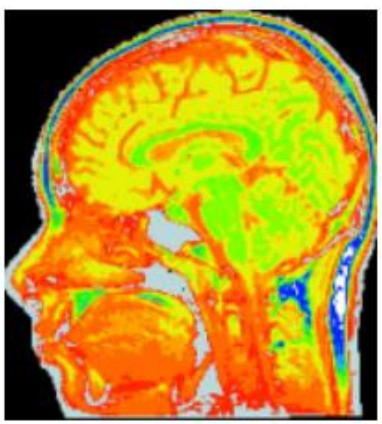
Module 4-Image Restoration

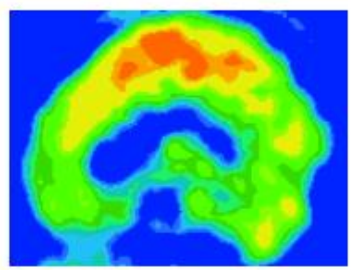
Registration

- **Spatial transform** that maps points from one image to corresponding points in another image

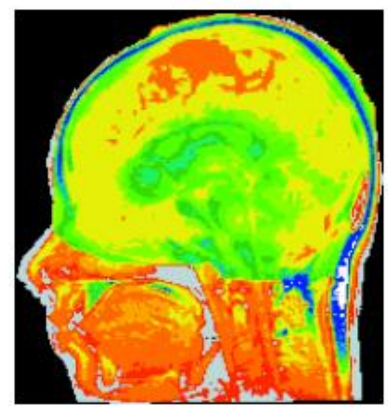
  matching two images so that corresponding coordinate points in the two images correspond to the same physical region of the scene being imaged also referred to as image fusion, superimposition, matching or merge



MR



SPECT



registered

- **Dimensionality**
  - 2D-2D, 3D-3D, 2D-3D
- **Nature of registration basis**
  - Image based
    - Extrinsic, Intrinsic
  - Non-image based
- **Nature of the transformation**
  - Rigid, Affine, Projective, Curved
- **Interaction**
  - Interactive, Semi-automatic, Automatic
- **Modalities involved**
  - Mono-modal, Multi-modal, Modality to model

- **Subject:**

  Intrasubject

  Intersubject

  Atlas
- **Domain of transformation**
  - Local, global
- **Optimization procedure**
- **Object**

Rigid Model

- A translation is a constant displacement over space. $x' = x + t_x \qquad y' = y + t_y$

  where t_x and t_y are the translation parameters in the x and y axes
- A rotation is

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

- These elementary transformations can be written in homogeneous coordinate matrix formulation as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- A rigid transformation involving a rotation followed by a translation can be defined as:

$$T_{rigid} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & t_x \\ -\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

  – The coefficients are found by solving the matrix using simple matrix inversion or Singular Value Decomposition (SVD)

- An affine transformation can be decomposed into a linear transformation and a translation.
  - An affine model maintains spatial relationships between points and asserts that lines that are parallel before transformation remain parallel after transformation .
- The general affine model has six independent parameters in 2D:

$$x' = a_{11}x + a_{12}y + t_x$$
$$y' = a_{21}x + a_{22}y + t_y$$

Affine Transformation

- Used to correct for scaling and shearing differences

  e.g. shear caused by the gantry-tilt of CT scanner

$$T_{affine} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Perspective Transformation

- The perspective model has eight independent parameters in 2D.
  - This model asserts that lines that are parallel before transformation can intersect after transformation.
  - In 2D there are two parameters that control perspective t_x and t_y:

$$T_{perspective} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ p_1 & p_2 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

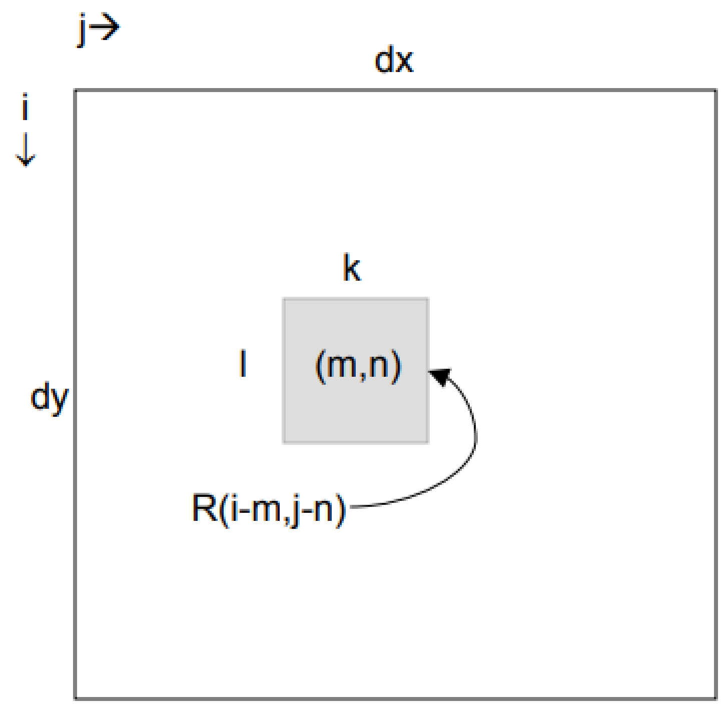
Similarity Measures

- Similarity-measures are a form of registration based on the intensity properties of the images.
  - Similarity measure use cost-functions which quantify the degree of similarity between two images.
  - Registration models based on these cost-functions simply adjust the parameters of an appropriate spatial-transformation model until the cost-function reaches a local optimum.

Algorithm: Similarity measure search strategies

1. Define the reference and search windows
2. Define a similarity measure cost-function
3. For all the positions of the reference window in the search window, calculate the similarity measure
4. Select the optimal location

- R(x,y) → reference window, C(x,y) → search window
  - $k/2<m<i-k/2$
  - $l/2<n<j-l/2$



Cross-Correlation

- Cross-correlation was one of the first cost-functions used for image registration.
  - This cost-function multiplies the intensities of the two images at each pixel and sums the resulting product.
  - The product is then normalized by dividing it by the product of the root-mean squared intensities of each of the two images.
  - Used for monomodal, temporal registration

SSD

- The sum of squared differences (SSD)
  - Minimized, SSD=0 if correctly matched
  - Used for monomodal, temporal registration

$$SSD(m,n) = \sum_{i}^{dx} \sum_{j}^{dy} \left[C(i,j) - R(i-m, j-n) \right]^2$$

Correlation Coefficient

- The correlation coefficient (CC) is an extension to correlation which assumes that the intensities in the two images are linearly related.
  - Maximized, CC=1 if correctly matched

$$CC(m,n) = \frac{1}{N} \frac{\sum_{i}^{dx} \sum_{j}^{dy} \left[C(i,j) - \bar{C} \right] \left[R(i-m, j-n) - \bar{R} \right]}{\left\{ \sum_{i}^{dx} \sum_{j}^{dy} \left[C(i,j)^2 - \bar{C} \right]^2 \cdot \sum_{i}^{dx} \sum_{j}^{dy} \left[R(i-m, j-n)^2 - \bar{R} \right] \right\}^{0.5}}$$

SAD

- The sum of absolute differences (SAD)
  - Minimized, SAD=0 if correctly matched
  - Used for monomodal, temporal registration

$$SAD(m,n)=\sum_{i}^{dx}\sum_{j}^{dy}\left|C(i,j)-R(i-m,j-n)\right|$$

  - Normalized SAD

$$SADN(m,n)=\sum_{i}^{dx}\sum_{j}^{dy}\left|\left[C(i,j)-\bar{C}\right]-\left[R(i-m,j-n)-\bar{R}\right]\right|$$

Other Similarity Measures

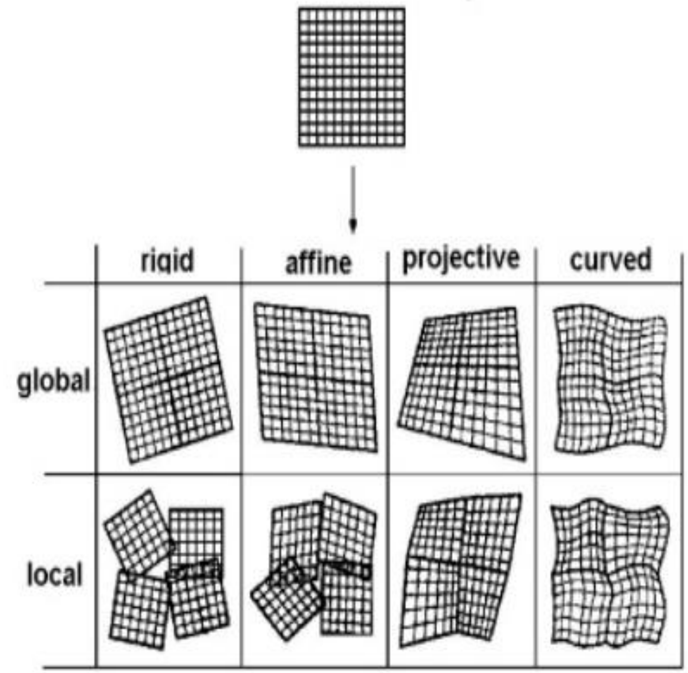
- There are various other similarity measures:
  - Variance of Intensity Ratios (VIR)
  - Variance of Differences (VOD)
    - The standard deviation of differences
  - Stochastic Sign Crossing (SSC)
    - Deterministic Sign Change (DSC)
  - Histogram of Differences based measures
    - Entropy of the Histogram of Differences (ENT)
    - Energy of the Histogram of Differences (EHD)

- **Mono-modality:**
  - ✓ A series of same modality images (CT/CT, MR/MR, Mammogram pairs,…).
  - ✓ Images may be acquired weeks or months apart; taken from different viewpoints.
  - ✓ Aligning images in order to detect subtle changes in intensity or shape

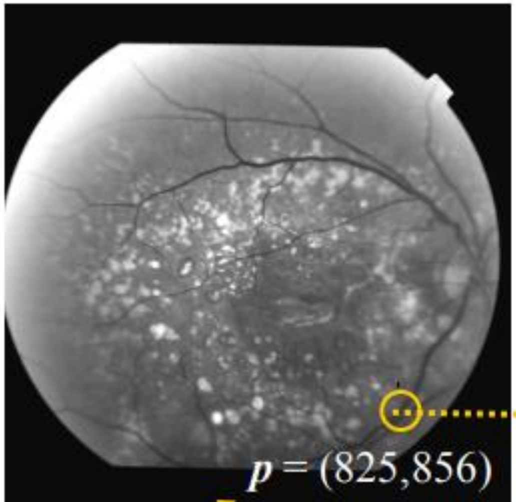
- **Multi-modality:**
  - ✓ Complementary anatomic and functional information from multiple modalities can be obtained for the precise diagnosis and treatment.
  - ✓ Examples: PET and SPECT (low resolution, functional information) need MR or CT (high resolution, anatomical information) to get structure information.

Method used to find the transformation

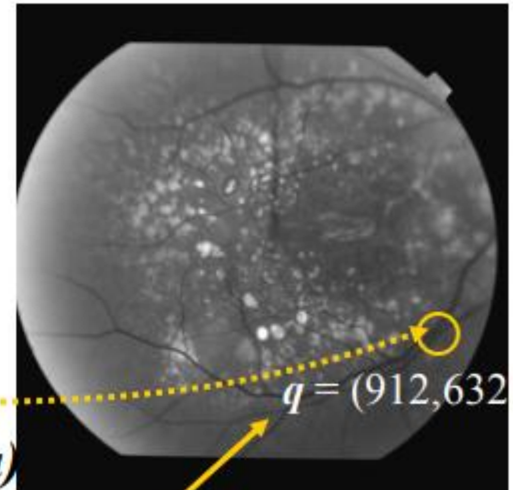
- **Rigid & affine**
  - Landmark based
  - Edge based
  - Voxel intensity based
  - Information theory based
- **Non-rigid**
  - Registration using basis functions
  - Registration using splines
  - Physics based
    - Elastic, Fluid, Optical flow, etc.



Registration is an alignment problem



$q = T(p;a)$

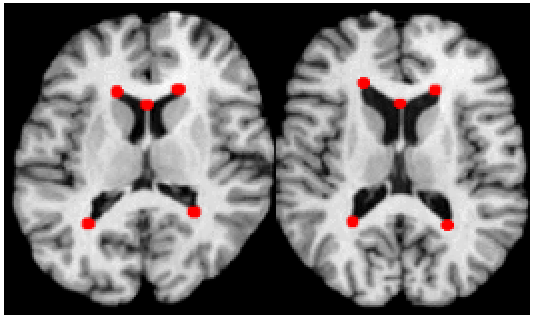


Pixel location in first image

Homologous pixel location in second image

Pixel location mapping function

Landmark Based



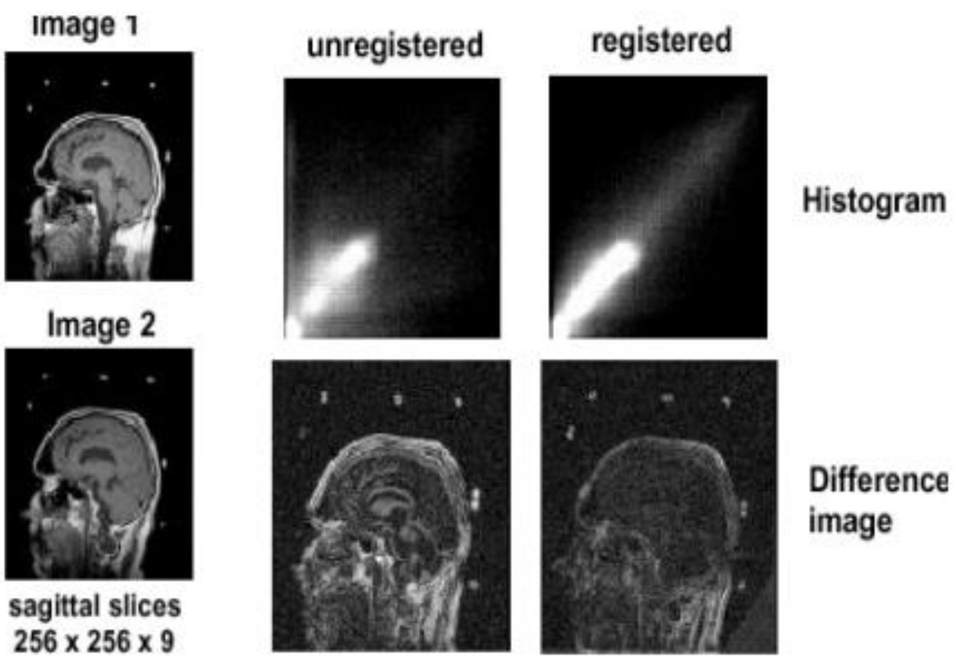
- Identifying corresponding points in the images and inferring the image transformation
- Types of landmarks
  - Extrinsic
    - artificial objects attached to the patient
  - Intrinsic
    - internal anatomical structures
- Computing the average or “centroid” of each set of points → translation
- Rotated this point set about the new centroid until the sum of the squared distances between each corresponding point pair is minimized

Surface Based

- Method
  - Extracting corresponding surfaces
  - Computing the transformation by minimizing some measure of distance between the two surfaces
- Algorithms used
  - The “Head and Hat” Algorithm
  - The Iterative Closest Point Algorithm
  - *Registration using crest lines*

Intensity Based

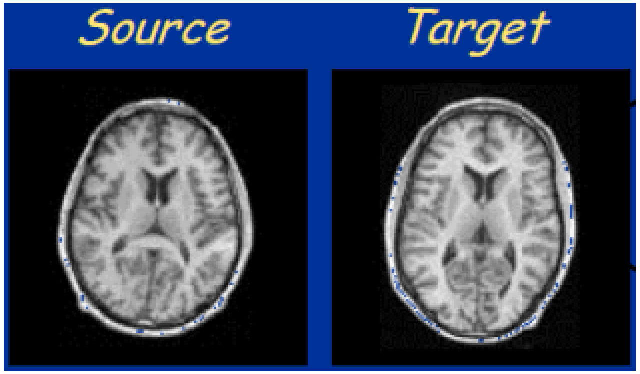
- Method
  - Calculating the registration transformation by optimizing some measure calculated directly from the voxel values in the images
- Algorithms used
  - Registration by minimizing intensity difference
  - Correlation techniques
  - Ratio image uniformity
  - Partitioned Intensity Uniformity

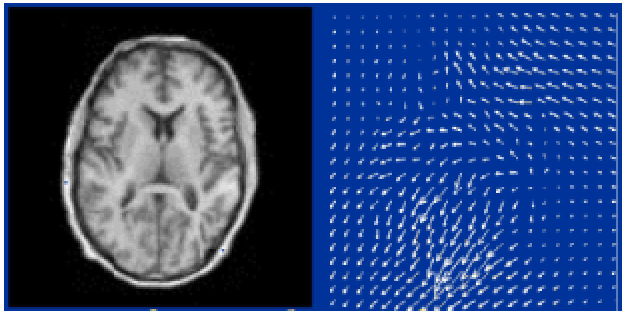


Intensity Based

- Intensity-based methods compare intensity patterns in images via correlation metrics

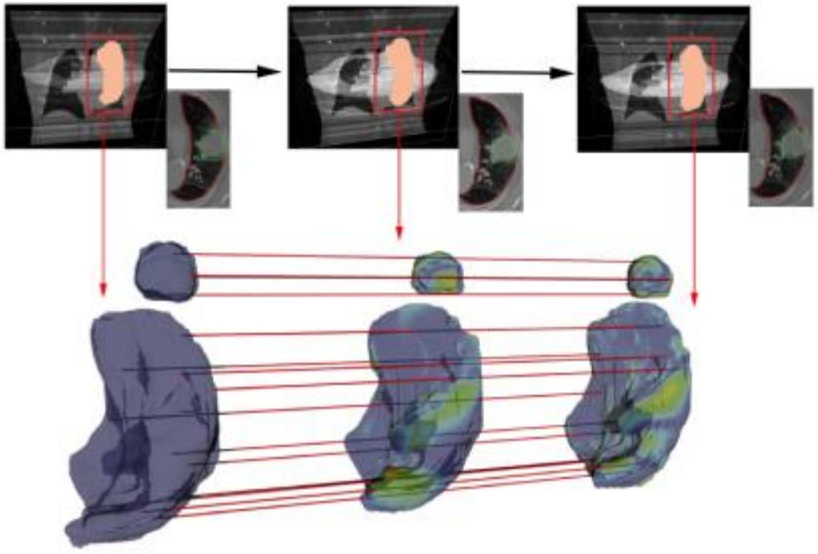
- Sum of Squared Differences
- Normalized Cross-Correlation
- Mutual Information





Feature Based

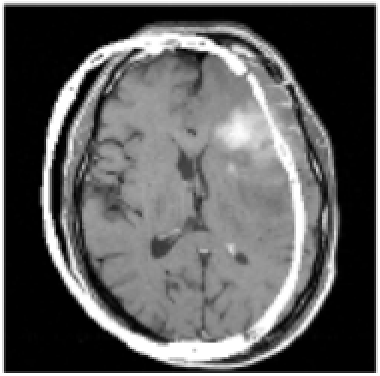
- Feature-based methods find correspondence between image features such as points, lines, and contours.
- Distance between corresponding points
- Similarity metric between feature values
  - e.g. curvature-based registration



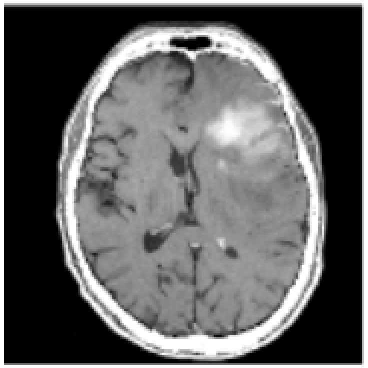
Information Theory Based

- Image registration is considered as to maximize the amount of shared information in two images
  - reducing the amount of information in the combined image
- Algorithms used
  - **Joint entropy**
    - Joint entropy measures the amount of information in the two images combined
  - **Mutual information**
    - A measure of how well one image explains the other, and is maximized at the optimal alignment
  - **Normalized Mutual Information**

Not registered



Registered



Mutual Information

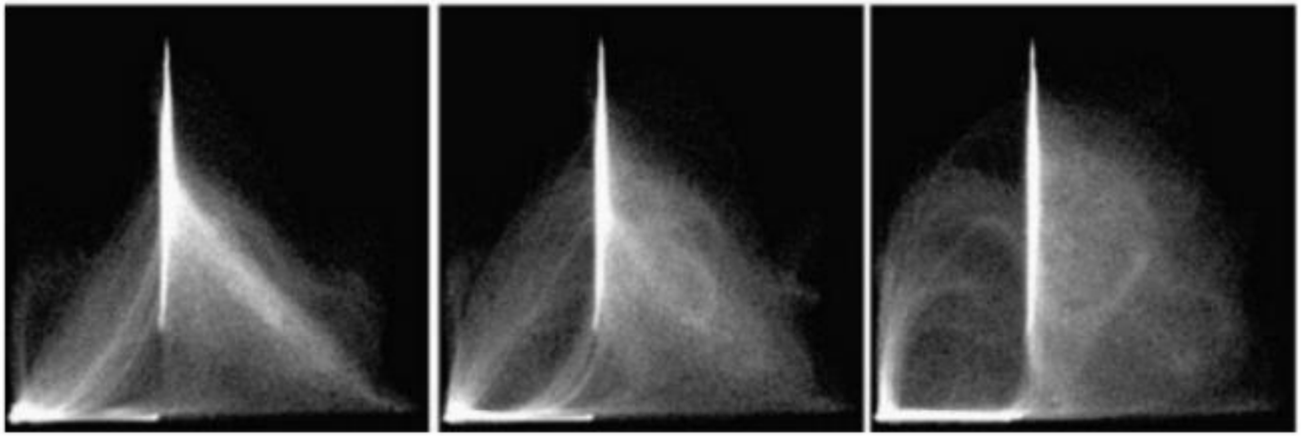
$$MI(I,J \mid T) = \sum_{i,j} p_{i,j} \log \frac{p_{i,j}}{p_i p_j}$$

Algorithms for maximising mutual information (between intensities) have been the most popular for medical image registration to date.

There are many refinements underway … not least using measurements of local phase instead of intensity*

*Mellor and Brady, Medical Image Analysis, 2004

Roger Woods' heuristic observation



Images perfectly aligned

2mm displacement of one image to the side

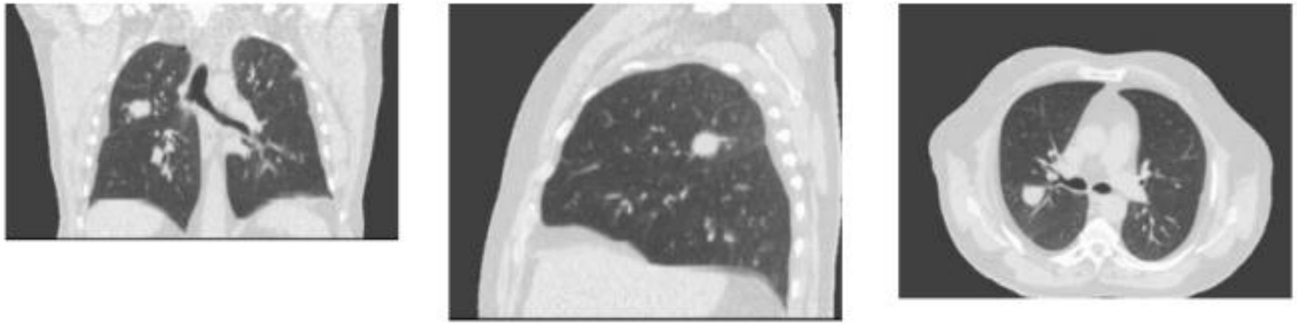
5mm displacement of one image to the side

Heuristic observation is that when the images are aligned, the joint histogram appears "sharpest" : "Woods' criterion"

Why this should be the case is still not certain!

Non-Rigid (Deformable) Registration

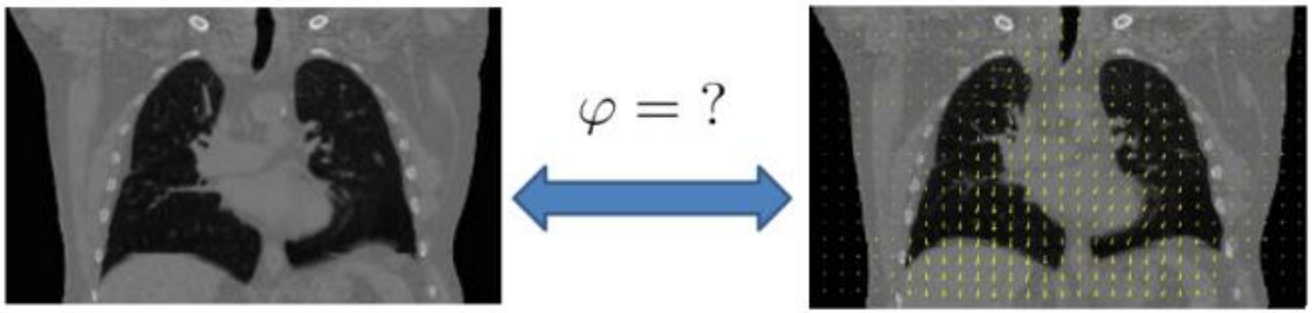
Reasons for Deformable Registration



- Patients move *(alignment of temporal series)*
- Patients change *(pre- / post-treatment images)*
- Patients differ *(creation of atlases)*

Deformable Registration

Computing a non-linear spatial transformation between ***corresponding*** structures in two images

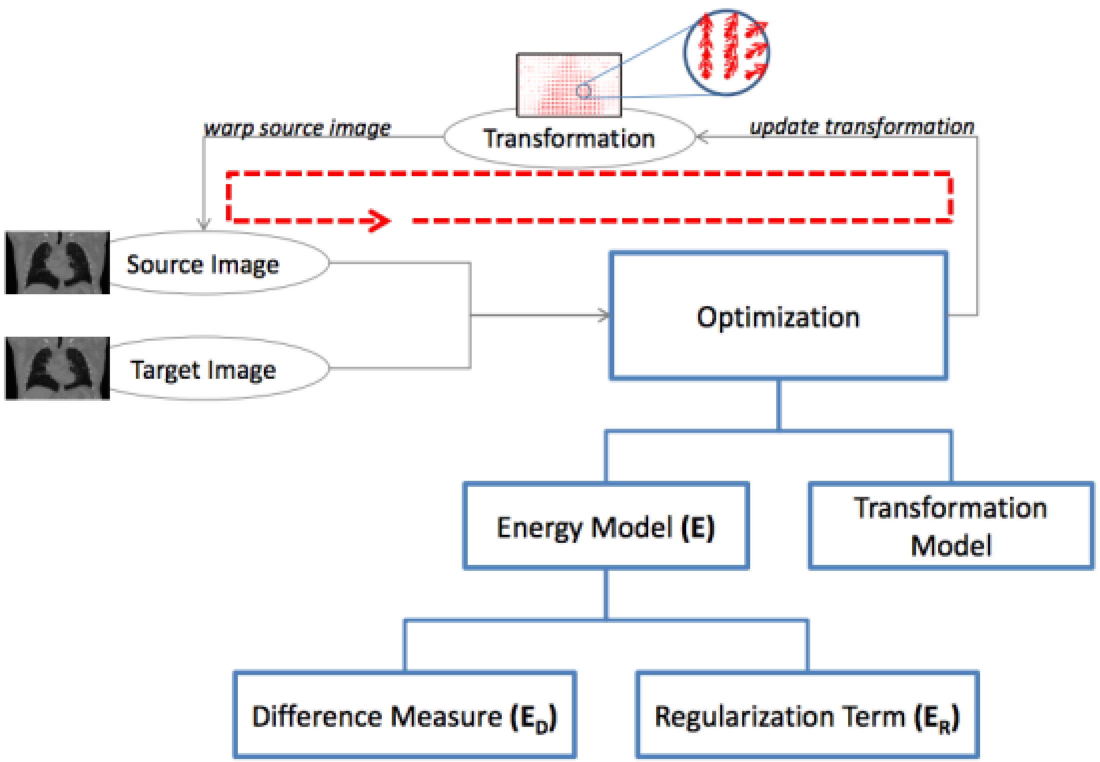


Intensity-based registration:
Minimize a difference term, based on the (pre-processed) image intensities.

No feature-based registration:

- extraction of distinct, sparsely located features
- matching of extracted features

Deformable Registration General Framework



Reference

- Dr. Ulas Bagci LECTURE 15: Medical Image Registration I (Introduction), SPRING 2016, HEC 221, Center for Research in Computer Vision (CRCV), University of Central Florida (UCF), Orlando, FL 32814.

Characteristics of Registration Methods

- Feature Space
- Similarity Metrics
- Search Strategy

Feature Spaces and Their Attributes
RAW INTENSITY - most information
EDGES - intrinsic structure, less sensitive to noise Edges [Nack 77] Contours [Medioni 84] Surfaces [Pelizzari 89]
SALIENT FEATURES - intrinsic structure, accurate positioning Points of locally maximum curvature on contour lines [Kanal 81] Centers of windows having locally maximum variances [Moravec 81] Centers of gravity of closed boundary regions [Goshtasby 86] Line intersections [Stockman 82] Fourier descriptors [Kuhl 82]
STATISTICAL FEATURES - use of all information, good for rigid transformations, assumptions concerning spatial scattering Moment invariants [Goshtasby 85] Centroid/principal axes [Rosenfeld 82]
HIGHER LEVEL FEATURES - uses relations and other higher level information, good for inexact and local matching Structural features: graphs of subpattern configurations [Mohr 90] Syntactic features: grammars composed from patterns [Bunke 90] Semantic networks: scene regions and their relations [Faugeras 81]
MATCHING AGAINST MODELS - accurate intrinsic structure, noise in one image only Anatomic atlas [Dann 89] Geographic map [Maitre 87] Object model [Terzopoulos 87]

Similarity Metric	*Advantages*
Normalized cross-correlation function [Rosenfeld 82]	accurate for white noise but not tolerant of local distortions, sharp peak in correlation space difficult to find
Correlation coefficient[Svedlow 76]	similar to above but has absolute measure
Statistical correlation and matched filters[Pratt 78]	if noise can be modeled
Phase-correlation [De Castro 87]	tolerant of frequency dependent noise
Sum of absolute differences of intensity [Barnea 72]	efficient computation, good for finding matches with no local distortions
Sum of absolute differences of contours [Barrow 77]	can be efficiently computed using "chamfer" matching, more robust against local distortions - not as sharply peaked
Contour/surface differences[Pelizzari 89]	for structural registration
Number of sign changes in pointwise intensity difference [Venot 89]	good for dissimilar images
Higher-level metrics: structural matching: tree and graph distances [Mohr 90], syntactic matching: automata [Bunke 90]	optimizes match based on features or relations of interest

Search Strategy	*Advantages and Reference Examples*
Decision Sequencing	Improved efficiency for similarity optimization for rigid transformations [Barnea 72]
Relaxation Labeling	Practical approach to find global transformations when local distortions are present, exploits spatial relations between features [Hummel 83], [Price 85], [Ranade 80], [Shapiro 90]
Dynamic Programming	Good efficiency for finding local transformations when an intrinsic ordering for matching is present [Guilloux 86], [Maitre 87], [Milios 89], [Ohta 87]
Generalized Hough Transform	For shape matching of rigidly displaced contours by mapping edge space into dual "parameter" space [Ballard 81], [Davis 82]
Linear Programming	For solving system of linear inequality constraints, used for finding rigid transformation for point matching with polygon-shaped error bounds at each point [Baird 84]
Hierarchical Techniques	Applicable to improve and speed up many different approaches by guiding search through progressively finer resolutions [Bajscy 89],[Bieszk 87],[Davis 82],[Paar 90]
Tree and Graph Matching	Uses tree/graph properties to minimize search, good for inexact and matching of higher level structures [Gmur 90],[Sanfeliu 90]

Point based methods

- If some set of corresponding point pairs can be identified a priori for a given pair of views, then registration can be effected by selecting a transformation that aligns the points.
- Because such points are taken as being reliable for the purposes of registration, they are called fiducial points, or fiducials.

- The determination of a precise point within a feature is called fiducial localization
- The transformation that aligns the corresponding fiducial points will then interpolate the mapping from these points to other points in the views.

- The fiducial localization process may be based on interactive visual identification of anatomical landmarks, such as the junction of two linear structures, e.g. the central sulcus with the midline of the brain or the intersection of a linear structure with a surface, e.g. the junction of septa in an air sinus, etc.

- the feature may be a marker attached to the anatomy and designed to be accurately localizable by means of automatic algorithms
- the chosen point will be inevitably displaced somewhat from its correct location.
- This displacement in the determination of the fiducial point, which cannot ordinarily be observed directly, is commonly called the *fiducial localization error* (**FLE**).
- Such errors will occur in both image spaces.

- Marker-based registration has the considerable advantage over landmark-based registration that the fiducial feature is independent of anatomy.
- Automatic algorithms for locating fiducial markers can take thus advantage of knowledge of the marker’s size and shape in order to produce a consistent localization point within it

- registration accuracy depends only on degree to which the chosen points correspond in the two views.
- Because it is not affected by the particular point chosen, and because the mean position relative to the marker can be expected to be the same in the two views, the effective mean displacement, <FLE> in a given view is zero

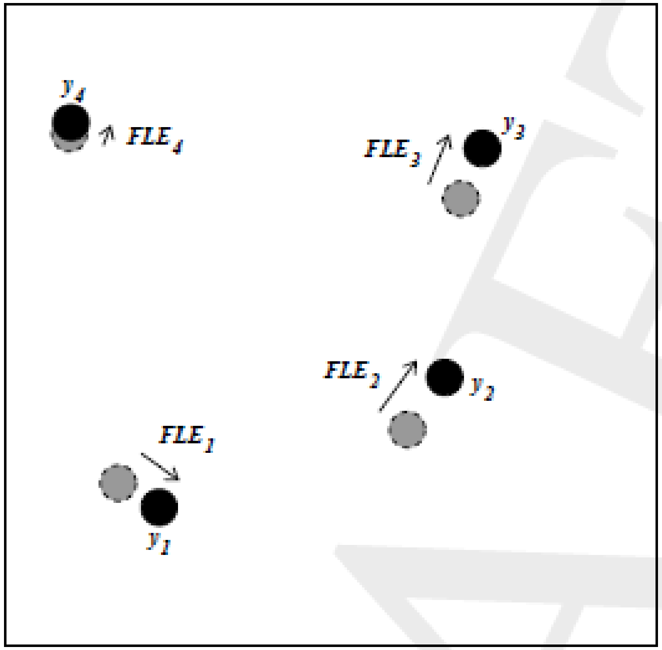
- Smaller markers, as measured in voxels, will be more poorly localized than larger ones.
- The shape of a marker that is smaller than a voxel cannot be represented at all, i.e., only one marker is bright.
- More importantly, with regard to registration accuracy, such a marker can be situated entirely within a voxel with the result that the brightness “pattern” is independent of the marker’s position within the voxel.

first order, ignoring noise and reconstruction artifacts, the function will be linear,

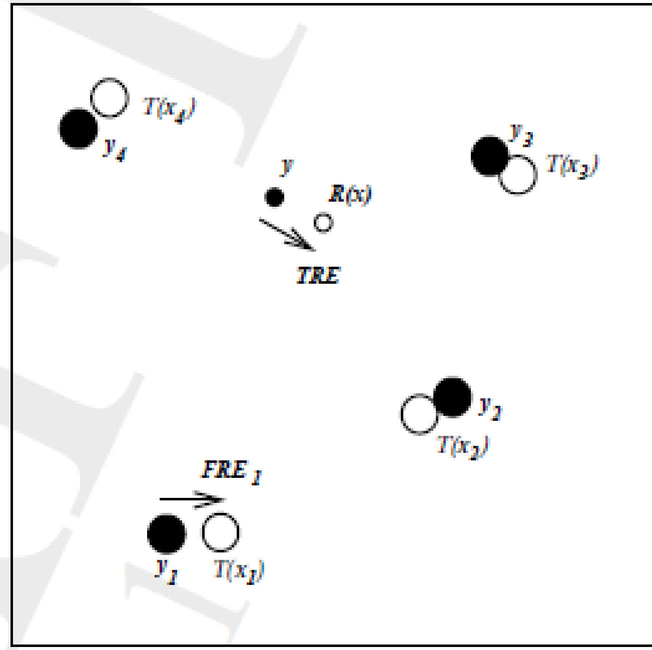
$$I = aV + I_0, \tag{8.18}$$

where I is the voxel intensity and V is the volume of intersection. I_0 is the intensity for an empty voxel, sometimes called a "background" voxel. While $I_0 = 0$ for most modalities, it is sometimes nonzero (CT, for example) and must be accounted for. If I_i is measured for one voxel i of a set of voxels that includes all those occupied by a given marker, an approximate centroid can be calculated by a weighted sum,

$$\mathbf{x} = \sum_i^n (I_i - I_0)\mathbf{x}_i / \sum_i^n (I_i - I_0). \tag{8.19}$$



(a) Schematic of point-based registration illustrating *Fiducial Localization Error* (FLE). Black circles represent positions y at which points are determined by the localization process in one of two spaces involved in the registration process. The light circles represent the actual positions.



(b) Schematic of point-based registration illustrating two measures of registration error. Black circles represent positions y in one space. The unfilled circles represent positions x in the other space after they have been mapped by the registering transformation $\mathcal{T}$. The larger, numbered circles are the points used to effect the registration. *Fiducial Registration Error* (FRE) is the alignment error between these. *Target Registration Error* (TRE) is the registration error at a point (smaller circles) not used to effect the registration.

the *fiducial registration error*, or FRE, is defined as follows. First we define an individual fiducial registration error,

$$\mathbf{FRE}_i = \mathcal{T}(\mathbf{x}_i) - \mathbf{y}_i, \tag{8.20}$$

where $\mathbf{x}_i$ and $\mathbf{y}_i$ are the corresponding fiducial points in views X and Y, respectively, belonging to feature i, as depicted in Fig. 8.5. Then we define FRE in terms of the magnitudes of the $\mathbf{FRE}_i$.

$$\mathrm{FRE}^2 = (1/N) \sum_i^N w_i^2 \mathrm{FRE}_i^2, \tag{8.21}$$

where N is the number of fiducial features used in the registration and w_i^2 is a non-negative weighting factor, which may be used to decrease the influence of less reliable fiducials. For example, if $\langle \mathrm{FLE}_i^2 \rangle$ is the expected squared fiducial localization error for fiducial i, then we may choose to set $w_i^2 = 1/\langle \mathrm{FLE}_i^2 \rangle$, where $\mathbf{FLE}_i$ is the fiducial localization error for fiducial i.

Fig. 8.5 also depicts *target registration error*, or **TRE**, which is, simply, registration error calculated at some point of interest,

$$\mathbf{TRE}(\mathbf{x}) = \mathcal{T}(\mathbf{x}) - \mathbf{y}. \tag{8.22}$$

The term "target" is meant to suggest that the point is the subject of some diagnosis or treatment.

8.3.1 *Points in rigid transformations*

If the transformation to be determined is constrained to be rigid, then, Eq. 8.21 can be written as

$$\mathrm{FRE}^2 = (1/N) \sum_{i}^{N} w_i^2 |R\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2. \tag{8.23}$$

If the $\mathbf{FLE}_i$ are random errors with zero means and isotropic distributions for all fiducials, then an optimum registration can be achieved by minimizing FRE^2 with $w_i = 1/\langle \mathrm{FLE}_i^2 \rangle$. The minimization of Eq. 8.23 is known as the "Orthogonal Procrustes" problem in the statistics literature.[2] Closed-form solutions for this problem have been available from that discipline since the first one was published by Green in 1952 [33]. The problem is also important in the theory of shape [34–36]. Algorithm 8.1 provides a simple, reliable method of solution.

Algorithm 8.1: Point-based, rigid registration

Find R and $\mathbf{t}$ to minimize $\sum_i^N w_i^2 |R\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2$:

1. Compute the weighted centroid of the fiducial configuration in each space:

$$\begin{aligned}\bar{\mathbf{x}} &= \sum_i^N w_i^2 \mathbf{x}_i / \sum_i^N w_i^2 \\ \bar{\mathbf{y}} &= \sum_i^N w_i^2 \mathbf{y}_i / \sum_i^N w_i^2.\end{aligned}$$

2. Compute the displacement from the centroid to each fiducial point in each space:

$$\begin{aligned}\tilde{\mathbf{x}}_i &= \mathbf{x}_i - \bar{\mathbf{x}} \\ \tilde{\mathbf{y}}_i &= \mathbf{y}_i - \bar{\mathbf{y}}.\end{aligned}$$

3. Compute the weighted fiducial covariance matrix:

$$H = \sum_i^N w_i^2 \tilde{\mathbf{x}}_i \tilde{\mathbf{y}}_i^t,$$

where the superscript t indicates transposition.

4. Perform singular value decomposition of H:

$$H = U \Lambda V^t,$$

where $U^t U = V^t V = I$, $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$.

5. $R = V\,\mathrm{diag}(1, 1, \det(VU))U^t$.

6. $\mathbf{t} = \bar{\mathbf{y}} - R\bar{\mathbf{x}}$.

The application of Algorithm 8.1 minimizes FRE^2, but, as indicated in Fig. 8.5, finite fiducial localization error, FLE can be expected to make both FRE and TRE nonzero. The relationships among the expected values of FLE^2, FRE^2, and TRE^2 are known to an excellent approximation for the case of equal FLE, and uniform weighting ($w_i^2 = 1$). The simplest relationship is that between the expected values of FLE and FRE,

$$\langle \mathrm{FRE}^2 \rangle \approx (1 - 2/N)\langle \mathrm{FLE}^2 \rangle, \tag{8.24}$$

fiducials' configuration. An estimate of $\langle \mathrm{FLE}^2 \rangle$ can be obtained for a given fiducial design and image acquisition protocol by performing a set of registrations involving pairs of images of possibly differing configurations and numbers of fiducials and forming the weighted average,

$$\langle \mathrm{FLE}^2 \rangle \approx (1/M) \sum_i^M N_i/(N_i - 2) \times \mathrm{FRE}_i^2, \tag{8.25}$$

where M is the number of registrations performed and N_i is the number of fiducials involved in registration i.

The relationship between $\langle \mathrm{TRE}^2 \rangle$ and $\langle \mathrm{FLE}^2 \rangle$ depends on both the configuration of markers and the target position. It is most easily stated in terms of quantities measured relative to the principal axes of the fiducial configuration:

$$\langle \mathrm{TRE}^2 \rangle \approx \frac{1}{N}\left(1 + \frac{1}{3}\sum_{k=1}^{3} \frac{d_k^2}{f_k^2}\right)\langle \mathrm{FLE}^2 \rangle, \tag{8.26}$$

where d_k is the distance of the target from principal axis k, and f_k is the RMS distance of the fiducials from the same axis [50]. This approximation, like that of

Points in scaling transformations

If the transformation to be determined includes isotropic scaling, i.e., is of the form of Eq. 8.7, then Eq. 8.21 becomes

$$\mathrm{FRE}^2 = (1/N)\sum_{i}^{N} w_i^2 |sR\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2. \tag{8.28}$$

A simple extension of Algorithm 8.1 determines the scaling s, rotation R, and translation $\mathbf{t}$ that minimize FRE^2. The extension is given in Algorithm 8.2 [54].

Algorithm 8.3: Point-based registration: Nonisotropic scaling

Find R, $\mathbf{t}$, and S to minimize $\sum_i^N w_i^2 |RS\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i|^2$:

1. Perform steps 1 and 2 of Algorithm 8.1 to determine the centroids, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$, in each space and the displacement of each fiducial point $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{y}}_i$ from the centroid in its respective space.
2. Set iterative count $n = 0$.
3. Choose an initial scaling matrix $S^{(0)}$.
4. Repeat the following steps:
   (a) Set $\tilde{\mathbf{x}}_i^{(n)} = S^{(n)} \tilde{\mathbf{x}}_i$.
   (b) Perform steps 3 through 5 of Algorithm 8.1 to find R.
   (c) Add one to n.
   (d) Determine new value of $S^{(n)}$.
5. Stop when FRE $<$ threshold or $n >$ maximum iteration count.
6. $\mathbf{t} = \bar{\mathbf{y}} - RS\bar{\mathbf{x}}$.

Points in perspective projections

The form of the transformation associated with these projections from three dimensions to two is the perspective transformation. That transformation is described in Section 8.2.2.4 in terms of the pinhole camera and is given by Eq. 8.12. Without loss of generality the problem can be treated by orienting the coordinate system so that $\hat{\mathrm{p}}$ is aligned with the z axis and by placing the origin at the pinhole. With these choices, we have $\mathbf{x} \cdot \hat{\mathbf{p}} = z$ and $\alpha = 0$. Eq. 8.12 then simplifies to

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{8.30}$$

which transforms all points to the $z' = f$ plane. Solving Eq. 8.30 for x, y, and z leads to $\mathbf{x} = \lambda \mathbf{x}'$, where λ is an arbitrary scalar. Thus, a given two-dimensional image point, x', y' lying in the $z' = f$ image plane is projected to a line in three-dimensional space that passes through the three-dimensional origin (the pinhole) and through the three-dimensional image point, x', y', f.

Surface-based methods

- Surface provides basic features for both rigid-body and nonrigid registration.
- A central and difficult question that must be addressed by any nonrigid surface-based registration algorithm is how deformation of the contents of an object is related to deformation of the surface of the object.
- Most of the surface-based registration algorithms that have been reported are concerned with rigid-body transformation, occasionally with isotropic or nonisotropic scaling.

- The approach for solving the surface-based registration problem that is frequently used and that is normally used in the medical image processing community, is to search for the transformation that minimizes some disparity function or metric between the two surfaces X and Y
- The disparity function is generally a distance.

Head and hat algorithm

- The first investigators to apply surface-based registration to a medical problem were Pelizzari, Chen, and colleagues
- They used their “head and hat” algorithm to register CT, MR, and PET images of the head.
- The “hat” is a skin surface point set {xj}
- The “head” is a polygon set model of the skin surface Y created by segmenting contours in contiguous transverse image slices

- They define yj as the intersection with head Y of a line joining through transformed “hat” point and the centroid of the head
- Transformation found using gradient descent
- The major limitations of this technique are due to the particular distance used, i.e., the distance from the surface point to the surface intersection along a line passing through the surface centroid.
- This definition of distance requires that the surface be approximately spherical.
- It also requires that a good initial transformation be supplied as input to the transformation parameter search.

8.4.3 Distance definitions

A more general definition of distance between a point and a surface is the distance between the point and the closest point on the surface. That is, the correspondence function C in Eq. 8.34 is the closest point operator, and $\mathbf{y}_j$ is the point on the surface Y closest to the transformed point $T(\mathbf{x}_j)$. The closest point and distance calculation depends on the surface representation. For example, a common representation is a triangle set. Let t be the triangle defined by the three vertices $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$. The distance between the point $\mathbf{x}_j$ and the triangle t is

$$d(\mathbf{x}_j, t) = \min_{u+v+w=1} ||u\mathbf{r}_1 + v\mathbf{r}_2 + w\mathbf{r}_3 - \mathbf{x}_j||, \tag{8.36}$$

where $u \in [0, 1]$, $v \in [0, 1]$, and $w \in [0, 1]$. The required closed-form computations are straightforward. Let $T = \{t_i\}$ for $i = 1, \ldots, N_t$ be a set of N_t triangles. The distance between the point $\mathbf{x}_j$ and the triangle set T is given by

$$d(\mathbf{x}_j, T) = \min_{i \in \{1, \ldots, N_t\}} d(\mathbf{x}_j, t_i). \tag{8.37}$$

Iterative closest point algorithm

All surface-based registration algorithms must search for the transformation $\mathcal{T}$ that minimizes the disparity function in Eq. 8.33 or a variation thereof. This is a general nonlinear minimization problem that is typically solved using one of the common gradient descent techniques (e.g., see [56]). The search will typically converge to, or very close to, the correct minimum of the disparity function minimum if the initial transformation is within about 20–30 degrees and 20–30 mm of the correct solution. To help minimize the possibility of the search getting stuck in a local minimum, many investigators perform the search in a hierarchical coarse-to-fine manner.

Algorithm 8.4: Weighted geometrical feature (WGF) rigid-body registration

Find the rigid-body transformation $\mathcal{T}$ that minimizes the disparity function in Eq. 8.40:

1. Initialization: $k = 1$, $\mathbf{x}_{ij}^{(0)} = \mathbf{x}_{ij}$, $\mathbf{x}_{ij}^{(1)} = \mathcal{T}^{(0)}(\mathbf{x}_{ij}^{(0)})$,

   where $\mathcal{T}^{(0)}$ is some initial transformation. The variable k and the superscript on $\mathbf{x}$ are iteration indices. The algorithm can be repeated using multiple initial transformations to solve the local minimum problem.

2. Iteratively apply the following steps, incrementing k after each loop, until convergence within a tolerance ϵ is achieved:

   (a) For each shape X_i, compute the closest points $\mathbf{y}_{ij}^{(k)} = \mathcal{C}_i(\mathbf{x}_{ij}^{(k)}, Y_i)$ for $j = 1, \ldots, N_{X_i}$.

   (b) Compute the transformation $\mathcal{T}^{(k)}$ between the initial point set, $\{\mathbf{x}_{ij}^{(0)}\}$, and the current set, $\{\mathbf{y}_{ij}^{(k)}\}$, using the weights $\{w_{ij}\}$. This step is effected by means of Algorithm 8.1 with the points for all shapes collected in each of the two point sets to produce two corresponding point sets.

   (c) Apply the transformation to produce registered points $\mathbf{x}_{ij}^{(k+1)} = \mathcal{T}^{(k)}(\mathbf{x}_{ij}^{(0)})$.

   (d) Terminate the iterative loop when $d(\mathcal{T}^{(k)}) - d(\mathcal{T}^{(k+1)}) < \epsilon$, where $d(\mathcal{T})$ is given by Eq. 8.40.

Intensity based methods

- SSD
- SAD
- CC