

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2024)10-2937-18

论文引用格式: Yang H Z, Shu W T and Wang M. 2024. Literature review of contextual information construction and applications in mixed reality. Journal of Image and Graphics, 29(10):2937-2954(杨浩中, 舒文桐, 汪淼. 2024. 混合现实中的上下文信息构建与应用进展综述. 中国图象图形学报, 29(10):2937-2954)[DOI:10.11834/jig.230750]

混合现实中的上下文信息构建与应用进展综述

杨浩中^{1#}, 舒文桐^{1#}, 汪淼^{1,2*}

1. 北京航空航天大学虚拟现实技术与系统全国重点实验室, 北京 100191; 2. 中关村实验室, 北京 100094

摘要: 随着信息技术的发展, 混合现实(mixed reality, MR)技术已应用于医疗、教育和辅助引导等众多领域。MR场景包含丰富的语义信息, 基于场景上下文信息的混合现实技术可以改善用户对场景的感知, 优化用户的交互操作、提升交互模型的准确度, 受到广泛关注。然而, 目前在该领域没有针对上下文信息进行调查的综述类文献, 缺乏梳理与分类。本文的研究对象是使用上下文信息的MR技术与系统。通过对MR领域的文献调研, 本文提出了3个研究问题, 并对国外近20年的33篇实证研究论文进行分析, 概述了使用上下文信息的MR技术的最新发展, 从3个维度出发进行分类学研究并分别提出分类标准, 如上下文信息种类、上下文知识库的构建方式和应用领域等。其中, 上下文信息的种类可以分为场景语义、对象语义、空间关系、群组关系、从属关系和运动信息6类, 知识库的构建按用户介入角度和基础技术类型划分, 应用领域从场景类型和生成式特性出发进行分类。通过对不同维度对研究对象的分类, 本文对提出的研究问题进行了回应, 并总结了现阶段的不足以及未来可能的研究方向。本综述可以辅助不同领域的研究人员对上下文信息的设计、选择和评估, 从而推动未来混合现实应用技术与系统的研发。

关键词: 虚拟现实(VR); 增强现实(AR); 感知与交互; 上下文信息; 场景语义

Literature review of contextual information construction and applications in mixed reality

Yang Haozhong^{1#}, Shu Wentong^{1#}, Wang Miao^{1,2*}

1. State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China;
2. Zhongguancun Laboratory, Beijing 100094, China

Abstract: With the development of information technology, mixed reality (MR) technology has been applied in various fields, such as healthcare, education, and assisted guidance. MR scenes contain rich semantic information, and MR technology based on scene context information can improve users' perception of the scene, optimize user interaction, and enhance the accuracy of interaction models. Therefore, they have quickly gained widespread attention. However, literature reviews specifically investigating context information in this field are limited, and organization and classification are lacking. This paper focuses on MR technology and systems that utilize context information. This study was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. First, keywords for the search were determined on the basis of three factors: research domain, study subjects, and research scenarios. Subsequently, searches were performed in two influential databases in the field of MR: ACM Digital Library and IEEE Xplore. A prelimi-

收稿日期: 2023-10-17; 修回日期: 2023-12-19; 预印本日期: 2023-12-26

共同一作; * 通信作者: 汪淼 miaow@buaa.edu.cn

基金项目: 国家自然科学基金项目(61932003); 中央高校基本科研业务费专项资金资助

Supported by: National Natural Science Foundation of China(61932003); Fundamental Research Funds for the Central Universities

nary screening was then executed, considering the types of journals and conferences to eliminate irrelevant and unpublished literature. Subsequently, the titles and abstracts of the articles were reviewed sequentially, eliminating duplicates and irrelevant results. Finally, a total of 210 articles were individually screened to select 29 papers for the review. Additionally, four more articles were included on the basis of expertise, resulting in a total of 33 articles for the review. Through a comprehensive literature review of MR databases, three research questions were formulated, and a dataset of research articles was established. The three research questions addressed in this paper are as follows: 1) What are the different types of scene context? 2) How is scene context organized in various MR technologies and systems? 3) What are the application areas of empirical research? On the basis of the evolution of scene context and the refinement of MR technologies and systems, we analyze the empirical research papers spanning nearly 20 years. This analysis involves summarizing previous research and providing an overview of the latest developments in systems that leverage scene context. We also propose potential classification criteria, such as types of scene context, construction methods of knowledge bases for contextual information, fundamental technologies, and application domains. Among the various types of scene context, we categorize them into six classes: scene semantics, object semantics, spatial relationships, group relationships, dependence relationships, and motion relationships. Scene semantics is the semantic information encompassed by various elements in the scene environment, including objects, characters, and texture information. In the categorization of object semantics, we consider information about the individual object itself, such as user information, type, attributes, and special content. Spatial relationship refers to numerical information, such as the relative position, angle, or arrangement between various objects in the scene. We analyzed spatial relationships in three ways: base spatial relationships, microscene spatial information, and real-scene spatial information. We consider a certain number of closely neighboring objects of the same category as a group. Group relations focus on information about the overall perspective such as intergroup relations and the number of groups. Dependence relationship is concerned with the dependencies and affiliations that may exist between different objects in the scene at the functional and physical levels. Motion information is a new type of scene context, including basic motion information and special motion information, which describes the dynamic information of scene objects. Through an analysis of the utilization of various types of scene context, we establish the relationship between research objectives and contextual information, providing guidance on the selection of contextual information. The construction of knowledge bases is examined from user-intervention perspectives and types of fundamental technologies. Knowledge bases established with user intervention typically rely on researchers' abstract analysis of scene objects rather than pre-existing databases. Conversely, knowledge bases built without user intervention rely on existing information, such as low-level raw data in databases or predefined scenarios. The underlying technologies in this context are categorized into virtual reality (VR) and augmented reality (AR). Conducting classification research from the dual perspectives of user intervention and fundamental technology facilitates a deeper understanding of how contextual information is organized in various MR systems. Application areas are investigated on the basis of the types of scenarios and whether they involve generative processes or not. The types of application scenarios are then categorized into six types: auxiliary guidance, AR annotation, scene reconstruction, medical treatment, object manipulation, and general purpose. Generative models can automatically generate target information, such as AR-annotated shadows based on the scene, whereas nongenerative models mainly focus on specific operations. Through analysis from these two perspectives, the advantages and disadvantages of MR systems and technologies in different application scenarios can be explored. Drawing upon the exploration and research in these three dimensions, we investigate the challenges associated with selecting, acquiring, and applying contextual information in MR scenarios. By classifying the research objects from different dimensions, we address the research questions and identify current shortcomings and future research directions. The aim of this review is to support researchers across diverse fields in designing, selecting, and evaluating scene context, ultimately fostering the advancement of future MR application technologies and systems.

Key words: virtual reality(VR); augmented reality(AR); perception and interaction; context information; scene semantics

0 引言

随着科技的不断发展,混合现实(mixed reality, MR)技术逐渐成为人们关注的焦点。混合现实技术可以看做是增强现实(augmented reality, AR)和虚拟现实(virtual reality, VR)的结合(Speicher等,2019)。通过将虚拟世界和现实世界相结合,用户可以在真实的环境中与虚拟的物体进行交互,从而获得一种全新的体验。任何混合现实系统都具备3个重要的特点(Rokhsaritalemi等,2020):1)真实世界和虚拟世界对象的结合;2)实时交互;3)虚拟对象和真实对象之间的映射。

人们已经对MR技术进行了各种研究,并创建了基于其特点的多种调研,如MR应用技术(Costanza等,2009; Bekele等,2018)、隐私安全(Acquisti等,2014; Roesner等,2014; Szczuko, 2014; De Guzman等,2019)和可视化(Sato等,1999; Loscos等,2000; Haller等,2003; Jacobs和Loscos,2006)等,并应用于医疗(Nifakos和Zary,2014; Moro等,2017; Stretton等,2018)和教育(Campbell等,2016; Murphy等,2018; Sala,2021)等多种领域。MR技术有着广阔的应用领域和可观的发展前景,已然成为新时代信息技术中的重要研究方向。

混合现实场景中各种技术的实现离不开信息的获取。在MR场景下客观存在着许多包含信息数据的内容,如物体、数值、关系和背景等,这些信息定义了场景上下文信息。早在1995年,Bukowski和Séquin(1995)利用场景对象的空间上下文信息,提出了一个虚拟环境中三维对象操作的框架。之后,Gösele和Stuerzlinger(1999)在场景操作中使用了语义信息进行约束,考虑了场景对象的从属关系。但此后场景上下文信息相关的论文数量并没有明显增长。2010年之后,随着混合现实技术的进一步发展,MR相关的各类应用场景逐渐丰富,人们开始重视场景中上下文内容所蕴含的信息。场景上下文信息频繁出现于各种涉及场景对象处理的工作中,如室内场景的重建(Chen等,2014)和三维对象的选择(Wang等,2021)等研究。上下文信息的应用领域也迎来快速发展,广泛应用于医疗、AR注释和场景重建等方向。然而,尽管目前场景上下文的研究已经取得可观进展,并已经在一些领域得到初步应用,但

研究人员在选择、获取和应用上下文信息时较为随意、混乱,缺乏系统性指导。由此,为了继续推动场景上下文信息研究方向持续发展,亟需对场景上下文信息进行全面调研,为研究人员提供系统性指导建议。

本文对混合现实场景下的上下文信息相关的研究进行了调研。作为第一项对该领域研究情况进行综述的工作,本文在上下文信息类型、知识库构建过程和应用领域等维度进行了讨论,并在每个维度进行了分类学研究,通过文献检索与调研,分析了场景与上下文信息的关系,并讨论了应用领域的现状。本工作主要有以下3项贡献:1)调查上下文信息有关的工作,进行详细梳理,并简述其发展历史。2)对目前在该领域的文献进行分类法研究,从3个维度分析并各给出一种分类标准。3)系统性地对研究人员进行指导,帮助研究人员了解混合现实场景上下文信息,使其能够更加高效、准确地在研究工作中运用场景上下文信息。

1 概念定义和研究问题的提出

本节首先明确研究的问题与背景。本文关注的是混合现实特定场景下利用上下文信息的工作。近些年来,不少混合现实方向的论文利用了场景中的上下文信息来辅助问题的研究。本文详细调查了涉及场景上下文信息的论文,分析相关问题,并进行分类学讨论。

1.1 对混合现实的定义

AR、VR和MR是3种不同的技术,它们的主要区别在于现实世界和虚拟世界之间的交互方式、程度以及应用场景。AR通过在现实世界中叠加虚拟元素来增强用户的感知体验,将计算机生成的材料,如文本、图像和视频,投射到用户对现实世界的感知上(Yuen等,2011)。该技术通常部署于手机、平板电脑或AR眼镜等电子设备,应用场景包括游戏、教育和医疗等领域。VR技术则是在虚拟世界中的沉浸式体验,用户可以通过佩戴虚拟现实头戴式设备访问虚拟世界,并使用手柄等设备进行交互(Anthes等,2016),其应用场景包括游戏、培训和模拟等领域。

关于MR,本文参考Speicher等人(2019)的定义。MR技术可以看做是AR和VR的集合体,它将

增强现实和虚拟现实结合起来,创造出一种新的现实感受。在这种情况下,MR表示在同一款应用程序或同一款设备上使用AR或VR的能力。在本文中,根据Milgram和Kishino(1994)以及Milgram等人(1995)的研究,首先定义MR的研究范围(如图1所示)。本文所考虑的MR处于现实到虚拟范畴之间,包括增强现实、增强虚拟现实和虚拟现实等。之后的研究中将根据定义的范畴对增强现实和虚拟现实进行讨论分类。

1.2 对上下文信息的定义

本文将上下文信息定义为:在混合现实特定场景下客观存在的物体、数值、关系和背景等空间或时间信息,这些信息可以直接或通过计算后间接获取。在英文中常以Context、Awareness、Cue和Guidance等词语抽象出MR中的环境、感知、线索和引导等信息。这些信息存在于场景中,为研究目标提供信息数据的支持,帮助研究人员更加高效精准地达成研究目标。上下文信息客观存在于特定场景内,并由研究主体所获取。故本文所研究的上下文信息并不包括物理感知,如听觉、触觉等反馈,以多模态作为输入的研究系统也不在本文的研究范围内。

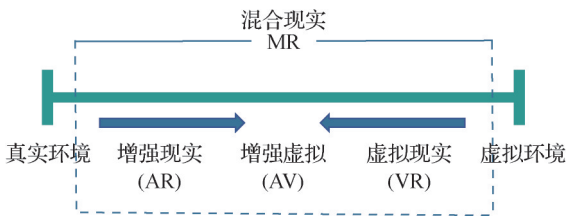


图1 MR研究范围

Fig. 1 Research scope of MR

1.3 研究问题

本工作的目的是回答以下研究问题:

问题1):MR中构建和应用的上下文信息有哪些类型?场景上下文信息涵盖着丰富的各类信息,不同场景中存在的信息也大不相同。本文探究在特定的混合现实场景中,研究者更倾向于哪些种类的上下文信息。更具体地,本文探究以下子问题:

(1)在MR场景中更倾向于使用哪种信息?

(2)不同上下文信息和研究问题之间存在着什么样的关系?

问题2):上下文信息在不同MR中是如何构建和应用的?本问题目标在于广泛了解MR中使用上

下文信息的不同系统的架构,探索各个系统中上下文信息的应用方法。更具体地,本文探究以下子问题:

(1)场景中上下文信息的知识库是如何建立的?

(2)不同场景下知识库的构建基于哪种基础技术?

问题3):实证研究的应用包括哪些领域?本问题研究使用场景上下文信息的文章所探索的应用方向,探究场景上下文信息在不同领域中应用的可能性。更具体地,本文探究以下子问题:

(1)混合现实场景中,基于上下文信息的工作主要应用于哪些方向?

(2)各方向的研究现状和发展趋势如何?

2 方法论

为了回答提出的问题,本文使用国际通用惯例,即系统评价和元分析方法(preferred reporting items for systematic reviews and meta-analyses, PRISMA)进行调查(Pickering等,2015)。首先提取研究的关键字,然后选择要搜索的数字图书馆,对限定搜索的文章经过一系列处理后最终建立研究文章的数据集。

2.1 关键词的选择

本文研究对象是MR领域,根据常见MR的实现场景和同义词,将VR,AR和MR一同作为文献搜索的基础。对于“上下文”,本文使用“指导”,“线索”和“感知”等词语进行组合来表示上下文信息对用户感知的增强现象。对于具体用例方面,本文综述对象是从房间级别或场景级别展开工作,因此使用“房间”和“场景”等词语。在得到这些关键词后本文进行了组合并在数据库上检索,关键词如图2所示。

2.2 时间范围

国内外关于MR的综述大多都是对近10年内的文章进行整理总结(丁文郁和单美贤,2023;平佳敏等,2021;Zhou等,2008),但由于本文是第一项对上下文信息的运用进行综述的工作,为了避免将有影响力的文章排除在外,本文将研究范围扩展到2000年以来的所有文章。同时从检索结果也观察到,近几年来随着Oculus Quest和PICO等设备的推出,该领域文章的数量出现较大增长。

2.3 数据库和搜索参数

为了避免噪声的增加(未出版作品、网页和重复

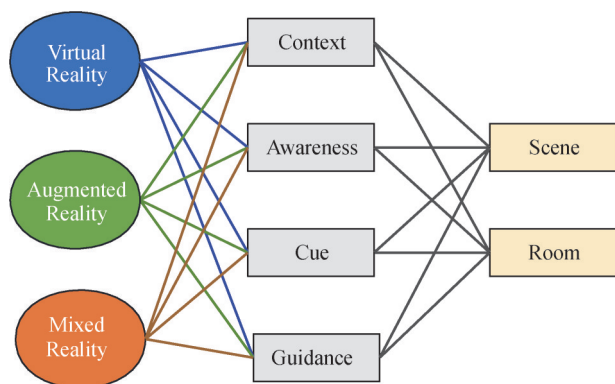


图2 检索文章的数据库中包含的关键词

Fig. 2 Keywords included in the databases that we retrieve the articles from

文献等), 本文没有选取谷歌学术等平台, 而是仅针对 ACM Digital Library 和 IEEE Xplore 这两个在 MR 领域影响力较大的数据库来进行检索。通过对每个数据库中的标题和摘要检索得到初步结果, 共 1 357 篇, 如图 3 所示。

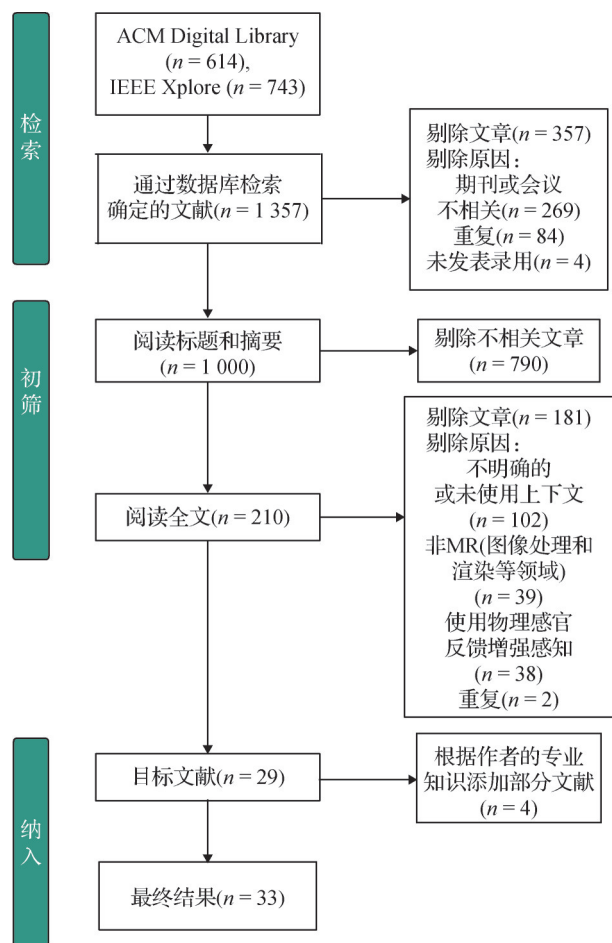


图3 PRISMA 流程图

Fig. 3 PRISMA flow

2.4 纳入和排除的标准

本文收录以英语为主要语言的论文, 并将重点放在经过同行审查的已发表的论文上, 这些论文不包括海报和演示。本文要求论文必须在特定场景下进行研究。论文不能仅仅是提出一种新的操作方式, 必须要提出完整的应用场景并进行实证研究。如在混合现实室内场景内, 利用场景上下文中的空间关系与从属关系进行三维物体选择问题的研究; 或在室外场景中, 利用场景类型约束与人群空间关系进行场景人群重建问题的研究。特定场景需应用于 VR 或 AR 环境下, 利用特定场景中客观存在的上下文信息分析问题。本文排除了那些通过物理增强并使用各种感官作为反馈来提升感知的方法, 将主要关注点集中在计算机图形学和人机交互等场景下, 如: IEEE VR、ACM CHI、ACM SIGGRAPH 和 IEEE TVCG 等。

2.5 相关文献的选取流程

图 3 显示了在 PRISMA 指导下的文献搜集流程。根据关键词初步搜索到的结果共有 1 357 篇, 其中 ACM Digital Library 614 篇, IEEE Xplore 743 篇。通过对这些文献进行初步筛选, 去除了重复、未发表和不相关的期刊或会议之后剩余 1 000 篇文献。由各个作者阅读了该结果中每一篇文章的标题和摘要, 剔除了不符合要求的文章, 进行此步骤后剩余 210 篇。接下来对 210 篇文献进行全文阅读, 保证每一篇文献被至少一位作者进行从头到尾的阅读并进行交流讨论。这一过程去除了不相关的 181 篇文献, 剩余 29 篇文献。最后根据专业知识增加了额外的 4 篇文献, 这也是 PRISMA 文献综述的常用手段 (Fidalgo 等, 2023)。本工作对最后的 33 篇文献进行阅读讨论并搜集了所有的相关数据, 在数据搜集过程中进一步细化了研究问题并应用了各种分类法。

3 分类法

本节在上下文信息类型、知识库的构建和应用领域 3 种维度进行了分类学研究, 并在每个维度提出了一种分类标准。针对各种分类方式进行汇总, 用不同的分类法解答提出的研究问题。本文分析了每个维度上的论文分布, 并将其与定性观察相结合, 深入了解不同研究间的差异。本工作根据年份梳理了各项文献和特性, 如表 1 所示。

表 1 MR 系统按照各自技术的分类

作者和时间	使用技术	应用场景	上下文信息种类
Reitmayr 和 Schmalstieg (2005)	VR&AR	AR 注释	场景语义
Irawati 等人 (2006)	AR	通用	从属关系
Oh 等人 (2006)	VR	物体操作	对象语义
Bichlmeier 等人 (2007)	AR	医疗	空间关系
Dehmeshki 和 Stuerzlinger (2008)	VR	物体操作	群组关系
Ennis 等人 (2008)	VR	场景重建	空间关系、对象语义
Kalkofen 等人 (2009)	AR	AR 注释	空间关系
Fisher 等人 (2011)	VR	场景重建	场景语义、从属关系
Merrell 等人 (2011)	VR	物体操作	空间关系、对象语义
Tutenel 等人 (2011)	VR	场景重建	空间关系、场景语义、从属关系
Yu 等人 (2011)	VR	场景重建	空间关系、从属关系
Cashion 等人 (2013)	VR	通用	群组关系、运动信息
Fujita 等人 (2013)	AR	AR 注释	空间关系、运动信息
Joseph 等人 (2013)	AR	通用	空间关系、场景语义
Chen 等人 (2014)	AR	场景重建	场景语义
Koulieris 等人, (2014)	VR	通用	空间关系、对象语义
Rumiński 和 Walczak (2014)	AR	通用	空间关系、对象语义
Rumiński (2015)	AR	通用	空间关系、对象语义
ElSayed 等人 (2016)	AR	AR 注释	场景语义
Barreira 等人 (2018)	AR	AR 注释	空间关系
Ma 等人 (2018)	VR	场景重建	空间关系、场景语义、从属关系
Wang 等人 (2018)	VR&AR	通用	空间关系
Gras 和 Yang (2019)	AR	医疗	空间关系、运动信息
Lang 等人 (2019)	AR	物体操作	场景语义
Rivu 和 Burschka (2019)	VR	物体操作	运动信息
Tahara 等人 (2020)	AR	场景重建	空间关系、场景语义、从属关系
He 等人 (2022)	VR	场景重建	空间关系、场景语义、对象语义
Köppel 等人 (2021)	AR	AR 注释	空间关系、运动信息
Wang 等人 (2021)	VR	物体操作	空间关系、对象语义、群组关系、从属关系
Yu 等人 (2021)	AR	物体操作	空间关系、对象语义
Li 等人 (2022)	AR	通用	场景语义
Rasla 和 Beyeler (2022)	VR	通用	场景语义
Li 等人 (2023)	AR	医疗	空间关系

3.1 上下文信息类型

针对所提问题,本文基于最后 33 篇论文中收集

到的信息,在上下文信息的类型维度建立了一种分类法。该分类法由 6 个部分组成,本节分析了每个

小维度上的论文分布,并提出了不同类型上下文与研究问题间存在的关系。

3.1.1 场景语义

在混合现实情景下,本文所讨论的场景语义是指场景环境下各种元素所包涵的语义信息。这些元素包括虚拟或真实场景中的物体、人物、功能可供性、材料和纹理等。该类别包括获取、分析使用场景语义的论文工作。如图 4 所示,这类工作通常通过扫描仪获取场景的 RGB-D 图像流信息(Tahara 等, 2020; Kang 等, 2020),进行处理后提取所需要的场景语义信息。

Chen 等人(2018)从 AR 的角度提出了一种基于上下文感知的 AR 框架。如图 5 所示,在获取场景输入后,该框架进行了整体的语义分割等一系列操作,并基于提取的语义信息,实现上下文感知的交互功能。

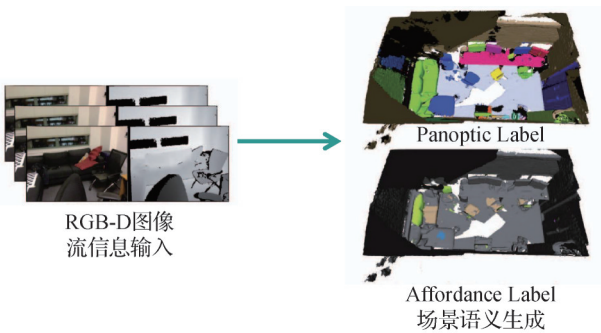


图 4 场景语义简易获取流程(Tahara 等, 2020)

Fig. 4 Scene semantics simple acquisition process (Tahara et al. , 2020)

Tahara 等人(2020)进一步丰富了 AR 表达,使用功能可供性作为额外的语义信息,将 3D 语义场景图和设计者定义的 AR 上下文场景图相结合,不仅识别到单个物体的属性,还识别了多个物体之间的关系。

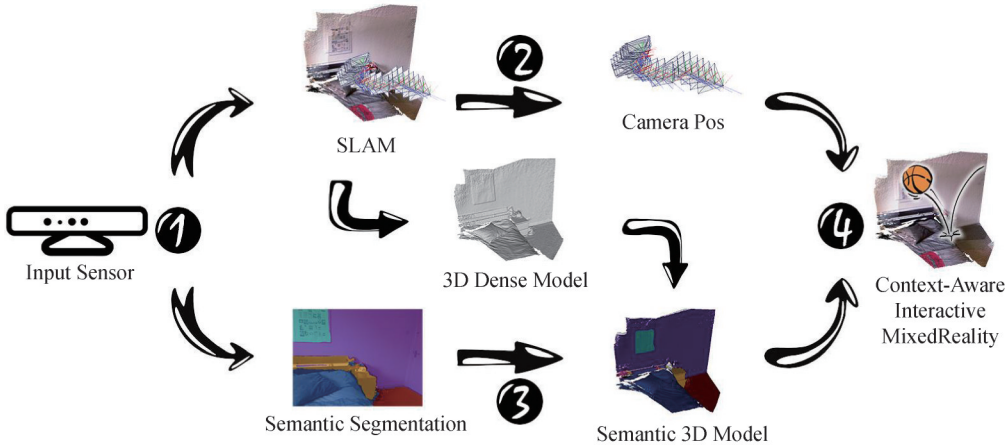


图 5 上下文感知 AR 框架流程图(Chen 等, 2018)

Fig. 5 AR context-aware framework flow(Chen et al. , 2018)

Tutenel 等人(2011)提出一个自动生成 VR 场景的程序过滤器,如图 6 所示,通过基于 Wordnet (Miller, 1995)提供的语义库查询操作,可以将“整洁的办公室”(图 6 左)和“派对”相结合,生成“派对后杂乱的办公室”(图 6 右)。通过语义查询实现场景中上下文信息的利用。语义查询是过滤过程中最重要和最强大的指令,主要功能为从场景中选择特定类的对象并查询它们的属性。

随着数字三维场景模型的发展,基于上下文的室内场景分析工作层出不穷。通过在 3D 场景数据库中学习所需的场景语义信息,进行场景内容的合成、分析、查询和比较等工作(Fisher 和 Hanrahan, 2010; Fisher 等, 2012; Xu 等, 2013)。Chen 等人

(2014)进一步解决了现实场景的识别问题,提出了室内场景自动语义建模的新解决方案。他们从 Google 3D Warehouse 中获得大量场景数据,构建 3D 场景数据库作为学习上下文关系的基础,将从虚拟场景学习到的关系应用到现实场景建模问题中。

一些研究者考虑以人为中心的环境理解与分析(Savva 等, 2014; Fisher 等, 2015; Fu 等, 2017)。这些工作通过学习场景语义信息,增强了场景中用户与虚拟代理之间的交互感知。Lang 等人(2019)利用场景语义,在混合现实应用中解决了虚拟代理情景一致化的放置问题。该工作在预处理阶段进行扫描,并使用 Mask R-CNN(He 等, 2017)方法获取场景与关键对象的信息。在优化阶段通过场景语义设置



图6 VR场景生成(Tutenel等,2011)
Fig. 6 VR scene generation(Tutenel et al. , 2011)

约束条件,设计成本函数迭代计算虚拟代理合适的位置和方向。

场景语义在场景识别、分析和重建等场景级别的研究中发挥着重要作用。通过从整个场景的视角提供场景的构造和元素等丰富的语义信息,场景语义能够为研究者提供全面的场景理解,进而指导混合现实系统更好地理解场景的内容与用户的行为意图,提供更加智能化的交互体验。

3. 1. 2 对象语义

混合现实场景中各种独立个体蕴含着大量的上下文信息。本文将场景对象定义为存在于场景内的独立个体,例如物体、用户和虚拟化身。如图7所示,在对象语义的分类中,本文考虑的是对象个体本身的信息,如用户信息、类型、属性和特殊内容等。这类工作偏重于关注某些个体或事物本身所具有的信息。对象的位置和相互关系并不在本种类型的范畴中,它们会在后续分类部分中进行阐述。

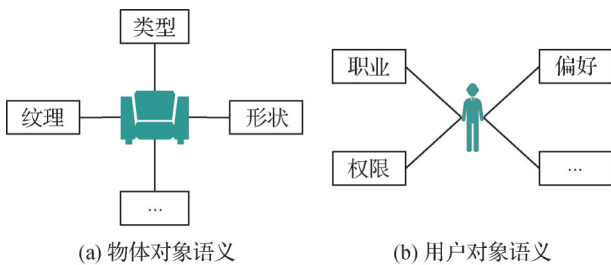


图7 对象语义
Fig. 7 Semantics of object
((a) item semantics;(b) user semantics)

在对象语义中,最常见的上下文信息是物体对象的类型和属性,如家具的类型、物体的可交互属性和几何属性等(He等,2022)。室内场景中的家具对象有着各自的类型(如椅子、桌子和床)和属性(如形

状、纹理),Yu等人(2021)提出了一个混合现实推荐系统。如图8所示,该工作考虑了家具类型信息和家具之间的空间信息,以互动的方式为用户提供上下文感知的家具推荐。



图8 AR家具推荐(Yu等,2021)
Fig. 8 AR furniture recommendation(Yu et al. , 2021)
((a) scene input;(b)object semantics(the first row) and spatial relationships(the second row))

对象语义另一个常用方向是用户对象,关注用户的职业、偏好和权限等信息。Rumiński和Walczak(2014)提出了一种上下文增强现实语言(contextual augmented reality language,CARL),用于解决交互式AR场景的动态合成问题。CARL考虑了用户偏好和特权等对象语义上下文信息,实现了复杂AR场景的动态建模与组合,并能够根据上下文和用户偏好来显示不同类型的信息。

对象语义易于获取且具有稳定性,通常不会发生变化。对象语义以独立个体的视角,提供了物体和用户对象的本质信息,是不同对象之间差异性的

重要评判标准。

3.1.3 空间关系

空间关系是场景上下文信息中最基础且最重要的类型,是研究者们最常用的场景上下文信息。空间关系是指场景中各种物体、地点或对象之间的相对位置、角度或排列方式等数值信息。在该类别的讨论中,本文将从3个方面简述空间关系:基础空间关系、细微场景空间信息以及真实空间信息。

基础空间关系包括:距离、位置、方向、角度和拓扑关系(Koulieris 等,2014)等,如图9所示。在场景语义类型中,Tahara 等人(2020)在提取场景空间关系时,使用定向边界框(oriented bounding boxe, OBB)(Gottschalk,2000)将物体进行划分,考虑了边界框之间的简单方向关系(前后左右)与物体中心点之间的距离信息。基础空间关系还多用于物体操作相关的工作中,如三维场景中的选择和移动技术、室内家具布置问题等。Wang 等人(2021)将上下文信息应用于室内虚拟现实场景中,提出了一种室内对象选择和移动的新方法。他们重点关注用户与室内家具之间的交互,考虑了室内场景中的空间、群组、从属和对象语义等多种上下文信息。

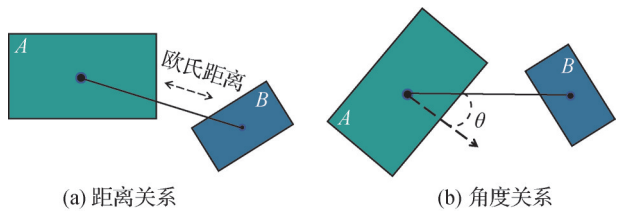


图9 基础空间关系
Fig. 9 Basic spatial relationship
((a) distance relation;(b) angular relation)

一些研究者的工作则更关注细微场景的空间信息,偏重研究物体对象微小部位的信息。Bichlmeier 等人(2007)提出了一种新的医学可视化方法,使用增强现实修改区域透明度。其可视化效果由活体皮

肤表面上的细微空间信息所决定,如皮肤碎片表面周围曲率、与相机间的夹角以及视线与皮肤表面的距离等。Li 等人(2023)则针对焦点和上下文进行渲染,根据表面法线和视角方向设置医学透视的不透明度。

在AR场景环境中,真实世界空间信息也是重要的内容。Barreira 等人(2018)研究了增强现实户外阴影的模拟渲染问题,借助场景中物体的位置、经纬度等信息来计算日光照明的模拟,从而生成情景一致的AR虚拟物体阴影。Köppel 等人(2021)开发了一个上下文响应框架来优化AR中的标签放置。其框架考虑了标签的法向量、GPS(global positioning system)位置和移动方向等空间信息,通过上下文实现空间的定位和映射,处理注释标签的遮挡问题,从而支持AR用户的导航和探索。

空间关系蕴含了场景对象的距离、位置、方向和角度等基础信息,以数值的形式描述关系,具有客观和精准的优点。作为场景上下文信息中最基础且最重要的类型,空间关系具有普适性,能够广泛应用于各种研究场景,是研究者们首选的信息类型。

3.1.4 群组关系

本文认为,一定数量紧密相邻的相同类别的对象视为一个群组。如图10所示(Wang 等,2021),VR场景中相邻的厨房用具、植物和座椅等各自被抽象为同一群组。对象群组关系考虑整体视角,如组间关系和群体数量等。Ennis 等人(2008)评估了位置和方向对行人结构合理性的影响,研究了人类对静态场景中的虚拟人群特征的感知情况。在对虚拟人群的研究中,他们从群体视角考虑了人群流量、位置方向、个体与群体的变化等群组关系,取得了不错的效果。Cashion 等人(2013)考虑了室内同种物体的场景密度,提出了一个基于实时场景上下文分析的自动三维选择技术分配框架。群组关系反映了一种对象群体关系,将同类别的对象群看成一个整体,分



图10 群组关系(Wang 等,2021)
Fig. 10 Group relationship(Wang et al. , 2021)

析整体所蕴含的关系、属性和变化趋势等情况。

3.1.5 从属关系

在混合现实场景中,某一类别的对象可能会与另一类别的对象存在着功能或物理依赖关系。从属关系关注场景中不同对象在功能与物理层面所可能存在的依赖和附属等关系。例如,椅子被放置在电脑的正前方,桌椅与电脑间存在着物理依赖的从属关系;而电脑由显示器、键盘和鼠标等对象构成,电脑作为一个整体,与其组件子对象之间存在着功能上的从属关系。通过考虑对象之间的从属关系,可以根据对象的依赖和附属关系将其划分为父子对象。通常,父对象可以有多个子对象,而子对象只能有一个父对象。

从属关系常用于场景结构关系的理解,也可用

于场景中三维对象的选择工作。Fisher等人(2011)使用图内核(Vishwanathan等,2010)方法来描述场景中的结构关系,在场景的关系图中考虑了两种从属关系:1)父子关系。若一个对象由其他若干对象组成,则该对象作为父对象,组件对象作为子对象,如父对象电脑由子对象显示器和键盘组成。2)表面接触关系。即考虑两个对象物理表面是否接触。Wang等人(2021)借助依赖关系对室内三维对象的选择进行了优化。如图11所示,该工作分别从卧室中的床和床头柜、客厅中的沙发和凳子、休息室中的咖啡桌和坐垫中抽象出从属关系。

从属关系描述对象对之间的相互关系,研究人员可使用从属关系构建场景对象的关系图,优化研究问题。



图 11 从属关系实例(Wang等,2021)

Fig. 11 An example of dependence relationship(Wang et al. , 2021)

3.1.6 运动信息

不同类型场景中的各类对象存在着丰富的运动信息。根据运动信息的泛用性,本文将其划分为基础运动信息和特殊运动信息。

基础运动信息是指对象在场景中进行位移、转向等位置和方向上的连续变化时所带来的各类信息,例如场景对象的运动速度、方向和轨迹等信息。Cashion等人(2013)在物体选择工作中考虑了光标的实时移动速度。Ennis等人(2008)在对虚拟人群的研究工作中考虑了人群移动的方向与趋势。基础运动信息易于获取、分布广泛,普遍存在于各类场景中,具有较好的泛用性,可以应用于多种研究场景。

而特殊运动信息是指场景对象由于个体属性与场景类型所特有的活动信息。例如,在聚会场景中,研究者可以利用说话的时间、交谈的频率、选择的对象等信息进行用户分析。Fujita等人(2013)设计了一种适合聚会的增强信息环境,使用地板显示器、墙

壁显示器和扬声器等多种环境信息刺激对话。该工作测量了参与者的非语言线索,考虑用户的说话时间、串话时间和手势的数量等活动信息,从而对用户的对话状态进行估计判断。而在使用手势的场景中,手势活动蕴含着重要的信息量。Rivu和Burschka(2019)对手势的活动信息进行研究,在动态环境下实现了具有上下文感知能力的三维可视化。如图12所示,当用户拿起茶杯时,借助用户的手势与手部动作信息,预测下一个可能的操作轨迹。特殊的运动信息取决于场景和对象的特质,在不同类型的场景任务下有着不小差异,在不同研究场景中应因地制宜地选择不同的运动信息。

运动信息描述了场景对象的动态信息,是一种新型的信息类型。基础运动信息易于获取且分布广泛,在各种应用场景中具有良好的泛用性。同时,研究者可根据场景研究对象的特性,考虑其所蕴含的特殊运动信息。总体而言,运动信息涵盖范围广泛,

可拓展空间宽阔,有着可观的发展潜力。



图12 运动信息示例(Rivu和Burschka,2019)

Fig. 12 An example of motion information
(Rivu and Burschka,2019)

3.1.7 分类法总结

在对6种上下文信息类型进行研究后,发现空间关系出现频率最高,其次是对象语义。空间几何信息和对象自身的属性构成了场景中最基础的上下文信息,蕴含丰富的信息量且易于获取。当研究者需要考虑研究对象几何关系、属性类别等信息时,空间关系和对象语义是最佳的选择。

场景语义以整个场景的视角,提供场景的构造、元素构成等丰富的语义信息,常用于研究场景识别、分析、重建等场景级别的研究工作中。若研究者需要将视野扩展到整个场景所蕴含的信息上时,通常会选择场景语义。调查发现,场景语义通常单独出现。在未来的工作中,可以考虑将场景语义和其他上下文信息类型相结合。

群组关系和从属关系蕴含着多个对象之间的关系,往往应用于一些特定需求的多对象工作中,如三维对象的选择、虚拟人群的研究等。而运动信息方向的研究较少,有着很大的发展空间。

研究者在考虑场景上下文信息时,可优先考虑空间关系和对象语义,根据场景研究内容判断是否要运用场景语义,并结合研究的需求和目标灵活选择群组关系、从属关系和运动信息等上下文信息。

3.2 上下文知识库的构建

为了探究不同混合现实系统中上下文知识库的构建方式,本节建立了新的分类法,如表2所示。将

知识库定义为场景中各个元素自身的属性和彼此之间的关系,如AR中的虚拟物体具有GPS位置等物理属性(Köppel等,2021),房间内的虚拟代理与各种物品存在“支撑依附”的关系(Lang等,2019)。

系统中知识库的构建方式从是否有用户介入的角度出发。用户介入下构建的知识库不依赖现有的信息数据库,一般通过研究人员对场景物体关系的抽象分析建立;无用户介入类型是一种客观场景信息到主观场景元素关系的学习转化,常依赖于已存在的原始场景上下文信息,如利用Google 3D Warehouse数据库分析或直接将空间信息作为输入。通过分析发现,40%($n=14$)的文章属于用户介入类型,通过手动设计数据库构建相应的模型并在限定条件下对场景进行实验。60%($n=21$)的文章属于无用户介入类型,这些系统通过建立自动化的模型识别上下文信息并进行处理。

3.2.1 用户介入角度

用户介入类型的论文主要集中在对房间物品关系知识库的建立和抽象上,如Irawati等人(2006)率先对物体之间的关系(如是否可放置物品)通过结构化查询语句建立数据库,优化了家具摆放问题。Rumiński和Walczak(2014)建立了结合上下文的增强现实建模语言CARL,该语言可自主设置上下文信息(如位置、用户偏好等)并在AR环境中建立环绕声来进行实证研究。Fisher等人(2011)对房间内物品类别进行抽象,建立知识库识别相似场景并进行重建。Rivu和Burschka(2019)针对常见房间类型下的物体构建了知识库作为场景动作的指导。Wang等人(2021)提出房间内物体的群组关系和从属关系,通过手动构建上下文知识库增强人机交互的体验感。

对于无用户介入类型,Bichlmeier等人(2007)通过建立皮肤的透明度模型来增强在医疗手术中的视觉。后续Gras和Yang(2019)以及Li等人(2023)在医疗AR注释的遮挡问题中提出不同模型并进行了进一步实验,但并未提出新的知识库构型。Chen等人(2014)将基于上下文的分类表述为一个通过图切割优化来解决的马尔可夫随机场问题(Li,1994),借助数据库中学习到的上下文关系来优化室内场景重建。He等人(2022)将大量的房间场景作为数据库进行学习,通过点云提取物体的相关性和布局模式进行房间重建。

表2 上下文信息获取方式和系统技术的分类
Table 2 Classification of contextual information acquisition methods and system technologies

上下文获取	场景类别		总和
	AR	VR	
人为	Irawati 等人(2006)	Ennis 等人(2008)	14
	Joseph 等人(2013)	Dehmeshki 和 Stuerzlinger(2008)	
	Rumiński 和 Walczak(2014)	Fisher 等人(2011)	
	Rumiński(2015)	Yu 等人(2011)	
	Li 等人(2022)	Ma 等人(2018)	
		Rivu 和 Burschka(2019)	
		Wang 等人(2021)	
		Reitmayr 和 Schmalstieg(2005)	
	Bichlmeier 等人(2007)		
	Kalkofen 等人(2009)		
非人为	Fujita 等人(2013)	Oh 等人(2006)	21
	ElSayed 等人(2016)	Tutenel 等人(2011)	
	Barreira 等人(2018)	Merrell 等人(2011)	
	Lang 等人(2019)	Cashion 等人(2013)	
	Gras 和 Yang(2019)	Koulieris 等人(2014)	
	Tahara 等人(2020)	Chen 等人(2014)	
	Köppel 等人(2021)	He 等人(2022)	
	Yu 等人(2021)	Rasla 和 Beyeler(2022)	
	Li 等人(2023)		
		Wang 等人(2018)	
总和	18	17	35

通过用户介入构建的知识库对场景上下文关系有更高层次的抽象作用,有助于提高对场景构建中物体关系更深层次的理解。但随着自学习等方法的兴起,自动提取上下文的工作逐渐占据主导地位。这些工作通过学习环境中的上下文信息来构建知识库,但用户介入模型的发展仍然是必要的,场景和物体粒度的进一步抽象是一个重要的研究内容。该方面的工作有助于提升上下文信息库的概括能力,并为各领域的研究人员提供基于上下文信息的辅助方案。

3.2.2 基础技术角度

本文从AR和VR两方面对上述工作进行了二次划分,以探究不同知识库构建方式与MR系统基础技术之间的关系。在VR方向,两种类型的文章数量不相上下。但在AR方向,无用户介入类型的文章比例为67%($n = 12$),仅有33%($n = 6$)的工作中由用户介入构建知识库。可以看到,AR技术中采取自动构建知识库的无用户介入的研究发展迅速,研

究人员通常利用真实世界中现有的上下文信息数据对关系进行学习。Elsayed 等人(2016)将真实世界物体的物理线索作为上下文,提出了不同的视觉模式以增强感知。然而,由用户介入构建知识库方面的工作较少。即使研究时间相对较早,Irawati 等人(2006)在2006年便尝试在AR中建立关系,但近些年来该方向的发展缓慢,产出文章数量上和自动构建知识库类型有着较大的差距。本文认为,用户介入类型的AR应用发展的难点是现实世界的多样性和真实性。现实世界中物体信息多样且真实,导致物体属性的设置变得异常困难。由此,针对AR下的用户介入设置知识库的工作较少。在未来的工作中,本文建议研究人员在AR特定场景下尝试进行人为设置知识库的研究,创新AR场景元素关系的构建方式,开辟新的研究领域。

3.3 应用领域

对于问题3),本文的目标在于确定上下文信息的具体应用场景。如表3所示,本文根据应用领域

以及系统是否为生成式对上述工作进行了划分,这有助于研究者更好地观察不同领域的发展趋势和应用前景。目前使用上下文信息的应用场景可以分为 6 种:辅助引导、AR 注释、场景重建、医疗、物体操作和通用。

其中 27%($n = 9$)的论文围绕场景重建展开,这些工作首先通过抽象研究场景与建立知识库构型来获取目标场景的上下文信息,然后分析上下文并提取场景中的规则进行重建。如 Ennis 等人(2008)通过对静态图像中虚拟人群分布的上下文进行分析,根据场景类型、附近物体存在情况和行走约束等信息重建场景,提高人群形成的感知度。Chen 等人

(2014)通过处理分析房间场景类型的数据库,学习了大量场景中的上下文关系,进而建立分类器来重建输出房间内的场景。

有 21%($n = 7$)的论文集中在通用方法的探索上,如 Cashion 等人(2013)在 VR 选择操作问题上提供了通用方案,Rumiński 和 Walczak(2014)提出 CARL 为上下文信息的运用提供了普适的编程解决方案。

AR 注释和物体操作类型均有 18%($n = 6$)的论文。AR 注释中的工作主要围绕 AR 下虚拟标签的遮挡问题和使用上下文增强可视区域展开:Barreira 等人(2018)通过计算虚拟物体的位置和经纬度等上下文信息来模拟日光并生成遮挡下的阴影;Kalkofen

表 3 应用领域划分
Table 3 Distribution of the articles by the application domains

应用领域	系统类别		总和
	非生成式	生成式	
辅助引导		Joseph 等人(2013) Rasla 和 Beyeler(2022)	2
AR 注释		Kalkofen 等人(2009) Fujita 等人(2013) ElSayed 等人(2016) Barreira 等人(2018) Köppel 等人(2021) Li 等人(2022)	6
场景重建	Tutenel 等人(2011)	Ennis 等人(2008) Fisher 等人(2011) Yu 等人(2011) Merrell 等人(2011) Chen 等人(2014) Ma 等人(2018) Tahara 等人(2020) He 等人(2022)	9
医疗		Bichlmeier 等人(2007) Gras 和 Yang(2019) Li 等人(2023)	3
物体操作	Oh(2006) Dehmeshki 和 Stuerzlinger(2008) Rivu 和 Burschka (2019) Wang 等人(2021)	Lang 等人(2019) Yu 等人(2021)	6
通用	Irawati 等人(2006) Cashion 等人(2013) Rumiński 和 Walczak (2014) Rumiński 等人(2015) Wang 等人(2018)	Reitmayr 和 Schmalstieg(2005) Koulieris 等人(2014)	7
总和	10	23	33

等人(2009)通过启发式规则建立图结构并分离焦点和上下文来增强视觉。物体操作则根据物体所处的上下文环境来增强用户对其操作的过程,如Wang等人(2021)。

此外,还有较少论文围绕医疗($n=3$)和辅助引导($n=2$)展开。医疗领域针对AR场景中皮肤区域的不透明度展开,Gras和Yang(2019)通过对注视点周围上下文信息的获取来决定全AR视图、部分AR视图或人眼视图的使用。辅助引导的工作探索如何使用上下文信息对特殊环境或特殊群体进行辅助,如Rasla和Beyeler(2022)测试了视障人群对空间中边缘或深度上下文信息的使用喜好。

本文还调查了应用的生成式性质,生成式的模型具有自动生成模型目标信息的能力,如生成目标重建场景或AR注释自动生成阴影;非生成式的模型则针对特定的操作和关系,构建数据库满足特定场景要求下的用户需求。通过分析系统搜集的文章,本文发现有70%($n=23$)的论文都提出了生成式模型。如Yu等人(2021)提出了一个混合现实推荐系统,根据上下文为推荐用户提供合适的家具。非生成式的模型则依赖特定的场景规划与需求。如Wang等人(2021)通过将室内物体划分成不同的组来构建从属关系,从而实现了在选择操作时的快速切换。

AIGC(AI-generated content)的发展给生成式系统注入了更多的活力。通过使用知识库中的上下文信息,生成式模型可以对虚拟人群、房间场景、增强视觉和辅助引导等背景下的应用生成目标视图。各个应用领域上的生成式工作发展迅速,但是在关于物体操作领域的生成式混合现实系统的工作仍然较少。这是因为物体的种类丰富,尤其是其可操作属性难以定义,抽象难度大,定义不同场景中不同物体之间的上下文关系尤为困难。本文建议研究者们抽象并提取物体本身和物体之间更有概括性的上下文信息,提出适应复杂真实世界的知识库并生成真实物体之间的关系语义。

4 结 论

本文对混合现实中的上下文信息进行了综述。首先在MR领域的若干数据库进行了详细的检索,然后根据PRISMA方法最终确定了33篇与混合现实

中的上下文信息有关的论文。通过对这些论文的分析研究,本文对场景和上下文信息提出了3个问题,进而在上下文信息类型、知识库的构建和应用领域3种维度进行了讨论,并在每个维度进行了分类学研究。基于不同分类法的结果,本文讨论了场景上下文信息的研究现状,分析了MR场景与上下文信息的关系,并研究了总体的应用方向趋势。研究者可以通过本文提出的分类法,根据工作方向选择恰当的场景上下文信息。

场景上下文信息为实现各个场景中的研究任务提供了丰富的场景信息,在近近年来逐渐受到混合现实场景研究人员的重视。不过,该方向仍处于起步阶段,未来仍然有许多亟需发展的方向,主要包括以下内容:

1)丰富上下文信息类型。VR和AR设备在场景内容的丰富性上逐步提高,导致更多种类上下文语义信息出现。在教育、会议和远程协助等新应用领域下应当考虑使用新型的具有场景特色的上下文信息来提升系统的沉浸感。同时,现有类型的上下文也存在大量可拓展内容,可以考虑使用情感线索和行为趋向等信息增强感知。

2)不同类型上下文信息的结合。大多数的研究工作仅使用了一种或两种场景的上下文信息,并没有充分利用场景所包含的上下文内容。随着场景信息的丰富,未来研究工作应考虑结合分析多种场景上下文信息,如结合空间关系、从属关系与对象语义分析场景对象,提高对场景上下文信息的利用率,从更加全面的角度研究场景内容。

3)知识库构型的革新。目前,在AR和VR中用户介入构建知识库的工作较少,且主要围绕场景物体的选择、操作和可放置等局部属性展开。随着使用场景的增加,现有的知识库构型描述全局上下文关系的能力亟待提升。未来的研究可以考虑结合图结构等方法对现有的知识库构型进行补充,对全局视角下的场景做出描述。

4)生成式MR系统的发展。随着AIGC的兴起,生成式模型发展越来越迅速。结合生成式模型的智能算法和MR系统的沉浸式特点,可以考虑从提高个性化和沉浸式体验的角度出发,从上下文提取出用户位置、偏好等信息,预测用户行为并生成智能化的虚拟体验。

5)多项技术结合。目前,上下文信息的使用仍

然局限于特定场景下,但实际的MR应用需要对复杂环境进行处理。因此,可以考虑结合大模型和深度学习学习方法学习大规模场景数据,提高对上下文的认知程度。对静态场景中的上下文关系和动态场景中的变化内容进行描述,准确把握用户使用情景并提供更真实和沉浸的体验。

参考文献(References)

- Acquisti A, Gross R and Stutzman F. 2014. Face recognition and privacy in the age of augmented reality. *Journal of Privacy and Confidentiality*, 6(2): 1-20 [DOI: 10.29012/jpc.v6i2.638]
- Anthes C, García-Hernández R J, Wiedemann M and Kranzlmüller D. 2016. State of the art of virtual reality technology//*Proceedings of 2016 IEEE Aerospace Conference*. Big Sky, USA: IEEE: 1-19 [DOI: 10.1109/AERO.2016.7500674]
- Barreira J, Bessa M, Barbosa L and Magalhães L. 2018. A context-aware method for authentically simulating outdoors shadows for mobile augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 24(3): 1223-1231 [DOI: 10.1109/TVCG.2017.2676777]
- Bekele M K, Pierdicca R, Frontoni E, Malinverni E S and Gain J. 2018. A survey of augmented, virtual, and mixed reality for cultural heritage. *Journal on Computing and Cultural Heritage*, 11(2): #7 [DOI: 10.1145/3145534]
- Bichlmeier C, Wimmer F, Heining S M and Navab N. 2007. Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality//*2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. Nara, Japan: IEEE: 129-138 [DOI: 10.1109/ISMAR.2007.4538837]
- Bukowski R W and Séquin C H. 1995. Object associations: a simple and practical approach to virtual 3D manipulation//*Proceedings of 1995 Symposium on Interactive 3D Graphics*. Monterey, USA: ACM: 131-ff [DOI: 10.1145/199404.199427]
- Campbell A G, Santiago K, Hoo D and Mangina E. 2016. Future mixed reality educational spaces//*Proceedings of 2016 Future Technologies Conference (FTC)*. San Francisco, USA: IEEE: 1088-1093 [DOI: 10.1109/FTC.2016.7821738]
- Cashion J, Wingrave C and LaViola J J. 2013. Optimal 3D selection technique assignment using real-time contextual analysis//*2013 IEEE Symposium on 3D User Interfaces (3DUI)*. Orlando, USA: IEEE: 107-110 [DOI: 10.1109/3DUI.2013.6550205]
- Chen K, Lai Y K, Wu Y X, Martin R and Hu S M. 2014. Automatic semantic modeling of indoor scenes from low-quality RGB-D data using contextual information. *ACM Transactions on Graphics*, 33(6): #208 [DOI: 10.1145/2661229.2661239]
- Chen L, Tang W, John N, Wan T R and Zhang J J. 2018. Context-aware mixed reality: a framework for ubiquitous interaction [EB/OL]. [2023-10-17]. <https://arxiv.org/pdf/1803.05541.pdf>
- Costanza E, Kunz A and Fjeld M. 2009. Mixed reality: a survey//Lalanne D and Kohlas J, eds. *Human Machine Interaction*. Berlin, Heidelberg: Springer: 47-68 [DOI: 10.1007/978-3-642-00437-7_3]
- De Guzman J A, Thilakarathna K and Seneviratne A. 2019. Security and privacy approaches in mixed reality: a literature survey. *ACM Computing Surveys*, 52(6): #110 [DOI: 10.1145/3359626]
- Dehmshki H and Stuerzlinger W. 2008. Intelligent mouse-based object group selection//*9th International Symposium on Smart Graphics*. Rennes, France: Springer: 33-44 [DOI: 10.1007/978-3-540-85412-8_4]
- Ding W Y and Shan M X. 2023. A systematic literature review of affective experience in virtual reality learning environment. *Software Guide*, 22(9): 32-44 (丁文郁, 单美贤. 2023. 虚拟现实学习环境情感体验研究系统文献综述. *软件导刊*, 22(9): 32-44) [DOI: 10.11907/rjdk.222144]
- ElSayed N A M, Smith R T and Thomas B H. 2016. Horus eye: see the invisible bird and snake vision for augmented reality information visualization//*2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. Merida, Mexico: IEEE: 203-208 [DOI: 10.1109/ISMAR-Adjunct.2016.0077]
- Ennis C, Peters C and O'Sullivan C. 2008. Perceptual evaluation of position and orientation context rules for pedestrian formations//*Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*. Los Angeles, California, USA: ACM: 75-82 [DOI: 10.1145/1394281.1394295]
- Fidalgo C G, Yan Y K, Cho H, Sousa M, Lindlbauer D and Jorge J. 2023. A survey on remote assistance and training in mixed reality environments. *IEEE Transactions on Visualization and Computer Graphics*, 29(5): 2291-2303 [DOI: 10.1109/TVCG.2023.3247081]
- Fisher M and Hanrahan P. 2010. Context-based search for 3D models//*ACM SIGGRAPH Asia 2010 Papers*. Seoul, Korea (South): ACM: #182 [DOI: 10.1145/1866158.1866204]
- Fisher M, Ritchie D, Savva M, Funkhouser T and Hanrahan P. 2012. Example-based synthesis of 3D object arrangements. *ACM Transactions on Graphics*, 31(6): #135 [DOI: 10.1145/2366145.2366154]
- Fisher M, Savva M and Hanrahan P. 2011. Characterizing structural relationships in scenes using graph kernels//*ACM SIGGRAPH 2011 Papers*. Vancouver, Canada: ACM: #34 [DOI: 10.1145/1964921.1964929]
- Fisher M, Savva M, Li Y Y, Hanrahan P and Nießner M. 2015. Activity-centric scene synthesis for functional 3D scene modeling. *ACM Transactions on Graphics*, 34(6): #179 [DOI: 10.1145/2816795.2818057]
- Fu Q, Chen X W, Wang X T, Wen S J, Zhou B and Fu H B. 2017. Adaptive synthesis of indoor scenes via activity-associated object

- relation graphs. *ACM Transactions on Graphics*, 36(6): #201 [DOI: 10.1145/3130800.3130805]
- Fujita K, Itoh Y, Takashima K, Nakajima K, Hayashi Y and Kishino F. 2013. Ambient party room: a room-shaped system enhancing communication for parties or gatherings//2013 IEEE Virtual Reality (VR). Lake Buena Vista, USA: IEEE: 1-4 [DOI: 10.1109/VR.2013.6549436]
- Gösele M and Stuerzlinger W. 1999. Semantic constraints for scene manipulation//Proceedings of 1999 Spring Conference on Computer Graphics. Comenius University: 140-146
- Gottschalk S A. 2000. Collision Queries Using Oriented Bounding Boxes. North Carolina, USA: The University of North Carolina at Chapel Hill
- Gras G and Yang G Z. 2019. Context-aware modeling for augmented reality display behaviour. *IEEE Robotics and Automation Letters*, 4(2): 562-569 [DOI: 10.1109/LRA.2019.2890852]
- Haller M, Drab S and Hartmann W. 2003. A real-time shadow approach for an augmented reality application using shadow volumes//Proceedings of the ACM Symposium on Virtual Reality Software and Technology. Osaka, Japan: ACM: 56-65 [DOI: 10.1145/1008653.1008665]
- He K M, Gkioxari G, Dollár P and Girshick R. 2017. Mask R-CNN//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2980-2988 [DOI: 10.1109/ICCV.2017.322]
- He Y, Liu Y T, Jin Y H, Zhang S H, Lai Y K and Hu S M. 2022. Context-consistent generation of indoor virtual environments based on geometry constraints. *IEEE Transactions on Visualization and Computer Graphics*, 28(12): 3986-3999 [DOI: 10.1109/TVCG.2021.3111729]
- Irawati S, Green S, Billingham M, Duenser A and Ko H. 2006. "Move the couch where?": developing an augmented reality multimodal interface//2006 IEEE/ACM International Symposium on Mixed and Augmented Reality. Santa Barbara, USA: IEEE: 183-186 [DOI: 10.1109/ISMAR.2006.297812]
- Jacobs K and Loscos C. 2006. Classification of illumination methods for mixed reality. *Computer Graphics Forum*, 25(1): 29-51 [DOI: 10.1111/j.1467-8659.2006.00816.x]
- Joseph S L, Zhang X C, Dryanovski I, Xiao J Z, Yi C C and Tian Y L. 2013. Semantic indoor navigation with a blind-user oriented augmented reality//Proceedings of 2013 IEEE International Conference on Systems, Man, and Cybernetics. Manchester, UK: IEEE: 3585-3591 [DOI: 10.1109/SMC.2013.611]
- Kalkofen D, Mendez E and Schmalstieg D. 2009. Comprehensible visualization for augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 15(2): 193-204 [DOI: 10.1109/TVCG.2008.96]
- Kang Z Z, Yang J T, Yang Z and Cheng S. 2020. A review of techniques for 3D reconstruction of indoor environments. *ISPRS International Journal of Geo-Information*, 9(5): #330 [DOI: 10.3390/ijgi9050330]
- Köppel T, Gröller M E and Wu H Y. 2021. Context-responsive labeling in augmented reality//2021 IEEE 14th Pacific Visualization Symposium (PacificVis). Tianjin, China: IEEE: 91-100 [DOI: 10.1109/PacificVis52677.2021.00020]
- Koulieris G A, Drettakis G, Cunningham D and Mania K. 2014. An automated high-level saliency predictor for smart game balancing. *ACM Transactions on Applied Perception*, 11(4): #17 [DOI: 10.1145/2637479]
- Lang Y N, Liang W and Yu L F. 2019. Virtual agent positioning driven by scene semantics in mixed reality//Proceedings of 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Osaka, Japan: IEEE: 767-775 [DOI: 10.1109/VR.2019.8798018]
- Li C Y, Li W W, Huang H K and Yu L F. 2022. Interactive augmented reality storytelling guided by scene semantics. *ACM Transactions on Graphics*, 41(4): #91 [DOI: 10.1145/3528223.3530061]
- Li R Y, Han B X, Li H W, Ma L F, Zhang X R, Zhao Z and Liao H. 2023. A comparative evaluation of optical see-through augmented reality in surgical guidance. *IEEE Transactions on Visualization and Computer Graphics*, 30(7): 4362-4374 [DOI: 10.1109/TVCG.2023.3260001]
- Li S Z. 1994. Markov random field models in computer vision//Proceedings of the 3rd European Conference on Computer Vision. Stockholm, Sweden: Springer: 361-370 [DOI: 10.1007/BFb0028368]
- Loscos C, Drettakis G and Robert L. 2000. Interactive virtual relighting of real scenes. *IEEE Transactions on Visualization and Computer Graphics*, 6(4): 289-305 [DOI: 10.1109/2945.895874]
- Ma R, Patil A G, Fisher M, Li M Y, Pirk S, Hua B S, Yeung S K, Tong X, Guibas L and Zhang H. 2018. Language-driven synthesis of 3D scenes from scene databases. *ACM Transactions on Graphics*, 37(6): #212 [DOI: 10.1145/3272127.3275035]
- Merrell P, Schkufza E, Li Z Y, Agrawala M and Koltun V. 2011. Interactive furniture layout using interior design guidelines. *ACM Transactions on Graphics*, 30(4): #87 [DOI: 10.1145/2010324.1964982]
- Milgram P and Kishino F. 1994. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12): 1321-1329
- Milgram P, Takemura H, Utsumi A and Kishino F. 1995. Augmented reality: a class of displays on the reality-virtuality continuum//Proceedings Volume 2351, Telemanipulator and Telepresence Technologies. Boston, United States: SPIE: 282-292 [DOI: 10.1117/12.197321]
- Miller G A. 1995. WordNet: a lexical database for English. *Communications of the ACM*, 38(11): 39-41 [DOI: 10.1145/219717.219748]
- Moro C, Štromberga Z, Raikos A and Stirling A. 2017. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical Sciences Education*, 10(6): 549-559 [DOI: 10.1002/ase.1696].

- Murphy K M, Cash J and Kellinger J J. 2018. Learning with avatars: Exploring mixed reality simulations for next-generation teaching and learning//Handbook of Research on Pedagogical Models for Next-Generation Teaching and Learning. Hershey, USA: IGI Global: 1-20 [DOI: 10.4018/978-1-5225-3873-8.ch001]
- Nifakos S and Zary N. 2014. Virtual patients in a real clinical context using augmented reality: impact on antibiotics prescription behaviors//e-Health—For Continuity of Care. [s.l.]: IOS Press: 707-711 [DOI: 10.3233/978-1-61499-432-9-707]
- Oh J Y, Stuerzlinger W and Dadgari D. 2006. Group selection techniques for efficient 3D modeling//3D User Interfaces (3DUI'06). Alexandria, USA: IEEE: 95-102 [DOI: 10.1109/VR.2006.66]
- Pickering C, Grignon J, Steven R, Guitart D and Byrne J. 2015. Publishing not perishing: how research students transition from novice to knowledgeable using systematic quantitative literature reviews. Studies in Higher Education, 40(10): 1756-1769 [DOI: 10.1080/03075079.2014.914907]
- Ping J M, Liu Y and Weng D D. 2021. Review of depth perception in virtual and real fusion environment. Journal of Image and Graphics, 26(6): 1503-1520 (平佳敏, 刘越, 翁冬冬. 2021. 虚实融合场景中的深度感知研究综述. 中国图象图形学报, 26(6): 1503-1520) [DOI: 10.11834/jig.210027]
- Rasla A and Beyeler M. 2022. The relative importance of depth cues and semantic edges for indoor mobility using simulated prosthetic vision in immersive virtual reality//Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology. Tsukuba, Japan: ACM: 27 [DOI: 10.1145/3562939.3565620]
- Reitmayr G and Schmalstieg D. 2005. Flexible parametrization of scene graphs//IEEE Proceedings. VR 2005. RealityVirtual, 2005. Bonn, Germany: IEEE: 51-58 [DOI: 10.1109/VR.2005.1492753]
- Rivu S R and Burschka D. 2019. Context-aware 3D visualization of the dynamic environment//Proceedings of 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). San Jose, USA: IEEE: 42-47 [DOI: 10.1109/MIPR.2019.00016]
- Roesner F, Kohno T and Molnar D. 2014. Security and privacy for augmented reality systems. Communications of the ACM, 57(4): 88-96 [DOI: 10.1145/2580723.2580730]
- Rokhsaritalemi S, Sadeghi-Niaraki A and Choi S M. 2020. A review on mixed reality: current trends, challenges and prospects. Applied Sciences, 10(2): #636 [DOI: 10.3390/app10020636]
- Rumiński D. 2015. Modeling spatial sound in contextual augmented reality environments//Proceedings of the 6th International Conference on Information, Intelligence, Systems and Applications (IISA). Corfu, Greece: IEEE: 1-6 [DOI: 10.1109/IISA.2015.7387982]
- Rumiński D and Walczak K. 2014. Dynamic composition of interactive AR scenes with the CARL language//Proceedings of 2014 IISA International Conference on Information, Intelligence, Systems and Applications. Chania, Greece: IEEE: 329-334 [DOI: 10.1109/IISA.2014.6878808]
- Sala N. 2021. Virtual reality, augmented reality, and mixed reality in education: a brief overview//Current and Prospective Applications of Virtual Reality in Higher Education. Hershey, USA: IGI Global: 48-73 [DOI: 10.4018/978-1-7998-4960-5.ch003]
- Sato I, Sato Y and Ikeuchi K. 1999. Acquiring a radiance distribution to superimpose virtual objects onto a real scene. IEEE Transactions on Visualization and Computer Graphics, 5(1): 1-12 [DOI: 10.1109/2945.764865]
- Savva M, Chang A X, Hanrahan P, Fisher M and Nießner M. 2014. SceneGrok: inferring action maps in 3D environments. ACM Transactions on Graphics, 33(6): #212 [DOI: 10.1145/2661229.2661230]
- Speicher M, Hall B D and Nebeling M. 2019. What is mixed reality?//Proceedings of 2019 CHI Conference on Human Factors in Computing Systems. Glasgow, UK: ACM: #537 [DOI: 10.1145/3290605.3300767]
- Stretton T, Cochrane T and Narayan V. 2018. Exploring mobile mixed reality in healthcare higher education: a systematic review. Research in Learning Technology, 26: #2131 [DOI: 10.25304/rlt.v26.2131]
- Szczuko P. 2014. Augmented reality for privacy-sensitive visual monitoring//Proceedings of the 7th International Conference on Multimedia Communications, Services and Security. Krakow, Poland: Springer: 229-241 [DOI: 10.1007/978-3-319-07569-3_19]
- Tahara T, Seno T, Narita G and Ishikawa T. 2020. Retargetable AR: context-aware augmented reality in indoor scenes based on 3D scene graph//2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Recife, Brazil: IEEE: 249-255 [DOI: 10.1109/ISMAR-Adjunct51615.2020.00072]
- Tutenel T, Van Der Linden R, Kraus M, Bollen B and Bidarra R. 2011. Procedural filters for customization of virtual worlds//Proceedings of the 2nd International Workshop on Procedural Content Generation in Games. Bordeaux, France: ACM: #5 [DOI: 10.1145/2000919.2000924]
- Vishwanathan S V N, Schraudolph N N, Kondor R and Borgwardt K M. 2010. Graph kernels. Journal of Machine Learning Research, 11: 1201-1242
- Wang L L, Cao A T, Li Z C, Yang X F and Popescu V. 2018. Effective free field of view scene exploration in VR and AR//2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Munich, Germany: IEEE: 97-102 [DOI: 10.1109/ISMAR-Adjunct.2018.00043]
- Wang M, Ye Z M, Shi J C and Yang Y L. 2021. Scene-context-aware indoor object selection and movement in VR//Proceedings of 2021 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Lisbon, Portugal: IEEE: 235-244 [DOI: 10.1109/VR50410.2021.00045]
- Xu K, Chen K, Fu H B, Sun W L and Hu S M. 2013. Sketch2Scene: sketch-based co-retrieval and co-placement of 3D models. ACM

Transactions on Graphics, 32(4): #123 [DOI: 10.1145/2461912.2461968]

Yu H F, Liang W, Song S H, Ning B and Zhu Y X. 2021. Interactive context-aware furniture recommendation using mixed reality//Proceedings of 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). Lisbon, Portugal: IEEE: 450-451 [DOI: 10.1109/VRW52623.2021.00107]

Yu L F, Yeung S K, Tang C K, Terzopoulos D, Chan T F and Osher S J. 2011. Make it home: automatic optimization of furniture arrangement. ACM Transactions on Graphics, 30(4): #86 [DOI: 10.1145/1964921.1964981]

Yuen S C Y, Yaoyuneyong G and Johnson E. 2011. Augmented reality: an overview and five directions for AR in education. Journal of Educational Technology Development and Exchange, 4(1): 119-140 [DOI: 10.18785/jetde.0401.10]

Zhou F, Duh H B L and Billinghurst M. 2008. Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR//2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. Cambridge, UK: IEEE: 193-202 [DOI: 10.1109/ISMAR.2008.4637362]

作者简介

杨浩中,男,硕士研究生,主要研究方向为虚拟/增强现实。
E-mail: haozhongYang@buaa.edu.cn

汪淼,通信作者,男,副教授,主要研究方向为虚拟/增强现实、可视媒体智能生成。E-mail: miaow@buaa.edu.cn

舒文桐,男,硕士研究生,主要研究方向为虚拟/增强现实。
E-mail: wentongshu@buaa.edu.cn