AMCC 6500G: Special Topics on Video Generation

Course Outline — Spring 2026

Course Information

- **Term:** Spring 2026
- **Time:** Friday 13:30 – 16:20
- **Location:** RM 1409
- **Instructor:** Prof. Chao (Harry) Yang
- **Teaching Assistant:** Han Zhu (mkhanzhu@ust.hk)
- **Contact:** yangharry@ust.hk

Course Description

This advanced research seminar treats video generation not merely as media synthesis, but as the foundation for **General World Models** and **Embodied Agents**. We will move beyond standard diffusion to explore the frontiers of **Flow Matching** and **Autoregressive Visual Transformers**. The curriculum emphasizes the convergence of video with robotics and interaction: specifically **Vision-Language-Action (VLA)** models, **drivable 3D avatars**, and **neural simulators**. Students will investigate how large-scale video pre-training enables "Playable Worlds," where AI agents can perceive, predict, and act within consistent, generated 4D environments.

Course Format & Tools

- **Format:** Lecture + Presentation + Projects
- **Tools:** PyTorch, Diffusers, ComfyUI, 3D Gaussian Splatting, Horizon/World Model codebases, MuJoCo

---

Weekly Schedule

Week 1: The New Thesis — Video Models as World Models

Week 2: Video Representation Learning Beyond Reconstruction

Week 3: Modern Video Generators — Diffusion vs. Autoregressive vs. "Video-LLM" Unification

Week 4: Playable / Action-Conditioned Video — From Prediction to Control

Week 5: VLA I — How Vision-Language Meets Control

Week 6: VLA II — Video as a Counterfactual Simulator for Agents

Week 7: Generative Agents in Video Worlds

Week 8: Talking Heads (Audio → Video) as a Controlled Generation Problem

Week 9: Drivable Avatars — From 2D Faces to 3D Gaussian/NeRF Heads

Week 10: 4D Dynamic Scenes — Text-to-4D, Dynamic NeRFs, and Gaussian Splatting in Motion

Week 11: Long-Horizon Consistency — Memory, Tokens, and Anti-Drift Methods

Week 12: Safety, Provenance, and Reality Defense (for Video + Avatars)

Week 13: Final Project

Assessment

Run an open source repository and make small improvements on:

1. **Project 1 (20%):** World Model
2. **Project 2 (20%):** VLA
3. **Project 3 (20%):** Agents
4. **Course Presentation (10%):** Present a chosen paper on relevant topics
5. **Final Project Presentation (30%):** Presentation and demo of one selected project developed during the semester