Optimization by Gaussian Smoothing with Application to Geometric Alignment

PhD Final Exam
Hossein Mobahi
Advised by Prof. Yi Ma
Department of Computer Science
University of Illinois at Urbana-Champaign

Outline

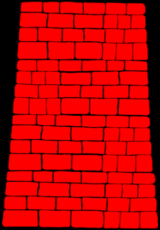
- Contributions
- Optimization & Homotopy
- Smoothing
- Asymptotic Analysis
- Transformation Kernels
- Application to 2D Alignment
- Application to 3D Alignment
- Conclusion

Contributions

- A Formal Framework for Asymptotic Analysis
- Easy to Check Condition for Asymptotic Convexity
- Simple Form for Asymptotic Minimizer
- Closed Form Kernels for Efficiently Smoothing Alignment Objective
- Formulation & Evaluation of 2D Alignment by Smoothing Method
- Formulation & Evaluation of 3D Point Cloud Alignment by Smoothing Method

Convex vs Nonconvex

Convex



Nonconvex

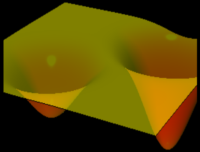
- Nonconvex optimization difficult in general
- Pressure to approximate by convex models
- Real world problems have regularity... may lead to tractable solutions
- Exploiting such structures: SOS, DC, smoothness, etc.

Homotopy Continuation

Smooth deformation of an easy problem into the actual problem, while tracing the solution

- Consider solving $\boldsymbol{f} : \mathcal{X} \to \mathcal{Y}$
- Embed $\boldsymbol{f}(\boldsymbol{x})$ into a parameterized family $\boldsymbol{g}(\boldsymbol{x}, t)$, where $\boldsymbol{g} : \mathcal{X} \times \mathcal{T} \to \mathcal{Y}$ and $\mathcal{T} = \{\, t \,|\, 0 \leq t \leq 1\}$
    - $\boldsymbol{g}(\boldsymbol{x}, 0)$ should be easy to solve
    - $\boldsymbol{g}(\boldsymbol{x}, 1) = \boldsymbol{f}(\boldsymbol{x})$

Homotopy Continuation

$$f(\boldsymbol{x}) \triangleq$$ 

$$k_\sigma(\boldsymbol{x}) \triangleq \frac{1}{(\sqrt{2\pi}\sigma)^{\dim(\boldsymbol{x})}} e^{-\frac{\|\boldsymbol{x}\|^2}{2\sigma^2}}$$

$$\boldsymbol{g}(\boldsymbol{x}, t) \triangleq t f(\boldsymbol{x}) + (1-t)(x^2 + y^2)$$

$$\boldsymbol{g}(\boldsymbol{x}, t) \triangleq [f \star k_{\frac{1}{t}-1}](\boldsymbol{x})$$

Where Does Smoothing Come From?

Describing Smoothing by PDEs

- Evolution of a function in a region $\mathcal{X}$ over time.
- Initial condition $g(\boldsymbol{x}, \sigma \rightarrow 0) = f(\boldsymbol{x})$.
- Boundary condition on $\mathcal{X}$ if any

Heat Equation

$$\frac{\partial}{\partial \sigma} g = \sigma \Delta g$$

Schrodinger's Equation

$$\frac{\partial}{\partial \sigma} g = i \sigma \Delta g$$

Where Does Smoothing Come From?

Why Heat Kernel for Optimization?

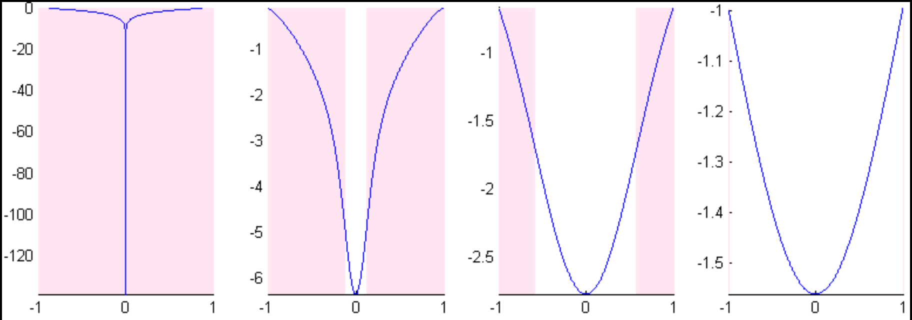
If $\mathcal{X} = \mathbb{R}^n$ and $k(\boldsymbol{x}; \sigma^2)$ is isotropic Gaussian, solution of heat equation is:

$$g(\boldsymbol{x}; \sigma) = [f \star k(\,.\,; \sigma^2)](\boldsymbol{x})\,.$$

Kills high frequencies, hence suppresses brittle local minima.

Consider $f(x) = e^{-\frac{t^2}{2\epsilon^2}} - e^{-\frac{t^2\epsilon^2}{2}}$ for $\epsilon > 0$

This functions resembles the ℓ_0 norm much better than ℓ_1. Except at its tip, it is concave everywhere. However, it is asymptotically convex!



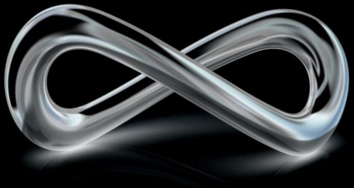
Popularity Across Disciplines

Known As

- Graduated Optimization (Computer Vision) [Blake & Zisserman 87]
- Optimization by Diffusion Equation (Chemistry) [Piela 89]
- Optimization by Homotopy Continuation (Numerical Computing) [Watson 88]
- Deterministic Annealing (Machine Learning) [Rose 98]

Despite its long age and mathematical roots, there is little understanding about its fundamental aspects.

Definition of Asymptotic Convexity

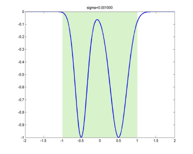


Definition

A real-valued continuous function $f(x)$ is called asymptotically convex if following statement holds:

$$\forall M > 0 \ , \ \exists \sigma^*(M) \ , \ \forall \boldsymbol{x_1} \in \mathcal{B}(\mathbf{0}, M) \ , \ \boldsymbol{x_2} \in \mathcal{B}(\mathbf{0}, M) \ , \ a \in [0,1]$$
$$\sigma \geq \sigma^*(M) \Rightarrow g(a\boldsymbol{x_1} + (1-a)\boldsymbol{x_2}; \sigma) \leq a g(\boldsymbol{x_1}; \sigma) + (1-a) g(\boldsymbol{x_2}; \sigma)$$

Asymptotic Convexity Example

For $f(x) =$ 

we have $M = 1, \sigma^* \approx 0.9$

Theoretical Contributions

Asymptotic Convexity

Under mild conditions, any function $f(\boldsymbol{x})$ satisfying $-\infty < \int_{\mathbb{R}^n} f(\boldsymbol{x})\, d\boldsymbol{x} < 0$ is asymptotically convex.

Asymptotic Minimizer

Under mild conditions, any function $f(\boldsymbol{x})$ that is asymptotically convex with $\int_{\mathbb{R}^n} f(\boldsymbol{x}) \neq 0$ has the asymptotic minimizer at the center of mass:

$$\boldsymbol{x}^* = \frac{\int_{\mathbb{R}^n} \boldsymbol{x} f(\boldsymbol{x})\, d\boldsymbol{x}}{\int_{\mathbb{R}^n} f(\boldsymbol{x})\, d\boldsymbol{x}}$$

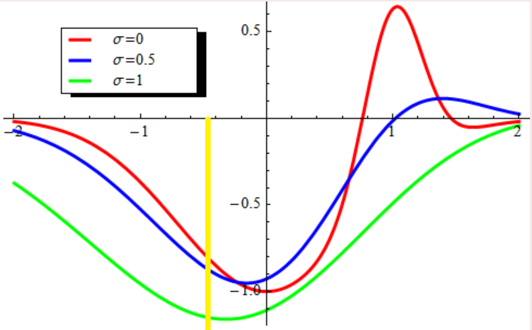
Nice Property

Both conditions are “derivative-free”.

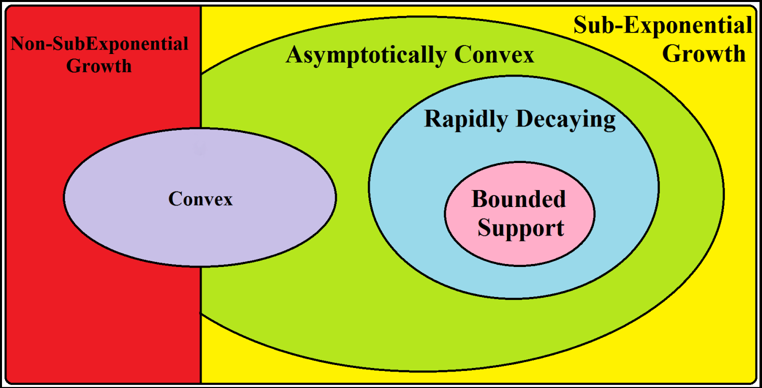
Example

Show that $f(x) = e^{-\frac{(x-1)^2}{0.1}} - e^{-x^2}$ is asymptotically convex and find its asymptotic minimizer.

$\int_{\mathbb{R}} f(x)\, dx \approx -1.21195$, thus $f(x)$ is asymptotically convex. Also $x^* = \frac{\int_{\mathbb{R}} x f(x) dx}{\int_{\mathbb{R}} f(x)\, dx} \approx -0.46247$.



Taxonomy for Functions $\{f : \mathbb{R}^n \to \mathbb{R}\}$



Applications to Image Alignment

Alignment as optimization

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \int_{\mathcal{X}} \Big(f_1(\boldsymbol{\tau}(\boldsymbol{x};\boldsymbol{\theta})) - f_2(\boldsymbol{x}) \Big)^2 d\boldsymbol{x}$$

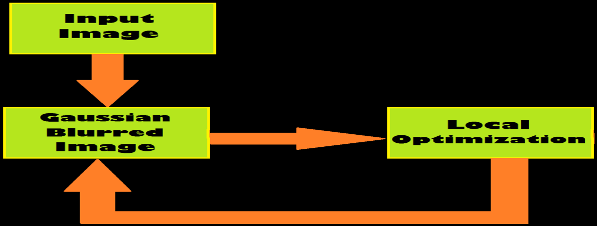
This is non-convex in variable $\boldsymbol{\theta}$.

Linearization

Linearize $f_1(\boldsymbol{x}+\boldsymbol{d})$ around $\boldsymbol{d}=\boldsymbol{0}$ to get a convex quadratic.

$$\hat{d} = \arg\min_{\boldsymbol{d}} \int_{\mathcal{X}} \Big(f_1(\boldsymbol{x}) + \boldsymbol{d}^T \nabla f_1(\boldsymbol{x}) - f_2(\boldsymbol{x}) \Big)^2 d\boldsymbol{x}$$

$$|\,\hat{f}_1(\boldsymbol{x}) - f_1(\boldsymbol{x})\,| \;\leq \frac{\Lambda}{2} \|\boldsymbol{d}\|^2$$



Motion Models other than Translation

Problem of Lucas-Kanade for Other Models

Consider scaling transformation $\boldsymbol{x} \to s\boldsymbol{x}$. Linearization of $f_1(s\boldsymbol{x})$ around $s = 1$:

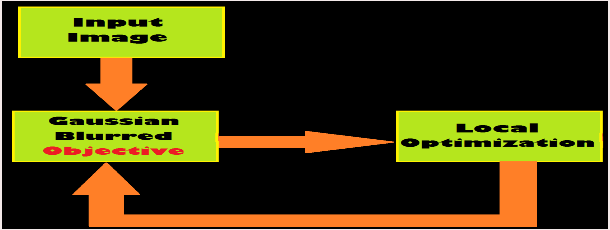
$$\hat{f}_1(s\boldsymbol{x}) = f_1(\boldsymbol{x}) + (\boldsymbol{s} - 1)(\boldsymbol{x}^T \nabla f_1(\boldsymbol{x}))$$

$$\hat{s} = \arg\min_{\boldsymbol{s}} \int_{\mathcal{X}} \Big(f_1(\boldsymbol{x}) + (\boldsymbol{s} - 1)(\boldsymbol{x}^T \nabla f_1(\boldsymbol{x})) - f_2(\boldsymbol{x}) \Big)^2 d\boldsymbol{x}$$

$$| \hat{f}_1(s\boldsymbol{x}) - f_1(s\boldsymbol{x}) | \leq \frac{\Lambda(s-1)^2}{2} \|\boldsymbol{x}\|^2$$

Error grows in $\|\boldsymbol{x}\|^2$ as well!

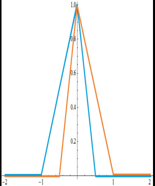
Smoothing the Objective



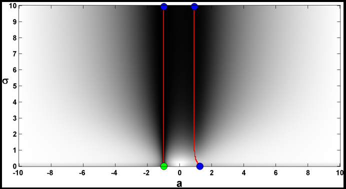
- Smoothing the objective automatically produces spatially varying blurs without any hack.
- People in computer vision have realized the advantage of spatially varying kernels for matching on heuristic basis, e.g. (Berg & Malik 2001)

Toy Example: 1-d Scale Alignment

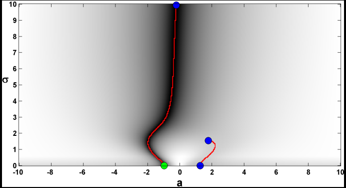
- **Actual Task** $\int_{\mathbb{R}} (f_1(ax) - f_2(x))^2 \mathbb{I}_{\|\boldsymbol{x}\| \leq 1} dx$
- **Signal Smoothing** $\int_{\mathcal{X}} \Big(\mathbb{I}_{\|\boldsymbol{x}\| \leq 1} \big([f_1 \overset{x}{\circledast} k_\sigma](ax) - [f_2 \overset{x}{\circledast} k_\sigma](x) \big) \Big)^2 dx$
- **Objective Smoothing**
  $[k_\sigma \overset{\ominus}{\circledast} \int_{\mathcal{X}} \Big(\mathbb{I}_{\|\boldsymbol{x}\| \leq 1} \big(f_1(. \times x) - f_2(x) \big) \Big)^2 dx](a)$



Signals



Signal Smoothing



Objective Smoothing

Efficient Computation of Smoothed Objective

Question

- Assume $\mathcal{X} = \mathbb{R}^n$ and $\Theta = \mathbb{R}^m$
- Given a domain transformation $\boldsymbol{\tau} : \mathcal{X} \times \Theta \to \mathcal{X}$.

Is there any $u_{\boldsymbol{\tau},\sigma} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ satisfying the following integral equation?

$$\forall f \quad : \quad [f(\boldsymbol{\tau}(\boldsymbol{x}, \cdot)) \star k(\cdot; \sigma^2)]\,(\boldsymbol{\theta}) = \int_{\mathcal{X}} f(\boldsymbol{y})\, u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y}$$

Applied Contributions

We do find such u and call it a transformation kernel.

Kernels for Common Transformations

Name	$\boldsymbol{\tau}(\boldsymbol{x}, \boldsymbol{\theta})$	$u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{y})$
Translation	$\boldsymbol{x} + \boldsymbol{d}$	$k(\boldsymbol{\tau}(\boldsymbol{x}, \boldsymbol{\theta}) - \boldsymbol{y}; \sigma^2)$
Translation+Scale	$\boldsymbol{a}^T \boldsymbol{x} + \boldsymbol{d}$	$K(\boldsymbol{\tau}(\boldsymbol{x}, \boldsymbol{\theta}) - \boldsymbol{y}; \sigma^2 \operatorname{diag}([1 + x_i^2]))$
Affine	$\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$	$k(\boldsymbol{\tau}(\boldsymbol{x}, \boldsymbol{\theta}) - \boldsymbol{y}; \sigma^2(1 + \|\boldsymbol{x}\|^2))$
Homography	$\frac{1}{1+\boldsymbol{c}^T\boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$	$n = 2 : p(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{y}, \sigma)\ e(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{y}, \sigma)$

Correctness

Name	$\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})$	$u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y})$
Translation	$\boldsymbol{x}+\boldsymbol{d}$	$k(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})-\boldsymbol{y};\sigma^2)$
Translation+Scale	$\boldsymbol{a}^T\boldsymbol{x}+\boldsymbol{d}$	$K(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})-\boldsymbol{y};\sigma^2\,\mathrm{diag}([1+x_i^2]))$
Affine	$\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}$	$k(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})-\boldsymbol{y};\sigma^2(1+\|\boldsymbol{x}\|^2))$
Homography	$\frac{1}{1+\boldsymbol{c}^T\boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b})$	$n=2: p(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y},\sigma)\ e(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y},\sigma)$

Heat Equation

$$\sigma\Delta_{\boldsymbol{\theta}}[f(\boldsymbol{\tau}(\boldsymbol{x},\,\cdot\,))\star k(\,\cdot\,;\sigma)](\boldsymbol{\theta})=(\partial/\partial\sigma)[f(\boldsymbol{\tau}(\boldsymbol{x},\,\cdot\,))\star k(\,\cdot\,;\sigma)](\boldsymbol{\theta})$$

Initial Condition

$$\lim_{\sigma\to 0^+}\int_{\mathcal{X}} f(\boldsymbol{y})u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y}=f(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta}))$$

Comparison against Geometric Blur

Name	$\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})$	$u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y})$
Translation	$\boldsymbol{x}+\boldsymbol{d}$	$k(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})-\boldsymbol{y};\sigma^2)$
Translation+Scale	$\boldsymbol{a}^T\boldsymbol{x}+\boldsymbol{d}$	$K(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})-\boldsymbol{y};\sigma^2\,\mathrm{diag}([1+x_i^2]))$
Affine	$\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}$	$k(\boldsymbol{\tau}(\boldsymbol{x},\boldsymbol{\theta})-\boldsymbol{y};\sigma^2(1+\|\boldsymbol{x}\|^2))$
Homography	$\frac{1}{1+\boldsymbol{c}^T\boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b})$	$n=2:p(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y},\sigma)\ e(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y},\sigma)$

Geometric Blur Kernel (Berg & Malik 2001)

$$u_\sigma(\boldsymbol{x},\boldsymbol{y})=k(\boldsymbol{y}-\boldsymbol{x};\sigma^2\|\boldsymbol{x}\|^2)$$

- Heuristic
- Identity Transformation
- Singular at Origin

Alignment By Smoothing Algorithm

Original Objective

$$h(\boldsymbol{\theta}) \triangleq \int_{\mathcal{X}} f_1(\boldsymbol{\tau}(\boldsymbol{x}, \boldsymbol{\theta}))\, f_2(\boldsymbol{x})\, d\boldsymbol{x}$$

Smoothed Objective

$$\begin{aligned} z(\boldsymbol{\theta}, \sigma) &= \int_{\mathcal{X}} \big(f_2(\boldsymbol{x}) [f_1(\boldsymbol{\tau}(\boldsymbol{x}, .)) \star k(\,\cdot\,, \sigma^2)](\boldsymbol{\theta}) \big)\, d\boldsymbol{x} \\ &= \underbrace{\int_{\mathcal{X}} \Bigg(f_2(\boldsymbol{x}) \underbrace{\Big(\int_{\mathcal{X}} f_1(\boldsymbol{y}) u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y} \Big)}_{\text{Blurred } f_1} \Bigg) d\boldsymbol{x}}_{\text{Inner Product}} \end{aligned}$$

Computation with Homography Kernel

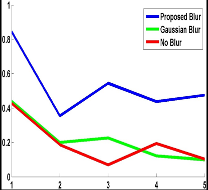
Approximation

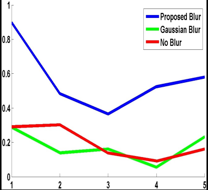
- Model image as piecewise constant so that,

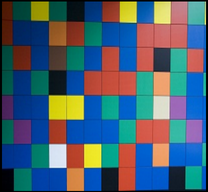
$$\int_{\mathcal{X}} f_1(\boldsymbol{y}) u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y})\, d\boldsymbol{y} = \sum_{i=1}^{W}\sum_{j=1}^{H} F_1(i,j) \int_{\underline{y_i}}^{\overline{y_i}} \int_{\underline{y_j}}^{\overline{y_j}} u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y})\, d\boldsymbol{y}\,.$$

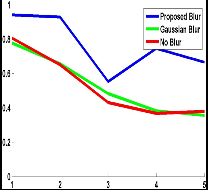
- Now computing the integral transform amounts to computing $\int_{\mathcal{X}_{ij}} u_{\boldsymbol{\tau},\sigma}(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{y})\, d\boldsymbol{y}$.
- We use Laplace approximation to compute integral above.

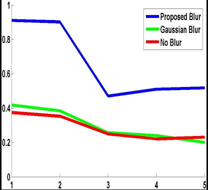
Quantitative 2D Alignment Results

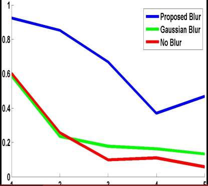












Qualitative 2D Alignment Results

A 3D Reconstruction Scenario

- In reality, a scene barely consists of a single planar surface
- A real example: 3D reconstruction of an octagonal building
  - Given eight uncalibrated and widely separated images
  - Each image covers a pair of adjacent facades
  - Segment an image into piecewise planar regions
  - Rectify a segment by a single homography
  - Dense pixel-wise match across pairs of facades
  - Bundle adjustment

Qualitative 2D Alignment Results

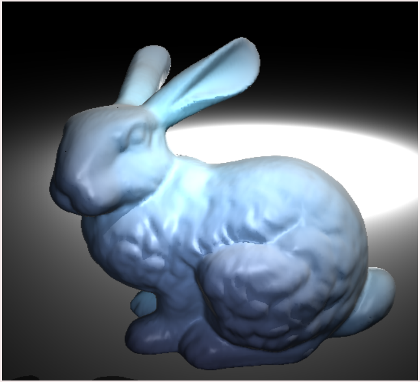
A 3D Reconstruction Scenario

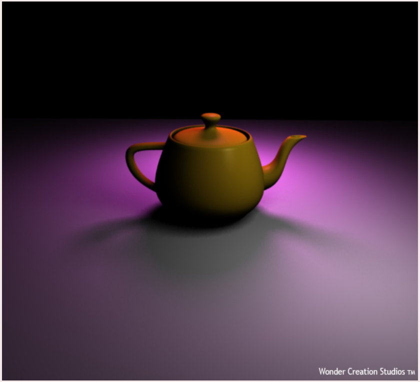
- Each segmented region may not share the same location and scale across images
- Refine their location and scale by alignment in scale+displacement space
- Use smoothed objective of normalized cross correlation (NCC)

Qualitative 2D Alignment Results

A 3D Reconstruction Scenario

3D Point Cloud Alignment





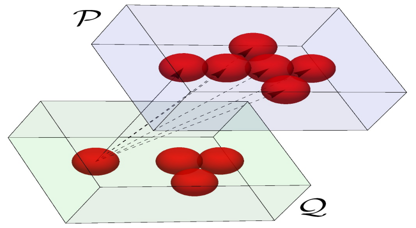
3D Point Cloud Alignment

Formulation

$$(\boldsymbol{\theta}^*, \boldsymbol{c}^*) = \arg\min_{\boldsymbol{\theta}, \boldsymbol{c}} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} \|\boldsymbol{\tau}(\boldsymbol{p}_i, \boldsymbol{\theta}) - \boldsymbol{q}_j\|^2$$

s.t. $\forall j \in \{1, \ldots, n\} \quad \sum_{i=1}^{m} c_{i,j} = 1$

$\forall i \in \{1, \ldots, m\} \forall j \in \{1, \ldots, n\} \quad c_{i,j} \in \{0, 1\}$



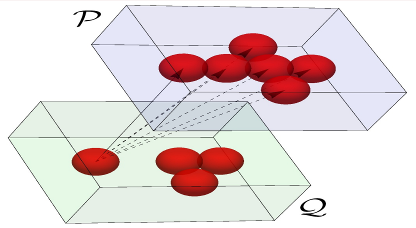
3D Point Cloud Alignment

Formulation

$$(\boldsymbol{\theta}^*, \boldsymbol{c}^*) = \arg\min_{\boldsymbol{\theta}, \boldsymbol{c}} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} \|\boldsymbol{\tau}(\boldsymbol{p}_i, \boldsymbol{\theta}) - \boldsymbol{q}_j\|^2$$

s.t. $\forall\, j \in \{1, \ldots, n\} \quad \sum_{i=1}^{m} c_{i,j} = 1$

$\forall\, i \in \{1, \ldots, m\} \forall\, j \in \{1, \ldots, n\} \quad c_{i,j}(1 - c_{i,j}) = 0$



3D Point Cloud Alignment

Approximate by Quadratic Penalty

$$
\begin{aligned}
(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{c}}) &= \arg\min_{\boldsymbol{\theta}, \boldsymbol{c}} h(\boldsymbol{\theta}, \{c_{i,j}\}) \\
h(\boldsymbol{\theta}, \boldsymbol{c}) &\triangleq \epsilon\Big(\sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\|\boldsymbol{\tau}(\boldsymbol{p}_i, \boldsymbol{\theta}) - \boldsymbol{q}_j\|^2\Big) \\
&\quad + \sum_{j=1}^{n}(1 - \sum_{i=1}^{m} c_{i,j})^2 + \sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}^2(1 - c_{i,j})^2
\end{aligned}
$$

ϵ is a small positive number.

3D Point Cloud Alignment

Smoothed Objective

$$
\begin{aligned}
z(\boldsymbol{\theta}, \boldsymbol{c}; \sigma) &\triangleq \Big[\Big(\big[h(\,.\,,\,.\,) \star k(\,.\,;\sigma^2)\big]\,(\boldsymbol{c})\Big) \star k(\,.\,;\sigma^2)\Big]\,(\boldsymbol{\theta}) \\
&= \epsilon \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} \big(\|\boldsymbol{\tau}(\boldsymbol{p}_i, \boldsymbol{\theta}) - \boldsymbol{q}_j\|^2 + 3\sigma^2(1 + \|\boldsymbol{p}_i\|^2)\big) \\
&\quad + \sum_{j=1}^{n} (1 - \sum_{i=1}^{m} c_{i,j})^2 \\
&\quad + \sum_{i=1}^{m} \sum_{j=1}^{n} (c_{i,j} - 1)^2 c_{i,j}^2 + 6\sigma^2 (c_{i,j} - \frac{1}{2})^2
\end{aligned}
$$

We assume $\boldsymbol{\tau}$ is an affine transformation, i.e. $\boldsymbol{\tau}\big(\boldsymbol{p}; (\boldsymbol{A}, \boldsymbol{b})\big) \triangleq \boldsymbol{A}\boldsymbol{p} + \boldsymbol{b}$

3D Point Cloud Alignment

Asymptotic Minimizer

- The objective is asymptotically convex.
- Its asymptotic minimizer is the simple form $\boldsymbol{A}^* = \boldsymbol{O}$, $\boldsymbol{b}^* = \mathbf{0}$, and $\boldsymbol{c}^* = \frac{1}{2}$.

3D Point Cloud Alignment

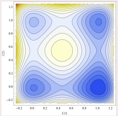
Illustrative Example

- Set $d = 1$, $\epsilon = 0.01$, $\mathcal{P} = \{p_1, p_2\}$, and $\mathcal{Q} = \{q_1, q_2\}$, where $p_1 = q_1 = -1$ and $p_2 = q_2 = 1$.
- The points are already aligned, thus fix $\boldsymbol{A}$ and $\boldsymbol{b}$ to their optimal values $\boldsymbol{A} = \boldsymbol{I}$ and $\boldsymbol{b} = \boldsymbol{0}$.
- Fix $c_{1,2}$ and $c_{1,2}$ to their optimal value $c^*_{1,2} = 1 - c_{1,1}$ and $c^*_{2,2} = 1 - c_{2,1}$.
- This leaves us with only two variables $c_{1,1}$ and $c_{2,1}$, whose optimal solution must be $c^*_{1,1} = 1$ and $c^*_{2,1} = 0$.

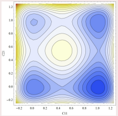
3D Point Cloud Alignment

Optimization Landscape of Smoothed Objective

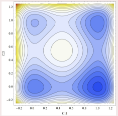
- At the non-smoothed function ($\sigma = 0$), global minimum is near $(c_{1,1}, c_{2,1}) = (1, 0)$ and three local minima near $(0, 1)$, $(0, 0)$ and $(1, 1)$.
- For large enough σ, the landscape becomes convex, with minimizer around $(c_{1,1}, c_{2,1}) = (\frac{1}{2}, \frac{1}{2})$, as anticipated by the asymptotic minimizer result.
- The path of minimizer form large to small σ converges to the global minimum



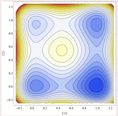
$\sigma = 0$



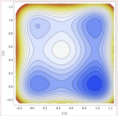
$\sigma = \frac{3}{88}$



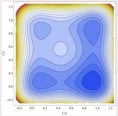
$\sigma = \frac{6}{88}$



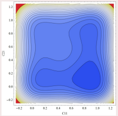
$\sigma = \frac{9}{88}$



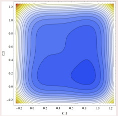
$\sigma = \frac{12}{88}$



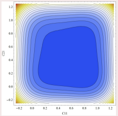
$\sigma = \frac{15}{88}$



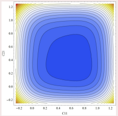
$\sigma = \frac{27}{88}$



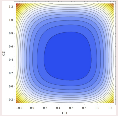
$\sigma = \frac{30}{88}$



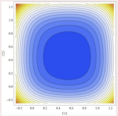
$\sigma = \frac{33}{88}$



$\sigma = \frac{18}{88}$



$\sigma = \frac{21}{88}$

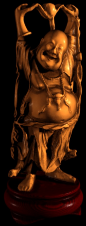


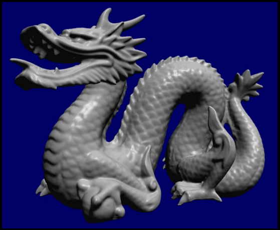
$\sigma = \frac{24}{88}$

3D Point Cloud Alignment

Qualitative Results

- Using point cloud data of some objects in Stanford 3D.
- Comparison against Iterative Closest Point (ICP) algorithm; that is alternation between:
  - Given transformation, establish correspondence between pair of points (of the two clouds)
  - Given correspondence, optimization the alignment transformation between the point clouds.

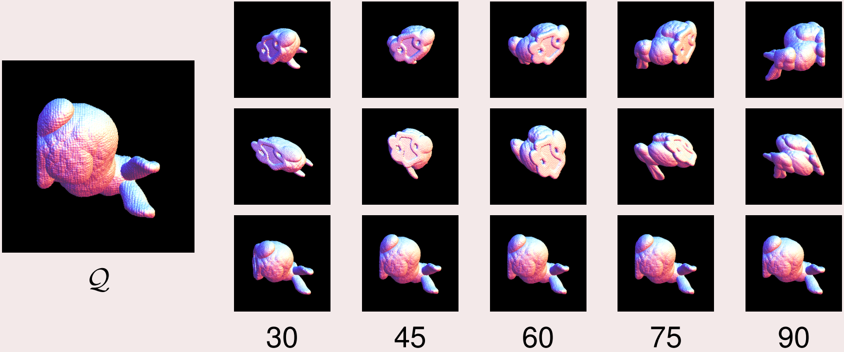




3D Point Cloud Alignment

Qualitative Results

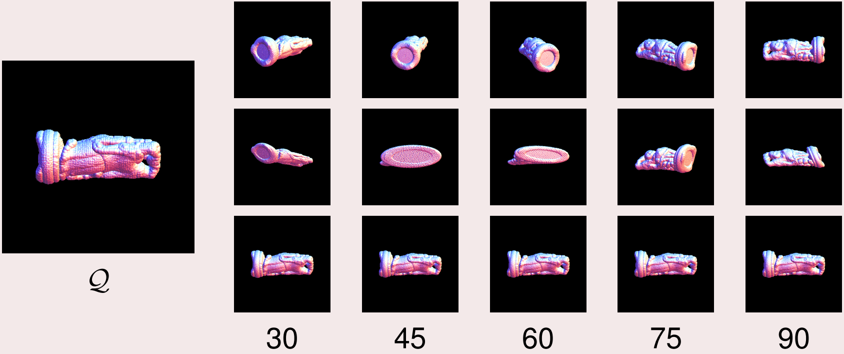
- Top: Input $\mathcal{P}$, which is a rotated version of $\mathcal{Q}$.
- Middle: Transformed $\mathcal{P}$ to match $\mathcal{Q}$ using ICP.
- Bottom Row: Transformed $\mathcal{P}$ to match $\mathcal{Q}$ using proposed method.



3D Point Cloud Alignment

Qualitative Results

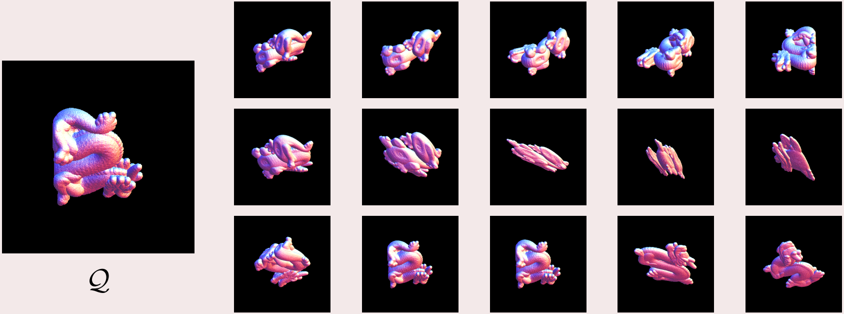
- Top: Input $\mathcal{P}$, which is a rotated version of $\mathcal{Q}$.
- Middle: Transformed $\mathcal{P}$ to match $\mathcal{Q}$ using ICP.
- Bottom Row: Transformed $\mathcal{P}$ to match $\mathcal{Q}$ using proposed method.



3D Point Cloud Alignment

Qualitative Results

- Top: Input $\mathcal{P}$, which is a rotated version of $\mathcal{Q}$.
- Middle: Transformed $\mathcal{P}$ to match $\mathcal{Q}$ using ICP.
- Bottom Row: Transformed $\mathcal{P}$ to match $\mathcal{Q}$ using proposed method.



Conclusion & Future Directions

Optimization by Smoothing

- Contributions
  - Rigorous definitions for asymptotic convexity.
  - Derivative free test for asymptotic convexity.
  - Derivative free form for asymptotic minimizer.
- Future Directions
  - Asymptotic analysis of non-Gaussian smoothing.
  - Conditions that guarantee a traceable path.
  - Conditions that guarantee reaching the global minimum.

Conclusion & Future Directions

Alignment

- Contributions
  - Showing that traditional Gaussian image blurring (e.g. Lucas-Kanade) is not suitable for non-displacement motions.
  - Derivation of spatially varying kernels required for objective smoothing.
  - Formulation of 2D and 3D alignment via objective smoothing.
- Future Directions
  - Exploring potential connections between our kernels and blur kernels for deblurring or motion from blur.
  - Heuristic spatially-varying kernels showed success in object detection & recognition. Our kernels may provide a principled framework for developing such kernels.
  - Exploiting smoothness and localized form of our kernels to compute integral transforms faster.


Thank you