3D Object Reconstruction with Heterogeneous Sensors

- combining camcorders & 3D depth cameras

Li Guan [1,2], lguan@cs.unc.edu

Jean-Sébastien Franco [3,4], Jean-Sebastien.Franco@labri.fr

Marc Pollefeys [1,2] marc.pollefeys@inf.ethz.ch

1. UNC-Chapel Hill, U.S.A.
2. ETH Zürich, Switzerland
3. LaBRI, INRIA Sud-Ouest, France
4. University of Bordeaux, France

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich

3D Depth Sensor (RIM sensors)

- Theory
  - Time-of-Flight

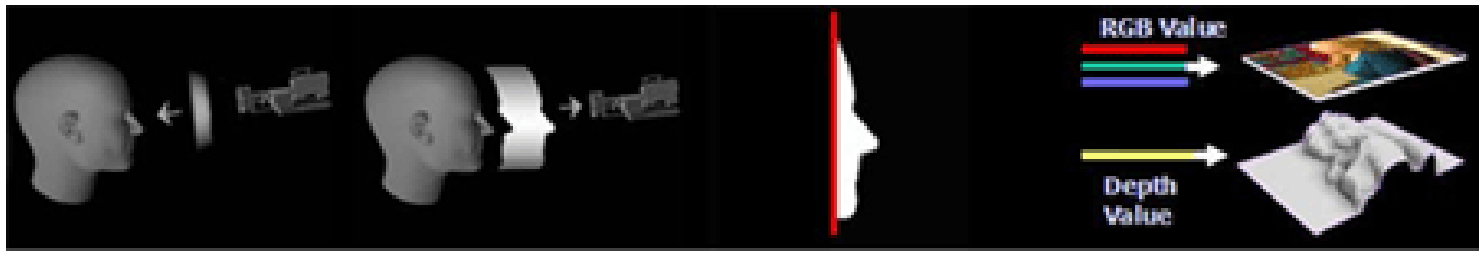
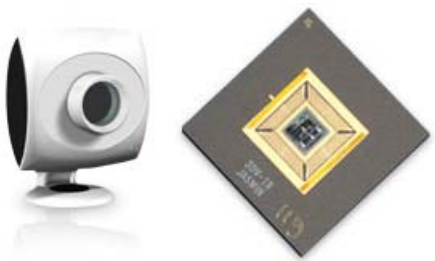


Fig. from 3DV system website

- Products
  - Canesta camera
  - Swiss Ranger (res. 176*144, up to 50fps)
  - PMD camera
  - ZCam (<$100)



3D Depth Sensor (cont.)

- Applications
  - Robotics planning
  - Gaming, HCI, Virtual Reality
  - Biometric/Medical
  - 3D object reconstruction

http://www.3dcgi.com/images/face_2d_3d.jpg

- Challenges
  - Low resolution
  - The depth precision is fine, but accuracy is low
    - Active sensing – reflectance (could be 40cm)

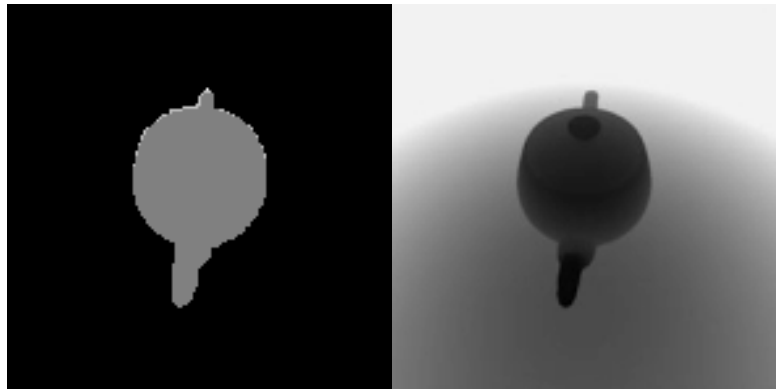
Fig from wikipedia

General Task

- Goal: 3D dynamic object (e.g. human) modeling with geometrically calibrated fixed pose sensors.
  - Video camcorder (silhouette cues)
  - 3D depth cameras (depth maps)
  - More…

Motivation

- Complimentary Sensor Observations



silhouette depth map

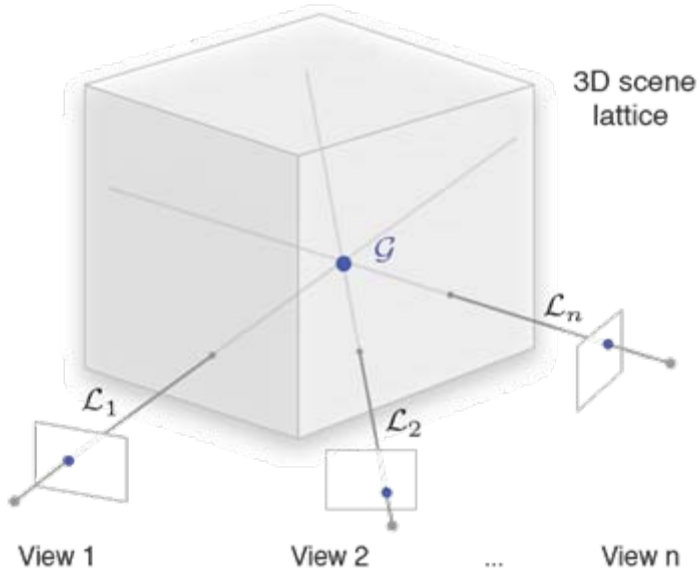
- Color Textures
- High fps enables dynamic scene reconstruction

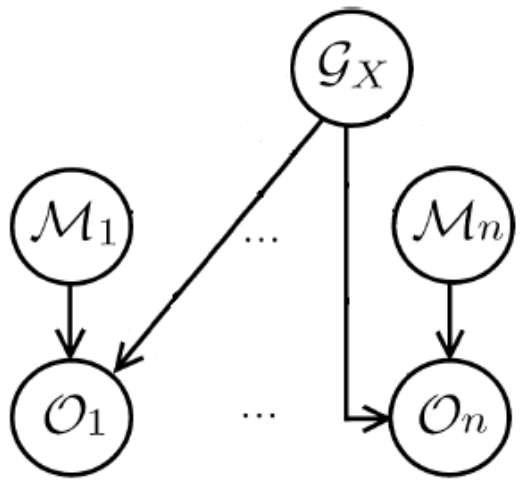
Previous Works

- Depth map based methods
  - Patch alignment, and hole filling
    - [Bajaj 95, Hilton 98, Davis 01, Casciola 05, Curless 06, ...]
- Silhouette based methods
  - Concepts & algorithms
    - [Baumgart 74, Laurentini 94, Szeliski 93, Franco 00, Lazebnik 07, ...]
  - Robust reconstruction
    - [Franco 05]
- Sensor fusion for reconstruction
  - Depth image with silhouette information
    - Surface representation [Li 96]
    - Volumetric representation [Sablatnig 02, Yemez 07]

Sensor Fusion Framework

- Notations
  - $\mathcal{G}$ as the binary state space
  - $\mathcal{M}_i$ as the sensor models
  - $\mathcal{O}_i$ as the sensor observations



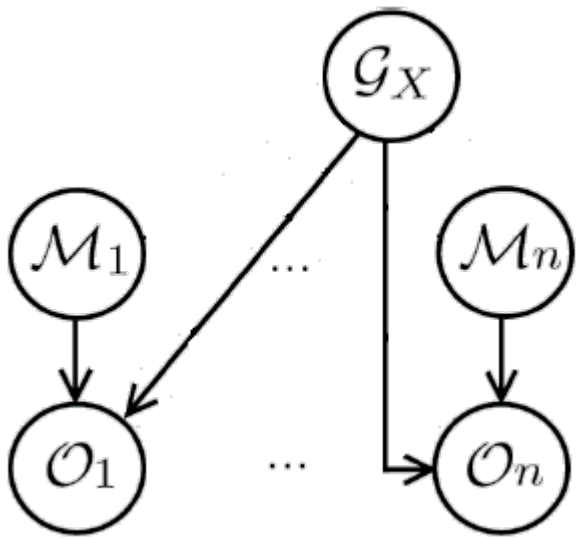


Main Formula

- Bayes rule

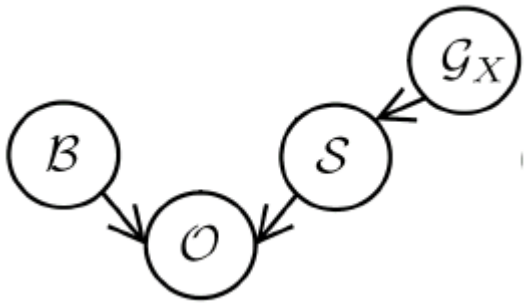
$$p(\mathcal{G}_X|\mathcal{O}^p_{1,\ldots,n},\mathcal{M}_{1,\ldots,n})$$

Observation Formation – Sensing model



Sensor Model for Video Camera

[Franco & Boyer, ICCV05]



$$p(\mathcal{O}_i^p|\mathcal{G}_X,\mathcal{M}_i^p) = \sum_{\mathcal{S}_i^p} p(\mathcal{O}_i^p|\mathcal{S}_i^p,\mathcal{B}_i^p)\,p(\mathcal{S}_i^p|\mathcal{G}_X)$$

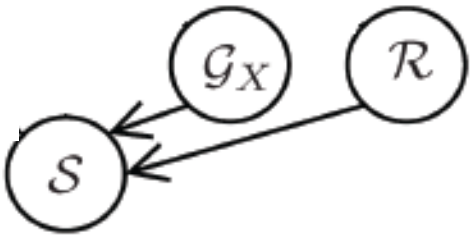
Image formation term

Silhouette formation term

- Background model $\mathcal{B}_i^p$
  - RGB Gaussian mode for every pixel $\mathcal{N}(\mathcal{O}_i^p, \mu_i^p, \sigma_i^p)$
- Binary silhouette hidden variable $\mathcal{S}_i^p$

Sensor Model for Video Camera (cont.)

[Franco & Boyer, ICCV05]



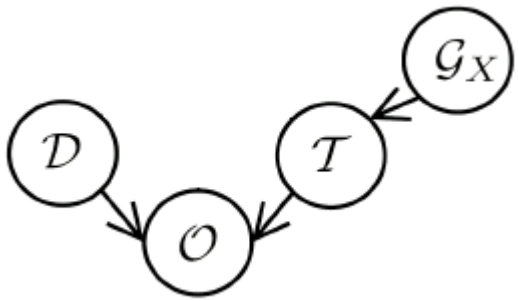
- External detection cause
  - Likelihood of other voxels on the same voxels
- Sampling variable
  - To model error of the “silhouette sensor” errors, due to geometric calibration, mis-synchronization, etc.

Sensor Model for 3D Depth Camera

$$p(\mathcal{O}_i^p|\mathcal{G}_X,\mathcal{M}_i) = \int_0^{d_{max}} p(\mathcal{O}_i^p|\mathcal{T}_i^p,\mathcal{D}_i^p)\,p(\mathcal{T}_i^p|\mathcal{G}_X)\,d\mathcal{T}_i^p$$

Depth measurement term

Surface formation term



- Continuous surface front hidden variable $\mathcal{T}_i^p \in [0, d_{max}]$
- Depth measurement term
  - Gaussian mode for every pixel $\mathcal{N}(\mathcal{T}_i^p, \bar{\sigma})$

precision

Sensor Model for 3D Depth Camera (cont.)

- Surface formation term, for *Gx*

$$p(\mathcal{T}_i^p|[\mathcal{G}_X=1]) = \begin{cases} 1/d_{max} & \text{if } \mathcal{T}_i^p < d_X - \epsilon \\ (1-d_X/d_{max})/\epsilon & \text{if } d_X - \epsilon \leq \mathcal{T}_i^p \leq d_X \\ 0 & \text{if } \mathcal{T}_i^p > d_X \end{cases}$$

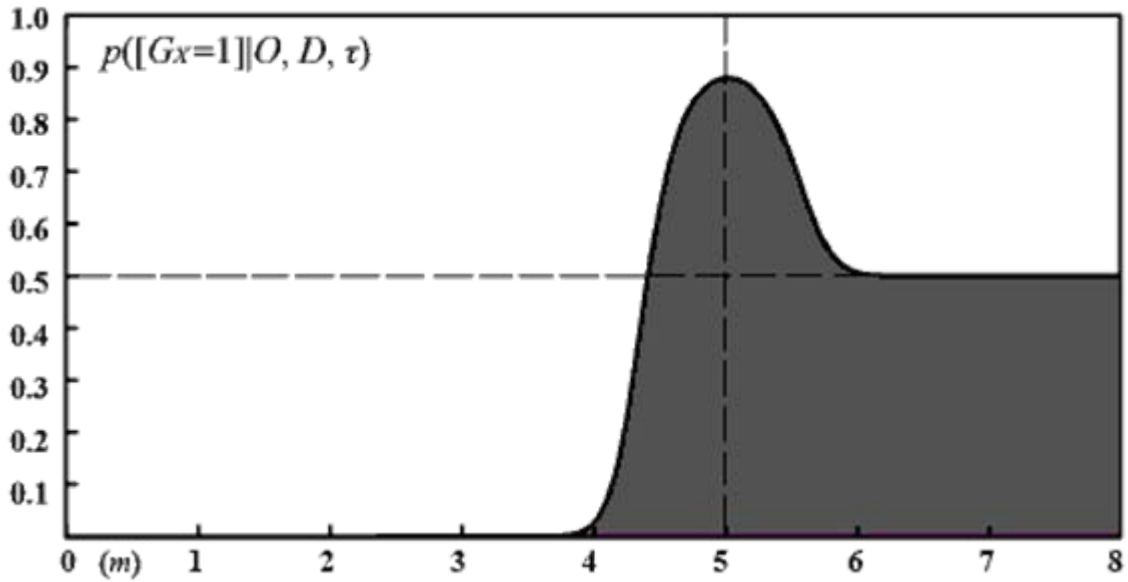
$$p(\mathcal{T}_i^p|[\mathcal{G}_X=0]) = 1/d_{max}$$

Synthetic example:

$\sigma = 0.3m$

$d_{max} = 8m$

$\mathcal{O} = 5.0m$



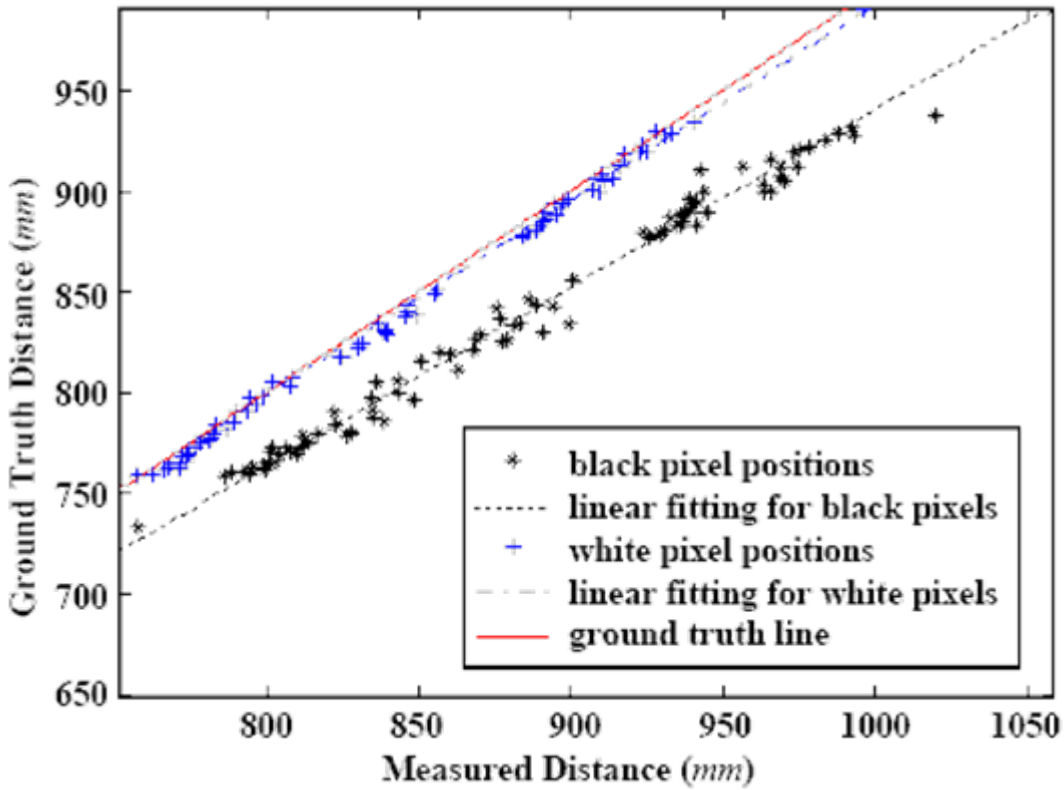
Sensor Specs.

- **Canon HG10 HDD camera**
  - 1920 by 1080 RGB
  - 25 fps

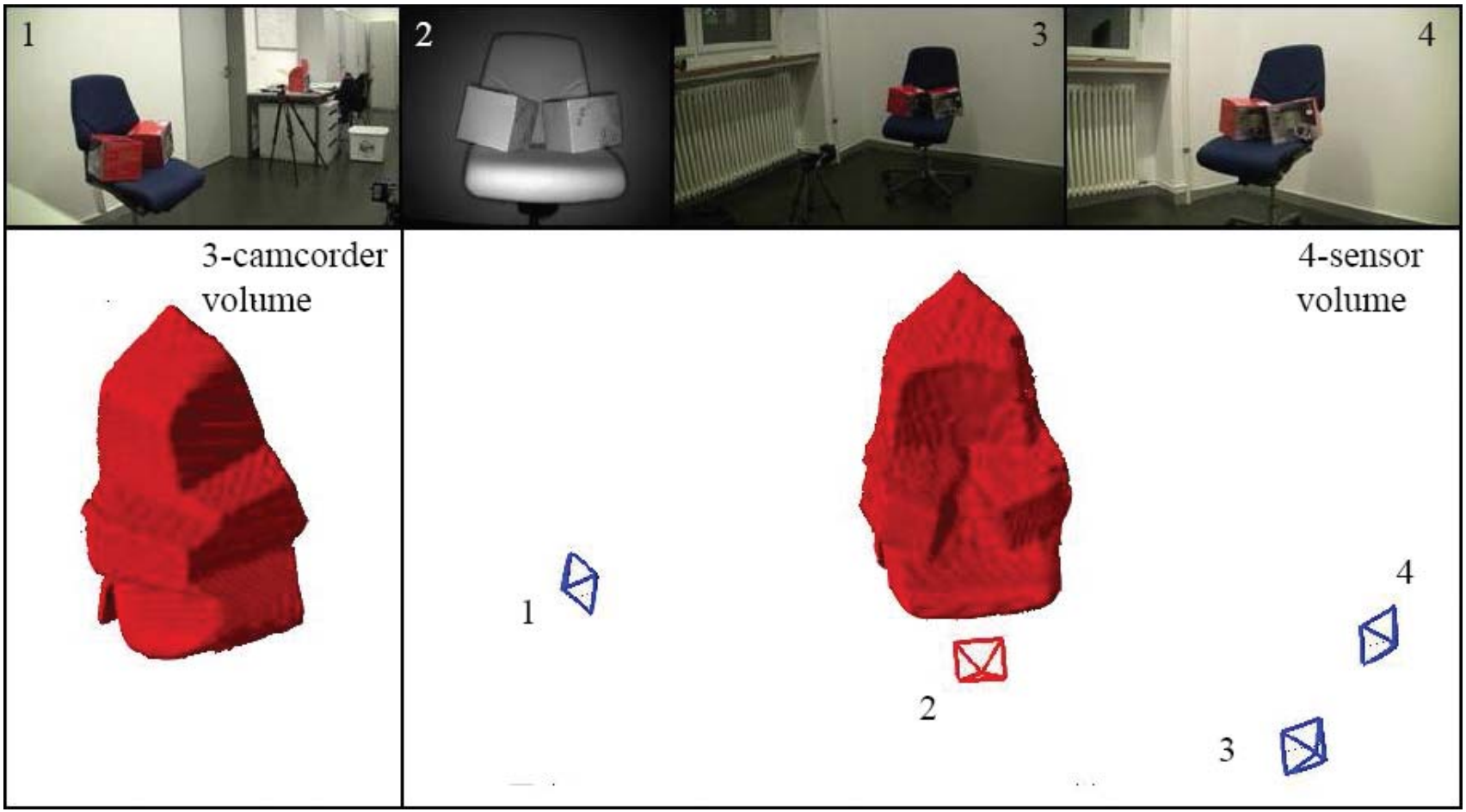
- **Swiss Ranger 3100**
  - 176 by 144
  - 5 fps
  - 19MHz, 20 MHz, 21 MHz (smallest maximum depth = 7.1 *m*)

Geometric Calibration

- Checkerboard calibration [Z.Zhang 00]
- Bundle adjustment considering radial distortions
- Linear Depth calibration refinement



Results

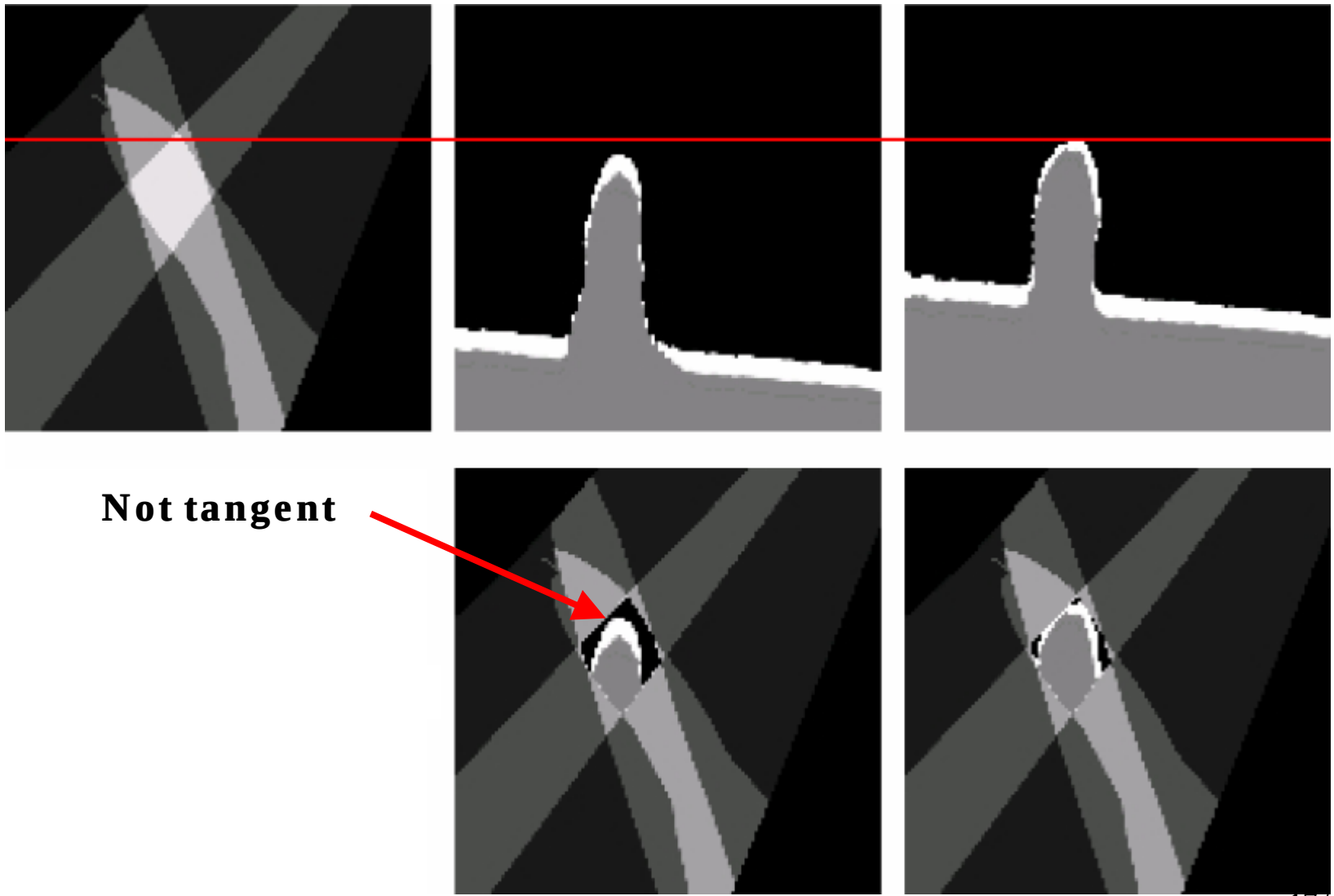


Results (cont.)

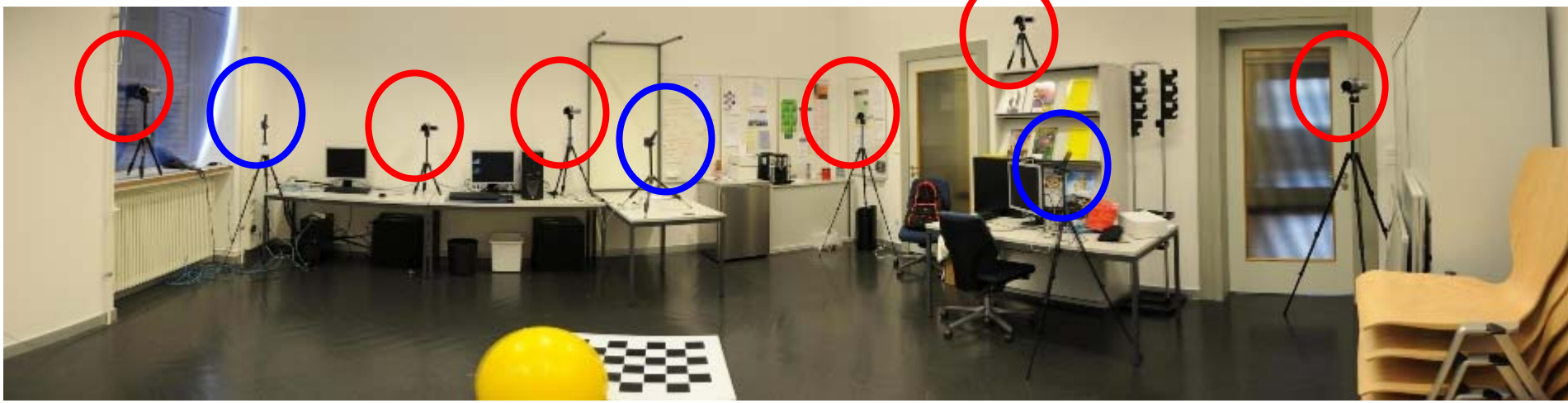
3-camcorder volume

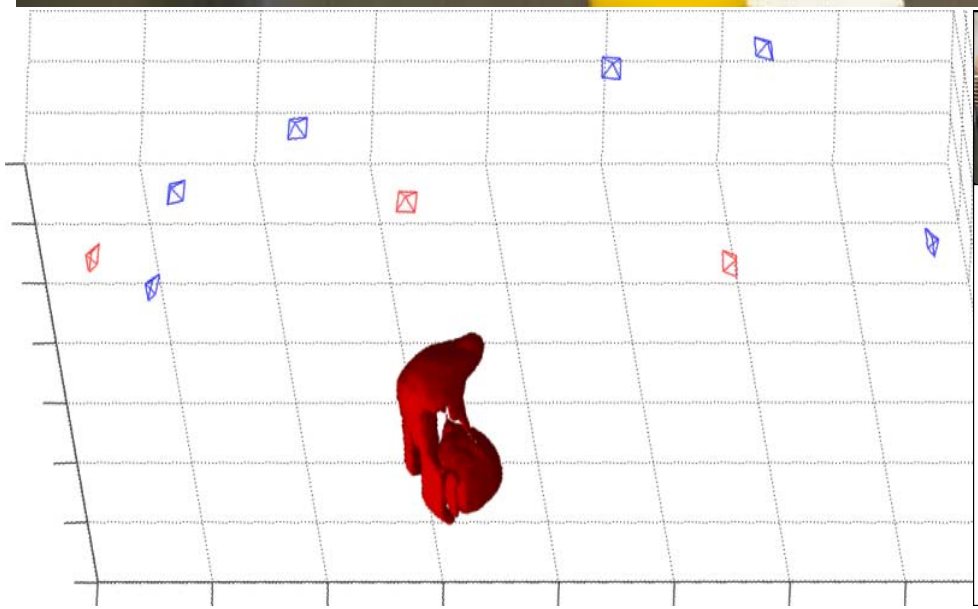
4-sensor volume

Results (cont.)



Results (cont.)





Future Work

- Geometric Calibration
  - Resolution
  - Depth calibration
- More general sensor fusion
  - Stereo camera
- Ultimate challenge of outdoor environment
- Synchronization and video processing
- Algorithm speedup

Summary

- General heterogeneous sensor fusion framework
- 3D depth camera sensor model for reconstruction
- Attempt to test depth calibration
  - Resolution
  - Sensor output