# 会议纪要

**会议主题：** 基于扩散流复用的视频生成技术

**会议时间：** 2024年12月16日19：00至20：00

**会议地点：** 西安电子科技大学北校区会议中心303-1会议室

**线下人员：**承楠孙瑞锦 尹志胜 贺靖超 王葳 沈京龙 王秀程 马龙飞 周新阳 权赟昊 张玉洁 胡陆莹 齐阁 邓川 孙兴栋 方忠盛 贾宏刚 郑雯馨 祝馨平 贾昊燏 张智杰 杨杰 邱子仪 刘永红 万佳林 韩松明 刘苏 尚佳瑶 陈梦豪 郑佩林 孙路路 王天宇 于凡迪 尉家豪 杨双宇 傅连浩 郭译凡 侯毓真 李青壮 许嘉洁 张岳 朱煜朋

**线上人员：**赵璇、黄蕾

**记 录 人：**陈梦豪

**会议内容：**

1. **论文分享：**

介绍了一种名为"扩散复用运动"的网络Dr. Mo，以加速潜在视频的生成。Dr. Mo架构能够显著加快视频任务中的扩散模型速度，同时提高视觉质量。

**王秀程**：radio map推演中，有视频生成任务，该技术可以提高生成视频帧的速度和精度

* **承楠：**这个技术提出的背景是什么，应用在什么领域

**韩松明：**应用在视频生成领域

* **承楠：**连续的十几帧都是通过去噪生成的吗

**韩松明：**如果没有这个技术的话就要连续去噪生成每一帧

* **承楠：**两帧都去噪的话怎么保证两帧是一个连续的画面呢，去噪怎么实现视频生成的

**韩松明：**用到了前面几帧的图片信息生成后面

**沈京龙：**现有常见的生成视频技术是对四维视频张量直接做去噪生成视频

**王秀程：**这个技术必须十几帧同时生成，现有的可灵模型可生成一个小时以上的视频

* **承楠：**AI生成的图片为什么一眼就能看出来

**王秀程：**数据集问题，好的数据集生成的是看不出来的

* **承楠：**人的照片一眼就能看出来，特别是人脸

**王秀程：**生成的人的手是不真实的，人脸看着真实，更多的还是靠看手指来判断是否是AI生成的

* **承楠：**radiodiff可以和这个技术结合吗

**王秀程：**可以结合，这个用的也是一种预训练的模型

* **承楠：**radiodiff参数也是提前训练好的吗，可以改到这个技术上吗

**王秀程：**没有代码很难复现

* **孙瑞锦：**这个是国外的工作吗

**韩松明：**一作是复旦，后面的作者是国外的

* **承楠：**为什么取Dr,Mo这个名字

**王秀程：是**difussion几个英语的缩写

* **承楠：**步数越大越靠前是吗，为什么步数越大计算的内容更多了，但是效果还会更差

呢，这个地方值得思考。这说明这个工作是有一个最优的计算的点，不是说越大越好。

**韩松明：**这个技术不是说t越小越好

* **承楠**：信息论和AI的结合现在人还不多，后续靠你们

**王秀程**：他的这个工作做的是相当复杂，但是其实不用这么复杂，有一些工作是为了凑篇幅加上去的

**权赟浩：**denoise那一步的step步数不是越高越好，在论文中有没有对计算效率和转换效率之间的关系进行具体研究。

**韩松明**：没有进行研究，直接取了一个

**郭译凡：**整个思路和Flash attention完全相反，Flash attention的思路是放弃一些中间变量，用计算资源代替存储资源。那这个技术的中间变量的存储是放在显存还是内存，如果存在内存中的话计算效率还是有影响的，从内存中取中间变量是影响计算效率的，说明这个技术相当于用存储资源换计算资源。

* **承楠：**小而专的论文最好是在小的论文分享会讲，组会更倾向于讲一些大家正在做或

者做过的工作，大家可以互相了解一下正在做什么，提一些意见。如果没有现有的工作再做论文的分享。安排下周王秀程进行汇报，下下周马龙飞和傅连浩进行汇报。

**二、承楠老师关于就业的一些建议、看法**

1、找工作的同学，形式不好，没有找到工作的先找一个保底的工作。

2、一定要提高自身的认知水平，提高认知的过程就是不断地接收外部刺激，并对该刺激进行思考。前些年用土地杠杆撬动城市发展，大家都可以得到好的工作。近几年就业的问题客观存在，需要对其进行应对。学历不代表一定可以找到工作，还是需要提高个人的能力。

3、企业招聘时如何选择：企业内部也在靠AI加速代码编写等工作，会写代码不代表有更大的优势。能够发现问题并抓住问题关键点进行解决才是重要的，只会写代码或者会一些算法是不行的。在科研中要培养自己解决问题的能力，主动发现研究热点并快速将思路形成工作。不能眼高手低，在做事情的时候去主动思考，放低期待，提高能力，深挖某个方向一定会有好的结果。

4、实验室已经提供了很好的科研环境，后续会和企业和研究所多建立项目，增加定向就业的可能性。

5、幸福的生活不一定只追求金钱和高薪的工作，高薪工作岗位本身不是特别多。现在二十多岁，后面的机会还有很多。人一辈子不一定只会从事一个工作，至少会切换一个行业或事情，从事目前工作的同时要观察下一个可切入的工作。

6、好的方向有几个特点：首先是行业足够大，不会被前面的人做完后面的人就没有办法再进入这个行业；其次要能够对上一个技术容纳，有包容性。当前来看AI会是未来比较好的一个方向。但是也需要注意本身的能力能不能够支撑自己从一个行业切换到另一个行业，部分工作不能够提升能力。

**三、实验室规章制度**

承楠老师对实验室纪律、毕业要求、补助等情况进行介绍。本实验室的同学应当认可老师的工作安排，有抵触心理主动提出，可以更换到其他团队。

**四、开题相关内容**

开题需要的材料要提前做好，会邀请本校和外校专家组同时评判，有无法通过开题的风险。

**会议总结：**

1、介绍了Dr. Mo技术的具体实现和实验结果，该技术通过结合精心设计的轻量级帧间运动模型，将这些粗粒度噪声传播到下一帧，从而消除了帧扩散模型中的大量计算冗余。讨论了该技术应用到当前工作的可能性以及该技术中的一些不足之处。

2、承楠老师发表关于就业以及行业选择的一些看法

3、学习实验室的规章制度，包括纪律要求、毕业要求、实验室补助等情况。

4、对研究生开题工作做了简单部署。