





第19卷第6期（总第218期）

2014年6月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 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

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@irsa.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司
(邮政信箱：北京399信箱 邮编：100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@irsa.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

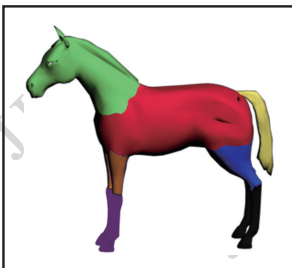
国内定价 50.00元



帧内复制粘贴视频伪造的盲检测(第825页)



基于流形学习的人体动作识别(第914页)



层次结构导向下的动画角色模型分割与标记(第955页)

综述

虹膜图像质量评价综述

李星光, 孙哲南, 谭铁牛

0813

图像处理和编码

帧内复制粘贴视频伪造的盲检测

杜振龙, 焦丽鑫, 李晓丽, 郭延文, 杨小健

0825

破损区域分块划分的图像修复

翟东海, 鱼江, 段维夏, 肖杰

0835

分布式视频编码中相关噪声分布的非参数估计

胡晓飞, 朱秀昌

0843

分数阶原始对偶去噪模型及其数值算法

田丹, 薛定宇, 杨雅婕

0852

结合感知特征和自然场景统计的无参考图像质量评价

贾惠珍, 孙权森, 王同罕

0859

图像置乱度评估的层次分析法

曹光辉, 胡凯, 张兴

0868

图像分析和识别

结合目标预测位置的压缩跟踪

罗会兰, 钟宝康, 孔繁胜

0875

标牌粘连字符自适应定位分割重建识别

洪涛, 梁伟建, 卢玉凤

0886

结合色度和纹理不变性的运动阴影检测

殷保才, 刘羽, 汪增福

0896

仿射不变的自适应局部线性嵌入

顾播宇, 孙俊喜, 李洪祚, 刘广文, 曹永刚

0906

基于流形学习的人体动作识别

王鑫, 沃波海, 管秋, 陈胜勇

0914

图像理解和计算机视觉

基于压缩感知理论的3维人脸快速重建

周大可, 吴子扬, 杨欣

0924

压缩感知跟踪中的特征选择与目标模型更新

石武祯, 宁纪锋, 颜永丰

0932

高动态场景的图像融合优化和实时应用

范逵, 周晓波

0940

计算机图形学

真实图像转换的水墨图像绘制模拟

陈添丁, 金炜炜, 陈英旦, 吴淦

0946

层次结构导向下的动画角色模型分割与标记

李琳, 陶涛, 刘晓平

0955

遥感图像处理

静止卫星图像热带气旋云系自动识别

耿晓庆, 李紫薇, 杨晓峰

0964

非下采样Contourlet域融合和参数化内核图割的SAR图像无监督水灾变化检测

李青松, 覃锡忠, 贾振红, 杨杰, 胡英杰

0971

地理信息技术

描述定性方向关系的复合表达模型

王中辉, 杨艳春

0979

CONTENTS

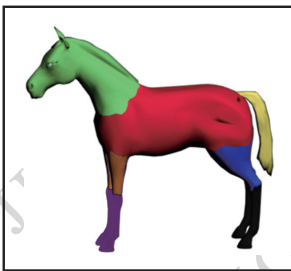
JOURNAL OF IMAGE AND GRAPHICS



Intraframe copy-paste forgery video blind detection based on dense SIFT flow(P825)



Human action recognition based on manifold learning(P914)



Animation character model segmentation and labeling approach under the guidance of hierarchical structures(P955)

Review

- Overview of iris image quality-assessment
Li Xingguang, Sun Zhenan, Tan Tieniu 0813

Image Processing and Coding

- Intraframe copy-paste forgery video blind detection based on dense SIFT flow
Du Zhenlong, Jiao Lixin, Li Xiaoli, Guo Yanwen, Yang Xiaojian 0825
- Image inpainting algorithm based on partition block of damaged region
Zhai Donghai, Yu Jiang, Duan Weixia, Xiao Jie 0835
- Non-parametric estimation for correlation noise distribution in distributed video coding
Hu Xiaofei, Zhu Xiuchang 0843
- Fractional-order primal-dual model and numerical algorithm for denoising
Tian Dan, Xue Dingyu, Yang Yajie 0852
- Blind image quality assessment based on perceptual features and natural scene statistics
Jia Huizhen, Sun Quansen, Wang Tonghan 0859
- Image scrambling degree evaluation based on analytic hierarchy process
Cao Guanghui, Hu Kai, Zhang Xing 0868

Image Analysis and Recognition

- Object tracking algorithm by combining the predicted target position with compressive tracking
Luo Huilan, Zhong Baokang, Kong Fansheng 0875
- Merged characters segmentation reconstruction and recognition of labels based on adaptive location
Hong Tao, Liang Weijian, Lu Yufeng 0886
- Moving shadow detection by combining chromaticity and texture invariance
Yin Baocai, Liu Yu, Wang Zengfu 0896
- Affine invariant adaptive locally linear embedding
Gu Boyu, Sun Junxi, Li Hongzuo, Liu Guangwen, Cao Yonggang 0906
- Human action recognition based on manifold learning
Wang Xin, Wo Bohai, Guan Qiu, Chen Shengyong 0914

Image Understanding and Computer Vision

- Fast 3D face reconstruction based on compressed sensing
Zhou Dake, Wu Ziyang, Yang Xin 0924
- Feature selection and target model updating in compressive tracking
Shi Wuzhen, Ning Jifeng, Yan Yongfeng 0932
- The optimization of image fusion and real-time application for HDR scenarios
Fan Kui, Zhou Xiaobo 0940

Computer Graphics

- Rendering simulation of real-image converting into chinese ink painting
Chen Tianding, Jin Weiwei, Chen Yingdan, Wu Di 0946
- Animation character model segmentation and labeling approach under the guidance of hierarchical structures
Li Lin, Tao Tao, Liu Xiaoping 0955

Remote Sensing Image Processing

- Tropical cyclone auto-recognition from stationary satellite imagery
Geng Xiaoqing, Li Ziwei, Yang Xiaofeng 0964
- Unsupervised detection of flood changes with SAR images combining nonsubsampling Contourlet domain fusion and parametric kernel graph cuts
Li Qingsong, Qin Xizhong, Jia Zhenghong, Yang Jie, Hu Raphael 0971

Geoinformatics

- Compound model for describing direction relations qualitatively
Wang Zhonghui, Yang Yanchun 0979

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2014)06-0843-09

论文引用格式: Hu X F, Zhu X C. Non-parametric estimation for correlation noise distribution in distributed video coding[J]. Journal of Image and Graphics, 2014, 19(6): 843-851. [胡晓飞, 朱秀昌. 分布式视频编码中相关噪声分布的非参数估计[J]. 中国图象图形学报, 2014, 19(6): 843-851.] [DOI:10.11834/jig.20140604]

分布式视频编码中相关噪声分布的非参数估计

胡晓飞, 朱秀昌

南京邮电大学图像处理与图像通信江苏省重点实验室, 南京 210003

摘要: 目的 针对分布式视频编码系统中相关噪声(CN)分布难以准确模拟的问题,提出了一种 CN 的非参数估计方法。方法 根据 CN 分布的特点,提出 CN 的非参数估计方法,建立了基于最优窗宽的核密度估计-均匀分布模型(KDEUDM),比较了变换域 Wyner-Ziv(TDWZ)系统中 CN 的参数估计法和非参数估计法所建立的噪声模型对系统性能的影响。结果 实验结果表明,非参数估计方法能较准确地模拟 CN 的分布,与参数估计法相比,用非参数估计法建立的噪声模型能使 WZ 帧编码在高码率下最高能节约 10% 的码率。结论 非参数估计法是 TDWZ 系统中有效的相关噪声估计方法。

关键词: 分布式视频编码;变换域 Wyner-Ziv;相关噪声分布;核密度估计;噪声模型

Non-parametric estimation for correlation noise distribution in distributed video coding

Hu Xiaofei, Zhu Xiuchang

Image Processing and Image Communication Key Laboratory, Nanjing University of Posts and Communications, Nanjing 210003, China

Abstract: **Objective** Since it is difficult to estimate the correlation noise distribution, a non-parametric estimation method for correlation noise is proposed. **Method** Non-parametric estimation method for correlation noise is proposed according to the feature of correlation noise, a kernel density estimation-uniform distribution model based on optimal bandwidth is proposed. The impact of the noise model using parameter estimation method and non-parametric estimation method on the system performance in the transform domain Wyner-Ziv system is compared. **Result** The experiments results show that non-parametric estimation method can simulate correlation noise distribution and the established noise model using non-parametric estimation method can save rate about 10% at most at high bit rate for WZ frame. **Conclusion** Non-parametric estimation method is an effective noise estimation method for TDWZ system.

Key words: distributed video coding; transform domain Wyner-Ziv; correlation noise distribution; kernel density estimation; noise model

0 引言

虽然以 MPEG-X 和 H. 26X 为代表的传统视频

编码器实现了视频数据的高效压缩,推动了视频编码技术的快速发展,但是由于采用了离散余弦变换(DCT)结合帧间预测技术的混合编码器,在编码端需要完成变换、量化、熵编码、以及运动估计和运动

收稿日期:2013-07-22;修回日期:2014-01-10

基金项目:国家自然科学基金项目(61071091,61271082);南京邮电大学科研基金项目(NY212101)

第一作者简介:胡晓飞(1975—),女,副教授,2013 年于南京邮电大学获信号与信息处理专业博士学位,研究方向为视频信号处理、分布式信源编码。E-mail: fhuxf@126.com

补偿等大量的运算,而解码器仅仅利用来自编码器的信息重建视频信号,这使得编码复杂度远远高于解码复杂度。

近年来兴起的无线视频监控、无线传感器网络、移动可视电话等面向上行链路、低功耗、低带宽的应用需求,其无线视频编码设备简单,且存储容量、计算能力和功率资源等方面都受到很大的限制,而解码设备一般无能量限制且有较强的处理能力。传统的复杂的编码器显然不适用于这种场合,复杂度不对称的结构需要逆转来适应新的应用需求。分布式视频编码(DVC)就是一种适合此类应用场合,最大限度地减小编码复杂度,在解码端挖掘视频信号的时间冗余度的全新的视频编码方案。由于DVC在编码效率上与传统的视频编码框架还有一定的距离,关于DVC技术的应用研究还处于探索阶段。

目前,DVC的研究主要集中在 Wyner-Ziv DVC(WZ DVC)^[1-2]和基于压缩感知的DVC两个方面^[3-4],其中WZ DVC的研究始于2002年^[5]。与基于压缩感知理论的DVC相比,WZ DVC更成熟,是DVC中的主流技术。本文主要以WZ DVC编码框架为基础进行讨论。

在WZ DVC系统中,最理想的情况是在解码端能够通过边信息(SI)帧准确重建WZ帧,但是实际上SI帧和WZ帧之间是存在误差的。这种情况可看做SI帧与信源WZ帧之间有一条虚拟信道,SI帧与WZ帧的误差是由于WZ帧经过该虚拟信道传输后引入了噪声而造成的,WZ帧与SI帧之间的误差,即相关噪声(CN)。如果能准确地模拟CN的分布,

Turbo码或低密度奇偶校验码(LDPC)解码器就能利用SI帧对WZ帧进行译码。CN模型越准确,Turbo或LDPC解码器成功解码需要的校验比特就越少,既降低了编码码率,又降低了解码器解码的复杂度。然而,在DVC系统中,CN的统计特性是未知的、随机的,并且WZ编码器只能准确地获得信源WZ帧的信息,WZ解码器只能准确地获得SI帧的信息,即无论是编码器还是解码器都不能同时准确地得到SI帧和WZ帧的信息,所以,CN分布的准确模拟是WZ DVC系统中一个重要的研究方向,CN的准确建模在WZ DVC系统中起着至关重要的作用。

关于DVC系统中CN分布模型的研究大多是先确定模型的概率密度分布函数,然后用参数估计的方法估计模型的参数。由于Laplace分布能够在复杂度和准确度之间达到较好的平衡,CN通常被描述成Laplace分布,其分布参数通过参数估计的方法得到^[6-7],也有学者利用其他的分布模型来模拟CN的分布,如指数模型^[8]、高斯模型^[9]等。

鉴于变换域Wyner-Ziv(TDWZ)系统与像素域Wyner-Ziv系统相比在编码效率上有了很大程度的提升,又没有过多增加实现的复杂度,以TDWZ系统为编码框架,在文献[10]的基础上,研究CN分布的非参数估计方法,并与参数估计方法相比较,分析其带来的编码性能增益的改善。

1 TDWZ 编码框架

图1是本文采用的TDWZ编码框架。在该编码

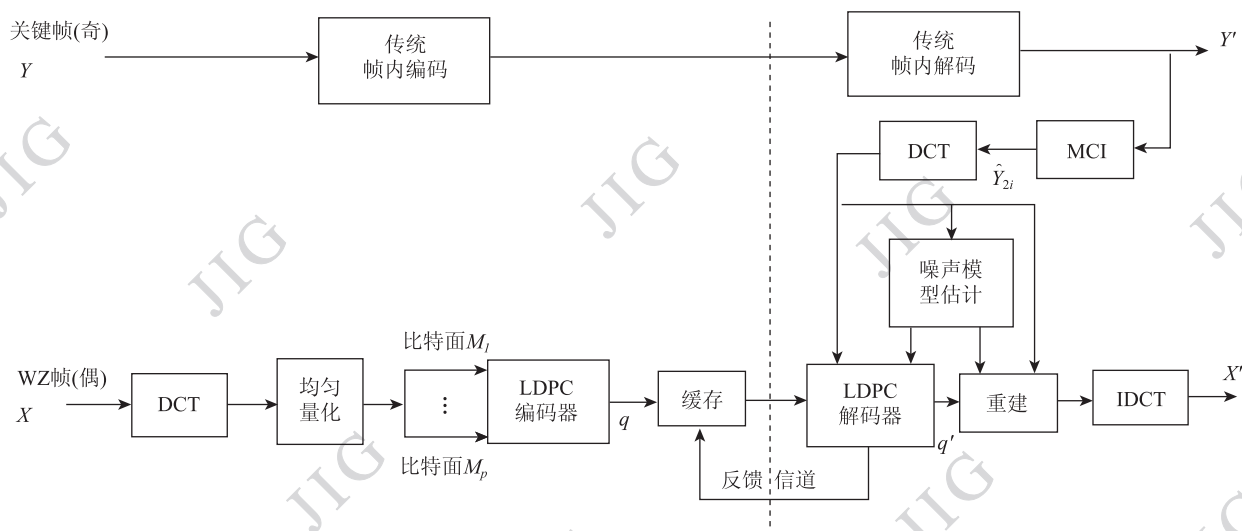


图1 基本的TDWZ编码框架

Fig. 1 Basic TDWZ coding framework

框架中, 图像组 (GOP) 长度固定为 2, 关键帧 (奇数帧) 采用传统的 H. 264 帧内编解码方法。编码 WZ 帧 X (偶数帧) 时将图像分成不重叠的 4×4 大小的块, 首先对每个子块进行 DCT 变换, 并对 DCT 系数均匀量化, 量化后的系数用格雷码表示, 然后输入 LDPC 编码器, 输出得到校验比特 q 。解码时, 先由解码的关键帧 Y'_{2i-1}, Y'_{2i+1} 进行运动补偿帧间插值 (MCI) 得到 WZ 帧的边信息 $\hat{Y}_{2i}$, 再用文献 [7] 的方法进行 CN 模型估计, LDPC 解码器在已知校验比特 q 和边信息 $\hat{Y}_{2i}$ 的条件下解码得到纠错解码系数 q' , 最后经重建和反 DCT 变换 (IDCT) 模块得到重建的 WZ 帧 X' 。

2 CN 分布的估计

2.1 CN 分布的参数估计法

统计学中, 参数估计和非参数估计共同构成了概率密度估计方法。参数密度估计方法首先假定样本分布的概率密度函数类型已知, 而决定概率密度函数的参数具体取值未知, 参数估计的目的是寻找参数的最优估计, 从而获得密度函数的估计。

Laplace 模型是普遍采用的一种 CN 模型, 即

$$p(T(u, v)) = \frac{\alpha}{2} \exp(-\alpha |T(u, v)|) \quad (1)$$

式中, α 表示模型参数, $\alpha = \sqrt{\frac{2}{\sigma^2}}$, σ^2 是 CN 系数的方差, $T(u, v)$ 表示每个子块 CN 的 (u, v) 处的 DCT 变换系数。这里把式 (1) 的 Laplace 分布模型称为单模 Laplace 分布。

虽然噪声模型可以通过求解模型参数 α 得到, 但是按照这种方法计算出来的参数也不能完全描述噪声残差的分布, 也就是说 CN 并不是严格意义上的 Laplace 分布。实际的 CN 比 Laplace 分布具有更高的尖峰和严重的拖尾现象^[11], 因此, 用 Laplace 模型来模拟 CN 分布并不理想。为了准确地模拟 CN 的拖尾和突变, 文献 [10] 提出基于期望最大化 (EM) 算法的二状态 Laplace 模型来模拟 CN 的分布, 即

$$p(x | \theta) = \pi_1 p_1(x | \theta_1) + \pi_2 p_2(x | \theta_2) \quad (2)$$

式中, π_1 和 π_2 是混合系数, 满足 $\pi_1 + \pi_2 = 1$, $p_1(x | \theta_1)$ 和 $p_2(x | \theta_2)$ 服从 Laplace 分布, $\theta_1 = [\mu_1 \ \sigma_1^2]^T$ 、

$\theta_2 = [\mu_2 \ \sigma_2^2]^T$ 分别是两个状态的 Laplace 分布参数, $\theta = [\theta_1 \ \theta_2]$ 。

2.2 CN 分布的非参数估计法

上述基于 EM 算法的 CN 估计方法是一种参数估计方法, 它总是假设概率密度函数的参数形式已知, 即服从二状态混合 Laplace 分布, 但是对于许多实际应用场合, 如 CN 的分布模型, 上述假设并不成立, 特别是所有经典的密度函数的参数形式都是单模的, 所以用参数估计方法去模拟复杂分布时的准确度得不到保证。非参数估计方法不必假设密度的参数形式已知, 它是通过数据点来获得概率密度函数 f 的估计 $\hat{f}$, 得到的密度函数估计的形状主要依赖于数据的结构, 因此能处理任意的概率分布。本文以核密度估计 (KDE) 为代表来讨论 CN 分布的非参数估计方法。

2.2.1 核密度估计^[12-13]

KDE 也被称为 Parzen 窗估计, 假设数据集 $X = \{x_1, \dots, x_N\}$ 为 d 维空间中的 N 个样本点, 核概率密度估计可以表示为

$$\hat{f}(x) = \frac{1}{Nh^d} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right) \quad (3)$$

式中, K 称为核函数 (Kernel function), 也被称为窗函数, h 称为窗宽或带宽。可以看出, 核函数是一种权函数, 核估计利用数据点 x 到 x_i 的距离 $(x - x_i)$ 来决定在 x_i 处估计点 x 的密度时所起的作用。

核函数可以有多种, 通常选择什么核函数不是密度估计中最关键的因素, 因为选择任何核函数都能保证密度估计具有稳定相合性, 最重要的是带宽对估计分布的光滑程度影响很大^[14], 自然地如何选择带宽成为 KDE 中最重要的问题。

2.2.2 CN 的自适应窗宽核密度估计

式 (3) 的 KDE 计算了在估计点周围一个窗体范围内样本点的加权局部平均, 如果 h 太小, 那么密度估计偏向于把概率密度分配得太局限于观测数据附近, 使估计结果有很多错误的峰值; 如果 h 太大, 那么密度估计在很大的邻域里求平均会光滑掉数据的一些重要特征。

一般而言, 带宽的选择应该依赖于样本的数量, 在样本点密集的区域采用较小的带宽, 在样本点稀疏的区域采用较大的带宽。最优带宽的选取一般通过计算概率密度估计与真实概率密度之间的误差, 然后通过最小化积分均方误差 (MISE) 或者渐近积分均方误差 (AMISE) 的方法来计算最优带宽。为

此,令 σ 为给定数据集 X 的总体标准偏差, N 为给定数据集 X 中样本点的数目,选择高斯窗作为窗函数,通过经验式(4)来估计 AMISE 意义下的最优带宽

$$h_{\text{opt}} = 1.06\sigma N^{-\frac{1}{5}} \quad (4)$$

对于偏差较大的数据,用式(4)的最优带宽将会导致拟合过于平滑,于是用式(5)的基于四分位距 R 的带宽最优化方法^[13],即

$$h_{\text{opt}} = 0.79RN^{-\frac{1}{5}} \quad (5)$$

相应地,式(4)就称为基于偏差 σ 的带宽最优化。

对于 Γ 族分布,式(4)显示较好的拟合性能,而式(5)适用于偏差大的有拖尾的数据,为了选择能在两种情况下都取得较好拟合性能的带宽,令

$$A = \min(\sigma, R/1.34) \quad (6)$$

$$h_{\text{opt}} = 0.9AN^{-\frac{1}{5}} \quad (7)$$

CN 分布取决于 SI 帧的质量,其概率密度分布曲线可能是比较复杂的形式,未必服从 Laplace 分布或高斯分布。鉴于 CN 的特殊性,采用高斯窗,选择式(7)的最优窗宽,用 KDE 的方法估计 CN 的分布情况。

2.3 性能分析与比较

2.3.1 拟合准确度比较

分别选取国际标准化图像格式 QCIF 格式的 Forman 序列和 Akiyo 序列第 22 帧作为 WZ 帧,取序列的第 21 帧和第 23 帧做 MCI 得到 SI 帧,从而得到噪声残差,把 CN 分割成 4×4 子块做 DCT 变换后,按 Zig-Zag 扫描顺序形成 16 个频带,任意取其中的直流(DC)系数频带和交流(AC)系数 AC1、AC9 两个频带,做出实际数据的直方图分布、单模 Laplace 分布、基于 EM 算法的二状态混合 Laplace 分布、基于式(7)的自适应窗宽 KDE 拟合曲线,如图 2、图 3 所示。

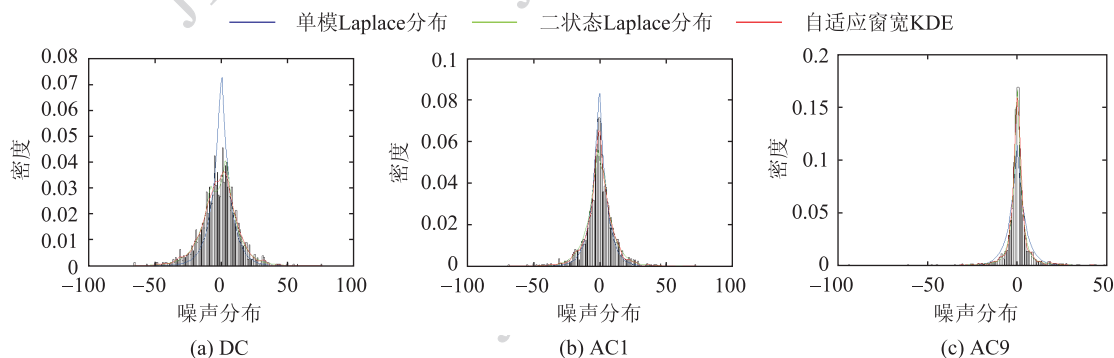


图 2 Forman 第 22 帧相应子带 CN 拟合情况

Fig. 2 Subband CN fitting for the 22nd frame of Forman

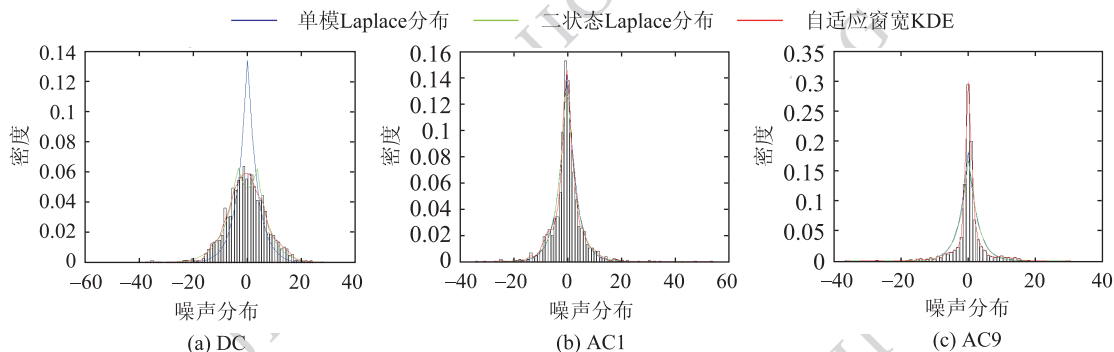


图 3 Akiyo 第 22 帧相应子带 CN 拟合情况

Fig. 3 Subband CN fitting for the 22nd frame of Akiyo

从图 2、图 3 可以看出, 由于低频带 CN 数据偏差较大、拖尾现象较严重, 二状态混合 Laplace 分布和基于式(7)的最优窗宽 KDE 在低频带的拟合精度明显高于单模 Laplace 分布曲线; 在高频带, CN 数据的分布较集中, 几种拟合方法在视觉上看不出太大的差距。

为了进一步比较拟合的准确性, 用 KL 距离 (Kullback-Leibler divergence)^[15] 定量地判断几种拟

合方法的准确度。

Forman 和 Akiyo 两个序列各频带用单模 Laplace 分布、二状态 Laplace 分布和自适应窗宽 KDE 几种拟合方法拟合, 与原始数据分布的 KL 距离如表 1 所示。从表 1 中的结果可以看出, 总的来说, 本文研究的基于 EM 算法的二状态 Laplace 分布和 KDE 算法的拟合准确度都优于单模 Laplace 分布, 并且绝大多数情况下, 用 KDE 法拟合 CN 的分布是最准确的。

表 1 Forman 和 Akiyo 序列 CN 几种拟合方法的 KL 距离比较
Table 1 CN fitting methods comparison of KL divergence for Forman and Akiyo

频带	Forman			Akiyo		
	单模 Laplace 分布	二状态 Laplace 分布	自适应窗宽 KDE	单模 Laplace 分布	二状态 Laplace 分布	自适应窗宽 KDE
DC	0.27	0.07	0.05	0.28	0.05	0.02
AC 1	0.13	0.08	0.04	0.11	0.06	0.03
AC 2	0.11	0.09	0.03	0.10	0.06	0.01
AC 3	0.12	0.06	0.02	0.15	0.21	0.03
AC 4	0.11	0.07	0.04	0.16	0.09	0.06
AC 5	0.11	0.08	0.03	0.12	0.07	0.01
AC 6	0.11	0.06	0.01	0.17	0.06	0.01
AC 7	0.12	0.11	0.02	0.26	0.21	0.06
AC 8	0.11	0.09	0.02	0.23	0.24	0.01
AC 9	0.22	0.06	0.05	0.22	0.21	0.01
AC10	0.14	0.02	0.05	0.29	0.04	0.06
AC11	0.12	0.08	0.02	0.30	0.12	0.10
AC12	0.15	0.06	0.04	0.20	0.05	0.02
AC13	0.13	0.13	0.01	0.33	0.10	0.11
AC14	0.12	0.09	0.04	0.37	0.08	0.09
AC15	0.17	0.04	0.01	0.33	0.05	0.37

2.3.2 算法复杂度分析

虽然基于 EM 算法的二状态 Laplace 分布和自适应窗宽 KDE 的拟合准确度都要优于单模 Laplace 分布, 但是它们都是以运算复杂度的提高为代价的。基于 EM 算法的二状态 Laplace 分布的运算复杂度主要和 EM 算法的迭代次数有关; KDE 实际上是对离散样本点在一个窗宽范围内插一个核函数, 由多个核函数加权得到光滑的分布曲线, 因此自适应窗宽的 KDE 的算法复杂度主要和样本点的多少有关。

以 Akiyo 序列的第 22 帧为例, 对相关噪声几个频带的 DCT 系数用基于 EM 算法的二状态 Laplace 分布和自适应窗宽的 KDE 拟合所需的运算代价如表 2 所示。表 2 列出了基于 EM 算法的二状态 Laplace 分布所需要的迭代次数和 KDE 算法需要的样本点数, 以及两种方法在 2.4 GHz 主频、2 GHz RAM 存储器的 PC 机上, 用 Matlab7.0 编译所需的时间。可见, 二状态 Laplace 分布的运算代价随着 EM 算法的迭代次数的增加而增加, KDE 方法的运算代价随着样本点数的增加而增加。

表2 两种拟合方法的复杂度比较

Table 2 Complexity comparison for the two fitting methods

频带	二状态 Laplace 分布		自适应窗宽 KDE	
	迭代次数	时间/s	样本点数	时间/s
DC	10	0.95	83	1.00
AC1	11	0.94	294	3.69
AC6	43	3.37	61	0.79
AC7	53	4.04	49	0.58

3 TDWZ 系统中的 CN 模型

文献[7]中提出的噪声模型通常被作为噪声模型研究的参照,它以式(1)的 Laplace 模型为依据,模型参数 α 采用在线估计的方法得到,即

$$\alpha(u, v) = \begin{cases} \sqrt{\frac{2}{\sigma_b^2}} & [D_{|b|}(u, v)]^2 \leq \sigma_{|b|}^2 \\ \sqrt{\frac{2}{[D_{|b|}(u, v)]^2}} & [D_{|b|}(u, v)]^2 > \sigma_{|b|}^2 \end{cases} \quad (8)$$

式中, $\sigma_{|b|}^2$ 表示 b 子带中 CN 绝对值的方差, $D_{|b|}(u, v)$ 表示 b 子带 (u, v) 处噪声的绝对值与该子带均值之差。

为了比较 CN 分布的参数估计方法与非参数估计方法对 WZ DVC 系统性能的影响,以及它们与文献[7]中所用的 CN 模型相比带来的视频编码器率失真性能的改善,采用文献[10]提出的基于 EM 算法的混合 Laplace-均匀分布模型 (MLUDM) 来描述 CN 的分布,即

$$p_{\text{MLUDM}}(T(u, v) | \theta) = \begin{cases} \pi_1 p_1(T(u, v) | \theta_1) + \pi_2 p_2(T(u, v) | \theta_2) & [D_{|b|}(u, v)]^2 \leq \sigma_{|b|}^2 \\ \frac{1}{Q(b)} & [D_{|b|}(u, v)]^2 > \sigma_{|b|}^2 \end{cases} \quad (9)$$

式中, $p_{\text{MLUDM}}(T(u, v) | \theta)$ 表示 MLUDM 分布模型, $Q(b)$ 表示 $b(b=0, 1, \dots, 15)$ 子带的量化等级数。

同时,为 CN 建立基于最优带宽的 KDEUDM,即

$$p_{\text{KDEUDM}}(x) = \begin{cases} \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right) & [D_{|b|}(u, v)]^2 \leq \sigma_{|b|}^2 \\ \frac{1}{Q(b)} & [D_{|b|}(u, v)]^2 > \sigma_{|b|}^2 \end{cases} \quad (10)$$

式中,核函数 K 选择高斯核函数、窗宽 h 按式(7)选取。

这样,在文献[7]的 CN 模型的基础上,式(9)的基于 EM 算法的 MLUDM 模型和式(10)的基于最优带宽的 KDEUDM,分别通过 EM 算法或自适应窗宽的 KDE 算法准确地模拟 CN 分布的尖峰和拖尾,通过均匀分布描述由 SI 估计引入的误差较大的子带分量的 CN 分布,这就为 CN 的 DCT 系数建立了全新的噪声分布模型。

4 实验结果和分析

选取标准 QCIF 格式的视频序列 Forman, Akiyo, Mobile, Coastguard, Carphone 和 Suzie 前 50 帧,用图 1 所示的 TDWZ 编码框架进行编码。比较 3 种 CN 模型,即文献[7]的 CN 模型、式(9)所示的基于 EM 算法的 MLUDM 模型以及式(10)提出的 KDEUDM 模型, WZ 帧编码的主观效果和客观效果。

实验采用图 4 所示的 4 个量化等级由低到高的量化矩阵,矩阵中的数字表示相应频带位置处 DCT 系数的量化等级,“0”表示此处 WZ 帧的系数直接由 SI 帧的系数代替,不需要进行 WZ 编码和传输纠错比特。每一 WZ 帧用图 4 所示的 4 个量化矩阵量化后,进行 WZVC 编码和解码,比较了在相同质量 SI 帧的前提下,采用不同 CN 分布模型时 DVC 系统的编码效率。重建 WZ 帧的主观视觉效果如图 5、图 6 所示,峰值信噪比 (PSNR) 和编码每一帧所需的比特数曲线如图 7 所示。

从图 5、图 6 中可以看出,图 5(b)(c)、图 6(b)(c) 采用改进的噪声模型后,重建的 WZ 帧图像质量都要高于图 5(a) 和图 6(a);虽然图 5(b)(c)、图 6(b)(c) 具有相近的视觉效果,但是从图 7(a) 的性能曲线可以看出,采用 KDEUDM 模型,编码每一帧所需的比特数要比采用 MLUDM 模型低近 10%。

32	16	8	0
16	8	0	0
8	0	0	0
0	0	0	0

(a) A

64	32	16	8
32	16	8	0
16	8	0	0
8	0	0	0

(b) B

128	32	32	16
32	32	16	8
32	16	8	0
16	8	0	0

(c) C

128	64	32	16
64	32	16	8
32	16	8	4
16	8	4	0

(d) D

图 4 量化矩阵

Fig. 4 Quantization matrix



PSNR=32.15 dB
(a) 文献[7]的CN模型



PSNR=32.34 dB
(b) MLUDM



PSNR=32.34 dB
(c) KDEUDM

图 5 Forman 第 22 帧 TDWZ 编解码结果(用量化矩阵 C)Fig. 5 TDWZ coding results for the 22nd frame of Forman (using quantization matrix C)

PSNR=32.25 dB
(a) 文献[7]的CN模型



PSNR=33.69 dB
(b) MLUDM



PSNR=33.69 dB
(c) KDEUDM

图 6 Forman 第 22 帧 TDWZ 编解码结果(用量化矩阵 D)Fig. 6 TDWZ coding results for the 22nd frame of Forman (using quantization matrix D)

其他 5 个序列 WZ 帧编码的性能曲线见图 7 (b)~(f)所示,从图 7 中可以看出,对 CN 模型改进后,高量化等级下编码的率失真性能增益最高能达到 1 dB 以上。其中在得到相同 WZ 帧图像质量的前提下,采用 KDEUDM 模型比 MLUDM 模型编码码率能节省 5%~10%。

对于两种噪声模型运算复杂度,测试了上述 6

个序列在 2.4 GHz 主频,2 GHz RAM 存储器的 PC 机上,用 Matlab7.0 建模所需的时间,结果如表 3 所示。可以看出在视频序列的运动量较小的情况下,由于边信息帧能较准确地预测,所以用于核密度估计的样本点数较少,建立 KDEUDM 模型所需的时间也较少,而其他大部分情况下建立 KDEUDM 模型的代价要高于 MLUDM 模型。

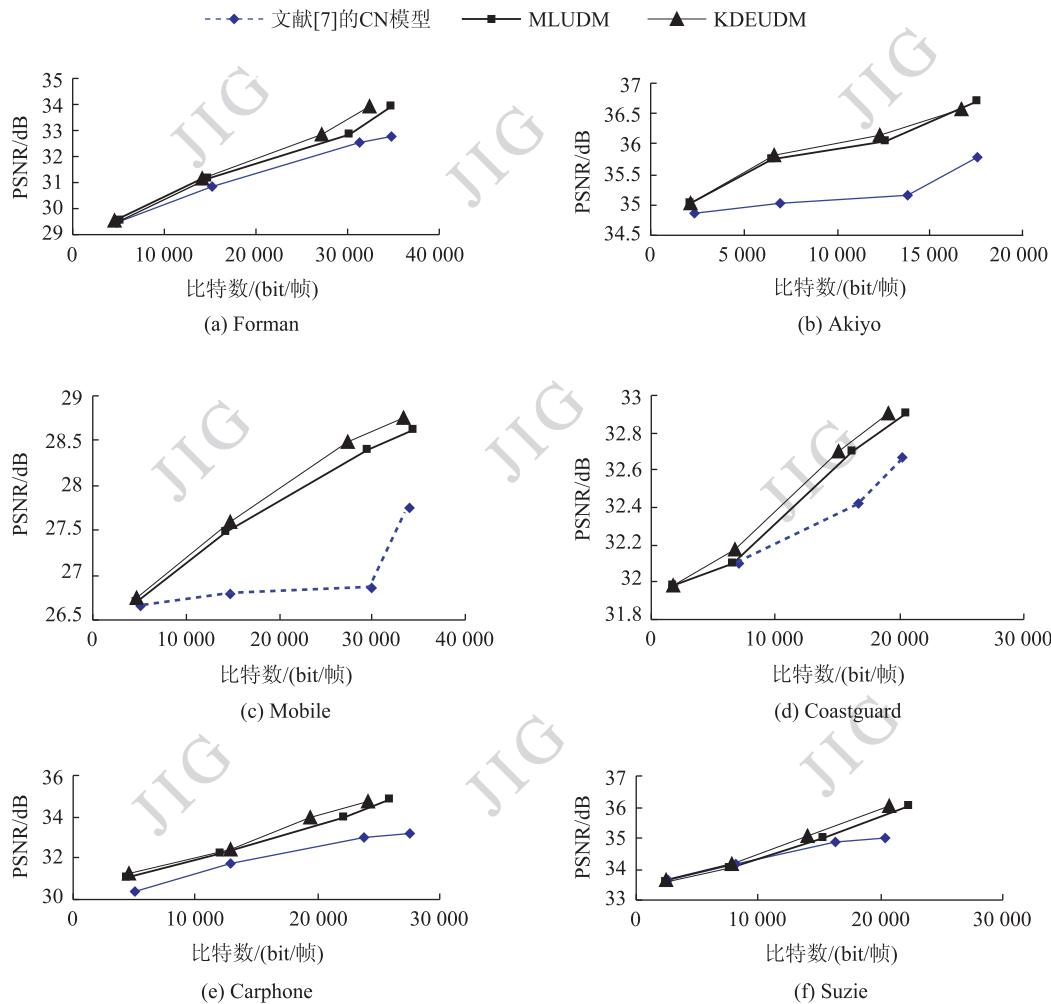


图7 测试结果
Fig. 7 Test results

表3 两种噪声模型的建模时间对比
Table 3 Modeling time comparison
of the two noise models

序列	MLUDM	KDEUDM
Forman	11.90	19.67
Akiyo	22.43	16.24
Mobile	15.21	25.87
Coastguard	16.73	18.14
Carphone	12.58	19.21
Suzie	18.33	12.24

码率和误码率,也影响着 WZ 帧重建的图像质量,是造成 WZ DVC 与传统的视频编解码系统性能差距的重要因素之一。本文讨论了 WZ DVC 系统中 CN 分布的参数估计和非参数估计方法,比较了 DVC 系统中采用 MLUDM 模型和 KDEUDM 模型对 WZ 帧编码性能的改善,并通过实验表明了得到相同图像质量的条件下,采用 KDEUDM 模型比 MLUDM 模型最高能节省 10% 的码率,非参数估计法是 TDWZ 系统中一种有效的噪声估计方法。

参考文献 (References)

[1] Girod B, Aaron A M, Rane S, et al. Distributed video coding [J]. Proceedings of the IEEE, 2005, 93(1): 71-83.
[2] Chiu C C, Chien S Y, Lee C H, et al. Distributed video coding: a promising solution for distributed wireless video sensors or not? [C]//Visual Communications and Image Processing. Piscat-

5 结 论

在 DVC 系统中,CN 分布模型联系着编码器和解码器,模型的准确度影响着 LDPC 编解码的编码

- away, NJ: IEEE, 2011: 1-4.
- [3] Kang L W, Lu C H. Distributed compressive video sensing [C] // Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE, 2009: 1169-1172.
- [4] Baig Y, Lai E M K, Punchihewa A. Distributed video coding based on compressed sensing [C] // Proceedings of IEEE International Conference on Multimedia and Expo Workshops. Piscataway, NJ: IEEE, 2012: 325-330.
- [5] Aaron A, Rui Z, Girod B. Wyner-Ziv coding of motion video [C] // Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: IEEE, 2002, 1: 240-244.
- [6] Luong H V, Huang X, Forchhammer S. Adaptive noise model for transform domain Wyner-Ziv video using clustering of DCT blocks [C] // Proceedings of the 13th IEEE International Workshop on Multimedia Signal Processing. Piscataway, NJ: IEEE, 2011: 1-6.
- [7] 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.
- [8] Thomas M, Jerome G, Beatrice P P, et al. Using an exponential power model for wyner ziv video coding [C] // Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing. Piscataway, NJ: IEEE, 2010: 2338-2341.
- [9] Yin M, Cai S T, Xie Y. Wyner-Ziv video coding based on Gaussian mixture model [J]. Chinese Journal of Computers, 2012, 35(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, 35(1): 173-182.]
- [10] Hu X F, Zhu X C. A Wyner-ziv video coding method utilizing mixture correlation noise model [J]. Journal of Electronics, 2012, 29(3/4): 197-203.
- [11] 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.
- [12] Zhang X G. Pattern Recognition [M]. 3rd ed. Beijing: Tsinghua University Press, 2010. [张学工. 模式识别 [M]. 3 版. 北京: 清华大学出版社, 2010.]
- [13] Silverman B W. Density Estimation for Statistics and Data Analysis [M]. London: Chapman & Hall, 1986.
- [14] Bors A G, Nasios N. Kernel bandwidth estimation for nonparametric modeling [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2009, 39(6): 1543-1555.
- [15] Kullback S, Leibler R A. On information and sufficiency [J]. The Annals of Mathematical Statistics, 1951, 22(1): 79-86.