

汽车人机交互产品用户满意度度量方法研究

钟建, 唐明清, 唐秋阳, 李文博, 许娜
(重庆大学, 重庆 400044)

摘要:目的 研究基于面部表情识别技术的用户满意度客观度量方法。方法 以两款车载信息系统为载体, 以面部表情识别与BP神经网络算法为技术手段, 设计用户分别与两款系统进行人机交互的实验, 建立用户面部表情与用户主观满意度的映射模型, 并进行用户满意度预测, 对比模型预测值与用户主观量表值, 分析得出模型的预测能力, 验证度量方法的可行性。结论 该模型对用户满意度的预测值与用户主观满意度值的整体均方误差为0.165, 实现了在较小误差范围内的准确预测。模型通过识别用户与产品进行人机交互时的面部表情, 能有效客观地度量用户对产品的满意度。

关键词: 面部表情识别; 人机交互; 可用性; 满意度; 度量模型; BP神经网络

中图分类号: TB472 文献标识码: A 文章编号: 1001-3563(2019)06-0239-06

DOI: 10.19554/j.cnki.1001-3563.2019.06.041

Method for Measuring User Satisfaction with Human-computer Interaction Products of Automobiles

ZHONG Jian, TANG Ming-qing, TANG Qiu-yang, LI Wen-bo, XU Na
(Chongqing University, Chongqing 400044, China)

ABSTRACT: The paper aims to study objective measurement method of user satisfaction based on facial expression recognition technology. Two vehicle-mounted information systems were taken as the carriers to design experiment for human-computer interaction with the two systems respectively through facial expression recognition and BP neural network algorithm. The user's facial data of the first system was used to establish the mapping model between user's facial expression and user's subjective satisfaction, and the model was used to predict the user's satisfaction of the second system. The model predictive value was compared with the user subjective scale value, and the predictive ability of the model was analyzed to verify feasibility of the above measures. The overall mean square error of the model for predicting user satisfaction and user subjective satisfaction is 0.165, which realizes the accurate prediction within the smaller error range. By identifying the facial expression information of users when interacting with the product, the model could effectively and objectively measure users' satisfaction with the product.

KEY WORDS: facial expression recognition; human-computer interaction; availability; satisfaction; metric model, BP neural network

经济的不断发展使汽车进入千家万户, 技术的不断进步让智能网联成为汽车发展的必然趋势。实现车与人、车、路、后台等智能信息交换共享, 被列入中国汽车工业协会对智能网联汽车的发展规划^[1]。车载信息系统作为汽车内外信息交互与娱乐的主要界面, 是智能网联汽车的重要组成部分。可用性好的系统不

仅能让用户驾驶更加安全, 更能让用户产生舒适、愉悦的交互体验。用户体验的好坏直接影响着产品的销售, 决定着产品的成败, 是产品设计重中之重。以用户为核心的设计理念, 已经得到汽车企业的认可^[2]。各大汽车企业纷纷成立了用户体验部门。满意度作为可用性的3个评价指标之一, 是用户在使用产品

收稿日期: 2018-11-12

作者简介: 钟建(1993—), 男, 四川人, 重庆大学硕士生, 主攻产品创新设计。

通信作者: 郭钢(1960—), 男, 重庆人, 重庆大学教授、博士生导师, 主要从事产品数字化设计与制造、驾驶行为研究。

后对产品的主观评价。研究以间接度量用户对车载信息系统人机交互满意度的方式，对提高产品的可用性，提升产品用户体验有重要意义。

1 用户满意度与情绪

对于产品而言，满意是用户在使用产品后的实际感受与对产品的事先期望的相对关系。根据 ISO 标准 9241—11，满意度被定义为“用户在使用产品时的舒适度和可接受程度”，而可用性的定义是“特定的用户在特定的使用场景下，为了达到特定的目标而使用产品时，所感受到的有效性、效率及满意度” [3]。

最初，Bennet^[4]和 Shackel^[5]等将可用性概念用于解释产品使用过程中有效性和效率的程度。随着技术的革新，用户关心的不仅是表现产品工作能力的属性，还越来越关心产品带来的主观感受和情感印象。有研究表明，吸引力、情感上的可用性、愉悦等主观因素影响人们对产品的使用情况^[6]。

有效性和效率可以通过用户使用产品完成特定任务的时间、步骤数、出错次数等客观指标来度量，而满意度是用户对产品的主观评价，难以通过客观的方式直接度量。目前用户满意度的获取主要通过用户进行自我报告，如问卷、量表、访谈等形式。其中 7 级李克特 (Likert) 量表被广泛应用于量化用户的满意状态，从 7 到 1 分为“非常满意、满意、比较满意、一般、比较不满意、不满意、非常不满意” [7]。自我报告不可能是精确的，因为每个人对各满意程度的标准是不同的。所以，在样本量较小的情况下，主观的自我报告并不能非常真实地反映用户的满意度。

根据情绪心理学中修正的詹姆斯—兰格理论，事件会引起人的认知评价，使人产生生理变化和行为，形成人的情绪感受^[8]，见图 1。用户使用产品进行人机交互，会引起用户包括满意度在内的认知评价，形成人的情绪感受。因此，用户满意度与情绪之间存在着内在的联系。研究表明，可用性较强的设备能带来更多的积极情绪，如满意、享受、兴奋，而可用性问题的设备则会带给用户挫败感^[9]，因此情绪可以间接反映用户的满意度。

事件 → 认知/评价 → 生理变化&行为 → 情绪感受

图 1 修正的詹姆斯—兰格理论
Fig.1 The modified James-Lange theory

目前，测量情绪主要依靠自我报告、生理测量、行为反应等方法。自我报告通过问卷、量表或访谈等形式得到量化的用户心理数据，得出用户对产品的主观情绪感受，如 PAD 情绪量表^[10]。用户情感具有复杂性、模糊性、变动等特性，通过主观问卷方法难以获取用户准确的情绪感受。生理测量是对随情绪唤醒而波动的变量的测量，如通过接触式情感信号采集装置测量穿戴者的呼吸、心率、血压、出汗、体温、肌

肉反应、皮肤电等信号的情感计算技术^[11]。生理信号采集装置需要花费大量时间进行穿戴和调试，并且穿戴的装置会在一定程度上影响用户的操作体验，会影响测量结果的准确性。

本文在总结上述研究方法的基础上，提出了基于用户面部表情的产品人机交互系统用户满意度度量方法。用户在人机交互过程中的面部特征变化，是随评价与情绪感受一起产生的行为反应。用户面部表情与用户的满意度之间必然存在着一定的联系。

2 研究框架

2.1 基于面部特征变化的情绪识别方法

面部表情识别技术根据面部特征在特定认知活动中的变化，将人类表情以具体的情绪分类，如高兴、悲伤和愤怒等^[12]。研究者可以从面部表情中提取很多有用信息，如点头、眨眼、抿嘴及其他细微面部特征。Lasa 等^[13]提出了“Eyeface”的测量方法，结合眼动和面部特征联合分析用户使用产品中的情绪。FaceReader 软件基于 Ekman 提出的面部运动编码系统 (FACS) 与表情自动分析技术^[14]，可用于实时采集用户在使用产品过程中的面部表情，自动分析得出任意时间节点高兴、悲伤、生气、惊讶、害怕、厌恶和自然状态这 7 种情绪强度值，并可导出整个视频样本各情绪强度的均值，其软件界面见图 2。该软件运用了上万个人脸样本训练面部表情模型，其中有 4000 多个来自中国、日本等东亚地区，其开发者在软件手册中报告的表情识别准确率达 89%。

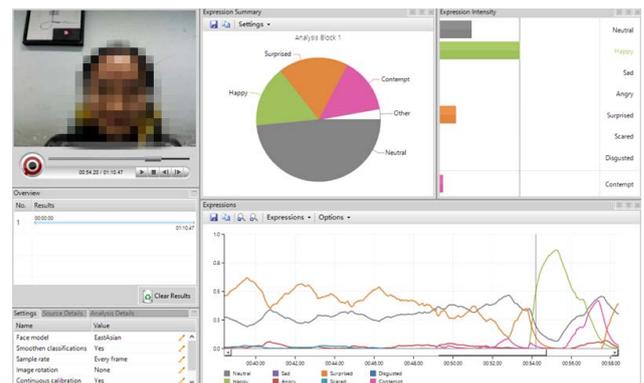


图 2 FaceReader 软件界面
Fig.2 Interface of software FaceReader

2.2 车载系统功能选择

选择 CA-Link 与 CarLife 两款车载互联系统做对比验证实验。两系统分别命名为 A、B 系统，见图 3—图 4。其中 CA-Link 是国内某汽车企业自主开发的系统，而 CarLife 是百度公司推出的。两款系统为同类型的车载互联系统，可通过蓝牙、USB 等方式实现手机与车机互联，是初级的车联网解决方案。由于由不同



图 3 A 系统主界面
Fig.3 Main interface of system A



图 4 B 系统主界面
Fig.4 Main interface of system B

行业的公司设计，它们在各级界面、功能逻辑、个性化设置等方面均有所区别，使人产生不同的交互体验。

实验开始前，进行了关于用户对车载互联系统功能重要度的问卷调查。结果显示，用户对各功能重要度排序为导航、电话、音乐、收音机、天气、语音、视频等，其中导航、电话、音乐的重要度远高于其他应用。因此，选取导航、电话、音乐这 3 个具有代表性的系统功能作为两款系统的研究对象。

2.3 基于面部表情的用户满意度模型

基于面部表情的用户满意度模型建立包括 3 个

过程，分别是用户使用车载互联系统的人机交互过程，数据的采集过程和将数据转化为数学模型的数据分析建模过程。基于面部表情的用户满意度模型建立过程见图 5。

人机交互过程包含了用户集和功能任务集，描述了用户使用目标产品完成不同任务产生实验数据的过程。数据采集包括对面部表情视频、手指操作视频和 7 级 Likert 量表的采集。数据分析建模包括：

- (1) 通过 Facereader 分析视频得出用户使用系统 3 个功能分别产生的 7 种情绪的强度值，将问卷进行统计分析得出用户对 3 个功能的满意度值；
- (2) 使用

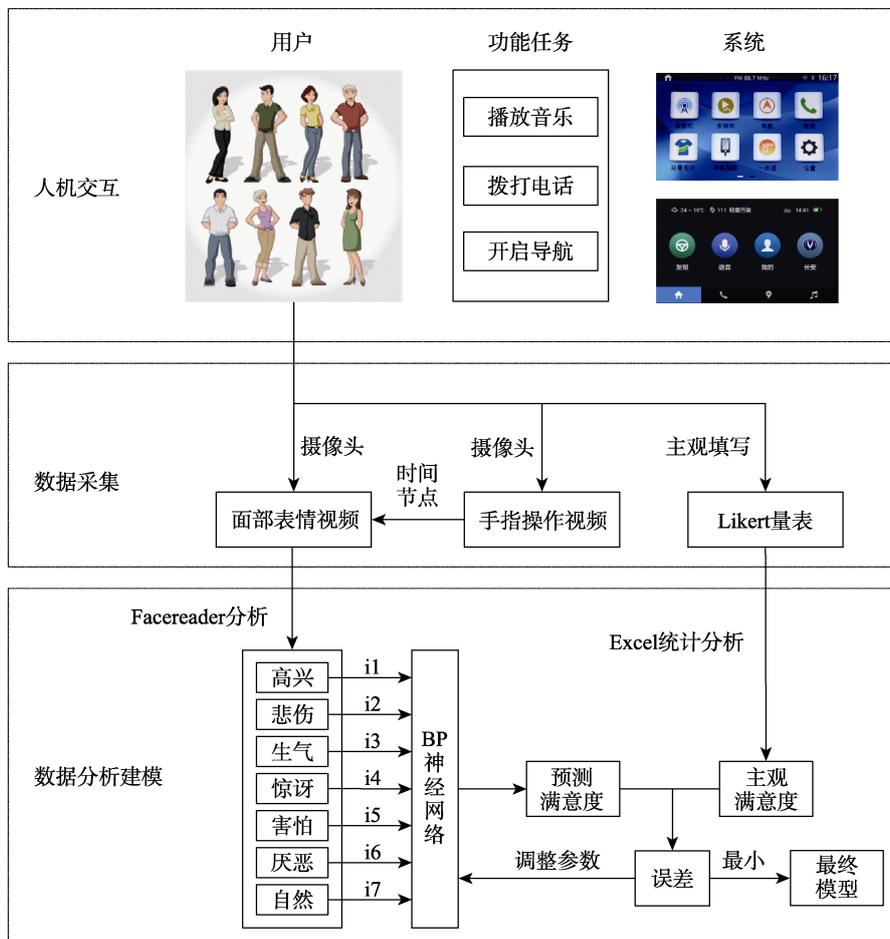


图 5 基于面部表情的用户满意度模型建立过程
Fig.5 Process of establishing the user satisfaction model based on facial expression

BP神经网络算法建立将对应情绪强度值转化为满意度值的数学模型^[15]。

3 研究方法

3.1 实验设计

受实验条件限制并考虑到实验安全,本文设计了基于实车的静态实验。在实验中,每个被试将分别用两款车载互联系统的音乐、电话、导航完成指定的任务。为了避免学习效应对满意度产生影响,部分被试被要求先做A系统的任务,后做B系统的任务,其他被试反之。实验数据分为两部分,A系统的实验数据用BP神经网络算法训练用户满意度模型,B系统的实验数据通过训练完成的模型预测被试对B系统功能元素的满意度,并与用户主观满意度作对比分析。若模型能有效预测用户的满意度,则说明对于不同的车载信息系统,用户使用系统时的面部表情与用户满意度的关系依然成立,即可证明基于面部表情来度量车载系统人机交互用户满意度的方法是可行的。

3.2 被试信息

经过招募和筛选,30名年龄在23~45岁的被试参与实验,其中女性5名、男性25名。被试均来自中国重庆,有两年以上驾驶经验,自愿参加本次实验。由于眼镜会对面部识别产生影响,筛选后的被试均视力正常。为排除使用经验对实验的影响,被试此前均未使用过本实验中的车载互联系统。

3.3 实验环境

实验在一辆车型为ChangAn CS75的SUV中进行。被试被指定坐在正驾驶的座位上操作车载互联系统,被试面部前方有摄像头记录被试面部表情,并有摄像头记录手指操作,两个摄像头型号为罗技C920型1080p高清摄像头。后排有两位测试人员负责使用电脑同时录制被试面部表情视频和手指操作视频,实验环境见图6。



图6 实验环境
Fig.6 Experiment environment

3.4 实验任务

根据用户认为重要度最高的3个功能为用户分别指定了3个任务,见表1。任务设置的目的是让被试完成各功能的使用,产生特定人机交互行为。

表1 任务列表
Tab.1 Tasks

编号	功能	任务详细描述
a	音乐	T1 打开网易云音乐播放器
		T2 选择《难忘的一天》进行播放
		T3 播放10秒后关闭音乐
b	电话	T1 进入拨打电话界面
		T2 输入“02365391307”并拨打
		T3 接通3秒后挂断
c	导航	T1 打开百度地图
		T2 输入“重庆大学”为目的地,并开始导航
		T3 在听见“开始导航”的提示音后,关闭导航并退出

3.5 数据分析建模

数据分析建模分为两个阶段,首先是分析视频与量表得出被试的情绪强度值和满意度值,然后将对应的情绪强度值和满意度值通过BP神经网络建立模型。

3.5.1 数据分析处理

手指操作视频和面部表情视频是同时录制的,从手指操作视频中确定被试使用每个系统功能对应的开始节点和结束节点,由此将每个被试面部表情视频截成3段。然后,将面部表情视频导入FaceReader软件分析得出被试对系统功能的情绪强度值,用 E_{mn} 表示。

$$E_{mn} = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7\} \quad (1)$$

式中: m 为表被试编号, n 为功能编号, $e_1 - e_7$ 分别表示高兴、悲伤、生气、惊讶、害怕、厌恶和自然状态这7种情绪的强度值。

将Likert量表结果统计分析得出每个被试对3个系统功能的主观满意度,用 S_{mn} 表示,其中, m 代表被试编号, n 代表功能编号。因此,情绪强度矩阵 E 和主观满意度矩阵 S 如式2和式3所示。

$$E = [E_{1a}, E_{2a}, E_{3a}, \dots, E_{30a}; E_{1b}, E_{2b}, E_{3b}, \dots, E_{30b}; E_{1c}, E_{2c}, E_{3c}, \dots, E_{30c}]^T \quad (2)$$

$$S = [S_{1a}, S_{2a}, S_{3a}, \dots, S_{30a}; S_{1b}, S_{2b}, S_{3b}, \dots, S_{30b}; S_{1c}, S_{2c}, S_{3c}, \dots, S_{30c}]^T \quad (3)$$

3.5.2 数据建模

使用matlab代码实现BP神经网络算法建模,使用的BP神经网络结构见图7。将A系统的情绪强度矩阵 E 作为BP神经网络的输入矢量,主观满意度矩阵 S 作为目标矢量,导入matlab训练BP神经网络

模型，其中剔除了一组情绪强度值丢失的数据。经过不断调整参数得出最优的模型，网络参数最终设置见表 2。

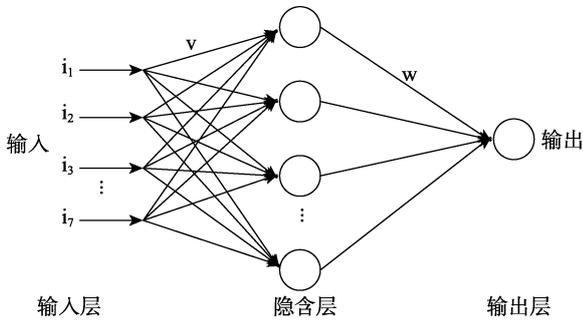


图 7 BP 神经网络结构
Fig.7 Structure of BP neural network

表 2 神经网络参数设置
Tab.2 Parameter setting of neural network

隐含层层数	隐含层 神经元个数	迭代次数	学习率	预设精度
1	11	10000	0.1	0.001

4 结果分析

4.1 模型预测结果

将 B 系统的情绪强度值 E'_{mm} 输入建立的模型得到预测满意度，模型输出的满意度矩阵用 A 表示，见式 4。音乐、电话、导航 3 个功能的用户满意度模型预测结果与 Likert 量表结果对比，见图 8—图 10。

$$A = [A_{1a}, A_{2a}, A_{3a}, \dots, A_{30a}; A_{1b}, A_{2b}, A_{3b}, \dots, A_{30b}; A_{1c}, A_{2c}, A_{3c}, \dots, A_{30c}]^T \quad (4)$$

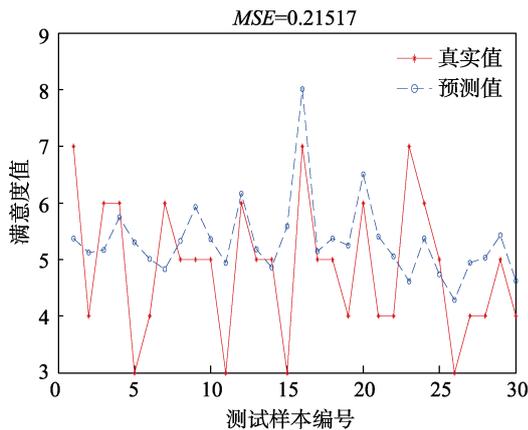


图 8 音乐功能预测值与真实值对比
Fig.8 Comparison between the prediction value of music function and the real value

均方误差是指预测值与真实值之差平方的期望值，记为 MSE ，其计算公式见式 (5)。

$$MSE = \frac{1}{N} \sum_{i=1}^N [x_{(i)} - \hat{x}_{(i)}]^2 \quad (5)$$

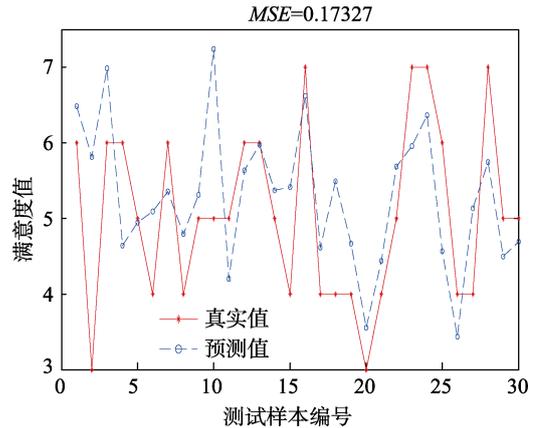


图 9 电话功能预测值与真实值对比
Fig.9 Comparison between the prediction value of telephone function and the real value

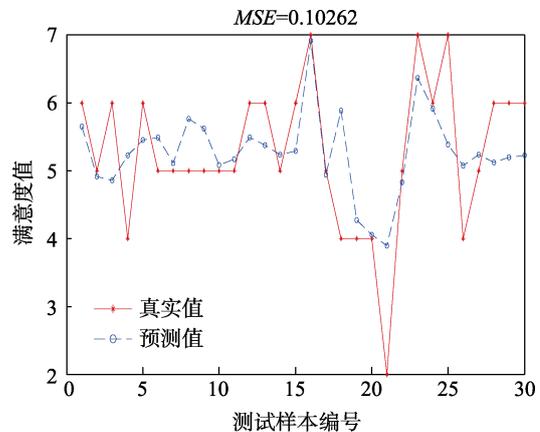


图 10 导航功能预测值与真实值对比
Fig.10 Comparison between the prediction value of navigation function and the real value

式中： N 表示数据个数， $x_{(i)}$ 表示实际量表值， $\hat{x}_{(i)}$ 表示预测值。显然，均方误差 MSE 越小，模型拟合或预测效果越好。

使用式 (5) 计算并保留 3 位有效数字后，得到音乐、电话、导航 3 个功能的均方误差 MSE 值分别为 0.215, 0.173, 0.106，整个 B 系统的均方误差 MSE 为 0.165。说明该模型能在较小误差范围内对用户满意度进行预测，模型有效。

4.2 误差分析

预测模型误差出现的原因，主要有以下 3 个客观因素：(1) 用户满意度的模糊性，没有精确的标准来衡量人的主观感受，模型通过整体样本数据训练所得，代表整体样本的平均标准，必然会产生个体误差；(2) 回归模型的特性，回归模型预测输出值为小数，主观量表所得为整数；用户不可能都是标准的“满意”、“比较满意”，可能是介于两者之间的，因此，通过大量样本训练后的回归模型更能反应用户的真实满意度；(3) Facereader 软件表情识别准确率为 89%，被试的小部分面部信息可能无法识别或识别错

误,导致误差的产生。除上述客观原因之外,数据维度较少且模型训练样本量较小,对模型的准确性也存在较大影响。

5 结语

本文提出了一种基于面部表情特征的满意度度量方法,通过研究用户面部表情与用户满意度之间的关系,建立了基于面部表情识别的车载系统人机交互满意度度量模型。改进后的模型已经实现了在一定误差范围内有效预测用户对不同车载系统人机交互产品功能的满意度。实验结果及模型的建立进一步证明:识别用户与产品交互时产生的面部表情能有效客观地度量其对产品的满意度。

下一步可以从以下几个方面进行更加深入的研究:(1)基于模拟驾驶平台、实车进行模拟路况和真实路况下的动态测试,获取能更加真实地反映车载信息系统满意度的数据,以此建立可靠性更高的模型;(2)获取更多不同车载信息系统的样本数据及数据维度,使用深度神经网络、卷积神经网络等深度学习算法,建立预测性能更好、适用范围更广的模型;(3)剖开面部表情深入面部细节特征,研究用户面部细节特征与用户满意度之间的关系。

参考文献:

- [1] 中国汽车工程学会. 节能与新能源汽车技术路线图[M]. 北京: 机械工业出版社, 2016.
CHINA-SAE. Technology Roadmap for Energy Saving and New Energy Vehicles[M]. Beijing: Mechanical Industry Press, 2016.
- [2] 唐帮备, 郭钢, 王凯, 等. 联合眼动和脑电的汽车工业设计用户体验评选[J]. 计算机集成制造系统, 2015, 21(6): 1449—1459.
TANG Bang-bei, GUO Gang, WANG Kai, et al. User Experience Evaluation and Selection of Automobile Industry Design with Eye Movement and Electroencephalograms[J]. Computer Integrated Manufacturing Systems, 2015, 21(6): 1449—1459.
- [3] 常言说. 基于模型的高可用性用户界面开发研究[D]. 西安: 西北大学, 2013.
CHANG Yan-shuo. The Research of Model-Based High Usability User Interface Development[D]. Xi'an: Northwest University, 2013.
- [4] BENNET J L. Managing to Meet Usability Requirements: Establishing and Meeting Software Development Goals[J]. Visual Display Terminals, 1984(1): 161—184.
- [5] SHACKEL B. The Concept of Usability[J]. Visual Display Terminals, 1984(1): 45—87.
- [6] 张婷. 产品设计的可用性评估方法的开发和验证:以手机为例[D]. 北京: 清华大学, 2009.
ZHANG Ting. Development and Validation of a Methodology for Evaluating Product Design at the Early Stage of Lifecycle: the Call Phone[D]. Beijing: Tsinghua University, 2009.
- [7] SEVA R R, DUH H, HELANDER M G. The Marketing Implications of Affective Product Design[J]. Applied Ergonomics, 2006, 38(6): 723—731.
- [8] SHIOTA M N, KALAT J W. 情绪心理学[M]. 北京: 中国轻工业出版社, 2016.
SHIOTA M N, KALAT J W. Emotional Psychology[M]. Beijing: Chinese Light Industry Press, 2016.
- [9] 杨超. 表情分析技术在可用性评估中的应用研究[D]. 杭州: 浙江理工大学, 2015.
YANG Chao. The Application Research of Facial Expression Analysis Method in Usability Evaluation[D]. Hangzhou: Zhejiang Sci-Tech University, 2015.
- [10] MEHRABIAN A. Pleasure Arousal Dominance: A General Framework for Describing and Measuring Individual Differences in Temperament[J]. Current Psychology, 1996, 14(4): 261—292.
- [11] VERMA G K, TIWARY U S. Multimodal Fusion Framework: A Multiresolution Approach for Emotion Classification and Recognition from Physiological Signals[J]. Neuroimage, 2014, 102(1): 162—172.
- [12] KHAN S A, HUSSAIN A, USMAN M. Facial Expression Recognition on Real World Face Images Using Intelligent Techniques: A Survey[J]. Optik-International Journal for Light and Electron Optics, 2016, 127(15): 6195—6203.
- [13] LASA G, JUSTEL D, RETEGI A. Eyeface: A New Multimethod Tool to Evaluate the Perception of Conceptual User Experiences[J]. Computers in Human Behavior, 2015(52): 359—363.
- [14] EKMAN P, FRIESEN W V. Manual for the Facial Action Coding System[M]. CA: Consulting Psychologists Press, 1978.
- [15] 刘艳, 任章. 基于神经网络的混合模型建模方法及应用[J]. 计算机仿真, 2007(2): 45—48.
LIU Yan, REN Zhang. A Mixed Model Modeling Method Based on Neural Network and Its Application[J]. Computer Simulation, 2007(2): 45—48.