Gaussian processes for non-rigid registration

-

Connections to medical image analysis

Marcel Lüthi

Graphics and Vision Research Group
Department of Mathematics and Computer Science
University of Basel

Aims of the talk

- Show how Analysis by Synthesis and Gaussian processes lead to a family of methods for non-rigid registration

- Provide an understanding of many common algorithms in terms of Gaussian processes

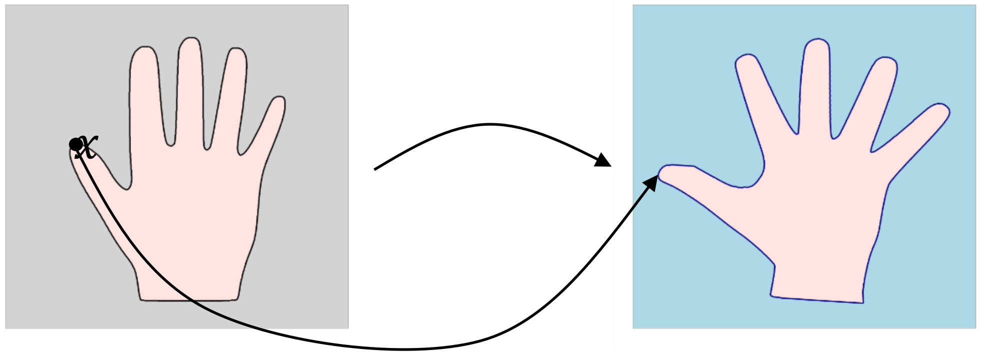
- Show how to derive new registration approaches using GPMMs and MCMC

Outline

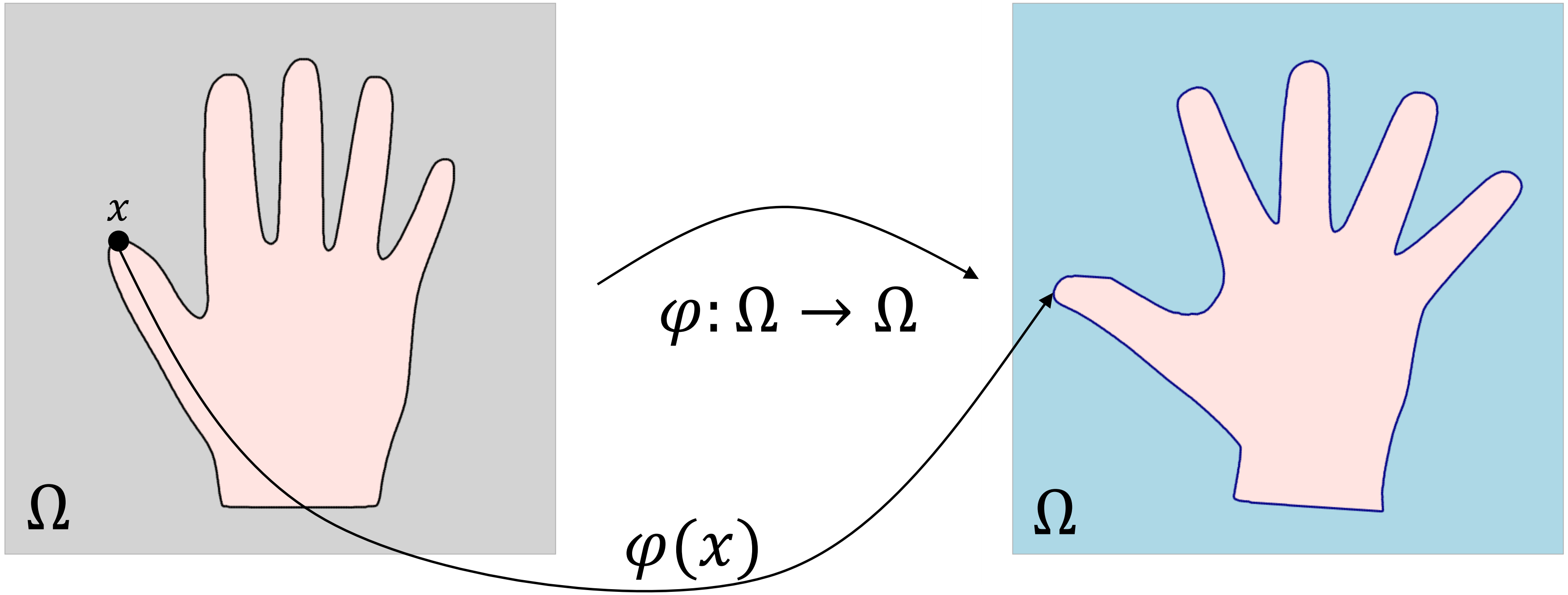
- The registration problem
  - Problem formulation
  - Registration as analysis by synthesis problem
  - An algorithm using Gaussian process priors
- Priors for registration
  - Spline-based models, Radial basis functions
  - Multis-scale and Spatially-varying models
  - Statistical deformation models
- Likelihood functions
  - Landmark registration
  - Image to image registration
  - Surface to image registration
- Advancing registration

Outline

- The registration problem
  - Problem formulation
  - Registration as analysis by synthesis problem
  - An algorithm using Gaussian process priors
- Priors for registration
  - Spline-based models
  - Radial basis functions
  - Statistical deformation models
- Likelihood functions
  - Landmark registration
  - Image to image registration
  - Surface to image registration
- Some ideas where to go from here



The registration problem



Reference:

$I_R: \Omega \to \mathbb{R}$

Target:

$I_T: \Omega \to \mathbb{R}$

The registration problem

Variational formulation

$$\varphi^* = \arg\min_{\varphi} D[I_R, I_T \circ \varphi] + \lambda R[\varphi]$$

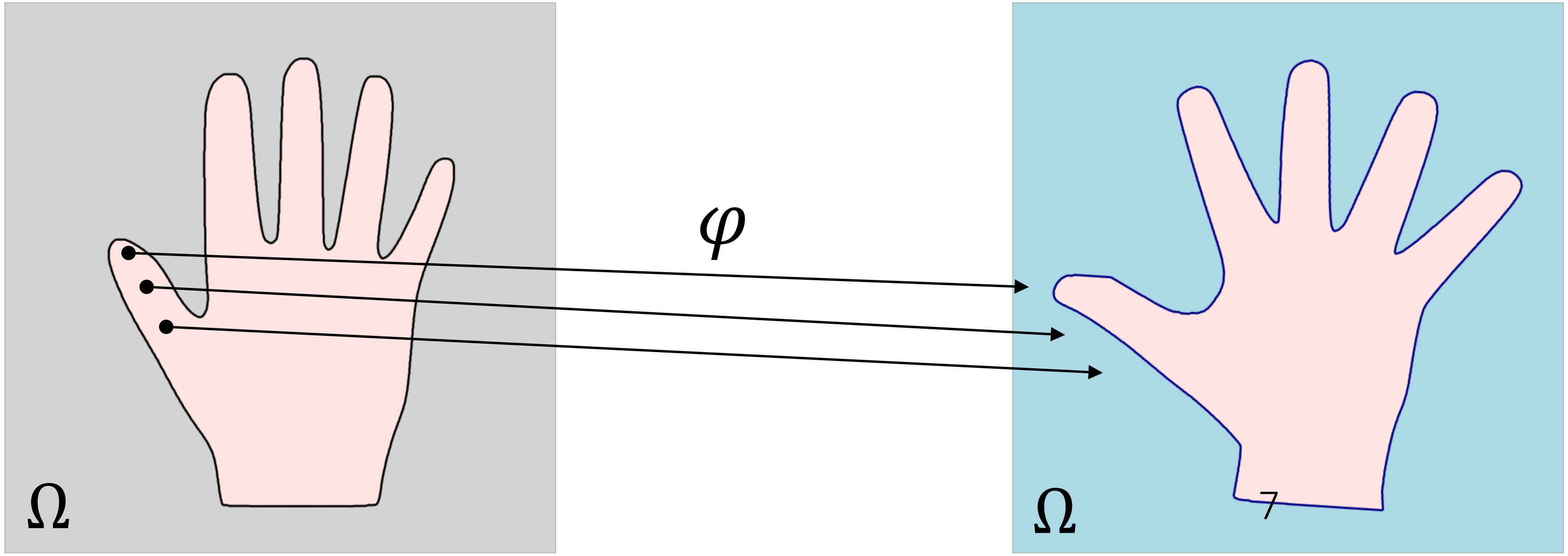
Mapping φ^* is trade-off that

- makes the images look similar (for similarity measure D)
- matches the prior assumptions (encoded by regularizer R)

The registration problem

Variational formulation

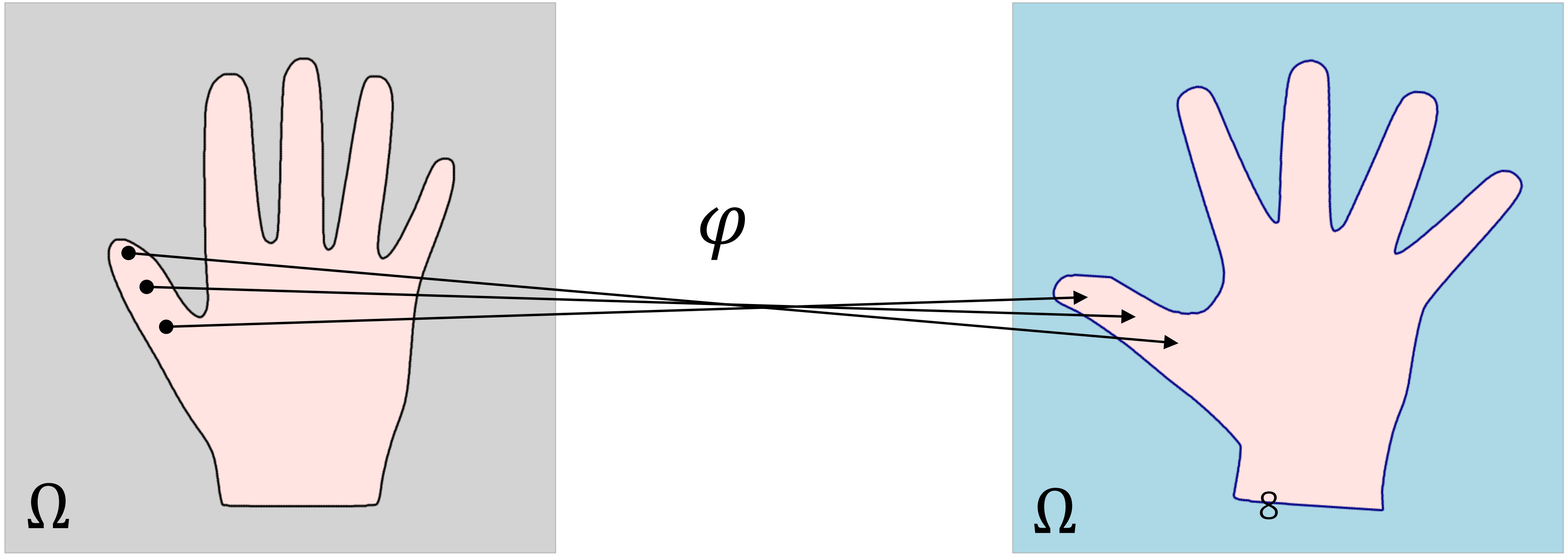
$$\varphi^* = \arg\min_{\varphi} D[I_R, I_T \circ \varphi] + \lambda R[\varphi]$$



The registration problem

Variational formulation

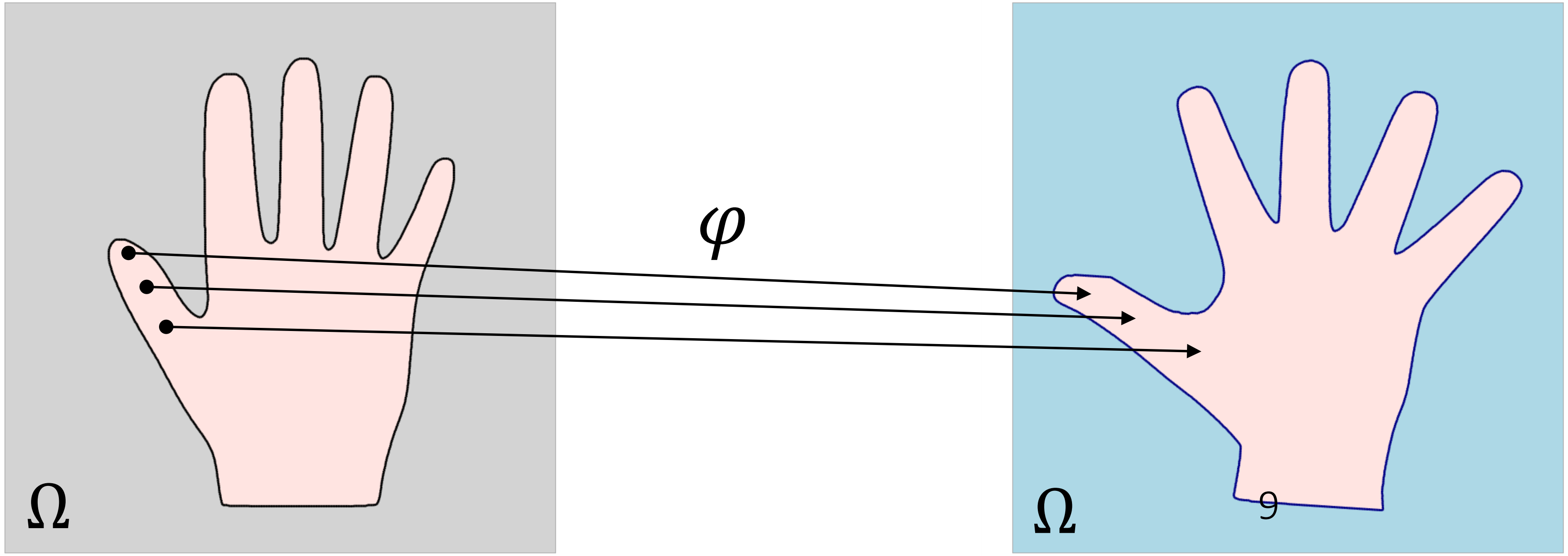
$$\varphi^* = \arg\min_{\varphi} D[I_R, I_T \circ \varphi] + \lambda R[\varphi]$$



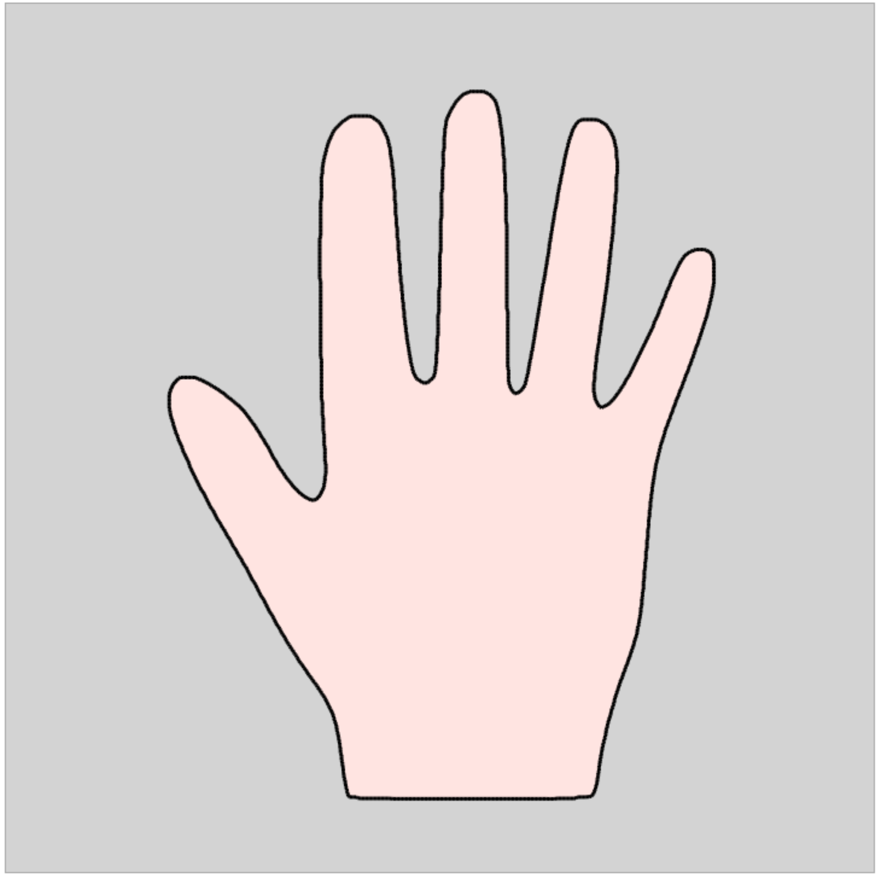
The registration problem

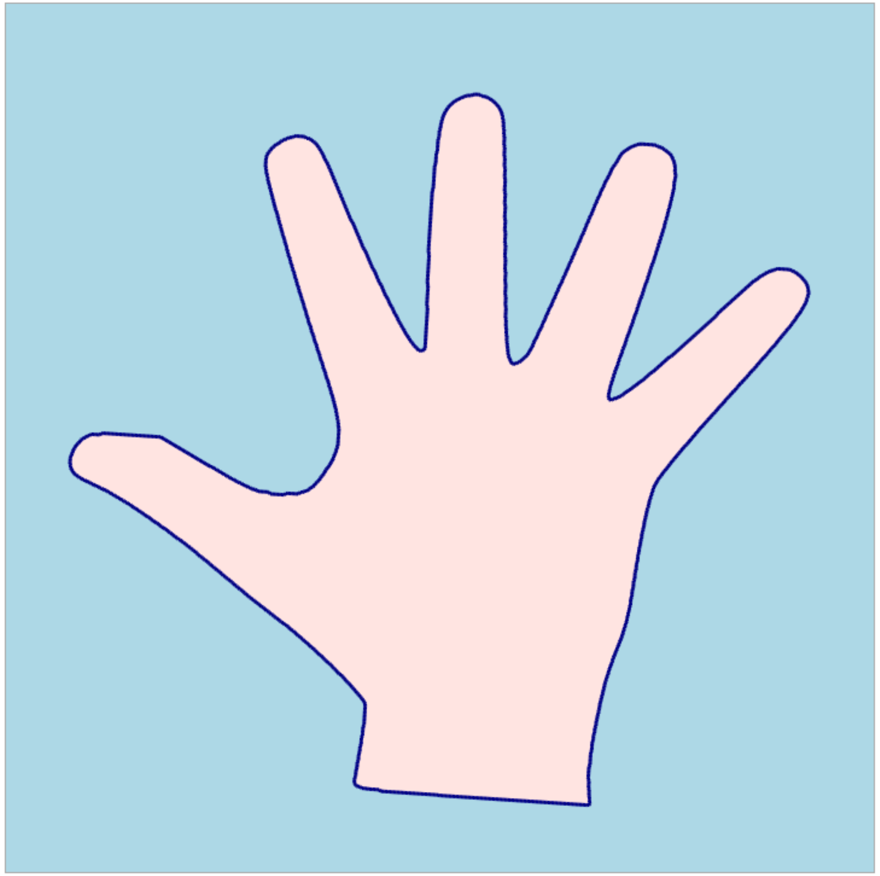
Variational formulation

$$\varphi^* = \arg\min_{\varphi} D[I_R, I_T \circ \varphi] + \lambda R[\varphi]$$

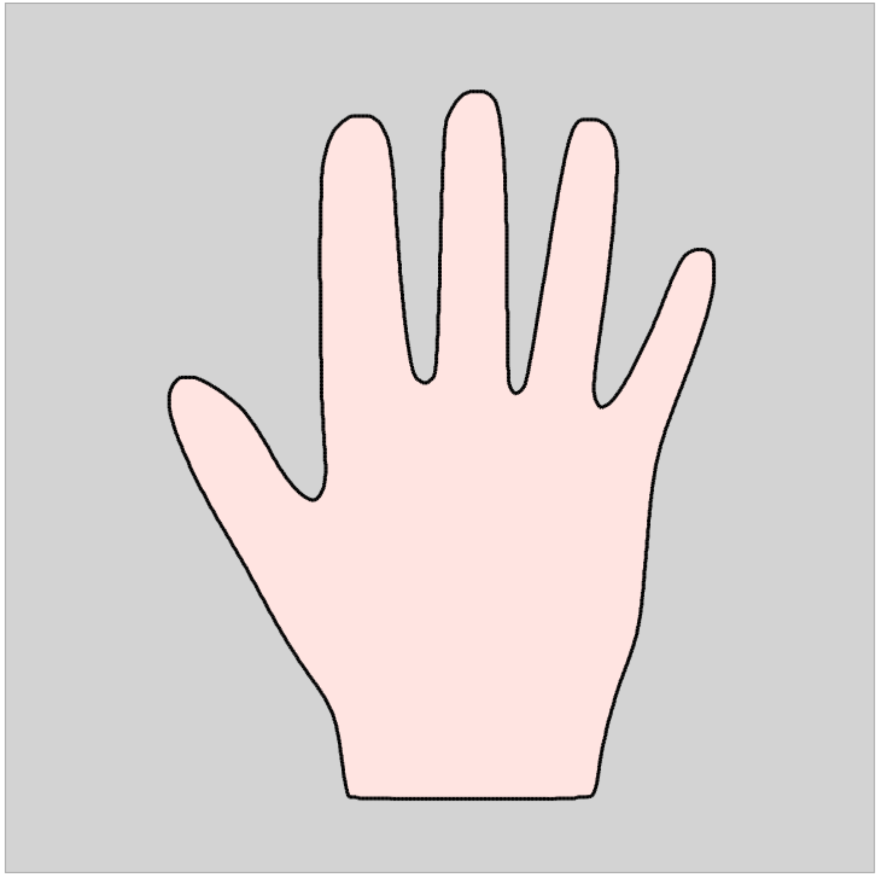


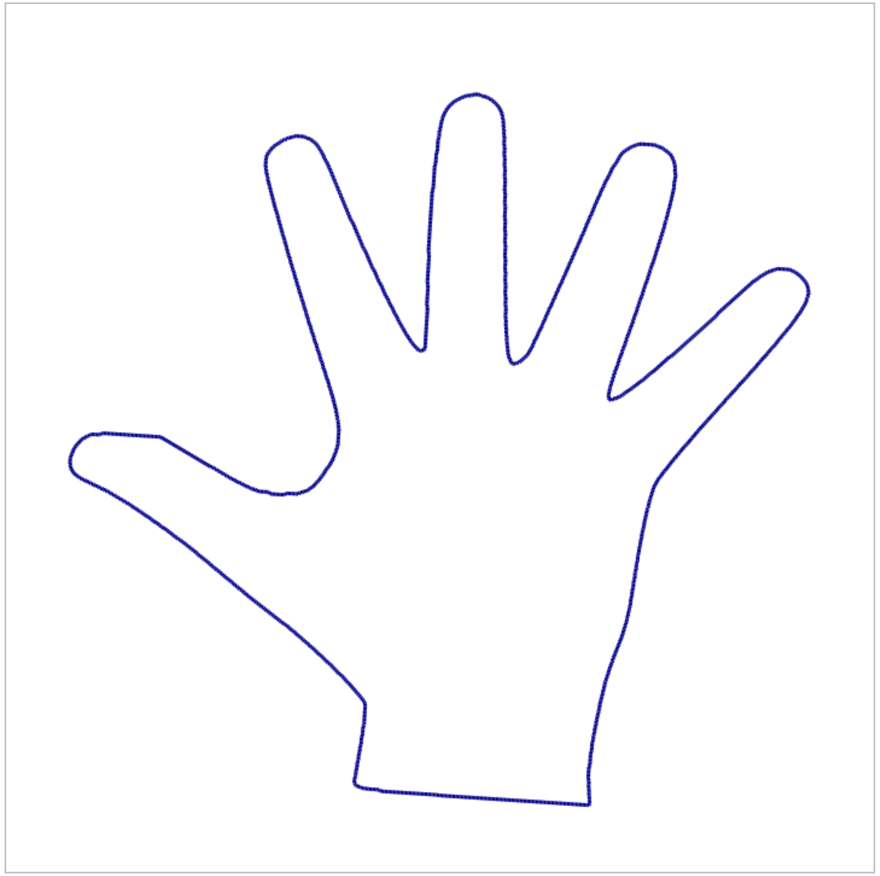
Representation of the mapping φ



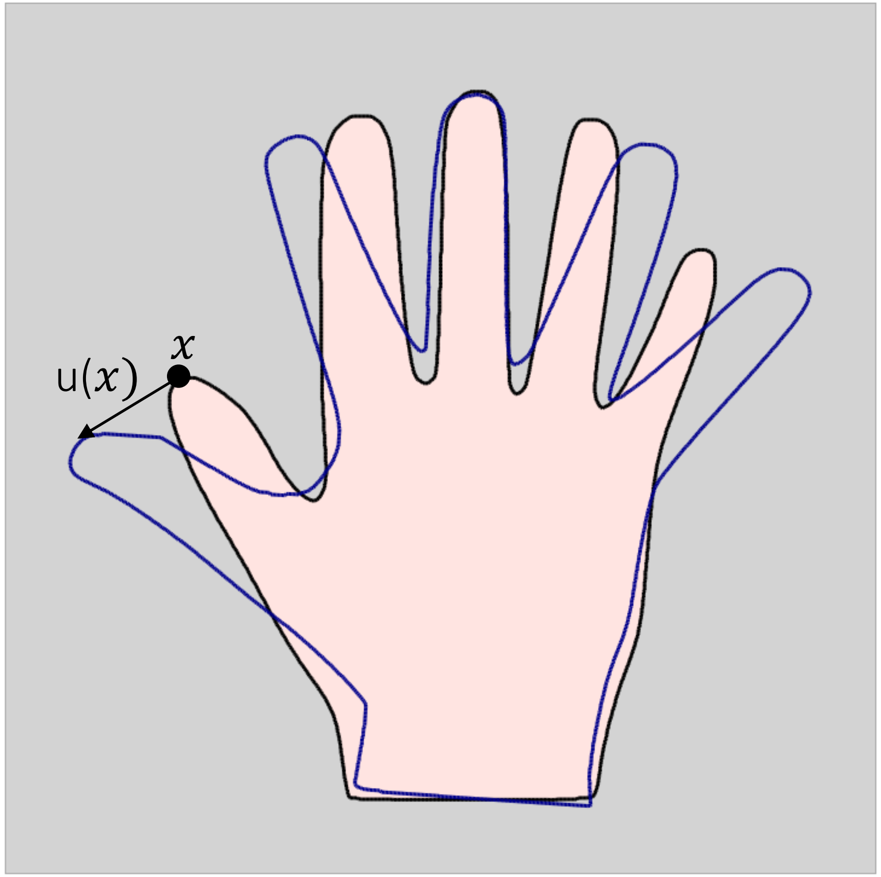


Representation of the mapping φ





Representation of the mapping φ

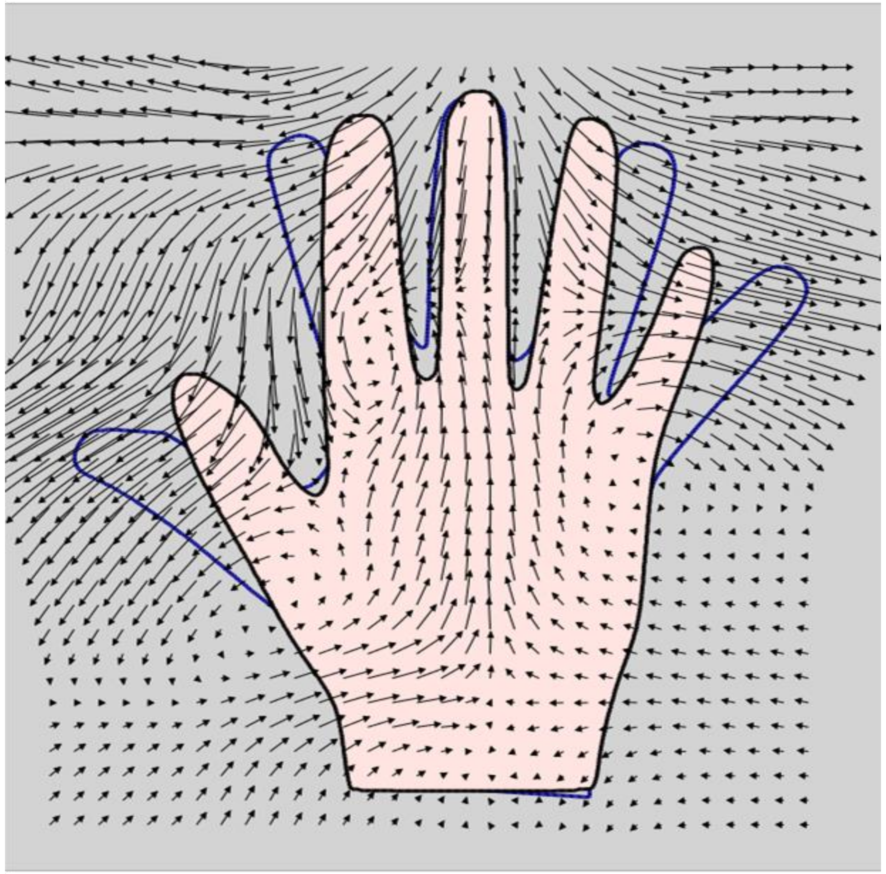


Assumption:
Images are rigidly aligned

- Mapping can be represented as a displacement vector field:

$$\varphi(x) = x + u(x)$$
$$u : \Omega \rightarrow \mathbb{R}^d$$

Representation of the mapping φ



Assumption:
Images are rigidly aligned

- Mapping can be represented as a displacement vector field:

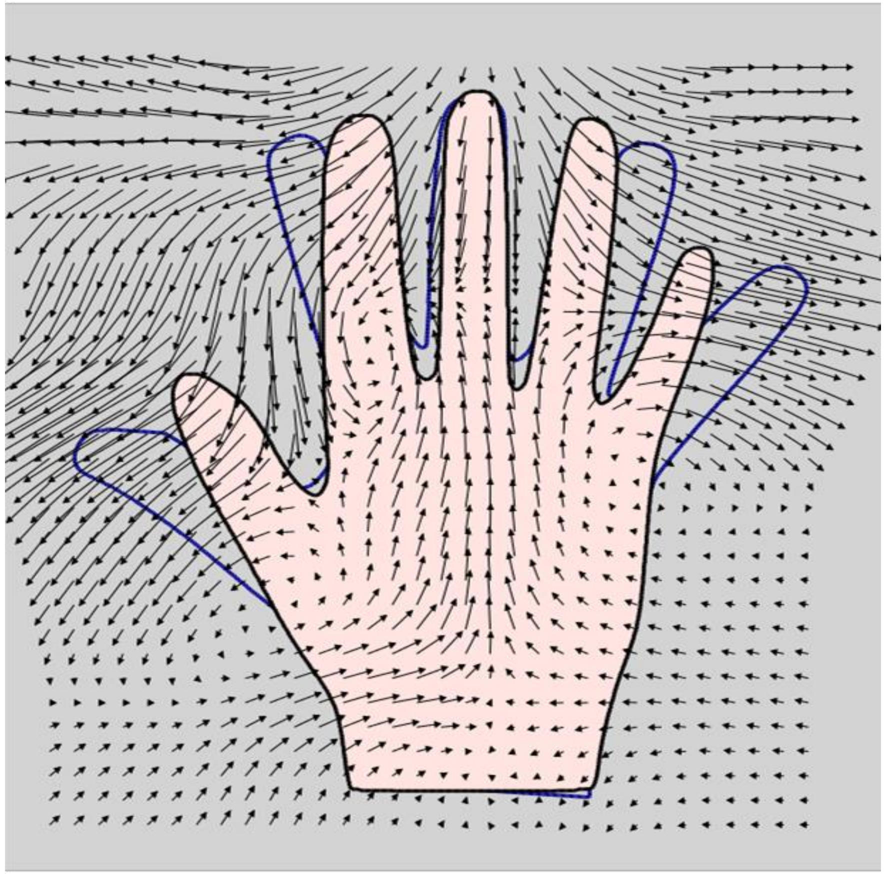
$$\varphi(x) = x + u(x)$$
$$u : \Omega \to \mathbb{R}^d$$

Further assumption:

- φ is parametric :

$$\varphi[\theta](x) = x + u[\theta](x)$$

Representation of the mapping φ



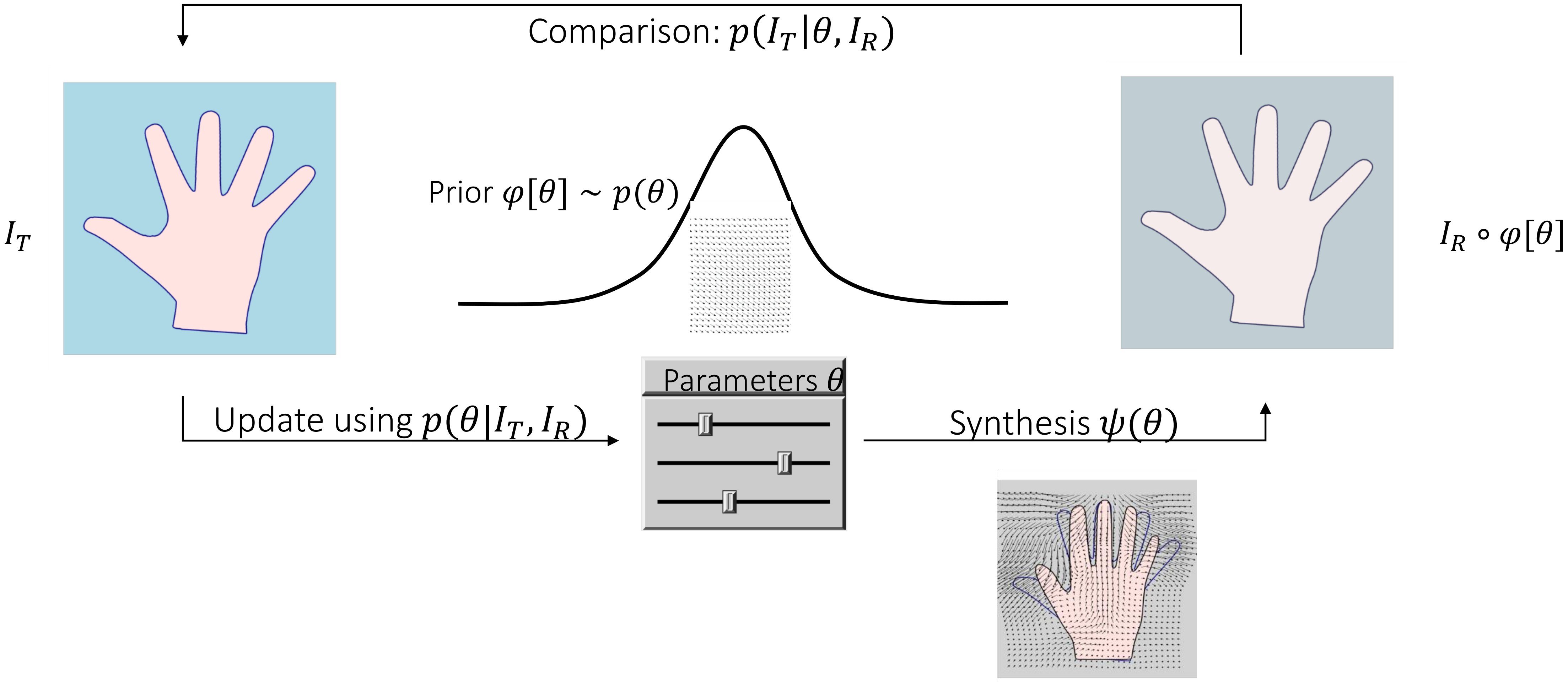
Mapping:

$$\varphi[\theta](x) = x + u[\theta](x)$$

Observation:

- Knowledge of θ and I_R allows us to synthesize target image I_T
  - (at least up to intensity differences)

Registration as analysis by synthesis



Probabilistic formulation of registration

Using Bayes rule: $P(\varphi[\theta]|I_T, I_R) = \frac{P(I_T|\varphi[\theta], I_R)P(\varphi[\theta], I_R)}{P(I_T)}$

MAP solution

$$\theta^* = \arg\max_\theta p(\varphi[\theta]|I_T, I_R) = \arg\max_\theta p(\varphi[\theta])p(I_T|I_R \circ \varphi[\theta])$$

Mapping θ^* is trade-off that defines a mapping $\varphi[\theta^*]$ which

- explains the data well (likelihood function)
- matches the prior assumptions (prior distribution)

Registration problem

$$\varphi^* = \arg\max_{\theta} p(\varphi[\theta]) p(I_T | I_R \circ \varphi[\theta])$$
$$= \arg\max_{\theta} \ln p(\varphi[\theta]) + \ln(p(I_T | I_R \circ \varphi[\theta]))$$
$$= \arg\min_{\theta} -\ln p(\varphi[\theta]) - \ln(p(I_T | I_R \circ \varphi[\theta]))$$

The registration problem

Variational formulation

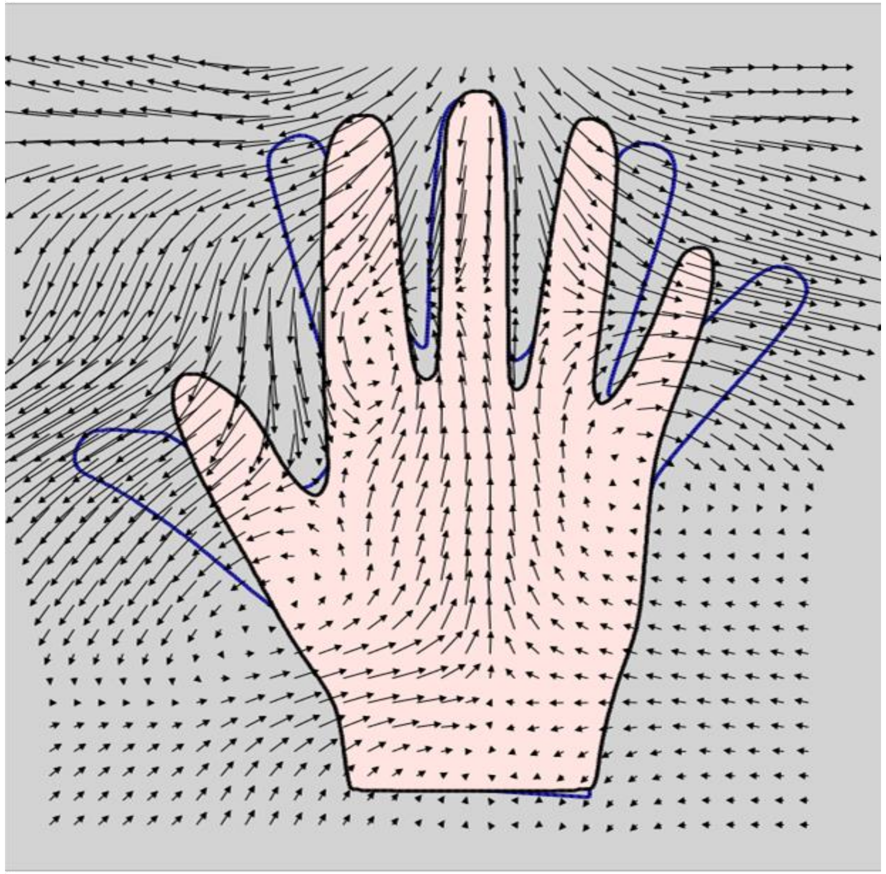
$$\varphi^* = \arg\min_{\varphi} D[I_T, I_R \circ \varphi] + \lambda R[\varphi]$$

Probabilistic formulation

$$\theta^* = \arg\min_{\theta} -\ln\left(p(I_T | I_R \circ \varphi[\theta])\right) - \ln p(\varphi[\theta])$$

Take home message: Registration is model fitting!!!

Gaussian processes



Define the Gaussian process

$$u \sim GP(\mu, k)$$

with mean function

$$\mu: \Omega \to \mathbb{R}^2$$

and covariance function

$$k: \Omega \times \Omega \to \mathbb{R}^{2\times 2} \ .$$

Parametric representation of Gaussian process

Represent GP using only the first r components of its KL-Expansion

$$u = \mu + \sum_{i=1}^{r} \alpha_i \sqrt{\lambda_i}\, \phi_i, \qquad \alpha_i \sim N(0,1)$$

- We have a finite, parametric representation of the process.
- We know the pdf for a deformation u

$$p(u) = p(\alpha) = \prod_{i=1}^{r} \frac{1}{\sqrt{2\pi}} \exp(-\alpha_i^2/2) = \frac{1}{Z} \exp(-\frac{1}{2} \|\alpha\|^2)$$

Registration problem

$$\varphi^* = \arg\min_\theta - \ln p(\varphi[\theta]) - \ln p(I_T | I_R \circ \varphi[\theta])$$

$$= \arg\min_\theta -\ln \frac{1}{Z} \exp(-\frac{1}{2} \|\theta\|^2) - \ln(p(I_T | I_R \circ \varphi[\theta])$$

$$= \arg\min_\theta -\ln \frac{1}{Z} + \frac{1}{2} \|\theta\|^2 - \ln(p(I_T | I_R \circ \varphi[\theta])$$

$$= \arg\min_\theta \frac{1}{2} \|\theta\|^2 - \ln(p(I_T | I_R \circ \varphi[\theta])$$

Summary: registration problem

$$\arg\min_{\theta} \frac{1}{2}\|\theta\|^2 + \ln(p(I_T | I_R \circ \varphi[\theta]))$$

- Variational and probabilistic formulation are closely related
  - Prior can be seen as regularizer
  - Likelihood term is an image similarity
- For a low-rank Gaussian process prior, the problem becomes parametric since

$$\varphi[\theta](x) = x + \mu(x) + \sum_{i=1}^{r} \theta_i \sqrt{\lambda_i}\, \phi_i(x)$$

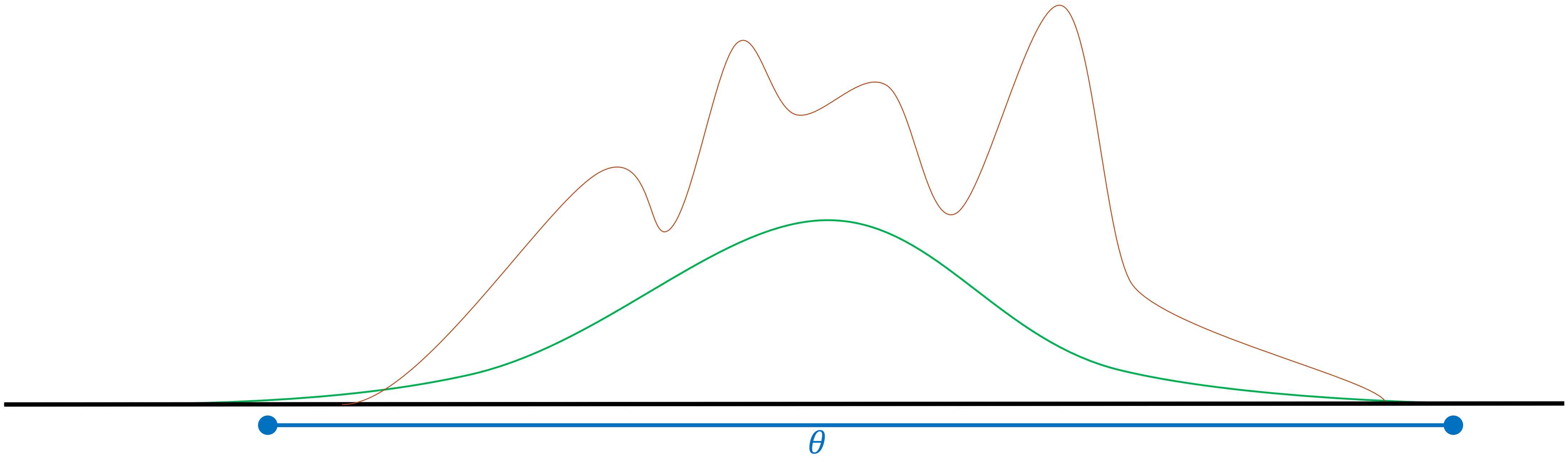
  - Can be optimized using gradient-descent schemes.
- All the regularization assumptions are encoded in the eigenfunctions ϕ_i

Outline

- The registration problem
  - Problem formulation
  - Registration as analysis by synthesis problem
  - An algorithm using Gaussian process priors
- Priors for registration
  - Spline-based models, Radial basis functions
  - Multis-scale and Spatially-varying models
  - Statistical deformation models
- Likelihood functions
  - Landmark registration
  - Image to image registration
  - Surface to image registration
- Advancing registration

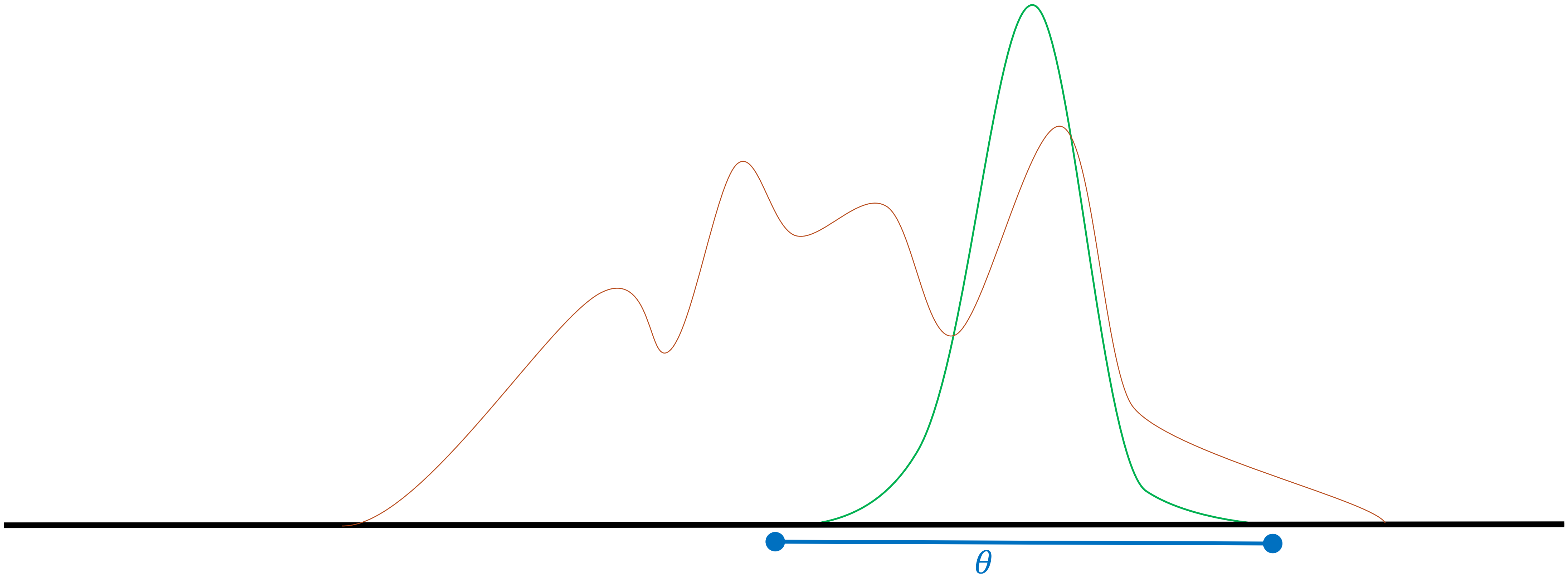
Why are priors interesting?

$$\theta^* = \arg\max_{\theta} p(\varphi[\theta])p(I_T|I_R \circ \varphi[\theta])$$



Why are priors interesting?

$$\theta^* = \arg\max_\theta p(\varphi[\theta])p(I_T|I_R \circ \varphi[\theta])$$



Defining a Gaussian process

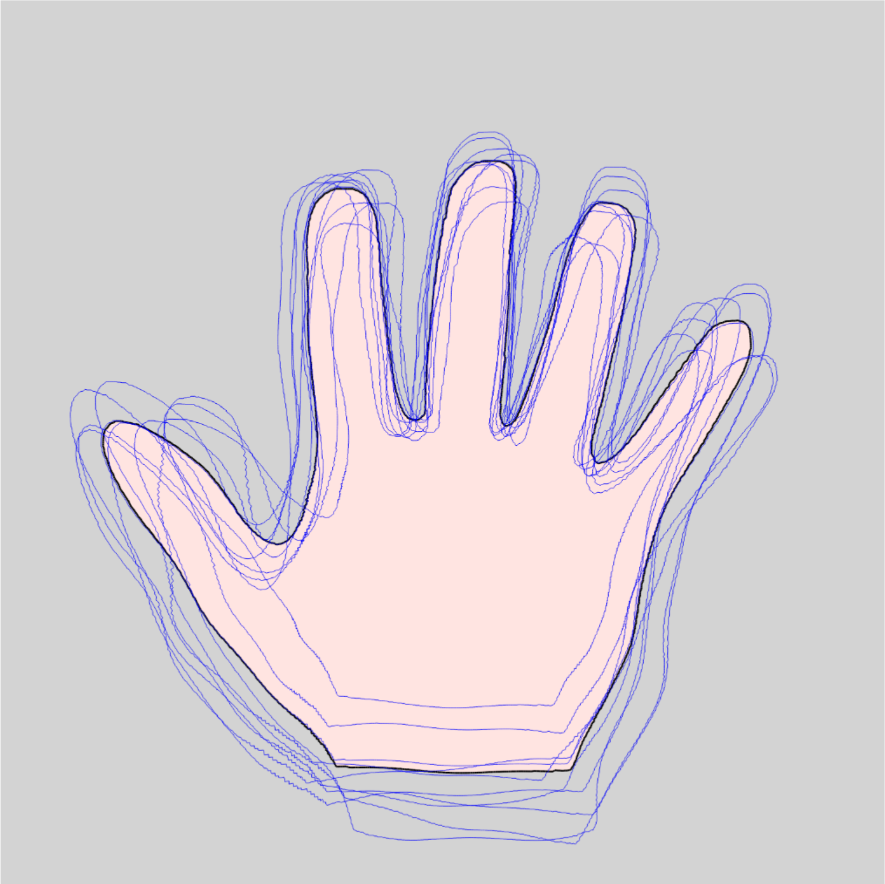
A Gaussian process

$$GP(\mu, k)$$

is completely specified by a mean function μ and covariance function (or kernel) k.

- $\mu: \Omega \to \mathbb{R}^d$ defines how the average deformation looks like
- $k: \Omega \times \Omega \to \mathbb{R}^{d \times d}$ defines how it can deviate from the mean
  - Must be positive semi-definite

The mean function



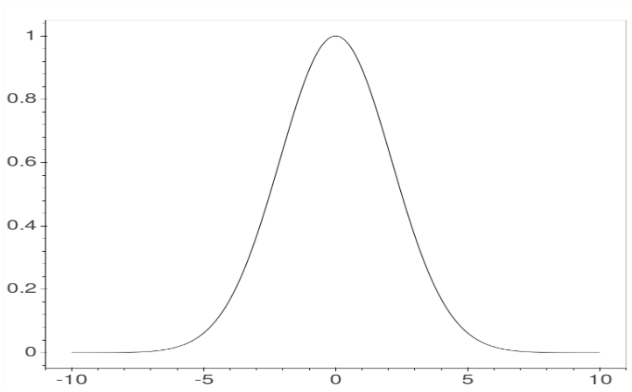
- Usual assumption:

$$\mu(x) = \begin{pmatrix} \mu_1(x) \\ \vdots \\ \mu_d(x) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

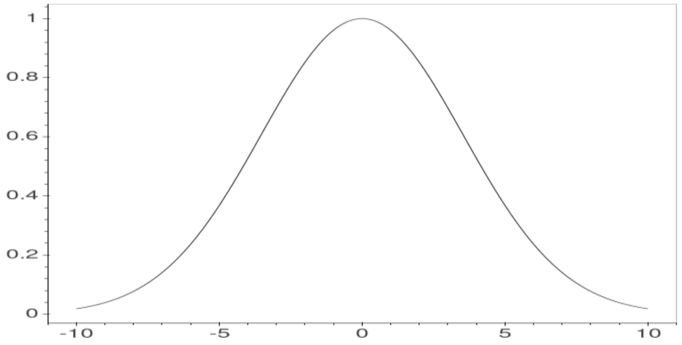
- The reference shape is an average shape.

Scalar-valued Gaussian kernel

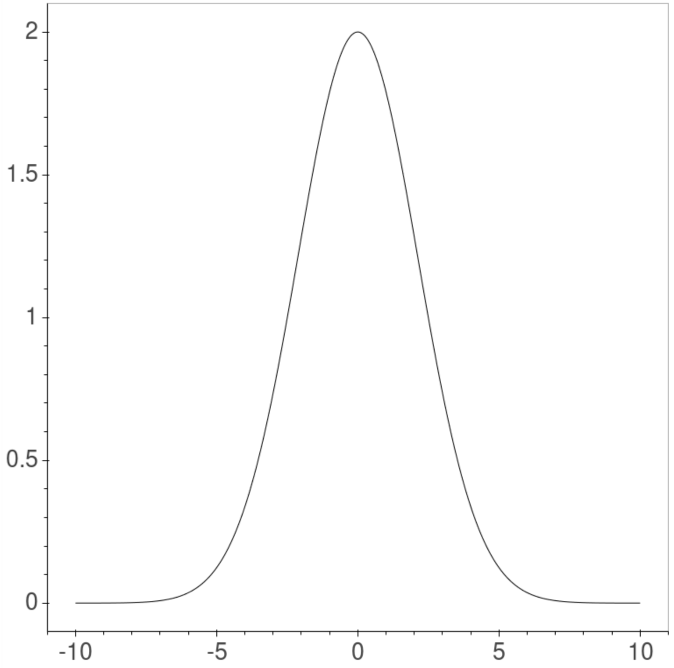
$$k(x, x') = s \exp(-\frac{\|x - x'\|^2}{\sigma^2})$$



$s = 1, \sigma = 3$



$s = 1, \sigma = 5$



$s = 2, \sigma = 3$

Diagonal kernel

$$k(x, x') = \begin{pmatrix} k^{(1)}(x, x') & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & k^{(d)}(x, x') \end{pmatrix}$$

- $k^{(1)}, \dots, k^{(d)}: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ are scalar-valued kernels
- $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{d \times d}$ becomes a matrix valued kernel.

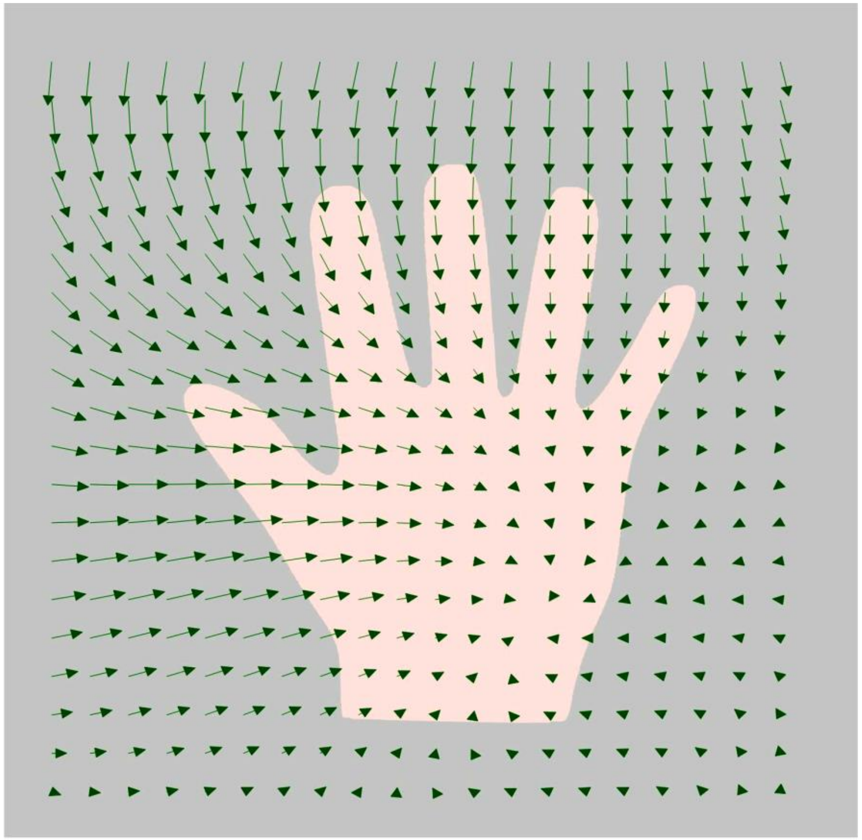
Assumption: Each dimension is modelled independently.

- the output-dimensions are uncorrelated.

A model for smooth 2D deformations

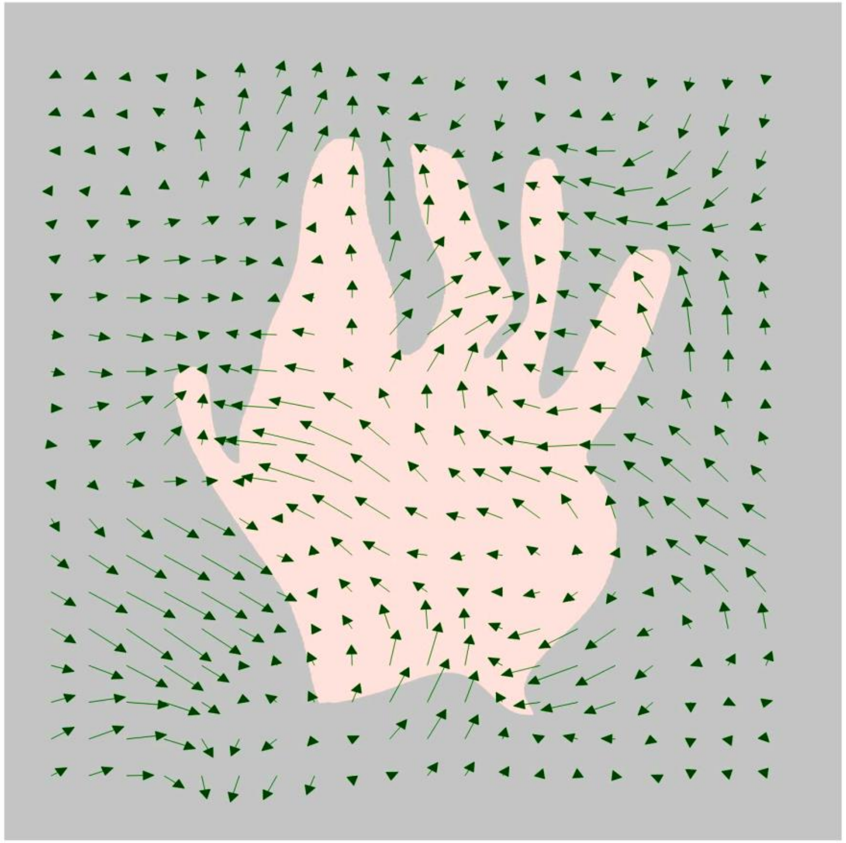
$$k(x, x') = \begin{pmatrix} s_1 \exp\left(-\frac{\|x - x'\|^2}{\sigma_1^2}\right) & 0 \\ 0 & s_2 \exp\left(-\frac{\|x - x'\|^2}{\sigma_2^2}\right) \end{pmatrix}$$

A model for smooth deformations



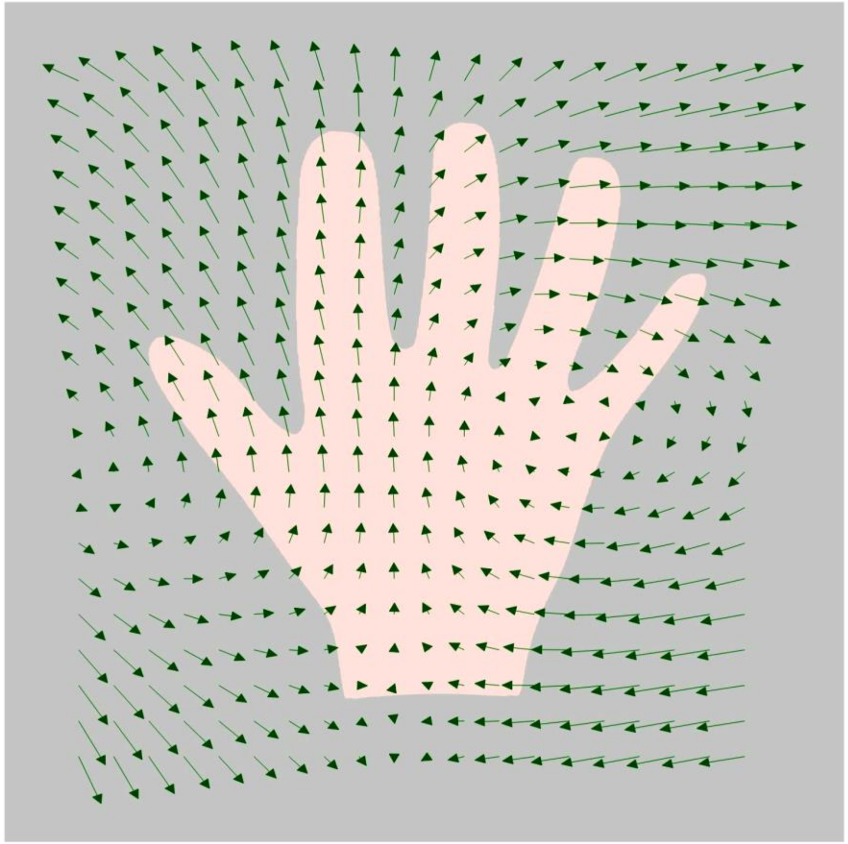
$s_1 = s_2$ small, $\sigma_1 = \sigma_2$ large

A model for smooth deformations



$s_1 = s_2$ small, $\sigma_1 = \sigma_2$ small

A model for smooth deformations



$s_1 = s_2$ large, $\sigma_1 = \sigma_2$ large

Matern class of kernels

$$k(x,x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(2\sqrt{2\nu} \frac{\|x-x'\|}{\rho} \right)^{\nu} K_\nu(\sqrt{2\nu} \frac{\|x-x'\|}{\rho})$$

- Γ is the Γ function, k_ν the modified Bessel function and ν, ρ are parameters
- The derivatives are $\nu - 1$ times differentiable

- Special cases:
  - $\nu = \frac{1}{2}$: $k(x,x') = \sigma^2 \exp(-\frac{\|x-x'\|}{\rho})$
  - $\nu = \frac{3}{2}$: $\mathrm{k}(\mathrm{x},\mathrm{x}') = \sigma^2 (1 + \frac{\sqrt{3}\|x-x'\|}{\rho}) \exp(-\frac{\sqrt{3}\|x-x'\|}{\rho})$
  - $\nu \to \infty$ Gaussian kernel

Thin-plate splines

- Minimize the bending energy of a metal sheet

$$R[\varphi] = \sum_{k=1}^{d} \int_{\Omega} \left(\nabla^T \nabla \varphi_k(x)\right)^2 dx$$

- Corresponding covariance function

$$k(x, x') = \frac{1}{12}(2\|x - x'\|^3 - 3R(\|x - x'\|^2 + R^3)$$
$$\text{where } R = \max_{x,x' \in \Omega} \|x - x'\|$$

Rohr, Karl, et al. "Landmark-based elastic registration using approximating thin-plate splines." *IEEE Transactions on medical imaging* 20.6 (2001): 526-534.

Elastic body splines

- Mechanical model of an elastic body or material
- Solution to the following PDE

$$\mu\nabla^2 u(x) + (\mu + \lambda)\nabla[\nabla \cdot u(x)] = c|x|$$

- Corresponding (matrix-valued) covariance function (may not be positive definite)

$$k(x, x') = (12(1 - \nu) - 1)|x|^2 I - 3xx^T$$

$$\text{where } \nu = \frac{\lambda}{2(\lambda + \mu)}$$

Kohlrausch, Jan, Karl Rohr, and H. Siegfried Stiehl. "A new class of elastic body splines for nonrigid registration of medical images." *Journal of Mathematical Imaging and Vision* 23.3 (2005): 253-280.

B-Splines

- Use B-Spline basis function

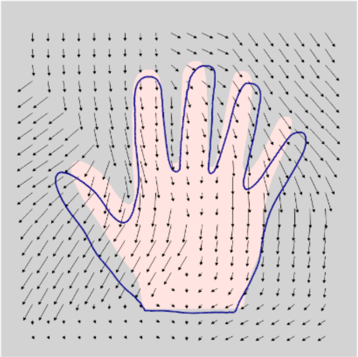
$$k(x, y) = \sum_{k \in \mathbb{Z}^d} \beta(sx - k)\beta(sy - k)$$

Where s is a scaling constant

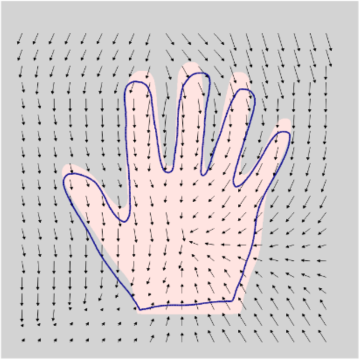
- Rueckert, Daniel, et al. "Nonrigid registration using free-form deformations: application to breast MR images." *IEEE transactions on medical imaging* 18.8 (1999): 712-721.
- Klein, Stefan, et al. "Elastix: a toolbox for intensity-based medical image registration." *IEEE transactions on medical imaging* 29.1 (2010): 196-205.

Statistical deformation models

Estimate mean and covariance function from data:

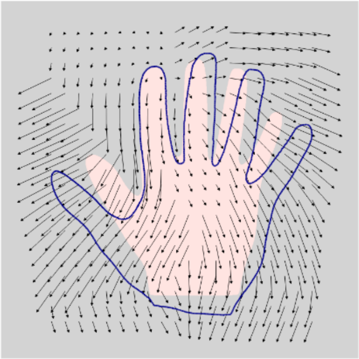


$u^1 : \Omega \to \mathbb{R}^2$



$u^2 : \Omega \to \mathbb{R}^2$

...



$u^n : \Omega \to \mathbb{R}^2$

$$\mu(x) = \overline{u}(x) = \frac{1}{n}\sum_{i-1}^{n} u^i(x)$$

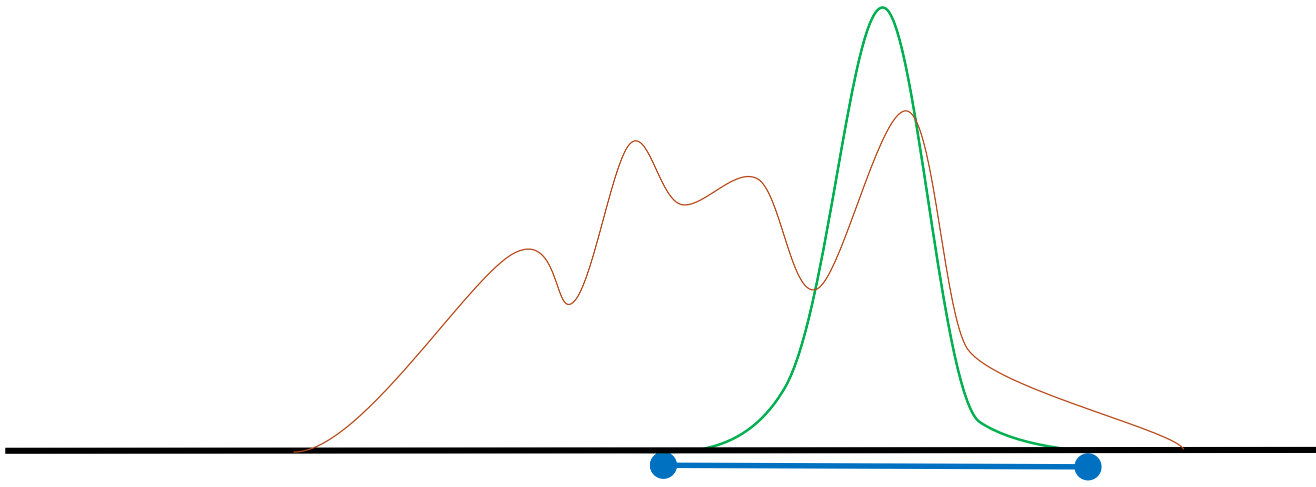
$$k_{SM}(x, x') = \frac{1}{n-1}\sum_{i}^{n}\left(u^i(x) - \overline{u}(x)\right)\left(u^i(x') - \overline{u}(x')\right)^T$$

Summary: Priors

- Leads to formulation of many standard transformation models in terms of Gaussian process
  - Improves understanding of methods
  - Let's us switch between "priors"

- Purely conceptual formulation
  - No algorithms

- Can sample and visualize deformations
  - Invaluable to check assumptions



Outline

- The registration problem
  - Problem formulation
  - Registration as analysis by synthesis problem
  - An algorithm using Gaussian process priors
- Priors for registration
  - Spline-based models, Radial basis functions
  - Multis-scale and Spatially-varying models
  - Statistical deformation models
- Likelihood functions
  - Landmark registration
  - Image to image registration
  - Surface to image registration
- Advancing registration

Landmark likelihood

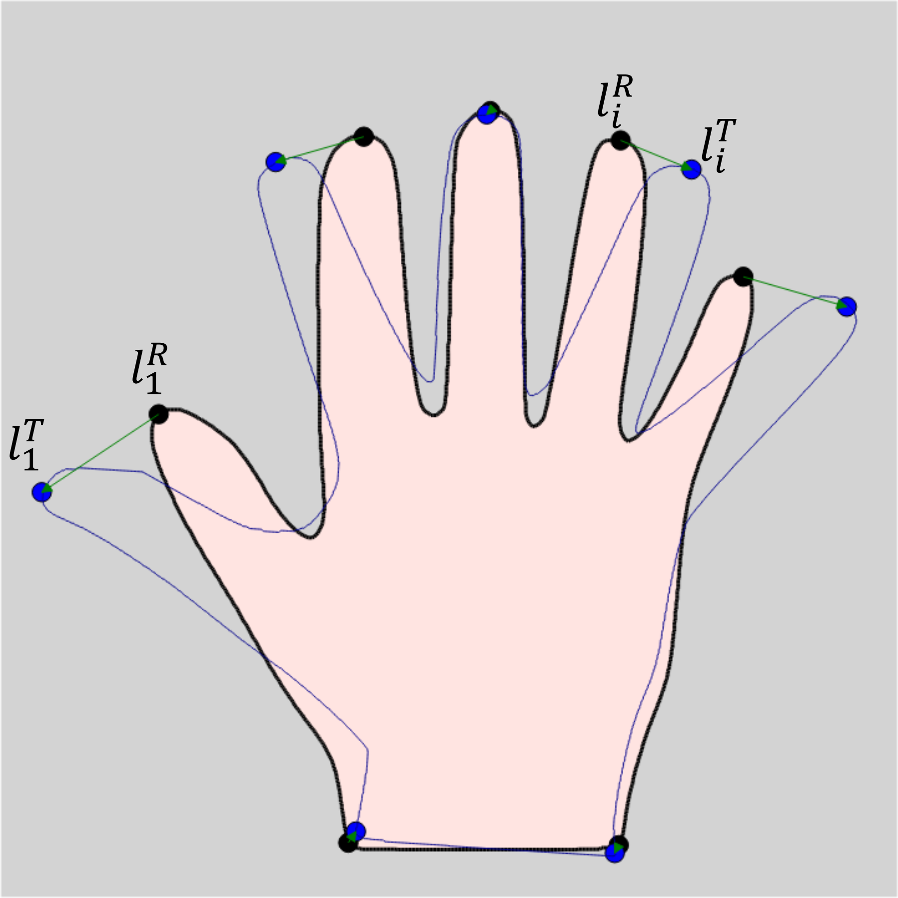
For one landmark pair (l_R, l_T):

$$p(l_T|\theta, l_R) = N(\varphi[\theta](l_R), \sigma^2)$$

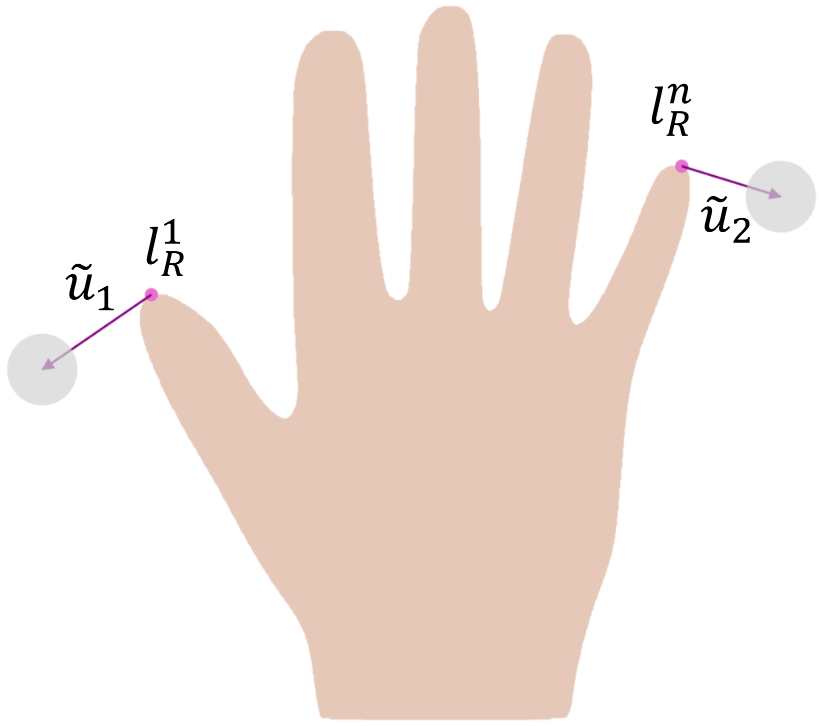
For many landmarks

$$L = ((l_R^1, l_T^1), \dots, (l_R^n, l_T^n))$$

$$p(l_1^T, \dots, l_n^T | \theta, l_R^1, \dots, l_R^n) = \prod_i N(\varphi[\theta](l_R), \sigma^2)$$



Landmark likelihood and GP Regression



Given:

- Gaussian process: $u \sim GP(\mu, k)$
- Observations: $\{(l_i^R, \tilde{u}_i), i = 1, \dots, m\}$

Assume:

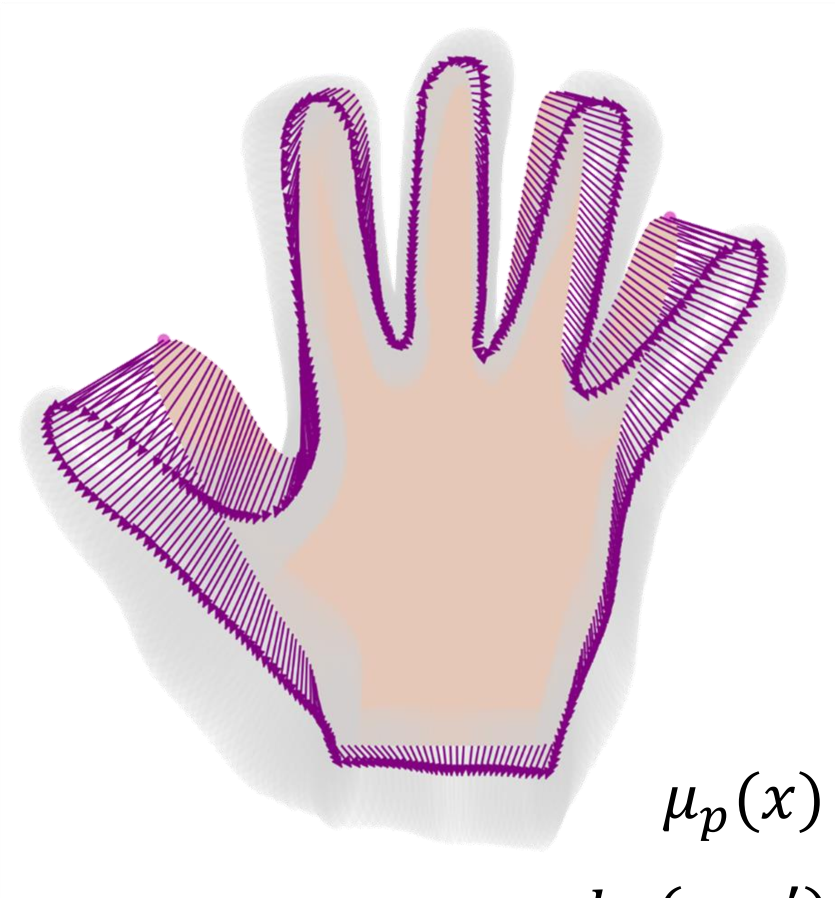
$$u(\tilde{x}_i) + \epsilon = \tilde{u}_i \text{ with } \epsilon \sim N(0, \sigma^2 I_{2\times 2}).$$

Goal:

- Find posterior distribution

$$u \mid l_1^R, \dots, l_n^R, \tilde{u}_1, \dots, \tilde{u}_m$$

Gaussian process regression



The posterior

$$u \,| l_1^R, \ldots, l_n^R, \tilde{u}_1, \ldots, \tilde{u}_m$$

is a Gaussian process

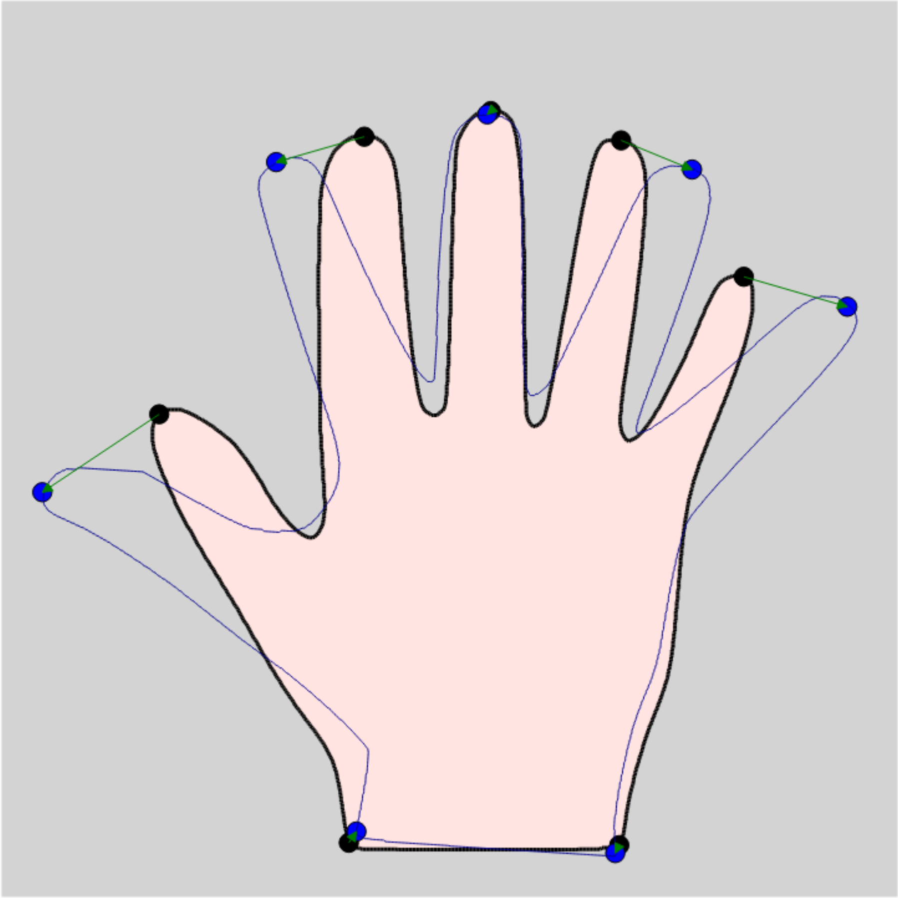
$$GP(\mu_p, k_p)$$

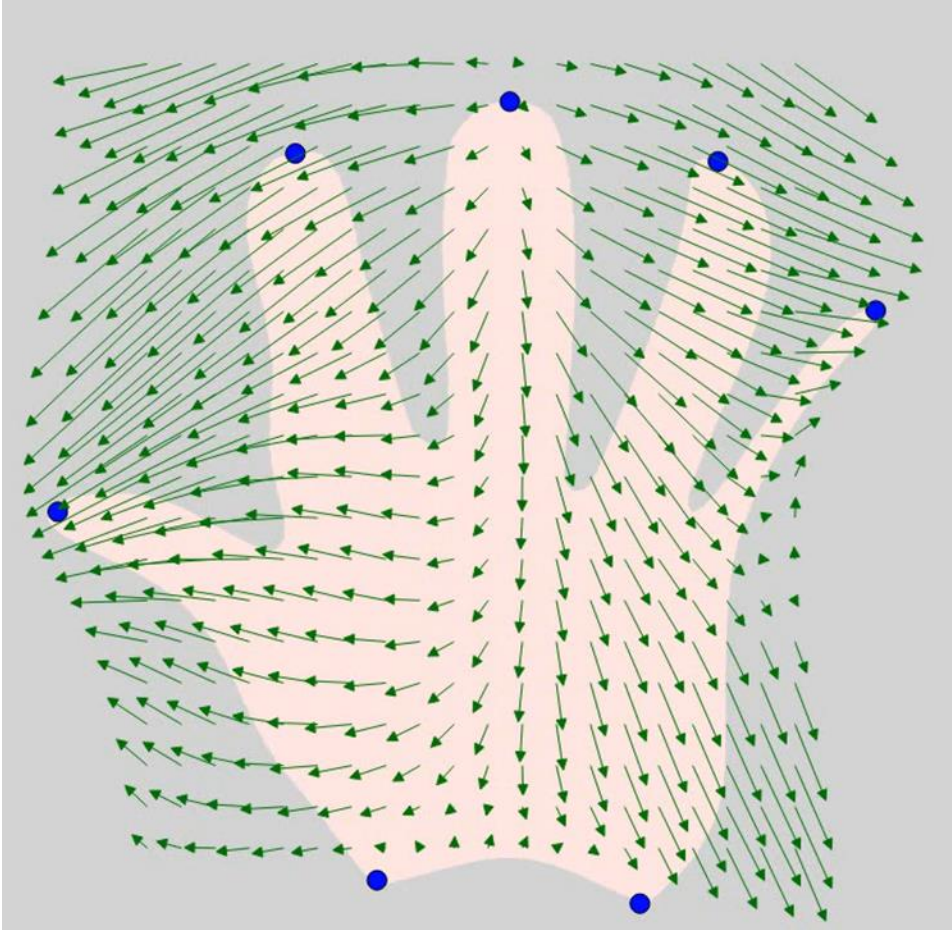
Its parameters are known analytically.

$$\mu_p(x) = \mu(x) + K(x, Y)(K(Y,Y) + \sigma^2 I_{2m\times 2m})^{-1}(\tilde{\boldsymbol{u}} - \mu(Y))$$

$$k_p(x, x') = k(x, x') - K(x, Y)\,(K(Y,Y) + \sigma^2 I_{2m\times 2m})^{-1} K(Y, x')$$

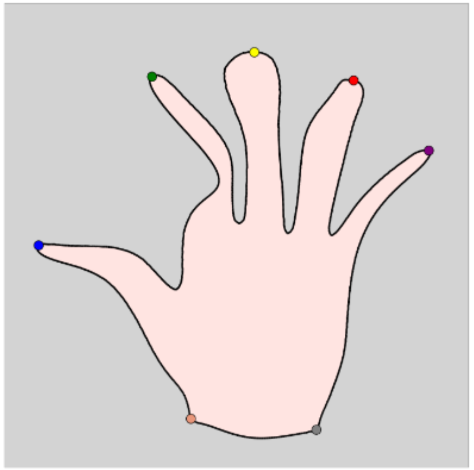
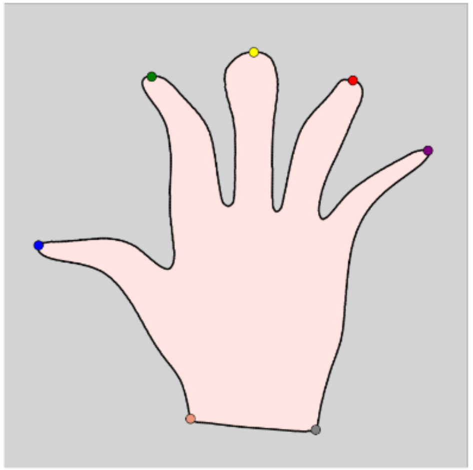
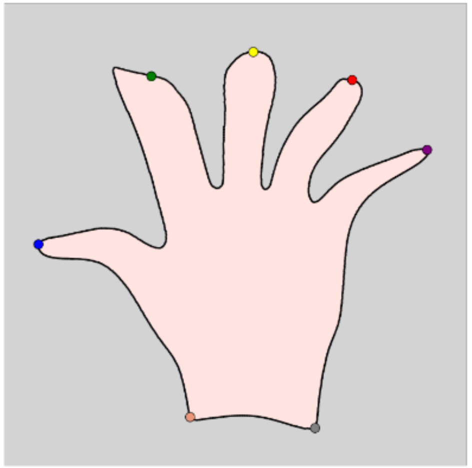
Landmark registration





Landmark likelihood: Some remarks

- Classical problems in registration
- Needs either many landmark points or good structure of prior to achieve good results

Rohr, Karl, et al. "Landmark-based elastic registration using approximating thin-plate splines." *IEEE Transactions on medical imaging* 20.6 (2001): 526-534.

Image to image registration

- What is a good synthesis function?

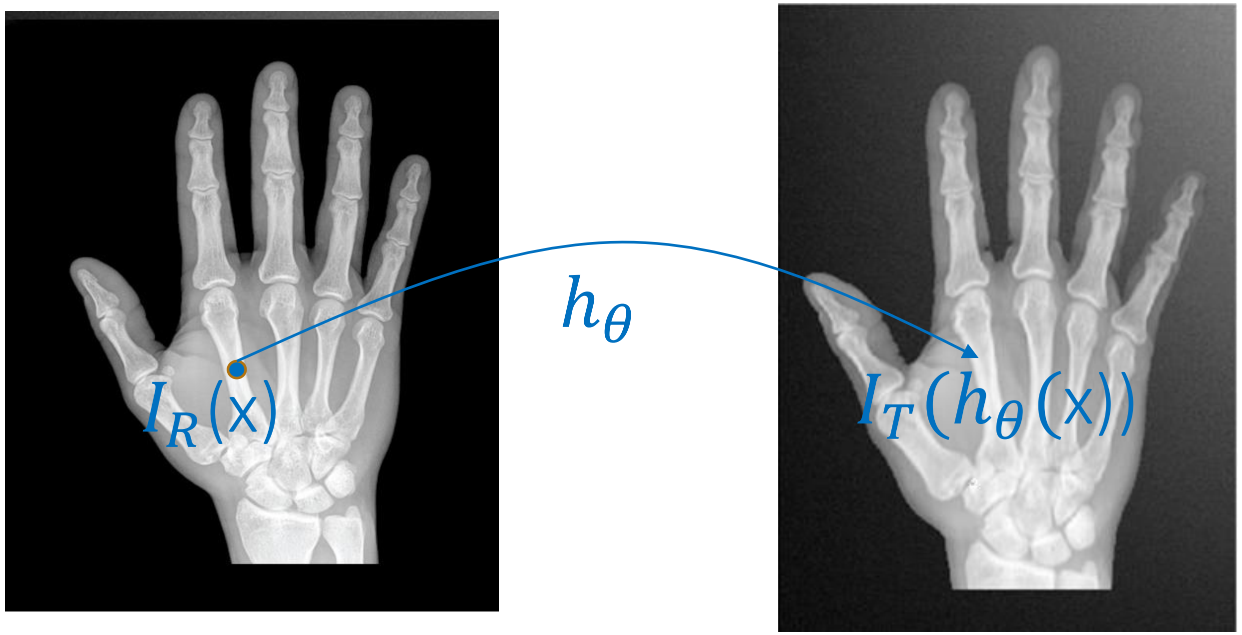
$$I_R \circ h_\theta^{-1} \longrightarrow$$

Simple choice: Use the warped reference image!

Image likelihood (single point)

- Probabilistic model: $I_T\big(h_\theta(x)\big) = I_R(x) + \epsilon,\ \ \epsilon \sim N(0, \sigma^2), x \in \Omega_\mathrm{R}$
- Likelihood for a single point x:

$$p(I_T(h_\theta(x))|\theta, I_R, x) \sim N(I_R(x), \sigma^2)$$



$I_R : \Omega \to \mathbb{R}$ $\qquad$ $I_T : \Omega \to \mathbb{R}$

Image likelihood (full image)

- Assuming that noise is independent at each point:

$$p\left(I_T \circ h_\theta \,\middle|\, \vec{\theta}, I_R\right) \sim \prod_{x \in \mathrm{I}_R} N(I_R(x), \sigma^2)$$

$$p\left(I_T \circ h_\theta \,\middle|\, \vec{\theta}, I_R\right) = \frac{1}{Z} \prod_{x \in I_R} \exp(-\frac{\left(I_R(x) - I_T\left(h_\theta(x)\right)\right)^2}{\sigma^2})$$

The sum of squared distance metric

$$\ln p\left(I_T \circ h_\theta \,\middle|\, \vec{\theta}, I_R\right) = \ln\left[\frac{1}{Z}\prod_{x\in I_R} \exp\left(-\frac{\left(I_R(x) - I_T\left(h_\theta(x)\right)\right)^2}{\sigma^2}\right)\right]$$

$$= -Z_1 \sum_{x\in I_R} \left(I_R(x) - I_T\left(h_\theta(x)\right)\right)^2$$

$$= D_{SSD}[I_R, I_T, h_\theta]$$

- The sum of squared differences implements an independence assumption
- The parameter σ^2 becomes a weighting constant.

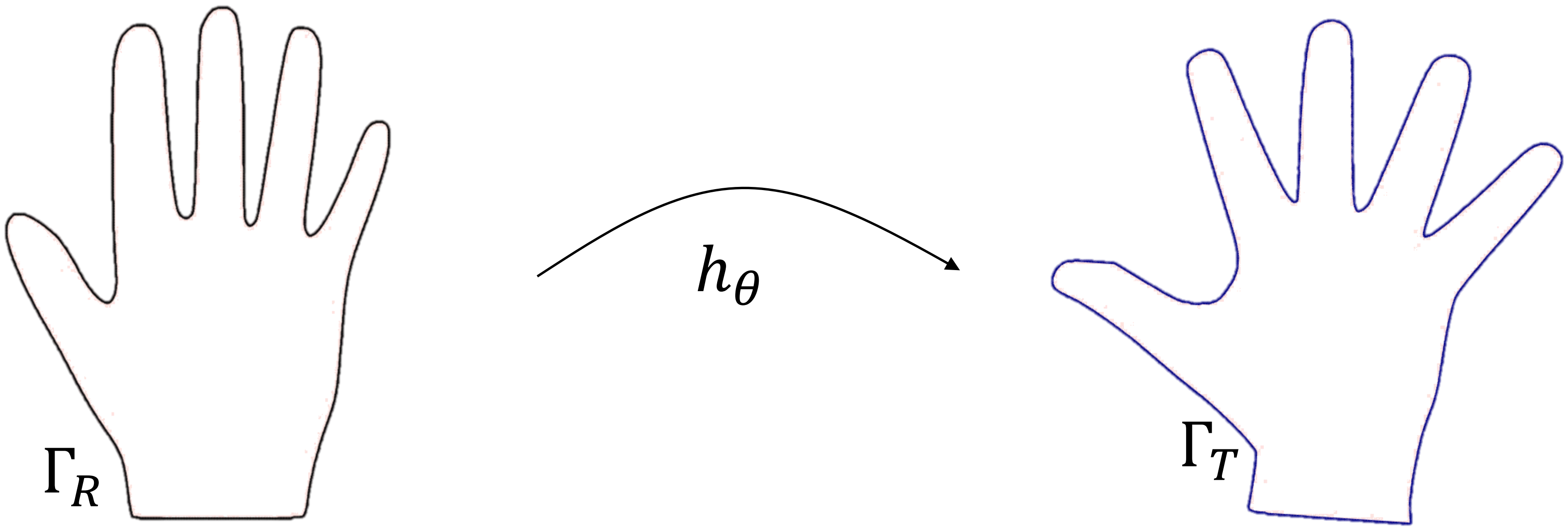
Likelihood from other metrics

- We can use any standard image metric $D[I_R, I_T, h_\theta]$ to define a likelihood function:

$$p(I_T|\theta, I_R) = \frac{1}{Z}\exp -(D[I_R, I_T, h_\theta])$$

- Examples:
  - Normalized cross correlations
  - Mutual information
  - …
- Makes it possible to reformulate any standard registration problem into the analysis by synthesis framework.
- Special case of collective likelihood

What about surface registration?



Reference (surface):
Γ_R

Target (surface):
Γ_T

A trick: Implicit definition of a surface

- Any surface Γ can be represented as the zero level set of a level set function Φ

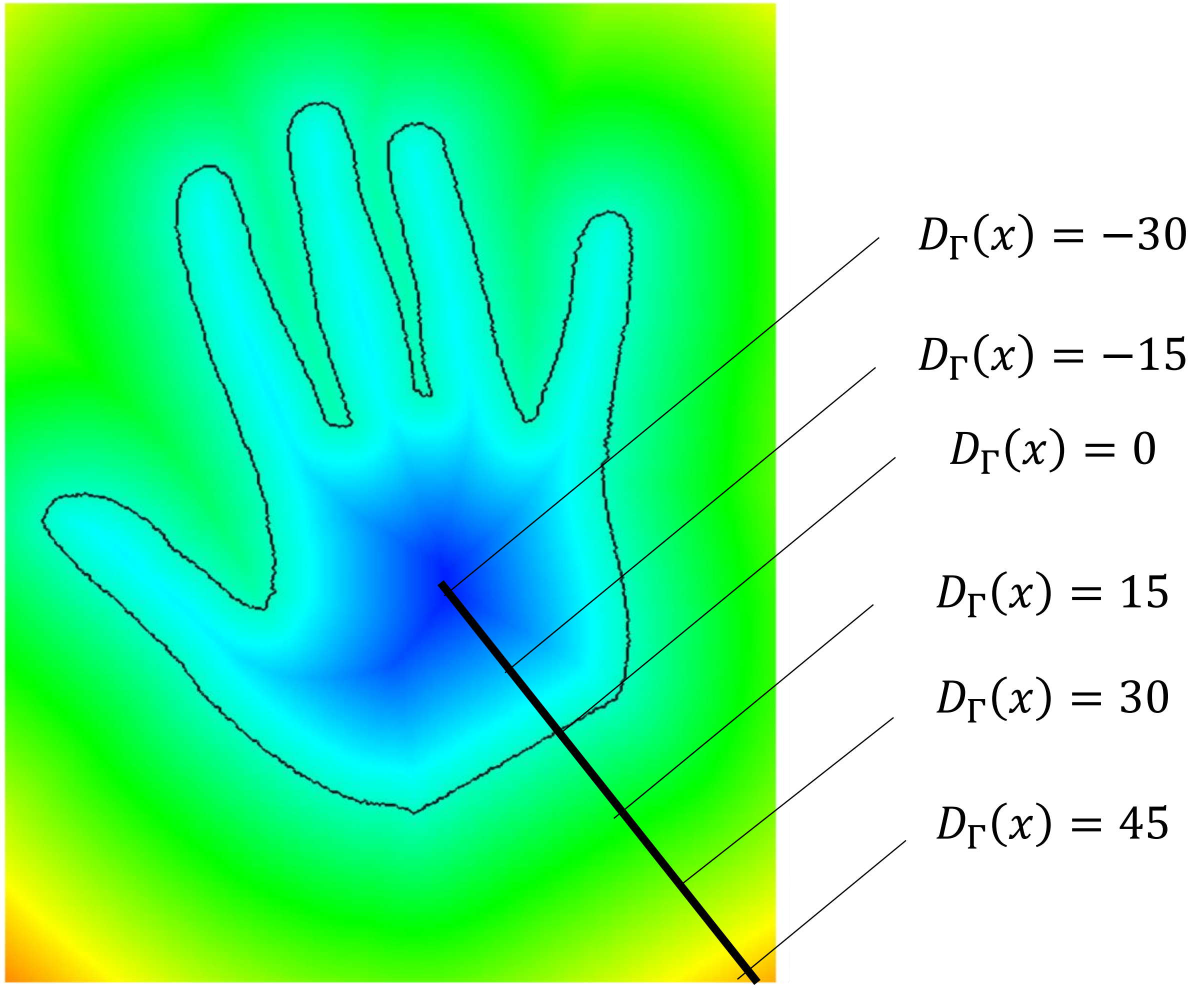
$$\Gamma = \{x \mid \Phi(x) = 0\}$$

- Popular choice is the signed distance function defined as

$$D_\Gamma(x) = \|\mathrm{CP}_\Gamma(x) - x\|$$

with

$$\mathrm{CP}_\Gamma(x) = \arg\min_{x' \in \Gamma} \|x - x'\|$$

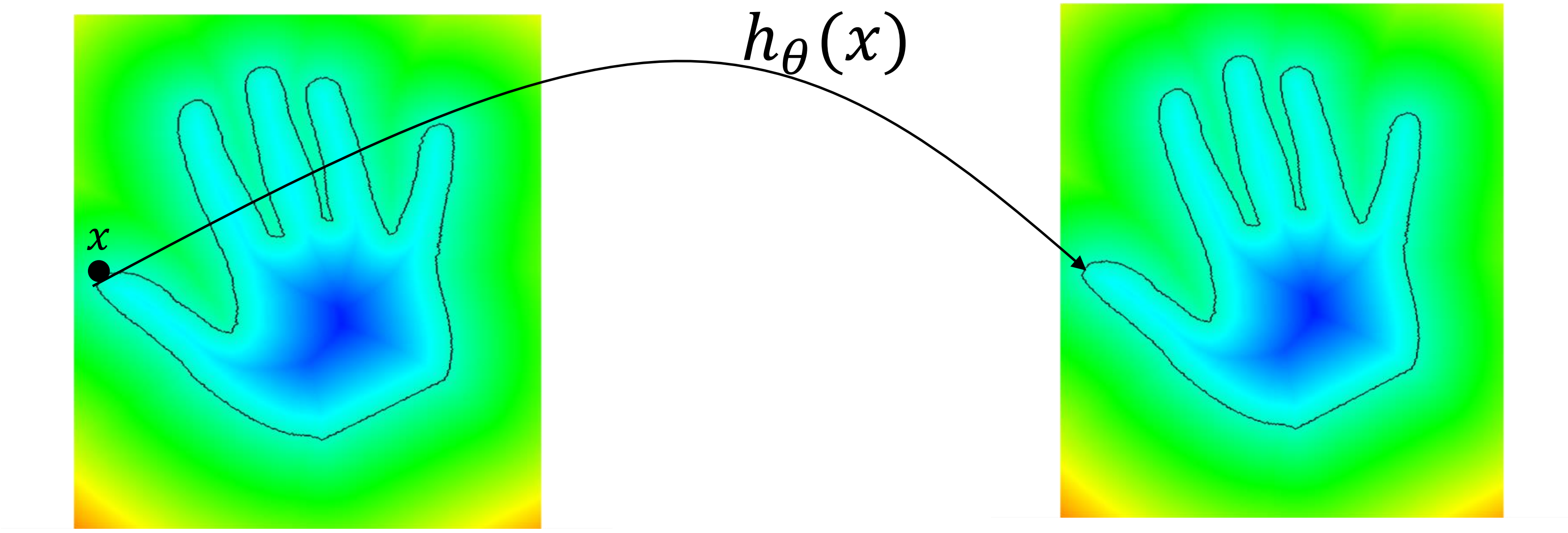


Surface registration as image registration

- We define the distance functions and use image to image likelihoods

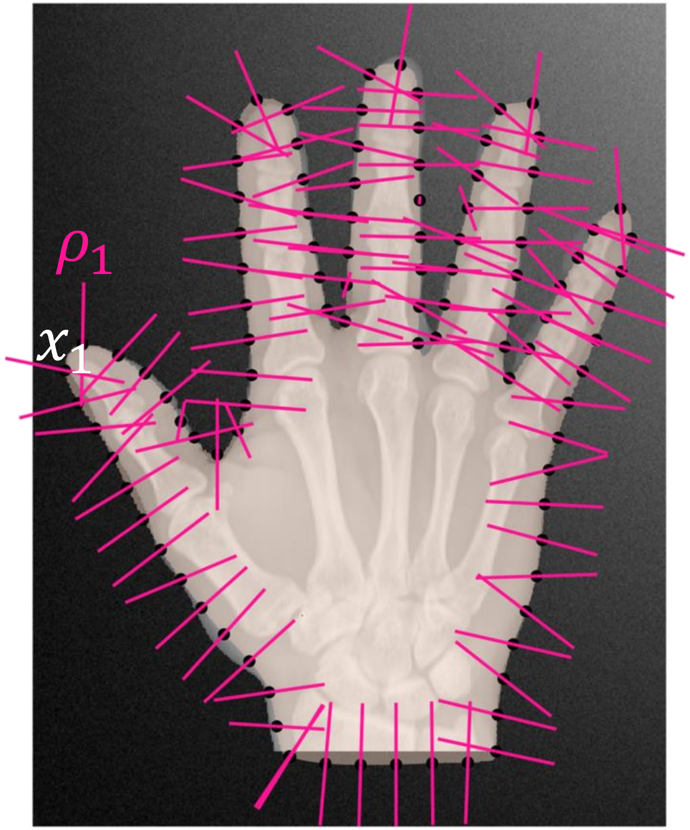
$$p\left(D_T(h_\theta(x)) \middle| \vec{\theta}, D_R, x\right) \sim N(D_R(x), \sigma^2)$$

- Most likely solution will map points on the zero level-sets to each other
  - Noise parameter σ^2 has geometric interpretation (variance of distance between the mapped points)



Reference $D_R: \Omega_R \rightarrow \mathbb{R}$ Target $D_T : \Omega_T \rightarrow \mathbb{R}$

Active shape models (surface to image registration)



- ASMs model each profile as a normal distribution

$$p(\rho_i) = N(\mu_i, \Sigma_i)$$

- Single profile point x_i:

$$p(\mathrm{I}_T(h_\theta(x_i))|\theta, x_i) = N(\mu_i, \Sigma_i)$$

- Image likelihood:

$$p(\mathrm{I}_T(h_\theta(x))|\theta, \Gamma_R) = \prod_i N(\mu_i, \Sigma_i)$$

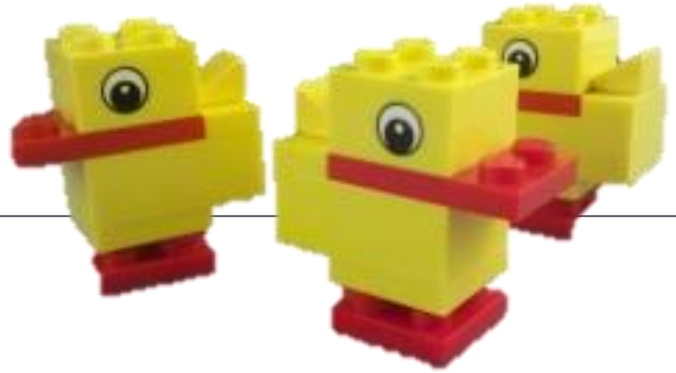
Summary: Likelihood functions

- Synthesis function is often just a warp of a reference
  - Works well if modality and dimensionality is the same
  - Leads to very simple systems
- We get probabilistic interpretations of some standard metrics
  - Makes assumptions more clear
- If we do not want full interpretation any metric can be turned into a likelihood function

Outline

- The registration problem
  - Problem formulation
  - Registration as analysis by synthesis problem
  - An algorithm using Gaussian process priors
- Priors for registration
  - Spline-based models
  - Radial basis functions
  - Statistical deformation models
- Likelihood functions
  - Landmark registration
  - Image to image registration
  - Surface to image registration
- Advancing registration
  - More expressive priors
  - Hybrid registration
  - ASM using the Metropolis Hastings algorithm

Constructing s.p.d. kernels



1. $k(x,x') = f(x)\, f(x')^T, f: X \to \mathbb{R}^d$
2. $k(x,x') = \alpha k(x,x'), \alpha \in \mathbb{R}_+$ (scaling)
3. $k(x,x') = B^T k(x,x') B, \quad \mathrm{B} \in \mathbb{R}^{r\times d}$ (lifting)
4. $k(x,x') = k_1(x,x') + k_2(x,x')$ (or relationship)
5. $k(x,x') = k_1(x,x') \cdot k_2(x,x')$ (and relationship)
6. $k(x,x') = k(\phi(x), \phi(x'))$ (domain warp)

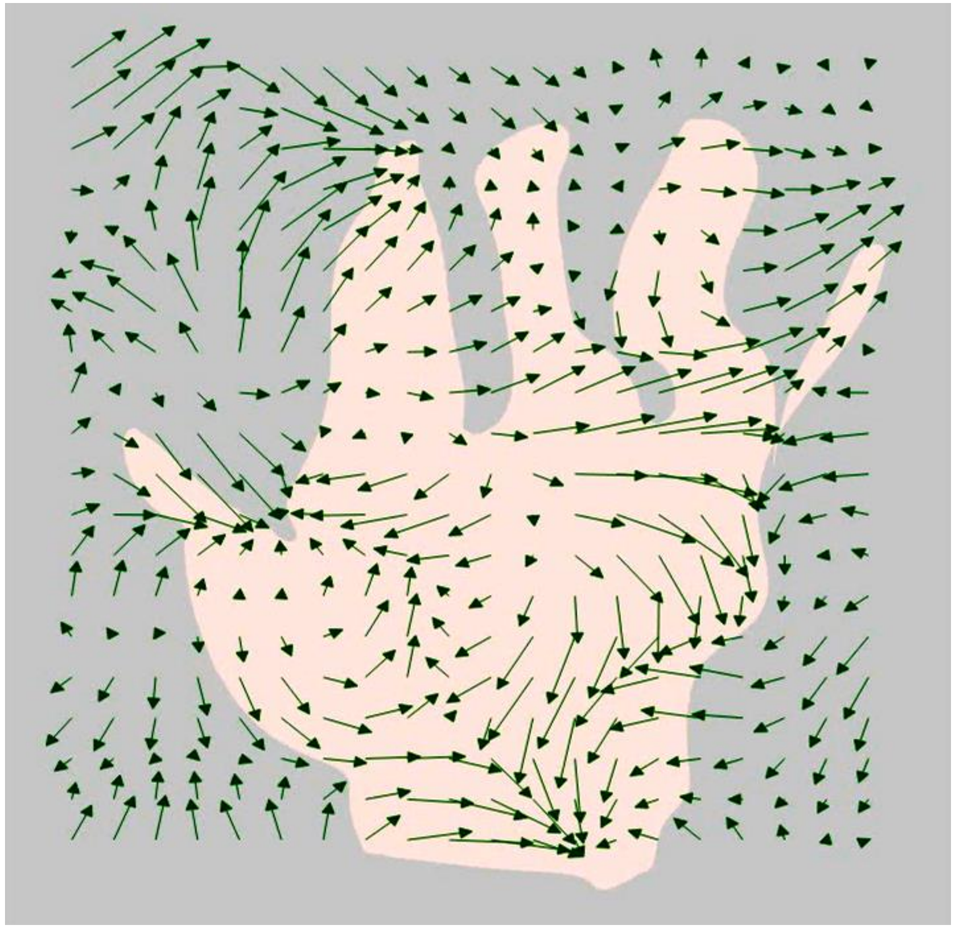
Multi-scale kernels

Add kernels that act on different scales:

$$k(x, x') = \sum_{i=0}^{n} \sum_{k \in \mathbb{Z}^d} \beta(2^{-i}x - k)\beta(2^{-i}y - k)$$

Opfer, Roland. "Multiscale kernels."
Advances in computational mathematics 25.4 (2006): 357-380.

Multi-scale kernel



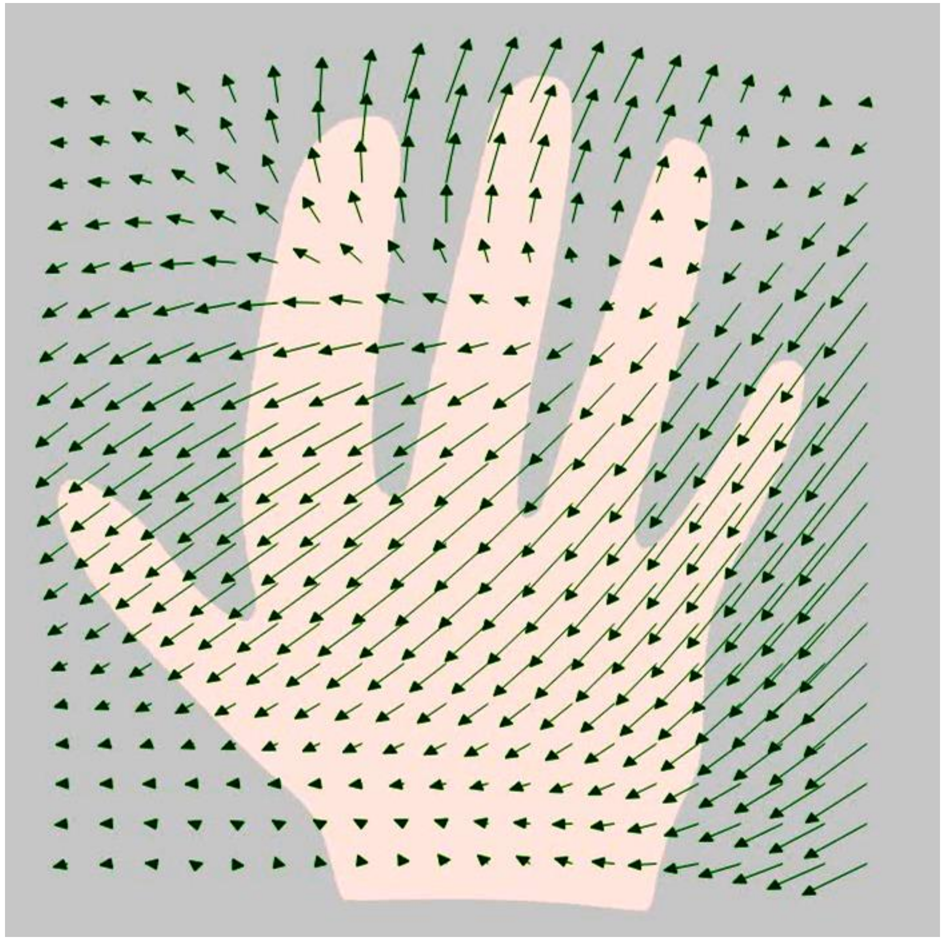
Anisotropic priors

Scale deformations differently in each direction

$$\mathrm{k}(x, x') = R^T \begin{pmatrix} \sqrt{s_1} & 0 \\ 0 & \sqrt{s_2} \end{pmatrix} k(x, x') \begin{pmatrix} \sqrt{s_1} & 0 \\ 0 & \sqrt{s_2} \end{pmatrix} R$$

- R is a rotation matrix
- k is scalar valued
- s_1, s_2 scaling factors

Anisotropic priors

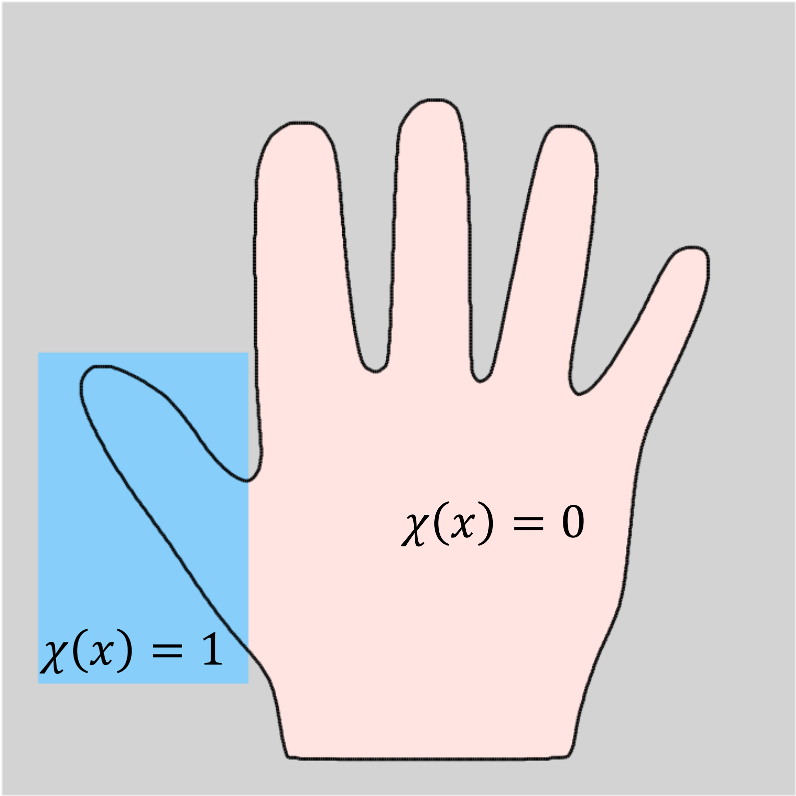


Spatially-varying priors

Use different models for different regions

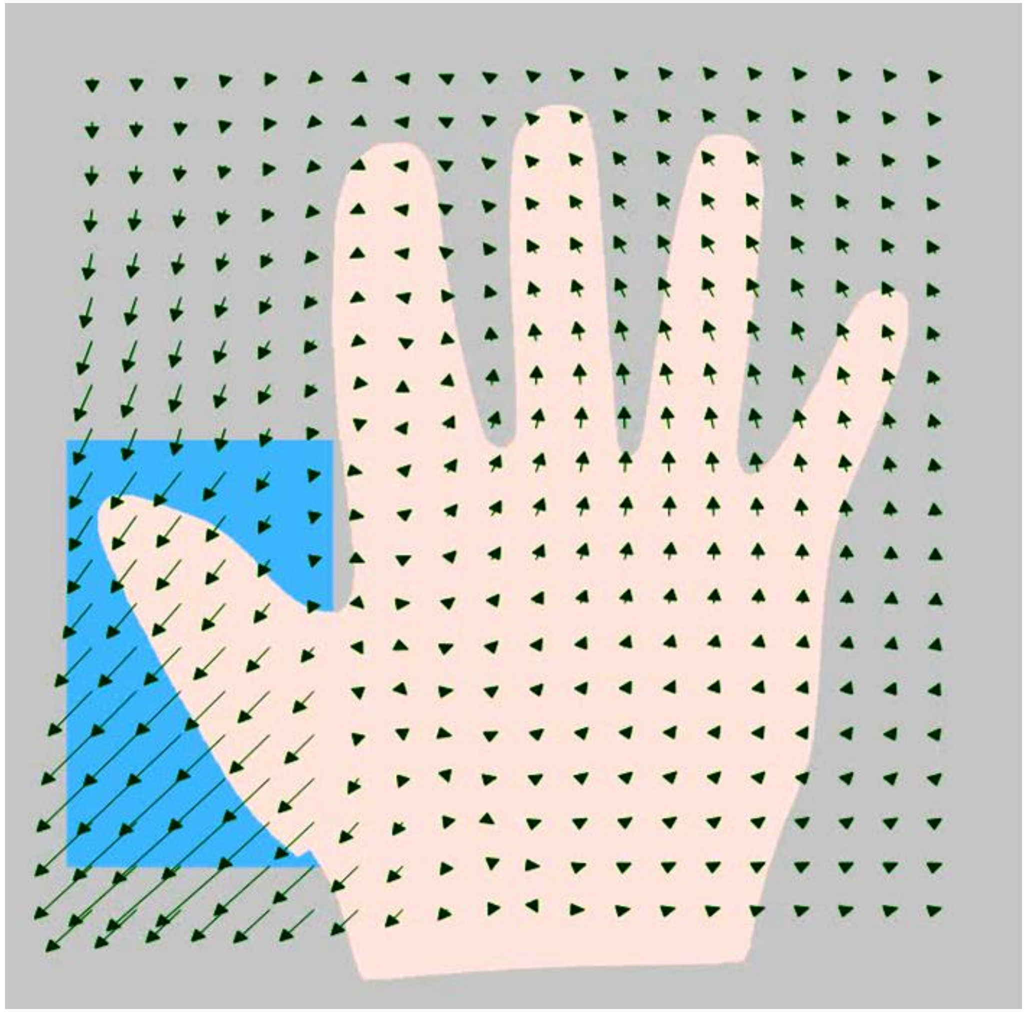
$$k(x,x') = \chi(x)\chi(x')k_1(x,x') + \big(1-\chi(x)\big)(1-\chi(x'))\, k_2(x,x')$$

$$\chi(x) = \begin{cases} 1 & \text{if } x \in \text{thumb region} \\ 0 & \text{otherwise} \end{cases}$$

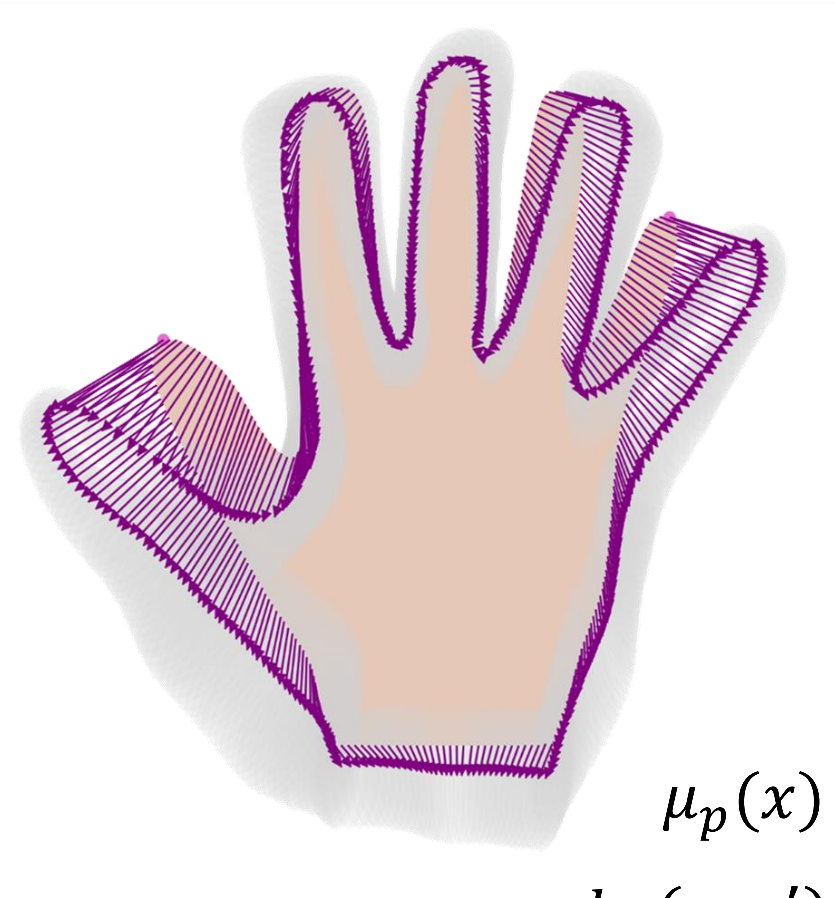


Freiman, Moti, Stephan D. Voss, and Simon K. Warfield. "Demons registration with local affine adaptive regularization: application to registration of abdominal structures." *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE, 2011.

Spatially-varying priors



Landmark registration using Gaussian processes



The posterior

$$u \mid \tilde{x}_1, \dots, \tilde{x}_m, \tilde{u}_1, \dots, \tilde{u}_m$$

is a Gaussian process

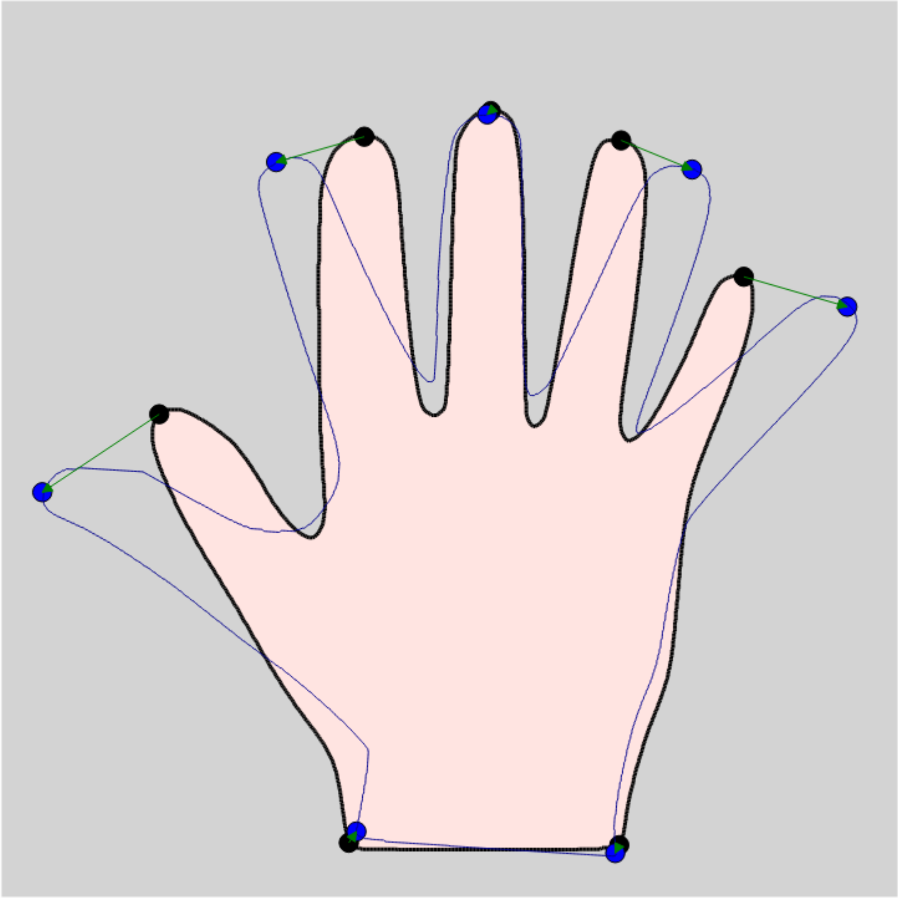
$$GP(\mu_p, k_p)$$

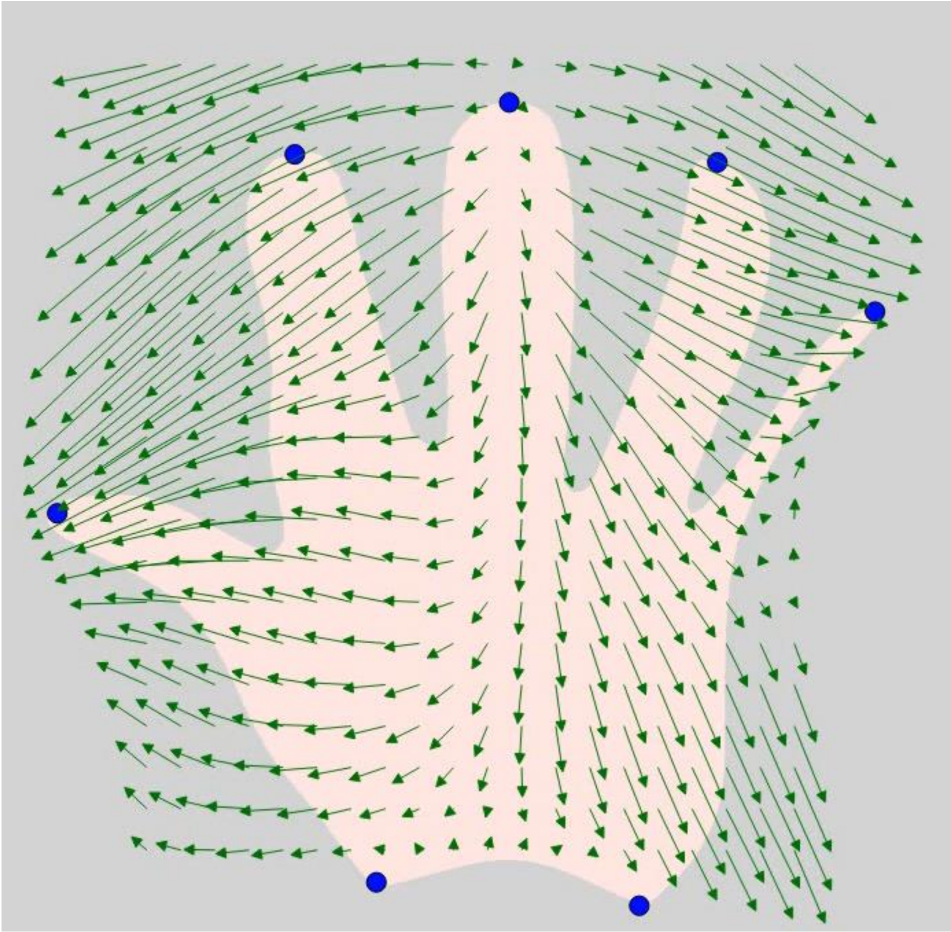
Its parameters are known analytically.

$$\mu_p(x) = \mu(x) + K(x, Y)(K(Y, Y) + \sigma^2 I_{2m \times 2m})^{-1}(\tilde{\boldsymbol{u}} - \mu(Y))$$

$$k_p(x, x') = k(x, x') - K(x, Y)\,(K(Y, Y) + \sigma^2 I_{2m \times 2m})^{-1} K(Y, x')$$

Landmark registration





Hybrid registration

- We can now combine landmark registration with intensity:
  1. Compute a posterior model using landmarks $GP(\mu, k)$
  2. Use $GP(\mu_p, k_p)$ as prior for registration with any image likelihood you like
- Example of Bayesian inference:

$$p(u) \rightarrow p(u|L_R, L_T) \rightarrow p(u|L_R, L_T, I_R, I_T)$$

- Elegant solution to hybrid registration

Wörz, Stefan, and Karl Rohr. "Hybrid spline-based elastic image registration using analytic solutions of the navier equation." *Bildverarbeitung für die Medizin 2007*. Springer Berlin Heidelberg, 2007. 151-155.

Lu, Huanxiang, Philippe C. Cattin, and Mauricio Reyes. "A hybrid multimodal non-rigid registration of MR images based on diffeomorphic demons." *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010.

Use Metropolis-Hastings for registration

Draw a sample x' from $Q(x'\|x)$	Propose
With probability $\alpha = \min\left\{\frac{P(\boldsymbol{x}')}{P(\boldsymbol{x})}, 1\right\}$ accept $\boldsymbol{x}'$ as new sample	Verify

- Frees us from "tyranny of differentiability"
  - Easy to integrate contours
  - Let's us model effects such as outliers, artifacts, … in principled ways
- Makes it possible to integrate results of bottom up proposals (landmark detectors)
- Let's us reason about uncertainty of a solution

Call to arms

- Our MCMC scheme was designed for really difficult problems
  - 3D => 2D
  - Complex illumination
  - No scale
  - Uncontrolled environment

- Let's start together to tackle the complicated problems in medical image analysis.

Call to arms

- Our MCMC scheme was designed for really difficult problems
  - 3D => 2D
  - Complex illumination
  - No scale
  - Uncontrolled environment
- Let's start together to tackle the complicated problems in medical image analysis

Now you know how!