

Telepresence Technologies

Thr, August 6 (Week 7.5)

Point Cloud vs. Skeleton

Point Cloud: directly sends all the RGBD pixels

Skeleton: after the mesh is prepared, only sends the position and rotation data of the bones for every frame.

Comparison:

Point cloud is the raw version, while the skeleton would be the one that requires computer vision to figure out the mesh and skeleton information. As a result, it requires less bandwidth and may have better visual quality.

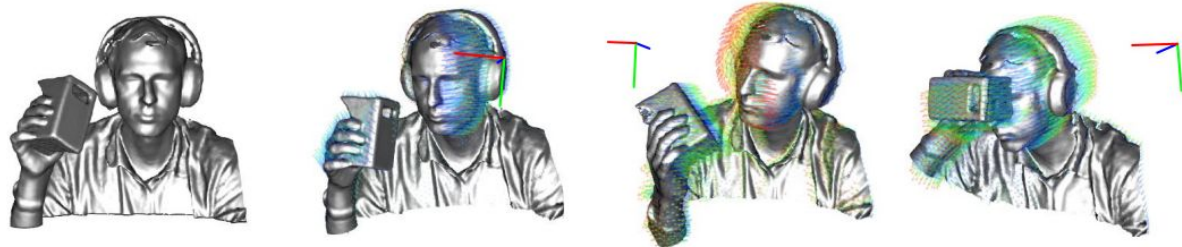
Dynamic Fusion (Newcombe et. al., 2015)

Building a canonical model with the voxel method and warping them for individual frames.



(a) Initial Frame at $t = 0s$

(b) Raw (noisy) depth maps for frames at $t = 1s, 10s, 15s, 20s$



(d) Canonical Model

(e) Canonical model warped into its live frame



Live Input Depth Map



Live Model Output



Live RGB Image (unused)



Canonical Model Reconstruction



Warped Model

Fusion4D (Dou et al., 2016)

Using multiple cameras but still with voxels. Relies on a function more sophisticated than TSDF.

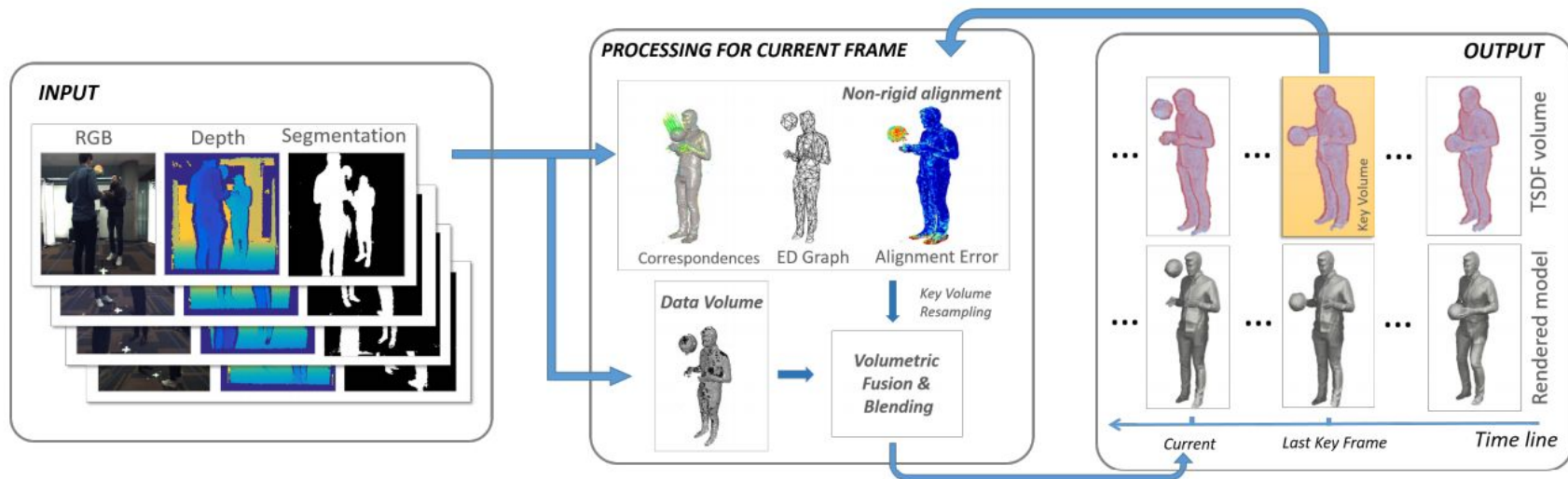


Figure 2: The Fusion4D pipeline. Please see text in Sec. 3 for details.



Fusion4D
(our approach)



Newcombe et al. 2015
(our reimplementation)



Zollhöfer et al. 2014

Usage of Reconstruction Meshes

Applying 3D reconstruction to the person in the telepresence system.

Advantage: The one with highest visual quality.

Disadvantage: Requires a large amount of computation.

Trade-off with the Number of Cameras

Adding more cameras can provide better visual quality; however, not without any cost.

Advantage: Visual Quality

Cost: Computational Complexity, Difficulty to Install the System

3D Reconstruction without External Cameras

AR headset with many cameras building the environment including yourself.

