# Computer Vision Systems Programming VO

## 3D Vision Applications

Christopher Pramerdorfer

Computer Vision Lab, Vienna University of Technology

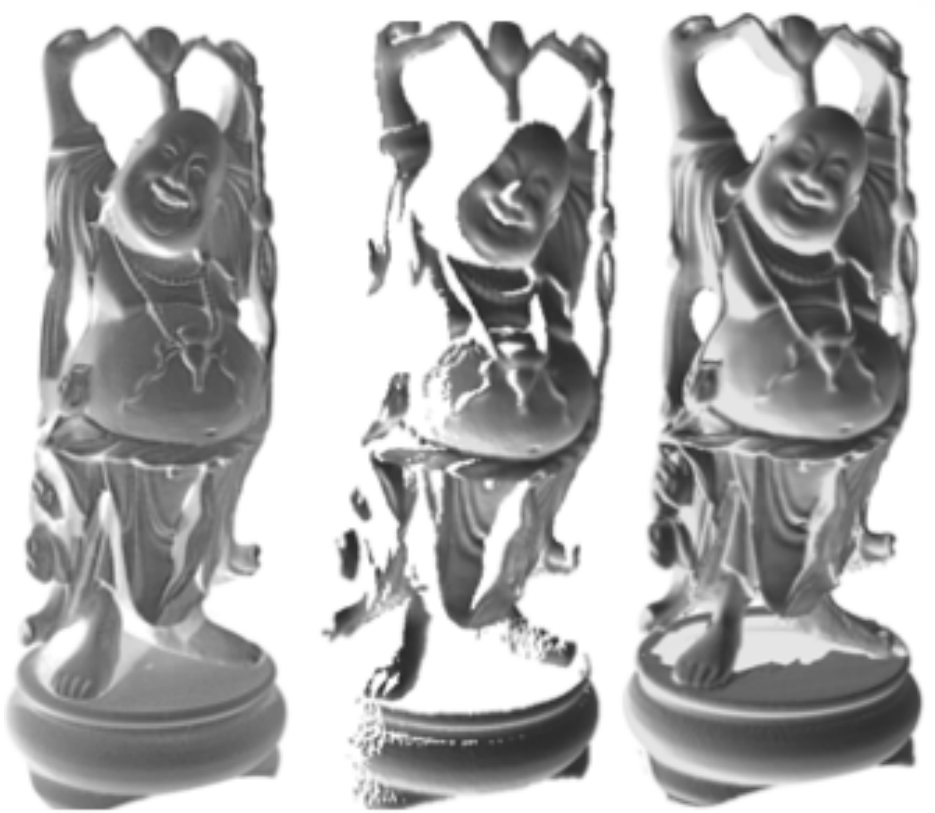
# Topics

CV applications utilizing scene geometry (3D data)

- Focus on those based on Kinect

Images by Ryuzo Okada, Shotton et al. 2011, Newcombe et al. 2011

# 3D Reconstruction



Images from Curless and Levoy 1996

# 3D Reconstruction

The obvious thing we can do is generate 3D models

- Usually involves combing multiple point clouds

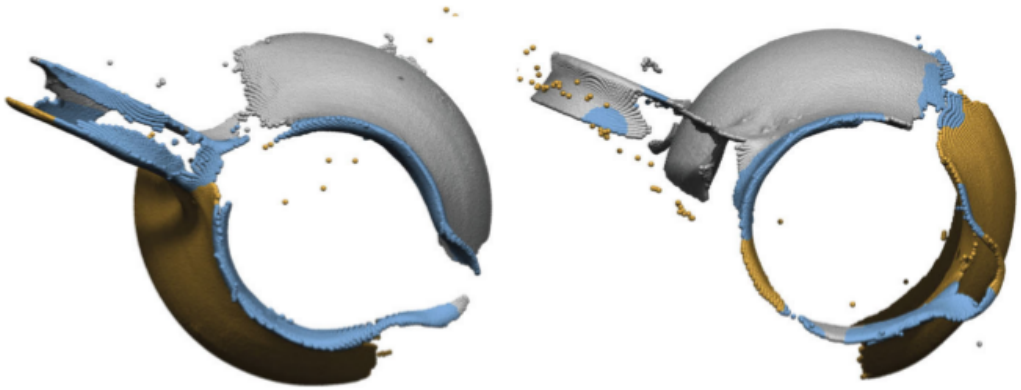
Accomplished in two steps

- Align range data
- Merge range data in a way that minimizes errors

# 3D Reconstruction
## Range Data Alignment – Iterative Closest Points

Popular method for aligning two point clouds $\{\mathbf{r}\}$, $\{\mathbf{s}\}$

- Goal is to find parameters $\psi$ of some transformation $\mathcal{T}$
- Usually assuming a rigid transformation



Images from Aiger, Mitra, and Cohen-Or 2008

Algorithm iterates between

- Finding point correspondences based on distance, $\{(r_n, s_n)\}_n$
- Finding the $\boldsymbol{\psi}$ that minimizes $\sum_n \|\mathbf{r}_{r_n} - \mathcal{T}(\mathbf{s}_{s_n}; \boldsymbol{\psi})\|_2^2$

Converges towards a local minimum

- Requires good initial estimate of $\boldsymbol{\psi}$

`https://www.youtube.com/watch?v=ii2vHBwlmo8`

Truncated signed distance functions (TSDFs)

- Similar to distance transforms in 3D (0 = surface)
- But distances are signed, measured along view rays

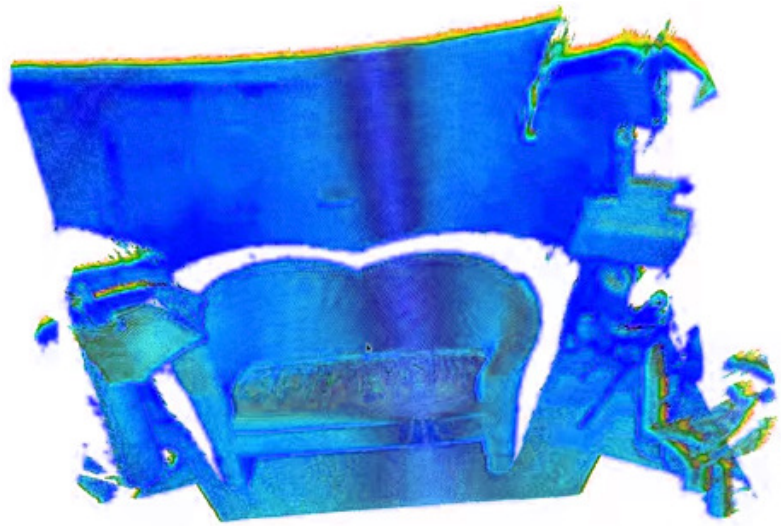
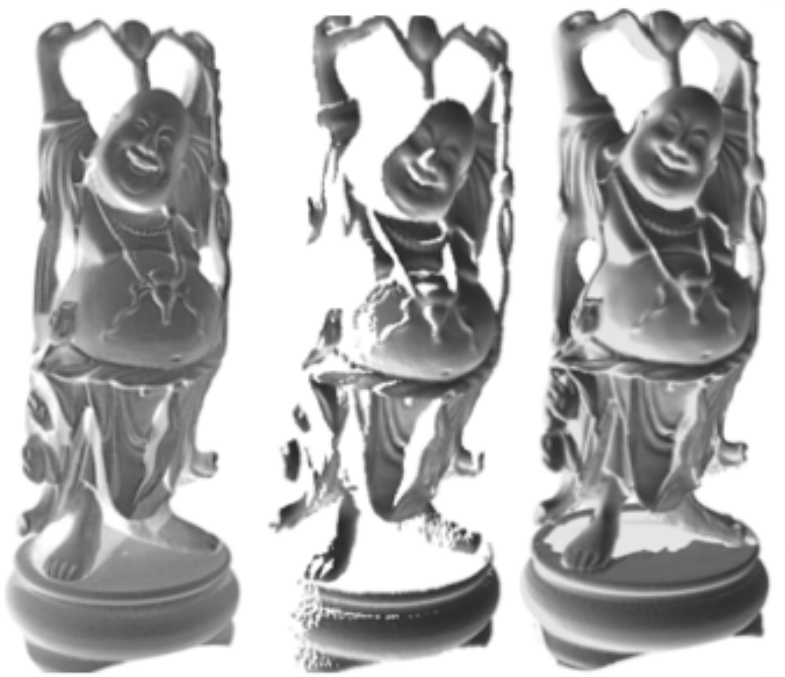


Image from https://www.youtube.com/watch?v=AjjSZufyprU

# 3D Reconstruction
## Range Data Merging – TSDF Fusion

Merged data = weighted average over aligned TSDF voxels

- Weights based on e.g. object distance, angle

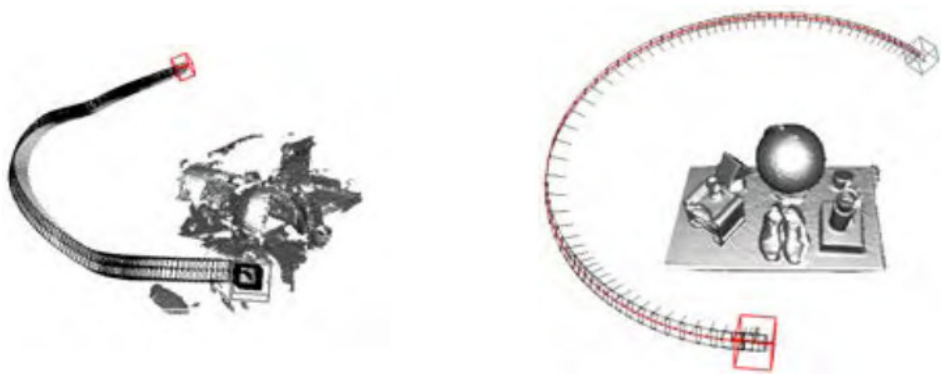


Images from Curless and Levoy 1996

Temporal fusion of Kinect depth maps

Based on the above methods (ICP & TSDF fusion)

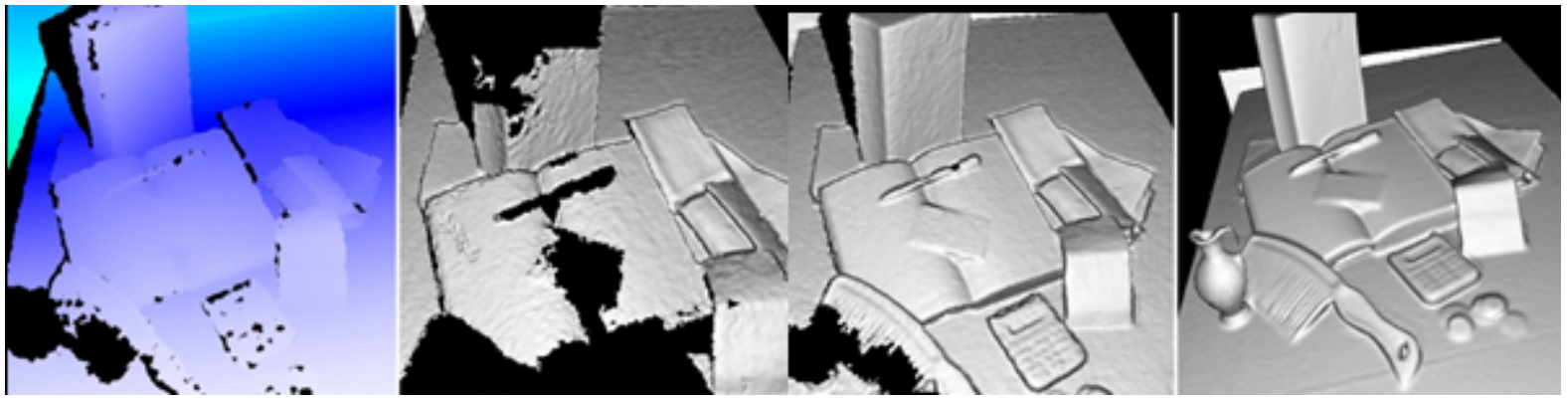
- But $\{\mathbf{r}\}$ synthesized from merged model
- Suppresses alignment error accumulation



Images from Newcombe et al. 2011

# 3D Reconstruction
## Kinect Fusion

https://www.youtube.com/watch?v=quGhaggn3cQ



Images from microsoft.com

# 3D Reconstruction
## Application Fields

Cultural heritage

Virtual and augmented reality

Architecture

...

# 3D Reconstruction
## Application Fields

[Video]

# Object Detection

3D data enables reliable object detection

- Metric information available
- Invariant to camera parameters

# Object Detection

## Breaking Assistance via Person Detection

https://www.youtube.com/watch?v=oU4XQvxO10k

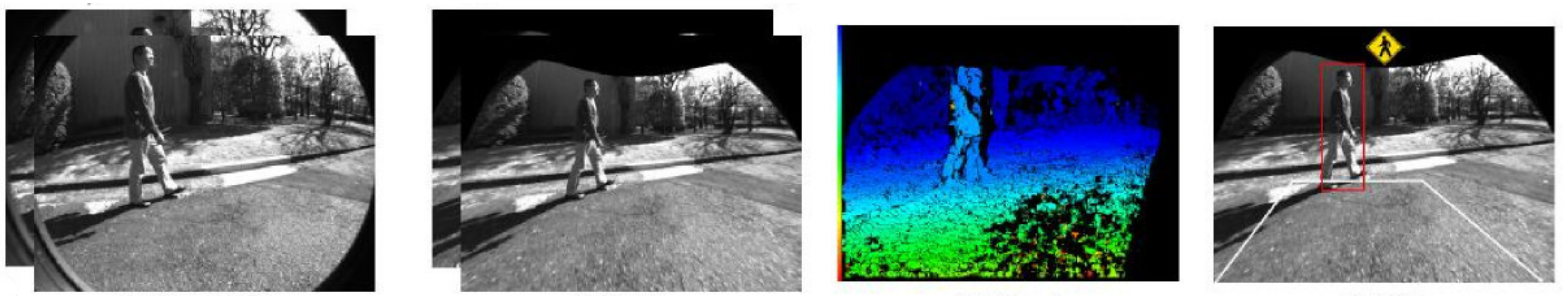


Image from Ryuzo Okada, Toyota

# Object Detection

## Interactive Art Installations


Image from ortlos.com

# Object Detection
## Fall Detection (fearless)

Fall detection system developed at CVL

- Uses data from a single Kinect sensor
- Runs on an inexpensive single-board computer


Image from `fearless-project.eu`

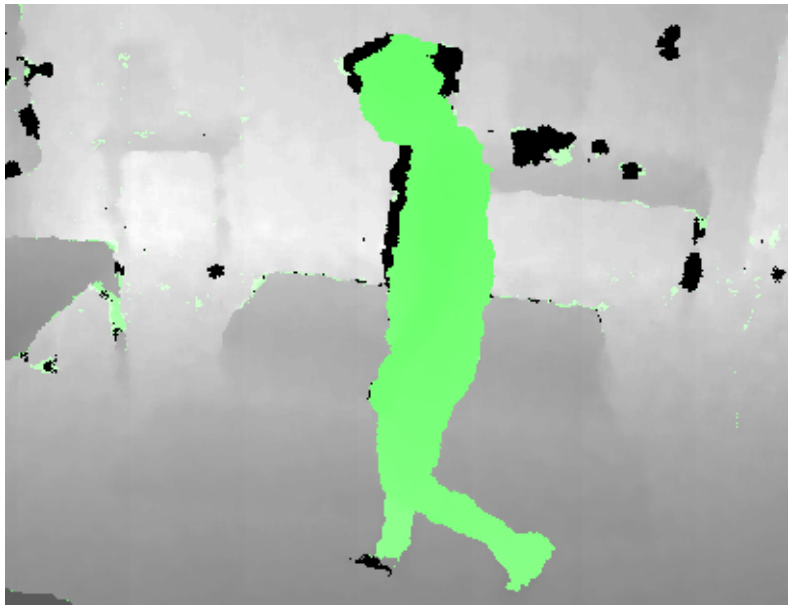
# Object Detection
## Fall Detection (fearless)

[Video]

# Object Detection
## Fall Detection (fearless) – Motion Detection

Motion detection via background subtraction

- Robust because not affected by illumination, clothing

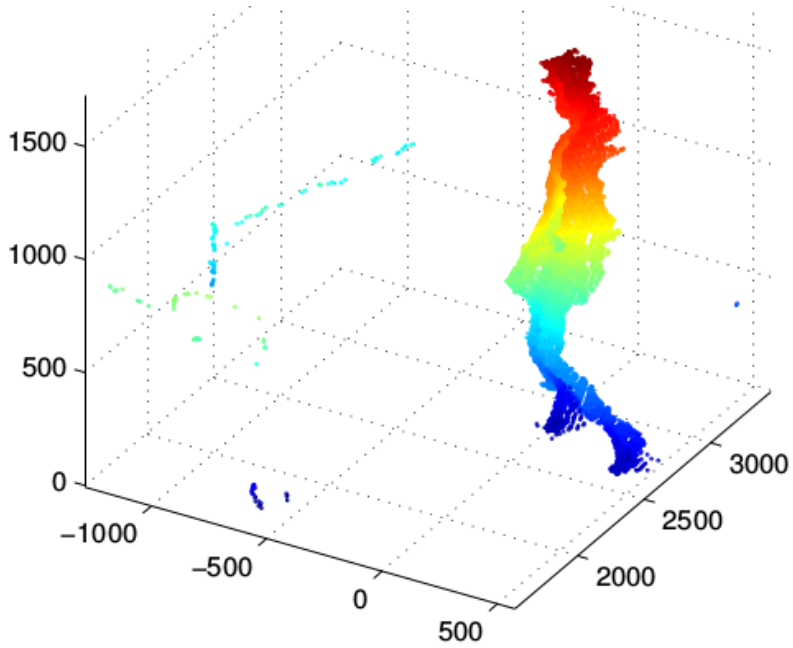


# Object Detection

## Fall Detection (fearless) – Person Detection

Moving areas projected to point cloud in world coordinates

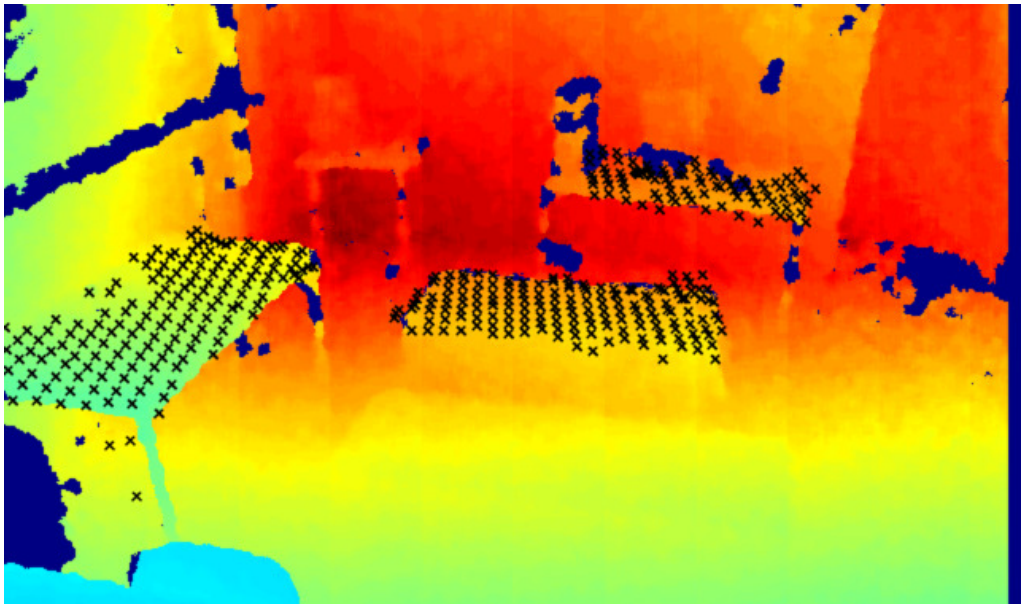
Person classification based on geometry



# Object Detection
## Fall Detection (fearless) – Person Detection

Fall detection based on temporal height change

Furniture detection to reduce false alarm rate



# To Be Continued

More next time!

# Bibliography


Aiger, Dror, Niloy J Mitra, and Daniel Cohen-Or (2008). **4-points congruent sets for robust pairwise surface registration**. ACM TOG.

Curless, Brian and Marc Levoy (1996). **A volumetric method for building complex models from range images**. Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. ACM, pp. 303–312.

Newcombe, Richard A et al. (2011). **KinectFusion: Real-time dense surface mapping and tracking**. ISMAR.

Shotton, Jamie et al. (2011). **Real-Time Human Pose Recognition in Parts from a Single Depth Image**. CVPR.