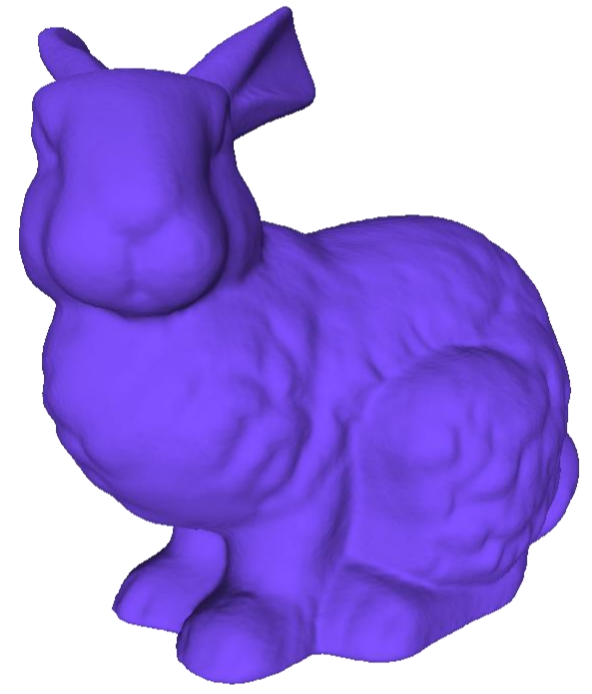


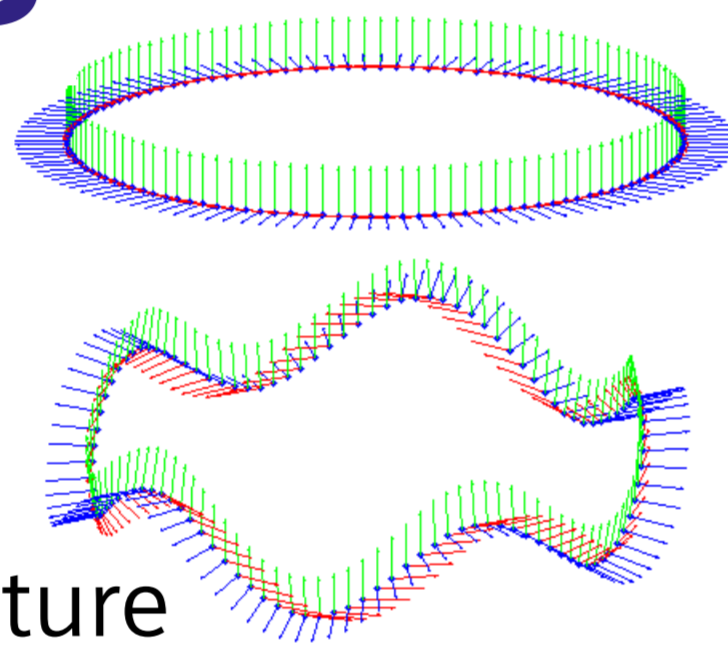
Abstract

We present an approach for highly accurate 3D object reconstruction with RGB-D sensors. The camera motion is tracked frame-to-frame in real-time via a novel direct implicit-to-implicit scheme between pairs of signed distance fields (SDFs), whereby expensive correspondence search is avoided. Our energy formulation further allows for incorporation of photometric and surface orientation constraints. We achieve more precise trajectory estimates and higher quality 3D models than related methods.

Advantages



- ✓ Extremely accurate
- ✓ Fully automatic
- ✓ Real-time
- ✓ Generic motion
- ✓ Generic geometry and texture



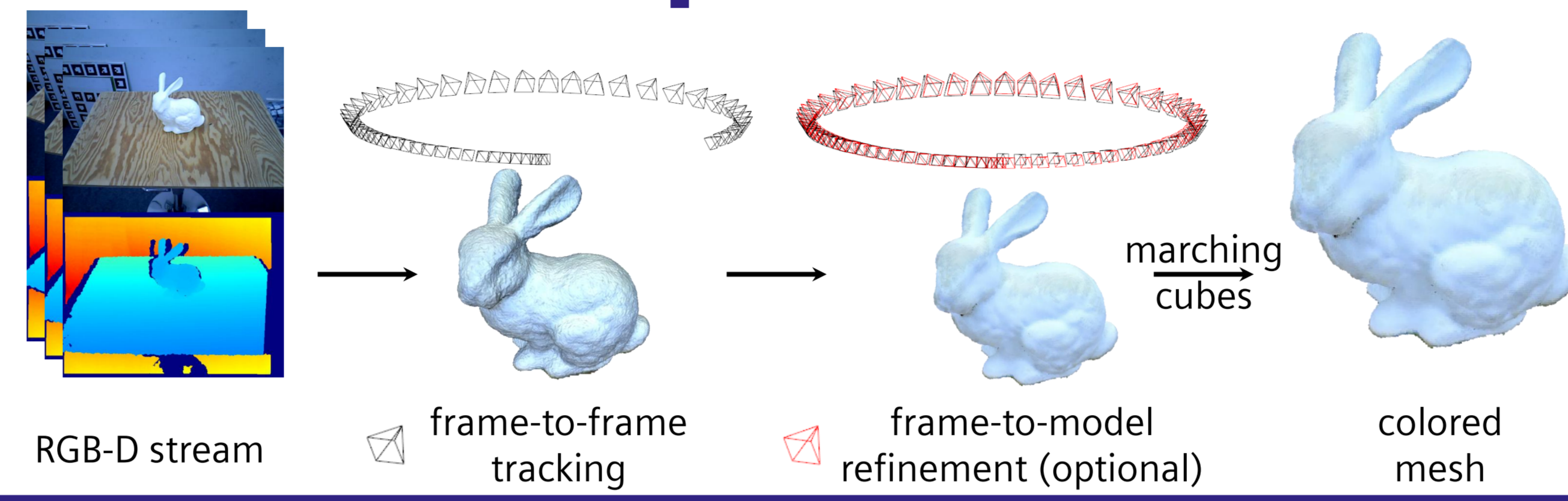
Related Work

- KinectFusion [6]:
 - frame-to-model point-plane ICP tracking
 - global model as weighted average [3] of TSDFs
- Point-to-Implicit [1,2]:
 - frame-to-model projection of point cloud onto global TSDF
- Limitations:
 - ICP constrained to regular motion, not robust to gross outliers
 - cannot remedy large error accumulation in the global model

References

- (1) Bylow, E., Sturm, J., Kerl, C., Kahl, F., Cremers, D.: *Real-Time Camera Tracking and 3D Reconstruction Using Signed Distance Functions* (RSS 2013)
- (2) Canelhas, D.R., Stoyanov, T., Lilienthal, A.J.: *SDF Tracker: A Parallel Algorithm for On-line Pose Estimation and Scene Reconstruction from Depth Images* (IROS 2013)
- (3) Curless, B., Levoy, M.: *A Volumetric Method for Building Complex Models from Range Images* (SIGGRAPH 1996)
- (4) Kehl, W., Navab, N., Ilic, S.: *Coloured Signed Distance Fields for Full 3D Object Reconstruction* (BMVC 2014)
- (5) Masuda, T.: *Registration and Integration of Multiple Range Images by Matching Signed Distance Fields for Object Shape Modeling* (CVIU 2002)
- (6) Newcombe, R.A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A.J., Kohli, P., Shotton, J., Hodges, S., Fitzgibbon, A.: *KinectFusion: Real-Time Dense Surface Mapping and Tracking* (ISMAR 2011)

Pipeline



SDF Generation

- Projective single-frame **TSDF** ϕ with **weights** ω , **color field** ζ [4] and **surface normal field** n [5] from RGB-D pair (I,D):

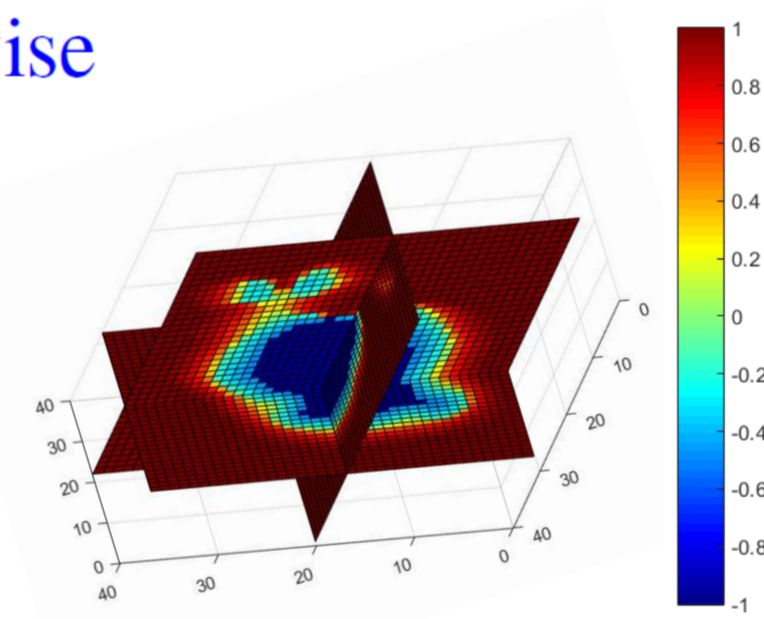
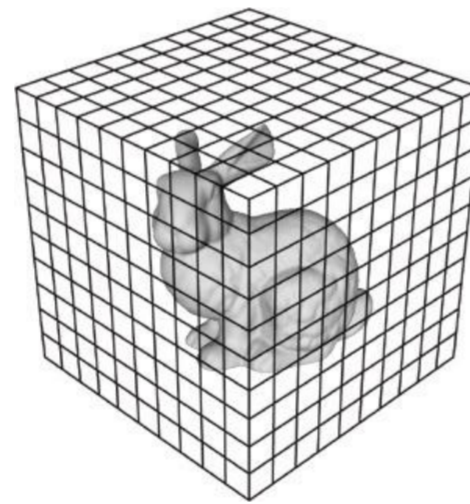
$$\phi_{true}(\mathbf{X}) = D(\pi(\mathbf{X})) - z$$

$$\phi(\mathbf{X}) = \begin{cases} \text{sgn}(\phi_{true}(\mathbf{X})) & , \text{if } |\phi_{true}(\mathbf{X})| \geq \delta \\ \frac{\phi_{true}(\mathbf{X})}{\delta} & , \text{otherwise} \end{cases}$$

$$\omega(\mathbf{X}) = \begin{cases} 1 & , \text{if } \phi_{true}(\mathbf{X}) > -\eta \\ 0 & , \text{otherwise} \end{cases}$$

$$\zeta(\mathbf{X}) = I(\pi(\mathbf{X}))$$

$$\bar{n}(\mathbf{X}) = \nabla_{\mathbf{X}}(\phi_{true}(\mathbf{X}))$$



Registration Energy

- Find 6 DoF camera pose ξ that best aligns two SDFs:

$$E_{geom}(\xi) = \sum_{\text{voxels}} (\phi_{ref}\omega_{ref} - \phi_{cur}(\xi)\omega_{cur}(\xi))^2$$

$$E_{RGB}(\xi) = \sum_{\text{surface voxels}} (\zeta_{ref} - \zeta_{cur}(\xi))^2$$

$$E_{norm}(\xi) = \sum_{\text{surface voxels}} (1 - \bar{n}_{ref} \cdot \bar{n}_{cur}(\xi))$$

$$E_{SDF}(\xi) = \frac{1}{2} (E_{geom}(\xi) + w_{RGB}E_{RGB}(\xi)) + w_{norm}E_{norm}(\xi)$$

- Solution: 6x6 system resulting from first-order Taylor approximation around current pose estimate:

$$\mathbf{A} = \sum_{\text{voxels}} \nabla_{\xi}^T \phi_{cur}(\xi^k) \nabla_{\xi} \phi_{cur}(\xi^k)$$

$$\mathbf{b} = \sum_{\text{voxels}} (\phi_{ref} - \phi_{cur}(\xi^k) + \nabla_{\xi} \phi_{cur}(\xi^k) \xi^k) \nabla_{\xi}^T \phi_{cur}(\xi^k)$$

$$\frac{dE_{geom}}{d\xi} = \mathbf{A}\xi - \mathbf{b} \Rightarrow \xi^* = \mathbf{A}^{-1}\mathbf{b} \Rightarrow \xi^{k+1} = \xi^k + \beta(\xi^* - \xi^k)$$

Qualitative Results



Quantitative Results

