

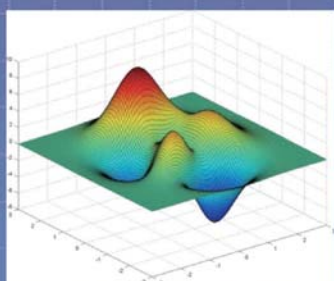


主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

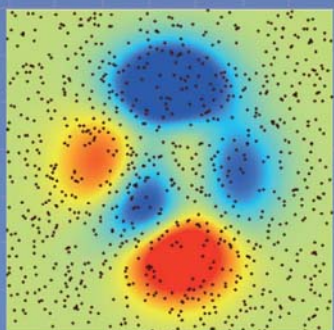
中国图象图形学报

2014
02
VOL.19

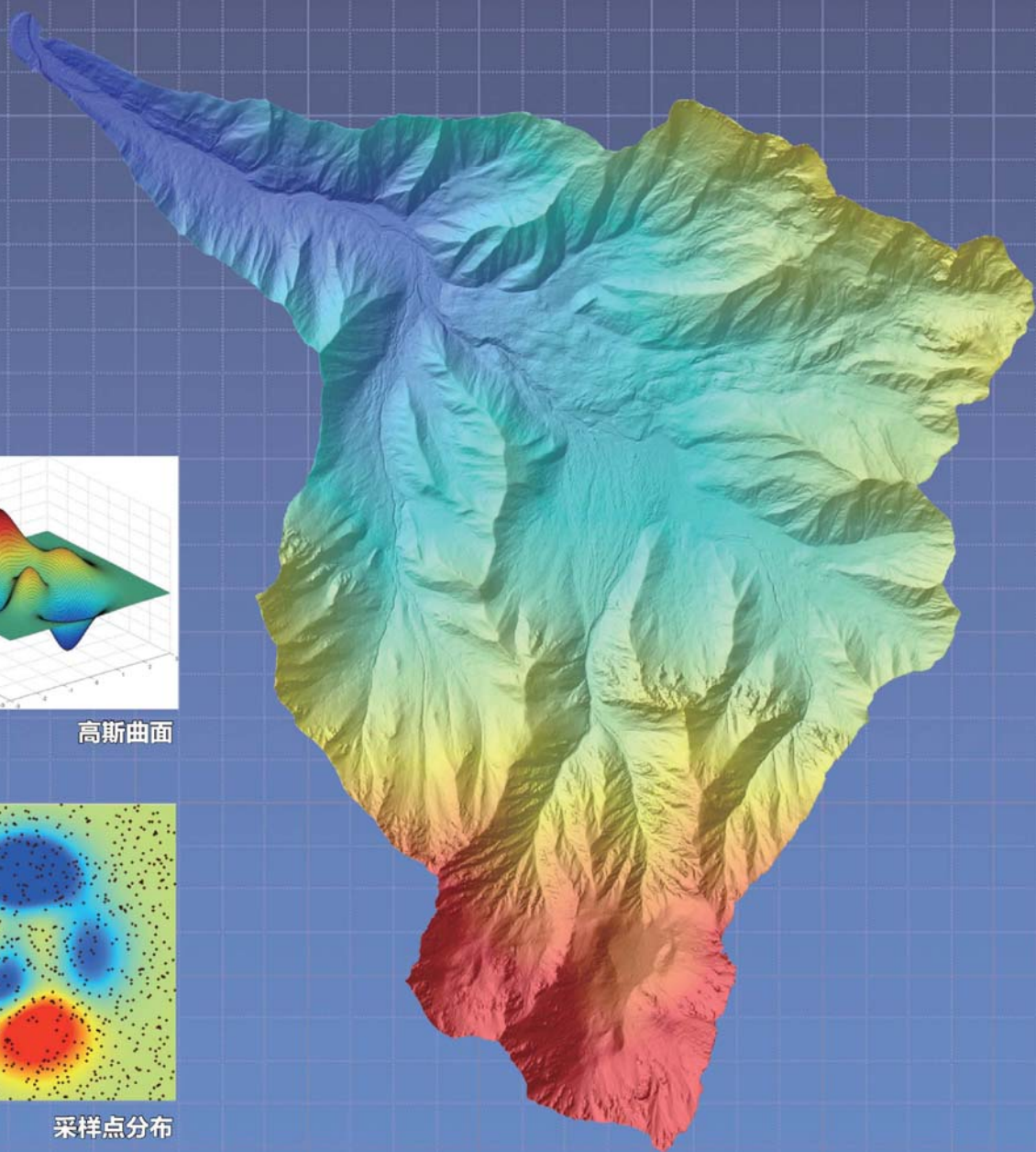
ISSN1006-8961
CN11-3758/TB



高斯曲面



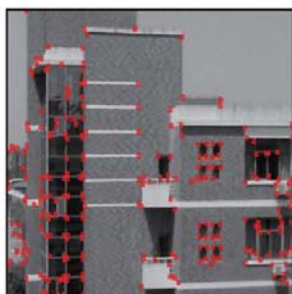
采样点分布



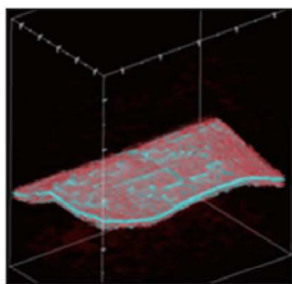
HASM地形模拟 P290



邻域最短距离法寻找最佳拼接缝(第227页)



曲率尺度空间与链码方向统计的角点检测(第234页)



目标特征指导的多分辨率体绘制算法(第283页)

计算机视觉前沿论坛

行为理解的认知推理方法

陶霖密, 杨卓宁, 王国建 0167

深度学习及其在目标和行为识别中的新进展

郑胤, 陈权崎, 章毓晋 0175

图像处理和编码

分布式视频编码中基于改进FCM聚类的相关噪声模型估计

杨春玲, 吴娟 0185

结合一阶自回归滑动平均和压缩感知的视频模型

王教余, 徐小红, 沈仁明, 廖重阳, 杨勋 0194

信息量加权的梯度显著度图像质量评价

徐少平, 杨荣昌, 刘小平 0201

优化加权TV的复合正则化压缩感知图像重建

费选, 韦志辉, 肖亮, 李星秀 0211

消除图像伪轮廓的各向异性自适应滤波

倪婧, 王朔中, 廖纯, 曾兴 0219

邻域最短距离法寻找最佳拼接缝

郑悦, 程红, 孙文邦 0227

图像分析和识别

曲率尺度空间与链码方向统计的角点检测

曾接贤, 李炜烨 0234

基于方向特征的手掌静脉识别

周宇佳, 刘娅琴, 杨丰, 黄靖 0243

复杂环境下高效物体跟踪级联分类器

江伟坚, 郭躬德 0253

图像理解和计算机视觉

均值规范化对比度的局部特征描述符

颜雪军, 赵春霞, 袁夏, 徐丹, 刘凡 0266

计算机图形学

三角域上Said-Ball基的推广渐近迭代逼近

张莉, 李园园, 杨燕, 檀结庆 0275

目标特征指导的多分辨率体绘制算法

郭思奇, 鲁才, 聂小燕 0283

地理信息技术

高精度曲面建模优化方案

赵明伟, 岳天祥, 赵娜 0290

医学图像处理

应用于人体关节缺损修复的建模与可视化

邢慧君, 杨健, 李勤 0297

小邻域统计信息核磁共振医学图像分割模型

张建伟, 方林, 陈允杰, 詹天明 0305

遥感图像处理

结合暗通道原理和双边滤波的遥感图像增强

周雨薇, 陈强, 孙权森, 胡宝鹏 0313

改进NSCT和IHS变换相结合的遥感影像融合

刘慧, 周可法, 王金林, 王珊珊 0322

分段2维主成分分析的超光谱图像波段选择

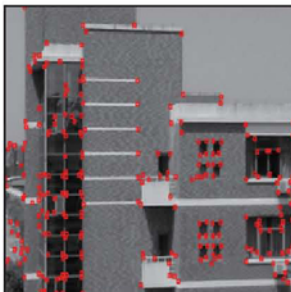
张婧, 孙俊喜, 阮光诗, 刘红喜 0328

CONTENTS

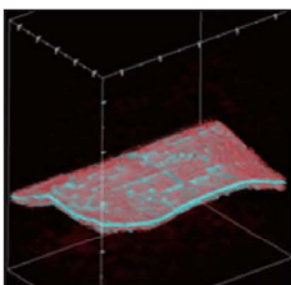
JOURNAL OF IMAGE AND GRAPHICS



Finding an optimal seam-line through the shortest distance in the neighborhood(P227)



Corner detection based on curvature scale space and chain code direction statistics(P234)



Target feature based multi-resolution volume rendering(P283)

Forum: Forefront of Computer Vision

Cognitive reasoning method for behavior understanding

Tao Linmi, Yang Zhuoning, Wang Guojian 0167

Deep learning and its new progress in object and behavior recognition

Zheng Yin, Chen Quanqi, Zhang Yujin 0175

Image Processing and Coding

Correlation noise modeling based on improved Fuzzy C-Means clustering in distributed video coding

Yang Chunling, Wu Juan 0185

Compressed sensing video model based on a first-order autoregressive moving average model

Wang Jiaoyu, Xu Xiaohong, Shen Renming, Liao Chongyang, Yang Xun 0194

Information content weighted gradient saliency structural similarity index for image quality assessment

Xu Shaoping, Yang Rongchang, Liu Xiaoping 0201

Compound regularized compressed sensing image reconstruction based on optimal reweighted TV

Fei Xuan, Wei Zhihui, Xiao Liang, Li Xingxiu 0211

False contour suppression with anisotropic adaptive filtering

Ni Jing, Wang Shuozhong, Liao Chun, Zeng Xing 0219

Finding an optimal seam-line through the shortest distance in the neighborhood

Zheng Yue, Cheng Hong, Sun Wenbang 0227

Image Analysis and Recognition

Corner detection based on curvature scale space and chain code direction statistics

Zeng Jiexian, Li Weiye 0234

Palm-vein recognition based on oriented features

Zhou Yujia, Liu Yaqin, Yang Feng, Huang Jing 0243

Efficient cascade classifier for object tracking in complex conditions

Jiang Weijian, Guo Gongde 0253

Image Understanding and Computer Vision

Local feature descriptor based on mean normalized contrast

Yan Xuejun, Zhao Chunxia, Yuan Xia, Xu Dan, Liu Fan 0266

Computer Graphics

Generalized progressive iterative approximation for Said-Ball based on triangular domains

Zhang Li, Li Yuanyuan, Yang Yan, Tan Jieqing 0275

Target feature based multi-resolution volume rendering

Guo Siqi, Lu Cai, Nie Xiaoyan 0283

Geoinformatics

HASM optimization based on the improved difference scheme

Zhao Mingwei, Yue Tianxiang, Zhao Na 0290

Medical Image Processing

Method of modeling and visualization for repairing of osteoarticular defect

Xing Huijun, Yang Jian, Li Qin 0297

Magnetic resonance medical images segmentation based on a local statistical information model

Zhang Jianwei, Fang Lin, Chen Yunjie, Zhan Tianming 0305

Remote Sensing Image Processing

Remote sensing image enhancement based on dark channel prior and bilateral filtering

Zhou Yuwei, Chen Qiang, Sun Quansen, Hu Baopeng 0313

Remote sensing image fusion based on an improved NSCT and IHS transformation

Liu Hui, Zhou Kefa, Wang Jinlin, Wang Shanshan 0322

Segmented 2DPCA algorithm for band selection of hyperspectral image

Zhang Jing, Sun Junxi, Ruan Guangshi, Liu Hongxi 0328

中图法分类号: TN919.8 文献标识码: A 文章编号: 1006-8961(2014)02-0185-09

论文引用格式: 杨春玲, 吴娟. 分布式视频编码中基于改进 FCM 聚类的相关噪声模型估计[J]. 中国图象图形学报, 2014, 19(2): 185-193.
 [DOI:10.11834/jig.20140203]

分布式视频编码中基于改进 FCM 聚类的相关噪声模型估计

杨春玲, 吴娟

华南理工大学电子与信息学院, 广州 510640

摘要: **目的** 在分布式视频编码中, 为了更加准确地描述相关噪声残差子带的变化特性, 提出一种基于改进的模糊 C 均值(FCM)聚类的模型估计方法。**方法** 本文算法针对每一解码子带选取不同的特征矢量; 利用改进的模糊 C 均值进行聚类; 采用阈值控制法求取相应的模型参数; 然后用重建子带更新下一解码子带的特征矢量, 直到一帧中所有子带解码完成。针对模糊 C 均值对初始聚类中心的敏感性, 采用随机生成隶属度矩阵的方法来缓解聚类陷入局部最优的问题。**结果** 从实验效果和算法复杂度角度考虑, 将残差样本聚为 8 类。实验结果表明, 本文聚类算法可以更加准确地模拟帧内不同区域的不同信道噪声特性, 对于运动越剧烈的序列效果越好, 相对于子带级拉普拉斯估计, 平均增益达 1 dB。**结论** 提出了一种新的相关噪声估计方法, 针对不同的子带选取不同的特征矢量, 并重建更新。实验结果表明, 本文算法能更好地描述相关噪声特性, 获得系统性能的提高。

关键词: Wyner-Ziv 分布式视频编码; 相关噪声模型; FCM 聚类算法; 特征矢量; 拉普拉斯参数

Correlation noise modeling based on improved Fuzzy C-Means clustering in distributed video coding

Yang Chunling, Wu Juan

School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China

Abstract: **Objective** In Wyner-Ziv (WZ) Distributed Video Codec, in order to describe the change characteristics of the correlation noise residual sub-band more accurately, we propose a new correlation noise-modeling algorithm based on improved-FCM (Fuzzy C-Means) clustering. **Method** In our proposed method, for each decoded sub-band, the residual coefficient eigenvectors are formed with the residual coefficients from adjacent sub-bands, and then they are clustered into different categories by an improved fuzzy C-means clustering algorithm. In order to avoid the overflowing problem caused by having only a few samples in one category, a threshold method is adopted to estimate the correlation noise parameter, which is useful to help decode the corresponding sub-band. Then the reconstructed sub-band is used to update the next subband eigenvectors to obtain more accurate eigenvectors. All the sub-bands are decoded using the same process. Fuzzy C-means algorithm is sensitive to the initial clustering centers. A method that is to produce a random membership degree matrix before iteration can solve this problem to some extent. **Result** Considering both the algorithm performance and complexity, the subband residual coefficients are clustered to eight classes. The experimental results show that this method can simulate accurately the different channel noise characteristics of different region in one frame. What's more, the more complex the video motion, the more obvious the performance superiority of the new algorithm. Compared with that of the subband level

收稿日期: 2013-05-23; 修回日期: 2013-09-03

基金项目: 国家自然科学基金项目(60972135)

第一作者简介: 杨春玲(1970—), 女, 教授, 于南京理工大学获电子工程通信与信息系统专业博士学位, 主要研究方向为图像/视频压缩、图像质量评价。E-mail: eeclyang@scut.edu.cn

Laplace method, the average online rate-distortion performance can be improved up to 1 dB. **Conclusion** A new correlation noise model based on an improved fuzzy C-means clustering algorithm is proposed in this paper. The experiment results show that the rate-distortion performance based on the new algorithm is better than that of sub-band Laplacian solution and the Laplacian-Cauchy mixture model, and the more complex video motion, the more obvious performance gain for the new algorithm.

Key words: Wyner-Ziv distributed video coding; correlation noise modeling; fuzzy C-means clustering algorithm; eigen-vectors; laplacian parameter

0 引言

与传统的视频编码技术(如 H. 264)相比,分布式视频编码(DVC)是一种全新的编码技术,它主要从 Slepian-Wolf 提出的分布式无损编码理论及 Wyner-Ziv 提出的分布式有损编码理论发展而来^[1-2]。分布式视频编码将运动估计和运动补偿等复杂运算从编码端移到解码端,这种方法对新近出现的计算能力、存储能力和功率等资源都受限的无线传输应用场合提供了很好的解决方案。因此近十几年来,分布式视频编码越来越受到关注。

基于 Turbo 码的变换域 Wyner-Ziv (TDWZ) 视频编码是一种被广泛关注的 DVC 编码方案。它将解码端边信息(SI)看做是原始 WZ 帧通过虚拟信道传输后的加噪信号,解码端通过纠错译码得到恢复的视频序列。纠错的过程中,原始信息与边信息之间相关噪声模型(CNM)对压缩性能影响很大。相关噪声模型估计越准确,Turbo 码成功解码需要的校验位就越少,编码效率就越高。所以相关噪声模型的精确与否很大程度上影响着分布式视频编码的率失真性能。

早前关于 CNM 模型的文献大多采用离线拉普拉斯噪声模型(理想情况下)来模拟 TDWZ 中的虚拟信道,但离线估计方法需要用到原始 WZ 帧信息,不能用于实际中。2008 年,Brites 等人^[3]提出的在线子带级拉普拉斯噪声模型及参数估计方法,由于其实际可行、算法简单,得到了学者的广泛认可,目前大多分布式视频编码的相关文献都针对残差子带估计相关噪声模型和参数^[4-8]。王凤琴等人^[4]发现残差变化后直流系数带统计分布相比零均值拉普拉斯分布存在一定的偏差,提出了一种基于变换域的相关噪声分布模型自适应参数估计方法,有效提高了编码效率。Li 等人^[5]针对边信息对相关噪声模型和重建质量的影响,提出了边信息改进方法和利用部分已解码的 WZ 帧对相关噪声模型进行更新估

计的算法,获得了系统性能的提高。Wang 等人^[6]深入研究相关噪声的动态特性,提出了一种自适应相关噪声估计算法来获得系统性能的提高。尹明等人^[7]研究发现相关噪声信息的分布并不满足某种单峰分布,提出了基于高斯混合模型和样本特征的期望最大算法的相关噪声模型,能更精确地描述相关噪声。本研究小组^[8]也针对变换域大残差和小残差系数统计分布与传统的拉普拉斯分布存在一定偏差的事实,提出了一种拉普拉斯-柯西混合分布(LCMD)相关噪声模型及参数估计算法,该算法计算简单,能够对相关噪声模型进行较好的估计。尽管围绕着如何提高噪声模型的精度做了不少研究,但是,视频信息固有的非平稳特性和解码端不能获得当前帧信息等原因,仅针对一个残差子带考虑并不能精确反映出相关噪声的空域运动变化特性。此外,对于边信息的估计并非对于单像素而言,而是利用了相邻的像素之间的关系来获取运动矢量,反映到 DCT 域上,子带间并非孤立的,而是具有某种相似的运动特性,对应的相关噪声分布同样有一定的相关性。特别是当相关噪声获取并不精确的情况下,对子带单独求取模型参数会造成更大的误差。

基于以上特点,Luong 等人^[9-10]试图利用子带之间的这种联系,对每一个子带,聚类时利用 4×4 块的 16 个 DCT 系数形成的 16 维特征矢量,聚类后求取不同类的模型参数,从而得到了较好的压缩性能。但由于在聚类时样本的特征矢量维数太高(16 维),每个子带需在解码前面子带得到新的样本特征矢量后重新聚类,且算法中还把每个系数得到的聚类后拉普拉斯参数和系数级拉普拉斯比较后做选择,导致文献[9]中算法复杂度相当高。此外,虽然 LCMD 模型^[8]通过对残差子带系数进行分类,在一定程度上改进了子带拉普拉斯噪声模型存在的对小系数赋予较低置信度而对大系数赋予较高置信度的缺陷。但是进一步的研究发现,仅对残差系数分为两类,并没有完全挖掘通过

残差系数分类来获得系统性能提高的潜力。为了弥补以上不足,进一步提高相关噪声模型和参数估计的准确度,且适当降低模型参数估计算法的复杂度,在前人研究的基础上,深入探讨利用前后子带之间的相互联系,提出了基于改进模糊 C 均值聚类 (FCM) 的模型参数估计方法,有效地提高了系统率失真性能。

1 TDWZ 中在线噪声模型估计

图 1 给出了 TDWZ 分布式视频编解码框图^[3],

本文算法改进内容如图中阴影方框部分所示。在此系统编码端,视频帧序列被分为关键帧 (Key 帧) 和 WZ 帧,对于 WZ 帧,经过 DCT 变换、量化、格雷码码平面分离等过程形成码流,进入到 Slepian-Wolf 编码器对码平面进行 Turbo 码编码,生成对应的校验信息。Key 帧采用传统的 H.264/AVC 帧内编码方法进行编码,在解码端对接收到的 Key 帧信息进行帧内解码,利用已解码的 Key 帧信息进行运动估计、插值得到边信息 SI 帧,利用生成的边信息和传送到解码端的校验位信息进行 Turbo 码解码,然后经过重建恢复出原始 WZ 帧。

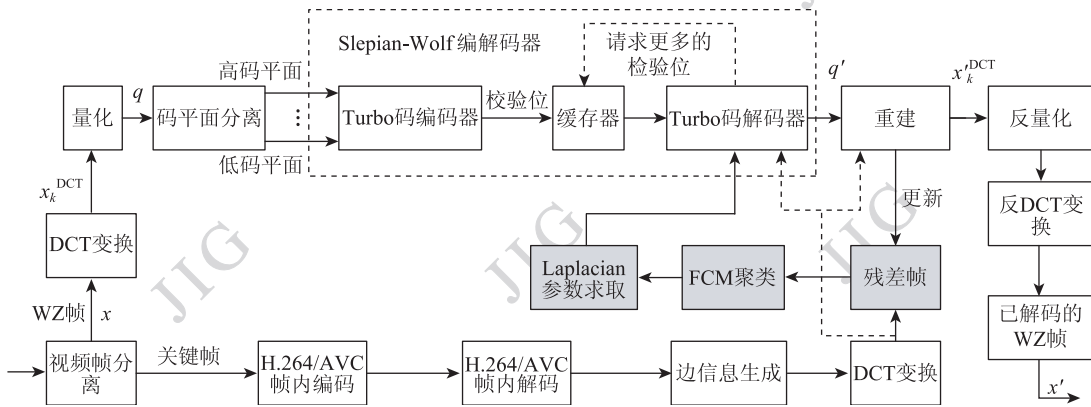


图 1 DCT 变换域的 Wyner-Ziv 视频编码系统框图

Fig. 1 DCT transform domain Wyner-Ziv video coding system framework

从实际应用角度考虑,解码端无法获得原始 WZ 帧信息,所以相关噪声模型基本围绕在线子带^[3]来建立,具体步骤为:

1) 利用边信息生成过程中得到的前后向补偿帧来获得残差帧 $N(x, y)$, 即

$$N(x, y) = X_f(x + dx_f, y + dy_f) - X_b(x + dx_b, y + dy_b) \quad (1)$$

式中, $X_f(x + dx_f, y + dy_f)$ 和 $X_b(x + dx_b, y + dy_b)$ 分别为前向和后向运动补偿帧。

2) 对 $N(x, y)$ 进行 DCT 变换, 得到 16 个子带的 DCT 残差系数集合 $R_j, j \in \{0, 1, \dots, 15\}$ 。 $R_j(u, v)$ 表示位 (u, v) 处的 DCT 子带残差系数。

3) 求取子带 j 的拉普拉斯模型参数

$$p(R_j) = \frac{\alpha_j}{2} \exp(-\alpha_j |R_j|) \quad (2)$$

$$\alpha_j = \frac{\sqrt{2}}{\sigma_j}, \sigma_j = \sqrt{E(|R_j|^2) - E(|R_j|)^2} \quad (3)$$

式中, $p(\cdot)$ 为拉普拉斯概率密度函数, α_j 为子带 j

拉普拉斯分布参数, σ_j 为残差子带 j 的标准差。

在子带级拉普拉斯分布噪声建模方法中,同一子带的残差系数对应相同的拉普拉斯参数,从而导致对小残差系数赋予了较低的置信度,而对大残差系数赋予了较高的置信度,因此得到的估计模型不准确,这从图 2 给出的 soccer 第 58 帧的 AC_1 子带直方图统计分布和相应子带级拉普拉斯分布可以看出。针对这一问题,文献[8]提出的拉柯混合模型,有效地弥补同一子带用同一拉普拉斯参数的不足,拉柯混合模型对于小残差系数能够获得更大的置信度,对于大残差系数可以获得较小的置信度,进一步地逼近残差系数的尖峰和重尾分布。但是,从图 2 中可以看出,对于小残差系数,拉柯模型所获得的分布参数跟真实分布仍存在差距;对于大残差系数,柯西模型可以延缓分布曲线的衰减,但是并不能非常精确地描述真实分布。然而,拉柯模型的思想是使得残差系数在 $[-TH, TH]$ 区间内的柯西分布积分等于拉普拉斯分布积分,这样的处理办法注定只

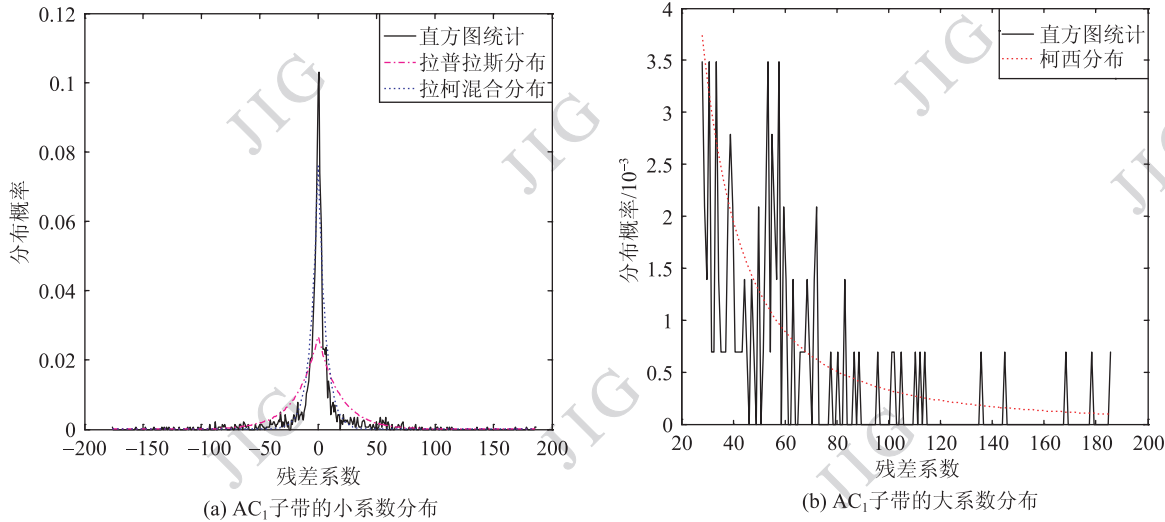


图2 soccer序列第58帧中 AC_1 系数统计特征及模型分布

Fig. 2 The 58th frame of soccer small and large coefficient of AC_1 subband statistical and distribution model

能对残差系数分为有限的两类 TH 表示大小系数的分界值,小系数用拉普拉斯模型建模,大系数用柯西建模。但是通过分析,可以将残差系数聚成更多的类别,对于小残差系数获得更小的方差,从而得到更大的分布参数来逼近真实分布;对于大残差系数,分成不同的类别获得不同程度的衰减,从而获得更精确的估计。

然而,同一帧边信息的质量随空间位置不同而变化,因此不同区域的信道噪声特性是不一样的,反应在 DCT 域上,信道噪声特性在同一子带不同位置区域也是不同的。但在视频帧内的运动过程中,像素与像素之间的运动特性某种程度上具有非独立性,DCT 子带与子带之间并不是孤立存在的,而是有着某种相互影响的关系^[10],所以如果仅对单个残差子带系数进行聚类的话,这样得出的分布参数并没有将子带间的关系利用起来。所以,将残差子带 DCT 变换之后按 zig-zag 的顺序排列,把相邻的子带中对应位置的残差系数组成特征矢量,然后对残差子带进行聚类,对聚类后的类别求取相应的拉普拉斯分布参数。

2 基于改进 FCM 聚类的相关噪声模型估计

首先详细介绍本文所提出的基于 FCM 聚类的相关噪声模型估计方法和步骤,然后阐述改进的 FCM 初始聚类中心方法,最后进一步分析算法对残差系数分布的改善。

2.1 FCM 聚类噪声模型估计算法

模糊 C 均值(FCM)聚类算法^[11-12]由于具有理论简单、易于理解、计算迅速等特点,且与 K-Means 聚类方法相比聚类效果更好,目前已被应用到多个领域^[13-14]。FCM 算法通过优化目标函数得到样本点对所有类别的隶属度,从而决定该样本点的类属,以达到自动对数据样本进行分类的目的。基于 FCM 聚类算法的以上优点,提出了利用 FCM 聚类算法来对分布式视频编码中在线子带残差系数进行聚类的思想,提高拉普拉斯模型参数精度,从而改善系统压缩性能。FCM 聚类噪声模型估计算法主要有以下 4 个步骤:

1) DCT 残差子带生成

把式(1)得到的残差帧分成 4×4 的不重叠块, DCT 变换后得到的 16 个 DCT 子带按照 zig-zag 顺序进行排列,如图 3 所示。第 j 个 DCT 子带用 R_j 表示,其中 $0 \leq j \leq 15$ 。每个子带由 N 个残差系数值组成, R_j^n 表示残差系数帧中第 j 个子带第 n 个系数值, $1 \leq n \leq N$ 。

0	1	5	6
2	4	7	12
3	8	11	13
9	10	14	15

图3 DCT子带 zig-zag 解码顺序

Fig. 3 DCT subband zig-zag decoding order

2) DCT 残差子带聚类

对于当前解码子带 j , 每个系数和相邻的两个子带中对应的系数一起组成一个 3 维的特征矢量。在对 DC 子带进行聚类时, 选取 AC_1 和 AC_2 组成特征矢量, 对最后一个子带, 选取前面两个子带组成特征矢量。这样做一方面是为了与提出的以前子带组成 3 维特征矢量保持一致; 另一方面是因为一个子带解码重建后, 会对该子带进行更新, 所以对于子带 0 和子带 1, 子带 T 和 $T-1$, 其特征矢量是不一样的, 考虑到子带间的相关性, 并且充分利用已经解码重建的子带。残差特征矢量组成具体为

$$\mathbf{E}_j^n = \begin{cases} (R_0^n, R_1^n, R_2^n) & j=0 \\ (R_{j-1}^n, R_j^n, R_{j+1}^n) & 1 \leq j \leq T-1 \\ (R_{T-2}^n, R_{T-1}^n, R_T^n) & j=T \end{cases} \quad (4)$$

式中, $T+1$ 表示 DCT 系数量化^[15]后所需编解码的子带数目。

得到所需解码子带 j 的残差特征矢量样本之后, 用提出的改进 FCM 方法聚类。

3) 聚类后拉普拉斯参数估计

对于子带 j 对应的残差特征矢量样本聚类后, 对于各类残差系数, 用式(5)求取其拉普拉斯参数, 即

$$\alpha_{jk} = \begin{cases} \sqrt{2} & \text{var}(\mathbf{R}_{jk}) \leq 1 \\ \sqrt{\frac{2}{\text{var}(\mathbf{R}_{jk})}} & \text{其他} \end{cases} \quad (5)$$

式中, α_{jk} 表示子带 j 中属于第 k 类残差系数的拉普拉斯分布参数值, $\mathbf{R}_{jk}$ 表示子带 j 中的第 k 类残差系数。

如果第 j 类样本数很少 (其中 $1 \leq j \leq c$, c 为聚类的类别数) 导致方差小于 1, 为了防止溢出, 则该类的拉普拉斯参数就设为 $\sqrt{2}$ 。大量实验结果表明此阈值设置能够有效地抑制某类别残差样本过少而参数过大现象。

4) 信息更新

为了进一步提高压缩效率, 解码下一个子带时, 利用上一个子带重建后的信息更新策略, 即当对子带 $j+1$ 聚类时, 其特征矢量为: $\mathbf{E}_{j+1}^n = (\hat{R}_j^n, R_{j+1}^n, R_{j+2}^n)$, 其中, $\hat{R}_j^n$ 表示重建子带 j 后更新的第 n 个残差系数, 即已重建子带 j 和边信息子带 j 的差。

更新完重建后的子带 j 残差系数之后, 回到步骤 2) 对子带 $j+1$ 对应的残差特征矢量进行聚类, 直

到所有的子带解码完毕。

2.2 改进的 FCM 算法

FCM 算法对于聚类中心的初值选择十分敏感, 如果聚类中心初值选择不当, 则整个聚类迭代过程很容易陷入局部最优, 这将严重影响到聚类最终结果, 从而影响整个聚类的精度。针对这一问题, 采用了一种改进方法, 即先随机产生隶属度矩阵, 归一化处理后, 利用等价迭代表达式求取初始聚类中心; 然后进行迭代聚类过程。实验结果表明此方法会从很大程度上避免聚类陷入局部最优, 本文算法能够对子带残差系数进行很好的聚类。

以子带 j 为例, 改进的 FCM 的聚类过程具体描述如下:

1) 初始化 初始化包括确定聚类类别数目 c 和各聚类中心的初始化。

首先用随机数对隶属度矩阵 μ 初始化, 并经归一化使其满足 $\sum_{i=0}^c \mu_{in}^k = 1$, μ_{in}^k 是子带 j 中第 n 个样本属于第 k 个聚类中心的隶属度; 利用随机生成的隶属度矩阵 μ , 得到各类初始聚类中心 $\mathbf{v}_{jk}$, 即

$$\mathbf{v}_{jk} = \frac{\sum_{n=1}^N (\mu_{jn}^k)^m \mathbf{E}_j^n}{\sum_{n=1}^N (\mu_{jn}^k)^m}, 1 \leq k \leq c \quad (6)$$

式中, $\mathbf{E}_j^n$ 为残差特征矢量。 $\mathbf{v}_{jk} = [v_{jk}^1, v_{jk}^2, v_{jk}^3]$ 表示残差系数子带 j 中第 k 个聚类中心, 同时令迭代次数 $r=1$ 。

2) 聚类

(1) 隶属度矩阵 μ 的更新 隶属度矩阵 μ 更新为

$$\mu_{jn}^k = \frac{\left(\frac{1}{\|\mathbf{E}_j^n - \mathbf{v}_{jk}\|^2} \right)^{1/(m-1)}}{\sum_{k=1}^c \left(\frac{1}{\|\mathbf{E}_j^n - \mathbf{v}_{jk}\|^2} \right)^{1/(m-1)}} \quad (7)$$

$$k=1, 2, \dots, c; n=1, 2, \dots, N$$

(2) 更新聚类中心 得到更新的隶属度矩阵 μ 之后, 再利用式(6)更新各类聚类中心 $\mathbf{v}_{jk}$ 。并利用式(8)判断聚类是否结束。

$$\max \| \mathbf{v}_{jk}^{(r)} - \mathbf{v}_{jk}^{(r-1)} \| < \varepsilon, 1 \leq k \leq c \quad (8)$$

若式(8)成立, 则停止迭代, 完成子带 j 对应残差特征矢量的聚类; 否则返回(1), 使得 $r=r+1$ 。本文迭代阈值设为 0.001。

2.3 模型分布效果分析

通过改进 FCM 聚类算法对残差子带进行聚类,

来获得更加精准的拉普拉斯分布参数。由于本文算法并非对单独的子带残差系数聚类,通过相邻子带之间的相关性来获取残差系数之间的所属类别。为了能用分布曲线直观地描述本文算法的效果,分别将与文献[3]和文献[8]一致的残差系数对应的模型参数描绘如图4所示。图4给出了 soccer 序列第58帧 AC_1 子带聚类后的模型分布曲线,从图4中可以看出,精细化地将小系数聚为一类,会得到更小的方差,相比子带级拉普拉斯分布和拉柯分布,聚类后的分布更加逼近其所具有的尖峰分布,从而可以获得更加精准的噪声模型估计。

为了验证残差系数聚类类别数对系统性能的影响,图5示出了将残差系数分别聚为2类、4类、8类和16类,对 foreman 序列在第8个RD点平均PSNR相当(40.55 dB)的条件下解码每帧所需比特分布图。从图5中可以看出,随着聚类类别数的增加,解码每帧所需的码率逐渐减少,但是当类别数为16时,码率的减少趋缓与聚为8类每帧所需码率相当,但是会带来很大计算复杂度。对于其他序列也有类似

情况,篇幅限制,不一一列出。所以折中考虑,将残差子带聚为8类,下一节将具体分析系统率失真性能。

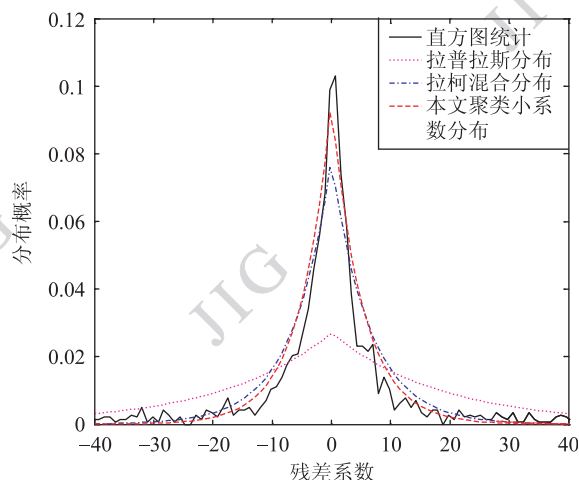


图4 soccer 序列第58帧 AC_1 子带本文算法聚类效果图

Fig. 4 Result of the 58th frame of soccer AC_1 subband proposed clustering algorithm

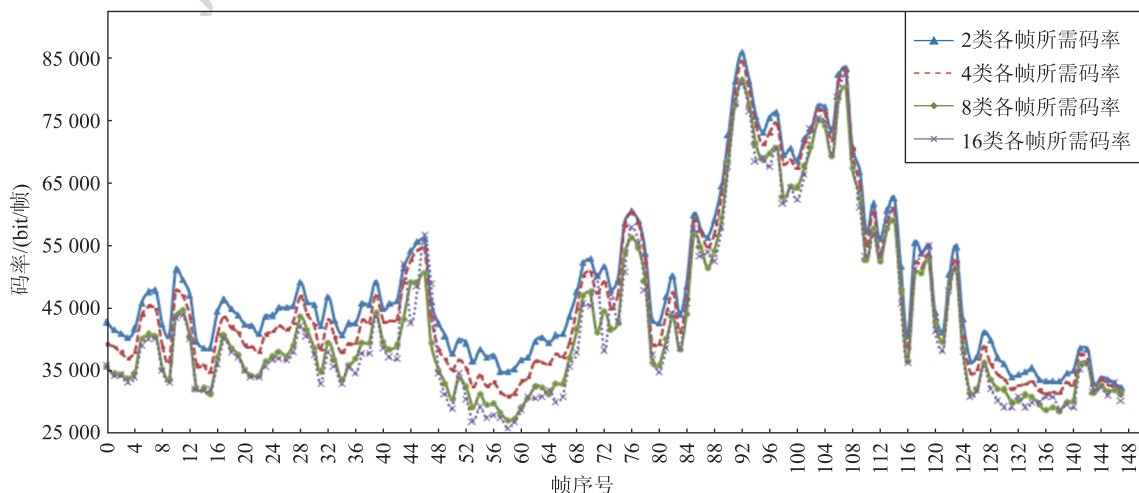


图5 foreman 序列聚为2类、4类、8类和16类各帧解码所需的码率

Fig. 5 The rate required by each foreman frame divided into 2, 4, 8, 16 classes

3 实验仿真结果及分析

为验证提出改进 FCM 聚类算法对 TDWZ 系统 RD 性能的改善,实验中选用3种具有典型代表性的测试序列: coastguard、foreman 和 soccer, 分别对应中等、高运动程度的视频序列,分辨率为 QCIF 格式,帧率为 30 Hz, GOP (group of pictures) 长度为 2, 编码前 299 帧(149 个 WZ 帧)。测试序列既包含运

动平缓部分又包含运动剧烈部分,使得结果更具有普遍性。Key 帧采用 H. 264/AVC Intra 编码,选取合适的量化参数 QP (quantization parameter) 值使得关键帧和 WZ 帧保持质量相当^[15]。边信息生成方法采用 SEB (side estimation bidirectional) 算法^[16]生成,解码端采用 VRS (variable reconstruction scheme) 方法^[17]进行重建。在线相关噪声建模由 Brites 等人^[3]提出,所以将仿真结果与其进行比较。由于本文算法是将前后子带组成特征矢量进行聚类,同时

也与本研究小组之前提出的分成两类的 LCMD (Laplacian-cauchy mixture distribution) 模型^[8]和基于块聚类^[9]的进行了压缩性能比较。最后, 基于本文算法与块聚类算法^[9]性能差异, 比较了不同算法的时间复杂度。

3.1 算法 RD 性能分析

各噪声分布模型的压缩性能比较如图 6(a) — (c) 所示。实验结果表明, 无论是相比于在线子带级^[3], 还是在线 LCMD^[8], 本文算法都能使率失真性能得到显著改善。

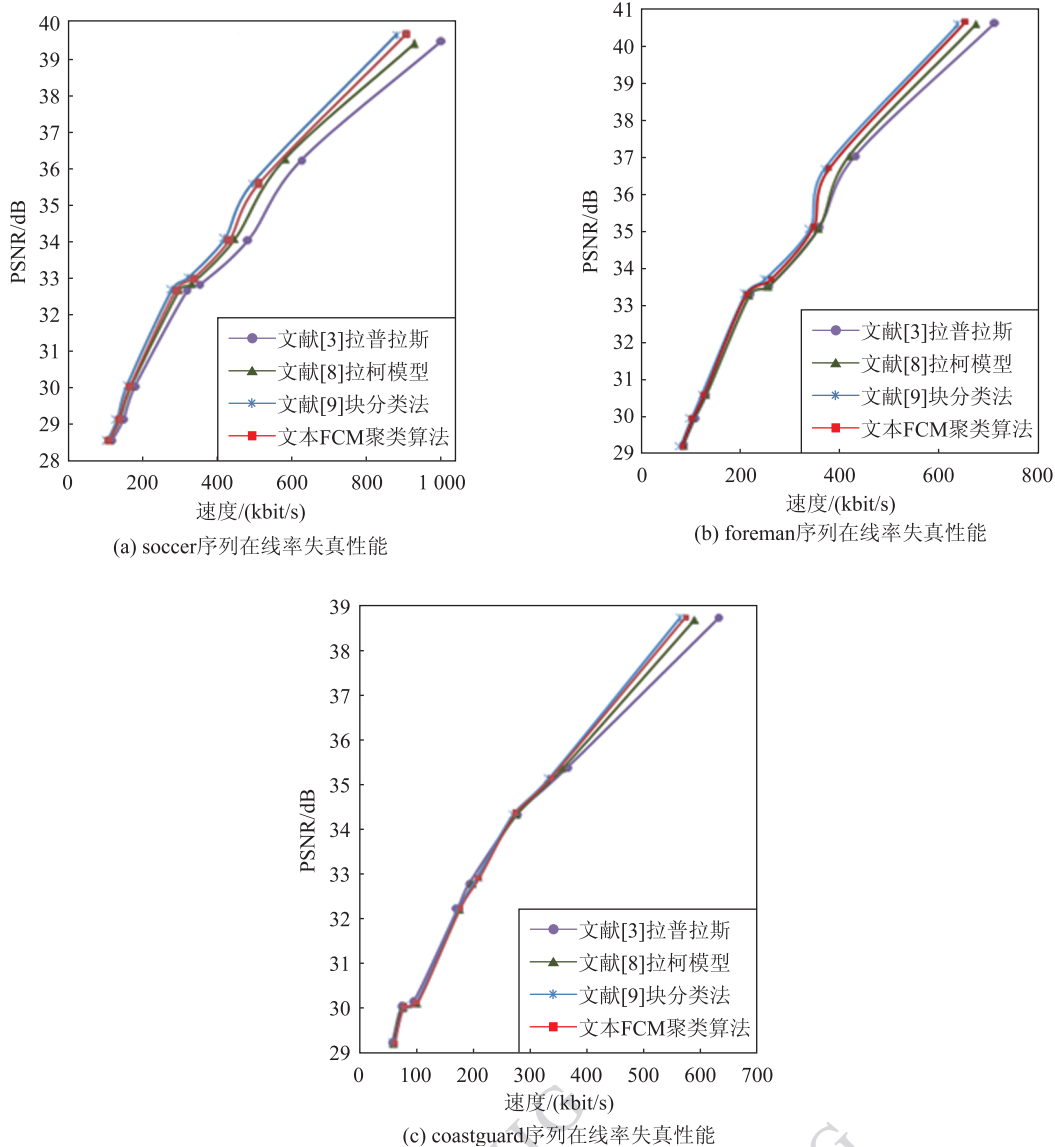


图6 各噪声模型的在线压缩性能比较

Fig. 6 The comparison of compression performance of the online noise models

由图 6(a) — (c) 结果可知, 对于不同运动程度的视频序列, 获得的增益不同。对于运动中等程度的 foreman 和 coastguard 序列, 相对于子带级^[3], 本文算法可以获得 0.8 dB 的性能提升; 对于运动程度剧烈的 soccer 序列, 可获得 1.3 dB 的提高。相对于拉柯混合模型^[8], 本文算法也能获得性能的提高, 对 foreman 和 coastguard 序列, 所获得的平均增益分别为

0.4 dB 和 0.2 dB; 对 soccer 序列, 相应地有 0.5 dB 的提高。从图 6 中可以看出, 一般对于运动程度越剧烈的序列, 获得的增益越大; 对于运动越平缓的序列, 本文算法所带来的增益越小。这主要是因为视频序列运动剧烈程度对边信息的影响程度不同造成的。对于运动相对平缓的序列, 边信息质量较高, 相应得到的子带残差系数方差就越小, 获得的残差较接近

真实尖峰的分布,从而改进空间较小。但是对于运动越剧烈的序列,由于生成的边信息质量较差,残差系数值较大,子带级的参数估计方法将会造成相当大的误差,而利用其前后子带间的相关性会获得更加可靠的估计,从而减小误差。从文献[8]得知,soccer 序列子带级拉普拉斯参数与真实残差直方图分布相对于运动中等的 foreman 序列差别更大,而本文算法聚类后参数分布更加接近真实的残差分布,获得的增益就越大。这也进一步说明本文算法将子带残差系数聚为 8 类比 2 类更加精准。

同时,从图 6 也可以看出,对于 foreman 和 coastguard 序列,随着量化步长的减小,本文算法获得的增益越发明显,一方面是因为对于大量化步长矩阵,关键帧的量化步长 Q_{step} 也相对较大,量化比较粗糙;另一方面是因为量化步长越大,所需传输的子带就越少,本文算法的优越性也就越不明显。但是对于较小量化步长,传输的子带较多,本文算法带来的优越性更加明显。

相对于较复杂的块聚类^[9]算法,本文算法对各

个序列有 0.1~0.2 dB 的差距,因为基于块分类更大程度上利用了不同子带之间的关系。但是相对于时间复杂度而言,它所获得的性能提高微乎其微。

3.2 模型复杂度分析

在 DVC 中,系统率失真性能是衡量算法优劣的一方面,而算法的复杂度对于算法能否应用于实际中也非常重要。表 1 给出了不同算法的解码时间复杂度对比。实验代码的软件环境为 Visual Studio 2008,使用 3.1 GHz 主频和 4.0 GB 内存,Windows7 操作系统的 PC 机。从表 1 可以看出,本文算法相对于子带级拉普拉斯参数估计的确由于迭代引入了较大的计算量,但是由于本文算法可以获得较大系统性能的提高,而 Wyner-Ziv 编码架构的主要目标是寻求“编码简单,解码复杂”,所适合的应用也允许这种情况(解码端通常是具有较强计算性能的中心节点)出现。然而,相比于块聚类^[9]算法,本文算法的运算复杂度要小很多,而文献[9]的压缩性能比本文算法的提高微乎其微。因此,对于系统性能提升和复杂度折中考虑,本文算法具有较大的优势。

表 1 不同算法解码时间比较
Table 1 Different algorithm time complexity

RD 点	soccer			coastguard			foreman		
	文献[3]	本文算法	块聚类 ^[9]	文献[3]	本文算法	块聚类 ^[9]	文献[3]	本文算法	块聚类 ^[9]
Q0	1.891	6.189	10.371	1.576	4.94	8.665	1.264	4.831	10.815
Q1	2.434	6.764	10.875	2.201	4.912	9.8	1.99	5.181	11.346
Q2	3.183	11.185	18.269	3.44	8.77	13.189	2.913	8.665	16.735
Q3	5.724	18.103	30.899	6.097	13.444	21.134	4.408	12.983	28.257
Q4	6.653	21.158	36.983	6.49	16.837	24.18	5.791	15.28	29.805
Q5	9.157	22.954	41.712	7.598	18.766	30.02	7.451	17.065	34.39
Q6	11.35	24.231	44.729	9.13	19.174	36.879	8.881	17.292	38.16
Q7	18.56	28.965	52.273	14.087	22.175	42.99	14.381	20.657	45.739

4 结 论

针对子带级 laplacian 模型不能准确描述相关噪声残差子带不同区域的变化特性的问题,以及考虑到现有算法未将相邻子带之间关系利用起来,提出了基于改进 FCM 聚类的噪声模型估计方法;针对 FCM 聚类效果对于初始聚类中心的敏感性这一问题,采用改进初始聚类中心方法。实验结果表明,本文算法获得的系统率失真性能得到明显改善,尤其

对于运动越剧烈的视频序列,增益更加明显。此外,寻求低复杂度的有效算法进行相关噪声估计也是本文需要进一步研究的工作。

参考文献 (References)

[1] Slepian D, Wolf J. Noiseless coding of correlated information sources [J]. IEEE Transactions on Information Theory, 1973, 19(4):471-480. [DOI:10.1109/TIT.1973.1055037]
[2] Wyner A, 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. [DOI: 10.1109/TIT.1976.1055508]
- [3] 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. [DOI:10.1109/TCSVT.2008.924107]
- [4] Wang F Q, Fan Y Y, Zhao J, et al. Correlation noise model for transform domain Wyner-Ziv video coding[J]. Journal of Data Acquisition & Processing, 2009, 24(5): 609-614. [王凤琴, 樊养余, 赵炯, 等. 基于变换域 Wyner-Ziv 视频编码的相关噪声模型[J]. 数据采集与处理, 2009, 24(5): 609-614.]
- [5] Li B, Wang Y M, Liu Y, et al. An improved method for side information generation and correlation parameter estimation for distributed video coding[C]//Proceedings of IEEE Globecom Workshops. Houston, TX: IEEE 2011; 40-44. [DOI:10.1109/GLOCOMW.2011.6162481]
- [6] Wang S, Cui L J, Stankovic L, et al. Adaptive correlation estimation with particle filtering for distributed video coding[J]. IEEE Transactions on circuits and systems for video Technology, 2012, 22: 649-658. [DOI:10.1109/TCSVT.2011.2171263]
- [7] Yin M, Cai S T, Xie Y. Wyner-Ziv video coding based on Gaussian mixture model[J]. Chinese Journal of Computers, 2012(1): 173-182. [尹明, 蔡述庭, 谢云. 基于高斯混合模型的 Wyner-Ziv 视频编码[J]. 计算机学报, 2012(1): 173-182.]
- [8] Yang C L, Xie X L. A Laplacian-Cauchy mixture model for improved correlation noise modeling in DVC[J]. Journal of Electronics and Information, 2012, 34(10): 2348-2352. [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.]
- [9] Van Luong H, Huang X, Forchhammer S. Adaptive noise model for transform domain Wyner-Ziv video using clustering of DCT blocks[C]//Proceedings of the 13th International Workshop on Multimedia signal processing (MMSP). Hangzhou, China: IEEE 2011; 1-6. [DOI:10.1109/MMSP.2011.6093774]
- [10] Huang X, Forchhammer S. Cross-band noise model refinement for transform domain wyner-ziv video coding[J]. Signal Processing: Image Communication, 2012, 27(1): 16-30. [DOI:10.1016/j.image.2011.06.008]
- [11] Bian Z Q, Zhang X G. Pattern recognition. [M]. 2nd ed. Beijing: Tsinghua University Press, 2000. 1: 273-283. [边肇祺, 张学工. 模式识别[M]. 2版. 北京: 清华大学出版社, 2000: 273-283.]
- [12] Wang Z F. Comparison of four kinds of fuzzy C-Means clustering methods[C]//Proceedings of 2010 Third International Symposium on Information Processing. Qingdao, China: IEEE, 2010: 563-566. [DOI:10.1109/ISIP.2010.133]
- [13] Zhu X Y, Yu Q Y, Huang J. Color image segmentation algorithm based on FCM clustering[D]. Jilin: Jilin University, 2011. [朱小雨, 虞强源, 黄晶. 基于FCM聚类的彩色图像分割算法研究[D]. 吉林: 吉林大学, 2011.]
- [14] Linda O, Manic M. General type-2 Fuzzy C-Means algorithm for uncertain fuzzy clustering[J]. IEEE Transactions on Fuzzy Systems, 2012: 883-897. [DOI:10.1109/TFUZZ.2012.2187453]
- [15] Artigas X, Ascenso J, Dalai M, et al. The DISCO-VER codec: Architecture, techniques and evaluation[C]//Proceedings of the Picture Coding Symposium. Lisbon, Portugal: EURASIP, 2007, 17(9): 1103-1120.
- [16] Li Z, Liu L, Delp E J. Rate distortion analysis of motion side estimation in Wyner-Ziv video coding[J]. IEEE Transactions on Image Processing, 2007, 16(1): 98-113. [DOI: 10.1109/TIP.2006.884934]
- [17] Yang C L, Lin Y G, Po L M, et al. Variable reconstruction scheme for distributed video decoding[C]//Proceedings of Tenccon 2008—2008 IEEE Region 10 Conference. Hyderabad, India: IEEE, 2008: 1-5. [DOI: 10.1109/TENCON.2008.4766520]