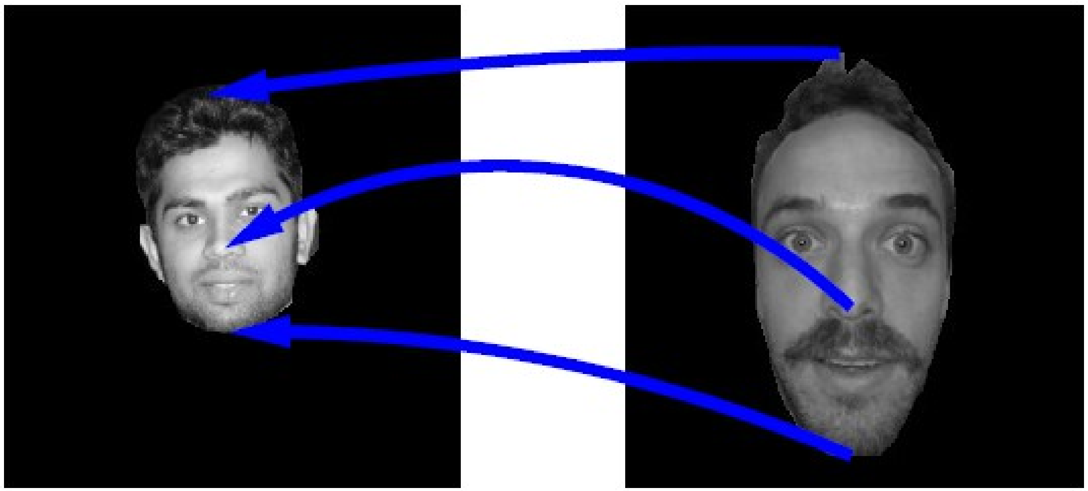
Image registration

Dana Cobzas

January 2016

Image registration

WHAT is image registration



Transform a “source” image to **match** a “target” image

Image registration

WHAT is image registration

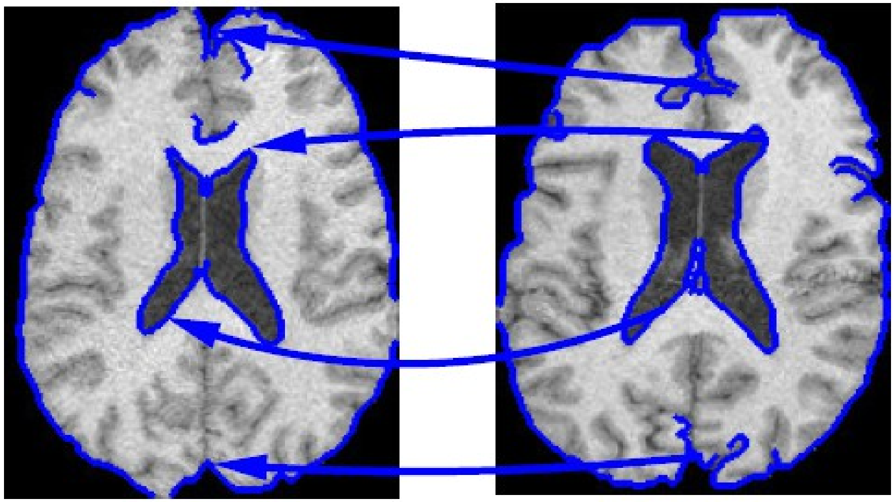
Transform a “source” image to **match** a “target” image

Medical image registration

WHAT is image registration

Transform a "source" image to **match** a "target" image

Medical image registration

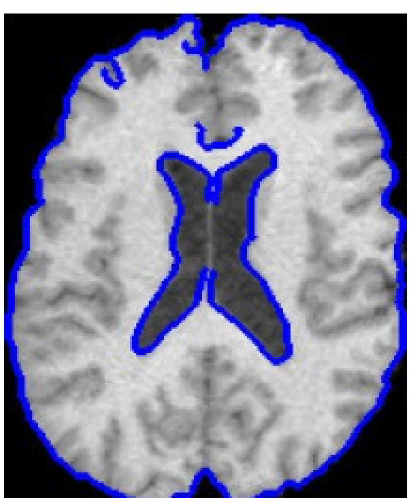


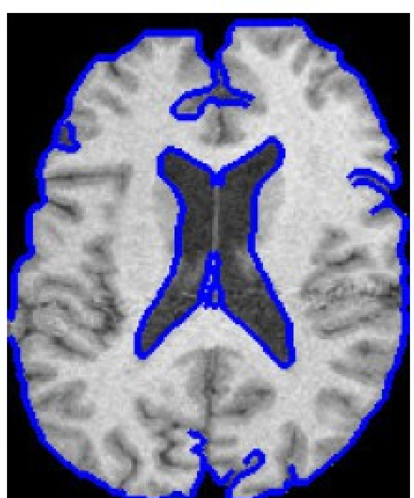
Medical image registration

WHAT is image registration

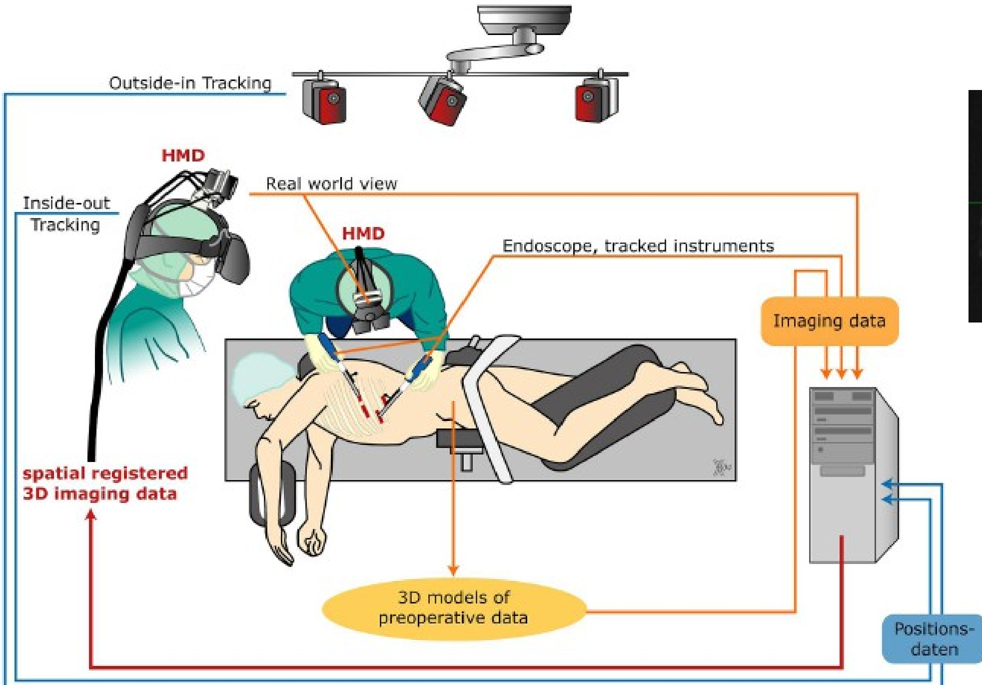
Transform a “source” image to **match** a “target” image

Medical image registration

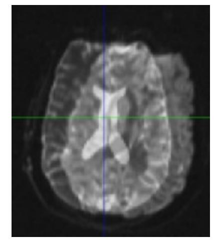




Medical applications

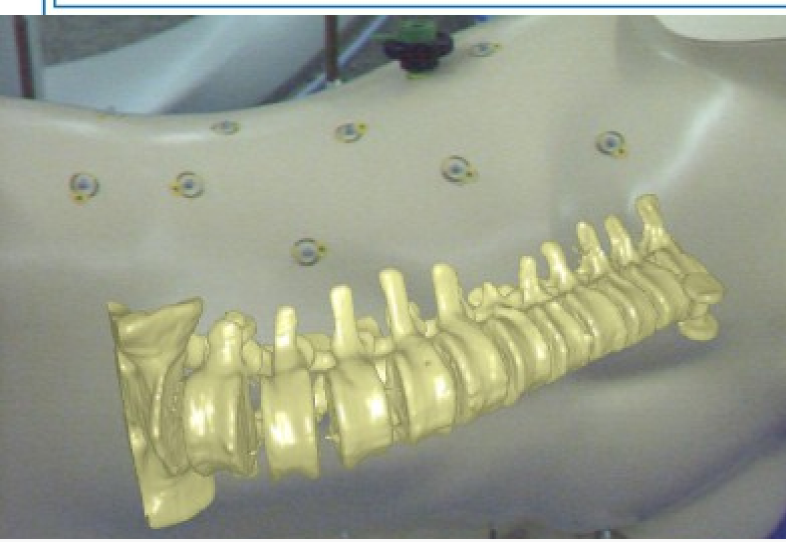


- **Data (source, target)**

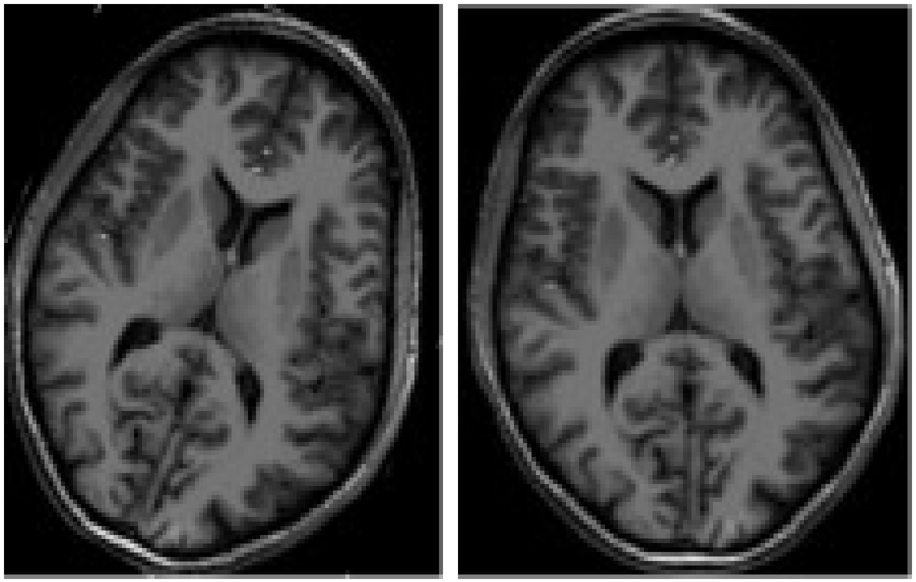


  - different medical images modalities (MRI, XRay, CT…)
  - pre-acquired medical images with real time images (video)
  - patient data with an atlas

- **For:**
  - atlas generation
  - augmented reality (surgery)
  - better diagnosis
  - data analysis



reference source



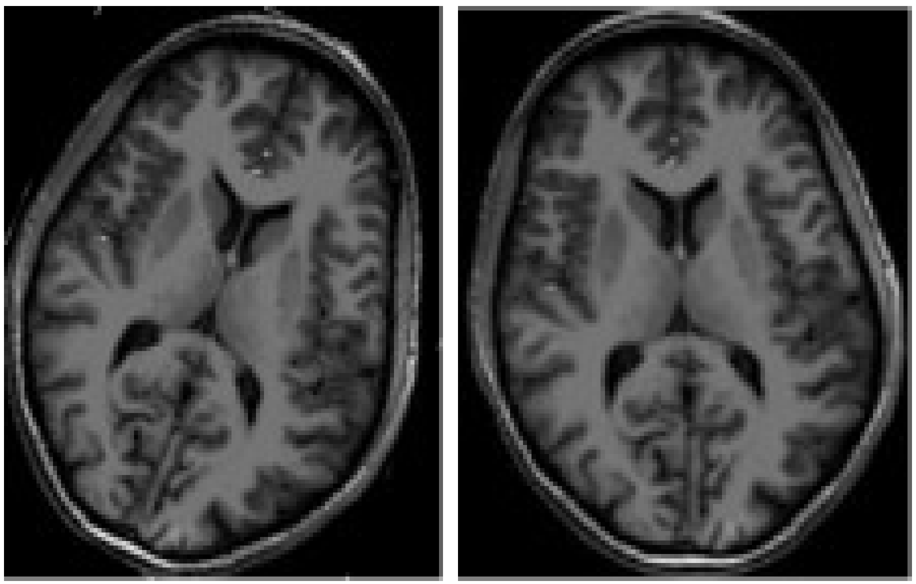
reference

registered
source

Affine registration
6DOF – 2D
12 DOF - 3D
Low number DOF

Similar to tracking

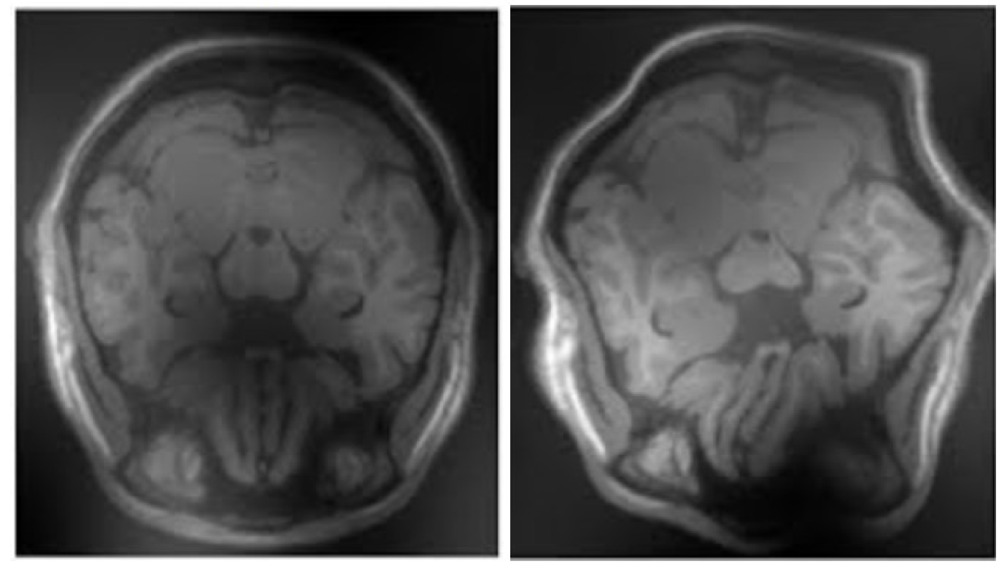
reference source



Affine registration
6DOF – 2D
12 DOF - 3D
Low number DOF

Similar to tracking

Reference image Source image



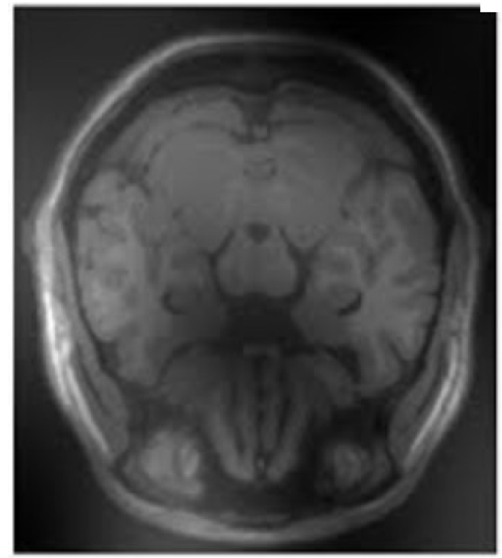
reference

source

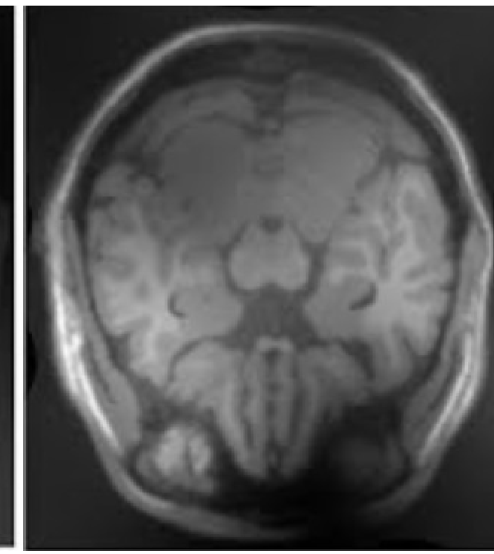
Affine registration
6DOF – 2D
12 DOF - 3D
Low number DOF

Similar to tracking

Reference image



Registered source

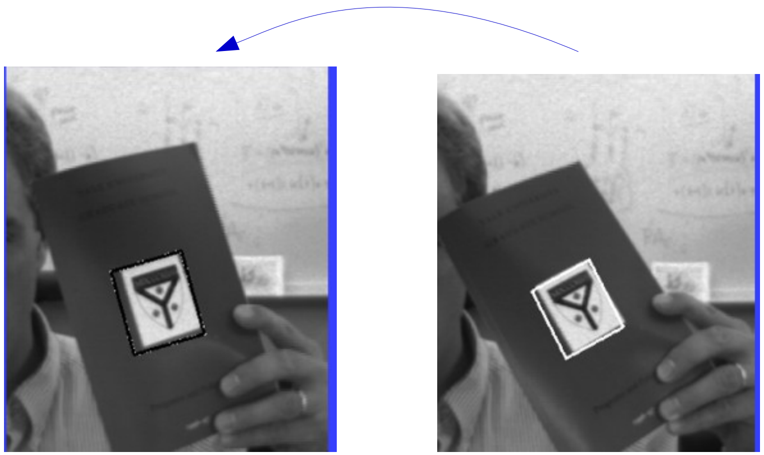


Deformable registration
Many DOF
Points move independently

Similar to optic flow

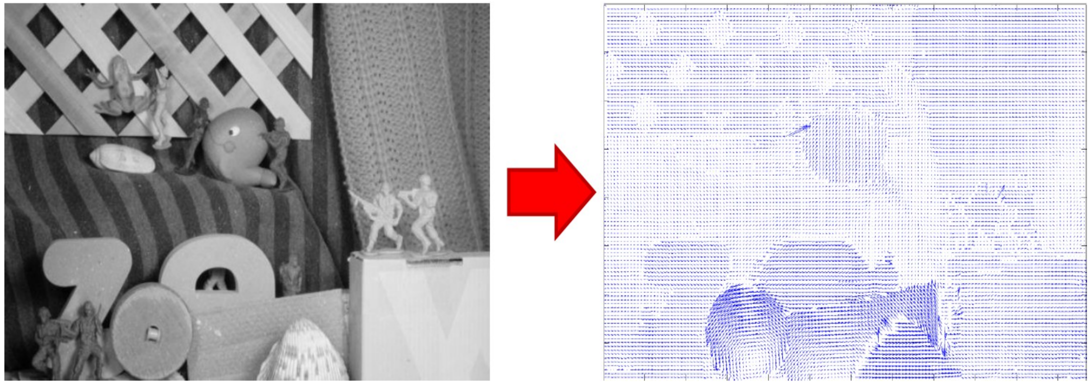
Formulation

Tracking



$$\min_p (I_t(T(p,x) - I_0(x))^2$$

Optic flow



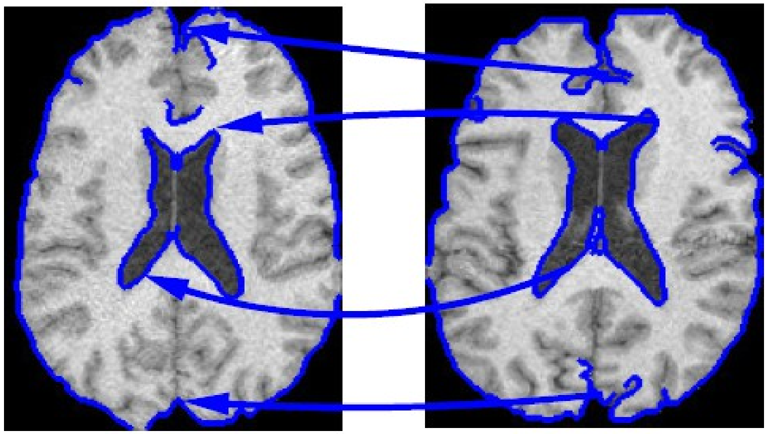
$$\min_p (I_t(x+p) - I_0(x))^2$$

$$\min_p (I_t(T(p,x)) - I_0(x))^2 \quad , \quad T(p,x) = x+p$$

Formulation

Registration

target source

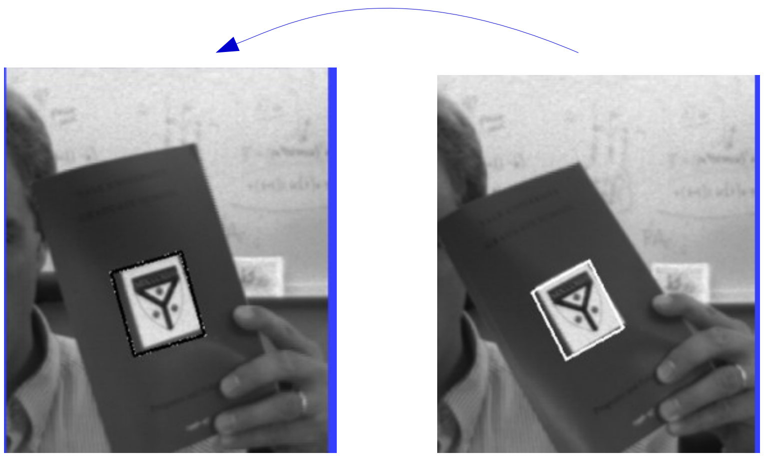


I_T I_S

$\min_p (I_S(T(p,x) - I_T(x))^2$

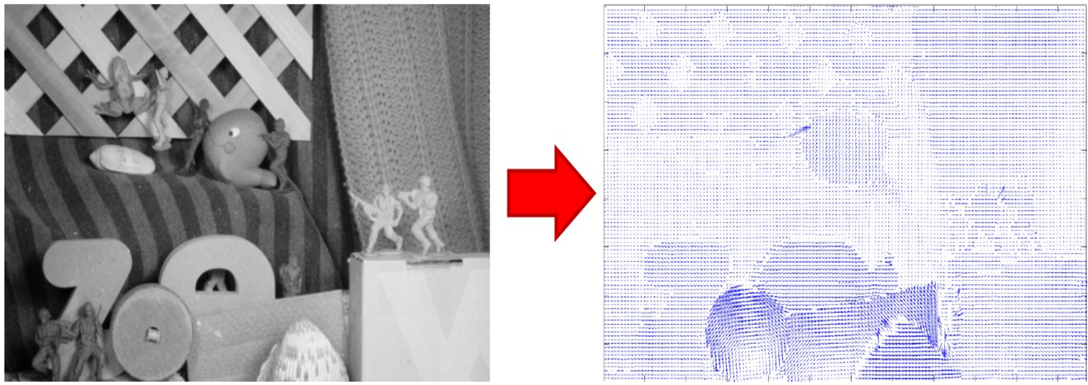
T = image transformation model

Tracking



$\min_p (I_t(T(p,x) - I_0(x))^2$

Optic flow



$\min_p (I_1(x+p) - I_0(x))^2$

$\min_p (I_1(T(p,x)) - I_0(x))^2 \quad , \quad T(p,x) = x+p$

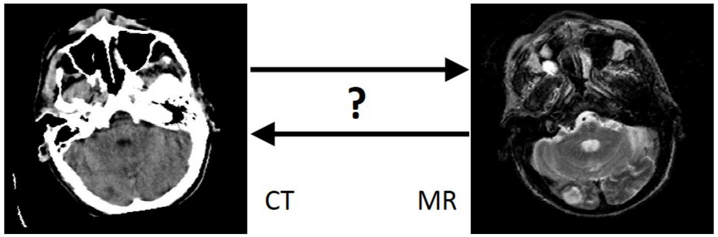
BUT ...

$\min_p \, (I_S(T(p,x) - I_T(x))^2$

BUT ...

$$\min_p (I_S(T(p,x) - I_T(x))^2$$

Source and target image can be very different



$\min_p$ sim$(I_S(T(p,x) – I_T(x))$

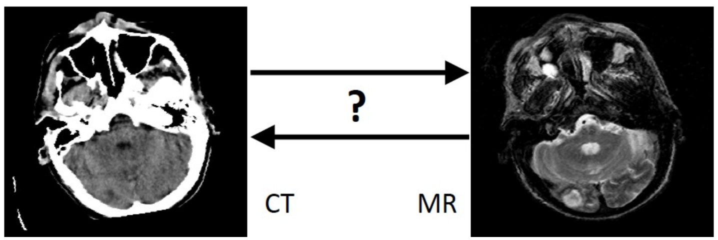
Similarity score : SSD, NCC, MI ...

Similarity score

BUT ...

$\min_p (I_S(T(p,x) - I_T(x))^2$

Source and target image can be very different



$\min_p \mathrm{sim}(I_S(T(p,x) – I_T(x))$

Similarity score : SSD, NCC, MI ...

Similarity score

Motion is large

No continuous flow of images

Transformation models

Linear registration – few DOF

Deformable registration – many DOF

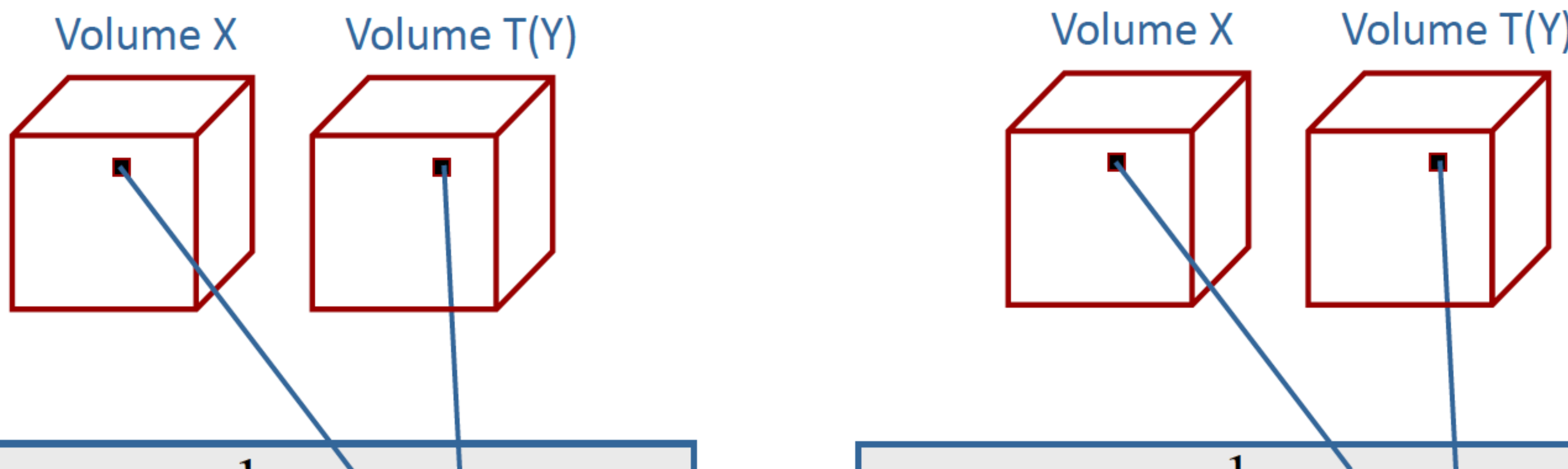
Optimization

Difficult optimization

Similarity scores

- SSD SAD
- Cross correlation
- Mutual information

SSD-SAD



$$SSD = \frac{1}{N}\sum_i (x_i - y_i)^2$$

Sum of Squared Differences

$$SAD = \frac{1}{N}\sum_i |x_i - y_i|$$

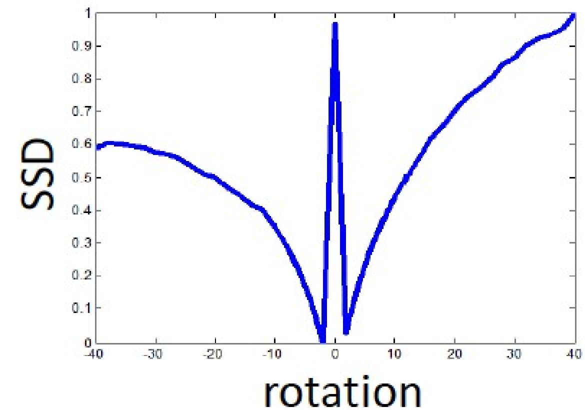
Sum of Absolute Differences:
Less sensitive on large intensity differences than SSD

[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Limitations of SSD

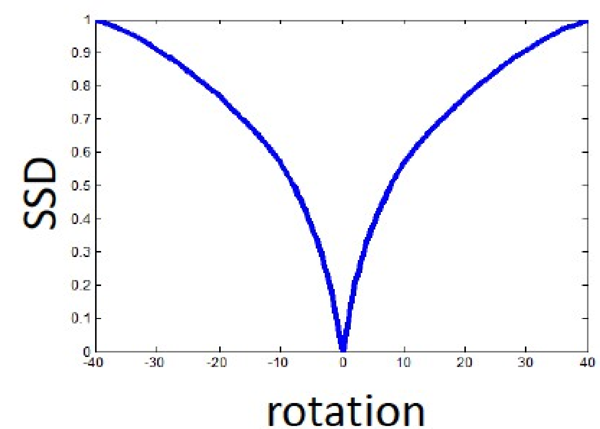
- Illumination change affects similarity function





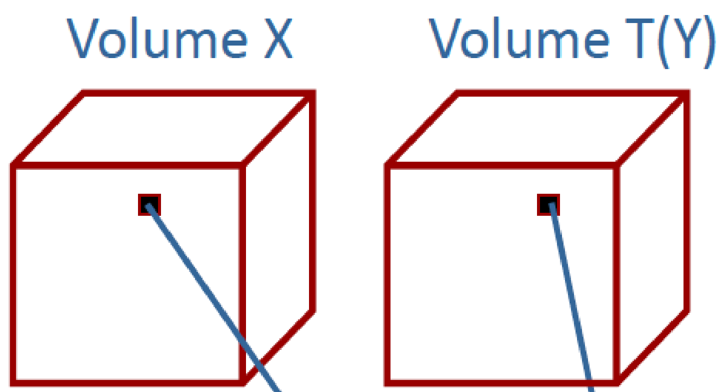
- Idea: normalization of images $\tilde{X} = \frac{X - \mathbb{E}[X]}{\sigma(X)}$





[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Normalized cross-correlation NCC



$\bar{x}$: Mean

σ_x : Standard deviation

N : Number of pixels

$$NCC = \frac{1}{N}\sum_i \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}$$

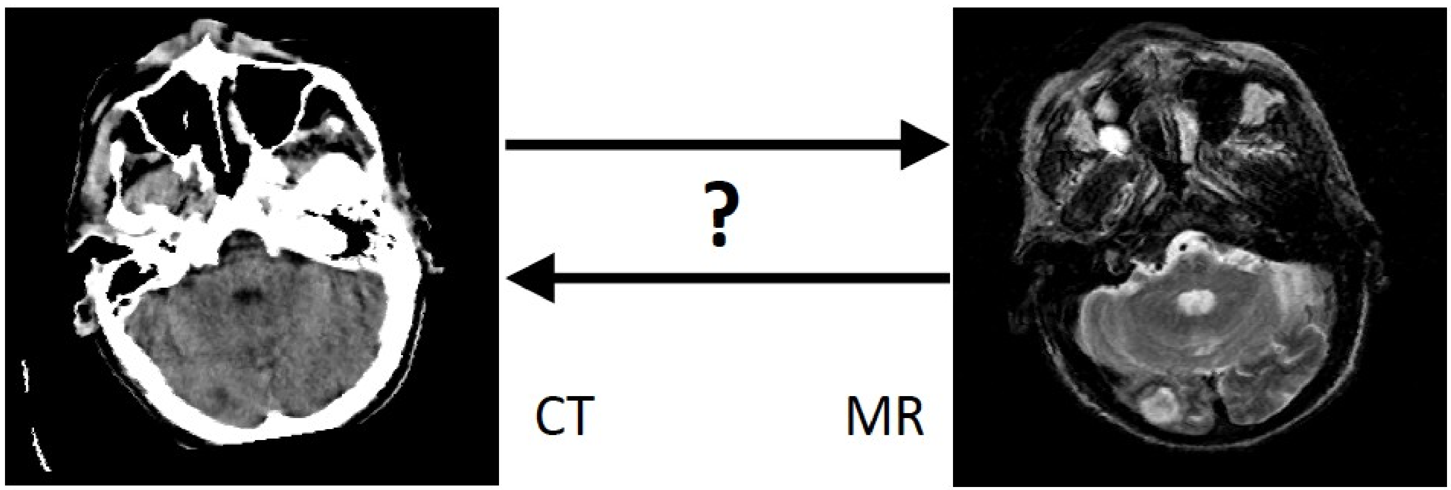
Normalized Cross Correlation:

Expresses the linear relationship between voxel intensities in the two volumes

[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Multi-modality registration

- More complex intensity relationship



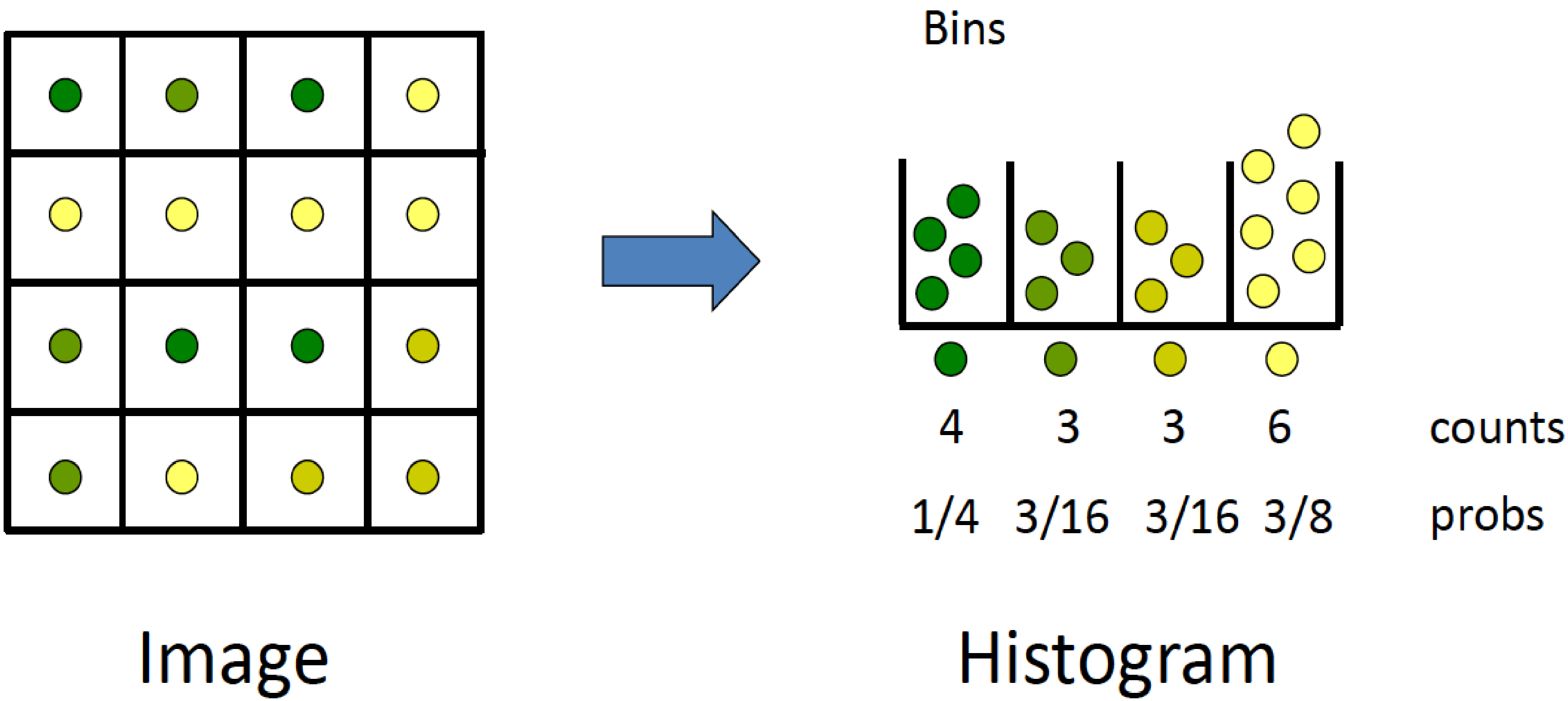
- Approaches:
  - Simulate one modality from the other one
  - Apply sophisticated similarity measures

[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Information theoretical approach

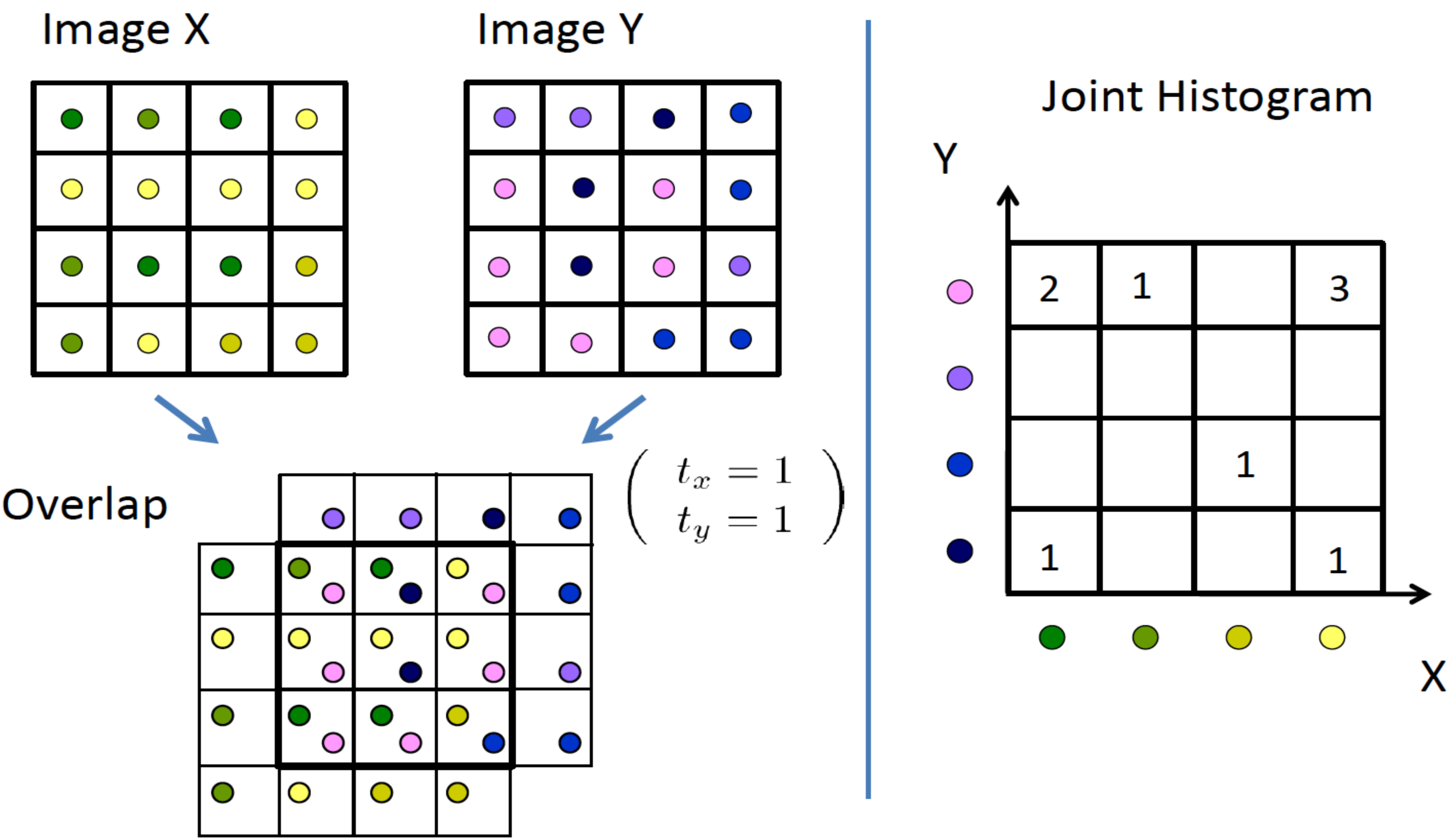
Information theoretical approach

Image histogram



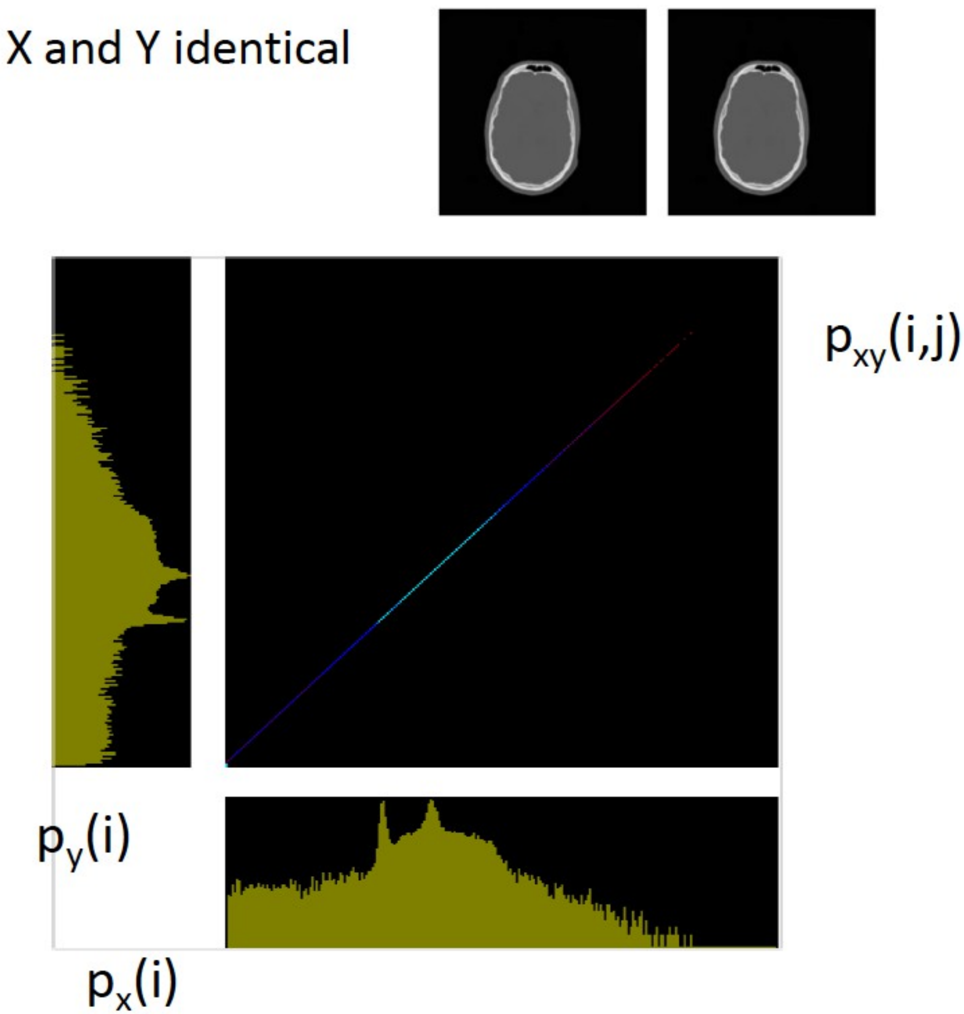
[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

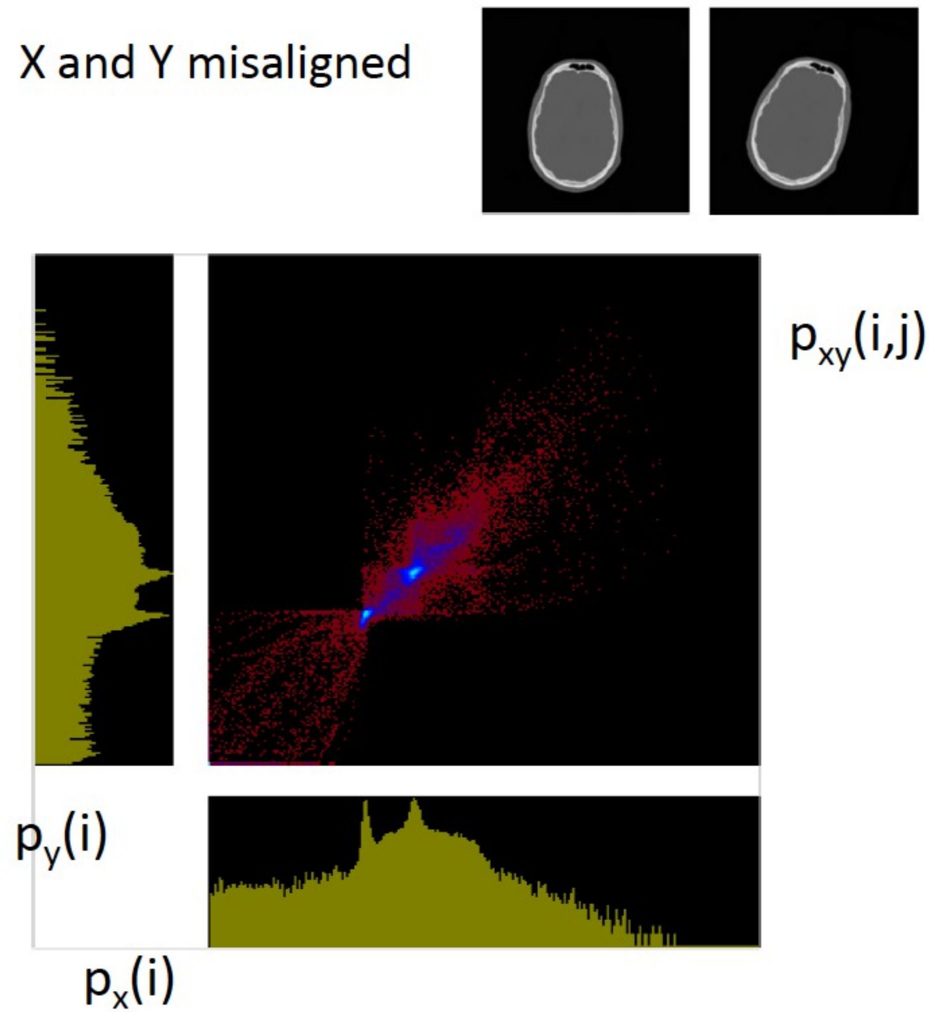
Joint histogram



[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Joint histogram

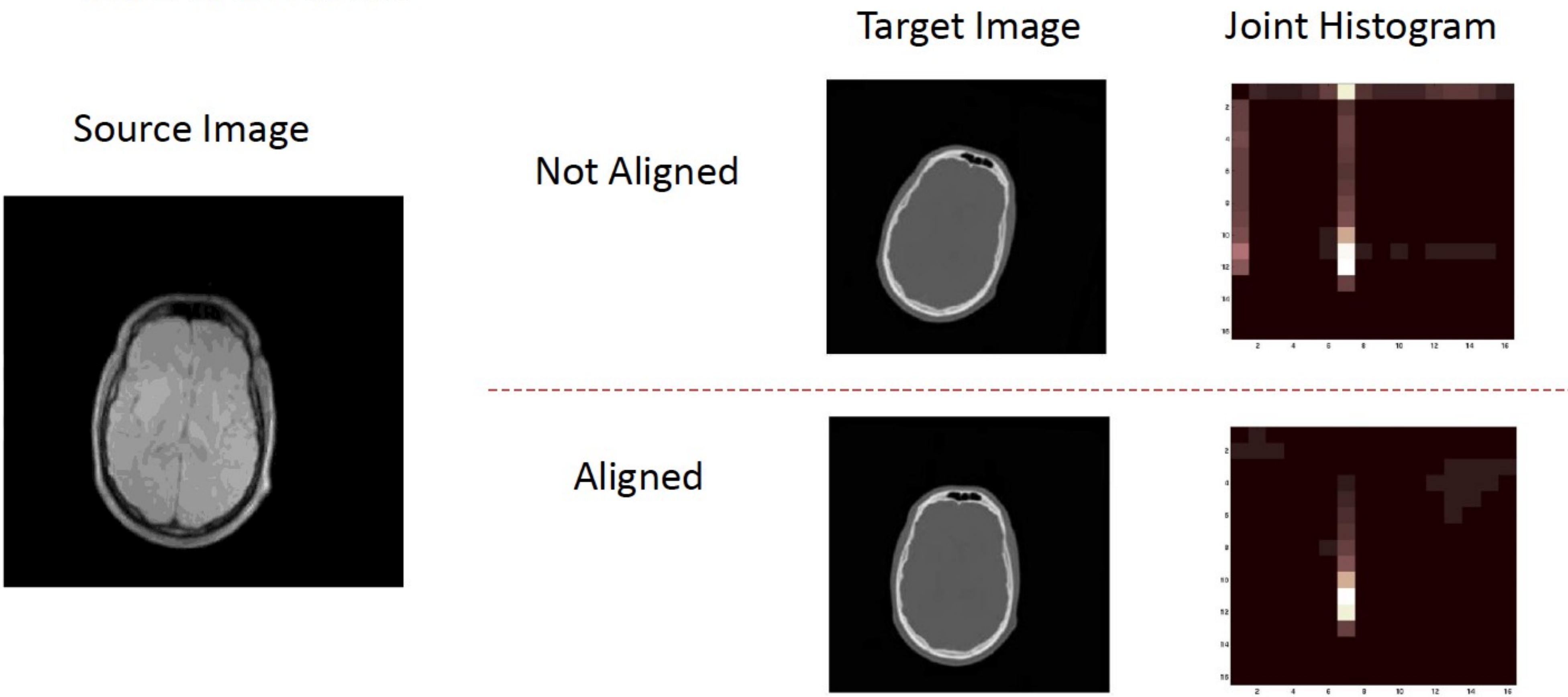




[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Joint histogram

- Histogram for images from different Modalities



Source: PhD Thesis, L. Zöllei

[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Information theoretical approach

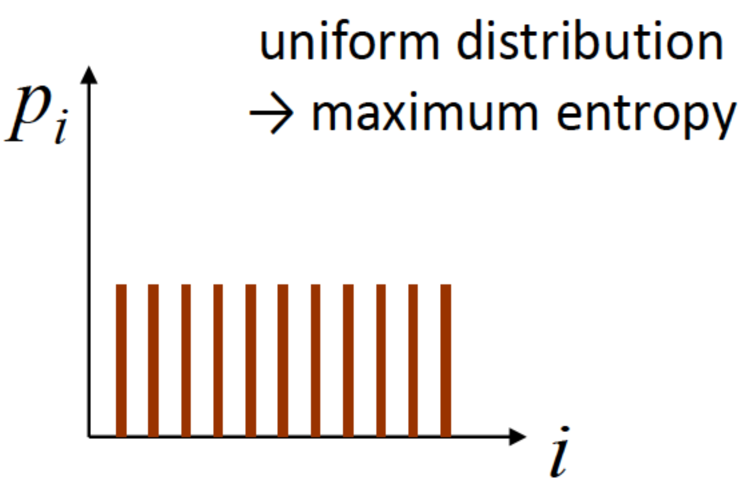
How to qualify quality of alignment between two images ?
> measure the *structure* of the joint distribution

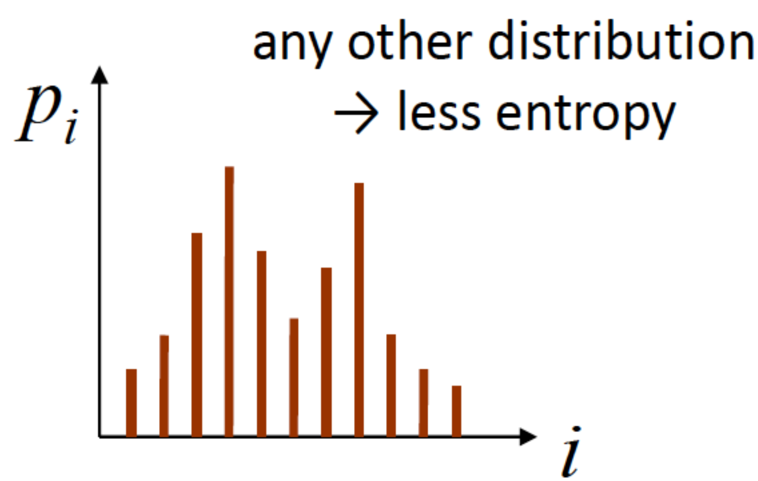
How to measure the structure ?
> Shannon Entropy

Entropy

Shannon Entropy, developed in the 1940s
(communication theory)

$$H = -\sum_i p_i \log p_i$$





[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Mutual information

$$MI(X,Y) = H(X) + H(Y) - H(X,Y)$$
$$= \sum_i \sum_j p_{xy}(i,j) \log \frac{p_{xy}(i,j)}{p_x(i)\, p_y(j)}$$

- Maximized if X and Y are perfectly aligned
- H(X) and H(Y) help to make the measure more robust
- Maximization of mutual information leads to minimization of joint entropy

[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Derivatives of similarity measures

- General form of derivative of similarity metrics

$$\frac{\partial\, \mathrm{SM}(X, Y(T_p))}{\partial p} = \boxed{\frac{\partial\, \mathrm{SM}(X, Y)}{\partial Y}} \quad \frac{\partial Y}{\partial T_p} \quad \frac{\partial T_p}{\partial p}$$

  - SSD:

$$\nabla \mathrm{SSD}(X, Y) = -2 \cdot (X - Y)$$

  - MI

$$\nabla \mathrm{MI}(X, Y) = G_\Psi * \frac{1}{|\Omega|} \left(\frac{\partial_2 p(X, Y)}{p(X, Y)} - \frac{p'(Y)}{p(Y)} \right)$$

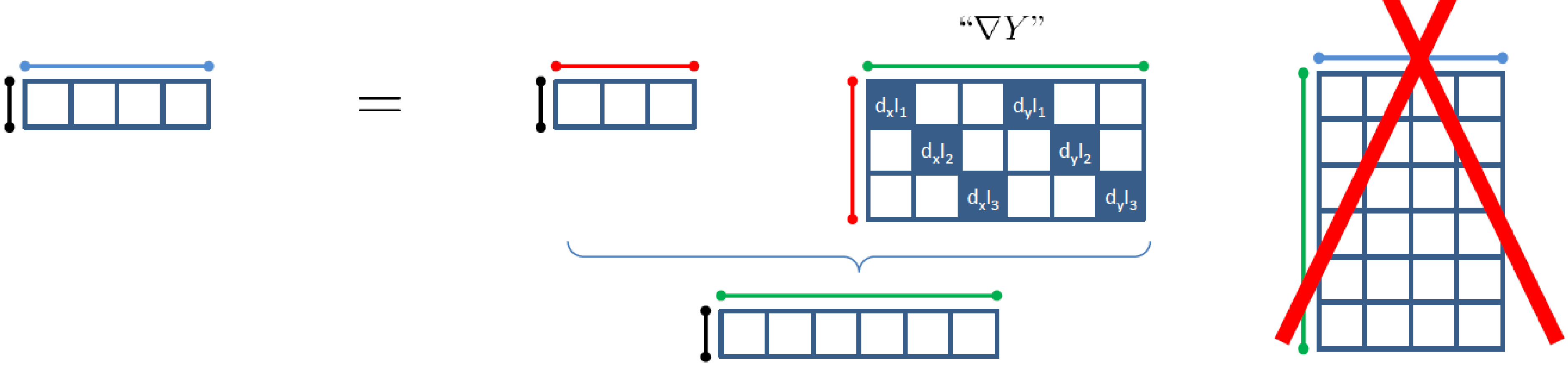
Hermosillo, G., Chefd'Hotel, C., Faugeras, O.: *Variational Methods for Multimodal Image Matching*, International Journal of Computer Vision 50(3) (2002)

[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Derivatives of similarity measures

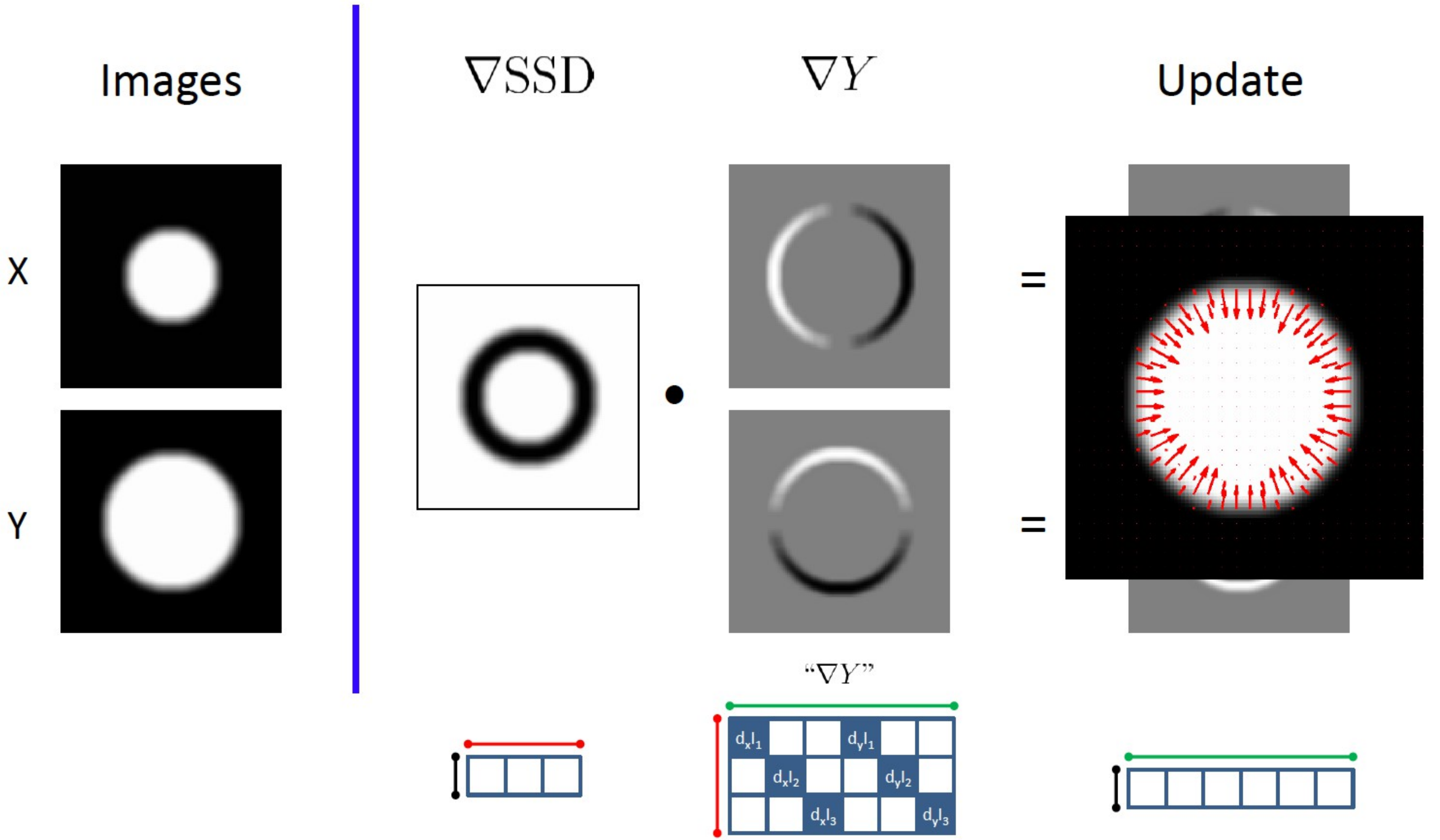
- General form of derivative of similarity metrics

$$\frac{\partial\,\mathrm{SM}(X, Y(T_p))}{\partial p} = \frac{\partial\,\mathrm{SM}(X, Y)}{\partial Y} \quad \frac{\partial Y}{\partial T_p} \quad \cancel{\frac{\partial T_p}{\partial p}}$$



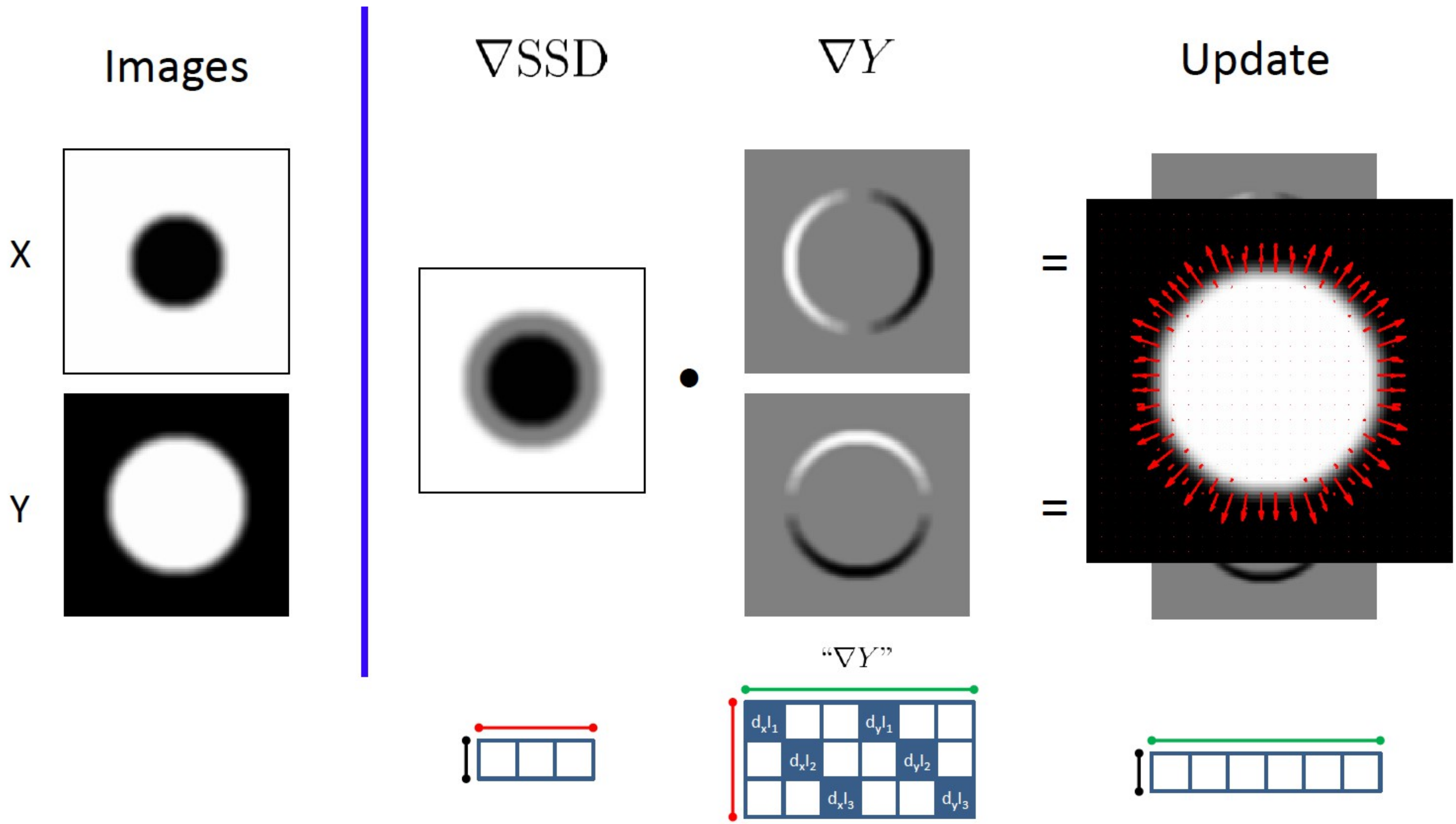
[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Derivatives of similarity measures



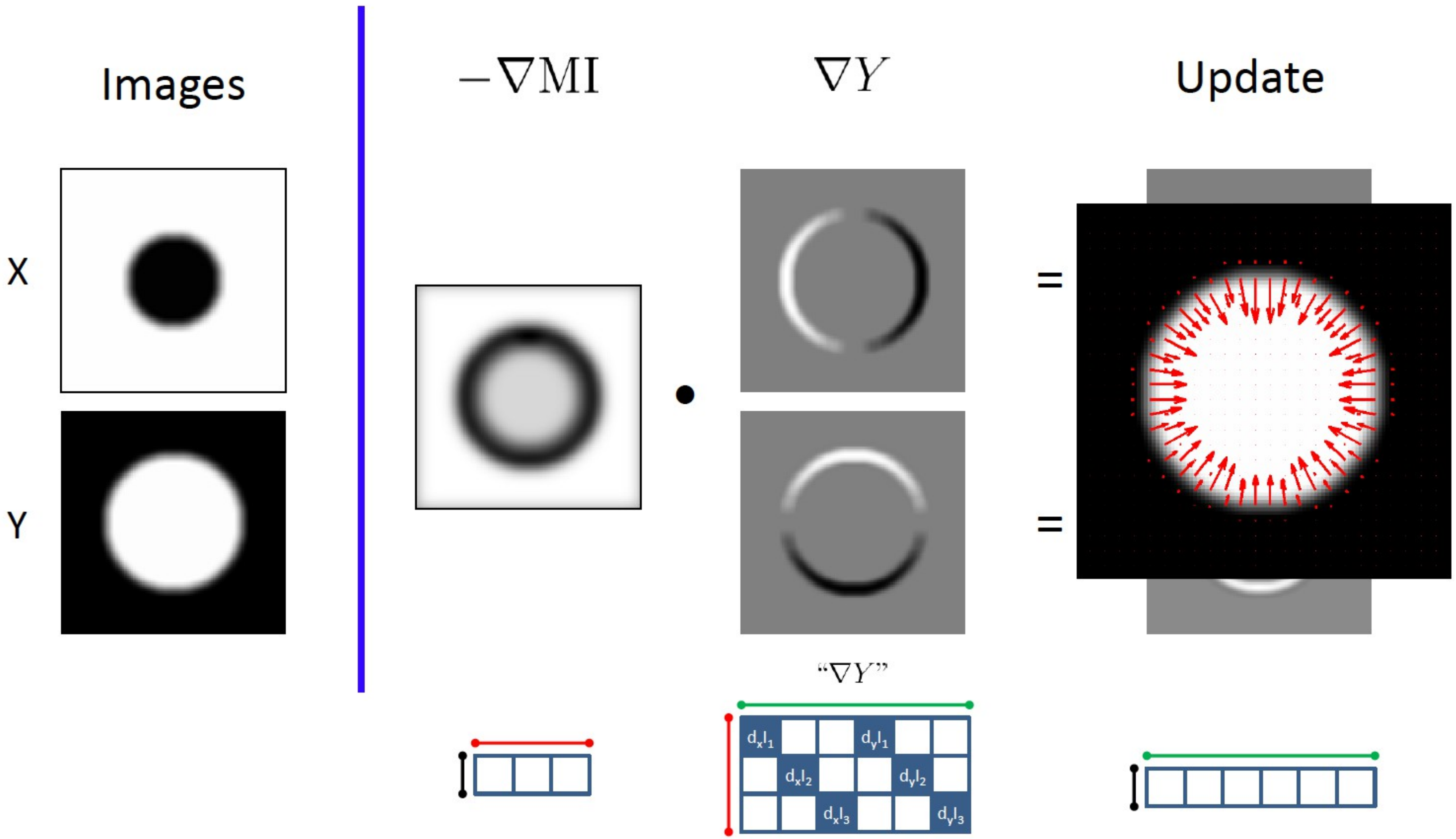
[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Derivatives of similarity measures



[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Derivatives of similarity measures

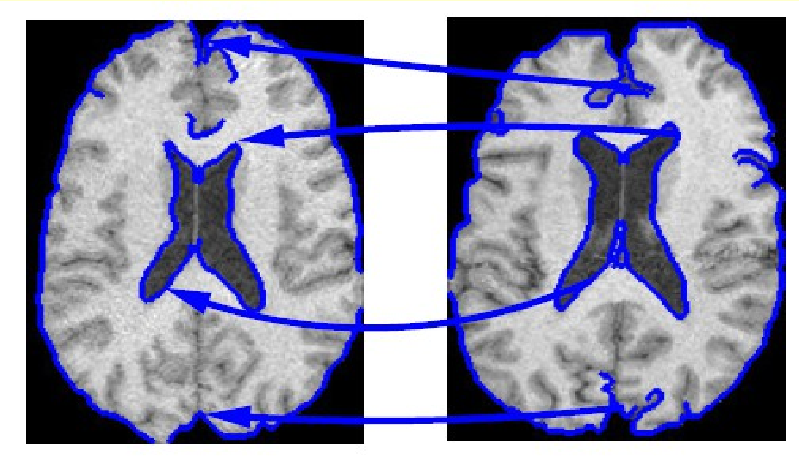


[Slides from Christian Wachinger MICCAI 2010 registration tutorial]

Transformation models and optimization

- Linear : rigid affine
- Interpolation models : Bsplines
- Continuous models and regularization

target source

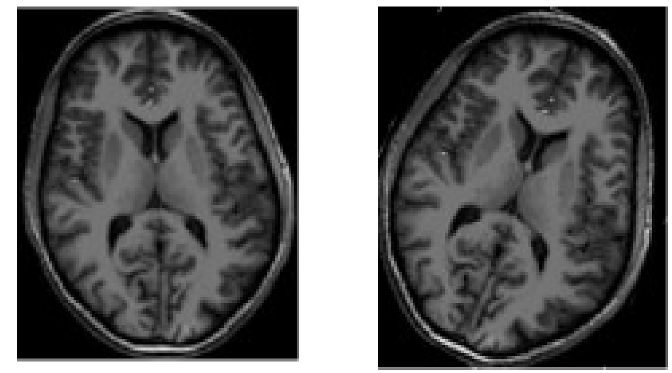


I_T I_S

$$\min_p (I_S(T(p,x) - I_T(x))^2$$

1. Linear registration :

> few (12) DOF
> nonlinear least square optimization
> very similar to tracking



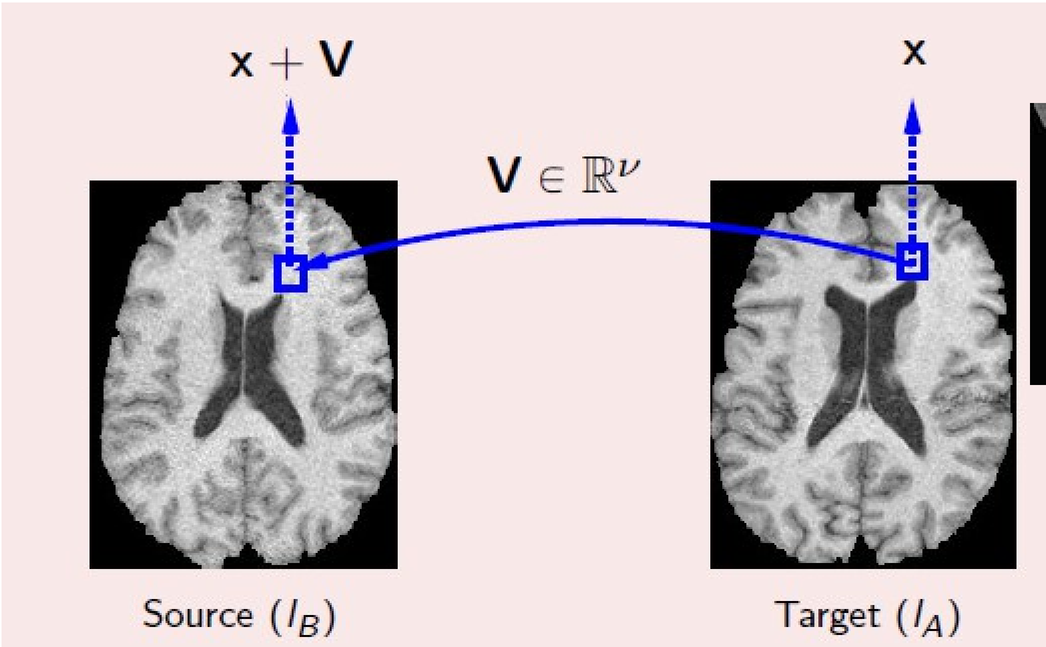
$p_{t+1} = p_t + \Delta p$

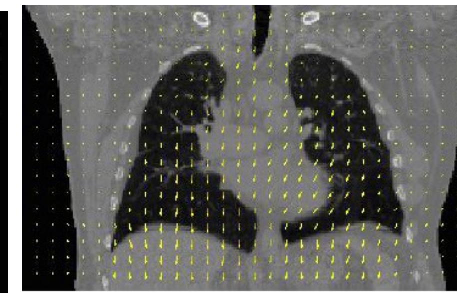
Gauss Newton : $\Delta p = -(J'J)^{-1}J'e$

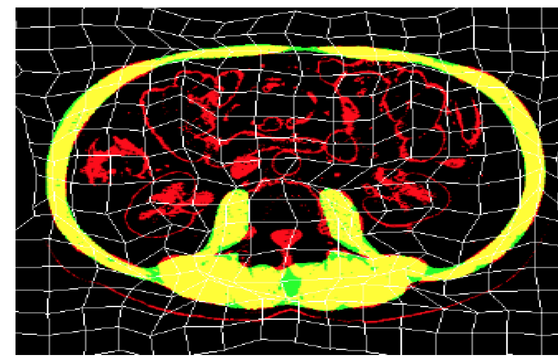
J = Jacobian of the warp $J = \nabla I_S (\partial T/\partial p)$

- "t" artificial time
- "e" defined over the whole image
- "far" from solution > multiresolution

2. Non-linear (deformable) registration





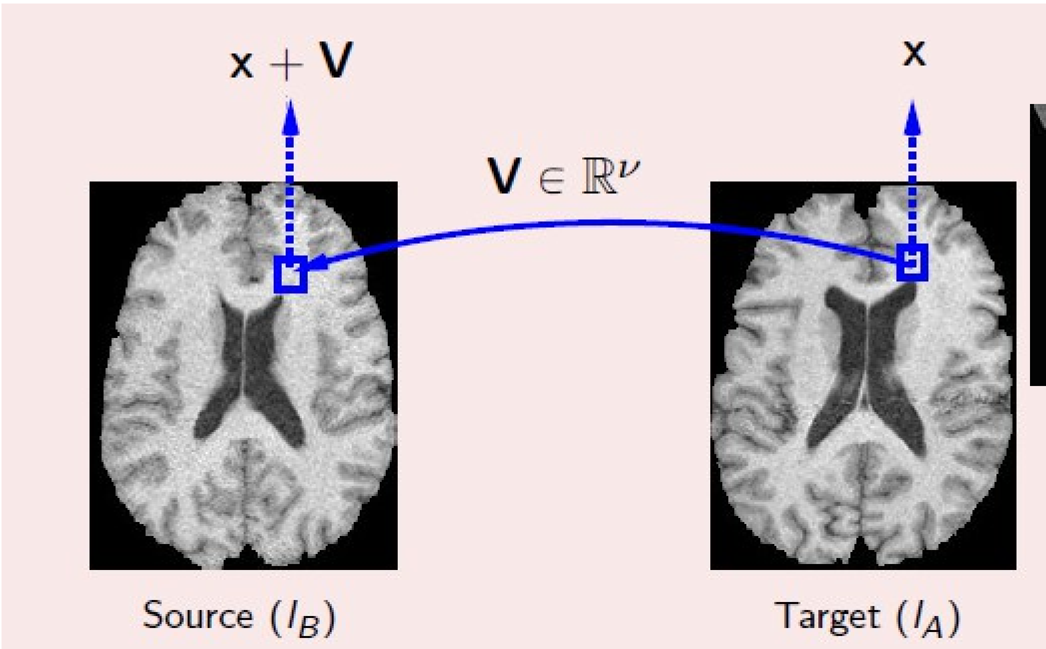


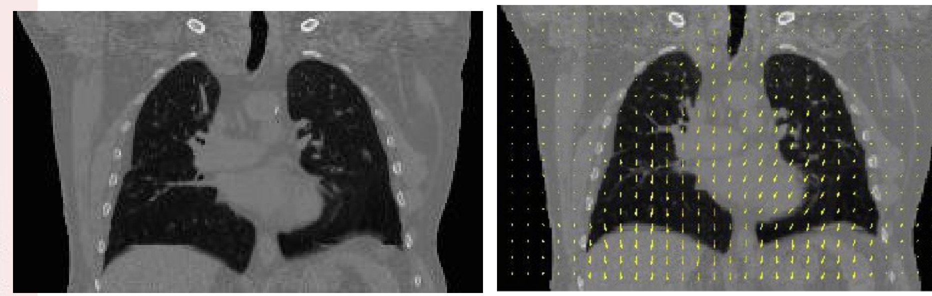
- many DOF : one vector for every image point

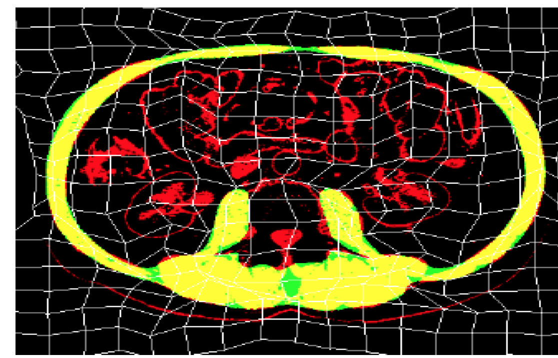
Looking for a deformation field (vector field) v that will “move” each voxel in image B (source) to the corresponding voxel in image A(target)

$$\min_{\mathbf{v}} \text{sum}_{x} (I_T(x) - I_S(x+\mathbf{v}))^2$$

2. Non-linear (deformable) registration







- many DOF : one vector for every image point

Looking for a deformation field (vector field) v that will “move” each voxel in image B (source) to the corresponding voxel in image A(target)

$$\min_{\mathbf{v}} \text{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}))^2$$

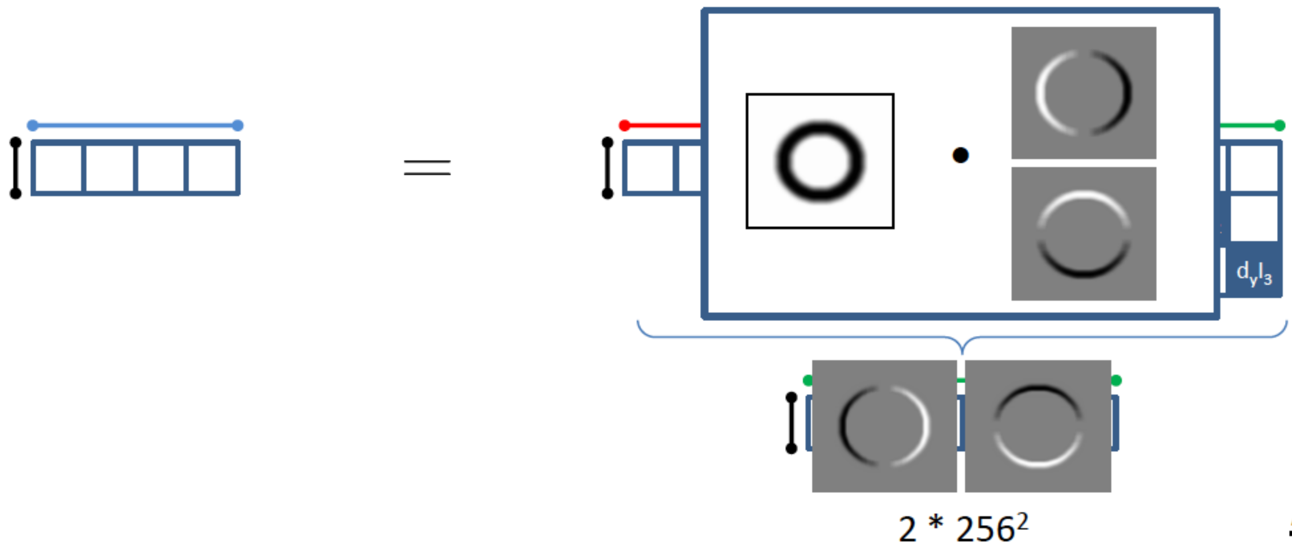
- ill-posed problem
- similar to optic flow but motion is large
- need for regularization of deformation field v

  > solution 1: reduce DOF – interpolation models (ex. Bsplines, FEM)

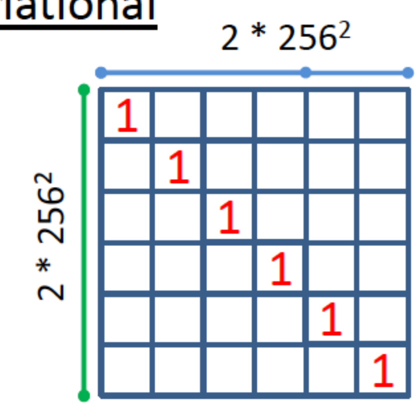
  > solution 2: add explicit regularization terms in the energy

Linear vs nonlinear registration

$$\frac{\partial\, \mathrm{SM}(X, Y(T_p))}{\partial p} = \frac{\partial\, \mathrm{SM}(X, Y)}{\partial Y} \quad \frac{\partial Y}{\partial T_p} \quad \frac{\partial T_p}{\partial p}$$

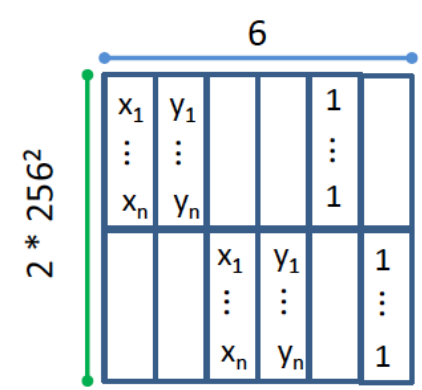


Variational



Affine

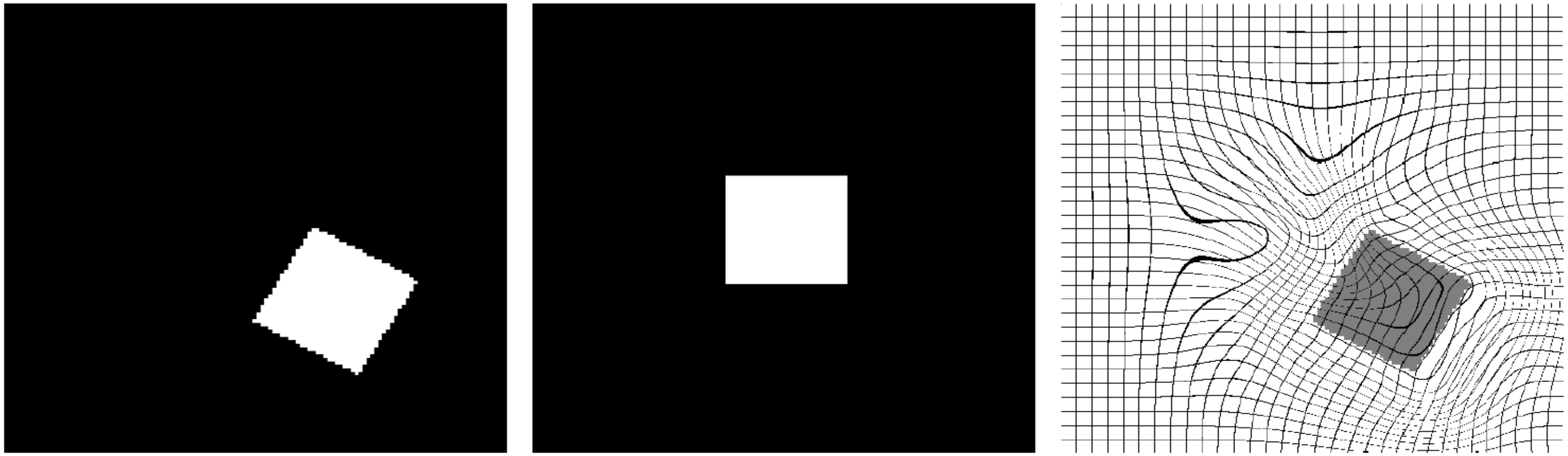
$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$



- Deformable registration is more sensitive to errors from poor modeling
- Averaged out in linear registration

Interpolation models
Parametric deformable models

- Bsplines (Free Form Deformations)
- Radial Basis Functions (RBF)
- Trigonometric functions (Discrete Fourier, Cosine basis)
- Finite element methods



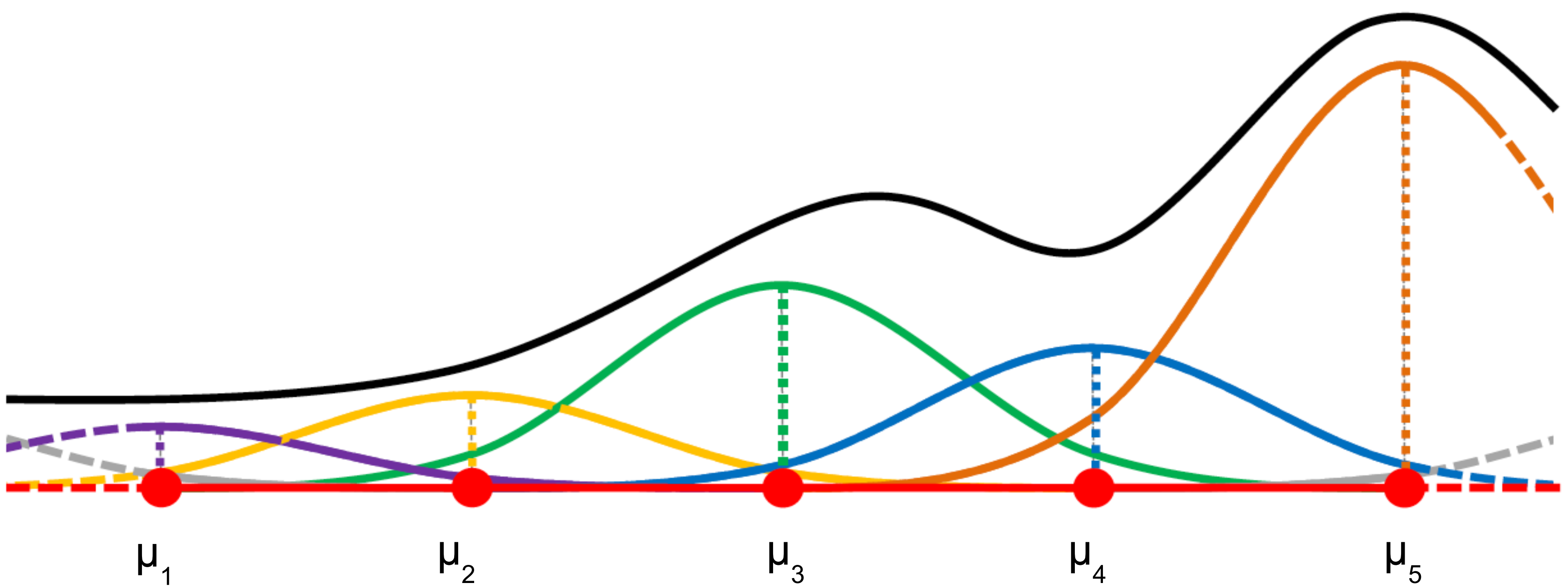
fixed moving deformed moving

Basis representation

$$u_p(x) = \sum_k \mu_k B_k(x)$$

Parameters: $\mu_k \in \mathbb{R}^d$

Basis functions: $B_k : \Omega \to \mathbb{R}$



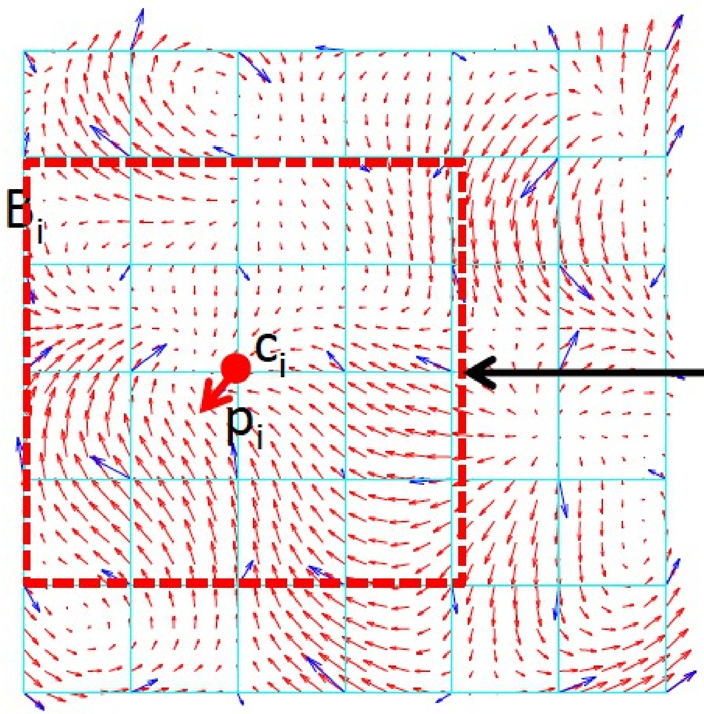
[Slides from Darko Zikic MICCAI 2010 registration tutorial]

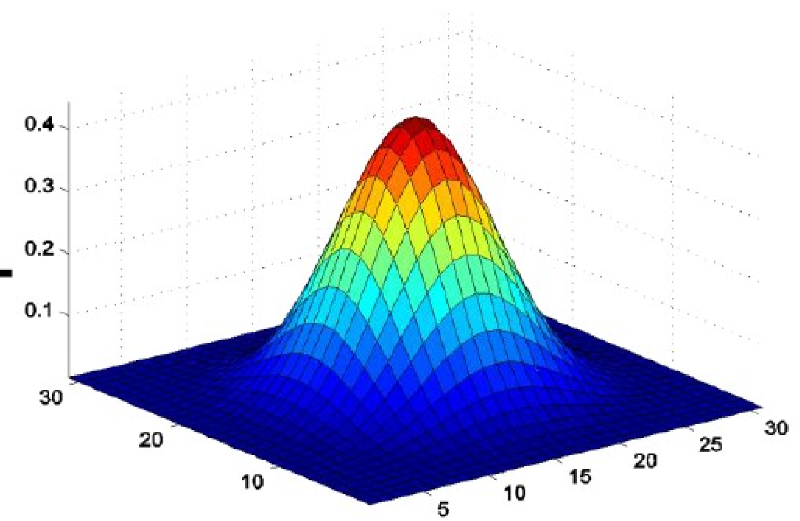
Basis representation

$$u_p(x) = \sum_k \mu_k B_k(x)$$

Parameters: $\mu_k \in \mathbb{R}^d$

Basis functions: $B_k : \Omega \to \mathbb{R}$

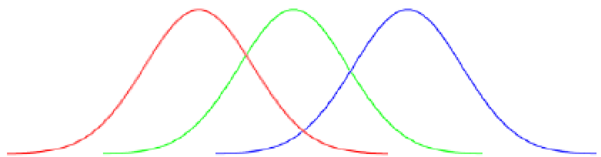




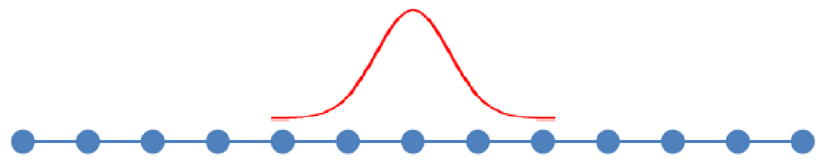
[Slides from Darko Zikic MICCAI 2010 registration tutorial]

B-spline parametrization

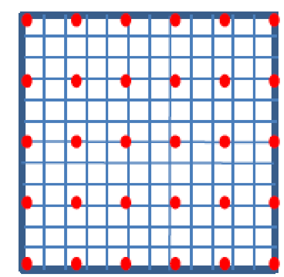
- **Rueckert 1999**
- **Cubic B-splines** (degree D = 3), basis functions all have same shape and are translated versions of each other



- **Compact support** of D+1 control points



- **Regular grid** of control points



- **More:** piecewise polynomial, inherent smoothness, differentiability, hierarchical

[Slides from Marius Staring MICCAI 2010 registration tutorial]

B-spline parametrization
Free Form Deformations (FFD)

- *Deformation* is modeled by B-splines

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

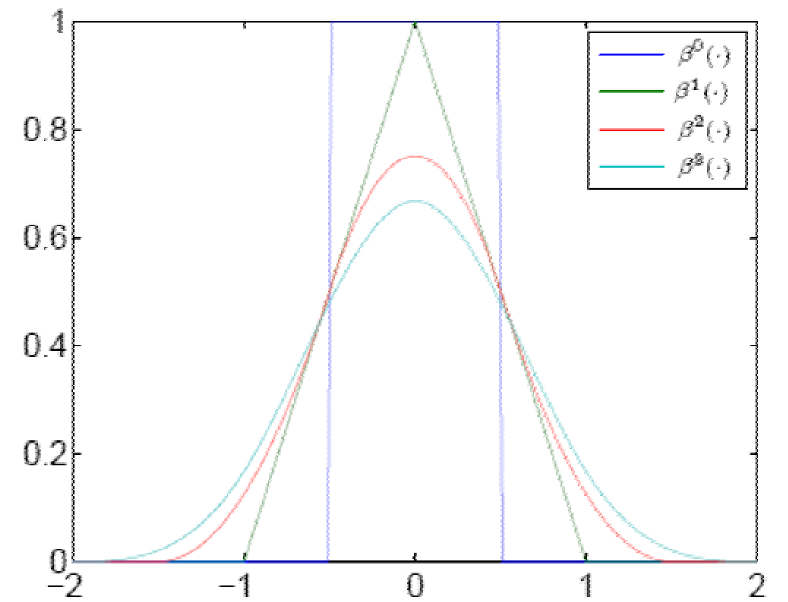
y_i control points

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

[Slides from Marius Staring MICCAI 2010 registration tutorial]

Basis functions

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



- Properties:
  - β^3 is twice differentiable $\rightarrow$ so is $\boldsymbol{T}_{\boldsymbol{\mu}}$
  - Convolution: $\beta^n(x) = \underbrace{\beta^0 * \cdots * \beta^0(x)}_{n+1 \text{ times}}$
  - Derivative can be done in terms of B-splines too:

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

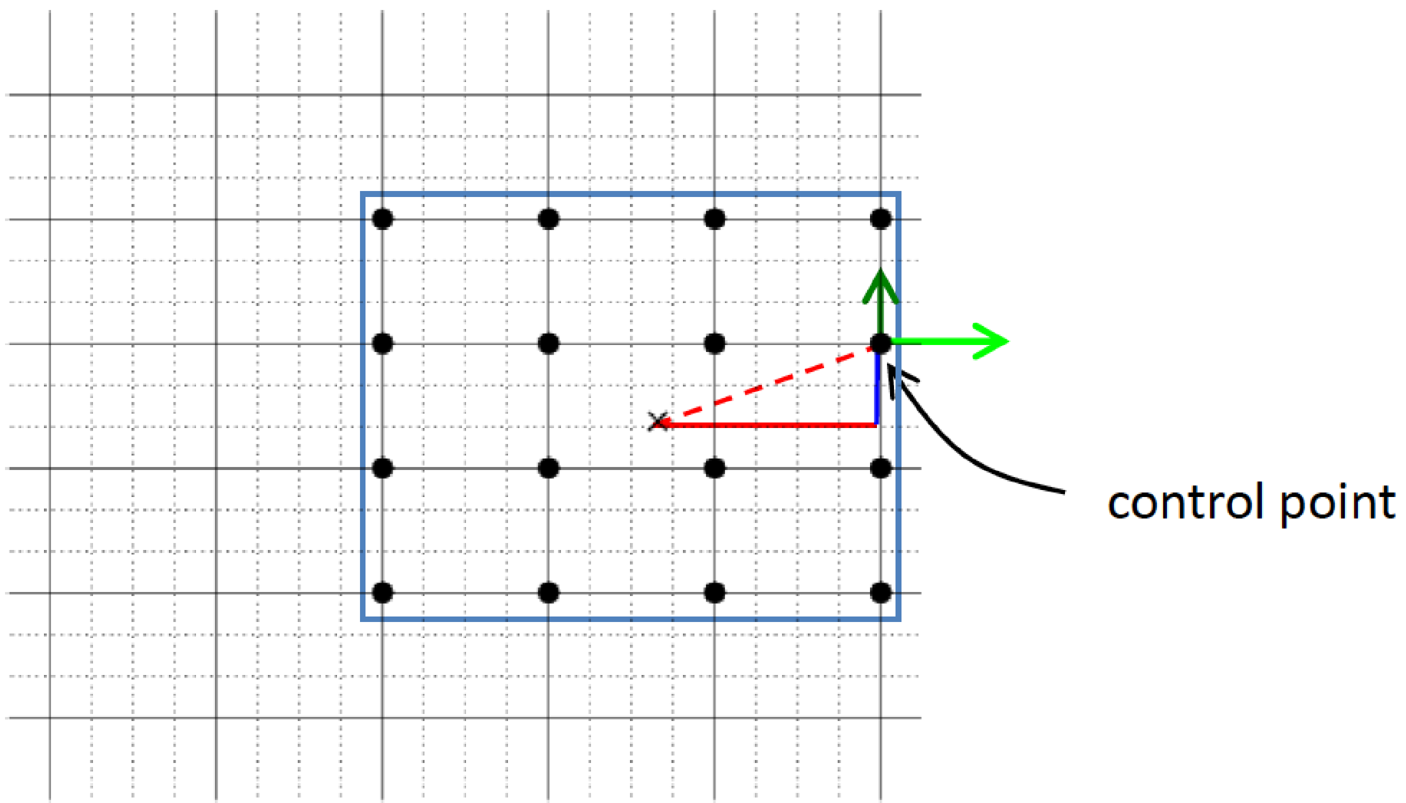
[Slides from Marius Staring MICCAI 2010 registration tutorial]

B-splines practically

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

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

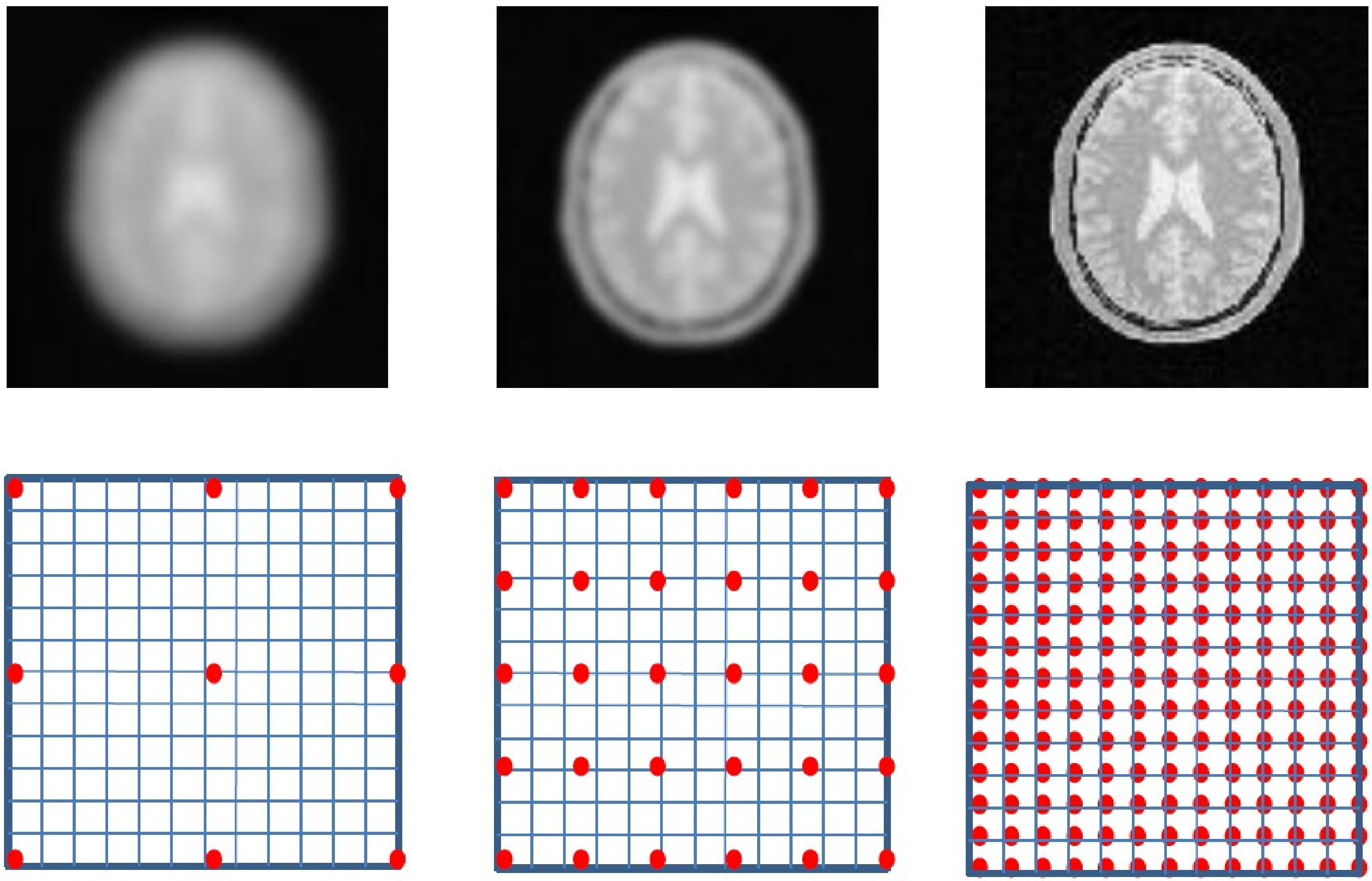
Translated versions of B



[Slides from Marius Staring MICCAI 2010 registration tutorial]

B-spline parametrization

- Multi-resolution: coarse to fine



[Slides from Marius Staring MICCAI 2010 registration tutorial]

Optimization with B-splines

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

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

Analytic derivatives of the warp

[Slides from Marius Staring MICCAI 2010 registration tutorial]

Optimization with B-splines

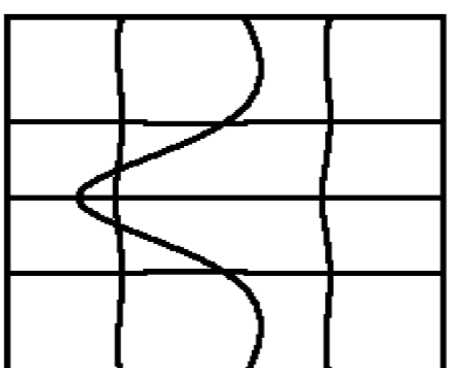
- $$\arg\min_{\boldsymbol{\mu}} \mathcal{C}(I_F, I_M, \boldsymbol{T}_{\boldsymbol{\mu}}), \qquad \boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - a_k \frac{\partial \mathcal{C}}{\partial \boldsymbol{\mu}_k}$$

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

Analytic derivatives of the warp

Regularization terms

folding



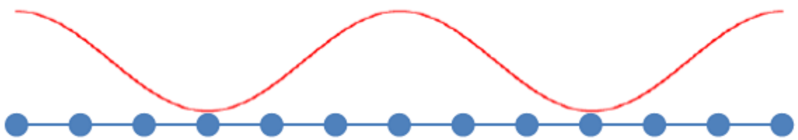
Valid deformation fields
Prevent fording
- Smoothness of T

$$P = f\left(\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{x}}, \frac{\partial^2 \boldsymbol{T}}{\partial \boldsymbol{x} \partial \boldsymbol{x}'}\right)$$

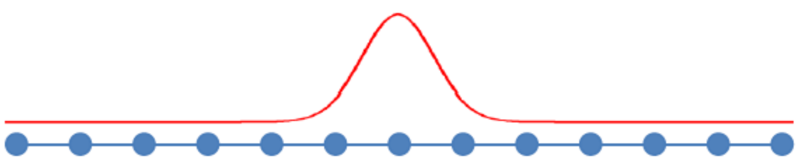
[Slides from Marius Staring MICCAI 2010 registration tutorial]

Other basis

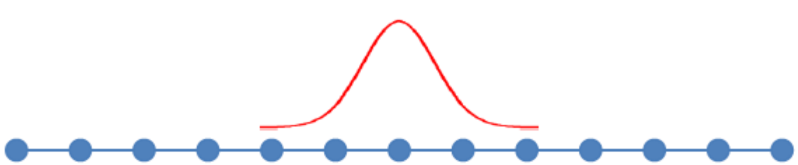
- Global Support
  - Fourier/Cosine Bases



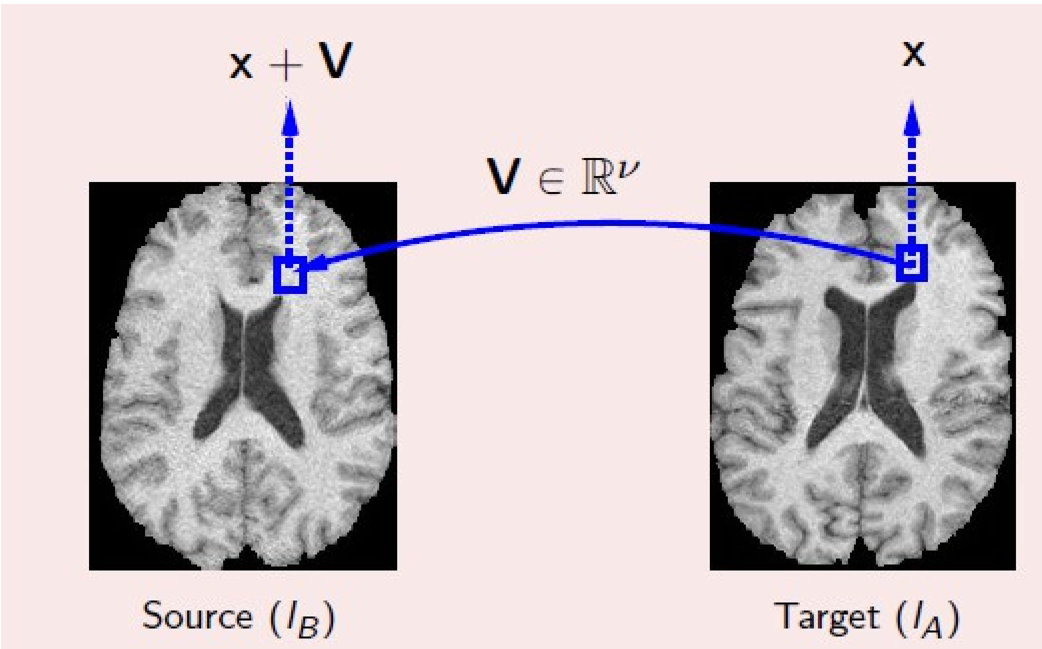
  - Radial basis functions RBFs (e.g. Thin-plate Splines (TPS))
  - Gaussians (in theory)

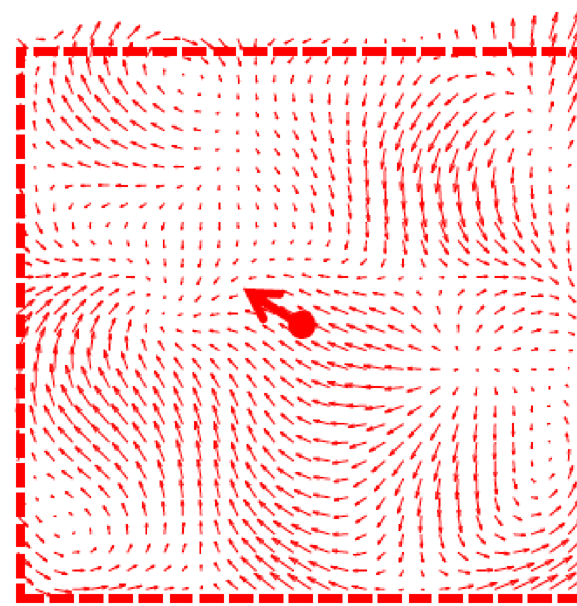


- Compact Support
  - B-Splines
  - Some RBFs
  - Gaussians (in practice)



Non-parametric models
Energy-based formulations





$$\min_{\mathbf{v}} \text{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}))^2$$

- Ill-posed > **regularization for v**
- Optimization : solve for v iteratively adding small updates delta **δv**

  Each step is similar to an **optic flow problem**

$$\min_{\delta\mathbf{v}} \text{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}+\delta\mathbf{v}))^2$$

Regularization and optimization

$$\min_{\mathbf{v}} D(I_T, I_S, v) + \alpha R(v)$$

SSD: $D(I_T, I_S, v) = \mathrm{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}))^2$

Diffusion : $| \nabla v|^2$

$$\mathbf{v = v + \delta v}$$

Regularization and optimization

$$\min_{\mathbf{v}} D(I_T, I_S, v) + \alpha R(v)$$

SSD: $D(I_T, I_S, v) = \text{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}))^2$

Diffusion : $|\nabla v|^2$

Iterative updates **$v = v + \delta v$**

1. Decouple data and regularization updates

Demons-like
Elastic regularization
[Thirion 1996/1998]
Regularize v

$\delta v = -\nabla D$
$v = v + \alpha \delta v$
$v = G * v$

Fluid-like
[Bro-Nielsen 1996, Pennec 1999]
Regularize δv

$\delta v = -\nabla D$
$\Delta v = G*(\alpha \delta v)$
$v = v + \delta v$

Regularization and optimization

$$\min_v D(I_T,I_S,v) + \alpha R(v)$$

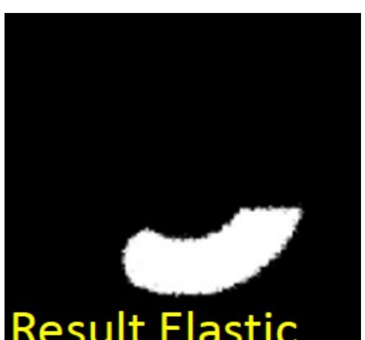
SSD: $D(I_T,I_S,v) = \text{sum}_x\, (I_T(x) - I_S(x+\mathbf{v}))^2$

Diffusion : $|\nabla v|^2$

Iterative updates $\mathbf{v} = \mathbf{v} + \boldsymbol{\delta}\mathbf{v}$

1. Decouple data and regularization updates

Images from [Christensen 1994]



Challenge to preserve regular deformations without folding

Regularization and optimization

$$\min_{\mathbf{v}} D(I_T, I_S, v) + \alpha R(v)$$

SSD: $D(I_T, I_S, v) = \text{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}))^2$

Diffusion : $|\nabla v|^2$

2. Optimize whole energy / Variational formulation

- Gradient descent update

$\delta v = -(\nabla D + \alpha \nabla R) = - \nabla E$

Regularization and optimization

$$\min_{\mathbf{v}} D(I_T, I_S, v) + \alpha R(v)$$

SSD: $D(I_T, I_S, v) = \text{sum}_x \, (I_T(x) - I_S(x+\mathbf{v}))^2$

Diffusion : $| \nabla v|^2$

2. Optimize whole energy > Variational formulation

- Gradient descent update

$\delta v = -(\nabla D + \alpha \nabla R) = - \nabla E$

Numerically unstable
Slow convergence

> Better optimization methods
ex. Gauss-Newton

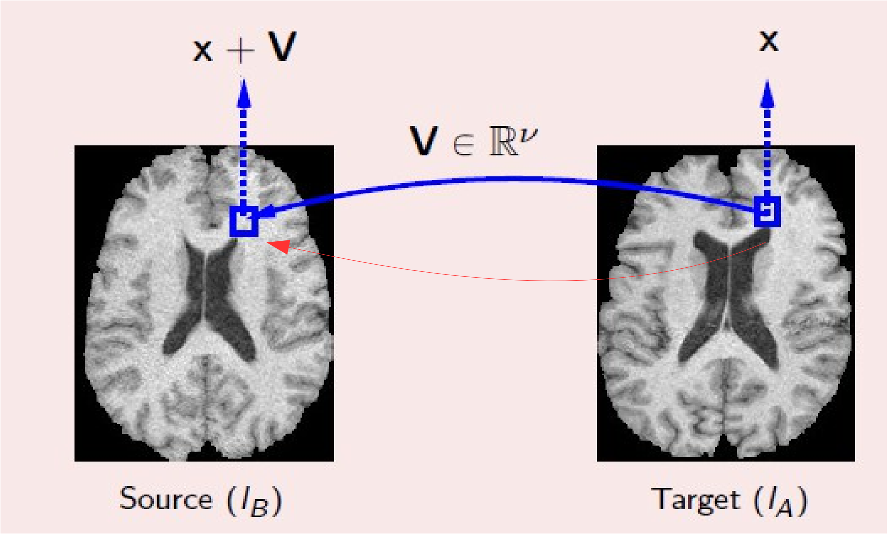
$\Delta v = -(J_e^T J_e)^{-1} \nabla E$

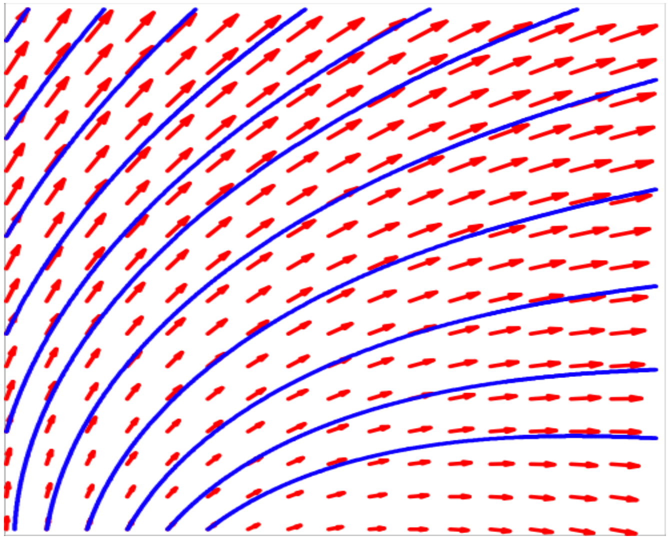
With $J_e^T J_e = \nabla I_S(v) \; \nabla I_S(v)^T - \alpha\Delta$

Diffeomorphic deformations

Diffeomorphic transformation
- Bijection – it can be inverted
- It's inverse is differentiable (smooth)

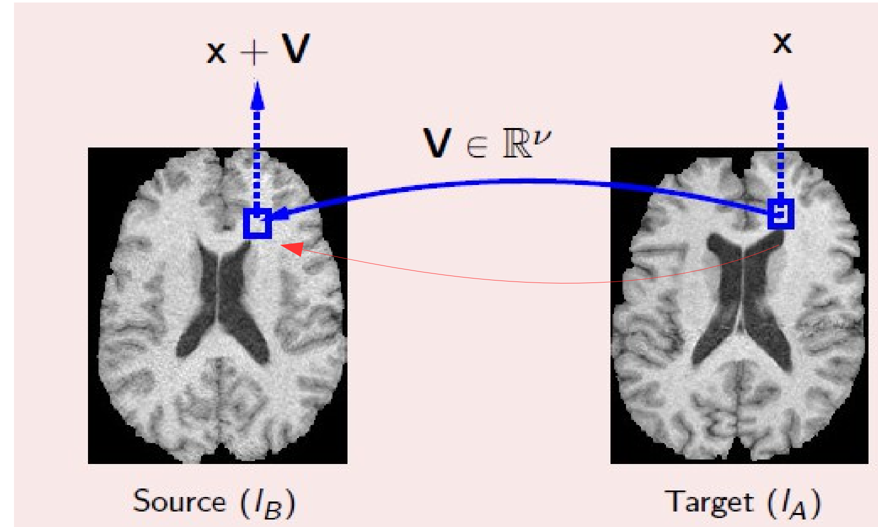
Practically
- No folding
- Preserves topology
- Essential for computational anatomy





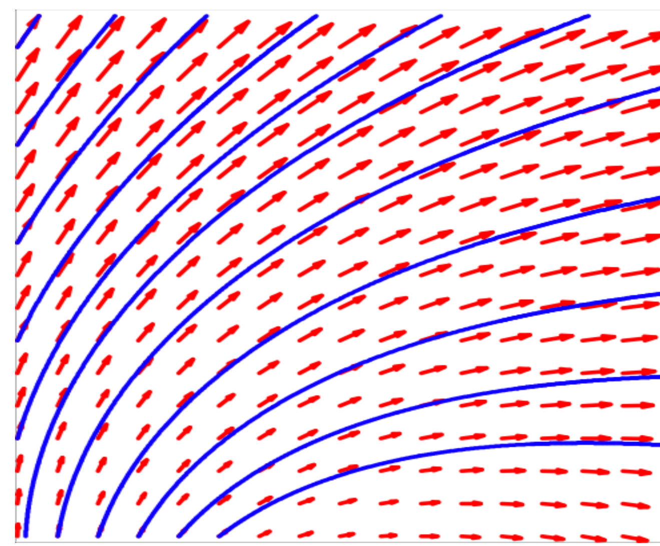
Diffeomorphic deformations

Diffeomorphic transformation
- Bijection – it can be inverted
- It's inverse is differentiable (smooth)

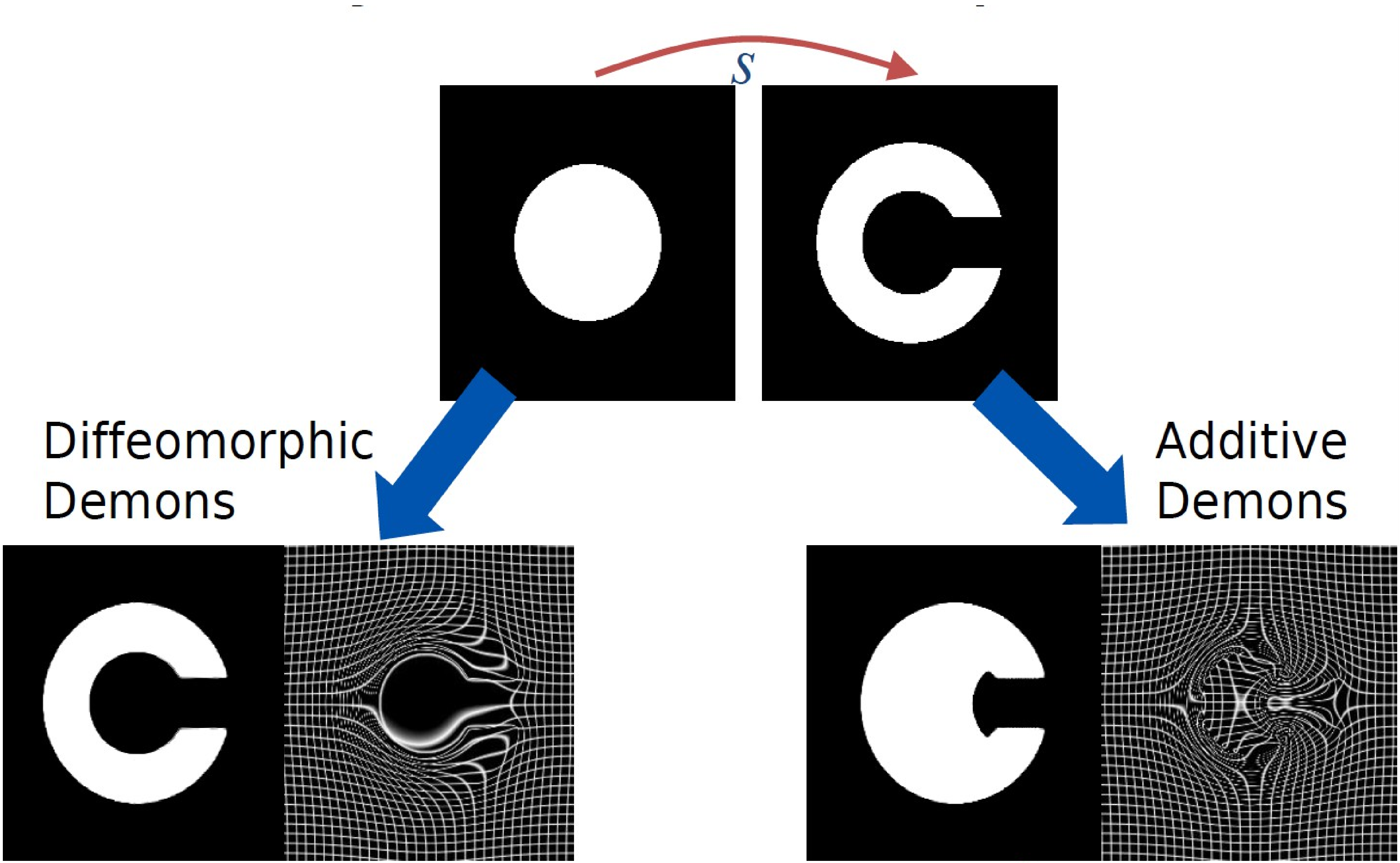


Practically
- No folding
- Preserves topology
- Essential for computational anatomy

Methods
- FFD: FFD are diffeomorphic by definition
- Diffeomorphic by formulation
  - Diffeomorphic Demons [Vercauterenet al., NeuroImage09]
  - LDDMM – Large Deformation Diffeomorphic Metric Mapping (deformable reg in Sobolev space) [Beg et al . IJCV 2005]
- Constrain field to be diffeomorphic after each optimization step [ANTS - SIM]

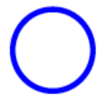


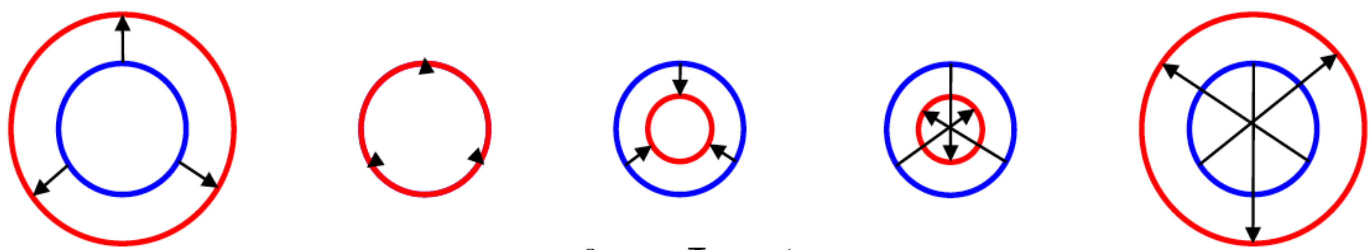
Diffeomorphic deformations



Jacobian of deformations

Effect of |J| change

 = volume x  = volume $T(x)$ $\quad \boldsymbol{J} = \det \nabla_x \boldsymbol{T}$



$\boldsymbol{J} > 1$ $\quad \boldsymbol{J} = 1$ $\quad 0 < \boldsymbol{J} < 1$ $\quad -1 < \boldsymbol{J} < 0$ $\quad \boldsymbol{J} < -1$

- Incompressibility:

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

$$\mathcal{R} = \sum_x \log \boldsymbol{J} \qquad \text{(Rohlfing 2003)}$$

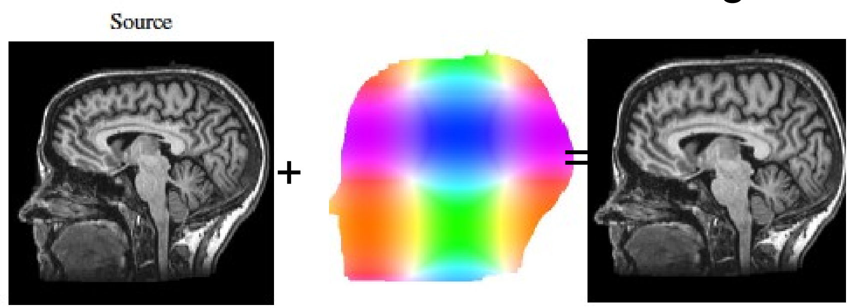
- Invertibility:

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

Evaluation methods for image registration

- Synthetic data

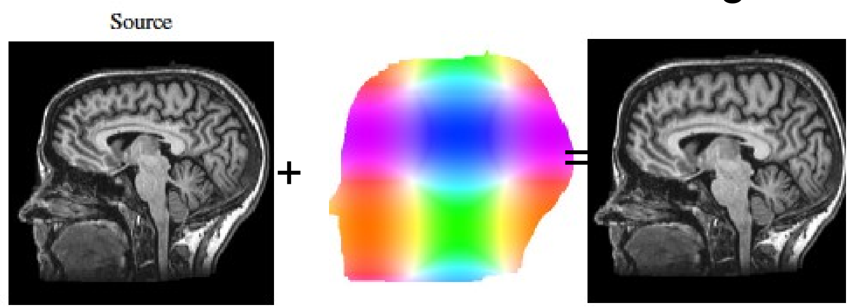
Source Deformation field Taget



Evaluation methods for image registration

- Synthetic data

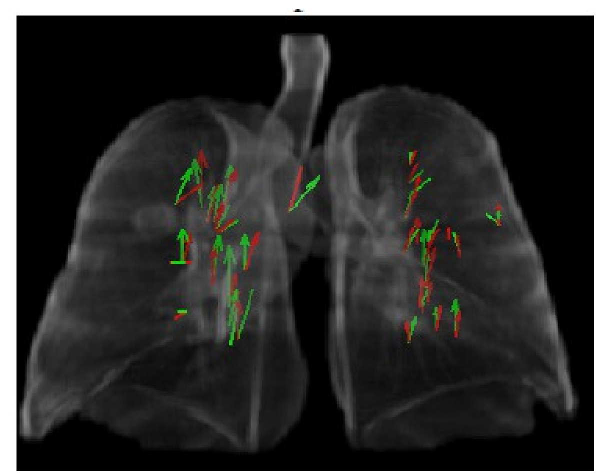
Source Deformation field Taget



- Manually annotated feature points in images

  Ex. POPI database

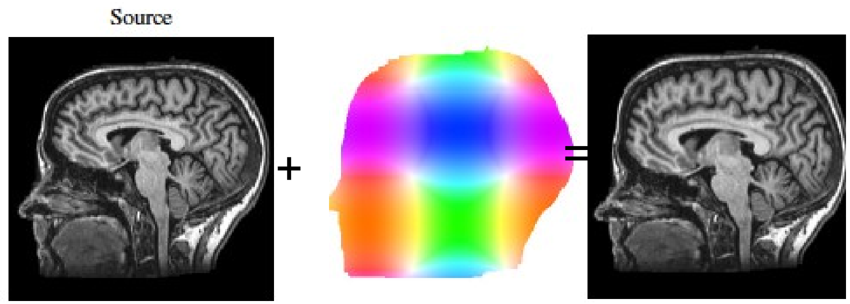
  10 volumes with 40 annotated landmarks



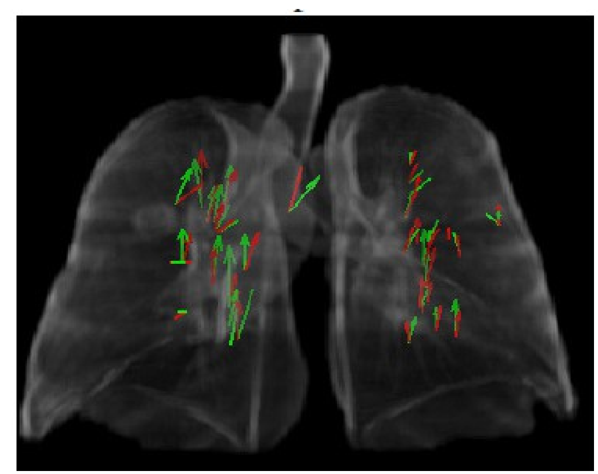
Evaluation methods for image registration

- Synthetic data

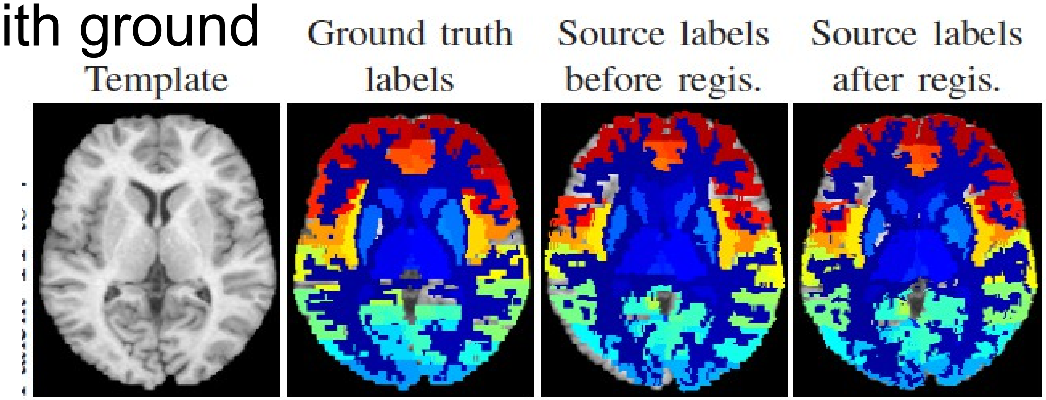
Source Deformation field Taget



- Manually annotated feature points in images
  Ex. POPI database
  10 volumes with 40 annotated landmarks



- Use matching of manual segmentations
  compare labels after registration with ground truth
  popular in brain registration
  Ex. CUMC12
  10 subjects, 128 regions



Practical considerations in image registration

$\min_p \; sim(I_S(T(p,x) - I_T(x)) \; [+ R(T)]$

Optimization Similarity Transformation

- Choose a suitable similarity score

SSD – same modality, no intensity variation
NCC – same modality , intensity variations
MI – different modalities

- Choose the transformation model

Linear, affine – same patient ; no additional distortion due to imaging (ultrasound-MRI)
Nonlinear – patient-atlas; inter-patients

- Optimization - Local methods:

Good initialization : always perform an affine registration first
Multiresolution approach
Smoothing helps

Software

Evaluation paper:
Compares several available software and methods
Klein et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration, Neuroimage 2009
Data at http://www.mindboggle.info/data.html

Software

Evaluation paper:
Compares several available software and methods
Klein et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration, Neuroimage 2009
Data at http://www.mindboggle.info/data.html

ANTS – Brian Avants U Penn
Very good linear and nonlinear diffeomorphic registration (SIM)

FSL
Easy to use

ITK – diffeomorphic demons
C++ library

Registration software based on ITK

Other “tools” that can do image registration: MedInria, 3D Slicer