



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 10 2005 051 745 A1 2006.05.24**

(12)

Offenlegungsschrift

(21) Aktenzeichen: **10 2005 051 745.5**

(22) Anmeldetag: **27.10.2005**

(43) Offenlegungstag: **24.05.2006**

(51) Int Cl.⁸: **G06T 1/40 (2006.01)**
G06T 5/50 (2006.01)

(30) Unionspriorität:

60/622,971 28.10.2004 US
11/259,560 26.10.2005 US

(71) Anmelder:

Siemens Corp. Research, Inc., Princeton, N.J., US

(74) Vertreter:

Patentanwälte Westphal Mussnug & Partner,
78048 Villingen-Schwenningen

(72) Erfinder:

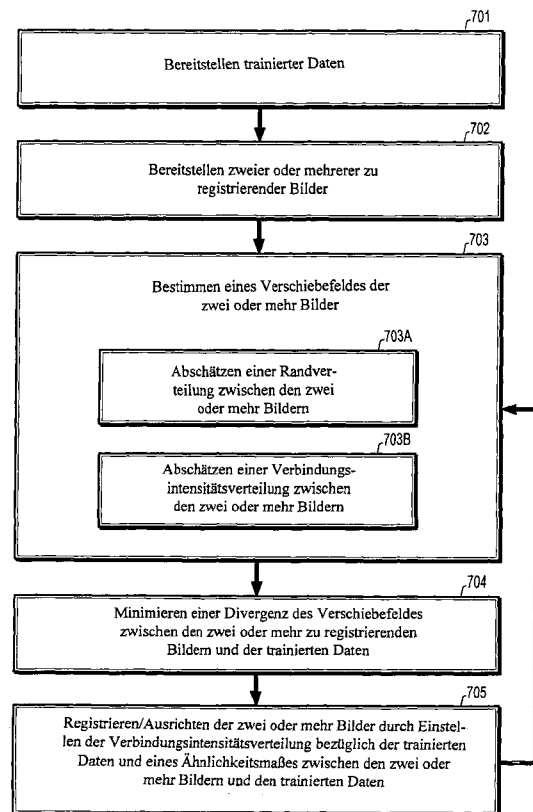
Guetter, Christoph, Princeton, N.J., US;
Hornegger, Joachim, 91096 Möhrendorf, DE; Xu,
Chenyang, Allentown, N.J., US

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

(54) Bezeichnung: **Unstarre multimodale Registrierung unter Verwendung statistischer Lernmethoden**

(57) Zusammenfassung: Ein Verfahren zur unstarren, multimodalen Registrierung enthält das Bereitstellen trainierter Daten entsprechend zuvor registrierter Bilder, die durch unterschiedliche Modalitäten erfasst wurden (701), das Empfangen zweier oder mehr Bilder, die durch die unterschiedlichen Modalitäten erfasst wurden (702), und das Registrieren der zwei oder mehr Bilder (705) in Übereinstimmung mit Kontextinformation, die auf der Grundlage der trainierten Daten bestimmt wird (703).



Beschreibung

Hintergrund der Erfindung

1. Technisches Gebiet

[0001] Die vorliegende Erfindung bezieht sich auf die multimodale Bildspeicherung und insbesondere auf ein System und ein Verfahren zur unstarren Speicherung multimodaler Bilddaten unter Verwendung statistischer Lernmethoden.

Stand der Technik

[0002] Eine unstarre, multimodale Bild/Volumen-Speicherung ist ein offenes Forschungsproblem und ein bedeutsamer Teil der laufenden Forschungen. Unstarre, multimodale Bildspeicherung in medizinischen Anwendungen ist für Ärzte besonders wichtig geworden. In einer solchen Umgebung sind Genauigkeit, Robustheit und Leistungsfähigkeit gefordert, um Diagnosen zuverlässig zu unterstützen. Die gegenwärtige unstarre, multimodale Speichertechnologie ist zunehmend kompliziert, jedoch kontextfrei.

[0003] Die Verschmelzung komplementärer Bildinformation hat gezeigt, dass sie bei der Diagnose besonders vorteilhaft ist. Weiterhin benötigten Abbildungstechniken, wie die molekulare Abbildung, eine multimodale Bildspeicherung zur Anzeige von funktionaler, anatomischer und/oder molekularer Bildinformation in einem einzigen verschmolzenen Bild.

[0004] Eine besondere Herausforderung an die unstarre, multimodale Speicherung ist die Handhabung unaufgelöster Korrespondenzen von Bild/Volumen-Orten, die aus Versuchen erwachsen, Funktionalbereiche mit Anatomie in Zuordnung zu bringen. Es ist ein offenes Problem, solche Korrespondenzen zuverlässig und in einer Weise zu finden, die für den Benutzer, beispielsweise den Arzt, aussagefähig ist. Vorhandene unstarre Bildspeicherungsmethoden behandeln unaufgelöste Korrespondenzen durch Angabe eines Energiefunktional, das allgemein optimal eine Lösung für das zu Grunde liegende Speicherproblem anbietet. Eine solches Energiefunktional kann beispielsweise durch ein Anziehungspotential, z. B. Ähnlichkeits/Distanz-Messung, und einen normalisierenden Term zusammengesetzt sein, der das Potential in einer etwas aussagekräftigen Weise einengt. Regularisierung ist wegen der schlechtgestellten Natur des Speicherproblems erforderlich. Eine vorgeschlagene Lösung kann entweder durch Entlangschreiten an einem Gradienten oder durch Lösung einer mit dem Problem einhergehenden PDE gefunden werden. Die Schlechtgestellttheit des Energiefunktional und/oder die Größe des Parameterraums sind noch immer problematisch beim Auffinden eines allgemeinen Optimums durch Verwenden einer Regularisierung. Die Regularisierung kann zu einer Fehlregistrierung oder Konvergenz in einem örtlichen Extremum führen.

Aufgabenstellung

[0005] Es besteht daher Bedarf an einem System und einem Verfahren zum unstarren Registrieren multimodaler Bilddaten unter Verwendung statistischer Lernmethoden.

Übersicht über die Erfindung

[0006] Gemäß einer Ausführungsform der vorliegenden Erfindung enthält ein rechnergestütztes Verfahren zur unstarren, multimodalen Registrierung das Anbieten trainierter Daten entsprechend zuvor gespeicherter Bilder, die durch unterschiedliche Modalitäten erfasst wurden, das Empfangen zweier oder mehr Bilder, die durch die unterschiedlichen Modalitäten erfasst wurden, und das Registrieren der zwei oder mehr Bilder entsprechend der Kontextinformation, die auf der Grundlage der trainierten Daten bestimmt wird.

[0007] Das Registrieren der zwei oder mehr Bilder ist automatisch.

[0008] Die trainierten Daten sind eine Verbindungswahrscheinlichkeitsfunktion der unterschiedlichen Modalitäten. Die Verbindungswahrscheinlichkeitsfunktion ist eine Einschränkung oder ein Regularisierungsfaktor in einem Optimierungsprozess der Registrierung. Die trainierten Daten sind ein durchschnittlicher Bildort eines Bildintensitätspaares. Die trainierten Daten sind Gestaltinformation, das Verfahren enthält das Bestimmen einer oder mehrerer Gestalten in den zwei oder mehr Bildern und die Klassifizierung der einen oder mehreren Gestalten als gültig oder ungültig entsprechend den trainierten Daten.

[0009] Eine Korrespondenz zwischen den zwei oder mehr Bildern wird unter Verwendung einer Kontextinformation von Gestalten in den verschiedenen Modalitäten der trainierten Daten bestimmt. Gemäß einer Ausführungsform der vorliegenden Offenbarung liefert ein rechnergestütztes Verfahren zum unstarren, multimodalen Registrieren trainierte Daten entsprechend einer früheren Registrierung und liefert zwei oder mehr Bilder, die unter Verwendung unterschiedlicher Modalitäten erfasst werden. Das Verfahren umfasst das Bestimmen eines Verschiebefeldes der zwei oder mehr Bilder, wobei ein erstes Bild der zwei oder mehr Bilder als ein Bezugsbild definiert ist und ein nachfolgendes Bild als ein Ausrichtbild definiert ist, umfassend das Abschätzen einer Randverteilung des Bezugsbildes und des Ausrichtbildes und das Abschätzen einer Verbindungsintensitätsverteilung zwischen dem Bezugsbild und dem Ausrichtbild. Das Verfahren umfasst das Ausrichten der zwei oder mehr Bilder durch Einstellen der Verbindungsintensitätsverteilung, bis sie im Wesentlichen zu einer gelernten Verbindungsintensitätsverteilung der trainierten Daten passt und gleichzeitig ein gewünschtes Ähnlichkeitsmaß zwischen den zwei oder mehr Bildern und den trainierten Daten ergibt.

[0010] Die Ausrichtung der zwei oder mehr Bilder umfasst die Regularisierung des Verschiebefeldes zwischen den zwei oder mehr Bildern und den trainierten Daten. Die Regularisierung verwendet eine Tycho-nov-Regularisierung mit dem folgenden Gradienten:

$$\nabla_u R[u] = \operatorname{div} \left(\frac{R'[u]}{\nabla u} \nabla u \right) = \operatorname{div}(\nabla u) = \Delta u,$$

wobei u das Verschiebefeld ist und R die Regularisierung von u definiert. Das Verfahren der unstarren, multimodalen Registrierung ist gegeben durch:

$$\mathcal{J}(u) = \alpha \mathcal{I}_{MI}(u) + (1 - \alpha) \mathcal{I}_{KL}(u) + \lambda R(u), \quad \alpha \in [0,1], \lambda \in \mathbb{R}_+$$

$$\hat{u} = \operatorname{argmin} \mathcal{J}(u),$$

wobei λ eine positive Konstante ist, die einen Regularisierungsumfang bestimmt, \mathcal{I}_{KL} eine Divergenz zwischen der Verbindungsintensitätsverteilung der zwei oder mehr Bilder und der Verbindungsintensitätsverteilung der trainierten Daten misst und \mathcal{I}_{MI} eine gegenseitige Information der Verbindungsintensitätsverteilung der zwei oder mehr Bilder bezeichnet.

[0011] Das wünschenswerte α -Ähnlichkeitsmaß zwischen den zwei oder mehr Bildern und den trainierten Daten ist ein Maximum.

[0012] Gemäß einer Ausführungsform der vorliegenden Offenbarung wird eine Programmspeichervorrichtung angegeben, die maschinenlesbar ist und greifbar ein Programm von Befehlen verkörpert, die durch die Maschine ausführbar sind, um Verfahrensschritte für unstarre, multimodale Registrierung auszuführen. Das Verfahren enthält das Anbieten trainierter Daten entsprechend zuvor registrierten Bildern, die durch unterschiedliche Modalitäten erfasst wurden, das Empfangen von zwei oder mehr durch die unterschiedlichen Modalitäten erfassten Bilder und das Registrieren der zwei oder mehr Bilder entsprechend der Kontextinformation, die auf der Grundlage der trainierten Daten bestimmt wird.

Ausführungsbeispiel

Kurzbeschreibung der Zeichnungen

[0013] Bevorzugte Ausführungsformen der vorliegenden Erfindung werden nachfolgend detaillierter unter Bezugnahme auf die begleitenden Zeichnungen erläutert.

[0014] Fig. 1 ist ein Diagramm eines Systems zur multimodalen Bildregistrierung gemäß einer Ausführungsform der vorliegenden Erfindung;

[0015] Fig. 2 ist ein Diagramm des Registrierungsblocks **102** in Fig. 1;

[0016] Fig. 3 ist ein Diagramm eines Rechnersystems für multimodale Bildregistrierung gemäß einer Ausführungsform der vorliegenden Offenbarung;

[0017] Fig. 4a–g sind Diagramme von Phantombildern gemäß einer Ausführungsform der vorliegenden Offenbarung;

[0018] Fig. 5a–f sind Schnittbilder gemäß einer Ausführungsform der vorliegenden Offenbarung;

[0019] Fig. 6a–d sind SPECT/CT-Bilder gemäß einer Ausführungsform der vorliegenden Offenbarung; und

[0020] Fig. 7 ist ein Flussdiagramm eines Verfahrens gemäß einer Ausführungsform der vorliegenden Offenbarung.

Detaillierte Beschreibung bevorzugter Ausführungsformen

[0021] Ein Verfahren zur unstarren, multimodalen Bildregistrierung verlangt A-priori-Fachwissen. A-priori-Fachwissen erhält man von registrierten Bildern, die dazu verwendet werden, statistisch Wahrscheinlichkeitsverteilungen zu lernen, bevor eine Registrierung an interessierenden Bildern ausgeführt wird. Unregistrierte Bilder werden durch ein unstarres Registrierungsverfahren unter Verwendung der trainierten Fachkenntnis ausgerichtet.

[0022] Gemäß Fig. 1 enthält ein System zur multimodalen Registrierung eine Registrierungskomponente, eine Abbildungskomponente, eine Trainingskomponente und gegebenenfalls eine Datenbank von ziel- und modalitätsspezifischen Statistiken. Die Registrierungskomponente **101** empfängt wenigstens zwei unregistrierte Bilder, die unter Verwendung unterschiedlicher Modalitäten aufgenommen wurden, z. B. Einzelphotonenemissions-CT (SPECT)/Computertomographie (CT), Positronenemissionstomographie (PET)/CT oder T1–T2 MR-Bilder. Die unregistrierten Bilder werden durch eine Abbildungskomponente **102** geliefert, beispielsweise Abbildungshardware/Software. Die Registrierungskomponente **101** verwendet Kontextinformation von Gestalten in den unregistrierten Bildern zur Ausrichtung der unregistrierten Bilder, die von der Abbildungskomponente **102** eingegeben werden. Die Kontextinformation wird von einer Trainingskomponente **103** geliefert. Die Trainingskomponente **103** führt statistische Lernmethoden zur Stützung der Registrierung oder einen Optimierungsprozess der Registrierungskomponente **101** aus. Trainingsdaten werden der Registrierungskomponente **101** durch eine Trainingskomponente **103** eingegeben. Die Trainingsdaten umfassen A-priori-Daten, die von zuvor registrierten Bildern bestimmt werden, die Modalitäten der unregistrierten Bilder darstellen; zuvor registrierte multimodale Datensätze. Eine Ausrichtung der vorregistrierten Bilder kann beispielsweise durch einen Techniker bestimmt werden. Gemäß Fig. 2 werden die Datensätze, z. B. Punktkorrespondenzen, dazu verwendet, Kontextinformation **203** zu lernen. Sobald die Kontextinformation gelernt ist, steht sie für die Registrierung **201** zur Verfügung.

[0023] Kontextkenntnis führt die automatisierte Registrierung. Kontext kann in den Registrierprozess **201** durch mehrere Strategien eingebracht werden. Eine Strategie ist es, statistische Lernmethoden **203** zu verwenden, wo vorregistrierte Datensätze die Kontextkenntnis bieten. Das Registrierungssystem lernt statistischen Information von den automatisch und/oder manuell vorregistrierten Datensätzen und erfasst Information, die als Grundwahrheit betrachtet werden kann, von der zukünftige Registrierungen Nutzen ziehen. Diese Information ist auch als eine A-priori-Kenntnis bekannt, weil sie vor der Registrierung erlangt wird.

[0024] Ein unstarres Registrierungsverfahren **201**, das von der Registrierungskomponente **101** ausgeführt wird, folgt einem allgemeinen Rahmen der Optimierung einer Energiefunktion, um die zu Grunde liegende Transformation zwischen den Bildern zu finden. Weiterhin enthält das unstarre Registrierungsverfahren **201** eine A-priori-Kenntnis **203**, um die Registrierung oder einen Optimierungsprozess zu führen. A-priori-Kenntnis kann entweder zum Betrieb des Registrierungsprozesses verwendet werden oder als eine Regularisierungsbeschränkung im Optimierungsprozess. Diese Kenntnis kann unterschiedlicher Art sein.

[0025] Die zusätzlich zur Einbindung von A-priori-Kenntnis **203** kann Information auch durch Benutzer-Feedback **202** erlangt werden. Mit einer Registrierung **201**, die aus einer A-priori-Kenntnis **203** lernt, nutzt man die Erfolgsrate der unstarren multimodalen Registrierung **201** aus. In manchen Fällen mag der Benutzer mit den Registrierungsergebnissen **204** aus verschiedenen Gründen nicht zufrieden sein. Das System lernt aus einer A-priori-Kenntnis, somit kann Feedback-Kenntnis **205** vom Benutzer eingebracht werden. Somit nutzt der Registrierungsprozess **201** nicht nur eine A-priori-Kenntnis **203** trainierter Datensätze **104** vorteilhaft aus, sondern auch eine Feedback-Schleife, die durch Information zugeführt wird, die vom Benutzer **202** gegeben wird.

[0026] Feedback **202** kann auch manuell bestimmt werden durch Korrektur eines Registrierungsergebnisses **204** beispielsweise an einigen gegebenen charakteristischen Positionen. Diese Korrekturinformation wird dann zur Lernkomponente **203** rückgemeldet **205**, um die zukünftige Registrierung **201** zu verbessern.

[0027] In multimodalen Fällen differiert die Registrierung, die an verschiedenen interessierenden Organen

ausgeführt werden, vom Ausmaß der Herausforderung. Dieses ist durch die Fähigkeiten der unterschiedlichen Abbildungstechniken bedingt, z. B. PET-Anwendungen am Kopf sind klarer hinsichtlich der Anatomie als solche beispielsweise des Thorax. Daher können organspezifische Lernmethoden durch die Registrierungskomponente unterstützt werden.

[0028] Eine manuelle oder automatische Datenbankklassifizierung verfolgt die A-priori-Information pro Organ und Modalitätskombination, die dann entsprechend in der Ausführung der Registrierung verwendet wird.

[0029] Jene Datenbanken (siehe [Fig. 1](#), Block **104**), die die Information über das Organ und die modalitätsspezifische statistische Information enthalten, können mit Systemen geteilt und/oder ausgeweitet werden, die von verschiedenen Benutzern beeinflusst/trainiert werden. Nimmt man Registrierungssysteme an, die in mehreren Krankenhäusern installiert sind, dann kann die unterschiedlich gelernte A-priori-Information zur Synergie verwendet werden. Somit kann ein System konstant lernen und sich verbessern auf der Grundlage des Feedback von vielen Benutzern, z. B. Spezialisten. Die Information kann über ein Netz, wie beispielsweise das Internet oder ein private Netz, zwischen den Krankenhäusern geteilt werden. Die Registrierungsqualität hat von einem solchen lernenden System einen großen Nutzen.

[0030] Lernende Verfahren **103** sind als überwachte lernende Verfahren beschrieben. Andere lernende Verfahren können verwendet werden. Beispielsweise im Falle unzureichender Datenverfügbarkeit, z. B. nicht genügend vorregistrierter Bilddaten, könnte eine unüberwachte Lerntechnik verwendet werden. Solche Techniken helfen bei der Bewältigung des Problems, dass man nicht weiß, wie eine Grundwahrheit aussieht.

[0031] Es versteht sich, dass die vorliegende Erfindung in zahlreichen Formen von Hardware, Software, Firmware, Spezialrechnern oder einer Kombination davon ausgeführt werden kann. In einer Ausführungsform kann eine regularisierte Bayes-Methode für die Bildwiederherstellung in Software als ein Anwendungsprogramm realisiert werden, das greifbar auf einer Programmspeichervorrichtung ausgebildet ist. Das Anwendungsprogramm kann hochgeladen und von einer Maschine ausgeführt werden, die jede geeignete Architektur aufweist.

[0032] Gemäß [Fig. 3](#) kann gemäß einer Ausführungsform der vorliegenden Offenbarung ein Rechnersystem **301** zur Ausführung eines Registrierungsverfahrens für multimodale Bilder unter anderem eine zentrale Prozessoreinheit (CPU) **302**, ein Speicher **303** und eine Eingabe/Ausgabe (I/O)-Schnittstelle **104** enthalten. Das Rechnersystem **301** ist im Allgemeinen über die I/O-Schnittstelle **304** mit einer Anzeigevorrichtung **305** und zahlreichen Eingabevorrichtungen **306**, wie beispielsweise einer Maus und einer Tastatur, verbunden. Die Unterstützungsschaltungen können Schaltungen, wie Cache, Netzteile, Taktschaltungen und einen Übertragungsbus einschließen. Der Speicher **303** kann einen Arbeitsspeicher (RAM), einen Festspeicher (ROM), ein Plattenlaufwerk, ein Bandlaufwerk usw. oder eine Kombination daraus enthalten. Die vorliegende Erfindung kann als eine Routine **307** ausgeführt werden, die in einem Speicher **303** gespeichert ist, und von der CPU **302** durchgeführt werden, um das Signal von der Signalquelle **308** zu verarbeiten. Als solches ist das Rechnersystem **301** ein Allzweckrechnersystem das zu einem Spezialrechnersystem wird, wenn es die Routine **307** der vorliegenden Erfindung ausführt.

[0033] Die Rechnerplattform **301** enthält auch ein Betriebssystem und einen Mikrobefehlscode. Die zahlreichen Prozesse und Funktionen, die hier beschrieben sind, können entweder Teil des Mikrobefehlscodes oder Teil des Anwendungsprogramms (oder eine Kombination davon) sein, das über das Betriebssystem ausgeführt wird. Außerdem können zahlreiche periphere Vorrichtungen mit der Rechnerplattform verbunden sein, wie beispielsweise eine zusätzliche Datenspeichervorrichtung und eine Druckvorrichtung.

[0034] Es versteht sich ferner, dass wegen einiger der Systembestandteile und Verfahrensschritte, die in den begleitenden Zeichnungen dargestellt sind und in Software ausgeführt werden können, die wirklichen Verbindungen zwischen den Systemkomponenten (oder die Prozessschritte) in Abhängigkeit von der Art und Weise differieren können, in der die vorliegende Erfindung programmiert ist. Mit den Lehren der vorliegenden Erfindung, die hier angeboten wird, an der Hand ist der einschlägige Fachmann in der Lage, diese und ähnliche Ausführungsformen oder Konfigurationen der vorliegenden Offenbarung zu entwerfen.

[0035] Gemäß [Fig. 7](#) kombiniert ein Verfahren gemäß einer Ausführungsform der vorliegenden Offenbarung beide Perspektiven **702** in eine vereinigte Formulierung durch gleichzeitige Anregung der beobachteten Verbindungsintensitätsverteilung, der erwarteten Verbindungsintensitätsverteilung zu ähneln, die a priori **701** gelernt wurde, und durch Maximierung eines Ähnlichkeitsmaßes. Dies kann intuitiv als Führung eines kontextfreien Ähnlichkeitsmaßes durch Vorkenntnis verstanden werden.

[0036] Ein kombiniertes Registrierungsrahmenwerk gemäß einer Ausführungsform der vorliegenden Offenbarung kann als eine Minimierung der vorliegenden Kostenfunktion definiert werden:

$$\mathcal{J}(u) = \alpha \mathcal{I}_{MI}(u) + (1 - \alpha) \mathcal{I}_{KL}(u) + \lambda R(u), \quad \alpha \in [0,1], \lambda \in \mathbf{R}_+$$

$$\hat{u} = \operatorname{argmin} \mathcal{J}(u), \quad (1)$$

wobei u ein Verschiebefeld ist, R die Regularisierung oder Glättung von u definiert und λ eine positive Konstante ist, die den Umfang der Regularisierung entscheidet. \mathcal{I}_{KL} misst die KL-Divergenz zwischen beobachteten und gelernten Daten, und \mathcal{I}_{MI} bezeichnet einen Ausdruck für gegenseitige Informationen (MI) der beobachteten Daten. Ein Verschiebefeld, das MI maximiert, wird durch Vorinformation gesteuert, um eine genaue Ausrichtung zu erzielen. Der Faktor α steuert den Führungsumfang durch Vorkenntnis. Für $\alpha = 0$ basiert das Registrierungsproblem nur auf der Vorinformation. Für $\alpha = 1$ ist die Registrierung als die klassische Optimierung von MI ohne jegliche Vorinformation definiert. Für $\alpha \in (0,1)$ wird die Maximierung von MI durch klinischen Kontext in Form von Vorkenntnis gelenkt, die durch die Minimierung der KL-Divergenz erfasst wurde. Diese Vorkenntnis kann auf mehrere Arten erlangt werden. Man kann die Fachkenntnis eines Arztes benutzen, der manuell Trainingsbilder ausrichtet, oder man kann die verschmolzenen Bilddaten wirksam einsetzen, die unter Verwendung der Dual-Modalität (PET/CT, SPECT/CT), auch als Hybridscanner bekannt, erlangt wurden. Letzteres liefert extensive Mengen an vorregistrierten Daten, was sehr wichtig zur Vermeidung patientenspezifischer Trainingsdaten ist. Um die Robustheit zu steigern, kann man eine Verbindungsdichteverteilung lernen, die durchschnittliche Vorinformation von n vorausgerichteten Bildern repräsentiert. Jedoch muss die mittlere Vorinformation sorgfältig untersucht werden, weil die meisten Scanner eine Fehlausrichtung aufgrund einer Organbewegung nicht korrigieren können.

[0037] Nachfolgend werden zwei Bilder, die durch die Funktionen zu registrieren sind, als $f_1: \Omega \subset \mathbf{R}^n \rightarrow \mathbf{R}$ und $f_2: \Omega \subset \mathbf{R}^n \rightarrow \mathbf{R}$ bezeichnet. Die Bilder werden registriert, indem ein grundlegendes Verschiebefeld **703** abgefragt wird. Mit gegebenen Bildern kann das Verschiebefeld durch eine Abbildung $u: \Omega \rightarrow \Omega$ modelliert werden. Ohne Verlust an Allgemeinheit können f_1 als das Bezugsbild und f_2 als das Ausrichtbild während des Registrierungsprozesses bezeichnet werden.

[0038] Der Randwert **703A** und die Verbindungsintensitätsverteilungen **703B** die aus $f_1(x)$ und $f_2(x + u(x))$ abgeschätzt werden, sind mit $p_1^0(f_1), p_2^0(f_2)$ und $p_u^0(f, f_2)$ bezeichnet. $p^1(f_1, f_2)$ ist ein Schätzwert für die Verbindungsintensitätsverteilung der Trainingsdaten. In der Praxis werden die Verteilungen unter Verwendung eines nicht-parametrischen Parzen-Fensterschätzers mit einer Gaussglocke als Fensterfunktion abgeschätzt. Vorkenntnis wird durch Minimierung der KL-Divergenz zwischen beobachteten und trainierten Daten **704** eingebaut. Die KL-Divergenz für ein gegebenes Verschiebefeld u kann ausgedrückt werden als

$$\mathcal{I}_{KL}(u) = \int_{\Omega} p_u^0(i_1, i_2) \ln \frac{p_u^0(i_1, i_2)}{p^1(i_1, i_2)} dx \quad (2)$$

wobei $i_1 = f_1(x)$ und $i_2 = f_2(x + u(x))$. Die MI-basierte Zielfunktion ist als das negierte MI zwischen dem Bezugsbild und dem Ausrichtbild, transformiert durch u , definiert und kann ausgedrückt werden als:

$$\mathcal{I}_{MI}(u) = - \int_{\Omega} p_u^0(i_1, i_2) \ln \frac{p_u^0(i_1, i_2)}{p_1^0(i_1) p_2^0(i_2)} dx \quad (3)$$

[0039] MI kann als die KL-Divergenz zwischen der beobachteten Verbindungsdichte und dem Produkt der beobachteten Randwerte betrachtet werden, während in \mathcal{I}_{KL} das Produkt der Randdichten durch die aus den Trainingsdaten gelernten Vorkenntnis ersetzt ist. Es ist anzumerken, dass das negierte MI hier zur Definition eines Aufwandes verwendet wird.

[0040] Das Minimum von (1) kann mittels Abweichungsberechnung gefunden werden. Man kann den Gradienten des kombinierten Funktionals bezüglich des Verschiebefelds herabsteigen. Der Gradient von (1) ist definiert als

$$\nabla_u \mathcal{J} = \alpha \nabla_u \mathcal{I}_{MI} + (1 - \alpha) \nabla_u \mathcal{I}_{KL} + \lambda \nabla_u R \quad (4)$$

[0041] Der Gradient von MI ist von Hermsillo et al. abgeleitet worden; der Gradient von MI ist bezüglich individueller Raumpunktverschiebungen unter Verwendung einer kontinuierlichen und differenzierbaren

Verbindungswahrscheinlichkeitsverteilung, die durch Parzen-Fensterung aufgebaut wurde, abgeleitet worden, was ein raumpunktwises Verschiebefeld ergibt. Der Gradient von MI kann beschrieben werden als:

$$\nabla_u \mathcal{I}_{MI} = -\frac{1}{N} \left[\left(\frac{\partial_2 p_u^0(i_1, i_2)}{p_u^0(i_1, i_2)} - \frac{\partial_2 p_u^0(i_2)}{p_u^0(i_2)} \right) * G_\sigma \right] (f_1(x), f_2(x+u(x)) \cdot \nabla f_2(x+u(x))) \quad (5)$$

[0042] Um den Gradienten der KL-Divergenz abzuleiten, kann die Definition für ein nicht-parametrisches Parzen-Dichtemodell verwendet werden. Nach einiger Manipulation kann $\nabla_u \mathcal{I}_{KL}$ wie folgt beschrieben werden:

$$\nabla_u \mathcal{I}_{KL} = -\frac{1}{N} \left[\left(\frac{\partial_2 p_u^0(i_1, i_2)}{p_u^0(i_1, i_2)} - \frac{\partial_2 p_u^1(i_2)}{p_u^1(i_2)} \right) * G_\sigma \right] (f_1(x), f_2(x+u(x)) \cdot \nabla f_2(x+u(x))) \quad (6)$$

[0043] Hier ist G_σ eine zweidimensionale Gaussglocke mit Standardabweichung σ , ∂_2 ist die partielle Ableitung einer Funktion bezüglich ihrer zweiten Variablen, und N ist eine Normierungskonstante. Der Ausdruck

$$\frac{\partial_2 p_u^0(i_1, i_2)}{p_u^0(i_1, i_2)} - \frac{\partial_2 p^1(i_1, i_2)}{p^1(i_1, i_2)}$$

ist eine Vergleichs-/Ähnlichkeitsfunktion eines Registrierungsverfahrens. Diese Vergleichsfunktion wird wiederholt bei der Registrierung ausgewertet, in diesem Falle auf ein Maximum. Ausrichtung wird durch wiederholte oder kontinuierliche Einstellungen des Verbindungsintensitätsmodells erreicht, bis es der gelernten Verbindungsintensitätsverteilung **705** gleicht. Weiterhin kann man durch diese Beurteilung sehen, dass in einer KL-basierten Lösung gemäß einer Ausführungsform der vorliegenden Offenbarung die beobachtete Verbindungsintensitätsverteilung nicht statisch bleibt.

[0044] Die Abweichungsberechnung ermöglicht die Bestimmung eines minimierenden Verschiebefeldes durch Absteigen längs des Gradienten $\nabla_u \mathcal{J}$. Der Gradientenfluss:

$$u_t = -\nabla_u \mathcal{J} \\ u(\cdot, 0) = u_0 \quad (7)$$

wobei u_0 eine geeignete Anfangsschätzung für das Verschiebefeld ist. Ein Tychonov-Modell kann für die Regularisierung **705** verwendet werden, z. B.

$$\nabla_u \mathcal{R}[u] = \operatorname{div} \left(\frac{\mathcal{R}'[u]}{|\nabla u|} \nabla u \right) = \operatorname{div}(\nabla u) = \Delta u,$$

Ihr Gradientenausdruck ist

$$\mathcal{R}[u] = \frac{1}{2} \int_\Omega |\nabla u(\mathbf{x})|^2 d\mathbf{x}.$$

wobei Δ den Laplace-Operator bezeichnet. Ausgehend von einer Anfangsschätzung wird eine Gradientenabsteigestrategie verfolgt, um eine Lösung für (1) zu finden. Um eine größere Klasse Deformationen zu entdecken, um die Berechnungskosten zu vermindern und um irrelevante Extrema des nicht-konvexen Funktionals zu vermeiden, kann ein Grob-nach-fein-Schema verwendet werden, z. B. aufeinander folgendes Glätten und Unterabtasten der Bilder.

[0045] Die trainierten Daten können beispielsweise eine Verbindungswahrscheinlichkeitsfunktion unterschiedlicher Modalitäten sein. Ein weiteres Beispiel trainierter Daten ist ein mittlerer Bildort eines Bildintensitätspaares. Ein noch weiteres Beispiel trainierter Daten ist Gestaltsinformation, wobei die Gestaltsinformation verglichen wird mit Gestaltsinformation, die zu den zu registrierenden Bildern bestimmt wurde. Die Gestaltsinformation kann dazu verwendet werden zu ermitteln, ob eine Registrierung entsprechend den Gestalten gültig ist, d. h. ob gewisse Organe gewisse erwartete Gestalten haben. Andere Beispiele trainierter Daten liegen für den Fachmann im Lichte der vorliegenden Offenbarung auf der Hand.

[0046] Es wurden Phantombilder geschaffen, um die Bedeutung der Verwendung kontext-spezifischer Infor-

mation hervorzuheben. Die [Fig. 4a–g](#) visualisieren eine mehrdeutige Einstellung für unstarre Registrierung. [Fig. 4a](#) ist ein Bezugsbild (512×512), [Fig. 4b](#) ist ein Ausrichtbild und [Fig. 4c](#) ist ein Differenzbild. Ein Kreis wird unstarrr auf einen anderen unterschiedlicher Intensität registriert. Sein Ort ist jedoch so gewählt, dass sich in einem Verbindungsbild eine Überlappung mit zwei anderen Kreisen, einem kleineren und einem größeren Kreis, ergibt ([Fig. 4a](#)). Diese Einrichtung deutet an, dass es wenigstens zwei äquivalente Optima für eine kontext-freie Distanzmessung zur Ausrichtung der Kreise gibt. Zwei Methoden wurden verglichen, z. B. Minimierung (1) mit $\alpha = 1$ bzw. $\alpha = 0$. Die Registrierung wurde so trainiert, dass der Kreis auf den kleinen Kreis ausgerichtet wurde. Die MI-Methode, $\alpha = 1$, findet ein Optimum beim Registrieren auf den großen Kreis, [Fig. 4\(d\)](#) und [4\(e\)](#). Die KL-Lösung, $\alpha = 0$, registriert auf den kleinen Kreis, [Fig. 4\(f\)](#) und [4\(g\)](#). Es ist anzumerken, dass bei Verwendung der KL-Lösung die Methode auch zur Ausrichtung auf den großen Kreis trainiert werden kann.

[0047] Die KL-Methode ($\alpha = 0$) wurde an einem simulierten magnetischen T1/T1-Resonanzbild-(MRI) Gehirndatensatz getestet, das von der Datenbank Brain Web Simulated Brain erhalten war. T1 und T2 repräsentieren zwei unterschiedliche Modalitäten von MRI-Abtastungen, wobei T1 die Spin-Gitter-Relaxationszeit ist und T2 eine Spin-Spin-Relaxationszeit ist. Die Coronalschnitte, [Fig. 5a](#) und [Fig. 5b](#), wurden für das Training verwendet, während die Registrierung an Sagittalschnitten ausgeführt wurde, [Fig. 5c](#) und [Fig. 5d](#).

[0048] Das T2-Bild ist durch ein künstlich geschaffenes Verschiebefeld deformiert worden. Dieses Experiment zeigt die Stärke des Trainings von Verbindungsintensitätsverteilungen, die erfolgreich für eine unstarre Registrierung verwendet werden.

[0049] Ein weiteres Experiment wurde an zwei korrespondierenden Schnitten eines SPECT/CT-Datensatzes (siehe [Fig. 6a](#)) ausgeführt, der durch einen SPECT/CT-Hybridscanner der Firma Siemens vom Typ Symbia T2 erhalten wurde. Vorkennntnis wurde aus diesen zwei Schnitten erzeugt und deformierte den SPECT-Schnitt durch ein künstliches Verschiebefeld (siehe [Fig. 6b](#)). MI ($\alpha = 1$) und KL ($\alpha = 0$) wurden auf Güte verglichen, und die endgültigen Registrierungsergebnisse sind in [Fig. 6c](#) bzw. [Fig. 6d](#) dargestellt. Da die Grundwahrheit verfügbar ist, kann der Unterschied zwischen den verworfenen SPECT-Bildern und dem Originalbild zur optischen Auswertung bestimmt werden. Obgleich eine Multi-Auflösungs-Strategie verwendet wurde, wird aus den Differenzbildern, die MI-basierende Lösung in einem irrelevanten örtlichen Minimum eingeschlossen, möglicherweise wegen ihrer Unempfindlichkeit gegenüber örtlicher Deformation. Dieses Experiment zeigt den möglichen Vorteil des Einbaus von Vorkennntnis bei der Registrierung in klinischen Anwendungsfällen.

[0050] Ein weiteres Experiment beschreibt eine PET/CT-Registrierung aus der klinischen Praxis unter Einsatz der visuellen Auswertung durch einen Experten. Die Bilddaten eines 70jährigen männlichen Patienten mit mehreren Läsionen in der Lunge wurden mit Hilfe eines Siemens Sensation 10 (CT) und eines Siemens Ecot 926 (PET) erhalten. Das PET wurde sechs Tage nach dem CT erzeugt. Gemäß der Auswertung durch einen erfahrenen Arzt wurden nur Teile des Volumens genau durch eine vorangehende manuelle Fusion registriert. Für Experimente wurden zwei Schnitte davon trainiert, die als gut registriert klassifiziert wurden, und es wurde ein Verfahren zur Registrierung eines fehlausgerichteten Schnitts ausgeführt. Der Arzt wertete die kombinierte Lösung aus, d. h. 40% MI und 60% KL, als die genaueste unter drei Registrierungsergebnissen (kombiniert, reines MI und reines KL) aufgrund ihrer Genauigkeit nicht nur bei der Ausrichtung des Herzens, sondern auch beim Abgleichen der Thorax-, Mediastinal- und Leber-Umriss. Variationen in Prozentsätzen der kombinierten Lösung sind möglich.

[0051] Gemäß einer Ausführungsform der vorliegenden Offenbarung verwendet unstarre, multimodale Bildregistrierung Vorinformation. Das Rahmenwerk erlaubt eine flexible Einstellung für die verfügbare Qualität von Vorkennntnis. Vorausgehende Experimente an synthetisch geschaffenen Phantomen und an realen MRI-SPECT/CT- und PET/CT-Daten zeigen, dass Vorkennntnis wichtig zum Ermitteln eines wünschenswerten grundlegenden Verschiebefelds sein kann.

[0052] Von Ausführungsformen eines Systems und eines Verfahrens zur multimodalen Bildregistrierung ist anzumerken, dass vom Fachmann im Lichte der obigen Lehren Modifikationen und Variationen vorgenommen werden können. Es versteht sich daher, dass Veränderungen an den speziellen Ausführungsformen der beschriebenen Erfindung gemacht werden können, die innerhalb des Umfangs und des Geistes der Erfindung liegen, wie sie von den beigefügten Ansprüchen definiert wird. Nachdem somit die Erfindung im Detail und speziell wie durch die Patentgesetze gefordert beschrieben ist, geben die beigefügten Ansprüche an, was durch das Patent beansprucht und geschützt werden soll.

Patentansprüche

1. Rechnergestütztes Verfahren zur unstarren, multimodalen Registrierung, umfassend:
Bereitstellen trainierter Daten entsprechend zuvor registrierter Bilder, die durch unterschiedliche Modalitäten erfasst wurden;
Empfangen zweier oder mehr Bilder, die durch die unterschiedlichen Modalitäten erfasst wurden; und
Registrieren der zwei oder mehr Bilder entsprechend Kontextinformation, die auf der Grundlage der trainierten Daten bestimmt wird.
2. Rechnergestütztes Verfahren nach Anspruch 1, bei dem das Registrieren der zwei oder mehr Bilder automatisch erfolgt.
3. Rechnergestütztes Verfahren nach Anspruch 1, bei dem die trainierten Daten eine Verbindungswahrscheinlichkeitsfunktion der unterschiedlichen Modalitäten sind.
4. Rechnergestütztes Verfahren nach Anspruch 3, bei dem die Verbindungswahrscheinlichkeitsfunktion eine Bedingung oder ein Regularisierungsfaktor in einem Optimierungsprozess der Registrierung ist.
5. Rechnergestütztes Verfahren nach Anspruch 1, bei dem die trainierten Daten ein mittlerer Bildort eines Bildintensitätspaares sind.
6. Rechnergestütztes Verfahren nach Anspruch 1, bei dem die trainierten Daten einer Gestaltsinformation sind, wobei das Verfahren umfasst:
Bestimmen einer oder mehrerer Gestalten in den zwei oder mehr Bildern; und
Klassifizieren der einen oder mehreren Gestalten als gültig oder ungültig gemäß den trainierten Daten.
7. Rechnergestütztes Verfahren nach Anspruch 1, bei dem eine Korrespondenz zwischen den zwei oder mehr Bildern unter Verwendung einer Kontextinformation von Gestalten in den unterschiedlichen Modalitäten der trainierten Daten bestimmt wird.
8. Rechnergestütztes Verfahren zur unstarren, multimodalen Registrierung, umfassend:
Bereitstellen trainierter Daten entsprechend einer Vorregistrierung;
Bereitstellen zweier oder mehrerer Bilder, die unter Verwendung unterschiedlicher Modalitäten erfasst wurden;
Bestimmen eines Verschiebefeldes der zwei oder mehr Bilder, wobei ein erstes Bild der zwei oder mehr Bilder als ein Bezugsbild definiert wird und ein nachfolgendes Bild als ein Ausrichtbild definiert wird, umfassend:
Abschätzen einer Randverteilung des Bezugsbildes und des Ausrichtbildes, und
Abschätzen einer Verbindungsintensitätsverteilung zwischen dem Bezugsbild und dem Ausrichtbild, und
Ausrichten der zwei oder mehr Bilder durch Einstellen der Verbindungsintensitätsverteilung, bis sie mit einer gelernten Verbindungsintensitätsverteilung der trainierten Daten im Wesentlichen übereinstimmt und gleichzeitig ein erwünschtes a-Ähnlichkeitsmaß zwischen zwei oder mehr Bildern und den trainierten Daten erhält.
9. Rechnergestütztes Verfahren nach Anspruch 8, bei dem die Ausrichtung der zwei oder mehr Bilder die Regularisierung des Verschiebefeldes zwischen den zwei oder mehr Bildern und den trainierten Daten umfasst.
10. Rechnergestütztes Verfahren nach Anspruch 9, bei dem die Regularisierung eine Tychonov-Regularisierung mit dem folgenden Gradienten verwendet:

$$\nabla_u R[u] = \operatorname{div} \left(\frac{R'[u]}{\nabla u} \nabla u \right) = \operatorname{div}(\nabla u) = \Delta u$$

wobei u das Verschiebefeld ist und R die Regularisierung an u definiert.

11. Rechnergestütztes Verfahren nach Anspruch 10, bei dem das Verfahren zum unstarren, multimodalen Registrieren gegeben ist durch:

$$\mathcal{J}(u) = \alpha \mathcal{L}_{MI}(u) + (1 - \alpha) \mathcal{L}_{KL}(u) + \lambda R(u), \quad \alpha \in [0, 1], \lambda \in \mathbb{R}_+$$

$$\hat{u} = \operatorname{argim} \mathcal{J}(u),$$

wobei λ eine positive Konstante ist, die einen Regularisierungsumfang bestimmt, \mathcal{L}_{KL} eine Divergenz zwi-

schen der Verbindungsintensitätsverteilung der zwei oder mehr Bilder und der Verbindungsintensitätsverteilung der trainierten Daten misst und \mathcal{I}_{MI} eine gegenseitige Information der Verbindungsintensitätsverteilung der zwei oder mehr Bilder bezeichnet.

12. Rechnergestütztes Verfahren nach Anspruch 8, bei dem das gewünschte a-Ähnlichkeitsmaß zwischen zwei oder mehr Bildern und den trainierten Daten ein Maximum ist.

13. Maschinenlesbare Programmspeichervorrichtung, die greifbar ein Programm von Befehlen ausbildet, das durch die Maschine ausführbar ist, um Verfahrensschritte für die unstarre, multimodale Registrierung durchzuführen, wobei das Verfahren die Schritte umfasst:

Bereitstellen trainierter Daten entsprechen zuvor registrierter Bilder, die durch unterschiedliche Modalitäten umfasst wurden;

Empfangen zweier oder mehr Bilder, die durch die unterschiedlichen Modalitäten erfasst wurden; und

Registrieren der zwei oder mehr Bilder entsprechend Kontextinformation, die auf der Basis der trainierten Daten bestimmt wird.

14. Verfahren nach Anspruch 13, bei dem das Registrieren der zwei oder mehr Bilder automatisch erfolgt.

15. Verfahren nach Anspruch 13, bei dem die trainierten Daten eine Verbindungswahrscheinlichkeitsfunktion der unterschiedlichen Modalitäten sind.

16. Verfahren nach Anspruch 15, bei dem die Verbindungswahrscheinlichkeitsfunktion eine Bedingung oder ein Regularisierungsfaktor in einem Optimierungsprozess der Registrierung ist.

17. Verfahren nach Anspruch 13, bei dem die trainierten Daten ein mittlerer Bildort eines Bildintensitätspaares sind.

18. Verfahren nach Anspruch 13, bei dem die trainierten Daten Gestaltsinformation sind, wobei das Verfahren umfasst:

Bestimmen einer oder mehrerer Gestalten in den zwei oder mehr Bildern; und

Klassifizieren der einen oder mehr Gestalten als gültig oder ungültig gemäß den trainierten Daten.

19. Verfahren nach Anspruch 13, bei dem eine Korrespondenz zwischen den zwei oder mehr Bildern unter Verwendung einer Kontextinformation von Gestalten in den unterschiedlichen Modalitäten der trainierten Daten bestimmt wird.

Es folgen 4 Blatt Zeichnungen

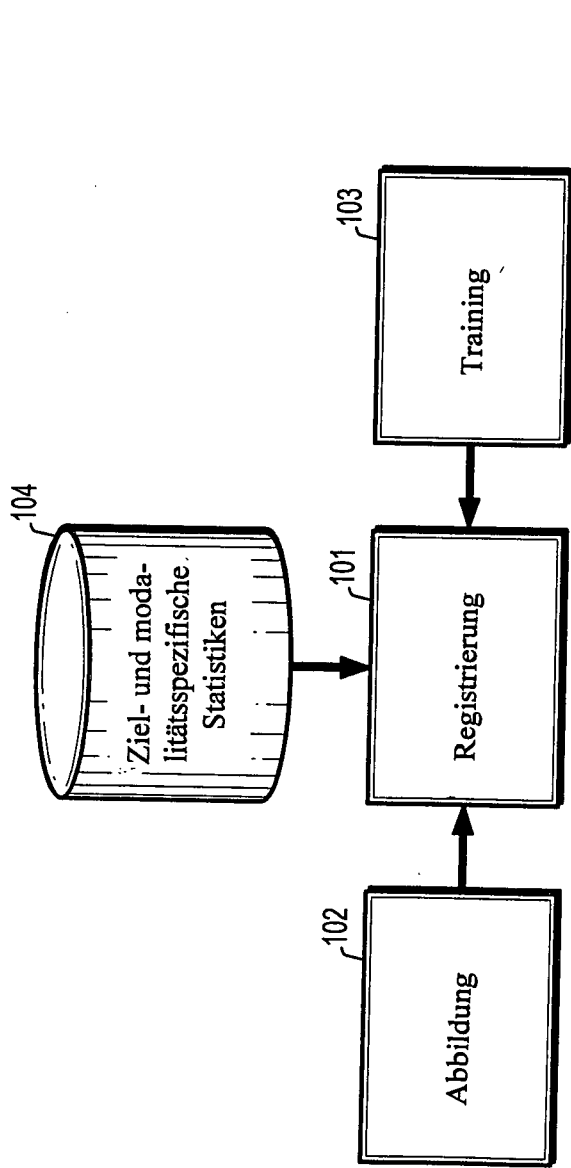


FIG. 1

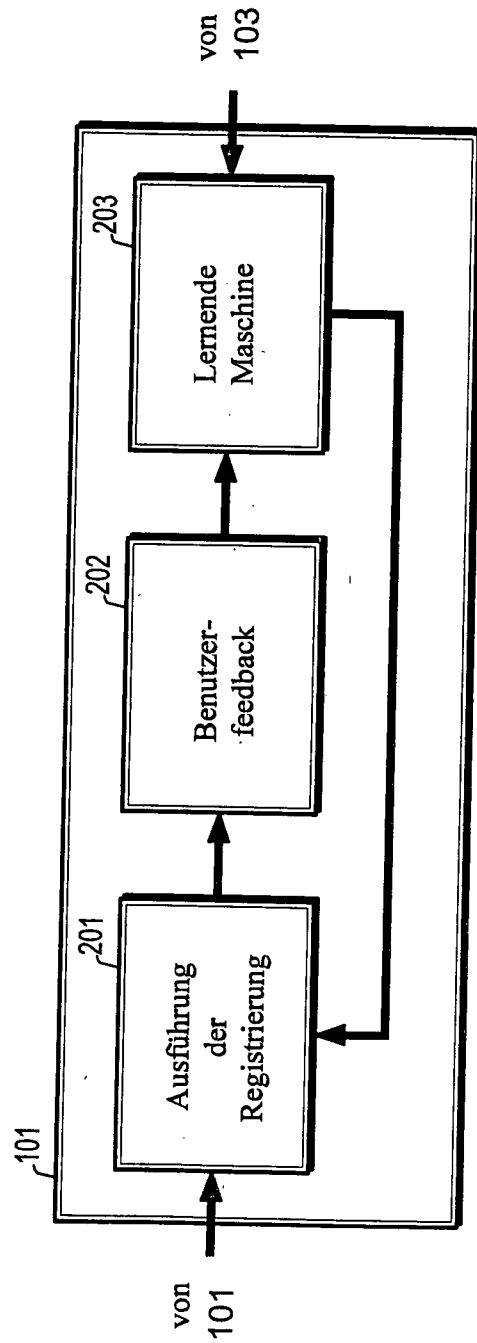


FIG. 2

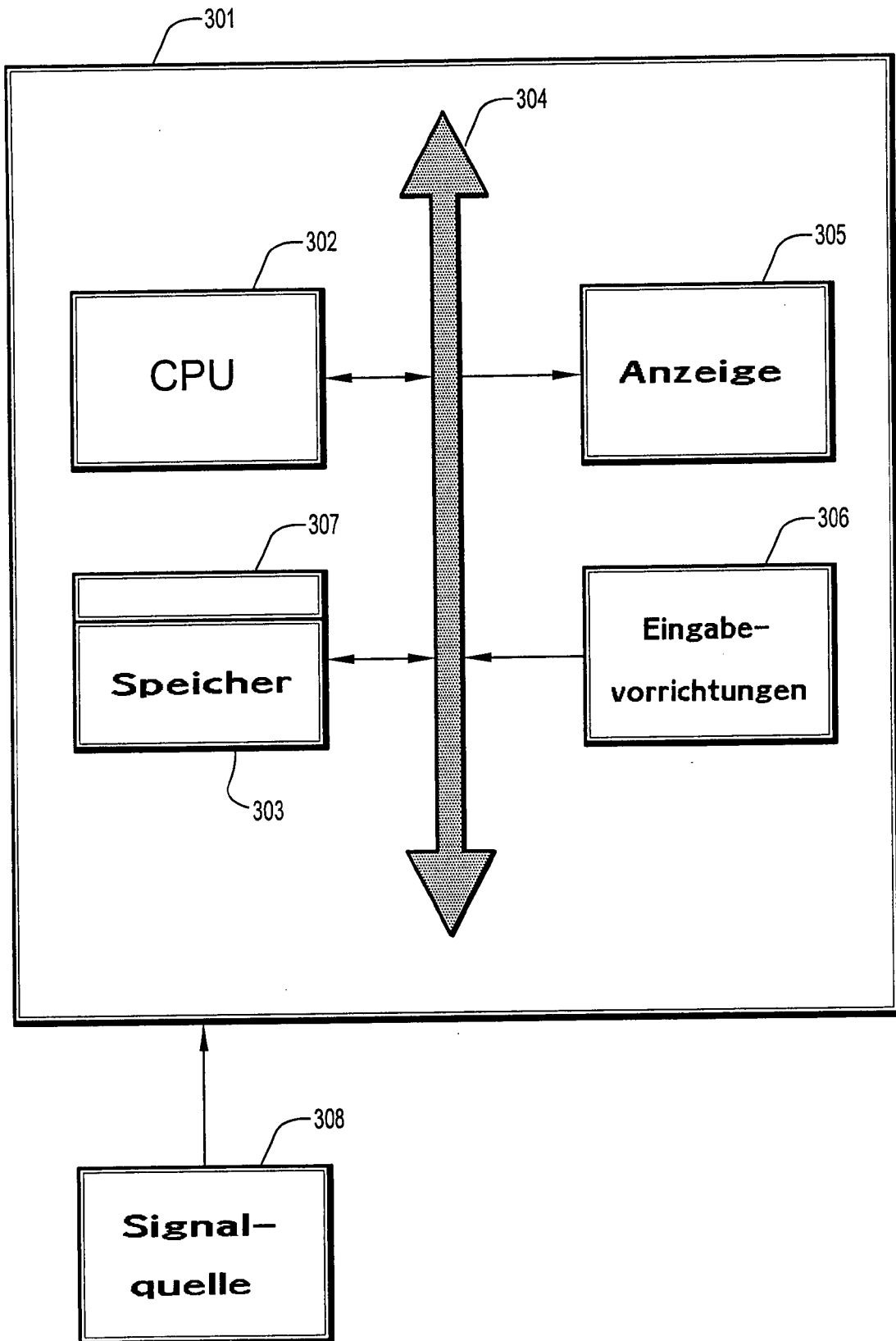


FIG. 3

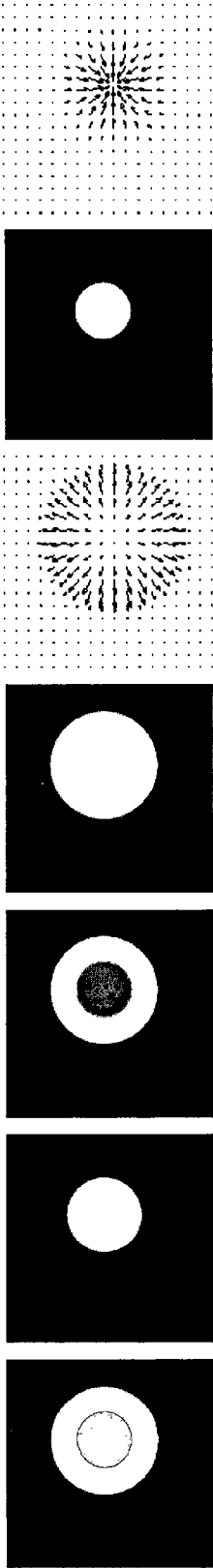


FIG. 4a FIG. 4b FIG. 4c FIG. 4d FIG. 4e FIG. 4f FIG. 4g

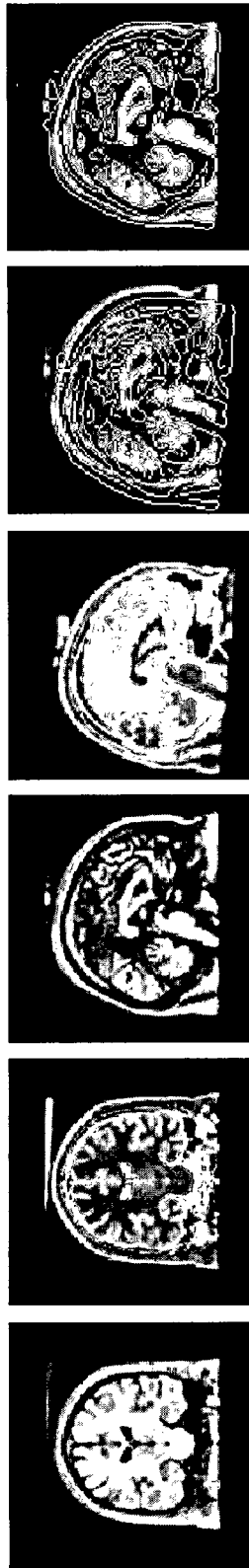


FIG. 5a FIG. 5b FIG. 5c FIG. 5d FIG. 5e FIG. 5f



FIG. 6a FIG. 6b FIG. 6c FIG. 6d

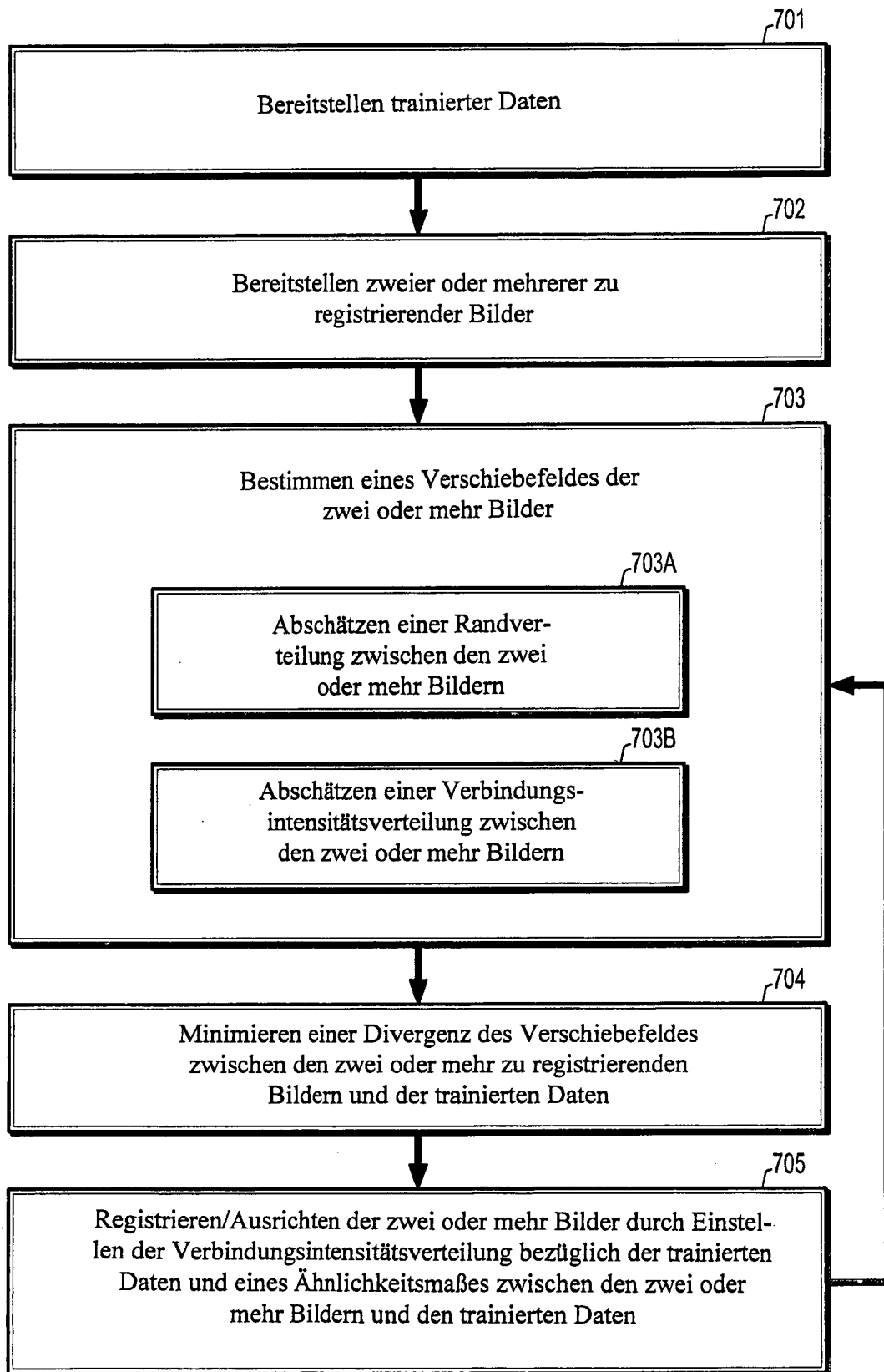


FIG. 7