# Medical Image Analysis

# Image Registration I

Mauricio Reyes, PhD

University of Bern
ARTORG Center for Biomedical Engineering Research

# Outline



- Why register images?
- The registration problem and its components
- Transformations
- How to measure a match
- Two classic algorithms: Procrustes and ICP
- Optimization, interpolation, and validation
- What comes next in *Medical Image Analysis*

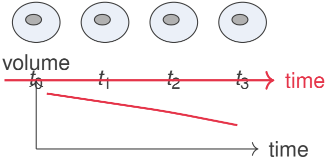
This lecture gives an overview. Advanced topics (B-splines, demons, diffeomorphisms, learning-based methods) are covered in detail in the *Medical Image Analysis* course.

# WHY REGISTER?

# Longitudinal monitoring



- Track **disease progression** over time in the same patient
- Quantify **treatment response**
- Requires aligning scans across time points so that change in the image reflects change in the patient — not motion

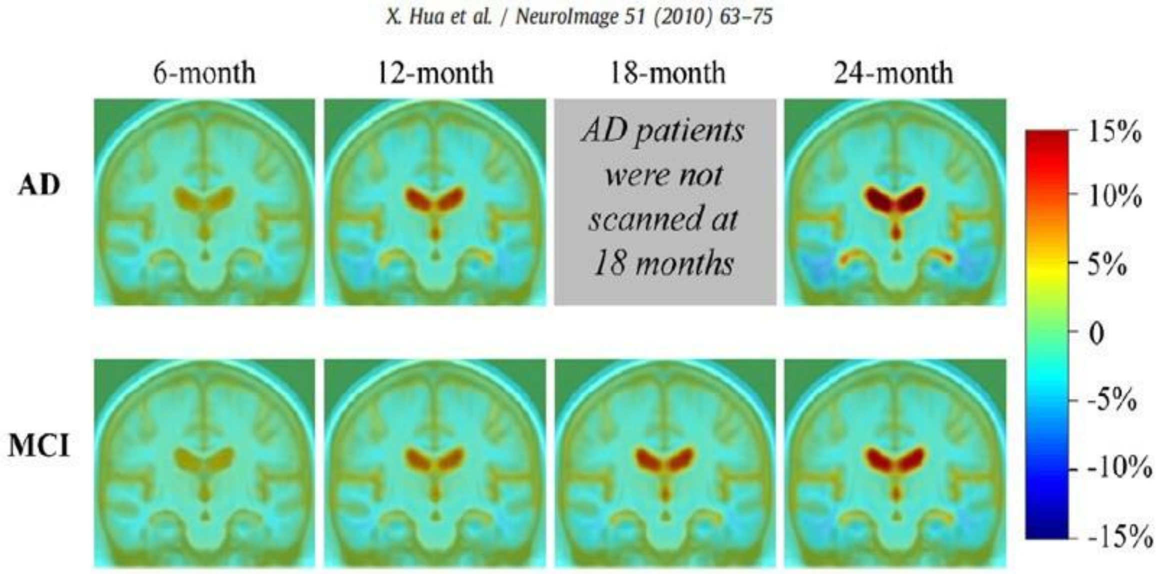



*Examples:* Alzheimer's disease atrophy maps, tumor growth, multiple sclerosis lesion evolution.

# Image registration: the global picture

## Image registration, what for?

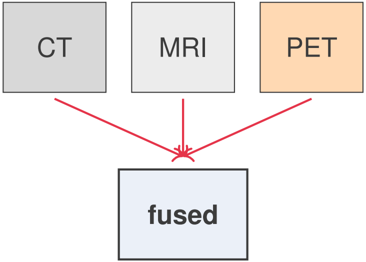
- Disease/treatment progress on a population



# Multi-modal fusion



- Different modalities reveal different information
  - **CT**: bone, dense tissue
  - **MRI**: soft tissue contrast
  - **PET**: metabolic activity
- Registration places them in a **common coordinate system** so we can read them together
- Standard in oncology, neurosurgery, cardiology



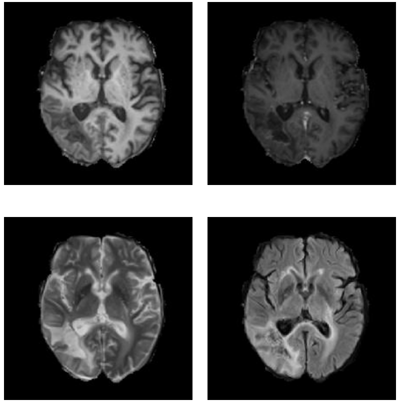

MR/PET, CT/PET fusion is now routine in clinical PET-CT scanners.

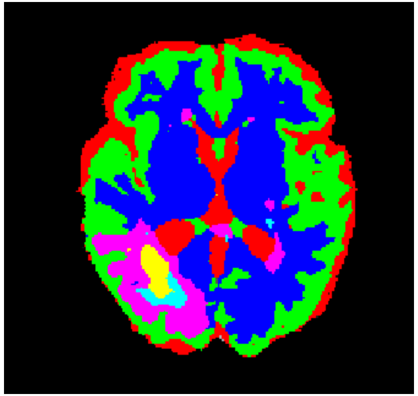
# Image registration: the global picture

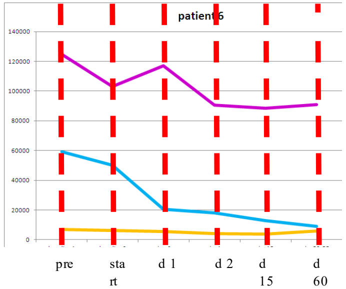
## Image registration, what for?

- Disease/treatment progress on a population

  Longitudinal brain tumor analysis

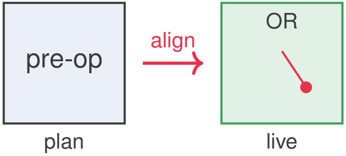






# Image-guided intervention

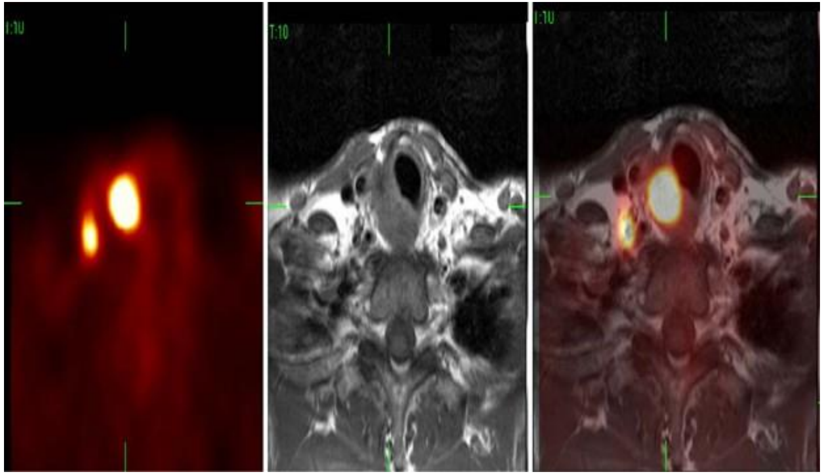
- **Pre-operative** planning images aligned with **intra-operative** reality
- Surgeon sees instrument position relative to a 3D plan
- Tumor resection, deep brain stimulation, orthopedic implants, radiation therapy
- Registration must be **fast** and **accurate** — millimeters matter



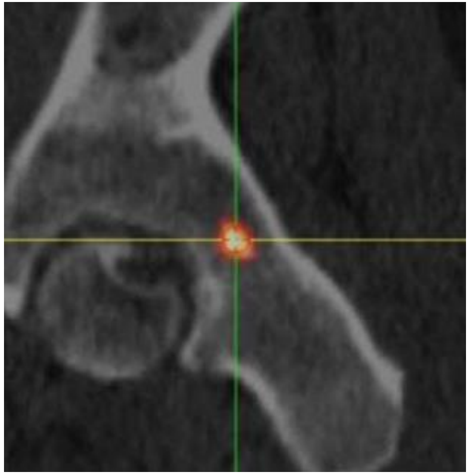
# Image registration: the global picture

## Image registration, what for?

- Correlating information across modalities



[Slomka et al. European Journal of Nuclear Medicine and Molecular Imaging, 2010]



Oliveira et al. CAS 2011

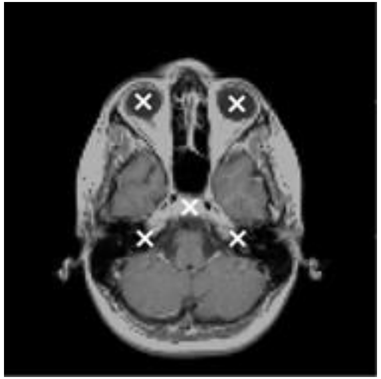
MRI/PET & CT/PET

# Image registration: the global picture

UNIVERSITÄT BERN

Image registration, what for?

- Pre-, and post-operative comparison (patient follow-up)
- Intra-operative image guided neurosurgery

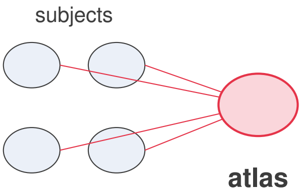


source: Abbassian, et al. 2001

# Atlas-based analysis and population studies



- An **atlas** is a reference image (or template) with anatomical labels
- Registering each subject to the atlas allows:
  - Automatic anatomical labeling
  - Statistical comparison across subjects
  - Group studies (control vs. patient)
- Foundation of computational neuroanatomy

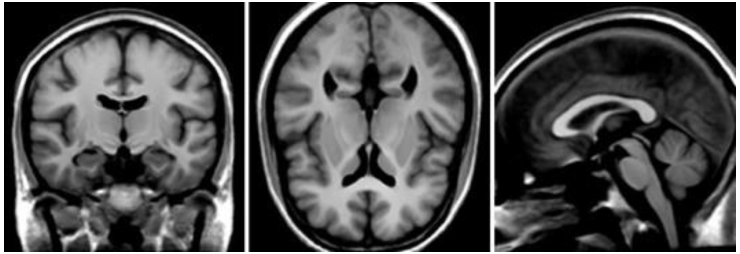


# Image registration: the global picture

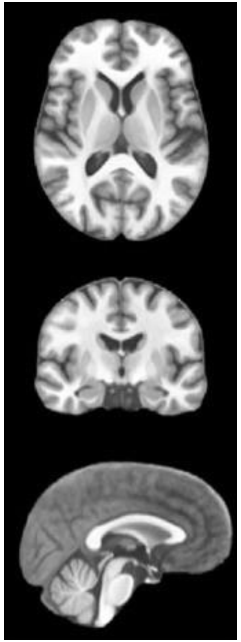
UNIVERSITÄT BERN

## Image registration, what for?

- Atlas creation



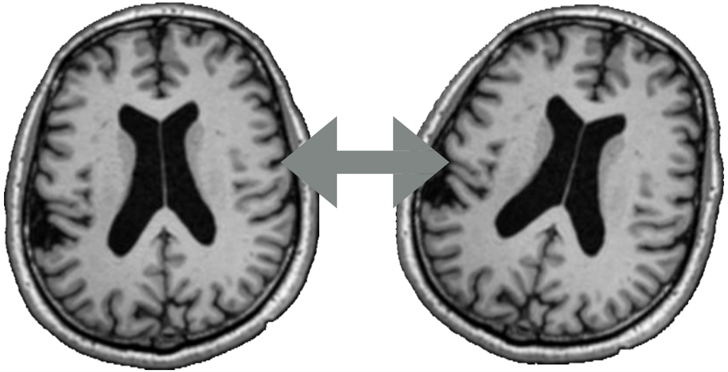
Rueckert et al. Miccai 2001



Rohlfing et al. Human Brain Mapping, 2010

# Image registration: the global picture

Image registration, how?



Apply a given **transformation** until a given **matching criteria** is fulfilled

Image registration is an optimization process!

# The registration problem



- Given two images $I$ (**fixed**) and $J$ (**moving**), find a transformation $T$ that aligns $J$ with $I$:

$$T^* = \arg\max_{T} \mathcal{S}\big(I, T(J)\big)$$

- $\mathcal{S}$ is a **similarity** measure — the higher, the better the match
- Equivalently, with a dissimilarity $\mathcal{D}$: $T^* = \arg\min_T \mathcal{D}\big(I, T(J)\big)$
- In voxel form, looking for a displacement $\mathbf{u}(\mathbf{x})$:

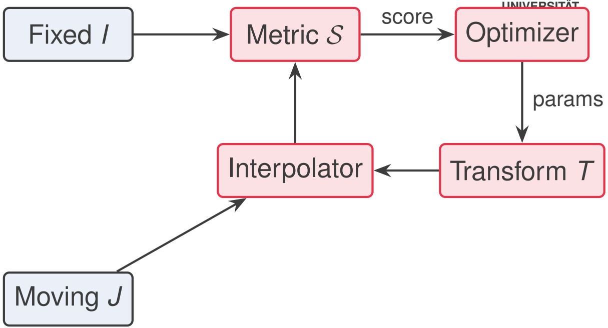
$$I(\mathbf{x}) \approx J\big(\mathbf{x} + \mathbf{u}(\mathbf{x})\big)$$

**Image registration is an optimization problem.**

# The four components

- **Transform** $T$ — the model of motion (rigid, affine, . . . )
- **Metric** $\mathcal{S}$ — how we score a match
- **Optimizer** — searches the parameter space
- **Interpolator** — resamples $J$ at non-grid locations

Every registration method is a particular choice for each of these four boxes.




Adapted from the ITK Software Guide.

# Registration vocabulary



- **Fixed image** $I$ — the reference; stays in place. *a.k.a.* target, reference
- **Moving image** $J$ — gets transformed. *a.k.a.* source, template
- **Image space** — discrete voxel indices $(i, j, k)$
- **Physical space** — real-world coordinates in mm where registration actually happens
- **Degrees of freedom (DOF)** — number of parameters in the transformation

Convention matters: most software (ITK, ANTs, SimpleITK, Elastix) maps points from the *fixed* into the *moving* image.

# Taxonomy of Image Registration

**Modality**: Mono-modal, multi-modal

**Subject**: Intra-subject, inter-subject, atlas

**Transform**: Rigid, similarity, (poly-)affine, non-rigid

**Source of information**: intensity-based, non-intensity-based

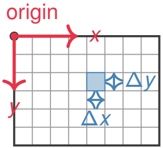
Interaction: manual, semi-, fully-automatic

**Dimensionality**: 2D-2D, 3D-3D, 2D-3D

**Domain:** local, global

# From voxel grid to physical space

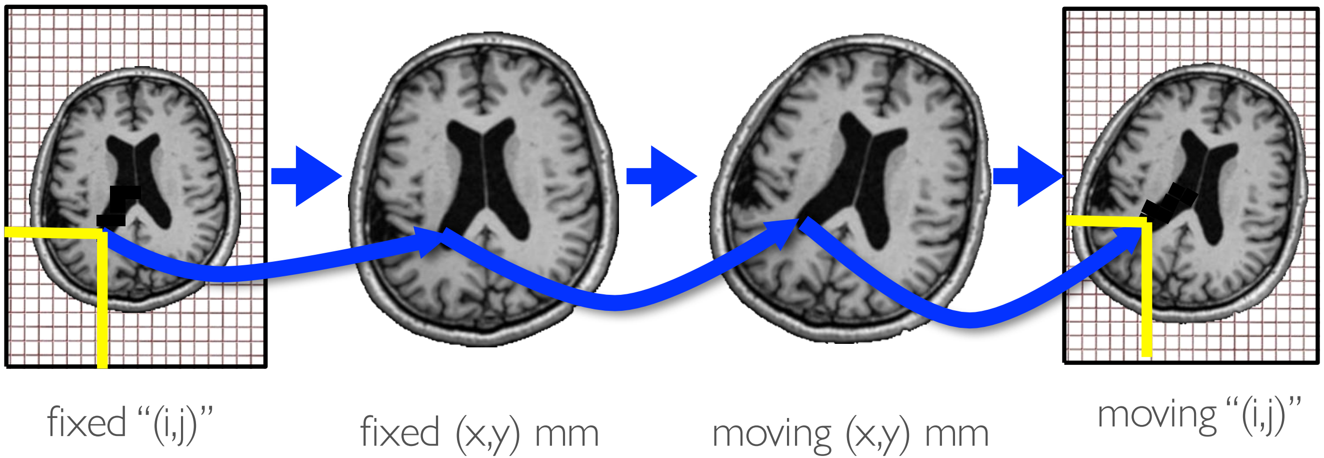
- A voxel index $(i, j)$ is just a label
- To register, we need **real positions in mm**
- Three pieces of metadata define the mapping:
  - **Origin** — physical coords of voxel $(0, 0)$
  - **Spacing** — voxel size in mm
  - **Direction** — orientation of the axes




Stored in NIfTI / DICOM headers. Mismatched metadata is the #1 cause of "my registration failed."

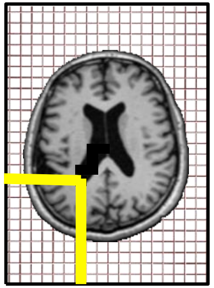
# Main components of image registration

## Physical and Discrete space

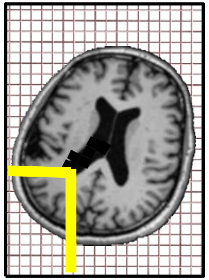


# Main components of image registration

## Eulerian frame



(i,j)



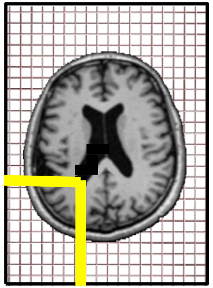
(k,l)

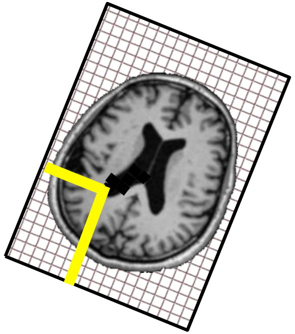
$$(i, j) = (k, l)$$

$$I(x(t = 1)) \sim J(x)$$

Flow of material over time

## Lagrangian frame

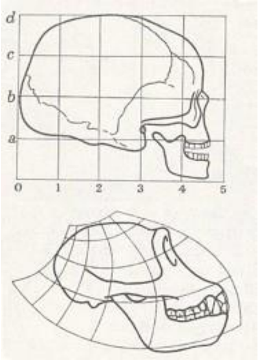




$$(i, j) = (k, l)$$

Material fix to grid

$$I(x + u(x)) \sim J(x)$$



D'ARCY THOMPSON 1917

UNIVERSITÄT

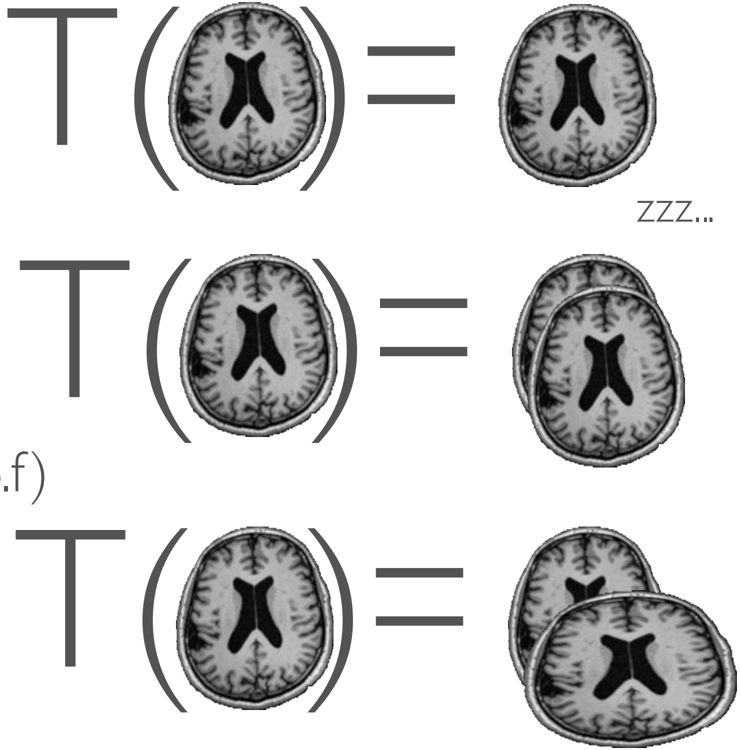
BERN

# TRANSFORMATIONS

# Main components of image registration

## Transformations

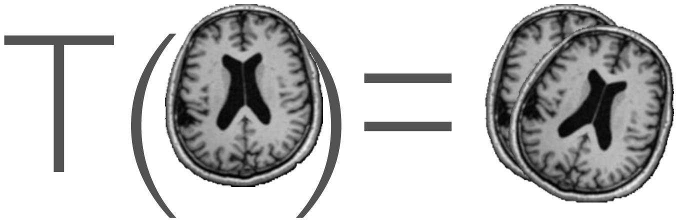
- Identity
- Translation (3 d.o.f)
- (Anisotropic) scaling (3 d.o.f)



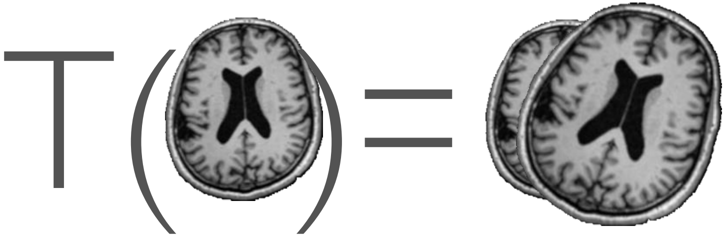
# Main components of image registration

**Transformations**

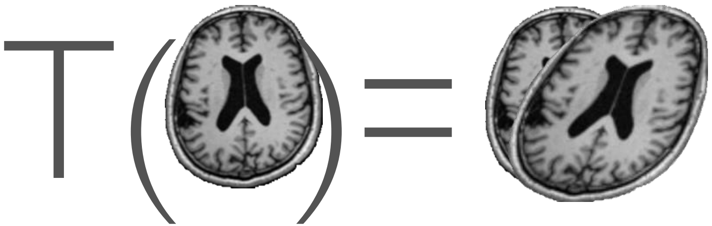
- Rigid (6 d.o.f)



- Similarity (7 d.o.f)



- Affine (12 d.o.f)

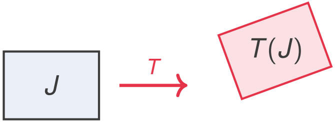


- Non-rigid (for later)

# Rigid transformation (6 DOF)



- 3 translations + 3 rotations
- **Preserves**: distances, angles, shapes
- Right model when the object is a rigid body
  - Bones, skull (no soft tissue)
  - Same patient, same session, head fixation



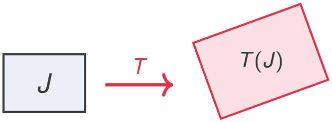

$$T(\mathbf{x}) \;=\; R\,\mathbf{x} + \mathbf{t}, \qquad R^\top R = I,\ \det R = 1$$

# Similarity transformation (7 DOF)



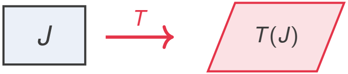
- Rigid + 1 **uniform** scale factor
- **Preserves**: angles, shape
- **Changes**: distances, by a constant ratio
- Useful when calibration differs slightly between scans, or across subjects of similar shape

$$T(\mathbf{x}) = s\,R\,\mathbf{x} + \mathbf{t}$$



# Affine transformation (12 DOF)






- Rigid + anisotropic scaling + shear
- **Preserves**: parallel lines, ratios along a line
- **Does not preserve**: angles or shape in general
- Common as a global pre-alignment step before non-rigid

$$T(\mathbf{x}) = A\mathbf{x} + \mathbf{t}, \qquad A \in \mathbb{R}^{3\times 3}$$

# Composition of transformations



- Multiple transforms can be chained:

$$T_1 \circ T_2\,(\mathbf{x}) \;=\; T_1\big(T_2(\mathbf{x})\big)$$

  $T_2$ is applied **first**, then $T_1$.
- **Order matters** — in general:

$$T_1 \circ T_2 \;\neq\; T_2 \circ T_1$$

- Inverses (when they exist) cancel:

$$T^{-1} \circ T \;=\; \mathrm{Id}$$

- Rigid, similarity, affine are all invertible. Non-rigid transforms generally need extra constraints (e.g. *diffeomorphisms*) to guarantee invertibility.

# Matrix form and homogeneous coordinates



- All linear transformations (rigid, similarity, affine) can be written as a matrix multiplication
- Translation alone is *not* linear — but the trick of **homogeneous coordinates** bundles it into a single matrix

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{M} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

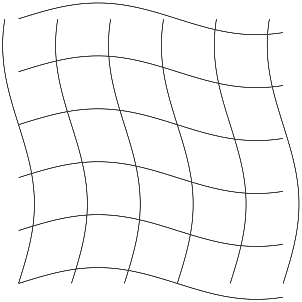
- Top-left $3 \times 3$ block: **rotation, shear, scale**
- Right column: **translation**
- Bottom row: keeps the 1

Composing transforms is just matrix multiplication: $M_{\text{total}} = M_1 \cdot M_2$.

# Beyond affine: non-rigid registration

UNIVERSITÄT
BERN

- Real anatomy deforms locally:
  - Breathing, heartbeat, peristalsis
  - Brain shift during surgery
  - Inter-subject anatomical variability
- Need a transformation with many local DOFs:
  - Free-form deformations (B-splines)
  - Demons / optical flow
  - Diffeomorphic methods (LDDMM, SyN)
  - Learning-based (e.g. VoxelMorph)



→ **Covered in detail in *Medical Image Analysis*.**

# HOW DO WE MEASURE A MATCH?

# Two families of similarity

UNIVERSITÄT BERN

### Feature-based

- Match **landmarks**, points, curves, surfaces
- Landmarks can be:
  - **Intrinsic** — anatomical (sulci, vessel bifurcations)
  - **Extrinsic** — attached fiducials (markers, frames)
- Similarity = Euclidean distance between corresponding points
- Sparse, fast, interpretable

### Intensity-based

- Use the **voxel intensities directly**
- No need to extract features
- Dense, automatic
- Choice of metric depends on the relationship between intensities of $I$ and $J$

Both families exist on a spectrum — many modern methods combine the two.

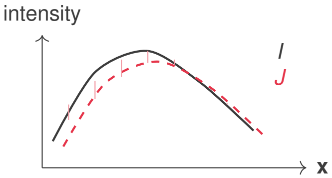
# Intensity metric 1 — Sum of squared differences (SSD)



- Simplest possible idea:

$$\mathrm{SSD}(I, J) = \sum_{\mathbf{x}} \left(I(\mathbf{x}) - J(\mathbf{x})\right)^2$$

- Assumes the same anatomy has the **same intensity** in both images
- Works well: **same modality**, same scanner, same session
- Fails: across modalities, with bias fields, contrast changes




Squared per-voxel differences, summed.

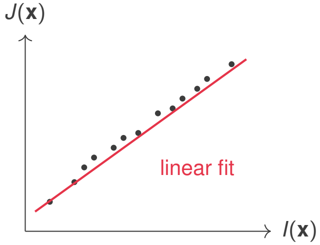
# Intensity metric 2 — Normalized cross-correlation (NCC)



- Measures **linear** relationship between intensities:

$$\mathrm{NCC}(I, J) = \frac{\sum_{\mathbf{x}} \left(I(\mathbf{x}) - \bar{I}\right)\left(J(\mathbf{x}) - \bar{J}\right)}{\sigma_I \, \sigma_J}$$

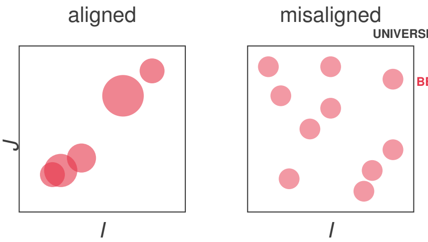
- Invariant to constant offset and scaling of intensities
- Works when $J = a\,I + b$ approximately
- Robust across same-modality scans with different brightness/contrast
- Range: $[-1, +1]$; we maximize




Pairs of corresponding intensities lie on a line.

# Intensity metric 3 — Mutual information (MI)

- For **multi-modal** registration, intensities are not even linearly related (think CT vs. MRI)
- **Mutual information**: measures *statistical dependence* between *I* and *J*
  - "Knowing the intensity in *I*, how well can I predict the intensity in *J*?"
- When images are aligned, the joint intensity distribution becomes **compact** (clusters)
- When misaligned, it spreads out

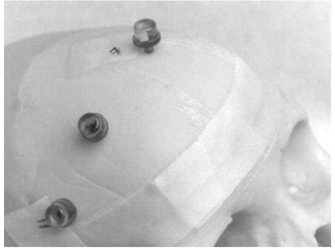


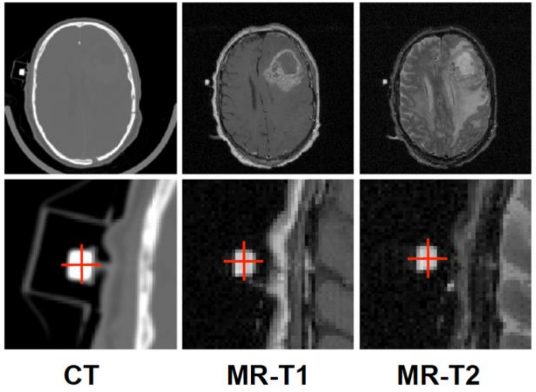

Joint intensity histograms.

Introduced for medical registration in 1995–97 (Viola & Wells; Maes et al.). Still widely used.

# Feature-based and intensity-based

- Feature-based
  - Uses points, landmarks, curves, etc. to compute a transformation
  - Intrinsic (part of the anatomy) or extrinsic (fiducial point)
  - Similarity is based on an Euclidean distance
- Intensity-based
  - Similarity constructed based on voxel intensity information





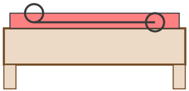
# TWO CLASSIC ALGORITHMS

# Procrustes



- Named after a robber in Greek mythology who offered travellers a “perfectly fitting” bed. . .
- . . . but it was the **guest** who was altered to fit the bed: stretched if too short, trimmed if too tall
- The hero Theseus eventually subjected Procrustes to his own method

**In registration: align two point sets by transforming one to fit the other.**



“one size fits all”

# Procrustes — the math



- Given two sets of $N$ **corresponding** points $\{\mathbf{x}_i\}$ and $\{\mathbf{y}_i\}$, find the rigid transform $(R, \mathbf{t})$ that minimizes:

$$E(R,\mathbf{t}) \;=\; \frac{1}{N}\sum_{i=1}^{N} \left\| \, R\,\mathbf{x}_i + \mathbf{t} \,-\, \mathbf{y}_i \, \right\|^2$$

- **Closed-form solution** for the rigid case — no iteration needed:
  - — Center both point sets on their means
  - — Build the cross-covariance matrix $H = \sum_i \tilde{\mathbf{x}}_i \tilde{\mathbf{y}}_i^\top$
  - — Decompose $H = U \Lambda V^\top$ via **SVD**
  - — $R = VDU^\top$ with $D = \mathrm{diag}(1, 1, \det(VU^\top))$
  - — $\mathbf{t} = \bar{\mathbf{y}} - R\,\bar{\mathbf{x}}$

The catch: it requires *known correspondences*. What if we don't have them?

# Iterative Closest Point (ICP)



**Problem**: align two surfaces / point clouds when correspondences are *unknown*.

**Idea**: alternate between guessing correspondences and solving Procrustes.

**Algorithm** (Besl & McKay, 1992):

1. For each point in $X$, find the closest point in $Y \to$ pseudo-correspondences
2. Solve Procrustes for $(R, \mathbf{t})$
3. Apply $(R, \mathbf{t})$ to $X$
4. Repeat until convergence

> Simple, fast, widely used

> Workhorse of **computer-assisted surgery** (registering pre-op CT to intra-op probe points)

> Sensitive to initialization and outliers — many robust variants exist

$\to$ **More in the *Computer Assisted Surgery* lecture.**

# ICP — visual convergence

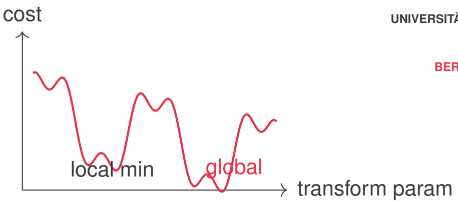



- Each iteration: re-pair, re-solve, re-apply
- Cost decreases monotonically $\rightarrow$ converges to a local minimum
- Good initialization is essential — ICP **cannot** recover from a bad starting pose

# Optimization in one slide



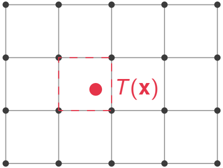
- Registration = searching the parameter space for the best $T$
- The cost surface is **not convex** — local minima are common
- Standard tactics:
  - **Good initialization** (often: center the images first)
  - **Multi-resolution pyramid** — start coarse, refine
  - Gradient-based methods (gradient descent, L-BFGS) when derivatives exist
  - Derivative-free methods (Powell, evolutionary) otherwise



# Interpolation in one slide



- After applying $T$, sample positions **don't fall on voxel centers**
- We need to estimate intensities at non-grid locations
- Common choices:
  - **Nearest neighbor** — fast, blocky, fine for label maps
  - **Linear** / **trilinear** — the default for intensities
  - **B-spline** — smoother, slower
- Choice of interpolator affects both accuracy and the smoothness of the cost surface




Off-grid sample needs interpolation from neighbors.

# Validating registration



- Ground truth is rare — so validation is hard
- Common strategies:
  - **Target Registration Error (TRE)** — distance between known landmark pairs after registration
  - **Overlap of propagated labels** — Dice / Jaccard between transformed segmentation and reference
  - **Visual inspection** — checkerboard overlays, difference images
  - **Plausibility** — e.g. Jacobian determinant of the transformation should stay positive (no folding)
- Synthetic experiments: apply a known transform, recover it, measure error

Beware: a low intensity-based metric does not guarantee a good anatomical match.

# What's next in *Medical Image Analysis*

UNIVERSITÄT
BERN

- **Non-rigid registration** in depth
  - Free-form deformations with B-splines
  - Demons / optical flow
  - Diffeomorphic methods (LDDMM, SyN)
- Mutual information — formal definition and entropy view
- Multi-resolution and regularization in detail
- **Learning-based registration** — VoxelMorph and successors; foundation models for medical imaging
- Specific applications: atlas construction, group studies, image-guided therapy

**This lecture gave you the vocabulary — the advanced course gives you the toolbox.**

# Summary

UNIVERSITÄT
BERN

- Registration aligns images by **optimizing a similarity** over a **transformation**
- Four building blocks: **transform, metric, optimizer, interpolator**
- Transformations form a hierarchy: identity $\rightarrow$ translation $\rightarrow$ rigid $\rightarrow$ similarity $\rightarrow$ affine $\rightarrow$ non-rigid
- Metrics depend on intensity relationship: **SSD** (same), **NCC** (linear), **MI** (any)
- **Procrustes** solves rigid alignment in closed form when correspondences are known
- **ICP** extends this when they aren't
- Validation requires ground truth surrogates — be careful with metric values

# References


- P. J. Besl & N. D. McKay, “A Method for Registration of 3-D Shapes,” *IEEE TPAMI*, 14(2), 1992.
- F. Maes et al., “Multimodality Image Registration by Maximization of Mutual Information,” *IEEE TMI*, 16(2), 1997.
- P. Viola & W. M. Wells, “Alignment by Maximization of Mutual Information,” *IJCV*, 24(2), 1997.
- D. Rueckert et al., “Nonrigid Registration Using Free-Form Deformations,” *IEEE TMI*, 18(8), 1999.
- B. B. Avants et al., “Symmetric Diffeomorphic Image Registration with Cross-Correlation (SyN),” *Medical Image Analysis*, 12(1), 2008.
- G. Balakrishnan et al., “VoxelMorph: A Learning Framework for Deformable Medical Image Registration,” *IEEE TMI*, 38(8), 2019.
- ITK Software Guide, `itk.org`.




## Optional homework:

- What is a *versor* and how does it parameterize a 3D rotation?
- How does a *projective* transformation differ from an affine one?
- Why does mutual information work across modalities while SSD does not?