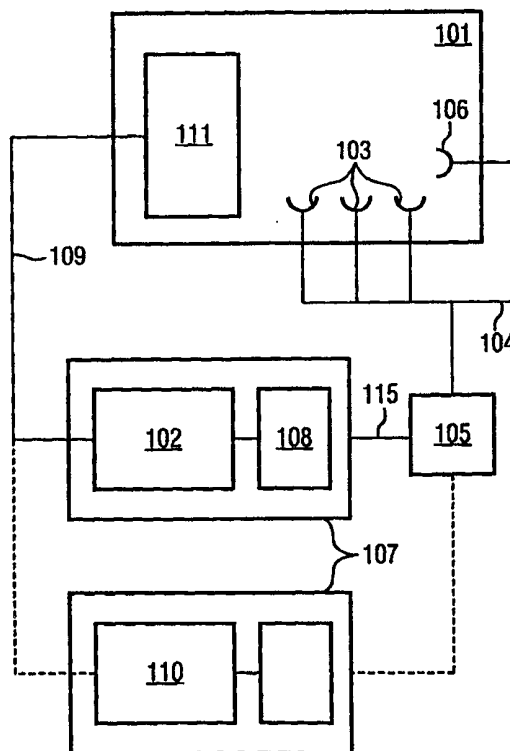



 INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation ⁶ : G06F 17/00	A2	(11) Internationale Veröffentlichungsnummer: WO 99/48020 (43) Internationales Veröffentlichungsdatum: 23. September 1999 (23.09.99)
(21) Internationales Aktenzeichen: PCT/DE99/00524 (22) Internationales Anmeldedatum: 26. Februar 1999 (26.02.99) (30) Prioritätsdaten: 198 11 906.2 18. März 1998 (18.03.98) DE (71) Anmelder (für alle Bestimmungsstaaten ausser US): SIEMENS AKTIENGESELLSCHAFT [DE/DE]; Wittelsbacherplatz 2, D-80333 München (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US): LANG, Bernhard [DE/DE]; Horemannstrasse 16/III, D-80636 München (DE). POPPE, Thomas [DE/DE]; Klematisstrasse 8, D-80689 München (DE). (74) Gemeinsamer Vertreter: SIEMENS AKTIENGESELLSCHAFT; Postfach 22 16 34, D-80506 München (DE).		(81) Bestimmungsstaaten: JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Veröffentlicht <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>
(54) Title: METHOD AND ARRANGEMENT FOR COMPUTER-ASSISTED DETERMINATION OF A REPRESENTATION SPECIFICATION		
(54) Bezeichnung: VERFAHREN UND ANORDNUNG ZUR RECHNERGESTÜTZTEN ERMITTLUNG EINER ABBILDUNGSVORSCHRIFT		
(57) Abstract The invention relates to a method and an arrangement for computer-assisted determination of a representation specification, describing a technical process in a predetermined space, wherein said technical process in a predetermined space is described by partial models. An approximative representation specification describing approximatively the technical process and including first partial models is determined. Subsequently, the representation specification is determined in such a way that the first partial models can be expanded with at least a second partial model, whereby the representation specification can be stabilized in comparison with the approximative representation specification. (57) Zusammenfassung Es wird ein Verfahren und eine Anordnung zur rechnergestützten Ermittlung einer Abbildungsvorschrift, welche einen technischen Prozeß in einem vorgegebenen Raum beschreibt, wobei der technische Prozeß in dem vorgegebenen Raum durch Teilmodelle beschrieben wird, angegeben. Dabei wird eine approximative Abbildungsvorschrift, welche den technischen Prozeß approximativ beschreibt und welche erste Teilmodelle umfaßt, ermittelt. Anschließend wird die Abbildungsvorschrift derart ermittelt, daß die ersten Teilmodelle mit mindestens einem zweiten Teilmodell erweitert werden, wobei die Abbildungsvorschrift im Vergleich zu der approximativen Abbildungsvorschrift stabilisiert wird.		



LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

Beschreibung

Verfahren und Anordnung zur rechnergestützten Ermittlung einer Abbildungsvorschrift

5

Die Erfindung bezieht sich auf die Ermittlung einer Abbildungsvorschrift. Durch eine solche Abbildungsvorschrift wird im Rahmen einer Abbildung durch ein neuronales Netz oder durch ein Fuzzy-System eine Eingangsgröße auf eine Ausgangsgröße abgebildet.

10

Eine Ermittlung einer derartigen Abbildungsvorschrift für eine Abbildung durch ein sogenanntes Radiale-Basisfunktionen-Netz (RBF-Netz) vom Typ Local Linear Map oder Multi Adaline ist aus [1] bekannt.

15

Grundlagen eines Radiale-Basisfunktionen-Netzes (RBF-Netz) sind aus [4] bekannt.

Ein Radiale-Basisfunktionen-Netz ist ein spezielles vorwärtsgerichtetes neuronales Netz, welches eine Abbildung beschreibt, durch die eine Eingangsgröße auf eine Ausgangsgröße abgebildet wird, und welches nur eine Schicht von verdeckten Neuronen besitzt. Ein verdecktes Neuron weist jeweils eine spezielle, radialsymmetrische (radiale) Zugehörigkeitsfunktion (Aktivierungsfunktion) auf.

20

25

Eine Zugehörigkeitsfunktion bzw. Aktivierungsfunktion ist mathematisch gesehen eine Basisfunktion eines Teilsystems (Teilmodell) eines Systems von Funktionen auf einem Raum (Gesamtmodell). Das Gesamtmodell wird für eine Abbildung, welche eine Eingangsgröße auf eine Ausgangsgröße in dem Raum abbildet, eingesetzt.

30

2

Dabei beschreibt eine Zugehörigkeitsfunktion bzw. Aktivierungsfunktion eine Zugehörigkeit eines Datenpunkts in dem Raum zu dem entsprechenden Teilmodell.

- 5 Das Teilmodell weist ein sogenanntes Zentrum auf, welches ein Datenpunkt in dem Raum ist, für welchen Datenpunkt die Zugehörigkeitsfunktion ein absolutes Maximum liefert. Mit zunehmendem Abstand eines ausgewählten Datenpunkts von dem Zentrum verringert sich der Wert der Zugehörigkeitsfunktion für den
10 ausgewählten Datenpunkt.

Aus [4] ist bekannt, als eine Zugehörigkeitsfunktion eine Gaußfunktion zu verwenden.

- 15 Ein RBF-Netz, wie es aus [4] bekannt ist, läßt sich wie folgt beschreiben:

Es sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion von $\mathbb{R}^n$ nach $\mathbb{R}$, die durch N Stützstellen gegeben ist. Jede Stützstelle i besteht aus einem n -dimensionalen Eingabevektor $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ und einem dazugehörigen reellwertigen Funktionswert $y_i \in \mathbb{R}$. Als Interpolationsbedingung gilt:

$$f(\mathbf{x}_i) = y_i \quad \text{für alle } i=1, \dots, N. \quad (1)$$

25

Das für eine Approximation verwendete Funktionensystem besteht aus radialsymmetrischen Basisfunktionen h_i , die jeweils einer Stützstelle als jeweiligem Zentrum zugeordnet sind. Sie sind wie folgt definiert:

30

$$h_i(\mathbf{x}) := f(||\mathbf{x} - \mathbf{x}_i||). \quad (2)$$

Die Basisfunktion h_i hat positive skalare Werte. Sie ist nur abhängig von einem Abstand eines Vektors $\mathbf{X}$ von der entsprechenden Stützstelle $\mathbf{X}_i$, die einem Neuron i zugeordnet ist, in einer beliebigen, vorgegebenen Norm. Meist wird ein euklidischer Abstand zwischen den Vektoren $\mathbf{X}$ und $\mathbf{X}_i$ im $\mathbf{R}^n$ als Norm verwendet. Damit erhält man folgende Darstellung für das RBF-Netz:

$$f(\mathbf{X}) = \sum_{i=1}^N c_i * h_i(\|\mathbf{X} - \mathbf{X}_i\|) \quad (3)$$

10

mit:

c_i : Gewichtungsfaktor für das Zentrum i .

In dieser Darstellung weist die radiale Basisfunktion h_i einen Index i auf, weil die radiale Basisfunktion h_i jeweils für ein zugeordnetes Zentrum i unterschiedliche Charakteristika, wie beispielsweise eine Breite oder eine Varianz einer Gaußfunktion, aufweist.

Ferner ist aus [4] ein Verfahren zu einem Training eines RBF-Netzes bekannt.

Das aus [1] bekannte RBF-Netz vom Typ Multi Adaline weist anstelle des Gewichtungsfaktors c_i ein sogenanntes lineares

Adaline act_k auf und läßt sich wie folgt beschreiben:

$$y = \sum_{k=1}^{nk} \chi_k * act_k \quad (4)$$

4

wobei:

$$\text{act}_k = \sum_{i=1}^{n_i} (x_i - \mu_{i,k}) * a_{i,k} + b_k \quad (5)$$

$$\chi_k = \frac{g_k}{\sum_{k=1}^{n_k} g_k} \quad (6)$$

$$g_k = \exp \left(- \frac{1}{2 * \sigma_k^2} * \sum_{i=1}^{n_i} (x_i - \mu_{i,k})^2 \right) \quad (7)$$

5 mit:

i, k Laufvariable

n_k Zahl der Zentrenn_i Wert der Dimensionalität eines Eingangsraumsa_{i,k} Steigung10 b_k Biaswertχ_k Partitionsfunktionμ_{i,k} Zentrumx_{i,y} Eingangsgröße, Ausgangsgrößeσ_k Varianz

15

Aus [2] und [3] ist ein sogenanntes Fuzzy-Clustering-Verfahren zur Datenanalyse bekannt.

Im Rahmen des Fuzzy-Clusterings werden c Cluster und entsprechende Zugehörigkeiten von Datenvektoren $\underline{x}_k$ derart bestimmt, daß Datenvektoren, die in einem Datenraum nahe bei einem Cluster liegen, eine möglichst hohe Zugehörigkeit und weit von dem Cluster entfernt liegende Datenvektoren $\underline{x}_k$ eine möglichst geringe Zugehörigkeit zu dem jeweiligen Cluster aufweisen.

25

5

Dies wird durch Minimierung einer Summe der mit Zugehörigkeiten u_{ik}^m gewichteten quadratischen euklidischen Abständen d_{ik}^2 erreicht. Es soll also eine Menge X von Datenvektoren $\underline{x}_k$

$$5 \quad X = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k, \dots, \underline{x}_n\}$$

in c Cluster (Untermengen der Menge von Datenvektoren) gruppiert werden.

10 Die Cluster werden beschrieben durch eine Zugehörigkeitsmatrix $\underline{U}$, welche c Zeilen und n Spalten aufweist. Jedes Element u_{ik} der Zugehörigkeitsmatrix $\underline{U}$ weist einen Wert innerhalb des Intervalls $[0, 1]$ auf und beschreibt eine Zugehörigkeit des Datenvektors $\underline{x}_k$ zu dem i -ten Cluster.

15

Die Summe der Zugehörigkeiten des Datenvektors $\underline{x}_k$ in den c Clustern muß folgender Vorschrift genügen:

$$\sum_{i=1}^c u_{ik} = 1 \quad \forall k = 1 \dots n. \quad (8)$$

20

Ein Cluster muß mindestens ein Element enthalten, so daß gilt:

$$\sum_{k=1}^n u_{ik} > 0 \quad \forall i = 1 \dots c. \quad (9)$$

25

Die Kostenfunktion J_m der Zugehörigkeitswerte wird gemäß folgender Vorschrift gebildet:

$$J_m = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \cdot d_{ik}^2. \quad (10)$$

30

Ein Abstand d_{ik} wird gebildet gemäß folgender Vorschrift:

$$d_{ik} = \|\underline{x}_k - \underline{v}_i\|_A = \sqrt{(\underline{x}_k - \underline{v}_i)^T \cdot \underline{A} \cdot (\underline{x}_k - \underline{v}_i)}. \quad (11)$$

Es wird mit $\underline{A}$ eine vorgebbare induzierte Norm des inneren Produkts gemäß Vorschrift (4) bezeichnet, die üblicherweise durch die Identitätsmatrix gegeben ist (Euklidischer Abstand).

Die Minimierung der Kostenfunktion J_m erfolgt unter Einsatz einer sogenannten Picard-Iteration.

Nacheinander werden Zugehörigkeitswerte u_{ik} und Clusterzentren $\underline{v}_i$ gemäß folgenden Vorschriften gebildet:

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}}, \quad (12)$$

$$\underline{v}_i = \frac{\sum_{k=1}^n u_{ik}^m \cdot \underline{x}_k}{\sum_{k=1}^n u_{ik}^m}. \quad (13)$$

Die Ermittlung der Zugehörigkeitswerte u_{ik} und der Clusterzentren $\underline{v}_i$ wird solange wiederholt, bis eine festgelegte Anzahl von Iterationen durchgeführt worden ist oder bis eine Änderung der Zugehörigkeitswerte u_{ik} und/oder bis eine Änderung der Clusterzentren $\underline{v}_i$ unter einem vorgegebenen Schwellenwert liegt.

Bei diesem oben beschriebenen, auch als Fuzzy-C-Means-Clustering bezeichneten Verfahren werden die Cluster durch ihre Clusterzentren $\underline{v}_i$ beschrieben.

Eine Abbildung, die durch eine bekannte Abbildungsvorschrift ein Systemverhalten eines technischen Systems bzw. Prozesses beschreibt, weist in vielen Fällen den Nachteil auf, daß die Abbildung mangelhaft stabil gegenüber einem fehlerhaften Datenpunkt oder Meßfehler ist.

Der Erfindung liegt das Problem zugrunde, ein Verfahren und eine Anordnung zur Ermittlung einer Abbildungsvorschrift einer Abbildung, welche einen technischen Prozeß beschreibt, anzugeben, mit welchen die Abbildung hinsichtlich der Stabilität gegenüber einem fehlerhaften Datenpunkt verbessert wird.

Das Problem wird durch das Verfahren sowie durch die Anordnung mit den Merkmalen gemäß den unabhängigen Patentansprüchen gelöst.

Bei einem Verfahren zur rechnergestützten Ermittlung einer Abbildungsvorschrift, welche einen technischen Prozeß in einem vorgegebenen Raum beschreibt, wobei der technische Prozeß in dem vorgegebenen Raum durch Teilmodelle beschrieben wird, wird eine approximative Abbildungsvorschrift, welche den technischen Prozeß approximativ beschreibt und welche erste Teilmodelle umfaßt, ermittelt. Anschließend wird die Abbildungsvorschrift derart ermittelt, daß die ersten Teilmodelle mit mindestens einem zweiten Teilmodell erweitert werden, wobei die Abbildungsvorschrift im Vergleich zu der approximativen Abbildungsvorschrift stabilisiert wird.

Eine Anordnung zur rechnergestützten Ermittlung einer Abbildungsvorschrift, welche einen technischen Prozeß in einem vorgegebenen Raum beschreibt, wobei der technische Prozeß in dem vorgegebenen Raum durch Teilmodelle beschrieben wird, weist einen Prozessor auf, der derart eingerichtet ist, daß eine approximative Abbildungsvorschrift, welche den technischen Prozeß approximativ beschreibt und welche erste Teilmo-

delle umfaßt, ermittelbar ist. Ferner ist der Prozessor derart eingerichtet, daß anschließend die Abbildungsvorschrift derart ermittelbar ist, daß die ersten Teilmodelle mit mindestens einem zweiten Teilmodell erweitert werden, wobei die
5 Abbildungsvorschrift im Vergleich zu der approximativen Abbildungsvorschrift stabilisiert wird.

Diese Anordnung ist insbesondere geeignet zur Durchführung des erfindungsgemäßen Verfahrens oder einer seiner nachfolgend erläuterten Weiterbildungen.
10

Der besondere Vorteil der Erfindung liegt darin, daß die Stabilität der Abbildung gegenüber einem fehlerhaften Datenpunkt erheblich verbessert wird. Gerade durch die Überlagerung der
15 ersten Teilmodelle mit dem zweiten Teilmodell wird eine Instabilität einer Abbildung, welche Abbildung nur erste Teilmodelle aufweist, und welche Instabilität insbesondere in einem Bereich in dem vorgegebenen Raum, der nur eine geringe Dichte von vorgegebenen Trainingsdatenpunkten aufweist, reduziert. Des Weiteren wird durch diese Vorgehensweise auch eine
20 Genauigkeit der Beschreibung des technischen Prozesses durch die Abbildungsvorschrift verbessert.

Bevorzugte Weiterbildungen der Erfindung ergeben sich aus den
25 abhängigen Ansprüchen.

Bevorzugt wird die Abbildungsvorschrift mit mehreren zweiten Teilmodellen erweitert und/oder das zweite Teilmodell mit mehreren ersten Teilmodellen überlagert.

30

Bei der Beschreibung komplexer technischer Prozesse sind/ist die Ausgangsgröße und/oder die Eingangsgröße mehrdimensionale/ein mehrdimensionale Größen/Größe.

Eine weitere Verbesserung der Beschreibung wird dadurch erreicht, daß die Überlagerung des zweiten Teilmodells unter Verwendung einer Gewichtungsfunktion durchgeführt wird.

- 5 In einer Weiterbildung wird die Gewichtungsfunktion unter Verwendung einer Summe von Gewichtungen ermittelt.

Bevorzugt wird die Abbildungsvorschrift im Rahmen eines neuronalen Netzes verwendet.

10

In einer Weiterbildung sind/ist das neuronale Netz und/oder das zweite Teilmodell vom Typ Multi Adaline.

- Bevorzugt wird die Abbildungsvorschrift wie folgt beschrieben:
15

$$y = \sum_{k=0}^{nk} \chi_k * act_k \quad (14)$$

wobei:

$$act_k = (x - \mu_k) * a_k + b_k \quad (15)$$

$$\chi_k = \frac{g_k}{\sum_{k=0}^{nk} g_k} = \frac{g_k}{g_0 + \sum_{k=1}^{nk} g_k} = \frac{g_k}{g_0 + SG} \quad (16)$$

20 $g_k = \exp \left(-\frac{1}{2 * \sigma_k^2} * (x - \mu_k)^2 \right) \quad (17)$

$$SG = \sum_{k=1}^{nk} g_k \quad (18)$$

mit:

- x, y Eingangsgröße, Ausgangsgröße
25 k Laufvariable
nk Zahl der Clusterzentren der ersten Teilmodelle
ak Steigung

	b_k	Biaswert
	SG	Summe von Gewichtungen
	g_k, g_0	Gewichtungsfunktionen
	χ_k	Partitionsfunktion
5	act_k	lineares Adaline
	μ_k	Clusterzentrum
	σ_k	Varianz

10 Bevorzugt sind/ist die Ausgangsgröße und/oder die Eingangsgröße eine Prozeßgröße und/oder eine Steuergröße einer technischen Anlage.

Einer Weiterbildung wird zur Überwachung und/oder Regelung einer technischen Anlage eingesetzt.

15 Bevorzugt ist die technische Anlage eine Stahlwalzanlage.

Eine vorteilhafte Weiterbildung wird im Rahmen eines Trainings eines neuronalen Netzes eingesetzt. Von besonderem Vorteil ist es, eine Weiterbildung bei einem Online-Lernen des neuronalen Netzes einzusetzen. Damit wird die Stabilität der Abbildung des technischen Prozesses im Betrieb des technischen Prozesses weiter verbessert. Dies ist dadurch möglich, daß bei dem Online-Lernen die Beschreibung des technischen

20 Prozesses im Betrieb des technischen Prozesses an ein sich änderndes Systemverhalten angepaßt wird.

25

Ausführungsbeispiele der Erfindung sind in Figuren dargestellt und werden im weiteren näher erläutert.

30

Es zeigen:

Figur 1 Schematische Darstellung von Komponenten einer durch ein neuronales Netz bzw. ein Fuzzy-System überwachten und gesteuerten Stahlwalzanlage

Figur 2 Verfahrensschritte für ein Training des neuronalen Netzes bzw. Fuzzy-Systems

Figur 3 Darstellung eines Abbildungsverhalten eines neuronalen Netzes

In Figur 1 sind schematisch Komponenten einer unter Verwendung eines neuronalen Netzes überwachten und gesteuerten Stahlwalzanlage zu einer Stahlverarbeitung dargestellt.

Figur 1 zeigt ein System der Stahlwalzanlage 101 zur Stahlverarbeitung (Walzprozeß). Das System der Stahlwalzanlage 101 wird unter Verwendung eines neuronalen Netzes 102 überwacht und gesteuert. Ein Systemverhalten der Stahlwalzanlage 101 wird durch das neuronale Netzes 102, welches vom Typ Radiale-Basisfunktionen-Netz ist, beschrieben.

Anstelle des neuronalen Netzes 102 kann auch ein Fuzzy-System 110 für die Überwachung und die Steuerung der Stahlwalzanlage 101 eingesetzt werden.

Ferner sind geeignete Meßmittel 103, beispielsweise Sensoren 103, dargestellt, mit denen Prozeßgrößen, die den Prozeß der Stahlverarbeitung beeinflussen, gemessen werden. Die Meßmittel 103 sind über einem Bus 104 mit einem Speicher 105 verbunden. Zu vorgebbaren Zeitpunkten werden die Prozeßgrößen gemessen und in dem Speicher 105 gespeichert.

Die gemessenen Prozeßgrößen sind chemische Prozeßgrößen, wie beispielsweise eine Kohlenstoffkonzentration oder eine Mangankonzentration, eine Endbandtemperatur eines gewalzten Stahlbandes, eine Endbanddicke des gewalzten Stahlbandes sowie eine Walzengeschwindigkeit.

Ferner ist ein weiteres Meßmittel 106 dargestellt, mit dem eine weitere Prozeßgröße, eine Walzkraft, die eine zu überwachende Größe und eine Steuergröße des Prozesses bzw. des Systems der Stahlwalzanlage 101 ist, gemessen wird. Das weitere Meßmittel 106 ist ebenfalls über den Bus 104 mit dem Speicher 105 verbunden. Die weitere Prozeßgröße wird ebenfalls an dem System Stahlwalzanlage 101 zu den vorgebbaren Zeitpunkten gemessen und in dem Speicher 105 gespeichert.

10

Der Speicher 105 ist über einen weiteren Bus 115 mit einer Verarbeitungseinheit 107, die einen Prozessor 108 aufweist, beispielsweise ein Rechner, verbunden. Das neuronale Netz 102 bzw. das Fuzzy-System 110 ist in Form von Software in der Verarbeitungseinheit 107 gespeichert. Der Prozessor 108 führt die Software aus.

15

Die Walzkraft ist eine Ausgangsgröße des neuronalen Netzes 102 bzw. des Fuzzy-Systems 110, welches das Systemverhalten der Stahlwalzanlage 101 beschreibt. Die Walzkraft wird aus Eingangsgrößen, die an das neuronale Netz 101 bzw. an das Fuzzy-System 110 angelegt werden, wobei das neuronale Netz 101 bzw. das Fuzzy-System 110 ein vorgegbares Abbildungsverhalten aufweist, bestimmt wird.

25

Für die Bestimmung der Eingangsgrößen des neuronalen Netzes 102 bzw. des Fuzzy-Systems 110 werden aus den gemessenen Prozeßgrößen solche Prozeßgrößen ausgewählt, die den Prozeß der Stahlverarbeitung maßgeblich beeinflussen. Der Einfluß jeweils einer Prozeßgröße auf den Prozeß wird durch eine Sensitivitätsanalyse bestimmt. Es werden folgende Eingangsgrößen ausgewählt:

30

- Kohlenstoffkonzentration des verarbeiteten Stahls
- Mangankonzentration des verarbeiteten Stahls
- Endbandtemperatur eines gewalzten Stahlbandes

35

- Endbanddicke des gewalzten Stahlbandes
- Walzengeschwindigkeit einer Stahlwalze

Die Eingangsgrößen werden an das neuronale Netz 102 bzw. an
5 das Fuzzy-System 110 angelegt. Das neuronale Netz 102 bzw.
das Fuzzy-System 110 bestimmt unter Verwendung der Eingangs-
größen die Ausgangsgröße Walzkraft. Die Ausgangsgröße wird an
dem neuronalen Netz 102 bzw. an dem Fuzzy-System 110 mittels
eines Mittels 108 abgegriffen und über eine Datenleitung 109,
10 welche das neuronale Netz 102 bzw. das Fuzzy-System 110 mit
dem System der Stahlwalzanlage 101 verbindet, an das System
der Stahlwalzanlage 101 übertragen.

Unter Verwendung der Ausgangsgröße Walzkraft wird das System
15 Stahlwalzanlage 101 unter Verwendung geeigneter Mittel 111
überwacht und gesteuert, wobei durch eine geeignete Einstel-
lung der Walzkraft eine vorgegebene Endbanddicke des Stahl-
bandes erzeugt wird.

20 Die einzustellende Walzkraft wird unter Verwendung des neuro-
nalen Netzes 102 bzw. des Fuzzy-Systems 110 für eine vorgege-
bene Endbanddicke des Stahlbandes bestimmt.

Figur 2 zeigt Verfahrensschritte, die im Rahmen eines Trai-
25 nings des neuronalen Netzes, mit welchem Training ein vorgeb-
bares Abbildungsverhalten bestimmt wird, durchgeführt werden.

Die Verfahrensschritte werden entsprechend bei dem Fuzzy-
Systems durchgeführt.

30

Für das Training des neuronalen Netzes werden Trainingsdaten
derart ermittelt, daß die ausgewählten Prozeßgrößen 201 und
die Walzkraft 202 an der Stahlwalzanlage zu vorgebbaren Zeit-
punkten gemessen, jeweils für einen Zeitpunkt als ein Trai-

14

ningsdatenvektor 203 zusammengefaßt und in dem Speicher gespeichert werden 204.

Die Trainingsdatenvektoren werden an das neuronale Netz angelegt 205.

Als eine Targetgröße für das Training des neuronalen Netzes wird ein unter Verwendung der Ausgangsgröße Walzkraft bestimmter sogenannter Walzkraftkorrekturfaktor ermittelt 206.

10

Unter Verwendung der Trainingsdatenvektoren und der Targetgröße wird das neuronale Netz mit einem bekannten Trainingsverfahren, wie es in [1] beschrieben ist, trainiert 207.

15 Das Training des neuronalen Netzes wird bevorzugt im Betrieb der Stahlwalzanlage durchgeführt, wobei die Trainingsdaten im Betrieb der Stahlwalzanlage gemessen werden und damit auch Anwendungsdaten für das neuronale Netz sind. Ein derartiges Training wird als ein sogenanntes Online-Training bezeichnet.

20

Das Online-Training weist den Vorteil auf, daß eine Änderung in dem Systemverhalten der Stahlwalzanlage ohne große Zeitverzögerung in dem Abbildungsverhalten des neuronalen Netzes berücksichtigt wird. Eine solche Änderung kann beispielsweise durch eine Abnutzung von einer Komponente der Stahlwalzanlage verursacht werden.

Im folgenden wird ein Abbildungsverhalten eines neuronalen Netzes, wie es für eine Überwachung und ein Steuerung eines Systems eines technischen Systems, beispielsweise das oben beschriebene System der Stahlwalzanlage, eingesetzt wird, näher beschrieben.

Figur 3 zeigt in einer vereinfachten Darstellung eine Abbildung durch ein neuronales Netz.

35

Für die vereinfachte Darstellung wird ein eindimensionaler Eingangsraum (x) als ein Eingangsraum des neuronalen Netzes mit einer eindimensionalen Eingangsgröße x verwendet. Im Allgemeinen wird eine Dimension des Eingangsraums durch eine Anzahl von Eingangsgrößen bestimmt. Das oben beschriebene neuronale Netz zur Überwachung und Steuerung der Stahlwalzanlage weist beispielsweise einen mehrdimensionalen Eingangsraum auf. Die Ausgangsgröße des neuronalen Netzes ist eine eindimensionale Ausgangsgröße y . Damit erfolgt durch das neuronale Netz, das ein Systemverhalten eines technischen Systems beschreibt, eine Abbildung der Eingangsgröße x_i auf die Ausgangsgröße y in einem Zustandsraum (x,y) 315.

Figur 3 zeigt den Zustandsraum (x,y) 315, in dem das Systemverhalten des Prozesses durch sogenannte Arbeitspunkte 304 des Prozesses beschrieben ist.

In Figur 3 sind weitere Arbeitspunkte 305 dargestellt, die fehlerhafte Arbeitspunkte 305 sind, wie sie beispielsweise durch eine falsche Messung von Prozeßgrößen auftreten können.

Liegen die fehlerhaften Arbeitspunkte 305 in einem Bereich in dem Zustandsraum (x,y) 315, in dem eine Dichte der Arbeitspunkte 304 gering ist, so wird durch ein Training des neuronalen Netzes, welches nur erste Teilmodelle umfaßt, eine Abbildungsvorschrift bzw. eine Abbildung 306 derart falsch bestimmt, daß die Abbildungsvorschrift bzw. die Abbildung 306 für Datenpunkte in dem Bereich instabil ist. Des Weiteren beschreibt die Abbildungsvorschrift bzw. die Abbildung 306 den technischen Prozeß in dem Bereich in dem Zustandsraum (x,y) 315 nicht mit einer ausreichender Genauigkeit.

Ferner sind in Figur 3 die drei ersten Teilmodelle durch jeweils ein einem ersten Teilmodell zugeordneten Clusterzentrum

μ_k 301, 302, 303, jeweils einer einem ersten Teilmodell zugeordneten Partitionsfunktion χ_k 309, 310, 311 und jeweils ein einem ersten Teilmodell zugeordneten lineares Adaline act_k 320, 321, 322 dargestellt.

5

Das Clusterzentrum μ_k 301, 302, 303 wird durch ein Clusteringverfahren, wie es in [2] und [3] beschrieben ist, bestimmt.

10 Die Stabilität der Abbildung und damit auch die Genauigkeit der Beschreibung des technischen Prozesses werden dadurch verbessert, daß die ersten Teilmodelle mit einem zweiten Teilmodell überlagert werden. Die verbesserte Beschreibung 308 ist ebenfalls in der Figur 3 dargestellt.

15

Das zweite Teilmodell umfaßt ein globales lineares Adaline act_0 323 mit einem Zentrum $\mu_0 = 0$.

Das neuronale Netz, welches ein Radiale-Basisfunktionen-Netz vom Typ Multi Adaline, wie es in [1] beschrieben ist, ist, läßt sich unter Berücksichtigung der Überlagerung mit dem zweiten Teilmodell wie folgt darstellen:

$$y = \sum_{k=0}^{nk} \chi_k * act_k \quad (14)$$

wobei:

$$25 \quad act_k = (x - \mu_k) * a_k + b_k \quad (15)$$

$$\chi_k = \frac{g_k}{\sum_{k=0}^{nk} g_k} = \frac{g_k}{g_0 + \sum_{k=1}^{nk} g_k} = \frac{g_k}{g_0 + SG} \quad (16)$$

$$g_k = \exp\left(-\frac{1}{2 * \sigma_k^2} * (x - \mu_k)^2\right) \quad (17)$$

$$SG = \sum_{k=1}^{nk} g_k \quad (18)$$

mit:

	x, y	Eingangsgröße, Ausgangsgröße
5	k	Laufvariable
	nk	Zahl der Clusterzentren der ersten Teilmodelle
	a_k	Steigung
	b_k	Biaswert
	SG	Summe von Gewichtungen
10	g_k, g_0	Gewichtungsfunktionen
	χ_k	Partitionsfunktion
	act_k	lineares Adaline
	μ_k	Clusterzentrum
	σ_k	Varianz

15

Die Summe von den Gewichtungen ist ein Maß für eine Zugehörigkeit eines ausgewählten Datenpunkts zu den ersten Teilmodelle. Kleine Werte von SG für einen ausgewählten Datenpunkt bedeuten, daß der ausgewählte Datenpunkt durch kein Clusterzentrum repräsentiert wird, wie beispielsweise die falschen Arbeitspunkte 305. Große Werte von SG für einen ausgewählten Datenpunkt bedeuten, daß der ausgewählte Datenpunkt einen geringen Abstand in dem Zustandsraum (x, y) 315 von einem der Clusterzentren 301, 302, 303 aufweist.

25

Die Zugehörigkeitsfunktion g_0 des zweiten Teilmodell ist eine Konstante. Die Konstante wird derart gewählt, daß folgende Abschätzungen gelten:

30 Für große Werte von SG :

$\chi_k \equiv \chi_1$ (für die ersten Teilmodelle, $1 \leq l \leq 3$)

$g_0 \ll SG$

Für kleine Werte von SG:

5 $\chi_1 \equiv 0$ (für die ersten Teilmodelle, $1 \leq l \leq 3$)

$\chi_0 \equiv 1$

$g_0 \gg SG$

Die Konstante g_0 wird in Abhängigkeit einer gegebenen Verteilung von SG gewählt. Absolute Werte von SG und die Verteilung von SG hängen von einer Datenverteilung in dem Zustandsraum (x, y) 315, einer Normierung, einer Lage und der Anzahl der Clusterzentren μ_h 301, 302, 303 und einem Smoothing-Wert ab. Die Wahl der Konstante erfolgt durch eine Optimierung der
15 Verteilung von SG. Beispielsweise wird $g_0 = 0.25$ gewählt.

Die beschriebene Abbildung durch das neuronale Netz wird als ein Sourcecode in einer C/C++ Bibliothek mit einer Schnittstelle zu MATLAB in einer Verarbeitungseinheit, beispielsweise ein Rechner mit einem Prozessor, implementiert.
20

Im folgenden werden Alternativen des Ausführungsbeispiels dargestellt.

25 Eine Überlagerung der ersten und der zweiten Teilmodelle kann auch derart erfolgen, daß jeweils ein erstes Teilmodell mit einem zugeordneten zweiten Teilmodell überlagert wird. Dabei können die zweiten Teilmodelle vom Typ lineares Adaline sein. Die Überlagerung kann, wie beschrieben, unter Verwendung der
30 Summe der Gewichte erfolgen.

Ebenso kann eine Überlagerung derart ausgestaltet sein, daß die ersten Teilmodelle mit einem zweiten Teilmodell und zugleich jeweils ein erstes Teilmodell mit einem zugeordneten weiteren zweiten Teilmodell überlagert wird.

5

Das zweite Teilmodell kann auch vom Typ Polynom sein. Ferner kann das zweite Teilmodell ein analytisches Modell, welches den technischen Prozeß oder Teile des technischen Prozesses beschreibt, oder eine andere Form von Vorwissen sein.

10

Die Abbildung kann auch durch ein Fuzzy-System beschrieben werden. Die oben beschriebene Zusammenhänge gelten entsprechend für das Fuzzy-System.

15

Anstelle des im Ausführungsbeispiel beschriebenen Fuzzy-Clustering-Verfahrens kann das Clustering unter Verwendung eines K-Means-Verfahrens (KM) oder durch ein Neural-Gas-Verfahren (NG) erfolgen.

20

Ferner kann die Erfindung, die bei dem beschriebenen Ausführungsbeispiel für eine Überwachung und eine Steuerung einer Stahlwalzanlage eingesetzt wird, auch für die Überwachung und Steuerung eines beliebigen technischen Systems, wie beispielsweise eine Papierwickelvorrichtung oder eine Kläranlage

25

eingesetzt werden.

Im Rahmen dieses Dokuments wurden folgende Veröffentlichungen zitiert:

- [1] O. Nelles, LOLIMOT - Lokale, lineare Modelle zur Identifikation nichtlinearer, dynamischer Systeme, in: Automatisierungstechnik, Bd. 45 (1997), S. 163 - 174, R. Oldenbourg Verlag, 1997
- [2] J. Hollatz und T. Runkler, Datenanalyse und Regelerzeugung mit Fuzzy-Clustering, Fuzzy-Systeme in Theorie und Anwendungen, in: Hellendoorn Adamy Prehn Wegmann und Linzenkirchner, Kapitel 5.6, Siemens AG, Nürnberg, 1997
- [3] J. C. Bezdek et al, Detection and Characterization of Cluster Substructure, II. Fuzzy c Varieties and Convex Combinations thereof SIAM Journal on Applied Mathematics, Volume 40, No. 2, Seite 358-370, 1981
- [4] A. Zell, Simulation Neuronaler Netze, S. 225-240, Addison-Wesley Publishing Company, 1994

Patentansprüche

1. Verfahren zur rechnergestützten Ermittlung einer Abbildungsvorschrift, welche einen technischen Prozeß in einem vorgegebenen Raum beschreibt, wobei der technische Prozeß in dem vorgegebenen Raum durch Teilmodelle beschrieben wird, dadurch gekennzeichnet, daß
 - eine approximative Abbildungsvorschrift, welche den technischen Prozeß approximativ beschreibt und welche erste Teilmodelle umfaßt, ermittelt wird, und
 - anschließend die Abbildungsvorschrift derart ermittelt wird, daß die ersten Teilmodelle mit mindestens einem zweiten Teilmodell erweitert werden, wobei die Abbildungsvorschrift im Vergleich zu der approximativen Abbildungsvorschrift stabilisiert wird.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Abbildungsvorschrift mit mehreren zweiten Teilmodellen erweitert wird.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß das zweite Teilmodell mit mehreren ersten Teilmodellen überlagert wird.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß die Ausgangsgröße und/oder die Eingangsgröße mehrdimensionale Größen sind/ist.
5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß die Überlagerung des zweiten Teilmodells unter Verwendung einer Gewichtungsfunktion durchgeführt wird.

6. Verfahren nach einem der Ansprüche 1 bis 5,
 d a d u r c h g e k e n n z e i c h n e t, d a ß die Gewich-
 tungsfunktion unter Verwendung einer Summe von Gewichtungen
 5 ermittelt wird.

7. Verfahren nach einem der Ansprüche 1 bis 6,
 d a d u r c h g e k e n n z e i c h n e t, d a ß die Abbil-
 dungsvorschrift im Rahmen eines neuronalen Netzes verwendet
 10 wird.

8. Verfahren nach einem der Ansprüche 1 bis 7,
 d a d u r c h g e k e n n z e i c h n e t, d a ß das neuro-
 nale Netz vom Typ Multi Adaline ist.
 15

9. Verfahren nach einem der Ansprüche 1 bis 8,
 d a d u r c h g e k e n n z e i c h n e t, d a ß das zweite
 Teilmodell vom Typ Adaline ist.

20 10. Verfahren nach einem der Ansprüche 1 bis 9,
 d a d u r c h g e k e n n z e i c h n e t, d a ß die Abbil-
 dungsvorschrift wie folgt beschrieben wird:

$$y = \sum_{k=0}^{nk} \chi_k * act_k \quad (14)$$

wobei:

25 $act_k = (x - \mu_k) * a_k + b_k \quad (15)$

$$\chi_k = \frac{g_k}{\sum_{k=0}^{nk} g_k} = \frac{g_k}{g_0 + \sum_{k=1}^{nk} g_k} = \frac{g_k}{g_0 + SG} \quad (16)$$

$$g_k = \exp\left(-\frac{1}{2 * \sigma_k^2} * (x - \mu_k)^2\right) \quad (17)$$

$$SG = \sum_{k=1}^{nk} g_k \quad (18)$$

mit:

- | | | |
|----|------------|--|
| | x, y | Eingangsgröße, Ausgangsgröße |
| 5 | k | Laufvariable |
| | nk | Zahl der Clusterzentren der ersten Teilmodelle |
| | a_k | Steigung |
| | b_k | Biaswert |
| | SG | Summe von Gewichtungen |
| 10 | g_k, g_0 | Gewichtungsfunktionen |
| | χ_k | Partitionsfunktion |
| | act_k | lineares Adaline |
| | μ_k | Clusterzentrum |
| | σ_k | Varianz |
- 15
11. Verfahren nach einem der Ansprüche 1 bis 10,
 d a d u r c h g e k e n n z e i c h n e t, d a ß die Aus-
 gangsgroÙe und/oder die EingangsgroÙe eine ProzeÙgroÙe
 und/oder eine SteuergröÙe einer technischen Anlage
 20 sind/ist.
12. Verfahren nach einem der Ansprüche 1 bis 11,
 g e k e n n z e i c h n e t durch den Einsatz zur Überwachung
 und/oder Regelung einer technischen Anlage.
- 25
13. Verfahren nach Anspruch 12,
 g e k e n n z e i c h n e t durch den Einsatz zur Überwachung
 und/oder Regelung einer Stahlwalzanlage.
- 30 14. Verfahren nach einem der Ansprüche 1 bis 13,

g e k e n n z e i c h n e t durch den Einsatz im Rahmen eines Trainings und/oder eines on-line Lernens des neuronalen Netzes.

5 15. Anordnung zur rechnergestützten Ermittlung einer Abbildungsvorschrift, welche einen technischen Prozeß in einem vorgegeben Raum beschreibt, wobei der technische Prozeß in dem vorgegebenen Raum durch Teilmodelle beschrieben wird, d a d u r c h g e k e n n z e i c h n e t, d a ß
10 die Anordnung einen Prozessor aufweist, der derart eingerichtet ist, daß folgende Schritte durchführbar sind:

- eine approximative Abbildungsvorschrift, welche den technischen Prozeß approximativ beschreibt und welche erste Teilmodelle umfaßt, ist ermittelbar, und
- 15 - anschließend ist die Abbildungsvorschrift derart ermittelbar, daß die ersten Teilmodelle mit mindestens einem zweiten Teilmodell erweitert werden, wobei die Abbildungsvorschrift im Vergleich zu der approximativen Abbildungsvorschrift stabilisiert wird.

20 16. Anordnung nach Anspruch 15,
d a d u r c h g e k e n n z e i c h n e t, d a ß die ersten Teilmodelle mit mehreren zweiten Teilmodellen erweiterbar sind.

25 17. Anordnung nach Anspruch 15 oder 16,
d a d u r c h g e k e n n z e i c h n e t, d a ß das zweite Teilmodell mit mehreren ersten Teilmodellen überlagerbar ist.

30 18. Anordnung nach einem der Ansprüche 15 bis 17,
d a d u r c h g e k e n n z e i c h n e t, d a ß die Ausgangsgröße und/oder die Eingangsgröße mehrdimensionale Größen sind/ist.

25

19. Anordnung nach einem der Ansprüche 15 bis 18,
dadurch gekennzeichnet, daß die Abbildungsvorschrift ein neuronales Netz beschreibt.

5 20. Anordnung nach einem der Ansprüche 15 bis 19,
dadurch gekennzeichnet, daß das neuronale Netz vom Typ Multi Adaline ist.

21. Anordnung nach einem der Ansprüche 15 bis 20,
10 dadurch gekennzeichnet, daß das zweite Teilmodell vom Typ Adaline ist.

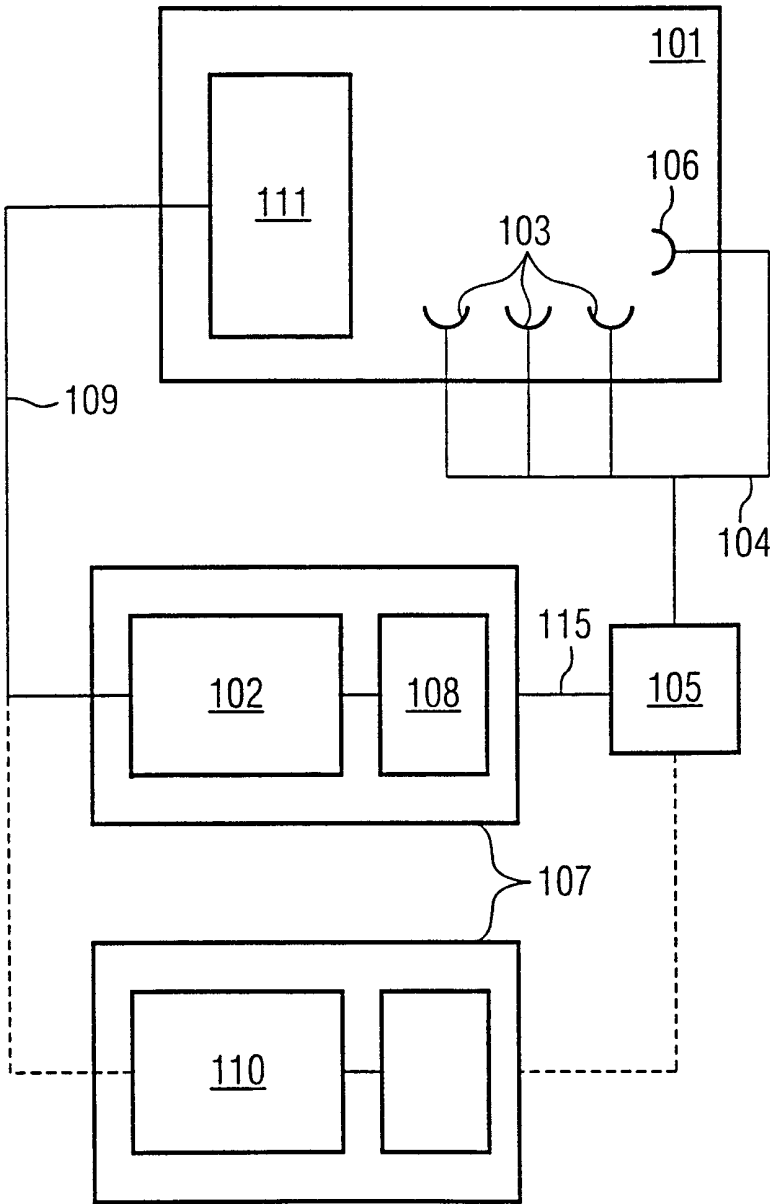
22. Anordnung nach einem der Ansprüche 15 bis 21,
dadurch gekennzeichnet, daß die Ausgangsgröße und/oder die Eingangsgröße eine Prozeßgröße
15 und/oder eine Steuergröße einer technischen Anlage sind/ist.

23. Anordnung nach einem der Ansprüche 15 bis 22,
20 gekennzeichnet durch den Einsatz zur Überwachung und/oder Regelung einer technischen Anlage.

24. Anordnung nach Anspruch 23,
gekennzeichnet durch den Einsatz zur Überwachung
25 und/oder Regelung einer Stahlwalzanlage.

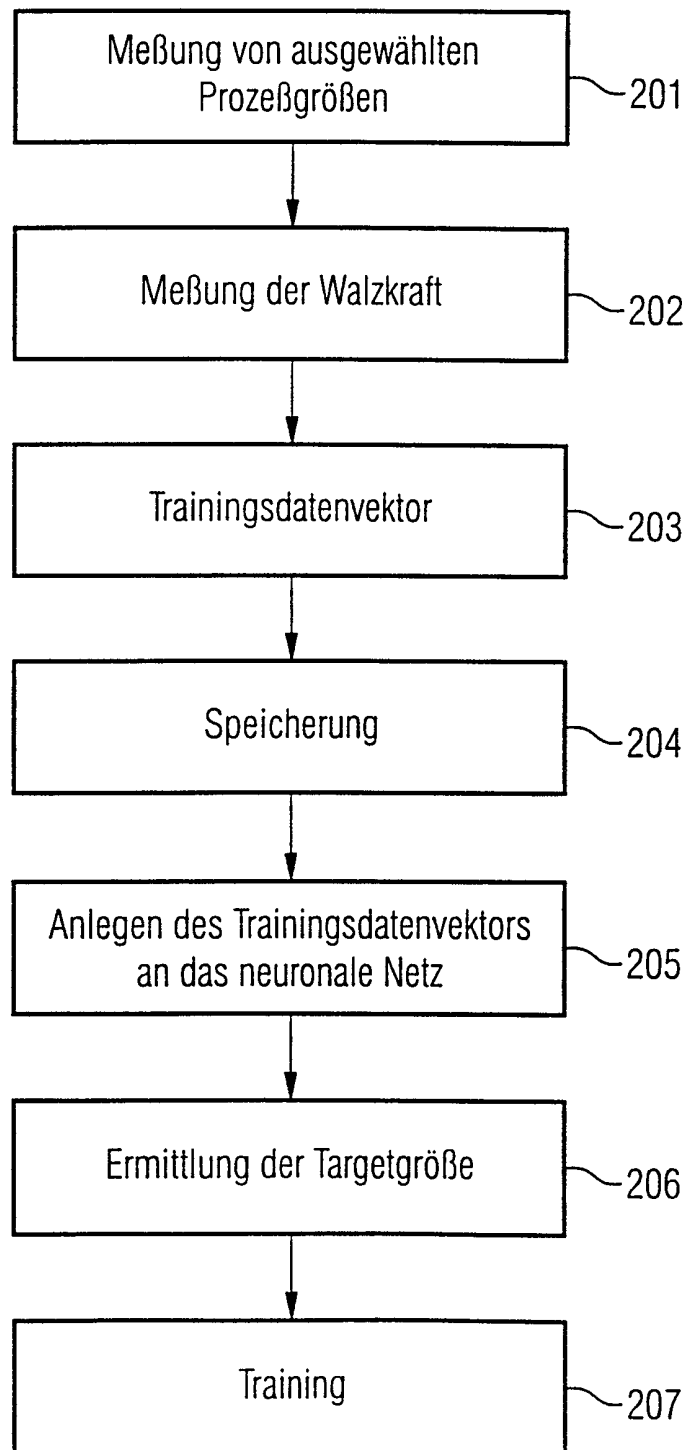
25. Anordnung nach einem der Ansprüche 15 bis 24,
gekennzeichnet durch den Einsatz im Rahmen eines Trainings und/oder eines on-line Lernens des neuronalen Net-
30 zes.

FIG 1



2/3

FIG 2



3/3

FIG 3

