

三网融合下基于压缩感知的 半参考视频质量评估方法研究^{*}

田玲¹, 廖紫纤², 何楚²

(1. 湖北邮电规划设计有限公司, 武汉 430023; 2. 武汉大学 电子信息学院, 武汉 430072)

摘要: 针对三网融合条件下的视频质量模型、评估体系和设计方案, 提出了一种基于压缩感知的半参考视频质量评估方法。该算法首先建立视频质量模型, 并通过压缩感知算法获取视频图像的感知系数; 然后利用视频图像的稀疏边缘分布去表征不同图像失真模型的相关性; 最后利用一个信息距离测度测量这些概率分布并进行量化。在标准测试图像库上的实验验证了该算法的有效性和对于不同失真模型的适应性, 特别是更适合于运营商在多屏、多制式、多终端、多分工界面等特点的三网融合复杂环境下的应用。

关键词: 三网融合; 视频质量评估; 半参考; 压缩感知

中图分类号: TP393.4 **文献标志码:** A **文章编号:** 1001-3695(2013)06-1892-03

doi:10.3969/j.issn.1001-3695.2013.06.077

Semi-reference video quality assessment method based on compressed sensing under triple-play

TIAN Ling¹, LIAO Zi-xian², HE Chu²

(1. Hubei Design Company of China Comservice, Wuhan 430023, China; 2. School of Electronic Information, Wuhan University, Wuhan 430072, China)

Abstract: Toward the video quality assessment problem in the field of triple-play, this paper proposed a semi-reference video quality assessment method based on compressed sensing. The algorithm first established the model of the video quality and obtained the perception coefficient of video images by compressed sensing. Then by using the sparse marginal distributions of video images, it characterized the correlation of different image distortion models. Finally, it used a distance measure of information to measure the probability distributions and quantify them. Experiments on a standard test image library confirmed the effectiveness of the algorithm and the adaptability for different distortion models. Moreover, it was more suitable for operators in a complicated environment application of multi-screen, multi-system, multi-terminal, multi-division interface in triple-play.

Key words: triple-play; video quality assessment; semi-reference; compressed sensing

0 引言

在国家政策、市场需求、技术发展等多种因素的推动下, 三网融合已经成为各大运营商下一步发展战略中的重要组成部分。三网融合是指电信网、有线电视网和互联网的融合, 其中基于高带宽能力并提供丰富灵活的多屏系统模式下的多媒体业务将是电信运营商提供差异化服务的核心竞争力之一^[1]。在多媒体业务当中, IPTV、网络视频、移动视频以及各类监控在内的视频业务成为最受关注 and 最具成长性的业务之一, 已成为全球大宽带的驱动业务。视频业务对于提升大宽带战略的整网价值、保护、促进和增值宽带以推动整网的盈利水平发挥了重要作用。

由于大数据量和实时性等特点, 视频业务是一个需要高带宽、低传输时延与低时延抖动的业务, 其服务质量极易受到目前具有分布式和开放环境的承载网络的稳定性的影响^[2,3]。视频业务在采集、处理、压缩、存储、传输和复制过程中都会受到各种因素的影响从而导致图像失真和质量下降。主要来源

可以分为三类: a) 视频采集中由于采集设备、环境等条件限制造成的质量问题, 或者从人感知角度觉得混乱或无法反映希望表达内容所体现的视频质量不佳; b) 在数字视频压缩中常用到离散余弦变换 (discrete cosine transform, DCT) 及其他视频变换, 并对所得系数进行量化, 尽管理论上 DCT 可以获得精确的变换和反变换, 在实际中编码器是无法获得完全精确的变换, 这些因素也会造成视频质量的下降; c) 视频在网络传输过程中造成的误码及丢码引起的视频失真与质量下降^[3]。因此, 提供视频质量检测、监控和保障体系, 为客户提供更为高效而准确的服务更具迫切性。

针对人类感知与客观参数描述的不同以及三网融合中多屏多参数视频图像等问题, 建立有效的视频质量评估模型, 研究客观质量评估指数体系和基于用户主观感知特点与情感倾向的视频质量评价方法, 对于开展面向三网融合的视频等多媒体业务的规划和咨询服务具有重要的基础意义, 也将在 IP 城域网等承载网设计、视频监控、可视电话等新业务领域发挥重要作用。同时国家三网融合政策的不确定造成的多种播放平台分工界面和片源的多样性, 也使得三网融合中视频质量

收稿日期: 2012-08-31; 修回日期: 2012-10-26 基金项目: 国家“863”信息通信重大资助项目

作者简介: 田玲, 数据设计经理, 硕士, 主要研究方向为数据业务网络; 廖紫纤, 硕士研究生; 何楚, 副教授, 博士 (chuhe@whu.edu.cn)。

评估方案需要考虑一些不确定性,具有更大的灵活性和一定的前瞻性。

基于上述背景,本文深入探讨了面向中国电信三网融合建设的视频质量评估方法,提出了一种基于压缩感知的半参考视频质量评估方法。a)针对中国电信三网融合示范网特点,对视频质量评估体系进行简单介绍并建立视频质量模型;b)将压缩感知应用于要评估的视频图像中,可以发现所得稀疏系数的边缘分布对于不同的图像失真类型呈现不同变化。因此,本文采用一个信息距离测度对这些概率分布进行测量来量化分布之间的不同。通过与主观图像质量评估方法进行对比,验证了该方法的有效性。由于并不局限于某一特定失真模型,该方法可以应用于广泛失真类型的图像质量评估。

1 视频质量评价体系

1.1 视频质量评估方法

用户对业务质量的主观评估 (quality of experience, QoE) [4] 主要分为两个方面:业务可靠性和业务舒适性。业务可靠性主要包含业务可用性与业务完整性;业务舒适性则包含业务请求延时与视频播放质量,如表1所示。其中,业务可用性、业务完整性以及业务请求延时的评估都相对容易获得,而视频播放质量的评估是目前的难点之一。视频质量评估方法通常分为主观评估方法和客观评估方法。

表1 业务质量评估方法

业务质量 评估方法	业务可靠性		业务完整性(业务异常中断率)	
	业务 舒适性	业务 请求时延	视频清晰度	视频播放质量
			视频损伤度(停顿、马赛克、画音不同步)	

主观评价方法让用户以打分等形式使主观体验以数据方法记录。ITU-R BT. 500-1 中特别制定了视频测量平均主观得分(mean opinion score, MOS)值的方法。客观评价方法包括全参考、无参考和半参考方法。

a)全参考方法结果较为准确。经典的均方误差(MSE)与峰值信噪比(PSNR)方法易计算,物理含义明确;基于小波的评估方法利用了小波分解的多分辨率信息;结构相似性(SSIM)方法基于结构信息提取,更能反映人类感知特点。但上述方法要求前提条件较为苛刻,在实用中往往难以大范围采纳。

b)半参考方法[5]通过提取图像频域或者小波域系数的特征,通过传递的原始图像少量信息对降质后图像进行评估,具有较大的应用价值。但文献[5]中的方法由于小波分析的系数较多,仅仅靠参数拟合以获得较大幅度的降维,难以获得准确结果。

c)无参考方法利用参考图像的部分信息对失真图像进行评价,在特征信息传输和质量评价准确性之间获得平衡[5],是当前三网融合建设值得重点考虑的质量评价方法。

1.2 视频质量评估设计方案

三网融合发展过程中诸多方面的成长性和不确定性,使得面向中国电信三网融合的视频质量评估设计方案必然要具有一定的前瞻性和灵活性。在前面论述的多种视频质量评价方法中,针对原始视频可以有完整视频信息、典型质量特征和无

参考等多种视频源端的采集方案。这也给设计方案中探测节点的布置方案带来了要求,可以针对不同的政策和网络现状采取不同的策略,从而进行不同的探测节点设置,并带给网络不同类型和数据量的质量评估所需信息的传输负载。建立起一套针对目前三网融合视频平台的多层次、多策略、有一定冗余的视频质量评估方案,可以对多种因素带来的视频质量下降进行可定位的实时监控与评估。

2 基于压缩感知的半参考视频质量评估方法

2.1 压缩感知

压缩感知是近年来数学和信息科学出现的理论研究和应用热点,在奈奎斯特采样定律等基础理论方面带来革命性突破。面向人们希望尽可能少地采集数据的需求,建立严格数学理论体系证明,通过随机信号等观测矩阵,可以以远低于奈奎斯特频率对具有一定稀疏性的信号进行采样,并能通过数值最优化问题准确重构出原始信号。

2.2 基于压缩感知的半参考视频质量评估方法

本文将压缩感知这一概念应用于视频质量的半参考评估方法,采用一个信息距离测度对发送端原始图像和接收端失真图像稀疏系数的概率分布进行测量来量化分布之间的不同。

用 $p(x)$ 和 $q(x)$ 分别表示两幅图像对应稀疏系数的概率密度函数, $X = \{x_1, x_2, \dots, x_N\}$ 表示一组随机独立选择的 N 个系数的集合。本文采用 KL 距离 (Kullback-Leibler distance, KLD) [6] 对两幅图像之间的差异进行测度:

$$d(p \parallel q) = \int p(x) \log \log \frac{p(x)}{q(x)} dx \quad (1)$$

假设 $p(x)$ 和 $q(x)$ 分别与原始参考图像和失真图像的稀疏系数概率密度函数相对应,为估计两幅图像之间的距离,必须同时获得原始图像和失真图像的系数直方图。失真图像的系数直方图可直接从接收端接收到的失真图像中获得,而在接收端获取参考图像的系数直方图则是难点。

自然图像统计的一个重要特征是系数的边缘分布可以与一个 2 参数的广义高斯密度模型 (generalized Gaussian density model, GGD) 拟合:

$$p_m(x) = \frac{\beta}{2\alpha(\frac{1}{\beta})} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} \quad (2)$$

其中, $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ 为 Gamma 函数。采用 GGD 模型后,仅需传送模型的两个参数 α 与 β 至接收端。与此同时,本文将 $p_m(x)$ 和 $p(x)$ 之间的 KLD,即预测误差,作为半参考特征的三个参数:

$$d(p_m \parallel p) = \sum_{i=1}^L p_m(i) \log \log \frac{p_m(i)}{p(i)} \quad (3)$$

其中: $p(i)$ 和 $p_m(i)$ 是直方图第 i 个单元对应的归一化值; L 表示直方图单元个数。

在接收端,由于失真后的图像不再满足自然图像的特征,图像与 GGD 无法拟合,因此在接收端直接将失真图像的稀疏系数作为特征,计算 $p_m(x)$ 和 $q(x)$ 之间的 KLD:

$$d(p_m \parallel q) = \int p_m(x) \log \log \frac{p_m(x)}{q(x)} dx \quad (4)$$

$p(x)$ 和 $q(x)$ 之间的 KLD 可以估计为

$$\hat{d}(p \parallel q) = d(p_m \parallel q) - d(p_m \parallel p) = \int p_m(x) \log \log \frac{p(x)}{q(x)} dx \quad (5)$$

即为原始参考图像与失真图像的失真度。

估计误差则为

$$d(p \parallel q) - \hat{d}(p \parallel q) = \int [p(x) - p_m(x)] \log \frac{p(x)}{g(x)} dx \quad (6)$$

图1给出了基于压缩感知的半参考视频质量评估方法的实现方式。在发送端,视频质量的特征提取包括三个步骤:a)将压缩感知应用于原始图像,得到一组稀疏系数;b)计算系数的直方图,利用GGD拟合获得特征参数 $\{\alpha, \beta, d(p_m \parallel q)\}$;c)对特征参数进行量化。 β 和 $d(p_m \parallel q)$ 均采用8 bit精度量化, α 采用11 bit浮点量化。在接收端视频质量的特征提取包括三个步骤:a)将压缩感知应用于失真图像,得到一组稀疏系数并获得对应系数直方图;b)根据直方图和RR特征参数 $\{\alpha, \beta\}$ 得到 $d(p_m \parallel q)$;c)由 $d(p_m \parallel q)$ 和第三个RR特征参数 $d(p_m \parallel p)$ 计算图像失真度 $\hat{d}(p \parallel q)$ 。

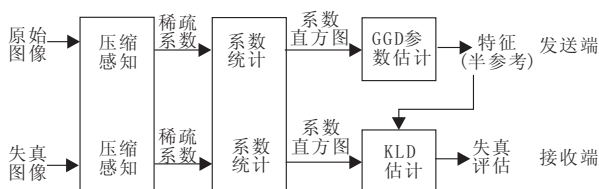


图1 基于压缩感知的半参考视频质量评估过程

3 实验

本文实验数据采用LIVE数据库^[7],该数据库包含的数据为对29幅高分辨率原始图像(大小为 768×512)进行五种类型的不同程度上的失真所获得的图像。这些失真图像被分为7组数据库,数据库1(87幅图像)和2(82幅图像)是JPEG2000压缩图像;数据库3(87幅图像)和4(88幅图像)是JPEG压缩图像;数据库5、6和7则是在JPEG2000比特流中采用快衰落瑞利信道模型分别进行高斯噪声、高斯模糊和传输错误失真。观察者在连续线性尺度上对图像质量进行感知,该线性尺度平均分为五个等级:bad, poor, fair, good以及excellent。每幅图像由20~25名观察者进行评估,每名观察者给出的原始分数被转换成Z-score并在每个数据集中进行重新调整以满足范围1~100。对每幅图像,计算平均意见得分(mean opinion score, MOS)和主观分数之间的标准方差。

为了定量描述本文所提出的质量评估方法,本文采用了三种评价指标。

a)根据视频质量专家组在第一阶段FR-TV测试中给出的建议,本文采用一个四参数的逻辑函数来提供客观和主观分数之间的非线性映射。经过非线性映射之后,计算预测得分和真实主观得分之间的相关系数,以评估预测准确率。

b)计算Spearman秩序相关系数以评估预测单调性。

c)采用离群率来评估预测一致性,其中,离群率定义为预测值在主观分数的 ± 2 标准方差范围以外的比例。

本文重复了文献[5]中的实验,该文献选择了一组主流质量评估算法作为对比实验,包括峰值信噪比(PSNR)、Lubin的Sarnoff模型^[8]、结构相似度指数(SSIM)^[9,10]以及基于小波的图像质量评估方法^[11]。其中Lubin的Sarnoff模型以及SSIM指数是最新广泛使用的全参考质量评估方法。

实验结果如表2所示,可以看出本文方法对于各种失真类型都获得了相当好的评估结果;同时,由于不包含复杂变换的

计算,使得该方法易于实现且计算效率高,特征参数也较少。

表2 实验结果

数据集	JP2(1)	JP2(2)	JPG(1)	JPG(2)	noise	blur	error
图像数	87	82	87	88	145	145	145
相关系数(预测准确率)							
本文方法	0.9418	0.9649	0.9314	0.9707	0.9492	0.9410	0.9435
文献[5]	0.9353	0.9490	0.8452	0.9695	0.8889	0.8872	0.9175
PSNR	0.9337	0.8948	0.9015	0.9136	0.9866	0.7742	0.8811
Sarnoff	0.9706	0.9650	0.9589	0.9837	0.9631	0.9480	0.9144
MSSIM	0.9675	0.9668	0.9646	0.9856	0.9706	0.9361	0.9439
小波方法 ^[11]	0.9393	0.9101	0.9091	0.9244	0.9045	0.9022	0.9133
秩序相关系数(预测单调性)							
本文方法	0.9340	0.9535	0.8830	0.8974	0.9098	0.9190	0.9256
文献[5]	0.9298	0.9470	0.8332	0.8908	0.8639	0.9145	0.9162
PSNR	0.9231	0.8816	0.8907	0.8077	0.9855	0.7729	0.8785
Sarnoff	0.9668	0.9565	0.9528	0.8904	0.9411	0.9381	0.9048
MSSIM	0.9566	0.9677	0.9572	0.9441	0.9718	0.9421	0.9497
小波方法 ^[11]	0.9291	0.9013	0.8012	0.8601	0.9012	0.9019	0.9024
离群率(预测一致性)							
本文方法	0.0397	0.0366	0.1042	0.0179	0.0852	0.1021	0.0602
文献[5]	0.0690	0.0366	0.1839	0.0341	0.1793	0.1172	0.0621
PSNR	0.0805	0.0976	0.0920	0.1818	0.0000	0.2069	0.1517
Sarnoff	0.0000	0.0366	0.0115	0.0000	0.0345	0.0276	0.0552
MSSIM	0.0000	0.0000	0.0000	0.0114	0.0000	0.0414	0.0345
小波方法 ^[11]	0.0701	0.0813	0.0601	0.3411	0.0134	0.0913	0.0931

4 结束语

针对目前的国内三网融合政策和网络现状,本文研究了视频质量模型与视频质量评估方法,视频质量评估方案应综合考虑有参考、无参考和半参考方法,并提出了一种基于压缩感知的半参考视频质量评估方法,在LIVE数据库上进行了一系列对比实验,实验结果验证了本文方法的有效性。

参考文献:

- [1] 中国电信广州研究院. 三网融合下的新业务[EB/OL]. (2010-10). <http://doc.mbalib.com/view/7851a92b1dd9345921eaa63d2530ccc.html>.
- [2] 姚良,奚溪. 三网融合下视频业务质量评估体系的研究[J]. 电信科学, 2011, 27(3): 27-33.
- [3] 范少芬. 基于网络损伤和流媒体特征的IPTV视频QoE评估方法的研究与实现[D]. 上海: 华东师范大学, 2010.
- [4] ITU-T P. 10/G. 100, vocabulary for performance and quality of service[S]. 2006.
- [5] WANG Zhou, SIMONCELLI E P. Reduced-reference image quality assessment using a wavelet-domain natural image statistic model[C]// Human Vision and Electronic Imaging X. 2005: 5666.
- [6] COVER T, THOMAS J. Elements of information theory[M]. New York: Wiley-Interscience, 1991.
- [7] SHEIKH H, WANG Zhou, BOVIK A, et al. Image and video quality assessment research at LIVE[EB/OL]. <http://live.ece.utexas.edu/research/quality/>.
- [8] LUBIN J. The use of psychophysical data and models in the analysis of display system performance[M]. Cambridge: The MIT Press, 1993: 163-178.
- [9] WANG Zhou, BOVIK A, SHEIKH H, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Trans on Image Processing, 2004, 13(4): 600-612.
- [10] WANG Zhou, BOVIK A, SHEIKH H, et al. The SSIM index for image quality assessment[EB/OL]. <http://www.cns.nyu.edu/~lcv/ssim/>.
- [11] 倪晓明. 基于小波变换图像质量评价新算法[D]. 厦门: 厦门大学, 2009.