AI in Medicine I

AI/ML for Medical Imaging – Image Registration

Veronika Zimmer
Chair for Computational Imaging and AI in Medicine
Departments of Informatics

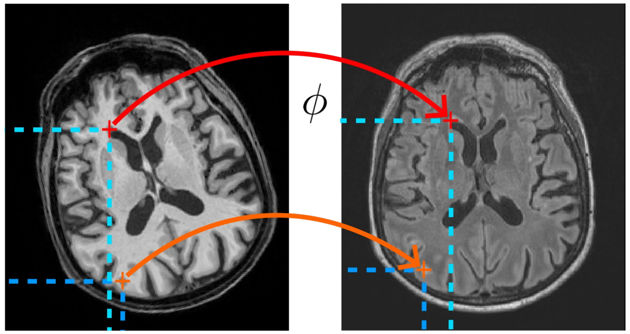
Overview

- Introduction to registration
  What is it? Why do we do this?
- Examples for registration
  What can we register? What are possible applications?
- Image Geometry and Image Warping
  How to spatially transform an image
- Registration Algorithms & Techniques
  How does it work? From classic to learning-based registration
- Evaluation of image registration
  How good is it?
- Learning-based image registration
  Can we learn it?
- Summary
  Take home messages

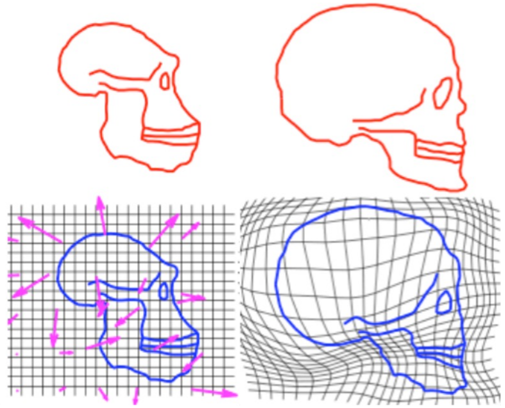
What is image registration?

- The aim is to establish spatial correspondences between two or multiple images.
- Image registration is the process of transforming sets of image data into one common coordinate system.

Establish anatomical correspondences



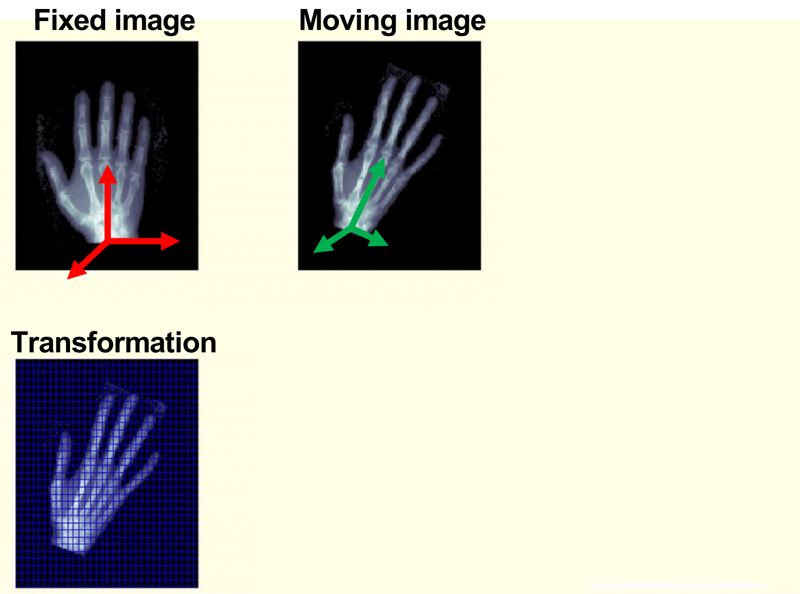
Deform underlying spatial domain



Durrleman, et al., *Morphometry of anatomical shape complexes with dense deformations and sparse parameters*, NeuroImage 2014.

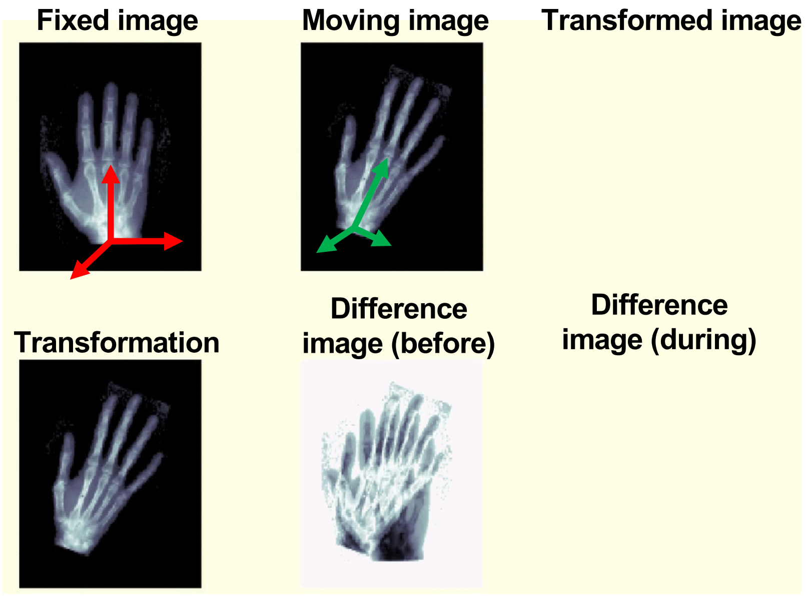
What is image registration?

- The aim is to establish spatial correspondences between two or multiple images.
- Image registration is the process of transforming sets of image data into one common coordinate system.



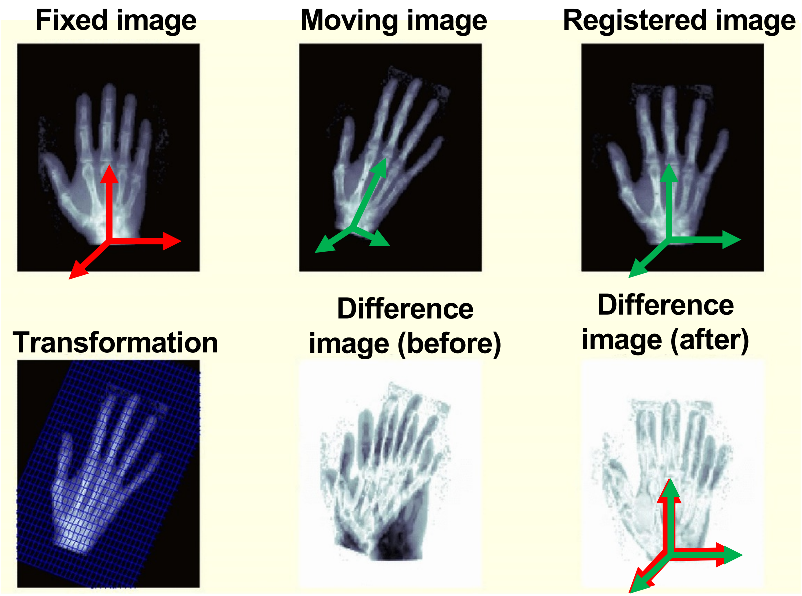
What is image registration?

- The aim is to establish spatial correspondences between two or multiple images.
- Image registration is the process of transforming sets of image data into one common coordinate system.



What is image registration?

- The aim is to establish spatial correspondences between two or multiple images.
- Image registration is the process of transforming sets of image data into one common coordinate system.



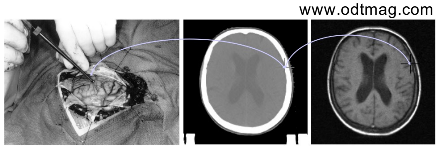
Clinical applications: When do we need it?

Finding **geometric**, **anatomical** and **functional** alignment for

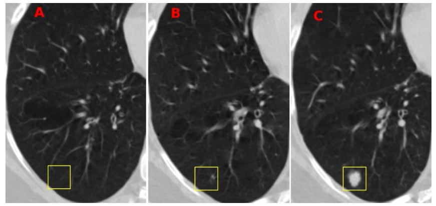
- Disease monitoring
- Motion analysis
- Growth analysis
- Assisted/Guided surgery
- Intervention and treatment planning
- Computer-aided diagnosis
- Population analysis

Image guided surgery

www.odtmag.com

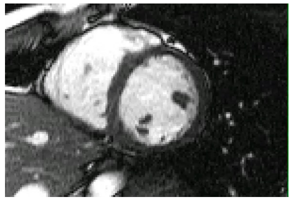


Disease monitoring

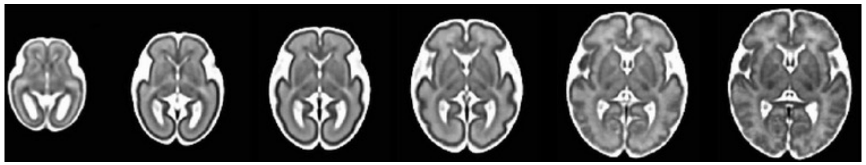


Tumour growth observed over 2 years from baseline

Cardiac motion



Healthy growth / development



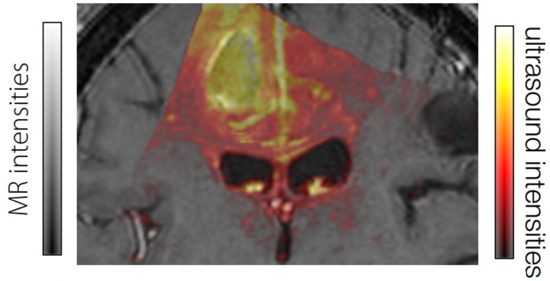
Dynamic movements

Intra-patient image registration

Registering images of the same patient.

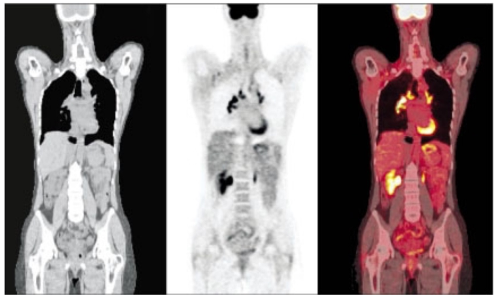
Applications:

(Multi-modal) image fusion

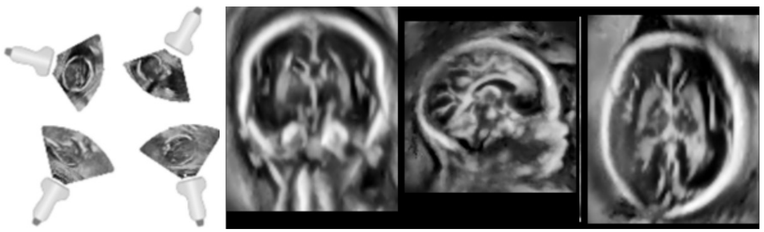


US-guided brain tumour surgery

Multimodal diagnostic by combining CT/PET scans



- Alignment of pre-interventional MRI to US (acquired during surgery)
- Combining CT and PET scans for diagnostic purposes
- Combining of multiple images for enhancing image quality



4D US image compounding

Intra-patient image registration

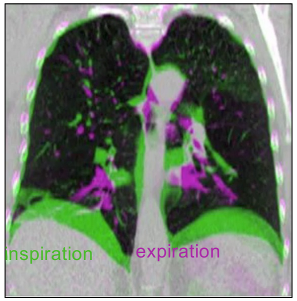
Registering images of the same patient.

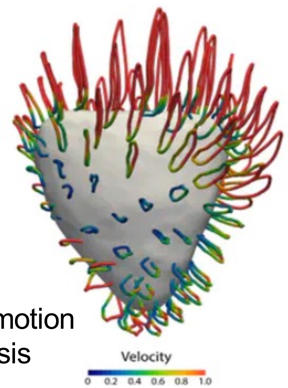
Applications:

Quantification of change

- Motion analysis
  (e.g., respiratory or cardiac motion)
- Dynamic registrations
  (e.g., for finding regions of contrast change in pre- and post-contrast breast MRI)

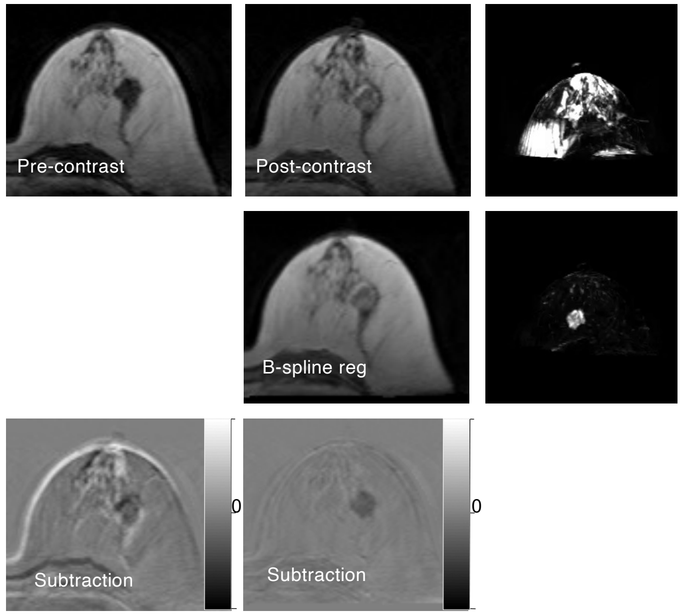
Estimation of respiratory motion





Cardiac motion analysis

Dynamic image registration



Inter-patient image registration

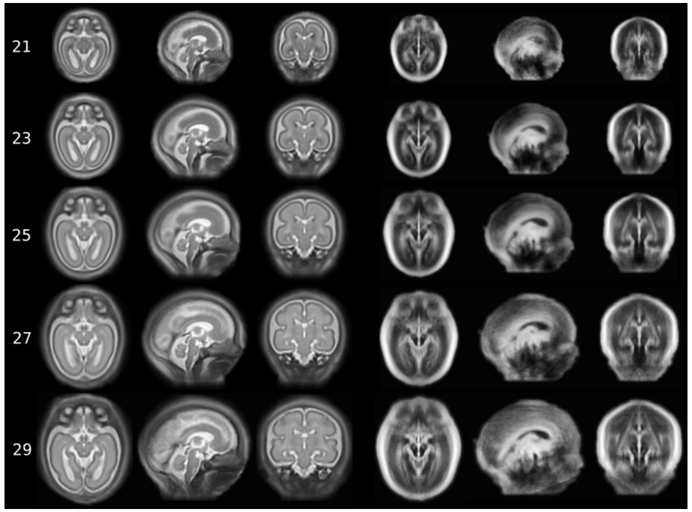
Registering images of different patients.

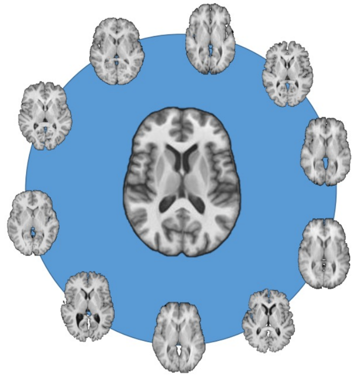
Applications:

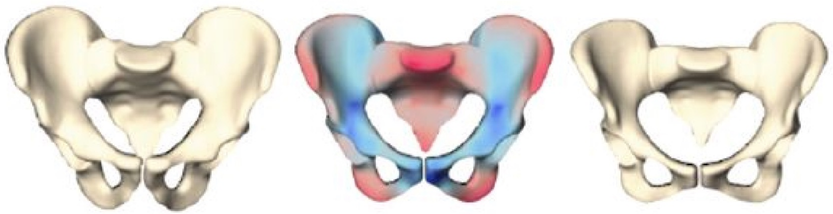
Population analysis

- Template construction of populations and/or subpopulations.
- Modelling of changes/main variations in a population, e.g., shape modelling.

Template / atlas construction







Shape modelling

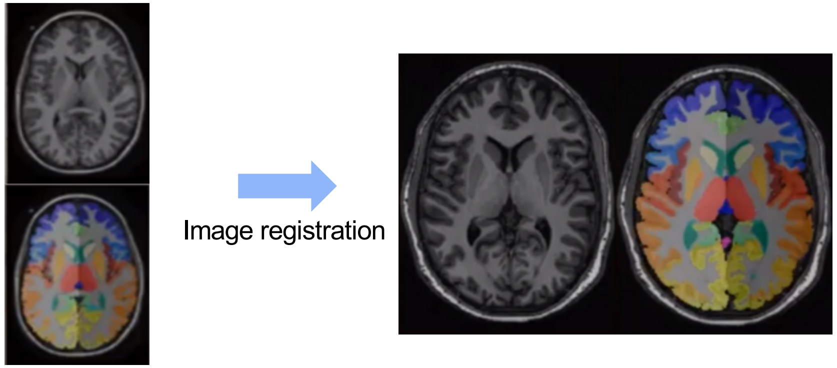
Inter-patient image registration

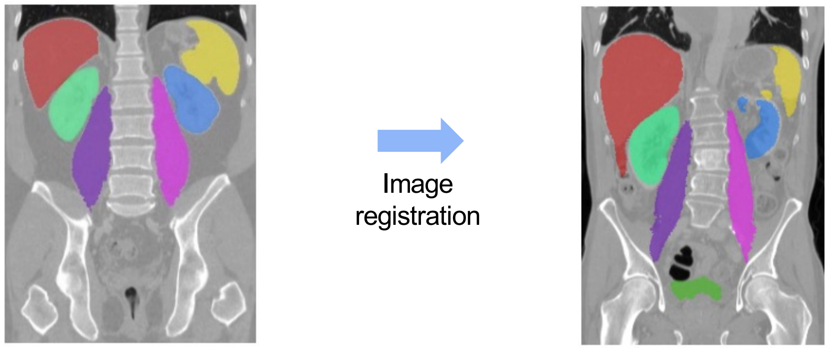
Registering images of different patients.

Applications:

Segmentation

- Multi-atlas segmentation





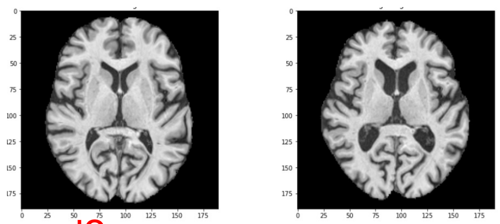
Atlas with expert segmentations

New patient scan

Main components

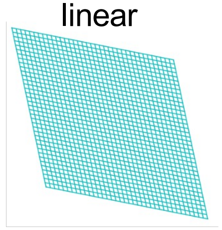
What do we register/align?

- **Fixed image** F defined on domain Ω_F
- **Moving image** M defined on domain Ω_M

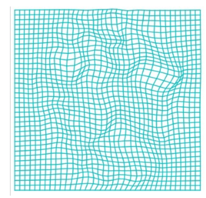


How to transform images?
What transformations are allowed?

linear



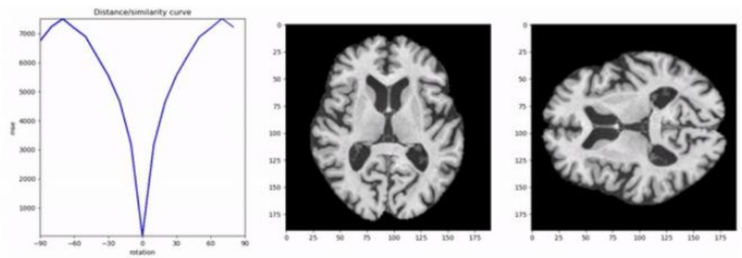
non- linear



- **Spatial transformation** ϕ mapping from Ω_F to Ω_M
- **Transformation model** Φ (space of transformations ϕ)
- Interpolation method

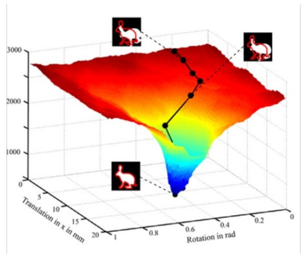
When are images similar?
How good is a transformation?

- **Objective / cost / energy function** $\mathcal{J}(\cdot,\cdot,\cdot)$
- Image similarity / distance measure $\mathcal{D}(\cdot,\cdot)$
- Regularization term $\mathcal{R}(\cdot)$



- **Optimization method**

How to find the optimal solution?



There exist many different notations in the field

Generic pairwise image registration algorithm

- Optimization problem

$$\phi^* = \underset{\phi}{\operatorname{argmin}}\, \mathcal{J}(\phi, F, M)$$

- Objective function

$$\mathcal{J}(\phi, F, M) = \mathcal{D}(F, M \circ \phi) + \alpha \mathcal{R}(\phi)$$

- Solved iteratively
- Often Gradient-based solvers
  → everything needs to be differentiable!

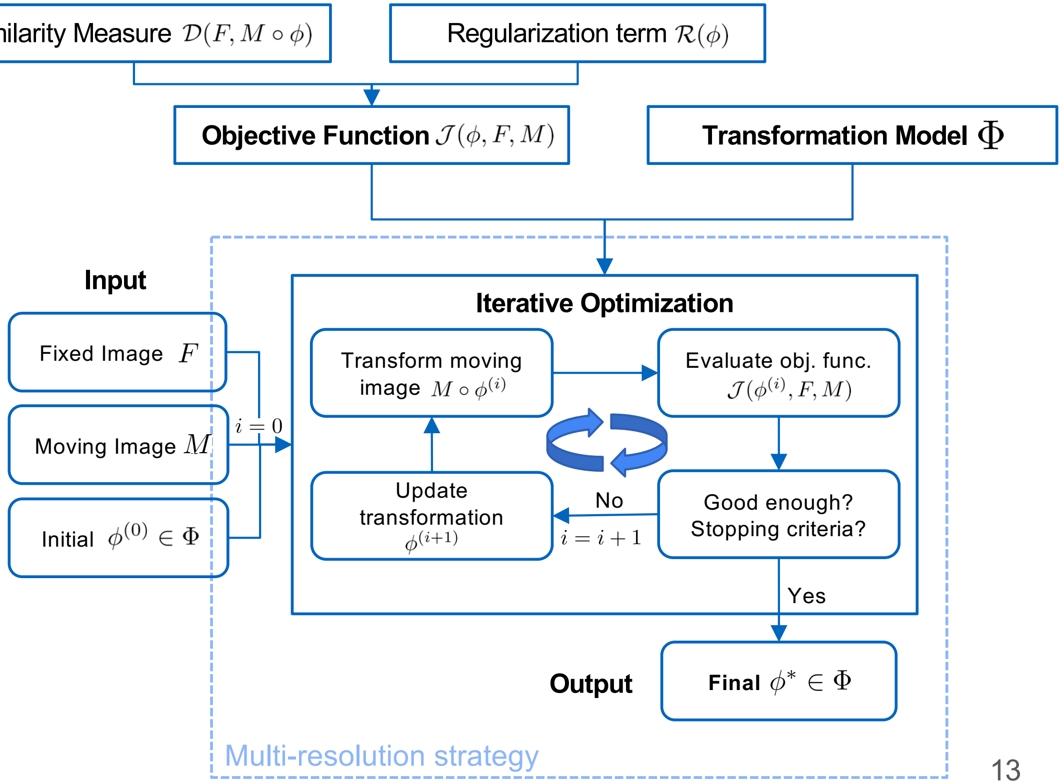


Image Geometry and Image Warping

Medical image representations

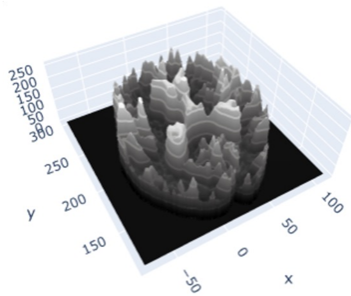
- **Continuous:** functions $F, M : \mathbb{R}^d \to \mathbb{R}$

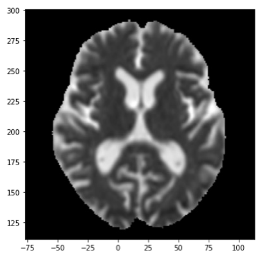
  $F(\mathbf{x}) \in \mathbb{R}, \mathbf{x} \in \Omega_F \subset \mathbb{R}^d$

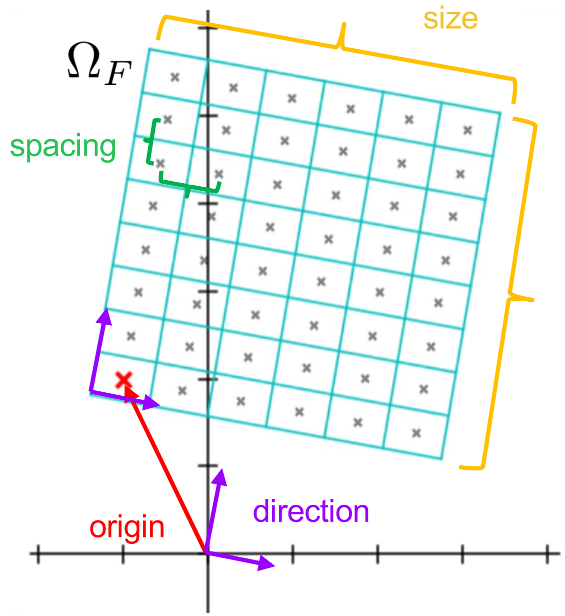
- **Discrete:** d-dim. Matrix defined on a regular grid $F, M \in \mathbb{R}^{m_1 \times \cdots \times m_d}$

  $\Omega_F = [0, m_1] \times \cdots \times [0, m_d]$

- **Meta information:** Pixel / voxel spacing
  Image origin
  Image direction / orientation







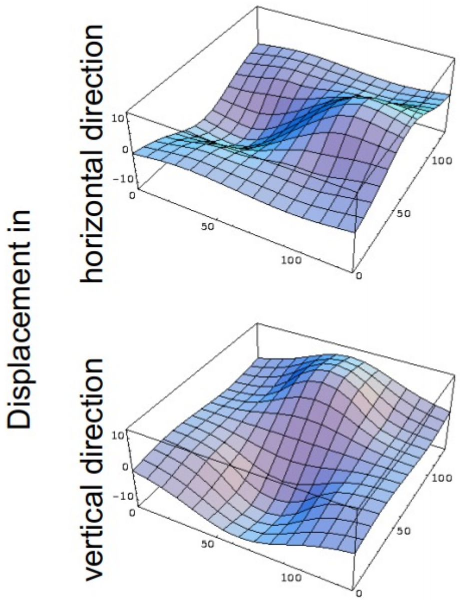
World coordinate system

Spatial transformations: Representations

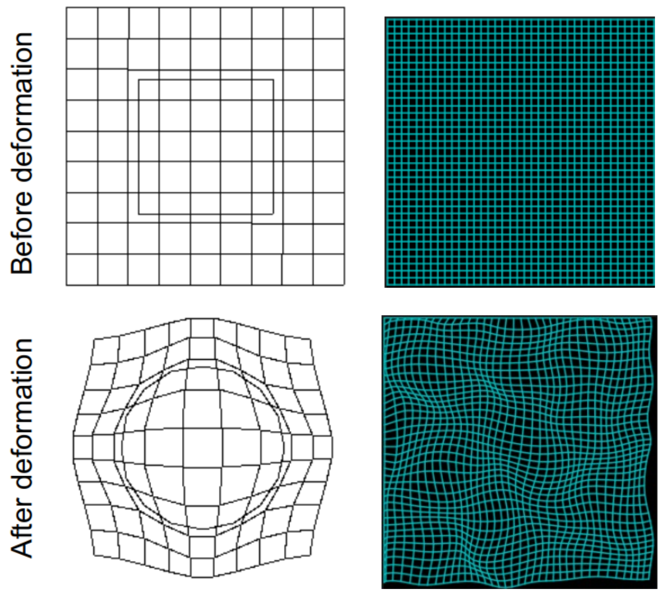
- Spatial transformations $\phi : \Omega_F \subset \mathbb{R}^d \rightarrow \Omega_M \subset \mathbb{R}^d$

$$M \circ \phi = M(\phi)$$

Visualization separate per dim



Visualization as deformed grid

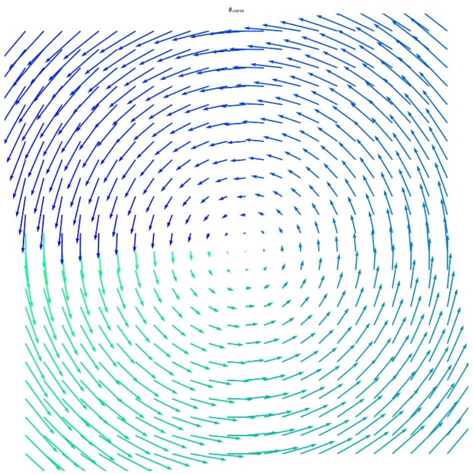


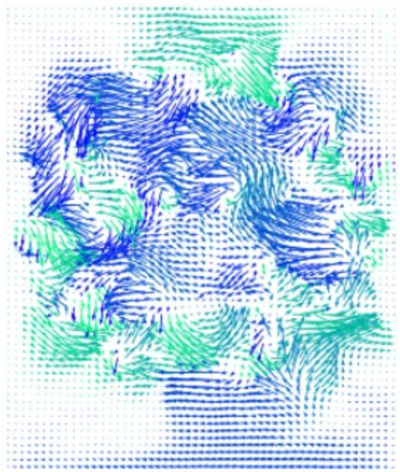
Spatial transformations: Representations

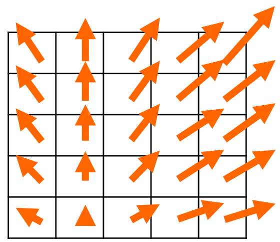
- Spatial transformations $\phi : \Omega_F \subset \mathbb{R}^d \rightarrow \Omega_M \subset \mathbb{R}^d$

$$M \circ \phi = M(\phi)$$

Visualization as vector field







Spatial transformations: Global/Linear

- Spatial transformations $\phi : \Omega_F \subset \mathbb{R}^d \rightarrow \Omega_M \subset \mathbb{R}^d$

$$M \circ \phi = M(\phi)$$

- **Global (or linear) transformations models (parametric)**

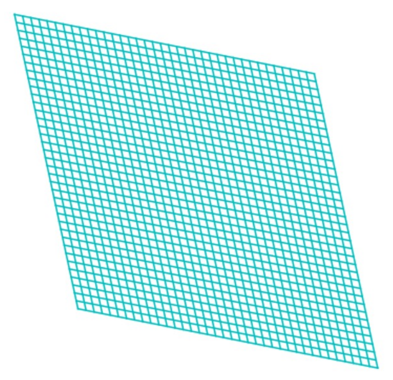
$$\phi : \mathbb{R}^p \times \mathbb{R}^d \rightarrow \mathbb{R}^d$$

(in homogenous coordinates) $\phi(\theta, \mathbf{x}) = A \cdot \mathbf{x} \quad A = \begin{pmatrix} A_{11} & A_{12} & b_1 \\ A_{21} & A_{22} & b_2 \\ 0 & 0 & 1 \end{pmatrix}$

Affine parameters in 2D:

$$\theta = (A_{11}, A_{12}, b_1, A_{21}, A_{22}, b_2)^T \in \mathbb{R}^6, (d = 2, p = 6)$$

linear

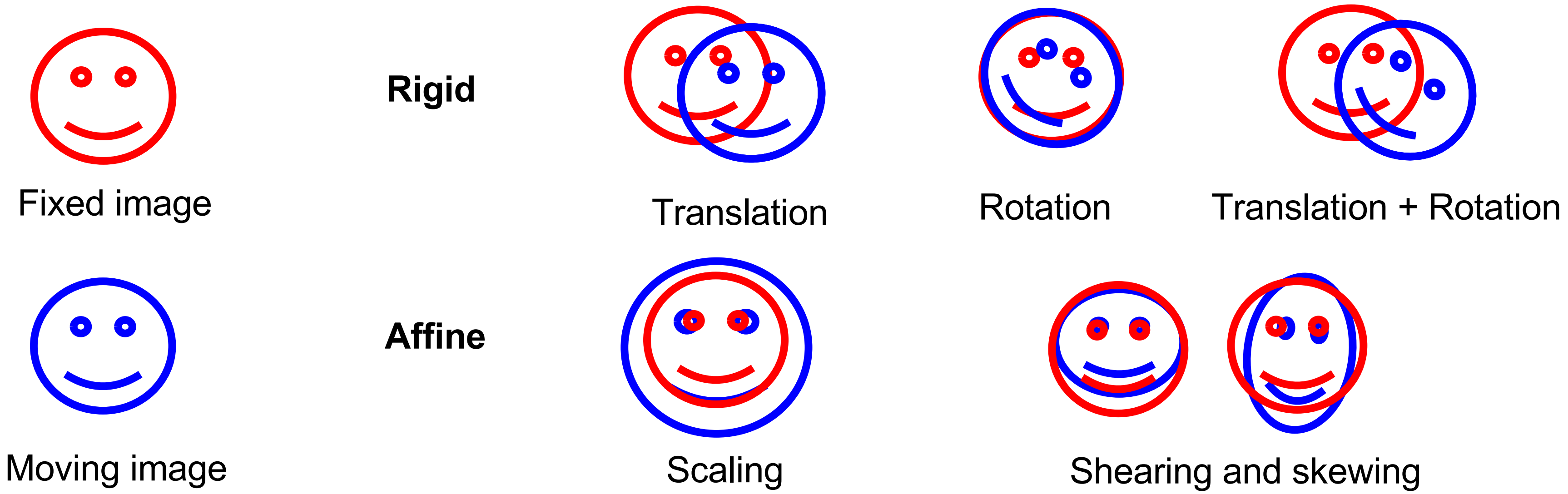


Spatial transformations: Global/Linear

- Spatial transformations $\phi : \Omega_F \subset \mathbb{R}^d \rightarrow \Omega_M \subset \mathbb{R}^d$

  $M \circ \phi = M(\phi)$

- **Global (or linear) transformations models (parametric)**



Rigid transformation model

$$\boxed{\phi_{\text{rigid3D}}(\mathbf{x};\mathbf{p}) = R_{xyz} \cdot \mathbf{x} + \mathbf{t}}$$

Voxel coordinate $\mathbf{x} \in \Omega \subset \mathbb{R}^3$

3D Rotation matrix $R_{xyz} \in \mathrm{SO}(3) \subset \mathbb{R}^{3\times 3}$

Translation vector $\mathbf{t} \in \mathbb{R}^3$

Parameter vector $\mathbf{p} \in \mathbb{R}^6$

3D Rotation matrix

$$R_{xyz} = R_z \cdot R_y \cdot R_x = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \quad R_y = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \quad R_z = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Parameter vector

in 3D!

$$\mathbf{p} = (\alpha, \quad \beta, \quad \gamma, \quad t_x, \quad t_y, \quad t_z)^T \in \mathbb{R}^6$$

Rigid transformation model

$$\phi_{\text{rigid3D}}(\mathbf{x};\mathbf{p}) = R_{xyz} \cdot \mathbf{x} + \mathbf{t}$$

Voxel coordinate $\mathbf{x} \in \Omega \subset \mathbb{R}^3$

3D Rotation matrix $R_{xyz} \in \mathrm{SO}(3) \subset \mathbb{R}^{3\times 3}$

Translation vector $\mathbf{t} \in \mathbb{R}^3$

Parameter vector $\mathbf{p} \in \mathbb{R}^6$

In homogeneous coordinates

$$\phi_{rigid3D}(\mathbf{x};\mathbf{p}) = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = R_{xyzt} \cdot \mathbf{x}$$

Parametric image registration

$$\boxed{\mathbf{p}^* = \underset{\phi_{\mathbf{p}}}{\operatorname{argmin}}\, \mathcal{J}(F, M \circ \phi_{\mathbf{p}})}$$

Affine transformation model

$$\boxed{\phi_{\text{affine3D}}(\mathbf{x};\mathbf{p}) = A \cdot \mathbf{x} + \mathbf{t}}$$

Voxel coordinate	$\mathbf{x} \in \Omega \subset \mathbb{R}^3$
3D Affine matrix	$A \in \mathbb{R}^{3\times 3}$
Translation vector	$\mathbf{t} \in \mathbb{R}^3$
Parameter vector	$\mathbf{p} \in \mathbb{R}^{12}$

3D Affine matrix (in homogeneous coordinates)

$$A = A_{\text{shear}} \cdot A_{\text{scale}} \cdot R_{xyzt} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

e.g., $A_{\text{shear}} = \begin{pmatrix} 1 & 0 & sh_x & 0 \\ 0 & 1 & sh_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

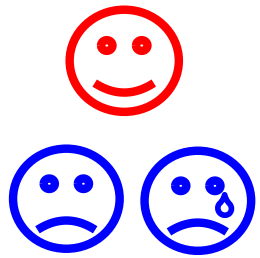
$$A_{\text{scale}} = \begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Spatial transformations: Local/Non-linear

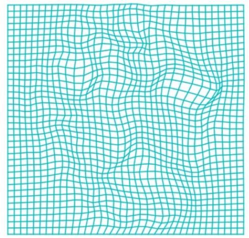
- Spatial transformations $\phi : \Omega_F \subset \mathbb{R}^d \rightarrow \Omega_M \subset \mathbb{R}^d$ $\quad\quad \phi(\mathbf{x}) = \mathbf{x} + \mathbf{u}$

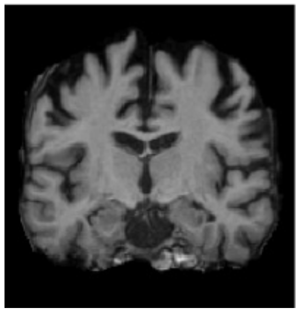
  $M \circ \phi = M(\phi)$

- **Local (or non-linear) transformation models**
- There may be local deformations **u** due to tissue changes, organ motion, tissue growth (healthy or pathological), tissue loss (due to atrophy or resection), …

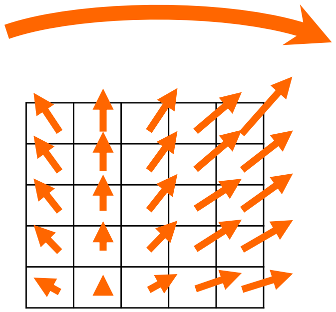


non- linear

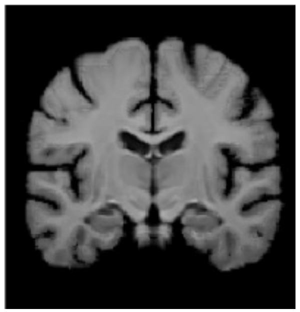




moving scan M



field ϕ



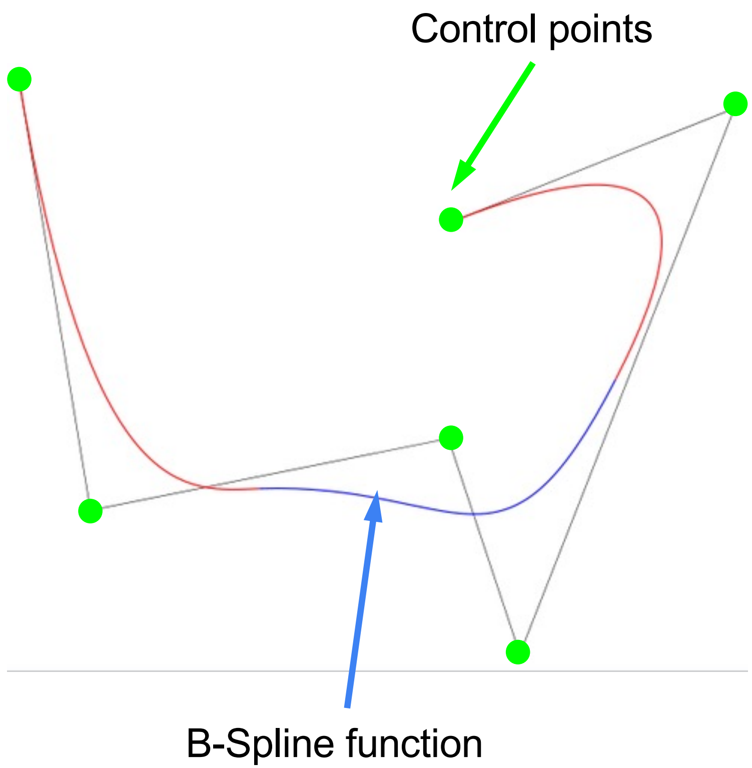
fixed scan F

Spatial transformations: Local/Non-linear

- **Parametric local** transformation models

  Free-Form-Deformations via B-Splines

- B-spline basis functions for smooth, continuous interpolations



$$B_0(s) = (1-s)^3 \qquad B_2(s) = (-3s^3 + 3s^2 + 3s + 1)/6$$

$$B_1(s) = (3s^3 - 6s^2 + 4)/6 \qquad B_3(s) = s^3/6$$

Spatial transformations: Local/Non-linear

- **Parametric local** transformation models

  Free-Form-Deformations via B-Splines

- B-spline basis functions for smooth, continuous interpolations
- Construct a regular mesh of control points and apply 3D tensor product of 1D B-splines
- Increasing mesh resolution for multi-level B-splines

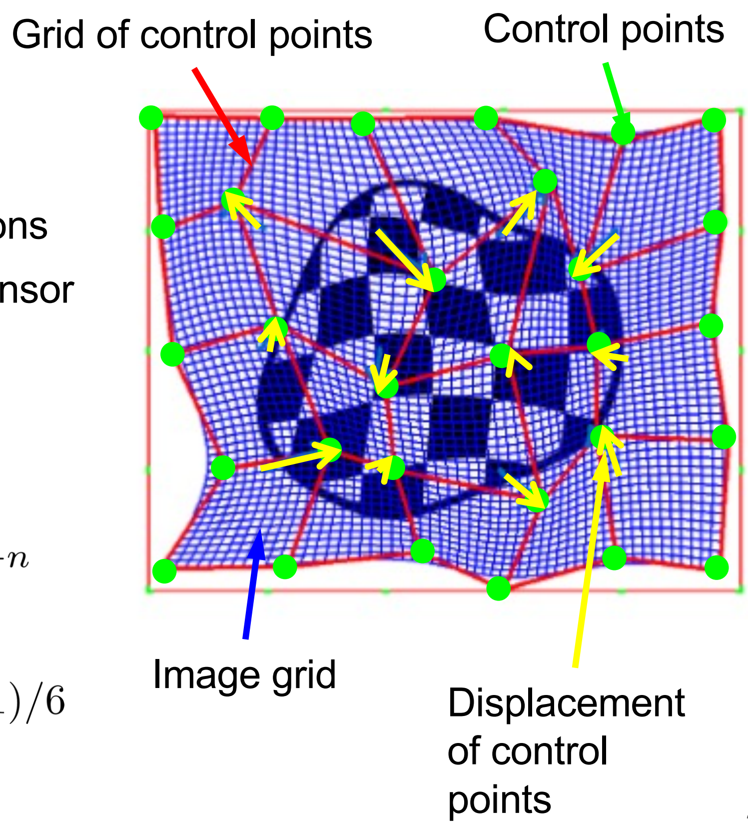
$$\mathbf{u}(\mathbf{x}) = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)c_{i+l,j+m,k+n}$$

$B_0(s) = (1-s)^3$

$B_1(s) = (3s^3 - 6s^2 + 4)/6$

$B_2(s) = (-3s^3 + 3s^2 + 3s + 1)/6$

$B_3(s) = s^3/6$

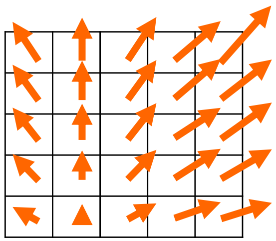


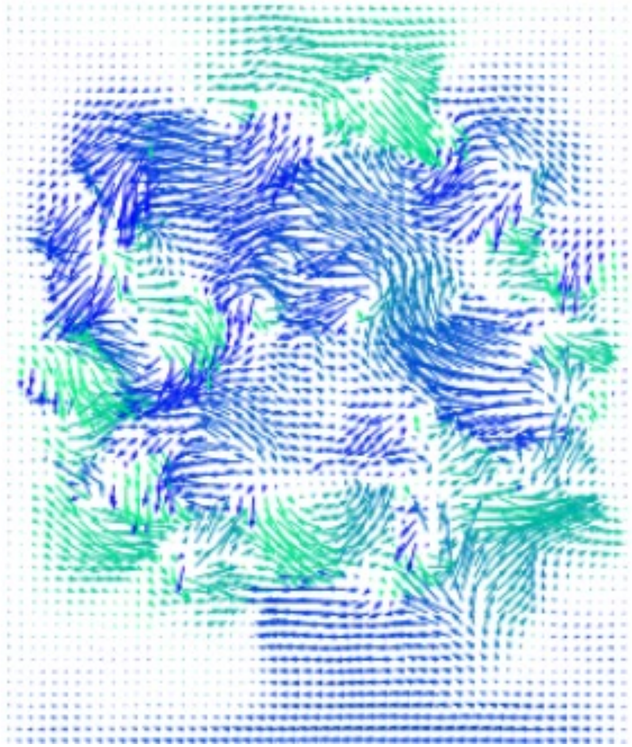
Spatial transformations: Local/Non-linear

- **Non-parametric** transformation models

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

- With displacements $\mathbf{u} \in \mathbb{R}^d$
- Dense displacement fields
- The transformation model **is** the deformation field





Dense vs. parametric deformation models

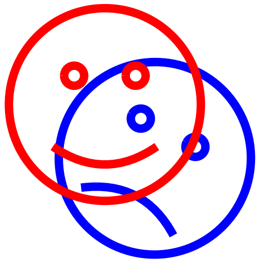
We can use a **dense deformation field**, ie one displacement vector per image pixel

- Lots of Degrees of Freedom (3x image dimension), very flexible locally, but not so much globally
- Needs a lot of regularisation to avoid collapsing / intersecting deformations
- Slow in optimisation, large in storage/memory

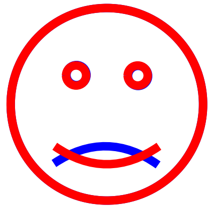
We can use a **compact, parametric representation of the deformations**, that can be densely sampled:

- Fewer DoFs, very flexible globally, not so much locally
- Often has built-in regularisation due to smooth model assumptions
- Fast in optimisation, compact in storage/memory

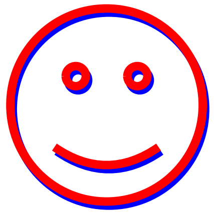
Combine global and local registration



Global + local motion



After global registration



After local registration

$$\text{Using addition: } \mathbf{x}' = T_{total}(\mathbf{x}) = T_{local}(\mathbf{x}) + T_{global}(\mathbf{x})$$

$$\boxed{\text{Using composition: } \mathbf{x}' = T_{total}(\mathbf{x}) = T_{local}(\mathbf{x}) \circ T_{global}(\mathbf{x})}$$

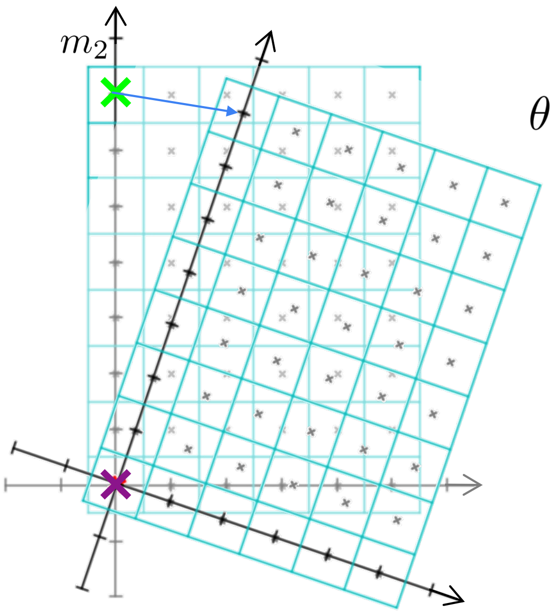
Parametric transformation models

	identity	rigid	rigid+scaling	affine	non-linear
3D	0	6	9	12	10’s – 100,000’s
2D	0	3	4	6	10’s – 10,000’s

Number of Parameters

Transformation of an image domain/grid

Image coordinate system



Center of rotation

$$\phi(\theta, \mathbf{x}) = A \cdot \mathbf{x}$$

$$\theta = (\alpha, t_x, t_z, c_x, c_y)$$
$$(20°, 0, 0, 1.5, 4.5)$$

Rotation by 20° around the point (1.5, 4.5)

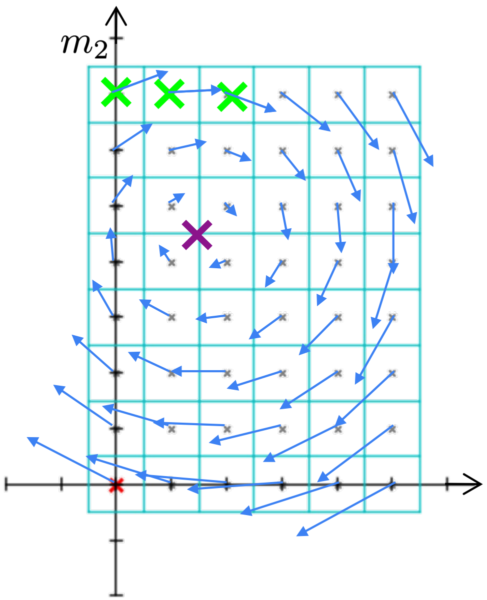
Rotation matrix in 2D

$$R = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Domain $\Omega = [0, m_1] \times [0, m_2]$

Transformation of an image domain/grid

Image coordinate system



Center of rotation

$$\phi(\theta, \mathbf{x}) = A \cdot \mathbf{x}$$

$$\theta = (\alpha, t_x, t_z, c_x, c_y)$$
$$(20°, 0, 0, 1.5, 4.5)$$

Rotation by 20° around the point (1.5, 4.5)

Rotation matrix in 2D

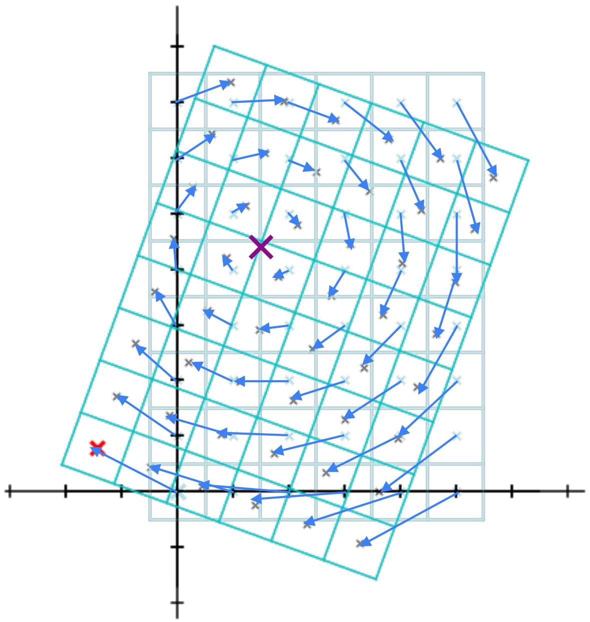
$$R = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 0 & t_x \\ 0 & 0 & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & -t_x \\ 0 & 0 & -t_y \\ 0 & 0 & 1 \end{pmatrix}$$

- Domain $\Omega = [0, m_1] \times [0, m_2]$

Transformation of an image domain/grid

Image coordinate system



Center of rotation

$$\phi(\theta, \mathbf{x}) = A \cdot \mathbf{x}$$

$$\theta = (\alpha, t_x, t_z, c_x, c_y)$$
$$(20°, 0, 0, 1.5, 4.5)$$

Rotation by 20° around the point (1.5, 4.5)

Rotation matrix in 2D

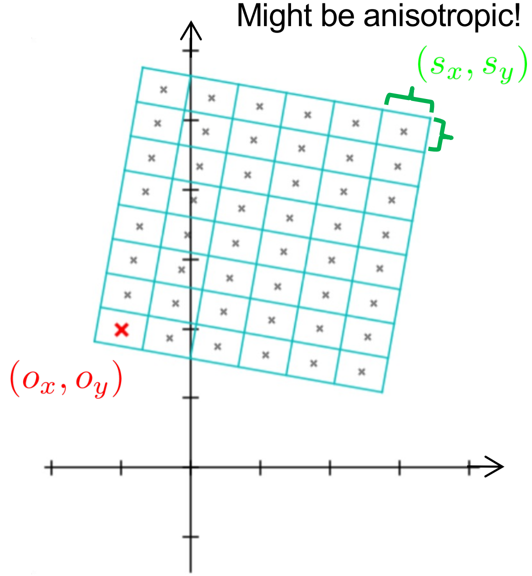
$$R = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 0 & t_x \\ 0 & 0 & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & -t_x \\ 0 & 0 & -t_y \\ 0 & 0 & 1 \end{pmatrix}$$

- Domain $\Omega = [0, m_1] \times [0, m_2]$

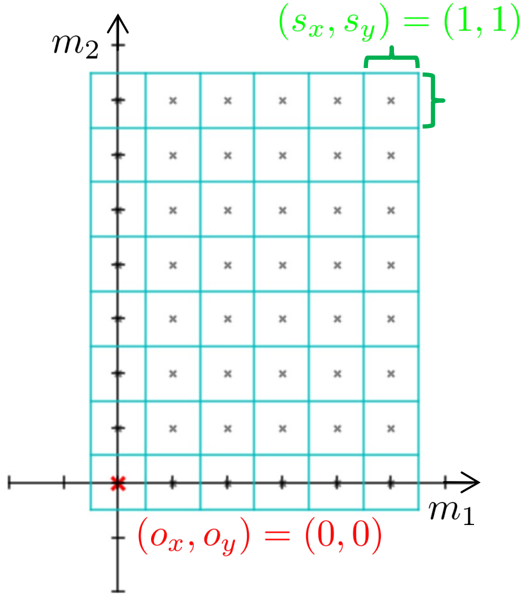
Image coordinate systems

World coordinate system



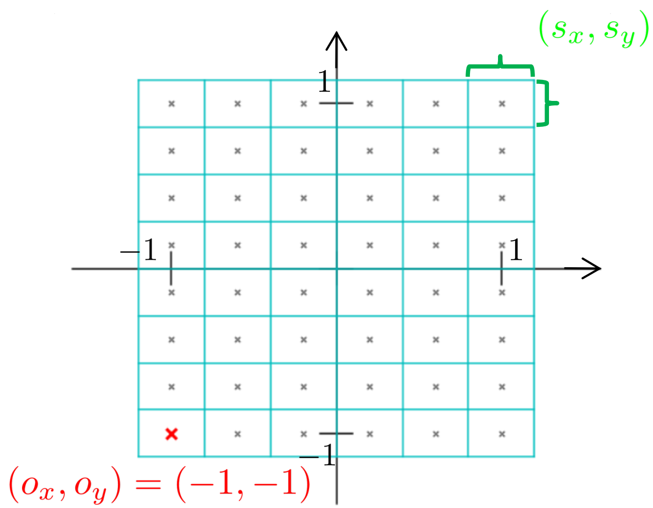
- Origin : (o_x, o_y)
- Spacing in mm
- Orientation
- Domain Ω

Image coordinate system



- Origin : (0,0)
- Spacing in px
- Domain $\Omega = [0, m_1] \times [0, m_2]$

Normalized coordinate system



- Origin : (-1,-1)
- Spacing : 2 / (sz-1)
- Domain

$$\Omega = [-1, 1] \times [-1, 1]$$

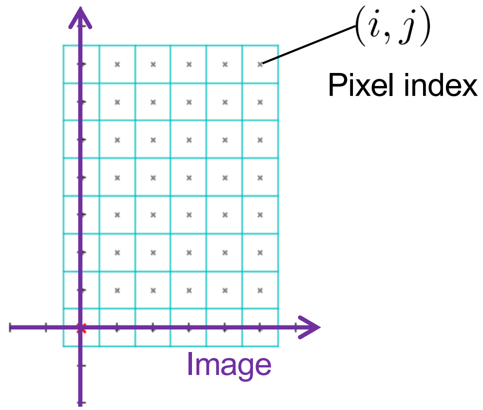
Change between coordinate systems

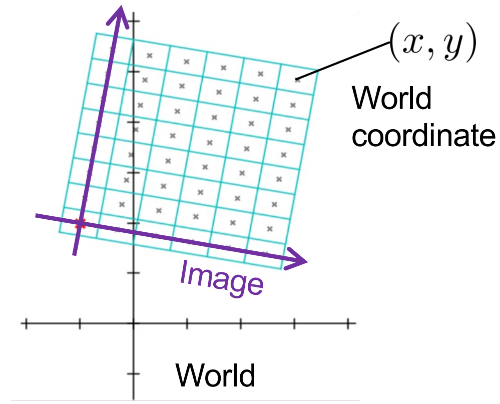
Image to world

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = T_{i2w} \begin{pmatrix} i \\ j \\ 1 \end{pmatrix}$$

$$T_{i2w} = \begin{pmatrix} 1 & 0 & \color{red}{\boldsymbol{o_x}} \\ 0 & 1 & \color{red}{\boldsymbol{o_y}} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \color{purple}{\boldsymbol{d_{xx}}} & \color{purple}{\boldsymbol{d_{yx}}} & 1 \\ \color{purple}{\boldsymbol{d_{xy}}} & \color{purple}{\boldsymbol{d_{yy}}} & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \color{green}{\boldsymbol{s_x}} & 0 & 0 \\ 0 & \color{green}{\boldsymbol{s_y}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Image origin — Image direction — Pixel spacing





World to image

$$T_{w2i} = T_{i2w}^{-1}$$

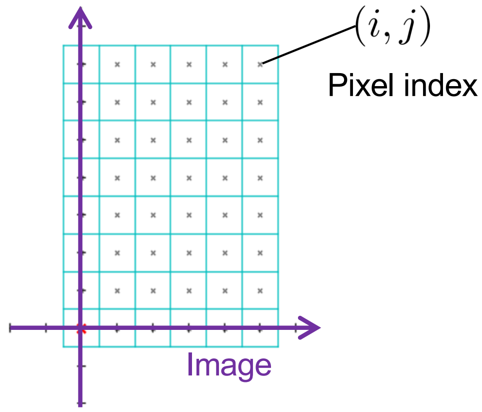
$$\begin{pmatrix} i \\ j \\ 1 \end{pmatrix} = T_{w2i} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Change between coordinate systems

Image to normalized domain

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = T_{i2n} \begin{pmatrix} i \\ j \\ 1 \end{pmatrix}$$

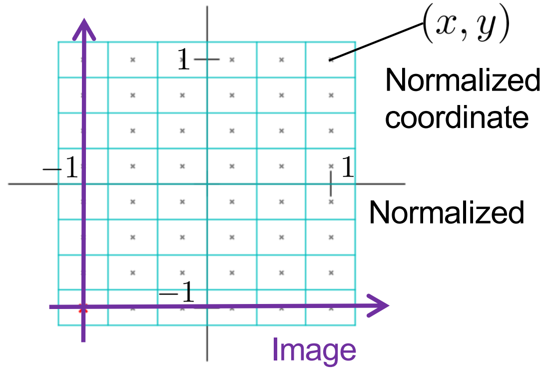
$$T_{i2n} = \begin{pmatrix} 1 & 0 & \mathbf{-1} \\ 0 & 1 & \mathbf{-1} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{1} & \mathbf{0} & 1 \\ \mathbf{0} & \mathbf{1} & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{\mathbf{2}}{\boldsymbol{m_1-1}} & 0 & 0 \\ 0 & \frac{\mathbf{2}}{\boldsymbol{m_2-1}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Normalized domain to image

$$T_{n2i} = T_{i2n}^{-1}$$

$$\begin{pmatrix} i \\ j \\ 1 \end{pmatrix} = T_{n2i} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$



More than one coordinate systems

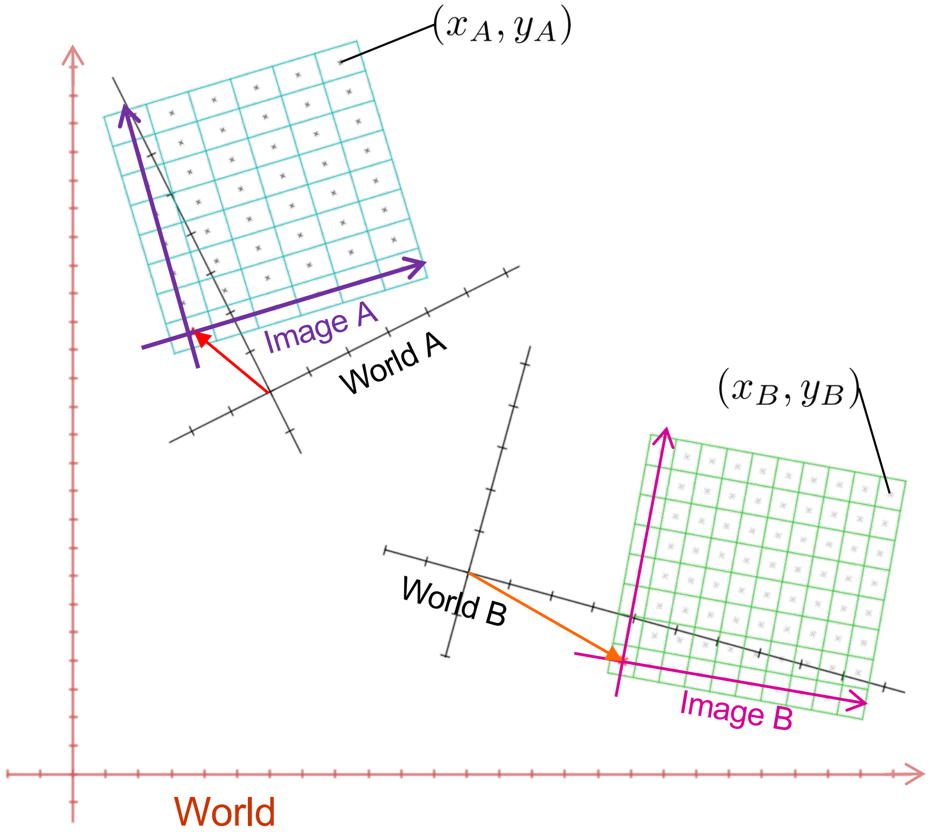
Image B to image A

$$\begin{pmatrix} x_A \\ y_A \\ 1 \end{pmatrix} = T^A_{i2w} \cdot T_{B2A} \cdot T^B_{w2i} \begin{pmatrix} x_B \\ y_B \\ 1 \end{pmatrix}$$

For this transformation (global or local) we need image registration!

Initializations:

- Align image centres
- Ignore origin and direction
- Align centres of intensity mass



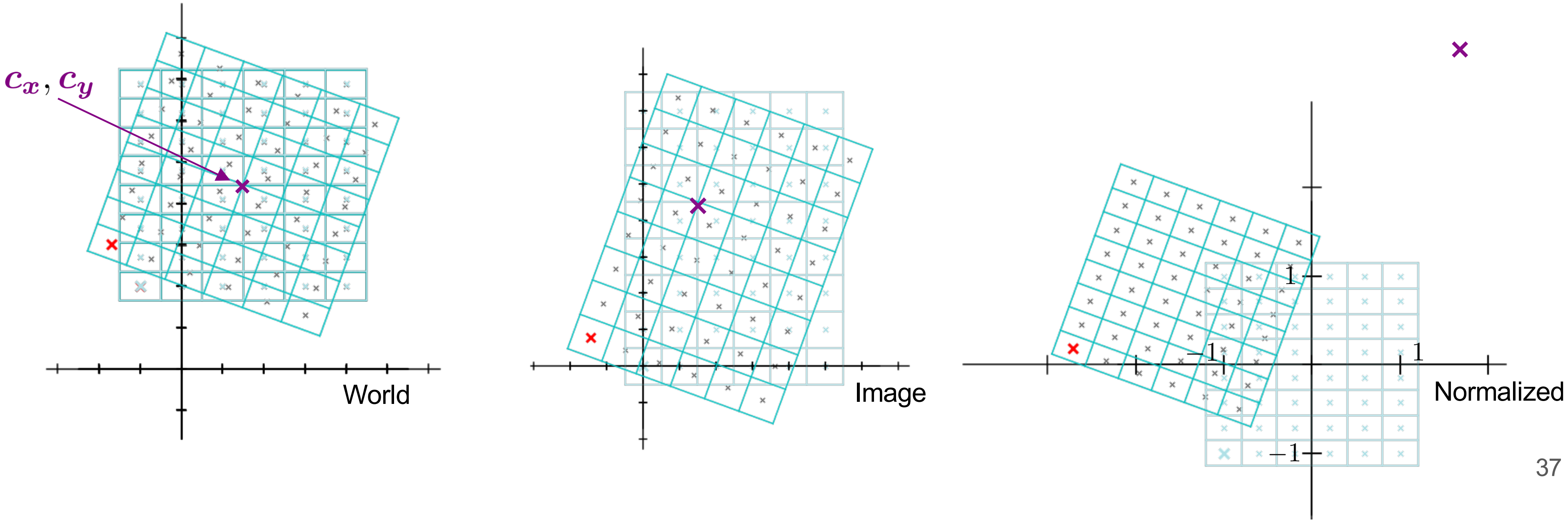
Spatial transformation in different coordinate systems

$$\theta = (\alpha, t_x, t_y)$$

Rigid transformation in world coordinate system

Translations t_x, t_y in mm
Center of rotation c_x, c_y in world coordinates

$$\phi(\theta) = \begin{pmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{pmatrix}$$



Spatial transformation in different coordinate systems

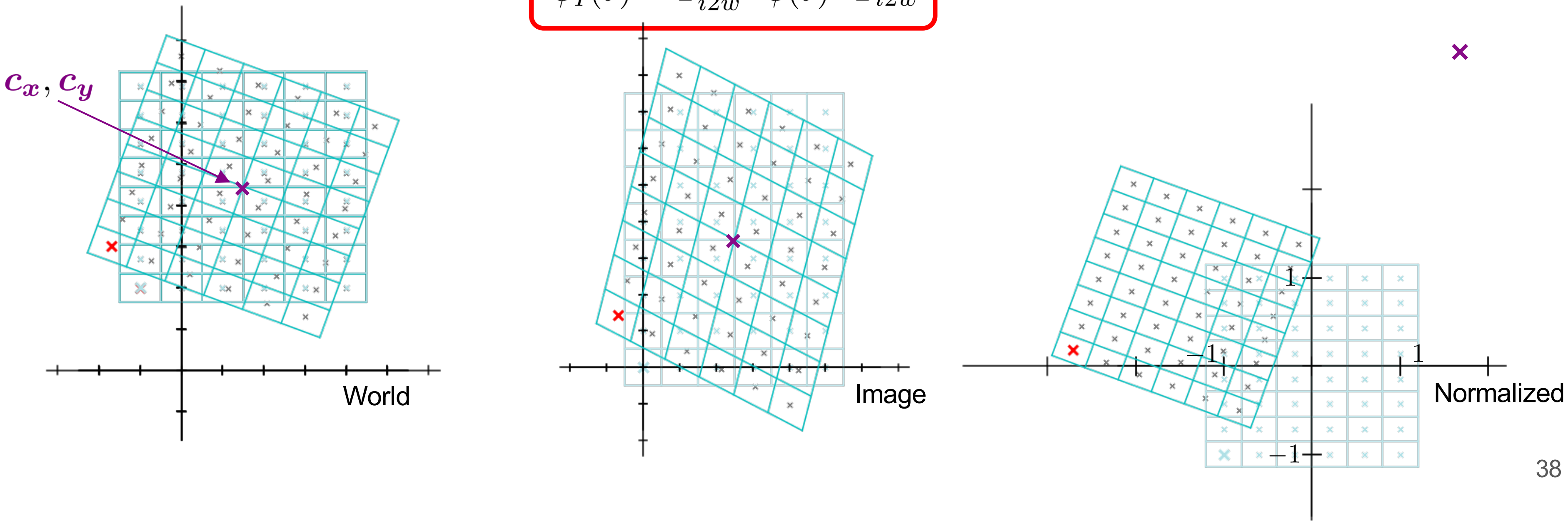
$$\theta = (\alpha, t_x, t_y)$$

Rigid transformation in world coordinate system

Translations t_x, t_y in mm
Center of rotation c_x, c_y in world coordinates

$$\phi(\theta) = \begin{pmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{pmatrix}$$

$$\boxed{\phi_I(\theta) = T_{i2w}^{-1} \cdot \phi(\theta) \cdot T_{i2w}}$$



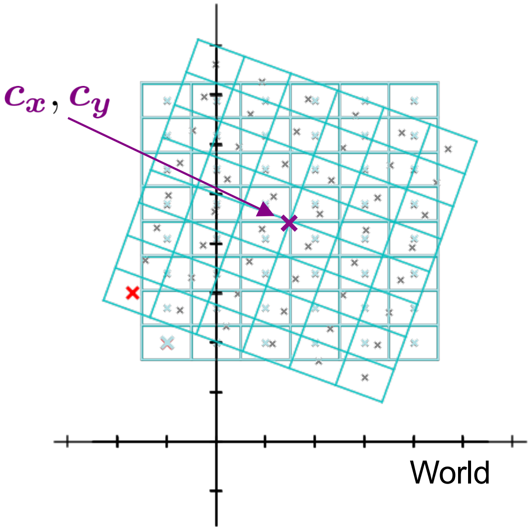
Spatial transformation in different coordinate systems

$$\theta = (\alpha, t_x, t_y)$$

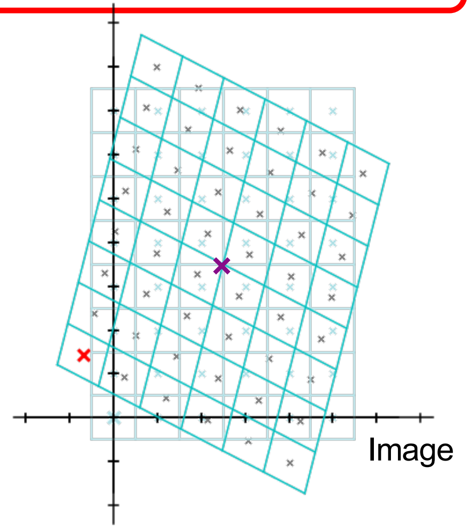
Rigid transformation in world coordinate system

Translations t_x, t_y in mm
Center of rotation c_x, c_y in world coordinates

$$\phi(\theta) = \begin{pmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{pmatrix}$$



$$\phi_I(\theta) = T_{i2w}^{-1} \cdot \phi(\theta) \cdot T_{i2w}$$



$$\phi_N(\theta) = T_{n2w}^{-1} \cdot \phi(\theta) \cdot T_{n2w}$$
$$T_{n2w} = T_{i2w} \cdot T_{i2n}^{-1}$$

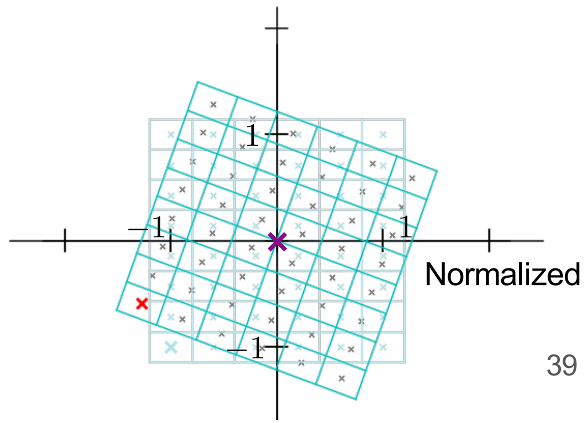


Image Interpolation

- Given a spatial transformation $\phi : \Omega_F \to \Omega_M$ between image M and F.
- The transformed point $\phi(x) = x'$ might fall between pixels and associated intensities.
- How do we find the new intensities of the pixel grid locations?

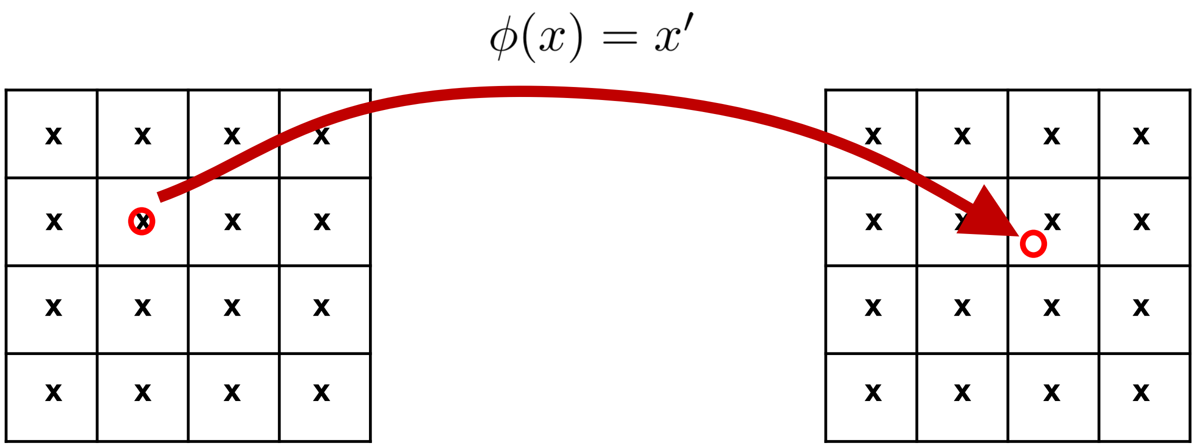
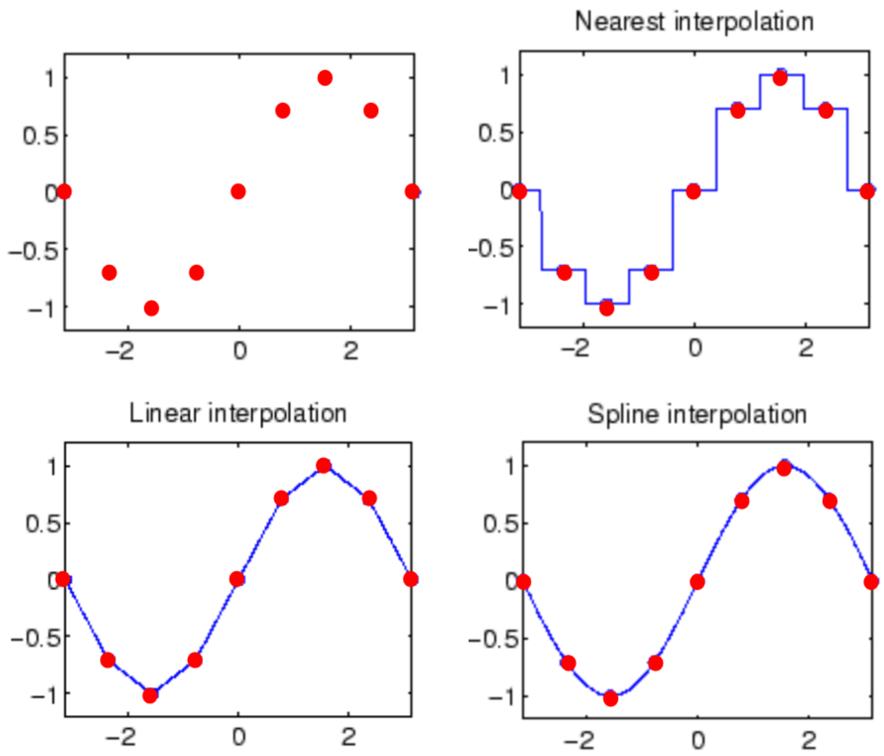
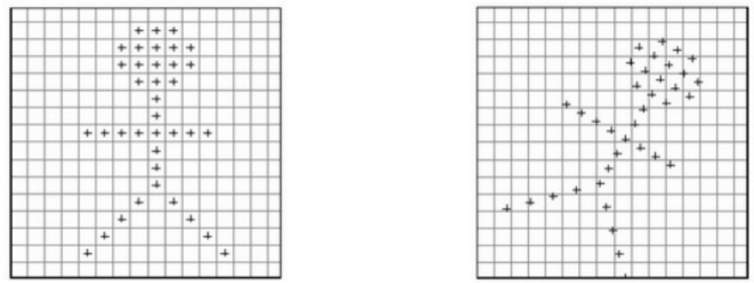


Image Interpolation

Interpolation in 1D



Interpolation in 2D

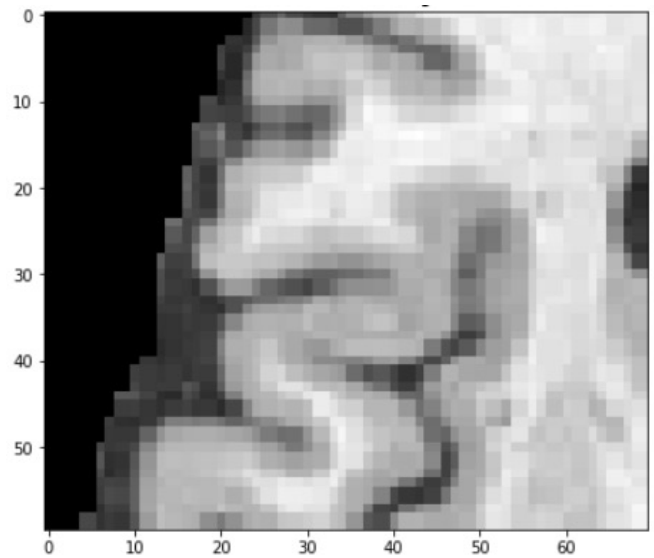


Source: https://climserv.ipsl.polytechnique.fr/documentation/idl_help/Interpolation_Methods.html

Image Interpolation

- Every resampling introduces errors.
  Avoid resampling many times!

Nearest Neighbors



Linear

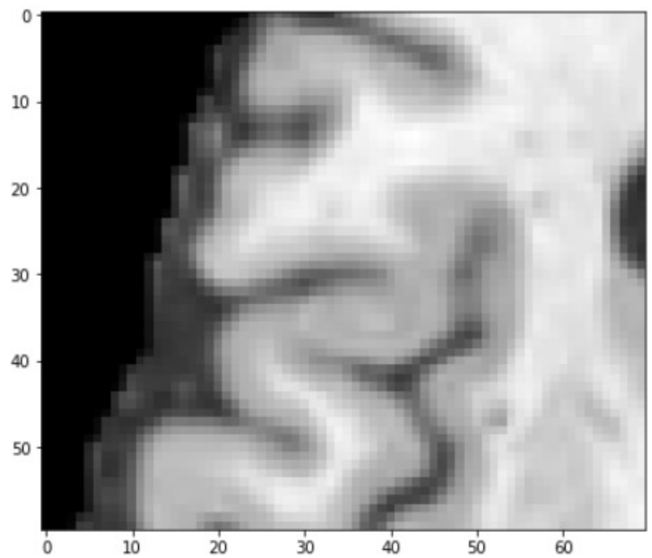
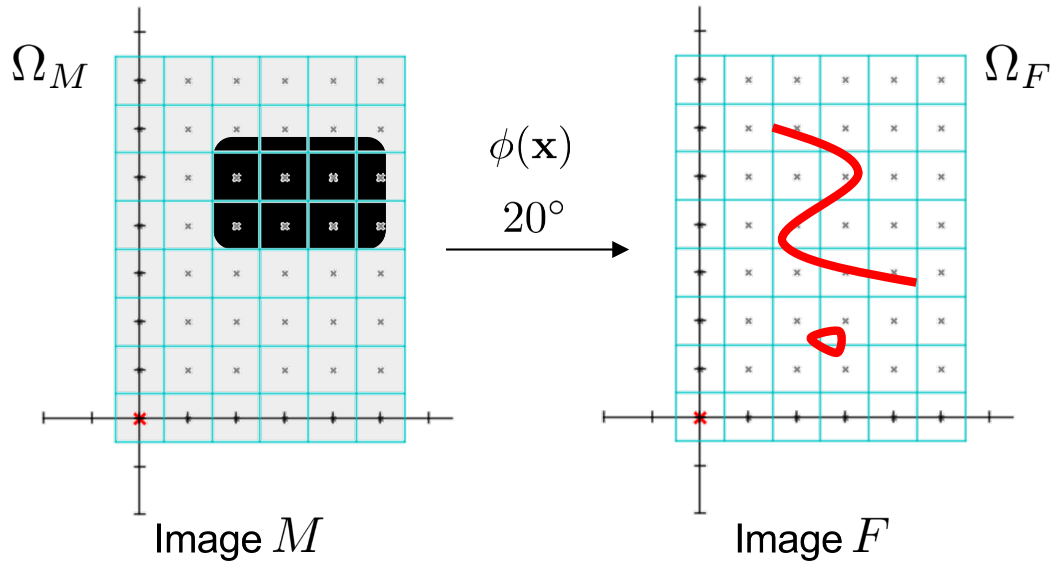


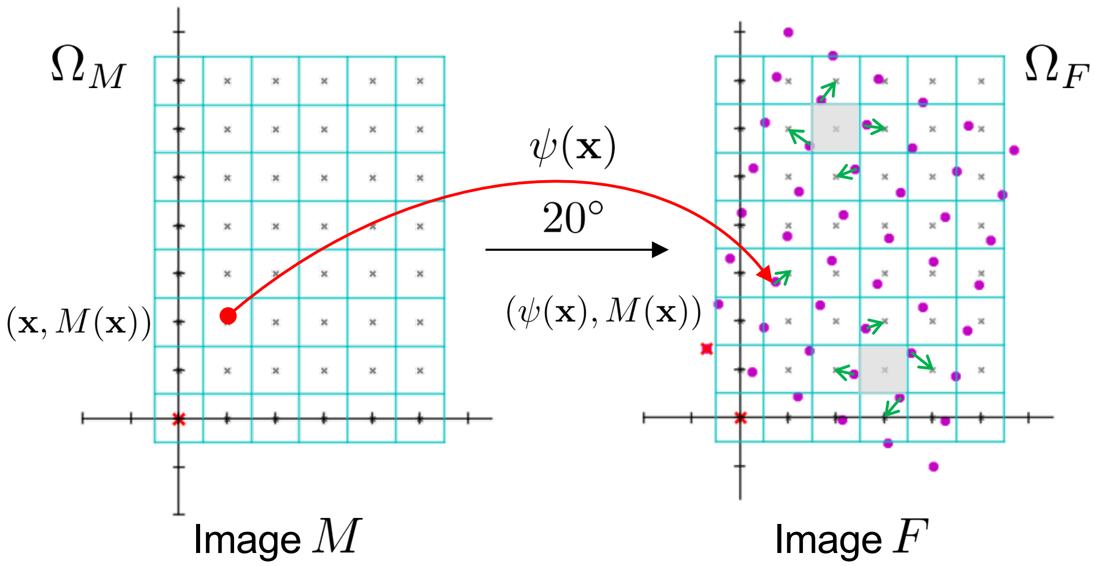
Image warping

- How to apply a transformation to an image?
- Two approaches: **Forward and backward mapping**



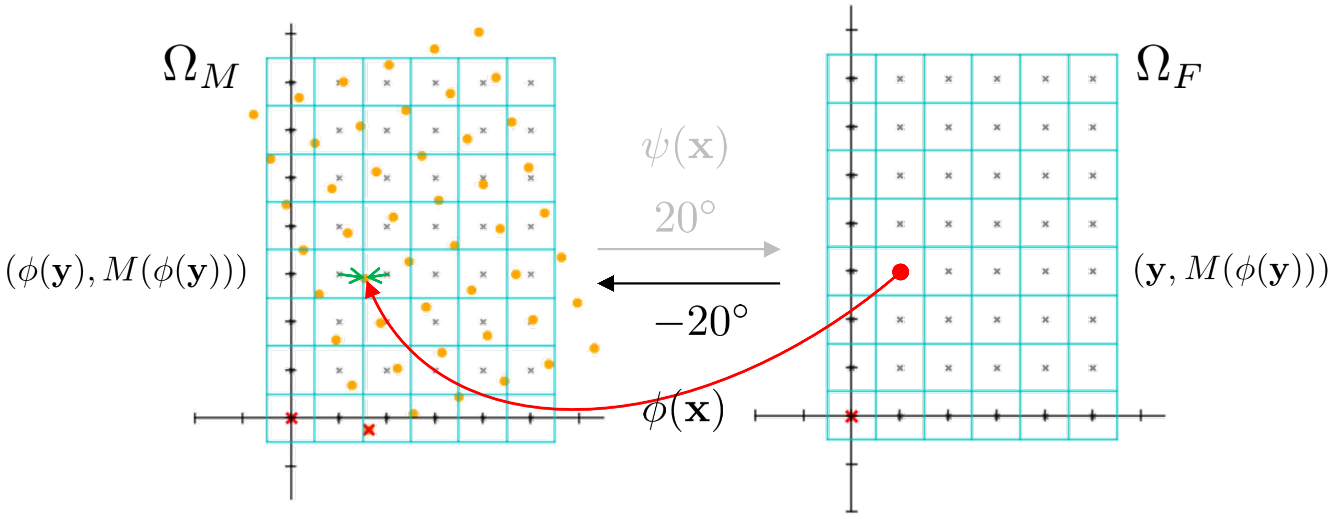
Forward mapping (Lagrange approach)

- Consider a grid point $\mathbf{x} \in \Omega_M$ in image M and a spatial transformation $\psi : \Omega_M \to \Omega_F$.
- The intensity value $M(\mathbf{x})$ at position $\mathbf{x}$ is moved to the new position $\psi(\mathbf{x})$.
- Interpolation to regular grid points.
- Forward mapping might result in holes in the warped image!



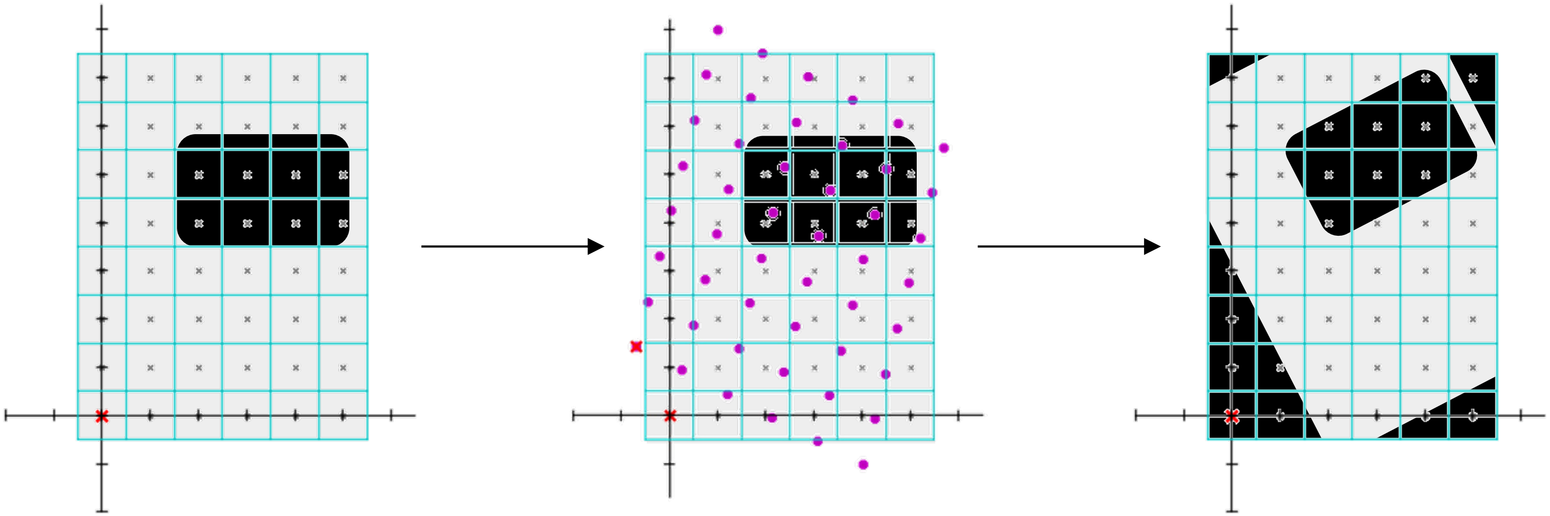
Backward mapping (Euler approach)

- Consider a grid point $\mathbf{y} \in \Omega_F$ in image $M \circ \phi$ and a spatial transformation $\phi : \Omega_F \to \Omega_M$.
- The intensity value at position $\mathbf{y}$ is set to the intensity value $M(\phi(\mathbf{y}))$ from position $\phi(\mathbf{y})$.
- Interpolation to get intensity value at non-grid points.
- The inverse transformation need to exist for backward mapping! $\phi = \psi^{-1}$



Backward mapping (Euler approach)

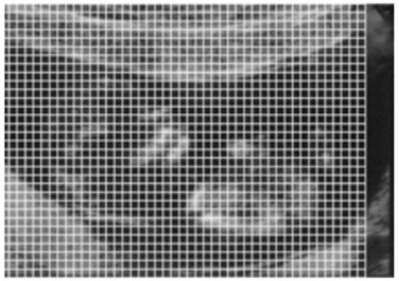
A rotation of 20° (of the domain Ω_M of the moving image) results in a -20° rotation of the moving image after resampling (interpolation to grid points)!

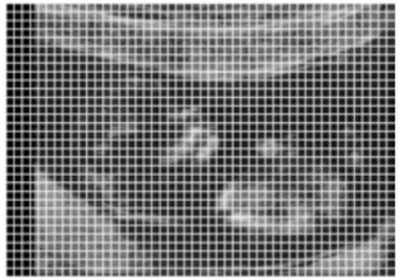


Transformation model

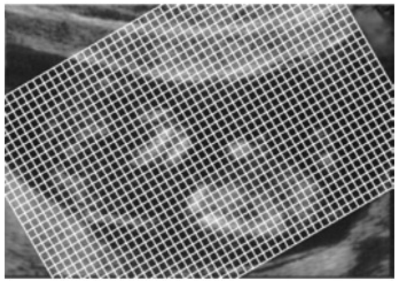
Image *M*,
to be transformed

Rigid





Affine



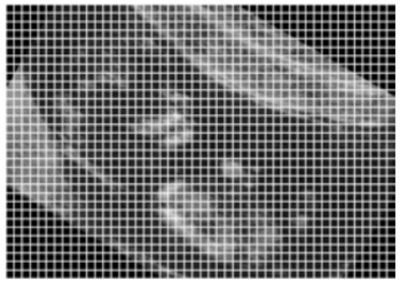


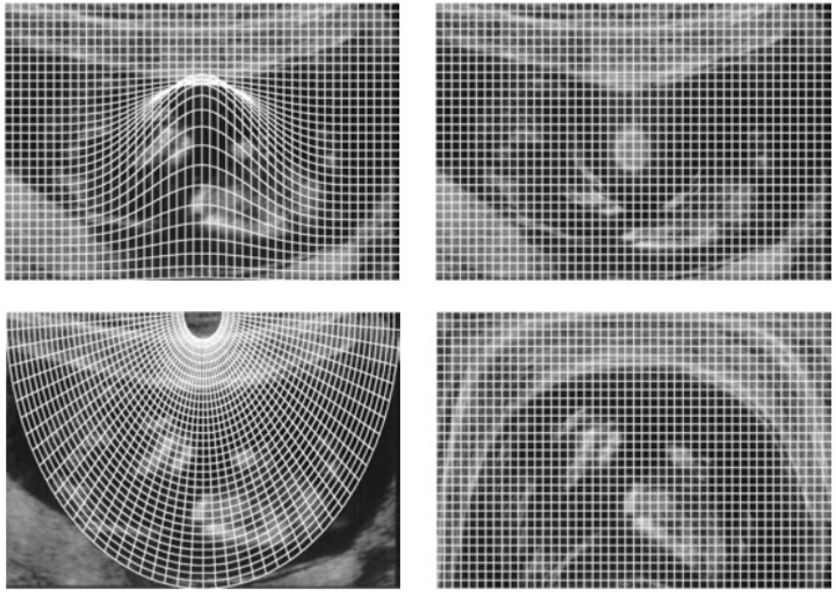
Image *M*, transformation ϕ

$M \circ \phi$

Source: Modersitzki, SIAM, 2009.

Transformation model

Image *M*,
to be transformed



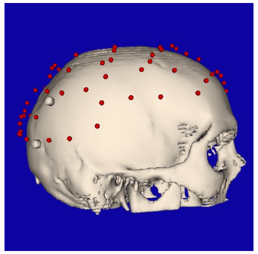
Source: Modersitzki, SIAM, 2009.

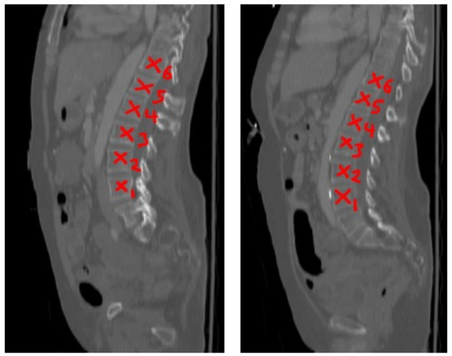
Objective function and Optimization

Types of image registration

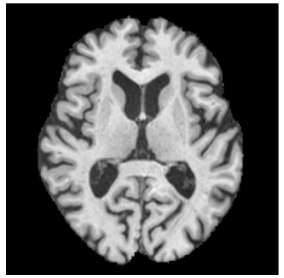
Two main approaches

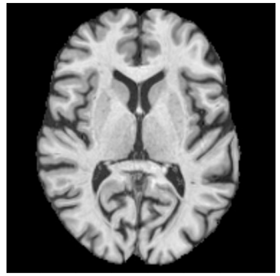
- **Feature-based matching**
  - Identify corresponding structures
  - Fit spatial transformation
- **Intensity-based matching** **In this lecture!**
  - Define notion of similarity between images
  - Find a spatial transformation which maximizes the similarity





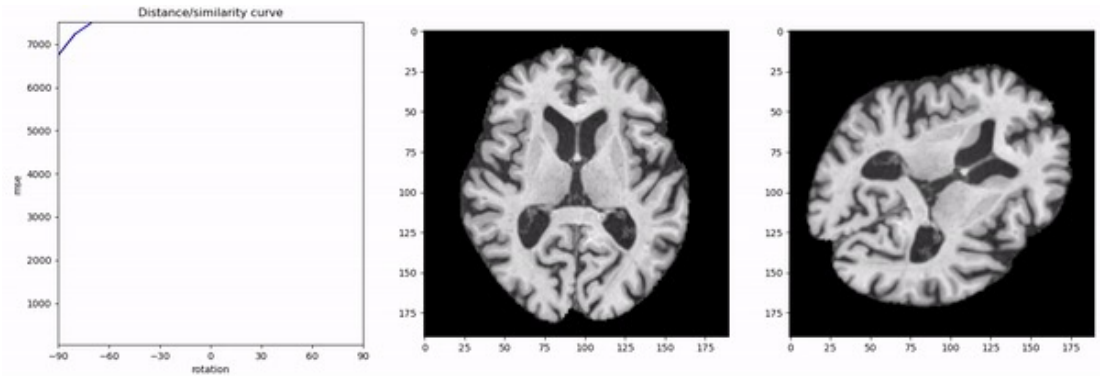
Mono-modal image similarity measures



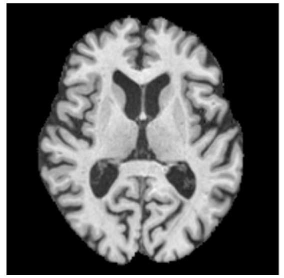


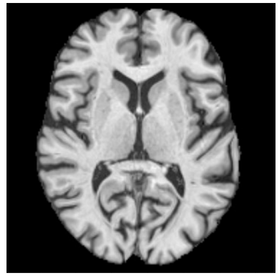
- Sum-of-square-differences (SSD)
  Mean-squares-error (MSE)
- Assumes identity relationship between image intensities in both images
- Optimal measure if the difference between both images is Gaussian noise
- Sensitive to outliers

$$\mathcal{D}_{\mathrm{SSD}}(F, M \circ \phi) = \frac{1}{N} \sum_{i=1}^{N} (F(\mathbf{x}_i) - (M \circ \phi)(\mathbf{x}_i))^2$$



Mono-modal image similarity measures



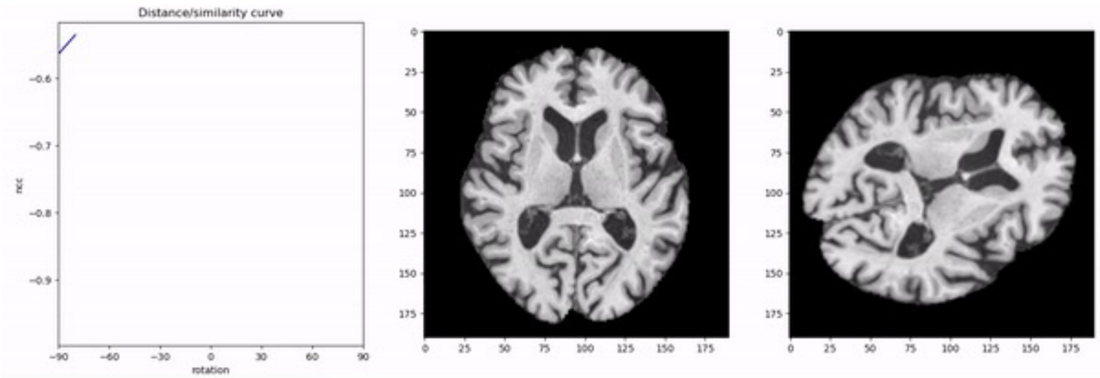


- Normalized Cross Correlation (NCC)

- Assumes linear relationship between image intensities in both images
- Useful if images have been acquired with different intensity windowing

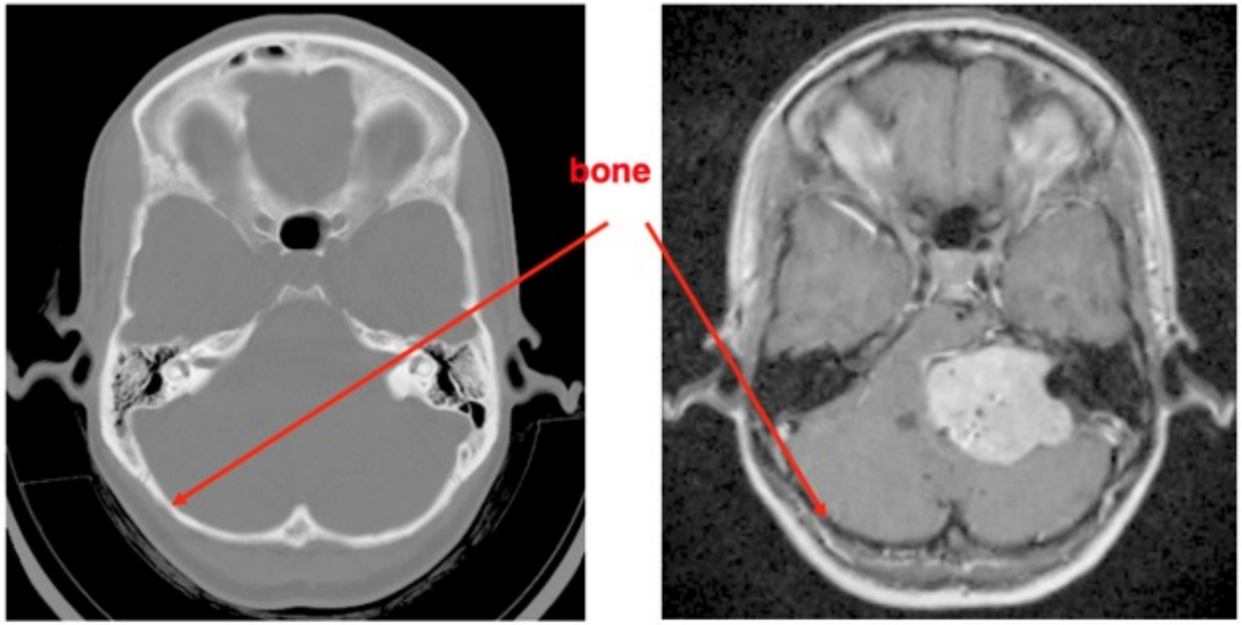
$$\mathcal{D}_{\mathrm{NCC}}(F, M \circ \phi) = \frac{\sum_i (F(\mathbf{x}_i) - \mu_F)((M \circ \phi)(\mathbf{x}_i) - \mu_M)}{\sqrt{\sum_i (F(\mathbf{x}_i) - \mu_F)^2 \sum_i ((M \circ \phi)(\mathbf{x}_i) - \mu_M)^2}}$$

μ_F, μ_M : Average intensities in images *F* and *M*, respectively



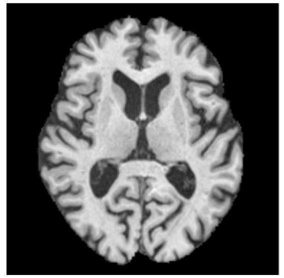
Multi-modal image registration

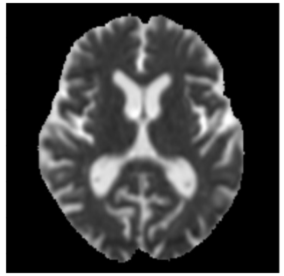
- No functional mapping possible



CT MR

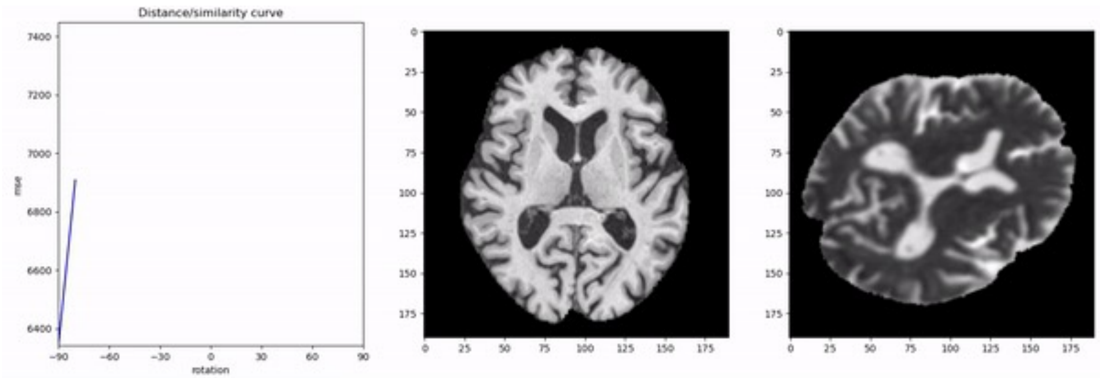
Multi-modal image similarity measures





- Sum-of-square-differences (SSD)
  Mean-squares-error (MSE)
- Assumes identity relationship between image intensities in both images
- Optimal measure if the difference between both images is Gaussian noise
- Sensitive to outliers

$$\mathcal{D}_{\mathrm{SSD}}(F, M \circ \phi) = \frac{1}{N}\sum_{i=1}^{N}(F(\mathbf{x}_i) - (M \circ \phi)(\mathbf{x}_i))^2$$



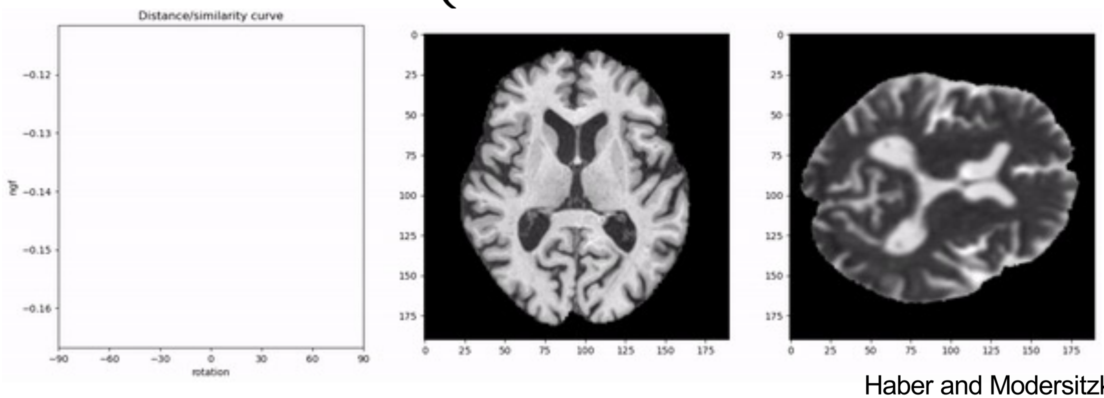
Multi-modal image similarity measures

- Normalized Gradient Fields (NGF)
- Assumes intensity changes at the same locations

$$\mathcal{D}_{\mathrm{NGF}}(F, M \circ \phi) = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{n}(F, \mathbf{x}_i) \times \mathbf{n}(M \circ \phi, \mathbf{x}_i))^2$$

With normalized gradients

$$\mathbf{n}(I, \mathbf{x}) = \begin{cases} \frac{\nabla I(\mathbf{x})}{\|\nabla I(\mathbf{x})\|}, & \nabla I(\mathbf{x}) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

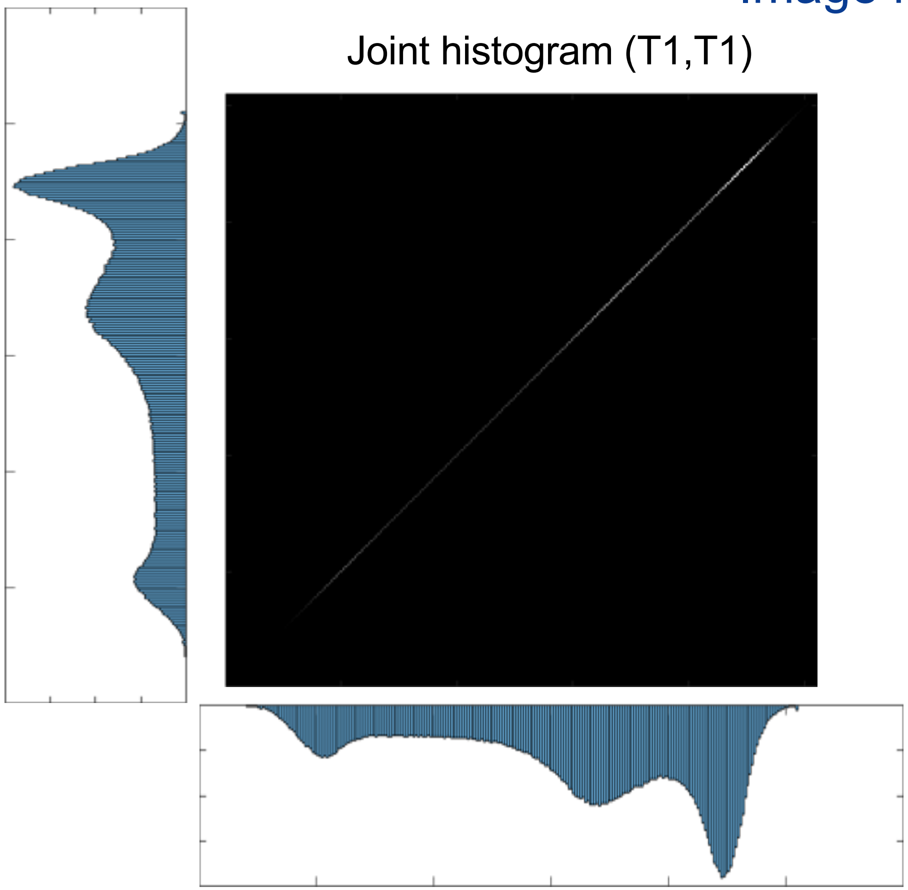


Haber and Modersitzki. *Methods of information in medicine, 2007*

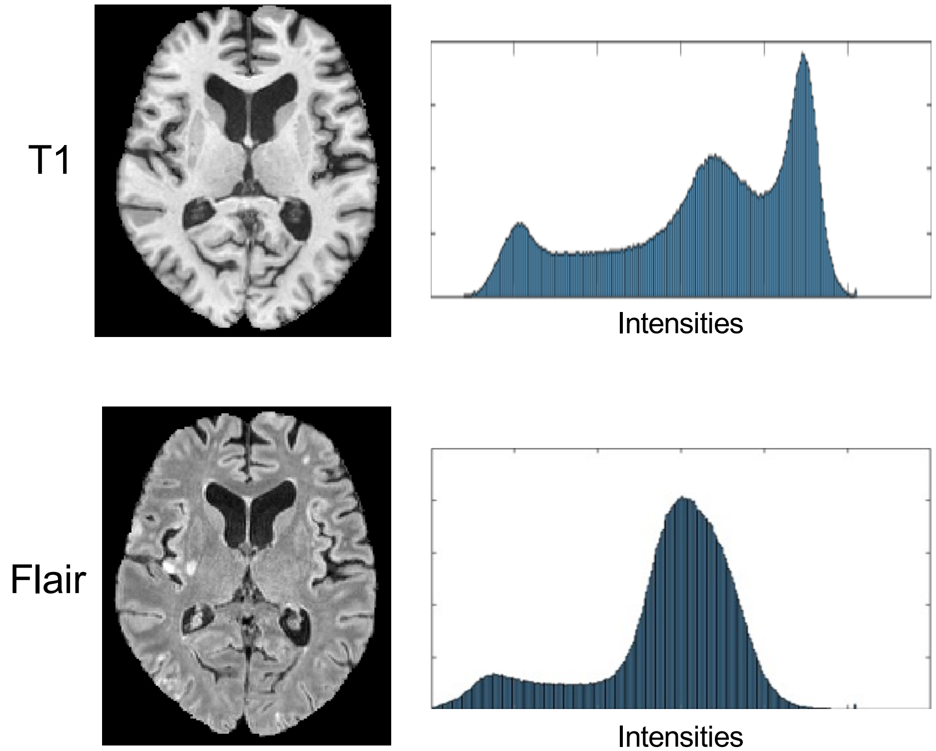
Multi-modal image similarity measures

Image histograms

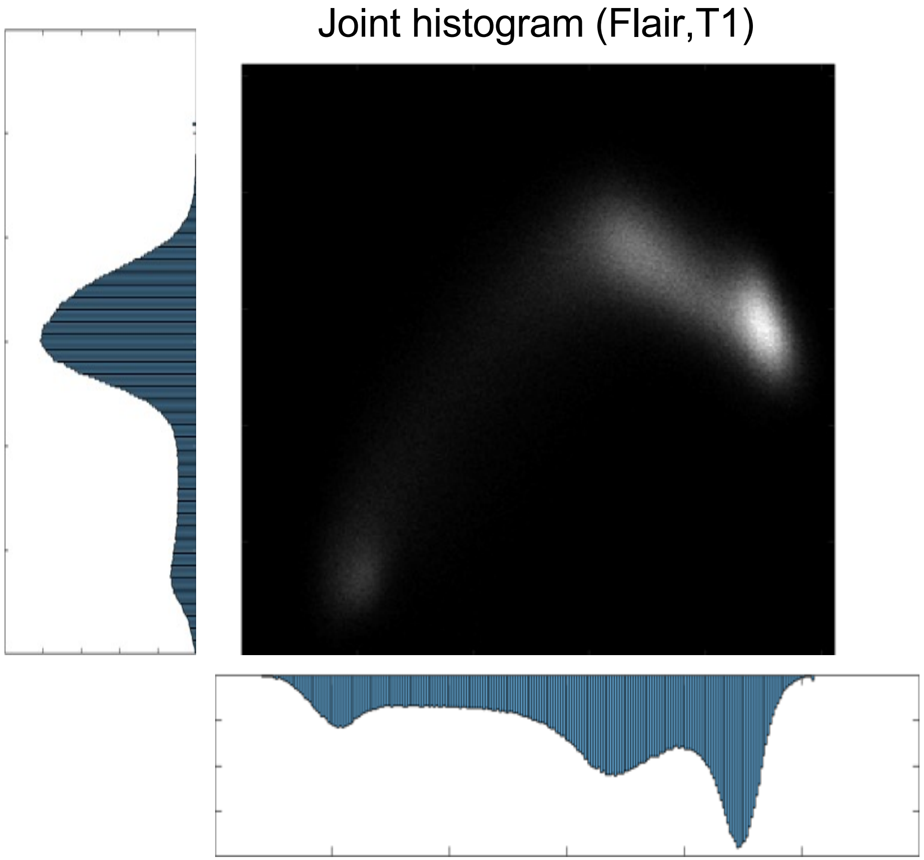
Joint histogram (T1,T1)

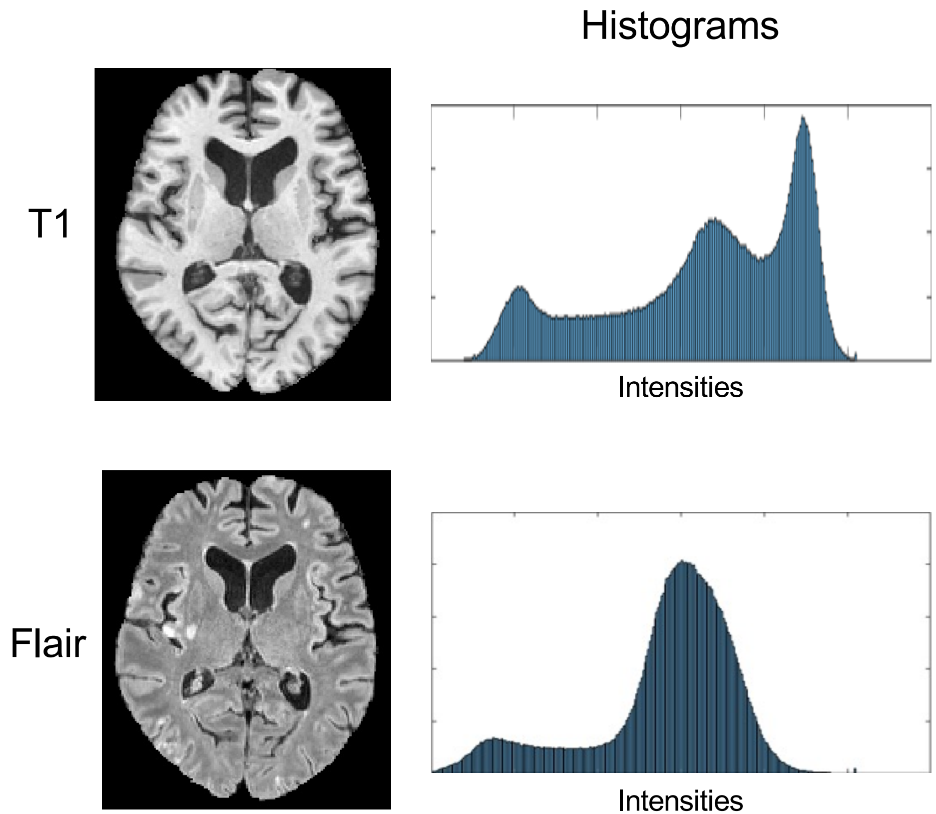


Histograms



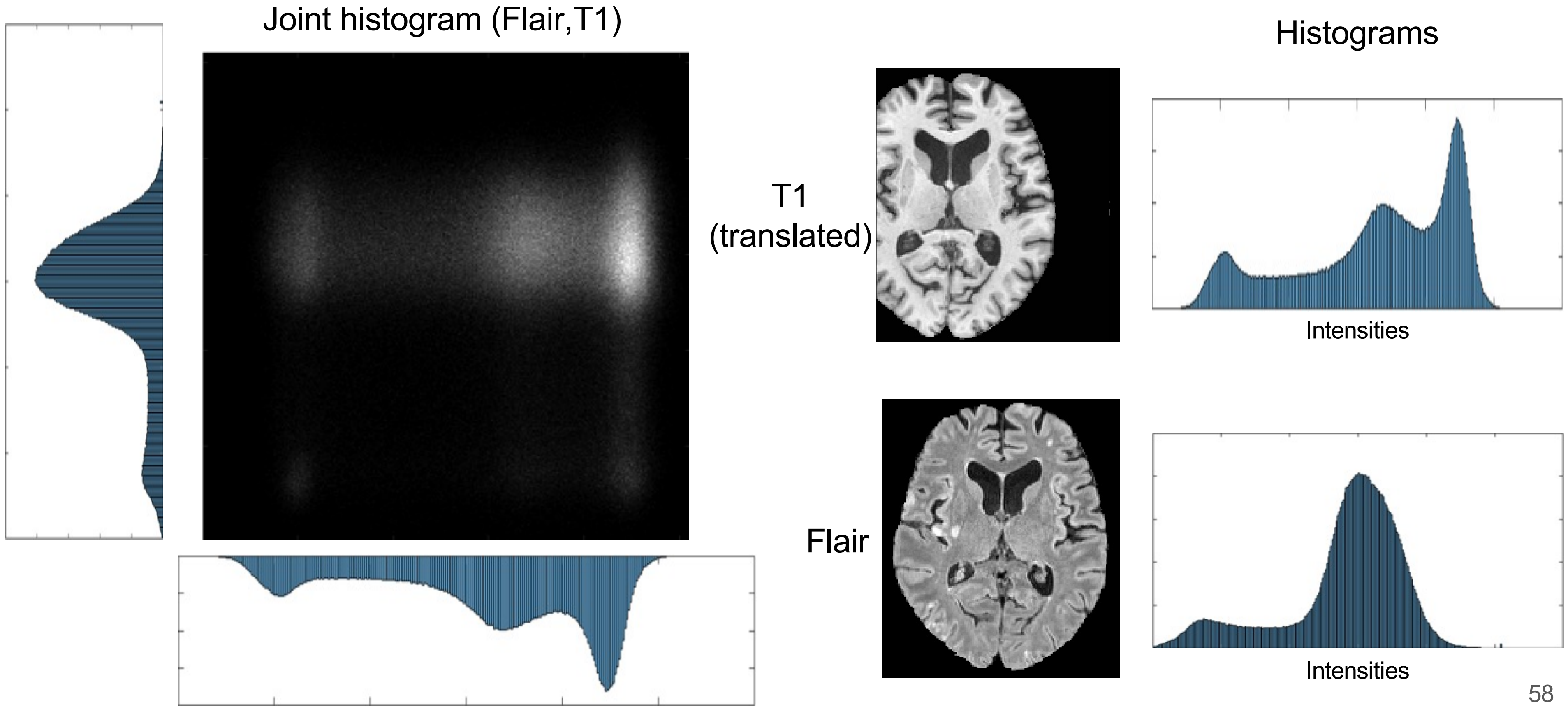
Multi-modal image similarity measures
Image histograms





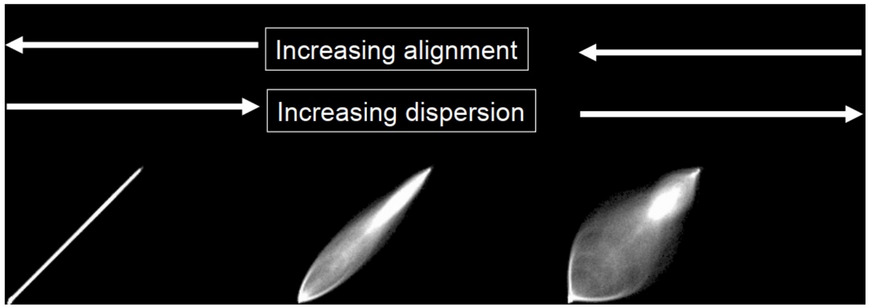
Multi-modal image similarity measures

Image histograms



Multi-modal image similarity measures
Joint image histograms

Mono-modal

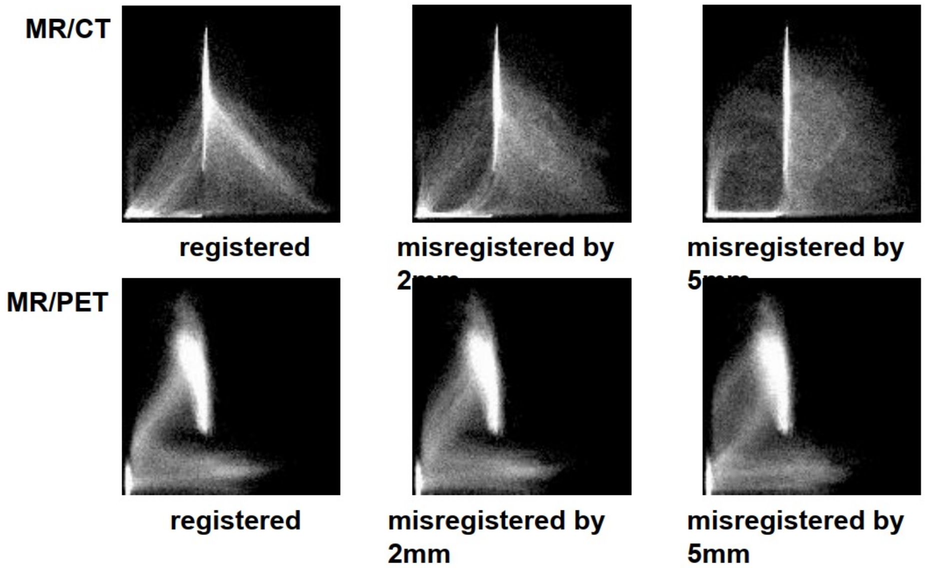


MR/MR registered

misregistered by 2mm

misregistered by 5mm

Multi-modal



Source:
Studholme et al. *Medical Physics, 1997.*
Studholme et al. *Medical Image Analysis, 1996.*

Multi-modal image similarity measures

Images as probability density functions

$$p(a,b) = \frac{h(a,b)}{N}$$

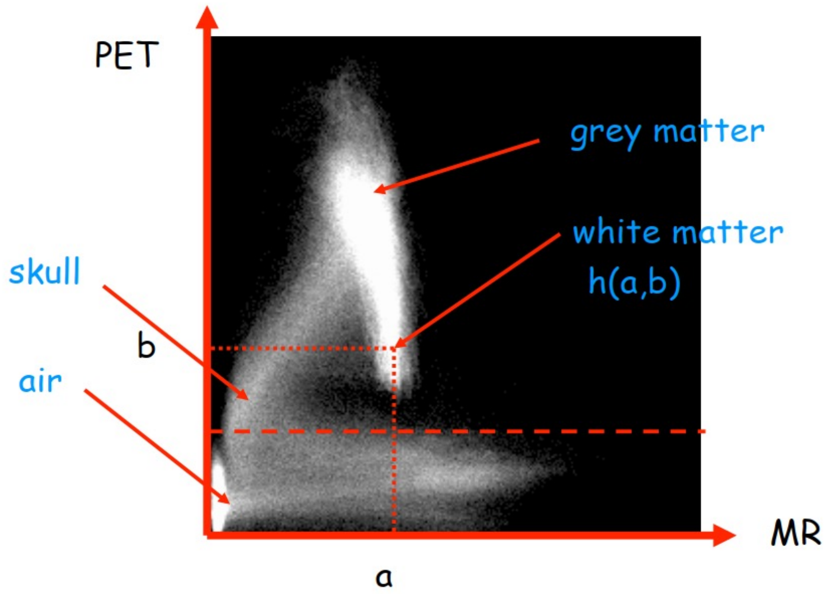
Joint probability of a voxel having intensity *a* in the first and intensity *b* in the second image

$$p(a) = \sum_b p(a,b)$$

Marginal probability of a voxel in the first image having intensity *a*

$$p(b) = \sum_a p(a,b)$$

Marginal probability of a voxel in the second image having intensity *b*



$$h(a,b)$$

Multi-modal image similarity measures

Images as probability density functions

Shannon Entropy

$$H(A) = -\sum_{a} p(a) \log p(a)$$

- Describes the amount of information in image A
- Entropy: how much disorder is in the system?
- The image content is
  - maximal (information-theoretic sense) if all intensities have equal probabilities
  - minimal (information-theoretic sense) if one intensity *a* has a probability of 1, i.e. $p(a) = 1$

Joint Entropy

$$H(A, B) = -\sum_{a} \sum_{b} p(a, b) \log p(a, b)$$

- Describes the amount of information in combined images A,B

Multi-modal image similarity measures

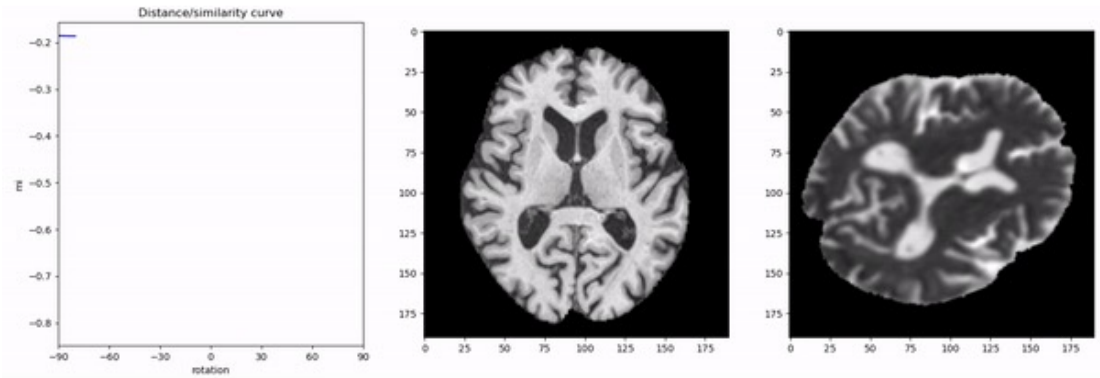
- Mutual Information (MI)

$$\mathcal{D}_{\mathrm{MI}}(F, M \circ \phi) = H(F) + H(M \circ \phi) - H(F, M \circ \phi)$$

- Normalized Mutual Information (NMI)

$$\mathcal{D}_{\mathrm{NMI}}(F, M \circ \phi) = \frac{H(F) + H(M \circ \phi)}{H(F, M \circ \phi)}$$

- Describes how well one image is explained by the other



Viola and Wells. *International journal of computer vision*, 1997.

Studholme, et al. *Pattern recognition, 1999.*

Regularization term

- Objective function $\mathcal{J}(\phi, F, M) = \mathcal{D}(F, M \circ \phi) + \boxed{\alpha\mathcal{R}(\phi)}$
- Well-posed problem:
  - a solution exists
  - the solution is unique
  - the solution's behaviour changes continuously with the initial conditions
- For non-parametric image registration problem is ill-posed
  - without further constraints, there are infinitely many solutions
  - the solutions might not be continuous, diffeomorphic, …
- Regularization term to convert problem into a well-posed problem

- In parametric registration: the search space is restricted by the transformation model
- In non-parametric registration: the problem itself is constraint

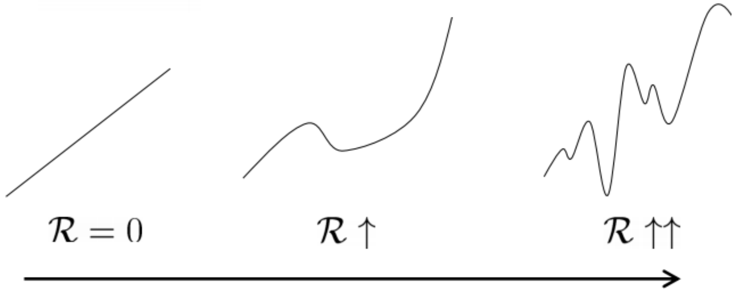
Regularization term

- Which regularization term to use?
  What kind of deformations are plausible and realistic for the application? Diffeomorphisms?
- **Smoothness**
  Diffusion regularization (1st derivatives)
  Modersitzki, 2004.

$$\mathcal{R}(\phi) = \frac{1}{2} \int_{\Omega_F} \sum_{j,k}^{d} \left(\frac{\partial u_k}{\partial x_j} \right)^2$$

  Bending energy (2nd derivatives)
  Rueckert, et al., 1999.

$$\mathcal{R}(\phi) = \frac{1}{2} \int_{\Omega_F} \sum_{i,j,k}^{d} \left(\frac{\partial^2 u_k}{\partial x_j x_i} \right)^2$$



Staring, Tutorial@MICCAI 2010.

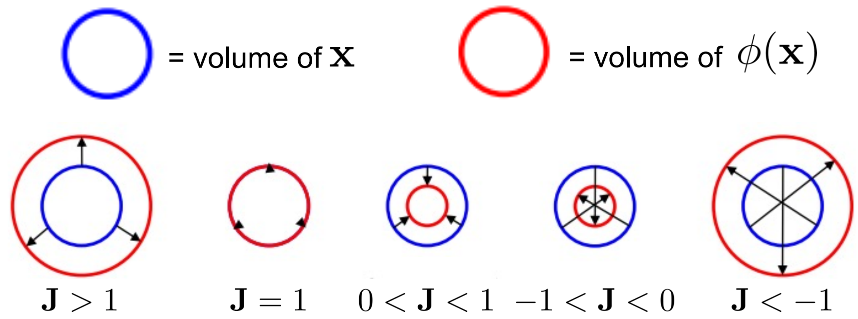
Regularization term

- Which regularization term to use?
  What kind of deformations are plausible and realistic for the application? Diffeomorphisms?
- Smoothness
  Diffusion regularization (1st derivatives)
  Fischer and Modersitzki, 2004.

  Bending energy (2nd derivatives)
  Rueckert, et al., 1999.
- **Jacobian**

$$\nabla\phi(x) = \left(\frac{\partial u_i}{\partial x_j}(\mathbf{x})\right)_{i,j=1}^{d,d}$$

$$\mathbf{J}\phi(\mathbf{x}) = \det \nabla\phi(\mathbf{x})$$



Staring, Tutorial@MICCAI 2010.

- Other
  Elastic regularization
  Curvature regularization
  Fischer and Modersitzki, 2004.

$$\mathcal{R}(\phi) = \sum_{\Omega_F} \log \mathbf{J}\phi(\mathbf{x})$$

(Incrompressibility)

Rohlfing, et al. 2003

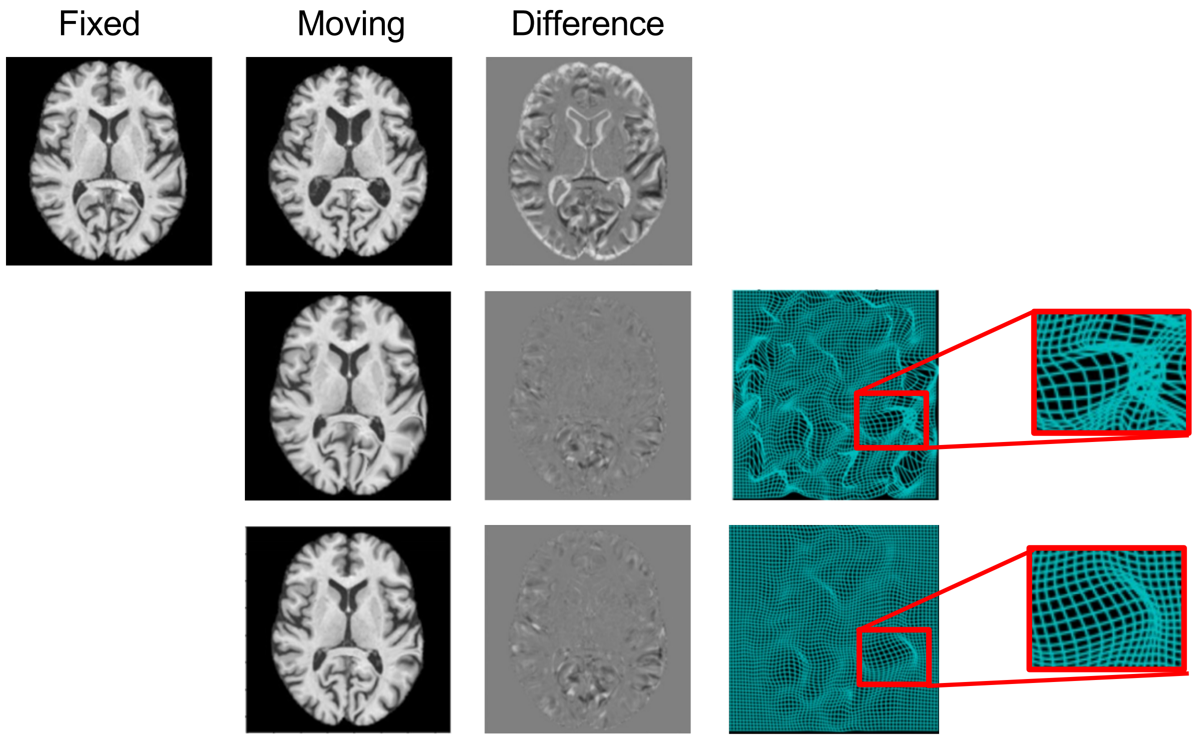
Regularization term: Example

Objective function

$$\mathcal{J}(\phi, F, M) = \mathcal{D}(F, M \circ \phi) + \alpha\mathcal{R}(\phi)$$

No regularization $(\alpha = 0)$

Regularization $(\alpha > 0)$



Optimization strategies

- Energy minimization

$$\mathcal{J}(\phi, F, M) = \mathcal{D}(F, M \circ \phi) + \alpha \mathcal{R}(\phi)$$

$$\phi^* = \underset{\phi}{\operatorname{argmin}}\, \mathcal{J}(\phi, F, M)$$

- Discrete optimization: search space is discrete domain (e.g., finite or integers)
- Continuous optimization: search space is continuous (e.g., real numbers)

- Discretization of images, transformations, objective function
- Discrete SSD: $\mathcal{SSD}(F, M, \phi) \approx \frac{1}{2} \sum_{x \in \Omega_F} (F(\mathbf{x}) - M(\phi(\mathbf{x})))^2$
- Derivatives: $\nabla_\phi \mathcal{D}, \nabla_\phi \mathcal{R}$ $\quad \nabla_\phi \mathcal{SSD}(F, M, \phi) = \sum_{x \in \Omega_F} (F(\mathbf{x}) - M(\phi(\mathbf{x}))) \nabla_\phi M(\phi(\mathbf{x}))$

  $\nabla_\phi M(\phi(\mathbf{x}))$

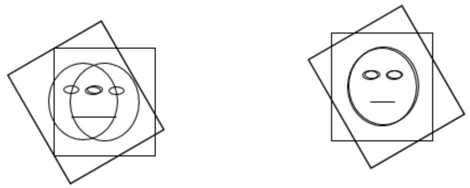
Optimization strategies

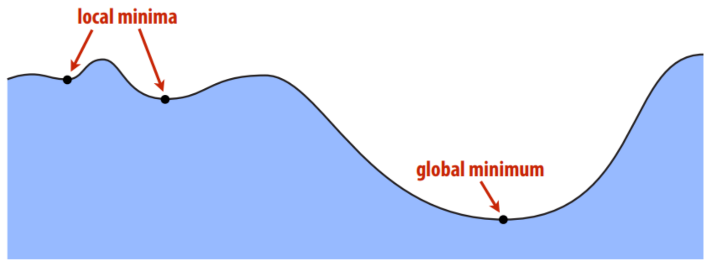
- **Global minimizer**:
  $\boldsymbol{x}^* \in \mathbb{R}^n$ is a global minimizer of function f, if

$$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}), \forall \boldsymbol{x}$$

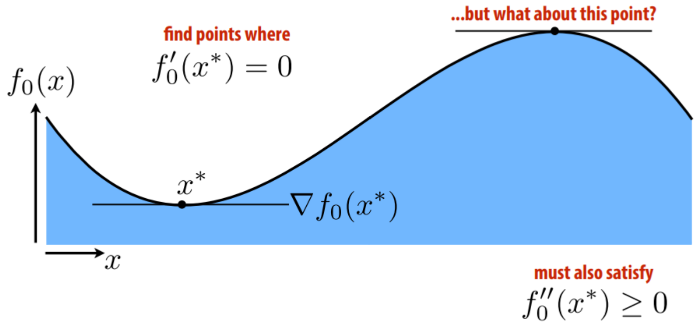
- **Local minimizer**:
  $\boldsymbol{x}^* \in \mathbb{R}^n$ is a local minimizer of function f, if there is a neighbourhood $\mathcal{N}$ around $\boldsymbol{x}^*$, such that

$$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathcal{N}$$





- We want the global minimizer!
- Note that we optimize over the transformation space!
  n is number of transformation parameters, or number of pixel in dense displacement field.
- Gradient-based approaches



Line search approaches

- Line search approaches:

  Initialization $\boldsymbol{x}_0 \in \mathbb{R}^n$

  Search direction $\boldsymbol{s}_k \in \mathbb{R}^n$

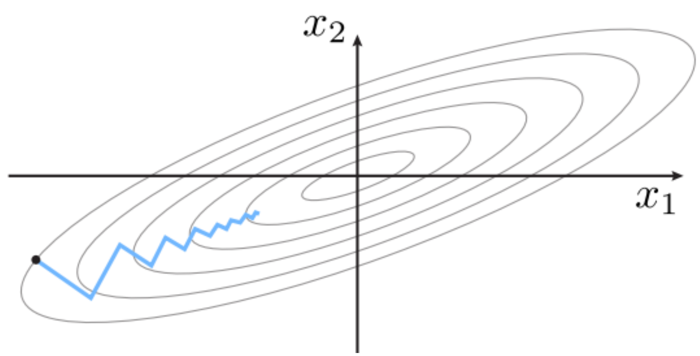
  Step length $\alpha_k \in \mathbb{R}^+$

Update rule:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{s}_k$$

- Gradient descent

$$\boldsymbol{s}_k = -\nabla f(\boldsymbol{x})$$



- Higher order derivatives:

  Newton

  Quasi-Newton

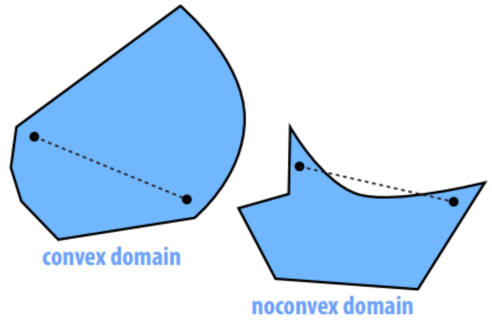
  L-BFGS

- Stochastic gradient descent: replaces the gradient by an estimate calculated from a randomly selected subset of the data Klein, et al. IJCV, 2009.
- Adam is a modified SGD!

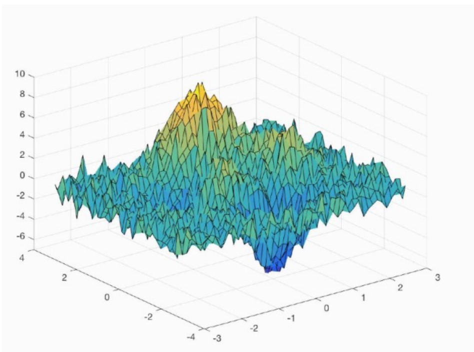
Gradient-based optimization

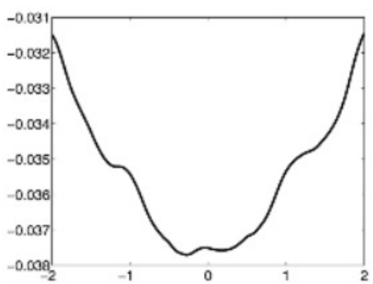
- Gradient-based approaches work well for convex optimization problems
- Convexity:
  $K \subset \mathbb{R}^n, K$ is convex, if

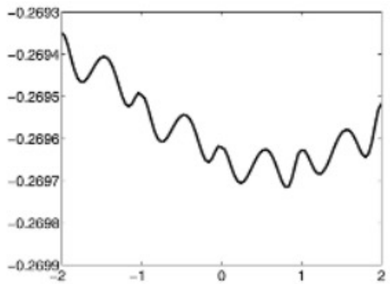
$$(1-\lambda)x + \lambda y \in K \qquad \forall x, y \in K, \lambda \in [0,1]$$

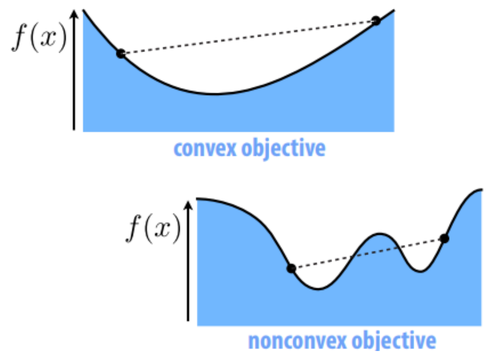


- The registration problem is typically highly non-convex!
- It needs a good initial transformation for convergence.





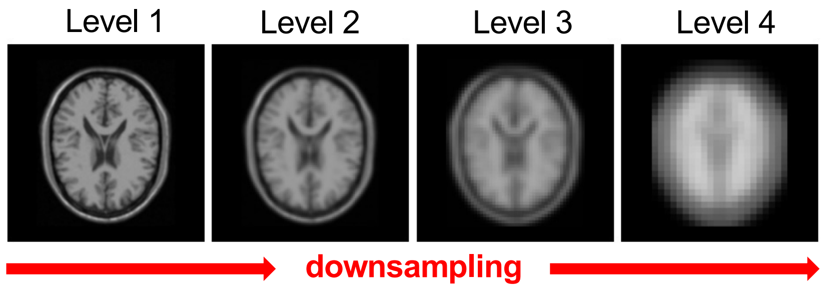


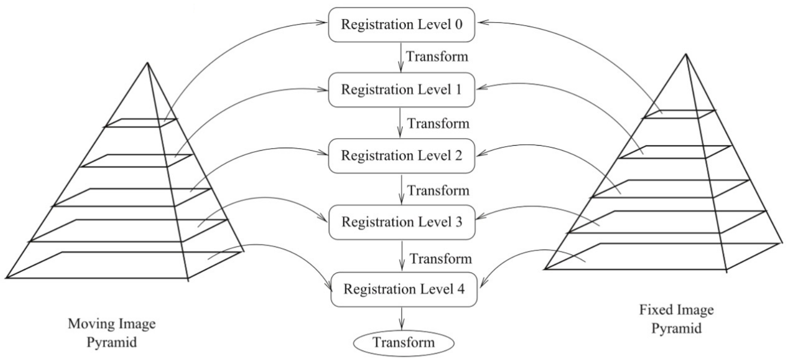


Which function is easier to optimize?

Multi-resolution strategies

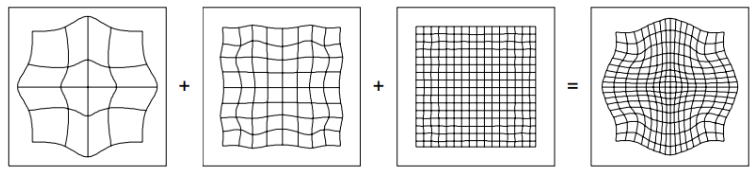
- For efficiency
- Better initialization at subsequent levels
- Avoid local minima
- Increase the speed
- Increase robustness
- Regularisation
- Large deformations





Resolution levels $\ell = 1, \ldots, L$ with coarsest level L.
Final transformation:

$$\phi = \phi_1 \circ \phi_2 \circ \cdots \circ \phi_L$$



Schnabel, et al. MICCAI, 2001.

Generic pairwise image registration algorithm

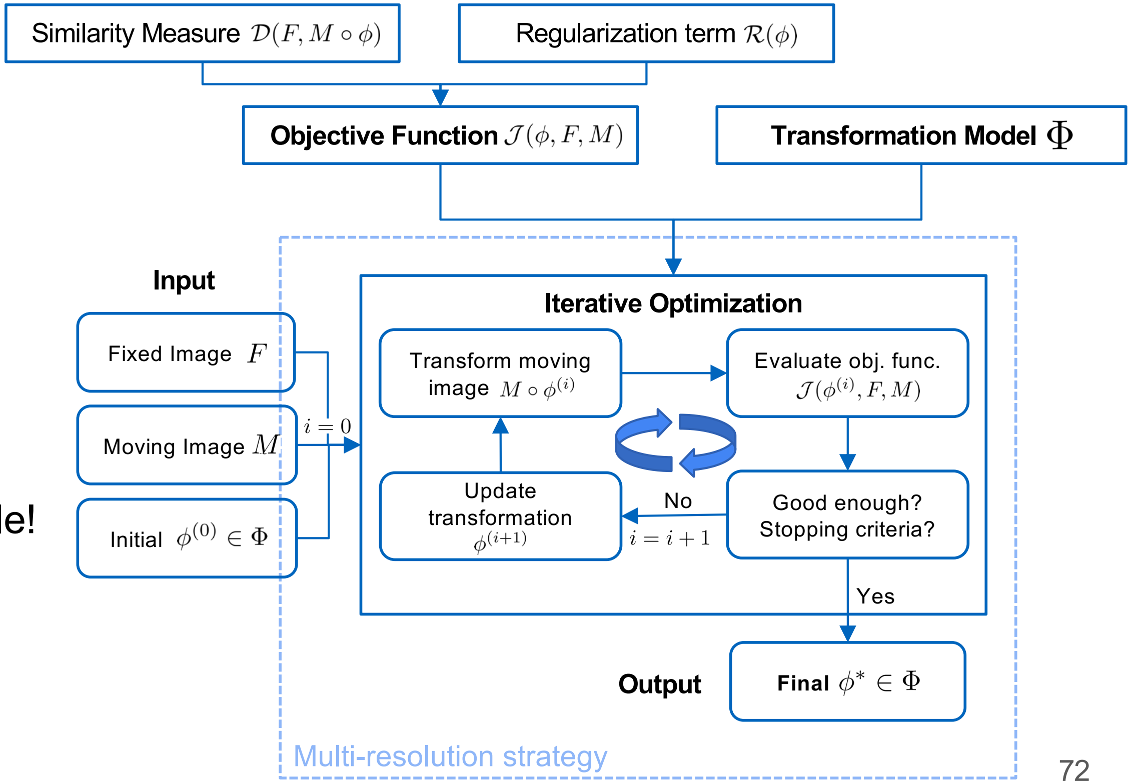
- Optimization problem

$$\phi^* = \underset{\phi}{\operatorname{argmin}}\, \mathcal{J}(\phi, F, M)$$

- Objective function

$$\mathcal{J}(\phi, F, M) = \mathcal{D}(F, M \circ \phi) + \alpha \mathcal{R}(\phi)$$

- Solved iteratively
- Often Gradient-based solvers
  $\longrightarrow$ everything needs to be differentiable!

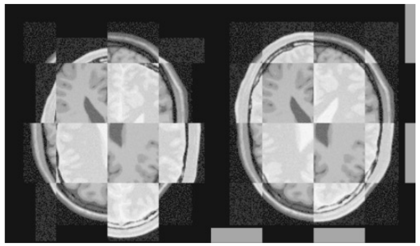


Evaluation

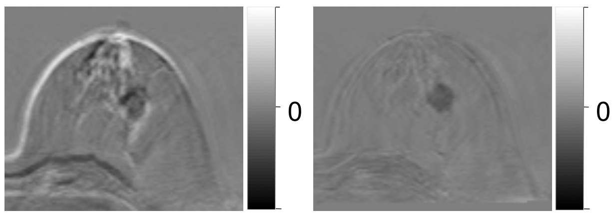
Evaluation of image registration

Two types of evaluation:
Qualitative and **Quantitative** evaluation.

Qualitative



Checkerboard before and after



Difference before and after

Quantitative

What is a correct registration?

How to define a ground truth?

Quantitative evaluation

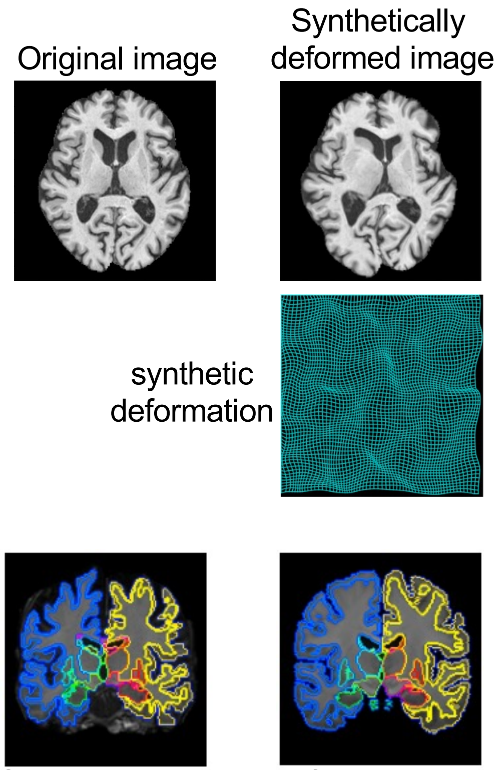
Requirement of **ground truth** or **surrogate measures**

- **Target registration error (TRE):** True error, often unknown
- **Ground truth:** Exact definition of correct transformation

  - Artificial transformations applied to the images
  - Simulations or phantoms

  - hard to define realistically deformed images
  - artificially deformed images are simplified representations of the true registration problem
- **Surrogate measures:** Approximation to correct transformation

  - Segmentation label overlap
  - Landmark error



Segmentation overlap for evaluation

Pluim, et al., *The truth is hard to make: Validation of medical image registration*. IEEE ICPR, 2016.

Quantitative evaluation measures

Corresponding points / anatomical landmarks
(linear registration)

Similarity measures (not used for registration!)

Segmentation label overlap

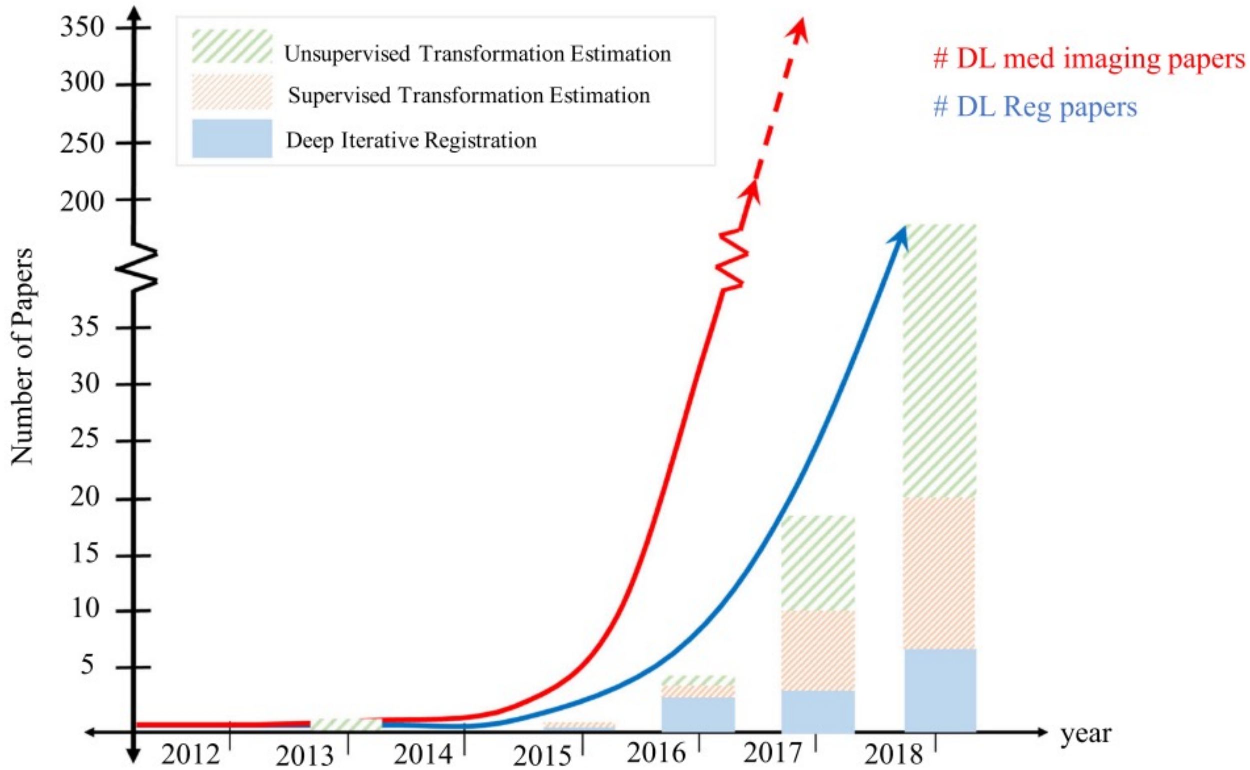
Quality of deformation field (e.g., consistency, smoothness, volume loss, topology, …)

Simulations
- Artificial transformations
- Physical models/ phantoms

Evaluation needs to be independent of registration features or cost function!

Learning-based Image Registration

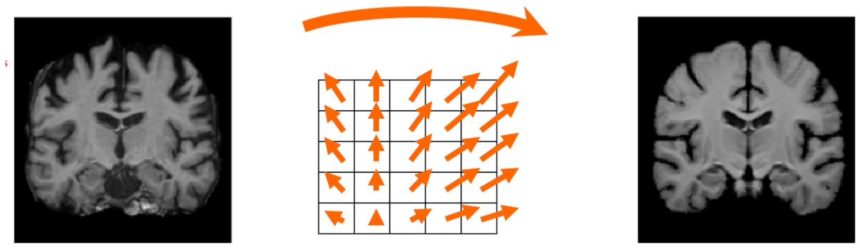
Learning-based approaches



Haskins, et al. Deep learning in medical image registration: a survey. *Machine Vision and Applications, 2020.*

Conventional pairwise registration

- significant development
- slow for two images



$$\phi^* = \underset{\phi}{\operatorname{argmin}}\, \underbrace{\mathcal{D}(F, M \circ \phi)}_{\text{Image matching}} + \alpha \underbrace{\mathcal{R}(\phi)}_{\text{Smooth field}}$$

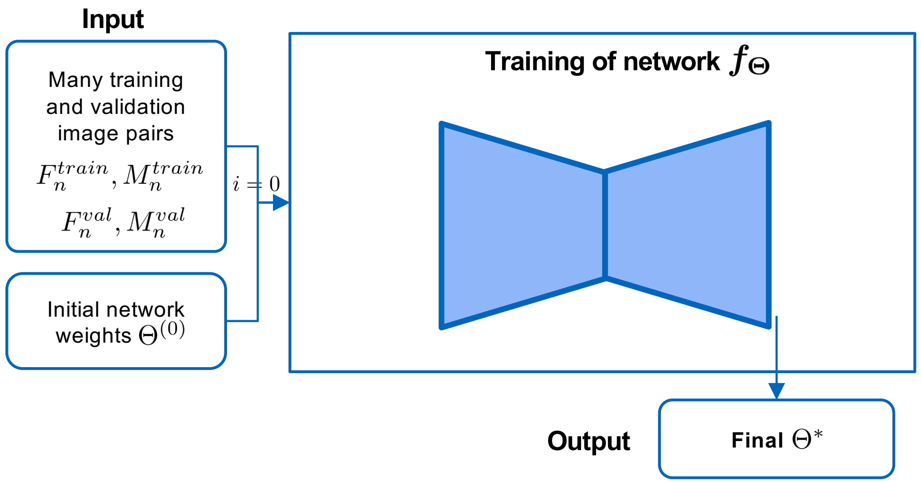
Optimal
deformation field

Learning-based approaches

- Learn the **spatial transformation**

- Exploit information of many image pairs
- Optimization of many training samples instead of over one image pair
- CNN f_Θ parametrised by Θ to encode transformation
- Differentiable image sampling
Jaderberg, et al., 2015.

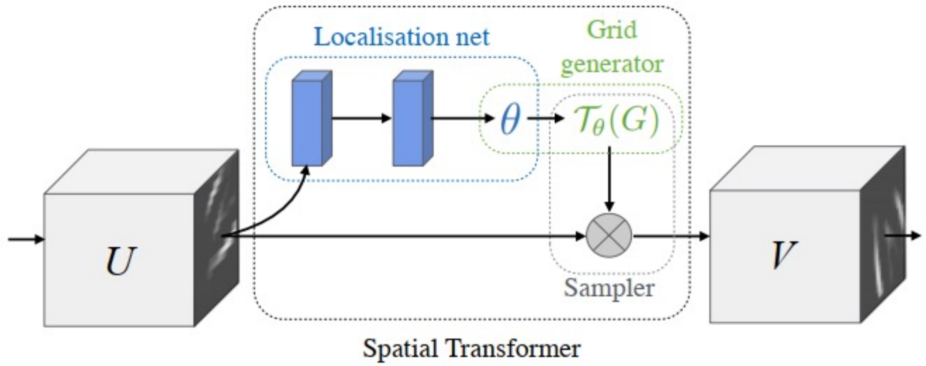
- **Loss? How to train?**

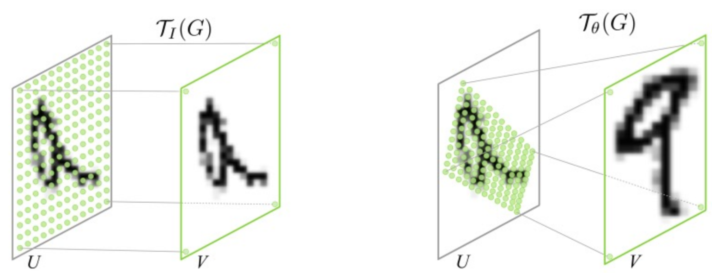


Spatial Transformer (STN)

- First method exploiting deep learning for image alignment
- STN are part of neural network for classification
- Task: spatially transform input images such that the classification task is simplified

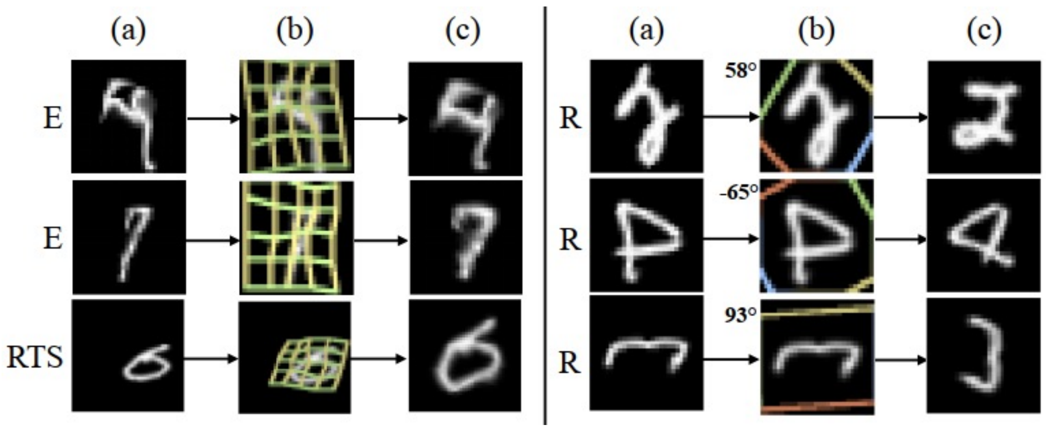
But:
- Not an explicit image registration method
- Image alignment is not guaranteed and only performed when it is beneficial for the classification task

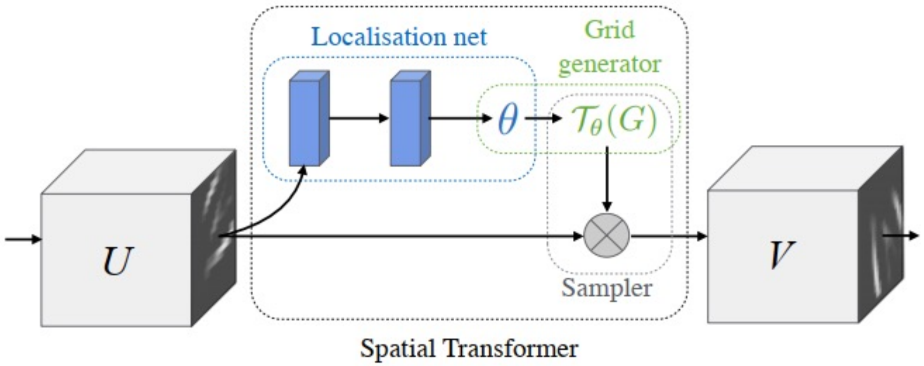


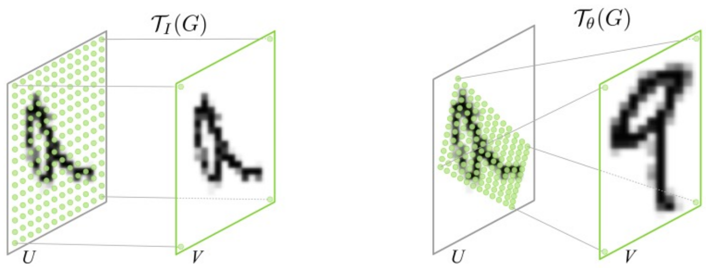


Jaderberg et al. Spatial transformer networks. *Advances in neural information processing systems*, 2015.

Spatial Transformer (STN)

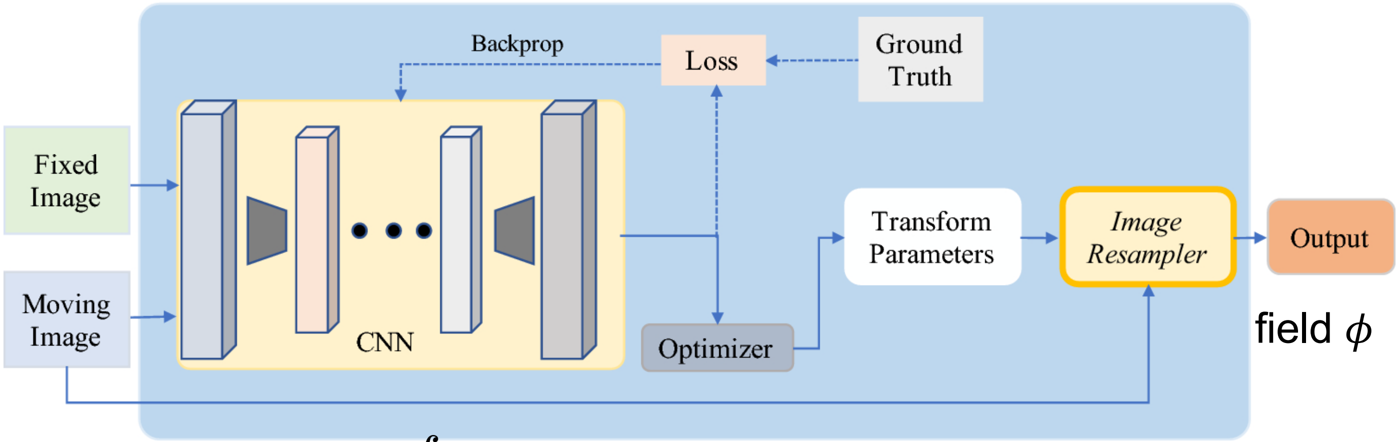






Jaderberg et al. Spatial transformer networks. *Advances in neural information processing systems*, 2015.

Supervised image registration - Training



field ϕ

network f_θ
with parameters θ

- Training data $\mathcal{F}, \mathcal{M}, \Phi$
  with triplets $\{F_i, M_i, \phi_i\}$
- ϕ from classical methods as 'ground truth', simulations, etc.
- External data (segmentations, landmarks, etc.)
- We are optimizing the network parameters, not the deformation ϕ

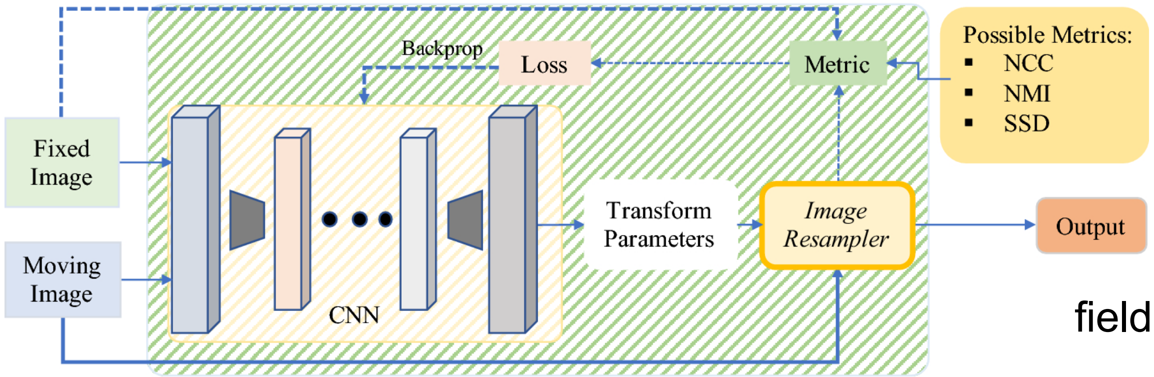
Loss function:

$$\mathcal{L}(\theta; \mathcal{F}, \mathcal{M}, \Phi) = \sum_i \| \phi_i - \underbrace{f_\theta(F_i, M_i)}_{\text{estimated } \hat{\phi}_i} \|_2^2$$

ϕ_i : Ground truth transformation

$f_\theta(\cdot, \cdot)$: Network with two inputs and weights θ

Unsupervised image registration - Training



field ϕ

network f_θ
with parameters θ

- Training data:
  only images $\mathcal{F}, \mathcal{M}$
  with image pairs $\{F_i, M_i\}$

$$\mathcal{L}(\theta;\mathcal{F},\mathcal{M}) = \sum_i \mathcal{J}(F_i, M_i, f_\theta(F_i, M_i))$$

$$= \sum_i \mathcal{D}(F_i, M_i \circ \underbrace{f_\theta(F_i, M_i)}_{\phi_i}) + \alpha\mathcal{R}(\underbrace{f_\theta(F_i, M_i)}_{\phi_i})$$

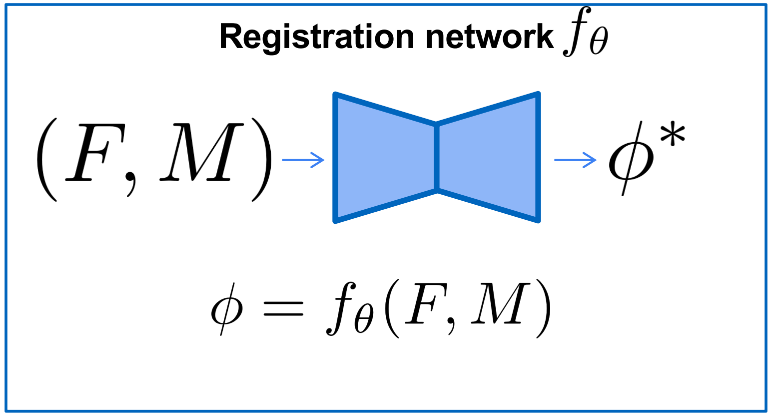
deformation between F_i and M_i

- SGD based techniques
- Each image pair contributes **slightly** to θ
  Classical optimization: slightly update ϕ for an image pair

Learning-based image registration - Inference

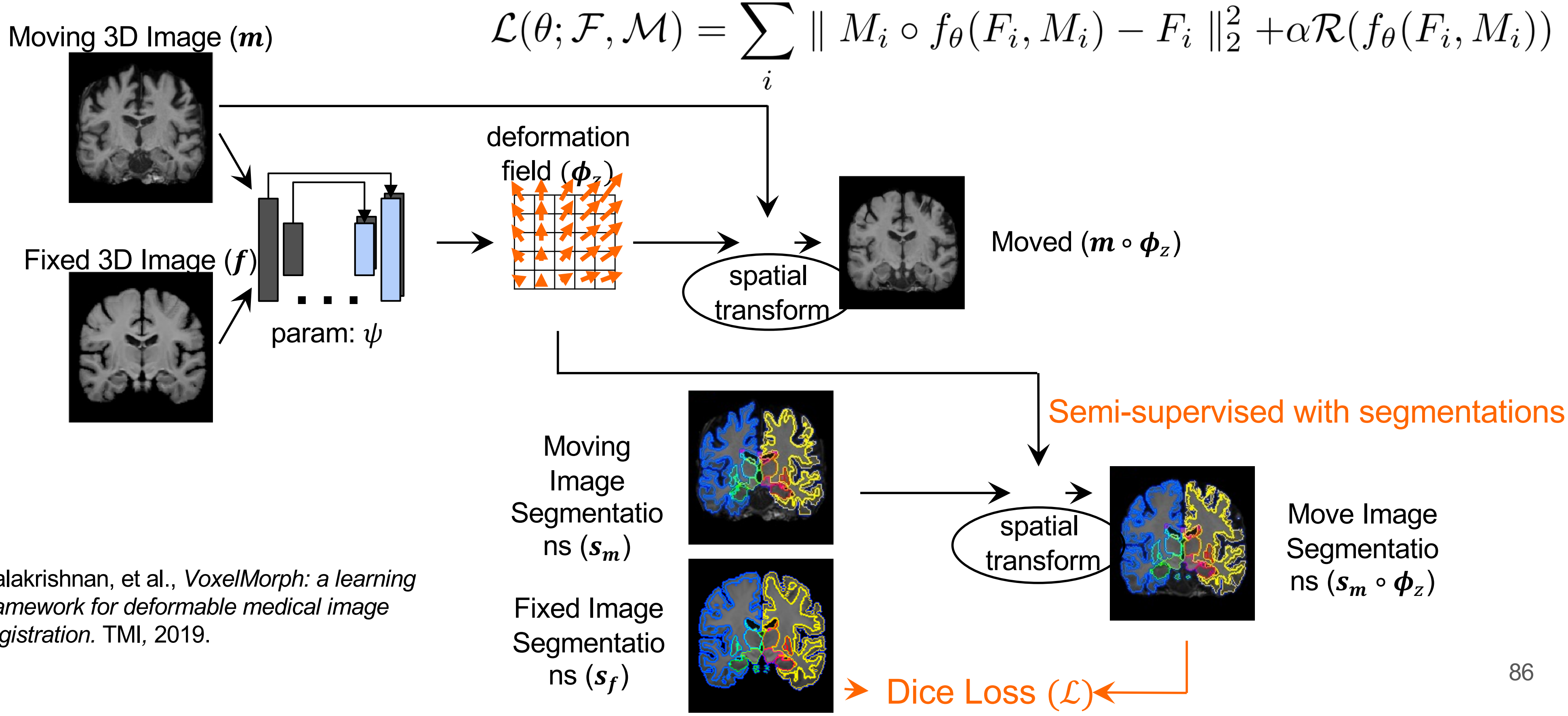
- After training, registration is trivial: a single forward pass through the network.
- At test time, only ~0.5sec for an image pair (GPU)
- *Note this might fail for an out-of-sample test pair*

Single forward pass



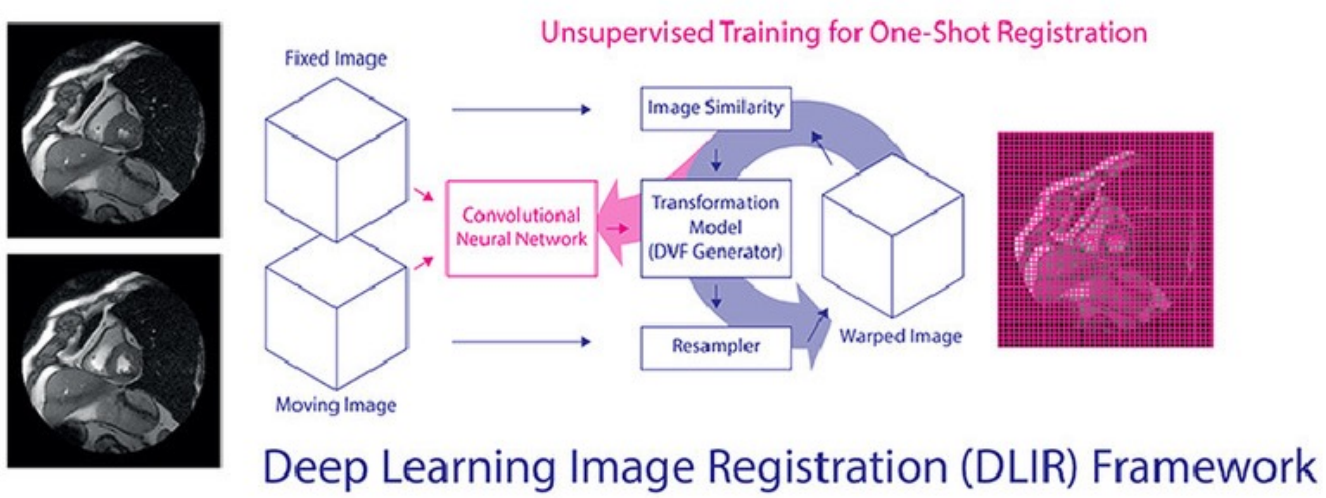
VoxelMorph

$$\mathcal{L}(\theta; \mathcal{F}, \mathcal{M}) = \sum_i \| M_i \circ f_\theta(F_i, M_i) - F_i \|_2^2 + \alpha \mathcal{R}(f_\theta(F_i, M_i))$$



Balakrishnan, et al., *VoxelMorph: a learning framework for deformable medical image registration.* TMI, 2019.

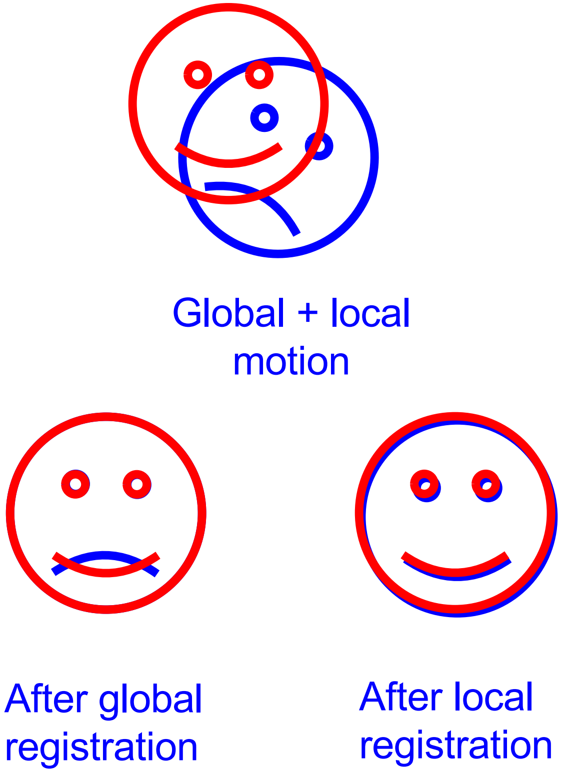
ConvNets for affine/nonlinear registration

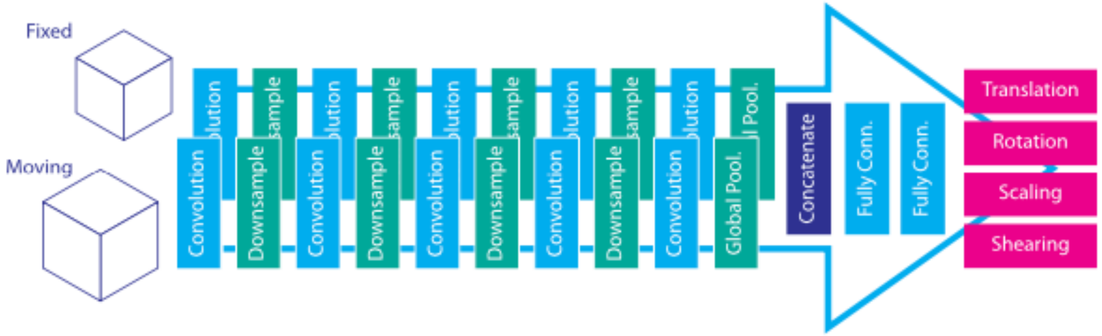


De Vos, et al., *A deep learning framework for unsupervised affine and deformable image registration*. MedIA, 2019.

ConvNets for affine/nonlinear registration

- First perform affine registration
- Uses separate pipeline for input images
- Can be of different sizes/resolutions
- Global average pooling, followed by decoding of of relative orientationsi into the 12 DOFs

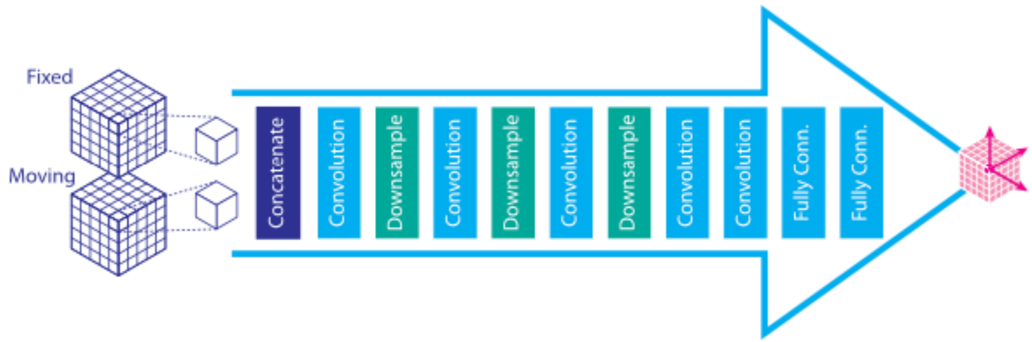


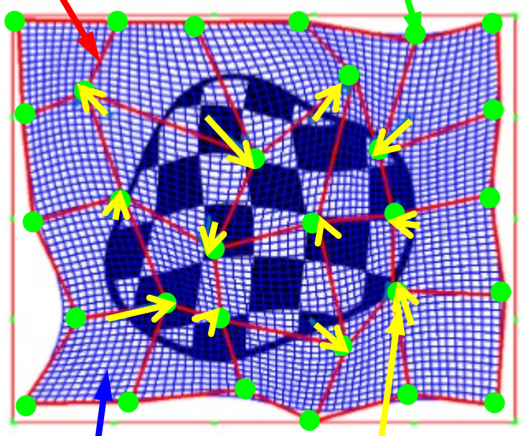


De Vos, et al., *A deep learning framework for unsupervised affine and deformable image registration*. MedIA, 2019.

ConvNets for affine/nonlinear registration

- Resample affine registration result to perform a deformable registration
- Uses single pipeline for (resampled) input images
- **Predict B-spline FFD control points**

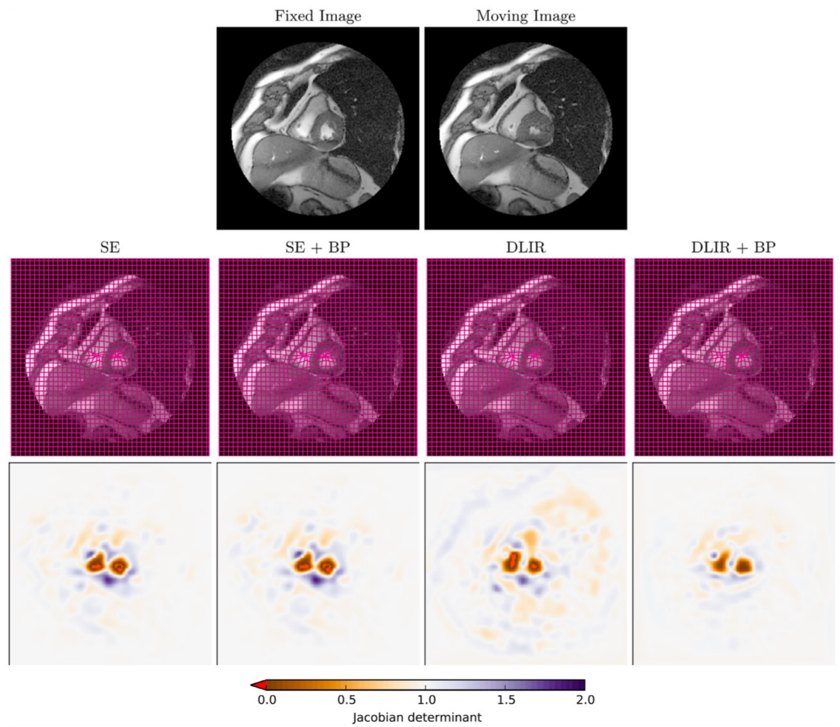




De Vos, et al., *A deep learning framework for unsupervised affine and deformable image registration*. MedIA, 2019.

ConvNets for affine/nonlinear registration

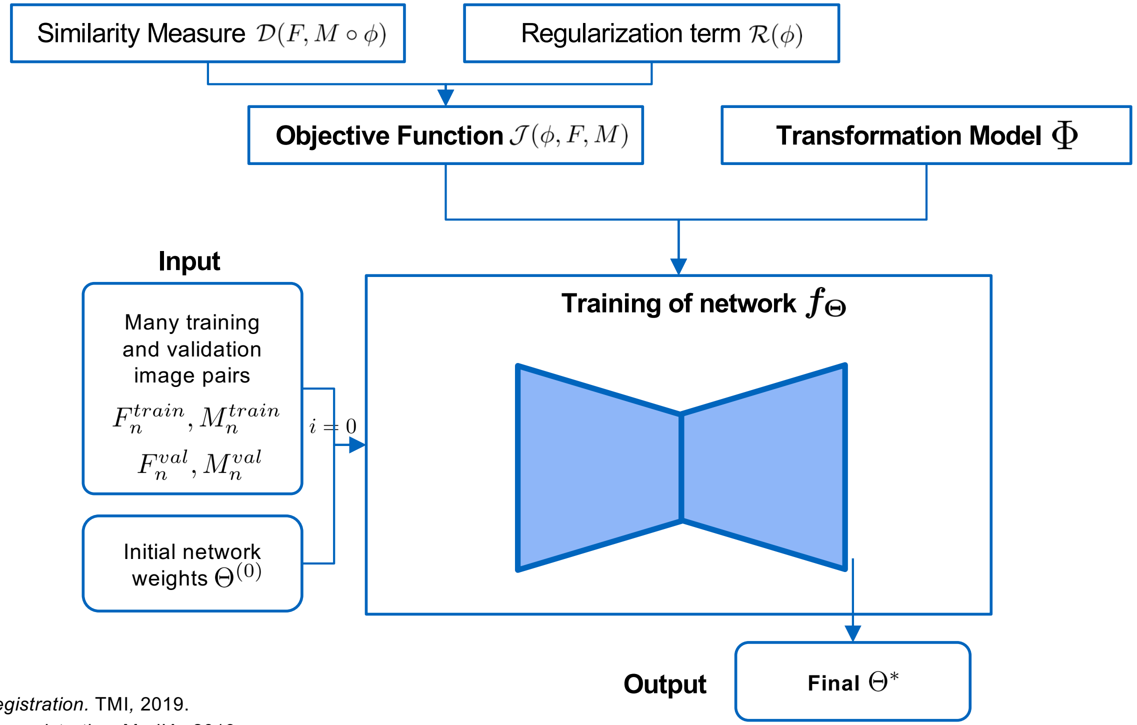
SE: baseline B-spline method ("Elastix")
DLIR: Learning-based method
BP: with/wihout bending penalty



De Vos, et al., *A deep learning framework for unsupervised affine and deformable image registration*. MedIA, 2019.

Learning-based image registration algorithm

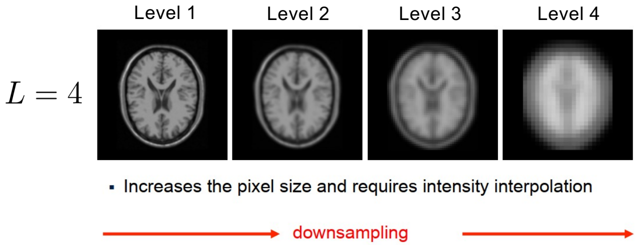
- Learn the **spatial transformation**
- Exploit information of many image pairs
- Optimization of many training samples instead of over one image pair
- CNN f_Θ parametrised by Θ to encode transformation
- Differentiable image sampling
  Jaderberg, et al., 2015.
- **Supervised**
- **Unsupervised**
- **Weakly supervised**



Unsupervised learning-based registration

Balakrishnan, et al., *VoxelMorph: a learning framework for deformable medical image registration*. TMI, 2019.
De Vos, et al., *A deep learning framework for unsupervised affine and deformable image registration*. MedIA, 2019.

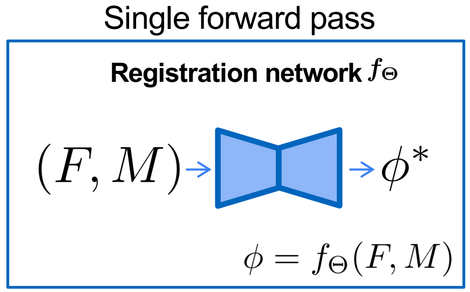
Multi-resolution networks



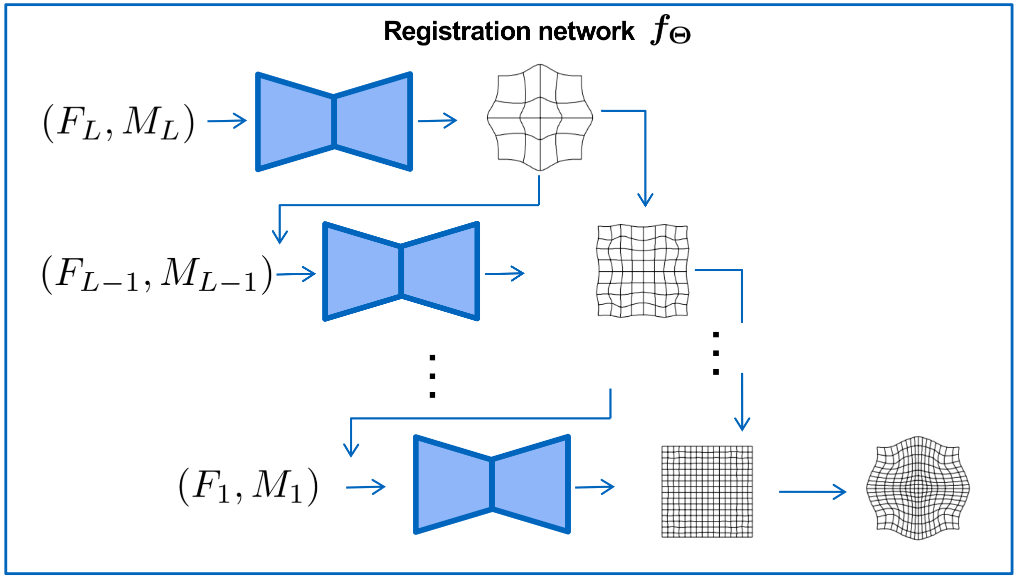
Resolution levels $\ell = 1, \dots, L$

$$\phi_\ell = f_{\Theta_\ell}(F_\ell, M_\ell)$$

$$\phi = \phi_1 \circ \phi_2 \circ \cdots \circ \phi_L$$



- Adapting strategy from conventional image registration.
- Learning transformation in multi-resolution fashion.
- Construction loss over multiple resolutions.
- Improved performance over single resolution.



de Vos, et al., A deep learning framework for unsupervised affine and deformable image registration, MedIA 2019.

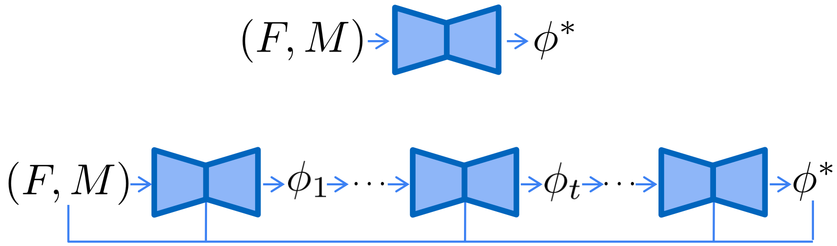
Hering, et al., mlVIRNET: Multilevel Variational Image Registration Network, MICCAI 2019.

Mok and Chung, Large Deformation Diffeomorphic Image Registration with Laplacian Pyramid Networks, MICCAI 2020.

Other learning-based approaches

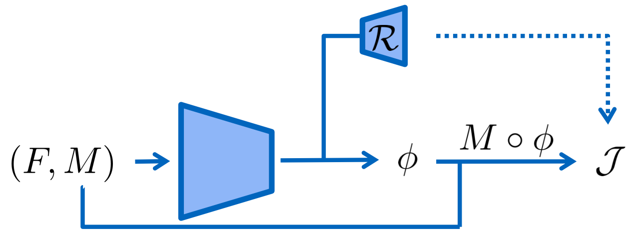
Learn the transformation in multiple steps

- A single forward pass is not accurate enough
- Often conventional approaches surpass learning-based methods



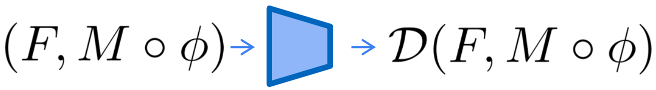
Learn the regularisation

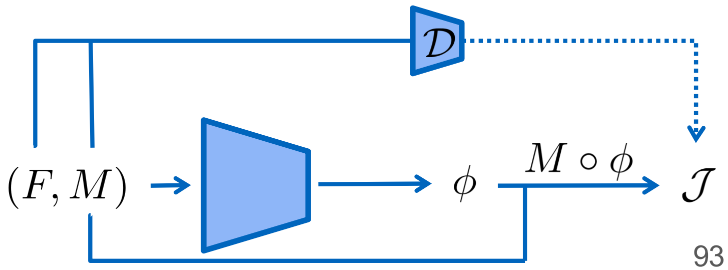
- Conventional regularisation is explicit, global, might be too restrictive.



Learn the image similarity:

- CNN f_Θ parametrised by Θ to encode distance between images



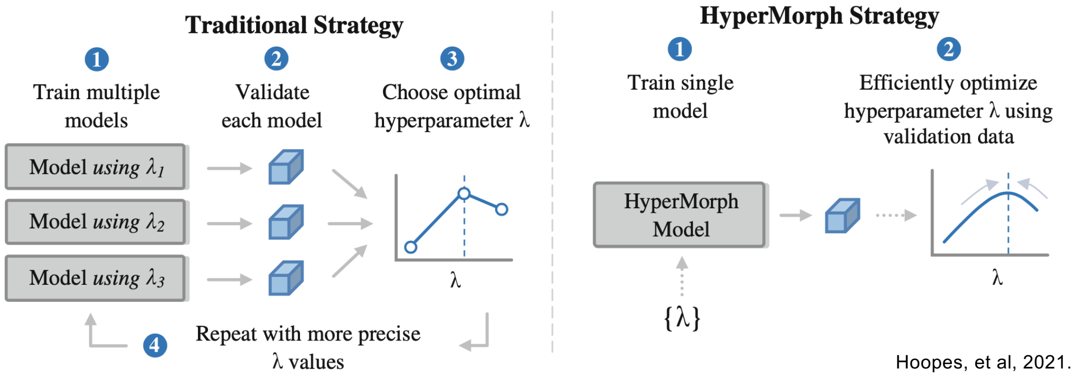


Learning regularisation: hyperparameter

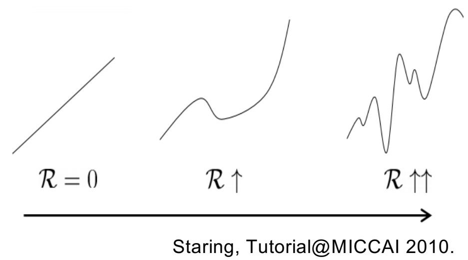
- The regularisation weight is a hyper-parameter that needs to be determined on the validation set
- Conventional tuning means one hyper parameter value one training, which is computationally costly

Loss function:

$$\mathcal{J}(\phi, F, M) = \mathcal{D}(F, M \circ \phi) + \alpha \cdot \mathcal{R}(\phi)$$



Hoopes, et al, 2021.

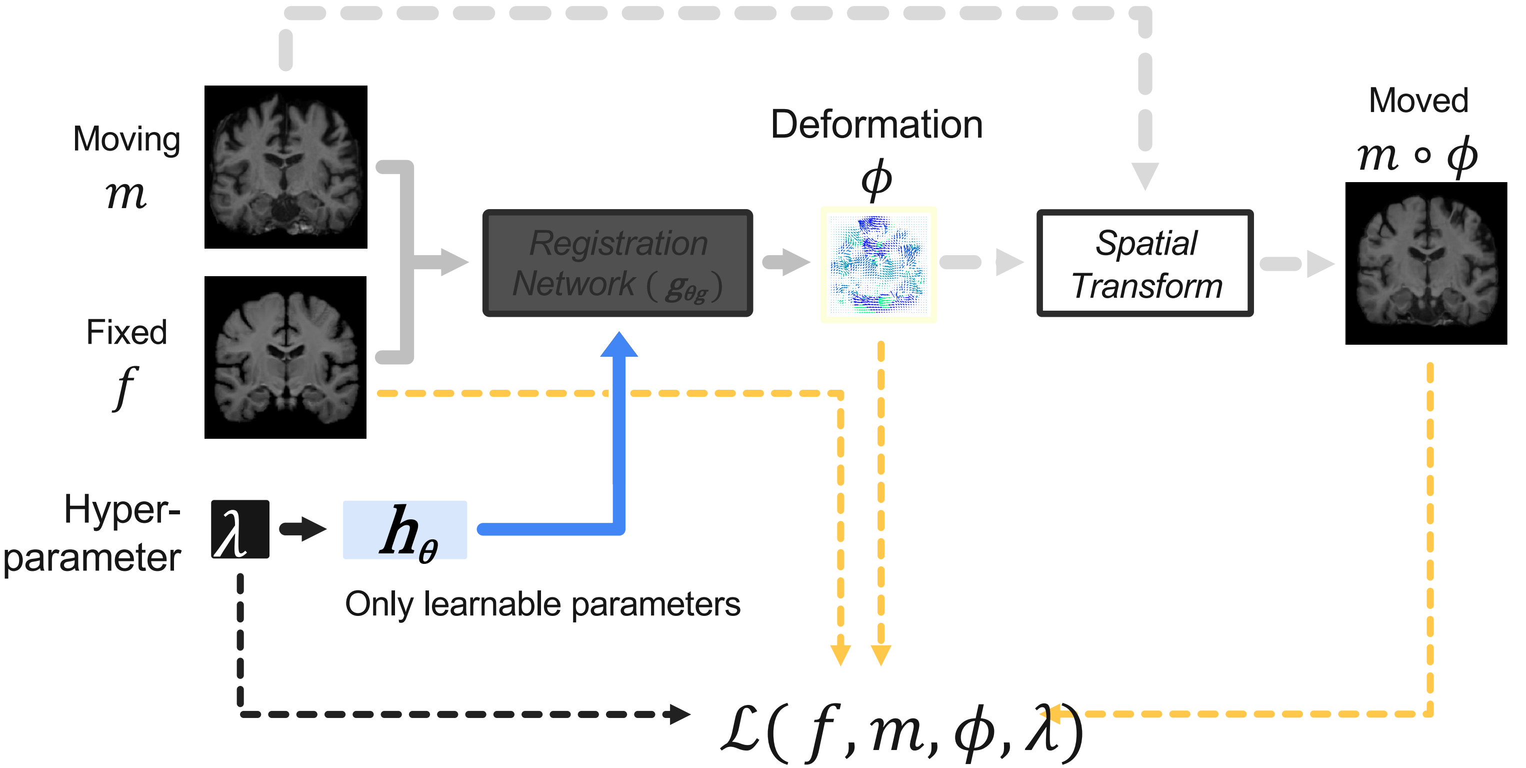


Staring, Tutorial@MICCAI 2010.

Hoopes, et al. HyperMorph: Amortized Hyperparameter Learning for Image Registration, IPMI 2021.

Mok and Chung. Conditional Deformable Image Registration with Convolutional Neural Network, MICCAI 2021.

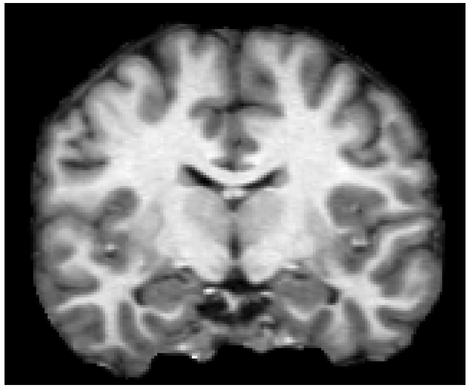
HyperMorph



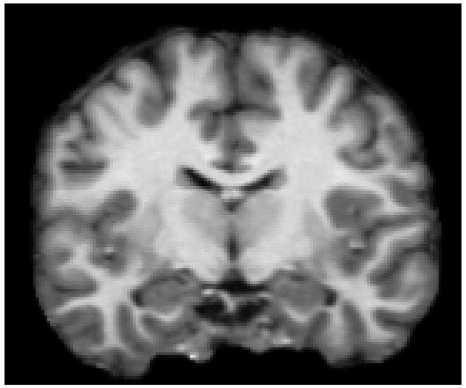
Hoopes, et al. HyperMorph: Amortized Hyperparameter Learning for Image Registration, IPMI 2021.

HyperMorph - Interactive tuning

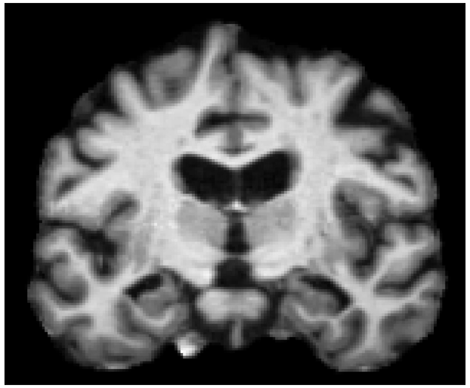
$\lambda = 1.00$



Moving Image



Moved Image



Fixed Image

Hoopes, et al. HyperMorph: Amortized Hyperparameter Learning for Image Registration, IPMI 2021.

Conventional vs. deep learning-based registration

Conventional registration:

Requires transformation model, cost function, optimisation method

− Slow for each image pair *(though discrete optimisation is reasonably fast)*

+ “pairwise optimisation”, no training data needed

Deep learning registration:

Requires network architecture design, loss function, hyperparameter tuning

− Needs many image pairs for (unsupervised) training, slow at training time

+ Very fast at inference time!

Conventional vs. deep learning registration

Similarities

- Conventional cost function becomes DL loss function
- Transformation models
- Multi-scale / -resolution processing

Differences

- Data requirements: DL models need many image pairs for training
- Speed: Iterative optimization is typically time consuming, whereas inference time for DL models can be very fast

Challenges for learning-based methods

- Out-of-distribution data: Still unclear how to treat very complex, discontinuous deformations, pathologies, anatomical variations
- Need to retrain/fine-tune for new image domains and clinical applications

Summary

Medical image registration is even harder than segmentation!

- Non-convex, ill-posed problem
- We only have surrogate evaluation measures: image similarity, label overlap, …
- *Beware that you your cost/loss function is not an objective evaluation criterion!*
- We do not have training data suitable for supervised approaches, so we mostly need to go **unsupervised** *(falling into some of the same pitfalls as for conventional registration)*
- Efforts in learning-based registration have only started in recent years, 2018 onwards
- Deep learning solutions go beyond predicting the transformation in a single forward pass.

Some literature – Before learning-based registration

PERGAMON

Pattern Recognition 32 (1999) 71–86

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

An overlap invariant entropy measure of 3D medical image alignment

C. Studholme[a,*], D.L.G. Hill[b], D.J. Hawkes[b]

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 18, NO. 8, AUGUST 1999

Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images

D. Rueckert,* L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes

Medical Image Analysis 16 (2012) 1423–1435

Contents lists available at SciVerse ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

ELSEVIER

MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration

Mattias P. Heinrich[a,b,*], Mark Jenkinson[b], Manav Bhushan[a,b], Tahreema Matin[d], Fergus V. Gleeson[d], Sir Michael Brady[c], Julia A. Schnabel[a]

Medical Image Analysis 33 (2016) 145–148

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

ELSEVIER

Editorial

Advances and challenges in deformable image registration: From image fusion to complex motion modelling

CrossMark

Julia A. Schnabel[a,c,*], Mattias P. Heinrich[b], Bartłomiej W. Papież[c], Sir J. Michael Brady[d]

Some literature – After learning-based registration

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 38, NO. 8, AUGUST 2019

VoxelMorph: A Learning Framework for Deformable Medical Image Registration

Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, John Guttag, and Adrian V. Dalca

Medical Image Analysis 52 (2019) 128–143

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

A deep learning framework for unsupervised affine and deformable image registration

Bob D. de Vos[a,*], Floris F. Berendsen[b], Max A. Viergever[a], Hessam Sokooti[b], Marius Staring[b], Ivana Išgum[a]

[a] Image Sciences Institute, University Medical Center Utrecht and Utrecht University, Utrecht, The Netherlands
[b] Division of Image Processing of the Leiden University Medical Center, Leiden, The Netherlands

Machine Vision and Applications (2020) 31:8
https://doi.org/10.1007/s00138-020-01060-x

ORIGINAL PAPER

Deep learning in medical image registration: a survey

Grant Haskins[1] · Uwe Kruger[1] · Pingkun Yan[1]

Received: 26 February 2019 / Revised: 20 December 2019 / Accepted: 14 January 2020 / Published online: 29 January 2020

References

Survey articles

- Rueckert, Daniel, and Julia A. Schnabel. "Medical image registration." *Biomedical image processing*. Springer, Berlin, Heidelberg, 2010. 131-154.
- Sotiras, Aristeidis, Christos Davatzikos, and Nikos Paragios. "Deformable medical image registration: A survey." *IEEE transactions on medical imaging* 32.7 (2013): 1153-1190.
- Viergever, Max A., et al. "A survey of medical image registration–under review." *Medical image analysis* 33 (2016): 140-144.
- Haskins, Grant, Uwe Kruger, and Pingkun Yan. "Deep learning in medical image registration: a survey." *Machine Vision and Applications* 31.1 (2020): 1-18.

Some software

- http://mirtk.github.io/ Medical Image Registration ToolKit (MIRTK)
- https://elastix.lumc.nl/ elastix
- https://github.com/airlab-unibas/airlab AIRLAB Autograd Image Registration Laboratory
- http://www.mpheinrich.de/software.html
- https://github.com/voxelmorph/voxelmorph
- https://github.com/BDdeVos/TorchIR