

# 3DQoE 评价方法及其模型研究进展综述

刘延伟<sup>1</sup>, 刘金霞<sup>2</sup>, 慈松<sup>1</sup>, 要瑞霄<sup>1</sup>, 赵平华<sup>1</sup>, 王遵义<sup>2</sup>

(1. 中国科学院声学研究所高性能网络实验室, 北京 100190; 2. 浙江万里学院, 浙江宁波 315100)

**摘要:** 随着 3D 视频技术的逐步发展, 3D 视频网络服务的用户体验质量(3DQoE)评价逐步成为学术界和工业界都极为关注的热点问题之一. 本文通过分析端到端的 3DQoE 影响因素, 总结基于 3D 视频与图像质量的 3DQoE 评价方法, 综述 3DQoE 模型的研究进展. 通过 3DQoE 评价与模型研究现状的分析, 本文给出了目前 3DQoE 评价与模型研究过程中存在的精确性, 可靠性以及通用性问题, 并指出针对大规模网络服务的 3DQoE 建模是未来可能的研究方向.

**关键词:** 3DQoE 模型; 3D 视频/图像质量评价; 3D 视频服务

**中图分类号:** TP919.72      **文献标识码:** A      **文章编号:** 0372-2112 (2015)03-0568-09

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

**DOI:** 10.3969/j.issn.0372-2112.2015.03.023

## Overview of 3DQoE Assessment and Modeling

LIU Yan-wei<sup>1</sup>, LIU Jin-xia<sup>2</sup>, CI Song<sup>1</sup>, YAO Rui-xiao<sup>1</sup>, ZHAO Ping-hua<sup>1</sup>, WANG Zun-yi<sup>2</sup>

(1. High Performance Network Lab, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China;

2. Zhejiang Wanli University, Ningbo, Zhejiang 315100, China)

**Abstract:** Due to the rapid advancements in 3D video technologies, 3D quality of experience (3DQoE) assessment for 3D video networking service attracts the researchers' attention in both academic and industry world. The contribution factors of 3DQoE in the end-to-end processing chain are analyzed. The 3D video/image quality based 3DQoE evaluation methodologies and the current research progresses in 3DQoE modeling are reviewed. Based on the analysis of the current research status of 3DQoE assessment and modeling, we point out the inadequacies in accuracy, reliability and universality of the current 3DQoE assessment and modeling research, and indicate that 3DQoE model towards massive networking service is possibly the future research direction.

**Key words:** 3D quality of experience model; 3D video/image quality assessment; 3D video service

## 1 引言

近年来,随着 3D 内容制作与 3D 显示技术的发展, 3D 视频以 3D 电视的形式已经逐步走进了家庭影院, 提供给家庭用户超强的立体感和临场感体验. 由于高清晰度 3D 视频的震撼体验, 高清 3D 视频将成为下一代媒体的主要代表形式之一. 相比于单纯的高清平面视频, 3D 视频的采集制作、压缩传输以及显示技术等都较为复杂. 因此, 目前 3D 视频的内容种类不够丰富, 体验质量不够理想.

3D 视频服务进入家庭主要是通过蓝光 3D 光盘、宽带互联网络、广播、有线电视网络等多种传播通道. 其中, 基于宽带互联网络的 3D 视频传输是一种既经济又能兼容现有平面视频分发的传输方式. 传统的网络视频服务质量评价大多基于 QoS (Quality of Service) 指标, 主要从网络的层面出发来评价服务的质量, 但是并不能客

观的反映用户的最终体验质量. 为此, 研究人员又提出了体验质量(QoE, Quality of Experiences)的概念<sup>[1,2]</sup>. QoE 是一个以用户为核心并考虑影响体验质量各种因素的一个总体概念. 目前, 针对平面视频服务的 QoE 研究工作已经很多. 相比于平面视频, 3D 视频涉及到深度感知特性. 因此, 针对 3D 视频服务的 3D 视频体验质量(3DQoE, 3D Quality of Experience)评价更为复杂<sup>[3]</sup>.

由于 3D 视频服务需求的逐步加大, 3D 视频服务体验质量的评价工作就显得格外重要. 但是, 目前 3DQoE 大多采用观看打分的主观评价模式<sup>[4]</sup>. 这种评价模式需要专业的评价环境和巨大的人力耗费, 并且无法自动执行. 因此, 研究可以自动执行、并且能够为 3D 视频服务质量优化提供指导功能的 3DQoE 模型是非常必要的. 在最近的文献中, 已经出现了一些关于 3DQoE 模型的研究工作<sup>[5~8]</sup>, 取得了一定的进展, 但是还存在着一些问题值得深入的剖析. 基于这一点, 本文通过分析基

于 3D 视频或图像质量的 3DQoE 评价方法,然后对现有的一些 3DQoE 模型进行归纳总结,尝试给出 3DQoE 评价与建模在未来的研究趋势。

## 2 3DQoE 影响因素分析

由于 3D 视频涉及到多个视点,即使进行了有效的压缩,3D 视频的数据量依然很大,而网络带宽资源有限,导致网络化 3D 视频服务的体验质量不高.特别是针对无线传输的 3D 视频应用,随着移动用户的增多,移动宽带网络下的资源显得更加匮乏,并且用户的移动性较强,这导致视频传输过程中丢包严重,从而影响用户的观看体验质量<sup>[9]</sup>.除了网络环节对 3D 产品观看体验质量的影响,3D 内容制作也影响 3D 视频的观看体验质量<sup>[10]</sup>.目前,由于 3D 采集和制作技术的限制,3D 视频内容产品类型不够丰富;3D 内容的后期显示要求高,3D 内容与显示设备的匹配不够理想.这些都严重影响用户观看的体验质量。

3DQoE 的影响因素很多.内容采集、压缩、网络传输、终端显示、用户的心理、生理因素<sup>[11]</sup>以及视频内容的上下文关系等多方面因素相互作用,共同影响 3DQoE<sup>[3]</sup>.图 1 给出了 3D 视频处理链条上对 3DQoE 造成关键影响的不同因素的分析.在内容采集方面,摄像机同步、颜色与亮度校正、深度范围的控制、视差感知舒适度的控制等因素直接影响 3D 观看体验.在压缩方面,有损压缩的量化参数控制、容错编码参数控制以及视点之间码率的分配等也影响接收端的 3D 体验质量.在网络传输方面,不同协议层参数控制、丢包控制、物理层信道编码控制等也直接影响 3D 视频传输过程中的视频失真.在 3D 视频后期制作方面,基于深度和纹

理的虚拟视点绘制在绘制过程中导致虚拟视点视频产生部分失真<sup>[12,13]</sup>,也直接影响双目 3D 视频的渲染与显示质量.在 3D 显示方面,分辨率大小、视频宽高比、深度范围、舒适度范围等直接与用户的观看质量相关.在用户方面,用户的心理因素、生理因素、观看 3D 引起的疲劳、以及视频内容是否适合用户的喜好等因素也直接影响用户的体验质量。

3D 影视的观看体验不同于平面视频体验,观看疲劳以及观看舒适度是人类生理和心理对 3D 内容观看的直接反映<sup>[14]</sup>.3D 观看疲劳与不舒适感直接来自于 3D 内容与 3D 显示以及人类视觉的协调程度.3D 内容失真造成的 3D 视差混乱会导致观看眩晕等不适.目前的 3D 显示技术使得用户观看时眼睛的聚焦频繁变化,观众需要进行大量的视觉切换,眼睛传递给大脑的信息是观看者在运动,而肌肉传递的信号是静止的,长时间会导致信号错乱,引起疲劳.因此,改善 3D 内容制作与 3D 显示技术环节,控制 3D 视差的合适范围,可以有效地提高用户的 3D 体验质量。

由于 3D 内容采集与制作本身很难控制,而用户本身的各种人文因素 (human factors) 也无法控制.一旦 3D 内容制作完成,从服务提供商的角度看,为了提高用户的观看体验质量,无法在内容采集和用户本身方面再做工作来改善 3D 视频通信的服务质量.只能从信源压缩、传输网络调整、用户观看条件等方面采用一定的技术手段来提高用户体验质量.这样,提高用户体验质量的关键因素主要集中在内容压缩、网络传输、用户观看的环节上.而这些因素的影响相互交叉,如果互相之间不够协调,将严重影响用户的观看体验质量。

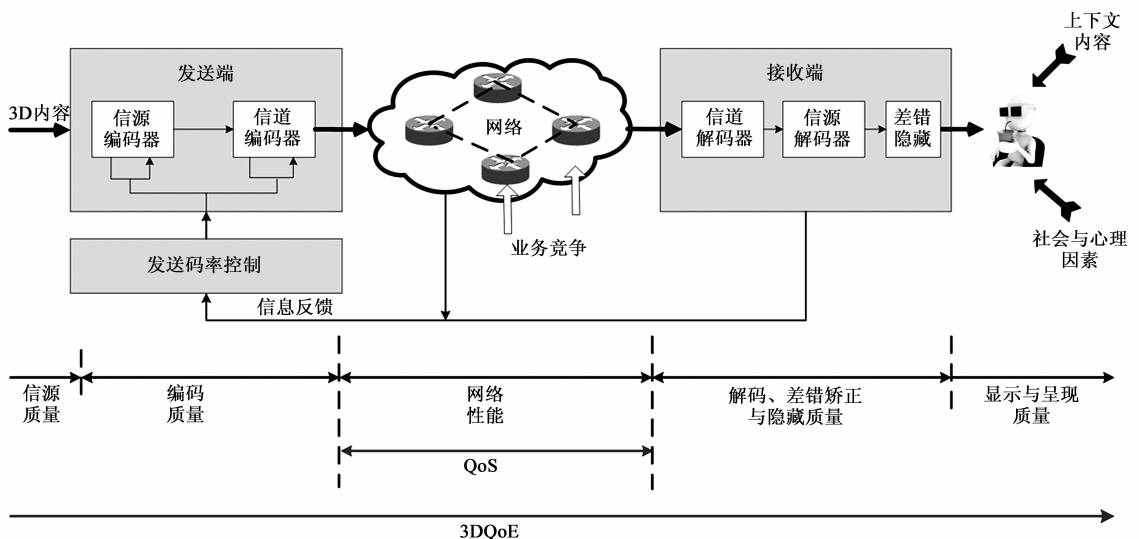


图1 3D处理链条上影响3D体验质量的因素分析

### 3 基于 3D 视频或图像质量的 3DQoE 评价方法

由于 3D 视频的视觉体验质量直接由观看得到,因此 3D 视频或图像的质量评价可以一定程度上反映 3D 视频的观看体验质量. 3D 视频或图像质量的评价通常不会考虑端到端处理过程中的影响因素,而只是对受损视频或图像进行质量评价. 但是在 3D 视频服务过程中对受损 3D 视频或图像的质量的评价可以在一定程度上间接反映 3D 视频服务体验的质量.

针对平面视频的质量评价主要分为主观评价和客观评价方法. 由于主观评价方法是以用户的总体视觉感受为目标,所以主观评价方法依然可以用来评价 3D 视频或图像. 主观视频质量评价是以平均意见评分 (Mean Opinion Score, MOS) 的方式进行量化评价的<sup>[15]</sup>. 由于主观评价方法耗费大量的人力和物力,测试时间较长,不适合应用在实时视频传输系统中. 基于这些考虑,研究人员提出了很多客观评价方法来试图替代主观评价在实时视频系统中的应用. 根据评价过程中用作参考的信息量多少,客观评价方法可以分为全参考信号评价方法 (full-reference quality evaluation), 部分参考信号评价方法 (reduced-reference quality evaluation) 和无参考信号评价方法 (no-reference quality evaluation). 具有代表性的全参考信号评价方法如峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 评价、结构相似度 (Structural Similarity, SSIM) 评价<sup>[16]</sup>和动态图像质量评估准则 (Moving Picture Quality Metric, MPQM)<sup>[17]</sup>. 部分参考信号评价方法通过利用部分参考信息对视频或图像质量进行评价<sup>[18]</sup>. 相比于全参考信号或部分参考信号评价方法,无参考信号评价方法主要是通过估计视频失真来评价视频质量<sup>[19]</sup>. 客观评价方法通过对视频质量进行客观量化,利用客观指标进行评价,避免了测试结果受测试人员个体因素的影响,也降低了测试成本. 客观评价方法对平面视频质量评价起到了一定的效果,但是由于 3D 视频涉及到一些新的特性,这些方法并不完全适用于 3D 视频应用<sup>[20]</sup>.

针对 3D 视频的特性,研究人员提出了新的针对 3D 视频或 3D 图像质量的评价方法<sup>[21~26]</sup>. 这些方法在 2D 视频或图像评价方法的基础上又考虑了 3D 视频质量中涉及到的新元素,例如深度感知质量以及双目图像融合感知质量等. 根据评价中是否有原始参考信号,3D 视频或图像的评价依然可以归为三类方法,即全参考信号评价方法、部分参考信号评价方法以及无参考信号评价方法.

#### 3.1 全参考 3D 视频或图像质量评价

全参考信号 3D 视频或图像质量评价方法利用原始 3D 视频或图像所有维度的信号作为参考,对受损信

号的质量进行对比评价. 人类视觉系统主要从两个维度去感知 3D 视频或图像质量,即左右视点图像的视网膜重叠图像 (cyclopean image) 视觉感知和深度感知. 全参考信号 3D 视频或图像质量评价流程如图 2 所示<sup>[27]</sup>. Cyclopean 图像可以直接由式 (1) 得到,即

$$CI(x, y) = W_L(x, y) \times I_L(x, y) + W_R((x + d), y) \times I_R((x + d), y), \quad (1)$$

其中,CI 代表左右视点图像的视网膜重叠 cyclopean 图像,  $I_L$  表示左视点图像,  $I_R$  表示右视点图像,  $d$  代表左右视点图像对应像素之间的视差,  $W_L(x, y)$  和  $W_R(x, y)$  分别表示描述双眼竞争 (binocular rivalry) 的加权系数,其可以通过归一化的 Gabor 滤波响应的幅度计算得到,

$$W_L(x, y) = \frac{GE_L(x, y)}{GE_L(x, y) + GE_R((x + d), y)} \quad (2)$$

$$W_R(x + d, y) = \frac{GE_R((x + d), y)}{GE_L(x, y) + GE_R((x + d), y)} \quad (3)$$

其中,  $GE_L$  和  $GE_R$  可以通过提取图像的亮度和色度特征进行滤波计算来得到<sup>[28]</sup>.

除了 cyclopean 图像的视觉感知,视差图是影响深度感知的直接因素. 在评价过程中需要构建视差图或深度图来评价受损 3D 视频或图像的深度感知质量. 视差图像可以通过立体匹配算法进行估计得到. 最后,根据原始的双目 3D 图像参考信号和受损的待评价双目 3D 图像信号分别得到相应的 cyclopean 图像和视差图,然后评价视网膜上的 cyclopean 图像质量 (包含基于视差线索的深度感知). 全参考信号 3D 视频或图像质量评价方法利用了原始参考信号,评价结果和主观评价结果具有较高的斯皮尔曼 (Spearman) 相关性系数,并且视差图的精度对评价结果的精度有一定影响,但是影响并不大<sup>[27]</sup>.

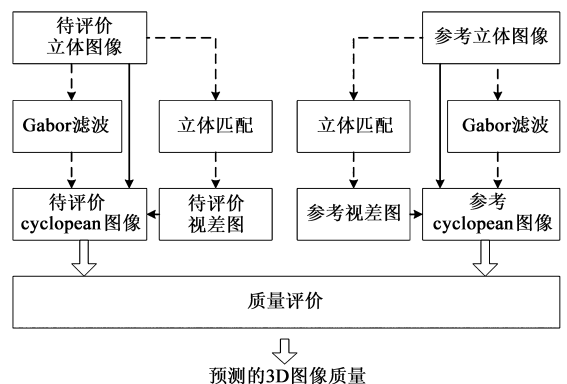


图2 基于全参考信号的3D视频或图像质量评价方法

#### 3.2 部分参考 3D 视频或图像质量评价

基于全参考信号的评价方法需要原始图像作为评价的参考. 在实际应用中,视频或图像的接收端很难得到原始图像进行参考,因此该方法的应用受到一定的

限制.在研究过程中,一些学者又提出了基于部分参考信号的 3D 视频或图像评价方法.所谓部分参考信号的评价就是指利用一些原始图像的关键特征信息作为参考,对受损视频或图像的关键特征信息进行评价,进而达到评价整幅图像的目的.图 3 给出了一种基于 edge 信息的针对深度加彩色纹理视频格式的 3D 视频评价方法<sup>[29,30]</sup>.作为部分被参考的特征信息对部分参考评价方法效果的影响至为关键.文献[30]利用 edge 特征图作为特征信息.首先,在视频发送端,通过 Sobel 滤波提取原始视点彩色纹理图像和深度图的 edge 结构特征信息,其中深度图可以通过立体匹配算法得到.在视频传输过程中,压缩的纹理视频和深度以及它们的原始 edge 结构图一并传输到接收端.在接收端,用同样的方法对受损的纹理视频图像和深度图像进行处理得到相应的 edge 结构特征图.然后对受损信号生成的 edge 结构特征图和原始的 edge 结构特征图分别利用平均结构相似性指标进行评价(Mean SSIM, MSSIM),进而达到对彩色纹理视频和深度进行评价的目的.该方法和全参考评价方法相比,尽管具有一定的优势,但是该方法并没有考虑到彩色纹理视频和深度在人类视频感知系统中的融合过程,缺少对 3D 认知过程的深入刻画.

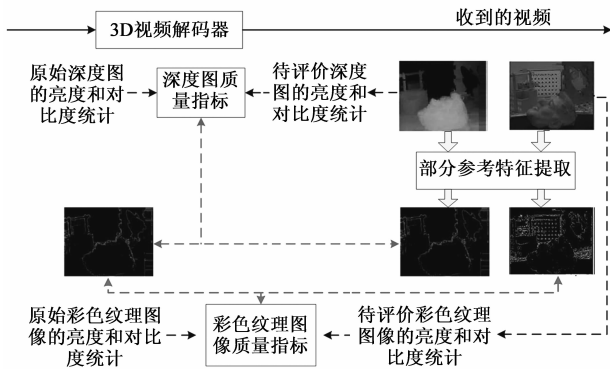


图3 基于部分参考信号的3D视频或图像评价方法

### 3.3 无参考 3D 视频或图像质量评价

为了进一步使得 3D 视频客观评价方法更符合实际应用条件,研究人员最近提出了针对 3D 视频或图像的无参考信号评价方法<sup>[31~33]</sup>.图 4 给出了针对自然采集的双目视频的无参考信号评价方法的结构框图<sup>[33]</sup>.与全参考或部分参考信号评价方法不同,无参考信号的评价方法直接提取受损图像的特征信息,然后根据特征信息的质量对图像进行评价.这种方法依然利用 cyclopean 图像视觉感知和基于视差的深度感知来评价 3D 视频或图像的 3D 感知质量.通过 Gabor 滤波构造 cyclopean 图像以及相应的二维特征,即 cyclopean 图像的频域统计分布特性;通过立体匹配算法构造视差图并提取视差图的三维特征信息,也就是视差图的频域

统计分布特性和视差估计的不稳定特性.得到这些二维和三维的特征信息后,基于失真类型识别评价(Distortion Identification-based Image Verity and Integrity Evaluation, DIIVINE)<sup>[34]</sup>指数来评价 3D 视频或图像的整体质量.基于无参考信号的 3D 视频或图像评价方法更符合实际应用环境,但是其评价过程需要基于失真类型匹配的无参考评价模式数据库,特别是评价结果的精度严重依赖于这个数据库.如果数据库的模式分类能够覆盖实际应用环境下的各种失真类型,则评价将会取得较好的结果.

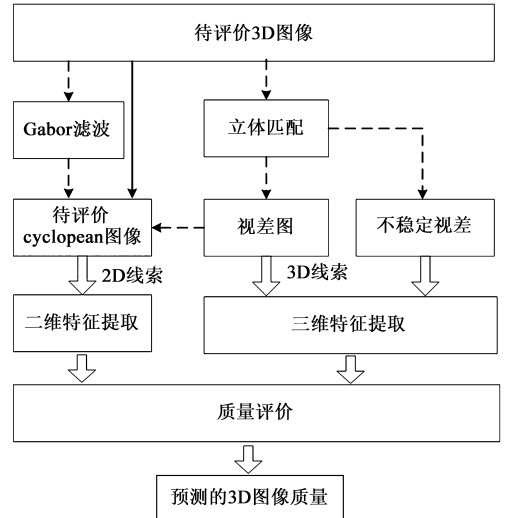


图4 基于无参考信号的3D视频或图像质量评价方法

目前,基于 3D 视频或图像质量的 3DQoE 客观评价方法的效率都需要主观评价进行验证.因此构建标准的具有多种损伤模式的 3D 视频或图像主观评价数据库(ground truth data)就显得尤为重要.特别是面向 3DQoE 的 3D 视频或图像评价,损伤模式更是代表了 3D 视频服务的不同影响因素.丰富的损伤模式以及相应的主观测试结果,可以为客观测试方法提供强有力的效率分析和验证.目前,典型的 3D 视频和图像测试数据库有视频质量评价专家组(VQEG) 3DTV 小组的针对 3DQoE 主观测试用的 3D 视频数据集<sup>[4]</sup>,洛桑联邦理工学院的具有不同摄像机间距的 3D 视频/图像数据库<sup>[35]</sup>,和坦佩雷大学构建的移动 3D 视频数据库<sup>[36]</sup>.这些视频数据库分别提供了不同的损伤特征,但是还缺少全面的损伤模式,特别是缺少针对虚拟视点方面的 3D 视频损伤数据.因此基于目前这些数据库进行验证的客观评价方法针对部分损伤特征有效,但还有待更全面损伤特征的有效性验证.

## 4 3DQoE 模型的研究进展

在 3D 视频通信中,3DQoE 是 3D 视频服务质量最

贴切的评价指标.3DQoE 与 3D 视频或图像质量评价既有相似之处,又存在着区别.相似之处在于 3D 视频或图像的质量间接反映了 3DQoE<sup>[37]</sup>,但是 3DQoE 又增加了新的元素:(1)3D 观看疲劳<sup>[38]</sup>、3D 深度范围以及舒适度问题;(2) 3DQoE 是和端到端视频处理相关联的,也就是和 3D 视频服务是息息相关的.3DQoE 主要是从用户的角度来评价服务质量的指标.而 3D 视频或图像评价仅仅是对受损视频或图像的评价.3D 视频或图像质量客观评价间接反映 3DQoE 质量,为 3DQoE 客观评价模型提供了一定的技术手段.

近年来,由于网络通信服务理念的发展,网络服务质量的评价由 QoS 延伸到 QoE.针对 3D 视频服务,除了 2D 视频服务的一些常规影响因素,3DQoE 更多的涵盖了 3D 视觉的一些特性.3D 视觉特性在端到端的网络服务过程中如何受到影响,直接体验在用户的观看体验质量上.根据这种理解,研究人员提出了一些 3DQoE 模型的研究工作,希望在没有评价人员干预的情况下,自动的评价客户端接收到的 3D 视频服务的质量.从目前的研究来看,针对 3DQoE 的预测模型主要是模拟人类视觉知觉的形成过程,并把这些过程和网络服务联系起来,通过网络服务影响因素和人类 3D 体验感知之间的映射关系来预测 3DQoE.

从 3D 体验感知类型的组合形式看,主要有两类 3DQoE 建模方法.一类是基于多种感知线性组合的 3DQoE 模型,包括图像质量、深度、观看舒适度以及串扰等不同类型感知的线性组合.另一类是基于非线性多种类型感知组合的 3DQoE 模型,主要包括基于智能学习等人工智能手段刻画多种感知复杂组合关系的模型.

#### 4.1 基于多种类型感知线性组合的 3DQoE 模型

从目前对人类 3D 视觉感知的理解来看,主要包括三种类型的感知,即双目图像形成的 cyclopean 图像质量的感知、基于深度信息的第三维度感知以及观看舒适度的体验感知.在 3D 视频服务中,这三类感知都会受到端到端不同因素的影响而形成失真.这些失真组合在一起,共同影响观看者接收到的 3DQoE.从 3DQoE 影响因素造成的直接后果看,这些因素导致双目图像信息产生失真而间接影响 3DQoE 中的三类感知的质量.这三类感知都来自于待评价双目图像提供的信息,如何通过分类拆解这些信息并将这些信息和端到端的影响因素关联起来,然后将这些信息感知和产生信息失真的因素有机的组合,这一过程就是 3DQoE 的建模过程.依据这三类感知的表示,3DQoE 模型又可以归纳为两种类型.一种是基于正向表示双目会聚 cyclopean 图像感知效果<sup>[5,39]</sup>的建模方法.还有一种是基于反向表

示双目会聚 cyclopean 图像感知效果<sup>[40]</sup>的建模方法,即基于串扰失真评价的建模方法.

(1)基于正向表示双目图像会聚质量的 3DQoE 建模方法

基于正向表示双目图像会聚质量的 3DQoE 建模方法主要是如何刻画三类不同类型感知和 3DQoE 之间的关系.图 5 给出了三种感知在 3DQoE 模型中的层次关系.其中双目会聚得到的 cyclopean 图像和形成的深度感知直接影响到 3D 感知过程中 3D 图像观看的自然特性,而自然感知效果也会一定程度的影响观看舒适度.将这三种感知线性组合,得到总体的 3DQoE 模型如下,

$$EC = w_1 \cdot IQ + w_2 \cdot D + w_3 \cdot VC \quad (4)$$

其中 EC(Evaluation Conception)表示用户体验评价的高层次概念,也就是指用户的 3DQoE, IQ 代表双目图像会聚产生的 cyclopean 图像质量, D 代表深度感知质量,而 VC 表示观看的舒适度感知质量,  $w_1$ 、 $w_2$ 、 $w_3$  分别代表三种感知在 3DQoE 形成过程中相应的权重.这些权重可以根据不同实验条件的配置进行 3D 视频主观评价实验而得到.在文献[5]中,  $w_1$ 、 $w_2$ 、 $w_3$  分别取值为 0.34、0.27、0.40.从式(4)中权重的取值可以看出,与其他两种感知相比,观看舒适度感知占据了总体感受的五分之二,这也说明 3D 视频体验与 2D 视频体验不同,舒适不疲劳的观看感觉是影响用户体验的关键因素之一.三种类型的感知和端到端影响因素的关联关系与传统的 2D 视频 QoE 和影响因素的关系基本相同.如图 1 所示,各种因素影响双目图像质量,但是对于具体哪些因素分别影响哪一种类型感知以及他们相互之间交互作用对总体 3DQoE 的影响,目前的研究还缺少相应的认识,因此这种 3DQoE 模型的预测精度受到一定的限制.特别是观看舒适度受到很多方面的影响,与观看对象的视觉心理等人文因素有很强的关联性,与其他两种类型感知也相互作用.而目前基于线性模型来刻画三种感知的相互关系,只能是对 3DQoE 进行一种近似的预测.

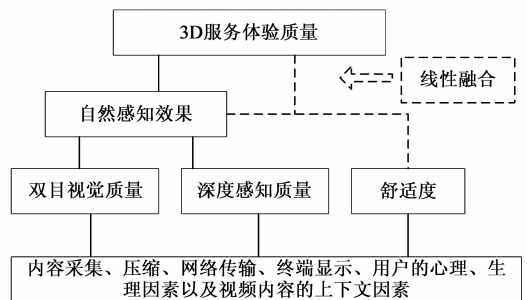


图5 三种感知在3DQoE模型中的层次关系

(2)基于串扰失真评价的 3DQoE 建模方法

双目图像在视网膜融合过程中,会产生一定的串

扰(crosstalk)效果<sup>[41]</sup>.而串扰失真从反向角度体现了双目图像会聚的质量,也就是串扰感知的质量.串扰感知可以通过描述不同影响因素造成的感知质量的变化而计算得到.具体的串扰感知影响因素总结于表 1<sup>[41]</sup>.在基于串扰失真评价的 3DQoE 模型中,深度感知与舒适度感知融合在一起,共同刻画 3D 观看的舒适度.

表 1 基于串扰失真的 3DQoE 感知类型与及相应的影响因素

| 感知类型       |                         | 相关影响因素  |
|------------|-------------------------|---|
| 串扰感知       | 串扰与原始视点信号的差别程度          | 串扰水平,场景内容对比度,串扰水平和场景内容对比度的相互作用                                |
|            | 串扰和原始视点信号之间的间隔距离        | 摄像机间距,场景内容深度,串扰水平和场景内容深度之间的相互作用                               |
|            | 串扰与原始视点信号差别和间隔距离之间的相互作用 | 串扰水平和摄像机间距之间的相互作用,场景内容深度和对比度同时作用,串扰水平、摄像机间距与场景内容深度和对比度之间的交互作用 |
|            | 空间位置                    | 可视串扰区域内的场景内容深度  |
| 深度关联的舒适度感知 | 视差                      | 摄像机间距,场景内容深度,屏幕分辨率,以及他们之间的相互作用                                |
|            | 舒适区域                    | 屏幕的像素宽度、观看位置  |

在基于串扰失真评价的 3DQoE 描述中,串扰感知质量刻画为

$$V_{cp} = \text{AVG}(L_s \times (1 - R_{pdis}/255)) \quad (5)$$

其中 AVG 表示平均操作,而“ $\times$ ”表示矩阵的元素相乘操作. $L_s$  表示系统引入的串扰信号  $L_c$  和没有串扰信号  $L_o$  之间的结构相似性评价图,即

$$L_s = \text{SSIM}(L_o, L_c) \quad (6)$$

在式(5)中, $R_{pdis}$  表示经过门限值滤波的视差图,而  $R_{dis}(i, j)$  表示基于平方误差和加上最小滤波(Sum of Squared Difference plus Min Filter, SSDMF)<sup>[32]</sup>得到的视差信息,其可以按如下计算得到.

$$R_{pdis}(i, j) = \begin{cases} R_{dis}(i, j), & \text{if } L_s(i, j) < 0.977 \\ 0, & \text{if } L_s(i, j) \geq 0.977 \end{cases} \quad (7)$$

$$R_{dis} = \text{SSDMF}(R_o, L_o) \quad (8)$$

在式(8)中, $L_o$  和  $R_o$  分别表示原始的左右视点图像信号,而左右视点图像的串扰信号可以根据串扰水平  $P$  计算如下:

$$\begin{cases} R_c = R_o + P \times L_o \\ L_c = L_o + P \times R_o \end{cases} \quad (9)$$

融合深度感知的舒适度感知刻画为

$$V_{devc} = \frac{\sum R_{dis}(i, j)}{D_{max} \times \text{Res}}, \text{if } R_{dis}(i, j) \leq D_{max} \quad (10)$$

其中 Res 表示视频的分辨率,而  $D_{max}$  表示为观看舒适区域内的最大的视差.

在得到串扰感知质量评价和融合深度感知的舒适度感知评价后,通过线性融合得到总体的 3DQoE,

$$V_{3DQoE} = (1 - \alpha) \times V_{cp} + \alpha \times V_{devc} \quad (11)$$

其中, $\alpha$  是一个平衡两种类型感知的权重因子,其通过主观评测实验可以取值为 0.84(这一参数取决于不同的实验设置).串扰感知从会聚的反向角度刻画了双目图像的会聚图像质量,但是由于该模型依然采用的是多种感知的线性组合,因此其预测精度依然局限在实验室环境,无法满足实际应用环境的要求.特别是线性关系的权重取值缺少动态的适配特性,在实际应用中很难精确描述 3DQoE 影响因素的动态变化特性.而且,串扰感知的刻画与 3D 显示设备的性能和观看条件息息相关,对显示设备依赖性较大,很难精确刻画网络服务因素对 3DQoE 的影响.

#### 4.2 基于非线性多种类型感知组合的 3DQoE 模型

由于人类的视觉感知系统非常复杂,目前还不能确定多种感知之间的关系是线性关系,特别是多种感知之间可能也存在着一定的相互作用,共同影响 3DQoE.因此,最近的研究工作提出了一些利用人工智能算法模拟人类视觉系统认知 3DQoE 过程的算法.这种模拟人类视觉系统认知过程的方法利用仿生智能技术,例如认知神经网络<sup>[42]</sup>以及机器学习<sup>[43]</sup>等来构建 3DQoE 模型.具体的基于神经网络(Random Neural Network, RNN)的 3DQoE 模型<sup>[8]</sup>,参考图 6.这种建模方法囊括了 3DQoE 的多种贡献因素,诸多的贡献因素构成了一个特征向量.基于 3D 感知分类的分析,并依据已有的特征向量构造预定可控的 3D 视频失真模式对智能模型进行训练,例如神经网络模型.训练后的模型被赋予相应的模型参数并用于预测实际的 3DQoE.在实际的应用中通过采集影响因素的样本,利用训练好的模型可以预测实际的 3DQoE.具体的 3DQoE 影响因素包括影响 3D 视差的一些因素,例如分辨率大小、网络传输丢包、压缩的码率等.基于人工智能的 3DQoE 模型能动态的刻画不同影响因素之间的相互作用以及他们对 3DQoE 的协同影响,能够取得很好的预测效果,更适合于实际的动态变化应用场景.但是这种 3DQoE 的预测精度一定程度上依赖于人工智能的训练过程和训练集合中失真模式的丰富程度<sup>[8]</sup>.如果有足够丰富的 3D 视频失真模式用于训练,则得到的 3DQoE 会取得很好的预测结果.

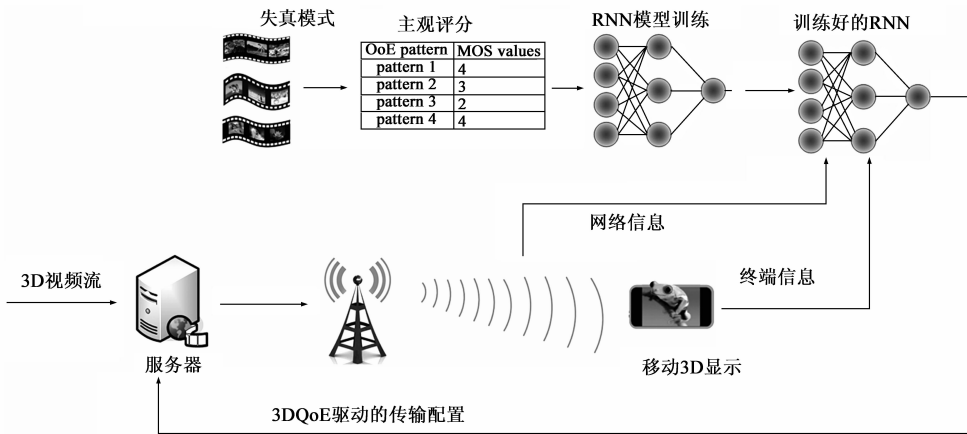


图6 面向3D视频传输的3DQoE建模方法

## 5 结论与展望

从目前关于3DQoE评价与模型的研究现状可见,基于3D视频或图像质量的3DQoE评价方法取得了一定的进展,3DQoE建模研究也获得了一定突破,关于3D感知类型的认识比较清楚,但是目前的3DQoE评价与模型研究在以下几个方面还存在着很大的提升空间。

(1)3DQoE的刻画精度与模型的通用性问题.目前的研究基本是基于主观实验统计分析,同时结合认知模型进行3DQoE的建模.由于3D体验影响因素的多样复杂的特性,想要3DQoE模型覆盖所有的影响因素是不现实的,因此目前的3DQoE模型的精度还不足以满足实际应用的需求.特别是目前的研究大多针对特定可控的实验序列进行研究,其涵盖的应用因素并不全面,得到的3DQoE模型的应用范围也不是适用于所有的3D视频服务.因此,针对通用、高精度的3DQoE模型研究工作还需要投入更多的精力。

(2)目前3DQoE评价与模型的研究工作存在着可靠性问题<sup>[44]</sup>.这一问题来源于目前缺少统一标准的受损3D视频数据库<sup>[3]</sup>,以及受控的3DQoE多种类因素干扰下的模式数据库.构建标准的3DQoE模式数据库可以有效的评价3DQoE的精度和适用范围以及明确的优势和劣势,对研究人员改进和推动3DQoE模型精度的提高具有重要的意义。

(3)目前3DQoE评价与模型的研究还局限在小规模的3D视频服务,和真正的3D视频应用系统结合并不紧密.随着未来网络的兴起,大数据背景下的3D视频服务将会占据主要的网络业务市场<sup>[45]</sup>.而大规模服务的数据体量和相应的3DQoE影响因素将会更为复杂,因此针对大规模3D视频网络服务的3DQoE评价与模型研究,对未来3D视频通信产业的发展将具有非常现实的意义。

## 参考文献

- [1] ITU-T Rec. G. 100/P. 10 Amendment 1-2007. New Appendix I -definition of Quality of Experience (QoE)[S].
- [2] 林闯,胡杰,孔祥震.用户体验质量的模型与评价方法综述[J].计算机学报,2012,35(1):1-15.  
Lin Chuang, Hu Jie, Kong Xiang Zhen, Survey on models and evaluation of quality of experience[J]. Chinese Journal of Computers, 2012, 35(1): 1-15. (in Chinese)
- [3] Hewage C, Martini M. Quality of experience for 3D video streaming[J]. IEEE Communications Magazine, 2013, 51(5): 101-107.
- [4] VQEG. Establishing a Ground Truth for Quality of Experience in 3D for Assessment Methodologies in 3D Video Quality Assessment[EB/OL]. <http://www.its.bldrdoc.gov/vqeg/projects/3dtv/3dtv.aspx>, 2014-5-2.
- [5] Chen W, Fournier J, Barkowsky M, Callet P. Quality of experience model for 3DTV[A]. SPIE Stereoscopic Displays and Applications XXIII[C]. USA: SPIE, 2012. 1-9.
- [6] Xing L, You J, Ebrahimi T, Perks A. An objective metric for assessing quality of experience on stereoscopic images[A]. International Workshop on Multimedia Signal Processing[C]. USA: IEEE, 2010. 373-378.
- [7] Chen W, Fournier J, Barkowsky M, Le Callet P. Exploration of quality of experience of stereoscopic images: binocular depth[A]. Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics-VPQM[C]. USA: IEEE, 2012. 1-6.
- [8] Liu Y, Ci S, Tang H, Ye Y, Liu J. QoE-oriented 3D Video transcoding for mobile streaming[J]. ACM Transactions on Multimedia Computing, Communications and Applications, 2012, 8(3s): article 42.
- [9] Liu Y, Ci S, Tang H, Ye Y. Application-adapted mobile 3D video coding and streaming-a survey[J]. 3D Research Journal, 2012, 3(1): article 3.

- [10] Smolic A, Kauff P, Knorr S, Hornung A, Kunter M, Muller M, Lang M. Three-dimensional video postproduction and processing[J]. *Proceeding of IEEE*, 2011, 99(4): 607 – 625.
- [11] Tam W, Stelmach L, Coriveau P. Psychovisual aspects of viewing stereoscopic video sequences[A]. *Proceeding of SPIE* [C]. USA: IEEE, 1998. 226 – 235.
- [12] 喻莉, 张军涛, 邓慧萍, 等. SSIM 度量虚拟视点绘制失真的深度图帧内编码[J]. *中国图象图形学报*, 2012, 17(11): 1353 – 1358.  
Yu L, Zhang J, Deng H, et al. Virtual view rendering distortion using the structural similarity index as quality metric for depth map intra coding[J]. *Journal of Image and Graphics* 2012, 17(11): 1353 – 1358. (in Chinese)
- [13] Liu Yan-wei, Huang Qing-ming, Ma Si-wei, et al. Joint video/depth rate allocation for 3D video coding based on view synthesis distortion model[J]. *Signal Processing: Image Communication*, 2009, 24(8): 666 – 681.
- [14] Urvoy M, Barkowsky M, Le Callet P. How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors[J]. *Annales des Télécommunications*, 2013, 68(11 – 12): 641 – 655.
- [15] ITU-R Recommendation BT. 500-11. Methodology for the Subjective Assessment of the Quality of Television Pictures [EB/OL]. <http://www.dii.unisi.it/~menegaz/DoctoralSchool2004/papers/ITUR-BT.500-11.pdf>, 2002.
- [16] Wang Z, Bovik A, Sheikh H, Simoncelli E. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600 – 612.
- [17] Branden Lambrecht C. Color moving pictures quality metric [A]. *Proceedings of the IEEE International Conference on Image Processing* [C]. USA: IEEE, 1996. 885 – 888.
- [18] 王新岱, 杨付正, 常义林. 基于数字水印的无参考视频质量评估方法[J]. *西安电子科技大学学报*, 2004, 31(2): 229 – 233.  
Wang X, Yang F, Chang Y. A no-reference video quality assessment method based on digital watermark[J]. *Journal of Xidian University*, 2004, 31(2): 229 – 233. (in Chinese)
- [19] Naccari M, Tagliasacchi M, Tubaro S. No-reference video quality monitoring for H. 264/AVC coded video[J]. *IEEE Transactions on Multimedia*, 2009, 11(5): 932 – 946.
- [20] Quan H, Le Callet P, Barkowsky M. Video quality assessment: From 2D to 3D-challenges and future trends[A]. *Proceeding of IEEE International Conference on Image Processing* [C]. USA: IEEE, 2010. 4025 – 4028.
- [21] Hewage C, Worrall S, Dogan S, Kondoz A. Quality evaluation of colour plus depth map based stereoscopic video[J]. *IEEE Journal of Selected Topics in Signal Processing: Visual Media Quality Assessment*, 2009, 3(2): 304 – 318.
- [22] Tam W, Speranza F, Yano S, Shimono K, Ono H. Stereoscopic 3D-TV: visual comfort [J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 335 – 346.
- [23] You J, Xing L, Perkis A, Wang X. Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis [A]. *Proceeding of International Workshop Video Processing and Quality Metrics for Consumer Electronics* [C]. USA: IEEE, 2010. 1 – 6.
- [24] Hewage C, Worrall S, Dogan S, Kondoz A. Prediction of stereoscopic video quality using objective quality models of 2-D video[J]. *Electron Lett*, 2008, 44(16): 963 – 965.
- [25] Bensalma R, Larabi M. A perceptual metric for stereoscopic image quality assessment based on the binocular energy[J]. *Multidimensional System Signal Processing*, 2012, 24(2): 281 – 316.
- [26] 蒋刚毅, 黄大江, 王旭, 等. 图像质量评价方法研究进展 [J]. *电子与信息学报*, 2010, 32(1): 219 – 226.  
Jiang G Y, Huang D J, Wang X, et al. Overview on image quality assessment methods[J]. *Journal of Electronics & Information Technology*, 2010, 32(1): 219 – 226. (in Chinese)
- [27] Chen M, Su D, Kwon C, Cormack L, Bovik A. Full-reference quality assessment of stereopairs accounting for rivalry [A]. *Forty-sixth Annual Asilomar Conference on Signals, Systems, and Computers* [C]. USA: IEEE, 2012. 1 – 5.
- [28] Field D J. Relations between the statistics of natural images and the response properties of cortical-cells[J]. *Journal of the Optical Society of America*, 1987, 4(12): 2379 – 2394.
- [29] Hewage C, Martini M. Reduced-reference quality metric for 3D depth map transmission[J]. *Proceeding of Capture, Transmission and Display 3D Video* [C]. Finland: IEEE, 2010. 1 – 4.
- [30] Hewage C, Martini M. Edge-based reduced-reference quality metric for 3-D video compression and transmission[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2012, 6(5): 471 – 482.
- [31] Akhter R, Baltes J, Parvez Sazzad Z, Horita Y. No reference stereoscopic image quality assessment [A]. *Proceeding of SPIE* [C]. USA: SPIE, 2010. 75240T.
- [32] Solh M, AlRegib G. A no-reference quality measure for dibr-based 3D videos [A]. *Proceeding of International Conference on Multimedia and Expo* [C]. USA: IEEE, 2011. 1 – 6.
- [33] Chen M, Cormack L, Bovik A. No-reference quality assessment of natural stereopairs [J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3379 – 3391.
- [34] Moorthy A, Bovik A. Blind image quality assessment: From natural scene statistics to perceptual quality [J]. *IEEE Transactions on Image Processing*, 2011, 20(12): 3350 – 3364.
- [35] EPFL-MMSPG. 3D Video Quality Assessment [EB/OL]. <http://mmspg.epfl.ch/3dvqa>, 2013-1-2/ 2014-7-30.

- [36] MOBILE 3DTV Project. Stereo-video Database [EB/OL]. <http://sp.cs.tut.fi/mobile3dtv/stereo-video/>, 2010-1-2/2014-7-30.
- [37] Lebreton P, Raake A, Barkowsky M, et al. Perceptual preference of S3D over 2D for HDTV in dependence of video quality and depth [A]. IEEE 11th IVMSWP Workshop [C]. USA: IEEE, 2013. 1 – 4.
- [38] Kim D, Sohn K. Visual fatigue prediction for stereoscopic image [J]. IEEE Transactions on Circuits System and Video Technology, 2011, 21(2): 231 – 236.
- [39] Wang K, et al. Perceived 3D TV transmission quality assessment; multilaboratory results using absolute category rating on quality of experience scale [J]. IEEE Transactions on Broadcasting, 2012, 58(4): 544 – 556.
- [40] Xing L, You J, Ebrahimi T, Perkis A. Assessment of stereoscopic crosstalk perception [J]. IEEE Transactions on Multimedia, 2012, 14(2): 326 – 337.
- [41] Xing L. Towards Reliable Stereoscopic 3D Quality Evaluation-Subjective Assessment and Objective Metrics [D]. Norway: Norwegian University of Science of Technology, 2013.
- [42] Mohamed S, Rubino G. A study of real-time packet video quality using random neural networks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(12): 1071 – 1083.
- [43] Menkovski V, Exarchakos G, Liotta A. Machine learning approach for quality of experience aware networks [A]. Proceeding of 2nd International Conference on Intelligent Networking and Collaborative Systems [C]. USA: IEEE, 2010. 461 – 466.
- [44] Goldmann L, Ebrahimi T. Towards reliable and reproducible 3D video quality assessment [A]. International Conference on Three-Dimensional Imaging, Visualization, and Display 2011 [C]. USA: SPIE, 2011. 1 – 6.
- [45] Kulik I, András K, Anh T, Bokor L. Analysis of the relationship between quality of experience and service attributes for 3D future Internet multimedia [A]. 4th IEEE International Conference on Cognitive Info Communications [C]. USA: IEEE, 2013. 641 – 646.

#### 作者简介



刘延伟 男, 1976 年出生于黑龙江省, 博士, 中国科学院声学研究所助理研究员, 主要研究方向为视频处理与通信、无线通信。

E-mail: liuyw@hpl.ac.cn



刘金霞 (通信作者) 女, 1969 年出生于黑龙江省, 现为浙江万里学院副教授, 主要研究方向为立体视频编码与通信。

E-mail: liujinxia1969@126.com