ROBOTICS & PERCEPTION GROUP

Universität Zürich[UZH]

Lecture 10
3D Reconstruction

Prof. Dr. Davide Scaramuzza

sdavide@ifi.uzh.ch

3D Reconstruction from Multiple views

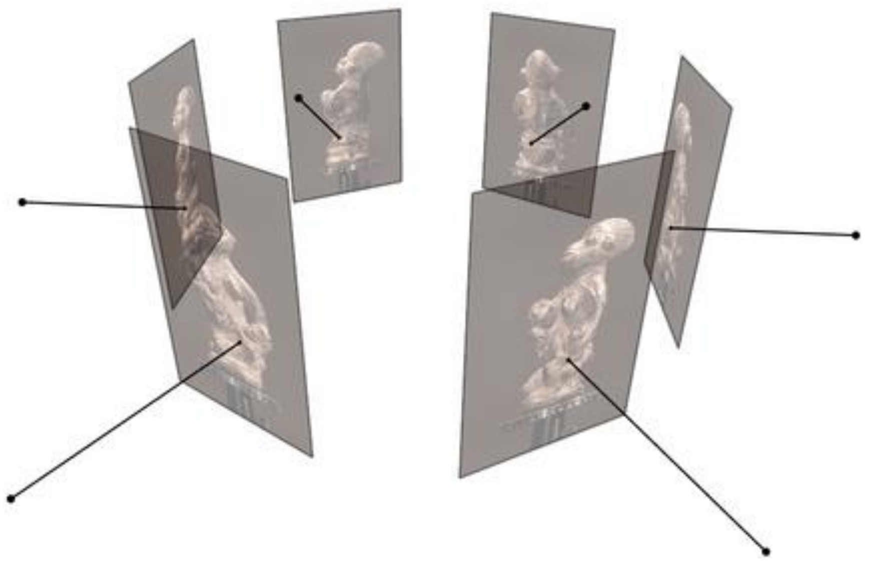
Assumption

- Cameras are calibrated
  - both **intrinsically**
    - $\boldsymbol{K}$ matrix for each camera is known
  - and **extrinsically**
    - relative positions $\boldsymbol{T}$ and orientations $\boldsymbol{R}$ between cameras are known, for instance through Structure from Motion

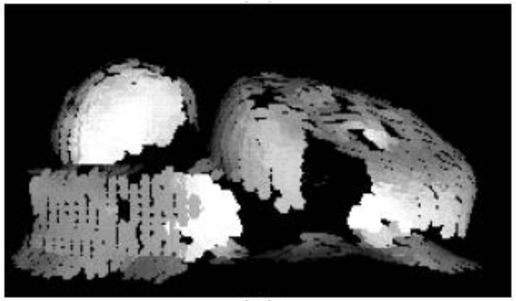
Multi-view stereo

Input: calibrated images from several viewpoints

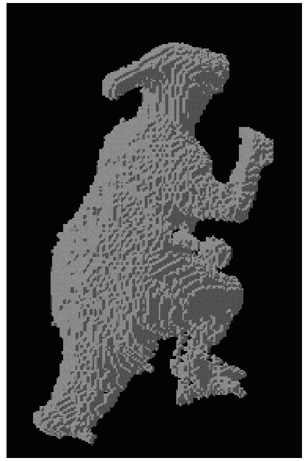
Output: 3D object model



Figures by Carlos Hernandez



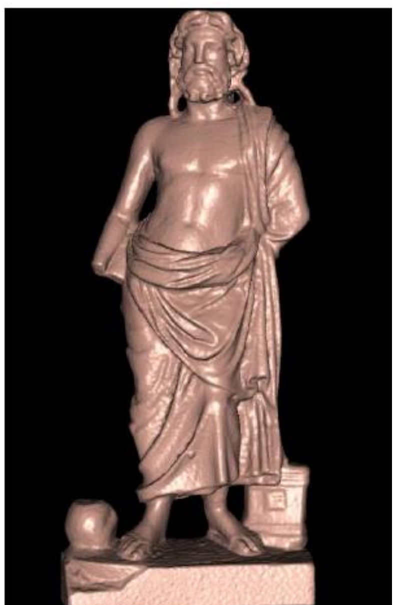
Fua
1995



Seitz, Dyer
1997

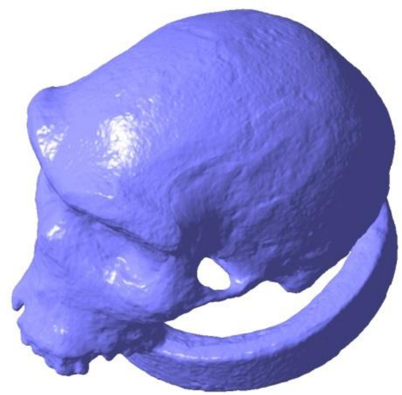
Narayanan, Rander, Kanade
1998

Faugeras, Keriven
1998

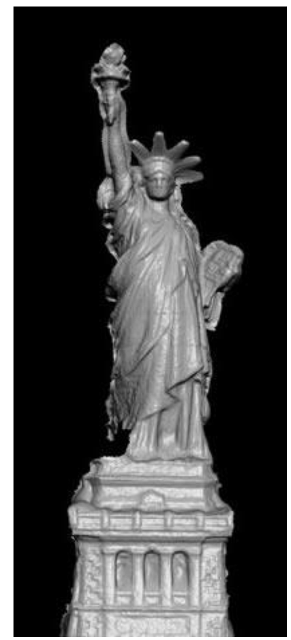


Hernandez, Schmitt
2004

Pons, Keriven, Faugeras
2005



Furukawa, Ponce
2006

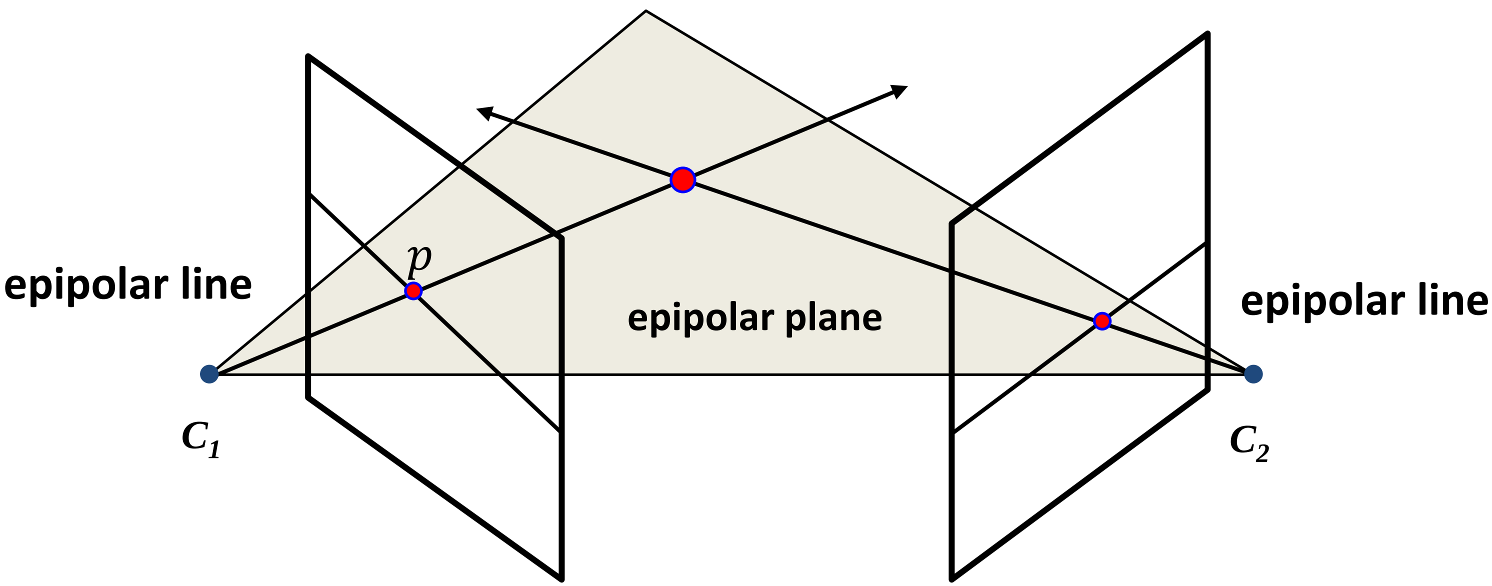


Goesele et al.
2007

Kolev, Brox, Cremers
2012

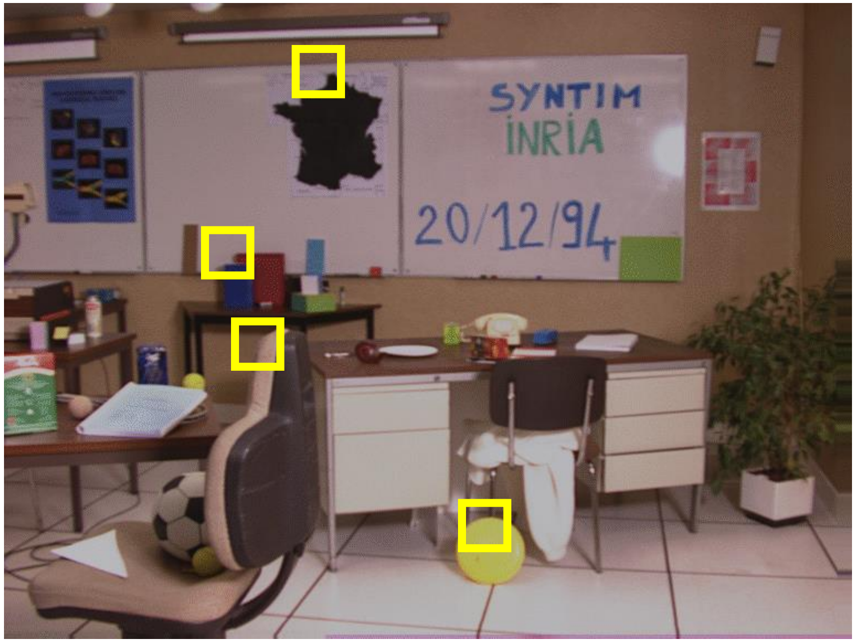
Review: The Epipolar Plane

The two camera centers and the feature p determine a plane called the "epipolar plane", which intersect each camera image plane into an epipolar line.

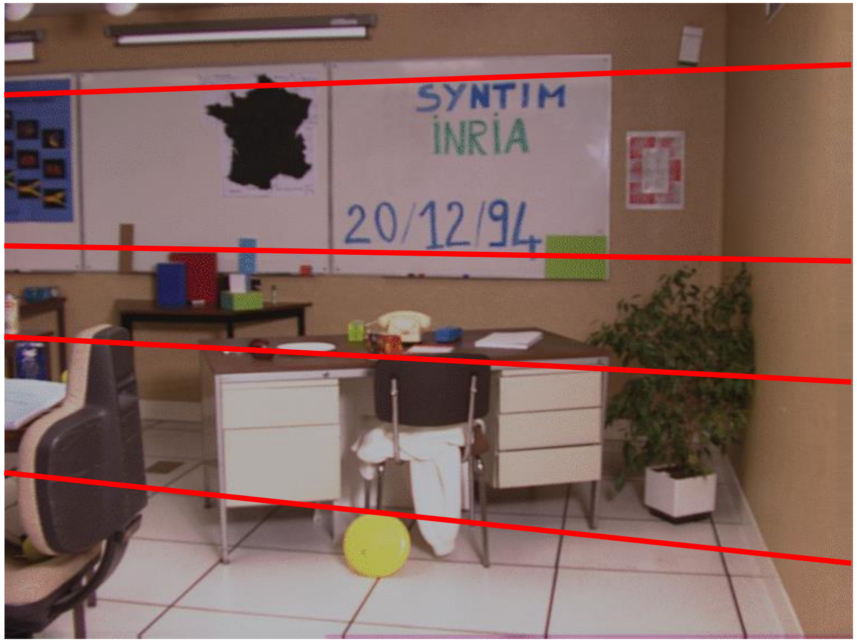


Review: Epipolar Lines for Correspondence Search

Thanks to the epipolar constraint, corresponding points only need to be searched along epipolar lines



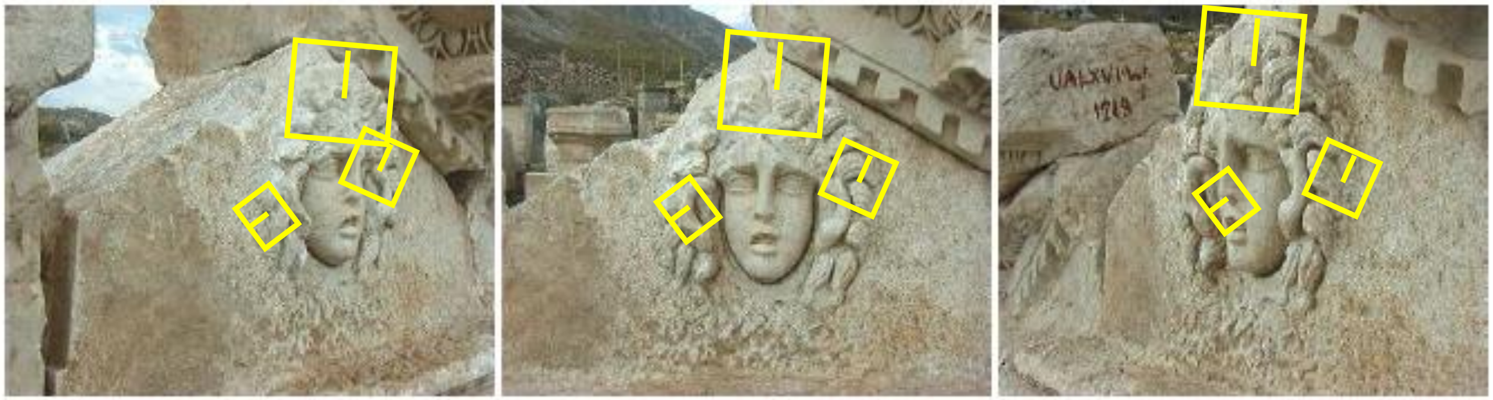
Left image

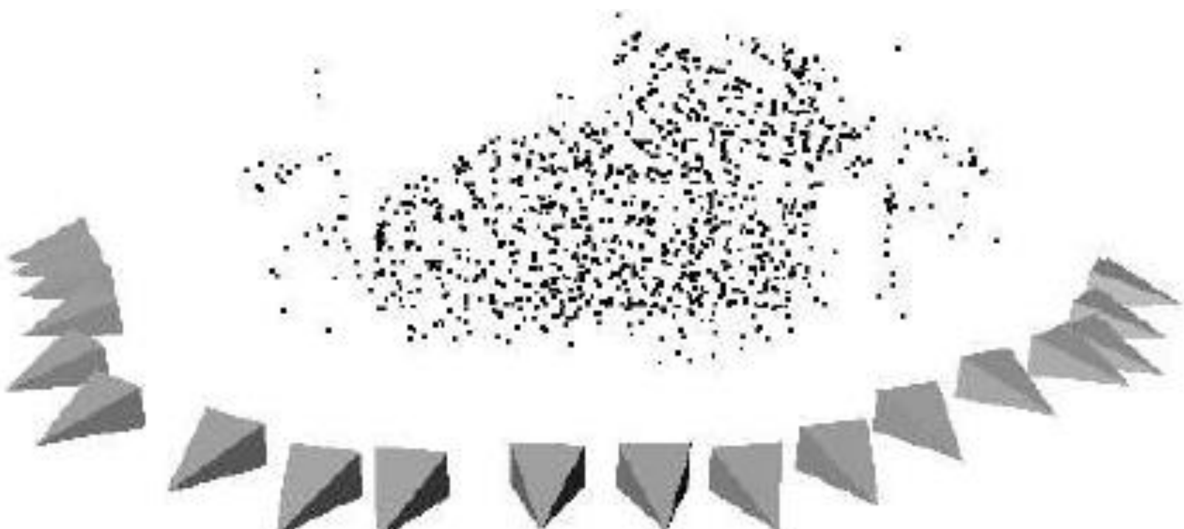


Right image

Sparse Reconstruction

- Estimate the structure from a “sparse” set of features





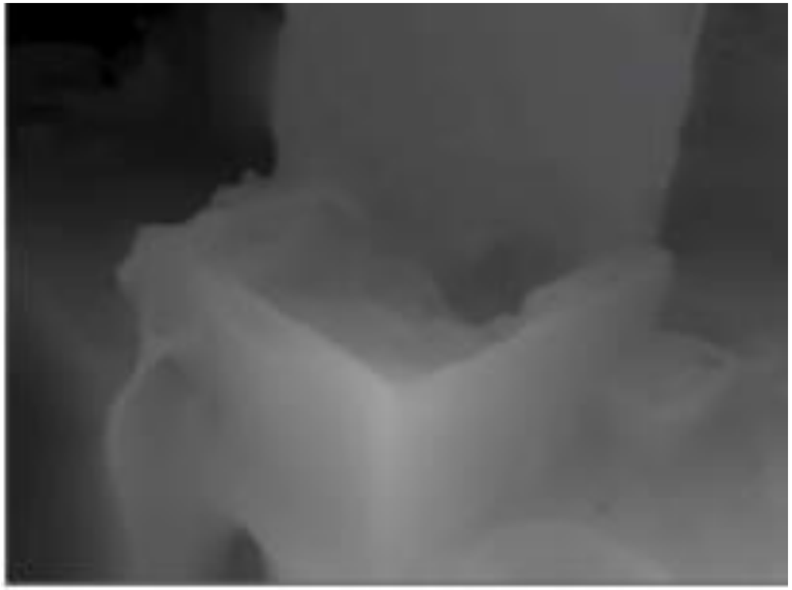
Dense Reconstruction

- Estimate the structure from a “dense” region of pixels

Dense reconstruction

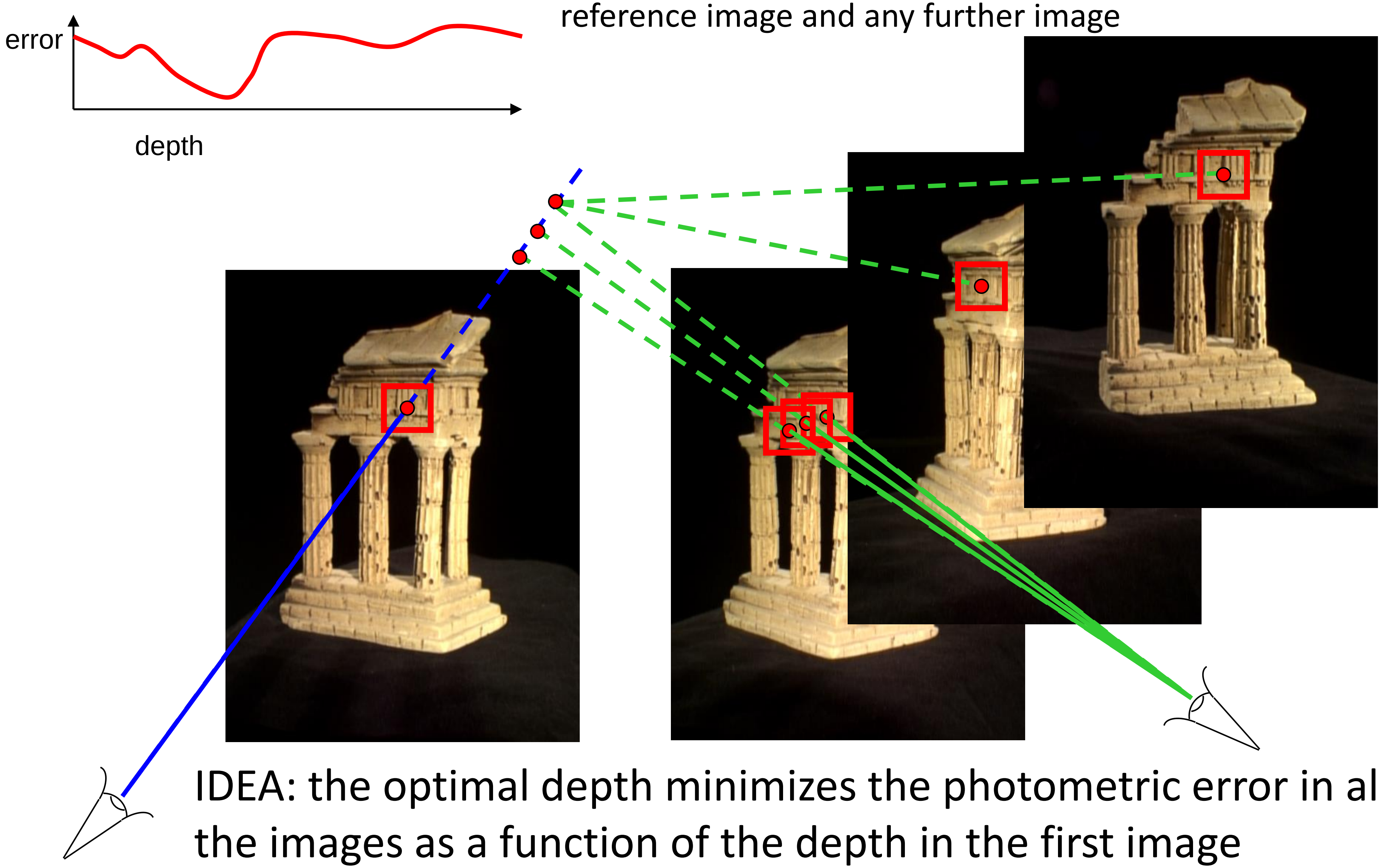
- **Local methods**
  - Estimate depth for every pixel independently

- **Global methods**
  - Refine the depth surface as a whole by enforcing smoothness constraint



Photometric error (e.g., SSD, SAD, ZNCC)

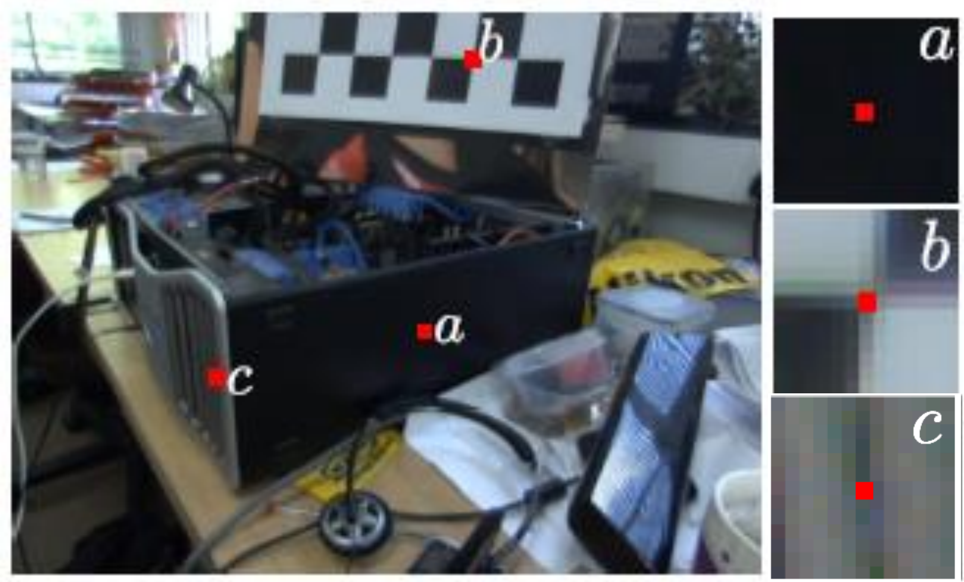
This error plot is derived for every combination of the reference image and any further image



IDEA: the optimal depth minimizes the photometric error in all the images as a function of the depth in the first image

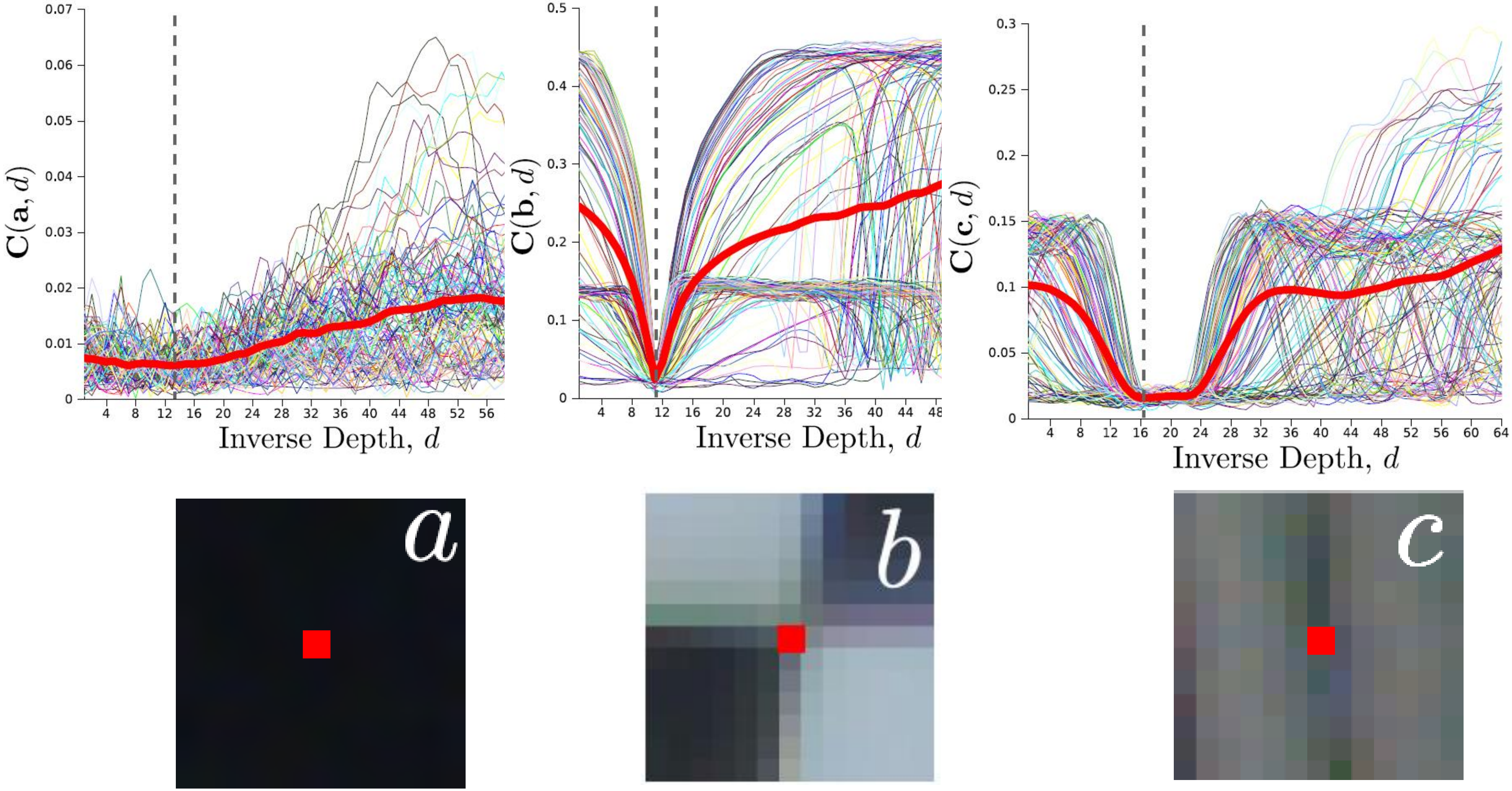
Aggregated photometric error

- Dense reconstruction requires establishing dense correspondences
- Correspondences are computed based on photometric error
  - Difference among pixel intensity values or
  - patch-based correlation
    - SAD, SSD, NCC
- Not all the pixels can be matched reliably
  - Viewpoint and illumination changes, occlusions
- Take advantage of many small baseline views where high quality matching is possible



[Newcombe et al. 2011]

Aggregated photometric error



- Photometric error shows multiple minima (because of noise, lack of textures or repetitive textures)
- For distinctive pixels (as in ***b***) the aggregated photometric error has one clear minimum.

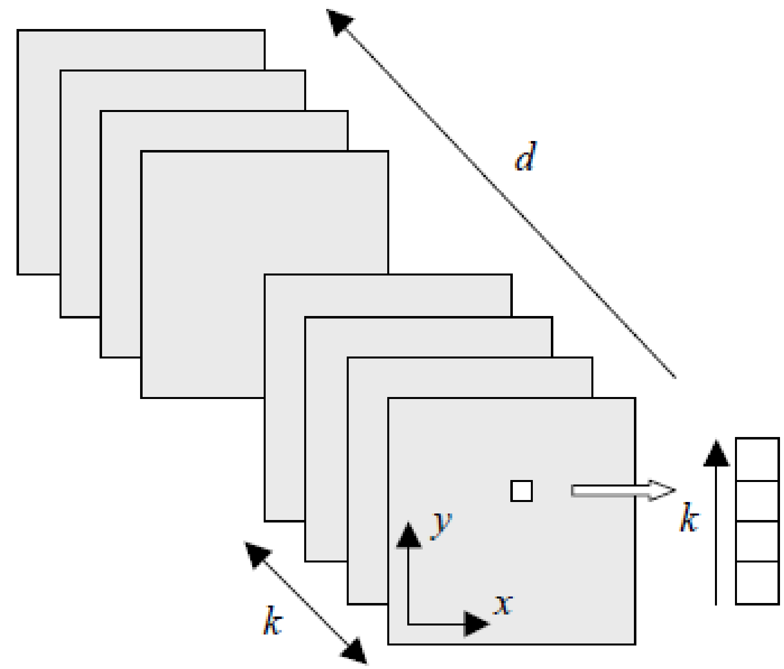
Generalized Disparity Space Image

- For discrete depth hypotheses the aggregate photometric error with respect to the reference image can be stored in the generalized disparity space image

$$C(u,v,d) = \sum_k \rho(\widetilde{I_k}(u',v',d) - I_r(u,v))$$

  $\widetilde{I_k}(u',v',d,k)$ is the pixel in the k-th image associated with the pixel (u',v') in the reference image I_r and depth hypothesis d

- $\rho(\cdot)$ is the photometric error (e.g., SSD, SAD)



[Szeliski and Golland 1999]

Solution

The solution to the depth estimation problem *is a function* in disparity space that best describes the shape of the surface in scene:

- *find a* ***surface embedded in the disparity space image*** *that presents some* ***optimality properties***
  - **Minimum aggregated photometric cost**

$$arg\min_{d} C$$

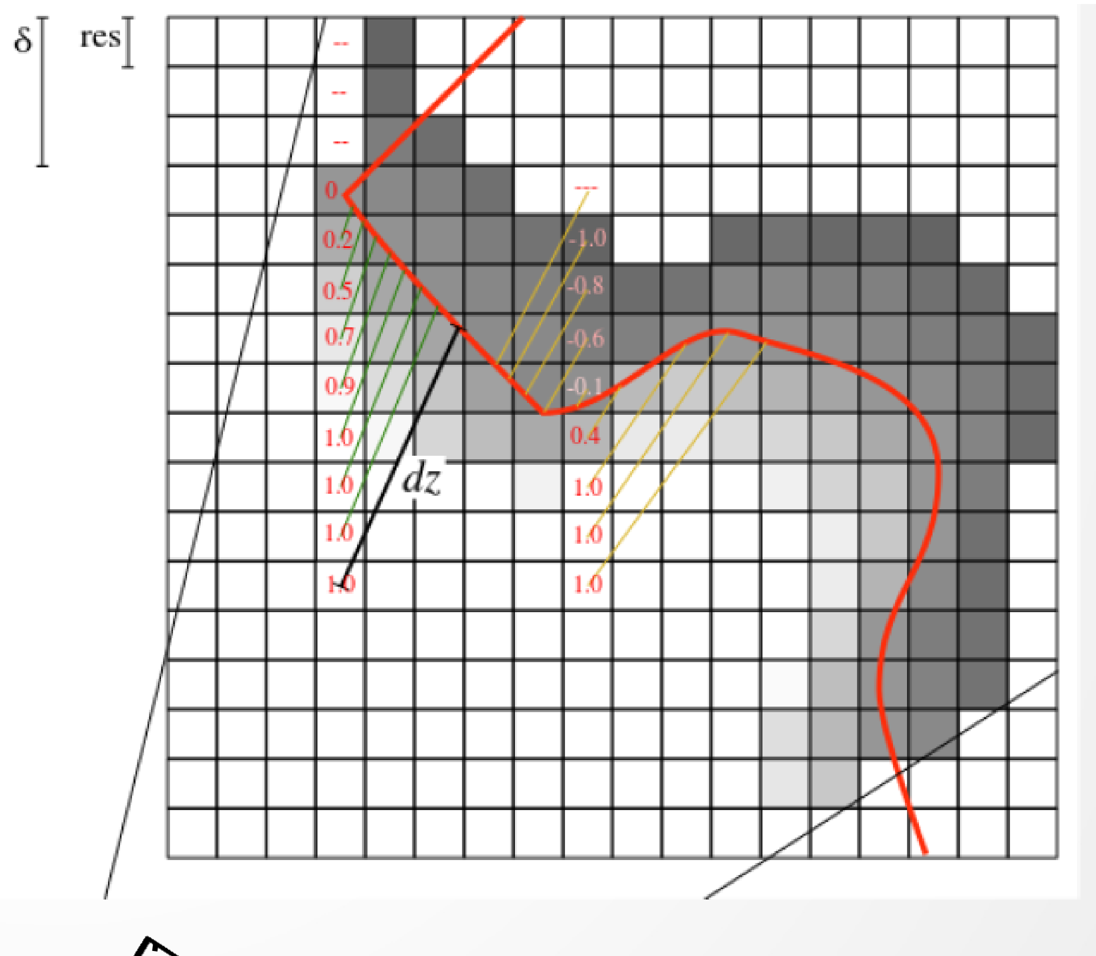
  - *AND (optionally)* **best piecewise smoothness** (global methods)

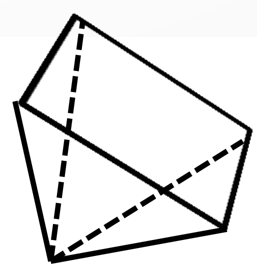
Solution

The solution to the depth estima

that best describes the shape of

- *find a* ***surface embedded in th***
  optimality properties
  - **Minimum aggregated pho**
  - *AND (optionally)* **best piec**





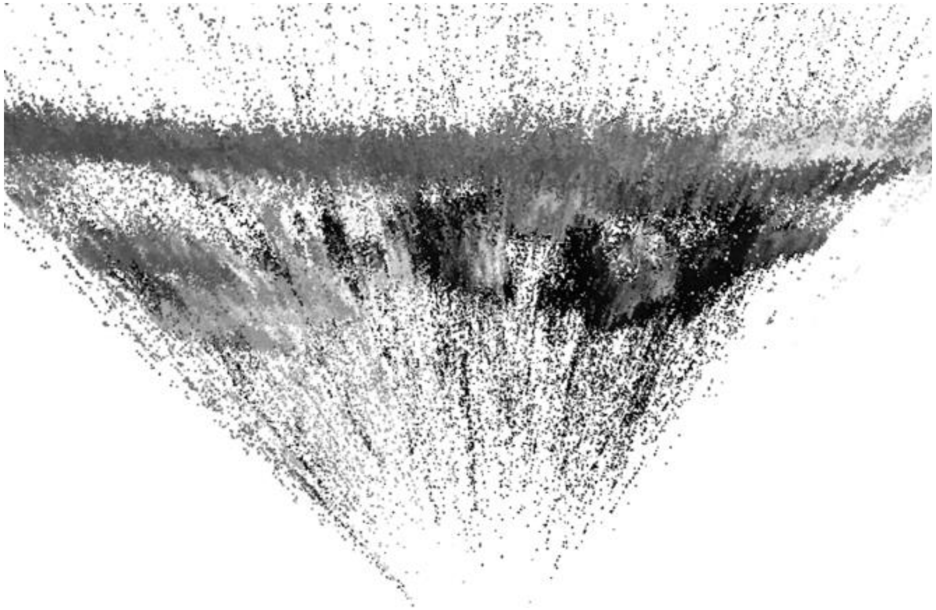
Solution

The solution to the depth estimation problem *is a function* in disparity space that best describes the shape of the surface in scene:

- *find a **surface embedded in the disparity space image** that presents some **optimality properties***
  - **Minimum aggregated photometric cost**

$$arg\min_{d} C$$

  - *AND (optionally)* **best piecewise smoothness** (global methods)



Solution

The solution to the depth estimation problem *is a function* in disparity space that best describes the shape of the surface in scene:

- *find a* ***surface embedded in the disparity space image*** *that presents some* ***optimality properties***
  - **Minimum aggregated photometric cost**

$$arg \min_d C$$

  - *AND (optionally)* **best piecewise smoothness** (global methods)

Solution

- **Global methods**
  - Formulated in terms of energy minimization
  - The objective is to find $d(u, v)$ that minimizes a global energy

$$E(d) = \underbrace{E_d(d)}_{\text{Data term}} + \lambda \underbrace{E_s(d)}_{\text{Regularization term}}$$

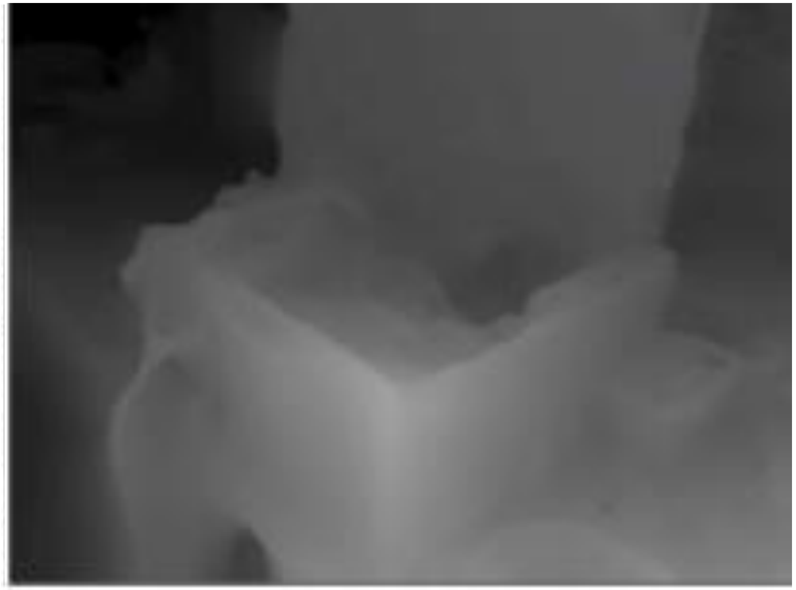
$$E_d(d) = \sum_{(u,v)} C(u, v, d(u, v))$$

$$E_s(d) = \sum_{(u,v)} \rho_d\big(d(u, v) - d(u + 1, v)\big) + \rho_d\big(d(u, v) - d(u, v + 1)\big)$$

- ρ_d is a norm (e.g. L_2, L_1 or Huber norm)
- λ controls the tradeoff data / regularization

Regularized depth maps

- **The regularization term** $E_s(d)$ acts as a ***prior*** in the minimization of the energy functional $E(d)$
  - ***Penalizes*** *non smooth surfaces* (results of noisy measurements)
  - "*Fills the holes*" by means of *a priori* information



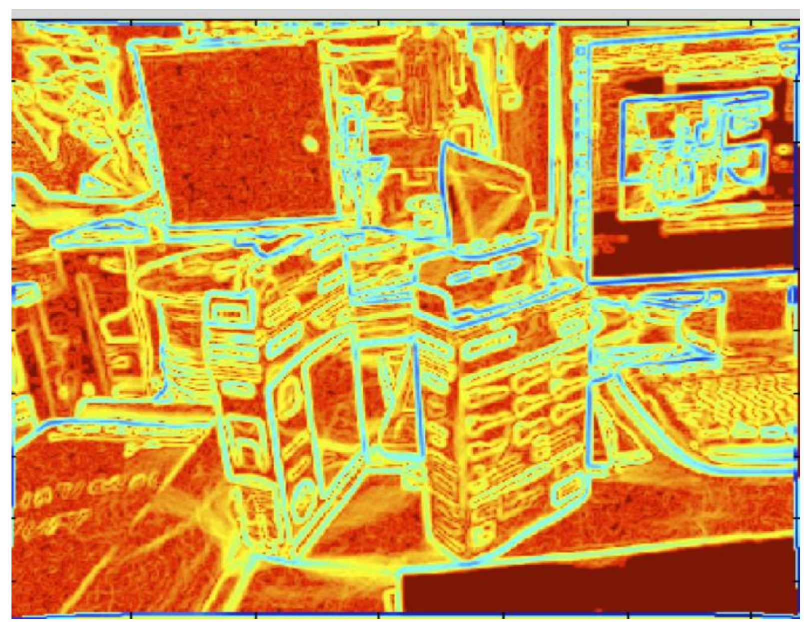
[Newcombe et al. 2011]

Regularized depth maps

- Popular assumption: *discontinuities in intensity* **coincide** with *discontinuities in depth*
- ***Control smoothness penalties*** according to image gradient

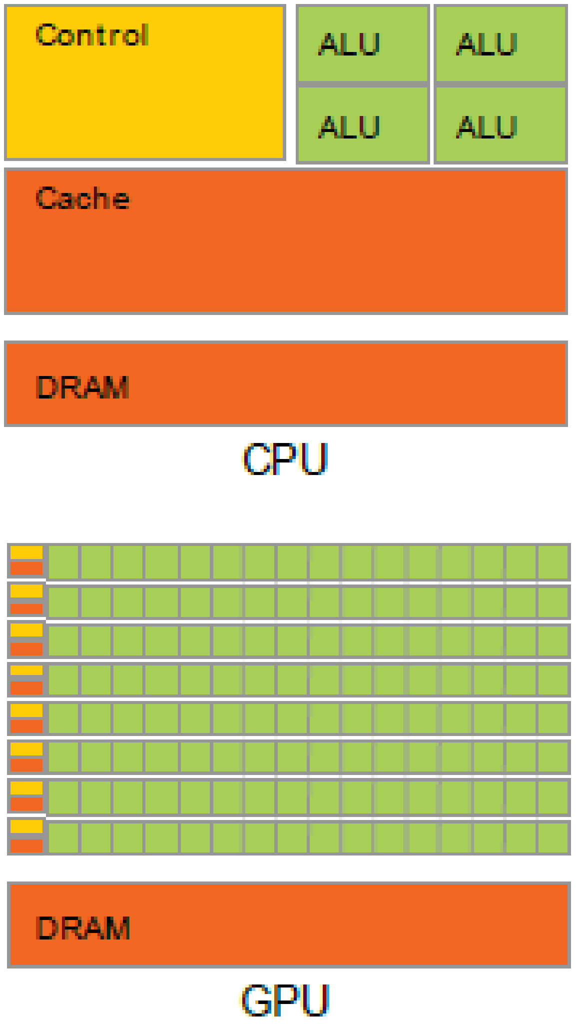
$$\rho_d\big(d(u,v) - d(u+1,v)\big) \cdot \rho_I\,(\|I(u,v) - I(u+1,v)\|)$$

- ρ_I is some monotonically decreasing function of intensity differences: ***lowers*** *smoothness costs at* ***high intensity gradients***



GPGPU

- ***GPGPU = General Purpose computing on Graphics Processing Unit***
- Perform demanding calculations on the GPU instead of the CPU
- On the GPU: high processing power *in parallel*
- More transistors devoted to data processing

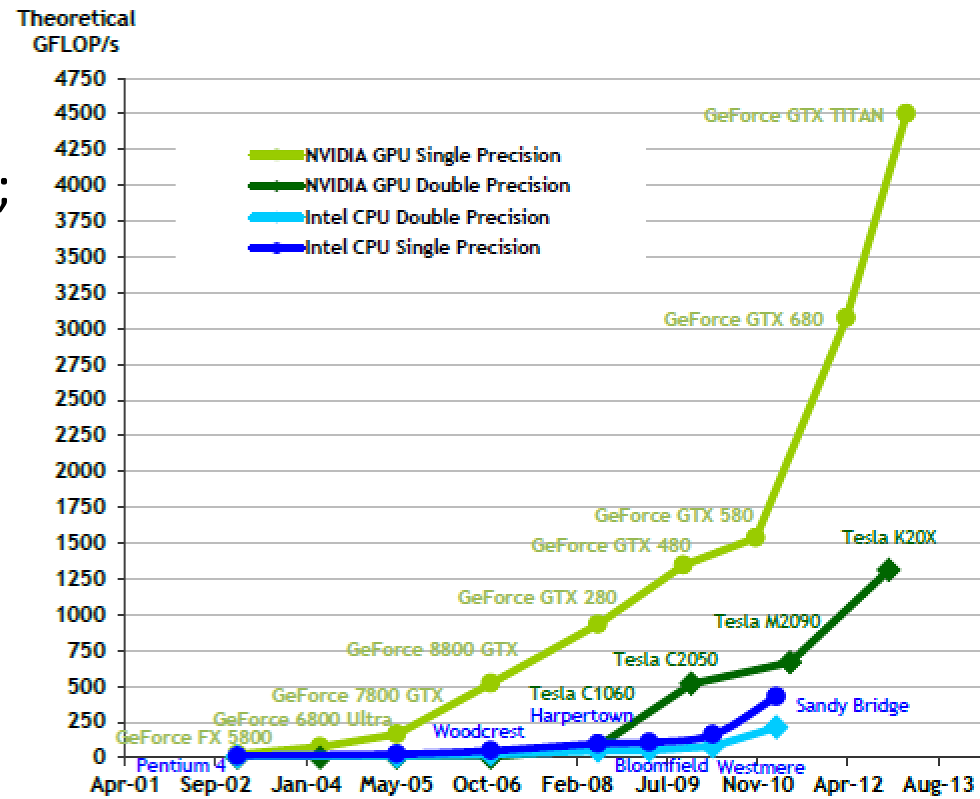


GPU Capabilities

- Fast pixel processing
  - Ray tracing, draw textures, shaded triangles faster than CPU
- Fast matrix / vector operations
  - Transform vertices
- Programmable
  - Shading, bump mapping
- Floating-point support
  - Accurate computations

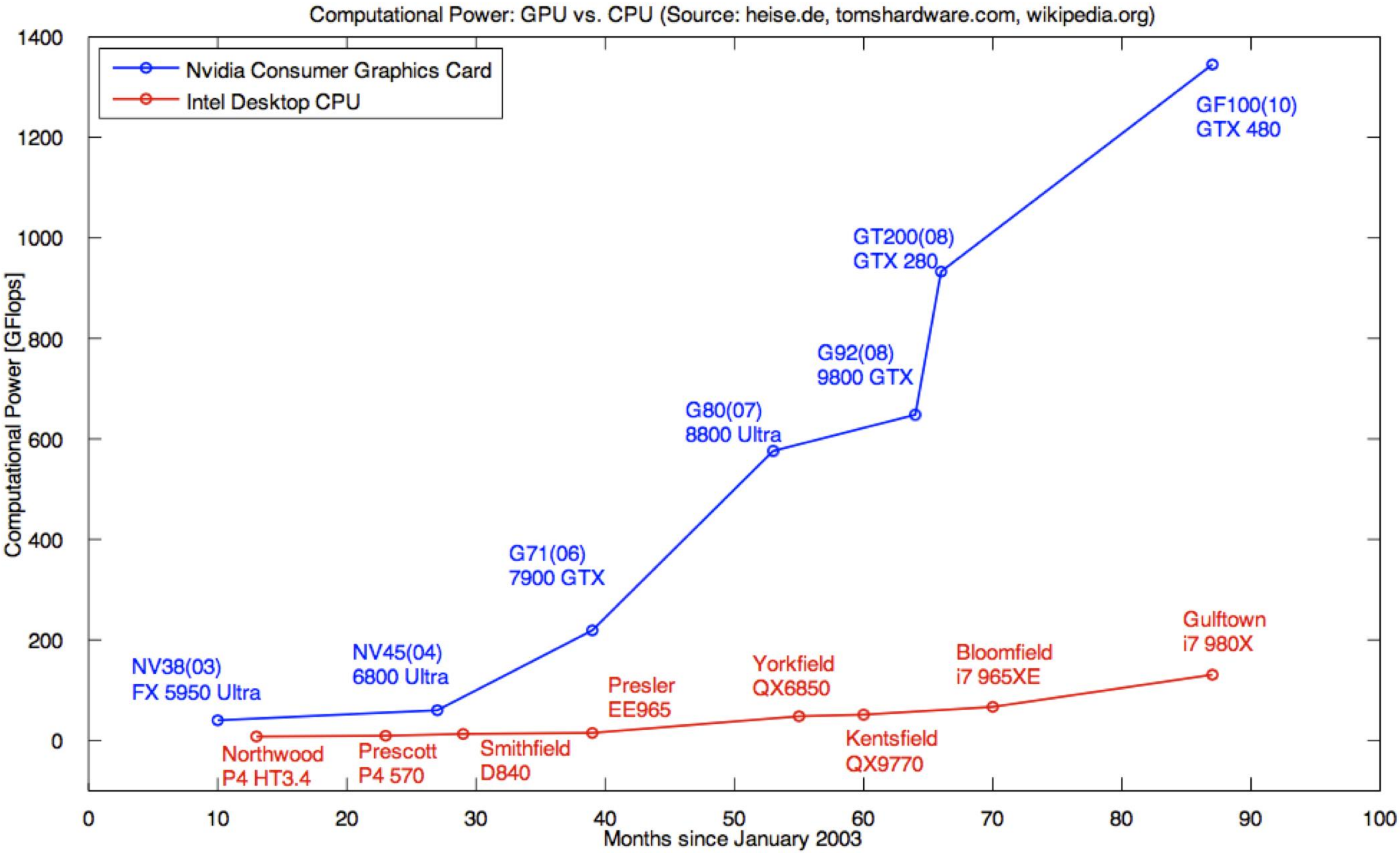
When GPGPU

- GPUs run *thousands of lightweight threads* ***in parallel***
  - *Typically* on consumer hardware: 1024 threads per multiprocessor; 30 multiprocessor => **30k threads**.
  - Compared to CPU: 4 quad core support 32 threads (with HyperThreading).
- Well suited for **data-parallelism**
  - The same instructions executed on multiple data in parallel
  - High **arithmetic intensity**: *arithmetic operations / memory operations*



[Source: nvidia]

Why GPGPU

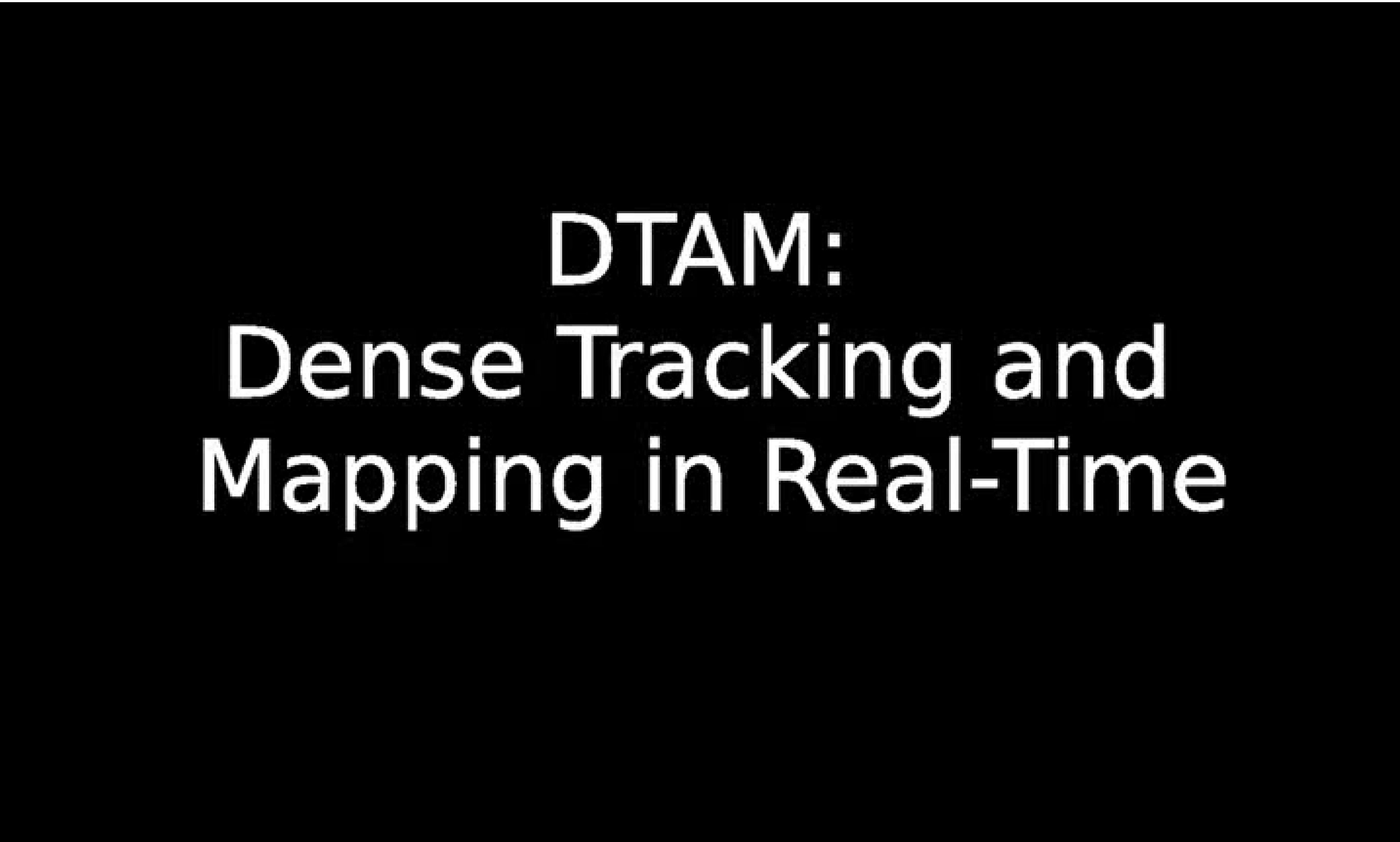


Computational Power: GPU vs. CPU (Source: heise.de, tomshardware.com, wikipedia.org)
Nvidia Consumer Graphics Card
Intel Desktop CPU
Computational Power [GFlops]
Months since January 2003
0
200
400
600
800
1000
1200
1400
0
10
20
30
40
50
60
70
80
90
100
NV38(03)
FX 5950 Ultra
NV45(04)
6800 Ultra
G71(06)
7900 GTX
G80(07)
8800 Ultra
G92(08)
9800 GTX
GT200(08)
GTX 280
GF100(10)
GTX 480
Northwood
P4 HT3.4
Prescott
P4 570
Smithfield
D840
Presler
EE965
Yorkfield
QX6850
Kentsfield
QX9770
Bloomfield
i7 965XE
Gulftown
i7 980X

GPGPU for 3D Reconstruction

- **Image processing**
  - Filtering
  - Warping
  - Feature extraction
- **Multi-view geometry**
  - Search for dense correspondence
    - *Pixel-wise* operations (correlation)
    - Matrix and vector operations (epipolar geometry)
  - Photometric Cost Aggregation
- **Global optimization**
  - *Variational methods*
    - ***Parallel, in-place*** operations for gradient / divergence computation

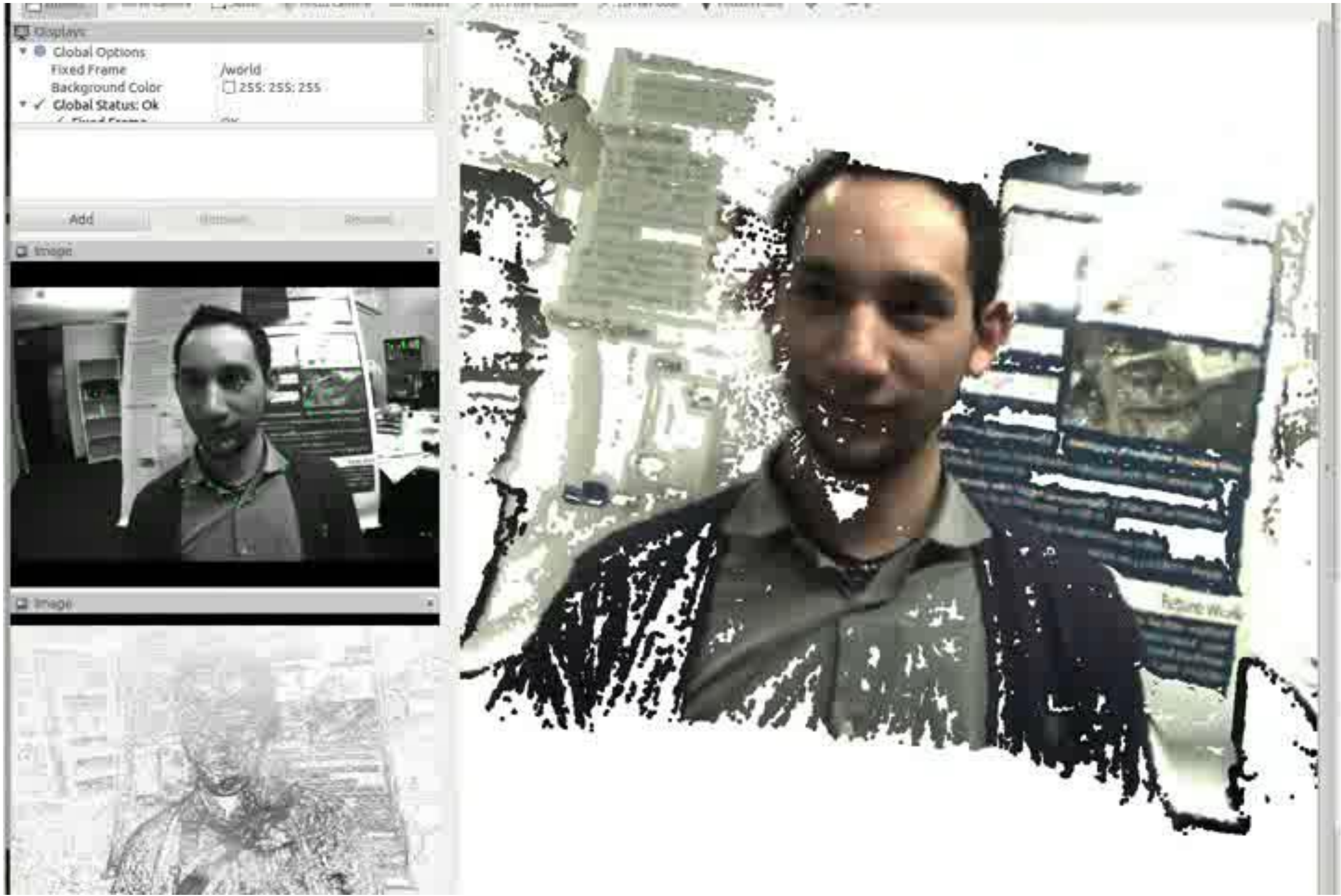
DTAM: Dense Tracking and Mapping in Real-Time, ICCV'11
by Newcombe, Lovegrove, Davison



REMODE:
Regularized Monocular Dense Reconstruction

[M. Pizzoli, C. Forster, D. Scaramuzza, REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time, IEEE International Conference on Robotics and Automation 2014]

REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time, ICRA’14, by Pizzoli, Forster, Scaramuzza

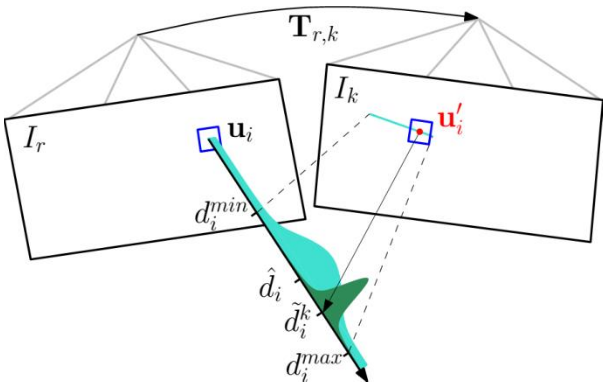


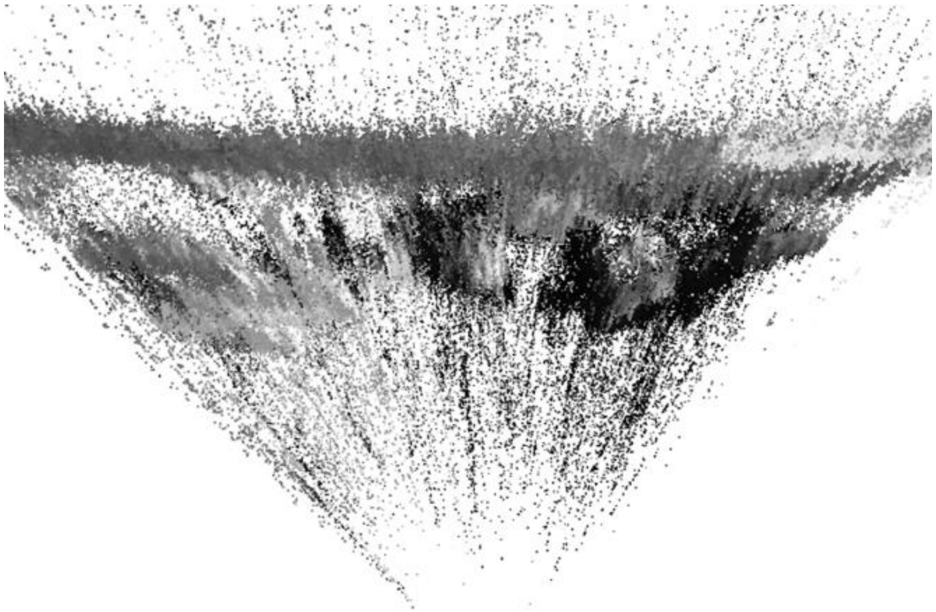
REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time, ICRA’14, by Pizzoli, Forster, Scaramuzza

[M. Pizzoli, C. Forster, D. Scaramuzza, REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time, ICRA’14]

REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time, ICRA'14, by Pizzoli, Forster, Scaramuzza

- Tracks every pixel (like DTAM) but **Probabilistically**
- Runs live on video streamed from MAV (50 Hz on GPU)
- Copes well with low texture surfaces

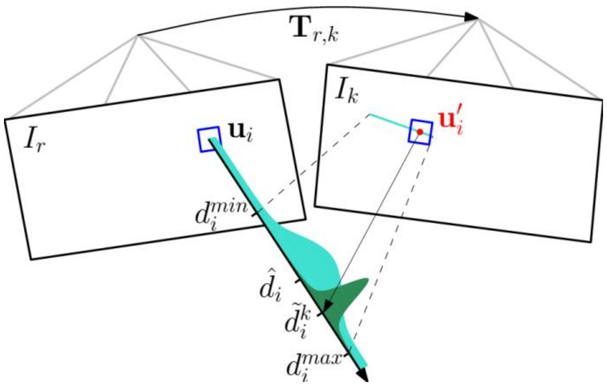




[Pizzoli, Forster, Scaramuzza, REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time ICRA'14]

REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time, ICRA'14, by Pizzoli, Forster, Scaramuzza

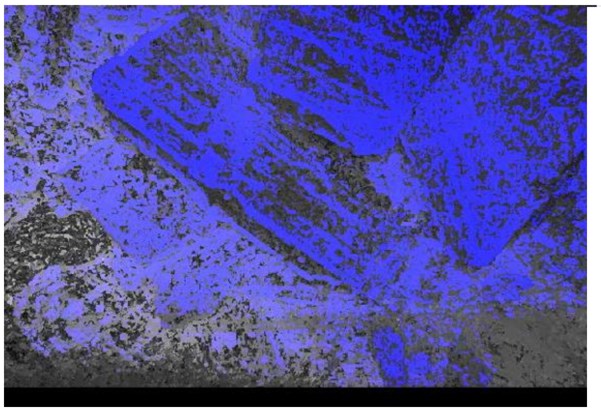
- Tracks every pixel (like DTAM) but **Probabilistically**
- Runs live on video streamed from MAV (50 Hz on GPU)
- Copes well with low texture surfaces

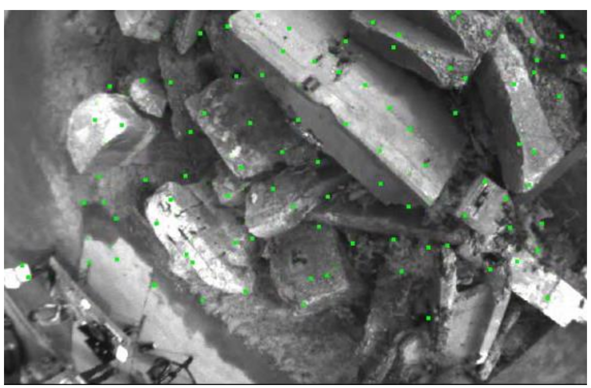


[Pizzoli, Forster, Scaramuzza, REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time ICRA'14]

REMODE applied to autonomous flying 3D scanning

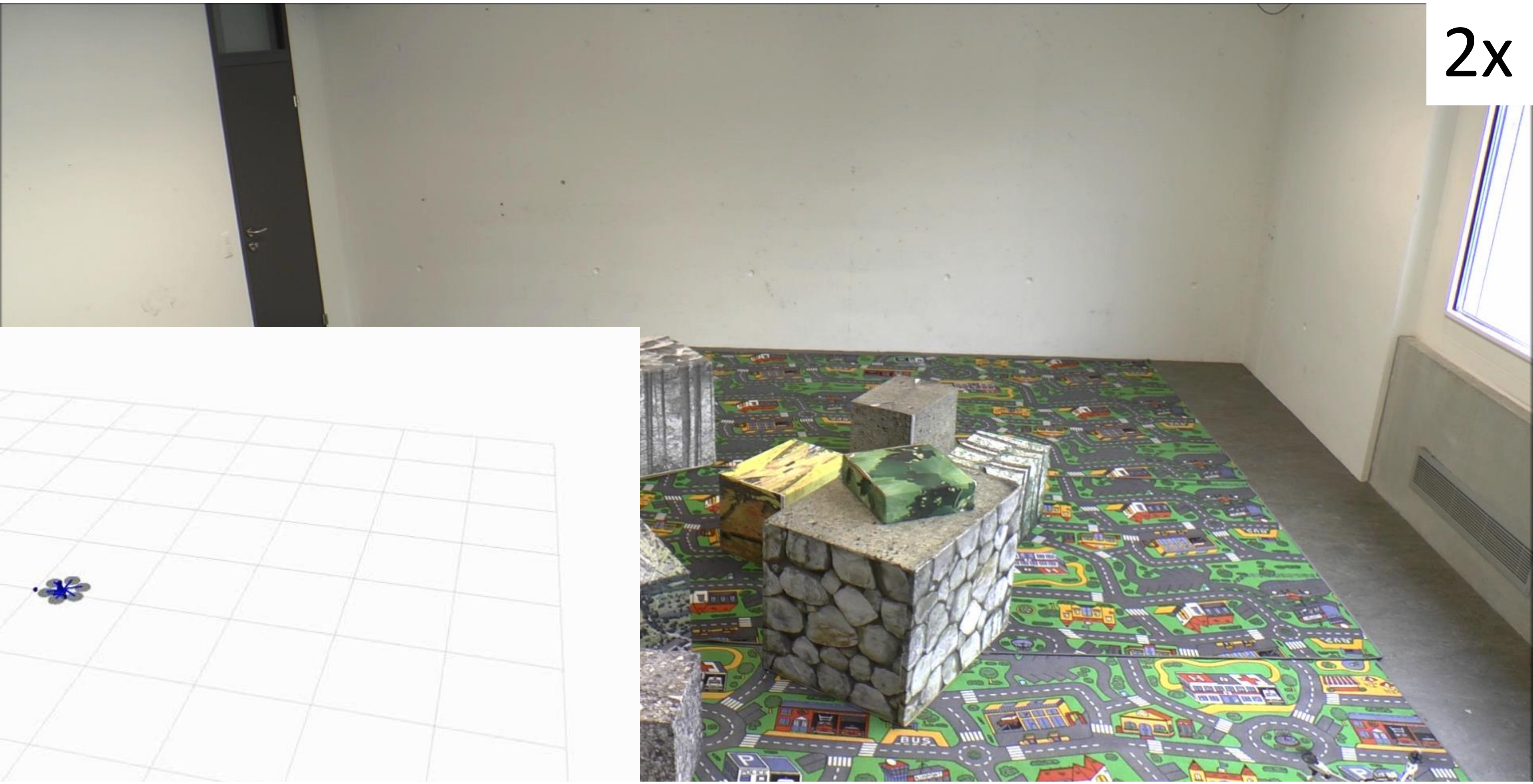
- Tracks every pixel (like DTAM) but **Probabilistically**
- Runs live on video streamed from MAV (50 Hz on GPU)
- Copes well with low texture surfaces





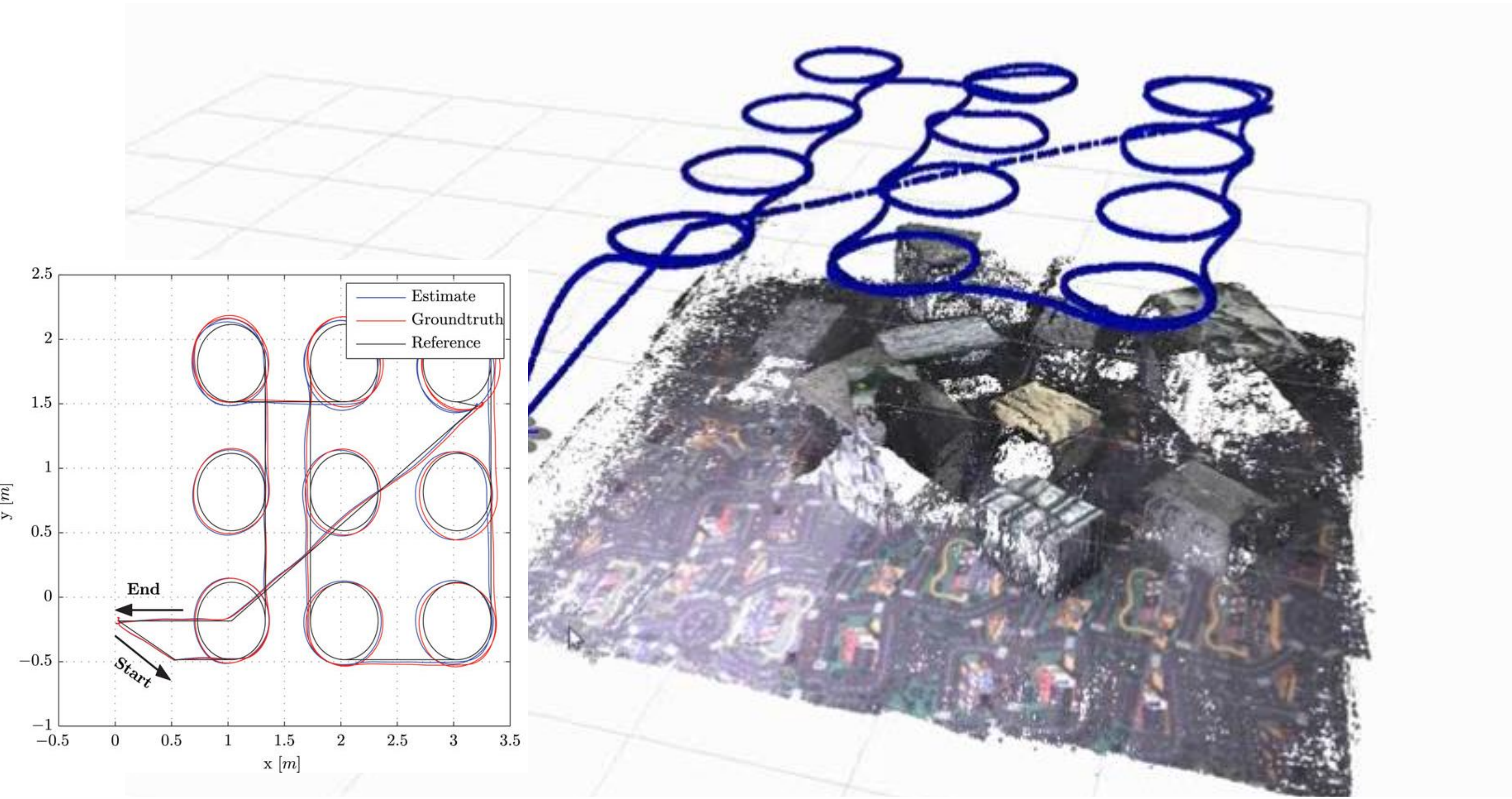
Live demonstration at the Firefighter Training Area of Zurich
Featured on ARTE Tv channel on November 22 and SRF 10vo10

REMODE applied to autonomous flying 3D scanning



Faessler, Fontana, Forster, Mueggler, Pizzoli, Scaramuzza, Autonomous, Vision-based Flight and Live Dense 3D Mapping with a Quadrotor Micro Aerial Vehicle, **Journal of Field Robotics, 2015**.

REMODE applied to autonomous flying 3D scanning



Faessler, Fontana, Forster, Mueggler, Pizzoli, Scaramuzza, Autonomous, Vision-based Flight and Live Dense 3D Mapping with a Quadrotor Micro Aerial Vehicle, **Journal of Field Robotics, 2015**.

REMODE applied to autonomous flying 3D scanning

Live demonstration at the Firefighter Training Area of Zurich
Featured on ARTE Tv channel on November 22 and SRF 10vo10

arte SRF

[Pizzoli, Forster, Scaramuzza, REMODE: Probabilistic, Monocular Dense Reconstruction in Real Time ICRA'14]

3DAround iPhone App

Dacuda