arXiv:2603.01238v1 [cs.HC] 1 Mar 2026

Proscenium：探索大型双层透明显示器上的分层信息体验设计空间

Chen Chen
chechen@fiu.edu 佛罗里达国际大学 迈阿密，佛罗里达州，美国 微软研究院 雷德蒙德，华盛顿州，美国

Michel Pahud
mpahud@microsoft.com 微软研究院 雷德蒙德，华盛顿州，美国

David Brown
dave.bro
wn@microsoft.com 微软研究院 雷德蒙德，华盛顿州，美国

Chuck Needham
chuck@chuckneedham.com
微软研究院
雷德蒙德, WA, USA

Balasaravanan T. Kumaravel
bala@tkbala.com
微软研究院
雷德蒙德, WA, USA

Andrew D. Wilson
awilson@microsoft.com
微软研究院
雷德蒙德, WA, USA

Ken Hinckley
kenneth.p.hinckley@gmail.com
微软研究院
雷德蒙德, WA, USA

Nicolai Marquardt
nicmarquardt@microsoft.com
微软研究院
雷德蒙德, WA, USA

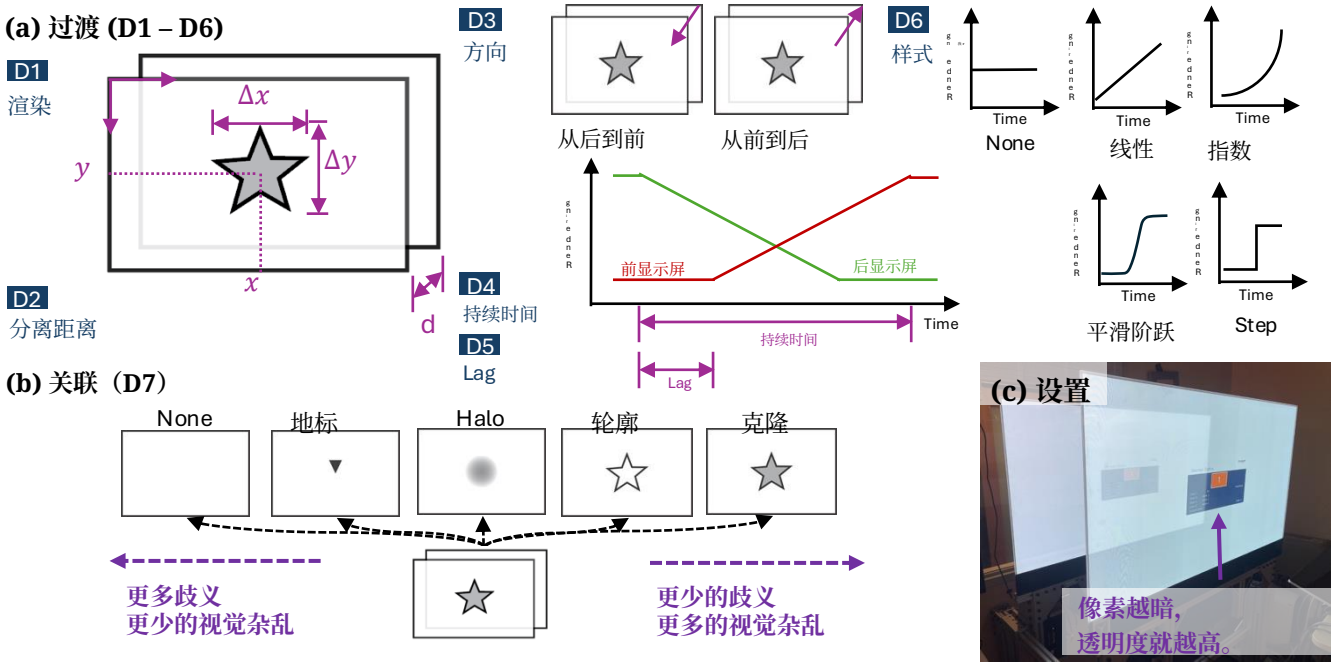

图1：Prosc enium 工作站的设计空间，包括过渡 (a)和关联 (b)。（c）Proscenium 工作站的设置 n.

摘要

分层信息空间是设计直观且引人入胜的交互体验的一种颇具前景的策略。尽管多层显示器通过图层之间的轻微分隔实现有限的深度感知，从而支持颇具前景的交互技术，但如何充分利用分层信息的独特优势来设计体验，仍不明朗。为此，我们介绍了 Proscenium——一套双层大型透明显示器工作空间设置

本作品采用知识共享署名 4.0 国际许可协议授权。*CHI EA '26*，西班牙巴塞罗那© 2026 版权归所有者/作者所有。ACM 国际标准书号
979-8-4007-2281-3/2026/04 https://doi.org/10.1145/3772363.3798304

，其图层间距可调。我们展示了初步的设计空间，重点探讨如何在不同显示器之间对渲染的信息进行过渡和关联，并展示了涵盖六个类别的 14 个推测性体验原型。

CCS 概念
• **以人为中心的计算** → **交互范式**。

关键词
大型显示器、透明显示器、推测性设计、体验原型设计

ACM 参考格式：
Chen Chen、MichelPahud、DavidBrown、Chuck Needham、Balasaravanan T. Kumaravel、Andrew D. Wilson、Ken Hinckley 和 Nicolai Marquardt。2026 年。Proscenium：探索大型双层透明显示器上的分层信息体验设计空间。载于《*2026*年*CHI*计算系统人因会议扩展摘要集》（*CHI EA '26*），*2026*年*4*月*13*日至*17*日，西班牙巴塞罗那。ACM，纽约，美国，9页。https://doi.org/10.1145/3772363.3798304

1 引言
信息分层是设计直观且引人入胜的交互体验的一项关键策略。它能让前景突出显示关键信息，同时保留背景中的情境感知。它还带来了新的设计机会，使实体能够根据现实情境在不同图层之间动态过渡。例如，智能白板 Range [17] 展示了，当用户的物理距离增大时，可以通过叠加透明背景层，将草图内容推至背景。

通过简单调整前景层和背景层的透明度，即可在传统二维显示器上呈现分层信息。然而，在单个显示器上渲染多个图层可能不够直观，并造成视觉杂乱。区分图层的另一种方法是使用**M**ulti-**L**ayer **D**isplays（多层显示，MLD），例如 Pure-Depth [30], ，它采用一台分层液晶显示器，并在第二块屏幕上叠加一层透明显示器。Prema *et al.* [29] 推测了在显示器上呈现物体的可能方法，但如何设计交互体验以充分利用有限的深度仍不明确。显示器之间的间距很小，也可能限制深度感知。使用分层透明显示器创造光学错觉的类似概念也已应用于现代艺术作品中，但交互体验尚未得到探索 [16, 34]。虽然混合现实头显和空间显示器（例如，[12]）能够轻松渲染分层信息，但它们依赖额外设备和/或计算成本，这可能限制其在许多应用中的使用。

为探索如何设计无头显分层信息体验，我们介绍了 *Proscenium*[1], ，这是一种大型透明显示器工作空间设置，采用双层显示器，且显示器之间的距离可调。与使用不透明后置显示器（例如 [30]）不同，叠加两台透明显示器既能提升空间利用率，也能让渲染出的物体在视觉上更自然地融入工作环境。尽管显示层的数量可以超过两层，无限多层则相当于体积显示器，但我们仅聚焦于双层设置——这是理解分层信息设计空间及其交互体验的第一步。未来的研究可以借鉴我们的见解，探索超越双层的体验。借助 Proscenium，我们展示了初步的设计空间，并呈现了涵盖六大类别的 14 个推测性体验原型。尽管本文聚焦于推测性设计和体验原型设计，未来的研究仍可与现实世界的利益相关者合作，对所提出的设计空间进行验证和评估。

[1]"Proscenium"最初是戏剧舞台设计中的一种隐喻，指利用垂直空间来突出主场景 [27]。

2 渲染技术的初步设计空间
Proscenium采用双层透明显示设置，由前层和后层组成，两台显示器之间的间距可调（图1c）。图1a–b展示了我们综合归纳的设计空间，分为过渡和关联两类，有助于理解如何设计跨两台显示器的体验，以及用户如何与之交互。

过渡（D1–D6）。过渡定义了特定渲染内容（如远程用户或共享内容）如何在前层与后层之间跨空间（D1–D3）和时间（D4–D6）进行过渡。值得注意的是，渐隐被视为一种特定类型的过渡，即实体逐渐显现（渐入）或消失（渐出）。

关联（D7）。由于两台显示器之间存在视觉不连续性 [29] ，我们借助关联这一概念，定义了帮助用户在认知上将前层和后层显示器上的两种渲染联系起来的潜在策略。我们推测可能的策略包括无、地标、光晕、轮廓和克隆。图1b展示了后层如何渲染来自前层的信息，以视觉方式帮助观众关联跨图层的信息。在极端情况下，将前层的渲染克隆到后层可以减少歧义，但也可能增加视觉杂乱。

3 Proscenium 工作站
我们通过迭代搭建了一套大型双层透明显示器设置，以便对各种体验和交互进行快速原型设计、迭代和实验。我们使用两台 Planar LookThru LO552 55 英寸透明 OLED 显示器，分辨率为全高清 [28]（图1c）。两台显示器平行安装在一个铝制框架上，彼此间距可调，使我们能够针对不同应用调整 D2。我们没有实现完整的端到端应用，而是使用预录制的视频和图像来实现快速体验原型设计。

4 体验原型设计
以体验原型设计方法为基础， [4],我们介绍了 Proscenium 所支持的 14 种体验，并将其归纳为六大类别。

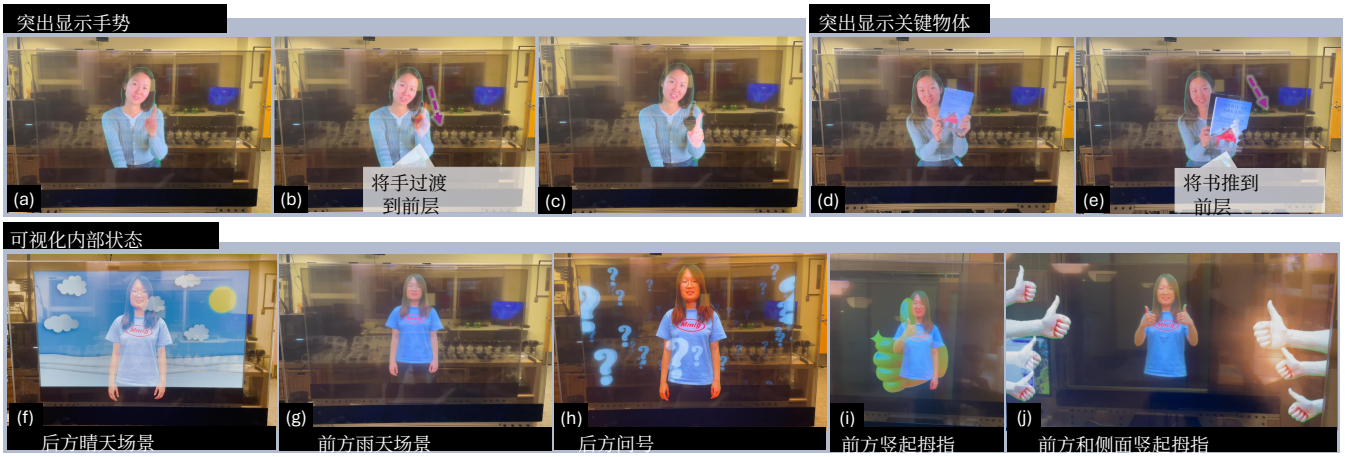

图 2：突出远程临场感中非言语线索的体验，包括 强调伸出的手（a - c）、关键物体（d - e）以及远程用户的内部状态（f - j）。

4.1 突出远程临场感中的非言语线索（E1 - E3）

我们探索如何利用 Proscenium 突显远程双人通信中的关键非语言线索。

E1 使用图层分离来强调伸出的手。我们展示了如何通过图层分离，突出远程用户说话时伸出手部关键手势等行为。图 2a - c 展示了一个示例：远程用户无意识地伸出手势表示 “1”，同时口头说道：“首先，*[...]*”。我们将分离距离（D2）设为 72 cm，约等于一位成年女性的臂长 [15]。在尝试了各种渲染参数（D1）后，我们决定仅使用透明度，因为这不会扭曲渲染后的远程用户的比例。图 2b 展示了一个短暂状态：手同时渲染在前显示屏和后显示屏上，并降低了透明度。手过渡到前层后，我们将后层上的手设为完全透明，以突出手势行为（图 2c ）。

E2 通过图层分离突出关键物体。与 E1 类似，我们对体验进行了原型设计，以配合口头信息突出显示关键物体。我们探索了如何通过图层分离增强实体物品的显著性。图 2d - e 展示了手持书籍如何被提取出来，并逐渐“推”至前显示屏，同时远程用户口头强调道：*“[...] 这是一本很棒的书，我读过 [...]”*。我们使用与 E1 相同的参数（D1 – D6）对这一体验进行了原型设计。当关键物体在物理上逐渐靠近 Proscenium 用户时，可能会营造出更真实、更具吸引力且更专注的交流体验。

E3 可视化内部状态。受二维视频会议应用中可视化用户内部状态的技术启发 [2, 23, 40]，我们展示了 Proscenium 如何在远程视频会议体验中突出显示内部状态。图 2f-j 展示了五种具有代表性的原型体验，其中叠加的动画会随着口述消息同步呈现，并在前显示屏和后显示屏之间过渡。具体而言，图 2f - h 展示了如何从现实生活经验中汲取灵感，在不同显示层上以隐喻方式呈现快乐、悲伤和困惑等心理状态。图 2i - j 展示了如何通过前显示屏上的竖起大拇指动画和多个竖起大拇指动画来增强远程用户的体验。

4.2 通过邻近性促进互动 E4 -E7

第二组体验原型设计探索了渲染信息与 Proscenium 用户之间物理距离的变化，如何促进（或阻碍）用户的互动意愿及参与度。

E4 拉动和推送共享应用设计，以开展协作式头脑风暴。在共享任务空间的视频会议体验中 [5]，我们探讨了任务与远程用户之间，以及任务与 Proscenium 用户之间的邻近性变化如何影响沟通体验。图 3a–e 展示了 Proscenium 如何促进用户与远程参与者协作，共同设计类似 Figma [10] 的智能手机应用。我们允许将共享设计拉动至 Proscenium 用户一侧（图 3d，例如，远程用户试图为 Proscenium 用户分配新的设计任务时），或推送给远程参与者（图 3e，例如，某位 Proscenium 用户试图强调某个设计的要点时）。这有助于用户参与设计讨论，并在渲染信息于社交空间与个人空间之间过渡时，提供相应的观看体验 [13]。实现这种体验需要三个图层：两个图层分别位于靠近和远离 Proscenium 用户的位置，另一个图层则置于两者之间。我们通过略微提高透明度、放大原型应用设计的尺寸并添加细微阴影，营造信息靠近 Proscenium 用户的视觉效果（图 2c–d） [39]。

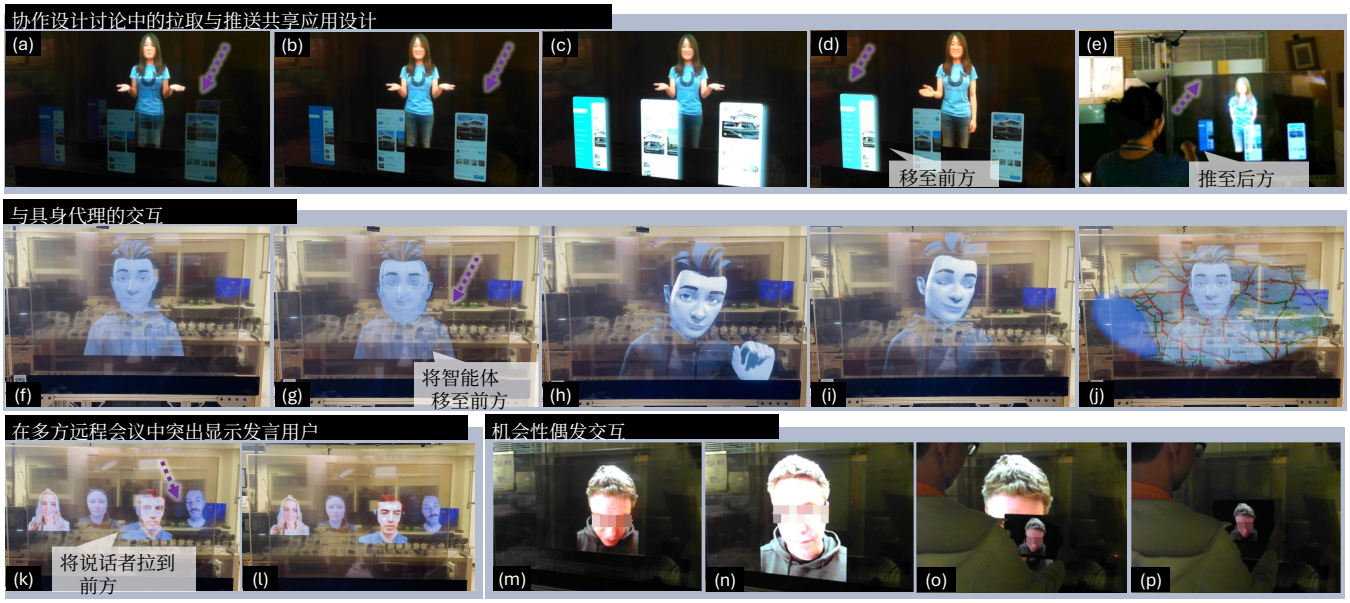

图 3：展示利用图层间隔所带来的邻近性促进交互的体验，包括在协作头脑风暴中拉动和推动共享应用设计（a–e）、与具身代理交互（f–j）、在多方视频会议中突出显示发言用户（k–l），以及机会性偶发交互（m–p）。

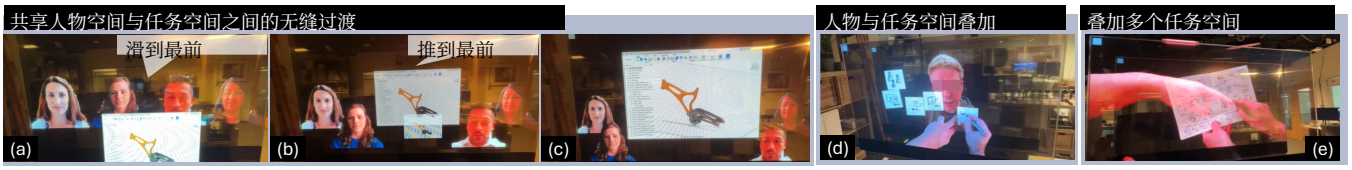

图 4：提供共享人物空间和任务空间管理支持的体验，包括共享人物空间与任务空间之间的无缝过渡（a - c）、人物空间和任务空间叠加（d），以及多个任务空间（e）。

E5 与具身代理的交互。我们展示了改变具身代理与 Proscenium 用户之间的邻近性如何促成更投入的交互。图 3f–h 展示了真人大小的具身代理如何过渡至前层、进入 Proscenium 用户的个人区域，并在 Proscenium 用户接近显示器、进入代理的个人空间时，主动发起问候等交互。

E6 在多方视频会议体验中突出显示发言用户。基于邻近交互的理念 [22]，，Proscenium 可通过突出显示当前发言者来增强多方视频会议体验。图 3k - l 展示了会议参与者开始发言时，如何将其拉近并过渡至前显示屏。缩短正在进行视频会议的发言参与者与 Proscenium 用户之间的邻近距离，同时将说话者从社交空间过渡到个人空间，可以增强用户对当前远程参与者的感知。

E7 机会性偶发交互。Proscenium 支持机会性偶发交互，这类交互通常自发发生，且无需事先规划。为了促进远程用户与 Proscenium 用户之间的此类交互，Proscenium 会将远程用户从公共空间带到社交空间，即将远程用户的视频流过渡至前显示屏（图 3m - n）。偶发交互结束后，Proscenium工作站允许 Proscenium 用户将远程用户移至个人空间，方法是将视频会议转移到个人设备上，例如，平板电脑（图 3o - p）。

4.3 管理人物空间和任务空间（E8 - E10）

Proscenium 可支持在视频会议体验中管理共享人物空间和任务空间 [5, 6]。

E8 共享人物空间与任务空间之间的无缝过渡。 在视频会议体验中，无缝过渡对于共享人物空间和任务空间至关重要 [6]。此体验展示了如何将共享内容推送到后

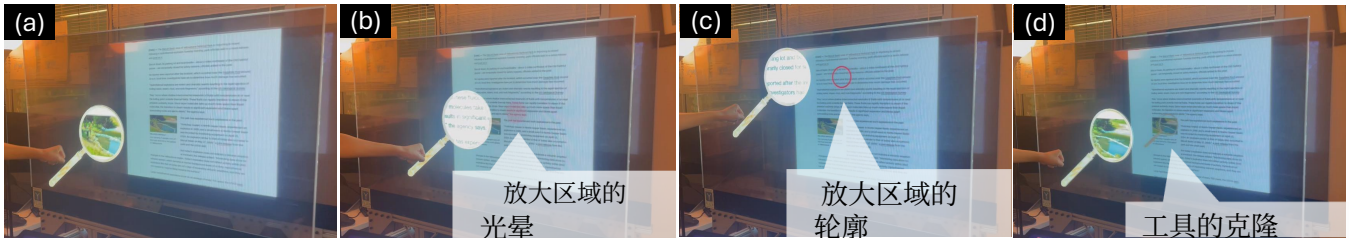

图 5：展示前显示屏上的焦点信息如何通过链接到后显示屏来呈现的体验，采用无（a）、光晕（b）、轮廓（c）和克隆（d）策略。

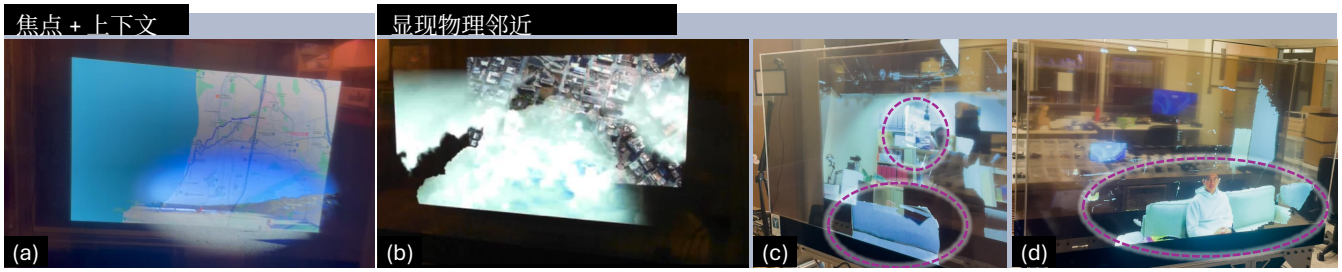

图 6：展示（a）概览 + 细节以及（b - d）物理邻近性的呈现的体验。从前显示屏提取的关键实体以紫色圆圈突出显示。

显示屏（图 4a - c），方法是调整中心位置（x 和 y）与缩放比例（Δx 和 Δy），以确保共享内容始终完全可见。这种细微的过渡让同地的 Proscenium 用户能够感知共享工作空间如何呈现并推送给远程会议参与者。

E9 叠加人物空间与任务空间。在进行某些共享任务时，Proscenium 用户可能还需要留意远程用户。能够在专注于任务空间的同时"读懂"远程会议参与者，已被认为是实现有效远程临场感的关键因素 [6]。图 4d 展示了这样一种体验：将工作空间中的活动（例如用便利贴头脑风暴设计构想）移至前显示屏，同时将远程用户保留在后层。利用物理分隔来区分任务空间与用户空间，能够营造逼真的观看体验，展现任务与用户在同处一地时各自所处的位置，使 Proscenium 用户能够专注于任务，同时留意远程协作者。

E10 叠加多个任务空间。Proscenium 工作站支持多名视频会议参与者在共享工作空间中协作。我们演示了如何通过叠加组合多个独立工作空间。多名远程参与者可以同时与共享工作空间交互，每位参与者的内容都会渲染在独立的显示层上（图 4e）。这为 Proscenium 用户营造了视觉上的分层体验，让远程参与者仿佛身处同一地点。

4.4 分离任务与工具（E11）

Proscenium工作站可用作专为同地会议设计的智能可视化工具。受 TouchTools [14]启发，我们使用两个不同的图层来渲染工具和任务。为了帮助用户在认知上将工具与任务联系起来，我们实现了四种关联 策略，分别是无、光晕、轮廓和克隆，如图 5a–d 所示。

4.5 概览 + 细节（E12）

概览 + 细节 [8] 技术已广泛应用于地图应用中，用户在浏览空间时，可以同时查看地图位置和详细的场景图像。这种体验展示了一种设计：前显示屏渲染场景，后显示屏显示相应的地图位置（图 6a）。用户可以了解特定感兴趣位置的情况，同时保持对其在地图上位置的认知。

4.6 物理邻近性的呈现（E13 - E14）

Proscenium 可以排列并渲染反映物理距离的信息。

E13 用于检查卫星云图的辅助工具。卫星云图能够直观查看特定区域的云量，这对于天气预报、气候研究和灾害管理至关重要。使用 Proscenium 时，前显示屏渲染云层，而卫星视图则保留在后显示屏上，展现了云层与地面之间的物理距离（图 6b）。

E14 2.5D 视频会议体验。Proscenium 可以通过放大并强调关键深度信息，促成 2.5D 视频会议体验。图 6c 展示了如何

书、手以及部分手臂可被实时提取到前显示屏，使 Proscenium 用户在交流过程中注意到关键物体。这种体验类似于 E1 - E2，尽管我们使用实际的人际距离来分隔信息层。图 6d 展示了另一个示例：远程参与者坐在沙发上的场景被提取到前显示屏，而仅将显示器等关键物体渲染在后显示屏上。移除远程参与者身后不太关键的背景信息，可以增强远程临场感体验，使人感觉远程参与者仿佛与 Proscenium 用户实际共处。

5 结论与未来工作

我们介绍并探索了围绕 Proscenium 的交互设计——一款双层大型透明 OLED 显示工作站。基于我们的推测性设计空间，我们快速制作了 14 种体验原型，以探索两块透明显示器的分离方式如何支持未来的交互体验设计。

我们的初步推测性设计和体验原型设计指向了未来工作的四个方向：

首先，尽管本文未包含用户研究（例如，实证评估、专家反馈和/或与替代方案的比较），但未来工作将与现实世界的利益相关者开展设计评审。这将使我们能够迭代完善设计，并进一步探索图1a–b中概述的关键参数。

第二，尽管我们的设计空间主要关注过渡和关联，但未来对 Proscenium 的研究也将探索更广泛的设计参数。例如，利用透视显示屏的优势，未来研究可以考察显示屏摆放位置的影响（例如，用作房间隔断或集成到窗户构件中）。更广泛而言，未来工作还可以探讨分层显示屏与周围物理环境之间的关系——现有的多种传感技术（例如 [1, 3, 7], ）或许有助于实现这一探索。未来工作还可以研究交互故障和意外视觉干扰的影响。

第三，尽管这张海报中展示的体验主要是通过录制和编辑的视频与图像素材制作原型的，我们也希望构建功能完整的端到端应用，以便在真实且生态效度良好的环境中进行评估。

最后一项长期工作是通过增加更多图层，了解其影响并探索相应设计。本文仅将双层设置作为理解分层信息体验设计与交互的第一步，而未来研究可以基于我们的见解拓展至双层以上。尽管 Proscenium 并非全息显示屏，未来工作仍可探索它如何支持更多样化的分层信息体验。

致谢

感谢匿名评审者提供的宝贵意见。我们衷心感谢微软研究院的 Project Transcendence 团队 [33] 给予的支持和反馈。我们还要感谢 Nathalie H. Riche、Kori Inkpen、John Tang、Ed Cutrell、MartezMott、Sean Rintel、Payod Panda、Lev Tankelevitch、Hugo Romat、Gon- zalo Ramos、Jialu Gao、Sneha Gathani、Jieun Kim、Xiajie Zhang，以及微软研究院的 Humphrey Curtis、Zheng Ning、Anna Offenwanger 和 Frederic Gmeiner，在本研究早期阶段提供的真知灼见。

参考文献

[1] Yuvraj Agarwal、Christopher Harrison、Gierad Laput、Sudershan Boovaraghavan、Chen Chen、Abhijit Hota、Bo Robert Xiao 和 Yang Zhang。2020。虚拟传感器系统。美国专利申请 16/591,987。[2] 苹果公司。无日期。使用 Reactions、Presenter Overlay 及其他效果进行 Mac 视频会议。https://support.apple.com/en-us/105117 访问日期：2024年9月24日。[3] Sudershan Boovaraghavan、Chen Chen、Anurag Maravi、Mike Czapik、Yang Zhang、Chris Harrison 和 Yuvraj Agarwal。2023。Mites：面向建筑物的通用感知基础设施的设计与部署。《ACM交互、移动、可穿戴与普适技术汇刊》7卷1期，文章2（2023年3月），32页。https: //doi.org/ 10 1145 3580865 . / Marion Buchenau 和 Jane Fulton Suri。[4] 2000。体验原型设计。载于第三届交互系统设计会议论文集：流程、实践、方法与技术（美国纽约州纽约市）（DIS '00）。计算机协会，美国纽约州纽约，424–433。https: //doi.org/ 10 1145 347642 347802 . / . Bill Buxton。无日期。媒体空间——意义空间——会议空间。[5] 2009 Springer London，英国伦敦， 217 231 10 1007 978 1 84882 483 6_13 – 。https://doi.org/ 。 / - - - - [6] William A. S. Buxton。无日期。远程临场感：整合共享任务空间与个人空间。载于 Graphics Interface '92 会议论文集（加拿大不列颠哥伦比亚省温哥华）。Morgan Kaufmann Publishers Inc.，美国加利福尼亚州旧金山， 123 129 10 5555 155294 155309 – 。https://doi.org/ 。 / 。Chen [7] Chen、Ke Sun 和 Xinyu Zhang。无日期。CapTag：迈向可打印的普适物联网：海报摘要。载于第18届嵌入式网络传感器系统会议论文集（日本，虚拟活动）（SenSys '20）。计算机协会，美国纽约州纽约，669–670。https://doi.org/ 10 1145 3384419 3430410 . / 。Andy Cockburn 、 Amy Karlson 和 Benjamin B. Bederson。2009。概览 + 细节、缩放和焦点+上下文界面的综述。《ACM计算综述》41卷1期，文章2（2009年1月），31页。https://doi.org/10.1145/ 1456650.1456652[9] Steven Dow、Blair MacIntyre、Jaemin Lee、Christopher Oezbek、Jay David Bolter 和 Maribeth Gandy。2005。在整个迭代设计过程中提供绿野仙踪支持。《IEEE普适计算》4卷4期（2005年10月），18–26。https://doi.org/10. 1109/MPRV.2005.93[10] Figma。无日期。Figma。https://figma.com 访问日期：2025年1月19日。[11] Jeff Foster。2014。《绿幕手册：真实场景制作技术》。Routledge。https://doi.org/ 10 4324 9781315770833 . / Looking Glass。无日期。无头显空间显示：人人皆可观看的 D 全息显示屏。[12] https://lookingglassfactory.com 访问日期：2024年9月24日。[13] Edward T. Hall。1990。《隐藏的维度》。Knopf Doubleday出版集团。[14] Chris Harrison、Robert Xiao、Julia Schwarz 和 Scott E. Hudson。2014。Touch-Tools：利用人们对实体工具的熟悉度与技能增强触摸交互。载于SIGCHI计算机系统中的人为因素会议论文集（加拿大安大略省多伦多）（CHI '14）。计算机协会，美国纽约州纽约， 2913 2916 10 1145 2556288 2557012 – 。[15] https://doi.org/ 。 / 。William J. Hennessey、Frank J.E. Falco、Gary Goldberg 和 Randall L. Braddom。1994。性别与臂长：对上肢神经传导参数的影响。《物理医学与康复档案》75卷3期（1994），265–269。https://doi.org/ 10 1016 0003 9993 94 90026 4 . / - ( ) - Homeli。无日期。分层绘图：[16] Nobuhiro Nakanishi 创作的 D 雕塑摄影。https://homeli.co.uk/layer-drawings-3d-sculptural-photography-by-nobuhiro-nakanishi 访问日期：2024年9月25日。[17] Wendy Ju、Brian A. Lee 和 Scott R. Klemmer。2008。Range：通过电子白板设计探索隐式交互。载于2008年ACM计算机支持的协同工作会议论文集（美国加利福尼亚州圣迭戈）（CSCW '08）。计算机协会，美国纽约州纽约，17–26。https://doi.org/ 10 1145 1460563 1460569 . / 。[18] 1976 Wolfgang Kabsch。无日期。确定关联两组向量的最佳旋转的一种解法。《晶体学报》A辑：晶体物理学、衍射、理论与普通晶体学32卷5期（1976），922–923。https://doi.org/10.1107/ S0567739476001873[19] Wolfgang Kabsch。1978。关于确定关联两组向量的最佳旋转解法的讨论。《晶体学报》A辑：晶体物理学、衍射、理论与普通晶体学34卷5期（1978），827–828。https: //doi.org/10.1107/S0567739478001680[20] Alexander Kirillov、Eric Mintun、Nikhila Ravi、Hanzi Mao、Chloe Rolland、Laura Gustafson、Tete Xiao、Spencer Whitehead、Alexander C. Berg、Wan-Yen Lo、Piotr Dollár 和 Ross Girshick。2023。Segment Anything。https://doi.org/10.48550/ arXiv.2304.02643 arXiv:2304.02643 [cs.CV]

[21] Camillo Lugaresi、Jiuqiang Tang、Hadon Nash、Chris McClanahan、Esha Uboweja、Michael Hays、Fan Zhang、Chuo-Ling Chang、Ming Guang Yong、Juhyun Lee、Wan-Teh Chang、Wei Hua、Manfred Georg 和 Matthias Grundmann。2019。MediaPipe：用于构建感知管线的框架。https://doi.org/10.48550/arXiv.1906.08172 [cs.DC] 。arXiv: 1906.08172 [cs.DC] /arXiv. 。[22] Nicolai Marquardt 和 Saul Greenberg。2015。邻近交互：从理论到实践。Morgan & Claypool Publishers。https://doi.org/10.1007/978-3-031- 02208-1[23] 微软。2024。在 Microsoft Teams 会议中使用实时反应表达自我。https://support.microsoft.com/en-us/office/express-yourself-in-microsoft-teams-meetings-with-live-reactions-a8323a40-3d07-4129-934b-305370a36e21 访问日期：2025年1月17日。[24] 微软。无日期。Surface Pro 平板电脑。https://www.microsoft.com/en-us/surface/ devices/surface-pro-11th-edition 访问日期：2025年2月3日。[25] Molly Jane Nicholas、Nicolai Marquardt、Michel Pahud、Nathalie Riche、Hugo Romat、Christopher Collins、David Ledo、Rohan Kadekodi、Badrish Chandramouli 和 Ken Hinckley。2023。Escapement：借助传感器介导的时间抽象，通过视频进行交互式原型设计的工具。载于2023年计算机系统人机交互大会论文集（德国汉堡）（CHI '23 799, 14 10 1145 3544548）。计算机协会，美国纽约州纽约，文章页码。https://doi.org/ 。 / 。[26] 3581119。Orbec。无日期。Femto Bolt 飞行时间相机。https://www.orbbec.com/products/tof- camera/femto-bolt 访问日期：2025年1月19日。[27] Richard Pilbrow。2016。场所的魔力：剧院建筑中一场未获认可的革命。载于表演艺术中心管理。Routledge，96–114。[28] Planar。无日期。Planar LookThru 55 英寸透明 OLED 显示屏。https://www.planar.com/products/transparent-oled-displays 访问日期：2025年1月8日。[29] Vijay Prema、Gary Roberts 和 BC Wuensche。2006。多层显示技术的三维可视化技术。载于新西兰国际图像与视觉计算会议（IVCNZ），第6卷。27–29。https://www.cs.auckland. ac.nz/~burkhard/Publications/IVCNZ06_PremaRobertsWuensche.pdf[30] PureDepth。无日期。PureDepth 多层显示屏。https://puredepth.com 访问日期：2024年9月24日。[31] R00-SG06。2015。BT.709：高清电视标准在制作和国际节目交换中的参数值。https://www.itu.int/rec/R-REC-BT.709 访问日期：2024年9月24日。[32] Nikhila Ravi、Valentin Gabeur、Yuan-Ting Hu、Ronghang Hu、Chaitanya Ryali、Tengyu Ma、Haitham Khedr、Roman Rädle、Chloe Rolland、Laura Gustafson、Eric Mintun、Junting Pan、Kalyan Vasudev Alwala、Nicolas Carion、Chao-Yuan Wu、Ross Girshick、Piotr Dollár 和 Christoph Feichtenhofer。2024。SAM 2：在图像和视频中进行 Segment Anything。https://doi.org/10.48550/arXiv.2408.00714 arXiv: 2408 00714 [cs.CV] . cs.CV 微软研究院。无日期。Transcendence：重新构想协同工作方式。https://microsoft.com/en-us/research/project/transcendence 访问日期：2026年2月21日。[34] Emma Taggart。2018。艺术家在墙上手绘三维视觉错觉，玩转深度与透视。https://mymodernmet.com/math-art-line-drawing- katy-ann-gilmore 访问日期：2024年9月25日。[35] Ultraleap。无日期。Leap Motion 控制器。https://www.ultraleap.com/leap-motion- controller-overview 访问日期：2025年1月19日。[36] S. Umeyama。1991。两组点模式之间变换参数的最小二乘估计。《IEEE 模式分析与机器智能汇刊》 13, 4 1991 376 380 10 1109 34 88573 ( ), – 。https://doi.org/ 。 / 。[37] B Wong。2003。无日期。ActualDepth 多层显示屏的技术可供性。技术报告。为 Deep Video Imaging Ltd 准备的报告，汉密尔顿。[38] Lihe Yang、Bingyi Kang、Zilong Huang、Xiaogang Xu、Jiashi Feng 和 Hengshuang Zhao。2024。Depth Anything：释放大规模无标注数据的力量。载于IEEE/CVF 计算机视觉与模式识别会议论文集。10371–10381。[39] Shumin Zhai、William Buxton 和 Paul Milgram。1996。部分遮挡效应：在三维人机交互中利用半透明。《ACM 人机交互汇刊》3卷3期（1996年9月），254–284。https: //doi.org/ 10 1145 234526 234532 . / 。Zoom。无日期。使用手势识别。https://support.zoom.com/hc/en/article? id=zm_kb&sysparm_article= KB0067842 访问日期：2025年1月17日。[40] 2024

A Proscenium工作站设置的补充细节

本节为第3节提供额外的补充材料，重点介绍Proscenium工作站的设置与实现。

A.1 硬件配置

Proscenium工作站使用两台现有的 Planar LookThru LO55255 英寸透明 OLED 显示器，分辨率为全高清 [28]。透明度由每个像素颜色的亮度决定，具体定义见 BT. 709 [31]，$\alpha = 0.2126r + 0.7152g + 0.0722b$ 标准 [30, 37]。虽然以往涉及多层显示的研究通常使用 LCD 显示屏，但我们选择了透明 OLED 显示器，因为它具有更大的可视角度、更准确的色彩表现，并且在渲染全黑像素时完全透明。Proscenium工作站还配备了第三台显示器，用于快速原型设计和探索不同的体验。虽然我们主要依赖绿野仙踪研究，由研究人员在用户不知情的情况下控制 Proscenium 上的渲染行为，但我们也将 Proscenium 与手部追踪传感器和 RGB-D 摄像头集成，以便针对一些体验开展快速原型设计。

A.2 支持快速体验原型设计的探索工具

我们设计并开发了一款基于 Unity 构建的 Windows 应用程序，用作快速原型设计工具。图7展示了软件设计中的关键组件。在体验原型设计过程中，我们的原型设计工具使设计人员能够快速、迭代地创建各种体验，并尝试设计空间中设想的不同参数（图1），而无需开发完整的端到端应用程序 [9]。研究人员可以配置档案，以定义渲染信息如何过渡、前景与背景信息如何在视觉上关联，以及如何使用传感器数据流。在初步的绿野仙踪研究中，我们的工具需要支持研究人员触发关键动画，以模拟虚拟的交互体验。

在双显示屏配置上渲染信息实体。为了探索 Proscenium 的体验及其设想的设计空间，我们选择并整理了能够体现 Proscenium 交互核心概念的工作空间应用程序（有关所探索体验的深入讨论，见第4节）。我们没有从头开发完整的端到端应用程序，而是收集了一系列图像和视频片段素材。随后，我们的原型设计工具根据用户的预配置渲染所选的多媒体素材。使用离线预录的多媒体素材制作交互体验原型，使我们能够快速探索各种观看体验、交互技术和应用场景，同时随着原型保真度的提升，迭代完善设计。为了在透明显示器上呈现信息，并实现过渡和关联的概念，我们使用两台配备定制着色器的 Unity 摄像机，在每台透明显示器上渲染信息；此外，还使用另一台摄像机在第三台显示器上渲染交互式图形用户界面控制器。

图层间的过渡。 在各种推测的体验中，过渡被确定为值得探索的一个关键维度。由于设计维度多样，且可能随时间和空间变化，利用图像和视频资产对过渡行为进行建模和实现颇具挑战。Proscenium 原型构建工具的设计应使研究人员能够快速构建图1所示过渡样式的原型。为此，我们设计了一条过渡曲线，用于定义渲染效果如何随

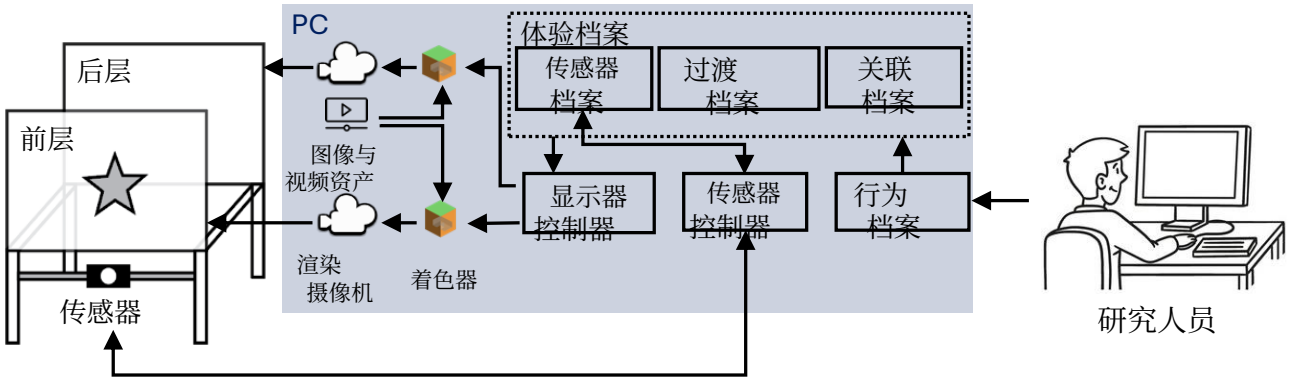

图7：Proscenium工作站的软硬件实现，设计师和/或研究人员可借此快速制作各种体验原型，并试验与图1所述过渡和关联相关的不同参数。

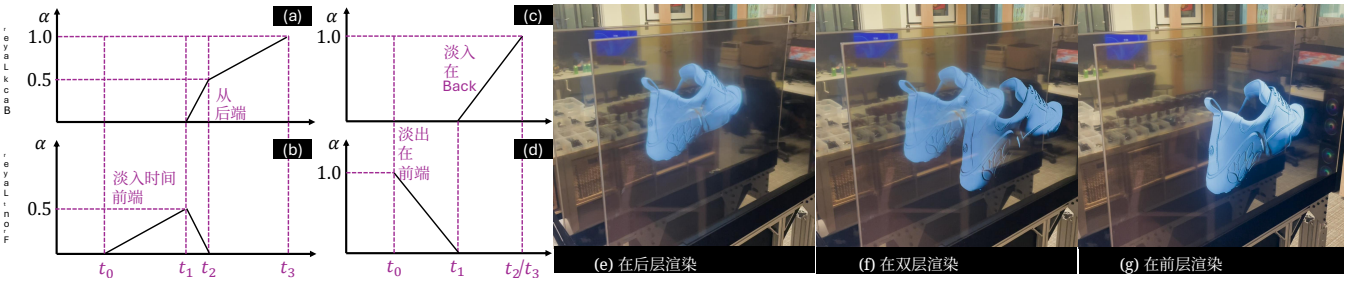

图8：前后显示屏的示例传递函数，通过调整透明度（α）实现过渡和渐隐效果；（a - b）信息渲染在前显示屏渐显、随后在后显示屏渐显时的传递函数；（c - d）信息渲染从前显示屏过渡到后显示屏时的传递函数；（e - g）渲染的鞋子位于后层、前后两层以及前层时的示例。

时间变化，并以三个可配置阶段为框架。研究人员可以为每个阶段定义过渡样式，从而对过渡行为进行原型构建，包括渐隐。图8a - b和图8c - d分别展示了将渲染的鞋子逐渐淡入前层，以及将鞋子从前层过渡到后层的示例。虽然 α 被用作D1中指定的渲染参数，但我们的工具也支持研究人员尝试其他类型的参数。图8e - g分别展示了渲染的鞋子位于后层、同时位于前层和后层，以及位于前层的示例。

配置并集成额外的传感器和设备。我们最初设计的Proscenium工作站采用两台平行堆叠的透明显示器，但不具备任何传感能力。尽管可以采用绿野仙踪法，由实验人员手动触发预先配置的多媒体素材，但对于某些体验的原型制作，可能仍需具备最低限度的传感能力，以便根据不断变化的情境（例如手势）动态调整观看体验。我们设计的 Proscenium工作站包含另外两套传感系统：

• **手部追踪**。Proscenium工作站原型设计中使用的透明显示器仅用于渲染信息，不具备触摸屏功能。为了追踪手部位置和手势，我们在前显示屏底部放置了一台 Leap Motion 摄像头 [35]，持续采集每只手在三维空间中的 27 个关键关节。在 Proscenium 初始化期间，我们首先使用 Kabsch-Umeyama 算法 [18, 19, 36]，计算 Leap Motion 摄像头系统与 Proscenium 原型设计工具之间的最佳平移矩阵和旋转矩阵。校准过程中，我们使用了前显示屏上的 12 个预设位置。这些变换矩阵随后会在后续的设计和研究过程中持续保存。在体验原型设计期间，我们的原型设计工具持续计算摄像头系统中追踪到的手部位置，以渲染预先配置的多媒体素材。

• **其他平板电脑和移动传感器的集成**。 除了通过手部追踪为 Proscenium 扩展触摸屏功能之外，我们还将另一台平板电脑集成到工作站中。我们的工具能够利用平板电脑广播的传感器数据流，对渲染信息行为进行原型设计。我们的实现使用一台 Microsoft Surface Pro 平板电脑 [24] 和 Escapement 工具 [25]来传输传感器数据。

• **远程用户深度增强视频流的传输与理解**。Proscenium 工作站集成了一台 RGB-D

摄像头，以便探索如何利用双透明显示器渲染和呈现远程视频会议参与者，并进行体验原型设计。虽然 Proscenium 的大多数体验原型都使用预先录制的图像和视频资产（第 A.3 节），但也有少数采用更高保真度开发，以提供更逼真的交互体验。尽管如今的视频会议应用需要无线传输摄像机视频，我们仍选择通过串行线缆将 Orbec Femto Bolt相机 [26] 连接到 Proscenium 工作站。探索和体验评估期间，相机将放置在另一个房间。我们的初始原型使用无线IP摄像机的视频流，并采用 Depth Anything [38]等近期预训练的 AI 模型来推断逐像素深度；但最终我们决定使用 RGB-D 摄像头，因为这大大简化了原型设计工具的开发，让我们能够专注于体验原型设计，而不是设计用于实时高质量视频流的网络架构。

A.3 资产创建与优化

我们并未开发功能完备的端到端应用，而是通过 Proscenium工作站进行体验原型设计，主要借助 WOz 流程重现设想中的体验。为支持这一流程，我们预先录制并编辑了一系列图像和视频资产，随后使用 Proscenium 工具，按照预定义的行为和体验档案进行渲染。这种原型设计方法无需投入完整应用开发所需的额外工作，便能快速探索多种技术与体验，同时也支持迭代优化设计。我们采用了绿幕等常见的电影拍摄技术 [11]，录制用户说话、展示关键物体以及做出不同手势行为的画面。这样，在透明显示器上渲染用户时，我们便能高效移除背景，营造出远程用户与 Proscenium 用户身处同一空间的视觉效果。我们还使用了 MediaPipe [21] 和 Segment Anything模型 [20, 32] 等现有 AI 工具提取手部和关键物体，并根据预定义的体验档案将它们渲染在独立的显示层上。