



图1 帧外插与补全系统流程图

本项目旨在构建一套面向移动端异构硬件（GPU+NPU）的高效、稳定、无传统光流依赖的实时帧外插与空洞补全系统。通过充分利用移动端渲染管线中已有的 MV 信息，有效避免了计算稠密光流的复杂性和高开。整体方案包含以下四个清晰界定的模块：

1. 基于渲染侧 MV 的前向时间重建模块

- 双向投影结构：以上一帧(t-1)和当前帧(t)的渲染侧MV为输入，直接进行未来帧(t+1)的前向投影，利用低分辨率深度缓冲 (Z-buffer)检测遮挡区域，生成时间预计小于 1ms，采用最近帧优先原则消除像素冲突。
- 三通道多域输入：Warp操作输出RGB颜色、遮挡掩码(Occlusion Mask)以及可选的投影残差(Residual MV，用于少量边缘区域校正)。

2. 空洞检测与 Patch-based 局部补全模块

- 简化空洞检测策略：利用MV长度差异与深度跳变，仅区分静态遮挡和动态遮挡两类空洞，降低了对物体实例信息的依赖。
- 局部Patch修复网络：设计轻量级CNN，以空洞局部的RGB像素块、遮挡掩码和为输入，输出纹理和几何局部一致的补全结果，无需跨帧轨迹估计，保持计算效率。

3. 内外插融合与级联式多帧插值模块

- 内外插融合：在帧外插基础上，引入帧内插并以Cascade推理方式实现。优先外插(t+1, t+2)，然后在空余时段逐步补推中间帧(t+0.5, t+1.5)。
- 容错与延迟控制策略：明确设定NPU利用率、温度与电量的阈值，极端条件下降级为单步外插，以确保系统的稳定性与实时性。

4. 轻量级跨帧色彩稳定模块

- 局部光照估计：利用Warp结果中的RGB与深度生成局部光照特征确保模块独立性。

- LUT调色网络：采用小型CNN网络生成低维色彩查找表，以应对动态光照下的非线性色彩变化。
- 时域卷积约束：使用轻量级卷积，实时保持连续帧的亮度与色彩稳定，抑制帧间闪动感。

此外，为进一步优化异构硬件的部署效率，提出了以下**部署策略**：

- GPU+NPU双线程解耦：GPU负责Warp与遮挡计算，基于NPU的插帧子线程以双缓冲方式完成 Warp → 检测 → 修复 → 调色流程，避免主线程阻塞；
- 参数约束与网络量化：整体网络参数控制在3M以内，支持INT8对称量化，兼容移动端NPU。
- 动态精度切换：在设备温度超过设定温度（比如45℃）或电量过低（比如低于15%）时，自动降低插帧精度或减少插帧数量，维持用户体验与设备稳定性的平衡。

（二）技术路线

项目将采用“算法开发—模型压缩—异构部署—场景评测”四阶段技术路线：

- 算法阶段：在桌面端完成 Warp + 修复网络的设计与训练，形成高保真插帧框架；
- 模型轻量化：采用量化、结构裁剪与知识蒸馏手段，压缩模型规模并适配移动端 NPU；
- 平台部署与调优：项目将在移动端平台上完成异构调度与 AI 插帧模块的高效部署与推理调优；
- 场景评测与优化：选择多个主流高帧率游戏与典型交互场景，进行画质、功耗、稳定性全维度评估。

（三）技术创新点

本项目针对移动端实时插帧与帧外插技术的瓶颈，对比上一代基于PC端的帧外插方法进行了如下五项关键创新：

1. 低分辨率深度缓冲重建：

相比传统全分辨率深度缓冲生成方法，本项目首次提出利用低分辨率（例如1/4或1/2分辨率）的深度缓冲，通过结合深度引导的空间超分算法，在极大降低GPU显存带宽占用与片元填充开销的同时，仍可保证外插帧的空间一致性和精度，有效缓解了移动端性能瓶颈。

2. 单次推理多帧外插：

相比上一代逐帧单独推理的方法，本项目提出一种一次网络推理同时外插多帧的框架设计，显著降低了推理开销与计算冗余，在同等硬件条件下显著提升了帧率，同时有效降低移动设备功耗。

3. 内外插混合帧生成机制：

与传统单一内插或外插方案相比，本项目设计了一种内插与外插交替组合的混合方案，既有效降低了整体延迟，又保障了插帧画面质量和时序一致性，特别适合对时延敏感的游戏与VR场景。

4. 移动端NPU裁剪蒸馏优化：

上一代模型架构通常较大且未针对移动端NPU优化，本项目首次将模型结构裁剪与知识蒸馏技术应用于插帧网络，生成了极致精简的网络架构，可高效部署于移动端NPU，实现极低的推理延迟和小体积模型，满足手机端低功耗实时运行需求。

5. 轻量级跨帧色彩稳定模块：

针对前代方案普遍存在的跨帧色彩漂移与不稳定问题，本项目创新性地提出了一种极轻量的跨帧色彩稳定模块，利用全局色调映射网络与时间一致性正则，有效提升了连续帧之间色彩的一致性和稳定性。

综合上述创新技术，本项目在保证画面质量的同时，可有效降低移动端插帧算法的延迟、算力需求和功耗，适合高端移动游戏、VR和低延迟流媒体等应用场景。