

课 题 目 录

一、影像算法

- 1.1 未来 AI 摄影发展方向与产品应用研究
- 1.2 空间影像双摄画质一致性研究
- 1.3 超大分辨率 AIGC 人像增强技术研究
- 1.4 轻量化文生图扩散模型技术研究
- 1.5 可组件化的 3D 场景生成和编辑
- 1.6 高效视频生成架构
- 1.7 面向摄影机器人的主动式相机控制系统
- 1.8 基于 AIGC 的低时延 SDR2HDR 算法研究
- 1.9 光线路径追踪的高效算法原型研究
- 1.10 高性能特征跟踪和 3D 空间推理

二、材料与工艺

- 2.1 移动终端产品散热技术创新
- 2.2 电池在线 EIS 检测算法及失效状态评估
- 2.3 6.75MHz 高功率可集成逆变整流控制系统
- 2.4 高性能广色域 BT2020 发光材料器件研究
- 2.5 7 系铝合金阳极掉膜优化研究
- 2.6 压铸铝合金做阳极外观
- 2.7 高强超薄盖板玻璃材料设计与制造工艺协同攻关
- 2.8 用于有机特种纤维的无热效应切割技术
- 2.9 持久耐脏和耐黄变的透明弹性体
- 2.10 折叠屏叠层结构均质化本构建模研究
- 2.11 温度驱动的湿气浓度扩散方程求解技术及其在 ABAQUS 实现
- 2.12 固态电池应用特性研究

三、光学、颜色及器件

- 3.1 基于人眼感知的覆盖真实场景的 HDR 影像与显示评价解决方案
- 3.2 润滑条件下的摩擦系统模型建立与仿真
- 3.3 高精度宽频谱防抖控制技术
- 3.4 EVS 信号分析及应用专题研究
- 3.5 基于图像传感器的任意频率 flicker 实时检测
- 3.6 基于 XR 的色彩模型及显示效果调试

四、网络通信

- 4.1 低功耗高性能 Wi-Fi 或 无线传输模式方式
- 4.2 城市高楼环境下的精准定位
- 4.3 基于可穿戴设备的水下通信技术研发
- 4.4 识别用户使用手机时不同的使用场景

五、运动与健康

- 5.1 睡眠干预方法研究
- 5.2 无创血糖检测探索性研究
- 5.3 高接触阻抗下的生物电阻抗测量技术
- 5.4 视觉敏感人群与显示需求研究

六、芯片技术

- 6.1 基于 AI Agent 的 AI 加速器架构探索平台
- 6.2 基于存算技术的端侧高效 AI 加速器的技术方案
- 6.3 面向高效低延迟 AI 加速器架构的敏捷开发技术方案

七、软件优化技术

- 7.1 移动终端设备调度算法优化研究
- 7.2 并行分块渲染算法能效比优化预研

八、声学技术

- 8.1 基于 AI 眼镜形态的声学泡泡算法模型开发
- 8.2 移动终端设备中扬声器后腔声波驱动电池盖结构振动的抑制策略与优化设计研究

九、机器人控制技术

- 9.1 基于 MR 遥操作的小样本动作学习及任务泛化能力研究
- 9.2 跨平台融合感知框架

一、影像算法

1.1 未来 AI 摄影发展方向与产品应用研究

技术背景

随 AIGC 技术快速发展，“AI 摄影”概念应运而生，由 AI 生产的图片内容所呈现的效果惊艳感甚至超越了日常传统摄影的观感和美感，但同时也带来关于内容真实性的争议，如 AI “虚空补图”（如 60 倍长焦细节伪造）、图片过度美化导致图像失真，此外生成内容的不可控性（内容扭曲畸变、违反物理规律等）也为该技术更大范围的产业应用带来阻碍。

本课题希望探索未来 AI 摄影发展趋势与潜在的未来产品应用形态，在技术维度探索 AI 解决真实性与创造性的平衡解决方案（解释：AI 生成图片的真实性缺陷源于算法对物理世界理解的不足（如光影逻辑、细节还原、建筑物还原等）；而创造性则因模型本质是“数据重组工具”，无法替代人类的真实情感表达和文化洞察），使技术真正服务于用户创作表达；产品应用维度探索 AI 作为摄影工具，服务于普通用户摄影创作应用场景有哪些，从而借助技术演进趋势，探索未来 AI 摄影工具的可能产品形态。

本课题的最终交付件是关于上述问题的一套技术解决方案与产品设计方案，其中技术解决方案核心解决“AI 真实性与创造性的平衡问题”，产品设计方案核心对应“未来 AI 摄影工具的可能产品形态”。

AI 技术促进摄影产业的变革。从拍摄到后期，生成式 AI 可覆盖全链路：

- 拍摄阶段：AI 实时识别场景并优化参数、拍摄画质等；
- 后期处理：支持 AI 消除杂物、扩展背景、风格迁移（如 vivo、小米 14 Ultra 的生成式编辑功能）；
- 个性化创作：通过用户行为数据本地训练，形成“养成式相机”（如 Capture One 的智能仿色功能）

当前现状

技术应用层面

领域	典型应用	代表案例
大众摄影	AI 一键美化、智能构图	vivo AI 视效（四季人像、四时人像等）、美图秀秀等软件
专业创作	AI 批量修图、风格仿色	Capture One 能仿色、Adobe Sensei26
行业场景	电商虚拟场景生成	AI 生成商品广告图

技术挑战

AI 生成对摄影技术能力赋能与局限性

1. 技术优势：效率与多样性

拍摄大片的门槛低：AI 简化了摄影技术门槛，使非专业用户也能参与摄影创作，获得极致的影像效果，推动摄影平权化时代；

AI 突破传统摄影创作天花板：AI 生成技术突破了传统硬件限制，结合摄影艺术的多样性，在传统影像场景下可以解决原有技术解决不了的需求和痛点，如旅行途中天气阴沉拍不出阳光普照的好天气，AI 技术可以一键获得。或 AI 可作为摄影创作的“灵感伙伴”，提供新视角和可能性，拍摄超越眼之所见的还原心中想象的画面，如望庐山瀑布曾经诗词中“飞流直下三千尺”的画面能够直接复现眼前；

2. 核心局限：惊艳感与真实世界关联性的均衡

AI 与真实世界的关联：

AI 拍摄或者后期处理的图片从效果惊艳感的观感来说，是超出日常拍摄设备和技能能达成的效果，但同时也会带来真实性争议，如 AI“虚空补图”（如 60 倍长焦细节伪造）、图片过度美化导致图像失真。且存在不可控性，如算法在元素组合时可能产生偏差（如“六根手指”问题），需人工干预修正；无统一规范界定 AI 生成内容真实性（如纪实摄影需区块链存证防篡改）

人类情感表达：AI 缺乏人类的情感和主观意识，其作品本质是算法对数据的重组，无法传递人类思维独特的生命体验或文化内涵。

技术诉求

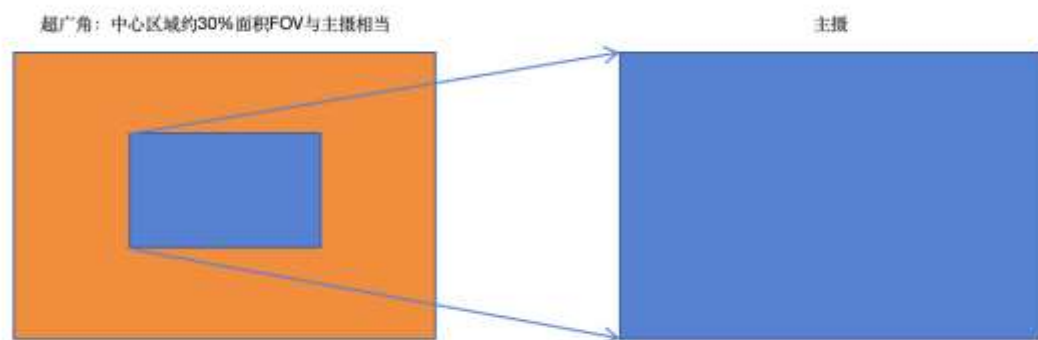
核心目标：本课题聚焦技术攻坚（算法/硬件）、产品创新（个性化/场景化）探索未来 AI 摄影发展趋势与产品应用形态，技术上探索 AI 解决真实性与创造性的平衡解决方案，使技术真正服务于用户创作表达。产品应用上探索 AI 作为摄影工具，服务于普通用户摄影创作应用场景有哪些，未来的 AI 摄影工具的产品形态应该是什么样子。

希望承接课题人的研究方向：视觉 AI、多模态生成以及面向产业应用的 GenAI 方向。

1.2 空间影像双摄画质一致性研究

技术背景

手机拍摄空间影像，分别使用主摄（23mm/35mm 焦距）和超广角（14mm 焦距），进行拍摄，经过 rectify 算法后，超广角中心约 30%左右的图像被取出并 upscale 到主摄分辨率，这导致主副摄在清晰度上有明显差距；另外，双摄之间焦距和光圈不同，导致景差距明显，在拍摄近距离物体时，双摄背景虚化程度差别明显；同时两个相机可能使用不同型号或厂家的 sensor，再叠加环境的光照因素，因此在成像画质方面存在较大的差距，影响空间影像的观看体验。因此希望能够进行双摄画质融合或迁移，使双摄在清晰度、亮度、影调、景深等基础画质层面高度一致。



副摄裁切中心区域与主摄 rectify，清晰度下降



主副摄景深差异，导致背景虚化程度差别明显



不同 sensor 成像基础画质差别明显

技术挑战

1. 双摄融合或迁移需要先进行双摄对齐，依赖深度估计的精度，以及双摄之间的同步，很难做到完全对齐
2. 对齐后存在遮挡关系造成的“空洞”，“空洞”区域缺少一个视角的数据
3. 处理时间长，尤其是清晰度、景深、影调、亮度等串行处理时
4. 处理视频是需要接近实时处理，并且需要考虑时间一致性

当前结果

1. 主副摄间清晰度的迁移已经完成，但存在 **bad case**，并且性能无法满足处理视频的需求；
2. 景深一致性上可以做到双摄同模糊，但模糊的 **pattern** 不太一致。希望能做到双摄同清晰。



副摄处理前，背景没有虚化



副摄处理后背景虚化，虚化方式接近主摄



主摄背景存在光学虚化

技术诉求

任务: 开发双摄画质迁移或融合经算法, 实现空间影像双摄画质之间的一致性。处理后, 使用手机主摄和广角相机拍摄的空间照片、视频左右目画质及画质一致性上可以对标同场景下相同双摄拍摄的空间照片、视频。

1. 算法处理后清晰度接近主摄, 并且画质与主摄接近
2. 算法处理后, 左右目具有景深一致性, 并且背景或前景虚化程度与广角镜头接近
3. 算法处理后, 左右目图像在亮度、影调上基本一致, 并与主摄接近。
4. 算法可以实时处理 4K60fps 视频, 且最终效果具有帧间的一致性, 无可见 flicker。

1.3 超大分辨率 AIGC 人像增强技术研究

技术背景

近年来 AIGC 大模型技术在图像生成领域取得了十分优异的效果, 因其高清的内容生成能力与强大的分布拟合能力, 逐渐成为主流的生成模型。但目前学界大多做的都是 512 或 1024 偏小分辨率输入的研究, 在实际用户场景中, 有时会遇到输入图本身分辨率很大且人脸占比很大的场景, 比如人脸大小达到 3kx4k, 如果通过插值下采到 1024 再进行增强生成, 在上采阶段会丢失清晰度和纹理细节, 如果通过网络后面阶段设计简单一个 SR 模块, 那么核心效果都在小图生成, 仅仅通过 SR 也很难达到效果上限, 本项目针对手机拍摄的超大分辨率人像场景如下图所示, 人脸大小范围在 3k-8K 的范围内, 旨在探索利用超大分辨率的 AIGC 人像增强技术研究达到单反级的人像效果。



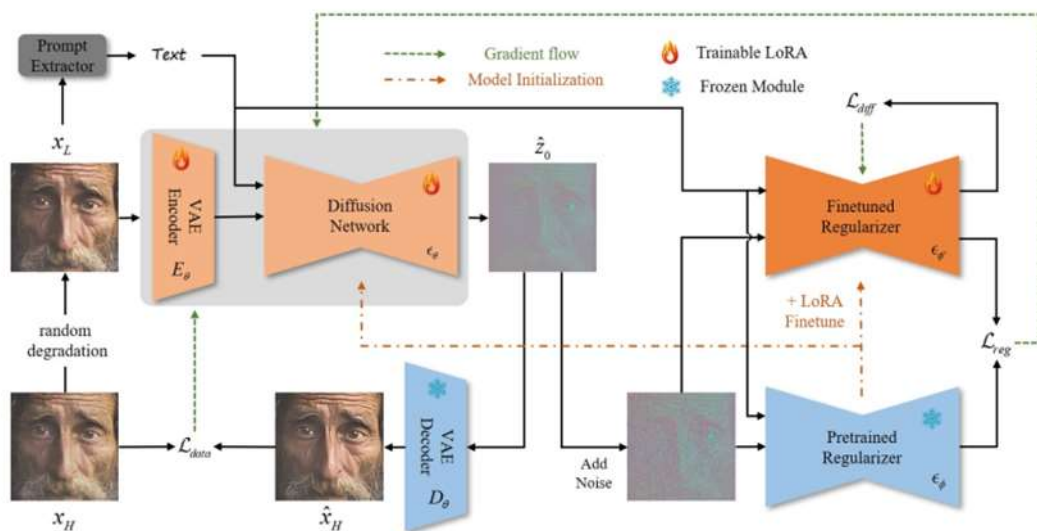
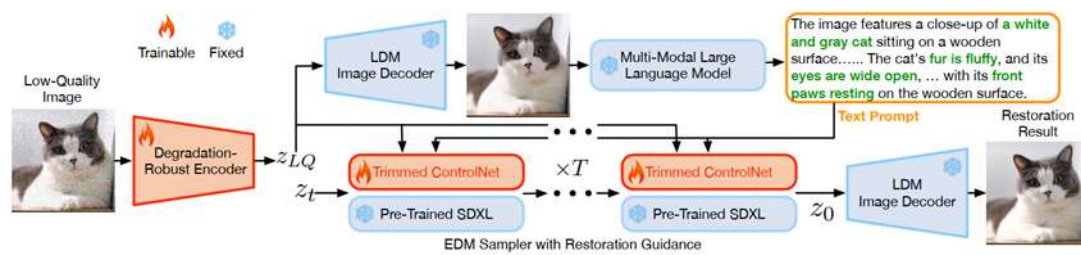
技术挑战

1. 人脸超大分辨率, 需要设计处理大分辨率的图像增强技术模型
2. 目前 AIGC 基模最大只有 1024 或 2048, 如何基于 AIGC 做好用户实际场景的 AIGC 超大分辨率效果
3. 超大分辨率在训练和推理时时间性能都是很有影响, 设计的方案做好效果与时间性

能的均衡

当前结果

目前业界 AIGC 增强的方案较多，比如 DiffBIR 【1】，SUPIR 【2】，OSDiff 【3】 等，但都是在较小分辨率如 512 上进行处理。



技术诉求

- 1.设计一种满足业务需求的超大分辨率 AIGC 人像增强技术，在实际用户场景测试集中可以达到 85%的成片率
- 2.设计的超大分辨率 AIGC 人像增强技术没有严重的 artifacts 问题。

参考文献：

- [1] Lin X, He J, Chen Z, et al. Diffbir: Towards blind image restoration with generative diffusion prior (2023) [J]. arXiv preprint arXiv:2308.15070.
- [2] Yu F, Gu J, Li Z, et al. Scaling up to excellence: Practicing model scaling for photo-realistic image restoration in the wild[C]//Proceedings of the IEEE/CVF

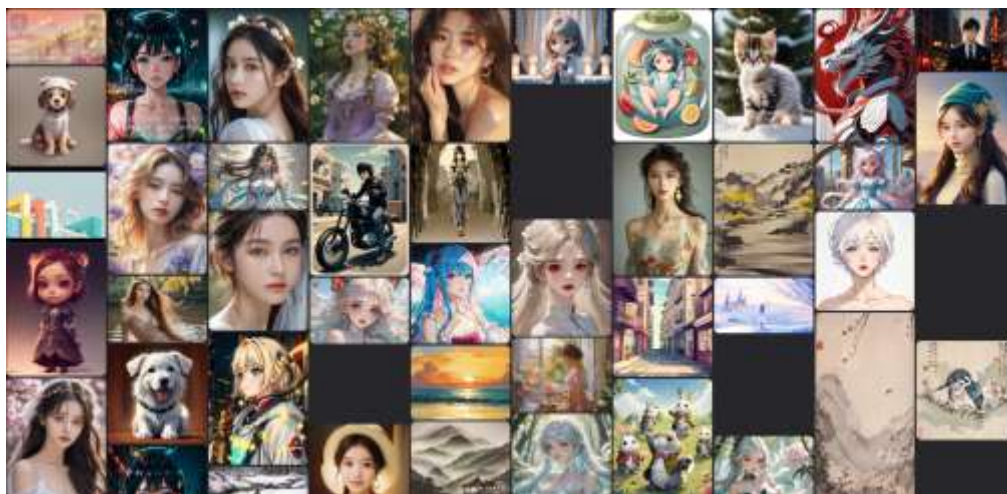
Conference on Computer Vision and Pattern Recognition. 2024: 25669–25680.

[3] Wu R, Sun L, Ma Z, et al. One-step effective diffusion network for real-world image super-resolution[J]. Advances in Neural Information Processing Systems, 2024, 37: 92529–92553.

1.4 轻量化文生图扩散模型技术研究

技术背景

大规模文本到图像扩散模型（文生图模型）在内容生成方面取得了显著的成功，为图像编辑和视频生成等许多应用提供支持。将文生图模型部署在云上会带来与数据安全性和服务器成本相关的问题。然而，在移动平台上部署的文生图模型往往伴随着模型体积巨大和运行时间缓慢的问题。首先，文本图像扩散模型中网络架构的复杂性涉及大量参数，通常达到数十亿个参数，导致计算成本高昂。其次，扩散模型的固有设计需要通过迭代去噪生成图像，需要对模型进行多次去噪。当前移动平台的文生图模型存在生成图片耗时长，功耗高，内存占用大等等诸多问题，限制了文生图模型在移动平台的应用。



技术挑战：手机计算资源受限，如何构建轻量高效的文生图模型，同时保证效果。

计算资源有限：移动设备的 CPU 和 GPU 性能相较于桌面或服务器设备较弱，难以高效运行复杂的扩散模型，导致处理速度缓慢。

高功耗和电池消耗：运行复杂的深度学习模型会显著增加设备的能耗，缩短电池续航时间，影响用户体验。

运行速度慢：移动应用通常需要即时响应用户操作，而扩散模型的多步生成过程往往导致较高的延迟，难以满足实时性需求。

内存和存储限制：移动设备的 RAM 和持久存储空间有限，难以容纳大型模型及其运行所需的数据，可能导致内存不足或存储压力。

散热和发热问题：长时间高强度的计算任务会导致设备过热，可能引发散热问题，影响设备性能和用户体验，甚至损害硬件。

图像生成尺寸小：由于内存，功耗等硬件限制，导致扩散模型通常只能进行 1024*1024 以下 size 的图像推理生成，而当前手机拍照的像素通常能达到 3000*4000 以上，移动设备上部署模型几乎无法处理如此高像素的图片，影响了文生图模型在移动设备的应用。

小模型出图质量差：业界开源的小规模文生图模型出图质量都比较差，很难达到 SDXL 等大规模模型的出图指标。

综上所述：需要构建一个适配移动平台，能够在移动平台功耗，内存，计算资源限制的场景下，保证出图质量的高性能高效率的文生图模型框架。

当前结果

当前业界针对移动平台的 Diffusion 模型有部分工作，如 mobile diffusion, BK-SDM, SnapFusion 等工作，但是效果距离云端的大型模型仍有很大差距。仍需要在保证高质量出图效果的基础上降低移动平台部署的难度。

技术诉求

一、任务：实现一个为移动平台设计的高性能 Diffusion 模型架构。（下列指标以 iphone15 pro 设备 GPU 平台为例）

- 1.diffusion 模型参数量：不大于 400M(diffusion+textencoder + vae 小于 600M)。
- 2.diffusion 模型 FLOPs (G): $\leq 182G$ 。（不超过 SDXL 的 1/4，参考 mobile diffusion）
- 3.diffusion 模型耗时：one step 生成 1024*1024 尺寸图像不超过 300ms。
- 4.VAE：生成 1024*1024 尺寸，耗时 $\leq 200ms$ ，GFLOPS $\leq 500G$ ，param $\leq 40M$
- 5.出图质量与图文一致性：优于 SDXL（FID，CLIP 等客观指标；人工主观评测）。

二、涉及到的研究技术方向

- 1、扩散模型的剪枝、量化和蒸馏、步数蒸馏等研究
- 2、扩散模型的创新结构设计
- 3、线性注意力结构在扩散模型中的研究

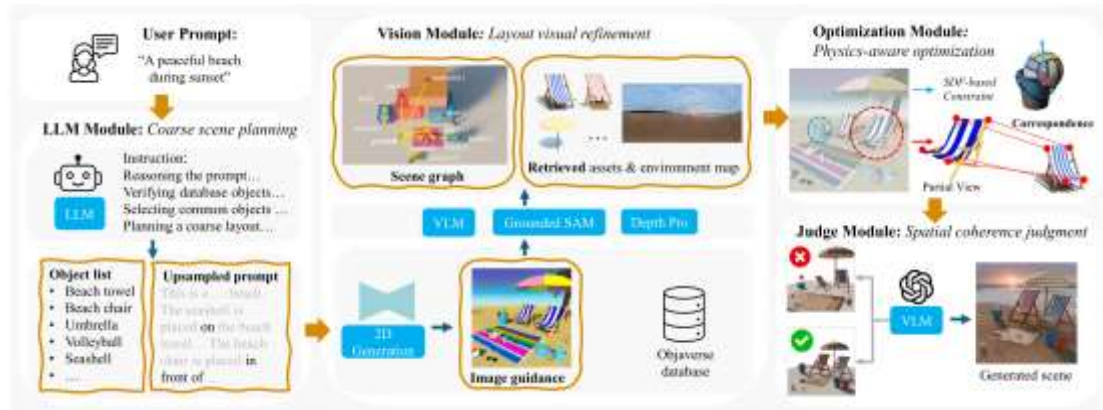
参考文献:

- [1] HU D, CHEN J, HUANG X, et al. SnapGen: Taming High-Resolution Text-to-Image Models for Mobile Devices with Efficient Architectures and Training[J]. 2024.
- [2] LI Y, WANG H, JIN Q, et al. SnapFusion: Text-to-Image Diffusion Model on Mobile Devices within Two Seconds[J].
- [3] KIM B K, SONG H K, CASTELLS T, et al. BK-SDM: A Lightweight, Fast, and Cheap Version of Stable Diffusion[J]. 2023.
- [4] ZHAO Y, XU Y, XIAO Z, et al. MobileDiffusion: Subsecond Text-to-Image Generation on Mobile Devices[J]. 2023.

1.5 可组件化的 3D 场景生成和编辑

技术背景

真实物理世界是以三维空间形态存在的，所有物体都具备几何结构、材质属性和动态交互特性。当前移动影像技术受限于二维平面媒介，难以完整记录三维场景和感知空间关系。基于此背景，我们期望研发基于空间感知的 3D 场景生成技术，并将场景转化为多种物体的组合，旨在突破传统摄影的维度限制，构建支持空间具身智能和虚拟现实的下一代影像基础设施。该技术以稀疏视角甚至单张 RGB 图像作为输入源，通过 3D 生成式大模型（如 3D 扩散生成或自回归生成模型）实现三维场景的解构与重建，结合 3D Gaussian Splatting 等前沿渲染技术，实现场景组件化，最终生成包含几何、纹理材质、光照信息的可交互三维场景。该技术的突破将推动三维影像成为继图像和视频后的新型手机影像媒介。既可以为用户在手机端实现场景生成及空间感知，并通过多模态（文本、图像等）实现虚拟物品生成及融合拍摄等创新交互，又可以为具身智能和 3D 运镜构图等任务提供低成本且真实的三维世界模拟，解决训练中数据样本不足的痛点。



可组件化的场景生成：从单目视频/单图生成由多个物体构成的 3D 场景，实现交互编辑和世界仿[4]

技术挑战

对于可组件化 3D 场景生成技术，当前学界方法大致可分为两类路线：1）基于神经渲染的场景重建和理解 [1]，3D Gaussian Splatting（3DGS，高斯溅射）是近一年新兴的场景重建和新视角渲染渲染技术，结合 VLM 等视觉语言大模型，对场景进行三维实例化建模以实现高效的三维语义理解，但由于 3DGS 的使用会引入额外的优化时间，且仅针对单一场景，泛化能力较弱。2）基于 AIGC 的方法生成 3D 场景 [2,3]，受限于小规模 3D 室内场景数据集，导致场景多样性和布局复杂性不足，且为文本驱动的，生成场景真实感较弱。



场景组件化：将 3D 场景转化为一个个物体组件，并建立物体间空间关系[5]

主要的技术挑战：

1. 场景组件化及空间关系缺失：现有场景生成方法缺乏对三维空间几何和物理规则的直观感知。在得到一个 3D 场景后，不能有效的将场景组件化，继而无法显式建模场景内的物体关系，比如物件间的支撑关系（“如花瓶放置在桌面上”）、功能关系（如“椅子围绕餐桌排列”）等。只有充分理解 3D 场景中各个物体关系，继而在虚拟物品融合拍摄等创新交互和

空间位置编辑等应用中更好的实现空间对齐。

2. 单图场景生成及交互编辑：现有的单图场景生成方法缺乏高质量高真实感的室内/室外 3D 场景，从而导致场景生成的多样性和保真性存在挑战。本项目旨在借助多模态大模型的空间理解能力和 3D 物体生成大模型的多样性和保真度，以单图作为输入，对图像中的物体进行解构，并借助 3D 物体生成大模型生成各个物体，优化物体间的空间布局，从而实现组件化的 3D 交互场景生成。

当前结果

vivo 对于 3D 物体生成技术[6]和 3D 场景重建技术[7]进行了独立的工作复现和研究，但相关的技术积累还不够全面和深入，当前学术前沿也缺乏增量式快速场景重建和理解方法的探索，增量式重建的好处可有效降低场景重建的算力需求，并可通过开放词汇的语义理解对场景实现高效的组件化。同时，基于 LLM 大模型实现文本驱动的 3D 交互场景生成技术[4]，虽引入了 LLM 使得该方法可生成多样化的 3D 交互场景，但整体场景保真性不够，与严肃摄影的高保真要求存在不小差距。

技术诉求

任务：可组件化的 3D 场景生成和编辑

增量式开放词汇下的场景重建、理解与交互：以视频为输入，结合 VLM 模型，优化重建速度，以数据增量的形式，实现场景的快速生成和语义理解。支持对 3D 场景的模块化编辑和组合，对组件化后物体所缺失内容进行补全，并实现光照编辑、纹理编辑、物理仿真等诸多世界模拟效果。

单图 3D 场景生成及编辑：以单图为输入，构建组件化的场景生成管线，借助多模态大模型（VLM，SAM，3D-AIGC 等），对图片内物体关系建图（scene graph）、生成对应物体，优化物体间的空间布局，能够生成符合物理规律且具备空间真实感的 3D 交互场景。并可通过文本等多模态信息驱动物体空间位置编辑、纹理编辑等交互编辑形式。

参考文献：

- [1] Zhai H, Li H, Li Z, et al. PanoGS: Gaussian-based Panoptic Segmentation for 3D Open Vocabulary Scene Understanding[J]. arXiv preprint arXiv:2503.18107, 2025.
- [2] Lin C, Mu Y. Instructscene: Instruction-driven 3d indoor scene synthesis with semantic graph prior[J]. arXiv preprint arXiv:2402.04717, 2024.

- [3] Yang Y, Jia B, Zhi P, et al. Physcene: Physically interactable 3d scene synthesis for embodied ai[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 16262-16272.
- [4] Ling L, Lin C H, Lin T Y, et al. Scenethesis: A Language and Vision Agentic Framework for 3D Scene Generation[J]. arXiv preprint arXiv:2505.02836, 2025.
- [5] Dong W, Yang B, Yang Z, et al. HiScene: Creating Hierarchical 3D Scenes with Isometric View Generation[J]. arXiv preprint arXiv:2504.13072, 2025.
- [6] Hunyuan3D 2.0: Scaling Diffusion Models for High Resolution Textured 3D Assets Generation
- [7] Kerbl B, Kopanas G, Leimkühler T, et al. 3d gaussian splatting for real-time radiance field rendering[J]. ACM Trans. Graph., 2023, 42(4): 139:1-139:14.

1.6 高效视频生成架构

技术背景

视频生成技术旨在从输入条件（文本、图像、动作序列等）合成高质量、连续的视频内容，在图像视频编辑、影视特效、游戏开发、虚拟现实、自动驾驶仿真等领域具有重要价值。但是当前视频生成模型推理速度较慢，HunyuanVideo 在 A100 GPU 上生成一个 160 帧 720p 的视频需要约 50 分钟时间，这主要是因为扩散模型（如 Video Diffusion Models）通过迭代去噪提升质量，但生成高分辨率视频时计算复杂度呈指数级增长（如高分辨视频需处理百万级像素 $\times$ 时序维度）。此外，当前模型非常依赖于高端 GPU（如 NVIDIA A100），难以在边缘设备或移动端实现实时推理。

技术挑战

1. 时间一致性建模：长视频生成中如何保持物体运动、光照变化的连贯性。
2. 计算复杂性：高分辨率视频（如 1080p）生成需处理百万级像素 $\times$ 时序维度，算力需求巨大。
3. 实时性瓶颈：扩散模型迭代去噪步骤多（通常 50~100 步），难以满足交互式应用需求。
4. 端侧部署：移动端/边缘设备资源受限，需模型轻量化与推理加速技术。

当前结果

当前行业上主要关注如何提升视频生成模型效果，对于模型大小和推理速度还没有太多深入的研究，一些初期的轻量化视频生成模型，例如 LTX-Video, SnapGen-V 等难以兼顾视

频生成质量和生成速度，基本上以大幅牺牲生成效果来达到快速推理，如何兼顾视频生成质量和推理效率是后续的一大研究方向。

技术诉求

核心目标：构建“高质-高效-轻量”三位一体的视频生成架构，突破实时生成与部署瓶颈。具体可以分为以下几点：

- 1) 高效推理速度：提升视频生成模型推理速度，减少推理迭代次数，在 1 步或者 4 步内达到 50 步推理效果。
- 2) 端侧友好算子：当前视频生成模型在端侧几乎没有落地可能性，算子较为复杂，内存操作连续性较差，需要进一步设计端侧友好的轻量化算子。
- 3) 轻量化模型设计：模型大小进一步减小，通过蒸馏等策略将大模型能力蒸馏到轻量化模型中。

参考文献

SnapGen-V: Generating a Five-Second Video within Five Seconds on a Mobile Device

https://huggingface.co/docs/diffusers/main/en/api/pipelines/ltx_video

Latent Consistency Models: Synthesizing High-Resolution Images with Few-Step Inference

1.7 面向摄影机器人的主动式相机控制系统

技术背景

传统计算摄影技术主要针对用户已拍摄的图像，通过算法优化画质、色彩和光影等核心摄影属性来提升视觉效果。然而在实际拍摄场景中，由于构图不当、逆光条件、镜头眩光或表面反光等问题，用户往往难以直接获得理想的拍摄效果。本项目创新性地提出了面向摄影机器人的主动式相机控制系统，旨在通过智能化的拍摄决策，主动获取场景中的最佳图像和视频画面。区别于被动式的后期优化，我们的系统从两个关键维度实现突破：

- 1) 最佳图像画面：专业级影像创作的核心在于对画面要素的系统性把控。从构图张力到光影叙事，每个视觉元素都直接影响作品的最终表现力。理想的画面呈现需要同时满足：
构图美学：通过精准的视角选择建立视觉秩序，将二维画面转化为具有纵深感的叙事空间；
光影控制：运用光线方向、强度和色温的变化，实现从纪实记录到艺术表达的跃升；技术规

避：智能识别并避免眩光、反光等常见成像缺陷。然而现实创作中存在显著的能力断层，专业摄影师具备三维空间的敏锐感知、深厚的视觉美学积淀，以及与拍摄对象建立精准的视觉对话能力。这些要素共同构成了影像创作的底层逻辑，也是区分业余与专业的关键所在。其核心挑战在于如何将专业摄影师的视觉智慧转化为可量化的控制参数，在动态拍摄中实时优化构图与光影，同时规避常见成像问题。这一突破将重新定义影像创作的准入门槛，实现从被动补救到主动创作的技术跃迁。

2) 最佳视频画面：视频运镜本质上是静态构图在时间维度上的延伸与演化。在传统影视创作中，摄影师与导演往往需要投入大量时间反复推敲取景方案，这种精益求精的创作方式不可避免地推高了制作成本。而在游戏引擎或混合现实等虚拟场景中，尽管取景流程得到一定简化，但定义符合美学标准的相机位姿与运动轨迹仍高度依赖人工操作：艺术家必须手动设定关键帧的空间坐标、相机朝向，并精确调整关键帧间的时间插值曲线。这种模式虽然能实现高度可控的视觉效果，但其效率瓶颈也日益凸显。因此，如何通过算法实现智能取景与自动化运镜，已成为突破行业生产力边界的关键课题。

技术挑战

1) 3D 空间感知：主动式相机控制的核心在于对场景的三维空间感知能力。当前主流技术方案仍主要依赖二维图像分析辅助相机位姿调整，难以实现精确的空间关系建模。

2) 基于 3D 空间感知的美学图像/视频检索：通过检索专业美学图像/视频可为最佳图像/视频画面提供精准参考，但如何结合三维空间感知能力实现更细粒度的图像/视频检索（如视角、景深、空间布局匹配）仍是行业亟待突破的技术瓶颈。

3) 依赖高质量 3D 场景的训练数据：高质量的室内/室外 3D 场景为主动式相机控制的算法训练提供了数据基础，然而行业中尚未有成熟的面向相机控制场景的合成数据。现有数据集（如 Matterport3D）缺乏运镜轨迹标注，如何收集并充分利用 3D 场景的合成数据值得深入研究。

4) 基于 VLM 的交互式相机控制：专业摄影依赖摄影师与模特/用户之间的细腻互动和清晰沟通，如何借助最新的 VLM 模型来更好的解析用户需求，进行对应的主动式相机控制是个行业未解决的技术问题。

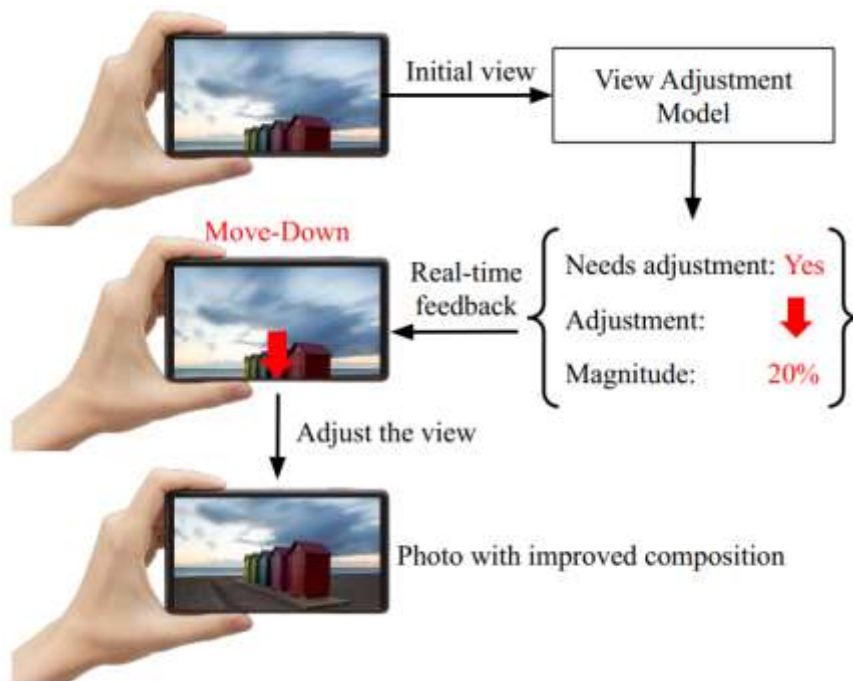
5) 技术路线的分歧：最佳视频画面的核心在于相机位姿的时序建模，当前学界存在两种技术范式：，1) 基于强化学习的动作预测（action prediction），依赖马尔可夫决策过程

（MDP）建模，以美学评估指标（如构图评分、运动平滑度）构建奖励函数，适用于连续动作空间优化；2）基于扩散模型的去噪过程生成相机轨迹。两种范式在轨迹稳定性（强化学习）与创意多样性（扩散模型）上存在根本性差异，需通过跨场景 benchmark 验证最优方案

6）3D 合成数据的 sim2real 挑战：最佳图像/视频画面问题的最终目的是运用在真实世界场景中，如何解决 3D 合成数据的 sim2real 问题，在实际场景中取得良好的泛化能力有待验证。

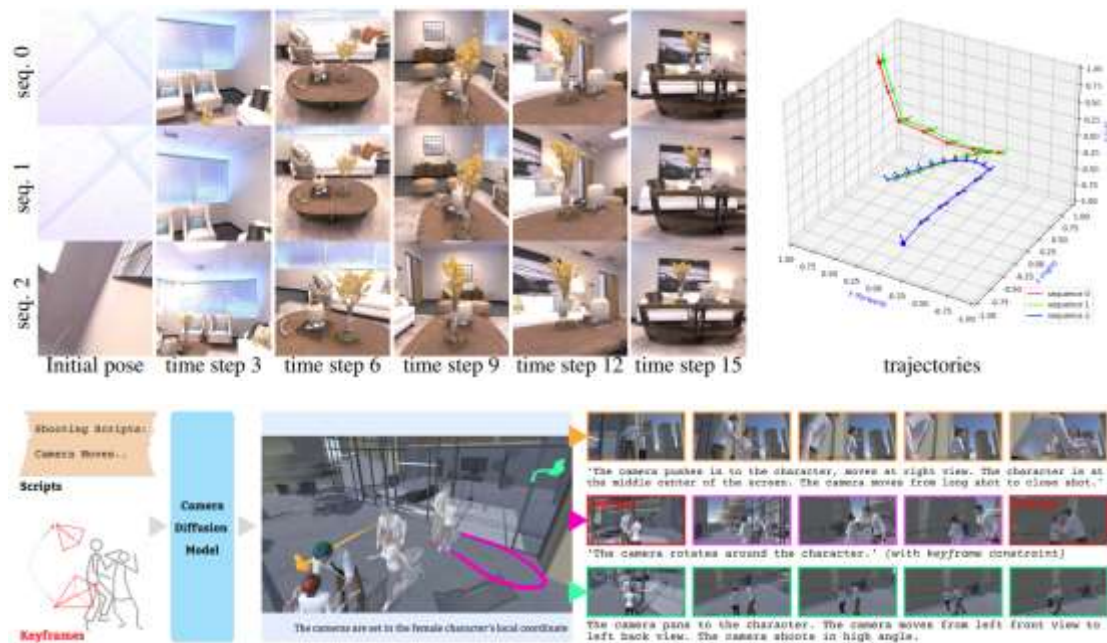
当前结果

最佳图像画面：该项目构建了一个多维度的技术体系，涵盖构图美学、光影控制、眩光抑制和反光消除等核心技术模块。当前行业主流的构图技术方案普遍采用基础的图像裁剪（image cropping）作为构图优化手段，但该技术存在显著局限性：其仅适用于通过简单裁切画面边缘即可优化构图的特定场景（如二次构图或比例调整）。在更复杂的实际拍摄场景中——特别是需要动态调整拍摄视角、机位或透视关系时——这种被动式的后期处理方案往往无法满足专业级构图需求，具体表现为：1）无法修正因拍摄角度导致的透视畸变；2）难以调整画面元素的空间分布关系；3）会损失原始图像的分辨率与画质。



最佳视频画面：相机运镜本质上是静态构图在时间维度上的动态拓展。既有视频运镜的算法虽能有效泛化至训练场景内的各类初始姿态，但在面对全新的、未曾见过的场景时，其泛化性能显著下降。因此，实现跨场景的稳健泛化已成为未来研究的核心目标，这将有希望使

视频运镜发展成为一种更具普适性的视角序列生成框架。然而，达成这一目标需要大幅提升计算资源、获取更大规模的场景数据集，并设计出具备良好可扩展性的深度强化学习算法。此外，现有的神经美学评估器主要关注单帧图像的美学质量，而帧与帧之间的平滑性则依赖于人工设计的启发式规则。因此，开发能够评估时序美学一致性的神经模型是本领域亟待突破的关键方向。据我们所知，此类模型目前尚未公开，故构建此类模型将成为未来研究的重点方向。

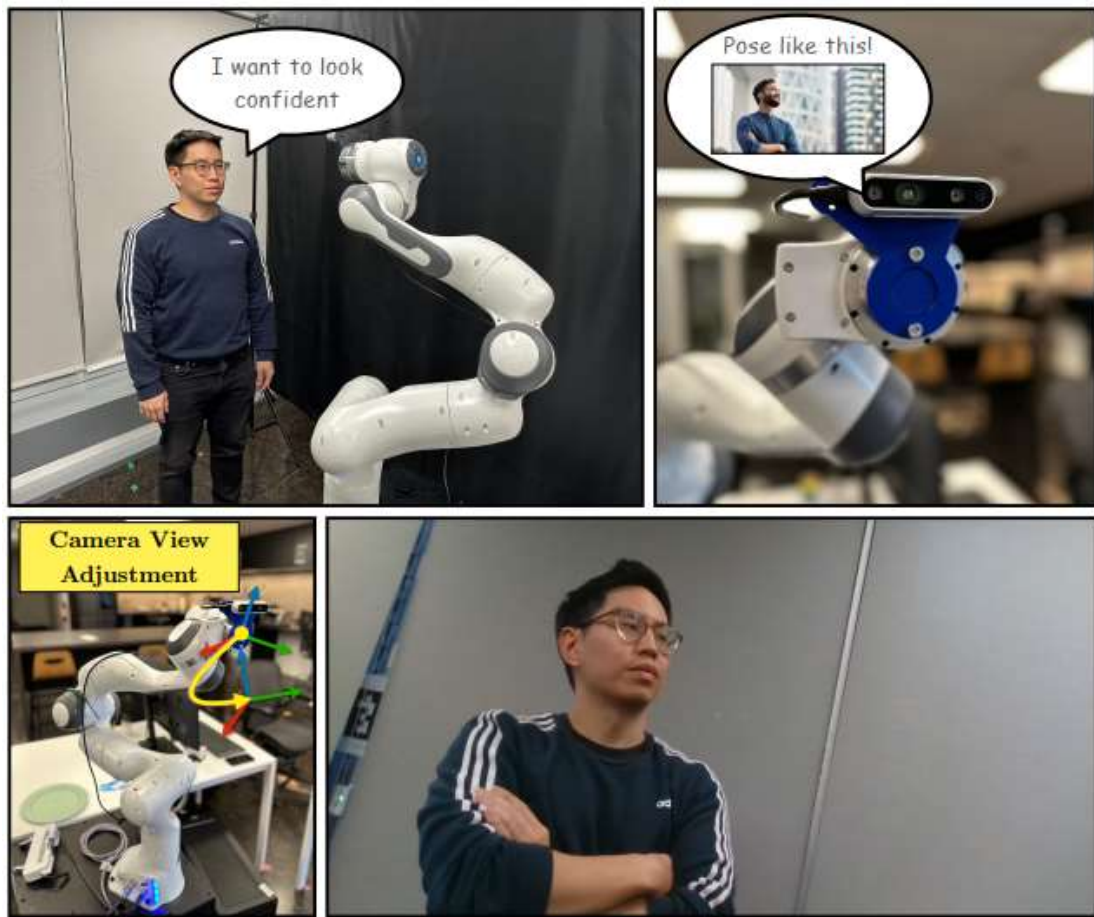


技术诉求

该项目分为二阶段目标

1. 最佳图像画面：

- 3D 空间感知能力：**对于图像的 3D 检测、识别、分割；基于视频流的连续 3D 空间感知能力。
- 基于 3D 控制感知的美学图像检索：**利用图像的精确 3D 感知信息，进行美学图像（Unsplash 等）的检索。
- 基于 VLM 的交互式相机控制：**根据用户输入的文本语言描述，利用检索到的美学图像做引导，进行针对性的图像构图&光影调节。
- 融合可移动平台底座+机械臂+VIVO 手机，**进行实际场景的自动化相机控制。



2. 最佳视频画面

1. 研发具备行业领先水平的视频运镜控制算法： 本项目旨在攻克视频运镜控制的关键技术瓶颈，开发出性能卓越、达到行业领先水平的智能算法，以实现虚拟摄像机运动的精准、流畅和富有艺术性的控制。

2. 构建面向视频运镜任务的高质量 3D 合成数据集： 为有效训练和评估所提出的视频运镜控制算法，本研究将着重构建一个包含多样化场景的高质量 3D 合成数据集，具体涵盖以下类型：室内场景（Indoor Scenes）；室外场景（Outdoor Scenes）；包含虚拟人的动态场景（Dynamic Scenes with Virtual Agents）

3. 探索算法在真实世界中的泛化性能： 在合成数据集上完成算法的开发和初步验证后，本研究将进一步评估所提出算法在真实世界数据上的泛化能力，考察其在复杂、未知的真实场景中的适应性和有效性。

4. 融合可移动平台底座+机械臂+VIVO 手机，进行实际场景的自动化视频运镜。

参考文献

- [1] Camera View Adjustment Prediction for Improving Image Composition
- [2] PhotoBot: Reference-Guided Interactive Photography via Natural Language
- [3] Beyond Image Borders: Learning Feature Extrapolation for Unbounded Image Composition
- [4] AutoPhoto: Aesthetic Photo Capture using Reinforcement Learning
- [5] GAIT: Generating Aesthetic Indoor Tours with Deep Reinforcement Learning
- [6] Cinematographic Camera Diffusion Model
- [7] ANSEL Photobot: A Robot Event Photographer with Semantic Intelligence

1.8 基于 AIGC 的低时延 SDR2HDR 算法研究

技术背景

HDR (High Dynamic Range, 高动态范围) 技术能够提供更广泛的亮度和更丰富的色彩, 可以给视觉效果带来显著的提升。近年来, 随着 HDR 技术的不断发展, 其在游戏创作领域的应用也日益广泛。越来越多的主机游戏开始制作 HDR 画质, 为玩家带来更加逼真的游戏体验。然而, 目前手机游戏因制作生态、成本等限制尚未能实现真正的 HDR 画质, 并且短期内这一情况难以改变。因为 HDR 画质对提升游戏视觉效果的明显优势, 所以我们需要开发一种低时延的 SDR2HDR 转换算法, 以优化游戏场景的显示效果。



图 1 HDR、SDR 游戏效果演示

当前的 SDR2HDR 算法主要分为传统方法和 AI 方法。传统方法通常能够提供稳定的 HDR 效果, 但其视觉效果的提升有限, 难以满足要求。相比之下, 基于 AI 的方法可以提供更优的 HDR 效果, 但其泛化能力较差, 容易导致负面效果。



图 2 现有 SDR2HDR 算法的细节重建能力演示

因此，需要进一步提升现有 AI 方法的能力，稳定地将 SDR 转换为优秀的 HDR 效果。

技术挑战

通过 AIGC 生成匹配游戏内容、能呈现更多细节、实现亮处不过曝,暗处有细节的 HDR 效果，主要有以下两个挑战点：

一、生成的细节要与场景匹配、融洽

HDR 与 SDR 的差异体现在三方面：动态范围、色域、数据精度。首先，动态范围的差异会导致在 SDR 内容中，超出动态范围的内容（高光细节、暗处细节）会因为裁剪和压缩处理导致细节丢失。其次，数据精度的差异使 HDR 内容可以存储更多的信息，减少在量化过程中丢失的细节。对于这些在 SDR 中不存在的细节，利用 AIGC 的能力生成新的细节来弥补，但是生成的细节需要匹配当前的游戏画面，避免导致负面效果。



图 3 游戏地平线西之绝境 HDR、SDR 高光细节差异

二、低时延、高性能

时延直接影响游戏的流畅性和玩家的整体体验。低时延意味着输入指令和游戏响应之间的时间差更小，确保动作即时反映在游戏中，尤其是在快节奏的游戏例如王者荣耀、和平精英。时延过高可能导致输入延迟、画面卡顿，这会影响玩家的表现和游戏体验。因此，需要将 SDR2HDR 的算法的运行时长控制在较短时间内，避免影响用户的使用体验。

当前结果

通过 AIGC 生成匹配游戏内容、能呈现更多细节、实现亮处不过曝，暗处有细节的 HDR 效果，同时 AIGC 生成的细节要能融入到游戏场景中，不突兀；处理时延满足低时延需求。

技术诉求

1. 设计一种满足业务需求的人像恢复增强技术，在实际用户场景测试中具有较好的定量及定性效果。
2. 设计的人像恢复增强算法效果没有严重的 artifacts 问题。

参考文献

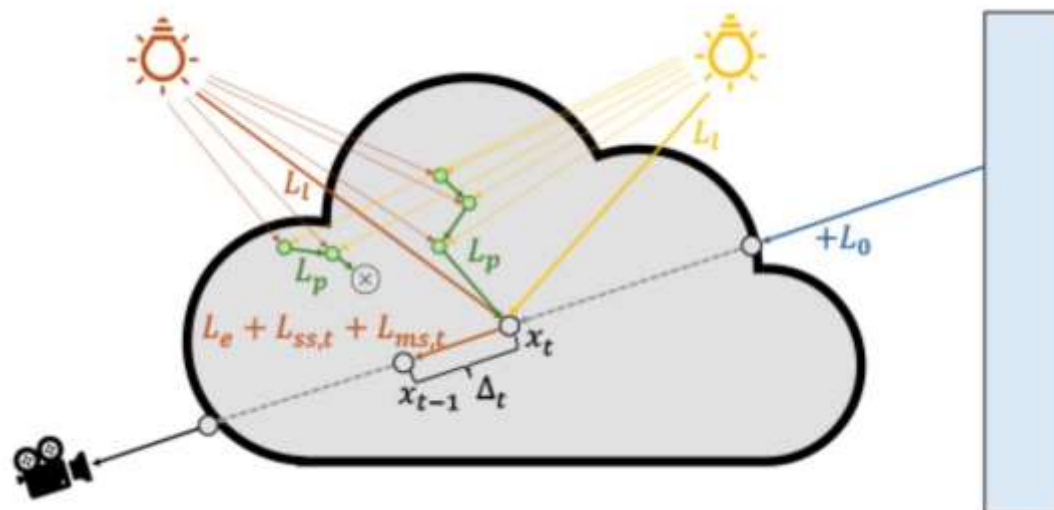
Hanji P. Comparison of single image HDR reconstruction methods — the caveats of quality assessment. In ACM SIGGRAPH 2022 Conference Proceedings (SIGGRAPH '22). Association for Computing Machinery, New York, NY, USA, Article 1, 1–8. <https://doi.org/10.1145/3528233.3530729>
<https://hdr10plus.org/gaming-2/>
<https://developer.baidu.com/article/detail.html?id=294006>

1.9 光线路径追踪的高效算法原型研究

技术背景

几十年来，光栅化一直是首选的渲染技术，因为其速度快且能生成高质量的图像。为追求更接近现实的渲染，光线追踪技术应运而生。光线追踪是一种光传输模拟，其核心思想源于几何光学：从视点（相机）向屏幕上的每个像素发射光线，追踪光线与场景几何体的交点。在该交点，根据物体材质属性（BRDF/BTDF）、光源信息等，计算该点的局部光照（如漫反射、镜面反射）。为了模拟全局光照效果（如软阴影、环境光遮蔽、间接漫反射、焦散等），

算法会在交点处根据概率分布采样生成新的“次级光线”（如阴影光线、反射光线、折射光线、散射光线），并递归或迭代地进行追踪，收集沿路径返回的光线贡献。最终，通过累加路径上各交点贡献的辐射亮度得到像素颜色。



物理基础：其理论基石是渲染方程（Rendering Equation），由 Kajiya 于 1986 年形式化。该积分方程描述了场景中任意一点到达视点的辐射亮度。

- 蒙特卡洛积分：路径追踪是解决渲染方程最核心、最主流的蒙特卡洛积分方法。它通过随机采样（选择光线方向、光路长度）来估算复杂的多维积分。

- 与传统光栅化的对比：主要处理直接光照的实时光栅化渲染（如 OpenGL/DirectX 的传统管线），光线追踪（尤其是路径追踪）能够更自然、物理准确地模拟光线的全局传播行为，产生照片级真实感图像，但计算代价巨大。近年来，硬件光追加速（如 RT Core）和 AI 降噪（如 DLSS/FSR）使其在实时应用（游戏）中变得可行。

- 路径追踪的演进：从 Whitted 风格的光线追踪（仅处理镜面反射和折射），发展到分布式光线追踪（Cook, 1984, 处理模糊效果），再到 Kajiya 提出的路径追踪（1986），以及后续的一些重大改进（如双向路径追踪 BDPT、Metropolis 光线追踪 MLT、光子映射 Photon Mapping、辐射度缓存 Irradiance Cache 等）。

近年来，图形硬件开始包含专用核心来加速光线路径追踪，有效增强了场景的真实感、赋予场景深度感等，使其更具沉浸和视觉吸引力，对光栅化的渲染解决方案逐步出现替代趋势。但计算效率、算法通用性、成本效益等仍是推动该技术从少数高端应用走向普及的关键诉求，需要学术界和工业界的持续协同创新。光线追踪代表了通往终极真实感渲染的一条重要且充满活力的算法路径。

技术挑战

光线路径追踪尽管是物理正确方法可行的框架，其应用仍面临巨大挑战：

1. 计算复杂度高（噪声与收敛）：蒙特卡洛方法的固有缺陷是噪声（方差）。在复杂光照场景（如焦散、漫反射主导的间接光照）或低光照区域，需要采集巨量样本才能获得低噪声图像。计算时间成为核心瓶颈。

2. 渲染方程求解难度：

2.1 高维积分：渲染方程涉及高维（波长、时间、空间方向）积分，蒙特卡洛积分效率对采样策略极其敏感。

2.2 复杂材质建模：对高度各向异性、多层、微表面结构材质（如金属拉丝、清漆、毛发布料）进行准确的 BRDF 建模和高效重要性采样是难点。

2.3 参与介质（体积光）：烟、雾、火焰等非均匀介质中的散射和吸收计算（需追踪穿过介质的光线），大大增加光线追踪路径的复杂度和计算量。

3. 场景几何复杂性：显示场景包含数百万乃至数十亿三角形，高效、快速地进行光线-三角形求交的复杂挑战，常依赖于 BVH、KD-Tree 等加速结构。加速结构本身的内存占用和构建/更新时间也构成挑战，同时动态场景更新加速结构开销更大。

4. 高效采样与降噪：

4.1 采样效率：设计有效的重要性采样策略（充分利用 BRDF、光源、材质等信息引导光线走向）是减少方差、加速收敛的核心，自适应采样对降低采样的开销也至关重要。

4.2 降噪：为了在有限样本数下获得可用图像，时域+空域结合的降噪成为实时光追的关键技术。当前基于机器学习的降噪（DLSS/FSR Ray Reconstruction）效果显著，但也依赖样本历史、运动矢量等信息，易出现鬼影、残影等问题，稳定性和通用性亦待提高。

5. 实时性能要求：为了在实时应用（如 60 帧以上）中需平衡处理帧率与保真效果，对硬件架构（光追加速单元）、软件算法优化和降噪技术提出了极高要求，需系统性提升。

当前结果

光线追踪相关技术在多个领域取得较为显著的成果，如电影与视觉特效中 V-Ray, Arnold, RenderMan, Redshift 等商业渲染器广泛采用路径追踪及其变种，为电影工业生成无噪点的照片级图像。在游戏的实时渲染中，NVIDIA RTX 系列 GPU（RT Core）、AMD RDNA 2/3（Ray Accelerators）提供专用硬件光追单元，显著提升了实时光追性能表现。软件生态上，Microsoft

DXR、Vulkan Ray Tracing、Khronos GLSL_ray_tracing 等 API 普及了软件生态。Unreal Engine 5（Lumen - 混合方案显著依赖路径追踪）、Frostbite（战地系列）、Decima（死亡搁浅）等主流引擎深度集成了路径追踪技术。同时，NVIDIA DLSS Ray Reconstruction（基于 AI 的超分辨率+时空降噪），AMD FidelityFX Super Resolution（FSR）等极大缓解了实时路径追踪的采样数瓶颈，以更少样本获得可接受画质。

基于当前的发展态势，出现了更多的机会点，算法的进步尤为重要：

1. 采样算法改进：多重重要性采样（MIS）结合材质和光源采样效果显著。自适应采样进一步发展。
2. 渲染框架优化：如 NVIDIA OptiX、Embree 等高度优化的光追框架充分利用硬件特性（RT Core）和并行计算。
3. 混合渲染架构（Hybrid Rendering）：结合光栅化完成主要画面（GBuffer）、使用光追处理特定全局光照效果（如反射、阴影、AO），是实现实时高画质的主流策略（UE5 Nanite + Lumen）。
4. 其他渲染方式：可微渲染（Differentiable Rendering）结合深度学习进行逆向渲染（材质估算、光照求解）、神经光场/辐射场（NeRF）等新兴领域也建立在路径追踪或其变体基础上。

技术诉求

为提升光线追踪技术的效能、普及度和应用范围，未来研究和技术发展主要关注以下核心诉求，提出相应高效率算法优化任务：

1. 优化采样策略：发展对复杂高维场景具有更强适应性的自适应性重要性采样算法，更高效地引导光线。
2. 提升降噪性能：开发通用性强、稳定性高、对运动场景鲁棒的 AI 降噪技术，进一步降低实时应用所需的最小有效样本数。探索少样本、零样本重建的可能性。
3. 提升计算性能：深度优化算法流程，让算法并行化/矢量化运行，提升存在大规模参与介质的场景的渲染速度，研发更高效的动态加速结构（BVH），发展能高效处理各类复杂物理材质模型及其采样方法的统一框架等。

参考文献：

1. Kajiya, J. T. (1986). The rendering equation. Proceedings of the 13th annual

- conference on Computer graphics and interactive techniques (SIGGRAPH '86), 143-150.
2. Whitted, T. (1980). An improved illumination model for shaded display. *Communications of the ACM*, 23(6), 343-349.
 3. Cook, R. L., Porter, T., & Carpenter, L. (1984). Distributed ray tracing. *Proceedings of the 11th annual conference on Computer graphics and interactive techniques (SIGGRAPH '84)*, 137-145.
 4. Veach, E., & Guibas, L. J. (1995). Optimally combining sampling techniques for Monte Carlo rendering. *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (SIGGRAPH '95)*, 419-428.
 5. Jensen, H. W. (2001). *Realistic image synthesis using photon mapping*. AK Peters/CRC Press.
 6. Hachisuka, T., & Jensen, H. W. (2009). Stochastic progressive photon mapping. *ACM Transactions on Graphics (TOG)*, 28(5), 1-8.
 7. Pharr, M., Jakob, W., & Humphreys, G. (2016). *Physically Based Rendering: From Theory to Implementation* (3rd ed.). Morgan Kaufmann.
 8. Wald, I., et al. (2019). RTX Beyond Ray Tracing: Experiences on Developing an Efficient AI-Accelerated Denoiser. *ACM SIGGRAPH 2019 Talks (SIGGRAPH '19)*, Article 68, 1-2. (OptiX & AI Denoising)
 9. Akenine-Möller, T., Haines, E., Hoffman, N., Pesce, A., Iwanicki, M., & Hillaire, S. (2018). *Real-Time Rendering* (4th ed.). A K Peters/CRC Press.
 10. Microsoft. (2023). *DirectX Raytracing (DXR) Documentation*. Microsoft Learn.
 11. Khronos Group. (2023). *Vulkan Ray Tracing Final Specification*. Khronos.org.
 12. NVIDIA. (2023). *NVIDIA DLSS Ray Reconstruction*. NVIDIA Developer.
 13. Epic Games. (2023). *Lumen Global Illumination and Reflections in Unreal Engine*. Unreal Engine Documentation.
 14. Boksansky, J., & Křivánek, J. (2021). Temporal Consistency for Real-Time Path Tracing. *Computer Graphics Forum*, 40(8), 41-54.
 15. Krizar A, Woda M. Proposal of an Optimised Path Tracing Algorithm for Computer Graphics[C]//International Conference on Dependability and Complex Systems. Springer, Cham, 2025: 117-128.
 16. DiDio T A. Optimized ray tracing for real time use without hardware accelerators[D]. Virginia Tech, 2025.
 17. Stadlbauer P, Tatzgern W, Mueller J H, et al. Adaptive Multi-view Radiance Caching for Heterogeneous Participating Media[C]//Computer Graphics Forum. 2025: e70051.

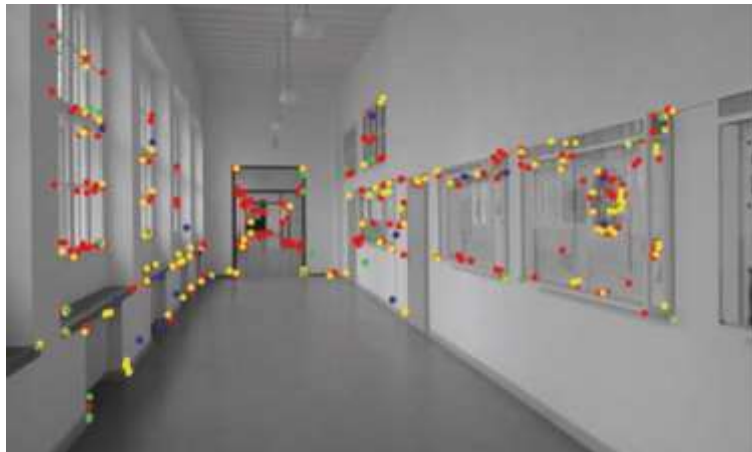
1.10 高性能特征跟踪和 3D 空间推理

技术背景

三维空间推理是一种对真实空间的理解、操作和推理空间关系的能力，其中包括空间理解和视觉推理。空间理解指的是系统通过感知环境中的几何结构、物体位置和空间关系，构

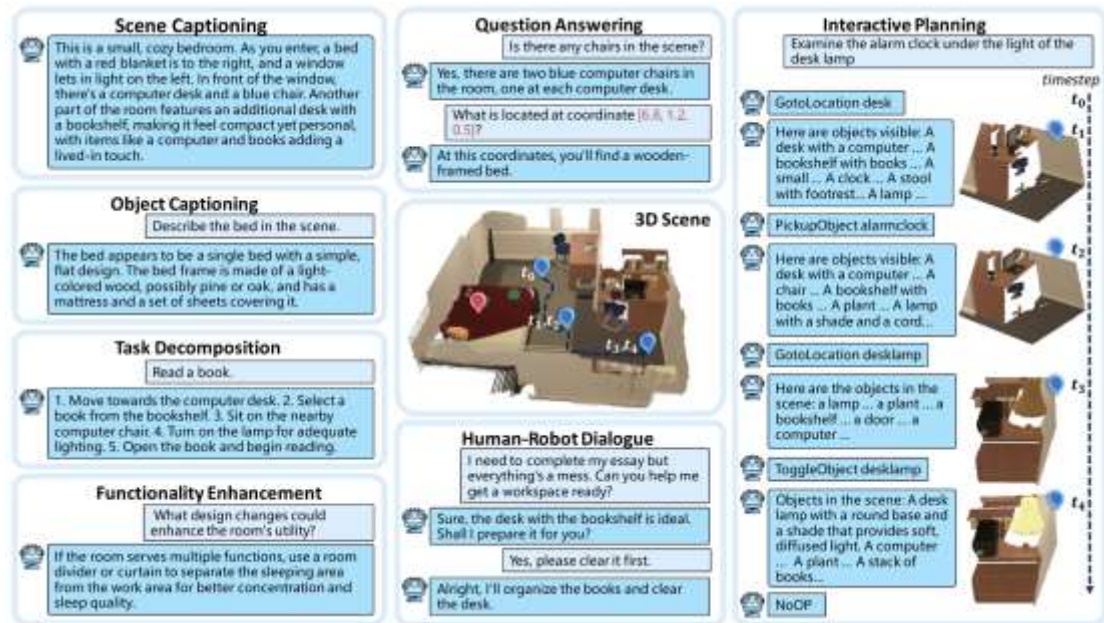
建出对三维世界的认知模型。这通常依赖深度图、点云、摄像头位姿等数据进行分析，结合同步定位与地图构建（SLAM）、空间语义分割和三维重建等算法，让机器能够感知「哪里是什么」。

其中，SLAM 是 XR 头盔实现空间定位功能的基础技术，XR 头盔上 SLAM 的实现依赖全局曝光（global shutter）的鱼眼相机和惯性测量单元（IMU）两个关键器件，通过信息融合的方法输出轨迹信息（6DOF，six degree of freedom），其中鱼眼相机的视觉处理和跟踪是实现高精度、高鲁棒性 SLAM 的关键技术。现阶段 XR 头盔的 slam 视觉跟踪，采用传统的特征点提取和跟踪方案，特征点提取通常采用角度提取的方法，跟踪采用光流跟踪和描述子全局跟踪的方法，如下图：



由于 XR 头盔的使用场景复杂，有白天、黑夜等光照变化差异大的场景，也有弱纹理、重复纹理等不同纹理的场景，还有商场、办公室、室外等场景大小不一的场景。技术上，传统的特征提取和跟踪方法很难实现全场景覆盖。采用点、线和结构特征多模态特征融合的方式可以显著提高复杂场景的特征鲁棒性。而基于深度学习的方法其表现更优于传统的算法。

另外，视觉推理则是在获取图像或视频的基础上进一步进行高层次的信息分析与逻辑判断，它不仅需要识别图像中的物体，还要理解这些物体之间的关系、交互方式，甚至推断出场景背后的语义信息或动作意图。技术上，它涉及目标检测、图神经网络、视觉问答系统（VQA）、以及用于关系建模的推理模块（如 Scene Graph 或 Relation Network）。两者常被结合应用于智能体（如机器人或 XR 设备），前者提供对环境的结构化建模，后者赋予系统对视觉内容的理解与决策能力，从而实现如 3D 场景问题、语义感知与语音交互、自动导航与任务规划、人机协作中的意图理解等更高级的智能行为。



子任务 1：基于深度神经网络的点、线、结构特征提取和描述方法

技术挑战

端测部署：深度网络浮点运算需求导致高功耗，模型压缩技术（如 8 位量化+通道剪枝等）可降低计算量，但效果也随之下降。在边缘设备（如 XR 眼镜或机器人）上高帧率实时运行，且保持良好效果与体验成为关键挑战

挑战场景的匹配跟踪：SLAM 需要特征能长时间的关联，线或结构特征的提取与描述没有那么稳定，容易跟踪丢失，难以对 SLAM 后端的 POSE 结算产生很好的约束

模型的泛化能力：XR 头显使用场景广泛，需满足不同纹理、光照、场景和动作等。

当前结果

深度学习点特征经典算法 Supperpoint，较传统点特征有更好的鲁棒性，已有厂家在嵌入式设改进后使用。PLNet 采用统一架构同时检测点线特征，结合自监督点特征学习和单阶段线段检测，在弱纹理场景中表现优于传统方法。特征提取与跟踪逐步有原来的点特征，向线、结构和多模态特征发展，可以更好的支撑 SLAM 在不同场景下的定位鲁棒性。目前需要还需要优化在嵌入式设备上耗时和功耗问题，以及全场景泛化等问题。

技术诉求

1、在光照变化，动态场景、弱纹理、重复纹理等场景鲁棒性更高；

单目特征数目>200，单帧匹配成功率>50%，正确率>90%，小范围运动连续 10 帧跟踪正确

匹配特征数>30

- 2、性能消耗更小，耗时更短，30Hz 帧率四目功耗<200mW,四目耗时<15ms;
- 3、在 DSP/NPU 平台上实现

子任务 2：基于深度神经网络的点、线、结构特征提取和描述方法

技术挑战

端侧部署：场景理解涉及大规模空间信息处理，若模型引入 3D Transformer 等结构，计算资源消耗大，难以部署在边缘设备（如 XR 眼镜或机器人）上实时运行；

缺乏物理一致性：模型可能生成违反物理逻辑的答案，如“椅子在天花板上”，缺乏对物理约束、支持关系、遮挡逻辑的内在建模；

缺乏时间一致性：当前模型多以单轮输入进行推理，难以在多轮对话中维护空间上下文，场景状态变化后，模型难以持续追踪，如“现在门开着了吗？”；

缺乏鲁棒性与安全性保障：对输入语言中的歧义或模糊描述，如“靠近窗户的那个桌子”敏感，易误解；对场景感知失败或遮挡处理能力不足，可能输出不可信推理结果；

场景结构泛化性差：模型可能在训练集中学会“投机策略”（shortcuts），但对新场景或新描述泛化性能弱，特别是在开放世界或未知物体场景中，易出现错误理解。

当前结果

当前 Scene-LLM 较具代表性，实现 AI 模型从“语言-视觉”走向“语言-视觉-空间”的关键进化，技术融合度高、应用潜力广，是未来人机协作、XR、认知机器人等场景的核心支撑。但现阶段仍需在几何建模、知识融合、多模态训练框架、以及推理可解释性方面进一步突破，才能实现真正“场景可理解、语义可交互、行为可执行”的智能系统。

技术诉求

可选择如下部分实现：

- 实现三维空间推理的端侧部署；
- 解决当前模型的物理和时间一致性问题；
- 提升模型的鲁棒性和泛化性。

参考文献

-
- [1] Fu, R., Liu, J., Chen, X., Nie, Y., & Xiong, W. (2025, February). Scene-llm: Extending language model for 3d visual reasoning. In 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (pp. 2195-2206).
 - [2] Man, Y., Gui, L. Y., & Wang, Y. X. (2024). Situational awareness matters in 3d vision language reasoning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 13678-13688).
 - [3] Ha, H., & Song, S. (2022). Semantic abstraction: Open-world 3d scene understanding from 2d vision-language models. arXiv preprint arXiv:2207.11514.
 - [4] Chen, S., Chen, X., Zhang, C., Li, M., Yu, G., Fei, H., ... & Chen, T. (2024). Ll3da: Visual interactive instruction tuning for omni-3d understanding reasoning and planning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 26428-26438).
 - [5] Liu, R., Wang, W., & Yang, Y. (2024). Volumetric environment representation for vision-language navigation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 16317-16328).

二、 材料与工艺

2.1 移动终端产品散热技术创新

技术背景

近年来，以手机为代表的移动终端产品，性能不断提升，发热问题也愈加突出，散热逐渐成为移动终端产品的一大瓶颈。如果散热较差，一方面会导致器件温度过高，影响可靠性，另一方面会导致表面温度过高，直接影响用户热体验。有时为了缓解发热问题，不得不对性能做些限制，这样又影响了产品的性能发挥。因此，亟需提升移动终端产品的散热能力，以保障移动终端产品满足用户持续提升的需求，从而打造更加优越的行业竞争力。

手机常用的散热材料如下图所示，主要包含均热板、石墨膜等，经过产品的不断迭代演进，其性能提升已经非常缓慢。另一方面，手机产品对整机外观、重量等有着较高的要求，如何利用更小的结构空间实现散热能力的提升也是一大挑战。因此，需要做更多的散热技术创新，来进一步提升手机的散热能力。

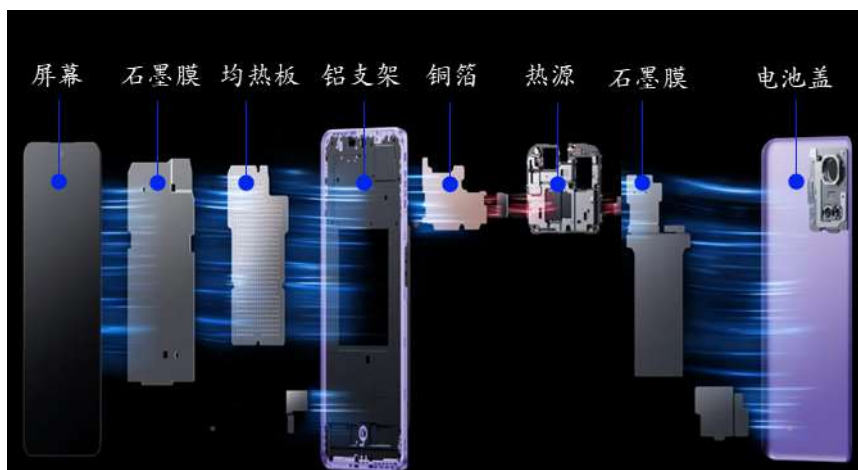


图 1. 手机散热材料组成示意图

技术挑战

1) 目前手机主流散热方案仍是被动散热。对于手机产品而言，其外观形态已较为稳定，多年来并无较大变化，这意味着手机产品的散热表面积基本是固定不变的，因此对于传统自然散热技术而言，必然存在着散热极限。随着超大均热板的应用，手机均温能力足够优秀之后，散热能力已愈加逼近该极限。各手机厂商在散热能力表现上也无法拉开较大差距，因此如何在保证不牺牲手机外观设计的情况下继续提升散热成为行业一大难题。

2) 主动散热是进一步提升移动终端产品散热能力的重要方向之一，但传统风扇存在一系列的负面影响，如外观（风扇尺寸大）、噪声、功耗（功耗大影响续航）、可靠性（防水、防尘）等方面，如何最大程度减小这些负面影响，是推广主动散热技术的主要瓶颈。

当前结果

目前主流手机厂商高端旗舰机的散热方案均为铝合金中框搭配均热板和石墨，各家的散热技术宣传也都集中在均热板和石墨上，包括一加宣传的天工散热系统（如图 2 所示），运用超大面积的均热板，以及将多个均热板叠加，以实现更好的均温，石墨的导热系数宣称可由传统的 1200 Wm-1K-1 导热系数提升至 2000 Wm-1K-1；小米宣传的环路热管（如图 3 所示）另辟蹊径，宣称均热板内部蒸汽与液体分离可实现更高的 VC 传热性能。但这些方案之间大同小异，均是提升均热表现，当均热面积足够大之后，散热表现已无法拉开显著差距。

除此之外，为了实现更极致的游戏性能发挥，突破当前自然散热瓶颈，红魔手机甚至采用了微型风扇来辅助手机散热（如图 4 所示）。通过微型风扇，将手机内部热量吹送至外界环境中实现降温。但受限于产品外观，微型风扇尺寸较小，所带来的散热提升作用仍不足以满足所有超重载游戏的散热需求。另一方面，微型风扇会引入可靠性、噪声等问题，行业内目前仍未普及。

为了应对更高的散热需求，除了手机本身的散热设计，近几年在手机配件上，也有一些散热新技术发布，如 OPPO 曾发布过冰肤散热保护壳（如图 5 所示，内部增加水凝胶延缓升温速度，但存在回水较慢的问题），华为发布了微泵液冷保护套等。但上述配件其实对温升改善效果仍不是很显著，且代价过高（成本、厚度等）。目前手机行业应用最多、效果较显著的配件仍是散热背夹（如图 6 所示），通过 TEC+散热器+轴流风扇的组合方式可实现 10°C+ 的降温效果，但散热背夹普遍体积过大、重量过重并且需外接电源，在用户端的普及上始终有很大阻力。



图 2 一加 ACE3 Pro 天工散热系统



图 3 小米环形冷泵



图 4 红魔手机内部加风扇提升散热



图 5 OPPO 联合武汉大学发布冰肤散热保护壳



图 6 散热背夹图示

技术诉求

任务:探索除上述散热技术外,能够进一步提升移动终端产品散热能力的新技术或材料,包含被动散热和主动散热,具体如下:

1. 被动散热技术创新:

1) 现有技术的显著改良:通过采用新技术,能够显著提升现有均热板、石墨等性能,或者保持性能不变,实现更轻薄,如均热板厚度减至 0.23mm 以下时等效导热系数仍有 5000 Wm-1K-1 以上)。目前手机行业均热板尺寸约 100mm*70mm*0.3mm (长*宽*厚),主要材质有不锈钢、铜合金、钛合金,等效导热系数约 8000 Wm-1K-1。石墨类导热系数目前最高为 2000 Wm-1K-1。

2) 除均热板、石墨类的常规均热技术外,其他形式的散热新技术:不限于用在产品内部,可单独作为散热配件。如水凝胶(通过蒸发水分降温并在水分蒸干之后主动从环境吸收水分)、特殊制冷技术(如通过物理变化或化学反应制冷),或者其他具有吸热作用的新技术,能够对手机快速降温(降温幅度 $\geq 2^{\circ}\text{C}$),持续作用 1h 以上,具有重复使用的可行性(非一次性,区别于退烧贴),且当降温功能失效后可人为或主动快速恢复至原状态,工作时无需外接电源,方便用户携带。

3) 提升热体验的新技术/材料:用于手机电池盖或保护套等,作为外观结构件,具有“冰感”的作用,即:同样温度下,相比其他材质(玻璃、皮革、塑胶)触摸起来更凉爽。

2. 主动散热相关技术创新：

1) 新形态微型风冷技术创新：新形态的风冷散热技术，相比目前行业所用的微型离心风扇，能够在性能（PQ）、功耗、体积、噪声、可靠性等方面有明显优势。期望该散热器件本身尺寸控制在 20mm*20mm*4mm（长*宽*高）左右（厚度尺寸需求优先级高于平面尺寸，越薄越好），功耗不高于 0.3W，噪声（0.3m）小于 30db（A），最大风量不低于 0.5CFM，最大静压不低于 100Pa（该要求仅供参考，如个别方面优势显著，可适当放宽其他方面的要求，综合表现领先微型离心风扇）。

2) 风冷降噪技术研究：能够用于手机内部的微型风道（宽约 10mm、高度 $\leq$ 2mm）内，不会占用过多空间，主要目的降低风扇 1-6K HZ 范围的噪声。

3、可折叠散热技术创新：导热系数或等效导热系数优于行业目前所用的柔性石墨（ $\geq 1400 \text{ Wm}^{-1}\text{K}^{-1}$ ），厚度不超过 0.25mm，耐弯折性好（用于折叠机），可实现 30W 次弯折（0-180°弯折，最小弯折半径 0.8mm）无明显损伤、导热系数衰减低于 10%

2.2 电池在线 EIS 检测算法及失效状态评估

技术背景

电池随时间使用会老化衰减，充放电性能下降，其中包括最大容量、峰值功率下降、电池阻抗增大，因此电池管理系统需要能够实时获取电池当前的最真实状态，以满足任何场景下的电池安全和性能要求。另外，随着硅体系电池的发展和应用，手机的续航得到了巨大的提升，但同时也伴随着硅损失和产气鼓包的风险。为了解决以上问题，我们需要在现有电池管理系统基础上拓展研究更加先进的电池检测方式，以提升状态估计精度、更精准的识别电池安全风险、优化电池充放电管理算法。电池电化学阻抗谱(EIS)是表征电池特征的重要参数，近些年来已经广泛应用于新能源汽车、储能和电子类产品的 BMS 开发，其快速、无损、原位表征等特点使其成为目前最先进的电池检测方法之一，对电池状态估计、失效预警和充放电控制都提供了关键的特征参数。

技术挑战

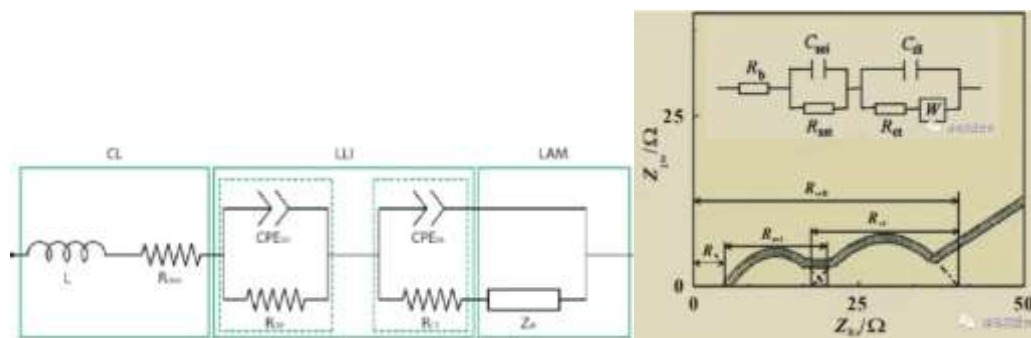
电池 EIS 阻抗检测的测量方法是对电池施加小幅度的正弦波电位或电流扰动，进而测量系统产生的相应电流或电位响应，进而计算得出电池阻抗图谱。但在整机环境下，正弦电

流/电位激励的实现有一定难度，受制于许多因素的影响如频率范围、电流幅度精度等，以及电压响应的测量存在系统本身带来的噪声、检测采样率不足，造成精度差等问题。

目前应用较为普遍的方式是采用小波变换或者 S 变换等时频转换方法，基于一段脉冲充电或者放电的工况，提取出频域的阻抗特征，进而获取电池 EIS。但是这种方法也存在诸多问题，比如频率范围较低、温度变化、采样周期&采样跳变等均会影响 EIS 估计精度。

技术现状

充电过程中，控制充电器实现恒流充电+停止充电一段时间制造脉冲工况，检测电流&电压波形，通过公式 $R=dV/dI$ 得到一定频率范围内的阻抗，用相对增长量来表征电池衰减情况。



技术诉求

任务一、基于电芯 EIS 实测数据和脉冲数据（采样周期 100ms~500ms），建立 EIS 在线检测模型。

要求：

1. 采用小波变换或 S 变换等时频转换算法，得到不同频率范围的 EIS，输出模型代码和算法说明文档；
2. 模型 EIS 频率范围至少达到 10Hz~0.01Hz；
3. 在任何老化状态、温度、SOC 条件下，模型 EIS 与实际 EIS 之间的均方根误差 $\leq 3\%$ 、不同老化路径电池的模型 EIS 均满足该误差要求；
4. EIS 的计算在充电、放电工况下均能够触发，输出不同场景下算法触发的对应工况说明；

任务二、基于 ECM 和机器学习等方法，研究电池失效&老化的阻抗特性、建立 EIS 与电池失效预警&老化估计模型

要求：

- 1.输出 EIS&DRT 与膨胀程度（膨胀体积 15%、膨胀速率或电芯压力）&析锂（电芯压力、面积和容量）的关联关系。
- 2.准确识别异常老化，同时输出异常老化电池归因分析：识别到最主要失效原因。
- 3.输出 EIS&DRT 与 SOH 关联关系。

2.3 6.75MHz 高功率可集成逆变整流控制系统

技术背景

无线充电在便携终端上具有广泛的应用。目前手机摄像头持续变大，凸包变厚，导致现在 Qi 协议无线充自由度差的劣势愈发明显：凸包严重，手机背部被顶起，导致无线充电距离变远、线圈设计偏下，与第三方无线充底座难于失配。车载无线充很难找到最佳充电位置。随着工信部进一步放开无线充频率、功率限制[1]，磁共振无线充技术获得了可发展的空间。工信部明确 6.78MHz 为无线充电可用频率，且最高功率限制提升至 80W[1,2]。磁共振技术为无线充发展方向之一，它具有多种优势：1.无线充电自由度大，可以解决线圈位置偏移，凸包带来的充电效率下降问题。2.工作频率提升，线圈电感量设计值小，内部占据空间变小。3.充电距离远，可实现“边玩边充”。

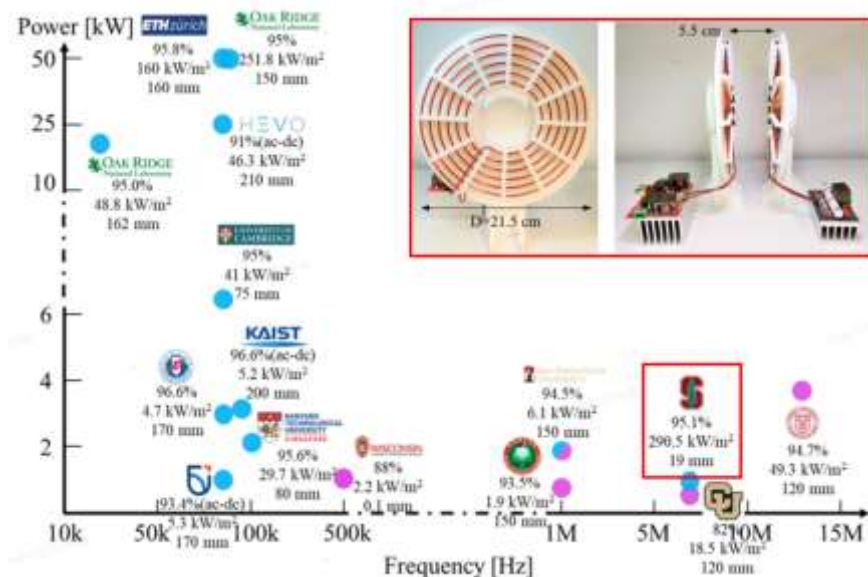


图 1. 磁共振无线充电性能

技术挑战

1. 逆变系统

不同于现有 Qi 协议无线充系统可以使用全桥逆变，射频高功率（6.75MHz）需要引入 E 类开关逆变电路等功率放大技术。随着工作频率提升至 6.78MHz，系统需要考虑射频阻抗匹配问题。随着接收线圈移动，发射电路的负载阻抗有较大波动，系统很难维持在最高效率。

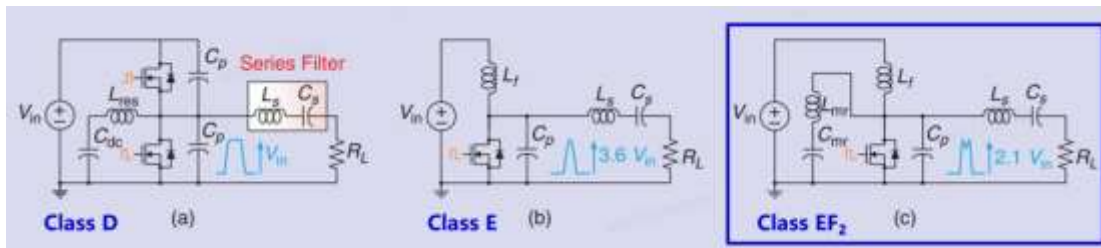


图 2. 射频功率放大器电路

2. 耦合系统

无线充电线圈的磁场作用范围内需要相对均匀的磁场分布。一方面传统磁屏蔽材料-纳米晶随着工作频率增加磁损急剧增加，更均匀的磁场分布可以降低磁损耗影响。另一方面更均匀的磁场分布可提高充电区域内移动功率稳定性，实现倾斜翻转等多种充电姿态。因此在耦合系统设计中，需要考虑充电线圈走线的设计，使得无线充电区域的磁场分布均匀。

3. 整流系统

在无线能量传输接收端，需要研究射频能量到直流能量转换的整流器设计。业界现有集成同步整流控制器最高只支持约 2MHz，6.78MHz 无线充电接收端整流效率低，功率无法达到现有 kHz 水平。由于在连续导通模式（CCM）下，整流电流的快速变化可能导致突变。在高 di/dt 条件下，同步整流控制器难以在电流为零时及时关闭同步整流（SR）开关管，这可能引起额外的开关损耗和增加可靠性风险。设计能够在 CCM 下快速响应电流突变的同步整流控制器面临着技术挑战。

4. 功率反馈控制系统

不同位置和充电姿态情况下耦合参数会发生大范围改变，为了维持负载端充电功率的稳定，需要逆变系统需要根据输出负载的变化来调整 Tx 输入功率、工作频率和信号占空比等参数来维持负载端充电功率的稳定。因此需要通过功率反馈控制技术进行动态响应调整，具有设计上的挑战。

技术现状

目前业界集成同步整流控制系统只能支持不超过 2MHz[3-5]。在磁共振技术方面还是只能维持在固定位置和姿态的高功率高效率充电，动态调整范围小，逆变整流系统效率低，未达到目前 Qi 协议无线充水平。

技术诉求（完成任一）

任务一：6.765-6.795MHz 工作频率范围的高效率高动态范围逆变控制器（switch-mode inverter）电路和控制算法。

1. 逆变开关管系统实现至少 30W 稳态输出功率，最高逆变效率不低于 93%。
2. 逆变系统有阻抗匹配调整能力，输出阻抗在 3-50 Ω 动态变换过程中，维持最高输出功率且效率不低于 80%，对输入阻抗做寻优研究。

任务二：6.765-6.795MHz 工作频率范围的高效可集成全桥同步整流控制器电路和控制算法

1. 控制器控制 SR 开关管导通占比大于 90%，二极管导通占比小于 10%，预期整流桥最高效率 $\geq 98\%$ @10V3A。
2. 控制器支持连续导通模式下电流非过零关断。
3. 5-30W 输入功率动态变化下输出电压不低于 9.9V，动态变化过程中最低效率不低于 80%。

任务三：设计在 6.765-6.795MHz 工作频率范围内工作的无线充线圈模组。[6]

1. Tx 大小约 200*200mm, Rx 大小约 60*60mm。线圈背部均需覆盖锰锌铁氧体，纳米晶等磁屏蔽材料。线圈背部放置金属不影响系统工作。
2. Rx 线圈模组厚度不超过 0.4mm，交流阻抗 Rs 小于 200m Ω 。对电流分布及磁场均匀性进行针对性研究。
3. Rx 和 Tx 最佳耦合位置 S21 大于 -0.8dB。Rx 在 Tx 范围内自由移动，倾斜（倾角大于 45°），移动过程中 S21 大于 -3dB，避免出线分频，零耦合现象。
4. 耦合谐振系统有调制解调的能力，可以做到带内实时通信，在上述自由度范围内通信不中断。

参考文献

- [1] https://www.gov.cn/zhengce/zhengceku/202305/content_6883940.htm
- [2] <https://www.itu.int/rec/R-REC-SM.2129-0-201908-I/en>
- [3] <https://www.ti.com/product/UCD7138>
- [4] <https://www.ti.com/product/UCC24612>
- [5] <https://www.ti.com/seclit/ml/slup379/slup379.pdf>
- [6] Y. Zhuang, A. Chen, C. Xu, Y. Huang, H. Zhao and J. Zhou, "Range-Adaptive Wireless Power Transfer Based on Differential Coupling Using Multiple Bidirectional Coils," in IEEE Transactions on Industrial Electronics, vol. 67, no. 9, pp. 7519-7528, Sept. 2020, doi: 10.1109/TIE.2019.2945304.

2.4 高性能广色域 BT2020 发光材料器件研究

技术背景

显示屏 RGB 三子像素显示的 CIE 组成的色域，决定了这个显示屏的色彩丰富程度，显示屏的色域越广，能够显示的颜色越多，屏幕越鲜艳，同时也能够更加真实的还原现实世界。当前主流的显示材料器件基本采用 DCI-P3(Digital Cinema Initiatives-Protocol 3)为标准，而 DIC-P3 对于指针色域(自然界内所有颜色的合集)的覆盖只有 86.9%(CIE1931 计算)。2012 年国际标准电信联盟定义 BT. 2020(也写作 Rec. 2020)色域，其对指针色域的覆盖能够达到 99.9%，是基于 CIE1931 色彩空间内三色显示最广的色域。

标准 DCI-P3，BT. 2020 以及当前旗舰手机屏幕(以典型 OLED 为例)的 CIE 坐标以及图示如下(CIE1931 Lxy 色彩空间):

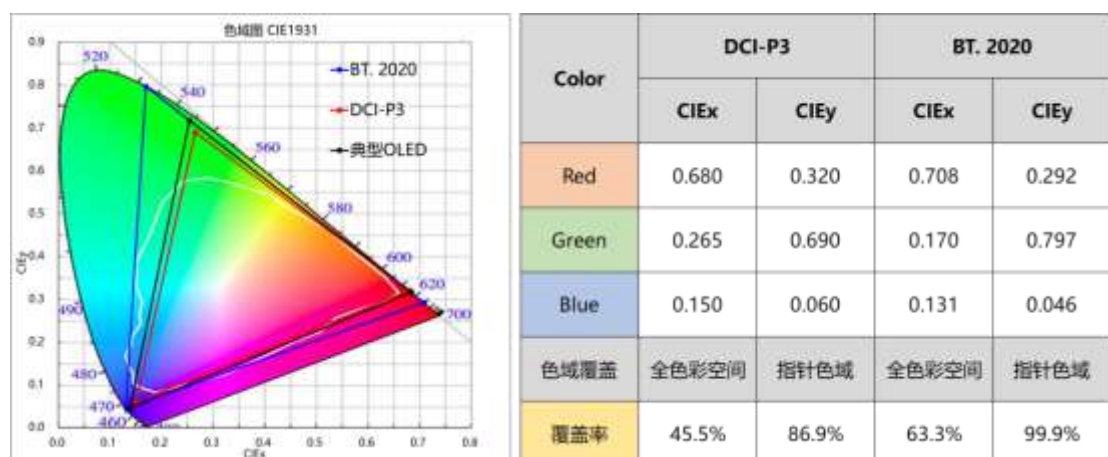


图 1 DCI-P3 和 BT2020 的图示以及对应的 CIE

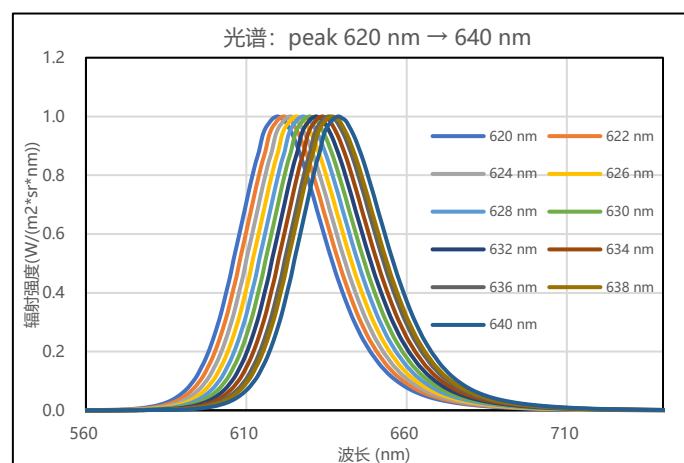
技术挑战

对于 BT.2020 绿光器件，其 CIE 为(0.170,0.797)，选点对应的是 532 nm 的单色光，要实现此绿光，对绿光光谱的半峰宽和对称性要求非常高，需要半峰宽在 15 nm 以下，同时光谱对称性(Peak 左边的面积/Peak 右边的面积)在 0.95 以上，才能保证 BT.2020 色域的覆盖率在 98%以上。除了对绿光发光光谱的高要求以外，同时还要保证器件的发光效率和寿命，对绿光器件而言，挑战巨大。

光谱形状	峰值 (nm)	半峰宽 (nm)	CIEx	CIEy	R/B为标准BT2020色点	
					BT2020	DCI-P3
高斯分布 (理想)	532	FWHM=1 nm	0.170	0.797	100.0%	100.0%
	532	FWHM=5 nm	0.171	0.796	99.8%	100.0%
	532	FWHM=10 nm	0.172	0.794	99.5%	100.0%
	532	FWHM=15 nm	0.175	0.790	98.8%	100.0%
	532	FWHM=20 nm	0.179	0.783	97.6%	100.0%
	532	FWHM=25 nm	0.184	0.775	96.2%	100.0%
	532	FWHM=30 nm	0.190	0.764	94.3%	99.9%

图 2 不同半峰宽的绿光器件对应的 CIE 以及色域

对于 BT. 2020 红光器件，其 CIE 为(0.708,0.292)，选点对应的是 630 nm 的单色光，由于人眼对于红光颜色的敏感性差于绿光，对于红光器件的发光光谱而言，只需要把 Peak 设计到 ≥ 630 nm，对应的半峰宽在 40 nm 即可。然而对于红光器件挑战最大的是其发光效率，由于人眼视锥细胞的特征，人眼对于波长越大的红光越不敏感，到 780 nm 以外人眼就无法察觉了，对应的视见函数计算表明，相同的辐射强度，634 nm 的红光点的发光效率比 622 nm 的红光下降 34%



主波长(nm)	620	622	624	626	628	630	632	634	636	638	640
L (nit)	12.52	11.77	11.05	10.35	9.67	9.01	8.38	7.77	7.19	6.91	6.37
CE%	106%	100%	94%	88%	82%	77%	71%	66%	61%	59%	54%
CIEx	0.6785	0.6827	0.6867	0.6903	0.6937	0.6969	0.6998	0.7026	0.7052	0.7064	0.7087
CIEy	0.3213	0.3171	0.3132	0.3096	0.3062	0.3030	0.3001	0.2973	0.2948	0.2935	0.2913

图 3 不同波长红光的发光亮度比以及 CIE

当前结果

当前学术界以及产业界业内有很多关于 BT.2020 发光材料器件的研究，一般的绿色器件的性能为 CIEx: 0.17, CIEy: 0.75, CE=180 Cd/A, LT95=500 hr (J=10 mA/cm2).

一般 BT.2020 红色发光材料器件的性能为 CIEx: 0.700, CIEy 为 0.298, CE=45 Cd/A, LT95=1000 hr-J10 (J=10 mA/cm2)。

技术诉求

任务：高性能广色域发光材料器件研究

- 1. 绿色器件 CIEx: 0.160~0.180, CIEy: ≥0.770, CE≥200 Cd/A, LT95≥1000 (J=10 mA/cm2).
- 2. 红色器件 CIEx: ≥0.700, CIEy: 0.280~0.320, CE≥55 Cd/A, LT95≥2000 hr-J10(J=10 mA/cm2).

特别说明：

- 1. 以上两个任务至少选一个.
- 2. 发光材料器件不限定于 OLED，量子点 QLED，Micro LED 等显示技术均可.

参考文献

https://tftcentral.co.uk/articles/pointers_gamut;
Pointer M R. The gamut of real surface colours[J]. Color Research & Application, 1980, 5(3): 145-155.
Series B T. Parameter values for ultra-high definition television systems for production and international programme exchange[C]//Proc. ITU-T, Bt. 2020. 2012: 1-7.
CIE. CIE 170-2: 2015. Fundamental Chromaticity Diagram with Physiological Axes--Part 2: Spectral Luminous Efficiency Functions and Chromaticity Diagrams[J]. 2016.

2.5 7 系铝合金阳极掉膜优化研究

技术背景

铝合金，凭借其成本低、轻量化等优势，成为中高端手机应用最广泛的中框材料。目前手机中框用铝合金材质的表面处理工艺为阳极氧化，如图 1 所示。

6 系铝合金（Al-Mg-Si 体系），由于强度高（屈服强度 $<500\text{MPa}$ ）、阳极外观效果好等优点，成为行业内手机（直板机）中框的主流材料。随着折叠屏手机的兴起，中框结构趋向于轻薄设计，厚度由 7-9mm 减薄至 4-5mm，6 系铝合金的强度已无法满足应用需求。7 系铝合金（Al-Zn-Mg 体系），由于具有更高的性能（屈服强度 $>500\text{MPa}$ ），逐渐被 3C 行业应用起来。



图 1 智能手机的铝合金中框

更高性能的 7 系铝合金应用于手机中框，虽然给产品可靠性带来显著提升，但目前存在一些瓶颈问题尚未解决，特别是 7 系铝合金阳极氧化后氧化膜脱落问题。数据显示：苹果、三星等品牌手机产品采用 7 系铝合金（7N03、7R03 牌号），均存在较高的市场失效率（1-2%），影响用户口碑。相比较 6 系铝合金，阳极膜脱落问题也成为制约 7 系铝合金应用的关键要素。



图 2 7 系铝合金手机中框使用后的掉膜现象

技术挑战

基于上述技术背景及问题，目前总结出的 3 大技术挑战：

(1) 7 系铝合金阳极膜脱落的机理

数据显示：7 系铝合金阳极膜脱落不良比例明显高于 6 系铝合金。而针对两种体系铝合金的失效机理尚不完全明确，比如合金成分、微观组织缺陷、阳极氧化工艺、使用环境等要素的影响。

(2) 7 系铝合金阳极膜脱落现象的表征手段

7 系铝合金阳极膜脱落失效主要来源于用户使用后的市场反馈。现有的常规测试项次，比如单体盐雾测试、耐手汗测试、高温高湿测试，以及整机循环测试等，均无法准确表征出 7 系铝合金阳极膜脱落市场失效情况，导致对优化该问题的导入措施无法准确评估。

(3) 7 系铝合金阳极氧化工艺优化

现有阳极氧化工艺主要采用纯硫酸体系，考虑到 7 系铝合金耐腐蚀性差，已逐步导入硫酸&草酸混合酸体系，但仍无法有效解决 7 系铝合金阳极膜脱落问题。

当前结果

目前行业内，终端折叠机产品均已开始应用 7 系铝合金，合金成分及性能如下表所示。成分整体差异不大，均面临阳极膜脱落问题。

材料牌号	屈服强度	Zn	Mg	Cu	Al
H	550MPa	5.5-6.5	1.2-2.2	0.10-0.50	bal.
R	530MPa	4.5-6.0	1.5-2.5	0-0.50	bal.
S	500MPa	5.0-6.0	1.5-2.2	0.25-0.50	bal.

技术诉求

根据本课题面临的技术挑战，提出以下 3 个技术诉求：

(1) 明确 7 系铝合金阳极膜脱落的失效机理及影响要素

对 7 系铝合金的阳极氧化膜脱落的演化全过程进行深入研究，明确材料、工艺及使用场景的关键要素对掉膜的影响。

(2) 设计快捷且能准确表征 7 系铝掉膜的测试方案

基于对 7 系铝合金的材料特性，设计准确表征 7 系铝合金阳极膜脱落的测试实验方案，可有效评估 7 系铝合金市场掉膜风险。

(3) 解决 7 系铝掉膜的可量产阳极氧化工艺：

基于深入的机理研究及有效的表征手法，对现有阳极工艺进行优化，解决 7 系铝合金阳

极膜脱落问题。优化的阳极氧化工艺需考虑量产应用以及外观效果。

2.6 压铸铝合金做阳极外观

技术背景

铝合金，由于其成本低、阳极色彩丰富等优点，广泛应用于手机中框、平板后盖等部件。常用的制备工序为：合金化熔铸铝锭→挤压成型制备铝板→CNC 加工铣出结构特征→阳极氧化获得外观效果。由于手机中框、平板后盖等部件内腔结构复杂，导致 CNC 加工周期长，产品的成本居高不下。



图 1 铝合金手机中框

高压压铸工艺，作为近成型工艺，常用于制备复杂结构件。采用压铸工艺制备手机中框、平板后盖等部件，可显著降低 CNC 加工时间，产生可观的降本收益。然而，采用压铸工艺存在两个技术问题：材料和外观缺陷。

（1）材料问题

常用的压铸铝合金材质为 Al-Si 系铝合金，由于高比例的 Si，材料组织内存在大量的初生 Si 相，阳极氧化过程中无法成膜，且表面产生大量黑灰。

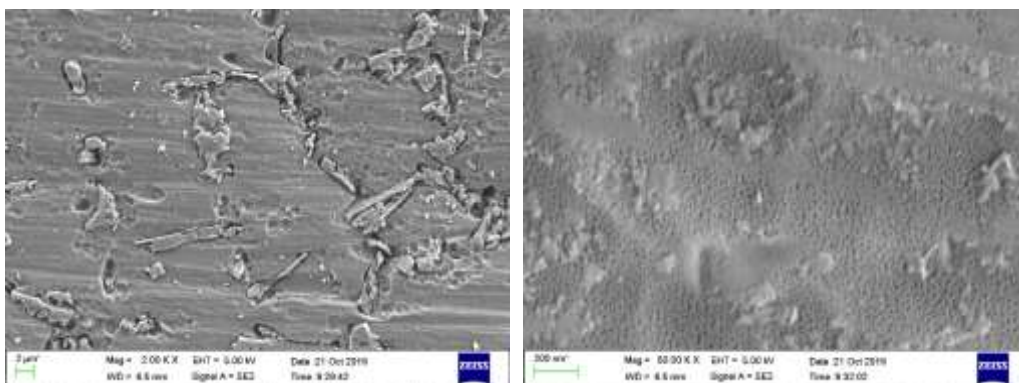


图 2 Al-Si 压铸铝合金微观组织和阳极后表面形貌

（2）外观缺陷问题

由于高压压铸工艺的局限性，容易产生气孔、砂眼等常见缺陷，影响产品外观效果，导致良率降低，无法达到降本收益。



图 3 Al-Mn 系压铸铝合金产品表面的砂孔缺陷

技术挑战

基于上述技术背景，目前总结出的两大技术挑战：

（1）兼顾高流动性、高强度、可阳极氧化的压铸铝合金材料

目前制备手机中框、平板后盖等外观件的铝合金材质主要为 6 系 Al-Mg-Si 铝合金（6013 牌号）和 7 系 Al-Zn-Mg 铝合金（7N03、7R03 等自研材料），材料强度高（屈服强度 $>340\text{MPa}$ ，延伸率 $>5\%$ ）、阳极外观效果好，但 6 系、7 系铝合金由于凝固特性差、热裂倾向高、黏膜等问题不适合压铸工艺。

Al-Si 系压铸铝合金，常应用于制备内置件。考虑到压铸流动性，Si 含量控制在 9~12% 区间，同步添加微量的 Mg、Cu 等元素，改善材料的强度（屈服强度 $>180\text{MPa}$ ，延伸率 $>4\%$ ）。由于高比例的 Si，不适宜阳极氧化表面处理工艺。

Al-Mn 系压铸铝合金，具有较好的压铸流动性，且能阳极氧化，但该体系合金强度低（屈服强度 $>70\text{MPa}$ ），且无法热处理强化，一般应用于制备结构强悍的厚壁件。手机中框、平板后盖等部件局部区域厚度 $<0.5\text{mm}$ ，应用 Al-Mn 系压铸铝合金，可靠性风险较大。

因此，开发兼顾高流动性、高强度、可阳极氧化的压铸铝材料是本课题一大关键挑战。

（2）压铸气孔、砂眼等缺陷瓶颈问题

压铸成形过程中，由于气体卷入或挥发，以及砂粒、氧化物等非金属氧化物渗入，在阳极后铸件表面易呈现肉眼可见的气孔、砂眼缺陷，且位置不固定。通过使用粗粒度的砂可部分优化该缺陷（家用电器领域有应用），但手机、平板等 3C 产品对外观精致度要求很高（砂

粒度 $>150\#$)，因此需改良现有的压铸工艺方案，或开发新的方案，降低或去除压铸过程中的缺陷。

技术现状

目前行业内可用于阳极氧化的压铸铝合金材料，多为 Al-Mn 体系。Al-Mn 体系压铸成形性较好，阳极外观效果接近于 6 系型材铝合金。但由于强度低，不可热处理强化，且当前的压铸工艺难以规避砂孔缺陷，因此多用于制备门把手、品牌 logo 以及化妆品等非 3C 产品外观件。



图 4 Al-Mn 体系可阳极压铸铝的应用

技术诉求

根据本课题面临的技术挑战，故提出以下 2 个技术诉求：

(1) 兼顾压铸成形性、高强度、可阳极氧化的压铸铝合金新材料开发

- ① 压铸成形性：能适应压铸成形薄壁件，大致尺寸 $50\text{mm} \times 100\text{mm} \times 0.5\text{mm}$ ；
- ② 高强度：压铸态屈服强度 $>180\text{MPa}$ ，伸长率 $>4\%$ ，可采用热处理时效；
- ③ 可阳极氧化：阳极氧化成膜、染色，外观效果接近当前 3C 产品喷砂阳极效果。

(2) 解决压铸气孔、砂眼及流纹缺陷的新型压铸工艺

现有的压铸工艺，难以避免成形过程中因卷气、材料氧化带来的外观缺陷，因此需要在现有的压铸工艺进行改良优化，以解决压铸工艺带来的气孔、砂眼问题。

要求铸件产品外观面的气孔直径 $\leq 30\mu\text{m}$ ，压铸流纹做 150#喷砂浅色阳极不可见。

2.7 高强超薄盖板玻璃材料设计与制造工艺协同攻关

技术背景

在消费电子领域，尤其是手机产品，轻薄化已成为重要的发展趋势。盖板玻璃作为直接

与外界接触并保护内部显示与触控组件的关键部件，其性能对手机整体品质和用户体验有着至关重要的影响。随着对手机轻薄化要求的不断提高，对盖板玻璃也提出了更为严苛的要求。

玻璃材料的密度在现有技术条件下基本保持不变，因此实现手机轻薄化的一个重要途径是减小盖板玻璃的厚度。然而，单纯地降低厚度会导致玻璃的强度等性能下降，影响其对手机屏幕的保护能力。为了在减薄的同时维持甚至提升玻璃的性能，需要对玻璃进行优化设计，其中包括对本征强度和化学强化强度的提升。

从玻璃的微观结构来看，玻璃网络结构主要由硅氧四面体等基本结构单元组成。当玻璃中存在较多非桥氧时，会破坏玻璃网络的连续性和稳定性，从而降低玻璃的性能。因此，优化玻璃网络结构，减少非桥氧的含量，对于提高玻璃性能至关重要。但这一优化过程往往会导致玻璃粘度增加，在玻璃熔制过程中容易产生气泡和条纹等缺陷。如图 1 所示，当存在较多非桥氧时（左图），网络结构变得松散和无序；正常的玻璃网络结构中（右图），硅氧四面体连接紧密且有序。

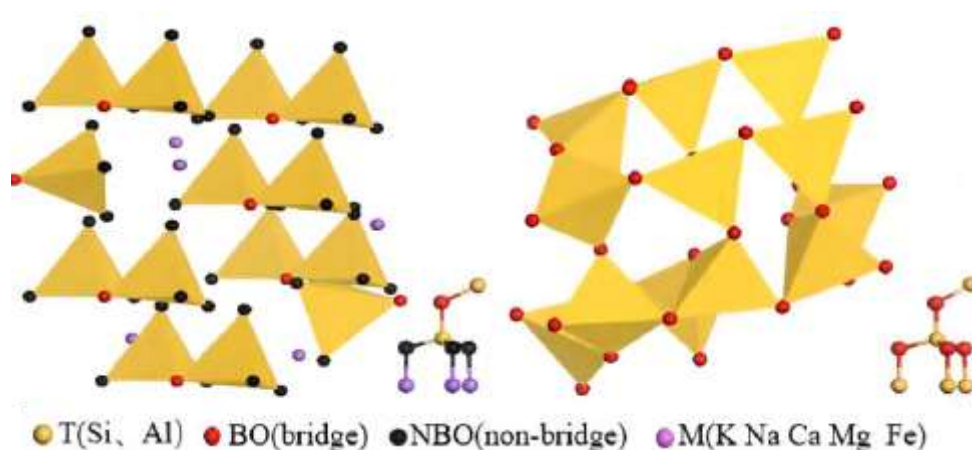


图 1：玻璃网络结构（左侧为高非桥氧结构（易熔但有缺陷），右侧为高桥氧结构（致密但粘度高））

玻璃厚度的减小和强度的提升，虽然赋予了盖板玻璃轻薄与坚固的特性，但也给其后加工带来了诸多挑战。在 CNC 加工过程中的裂纹和划伤问题（图 2）。因此，盖板玻璃不仅要具备优异的强度和轻薄特性，还必须在制造工艺上能够满足大规模生产以及复杂后加工的要求。任何制造工艺上的问题，如产品缺陷等，都可能降低产品的良品率和生产效率，进而增加生产成本。

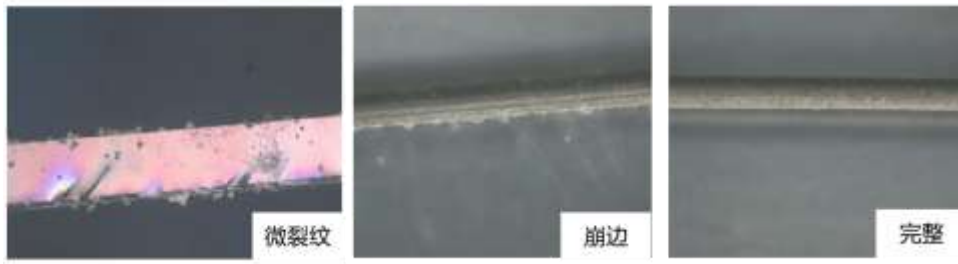


图 2：玻璃加工缺陷

技术挑战

1、玻璃性能优化与厚度减小的矛盾

在减小盖板玻璃厚度的情况下，玻璃内部缺陷对其强度的影响会被放大。例如，玻璃中的微小裂纹、杂质等缺陷在承受外力时，容易成为应力集中点，导致玻璃过早破裂。同时，厚度减薄使得玻璃整体的承载能力下降，如何在这种情况下提升本征强度和化学强化强度成为一大挑战。如图 3 所示，较厚的玻璃在受到相同外力时，能更好地分散应力；而较薄的玻璃由于厚度小，应力集中更为明显，更容易破裂。

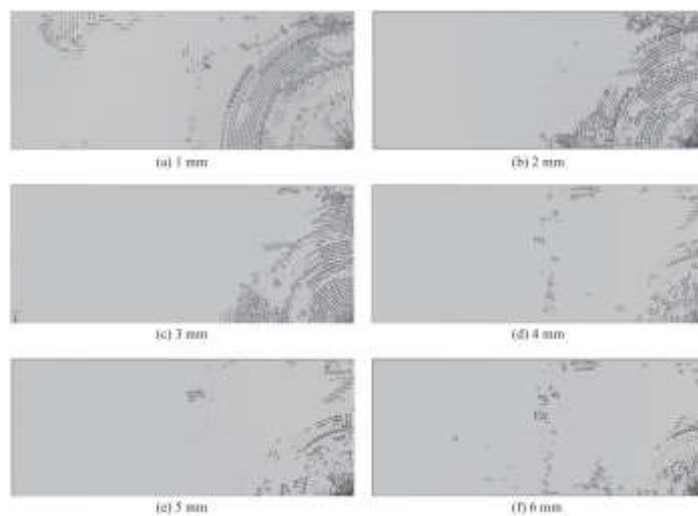


图 3：不同厚度玻璃受冲击破裂

2、玻璃网络结构优化带来的制造问题

优化玻璃网络结构，减少非桥氧以提升玻璃性能时，会导致玻璃粘度增加。在玻璃熔制过程中，粘度增加使得玻璃液中的气泡难以排出，同时容易产生条纹缺陷。气泡的存在会降低玻璃的透明度和强度，而条纹则会影响玻璃的光学均匀性，如图 4 所示，玻璃中的气泡（左图）和条纹（右图）会对光线传播产生干扰。

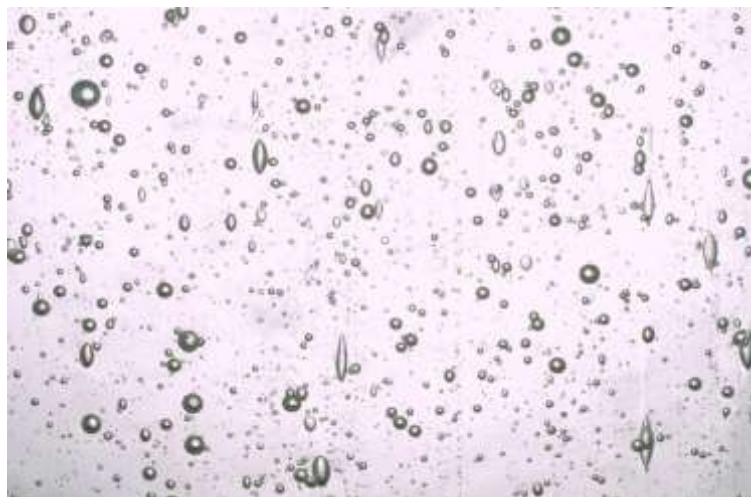


图 4：玻璃中的气泡和条纹

3、强度提高与加工性能的冲突

随着玻璃强度的提高，其硬度也相应增加，这给 CNC 加工带来了困难。在 CNC 加工过程中，过高的硬度容易导致刀具磨损加剧，同时在玻璃表面产生微裂纹和划伤等缺陷。这些微裂纹和划伤在后续使用过程中，可能会在外界应力作用下扩展，最终导致玻璃破裂，影响产品质量。

技术现状

在高强超薄盖板玻璃研发进程中，通过调整玻璃配方，减少非桥氧含量，增加了玻璃的本征强度，然而，玻璃配方调整致使玻璃熔体粘度上升，熔制过程中气泡排出效率降低，产品良率下降。

后加工环节同样面临挑战，高强度玻璃在 CNC 切割、研磨过程中，微裂纹、擦伤缺陷发生率提高，影响玻璃强度及加工良率。

技术诉求

任务 1：开发可量产的玻璃配方和 DOE 工艺

深入探究不同氧化物对玻璃性能的影响规律，通过理论计算和实验验证相结合的方法，设计出新型玻璃配方，使玻璃在满足高强度和良好化学强化性能的同时，具备适合大规模生产的工艺性能。

针对轻薄玻璃的化学强化深度和应力低的结果，进行系统的化学强化工艺设计，例如：使 0.55mm 的轻薄玻璃 DOL 深度达到 110 μ m 以上，CS₃₀ 达到 150MPa 以上，增强轻薄玻

璃的抗跌落性能。

任务 2：优化生产和加工性能

针对玻璃粘度增加导致的气泡和条纹问题，研究开发新的熔制和澄清技术，如采用超声波辅助澄清、电磁搅拌等技术，提高玻璃液的质量。目标是将气泡和条纹缺陷降低 50% 以上。

解决后加工引起的缺陷问题，研究表面处理技术，如化学抛光、等离子体处理等，改善玻璃表面微观结构，减少玻璃表面擦划伤，提升玻璃强度，同时具备较高的加工良率。

2.8 用于有机特种纤维的无热效应切割技术

技术背景

有机特种纤维（UPE/PBO/聚酰亚胺纤维等）凭借其高比强度、高比模量以及透波性好等优势（UPE 纤维模量达 150GPa，密度仅 0.97g/cm³，PBO 纤维强度可以达到 5.8GPa，密度仅为 1.54），在军工、航天复合材料等领域应用广泛。近几年随着消费电子产品轻量化需求越发突出，这些有机特种纤维也开始应用于智能手机产品。目前行业内常用的是采用红外激光（如 CO₂激光）通过光热转化实现切割，但热影响区（HAZ）可达 200-500μm [1]，导致边缘碳化。或者使用紫外激光（如 355nm），虽能缩小 HAZ 至 50μm [2]，但对曲面构件存在焦深限制问题。

技术挑战

热损伤控制：相比于碳纤维，UPE 等有机纤维的导热系数本身很低，激光切割比较容易产生热聚集。同时这些纤维热分解温度窗窄（如 PBO 纤维分解温度在 650℃，耐热温度不高于 300℃，UPE 纤维的熔点甚至只有 130℃左右），现有业内产业化技术难以避免化学键断裂引起的黄变甚至碳化，具体现象如下图所示。



a. UPE截面发黑 b. 切割边缘发黑异色 c. 切割边缘UPE卷边

三维切割瓶颈：常用的二氧化碳激光，光束聚焦系统对曲面适应差（ $\pm 5^\circ$ 倾斜角即导致能量密度下降 40%），制件厚度 $> 3\text{mm}$ 时出现锥度效应（切口锥角 $> 8^\circ$ ）。

技术现状

目前技术路线代表性成果的局限性：

- ① 水导激光：德国 IPG 公司实现 UPE 切割无黑化[3]，设备成本增加 300%，且有水渍污染问题；
- ② 飞秒激光：中科院实现 $5\mu\text{m}$ 热影响区[4]，加工效率仅 0.2m/min ；
- ③ 等离子体辅助：MIT 开发低温等离子体切割技术[5]适用于单层薄膜，初步评估对于 $0.1\text{mm}\sim 0.5\text{mm}$ 连续纤维复合材料制件可行性较低。

技术诉求

- ① 开发适用于高强度有机特种纤维的冷加工（低热效应）工艺（以 UPE 纤维为例，基材加工温度 $< 120^\circ\text{C}$ ）；
- ② 热效应区域 HAZ 小于 $20\mu\text{m}$ ，无明显熔融瘤状物，无明显黄变和碳化；
- ③ 兼容 $0.1\text{--}0.5\text{mm}$ 厚度的基材；
- ④ 满足 3D 结构件切割，尺寸参考：弧面切角 120° ，弧高 1.5mm ，R 角 0.6mm 。

参考文献

- [1] Wang L. Laser Processing of Polymers, Springer 2024.
- [2] US Patent US20240123456A1
- [3] IPG Technical Report Water Jet Guided Laser Cutting, 2024.
- [4] 亚微米级热影响区控制，飞秒激光加工技术, 2024.
- [5] Low-Temperature Plasma Cutting, MIT Plasma Lab, 2023.

2.9 持久耐脏和耐黄变的透明弹性体材料

技术背景

TPU 手机保护壳对手机的保护功能极佳，可以延缓手机使用寿命，深受用户认可，应用最为广泛。但是透明 TPU 保护壳在用户使用过程中有发黄、变脏污问题，又会影响用户的使用体验。

技术挑战

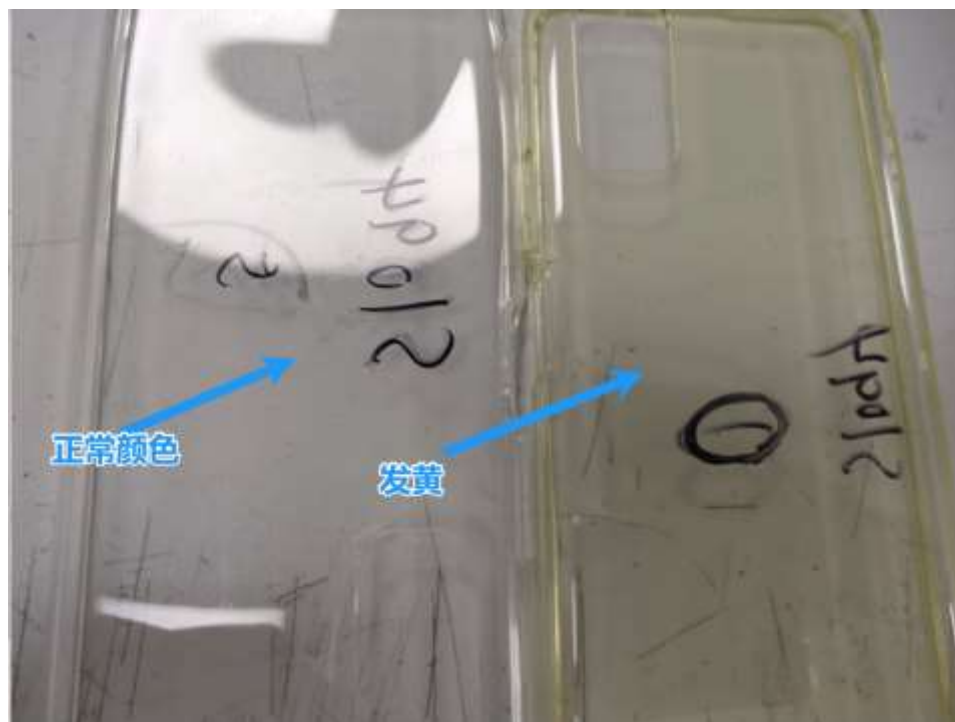
TPU 材料受到光照、温湿度环境老化，在用户使用过程中，接触到的手汗、化妆品等外界物质容易渗入，会造成保护壳黄变，发黑等问题。

1、聚酯型 TPU 材料在温湿度环境中会老化降解；

2、同时 TPU 材料为极性弹性体材质，常温下高于其 T_g 点，分子链没有完全冻结，造成其容易被浸染，外界物质也容易渗入，TPU 材料保护壳黄变、脏污是行业难题，急待攻克。

技术现状

1、高温高湿/UV 光照等测试发黄：



2、用户使用一段时间发黄发黑：



技术诉求

任务：寻找/研究开发抗黄变、耐脏污的高透弹性体材料

1. 高透弹性体材质，不限于 TPU 材料，邵氏硬度 95A 左右，可以注塑成型；
2. 材料稳定、对环境不敏感（光照、高温高湿环境、化妆品等）不容易老化，不黄变；
3. 对于外界物质有持续的低浸润能力，不易被浸染，（比如疏水、疏油能力）外界污渍能轻易的擦拭掉。

2.10 折叠屏叠层结构均质化本构建模研究

技术背景

随着电子设备日趋轻薄化、便携化，以及用户对更多互动体验的需求，折叠屏技术逐渐成为市场的焦点。折叠屏是通过多层材料的复杂叠加来实现的，如图 1 所示，其中包括硬质膜材如 PET（聚对苯二甲酸乙二醇酯）和 PI（聚酰亚胺），以及软体 OCA/PSA 胶。这些材料的独特特性使得折叠屏在使用中能够灵活转换形态，同时也带来了诸多挑战。

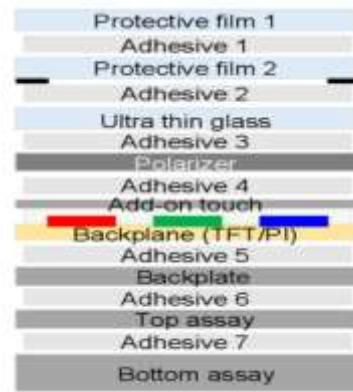


图 1 折叠屏叠层示意图

折叠屏的组装不仅是材料的简单叠加，还涉及到这些材料的复杂力学性能的相互耦合。硬质膜材 PET/PI 具有显著的非线性粘弹性特点，容易产生不可回复的变形，如图 2 所示 PET/PI 部分力学性能特点，且这些硬质膜材弹性模量跨度达 2~8GPa；OCA/PSA 胶力学性能则表现出强烈的蠕变和松弛等时间相关力学特性，如图 3 所示 OCA/PSA 胶部分力学性能特征，且其具有超低剪切模量(0.001~0.1MPa 量级)，高体积模量（1~10GPa 量级）特点，导致其极易形状变形，不易体积变形。可以明显发现折叠屏材料力学性能复杂且不同材料之间差异巨大，最终造成折叠屏力学响应具有方向各向异性、粘弹性及粘塑性等特征，这些特征导致折叠屏在静态载荷条件下出现如折痕和扭矩滞回环等力学现象，如图 4 所示，在高速跌落冲击条件下表现出明显的吸能或材料硬化现象。

为了准确评估折叠屏的耐久性和使用性能，目前的仿真分析多采取详细的本构建模表征，对所有层材料进行精确建模仿真。但此种方法导致模型网格数量庞大，计算效率较低，不利于缩短研发周期。

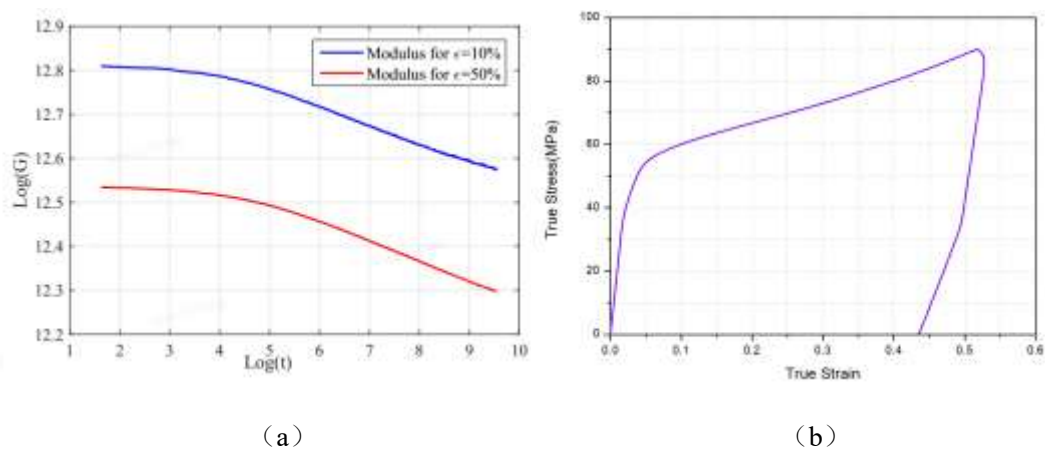


图 2 硬质膜材 PET/PI 力学性能：(a) 不同应变幅值下的动态测试模量；(b) 加卸载产生

的不可回复变形

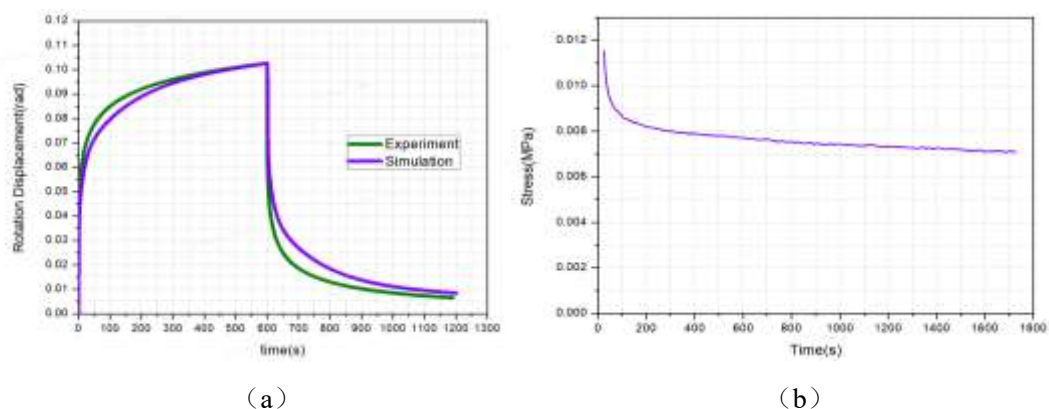


图 3 OCA/PSA 胶力学性能：(a)OCA 胶的蠕变；(b) OCA 胶的应力松弛

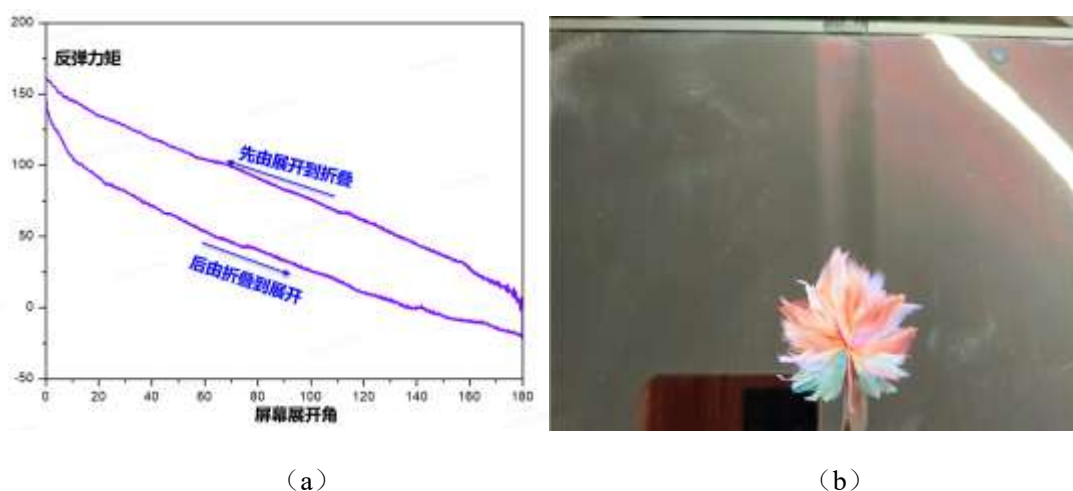


图 4 OCA/PSA 胶力学性能：(a)折叠屏开合扭矩滞回环；(b) 折叠屏折痕效果

技术挑战

在折叠屏的仿真分析中，最大的挑战莫过于如何在保证准确性和合理性前提下提高计算效率。传统的仿真方法需要对每一层膜材及胶材进行独立建模，并且需要分别划分细致的网格以反映其复杂的力学行为，然而这需要巨大的计算资源以及较长的仿真分析时间。

相较于传统方法，均质化建模方案希望能够将所有叠层视作一个整体进行统一仿真，从而减少仿真模型的复杂度，提高计算效率。此种方案的最大难点在于如何在均质化过程中反应屏幕整体响应的各向异性、以及长时和瞬时条件下粘弹性、粘塑性等复杂力学特征，这些都是影响折叠屏性能的关键因素。

均质化建模的过程中需要准确地将不同材料属性综合到一个统一材料体系内，确保仿真结果能够真实反映实际使用情况。考虑到材料的多尺度、跨时间响应行为，传统数值方法可能无法满足精确度要求。因此，对于均质化模型定义、数值算法优化以及结果验证都是非常

具有挑战性的技术问题，需要深入的理论探讨和实验支持。

技术现状

目前，在折叠屏材料本构建模方面，我们通过各种实验测试确定了 PET/PI 硬质膜材和OCA/PSA 胶的基本力学参数，可通过折叠屏叠层结构的详细建模方法，对折叠屏在静态/准静态长时工况下的性能指标做出仿真评估，也可对跌落冲击瞬时工况条件进行仿真分析。然而由于材料的复杂性和使用场景的多样性，这些参数在多层叠加使用中仍面临较大的挑战，例如仿真分析计算效率低,所需要的计算资源巨大，无法考虑折叠态的预应力等问题。

为了提升折叠屏的仿真计算效率，对折叠屏均质化建模方面也进行了尝试，但这些方法对于表征折叠屏的复杂力学性能仍然存在各种问题，仍无法满足项目的实际应用需求。

技术诉求

基于当前研究所面临的技术挑战，我们希望通过与高校合作，集合更多的研究资源和智慧，突破折叠屏均质化建模的瓶颈。我们的具体技术诉求如下：

理论研究：深入研究折叠屏材料的力学性能特点，探索均质化模型建立的理论基础。特别是针对材料的粘弹性、粘塑性及各向异性特征，提出适合的均质化方案。均质化本构模型应当支持跨时间尺度的仿真，既要能表征折叠屏在长时载荷下的响应，又要能表征折叠屏的大变形瞬时力学响应，最终形成一套完整且实用的各向异性粘弹粘塑性本构模型理论。

实验验证：建立相应的实验平台，对均质化模型理论进行验证，确保理论计算预测结果与试验结果相符。通过实际实验数据支持，优化模型参数，形成完整的理论模型预测和试验校正的规范和流程。

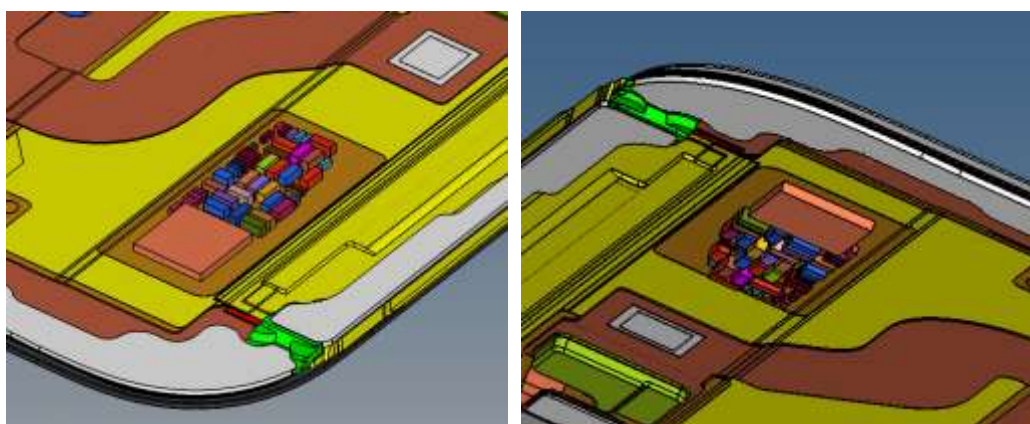
本构模型数值化：引入先进的数值计算方法，基于 ABAQUS 仿真平台将所研究的跨时间尺度各向异性粘弹粘塑性本构模型推广至三维复杂应力状态，并将其数值化为 VUMAT/UMAT 代码，实现数值化应用。

实际项目仿真验证：在折叠屏叠层整体建模少网格数量条件下，实现折叠屏折痕，扭矩滞回环等长时工况以及跌落、钢球冲击等瞬时工况的准确及高效仿真分析。

2.11 温度驱动的湿气浓度扩散方程求解技术及其在 ABAQUS 实现

技术背景

智能手机在日常使用中不可避免地会暴露在高温高湿的严苛环境中。无论是用户长时间运行大型游戏导致的机身发热，还是户外工作者在沿海地区等高湿度场景下的使用，抑或是将手机遗忘在密闭汽车内形成的"蒸笼效应"，都会使内部温度突破 50°C 并伴随 90% 以上的相对湿度。这种双重环境压力将导致精密电子器件的加速劣化：金属触点受盐雾腐蚀引发接触不良，PCB 基板吸潮膨胀造成线路微短路，锂电池电解液分解加速导致鼓包风险，甚至连屏幕 OCA 光学胶也会因湿热老化出现脱胶现象。



技术挑战

整机及其器件需要放置于变化的温度场中进行仿真分析，与此同时，该结构还需要考虑周围湿气的作用。湿气具有独立的自由度，且该变量的变化遵循扩散方程，仿真分析求解湿气影响时，需要中考虑温度边界的前提下，求解其扩散方程。

技术现状

当前仿真软件中，尚无法做到温度场与湿气扩散方程耦合求解，需要开发新的单元及其求解技术。

技术诉求

任务：基于温度场的湿气浓度扩散方程数值求解并在 ABAQUS 用隐式与显式相场法（UEL 与 VUEL）实现

1. 根据 vivo 描述的背景信息，建立受温度场影响的湿气浓度扩散方程，进一步推导其有限元格式；

2. 根据相场法原理，编写、调试可求解湿气浓度扩散方程的用户单元子程序，包含 UEL 与 VUEL，并对子程序进行验证，要求与实验数据一致。

2.12 固态电池应用特性研究

技术背景

固态电解质对比液态电解液的优势：固态电解质可以有效降低传统电池起火爆炸风险，且改善低温性能和电池健康度。

传统液态锂离子电池自 1991 年商业化以来，凭借其高能量密度、长循环寿命以及快速充放电能力，已成为便携式电子设备、电动汽车以及储能系统的核心能源载体。然而传统锂离子电池由于电解液的高反应活性，常面临安全挑战，如图 1 所示。为解决当前安全问题，通过使用固态电解质代替液态电解液降低其反应放热程度，可有效降低锂离子电池起火爆炸等风险。



图 1：近年锂离子电池起火导致事故部分汇总

固态电解质技术路线：聚合物固态电解质目前最适合消费类电池

目前固态电解质主要分为氧化物固态电解质、硫化物固态电解质卤化物固态电解质和聚合物固态电解质四大类。以下进行详细介绍：

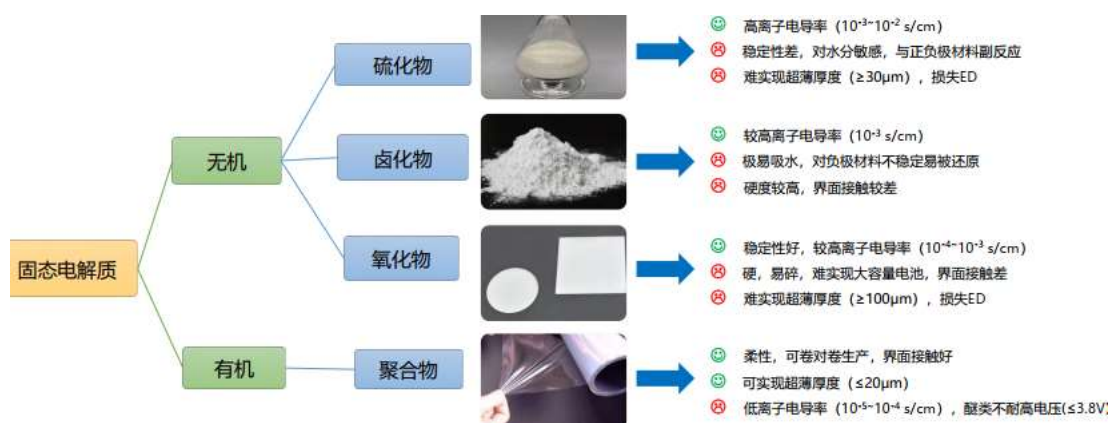


图 2：固态电解质分类及优缺点

氧化物固态电解质具有较高的锂离子传导效率，较好的高电压耐受窗口，曾经受到广泛关注，如早期清陶、卫蓝等均采用此路线。但该路线存在较大科学问题暂无法解决，即无机氧化物固态电解质陶瓷硬度较高，只能与钴酸锂等活性物质形成点接触（液态电解液为面接触），接触面积小导致锂离子传输通道不畅，阻抗大。目前科学界仍无有效方法解决。

硫化物固态电解质相比氧化物固态电解质锂离子传导效率最高，且硫化物质地柔软，容易与活性材料如钴酸锂形成更多的接触面，避免接触面积过小而导致的极化问题产生，阻抗小。但硫化物固态电解质同样存在明显的劣势，即硫化物自身热安全性能较差，无法根本有效解决传统锂离子电池起火爆炸的问题，且高电压窗口兼容性不足

卤化物固态电解质离子电导率、硬度、高电压兼容性等均介于硫化物和氧化物之间，并未解决根本问题。

无论氧化物、卤化物还是硫化物，均需要大幅改变常规锂离子电池生产方式，前期设备投入较大。

早期聚合物基固态电解质相比氧化物和硫化物固态电解质，虽然锂离子电导率最低，且不具备高电压耐受性，但由于其兼容常规锂离子电池生产产线，经推出后现已小规模量产，展现出了其应用潜力。近来，高分子设计的进步使聚合物基固态电解质得到快速发展，学术界已出现通过分子设计指导克服聚合物基固态电解质固有缺点的技术手段，故未来聚合物基固态电解质在锂离子电导率、硬度、电压耐受性和生产设备兼容性方面展示出极大的潜力

技术挑战

提升锂离子电导率和高电压耐受性是关键。

1、提升聚合物固态电解质锂离子电导率：

锂离子电导率的提升是通过改善锂盐的解离程度来实现的。聚合物单体中特定官能团通过与锂盐阴离子的强相互作用，如氢键、配位键等，使锂盐解离平衡向溶解方向移动，释放出更多自由锂离子，提高了锂离子在电解质中的传输效率。

2、提升聚合物固态电解质的高电压耐受性：

固态电解质的高电压耐受特性是通过改善聚合物单体的电子轨道能级来实现的。通过在聚合物单体中加入部分官能团，如-CF₃ 三氟甲基的强吸电子诱导效应来稳定电子防止其被正极侧夺走，继而改善聚合物固态电解质在正极侧的稳定性。

技术诉求（任务一必须完成，任务二和三挑选其一完成）

任务一： 筛选聚合物最优聚合物高分子结构，建立构效关系：

1. 强吸电子基团改性，分子链支链化改性等，与高电压窗口之间关系
2. 空间位阻基团改性，弱溶剂化基团改性，与快充倍率之间的关系
3. 聚合物电解质分子量与聚合物电解质柔性之间的关系
4. 整合后，聚合物电解质与长循环性能之间的联系，并明确与液态电解液相比失效机理差异

任务二： 基于聚合物电解质制备 demo 电池。

电池 demo 制备要求：

1. 载体：8Ah；化学体系：正极 4.55V LCO（钴酸锂）+负极 25%沉积硅碳; EDtyp（能量密度典型值）≥950Wh/L
2. 充电功率>90W 且首段大倍率充电时间>2min
3. -10℃低温放电 3.0V vs 常温低温放电>76%
4. 循环寿命最低要求：以 90W 功率对应电流充电且初始首步电流>2min 条件的初始条件+vivo 指定循环中抬压降压降流策略条件下，常温循环 1000cls 应保证容量保持率≥75%，厚度膨胀应不高于初始半电态厚度的 13.5%，高温循环 700cls 应保证容量保持率≥75%，厚度膨胀应不高于初始半电态厚度的 13.5%。（初始半电态厚度为 3.97V 条件下电芯厚度）

任务三： 聚合物电解质固态电池的应用特性研究以及和电解液电池的对比研究。

载体：Base 组：任务一中 demo 电池；对照组：电解质改为液态电解液；其余均完全相

同，要求正极使用 4.55V LCO+负极 25%沉积硅碳，电池 8Ah。

要求:

1. 体积能量密度对比研究，要求体积能量密度 Base 组不低于对照组
2. 充电析锂窗口对比研究，要求 Base 组对比对照组 Loss<0.3C
3. -10°C低温放电 3.0V 容量保持率对比研究，要求低温容量保持率不低于对照组
4. 循环寿命对比研究。以 90W 功率对应电流充电且初始首步电流>2min 条件的初始条件以及 vivo 指定循环制式中抬压降压降流策略条件下 1000cls 的容量保持率及厚度膨胀对比，要求 Base 对比对照组需提升 100cls 以上。
5. 电化学阻抗谱（EIS）以及弛豫时间分布（DRT）随着电池衰减的变化对比（vivo 指定循环中抬压降压降流策略条件下 1000cls）。
6. 安全性能对比。包含重物冲击、针刺、刀剪、200°C热箱 60min 等失效过程对比。

三、 光学、颜色及器件

3.1 基于人眼感知的覆盖真实场景的 HDR 影像与显示评价解决方案

技术背景

高动态（HDR: high dynamic range）影像与显示技术在手机行业已经变得越来越流行，相较于标准动态（SDR: standard dynamic range）技术，HDR 技术具有更好的亮度动态范围、更大的色域、更高的位宽的特点，如图 1-（a）。这些技术特点可以极大地还原真实场景并且在此基础上进行更大余地的美的渲染，大幅提升用户体验，如图 1-（b）。



图 1 (a) SDR 与 HDR 参数比较；(b) SDR 与 HDR 效果比较示例。

目前在影像端以及显示端都已经开发了多种方法进行 HDR 效果渲染，特别是在显示端，已经形成了相关标准，例如 HDR10+、dolby vision 等。但是目前缺乏基于人眼感知的或可以真实场景的 HDR 效果评价方法。这对于 HDR 渲染算法的优化造成了很大困难，业界没有明确的目标实现更加真实或者更加符合审美的 HDR 效果。

技术挑战

- 1) 人眼感知的复杂性：HDR 人因实验数据较少，缺乏建模所需必要数据。
- 2) 真实场景多样性：真实 HDR 场景种类多样，例如夜景、日光场景，难以全面准确评估。
- 3) 实验室仿真场景难以复现：HDR 场景所需的高动态、大色域环境较难在实验室复现。

当前结果

目前已经有一些团队在开展 HDR 评价相关工作，在影像端，浙江大学团队研究了颜色

空间选择以及色彩渲染路径、指标对于 HDR 效果的影响，如图 2-（a）。但是该类研究的影像处理手法与手机行业存在一定差异，本质上是评价颜色科学领域的色貌模型的效果，与实际工程需求存在一定偏差，方法泛化性不足。在显示端，dolby 团队研究了 HDR 动态对于用户喜好性感知的的影响，如图 2-（b）。但是该研究涵盖的场景较少，且没有进一步研究 tone/gamut mapping 对于显示效果的影响，此外该实验设置为大屏，而非手机小屏，应用于手机领域的结论还有待验证。国际电信联盟(ITU)也推出了全参考 HDR 色差计算模型，如图 2-（c）。但是该研究针对颜色空间的色彩，对于图片的整体色差无法做出类似人眼的权衡判断。综上，他们均没有完整的解决我们的核心技术诉求。

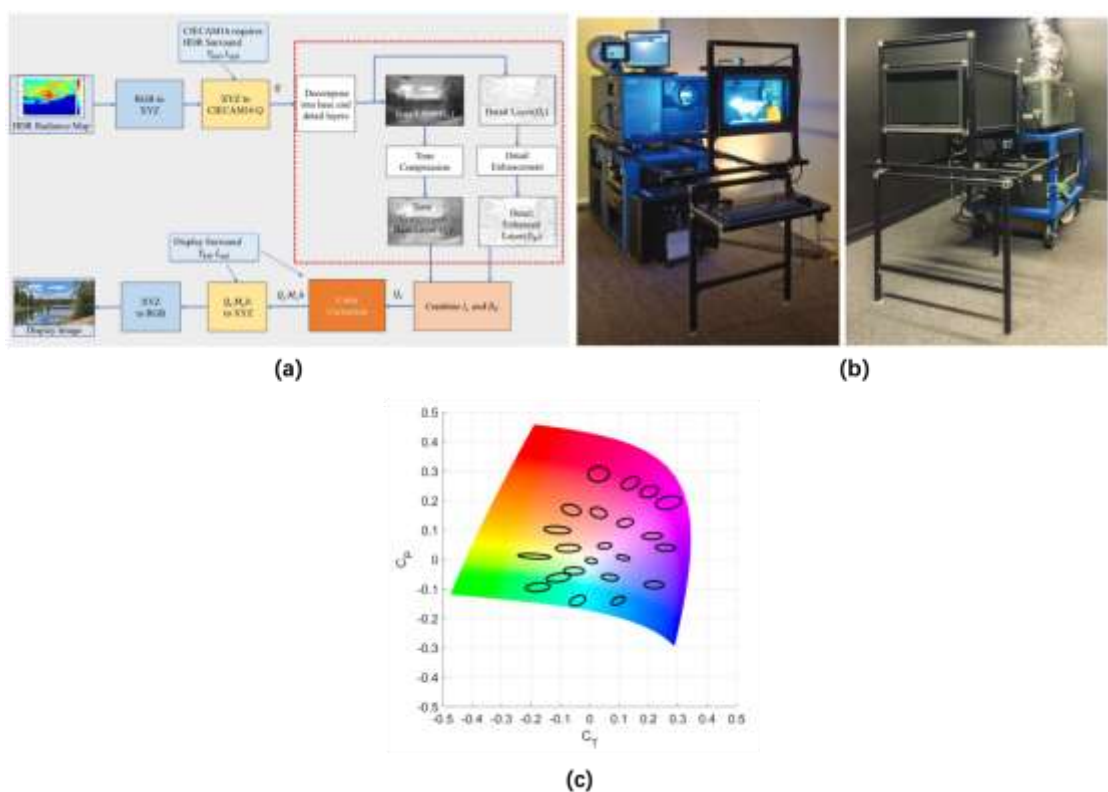


图 2 (a) 浙大团队 HDR tone/gamut 映射算法流程；(b) Dolby 团队 HDR 显示研究设备；(c) ICtCp 色彩空间恰可察觉色差示例。

技术诉求

核心目标：基于人眼感知的覆盖真实场景的 HDR 影像与显示评价解决方案：

1) 基于人因实验，研究人眼的 HDR 影像与显示感知模型；

a) 例如基于 HDR 场景、显示环境，基于心理物理学实验，获取真实度、舒适度、喜好度评价数据，并建立人因模型；

- 2) 搭建实验室仿真场景（或基准 HDR 显示系统）并定义实景用于 HDR 评价；
 - a) 例如采用 LED 灯箱模拟高动态广色域实景；
- 3) 建立 HDR 影像与显示评价模型，输出真实度、喜好度等指标或者建立基准参考数据集，并且开发全参考差异度量方法。
 - a) 可以基于人因模型输出无参考评价模型，也可基于参考场景、样张、显示方案给出全参考评价模型，判断与基准的差异。

参考文献

Mehmood I, Shi X, Khan M U, et al. Perceptual Tone Mapping Model for High Dynamic Range Imaging[J]. IEEE Access, 2023, 11: 110272-110288.

Daly S, Kunkel T, Sun X, et al. Viewer preferences for shadow, diffuse, specular, and emissive luminance limits of high dynamic range displays, SID Symposium Digest of Technical Papers (ISSN: 2168-0159) 44 (1)(2013) 563-566[EB/OL].

https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2124-0-201901-1!!PDF-E.pdf

https://professional.dolby.com/siteassets/pdfs/ictcp_dolbywhitepaper_v071.pdf

3.2 润滑条件下的摩擦系统模型建立与仿真

技术背景

手机摄像的对焦及防抖需要驱动马达实现，其中接触摩擦形态的摄像头马达是其中的主要实现方式，随着摄像头产品的精度要求的提升，大镜头负载，长行程的需求日益增加，对于现有的摩擦形态马达的接触形态需要进行准确预估，同时在摩擦系统下的润滑特性也需要融合考量到接触模型中，才能更加准确的将系统模型建立，输出最终正确的设计参数，为马达性能设计，调试优化提供依据。

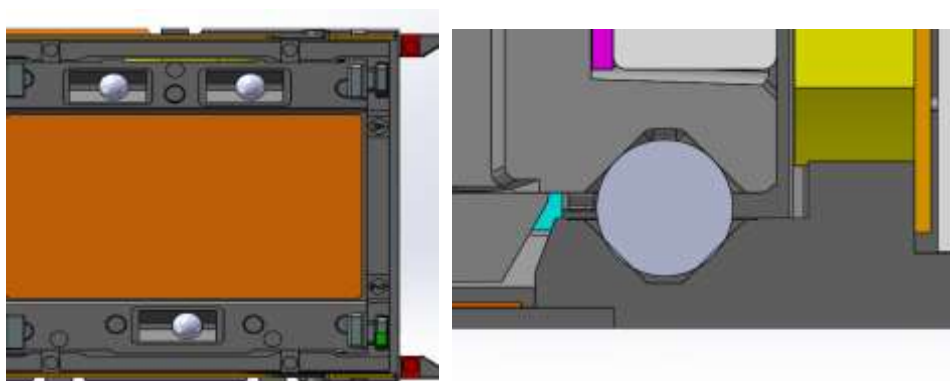


图 1 滚珠滚动摩擦结构及截面形态（图中三个滚珠）

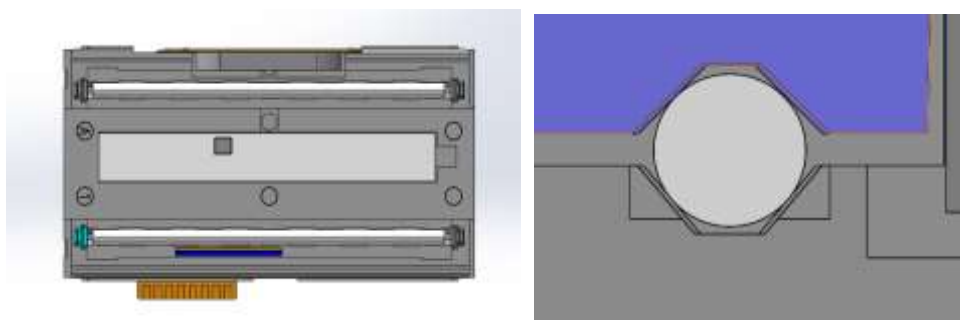


图2 导杆滑动摩擦结构截面形态（图中两根导杆）

技术挑战

由于该摩擦系统中的材料种类较多，有塑胶与陶瓷接触，塑胶与金属接触，塑胶与塑胶接触，陶瓷与金属接触等接触方式，接触面材料由于成型加工存在表面粗糙度，同时要增加接触面的润滑特性，在接触面还需要涂布润滑油脂。由于手机在使用过程中存在摔落等冲击环境，会对塑胶或金属表面造成一定深度凹坑，导致接触面形态发生变更，造成局部位置摩擦系数变更，影响最终驱动效果，甚至造成产品失效。，因此如何通过仿真模型的建立与仿真实现摩擦系统摩擦系数的准确预估是一个重大突破。

当前结果

当前的接触摩擦形态马达的摩擦系数主要通过过往案例中的实测数据得知，但是实际测试无法准确得知产品整个运行状态每个位置的摩擦数据，由于润滑油存在一定粘滞特性，也无法准确得出实际摩擦力大小，因此现有情况下的摩擦系数均为经验数值，不同材料与不同接触结构间的摩擦系数无法准确给出。

技术诉求

基础任务 1: 建立多种材料基本接触条件模型，搭建理论测试模型，结合实际测试数据，输出一段运动距离的摩擦接触的摩擦系数；

基础任务 2: 建立不同接触面形态，考量微型凹坑变化对实际摩擦形态造成的影响；

进阶任务 1: 结合不同润滑油的润滑及粘滞特性，构建仿真模型，输出一段运动距离的摩擦接触的摩擦系数；

3.3 高精度宽频谱防抖控制技术

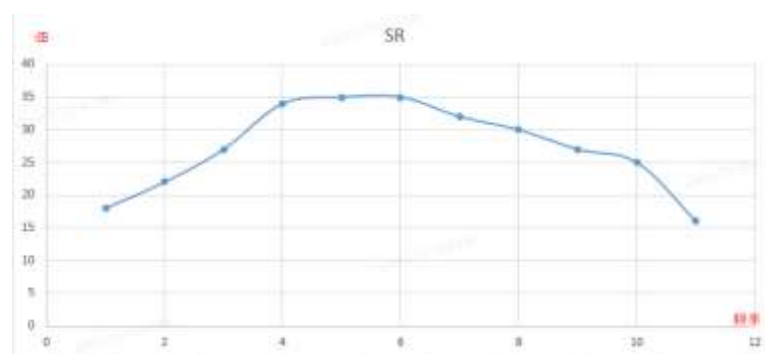
技术背景

光学防抖（Optical Image Stabilization, OIS）是相机和摄像头中一种有效减少手抖或其他震动引起的图像模糊的技术。其原理是通过陀螺仪侦测微小的手抖或其他震动，信号传至微控制器，控制器（MCU）计算需要补偿的位移量，移动镜头或传感器来补偿，从而有效地减少外界抖动引起的图像模糊，拍摄出清晰的图像。



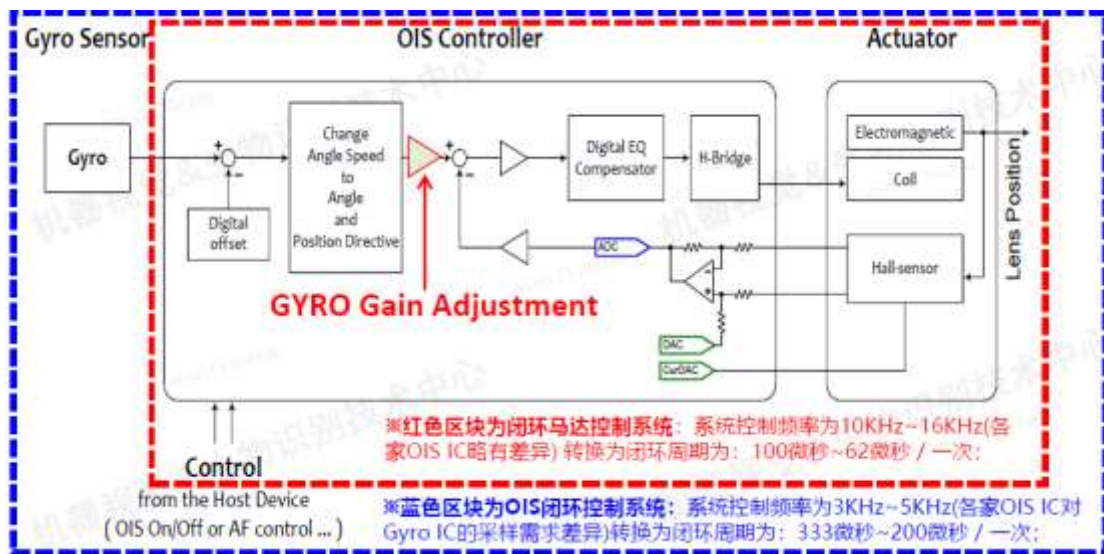
技术挑战

在用户更加聚焦影像效果的背景下，手机摄像头等对 OIS 防抖有了更高的需求，尤其是长焦摄像头，在 Zoom 下，控制精度的误差也会被成倍的放大。用户手持手机拍照抖动频率范围 4~6Hz，当前整机 OIS 系统方案针对用户手持场景，在 4Hz~8Hz 频率范围内防抖效果比较好；但是随着更多拍照，录像场景的需求，抖动的频率和幅度有了更多需求；因此，更高精度（微小的抖动），更宽范围频率和更大幅度防抖的控制技术，是重要的技术挑战。



当前结果

当前整机 OIS 系统框架拆解如下：在整机相机系统方案中，结合电子防抖（Electronic Image Stabilization, EIS）软件算法对图像进行后处理，进一步减少图像抖动，但仅在中低频效果较佳。



	用户使用场景	当前状态
低频震动	手持	有效减少画面模糊和跳动，但是潜望 zoom 下还需要提升
中频震动	步行，骑车，乘车等	当前 OIS 系统补偿效果较差，结合 EIS 补偿有一定效果
高频震动	快速振动，瞬时冲击	当前 OIS 系统不能有效补偿，结合 EIS 补偿也效果不佳

技术诉求

主要技术诉求包括二方面：

1) 精度更高的控制系统：在现有马达结构设计上，改良硬件结构设计，提升控制精度，目标做到全行程区域内闭环反馈精度 $<0.5\mu\text{m}$ ；并提升马达的 Plant FRA 一致性，目标做到运动状态下跟随 Blur 量 $<3\mu\text{m}$ 。

2) 优秀的闭环控制算法研究：在现有的马达运动特性基础上，进行闭环控制算法优化，提升闭环响应速度。目标闭环控制算法同时满足防抖（连续跟踪目标）与阶跃响应需求。

a. 阶跃响应性能：小步进： $5\text{ms}@10\mu\text{m}$ /大步进： $5\text{ms}@500\mu\text{m}$ ，Tolerance 标准均为 $\pm 2\mu\text{m}$ ）

b. 连续跟踪性能：实现频率在 $1\sim 60\text{Hz}$ 下进行目标跟踪，Gain 差异 $<1\text{dB}$

通过上述软硬结合方案实现多场景下更好的防抖效果。

3.4 EVS 信号分析及应用专题研究

技术背景

基于事件的相机是受仿生技术启发的异步传感器，其关键优势包括微秒级时间分辨率、低延迟、高动态范围和低功耗。正因如此，基于事件的相机开辟了传统（基于帧的）相机无法想象的领域，而传统相机在过去 60 年中一直是主流传感技术。EVS 传感器高达 1000 samples/sec 的时间分辨率，使得其能捕捉时间轴上场景内的精细亮度变化，近年来关于 RGB 与 EVS 融合的相关研究及应用方案趋势显示，EVS 有着重要的应用价值与远大前景，EVS 辅助能改善许多传统 RGB 传感器无法很好处理的场景并有效提升用户体验。

技术挑战

EVS 传感器与 RGB 传感器的信号融合，非常仰赖信号层面上时间空间的对齐同步，因此，关于 EVS 信号的各种特性，如延迟性，事件阈值，事件噪声，信号频宽等，皆会影响与 RGB 融合计算精度及应用的最终效果。目前对于 EVS 信号特性的准确度与信号品质测量技术仍在发展的阶段尚未成熟，并且 EVS 在行业上的制造商以及 sensor 的规格各家发展皆有差异，并无统一标准。因此 EVS 的信号分析及品质上的评估，与信号融合后精度，是 EVS 目前于商业应用的重要挑战。

此外，EVS 辅助可以大大改善传统 RGB 传感器影像品质无法很好处理的场景，如影像去模糊，影像超分，视频插帧、计算摄影等，但是在基于事件的表示、事件信号处理、时间同步控制以及与 RGB 信号的交互融合算法研究等方面，仍有大量的工作和挑战。

当前结果

目前 EVS 传感器引起了越来越多的业界和学术界的关注，并推动了该领域令人振奋的进展，学术科研上论文数量或相关研究都有逐年增加的趋势，但在消费类电子市场上，实际有效的 EVS 传感器商用尚未成功落地，技术复杂度和场景普适性等方面仍存在技术难点需要攻克。涵盖传感硬件以及充分利用新型事件相机所需的处理、数据和学习方法方面，仍存在巨大的技术挑战。

技术诉求

任务：EVS 信号分析及应用研究

1. 在影像去模糊、视频插帧等典型应用场景，EVS 与 RGB 数据信号融合增强的算法研究

包括融合算法模型、网络架构设计、合成或真实场景数据集建立、损失及评估函数定义等；

2. EVS 事件信号特性分析、事件信号处理及事件仿真算法方面的研究

包括对传感器硬件层面的信号延迟性，事件阈值，事件噪声，信号频宽等特性及其对后期应用的影响进行研究，以及事件信号的处理，如时间同步、事件噪声去除、事件信号表达等方面的研究；

参考文献

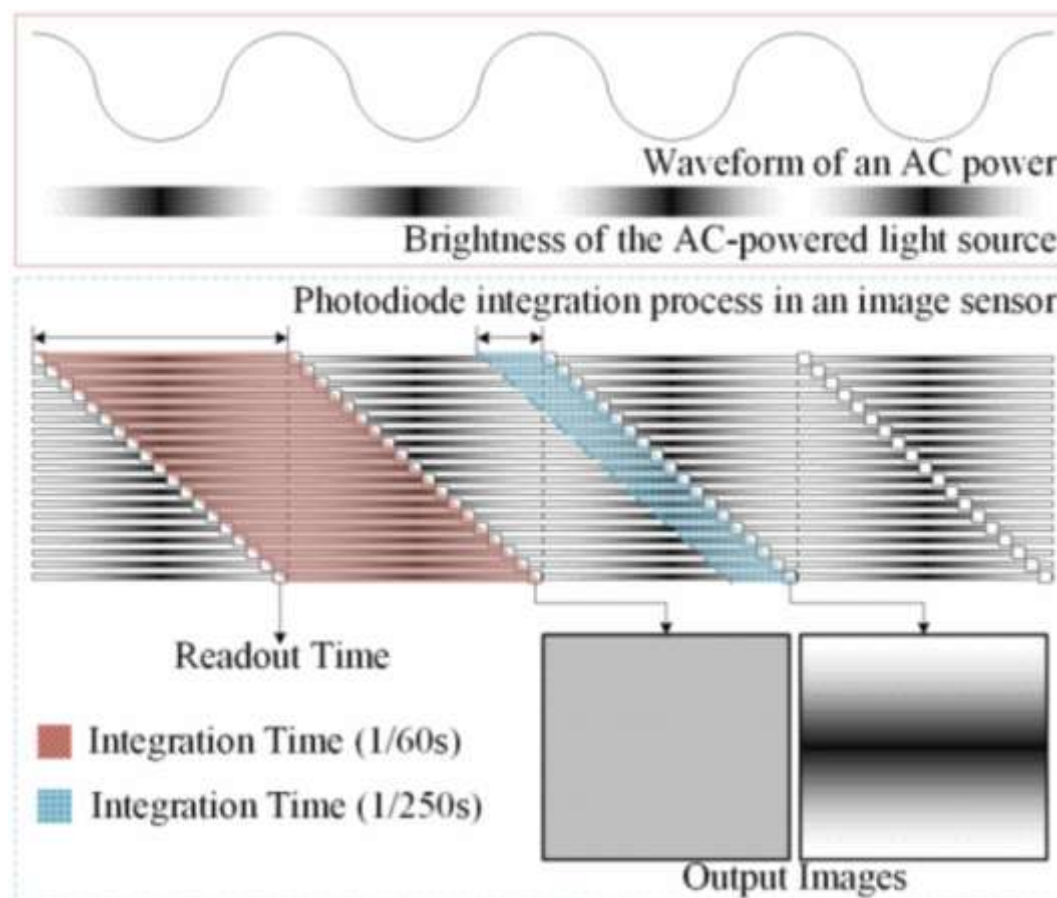
<https://www.sony-semicon.com/en/technology/industry/evs.html>

<https://www.prophesee.ai/event-based-vision-applications/>

3.5 基于图像传感器的任意频率 flicker 实时检测

技术背景

Flicker（闪烁）现象主要是由于光源的频率与拍摄时间不匹配导致在影像画面上出现明暗条纹或者闪烁等现象，对于成像以及录像的画面造成比较大的干扰。





移动影像逐步发展之后，出现了更多容易出现 flicker 的场景，如电子屏幕（常见电脑屏幕、户外大屏幕、舞台荧幕等等）。

技术挑战

因为电子屏幕的指标不尽相同，对于高精度、多频率的检测要求较高；

同时，在实际场景中，屏幕往往只占据部分画面，因此对于区域性的检测有要求；

现在常规的单独 flicker 检测传感器在上述场景，无法较好地满足需求。

如何实现高精度检测、同时针对实际可能的不同区域进行区域检测是当前重要的技术挑战课题。

当前结果

当前为单独 flicker 检测传感器，检测范围在 70~500Hz，检测精度 2Hz。

提升检测范围、检测精度（特别混合光源下），支持区域性检测，能够与常规的图像传感器做整合是需要研究探索的课题。

技术诉求

任务：基于相机图像传感器检测 200Hz 内任意频率的 flicker，包括照明光源，电脑荧幕，户外大屏幕，舞台大屏幕应用场景。

1. 检测范围：<200Hz

2. 检测精度：1Hz

3. 分区检测精度：12x9 block size
4. 感度/灵敏度：满足 DC+AC 弱交流 0.1%
5. 实时检测速率：至少满足 120fps

3.6 基于 XR 的色彩模型及显示效果调试

技术背景

随着 XR（扩展现实，Extended Reality）技术在游戏娱乐、工业仿真、医学培训等领域的深入发展，用户对视觉质量、沉浸感和真实感的要求不断提高。XR 系统作为一种高度集成的视觉交互平台，其显示效果不仅依赖于传统的显示面板技术，还受到复杂光学路径、观看条件和人眼感知机制的共同影响。因此，如何实现色彩还原准确、显示效果自然、感知体验统一的 XR 图像输出，成为当前研究的关键问题之一。

技术挑战

1、与手机、平板、电视等传统 2D 显示终端相比，XR 设备具有显著不同的观看条件和光学结构。一方面，用户需佩戴头戴式设备（如 VR 头显或 AR 眼镜）进入沉浸式体验，这种封闭式设计在物理上隔绝了环境光，形成了几乎全黑的观看环境。在这种低环境光甚至全暗条件下，用户对画面亮度、色彩饱和度与对比度的敏感性显著提高，任何显示参数的微小偏差都可能被感知放大，从而影响整体的沉浸感。

2、另一方面，XR 设备普遍采用大视场角（FOV）设计，主流设备可达到 80°至 120°，远超传统显示终端的 30°左右。这种广角视域虽然增强了空间沉浸感，但也对图像的亮度均匀性、色彩一致性以及边缘显示质量提出更高要求。传统色彩模型如 sRGB、CIELAB 等，多数是在标准观察条件（2°视角、D65 光源、平面显示）下构建的，其假设前提并不适用于大视场角与暗室条件下的 XR 应用。因此，需要针对 XR 特有的观看条件构建新的视觉感知模型，以支持色彩准确传递与感知一致性优化

3、此外，XR 显示系统往往采用透镜光学成像系统将图像投射至人眼，而这种光学路径在现实中不可避免地引入畸变、模糊、亮度衰减及色散等问题。这些问题直接影响用户对色彩清晰度、饱和度与对比度的主观感知。例如，图像边缘区域因清晰度下降，可能出现色彩饱和度感知减弱的现象，从而造成显示不一致性。因此，仅依赖物理显示参数调节已难以满

足高质量视觉输出的需求，亟需引入感知驱动的色彩补偿模型，实现光学成像前后主观色彩体验的一致性。



当前结果

直接使用传统的色彩理论进行显示色彩调试。

技术诉求

1. 适用于暗环境与大视场角的色彩感知模型
2. 建立适配 XR 显示的最佳亮度映射曲线与白点模型
3. 针对光学成像导致的图像色彩质量衰减的主观色彩补偿机制

四、 网络通信

4.1 低功耗高性能 Wi-Fi 或 无线传输模式方式

技术背景

在智能终端品类中，多设备互联、传输的需求比较大，使用蓝牙传输有很大的带宽瓶颈，使用 Wi-Fi 的诉求变得突出。

使用 Wi-Fi 的场景中，主要包括低速率视频推流、高速率传输文件等；两者对 Wi-Fi 都是强诉求；对于可移动设备，电池容量有限，导致产品的续航表现就会受影响。对于用户来说，提升 Wi-Fi 的功耗表现，就是提升用户对续航的直观体验。

技术挑战

1. 如何在低速率传输中，通过睡眠等策略的优化，既能满足数据传输一定的实时性；又能尽可能地满足低功耗的表现；
2. 在高速率传输文件的场景中，识别出功耗的主要消耗点，并提供策略有针对性地对该场景进行功耗极限优化。

技术现状

启用高速率高带宽模式去传输文件时，Wi-Fi 芯片整体的功耗非常大，对系统续航的影响大，即便是低速率传输场景中，由于 Wi-Fi 射频工作的时间持续，导致功耗也不能做好很好的降低。

作为参考，限定 Wi-Fi 芯片规格，1x1 Wi-Fi6 80MHz 频宽，PHY 带宽上限 600Mbps 这种 Wi-Fi 规格：

1. 在直播推流场景下，使用低速率（10Mbps）推送视频流，Wi-Fi 整体功耗大约在 50mA
2. 在大文件告诉传输场景下，使用 200Mbps 传输文件，整体 Wi-Fi 部分的功耗大概在 250mA

技术诉求

任务：Wi-Fi 芯片规格，1x1 Wi-Fi6 80MHz 频宽，PHY 带宽上限 600Mbps；对高带宽 Wi-Fi 的功耗做优化，满足产品极致续航的需求；具体诉求如下：

1. 在直播推流场景下，使用低速率（10Mbps）推送视频流，Wi-Fi 整体功耗控制在 30mA 以内；对比传统同等 Wi-Fi 规格的 50mA，功耗表现提升 40%以上；

2. 在大文件告诉传输场景下，使用 200Mbps 传输文件，整体 Wi-Fi 部分的功耗控制在 150mA 以内；对比传统同等 Wi-Fi 规格的 250mA，功耗表现提升 40%以上。

4.2 城市高楼环境下的精准定位

技术背景

在繁华都市的核心区域，高楼大厦鳞次栉比，街道如同狭窄的峡谷。在城市高楼下使用定位功能，不同用户的使用需求和行为各不相同，比如：

- ① 行车导航、步行导航、骑行导航
- ② 打车、分享实时位置、用户之间找人
- ③ 找特定的地点（商场入口、外卖导航等）

用户在这类环境下的定位和导航的需求越来越频繁，并对定位的精度和实时性的要求越来越高。

现有技术主要是通过 GNSS 卫星定位来获取位置的。在城市高楼这类复杂多样环境下，一方面，当 GNSS 信号被建筑物外墙大量反射，接收机接收到众多反射后的多径信号，另外一方面，部分信号被建筑物完全阻挡，形成信号“死角”，最终导致定位误差波动大，定位结果容易出现漂移或延迟更新。

在此基础上叠加 RTK（实时差分定位）等方式来提升定位精度的方案，在空旷道路场景下有较好的效果，但是在城市高楼下效果无明显改善。

技术挑战

1、因为手机接收到的卫星信号受到折射和反射等的影响，导致常见的定位算法(包括 L1 和 L5 双频的使用)在城市峡谷环境下恶劣的区域无法计算出高精度位置(<5m)；

2、使用周边基站、Wi-Fi 等网络定位的方式，严重依赖于周边基站、Wi-Fi 的覆盖情况，定位精度无法得到保证；

3、手机的天线性能、传感器精度受限（无法达到一个很高的标准），并且用户使用手机的行为多样，包括使用车载支架、手握、摆臂、裤兜、背包等多种形态，对手机接收到的信

号会产生一定的影响，并且对惯导等位置推算算法的精度也会造成一定的挑战。

当前结果

在城市中心高楼密集的区域（城市峡谷），GNSS 定位精度差，定位和导航的漂移情况严重。

技术诉求

结合当前手机能获得的信息（包括 GPS 卫星信号、各类传感器信号、Wi-Fi/蓝牙/基站信号，以及通过网络从服务器获取的数据如差分数据等），自行定义和采集验证所需数据集，设计高精度定位方案，能做到准确识别出当前的环境是否为城市峡谷卫星信号遮挡/多径严重的区域，并能在这类场景中进行高精度定位。

注：

1、 要求所设计的高精度定位方案具有普适性，能够有效适用于全国范围内的各种城市峡谷场景（不局限于当前验证的个别特定场景）

2、 对于需要采集用于验证/学习训练的数据，如果是需要大批量大范围采集的数据，则需要提供对应可行的众包采集方案（如何确保采集数据的可靠性和效率，以及方案是否原则上不违反正常用户使用习惯）

相关指标如下：

1、首个位置点定位精度 $CEP_{95} < 5m$ （即有 95% 的概率其定位结果会在这个半径 5m 的圆内），定位速度 $ttff$ （首次定位时间） $< 5s$ （定位速度以方案实际应用于目前市场上主流旗舰手机型号上的实测效果为准）

注：CEP 是 Circular Error Probability，圆概率误差

2、能够持续实时定位，每秒更新位置结果，轨迹出现漂移的精度 $< 5m$

3、场景识别准确率 $> 80\%$ （不同场景的总样本量 > 50 ，需要包含空旷、部分遮挡、严重遮挡等各种场景，即在众多场景中识别出城市峡谷场景的准确率 $> 80\%$ ）

4.3 基于可穿戴设备的水下通信的技术研究

技术背景

随着技术创新、军事需求与民用普及，人类在对水下资源探索的需求越来越强烈。水下通信是实现海洋/湖泊资源开发、环境监测、潜水活动等应用的关键技术。由于传统的无线通信的载体——电磁波在水中衰减严重（频率越高，水中衰减越大），其中海水比淡水对电磁波的衰减更大，高频的短波几乎无法穿过海水传播，因此传统的无线电通信不适合用于水下通信。

当前主流技术是水声通信（利用声波），其传播距离可达千米级，但存在带宽窄、延迟高、多径效应等问题。其他技术如光通信（方向性要求高）、电磁波低频通信（需巨型天线）和磁感应通信（新兴技术）也在探索中。

技术挑战

当前水下通信存在的问题：

1. 带宽受限与速率低：声波可用带宽随距离增加急剧下降，导致通信速率受限；
2. 多径效应与时延扩展：声波经海面、海底反射形成多路径信号，导致码间干扰；
3. 多普勒效应与信道时变：收发端相对运动或海水参数（温度、盐度）变化引起信号失真；
4. 环境噪声干扰：海洋生物、船舶、风浪等噪声会降低信噪比，从而导致丢包。

另外，还存在水下与水上的跨介质通信的难题：电磁波在跨空气-海水界面时衰减严重，光波易被吸收或散射，而声波难以跨越介质界面。现有技术如极低频（ELF）电磁波虽穿透性强，但需庞大天线和高功耗。

本课题需要评估清楚以上技术的可行性，还需要将可行的技术转换落地到小型化的可穿戴设备里面，同时需要具备低功耗的能力，以便于未来商用普及，最终形成水下通信网络。

当前结果

手机（2.4G 频段的无线电通信）在 LOS（空口无遮挡）场景下最远的传输距离有数百米，但是将其放到水下测试，实测传输距离小于 1 米。

技术诉求

基于现有的手机或手表的硬件（可以做适当的器件定制），第一阶段的目标：

- 1) 实现收发两端都在水下，稳定的双向传输距离大于 10 米；
 - 2) 解决跨介质信号衰减问题：实现收发一端在水下，另外一端在水上，稳定的双向传输距离大于 10 米；
 - 3) 以上两个场景的传输延迟需要小于 1 秒
- 基于现有的手机或手表的硬件（可以做适当的器件定制），第二阶段的目标：
- 1) 实现收发两端都在水下，稳定的双向传输距离大于 30 米；
 - 2) 解决跨介质信号衰减问题：实现收发一端在水下，另外一端在水上，稳定的双向传输距离大于 30 米；
 - 3) 提升通信性能：在相同的传输距离下，传输速率提升至第一阶段的 2 倍以上

4.4 识别用户使用手机时不同的使用场景

技术背景

一、用户不同使用场景下天线性能优化，提升通信体验

- 1、当前手机设计天线与结构件共体，紧凑的布局在手机四周（见图 1）
- 2、当用户在使用手机的时候，手机天线会靠近人体，人体会导致手机天线性能有一定恶化（①对天线效率有吸收 ②影响天线阻抗，导致天线有频偏从而失配，见图 2），从而影响手机的无线发射功率
- 3、如果可以识别用户不同使用场景，那么就可以通过天线调谐开关在对应使用场景调谐天线阻抗，减少失配影响，优化天线性能

二、降 SAR 方案优化，提升通信体验

- 1、当前手机需要满足不同市场 SAR（比吸收率）法规要求，从而对不同场景手机发射功率有一定要求（手机发射功率越高，SAR 越高）
- 2、现有方案为解决 SAR 与手机发射功率之间的矛盾，解决方案之一是通过电容传感器（Sar Sensor）识别到有一定介电常数的物体（包括人体、金属等）靠近后传感器触发（技术原理见图 3），手机降功率满足 SAR 法规；在自由场景（无介电常数物体靠近）提升手机发射功率，优化通信体验
- 3、但是现有的电容传感器无法识别是靠近人体还是靠近其他物体；如果可以识别人体还是非人体（金属等），那么在非人体场景就可以不用降功率，进一步提升通信性能

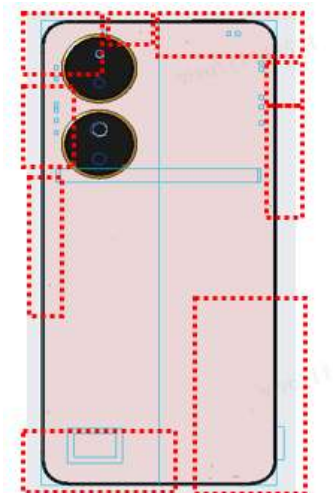


图 1 手机天线布局



图 2 用户使用手机时对天线频偏影响

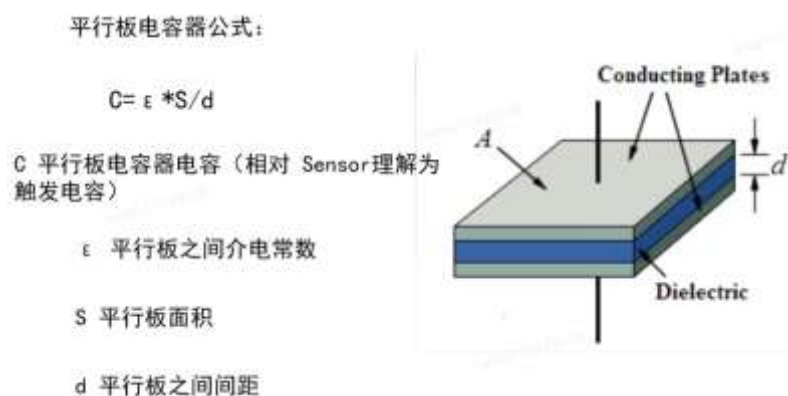


图 3 电容传感器基本原理

技术挑战

- 1、基于手机现有的硬件识别用户不同的用户使用场景（单手左手握、单手右手握、横竖屏、用户跑步、走路等场景），且作为通用方案，能够适用所有机型
- 2、基于手机现有的硬件识别是人体还是非人体靠近手机，作为电容传感器辅助判断，且作为通用方案，能够适用所有机型

当前结果

- 1、可以识别横竖屏场景，其他场景暂时无法识别
- 2、有方案可以识别人体还是非人体，但是当前方案识别精度无法满足需求，或者无法适用所有机型

技术诉求

- 1.通过手机现有硬件识别用户使用场景：单手左右/右手场景使用手机、横屏/竖屏场景使用手机、走路/跑步等场景使用手机；其中识别率>95%，误识别率为零（可以有一定程度不识别率，但是识别场景不会误判）
- 2.通过手机现有硬件识别是否是人体靠近还是非人体靠近，其中识别率>95%，误识别率为零（可以有一定程度不识别率，但是识别场景不会误判）

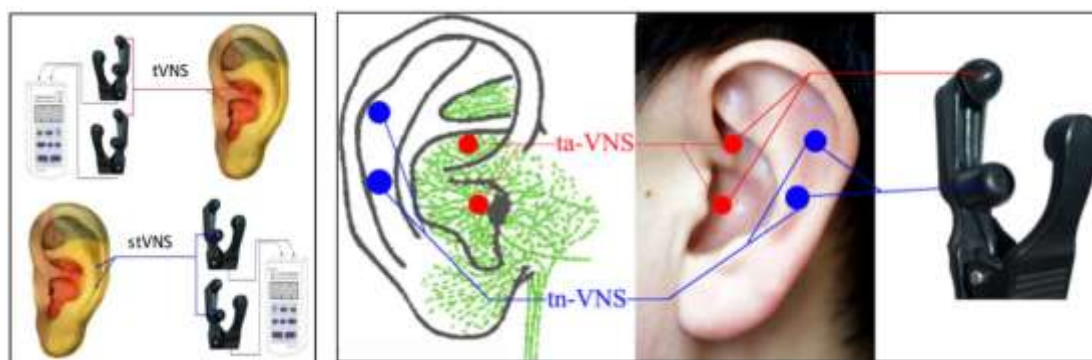
五、运动与健康

5.1 睡眠干预方法研究

技术背景

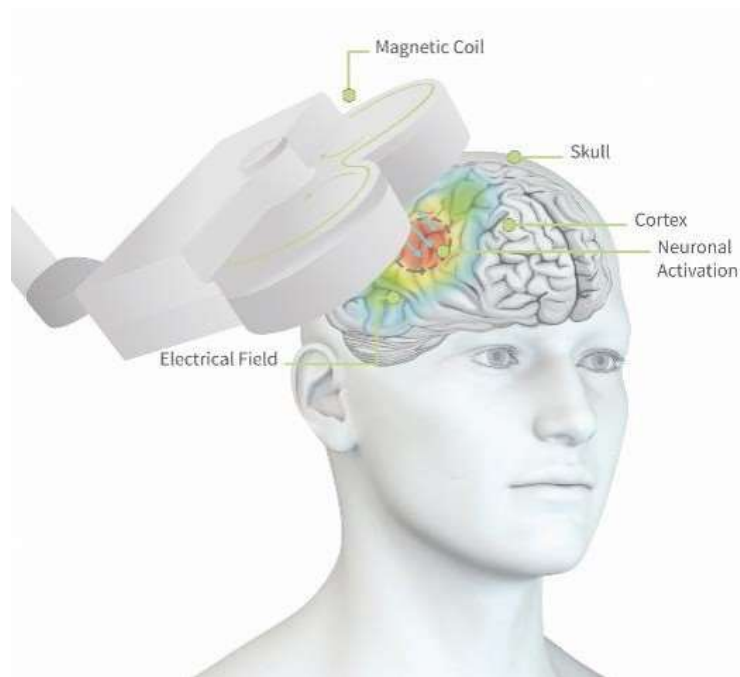
随着社会工作节奏的加快和生活压力的增加，睡眠问题已成为全球公共卫生领域的重大挑战。研究表明，全球约有 30% 的人群受到不同程度的睡眠障碍困扰，例如入睡困难、睡眠质量下降和早醒等。这些问题不仅影响个体身心健康，还与认知功能下降、生产力降低以及慢性疾病风险增加密切相关。传统睡眠干预方法包括心理干预和药物干预。心理干预，如认知行为疗法（CBT-I），虽然效果显著，但见效较慢，且需要专业指导，普及性受限；药物干预虽然见效快，但存在药物依赖、耐药性以及副作用等风险。因此，非药物物理干预因其无创性、低副作用和易用性，逐渐受到关注，成为研究热点。

当前，非药物物理干预技术包括经皮迷走神经刺激（tVNS）、脉冲电磁场疗法（PEMF）、经颅磁刺激（TMS）等已有一定研究基础的方案，以及光遗传刺激（Optogenetics）和脑机接口（BCI）等前沿技术方向。这些技术通过调控神经系统或脑电活动，试图改善睡眠质量。然而，现有技术在有效性、便捷性和普及性等方面仍存在显著不足，尚未形成公认的标准化干预方法或设备。



tVNS 示意图

(<https://sukhrajshothi.medium.com/tvns-therapy-electrocuting-your-ears-60c038c147f>)



经颅磁刺激（TMS）示意图

技术挑战

开发高效、便捷的非药物睡眠干预技术面临以下关键挑战：

1. 有效性验证

现有技术如 tVNS、PEMF 和 TMS 在改善睡眠方面的机制尚未完全明确，临床试验数据有限，缺乏大规模、多中心的验证研究。

2. 个体差异

睡眠问题的成因复杂，个体神经生理特征差异显著，单一干预方法难以适应不同人群的需求。

3. 设备便捷性

现有设备多为实验室级别，体积较大、操作复杂，难以满足家庭使用的便捷性需求。

4. 安全性与长期效应

非药物干预技术的长期安全性尚未充分验证，例如 TMS 可能引发的头痛或神经兴奋性改变，需进一步研究。

5. 技术成熟度

光遗传刺激和 BCI 等前沿技术目前仍处于理论或动物实验阶段，距离临床应用尚有较大差距。

当前结果

领域当前现状举例

1. 经皮迷走神经刺激 (tVNS)

tVNS 通过刺激耳部迷走神经，调节自主神经系统活动，增强副交感神经功能，促进睡眠。研究表明，tVNS 在改善慢性失眠和原发性失眠患者的睡眠质量方面具有显著效果。例如，Yap et al. (2024) 的一项随机对照试验发现，tVNS 组的匹兹堡睡眠质量指数 (PSQI) 评分较基线降低约 20% ($p < 0.05$)，表明睡眠质量显著改善 (Transcutaneous Auricular Vagus Nerve Stimulation)。另一项研究 (Luo et al., 2022) 显示，tVNS 治疗组的睡眠质量有效率 (PSQI 评分降低 $\geq 50\%$) 达 73%，显著高于对照组 ($p = 0.027$) (Transcutaneous Vagus Nerve Stimulation)。尽管 tVNS 显示出潜力，但其长期效果和最佳刺激参数仍需进一步优化，且当前设备的便携性和普及性有待提高。

2. 脉冲电磁场疗法 (PEMF)

PEMF 通过低频电磁场调控脑电活动，理论上可促进慢波睡眠和神经修复。早期研究为 PEMF 在改善失眠症状方面提供了初步证据。例如，Pelka et al. (2001) 的一项双盲安慰剂对照研究发现，8 周 PEMF 治疗使失眠患者的入睡时间缩短约 30 分钟，70% 的受试者报告显著或完全缓解 (Pulsed Electromagnetic Fields)。另一项研究 (Szent-Gyorgyi et al., 1996) 显示，低能量发射治疗 (一种 PEMF 形式) 可增加睡眠周期约 30% ($p = 0.0001$)，并显著减少入睡延迟 (Low Energy Emission Therapy)。然而，这些研究较早，当前缺乏近期大规模研究来确认其有效性，且设备的便携性和家用普及性仍需改进。

3. 经颅磁刺激 (TMS)

TMS 通过磁场刺激大脑皮层，调节神经兴奋性，已被证明在改善伴有抑郁症的失眠症状方面具有显著效果。例如，Croarkin et al. (2022) 发现，TMS 治疗后抑郁症患者的睡眠质量显著改善 (PSQI 评分降低)，且这种改善独立于情绪变化 (Effects of Transcranial Magnetic Stimulation)。Liu et al. (2025) 的一项系统综述分析进一步证实，重复性 TMS (rTMS) 显著改善了超过轻度抑郁情绪患者的睡眠质量 ($p < 0.05$) (The Effect of Repetitive Transcranial Magnetic Stimulation)。尽管 TMS 在抑郁相关失眠中的应用较为广泛，但其在非抑郁症失眠患者中的效果仍需进一步研究。此外，TMS 设备成本高昂，且可能出现副作用如头痛，限制了其广泛应用。

4. 光遗传刺激 (Optogenetics)

光遗传刺激通过光控神经元活动实现精准调控,主要用于动物实验以研究睡眠的神经机制。例如, Boyce et al. (2011)使用光遗传技术干扰小鼠的睡眠连续性,发现这会显著损害记忆整合,强调了睡眠连续性对认知功能的重要性 (Optogenetic Disruption of Sleep Continuity)。Jego et al. (2013)则通过光遗传技术识别了下丘脑中黑色素浓缩激素 (MCH) 神经元组成的一个回路,该回路在调控快速眼动 (REM) 睡眠中发挥关键作用 (Optogenetic Identification of REM Sleep Circuit)。这些研究为理解睡眠调控的神经回路提供了重要见解,但光遗传刺激目前仍限于实验室研究,临床应用面临伦理和技术障碍。

5. 脑机接口 (BCI)

BCI 通过记录和调控脑电信号,理论上可用于睡眠优化。Van der Werf et al. (2012)的研究发现,通过 BCI 训练可以显著增加睡眠梭子 (与记忆整合相关的脑电波) 的密度,表明 BCI 可能通过神经反馈影响睡眠质量 (Sleep Spindles Modulated by BCI)。MNEMOSYNE 项目 (2024) 则探索了在睡眠期间使用 BCI 操纵厌恶记忆,以治疗创伤后应激障碍等与恐惧或焦虑相关的疾病 (Brain Computer Interface for Memory Manipulation)。尽管 BCI 在睡眠调控方面的潜力正在被探索,但其作为睡眠干预工具的直接应用仍处于早期阶段,技术复杂性和设备依赖性是当前的主要挑战。

总体而言,现有技术在实验室环境中显示出一定潜力,但缺乏统一的标准化设备和广泛的临床验证,限制了其在实际场景中的推广。

当前技术积累

当前 vivo 在该方面技术积累相当有限,睡眠技术相关的积累基本都集中于监控测。虽然在应用层面存在一些基础功能,如助眠音乐或者科普性的睡眠卫生知识等,但实际并无坚实的有效性或原理性研究保证,其严谨性和客观价值乏善可陈。

因此,我们破解去需要构建优先干预睡眠的技术体系,理论体系,和针对有效性进行实验性验证的实验体系以更好,更科学的服务我们的用户。

技术诉求

任务: 开发可以镇静安眠的睡眠干预方法

1. 有效性: 具有原理上的可行性, 且具有初步的研究结果证明其有效性并且该实验是

未来在合作中可有效复现的。

2. 形态：最好适合手表、耳机等设备形态，也可以其他可穿戴或非接触形态，考虑到未来潜在的装备形态仍然具有很大的发展空间，形态上不做关键限制。

3. 安全性：对用户心理与生理无显著副作用。

参考文献

1. Boyce, R., Glasgow, S. D., Williams, S., & Adamantidis, A. (2011). Optogenetic disruption of sleep continuity impairs memory consolidation. *Proceedings of the National Academy of Sciences*, 108(36), 14915 – 14920. <https://doi.org/10.1073/pnas.1015633108>
2. Croarkin, P. E., Nakonezny, P. A., Deng, Z.-D., Romanowicz, M., Voort, J. L. V., Camsari, D. D., Lee, J., & Lewis, C. P. (2022). Effects of transcranial magnetic stimulation on sleep quality and mood in patients with major depressive disorder. *Journal of Clinical Sleep Medicine*, 18(5), 1341 – 1348. <https://doi.org/10.5664/jcsm.9948>
3. Jegu, S., Glasgow, S. D., Herrera, C. G., Ekstrand, M., Reed, S. J., Boyce, R., Friedman, J., Burdakov, D., & Adamantidis, A. R. (2013). Optogenetic identification of a rapid eye movement sleep modulatory circuit in the hypothalamus. *Nature Neuroscience*, 16(11), 1637 – 1643. <https://doi.org/10.1038/nn.3539>
4. Liu, Y., Zhang, X., Yang, L., & Wang, Z. (2025). The effect of repetitive transcranial magnetic stimulation on sleep quality in patients with more than mild depressive mood: A systematic review and meta-analysis. *Frontiers in Psychiatry*, 16, 1511930. <https://doi.org/10.3389/fpsyt.2025.1511930>
5. Luo, Y., Qu, D., Song, F., Li, X., Liu, S., & Zhang, X. (2022). Transcutaneous vagus nerve stimulation could improve the effective rate on the quality of sleep in the treatment of primary insomnia: A randomized controlled trial. *Brain Sciences*, 12(10), 1296. <https://doi.org/10.3390/brainsci12101296>
6. MNEMOSYNE Project. (2024). Brain computer interface to study and manipulate memories of aversive experience during sleep. CORDIS European Commission. Retrieved from <https://cordis.europa.eu/project/id/726169>
7. Pelka, R. B., Jaekel, M., & Fitzner, R. (2001). Impulse magnetic-field therapy for insomnia: A double-blind, placebo-controlled study. *Advances in Therapy*, 18(6), 255 – 260. <https://doi.org/10.1007/BF02850303>
8. Szent-Gyorgyi, A., et al. (1996). Low energy emission therapy (LEET) in the treatment of sleep disorders. *Sleep*, 19(4), 327 – 336. <https://doi.org/10.1093/sleep/19.4.327>
9. Van der Werf, Y. D., Van der Helm, E., Schoonheim, M. M., Ridderikhoff, A., & Van Someren, E. J. W. (2012). Sleep spindles are locally modulated by training on a brain-computer interface. *Proceedings of the National Academy of Sciences*, 109(49), 20467 – 20472. <https://doi.org/10.1073/pnas.1119112109>
10. Yap, J. Y. Y., Keatch, C., Lambert, A. J., Woods, W., Stansfeld, J. L., & Baron-Cohen, S. (2024). Transcutaneous auricular vagus nerve stimulation for chronic insomnia disorder: A randomized clinical trial. *JAMA Network Open*, 7(12), e2444221. <https://doi.org/10.1001/jamanetworkopen.2024.44221>

5.2 无创血糖检测探索性研究

技术背景

糖尿病是一种全球范围内常见的慢性疾病，血糖监测在糖尿病的治疗与管理中具有重要意义。目前主流的血糖检测方式主要依赖指尖采血，给患者带来了明显的痛苦与心理负担，也存在感染等卫生问题。因此，发展无创血糖检测技术成为迫切的研究方向。无创血糖的研究已有相当的历史，其难度与应用挑战已经在过去的研究与实践中有了充分讨论，但近期以拉曼光谱技术血糖监测为首的新研究进展显示了该技术领域在成熟度，可行性以及小型化机会上可能迈上了一个新的台阶。因此，该技术有必要和需求展开探索性研究。

无创血糖监测技术涉及众多物理和化学原理，主流方案大致可分为近红外光谱法、拉曼光谱法、体液电化学法、光声光谱法等。近红外光谱法通过检测皮肤穿透后的近红外光吸收特性来估算血糖浓度，具有快速、便捷的优点；拉曼光谱法利用激光照射后散射光的频移来分析葡萄糖分子特征，灵敏度较高，适合微量检测；体液电化学法则通过检测汗液、泪液等体液中葡萄糖含量，借助电化学传感器实现血糖推算，兼具便携性与较强的实时监测能力；光声光谱法利用组织对调制光吸收后产生的声波信号来间接测量血糖，具有较强的组织穿透性和选择性。这些方法各有优劣，正在不断发展以提高准确性和实用性，但整体上，离实际应用还有一定的距离。

技术挑战

1. 近红外光谱法（NIR）

近红外光谱法通过测量皮肤对近红外光的吸收特性来估算血糖浓度。然而，该方法面临多重挑战：

组织散射与水吸收干扰：人体组织对近红外光的强烈散射以及水分的高吸收率会显著降低信号的质量，影响测量的准确性。

个体差异性：皮肤厚度、色素含量和水合状态等个体差异会导致测量结果的不一致性，增加了算法校准的复杂性。

信号稳定性：环境温度、压力等外部因素可能影响测量的稳定性，限制了该方法在实际应用中的可靠性。

2. 拉曼光谱法

拉曼光谱法利用激光照射后散射光的频移来分析葡萄糖分子的特征。其主要技术挑战包括：

信号强度弱：葡萄糖的拉曼散射信号极其微弱，容易被背景噪声淹没，导致检测灵敏度低。

背景荧光干扰：皮肤中的蛋白质等物质产生的背景荧光可能与拉曼信号强度相当，增加了信号提取的难度。

设备复杂性：为了提高信噪比，需要使用高功率激光器和高灵敏度的探测器，导致设备体积大、成本高，不利于便携化和居家使用。

尽管如此，近年来的研究，如多重微空间拉曼光谱成像技术（ μSORS ），已在提高信噪比和降低个体差异影响方面取得了进展。

3. 体液电化学法

体液电化学法通过检测汗液、泪液等体液中的葡萄糖含量，借助电化学传感器实现血糖推算。其面临的技术挑战包括：

体液采集难度：汗液和泪液的分泌量受多种因素影响，采集稳定性差，可能导致测量结果不准确。

葡萄糖浓度低：体液中的葡萄糖浓度通常低于血液，检测灵敏度要求高，增加了传感器设计的难度。

干扰物质影响：体液中存在多种电活性物质，可能干扰葡萄糖的检测，影响结果的准确性。

4. 光声光谱法（PAS）

光声光谱法利用组织对调制光吸收后产生的声波信号来间接测量血糖。其主要技术挑战包括：

信号微弱：光声信号强度较低，容易受到环境噪声的干扰，导致信噪比低，影响检测的可靠性。

设备复杂性：为了提高信号强度，需要设计复杂的共振腔结构，增加了设备的体积和成本，不利于便携化。

个体差异性：皮肤厚度、血流量等个体差异会影响光声信号的产生和传播，增加了

算法校准的复杂性。

综上所述，尽管无创血糖监测技术在理论上具有广阔的应用前景，但在实际应用中仍面临诸多技术挑战。未来的研究需要在提高检测灵敏度、降低设备复杂性、减少个体差异影响等方面取得突破，以实现无创血糖监测技术的广泛应用。

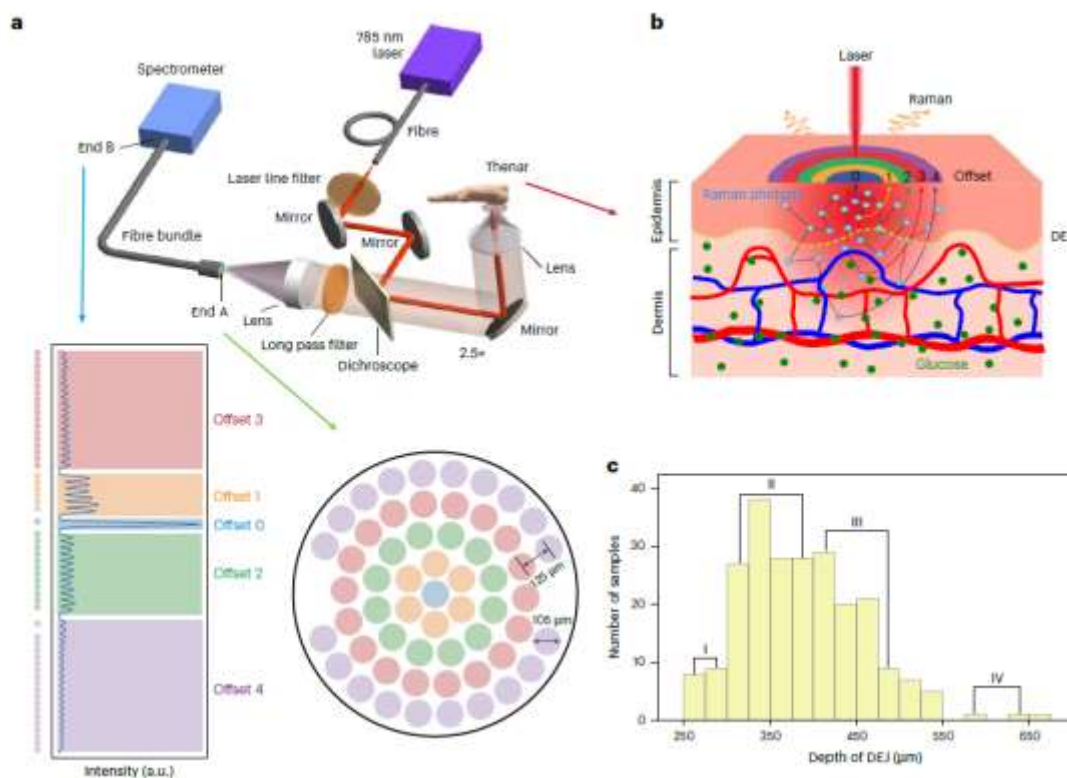
当前结果

1. 近红外光谱法

近红外光谱法利用 750–2500 nm 波段光对皮肤组织的穿透及反射特性测量血糖含量。由于近红外光在组织中水、血红蛋白等的吸收较弱，穿透性较强，可以获得葡萄糖分子的特征吸收信号 [1]。近年来，该方法主要通过选取最佳波长组合并结合多元校正模型或机器学习算法来提高精度。例如，有研究采用反射模式近红外测量手指静脉或舌下组织，应用偏最小二乘回归(PLSR)等方法进行标定，实验结果显示平均绝对相对偏差(mARD)可达 7.6%、95%以上的测量值分布在 Clarke 误差网格 A 区 [2]。尽管取得一定进展，但近红外测量仍受组织散射、皮肤水分和生理多样性影响，需要进一步扩大样本规模和改进算法，以提高算法泛化能力和实时连续监测性能 [1][2]。

2. 拉曼光谱法

拉曼光谱法通过检测散射光中葡萄糖分子的振动指纹信号实现浓度估计，具有化学特异性高的特点。常见研究包括共聚焦拉曼和空间偏置拉曼(SORS)等技术。2025 年上海交通大学团队提出的多重微空间拉曼光谱(m μ SORS)方法，可深度选择性地扫描皮肤下真皮层边界处，获得稳定的葡萄糖拉曼指纹 [3]。该方法在 230 名糖尿病患者临床测试中，构建的光谱回归模型实现了平均绝对相对偏差 14.6%，99.4%的预测值落在 Clarke 网格 A/B 区 [3]。此外，丹麦等团队开发的 830 nm 光源共聚焦拉曼仪器经两日校准后，在 50 名受试者中实现了 12.8% MARD 和 100%的 A/B 区可靠性 [4]。这些结果表明拉曼技术在不需要个性化长期标定下具有良好的监测潜力。另一类研究利用表面增强拉曼(SERS)效应，例如银纳米结构微针阵列，可将测量点直接置入表皮下提高信噪比。不过此类方法涉及微创微针，尚属准无创范畴。总体而言，拉曼法以其分子选择性与光谱丰富度为优势，最新进展集中在深度选择性采样、便携化仪器和短期自动标定方案上 [3][4]。



微型空间偏移拉曼光谱成像（mSORS）示意图

体液电化学法

体液电化学法主要指从汗液、泪液等体液中检测葡萄糖，借助可穿戴电化学传感器（酶法或非酶法）实时测量信号。近期研究多聚焦于汗液传感。重庆医科大学等团队开发了一种柔性非酶电化学汗液葡萄糖传感器，将 Au 纳米粒子修饰的多壁碳纳米管（AMWCNTs/AuNPs）掺杂高分子胶结合到屏幕印刷电极上，线性检测范围 50 – 600 $\mu\text{mol/L}$ 、检测限 3.2 $\mu\text{mol/L}$ ，可检测极低浓度，并在 500 次 180°反复弯折后性能无明显退化 [5]。该传感器在体表贴附测试中成功监测到汗液葡萄糖随血糖波动的变化，其结果与正常生理血糖波动趋势一致 [5]。以上工作表明，非酶可挠传感器具有灵敏度高、柔性好和耐弯折等优点，可用于连续无创汗液葡萄糖监测。其他研究还报道了基于碳纳米复合墨水的印刷式可穿戴传感器、多孔 Pt/MXene 电极等方案。总体看，体液电化学法在灵敏度和服用性上进展显著，但要解决汗液样本量和校准一致性问题，还需进一步研发更稳定的传感阵列和多点校准模型 [5]。

光声光谱法

光声光谱法利用中红外或近红外激光照射人体组织，葡萄糖吸收光能产生声波信号，超

声换能器检测声学信号并定量分析葡萄糖浓度。这一方法结合了红外光谱的特异吸收与声学放大作用，被认为对血糖测量非常有前景 [6]。其核心挑战在于光声信号弱，需要设计高品质因子的光声谐振腔提高信噪比。2023 年 Aloraynan 等开发了双波长量子级联激光 (QCL) 光声系统，在人体皮肤模拟模型上实现了 ± 12.5 mg/dL 的检测灵敏度，并利用机器学习分类使 96.7% 的预测值落在 Clarke A/B 区，符合 FDA 和加拿大标准 [7]。2024 年 Uluç 等提出了深度门控中红外光声传感器 (DIROS)，通过时间门控聚焦血管丰富层，抑制表皮干扰，使最终 88% 的测量值位于误差网格 A 区（无门控时仅 65%） [8]。此外，2024 年综述也报告了多种光声谐振腔设计和系统方案，并指出合理选择谐振器几何结构、激光调制频率等对提高检测灵敏度至关重要 [6][9]。总体来看，光声法结合了直吸收和声学放大的优点，新一代中红外激光器和深度选择技术使得无创糖监测的准确度获得显著提升。未来该技术需在安全功率、人体实验和系统集成方面持续优化，以实现在真实临床环境中的连续监测应用。

技术诉求

技术方案：包含但不限于前述提及的无创血糖监测方案，考虑到当前领域研究，拉曼光谱族可以作为优先探索的技术方案，但本质上不做硬性要求。

构建可靠而且可进行小规模验证的技术原理以及验证方案

如果是前面提及的方案，需要说明应如何应对前面提及的技术挑战，比如信号 SNR 低、皮肤测量环境复杂，测量不稳定，人群差异大等问题

如何实现小型化和控制成本，该部分可以作为长线技术合作计划提出

优先考虑手表、耳机等穿戴形态，但考虑到未来产品形态仍具有相当的发展空间，对形态不做硬性要求。

参考文献

- [1] A. Hina and W. Saadeh, “Noninvasive Blood Glucose Monitoring Systems Using Near-Infrared Technology—A Review,” *Sensors*, vol. 22, no. 13, Art. 4855, 2022.
- [2] Z. Alian et al., “In vivo near-infrared noninvasive glucose monitoring using tongue and finger reflectance,” *Journal of Biomedical Optics*, vol. 27, no. 8, 2022.
- [3] Y. Zhang et al., “Subcutaneous depth-selective spectral imaging with μ SORS enables noninvasive glucose monitoring,” *Nature Metabolism*, vol. 7, pp. 421–433, 2025.
- [4] A. Pors et al., “Calibration and performance of a Raman-based device for

- non-invasive glucose monitoring in type 2 diabetes,” *Scientific Reports*, vol. 15, Art. 10226, 2025.
- [5] Y. Chen et al., “A wearable non-enzymatic sensor for continuous monitoring of glucose in human sweat,” *Talanta*, vol. 278, Art. 126499, 2024.
- [6] R. A. M. Emon et al., “Photoacoustic Resonators for Non-Invasive Blood Glucose Detection Through Photoacoustic Spectroscopy: A Systematic Review,” *Sensors*, vol. 24, no. 21, Art. 6963, 2024.
- [7] A. Aloraynan et al., “Dual quantum cascade lasers for noninvasive glucose detection using photoacoustic spectroscopy,” *Scientific Reports*, vol. 13, Art. 7927, 2023.
- [8] N. Uluç et al., “Non-invasive measurements of blood glucose levels by time-gating mid-infrared optoacoustic signals,” *Nature Metabolism*, vol. 6, pp. 678 – 686, 2024.
- [9] A. Singh et al., “Advanced resonator designs for improving mid-infrared photoacoustic glucose sensors,” *Sensors and Actuators B: Chemical*, vol. 416, Art. 134019, 2024.

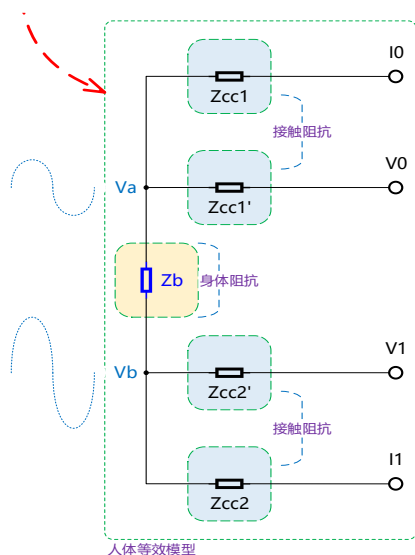
5.3 高接触阻抗下的生物电阻抗测量技术

技术背景

随着人们对健康管理重视程度的不断提升，大家开始关注身体的多项健康指标(比如心率、心电、血氧、体脂、肌肉含量等指标)，因此智能手表设备在日常生活中得到了广泛应用。其中，人体成分测量的核心原理是生物电阻抗法(BIA)：向人体施加一个安全的低频交流电流，通过测量人体的电阻抗值来推算身体成分。当电流通过人体时，由于不同身体组织的导电特性不同，如脂肪组织的导电能力较差，而肌肉、血液等组织的导电能力较强，因此电阻抗值与身体成分之间存在一定的相关性。

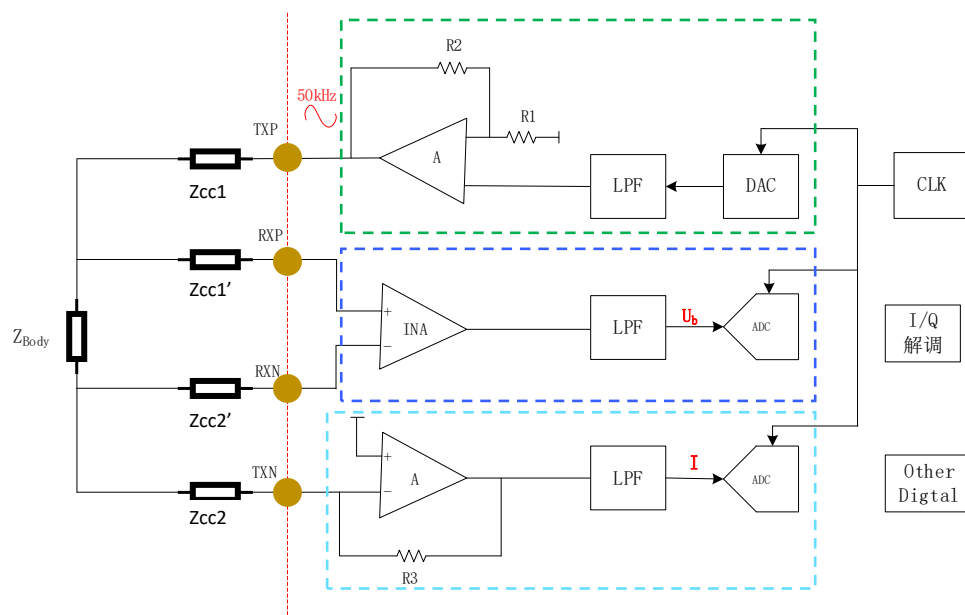
在实际使用过程中，设备会面临高接触阻抗的问题。接触阻抗主要来源于人体皮肤与检测电极之间的接触情况。皮肤表面的角质层、汗液分泌量、皮肤湿度、接触面积以及手表佩戴的松紧度等因素都会影响接触阻抗的大小。当接触阻抗较高时，施加到人体的电流会显著减小，导致测量信号变得微弱，从而使得难以准确获取人体的电阻抗值。

随着用户对身体成分测量精度要求的不断提高，解决高接触阻抗下的准确测量问题已成为检测技术的关键挑战之一。



技术挑战

BIA 测量电路架构如下：



目前在高接触阻抗环境中，接触电极获取的电流信号极其微弱，容易受到外界电磁干扰和电路噪声的影响。如何设计高灵敏度、低噪声的信号采集电路，以准确提取包含人体电阻抗信息的微弱信号，是首要技术挑战。除此以外，在接触阻抗的实时监测与补偿、复杂接触条件下的测量模型优化、硬件设计与材料选择等方面依旧存在设计瓶颈。例如，接触阻抗监测的实时性和准确性还有待提高，复杂接触条件下的补偿模型还需要进一步优化，新型电极材料的耐用性和成本问题也需要进一步研究。未来，需要结合多学科的理论和技术，开发更加有效的解决方案，以提高实际应用中的测量准确性和可靠性。

当前结果

针对上述挑战，可以从以下 4 个方面进行研究与解决。下列方向供参考，鼓励参与者提出其他创新方案：

1.高接触阻抗下的信号采集与处理

通过采用高分辨率的 ADC、低噪声运算放大器，结合数字处理和滤波，提高对微弱信号的采集能力和抗干扰性能。同时采用多频检测技术，提升激励频率，改善接触阻抗的影响。

2.接触阻抗的实时监测与补偿

准确实时地监测接触阻抗的大小，并根据接触阻抗的变化对测量过程进行补偿。同时，针对不同的接触阻抗值，需要建立相应的补偿模型，对测量信号进行修正，以消除接触阻抗对人体电阻测量的影响。

3.复杂接触条件下的测量模型优化

人体皮肤与电极的接触情况复杂多变，不同用户的皮肤特性、接触方式以及使用场景都会导致接触阻抗存在显著差异。需要结合人体生理特征和接触界面的电学 or 生物学特性，建立更精确的测量模型，通过优化模型参数和算法，提高设备在复杂接触条件下的测量准确性。

4.硬件设计与材料选择

电极设计和材料选择直接影响接触阻抗的大小。常用的电极材料如金属电极在皮肤表面容易形成较高的接触阻抗，而导电凝胶等材料虽然具有较好的导电性，但耐用性较差。可以研究新型的电极材料，具有低接触阻抗、高稳定性以改善皮肤与电极之间的接触性能。也可以在电极的结构设计方面，优化电极的形状、尺寸、皮肤接触形式等，从而降低接触阻抗的影响。

技术诉求

任务：课题申请人可在主要技术诉求和创新技术诉求中任选其一，也可两者都选。

1.主要技术诉求：

5KHz-1000KHz 范围激励信号下，动态范围 10Ω~1500Ω，

准确度要求如下：

动态范围 10-300Ω，准确度：≤±3Ω；

动态范围 300-1500 Ω ，准确度： $<\pm 1\%$ ；

接触阻抗要求：覆盖阻抗和容抗，例如：50KHz 频率，等效阻抗需要支持 $\geq 100K\Omega$

2. 创新技术诉求：

不限制产品形态，创新性提出一种便于准确测量身体生物电阻抗的设备，需具备易用性、便携性、家用属性、成本经济性，实现可以高接触阻抗下的身体成分测量功能(包含体脂、肌肉量、骨盐量等信息)。

也可以尝试基于现有常见品类增加配件的形式，提出优雅、易用的解决方案，满足上述需求并解决痛点。

5.4 视觉敏感人群与显示需求研究

技术背景

在消费电子行业迅猛发展的今天，手机产品已深度渗透至人们的生活与工作场景。然而近年来，众多手机产品发布后，视觉舒适度失效率一直居高不下。用户频繁反馈，使用产品后出现眼睛干涩、疲劳、头晕等症状，部分用户甚至表示，使用手机 30 分钟内便会产生眼部不适。这一问题不仅严重影响了用户体验，降低了产品满意度，还对终端品牌形象造成了负面影响。因此，深入研究敏感人群的视觉敏感因子和需求，开发有针对性的显示模式和策略，成为提升产品竞争力、保障用户视觉健康的关键所在。

技术挑战

1. 敏感因子多样性：视觉敏感因子种类繁多，涵盖光、色彩、动态变化等多个维度，且不同个体对各因子的敏感程度差异显著。如部分用户对光敏感，而另一部分则对屏幕频闪更为敏感，如何全面且准确地识别这些敏感因子是首要难题。

2. 人群特征分析：明确各敏感因子对应的人群占比与影响权重难度较大。目前缺乏大规模、多维度的调研数据，难以精准量化不同敏感类型在整体用户群体中的分布情况，以及其对用户视觉体验和健康的影响程度。

3. 检测方法局限：现有的敏感因子检测或辨别方法存在局限性，无论是主观的用户自我报告，还是客观的仪器检测，都无法完全准确地判断个体的敏感类型和程度，亟需开发更科学、高效的检测手段或分析方法。

当前结果

目前已初步发现一些视觉敏感现象，例如干眼症患者往往对光线更为敏感，容易出现畏光症状；部分人群对屏幕频闪感知明显，长时间观看可能引发头晕、恶心等不适。但这些发现较为零散，缺乏系统性的研究和深入分析，尚未形成完整的理论体系和实践指导。

技术诉求

1、对视觉敏感人群进行定义与分类

视觉敏感人群是指因生理(含生理老化)、病理或个体差异等因素，对显示设备特定视觉参数产生暂时性或持续性的过度敏感反应，进而导致视觉不适或功能障碍的群体。

2、提供敏感人群视觉的显示策略建议或改善方案

举例：XX 敏感人群，通过 XX 方式辨别，人群占比 XX，对 XX 容易感到不适，在 XX 场景下，建议屏幕显示 XX 设置在 XX，采取 XX 策略，对敏感人群将更为友好。

以上诉求可通过文献、用户调研、人因研究来完成，我司亦可提供研究环境与相关资源支持，具体课题范围可进一步交流讨论

参考文献

1. 年龄因素对视觉敏感度影响-深度研究
<https://www.renrendoc.com/paper/397473768.html>
2. 分析不同类型干眼症患者的临床表现及其特点 庄博 中国实用医药 2017

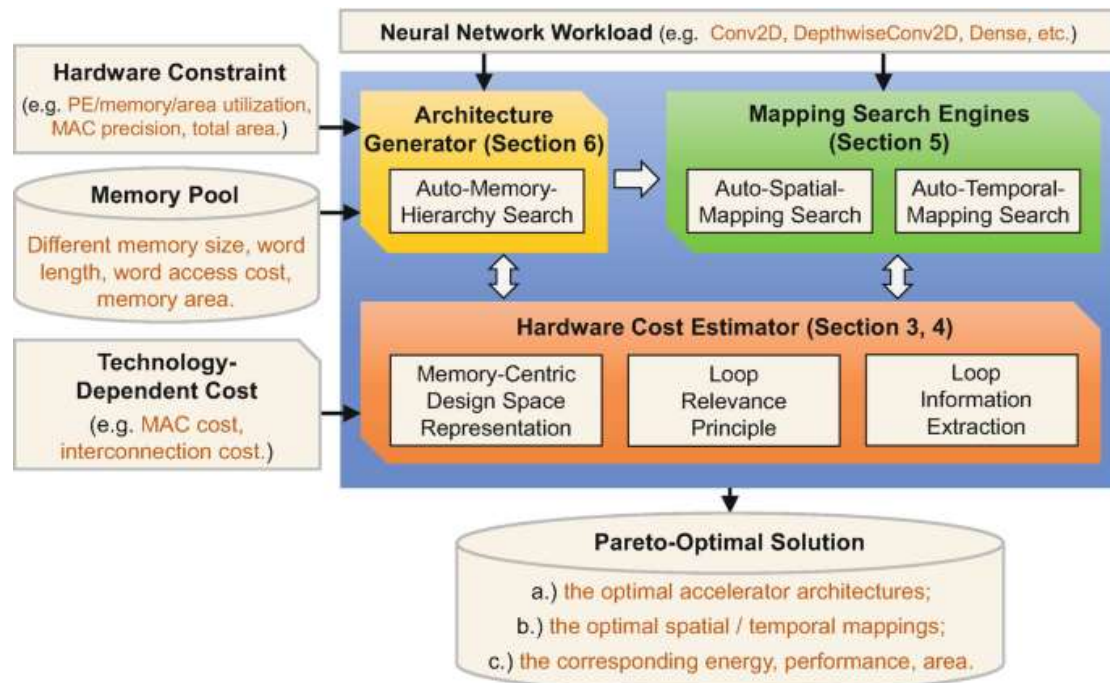
六、芯片技术

6.1 基于 AI Agent 的 AI 加速器架构探索平台

技术背景

AI 加速器架构探索平台是针对 AI 加速器的设计与优化工具，旨在通过仿真、资源调度分析、能效评估等技术，高效探索不同架构配置的性能表现。其核心功能包括：

- 参数化建模与仿真：支持可配置的计算单元、内存层次等硬件模块的快速建模与性能模拟。
- 多目标优化：平衡算力、功耗、面积等指标，动态调整架构参数。



AI Agent 技术在电子设计自动化（EDA）领域的应用正经历快速发展期。强化学习驱动的芯片设计已在多个关键环节实现突破：Google 的 Circuit Training 使用深度强化学习优化 TPU 芯片布局，实现显著的 PPA（性能、功耗、面积）优化；NVIDIA 的 PrefixRL 在并行前缀电路设计中比传统 EDA 工具减少 25% 面积且速度相当。商业化 EDA 工具正快速集成 AI 功能。Synopsys DSO.ai 使用强化学习技术自主搜索万亿级设计空间，实现 3 倍生产力提升；Cadence Cerebrus 基于机器学习的芯片优化系统自动优化数字设计流程。机器学习增强的设计空间探索已从概念验证转向产业部署，AI 驱动的拥塞预测、时序优化和功耗分析工具显著减少了设计迭代周期。

技术挑战

自动化架构搜索与参数优化

- 强化学习的动态决策：使用强化学习策略 RL Agent 或者 LLM Agent 探索最优参数组合。例如，根据推理延迟、能耗等目标函数动态调整计算核数量或内存带宽分配。
- 搜索空间优化：AI Agent 根据架构约束和实现的限制，对搜索空间进行多层次优化，缩小搜索空间，提取有效搜索参数。
- 多任务协同优化：多 Agent 框架可协同处理复杂设计约束（如制程限制、散热需求），在参数空间中高效搜索可行解。
- 可解释性与可信度：Agent 的决策需提供可视化解释（如架构变更对能效的影响路径）。

跨层级协同优化

- 软硬件联合设计：AI Agent 分析模型计算图特征（MLLM/AIGC），推荐匹配的加速器架构参数（如计算阵列规模），同时生成适配的驱动代码或算子优化策略。
- 生态工具链集成：基于标准化的函数调用协议（参考 Google 工具链设计），Agent 可调用 HLS/Chisel/LLM 工具链实现架构设计的自动化代码生成与验证。

仿真加速与智能验证

- 自动化仿真流程：利用代码生成 Agent，自动生成硬件仿真脚本与分析报告，替代传统手动调试环节。
- 仿真效率瓶颈：大规模架构搜索的计算开销过高，需结合轻量化仿真模型与 Agent 驱动的增量优化策略。

技术现状

相对较多的学术论文和 EDA 公司关注于 AI assisted RTL to GDS 的过程，比如 Synopsys 的 DSO，Nvidia 的 ChipNeMo，PrefixRL，Graph Cluster。对于 AI 加速器的架构探索框架，学术界开源了较多的框架，比如 MICAS 的 Zigzag，MIT 的 Timeloop/Accelergy，UCB 的 VAESA/COSA。但很少框架引入 AI Agent 进行自动化架构搜索和参数优化。

技术诉求

可选择如下任务进行探索：

1. 采用 LLM Agent 指导传统 MOO 优化算法进行探索，并提供可视化解释；
2. 实现多 Agent DSE 框架进行协同优化，包含（搜索空间优化 Agent，MOO 监督 Agent，代码生成 Agent，性能功耗预测 Agent 等）；

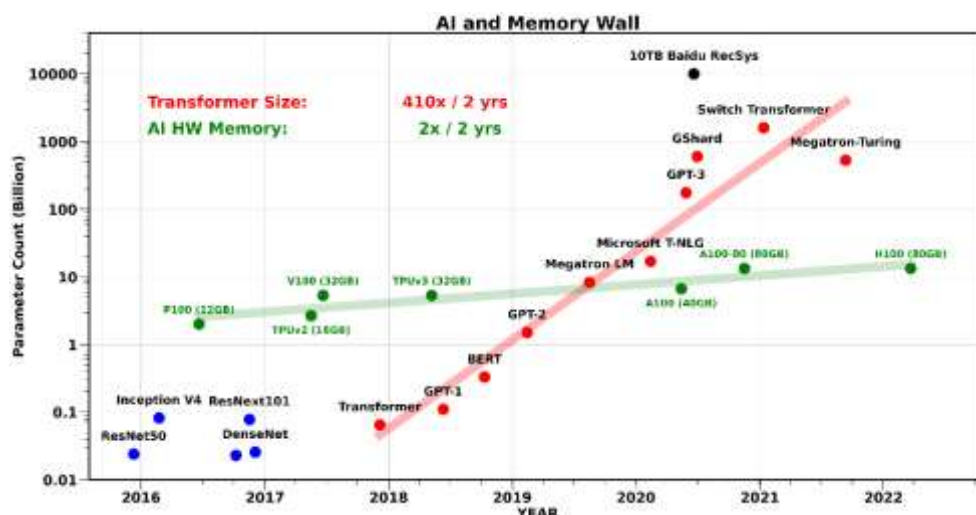
参考文献

1. L. Mei, P. Houshmand, V. Jain, S. Giraldo, and M. Verhelst, "Zigzag: A memory-centric rapid dnn accelerator design space exploration framework," arXiv preprint arXiv:2007.11360, 2020.
2. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," IEEE transactions on evolutionary computation, vol. 6, no. 2, pp. 182 - 197, 2002.
3. A. Parashar, P. Raina, Y. S. Shao, Y.-H. Chen, V. A. Ying, A. Mukkara, R. Venkatesan, B. Khailany, S. W. Keckler, and J. Emer, "Timeloop: A systematic approach to dnn accelerator evaluation," in 2019 IEEE international symposium on performance analysis of systems and software (ISPASS). IEEE, 2019, pp. 304 - 315.

6.2 基于存算技术的端侧高效 AI 加速器的技术方案

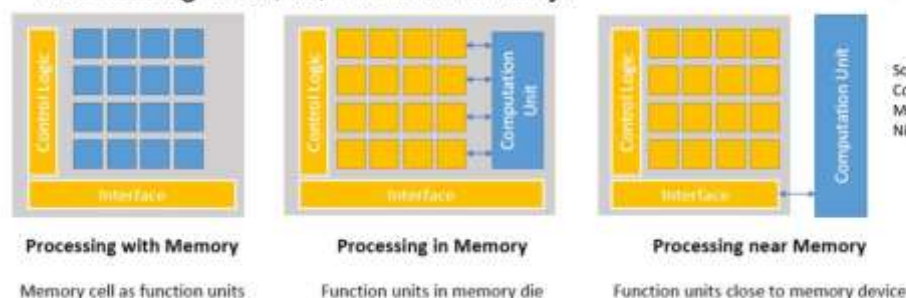
技术背景

人工智能技术的快速发展，特别是大规模语言模型（LLM）和 AI 生成内容（AIGC）的兴起，对传统芯片架构提出了前所未有的挑战。当前 AI 芯片设计面临从"计算受限"向"内存受限"的根本性转变，这一转变重新定义了芯片架构设计的核心考量。大模型推理和训练呈现出显著的内存密集型特征，参数量从 GPT-3 的 175B 增长至超万亿级别，远超单个 GPU 的 HBM 容量限制。内存墙问题已成为制约性能的主要瓶颈：内存带宽需求增长速度是计算能力提升的 3-5 倍，导致计算单元长时间等待数据传输。同时，Transformer 架构中 self-attention 机制的 $O(N^2)$ 时间复杂度和 85% 以上的矩阵乘法操作占比，要求芯片架构在保持高算力的同时，必须优化数据流和存储层次结构。AIGC 应用的多模态特征进一步复杂化了硬件需求。文本生成、图像合成、音频处理等不同模态需要差异化的计算原语和数据流模式，传统固定架构难以高效适配。混合精度计算需求（FP32、FP16、BF16、INT8、FP8）和动态负载均衡要求芯片具备更强的可配置性和自适应能力。



传统冯·诺依曼架构因存储与计算单元分离导致的“内存墙”问题（数据搬运功耗占总功耗 60%~70%），成为 AI 计算能效的核心瓶颈。存算一体（In-Memory Computing, IMC）通过将计算嵌入存储单元，直接在内存中完成矩阵运算，大幅减少数据移动，理论上能效比可提升 10-100 倍，尤其适合端侧低功耗场景。智能手机、AR/VR、机器人等端侧设备需实时处理视觉、语音和多模态 AI 任务（如实时目标检测、语音识别），但受限于算力（<10TOPS）和功耗（<5W），亟需通过存算技术突破能效瓶颈。新存储器技术（ReRAM、MRAM、PCM）和模拟计算电路的发展，为存算一体提供硬件基础。例如，ReRAM 交叉阵列可实现并行模拟矩阵乘法，显著加速神经网络计算。

• Processing near, in, or with memory:



技术挑战

• 模拟计算的非理想性

存储器件的工艺波动（如电阻漂移、噪声）导致计算精度下降，影响 AI 模型的推理准确率（如 ImageNet 分类误差增加 5%~10%）。

• 高精度 AI 任务的支撑不足

大模型（如 Transformer）需 16/32 位浮点计算，而模拟存算多支持低位整型（4/8-

bit)，需突破多位精度与非线性激活函数的硬件实现。

- **系统集成复杂度高**

存算模块需与数字逻辑电路（控制、量化、编码）协同工作，但混合信号设计（ADC/DAC 开销）可能抵消存算的能效优势。

- **算法与硬件的协同优化缺失**

现有模型(MLLM/AIGC)未针对存算架构优化（如权重稀疏性、数据分布），导致硬件利用率低，算法精度与能效难以兼顾。

技术现状

近年学术界提出许多采用存算一体技术的 Transformer 模型加速器，达到了较高的能效。但未能完全做到灵活性以及支持多精度。在灵活性，准确性，高能效多个维度未能达到较好的平衡点。

技术诉求

任务：基于存算一体技术的 AI 加速器架构需保证可扩展性，灵活性，高能效，多精度。

1. 实现基于存算一体技术的 LLM 加速器方案；
2. 实现基于存算一体技术的 AIGC 加速器方案。

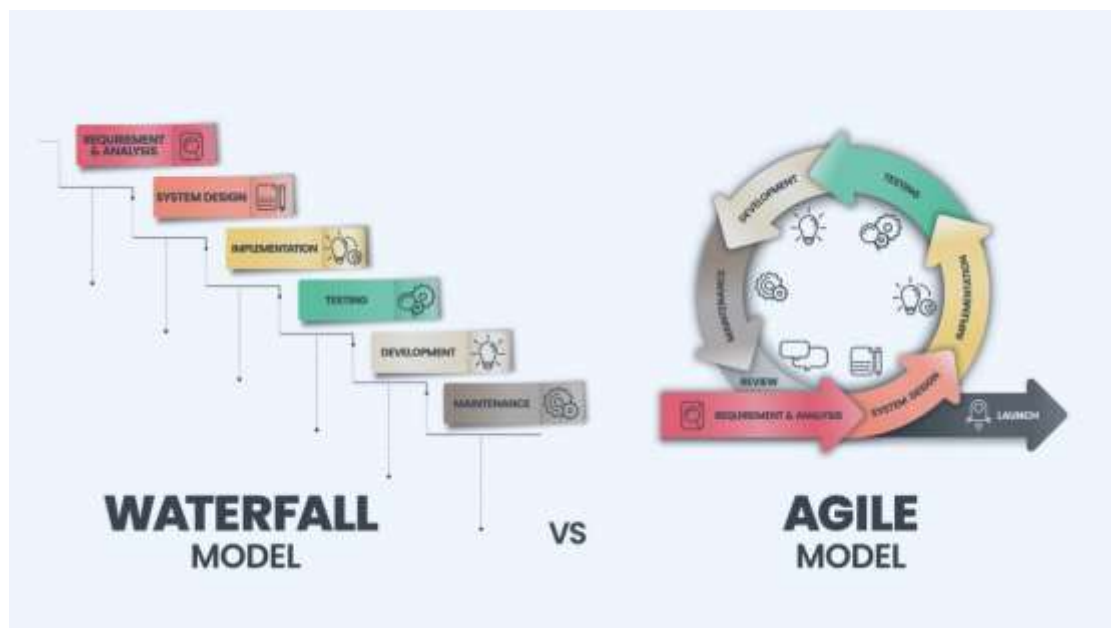
参考文献

1. Chih, Y. D., Lee, P. H., Fujiwara, H., Shih, Y. C., Lee, C. F., Naous, R., ... & Chang, T. Y. J. (2021, February). 16.4 An 89TOPS/W and 16.3 TOPS/mm² all-digital SRAM-based full-precision compute-in memory macro in 22nm for machine-learning edge applications. In 2021 IEEE International Solid-State Circuits Conference (ISSCC) (Vol. 64, pp. 252-254). IEEE.
2. Y. Yuan, Y. Yang, X. Wang, X. Li, C. Ma, Q. Chen, M. Tang, X. Wei, Z. Hou, J. Zhu et al., "34.6 A 28nm 72.12 TFLOPS/W Hybrid-Domain Outer-Product Based Floating-Point SRAM Computing-in-Memory Macro with Logarithm Bit-Width Residual ADC," in 2024 IEEE International Solid-State Circuits Conference (ISSCC), vol. 67. IEEE, 2024, pp. 576 - 578.
3. M.-E. Shih, S.-W. Hsieh, P.-Y. Tsai, M.-H. Lin, P.-K. Tsung, E.-J. Chang, J. Liang, S.-H. Chang, C.-L. Huang, Y.-Y. Nian et al., "20.1 NVE: A 3nm 23.2 TOPS/W 12b-Digital-CIM-Based Neural Engine for High-Resolution Visual-Quality Enhancement on Smart Devices," in 2024 IEEE International Solid-State Circuits Conference (ISSCC), vol. 67. IEEE, 2024, pp. 360 - 362.
4. Hager, P. A., Moons, B., Cosemans, S., Papistas, I. A., Rooseleer, B., Van Loon, J., ... & Eleftheriou, E. (2024, February). 11.3 Metis AIPU: A 12nm 15TOPS/W 209.6 TOPS SoC for Cost- and Energy-Efficient Inference at the Edge. In 2024 IEEE International Solid-State Circuits Conference (ISSCC) (Vol. 67, pp. 212-214). IEEE.

6.3 面向高能效低延迟 AI 加速器架构的敏捷开发技术方案

技术背景

随着 AI 任务场景多样化（如自动驾驶实时决策、AR/VR 动态渲染），AI 加速器设计需兼顾高能效（TOPS/W）与低延迟（<10ms），但传统硬件开发流程（RTL 设计→流片→验证）耗时长达 12-24 个月，难以快速响应算法迭代需求。据麦肯锡报告，90% 的 AI 芯片项目因开发周期过长错过市场窗口。“敏捷开发”旨在通过高层次抽象（如基于 Chisel 的硬件构造语言）、自动设计空间探索（DSE）和虚拟原型仿真，将开发周期压缩至 3-6 个月。大语言模型在 HDL 代码生成领域取得重要进展。NVIDIA 的 ChipNeMo 在工程助手聊天机器人、EDA 工具脚本生成和错误总结分析方面超越标准 LLM；RTLCoder 在 Verilog 生成任务中准确率达 70.6%；AutoChip 通过编译错误和仿真消息反馈实现 89.19% 的 Verilog 设计通过率。



核心目标：

- 快速迭代架构：支持动态调整计算单元、内存层次、互联拓扑。
- 软硬件协同验证：加速器与编译器/算法的同步优化（如 TVM+HLS 联动）。
- 参数化 IP 复用：通过模块化设计（PE 阵列、NoC 组件）复用已验证 IP 核。

关键驱动力：

- 架构异构化：需适配多种计算范式（矩阵、张量、图计算）。
- 制程逼近极限：3nm 以下工艺的物理效应复杂化，需早期预验证。
- 市场窗口缩短：AI 赛道竞争激烈，需支持“从模型到芯片”全流程快速迁移（如 Stable

Diffusion 变体快速硬件化)。

技术挑战

- **设计空间爆炸**

计算单元类型 (MAC 阵列、脉动阵列)、并行度 (SIMD 宽度)、存储层次 (共享缓存/分布式 SRAM) 的组合可能超过 10^6 种, 传统试错法无法高效收敛。

- **精度-效率的权衡验证**

高能效需采用近似计算 (如 6-bit 乘法器)、动态电压调节 (DVFS), 但功能验证和时序收敛耗时占比超 50%。

- **软硬件协同延迟优化**

编译器优化 (如算子融合) 需与硬件流水线深度耦合, 而传统瀑布式开发导致跨团队迭代缓慢。

- **虚拟原型精度与速度的矛盾**

RTL 仿真精度高但速度慢 (MHz 级), 高频模型 (Transaction-Level Modeling) 速度快但忽略关键时序细节, 难以支撑早期优化。

技术现状

有部分学术团队将敏捷开发方法学应用于 CPU 的开发, 较少 EDA 公司提供敏捷开发的解决方案, 如有提供, 一般针对特定领域或者特定架构, 缺乏灵活性和可扩展性。

技术诉求

任务: 针对 AI 加速器, 建立敏捷开发的框架, 考虑 Chisel/HLS/LLM 等快速开发语言或者 Agent。

1. 模块化 IP 生态系统, 建立参数化计算核, 标准化存储层次接口;
2. 开发 CA Model, 在可接受误差内实现 MHz 级仿真, 建立性能功耗模型。

七、软件优化技术

7.1 移动终端设备调度算法优化研究

技术背景

移动终端 SoC 存在多个器件，如 CPU，GPU，缓存，显示，影像，视频，音频等等，由于移动终端对功耗比较敏感，各个器件都会支持独立做 DVFS(Dynamic Voltage Frequency Scaling)，对于 CPU 这种器件会存在多个不同类型的 Core(性能 core 和能效 core)，如何选择使用性能 core 还是能效 core，以平衡性能和功耗，移动端的应用会同时使用多个器件才能更好的工作，如何在满足移动端应用性能体验的前提下，更好的把这些器件之间进行最佳的调度优化成为当前业界的难题。

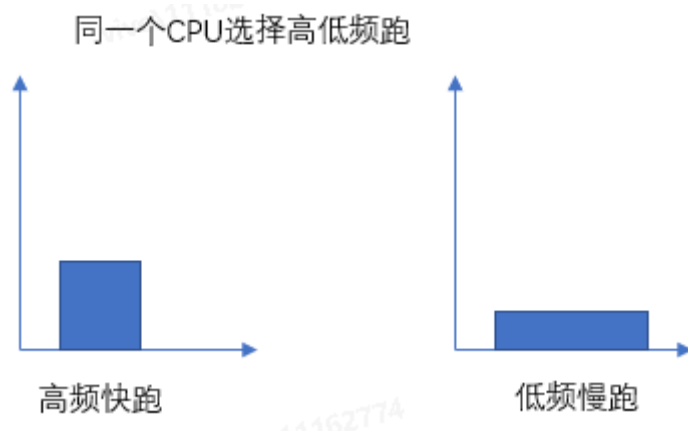
技术挑战

1、精准负载追踪调度算法研究：

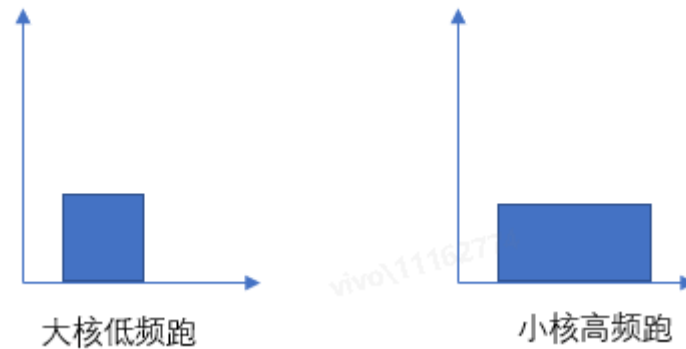
对于 CPU 器件，既做 DVFS 调度又做大小核之间的选择调度问题，不管是哪种调度，都依赖对 CPU/Task 负载的计算才能更好的决策做 DVFS 或者大小核调度，因此如何更精准的追踪 CPU/Task 的负载，从而能够给出最佳的 CPU 算力(通过 DVFS 和大小核调度选择)，兼顾系统性能和功耗的需求。

2、CPU 选核和调频调度算法研究：

CPU 多个 Core 之间如何做调度优化，包括如何选择大小核和高低频之间的最优解问题；如下示例，在满足业务性能需求的前提下，如下哪种组合调度功耗会最优：

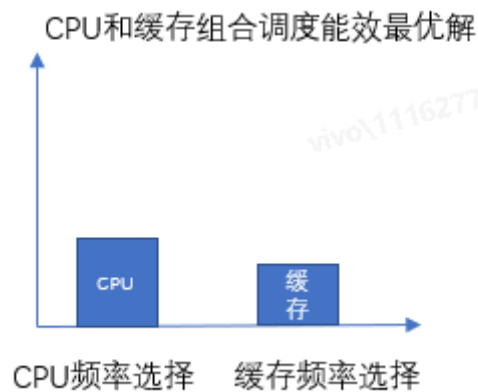


如何选择不同CPU及选择不同的频率跑



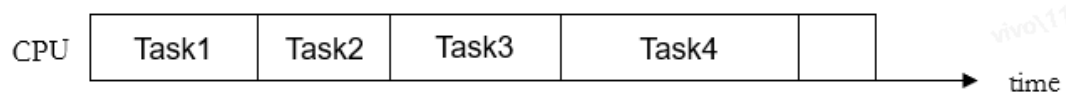
3、CPU 和各级缓存之间的联动调度算法研究：

CPU 和缓存之间如何做组合调度优化；如下所示，CPU 和缓存组合调度在满足业务性能需求的前提下，获得最优的 CPU 和缓存的功耗最优解问题；



技术现状

1、精准负载追踪调度算法研究



如上图所示，当前业界普遍使用基于 CPU/Task 的 wall time 为输入的负载计算方法，主要的算法有 2 种分别如下：

- 1) PELT (Per-Entity Load Tracking)
- 2) WALT (Windows-Assist Load Tracing)

以上 2 种算法都无法满足业务的负载跟踪需求，需要继续探索一种能够高精度的负载

跟踪算法

2、CPU 选核和调频调度算法研究

CPU 之间的调度目前倾向与小核调度及各个核的低频调度，但是发现在一些业务下，往大核或者高频调度，快速做完事情，让 CPU 下电反而能效结果更好；

3、CPU 和各级缓存之间的联动调度算法研究

CPU 和缓存直接目前没有很好的组合调度，目前主要存在 2 种方法，一种是 CPU 和缓存的频率做绑定调度，即 CPU 的频率会对应缓存的频率；另一种是利用缓存的带宽和 latency 的一些指标用于对缓存的调度，但是没有反过来对 CPU 做调度，没有达到很好的组合调度的效果。

技术诉求

1. 精准负载追踪调度算法研究

- 1)基于 task 粒度的高精度负载跟踪算法
- 2)基于 CPU 粒度的高精度负载跟踪算法

2. CPU 选核和调频调度算法研究

CPU 多个 core 之间如何组合调度，包括不同 CPU 的大小核调度及同 CPU 的高低频调度及两者之间的调度，即希望通过建模的方法指导如上如何做组合调度才能获取最优的 CPU 能效(CPU 快跑下电还是持续慢跑如何选择)，能够提供一套理论和建模的方法给出不同的组合调度方式即可。

3. CPU 和各级缓存之间的联动调度算法研究

CPU 和缓存之间的组合调度问题，如 CPU 在运行时，会存在 2 种状态，CPU 计算密集型在做大量的计算任务，或者 CPU 在等待缓存的读写数据，2 种不同的 CPU 运行状态会给出 CPU 不同的调度方式，同时也会作用到缓存的调度方式。即当 CPU 处于计算密集型时 CPU 可以给更高的算力以便及时完成计算任务，缓存可以给较低的频率；当 CPU 处于读写缓存数据时，CPU 可以给较少的算力，缓存需要更高的频率需求。可以参考以上示例，不仅局限计算密集型还是缓存密集型，能够提供一套完整的理论建模方法指导如何做 CPU 和缓存的组合调度最优解的问题。

7.2 并行分块渲染算法能效比优化预研

技术背景

Tile-Based Rendering (TBR) 架构：TBR 架构将屏幕划分为多个小的矩形区域（tile），每个 tile 逐步进行处理。这种架构的优势在于能够利用片上缓冲区（如 tile-sized on-chip buffers）来存储中间结果，从而减少对外部 DRAM 的访问，降低功耗和内存带宽需求，特别适用于移动 GPU 等低功耗和内存带宽受限的系统。

性能瓶颈与内存压力：现代 GPU 性能随着用户对图形质量需求的提升而不断提升，但增加计算单元（如着色器核心）可能导致内存子系统成为瓶颈，尤其是在处理精细几何模型和丰富纹理的场景时。TBR 架构虽然能减少内存访问，但传统方式一次只处理一个 tile，计算资源可能因 tile 工作负载不足而利用率低。

技术挑战

并行渲染中的内存压力：并行渲染多个 tile 虽然可以充分利用计算资源，但会增加对共享内存子系统的访问请求，导致内存拥塞，增加内存访问延迟，无法有效利用多线程来隐藏延迟，影响性能。

tile 间内存访问差异：屏幕中不同 tile 的内存访问需求差异较大，一些区域（如背景部分）可能内存访问需求低，而另一些区域（如有多个重叠对象的区域）内存访问需求高。这种差异使得如何合理调度 tile 成为避免内存压力的关键。

帧间一致性利用：充分利用帧间一致性来预测每个 tile 的内存压力，以便在并行渲染中合理分配资源，保证内存访问的均衡性，同时避免过多缓存未命中。

保持局部性：在并行渲染过程中，如何保持 tile 间的数据局部性（如纹理数据的局部性），以减少缓存未命中率，提高缓存效率，也是一个挑战。

技术现状

有论文结果表明：

性能提升：LIBRA 平均性能提升 20.9%，帧率（FPS）提升 11.4%。

内存延迟降低：平均内存延迟减少 13.5%。

能效改进：GPU 总能耗减少 9.2%。

局部性提升：纹理缓存命中率平均提升 10.6%。

技术诉求

并行 tile 渲染架构设计：设计一种能够有效支持多个 tile 并行渲染的 GPU 架构，包括多个 Raster Unit 的配置和协作机制。确保每个 Raster Unit 具有足够的私有资源（如 Z-Buffer 和 Color Buffer），以避免资源竞争和数据不一致问题。

智能调度算法开发：开发基于 tile 温度感知的智能调度算法，根据 tile 的内存访问需求（高温 tile 和低温 tile）合理分配到不同的 Raster Unit，以平衡内存压力。利用帧间一致性来预测下一帧的内存压力，并动态调整调度策略。

超级 tile 管理机制：设计一种超级 tile（supertile）的管理机制，将 tile 组合成更大的超级 tile，以保持纹理数据的局部性，减少缓存未命中率。同时，动态调整超级 tile 的大小，以适应不同应用和场景的需求。

动态调整策略：建立一种动态调整机制，根据实时性能监控和缓存命中率等指标，自动切换调度策略（如 Z-order 调度和温度感知调度），并动态调整超级 tile 的大小，以实现最佳的性能和能效平衡。

硬件开销优化：确保上述架构和算法的硬件开销极低，不影响 GPU 的整体功耗和性能。优化硬件设计，使得额外的存储和逻辑电路对 GPU 的面积和功耗影响最小化。

性能评估与优化：通过详细的性能评估和模拟实验，验证并行 tile 渲染架构和智能调度算法的有效性，并进一步优化其性能。分析不同应用和场景下的性能瓶颈，针对性地进行优化。

八、 声学技术

8.1 基于 AI 眼镜形态的声学泡泡算法模型开发

技术背景

在众多可穿戴设备中，AI 眼镜凭借其便捷性和实时性的特点，逐渐成为研究热点。AI 眼镜通常配备多个麦克风，这些麦克风为实现复杂的声学处理功能提供了硬件基础。

从实际应用场景来看，在嘈杂的社交场合（如餐厅、聚会场所）、多人会议环境以及开放式办公空间等场景中，人们往往面临着声音混杂的困扰，佩戴者在进行通话或者与设备交互时，难以抑制周围的人声。此时，通过利用 AI 眼镜上的多个麦克风，开发能够识别讲话人距离，并依据距离设定来实现声音筛选的声学泡泡算法模型具有重要的现实意义。该模型旨在佩戴者周围构建一个虚拟的“声学泡泡”，保留泡泡内讲话人的声音，同时有效遮掩泡泡外其他讲话人的声音，从而显著提升语音交互的质量和用户体验。



技术挑战

（一）复杂环境下的声音信号处理

实际应用场景中，声音环境极为复杂，存在多种干扰因素。不同频率、强度的声音相互交织，除了讲话人的声音外，还包含各种环境噪音（如风声、交通噪音、设备运转声）以及多个讲话人声音的相互干扰。AI 眼镜上的麦克风接收到的是混合了这些复杂信号的声音数据，如何从这些嘈杂的混合信号中准确提取出目标讲话人的声音，并精确识别其距离，是算法开发面临的一大挑战。此外，声音在传播过程中会受到反射、折射和衰减等因素的影响，尤其在室内环境中，声音会在墙壁、家具等物体表面多次反射，形成复杂的反射声，这进一步增加了声音信号处理的难度，使得准确判断声音的原始来源和距离变得更加困难。

（二）麦克风阵列的优化与校准

AI 眼镜的多个麦克风组成的麦克风阵列是实现声音距离识别和声学泡泡功能的关键

硬件部分。然而，麦克风阵列的性能受到多种因素的制约。首先，麦克风的数量、位置布局以及它们之间的距离需要进行优化设计。不同的应用场景和声学需求对麦克风阵列的布局有不同的要求，布局不合理可能导致声音采集的盲区或对某些方向声音的敏感度不足。其次，麦克风之间的校准也是一个重要问题。由于制造工艺和个体差异，每个麦克风的灵敏度、频率响应等特性可能存在细微差别，这会影响到整个麦克风阵列对声音信号采集的一致性和准确性。若不能对这些差异进行有效的校准和补偿，将严重影响算法对声音距离识别的精度，进而影响声学泡泡算法模型的性能。

（三）实时性与计算资源的平衡

声学泡泡算法模型需要实时处理麦克风采集到的大量声音数据，以实现声音的实时筛选和处理，为用户提供即时的听觉体验。这就要求算法具有高度的实时性，能够在极短的时间内完成复杂的声音信号分析、距离识别和声音增强抑制等操作。然而，实现这些功能往往需要大量的计算资源，如复杂的神经网络运算、大量的数据存储和快速的数据传输等。而 AI 眼镜作为一种可穿戴设备，其硬件资源（如处理器性能、内存容量、电池续航能力）相对有限，难以满足高强度计算的需求。因此，如何在有限的计算资源条件下，优化算法结构和计算流程，提高算法的运行效率，实现实时性与计算资源的良好平衡，是该技术开发过程中必须攻克的难题。

（四）模型的适应性与泛化能力

声学泡泡算法模型需要能够适应各种不同的应用场景和环境条件，包括不同的室内外场景（如会议室、教室、街道、公园等）、不同的声音环境特征（如嘈杂程度、声音反射特性等）以及不同用户的使用习惯和佩戴方式。然而，目前的模型在训练过程中往往受到训练数据的局限性，难以涵盖所有可能的场景和情况。这就导致模型在实际应用中，当遇到与训练环境差异较大的新场景时，可能出现性能下降、声音距离识别不准确、声学泡泡效果不理想等问题，即模型的适应性和泛化能力不足。为了解决这一问题，需要收集更加丰富多样的声音数据，采用更有效的数据增强和模型训练方法，提高模型对各种复杂场景的适应能力和泛化性能。

技术现状

1、在相关领域，已有一些类似技术的研究成果。例如，华盛顿大学开发的耳机原型利用人工智能算法实现了“声音泡泡”效果，能让佩戴者仅听到周围 3 至 6 英尺范围内的人

说话，泡泡外的声音平均降低 49 分贝。但该研究针对的是耳机设备，与 AI 眼镜在硬件结构和应用场景上存在差异。而 AI 眼镜由于其佩戴位置和形态特点，能够获取更丰富的空间声音信息，为开发更精准的声学泡泡算法提供了可能。



2、AI 在语音识别、声纹识别等相关领域的技术进步，如深度学习、监督学习方法在声纹识别中的应用，为声学泡泡算法模型的开发提供了有力的技术支撑和理论基础。

技术诉求

- 1、 基于 AI 眼镜形态，开发声学泡泡算法模型，在一定距离内保留对应的人声。
- 2、 探索不同麦克风组合形态下，声学泡泡效果受讲话者人数、距离以及各类环境噪声下的约束边界。
- 3、 鉴于 AI 眼镜硬件资源有限的现状，需要进行硬件与算法的协同优化，在效果和性能之间找到合适的平衡点。
- 4、 构建大规模多样化的声学数据集。

参考文献：

- 1、 华盛顿大学--Sound bubbles on hearables
- 2、 Neural Speech Separation Using Spatially Distributed Microphones
- 3、 IPDnet: A Universal Direct-Path IPD Estimation Network for Sound Source Localization.

8.2 移动终端设备中扬声器后腔声波驱动电池盖结构振动的抑制策略与优化设计研究

技术背景

由于智能手机追求轻薄手感、长续航、多功能，导致手机内部用于扬声器及其音腔的空间愈发极致，导致腔体太小而无法产生足够的低频。为了给用户提供更空间声和更大的音量，

目前高端智能手机大多采用了双扬声器的设计，即上下各一个扬声器，其中空间更为受限的是上扬声器，因为手机顶部布局了前后摄像头、各类传感器、主板等一系列核心器件，上扬声器及其音腔会侵占主板的空间，如图 1 所示，这导致目前上扬声器的堆叠空间越来越极致，音腔大小越来越难以保证，进而造成手机终端的音质难以进一步做好。

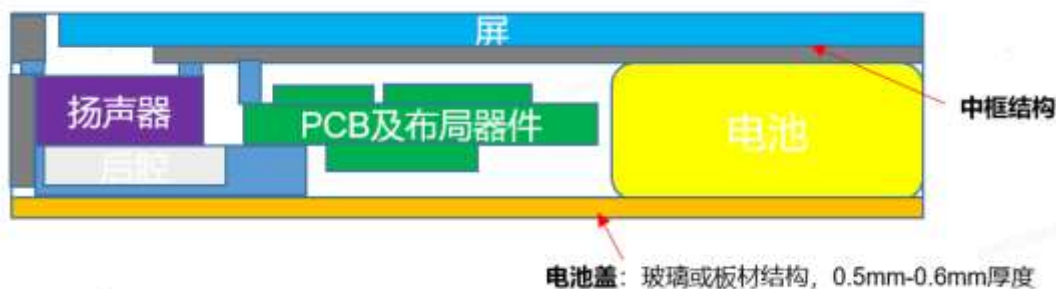


图 1 封闭后腔的上扬声器堆叠设计

在双扬没有普及的时代，手机上部只放置了一个受话器，并将手机整机内部空气间隙作为后腔，手机整机内部间隙的空间达到 5cc 左右，足够受话器发挥其低频潜力。然而，为了实现双扬立体声，我们将受话器（ $\approx 50\text{mW}$ ）替代为大功率的扬声器（ $\approx 1\text{W}$ ）时，如图 2 所示，我们发现扬声器振膜背部的声波传递到手机内部时，会带动电池盖振动，由于电池盖较薄，电池盖振动会产生很明显的震手的体验。通过检测 VOC 数据（客诉信息源），用户对于震手的客诉较多，造成很不好的体验。因此，目前我们已经采用图 1 所示的封闭后腔设计替代图 2 所示的开放后腔设计，但这样造成低频变差、整机厚度增加等后果。



图 2 开放后腔的上扬声器堆叠设计

行业内做开放式后腔都需要解决电池盖振动的问题，如图 3 是市面某上扬为开放式后腔的机型，无需封闭腔体的好处显而易见，PCB 布板面积增加，同时后腔容量大幅增加，甚至宣传低音效果。然而，我们拆机分析后发现，即便其在电池位置铺设了大量的减震泡棉，以我们测试环境测试的电池盖震感仍然很高，远远超出我们品控可接受的标准，voc 数据显示客诉“震手”投诉很多。



图 3 市面某机型的开放式设计和减震设计

技术挑战

目前采用的封闭式后腔方案震感较低，但封闭式后腔会侵占主板空间，在当前轻薄化的设计导向下，腔体大小受限，低频不足。如果按图 2 的开放式设计，可以利用整机间隙作为后腔，腔体大小足够，音质可以得到改善，但震感问题需要解决。当前工程上通过缓冲泡棉/吸音棉来减少壳体振动，但远远不足以解决问题，壳体震手的问题依然很严重。

技术现状

目前我们通过采用加速度计测量电池盖加速度最大值来检测震感大小，如图 3 所示，加速度计粘于电池盖上，然后从机器上读取加速度值。图 4 是我司不同类型的项目测得的数值，图 4 可以看出，开放式后腔结构（即利用整机间隙作后腔）的项目，在加泡棉、点胶、局部密封等常规手段改善后，震感仍然远远大于封闭式后腔的项目。



图3 测试方法

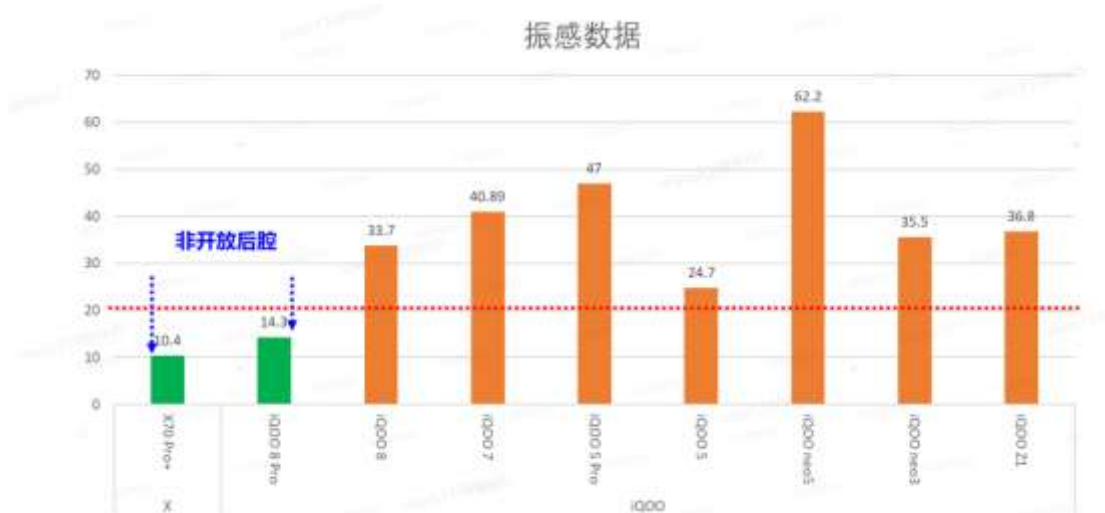
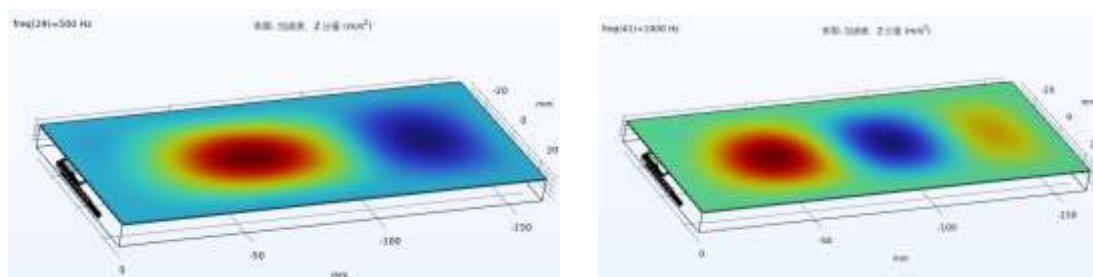


图4 相关项目的测试数值

我们建立了简易的仿真模型来分析震感，仿真结果如图5所示，可以看出电池盖材料不同频率会有模态，影响振动值（仿真模型仅供参考）。实测显示产生较大振动的频率主要在1000Hz以下。



技术诉求

任务：探索解决开放后腔带来的震感问题，助攻手机轻薄化设计和音质提升。使得采用开放后腔方案设计的手机电池盖震感大幅降低，测试最大加速度控制在 18m/s^2 以下，或相对常规设计将电池盖最大振动加速度降低 60%以上。

通过多方讨论，大致可以沿以下几个方向

1.常规方案下，引入新材料/复合材料替代原来的电池盖材料，在保证轻薄的基础上，提供新材料/复合材料设计方案，并指导我司指定加工实现的方案。电池盖整体厚度在 0.5-0.6mm。或利用压缩后不厚于 0.3mm 的辅料，通过隔绝声波/形成特定音腔来实现电池盖振动加速度降低。

2.研究吸振设计，即通过电池盖上增加附属结构，吸收电池盖振动能量转换为热能的技术。不显著增加质量，厚度控制在 1mm 以内，吸振模块面积控制在 $20*20\text{mm}$ 以内。

3.通过控制后腔声波能量来调控声波对电池盖的激励，即设计部分后腔结构，经过该段腔道后能量显著被吸收，从而在整机其它部分的声波能量大幅降低。该种设计应保证低频性能接近纯开放式后腔，避免接近完全堵塞后腔的效果。达到的效果应该类似于“声学黑洞”的效果，即通过吸收大部分声波来达到接近无限大后腔的效果，同时释放给整机内部间隙的声波能量较小。该结构可以围绕扬声器四周或背部设计，参考的扬声器尺寸在 $12*10*2.7\text{mm}$ ，或者更大尺寸的 $16*11*2.4\text{mm}$ ，相应的声波调控结构可布局在四周或背面，不超过扬声器范围 10mm。

4.引入声学超材料/结构超材料，在保证轻薄化的基础上，通过调控声波/结构振动来有效抑制电池盖振动，实现电池盖振动加速度的降低。

5.其它通过调节声固耦合机制实现的轻薄化设计。

备注：

（1）以上关于改进型材料/技术，需提供技术创新说明及理论支撑，以区别于现有常规技术方案；

（2）以上技术需求，为期望水平，如有未达到上述指标的技术研究，但较目前行业水平仍有一定创新性的提升，也可通过后续交流确认具体技术价值，共同协商开发目标。

九、 机器人控制技术

9.1 基于 MR 遥操作的小样本动作学习及任务泛化能力研究

技术背景

近年来，机器人通用策略研究在视觉表示学习与视频生成模型推动下取得显著进展。基于视觉的机器人策略（如 RT-1、Diffusion Policy）通过自监督学习从像素输入中提取语义与空间特征，但在动态场景理解与长时任务规划中仍面临挑战。

视频扩散模型（Video Diffusion Models, VDMs）通过建模连续帧的物理动态，展现出对复杂场景的预测能力。研究表明，将视频预测模型的隐式动态知识迁移至机器人策略（如 VPP 方法），可通过预测未来视觉表示指导动作生成，显著提升多任务泛化性。然而，现有方法依赖固定视角的机器人操作数据，缺乏对真实用户交互场景的适应性，且在家庭清洁等任务中难以实现小样本下的语义规划。

技术挑战

1. 数据瓶颈：现有机器人训练数据多依赖固定摄像头或仿真环境，缺乏真实用户第一视角的动态交互数据（如手眼协调、桌面整理动作），导致策略难以泛化至复杂家庭场景。
2. 动态场景建模不足：传统视觉编码器仅提取当前帧特征，忽略时序动态信息，难以预测长时任务中的物体状态变化（如餐具叠放、水流冲洗过程）。
3. 小样本适应能力弱：家庭任务需快速适应新物体、新布局，但现有扩散策略依赖大规模标注数据，对少量演示样本的利用效率低。
4. 可解释性缺失：基于端到端黑箱模型的决策缺乏透明性，难以满足家庭场景中安全性与用户信任需求。

技术现状

前沿研究（如 VPP、GR-1）已验证视频预测模型在机器人策略中的潜力：

- 预测表示学习：通过视频扩散模型（如 Stable Video Diffusion）的隐式特征，编码未来物体运动轨迹与机器人动作关联性，实现高成功率的多任务控制（MetaWorld 基准任务成功率 68.2%）。
- 高效逆动力学建模：结合单步前向预测与扩散策略，在真实机器人平台（Franka

Panda、Dexterous Hand）中实现 30 多种复杂任务的闭环控制。

- 多模态对齐：利用 CLIP 语言特征指导视频生成，初步验证语义指令与动作规划的协同能力。

然而，现有方法在家庭场景中仍存在局限：数据采集依赖机器人本体传感器，缺乏用户视角动作捕捉；语义规划框架未针对清洁任务优化（如物体分类、任务分解）。

技术诉求

为突破上述瓶颈，提出以下合作研发方向：

1. 第一视角动态数据采集与训练集构建

数据采集：基于 AR/MR 眼镜眼镜，采集用户执行家庭任务（桌面整理、洗碗）的第一视角视频流、手部动作与眼动数据，构建多模态数据集（RGB-D、IMU、语音指令）。

数据增强：结合视频扩散模型的场景生成能力（如 Stable Video Diffusion），合成多样化家庭环境（厨房布局、餐具类型）与异常工况（水流干扰、物体遮挡），提升数据覆盖性。

2. 小样本驱动的可解释语义规划框架

动态表示学习：通过视频扩散模型的单步预测特征、粗糙的未来轨迹，提取任务相关的关键语义标记（如“待清洗餐具位置”“桌面收拾物体的位置”），构建可解释的状态空间。

分层扩散策略：

- 高层语义规划：例如基于 CLIP（Contrastive Language-Image Pre-training）与视觉语言模型（如 Video-LLaMA）等方法，将用户指令（如“清理桌面”）分解为原子动作序列（抓取→分类→放置）。
- 底层动作生成：例如采用扩散 Transformer（DiT）建模动作-状态关联等方法，通过小样本微调适应新工具（洗碗刷）与新物体（异形餐具）。

可解释性验证：设计注意力可视化模块，追踪扩散策略头在动作生成中对语义标记的依赖关系（如“抓取勺子”决策关联未来水流预测）。

3. 家庭场景清洁任务验证

用户戴上标准训练眼镜，做若干次(<10)第一视角示意操作的视频教学后，机器人/臂可以根据训练模型完成清理桌面，洗碗、洗盘子并分类整理的动作。成功率>90%

- 场景：桌面整理（分类杂物→回收垃圾→收纳物品）、洗碗（抓取→冲洗→摆放），等 BEHAVIOR-1K 数据集中定义的任务。

- 嵌入式部署优化：
 - 开发手机级 ARM 平台（如骁龙 8 Gen3）的推理引擎，实现 30ms 内轨迹生成与 >20Hz 闭环控制。
 - 设计自适应碰撞点密度算法，根据机械臂速度动态调整感知分辨率，算力占用 <1TOPS。

9.2 跨平台融合感知框架

技术背景

XR 头显、智能眼镜以及机器人等前沿技术的发展，正引领我们步入一个以多模态感知为导引，藉由智能交互所驱动的沉浸式世界，而所有这些设备的核心功能都高度依赖于其对周围环境的准确感知与理解。当前不同平台上的感知系统受限于现有架构的限制，缺乏统一的架构和算法范式，这不仅限制了数据和算法的复用性，也阻碍了设备间协同工作的潜力。

因此，“跨平台感知融合”成为当前亟待解决的关键问题，旨在构建一个统一的、高效的感知框架，以支持不同智能设备在复杂环境中实现鲁棒、精准的感知能力。此外，针对不同器件设备以及系统特性，也应充分发挥各平台的优势，考察在多设备协作场景下的相互支援，达成优势互补，从而实现复合应用场景下用户体验和系统性能方面的提升。

技术挑战

平台差异

传感器配置差异

- XR 头显：通常配备鱼眼跟踪相机、高分辨率 RGB 相机、IMU，结构光或 TOF 深度传感器、眼动追踪相机或传感器、麦克风阵列、环境光传感器
- 智能眼镜：往往追求轻量化和低功耗，传感器通常更精简，可能只有 RGB 摄像头、IMU，甚至部分产品仅依赖一个或两个小型传感器
- 机器人：传感器配置最为多样，从简单的轮式里程计、IMU，到复杂的激光雷达（2D/3D）、毫米波雷达、超声波传感器、多视角摄像头、力传感器、编码器等

计算能力差异

- XR 头显：通常搭载高性能移动 SoC（如高通 XR2 Gen2+），拥有强大的 AI 加速能力

- 智能眼镜：计算能力受限于电池续航和散热，通常为低功耗芯片（如高通 AR1）
- 机器人：从嵌入式 MCU 到高性能 GPU 再到定制芯片，计算能力跨度较大

潜在方向

- 如何设计一个通用的感知框架，能够适应如此广泛的传感器类型和计算能力
- 在框架设计上如何对抽象出可配置的多模态输入输出，以支援不同的设备组合
- 推理模型或算法模块如何应对不同的输入输出组合
- 如何针对不同平台的特性分配更具适应性的任务或模块...

数据同步

潜在方向

- 如何处理同设备传感器同步：不同厂商、不同型号的传感器输出数据格式各异（例如，深度数据是 Z 轴距离还是 X/Y/Z 点云坐标，图像分辨率和畸变参数不同）、数据帧率各异，如何设计统一、自动化、鲁棒的校准方法

- 如何处理跨设备数据同步：不同设备各自有自己本地的时间系统，当接入同一个协作网络时，如何进行无线连接下的时间系统同步；当传输出现波动时，如何实现重新对准以及实时修正

- 高帧率传感器数据传输中面临的动态编解码算法...

跨设备融合推理

感知范围差异

- XR 头显：主视场与人眼感知范围基本相当，跟踪视场视相机布局有所不同，一般可覆盖到水平 220°以及俯仰 120°的水平。深度传感器典型 FOV

- 智能眼镜：器件配置更为有限，总体小于 XR 头显的感知范围，特别是彩色相机视场

- 机器人：不同系统的差异性较大，除常见的头部主相机类似 XR 头显外，会额外配置相当数量的激光雷达、测距传感器等其他设备，相对前两者而言能够提供更大范围的全景感知能力

感知能力倾向

- XR 头显：主打影音呈现，主要感知对象是用户自身状态（位姿、动作、操作...），对外界的环境感知以定位跟踪为主，附带一定空间理解特性

- 智能眼镜：主打智能体验，主要感知对象是在用户个人信息上下文和环境语义基础

上的意图理解

- 机器人：主打自主驱动，主要感知对象是本机状态和周遭环境，经过推理后形成任务决策和与之相对应的控制流

系统融合层级

- 决策层：各设备完成自己的感知任务，然后将结果交由一个“融合器”完成结果的处理并映射回各设备自己的空间

- 特征层：各设备仅完成本地特征的处理，然后依照不同任务将特征信息进行跨设备共享，各设备在获取到补充的额外特征后再完成推理

- 数据层：设备间直接进行原始数据的分享，利用更丰富的数据进行感知推理

端云结合

- 决策融合：譬如，端侧完成子任务推理，云端进行融合后下发处理结果

- 特征融合：譬如，端侧完成特征提取和上传，云端进行综合推理后下发结果

- 数据融合：譬如，端侧仅上传数据流，云端完成所有推理任务

潜在方向

- 依据任务不同对“设备/特征/数据”进行动态组合，在满足精度需求的前提下挑选或决策最小代价的组合以及与之对应的模型结构

- 不同设备平台但同源架构下的信息耦合，以及相应的轻量化模型设计

- 跨设备公共特征空间的构建与融合，以及在此基础上的多模态多任务推理

- 仿真数据集构建与评估体系

技术诉求

如何设计一套通用框架赋能不同设备平台

设备平台间的差异如何“抹平”

多设备间如何更好的实现相互支援和协作