

# Regularization in Deformable Registration

Marius Staring  
MICCAI Registration Tutorial  
24 September 2010, Beijing

# Overview

- Why?
- How?
- Popular terms + extensive example
- Miscellaneous remarks
- So?

Why ?

# Why?

- Image registration is ill-posed
- Transformation should be plausible
  - Problems: singularities, folding, invertibility, non-smoothness
- You want to!
  - Tissue characteristics
  - Application wishes

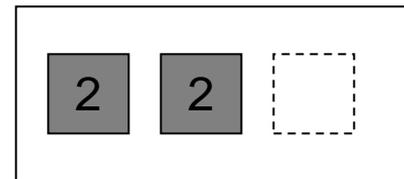
# ill/well-posedness

- Well-posedness (Hadamard):
  - A solution exists
  - The solution is unique
  - The solution depends continuously on the data

fixed

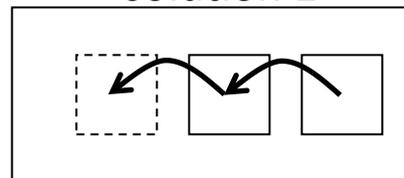


moving



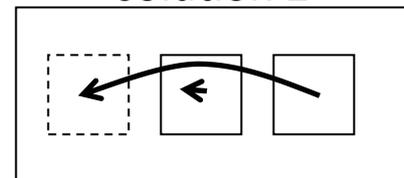
Uniqueness:

solution 1



$$|J|=1$$

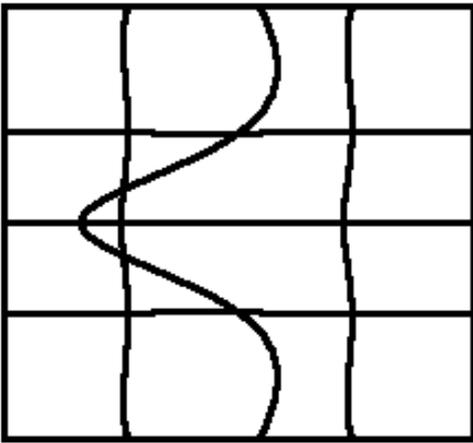
solution 2



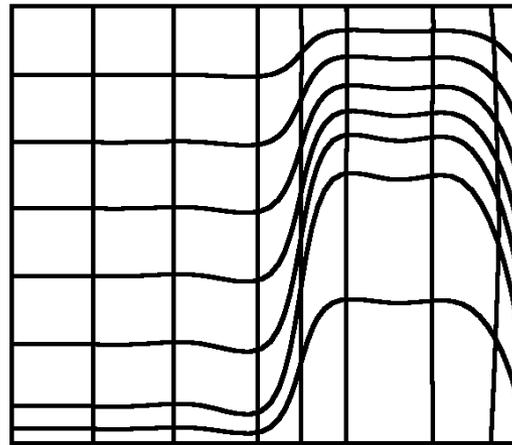
$$|J|=-2$$

# (im)plausible T

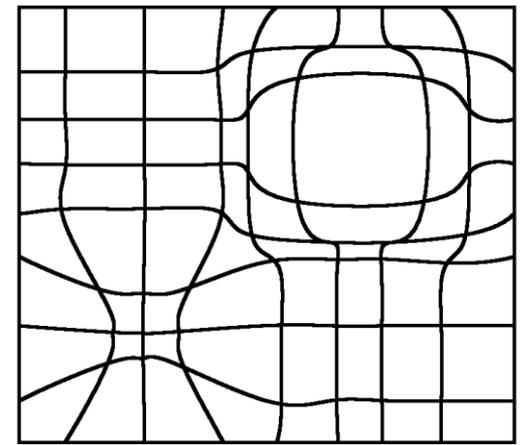
folding



irregular



expansion / compression



# Why?

- Tissue characteristics
  - Some tissue bends easier than others
  - Bone is rigid
  - Breast tissue does not compress easily
- Application wishes
  - Keep nodule rigid to visualize growth (not a physical reason)
  - Structures that appear only in one image are reduced in size, e.g. DSA data

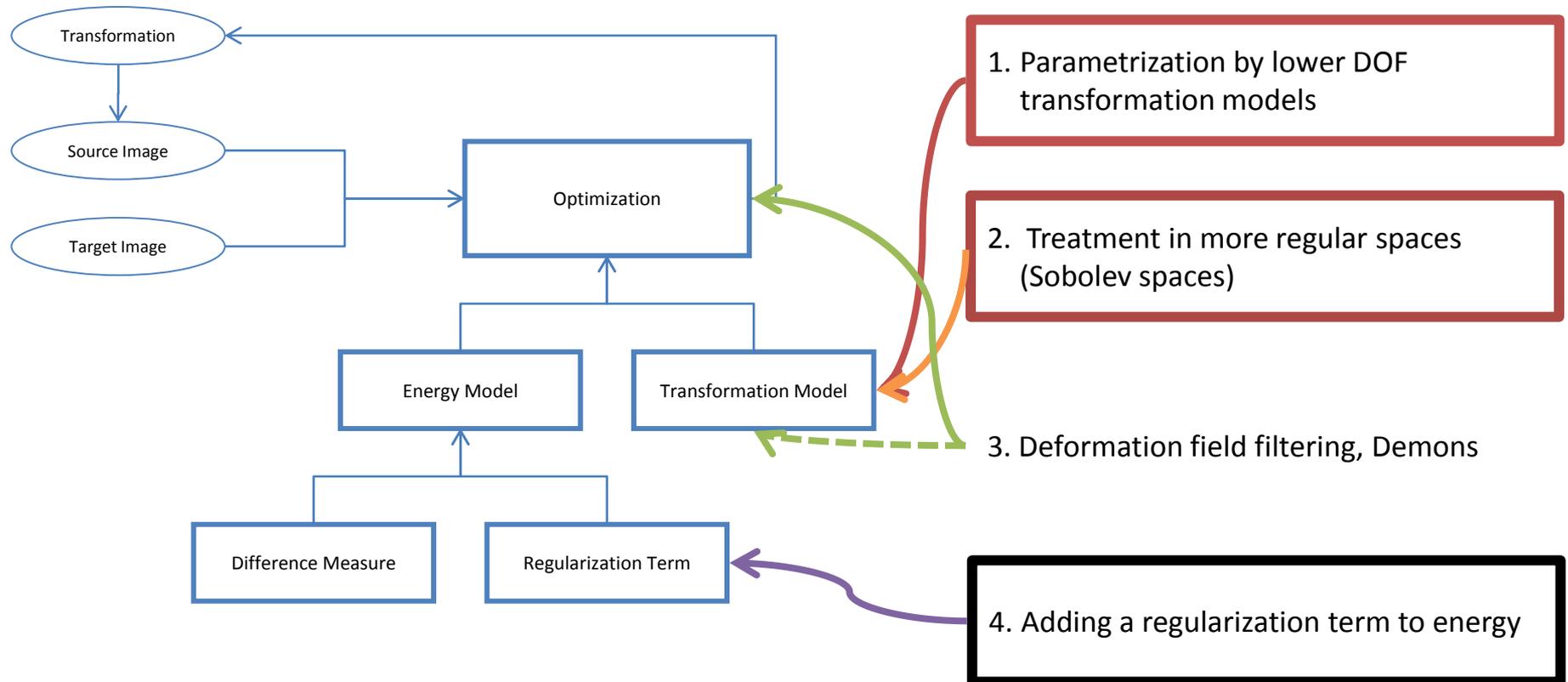
# How ?

# How?

1. Restricting the space of  $T$ , by parameterization
2. Restricting the space of  $T$ , by looking at special (Sobolev) spaces
3. By deformation field filtering
4. By penalty terms / regularizers

Implicit and explicit regularization

# Regularization Strategies: overview

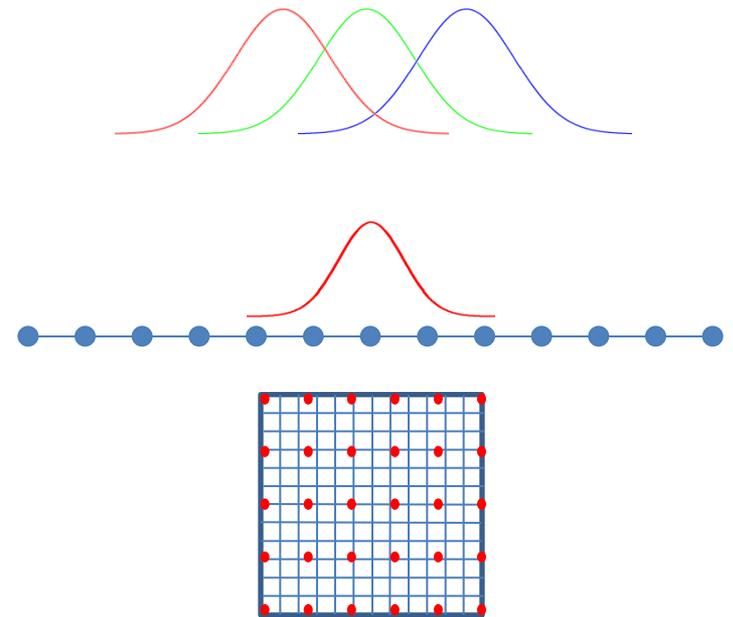


# 1: Parameterization

- Transformation using linear models
  - **Free Form Deformations (FFD), B-splines**
  - Trigonometric Functions  
(Discrete Fourier/Cosine Transformation Bases)
  - Radial Basis Functions (RBF)

# B-spline parametrization

- **Rueckert 1999**
- **Cubic B-splines** (degree  $D = 3$ ), basis functions all have same shape and are translated versions of each other
- **Compact support** of  $D+1$  control points
- **Regular grid** of control points
- **More:** piecewise polynomial, inherent smoothness, differentiability, hierarchical



# B-spline parametrization

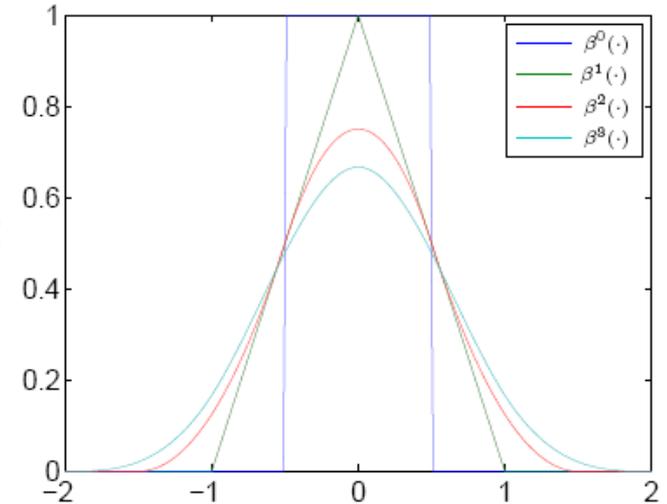
- *Deformation* is modeled by B-splines

$$\mathbf{T}_\mu = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} u_1(\mathbf{x}) \\ u_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \sum_i \mu_{i1} \beta^3 (x_1 - y_{i1}) \beta^3 (x_2 - y_{i2}) \\ \sum_i \mu_{i2} \beta^3 (x_1 - y_{i1}) \beta^3 (x_2 - y_{i2}) \end{bmatrix}$$

- $\mathbf{T} \Rightarrow \mathbf{T}_\mu$
- $\arg \min_{\mathbf{T}} \mathcal{C}(I_F, I_M, \mathbf{T}) \Rightarrow \arg \min_{\mu} \mathcal{C}(I_F, I_M, \mathbf{T}_\mu)$

# Basis functions

- $$\beta^3(x) = \begin{cases} -|x|^2 + \frac{1}{2}|x|^3 + \frac{2}{3} & 0 \leq |x| \leq 1 \\ \frac{1}{6}(2 - |x|)^3 & 1 \leq |x| \leq 2 \\ 0 & \text{otherwise} \end{cases}$$



- ## Properties:

- $\beta^3$  is twice differentiable  $\rightarrow$  so is  $T_\mu$

- Convolution: 
$$\beta^n(x) = \underbrace{\beta^0 * \dots * \beta^0}_{n+1 \text{ times}}(x)$$

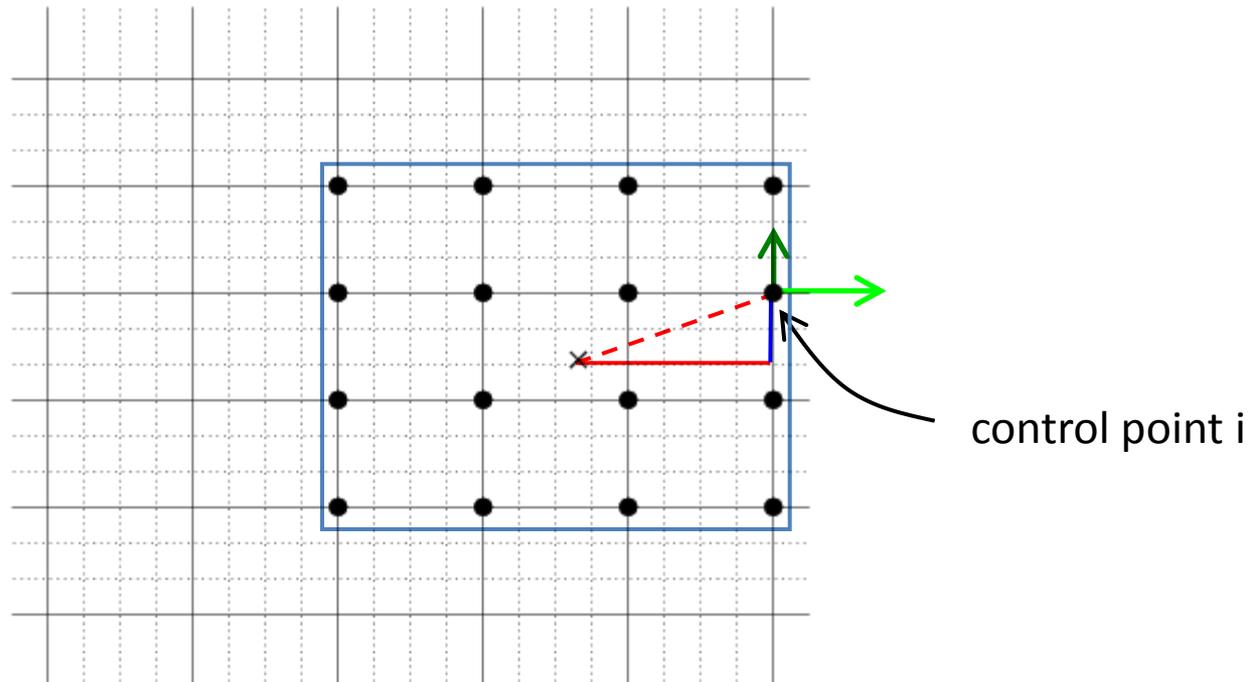
- Derivative can be done in terms of B-splines too:

$$\frac{d\beta^n(x)}{dx} = \beta^{n-1}\left(x + \frac{1}{2}\right) - \beta^{n-1}\left(x - \frac{1}{2}\right)$$

# B-splines practically

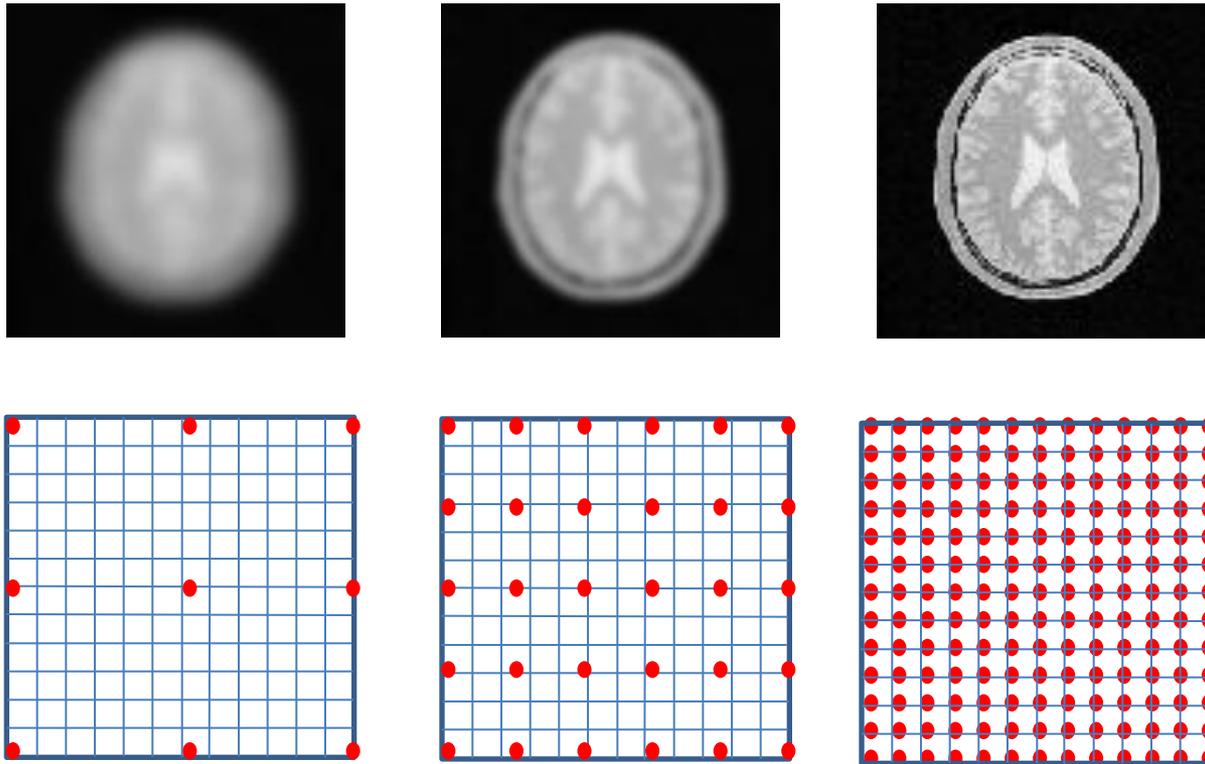
- control point
- × world coordinate x
- support region S(x)

$$\mathbf{T}_\mu = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \sum_i \underline{\mu_{i1}} \beta^3 (\underline{x_1 - y_{i1}}) \beta^3 (\underline{x_2 - y_{i2}}) \\ \sum_i \underline{\mu_{i2}} \beta^3 (\underline{x_1 - y_{i1}}) \beta^3 (\underline{x_2 - y_{i2}}) \end{bmatrix}$$



# B-spline parametrization

- Multi-resolution: coarse to fine



# Optimization with B-splines

- $\arg \min_{\boldsymbol{\mu}} \mathcal{C}(I_F, I_M, \mathbf{T}_{\boldsymbol{\mu}}), \quad \boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - a_k \frac{\partial \mathcal{C}}{\partial \boldsymbol{\mu}_k}$

- $\frac{\partial \mathcal{C}}{\partial \boldsymbol{\mu}} = f \left( \dots, \frac{\partial I_M}{\partial \mathbf{x}}, \frac{\partial \mathbf{T}}{\partial \boldsymbol{\mu}} \right)$

- $\mathbf{T}_{\boldsymbol{\mu}}(\mathbf{x}) = f(\boldsymbol{\mu}, \beta^3) \Rightarrow \frac{\partial \mathbf{T}}{\partial \boldsymbol{\mu}} = f(\beta^3)$

- **Penalty terms:**  $P = f \left( \frac{\partial \mathbf{T}}{\partial \mathbf{x}}, \frac{\partial^2 \mathbf{T}}{\partial \mathbf{x} \partial \mathbf{x}'} \right)$

$$\frac{\partial T_1}{\partial x_2}(\mathbf{x}) = \sum_{y_i \in \mathcal{S}(\mathbf{x})} \mu_{i1} \beta^3(x_1 - y_{i1}) \frac{\partial}{\partial x_2} \beta^3(x_2 - y_{i2})$$

## 2: How: By deformation field filtering

- Framework
- Filter choices
- Rigidity filter example

# Framework

- Demons like

- Thirion 1996/1998
- Original: Get gradient field, take step, filter  
Alternative: Get gradient field, filter, take step

Fluid version  
(Bro-Nielsen 1996, Pennec 1999)

Loop until convergence:

$$f \approx -\nabla E_{\text{SAD}}(u)$$

$$u = u + \tau f$$

$$u = G^\sigma * f$$

end

Loop until convergence:

$$f \approx -\nabla E_{\text{SAD}}(u)$$

$$h = G^\sigma * \tau f$$

$$u = u + h$$

end

# Filter choices

- Gaussian blurring (symmetric kernels)
  - Thirion (1998), with globally fixed std
  - Cachier 1999, Demons, with varying std, independent on tissue type
  - Stefanescu (2004), with varying std dependent on tissue type
- Special purpose filtering
  - Staring 2007, *asymmetric* filtering kernel, especially for rigidity

# Adaptive filtering for rigidity

- Adaptive filter:

$$m(\mathbf{x}) = \sum_j c(\mathbf{x}_j) \mathbf{u}(\mathbf{x}_j) / \sum_j c(\mathbf{x}_j)$$
$$\mathcal{F}(\mathbf{u}(\mathbf{x})) = (1 - c(\mathbf{x})) \mathbf{u}(\mathbf{x}) + c(\mathbf{x}) m(\mathbf{x})$$

$m$  = weighted mean

$u$  = deformation field

$c$  = tissue stiffness coefficient

$i$  = current voxel

$j$  = neighboring voxels

- Properties:
  - Adaptable to tissue type
  - No `cross-talk', e.g. on locations where rigid and nonrigid tissue meet

# Adaptive filtering for rigidity

- Framework

- deformation modeled by B-spline parameters  $\boldsymbol{\mu}^{\text{BS}}$

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{u}^{\text{BS}}(\boldsymbol{x}; \boldsymbol{\mu}^{\text{BS}})$$

- The optimizer searches for incremental parameters  $\boldsymbol{\Delta}_t^{\text{BS}}$

$$\boldsymbol{\mu}_{t+1}^{\text{BS}} = \boldsymbol{\mu}_t^{\text{BS}} + \boldsymbol{\Delta}_t^{\text{BS}}$$

- Insert the filtering step

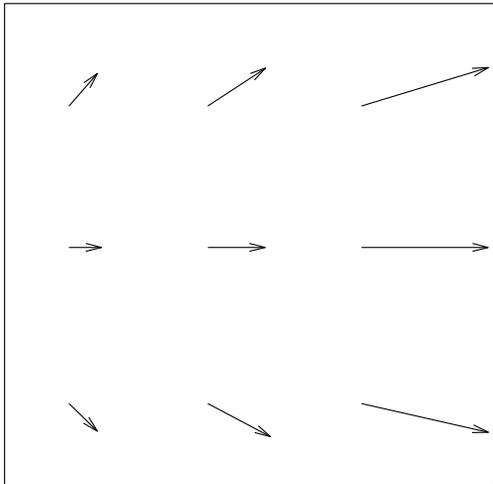
$$\boldsymbol{u}_{t+1}(\boldsymbol{x}) = \boldsymbol{u}_t(\boldsymbol{x}) + \boldsymbol{u}^{\text{BS}}(\boldsymbol{x}; \boldsymbol{\Delta}_t^{\text{BS}})$$

⇓

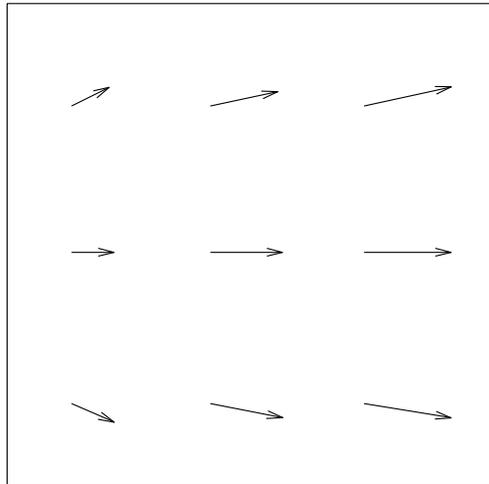
$$\boldsymbol{u}_{t+1}(\boldsymbol{x}) = \mathcal{F}(\boldsymbol{u}_t(\boldsymbol{x}) + \boldsymbol{u}^{\text{BS}}(\boldsymbol{x}; \boldsymbol{\Delta}_t^{\text{BS}}))$$

# Adaptive filter for rigidity

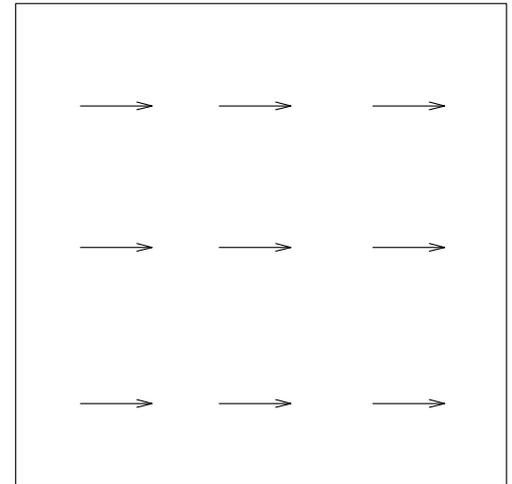
Initial deformation field



After 1 iteration



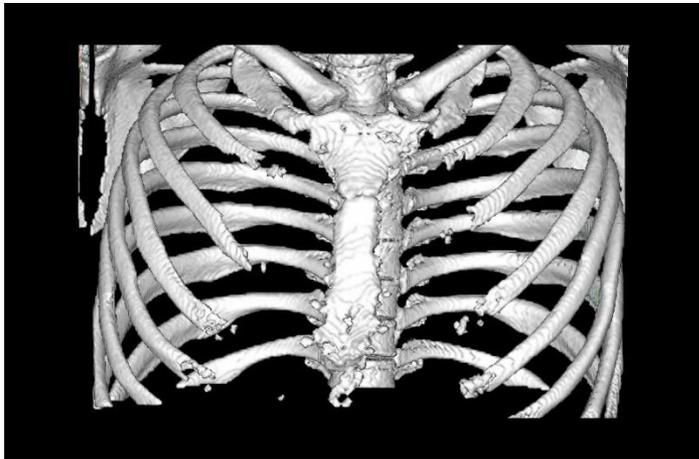
After 50 iterations



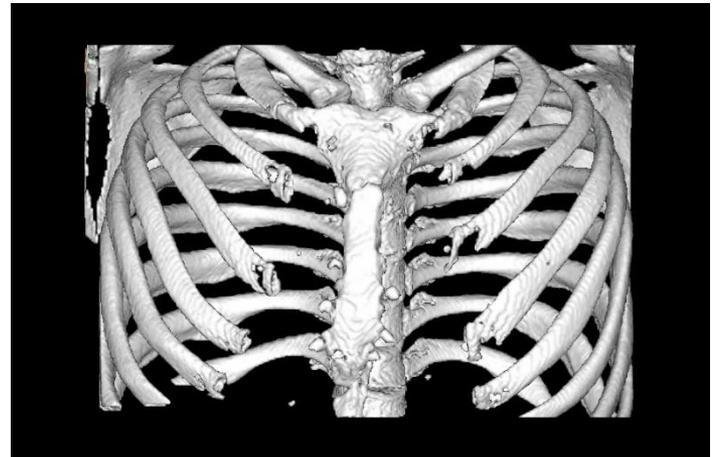
Downside: This adaptive filter converges to a local translation

# On the ribcage

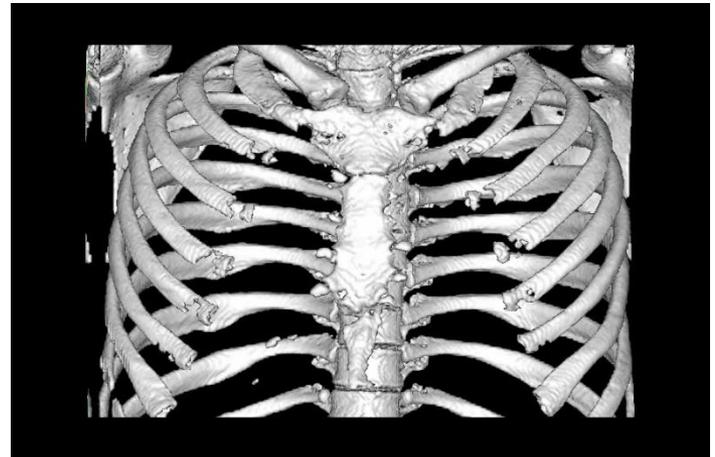
inspiration



B-spline →



B-spline with filtering →



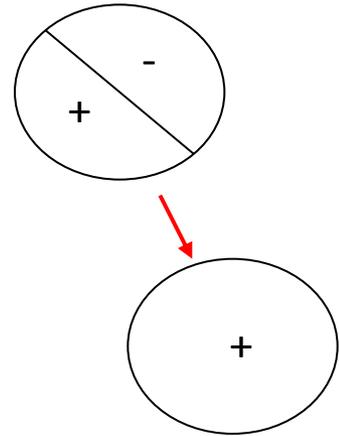
# 3: Sobolev spaces

- Intuition

- Instead of adding regularization to  $T$
- Define  $T$  in a regularized space

- References

- Darko's talk this morning
- Trounev 1998, Diffeomorphisms groups and pattern matching in image analysis
- Chéfd'hotel 2002, Flows of diffeomorphisms for multimodal image registration. And his thesis.



# 4: By penalty terms

- Framework:
  - Soft constraint
  - Hard constraint
- Several choices for  $\mathcal{R}$ 
  - Bending energy, smoothness, linear elasticity
  - Incompressibility
  - Rigidity
  - Invertibility (box constraints for B-spline)
  - Diffusion / Tikhonov / 1st-order
  - Curvature / Bending Energy / 2nd-order, fluid
- Locality
- Examples from own work

# How? Framework

- Soft constraint:

$$\arg \min_{\mathbf{T}} \mathcal{C} = \arg \min_{\mathbf{T}} \mathcal{S}(I_F, I_M, \mathbf{T}) + \alpha \mathcal{R}(\mathbf{T})$$

- Hard constraint:

$$\begin{aligned} \arg \min_{\mathbf{T}} \mathcal{S}(I_F, I_M, \mathbf{T}) \\ \text{s.t. } \mathcal{R}(\mathbf{T}) \end{aligned}$$

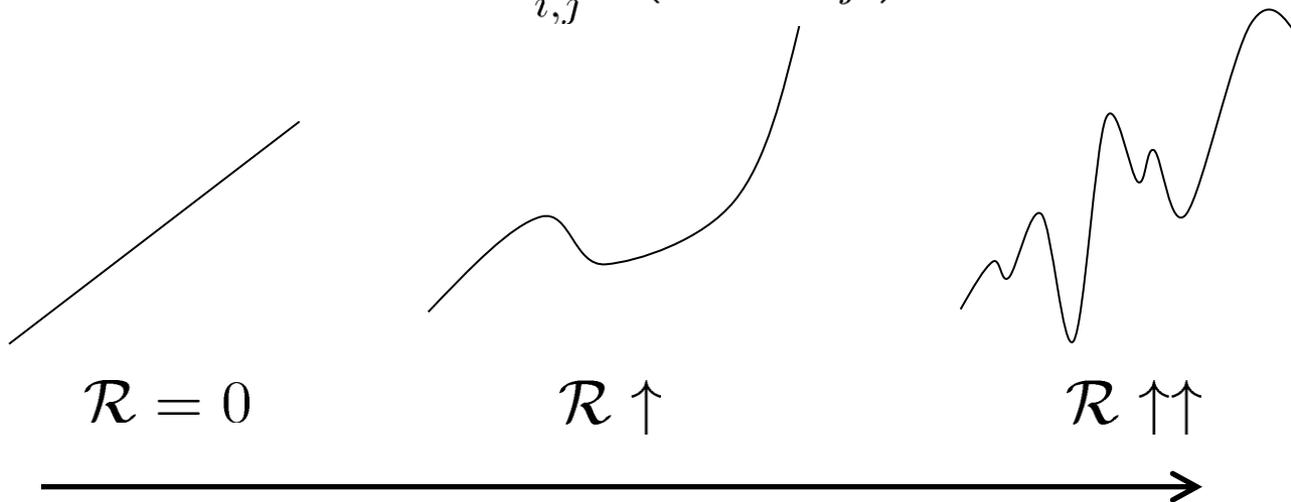
– See e.g. Haber and Modersitzki, 2004.

# Popular regularizers

# Smoothness

- Mathematical interpretation: differentiable

$$\mathcal{R} = \|\nabla_{\mathbf{x}} \mathbf{T}\|^2 = \sum_{i,j} \left( \frac{\partial^2 \mathbf{T}(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 \quad \text{should be small}$$



Also fights folding

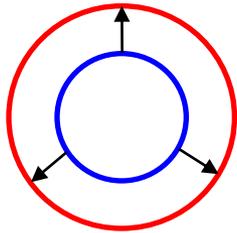
# Jacobian

- If  $J > 0$  then no folding
- How about inversion? By the inverse function theorem, it is sufficient for  $T$  to
  - be continuously differentiable, and
  - have positive Jacobian determinant
- Volume change control (incompressibility), folding
- You can compute volume change from  $T$ !

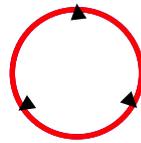
$$\text{vol}(V, \mathbf{T}) \equiv \int_{\mathbf{T}(V)} d\mathbf{y} = \int_V \det(\mathbf{I} + \nabla \mathbf{u}) d\mathbf{x}$$

# Effect of $|J|$ change

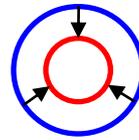
 = volume  $x$      
  = volume  $T(x)$      
  $J = \det \nabla_x T$



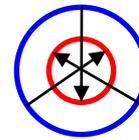
$J > 1$



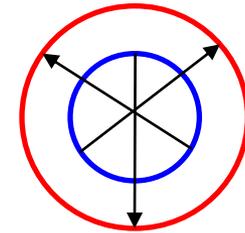
$J = 1$



$0 < J < 1$



$-1 < J < 0$



$J < -1$

- Incompressibility:

$$\mathcal{R} = \sum (\mathbf{J} - 1)^2$$

$$\mathcal{R} = \sum_x \log \mathbf{J} \quad (\text{Rohlfing 2003})$$

- Invertibility:

$$\mathcal{R} = \sum_x \exp(-\mathbf{J}) \quad (\text{Kybic 2000})$$

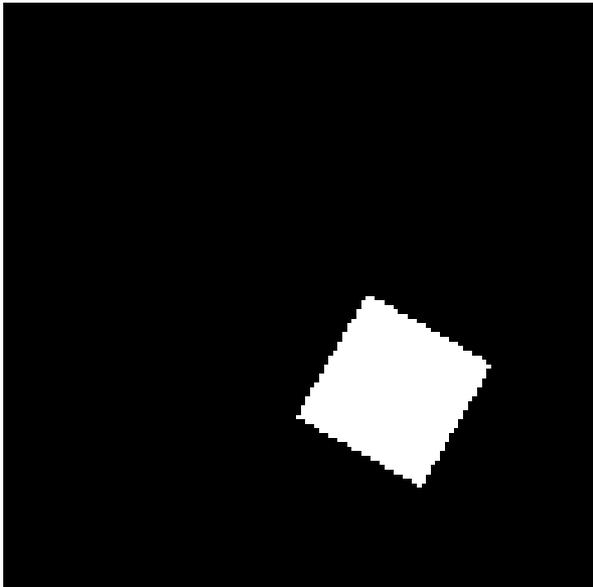
# Others

- Curvature:  $\mathcal{R} = \sum_x \|\Delta \mathbf{u}\|^2$  (Fischer 2004)
- Fluid, diffusion, see a.o. Fischer 2004
- Some work not based on derivatives
  - see Choi 2000, Rueckert 2006, several papers from Fessler et al
  - Control point manipulation
  - Mostly to guarantee invertibility
  - Better performance in speed and memory

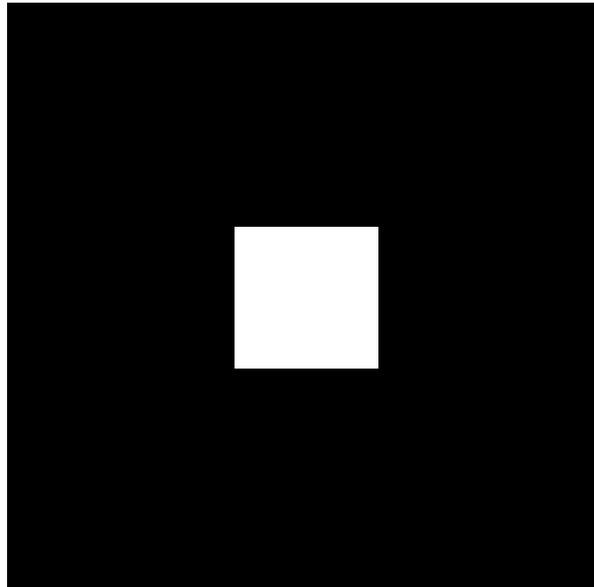
# Big example

# Rigidity penalty – Synthetic example

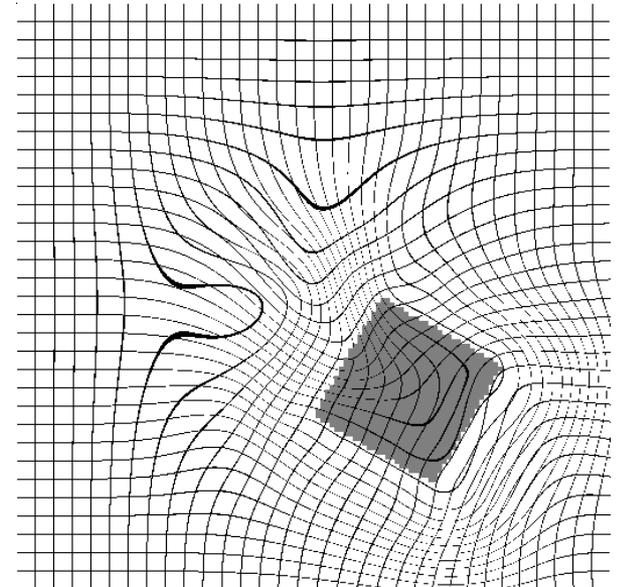
Purely nonrigid registration:



fixed



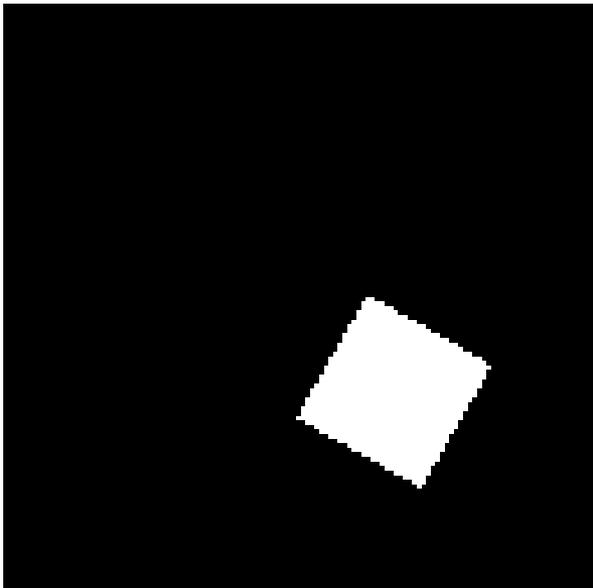
moving



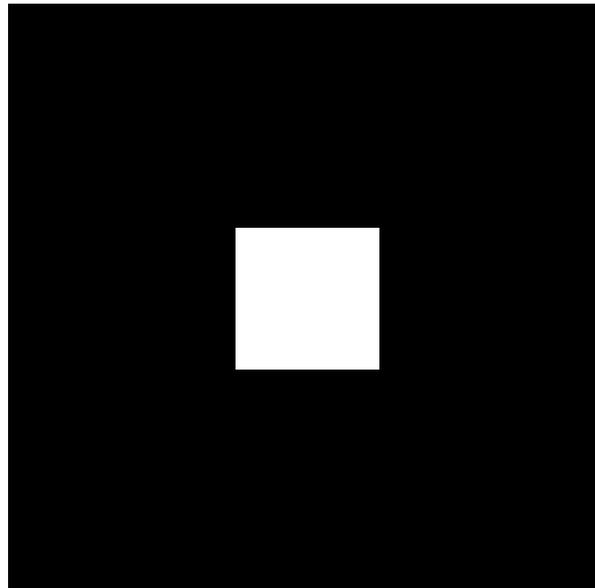
deformed moving

# Rigidity penalty – Synthetic example

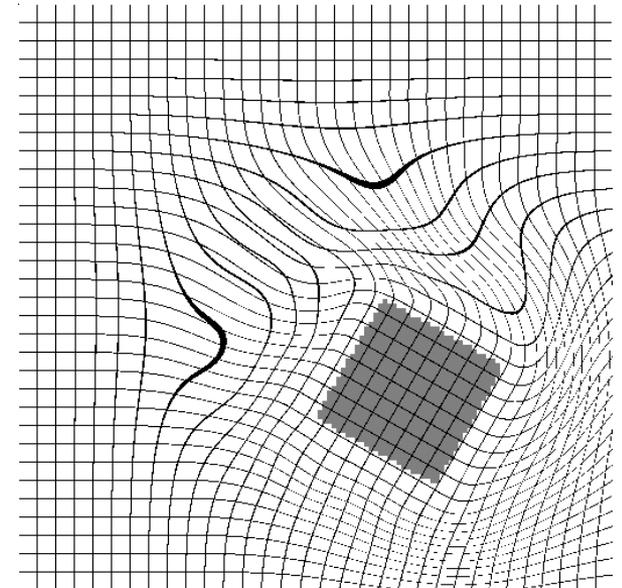
Nonrigid registration with rigidity penalty term:



fixed



moving



deformed moving

# Rigidity penalty – design

- What is rigidity of a deformation?

$$\mathbf{u}(\mathbf{x}) = \mathbf{R}\mathbf{x} - \mathbf{x} + \mathbf{t} \quad \Rightarrow \quad \frac{\partial u_i}{\partial x_j} = r_{ij} - \delta_{ij}$$

- 3 conditions:
  - $\mathbf{u}$  is affine
  - orthonormality of  $\mathbf{R}$
  - properness of  $\mathbf{R}$

# Rigidity penalty – design

- Affinity: second order derivatives are zero

$$AC_{kij} : \frac{\partial^2 u_k}{\partial x_i \partial x_j} = 0 \quad k, i, j = 1, 2$$

- Properness: rotation matrix  $\mathbf{R}$  is proper:

$$PC : \det \mathbf{R} - 1 = 0$$

⇓

$$r_{22} - r_{12}r_{21} - 1 = 0$$

# Rigidity penalty – design

- Rotation matrix  $R$  is orthonormal:

$$\text{OC}_{ij} : \sum_{k=1}^2 r_{ki} r_{kj} = \delta_{ij} \quad \forall i, j = 1, 2$$

$$\text{OC}_{ij} : \sum_{k=1}^2 \left( \frac{\partial u_k}{\partial x_i} + \delta_{ki} \right) \left( \frac{\partial u_k}{\partial x_j} + \delta_{kj} \right) - \delta_{ij} = 0 \quad \forall i, j = 1, 2$$

# Rigidity penalty – design

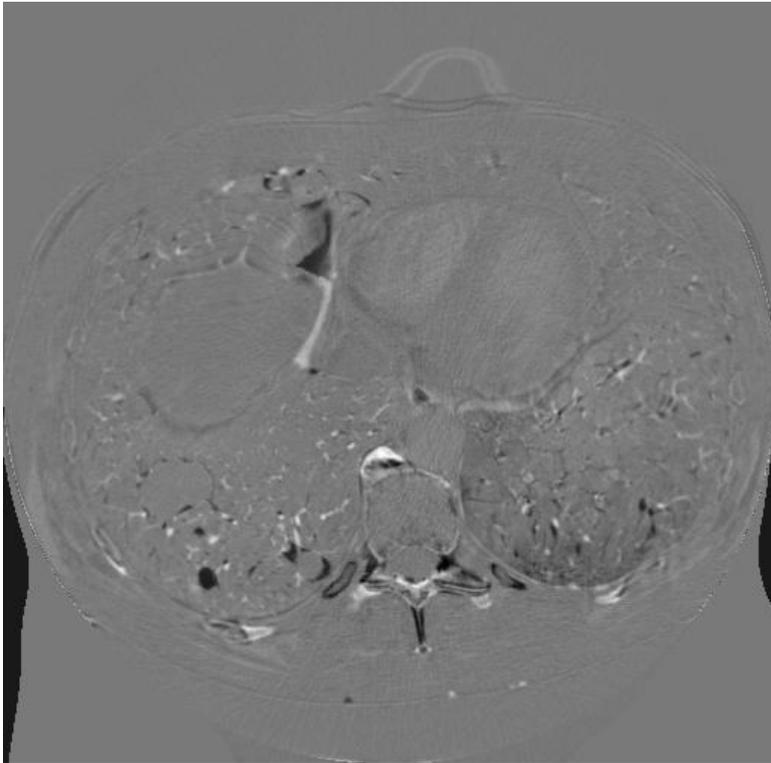
- Total constraint:

$$\mathcal{R}^{\text{rigid}}(\mathbf{u}) = \frac{1}{c(\mathbf{T}(\mathbf{x}))} \sum_{\mathbf{x}} c(\mathbf{T}(\mathbf{x})) \left\{ c_{\text{AC}} \sum_{k,i,j} \text{AC}_{kij}^2 + c_{\text{OC}} \sum_{i,j} \text{OC}_{ij}^2 + c_{\text{PC}} \text{PC}^2 \right\}$$

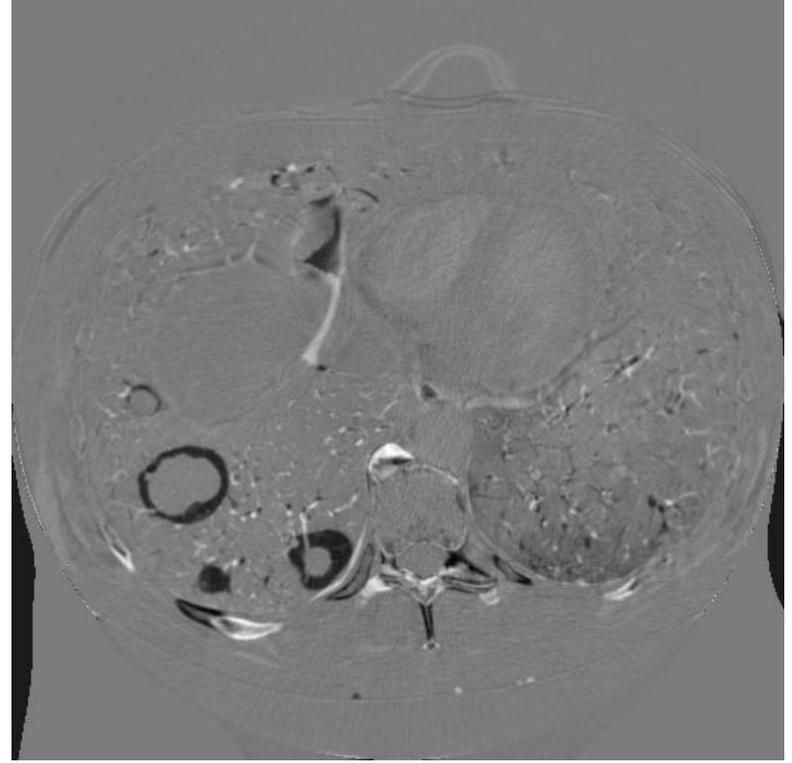
- $c(\mathbf{x})$  a rigidity coefficient dependent on tissue properties
  - $c_{\text{AC}}, c_{\text{OC}}, c_{\text{PC}}$  determine the relative strength for each term ( $\approx 100:1:10$ ).
- 
- Theorem:  $\mathcal{R}^{\text{rigid}}(\mathbf{u}) = 0$  iff  $\mathbf{T}(\mathbf{x})$  is rigid

# Rigidity penalty – thorax CT

Difference images



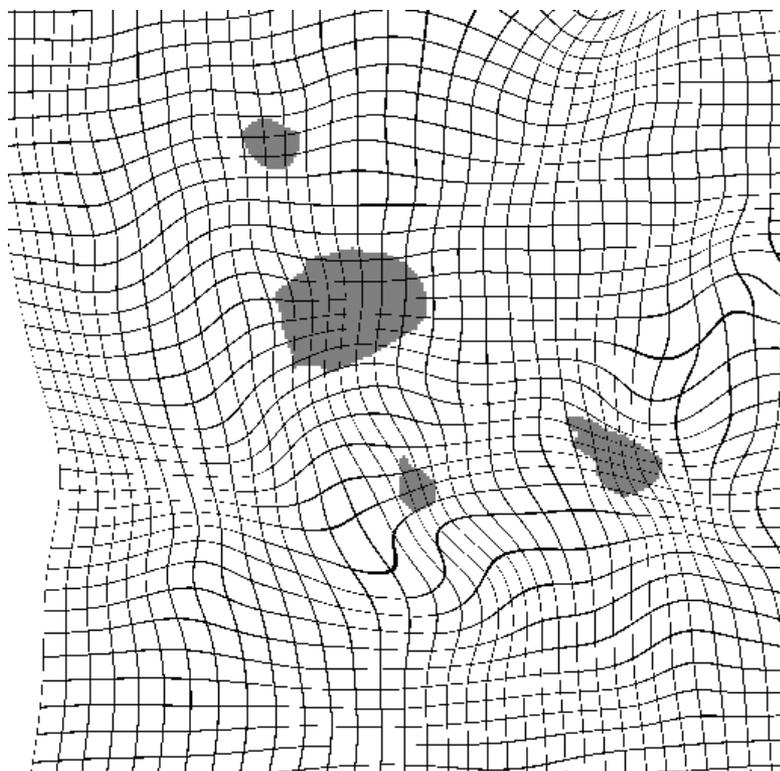
Without constraint



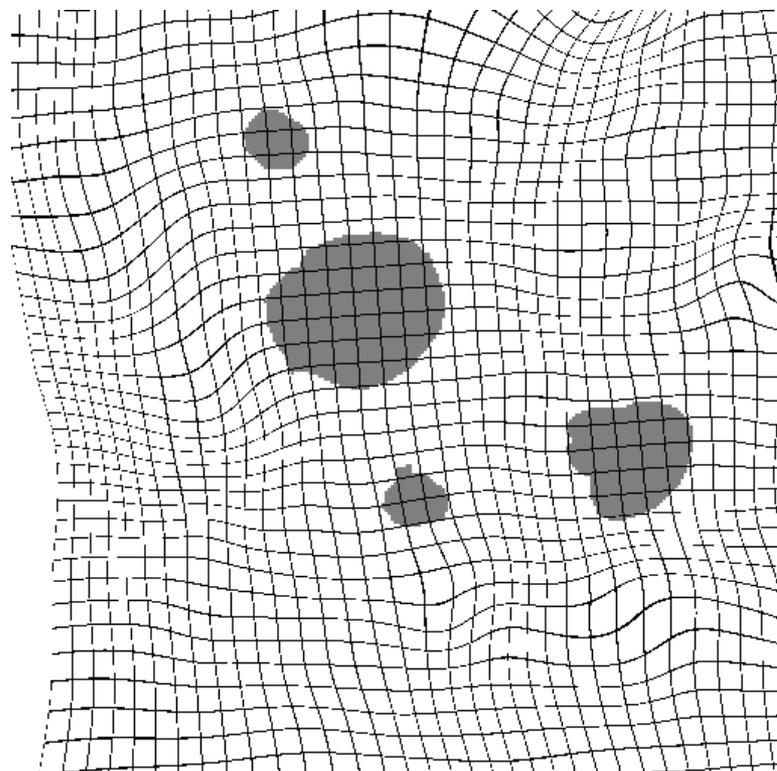
With constraint

# Rigidity penalty – thorax CT

## Deformation fields

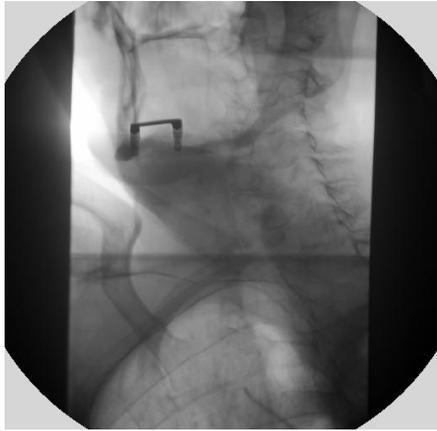


Without constraint



With constraint

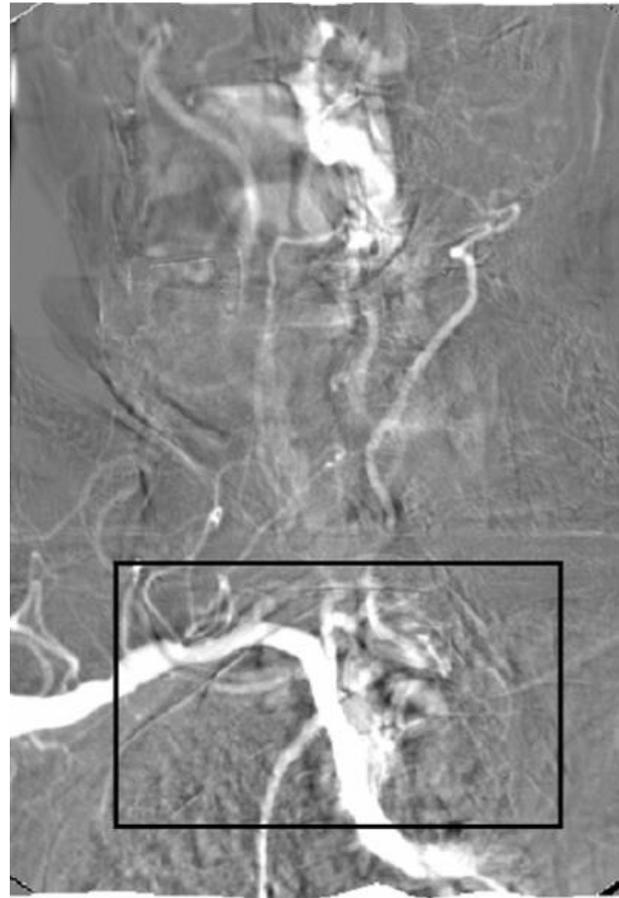
# Rigidity penalty – DSA



without contrast



with contrast



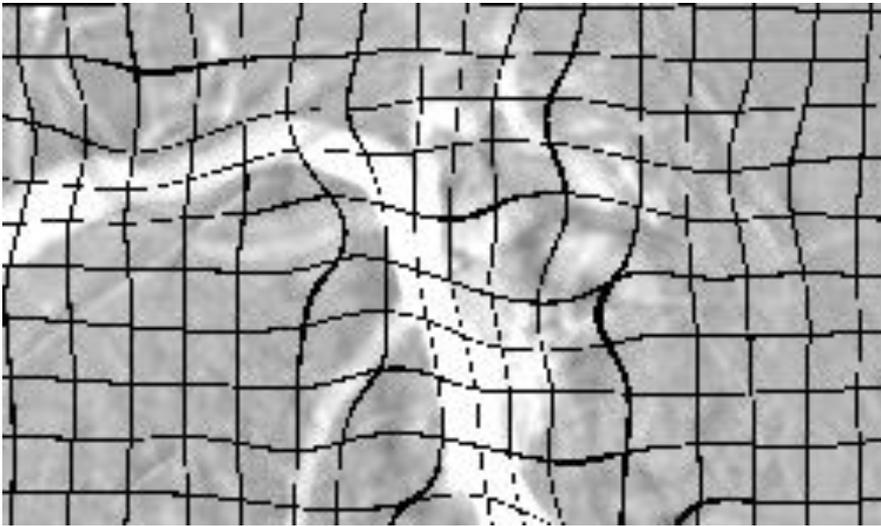
without constraint



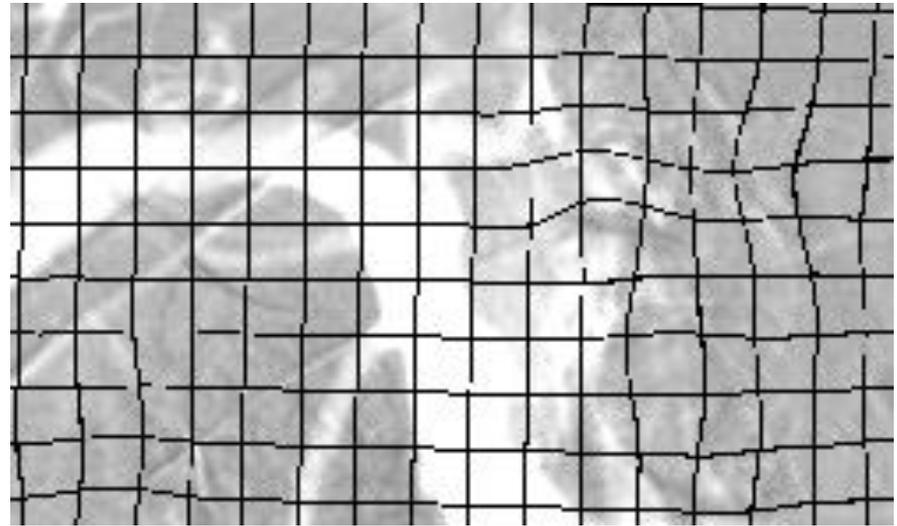
with constraint

# Rigidity penalty – DSA

Deformation fields



without constraint



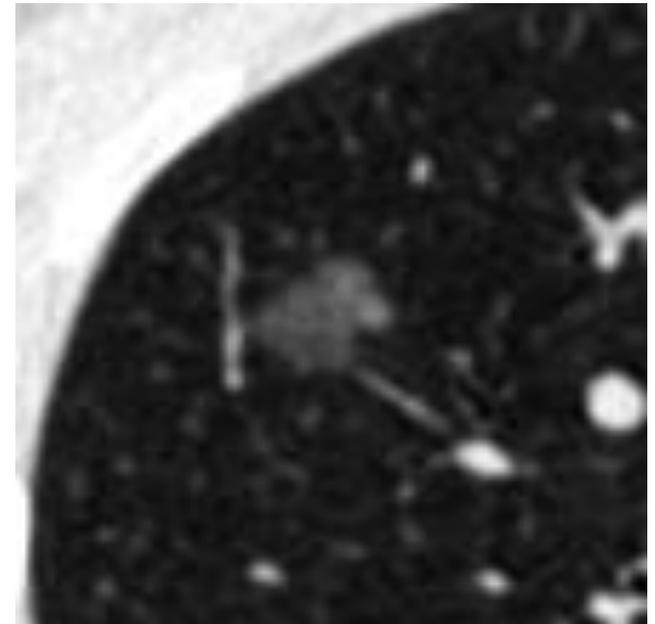
with constraint

# Rigidity penalty

- **Conclusions** (Staring 2007):
  - The rigidity penalty term is capable of achieving local rigidity, where standard registration is not.
  - Affinity is not (only regularization), properness to a large extent (due to incompressibility), orthonormality is the most dominant term. The full term is the best.

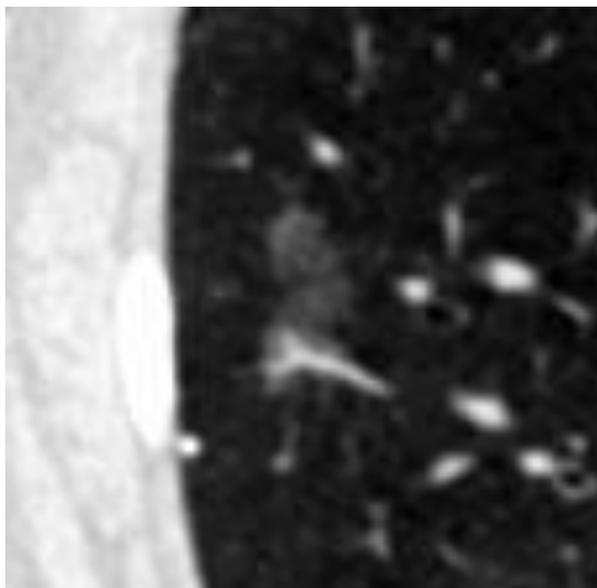
# Ground glass opacities

- In 2002 lung cancer caused 1.2 million deaths.
  - Early malignancy detection can improve prognosis.
  - Size and density change are important factors for malignancy.
- 
- Purpose: To demonstrate that image subtraction after image registration improves detection of change in pulmonary ground-glass nodules identified on chest CT scans.

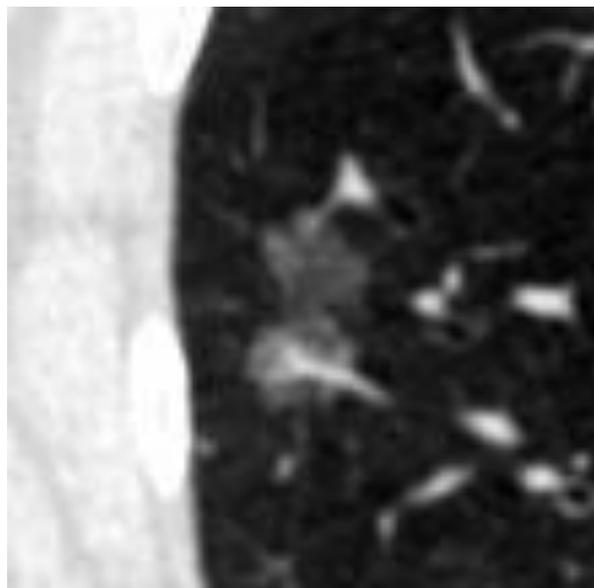


# Ground glass opacities: Idea

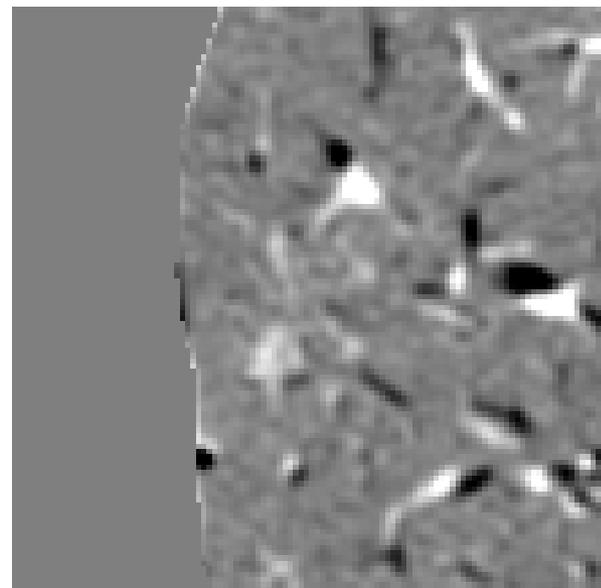
Change detection of ground glass nodules:



base line scan



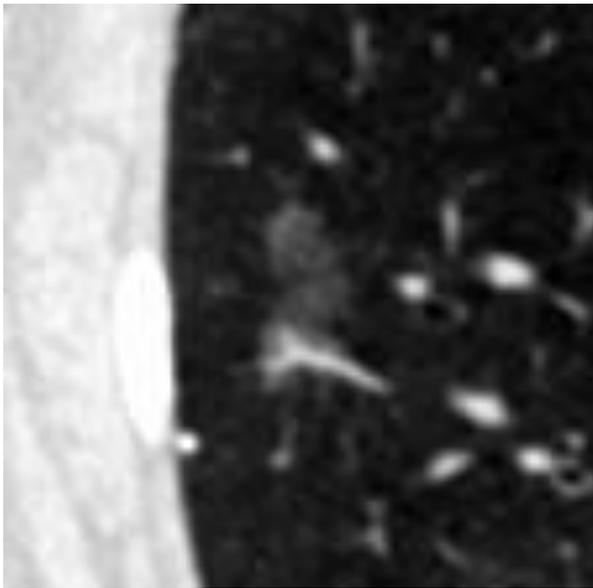
follow-up scan



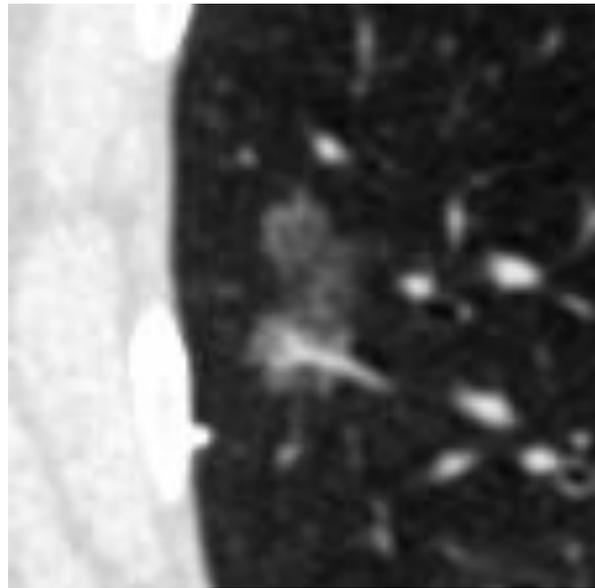
subtraction image

# Ground glass opacities: Idea

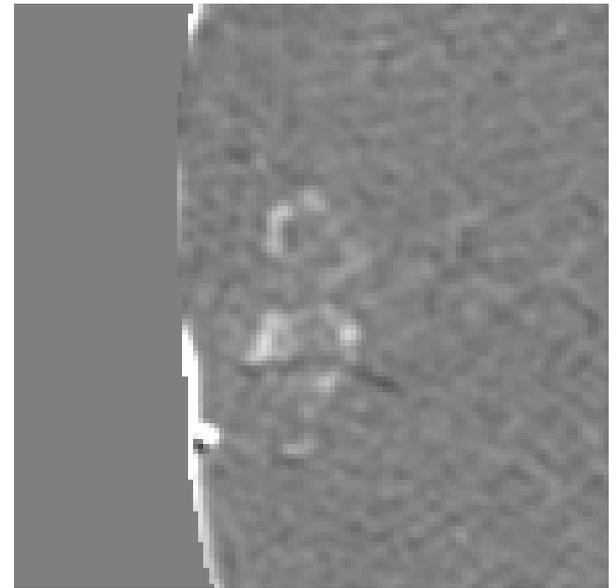
Nonrigid registration, but rigid around nodule:



base line scan



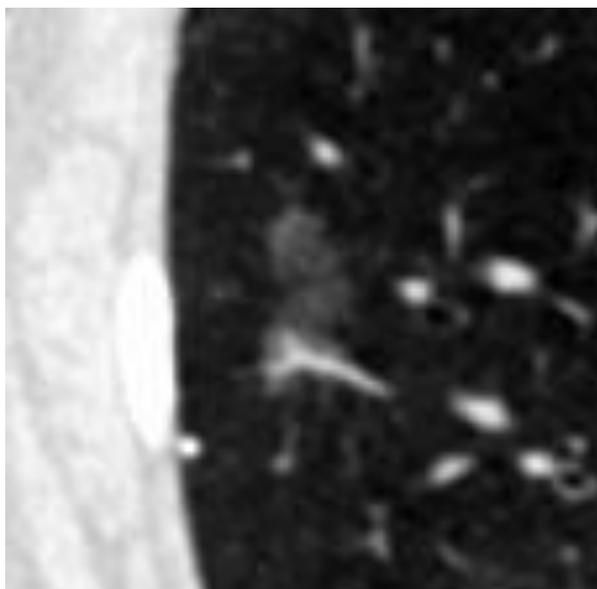
follow-up scan



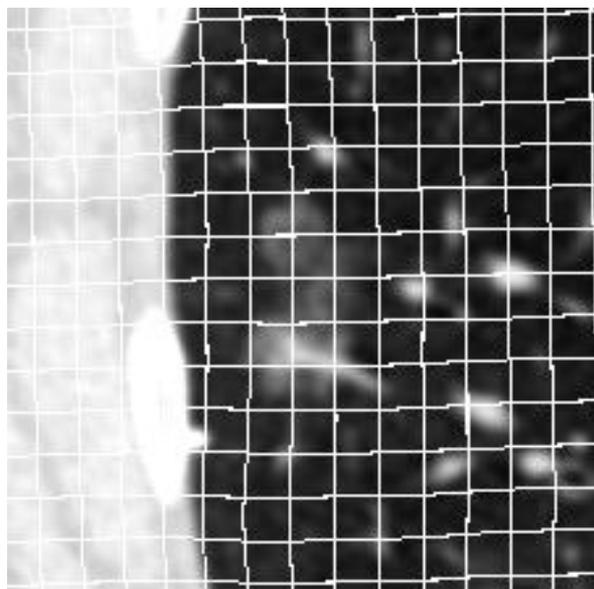
subtraction image

# Ground glass opacities: Idea

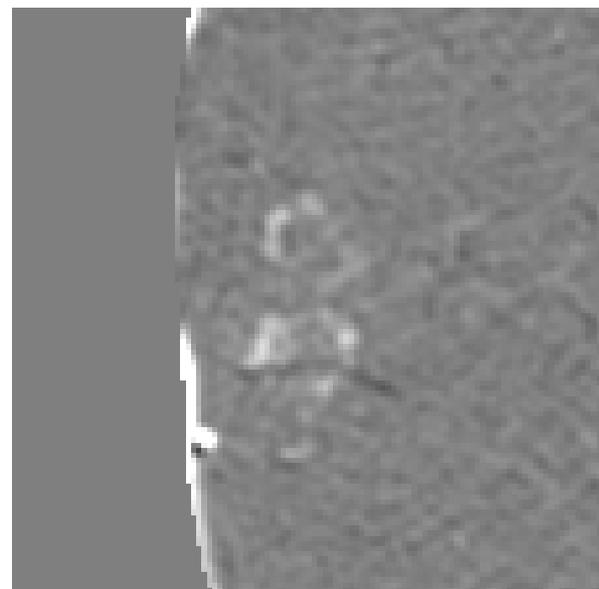
Nonrigid registration, but rigid around nodule:



base line scan



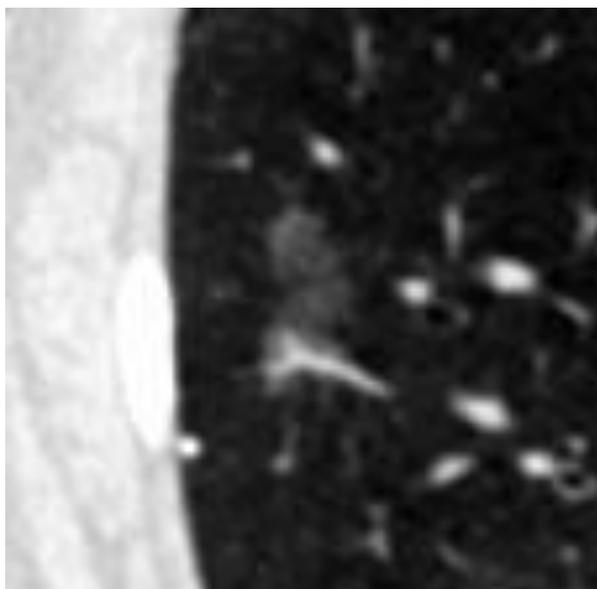
follow-up scan



subtraction image

# Ground glass opacities: Idea

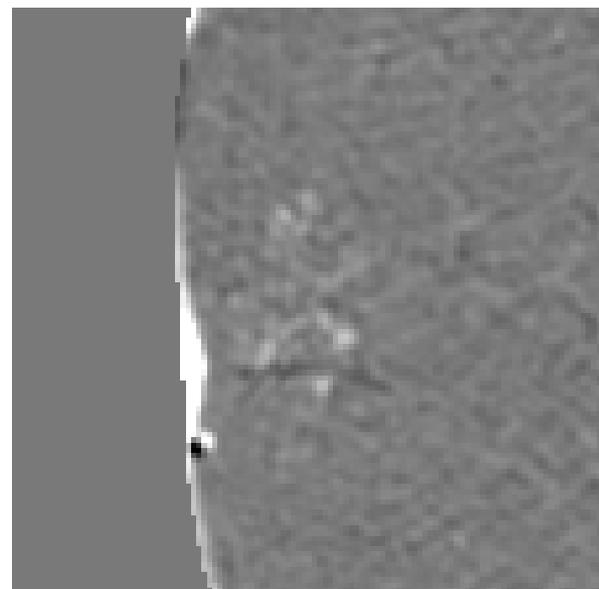
Otherwise nodule growth may be underestimated:



base line scan



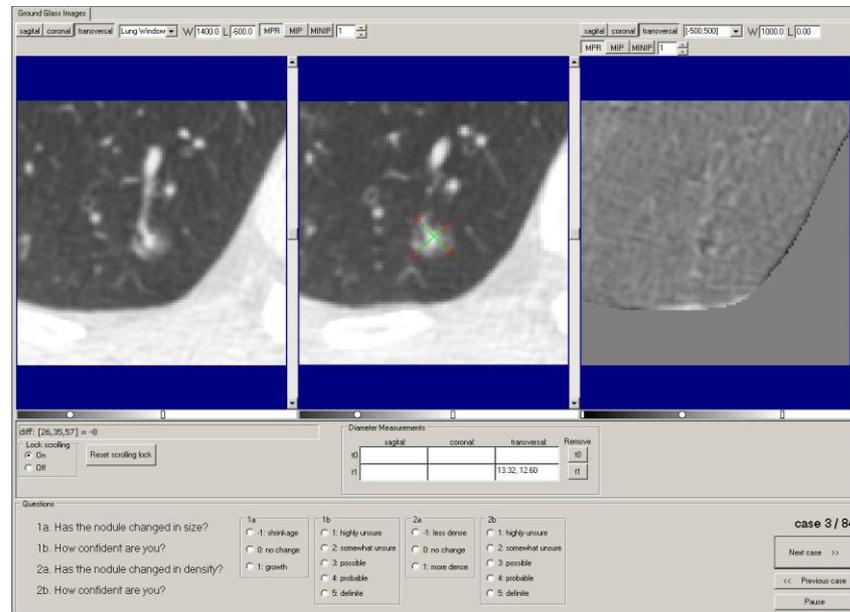
follow-up scan



subtraction image

# Ground glass opacities: Observer study

- Registration with local rigidity, to avoid concealment of nodule growth.
- Observer study: show images, with and without subtraction.
- Measure agreement for both methods with kappa statistic.

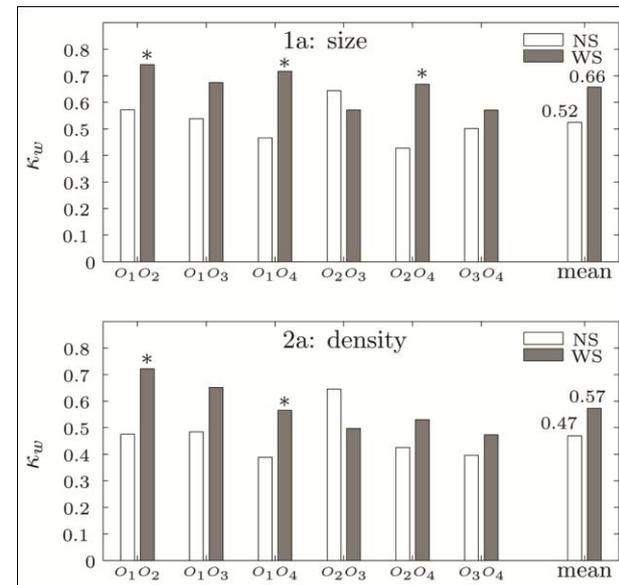


# Ground glass opacities: Results

Agreement with SoR		observer 1	observer 2	observer 3	observer 4	mean
size	standard	0.64	0.61	0.58	0.60	0.61
	subtraction	0.72	0.70	0.87 *	0.74 *	0.76
density	standard	0.57	0.52	0.52	0.51	0.53
	subtraction	0.68	0.60	0.74 *	0.54	0.64

- Category 1: highest confidence
- Category 2: lower confidence

		category 1	category 2
size	standard	0.86	0.26
	subtraction	0.91	0.57
density	standard	0.76	0.19
	subtraction	0.78	0.47

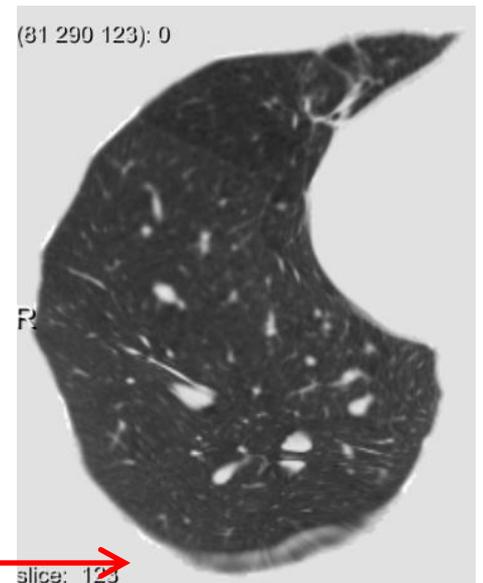


**Conclusion:** Better observer performance for low-confidence cases, better inter-observer variability. Staring 2009.

# Miscellaneous

# How to?

- How to choose  $\alpha$ ? Unsolved
- How to choose locality
  - (Manual) segmentations, assigning material properties
  - Based on HU (Ruan 2006)
  - Content-based / adaptive. Tang 2010: more in noisy parts, less near structures
- How to detect problems? Look at smearing, compute & check  $J$
- Avoid problems: use multi-resolution, multi-grid.



# What of this is in **elasti** ?

- T parameterized
  - ✓ B-splines (!)
  - ✓ Thin Plate Splines
- ✗ T non-parameterized
- ✗ Sobolev
- T filtering
  - ✓ Adaptive filtering for rigidity
  - ✗ Gaussian, anything else
- Penalty terms
  - ✓ Bending energy
  - ✓ Incompressibility
  - ✓ Rigidity

So ?

# So? Conclusions

- Regularization is needed!
- Parameterized: some implicit regularization for free  
Non-parameterized: always explicit
- Tailored  $R$ 's exist
- Can device your own, usually based on (derivatives of)  $T$
- Choosing weight not trivial (put your research efforts into it 😊)

# Questions ?