

基于生理电的移动网页时延对体验质量的影响

谭浩¹, 魏旭一¹, 孙家豪¹, 王斌²

(1.湖南大学, 长沙 410082; 2.上海华为技术有限公司, 上海 201206)

摘要: **目的** 从用户角度出发, 研究移动设备网页浏览过程中网页时延与用户体验质量之间的关系。**方法** 通过记录多类网页浏览过程中不同时延下用户皮电、心电两类生理指标的变化, 利用方差分析以及 t 检验计算其特征值发生显著性差异时对应的时延长度并对该时延长度进行标记, 结合 MOS 评分量表, 对用户的主观感受作出评价。**结果** 实验得到了较为可靠的结论, 研究发现不同移动网页时延下用户的体验质量是不同的。用户在打开网页时的主观体验与网页的时延长度整体存在非线性负相关关系。**结论** 当移动网页缓冲时延小于 1 s 时, 用户体验良好; 当时延超过 1 s 时, 用户体验质量在一定时延范围内随时延增加而下降且该变化呈现较大的个人差异性; 当时延超过 4 s 时, 用户负面体验较为严重。

关键词: 用户体验; 响应时延; 生理电测量; MOS 量表; 移动网页体验

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2018)14-0077-05

DOI: 10.19554/j.cnki.1001-3563.2018.14.015

Impact of Response Latency on User QoE in Mobile Web Browsing

TAN Hao¹, WEI Xu-yi¹, SUN Jia-hao¹, WANG Bin²

(1.Hunan University, Changsha 410082, China; 2.Huawei Technologies Co., LTD., Shanghai 201206, China)

ABSTRACT: From the perspective of users, the relationship between web page delay and user experience quality during mobile web browsing is studied. By recording the changes of two kinds of physiological indexes of the user's skin electricity and ECG during the multi-type web browsing process, the time delay length corresponding to the significant difference of its eigenvalues and the length of the time delay are marked by variance analysis and t test, and the subjective feelings of the users are evaluated by the MOS rating scale. The experimental results show that the quality of user experience is different under different mobile web pages. There is a nonlinear negative correlation between the subjective experience of users when opening web pages and the delay length of web pages. When the time delay of the mobile web page is less than 1 s, the user experience is good. When the time delay is more than 1 s, the user experience quality decreases at any time in a certain time delay, and the change presents a larger individual difference. At that time, the negative experience of the user is more serious than that of 4 s.

KEY WORDS: user experience; response latency; electrophysiological measurement; MOS scale; mobile web browsing experience

移动通信技术的发展以及智能终端的普及使得越来越多的服务转移到移动互联网络, 并使得用户对系统服务质量的要求不断提升, 如何衡量并保证用户移动设备上网浏览体验质量成为了网络服务商关

注的问题。目前, 国内对移动互联网体验质量的研究主要集中在视频和音频业务, 对网页浏览业务的研究较少。而在网页浏览业务中, 网页时延是影响用户体验的关键因素, 探究其与用户体验质量之间的定量关

收稿日期: 2018-03-12

基金项目: 国家自然科学基金项目 (61402159, 60903090); 国家科技支撑项目 (2015BAH22F00); 国家国际科技合作专项资金资助 (2012DFG70310); 华为研究院《基于人体感知的业务 MOS 分研究》项目

作者简介: 谭浩 (1977—), 男, 四川人, 博士, 湖南大学教授, 主要从事为人机交互、人机工程、交通工具造型设计与设计学基础理论研究等。

系对优化服务质量、提高用户体验具有重要意义。

1 响应时延与用户体验质量

用户在与服务或应用交互的过程中产生的对服务的主观感受，称为体验质量^[1]。目前采用最为广泛的 QoE 度量标准是基于性能指标的服务质量评价体系^[2]。然而，仅仅从技术层面对 QoE 进行评价往往会忽视用户的主观因素，而不能很好地反映用户的满意度和认可率^[3]，因此，需要关注用户的心理体验，在用户主观感受和技术指标之间建立映射关系，以帮助网络服务商在成本控制情况下优化用户体验。

在 QoS 的多项技术指标中，响应延迟是最常见的影响网页浏览体验的因素^[4]。国外相关领域的研究开始较早，Robert Miller 在 1968 年最先提出系统响应时延这一概念^[5]。在此之后，对 SRT 的研究主要集中在探究能够保证用户良好体验的时延长度以及影响时延长度感知的因素两方面。目前影响最广泛的理想响应时延标准分别由 Shneiderman 和 Seow 等在 1987 年和 2008 年提出^[6-7]。Shneiderman 将系统理想响应时延按照任务复杂度进行划分，而 Seow 则按照用户期望对响应时延建立了标准，Shneiderman 的理想响应时延标准见表 1，Seow 的理想响应时延标准见表 2。

表 1 Shneiderman 的理想响应时延标准
Tab.1 Shneiderman's standard system response time

任务类型	理想响应时延
打字及鼠标移动	50-150 ms
简单任务	1 s
一般任务	2-4 s
复杂任务	8-12 s

表 2 Seow 的理想响应时延标准
Tab.2 Seow's standard system response time

用户期望反馈类型	理想响应时延
即时	100~200 ms
迅速	0.5~1 s
连续	2~5 s
沉浸	7~10 s

2 用户体验质量度量方法

2.1 情绪与生理信号测量

用户体验体现在情绪之中。这里对用户体验质量的衡量主要基于情绪与生理信号测量方法。19 世纪 80 年代提出的情绪外周理论认为情绪刺激引起身体的生理反映^[8]。生理反映体现在生理指标的变化之中，这些变化由自主神经系统和内分泌系统支配，不

受人的主观控制，因此，利用生理信号可以真实、客观地评价用户情绪状态的变化，给用户体验度量提供了新的思路和方向。例如，Picard 等通过心电、皮电、肌电和呼吸作用探究了驾驶情境下司机的压力状况^[9]；Ward 等通过皮肤电反应和心率测量用户对同一网站不同设计版本的反应，发现设计不合理的网页会加剧用户的紧张反应^[10]。

情绪的维度与极性是情绪的固有属性，在同一情绪维度上存在不同的两极。当人处于平静状态时，认为其处于情绪的中常状态。在这里，实验素材均为负向情感诱发素材，即用户的情绪变化仅限于情绪零感知状态到严重负向情绪这一单一极向变化范围内，研究中最佳用户体验对应心理学领域中的情绪中常零感知状态，情绪状态研究范围见图 1。

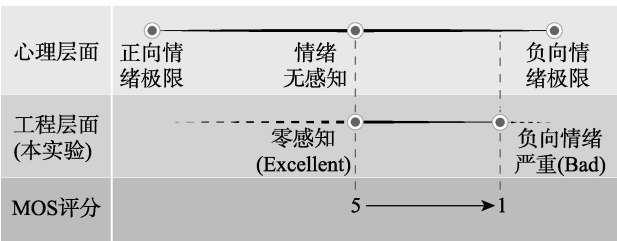


图 1 情绪状态研究范围
Fig.1 The interested range of emotional state in this study

在人机交互领域，皮电和心电及相关指标由于其对心理负荷的高敏感度且表现出相对稳定的结果成为用户体验研究中用来表征情绪变化的常用指标^[11]，也是这里研究用户浏览网页感受所采用的生理指标。

2.2 MOS 分值量表

在用户体验研究中，评分量表是最直接、常用的获得用户主观感受的方法^[12]。就 QoE 评价来说，国际电信联盟建议的“平均评估分值”是使用最广泛的评分量表，MOS 分值见表 3，它将用户的主观感受分为 5 个层次并对每个层次进行了定义。以生理信号分析为主，辅以行为观察和情绪主观报告识别多个被试的情绪，发现结果优于单纯利用生理信号的情绪识别结果^[13]，因此，利用 MOS 分值量表作为生理数据的合理补充，能够有效提高本实验结论的可靠性。

表 3 MOS 分值
Tab.3 MOS scale

MOS 分值	QoE	损害程度
5	优	不能察觉
4	良	可察觉但不严重
3	中	轻微
2	次	严重
1	劣	非常严重

3 实验方法

3.1 实验准备与被试

该实验在实验室环境中进行, 保证光线柔和、室温适宜。测试设备为华为 Mate7 手机, 并采用 MindWare 多导生理信号记录仪全程记录被试皮电与心电数据。

实验被试需身心健康, 且拥有至少 1 台移动设备并具备多次使用移动设备浏览网页的经验。最终实验被试选择了 30 名湖南大学师生, 年龄在 18 岁至 35 岁之间, 男女性各 15 名。

3.2 实验素材

本实验时延范围参考 Sheiderman 和 Seow 的响应时延标准, 结合目前网络服务质量水平, 测试网页时延范围为 0 s 至 7 s。在该时间范围内, 着重前 2 s 中的用户浏览体验变化, 因此, 在 0 s 至 2 s 范围内, 梯度设置为 0.2 s, 在 2 s 至 7 s 范围内, 设置梯度为 0.5 s。在测试网页选择方面, 为了避免单一网页测试可能造成的实验结果对某一类网页的强依赖性, 根据用户使用网页的目的将网页分为 4 类, 分别为搜索、购物、资讯和娱乐。每一时延长度测试素材均包括 4 类不同的网页, 在某一次网页任务中, 时延会随机出现在某一次页面跳转的过程中。根据真实网页模型, 将网页时延分为两部分, 分别为网页响应时延和网页等待时延。其中, 网页响应时延是指用户发出指令请求后, 请求页面首次出现提示信息的时间; 网页等待时延是指从页面出现首次提示信息到页面内容加载完成的时延。响应时延与等待时延的长度随机, 但两者之和需要为该次测试时延的长度。

3.3 实验过程

实验开始前, 被试需要了解实验流程及注意事项, 并穿戴传感器和电极。在这一过程中, 主试播放一段轻缓的音乐, 当被试放松至安静状态后, 实验人员开始记录被试的生理信号基线。之后被试开始实验任务, 被试任务流程见图 2。针对每一个时延, 用户需要在 4 个不同的网页上完成任务, 分别为: 百度——进入首页, 搜索指定名词, 并进入 1 个二级页面查看内容; 淘宝——进入首页, 搜索指定商品并将选定商品加入购物车; 腾讯新闻——打开腾讯新闻, 查看最新两则新闻; 优酷——进入首页, 搜索指定视频, 并打开直至视频开始播放。具有同样时延的 4 个网页为一组。每完成一组任务, 被试应用 MOS 量表对体验按照 1 分至 5 分打分。打分结束后, 被试休息 30 s, 待恢复平静状态后开始进行下一组任务, 直至 21 组任务全部完成。在此过程中, 持续记录被试的生理数据, 当用户开始或结束 1 次分任务时, 主试在其生理信号的对应时间点作出标记, 以便后期数据处理。

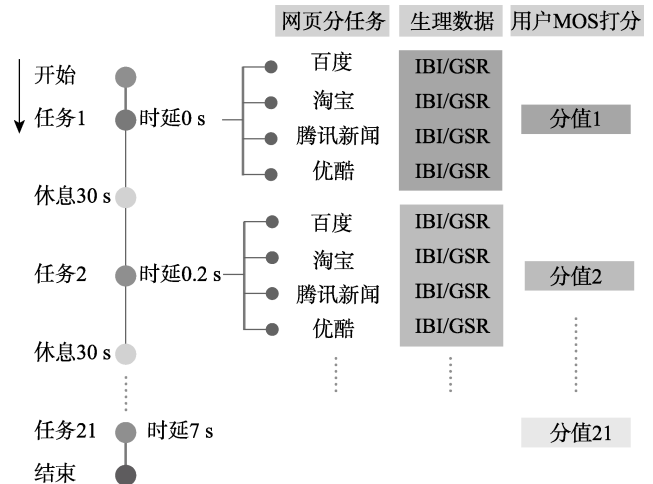


图 2 被试任务流程
Fig.2 Subjects' task flow

4 数据处理及实验结果

4.1 生理信号处理

为了探究随着网页时延增加被试情绪的变化以及对时延的忍耐限度, 需要根据生理数据找出每位被试情绪发生可察觉变化时对应的时延长度, 该类可察觉的情绪变化定义为在不同的时延长度下其皮电和心电信号的特征值与前一状态的特征值相比均发生了显著性差异。生理数据的处理步骤如下。

4.1.1 生理信号预处理

生理仪得到原始心电及皮电数据存在一定程度的噪声干扰, 因此, 首先需要对其进行预处理, 对皮电信号的处理过程主要包括滤波、去除基线和无用信息以及平滑处理。对心电信号的处理主要包括去噪和波形检测。预处理后的生理信号图像示例见图 3。

4.1.2 计算生理信号特征值

实验共得到有效数据 29 组, 各生理参数均符合正态分布。针对每个时延, 分别求取对应的皮电、心电信号特征值, 以判断其在时延增加过程中的变化情况。

选取皮电信号的均值作检验显著性差异的特征值, 其计算公式为: $gsr_{mean} = \frac{1}{N} \sum_{i=1}^N G(i)$, 其中 N 代表被试完成某次网页分任务采集的有效皮电数据点数, 是 GSR 样本值。对心电信号, 选取 IBI 相邻差绝对值的均值作为特征值, 其计算公式为 $\delta(IBM) = \frac{1}{N-1} \sum_{i=1}^{N-1} |IBM_{i+1} - IBM_i|$, 其中 N 代表在同一时延下被试完成某次任务时的心电数据点数, IBM_i 是根据心电信号得到的 IBI 样本值, 它表示两个 R 波之

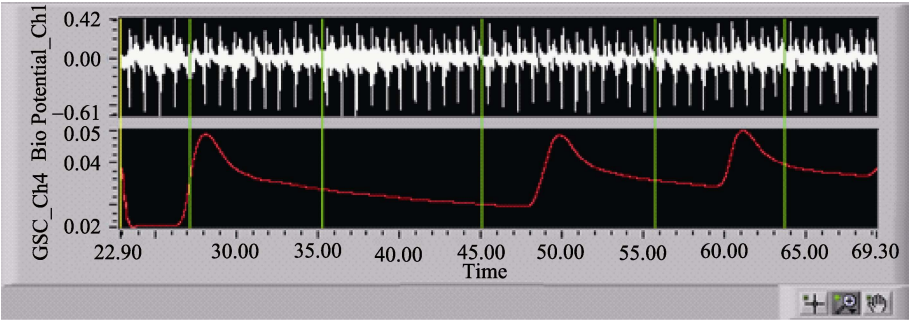


图 3 预处理后的生理信号图像示例
Fig.3 Sample graph of processed physiological signal

间的间期，即 t_{R-R} 。

由于每个时延长度下每位被试分别进行了 4 次网页任务，因此，对该 4 次任务过程中皮电、心电数据分别计算特征值，就得到某一时延长度下被试皮电、心电的特征值样本各一组。实验共得到皮电、心电特征值样本各 21 组。

4.1.3 对特征值进行显著性差异检验

在得到皮电与心电生理数据特征值之后，需要分别对它们进行显著性差异检验，生理信号特征显著性检验流程见图 4。

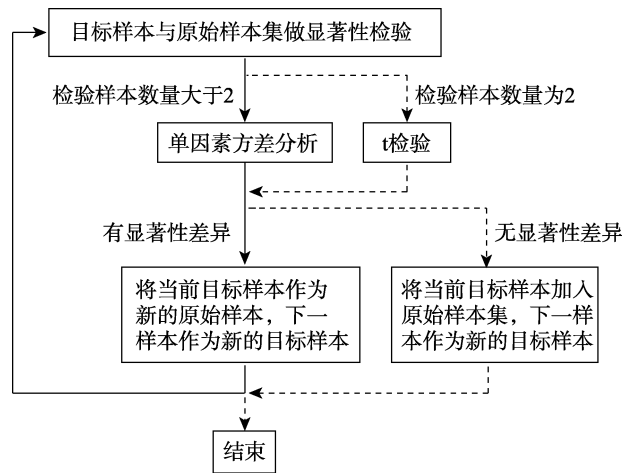


图 4 生理信号特征显著性检验流程
Fig.4 Process of significant difference test

以皮电数据为例，以 21 组特征值样本为材料进行显著性检验。针对某一个被试的皮电数据，从第一组任务（时延 0 s）得到的特征值开始，将其作为目标样本与原始样本集进行显著性检验，其中第一次的原始样本为被试在平静状态下采集到的皮电基准值，给定显著性水平 $\alpha = 0.05$ 。当某次检验结果显示差异显著时，对该时延值进行标记，且该次检验的目标样本成为新的原始样本；如果某次检验中样本之间的差异性不显著，则保留该次检验的所有样本，并将下一组任务数据作为新的目标样本，加入到当前的显著性差异计算中，直到 21 组特征值都参与过一次显著性差异检验。在该过程中，当某次检验的样本数量为 2

时，使用 T 检验；当样本数量大于 2 时，使用单因素方差分析来进行显著性检验。同理可进行心电数据的显著性差异计算。

4.1.4 寻找情绪变化可察点

针对同一被试，将皮电与心电信号中发生显著性差异的点进行一一比对，如果在某一时延长度下，皮电与心电信号均发生了显著性差异，则认为在该时延长度下被试情绪积累发生了可察觉变化，标记该时延长度为该被试的情绪变化可察点。

结果发现，对每位被试而言，情绪变化可察点的数量不尽相同，范围在 2 到 5 之间。数据有效的 30 位被试情绪变化可察点分布情况见图 5。

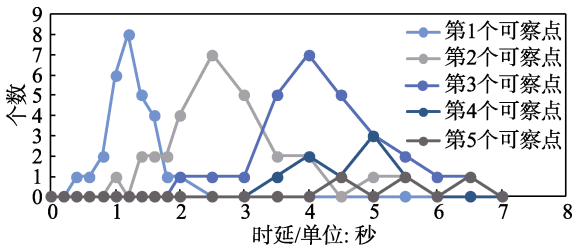


图 5 30 位被试情绪变化可察点分布情况
Fig.5 Distribution of detectable emotional changing points of 30 subjects

4.2 MOS 量表数据处理及结果

实验得到 30 位被试针对 21 组网页时延浏览体验的打分。将这些分值求均值和方差，得到的实验数据被试 MOS 打分均值见图 6 和被试 MOS 打分方差见图 7。

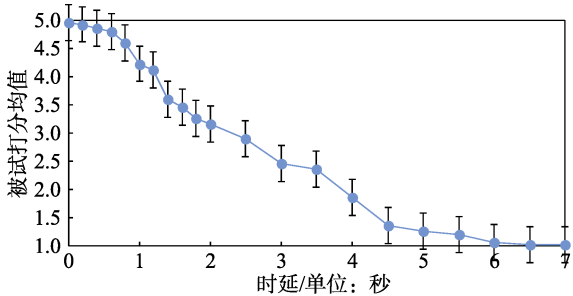


图 6 被试 MOS 打分均值
Fig.6 Average Value of MOS marks

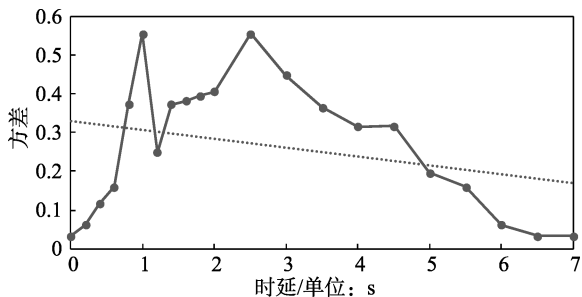


图 7 被试 MOS 打分方差
Fig.7 Variance of MOS marks

5 实验结论

被试的生理数据发生显著性变化时对应的时延长度集中在 3 个区间, 分别为第一次[1 s, 1.4 s], 峰值 1.2 s; 第二次[2.5, 3 s], 峰值 2.5 s; 第三次[3.5 s, 4.5 s], 峰值 4 s。即在该 3 个时延区间内, 用户发生生理上可察觉情绪变化的可能性最高, 浏览体验逐次下降。由于仅有少量用户存在第 4、5 个可察点, 很难通过统计数据作出一致结论。

当时延小于 1 s 时, 被试打分平均值均在 4.5 分以上, 表明被试网页浏览体验良好, 且随着时延长度的增加情绪变化较为缓和。在时延长度 1 s 到 4.5 s 范围内, 用户情绪变化较大, MOS 均值由 4.5 分以上降到 1.5 分以下; 当时延大于 4.5 s 时, 由于用户体验已经达到较差水平, 用户的情绪变化趋势又逐渐缓和。当时延在 0.8 s 到 4.5 s 之间时, 被试对浏览体验的打分结果差异较大; 当时延小于 1 s 或大于 4.5 s 时, 被试对浏览体验的评价较为一致。

综合两者, 用户浏览网页时负面体验的增长主要集中在[1 s, 4.5 s]区间内: 当时延超过 1 s 时, 用户开始出现负向情绪; 当时延超过 2.5 s 时, 用户负向情绪加重; 当时延超过 3.5 s 时, 大部分用户认为此时浏览体验非常糟糕。

6 结语

生理测量方法在国内的用户研究领域仍然处于探索阶段, 这里将生理电技术引入网页体验质量的评价中, 可以弥补国内该领域前人研究中单纯利用主观评价方法的不足, 具有可行性和必要性, 且以主观量表作为辅助验证手段, 得到了用户体验分级的时延区间, 在网页时延对用户体验的影响研究方面具有创新意义。但实验过程中被试的情绪可能受到接触式生理信号采集过程以及时间累积效应的影响, 在一定程度上降低了实验结果的准确性, 但生理信号的客观性和连续性等优势仍将使其成为用户体验研究领域的重要方向之一。

参考文献:

- [1] 林闯, 胡杰, 孔祥震. 用户体验质量(QoE)的模型与评价方法综述[J]. 计算机学报, 2002, 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, 2002, 35(1): 1—15.
- [2] FIEDLER M, HOSSFELD T, TRANG P. A Generic Quantitative Relationship between Quality of Experience and Quality of Service[J]. IEEE Network, 2010, 24(2): 36—41.
- [3] 郝益勇. 提升移动网络用户体验质量的理论与方法研究[D]. 北京: 北京邮电大学, 2012.
HAO Yi-yong. Research on Promoting Quality of Experience in Mobile Network[D]. Beijing: Beijing University of Posts and Telecommunications, 2012.
- [4] BHATTI N, BOUCH A, KUCHINSKY A. Integrating User-perceived Quality into Web Server Design[J]. Computer Networks, 2000, 33(1): 1—6.
- [5] DABROWSKI J, MUNSON E. 40 Years of Searching for the Best Computer System Response Time[J]. Interacting with Computers, 2011, 23(5): 555—564.
- [6] SHNEIDERMAN B. Designing the User Interface: Strategies for Effective Human-Computer Interaction[J]. Journal of the American Society for Information Science, 1988, 39(1): 22—34.
- [7] SEOW S. Designing and Engineering Time: The Psychology of Time Perception in Software[M]. New Jersey: Addison-Wesley, 2008.
- [8] 刘光远, 温万惠, 陈通. 人体生理信号的情感计算方法[M]. 北京: 科学出版社, 2014.
LIU Guang-yuan, WEN Wan-hui, CHEN Tong. Affective Computing Methods of Human Physiological Signals[M]. Beijing: Science Press, 2014.
- [9] HEALEY J, PICARD R. Detecting Stress During Real-world Driving Tasks Using Physiological Sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(2): 156—166.
- [10] WARD R, MARSDEN P. Physiological Responses to Different WEB Page Designs[J]. International Journal of Human-Computer Studies, 2003, 59(1/2): 199—212.
- [11] 葛燕, 陈亚楠, 刘艳芳. 电生理测量在用户体验中的应用[J]. 心理科学进展, 2014, 22(6): 959—967.
GE Yan, CHEN Ya-nan, LIU Yan-fang. Electrophysiological Measures Applied in User Experience Studies[J]. Advances in Psychological Science, 2014, 22(6): 959—967.
- [12] TULLIS T, ALBERT B. 用户体验度量[M]. 北京: 电子工业出版社, 2016.
TULLIS T, ALBERT B. Measuring the User Experience[M]. Beijing: Electronic Industry Press, 2016.
- [13] 陈继华, 李岚, 钱坤喜. 基于多生理信号的情绪初步识别[J]. 生物医学工程研究, 2006, 25(3): 141—146.
CHEN Ji-hua, LI lan, QIAN Kun-xi. Emotion Recognition by Physiological Signals[J]. Journal of Biomedical Engineering Research, 2006, 25(3): 141—146.