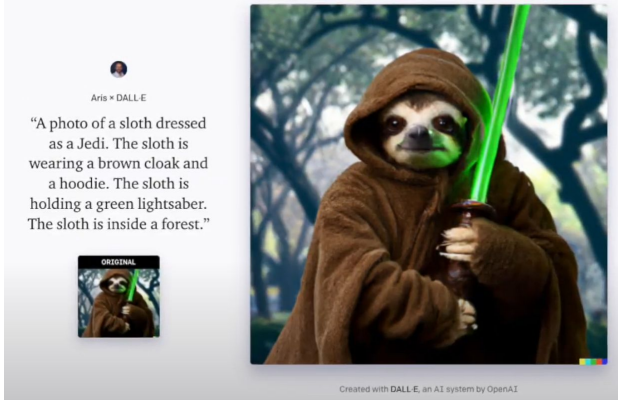
Master Seminar - Recent Advances in 4D Computer Vision

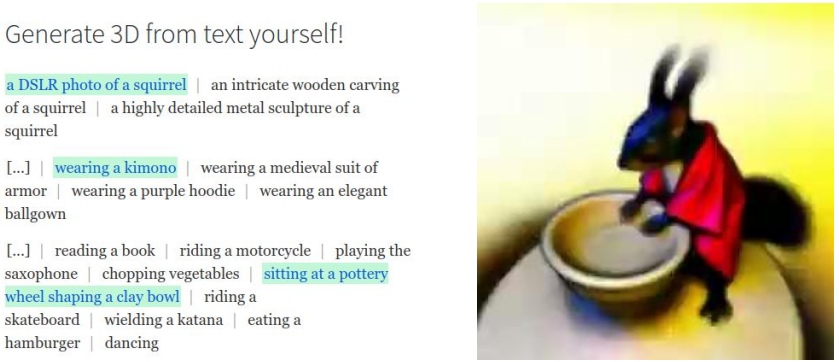
Cecilia Curreli, Mariia Gladkova

Why 4D?

2D



3D



Understanding 4D reality is an exciting direction of great research interest!

[1] Ramesh, et al. DALL-E 2"Hierarchical Text-Conditional Image Generation with CLIP Latents." (2022).
[2] Poole, Ben, et al. "Dreamfusion: Text-to-3d using 2d diffusion." (2022).

Text-to-4D Generation

A turtle swimming.

Waves crashing against a lighthouse.

A bee fluttering its wings fast.

A cat singing, best quality, 4K, HD

Generating 4D is a hot topic in Computer Vision research community!

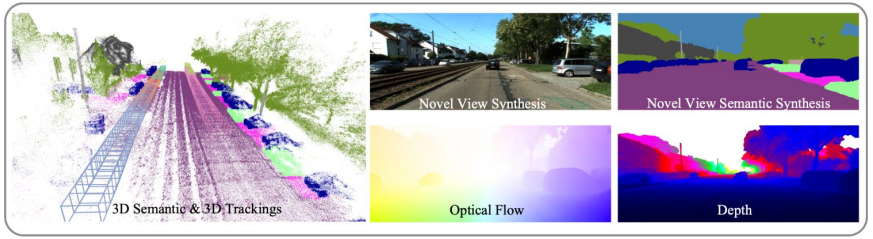
Ling et al. "Align Your Gaussians: Text-to-4D with Dynamic 3D Gaussians and Composed Diffusion Models" CVPR (2024).

4D Novel View Synthesis

We live in a 4D world :)

[2] Li et al. "Spacetime Gaussian Feature Splatting for Real-Time Dynamic View Synthesis" CVPR (2024).
[1] Wu et al. "4D Gaussian Splatting for Real-Time Dynamic Scene Rendering" CVPR (2024).

4D Autonomous Driving



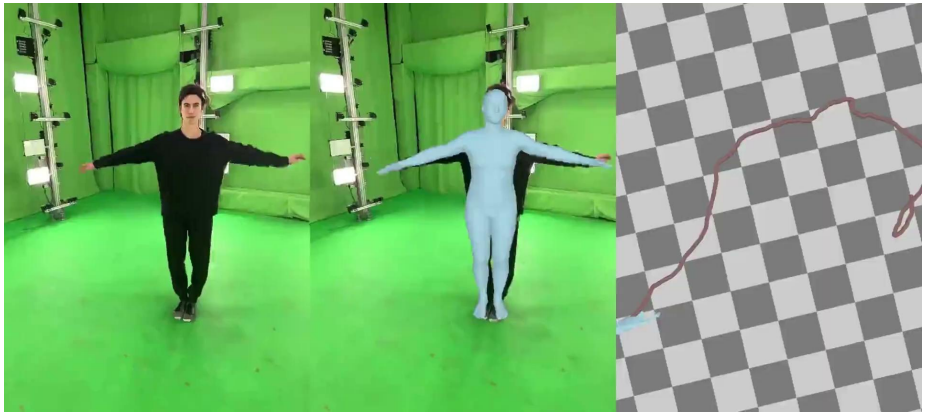
3D HD-Maps? Why not 4D!

Zhou et al. "HUGS: Holistic Urban 3D Scene Understanding via Gaussian Splatting" CVPR (2024).

4D Humans

Source Actor

Animated Avatar

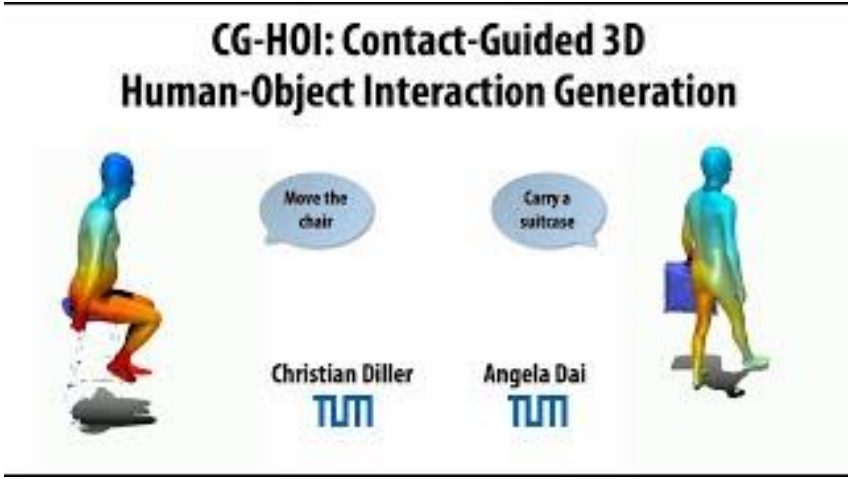


Multi-billion net-worth companies invest heavily in human avatars …

[1] Quian et al. "GaussianAvatars: Photorealistic Head Avatars with Rigged 3D Gaussians." CVPR (2024).
[2] Shin, Kim et al. "WHAM: Reconstructing World-grounded Humans with Accurate 3D Motion." CVPR (2024).

4D Human Motion

Human Motion **Composition**
with FlowMDM



..and look closer into their realistic motion and interactions!

[1] Barquero et al. "FlowMDM: Seamless Human Motion Composition with Blended Positional Encodings." CVPR (2024).
[2] Diller et al. "CG-HOI: Contact-Guided 3D Human-Object Interaction Generation." CVPR (2024).

4D Computer Vision

Generation

Reconstruction

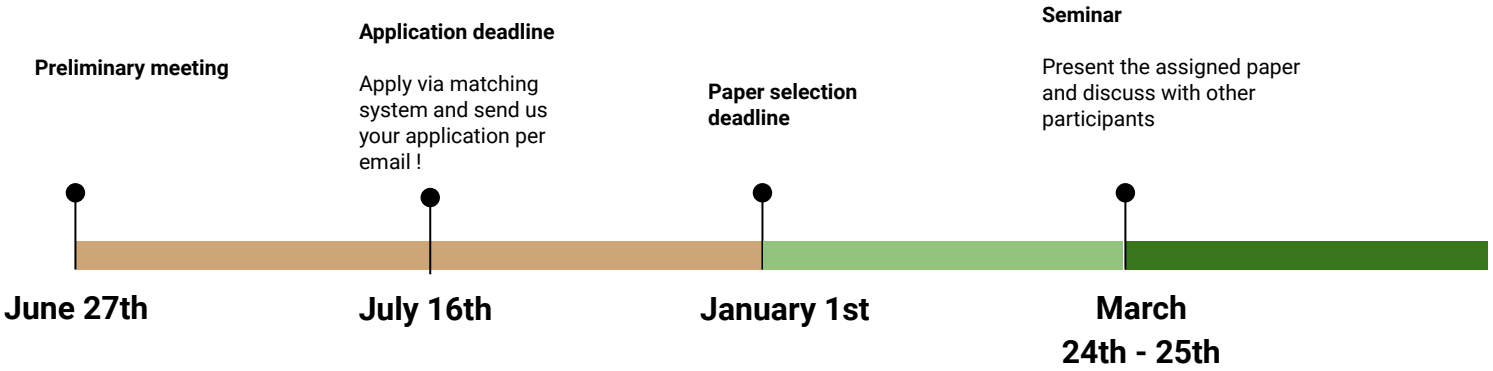
Understanding 4D reality is an exciting direction of great research interest!

[1] Ling et al. "Align Your Gaussians: Text-to-4D with Dynamic 3D Gaussians and Composed Diffusion Models" CVPR (2024).
[2] Wu et al. "4D Gaussian Splatting for Real-Time Dynamic Scene Rendering" CVPR (2024).

What are the topics discussed?

- Generation of dynamic scenes
  - Text-to-4D generation
- Estimation
  - 4D reconstruction & novel view synthesis
  - Motion estimation
  - Human-object interaction
- Various application domains
  - Autonomous driving
  - Robotics & XR
- … and many more!

How does the schedule look like?



March 2025, mark your calendars ;)

Further orga stuff

- Meet your supervisor
  - At least 1 week before presentation: paper and slides discussion
- Write a report
  - LaTeX template would be provided
  - 4 pages summarizing the paper and providing your high level insights
  - Report deadline is due in 3 weeks after your presentation
- All meetings and seminar sessions are mandatory to attend

How does the grading look like?

- Presentation: 50%
- Report: 40%
- Active participation and questions during seminar: 10%

What do I need to know beforehand?

- Previously attended DL lectures, such as I2DL, CV III, 3D Scanning & Motion Capture, and other
- Knowledge of linear algebra, probabilities, non-linear optimization is highly beneficial

We focus on learning methods, so brushing off the dust from DL notes will help!

I'm interested, how do I apply?

- Apply via matching system and assign our seminar high priority
- Send an email to 4dvision-ws24@vision.in.tum.de (example on the next slide) with
  - Short info about your background (see template)
  - Transcript of records
  - Resume

Two-step verification: matching system + email

Any tips for the email?

To 4dvision-ws24@vision.in.tum.de

Cc

Application 01234567

Hi Cecilia and Mariia,

My name is John Doe, and here are my details:

- Matriculation number: 01234567
- Bachelor grade: 2.0
- Master grade: 2.0
- DL courses: I2DL (2.3), Advanced DL for CV (1.3)

Make it concise and relevant to the seminar scope.

Best,
John

Any remaining questions?

Reach out via email for lost+found questions!