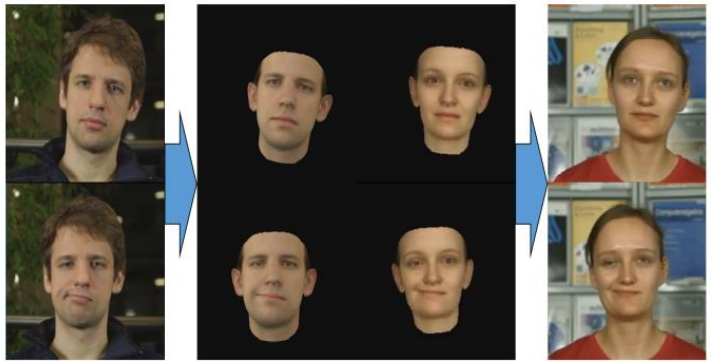
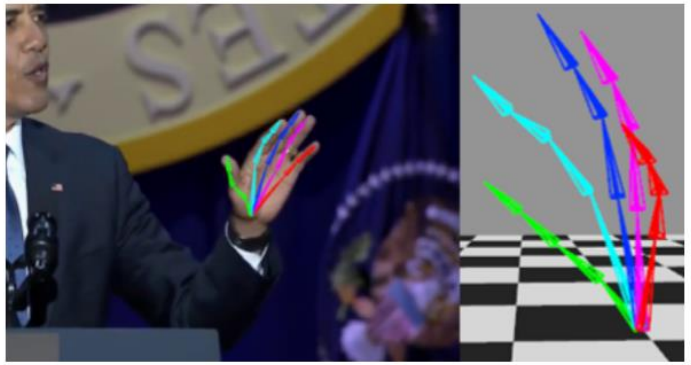



Computer Vision and Machine Learning for Computer Graphics

SS2019



Christian Theobalt
Mohamed Elgharib
Vladislav Golyanik



Graphics, Vision and Video Group, MPI Informatik

Graphics Vision & Video

max planck institut informatik

Overview

- Organization
- Introduction
- Topics
- Summary

Overview

- Organization
- Introduction
- Topics
- Summary

Organizers

Christian Theobalt
MPI Informatik, room 228
theobalt@mpi-inf.mpg.de

Mohamed Elgharib
MPI Informatik, room 218
elgharib@mpi-inf.mpg.de

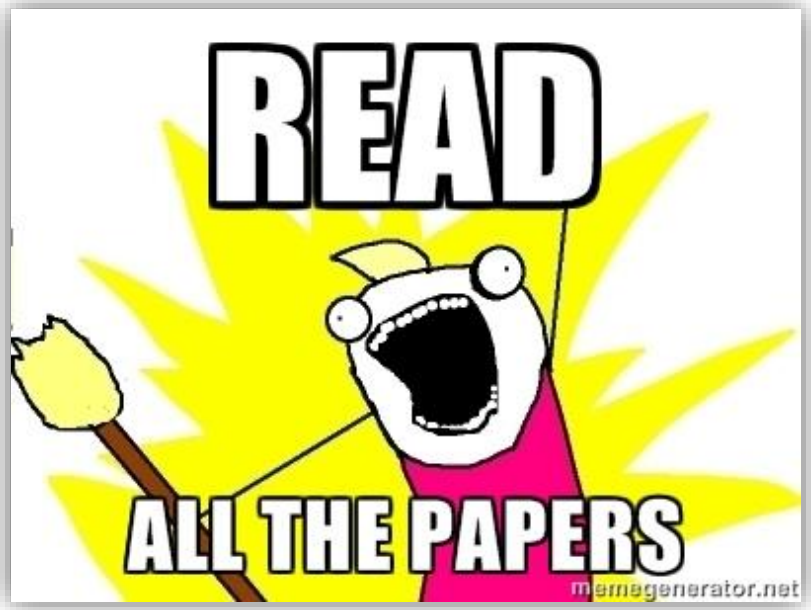
Vladislav Golyanik
MPI Informatik, room 219
golyanik@mpi-inf.mpg.de

Basic Coordinates

- Time: Thursdays, 14:15 – 15:45
- Place: MPI Informatik (E1 4), room 021
- Website:
  - http://gvv.mpi-inf.mpg.de/teaching/gvv_seminar_2019/

Formal requirements in a nutshell

- You read all the papers
- Your presence is required!
  - We will monitor attendance.
- Submit questions & participate in discussion
- One topic is “Your Topic” (2 papers):
  - Deliver a 40 minute presentation
  - Write a 5–7 page report
- Grade: talk 30%, discussion 30%, report 40%



Prior knowledge

- Not for beginners in visual computing
- You need experience in:
  - computer vision
  - computer graphics
  - geometric modeling
  - basic numerical methods
- Examples: you should know how …
  - … a camera is modeled mathematically
  - … 3D transformations are described
  - … a system of equations is solved, etc.

Organization

- 26 topics to choose from
  - listed on seminar website + introduced later today
- Maximum of 11 presentation slots:
  - First presentation: Thursday, 25 April 2019
  - Each week until Thursday, 18 July 2019 (including)
- Each topic comes with a supervisor:
  - You can ask questions by e-mail at any time
    - about your topic, the papers, your presentation and report
  - Up to one office hour per week

Presentations

- Order of presentation will be determined after topic assignment
  - Slots can be swapped if necessary: talk to other participants first
- About 40 minutes long:
  - Introduction (about 5 minutes):
    - summary of previous week
    - finding themes that join the two papers
  - Technical content (about 35 minutes):
    - presentation of the two papers
    - again finding the common links between the papers
- Public feedback from other students after discussion

Suggested presentation preparation

- Schedule two meetings with your supervisor:
  - First meeting: 2–3 weeks before presentation
    - Read the papers for this meeting
    - Ask questions if you have difficulties
    - Discuss your plans for presentation
  - Second meeting: 1 week before presentation
    - Prepare a preliminary presentation (not a full rehearsal)
    - We can provide feedback
  - It is your responsibility to arrange the meetings
  - Do not rely on us providing last-minute feedback

Discussion (45–60 minutes)

- Day before the seminar:
  - Submit 2+ questions for discussion to golyanik@mpi-inf.mpg.de
  - Important: your contribution will be marked
- At the seminar:
  - One person assigned in advance to lead the discussion
  - Will get the collected questions submitted before the seminar
  - Gives summary of the talk
  - Moderates and guides discussion
  - Raises open questions that remain
  - Discussion of the strengths and weaknesses of the two papers
  - This will also be marked

Report

- 5–7 page summary of the major ideas in your topic:
  - 3–4 pages on the two papers
  - 3–4 additional paper references
  - 2–3 pages with your own ideas, for example:
    - Limitations not mentioned in the paper + sketch of potential solution
    - Try to suggest improvements
    - Novel ideas based on content described in the papers
    - Can be the result of the discussion after your presentation
- The idea is that you get a feeling for your specific topic surpassing the level of simply understanding a paper.

Report

- Due date: **Friday, 15 August 2019**
  (4 weeks after the last seminar)
- Send PDF by e-mail
- We will provide a LaTeX template on the seminar website
  - If you use other software, make it look like the LaTeX template
    - this is your responsibility
  - Strongly recommended to learn LaTeX
    - used by nearly all research papers in visual computing

Grading scheme

- **Presentation** (overall: 30%)
  - Form (30%): time, speed, structure of slides
  - Content (50%): structure, story line and connections, main points, clarity
  - Questions (20%): answers to questions
- **Discussion** (overall: 30%)
  - Submitted questions (33%): insight, depth
  - Participation (33%): willingness, debate, ideas
  - Moderation (33%): strengths and weaknesses, integration of questions
- **Report** (overall: 40%)
  - Form (10%): diligence, structure, appropriate length
  - Context (20%): the big picture, topic in context
  - Technical correctness (30%)
  - Discussion (40%): novelty, transfer, own ideas / in own words

Benefits to you

- Practise important skills in research
  - Read and understand technical papers
  - Present scientific results and convince other people
  - Analyse and develop new ideas through discussions
- Discussion is essential:
  - If you don’t participate, you miss a big chance
  - Most ideas are developed in discussions about other papers
- Therefore:
  - Prepare for the seminar classes
  - Participate actively in the discussions
  - Benefit from the interaction in the group

What this seminar is not …

- A course to just sit and listen
  - Come prepared
  - Read all papers before class, think about problems, submit questions and discuss them in class
  - Your participation benefits everyone
    - the group makes the seminar
- "Cheap" credit points
  - Don't underestimate the time it takes to understand a paper, prepare a talk, and write a report
  - So please do take it seriously!

Schedule

- **11 April** – Introduction ◄ You are here
- **18 April** – Lectures:
  - "How to read an academic paper"
  - "How to give a good talk"
- **25 April** – First presentation by a student
- … 10 more weekly presentations …
- **18 July** – Last presentation by a student
- **15 August** – Report deadline

Overview

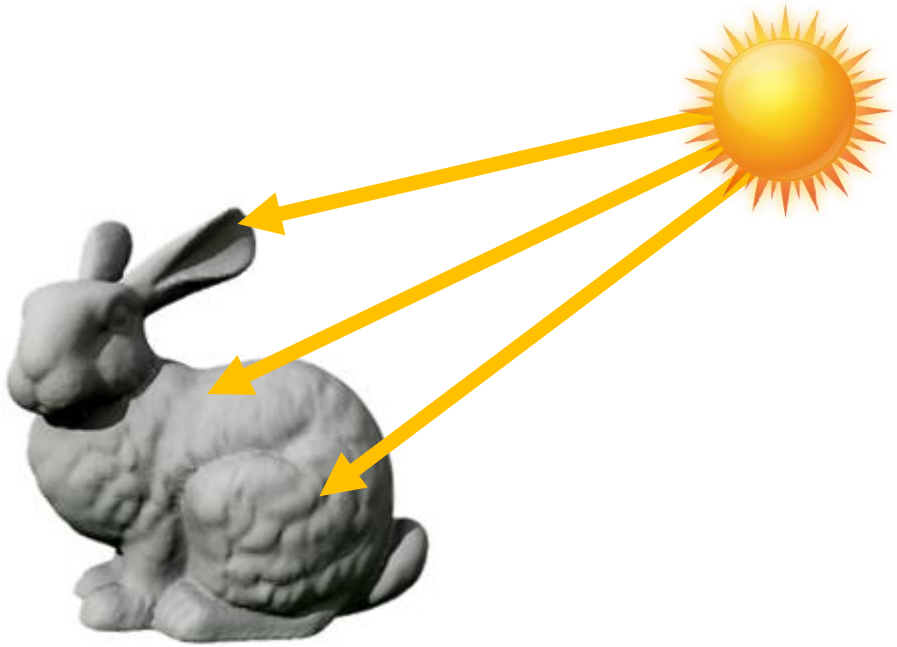
- Organization
- Introduction
- Topics
- Summary

Basics (Image Formation)

Basics (Image Formation)

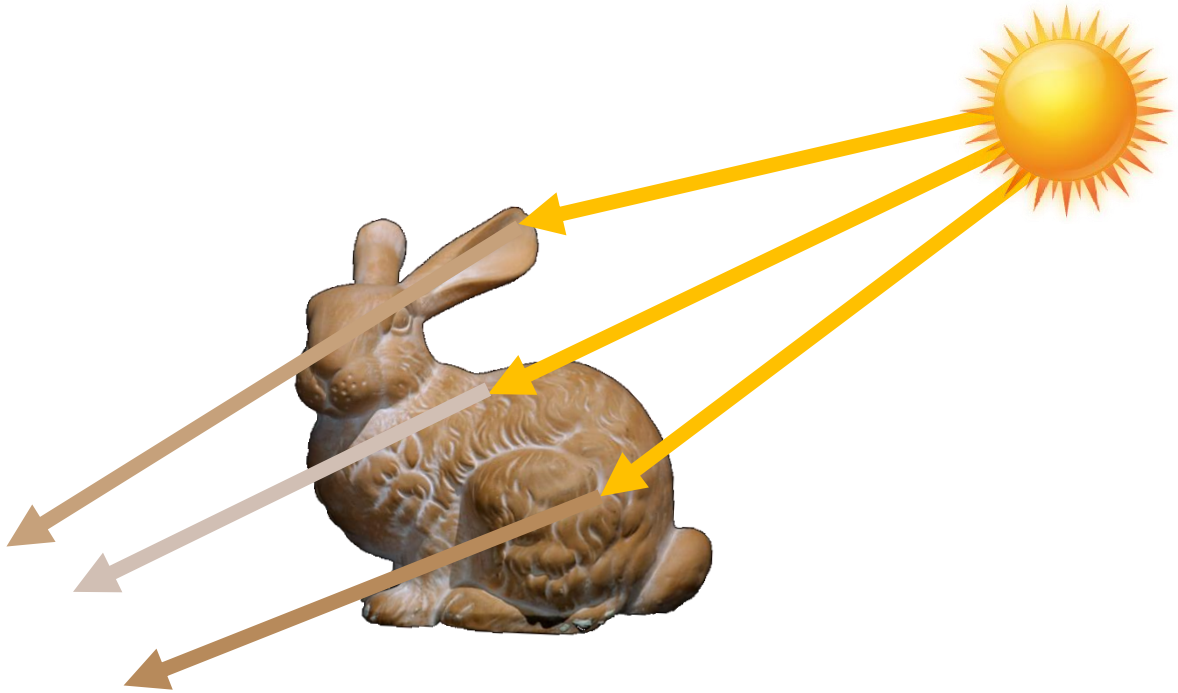
- **Geometry**

Basics (Image Formation)



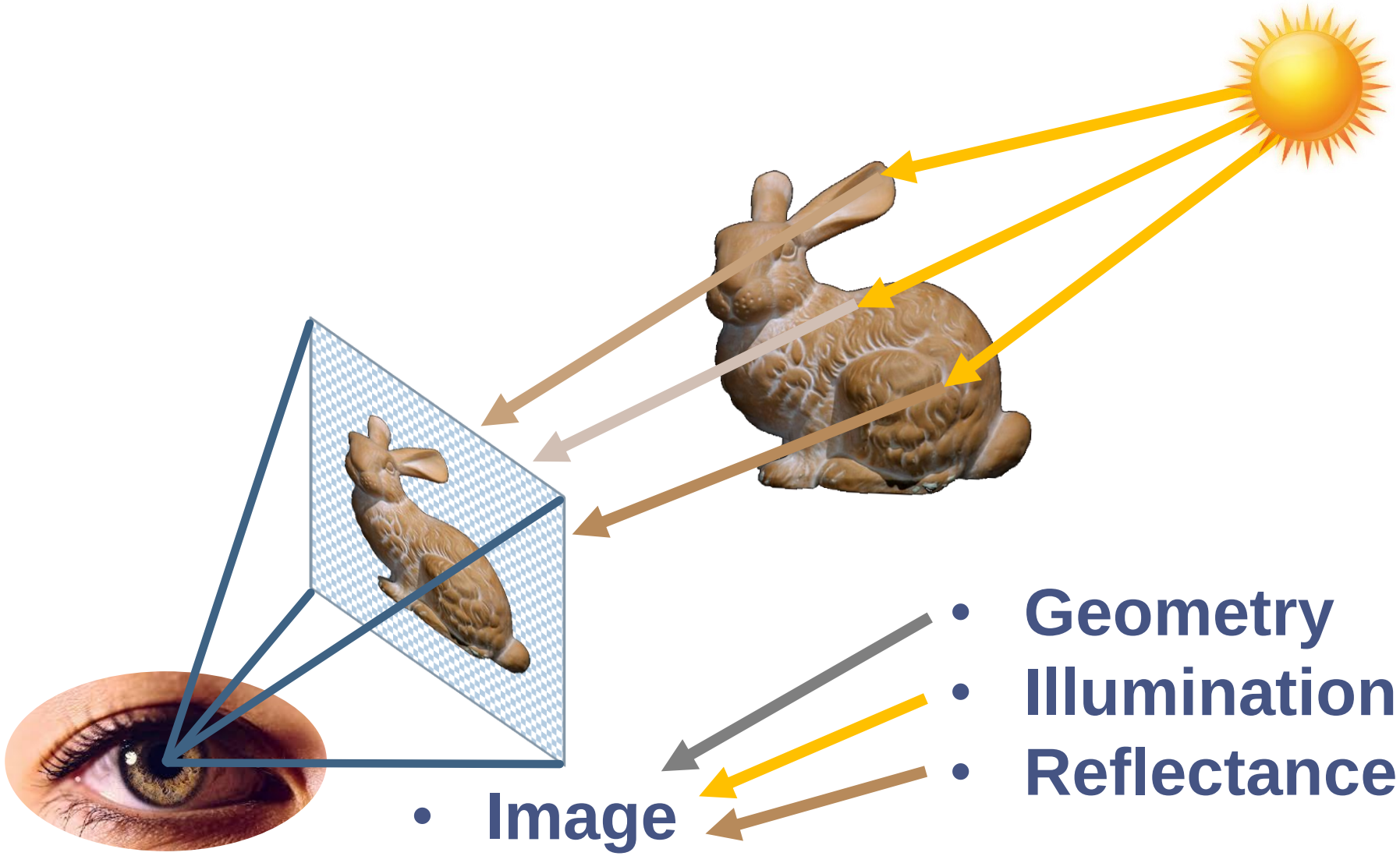
- **Geometry**
- **Illumination**

Basics (Image Formation)



- **Geometry**
- **Illumination**
- **Reflectance**

Basics (Image Formation)



- Geometry
- Illumination
- Reflectance

Basics (Image Formation)

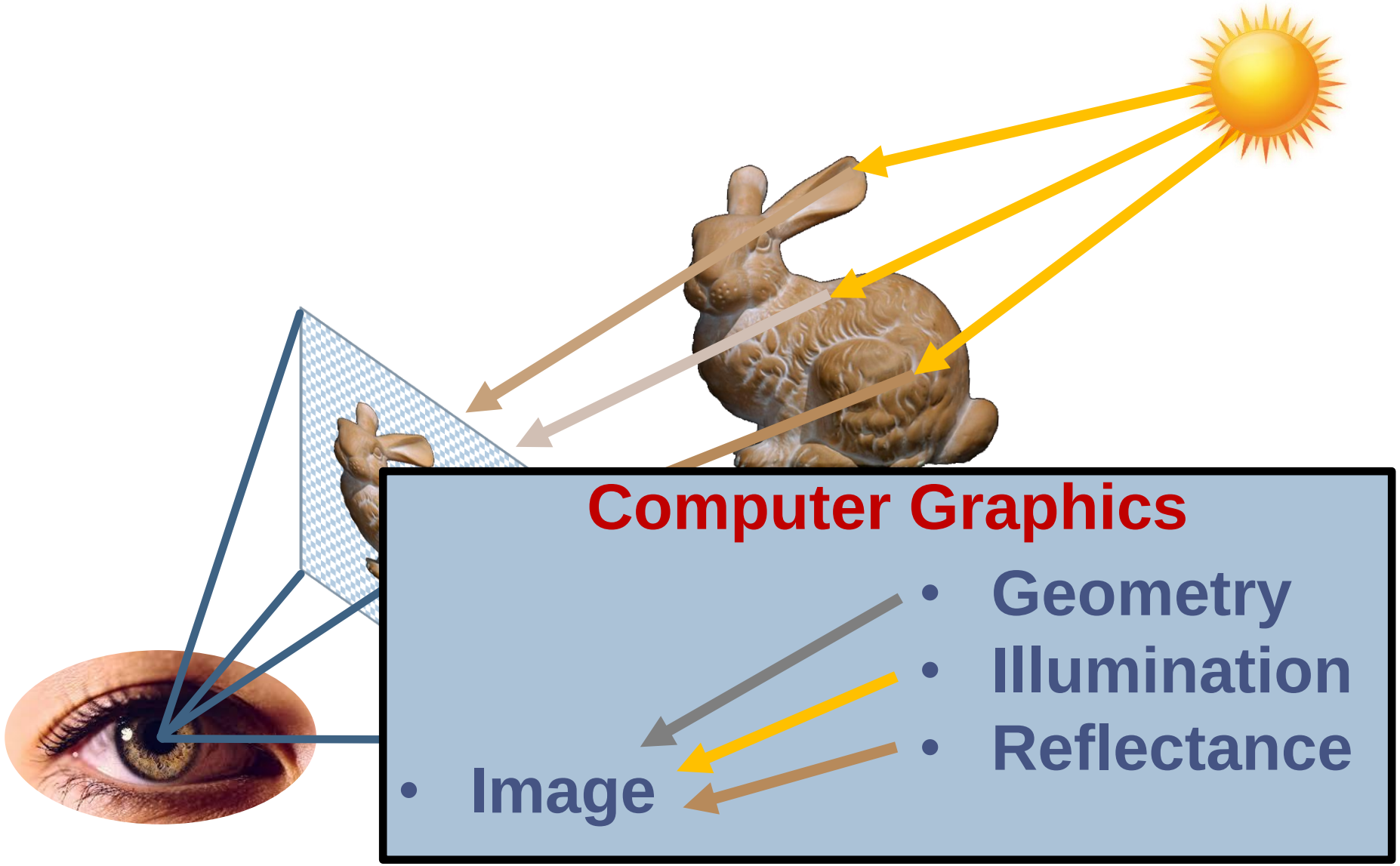
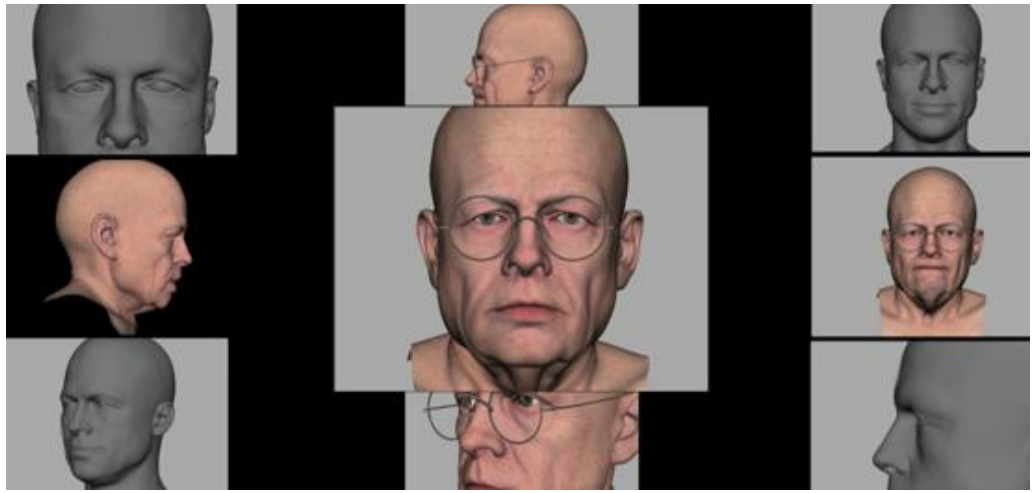


Photo-real virtual humans

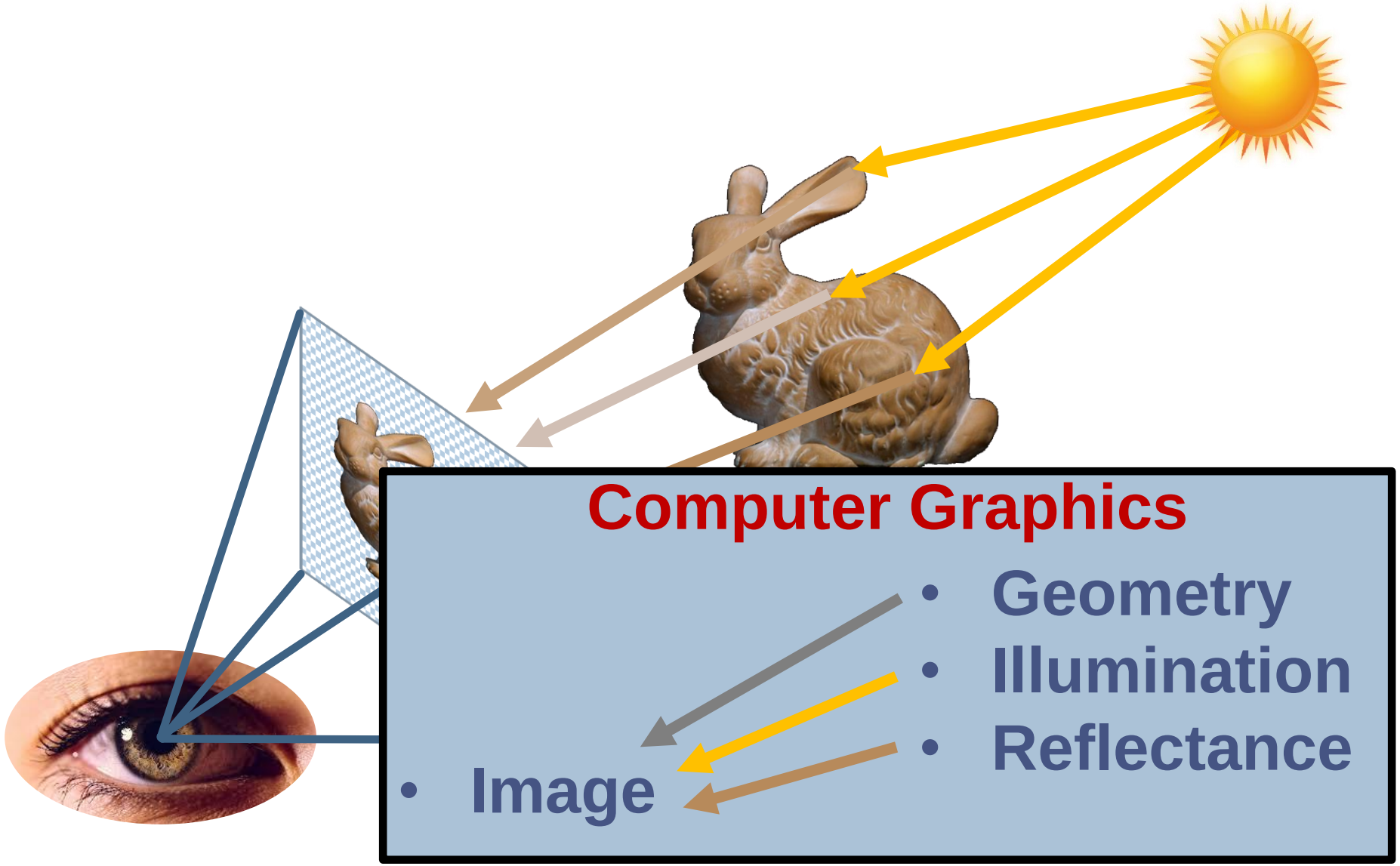
The Curious Case of Benjamin Button, 2008



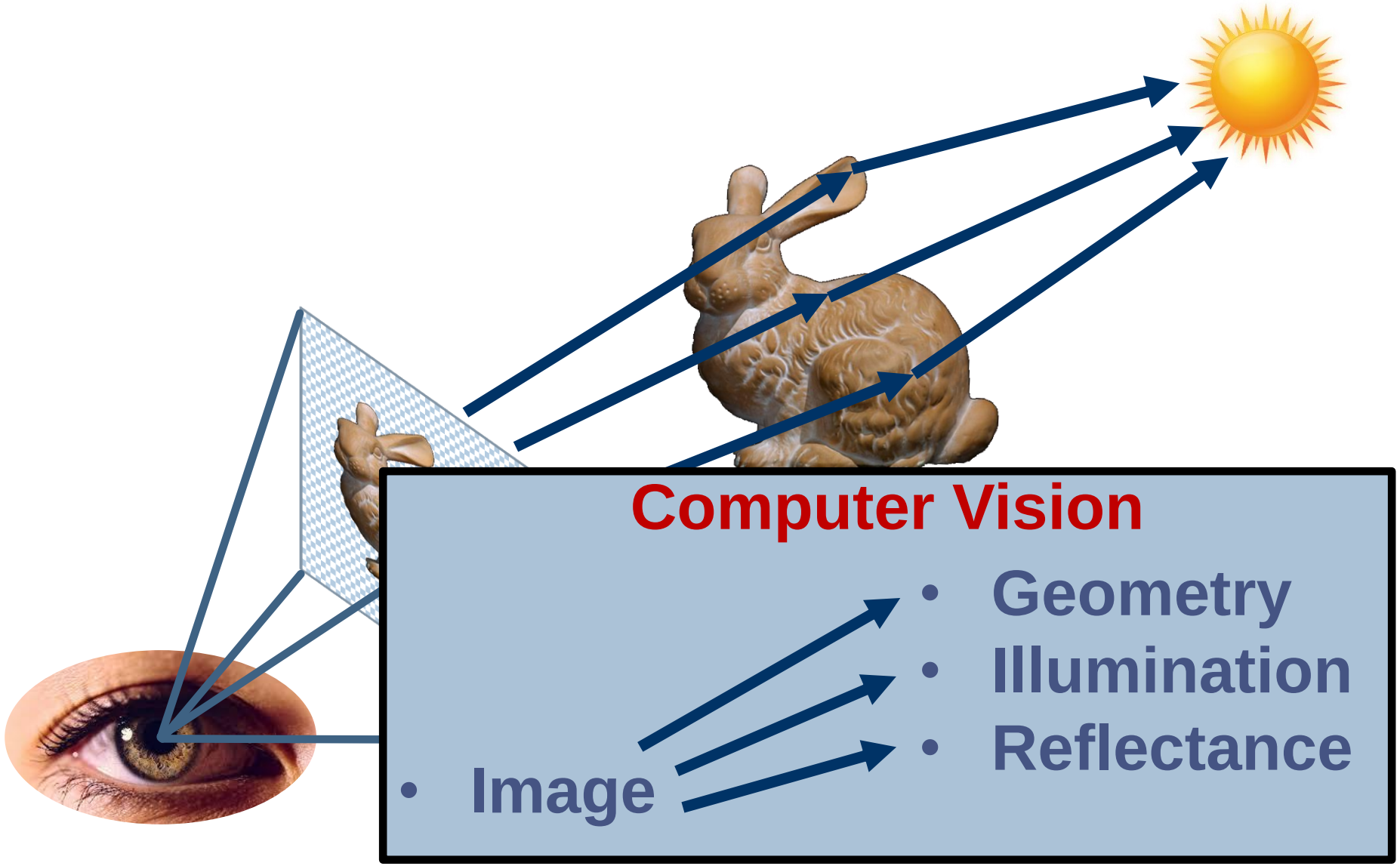
Real or rendered?

Real or rendered?

Basics (Image Formation)



Basics (Image Formation)



Computer Graphics / Computer Vision

Real world

- Images
- Videos
- Sensor data
- …

Computer Graphics

Computer Vision

Scene model

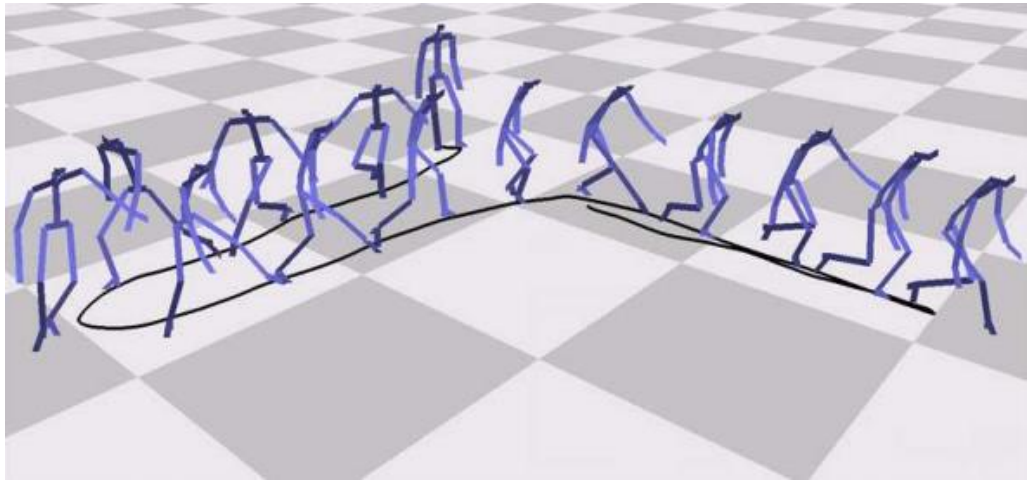
- Geometry
- Illumination
- Reflectance
- Motion

Overview

- Organization
- Introduction
- Topics
- Summary

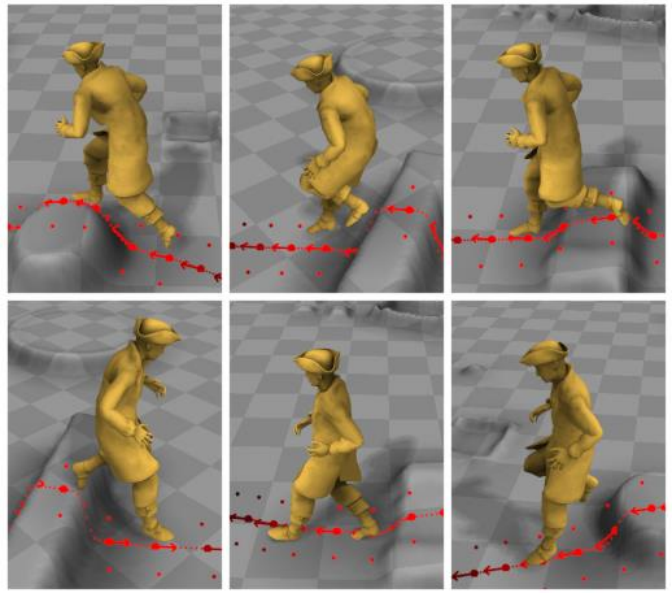
Human motion generation and control

Motion Graphs
(Kovar et al., SIGGRAPH 2002)



Motion control by finding closest transition points in the database

Phase-Functioned Neural Networks for Character Control
(Holden et al., SIGGRAPH 2017)

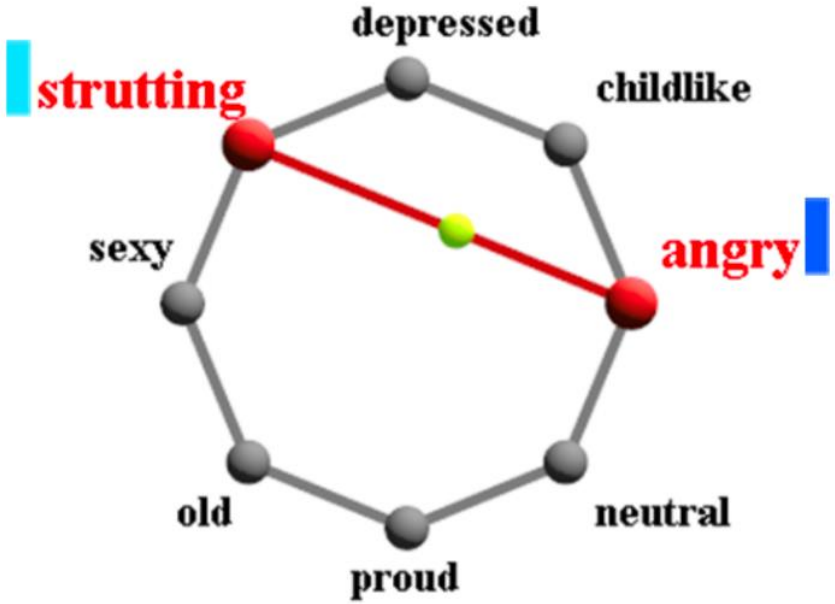


NN-based motion synthesis with a novel disambiguation approach to allow real-time control on various terrains

Style Transfer for Human Motion

Realtime Style Transfer for Unlabeled Heterogeneous Human Motion

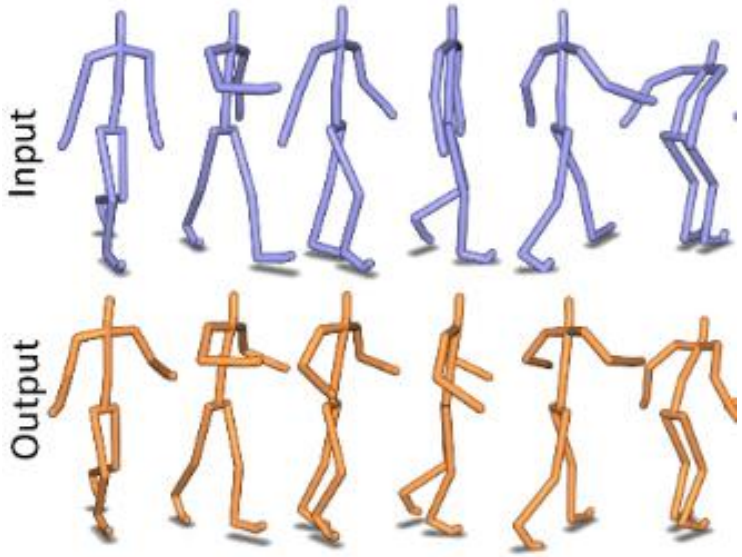
(Xia et al., SIGGRAPH 2015)



Real-time motion style transfer using a mixture of autoregressive models based on temporally local nearest neighbors

Spectral Style Transfer for Human Motion between Independent Actions

(Yumer and Mitra, SIGGRAPH 2016)



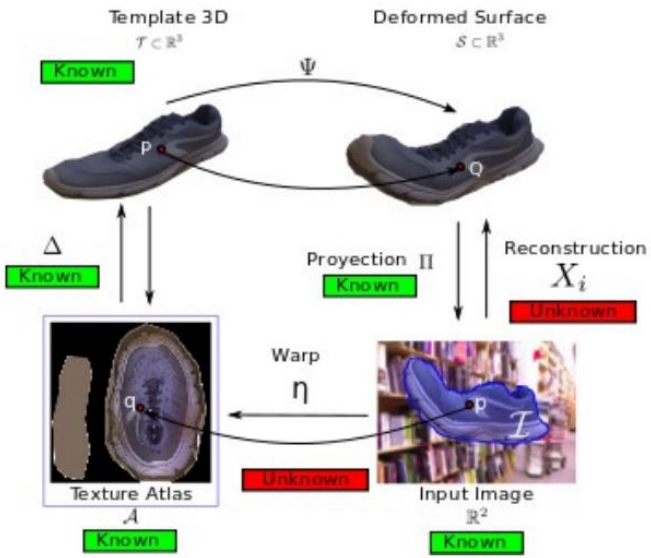
Motion style transfer by exploiting correlation of the difference between stylized motion in the spectral domain

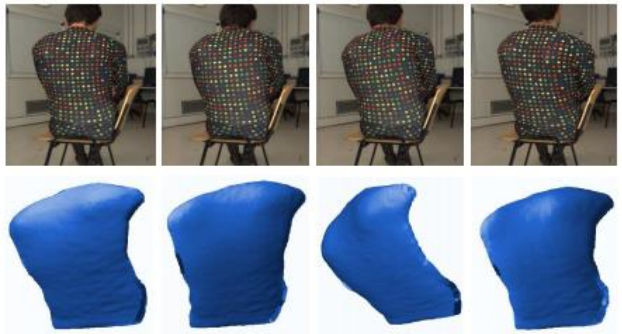
Monocular Non-Rigid 3D Reconstruction

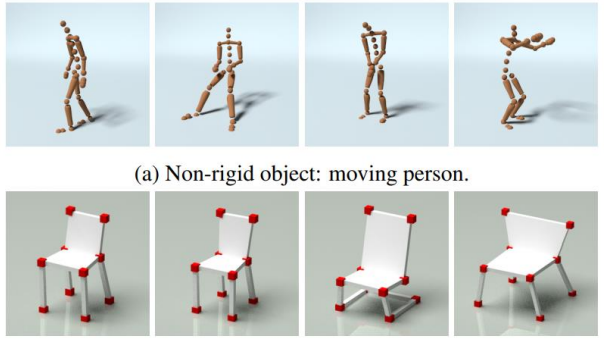
Scalable Dense Non-rigid Structure-from-Motion: A Grassmannian Perspective
(Kumar *et al.*, CVPR 2018)

Deep Shape-from-Template: Wide-Baseline, Dense and Fast Registration and Deformable Reconstruction from a Single Image
(Fuentes-Jimenez *et al.*, ArXiv.org, 2018)

Deep Interpretable Non-Rigid Structure from Motion
(Kong and Lucey, ArXiv.org, 2019)







(a) Non-rigid object: moving person.

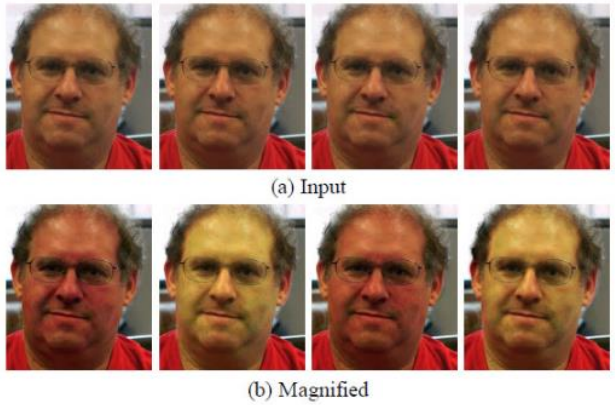
(b) Object category: chair.

Supervisor: Vladislav

Video Motion Magnification

Eulerian Video Magnification for Revealing Subtle Changes in the World

(Wu *et al.*, SIGGRAPH 2012)

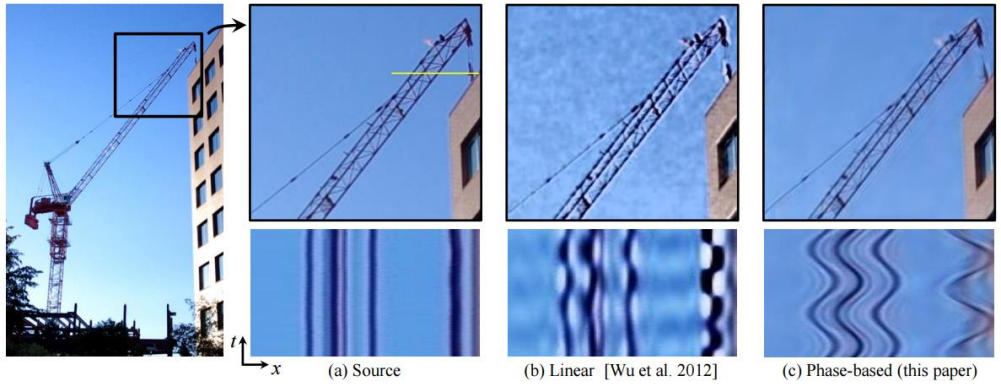


Phase-Based Video Motion Processing

(Wadhwa *et al.*, SIGGRAPH 2013)

Learning-based Video Motion Magnification

(Oh *et al.*, ECCV 2018)

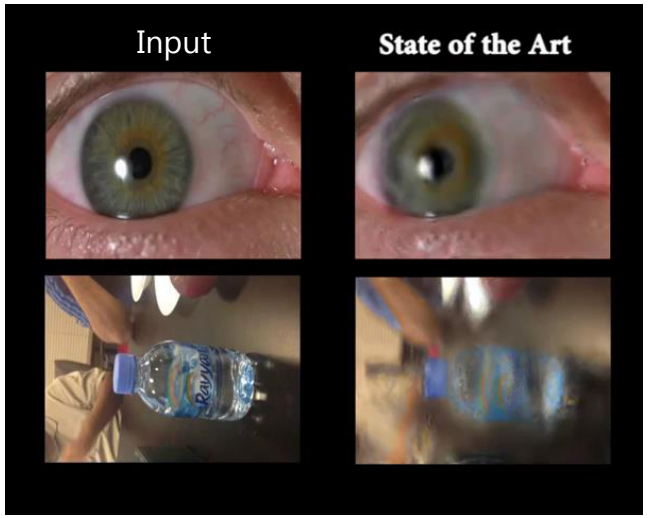


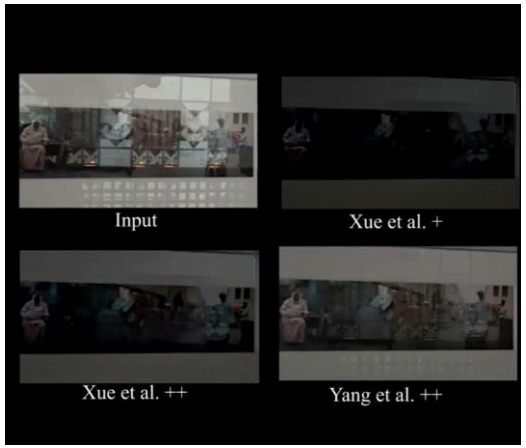
Supervisors: Vladislav, Mohamed

Motion Utilization for Computational Videography

Motion Magnification in Presence of Large Motions

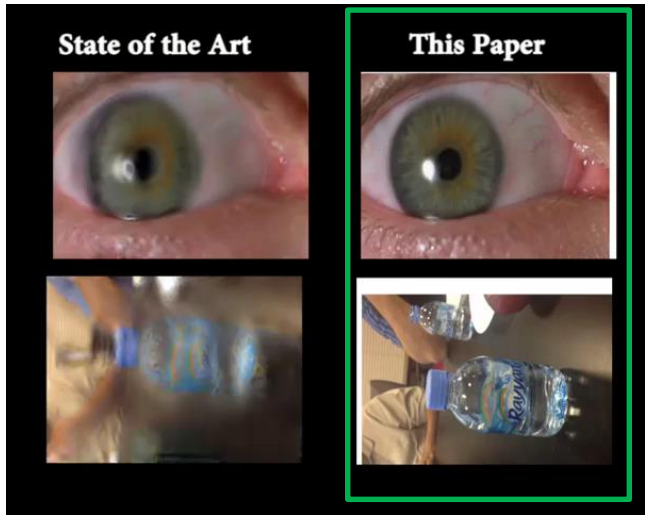
(Elgharib *et al.*, CVPR 2015)

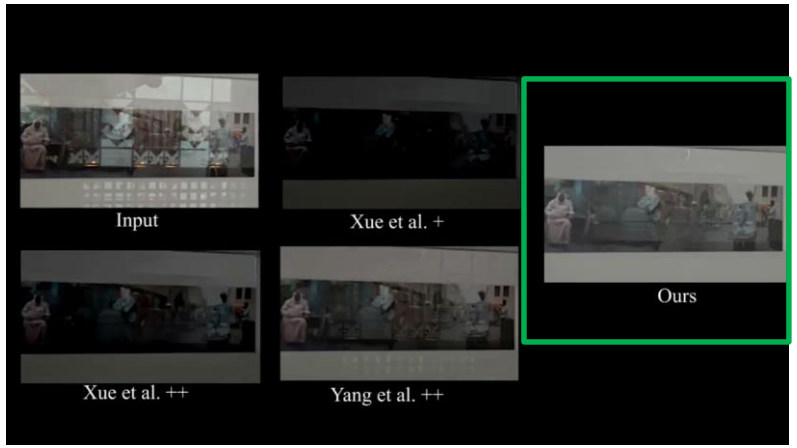




Video Reflection Removal through Spatio-temporal Optimization

(Nandoriya and Elgharib *et al.*, ICCV 2017)



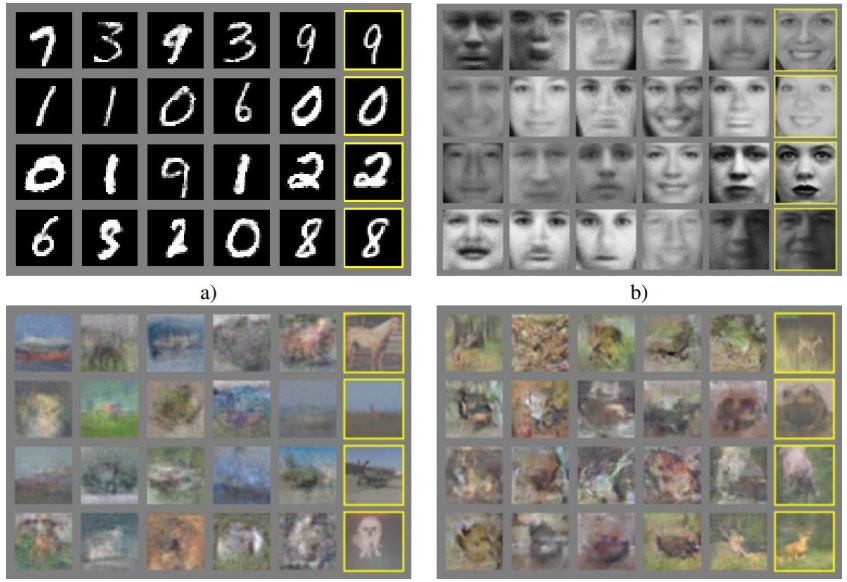


Supervisor: Mohamed

Generative Adversarial Networks and their Application for Video Face Manipulation

- **Generative Adversarial Nets**
  - *Goodfellow et al.*, NIPS 2014

- **Deep Video Portraits**
  - *Kim et al.*, SIGGRAPH 2018

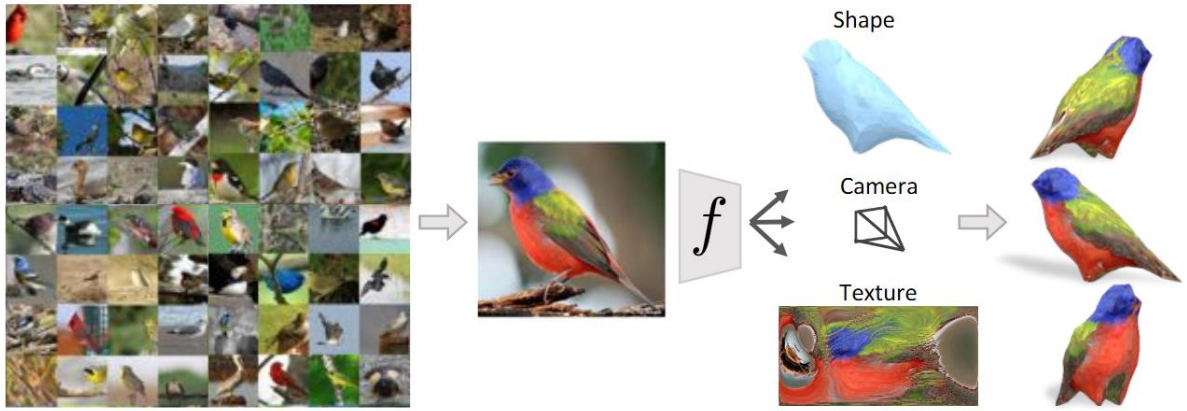


Input Output

General Mesh Reconstruction from Single Images

Learning Category-Specific Mesh Reconstruction from Image Collections

(Kanazawa et al., ECCV 2018)



Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images

(Wang et al., ECCV 2018)

Supervisor: Edgar

Real-time Non-rigid General Reconstruction

Graphics Vision & Video

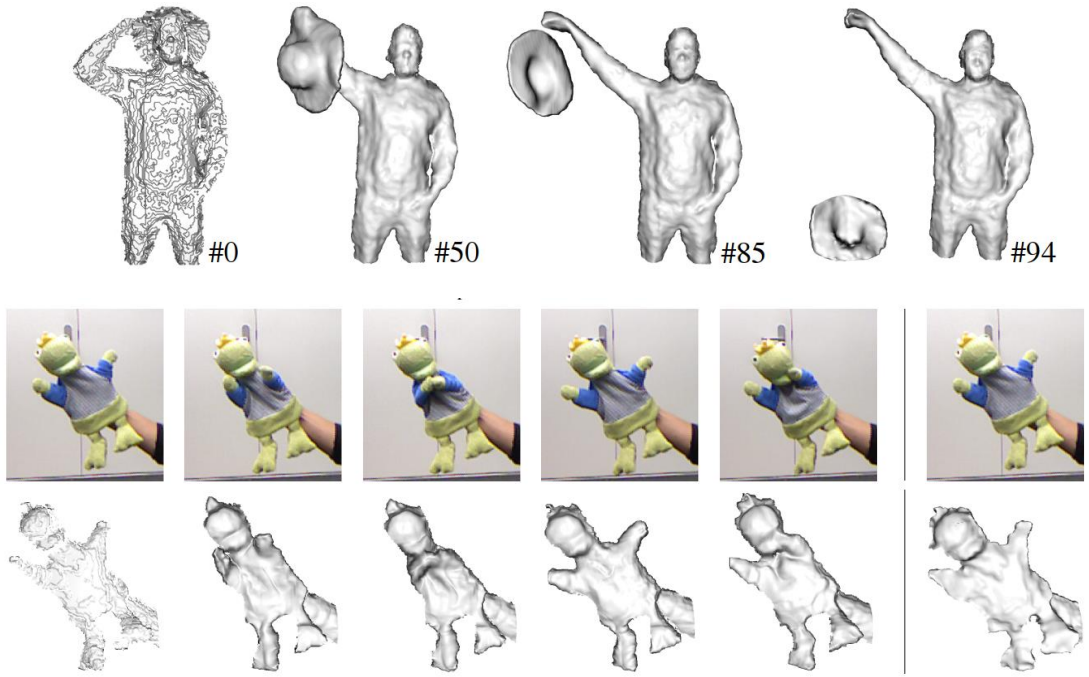
max planck institut informatik

VolumeDeform: Real-time Volumetric Non-rigid Reconstruction

(Innmann et al., ECCV 2016)

KillingFusion: Non-rigid 3D Reconstruction without Correspondences

(Slavcheva et al., CVPR 2017)



Supervisor: Edgar

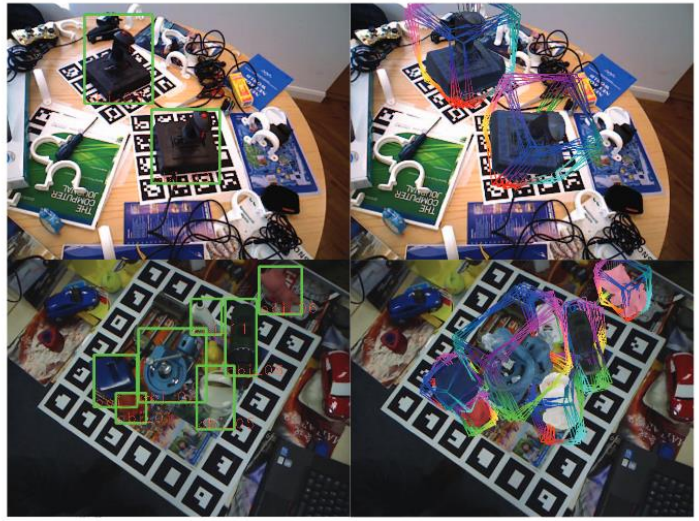
Object Pose Estimation with Neural Networks

Graphics Vision & Video

max planck institut informatik

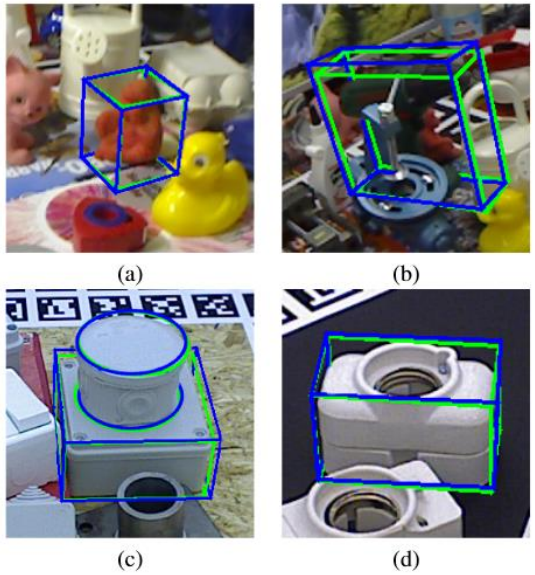
SSD-6D: Making RGB-Based 3D Detection and 6D Pose Estimation Great Again

(Kehl et al., ICCV 2017)



BB8: A Scalable, Accurate, Robust to Partial Occlusion Method for Predicting the 3D Poses of Challenging Objects without Using Depth

(Rad et al., ICCV 2017)

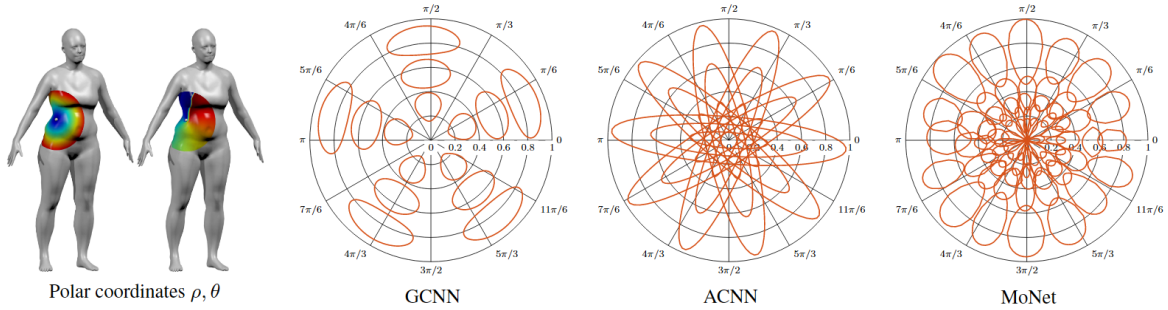


Supervisor: Edgar

Graph Convolutions in Neural Networks

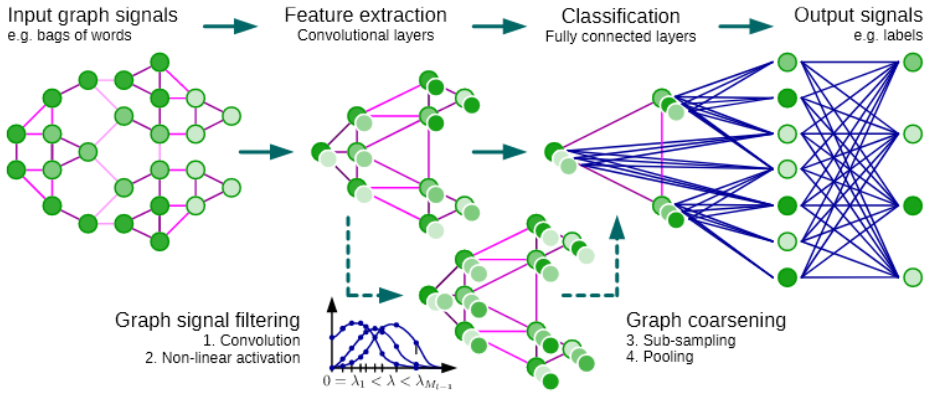
Geometric Deep Learning on Graphs and Manifolds Using Mixture Model CNNs

(Monti et al., CVPR 2017)



Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

(Defferrard et al., NIPS 2016)



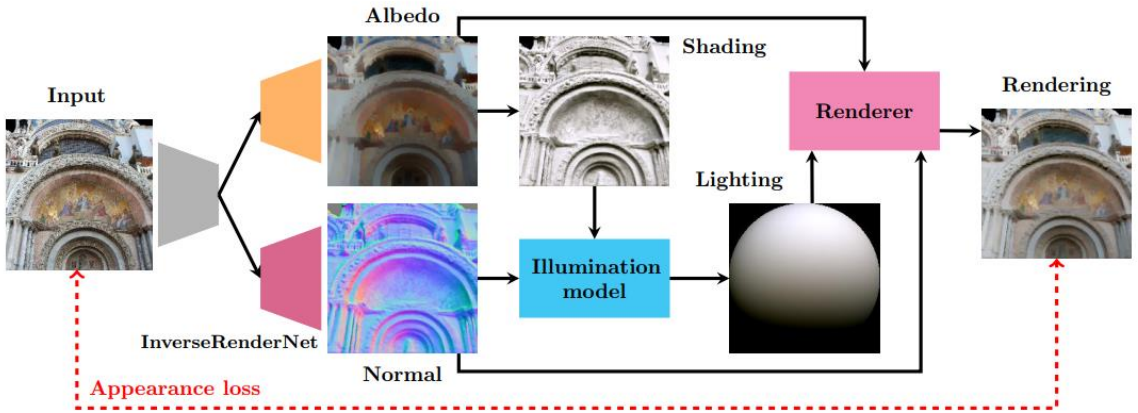
Supervisor: Edgar

Learning Lighting and Appearance

- "*Learning Intrinsic Image Decomposition from Watching the World*", Li and Snavely, CVPR 2018
- "*InverseRenderNet: Learning Single Image Inverse Rendering*", Yu and Smith, CVPR 2019

  - Learning to separate the components of light transport – lighting, reflectance and geometry
  - Data-driven representation learning for the appearance parameters
  - Photorealistic modification for augmented and virtual reality applications

Decomposing images into reflectance and shading by watching timelapse images

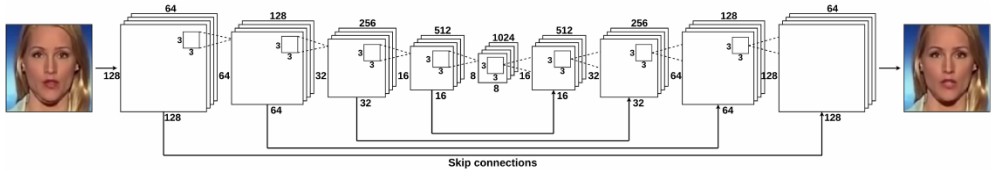


Decomposing images into reflectance and shading by watching Multiview images

Detection

FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces

(Rössler et al., 2018)



FaceForensics++: Learning to Detect Manipulated Facial Images

(Rössler et al., 2019)

Supervisor: Gereon

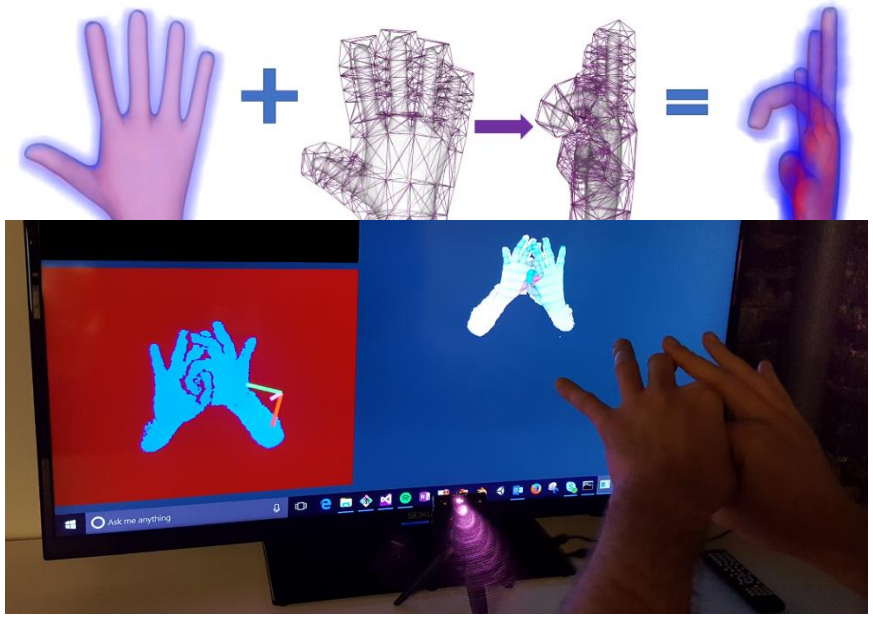
Reconstruction of Two Interacting Hands

Graphics Vision & Video

max planck institut informatik

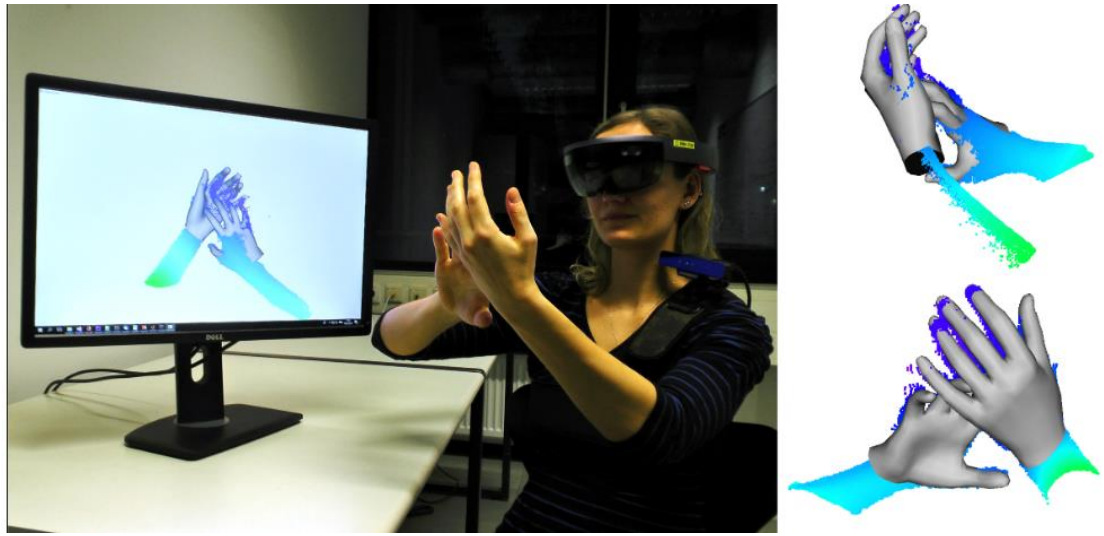
Articulated Distance Fields for Ultra-fast Tracking of Hands Interacting

(Taylor et al., SIGGRAPH Asia 2017)



Real-time Pose and Shape Reconstruction of Two Interacting Hands With a Single Depth Camera

(Mueller et al., SIGGRAPH 2019)

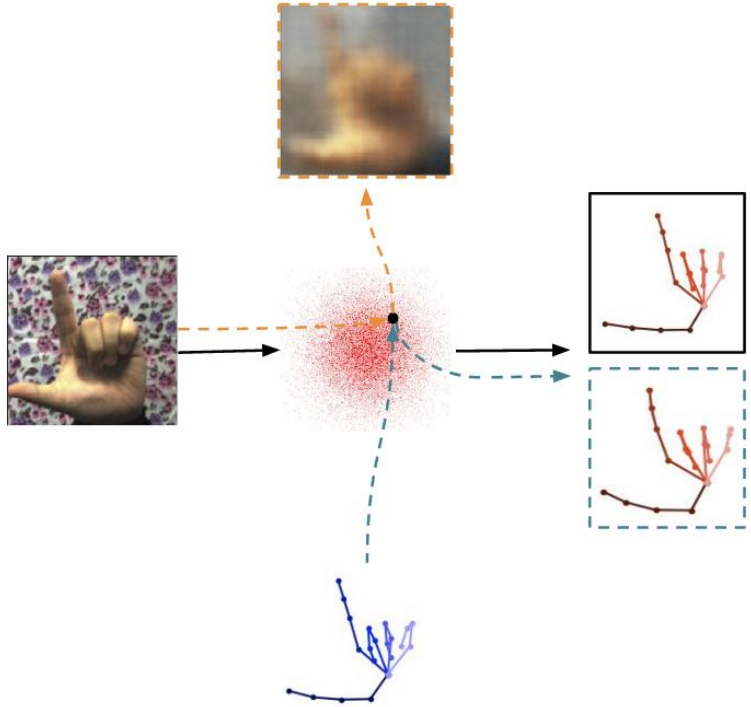


Supervisor: Franzi

Hand Pose Estimation from Monocular RGB

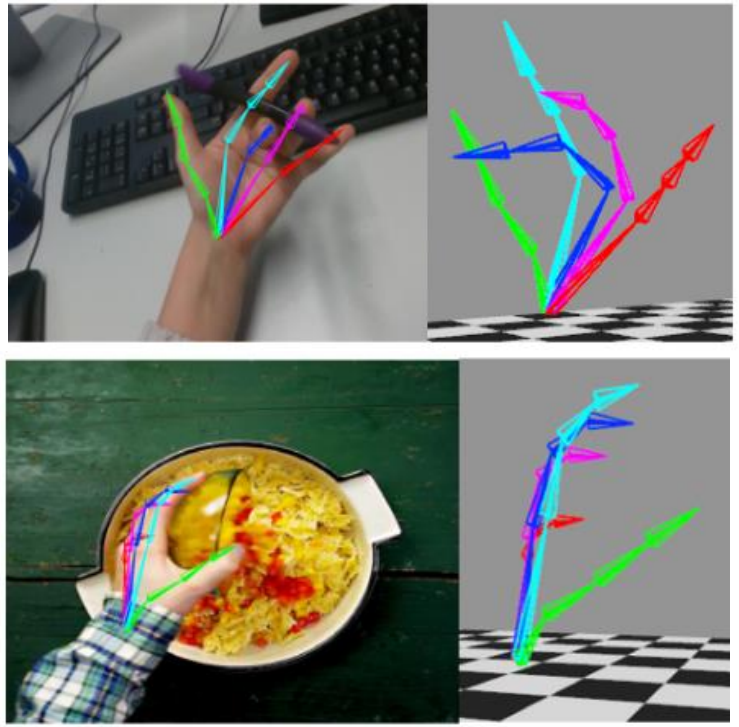
Cross-modal Deep Variational Hand Pose Estimation

(Spurr et al., CVPR 2018)



GANerated Hands for Real-time 3D Hand Tracking from Monocular RGB

(Mueller et al., CVPR 2018)

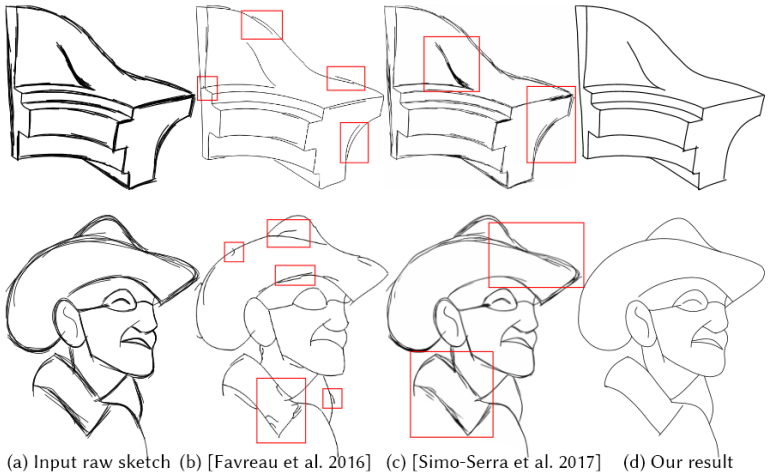


Supervisor: Franzi

Cleaning Sketches

StrokeAggregator: Consolidating Raw Sketches into Artist-Intended Curve Drawings

(Liu et al. Siggraph 2018)



(a) Input raw sketch (b) [Favreau et al. 2016] (c) [Simo-Serra et al. 2017] (d) Our result

Mastering Sketching: Adversarial Augmentation for Structured Prediction

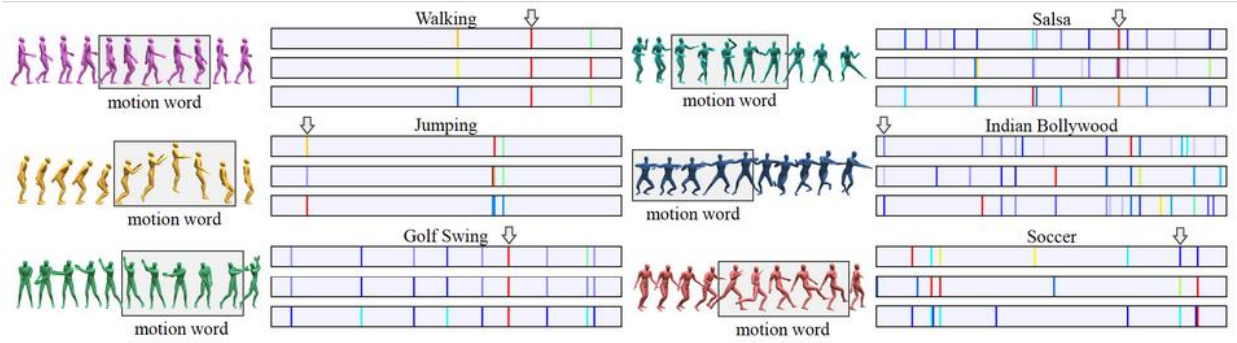
(Simo-Serra et al. TOG 2018)

Supervisor: Dushyant Mehta

How People Move

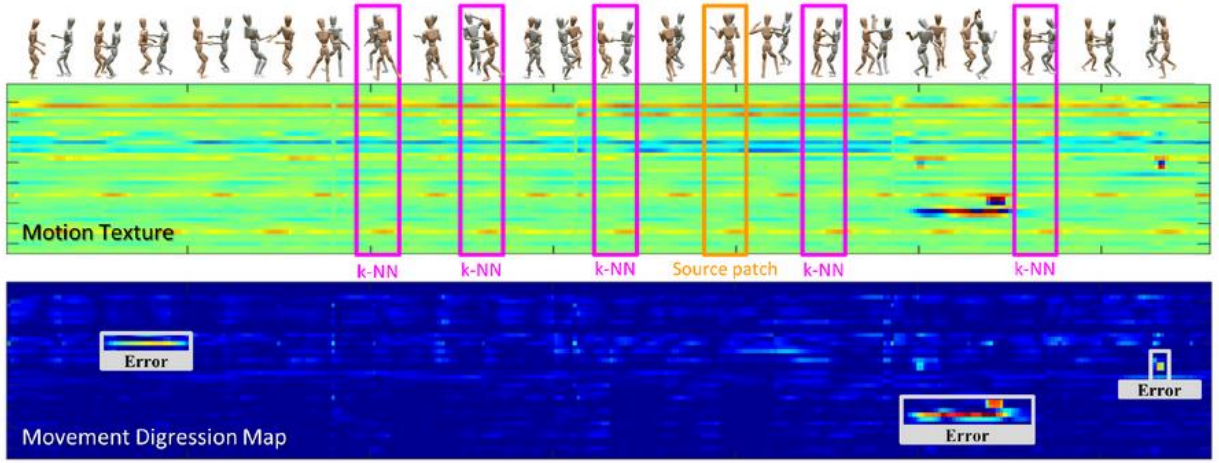
Deep Motifs and Motion Signatures

(Aristidou et al. Siggraph Asia 2018)



Self-similarity Analysis for Motion Capture Cleaning

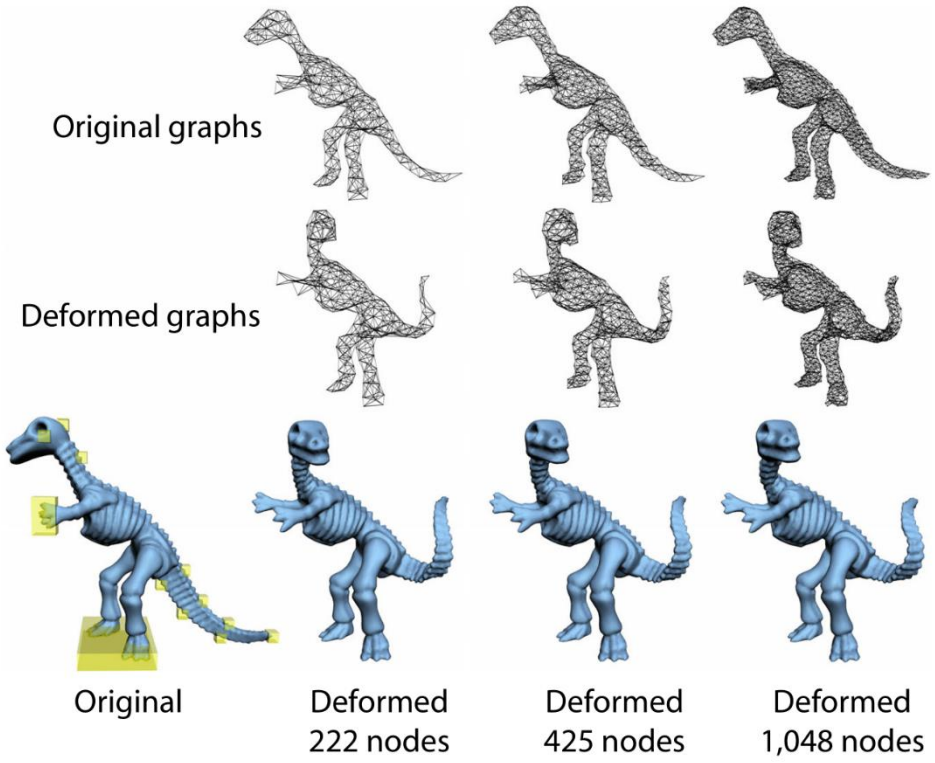
(Aristidou et al. Eurographics 2018 Special Issue)



Supervisor: Dushyant Mehta

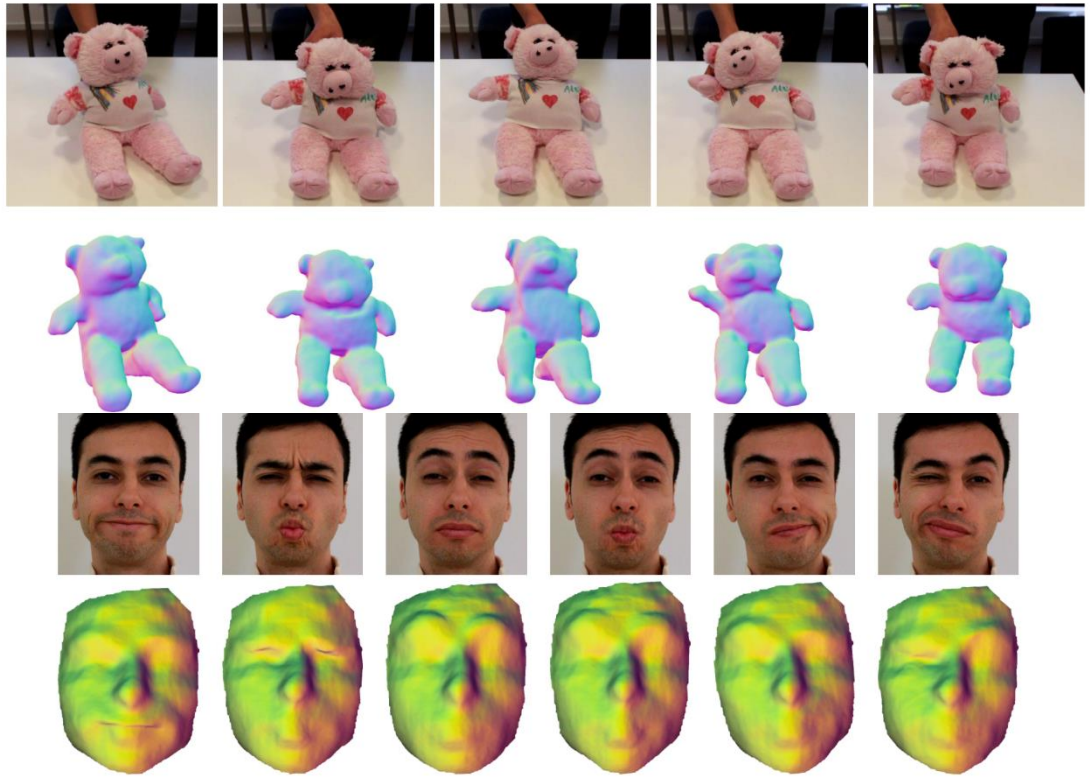
Non-rigid deformations – Constraints & Tracking

Embedded Deformation for Shape Manipulation
(Sumner et al., SIGGRAPH 2007)



Source: Sumner et al. 2007

Direct, Dense, and Deformable: Template-Based Non-Rigid 3D Reconstruction from RGB Video
(Yu et al., ICCV 2015)

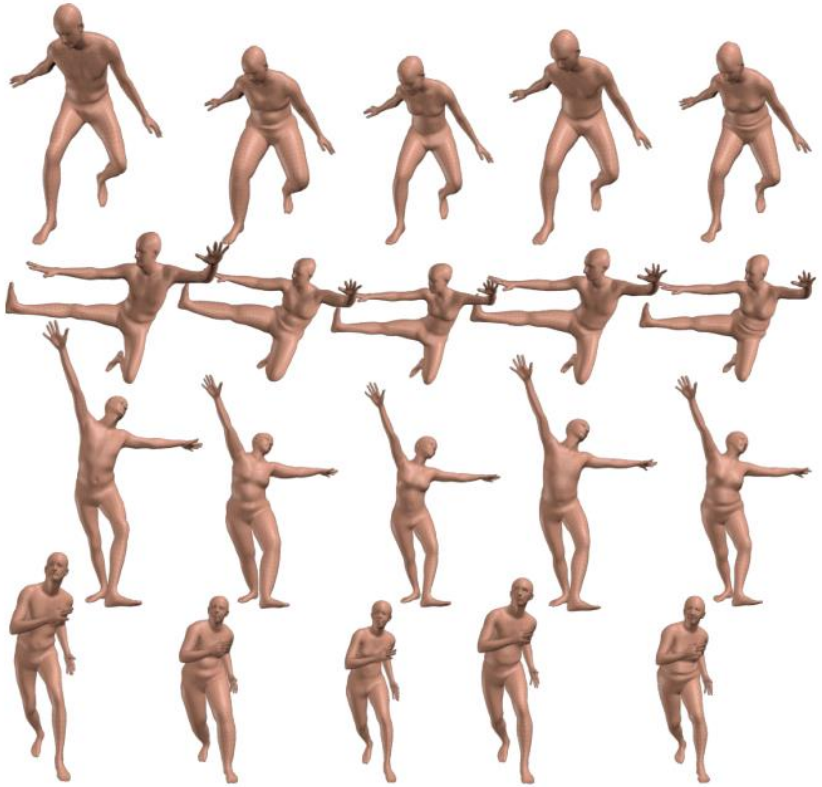


Source: Yu et al. 2015

Supervisor: Marc

Body & Clothing - Modelling & Tracking

SMPL: A Skinned Multi-Person Linear Model
(Loper et al., TOG 2015)



Source: Loper et al. 2015

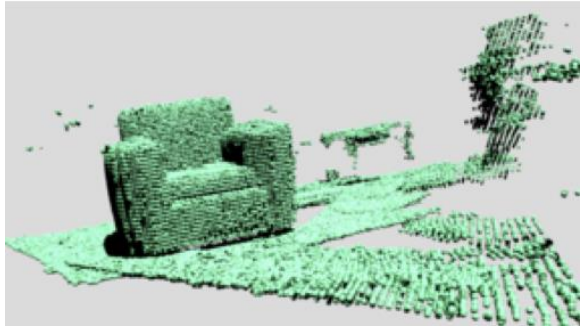
ClothCap: Seamless 4D Clothing Capture and Retargeting
(Pons-Moll et al., TOG 2017)

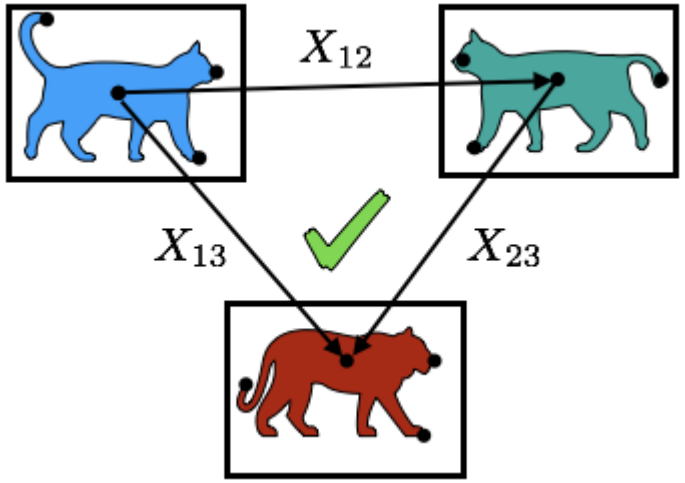
Source: Pons-Moll et al. 2017

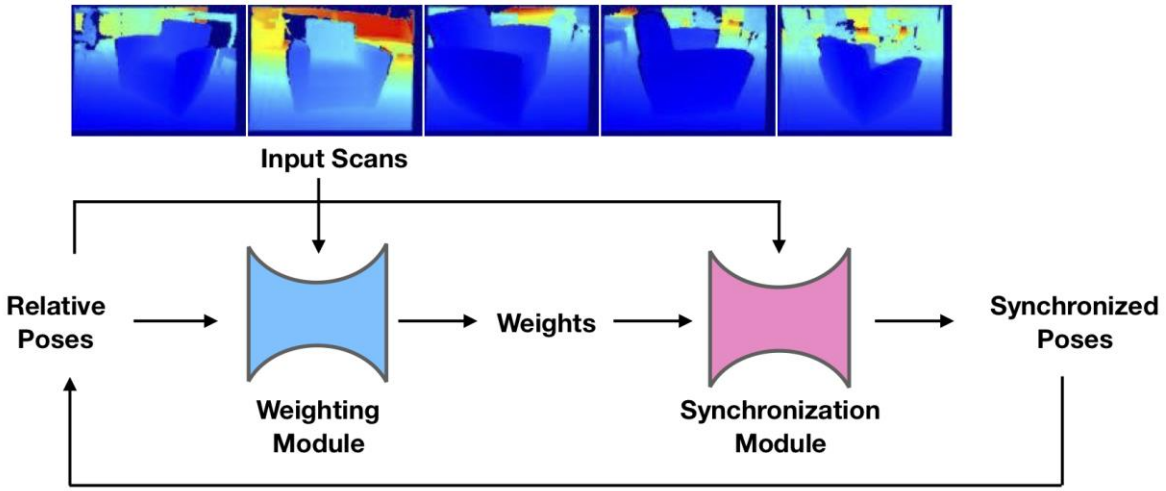
Supervisor: Marc

Transformation Synchronisation

- **Research question**: How to utilise cycle-consistency for robust mult-alignment
- **Required background**: linear algebra, optimisation
- **Papers**:
  - A Solution for Multi-Alignment by Transformation Synchronisation, **CVPR 2015**
    F. Bernard, J. Thunberg, P. Gemmar, F. Hertel, A. Husch, J. Goncalves
  - Learning Transformation Synchronization, **CVPR 2019**
    Xiangru Huang, Zhenxiao Liang, Xiaowei Zhou, Yao Xie, Leonidas Guibas, Qixing Huang







Supervisor: Florian

Learning to Match by Differentiable Programming

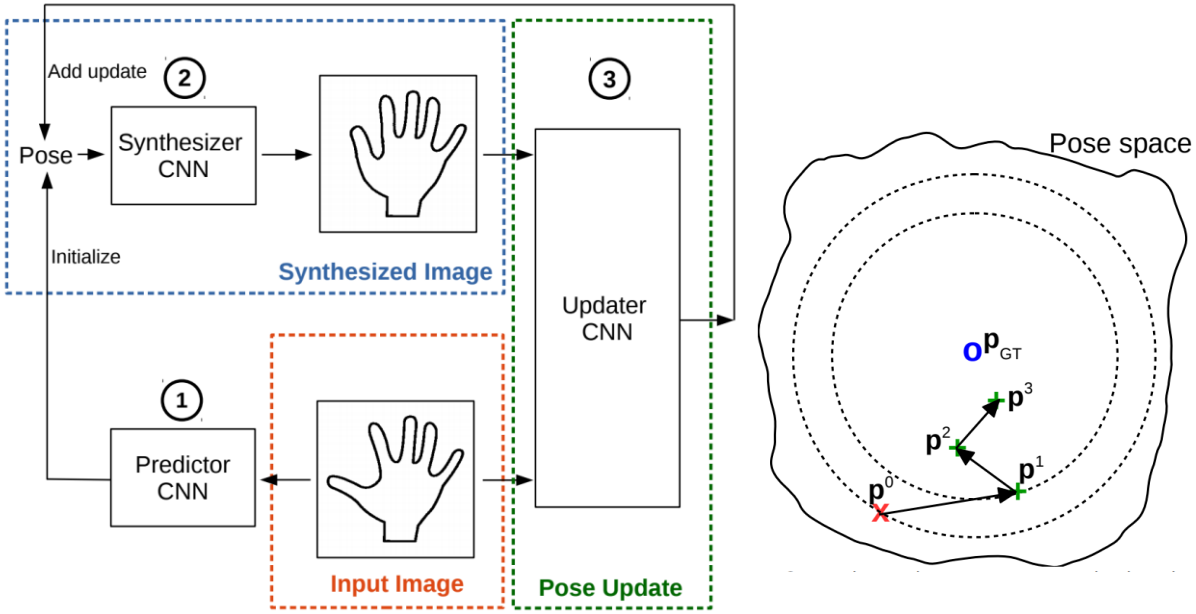
- **Research question**: How to develop deep learning systems that *backpropagate through algorithms*
- **Required background**: linear algebra, optimisation
- **Papers**:
  - Learning Latent Permutations with Gumbel-Sinkhorn Networks, **ICLR 2018**
    Gonzalo Mena, David Belanger, Scott Linderman, Jasper Snoek
  - Deep Learning of Graph Matching, **CVPR 2018**
    Andrei Zanfir and Cristian Sminchisescu

Supervisor: Florian

Utilizing Hand Synthesizer

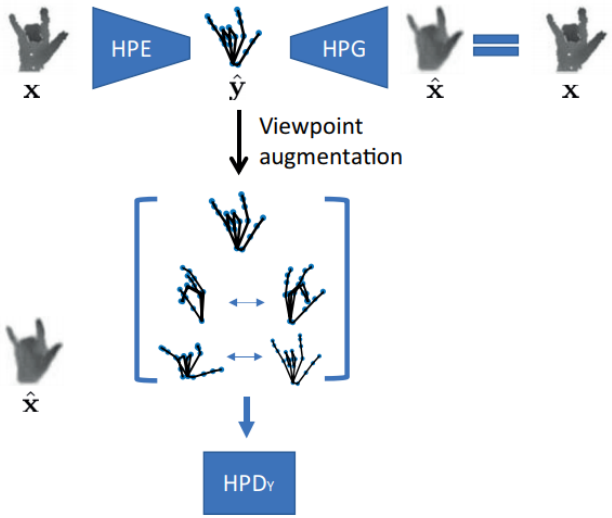
Generalized Feedback Loop for Joint Hand-Object Pose Estimation

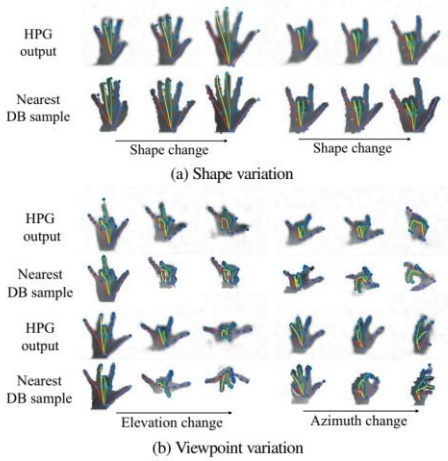
(Oberweger et al., TPAMI 2019)



Augmented Skeleton Space Transfer for Depth-based Hand Pose Estimation

(Baek et al., CVPR 2018)



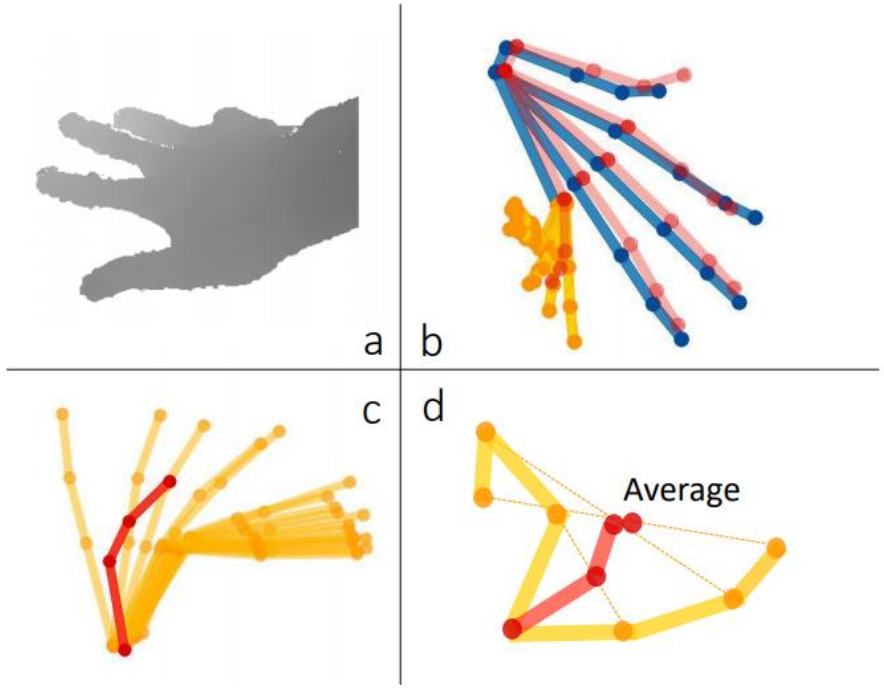


Supervisor: Jiayi

Leveraging Pose Uncertainty

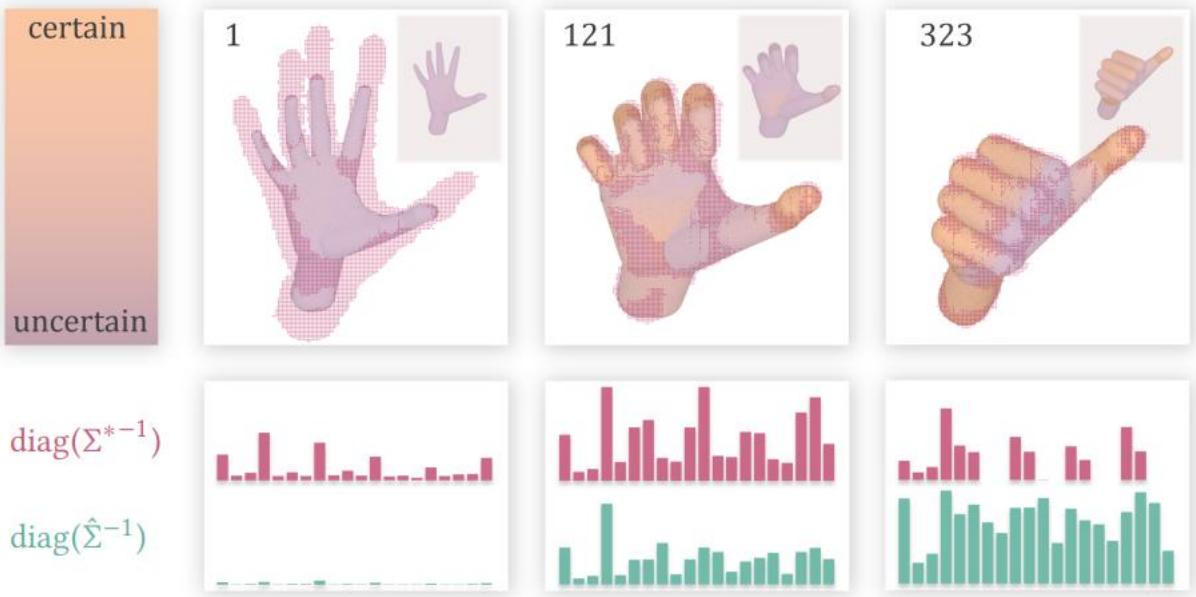
Occlusion-aware Hand Pose Estimation Using Hierarchical Mixture Density Network

(Ye et al., ECCV 2018)



Online Generative Model Personalization for Hand Tracking

(Tkach et al., SIGGRAPH Asia 2017)

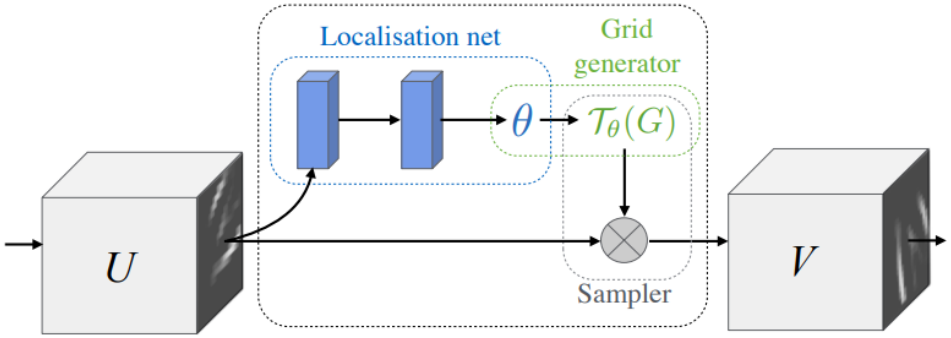


Supervisor: Jiayi

Deep Feature Invariance

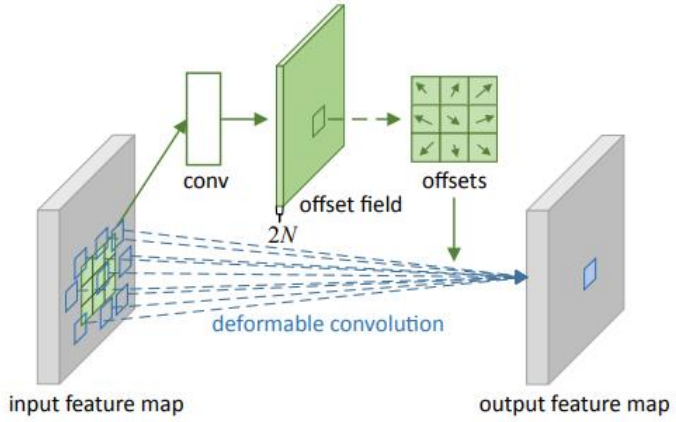
Spatial Transformer Networks

(*Jaderberg et al.* NIPS 2015)



Deformable Convolutional Networks

(*Dai et al.* ICCV 2017)

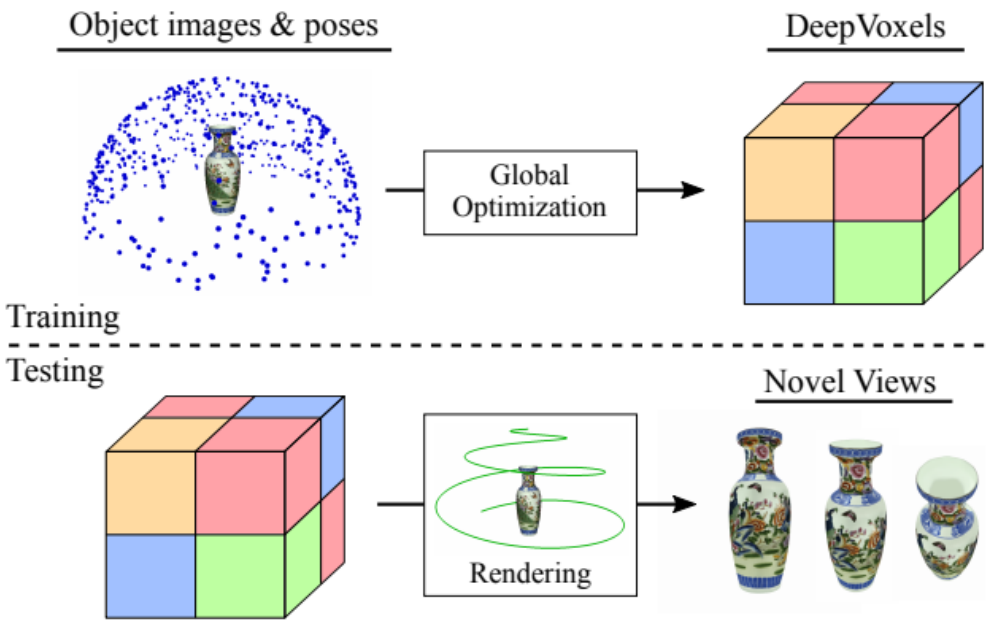


Supervisor: Mallikarjun

Latent Feature Embedding and Modification

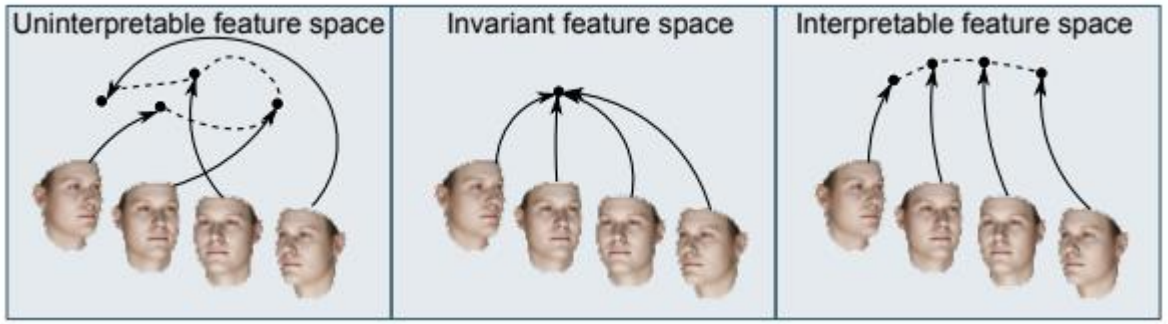
DeepVoxels: Learning Persistent 3D Feature Embeddings

(*Sitzmann et al.*, CVPR 2019)



Interpretable Transformations with Encoder-Decoder Networks

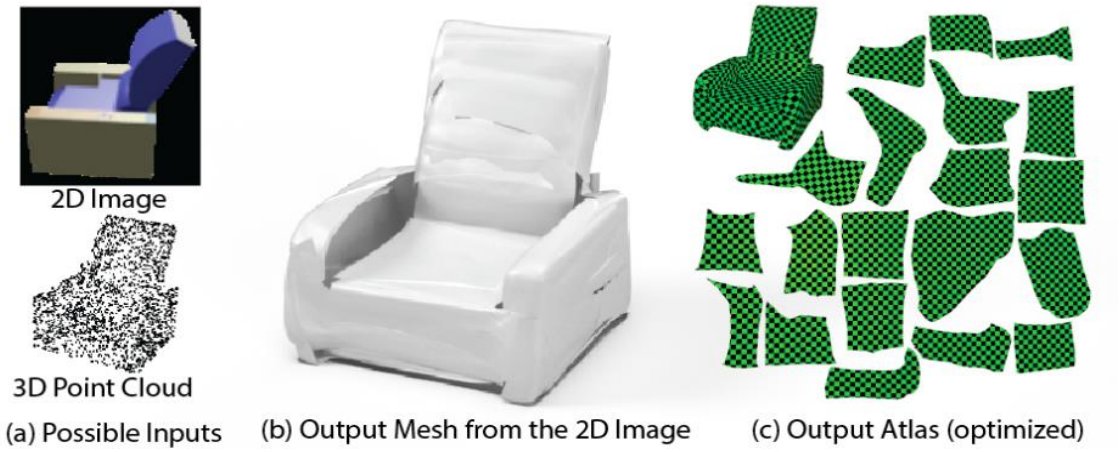
(*Worall et al.* ICCV 2017)



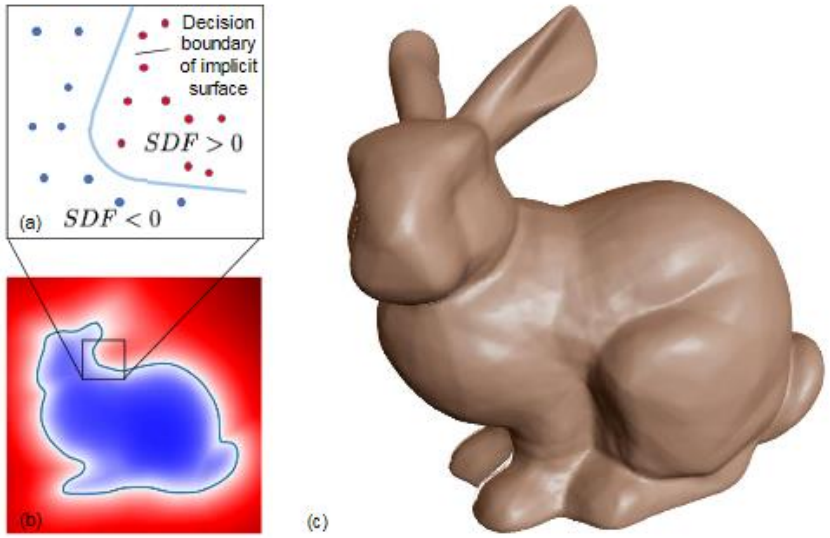
Supervisor: Mallikarjun

3D Surface Representations

AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation
(Groueix et al., CVPR 2018)



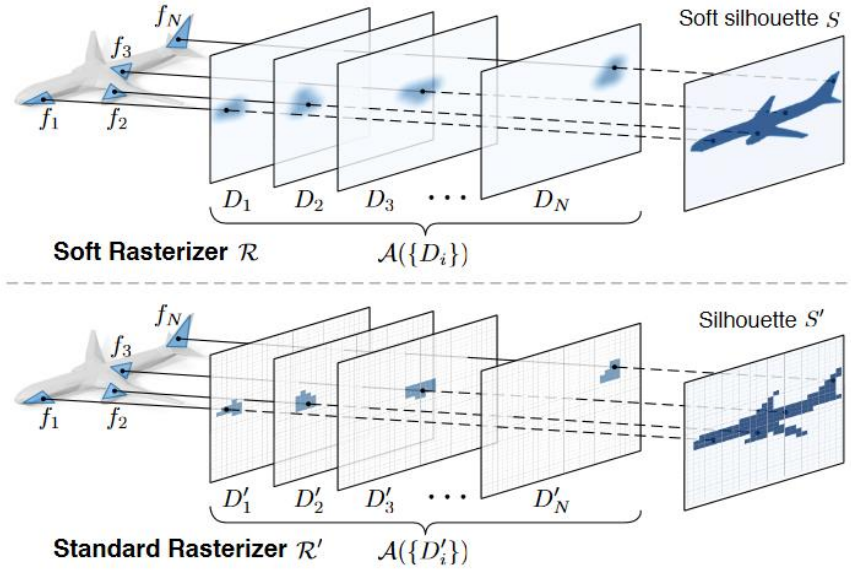
DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation
(Park et al., CVPR 2019)



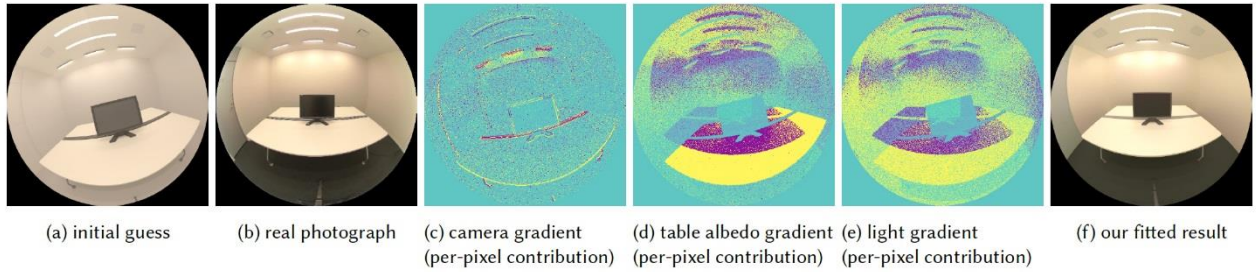
Supervisor: Ayush

“Differentiable” Rendering

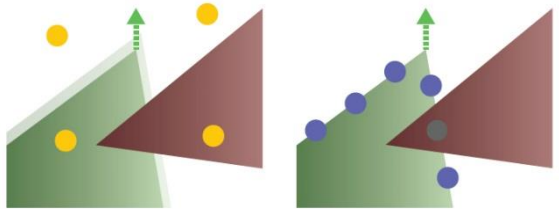
Soft Rasterizer: Differentiable Rendering for Unsupervised Single-View Mesh Reconstruction
(Liu et al., arXiv 2019)



Differentiable Monte Carlo Ray Tracing through Edge Sampling
(Li et al., SIGGRAPH Asia 2018)



(a) initial guess (b) real photograph (c) camera gradient (per-pixel contribution) (d) table albedo gradient (per-pixel contribution) (e) light gradient (per-pixel contribution) (f) our fitted result



(a) area sampling (b) edge sampling

Supervisor: Ayush

Overview

- Organization
- Introduction
- Topics
- Summary

Summary

- Topic assignment:
  - Email us to get a slot (first-come, first-served)
  - Send a list of 3 topics (in order of preference) until **Tuesday, 16 April 2019**
  - We will try to accommodate wishes as much as possible
  - Topics will be assigned on **Thursday, 18 April 2019**
- First topic presentation: **Thursday, 25 April 2019**
- Two lectures next week:
  - "How to read an academic paper"
  - "How to give a good scientific talk"

Thanks!

Any questions?