

## 3.4 Registrierung

In diesem Abschnitt beschäftigen wir uns mit dem Problem der *Registrierung* von Bildern, d.h. der Aufgabe zwei Bilder derselben Szene, oder zumindest ähnlicher Szenen, bestmöglich in Übereinstimmung miteinander zu bringen. Registrierungsverfahren werden in der biomedizinischen Bildverarbeitung häufig eingesetzt. Zu registrierende Bilder unterscheiden sich dabei in der Regel dadurch, dass sie von unterschiedlichen Positionen, mit unterschiedlichen Sensoren oder zu unterschiedlichen Zeitpunkten aufgenommen wurden.

### 3.4.1 Problemstellung

Bei der Registrierung werden zwei Bilder betrachtet, von denen eines üblicherweise als *Referenzbild*

$$f_R : \Omega \rightarrow \mathbb{R}$$

und das andere als *Templatebild*

$$f_T : \Omega' \rightarrow \mathbb{R}$$

bezeichnet wird. Die Aufgabe der Registrierung besteht darin, eine *optimale* Transformation

$$y : \Omega \rightarrow \Omega'$$

zu finden, so dass das transformierte Templatebild  $f_T(y(x))$  dem Referenzbild  $f_R(x)$  möglichst *ähnlich* ist. In welcher Weise man dabei von einer optimalen Transformation und ähnlichen Bildern sprechen kann, wird auch bei dieser Bildverarbeitungsaufgabe im Wesentlichen durch a-priori Informationen und eine Modellierung z.B. mit Variationsmethoden festgelegt.

### 3.4.2 Transformationen

Verfahren zur Registrierung lassen sich u.a. klassifizieren nach der Art der zugelassenen Transformationen. Man unterscheidet zwischen parametrischen und nicht-parametrischen (freien) Transformationen. Zu den parametrischen Transformationen zählen u.a. rigide, affine und projektive Transformationen. Nach Wahl einer bestimmten Anzahl an zugehörigen Transformationsparametern  $w_1, w_2, \dots$  sind diese eindeutig bestimmt. Die Transformationsklassen sind ihrer Nennung nach jeweils in der nachfolgenden Klasse enthalten.

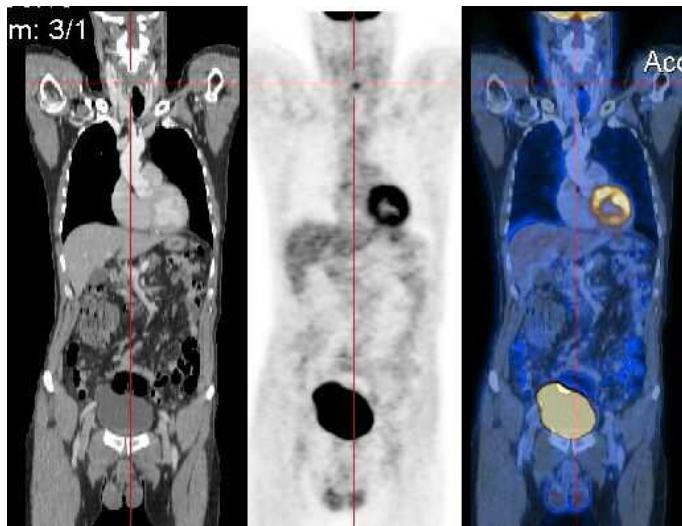


Figure 3.4: Beispiel einer PET-CT Bildregistrierung, Quelle: Wikipedia

Eine der einfachsten Transformation ist eine Translation. In 2D kann mit Hilfe des Translationsvektors  $w = (w_1, w_2) \in \mathbb{R}^2$  eine Translation geschrieben werden als

$$y(x; w) = x + Q(x)w, \quad \text{mit } Q(x) := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Bei der Klasse der rigiden Transformationen, manchmal auch starre Transformationen genannt, werden ausschließlich Translationen und Rotationen für die Koordinatentransformation angewandt. Diese Form der starren Registrierung ist sehr restriktiv, da transformierte Bildobjekte als starre Körper aufgefasst werden. Jegliche Art der Dichteänderung eines Objektes wird dabei a priori ausgeschlossen. Die Komponenten einer rigiden Transformation in 2D können mit 3 Parametern definiert werden als

$$\begin{aligned} y_1 &= \cos(w_1)x_1 - \sin(w_1)x_2 + w_2 \\ y_2 &= \sin(w_1)x_1 + \cos(w_1)x_2 + w_3. \end{aligned}$$

Eine rigide Transformation lässt sich kompakt schreiben als

$$y(x; w) = Q(x)g(w), \quad \text{mit } Q(x) := \begin{pmatrix} x_1 & x_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & 1 \end{pmatrix} \quad (3.36)$$

und

$$g(w) := (\cos w_1, -\sin w_1, w_2, \sin w_1, \cos w_1, w_3)^T.$$

Man beachte die Nichtlinearität bzgl.  $w$ . Eine rigide Transformation ist ein Spezialfall affiner Transformationen. Im Vergleich zu rigiden Transformationen mit Translationen und Rotationen erlauben affine Transformationen zusätzlich auch Scherungen

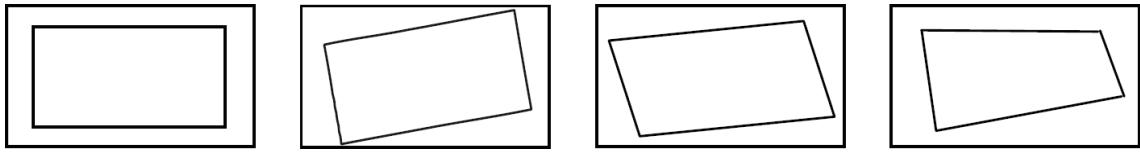


Figure 3.5: Von links nach rechts: Deformation eines Rechtecks mittels rigider, affiner und projektiver Gittertransformation.

und Skalierungen, siehe z.B. Abbildung 3.4.2. Affine Transformationen können mit 6 Parametern in 2D definiert werden als

$$y(x; w) = Q(x)w,$$

mit  $Q$  aus (3.36) und  $w \in \mathbb{R}^6$ . In 3D benötigt man analog 12 Parameter.

Darüber hinaus werden bei der Klasse der projektiven Transformationen neben affinen Transformationen noch projektive Verzerrungen berücksichtigt. In 3D können projektive Transformationen beispielsweise durch 15 Parameter festgelegt werden.

Ein Vorteil von parametrischen Transformationen ist, dass die Dimension des Suchraums in den meisten Fällen sehr klein ist. Allerdings lassen sich in der biomedizinischen Bildgebung natürlich auftretende Gittertransformationen mit parametrischen Modellen häufig nicht ausreichend abbilden. Allgemeinere, nicht-rigide Transformationen basieren beispielsweise auf Krümmungseigenschaften, Elastizität oder auch Hyper-Elastizität. Diese Konzepte lassen sich mit Variationsmethoden realisieren.

### 3.4.3 Variationsmethoden zur Bildregistrierung

Wie in der Einführung schon angedeutet, besteht das Problem der Registrierung von zwei Bildern darin, eine vernünftige Transformation zu finden, so dass die transformierte Version eines gegebenen Template-Bildes möglichst ähnlich ist zu einem gegebenen Referenzbild. Im Sinne von Variationsmethoden kann man die Aufgabe der Bildregistrierung über die Minimierung des folgenden Funktionalen modellieren

$$J(y) = D(f_T(y), f_R) + \alpha R(y) \rightarrow \min_y , \quad (3.37)$$

wobei  $D$  analog zu den Bildverarbeitungsaufgaben aus den vorherigen Abschnitten auch hier einen Datenterm darstellt. Es ist ein Distanzterm, der Unterschiede zwischen dem transformierten Template-Bild und dem Referenzbild in einem festzulegenden Maß bestraft. Die Distanzberechnung muss dabei nicht notwendigerweise direkt

auf allen Bildintensitäten basieren. Es kann auch sinnvoll sein nur bestimmte Funktionswerte zu verwenden oder bestimmte Features wie die Gradienten der Funktionen zu vergleichen. Mit dem Regularisierungsterm  $R$  können wir a priori Informationen über die Plausibilität einer Gittertransformation  $y$  einfließen lassen. Typischerweise wählt man Terme, die glatte oder elastische Gitterverformungen bevorzugen.  $\alpha$  bezeichnet auch hier einen Regularisierungsparameter, der die Gewichtung zwischen Datenterm und Regularisierung festlegt.

### 3.4.4 Datenterme zur Registrierung

Einer der bekanntesten Datenterme zur Bildregistrierung ist die  $L^2$ -Distanz zwischen dem transformierten Templatebild  $f_T(y(x))$  und dem Referenzbild  $f_R(x)$ , d.h.

$$D(f_T(y), f_R) = \frac{1}{2} \|f_T(y) - f_R\|_{L^2(\Omega)}^2 = \frac{1}{2} \int_{\Omega} (f_T(y(x)) - f_R(x))^2 dx .$$

Natürlich sind analog zur Entrauschung oder Rekonstruktion auch andere  $L^p$ -Normen bzw. Distanzen für unterschiedliche Rauschmodelle denkbar. Datenterme dieser Art, wie die  $L^2$ -Distanz vergleichen die Intensitäten des Referenzbildes und des transformierten Templatebildes direkt. Man nimmt also an, dass Intensitäten korrespondierender Gitterpunkte  $x$  und  $y(x)$  vergleichbar sind. Gerade im Hinblick auf multimodale Registrierung, siehe Abbildung 3.4.4, ist eine Gleichheit

$$f_T(y) = f_R$$

unter Umständen gar nicht erwünscht. Deshalb könnte man den Datenterm erweitern, indem man lediglich eine *lineare* Abhängigkeit zwischen transformiertem Templatebild und Referenzbild annimmt. Im Vergleich zum Integranden in der  $L^2$ -Distanz

$$(f_T(y(x)) - f_R(x))^2 = f_T^2(y(x)) - 2f_T(y(x))f_R(x) + f_R^2(x)$$

können wir uns dazu auf das Skalarprodukt

$$\langle f_T(y), f_R(x) \rangle = \int_{\Omega} f_T(y) f_R(x) dx$$

konzentrieren, das man auch als Kreuzkorrelation bezeichnet. Um zusätzlich noch eine Skalierung der Intensitäten des Templatebildes zu gestatten, kann man übergehen zu einer normalisierten Variante. Als Datenterm erhält man die sogenannte normalisierte Kreuzkorrelation (NCC),

$$D(f_T(y), f_R) = 1 - \frac{\langle f_T(y), f_R(x) \rangle^2}{\|f_T(y)\|^2 \|f_R\|^2} .$$

Ein weiteres häufig verwendetes Distanzmaß für die Bildregistrierung ist *mutual information* (deutsch: Transinformation oder gegenseitige Information). Das Konzept stammt ursprünglich aus dem Bereich der Informationstheorie und nutzt dabei die Stärke des statistischen Zusammenhangs zweier Zufallsgrößen aus. Mutual information ist die normalisierte Entropie einer gemeinsamen Dichte  $\rho_{[f_T, f_R]}$ , die die Koinzidenzen aller möglichen Paare summiert, d.h.

$$\begin{aligned} D(f_T(y), f_R) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \rho_{[f_T, f_R]} \log \left( \frac{\rho_{[f_T, f_R]}}{\rho_{[f_T]} \rho_{[f_R]}} \right) \\ &= H(f_T) + H(f_R) - H(f_T, f_R) , \end{aligned}$$

wobei  $H(f)$  eine Entropie von  $f$ ,

$$H(f) = - \int_{\mathbb{R}} \rho_{[f]} \log \rho_{[f]}$$

und  $H(f_T, f_R)$  eine gemeinsame Entropie,

$$H(f_T, f_R) = - \int_{\mathbb{R}} \int_{\mathbb{R}} \rho_{[f_T, f_R]} \log \rho_{[f_T, f_R]}$$

bezeichnet. Mutual information ist hochgradig nicht-konvex.

Eine Art Kompromiss zwischen der auf Intensitäten basierenden  $L^2$ -Distanz und der sehr allgemeinen mutual information ist ein Datenterm basierend auf *normalisierten Gradientenfeldern*. Gerade bei der multimodalen Registrierung können Gradienteninformationen (Kanten) eine sehr entscheidende Rolle spielen, wenn Intensitäten ansonsten schwer vergleichbar sind. Der Datenterm ist definiert als

$$D(f_T(y), f_R) = -\frac{1}{2} \int_{\Omega} \langle n(f_T, y(x)), n(f_R, x) \rangle^2 dx ,$$

wobei

$$n(f, x) := \begin{cases} \frac{\nabla f(x)}{\|\nabla f(x)\|} & , \text{ falls } \nabla f(x) \neq 0 \\ 0 & \text{sonst} \end{cases} .$$

Die Annahme bei diesem Datenterm ist, dass auch wenn sich die Intensitäten von Referenz- und Templatebild unterschiedlicher Modalitäten stark unterscheiden, ihre Intensitätsveränderungen an korrespondierenden Gitterpunkten vergleichbar sind. Ein weitere speziell Variante der Bildregistrierung ist die *landmarkenbasierte Registrierung*. Dabei wird eine Anpassung von Referenz- und Templatebild anhand ausgewählter Landmarken durchgeführt. Unter Landmarken versteht man ausgewählte anatomische Punkte. Sie können interaktiv durch den Benutzer festgelegt werden oder automatisch anhand besonderer geometrischer Eigenschaften, wie z.B. lokale



Figure 3.6: Von links nach rechts:  $f_R$  mit Landmarken,  $f_T$  mit Landmarken und  $f_T$  mit Landmarken-Trajektorien. Quelle: FAIR, Modersitzki

Krümmung, bestimmt werden. Bei computergestützten Operationen werden oftmals auch externe Marker, wie Schrauben, als Landmarken für eine Registrierung und perspektivischen Orientierung eingesetzt. In Vergleich zur intrinsische Registrierung spricht man hierbei von extrinsischer (äußerlichen) Registrierung. Anstatt die Distanz zweier vollständiger Dichtebilder zu minimieren, verwendet man bei der Landmarken-basierten Registrierung Datenterme, die lediglich Landmarken in Referenz- und Templatebild in Übereinstimmung bringen. Für  $n$  Landmarken-Paare  $(r_i, t_i)$ ,  $i = 1, \dots, n$  betrachtet man folgenden Distanzterm

$$D(y) = \sum_{i=1}^n \|y(r_i) - t_i\|^2 .$$

### 3.4.5 Regularisierungsterme zur Registrierung

Wie man an einfachen Beispielen mit Rauschen nachvollziehen kann, handelt es sich auch bei der Bildregistrierung um ein schlecht gestelltes Problem. Um eine sinnvolle Gitterverformung als Lösung zu erhalten, bedarf es einer gewissen Regularität. Mit Hilfe von Regularisierungstermen ist es auch bei dieser Bildverarbeitungsaufgabe möglich, a priori Wissen über das Problem einzubringen. In der Regel hängt die Wahl der Regularisierung vom speziellen Problem und den gegebenen Daten ab. Die am häufigsten verwendeten Regularisierungstechniken in diesem Bereich basieren auf Diffusion, Elastizität und Krümmung. Regularisierungsterme, die auf der  $L^2$ -Norm basieren lassen sich relativ allgemein über allgemeine Differentialoperatoren  $B$  in Abhängigkeit von  $y$  beschreiben,

$$R(y) = \frac{1}{2} \int_{\Omega} |B(y)|^2 dx .$$

Analog zur Entrauschung bei Bildern mittels Diffusion, z.B. quadratische Bestrafung bei Abweichungen von der Norm des Gradienten, erhalten wir in 2D auch hier mit

$$B := \begin{pmatrix} \nabla & 0 \\ 0 & \nabla \end{pmatrix}$$

eine Regularisierung mittels Diffusion

$$R(y) = \frac{1}{2} \int_{\Omega} (|\nabla y_1| + |\nabla y_2|)^2 .$$

Dieser Term ist konvex. Als Ergebnis erwarten wir glatte, homogene Gitterverformungen in Abhängigkeit von einem Regularisierungsparameter  $\alpha$ .

Bei der elastischen Regularisierung nehmen wir an, dass sich Objekte im Datensatz elastisch verformen. Mit dem Differentialoperator

$$B := \begin{pmatrix} \sqrt{\mu} \nabla & 0 \\ 0 & \sqrt{\lambda + \mu} \nabla \end{pmatrix}$$

$$R(y) := \frac{1}{2} \int_{\Omega} \mu \langle \nabla y, \nabla y \rangle + (\lambda + \mu)(\nabla \cdot y)^2 dx$$

Die Parameter  $\mu, \lambda \in \mathbb{R}$  bezeichnet man als Lamé-Konstanten. Man spricht bei dieser Regularisierung von linearer Elastizität. Eine Erweiterung ist Hyperelastizität. Damit lassen sich z.B. Verformungen modellieren, wie man sie von einem Radiergummi kennt.

Im Wesentlichen basieren diese Regularisierer auf Ableitungen erster Ordnung, siehe  $B$ . In dem Moment, wo wir Krümmungen in den Gitterverformungen berücksichtigen wollen, gehen wir über zu Ableitungen zweiter Ordnung. Mit

$$B := \begin{pmatrix} \Delta & 0 \\ 0 & \Delta \end{pmatrix}$$

erhalten wir einen Regularisierungsterm mit Krümmung

$$R(y) := \frac{1}{2} \int_{\Omega} (\Delta y_1)^2 + (\Delta y_2)^2 dx .$$

Dieser Term bestraft zu starke Krümmungen in der Transformation. Eine wichtige Eigenschaft ist hierbei die Rotationsinvarianz des Laplace-Operators. Der Krümmungsoperator besitzt außerdem einen unendlich-dimensionalen Nullraum und enthält alle harmonischen Vektorfelder, d.h. für die gilt  $\Delta y_i = 0$ .

Die vorgestellten Regularisierungsterme lassen sich analog im Sinne von allgemeinen Tensoren auf noch höhere Ordnungen erweitern.

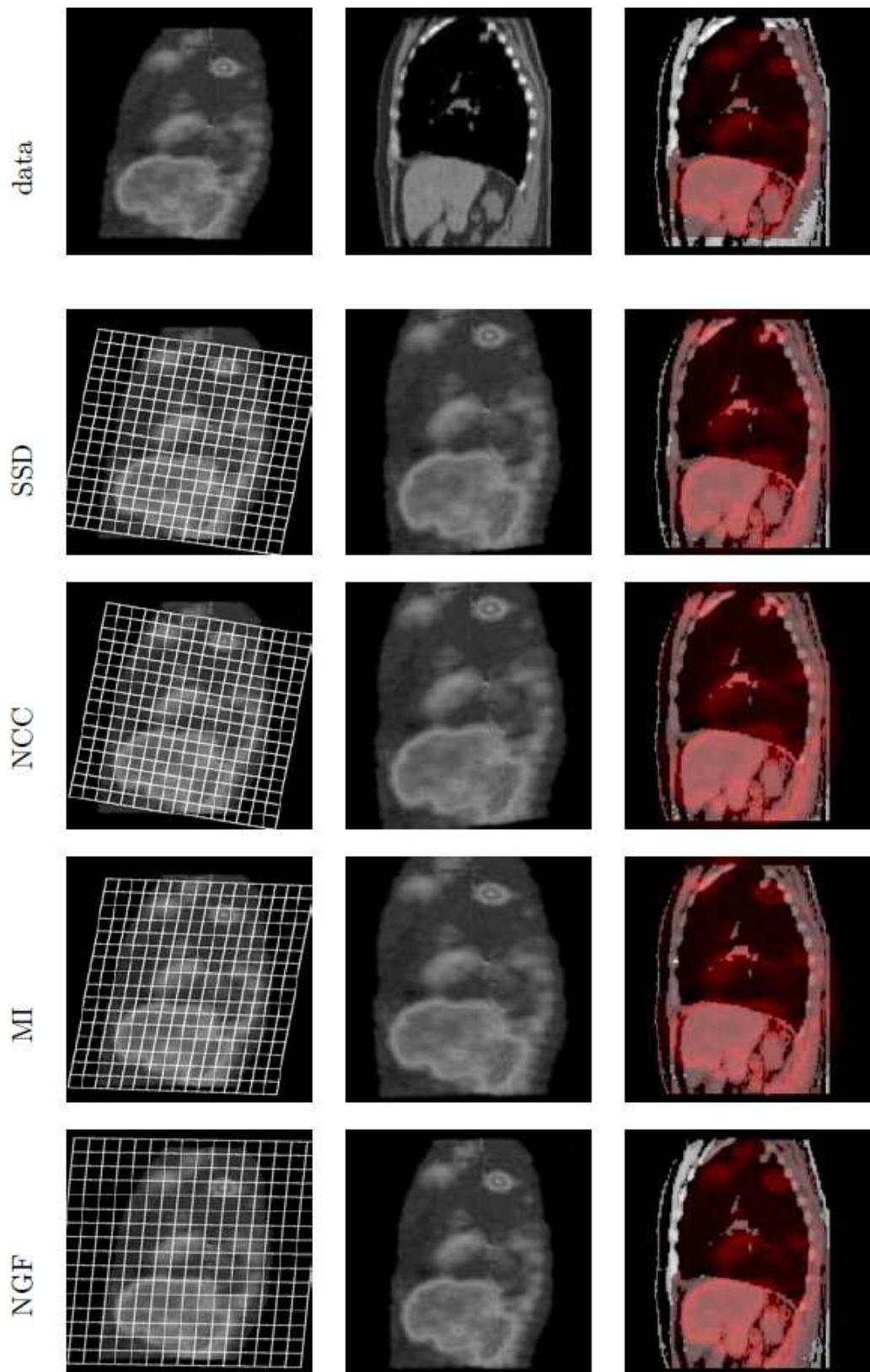


Figure 3.7: Multimodale PET/CT Registrierung mit unterschiedlichen Datentermen. links nach rechts: PET (mit Gittertransformation), CT, PET+CT.  
Quelle: FAIR, Modersitzki