Lab Tour:
Capture and Modeling of 3D Humans in 3D Scenes

Korrawe Karunratanakul

PhD student

Computer Vision and Learning Group (VLG)
ETH Zurich, Switzerland

ETH zürich

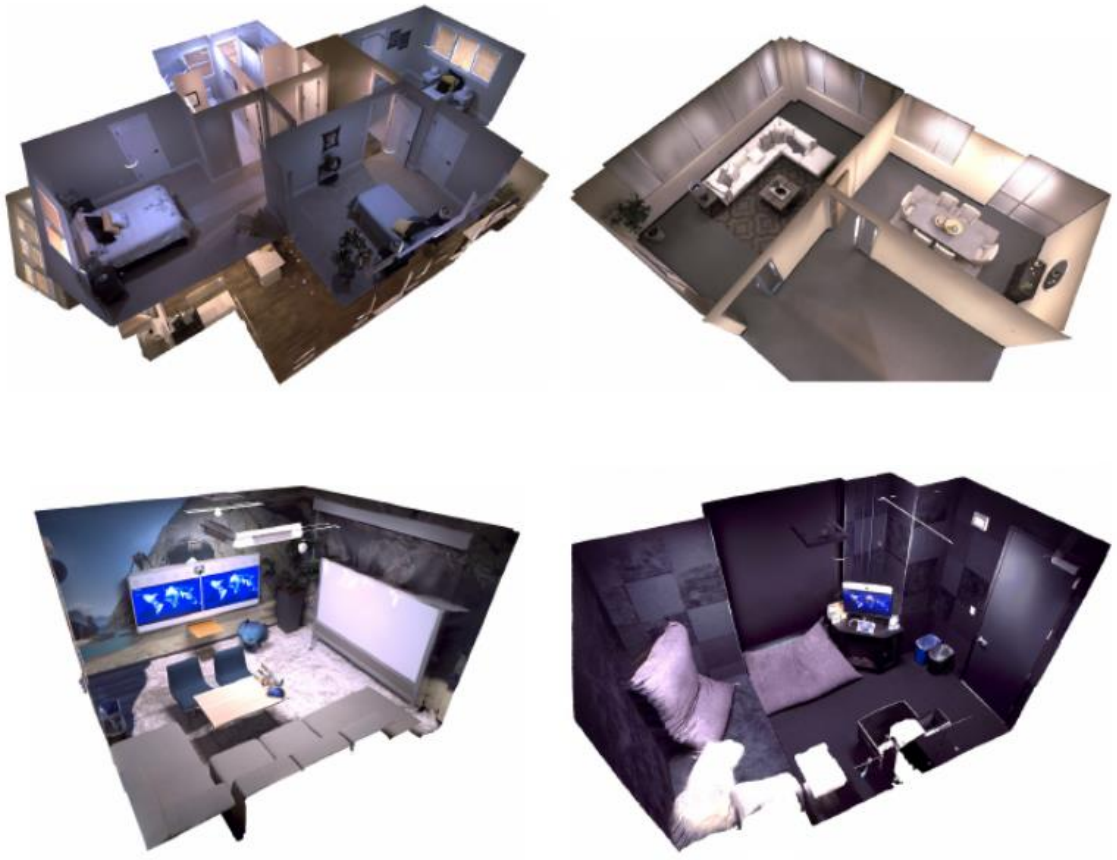
VLG Computer Vision and Learning Group

An overview of our group
(and our fields)

From my understanding …

Without verifying with my advisor first …

A Digital Replica of the World



Replica [1]

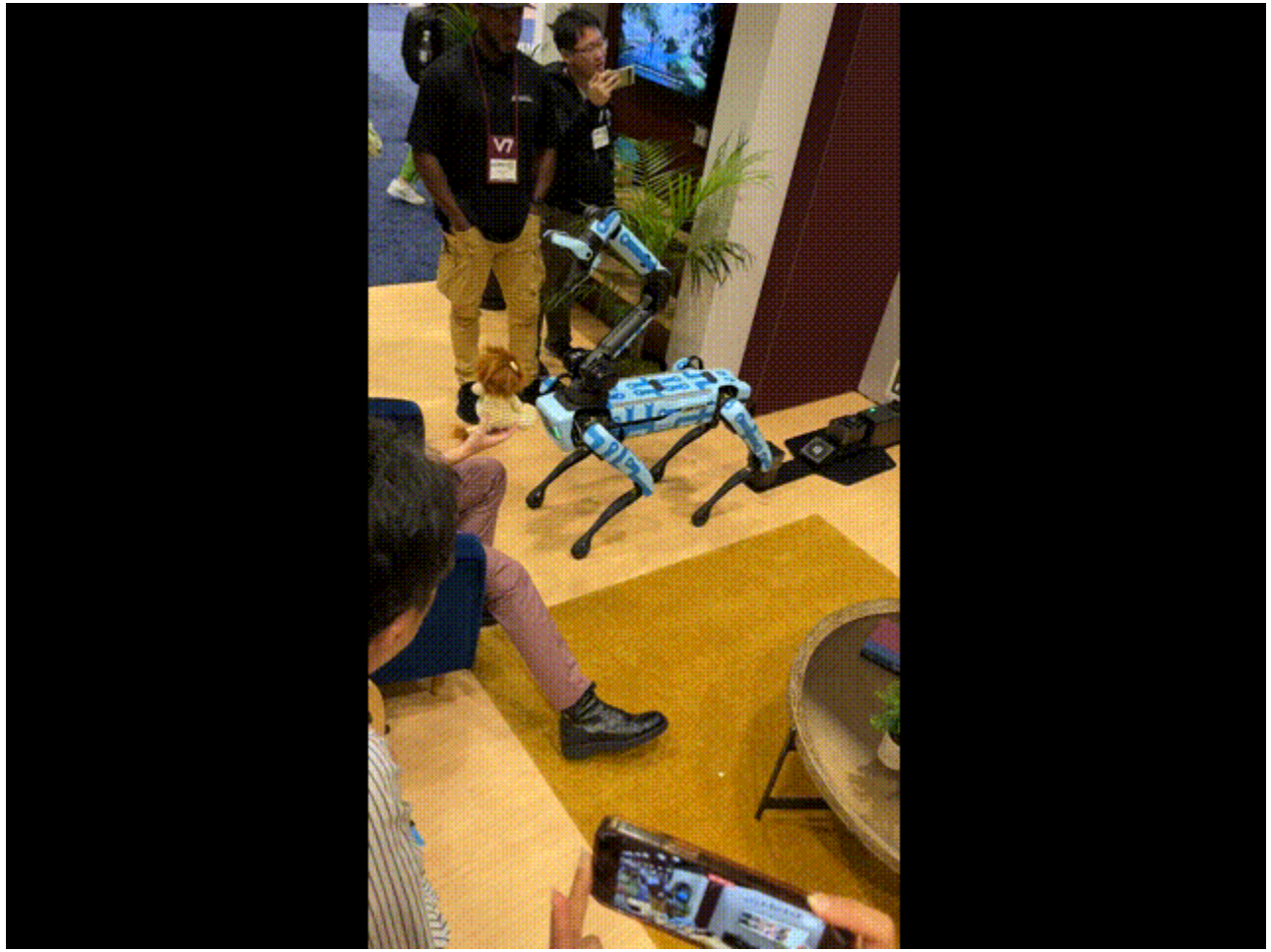
Matterport3D [2]

[1] Straub *et al.*, The Replica dataset: A Digital Replica of Indoor Spaces, *arXiv 2019*
[2] Change *et al.*, Matterport3D: Learning from {RGB-D} Data in Indoor Environments, *3DV 2017*

Embodied Agent in the Digital World

Khanna *et al*., Habitat Synthetic Scenes Dataset, *arXiv 2023*

Embodied Agent in the Digital World



Embodied Agent in the Digital World

Khanna *et al*., Habitat Synthetic Scenes Dataset, *arXiv 2023*

Generating Realistic Human Behaviors is Hard

Puig *et al.*, VirtualHome: Simulating Household Activities via Programs, *CVPR 2018*

Generating Realistic Human Behaviors is Hard

Hassan *et al*., Resolving 3D Human Pose Ambiguities with 3D Scene Constraints, *ICCV 2019*

Generating Realistic Human Behaviors is Hard

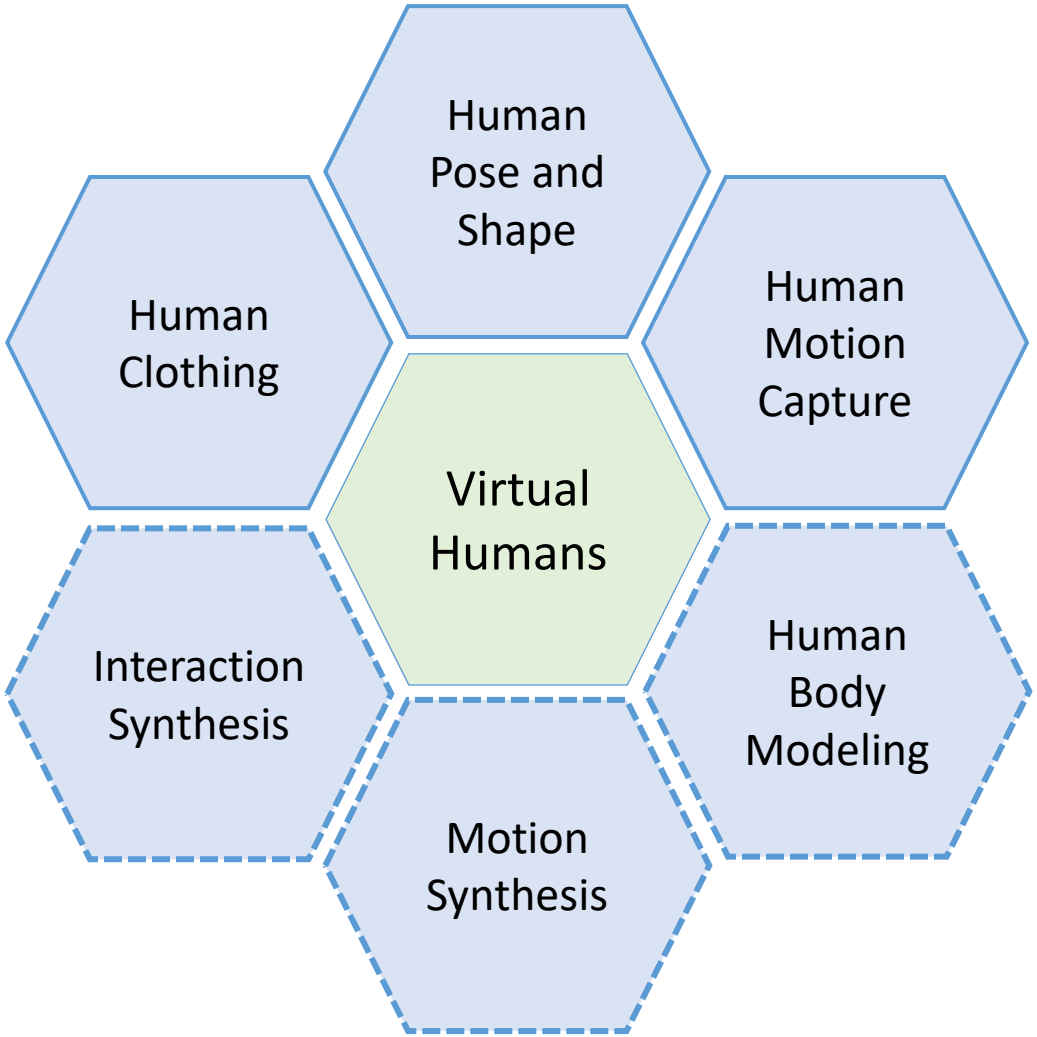
- Body Shapes
- Articulation
- 3D scenes and objects
- Physical plausibility
- Semantics

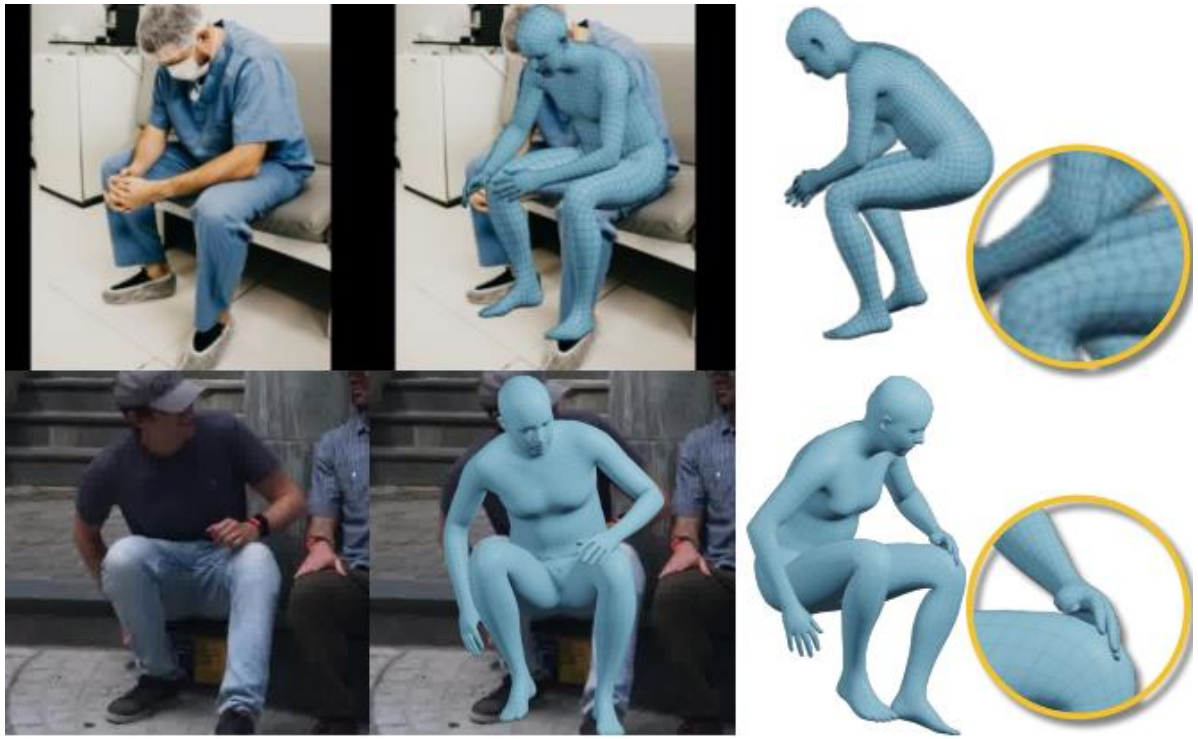
Hassan *et al.*, Resolving 3D Human Pose Ambiguities with 3D Scene Constraints, *ICCV 2019*

Our Research: Capture and model 3D humans in 3D scenes

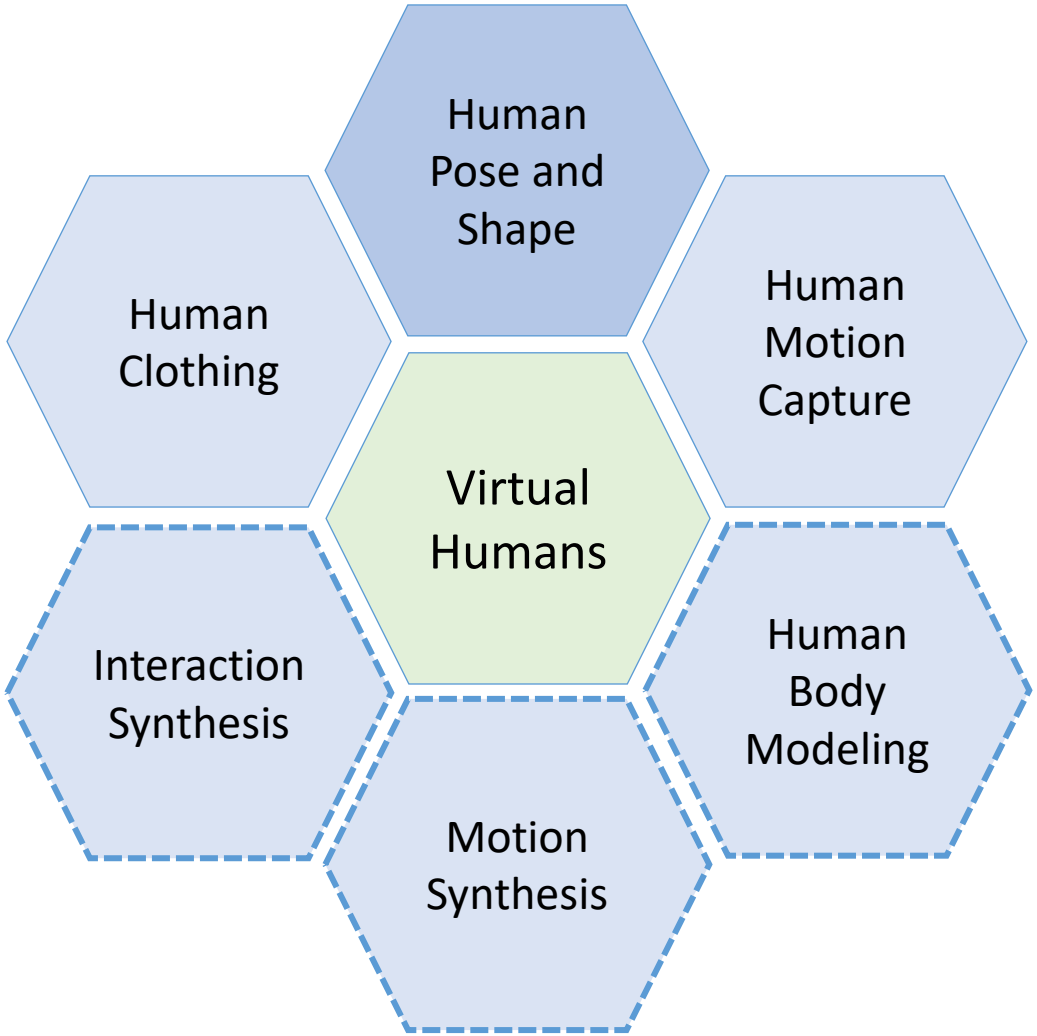
“Our goal is to build virtual humans that look, move, and behave like real ones”

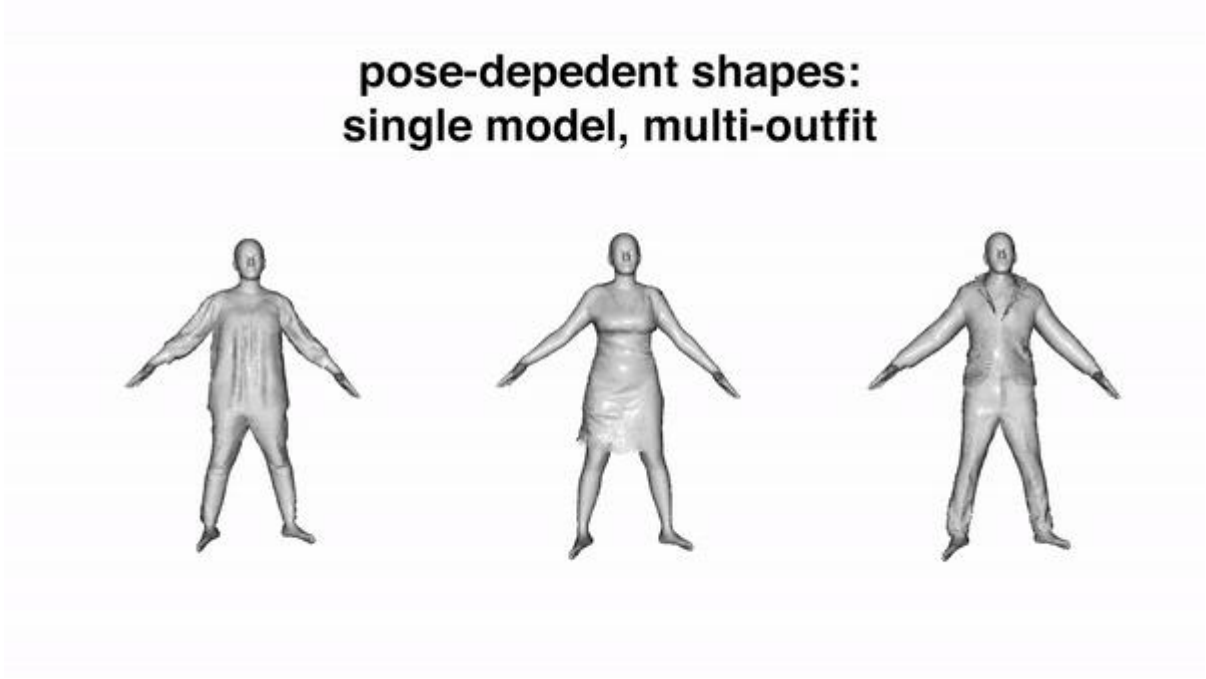
– Siyu Tang, from some talk.


Human
Pose and
Shape
Human
Clothing
Human
Motion
Capture
Virtual
Humans
Interaction
Synthesis
Human
Body
Modeling
Motion
Synthesis

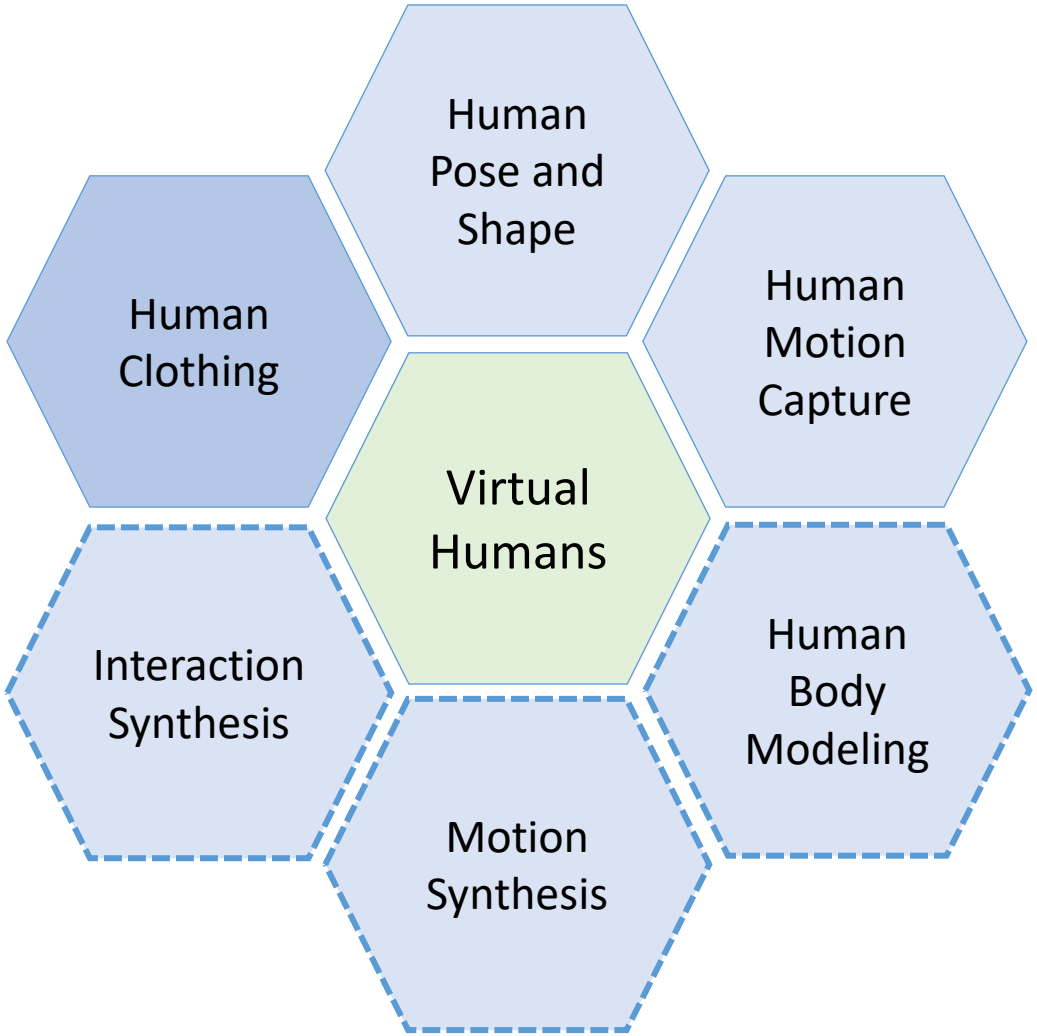


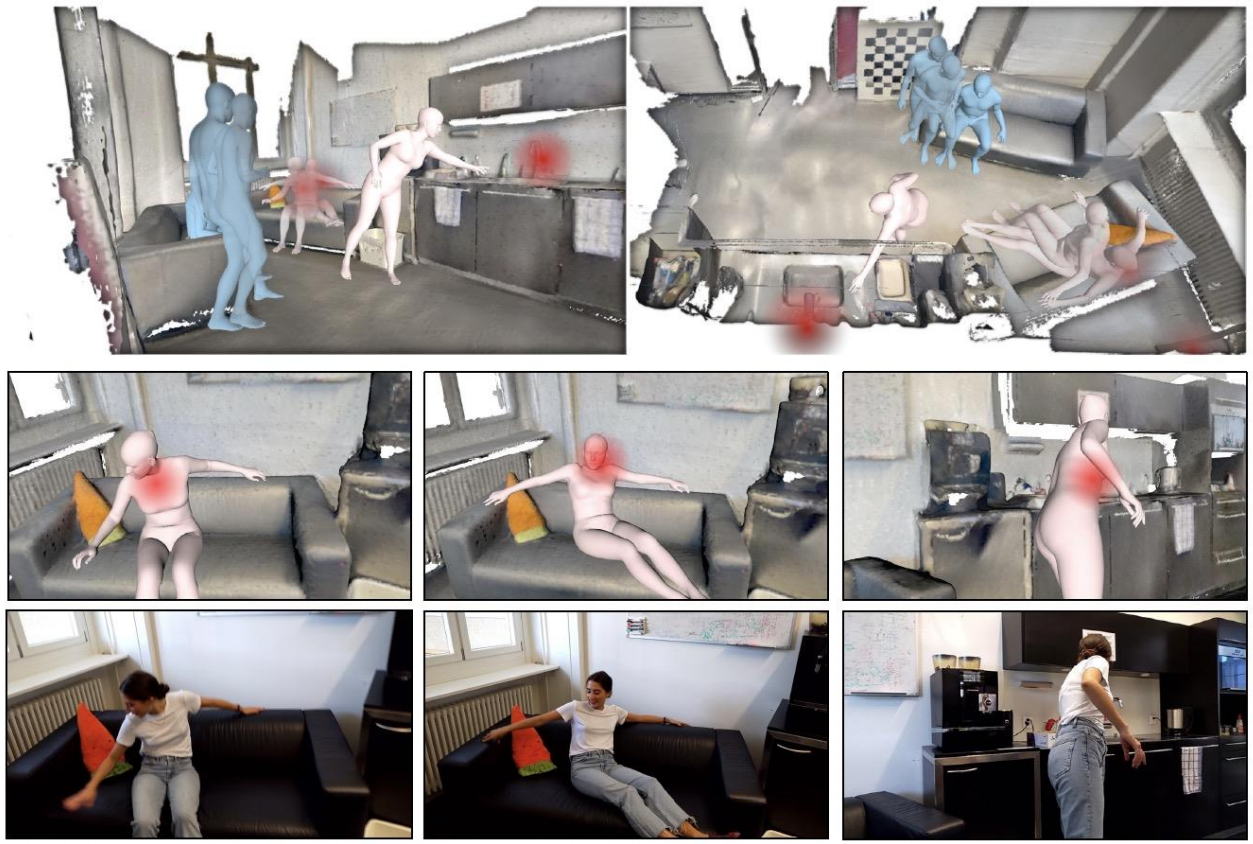
[1] PTF, *CVPR 2021*
[2] TUCH, *CVPR 2021*



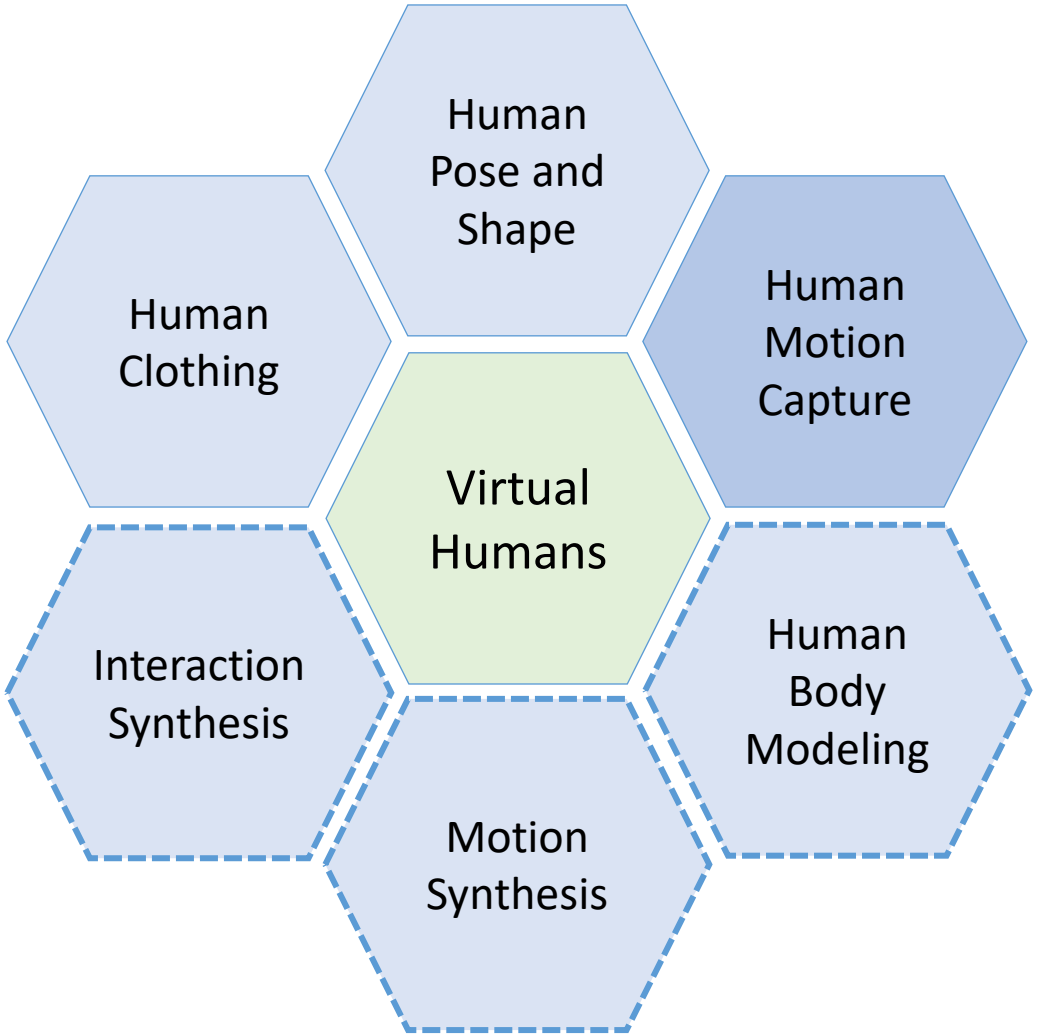


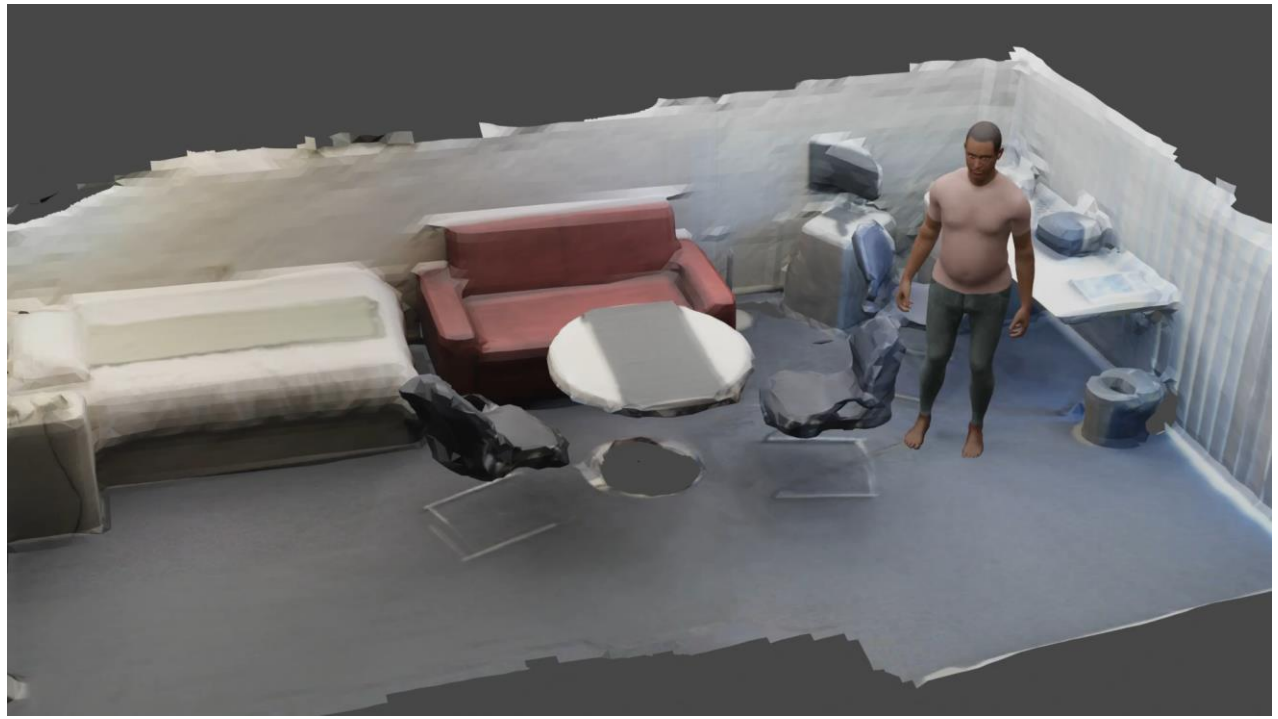
[1] CAPE, *CVPR 2020*
[2] SCALE, *CVPR 2021*
[3] POP, *ICCV 2021*

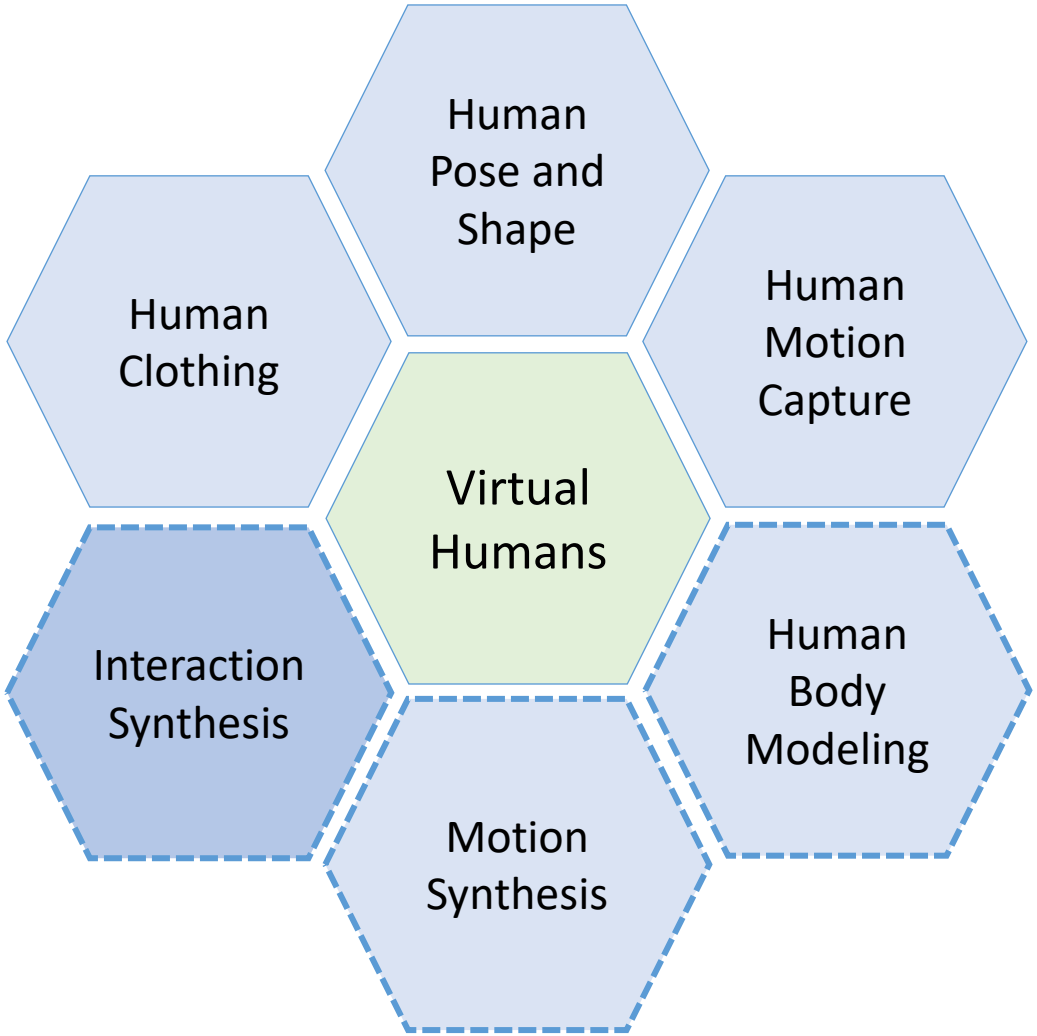




[1] LEMO, *CVPR 2021*
[2] EgoBody, *ECCV 2022*

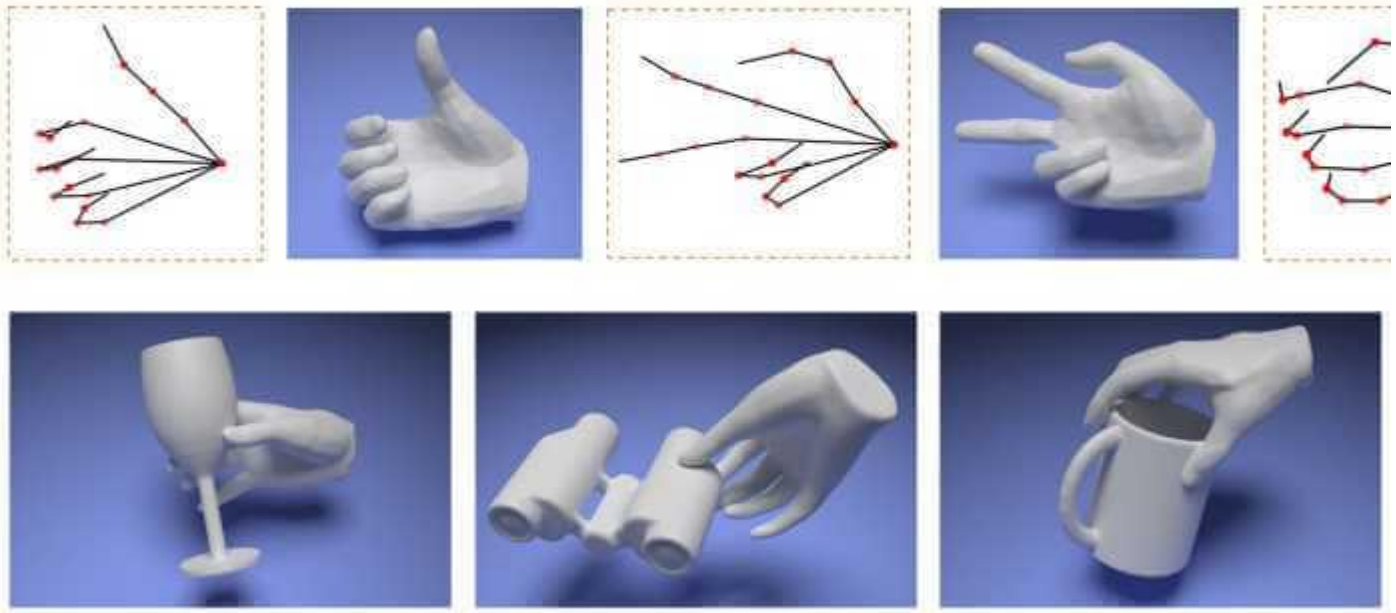


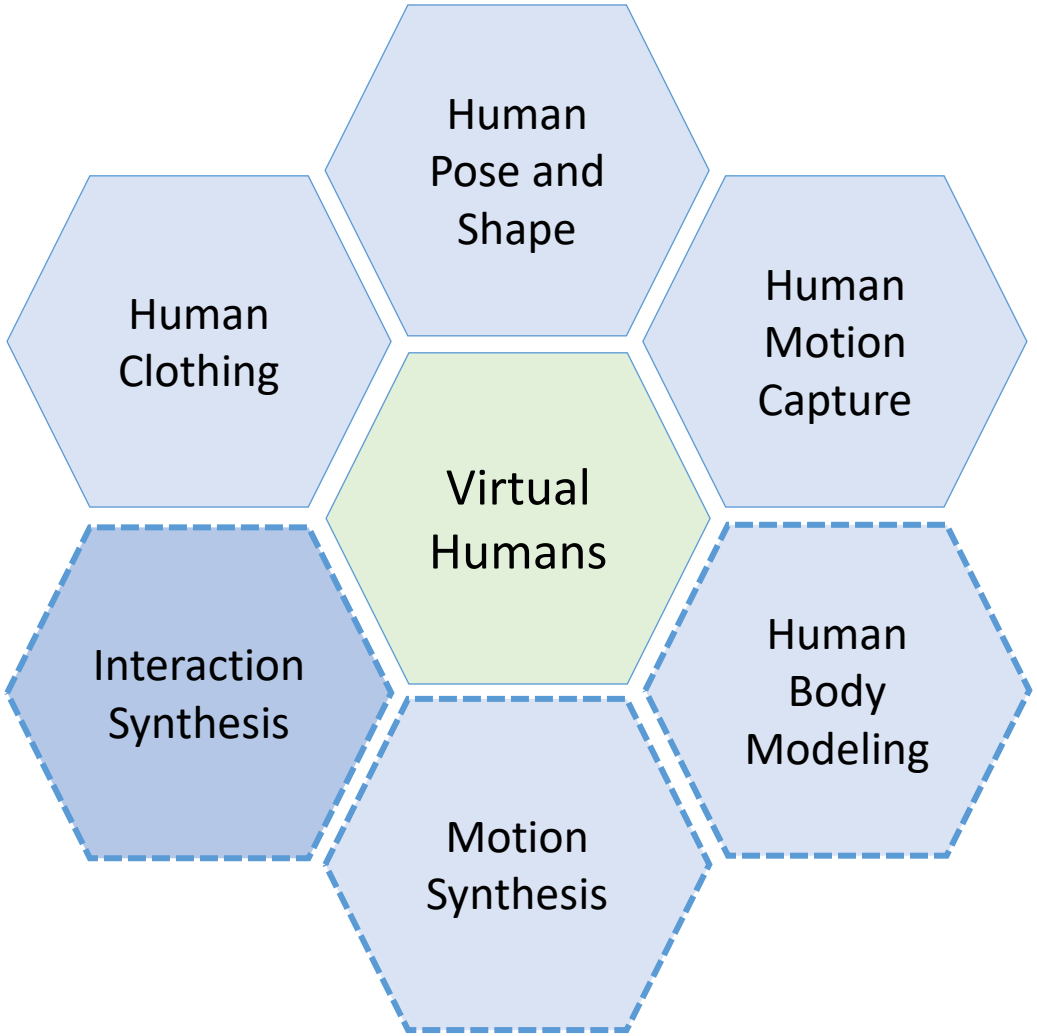




[1] DIMOS, *ICCV 2023*
[2] SAGA*, ECCV 2022*
[3] HALO, *3DV 2021*
[4] GraspingField, *3DV 2020*

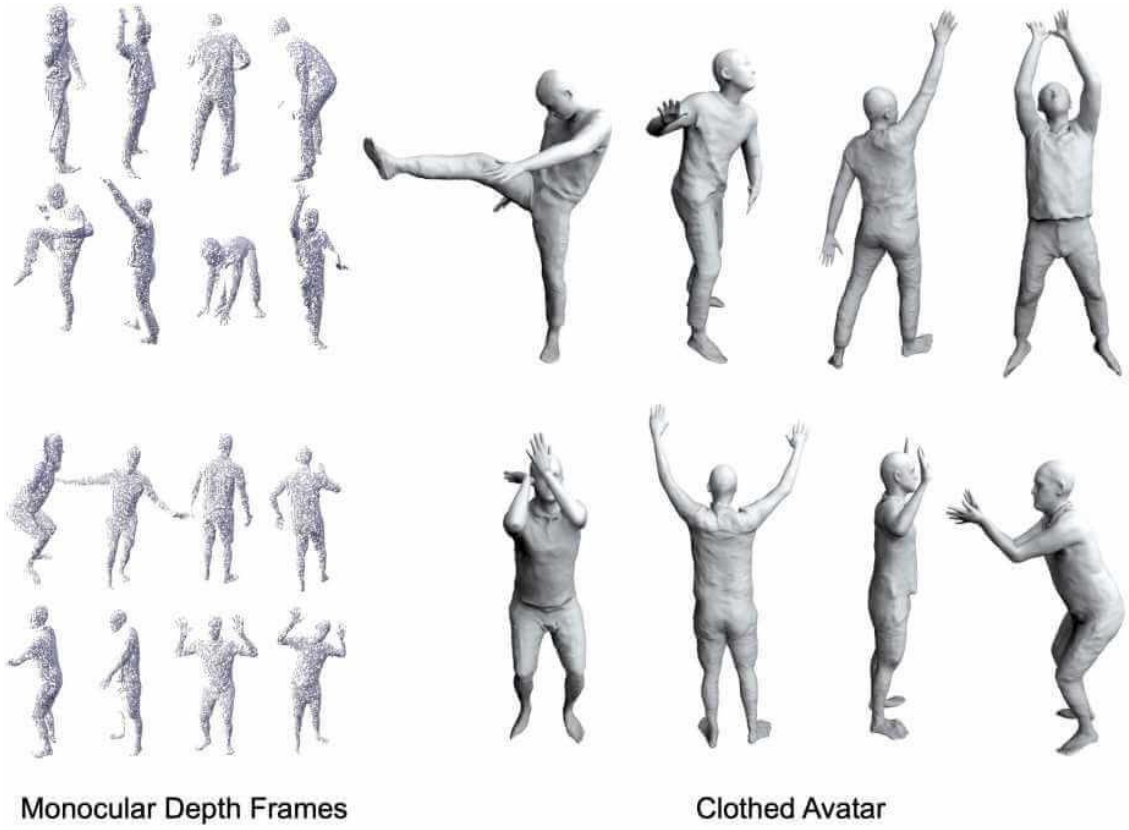
Not only body-scene, but also hand-object interaction!



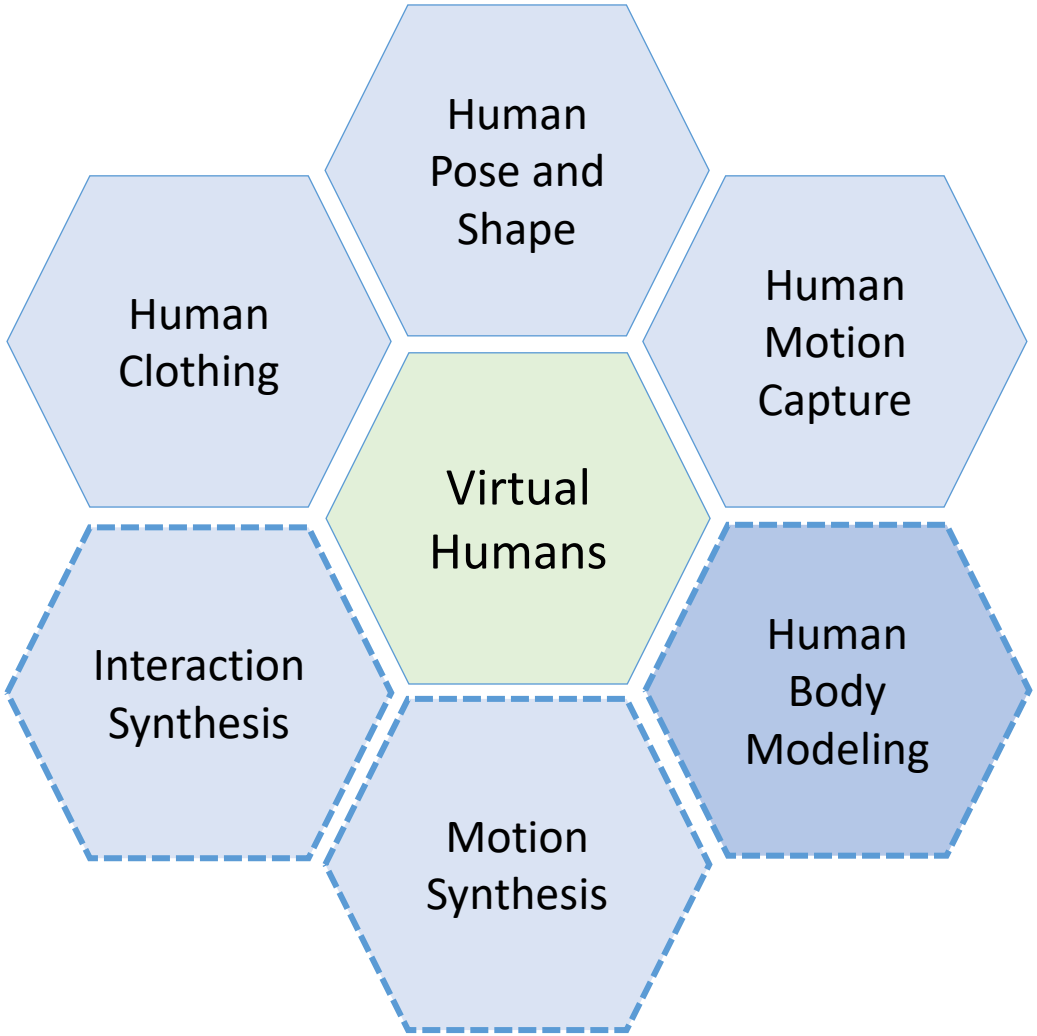


[1] DIMOS, *ICCV 2023*
[2] SAGA, *ECCV 2022*
[3] HALO, *3DV 2021*
[4] GraspingField, *3DV 2020*

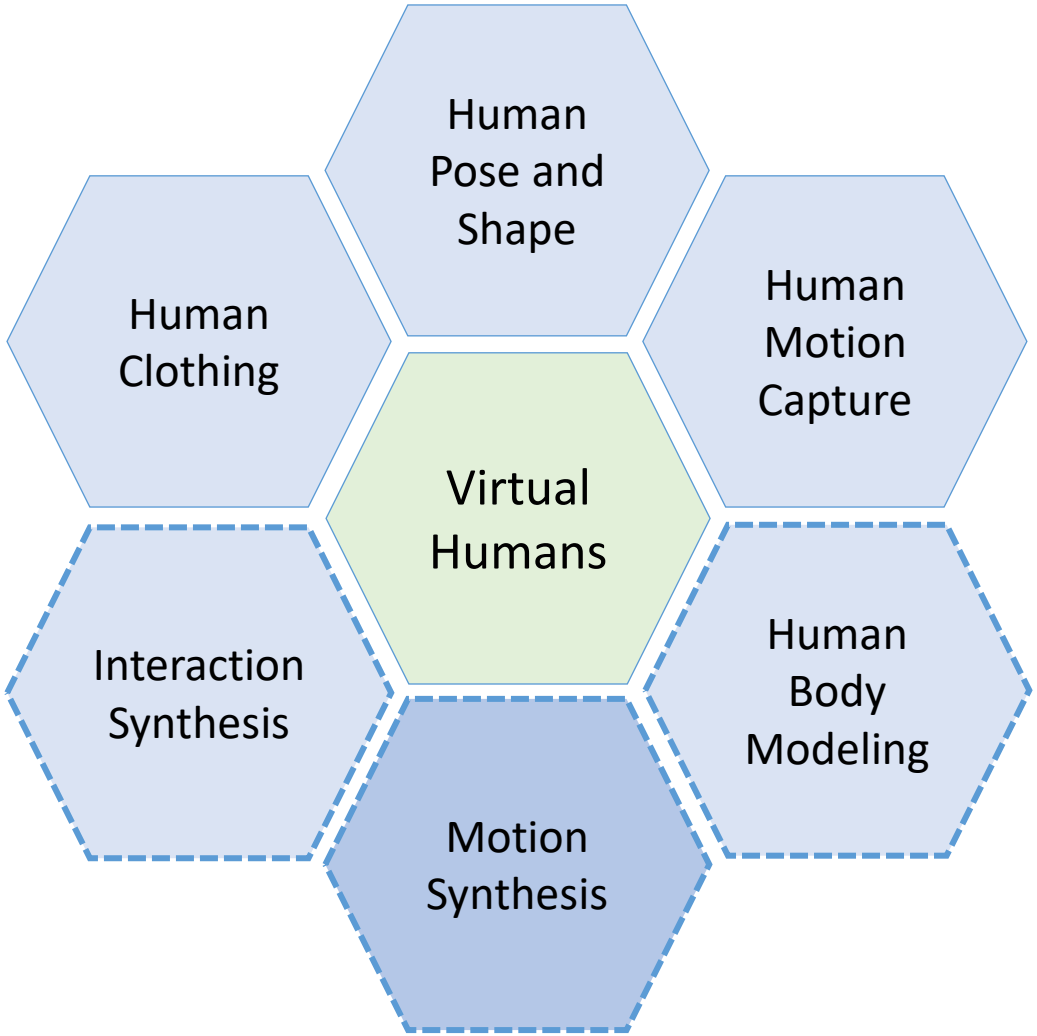
My work!



[1] MetaAvatar, *NeurIPS 2021*



“a person **stands up** and **walk clockwise** in circle then **sit** back down”

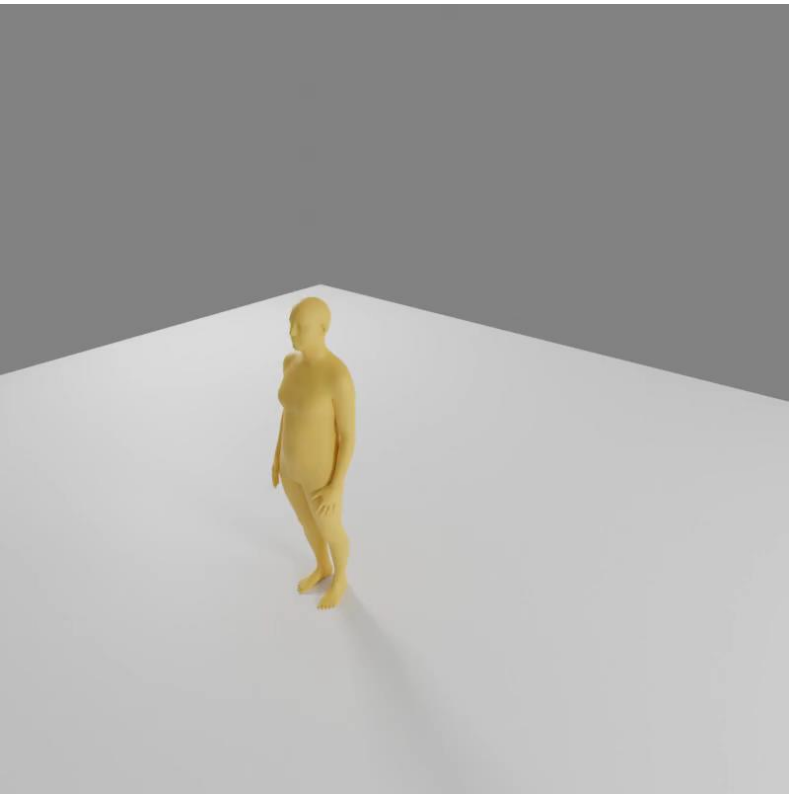


[1] GMD, *ICCV 2023*
[2] MOJO, *CVPR 2021*

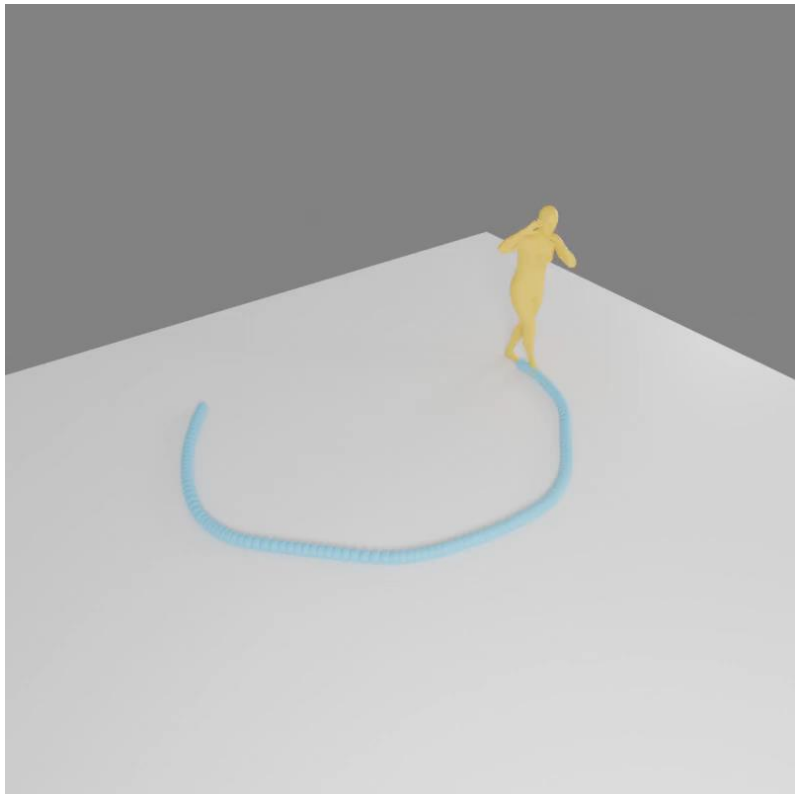
Guided Motion Diffusion

Synthesize motion from text with spatial objectives

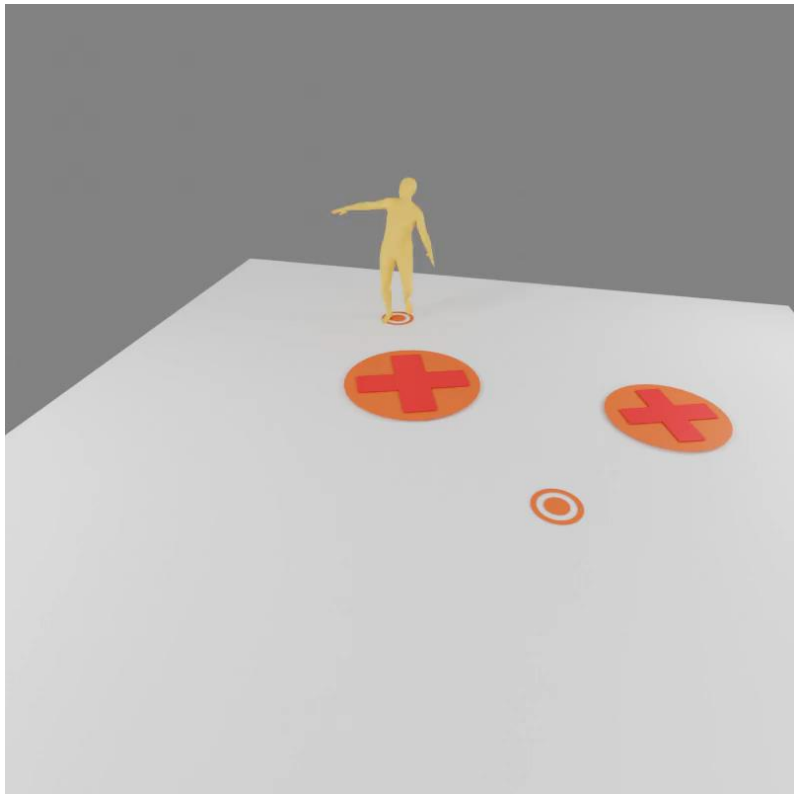
korrawe.github.io/gmd-project/



“a person **walks backward**”



“walk with hands up”



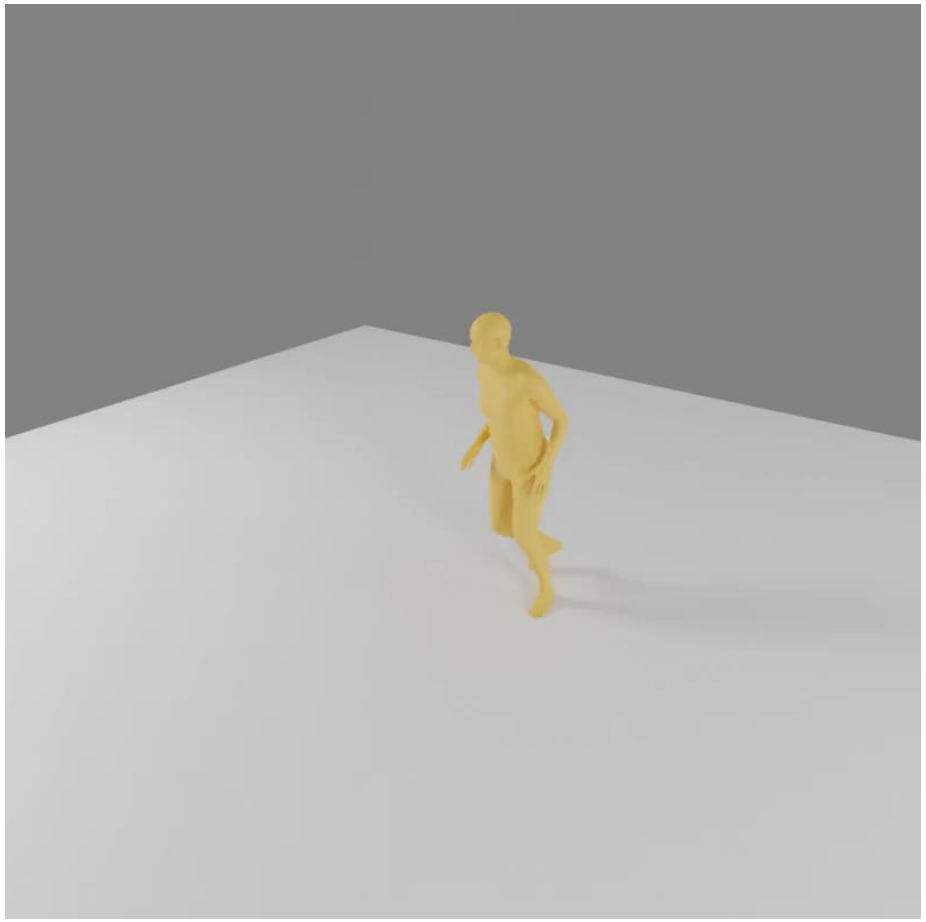
“dance”

[1] GMD, *ICCV 2023*

Guided Motion Diffusion

korrawe.github.io/gmd-project/

Try it for yourself at



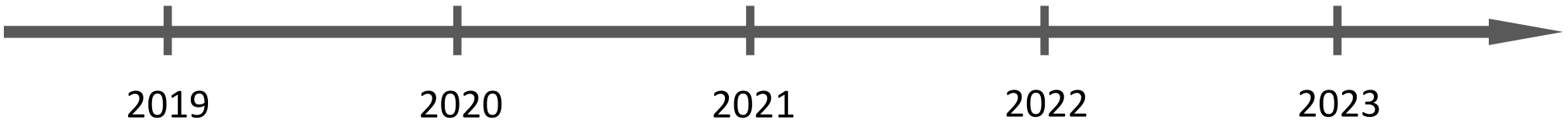
“a person runs in circle”

Now, let us put these projects in a wider context

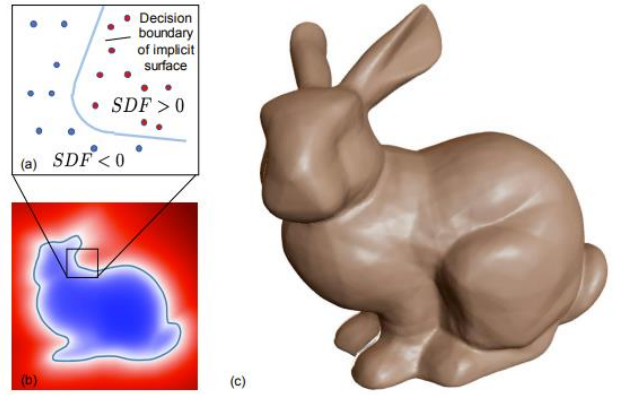
What else had happened in the fields?

What People Find Interesting in 3D Compute Vision

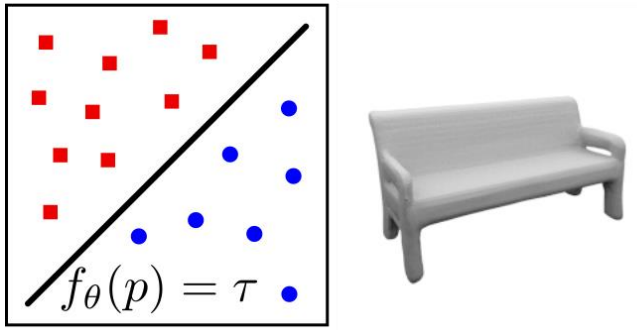
(From my perspective and only in my related fields)



- Implicit surface
  - 3D modeling
  - SDF
  - Occupancy Field



DeepSDF



Occupancy Networks

2019 2020 2021 2022 2023

- Implicit surface
  - 3D modeling
  - SDF
  - Occupancy Field

- Deformable
  Implicit surface

- Nerf
  - 3D modeling
  - Density function
    with color

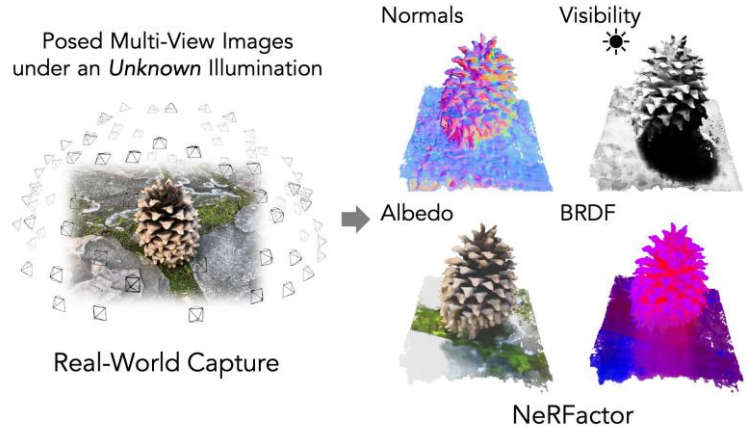
2019 | 2020 | 2021 | 2022 | 2023

- Deformable
  Implicit surface

Avatar Creation (with colors)

- Nerf
  - 3D modeling
  - Density function with color

Nerf with X



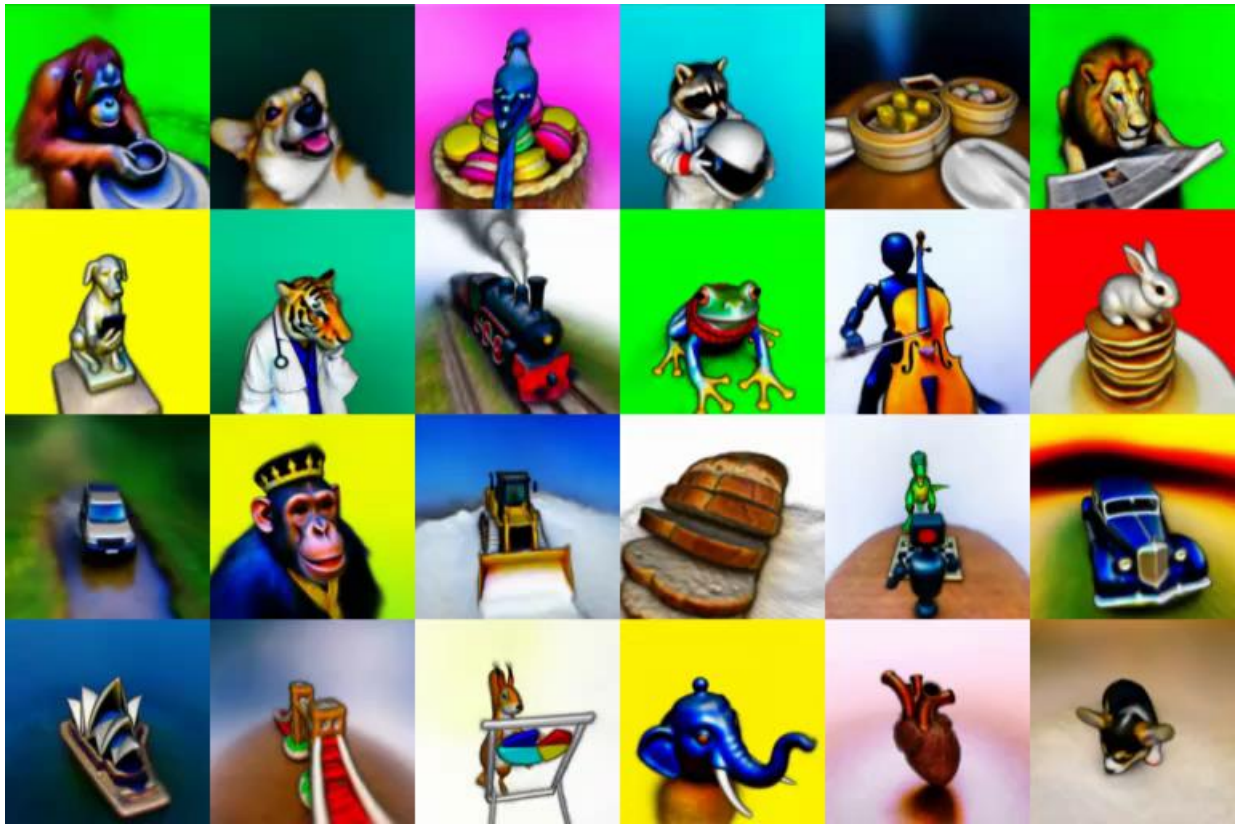
Diffusion models

2019 2020 2021 2022 2023

Diffusion models

Diffusion models for 3D

Diffusion models for motion



DreamFusion

2019 2020 2021 2022 2023

Open Challenges?

Open Challenges

- Avatar creation
  - Human with clothes
- Motion modeling
  - Human-human interaction
  - Human-scene interaction
  - Hand-object interaction
  - Describing “free” motions
- Motion capture
  - Out-of-frame movement
  - Occlusion
- Synthetic data
- And many more …

Open Challenges

- Avatar creation
  - Human with clothes
- Motion modeling
  - Human-human interaction
  - Human-scene interaction
  - Hand-object interaction
  - Describing “free” motions
- Motion capture
  - Out-of-frame movement
  - Occlusion
- Synthetic data
- And many more …

Open Challenges

- Avatar creation
  - Human with clothes
- Motion modeling
  - Human-human interaction
  - Human-scene interaction
  - Hand-object interaction
  - Describing “free” motions
- Motion capture
  - Out-of-frame movement
  - Occlusion
- Synthetic data
- And many more …

Open Challenges

- Avatar creation
  - Human with clothes
- Motion modeling
  - Human-human interaction
  - Human-scene interaction
  - Hand-object interaction
  - Describing “free” motions
- Motion capture
  - Out-of-frame movement
  - Occlusion
- Synthetic data
- And many more …

VirtualHome

Open Challenges

- Avatar creation
  - Human with clothes
- Motion modeling
  - Human-human interaction
  - Human-scene interaction
  - Hand-object interaction
  - Describing “free” motions
- Motion capture
  - Out-of-frame movement
  - Occlusion
- Synthetic data
- And many more …

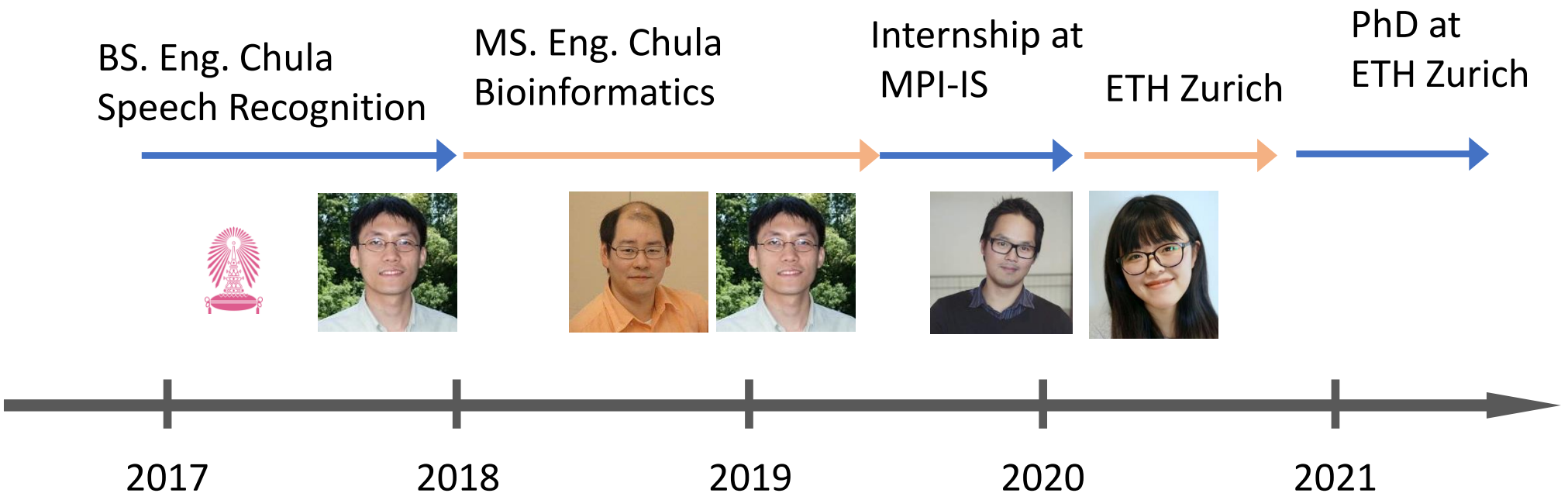
Question?

Non-technical Q&A

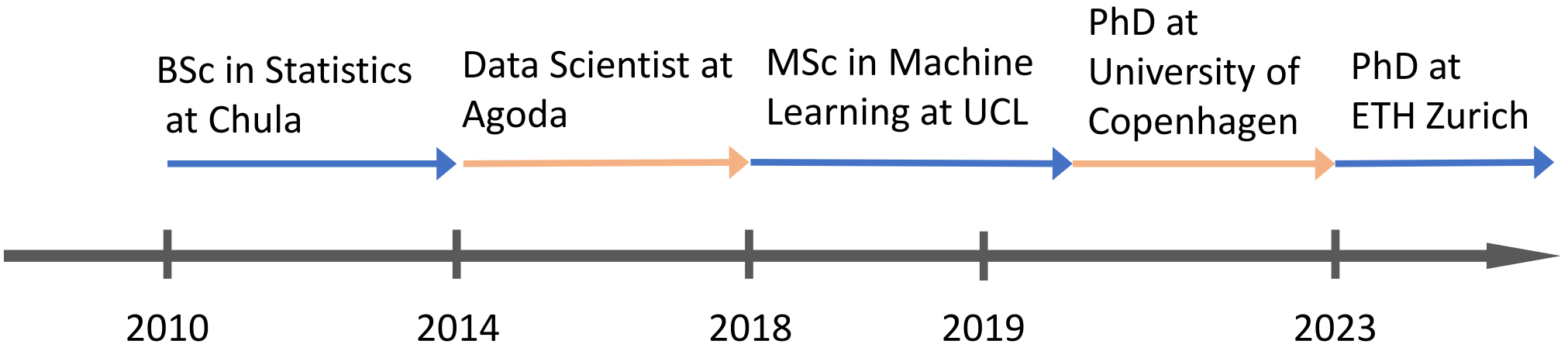
How Did I End Up Here?

- Why this lab? I don’t know…

Computer Vision



How Did I End Up Here?



Day-to-day Life

- Your PhD life is 80% depends on your advisor
- Everything is good as long as your work is done
- More supervision in 1st-2nd year (weekly 1-1 meeting)

For me:

- Flexible working hours (meaning: 12pm until sleep)
- Mostly remote
- Biking, gym, cooking

Struggles

- Language
- Loneliness?
- Depression
- Work-life balance
- Handling deadlines
- Handling paper rejections

The first paper is usually the hardest one. It will get better from there

Research Culture

Differences and suggestions

- Own your project project
- Take initiative. Find a way out
- Find someone to discuss ideas with

Soft Skills

Differences and suggestions

- Know how to ask others for help
  - But no one is responsible for your life!
- Making connections
  - Emailing people, GitHub, attending conferences
- Presentation skill
  - PhDs are usually the least busy person in the food chain so respect other's time
  - Understand your audience
- Time management
- Know when to take a break

Applying for Positions (including internship)

You either have

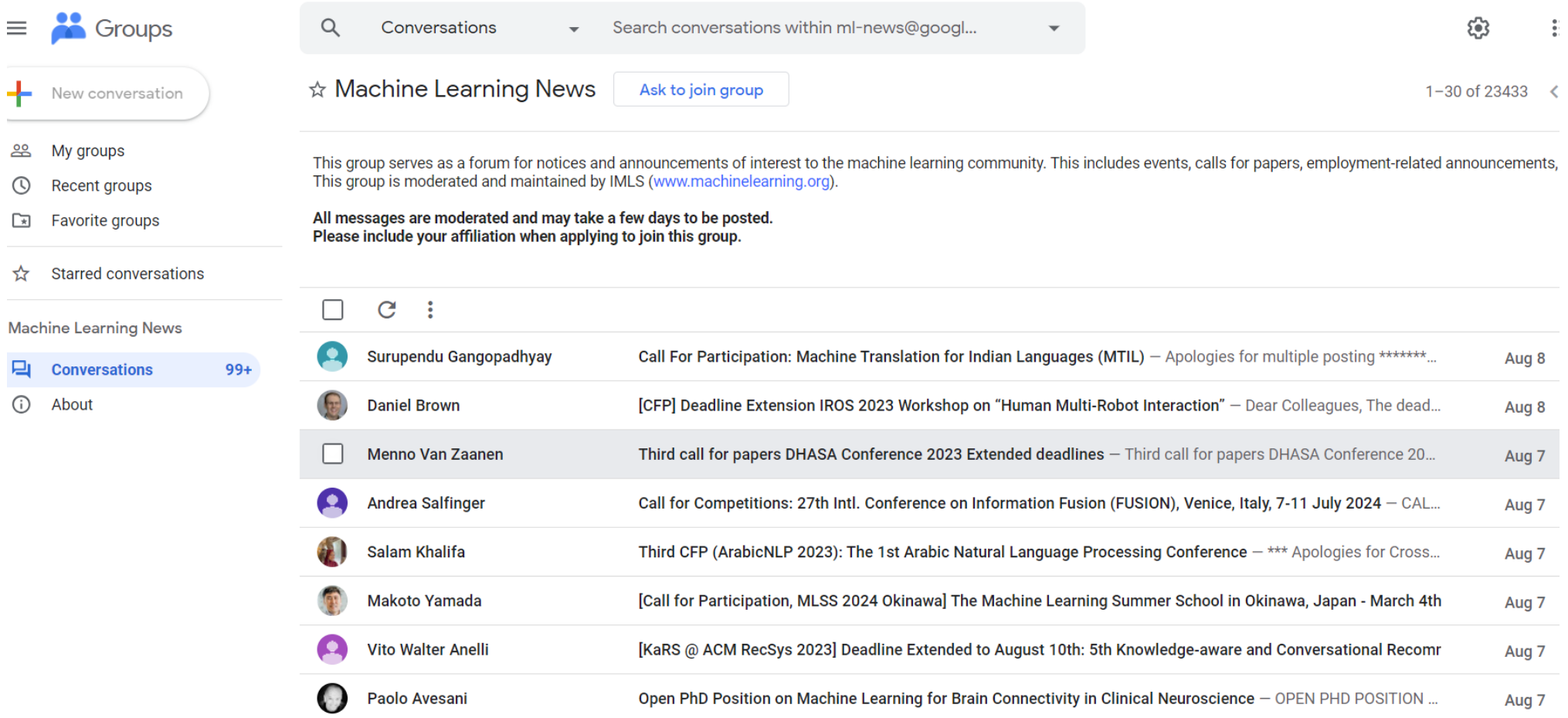
- Paper(s)
- Or connections and someone credible who can recommend you

Then apply and email them

The applying process differs from place to place.
(so there is no standard process)

Keep an eye on: https://groups.google.com/g/ml-news

Applying for Positions (including internship)



Keep an eye on: https://groups.google.com/g/ml-news

Applying for ETH

- Master's degree
  - 2 years
  - Strongest University in Continental Europe (we don't care about islands here) (by some ranking)
  - Essentially free
  - Cost of living around 1500 CHF ($1700 / 60k THB) per month

- PhD
  - Strong labs in ML (Robotics, Computer Vision, theory, etc.)
  - Comfortable life (in term of money)
  - Very few mandatory course works
  - To apply: Contact advisors first

Applying for ETH

- ETH AI Center
- Max Planck ETH Center for Learning Systems
- Zurich Graduate School in Mathematics