3D Scanning for Virtual Shopping

Dr. Jürgen Sturm
Group Leader RGB-D
metaio GmbH

Augmented Reality with the Metaio SDK

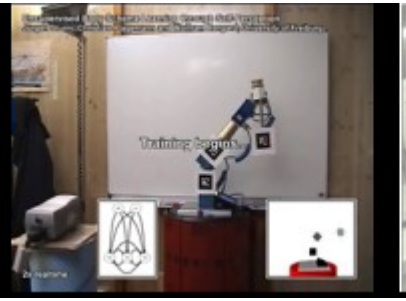
Metaio SDK

- SDK for developers to create augmented reality apps
- Supports iOS, Android, Windows, Unity
- Provides efficient implementations for
  - Marker-based 3D tracking
  - Template-based 3D tracking
  - Sparse visual odometry
  - Sparse SLAM (feature-based, local+global bundle adjustment, relocalization, uses depth if available)
  - Edge-based tracking (known CAD model)
- **This talk: Cutting edge research at Metaio**
  - **Dense RGBD-based camera tracking and 3D reconstruction**
  - **Face detection, tracking, and shape alignment**
  - **Visual inertial sparse odometry**

My Research Background

- Visual navigation for mobile robots


RoboCup

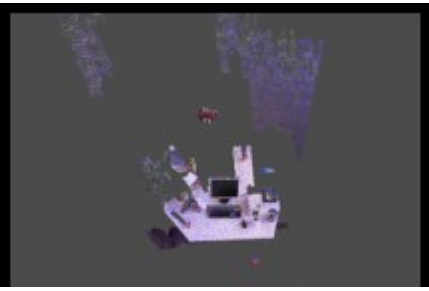

Kinematic Learning

Articulated Objects


Quadrotors

- Camera tracking and 3D reconstruction


RGB-D SLAM


Visual Odometry


Large-Scale Reconstruction


3D Scanning

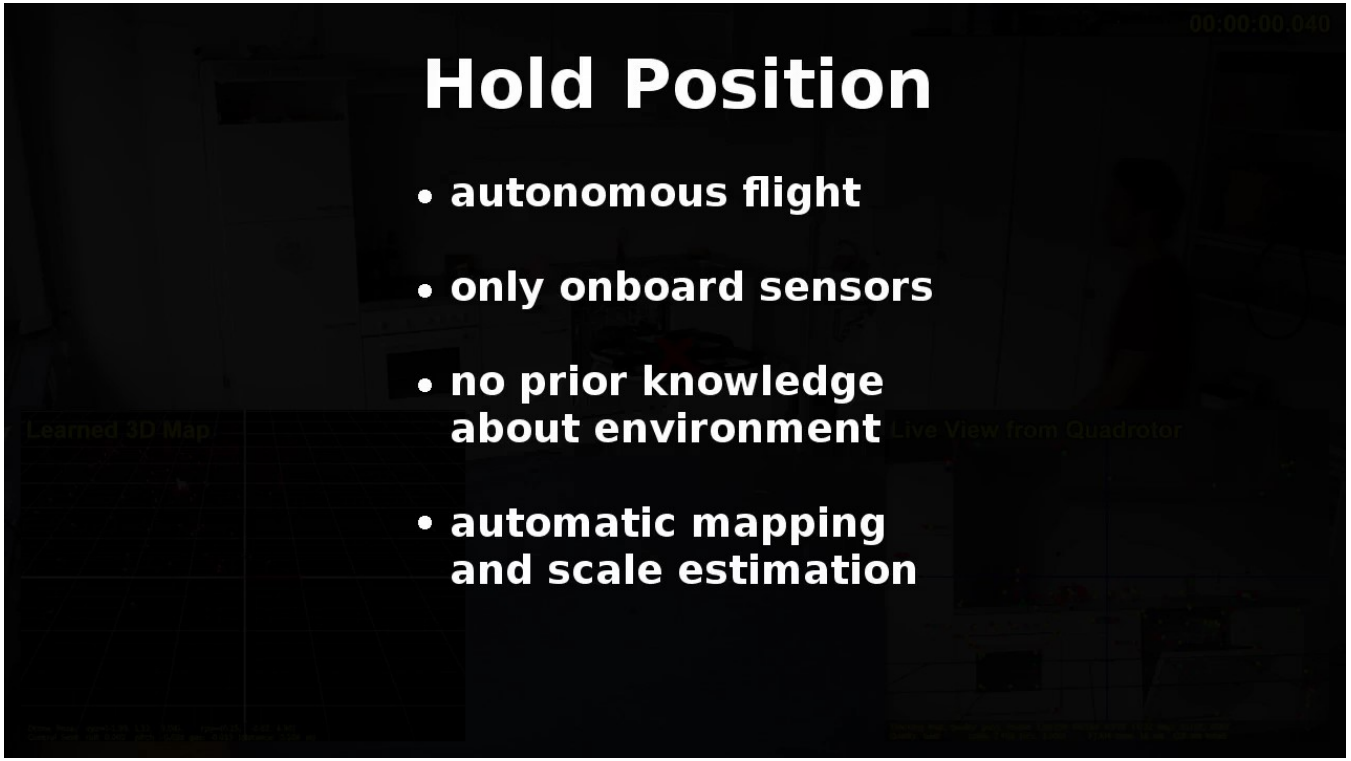
My Research Background

Camera-based Navigation of a Low-Cost Quadrotor

[IROS ’12, RSS ’13, UAV-g ‘13, RAS ’14]
TUM TeachInf Best Lecture Award 2012 and 2013
EdX Course AUTONAVx with 20k participants

My Research Background

Camera-based Navigation of a Low-Cost Quadrotor



[IROS ’12, RSS ’13, UAV-g ‘13, RAS ’14]
TUM TeachInf Best Lecture Award 2012 and 2013
EdX Course AUTONAVx with 20k participants

My Research Background

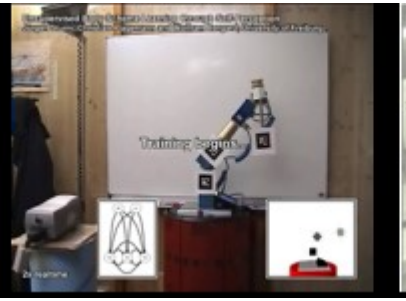
EdX Course „Autonomous Navigation for Flying Robots“

[IROS ’12, RSS ’13, UAV-g ‘13, RAS ’14]
TUM TeachInf Best Lecture Award 2012 and 2013
EdX Course AUTONAVx with 20k participants

My Research Background

- Visual navigation for mobile robots


RoboCup

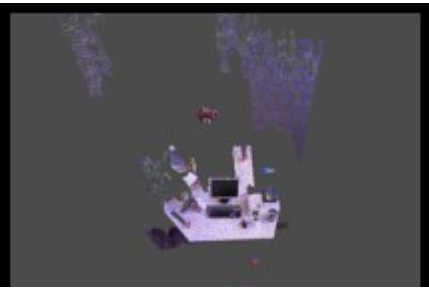

Kinematic Learning

Articulated Objects


Quadrotors

- Camera tracking and 3D reconstruction


RGB-D SLAM

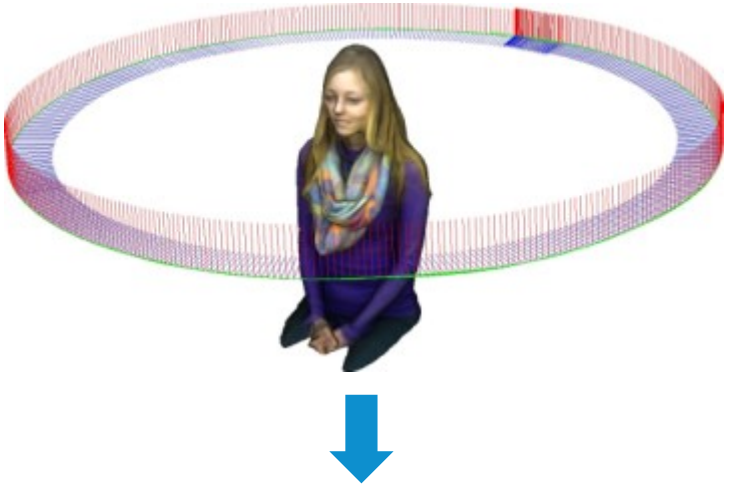

Visual Odometry


Large-Scale Reconstruction


3D Scanning

Motivation

Wouldn‘t it be cool to have a 3D photo booth?



Questions:

- How to scan a person in 3D?
- How to prepare the model for 3D printing?

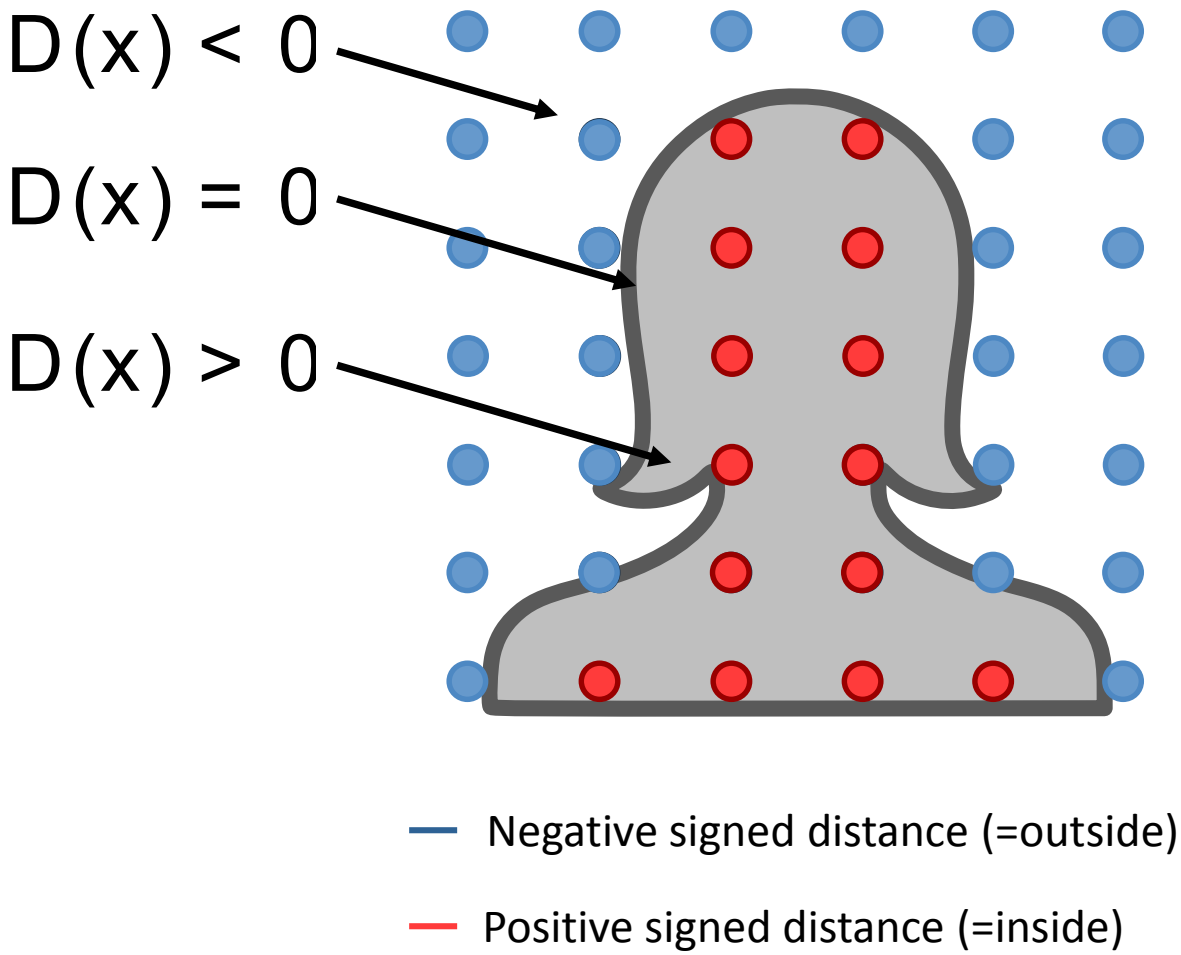
Problem Description

- Setup:
  Static RGB-D camera, person sitting on a swivel chair

- Given: A sequence of color and depth images
- Wanted: Accurate, watertight 3D model

Signed Distance Function (SDF)

[Curless and Levoy, ’96]

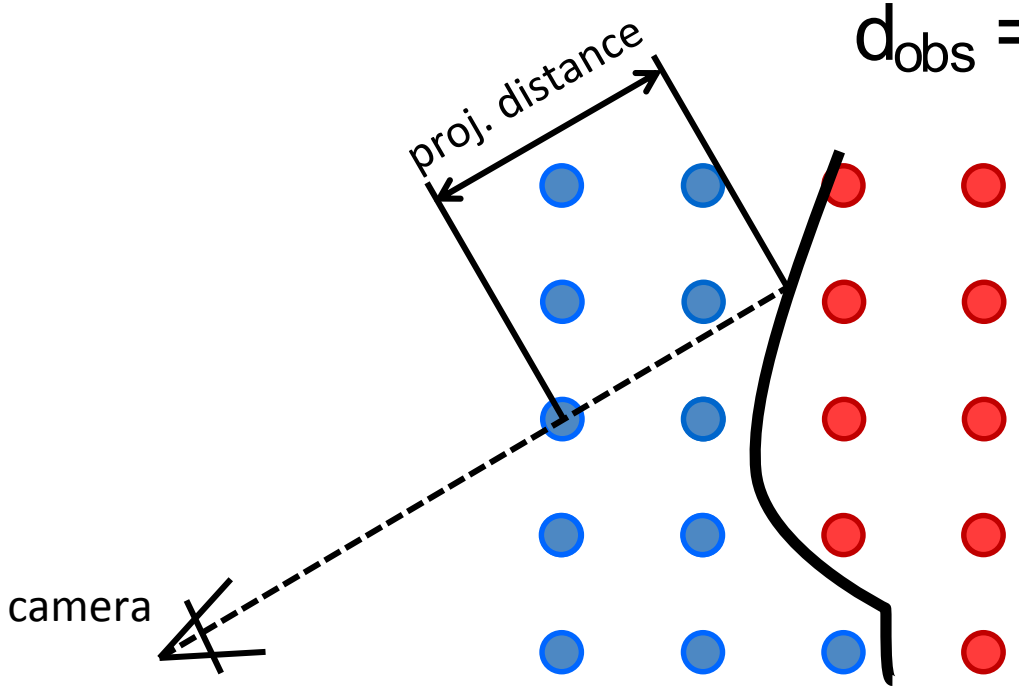


Signed Distance Function (SDF)

[Curless and Levoy, ’96]

- Compute SDF from a depth image
- Measure distance of each voxel to the observed surface
- Can be done in parallel for all voxels (→ GPU)

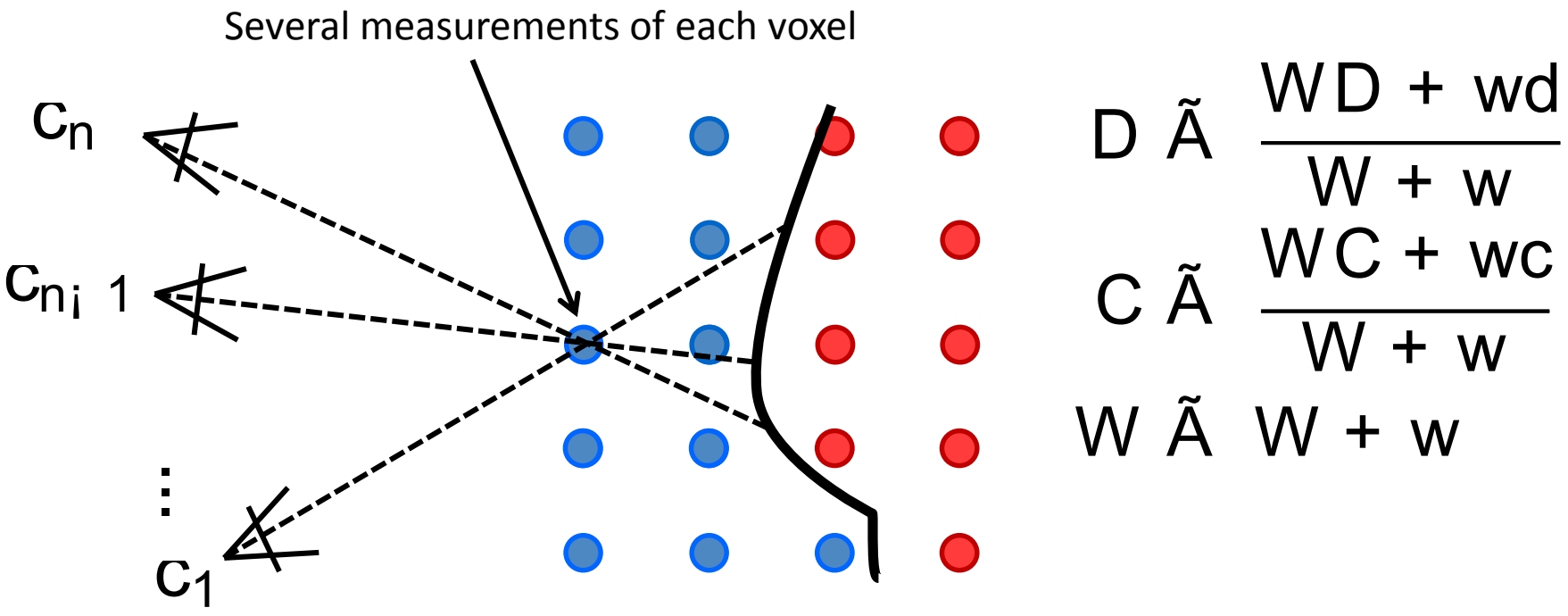
$$d_{obs} = z - l_z(p(x; y; z))$$



Signed Distance Function (SDF)

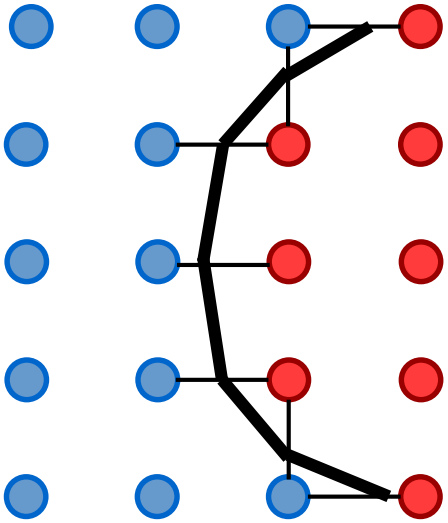
[Curless and Levoy, '96]

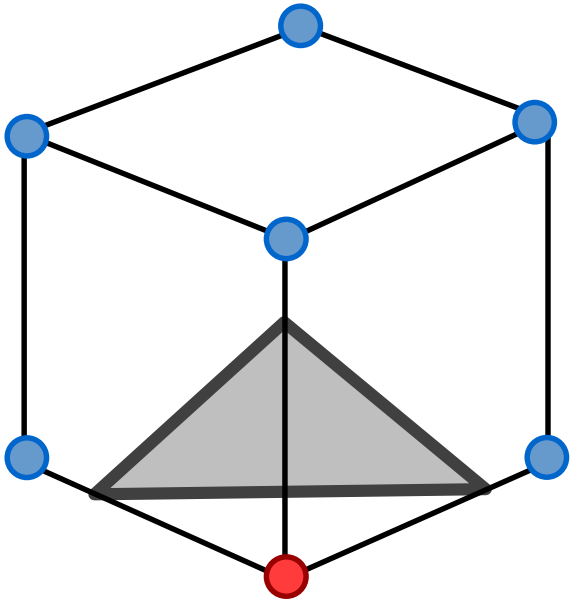
- Calculate weighted average over all measurements
- Assume known camera poses (for now)



Mesh Extraction using Marching Cubes

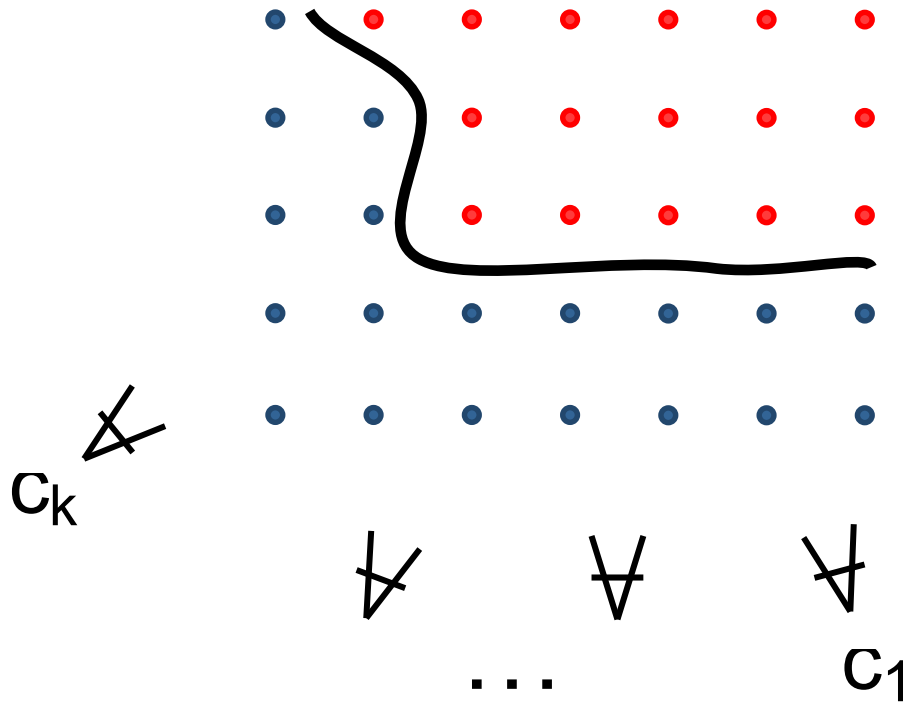
- Find zero-crossings in the signed distance function by interpolation





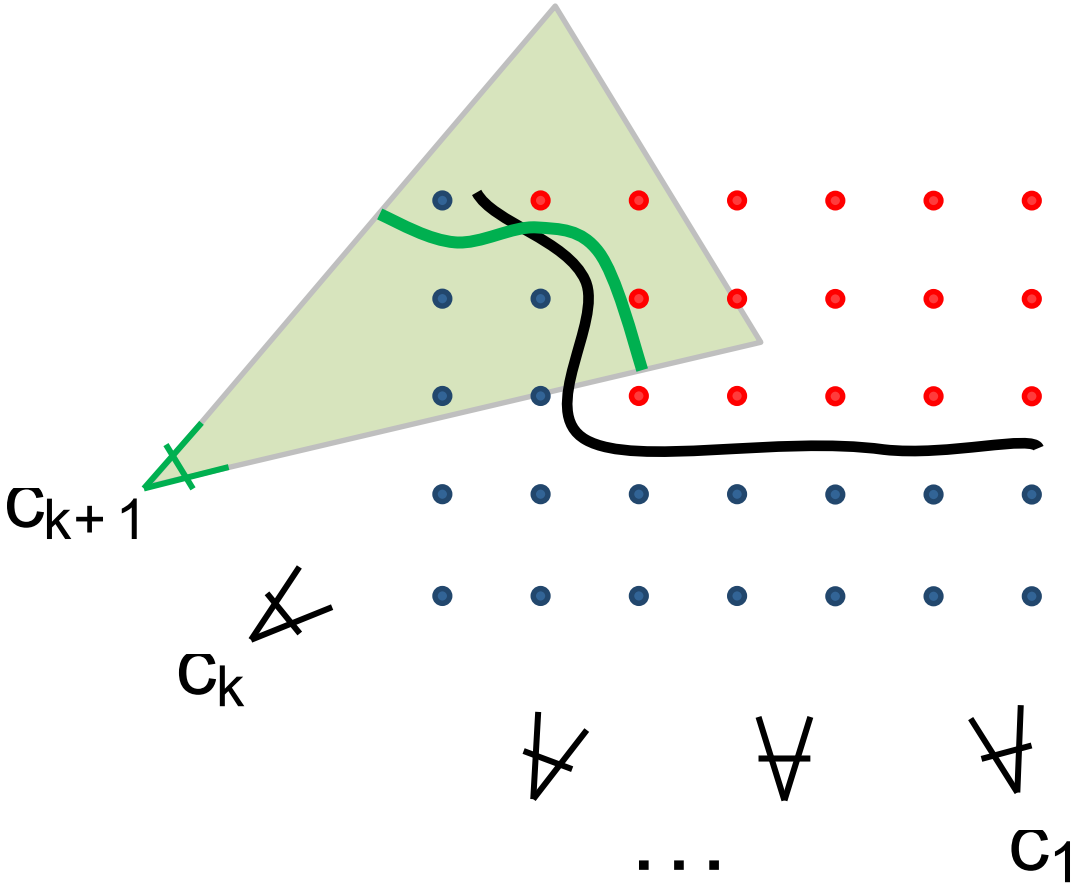
Estimating the Camera Pose

- SDF built from the first k frames



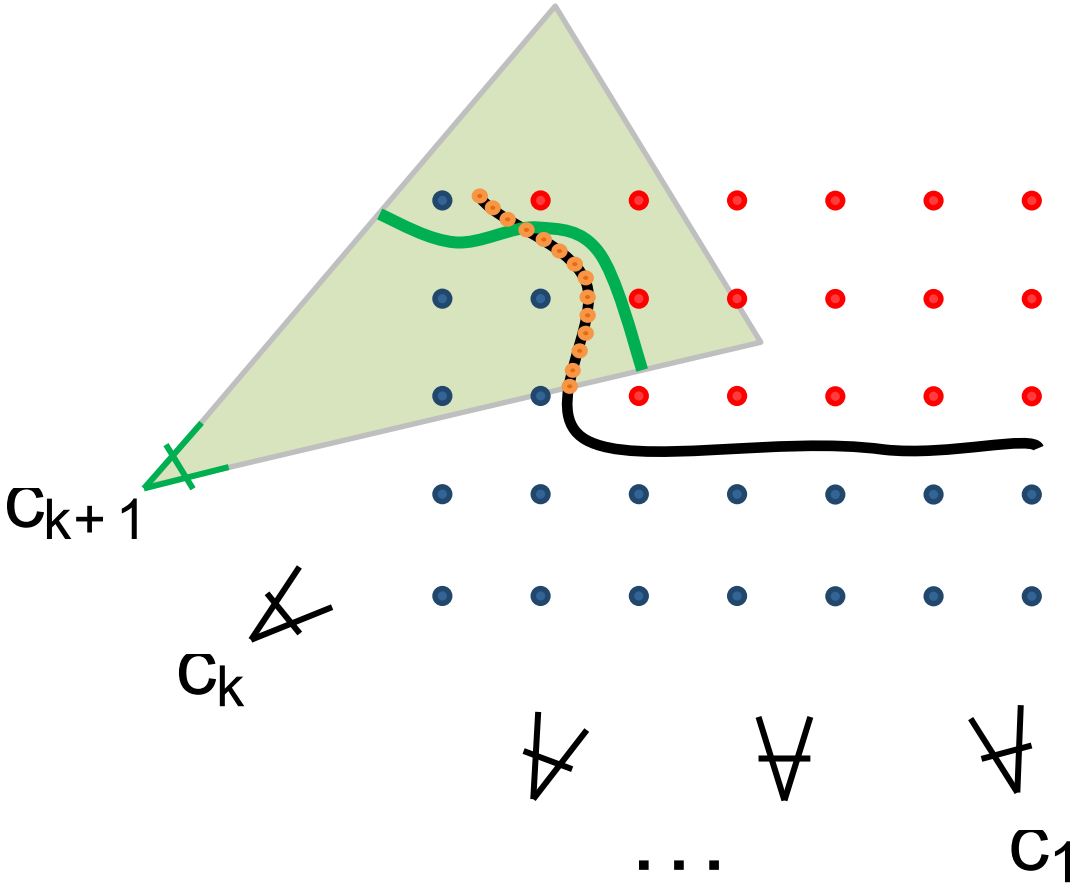
Estimating the Camera Pose

- We seek the next camera pose (k+1)



Estimating the Camera Pose

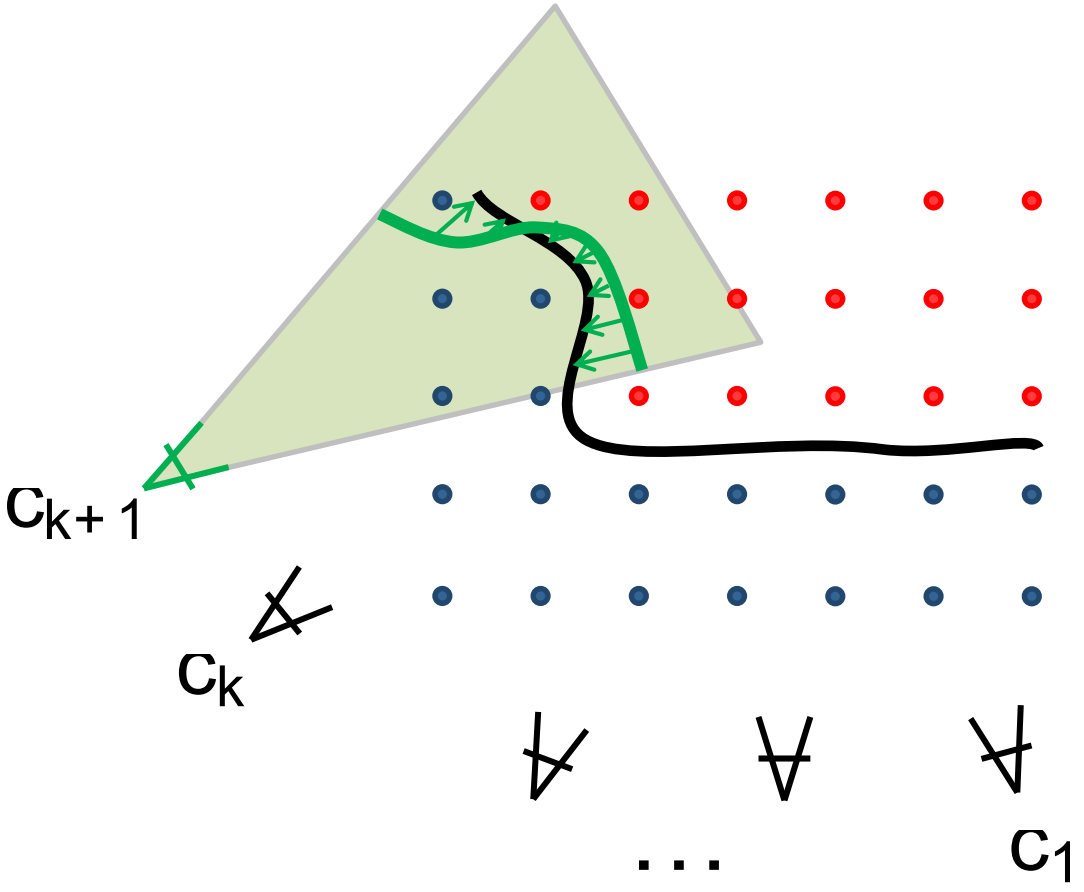
- KinectFusion generates a synthetic depth image from SDF and aligns it using ICP



Estimating the Camera Pose

[Bylow, Sturm, Kerl, Kahl, Cremers; RSS 2013]

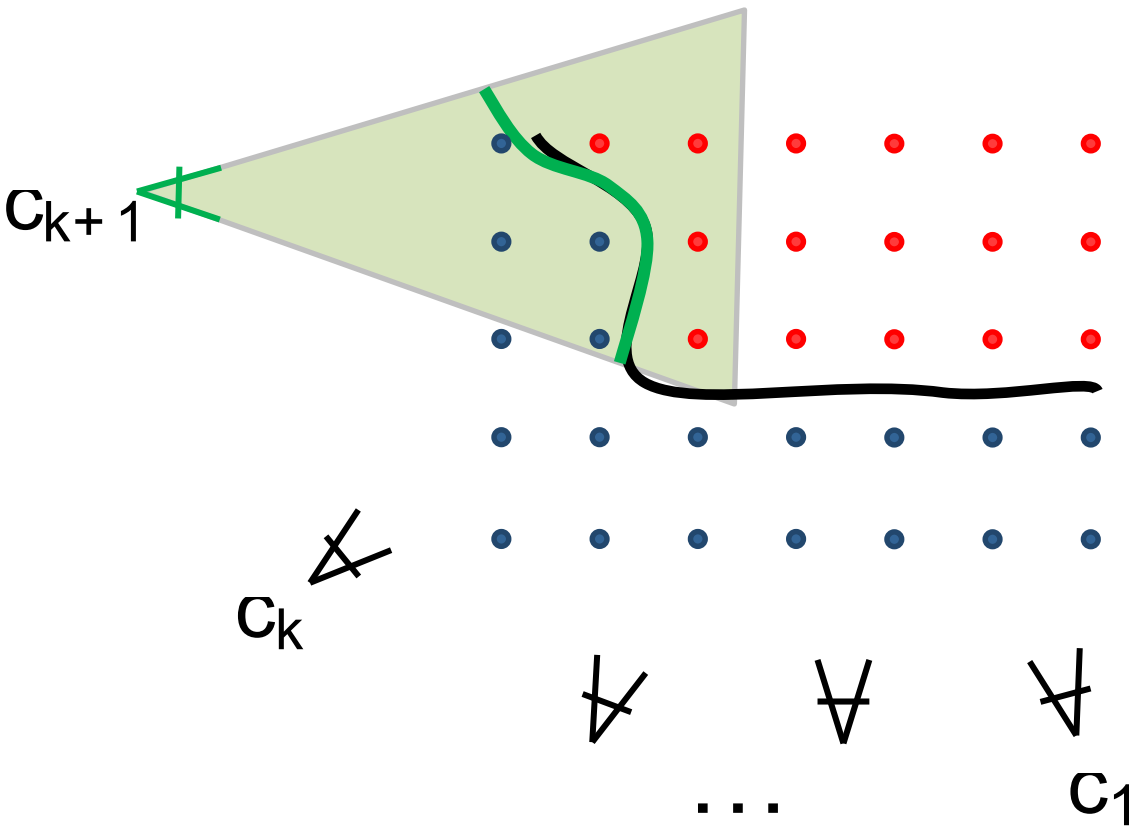
- Our approach: Use SDF directly during minimization



Estimating the Camera Pose

[Bylow, Sturm, Kerl, Kahl, Cremers; RSS 2013]

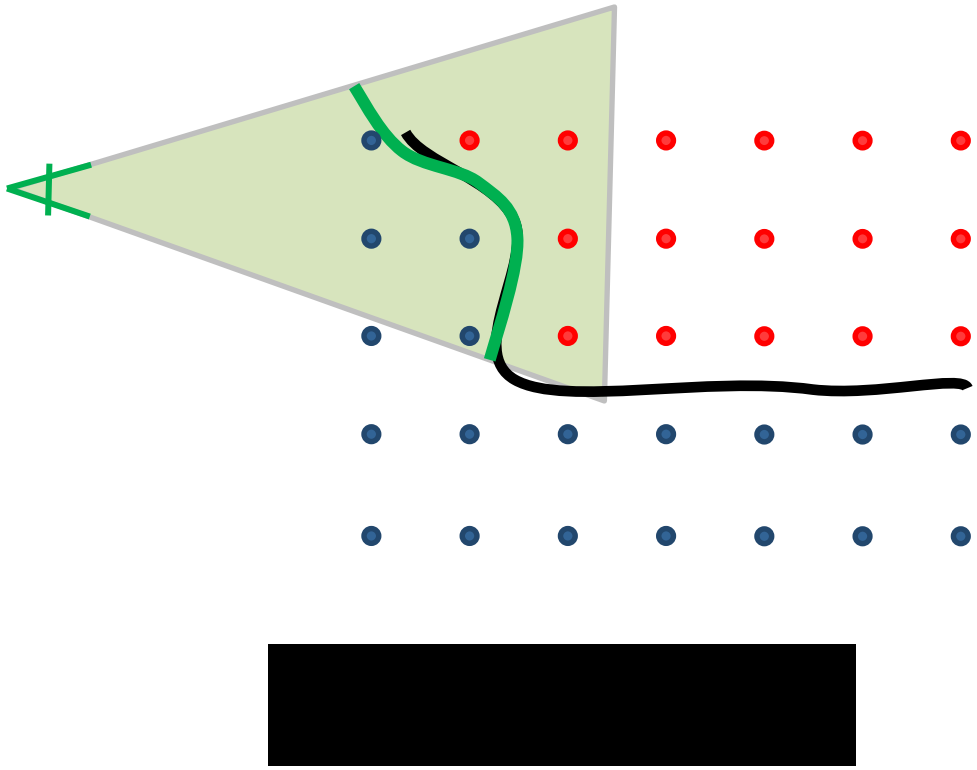
- Our approach: Use SDF directly during minimization



Estimating the Camera Pose

[Bylow, Sturm, Kerl, Kahl, Cremers; RSS 2013]

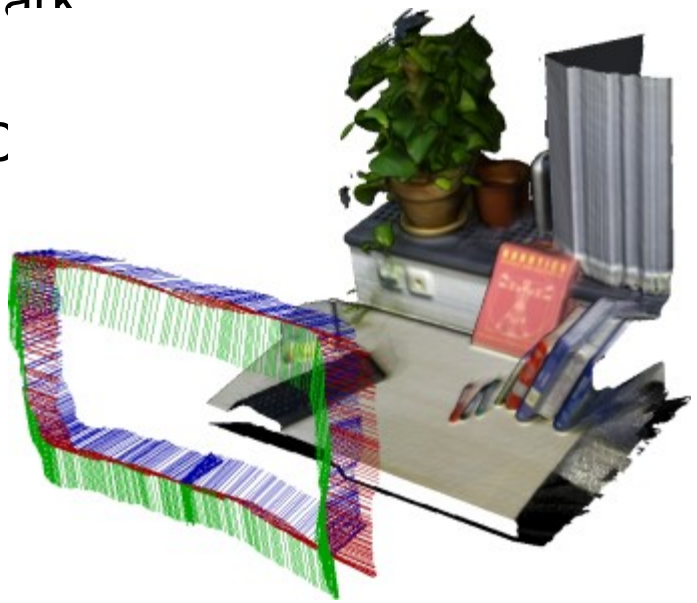
- Our approach: Use SDF directly during minimization



Evaluation on Benchmark

[Bylow, Sturm, Kerl, Kahl, Cremers; RSS 2013]

- Thorough evaluation on TUM RGB-D benchmark
- Comparison with KinFu and RGB-D SLAM
- Significantly more accurate and robust than IC

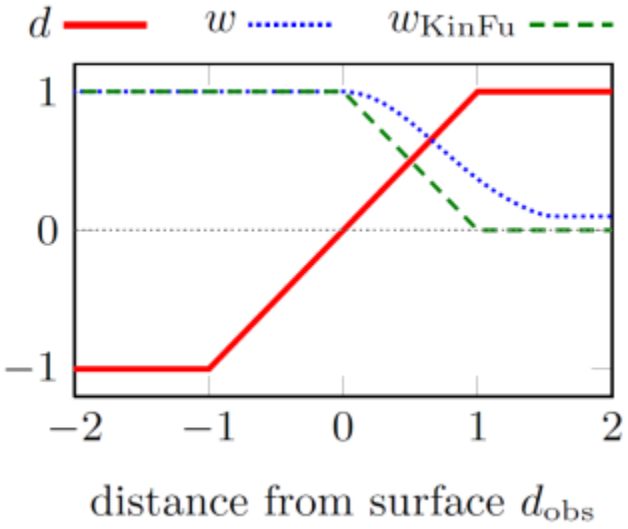


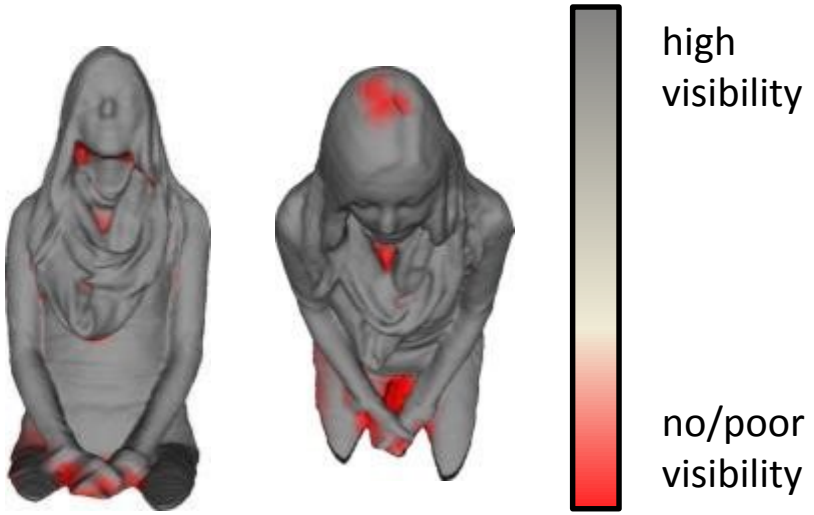
Algorithm	Resolution	Teddy (RMSE)	Desk (RMSE)	Plant (RMSE)
KinFu	256	0.156 m	0.057m	0.598 m
KinFu	512	0.337 m	0.068 m	0.281 m
Our	256	0.086 m	0.038 m	0.047 m
Our	512	0.080 m	0.035 m	0.043 m

Automatically Close Holes

[Sturm, Bylow, Kahl, Cremers; GCPR 2013]

- Certain voxels are never observed in near range
- Regions with no data result in holes
- Idea: Truncate weights to positive values

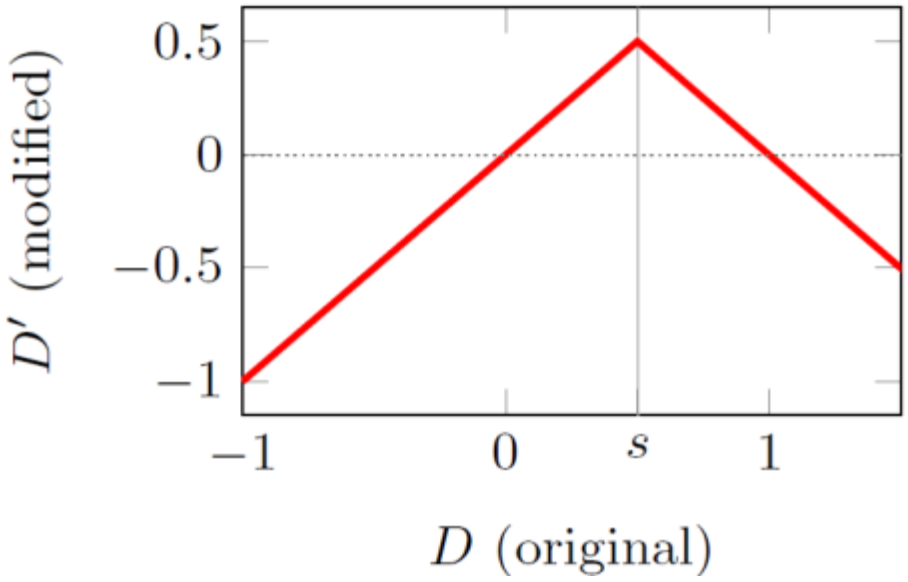




Hollowing Out

[Sturm, Bylow, Kahl, Cremers; GCPR 2013]

- Printing cost is mostly dominated by volume
- Idea: Make the model hollow



before

after

Video (real-time)

[Sturm, Bylow, Kahl, Cremers; GCPR 2013]

Examples of Printed Figures

[Sturm, Bylow, Kahl, Cremers; GCPR 2013]

More Examples

[Sturm, Bylow, Kahl, Cremers; GCPR 2013]




- Still need a Christmas present?

How Can Use This for Virtual Shopping?

- Virtual Try-On of
  - Sunglasses
  - Earrings
- Target devices:
  - Kiosk applications
  - Smartphones

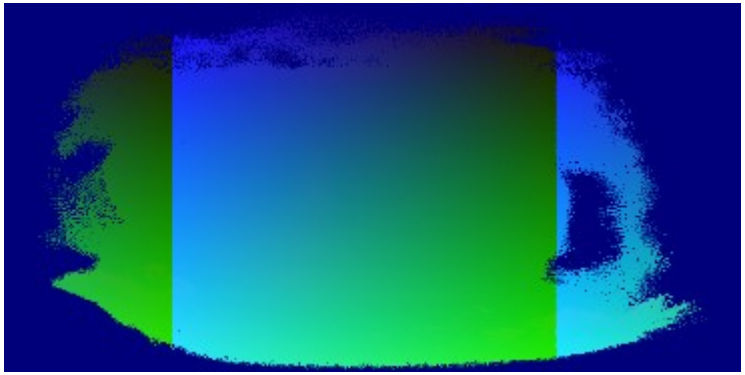
Productification challenges that we faced at Metaio:

- Dense SDF representation is not efficient
  - High memory consumption [redacted]
  - Computationally intense (strong GPU needed for real-time processing)
- Make it so that it always works
  - Initialization / recovery / re-localization
  - Automatic object placement
- Smartphones do not have depth cameras (yet)
  - Can we do the same with a monocular (user-facing) camera of a smartphone?

Efficient 3D Face Reconstruction

- How to reduce memory and CPU utilization?
- Specialized representation for faces
- Spherical base shape + bump map + 2D texture map

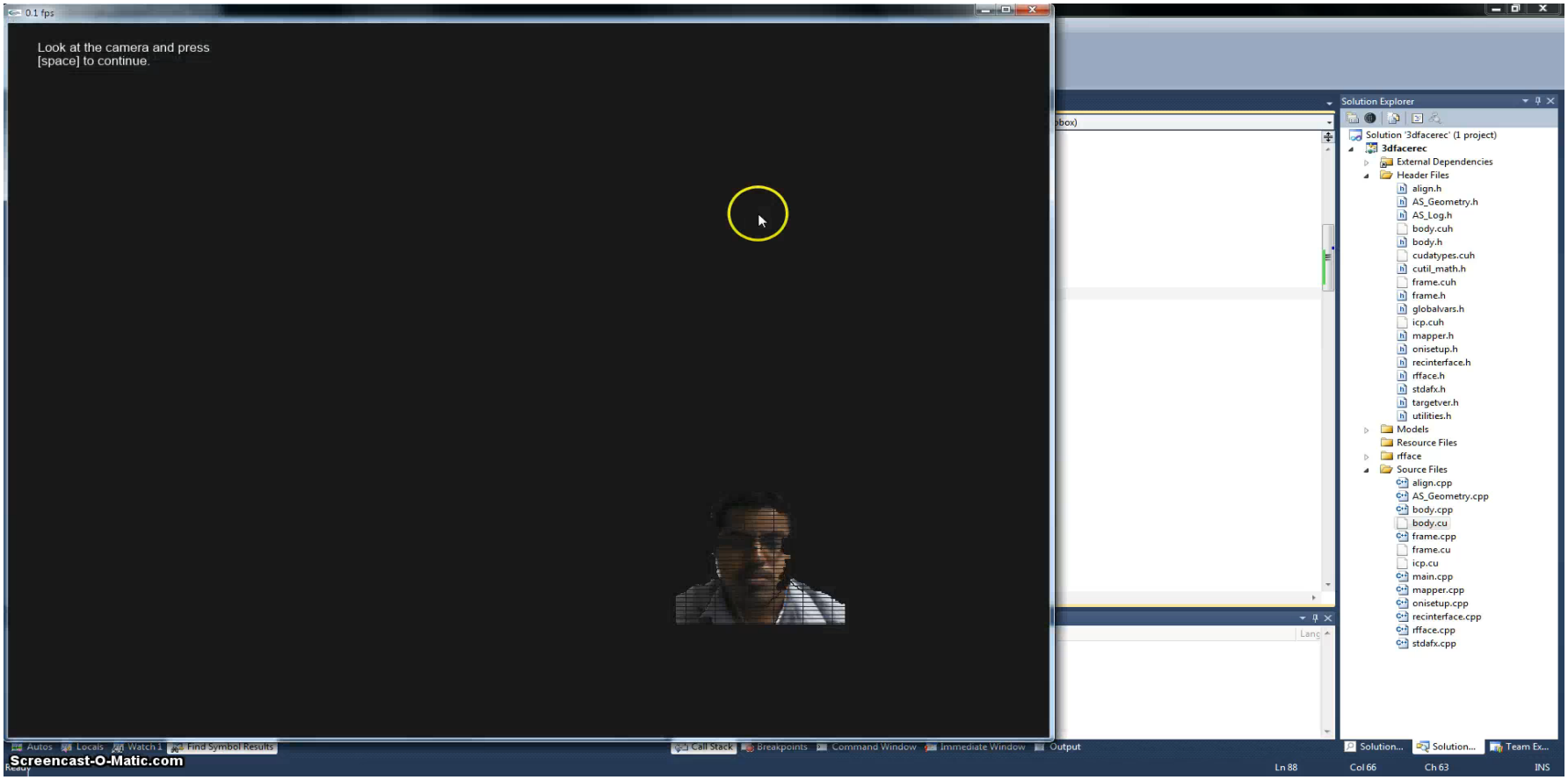
Bump Image with Texture Information



Bump Image Recording Local Geometry

3D Face Reconstruction using Bump Maps

- Bump map update takes 2ms on CPU
- Tracking is currently being ported from ICP to DVO, expected <5ms



3D Face Reconstruction using Bump Maps

- Reconstruct head while user turns left and right
- Use 3D geometry as an occlusion model

Physics Simulation

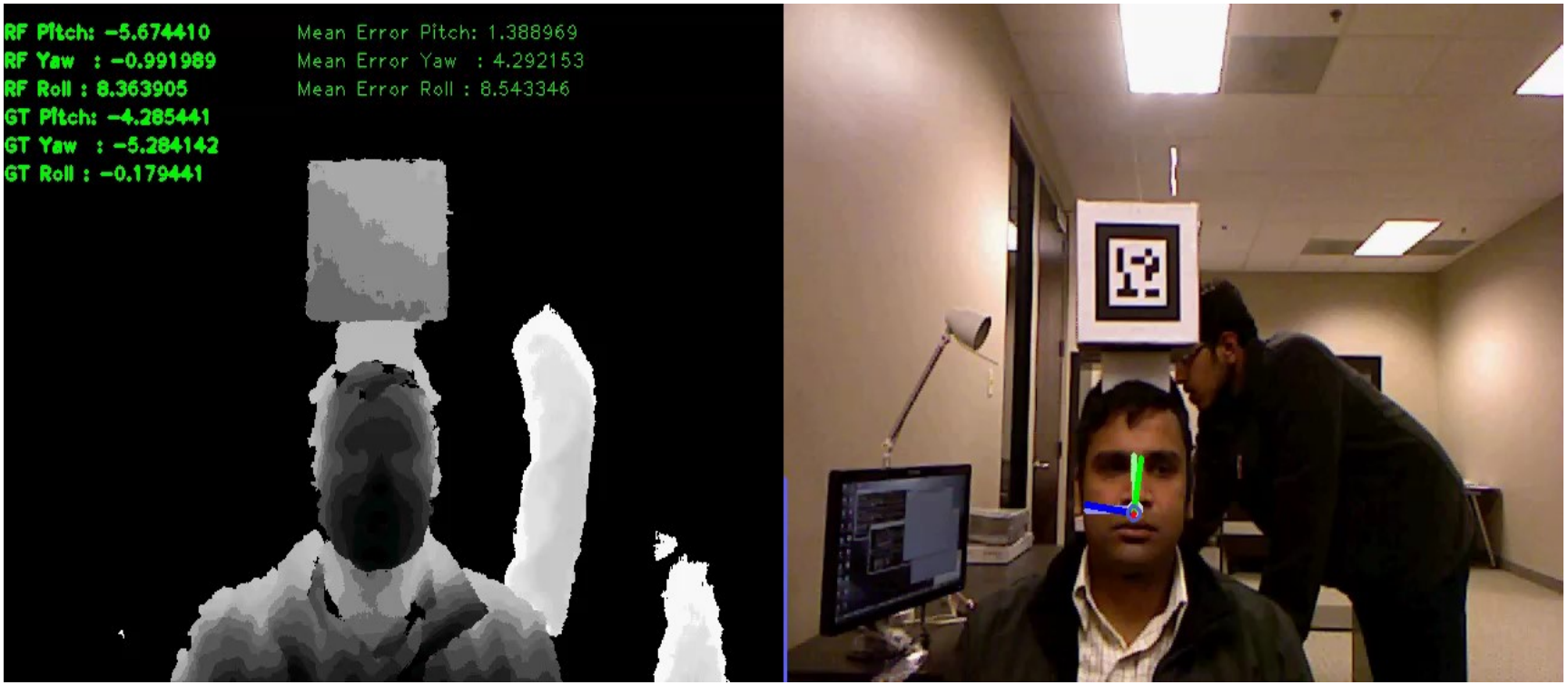
- Add physics / gravity to the models
- Unity engine

Initialization and Automatic Recovery

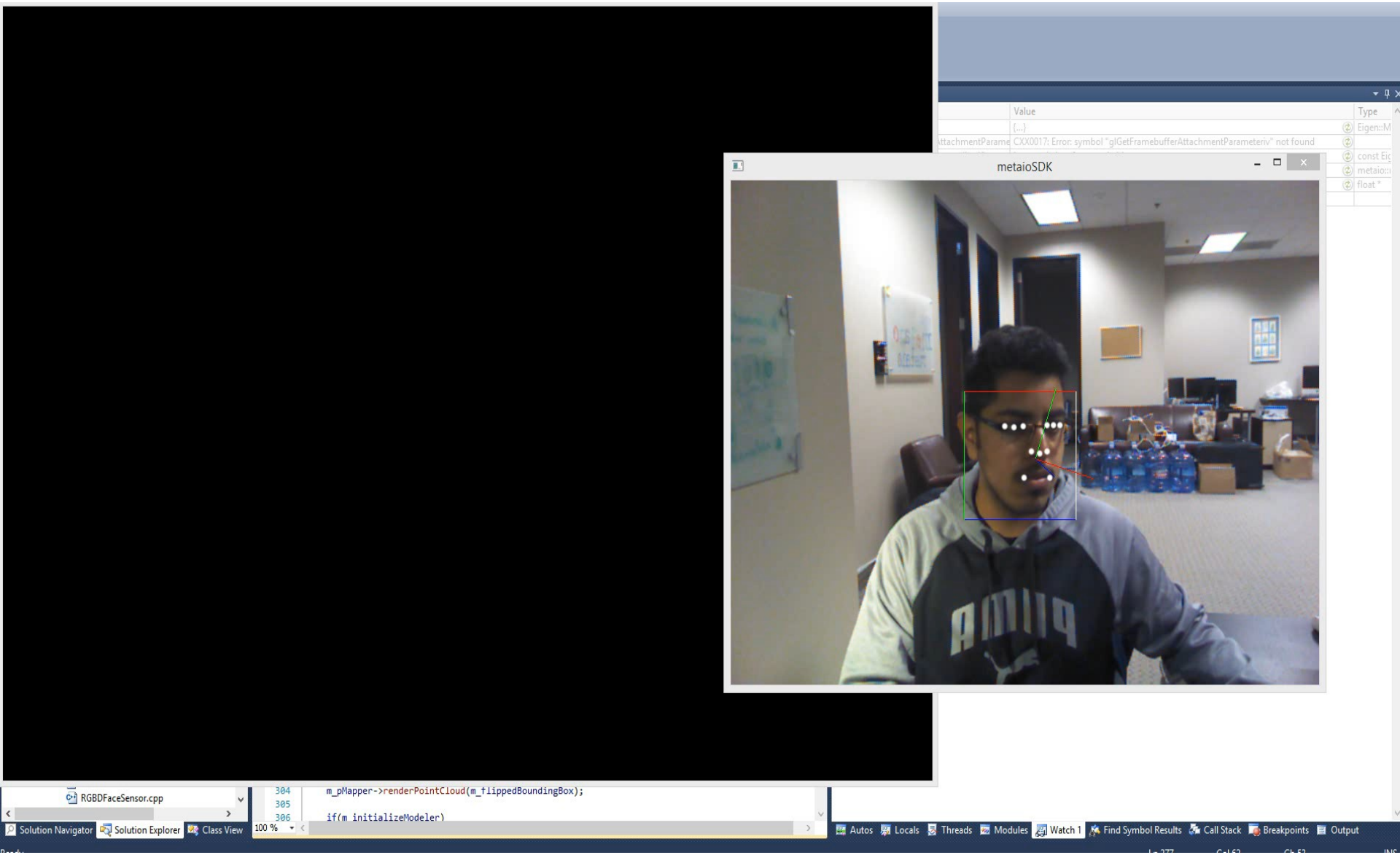
- How do we know what to reconstruct?
  - 2D face detection on the image
- How do we initialize the first pose?
  - 3D face pose estimation using random forests
- How can we recognize tracking loss?
  - Evaluate inlier ratio after ICP step

3D Head Pose Estimation using Random Forests

- Train a random forest that predicts 3D head pose from depth images
  - Pixel-wise classification head/no-head
  - Head pixels vote for center + head orientation
  - Estimate head center using mean shift and average inlier votes
- Recorded training dataset: 26 persons with marker on head



Re-initialization and non-frontal initialization



Demo

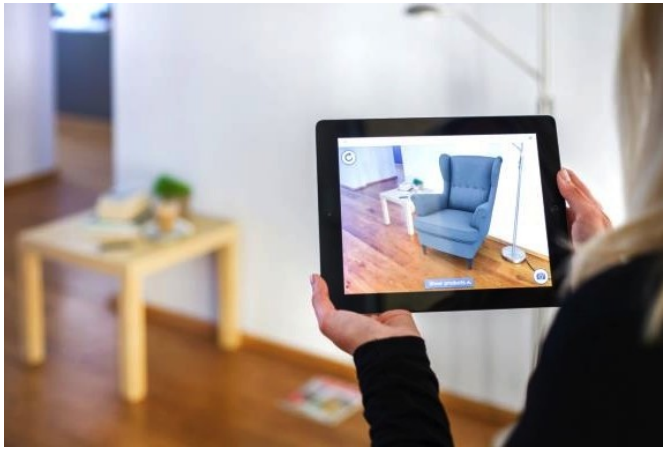
Can we do the same with a monocular camera?

- Shape alignment using random regression forests + ridge regression
- 68 fiducials (landmarks) per face
- Live demo

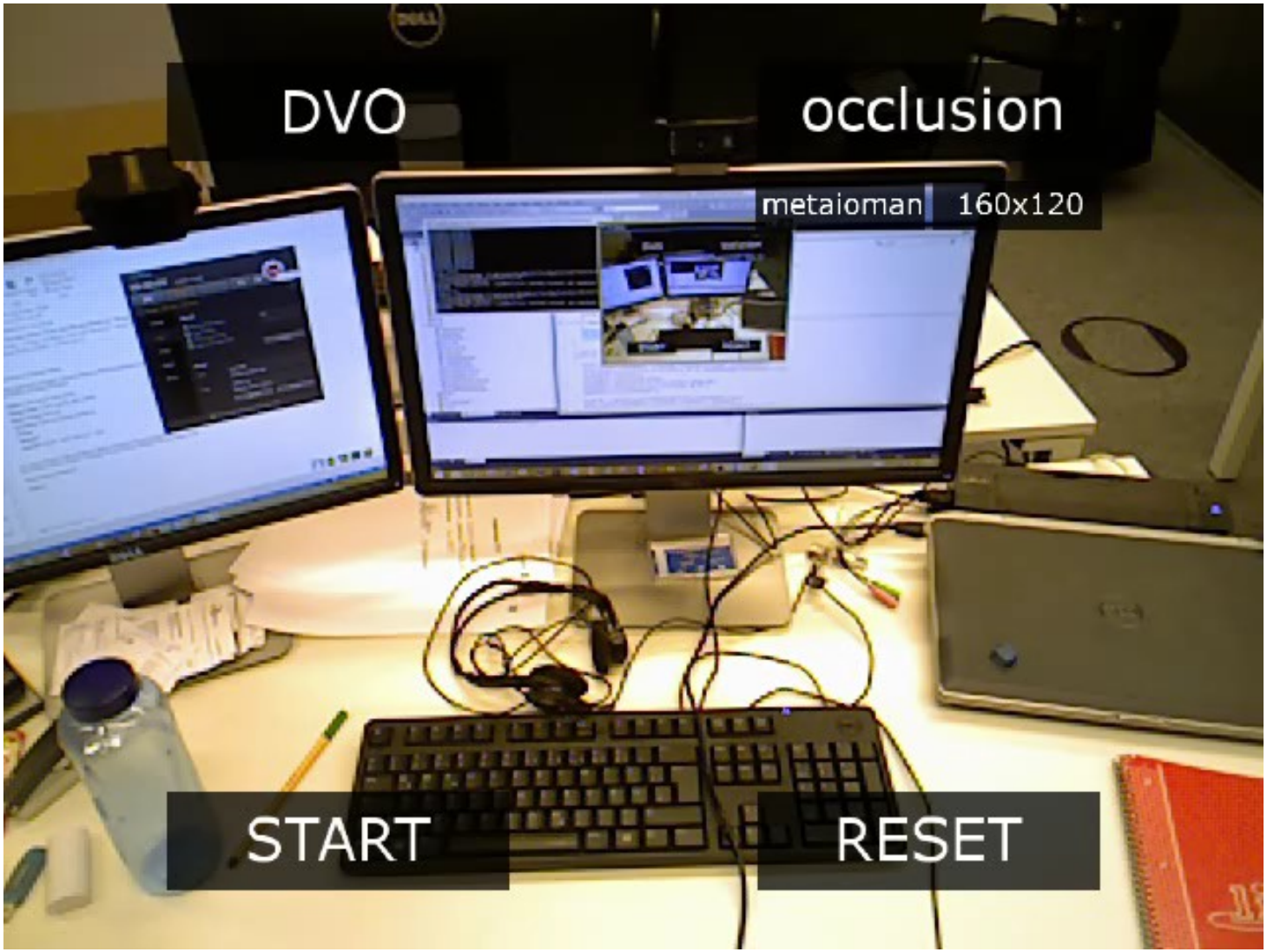
Face Alignment Demo

Virtual Shopping: Place in your room

- IKEA smartphone app
- Allows customers to place virtual furniture in their room
- Uses catalog as marker for SLAM & scale initialization
- **How can we improve the shopping experience?**
  - Occlusion modelling
  - Light estimation
  - Re-texturing of surfaces
  - 3D room scanning
  - Automatic room measurements



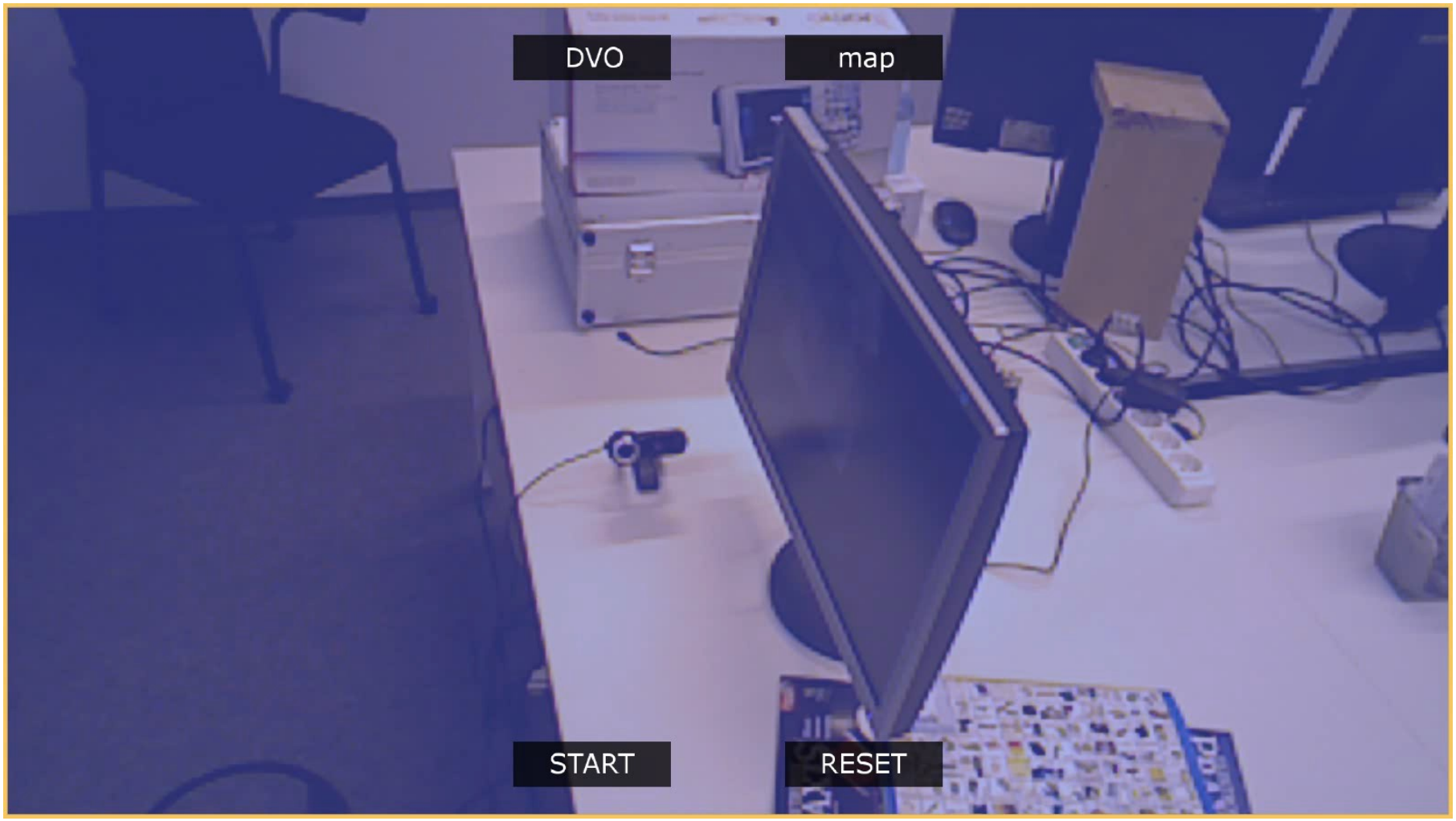
Occlusion Modeling



Re-Texturing of Surfaces

3D Room Scanning

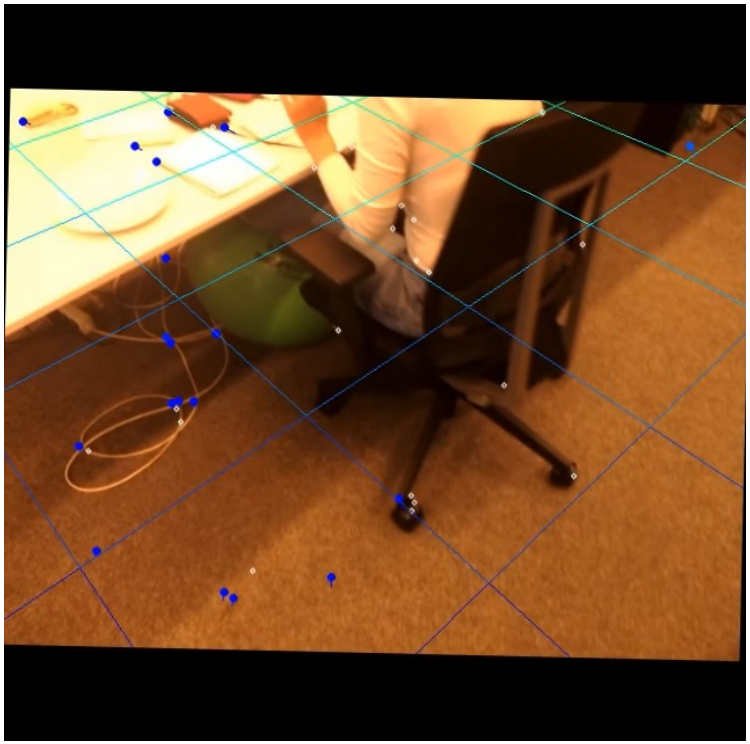
3D Room Scanning

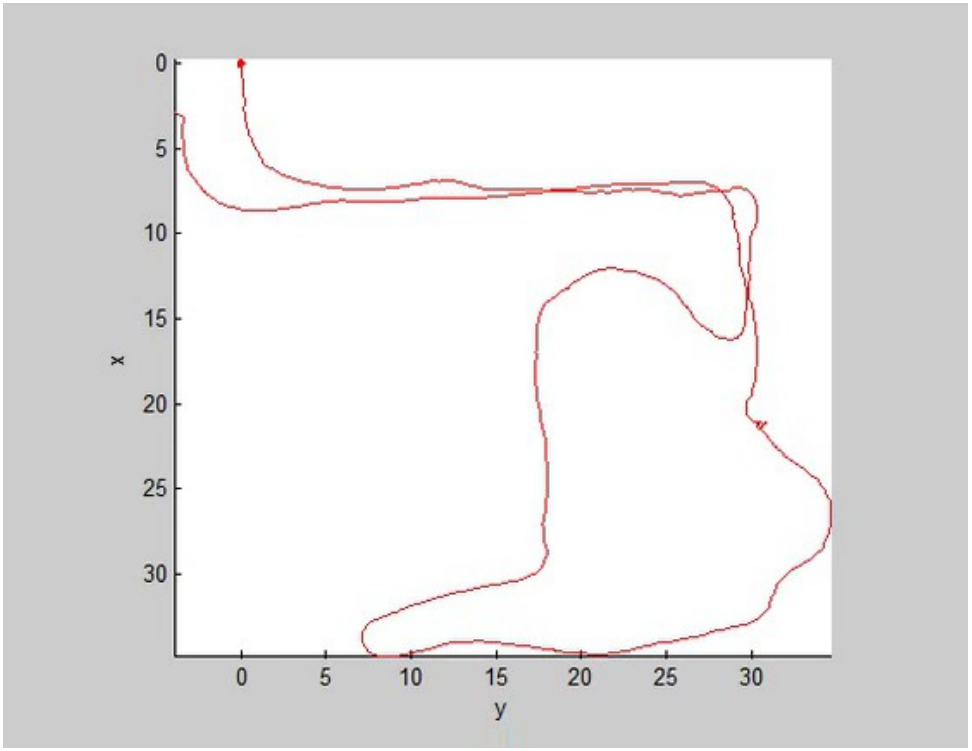


DVO Demo

Sparse inertial visual odometry (SIVO)

- Kalman filter-based approach
  - IMU is used in prediction step
  - Feature tracks are used in correction step
- Very robust, efficient, metrically correct





Summary

- Exciting field
- Hot topics
  - Dense methods for 3D tracking and reconstruction
  - Deep learning for face detection, pose estimation and alignment of fiducials
  - Approaches for tight IMU integration and data fusion
- Live demos
  - Virtual try-on of earrings
  - Face alignment
  - 3D room scanning / occlusion demo / re-texturing
- We‘re always looking for interesting speakers at our weekly research colloqium
- We‘re hiring!
  - Computer Vision
  - Machine Learning
  - Sensor/IMU data fusion
  - http://www.metaio.com/careers/

We're hiring!

Metaio
Phone (EMEA): +49-89-5480-198-0
Phone (US): +1-415-814-3376
info@metaio.com
www.metaio.com

http://www.facebook.com/metaio

@twitt_AR
http://twitter.com/#!/twitt_AR

http://augmentedblog.wordpress.com/

http://www.youtube.com/user/metaioAR