



南京邮电大学
Nanjing University of Posts and Telecommunications

基于潜在扩散模型的虚拟视点生成

Virtual viewpoint generation based on latent diffusion model

汇报人：方佳璇



目录

Contents

01

选题背景及意义

02

研究内容与思路

03

研究方法与实验

04

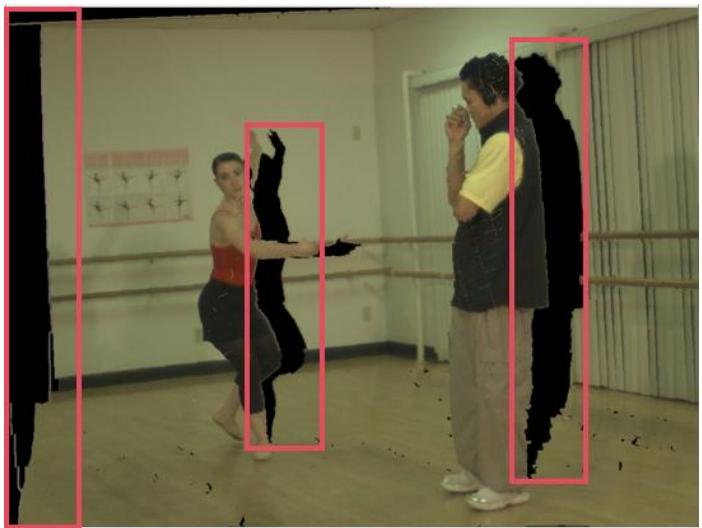
研究总结与展望



01

选题背景及意义

RESEARCH BACKGROUND AND SIGNIFICANCE



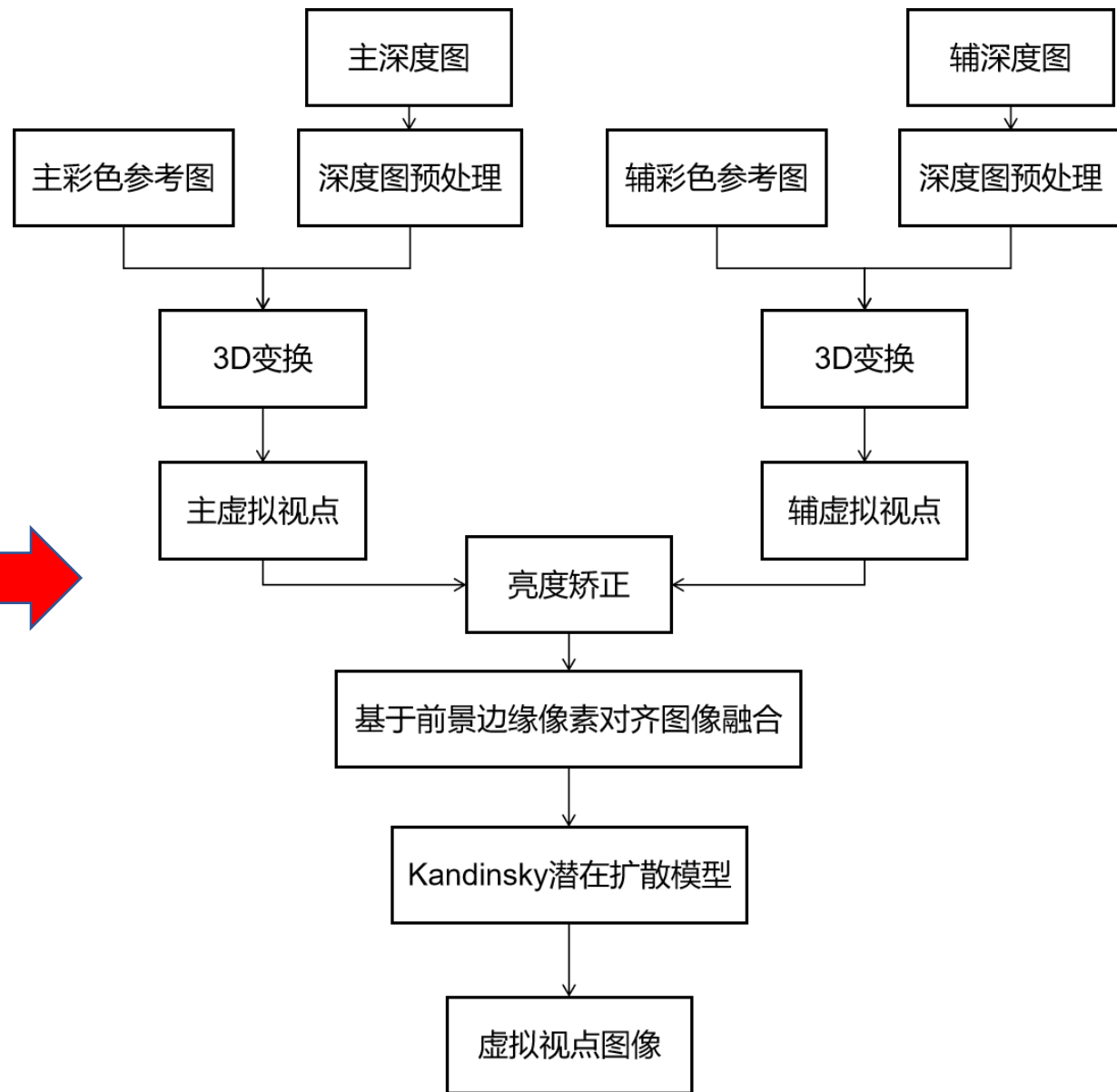
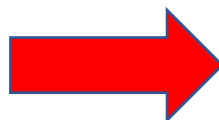
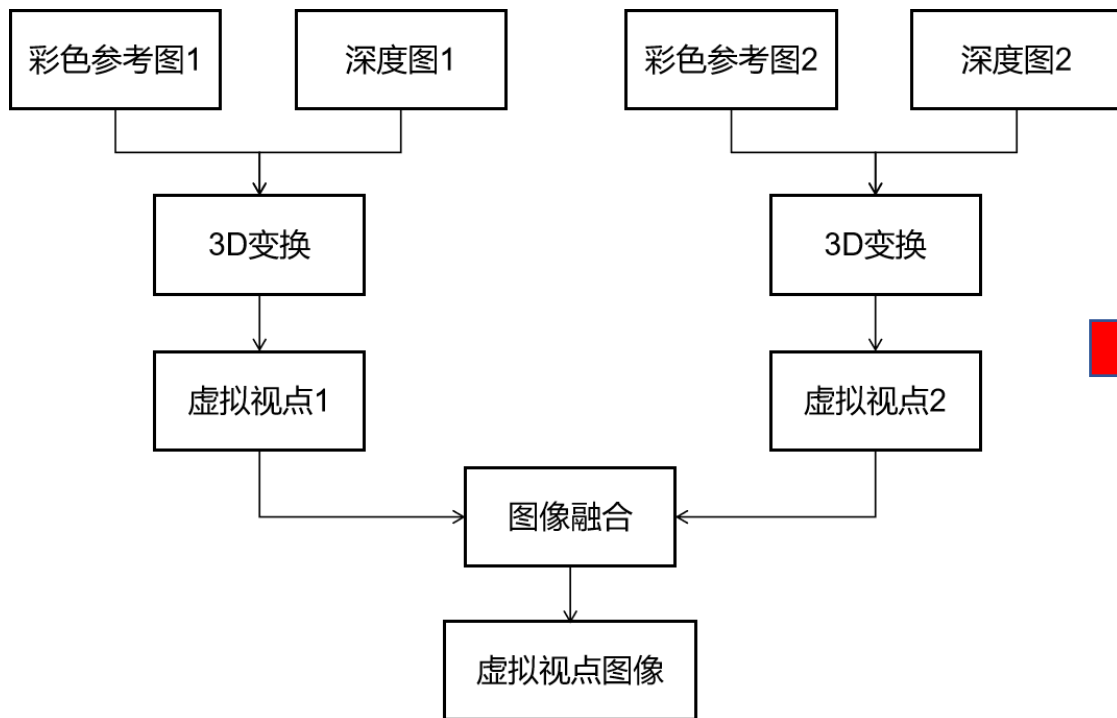
自由视点视频是一种可交互视频播放模式，用户可以选择视频播放位置和角度。为实现这种效果需要采集视频场景中大量的视点信息，这会导致成本过高。因此，自由视点视频可采用基于深度图绘制(Depth Image Based Rendering, DIBR)来减少视点信息的采集。但是，通过DIBR技术从已知视点合成新视点时会出现裂缝、伪影、空洞等问题，导致虚拟视点图像质量不佳，造成用户体验质量下降。因此，本课题针对基于深度图绘制中存在问题展开研究，旨在提升虚拟视点质量。



02

研究内容与思路

RESEARCH CONTENT AND IDEAS





03

研究方法与实践

RESEARCH METHODS AND EXPERIMENTS

1、深度图预处理

OSTU阈值算法根据图像的灰度分布确定一个最优的阈值，将图像分为前景和背景两部分。将深度图作为输入图像，通过OSTU阈值算法计算出一个最佳的阈值 T 。 T 的阈值公式如下所示

$$T = \text{Ostu}(P(x, y)) - C$$

在确定了阈值 T 后，计算膨胀后的深度图与原始深度图之间的差值，通过比较深度值差与阈值 T 的大小，即可求出前景边缘像素。

$$E(x, y) = \begin{cases} 1 & p(x, y) \oplus k - p(x, y) > T \\ 0 & p(x, y) \oplus k - p(x, y) \leq T \end{cases}$$

$$k = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

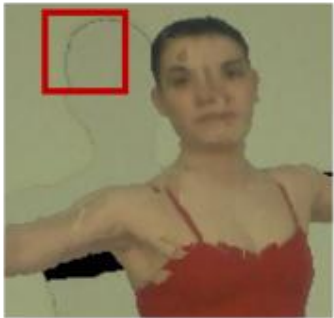
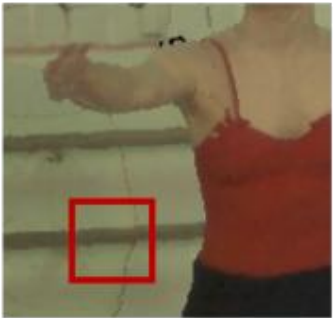
对于每个前景边缘像素，查找其四邻域像素的深度值，并将最大的深度值赋予给前景边缘像素以实现深度矫正。

$$E_{final}(x, y) = \begin{cases} \text{Max}(F(p(x, y))) & E(x, y) = 1 \\ p(x, y) & E(x, y) = 0 \end{cases}$$

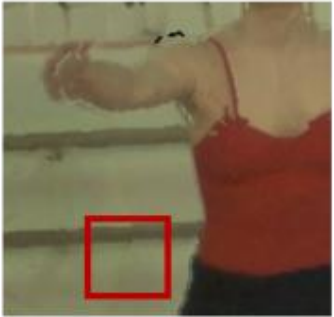


1、深度图预处理

矫正前



矫正后



第0帧

第33帧

第85帧

第45帧

第83帧

第99帧

2、亮度矫正

在Lab颜色空间中，分别分离出两个虚拟视点图像的 L 、 a 和 b 通道。然后找出每个通道中的有效像素，即非空洞区域，计算有效像素的数量和每个通道的平均值

$$\bar{L} = \frac{1}{N} \sum L(x, y), (x, y) \subseteq A$$

根据两个虚拟视点图像各通道的平均值，计算出每个通道的整体平均值

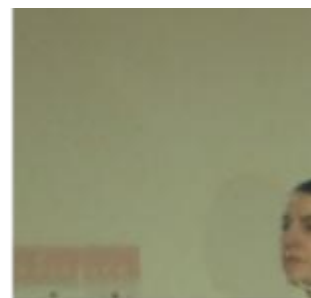
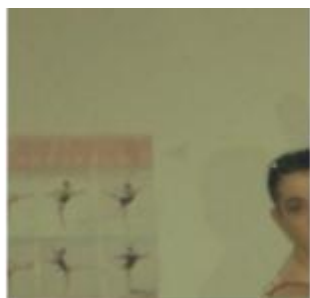
$$\bar{L}_{m,a} = \frac{\bar{L}_m + \bar{L}_a}{2}$$

然后，对于每个虚拟视点图像的有效像素，根据整体的平均值和该图像通道的平均值的比例关系，对该像素的通道值进行调整

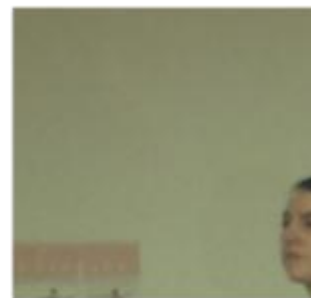
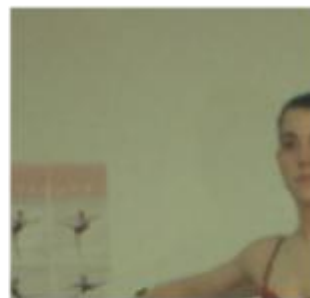
$$\bar{L}_{correct-m} = \frac{\bar{L}_{m,a}}{\bar{L}_m} L_m(x, y), (x, y) \subseteq A$$

2、亮度矫正

亮度矫正前



亮度矫正后



第0帧

第16帧

第69帧

第85帧

第99帧

3、基于前景边缘对齐的视图融合方法

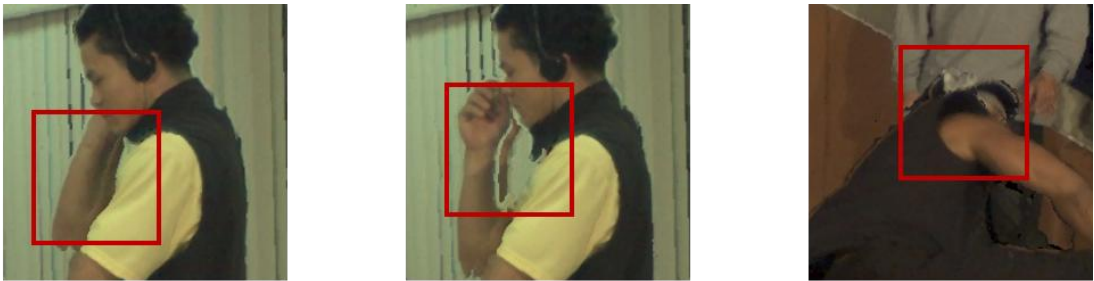
由深度图处理可得到前景边缘掩模图，标记为1。然后对两个虚拟视点图像的深度图进行对齐。通过计算两个深度图的前景边缘，使用估计误差找到最佳的位移量，使得两个深度图的前景边缘尽可能地对齐。误差公式如下

$$E = \sum_{i=0}^{r-1} \sum_{j=0}^{c-1} |p_l(x, y) - p_{move}(x, y)|$$

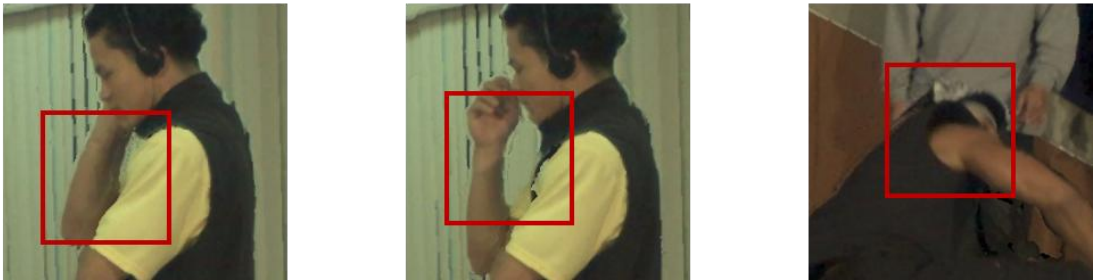
在完成深度图和彩色图像的对齐后，采用一种简单的视图融合策略。如下所示

$$p(x, y) = \begin{cases} p_a(x, y) & p_m(x, y) = 0 \\ p_m(x, y) & p_m(x, y) \neq 0 \\ 0 & p_m(x, y) = 0 \& p_a(x, y) = 0 \end{cases}$$

改进前



改进后



3、基于前景边缘对齐的视图融合方法

第49帧



第69帧



第85帧



第96帧



(a)双向DIBR

(b)改进后的双向DIBR

(c)原图

3、基于前景边缘对齐的视图融合方法

第33帧



第45帧



第83帧



第92帧



(a)双向DIBR

(b)改进后的双向DIBR

(c)原图

4、基于Kandinsky潜在扩散模型的空洞填充

例：

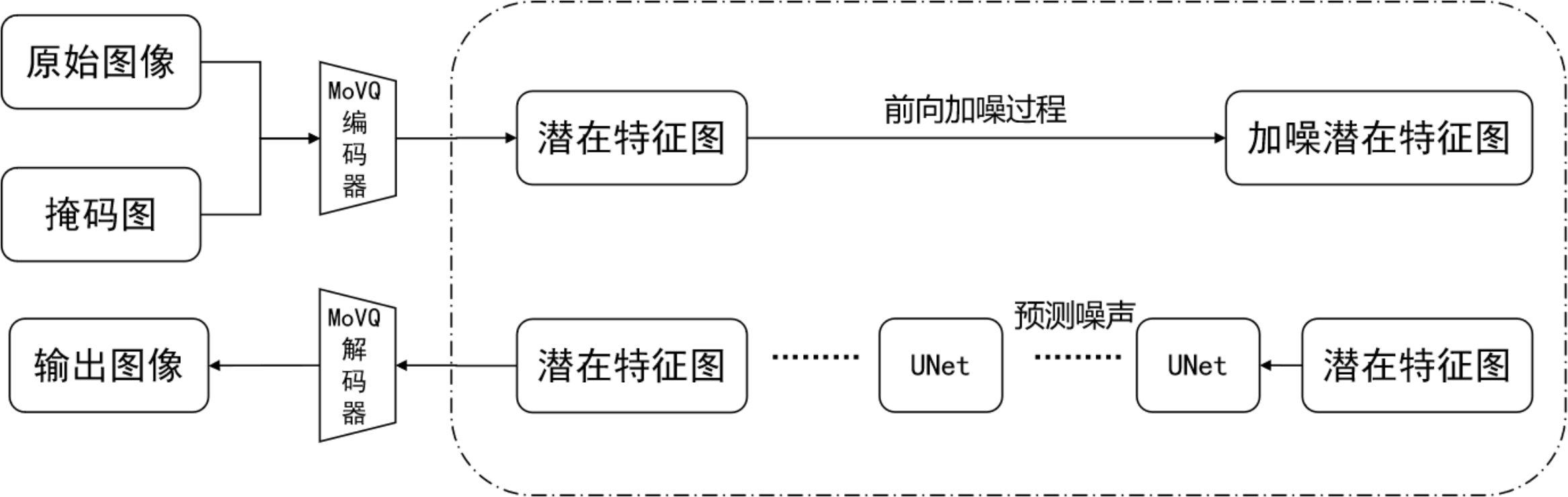


合成后的虚拟视图



掩码图

4、基于Kandinsky潜在扩散模型的空洞填充



4、基于潜在扩散模型的空洞填充

(a) 芭蕾修复图像



(b) 芭蕾原图



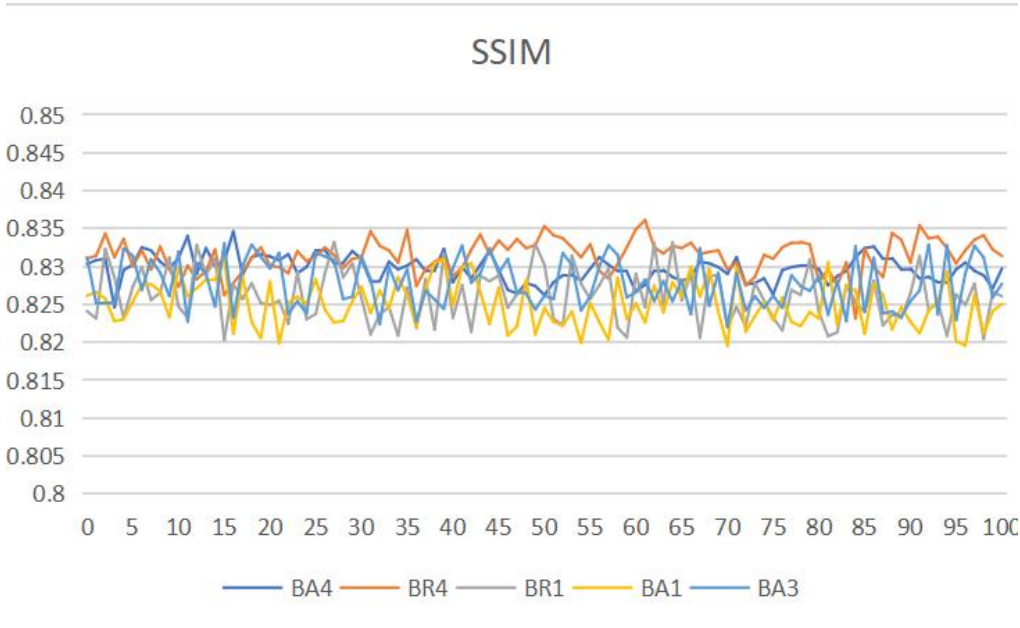
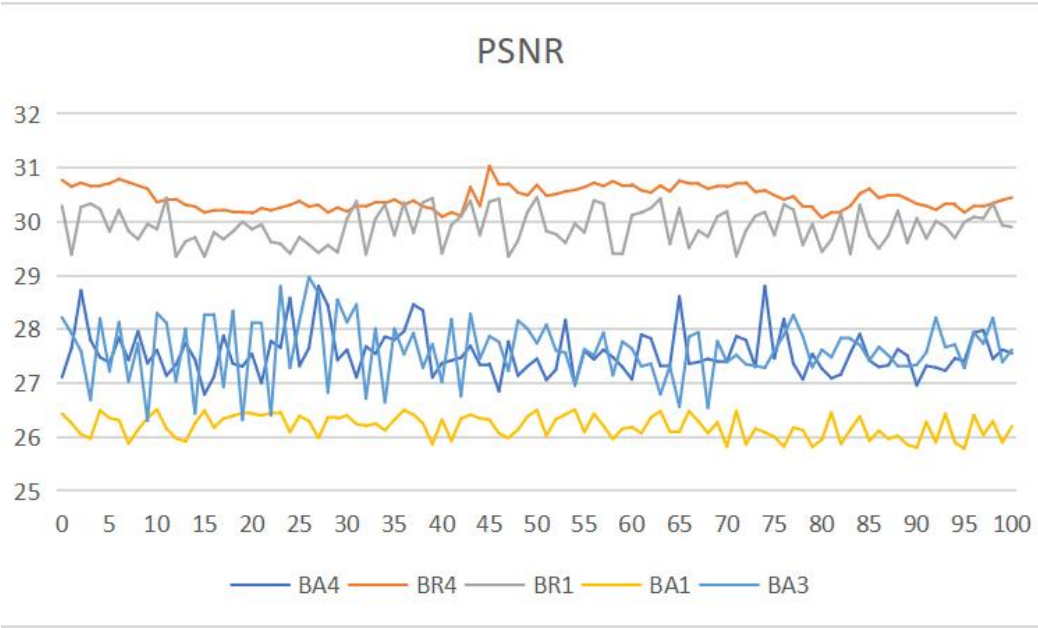
(c) 街舞修复图像



(d) 街舞原图



4、基于潜在扩散模型的空洞填充





04

研究总结与展望

RESEARCH SUMMARY AND PROSPECTS

(1) 运用OSTU阈值分割以及形态学处理来提取前景边缘像素，并且对前景边缘像素进行深度矫正，让虚拟视点图像在边缘区域有更好的结构连续性以及深度一致性，减少伪影的出现。

(2) 针对主虚拟视图和辅虚拟视图亮度差异导致融合不自然的问题，本文采用了一种基于Lab颜色空间的亮度矫正，依靠对虚拟视点图像亮度通道的处理，提高了合成图像在视觉上的亮度一致性，有效缓解了交界处的突变问题。

(3) 在图像融合方面，本文设计了一种基于前景边缘像素对齐的双向DIBR技术，借助计算前景边缘像素位置的差异值，预测最优像素偏移量以实现图像对齐，在避免结构冲突的保留了前景纹理的清晰度和完整性。

(4) 在剩余空洞的修复部分，本文使用Kandinsky潜在扩散模型用于空洞填充任务，凭借生成虚拟视图的空洞掩模图并结合潜在扩散模型，实现了对复杂结构和纹理细节的修复。



1、拓展用于时序连续帧的视点生成任务

在实际的FVV系统中，连续多帧生成的时间一致性对视觉连贯性至关重要。目前已有此类研究，后续研究可继续改进时间维度建模方法，从而增强生成视图间的帧间一致性。



2、增强模型对复杂场景的适应能力

本文所使用的数据集场景相对理想，未来可以将模型部署到更复杂的真实场景中，测试其对动态物体、非刚体变形、光照变化等复杂条件下的鲁棒性表现。



南京邮电大学
Nanjing University of Posts and Telecommunications

感谢倾听