3DMatch: Learning Local Geometric Descriptors from RGB-D Reconstructions

PRINCETON VISION & ROBOTICS

Andy Zeng Shuran Song Matthias Nießner Matthew Fisher Jianxiong Xiao Thomas Funkhouser

Code & Data
3dmatch.cs.princeton.edu

Motivation

Matching geometry in 3D scans is hard:

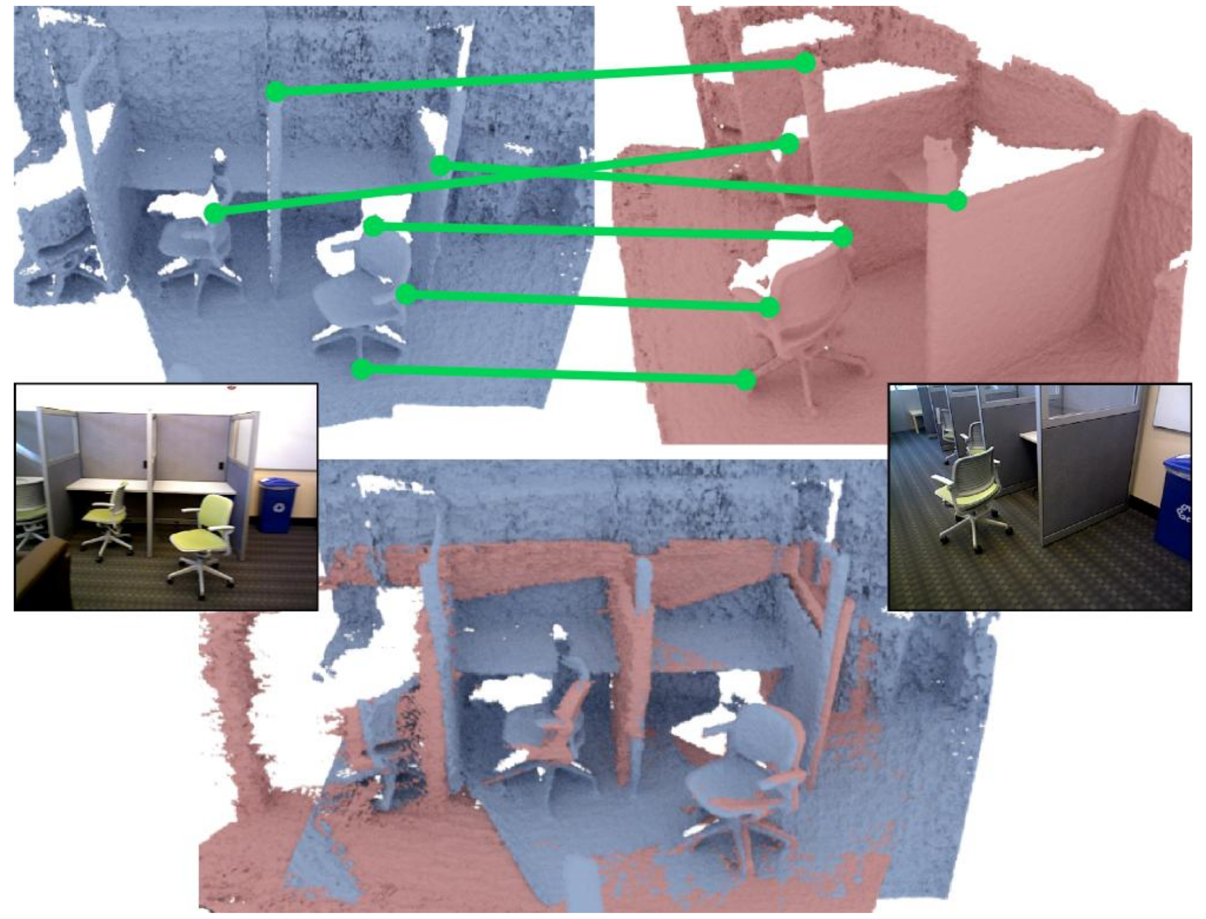
- sensor noise
- low resolution
- viewpoint differences
- partial surfaces

Hand-crafted geometric features:

✓ work well for 3D models with complete surfaces
✗ sensitive to noise and resolution
✗ unstable for partial scans
✗ difficult to generalize to new datasets

Learning a data-driven 3D descriptor?

It is difficult to collect sufficient training data - manually labeling correspondences in 3D scans is not only time consuming, but also prone to errors.

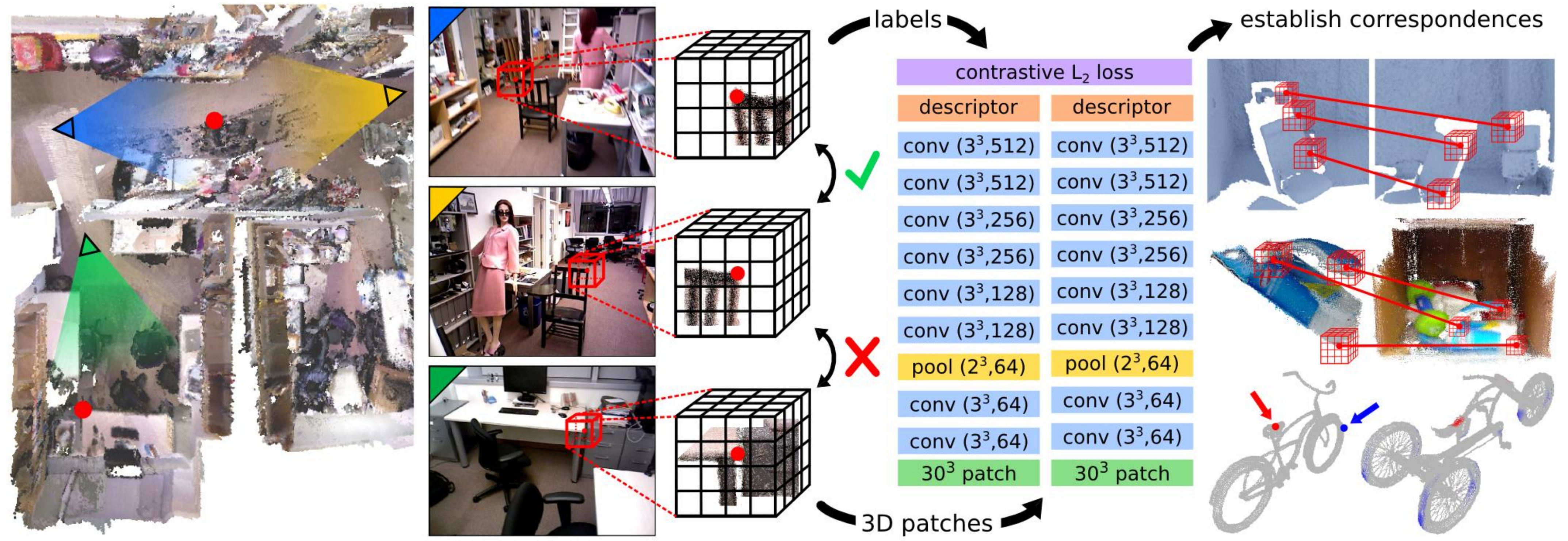


color images for visualization only

Our solution:

We present 3DMatch, a data-driven local 3D descriptor for matching geometric features in noisy and partial 3D scanning data. We amass training data by leveraging the free, long-range correspondence labels found in completed RGB-D scene reconstruction datasets.

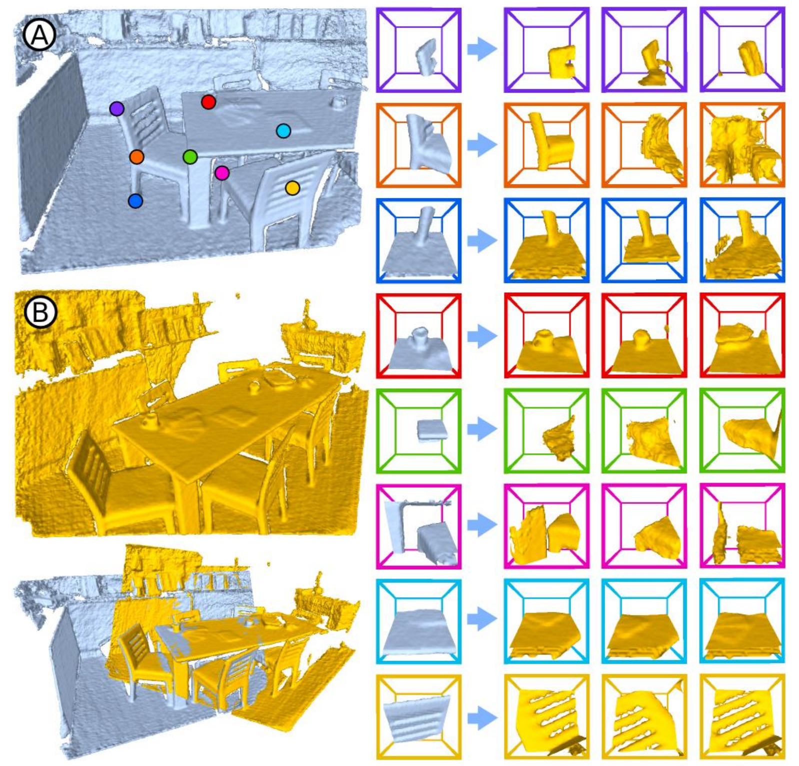
Self-supervised Learning from RGB-D Reconstructions



(a) RGB-D reconstructions (b) frames (c) local patches (d) siamese network (e) applications

From existing RGB-D reconstructions (a), we extract local 3D patches and correspondence labels from scans of different views (b). We collect pairs of matching and non-matching local 3D patches and convert into a volumetric representation (c) to train a 3D ConvNet-based descriptor (d). This geometric descriptor can be used to establish correspondences for matching 3D geometry in various applications (e) such as reconstruction, model alignment, and surface correspondence.

Keypoint Matching

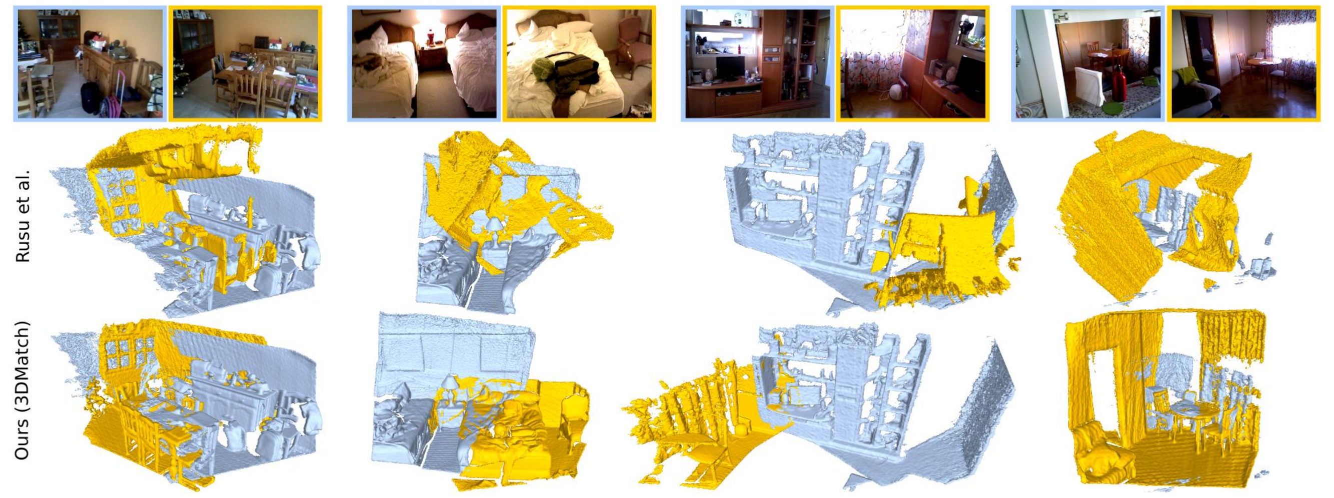


(Left) two scans (A and B) from different view angles. (Right) each row shows a local 3D patch from A, and three nearest neighbor local 3D patches from B using 3DMatch.

Geometric Registration

Method	Recall (%)	Precision (%)
Drost *et al.* [10]	5.3	1.6
Mellado *et al.* [24]	17.8	10.4
Rusu *et al.* [28]	44.9	14.0
Choi *et al.* [5]	59.2	19.6
Zhou *et al.* [45]	51.1	23.2
Rusu *et al.* [28] + RANSAC	46.1	19.1
Johnson *et al.* [19] + RANSAC	52.0	21.7
Ours + RANSAC	**65.1**	**25.2**

When combined with RANSAC, 3DMatch is able to outperform many other state-of-the-art geometric registration algorithms.

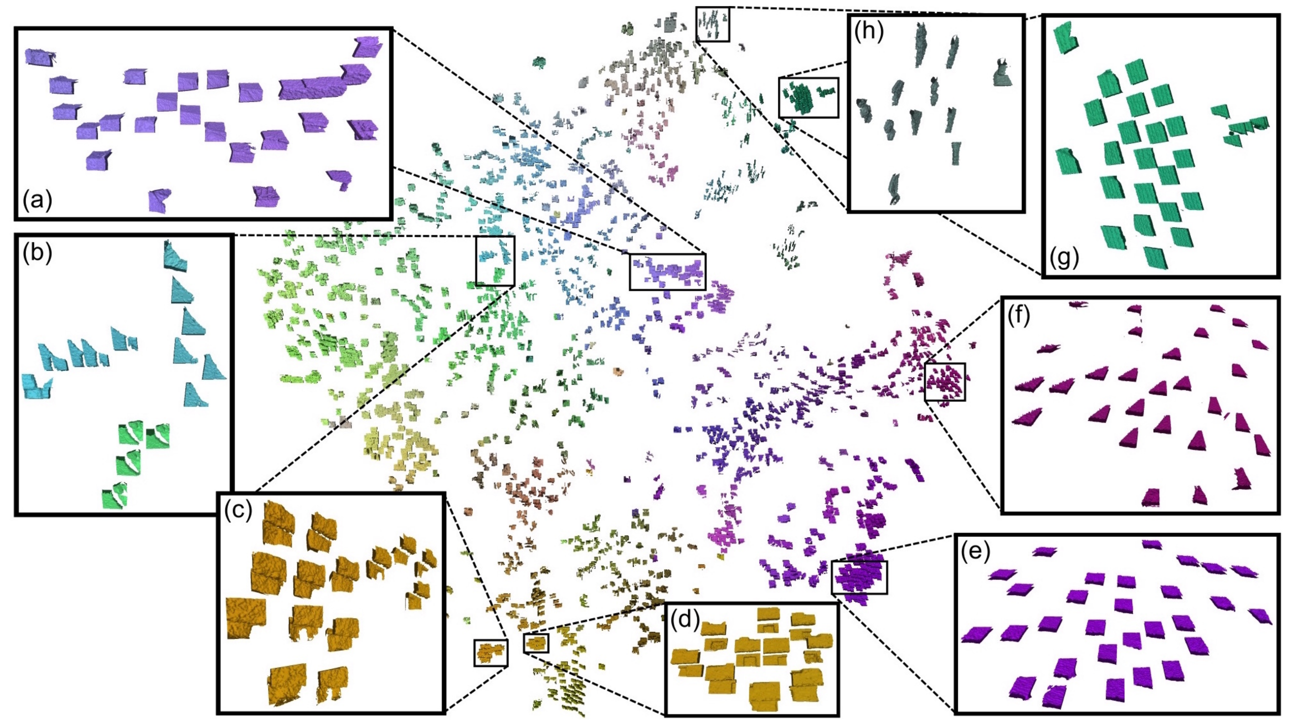


While Rusu et al. fails at aligning scans in challenging cases of loop closures, 3DMatch is able to successfully align each pair of scans by matching local geometric features.

Supersizing Training Data

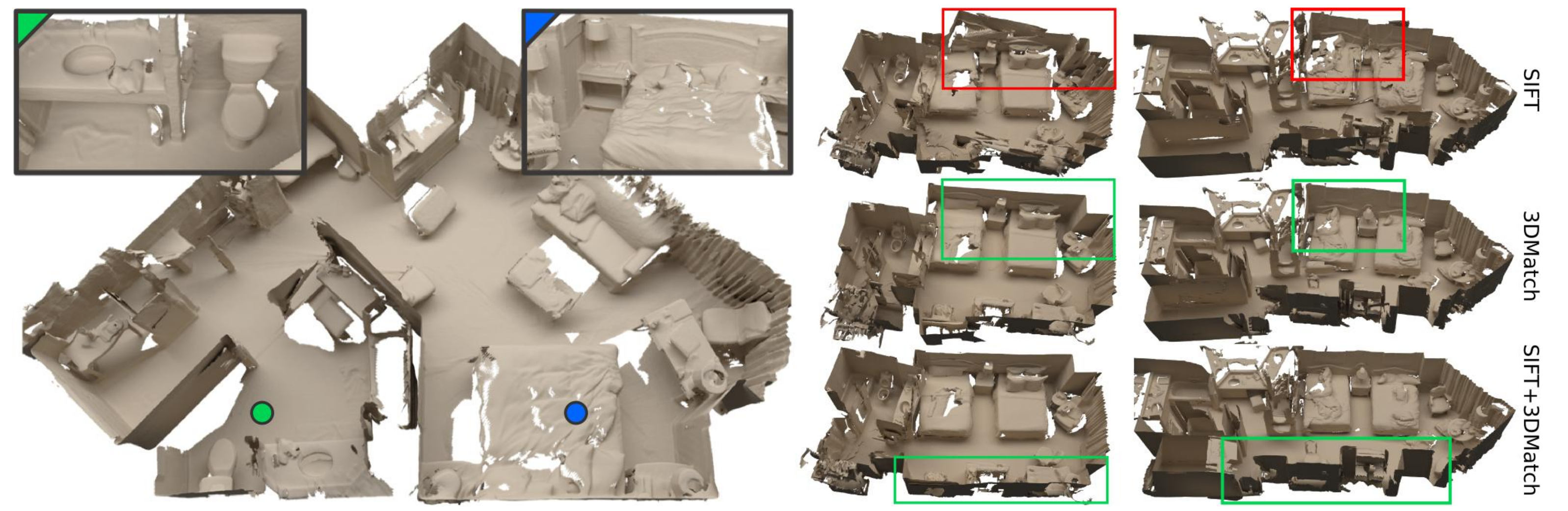
We make use of multiple RGB-D reconstruction datasets to train 3DMatch for our experiments. Each dataset contains depth scans of different environments with local geometries at varying scales, registered together by different reconstruction algorithms. These datasets provide a diverse surface correspondence training set with varying levels of sensor noise, viewpoint variance, and occlusion patterns.

Feature Visualization



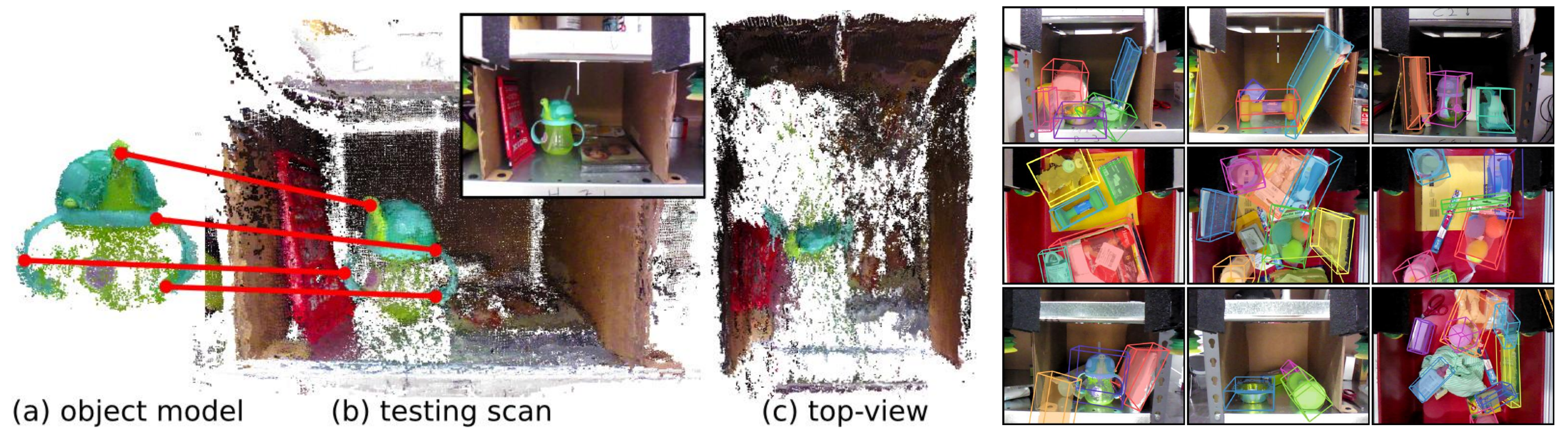
t-SNE embedding of 3DMatch descriptors for local 3D patches, which suggests that our 3DMatch ConvNet is able to cluster local 3D patches based on local geometric structures such as edges, planes, and corners.

3DMatch for 3D Reconstructions



(Left) reconstruction of an apartment using only 3DMatch. (Right) comparing of reconstructions results: 3DMatch is complementary to color features.

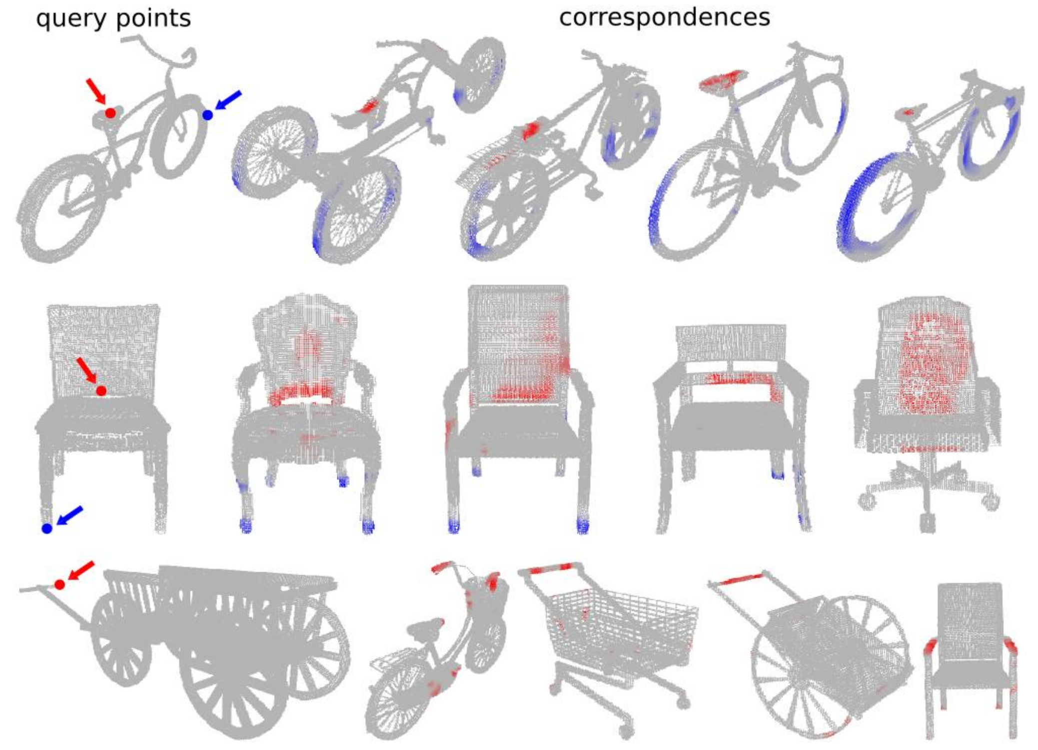
3DMatch for Object Pose Estimation



(a) object model (b) testing scan (c) top-view

(Left) 3DMatch can be used for 6D pose estimation by aligning object models (a) to noisy and partial 3D scans (b,c). (Right) predicted object poses using 3DMatch + RANSAC on the Shelf & Tote Benchmark.

3DMatch for Mesh Correspondence



Without any finetuning, 3DMatch can generalize to find geometrically similar correspondences on complete 3D meshes of the same object category (top and middle rows), and across different object categories (bottom row).