拉斯和高斯三种分布下的概率分布。

设子带系数在柯西、拉普拉斯和高斯三种分布下的概率分布分别为 $\rho_m^C(\chi_k)$ 、 $\rho_m^L(\chi_k)$ 和 $\rho_m^G(\chi_k)$, 它们可通过式(5)~(7)进行计算:

$$\rho_m^C(\chi_k) = \frac{1}{\pi} \frac{C_m}{C_m^2 + \chi_k^2}, \quad \chi_k \in B_m \quad (5)$$

$$\rho_m^L(\chi_k) = \frac{L_m}{2} e^{-L_m |\chi_k|}, \quad \chi_k \in B_m \quad (6)$$

$$\rho_m^G(\chi_k) = \frac{1}{\sqrt{2\pi} G_m} e^{-\frac{\chi_k^2}{G_m^2}}, \quad \chi_k \in B_m \quad (7)$$

步骤 4 分别计算当前子带在柯西、拉普拉斯和高斯三种分布下的信息熵。

利用步骤 3 的结果,通过式(4)分别计算当前子带在柯西、拉普拉斯和高斯三种概率分布下的信息熵 H_m^C 、 H_m^L 和 H_m^G 。

步骤 5 分别计算当前子带的信息熵与该子带在上述三种分布下的信息熵之差的绝对值。

设当前子带的信息熵及其在上述三种分布下的信息熵之差的绝对值分别为 ΔH_m^C 、 ΔH_m^L 和 ΔH_m^G , 它们可通过式(8)~(10)来计算:

$$\Delta H_m^C = |H_m^C - H_m| \quad (8)$$

$$\Delta H_m^L = |H_m^L - H_m| \quad (9)$$

$$\Delta H_m^G = |H_m^G - H_m| \quad (10)$$

步骤 6 确定当前子带的最佳概率分布模型。

由信息熵理论^[14]可知,当两种概率分布的信息熵越接近时,它们所携带的信息量越相近,且分布的形状越相似。本文通过对步骤 5 中信息熵之差绝对值的大小进行比较,选择绝对值最小的对应的分布作为当前子带的最佳的概率分布模型。

3.2 概率分布参数估计

3.2.1 最佳熵法^[15]

在确定某种概率分布(如拉普拉斯或柯西分布)的条件下,最佳熵的具体过程如下:首先采用式(3)计算子带系数的概率分布。然后,在设定概率分布参数的可能取值中逐一取出参数,计算当前子带在确定概率分布下的条件熵。最后,找出最小的条件熵,该熵所对应的分布参数作为最终的估计参数。

3.2.2 最小概率分布法

本文通过实验测试发现,采用最小概率分布法计算高斯分布的参数较最佳熵法更为准确。在高斯分布下,利用最小概率分布法的估算分布参数的过程如下:

首先,采用式(3)计算子带系数的概率分布 $\rho_m(\chi_k)$, 并设定分布参数 G_m 的取值范围和变化步长(即设定候选参数)。接着,取出当前候选参数值 G_m^i , 利用式(7)获得子带系数在高斯分布下的概率分布 $\rho_{m,i}^G(\chi_k)$ 。然后,利用式(11)计算当前子带中概率分布 $\rho_{m,i}^G(\chi_k)$ 和 $\rho_m(\chi_k)$ 的欧式距离,即

$$E_m(i) = \sum_{\chi_k \in B_m} |\rho_{m,i}^G(\chi_k) - \rho_m(\chi_k)| \quad (11)$$

遍历所有候选参数,选取 $E_m(i)$ 最小时的分布参数作为最终的估计参数。

3.3 在线的相关噪声建模方法

在 DVC 系统的解码端,由于缺少原始的 WZ 帧数据,需要先利用已解码的关键帧(K 帧)估计相关噪声。本文利用文献[9]提出的通过计算前后两个 K 帧的运动补偿帧之差来实现相关噪声的估计。设 $R_E(x, y)$ 为前后两个 K 帧的运动补偿帧之差,则其计算方法为:

$$R_E(x, y) = X_f(x + d_{xf}, y + d_{yf}) - X_b(x + d_{xb}, y + d_{yb}) \quad (12)$$

其中, $X_f(x + d_{xf}, y + d_{yf})$ 和 $X_b(x + d_{xb}, y + d_{yb})$ 分别表示前向补偿帧和后向补偿帧。

当相关噪声被估计出来后,在线 CNM 方法的剩余步骤与离线 CNM 方法的步骤一致。

4 实验结果及分析

本文采用基于 DCT 域的 WZ 编解码系统(采用 LDPC 码)^[15]作为实验平台,并在该平台上实现了提出的基于多概率混合分布的 CNM 方法。实验采用 Foreman、

Carphone、Coastguard 和 Stefan 四个视频序列的前 100 帧进行测试,视频帧的分辨率均为 QCIF 格式(176 × 144),采样格式为 4:2:0,帧率设为 30fps.实验仅考虑视频帧的亮度分量,编码结构设为 IPIPIPI 格式,K 帧和 WZ 帧均采用 JPEG 量化码表,量化参数分别为 {0.5, 1.0, 2.0, 4.0},而视频质量则采用峰值信噪比(Signal Noise Ratio, PSNR)来衡量.采用最小概率分布法估算高斯分布参数时,设定分布参数的取值范围是 [0, 5],变化步长为 0.01.为了衡量本文方法的性能,实验比较了采用文献[9]提出的基于拉普拉斯分布的 CNM 方法,以及文献[13]提出的基于拉普拉斯和柯西混合分布的 CNM 方法时系统的性能.

4.1 离线 CNM 方法的性能

WZ 视频编解码系统采用上述三种离线 CNM 方法的 R-D 曲线如图 3 所示.由图 3 可看出,对于不同的视频序列,本文提出的离线 CNM 方法相对于其他两种 CNM 方法而言,系统的 R-D 性能均有明显的提高.其中,在相同码率的条件下,本文的方法较文献[9]的方法而言,重构视频的平均 PSNR 值提高至少 0.61dB,最高可达 0.92dB;而相对于文献[13]的方法而言,重构视频的平均 PSNR 值提高至少 0.67dB,最高可达 1.15dB.上述数据表明,本文提出的离线 CNM 方法有效地提高了相关噪声模型的准确度,使得 WZ 编解码系统的 R-D 性能更优.

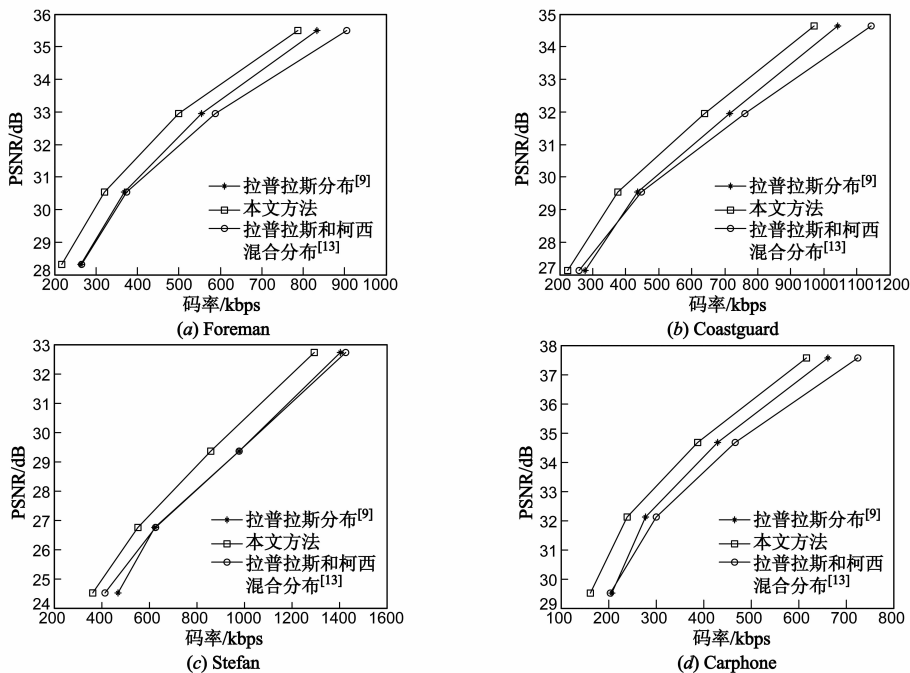


图3 系统采用离线CNM方法的R-D曲线比较

表1 Coastguard 和 Stefan 序列在不同量化参数下平均码率和平均 PSNR 值

QCIF Sequence @ 30Hz	Coastguard				Stefan			
	$Q = 0.5$	$Q = 1.0$	$Q = 2.0$	$Q = 4.0$	$Q = 0.5$	$Q = 1.0$	$Q = 2.0$	$Q = 4.0$
	PSNR = 34.64dB	PSNR = 31.94dB	PSNR = 29.53dB	PSNR = 27.13dB	PSNR = 32.71dB	PSNR = 29.36dB	PSNR = 26.75dB	PSNR = 24.52 dB
Rate [kbps]	Rate [kbps]	Rate [kbps]	Rate [kbps]	Rate [kbps]	Rate [kbps]	Rate [kbps]	Rate [kbps]	
①本文方法	969.78	638.79	376.70	224.03	1296.72	860.35	553.68	362.07
②拉普拉斯分布 ^[9]	1042.96	714.89	436.82	277.61	1405.47	979.47	625.04	469.47
③拉普拉斯和柯西混合分布 ^[13]	1143.83	759.92	448.75	259.38	1428.89	978.11	627.75	412.95
Rate(②-①)	73.18	76.10	60.12	53.58	108.75	119.12	71.37	107.40
Rate(③-①)	174.05	121.13	72.05	35.35	132.17	117.76	69.80	50.88

为了更详细地比较上述三种离线 CNM 方法的性能,表 1 给出了 Coastguard 和 Stefan 序列在不同的量化参数下系统的平均码率和平均 PSNR 值,以及采用上述三种离线 CNM 方法后 WZ 编解码系统的码率比较.由表 1 可看出,在重构视频质量相同的情况下,系统采用本文提出的离线 CNM 方法在不同的量化参数下,获得的平均码率均小于其他的两种 CNM 方法.其中,与文献[9]的方法相比,本文的方法可降低的最大的平均码率为 119.12kbps,最小的平均码率为 53.58kbps;与文献[13]的方法相比,本文的方法可降低的最大平均码率为 174.05kbps,最小平均码率为 35.35kbps.上述情况表明,本文提出的离线 CNM 方法与其它两种 CNM 方法相比,提高了相关噪声模型的准确度,减少了解码时所需要的校验比特,因而获得更低的系统码率.

4.2 在线 CNM 方法的性能

WZ 视频编解码系统的量化参数设为 0.5,采用上述三种在线 CNM 方法对 Foreman 和 Carphone 两个序列中各帧码率进行比较,其结果如图 4 所示.由于序列中各帧均为完全解码,因而尽管采用不同的 CNM 方法,序号相同的视频帧的重构质量是相同的.由图 4 可看出,在视频重构质量相同的情况下,本文提出的在线 CNM 方法与其他两种方法相比,相关噪声模型的准确度更高,系统解码时所需要的码率更少,系统 R-D 性能更优.

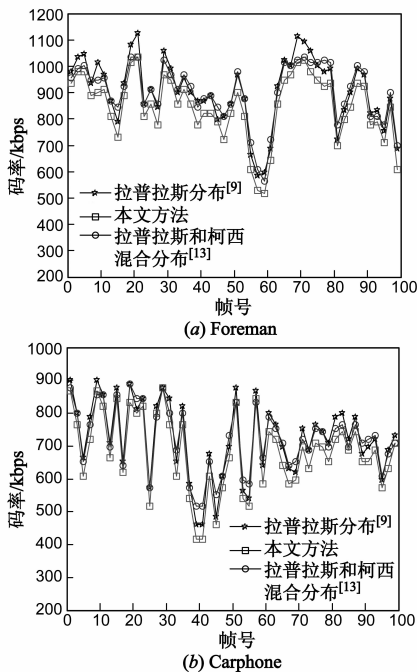


图4 系统采用在线CNM方法时每帧码率的比较

4.3 CNM 方法的复杂度分析

本文提出的离线和在线 CNM 方法涉及对拉普拉斯、高斯和柯西三种分布参数的拟合计算、信息熵计算

以及信息熵相似度的比较等运算,其计算复杂度明显高于文献[9]中基于拉普拉斯分布的 CNM 方法,也略高于文献[13]中基于拉普拉斯和柯西混合分布的 CNM 方法.然而,在实际的 DVC 编解码系统中,相关噪声建模通常是由解码端在线完成的,而解码端通常有足够的系统资源来完成相应的解码运算,所以计算复杂度略高并不会影响实际的应用.

5 结束语

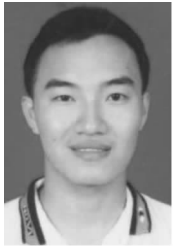
本文在分析相关噪声子带概率分布特征的基础上,提出了全新的基于多概率混合分布的 CNM 方法.该方法通过比较子带的信息熵与子带在柯西、拉普拉斯和高斯三种分布下信息熵的相似度,搜索最佳的概率分布对子带进行建模,以提高 CNM 的准确度.实验结果表明,采用本文提出的离线和在线 CNM 方法时,WZ 视频编解码系统的 R-D 性能明显优于基于拉普拉斯分布以及基于柯西和拉普拉斯混合分布的 CNM 方法.本文的方法为 DVC 系统中基于 DCT 子带层次的相关噪声建模提供了一个新的方案.

参考文献

- [1] Girod B, Aaron A, Rane S, et al. Distributed video coding[J]. Proceedings of the IEEE, 2005, 93(1): 71 - 83.
- [2] Xu Q, Xiong Z X. Layered Wyner-Ziv video coding[J]. IEEE Transactions on Image Processing, 2006, 15(12): 3791 - 3803.
- [3] Puri R, Ramchandran K. PRISM: A video coding paradigm with motion estimation at the decoder[J]. IEEE Transactions on Image Processing, 2007, 16(10): 2436 - 2448.
- [4] Artigas X, Ascenso J, Dalai M, et al. The DISCOVER codec: architecture, techniques and evaluation[A]. The 26th Picture Coding Symposium[C]. Lisbon, Portugal: IEEE Press, 2007. 1103 - 1120.
- [5] 干宗良, 齐丽娜, 朱秀昌. 一种空间域 Wyner-Ziv 视频编码系统的性能改进算法[J]. 电子学报, 2007, 35(10): 2014 - 2018.
GAN Zong-liang, QI Li-na, ZHU Xiu-chang. A novel performance improved method for pixel-domain Wyner-Ziv video coding system[J]. Acta Electronica Sinica, 2007, 35(10): 2014 - 2018. (in Chinese)
- [6] Slepian D, Wolf J K. Noiseless coding of correlated information sources[J]. IEEE Transactions on Information Theory, 1973, 19(4): 471 - 480.
- [7] Wyner A D, Ziv J. The rate-distortion function for source coding with side information at the decoder[J]. IEEE Transactions on Information Theory, 1976, 22(1): 1 - 10.
- [8] Aaron A, Zhang R, Girod B. Wyner-Ziv coding of motion video[A]. The 36th Asilomar Conference on Signals, Systems, and

- Computers[C]. Asilomar, CA, USA: IEEE Computer Society Press, 2002. 240 – 244.
- [9] Brites C, Pereira F. Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(9): 1177 – 1190.
- [10] Song J, Wang K, Liu H, et al. Progressive correlation noise refinement for transform domain Wyner-Ziv video coding[A]. 18th IEEE International Conference on Image Processing[C]. Brussels, Belgium: IEEE Press, 2011. 2625 – 2628.
- [11] Luong H V, Forchhammer S. Noise residual learning for noise modeling in distributed video coding[A]. Picture Coding Symposium[C]. Krakow, Poland: IEEE Press, 2012. 157 – 160.
- [12] Maugey T, Gauthier J, Pesquet-Popescu B, et al. Using an exponential power model for wyner-ziv video coding[A]. IEEE International Conference on Acoustics Speech and Signal Processing[C]. Dallas, Texas, USA: IEEE Press, 2010. 2338 – 2341.
- [13] 杨春玲, 谢小兰. 分布式视频编码中拉普拉斯-柯西混合分布相关噪声模型研究[J]. 电子与信息学报, 2012, 34(10): 2348 – 2352.
Yang Chun-ling, Xie Xiao-lan. A Laplacian-Cauchy mixture model for improved correlation noise modeling in distributed video coding[J]. Journal of Electronics & Information Technology, 2012, 34(10): 2348 – 2352. (in Chinese)
- [14] 傅祖芸. 信息论—基础理论与应用[M]. 北京: 电子工业出版社, 2007.
- [15] Chen D, Varodayan D. Unsupervised Learning of Motion for Distributed Video Coding[DB/OL]. <http://msw3.stanford.edu/~dchen/DVC,2014-02-14>.

作者简介



唐振华 男, 1979 年 12 月生于广西玉林, 博士, 副教授, 硕士生导师. 研究方向为分布式视频编码、基于内容的可变分辨率图像/视频编码等.

E-mail: tangedward@126.com



梁祥严 男, 1988 年 5 月生于广西玉林, 硕士. 研究方向为分布式视频编码.

E-mail: Laxiyan@163.com



覃团发 男, 1966 年 7 月生于广西宾阳, 博士, 教授, 博士生导师. 研究方向为多媒体通信、网络编码等.

E-mail: tfqin@gxu.edu.cn



常侃 男, 1983 年 6 月生于广西南宁, 博士, 副教授, 硕士生导师. 研究方向为视频编码、压缩感知等.

E-mail: pandack0619@163.com