Adaptive Direct RGB-D Registration and Mapping for Large Motions

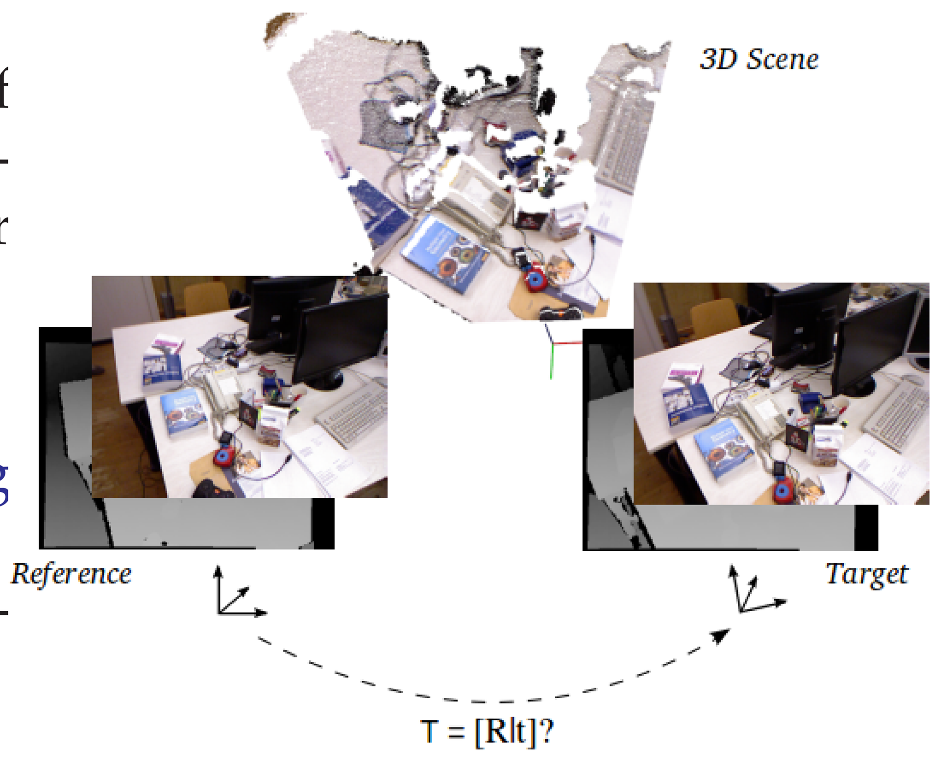
Renato Martins, Eduardo Fernandez-Moral and Patrick Rives

Inria Sophia Antipolis, France

{renato-jose.martins, eduardo.fernandez-moral, patrick.rives}@inria.fr

Problem and Motivation

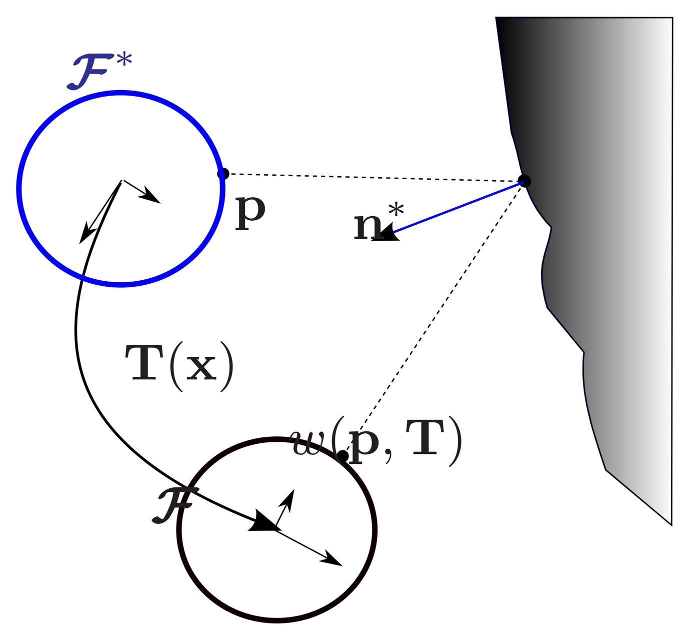
- **Direct RGB-D registration limitation:** models are valid for small motions.
- **Main objective**: the design of a **robust/efficient** direct RGB-D registration technique for large motions.
- **Multiple applications**:
  - Visual odometry, mapping and SLAM;
  - Navigation and visual servoing;
  - Augmented reality.



Background & Related Works

Classic RGB-D formulation: Find pose $\mathbf{T}(\mathbf{x}) \in \mathbb{SE}(3)$ that minimizes

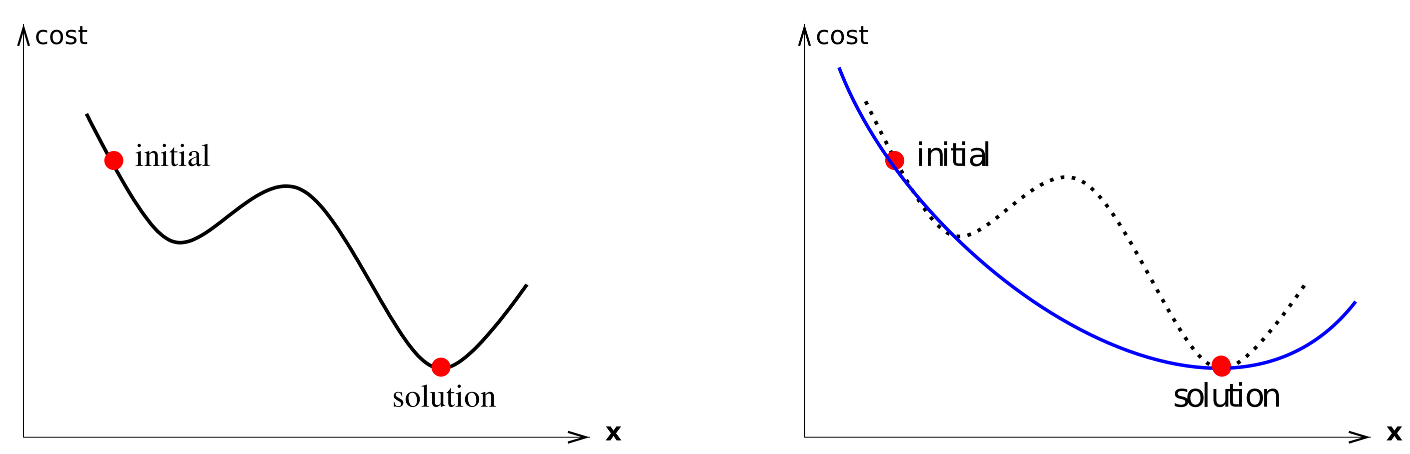
$$C(\mathbf{x}) = C_I(\mathbf{x}) + \mu^2 C_D(\mathbf{x})$$



- $C_I(\mathbf{x}) = \rho(\mathcal{I}(w(\mathbf{p}, \mathbf{T}(\mathbf{x}))) - \mathcal{I}^*(\mathbf{p}))$: SSD of pixel intensities **(photometric term)**;
- $C_D(\mathbf{x}) = \rho((\mathbf{R}\mathbf{n}^*(\mathbf{p}))^T(g(w(\mathbf{p}, \mathbf{T}(\mathbf{x}))) - \mathbf{T}(\mathbf{x})g^*(\mathbf{p})))$: direct point-to-plane error **(geometric term)**;

Denoting: $\mathbf{n}^*$: normal vector; $g(\bullet)$: 3D point; $\mathbf{w}(\bullet)$: warping; $\rho(\bullet)$: Tukey's robust function.

- Scaling factor μ:
  - Heuristically set;
  - μ based on covariance of each point [C. Kerl & D. Cremers, ICRA'13];
  - μ scaling pixels to meters [T. Tykkala et al, ICCV'11].



- **How to identify the neighbourhood where the RGB cost term is more discriminant?**

Method: Adaptive Formulation

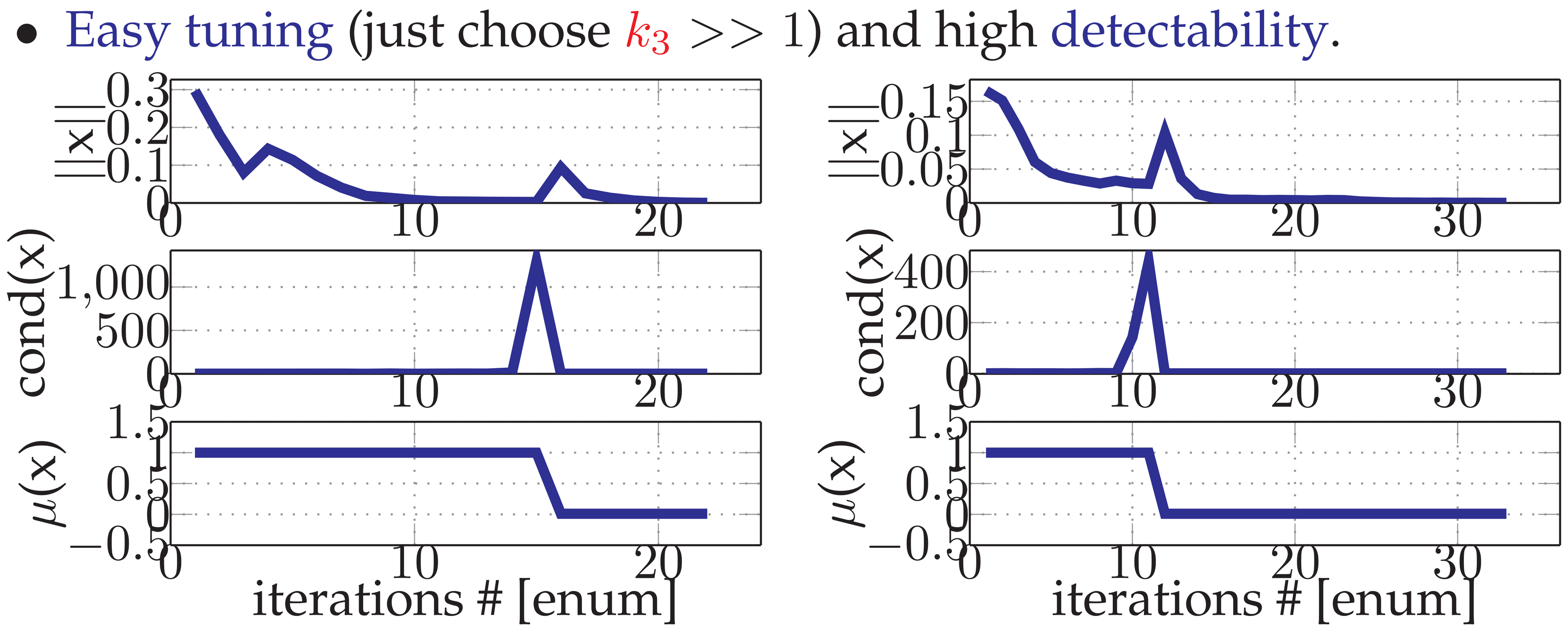
- **Approach:** to explore the relative variation of the RGB (C_I) and geometric (C_D) costs – conditioning:

$$\mu(\mathbf{x}) = \begin{cases} k_1 + k_2, \text{ if } cond_{\mathbf{x}}(C_I(\mathbf{x}))/cond_{\mathbf{x}}(C_D(\mathbf{x})) < k_3 \\ k_1, \text{ otherwise.} \end{cases}$$

- With the relative conditioning of a function as:

$$cond_{\mathbf{x}}(C(\mathbf{x})) = \left| \frac{C(\mathbf{x}_0 \circ \mathbf{x}) - C(\mathbf{x}_0)}{C(\mathbf{x}_0)} \right| / \frac{||\mathbf{x}||}{||\mathbf{x}_0||}$$

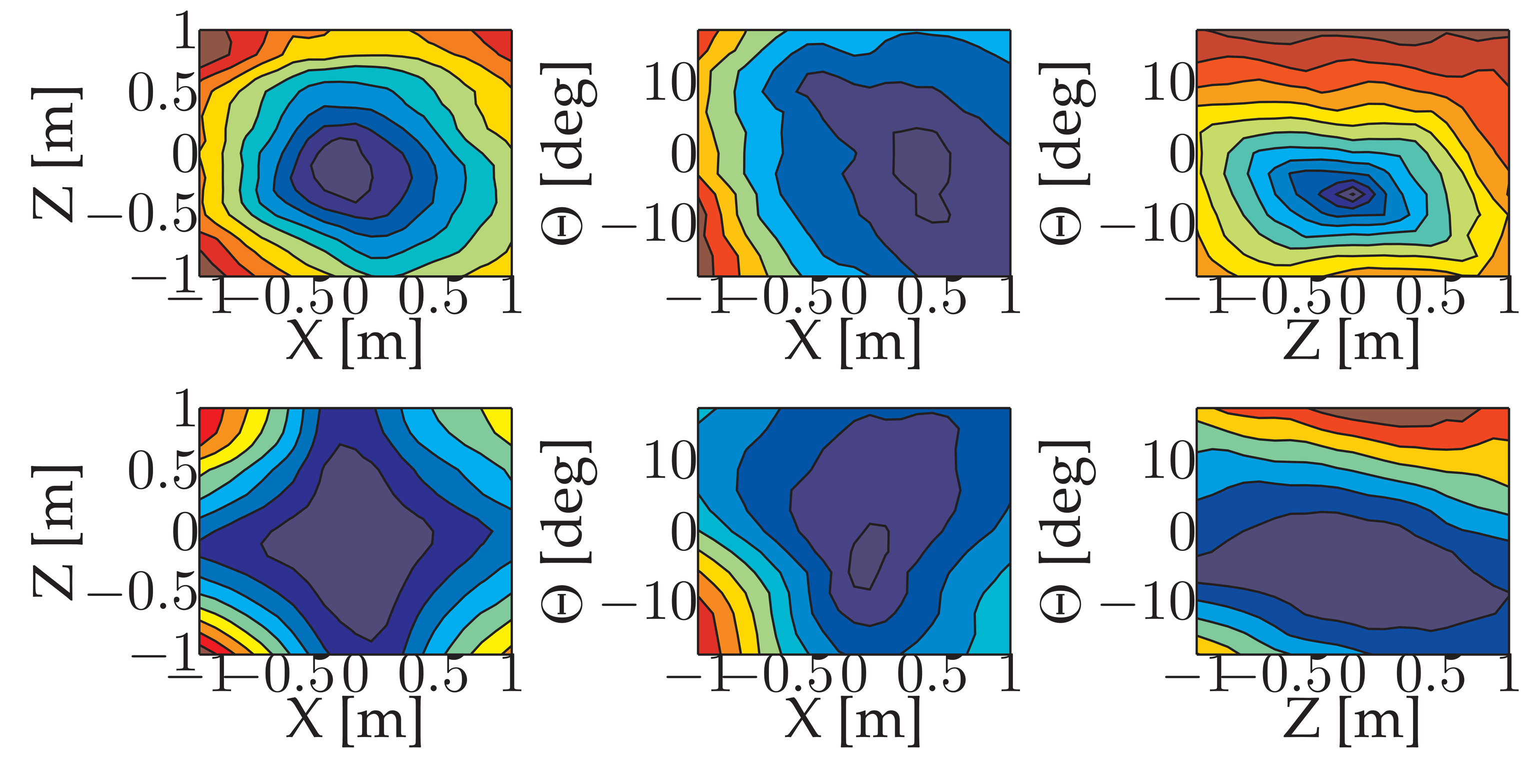
- Easy tuning (just choose $k_3 >> 1$) and high detectability.



Convergence Domain

Photo:

Geo:



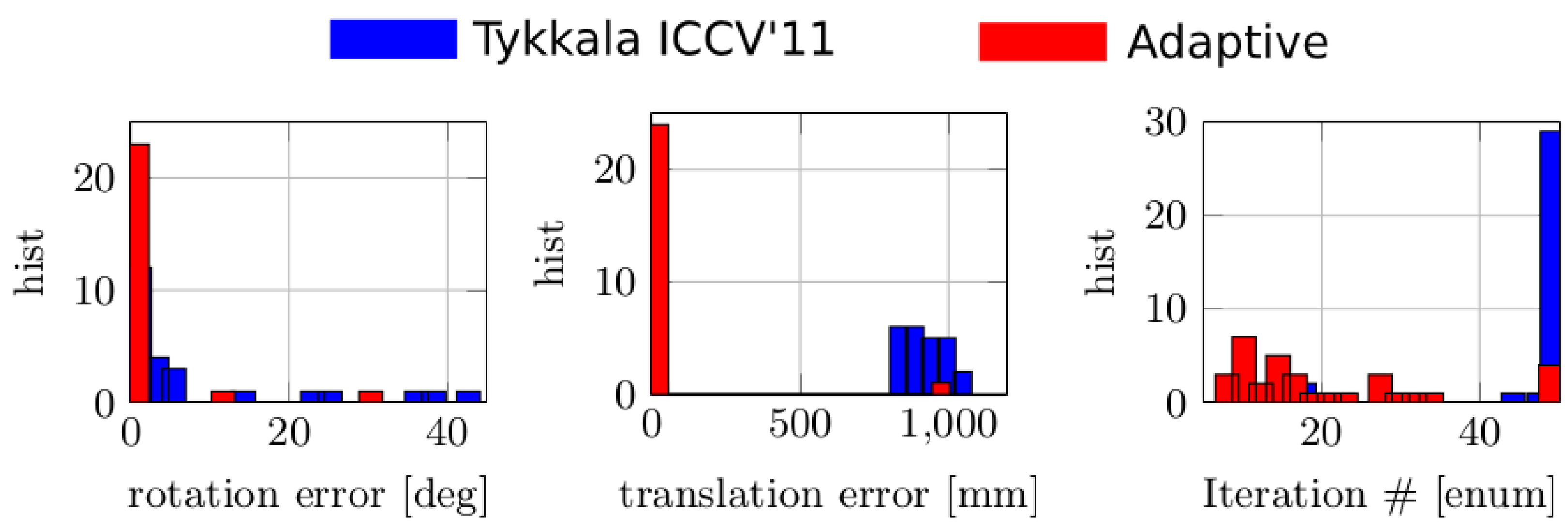
- 1) Geom. cost is flatter than RGB in the neighbourhood of the solution;
- 2) Do not guarantee sub-pixel precision from intensity only cost term.

Registration Results

Experimental set-up Sponza Atrium Sequence:

- Spherical sensor model;

- Test with gaps of 15 frames (≈ 1.2 meters between frames).



Experimental set-up KITTI VO/SLAM Sequence 00:

- Multi-resolution: pyramid of four levels;
- Tests with sub-sampling (gaps) of 1,2 or 3 frames.

KITTI outdoor sequence: average RRE[deg]/RTE[mm]

	Gap = 1	*Gap = 2*	*Gap = 3*
[Tykkala, ICCV'11]	0.08/23.1	0.78/268	3.68/1059
Adaptive	**0.06/16.4**	**0.37/47.5**	**1.05/238**

Conclusions & Perspectives

- Adaptive formulation that explores convexity and convergence properties of intensity and geometric data terms;
- Exploit more geometric term when further of the minimum; End up with classic RGB term near the solution;
- **20 times** faster in simulated sequences and at least as **three** times fast in real sequences (fixed resolution);

Next Step: add planes, edgelets/lines and image moments.

Main References

[1] Comport, A., Malis, E., Rives, P.: Real-time quadrifocal visual odometry. IJRR (2010).

[2] Tykkala, T., Audras, C., Comport, A.: Direct iterative closest point for real-time visual odometry. In ICCV Workshops (2011).

[3] Morency, L., Darrell, T.: Stereo tracking using ICP and normal flow constraint. In ICPR (2002).