



(12) **Offenlegungsschrift**

(21) Aktenzeichen: **10 2017 008 836.5**

(51) Int Cl.: **B25J 9/16 (2006.01)**

(22) Anmeldetag: **20.09.2017**

(43) Offenlegungstag: **29.03.2018**

(30) Unionspriorität:
2016-188857 **27.09.2016** **JP**

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

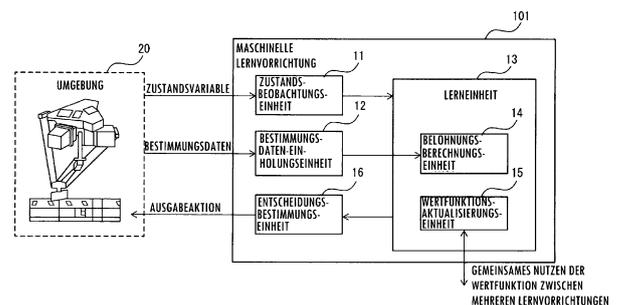
(71) Anmelder:
**FANUC CORPORATION, Oshino-mura,
Yamanashi, JP**

(72) Erfinder:
Ooba, Masafumi, Oshino-mura, Yamanashi, JP

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Maschinelle Lernvorrichtung und maschinelles Lernverfahren zum Lernen eines optimalen Objekt-Greifwegs**

(57) Zusammenfassung: Eine maschinelle Lernvorrichtung gemäß der vorliegenden Erfindung lernt einen Betriebszustand eines Roboters, der eine Mehrzahl von auf einer Trägervorrichtung angeordneten Objekten unter Verwendung einer Hand zum Greifen der Objekte in einem Behälter verstaut. Die maschinelle Lernvorrichtung umfasst eine Zustandsbeobachtungseinheit zum Beobachten der Position und Stellungen der Objekte und einer Zustandsvariablen, die mindestens eines von einer Zykluszeit zum Verstauen der Objekte in dem Behälter und einem Drehmoment und einer Vibration, die auftreten, wenn der Roboter die Objekte während des Betriebs des Roboters greift, umfasst; eine Bestimmungsdaten-Einholungseinheit zum Einholen von Bestimmungsdaten zum Bestimmen einer Spanne von jeweils der Zykluszeit, dem Drehmoment und der Vibration in Abhängigkeit von einem zulässigen Toleranzwert; und eine Lerneinheit zum Lernen des Betriebszustands des Roboters in Übereinstimmung mit einem Trainingsdatensatz, der aus einer Kombination aus der Zustandsvariablen und der Bestimmungsdaten gebildet wird.



Beschreibung

HINTERGRUND DER ERFINDUNG

1. Gebiet der Erfindung

[0001] Die vorliegende Erfindung betrifft eine maschinelle Lernvorrichtung und ein maschinelles Lernverfahren und betrifft insbesondere eine maschinelle Lernvorrichtung und ein maschinelles Lernverfahren zum Lernen eines optimalen Objekt-Greifwegs, wenn ein Roboter auf einer Trägervorrichtung angeordnete Objekte greift.

2. Beschreibung des Standes der Technik

[0002] Objektträgersysteme, die eine Mehrzahl von auf einer Trägervorrichtung, d. h., einem Förderer, beförderte Objekte sequenziell eines nach dem anderen mit einer Hand (mehrfingriger Hand), die mehrere Objekte halten kann, greifen und die die Objekte in Behälter legen, die gleichermaßen durch einen anderen Förderer befördert werden, sind bekannt. Herkömmlicherweise werden die Objekte beim Greifen und Herausnehmen der Objekte grundsätzlich in der Reihenfolge gegriffen, in der sich die Objekte stromabwärts von dem Förderer befinden.

[0003] Beispielsweise wird ein Verfahren vorgeschlagen, in dem ein Förderer entlang einer Förderrichtung von Objekten zweigeteilt ist, und die Objekte in der Reihenfolge, in der sich die Objekte stromabwärts von dem Förderer befinden, gegriffen und in Behälter gelegt werden (zum Beispiel die ungeprüfte Japanische Patentschrift (Kokai) Nr. 2014-104524, nachfolgend als „Patentdokument 1“ bezeichnet).

[0004] **Fig. 1** zeigt ein herkömmliches Verfahren zum Greifen einer Mehrzahl von auf einem Förderer **10** angeordneten Objekten (a1 bis a3 und b1 bis b3) mit einer an einem Roboter **20** vorgesehenen (nicht gezeigten) Hand und Ablegen der Objekte in (nicht gezeigte) Behälter. Die Objekte sind auf einer Ebene des Förderers **10** in willkürlichen Orientierungen (Stellungen) angeordnet, wie durch die Pfeile der Objekte selbst (a1 bis a3 und b1 bis b3) angezeigt.

[0005] Die Hand hat die Funktionen des Greifens von drei Objekten und des Ablegens der Objekte in einen Behälter. Wenn sich der Förderer **10** von der linken nach der rechten Seite von **Fig. 1** bewegt, befinden sich Objekte stromabwärts von dem Förderer **10** in der Reihenfolge a1, a2 und a3. Daher greift die Hand die Objekte a1, a2 und a3 in dieser Reihenfolge und legt die Objekte in einen Behälter und greift danach die Objekte b1, b2 und b3 in dieser Reihenfolge und legt die Objekte in einen anderen Behälter.

[0006] Da herkömmlicherweise die Objekte in der Reihenfolge gegriffen werden, in der sich die Objekte stromabwärts befinden, kann sich die Hand vor und zurück bewegen, wie durch die Pfeile in **Fig. 1** angezeigt.

[0007] Da weiterhin die Stellungen der Objekte nicht berücksichtigt werden, kann sich die Hand des Roboters **20** beträchtlich drehen, wie in **Fig. 1** gezeigt. Da beispielsweise die Objekte a1 und 3 ähnliche Stellungen aufweisen, dreht sich die Hand minimal beim Greifen der Objekte in der Reihenfolge a1, a2 und a3. Da jedoch die Hand die Objekte in der Reihenfolge greift, in der sich die Objekte stromabwärts befinden, greift die Hand die Objekte in der Reihenfolge a1, a2 und a3, wie in **Fig. 1** gezeigt. Infolgedessen dreht sich die Hand des Roboters **20** um 180 Grad, um das Objekt a2 nach dem Greifen des Objekts a1 zu greifen, und dreht sich danach um 180 Grad, um das Objekt a3 zu greifen.

[0008] Das Verfahren von Patentdokument 1, wie oben beschrieben, greift Objekte in der Reihenfolge, in der sich die Objekte stromabwärts befinden, ohne die Positionen der Objekte in der Breitenrichtung des Förderers **10** und die Orientierungen der Objekte zu berücksichtigen. Daher variiert die Bewegungszeit der Hand stark, und in einigen Fällen kann es dem Roboter misslingen, die Objekte in einen Behälter zu legen, während der Behälter vor dem Roboter vorbeiläuft.

[0009] Es ist denkbar, den Förderer für die Behälter jedes Mal zu stoppen, wenn der Roboter Objekte greift. Dies ist jedoch in aktuellen Produktionsstätten schwierig zu übernehmen, weil es Fälle gibt, in denen das Produktionsvolumen innerhalb einer bestimmten Zeitdauer vorbestimmt wird und das Fördern der Behälter aufgrund einer Beziehung zu nachfolgenden Schritten nicht gestoppt werden kann.

[0010] Patentdokument 1 beschreibt ein Verfahren, in dem ein Bereich weiter in kleinere Teilbereiche aufgeteilt wird und das Greifen in jedem Teilbereich ausgeführt wird, um die Tragestrecke zu reduzieren. Wenn jedoch der

Förderer breit ist, dürfte das Aufteilen des Bereichs wenig Auswirkung haben. Da weiterhin die Orientierungen (Stellungen) der Objekte nicht berücksichtigt werden, kann der Roboter ein Objekt greifen, das eine von einer zum Greifen günstigen Stellung völlig unterschiedliche Stellung aufweist.

[0011] Weiterhin berücksichtigt das Verfahren von Patentdokument 1 die dem Roboter eigene Fähigkeit (zum Beispiel die Stärke eines Mechanismus und dergleichen) und den Unterschied der Tragfähigkeit aufgrund von Unterschieden der Stellen der Objekte auf dem Förderer nicht.

KURZDARSTELLUNG DER ERFINDUNG

[0012] Die vorliegende Erfindung zielt darauf ab, eine maschinelle Lernvorrichtung und ein maschinelles Lernverfahren zum Reduzieren der Last auf einen Roboter, der eine Hand aufweist, die die Funktion des Greifens einer Mehrzahl von Objekten hat, bereitzustellen, sowie die Zykluszeit zu minimieren, die die Zeit ist, die der Roboter benötigt, um die Objekte in einem Behälter zu verstauen.

[0013] Eine maschinelle Lernvorrichtung gemäß einer Ausführungsform der vorliegenden Erfindung lernt einen Betriebszustand eines Roboters, der eine Mehrzahl von auf einer Trägervorrichtung angeordneten Objekten unter Verwendung einer Hand zum Greifen der Objekte in einem Behälter verstaut. Die maschinelle Lernvorrichtung umfasst eine Zustandsbeobachtungseinheit zum Beobachten, während des Betriebs des Roboters, der Positionen und Stellungen der Objekte und einer Zustandsvariablen, die mindestens eines von einer Zykluszeit zum Verstauen der Objekte in dem Behälter und eines Drehmoments und einer Vibration, die auftreten, wenn der Roboter die Objekte greift, umfasst; eine Bestimmungsdaten-Einholungseinheit zum Einholen von Bestimmungsdaten zum Bestimmen einer Spanne von jeweils der Zykluszeit, dem Drehmoment und der Vibration in Abhängigkeit von einem zulässigen Toleranzwert; und eine Lerneinheit zum Lernen des Betriebszustands des Roboters in Übereinstimmung mit einem Trainingsdatensatz, der aus einer Kombination aus der Zustandsvariablen und den Bestimmungsdaten gebildet wird.

[0014] Eine maschinelle Lernvorrichtung gemäß einer Ausführungsform der vorliegenden Erfindung lernt einen Betriebszustand eines Roboters, der eine Mehrzahl von auf einer Trägervorrichtung angeordneten Objekten unter Verwendung einer Hand zum Greifen der Objekte in einem Behälter verstaut. Das maschinelle Lernverfahren umfasst die Schritte des Beobachtens, während des Betriebs des Roboters, der Positionen und Stellungen der Objekte und einer Zustandsvariablen, die mindestens eines von einer Zykluszeit zum Verstauen der Objekte in dem Behälter und eines Drehmoments und einer Vibration, die auftreten, wenn der Roboter die Objekte greift, umfasst; Einholen von Bestimmungsdaten zum Bestimmen einer Spanne von jeweils der Zykluszeit, dem Drehmoment und der Vibration in Abhängigkeit von einem zulässigen Toleranzwert; und Lernen des Betriebszustands des Roboters in Übereinstimmung mit einem Trainingsdatensatz, der aus einer Kombination aus der Zustandsvariablen und den Bestimmungsdaten gebildet wird.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0015] Die Aufgaben, Merkmale und Vorteile der vorliegenden Erfindung werden aus der folgenden ausführlichen Beschreibung von Ausführungsformen zusammen mit den beigefügten Zeichnungen ersichtlicher werden. In den beigefügten Zeichnungen:

[0016] Fig. 1 ist eine Zeichnung, die ein Verfahren zum Greifen einer Mehrzahl von auf einem Förderer angeordneten Objekten mit einer an einem Roboter vorgesehenen Hand, um die Objekte in Behälter zu legen, gemäß der herkömmlichen Technik zeigt;

[0017] Fig. 2 ist ein Blockschaltbild einer maschinellen Lernvorrichtung gemäß einer ersten Ausführungsform der vorliegenden Erfindung;

[0018] Fig. 3 ist eine Zeichnung, die ein Verfahren zum Greifen einer Mehrzahl von auf einem Förderer angeordneten Objekten mit einer an einem Roboter vorgesehenen Hand, um die Objekte durch die maschinelle Lernvorrichtung der ersten Ausführungsform der vorliegenden Erfindung in Behälter zu legen, zeigt;

[0019] Fig. 4 ist ein Ablaufdiagramm eines Prozesses des die Objekte greifenden Roboters unter Verwendung der maschinellen Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung;

[0020] Fig. 5 ist eine Zeichnung, die ein Verfahren zum Greifen der auf dem Förderer angeordneten Objekte mit der an dem Roboter vorgesehenen Hand, um die Objekte in die Behälter zu legen, nachdem die maschinelle

Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung ein Lernen ausgeführt hat, zeigt;

[0021] Fig. 6 ist ein Ablaufdiagramm, das ein Verfahren zum Berechnen einer Belohnung in der maschinellen Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung zeigt;

[0022] Fig. 7 ist ein schematisches Schaubild eines Neuronenmodells;

[0023] Fig. 8 ist ein schematisches Schaubild eines dreischichtigen neuronalen Netzwerkmodells;

[0024] Fig. 9 ist ein Blockschaltbild einer maschinellen Lernvorrichtung gemäß einer zweiten Ausführungsform der vorliegenden Erfindung;

[0025] Fig. 10 ist ein Blockschaltbild einer maschinellen Lernvorrichtung gemäß einer dritten Ausführungsform der vorliegenden Erfindung; und

[0026] Fig. 11 ist ein Blockschaltbild einer maschinellen Lernvorrichtung gemäß einer vierten Ausführungsform der vorliegenden Erfindung.

AUSFÜHRLICHE BESCHREIBUNG DER ERFINDUNG

[0027] Eine maschinelle Lernvorrichtung und ein maschinelles Lernverfahren gemäß der vorliegenden Erfindung werden nachfolgend mit Bezug auf die Zeichnungen beschrieben.

[Erste Ausführungsform]

[0028] Eine maschinelle Lernvorrichtung gemäß einer ersten Ausführungsform wird mit Bezug auf die Zeichnungen beschrieben. Fig. 2 ist ein Blockschaltbild einer maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform der vorliegenden Erfindung. Fig. 3 ist eine Zeichnung, die ein Verfahren zum Greifen einer Mehrzahl von auf einer Trägervorrichtung (Förderer) angeordneten Objekten mit einer an einem Roboter vorgesehenen Hand durch die maschinelle Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung zeigt. Die maschinelle Lernvorrichtung **101** gemäß der ersten Ausführungsform der vorliegenden Erfindung lernt einen Betriebszustand eines Roboters **20**, der eine Mehrzahl von auf einer Trägervorrichtung **10** angeordneten Objekten p1 bis p6 unter Verwendung einer (nicht gezeigten) Hand zum Greifen der Objekte in (nicht gezeigten) Behältern verstaut.

[0029] Die maschinelle Lernvorrichtung **101** umfasst eine Zustandsbeobachtungseinheit **11**, eine Bestimmungsdaten-Einholungseinheit **12** und eine Lerneinheit **13**.

[0030] Die Zustandsbeobachtungseinheit **11** beobachtet, während des Betriebs des Roboters **20**, die Positionen und Stellungen der Objekte (p1 bis p6), und eine Zustandsvariable, die mindestens eines von einer Zykluszeit zum Verstauen der Objekte in dem Behälter und eines Drehmoments und einer Vibration, die auftreten, wenn der Roboter die Objekte greift, umfasst. Die Positionen und Stellungen der Objekte können basierend auf einem durch eine (nicht gezeigte) Kamera aufgenommenen Bild analysiert werden. In diesem Fall wird die Analyse der Positionen und Stellungen der Objekte bevorzugt zu dem Zeitpunkt abgeschlossen, wenn der Roboter **20** beginnt, die Objekte (p1 bis p6) zu greifen. Die Kamera befindet sich dabei bevorzugt stromaufwärts des Förderers **10** von dem Roboter **20**. Es wird angemerkt, dass der Förderer **10** die Objekte von der linken Seite zur rechten Seite von Fig. 3 und Fig. 5 mit einer konstanten Geschwindigkeit befördert.

[0031] Die Zykluszeit bezieht sich auf die Zeit, ab der der Roboter beginnt, eine vorgeschriebene Anzahl von Objekten in einem Behälter zu verstauen, bis der Roboter das Verstauen der Objekte in dem Behälter abschließt. Die vorgeschriebene Anzahl von Objekten bezieht sich auf Objekte, deren Greifreihenfolge (Weg) zu bestimmen ist, und auf in einem bestimmten Bereich **30** von Fig. 3 enthaltene Objekte. Der Roboter **20** weist bevorzugt einen Zeitgeber zum Messen der Zykluszeit auf.

[0032] Das Drehmoment tritt auf, wenn sich die Hand zu den Positionen der Objekte bewegt und wenn sich die Hand abhängig von den Stellungen der Objekte dreht. Das Drehmoment wird basierend auf durch Motoren zum Antreiben der Hand und eines (nicht gezeigten) Arms des Roboters **20** fließende Ströme berechnet. Der Roboter **20** weist bevorzugt Amperemeter zum Messen der durch die Motoren fließenden Ströme auf. Es wird angemerkt, dass nachdem die Hand ein Objekt gegriffen hat, sich die Hand dreht, während sie sich an eine

unterschiedliche Position bewegt, um ein weiteres Objekt zu greifen. Mit anderen Worten, nachdem die Hand ein Objekt gegriffen hat, bewegt sich die Hand, während sie sich dreht, um einen geeigneten Winkel zum Greifen des nächsten Objekts zu haben.

[0033] Vibration tritt auf, wenn sich die Hand zu Positionen des Objekts bewegt und stoppt, und wenn sich die Hand dreht und das Drehen abhängig von den Stellungen des Objekts stoppt. Um die Vibration zu messen, weist die Hand bevorzugt einen Beschleunigungssensor auf. Die Vibration wird basierend auf einer durch den Beschleunigungssensor detektierten Beschleunigung berechnet.

[0034] Die Bestimmungsdaten-Einholungseinheit **12** holt Bestimmungsdaten zum Bestimmen einer Spanne von jeweils der Zykluszeit, dem Drehmoment und der Vibration in Abhängigkeit von deren zulässigem Toleranzwert ein. Die zulässigen Toleranzwerte der Zykluszeit, des Drehmoments und der Vibration können in einem (nicht gezeigten) Speicher gespeichert werden. Die Zykluszeit, das Drehmoment und die Vibration sind bevorzugt allesamt gleich oder kleiner als die zulässigen Toleranzwerte.

[0035] Die Lerneinheit **13** lernt den Betriebszustand des Roboters in Übereinstimmung mit einem Trainingsdatensatz, der aus einer Kombination aus der Zustandsvariablen und den Bestimmungsdaten gebildet wird. Wenn die Zykluszeit, das Drehmoment und die Vibration allesamt gleich oder kleiner als die zulässigen Toleranzwerte sind, werden die Objekte bevorzugt in der Reihenfolge gegriffen, in der die Zykluszeit minimiert wird.

[0036] Als Nächstes wird der Arbeitsgang des Greifens von Objekten mit einem Roboter durch die maschinelle Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung mit Bezug auf das Ablaufdiagramm von **Fig. 4** beschrieben. Zuerst werden in Schritt S101 die derzeitigen Positionen und Stellungen einer Mehrzahl von Objekten (p1 bis p6) eingeholt. Die Objekte, deren Greifreihenfolge zu bestimmen ist, sind innerhalb eines bestimmten Bereichs **30** eines Förderers **10** angeordnet, wie durch die gestrichelten Linien von **Fig. 3** angedeutet. In dem Beispiel von **Fig. 3** ist die Greifreihenfolge der Objekte (p1 bis p6) zu bestimmen. Der bestimmte Bereich **30** des Förderers **10** ist bevorzugt ein Bereich, der eine optimale Anzahl von Objekten basierend auf einem Gleichgewicht zwischen der erforderlichen Zeit zum Berechnen der Zykluszeit, des Drehmoments und der Vibration in jeder Kombination der Greifreihenfolge der Objekte und der Geschwindigkeit des Förderers, um die Objekte zu bewegen, enthält.

[0037] Als Nächstes wird in Schritt S102 die Greifreihenfolge der Objekte basierend auf Lernergebnissen zugeordnet. Als Nächstes wird in Schritt S103 in Erwiderung einer Anfrage von dem Roboter **20** die Greifreihenfolge der Objekte von der maschinellen Lernvorrichtung **101** zum Roboter **20** gesendet.

[0038] **Fig. 5** ist eine Zeichnung, die ein Verfahren zeigt zum Greifen der auf dem Förderer angeordneten Objekte mit der an dem Roboter vorgesehenen Hand, um die Objekte in den Behältern zu verstauen, nachdem die maschinelle Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung einen Lernen ausgeführt hat. Als Ergebnisse des Lernens, wie durch Pfeile angedeutet, wird das Objekt p1 als Erstes gegriffen, das Objekt p3 wird als Nächstes gegriffen und das Objekt p2 wird zuletzt gegriffen.

[0039] Im Vergleich zu dem Fall der in **Fig. 1** als ein Beispiel gezeigten herkömmlichen Technik ermöglicht das Greifen der Objekte in dieser Reihenfolge eine Reduktion der Bewegungsstrecke der Hand, um die Objekte zu greifen. Somit ist die zum Greifen der drei Objekte (p1 bis p3) erforderliche Zeit reduziert. Daher ist es möglich, die Zykluszeit, d. h. die zum Verstauen aller Objekte in dem Behälter erforderliche Zeit, zu verkürzen.

[0040] Die Hand greift drei Objekte in dieser Ausführungsform, ist jedoch nicht darauf beschränkt; die Anzahl der durch die Hand gegriffenen Objekte kann zwei oder vier oder mehr betragen. „Greifen“ umfasst ein „Ansaugen“ durch die Hand.

[0041] Gemäß der vorliegenden Erfindung dreht sich die Hand, um das Objekt p3 zu greifen, nachdem das Objekt p1 gegriffen wurde, um einen kleineren Winkel als der Winkel, um den sich die Hand dreht, um das Objekt a2 zu greifen, nachdem das Objekt a1 gegriffen wurde, wie in der herkömmlichen Technik beschrieben (siehe **Fig. 1**). Folglich ist es möglich, das Drehmoment und die Vibration, die auftreten, während die maschinelle Lernvorrichtung die drei Objekte (p1 bis p3) greift, zu reduzieren. Daher ist es möglich die Gesamtwerte des Drehmoments und der Vibration, die auftreten, während die maschinelle Lernvorrichtung alle Objekte in dem Behälter verstaut, zu reduzieren.

[0042] Als Nächstes wird die Konfiguration der Lerneinheit **13** beschrieben. Wie in **Fig. 2** gezeigt, umfasst die Lerneinheit **13** eine Belohnungsberechnungseinheit **14** zum Berechnen einer Belohnung basierend auf

den Bestimmungsdaten und eine Wertfunktion-Aktualisierungseinheit **15**, die eine Wertfunktion aktualisiert, die zum Schätzen der Greifreihenfolge von Objekten verwendet wird, um mindestens eines von der Zykluszeit, dem Drehmoment und der Vibration basierend auf der Belohnung zu reduzieren.

[0043] Die Lerneinheit **13** aktualisiert eine Aktionswerttabelle entsprechend der Greifreihenfolge von Objekten basierend auf der Zustandsvariablen, die mindestens eines von der Zykluszeit, dem Drehmoment, der Vibration und der Belohnung umfasst.

[0044] Die Lerneinheit **13** kann eine Aktionswerttabelle, die mindestens einem von der Zykluszeit, dem Drehmoment und der Vibration entspricht, wenn ein anderer Roboter mit derselben Konfiguration wie der Roboter **20** andere Objekte in einem Behälter verstaut, basierend auf der Zustandsvariablen und der Belohnung des anderen Roboters aktualisieren.

[0045] Die Belohnungsberechnungseinheit **14** berechnet eine Belohnung basierend auf mindestens einem von der Zykluszeit, dem Drehmoment und der Vibration. Des Weiteren kann, wenn das Greifen eines Objekts misslingt, oder mit anderen Worten, wenn ein Greiffehler auftritt, die Belohnungsberechnungseinheit **14** eine negative Belohnung vorsehen.

[0046] Die Lerneinheit **13** umfasst ferner bevorzugt eine Entscheidungsbestimmungseinheit **16**, die die Greifreihenfolge von Objekten in Übereinstimmung mit Ergebnissen des Lernens in Übereinstimmung mit einem Trainingsdatensatz bestimmt.

[0047] Als Nächstes wird ein Verfahren zum Berechnen einer Belohnung beschrieben. **Fig. 6** ist ein Ablaufdiagramm, das ein Verfahren zum Berechnen einer Belohnung in der maschinellen Lernvorrichtung gemäß der ersten Ausführungsform der vorliegenden Erfindung zeigt. Zuerst holt in Schritt S201 die Zustandsbeobachtungseinheit **11** Daten zu jeweils der Zykluszeit, dem Drehmoment und dem Vibrationsbetrag ein.

[0048] Als Nächstes bestimmt in Schritt S202 die Belohnungsberechnungseinheit **14**, ob die Zykluszeit kürzer als ein Referenzwert ist oder nicht. Wenn die Zykluszeit kürzer als der Referenzwert ist, wird in Schritt S203 eine positive Belohnung vorgesehen. Wenn andererseits die Zykluszeit gleich groß oder länger als der Referenzwert ist, geht der Arbeitsgang weiter zu Schritt S201, und es wird keine Belohnung vorgesehen. Der Referenzwert der Zykluszeit ist der Durchschnittswert der Zykluszeit, wenn der Roboter für eine bestimmte Zeitdauer in der Vergangenheit betrieben wurde. Des Weiteren kann der Referenzwert in Übereinstimmung mit Lernergebnissen angepasst werden, wobei der Durchschnittswert als ein Anfangswert verwendet wird.

[0049] Als Nächstes bestimmt in Schritt S204 die Belohnungsberechnungseinheit **14**, ob sich das Drehmoment erhöht hat oder nicht. Wenn das Drehmoment gleich oder kleiner als ein Referenzwert ist, geht der Arbeitsgang weiter zu Schritt S205 und es wird keine Belohnung vorgesehen. Wenn andererseits das Drehmoment größer als der Referenzwert ist, wird eine negative Belohnung in Schritt S207 vorgesehen. Der Referenzwert des Drehmoments ist der Durchschnittswert des Drehmoments, wenn der Roboter für eine bestimmte Zeitdauer betrieben wurde. Des Weiteren kann der Referenzwert in Übereinstimmung mit Lernergebnissen angepasst werden, wobei der Referenzwert als ein Anfangswert verwendet wird.

[0050] Als Nächstes bestimmt die Belohnungsberechnungseinheit **14** in Schritt S206, ob sich der Vibrationsbetrag erhöht hat oder nicht. Wenn der Vibrationsbetrag gleich groß oder kleiner als ein Referenzwert ist, geht der Arbeitsgang weiter zu Schritt S205, und es wird keine Belohnung vorgesehen. Wenn andererseits der Vibrationsbetrag größer als der Referenzwert ist, wird eine negative Belohnung in Schritt S207 vorgesehen. Der Referenzwert des Vibrationsbetrags ist der Durchschnittswert des Vibrationsbetrags, wenn der Roboter für eine bestimmte Zeitdauer in der Vergangenheit betrieben wurde. Des Weiteren kann der Referenzwert in Übereinstimmung mit Lernergebnissen angepasst werden, wobei der Durchschnittswert als ein Anfangswert verwendet wird.

[0051] Als Nächstes wird in Schritt S208 eine Belohnung berechnet. Wenn R_c eine Belohnung basierend auf der Zykluszeit darstellt, R_r eine Belohnung basierend auf dem Drehmoment darstellt und R_v eine Belohnung auf der Vibration darstellt wird der Gesamtwert R der Belohnungen durch $R = \alpha R_c + \beta R_r + \gamma R_v$ berechnet, wobei vorgeschriebene Koeffizienten verwendet werden, die Gewichte definieren.

[0052] Als Nächstes aktualisiert die Lerneinheit **13** in Schritt S209 eine Aktionswerttabelle, die der Greifreihenfolge von Objekten entspricht, basierend auf einer Zustandsvariablen von mindestens einem von der Zykluszeit, dem Drehmoment, der Vibration und der Belohnung.

[0053] Die Lerneinheit **13** führt bevorzugt eine Berechnung einer durch die Zustandsbeobachtungseinheit **11** in einer mehrschichtigen Struktur beobachteten Zustandsvariablen durch und aktualisiert eine Aktionswertta-
belle in Echtzeit. Als ein Verfahren zum Ausführen einer Berechnung einer Zustandsvariablen in einer mehr-
schichtigen Struktur kann beispielsweise ein mehrschichtiges neuronales Netzwerk, wie in **Fig. 8** gezeigt, ver-
wendet werden.

[0054] Die in **Fig. 2** gezeigte maschinelle Lernvorrichtung **101** wird ausführlich beschrieben. Die maschinelle
Lernvorrichtung **101** hat die Funktion des Lernens von Wissen, sowie die Funktionen des Extrahierens von
nützlichen Regeln, Wissensrepräsentationen, Kriterien und dergleichen aus einer in die Vorrichtung durch Ana-
lyse eingegebene Datengruppe, und des Ausgebens von Bestimmungsergebnissen. Hierfür gibt es verschie-
dene Verfahren, die grob in „überwachtes Lernen“ „unüberwachtes Lernen“ und „bestärkendes Lernen“ unter-
teilt werden. Darüber hinaus wird beim Verwirklichen dieser Verfahren ein als „Deep-Learning“ bezeichnetes
Verfahren zum Lernen der Extraktion von den Merkmalen selbst eingesetzt.

[0055] Beim „überwachten Lernen“ wird eine große Menge an Datenpaaren einer bestimmten Eingabe und
ein Ergebnis (Label) einem Lernsystem bereitgestellt. Das Lernsystem lernt Merkmale aus dem Datensatz
und erhält heuristisch ein Modell, um ein Ergebnis aus einer Eingabe, d. h. der Beziehung dazwischen, vor-
herzusagen. „Überwachtes Lernen“ kann durch Verwenden eines Algorithmus, wie ein später beschriebenes
neuronales Netzwerk, verwirklicht werden.

[0056] Beim „unüberwachten Lernen“ wird nur eine große Menge von Eingabedaten dem Lernsystem bereit-
gestellt. Das Lernsystem lernt die Verteilung der Eingabedaten und wendet Komprimierung, Klassifizierung,
Ausrichtung oder dergleichen auf die Eingabedaten an, ohne mit entsprechenden Ausgabedaten als Überwa-
cher versorgt zu werden. Die Merkmale des Datensatzes können einem Clustern durch Analogie unterworfen
werden. Mit Verwendung dieses Ergebnisses wird, während ein bestimmtes Kriterium bereitgestellt wird, eine
Ausgabe zugeordnet, um das Kriterium zu optimieren, und dies erlaubt eine Vorhersage der Ausgabe. Es gibt
ebenfalls ein als „teilüberwachtes Lernen“ bezeichnetes Verfahren als eine Zwischenproblemlösung zwischen
dem „überwachten Lernen“ und dem „unüberwachten Lernen“, in dem ein Teil der Daten Paare einer Eingabe
und einer Ausgabe umfassen, während der andere Teil nur Eingabedaten umfasst. In dieser Ausführungsform
verwendet das unüberwachte Lernen Daten, die ohne aktuelles Betreiben des Roboters eingeholt werden kön-
nen, um die Lerneffizienz zu verbessern.

[0057] Probleme beim bestärkenden Lernen werden wie folgt bestimmt.

- Der Roboter beobachtet einen Umgebungszustand und bestimmt eine Aktion.
- Die Umgebung ändert sich in Übereinstimmung mit einigen Regeln, und manchmal ändern die Aktionen
des Roboters selbst die Umgebung.
- Ein Belohnungssignal kehrt immer dann zurück, wenn eine Aktion ausgeführt wird.
- Das Ziel ist, eine (diskontierte) Gesamtbelohnung in der Zukunft zu maximieren.
- Das Lernen wird ab einem Zustand begonnen, in dem ein Ergebnis, das durch eine Aktion bewirkt wird,
nicht oder nur unzureichend bekannt ist. Der Roboter erhält das Ergebnis als Daten erst nachdem der Ro-
boter tatsächlich betrieben wird. Mit anderen Worten, es ist notwendig, eine optimale Aktion durch Auspro-
bieren zu suchen.
- Das Lernen kann ab einem günstigen Startpunkt begonnen werden, indem ein herkömmliches Lernen
ausgeführt wird (mittels dem oben beschriebenen überwachten Lernen oder inversen bestärkenden Ler-
nen), um in einem Ausgangszustand die Aktion eines Menschen zu imitieren.

[0058] Das „bestärkende Lernen“ ist ein Verfahren zum Lernen von optimalen Aktionen basierend auf Inter-
aktionen zwischen einer Aktion und einer Umgebung durch Lernen von Aktionen, sowie Bestimmen und Klas-
sifizieren; mit anderen Worten, es ist ein Lernverfahren, um eine in der Zukunft erhaltene Gesamtbelohnung zu
maximieren. In dieser Ausführungsform zeigt dies an, dass Aktionen, die eine Wirkung auf die Zukunft haben,
eingeholt werden können. Die folgende Beschreibung nimmt Q-Lernen als ein Beispiel, ist jedoch nicht darauf
beschränkt.

[0059] Das Q-Lernen ist ein Verfahren zum Lernen eines $Q(s, a)$ -Werts zum Wählen einer Aktion „a“ in einem
gewissen Umgebungszustand „s“. Mit anderen Worten, in einem gewissen Zustand „s“ wird eine Aktion „a“ mit
dem höchsten $Q(s, a)$ -Wert als eine optimale Aktion gewählt. Jedoch ist ein korrekter $Q(s, a)$ -Wert bezüglich
einer Kombination eines Zustands „s“ und einer Aktion „a“ zu Beginn überhaupt nicht bekannt. Daher wählt ein
Agent verschiedene Aktionen „a“ in einem gewissen Zustand „s“, und erhält eine Belohnung für jede Aktion
„a“. Dadurch lernt der Agent eine bessere Aktion, d. h. einen korrekten $Q(s, a)$ -Wert, zu wählen.

[0060] Das Ziel ist, eine in der Zukunft zu erzielende Gesamtbelohnung zu maximieren, d. h., $Q(s, a) = E[\sum \gamma^t r_t]$ als ein Ergebnis von Aktionen zu erzielen (es wird ein erwarteter Wert genommen, wenn sich ein Zustand in Übereinstimmung mit optimalen Aktionen ändert; die optimalen Aktionen sind selbstverständlich noch nicht bekannt und müssen somit während des Lernens gefunden werden). Beispielsweise wird eine Aktualisierungsgleichung für einen $Q(s, a)$ -Wert wie folgt dargestellt:

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

wobei s_t einen Umgebungszustand zum Zeitpunkt t darstellt und a_t eine Aktion zum Zeitpunkt t darstellt. Durch Ausführen der Aktion a_t ändert sich der Zustand zu s_{t+1} . „ r_{t+1} “ stellt eine durch die Zustandsänderung bereitgestellte Belohnung dar. Ein Begriff „max“ stellt das Produkt aus einem Q -Wert, wenn eine Aktion „ a “ mit dem zu diesem Zeitpunkt bekannten höchsten Wert Q in dem Zustand s_{t+1} gewählt wird, und γ dar. „ γ “ ist ein Parameter im Bereich von $0 < \gamma \leq 1$ und wird Diskontierungsfaktor genannt. „ α “ ist eine Lernrate im Bereich von $0 < \alpha \leq 1$.

[0061] Diese Gleichung bezeichnet ein Verfahren zum Aktualisieren eines $Q(s_t, a_t)$ -Werts einer Aktion a_t in einem Zustand s_t basierend auf einer Belohnung r_{t+1} , die als ein Ergebnis der Aktion a_t zurückgesendet wurde. Diese Aktualisierungsgleichung zeigt an, dass, im Vergleich zu einem $Q(s_t, a_t)$ -Wert einer Aktion „ a “ in einem Zustand „ s “, wenn im nächsten Zustand ein von einer Belohnung r_{t+1} plus der Aktion „ a “ abgeleiteter Wert $Q(s_{t+1}, \max_a a_{t+1})$ einer optimalen Aktion „max a “ höher ist, $Q(s_t, a_t)$ erhöht wird. Wenn nicht, wird $Q(s_t, a_t)$ verringert. Mit anderen Worten, der Wert einer Aktion in einem bestimmten Zustand wird an einen optimalen Aktionswert im nächsten Zustand angenähert, der von einer sofort als ein Ergebnis der Aktion zurückkehrenden Belohnung und der Aktion selbst abgeleitet wird.

[0062] Es gibt zwei Verfahren zum Darstellen von $Q(s, a)$ -Werten in einem Computer, d. h. ein Verfahren, in dem Q -Werte von allen Zustands-Aktions-Paaren (s, a) in einer Tabelle (Aktionswerttabelle) gespeichert werden, und ein Verfahren, in dem eine Funktion zum Annähern von $Q(s, a)$ -Werten erstellt wird. Bei letzterem Verfahren kann das oben beschriebene Aktualisierungsverfahren verwirklicht werden, indem ein Parameter für eine Näherungsfunktion mittels eines stochastischen Gradientenverfahrens oder dergleichen angepasst wird. Als die Näherungsfunktion kann ein wie später beschriebenes neuronales Netzwerk verwendet werden.

[0063] Als ein Näherungsalgorithmus für eine Wertefunktion im überwachten Lernen, dem unüberwachten Lernen und dem bestärkenden Lernen kann ein neuronales Netzwerk verwendet werden. Das neuronale Netzwerk besteht beispielsweise aus einer Arithmetik-Einheit, einem Speicher und dergleichen, die ein Modell eines Neurons imitieren, wie in **Fig. 7** gezeigt. Die **Fig. 7** ist eine schematische Ansicht eines Modells eines Neurons.

[0064] Wie in **Fig. 7** gezeigt erzeugt ein Neuron eine Ausgabe y in Erwiderung auf einer Mehrheit von Eingaben x (zum Beispiel Eingaben x_1 bis x_3). Jede der Eingaben x_1 bis x_3 wird mit einem der Eingabe x entsprechenden Gewicht w (w_1 bis w_3) beaufschlagt. Somit erzeugt das Neuron die durch die folgende Gleichung dargestellte Ausgabe y . Es wird angemerkt, dass die Eingabe x , die Ausgabe y und das Gewicht w allesamt Vektoren sind.

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

wobei θ ein Bias (Fehlerfaktor) ist und f_k die Aktivierungsfunktion ist.

[0065] Als Nächstes wird ein dreischichtiges neuronales Netzwerk mit Gewichten von drei Schichten, das aus einer Kombination von den oben beschriebenen Neuronen gebildet wird, mit Bezug auf **Fig. 8** beschrieben. **Fig. 8** ist ein schematisches Schaubild eines neuronalen Netzwerks mit Gewichten von drei Schichten D1 bis D3.

[0066] Wie in **Fig. 8** gezeigt werden eine Mehrzahl von Eingaben x (zum Beispiel Eingaben x_1 bis x_3) von der linken Seite des neuronalen Netzwerks eingegeben und ein Ergebnis y (zum Beispiel Ergebnisse y_1 bis y_3) wird von dessen rechten Seite ausgegeben.

[0067] Genauer gesagt werden die Eingaben x_1 bis x_3 in jede der drei Neuronen N11 bis N13 eingegeben, während sie entsprechend gewichtet werden. Die den Eingaben auferlegten Gewichte werden insgesamt als W_1 bezeichnet.

[0068] Die Neuronen N11 bis N13 geben die jeweiligen Vektoren Z_{11} bis Z_{13} aus. Die Vektoren Z_{11} bis Z_{13} werden insgesamt durch einen Merkmalsvektor Z_1 bezeichnet, der als ein Vektor betrachtet wird, der einen

Merkmalsbetrag aus dem Eingabevektor extrahiert. Der Merkmalsvektor Z1 ist ein Merkmalsfaktor zwischen dem Gewicht W1 und einem Gewicht W2.

[0069] Die Vektoren Z11 bis Z13 werden in jede der zwei Neuronen eingegeben, während sie entsprechend gewichtet werden. Die den Merkmalsvektoren auferlegten Gewichte werden insgesamt als W2 bezeichnet.

[0070] Die Neuronen N21 und N22 geben die j Vektoren Z21 beziehungsweise Z22 aus. Die Vektoren Z21 und Z22 werden insgesamt durch einen Merkmalsvektor Z2 bezeichnet.

[0071] Der Merkmalsvektor Z2 ist ein Merkmalsvektor zwischen dem Gewicht W2 und einem Gewicht W3.

[0072] Die Merkmalsvektoren Z21 und Z22 werden in jede der drei Neuronen N31 bis N33 eingegeben, während sie entsprechend gewichtet werden. Die den Merkmalsvektoren auferlegten Gewichte werden insgesamt als W3 bezeichnet.

[0073] Zum Schluss geben die Neuronen N31 bis N33 die jeweiligen Ergebnisse y1 bis y3 aus.

[0074] Das neurale Netzwerk besitzt einen Lernmodus und einen Wertvorhersagemodus. Im Lernmodus wird das Gewicht W unter Verwendung eines Lerndatensatzes gelernt. Im Wertvorhersagemodus wird die Aktion des Roboters unter Verwendung des in dem Lernmodus erhaltenen Parameters bestimmt (das Wort „Vorhersage“ wird er Einfachheit halber verwendet, es können jedoch verschiedene Aufgaben ausgeführt werden, die Detektion, Klassifizierung, Schlussfolgerung und dergleichen umfassen).

[0075] Im Wertvorhersagemodus können Daten, die durch einen tatsächlichen Betrieb des Roboters erhalten werden, sofort gelernt und in der nächsten Aktion wiedergegeben werden (Online-Lernen). Ferner kann das Lernen gemeinsam unter Verwendung einer vorab gesammelten Datengruppe ausgeführt werden, und ein Detektionsmodus kann danach unter Verwendung des Parameters ausgeführt werden (Batch-Lernen). Auf eine dazwischenliegende Weise kann der Lernmodus immer dann ausgeführt werden, wenn sich eine bestimmte Menge an Daten angesammelt hat.

[0076] Die Gewichte W1 bis W3 können unter Verwendung eines Fehlerrückführungsalgorithmus (Rückführungsalgorithmus) gelernt werden. Informationen über einen Fehler treten von rechts ein und breiten sich nach links aus. Der Fehlerrückführungsalgorithmus ist ein Verfahren, in dem die Gewichte in Bezug auf jedes Neuron angepasst (gelernt) werden, um den Unterschied zwischen einer Ausgabe y und einer Ist-Ausgabe y (Überwacher) in Erwiderung einer Eingabe x zu minimieren.

[0077] Ein solches neurales Netzwerk kann mehr als drei Schichten aufweisen (als Deep-Learning bezeichnet). Eine Arithmetik-Einheit, die eine Merkmalextraktion aus Eingaben in Stufen und der Regression von Ergebnissen ausführt, kann automatisch nur von Überwachungsdaten erfasst werden.

[0078] Um demzufolge das oben beschriebene Q-Lernen auszuführen, wie in **Fig. 2** gezeigt, umfasst die maschinelle Lernvorrichtung **101** gemäß dieser Ausführungsform die Zustandsbeobachtungseinheit **11**, die Lerneinheit **13** und die Entscheidungsbestimmungseinheit **16**. Ein auf die vorliegende Erfindung angewendetes maschinelles Lernverfahren ist jedoch nicht auf das Q-Lernen beschränkt. Beispielsweise entspricht im Falle der Anwendung des überwachten Lernens eine Wertfunktion einem Lernmodell, und eine Belohnung entspricht einem Fehler.

[0079] Wie in **Fig. 2** gezeigt umfasst der Zustand des Roboters **20** einen durch eine Aktion indirekt geänderten Zustand und einen durch eine Aktion direkt geänderten Zustand. Der durch eine Aktion indirekt geänderte Zustand umfasst die Zykluszeit, das Drehmoment und die Vibration. Der durch eine Aktion direkt geänderte Zustand umfasst einen Stromwert.

[0080] Die Lerneinheit **3** aktualisiert einen Aktionswert, der einer aktuellen Zustandsvariablen entspricht, und eine in der Aktionswerttabelle zu ergreifenden Aktion basierend auf der Aktualisierungsgleichung und der Belohnung.

[0081] In dem Beispiel von **Fig. 2** wird die eigene Aktionswerttabelle aktualisiert, indem die durch die Lerneinheit der eigenen maschinellen Lernvorrichtung aktualisierte Aktionswerttabelle verwendet wird, ist jedoch nicht darauf beschränkt. Die eigene Aktionswerttabelle kann aktualisiert werden, indem eine durch eine Lerneinheit

einer von der eigenen maschinellen Lernvorrichtung unterschiedlichen anderen maschinellen Lernvorrichtung aktualisierte Aktionswerttabelle verwendet wird.

[Zweite Ausführungsform]

[0082] Es wird eine maschinelle Lernvorrichtung gemäß einer zweiten Ausführungsform der vorliegenden Erfindung beschrieben. **Fig. 9** ist ein Blockschaltbild der maschinellen Lernvorrichtung gemäß der zweiten Ausführungsform der vorliegenden Erfindung. Der Unterschied zwischen einer maschinellen Lernvorrichtung **102** gemäß der zweiten Ausführungsform der Erfindung und der maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform besteht darin, dass die maschinelle Lernvorrichtung **102** mit einem ersten Roboter **21** und einem zweiten Roboter **22** durch eine Kommunikationseinheit **18**, d. h., einem Netzwerk verbunden ist und die Zustandsbeobachtungseinheit **11** eine aktuelle Zustandsvariable durch das Netzwerk erhält. Die anderen Konfigurationen der maschinellen Lernvorrichtung **102** gemäß der zweiten Ausführungsform sind dieselben wie die der maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform, und daher wird auf deren ausführliche Beschreibung verzichtet.

[0083] Die maschinelle Lernvorrichtung **102** ist bevorzugt in einem Cloud-Server installiert.

[0084] Es kann aufgrund einer hohen Verarbeitungslast für einen Roboter schwierig werden, ein Lernen während eines Hochgeschwindigkeitsbetriebs auszuführen. Gemäß der Konfiguration der maschinellen Lernvorrichtung der zweiten Ausführungsform der vorliegenden Erfindung wird das Lernen durch eine von der Robotersteuerung unterschiedlichen Vorrichtung ausgeführt, womit die Belastung für den Roboter reduziert wird.

[Dritte Ausführungsform]

[0085] Es werden maschinelle Lernvorrichtungen gemäß einer dritten Ausführungsform der vorliegenden Erfindung beschrieben. **Fig. 10** ist ein Blockschaltbild der maschinellen Lernvorrichtung gemäß der dritten Ausführungsform der vorliegenden Erfindung. Der Unterschied zwischen den maschinellen Lernvorrichtungen **103-1** und **103-2** gemäß der dritten Ausführungsform der vorliegenden Erfindung und der maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform besteht darin, dass eine maschinelle Lernvorrichtung in jeder einer Mehrzahl von durch einen Netzknoden **19** verbundenen Robotersteuerungen vorgesehen wird. Die anderen Konfigurationen der maschinellen Lernvorrichtungen **103-1** und **103-2** gemäß der dritten Ausführungsform sind dieselben wie die der maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform, und daher wird auf deren ausführliche Beschreibung verzichtet.

[0086] Eine erste maschinelle Lernvorrichtung **103-1** ist in einer ersten Robotersteuerung **201** zum Steuern eines ersten Roboters **21** vorgesehen. Eine zweite maschinelle Lernvorrichtung **103-2** ist in einer zweiten Robotersteuerung **202** zum Steuern eines zweiten Roboters **22** vorgesehen.

[0087] Diese Struktur ermöglicht die gemeinsame Nutzung einer durch eine der maschinellen Lernvorrichtungen unter den maschinellen Lernvorrichtungen erstellte Aktionswerttabelle und verbessert somit die Effizienz des Lernens.

[Vierte Ausführungsform]

[0088] Es wird eine maschinelle Lernvorrichtung gemäß einer vierten Ausführungsform der vorliegenden Erfindung beschrieben. **Fig. 11** ist ein Blockschaltbild der maschinellen Lernvorrichtung gemäß der vierten Ausführungsform der vorliegenden Erfindung. Der Unterschied zwischen einer maschinellen Lernvorrichtung **104** gemäß der vierten Ausführungsform der vorliegenden Erfindung und der maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform besteht darin, dass die maschinelle Lernvorrichtung **104** die Positionen und Stellungen von auf einem Förderer angeordneten Objekten aufzeichnet und sammelt und die Positionen und Stellungen der Objekte (p1 bis p6) auf dem Förderer **10** offline reproduziert, um die optimale Greifreihenfolge der Objekte zu lernen. Die anderen Konfigurationen der maschinellen Lernvorrichtung **104** gemäß der vierten Ausführungsform sind dieselben wie die der maschinellen Lernvorrichtung **101** gemäß der ersten Ausführungsform, und daher wird auf deren ausführliche Beschreibung verzichtet.

[0089] Eine erste Robotersteuerung **201** steuert einen ersten Roboter **21**, während eine zweite Robotersteuerung **202** einen zweiten Roboter **22** steuert. Die maschinelle Lernvorrichtung **104** ist außerhalb der ersten Robotersteuerung **201** und der zweiten Robotersteuerung **202** vorgesehen. Die maschinelle Lernvorrichtung **104** empfängt Daten („LOG“) über die Positionen und Stellungen der Objekte, die aus durch (nicht gezeigte)

Kameras aufgenommenen Bildern eingeholt werden, wobei die Kameras jeweils in der Nähe von jeweils dem ersten Roboter **21** und dem zweiten Roboter **22** vorgesehen sind, und lernt die optimale Greifreihenfolge der Objekte. Die Lernergebnisse werden an die erste Robotersteuerung **201** und die zweite Robotersteuerung **202** gesendet, sodass der erste Roboter **21** und der zweite Roboter **22** die Objekte in der optimalen Reihenfolge greifen können.

[0090] Gemäß der maschinellen Lernvorrichtung und dem maschinellen Lernverfahren der Ausführungsformen der vorliegenden Erfindung ist es möglich, sowohl die Belastung des Roboters zu reduzieren als auch die Zykluszeit zu minimieren, die die Zeit ist, die von dem Roboter benötigt wird, um die Mehrzahl von Objekten unter Verwendung der die Funktion des Greifens der Objekte aufweisenden Hand in Behältern zu verstauen.

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

- JP 2014-104524 [0003]

Patentansprüche

1. Maschinelle Lernvorrichtung zum Lernen eines Betriebszustands eines Roboters, der eine Mehrzahl von auf einer Trägervorrichtung angeordneten Objekten unter Verwendung einer Hand zum Greifen der Objekte in einem Behälter verstaut, wobei die maschinelle Lernvorrichtung umfasst:
eine Zustandsbeobachtungseinheit (**11**) zum Beobachten, während des Betriebs des Roboters, der Positionen und Stellungen der Objekte und einer Zustandsvariablen, die mindestens eines von einer Zykluszeit zum Verstauen der Objekte in dem Behälter und eines Drehmoments und einer Vibration, die auftreten, wenn der Roboter die Objekte greift, umfasst;
eine Bestimmungsdaten-Einholungseinheit (**12**) zum Einholen von Bestimmungsdaten zum Bestimmen einer Spanne von jeweils der Zykluszeit, dem Drehmoment und der Vibration in Abhängigkeit von einem zulässigen Toleranzwert; und
eine Lerneinheit (**13**) zum Lernen des Betriebszustands des Roboters in Übereinstimmung mit einem Trainingsdatensatz, der aus einer Kombination aus der Zustandsvariablen und den Bestimmungsdaten gebildet wird.
2. Maschinelle Lernvorrichtung nach Anspruch 1, wobei die Zykluszeit die Zeit ist, ab der der Roboter beginnt, eine vorgeschriebene Anzahl von Objekten in einem Behälter zu verstauen, bis der Roboter das Verstauen der Objekte in dem Behälter abschließt.
3. Maschinelle Lernvorrichtung nach Anspruch 1 oder 2, wobei das Drehmoment basierend auf einem durch einen Motor zum Antreiben des Roboters fließenden Strom berechnet wird.
4. Maschinelle Lernvorrichtung nach einem der Ansprüche 1 bis 3, wobei die Vibration basierend auf einer durch einen in der Hand vorgesehenen Beschleunigungssensor detektierten Beschleunigung berechnet wird.
5. Maschinelle Lernvorrichtung nach einem der Ansprüche 1 bis 4, wobei die Lerneinheit (**13**) umfasst:
eine Belohnungsberechnungseinheit (**14**) zum Berechnen einer Belohnung basierend auf den Bestimmungsdaten; und
eine Wertfunktions-Aktualisierungseinheit (**15**) zum Aktualisieren einer Wertfunktion, die zum Schätzen der Greifreihenfolge der Objekte verwendet wird, um mindestens eines von der Zykluszeit, dem Drehmoment und der Vibration basierend auf der Belohnung zu reduzieren.
6. Maschinelle Lernvorrichtung nach Anspruch 5, wobei die Lerneinheit (**13**) eine Aktionswerttabelle, die der Greifreihenfolge der Objekte entspricht, basierend auf der Zustandsvariablen von mindestens einem von der Zykluszeit, dem Drehmoment und der Vibration und der Belohnung aktualisiert.
7. Maschinelle Lernvorrichtung nach Anspruch 6, wobei die Lerneinheit (**13**) eine Aktionswerttabelle, die mindestens einem von der Zykluszeit, dem Drehmoment und der Vibration entspricht, wenn ein anderer Roboter mit derselben Konfiguration wie der Roboter andere Objekte in einem Behälter verstaut, basierend auf einer Zustandsvariablen des anderen Roboters und der Belohnung aktualisiert.
8. Maschinelle Lernvorrichtung nach einem der Ansprüche 5 bis 7, wobei eine Belohnungsberechnungseinheit (**14**) die Belohnung basierend auf mindestens einem von der Zykluszeit, dem Drehmoment und der Vibration berechnet.
9. Maschinelle Lernvorrichtung nach einem der Ansprüche 1 bis 8, ferner umfassend eine Entscheidungsbestimmungseinheit (**16**) zum Bestimmen der Greifreihenfolge der Objekte basierend auf einem Lernergebnis durch die Lerneinheit (**13**) in Übereinstimmung mit dem Trainingsdatensatz.
10. Maschinelle Lernvorrichtung nach einem der Ansprüche 1 bis 9, wobei die maschinelle Lernvorrichtung (**102**) mit dem Roboter durch ein Netzwerk verbunden ist, und die Zustandsbeobachtungseinheit (**11**) eine aktuelle Zustandsvariable durch das Netzwerk erhält.
11. Maschinelle Lernvorrichtung nach einem der Ansprüche 1 bis 10, wobei die maschinelle Lernvorrichtung (**102**) in einem Cloud-Server installiert ist.
12. Maschinelles Lernverfahren zum Lernen eines Betriebszustands eines Roboters, der eine Mehrzahl von auf einer Trägervorrichtung angeordneten Objekten in einem Behälter unter Verwendung einer Hand zum Greifen der Objekte verstaut, wobei das maschinelle Lernverfahren die Schritte umfasst:

Beobachten, während des Betriebs des Roboters, der Positionen und Stellungen der Objekte und einer Zustandsvariablen, die mindestens eines von einer Zykluszeit zum Verstauen der Objekte in dem Behälter und eines Drehmoments und einer Vibration, die auftreten, wenn der Roboter die Objekte greift, umfasst;
Einholen von Bestimmungsdaten zum Bestimmen einer Spanne von jeweils der Zykluszeit, dem Drehmoment und der Vibration in Abhängigkeit von einem zulässigen Toleranzwert; und Lernen des Betriebszustands des Roboters in Übereinstimmung mit einem Trainingsdatensatz, der aus einer Kombination aus der Zustandsvariablen und den Bestimmungsdaten gebildet wird.

Es folgen 10 Seiten Zeichnungen

Anhängende Zeichnungen

FIG. 1

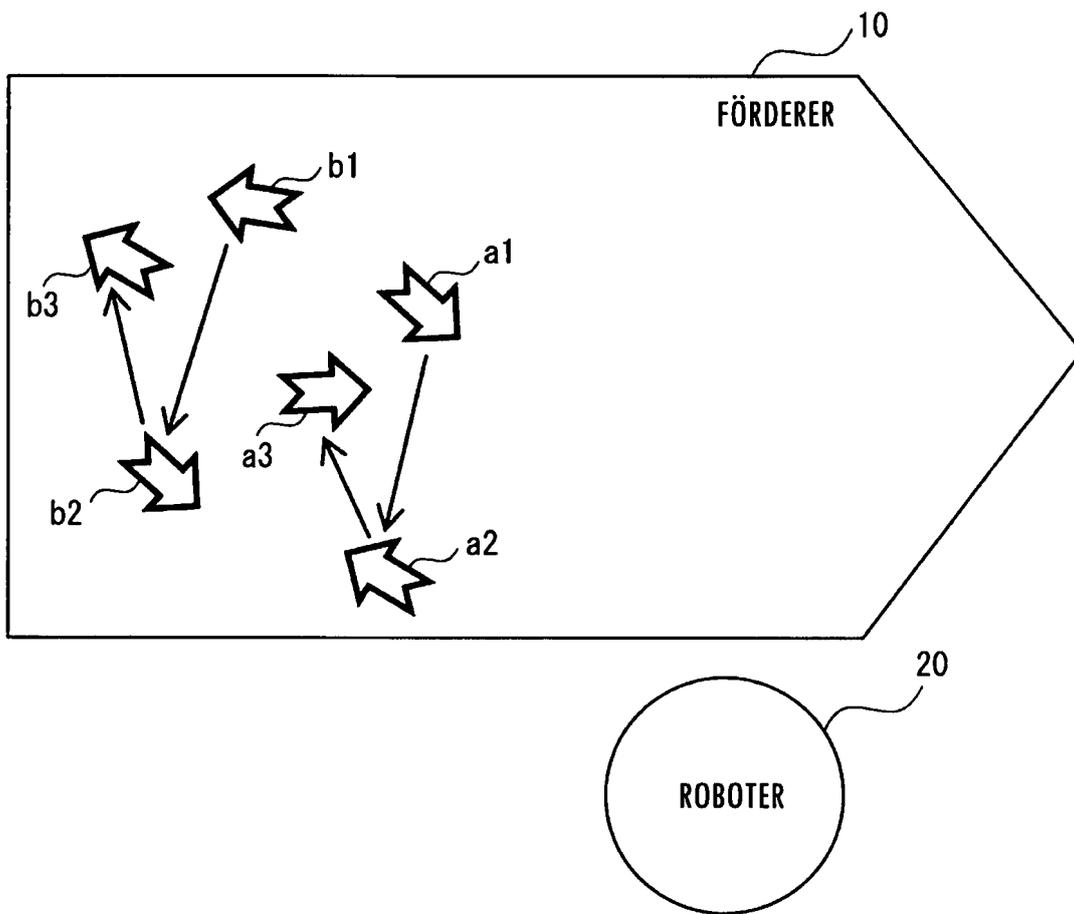


FIG. 2

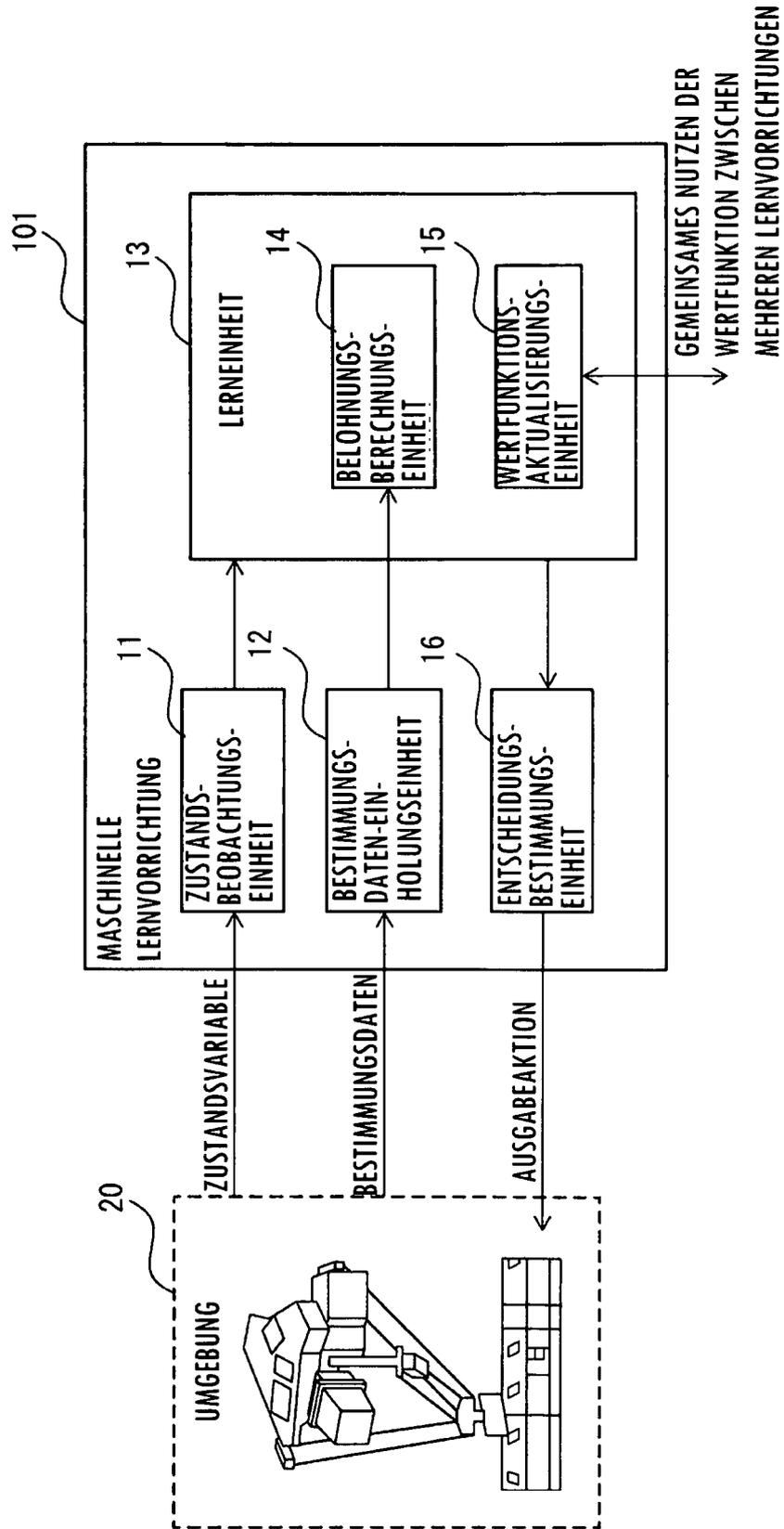


FIG. 3

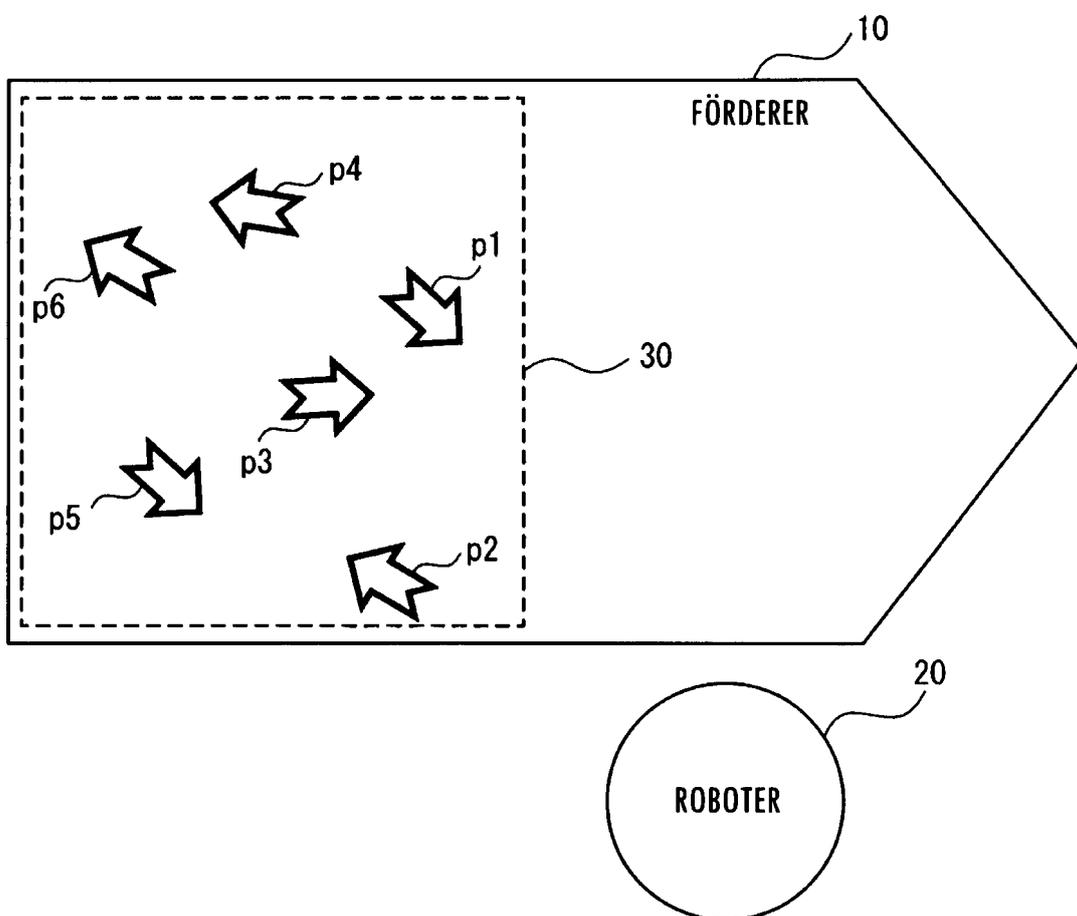


FIG. 4

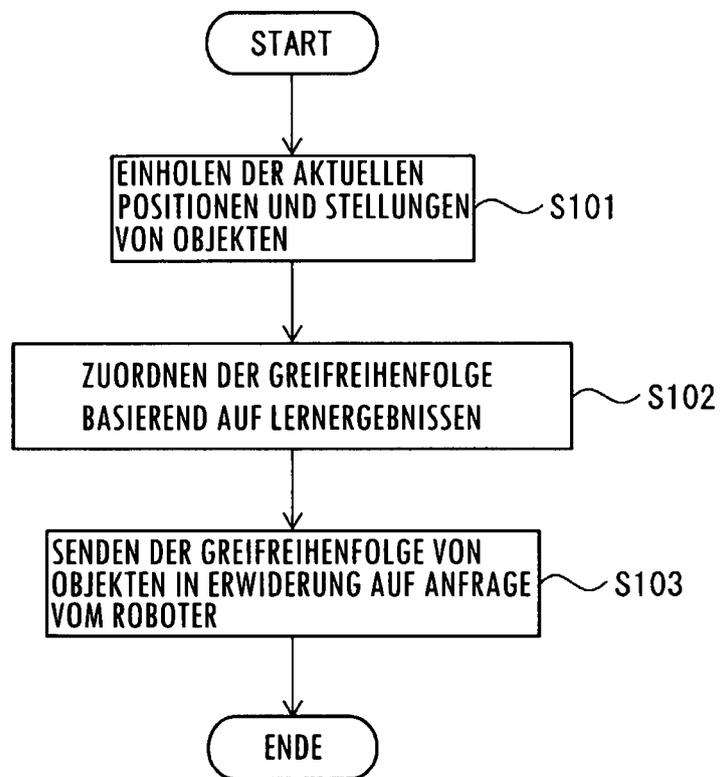


FIG. 5

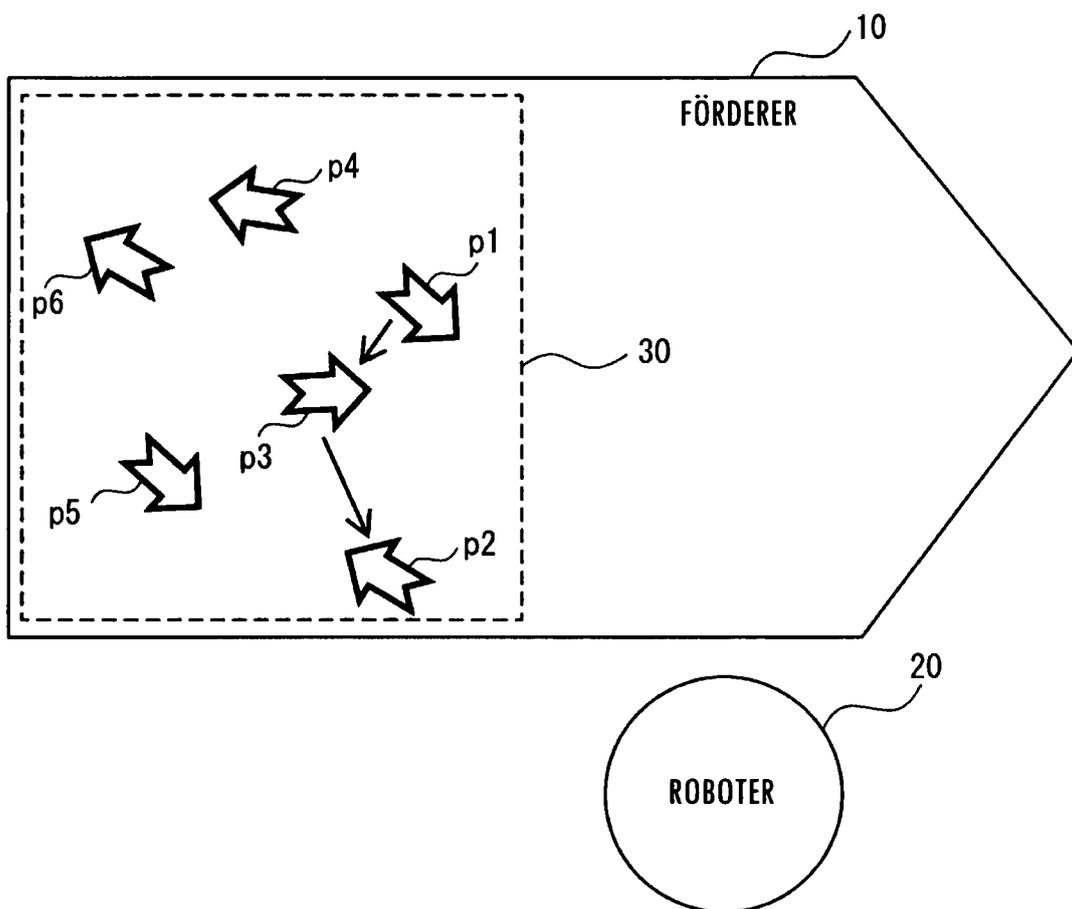


FIG. 6

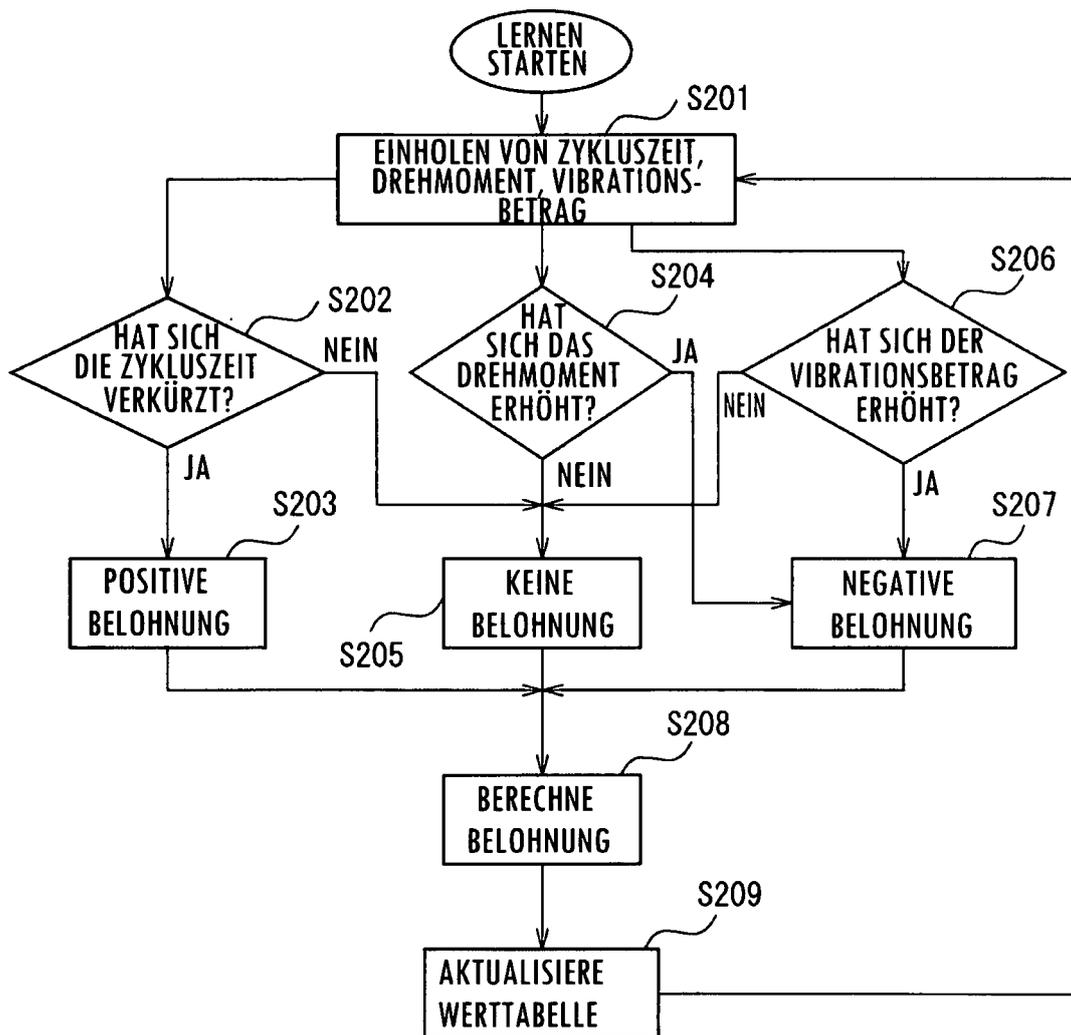


FIG. 7

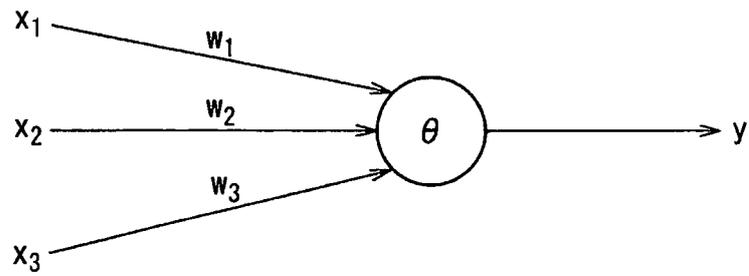


FIG. 8

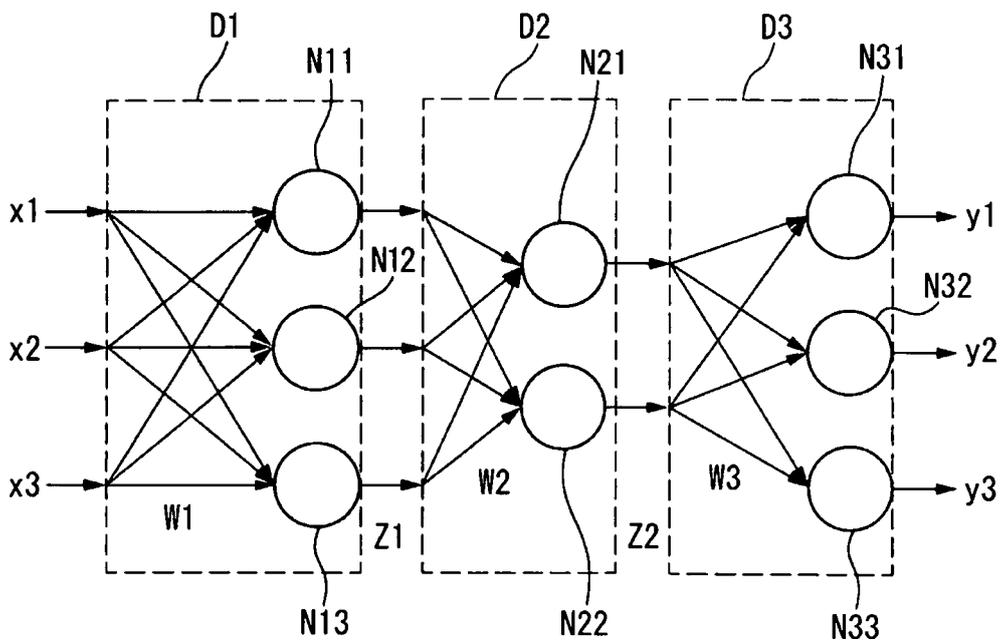


FIG. 9

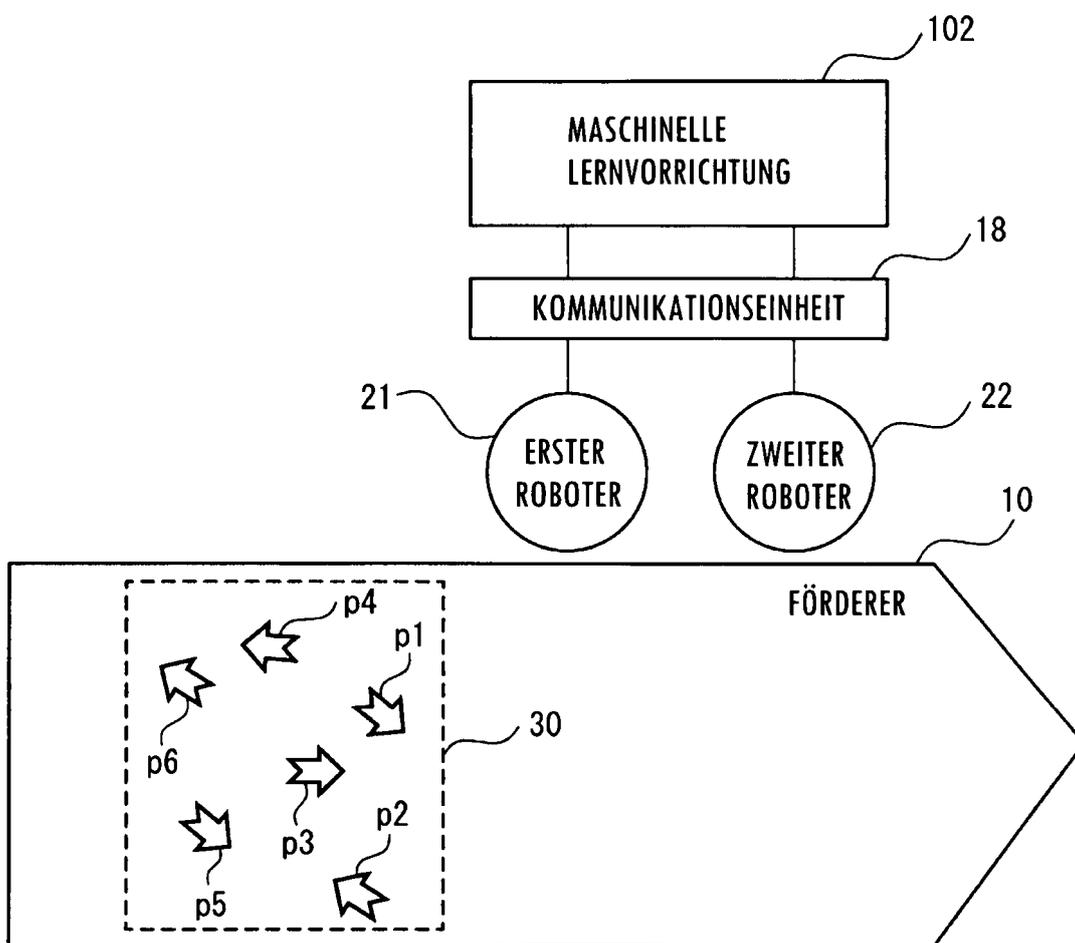


FIG. 10

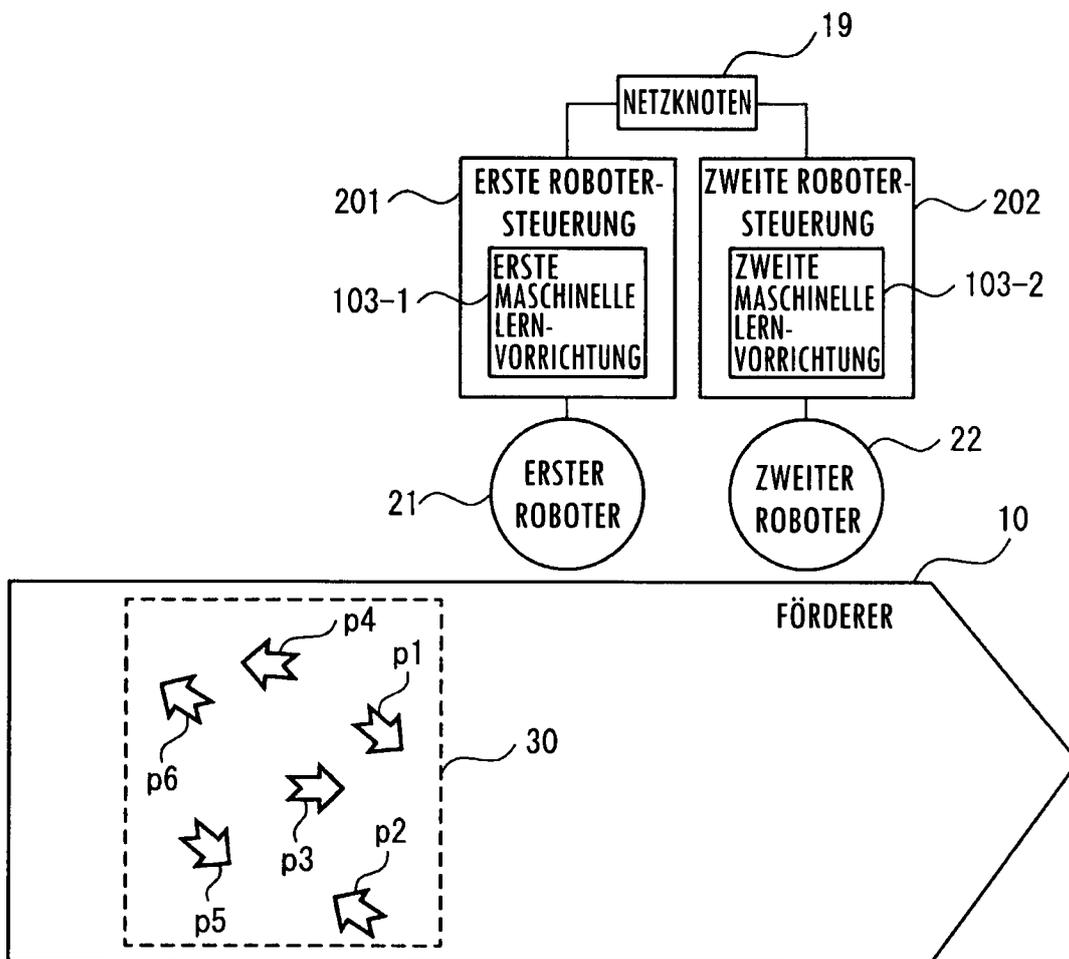


FIG. 11

