Clustering in Registration of 3D Point Clouds

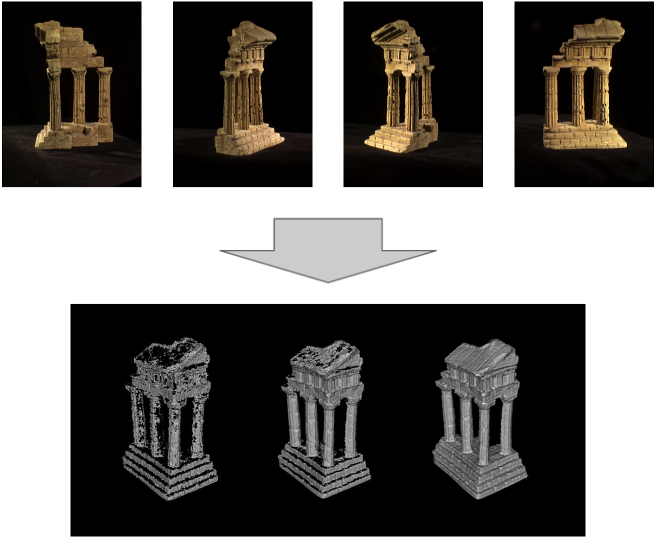
Sergey Arkhangelskiy [1] Ilya Muchnik [2]

[1]Google Moscow

[2]Rutgers University, New Jersey, USA

International Workshop "Clusters, orders, trees: Methods and applications" in honor of Professor Boris Mirkin
December 12-13, Moscow

What is 3D reconstruction?



What is 3D Point Cloud?

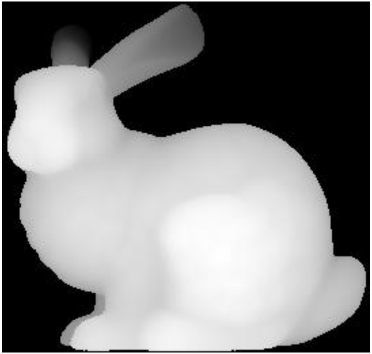
LIDARs (Light Detection and ranging)

And consumer devices like Microsoft Kinect

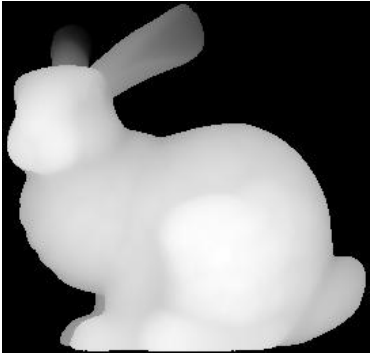
See the world as **depth** image.
In depth image every pixel represents distance from the camera to the scene.

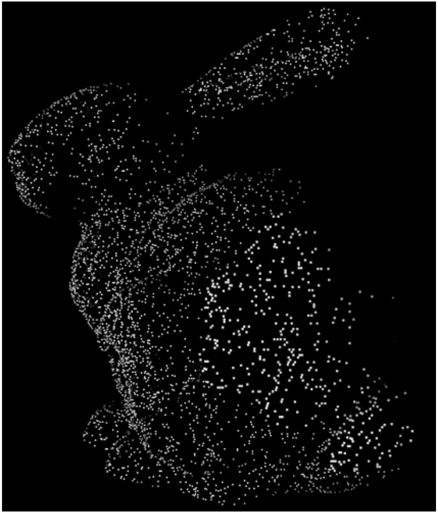
What is 3D Point Cloud?

What is 3D Point Cloud?



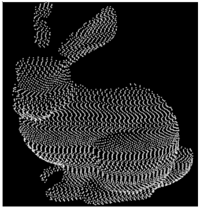
What is 3D Point Cloud?

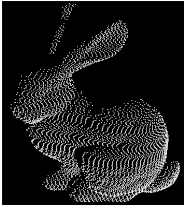


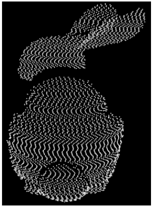


Reconstruction from 3D point clouds

- We have partial views of an object.
- We want to merge them and get integrated representation.



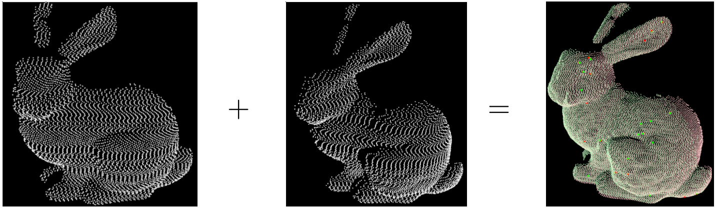




- To integrate we are to find coordinate system transformations to common coordinate frame for each cloud.

Registering two point clouds

- Registering multiple point clouds can be reduced to registering two clouds.
  - $Reconstruction(C_1, C_2) = C_1 + C_2$
  - $Reconstruction(C_1, C_2, C_3) = C_3 + Reconstruction(C_1, C_2)$
  - ...
- For two clouds we have to find just one coordinate frame transformation: $T : C_1 \to C_2$
- How? - that's the question.



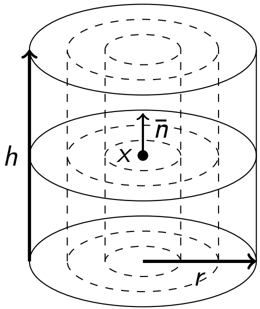
Looking for matches

- Match - pair of points $[x, y]$, $x \in C, y \in D$.
- We say match is **correct** if both points are images of the same scene point.
- The hardest part of reconstruction problem is extracting the subset of correct matches.
- Number of all matches is $N_1 \cdot N_2$ ($N_1 = |C|$, $N_2 = |D|$).
- We propose the clustering-based method for this.
- To compute transformation for given correct matches M_0 we solve least-squares problem.

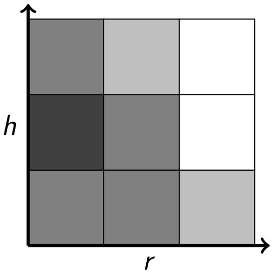
$$\sum_{[x_i, y_i] \in M_0} ||Tx_i - y_i||^2 \to \min_T$$

Local point descriptor. Spin images [JH99].

In every point we build cylinder of fixed radius and height.



We compute histogram for number of nearest points in each ring.



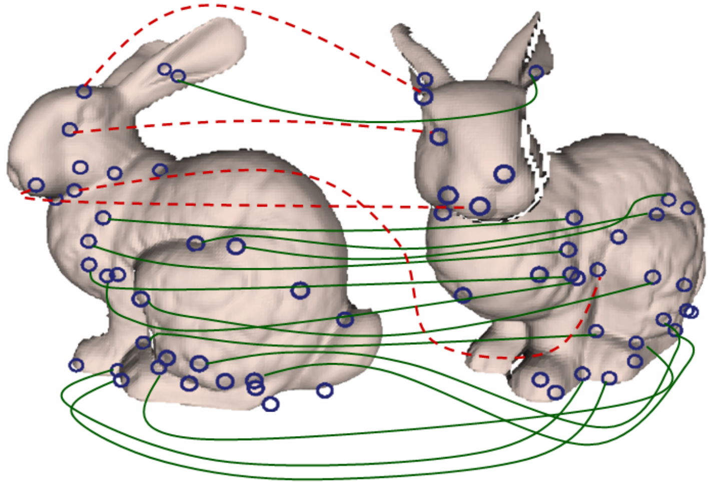
This histogram represents the surface in the neighborhood of x and can be written as a vector.

Descriptor-based matching

- For every $x \in C$ find k points from D with the closest descriptors (k is an external parameter). We obtain $k \cdot N_1$ matches.
- For every $y \in D$ find k closest (in descriptor space) points from C. It gives $k \cdot N_2$ more matches.
- Intersect these sets. Now we have less than $k \cdot \min(N_1, N_2)$ matches.

Descriptor-based matching

Some of them are still mismatches



[CCFM08]

Geometric consistency of matches.

Suppose $[x_1, y_1]$ and $[x_2, y_2]$ are "correct" matches. It means there are points z_1 and z_2 on scene surface that

$$\begin{aligned} x_1 &= T_x z_1, \quad & x_2 &= T_x z_2, \\ y_1 &= T_y z_1, \quad & y_2 &= T_y z_2, \end{aligned}$$

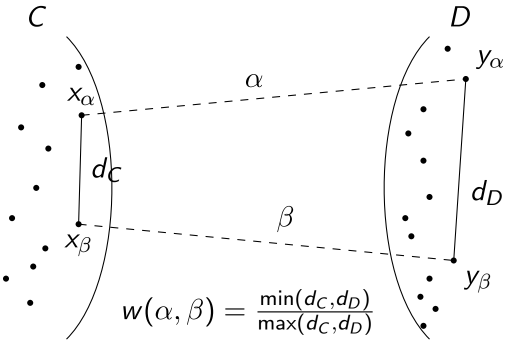
where T_x and T_y are (unknown) rigid transformations mapping object coordinate frame to the cloud frames.
Then

$$||x_1 - x_2|| = ||z_1 - z_2|| = ||y_1 - y_2||$$

Inconsistency function

Let $\alpha = [x_\alpha, y_\alpha]$ and $\beta = [x_\beta, y_\beta]$ be matches. The inconsistency function $w(\alpha, \beta)$:

$$w(\alpha,\beta) = 1 - \frac{\min\{||x_\alpha - x_\beta||, ||y_\alpha - y_\beta||\}}{\max\{||x_\alpha - x_\beta||, ||y_\alpha - y_\beta||\}}$$

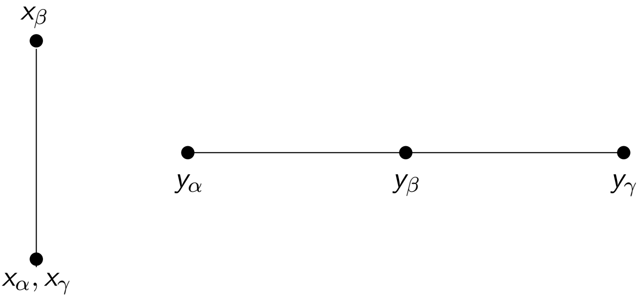


Inconsistency function properties

- $w(\alpha, \alpha) = 0$
- $w(\alpha, \beta) = w(\beta, \alpha)$
- ...

Inconsistency function properties

- $w(\alpha, \alpha) = 0$
- $w(\alpha, \beta) = w(\beta, \alpha)$
- No triangle inequality,
  $w(\alpha, \beta) = 0,\ w(\beta, \gamma) = 0,\ w(\alpha, \gamma) = 1$
  $w(\alpha, \beta) + w(\beta, \gamma) < w(\alpha, \gamma)$



Geometric consistency of the matches

But the good news is that if $[x_1, y_1], \ldots, [x_n, y_n]$ are good matches, then

$$w([x_i, y_i], [x_j, y_j]) \ll 1$$

for all i and j.

That gives us clue to setting up clustering problem...

Matches graph.

We define special undirected graph $G = \langle V, E \rangle$

- Nodes set $V =$ matches set M,
- All nodes are connected: $E = V \times V$,
- All edges are weighted with $w(\alpha, \beta)$.

Looking for a **sparse** subgraph $\langle V', E' \rangle$

- $V' \subseteq V$,
- $E' = V' \times V'$,
- All the edge weights are small $w(\alpha, \beta) \ll 1$.

We can consider this as **single cluster clustering problem**.

Layered clusters [MM02]

For $\alpha \in M$, $H \subseteq M$ we define

$$\pi(\alpha, H) = \sum_{\beta \in H} w(\alpha, \beta)$$

We build the subsets and elements sequences $\{H_i\}$, $\{\alpha_i\}$:

$$\begin{aligned} H_0 &= M, \\ \alpha_i &= \max_{\alpha \in H_{i-1}} \pi(\alpha, H_{i-1}), \\ H_i &= H_{i-1} \setminus \{\alpha_i\}. \end{aligned}$$

The biggest H_i that has average distance between nodes less than given threshold Δ is considered as the final set of matches.

$$\sum_{\alpha, \beta \in H_i} w(\alpha, \beta) < \Delta \cdot |H_i| \cdot (|H_i| - 1)$$

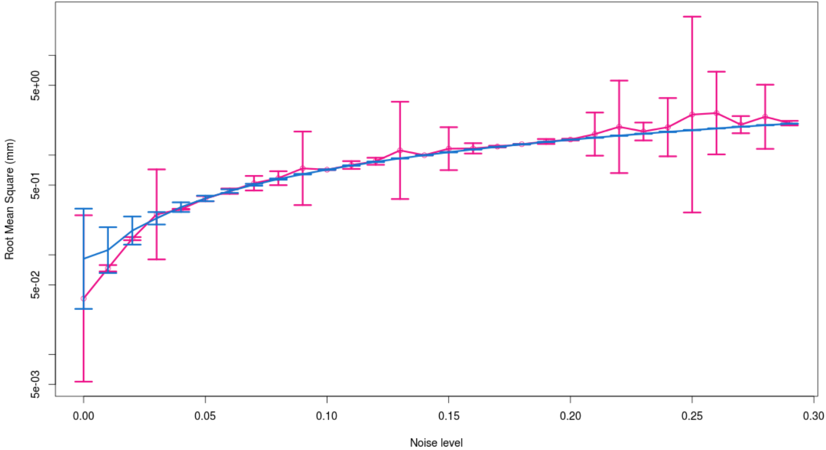
Experimental results

Experiment #1

- We take one depth image of Armadillo model [Sta11].
- For a noise parameter $a \in [0, 0.29]$ we generate noised model. Gaussian noise with deviation $a \cdot \delta$ added to each coordinate of every point. δ is median distance between neighboring points - model resolution parameter.
- Rotate and shift the model randomly.
- Register original and modified models.
- Measure RMS error (root means square, $\frac{1}{n}\sum ||Tx_i - y_y||^2$), rotation and shift error.
- Compare with RANSAC (Random Sample Consensus) registration [FB81].

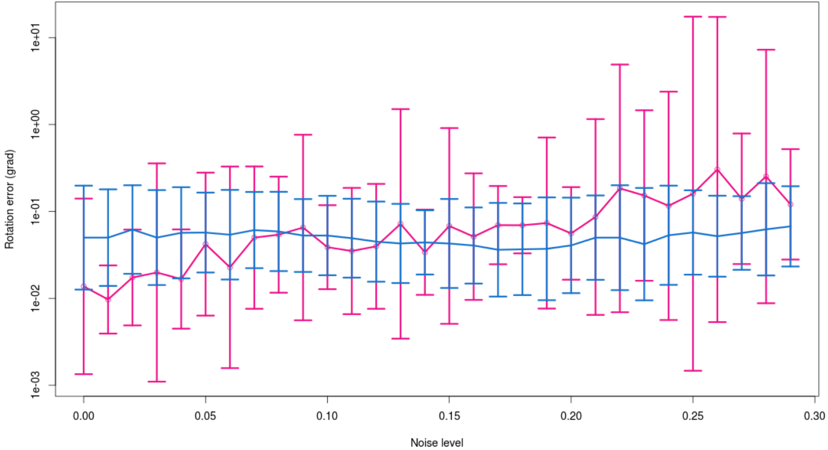
Experimental results

RMS error (mm) for our method and RANSAC-based method on different noise levels.



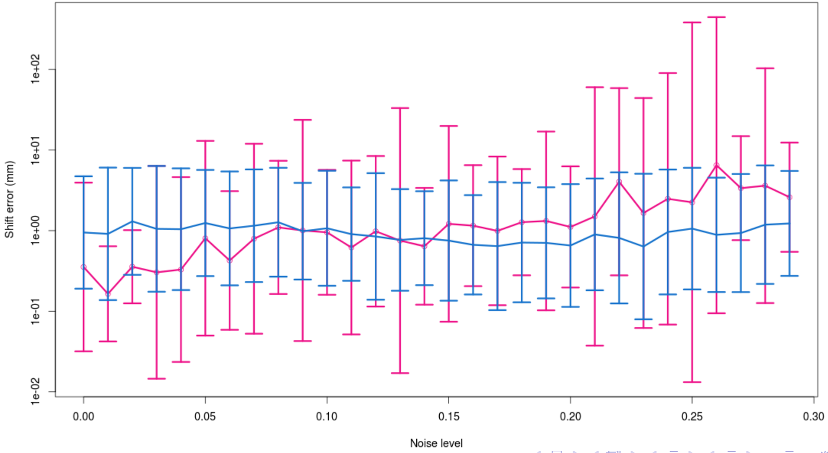
Experimental results

Rotation error (grad) for our method and RANSAC-based method on different noise levels.



Experimental results

Translation error (mm) for our method and RANSAC-based method
on different noise levels.



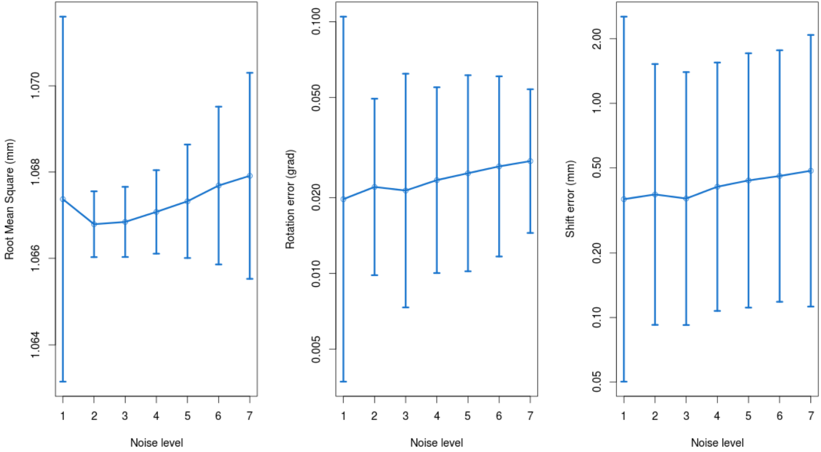
Experimental results

Experiment #2

- Fix noise parameter $a = 0.15$.
- Add noise and random transformation to the model.
- For $k \in \{1, \ldots, 7\}$ (number of closest descriptors) register noised and original model.
- Measure RMS, rotation and shift errors.

Experimental results.Proposed method performance for different k.

Proposed method performance for different k values.



Results

- Special graph representation of point clouds matches was proposed.
- The problem of extracting correct matches was reduced to single cluster clustering problem on this graph.
- Layered clusters method was applied to this problem.
- Performance and stability of the proposed technique is better or comparable to state-of-arts registration methods.

Future work and research directions

- More experiments to understand the parameters influence.
- Theoretical statements, i.e. what properties of extracted cluster guarantee correct registration.
- Examine other clustering procedures, e.g. k-means.
- Simultaneous registration of several point clouds.
- Registering images.

References I

U. Castellani, Marco Cristani, S. Fantoni, and V. Murino.
Sparse points matching by combining 3D mesh saliency with statistical descriptors.
Computer Graphics Forum, 27(2):643–652, April 2008.

Martin A Fischler and Robert C Bolles.
Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography.
Communications of the ACM, 24(6):381–395, 1981.

M Goesele, B Curless, and S M Seitz.
Multi-View Stereo Revisited.
2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Volume 2 CVPR06, 2:2402–2409, 2006.

References II

A.E. Johnson and Martial Hebert.
Using spin images for efficient object recognition in cluttered 3D scenes.
Pattern Analysis and Machine Intelligence, IEEE Transactions on, 21(5):433–449, 1999.

Boris Mirkin and Ilya Muchnik.
Layered clusters of tightness set functions.
Applied mathematics letters, 15, 2002.

S M Seitz, B Curless, J Diebel, D Scharstein, and R Szeliski.
A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms.
2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Volume 1 CVPR06, 1(c):519–528, 2001.

References III

The Stanford 3D Scanning Repository, 2011.

Questions

Thank you for your attention.

Questions?

mc.vertix@gmail.com