

具身理论视野下的智能虚拟助理可视化设计研究

李佳佳, 于沁宁

(青岛大学, 青岛 266000)

摘要: **目的** 具身认知和存在现象学等理论为虚拟助理设计研究提供了新的视野, 重构虚拟助理的可视化设计方法论, 提升用户在交互过程中依托身体参与产生的认知理解和社交体验。**方法** 通过梳理具身理论及其在虚拟环境下的用户研究, 结合案例分析, 指出用户的具身感知与虚拟助理建构社交关系的重要意义, 并将这一论点融入虚拟助理的可视化设计原则及其方法中, 进一步推演出具身交互的设计范式。**结论** 依托身体参与产生意义的具身体验成为用户与虚拟助理间最直接有效的交互手段, 虚拟助理的具身性构建了人机社交关系, 促使虚拟助理的可视化设计应结合其具身本质及其社交化特征, 从虚拟人格塑造(形)、非语言行为的视觉表现(动)到社交化的视觉反馈(情)3个层次阐释并建立虚拟助理的可视化设计范式。

关键词: 具身设计; 虚拟助理; 社交效应; 用户感知

中图分类号: TB47 **文献标识码:** A **文章编号:** 1001-3563(2021)18-0269-05

DOI: 10.19554/j.cnki.1001-3563.2021.18.032

Visual Design of Intelligent Virtual Assistant from Perspective of Embodiment

LI Jia-jia, YU Qin-ning

(Qingdao University, Qingdao 266000, China)

ABSTRACT: Theories of Embodied Cognition provides a new vision for the design methods of Virtual Assistants, re-constructing the visual design methodology of virtual assistants, and promoting users' perception and social experience based on body in the process of interaction. Through the analysis of embodied theories and its user research under the virtual environment, and combining with case analysis, this paper points out the significance of social relationship between users' embodied perception and Virtual Assistants, and integrates this argument into its visual design principles and detailed design approaches to further deduce the design paradigm of Embodied Interaction. Based on the involvement of bodies, Embodied Experience becomes the most direct and effective approach to build the human-computer social relationship. The visual design of Virtual Assistants should focus on its embodied traits and its social attribute. The design paradigm of Virtual Assistants is interpreted and built from three levels including shaping characteristics, building non-verbal behaviors and designing the visual feedback of sociability.

KEY WORDS: embodied design; virtual assistant; social effect; user perception

人际交流是个体之间通过语言、肢体动作、表情等手段进行信息传递和情感互通的过程。在人机交互领域,为了能让机器与人类产生和谐的互动效果,使机器具备像人一样的“社交”能力,成为当下人机交互理论研究的新热潮。而智能虚拟助理更是凭借着“具身化”的外形、语言和动作成为机器社交化的重

要研究对象。

自2014年起,各大科技公司纷纷推出了具有代表性的智能虚拟助理,包括苹果Siri,微软Cortana、谷歌Assistant、亚马逊Alexa等。国内公司也加紧了步伐,紧锣密鼓地发布了百度度秘、阿里小蜜、腾讯叮当等,力图从智能虚拟助理的场景切入,完成在未

收稿日期: 2021-06-05

基金项目: 山东省社会科学规划研究专项(19CKPJ08); 山东省艺术科学重点课题(ZD201906108)

作者简介: 李佳佳(1988—),女,山东人,博士,青岛大学副教授,主要研究方向为虚拟现实与智能体设计。

来人工智能市场的部署。这些智能虚拟助理通过具身化的属性,不仅呈现了自身的特征和意图,更赋予了其社会化的角色,实现了人机的社交互动。然而,鲜有从艺术设计领域考察智能虚拟助理具身设计问题的复杂性和多样性。因此,本研究通过梳理“具身化”所带来的社交效应,提出在社交范式下“具身化”智能虚拟助理的可视化设计原则及其设计方法,以期给未来的交互设计师在创造智能虚拟助理的“具身化”特征时,提供相应的设计指南和理论指导。

1 虚拟环境中的具身感知

具身思想并不是一个全新的概念,它有深刻的哲学渊源。其理论研究和哲学探索源自海德格尔的“存在论”和莫里斯·梅洛·庞蒂的直觉现象学,强调了人类的感知体验都可以通过身体来实现。在海德格尔现象学研究的“存在”概念中,是需要通过一种特殊的存在者——此在,即人来进行的^[1]。它强调的是一种参与状态,在这种状态下,人认识世界的方式是用人的身体以合适的方式与世界其他事物进行互动,在互动过程中获得对世界的认知^[2]。即是说,此在不是简单地嵌入在世界中,而是具身化的存在^[3]。实际上,莫里斯·梅洛·庞蒂的“身体-主体论”的概念与具身理论更为相关,他主张知觉的主体是身体,人是通过身体而不是意识去构建周围的环境和空间,并与其他事物进行互动。更重要的是,庞蒂强调了人与世界的关系不是封闭对立,而是开放交流。在感知世界里,人不仅能够感知自我,也能感知到外物的情况,感知知觉是来自双方,这一理论为后来具身认知概念的出现奠定了基础。

具身认知强调了身体在认知过程中所发挥的关键作用,是认知心理学的研究范畴。近些年,这一概念越来越成为人机交互领域的研究热点,尤其是在计算机技术蓬勃发展的今天,如人工智能、机器学习、虚拟现实等,这些技术使计算机不再是一个外在的存在,而是成为了人的身体的延伸,人机交互则在人的身体感知下与机器共同完成具身化的体验。本研究着重探讨的对象是在智能虚拟环境中,一个典型的具身化群体——智能虚拟助理(Intelligent Virtual Agents, IVA),指的是在虚拟环境中具有生命的对象 Agents,以及由用户化身的对象 Avatar。不少研究将 Agents 与 Avatar 统称为助理 Agents,而事实上两者有着鲜明的区别,Avatar 通常指在虚拟环境中,用户通过角色化身参与到虚拟世界的互动中,即指由用户扮演的虚拟角色。Avatar 具备一定的智能化,通过计算机算法和设定可以自主的完成具体的任务,但实际上仍是由用户进行操作,如游戏《第二人生》中创建的虚拟角色。Agents 则指代那些具备独立智能化,由计算机生成的,用来模拟真实世界中的人或事物行为和动作的虚拟角色,如 SpeakaZoo 的虚拟动物、微软“小冰”,

见图 1。

对于两种虚拟助理的划分,将导致用户在虚拟环境中不同的具身化感知。虚拟助理 Avatar 的具身感知可分为三个部分:自定位感觉、代理感觉、身份认知^[4]。强调了用户凭借具身化的角色化身,将参与感、沉浸感、具身感等心理的认知状态投射到物理的身体感知层面,事实上感知的是自我。而虚拟助理 Agents 由于其智能化的独立性,使用户能够感知外物的“具身化”存在,感知知觉是来自双方。这种交互式的具身感知,将有助于用户与虚拟助理 Agents 建立一种社交范式,即巴克斯顿(Buxton)于 1987 年提出技术反映的人类三个层面能力:物理感知,认知技能,和社会关联^[5]。

2 具身理论研究及其社交情感可视化

2.1 智能虚拟助理的“具身”意图

智能虚拟助理所涉猎的领域甚广,为人所熟知的领域诸如计算机游戏、虚拟训练环境、虚拟主持人、虚拟导购等,智能个人助理也是其中一类,是具备较高水平智能化的虚拟助理,包括苹果 Siri,微软“小冰”,百度度秘等,其共同点是拥有“具身化”的特征。詹姆斯·莱斯特(James C Lester)称之为“人格效应”(Persona Effect),他认为“拟人化”的人格塑造能够提升用户的认知,帮助用户参与到交互过程中,并改善交互结果^[6]。Botanic 虚拟角色公司总裁马克·梅多斯(Mark Meadows)在谈及虚拟助理时表示“虚拟助理需要人格,这样我们才知道如何与他们互动。这些人格会体现在他们的外表上。”^[7]事实上,用户会将计算机人格化,并赋予它们人的特征和属性是社会的、自然发生的,即使用户知道这样做不合理^[8]。这意味着,人对于计算机的人格化反应是一种本能的自然状态,当计算机具备更多的拟人特征时,这种人格化的反应会愈发强烈,丹尼尔·丹尼特(Daniel Dennett)称之为“意向立场”(Intentional Stance),是一种理解世界存在和其他个体行为的方式^[9]。这些像

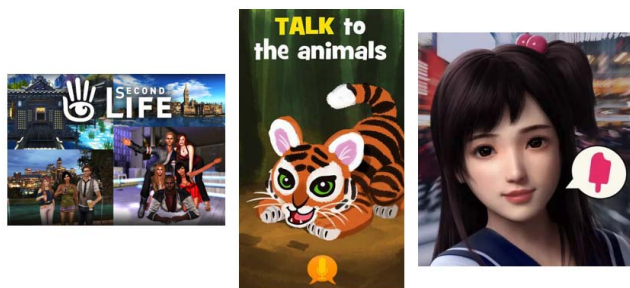


图 1 游戏《第二人生》、SpeakaZoo 及微软“小冰”3D 形象设计

Fig.1 The 3D image design in Second Life, SpeakaZoo and Microsoft “Xiaobing”

人的“具身化”特征所带来的心理认知技能,将有助于用户理解智能虚拟助理及其所交互的信息,也能够表达智能虚拟助理其本身的意图和情感。

2.2 智能虚拟助理的社交情感可视化

人作为一个社会群体和存在,是“具身化”和社会化的角色。当智能虚拟助理凭借其“具身化”的特质存在时,用户在交互过程中不仅是获取信息,也在完成社会活动,建立社交关系,只是社交对象由真实的人转换为虚拟的智能个体^[10]。在人际交流中,为了维持好的社交关系,往往需要人们在了解自我意图和情绪的基础上,还要理解交互对象的意图和情绪,以给予正确而合理的情感反馈。戈尔曼(Goleman)教授称之为“社交智能”(Social Intelligence),他认为建立和发展社交智能的主要影响因素之一就是情感化表现。

尽管会话功能一直被视为决定虚拟助理优劣的至关因素,但维持长期的社交情感互动需要可视化设计元素的介入。虚拟助理的情感化表征,主要通过语言、外形、动作、神态等元素呈现,而这些情感表现元素将成为虚拟助理进行可视化设计的主要特色,并成为与用户建立信任感和融洽感的手段,辅助用户通过这些情感可视化的特征来了解智能虚拟助理的具体技能,从而减轻认知负担。例如,瓦尔克(Walker)等^[11]的研究表明,与文字界面相比,用户跟一张会说话的脸互动,会花更多的时间与之交互,在互动过程中犯错误也更少。雷特(Leite)等的实验则验证了具备相同可视化外形的机器人,具备情感反馈的机器人会比没有情感表现的机器人更受欢迎,使人感到愉悦。因此,这赋予了智能虚拟助理的可视化设计以极大的灵活性,开辟了一个新的“具身”化设计新空间,即通过“具身”化所带来的社交情感效应建立智能虚拟助理的可视化模型及感知特性。

3 智能虚拟助理的可视化设计方法

3.1 用户创建意义

智能虚拟助理作为一种交互式产品,是需要用户在交互过程中发挥作用的。换言之,只有用户参与到会话过程中,虚拟助理的任务才能够完成。用户参与会话过程的程度,受到用户对虚拟助理喜爱(Likability)程度的影响。这意味着,将用户调研放在对于智能虚拟助理的前期可视化设计阶段,有助于设计师设计符合用户需求和使人喜爱的虚拟助理形象。因此,在虚拟助理的可视化设计过程中,需要由用户而不是设计师来决定虚拟助理的“具身”特征。如目前绝大多数的智能虚拟助理几乎都是女性身份,包括苹果的 Siri、微软“小冰”、亚马逊 Echo,还有机器人“索菲亚”“佳佳”等也被设计为女性形象居多,这并非偶然,而是在一份调研中,用户在与虚拟助理的交互过程中

更愿意接受来自女性声音的帮助,这也成为多数虚拟助理是女性角色的原因,见图 2。

瓦尔克(Walker)在早期也进行了关于用户对虚拟助理的交互反应研究,实验中对比了三种虚拟助理的可视化形象:一个文本界面、一张面无表情会说话的脸和一张表情严肃会说话的脸。实验表明,一张表情严肃的脸不仅能够激发用户进行更多的会话讨论,还创造了让人不愉快的交互体验。因此,设计者在进行虚拟助理的可视化形象设计之初,需要适当进行用户调研,即便用户与虚拟助理的交互过程是合作完成的,但对虚拟助理的“第一印象”仍然掌握在用户手中,并决定着用户是否愿意将互动进行下去。

3.2 人格塑造

3.2.1 建立人格特征

按照科恩(Cohen)在《语音用户界面设计》一书中对虚拟助理人格化的定义:“我们用来向我们或其他人,表达有意识意图的假象角色。”^[12]在虚拟助理“具身化”设计领域,“人格”等同于“角色”,即用户通过虚拟助理的“具身”化特征,判断出人格或角色的标准化心理形象。在这里之所以称之为“心理形象”,是为了涵盖具有完整的拟人、非拟人化形象以及无可可视化形象的虚拟助理。事实上,一个没有可视化形象的虚拟助理,仍可以通过它的其他输出模式诸如语音进行人格判断,在心里塑造出一个虚拟的可视化形象。“具身化”帮助塑造了虚拟助理的人格特征,这些具有“人格”的虚拟助理形象不仅能够建立用户与虚拟助理的信任感、融洽感、同理心等移情作用,也能促进用户对虚拟助理的具体能力和可执行任务的认知。

3.2.2 “一致性”原则与规避“恐怖谷”

在“人格”的设计过程中,并不需要一味地为虚拟助理增加不必要的人格特征。人格的一致性,不仅取决于虚拟助理各交互模态的一致性,如语言、动作、神态、表情、外形等,也取决于虚拟助理的执行任务情境以及用户的特征,如个性、年龄、性别等。因此,人格特征设计的一致性至关重要,它让使用户在与虚拟助理的沟通中预测接下来会发生什么。

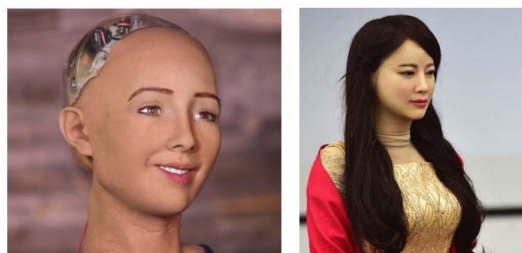


图 2 机器人“索菲亚”以及中国科学技术大学研发的交互机器人“佳佳”

Fig.2 The Robot “Sophia” and “Jiajia” developed by University of Science and Technology of China

人们擅长于识别图标,并且能很快地识别出常见的图标。因此,为了实现用户与智能虚拟助理的和谐交流,增加其“拟人化”的视觉属性,将利于用户与之建立社交情感,甚至信任感。然而,在拟人化的过程中,无可避免地会遇到“恐怖谷”,见图3,即当随着机器人的拟人化程度越来越高时,人们的情感反应会遭遇一个急剧的下滑状态,产生恐怖感。石黑浩(Hiroshi Ishiguro)一直致力于验证“恐怖谷”效应的真实性,并提出了“如果一个不怎么像人的机器人,具有符合人性的社会化动作,会比那些长得像人的机器人更拟人化。”^[13]因此,社交特性将成为虚拟助理规避“恐怖谷”的有力武器,虚拟助理的外形也许并非一定要“像人”,也许只是一个抽象图像或者一个非拟人化的形象,但其一定具备着人格特征并努力与用户产生社交互动。

3.3 社交智能的视觉反馈

3.3.1 具身化的形象设计

具身化的形象设计有助于智能虚拟助理获得独立的“人格特征”,提升用户的参与度,即使是简单的面孔也能唤起用户的情绪反应。智能虚拟助理的“具身”化形象既可以是静态的图像、卡通形象,或是一段真人的视频,一些非拟人化的虚拟形象,甚至可能只是一个抽象的可视化动态图标。目前在市场投入使用的虚拟助理类产品中,具身化的形象大致可分

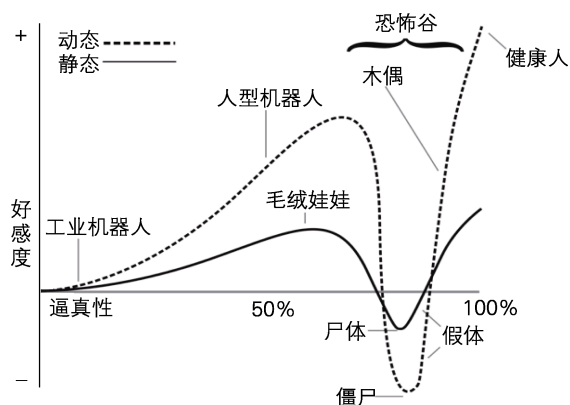


图3 Mori的“恐怖谷”理论

Fig.3 “Uncanny Valley” Theory from Mori

为两类:抽象化视觉符号和具象化视觉符号。前者包括苹果的Siri、亚马逊Alex等没有虚拟形象的智能助理,以简单的视觉符号提示作为信息反馈,后者则包括微软“小冰”“小i”机器人等具备一个完整虚拟形象的智能助理。其可视化的呈现形式主要区分为自然拟物化(Naturalism)和图标化(Iconic)、二维和三维,以及虚拟屏幕和有形实体。

更为重要的是,不同的呈现形式在不同场景下也会发生变化。对于一个成功的智能虚拟助理而言,是否需要拥有一个“具身化”的形象,取决于具体的使用场景、功能定位、用户及市场需求等因素,而它们构成了设计前期独有的“形象模型”,以辅助设计师设计与之对应的形象。譬如,微软虚拟助理Cortana的设计师就特意避免了一些具象化的形象设计,用视觉反馈来表现虚拟助理的信息接收和回复,但后期为了使Cortana的人格特征更为鲜明、情感表达更为丰富,设计团队与游戏《Halo》的开发团队进行了合作,将Cortana定义为该游戏里的一个女性角色,见图4。因此,Cortana视觉反馈的色彩效果也与游戏中女主角特有的蓝色相关,即以蓝色的圆环动效作为虚拟助理Cortana倾听和回复的状态。

当确定智能虚拟助理需要一个具象化的形象时,为了避免“恐怖谷”以及拟人化特征所带来的诸如种族、性别、年龄等在角色上的属性差异,建议虚拟形象选择卡通或动物等无社会属性的角色,如获得较高形象评价的Jibo、Kuri、Rokid等虚拟助理均有着相似的造型特征,包括圆弧形的设计、单色及光滑的外表、卡通式的头部与身体分配比例。这些形象特征如前文所述,取决于对使用场景、功能、用户及市场的定位,正如Kuri的设计师Kaijen Hsiao所言:“我们很难去打造让所有人喜欢的装饰品,而最接近大众的选择,就是极简主义的风格加上简单的白色。”

3.3.2 非语言行为与社交反馈

动作、表情、神态等非语言行为同语言行为一样能够对信息进行交换和传播,甚至所传递的信息更多。智能虚拟助理的非语言行为不仅为用户提供了符合自然用户习惯的多模态交互方式,也激发了用户与



图4 Cortana、小i机器人以及Jibo的形象设计

Fig.4 The visual design of Cortana, Xiao-i and Jibo

虚拟助理的社交关联。迪卡·戴根恩 (Tickle-Degnen) 和罗伯特·罗森塔尔 (Robert Rosenthal) 将其称为非语言行为促发的融洽关系, 即交互双方通过表现积极的情绪 (点头、微笑), 互相凝视以及配合 (同步的展现动作) 可以增加双方具备融洽感的社交关系。好的非语言行为表现能够弥补智能虚拟助理在会话或形象上的不足, 在南加州大学 ICT 实验室的一项研究表明, 与动态虚拟人物交流更能提升用户的参与感和信任感。对于有具体形象的智能虚拟助理而言, 非语言行为具体表现在拟人化的动作、表情、神态等方面; 对于不具备形象的智能虚拟助理而言, 非语言行为主要体现在可视化的视觉反馈效果, 使用户实时获悉系统正在倾听和接收的信息。

以社交过程为例: 首先, 用户与虚拟助理的交互是需要一个信息的接收口和一个信息传达的出口, 这意味着两者需要在社交会话中呈现出倾听和反馈的状态。其表现出来的输入和输出接受状态是区别于常态状态的一种强化视觉表现, 如抽象类的虚拟助理形象, 倾听和反馈的状态更倾向于用动态的音波呈现, 而拟人化程度更高的虚拟助理倾向于用耳朵的动态, 或用一些“点头”“微笑”等表示赞同或倾听的动作表达信息已被接收, 亦更符合人际交往的互动体验^[14], 见图 5。其次, 在用户与虚拟助理交互过程中, 需要可视化的图像或动作表示交流过程的开始与结束, 也就是需要表明明确的界限感, 以保证用户能够准确地接收信息或者实时地进行信息反馈。最后, 虚拟助理可视化表达的多样性、随机性将更好地体现智能化的水平, 重复的、循环的可视化形态会导致苍白和低效, 使用户对虚拟助理的智能程度产生怀疑, 甚至会因为自我效能感的降低而对技术产生抗拒感。

综上所述, “具身化”智能虚拟助理的可视化设计方法遵循一个“回路”模型, 见图 6。从用户参与设计过程入手, 进入“上层”的形象及人格塑造, 最终通过这些“具身”的可视化设计攀升“顶层”的情感领域, 实现用户与智能虚拟助理的社交体验, 而其所建立的社交体验又形成了“回路”, 影响着用户对于虚拟助理的具身认知。

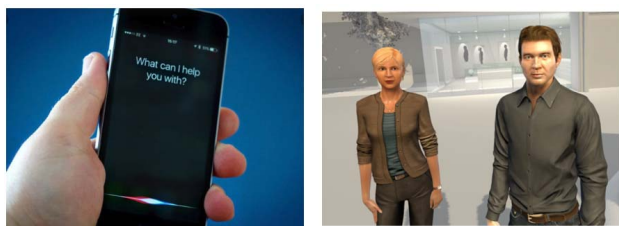


图 5 苹果 Siri 接收信息时的音波动效和南加州大学 ICT 实验室开发的虚拟助理

Fig.5 The sound wave motion effect when Apple Siri receives information and the virtual assistant developed by the ICT Laboratory of the University of Southern California

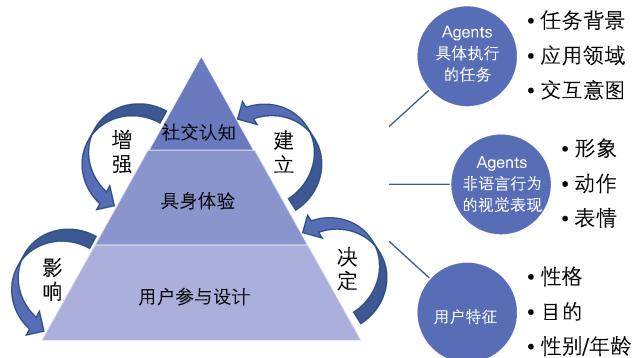


图 6 “具身”智能虚拟助理的可视化设计模型

Fig.6 The visual design model of “Embodied” Intelligent Virtual Assistants

4 结语

人类天生就喜欢人格化, 将身边的事物当成另一个“人”的存在, 而这种赋予“人格化”的能力存在人们的潜意识中。智能虚拟助理凭借具身化的特征, 将用户带入了一个由数字创造的赛博空间, 使用户意识到除了“我”之外的另一个具备智能、情感的独立个体存在, 打破了以往人们对智能助理类产品的工具属性认知, 并尝试与其建立人机社会关系, 使其不断融入人们的日常生活中^[15]。从具身视野洞悉智能虚拟助理可视化设计的多样性和复杂性, 融入其特有的具身本质以及社交属性, 促使设计从单一的形象处理走向多模态可视化的情感转型, 将对构建虚拟助理的具身设计范式有着指导性意义。

参考文献:

- [1] 朱立元. 现代西方美学史[M]. 上海: 上海文艺出版社, 1996.
ZHU Li-yuan. The History of Modern Western Aesthetics[M]. Shanghai: Shanghai Literature and Art Publishing House, 1996.
- [2] 叶浩生. 具身认知: 认知心理学的新取向[J]. 心理科学进展, 2010(5): 705-710.
YE Hao-sheng. Embodied Cognition: A New Approach in Cognitive Psychology[J]. Advances in Psychological Science, 2010(5): 705-710.
- [3] 姚争为, 杨琦, 潘志庚, 等. 具身交互与全身交互的比较[J]. 计算机辅助设计与图形学学报, 2018(12): 2366-2376.
YAO Zheng-wei, YANG Qi, Pan Zhi-geng, et al. Comparison of Embodied Interaction and Whole Body Interaction[J]. Journal of Computer-Aided Design & Computer Graphics, 2018(12): 2366-2376.
- [4] 王楠. 基于具身视角的 VR 电影场景叙事[J]. 当代电影, 2018(12): 111-115.
WANG Nan. On VR Film Scenario Narrative Based on Embodied Angle of View[J]. Contemporary Cinema, 2018(12): 111-115.

(下转第 281 页)