

基于数据驱动的用户需求智能获取方法研究

曹家港¹, 刘键¹, 徐悬²

(1.北京工业大学, 北京 100124; 2.北京理工大学, 北京 100081)

摘要: **目的** 数字化浪潮下人们的消费行为也发生了改变。线上购物伴生着大量用户评论, 具有实时、真实、全面等特点, 这是挖掘用户需求的绝佳资料。本文利用线上评论, 结合感性工学和自然语言处理技术, 提取其中对应于产品元素的用户需求, 为后来的产品改进或开发提供指导。**方法** 本文通过爬虫工具获取目标产品大量评论, 对文本进行词频分析筛选出要分析的产品元素, 根据筛选结果将评论归类, 从中提取形容词, 并经过专家汇总, 最后对含有形容词的语句进行情感分析, 获取用户对产品的需求信息。最后依靠感性工学指导后续设计方案生成。**结论** 本文用亚马逊网站上山地自行车产品进行实例验证, 结果证明该方法能够快速高效地获得用户对产品的准确评价, 是一种切实可行的获取用户需求的智能化设计方法。

关键词: 数据驱动设计; 网络爬虫; 情感分析; 感性工学

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2021)24-0129-11

DOI: 10.19554/j.cnki.1001-3563.2021.24.014

Intelligent Acquisition Method of User Requirements Based on Data Driven

CAO Jia-gang¹, LIU Jian¹, XU Xuan²

(1.Beijing University of Technology, Beijing 100124, China; 2.Beijing Institute of Technology, Beijing100081, China.)

ABSTRACT: Under the digital wave, people's consumption behavior has also changed. Online shopping is accompanied by a large number of user reviews, which are real-time, true, and comprehensive. This is an excellent data for mining user needs. This research uses online reviews, combined with Kansei engineering and natural language processing technology, to extract user needs corresponding to product elements, and provide guidance for subsequent product improvement or development. This article first obtains a large number of reviews of the target product through a crawler tool, performs word frequency analysis on the text to filter out the product elements to be analyzed, and then categorizes the reviews according to the screening results, extracts adjectives from them, and summarizes them by experts, and then conducts the sentence containing adjectives Sentiment analysis to obtain information about the needs of users for the product. Finally, rely on Kansei engineering to guide the generation of subsequent design schemes. This paper uses the mountain bike product on Amazon website to verify the actual situation, and the result proves that this method can quickly and efficiently obtain the user's accurate evaluation of the product, and it is a practical and feasible intelligent design method to obtain the user's needs.

KEY WORDS: data-driven design; web crawler; sentiment analysis; Kansei engineering

设计学经过多年的发展如今已脱胎于艺术学成为一门新兴的交叉学科——采用科学的方法进行设计研究的科学^[1]。在设计学中, 用户需求获取是前期

产品概念设计的重要步骤。传统用户调研方法包括问卷调研、用户访谈、田野调查、焦点小组等。如 P Gringras^[2]为了描述儿科医生对中轻度发育迟缓儿童

收稿日期: 2021-09-22

基金项目: 北京市教委重点项目 (SZ201910005001)

作者简介: 曹家港 (1996—), 男, 河南人, 北京工业大学硕士生, 主攻数据驱动设计。

通信作者: 刘键 (1987—), 男, 山东人, 博士, 北京工业大学副教授, 主要研究方向为智能服务设计、智能交互设计。

进行的调查的范围和费用,对泰晤士河地区 79 名儿科医生发送邮件问卷,并有 86%的医生做了回复;GORTON E^[3]等人为了了解女性对尿动力学检查的耐受性,设计了前瞻性的调查问卷,最终收回了 297 位女性的调研结果。这类方法能直接从用户处获取信息,发掘用户的显、隐性需求,但局限性在于效率低下,面对需要大规模用户信息的设计任务时,高昂的调研支出对很多小微企业或独立设计师都是难以承受的负担;且为保证结果可靠,这类方法往往需要较高的技巧性,来排除用户情绪影响或主观偏见等因素的干扰。

如今大众生活方式受到数字化浪潮的洗礼,网上消费成为很多人的首选,也相应诞生很多线上消费者社区。在电商平台、社区里用户发表了大量关于产品的评论。用户线上评价阶段情绪相对放松自在,因而更能自由表达真实情感。因此在线产品评论被认为是对潜在客户和产品制造商都有用的重要信息资源^[4]。但线上评论数量巨大,人工浏览明显不现实。随着计算机科学技术的飞速发展,各种智能化方法,如卷积神经网络、词频-逆文档频率算法(Term Frequency-Inverse Document Frequency, TF-IDF)等为从评论中提取信息提供了可能。这也为数据驱动的设计打下了基础。顾名思义,数据驱动即为从数据中提取关键信息,并将其作为产品开发或项目进行的主要依据,即将数据作为主要驱动力推动项目进行。该方法能够排除人的感性干扰,结果更准确、快速、客观。如商业智能(Business Intelligence,简称 BI)就是采用数据驱动的新商业模式。学界也有相关研究。如陆蔚华^[5]等人提出了一种多模态人机数据驱动的服务设计方法,有效优化了民航值机体验服务的设计。

1 文献综述

传统调研方法存在一定局限性,因此,本研究意图展开一种基于数据驱动的设计方法,主要涉及情感分析和感性工学等理论,以下分别进行阐述:

1.1 基于自然语言处理的情感分析

自然语言处理(Natural Language Processing,简称 NLP)是人工智能技术中的一个重要分支。顾名思义,该技术是让计算机能理解人类语言并作出反应。从 20 世纪到当下,自然语言处理的实现经历了“基于规则→基于统计→基于神经网络”的转变。现在基于深度学习的自然语言处理已经实现了较高的正确率,并作为人机交互的重要组成部分,应用在人机对话、舆情监测、机器翻译等场景。情感分析是 NLP 的功能之一,是通过分析语句中词的词性及顺序来推测语句的情感倾向,量化为分值,并根据分值将文本判定为积极、消极或中性等情感,是一种分类任务,

也叫作情感挖掘。如今,专家依然在不断改进算法类型来尝试提升其准确率或适应不同场景,如:张瑾^[6]等人针对用户评论方面级别细粒度情感分析任务,提出了 BiGRU-Attention 与门控机制(gated mechanisms)相结合的文本情感分类模型,并用 AI challenger 2018 细粒度情感分析中文数据集做验证,获得了很好的实验效果。

1.2 设计与自然语言处理

基于自然语言处理可以高效处理语言信息的特性,该技术也被广泛应用于学术研究。如在设计学领域,该技术帮助设计师从评论中获取有用信息:如 Robert Ireland^[7]等人介绍了一个用于分析在线产品评论的计算系统,并用科尔曼椅子的评论做实践,让人类设计师分析相同的语料来验证该模型的准确性;KIM W^[8]等人对收集到的产品评论预处理后,用 self-organizing map(SOM,自组织映射)对语料进行聚类分析,并将结果映射在 26×38 的地图上,直观得到了用户情感热点,帮助设计师捕捉用户使用产品时的情绪;张国方^[9]等人利用 TF-IDF 算法及依存句法分析提取产品特征,基于 BERT(Bidirectional Encoder Representation from Transformers)预训练模型进行文本情感分析,并结合其他设计学方法,开发了一种评论数据驱动的产品规划方法,是数字驱动设计的重要尝试。除了简单的情感分析,计算机技术能够调整算法类型,从评论中辨别用户隐性需求等深层次、多维度的信息:如王新^[10]等人使用云平台的隐喻提取技术,基于 Synonyms 的相似度计算及改进的 k-means 算法实现的聚类,格式塔逻辑系统方法将隐式需求映射到产品技术特性上。建立了一个系统的方法来有效地获取客户的隐式需求;纪雪^[11]等人采用基于层次主题模型的半监督算法,依存句法等方法,提出一种考虑产品属性层次性的评论挖掘及需求获取方法,弥补了该领域研究忽视产品层次性的短板。

1.3 感性工学

随着企业越来越重视用户的因素,感性工学(Kansei Engineering)日益受到多方面的瞩目;近年来随着服务设计的兴起,感性工学在服务设计领域也迅速引起了人们的关注^[12],如今已是设计学科中常用的方法之一。如果说人机工程学是让产品更符合人体尺寸的学科,那么感性工学则是让产品更符合人的“心理尺寸”的学科。感性工学体现了以人为本的设计策略,它采用工程学的测量方法研究人的感性,结合人的感性信息和产品设计要素,并进行量化^[13]。在一定规范内定量地表达用户对外界刺激产生的直觉反应,用确切的数字表示用户对产品元素的情感。量化的方式包括语意差异法,眼动测量,肌动测量等。这些量化信息被设计师采用,从中挖掘用户真正喜欢

的产品类型,为后续产品开发作出指导。感性工学在设计领域已经有成熟的应用,如:王飞^[14]等人,在感性工学框架中,通过 SPSS 25.0 软件的使用,确定出拖拉机具有代表性的感性词汇,为拖拉机的创新设计提供新的思路;FENECH A^[15]等人提出了一种感性工学和集成化的方法来支持多周期的产品进化。

1.4 感性工学与自然语言处理

感性工学中将个人情感化为分值是从感性到理性的关键环节。传统情况下这时需要被试者参与,这样得到的结果最为直接,但无疑成本也较高。其实用户打分也是将自己的情感起伏映射为数值的过 程,这与情感分析的过程不谋而合。近年来学术界有了感性工学与计算机科学技术融合的研究趋势^[16],通过准确高效的算法获得数据结果,经过感性工学承接并转换为可直接作用于人的服务或产品:CHIU M C^[17]等人就用文本挖掘分析顾客的线上评论并提取喜好,并依靠形态学分析和感性工学为特定产品识别未来的趋势,探索重大机遇;YEH C T^[18]等人提出了一种将感性工学与数据挖掘技术相结合的服务设计方法,收集客户的相关感知词汇和服务属性,然后通过数据挖掘技术量化服务属性、感知响应和使用意图之间的关系。

综上所述,结合数字技术的智能设计方法吸引了国内外众多学者去研究。在全球化背景下,先进的设计方法有助于中国制造业升级逐步实现从“代工生产”向“委托设计与制造”再向“自有品牌生产”的转化^[19]。不仅如此,随着数据处理在设计中的比重日益加大,也加快了智能设计方法从“传统设计方法”到“基于程序的设计”再到“数据驱动的设计”的转变^[20]。基于数据处理的新设计流程模型正在形成(见图 1)。在这种模型下,智能技术、数理统计方法等数据处理技术帮助设计师处理大量数据素材,从中发掘数据规律或是捕捉有用的信息,在感性工学、卡诺模型等一系列产品开发技术的加持下,生成最终的产品或服务。当然在此模型中,人类设计师依然处于关键地位,但相较于传统设计开发流程已经是较大的革新。数据是新时代的核心资源,对设计来说也是如此。合理的数据分析能让设计师更加贴近用户痛点,且过程更有说服力,设计过程不像灵感乍现那样不可再现。设计开始向数据驱动靠拢。

2 研究方案

有很多学者对利用情感分析的设计流程进行了探索,但大多是基于固定模式的设计,并不一定适用于多样的产品,泛化性较弱。基于此,本研究建立一种通用性较强的设计流程:基于情感分析与文本处理技术,建立评论内容与产品元素的直接连接,并将情感分析结果用于感性工学分析,开发了一种由数据驱

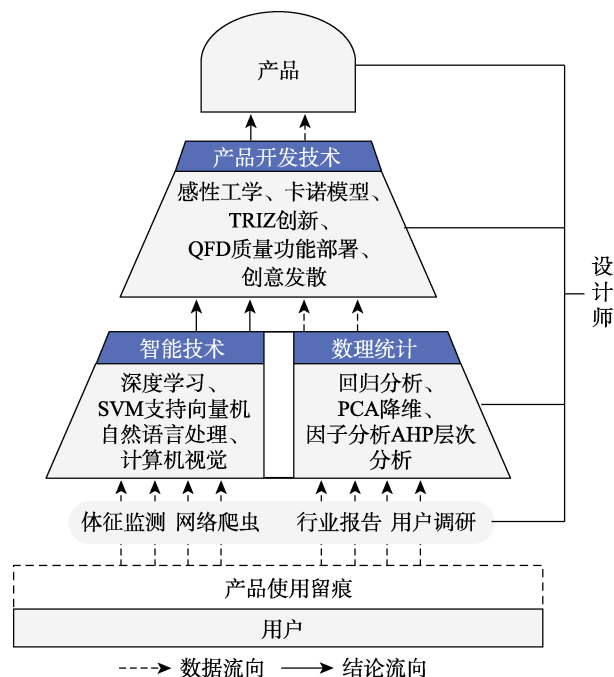


图 1 基于数据处理的新设计流程模型

Fig.1 New design process model based on data processing

动的设计流程。再者,很多相关研究将重心放在跨界设计的原理解释上,对实际实验着墨并不多,这使得初学者难以下手,本文给出了相关技术的实现流程及每个过程的代码步骤,希望能对其他设计学研究者提供学术参考。实验步骤如下。

(1) 利用爬虫工具,收集目标产品的大量在线评论;(2) 汇总所有评论,用 python 程序对其进行词频分析。按照文本词语出现频率降序排列,找出用户最为关注的产品元素并筛选;(3) 进一步应用 python 程序文本处理,找出评论中形容目标元素的句子,并提取其中的形容词;(4) 专家对形容词进行汇总分类,对应句子也相应归类分组;(5) 利用自然语言处理开源工具库对不同组的句子进行情感分析,获得情感参数并取其平均值,作为该形容词的评价得分。

以上 5 个步骤分别为语料搜集→提取产品元素→获取形容词→专家评审→计算情感得分,情感分析的目标元素直接由评论内容决定。通过以上步骤的分析计算,可以得出在同一产品品类下,用户对不同类型配件的评价,并定量地表示出来。这可以为进一步的产品改进或是用户个性化定制提供依据。值得一提的是,本研究聚焦于形态研究,因此并没有将色彩作为评价标准。

3 实验过程

在本实验中,以山地自行车为目标产品。因为该产品复杂度适中,便于分析;且该产品专业性较强,用户对其有较高的要求。评论来源选择电商平台亚马

逊,因为只有买家才能够发言,避免了“云买家”的误导言论。

3.1 语料搜集

亚马逊平台上自行车有“折叠自行车”“公路自行车”“电动自行车”等多个种类。在语料搜集时,为确保研究的准确性,保证用户评价语境的一致,目

标产品选为山地自行车。

在亚马逊平台上,搜集“mountain bike”品类的产品,然后利用谷歌商店的“Instant Data Scraper”爬虫插件,爬取“Mongoose”“Schwinn”等品牌的自行车产品评论共四千余条。为后续分类方便,将每款产品评论单独命名并独立储存。

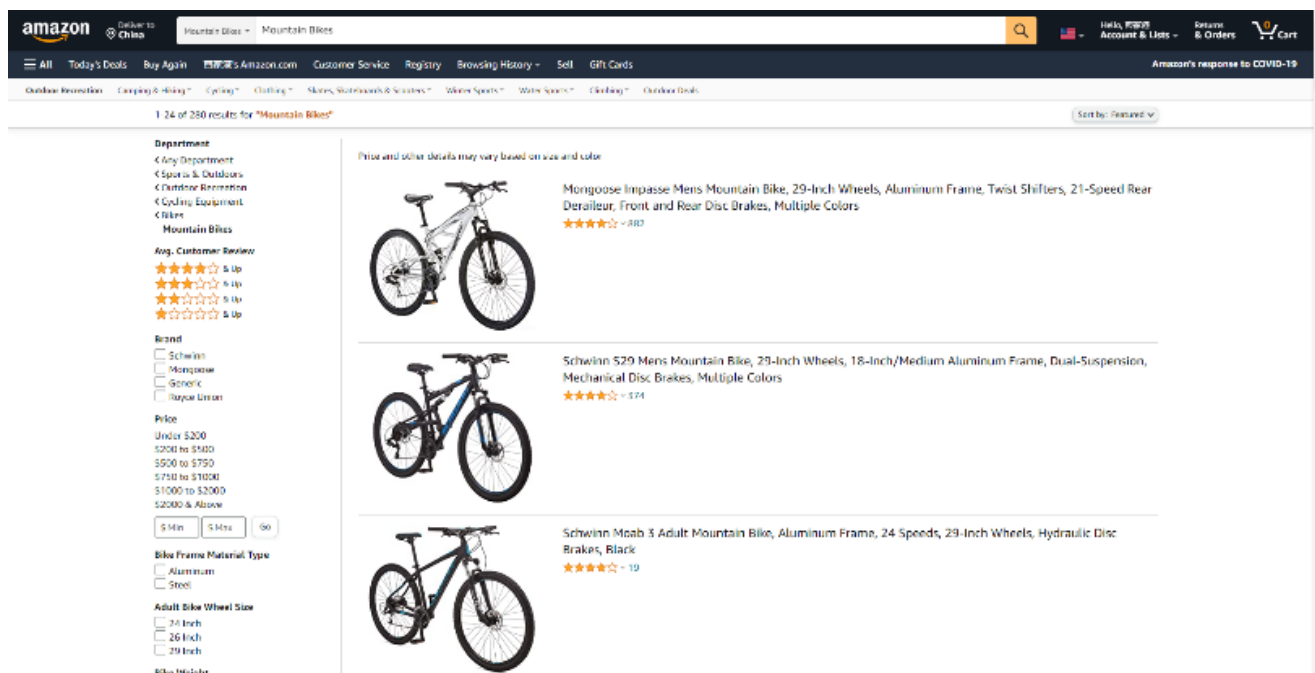


图 2 亚马逊搜索界面
Fig.2 Amazon's search page

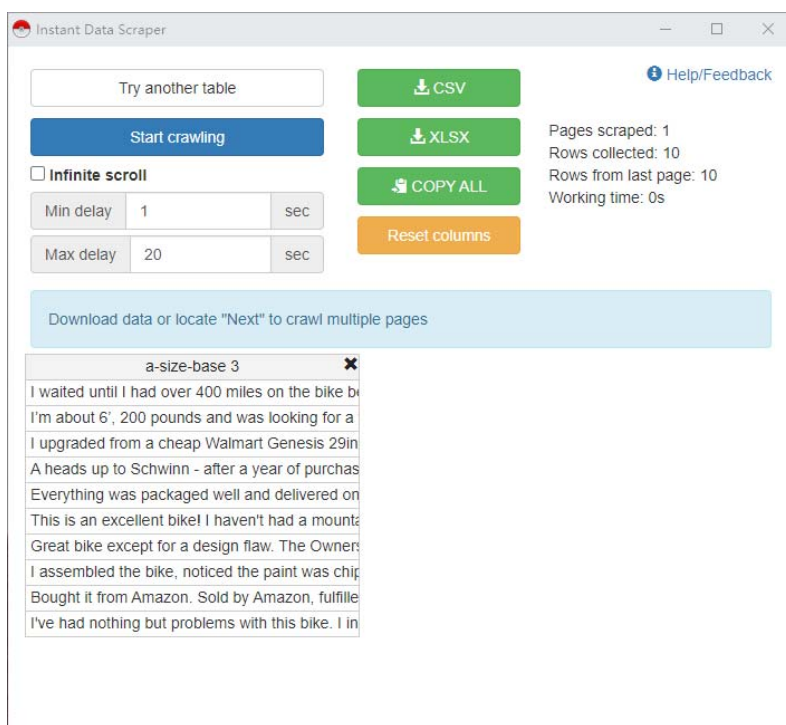


图 3 爬虫插件界面
Fig.3 Interface of crawler extender



图 4 分词代码

Fig.4 Code for word separation

3.2 提取产品元素

将所有评论汇总，用 python3.8 编程对其进行词频分析，并输出频率最高的前 400 个单词。代码如下：

该段代码首先编译了“get_text”函数，遍历所有评论内容，去除了文本中的无效字符，并将文本内容以词为单位分隔开；随后又遍历所有单词，对每个单词出现次数进行计数；得出每个单词出现频次后按照降序排列，最后输出出现最多的前 400 个单词。删除掉“the”“and”等没有分析价值的定冠词、连词等，留下产品元素，出现最多的 11 个产品元素结果汇集见表 1。

其中已经将表示同一元素的不同表达合并，如“Derailleur”和“Shifter”都表示变速器；“Handle”，“Handlebars”和“Bars”都表示车把手。由表 1 可直观看出，基于提取到的语料，用户最关注的元素。但这 11 个元素并不都适合做外观上的分析。首先，对自行车产品的评价并不是完全基于外观，更多的是功能实现方面。如：Further riding indicates that the gears just don't shift quickly/smoothly relative to better sets. Adjusting cable tension helped only incrementally but still could no solidly land a gear, especially when moving across several sprockets quickly。

表 1 用户评论词频统计结果

Tab.1 Word frequency statistics of user comments

产品元素	出现频次
Seat 座位	1251
Brake 刹车	1238
Tire 轮胎	1060
Gear 齿轮	1041
Wheel 车轮	1014
Handle+Handlebar+Bar 把手	834
Pedal 踏板	619
Frame 车架	481
Derailleur+Shifter 变速器	385
Chain 车链条	366
Fender 挡泥板	282

本研究重点在于产品外观的设计，因此需要排除外观占比较低的元素如“Brake”“Gear”“Derailleur”“Chain”等。因为用户评价这些元素更多是基于功能方面的考量，而它们的功能实现在于他们的材质、机械设计等方面。

其次，“Wheel”“Fender”在后续的分析中也被弃用。虽然这两者出现频率较高，但是观察目标产品群特征，这两者差别不大。“Wheel”车轮均为金属自行车车轮辐条，但是胎面有区别，因此舍弃“Wheel”但保留“Tire”为分析的目标元素。而搜集的所有 21 款产品中，只有 2 款拥有“Fender”挡泥板，同时因为挡泥板会影响到山地车的整体性能和稳定性，破坏车手的舒适性^[21]，因此也舍弃“Fender”。

最后舍弃掉“Pedal”踏板，因为通过观察案例库，踏板也并没有明显的外观差别。通过抽样调查评论内容，关于踏板的评价大多着眼于尺寸的问题，如：When you are going slow-maneuvering and turn the front wheel it hits the pedal-yes the wheel rubs on the pedal.

I'm only 5'11 and with the seat at maximum height, my knee is still bent when the pedal is fully extended.

这是车部件之间的协调性问题，不属于外观的范畴，故舍去。因此，为接下来的实验留下“Seat”“Tire”“Handlebar”“Frame”4 个关键元素。其实粗略地看，山地自行车最为明显的元素也的确是这 4 个。

3.3 获取形容词

3.3.1 文本归类

确定分析目标后，结合 Zwicky 开发的形态学分析方法对产品元素进行分析^[22]，根据分析结果将产品元素区分为不同款式，并将对应评论分类。款式分类结果如表 2。

将评论内容根据样本类型划分，也对应分开。这样，同组评论内容的形容目标都是相同类型的产品。

3.3.2 截取句子、提取形容词

一段评论中可能会提到多个元素，在分析时需要一一对应，防止混淆。因此在提取形容词之前需要定

表 2 产品元素分类结果
Tab.2 Classification results of product elements

	类别一	类别二	类别三
Seat 车座	 宽座	 窄座	
Tire 轮胎	 凹胎	 凸胎	
Handlebar 车把手	 平把	 燕把	 弯把
Frame 车架	 减震架	 直架	 弯架

位到提到产品元素的那句话，再提取该句话中的形容词。操作步骤为：评论分句、目标定位、提取形容词三个环节。以“框架”为例，截取评论中带有“frame”“frames”的句子过程如下所示。

1）评论分句（见图 5）。该段代码首先读取准备好的评论 txt 文件，用 strip 函数去除文本前后空格，然后以句号“.”为标准将评论分割，并逐行写入新建的“分句”文件夹中。文本格式变化示意图 6。

2）目标定位（见图 7）。为方便后续提取形容词，这一步将 txt 中的文本内容逐行转移至 csv 文件；接着编译了 cleantxt 函数，调用 nltk 库中的英文停用词，来对文本进行清洗，提高后续分析的准确性；最后用 contains 命令保留含有关键词“frame”和“frames”的句子。为保证与下一阶段的连贯性，这一步的结果不再另外导出。

3）提取形容词。该代码承接于“目标定位”的部分（见图 8），首先编译了 reviewblob、tagsdict、

```
import codecs
txt=open(r"C:\Users\ok\Desktop\尝试.txt",encoding='utf-8')
list=""
text=[]
for line in txt:
    li=line.strip("\n")
    list=list+li
    list=list+' '
    list.strip(' ')
    list=list.split(" ")
    file_write_obj = open(r"C:\Users\ok\Desktop\分句.txt", "w",encoding='utf-8')
    for h in list:
        if h == '':
            continue
        h.strip(' ')
        h=h+'.'
        file_write_obj.write(h)
        file_write_obj.write('\n')
    file_write_obj.close()
print("保存文件成功")
```

← 打开评论文件

← 拆分文本内容

← 创建储存文件

← 将句子集中写入新文件中

图 5 评论分句代码

Fig.5 Code for breaking comments into sentences

"For the price, this is a decent bike.I bought it (in white) for my 9yr old daughter. It is he
more akin to a track bike or a triathletes lower "aero" position.A full size adult might
This is the next size up for my 10 year old to grow into. He can just touch the ground c
This is a great bike. We are a biking family. I have 5 bikes myself and everyone else has
Nice looking bike. My son hasn' t ridden it yet because (after 15 hours spent trying to e
"As long as your probably going to want to buy a new saddle I also recommend purcha
If buying for a preteen or teen boy (who' s hard on things), I wouldn' t suggest this. B
Overall the bike is decent though sent with major back tire frame defect
This bike is great. easy to assemble and adjust if you've owned bikes before and are sor
"Listed as a "boys" bike with the picture showing a boys top tube on the amazon site.
Great bike for the price. Assembled in 20 min. Front rim was slightly warped, but for 9
My grandson loves the bike ☺☺
Very please with this purchase. I bought for my 11 year old son who is on the larger sid
Great Bike, purchase with confidence
I gave it a 4 because I had to pay to have them adjust the brake as it was not setup pro
Purchased this bike for my son. He used it a total of 3 THREE times and it broke the thr
My son loves this cycle.Pros:very strong, durable, comfortable to drive, switching gear is
My son loves this bike! Easy to put together and great quality!
Did take my husband a while to put together but it seems pretty solid and our 10 year c
Great bike. well packaged, delivered on time and a fun project to assemble with my sor
Got this for my son for Christmas. It was his second ride on it today and the rear derail
Bike didn' t come with all parts for assembly. Contacted Pacific Cycling and part was s
Good bike for the money.
My boy loves this bike
Good bike at a fair price.
The bike is good but the shifting mechanism is horrible, bad quality, the brake system is v

分句

"For the price, this is a decent bike.
I bought it (in white) for my 9yr old daughter.
It is her first "big kid's" bike.
The white with pink accented decals fits the "girl" style fine.
though I don't know why the listing says "boys".
The drop-style top tube allows for easy on and off.
I'd give it 5 stars except.
The brakes and shifters needed to be reinstalled to work properly.
Shifters were upside down on wrong ends of the bar.
That force brake levers way too far down to work with smaller hands.
Be prepared to take this bike to a shop and have a once over before you put your kid o
And that may not be cheap.
Also.
this is a 24" wheel bike.
By definition it is designed for smaller bodies.
Why Mongoose puts a full reach handle bar stem on this.
and a full size grownup's-butt seat on it.
is beyond me.
This forces a smaller person to have to reach WAY out to get hands in position and its
more akin to a track bike or a triathletes lower "aero" position.
A full size adult might fit the seat to handlebar distance.
but a full size adult would go to a 26" wheel (or larger) frame.
Mongoose is not the only company that makes this mistake.
Have seen it on many mid to low end bikes.
Simply not thought thru well enough.
And yet the answers so simple.

图 6 文本格式变化示意

Fig.6 Change of text format

<pre>import textblob import pandas as pd import numpy as np import nltk from nltk.corpus import stopwords from nltk import word_tokenize, PorterStemmer, WordNetLemmatizer import csv</pre>	← 调用需要的包
<pre>f = open(r'C:\Users\ok\Desktop\新建文件夹\分句.csv', 'w', newline='', encoding='utf-8') fieldnames = list(range(1)) csv_dict = csv.DictWriter(f, fieldnames=fieldnames) csv_dict.writeheader() writer = csv.writer(f, dialect='excel')</pre>	← 创建CSV文件并设置参数
<pre>with open(r'C:\Users\ok\Desktop\新建文件夹\分句.txt', 'r', encoding='utf-8') as f2: line = f2.readline() while line!="": tmp = [[element for element in line.strip('\n').split('\n')]] writer.writerow(tmp) line = f2.readline() print("OK") f2.close() f.close()</pre>	← 将分句内容写入CSV文件
<pre>def cleantxt(test_text): words = word_tokenize(test_text) stopwords = nltk.corpus.stopwords.words('english') words_clean = [w for w in words if w.lower() not in stopwords and w.isalpha()] lemmatizer = WordNetLemmatizer() words_lemmas = [lemmatizer.lemmatize(t) for t in words_clean] clean_words = ' '.join(words_lemmas) return clean_words</pre>	← 文本清理，去除停用词等
<pre>data = pd.read_csv(r'C:\Users\ok\Desktop\新建文件夹\分句.csv') data['review_clean'] = data.iloc[:, 0].apply(lambda x: cleantxt(x)) sentence_all = data['review_clean'].str.split(' ', expand=True) sentence = sentence_all[sentence_all.iloc[:, 0].str.contains('frame frames')]</pre>	← 选出包含目标元素的句子

图 7 目标定位代码

Fig.7 Target location code

<pre>def reviewblob(text): blob = textblob.TextBlob(text) tags = blob.tags return tags</pre>	← 创建函数：词性标注
<pre>def tagsdict(text): tagsdict = dict(text) return tagsdict</pre>	← 创建函数：建立字典
<pre>def get_key2(dct, value): return [k for (k,v) in dct.items() if v == value]</pre>	← 创建函数：寻找内容
<pre>sentence['tags'] = sentence.iloc[:, 0].apply(lambda x: reviewblob(x)) sentence['tags_list'] = sentence['tags'].apply(lambda x: tagsdict(x)) sentence['adj'] = sentence['tags_list'].apply(lambda x: get_key2(x, 'JJ')) sentence.iloc[:, [0, 3]].to_csv(r'C:\Users\ok\Desktop\新建文件夹\分句_形容词.csv', index=False) print('结束')</pre>	← 调用函数，找出句子中的形容词

图 8 提取形容词代码

Fig.8 Code for extracting adjectives

get_key2 等 3 个函数，作用分别为词性标注、建立词典、寻找关键词；接着根据这 3 个函数，对上文筛选出的“sentence”内容进行依次处理就可以找出文本中含有的形容词汇，将结果写入名为“分句_形容词”的 csv 文件。形容词提取结果见图 9。

3.4 专家评审

用户的评价用语具有口语化、零碎化的特点。和感性工学结合，需要设计师把用户口语化的形容方式化为专业词汇。提取出的评论形容词表述及专家总结归纳后的形容词见表 3。

分完形容词后，将评论内容也跟随含有的形容词归属，进行区分或汇合。为后续情感分析做铺垫。

3.5 情感得分

对文本进行情感分析。自然语言处理发展迅速，如今已经有很多互联网公司推出了 NLP 的商业服务，如百度大脑的 AI 开放平台推出的语言处理应用技术；阿里云推出的自然语言处理等。此外，还有专业程序开发人员开发的开源 NLP 库，如“Spark NLP”“NLTK”“OpenNLP”等。“库”是封装实现功能的底层代码而来，提供完成任务所需的 API 接口，可以视为技术黑盒子。使用者只需要按照一般流程调用该库，并按照库的使用方式，输入它所需的参数，库便会自动运行得到结果。这为后来使用者提供了便利，是技术传承的重要方式。

	A	B	C	D	E	F	G	H	I	J	K
1		0 adj									
2	nice solid frame bike great tread tire easy shift throttle control gear selector	[nice, 'solid', 'great', 'tread', 'easy', 'shift']									
3	two Preset inside frame mount water bottle holder	[]									
4	Overall bike decent though sent major back tire frame defectThis bike great	[Overall, 'sent', 'major', 'great']									
5	rear wheel get pulled left pedal hard way tune frame also frame rear wheel arm inch	[rear, 'pedal', 'hard', 'mm', 'side']									
6	frame accommodates either fender	[]									
7	got bike sale would pay MSRP knowing frame skewed enough ruin wheel alignment	[bike, 'enough', 'stable', 'strong', 'solid']									
8	frame bit smaller GT XCR professional dirt trail bike GT also inch	[professional]									
9	already guessed handle bar stem come fitted frame upside presume save half inch	[handle, 'frame', 'upside', 'inch', 'necessary', 'dilemma', 'new', 'top', 'send', 'piece', 'cheap', 'dismantle', 'whole']									
10	frame nice shifter bolt junk	[nice]									
11	part seems correctly fork stem still move inside frame assembly finished	[fork]									
12	anyone help would appreciate greatly bike arrived dent frame box good shape dou	[good]									
13	frame probably weight lb dual front rear disk brake nice	[lb, 'dual', 'rear', 'nice']									
14	Mountain Bike Frame MaterialsThe frame influence bike weight strength longevity ric	[professional]									
15	model lighter aluminum frame result manufacturer expending dollar effort selection	[]									
16	frame material include steel titanium carbon fiber	[frame]									
17	Good bike took like day get put together going buy bigger Nabi tire cuz go trail wo	[Good, 'nubbe', 'fast', 'lil', 'foot', 'worn', 'already', 'awesome', 'big', 'beautiful', 'wrapped', 'lol', 'great', 'real', 'axel']									
18	picked bike today still rub frame sooo lm going order wheel mean time hopefully wv	[sooo, 'mean', 'new', 'nice', 'bad', 'serious']									
19	Even frame medium big	[frame, 'big']									
20	tire even frame feel like big bike	[frame, 'big']									
21	crotch lightly hit frame still painful enough hitting root rock stop abruptly	[painful]									
22	mile road maybe mile trail rear axle broke disk brake rubbed bad regardless positio	[mile, 'axle', 'disk', 'bad', 'regardless', 'due']									
23	frame beefy	[]									
24	fan frame cable clip keep pulling likely change put small black wire tire alternative	[likely, 'small', 'black']									
25	cable operates rear brake supposed secured bike frame small plastic clip see photo	[small]									
26	frame accommodates standard water bottle holder purchased separately installed	[standard]									
27	someone would expect mountain bike aluminum frame wheel tire heavier lightweight	[lightweight]									
28	frame warranted life owner	[]									
29	plastic clip hold rear brake cable frame well secured	[rear]									
30	Pedals super cheap plastic frame lacked free spinning bearing	[cheap, 'free']									

图 9 形容词提取结果

Fig.9 The results of adjective extraction

表 3 形容词归纳汇总

Tab.3 Summary of adjective conversion

产品元素		原形容词	归纳为
把手	平把	耐用的, 坚固的, 很棒的..... Durable, sturdy, Awesome.....	耐用 Durable
		光滑, 有缺陷的, 不舒服的..... Smooth, defective, uncomfortable.....	易滑的 Skidding
		时髦的, 倾斜的, 流行的..... Funky, lean, pop.....	时髦的 Funky
	燕把	可靠的, 必要的, 坚固的..... Reliable, necessary, sturdy.....	可靠的 Reliable
		蹩脚的, 愚蠢的, 难以接受的..... Crappy, stupid, unacceptable.....	蹩脚的 Crappy
	弯把	尴尬的, 失望的, 缺陷的..... Awkward, disappointed, defect.....	尴尬的 Awkward
车架	直架	出色的, 经典的, 完美..... Outstanding, classic, perfect.....	出色的 Outstanding
		可靠的, 强硬的, 专业的..... Reliable, tough, specialized.....	可靠的 Reliable
		技术性的, 无与伦比, 与众不同..... Technical, unparalleled, different.....	技术感 Technical sense
	弯架	沉重的, 有缺陷的..... Heavy, defective.....	沉重的 Heavy
		专业的, 合理的, 坚毅的..... Professional, reasonable, gritty.....	专业的 Professional
		笨重的, 强壮的, 沉重的..... Unwieldy, strong, heavy.....	笨拙的 Unwieldy
轮胎	凹面	弯曲的, 方便的, 耐用的..... Flex, handy, durable.....	灵活的 Flexible
		不适的, 脆弱的, 不安全..... Upsetting, fragile, unsafe.....	不适的 Upsetting
		专业的, 很棒, 极好的..... Professional、great、awesome.....	出色的 Outstanding
	凸面	棘手的, 古怪的..... Tricky、whack.....	不舒服的 Uncomfortable
		可忍受的, 得体的, 合理的..... Tolerable、decent、reasonable.....	稳健的 Moderate
		印象深刻的, 特别的, 不可思议..... Impressed, particular, incredible.....	印象深刻的 Impressive
车座	宽座	攻击性的, 古怪的, 怪异的..... Aggressive, odd, weird.....	攻击性 Aggressive
		圆滑的, 领先的, 精彩的..... Sleek, leading, wonderful.....	时尚的 Fashion
		适应的, 轻松的, 方便的..... Accommodate, cushy, handy.....	可调节的 Adjustable
	窄座	笨拙的, 糟糕的, 噱头..... Unwieldy, terrible, gimmicky.....	笨拙的 Unwieldy
		经典的, 直观的, 规律的..... Classic, intuitive, regular.....	经典的 Classic
		了不起的, 人机工程, 普遍的..... Fantastic, ergonomic, pandemic.....	流行的 Popular
		蹩脚的, 糟糕的, 可怕的..... Crappy, awful, horrible.....	蹩脚的 Crappy
		可靠的, 可接受的, 新颖的..... Reliable, acceptable, original.....	可靠的 Reliable

```
import pandas as pd
import textblob

sentence = pd.read_csv(r'C:\Users\ok\Desktop\类别-分句.csv')
sentence_topic = sentence[sentence.iloc[:,0].str.contains('professional|reasonable|gritty|standard|sturdy|reasonable|reliable')]

def reviewsentiment(text):
    blob = textblob.TextBlob(text)
    result_sentiment = blob.sentiment.polarity
    return result_sentiment

sentence_topic['sentiment']=sentence_topic.iloc[:,0].apply(lambda x: reviewsentiment(x))
sentence_topic['label'] = sentence_topic['sentiment'].apply(lambda x : "neg" if x<0 else "neu" if x>=0.0 and x<=0.2 else "pos")
sentence_topic.to_csv(r'C:\Users\ok\Desktop\类别-专业情感.csv',index=False)
```

图 10 情感分析代码
Fig.10 Code for sentiment analysis

	A	B	C	D
1	0 adj		sentiment	label
2	nice solid frame bike g	['nice', 'solid', 'great', 'tread',	0.458333	pos
3	got bike sale would pe	['bike', 'enough', 'stable', 'str	0.144444	neu
4	frame bit smaller GT X	['professional']	0.05	neu
5	already guessed handl	['handle', 'frame', 'upside', 'ir	0.178451	neu
6	Good bike took like d	['Good', 'nubbe', 'fast', 'lil', 'fc	0.383333	pos
7	frame accommodates	['standard']	0	neu
8	Pros bike frame feel m	['much', 'smooth', 'cheap', 'lk	0.272619	pos
9	aluminum frame sturd	['sturdy', 'heavy']	-0.2	neg
10	thing noticed right qui	['right', 'quick', 'light', 'simple	0.118899	neu
11	frame relatively light ri	['light', 'rigid', 'reliable', 'norm	0.275	pos
12	size frame perfect mig	['perfect', 'big']	0.270833	pos
13	one bit odd really fit tr	['odd', 'fit', 'traditional', 'able	0.122222	neu
14	frame perfect lb	['perfect']	1	pos
15	hoping ride frame ben	['bent', 'professional', 'worth	0.2	neu
16	get pay frame tire bike	['reasonable']	0.011111	neu
17			0.219016	
18				

图 11 情感分析结果
Fig.11 Results of sentiment analysis

本研究主要使用 TextBlob 进行情感分析。TextBlob 是用 python 编写的开源文本处理库，基于 NLTK 和 pattern 开发，可以对文本进行词性标注，情绪分类，名词提取等操作。调用 Textblob 包对语句进行情感分析代码如下，将结果写入特定文件。

这一部分情感分析代码先利用 contains 命令选出带有特定形容词的语句（见图 10），存储在“sentence_topic”中。接着编译 reviewsentiment 函数对文本进行情感极性判断。每个句子都对应情感参数 x ， x 取值范围为-1 至 1。经过调试，这里设定：若 $x<0$ ，则判定该句为负面情绪“negative”；若 $0\leq x\leq 0.2$ ，则判定该句为中立“neutral”；若 $x>0.2$ ，则为积极情绪“positive”。最后将结果写入名为“类别-专业情感”的 csv 文件中，情感分析结果见图 11。其中 A 列是原句子，B 列是包含的形容词，C 列是情感分数，D 列是情感极性。整合所有情感分数取平均值，作为产品该特点的情感参数。

4 结果及结论

对各个产品元素下的所有形容词重复以上步骤，汇总得到情感分析结果见表 4。

分析结果能清晰地看出用户对产品元素的需求，以把手为例，用户更看重耐用性、粗糙度及外观特征等特性，这为以后的产品改进提供了理论支撑。同时在表中能直观地表达用户对产品不同类型元素的情感。“特点”表示对产品元素的定性看法，“得分”

表 4 情感分析结果				
Tab.4 Results of sentiment analysis				
元素	类型	特点	得分	
车架	平把	耐用的 Durable	0.070437	
		易滑的 Skidding	0.027018	
		时髦的 Funky	0.055	
	燕把	可靠的 Reliable	0.167049	
		蹩脚的 Crappy	0.116167	
		尴尬的 Awkward	-0.18321	
	弯把	出色的 Outstanding	0.394938	
		可靠的 Reliable	0.163	
	减震架	技术感 Technical sense	0.054524	
		沉重的 Heavy	0.154523	
轮胎	直架	专业的 Professional	0.219016	
		笨拙的 Unwieldy	0.178908	
		灵活的 Flexible	0.288249	
	凸面	不适的 Upsetting	0.248157	
		出色的 Outstanding	0.043198	
		不舒服的 Uncomfortable	0.025833	
车座	宽座	稳健的 Moderate	0.173148	
		印象深刻的 Impressive	0.379262	
		攻击性的 Aggressive	0.014157	
	窄座	时尚的 Fashion	0.029365	
		可调节的 Adjustable	0.343561	
		笨拙的 Unwieldy	0.062175	
		经典的 Classic	0.355688	
		流行的 Popular	0.256813	
		蹩脚的 Crappy	-0.21764	
		可靠的 Reliable	0.319444	

则定量地表示体现该看法的程度，得分越大表示该特点越明显，得分为负则表示更接近该特点的对立面，即正负表示该情感的极性。如对窄座总结出的特点中，有“Crappy”，意为“蹩脚的”，但该评分为负，表示该产品元素是“不”蹩脚的，是正面评价。因此该表信息应该结合“评价”和“得分”两栏一起参考。

基于该方法，若某厂商想开发运动感强，面向年轻群体的自行车产品，就可以选择“reliable”“Popular”“Impressive”“Technical sense”等产品风格，结合表 4 对应的产品类型，选择燕把、窄座、减震框架、凸面胎等元素，快速形成产品方案见图 12。

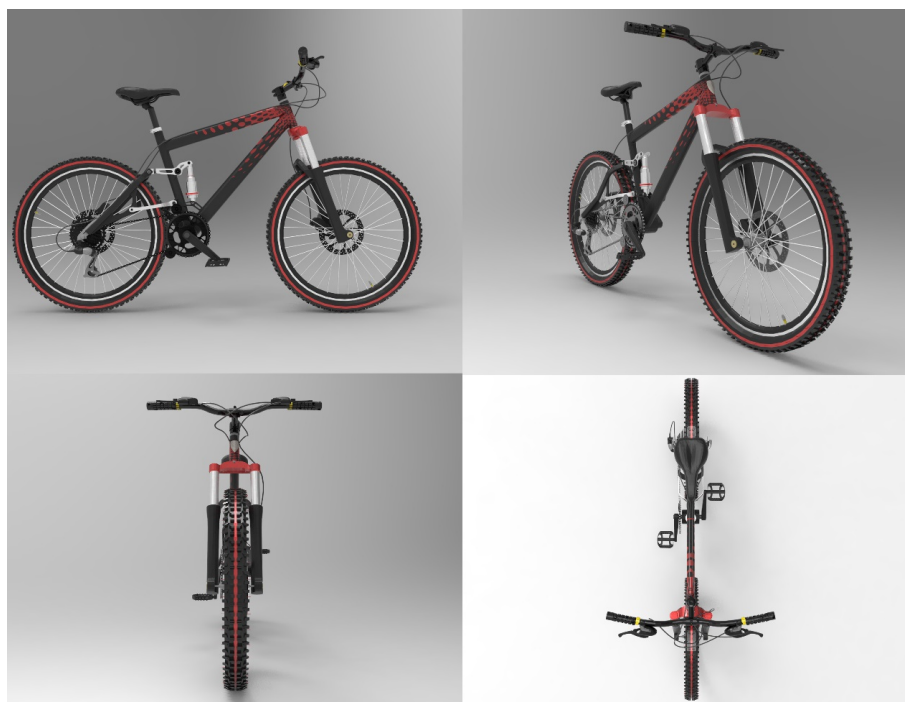


图12 产品方案
Fig.12 Product plan

5 结语

本研究给出了一种评论驱动的设计方法,获得用户对于不同产品类型的评价,并量化为情感分值。相比传统问卷调研、用户访谈等方式,该方法基于在线评论,可以帮助设计师提升创新效率,在海量数据中快速挖掘出更有代表性的设计洞见。另外,研究的主要分析过程基于自然语言处理,排除了个人情绪影响,结果更加客观理性。该方法基于数据驱动,对产品类型没有特别要求,适用性更广;操作迅捷,以调用工具包的形式完成关键任务,便于他人理解设计过程。

但是不得不承认该方法还是有不足之处,首先,该方法对所有评论内容“一视同仁”,不同评论间没有权重区分,使得分析结果“中庸化”,某些重要的评论信息可能会被淹没;其次,截取句子方式稍显武断,可能会将多余的信息参与分析,也可能忽略重要信息,影响结果准确率;最后,评论可能评价的是目标元素的多个方面(功能感受,审美感受等),该方法忽视了这一点,这也是该方法局限性之一。

由于审美和情感感知过程的复杂性,从互动的语境中获取想要的设计信息是具有挑战性的^[16],该方法的不断完善,离不开以人工智能为代表的计算机技术的持续进步和设计师将其纳入设计系统的不断尝试。

参考文献:

- [1] 席涛,周芷薇,余非石. 设计科学研究方法探讨[J]. 包装工程, 2021, 42(8): 63-78.
- [2] XI Tao, ZHOU Zhi-wei, YU Fei-shi. Design Science Research Methods[J]. Packaging Engineering, 2021, 42(8): 63-78.
- [3] GRINGRAS P. Choice of Medical Investigations for Developmental Delay: a Questionnaire Survey[J]. Child: Care, Health and Development, 1998, 24(4): 267-277.
- [4] GORTON E, STANTON S. Women's Attitudes to Urodynamics: a Questionnaire Survey[J]. British Journal of Obstetrics and Gynaecology, 1999, 106(8): 851-856.
- [5] LI Y, QIN Z, XU W R, et al. A Holistic Model of Mining Product Aspects and Associated Sentiments From Online Reviews[J]. Multimedia Tools and Applications, 2015, 74(23), 10177-10194.
- [6] 陆蔚华,姜冠岳,薛毅,等. 多模态人机数据驱动的值机体验服务设计[J]. 计算机集成制造系统, 2021, 1-18.
- [7] LU Wei-hua, JIANG Guan-yue, XUE Yi, et al. Multimodal Man-machine Data-driven Service Design for Check-in Experience[J]. Computer Integrated Manufacturing Systems, 2021, 1-18.
- [8] 张瑾,段利国,李爱萍,等. 基于注意力与门控机制相结合的细粒度情感分析[J]. 计算机科学, 2021, 48(8): 226-233.
- [9] ZHANG Jin, DUAN Li-guo, LI Ai-ping, et al. Fine-grained Sentiment Analysis Based on Combination of Attention and Gated Mechanism[J]. Computer Science, 2021, 48(8): 226-233..
- [10] ROBERT I, ANG L. Application of Data Analytics for Product Design: Sentiment Analysis of Online Product Reviews[J]. CIRP Journal of Manufacturing Science and

- Technology, 2018(23): 128-144.
- [8] WONJOON K, TAEHOON K, Ilsun R, et al. Mining Affective Experience for A Kansei Design Study on A Recliner[J]. *Applied Ergonomics*, 2019(74): 145-153.
- [9] 张国方, 寇姣姣, 陈令华. 网络评论文本驱动的汽车设计规划方法[J]. *机械设计*, 2021, 38(2): 139-144.
ZHANG Guo-fang, KOU Jiao-jiao, CHEN Ling-hua. Method of Automobile Design Planning Driven by Web Review Text[J]. *Journal of Machine Design*, 2021, 38(2): 139-144.
- [10] 王新, 乔文文. 基于云平台的用户隐式需求分析方法研究[J]. *机械设计与研究*, 2020, 36(5): 8-11.
WANG Xin, QIAO Wen-wen. Research on User Implicit Demand Analysis Based on Cloud Platform[J]. *Machine Design & Research*, 2020, 36(5): 8-11.
- [11] 纪雪, 高琦, 李先飞, 等. 考虑产品属性层次性的评论挖掘及需求获取方法[J]. *计算机集成制造系统*, 2020, 26(3): 747-759.
JI Xue, GAO Qi, LI Xian-fei, et al. Reviews Mining and Requirements Elicitation Methodology Considering Product Attributes' Hierarchy[J]. *Computer Integrated Manufacturing Systems*, 2020, 26(3): 747-759.
- [12] MARKUS H. The Modified Kansei Engineering-Based Application for Sustainable Service Design[J]. *International Journal of Industrial Ergonomics*, 2020(79): 102985.
- [13] 罗丽弦, 洪玲. 感性工学设计[M]. 北京: 清华大学出版社, 2015.
LUO LI-xian, HONG Ling. *Kansei Engineering Design* [M]. BeiJing: Tsinghua University Press, 2015.
- [14] 王飞, 戚彬, 印祥, 等. 基于感性工学的拖拉机造型设计研究[J]. *中国农机化学报*, 2020, 41(11): 83-88.
WANG Fei, QI Bin, YIN Xiang, et al. Research on Tractor Modeling Design Based on Kansei Engineering[J]. *Journal of Chinese Agricultural Mechanization*, 2020, 41(11): 83-88.
- [15] FENECH A, FRANCALANZA E, ANTHONY M A, et al. Kansei Engineering Over Multiple Product Evolution Cycles: An Integrated Approach[J]. *Procedia CIRP*, 2019(84): 76-81.
- [16] GAN Y, JI Y R, JIANG S, et al. Integrating Aesthetic and Emotional Preferences in Social Robot Design: An Affective Design Approach with Kansei Engineering and Deep Convolutional Generative Adversarial Network[J]. *International Journal of Industrial Ergonomics*, 2021(83), 103128.
- [17] CHIU M C, LIN K Z. Utilizing Text Mining and Kansei Engineering to Support Data-Driven Design Automation at Conceptual Design Stage[J]. *Advanced Engineering Informatics*, 2018(38): 826-839.
- [18] YEH C, CHEN M C. Applying Kansei Engineering and Data Mining to Design Door-To-Door Delivery Service [J]. *Computers & Industrial Engineering*, 2018(120): 401-417.
- [19] 徐悬, 刘键, 严扬, 等. 智能化设计方法的发展及其理论动向[J]. *包装工程*, 2020, 41(4): 10-19.
Xu Xuan, LIU Jian, YAN Yang, et al. Development and Theoretical Trend of Intelligent Design Methods[J]. *Packaging Engineering*, 2020, 41(4): 10-19.
- [20] 刘键, 黄赛, 曹家港, 等. 基于用户生成内容的产品服务系统设计方法[J]. *包装工程*, 2020, 41(24): 118-125.
LIU Jian, HUANG Sai, CAO Jia-gang, et al. Design Method of Product Service System Design Based on User-generated Content[J]. *Packaging Engineering*, 2020, 41(24): 118-125.
- [21] 秦磊, 王跃琦. 山地车可拆卸挡泥板创新设计[J]. *现代商贸工业*, 2010, 22(16): 383.
QIN Lei, WANG Yue-qi. Innovative Design of Detachable Fender for Mountain Bike[J]. *Modern Business Trade Industry*, 2010, 22(16): 383.
- [22] ZWICKY F. The Morphological Approach to Discovery, Invention, Research and Construction[C]. Berlin: Springer Berlin Heidelberg, 1967.