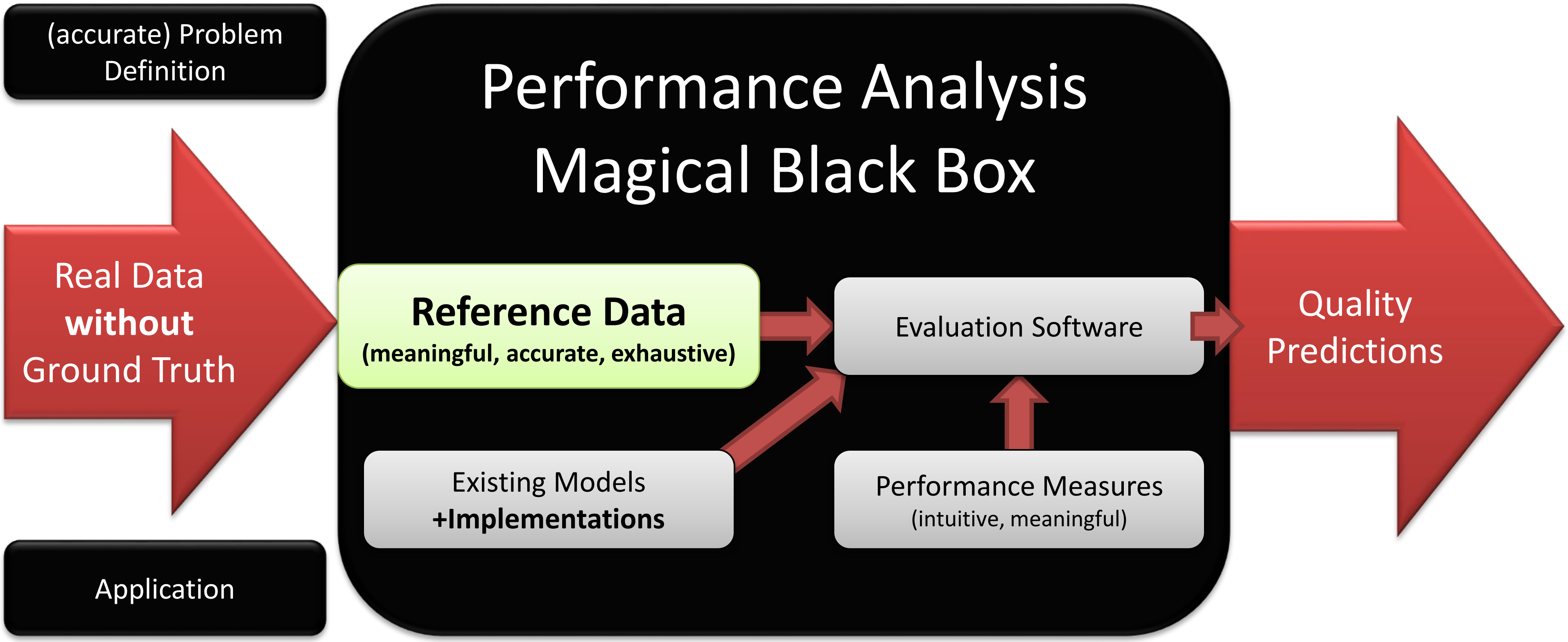
Heidelberg Collaboratory for Image Processing

Ground Truth Design Principles

~ Challenges, Approaches and The Future ~

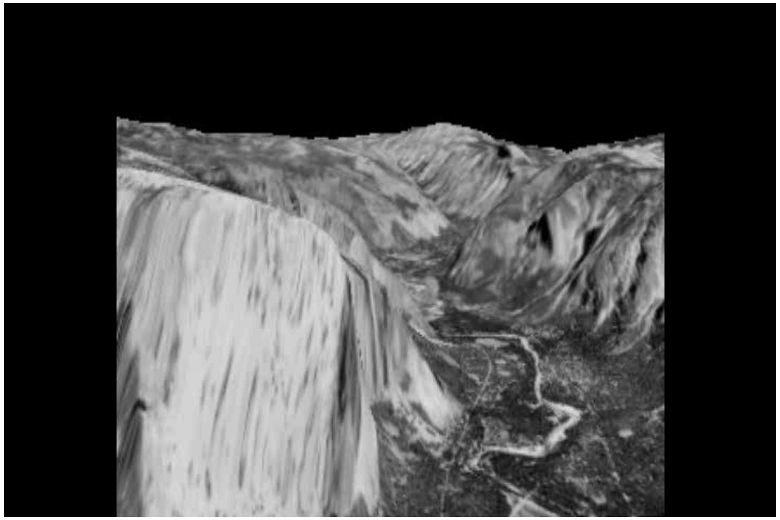
What is Performance Analysis?



„On Performance Analysis of Optical Flow Algorithms“: D. Kondermann, S. Abraham, G. Brostow, W. Förstner, S. Gehrig, A. Imiya, B. Jähne, F. Klose, M. Magnor, H. Mayer, R. Mester, T. Pajdla, R. Reulke, and H. Zimmer Dagstuhl LNCS Proceedings, Springer, 2012

Optical Flow Ground Truth 1994: Yosemite Sequence

Test Scene*

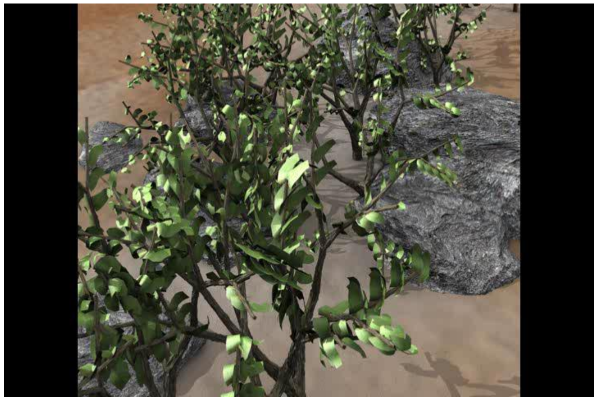


Real Scene

*J. Barron, D. Fleet, and S. Beauchemin. **Performance of optical flow techniques**. International Journal of Computer Vision, 12(1), 1994.

2007: Grove Sequence

Test Scene*



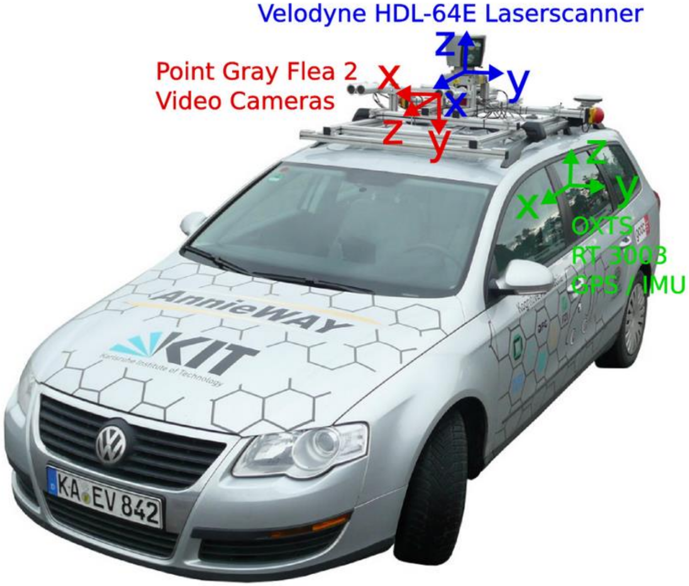
Real Scene

Crysis Scene (Game, 2007)

*Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M., Szeliski, R.: **A database and evaluation methodology for optical flow**. In: ICCV (2007)

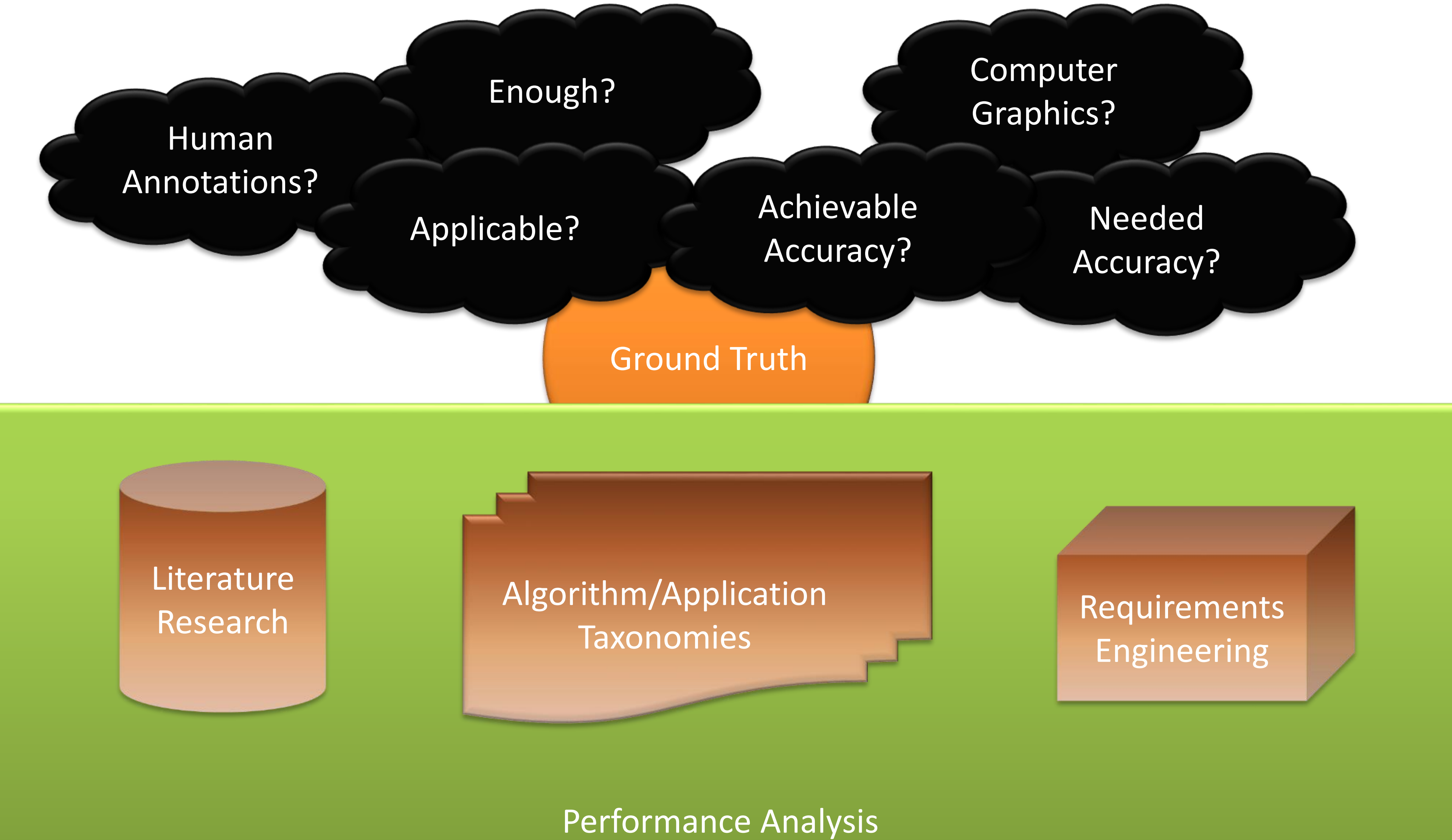
2012: New Datasets

MPI-Sintel - http://sintel.is.tue.mpg.de/



KITTI - http://www.cvlibs.net/datasets/kitti/

Central Questions



Reference Data

Without
Ground Truth

- Experts

With **weak**
Ground Truth

- Annotations
- More Data

With
Ground Truth

- Measurements
- Graphics

Reference Data **Without** Ground Truth

- Application Scenario:
  - Algorithms seriously fail all the time
  - Ground Truth acquisition currently impossible
  - Outdoor, fully dynamic environments
  - Extremely large amounts of data needed

Reference Data **Without** Ground Truth

- ECCV 2012 Workshop on
  - Unsolved Problems in Stereo and Optical Flow Estimation
- Robust Vision Challenge: eyeballed by expert jury
- http://hci.iwr.uni-heidelberg.de/Static/challenge2012

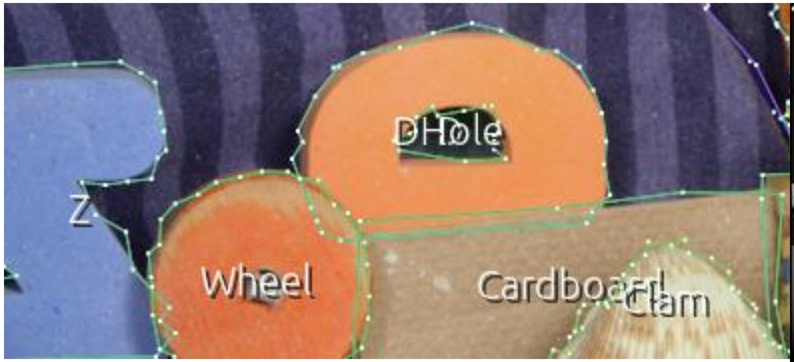
Heidelberg Collaboratory for Image Processing

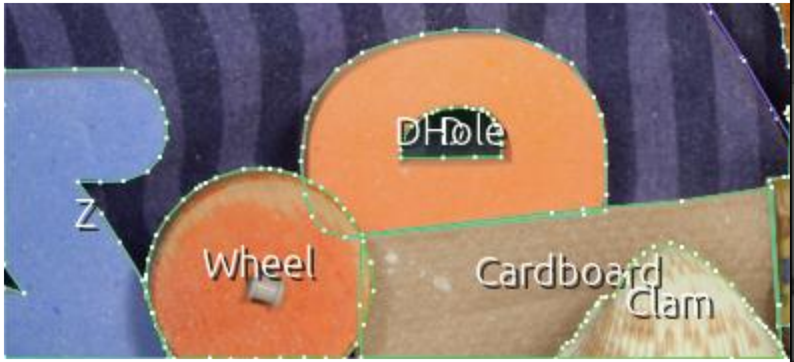
Reference Data With **Weak** Ground Truth

- Optical Flow Based on Human Annotations!?
  - Cost per frame has to be very small
  - Cannot ask users to annotate pixelwise flows
- Application Scenario:
  - No accurate measurement devices exist
  - Human brain performs better than algorithms
  - Very large amounts of data needed
  - Bias and other systematic errors less important

Weak GT With MTurk:

Input based on Ce Liu's work





Annotations by Crowd

Endpoint Error (0..2px)

Axel Donath, ICVS 2013 (Best Paper Award)

Reference Data **With** Ground Truth

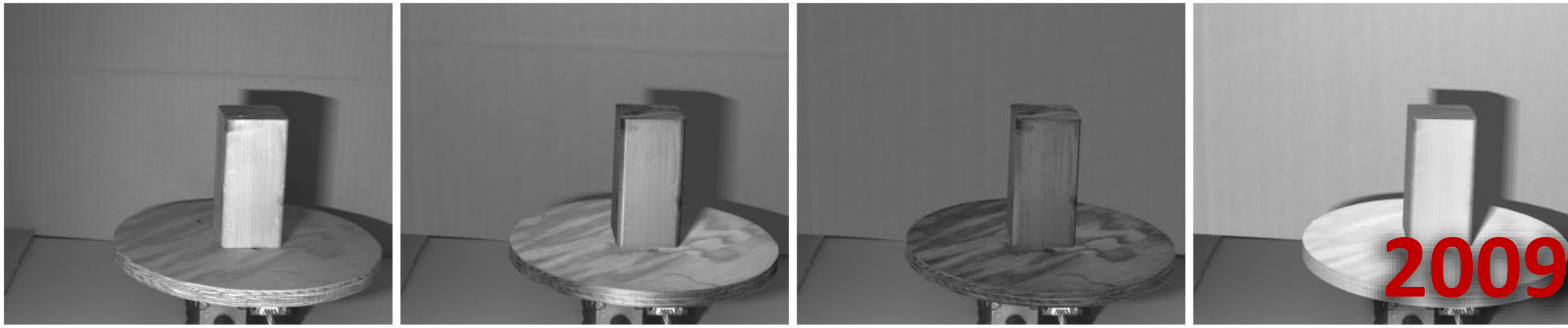
- Application Scenario
  – Very accurate data needed, parameters tuneable
  – Simulations are realistic enough
- **Option 1**: computer graphics
  – Perfect GT, but no real images!
- Option 2: measurement devices
  – One magnitude more accurate
  – Not existing for most low-level vision problems

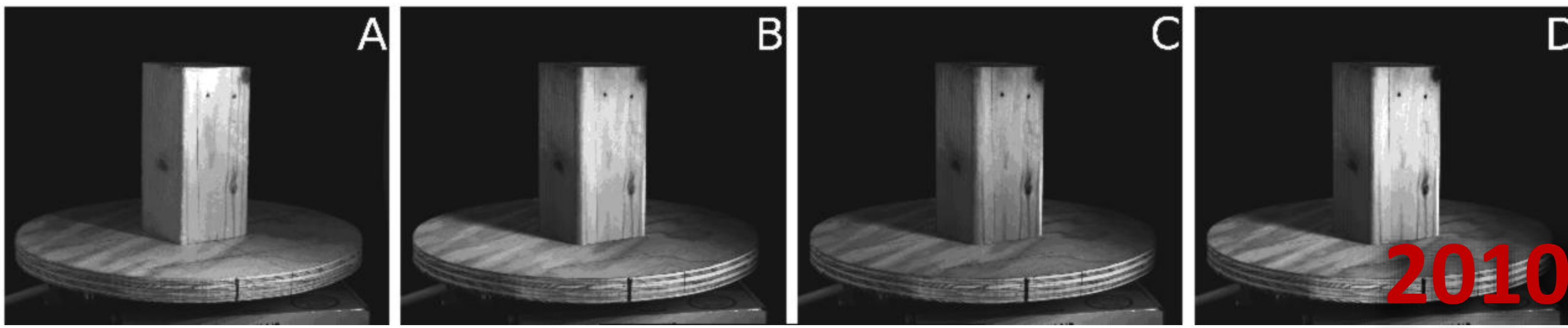
Real vs. Synthetic Images

Real world → Geometry → „GT“ with real images | Dynamics?

Other Questions:
- Suitability?
- Content?

Reference Data **With** Ground Truth

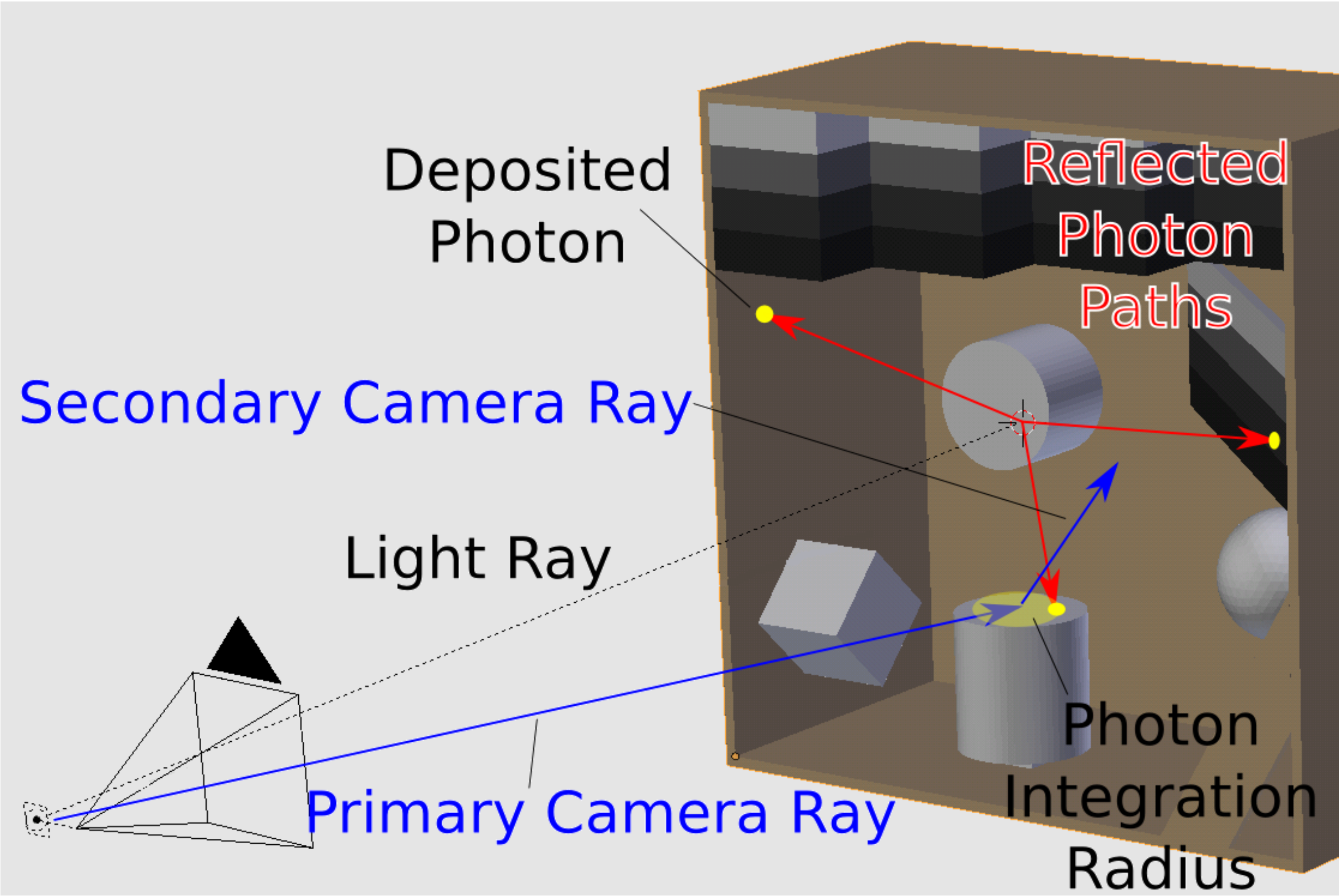




New Experiments (Lightfields)

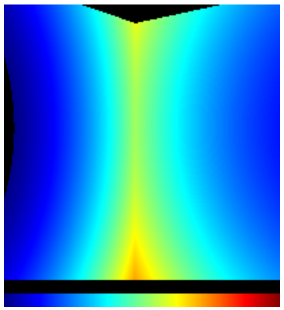
Multipath Simulation

- Global illumination to simulate light propagation

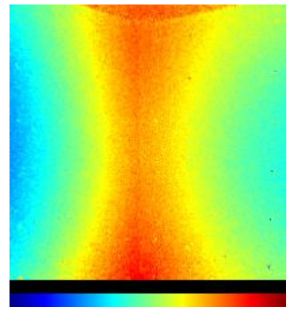


Heidelberg Collaboratory for Image Processing

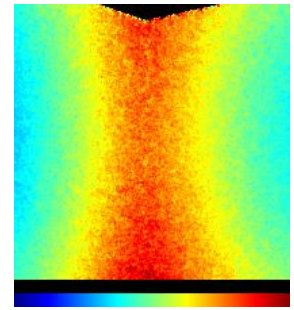
Results of ToF Simulation



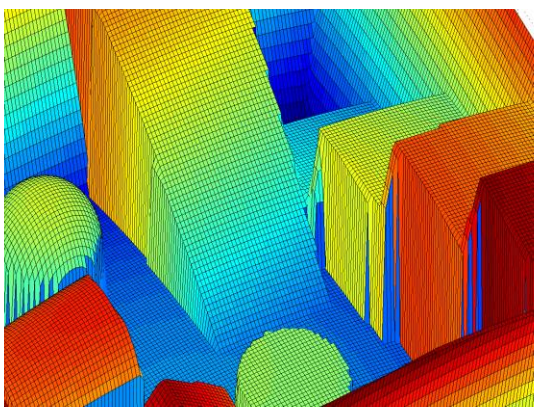
Real depth



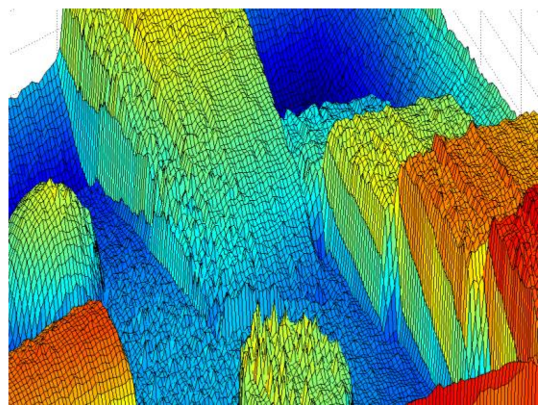
Measured depth



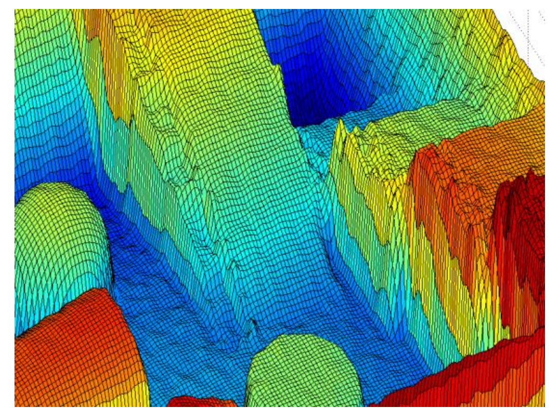
Simulated depth



Real geometry



Measured geometry



Simulated geometry

Reference Data **With** Ground Truth

- Application Scenario
  - Very accurate data needed, parameters tuneable
  - Simulations are realistic enough
- Option 1: computer graphics
  - Perfect GT, but no real images!
- **Option 2**: measurement devices
  - One magnitude more accurate
  - Not existing for most low-level vision problems

Main Problem: Geometry

<u>Tried</u>

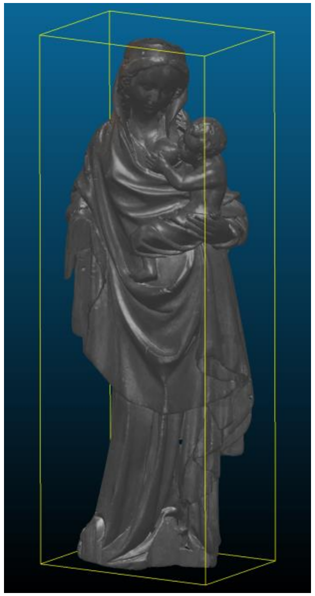
Manual Computer Graphics (Blender, Maya)
Games (CryEngine, Source Engine)
High-end 3D Scanners (Lidar, Structure Light)
Manual Measurements

…usually way too slow!

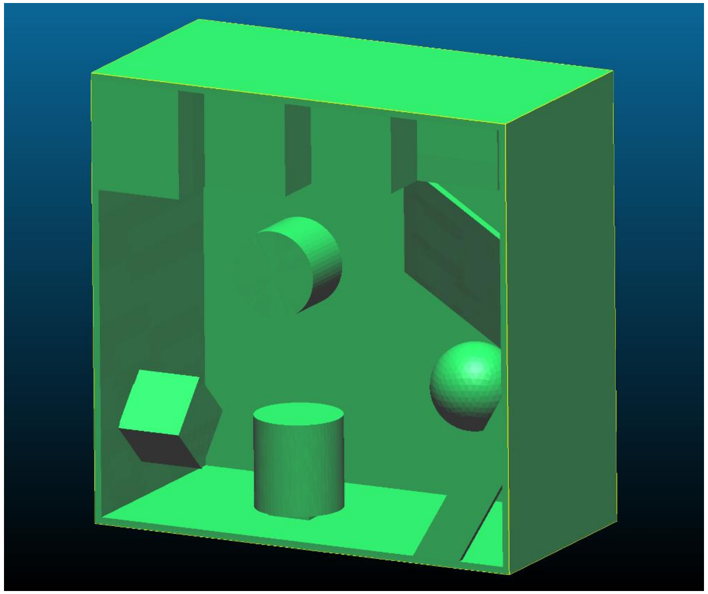
Flashback: **Weak** Ground Truth

- Low-end 3D Scanners (Kinect, TOF, …)
  - Idea: known camera track and geometry
  - Works for static scenes with moving cameras only
  - Accuracy of system needs to be known very well
- Application Scenario
  - Mostly controlled environments
  - Small but accurate amounts of data needed

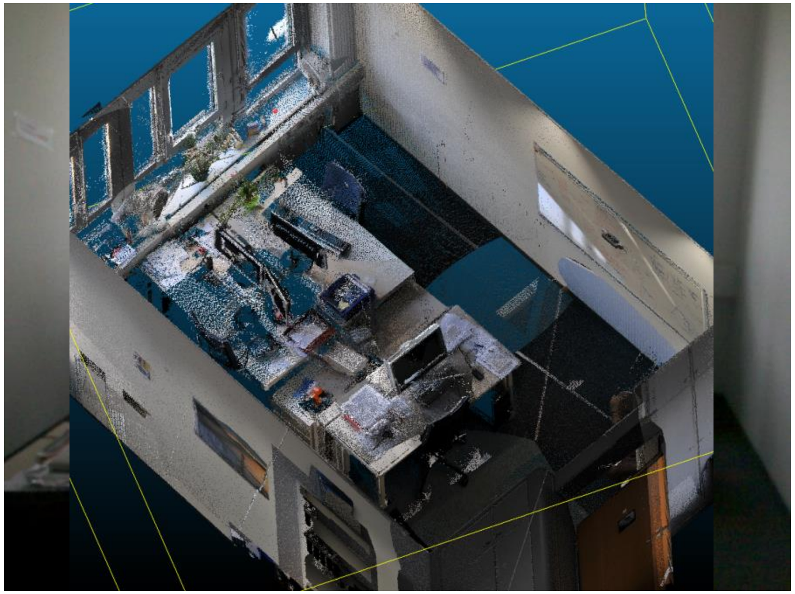
Fast Geometry: KinectFusion



@0,05mm
30cm



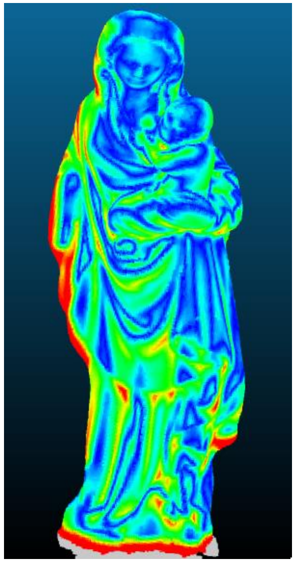
@1mm
150cm

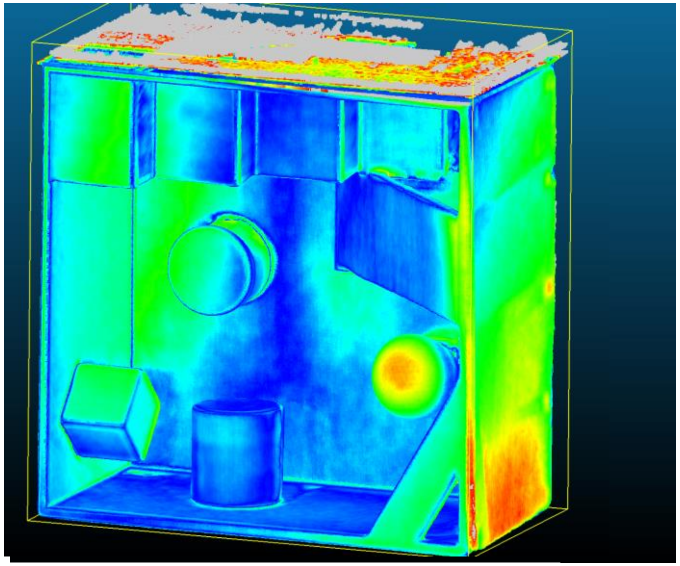


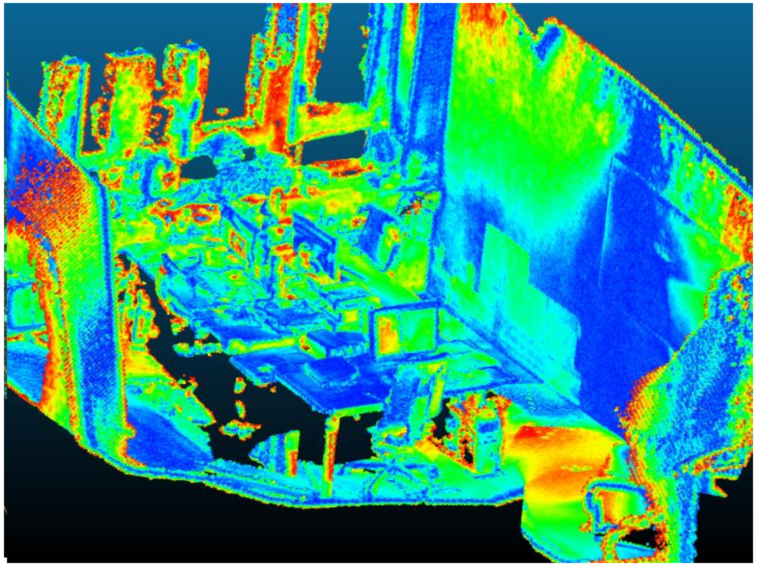
@5mm
700cm

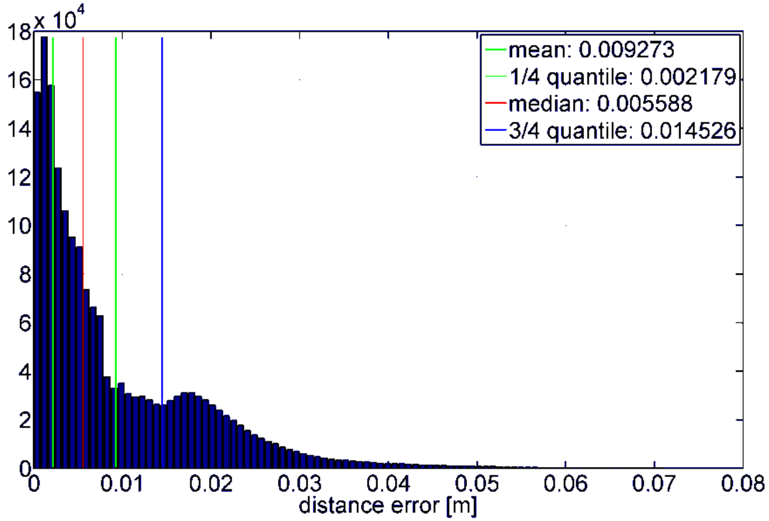
Heidelberg Collaboratory for Image Processing

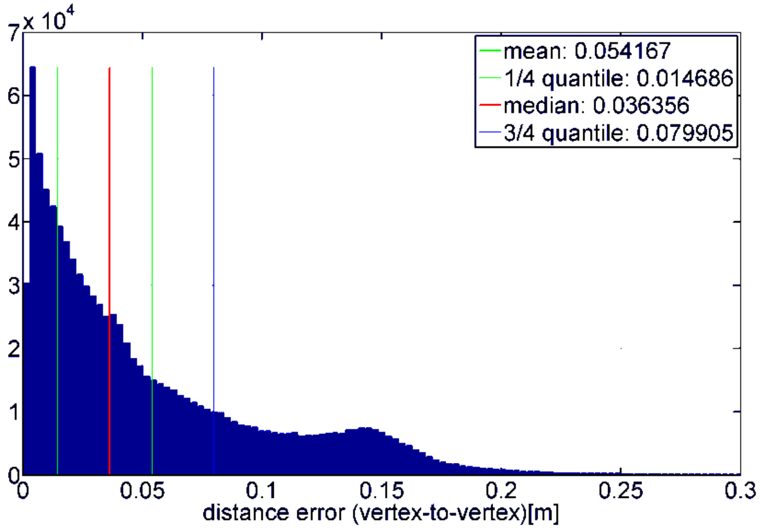
Fast Geometry: KinectFusion











Heidelberg Collaboratory for Image Processing

Conclusion

- Ground Truth generation:
  - Accuracy, Cost of Acquisition, Amount of Data
    - Needs Geometry, Radiometry, System/Camera Model
  - Content
    - Problem Definition, Amount of Data, Type of Data
  - Hardware vs. Software
  - Humans vs. Measurement Devices
  - Real vs. Synth

Heidelberg Collaboratory for Image Processing

Pallas Ludens

Need Ground Truth or Training Data?

Low Cost
Good Accuracy
Large Amounts
Full Service

daniel.kondermann@iwr.uni-heidelberg.de


DANIEL KONDERMANN