Introduction to medical image registration

Pietro Gori

Enseignant-chercheur
Equipe IMAGES - Télécom Paris
pietro.gori@telecom-paris.fr

October 02, 2019
Summary

1 Introduction

2 Geometric global transformations

3 Image warping and interpolations

4 Intensity based registration

5 Landmark based registration
   - Affine registration
   - Procrustes superimposition
   - Non-linear registration (small displacement)
   - Non-linear registration (diffeomorphism)
### Definition

- **Geometric**: find the *optimal* parameters of a *geometric transformation* to spatially align two different images of the same object. It establishes *spatial correspondence* between the pixels of the source (or moving) image with the ones of the target image.
- **Photometric**: Modify the intensity of the pixels and not their position.

### Applications

- Compare two (or more) images of the same modality (e.g. T1-w MRI of the brain of two different subjects).
- Combine information from multiple modalities (e.g. PET, DWI, T1-w MRI of the brain of the same subject).
- Longitudinal studies (e.g. monitor anatomical or functional changes of the brain over time).
- Relate preoperative and postoperative images after surgery.
Medical Image registration - Applications

- Same modality, different subjects
- Different modalities, same subject
- Longitudinal study
- Pre and postoperative
Image registration components

- **Dimensionality**: 2D/2D, 3D/3D, 2D/3D
- **Transformation** (linear / non-linear)
- **Similarity metric** (e.g. intensities, landmarks, edges, surfaces)
- **Optimization procedure**
- **Interaction** (automatic / semi-automatic / interactive)
- **Modalities** (mono-modal / multi-modal)
- **Subjects** (intra-subject / inter-subject / atlas construction)
Let $I$ and $J$ be the source and target images. They show the same anatomical object, most of the time with a different field of view and resolution (sampling).

$I(x, y)$ and $J(u, v)$ represent the intensity values of the pixels located onto two regular grids: \{\(x, y\)\} $\in \Omega_I$ and \{\(u, v\)\} $\in \Omega_J$.

For the same subject, the same anatomical point $z$ can be in position $x_z, y_z$ in $I$ and in $u_z, v_z$ in $J$. 
Mathematical definition

- Both $I$ and $J$ are functions:

$$I(x, y) : \Omega_I \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$(x, y) \rightarrow I(x, y)$$

- We look for a geometric transformation $T$, which is a 2D warping parametric function that belongs to a certain family $\Gamma$:

$$T_\phi(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$(x, y; \phi) \rightarrow T_\phi(x, y)$$

- $\phi$ is the vector of parameters of $T$. We look for a transformation that maps $(x_z, y_z)$ to $(u_z, v_z)$
Mathematical definition

- Most of the time, $I$ and $J$ are simply seen as matrices whose coordinates $(x, y)$ and $(u, v)$ are thus integer-valued (number of line and column).
- The values of the intensities of the pixels can be real numbers $\mathbb{R}$ (better for computations) or integer, usually in the range $[0, 255]$ for a gray-scale image.
- There are several kind of transformations:
Summary

1. Introduction

2. Geometric global transformations

3. Image warping and interpolations

4. Intensity based registration

5. Landmark based registration
   - Affine registration
   - Procrustes superimposition
   - Non-linear registration (small displacement)
   - Non-linear registration (diffeomorphism)
Global transformations

- Global means that the transformation is the same for any points $p$
- **Scaling** - multiply each coordinate by a scalar

$$
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
$$

(1)
**Global transformations**

- **Rotation** - WRT origin. Let \( p = [x \ y]^T \) and \( p' = [u \ v]^T \)

\[
\begin{bmatrix}
x \\
y
\end{bmatrix} = \begin{bmatrix}
r \cos(\alpha) \\
r \sin(\alpha)
\end{bmatrix}
\]
\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix}
r \cos(\alpha + \theta) \\
r \sin(\alpha + \theta)
\end{bmatrix} = \begin{bmatrix}
r(\cos(\alpha) \cos(\theta) - \sin(\alpha) \sin(\theta)) \\
r(\sin(\alpha) \cos(\theta) + \cos(\alpha) \sin(\theta))
\end{bmatrix}
\]
\[
\begin{bmatrix}
x \cos(\theta) - y \sin(\theta) \\
y \cos(\theta) + x \sin(\theta)
\end{bmatrix} = \begin{bmatrix}
\cos(\theta) & -\sin(\theta) \\
\sin(\theta) & \cos(\theta)
\end{bmatrix} \begin{bmatrix}
x \\
y
\end{bmatrix}
\]

- \( R = \begin{bmatrix}
\cos(\theta) & -\sin(\theta) \\
\sin(\theta) & \cos(\theta)
\end{bmatrix} \)
- \( \det(R) = 1 \)
- \( R^{-1} = R^T \)
Global transformations

- **Reflection**
- **Horizontal (Y-axis)**

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix} r \cos(\pi - \alpha) \\
r \sin(\pi - \alpha) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\
0 & 1 \end{bmatrix} \begin{bmatrix} x \\
y \end{bmatrix}
\]

- **Vertical (X-axis)**

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix} r \cos\left(\frac{3}{2}\pi + \alpha\right) \\
r \sin\left(\frac{3}{2}\pi + \alpha\right) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\
0 & -1 \end{bmatrix} \begin{bmatrix} x \\
y \end{bmatrix}
\]

- \(\det(R) = -1\)
- \(R^{-1} = R^T\)
Global transformations

- **Shear** - Transvection in French

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix} 1 & \lambda_x \\ \lambda_y & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]
2D Linear transformations

\[
\begin{bmatrix}
u \\ v
\end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]

- Linear transformations are combinations of:
  - scaling
  - rotation
  - reflection
  - shear

- Properties of linear transformations:
  - origin is always transformed to origin
  - parallel lines remain parallel
  - ratios are preserved
  - lines remain lines
It does not have a fixed point $\rightarrow$ no matrix multiplication

\[
\begin{bmatrix}
  u \\
  v
\end{bmatrix} = \begin{bmatrix}
  x \\
  y
\end{bmatrix} + \begin{bmatrix}
  t_x \\
  t_y
\end{bmatrix} = \begin{bmatrix}
  x + t_x \\
  y + t_y
\end{bmatrix}
\]
Homogeneous coordinates - 2D affine transformation

- Instead than 2D matrices we use 3D matrices.

- **Affine transformation**: combination of linear transformations and translations

- Let \( \mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}^T \) and \( \mathbf{p}' = \begin{bmatrix} u \\ v \end{bmatrix}^T \), we obtain \( \mathbf{p}' = \mathbf{Tp} \)

\[
\begin{bmatrix}
  u \\
  v \\
  1
\end{bmatrix} =
\begin{bmatrix}
  a & b & t_x \\
  c & d & t_y \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix} \\
\mathbf{T}
\]

- Properties of affine transformations:
  - origin is *not* always transformed to origin
  - parallel lines remain parallel
  - ratios are preserved
  - lines remain lines
2D Projective transformations (homographies)

\[
\begin{bmatrix}
  u \\
  v \\
  1
\end{bmatrix} = \begin{bmatrix}
  a & b & g \\
  c & d & h \\
  e & f & i
\end{bmatrix} \begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
\]

- Properties of projective transformations:
  - origin is \textit{not} always transformed to origin
  - parallel lines do \textit{not} necessarily remain parallel
  - ratios are \textit{not} preserved
  - lines remain lines
Examples

A square transforms to:

Projective 8dof
\[
\begin{bmatrix}
h_{11} & h_{12} & h_{13} \\
h_{21} & h_{22} & h_{23} \\
h_{31} & h_{32} & h_{33}
\end{bmatrix}
\]

Affine 6dof
\[
\begin{bmatrix}
a_{11} & a_{12} & t_x \\
a_{21} & a_{22} & t_y \\
0 & 0 & 1
\end{bmatrix}
\]

Similarity 4dof
\[
\begin{bmatrix}
sr_{11} & sr_{12} & t_x \\
sr_{21} & sr_{22} & t_y \\
0 & 0 & 1
\end{bmatrix}
\]

Euclidean 3dof
\[
\begin{bmatrix}
r_{11} & r_{12} & t_x \\
r_{21} & r_{22} & t_y \\
0 & 0 & 1
\end{bmatrix}
\]
Summary

1. Introduction

2. Geometric global transformations

3. Image warping and interpolations

4. Intensity based registration

5. Landmark based registration
   - Affine registration
   - Procrustes superimposition
   - Non-linear registration (small displacement)
   - Non-linear registration (diffeomorphism)
Forward warping

\[ T_\phi(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \]

Ideally, if \( \Omega_I \) and \( \Omega_J \) were continuous domains, we would simply map \((x, y)\) to \((u, v) = T_\phi(x, y)\) and then compare \(I(x, y)\) with \(J(u, v)\)

However, \( \Omega_I \) and \( \Omega_J \) are regular grids! What if \(T_\phi(x, y)\) is not on the grid \(\Omega_J\)? \(\rightarrow\) Splatting, add contribution to neighbor pixels.
Inverse warping

- Find the pixel intensities for the deformed image $I_T$ starting from $\Omega_J$: 
  $$(x, y) = T^{-1}_\phi(u, v)$$
- Assign to $I_T(u, v)$ the pixel intensity in $I(x, y)$
- What if $T^{-1}_\phi(u, v)$ is not on $\Omega_I$? $\rightarrow$ Interpolation!
**Goal**: estimate the intensity value on points not located onto the regular grids
- nearest neighbor
- bilinear
- cubic
- lanczos
- ...

![Diagram of interpolation](image)
**Interpolation**

- **Nearest neighbor**: \( J(u, v) = I(\text{round}(x), \text{round}(y)) \)

- **Bilinear**: \( \Delta y \)

\[
\frac{f(x, q) - f(x, y)}{\Delta y} = \frac{f(x, y+1) - f(x, q)}{1 - \Delta y} \quad \rightarrow \quad f(x, q) = (1 - \Delta y)f(x, y) + f(x, y + 1)\Delta y. \\
\text{Similarly,} \quad f(x + 1, q) = (1 - \Delta y)f(x + 1, y) + f(x + 1, y + 1)\Delta y. \\
\text{Then,} \quad f(p, q) = f(x, q)(1 - \Delta x) + f(x + 1, q)\Delta x
\]
Interpolation examples

Every time we deform an image, we need to interpolate it. For instance, this is the result on Lena after 10 rotations of 36 degrees:

![Interpolation Examples](http://bigwww.epfl.ch/demo/jaffine/index.html)

**Figure 1:** Original image - Nearest Neighbour - Bilinear

Source: [http://bigwww.epfl.ch/demo/jaffine/index.html](http://bigwww.epfl.ch/demo/jaffine/index.html) (Michael Unser)
Recipe:

- Given a source image $I$ and a global affine transformation $T$, compute the forward warping of the extremities of $I$. This gives the bounding box of the deformed image $I_T$
- Given the bounding box of $I_T$, create a new grid within it with, for instance, the same characteristics of $\Omega_J$ to allow comparison between $I_T$ and $J$
- Use the inverse warping and interpolation to compute the intensity values at the grid points of the warped image $I_T$ (we avoid holes)
- Be careful! During the inverse warping, points that are mapped outside $\Omega_I$ are rejected.
Given a transformation $T$ defined by a set of parameter $\theta$ and two images $I$ and $J$, how do we estimate $\theta$?

By minimizing a cost function:

$$\theta^* = \arg\min_{\theta} d(I_T, J) \quad (3)$$

The similarity measure $d$ might be based on the pixel intensities and/or on corresponding geometric objects such as control points (i.e. landmarks), curves or surfaces.
Given a transformation $\mathbf{T}$ defined by a set of parameter $\theta$ and two images $I$ and $J$, how do we estimate $\theta$?

By minimizing a cost function:

$$\theta^* = \arg \min_{\theta} d(I_T, J)$$

The similarity measure $d$ might be based on the pixel intensities and/or on corresponding geometric objects such as control points (i.e. landmarks), curves or surfaces.
Summary

1. Introduction
2. Geometric global transformations
3. Image warping and interpolations
4. Intensity based registration
5. Landmark based registration
   - Affine registration
   - Procrustes superimposition
   - Non-linear registration (small displacement)
   - Non-linear registration (diffeomorphism)
Intensity based registration

**Same modality**

- Sum of squared intensity differences (SSD). Best measure when $I$ and $J$ only differ by Gaussian noise. Very sensitive to “outliers” pixels, namely pixels whose intensity difference is very large compared to others.

$$d(I_T, J) = \sum_u \sum_v (J(u, v) - I(T^{-1}_\phi(u, v)))^2 = (J(u, v) - I_T(u, v))^2$$

(4)

- Correlation coefficient. Assumption is that there is a linear relationship between the intensity of the images

$$d(I_T, J) = \frac{\sum_u \sum_v (J(u, v) - \bar{J})(I_T(u, v) - \bar{I}_T)}{\sqrt{\sum_u \sum_v (J(u, v) - \bar{J})^2 \sum_u \sum_v (I_T(u, v) - \bar{I}_T)^2}}$$

(5)
Intensity based registration

Multi modality

- **Mutual information**. We first need to define the joint histogram between $I_T$ and $J$. The value at location $(a, b)$ is equal to the number of intensity values that have intensity $a$ in $I_T(u,v)$ and intensity $b$ in $J(u,v)$. For example, a joint histogram which has the value of 2 in the position (7,3) means that we have found two locations $(u,v)$ where the intensity of the first image was 7 ($I_T(u,v) = 7$) and the intensity of the second was 3 ($J(u,v) = 3$). By dividing by the total number of pixels, we obtain a joint probability density function (pdf) $p_{I_T,J}$.

The definition of joint entropy is:

$$H(I_T, J) = - \sum_a \sum_b p_{I_T,J}(a,b) \log(p_{I_T,J}(a,b))$$

where $a$ and $b$ are defined within the range of intensities in $I_T$ and $J$ respectively.
Intensity based registration

Figure 2: Joint histogram of a) same modality (IRM) b) different modality (MR-CT) c) different modality (MR-PET). First column, images are aligned. 2nd and 3rd columns images are translated. Taken from [2].
Intensity based registration

- The definition of Mutual information is:

\[ M(I_T, J) = H(I_T) + H(J) - H(I_T, J) \]  \( (7) \)

- where \( H(I_T) = -\sum_a p_{I_T}(a) \log(p_{I_T}(a)) \) and
  \( H(J) = -\sum_b p_J(b) \log(p_J(b)) \). The two marginal pdf \( p_{I_T} \) and \( p_J \) are simply the accumulations over the columns and rows of \( p_{I_T, J} \) respectively. It results:

\[ M(I_T, J) = \sum_a \sum_b p_{I_T, J}(a, b) \log \frac{p_{I_T, J}(a, b)}{p_{I_T, J}(a)p_J(b)} \]  \( (8) \)
Intensity based registration

\[
\]  

(9)

- where the conditional entropy
  \[
  H(J|I_T) = - \sum_a \sum_b p_{I_T,J}(a, b) \log p_{J|I_T}(b|a).
  \]

- Mutual information measures the amount of uncertainty about \( J \) minus the uncertainty about \( J \) when \( I_T \) is known, that is to say, how much we reduce the uncertainty about \( J \) after observing \( I_T \). It is maximized when the two images are aligned. [8]

- Maximizing \( M \) means finding a transformations \( T \) that makes \( I_T \) the best predictor for \( J \). Or, equivalently, knowing the intensity \( I_T(u, v) \) allows us to perfectly predict \( J(u, v) \).
Pixel-based similarity measures need an **iterative** approach where an initial estimate of the transformation is gradually refined using the gradient and, depending on the method, also the Hessian of the similarity measure with respect to the parameters $\theta$.

Possible algorithms: gradient descent, Newton-Raphson, Levenberg-Marquardt.

Problem of the “local minima” → stochastic optimization, line search, trust region, multi-resolution (first low resolution and then higher resolution).

**Validation** → visual inspection, alignment of manually segmented objects, value of similarity measure.
Summary

1. Introduction

2. Geometric global transformations

3. Image warping and interpolations

4. Intensity based registration

5. Landmark based registration
   - Affine registration
   - Procrustes superimposition
   - Non-linear registration (small displacement)
   - Non-linear registration (diffeomorphism)
Anatomical landmarks

Definition: anatomical landmark

An anatomical landmark is a point precisely defined onto an anatomical structure which establishes a correspondence among the population of homologous anatomical objects.

Figure 3: Example of manually labeled landmarks
Anatomical landmarks

• Given a set of $N$ landmarks $i$ defined on $I$ and $N$ corresponding landmarks $j$ defined on $J$, we seek to minimize:

$$\theta^* = \arg \min_{\theta} \sum_{p=1}^{N} ||T(i_p) - j_p||^2$$

(10)

• where $T(i_p)$ means that we apply the deformation $T$ to the $p$-th landmark $i_p$

• We suppose that all landmarks belong to $\mathbb{R}^2$ (it would be similar for $\mathbb{R}^3$)

• The metric is the Euclidean norm (or Frobenius when using matrices)
Landmark based affine registration

- We define:

\[ T(i_p) = A_i p + t \quad \forall i \in [1, N] \quad (11) \]

- Thus, \( \theta = A, t \)

\[ A^*, t^* = \arg \min_{A, t} f(A, t) = \sum_{p=1}^{N} ||A_i p + t - j_p||^2 \quad (12) \]

- From which it results:

\[ \frac{\partial f}{\partial t} = 2 \sum_{p=1}^{N} (A_i p + t - j_p) = 0 \rightarrow t^* = \bar{j} - A\bar{i} \quad (13) \]

- We notice that if we center the data (i.e. \( \tilde{i}_p = i_p - \bar{i} \) and \( \tilde{j}_p = j_p - \bar{j} \)) then \( t^* = 0 \). The criterion thus becomes

\[ f = \sum_{p=1}^{N} ||A\tilde{i}_p - \tilde{j}_p||^2 \]
Now we differentiate wrt $A$:

$$\frac{\partial f}{\partial A} = \sum_{p=1}^{N} \frac{\partial \|A\tilde{i}_p\|^2}{\partial A} - 2 \frac{\partial \langle A\tilde{i}_p, \tilde{j}_p \rangle}{\partial A}$$

$$= 2 \sum_{p=1}^{N} A\tilde{i}_p \tilde{i}_p^T - \tilde{j}_p \tilde{j}_p^T = 2 \sum_{p=1}^{N} (A\tilde{i}_p - \tilde{j}_p)\tilde{i}_p^T = 0$$

It results:

$$A^* = \left( \sum_{p=1}^{N} \tilde{j}_p \tilde{i}_p^T \right) \left( \sum_{p=1}^{N} \tilde{i}_p \tilde{i}_p^T \right)^{-1}$$

The matrix $\left( \sum_{p=1}^{N} \tilde{i}_p \tilde{i}_p^T \right)$ is invertible if the landmarks are not all aligned on a straight line.
From a computational point of view, it is easier to use homogeneous coordinates:

\[
T^* = \arg \min_T ||xT - y||^2_F
\]  

where we define

\[
\begin{bmatrix}
 x_1 & y_1 & 1 & 0 & 0 & 0 \\
 \vdots & : & & & & \\
 x_N & y_N & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & x_1 & y_1 & 1 \\
 \vdots & : & & & & \\
 0 & 0 & 0 & x_N & y_N & 1 \\
\end{bmatrix}
\begin{bmatrix}
 a_1 \\
 a_2 \\
 t_1 \\
 a_3 \\
 a_4 \\
 t_2 \\
\end{bmatrix}
- 
\begin{bmatrix}
 u_1 \\
 u_2 \\
 v_1 \\
 v_2 \\
 \vdots \\
 v_N \\
\end{bmatrix}
\]

\[
T^* = (x^T x)^{-1} x^T y
\]
We seek to minimize:

\[(s, R, t)^\ast = \arg \min_{s, R, t} \sum_{p=1}^{N} ||sRi_p + t - j_p||_2^2\]  \hspace{1cm} (19)

where \(s\) is a uniform scaling factor (scalar) and \(R\) is a rotation matrix. The translation vector \(t\) is, as before, equal to \(t^\ast = \bar{j} - s\bar{R}\bar{i}\). Thus, by centering the data \((X_c, Y_c)\), we obtain:

\[(s, R)^\ast = \arg \min_{s, R} f(s, R) = \sum_{p=1}^{N} ||s\tilde{R}_p - \tilde{j}_p||_2^2 = ||sX_cR - Y_c||_F^2\]  \hspace{1cm} (20)

Remember that: \(\sum_{p=1}^{N} ||\tilde{i}_p - \tilde{j}_p||_2^2 = ||X_c - Y_c||_F^2\) where \(X = [\tilde{i}_1^T; \tilde{i}_2^T; \ldots; \tilde{i}_N^T]\) and \(||X||_F^2 = \text{Tr}(X^TX)\)
We minimize wrt $s$:

$$f(s, R) \propto s^2 \|X_c R\|_F^2 - 2s \langle X_c R, Y_c \rangle_F$$

$$\frac{\partial f(s, R)}{\partial s} = 2s \|X_c\|_F^2 - 2 \langle X_c R, Y_c \rangle_F$$

$$s^* = \frac{\langle X_c R, Y_c \rangle_F}{\|X_c\|_F^2}$$

where we use the fact that $R^T R = R R^T = I$. Substituting into $f$:

$$R^* = \arg \min_R f(R) = - \frac{(\langle X_c R, Y_c \rangle_F)^2}{\|X_c\|_F^2}$$

$$= \arg \max_R |\langle X_c R, Y_c \rangle_F| = |\langle R, X_c^T Y_c \rangle_F|$$

$$= \arg \max_R |\langle R, U \Sigma V^T \rangle_F| = |\langle U^T RV, \Sigma \rangle_F| = |\langle Z, \Sigma \rangle_F|$$

where we use the SVD decomposition $X_c^T Y_c = U \Sigma V^T$ and the definition of the trace $\langle X_c R, Y_c \rangle_F = \text{Tr}(R^T X_c^T Y_c) = \langle R, X_c^T Y_c \rangle_F$
Procrustes superimposition (similarity transformation)

- Notice that $Z = U^T RV$ is an orthogonal matrix since it is the product of orthogonal matrices. Thus $Z^T Z = I$ and $z_j^T z_j = 1$. It follows that $z_{ij} \leq 1$.

$$R^* = \arg \max_R |\langle Z, \Sigma \rangle_F| = |\text{Tr}(\Sigma^T Z)| =$$

$$= |\text{Tr} \left( \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \right) | = \sum_{d=1}^{2} \sigma_d z_{dd} \leq \sum_{d=1}^{2} \sigma_d$$

(23)

- The maximum is obtained when $z_{dd} = 1$ \ \forall d, which means when:

$$Z = U^T RV = I \rightarrow R^* = UV^T$$

(24)

- In order to be sure that $R$ is a rotation matrix ($\det(R) = 1$), we compute [5]:

$$R^* = U \begin{bmatrix} 1 & 0 \\ 0 & \det(UV^T) \end{bmatrix} V^T = USV^T$$

(25)
Procrustes superimposition (similarity transformation)

- To recap [5]:

\[
R^* = USV^T \\
S^* = \frac{\langle R, Y_c X_c^T \rangle_F}{\|X_c\|_F^2} = \frac{\text{Tr}(S \Sigma)}{\|X_c\|_F^2} \\
t^* = \bar{j} - \frac{1}{N} \sum_{p=1}^{N} sR_{ip}
\] (26)
Non-linear registration (small displacement)

- We define the deformation of a pixel at location \( z = (x, y) \) as:

\[
T(z) = z + v(z) \quad \text{with} \quad v(z) = \sum_{p=1}^{N} K(z, i_p) \alpha_p
\]  \hspace{1cm} (27)

- where \( K(z, i_p) \) is a kernel, for instance \( K(z, i_p) = \exp \left( - \frac{||z-i_p||^2}{\lambda^2} \right) \) and \( \alpha_p \) is a 2D vector which need to be estimated.

- The displacement at any point \( z \) depends on the displacement of the neighbor landmarks.
Non-linear registration (small displacement)

- We minimize

\[
\alpha^* = \arg \min_{\alpha} f(\alpha; \lambda) = \sum_{p=1}^{N} \| i_p + v(i_p) - j_p \|^2_F
\]

\[
= \sum_{p=1}^{N} \| i_p + (\sum_{d=1}^{N} K(i_p, i_d) \alpha_d) - j_p \|^2_F
\]

\[
= \| i + K\alpha - j \|^2_F
\]

(28)

where \( K = \begin{bmatrix}
1 & K(i_1, i_2) & \ldots & K(i_1, i_N) \\
K(i_2, i_1) & 1 & \ldots & K(i_2, i_N) \\
\vdots & \vdots & \ddots & \vdots \\
K(i_N, i_1) & K(i_N, i_2) & \ldots & 1
\end{bmatrix} \) and

\( \alpha = [\alpha_1^T; \ldots; \alpha_N^T] \)
By differentiating wrt $\alpha$:

$$\frac{\partial \|i + K\alpha - j\|_F^2}{\partial \alpha} = 2K^T(i + K\alpha - j) = 0$$

$$\alpha^* = K^{-1}(j - i) \quad (29)$$

The matrix $K$ might not always be invertible (if $\lambda$ is too big for instance). We need to regularize it. A possible solution is to use a Tikhonov matrix such as $\alpha^T K \alpha$, thus obtaining:

$$\alpha^* = \text{arg min}_{\alpha} f(\alpha; \lambda, \gamma) = \|i + K\alpha - j\|_F^2 + \gamma \alpha^T K \alpha \quad (30)$$

$$\frac{\partial f(\alpha; \lambda, \gamma)}{\partial \alpha} = 2K^T(i + K\alpha - j) + 2\gamma K\alpha = 0$$

$$\alpha^* = (K + \gamma I)^{-1}(j - i) \quad (31)$$
Non-linear registration (small displacement)

- Why only small displacement? → We could approximate the inverse of $T(z)$ as $T^{-1}(z') = z' - v(z')$ obtaining:

\[
T(T^{-1}(z')) = z' - v(z') + v(z' - v(z')) \neq z' \quad (32)
\]

- The error is small only if $v(z' - v(z')) - v(z')$ is small, which is the case only when the displacement is small!

- We might have intersections, holes or tearing in area where the displacement is large
Small-displacement registration

Figure 4: First row: forward and inverse transformation. The first one is a one-to-one mapping whereas the second one presents intersections. Second row: composition of transformations. They should be the identity transforms. Image taken from [7].
Instead than using small-displacement transforms we should use **diffeomorphisms**

A diffeomorphism is a differentiable (smooth and continuous) bijective transformation (one-to-one) with differentiable inverse (i.e. nonzero Jacobian determinant)

Using diffeomorphic transformations we can preserve the topology and spatial organization, namely no intersection, folding or shearing may occur
Figure 5: First row: forward and inverse diffeomorphic transformation (both are one-to-one). Second row: composition of forward and inverse transformations. The result is the identity transform (i.e. no deformation). Image taken from [7].
Diffeomorphism

- One of the most used algorithm in medical imaging to create diffeomorphic deformations is called LDDMM: Large Deformation Diffeomorphic Metric Mapping
- Deformations are built by integrating a time-varying vector field $v_t(x)$ over $t \in [0, 1]$ where $v_t(x)$ represents the instantaneous velocity of any point $x$ at time $t$ (and no more a displacement vector !)

$$v_t(x)$$

\[ \phi_t(x) \]
Calling $\phi_t(x)$ the position of a point at time $t$ which was located in $x$ at time $t = 0$, its evolution is given by: $\frac{\partial \phi_t(x)}{\partial t} = v_t(\phi_t(x))$ with $\phi_0(x) = x$.

Integrating $\frac{\partial \phi_t(x)}{\partial t} = v_t(\phi_t(x))$ between $t \in [0, 1]$ produces a flow of diffeomorphisms (if $v$ is square integrable). The last diffeomorphism is the one we are interested into.
Figure 6: Image taken from T. Mansi - MICCAI - 2009

Rigid alignment | Non-linear registration to the template
Diffeomorphism

http://www.deformetrica.org/
The flow of diffeomorphisms produces a dense deformation of the entire 3D space. We know how to deform every point in the space.

The last diffeomorphism is parametrised by the initial velocity $v_0$.

From a mathematical point of view, to register a source image or mesh $I$ to a target image or mesh $J$ we minimize:

$$\arg\min_{v_0} D(\phi_1(I),T) + \gamma \text{Reg}(v_0)$$

(33)

where $D$ is a data term, $\text{Reg}$ is a regularization term and $\gamma$ their trade-off. We use an optimization scheme (e.g. gradient descent) to estimate the optimal deformation parameters $v_0$. 


4. The Matrix Cookbook


