A Multigrid Preconditioner for Hyperelastic Image Registration

Mini Symposium on Preconditioning Methods in Large-Scale Ill-Posed Inverse Problems

Lars Ruthotto\textsuperscript{1}, Chen Greif\textsuperscript{2} and Jan Modersitzki\textsuperscript{3}

\textsuperscript{1}Emory University, Atlanta, GA
\textsuperscript{2}The University of British Columbia, Canada
\textsuperscript{3}University of Lübeck, Germany

lruthotto@emory.edu
Introduction
Image Registration

Given a template image $\mathcal{T}$ and a reference image $\mathcal{R}$ find a reasonable transformation $y$, such that the transformed image $\mathcal{T}[y]$ is similar to $\mathcal{R}$, i.e., solve

$$\text{distance}[\mathcal{T}[y], \mathcal{R}] + \text{regularizer}[y] \xrightarrow{y} \min$$

subject to $\kappa_m \leq \text{constraints}[y] \leq \kappa_M$. 

Image Registration (Modersitzki 2009)
Image Registration

Given a template image $\mathcal{T}$ and a reference image $\mathcal{R}$ find a reasonable transformation $\gamma$, such that the transformed image $\mathcal{T}[^\gamma]$ is similar to $\mathcal{R}$, i.e., solve

$$\text{distance}[\mathcal{T}[^\gamma], \mathcal{R}] + \text{regularizer}[^\gamma] \xrightarrow{\gamma} \text{min}$$

subject to $\kappa_m \leq \text{constraints}[^\gamma] \leq \kappa_M$. 

Image Registration (Modersitzki 2009)
Image Registration (Modersitzki 2009)

Given a template image $\mathcal{T}$ and a reference image $\mathcal{R}$ find a reasonable transformation $y$, such that the transformed image $\mathcal{T}[y]$ is similar to $\mathcal{R}$, i.e., solve

$$
\mathcal{D}[\mathcal{T}[y], \mathcal{R}] + \alpha \mathcal{S}[y] \xrightarrow{y} \min
$$

subject to $\kappa_m \leq C[y] \leq \kappa_M$. 

reference $\mathcal{R}$  template $\mathcal{T} + \text{grid } y$  $\mathcal{T}[y]$
Hyperelastic Image Registration
Controlling Volume Change

Observation: Compressibility of tissue is limited in most applications.

Example: $\Omega = (-1, 1)^2$, $y_n : \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto (1/2)^n x$, for $n \in \mathbb{N}$.
Controlling Volume Change

Observation: Compressibility of tissue is limited in most applications.

Example: \( \Omega = (-1, 1)^2 \), \( y_n : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \), \( x \mapsto (1/2)^n x \), for \( n \in \mathbb{N} \).
Controlling Volume Change

**Observation:** Compressibility of tissue is limited in most applications.

Example: \( \Omega = (-1, 1)^2, \ y_n : \mathbb{R}^2 \to \mathbb{R}^2, \ x \mapsto (1/2)^n x, \ \text{for} \ n \in \mathbb{N}. \)
Controlling Volume Change

Observation: Compressibility of tissue is limited in most applications.

Example: $\Omega = (-1, 1)^2$, $y_n : \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto (1/2)^n x$, for $n \in \mathbb{N}$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>elastic$[y_n]$</th>
<th>hyperelastic$[y_n]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6.75</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>$\infty$</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
Controlling Volume Change

**Observation:** Compressibility of tissue is limited in most applications.

Example: $\Omega = (-1, 1)^2$, $y_n : \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto (1/2)^n x$, for $n \in \mathbb{N}$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>elastic $[y_n]$</th>
<th>hyperelastic $[y_n]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.00</td>
<td>82.20</td>
</tr>
<tr>
<td>2</td>
<td>6.75</td>
<td>3166.76</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$\infty$</td>
<td>12</td>
<td>$\infty$</td>
</tr>
</tbody>
</table>
Definition of our Hyperelastic Regularization Functional

**Definition (Hyperelastic Regularization Functional)**

For regularization parameters $\alpha_l, \alpha_a, \alpha_v > 0$ the hyperelastic regularization functional $S_{\text{hyper}} : \mathcal{A} \to \mathbb{R}^+$ is defined as

$$S_{\text{hyper}}[y] = \alpha_l S_{\text{length}}[y] + \alpha_a S_{\text{area}}[y] + \alpha_v S_{\text{vol}}[y]$$

$$= \int_\Omega \alpha_l \|\nabla u(x)\|_\text{Fro}^2 + \alpha_a \varphi_c(\text{cof}\nabla y(x)) + \alpha_v \psi(\det \nabla y(x)) \, dx$$

with the convex functions

$$\varphi_c(C) = \sum_{i=1}^3 \max\{\sum_{j=1}^3 C_{ji}^2 - 1, 0\}^2 \text{ and } \psi(v) = (v - 1)^4/v^2.$$
Definition of our Hyperelastic Regularization Functional

**Definition (Hyperelastic Regularization Functional)**

For regularization parameters $\alpha_l, \alpha_a, \alpha_v > 0$ the hyperelastic regularization functional $S^{\text{hyper}} : \mathcal{A} \rightarrow \mathbb{R}^+$ is defined as

$$S^{\text{hyper}}[y] = \alpha_l S^{\text{length}}[y] + \alpha_a S^{\text{area}}[y] + \alpha_v S^{\text{vol}}[y]$$

$$= \int_{\Omega} \alpha_l \|\nabla u(x)\|_\text{Fro}^2 + \alpha_a \varphi_c(\text{cof} \nabla y(x)) + \alpha_v \psi(\det \nabla y(x)) \, dx$$

with the convex functions

$$\varphi_c(C) = \sum_{i=1}^3 \max\{\sum_{j=1}^3 C_{ji}^2 - 1, 0\}^2 \quad \text{and} \quad \psi(v) = (v - 1)^4 / v^2.$$
Definition of our Hyperelastic Regularization Functional

**Definition (Hyperelastic Regularization Functional)**

For regularization parameters \(\alpha_l, \alpha_a, \alpha_v > 0\) the hyperelastic regularization functional \(S^\text{hyper} : A \to \mathbb{R}^+\) is defined as

\[
S^\text{hyper}[y] = \alpha_l S^\text{length}[y] + \alpha_a S^\text{area}[y] + \alpha_v S^\text{vol}[y]
= \int_\Omega \alpha_l \|\nabla u(x)\|_\text{Fro}^2 + \alpha_a \varphi_c(\text{cof}\nabla y(x)) + \alpha_v \psi(\det \nabla y(x)) dx
\]

with the convex functions

\[
\varphi_c(C) = \sum_{i=1}^{3} \max\{\sum_{j=1}^{3} C_{ji}^2 - 1, 0\}^2 \text{ and } \psi(v) = (v - 1)^4/v^2.
\]

For a domain \(\Omega \subset \mathbb{R}^3\) bounded by \(M > 0\) the admissible set is

\[
A := \{y \in W^{1,2}(\Omega, \mathbb{R}^3) : \text{cof}\nabla y \in L^4(\Omega, \mathbb{R}^{3 \times 3}), \det \nabla y \in L^2(\Omega, \mathbb{R}),
|\int y(x)dx| \leq \text{vol}(\Omega)(M + \text{diam}(\Omega)), \det \nabla y > 0 \text{ a.e.}\}.
\]
Existence Result for Unconstrained Image Registration

Assumptions:

**A1:** There exists \( g_D : \Omega \times \mathbb{R}^3 \times \mathbb{R}^{3\times3} \times \mathbb{R} \to \mathbb{R}^+ \) that is continuously differentiable, convex in \( \det \nabla y \), measurable in \( x \), and satisfies

\[
D[y] = \int_\Omega g_D(x, y(x), \det \nabla y(x)) \, dx
\]

**A2:** \( J[Id] < \infty \) for \( Id(x) = x \) on \( \Omega \)

**A3:** \( \Omega \) has a \( C^1 \) boundary denoted by \( \partial \Omega \)
Existence Result for Unconstrained Image Registration

Assumptions:

**A1:** There exists $g_D : \Omega \times \mathbb{R}^3 \times \mathbb{R}^{3\times3} \times \mathbb{R} \to \mathbb{R}^+$ that is continuously differentiable, convex in $\det \nabla y$, measurable in $x$, and satisfies

$$D[y] = \int_\Omega g_D(x, y(x), \det \nabla y(x)) \, dx$$

**A2:** $J[Id] < \infty$ for $Id(x) = x$ on $\Omega$

**A3:** $\Omega$ has a $C^1$ boundary denoted by $\partial \Omega$

---

**Theorem (Ball 1976, Ciarlet 1988, Droske & Rumpf 2003, LR 2012)**

Given images $\mathcal{T}, \mathcal{R} \in \text{Img}(\Omega)$, a distance functional $D$, $S^{\text{hyper}}$ and $A$ as on previous slide. Assume (A1)–(A3). Then

$$J[y] := D[y] + S^{\text{hyper}}[y] \xrightarrow{y} \min$$

admits a solution in $A$. 
Discretize $\rightarrow$ Optimize

Discretization $\mapsto$ finite dimensional problem:

$$\mathcal{D}^h(y^h) + \mathcal{S}^{\text{hyper}}(y^h) \xrightarrow{y^h} \min, \ y^h \in \mathcal{A}_h, \quad h \rightarrow 0.$$
Discretize $\rightarrow$ Optimize

Discretization $\rightsquigarrow$ finite dimensional problem:

$$
\mathcal{D}^h(y^h) + \mathcal{S}^{\text{hyper}}(y^h) \xrightarrow{y^h} \min, \ y^h \in A_h, \ h \rightarrow 0.
$$

- efficient optimization schemes (Gauss Newton, SQP, Augmented Lagrangian)
- line search (Armijo, ... with projection/backtracking)
- large steps
Discretize → Optimize

Discretization \(\leadsto\) finite dimensional problem:

\[
D_h(y_h) + S_{\text{hyper}}(y_h) \xrightarrow{y_h} \text{min}, \ y_h \in A_h, \ h \rightarrow 0.
\]

- efficient optimization schemes (Gauss Newton, SQP, Augmented Lagrangian)
- line search (Armijo, \ldots with projection/backtracking)
- large steps
- all parts have to be differentiable (data model)
Discretize $\rightarrow$ Optimize

Discretization $\sim$ finite dimensional problem:

$$D^h(y^h) + S^{\text{hyper}}(y^h) \rightarrow_{y^h} \min, \ y^h \in A_h, \ h \rightarrow 0.$$ 

- efficient optimization schemes (Gauss Newton, SQP, Augmented Lagrangian)

- line search (Armijo, ... with projection/backtracking)

- large steps

- all parts have to be differentiable (data model)

- need to solve linear system of type $H(y^h)\delta y = -\text{rhs},$

  $$H(y^h) \approx d_2D^h(y^h) + d_2S^{\text{hyper}}(y^h), \ \ \text{rhs} = dD^h(y^h) + dS^{\text{hyper}}(y^h).$$

- discretization not straightforward
Finite Element Space

Let $T \in \mathcal{M}$ be a triangular partition of the domain $\Omega \subset \mathbb{R}^d$ with $\text{vol}(T) \leq h$, $\forall T \in \mathcal{M}$. Consider the finite element space

$$
\mathcal{A}_h = \{ y \in C(\Omega, \mathbb{R}^d) : y|_T \text{ is affine linear } \forall T \in \mathcal{M} \} \subset W^{1,2}(\Omega, \mathbb{R}^d).
$$
Finite Element Space

Let $T \in \mathcal{M}$ be a triangular partition of the domain $\Omega \subset \mathbb{R}^d$ with $\text{vol}(T) \leq h$, $\forall T \in \mathcal{M}$. Consider the finite element space

$$\mathcal{A}_h = \{ y \in C(\Omega, \mathbb{R}^d) : y|_T \text{ is affine linear } \forall T \in \mathcal{M} \} \subset W^{1,2}(\Omega, \mathbb{R}^d).$$

$y^h|_T$ affine linear $\rightsquigarrow \nabla y^h|_T$ constant $\rightsquigarrow S^\text{hyper}[y^h]$ computed exactly.
Example: Linear vs. Hyperelastic
Example: Linear vs. Hyperelastic
Example: Linear vs. Hyperelastic

- Template
- Linear elastic
- Hyperelastic
- Data
- Transformation
- Final template
- \( \text{det} \nabla y \)
Motion Correction of Cardiac PET

- Fabian Gigengack, Martin Burger, European Institute of Molecular Imaging, University of Münster
- Otmar Schober, Department of Nuclear Medicine, University of Münster

Thoracic FDG\textsuperscript{18} PET, acquisition time $\approx 20$ min


*Motion correction in dual gated cardiac PET using mass-preserving image registration.*


*A Simplified Pipeline for Motion Correction in Dual Gated Cardiac PET.*
Motion Correction of Gated Cardiac PET

single gates $I_i$ without motion correction

Mass-preserving Image Registration

Given gates $I_1, \ldots, N$ and a reference gate $I_{\text{ref}}$ find reasonable transformations $y_1, \ldots, N$ solving

$$
\frac{1}{2} \left\| I_i[y_i] \cdot \det \nabla y_i - I_{\text{ref}} \right\|^2 + S^{\text{hyper}}[y_i] \xrightarrow{y_i} \min.
$$
Motion Correction of Gated Cardiac PET

Single gates $\mathcal{I}_i$ without motion correction with motion correction

Mass-preserving Image Registration

Given gates $\mathcal{I}_1, \ldots, \mathcal{I}_N$ and a reference gate $\mathcal{I}_{\text{ref}}$ find reasonable transformations $y_1, \ldots, y_N$ solving

$$\frac{1}{2} \left\| \mathcal{I}_i[y_i] \cdot \det \nabla y_i - \mathcal{I}_{\text{ref}} \right\|^2 + S^{\text{hyper}}[y_i] \xrightarrow{y_i} \text{min}.$$
Motion Correction of Gated Cardiac PET

Mass-preserving Image Registration

Given gates $I_1, \ldots, I_N$ and a reference gate $I_{\text{ref}}$ find reasonable transformations $y_1, \ldots, y_N$ solving

$$\frac{1}{2} \left\| I_i[y_i] \cdot \det \nabla y_i - I_{\text{ref}} \right\|^2 + S^{\text{hyper}}[y_i] \rightarrow \min.$$
Multigrid for Hyperelastic Image Registration
Examining the Hessian of $S^{\text{hyper}}$

Consider $S^{\text{hyper}}[y^h] = \alpha_l S^{\text{length}}[y^h] + \alpha_a S^{\text{area}}[y^h] + \alpha_v S^{\text{vol}}[y^h]$, for $y^h$ on $4^3$ grid.

\[
H^{\text{length}} = \alpha_l h^3 I_3 \otimes \hat{\Delta}
\]
Examining the Hessian of $S^{\text{hyper}}$

Consider $S^{\text{hyper}}[y^h] = \alpha_l S^{\text{length}}[y^h] + \alpha_a S^{\text{area}}[y^h] + \alpha_v S^{\text{vol}}[y^h]$, for $y^h$ on $4^3$ grid.

$$H^{\text{length}} = \alpha_l h^3 I_3 \otimes \Delta$$

$$H^{\text{area}}[y^h] = \alpha_a h^3 (d\text{cof}\nabla y^h)^T \text{diag}[\varphi''_c(\text{cof}\nabla y^h)] d\text{cof}\nabla y^h$$
Examining the Hessian of $S^{\text{hyper}}$

Consider $S^{\text{hyper}}[y^h] = \alpha_l S^{\text{length}}[y^h] + \alpha_a S^{\text{area}}[y^h] + \alpha_v S^{\text{vol}}[y^h]$, for $y^h$ on $4^3$ grid.

$$H^{\text{length}} = \alpha_l h^3 I_3 \otimes \Delta$$
$$H^{\text{area}}[y^h] = \alpha_a h^3 (d\text{cof}\nabla y^h)^T \text{diag}[\varphi''(\text{cof}\nabla y^h)] d\text{cof}\nabla y^h$$
$$H^{\text{vol}}[y^h] = \alpha_v h^3 (d \det \nabla y^h)^T \text{diag}[\psi''(\det \nabla y^h)] d \det \nabla y^h$$
$$\psi''(v) = 2(v^4 - 4v + 3)/v^4$$
Examining the Hessian of $S^{\text{hyper}}$

Consider $S^{\text{hyper}}[y^h] = \alpha_l S^{\text{length}}[y^h] + \alpha_a S^{\text{area}}[y^h] + \alpha_v S^{\text{vol}}[y^h]$, for $y^h$ on $4^3$ grid.

- $H^{\text{length}}$, #nz:10,575
- $H^{\text{area}}[y^h]$, #nz:42,157
- $H^{\text{vol}}[y^h]$, #nz:36,653

- Coupling between components of $y^h$
- Many non-zeros $\Rightarrow$ costly in terms of memory and computation
- $H^{\text{area}}$ and $H^{\text{vol}}$ change at every iteration (matrix-free implementation)
- $\psi''(v) \to \infty$ where $v \to 0^+$ $\Rightarrow$ convergence of iterative solver?
Local Fourier Analysis - 1

**Goal:** Quantify smoothing properties of $H(y) \approx d_2 D^h(y) + d_2 S^{\text{hyper}}(y)$
Local Fourier Analysis - 1

**Goal:** Quantify smoothing properties of $H(y) \approx d_2 D^h(y) + d_2 S^{\text{hyper}}(y)$

**Assumptions:**
- infinite domain in 2D, rectangular grid, fixed $h \in \mathbb{R}^+$
- regular triangulation
- no contribution of distance term to Hessian, i.e. $d_2 D^h(y_c) = 0$
- constant coefficients $\sim$ consider $y_c : \mathbb{R}^2 \to \mathbb{R}^2, y_c(x) = c \cdot x$ for $c \in \mathbb{R}^+$. 

Hessian:

$$H(c, \alpha_l) = \alpha_l H_{\text{length}} + \psi''(c^2) H_{\text{vol}}$$

with $\psi'(v) = 2(v^4 - 4v^2 + 3)/v^4$
Local Fourier Analysis - 1

**Goal:** Quantify smoothing properties of $H(y) \approx d_2 D^h(y) + d_2 S^{\text{hyper}}(y)$

**Assumptions:**
- infinite domain in 2D, rectangular grid, fixed $h \in \mathbb{R}^+$
- regular triangulation
- no contribution of distance term to Hessian, i.e. $d_2 D^h(y_c) = 0$
- constant coefficients $\Rightarrow$ consider $y_c : \mathbb{R}^2 \to \mathbb{R}^2, y_c(x) = c \cdot x$ for $c \in \mathbb{R}^+$.

**Hessian:**

$$H(c, \alpha) = \alpha_1 H^{\text{length}} + \psi''(c^2) H^{\text{vol}} \text{ with } \psi''(\nu) = 2(\nu^4 - 4\nu + 3)/\nu^4$$

with

$$H^{\text{length}} = \begin{pmatrix} D_1^T D_1 + D_2^T D_2 \\ D_1^T D_1 + D_2^T D_2 \end{pmatrix} \text{ and } H^{\text{vol}} = c^2 \begin{pmatrix} D_1^T D_1 & D_2^T D_1 \\ D_1^T D_2 & D_2^T D_2 \end{pmatrix}$$
Local Fourier Analysis - 2

Consider grid functions $\varphi(\theta, x) = \exp(\imath \theta \cdot x / h) \cdot [1, 1]^T$ for $\theta \in [-\pi/2, 3\pi/2]^2$. 
Local Fourier Analysis - 2

Consider grid functions $\varphi(\theta, x) = \exp(\nu \theta \cdot x/h) \cdot [1, 1]^T$ for $\theta \in [-\pi/2, 3\pi/2]^2$. 

$h$-ellipticity

![h-ellipticity diagram](image)
Local Fourier Analysis - 2

Consider grid functions $\varphi(\theta, x) = \exp(i \theta \cdot x/h) \cdot [1, 1]^T$ for $\theta \in [-\pi/2, 3\pi/2]^2$. 

$h$-ellipticity

smoothing factor Vanka
Consider grid functions $\varphi(\theta, x) = \exp(n\theta \cdot x/h) \cdot [1, 1]^T$ for $\theta \in [-\pi/2, 3\pi/2]^2$.

- ineffective ($\mu \gg 1$) for large compression / expansion of tissue
Consider grid functions $\varphi(\theta, x) = \exp(i\theta \cdot x/h) \cdot [1, 1]^T$ for $\theta \in [-\pi/2, 3\pi/2]^2$.

- ☁ ineffective ($\mu \gg 1$) for large compression / expansion of tissue
- ☀ effective ($\mu \approx 0.14$) for small compression/expansion and well chosen $\alpha_l$
Coloured Vanka Smoother

Hessian

E. Haber, R. Horesh, J. Modersitzki
Numerical Optimization for Constrained Image Registration.
LR, C. Greif, J. Modersitzki
A Multigrid Solver for Hyperelastic Image Registration.
Coloured Vanka Smoother

Breaks coupling $\implies$ parallel computing

E. Haber, R. Horesh, J. Modersitzki
*Numerical Optimization for Constrained Image Registration.*

LR, C. Greif, J. Modersitzki
*A Multigrid Solver for Hyperelastic Image Registration.*
Coloured Vanka Smoother

Hessian

color scheme

Hessian for yellow blocks

Breaks coupling $\leadsto$ parallel computing


Experimental Smoothing Analysis

- 2D registration (Disc → C)
- compute transformation $y$ with $\det \nabla y \in [0.05, 3.20]$
- build $H(y)$ and perform one smoothing step for grid functions $\varphi(\theta, x)$

![Reference Data Template + Grid Transformed Template](image)
Experimental Smoothing Analysis

- 2D registration (Disc $\rightarrow$ C)
- compute transformation $y$ with $\det \nabla y \in [0.05, 3.20]$
- build $H(y)$ and perform one smoothing step for grid functions $\varphi(\theta, x)$

![Diagram of registration results for a 2D academic test problem of transforming a disc to a C-shaped object as proposed by Christensen [6]. As guaranteed by the theory, the hyperelastic regularization problem yields an invertible transformation that introduces large volume changes ($\det r_y^2 \in [0.39, 5.22]$).]

![Impact of the stabilization proposed in Section 3.4 on experimentally determined smoothing factors. Smoothing factors for grid functions at different frequencies are shown for the original Hessian (left) and the stabilized Hessian (right) at a Gauss-Newton iteration at which the current iterate, $y_k$, introduces large volume changes ($\det r_{y_k}^2 \in [0.05, 3.20]$). While the Vanka smoother is ineffective for the original Hessian ($\mu = 1.26$), it is effectively stabilized for the hyperelastic regularization problem with $\mu = 0.71$.]

Example 4.1. To illustrate the challenging effect of large compressions and expansions we first consider the 2D test problem of transforming a disc to a C-shaped object. We use the SSD distance measure (2.2) and perform a multi-level registration using four levels with base meshes of size $16^2$, $32^2$, $62^2$, and $128^2$. We use both a direct solver and the proposed Multigrid PCG scheme with a stabilized Hessian ($s = 100$ in (3.12)). When using the direct method, we needed to perform 42, 54, 4, and 9 iterations on the respective levels and reduced the distance measure from 100% to 0.31%. A comparable number of iterations was required for the proposed multigrid method with stabilized Hessian (40, 8, 5, and 5) yielding an almost identical reduction of the distance measure.
Experimental Smoothing Analysis

- 2D registration ($\text{Disc} \rightarrow \text{C}$)
- Compute transformation $y$ with $\det \nabla y \in [0.05, 3.20]$
- Build $H(y)$ and perform one smoothing step for grid functions $\varphi(\theta, x)$

$$\log_{10}(\min(\psi''(\det \nabla y)), 100\alpha_l/\alpha_v)$$
Experimental Smoothing Analysis

- 2D registration (Disc → C)
- compute transformation $y$ with $\det \nabla y \in [0.05, 3.20]$
- build $H(y)$ and perform one smoothing step for grid functions $\varphi(\theta, x)$

$$\log_{10}(\min(\psi''(\det \nabla y)), 100\alpha_l/\alpha_v)$$

smoothing factor, $\mu = 0.71$
Experimental Smoothing Analysis

- 2D registration (Disc $\rightarrow$ C)
- compute transformation $y$ with $\det \nabla y \in [0.05, 3.20]$
- build $H(y)$ and perform one smoothing step for grid functions $\varphi(\theta, x)$

![registration result](image)

$\log_{10}(\min(\psi''(\det \nabla y)), 100\alpha_l/\alpha_v)$

smoothing factor, $\mu = 0.71$

☀ Effective stabilization. No dramatic increase in outer iterations!
Hyperelastic Registration of a Bending Knee

Large-scale test problem \( \approx 6 \) million unknowns
Hyperelastic Registration of a Bending Knee

Large-scale test problem \( \approx 6 \) million unknowns

reference  template  final template

slice projections
Hyperelastic Registration of a Bending Knee

Large-scale test problem $\approx 6$ million unknowns

- Slice projections
- Reference
- Template
- Final template
- Absolute difference
### Hyperelastic Registration of a Bending Knee - 2

<table>
<thead>
<tr>
<th>GIter</th>
<th>Jacobi-CG</th>
<th>MG-CG</th>
<th>stabilized MG-CG</th>
<th>stab MG-CG (block Jacobi)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J #iter relres</td>
<td>J #iter relres</td>
<td>J #iter relres</td>
<td>J #iter relres</td>
</tr>
<tr>
<td>coarse level</td>
<td>-1</td>
<td>8.7e7</td>
<td>8.7e7</td>
<td>8.7e7</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1.3e7</td>
<td>1.3e7</td>
<td>1.3e7</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6.1e6 41 9.94e-3</td>
<td>6.1e6 2 8.93e-3</td>
<td>6.1e6 2 8.93e-3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.4e6 86 7.77e-3</td>
<td>5.4e6 3 7.80e-3</td>
<td>5.4e6 2 7.79e-3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.3e6 71 9.97e-3</td>
<td>5.3e6 3 7.41e-3</td>
<td>5.3e6 2 8.95e-3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5.2e6 62 9.58e-3</td>
<td>5.2e6 3 1.31e-3</td>
<td>5.2e6 2 8.65e-3</td>
</tr>
<tr>
<td>medium level</td>
<td>-1</td>
<td>1.2e8</td>
<td>1.2e8</td>
<td>1.2e8</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1.0e7</td>
<td>1.0e7</td>
<td>1.0e7</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>8.0e6 37 9.49e-3</td>
<td>8.0e6 2 5.38e-3</td>
<td>8.0e6 2 5.37e-3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.8e6 56 8.49e-3</td>
<td>6.8e6 3 5.61e-3</td>
<td>7.2e6 2 6.35e-3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6.6e6 70 9.52e-3</td>
<td>6.6e6 4 8.45e-3</td>
<td>6.9e6 2 7.63e-3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6.5e6 102 9.65e-3</td>
<td>6.5e6 7 9.47e-3</td>
<td>6.6e6 2 9.00e-3</td>
</tr>
<tr>
<td>fine level</td>
<td>-1</td>
<td>1.5e8</td>
<td>1.5e8</td>
<td>1.5e8</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>9.5e6</td>
<td>9.5e6</td>
<td>9.5e6</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>8.3e6 51 9.49e-3</td>
<td>8.3e6 2 7.79e-3</td>
<td>8.3e6 2 6.41e-3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.0e6 96 9.98e-3</td>
<td>8.0e6 3 7.55e-3</td>
<td>7.9e6 2 8.07e-3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8.0e6 96 9.88e-3</td>
<td>8.0e6 3 6.05e-3</td>
<td>7.8e6 3 2.43e-03</td>
</tr>
</tbody>
</table>

- Stabilized multigrid preconditioner reduces number of CG-iterations by a factor of \( \approx 30 \).  
- Vanka step is about 20 times more expensive than Jacobi per iteration!
Hyperelastic Registration of a Bending Knee - 2

<table>
<thead>
<tr>
<th>GNiter</th>
<th>Jacobi-CG</th>
<th>MG-CG</th>
<th>stabilized MG-CG</th>
<th>stab MG-CG (block Jacobi)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J #iter relres</td>
<td>J #iter relres</td>
<td>J #iter relres</td>
<td>J #iter relres</td>
</tr>
<tr>
<td>coarse level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>8.7e7</td>
<td>8.7e7</td>
<td>8.7e7</td>
<td>8.7e7</td>
</tr>
<tr>
<td>0</td>
<td>1.3e7</td>
<td>1.3e7</td>
<td>1.3e7</td>
<td>1.3e7</td>
</tr>
<tr>
<td>1</td>
<td>6.1e6 41 9.94e-3</td>
<td>6.1e6 2 8.93e-3</td>
<td>6.1e6 2 8.93e-3</td>
<td>6.0e6 3 9.99e-3</td>
</tr>
<tr>
<td>2</td>
<td>5.4e6 86 7.77e-3</td>
<td>5.4e6 3 7.80e-3</td>
<td>5.4e6 2 7.79e-3</td>
<td>5.4e6 4 4.67e-3</td>
</tr>
<tr>
<td>3</td>
<td>5.3e6 71 9.97e-3</td>
<td>5.3e6 3 7.41e-3</td>
<td>5.3e6 2 8.95e-3</td>
<td>5.2e6 4 3.00e-3</td>
</tr>
<tr>
<td>4</td>
<td>5.2e6 62 9.58e-3</td>
<td>5.2e6 3 1.31e-3</td>
<td>5.2e6 2 8.65e-3</td>
<td>5.2e6 4 2.62e-3</td>
</tr>
<tr>
<td>medium level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.2e8</td>
<td>1.2e8</td>
<td>1.2e8</td>
<td>1.2e8</td>
</tr>
<tr>
<td>0</td>
<td>1.0e7</td>
<td>1.0e7</td>
<td>1.0e7</td>
<td>1.0e7</td>
</tr>
<tr>
<td>1</td>
<td>8.0e6 37 9.49e-3</td>
<td>8.0e6 2 5.38e-3</td>
<td>8.0e6 2 5.37e-3</td>
<td>8.0e6 3 7.32e-3</td>
</tr>
<tr>
<td>2</td>
<td>6.8e6 56 8.49e-3</td>
<td>6.8e6 3 5.61e-3</td>
<td>7.2e6 2 6.35e-3</td>
<td>7.2e6 3 8.14e-3</td>
</tr>
<tr>
<td>3</td>
<td>6.6e6 70 9.52e-3</td>
<td>6.6e6 4 8.45e-3</td>
<td>6.9e6 2 7.63e-3</td>
<td>6.6e6 4 3.92e-3</td>
</tr>
<tr>
<td>4</td>
<td>6.5e6 102 9.65e-3</td>
<td>6.5e6 7 9.47e-3</td>
<td>6.6e6 2 9.00e-3</td>
<td>6.5e6 4 4.88e-3</td>
</tr>
<tr>
<td>fine level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.5e8</td>
<td>1.5e8</td>
<td>1.5e8</td>
<td>1.5e8</td>
</tr>
<tr>
<td>0</td>
<td>9.5e6</td>
<td>9.5e6</td>
<td>9.5e6</td>
<td>9.4e6</td>
</tr>
<tr>
<td>1</td>
<td>8.3e6 51 9.49e-3</td>
<td>8.3e6 2 7.79e-3</td>
<td>8.3e6 2 6.41e-3</td>
<td>8.2e6 3 8.01e-3</td>
</tr>
<tr>
<td>2</td>
<td>8.0e6 96 9.98e-3</td>
<td>8.0e6 3 7.55e-3</td>
<td>7.9e6 2 8.07e-3</td>
<td>8.0e6 4 6.36e-3</td>
</tr>
<tr>
<td>4</td>
<td>8.0e6 96 9.88e-3</td>
<td>8.0e6 3 6.05e-3</td>
<td>7.8e6 3 2.43e-03</td>
<td></td>
</tr>
</tbody>
</table>

Stabilized multigrid preconditioner reduces number of CG-iterations by a factor of $\approx 30^\ast$
## Hyperelastic Registration of a Bending Knee - 2

<table>
<thead>
<tr>
<th>GNiter</th>
<th>Jacobi-CG</th>
<th>MG-CG</th>
<th>stabilized MG-CG</th>
<th>stab MG-CG (block Jacobi)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J #iter relres</td>
<td>J #iter relres</td>
<td>J #iter relres</td>
<td>J #iter relres</td>
</tr>
<tr>
<td>coarse level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>8.7e7</td>
<td>8.7e7</td>
<td>8.7e7</td>
<td>8.7e7</td>
</tr>
<tr>
<td>0</td>
<td>1.3e7</td>
<td>1.3e7</td>
<td>1.3e7</td>
<td>1.3e7</td>
</tr>
<tr>
<td>1</td>
<td>6.1e6 41 9.94e-3</td>
<td>6.1e6 2 8.93e-3</td>
<td>6.1e6 2 8.93e-3</td>
<td>6.0e6 3 9.99e-3</td>
</tr>
<tr>
<td>2</td>
<td>5.4e6 86 7.77e-3</td>
<td>5.4e6 3 7.80e-3</td>
<td>5.4e6 2 7.79e-3</td>
<td>5.4e6 4 4.67e-3</td>
</tr>
<tr>
<td>3</td>
<td>5.3e6 71 9.97e-3</td>
<td>5.3e6 3 7.41e-3</td>
<td>5.3e6 2 8.95e-3</td>
<td>5.2e6 4 3.00e-3</td>
</tr>
<tr>
<td>4</td>
<td>5.2e6 62 9.58e-3</td>
<td>5.2e6 3 1.31e-3</td>
<td>5.2e6 2 8.65e-3</td>
<td>5.2e6 4 2.62e-3</td>
</tr>
<tr>
<td>medium level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.2e8</td>
<td>1.2e8</td>
<td>1.2e8</td>
<td>1.2e8</td>
</tr>
<tr>
<td>0</td>
<td>1.0e7</td>
<td>1.0e7</td>
<td>1.0e7</td>
<td>1.0e7</td>
</tr>
<tr>
<td>1</td>
<td>8.0e6 37 9.49e-3</td>
<td>8.0e6 2 5.38e-3</td>
<td>8.0e6 2 5.37e-3</td>
<td>8.0e6 3 7.32e-3</td>
</tr>
<tr>
<td>2</td>
<td>6.8e6 56 8.49e-3</td>
<td>6.8e6 3 5.61e-3</td>
<td>7.2e6 2 6.35e-3</td>
<td>7.2e6 3 8.14e-3</td>
</tr>
<tr>
<td>3</td>
<td>6.6e6 70 9.52e-3</td>
<td>6.6e6 4 8.45e-3</td>
<td>6.9e6 2 7.63e-3</td>
<td>6.6e6 4 3.92e-3</td>
</tr>
<tr>
<td>4</td>
<td>6.5e6 102 9.65e-3</td>
<td>6.5e6 7 9.47e-3</td>
<td>6.6e6 2 9.00e-3</td>
<td>6.5e6 4 4.88e-3</td>
</tr>
<tr>
<td>fine level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.5e8</td>
<td>1.5e8</td>
<td>1.5e8</td>
<td>1.5e8</td>
</tr>
<tr>
<td>0</td>
<td>9.5e6</td>
<td>9.5e6</td>
<td>9.5e6</td>
<td>9.4e6</td>
</tr>
<tr>
<td>1</td>
<td>8.3e6 51 9.49e-3</td>
<td>8.3e6 2 7.79e-3</td>
<td>8.3e6 2 6.41e-3</td>
<td>8.2e6 3 8.01e-3</td>
</tr>
<tr>
<td>2</td>
<td>8.0e6 96 9.98e-3</td>
<td>8.0e6 3 7.55e-3</td>
<td>7.9e6 2 8.07e-3</td>
<td>8.0e6 4 6.36e-3</td>
</tr>
<tr>
<td>4</td>
<td>8.0e6 96 9.88e-3</td>
<td>8.0e6 3 6.05e-3</td>
<td>7.8e6 3 2.43e-03</td>
<td></td>
</tr>
</tbody>
</table>

Stabilized multigrid preconditioner reduces number of CG-iterations by a factor of \( \approx 30^* \)

* Vanka step is about 20 times more expensive than Jacobi per iteration!
Image Registration as Optimal Control Problem
Optimal Control Formulation of Image Registration

Optimal Control Problem

Given a template image $\mathcal{T}$ and a reference image $\mathcal{R}$ find a smooth velocity field $\nu$, such that the final state $u(\cdot, 1)$ is similar to $\mathcal{R}$, i.e., solve

$$\minimize_{\nu, u} \text{distance}[u(\cdot, 1), \mathcal{R}] + \text{regularizer}[\nu]$$

subject to $u(\cdot, 0) = \mathcal{T}$ and

$$\begin{cases} \partial_t u + \nu \cdot \nabla u = 0, \\ \partial_t u + \nabla \cdot (uv) = 0, \end{cases}$$

intensity preservation

mass-preservation.
Optimal Control Formulation of Image Registration

Given a template image $\mathcal{T}$ and a reference image $\mathcal{R}$ find a smooth velocity field $v$, such that the final state $u(\cdot, 1)$ is similar to $\mathcal{R}$, i.e., solve

$$\begin{align*}
\text{minimize}_{v, u} & \quad \text{distance}[u(\cdot, 1), \mathcal{R}] + \text{regularizer}[v] \\
\text{subject to} & \quad u(\cdot, 0) = \mathcal{T} \quad \text{and} \quad \begin{cases} 
\partial_t u + v \cdot \nabla u = 0, \\
\partial_t u + \nabla \cdot (uv) = 0,
\end{cases}
\end{align*}$$

intensity preservation

mass-preservation
Regularizing the Flow Field

For non stationary velocity fields, consider $H^1$ regularizer

$$S_{\text{diff}}(v) = \frac{1}{2} \int_0^1 \int_\Omega \sum_{k=1}^d |\nabla v_k(x, t)|^2 dx dt$$

or $H^2$ regularizer

$$S_{\text{curv}}(v) = \frac{1}{2} \int_0^1 \int_\Omega \sum_{k=1}^d |\Delta v_k(x, t)|^2 dx dt$$
Regularizing the Flow Field

For non stationary velocity fields, consider $H^1$ regularizer

\[ S^{\text{diff}}(v) = \frac{1}{2} \int_0^1 \int_{\Omega} \sum_{k=1}^d |\nabla v_k(x, t)|^2 dx dt \]

or $H^2$ regularizer

\[ S^{\text{curv}}(v) = \frac{1}{2} \int_0^1 \int_{\Omega} \sum_{k=1}^d |\Delta v_k(x, t)|^2 dx dt \]

Remarks:

- smoothness of transformation also depends on regularity of images
- regularizers can be shown to be sufficient to yield diffeomorphisms
- regularizers are quadratic and separable w.r.t. components in $v$.


Illustration: Lagrangian Methods

Let $I(\cdot) \in \mathbb{R}^{m \times n}$ be an interpolation matrix ($I \geq 0$, sparse, rows sum to one), $y_0$ be fixed(!) grid point, $v : \Omega \times [0, 1] \rightarrow \mathbb{R}^d$ velocity

Advection, $u(y_1, 1) = I(y(v, y_1, 1, 0))u_0$
Illustration: Lagrangian Methods

Let $I(\cdot) \in \mathbb{R}^{m \times n}$ be an interpolation matrix ($I \geq 0$, sparse, rows sum to one), $y_0$ be fixed(!) grid point, $v : \Omega \times [0, 1] \to \mathbb{R}^d$ velocity

advection, $u(y_1, 1) = I(y(v, y_1, 1, 0))u_0$

continuity, $u(y_0, 1) = I(y(v, y_0, 0, 1))^T u_0$
Illustration: Lagrangian Methods

Let \( I(\cdot) \in \mathbb{R}^{m \times n} \) be an interpolation matrix (\( I \geq 0 \), sparse, rows sum to one), \( y_0 \) be fixed(!) grid point, \( v : \Omega \times [0, 1] \rightarrow \mathbb{R}^d \) velocity

Advection, \( u(y_1, 1) = I(y(v, y_1, 1, 0))u_0 \)

Continuity, \( u(y_0, 1) = I(y(v, y_0, 0, 1))^\top u_0 \)

J. Fohring, E. Haber, LR.
Example: Disc-to-C - Performance of Preconditioner

setup: coarse mesh \((32 \times 32)\) \(v\) non stationary (3 time intervals), diffusion regularizer, \(\alpha = 4000\), RK4 with 10 time steps

\[\mathcal{T} + y\]

Sparsity of \(H(v)\)

PCG convergence

- CG
- PCG-Jac
- PCG-SGS
- PCG-Spec
Summary
Summary: Multigrid for Hyperelastic Image Registration

Features

▶ physically motivated
▶ existence and (local) invertibility of solutions
▶ allows for large deformations

LR, C. Greif, J. Modersitzki
A Multigrid Solver for Hyperelastic Image Registration.
Summary: Multigrid for Hyperelastic Image Registration

Features

▶ physically motivated
▶ existence and (local) invertibility of solutions
▶ allows for large deformations

Computational Challenges

▶ mimetic discretization → computationally very intense
▶ tightly coupled linear system with locally varying coefficients
▶ smoothing analysis: poor smoothing quality for large volume changes
▶ stabilization of Hessian

LR, C. Greif, J. Modersitzki
Summary: Multigrid for Hyperelastic Image Registration

Features

▶ physically motivated
▶ existence and (local) invertibility of solutions
▶ allows for large deformations

Computational Challenges

▶ mimetic discretization ⇞ computationally very intense
▶ tightly coupled linear system with locally varying coefficients
▶ smoothing analysis: poor smoothing quality for large volume changes
▶ ⚪ stabilization of Hessian

LR, C. Greif, J. Modersitzki
A Multigrid Solver for Hyperelastic Image Registration.