



(12) **Patentschrift**

(21) Aktenzeichen: **10 2018 006 946.0**
 (22) Anmeldetag: **30.08.2018**
 (43) Offenlegungstag: **13.12.2018**
 (45) Veröffentlichungstag
 der Patenterteilung: **08.10.2020**

(51) Int Cl.: **B25J 9/18 (2006.01)**
B25J 9/22 (2006.01)

Innerhalb von neun Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

(30) Unionspriorität:
2017-112191 07.06.2017 JP

(73) Patentinhaber:
FANUC Corporation, Oshino-mura, Yamanashi, JP

(74) Vertreter:
Wuesthoff & Wuesthoff, Patentanwälte PartG mbB, 81541 München, DE

(72) Erfinder:
Ueda, Tetsuji, Oshino-mura, Yamanshi, JP

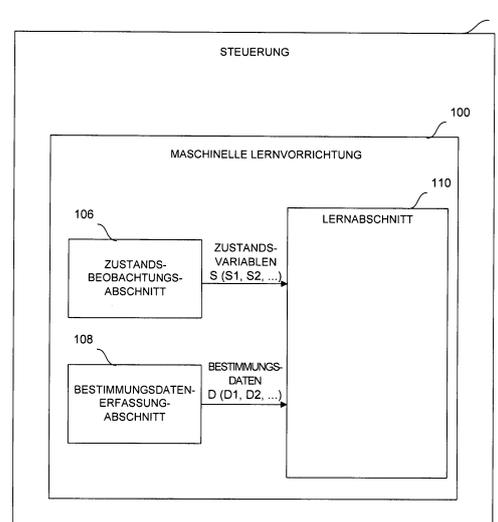
(56) Ermittelter Stand der Technik:

DE	10 2016 008 987	A1
DE	10 2016 009 030	A1
DE	10 2016 009 113	A1
DE	10 2016 014 155	A1
DE	699 25 204	T2
DE	11 2016 005 365	T5
JP	H06- 284 886	A

(54) Bezeichnung: **Steuerung und maschinelle Lernvorrichtung**

(57) Hauptanspruch: Steuerung (1; 2), die einen Kompensationsbetrag einer Einlernposition bei der Steuerung eines Roboters (160; 160') gemäß der in den Einlernpositionen enthaltenen Einlernposition bestimmt, wobei die Steuerung (1; 2) umfasst:
 eine maschinelle Lernvorrichtung (100; 120), die einen Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition lernt, wobei die maschinelle Lernvorrichtung (100; 120) umfasst:
 einen Zustandsbeobachtungsabschnitt (106), der als Zustandsgrößen, die einen aktuellen Zustand einer Umwelt ausdrücken, Einlernpositionskompensationsbetragsdaten, die den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition anzeigen, Motorstörungswertdaten, die einen Störungswert von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (106; 160) anzeigen, und Einlernpositionsdaten beobachtet, die als die Zustandsgrößen die Einlernposition der Einlernpositionen umfassen,
 einen Bestimmungsdaten-Erfassungsabschnitt (108), der Bestimmungsdaten erlangt, die ein Eignungsbestimmungsergebnis des Störungswertes von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (160; 160') anzeigen, und
 einen Lernabschnitt (110), der den Kompensationsbetrag der Einlernposition des Roboters (160; 160') in Zuordnung zu den Motorstörungswertdaten und der Einlernposition

unter Verwendung der Zustandsgrößen und der Bestimmungsdaten lernt.



Beschreibung

[0001] Die vorliegende Erfindung betrifft eine Steuerung und eine maschinelle Lernvorrichtung und insbesondere eine Steuerung und eine maschinelle Lernvorrichtung, die eine Einlernposition optimieren.

[0002] Allgemeine Industrieroboter werden gemäß einem zuvor erstellten Betriebsprogramm angetrieben oder so angetrieben, dass sie einen zuvor von einem Programmierhandgerät oder dergleichen eingelernten Einlernpunkt durchlaufen. Das heißt, die Roboter werden entlang einer vorgegebenen Bahn gefahren. Zum Beispiel offenbart die Druckschrift JP H06-284 886 A als verwandte Technik, die sich mit dem Einlernbetrieb eines Roboters befasst, eine Erfindung, die eine statistische Verarbeitung bezüglich der Abweichung zwischen tatsächlichen Positionsdaten und einem Positionierungszielwert von einer Steuerung durchführt und einen Kompensationsbetrag bestimmt, um die Abweichung aufzulösen.

[0003] Industrieroboter werden für verschiedene Zwecke eingesetzt, beispielsweise für die Installation von Werkstücken, die Entnahme von bearbeiteten Werkstücken und die Befestigung von Werkzeugen an Werkzeugmaschinen. Wenn ein Werkstück unter Verwendung eines Industrieroboters in einer Werkzeugmaschine installiert wird, wird der Vorgang des Haltens des Werkstücks und des Installierens des Werkstücks in einer Spannvorrichtung in einem Bearbeitungsbereich eingelernt, indem dem Roboter, wie oben beschrieben, ein Lehrpunkt eingelernt wird. Wenn zum Beispiel ein zylindrisches Werkstück in Bezug auf eine konvex geformte Spannvorrichtung, wie in **Fig. 9** gezeigt, eingestellt wird, kann das Werkstück so eingestellt werden, dass es von einer Annäherungsposition zu einer Zielposition in eine Richtung parallel zu der Richtung der Spannvorrichtung nach unten bewegt wird, wobei das Loch des Werkstücks parallel zu der Richtung der Spannvorrichtung ausgerichtet wird.

[0004] Wenn jedoch ein Bediener, der dem Roboter den Vorgang einlernt, ein Einlernbeginner ist, kann das Werkstück möglicherweise nicht präzise an der Annäherungsstelle positioniert werden, wird die Richtung des Werkstücks an der Annäherungsposition geneigt, und bleibt das Werkstück aufgrund der Abweichung einer Richtung, in die das Werkstück von der Annäherungsposition zu der Zielposition nach unten bewegt wird, an einer Spannvorrichtung hängen oder wird aufgrund des Kontakts zwischen dem Werkstück und der Vorrichtung eine Reibungskraft erzeugt, wodurch eine Störung (Last) in dem Handgelenkschaft oder dergleichen des Roboters auftreten könnte, der das Werkstück hält. Durch ein solches Einlernen wird jedes Gelenk des Roboters jedes Mal, wenn das Werkstück auf der Grundlage des eingelernten Vorgangs installiert wird, belastet, was die Ursache für Probleme oder einen Fehler mit dem Roboter darstellen kann, wenn der Betrieb des Roboters fortgesetzt wird.

[0005] Das Dokument DE 10 2016 014 155 A1 offenbart ein System mit einem Roboter zum Zuführen und Herausnehmen eines Werkstücks aus bzw. in eine(r) Werkzeugmaschine. Es wird für jede einer Mehrzahl von Roboterachsen jeweils ein Störeinfluss bestimmt, während das Werkstück zugeführt oder entnommen wird. Basierend darauf kann eine Roboterbewegung gestoppt und die Werkzeugmaschine bewegt werden, um letztendlich den Störeinfluss zu unterbinden.

[0006] Die Offenlegungsschrift DE 10 2016 009 113 A1 beschreibt eine maschinelle Lernvorrichtung für einen Roboter. Der Roboter beobachtet einen Umgebungszustand, um über seine Handlung zu entscheiden. Jedes Mal, wenn eine Handlung erfolgt ist, wird ein entsprechendes Belohnungssignal zurückgegeben. Hieraus kann der Roboter derart eingelernt werden, dass nur noch vorteilhafte Handlungen durchgeführt werden. Dies soll insbesondere das Zusammenarbeiten des Roboters mit einer Person erleichtern und effizienter gestalten.

[0007] Das Dokument DE 10 2016 009 030 A1 offenbart eine Vorrichtung für maschinelles Lernen. Eine Lage mehrerer ungeordnet übereinanderliegender Werkstücke wird mit einer dreidimensionalen Messvorrichtung ermittelt. Außerdem wird ein Zustand eines Roboters erfasst, der zum Greifen der Werkstücke ausgebildet ist. Basierend auf dem Erfolg oder Misserfolg eines Greifvorgangs wird eine Stellgröße gelernt, auf deren Basis der Roboter zum sicheren Greifen eines Werkstücks gesteuert werden kann.

[0008] Das Dokument DE 10 2016 008 987 A1 beschreibt ein Fehlervorhersagesystem mit einer Maschinenlernvorrichtung. Ein Zustand eines Roboters wird beobachtet und es wird überwacht, ob in dem Roboter ein Fehler aufgetreten ist und. Wenn dies der Fall ist, wird auch ein Fehlergrad bestimmt. Der beobachtete Zustand kann anschließend mit dem Fehler bzw. Fehlergrad von der Maschinenlernvorrichtung korreliert werden, um eine fehlerlose Bewegung des Roboters zu ermöglichen.

[0009] Die Druckschrift DE 11 2016 005 365 T5 offenbart ein Lernverfahren eines Roboters mit einem Master-Arm und einem Slave-Arm. Es wird eine Schwerkraftkompensation durchgeführt, sodass eine Person den Master-Arm auf einfache Weise in eine gewünschte Position bewegen kann. Die Lage des Roboters in der gewünschten Position wird gespeichert, wodurch bei einer späteren Steuerung des Roboters das Anfahren der gewünschten Position auch ohne Eingriff der Person ermöglicht wird.

[0010] Das Dokument DE 699 25 204 T2 offenbart eine Roboter-Steuerinrichtung. Eine Ausrichtung wird gespeichert als relative Beziehung zwischen einem Werkzeug-Koordinatensystem und einem Arbeits-Koordinatensystem der Umgebung. Durch Transformationen kann ein Bezug zwischen einem Roboter-Referenz-Koordinatensystem und den beiden anderen Koordinatensystemen bestimmt werden. Dies ermöglicht das Bewegen des Roboters entsprechend der gespeicherten Ausrichtung.

[0011] Angesichts des vorstehend beschriebenen Problems liegt der vorliegenden Erfindung die Aufgabe zugrunde, eine Steuerung und eine maschinelle Lernvorrichtung bereitzustellen, die in der Lage sind, eine Einlernposition zu optimieren, ohne eine unnötige Störung bei der Bedienung eines Roboters zu verursachen.

[0012] Zur Lösung des vorstehend beschriebenen Problems ist erfindungsgemäß eine Steuerung gemäß Anspruch 1 und eine maschinelle Lernvorrichtung gemäß Anspruch 8 vorgesehen.

[0013] Gemäß einer Ausführungsform führt die Steuerung das maschinelle Einlernen eines Kompensationsbetrags einer Einlernposition eines Roboters bezüglich einer Störung aus, die in einem Motor erzeugt wird, der jedes Gelenk des Roboters antreibt, und kompensiert und steuert die Einlernposition zur Reduzierung der Störung, wenn sich der Roboter in die Einlernposition bewegt, auf Grundlage eines Ergebnisses des maschinellen Lernens.

[0014] Eine Steuerung gemäß einer Ausführungsform der vorliegenden Erfindung bestimmt einen Kompensationsbetrag einer Einlernposition bei der Steuerung eines Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition. Die Steuerung umfasst eine maschinelle Lernvorrichtung, die einen Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters gemäß der Einlernposition lernt. Die maschinelle Lernvorrichtung umfasst einen Zustandsbeobachtungsabschnitt, der als Zustandsgrößen, die einen aktuellen Zustand einer Umwelt ausdrücken, Einlernpositionskompensationsbetragsdaten, die den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters gemäß der Einlernposition anzeigen, und Motorstörungswertdaten, die einen Störungswert von jedem der Motoren des Roboters bei der Steuerung des Roboters anzeigen, beobachtet, einen Bestimmungsdaten-Erfassungsabschnitt, der Bestimmungsdaten erlangt, die ein Eignungsbestimmungsergebnis des Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters anzeigen, und einen Lernabschnitt, der den Kompensationsbetrag der Einlernposition des Roboters in Zuordnung zu den Motorstörungswertdaten unter Verwendung der Zustandsgrößen und der Bestimmungsdaten lernt.

[0015] Der Zustandsbeobachtungsabschnitt kann ferner als die Zustandsgrößen Einlernpositionsdaten beobachten, die die Einlernposition der Einlerndaten umfassen, und der Lernabschnitt kann die von dem Zustandsbeobachtungsabschnitt beobachteten Daten in Zuordnung zu den Motorstörungswertdaten lernen.

[0016] Die Bestimmungsdaten können, außer einem Eignungsbestimmungsergebnis des Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters, mindestens eines von einem Eignungsbestimmungsergebnis einer Einlernposition, in die sich der Roboter schließlich bewegt, einem Eignungsbestimmungsergebnis eines von einem Sensor erfassten Werts und einem Eignungsbestimmungsergebnis von Zykluszeit bei der Steuerung des Roboters gemäß der Einlernposition nach der Kompensation umfassen.

[0017] Der Lernabschnitt kann einen Belohnungsberechnungsabschnitt, der eine Belohnung in Zuordnung zu dem Eignungsbestimmungsergebnis berechnet, und einen Wertfunktion-Aktualisierungsabschnitt, der unter Verwendung der Belohnung eine Funktion aktualisiert, die einen Wert des Kompensationsbetrags der Einlernposition relativ zu dem Störungswert von jedem der Motoren des Roboters bei der Steuerung des Roboters ausdrückt, haben.

[0018] Der Lernabschnitt kann eine Berechnung der Zustandsgrößen und der Bestimmungsdaten auf Grundlage einer mehrschichtigen Struktur durchführen.

[0019] Die Steuerung kann ferner einen Entscheidungsfindungsabschnitt umfassen, der einen Sollwert auf Grundlage des Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß der Einlernposition auf Grundlage eines Lernergebnisses des Lernabschnitts ausgibt.

[0020] Der Lernabschnitt kann den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters gemäß der Einlernposition in jedem von einer Mehrzahl von Robotern unter Verwendung der für jeden der Mehrzahl von Robotern erlangten Zustandsgrößen und Bestimmungsdaten lernen.

[0021] Die maschinelle Lernvorrichtung kann in einem Cloud-Server vorhanden sein.

[0022] Eine maschinelle Lernvorrichtung gemäß einer weiteren Ausführungsform der vorliegenden Erfindung lernt einen Kompensationsbetrag einer Einlernposition bei der Steuerung des Roboters gemäß der Einlernposition, die in den Einlerndaten enthalten ist. Die maschinelle Lernvorrichtung umfasst: einen Zustandsbeobachtungsabschnitt, der als Zustandsgrößen, die einen aktuellen Zustand einer Umwelt ausdrücken, Einlernpositionskompensationsbetragsdaten, die den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters gemäß der Einlernposition anzeigen, und Motorstörungswertdaten, die einen Störungswert von jedem der Motoren des Roboters bei der Steuerung des Roboters anzeigen, beobachtet; einen Bestimmungsdaten-Erfassungsabschnitt, der Bestimmungsdaten erlangt, die ein Eignungsbestimmungsergebnis des Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters anzeigen; und einen Lernabschnitt, der den Kompensationsbetrag der Einlernposition des Roboters in Zuordnung zu den Motorstörungswertdaten unter Verwendung der Zustandsgrößen und der Bestimmungsdaten lernt.

[0023] Gemäß einer Ausführungsform der vorliegenden Erfindung ist es möglich, einen Störungswert, der in einem Motor erzeugt wird, der jedes Gelenk eines Roboters antreibt, zu reduzieren und Probleme, wie zum Beispiel das Auftreten eines Fehlers in dem Roboter, durch die Kompensierung einer Einlernposition des Roboters auf Grundlage eines Lernergebnisses zu verhindern.

Fig. 1 ist ein schematisches Hardware-Konfigurationsdiagramm einer Steuerung gemäß einer ersten Ausführungsform;

Fig. 2 ist ein schematisches Funktionsblockschaltbild der Steuerung gemäß der ersten Ausführungsform;

Fig. 3 ist ein schematisches Funktionsblockschaltbild, das eine Ausführungsform der Steuerung zeigt;

Fig. 4 ist ein schematisches Flussdiagramm, das die Ausführungsform eines maschinellen Lernverfahrens zeigt;

Fig. 5A ist ein Diagramm zur Beschreibung eines Neurons;

Fig. 5B ist ein Diagramm zur Beschreibung eines neuronalen Netzes;

Fig. 6 ist ein schematisches Funktionsblockschaltbild einer Steuerung gemäß einer zweiten Ausführungsform;

Fig. 7 ist ein schematisches Funktionsblockschaltbild, das eine Ausführungsform eines Systems mit einer Mehrzahl von Robotern zeigt; und

Fig. 8 ist ein schematisches Funktionsblockschaltbild, das eine weitere Ausführungsform eines Systems mit einer Mehrzahl von Robotern zeigt.

Fig. 9 ist ein Diagramm zur Beschreibung eines Problems beim Einlernen gemäß einer verwandten Technik.

[0024] **Fig. 1** ist ein schematisches Hardware-Konfigurationsdiagramm, das eine Steuerung und die wesentlichen Teile einer von der Steuerung gemäß einer ersten Ausführungsform gesteuerten Werkzeugmaschine zeigt.

[0025] Eine Steuerung **1** kann z.B. als Steuerung zum Steuern eines Industrieroboters (nicht gezeigt) montiert werden, der die Installation eines Werkstücks, die Entnahme eines maschinell bearbeiteten Werkstücks, die Befestigung eines Werkzeugs an einer Werkzeugmaschine oder dergleichen durchführt. Eine Zentraleinheit (CPU) **11** der Steuerung **1** gemäß der Ausführungsform ist ein Prozessor, der die Steuerung **1** vollständig steuert. Die CPU **11** liest ein in einem Nur-Lese-Speicher (ROM) **12** gespeichertes Systemprogramm über einen Bus **20** und steuert die gesamte Steuerung **1** gemäß dem Systemprogramm. Ein Direktzugriffsspeicher (RAM) **13** speichert temporäre Berechnungsdaten oder Anzeigedaten und verschiedene Daten oder dergleichen, die von einem Bediener über ein später beschriebenes Programmierhandgerät **60** eingegeben werden.

[0026] Ein nichtflüchtiger Speicher **14** ist als ein Speicher ausgebildet, der seinen Speicherzustand beispielsweise durch Datensicherung oder dergleichen mit einer Batterie (nicht gezeigt) auch bei ausgeschalteter Steuerung **1** beibehält. Der nichtflüchtige Speicher **14** speichert die von dem Programmierhandgerät **60** über eine Schnittstelle **19** eingegebenen Einlerndaten, ein über eine Schnittstelle (nicht dargestellt) eingegebenes, den Roboter steuerndes Programm, oder dergleichen. Programme oder verschiedene Daten, die in dem nichtflüchtigen Speicher **14** gespeichert sind, können bei Ausführung/Verwendung in den RAM **13** entwickelt werden. Weiterhin speichert der ROM **12** vorab verschiedene Systemprogramme (ein Systemprogramm umfassend, um die Kommunikation mit einer maschinellen Lernvorrichtung **100**, die später beschrieben wird, zu steuern) zur laufenden Verarbeitung für die Steuerung eines Roboters oder das Einlernen einer Einlernposition oder dergleichen.

[0027] Das Programmierhandgerät **60** ist eine manuelle Dateneingabevorrichtung mit einer Anzeige, einem Griff, einem Hardware-Schlüssel oder dergleichen. Das Programmierhandgerät **60** empfängt über die Schnittstelle **19** Information von der Steuerung **1**, um diese anzuzeigen, und gibt Impulse, Befehle und verschiedene Dateneingaben von dem Griff, dem Hardware-Schlüssel oder dergleichen an die CPU **11** weiter.

[0028] Ein Achssteuerkreis **30** zur Steuerung der Achse eines Gelenks oder dergleichen eines Roboters empfängt von der CPU **11** einen Bewegungssollbetrag der Achse und gibt einen Befehl zum Bewegen der Achse an einen Servoverstärker **40** aus. Der Servoverstärker **40** steuert bei Erhalt des Befehls einen Servomotor **50** an, der die Achse des Roboters bewegt. Der Servomotor **50** für die Achse umfasst eine Positions-/Drehzahlerfassungsvorrichtung und gibt ein Positions-/Drehzahlrückmeldesignal von der Positions-/Drehzahlerfassungsvorrichtung an den Achssteuerkreis **30** zurück, um eine Positions-/Drehzahlregelung durchzuführen. Es wird angemerkt, dass der Achssteuerkreis **30**, der Servoverstärker **40** und der Servomotor **50** im Hardware-Konfigurationsdiagramm von **Fig. 1** einzeln dargestellt sind, aber tatsächlich entsprechend der Anzahl der Achsen eines zu steuernden Roboters vorgesehen sind. Bei einem Roboter mit sechs Achsen sind beispielsweise für jede der sechs Achsen der Achsregelkreis **30**, der Servoverstärker **40** und der Servomotor **50** vorgesehen.

[0029] Eine Schnittstelle **21** ist eine Schnittstelle, um die Steuerung **1** und die maschinelle Lernvorrichtung **100** miteinander zu verbinden. Die maschinelle Lernvorrichtung **100** umfasst einen Prozessor **101**, der die gesamte maschinelle Lernvorrichtung **100** steuert, einen ROM **102**, der ein Systemprogramm oder dergleichen speichert, einen RAM **103**, der Daten in jeder dem maschinellen Lernen zugeordneten Verarbeitung zwischenspeichert, und einen nichtflüchtigen Speicher **104**, der zum Speichern eines Lernmodells oder dergleichen verwendet wird. Die maschinelle Lernvorrichtung **100** kann jede Information (z.B. Positionsinformation oder den aktuellen Wert des Servomotors **50**, und Einstellinformation zu einem laufenden Programm, oder Einlerninformation oder dergleichen, die im RAM **13** oder dergleichen gespeichert sind) beobachten, die von der Steuerung **1** über die Schnittstelle **21** erfasst werden kann. Weiterhin führt die Steuerung **1** bei Empfang von Befehlen zur Steuerung des Servomotors **50** und der Peripherievorrichtung eines Roboters, die von der maschinellen Lernvorrichtung **100** ausgegeben werden, die Kompensation oder dergleichen eines Befehls zur Steuerung des Roboters basierend auf einem Programm oder Einlerndaten durch.

[0030] **Fig. 2** ist ein schematisches Funktionsblöckschaltbild der Steuerung **1** und der maschinellen Lernvorrichtung **100** gemäß der ersten Ausführungsform.

[0031] Die maschinelle Lernvorrichtung **100** umfasst Software (wie einen Lernalgorithmus) und Hardware (wie den Prozessor **101**) zum spontanen Erlernen eines Kompensationsbetrages einer Einlernposition eines Roboters bezüglich eines Störungswerts, der in einem Motor erzeugt wird, der jedes Gelenk des Roboters antreibt, durch sogenanntes maschinelles Lernen. Eine von der maschinellen Lernvorrichtung **100** der Steuerung **1** zu erlernende Aufgabe entspricht einer Modellstruktur, die die Korrelation zwischen einem Störungswert, der in einem Motor erzeugt wird, der jedes Gelenk eines Roboters antreibt, und einem Kompensationsbetrag einer Einlernposition des Roboters ausdrückt.

[0032] Wie in dem Funktionsblock von **Fig. 2** dargestellt, umfasst die maschinelle Lernvorrichtung **100** der Steuerung **1** einen Zustandsbeobachtungsabschnitt **106**, einen Bestimmungsdaten-Erfassungsabschnitt **108** und einen Lernabschnitt **110**. Der Zustandsbeobachtungsabschnitt **106** beobachtet Zustandsgrößen **S**, die den aktuellen Zustand einer Umwelt ausdrücken, die Einlernpositionskompensationsbetragsdaten **S1**, die einen Kompensationsbetrag einer Einlernposition von jedem von Motoren eines Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition anzeigen, und Motorstörungswertdaten **S2**, die einen Störungswert von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition anzeigen, umfassen. Der Bestimmungsdaten-Erfassungsabschnitt **108** erlangt Bestimmungsdaten **D**, die ein Eignungsbestimmungsergebnis des Störungswerts des

entsprechenden Motors bei der Steuerung des Roboters gemäß der kompensierten Einlernposition enthalten. Unter Verwendung der Zustandsgrößen **S** und der Bestimmungsdaten **D** lernt der Lernabschnitt **110** den Störungswert des entsprechenden Motors und die Einlernpositionskompensationsbetragsdaten **S1** in Zuordnung zueinander.

[0033] Der Zustandsbeobachtungsabschnitt **106** kann beispielsweise als eine der Funktionen des Prozessors **101** oder als in dem ROM **102** gespeicherte Software zum Betreiben des Prozessors **101** konfiguriert sein. Unter den Zustandsgrößen **S**, die von dem Zustandsbeobachtungsabschnitt **106** beobachtet werden, können die Einlernpositionskompensationsbetragsdaten **S1** als ein Kompensationsbetrag bezüglich einer Einlernposition (einem Satz von Koordinaten von jedem der Motoren eines Roboters) erlangt werden, die in den Einlerndaten enthalten sind. Die Einlernpositionskompensationsbetragsdaten **S1** können nur ein Kompensationsbetrag einer Einlernposition sein, in die sich ein Roboter schließlich bewegt (eine Endposition, wenn ein Werkstück in dem Beispiel in **Fig. 9** in eine Spannvorrichtung eingebracht wird). Abgesehen davon können die Einlernpositionskompensationsbetragsdaten **S1** einen Kompensationsbetrag einer Annäherungsposition enthalten, die eine Einlernposition anzeigt, bevor sich der Roboter schließlich bewegt. Außerdem kann, als die Einlernpositionskompensationsbetragsdaten **S1**, der Höchstwert eines Störungswerts, der in jedem Motor während der Bewegung in eine Einlernposition erzeugt wird, oder ein Serienwert, der erlangt wird, wenn der in dem entsprechenden Motor erzeugte Störungswert während der Bewegung in die Einlernposition für jeden vorgeschriebenen Zyklus (zum Beispiel 10 ms) erlangt wird, verwendet werden.

[0034] Als die Einlernpositionskompensationsbetragsdaten **S1** kann zum Beispiel ein Kompensationsbetrag einer Einlernposition von jedem der Motoren eines Roboters verwendet werden, der von einem Facharbeiter berichtet wird und in die Steuerung **1** eingegeben wird. Hier umfasst der Kompensationsbetrag der Einlernposition von jedem der Motoren des Roboters eine Motorkompensationsrichtung (einen positiven/negativen Wert), in die der Motor korrigiert wird. Außerdem kann, als die Einlernpositionskompensationsbetragsdaten **S1**, ein Kompensationsbetrag einer Einlernposition von jedem der Motoren eines Roboters, die durch die maschinelle Lernvorrichtung **100** im vorherigen Lernzyklus basierend auf einem Lernergebnis des Lernabschnitts **110** bestimmt wird, verwendet werden, wenn das Lernen bis zu einem gewissen Grad vorangeschritten ist. In einem solchen Fall kann die maschinelle Lernvorrichtung **100** vorab für jeden Lernzyklus einen bestimmten Kompensationsbetrag einer Einlernposition jedes der Motoren eines Roboters in dem RAM **103** zwischenspeichern, sodass der Zustandsbeobachtungsabschnitt **106** aus dem RAM **103** den Kompensationsbetrag der durch die maschinelle Lernvorrichtung **100** im vorherigen Lernzyklus ermittelten Einlernposition jedes der Motoren des Roboters erfasst.

[0035] Ferner kann, als die Motorstörungswertdaten **S2**, aus den Zustandsgrößen **S** zum Beispiel ein Ladungswert oder dergleichen, der aus einem Stromwert indirekt berechnet wird, der durch den Servomotor **50** fließt, bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition verwendet werden. Die Motorstörungswertdaten **S2** können direkt unter Verwendung einer Drehmomentmessung oder dergleichen berechnet werden, die an einem Motor befestigt ist.

[0036] Der Bestimmungsdaten-Erfassungsabschnitt **108** kann beispielsweise als eine der Funktionen des Prozessors **101** oder als in dem ROM **102** gespeicherte Software zum Betreiben des Prozessors **101** konfiguriert werden. Als Bestimmungsdaten **D** kann der Bestimmungsdaten-Erfassungsabschnitt **108** einen Eignungsbestimmungswert **D1** bezüglich eines Störungswerts von jedem Motor bei der Steuerung eines Roboters gemäß einer kompensierten Einlernposition verwenden. Unter Verwendung der gleichen Einrichtung, die verwendet wird, wenn der Zustandsbeobachtungsabschnitt **106** die Motorstörungswertdaten **S2** beobachtet, kann der Bestimmungsdaten-Erfassungsabschnitt **108** einen Störungswert von jedem Motor bei der Steuerung eines Roboters gemäß einer kompensierten Einlernposition erlangen. Die Bestimmungsdaten **D** sind ein Index, der ein Ergebnis ausdrückt, das erhalten wird, wenn ein Roboter gemäß einer Einlernposition unter den Zustandsgrößen **S** gesteuert wird.

[0037] Bezogen auf den Lernzyklus des Lernabschnitts **110** sind die gleichzeitig in den Lernabschnitt **110** eingegebenen Zustandsgrößen **S** diejenigen, die auf Daten des vorherigen Lernzyklus basieren, in dem die Bestimmungsdaten **D** erfasst wurden. Während, wie oben beschrieben, die maschinelle Lernvorrichtung **100** der Steuerung **1** das maschinelle Lernen vorantreibt, wird die Erfassung der Motorstörungsdaten **S2**, die Implementierung des Steuerns eines Roboters gemäß einer basierend auf den Einlernpositionskompensationsbetragsdaten **S1** kompensierten Einlernposition und die Erfassung der Bestimmungsdaten **D** in einer Umwelt wiederholt durchgeführt.

[0038] Der Lernabschnitt **110** kann beispielsweise als eine der Funktionen des Prozessors **101** oder als in dem ROM **102** gespeicherte Software zum Betreiben des Prozessors **101** konfiguriert werden. Gemäß einem beliebigen als maschinelles Lernen bezeichneten Lernalgorithmus lernt der Lernabschnitt **110** die Einlernpositionskompensationsbetragsdaten **S1** bezüglich eines Störungswerts von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition. Der Lernabschnitt **110** kann das Lernen basierend auf einem Datensatz mit den oben beschriebenen Zustandsgrößen **S** und den Bestimmungsdaten **D** wiederholt ausführen.

[0039] Wenn der Zyklus des Lernens der Einlernpositionskompensationsbetragsdaten **S1** bezüglich eines Störungswerts von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition wiederholt ausgeführt wird, entsprechen, von den Zustandsgrößen **S**, die Motorstörungswertdaten **S2** und die Einlernpositionskompensationsbetragsdaten **S1** einem Störungswert von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition, die in dem vorherigen Lernzyklus erlangt werden, wie vorstehend beschrieben, bzw. einem Kompensationsbetrag einer Einlernposition von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten, die auf Grundlage von vergangenen Lernergebnissen erlangt werden, enthaltenen Einlernposition. Außerdem entsprechen die Bestimmungsdaten **D** einem Eignungsbestimmungsergebnis bezüglich eines Störungswerts von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer (kompensierten) Einlernposition in einem aktuellen Lernzyklus in einem Zustand, in dem die Einlernposition auf Grundlage der Einlernpositionskompensationsbetragsdaten **S1** kompensiert wurde.

[0040] Durch wiederholtes Ausführen eines solchen Lernzyklus' kann der Lernabschnitt **110** automatisch ein Merkmal identifizieren, das die Korrelation zwischen einem Störungswert (den Motorstörungswertdaten **S2**) von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition und einem Kompensationsbetrag bezüglich der Einlernposition nahelegt. Obgleich die Korrelation zwischen den Motorstörungswertdaten **S2** und einem Kompensationsbetrag bezüglich einer Einlernposition zu Beginn eines Lernalgorithmus im Wesentlichen unbekannt ist, identifiziert der Lernabschnitt **110** schrittweise ein Merkmal, das die Korrelation anzeigt, und interpretiert die Korrelation, während das Lernen fortschreitet. Wenn die Korrelation zwischen den Motorstörungswertdaten **S2** und einem Kompensationsbetrag bezüglich einer Einlernposition bis zu einem gewissen zuverlässigen Grad interpretiert wird, können die von dem Lernabschnitt **110** wiederholt ausgegebenen Lernergebnisse zur Auswahl der Aktion (d.h., Entscheidungsfindung) verwendet werden, um zu bestimmen, in welchem Maße ein Kompensationsbetrag einer Einlernposition in Bezug auf einen aktuellen Zustand (d.h. einen Störungswert von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition) angepasst wird. Das heißt, wenn ein Lernalgorithmus fortgeschritten ist, kann sich der Lernabschnitt **110** der Korrelation zwischen einem Störungswert eines jeden der Motoren des Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition und der Aktion des Bestimmens, in welchem Maße ein Kompensationsbetrag der Einlernposition des Roboters in Bezug auf den Zustand auf eine optimale Lösung angepasst wird, schrittweise annähern.

[0041] Wie oben beschrieben, lernt der Lernabschnitt **110** in der maschinellen Lernvorrichtung **100** der Steuerung **1** einen Kompensationsbetrag einer Einlernposition eines Roboters gemäß einem maschinellen Lernalgorithmus unter Verwendung der Zustandsgrößen **S**, die von dem Zustandsbeobachtungsabschnitt **106** beobachtet werden, und der Bestimmungsdaten **D**, die von dem Bestimmungsdaten-Erfassungsabschnitt **108** erfasst werden. Die Zustandsgrößen **S** setzen sich aus den Einlernpositionskompensationsbetragsdaten **S1** und den Motorstörungswertdaten **S2** zusammen, die durch eine Störung kaum beeinflusst werden. Außerdem werden die Bestimmungsdaten **D** durch die Erfassung eines aktuellen Werts des Servomotors **50**, der tatsächlich von der Steuerung **1** oder dergleichen gemessen wird, eindeutig berechnet. Dementsprechend kann die maschinelle Lernvorrichtung **100** der Steuerung **1** unter Verwendung von Lernergebnissen des Lernabschnitts **110** automatisch und genau einen Kompensationsbetrag einer Einlernposition eines Roboters gemäß einem Störungswert von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß den in den Einlerndaten des Roboters enthaltenen Einlernposition berechnen, ohne sich auf Berechnung oder Schätzung zu verlassen.

[0042] Wenn es möglich ist, einen Kompensationsbetrag einer Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition zu berechnen, ohne sich auf eine Berechnung oder Schätzung zu verlassen, kann ein geeigneter Wert des Kompensationsbetrags der Einlernposition des Roboters schnell bestimmt werden, indem nur ein Störungswert (die Motorstörungswertdaten **S2**) von jedem Motor des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des

Roboters enthaltenen Einlernposition erlangt werden. Dementsprechend kann ein Kompensationsbetrag einer Einlernposition eines Roboters effektiv bestimmt werden.

[0043] Als ein erstes modifiziertes Beispiel der maschinellen Lernvorrichtung **100** der Steuerung **1** kann der Zustandsbeobachtungsabschnitt **106** ferner Einlernpositionsdaten **S3** beobachten, die als die Zustandsgrößen **S** eine Einlernposition von Einlerndaten umfassen. Die Einlernpositionsdaten **S3** können als eine Einlernposition (ein Satz von Koordinaten-Werten von jedem der Motoren eines Roboters) erlangt werden, die in den Einlerndaten enthalten sind. Die Einlernpositionsdaten **S3** können z.B. nur eine Einlernposition sein, in die sich ein Roboter schließlich bewegt (eine Endposition, wenn ein Werkstück in einer Spannungsvorrichtung in dem Beispiel aus **Fig. 9** eingebracht wird). Abgesehen davon können die Einlernpositionsdaten **S3** eine Annäherungsposition umfassen, die eine Einlernposition anzeigt, bevor sich der Roboter schließlich bewegt.

[0044] In dem vorstehend beschriebenen modifizierten Beispiel kann die maschinelle Lernvorrichtung **100** einen Kompensationsbetrag bezüglich einer Einlernposition für sowohl die Einlernposition als auch einen Störungswert von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition lernen. Dadurch kann, da es möglich ist, auf Grundlage von Einlerndaten einen Kompensationsbetrag einer Einlernposition in einen geeigneten Wert zu ändern, während die Haltung eines Roboters bei der Steuerung des Roboters berücksichtigt wird, die Störung des Motors des Roboters geeigneter reduziert werden.

[0045] Als ein zweites modifiziertes Beispiel der maschinellen Lernvorrichtung **100** der Steuerung **1** kann der Bestimmungsdaten-Erfassungsabschnitt **108**, zusätzlich zu dem Eignungsbestimmungswert **D1** bezüglich eines Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition, als die Bestimmungsdaten **D** einen Eignungsbestimmungswert **D2** einer Einlernposition, in die sich ein Roboter schließlich bewegt, einen Eignungsbestimmungswert **D3** eines Wertes, der von einem Sensor, wie beispielsweise einem Schwingungssensor und einem Geräuschsensor, erfasst wird, einen Evaluationswert **D4** von Zykluszeit bei der Steuerung des Roboters gemäß einer kompensierten Einlernposition oder dergleichen verwenden.

[0046] Gemäß dem vorstehend beschriebenen modifizierten Beispiel kann die maschinelle Lernvorrichtung **100** einen Abweichungsgrad einer Einlernposition, das Auftreten von Vibrationen oder anormalen Geräuschen, eine Zunahme der Zykluszeit oder dergleichen beim Lernen eines Kompensationsbetrags einer Einlernposition eines Roboters bezüglich eines Störungswerts von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition berücksichtigen.

[0047] Als ein drittes modifiziertes Beispiel für die maschinelle Lernvorrichtung **100** der Steuerung **1** kann der Lernabschnitt **110** unter Verwendung der ermittelten Zustandsgrößen **S** und Bestimmungsdaten **D** für jeden der Mehrzahl von Robotern, der die gleiche Arbeit ausführt, einen Kompensationsbetrag einer Einlernposition bei der Ausführung von Steuerung gemäß der in den Einlerndaten enthaltenen Einlernposition einer Mehrzahl von Robotern lernen. Gemäß der Konfiguration ist es möglich, eine Menge eines Datensatzes, der die in einem bestimmten Zeitraum erfassten Zustandsgrößen **S** und Bestimmungsdaten **D** umfasst, zu erhöhen. Daher kann die Geschwindigkeit und die Zuverlässigkeit des Erlernens eines Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß der in den Einlerndaten eines Roboters enthaltenen Einlernposition mit einem Satz von vielfältigeren Daten als Eingaben verbessert werden.

[0048] In der maschinellen Lernvorrichtung **100** mit der obigen Konfiguration ist ein von dem Lernabschnitt **110** ausgeführter Lernalgorithmus nicht speziell eingeschränkt. Beispielsweise kann ein als maschinelles Lernen bekannter Lernalgorithmus eingesetzt werden. **Fig. 3** zeigt als eine Ausführungsform der in **Fig. 1** gezeigten Steuerung **1** eine Konfiguration, die den Lernabschnitt **110** umfasst, der das verstärkende Lernen als Beispiel für einen Lernalgorithmus durchführt.

[0049] Das verstärkende Lernen ist ein Verfahren, bei dem, während der aktuelle Zustand (d.h. eine Eingabe) einer Umwelt, in der ein Lernziel existiert, beobachtet wird, eine vorgeschriebene Aktion (d.h. eine Ausgabe) in dem aktuellen Zustand durchgeführt wird und der Zyklus des Vergebens einer Belohnung der Aktion wiederholt durch Ausprobieren durchgeführt wird, um Maßnahmen zu lernen (ein Kompensationsbetrag einer Einlernposition bei der Steuerung eines Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition in der maschinellen Lernvorrichtung der vorliegenden Anwendung), um die Summe der Belohnungen als eine optimale Lösung zu maximieren.

[0050] In der maschinellen Lernvorrichtung **100** der in **Fig. 3** gezeigten Steuerung **1** umfasst der Lernabschnitt **110** einen Belohnungsberechnungsabschnitt **112** und einen Wertfunktion-Aktualisierungsabschnitt **114**. Der Belohnungsberechnungsabschnitt **112** berechnet eine Belohnung R , die einem Eignungsbestimmungsergebnis (entsprechend den im nächsten Lernzyklus, in dem die Zustandsgrößen \mathbf{S} erfasst wurden, verwendeten Bestimmungsdaten \mathbf{D}) eines Störungswerts jedes der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer kompensierten Einlernposition basierend auf den Zustandsgrößen \mathbf{S} zugeordnet wird. Der Wertfunktion-Aktualisierungsabschnitt **114** aktualisiert unter Verwendung der berechneten Belohnung R eine Funktion Q , die einen Wert eines Kompensationsbetrags einer Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition ausdrückt. Der Lernabschnitt **110** lernt einen Kompensationsbetrag einer Einlernposition eines Roboters bezüglich eines Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition derart, dass der Wertfunktion-Aktualisierungsabschnitt **114** die Funktion Q wiederholt aktualisiert.

[0051] Ein Beispiel für einen Lernalgorithmus des verstärkenden Lernens, der von dem Lernabschnitt **110** durchgeführt wird, wird beschrieben. Der Algorithmus in diesem Beispiel ist als Q -Lernen bekannt und drückt ein Verfahren aus, bei dem ein Zustand s eines Aktionssubjekts und eine möglicherweise von dem Aktionssubjekt im Zustand s durchgeführte Aktion a als unabhängige Größen angenommen werden und eine Funktion $Q(s, a)$, die einen Aktionswert ausdrückt, wenn die Aktion a im Zustand s ausgewählt wird, gelernt wird. Die Auswahl der Aktion a , bei der die Wertfunktion Q im Zustand s die größte wird, führt zu einer optimalen Lösung. Durch Starten des Q -Lernens in einem Zustand, in dem die Korrelation zwischen dem Zustand s und der Aktion a unbekannt ist, und wiederholtes Ausführen der Auswahl verschiedener Aktionen a durch Ausprobieren in einem beliebigen Zustand s , wird die Wertfunktion Q wiederholt aktualisiert, um einer optimalen Lösung angenähert zu werden. Wenn sich hierbei eine Umwelt (d.h. der Zustand s) ändert, während die Aktion a in dem Zustand s ausgewählt wird, wird eine Belohnung (d.h. die Gewichtung der Aktion a) r entsprechend der Änderung erhalten und wird das Lernen darauf ausgerichtet, eine Aktion a auszuwählen, durch die eine höhere Belohnung r erhalten wird. So kann die Wertfunktion Q in einem relativ kurzen Zeitraum einer optimalen Lösung angenähert werden.

[0052] Im Allgemeinen kann die Aktualisierungsformel der Wertfunktion Q wie die folgende Formel (1) ausgedrückt werden. In der Formel (1) drücken s_t und a_t einen Zustand bzw. eine Aktion zum Zeitpunkt t aus, und ändert sich der Zustand zu s_{t+1} mit der Aktion a_t . r_{t+1} drückt eine Belohnung aus, die erhalten wird, wenn sich der Zustand von s_t zu s_{t+1} ändert. Der Begriff $\max Q$ drückt Q in einem Fall aus, in dem eine Aktion a ausgeführt wird, durch die der maximale Wert Q zum Zeitpunkt $t + 1$ (der zum Zeitpunkt t angenommen wird) erreicht wird. α und γ drücken einen Lernkoeffizienten beziehungsweise einen Diskontierungsfaktor aus und werden willkürlich so eingestellt, dass sie in $0 < \alpha \leq 1$ beziehungsweise $0 < \gamma \leq 1$ fallen.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right) \quad (1)$$

[0053] Wenn der Lernabschnitt **110** das Q -Lernen durchführt, entsprechen die von dem Zustandsbeobachtungsabschnitt **106** beobachteten Zustandsgrößen \mathbf{S} und die von dem Zustandsdaten-Erfassungsabschnitt **108** erfassten Bestimmungsdaten \mathbf{D} dem Zustand s in dieser Aktualisierungsformel, die Aktion des Bestimmens eines Kompensationsbetrags einer Einlernposition eines Roboters in Bezug auf einen aktuellen Zustand (d. h., einen Störungswert von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition) entspricht der Aktion a in der Aktualisierungsformel und die durch den Belohnungsberechnungsabschnitt **112** berechnete Belohnung R entspricht der Belohnung r in der Aktualisierungsformel. Dementsprechend aktualisiert der Wertfunktion-Aktualisierungsabschnitt **114** wiederholt die Funktion Q , die einen Wert eines Kompensationsbetrags einer Einlernposition eines Roboters in Bezug auf einen aktuellen Zustand durch das Q -Lernen ausdrückt, unter Verwendung der Belohnung R .

[0054] Wenn der Roboter gemäß der Einlernposition gesteuert wird, die auf Grundlage eines Kompensationsbetrags der Einlernposition kompensiert wird, die nach der Bestimmung des Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition bestimmt wird, kann die von dem Belohnungsberechnungsabschnitt **112** berechnete Belohnung R positiv sein, zum Beispiel, wenn ein Eignungsbestimmungsergebnis des Betriebs des Roboters als „geeignet“ bestimmt wird (zum Beispiel wenn ein Störungswert von jedem der Motoren eines Roboters in einen zulässigen Bereich fällt oder wenn eine Einlernposition, in die sich der Roboter schließlich bewegt, in einen zulässigen Bereich fällt) oder kann negativ sein, zum Beispiel, wenn das Eignungsbestimmungsergebnis des Betriebs

des Roboters als „ungeeignet“ bestimmt wird (zum Beispiel wenn ein Störungswert von jedem der Motoren des Roboters über den zulässigen Bereich hinausgeht oder wenn eine Einlernposition, in die sich der Roboter schließlich bewegt, über den zulässigen Bereich hinausgeht).

[0055] Die Absolutwerte der positiven und negativen Belohnungen R können bei der Bestimmung der Belohnungen R gleich oder verschieden sein. Außerdem können als Bestimmungsbedingungen eine Mehrzahl von in den Bestimmungsdaten D umfassten Werten miteinander kombiniert werden, um eine Bestimmung durchzuführen.

[0056] Außerdem kann ein Eignungsbestimmungsergebnis des Betriebs eines Roboters nicht nur „geeignete“ und „ungeeignete“ Ergebnisse, sondern auch eine Mehrzahl von Ergebnisebenen umfassen. Wenn beispielsweise ein Maximalwert innerhalb eines zulässigen Bereichs eines Störungswerts jedes der Motoren als G_{\max} angenommen wird, wird die Belohnung $R = 5$ vergeben, wenn der Störungswert V von jedem der Motoren des Roboters innerhalb von $0 \leq V < V_{\max}/5$ fällt, wird die Belohnung $R = 2$ vergeben, wenn der Störungswert innerhalb von $V_{\max}/5 \leq V < V_{\max}/2$ fällt, und wird die Belohnung $R = 1$ vergeben, wenn der Störungswert V innerhalb von $V_{\max}/2 \leq V \leq V_{\max}$ liegt. Außerdem kann V_{\max} in der Anfangsphase des Lernens relativ größer eingestellt werden und so eingestellt werden, dass er mit zunehmendem Lernfortschritt abnimmt.

[0057] Der Wertfunktion-Aktualisierungsabschnitt **114** kann eine Aktionswerttabelle aufweisen, in der die Zustandsgrößen S , die Bestimmungsdaten D und die Belohnungen R in Zuordnung zu den durch die Funktion Q ausgedrückten Aktionswerten (z.B. numerische Werte) organisiert sind. In diesem Fall ist die Aktion des Aktualisierens der Funktion Q mit dem Wertfunktion-Aktualisierungsabschnitt **114** gleichbedeutend mit der Aktion des Aktualisierens der Aktionswerttabelle mit dem Wertfunktion-Aktualisierungsabschnitt **114**. Zu Beginn des Q -Lernens ist die Korrelation zwischen dem aktuellen Zustand einer Umwelt und einem Kompensationsbetrag einer Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition unbekannt. Daher werden in der Aktionswerttabelle verschiedene Arten der Zustandsgrößen S , die Bestimmungsdaten D und die Belohnungen R in Zuordnung zu Werten (Funktion Q) von zufällig festgelegten Aktionswerten aufbereitet. Es wird angemerkt, dass der Belohnungsberechnungsabschnitt **112** die Belohnungen R entsprechend den Bestimmungsdaten D sofort berechnen kann, wenn die Bestimmungsdaten D bekannt sind, und die Werte der berechneten Belohnungen R werden in die Aktionswerttabelle geschrieben.

[0058] Wenn das Q -Lernen unter Verwendung der Belohnung R , die einem Eignungsbestimmungsergebnis des Störungswerts von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition entspricht, vorangetrieben wird, wird das Lernen darauf ausgerichtet, die Aktion des Erhaltens einer höheren Belohnung R auszuwählen. Dann werden Werte (Funktion Q) von Aktionswerten für eine in einem aktuellen Zustand ausgeführte Aktion erneut eingeschrieben, um die Aktionswerttabelle entsprechend dem Zustand einer Umwelt (d.h. den Zustandsgrößen S und den Bestimmungsdaten D), die sich ändert, zu aktualisieren, wenn die ausgewählte Aktion im aktuellen Zustand ausgeführt wird. Durch wiederholtes Ausführen der Aktualisierung werden die Werte (die Funktion Q) der in der Aktionswerttabelle angezeigten Aktionswerte erneut größer eingeschrieben, wenn eine Aktion geeigneter ist. So wird der Zusammenhang zwischen einem aktuellen Zustand (einem Störungswert von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition) in einer unbekanntem Umwelt und einer entsprechenden Aktion (Bestimmung eines Kompensationsbetrags der Einlernposition des Roboters) schrittweise offensichtlich. Das heißt, durch die Aktualisierung der Aktionswerttabelle wird das Verhältnis zwischen einem Störungswert von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition und der Bestimmung eines Kompensationsbetrags der Einlernposition des Roboters schrittweise einer optimalen Lösung angenähert.

[0059] Der Ablauf des obigen Q -Lernens (d.h. die Ausführungsform eines maschinellen Lernverfahrens) durch den Lernabschnitt **110** wird mit Bezugnahme auf **Fig. 4** näher beschrieben.

[0060] Zunächst wählt der Wertfunktion-Aktualisierungsabschnitt **114** in Schritt **SA01** unter Bezugnahme auf eine Aktionswerttabelle zu diesem Zeitpunkt willkürlich einen Kompensationsbetrag einer Einlernposition eines Roboters als eine Aktion aus, die in einem aktuellen Zustand ausgeführt wird, der durch die Zustandsgrößen S angezeigt wird, die durch den Zustandsbeobachtungsabschnitt **106** beobachtet werden. Anschließend importiert der Wertfunktion-Aktualisierungsabschnitt **114** die Zustandsgröße S in dem durch den Zustandsbeobachtungsabschnitt **106** in Schritt **SA02** beobachteten aktuellen Zustand und importiert die Bestimmungsdaten D in dem durch den Bestimmungsdaten-Erfassungsabschnitt **108** in Schritt **SA03** erfassten aktuellen Zustand. In Schritt **SA04** bestimmt dann der Wertfunktion-Aktualisierungsabschnitt **114** basierend auf den

Bestimmungsdaten **D**, ob der Kompensationsbetrag der Einlernposition des Roboters geeignet ist oder nicht. Wenn der Kompensationsbetrag geeignet ist, wendet der Wertfunktion-Aktualisierungsabschnitt **114** in Schritt **SA04** eine durch den Belohnungsberechnungsabschnitt **112** berechnete positive Belohnung R auf die Aktualisierungsformel der Funktion Q an. Als nächstes aktualisiert in Schritt **SA06** der Wertfunktion-Aktualisierungsabschnitt **114** die Aktionswerttabelle unter Verwendung der Zustandsgröße **S** und der Bestimmungsdaten **D** in dem aktuellen Zustand, der Belohnung R und einem Wert (aktualisierte Funktion Q) eines Aktionswertes. Wenn in Schritt **SA04** bestimmt wird, dass der Kompensationsbetrag der Einlernposition des Roboters ungeeignet ist, wendet der Wertfunktion-Aktualisierungsabschnitt **114** in Schritt **SA07** eine von dem Belohnungsberechnungsabschnitt **112** berechnete negative Belohnung R auf die Aktualisierungsformel der Funktion Q an. Im Schritt **SA06** aktualisiert der Wertfunktion-Aktualisierungsabschnitt **114** dann die Aktionswerttabelle unter Verwendung der Zustandsgröße **S** und der Bestimmungsdaten **D** im aktuellen Zustand, der Belohnung R und des Werts (aktualisierte Funktion Q) des Aktionswertes. Der Lernabschnitt **110** aktualisiert die Aktionswerttabelle erneut, indem er wiederholt die obige Verarbeitung der Schritte **SA01** bis **SA07** durchführt und das Erlernen des Kompensationsbetrags der Einlernposition des Roboters vorantreibt. Es wird angemerkt, dass die Verarbeitung der Berechnung der Belohnungen R und die Verarbeitung der Aktualisierung der Wertfunktion in den Schritten **SA04** bis **SA07** für alle der in den Ermittlungsdaten **D** enthaltenen Daten durchgeführt werden.

[0061] Um das oben beschriebene Verstärkungslernen voranzutreiben, kann anstelle von beispielsweise dem Q -Lernen ein neuronales Netzwerk verwendet werden. **Fig. 5A** zeigt schematisch ein Neuronenmodell. **Fig. 5B** zeigt schematisch das Modell eines neuronalen Netzwerks mit drei Schichten, in denen die in **Fig. 5A** gezeigten Neuronen miteinander kombiniert sind. Das neuronale Netzwerk kann beispielsweise einem Neuronenmodell folgend durch eine Recheneinheit oder eine Speichereinheit konfiguriert werden.

[0062] Das in **Fig. 5A** dargestellte Neuron gibt ein Ergebnis y in Bezug auf eine Mehrzahl von Eingaben x aus (hier beispielsweise Eingaben x_1 bis x_3). Die Eingaben x_1 bis x_3 werden jeweils mit entsprechenden Gewichten w (w_1 bis w_3) multipliziert. Das Neuron gibt also das durch die folgende Formel 2 ausgedrückte Ergebnis y aus. Es wird angemerkt, dass in der folgenden Formel 2 eine Eingabe x , ein Ergebnis y und ein Gewicht w alle Vektoren sind. Zusätzlich drückt θ einen Bias aus und f_k drückt eine Aktivierungsfunktion aus.

$$y = f_k \left(\sum_{i=1}^n x_i w_i - \theta \right) \quad (2)$$

[0063] In dem neuronalen Netz mit den drei in **Fig. 5B** gezeigten Schichten werden eine Mehrzahl von Eingaben x (hier die Eingaben x_1 bis x_3 als Beispiel) von der linken Seite des neuronalen Netzwerks eingegeben und die Ergebnisse y (hier die Ergebnisse y_1 bis y_3 als Beispiel) von der rechten Seite des neuronalen Netzwerks ausgegeben. In dem in **Fig. 5B** gezeigten Beispiel werden die Eingaben x_1 bis x_3 mit entsprechenden Gewichten (insgesamt als w_1 ausgedrückt) multipliziert und in drei Neuronen **N11** bis **N13** eingegeben.

[0064] In **Fig. 5B** werden die jeweiligen Ausgaben der Neuronen **N11** bis **N13** insgesamt als z_1 ausgedrückt. Die Ausgaben z_1 können als Merkmalsvektoren betrachtet werden, die durch Extraktion von Merkmalsgrößen der Eingabevektoren erhalten werden. In dem in **Fig. 5B** gezeigten Beispiel werden die jeweiligen Merkmalsvektoren z_1 mit entsprechenden Gewichten (insgesamt als w_2 ausgedrückt) multipliziert und entsprechend in zwei Neuronen **N21** und **N22** eingegeben. Die Merkmalsvektoren z_1 drücken die Merkmale zwischen den Gewichten w_1 und den Gewichten w_2 aus.

[0065] In **Fig. 5B** werden die jeweiligen Ausgaben der Neuronen **N21** und **N22** insgesamt als z_2 ausgedrückt. Die Ausgaben z_2 können als Merkmalsvektoren betrachtet werden, die durch Extraktion von Merkmalsbeiträgen der Merkmalsvektoren z_1 erhalten werden. In dem in **Fig. 5B** gezeigten Beispiel werden die jeweiligen Merkmalsvektoren z_2 mit entsprechenden Gewichten (insgesamt als w_3 ausgedrückt) multipliziert und in drei Neuronen **N31** bis **N33** eingegeben. Die Merkmalsvektoren z_2 drücken die Merkmale zwischen den Gewichten **W2** und den Gewichten **W3** aus. Schließlich geben die Neuronen **N31** bis **N33** die entsprechenden Ergebnisse y_1 bis y_3 aus.

[0066] Es wird angemerkt, dass es möglich ist, das sogenannte Deep Learning einzusetzen, bei dem ein neuronales Netzwerk, das drei oder mehr Schichten bildet, verwendet wird.

[0067] In der maschinellen Lernvorrichtung **100** der Steuerung **1** führt der Lernabschnitt **110** die Berechnung der Zustandsgrößen **S** und der Bestimmungsdaten **D** als Eingaben x in einer Mehrschichtstruktur gemäß dem obigen neuronalen Netzwerk durch, sodass der Lernabschnitt **110** einen Kompensationsbetrag (Ergebnis y) einer Einlernposition eines Roboters ausgeben kann. Außerdem verwendet der Lernabschnitt **110** in der ma-

schinellen Lernvorrichtung **100** der Steuerung **1** ein neuronales Netzwerk als eine Wertfunktion beim verstärkenden Lernen und führt die Berechnung der Zustandsgrößen **S** und der Aktion **a** als Eingaben **x** in einer Mehrschichtstruktur gemäß dem obigen neuronalen Netz durch, sodass der Lernabschnitt **110** einen Wert (Ergebnis **y**) der Aktion in dem Zustand ausgeben kann. Es wird angemerkt, dass der Aktionsmodus des neuronalen Netzwerks einen Lernmodus und einen Wertevorhersagemodus umfasst. Beispielsweise ist es möglich, unter Verwendung eines Lerndatensatzes im Lernmodus ein Gewicht **w** zu lernen und mit dem gelernten Gewicht **w** im Wertevorhersagemodus einen Aktionswert zu bestimmen. Es wird angemerkt, dass Erkennung, Klassifizierung, Abzug oder ähnliches im Wertevorhersagemodus durchgeführt werden kann.

[0068] Die Konfiguration der obigen Steuerung **1** kann als ein durch den Prozessor **101** ausgeführtes maschinelles Lernverfahren (oder Software) beschrieben werden. Das maschinelle Lernverfahren ist ein Verfahren zum Erlernen eines Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition. Das maschinelle Lernverfahren umfasst:

- einen Schritt des Beobachtens, durch die CPU eines Computers, von Einlernpositionskompensationsbetragsdaten **S1**, die einen Kompensationsbetrag einer Einlernposition bei der Steuerung des Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition anzeigen, und von Motorstörungswertdaten **S2**, die einen Störungswert von jedem der Motoren eines Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition anzeigen, als Zustandsgrößen **S**, die den aktuellen Zustand einer Umwelt ausdrücken, in der der Roboter gesteuert wird;
- einen Schritt des Erfassens von Bestimmungsdaten **D**, die ein Eignungsbestimmungsergebnis des Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß der kompensierten Einlernposition anzeigen; und
- einen Schritt des Lernens der Motorstörungswertdaten **S2** und des Kompensationsbetrags der Einlernposition des Roboters in Zuordnung zueinander unter Verwendung der Zustandsgrößen **S** und der Bestimmungsdaten **D**.

[0069] Fig. 6 zeigt eine Steuerung **2** gemäß einer zweiten Ausführungsform.

[0070] Die Steuerung **2** umfasst eine maschinelle Lernvorrichtung **120** und einen Zustandsdatenerfassungsabschnitt **3**. Der Zustandsdatenerfassungsabschnitt **3** erlangt, als Zustandsdaten **S0** von Zustandsgrößen **S**, die von einem Zustandsbeobachtungsabschnitt **106** beobachtet werden, Einlernpositionskompensationsbetragsdaten **S1** und Motorstörungswertdaten **S2**. Die Zustandsdaten **S0**, die von dem Zustandsdatenerfassungsabschnitt **3** erlangt werden, können auch Einlernpositionsdaten **S3** umfassen, die eine Einlernposition von Einlerndaten oder dergleichen umfassen.

[0071] Der Zustandsdatenerfassungsabschnitt **3** kann die Zustandsdaten **S0** von jedem Abschnitt der Steuerung **2**, von verschiedenen Sensoren einer Werkzeugmaschine, von geeigneten von einem Arbeiter eingegebenen Daten oder dergleichen erlangen.

[0072] Die maschinelle Lernvorrichtung **120** der Steuerung **2** umfasst neben Software (wie einen Lernalgorithmus) und Hardware (wie einen Prozessor **101**) zum spontanen Erlernen eines Kompensationsbetrags einer Einlernposition bei der Steuerung des Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition durch maschinelles Lernen, Software (wie einen Berechnungsalgorithmus) und Hardware (wie einen Prozessor **101**) zum Ausgeben des erlernten Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition als einen Befehl für die Steuerung **2**. Die maschinelle Lernvorrichtung **120** der Steuerung **2** kann so konfiguriert werden, dass ein gemeinsamer Prozessor die gesamte Software, wie etwa einen Lernalgorithmus und einen Berechnungsalgorithmus, ausführt.

[0073] Ein Entscheidungsfindungsabschnitt **122** kann beispielsweise als eine der Funktionen des Prozessors **101** oder als in dem ROM **102** gespeicherte Software zum Betreiben des Prozessors **101** konfiguriert werden. Basierend auf einem Lernergebnis des Lernabschnitts **110** erzeugt und gibt der Entscheidungsfindungsabschnitt **122** einen Sollwert **C** aus, der einen Befehl zum Bestimmen eines Kompensationsbetrags einer Einlernposition eines Roboters bezüglich eines Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition umfasst. Wenn der Entscheidungsfindungsabschnitt **122** den Sollwert **C** an die Steuerung ausgibt, ändert sich der Zustand einer Umwelt entsprechend.

[0074] Der Zustandsbeobachtungsabschnitt **106** beobachtet in einem nächsten Lernzyklus Zustandsgrößen **S**, die sich nach der Ausgabe des Sollwerts **C** an eine Umwelt durch den Entscheidungsfindungsabschnitt **122** geändert haben. Der Lernabschnitt **110** aktualisiert beispielsweise eine Wertfunktion **Q** (das heißt, eine Aktionswerttabelle) unter Verwendung der geänderten Zustandsgrößen **S**, um einen Kompensationsbetrag einer Einlernposition bei der Steuerung eines Roboters gemäß der in den Einlerndaten eines Roboters enthaltenen Einlernposition zu lernen. Es wird angemerkt, dass der Zustandsbeobachtungsabschnitt **106** die Einlernpositionskompensationsbetragsdaten **S1** aus einem RAM **103** der maschinellen Lernvorrichtung **120**, wie in der ersten Ausführungsform beschrieben, erlangen kann, anstatt sie aus den durch den Zustandsdaten-Erfassungsabschnitt **3** erfassten Zustandsdaten **S0** zu erfassen.

[0075] Auf Grundlage eines erlernten Kompensationsbetrags einer Einlernposition bei der Steuerung eines Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition gibt der Entscheidungsfindungsabschnitt **122** den Sollwert **C** gemäß den Zustandsgrößen an die Steuerung **2 S** aus.

[0076] Durch wiederholtes Ausführen des Lernzyklus' treibt die maschinelle Lernvorrichtung **120** das Lernen eines Kompensationsbetrags einer Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition voran und verbessert zunehmend die Zuverlässigkeit des Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition, die von der maschinellen Lernvorrichtung **120** selber bestimmt werden.

[0077] Die maschinelle Lernvorrichtung **120** der Steuerung **2** mit der obigen Konfiguration erzeugt die gleiche Wirkung wie die oben beschriebene maschinelle Lernvorrichtung **100**. Insbesondere kann die maschinelle Lernvorrichtung **120** mit der Ausgabe des Entscheidungsfindungsabschnitts **122** den Zustand einer Umwelt verändern. Andererseits kann die maschinelle Lernvorrichtung **100** ein externes Gerät nach einer Funktion fragen, die dem Entscheidungsfindungsabschnitt entspricht, um Lernergebnisse des Lernabschnitts **110** in einer Umwelt wiederzugeben.

[0078] Fig. 7 zeigt ein System **170**, das Roboter **160** gemäß einer Ausführungsform umfasst.

[0079] Das System **170** umfasst eine Mehrzahl von Robotern **160** und **160'**, die mindestens den gleichen Betriebsvorgang ausführen, und ein drahtgebundenes/drahtloses Netzwerk **172**, das die Roboter **160** und **160'** miteinander verbindet. Mindestens einer der Mehrzahl von Robotern **160** ist als Roboter **160** mit der obigen Steuerung **2** konfiguriert. Außerdem kann das System **170** Roboter **160'** aufweisen, die die Steuerung **2** nicht umfassen. Die Roboter **160** und **160'** haben einen Mechanismus, der zum Durchführen eines Betriebsvorgangs für den gleichen Zweck benötigt wird, ohne sich auf eine Berechnung oder Schätzung zu verlassen.

[0080] In dem System **170** mit der obigen Konfiguration können die die Steuerung **2** umfassenden Roboter **160** von der Mehrzahl von Robotern **160** und **160'** automatisch und genau einen Kompensationsbetrag einer Einlernposition eines Roboters bezüglich eines Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition unter Verwendung der Lernergebnisse des Lernabschnitts **110** berechnen, ohne sich auf eine Berechnung oder Schätzung zu verlassen. Außerdem kann die Steuerung **2** von mindestens einem der Roboter **160** basierend auf den für jeden der anderen Roboter **160** und **160'** erhaltenen Zustandsgrößen **S** und Bestimmungsdaten **D** einen Kompensationsbetrag einer Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition, der allen Robotern **160** und **160'** gemeinsam ist, lernen, sodass die Lernergebnisse zwischen allen Robotern **160** und **160'** geteilt werden. Dementsprechend ermöglicht das System **170**, die Geschwindigkeit und die Zuverlässigkeit des Erlernens eines Kompensationsbetrags einer Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition mit einem breiteren Spektrum von Datensätzen (die Zustandsgrößen **S** und Bestimmungsdaten **D** enthalten) als Eingaben zu verbessern.

[0081] Fig. 8 zeigt ein System **170'** mit einer Mehrzahl von Robotern **160'** gemäß einer weiteren Ausführungsform.

[0082] Das System **170'** umfasst die maschinelle Lernvorrichtung **120** (oder **100**), die Mehrzahl von Roboter **160'** mit der gleichen Maschinenkonfiguration und ein drahtgebundenes/drahtloses Netzwerk **172**, das die Roboter **160'** und die maschinelle Lernvorrichtung **120** (oder **100**) miteinander verbindet.

[0083] In dem System **170'** mit der obigen Konfiguration kann die maschinelle Lernvorrichtung **120** (oder **100**) einen Kompensationsbetrag einer Einlernposition eines Roboters bezüglich eines Störungswerts von jedem der Motoren des Roboters bei der Steuerung des Roboters gemäß der in den Einlerndaten enthaltenen Einlernposition, die allen Robotern **160'** gemeinsam ist, basierend auf Zustandsgrößen **S** und Bestimmungsdaten **D**, die für jeden der mehreren Roboter **160'** erhalten wurden, lernen und automatisch und genau den Kompensationsbetrag der Einlernposition des Roboters in Bezug auf den Störungswert eines jeden der Motoren des Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition unter Verwendung der Lernergebnisse berechnen.

[0084] In dem System **170'** kann die maschinelle Lernvorrichtung **120** (oder **100**) eine Konfiguration aufweisen, die in einem Cloud-Server oder dergleichen im Netzwerk **172** vorhanden ist. Gemäß der Konfiguration kann eine gewünschte Anzahl der Roboter **160'** mit der maschinellen Lernvorrichtung **120** (oder **100**) verbunden werden, unabhängig von den vorhandenen Standorten und den Zeiten der Mehrzahl von Robotern **160'**.

[0085] Arbeiter, die in die Systeme **170** und **170'** eingreifen, können eine Bestimmung dazu durchführen, ob der Erreichungsgrad (die Zuverlässigkeit des Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition) des Erlernens eines Kompensationsbetrags einer Einlernposition eines Roboters bei der Steuerung des Roboters gemäß einer in den Einlerndaten des Roboters enthaltenen Einlernposition mit der maschinellen Lernvorrichtung **120** (oder **100**) durch die maschinelle Lernvorrichtung **120** (oder **100**) ein erforderliches Niveau zu einer geeigneter Zeit nach dem Beginn des Lernens erreicht hat.

[0086] Die Ausführungsformen der vorliegenden Erfindung sind oben beschrieben. Die vorliegende Erfindung ist jedoch nicht auf die Beispiele der oben genannten Ausführungsformen beschränkt und kann in verschiedenen Modi durch Hinzufügen von geeigneten Modifikationen durchgeführt werden.

[0087] Beispielsweise sind ein durch die maschinellen Lernvorrichtungen **100** und **120** ausgeführter Lernalgorithmus, ein durch die maschinelle Lernvorrichtung **120** ausgeführter Berechnungsalgorithmus und ein durch die Steuerungen **1** und **2** ausgeführter Steuerungsalgorithmus nicht auf die oben genannten Algorithmen beschränkt, sondern es können verschiedene Algorithmen eingesetzt werden.

[0088] Außerdem beschreiben die obigen Ausführungsformen eine Konfiguration, bei der die Steuerung **1** (oder **2**) und die maschinelle Lernvorrichtung **100** (oder **120**) eine unterschiedliche CPU aufweisen. Die maschinelle Lernvorrichtung **100** (oder **120**) kann jedoch durch die CPU **11** der Steuerung **1** (oder **2**) und einem in dem ROM **12** gespeicherten Systemprogramm realisiert werden.

[0089] Außerdem beschreiben die vorstehend beschriebenen Ausführungsformen eine Konfiguration, in der die Steuerung **2** einen Kompensationsbetrag einer Einlernposition bei der Steuerung des Roboters gemäß der in den Einlerndaten des Roboters enthaltenen Einlernposition auf Grundlage des Sollwerts **C** steuert, der von der maschinellen Lernvorrichtung **120** ausgegeben wird. Der Sollwert **C**, der von der maschinellen Lernvorrichtung **120** ausgegeben wird, kann jedoch auf der Anzeigevorrichtung eines Programmierhandgeräts als ein Bezugswert angezeigt werden, wenn ein Arbeiter das Lernen ausführt.

Patentansprüche

1. Steuerung (1; 2), die einen Kompensationsbetrag einer Einlernposition bei der Steuerung eines Roboters (160; 160') gemäß der in den Einlerndaten enthaltenen Einlernposition bestimmt, wobei die Steuerung (1; 2) umfasst:

eine maschinelle Lernvorrichtung (100; 120), die einen Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition lernt, wobei die maschinelle Lernvorrichtung (100; 120) umfasst:

einen Zustandsbeobachtungsabschnitt (106), der als Zustandsgrößen, die einen aktuellen Zustand einer Umwelt ausdrücken, Einlernpositionskompensationsbetragsdaten, die den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition anzeigen, Motorstörungswertdaten, die einen Störungswert von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (106; 160) anzeigen, und Einlernpositionsdaten beobachtet, die als die Zustandsgrößen die Einlernposition der Einlerndaten umfassen,

einen Bestimmungsdaten-Erfassungsabschnitt (108), der Bestimmungsdaten erlangt, die ein Eignungsbestimmungsergebnis des Störungswertes von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (160; 160') anzeigen, und

einen Lernabschnitt (110), der den Kompensationsbetrag der Einlernposition des Roboters (160; 160') in Zuordnung zu den Motorstörungswertdaten und der Einlernposition unter Verwendung der Zustandsgrößen und der Bestimmungsdaten lernt.

2. Steuerung (1; 2) nach Anspruch 1, wobei die Bestimmungsdaten, außer einem Eignungsbestimmungsergebnis des Störungswertes von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (160; 160'), mindestens eines von einem Eignungsbestimmungsergebnis einer Einlernposition, in die sich der Roboter (160; 160') schließlich bewegt, einem Eignungsbestimmungsergebnis eines von einem Sensor erfassten Wertes und einem Eignungsbestimmungsergebnis von Zykluszeit bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition nach der Kompensation umfassen.

3. Steuerung (1; 2) nach Anspruch 1 oder 2, wobei der Lernabschnitt (110) einen Belohnungsberechnungsabschnitt (112), der eine Belohnung in Zuordnung zu dem Eignungsbestimmungsergebnis berechnet, und einen Wertfunktion-Aktualisierungsabschnitt (114), der unter Verwendung der Belohnung eine Funktion aktualisiert, die einen Wert des Kompensationsbetrags der Einlernposition relativ zu dem Störungswert von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (160; 160') ausdrückt, hat.

4. Steuerung (1; 2) nach einem der Ansprüche 1 bis 3, wobei der Lernabschnitt (110) eine Berechnung der Zustandsgrößen und der Bestimmungsdaten auf Grundlage einer mehrschichtigen Struktur durchführt.

5. Steuerung (1; 2) nach einem der Ansprüche 1 bis 4, die ferner umfasst: einen Entscheidungsfindungsabschnitt (122), der einen Sollwert auf Grundlage des Kompensationsbetrags der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition auf Grundlage eines Lernergebnisses des Lernabschnitts (110) ausgibt.

6. Steuerung nach einem der Ansprüche 1 bis 5, wobei der Lernabschnitt (110) den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition in jedem von einer Mehrzahl von Robotern (160, 160') unter Verwendung der für jeden der Mehrzahl von Robotern (160, 160') erlangten Zustandsgrößen und Bestimmungsdaten lernt.

7. Steuerung nach einem der Ansprüche 1 bis 6, wobei die maschinelle Lernvorrichtung (100; 120) in einem Cloud-Server vorhanden ist.

8. Maschinelle Lernvorrichtung (100; 120), die einen Kompensationsbetrag einer Einlernposition bei der Steuerung eines Roboters (160; 160') gemäß der Einlernposition, die in den Einlerndaten enthalten ist, lernt, wobei die maschinelle Lernvorrichtung (100; 120) umfasst: einen Zustandsbeobachtungsabschnitt (106), der als Zustandsgrößen, die einen aktuellen Zustand einer Umwelt ausdrücken, Einlernpositionskompensationsbetragsdaten, die den Kompensationsbetrag der Einlernposition bei der Steuerung des Roboters (160; 160') gemäß der Einlernposition anzeigen, Motorstörungswertdaten, die einen Störungswert von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (160; 160') anzeigen, und Einlernpositionsdaten beobachtet, die als die Zustandsgrößen die Einlernposition der Einlerndaten umfassen; einen Bestimmungsdaten-Erfassungsabschnitt (108), der Bestimmungsdaten erlangt, die ein Eignungsbestimmungsergebnis des Störungswertes von jedem der Motoren (50) des Roboters (160; 160') bei der Steuerung des Roboters (160; 160') anzeigen; und einen Lernabschnitt (110), der den Kompensationsbetrag der Einlernposition des Roboters (160; 160') in Zuordnung zu den Motorstörungswertdaten und der Einlernposition durch Verwendung der Zustandsgrößen und der Bestimmungsdaten lernt.

Es folgen 9 Seiten Zeichnungen

Anhängende Zeichnungen

FIG.1

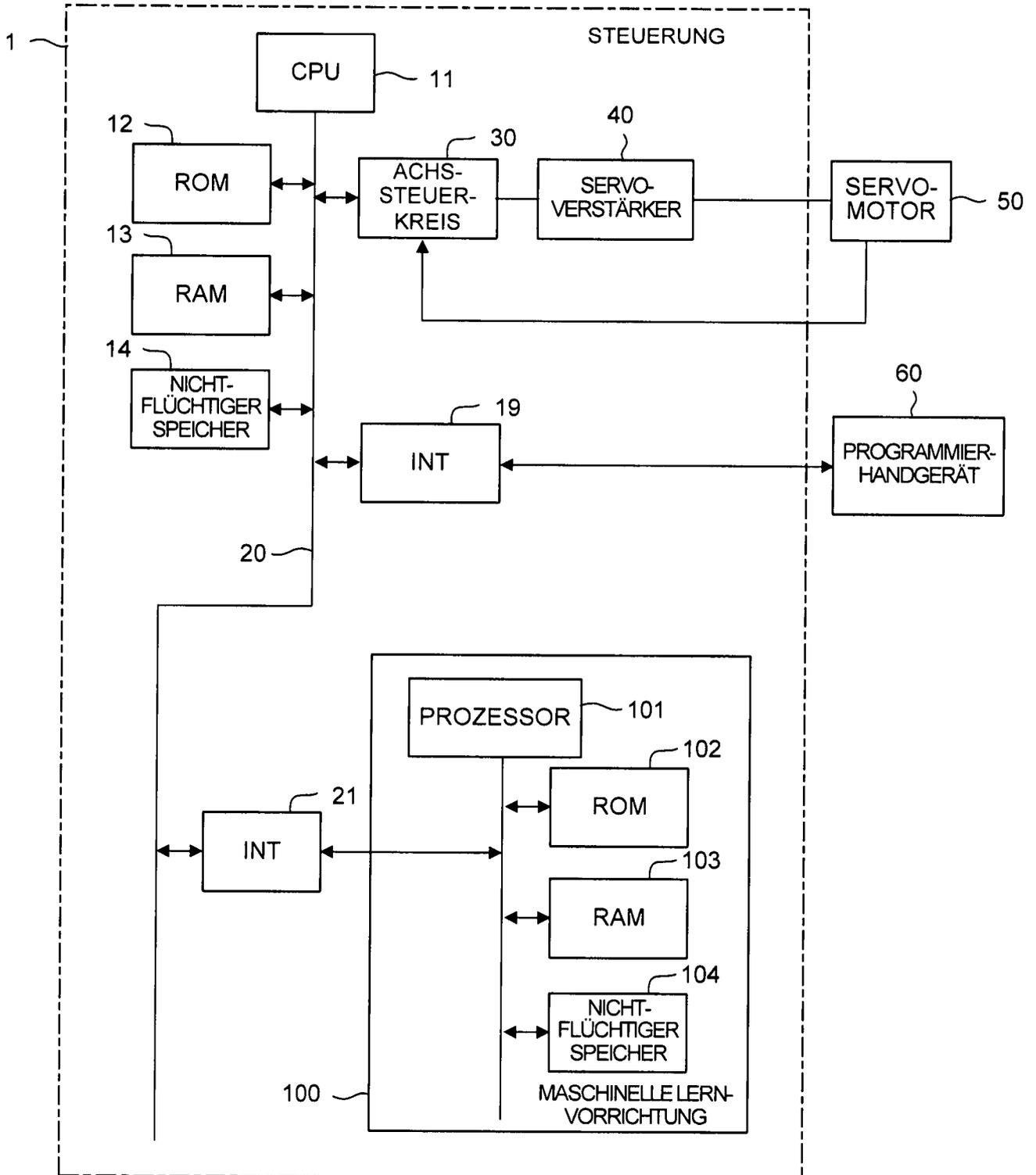


FIG.2

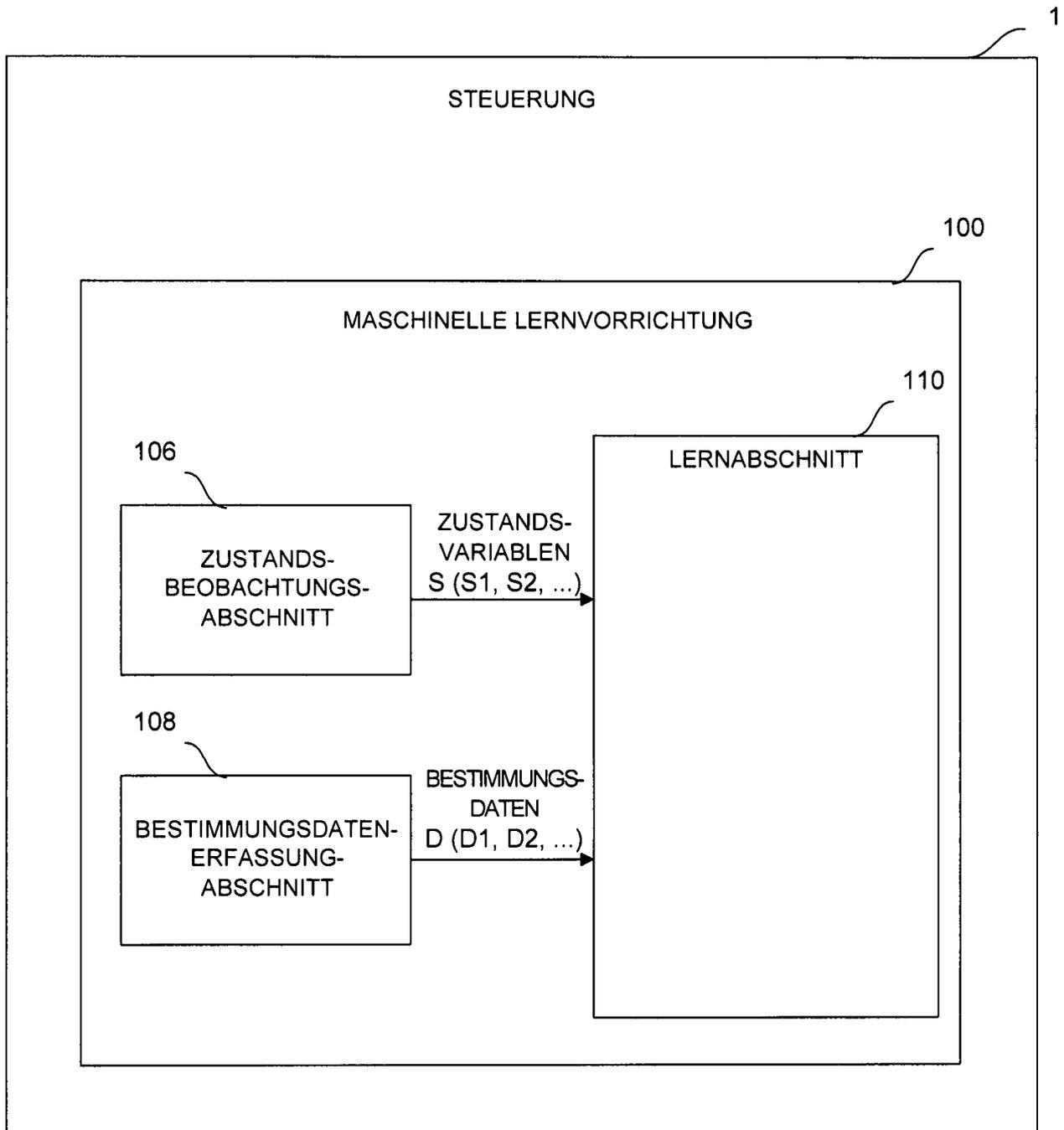


FIG.3

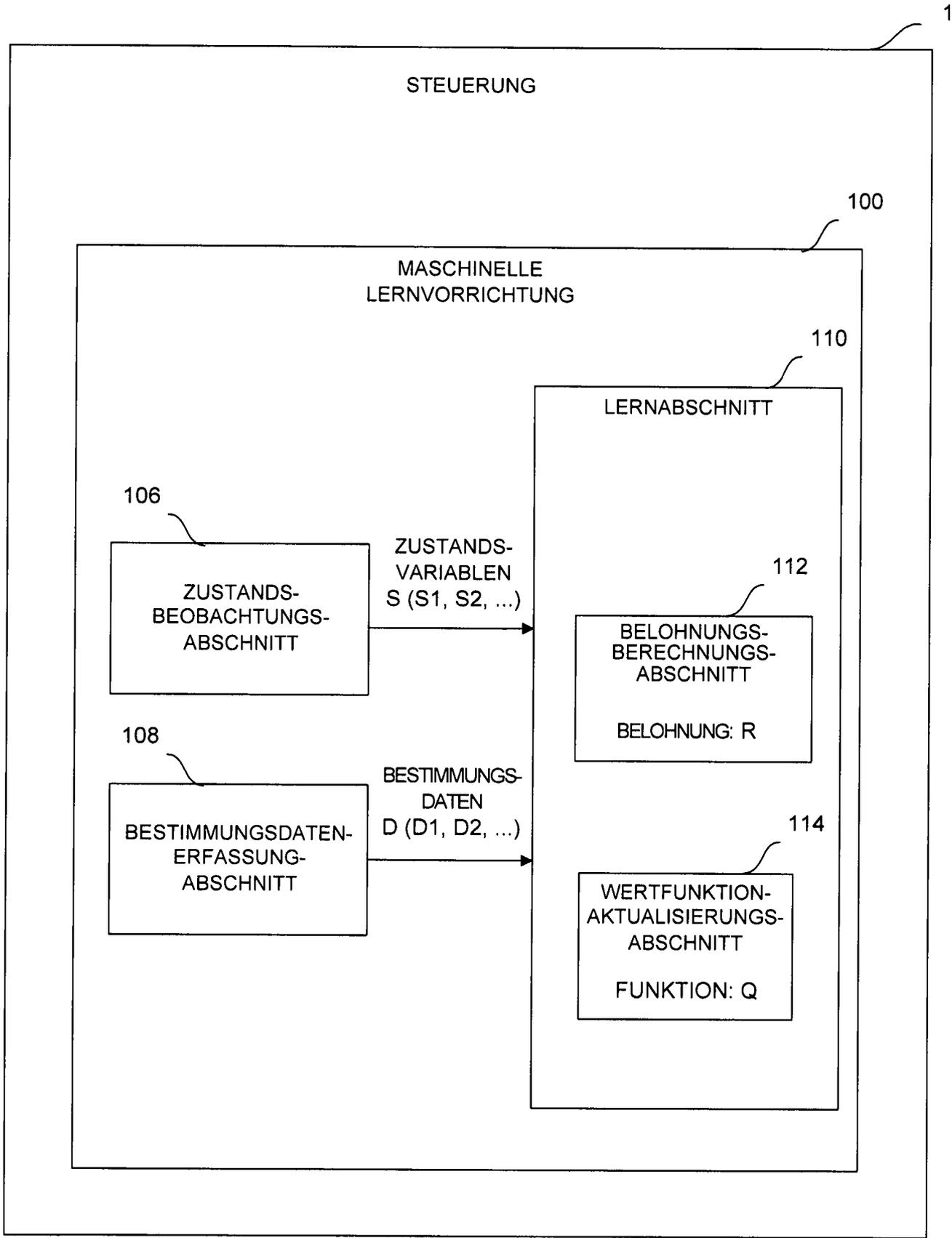


FIG.4

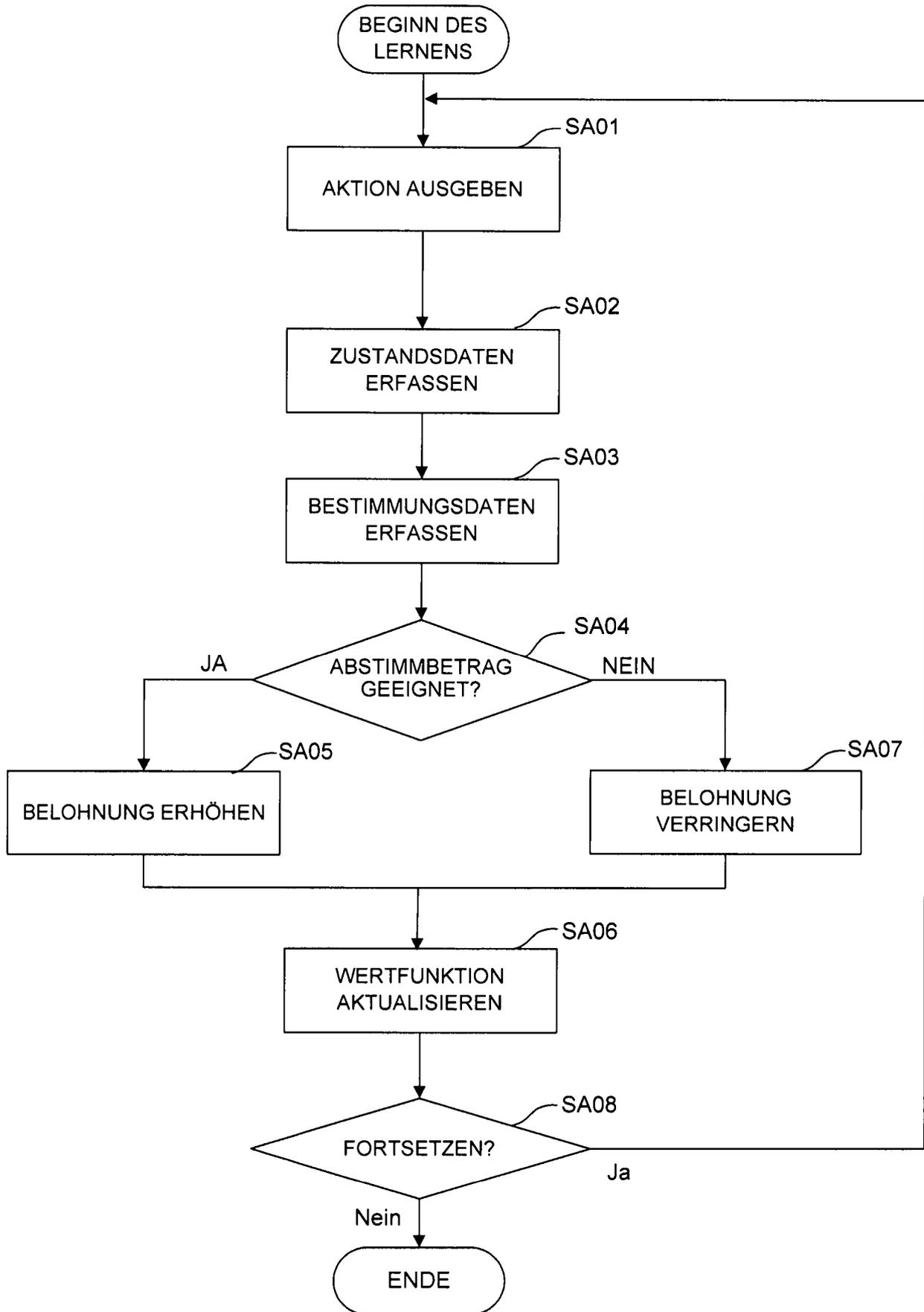


FIG.5A

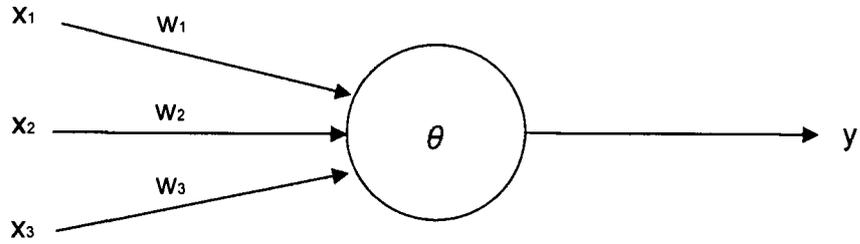


FIG.5B

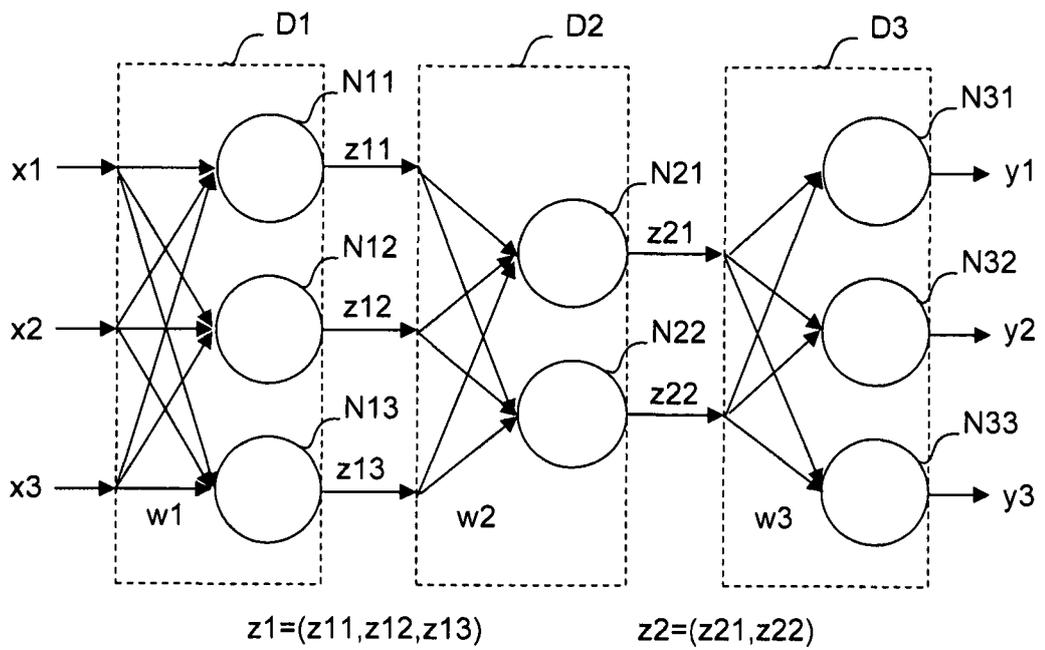


FIG.6

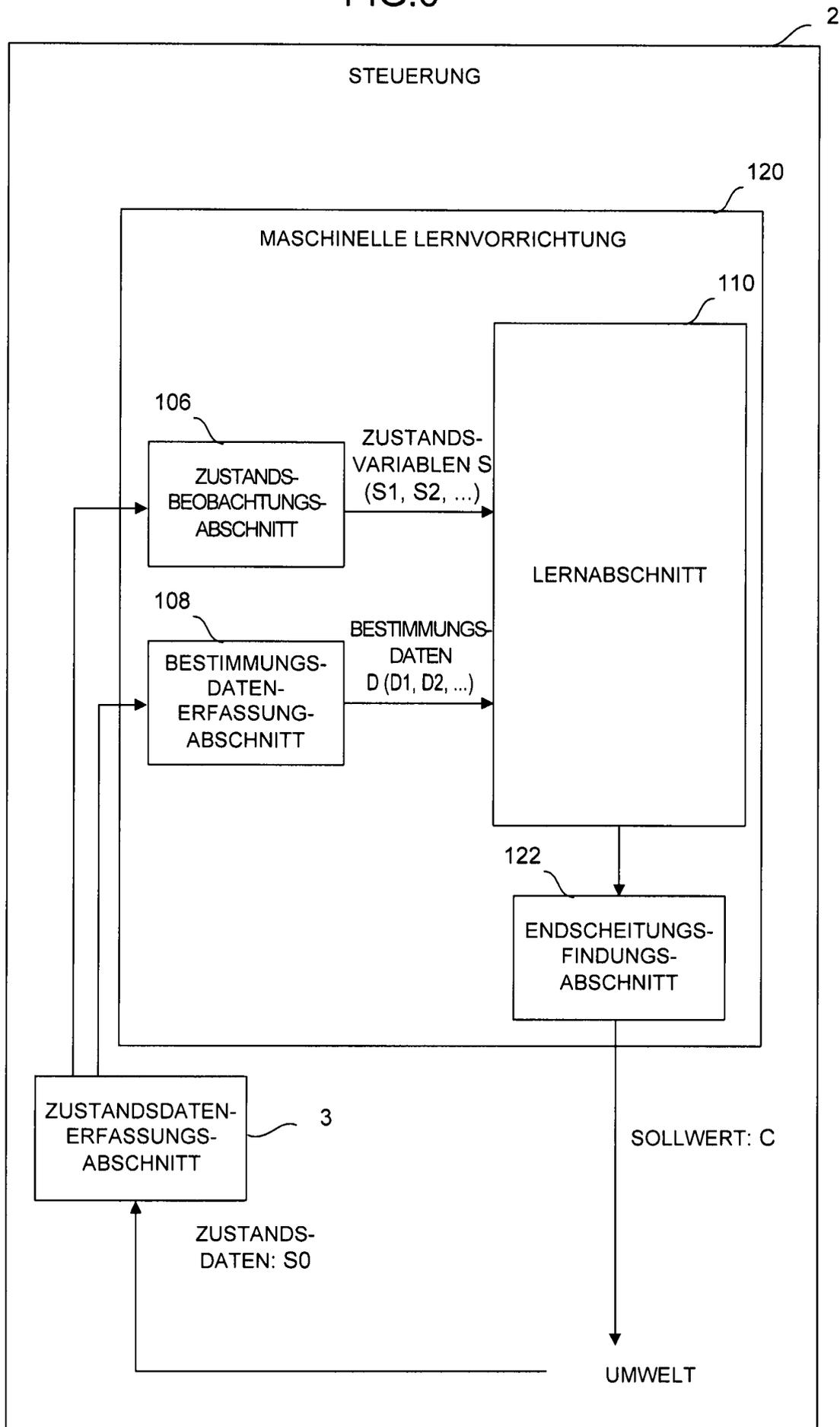


FIG.7

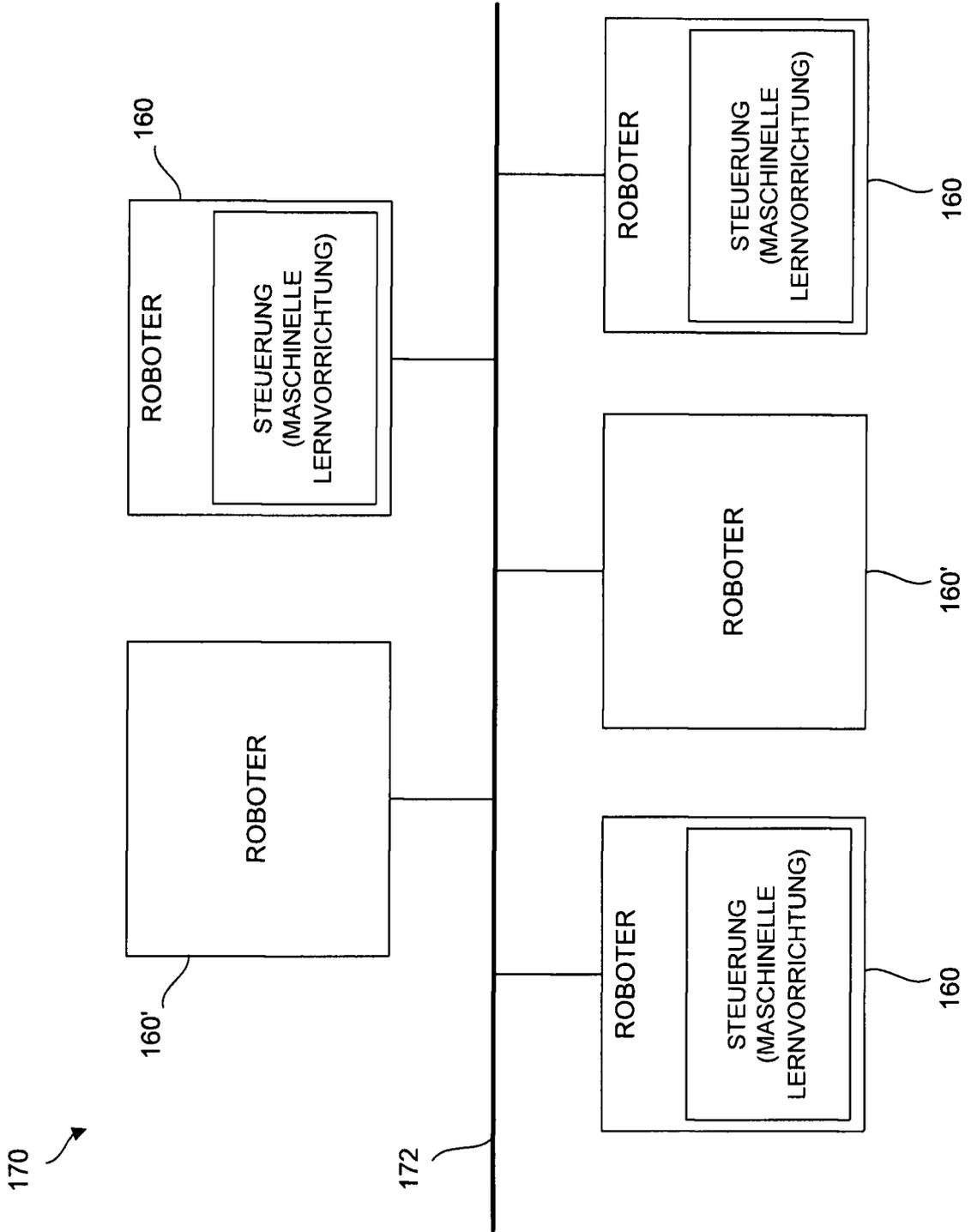


FIG.8

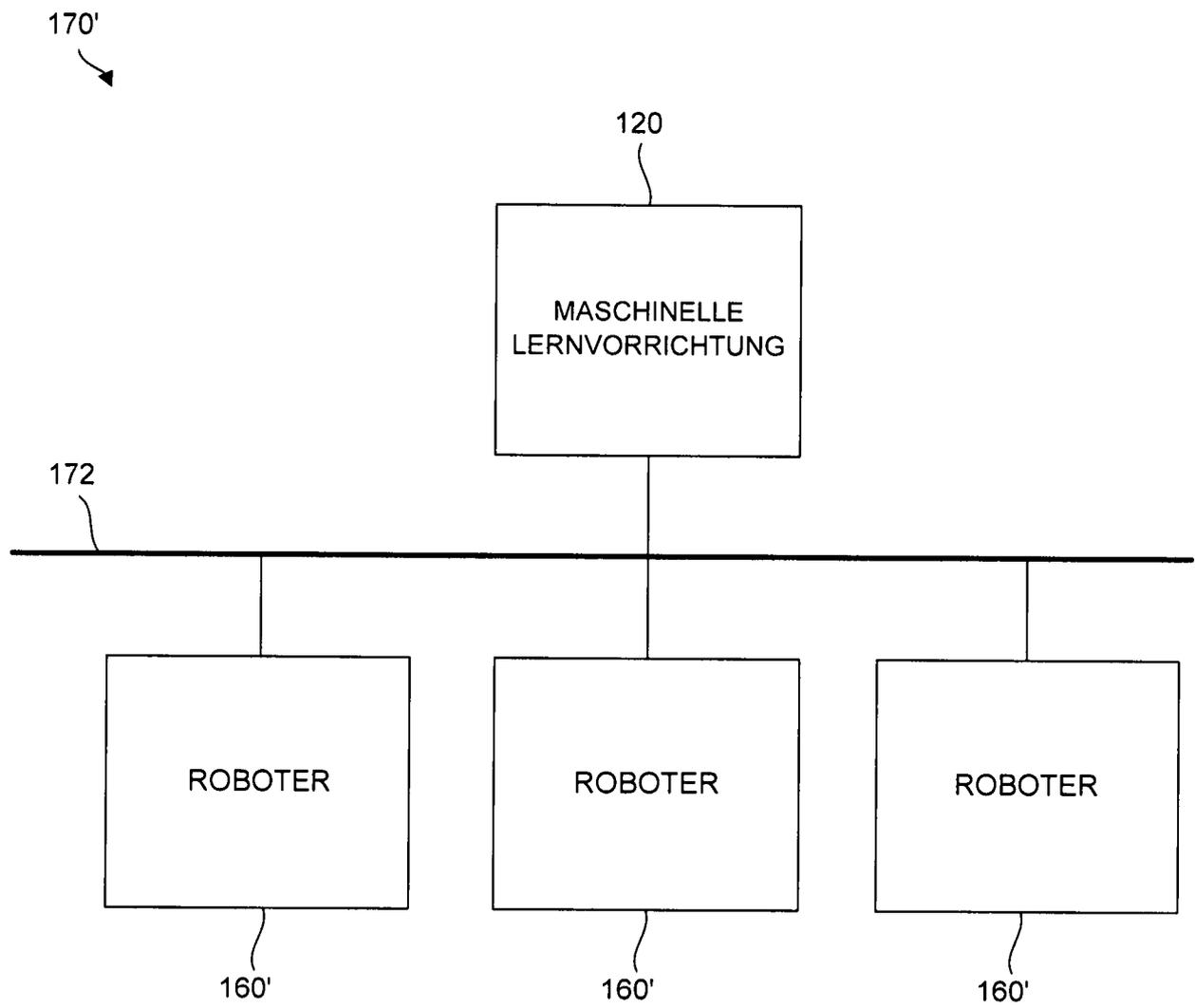


FIG.9

