

AI for Video

Ming Tang
Professor, DAAP, UC
Director of XR-Lab

Open/open-weight video AI models include [Wan 2.2](#), [LTX-2](#), [HunyuanVideo](#), [CogVideoX](#), [Mochi 1](#), [Open-Sora](#), [AnimateDiff](#), and [Stable Video Diffusion](#).

Closed-source/commercial platforms include [Veo 3.1](#), [Runway](#), [Kling AI](#), [Seedance](#), [Dream Machine](#), [Adobe Firefly Video](#), [Pika](#), [Hailuo AI](#), [PixVerse](#), and [Vidu](#).

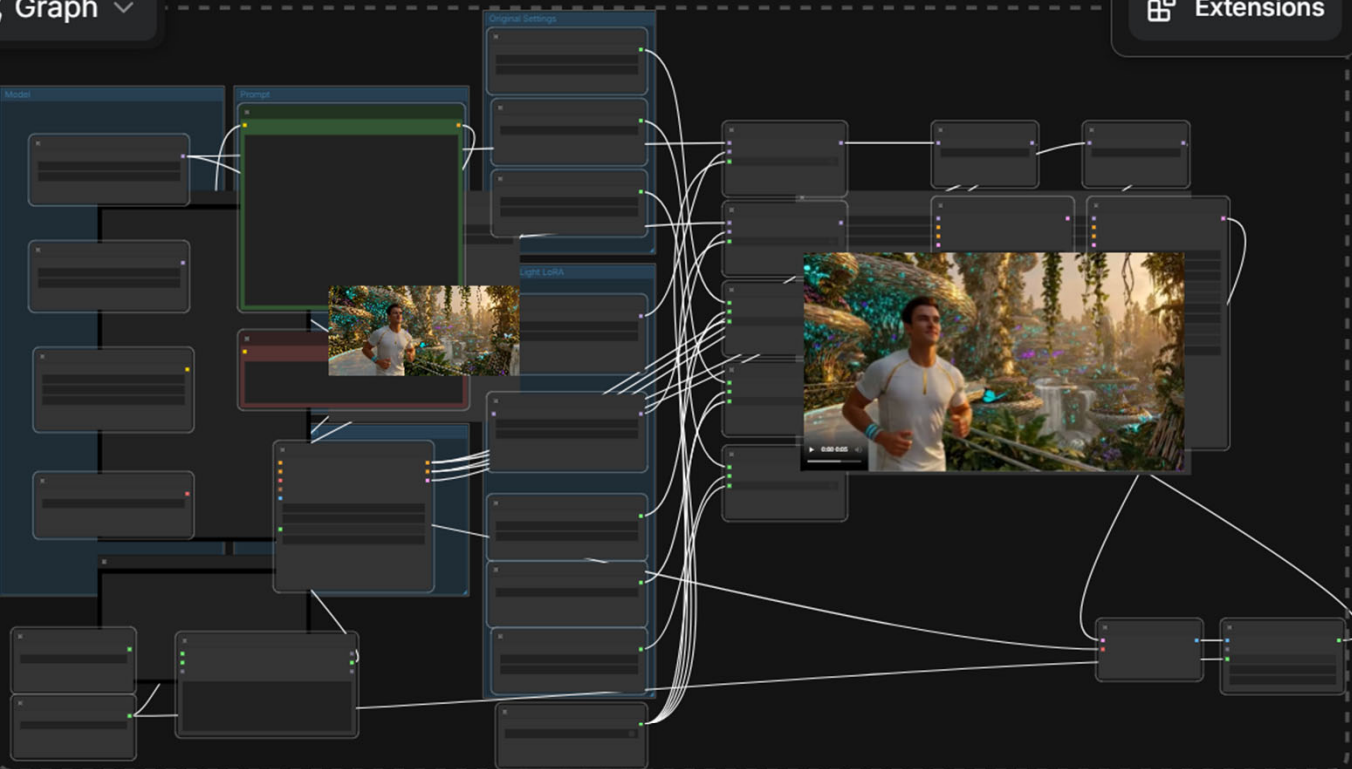
Graph

Model

Prompt

Original Settings

Light Lora



Extensions

1

Run

0 active

Job Queue

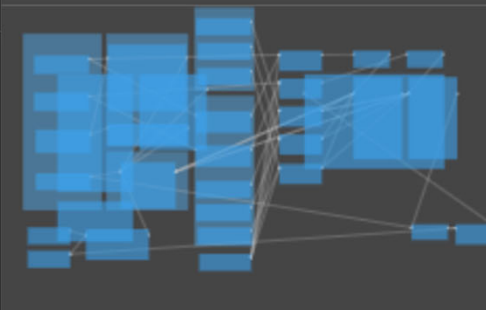
Clear queue

All

Completed

Today

Wan2.2 i2v 00006 .mp4
933.08



21%

AI Router

Hunyuan3D · Wan/ComfyUI · A1111

arch (user)

Ready

Docs

Refresh

Log out

Video

Image

Jobs

Workers

Account

Mode

Image to video

Seed

2071100

Prompt

Slow cinematic camera push forward. Smooth, steady movement. Architecture remains stable and unchanged. Subtle atmospheric movement and consistent lighting.same person, consistent facial features, detailed natural face, sharp eyes, stable facial identity, realistic skin

Negative prompt

warping, morphing, distorted architecture, deformed buildings, unstable geometry, changing building shapes, flickering, jitter, camera shake, sudden movement, rotation, blurry, artifacts, deformed face, distorted face, changing face, blurry face, asymmetrical eyes, malformed eyes, duplicate face, facial artifacts, morphing

Source image

Choose File Screenshot 2026-09-14 225023.jpg

AI servers

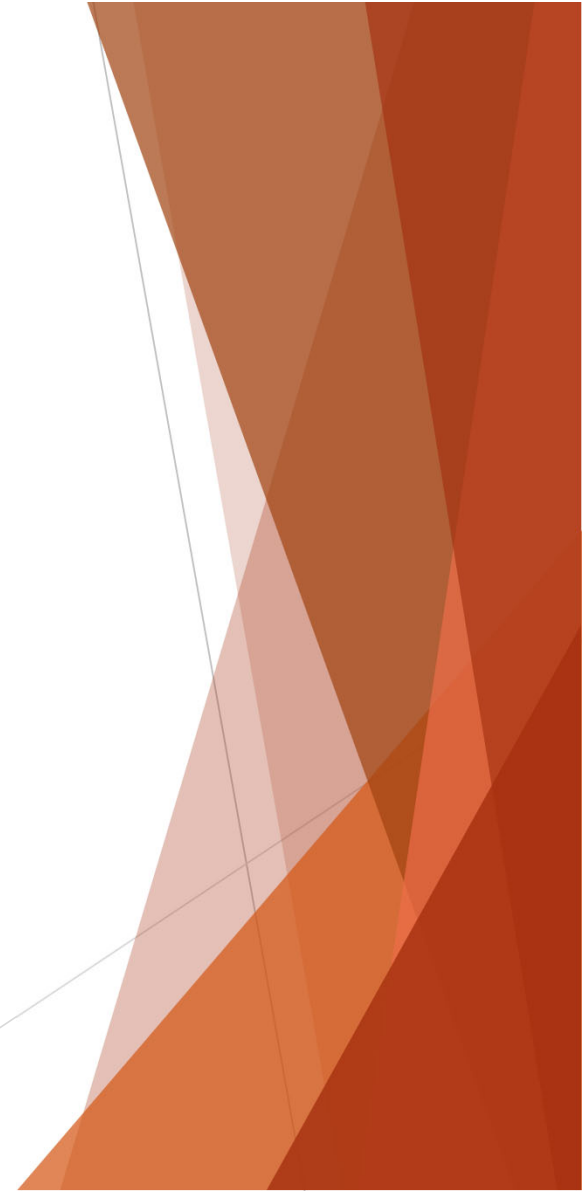
CFG means Classifier-Free Guidance

It controls how strongly the generated video follows your text prompt.

Lower CFG (e.g., 2-4): More freedom/natural variation; motion may look smoother, but the result can drift from the prompt.

Medium CFG (e.g., 4-6): Good balance between prompt adherence and natural-looking video.

Higher CFG (e.g., 7-10+): Follows the prompt more aggressively, but can introduce unnatural motion, oversharpening, flickering, or other artifacts.



Recommended Settings for Image-to-Video Test

- **Resolution:** 960×544
- **Frames / FPS:** 81 frames / 16 fps
- **Steps:** 25
- **CFG:** 4
- After generation, upscale 2× using another upscaling method.

Final

- **Resolution:** 1280×720
- **Frames / FPS:** 81 frames / 16 fps
- **Steps:** 30
- **CFG:** 4

Prompt: The characters walking down the street.



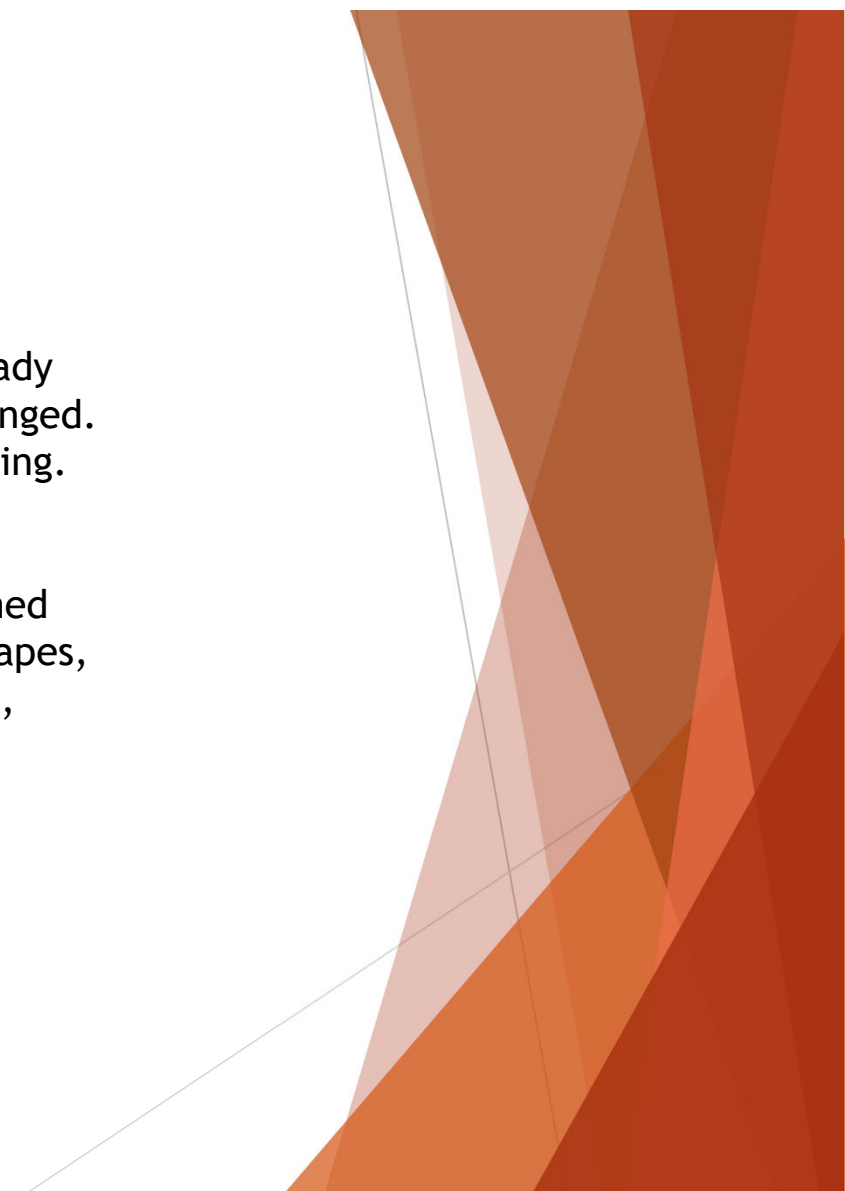
960 X 544, 16fps
Rendering time. 933s

Prompt

Slow cinematic camera push forward. Smooth, steady movement. Architecture remains stable and unchanged. Subtle atmospheric movement and consistent lighting.

Negative Prompt

warping, morphing, distorted architecture, deformed buildings, unstable geometry, changing building shapes, flickering, jitter, camera shake, sudden movement, rotation, blurry, artifacts



If There Is a Human Face in the Scene

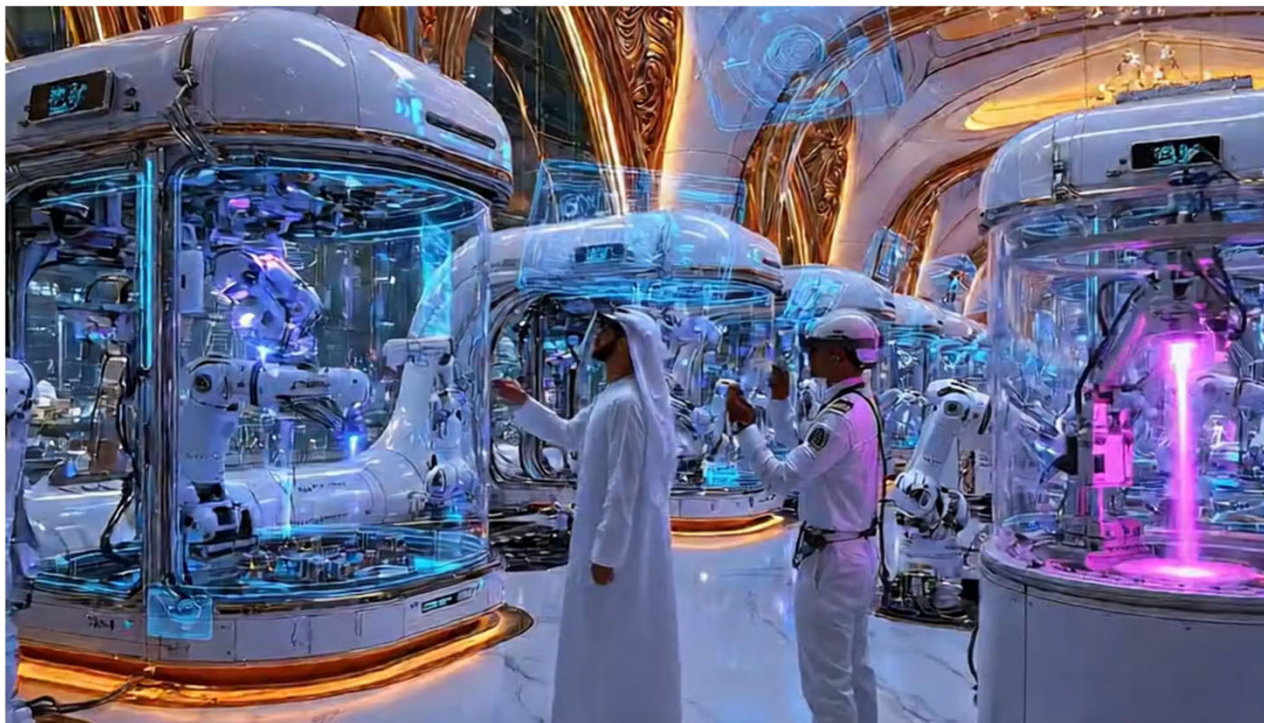
Add to the **prompt**:

Slow cinematic camera push forward. Smooth, steady movement. Architecture remains stable and unchanged. Subtle atmospheric movement and consistent lighting. same person, consistent facial features, detailed natural face, sharp eyes, stable facial identity, realistic skin

Add to the **negative prompt**:

warping, morphing, distorted architecture, deformed buildings, unstable geometry, changing building shapes, flickering, jitter, camera shake, sudden movement, rotation, blurry, artifacts, deformed face, distorted face, changing face, blurry face, asymmetrical eyes, malformed eyes, duplicate face, facial artifacts, morphing





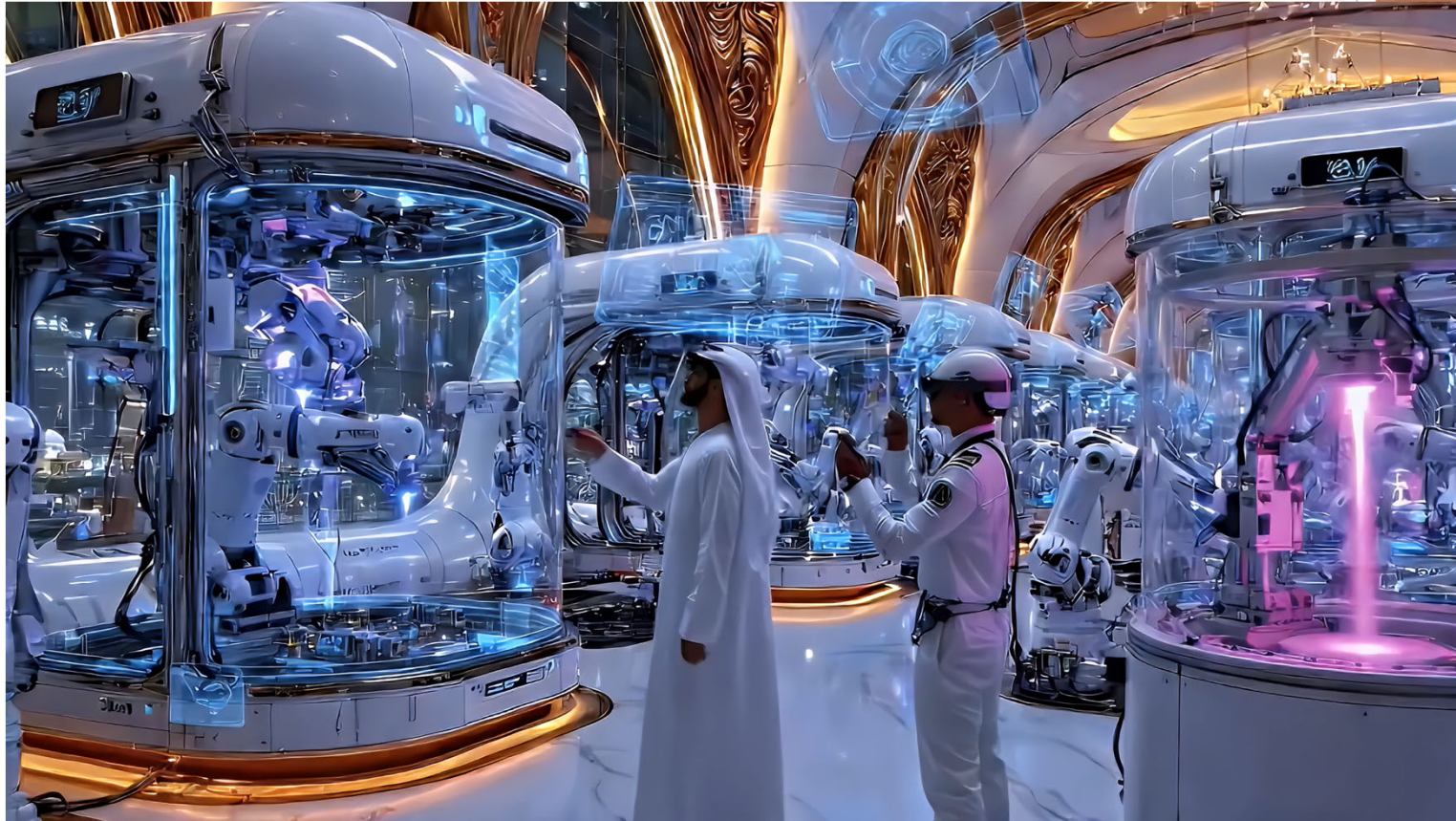
960 X 544, 16fps
569s render time



Upscale X 2, frame X 2



1280 X 720, 16fps
1020s render time



Upscale X 2, frame X 2 as slow motion



960 X 544, 16fps



Upscale X 2, frame X 2



960 X 544, 16fps



Upscale X 2, frame X 2



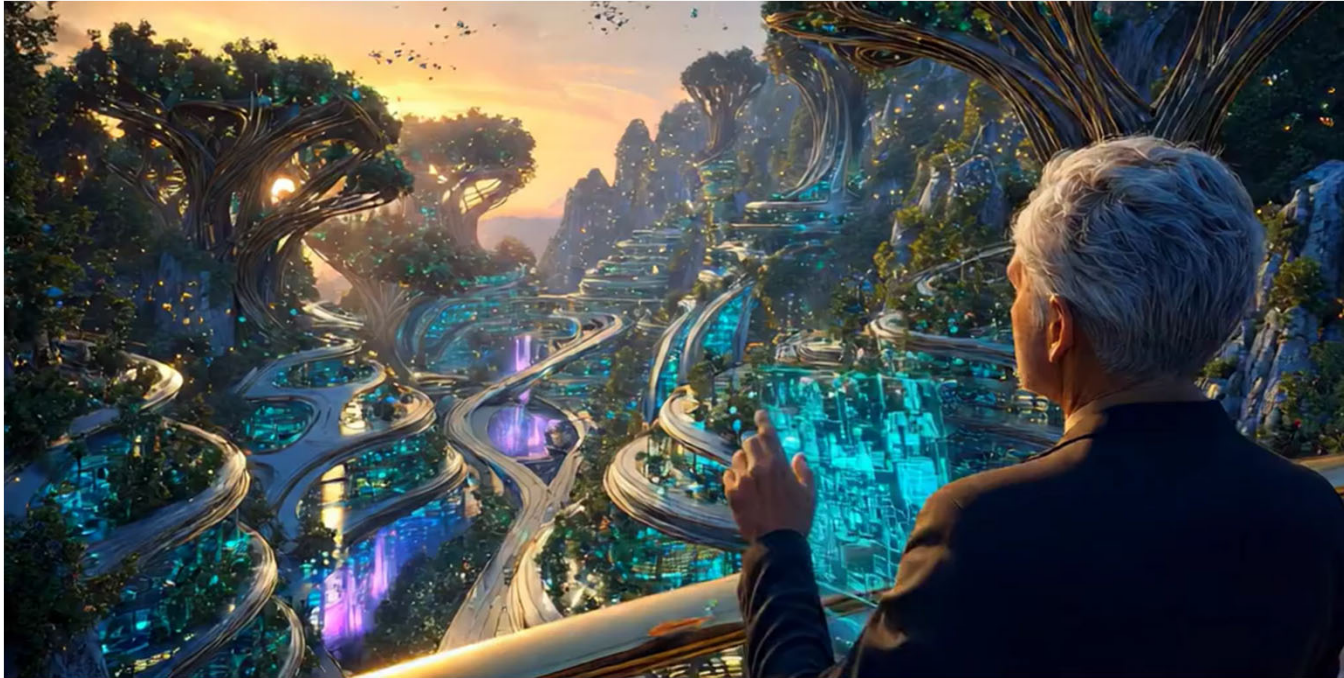
960 X 544, 16fps



1280 X 720, 16fps



1024 X 512, 16fps



1024 X 512, 16fps



Upscale X 2, frame X 2



960 X 544, 16fps



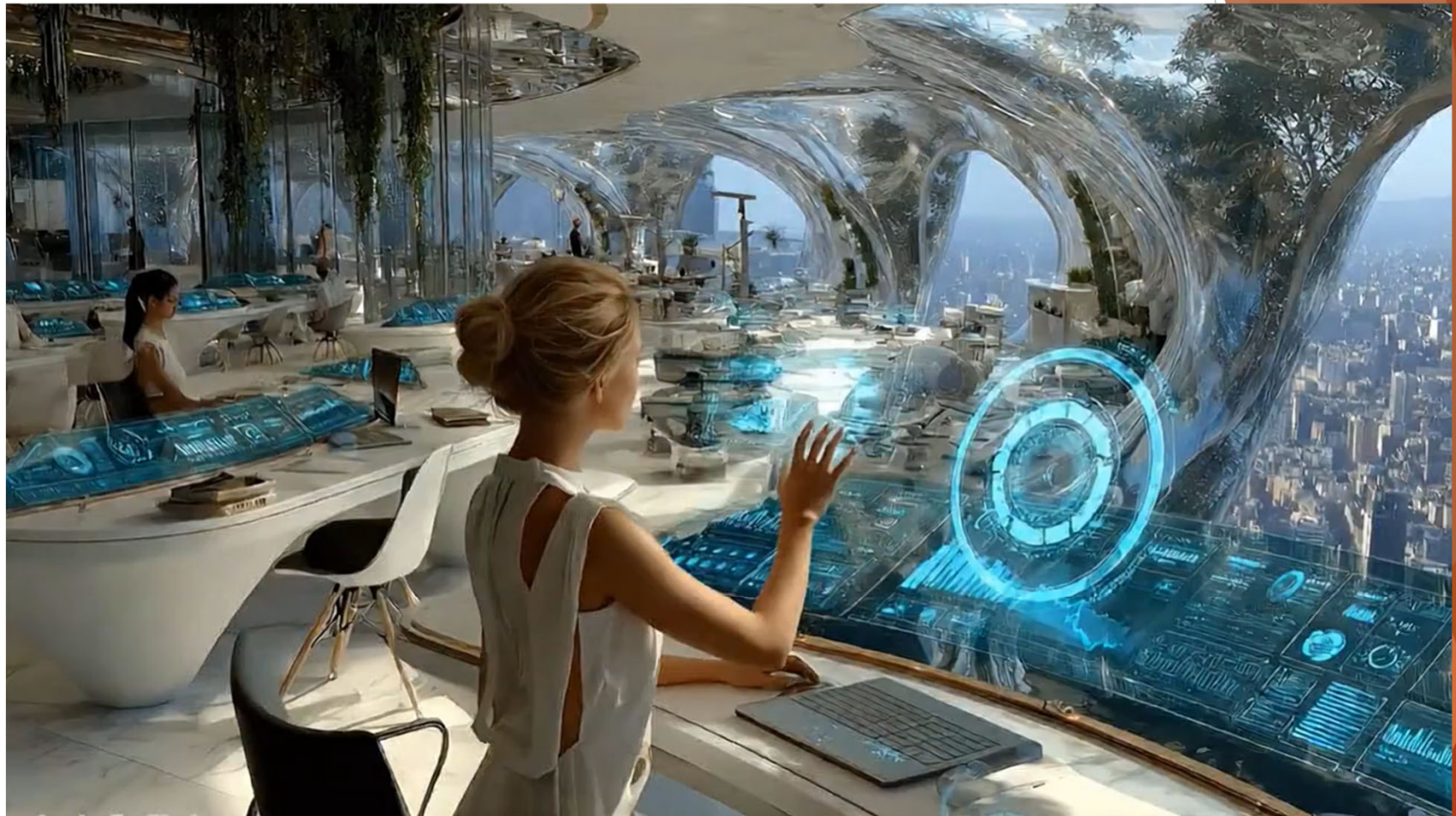
1920 X 1080, 16fps



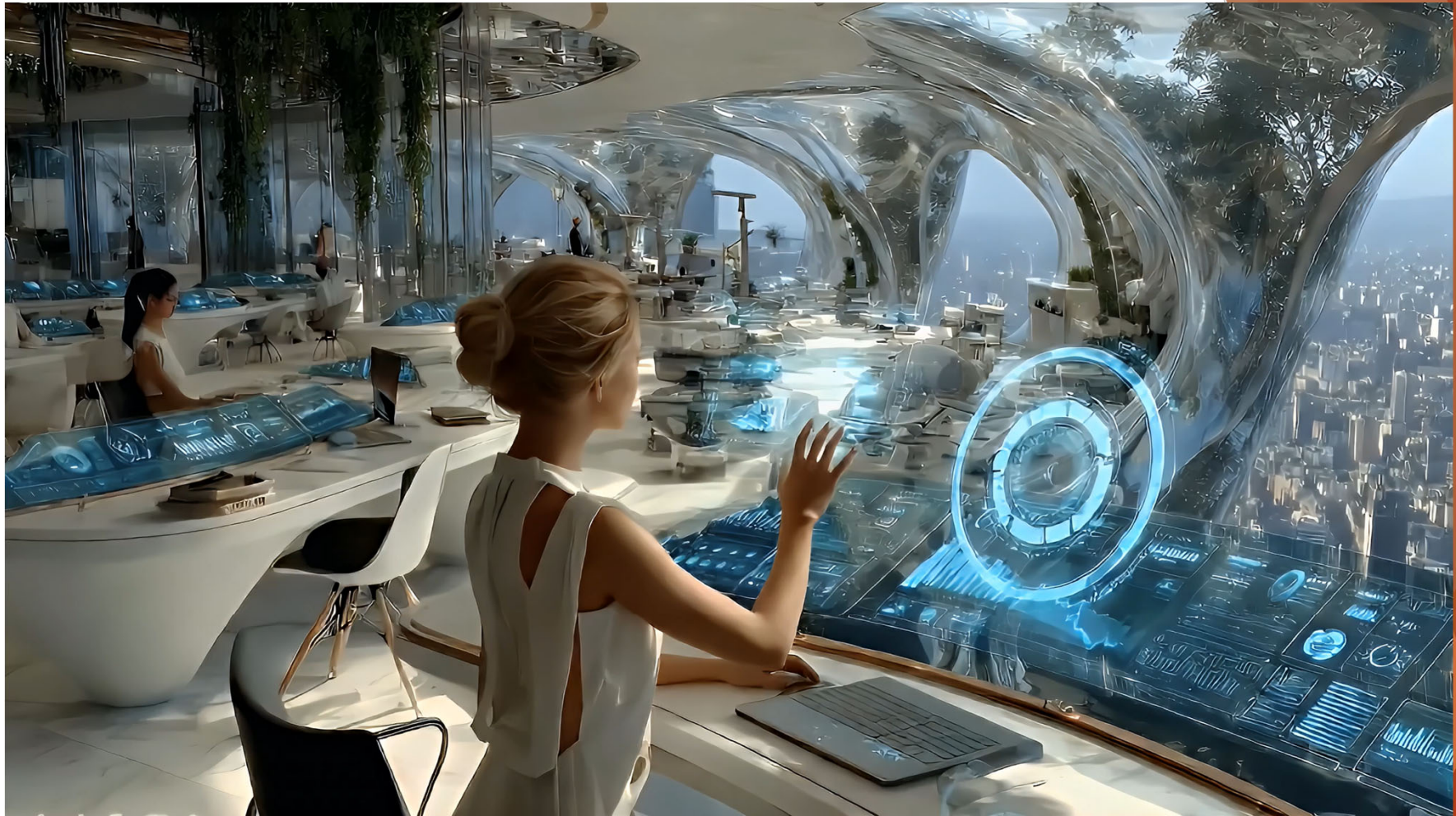
Upscale X 2, frame X 2



Upscale X 2, frame X 2, slowmotion



1280 X 720 16fps



Upscale X 2, frame interpolation to 32fps

Submit jobs between 6pm - 8AM when no one use computers in xr-lab.

Recommended Settings for Image-to-Video Test

- **Resolution:** 960 × 544
- **Frames / FPS:** 81 frames / 16 fps
- **Steps:** 25
- **CFG:** 4
- After generation, upscale 2× using another upscaling method.

Final

- **Resolution:** 1280 × 720
- **Frames / FPS:** 81 frames / 16 fps
- **Steps:** 30
- **CFG:** 4

Rule 1: Use Image-to-Video

Use **Image-to-Video** whenever possible. It gives you more control over the final output. Avoid **Text-to-Video**, which is less predictable and less efficient in terms of computing time.

Prepare your input image at the **exact pixel dimensions** you want for the final video, such as **960 × 544** or **1280 × 720**.

Rule 2: Keep Resolution at 1280 × 720 or Below

Do not go beyond **1280 × 720**. There is usually little benefit to generating AI video at very high resolutions, and it significantly increases rendering time. You can use other tools to upscale the video afterward.

Rule 3: Keep Videos at 5 Seconds or Less

Do not go beyond **5 seconds (81 frames)**. The default frame rate is **16 fps**. After generating the video, you can slow it down or use frame interpolation to increase it to **32 fps** for smoother motion.

Rule 4: Keep Render Time Under 60 Minutes

Our system will automatically cancel any job with an estimated or actual render time of more than **60 minutes**. Adjust the resolution, number of frames, or other settings if necessary to stay within this limit.