A Global Feature-less Scan Registration Strategy Based on Spherical Structural Representation

Bo Sun

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics
Technical Aspects of Multimodal Systems

03. December 2013

Outline

Scan Registration

- **Definition:**
  transforms multiple 3D datasets into the same coordinate system so as to align overlapping components of these datasets
- **Motivation:**
  - robotic mapping, SLAM, path planning, robotic navigation
  - data fusion, object recognition, shape retrieval

State of the art

- **Local methods:**
  - Iterative Closest Point (ICP)
  - Normal Distributions Transform (NDT)
- **Feature-based methods:**
  Program Flowchart:
  key-point extraction $\rightarrow$ feature description $\rightarrow$ feature matching $\rightarrow$ estimation of rigid transformation $\rightarrow$ refinement
  - traditional features
  - large planar patches based (robotic mapping)
- **No initial guess, no features?**
  - Keep listening...

Overview of the algorithm

- Compute the Spherical Structural Representation of scans to detach the rotation and translation information
- Estimate the rotation information between scans based on Spherical Structural Representation using Generalized Convolution Theorem
- Determine the translation information between scans using Phase Only Matched Filtering(POMF)

Characteristics of Sphere S^2

- Sphere could be regarded as critical point between 3D space and 2D surface
- In Group Theory, S^2 is the quotient of $SO(3)$ and $SO(2)$
- 3D rotation of Sphere could be converted to the translation of 2D Sphere Surface

Motivation

- If the structure of the scan could be represented by a graph on S^2, the rotation of the scan could be converted to the translation of 2D Sphere Surface
- If the spherical structural representation is translational invariant, the rotation of the scan could be decoupled from the translation information

Objective

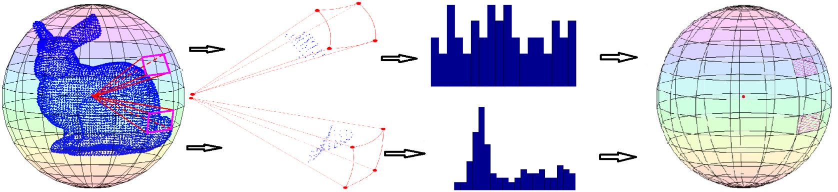
The purpose of computing the Spherical Structural Representation is to maintain the rotation information of the scan by a spherical graph and abandon the translation information.

Requirement

- Independent of the depths of points (translational invariant)
- Robust to outlier:
- Structural discrimination (different structures lead to different values)
- Locality preservation (similar structures result in similar values)

Spherical Entropy Image

Flowchart



Details

- divide the 3D point cloud into several patches (Spherical Grids)
- for each patch, build the histogram and compute the probability density function of variable P
- compute the **entropy** of variable P based on the probability distribution:
  $(P) = -\sum_{d \in D} p(P = d) \cdot log\{p(P = d)\}$
- The structural representation of the 3D point cloud is achieved by computing the entropy of patches in a dense manner, and we call this type of representation as Spherical Entropy Image (SEI)

Spherical Entropy Image

Potential Problem

The scans are divided by several 'analogous square pyramid', and the ways to divide scans are different between different scans.That means given a patch in scan 'A', maybe (generally) you could not find a patch contains the identical structure in scan 'B'.

Explanation

- The objective of SEI is not to extract any structures (features). It is used to describe the original scan. SEI is computed in a dense manner.
- Because the SEI is continuous, which means 'similar structure leads to similar value of SEI', maybe we could not find the precisely identical structure, but at least we could find the most similar structure.
- This problem is common in the algorithms involve division.(NDT,Zebedee Handheld,. . .)
- However, this is error-prone part of this algorithm.

Matching of Patterns on S^2

Problem

- Obtaining the SEIs of two scans, the rotation between the original data could be estimated by aligning their corresponding SEIs.
- How to match two graphs on S^2?

Math Express

- *Given a function h on the sphere, and its rotated version f for some* $g \in SO(3)$*:* $f = \wedge(g) \cdot h$
- Registration of the two functions could be achieved by correlation: $C(g) = \int_{S^2} f(\omega) \bullet \overline{\wedge(g) \cdot h(\omega)} d\omega$

Convolution Theorem

- In mathematics, the convolution theorem states that the Fourier Transform of a convolution is the point-wise product of Fourier Transforms.

$$\mathfrak{F}\{f * g\} = \mathfrak{F}\{f\} \cdot \mathfrak{F}\{g\}$$

- In other words, the convolution in time domain equals point-wise multiplication in frequency domain, then the original problem could be solved in frequency domain more efficiently.

Generalize ?

- Is Convolution Theorem applicable for the functions defined on S^2 ?
- If the answer of last question is 'yes', what is the Fourier Transform of function defined on S^2 ?

Spherical Fourier Transform

- The Fourier Transform decomposes a function into sum of sinusoids. That is, the Fourier Transform gives us another way to represent a function.
- Why 'sinusoids'? The sinusoids are eigenfunctions of the Laplacian operator, so they are maintain fidelity for most real systems.
- The sinusoids are eigenfunctions of Laplacian operator in Cartesian coordinate system. By the same token, the Laplacian operator also has effective forms in spherical coordinate system.

Spherical harmonics

- The spherical Laplacian operator's eigenfunctions are named spherical harmonic functions:
$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(cos\theta)e^{im\phi}$$
- The function defined on S^2 could be expanded as a linear combination of spherical harmonics:
$$f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} f_l^m Y_l^m(\theta,\phi)$$

Generalized Convolution Theorem

$$C(g) = \int_{S^2} f(\omega) \bullet \overline{\wedge(g) \cdot h(\omega)} d\omega \tag{1}$$

$$f(\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_l^m Y_l^m(\omega) \tag{2}$$

$$h(\omega) = \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} b_{l'}^{m'} Y_{l'}^{m'}(\omega) \tag{3}$$

$$C(g) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_l^m \overline{b_l^{m'}} \, \overline{D_{mm'}^l(g)} \tag{4}$$

Substituting equation (2)(3) into equation (1), and utilizing the 'Separation of Variables' technique and orthogonality between basis functions.

Conclusion

- The Generalized Convolution Theorem could convert the convolution of two functions on S^2 into the point-wise product of their corresponding Spherical Fourier Coefficients.
- The Generalized Convolution Theorem are adopted to match the two Spherical Entropy Images; in this way, the rotation between two scans are recovered.

Deduction of Generalized Convolution Theorem

$$C(g) = \int_{S^2} f(\omega) \bullet \overline{\wedge(g) \cdot h(\omega)} d\omega$$

(substitute the Spherical Fourier Expansion into the above equation)

$$= \int_{S^2} [\sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_l^m Y_l^m(\omega)] \bullet \overline{\wedge(g) \cdot [\sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} b_{l'}^{m'} Y_{l'}^{m'}(\omega)]} d\omega$$

(apply the 'sepration of variable' technique, terms that do not depend on ω are factored out)

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} a_l^m \overline{b_{l'}^{m'}} \int_{S^2} Y_l^m(\omega) \cdot \overline{\wedge(g) Y_{l'}^{m'}(\omega)} d\omega$$

(according to the orthogonality of spherical harmonics,the integral equals 0 unless $l' = l$)

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_l^m \overline{b_l^{m'}} \int_{S^2} Y_l^m(\omega) \cdot \overline{\wedge(g) Y_l^{m'}(\omega)} d\omega$$

(since the probability of spherical harmonic:$\wedge(g) Y_l^m(\omega) = \sum_{|k| \leq l} Y_l^k(\omega) D_{km}^l(g)$)

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_l^m \overline{b_l^{m'}} \int_{S^2} Y_l^m(\omega) \cdot \overline{\sum_{|k| \leq l} Y_l^k(\omega) D_{km'}^l(g)} d\omega$$

Deduction of Generalized Convolution Theorem

(apply the 'sepration of variable' technique, terms that do not depend on ω are factored out)

$$C(g) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} \sum_{k=-l}^{l} a_l^m \overline{b_l^{m'}} \, \overline{D_{km'}^l(g)} \int_{S^2} Y_l^m(\omega) \overline{Y_l^k(\omega)} d\omega$$

(according to the orthogonality of spherical harmonics,the integral equals 0 unless $k = m$, and the integral equal 1 when $k = m$.)

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_l^m \overline{b_l^{m'}} \, \overline{D_{mm'}^l(g)}$$

Wigner-D function

- $D_{mm'}^l(g)$ is called **Wigner-D function**; we can write the **Wigner-D function** using the Euler angle decomposition: $D_{mm'}^l(\alpha, \beta, \gamma) = \mathrm{e}^{-\mathrm{i}M\alpha} d_{mm'}^l(\beta) \mathrm{e}^{-\mathrm{i}M'\gamma}$
- In some sense, the $D_{km}^l(g)$ could be interpreted as the $k - th$ component of $\wedge(g)$ acting on $Y_l^m(\omega)$

Conclusion

- The Generalized Convolution Theorem could convert the convolution of two functions on S^2 into the point-wise product of their corresponding Spherical Fourier Coefficient.
- The Generalized Convolution Theorem are adopted to match the two Spherical Entropy Images; in this way, the rotation between two scans are recovered.

Translation determination

Problem

The translational invariant Spherical Entropy Image decouples the rotation and translation information of scans. Achieved the rotation, the scans could be rotated according to the recovered rotation. After that two scans are shifted cousins.

Solution (POMF)

Let $f_1(x, y, z)$ and $f_2(x, y, z)$ be two shifted signals, and $\mathcal{F}_1(u, v, k)$ and $\mathcal{F}_2(u, v, k)$ be their corresponding Fourier spectra. The shift between these two translated signals could be solved by the following equations:

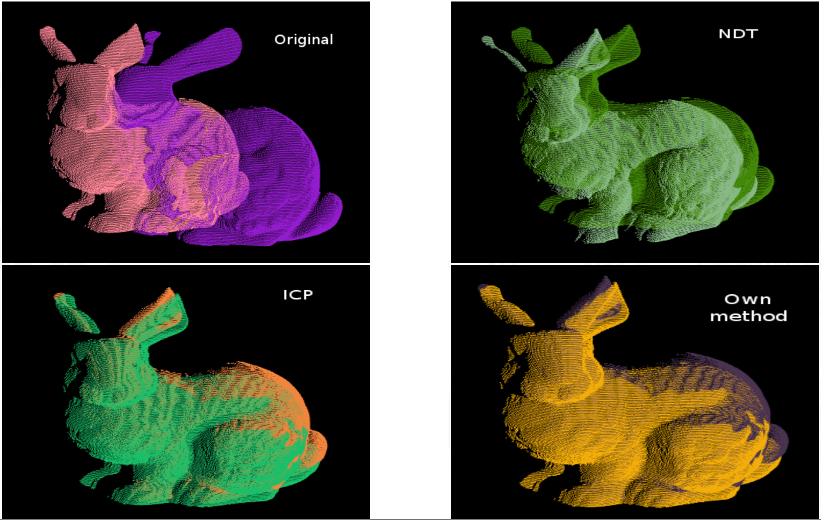
$$
\begin{aligned}
S(u, v, k) &= \frac{\mathcal{F}_1(u, v, k)^*}{|\mathcal{F}_1(u, v, k)|} \bullet \frac{\mathcal{F}_2(u, v, k)}{|\mathcal{F}_2(u, v, k)|} \\
s(x, y, z) &= \mathcal{F}^{-1}\{S(u, v, k)\} \\
(x_p, y_p, z_p) &= \arg\max_{(x,y,z)} s(x, y, z)
\end{aligned}
$$

Refinement

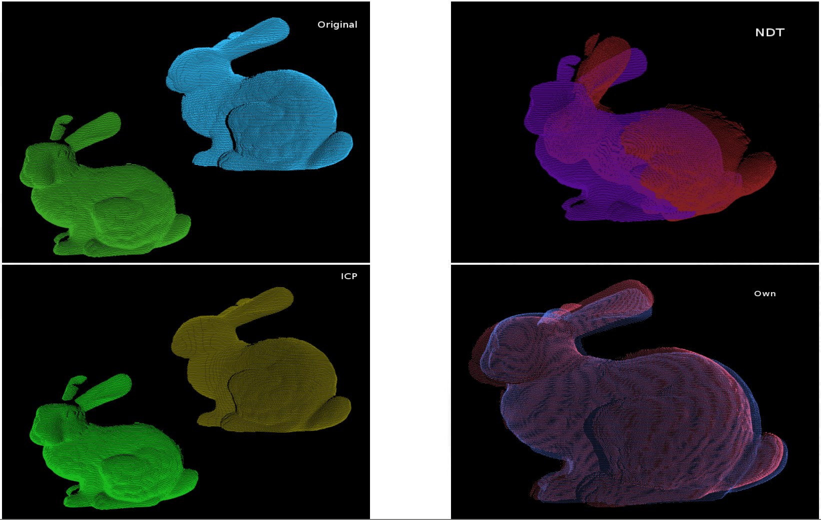
- Almost all the global methods need the refinement
- Normally the refinement is achieved by local methods
- the runtime of local methods depend on the crude alignment

- Global Methods: Bad $\Longrightarrow$ Ok
- Local Methods: Ok $\Longrightarrow$ Good

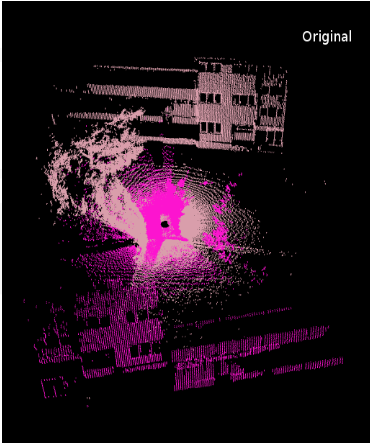
Bunny

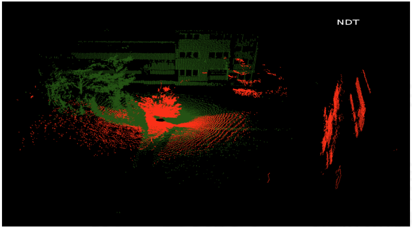


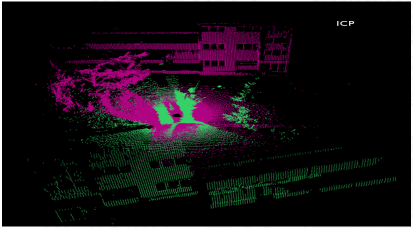
Bunny(large translation)



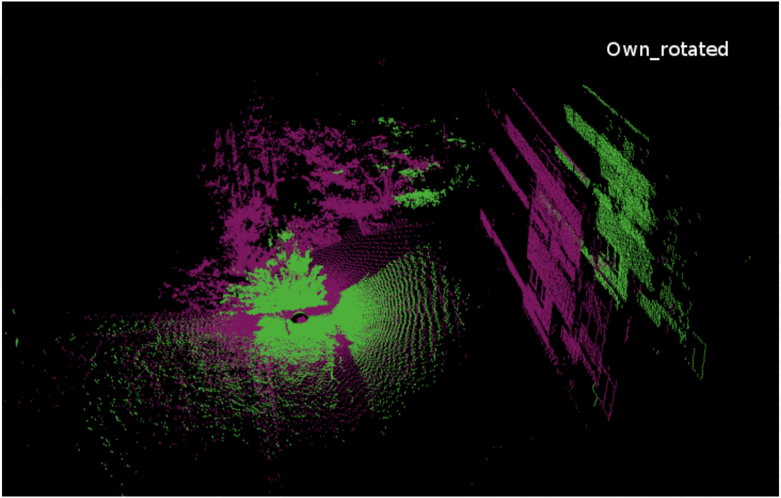
Campus



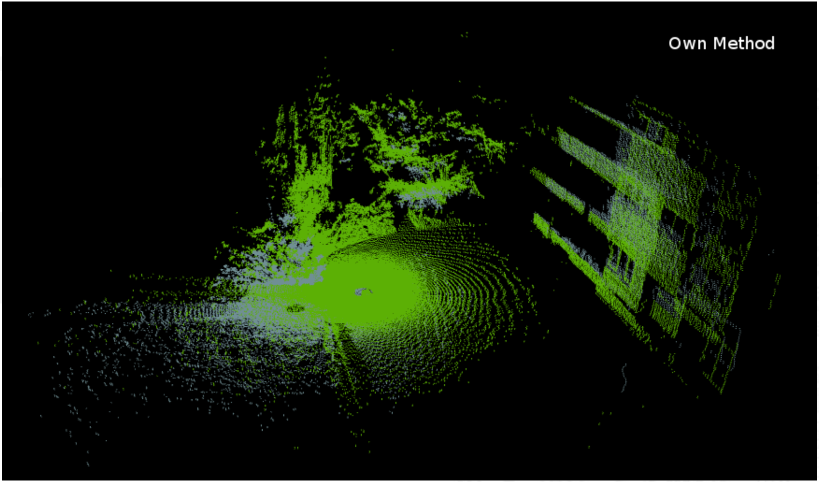




Campus



Campus



Conclusion

- Novel scan registration framework
- Novel Structural Representation (Spherical Entropy Images)
- Spherical Fourier Transform & Generalized Convolution Theorem

Future Work

- refine the translation estimation procedure
- more valid Spherical Structural Representation
- capture more data
- more experiments

. . . Thanks

Bo Sun

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics

Technical Aspects of Multimodal Systems

Iterative Closest Point (ICP)

Main components

- point clouds can be filtered, for example, to remove redundant points or compute descriptors like normals.
- a matching function needs to be applied to associate elements from a reading point cloud to a reference point cloud. This association is usually done in the Euclidean space using kd-tree to accelerate the search.
- remove mismatches or outliers
- the remaining points can be used to minimize the alignment error.

Comments

- requires initial estimate
- easy to get trapped in local minima
- When ICP is applied to robotics, special care needs to be taken to properly handle mismatches or outliers, like removing the higher-distance quantile of all paired points.

Normal Distribution Transform (NDT)

Main components

- subdivide the space occupied by the scan into a grid of cells
- a PDF is computed for each cell, based on the point distribution within the cell
- the PDF in each cell can be interpreted as a generative process for surface points within the cell
- when using NDT for scan registration, the goal is to find the pose of the current scan that maximizes the likelihood that the points of the current scan lie on the reference scan surface.

Comments

- The NDT can be described as a method for compactly representing a surface.
- The transform maps a point cloud to a smooth surface representation, described as a set of local PDFs, each of which describes the shape of a section of the surface.