dyson Robotics Lab Imperial College London

The Future of Real-Time SLAM

# ElasticFusion: real-time dense SLAM without a pose graph

**Stefan Leutenegger (representing Tom Whelan)**

18th December 2015 (ICCV Workshop)

# State of the art in real-time dense SLAM

## Plenty have limitations

- No (online) loop closure
- Only estimates trajectory
  - » Raw point cloud back projections
  - » Key frames
- Non-scalable
- Non-robust pose estimation

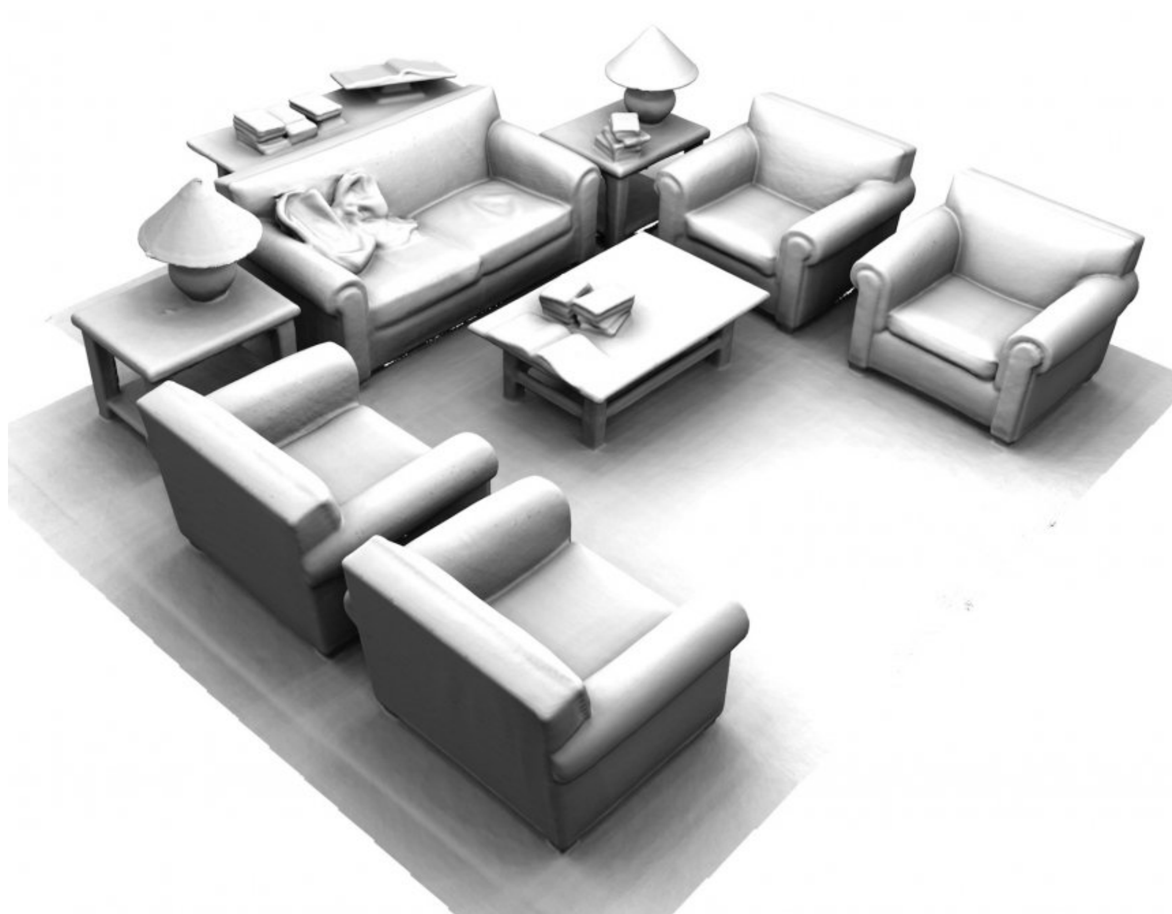
## Some of these include;

- Henry et al., Endres et al., Meilland & Comport, Kerl et al., Keller et al., Chen et al, Nießner et al., Newcombe et al., Steinbruecker et al., Stueckler et al.

# State of the art

## Offline approaches

- Qian-Yi Zhou
  Amazing results with RGB-D, strictly offline (>1hr processing time)



# State of the art

## Online approaches

- Kintinuous arguably the state of the art
  - » Online loop closure
  - » Estimates full 3D surface and trajectory
  - » Scalable (100's of metres)
  - » Full RGB and Depth pose estimation
- However, still not the "perfect" system
  - » A number of limitations

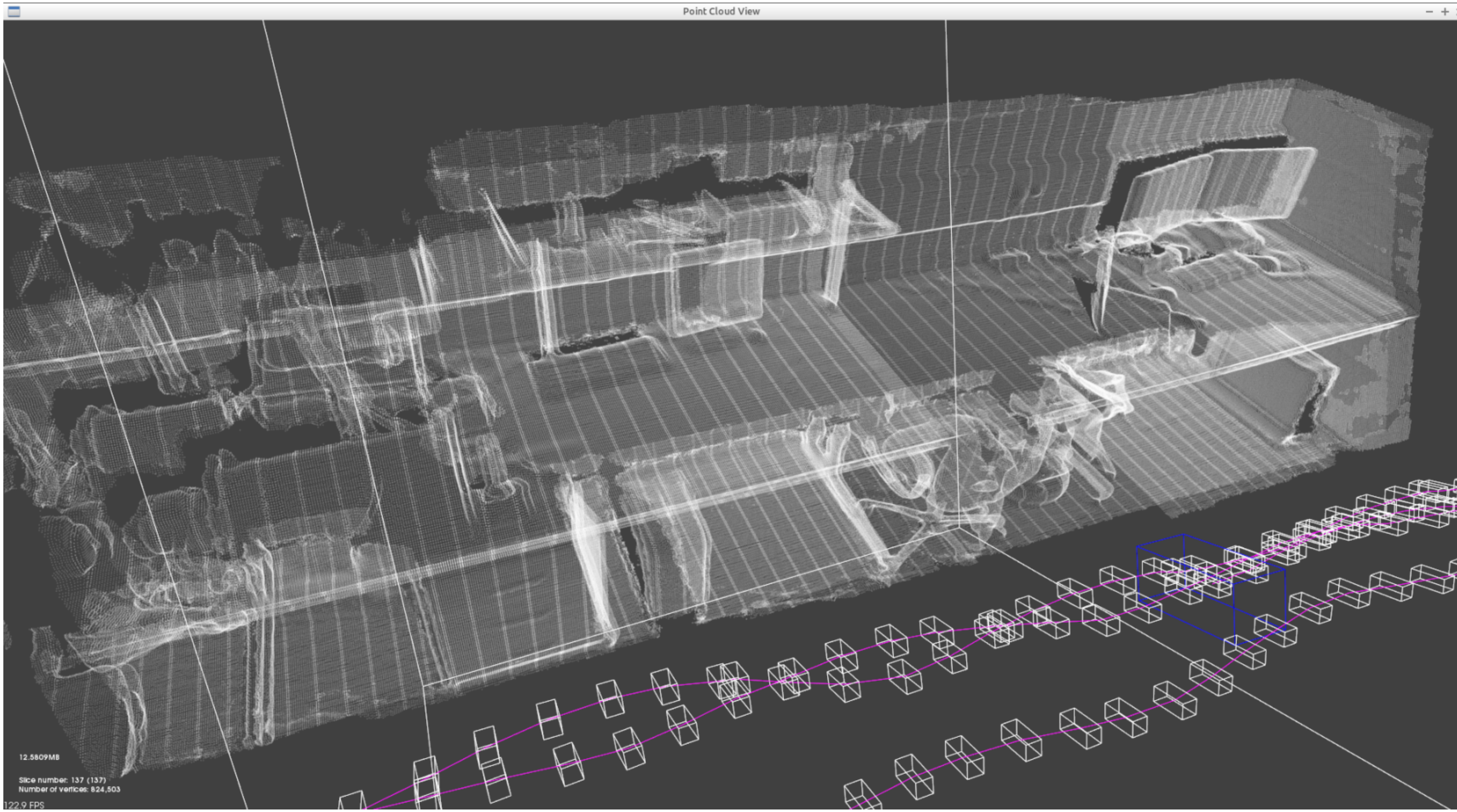
# State of the art

## Kintinuous

- Surface aliasing

# State of the art

## Kintinuous



# We want a SLAM system that...

Robustly estimates camera pose

- Geometry + photometry

Reliably estimates the surface

- Fused representation to remove noise

Scales well

- Room, house scale

Is completely closed loop (updating)

- Update revisited areas

Real-time

- Globally consistent map available at any point in time

Non-restrictive of motion

- Happy to deal with extremely loopy motion and many such loop closures

# Introducing ElasticFusion

Robustly estimates camera pose

- Full RGB and Depth frame-to-model tracking

Reliably estimates the surface

- Point-based fusion is good quality and a nice representation

Scales well

- Room scale seems doable, at a minimum

Is completely closed loop (updating)

- No separation between front end and back end

Real-time

- Strictly

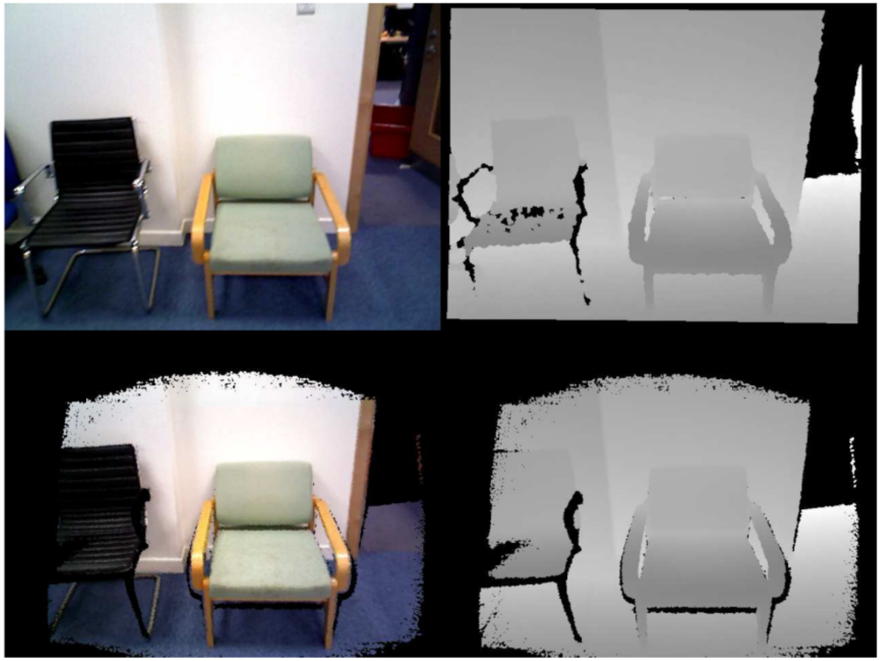
Non-restrictive of motion

- Since the front end and back end are one and the same, it is less restrictive given the full frame-to-model tracking

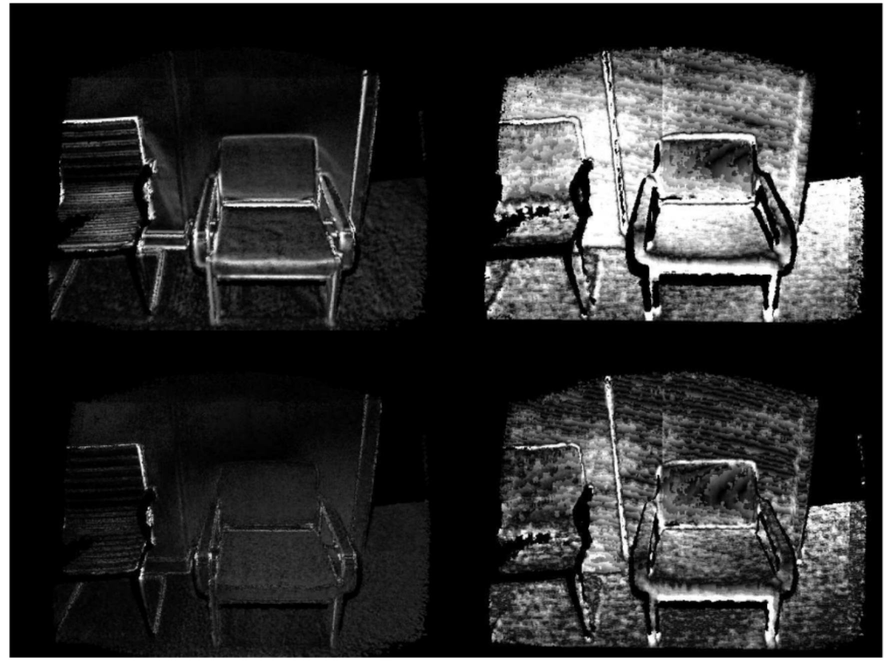
# How it works

1. Reconstruct surfel-based map of environment
2. Split into active/inactive regions
3. Directly register multiple passes of the same surface together
4. Reflect this in the map with a non-rigid space deformation
5. Use fern encoded key frames for global loop closures

# How it works: Tracking



Data terms



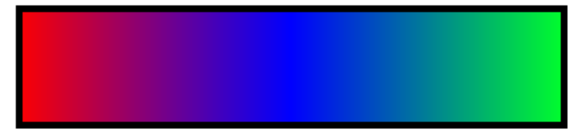
Residuals

$$E_{rgb} = \sum_{\mathbf{u}\in\Omega} \left( I(\mathbf{u}, \mathcal{C}_t^l) - I\left(\boldsymbol{\pi}(\mathbf{K}\exp(\hat{\boldsymbol{\xi}})\mathbf{T}\mathbf{p}(\mathbf{u}, \mathcal{D}_t^l)), \hat{\mathcal{C}}_{t-1}^a\right)\right)^2$$

$$E_{icp} = \sum_{k} \left( \left( \mathbf{v}^k - \exp(\hat{\boldsymbol{\xi}})\mathbf{T}\mathbf{v}_t^k \right) \cdot \mathbf{n}^k \right)^2$$

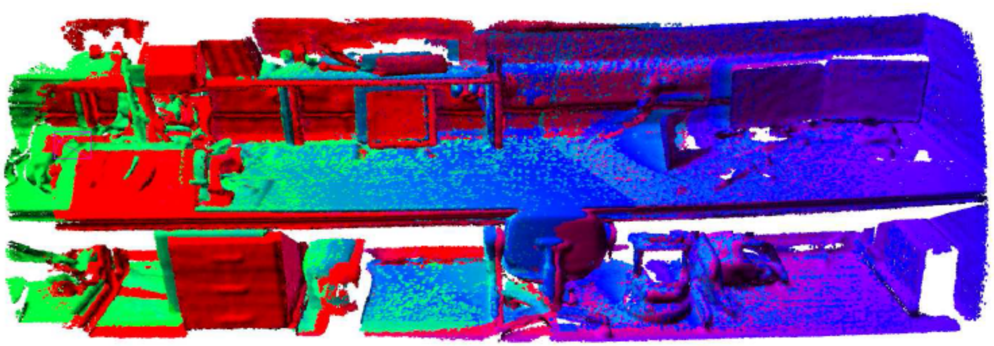
# How it works: Building a Deformation Graph

Mapping left to right

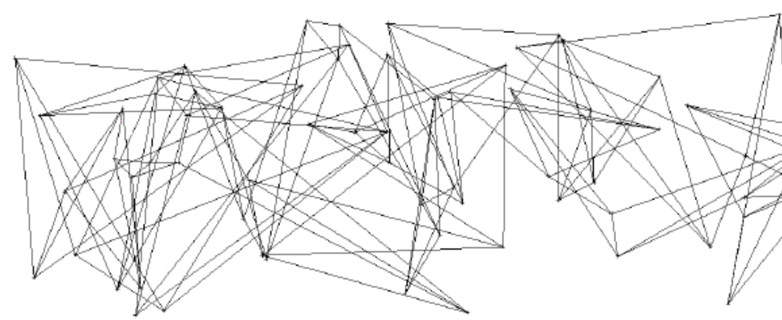


0 t

Time scale

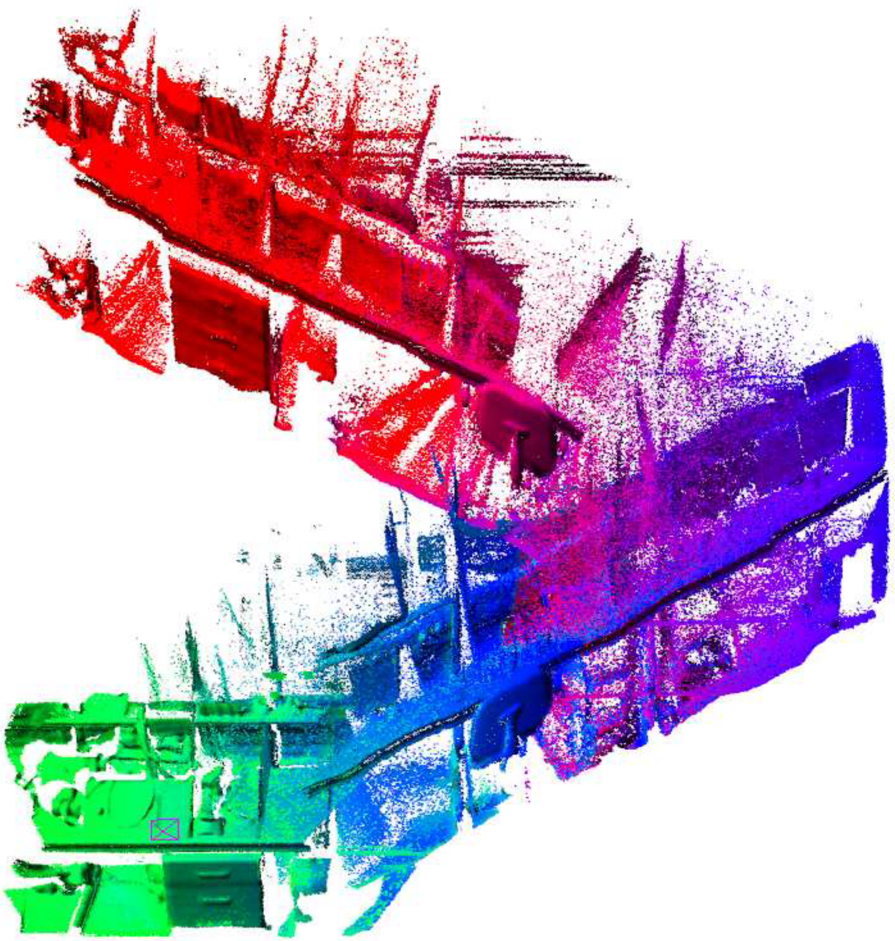


Mapping right to left

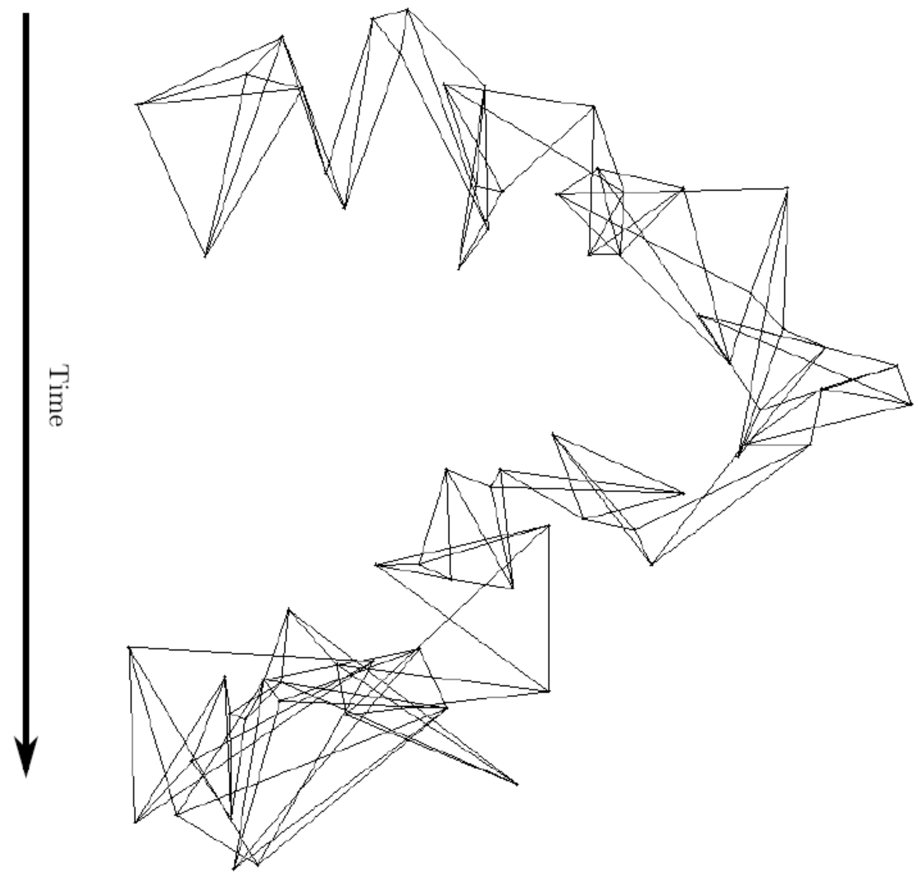


Deformation graph

# How it works: Time Stretched Visualisation



Time stretched map




Time stretched graph

# How it works: Loop Closure

$$E_{rot} = \sum_{l} \left\| {\mathcal{G}^{l}_{\mathbf{R}}}^{\top} \mathcal{G}^{l}_{\mathbf{R}} - \mathbf{I} \right\|_F^2$$ As-rigid-as-possible

$$E_{reg} = \sum_{l} \sum_{n \in \mathcal{N}(\mathcal{G}^l)} \left\| \mathcal{G}^{l}_{\mathbf{R}}(\mathcal{G}^{n}_{\mathbf{g}} - \mathcal{G}^{l}_{\mathbf{g}}) + \mathcal{G}^{l}_{\mathbf{g}} + \mathcal{G}^{l}_{\mathbf{t}} - (\mathcal{G}^{n}_{\mathbf{g}} + \mathcal{G}^{n}_{\mathbf{t}}) \right\|_2^2$$ Smoothness regulariser

$$E_{con} = \sum_{p} \| \phi(\mathcal{Q}^{p}_{\mathbf{s}}) - \mathcal{Q}^{p}_{\mathbf{d}} \|_2^2$$ New-to-old loop closure constraints

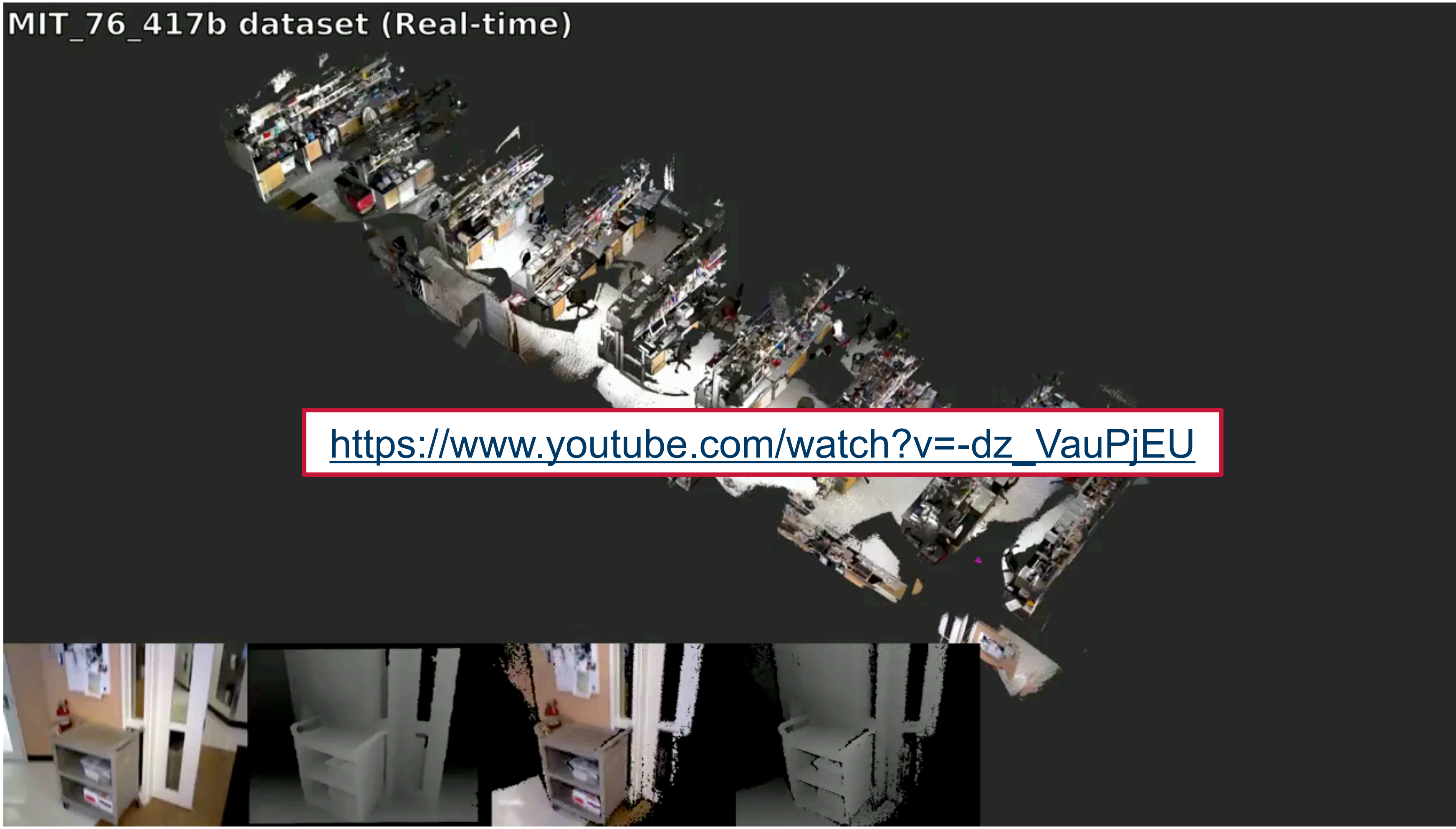
$$E_{pin} = \sum_{p} \| \phi(\mathcal{Q}^{p}_{\mathbf{d}}) - \mathcal{Q}^{p}_{\mathbf{d}} \|_2^2$$ Old-to-old anchoring constraints

$$E_{rel} = \sum_{p} \| \phi(\mathcal{R}^{p}_{\mathbf{s}}) - \phi(\mathcal{R}^{p}_{\mathbf{d}}) \|_2^2$$ Relative constraints (previous loop closures)

# ElasticFusion


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

# ElasticFusion – Extras




https://www.youtube.com/watch?v=-dz_VauPjEU

# Quantitative Results

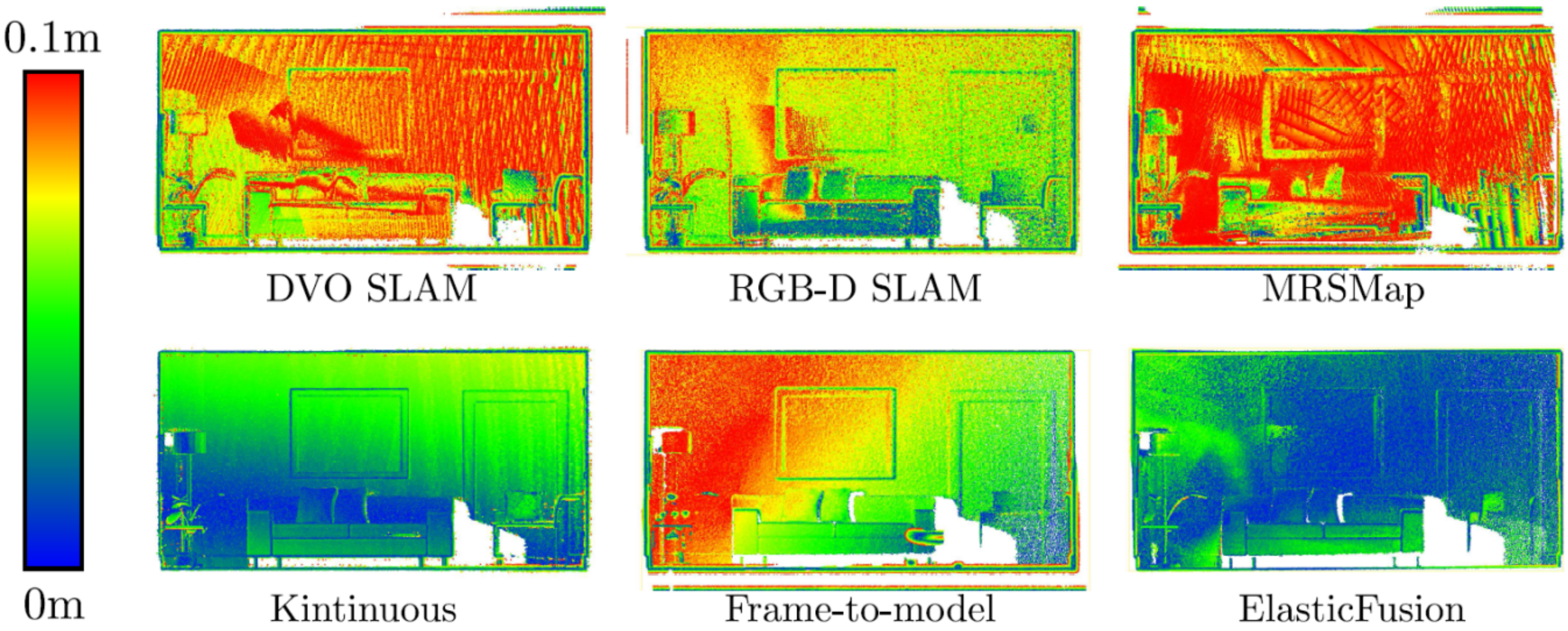
| System | fr1/desk | fr2/xyz | fr3/office | fr3/nst |
|---|---|---|---|---|
| DVO SLAM | 0.021m | 0.018m | 0.035m | 0.018m |
| RGB-D SLAM | 0.023m | **0.008m** | 0.032m | 0.017m |
| MRSMap | 0.043m | 0.020m | 0.042m | 2.018m |
| Kintinuous | 0.037m | 0.029m | 0.030m | 0.031m |
| Frame-to-model | 0.022m | 0.014m | 0.025m | 0.027m |
| ElasticFusion | **0.020m** | 0.011m | **0.017m** | **0.016m** |

TUM RGB-D Trajectory Error

| System | kt0 | kt1 | kt2 | kt3 |
|---|---|---|---|---|
| DVO SLAM | 0.032m | 0.061m | 0.119m | 0.053m |
| RGB-D SLAM | 0.044m | 0.032m | 0.031m | 0.167m |
| MRSMap | 0.061m | 0.140m | 0.098m | 0.248m |
| Kintinuous | 0.011m | 0.008m | 0.009m | 0.150m |
| Frame-to-model | 0.098m | **0.007m** | 0.011m | 0.107m |
| ElasticFusion | **0.007m** | **0.007m** | **0.008m** | **0.028m** |

ICL-NUIM Surface Error

# Quantitative Results



# Main Advances

Real-time deformation

- Great for overcoming the drift problem in a dense map

Fully closed loop

- No frontend/backend division opens up many possibilities

Open source

- https://github.com/mp3guy/ElasticFusion

# Light Source Estimation

Reflectance-driven

- Detect speculars

Why?

- It's cool
- Convincing AR effects
- Can be used to improve tracking
- Aids in path planning (i.e. avoid bright areas)

# Light Source Estimation

Reconstruct diffuse appearance

Bright raw observations are reflected rays

- Vote in voxel space (hough-like scheme)

Intersections of high votes and geometry are potentially light sources

- Geometry helps determine extent of light source, directionality and removes need for comprehensive reflected ray coverage

# Light Source Estimation