

# **Medizinische Bildverarbeitung VO**

## **Zusammenfassung**

### **residual = Residuum:**

...speziell in der Statistik die Differenz zwischen einer Regressionsgeraden und den Messwerten

...speziell in der Numerik die Größe, um die eine Gleichung nicht erfüllt ist, wenn man eine Näherung der Lösung einsetzt, im Gegensatz zum Fehler, also den Abstand der Näherung zur Lösung

### **generative model = Generatives Modell:**

...auffassen als statistische Modellierung eines unbekannten Datenerzeugungsprozesses

...Ein Generatives Modell modelliert eine vorgegebene Datenquelle durch einen stochastischen Prozess, der Datenwerte hervorbringt, die nach Struktur und Verteilung die vorgegebene Datenquelle approximieren.

...Eine Hypothese ist, dass das Gehirn sich beim Sammeln von Sinneseindrücken deren Statistik sozusagen selbst zu Eigen macht und seine internen Strukturen so anpasst, dass sie die Eigenschaften der einströmenden Reize nachahmen. Theoretische Modelle, die nach diesem Prinzip funktionieren, nennt man „generative Modelle“. Wenn sie mit Daten gefüttert werden, verändern sie in jedem Rechenschritt ihre eigenen Einstellungen so, dass sie selber den Daten immer ähnlicher werden. Die Modelle heißen „generativ“, weil man sie dazu benutzen kann, selbst Daten zu erzeugen – eine Art naturgetreue „Phantasie“-Daten.

### **Simulated Annealing:**

Die simulierte Abkühlung ist ein heuristisches Optimierungsverfahren. Das Verfahren wird zum Auffinden einer approximativen Lösung von Optimierungsproblemen eingesetzt.

Übertragen auf das Optimierungsverfahren entspricht die Temperatur einer Wahrscheinlichkeit, mit der sich ein Zwischenergebnis der Optimierung auch verschlechtern darf (je länger Algorithmus dauert, desto geringer ist die Temperatur/Wahrscheinlichkeit → Abkühlung).

Im Gegensatz zu einem Lokale-Suche-Algorithmus kann das Verfahren ein lokales Optimum wieder verlassen und ein besseres finden.

## **1.Foliensatz:**

- Rheumatoide Arthritis
  - joint space widths
  - erosions
- **PCA (Principal Component Analysis)**
  - Statistische Methode (Anisotropie bezeichnet die Richtungsabhängigkeit einer Eigenschaft oder eines Vorgangs)
  - Eigenwert: Ein Eigenvektor einer Abbildung ist in der linearen Algebra ein vom Nullvektor verschiedener Vektor, dessen Richtung durch die Abbildung nicht verändert wird. Ein Eigenvektor wird also nur gestreckt, und man bezeichnet den Streckungsfaktor als Eigenwert der Abbildung. → siehe auch Eigenwertproblem:  
<http://www.chemgapedia.de/vsengine/popup/vsc/de/glossar/e/ei/eigenwertproblem.glos.html>
  - Faktoren/Dimensionen reduzieren, Daten mit weniger Dimensionen beschreiben (aber trotzdem „ausreichend gut“ annähern)
  - Eigenvektor mit höchstem Eigenwert an erster Stelle, usw. → Die Varianz der Daten ist in diese Richtung am höchsten → Wenn Daten nur mit einer Dimension repräsentiert werden müssten, sollte man Sie mit diesem Eigenvektor/dieser Achse (mit dem höchsten Eigenwert) beschreiben
- **ASM (Active shape models)**
  - Initialisierung (mit Trainingsdaten Mean-Model erstellen, lokale Texturmodelle, etc.)
  - Shape zur Textur anpassen (neue Landmarkpositionen x)
  - Shape-Model-Instance anpassen (entsprechend den neuen Positionen)
  - Schritt 2&3 wiederholen bis Konvergenz eintritt

## 2.Foliensatz:

- **AAM (Active Appearance Models)**
  - Erweiterung der ASM-Methode
  - Shape-Model: Mean Shape Model + Modes (Eigenvektoren, die variiert plausible Abweichungen/Deformationen ergeben)
  - Texture-Model: „gesamte“ Textur der Trainingssets wird berücksichtigt
  - Verwendung eines combined-model
  - Gesamte Textur innerhalb der Landmarks wird verwendet (alle Training-Shapes werden trianguliert und auf den Mean-Shape gemapped, aus jedem Training-Example wird ein Texturvektor erstellt, alle Pixeln werden ausgelesen → viele Einträge im Vektor (bei 10x30-Bild: 300 Einträge))
  - Training notwendig (Zusammenhang: Differenzbild  $\leftrightarrow$  Korrekturparameter)
  - Es wird der residual-error (=“Restfehler“) zwischen betrachtetem Bild und generierter Modell-Instanz minimiert, um das Modell dem Bild anzupassen
- **CCA-AAM (Canonical component analysis - AAM)**
  - Mithilfe der CCA kann man Linearkombinationen zwischen zwei Variablen-Sets mit maximaler Korrelation finden. Somit kann die Anzahl der Einträge in der Regressionsmatrix verringert werden, da nur diejenigen Differenzbilder verwendet/gespeichert werden, die eine hohe Korrelation zu den Modellparametern haben (Rauschen in den Testdatensätzen wird somit ignoriert, bzw. hat nicht so einen großen Einfluss). **Bei der Suche ist die Frage: Werden Parameter abgeschätzt, oder auch ausgelesen????**
  - Das Training UND die Suche werden somit verbessert.
- **AAM-Suche:**
  1. Initialisierung (Kombinierten Eigenspace erstellen, Training: Regressionsmatrix erstellen oder CCA verwenden)
  2. Differenzbild berechnen
  3. Parameterupdate (mithilfe der Regressionsmatrix oder der CCA aus Differenzbild die Korrekturparameter auslesen/abschätzen)
  4. Modelupdate mithilfe der Korrekturparameter (Rotation, Skalierung, Translation)
  5. 2-4 Wiederholen bis die Suche konvergiert
- **AFM (Active Feature Models)**
  - Es wird nur Texturinformation in der Nähe der Landmarks verwendet
  - CCA wird beim Training/der Suche verwendet
  - ansonsten wie bei AAM

### 3.Foliensatz:

- Limitationen
  - große Datenmengen (mehrere TB) → Mühsam die richtigen/gewünschten Daten zu finden
  - genügend großes Trainingsset: es kann u.U. schwierig sein eine entsprechende Anzahl von gewünschten Bildern (spezielle Erkrankung, unverletzte anatomische Bereiche, etc.) zu finden/auszuwählen → oft nur ähnliche Pathologien vorhanden
  - Landmarks müssen gesetzt werden (annotation notwendig) und miteinander korrespondieren
- 2 Ansätze: **Model learning** (ohne/minimale Aufsicht/Überwachung), **Retrieval**
- MDL (minimum description length): Die MDL ist ein Kriterium dafür, wie effektiv ein Modell Daten kodieren kann:
  - das Modell selbst
  - die Daten kodiert durch das Modell
  - den Fehler des Modells („residual error“)  
$$\mathcal{L} = L(\mathcal{M}) + L(D|\mathcal{M}) + R$$
  - → Je kleiner  $L$ , desto besser das Modell.
- **Learn model from contours (shape):**
  - Landmark-Übereinstimmung soll gefunden werden (Wo soll Landmark auf der Kontur gesetzt werden? Welche Landmark entspricht welcher Landmark im anderen Example?)
  - Landmark-Positionen entlang kontinuierlicher Konturen anpassen → Dazu wird als Kriterium MDL verwendet. Die korrespondierenden (=übereinstimmenden) Landmarks der Examples (z.B. LM 1) auf allen Trainingsbildern entlang der Kontur so lange verschieben, bis sie bestmöglich übereinstimmen (also die DL minimiert wird).
  - Einzige Bedingung: Landmarks müssen auf der contour sein
    1. Landmarks initiieren
    2. Ein Modell der Shape-Vektoren bilden und DL berechnen
    3. DL minimieren = reparametrisieren der Kontour = Landmarkpositionen verschieben → dadurch werden Variationen die nicht notwendig sind vermieden/minimiert
  - ABER: Was ist wenn man keine annotierten kontinuierlichen Konturen hat?
- **Learn model from non-annotated data (shape and appearance):** Bilder sollen automatisch annotiert werden
  - Erfassung von Eigenschaften von Punkten (jeder Punkt des Bildes ist ein Kandidat, dessen Eigenschaften mithilfe eines Deskriptors erfasst werden)
  - Registrierung der Trainingsbilder (verschieben, rotieren, skalieren um eine grobe Deckung herzustellen falls die Trainingsbilder nicht normiert sind)
  - Interest Points in den Bildern finden: Position + Bildinformationen (lokale Textur-Eigenschaften wie gradient information, etc.) von interest-point-positions werden gespeichert

- Initiale Korrespondenzen: Paarweises Matchen der Interesting-Points + robust homography estimation (→ hierbei wird jeder Punkt auf genau einen Punkt abgebildet). Korrespondenzen (= Verbindungen von zusammengehörenden Landmarks) werden in den Interest-point-Listen als Indizes gespeichert, welche zusammen eine Matrix G ergeben.
- MDL: Optimierung der Landmark-Positionen (Kriterien: Kodierungs-Kosten für Shape, Textur und „elasticity regularization“). Optimierung durch anpassen von G:
  - Genetic Algorithm Search: Landmark-Zuordnungen werden zufällig verändert. Jene neuen Lösungen die besser bewertet werden, werden weiter „mutiert“.
  - Direct Fine Search: Ein Example herausnehmen und ins Modell projizieren, welches aus restlichen Examples gebildet wird. Mithilfe des gebildeten Modells das projizierte Example verbessern (Landmark-Zuordnungen werden gegebenenfalls verändert)
  - → Nach Konvergenz des Algorithmus müssen noch fehlende Landmarks als solche definiert/markiert werden. Dies wird dann gemacht, wenn die a-posteriori-Wahrscheinlichkeit (die Wahrscheinlichkeit, dass die betrachtete Landmark zu einer bestimmten Gruppe gehört, unter der Bedingung, dass sie dieser bestimmten „Gruppe“ zugeordnet wurde – was ja im vorangegangenen Algorithmus passiert ist) einer Landmark gering ist.
- Fehler: Konturfehler ca. 2px
- Mit dieser Methode ist es auch möglich Bewegungen nach zu verfolgen („track the movement“). Korrespondenzen werden zwischen Interesting-Points gebildet → Ein bestimmter Punkt wird durch Bildsequenz „verfolgt“. Ergebnis sind Bahnkurven/Trajektorien („trajectories“). → Damit können z.B. die Deformationseigenschaften eines Materials getestet/untersucht werden.
- **Learn the structure of models:**
  - Hier geht es darum die gefundenen Landmarks zu gruppieren (z.B. Landmarks anatomischen Strukturen zuordnen).
  - Mit Simulated Annealing die Gruppierung optimieren mit MDL als Kriterium.
  - Nach der erfolgten Gruppierung werden (wenn notwendig) neue Landmarks generiert, um die Textur besser abzudecken, bzw. die anatomische Struktur besser nachzubilden. Diese Landmarks werden dann Trianguliert um eine konvexe Hülle zu bilden.
  - Mit dieser Model-Clique-Methode wird der Rekonstruktionsfehler verringert und die erfasste Varianz erhöht (also mehr Daten werden erfasst) → Model-Clique besser/genauer!!!

## 4.Foliensatz:

- **Retrieval:** Einen ähnlichen schon bewerteten/analysierten/diagnostizierten Fall finden, anstatt den aktuellen vollkommen neu zu bewerten/analysieren/diagnostizieren.
- es werden Maße benötigt, um Bilder ranken/vergleichen zu können
- Probleme:
  - kleine Unterschiede im Bild können pathologisch großen Unterschied ausmachen
  - es besteht eine (natürliche) beträchtliche Variabilität in der Population (nicht alle Menschen sind gleich)
  - Bilder=beschränkte Domäne
  - viele verschiedene Bildgebungsmodalitäten
- Holistic (Ganzheitlicher Ansatz): Miniature Alignment
  - Entscheidend ist der „charakteristische Gesamteindruck“
  - Man hat ein Trainingsset aus dem Miniaturbilder erstellt werden; auf diese wird dann die PCA angewendet und ein kD-Baum erstellt
  - Das Query-Bild wird dann ebenfalls zu einem Miniaturbild umgewandelt
  - Die Miniaturbilder aus dem Trainingsset werden gerankt (indem das Miniatur-Query-Bild mit dem Miniatur-Trainingsset abgeglichen wird) mithilfe von:
    - Euklidische Distanz
    - Registrierung + anschließende Normierte Korrelation
  - 3D-Daten werden als Slices gespeichert – Ranking mithilfe von:
    - Registration der Miniaturbilder geschieht paarweise (Slice mit Slice) + anschließende normierte Korrelation
    - Binäre Masken + Gaussian Smoothing + sum(abs(A-B)) → hierbei werden die Unterschiede aufsummiert
    - Binäre Masken + Gaussian Smoothing + SIFT-like (chi-squared Distance) → charakteristische Punkte werden verglichen
- Based on Vocabulary (Ansatz basierend auf Vokabular): Hier wird für eine spezielle anatomische Region ein Vokabular von Merkmalen (Textur-Vokabular) angelernt
  - Wenn kleine Variabilitäten der erlernten Merkmale (statistische Eigenschaften) auftreten, sollen diese erfasst werden
  - Basis-Merkmale aus einer großen Datenmenge erfassen (die wenig bis nicht annotiert ist)
  - Descriptor soll ein bestimmtes Vokabular kennen, mit dem er das Organ (bzw. Krankheiten des Organs) beschreiben kann → für jede Merkmals-Region wird ein Textur-Histogramm erstellt, welches die Region charakterisiert (Jede Region hat ein Histogramm mit einer charakteristischen Frequenz)
  - Query: für das Query-Bild wird für eine Region ebenfalls ein Histogramm erstellt. Durch den Vergleich mit den Histogrammen aus dem Trainingsset kann das Trainingsset gerankt werden (Diffusion distance, Chi-Squared distance)
- Grundsätzliches zu Retrieval – Finden von ähnlichen Fällen:
  - Merkmalsbeschreibung von Bildern oder Regionen in Bildern

- Distanz zwischen Merkmalen bestimmen
- Ranking der Ergebnisse basierend auf Ähnlichkeit („Similarity“)

## 5.Foliensatz:

- **CT**
  - Röntgenstrahlung → Abschwächung wird gemessen, aus Abschwächungsprofilen (jeder Punkt im Abschwächungsprofil ist das Ergebnis der Abschwächung durch das Gewebe zwischen Röntgenquelle und Detektor) wird dann Bild rekonstruiert
  - Abschwächung kann durch „Lambeert-Beers law“ beschrieben werden
  - Jedes Gewebe hat einen bestimmten Absorptionskoeffizienten
  - Houndsfield Unit (Luft: -1000, Wasser: 0), HU liegt ca. zwischen -1000 und 3000
  - CT Rekonstruktion: Aus der Radon-Transformation der Projektionen/Röntgenabsorption wird das ursprüngliche Bild mittels gefilterter Rückprojektion(„filtered backprojection“) gefunden → Ziel: Von jedem Voxel/Punkt im Gewebe einen Absorptionskoeffizienten herausfinden
  - Unterschiedliche Beams: Parallel-, Fan-, oder Cone-Beam
  - Gated CT → Bildsequenzen einer sich periodisch bewegenden Struktur (z.B. Herz)
    - 1. Periode: vom 1. Ort/Teilvolumen zu versch. Zeitpunkten Aufnahmen machen
    - 2. Periode: vom 2. Ort/Teilvolumen zu den gleichen versch. Zeitpunkten Aufnahmen machen
    - usw. (bis man das gesamte Volumen hat)
    - → dadurch erhält man letztendlich zu jedem Zeitpunkt ein Bild vom Gesamten Volumen
- Segmentation ohne Modell (ohne a priori Wissen): Active Contours, Level Sets, Graph Cuts
- **Active Contours – Snakes (explicit)**
  - Ansatz:
    - Gradient an Übergängen/Grenzen größer
    - Grenzen haben eine gewisse smoothness (=Gleichmäßigkeit)
    - → bestimmtes Kriterium formulieren → Optimierungsansatz finden  
→ Kriterium optimieren
  - Kontur wird explizit durch Punkte beschrieben, die miteinander verbunden sind (Positionen der Punkte werden mithilfe eines Kriteriums gefunden)
  - Kontur wird optimiert, indem ein bestimmtes Kriterium (=Energiefunktion) minimiert wird → Kurve kann parametrisiert (= verändern der Parameter(=Koordinaten) der Punkte = verschieben der Punkte) werden.
  - Energiefunktion: Bestimmt durch
    - Interne Energie: Form der Kurve (Elastizität (1. Ableitung der Kurve): Punkte nicht so weit auseinander → Kontur wird tendenziell kleiner, Rigidität (2. Ableitung der Kurve): nicht so starke Krümmungen)
    - Externe Energie: Zusammenspiel von Bildinhalt und Kurve → z.B. Gradienten entlang der Kurve (gesucht: hoher Gradientenwert, lokales

Maximum → Energie minimal), Wege entlang niedriger Gradientenwerte werden „bestraft“

$$E = \int_0^1 \frac{1}{2} (\alpha |\mathbf{x}'(s)|^2 + \beta |\mathbf{x}''(s)|^2) + E_{ext}(\mathbf{x}(s)) ds$$

- α, β... Gewichtungsfaktoren (zusätzlich könnte auch die externe Energie Gewichtet werden)
- Limitationen durch explizite Beschreibung:
  - endliche Anzahl von Kontrollpunkten (Auflösung ist fixiert, kann dann möglicherweise zu niedrig sein)
  - begrenzte topologische/räumliche Struktur (Anzahl der Segmente muss ebenfalls vorher festgelegt werden und ist dann nicht mehr änderbar)
  - kann für das annotieren von Trainingssets verwendet werden
  - in 3D: gleiches Kriterium, aber für eine Oberfläche
  - Wird oft auf interaktive Art verwendet
  - 3D-Beispiel Gefäße: aus CT-Daten Gefäß-Modell bilden, dann im Originalbild registrieren
    - Sensor (eine deformierbares Modell/Oberfläche, die durch die Gefäße wandert) → Explorer-Balloon minimiert Energiefunktion (Interne Energie + Externe Energie + Pressure Force)
    - Sensor wandert durch Gefäße und speichert in Regelmäßigen Abständen Mittelpunkt+Radius der Gefäße
    - Abzweigungen erkennen: bei Abzweigung ist Segment fertig und neue Segmente beginnen (Segmente werden zum Schluss zu einem Gefäßbaum zusammengefügt)
- **Active Contours – Level sets (implicit):** keine Landmarks, keine Parametrisierung (anpassen der Kontur-Punkt-Koordinaten), keine fixen Topologien(zum Beispiel die Anzahl der zusammenhängenden Gebiete)
  - eine Funktion wird auf dem ganzen Bild/Volumen definiert → level-set function Φ
  - ein level-set ist ein „Schnitt“ durch diese Funktion bei einem bestimmten Wert
  - In einem Level Set (Niveaumenge) sind alle Punkte für die Φ den gleichen Wert liefert zusammengefasst. Die Menge (engl. set) aller Pixel mit dem Wert Null nennt man Zero-Level-Set. Sie trennt das Innere und Äußere der Kontur und ist immer geschlossen. Punkte die auf der Kontur liegen haben den Wert Null. Das Innere ist negativ, bzw. das Äußere positiv.
  - dann wird die Kontur so lange verändert (der Wert des Level-Sets verändert), bis das Konvergenzkriterium erreicht wird (=das passende Level-Set gefunden wurde)
  - Dense Field Levelset Algorithmus
    1. Initialisiere geschlossene Startkontur, die den Problemraum eindeutig in konturinnere und -äußere Regionen teilt
    2. Berechne eine vorzeichenbehaftete Distanztransformation auf dem Gitter in das Innere und das Äußere der Startkontur (Ich stelle mir das so vor: die Kontur wird ins Innere/Äußere „verschoben/bewegt“)

3. Solange das Konvergenzkriterium nicht erreicht ist:

1. Berechne die Änderung des Wertes der Levelsetgleichung auf allen Gitterpunkten mit einer numerischen Näherung der Evolutionsgleichung
2. Aktualisiere das Gitter mit den neuen Werten

## Die Levelsetgleichung

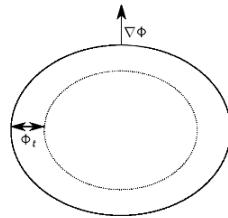
### Levelset Evolutionsgleichung

$$\Phi_t + F \cdot |\nabla \Phi| = 0$$

- $\Phi_t$  ist eine andere Schreibweise für  $\dot{\Phi}$  also die Ableitung nach der Zeit.  
Also wie hat sich  $\Phi$  vom Zeitpunkt  $t_1$  zum Zeitpunkt  $t_2$  verändert.

$$\Phi_t(x, y) = \Phi(x, y, t_2) - \Phi(x, y, t_1)$$

- $F$  ist eine Geschwindigkeitsfunktion
- $\nabla \Phi = \begin{pmatrix} \frac{\partial \Phi}{\partial x} \\ \frac{\partial \Phi}{\partial y} \end{pmatrix}$  ist der Gradient von  $\Phi$   
(Normale auf der Kurve im Punkt  $p_{xy}$ )



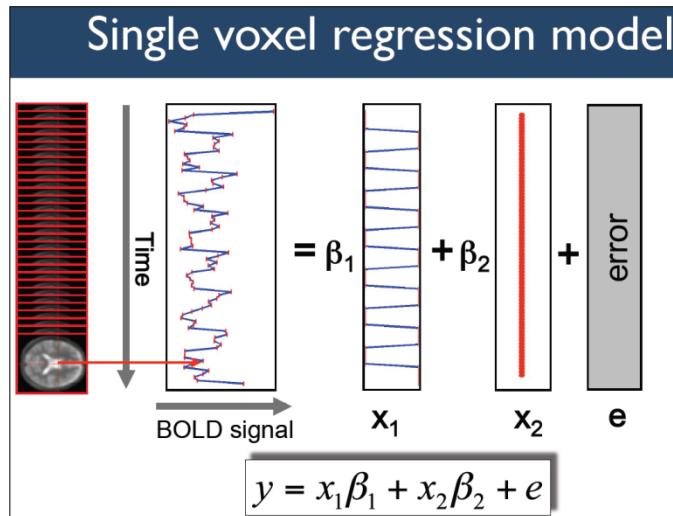
## **6.Foliensatz:**

- MRI: viel Noise, Auflösung steigt mit Tesla-Wert, gut für Weichgewebe
  - Wasserstoffatome (Protonen) haben einen Spin (magnetischer Moment) → richten sich am äußeren magnetischen Feld aus.
  - Störfeld wird appliziert: Radio frequency impulse (muss der Eigenfrequenz der Atome entsprechen: Resonanzfrequenz) lenkt Protonen ab (diese absorbieren die zugeführte Energie, werden abgelenkt und sind dann nicht mehr am äußeren Feld ausgerichtet)
  - Störfeld wird wieder abgeschalten: Protonen richten sich wieder am äußeren magnetischen Feld aus → die Zeit für dieses „Wiederausrichten“ kann gemessen werden: Relaxationszeit (T1: spin-lattice, T2: spin-spin)
  - Bei der Relaxation „senden“ die Protonen Energie (sie fallen auf ein niedrigeres Energieniveau zurück) → diese Energie kann gemessen werden
  - Da die Relaxationszeit von der chemischen Umgebung (also von der Art des Gewebes) abhängt, wird zum selben Zeitpunkt in unterschiedlichen Geweben ein unterschiedliches Signal gemessen → Kontrast
  - spatial localization (durch Gradientenfelder kann die Larmor-Frequency in einer bestimmten slice abgeändert werden → nur diese slice wird erregt)
- **Markov Random Fields (MRF)**
  - Bild/Problem mit Graphen darstellen
  - Knoten mit mehreren möglichen Labels
  - Ziel ist es die optimale Zuordnung von Labels (Bezeichnungen wie z.B. Hintergrund, Objekt A) zu den Knoten zu erhalten → dazu Kostenfunktion minimieren (Kosten minimieren: Kosten per Knoten & per Kante(zwischen 2 Knoten) minimieren)
- **Segmentation** by Graph Cuts: Semi-Automatisch (Benutzer definiert Seed-Points, die er der Quelle oder der Senke zuweist)
  - Pixel- oder Voxel-Grid als Graph interpretieren, jeder Pixel/Voxel entspricht einem Knoten, Verbindungen zwischen Pixeln/Voxeln sind Kanten
  - Zuweisen von Kantengewichten (abhängig vom Bildinhalt, z.B. Gradienten → identische Werte von Nachbarknoten=hohes Kantengewicht) und Knotengewichten (basierend auf Appearance oder manuellen Input) → Somit wird die Energiefunktion in den Graph „eingebettet“
  - Shortest path approach (verbinden von Punkten) VS. graph cuts approach (trennen/separieren von Regionen)
    - On 2D grids: graph cuts and shortest paths give optimal 1D contours
    - On N-D grids: Shortest paths still give optimal 1-D contours, Min-cuts give optimal hyper-surfaces (graph cuts trennen also auch im höherdimensionalen das Bild in Segmente)
  - Es wird nach dem billigsten (Kostenfunktion minimal) Schnitt/cut durch den Graphen gesucht, der Saatpunkte trennt (In Folien sind es 2 Terminalknoten → Quelle/Objekt und Senke/Hintergrund)
  - Dieser minimale cut definiert eine Segmentierung des Bildes/Volumens

- st-mincut wird mithilfe des maximalen Flusses berechnet (min-cut =max-flow)  
→ Der maximale Fluss (von Quelle zu Senke) ist gleich der Kapazität entlang des minimalen Schnittes
  - Globales Minimum wird gesucht/gefunden → relativ geringe Noise-Empfindlichkeit
  - ERWEITERUNG, da der minimale cut geometrisch „ungünstig“ sein kann: mitberücksichtigen der geometrischen Länge → “geodesic contour”: die Kostenfunktion ist jetzt nicht nur abhängig von den Kantengewichten, sondern auch von der Länge (in 2D), bzw. der Fläche (in 3D)
  - MULTI-WAY-GRAFH-CUTS: Hier hat man mehrere Saatpunkte, die man voneinander trennen will. Saatpunkte die ein Objekt ergeben sollen, sind mit den Terminalpunkten (t-links) verbunden. Bsp: 4 Objekte + Hintergrund → 5 t-links; als n-links werden die Kanten bezeichnet, die geschnitten werden
- **Localization** by Markov Random Fields: lokalisieren einer gesuchten anatomischen Struktur (auffinden der Position)
  - Knoten=landmarks, Labels=mögliche Positionen im Bild (=Interest Points, welche in einem Vorverarbeitungsschritt „herausgefunden“ werden) →lösen des MRF: jedem Knoten ein Label zuweisen (also jeder Landmark eine Position zuweisen)
  - Modell-basierter Ansatz: Extrahieren der Informationen aus einem Trainingsset (in  $n$  Trainingsbildern je  $m$  Landmarks ausgewählt):
    - Knotengewichte: basierend auf lokalem Aussehen/Erscheinen (“local appearance”) wie z.B. Textur
    - Kantengewichte: basierend auf räumlicher Beziehung/Anordnung der Knoten zueinander (Winkel,- und Längenstatistik)
  - Training des Models mithilfe eines annotierten Trainingssets

## 7.Foliensatz:

- **fMRI (functional Magnetic Resonance Imaging):** erfassen von Aktivität und der Bereich der Aktivität
  - räumliche Auflösung: ~4 mm
  - zeitliche Auflösung: ~1 sec
  - Es wird das Sauerstofflevel im Blut gemessen: Blood oxygen level dependent contrast (BOLD oder hemodynamic response), also „Kontrast abhängig vom Blutsauerstoffgehalt“
  - hemodynamic response: bei erhöhter Aktivität einer Gehirnregion besteht ein erhöhter Sauerstoffbedarf → dieser führt zu einer Dilatation der Versorgenden Gefäße (+ erhöhter Sauerstoffgehalt im Gefäß) → BOLD Signal wird gemessen, braucht 3-4 Sekunden bis es ansteigt und erreicht nach 5-7 Sekunden den Höchstwert (Hämodynamische Response-Funktion)
  - Als BOLD-Kontrast bezeichnet man in der Magnetresonanztomographie (MRT) die Abhängigkeit des (Bild-)Signals vom Sauerstoffgehalt in den roten Blutkörperchen. → unterschiedliche magnetische Eigenschaften des Hämoglobins (mit/ohne Sauerstoff)
- **General Linear Model (GLM) – univariat** (Jeder Voxel wird einzeln mit Stimulus verglichen)
  - Single-Voxel Regressionsmodell
  - MRI-Bilder werden
    1. aufeinander angepasst & normalisiert (Registrierung mit Referenzmodell → also mit einem Bild von einem Gehirn)
    2. geglättet und
    3. dann mit GLM analysiert
  - $Y = X\beta + e$ 
    - Y... Gemessene Werte (VEKTOR), BOLD Signal eines Voxels zu verschiedenen Zeitpunkten
    - X... DesignMATRIX, die den Erwartungswert von Y definiert
    - $\beta$ ... enthält Regressor/Regressoren und gibt den Einfluss von X auf die Gleichung an (wenn  $\beta=0$ : Y ist nur Rauschen, also kein Zusammenhang;  $\beta \neq 0$ : Y wird zumindest ein wenig von X beeinflusst, eventuell ein Zusammenhang) (VEKTOR)
    - e... Fehler (Unterschied zwischen Y und  $X\beta$ ) der bei optimalem  $\beta$  minimal ist (Methode der kleinsten Quadrate), Annahme: Fehler ist normalverteilt (VEKTOR)
    - Hier:



- Analyse der Frage „Ist eine Aktivierung des Gehirnbereichs(untersuchter Voxel) im Zusammenhang mit der Bedingung/untersuchten Aktivität (z.B. Fingertippen) gegeben?“
  1. Messen des BOLD Signals + Designmatrix definieren + [Faltung (HRF mit Designmatrix) und Hoch-Pass-Filtern,...]
  2. Nullhypothese: Keine Korrelation  $\rightarrow \beta_1=0$ .  **$\beta_2$  gibt den Wert der Baseline an!**
  3. optimales  $\beta_1$  wird mithilfe der „Methode der kleinsten Quadrate“ berechnet ( $e$  wird minimiert)
  4. Wenn es unwahrscheinlich ist, dass  $\beta_1=0$ : Nullhypothese verwerfen, Alternativhypothese „zusammenhängende Aktivität gegeben“ angenommen
- Faltung: Bevor die GLM angewendet werden kann, muss der erwartete Response (Designmatrix  $\rightarrow$  erwartete Werte) mit der HRF (Hämodynamischen Response Funktion  $\rightarrow$  Funktion, die allgemein den Verlauf eines Responses angibt) gefaltet werden + Hochpass-Filtrern. Warum?  $\rightarrow$  Welt ist nicht perfekt:
  - BOLD-Signal tritt verspätet auf
  - BOLD-Signal enthält ein niedrig-Frequenz-Rauschen (low-frequency noise  $\rightarrow$  low frequency drift), das bedeutet der Kontrast wird mit der Zeit immer schwächer. Wenn man das nicht mitberücksichtigt, kann ein Zusammenhang übersehen werden.
  - BOLD-Responses sind korreliert
- Probleme/Nachteile:
  - schlechte Registrierung durch Bewegung während der Bildgewinnung
  - Atmung
  - pulsierende Aterie
  - Trainingseffekte (bei den Probanden)
- Registrierung
  - Problem: hohe Signal-Noise-Rate
  - Prozesse von verschiedenen Subjekten  $\rightarrow$  um diese zu vergleichen: Reference-Framework  $\rightarrow$  „Talairach Atlas“ wird als Referenz verwendet

- Bevor Bilder analysiert werden, müssen Sie auf ein Referenz-Bild registriert werden
  - Brodmann-Areale (Einteilung der Großhirnrinde in Felder) sollen verglichen(bzw. davor aufeinander abgeglichen) werden (aber schwierig aus Anatomie abzuleiten)
- **Multi Voxel Pattern Analysis (MVPA) – multivariat** (mehrere Voxel werden untersucht – auch Beziehung untereinander – Ziel ist es eine bestimmte Aktivität zu erkennen (Finger tippen, Gesichter erkennen,...))
  - Frage: Korrelieren Voxel miteinander? So kann ein bestimmter Voxel bei einer Aktivität immer „nicht-aktiviert“ sein, vorausgesetzt ein anderer ist aktiviert.  
→ Das könnte mit GLM nicht erkannt werden, mit MVPA schon!
  - Brain-Response ist Verteilt und eventuell multivariat → deswegen multivariate Analyse notwendig → komplexer/schwieriger
  - Frage: Gibt es ein Muster(pattern) in den beobachteten Voxels (also ein sub-set der beobachteten Voxels), das den Versuchsbedingungen entspricht (Also mit dem eine bestimmte Aktivität erkannt werden kann)?
  - Wie funktioniert das ganze?
    - Man betrachtet mehrere Voxel (wichtig: relevante Voxel müssen identifiziert werden)
    - man trainiert einen Klassifikator an, der eine bestimmte Aktivität erkennen/vorausberechnen soll (ein bestimmtes sub-set steht für eine bestimmte Versuchsbedingung/Aktivität) → Klassifikator analysiert ein bestimmtes response pattern (=Reaktions-Muster: eine bestimmte Kombination von Voxel-Signalen, die ein sub-set der gesamten betrachteten Voxel darstellen)
    - jeder Klassifikator hat features (Voxel-Signale, also die relevanten Voxel mit ihren Signalen) und labels (Bezeichnung, für welche Versuchsbedingung der Voxel steht → z.B.: blau für Ruhezustand und rot für Fingertippen)
  - relevante Regionen müssen identifiziert werden, damit MVPA angewendet werden kann
- MVPA – localization (of distributed activation patterns) – **Random Forest**
  - identifizieren der Voxel, die für einen bestimmten Task(=Aufgabe) relevant sind um ihn erkennen zu können.
  - Random Forest:
    1. bilden von Entscheidungsbäumen mit zufällig gewählten Voxel
      - jedem Voxel wird ein Vorhersagewert zugeordnet (Ich glaube: Voxel mit geringem Vorhersagewert werden wieder verworfen)
      - Training: Voxel mit höchstem Vorhersagewert an der Wurzel; Voxel mit 2t-höchstem Vorhersagewert 1.Kind, usw. (Bsp. Voxel an der Wurzel „sagt“ Aktivität → weiter nach rechts, nächsten Voxel analysieren, usw. bis Ergebnis feststeht ODER Voxel an der Wurzel „sagt“ keine Aktivität → weiter nach links, nächsten Voxel analysieren, usw. bis Ergebnis feststeht;

Ergebnis ist in diesem Fall ob eine bestimmte Aktivität/Aufgabe gerade durchgeführt wird oder nicht, bzw. welche)

2. dann wird jeder Baum im Wald getestet (Wald-Fehlerquote, etc. berechnet) → wenn Wald ok, wird er verwendet
  3. Die Klassifikation, die am meisten Stimmen erhält wird gewählt (also in unserem Fall ob eine bestimmte Aktivität/Aufgabe gerade durchgeführt wird oder nicht, bzw. welche): Voting
- MVPA – relationship among stimuli → Beziehung zwischen verschiedenen Stimuli erkennen/untersuchen → schwieriger zu messen da komplexer
  - Connectivity (Konnektivität, Vernetzungsgrad, Verschaltung)
    1. strukturell: es gibt Neuronen/Axone
    2. funktional: statistische Abhängigkeiten zwischen lokalen Signalen
    3. effektiv: kausale Interaktion zwischen Neuronen
  - **Functional Networks**
    - gesucht: globales funktionales Konnektivitäts-Muster
    - z.B. Resting-State: bestimmte Netzwerke können erkannt werden, die auch bei Tasks relevant sind (bis jetzt nicht völlig aufgeklärt/verstanden), es sind auch im Ruhezustand bestimmte Gehirnbereiche/Netzwerke aktiv → diese sind konsistent zwischen Subjekten und Studien!
    - Paarweise-Korrelation zwischen Voxeln analysieren → mit Korrelationsanalyse: Einteilung des Gehirns in Cluster
    - Netzwerke als reference framework für Registrierung, da Registrierung basierend auf Anatomie nicht so gut funktioniert (z.B. unterschiedliche anatomische Struktur, anatomische Veränderung nach einer Schädigung während funktionale Struktur gleich bleibt) → dafür benötigt man ein globales Interaktionsmuster (Registrierung basierend auf Funktion)
  - **Markov Chain:** Markov-Kette ist ein stochatischer Prozess (eine Folge von Zufallsvariablen), wobei die Zukunft des Prozesses nur vom letzten Wert abhängt (und unabhängig von seiner Vergangenheit ist → "gedächtnislos");  
Wenn man einen zufälligen Spaziergang durch die Daten macht, ist es wahrscheinlicher einen näheren Datenpunkt zu passieren als einen, der weiter weg ist. Basierend darauf wird die Konnektivität zwischen 2 Datenpunkten, x und y, definiert als die Wahrscheinlichkeit beim zufälligen Spaziergang (Bildung einer Markov-Kette) in einem Schritt (oder mehreren) von x nach y zu gelangen.
  - Erstellen einer Funktionalen Karte; Transformation zu einer neuen Geometrie:
    - Um den Zusammenhang zwischen Tätigkeiten und der Aktivierung im Gehirn auch bei Variabilität der Studienpopulation bestimmen zu können, macht es Sinn auf eine alternative Geometrie auszuweichen. Dabei stehen nicht die anatomischen Bereiche im Vordergrund, sondern der funktionale Zusammenhang zwischen diesen Bereichen.
    - Mittels Markov Chains wird hier zuerst der paarweise Zusammenhang/Affinität zwischen den Punkten im Gehirn bestimmt, und die Diffusionsdistanz bestimmt (Diffusionsdistanz ist ein Maß für den Zusammenhang zwischen 2 Punkten → Diffusionsdistanz  $D_t$  zwischen 2 Punkten in einem Graph ist die

Summe der Transaktionskosten aller möglichen Pfade zwischen diesen Punkten mit Pfadlänge  $\leq t$ ).

- [http://en.wikipedia.org/wiki/Diffusion\\_map#Diffusion\\_distance](http://en.wikipedia.org/wiki/Diffusion_map#Diffusion_distance)
- Die Eigenwertzerlegung des Diffusionsoperators ermöglicht nun das Mapping der Diffusionsdistanz auf die euklidische Distanz. Die daraus entstehende funktionale Karte lässt sich nun relativ gut zwischen mehreren Subjekten/Populationen vergleichen. Hier ist die Distanz der Punkte zueinander nicht anatomisch, sondern durch deren Korrelation/den Zusammenhang bestimmt (Punkte/Voxel deren Zusammenhang größer ist, sind dann auch nahe beieinander – Punkte/Voxel mit niedrigem Zusammenhang sind weiter voneinander entfernt).
- Atlas:
  - Registrierung mit Funktionalem Atlas (in funktionaler Geometrie eng zusammenliegende Bereiche werden als Cluster definiert) liefert tendenziell besseres Ergebnis als Registrierung mit Anatomischem Atlas
    1. Zuerst Transformation der Subjekte in funktionale Geometrie
    2. Dann übereinanderlegen der Daten
    3. dann eventuell Rücktransformation in den „normalen“ Raum

## **8.Foliensatz:**

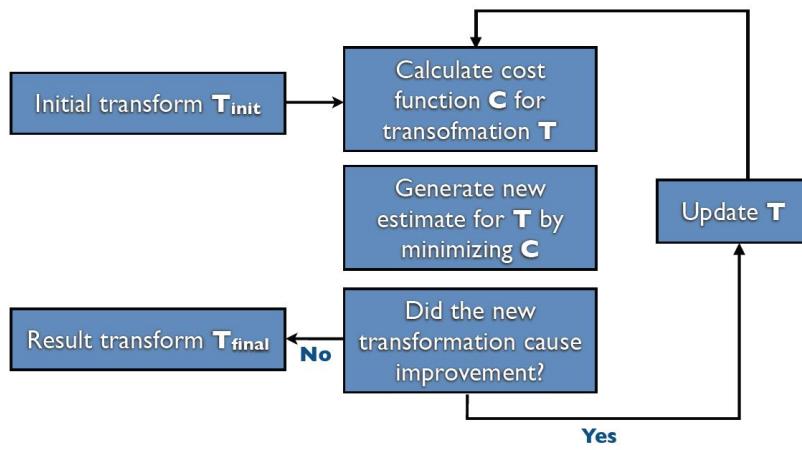
- **Anisotropic Diffusion Filtering**
  - Glätten innerhalb einer Region fördern
  - Glätten in Zwischenbereichen hemmen/verhindern
  - Basierend auf Bildinformation und Gradienten
    - unterschiedliche Glättungsstärke für unterschiedliche lokale Gradienten
  - Anwendungsbereiche:
    - Denoising (Entrauschen)
    - Segmentierung
  - Details + Formeln:
    - Bild:  $I(x, t)$
    - Diffusions-Funktion:  $c(x, t) = f(|\nabla I(x, t)|)$  → auf Folien: je größer Gradient, desto kleinerer Wert der Diffusionsfunktion (Abhängig von einer Diffusionskonstante  $\kappa$ )
    - Flow-Function:  $\Phi(x, t) = c(x, t)\nabla I(x, t)$  → Fluss hat Maximum wenn Gradientwert ähnlich der Diffusionskonstante  $\kappa$  (→kleiner Effekt bei Regionen mit kleinem/großem Gradient. Starker Effekt für Gradienten nahe zur Diffusionskonstanten  $\kappa$ .)
    - Gradient ergibt sich aus Intensitätsunterschieden zwischen benachbarten Pixeln
- **Diffusion Tensor Imaging (DTI)**
  - Diffusion von (Wasser-)Molekülen wird vom MRI-Scanner erfasst
  - Diffusion von (Wasser-)Molekülen wird dazu verwendet um Fasern zu lokalisieren
  - Diffusion ist Richtungsabhängig (=anisotrop). Die Richtungsabhängigkeit der Diffusion erlaubt Rückschlüsse auf den Verlauf der großen Nervenfaserbündel.
  - Wassermoleküle bewegen sich ständig und sind nach einer Zeit T entlang einer Achse verteilt:  $\sigma^2 = 2DT$  (wobei D der Diffusionskoeffizient ist)
  - Typischer Ablauf:
    1. Aufnehmen von Diffusion weighted images: Durch die Bewegung bzw. Diffusion der Wassermoleküle wird eine Randomisierung der Spinphase (Spin wird verschoben) verursacht und dadurch eine Abschwächung des empfangenen Signals ausgelöst (bei der Relaxation der Moleküle). Die Größe des Signalverlustes ermöglicht so eine quantitative Messung der Diffusion in der Gradientenrichtung (um den Tensor berechnen zu können werden mindestens 6 verschiedene Richtungen benötigt).
    2. Ermitteln des Diffusionstensor (Berechnen der Eigenwerte und Eigenvektoren): 3x3 Matrix, Eigenvektoren=Hauptachsen, Eigenwerte=Länge der Hauptachsen → Diffusionstensor beschreibt das dreidimensionale Diffusionsverhalten
    3. Analysieren der Anisotropie
    4. Nachverfolgen der Fasern

- **Registrierung**
  - Bei Modellbasierte-Suche wird ein Modell (trainiert und) verwendet, um eine bestimmte Struktur zu finden
  - Bei Registrierung werden 2 Bilder/Volumen gematched (durch deformieren/transformieren eines Bildes), um Übereinstimmungen/Analogien zwischen Positionen in den 2 Bildern zu finden

Registration vs. Models	
Registration	Model search
<ul style="list-style-type: none"> <li>• No training necessary</li> <li>• Very flexible</li> <li>• Slow</li> <li>• Over registration i.e., too flexible</li> <li>• Knowledge can be introduced by prototype annotation</li> </ul>	<ul style="list-style-type: none"> <li>• Training necessary</li> <li>• Fast</li> <li>• constraint by training set</li> <li>• Knowledge is encoded in the model, and can be annotated on prototype</li> </ul>

- Was brauchen wir?
  - 2 Bilder (source und target)
  - Transformation function (rigid/non-rigid, affine,...)
  - Similarity function (um source und target zu vergleichen)
- similarity-measure
  - Registrierung basierend auf geometrischen Merkmalen („**geometric features**“):
    - es müssen Merkmale (Punkte, Linien, Oberflächen) extrahiert werden  
→ abhängig von guter Merkmals-Lokalisierung
    - unabhängig von der Modalität
  - Registrierung basierend auf **Voxel-Similarity-Measure**
    - Verwendet eine Messung basierend auf Voxelintensitäten
    - Annahme, dass ein Zusammenhang zwischen den Intensitäten in beiden Bildern existiert
    - kein Extrahieren von Merkmalen
    - es müssen die Modalitäten berücksichtigt werden
  - Optimierung der (voxelbasierten) similarity: die optimale Transformation wird schrittweise gefunden, indem iterativ die Voxel-based-similarity-measure maximiert wird

# Registration algorithm



$C$  ist eine Funktion von: Referenzbild A, Quelbild B, Transformation T (T wird auf B angewendet)

- Single-Modal (Voxel-Based-)Similarity-Measures:
  - **SSD** (Sums of squared differences)
    - single-modal
    - nimmt an, dass Intensitäten in beiden Bildern gleich sind („Identitäts-Zusammenhang“)
    - optimal wenn Unterschied in Bildern nur Gaussian-Noise ist
    - empfindlich auf Ausreißer!
  - **SAD** (Sum of absolute differences):
    - single-modal
    - nimmt an, dass Intensitäten in beiden Bildern gleich sind („Identitäts-Zusammenhang“)
    - weniger empfindlich auf Ausreißer
  - **Correlation**:
    - single-modal
    - nimmt an, dass ein linearer Zusammenhang zwischen Bildintensitäten besteht
    - empfindlich auf große Intensitätswerte
    - sinnvoll wenn Bilder mit unterschiedlichen intensity-windows aufgenommen wurden
  - **CC** (Normalized Cross Correlation)
    - single-modal
    - nimmt an, dass ein linearer Zusammenhang zwischen Bildintensitäten besteht
    - sinnvoll wenn Bilder mit unterschiedlichen intensity-windows aufgenommen wurden
    - nicht mehr empfindlich auf hohe Intensitätswerte

- Multi-modale Bildregistrierung (Registrierung von Daten, die das gleiche Objekt zeigen, aber mit unterschiedlichen Modalitäten (=Bildgebungsverfahren) aufgenommen wurden) → Bildintensitäten haben ungewissen oder statistischen Zusammenhang, Zusammenhang ist a priori nicht bekannt, Zusammenhang kann mithilfe eines „2D-Histogramms“ oder einer „co-occurrence matrix“ untersucht werden
  - Intensity Remapping (CT intensities are remapped so they match approximately MR), Gradient correlation
  - 2D-Histogramm: bei korrekter Übereinstimmung ist die Streuung (=scattering) geringer als bei einer nicht-korrekt Verrückung/Verschiebung. Punkt in Histogramm wird als  $h(a,b)$  angegeben
  - Das Bild kann als Wahrscheinlichkeitsverteilung betrachtet werden
    - Wahrscheinlichkeitsverteilung kann mit 2D-Histogramm abgeschätzt werden (Verbund-, und Randwahrscheinlichkeit)
    - Verbundwahrscheinlichkeit:  $p(a,b) = h(a,b)/N \rightarrow$  Wahrscheinlichkeit eines Voxels den Wert a im Bild A zu haben und b in B.
    - Randwahrscheinlichkeit:  $p(a) = \sum_b p(a,b) \rightarrow$  Wahrscheinlichkeit eines Voxels in Bild A den Wert a zu haben.
  - **joint entropy** (gemeinsame Entropie/Informationsgehalt)
    - beschreibt die Menge der Information in den kombinierten Bildern A & B:  $H(A,B)$
    - Wenn Bilder keinen Zusammenhang haben („schlecht übereinandergelegt sind“):  $H(A,B) = H(A) + H(B)$
    - Wenn Zusammenhang besteht:  $H(A,B) \leq H(A) + H(B)$
    - Wir wollen den Zusammenhang bei Registrierung maximal machen → joint entropy zwischen zwei Bildern minimieren
    - Problem: empfindlich bezüglich der Überdeckung (bzw. der Überdeckungsfläche)
  - **mutual information**
    - Beschreibt wie gut ein Bild vom Anderen beschrieben wird:  $I(A,B) = H(A) + H(B) - H(A,B)$
    - Für Registrierung: „mutual information“ maximieren
  - **normalized mutual information**
    - $I(A,B) = (H(A) + H(B)) / H(A,B)$
    - unabhängig bezüglich der Überdeckung (bzw. der Überdeckungsfläche) → sonst könnte Überdeckung falsch interpretiert werden
  - All similarity measures can be calculated from a 2D-histogram of the two images.
- Transformation
  - Interpolation notwendig (nearest neighbour, bilinear in 2D, trilinear in 3D)
  - Rigid (starr): Translation, Rotation
  - Affine: Translation, Rotation, Skalierung, Scherung
    - $y = Ax + b$
    - $y \dots$  Position in Quellbild

- x...Position in Zielbild
  - b...definiert Translation
  - A...definiert lineare Transformation (Rotation, Skalierung, Scherung)
  - Transformation ist bijektiv(funktioniert in Beide Richtungen), wenn Matrix A invertierbar ist
- Non-rigid: alle Transformationen möglich (auch non-linear)
  - Free form deformation (FFD)
  - Thin plate splines (TPS): mapping basierend auf Kontrollpunkten (jeder Kontrollpunkt ist gewichtet, je nach Gewicht/Wichtigkeit hat er Einfluss auf Transformation und auf andere Kontrollpunkte), Wenn Kontrollpunkte transformiert werden wird Rest auch transformiert
- Optimization
  - Optimierung der Ähnlichkeitsfunktion (z.B. Normalized Mutual Information) meist iterativ
  - globales Optimum meist nicht möglich → lokale Suche
  - **Multi-Resolution-Strategy**
    - Dabei werden unterschiedliche Auflösungen der Bilder verwendet, um nicht in lokalen Maxima gefangen zu werden
- Anwendungen:
  - Langzeitstudien: Registrierung von Bildern des gleichen Patienten mit gleicher Modalität zu verschiedenen Zeitpunkten
  - Fusion verschiedener Modalitäten
  - Atlas building

## Summary

- Image registration finds a transformation that maps one image onto another image, so that *corresponding* voxels match.
- Choice of **similarity function**
- Single-modal, multi-modal
- **Transformation:** rigid, or non-rigid
- **Optimization:** multiresolution, ...
- Applications: longitudinal studies, atlas building, atlas based segmentation, image guided surgery, disease progression quantification.

## **9.Foliensatz:**

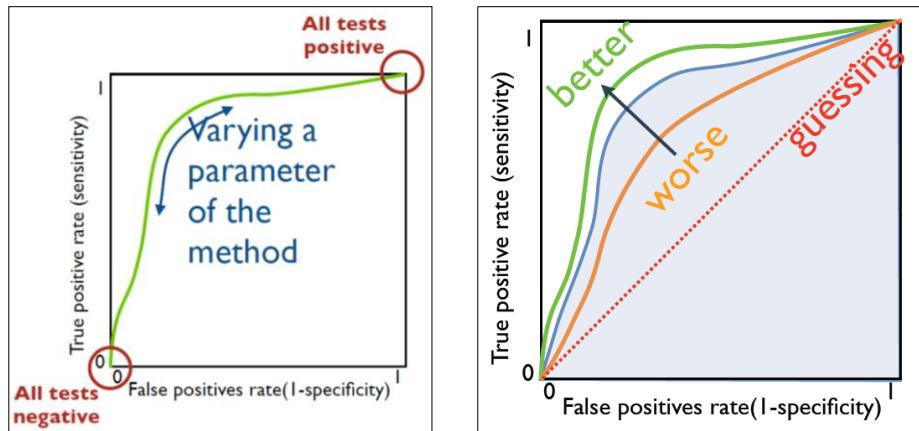
- Erkennen von Erosionen
  - Klassifizierungsaufgabe: man teilt die Punkte auf der Kontur als Erosion oder Non-Erosion ein → führt zu einer Messung des Ausmaßes der erosiven Destruktion auf der Kontur
  - GMM (Gaussian Mixture Model) → Generative Model, ist eine Summe von gewichteten Normalverteilungen. Weil wir d-dimensionale Räume betrachten, brauchen wir auch mehrdimensionale Normalverteilungen. → Modelle für gesunde und krankhafte „bone-patches“
  - Klassifizierung: A-Posteriori-Wahrscheinlichkeit: Wird bei Klassifizierung verwendet → (*Wahrscheinlichkeit für Erosion \* A-Posteriori-Wahrscheinlichkeit der GMMs (für Erosion) wenn Erosion gewählt werden würde*) wird Verglichen mit (*Wahrscheinlichkeit für Non-Erosion \* A-Posteriori-Wahrscheinlichkeit der GMMs (für Non-Erosion) wenn Non-Erosion gewählt werden würde*) → je nach dem welche Klasse höhere Wahrscheinlichkeit hat, wird der Punkt Erosion oder Non-Erosion zugewiesen
  - Klassifizierung anschaulich: Es wird mit dem Modell für gesunde Patches versucht den betrachteten Patch nachzubilden (dasselbe mit Modell für krankhafte Patches) → die Nachbildung mit geringerem Residual-Error wird gewählt (also jenes Modell, welches den betrachteten Patch besser beschreiben kann)
  - Wichtige Punkte die beachtet werden müssen:
    - Pathologien müssen fehlerfrei/genau/korrekt erkannt werden
    - es muss ein gutes Interface und eine gute Visualisierung der korrekten Ergebnisse vorhanden sein, damit der klinische Experte die Ergebnisse überprüfen und bestätigen kann
- Alzheimer und MCI („Mild Cognitive Impairment“ = leichte kognitive Beeinträchtigung)
  - Ziel: kleine Veränderungen über die Zeit am Hippocampus erkennen
  - Atlas-basierte-Segmentierung von Gehirnregionen
    - nicht einfach, da Population zu heterogen ist, als dass man Fälle „einfach so“ vergleichen kann
    - Similarity-function (Normalized Mutual Information) misst Ähnlichkeit zwischen „regions of interest“ (Bereich um Hippocampus im jeweiligen Bild) → ergibt Ähnlichkeit zwischen allen möglichen Kombinationen
- Validation Strategies
  - Notwendig um...
    - ...verschiedene Methoden zu vergleichen
    - ...Performance und Verlässlichkeit angeben zu können
    - ...den Output einer Methode mit klinisch relevanter Information verknüpfen zu können (z.B. „Patient ist nach 5 Jahren gesund“)
  - Was kann validiert werden?

- Messungen einer quantitativen Größe (z.B. joint space width)
  - Klassifizierungen (z.B.: ist im Bild ein maligner Tumor zu sehen? JA/NEIN; Klassifizierung Erosion/Non-Erosion)
  - Segmentierungen, Registrierungen (z.B. Segmentierung eines Knochens, Registrierung zweier Datensätze)
  - →Messung, Klassifizierung und Segmentierung sind eng miteinander verknüpft (z.B. joint space width-Messungen und Erosion-Detektion bei rheumatoider arthritis)
- Verschiedene Merkmale können verglichen werden:
  - Durchführungszeitdauer
  - Korrektheit des Ergebnisses
- Was kann als Referenz (mit dem das Ergebnis verglichen wird) angegeben werden?
  - Ground truth: Messung wird mit Messung an der Gewebeprobe verglichen
  - Relative ground truth: Messung wird mit Messung eines klinischen Experten (oder einer Expertengruppe mit gleicher Erfahrung) verglichen
  - Standard of reference (=Referenzstandard): Vergleich mit einer Standardmethode, die als das „momentan beste“ gilt
- Wenn keine Referenz vorhanden ist (z.B. wenn neue Methode besser als Standard of Reference ist), kann trotzdem...
  - ...die Präzision evaluiert werden, also wiederholbare Messungen bereitzustellen (→die Fähigkeit bei den gleichen Daten die gleichen Ergebnisse zu erhalten)
  - ...die Fähigkeit zur Unterscheidung evaluiert werden (zwischen unterschiedlichen Behandlungsstrategien, die nachweislich unterschiedliche Effekte haben)
- Was genau wird validiert?
  - Accuracy: Fähigkeit einer Methode Messungen oder Diagnosen des Referenzstandards widerzuspiegeln.
  - Precision: Fähigkeit einer Methode wiederholbare Messungen bereitzustellen (z.B: erhöht geringe Variabilität aufgrund von Noise die kleinste Einheit, die noch verlässlich gemessen werden kann)
  - Sensitivity: Anteil der positiven Fälle, die korrekt als positiv klassifiziert wurden
  - Specificity: Anteil der negativen Fälle, die korrekt als negativ klassifiziert wurden
- **Validating Measurements**
  - Durch Messung eines bestimmten Objektes (joint space width,...) soll etwas über den Status einer Krankheit herausgefunden werden → Annahme, dass Messung in Zusammenhang mit einer Eigenschaft der Krankheit steht. Zwei Möglichkeiten um Messungen zu beurteilen:
  - 1.Möglichkeit: Nur die Messungen selbst werden berücksichtigt (z.B. Vergleich von wiederholten Messungen):

- **Smallest detectable difference (SDD):** Der kleinste Unterschied der verlässlich detektiert werden kann
  - **Coefficient of variation (CV):** Menge der Messungs-Streuung verglichen mit der mittleren Messung →  $CV = \sigma/\mu$
  - **Relative error:**  $relative\ error = (measurement - standard\ of\ reference) / standard\ of\ reference$
- 2.Möglichkeit: Anstatt Referenz-Daten werden andere Informationen berücksichtigt (Krankheitsstatus,...) um einen Erwartungswert für die Messung zu erhalten.
  - Normalerweise steht keine „ground truth“ zur Verfügung, sondern nur ein Standard of Reference → diesem sollte man allerdings nur bis zu einem gewissen Grad vertrauen
  - Wenn auch keine s.o.r. vorhanden ist, kann PreWissen über Krankheit verwendet werden. Dabei wird die Fähigkeit der Messmethode, Daten zu differenzieren verwendet. z.B. Man hat zwei Gruppen, bei denen man weiß, dass sie einen Unterschied aufweisen, und beobachtet mit welcher Signifikanz die Messmethode zwischen den 2 Gruppen unterscheiden kann. (hier muss man sehr aufpassen!)
- **Validating Classifications (Pathologiedektion)**
  - Beurteilen von Klassifikatoren bezogen auf true/false positives/negatives:
 

		Pathology / condition	
		True	False
Test outcome	Positive	True positive (TP)	False positive (FP)
	Negative	False negative (FN)	True negative (TN)
  - **Sensitivity:**  $TP/(TP+FN)$
  - **Specificity:**  $TN/(FP+TN)$
  - **False positive rate:**  $FP / (FP + TN) = 1 - specificity$
  - **False negative rate:**  $FN / (TP + FN) = 1 - sensitivity$
  - **Diagnostic accuracy:**

$$\text{diagnostic accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$
  - Im Idealfall sind Sensitivität und Spezifität beide 1. Man muss immer beide Werte angeben, sonst sind sie (einzelne) bedeutungslos.
  - **ROC-Kurve (Receiver operator characteristic):** Ist eine Kurve um alle möglichen Abwägungen zwischen Sensitivität und Spezifität zu beurteilen. Dabei wird ein Parameter der Methode variiert, der den trade-off zwischen Sensitivität und Spezifität am besten steuert:



- **AUC (area under the curve):** Je größer die Fläche unter der Kurve ist, desto größer ist die Wahrscheinlichkeit die richtige Entscheidung zu treffen. (AUC größer  $\leftrightarrow$  Methode besser). Wenn 2 ROC-Kurven denselben AUC-Wert haben, kann man die ROC-Kurven in der klinisch relevanten Region vergleichen.

- **Validating Segmentations**

- Segmentierung
  - Bild wird in verschiedene Sub-Regionen unterteilt
  - diese Regionen überschneiden sich nicht
  - Pixel/Voxel innerhalb einer Region erfüllen ein bestimmtes Ähnlichkeitskriterium
- **Segmentation vs. S.O.R.:** 5 Zustände von Unterschieden zwischen Segmentierung (die bewertet wird) und dem Referenzstandard (=s.o.r.), ausgedrückt mit Parameter T:
  - **Correct Segmentation:** die gemeinsame Überdeckung zwischen Segmentierung und SOR erfüllt T.
  - **Over-Segmentation:** Eine Region in SOR entspricht mehreren Regionen in der Segmentierung. Jede Seg-Region hat mindestens T% Überdeckung mit SOR-Region UND SOR-Region hat mindestens T% Überdeckung mit der Seg-Regionen-Kollektion.
  - **Under-Segmentation:** Gegenteil von Over-Segmentation. Eine Region in Segmentierung überdeckt sich mit einer Kollektion/Ansammlung von SOC-Regionen.
  - **Missed Region:** Es gibt eine Region in der SOC aber keine Region in der Segmentierung, mit einer ausreichenden Übereinstimmung (T).
  - **Noise Region:** Es gibt eine Region in der Segmentierung aber keine Region in der SOC, mit einer ausreichenden Übereinstimmung (T).
- Oft ist die Segmentierung einer bestimmten Region von Interesse:
  - % der gemeinsamen Überdeckung zwischen Segmentierter Region und SOR-Region
  - **MSE (Mean square error)** of boundaries(=Grenzen/Ränder)

$$\frac{\text{Area}(A \cup M) - \text{Area}(A \cap M)}{\text{Area}(M)}$$

M...S.O.R.-Region

- **Cross-Validation** (Kreuzvalidierung: Validierung von Modellen):
  1. Daten werden in k Teilmengen aufgeteilt
  2. Daten werden iterativ in 2 Mengen aufgeteilt: (größere Menge=Trainingsmenge=besteht aus k-1 Teilmengen, kleinere Menge=Testmenge=besteht aus 1 Teilmenge)
  3. ein Teil der Daten wird zu Trainingszwecken benutzt  
(Segmentierungsverfahren wird mit Trainingsdaten trainiert) und die restlichen Daten (Testmenge) zur Validierung des Modells verwendet
  4. 2-3 wiederholen bis jede Teilmenge einmal als Testdaten verwendet wurde
  5. man benötigt ein zusätzliches test-set zur Validierung des Segmentierungsverfahrens, weil: die Parameter des Verfahrens/Algorithmus verändern sich normalerweise während der Kreuzvalidierung → validation-sets haben die Wahl der Parameter beeinflusst
    - Cross-Validation gibt trotzdem eine „gute Ahnung“ über die Performance eines Algorithmus
- Probleme/Fallen:
  - es werden nur synthetische Daten verwendet
  - durch Kenntnis des Algorithmus kann die Wahl der Daten beeinflusst werden sein → tendenziell einfache/simplere Daten
  - aufgrund von Sensoren: Artefakte in den Daten
  - Definition des Kriteriums/der Entscheidungsgrößen
  - Zuverlässigkeit des standard of reference?
  - es kann passieren, dass etwas gemessen wird, was man gar nicht messen will → Artefakte aufgrund des Experiment-Designs