



开放科学  
(资源服务)  
标识码  
(OSID)

# 基于影评内容的中国电影海外需求与满意度研究

王善玲 李敏 韦琪

南京航空航天大学经济与管理学院 南京 211106

**摘要:** [目的/意义] 正确理解不同文化群体的观影需求和满意度有助于中国电影企业做好产品制作及出口选择。[方法/过程] 采用网络爬虫技术获取2015-2019年中国出口北美的127部电影样本影评数据,利用依存句法分析、FP-growth算法以及情感字典揭示和量化了北美观众的观影需求和情感态度,并对结果进行可视化展示和解读。[结果/结论] 研究表明北美观众中国电影观影需求表现在中国故事、人物、剧情、中国功夫、电影特效、演员表现及场景设计;其中,中国功夫、电影特效和演员表现北美口碑较高,而故事讲述、剧情结构及人物塑造北美口碑较差。因此,提高中国电影北美市场的吸引力在于提高中国故事讲述,情节设计和人物塑造能力。

**关键词:** 观影需求;情感分析;文本挖掘;依存句法

**中图分类号:** N99, F740.22, G35

## The Analysis of Overseas Audience Demand and Satisfaction of China Movies Based on Movie Reviews

WANG Shanling LI Min WEI Qi

School of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

**Abstract:** [Objective/ Significance] Precisely capturing and correctly understanding the demand and satisfaction of audience with different cultural backgrounds will help movie producers and distributors perform well in movie-making and export choice. [Methods/process] this paper uncovers demand and satisfaction of moviegoers coming from North America by investigating their online reviews. From the foremost online review communities, moviegoer reviews have been retrieved, quantified, text-

**基金项目** 国家社会科学基金项目“泛媒体网络环境下中国电影跨文化绩效提升途径研究”(16BXW038)。

**作者简介** 王善玲(1996-), 研究生, 研究方向为文本挖掘, 电影跨文化传播, E-mail: shan\_ling\_wang@163.com; 李敏(1964-), 博士, 副教授, 研究方向为电影跨文化传播; 韦琪(1995-), 研究生, 研究方向为组织行为和人力资源管理。

mined, and interpreted using dependency parsing, FP-growth, and sentiment analysis. [Results/conclusions] Results demonstrate that North American audiences are less satisfied with the story, characters and plots of Chinese movies, and are more strongly fascinated by the actors, special effects and Chinese kung fu. Therefore, improving the appeal of Chinese movies in the North America lies in improving capability of Chinese storytelling, plot design and character building.

**Keywords:** Moviegoing demand; sentiment analysis; text mining; dependency parsing

## 引言

电影票房成功取决于电影属性与观众需求偏好的匹配程度,观众需求偏好是电影产品制作和营销开展的基础。然而,由于每部电影的独特性及短生命周期的特点使得其用户需求难以把握。对于外来电影,跨文化背景无疑会增加观众需求分析的复杂性<sup>[1]</sup>。海外消费者不同的价值观和思维方式及电影类短生命周期的特点使得消费需求变得多变、复杂<sup>[2]</sup>。

近年来,关于电影需求的研究通常采用计量模型,选取可量化的电影特征(如体裁,电影制式等)为主<sup>[1,3,4]</sup>,而那些同样会传达观众需求,影响票房表现的难以量化的电影特征(如,情节、人物等)往往被忽略<sup>[5]</sup>。基于问卷、访谈等方法<sup>[6]</sup>一定程度上弥补了上述方法不足,但问卷依赖于事先确定的特征变量,无法对消费者感知到的电影特征进行评估<sup>[7]</sup>,且存在样本规模较小<sup>[8]</sup>或通用性不强<sup>[9]</sup>等局限性。随着诸如IMdb、豆瓣等影评网站的建立,极大地方便了观众的意见表达。其提供的准确、及时且低成本、不受地域限制的信息反馈<sup>[10]</sup>成为外来企业分析观众需求的重要信息来源<sup>[11]</sup>。其中的贬褒评论隐含了观众需求偏好和情感态度<sup>[10,12]</sup>,但由于非结构化的特点使得其在电影研究中常常被忽略。鉴于此,基于当前电影跨文化研究的

不足,及隐藏在影评中大量的需求信息,本研究旨在通过考察不同文化背景观众的在线影评揭示他们对中国电影的观影需求和情感态度,为中国电影的海外传播及营销开展提供理论借鉴与管理启示。

## 1 理论基础

### 1.1 在线评论产品特征提取

用户需求获取的关键在于从在线评论中提取产品特征。产品特征通常由两类组成:一类是制造商提供的产品说明,往往不包含任何关于特征的具体信息(如产品设计的质量、易用性等);第二类特征则是消费者在使用或体验后表达观点的对象<sup>[13]</sup>。产品特征提取就是在既有的表达观点的文本中进行特征识别。

关于特征提取方法学者们展开了大量研究。主要分为基于统计的算法及基于句法分析的算法。最朴素的基于统计的算法是通过频率统计识别特征词,再采用邻近原则判断相应观点。例如,Liu<sup>[14]</sup>首次利用Apriori算法从自然语言中进行产品特征识别,但该方法仅限于检测与单个名词有强烈关联的特征且只考虑与邻近特征词搭配的形容词作为观点表达。Popescu和Etzioni<sup>[15]</sup>对其进行改进,使用候选特征和产品

类别之间的共现来提取产品特征,并使用点互信息表示关联程度。Hu 等<sup>[16]</sup>考虑文本特征的相似性,使用聚类算法对文本特征进行聚类,并使用类的中心作为产品特征。

基于句法分析的算法则通过建立语法规则、语言模式、或利用情感词典来挖掘产品特征及其相关观点。如 Zhuang 等<sup>[17]</sup>利用句法依存模板提取电影评论特征词与相应观点词之间的显式和隐式关系,Zhang 等<sup>[18]</sup>根据句子间的依存关系,结合情感字典定义一系列的提取规则。近年来,也有部分学者将统计与句法分析结合进行产品特征与相关观点挖掘。Poria 等<sup>[19]</sup>结合卷积神经网络和语言规则设计的集成分类器和词嵌入模型,在对笔记本电脑与餐厅评论文本特征提取任务中取得了很好的准确率。彭云<sup>[20]</sup>从中文商品评论的语言结构和特点出发,将语义关系知识嵌入到 LDA 模型,提出了适用于细粒度产品特征词和情感词提取的 SRC-LDA 模型。

关于如何从自然语言中提取特征词的方法,目前学术界尚未达成共识。学者们提出的算法各有其优点与不足,适用对象也不尽相同。基于统计的方法需要大量的向量特征,基于语法依存关系的方法则需要手工定义抽取规则,且很大程度上依赖于句子的语法准确性。根据本文的研究目的,拟结合基于句法分析与基于统计的算法对影评内容的特征观点对进行识别。

## 1.2 情感分析

情感分析又称意见挖掘。当前大部分关于情感分析的研究都是基于情感词典展开。情感词典是标注了情感倾向的词或短语的集合,基

于情感词典的方法可以挖掘出文本中的情感倾向<sup>[21]</sup>,在语义导向的情感分类中具有重要作用。近年来,情感词典被广泛应用于酒店、政治、电影、旅游等领域的情感分析任务<sup>[22-24]</sup>。

General Inquiry (GI)是最古老的人工情感词典之一,该词典诞生于1966年后便被不断改进。GI词典收录了约11789个词语,每个词按照极性、强度、词性等被打上不同标签,被广泛应用于社会学、心理学、经济学、人类学等领域。但是GI词典忽略了词语的情感强度,因此并不适用于细粒度的情感分析任务。为了解决在线评论中常见的新词语、符号及拼写错误等问题,Thelwall 等<sup>[25]</sup>进行手工编码,创建了 SentiStrength 情感词典定义了社交网络在线评论中的情感倾向。Severyn 等<sup>[26]</sup>提出了一种浅层的句法结构,并使用 SentiStrength 词典对 YouTube 上的意大利语在线评论进行情感分类。手工创建情感词典成本过高且主观性强,为了克服该方法的局限,词汇扩展法被提出。Liu<sup>[14]</sup>利用30个种子词在 WordNet 上建立了一个基于同义词和反义词关系的情感词汇库,编辑了大约6800个情感词。Khan 等<sup>[27]</sup>采用 SentiWordNet 和互信息算法创建了情感词典 SentiMI。在对康奈尔电影数据集情感分类任务里,准确率提高了7%。为了适用不同领域研究,领域词典逐步被开发。如 Zhang 等<sup>[24]</sup>基于大连理工大学开发的汉语情感词汇本体库,创建完整的微博情感词库。Nie 等<sup>[23]</sup>基于 HowNet 词典,采用 Word2Vec 算法创建了适用于酒店在线评论情感分析的语义分区情感词典。但是,目前还没有特别比较权威地适用于电影领域的情感词典。因此,本文借用 Khan 等创建的通用情感词典

SentiMI 进行影评内容的情感分析。

## 2 研究框架与研究方法

基于影评内容挖掘的观众原创内容分析是

探索观众需求和情感态度的有效途径。掌握跨文化观众的观影需求偏好和情感倾向,对中国电影产品出口选择与营销开展具有一定的指导意义。本文构建的基于影评内容的需求识别与情感分析模型,如图1所示。

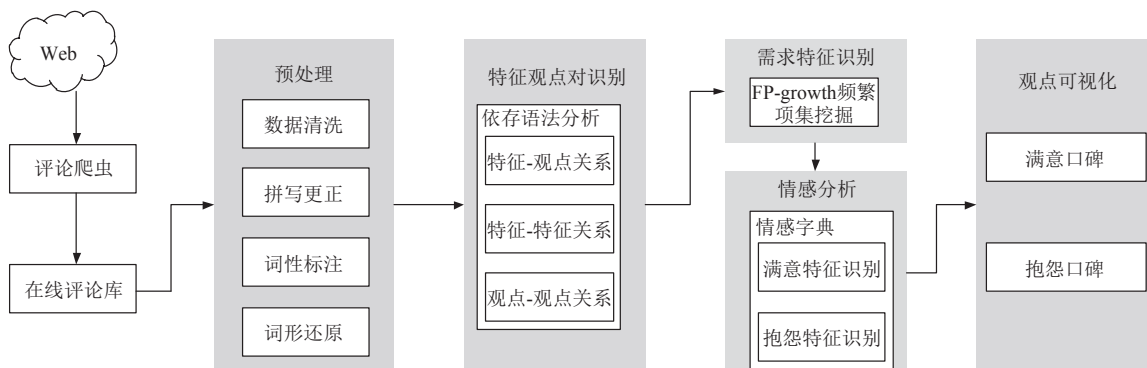


图1 基于影评内容的需求识别与情感分析模型

该模型中,首先通过网络爬虫从影评网站收集评论数据构建在线评论库;其次对收集到的数据进行清洗过滤,拼写更正,词性标注及词形还原等一系列预处理;再次基于依存句法关系进行特征观点对抽取,并利用 FP-growth 算法识别产品需求特征;从次基于情感字典对识别到的特征观点对进行情感分析;最后对得到的观点词汇归类并进行可视化。

### 2.1 数据采集及预处理

北美作为全球最大的电影市场,北美观众为全球电影产业做出了重要贡献。因此,本研究利用 python 编程从北美地区最知名影评网站之一 IMDb 采集 2015-2019 年中国出口北美 142 部电影的所有影评文本,剔除评论为 0 的影片,最终得到 127 部电影共 4959 条在线评论数据。获取数据字段包括用户名、评论日期、评论内容及在线评分等。由于电影产品属于短生命周期体验品,销售量波动极大导致需求特征难以

挖掘<sup>[28]</sup>。因此本文按照票房销售总额将出口北美的中国电影分为高、中、低三个区间进行分析,其中高票房指百万至千万美元票房区间共 18 部影片,合计 2854 条影评数据;中票房指十万至百万美元票房区间,共 69 部影片,1399 条影评数据;低票房指票房总额在十万美元以下区间,共 40 部影片,706 条影评数据。为提高数据质量以进行更准确的分析,首先需要对收集到的评论文本进行预处理。经过清洗,拼写更正、词性标注及词形还原等一系列的预处理工作后,进行特征观点对抽取并识别需求特征。

### 2.2 影评特征观点对提取及需求特征识别

本文主要借鉴文献[18]提出的联合观点挖掘算法进行文本挖掘工作,利用依存语法关系识别并确认观点词和特征词之间的内在规律以定义抽取规则,并根据情感词典抽取三元组信息:特征-特征关系、观点-观点关系及特征-观点关系。依存语法(dependency grammar)是

指根据词与词之间的直接依存关系来描述语言结构的框架，核心在于假设一个句子的层次结

构主要是由词与词之间二元不对称的依存关系构成<sup>[29]</sup>。

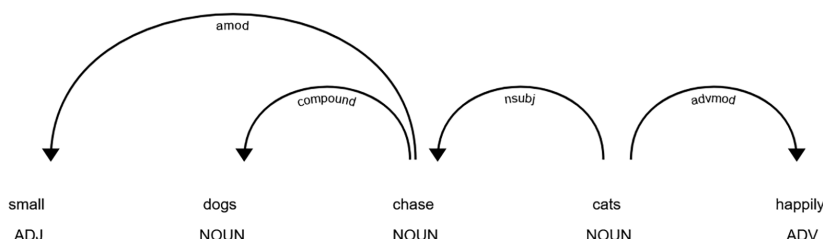


图2 句子 <small dogs chase cats happily> 的依存句法树

如图2展示了句子 < small dogs chase cats happily > 经过 spaCy 依存句法分析后的结果。其中名词 dogs 充当动词 chase 的主语 (compound)；而名词 cats 充当直接宾语 (nsubj)。形容词 small 是狗的定语修饰 (amod)；副词 happy 是 chase 的状语修饰 (advmod)。

为方便起见，首先定义关系符：FO 表示特

征 - 观点关系，OO 表示观点 - 观点关系，FF 表示特征 - 特征关系。本文基于情感词典 SentiMI 定义了基本提取任务进行信息提取：(1) 利用观点词提取产品特征词 (FO)；(2) 利用提取的特征词提取特征词 (FF)；(3) 基于已知观点词提取观点词 (OO)。基于影评内容定义了本文的依存关系抽取规则，具体如下 (表1) 所示。

表1 影评文本特征观点抽取规则

| 规则     | 输入 | 抽取公式                                                                 | 输出  | 例子                                                                                       |
|--------|----|----------------------------------------------------------------------|-----|------------------------------------------------------------------------------------------|
| Rule11 | O  | O=>F;<br>st.O ∈ {O}, O-dep ∈ {MR};<br>POS F ∈ {NN, NNS}              | f=F | It is such a spectacular movie.<br>(spectacular→amod→movie)                              |
|        |    |                                                                      | FO  | I really enjoy the movie.<br>(enjoy→dobj→movie)                                          |
|        |    |                                                                      |     | The story is interesting<br>(interesting←nsubj←story)                                    |
| Rule12 | O  | O=>H=>F;<br>st.O ∈ {O}, O/F-dep ∈ {MR};<br>POS F ∈ {NN, NNS}         | f=F | It just felt so juvenile, and distract from the story.                                   |
| Rule21 | F  | F=>F;<br>st.F ∈ {O}, F-dep ∈ {conj};<br>POS F ∈ {NN, NNS}            | f=F | The landscape and effect are bright and beautiful<br>(landscape→conj→effect)             |
|        |    |                                                                      | FF  |                                                                                          |
| Rule22 | F  | F=>F;<br>st.F ∈ {O}, F-dep ∈ {compound};<br>POS F ∈ {NN, NNS}        | f=F | Initial character design is some weird                                                   |
| Rule23 | F  | F=>H=>F;<br>st.F ∈ {O}, F-dep ∈ {pobj};<br>POS F ∈ {NN, NNS}         | f=F | I hope the translation can get the essence of the story.<br>(essence←prep←of←pobj←story) |
|        |    |                                                                      | FF  |                                                                                          |
| Rule31 | O  | O=>O;<br>St.O ∈ {O}, O-dep ∈ {advmod, conj};<br>POS O ∈ {RB, JJ, VB} | o=O | I love it and strongly recommend<br>(strongly→advmod→recommend)                          |
|        |    |                                                                      | OO  | (love→conj→recommend)                                                                    |
|        |    |                                                                      |     | The landscape and effect are bright and beautiful<br>(bright→conj→beautiful)             |

注：表1中，o (或 f) 表示得到的观点词 (或特征词)。O (或 F) 是给定的或得到的已知观点词 (或特征词) 的集合。{NN, NNS, RB, JJ, VB} 是词性标签。O-dep 表示观点词 O 通过 O-dep 关系依赖于第二个词。F-dep 指特征词 F 通过 F-dep 关系依赖于第二个词。MR={amod, dobj, nsubj, pobj}。



用户对某个话题越是在意,则越有可能在评论中提及该话题<sup>[30,31]</sup>,评论中出现的高频产品特征及其主流观点反映了用户对产品功能实现的评价和对产品功能期望的设想。因此,在获得了电影产品特征词以后,利用 FP-Growth 算法挖掘产品特征词中的频繁项集。FP-growth 算法挖掘产品特征频繁集的具体过程如下:

(1) 扫描数据集 F, 设定 minsup, 筛选出频繁 1- 项集并降序排列;

(2) 再次扫描数据集 F, 将每个事务中的元素按照频繁 1- 项集顺序进行重新排序, 形成事务数据库 D; 根据事务数据库 D 创建项头表及 FP-tree 树;

(3) 遍历项头表, 按照从下往上的顺序从 FP-tree 中找出满足 minsup 的频繁项集。

### 2.3 情感值量化及可视化

根据上文获取的特征观点词, 利用情感字典统计观众对需求特征的情感倾向, 实现观众情感的量化分析。根据情感分析结果将需求特征的观点词分为积极与消极两大类。并采取手

工方式对某些观点词的情感极性进行修正, 如 <predictable> 通常为积极词性, 而在电影领域则为消极词性。之后按照观点词频次排序, 取出频次最高的前 100 观点词聚类并进行可视化展示, 从而得到需求特征积极与消极口碑的具体原因, 有助于电影企业从观众角度识别需求特征具体优劣, 更好地进行产品改进以满足观众需求。

## 3 研究结果分析

基于上述定义的影评特征观点对抽取规则及情感字典 SentiMI, 本研究从影评内容中共获取特征观点对 74589 对, 其中高票房特征观点对 34751 对, 中票房特征观点对 26130 对, 低票房特征观点对 13708 对, 电影评论特征观点对结构化数据形式见表 2。将挖掘到的候选特征词作为 FP-growth 算法的输入, 挖掘观众对电影需求特征的重视程度; 将需求特征和观众情感汇聚, 利用情感字典计算不同需求特征的观众情感倾向, 实现观众评论细粒度的情感分析。

表 2 影评文本特征观点对结构化数据 (部分)

| 列 | 副词   | 观点词         | 候选特征词     | 情感值       | 情感极性 | 评论语句 |
|---|------|-------------|-----------|-----------|------|------|
| 1 | very | predictable | plot      | -4.040091 | 消极   | 1    |
| 2 | not  | like        | story     | -2.007    | 消极   | 1    |
| 3 |      | old         | action    | -0.994    | 消极   | 3    |
| 4 |      | great       | actor     | 2.007     | 积极   | 18   |
| 5 |      | weird       | character | -2.41     | 消极   | 5    |

### 3.1 中国电影海外观众需求感知分析

将挖掘到的候选特征词作为 FP-growth 算法的输入, 经过多次调试和对结果的观察, 最

终确定了高票房 (259)、中票房 (230) 及低票房 (111) 特征词的最小支持度。如表 3 展示了各票房区间在最小支持度下挖掘到的需求特征词。

表 3 各票房区间需求特征及支持度

| 高票房       |      | 中票房       |      | 低票房       |     |
|-----------|------|-----------|------|-----------|-----|
| 频繁项       | 支持度  | 频繁项       | 支持度  | 频繁项       | 支持度 |
| movie     | 1653 | movie     | 1237 | movie     | 576 |
| film      | 1098 | film      | 1263 | film      | 498 |
| story     | 631  | story     | 623  | story     | 342 |
| character | 544  | scene     | 562  | character | 294 |
| scene     | 430  | character | 471  | scene     | 236 |
| action    | 375  | action    | 454  | action    | 201 |
| effect    | 345  | art       | 347  | art       | 115 |
| plot      | 292  | actor     | 280  | actor     | 111 |
|           |      | effect    | 254  |           |     |
|           |      | plot      | 234  |           |     |

从表 3 可以看出,不同票房区间下,海外观众对中国电影的需求特征基本一致,主要体现在整体评价电影(film、movie)及关于电影故事(story)、人物(character)、场面(scene)、动作(action)。不同的是,仅有高票房和中票房区间的海外观众提到了电影特效(effect)和电影情节(plot),而中票房和低票房区间影片则提到电影艺术(art)和演员(actor)。这种不同可能源于高票房区间的影片大多以科幻(如《流浪地球》)、动画(《哪吒》)或大制作的动作片(如《长城》、《战狼 2》)为主,而低票房区间则主要为低成本的爱情及喜剧类型的小电影。

由于整体评价只能得知观众对于电影的整体印象无法挖掘出观众的需求特征,过滤观众对于电影的整体评论对象 film、movie,得出本文结论一:北美观众对中国电影需求感知表现在电影故事(story)、人物(character)、场面(scene)、动作(action)、特效(effect)、情节(plot),艺术(art)和演员(actor)。因

此,本文的第一个启示在于,中国电影企业在确定出口影片时,可以考虑从这 8 个需求特征维度进行综合评估,选择合适影片,以吸引更多海外观众。

### 3.2 中国电影海外观众需求满意度分析

根据情感字典和电影评论特征观点对结构化数据可以挖掘文本中的情感倾向,量化观众感知需求的情感态度,识别观众满意度。本文首先分析了整体票房区间,对情感倾向进行全局把握。图 3 展示的是 2015-2019 年 127 部电影特征词积极词汇、消极词汇情感值绝对值。可以看出,art、effect、actor 及 action 的积极情感值明显高于 story、character 及 plot,前者的消极情感值也明显低于后者。从数据基本可以看出,中国电影的 art、actor、effect 以及 action 在北美市场观众口碑较好,观众满意度较高,而 story、character 以及 plot 观众口碑不佳,观众满意度较低。

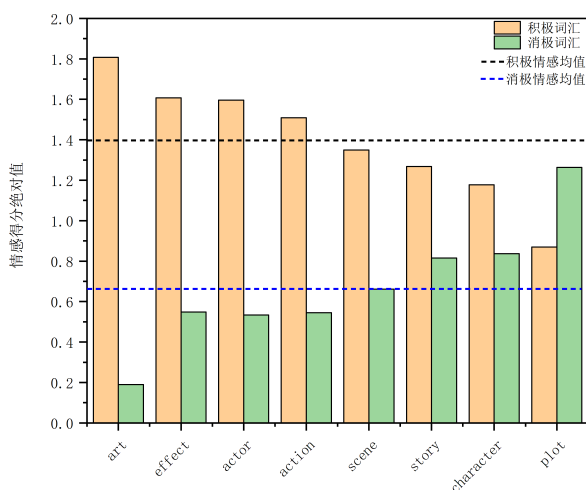


图 3 2015-2019 年中国电影北美观众需求特征情感得分

图 4 展示了 18 部高票房影片观众需求的情感值。可以看出高票房区间影评中, effect、

art、action、actor 及 scene 积极情感值高于平均水平, 消极情感值低于平均水平; plot、character、story 的积极情感值明显低于平均水平, 消极情感值明显高于平均水平。说明 effect、art、action、actor 及 scene 在北美口碑较好, 满意度较高; 而 plot、character、story 口碑一般, 满意度较低。

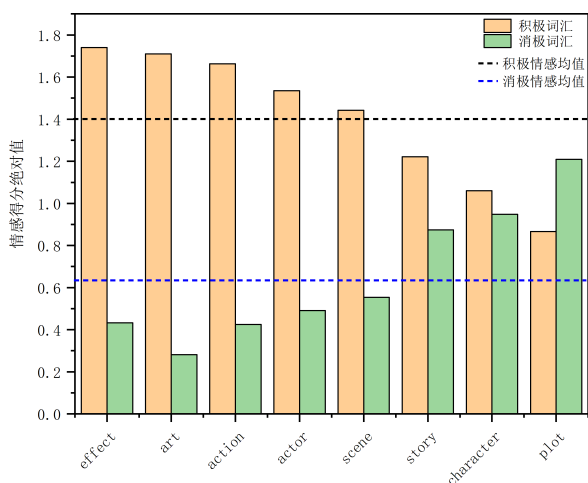


图 4 2015-2019 年高票房区间需求特征情感得分

图 5 显示了中票房区间积极词汇、消极词汇情感值的绝对值。从图 5 可以看出 art、actor、action 及 effect 积极情感值高于 scene、story、character 及 plot, 前者的消极情感值也明显低于后者; 而 scene、story、character 及 plot 积极情感值低于平均水平, 消极情感值高于平均水平。

图 6 显示的是低票房区间显著需求特征的情感得分。由图 6 可以看出, 低票房区间 art 积极情感值最高, 消极情感值最低; plot 消极情感值最低, 积极情感值最高。其他特征词无论是积极、消极还是整体情感得分, 基本一致较低。这也可以进一步解释票房不佳是由于观众各方面需求得不到满足。

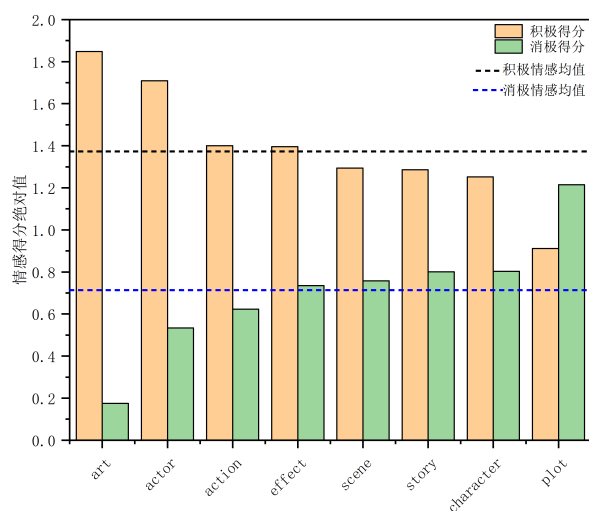


图 5 2015-2019 年中票房区间需求特征情感得分

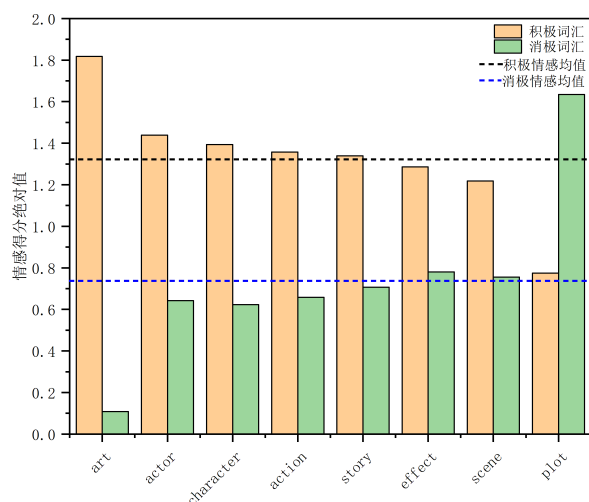


图 6 2015-2019 年低票房区间需求特征情感得分

根据上述分析, 得到各票房区间观众需求特征的满意程度, 获得电影产品的满意特征与抱怨特征 (见表 4)。对比高票房、中票房、低票房及全部票房区间, 中国电影在艺术 (art)、演员 (actor)、动作 (action) 及特效 (effect) 方面海外观众口碑基本一致, 观众需求得到满足属于满意特征。而故事 (story)、人物塑造 (character) 及情节设计 (plot) 各票房区间表现不佳, 海外观众需求无法满足属于抱怨特征。因此, 中国电影的艺术、演员、特效及动作这

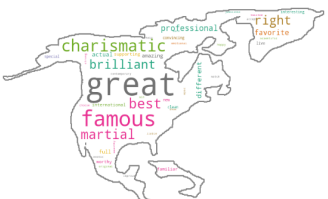


些特征可以作为海外营销的优势和亮点，以此提高中国电影海外市场的吸引力。中国故事讲述，情节设计和人物塑造能力属于劣势，电影制作者应着重对其改进。而场景设计观众满意度一般，说明该需求特征是中国电影企业努力提升的特征。

表 4 2015-2019 年各票房区间观众满意特征与抱怨特征

| 票房区间 | 满意特征        | 抱怨特征                 |
|------|-------------|----------------------|
| 全部票房 | 艺术、演员、特效、动作 | 故事、人物、情节             |
| 高票房  | 特效、艺术、动作、演员 | 故事、人物、情节             |
| 中票房  | 艺术、演员、动作、特效 | 场景、故事、人物、情节          |
| 低票房  | 艺术          | 演员、人物、动作、故事、特效、场景、情节 |

表 5 满意特征情感可视化分析

| 满意特征 | 积极口碑                                                                                                                                                  | 词云可视化                                                                                 |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| 演员   | (1) 阵容强大 (famous, favorite) ;<br>(2) 演技专业 (professional、brilliant、right、great、best) ;<br>(3) 富有魅力 (charismatic) ;<br>(4) 打斗功夫 (martial) ;             |  |
| 艺术   | (1) 武术 (martial art) ;<br>(2) 刺激有趣 (great、enjoy、interesting)                                                                                          |  |
| 动作   | (1) 功夫 (great、love、martial、best) ;<br>(2) 刺激有趣 (entertaining、intense、live) ;<br>(3) 大制作 (enough、big)                                                  |  |
| 特效   | (1) 制作精良 (great、spectacular、beautiful) ;<br>(2) 技术高超 (amazing、awesome、enjoy、excellent) ;<br>(3) 逼真 (realistic、decent) ;<br>(4) 独创 (original、unique) 。 |  |

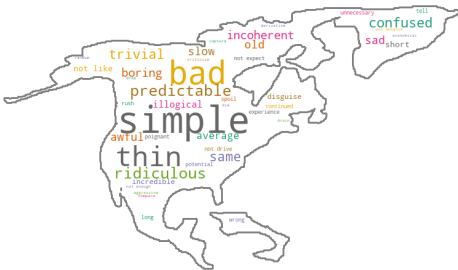
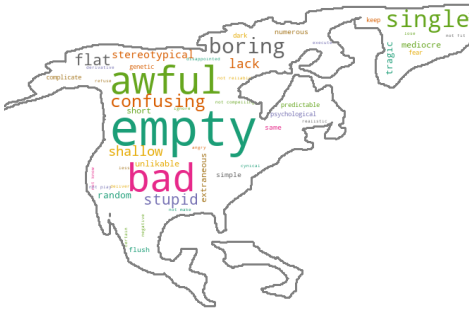
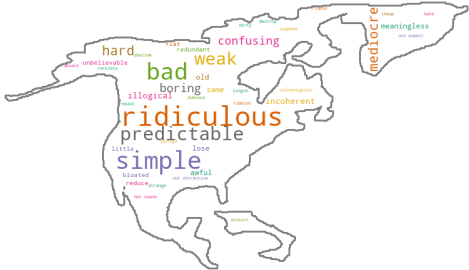
### 3.3 电影需求特征口碑可视化及分析

观点词是观众对于需求特征的意见表达和信息反馈，隐含了需求特征的优点与不足。根据观点词将电影产品的满意特征与抱怨进一步分解成各类维度，利用词云可视化特征的具体细节。表 5 展示了近年来北美观众满意特征的具体表现。通过词云可视化可以看出，中国电影演员方面的积极口碑主要在于演员阵容、表演专业，富有魅力及打斗功夫；而艺术呈现和动作设计方面更多描述的是中国功夫刺激有趣及大场面；制作精良，技术水平高超，逼真以及创新性的电影特效也是获得观众满意的重要原因。

表 6 展示了中国电影故事、人物以及情节在北美市场获得消极口碑的具体原因。从对应的词云图可以看出中国电影的故事劣势具体表现在内容空洞, 缺乏新意及逻辑不当。人物塑

造方面的劣势在于人物肤浅无趣, 形象单一刻板, 人物表现迷惑不解。而设计荒诞、缺乏新意、逻辑不当的情节也是中国电影北美失意的重要原因。

表 6 抱怨特征情感可视化分析

| 抱怨特征 | 消极口碑                                                                                                                                                               | 词云可视化                                                                                |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| 故事   | <p>(1) 内容空洞 (simple、thin、trivial、boring、bad);</p> <p>(2) 缺乏新意 (same、predictable);</p> <p>(3) 逻辑不当 (ridiculous、confused、incoherent);</p>                            |    |
| 人物   | <p>(1) 肤浅无趣 (empty、shallow、stupid、boring);</p> <p>(2) 单一刻板 (single、flat、stereotypical);</p> <p>(3) 迷惑不解 (confusing);</p>                                           |   |
| 情节   | <p>(1) 剧情荒诞 (ridiculous);</p> <p>(2) 缺乏新意 (simple、predictable mediocre、boring);</p> <p>(3) 迷惑不解 (weak、hard、confusing);</p> <p>(4) 逻辑不当 (illogical、incoherent);</p> |  |

综上所述, 强大的演员阵容、精彩的打斗功夫以及制作精良的电影特效是中国电影获取海外吸引力的重要原因, 而中国电影的故事讲述混乱、情节设计缺乏新意及人物形象单一刻板等都是饱受海外观众诟病的原因。因此, 通过对观点词可视化, 有利于中国电影企业从观

众角度识别需求特征的具体优劣, 明确电影产品的保持和改进方向, 更好地满足观众需求。

## 4 结论及启示

本文采用文本挖掘技术对 2015-2019 年中

国出口北美影片的在线影评进行需求挖掘,通过细粒度的情感分析深层次地挖掘观众感知需求与感知价值,主要得出如下结论:

(1) 北美观众对中国电影需求感知表现在演员、特效及中国功夫,这与已有研究一致<sup>[1, 32, 33]</sup>。本文新发现在于,难以量化的故事、情节、人物及场景也是北美观众可感知的重要需求。新奇有趣的故事、逻辑缜密的情节等特征是北美观众的诉求所在,这是以往定量模型无法分析的。因为一部电影通常被细分为40~60个由人物对白、动作等构成的场景,所有场景切换推动情节发展,这是一个难以量化的过程。基于影评内容的文本挖掘可以直接获得观众对于电影产品本身的感知需求。因此,本文的研究方法有助于全面评估观众对于电影产品的需求感知。

(2) 北美观众对中国电影的故事、人物及情节满意度较低。而演员、特效及中国功夫口碑满意度较高,对北美观众吸引力较强。

电影本质上是文化产品,根植于特定的民族文化背景中。观影者与创作者之间的文化距离会使得电影类体验品在国外推广时面临着“文化折扣”的挑战,即电影在跨越文化界限时的价值损失。由于北美观众与中国电影创作者之间缺乏共同的隐性背景、思维方式及经验积累,使得带有异域文化元素的故事内容、行为逻辑等难以被北美观众充分欣赏。因此,大多北美观众无法理解中国电影的在故事讲述、情节设计及人物塑造方面的行为逻辑,只能表示迷惑不解和逻辑不当。

但这种文化折扣并不是无法克服。本文发现,具有高度文化异质性的演员及中国功夫也

获得了较高满意度。通过对演员相关观点词聚类发现,北美观众认为中国电影演员阵容强大、表演专业、富有魅力,这些因素共同克服了演员特征在跨越文化界限时的价值损失,从而获得好评。此外,已有研究表明,当国家间的文化距离足够大时,产品中内在的文化元素与鲜明的民族风格会激发消费者的好奇心带来文化溢价<sup>[34]</sup>。正如北美观众对中国功夫的评价:刺激有趣。中国功夫这一独特的东方文化景观一定程度上满足海外观众对中华文化的好奇心。

本文的研究发现为中国电影产品的跨国营销提供了有益的管理启示:

(1) 文化差异导致当地观众对外来电影的反应不同于本土产品,因此为目标受众选择正确的电影产品至关重要。进军国际的电影企业可以开发自己的需求挖掘模型,挖掘目标受众的显著需求及感知价值,制作、选择合适的电影产品在外发行,迎合当地观众的需求偏好,提升海外观众电影接受水平。例如,对于中国电影分销商而言,应发挥明星演员、电影特效及中国功夫在吸引北美观众接受方面的优势。具体措施包括邀请北美观众熟悉的国际影星参与电影演出,选择制作优良的功夫类体裁影片或具有高超特效技术的科幻类体裁影片发行北美等。

(2) 电影消费的根本是内容消费。作为艺术媒介,中国电影首先应该明确如何在有限的空间和时间中提高中国故事讲述,情节设计和人物塑造能力。电影制片方可尝试将中国叙事、普世价值与具有本国鲜明特色的,如音乐、服装等文化元素巧妙融合,以增强电影美学的特异性和吸引力。此外,国际合拍可能会有效减

少电影产品的文化折扣,统计发现,近五年内中国电影北美票房超千万美元影片均为中美合拍。因为合拍电影通常以国际视角进行艺术阐释,运用“境外人士讲中国故事”的方式,让中国文化得到越来越多海外观众的认可及认同。

(3)通过社交媒介、影评网站等对电影发行、放映过程中的专家、观众发布的有利于提升消费者购买意愿的影评内容进行推送。同时既要重视和广泛收集海外观众的负面反馈,又要能及时对症下药,回应并解决海外观众的问题与不满,形成不断改善电影产品及绩效的良性途径。

本文还存在如下不足:首先,本文仅选择了自愿在网上发布评论的观众作为研究样本,而那些不太活跃于网络的观众被忽视了。因此不能确定所研究对象是否能完全代表北美观众。未来研究可进一步结合线上和线下观众的意见进行分析。其次,本文仅研究了北美观众对中国电影的感知需求和感知价值,并未研究中国观众对中国电影的感知需求,因此无法确定中国电影在北美满意特征与抱怨特征是由于国家间的文化差异形成还是电影本身形成,下一步可进行中国电影跨文化需求特征的对比研究。

## 参 考 文 献

- [1] Elliott C, Konar A P, LING H, et al. Behind film performance in China's changing institutional context:The impact of signals[J]. Asia Pacific Journal of Management. 2017, 35(1):63-95.
- [2] Brooks S. Does personal social media usage affect efficiency and well-being?[J]. Computers in Human Behavior, 2015(46):26-37.
- [3] Moon S, Song R. The Roles of Cultural Elements in International Retailing of Cultural Products:An Application to the Motion Picture Industry[J]. Journal of Retailing, 2015, 91(1):154-170.
- [4] 张律疏,李敏,张捷.在线评分对中国电影北美票房的影响研究[J].电影艺术,2017(5):155-160.
- [5] Chung J, Rao V R. A General Consumer Preference Model for Experience Products: Application to Internet Recommendation Services[J]. Journal of marketing research, 2012, 49(3):289-305.
- [6] Lee J, Chen C, Song H, et al. Consumption of Movie Experience:Cognitive and Affective Approaches[J]. Journal of Quality Assurance in Hospitality & Tourism, 2017, 18(2):173-199.
- [7] 漆谦.国外受众对中国电影的需求与期待[J].电影艺术,2013(6):50-54.
- [8] Zeng H, Hao L. Cross-cultural examination of the effects of promotional framing on consumers' responses:A comparison of China and Pakistan[J]. International Business Review, 2016, 25(5):1020-1029.
- [9] Tang L. Mine your Customers or Mine your Business:The Moderating Role of Culture in Online Word-of-Mouth Reviews[J]. Journal of International Marketing, 2017, 25(2):88-110.
- [10] Berger J, Humphreys A, Ludwig S, et al. Uniting the Tribes:Using Text for Marketing Insight[J]. Journal of Marketing, 2020, 84(1):1-25.
- [11] Guo W, Liang R, Wang L, et al. Exploring sustained participation in firm-hosted communities in China:the effects of social capital and active degree[J]. Behaviour & Information Technology, 2017, 36(3):223-242.
- [12] 赵宇晴,阮平南,刘晓燕,等.基于在线评论的用户满意度评价研究[J].管理评论,2020,32(3):179-189.
- [13] Archak N, Ghose A, Ipeirotis P G. Deriving the Pricing Power of Product Features by Mining Consumer Reviews[J]. Management Science, 2011, 57(8):1485-1509.
- [14] Liu M H A B. Mining and Summarizing Customer Reviews[C]. In Proceedings of ACM-KDD, 2004:168-177.
- [15] Popescu A, Etzioni O. Extracting Product Features

- and Opinions from Reviews[M]. London: Springer London, 2007:9-28.
- [16] Hu Y, Chen Y, Chou H. Opinion mining from online hotel reviews—A text summarization approach[J]. Information Processing & Management, 2017, 53(2):436-449.
- [17] Zhuang L, Jing F, Zhu X. Movie review mining and summarization[C]. ACM, 2006.
- [18] Zhang H, Sekhari A, Ouzrout Y, et al. Jointly identifying opinion mining elements and fuzzy measurement of opinion intensity to analyze product features[J]. Engineering Applications of Artificial Intelligence, 2016(47):122-139.
- [19] Poria S, Cambria E, Gelbukh A. Aspect extraction for opinion mining with a deep convolutional neural network[J]. Knowledge-Based Systems, 2016(108):42-49.
- [20] 彭云, 万常选, 江腾蛟, 等. 基于语义约束 LDA 的商品特征和情感词提取 [J]. 软件学报, 2017, 28(3):676-693.
- [21] Singh S K, Sachan M K. SentiVerb system: classification of social media text using sentiment analysis[J]. Multimedia Tools and Applications, 2019, 78(22):32109-32136.
- [22] Song C, Wang X, Cheng P, et al. SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis[J]. Knowledge-Based Systems, 2020(194):105572.
- [23] Nie R, Tian Z, Wang J, et al. Hotel selection driven by online textual reviews: Applying a semantic partitioned sentiment dictionary and evidence theory[J]. International Journal of Hospitality Management, 2020(88):102495.
- [24] Zhang W, Wang M, Zhu Y. Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? From the perspective of cognitive big data analytics[J]. International Journal of Information Management, 2020(50):498-514.
- [25] Thelwall M, Buckley K, Paltoglou G. Sentiment strength detection for the social web[J]. Journal of the American Society for Information Science and Technology, 2012, 63(1):163-173.
- [26] Severyn A, Moschitti A, Uryupina O, et al. Multilingual opinion mining on YouTube[J]. Information Processing & Management, 2016, 52(1):46-60.
- [27] Khan F H, Qamar U, Bashir S. SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection[J]. Applied Soft Computing, 2016(39):140-153.
- [28] 唐中君, 崔骏夫, 唐孝文, 等. 融合内容分析和关联分析的短生命周期体验品需求特征模式挖掘方法研究 [J]. 中国管理科学, 2019, 27(11):166-175.
- [29] De marneffe M, Nivre J. Dependency Grammar[J]. Annual Review of Linguistics, 2019, 5(1):197-218.
- [30] Xiao L. Analyzing consumer online group buying motivations: An interpretive structural modeling approach[J]. Telematics & Informatics, 2018, 35(4):629-642.
- [31] Jia S S. Motivation and satisfaction of Chinese and U.S. tourists in restaurants: A cross-cultural text mining of online reviews[J]. Tourism Management, 2020(78):104071.
- [32] Elberse, Anita. The Power of Stars: Do Star Actors Drive the Success of Movies?[J]. Journal of Marketing. 2007, 71(4):102-120.
- [33] 黄会林, 孙子荀, 王超, 等. 中国电影与国家形象传播——2017 年度中国电影北美地区传播调研报告 [J]. 现代传播 (中国传媒大学学报), 2018, 40(1):22-28.
- [34] Moon S, Mishra A, Mishra H, et al. Cultural and Economic Impacts on Global Cultural Products: Evidence from U.S. Movies[J]. Journal of International Marketing, 2016, 24(3):78-97.