arXiv:2603.02471v1 [cs.HC] 2 Mar 2026

打破窗口：探索扩展现实中的网页空间分解

Chenyang Zhang 佐治亚理工学院 亚特兰大，佐治亚州，美国
chenyang.zhang@gatech.edu

Tianjian Wei 佐治亚理工学院 亚特兰大，佐治亚州，美国 twei79@gatech.edu

Haoyang Yang 佐治亚理工学院 亚特兰大，佐治亚州，美国
alexanderyang@gatech.edu

Mar Gonzalez-Franco 谷歌 西雅图，华盛顿州，美国
margon@google.com

Yalong Yang*佐治亚理工学院 亚特兰大，佐治亚州，美国
yalong.yang@gatech.edu

Eric J Gonzalez*谷歌 西雅图，华盛顿州，美国
ejgonz@google.com

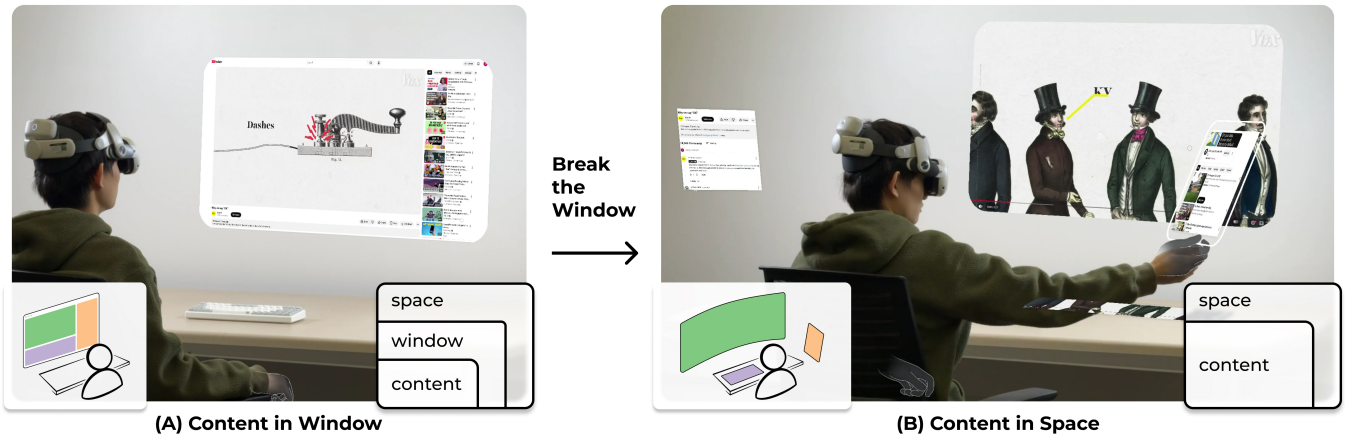


图 1：Break-the-Window（BTW）。传统的扩展现实网页体验会将完整的二维浏览器窗口置于三维空间中，保留一种内容 → 窗口 → 空间 **设计层级，使网页内容仍受窗口限制，无法充分利用空间属性（左）。BTW 重新思考了这种关系，打破窗口本身，并将网页分解为分布在空间中的用户界面片段，从而形成一种**内容 → 空间 **设计层级（右）。这一范式转变使用户能够在空间中更直接地接触网页信息，支持空间组织与具身交互，并促进网页结构语义与周围环境空间语义之间的直接对话。**

摘要

大多数扩展现实网页浏览器仍将网页呈现为单个浮动窗口，把桌面设计假设沿用到沉浸式空间中。我们探索了一种替代方案：打破浏览器窗口，并将网页拆分为混合现实工作空间中的空间用户界面片段。我们介绍了*Break-the-Window*（BTW），这是一种探索性原型，可将实时、功能完整的网页进行空间分解，形成可移动面板，支持在空中和附着于表面的布置，以及直接触控和射线交互。通过与扩展现实从业者开展的形成性研究，以及一项有 15 名参与者参与的探索性质性研究，我们观察到空间分解如何支持分布式注意力和空间意义建构，同时也揭示了协调成本、交互精确操作以及缺乏通用空间 UI 约定等方面的挑战。本研究引发了关于如何突破单窗口范式、为空间计算重新构想网页界面的讨论。

*这些作者对本研究贡献相同，并共同担任资深作者。

本作品采用 CreativeCommons Attribution4.0 InternationalLicense 授权。
CHI EA '26，西班牙巴塞罗那 © 2026 版权所有：所有者/作者。ACM ISBN 979-8-4007-2281-3/2026/04 https://doi.org/10.1145/3772363.3798474

CCS 概念

• **以人为中心的计算** → **混合/增强现实**；用户界面设计； 交互范式；人机交互实证研究。

关键词

空间网页浏览、空间用户界面、网页界面分解、扩展现实工作空间

ACM 参考文献格式：
Chenyang Zhang、Tianjian Wei、Haoyang Yang、Mar Gonzalez-Franco、Yalong Yang 和 Eric J Gonzalez。2026年。打破窗口：探索扩展现实中的网页空间分解。载于*2026*年计算机系统中人为因素会议扩展摘要（*CHI EA '26*），

2026年4月13日至17日，西班牙巴塞罗那。ACM，美国纽约州纽约市，6页。https://doi.org/10.1145/3772363.3798474

1 引言

网页浏览仍是最常见的数字活动之一，但如今的扩展现实网页体验大多仍沿用一种过时的交互模式：一个漂浮在三维空间中的二维浏览器窗口。这种以窗口为中心的范式植根于 PC/WIMP 传统 [10]，保留了内容 → 窗口 → 空间 层级，并限制了用户在处理多方面的网页内容时利用空间组织、具身交互和空间记忆的能力。

新兴的空间计算平台能够将信息分布在周围环境中，并通过手势、注视和射线交互等具身输入来访问信息 [9]。以往研究探索了扩展现实中媒体和信息的空间呈现 [11, 12]，以及基于相关性或上下文的自适应放置策略 [6, 13]。然而，扩展现实中的常规网页浏览仍大多受限于单窗口隐喻：整张网页被放置在空间中，成为平面，而非被重新构想为空间结构。

在这项工作中，我们提出这样一个问题：**如果我们打破浏览器窗口本身，会怎么样？**我们不再将网页视为不可分割的表面，而是探索将其拆分为功能性 UI 模块，并将这些模块分布到空间工作区中，使网页中已有的结构语义（例如主要内容、控件和外围信息）能够直接与扩展现实的空间语义相结合（图 1）。

我们介绍了***Break-the-Window*** **(BTW)**，这是一种将这一理念付诸实践的定位式研究原型。BTW 会将功能完整的网页拆分为多个空间面板，用户可以在以桌面为中心的混合现实工作区中放置、移动、调整大小并重新排列这些面板。系统支持空中布局和附着于表面的布局，也支持直接触控和射线交互，同时保留现有网页的功能，无需重新设计或修改底层网页界面。

为了深入探索这一主题，我们与六位扩展现实从业者开展了形成性研究，考察专家如何构想网页拆分、空间角色以及放置与交互之间的关系。随后，我们与 15 位参与者开展了定性研究，让他们体验 BTW，并以桌面网页浏览和单窗口扩展现实浏览器作为参照。研究结果表明，打破窗口会将浏览从与单一视图交互转变为置身于物理—数字融合工作区：注意力和交互分布于空间之中，布局成为一种语义资源，而物理锚定既能稳定交互，也会使交互变得复杂。与此同时，参与者也指出了协调成本、具身操作精度、生产力权衡以及缺乏通用空间 UI 约定等方面的摩擦。

综上，这张海报将 BTW 作为一个具体的设计原型，借此探讨在单窗口范式之外，网页浏览的空间 *UI* 语法可能是什么样。

2 相关工作

先前研究探索了在不同设备和空间环境中重新分布界面元素。增强现实增强型小部件和移动端到增强现实的技术将智能手机内容延伸至周围



图 2：我们在形成性研究中，对九个常用网站进行空间分解和布局标注的示例。这些网站涵盖购物、导航、媒体和内容创作等多种浏览目的。

空间，以支持精确操作、比较和书签标记 [5, 17]。XDBrowser [15] 和WinCuts [16]等跨设备系统展示了用户如何在不同屏幕间拆分和复制界面区域，揭示了细粒度碎片化与同步的优势。在沉浸式环境中，近期研究考察了如何在混合现实中明确呈现多个视图之间的关系，并综合提出了一个视觉关联技术设计空间，以支持空间分布式视图之间的协调与意义建构 [14]。总体而言，这些方法凸显了打破单体式界面的价值，但其重点主要在于分布和关联，而非在扩展现实中重构网页内容本身。

空间计算领域的研究进一步将空间定位为一种组织性和语义资源。Spatial- strates [4] 和 InteractionAdapt [7] 等系统支持可组合的跨扩展现实工作空间，并基于物理可供性实现布局适配。关于大型显示器和多显示器环境的研究表明，实体导航和空间分区能够提升性能与周边感知 [3, 8], ; 行业指南也倡导在扩展现实中将内容分布到空间面板中 [1]。Apple 的 visionOS 等商业系统引入了空间浏览，将网页扩展为沉浸式呈现，同时在很大程度上保留页面作为统一表面的形式 [2]。基于这些方向，BTW 探讨如何将功能性网页分解为按空间排列的用户界面片段，并研究这些空间化组件如何通过其在混合现实工作空间中的位置、关联和持久性获得新的语义含义。

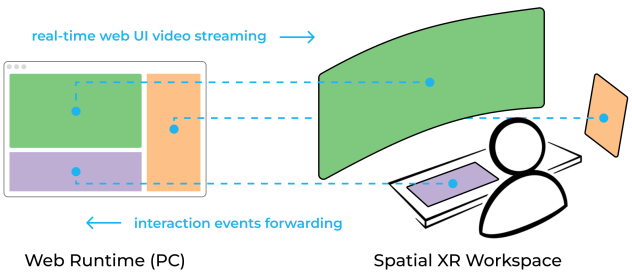


图 3：BTW 数据流概览。实时网页在标准浏览器中渲染，并以多个空间面板的形式流式传输到扩展现实中；与此同时，扩展现实中的用户交互会实时转发回原始网页。

3 形成性研究与设计原理

为了奠定空间网页浏览探索的基础，我们开展了一项形成性研究，考察扩展现实从业者如何分解现有网页并对其进行空间排列。六名经验丰富的扩展现实研究人员对九个常用网站的屏幕截图进行了标注，将每个页面划分为功能区域，并在出声思考的同时将这些区域分配到不同的空间位置（例如附着于表面、空中中央区域、空中侧边区域）。在各个网站中，参与者大体保留了熟悉的网页结构，同时根据功能角色和空间注意力重新分配组件（图 2），从而揭示出三种反复出现的模式：

R1. **角色感知分解**。参与者将网页视为由多个功能区域（例如主要内容、控件和上下文）组成的集合，而非不可分割的屏幕，因此倾向于采用粗粒度且语义明确的区块。R2. **有目的的空间角色**。空间布局反映了注意力与交互需求：空中中央空间用于主要内容的观看，表面用于频繁或精确的交互，外围区域用于次要内容。R3. **布局与交互耦合**。参与者期望交互方式与空间布局相匹配，将表面与直接触控、空中内容与间接输入联系起来，并且常通过移动内容来改变交互方式。

4 原型设计：Break-the-Window（BTW）

我们开发了 *Break-the-Window*（BTW），这是一个面向 Meta Quest 3 的探索性空间网页浏览原型。BTW 将实时网页分解为多个空间面板，这些面板可以放置在混合现实工作空间中并与之交互，同时保留原网页的功能。

4.1 SpatialWeb 分解与具身交互

为在不修改网页代码的情况下支持角色感知分解（R1），BTW 采用像素镜像方法。未修改的桌面浏览器会实时渲染网页，并将其视觉输出串流至扩展现实中。多个空间面板展示同一网页的不同裁剪区域，而用户在扩展现实中执行的交互（例如点击、拖动）会传回原始网页。这样便能将一个网页分解为多个面板，同时保持功能同步，使 BTW 成为灵活的原型，而非经过全面重新设计的浏览器。

BTW 的前端支持灵活的空间放置以及放置与交互之间的耦合（R2、R3）。面板可以自由定位、调整方向和缩放，也可以悬置在空中或附着于实体表面。用户通过抓取面板边缘来操控面板，面板会吸附到附近的水平表面，以支持基于桌面的交互。交互通过两种互补的技术实现：对附近面板使用直接触控，对远处面板使用射线交互。两种技术均支持标准网页交互，并会根据可触及性自动处理二者之间的切换，使用户能够根据任务需求赋予面板不同的空间角色，而不受固定布局的限制。

4.2 空间布局与代表性网页

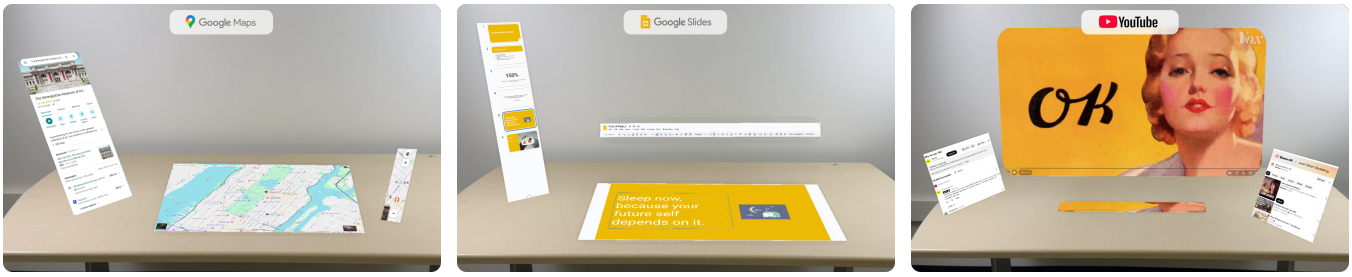
BTW 提供初始空间布局，将其视为设计假设，而非规定性配置。这些布局保留了原始网页中熟悉的结构分组，同时根据访问和交互需求区分空间角色：交互频繁或对精确操作要求较高的组件通常放置在更靠近用户或桌面的位置，而主要内容或供参考的内容则置于空中。参与者在使用过程中可以自由调整这些布局。

我们将 BTW 应用于三个常用网页——谷歌地图、Google 幻灯片和 YouTube，以探索空间分解如何呈现在不同的网页活动中。这些示例涵盖了连续与离散交互、观看与操控，以及低精度与高精度要求。**谷歌地图**将地图画布、信息面板和控件分开，并把地图放在靠近桌面的位置，以支持频繁拖动。**Google 幻灯片**将幻灯片画布、缩略图和工具栏分解开来，并优先将需要精确交互的内容放在靠近用户的位置。**YouTube**将观看区域、控件、评论和推荐内容分开，把视频播放器置于空中中央，同时将交互频繁的面板放在靠近身体的位置。总而言之，这些示例说明，BTW 能够支持适应任务需求的空间组织，而无需针对单一领域定制布局。

5 用户研究

我们开展了一项探索性定性研究，以考察人们如何与扩展现实中的空间分解式网页内容进行交互并理解其含义。研究重点关注交互策略和空间组织，而非性能或效率。

15 名参与者使用*Break-the-Window (BTW)*，并以两种体验参照为对比——桌面浏览器和单窗口扩展现实浏览器——在三个具有代表性的网站（谷歌地图、Google 幻灯片、YouTube）中任选两个完成基于任务的活动。研究在桌面工作区内的透视式虚拟现实环境中开展，参与者可以将空间化网页面板与键盘、纸张和笔等实体工具结合使用（图 5-A）。完成任务后，参与者接受了半结构式访谈。研究人员采用反思性主题分析对访谈数据进行分析，以识别空间组织、注意力和交互方面反复出现的模式，这些模式构成了下文报告的研究发现的基础



（a）谷歌地图 （b）Google 幻灯片 （c）YouTube

图 4：使用 *Break-the-Window*（BTW）进行空间化的三个代表性网页。

。有关研究方案、任务和定性分析过程的更多细节，请参阅补充材料。

6 项关键发现

我们的用户研究揭示了五种反复出现的模式，展现了参与者在 BTW 中如何组织和交互空间分解的网页内容。这些发现共同表明，将浏览器窗口拆分后，浏览便转变为一种空间化、具身化的工作空间，同时也带来了新的交互阻力。

F1. 注意力与交互分散于空间之中。不同于切换标签页和滚动，BTW 让多个面板同时可见且彼此分隔。参与者借助空间维持周边感知，同时专注于主要视图；他们常说，这样可以“屏蔽其他信息”，而不必将其关闭（P6）。这一转变与具身导航密切相关：用户会调整头部、双手和躯干的朝向，在不同内容区域之间移动（图 1-B）。与此同时，分散内容也增加了协调成本：关注偏离视线中心的面板需要付出体力，周边更新更容易被忽略，反复转身和伸手也会累积疲劳。参与者采用了距离相关的输入策略——近处面板使用直接触控，远处面板或需要精确操作时使用射线——但这两种方式都容易受到目标过小和意外触发的影响。

F2. 空间成为语义基底。参与者认为，空间布局不仅仅是视觉编排：位置还传达了功能角色、优先级和相互关系（图 5-bc）。近处的面板被视为“活动状态”，需要频繁交互；远处的面板则逐渐成为背景。随着时间推移，参与者建立了位置与功能之间的联系（例如，“记住空间，而非内容”，P9），并借助空间记忆检索信息。然而，放置不当的代价不小：面板看起来过近或过远，或频繁移动，都会打乱预期，导致参与者犹豫。

F3. 物理锚定既能稳定交互，也会使交互变得复杂。参与者经常借助物理隐喻（图 5-de），将面板像纸张一样放在桌上，或把表面面板视作“虚拟 iPad”（P13）。锚定到真实表面后，稳定的参照物提升了参与者感知到的控制力和精确操作能力。与此同时，与物理工作空间的耦合也带来了冲突：虚拟面板会遮挡纸张和工具，书写手势会被误解为输入操作，而表面稍有错位就会削弱人们对触控交互的信任。（图 5-f）

F4. 沉浸式体验带来的投入感并不意味着更高的生产力。参与者普遍认为，BTW 新颖且引人入胜，尤其适用于浏览和媒体消费等探索性、视觉化场景。然而，参与者并未一致认为自己因此工作得更快或更高效。对于工具繁多或要求精确操作的任务，具身输入会降低工作效率：小型控件难以选中，射线交互需要谨慎瞄准，而且体力消耗和注意力负担会随着时间推移而增加（图 5-gh）。参与者经常将 BTW 与桌面电脑使用习惯（打字、精确的光标控制、快速切换）进行对比，并认为桌面设备在常规的高精度任务中更加可靠。

F5. 空间网页浏览缺乏通用的用户界面语法。浏览不同网站时，参与者自行摸索布局规则，却一再指出缺少惯例（例如，“东西该放在哪儿没有规则”，P1）。随着工作空间不断扩大，这导致杂乱、与物理空间争夺空间、意外选择，以及脆弱的心智模型。参与者希望有相应机制来维持稳定性（例如锁定布局）、支持清晰易读的层级结构，并提供符合具身操作精度的交互反馈。

7 讨论：迈向网页浏览的空间 UI 语法

综合来看，这些发现表明，当网页界面不再局限于单一窗口时，空间便成为界面的一部分。我们并不提供规定性的指导方针，而是概述了四种结构性压力，它们指向空间网页浏览的新兴设计方向。

D1. 从自由放置到可解读的空间语义。空间自由度鼓励用户将意义外化到布局中，但随着空间和任务的变化，也增加了歧义和不一致性的风险。未来的空间网页系统可能需要轻量级结构，例如功能分组、一致的空间层级结构以及稳定的位置—功能映射，以确保布局在扩展时仍清晰易读。

D2. 从静态面板到自适应交互的表示。Web 用户界面元素假定用户以平面观看方式进行操作，并具备鼠标级精度。在空间工作区中，距离和姿势会持续变化，使固定尺寸和悬停式反馈变得脆弱。空间浏览器可能需要采用距离和模式感知表示（例如，自适应目标尺寸和灵敏度、深度感知反馈，或将精确输入与远距离观看解耦）。

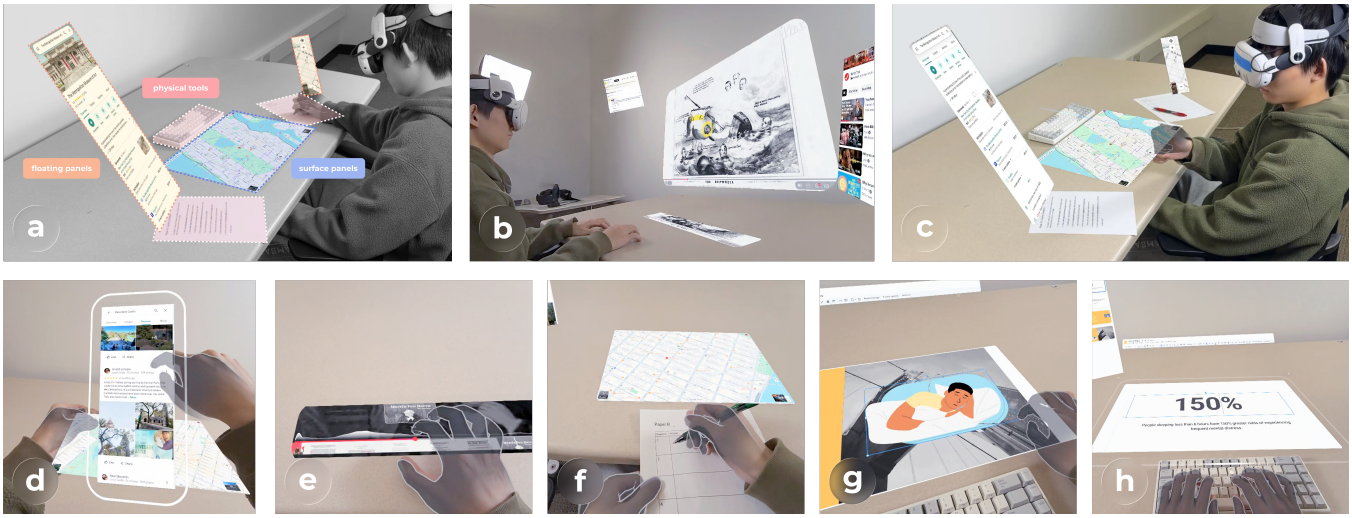


图 5：*Break-the-Window* (BTW) 系统的一组运行时快照。

D3. 从空间扩展到受控的空间容量。尽管扩展现实提供的显示范围比屏幕更大，但参与者仍将空间体验为有限且代价高昂：更大的布局增加了导航耗费和杂乱。这促使人们开发随着时间管理空间容量的机制——例如，分层、折叠，以及在紧凑状态和扩展状态之间进行过渡——同时保留用户的心智模型。

D4. 从界面到混合工作空间。当面板与桌子、纸张和双手共存时，交互会持续发生，边界也变得模糊。系统应更好地协调实体意图与数字意图（例如，通过模式切换或意图推断），并利用环境约束（表面结构、身体朝向）作为稳定锚点，而非冲突来源。

8 结论

这张海报探讨了当网页不再被视为单一窗口，而是作为嵌入混合工作空间中的空间组件时，会发生哪些变化。通过 *Break-the-Window (BTW)* 及探索性研究，我们发现，空间分解能够通过布局支持分布式注意力与意义建构，并引发物理隐喻（例如以桌面为基础的布置）。与此同时，参与者也指出了一些实际摩擦，例如跨空间协调的成本、具身输入下精确操作的挑战，以及虚拟面板与实体工具之间的冲突；他们还反复表达了对更清晰惯例和稳定性的期望。我们并未提出定型的解决方案，而是将 BTW 作为一个具体的研究原型，并以此为起点，邀请社区反馈未来的扩展现实浏览器应如何在空间自由度、可解释结构、交互适应性与易于管理的工作空间边界之间取得平衡。

项目获取。 Break-the-Window（BTW）原型已开源，地址：https://github.com/BTW-XR/Break-the-Window。

致谢

本材料部分基于美国国家科学基金会资助的研究工作，资助编号为 IIS-2441310。

参考文献

[1] Android Developers。2025。空间用户界面。https://developer.android.com/design/ui/ xr/guides/spatial-ui。设计与规划 > 用户界面设计 > XR头显与有线XR眼镜 > 指南。最后更新于2025年5月16日。访问于2026年1月。[2] Apple。2025。visionOS 26 为 Apple Vision Pro 带来强大的全新空间体验。https://nr.apple.com/DA6k2T2Ax1。新闻稿。访问于2026年1月。[3] Robert Ball、Chris North 和 Doug A Bowman。2007。Move to improve：通过促进物理导航提升用户在大型显示器上的表现。载于 *SIGCHI*人机交互系统中的人为因素会议论文集。191–200。[4] Marcel Borowski、Jens Emil Sloth Grønbæk、Peter WS Butcher、Panagiotis D Ritsos、Clemens Nylandsted Klokmose 和 Niklas Elmqvist。2025。Spatialstrates：通过空间超媒体实现跨现实协作。载于 第*38*届*ACM*用户界面软件与技术年度研讨会论文集。1–14。[5] Eugenie Brasier、Emmanuel Pietriga 和 Caroline Appert。2021。面向以智能手机为中心的交互的增强现实增强型小部件。载于 第*23*届移动人机交互国际会议论文集。1–12。[6] Yifei Cheng、Yukang Yan、Xin Yi、Yuanchun Shi 和 David Lindlbauer。2021。Semanticadapt：利用虚拟与物理语义连接，对混合现实布局进行基于优化的适配。载于 第*34*届*ACM*用户界面软件与技术年度研讨会。282–297。[7] Yi Fei Cheng、Christoph Gebhardt 和 Christian Holz。2023。Interactionadapt：面向情境化虚拟现实环境的交互驱动工作空间适配。载于 第*36*届*ACM*用户界面软件与技术年度研讨会论文集。1–14。[8] Jonathan Grudin。2001。数字世界分区：多显示器使用中的焦点与外围觉察。载于 *SIGCHI*人机交互系统中的人为因素会议论文集。458–465。[9] Julia Hertel、Sukran Karaosmanoglu、Susanne Schmidt、Julia Bräker、Martin Semmann 和 Frank Steinicke。2021。基于迭代式文献综述的沉浸式增强现实交互技术分类体系。载于 *2021 IEEE*混合与增强现实国际研讨会（*ISMAR*）。IEEE，431–440。[10] Robert JK Jacob、Audrey Girouard、Leanne M Hirshfield、Michael S Horn、Orit Shaer、Erin Treacy Solovey 和 Jamie Zigelbaum。2008。基于现实的交互：后WIMP界面的框架。载于 *SIGCHI*人机交互系统中的人为因素会议论文集。201–210。[11] Joseph J LaViola Jr、Ernst Kruijff、Ryan P McMahan、Doug Bowman 和 Ivan P Poupyrev。2017。三维用户界面：理论与实践。Addison-Wesley Professional。[12] Tao Lu、Qian Zhu、Tiffany Ma、Wong Kam-Kwai、Anlan Xie、Alex Endert 和 Yalong Yang。2025。自我视角与外部视角、主动与被动：探究视点与导航对空间沉浸感及沉浸式叙事理解的影响

叙事。*arXiv*预印本 *arXiv:2502.04542* （2025）。[13] Weizhou Luo、Anke Lehmann、Hjalmar Widengren 和 Raimund Dachselt。2022。我们应该把它放在哪里？用于协作意义建构的增强现实文档布局与放置策略。载于 *2022*年*CHI*人机交互系统中的人为因素会议论文集。1–16。[14] Weizhou Luo、Rufat Rzayev、Benjamin Russig、Sivanon Visutarporn、Marc Satkowski、Stefan Gumhold 和 Raimund Dachselt。2026。超越链接：探索混合现实中的多视图关系视觉表征。载于*2026*年*CHI*人机交互系统中的人为因素会议论文集 （西班牙巴塞罗那）（*CHI '26*）。美国计算机协会，美国纽约州纽约市。
doi:10.1145/3772318.3791398

[15] Michael Nebeling 和 Anind K Dey。2016 年。XDBrowser：用户定义的跨设备网页设计。载于 *2016* 年人机交互*CHI*会议论文集。5494–5505。[16] Desney S Tan、Brian Meyers 和 Mary Czerwinski。2004 年。WinCuts：操控任意窗口区域，以更有效地利用屏幕空间。载于 *CHI'04* 人机交互计算系统扩展摘要。1525–1528。[17] Jonathan Wieland、Hyunsung Cho、Sebastian Hubenschmid、Akihiro Kiuchi、Harald Reiterer 和 David Lindlbauer。2024 年。Push2ar：利用增强现实提升移动列表交互。载于 *2024* 年 *IEEE*混合与增强现实国际研讨会（*ISMAR*）。IEEE，671–680。