

“Video Generation Techniques for Novel View Synthesis with Flow-Matching Transformers”

By Xiuyuan Qiu

Abstract:

Novel view synthesis (NVS) aims to generate images of a scene from unseen camera viewpoints. Recent work, such as Stable Virtual Camera, shows that large-scale image diffusion models like Stable Diffusion can be adapted for pose-conditioned view synthesis by incorporating video-generation techniques with camera conditioning. In this thesis, we introduce MVFlow, a new NVS model that extends this approach to a different image generation architecture: a flow-matching diffusion transformer, specifically FLUX.1, which has demonstrated strong performance in image synthesis. We evaluate MVFlow under varying input view counts and pose distance settings. Our results show that this architectural transfer is feasible; however, the current architecture and conditioning strategy do not yet match the performance of a mature baseline such as Stable Virtual Camera.

Date: Monday, March 16th, 2026

Time: 11:00 AM – 1:00 PM

Location: 14-238b

Committee: Dr. Jonathan Ventura, Dr. Alex Dekhtyar, Dr. Austin Wright