



**PCT** WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro  
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<b>(51) Internationale Patentklassifikation<sup>6</sup> :</b> <b>G06F 15/80</b>	<b>A1</b>	<b>(11) Internationale Veröffentlichungsnummer:</b> <b>WO 98/10352</b>
		<b>(43) Internationales Veröffentlichungsdatum:</b> 12. März 1998 (12.03.98)

<p><b>(21) Internationales Aktenzeichen:</b> PCT/DE97/01497</p> <p><b>(22) Internationales Anmeldedatum:</b> 15. Juli 1997 (15.07.97)</p> <p><b>(30) Prioritätsdaten:</b> 196 35 758.6      3. September 1996 (03.09.96)      DE</p> <p><b>(71) Anmelder (für alle Bestimmungsstaaten ausser US):</b> SIEMENS AKTIENGESELLSCHAFT [DE/DE]; Wittelsbacherplatz 2, D-80333 München (DE).</p> <p><b>(72) Erfinder; und</b> <b>(75) Erfinder/Anmelder (nur für US):</b> NEUNEIER, Ralf [DE/DE]; Utzschneiderstrasse 10, D-80469 München (DE). ZIMMERMANN, Hans-Georg [DE/DE]; Schiffbauerweg 6 A, D-82319 Starnberg (DE).</p>	<p><b>(81) Bestimmungsstaaten:</b> JP, US, europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p><b>Veröffentlicht</b> <i>Mit internationalem Recherchenbericht.</i></p>
--	---

**(54) Title:** METHOD AND DEVICE FOR COMPUTER ASSISTED GENERATION OF AT LEAST ONE ARTIFICIAL TRAINING DATA VECTOR FOR A NEURONAL NETWORK

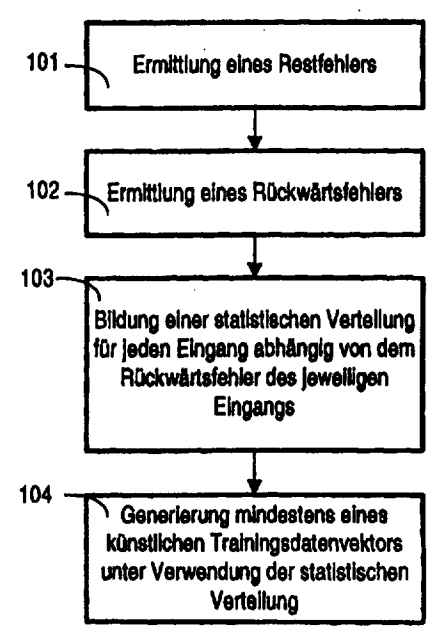
**(54) Bezeichnung:** VERFAHREN UND VORRICHTUNG ZUR RECHNERGESTÜTZTEN GENERIERUNG MINDESTENS EINES KÜNSTLICHEN TRAININGSDATENVEKTORS FÜR EIN NEURONALES NETZ

**(57) Abstract**

A residual error (101) is determined after neuronal network training has occurred. A backward error (102) is determined from the residual error. Artificial training data vectors are generated from a statistic random process with an underlying statistic distribution, wherein the respective backward error is considered for a neuronal network input (103, 104).

**(57) Zusammenfassung**

Nach erfolgtem Training eines neuronalen Netzes wird ein Restfehler ermittelt (101). Aus dem Restfehler wird ein Rückwärtsfehler ermittelt (102). Künstliche Trainingsdatenvektoren werden aus einem statistischen Zufallsprozeß generiert, dem eine statistische Verteilung zugrundeliegt, bei der der jeweilige Rückwärtsfehler für einen Eingang des neuronalen Netzes berücksichtigt wird (103, 104).



- 101 ... DETERMINATION OF AN UNDETECTED ERROR
- 102 ... DETERMINATION OF A BACKWARD ERROR
- 103 ... FORMATION OF A STATISTIC DISTRIBUTION FOR EACH INPUT DEPENDENT ON THE BACKWARD ERROR OF THE RESPECTIVE INPUT
- 104 ... GENERATION OF AT LEAST ONE ARTIFICIAL TRAINING DATA VECTOR USING STATISTIC DISTRIBUTION

**LEDIGLICH ZUR INFORMATION**

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

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

Beschreibung

- 5 Verfahren und Vorrichtung zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein neuronales Netz

### 1. Technischer Hintergrund

10

Neuronale Netze lernen mit Hilfe von Trainingsdaten. Die Trainingsdaten sind in vielen Anwendungsgebieten sehr verrauscht, beispielsweise bei der Modellierung von Finanzdaten wie Aktien- oder Währungskursen. Somit erhalten  
15 die Trainingsdaten zufällige Störungen, die nichts mit der eigentlich zu modellierenden Systemdynamik zu tun haben.

Durch die Approximationsfähigkeit der neuronalen Netze kann aber auch die transiente Struktur des zufälligen Rauschens  
20 gelernt werden. Dieses Phänomen wird als Übertrainieren des neuronalen Netzes bezeichnet. Durch ein übertrainiertes neuronales Netz wird in sehr verrauschten Systemen der Lernvorgang des neuronalen Netzes erheblich behindert, da die Verallgemeinerungsfähigkeit des neuronalen Netzes negativ  
25 beeinflußt wird.

Dieses Problem gewinnt in Anwendungsgebieten an Bedeutung, in denen nur eine geringe Anzahl von Trainingsdatenvektoren zur Adaption des neuronalen Netzes an die Anwendung, d.h. die zu  
30 modellierende, durch die Trainingsdatenvektoren repräsentierte Funktion, verfügbar ist.

Vor allen in diesen Anwendungsgebieten, aber auch allgemein in einem Trainingsverfahren eines neuronalen Netzes ist es  
35 vorteilhaft, zusätzliche Trainingsdatenvektoren künstlich zu

generieren, um somit eine größere Trainingsdatenmenge zu erhalten.

## 2. Stand der Technik

5 Es ist bekannt, die Generierung der künstlichen Trainingsdatenvektoren durch Verrauschen der verfügbaren Trainingsdatenvektoren der Trainingsdatenmenge durchzuführen. Es ist in diesem Zusammenhang aus dem Dokument [1] bekannt, die Trainingsdatenmenge mit gaußförmigem Rauschen mit dem Mittelwert  
10 0 und einer Varianz  $\sigma$ , die für alle Eingänge des neuronalen Netzes auf den gleichen Wert gesetzt wird, zu ermitteln.

Aus [4] ist es bekannt, Trainingsdaten durch Einführung zusätzlichen Rauschens zu generieren. Dabei ist es bekannt, die  
15 sogenannte Jackknife-Prozedur einzusetzen.

Dieses Verfahren weist jedoch einige Nachteile auf.

Dadurch, daß zur Generierung der zusätzlichen Trainingsdatenvektoren als statistische Verteilung, die zur Generierung  
20 verwendet wird, ein gaußförmiges Rauschen mit einer Varianz verwendet wird, die für alle Eingänge des neuronalen Netzes auf den gleichen Wert gesetzt wird, werden Trainingsdatenvektoren neu generiert, die keinerlei Aussage über das zu modellierende System enthalten. Die Trainingsdatenvektoren enthalten  
25 ferner keinerlei Information über das tatsächliche, dem System zugrundeliegende Rauschen. Somit wird zwar die Trainingsdatenmenge vergrößert, diese muß aber nicht den Lernvorgang unterstützen, da ein fest vorgegebenes Rauschen, das mit  
30 der eigentlichen Systemdynamik nichts zu tun hat, zum Trainieren des neuronalen Netzes verwendet wird. Damit kann es dann trotzdem zum Übertrainieren kommen.

Grundlagen über neuronale Netze sind beispielsweise aus dem  
35 Dokument [2] bekannt.

Grundlagen über Verwendung der neuronalen Netze in der Ökonomie sind beispielsweise aus dem Dokument [3] bekannt.

### 3. Kurzbeschreibung der Erfindung

5

Somit liegt der Erfindung das Problem zugrunde, künstlich neue Trainingsdatenvektoren für ein neuronales Netz zu bilden, wobei jedoch ein Übertrainieren des neuronalen Netzes vermieden wird.

10

Das Problem wird durch das Verfahren gemäß Patentanspruch 1 sowie durch die Vorrichtung gemäß Patentanspruch 9 gelöst.

15

Bei dem Verfahren wird nach dem Training des neuronalen Netzes mit verfügbaren Trainingsdatenvektoren einer Trainingsdatenmenge ein Restfehler ermittelt. Aus dem Restfehler wird beispielsweise unter Verwendung eines Gradientenabstiegs-Verfahrens ein eingangsbezogener Rückwärtsfehler ermittelt. Die Ermittlung des Rückwärtsfehlers entspricht der üblichen Vorgehensweise während des Trainings eines neuronalen Netzes zur Adaption der einzelnen Gewichte des neuronalen Netzes. Ist der eingangsbezogene Rückwärtsfehler ermittelt worden, wird unter Berücksichtigung des jeweiligen Rückwärtsfehlers eine dem jeweiligen Eingang zugeordnete statistische Verteilung generiert, und der künstliche Trainingsdatenvektor wird unter Berücksichtigung der jeweiligen statistischen Verteilung an den Eingängen des neuronalen Netzes erzeugt.

20

25

30

Mit diesem Verfahren ist es möglich, zusätzliche Trainingsdatenvektoren zu generieren, die eine Information über das neuronale Netz und die Struktur des neuronalen Netzes aktuell nach dem Training des neuronalen Netzes mit den verfügbaren Trainingsdatenvektoren enthalten.

35

Dies führt dazu, daß die künstlich erzeugten Trainingsdatenvektoren von dem Rückwärtsfehler, der nach dem Training des neuronalen Netzes noch existiert, abhängig sind und somit von

der Performanz des neuronalen Netzes abhängen. Dies führt dazu, daß die erzeugten Trainingsdatenvektoren Information über das zu modellierende System enthalten. Auf diese Weise kann ein Übertrainieren des neuronalen Netzes durch die zusätzlichen Trainingsdatenvektoren vermieden werden.

Die Vorrichtung gemäß Patentanspruch 9 weist eine Recheneinheit auf, die derart eingerichtet ist, daß die oben beschriebenen Verfahrensschritte durchgeführt werden.

10

Vorteilhafte Weiterbildungen der Erfindung ergeben sich aus den abhängigen Ansprüchen.

15

Es ist vorteilhaft, bei einer Online-Approximation des neuronalen Netzes, die auch als Online-Training bezeichnet wird, auch die jeweilige statistische Verteilung dem veränderten Trainingsdatensatz anzupassen. Dadurch wird das zu modellierende System noch genauer durch das neuronale Netz modelliert.

20

#### 4. Kurze Figurenbeschreibung

In den Figuren ist ein Ausführungsbeispiel der Erfindung dargestellt, welches im weiteren näher erläutert wird.

25

Es zeigen

Fig. 1 ein Ablaufdiagramm, in dem die einzelnen Verfahrensschritte des Verfahrens dargestellt sind;

30

Fig. 2 eine Skizze, in der eine Rechneranordnung dargestellt ist, mit der das Verfahren durchgeführt werden kann.

35

#### 5. Figurenbeschreibung

In Fig. 1 sind die einzelnen Verfahrensschritte des Verfahrens dargestellt.

In einem ersten Schritt (101) wird nach abgeschlossenem Training des neuronalen Netzes NN mit einer beliebigen Anzahl von Trainingsdatenvektoren TDVi einer Trainingsdatenmenge TDM von einem Rechner R ein Restfehler RE ermittelt.

Der Restfehler RE wird beispielsweise bei m Trainingsdatenvektoren TDV mit den Trainingsdatenvektoren TDV zugeordneten Sollwerten  $t_i$ , wobei mit einem Index i jeweils der Trainingsdatenvektor TDVi eindeutig gekennzeichnet wird, auf folgende Weise ermittelt:

$$15 \quad RE = \sum_{i=1}^m (t_i - y_i)^2 \quad (1).$$

Es kann jedoch selbstverständlich jede beliebige andere Vorschrift zur Bildung des Restfehlers RE nach dem Training des neuronalen Netzes NN verwendet werden.

20

Nach Ermittlung des Restfehlers wird ein Rückwärtsfehler RFj ermittelt (102). Die Bestimmung des Rückwärtsfehlers RF kann beispielsweise unter Verwendung eines Gradientenabstiegs-Verfahrens nach den Eingangssignalen der neuronalen Netze NN erfolgen.

25

Im weiteren wird folgende Nomenklatur für den Trainingsdatenvektor TDVi, sowie die Eingangsdaten des neuronalen Netzes NN, die sich aus den Trainingsdatenvektoren TDVi und einem Rauschtermvektor  $\underline{e}$  ergeben, verwendet:

30

Trainingsdatenvektor TDVi:  $\underline{x} = (x_1, \dots, x_j, \dots, x_n)$

Rauschtermvektor:  $\underline{e} = (e_1, \dots, e_j, \dots, e_n)$

35

Eingangsdatenvektor:  $\underline{z} = (x_1+e_1, \dots, x_j+e_j, \dots, x_n+e_n)$

Mit  $n$  wird eine Anzahl von Komponenten der Trainingsdatenvektoren  $TDV_i$ , der Rauschtermvektoren  $\underline{e}$  bzw. der Eingangsdatenvektoren  $\underline{z}$  bezeichnet.

Der Rückwärtsfehler  $RF_j$  wird für jeden Eingang  $j$  des neuronalen Netzes  $NN$  individuell ermittelt, wobei mit einem Index  $j$  jeweils eine Komponente des Eingangsdatenvektors  $\underline{z}$  bzw. ein Eingang des neuronalen Netzes  $NN$  eindeutig bezeichnet wird, .

Dies erfolgt beispielsweise nach dem bekannten Gradientenabstiegs-Verfahren zum Trainieren des neuronalen Netzes  $NN$ . Somit ergibt sich der Rückwärtsfehler  $RF$  des Eingangsdatenvektors  $\underline{z}$  aus den jeweiligen partiellen Ableitungen des Restfehlers  $RE$  nach den einzelnen Eingangssignalen  $z_j$ .

$$RF_j = \frac{\partial E^t}{\partial z_j^t} \quad (2).$$

Mit dem Symbol  $t$  wird jeweils ein Zeitpunkt eindeutig gekennzeichnet, zu dem die Ermittlung des Rückwärtsfehlers  $RF_j$  erfolgt.

Unter Berücksichtigung des Rückwärtsfehlers  $RF_j$  wird nun in einem dritten Schritt (103) eine statistische Verteilung  $S_j$  für jeden Eingang  $j$  des neuronalen Netzes  $NN$  gebildet.

Als statistische Verteilung  $S_j$  kann allgemein jede beliebige statistische Verteilung zur Beschreibung eines Zufallsprozesses verwendet werden.

Im folgenden werden zur einfacheren Darstellung jedoch nur eine uniforme statistische Verteilung und eine gaußförmige statistische Verteilung  $S_j$  näher erläutert.

Für die uniforme Verteilung ergibt sich eine uniforme Verteilungsbreite  $s_j^t$  beispielsweise nach folgender Vorschrift:

$$s_j^t = \frac{1}{t} \sum_{k=1}^t \left| \frac{\partial E^k}{\partial z_j^k} \right| \quad (3).$$

5

Der Rauschterm  $e_j^t$ , der zur Bildung mindestens eines künstlichen Trainingsdatenvektors KTDV verwendet wird, liegt im Bereich des Intervalls:

$$e_j^t \in [-s_j^t, s_j^t] \quad (4).$$

10

Bei einer gaußförmigen Verteilung ergibt sich die Rauschbreite  $s_j^t$  nach der folgenden Vorschrift:

$$s_j^t = \frac{1}{t} \sum_{k=1}^t \left( \frac{\partial E^k}{\partial z_j^k} \right)^2 \quad (5).$$

15

Der Rauschterm  $e_j^t$  ergibt sich für diesen Beispielsfall nach folgender Vorschrift:

$$e_j^t \in N\left(0, \sqrt{s_j^t}\right) \quad (6).$$

20

Mit  $N\left(0, \sqrt{s_j^t}\right)$  wird ein normal verteilter Zufallsprozeß mit einem Mittelwert 0 und der Standardabweichung  $\sqrt{s_j^t}$  bezeichnet.

25

Unter Verwendung der jeweiligen statistischen Verteilung  $S_j$  wird der mindestens eine künstliche Trainingsdatenvektor KTDV generiert (104).

Anschaulich bedeutet dies, daß der künstliche Trainingsdatenvektor KTDV durch den jeweils durch die statistische Verteilung  $S_j$  beschriebenen Zufallsprozeß generiert wird.

- 5 Durch diese Vorgehensweise haben die künstlichen Trainingsdatenvektoren KTDV statistisch die gleichen Eigenschaften wie die original verfügbaren Trainingsdatenvektoren TDVi.

Dies kann durch Verrauschen mit einem einheitlichen  
10 Rauschprozeß, wie dieser aus dem Dokument [1] bekannt ist, nicht erreicht werden.

Anschaulich läßt sich das Verfahren folgendermaßen beschreiben. Die Eingangsdaten, repräsentiert durch die original verfügbaren Trainingsdatenvektoren TDVi, werden mit einem Zufallsprozeß verrauscht, der für jeden Eingang abhängig ist  
15 von der jeweiligen Gradienten-Fehlerfunktion nach den Eingangssignalen.

- 20 Die jeweilige Rauschbreite  $s_j^t$  kann anschaulich folgendermaßen interpretiert werden:

- Kleine Werte von  $s_j^t$  werden durch Eingänge erzeugt, die eine gute Abbildung auf die Trainingsdaten erzeugen.

25

- Kleine Werte der Rauschbreite  $s_j^t$  können aber auch auf unwichtige, für das Training des neuronalen Netzes NN nicht relevante Eingänge hinweisen.

- 30 In beiden Fällen ist es im Sinne der Aufgabe, daß kein oder nur wenig Rauschen zu den ursprünglich vorhandenen Trainingsdatenvektoren TDVi hinzugefügt wird.

- Große Werte der Rauschbreite  $s_j^t$  deuten darauf hin, daß der  
35 Eingang j erhebliche Bedeutung aufweist, aber das neuronale

Netz NN bisher nicht in der Lage ist, eine adäquate Abbildung zu lernen.

- Rauschen auf Eingängen mit einer großen Rauschbreite  $s_j^t$   
5 "versteift" das neuronale Netz NN, wodurch ein Übertrainieren verhindert wird, was zu einer besseren Verallgemeinerungsfähigkeit des neuronalen Netzes NN führt.

10 Dieses Verfahren kann besonders vorteilhaft in Anwendungssituationen verwendet werden, in denen nur eine relativ geringe Anzahl von Trainingsdatenvektoren  $TDV_i$  zum Trainieren des neuronalen Netzes NN verfügbar ist.

15 Durch das Verfahren kann die Trainingsdatenmenge TDM künstlich erheblich erweitert werden zu einer erweiterten Trainingsdatenmenge KTDM, ohne die eigentliche Systemdynamik zu verfälschen, da die statistischen Eigenschaften der Trainingsdatenmenge TDM auch in den künstlichen Trainingsdatenvektoren KTDV enthalten sind.

20

Eine typische Anwendungssituation des Verfahrens liegt in der Analyse von Finanzmärkten, beispielsweise von Aktienmärkten oder auch Wertpapiermärkten.

- 25 Wird zum Trainieren des neuronalen Netzes NN ein sog. Online-Lernverfahren eingesetzt, bei dem während der Generalisierungsphase des neuronalen Netzes NN weitere Trainingsdatenvektoren TDV ermittelt werden, mit denen das neuronale Netz NN in der Generalisierungsphase weiter trainiert wird, so ist  
30 es in einer Weiterbildung des Verfahrens vorgesehen, abhängig von der Änderung der Trainingsdatenmenge TDM, die durch die Trainingsdatenvektoren  $TDV_i$  gebildet wird, auch die statistischen Verteilungen  $S_j$  anzupassen. Dies kann beispielsweise  
35 für eine uniforme Verteilung nach folgender Vorschrift erfolgen:

$$s_j^{t+1} = (1 - \alpha) \cdot s_j^t + \alpha \cdot \left| \frac{\partial E^k}{\partial z_j^k} \right| \quad (7).$$

Für eine gaußförmige statistische Verteilung  $S_j$  kann beispielsweise folgende Adaptionvorschrift eingesetzt werden:

5

$$\left( s_j^{t+1} \right)^2 = (1 - \alpha) \cdot \left( s_j^t \right)^2 + \alpha \cdot \left( \frac{\partial E^k}{\partial z_j^k} \right)^2 \quad (8).$$

Mit  $\alpha$  wird ein Adaptionfaktor bezeichnet.

- 10 Es hat sich als vorteilhaft herausgestellt, als Wert für den Adaptionfaktor  $\alpha$  den Kehrwert der Anzahl der Trainingsdatenvektoren  $TDV_i$  zu verwenden.

Im Rahmen dieses Dokumentes wurden folgende Veröffentlichungen zitiert:

- 5 [1] C. M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, U. K., ISBN 0-19-853864-2, S. 346 - 349, 1994
- 10 [2] G. Deco und D. Obradovic, An Information-Theoretic Approach to Neural Computing, Springer Verlag, New York, USA, ISBN 0-387-94666-7, S. 23 - 37, 1996
- 15 [3] H. Zimmermann und Rehkugler, Neuronale Netze in der Ökonomie, Kapitel 1, Neuronale Netze als Entscheidungskalkül, Vahlen Verlag, München, ISBN 3-8006-1871-0, S. 3 - 20, 1994
- [4] L. Holmström und P. Koistinen, Using Additive Noise in Back-Propagation Training, IEEE Transactions on Neural Networks, Vol. 3, No. 1, S. 24 - 38, Januar 1992

## Patentansprüche

1. Verfahren zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein Neuronales Netz (NN),  
5 (NN),
  - bei dem ein Restfehler (RE) nach einem Training des Neuronalen Netzes (NN) mit Trainingsdatenvektoren (TDVi) einer Trainingsdatenmenge (TDM) ermittelt wird (101),
  - bei dem aus dem Restfehler (RE) für mindestens einen Eingang (j) des Neuronalen Netzes (NN) ein Rückwärtsfehler (RFj)  
10 ermittelt wird (102),
  - bei dem jeweils eine dem Eingang (j) zugeordnete statistische Verteilung ( $S_j$ ) generiert wird, die von der Größe des jeweiligen Rückwärtsfehlers (RFj) abhängig ist (103),
  - 15 - bei dem der künstliche Trainingsdatenvektor generiert wird unter jeweiliger Verwendung der einem Eingang (j) zugeordneten statistischen Verteilung ( $S_j$ ) (104).
  
2. Verfahren nach Anspruch 1,  
20 bei dem der Rückwärtsfehler (RFj) mit einem Gradientenabstiegsverfahren ermittelt wird.
  
3. Verfahren nach Anspruch 2,  
bei dem das Gradientenabstiegsverfahren mit einem Backpropagation-Verfahren realisiert wird.  
25
  
4. Verfahren nach einem der Ansprüche 1 bis 3,  
bei dem die statistische Verteilung ( $S_j$ ) durch eine uniforme Verteilung gebildet wird, wobei eine Verteilungsbreite der  
30 uniformen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.
  
5. Verfahren nach einem der Ansprüche 1 bis 3,  
bei dem die statistische Verteilung ( $S_j$ ) durch eine gaußförmige Verteilung gebildet wird, wobei eine Verteilungsbreite  
35 der gaußförmigen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

6. Verfahren nach Anspruch 4 oder 5,  
bei dem bei einer Online-Approximation des Neuronalen Netzes  
(NN) nach einer beliebigen Anzahl neuer Trainingsdatenvektoren  
(TDVi) die statistische Verteilung ( $S_j$ ) an einen neu er-  
mittelten Rückwärtsfehler angepaßt wird.
7. Verfahren nach einem der Ansprüche 1 bis 6 eingesetzt zur  
Modellierung einer nichtlinearen Funktion, die mit einer ge-  
ringen Anzahl von Trainingsdatenvektoren beschrieben wird.
8. Verfahren nach einem der Ansprüche 1 bis 6 eingesetzt in  
der Analyse eines Finanzmarktes.
9. Vorrichtung zur rechnergestützten Generierung mindestens  
eines künstlichen Trainingsdatenvektors für ein Neuronales  
Netz (NN),  
mit einer Recheneinheit, die derart eingerichtet ist, daß  
- ein Restfehler (RE) nach einem Training des Neuronalen Net-  
zes (NN) mit Trainingsdatenvektoren (TDVi) einer Trainingsda-  
tenmenge (TDM) ermittelt wird (101),  
- aus dem Restfehler (RE) für mindestens einen Eingang (j)  
des Neuronalen Netzes (NN) ein Rückwärtsfehler (RFj) ermit-  
telt wird (102),  
- jeweils eine dem Eingang (j) zugeordnete statistische Ver-  
teilung ( $S_j$ ) generiert wird, die von der Größe des jeweiligen  
Rückwärtsfehlers (RFj) abhängig ist (103),  
- der künstliche Trainingsdatenvektor generiert wird unter  
jeweiliger Verwendung der einem Eingang (j) zugeordneten sta-  
tistischen Verteilung ( $S_j$ ) (104).
10. Vorrichtung nach Anspruch 9,  
bei der die Recheneinheit derart eingerichtet ist, daß der  
Rückwärtsfehler (RFj) mit einem Gradientenabstiegsverfahren  
ermittelt wird.
11. Vorrichtung nach Anspruch 9 oder 10,

bei der die Recheneinheit derart eingerichtet ist, daß die statistische Verteilung ( $S_j$ ) durch eine uniforme Verteilung gebildet wird, wobei eine Verteilungsbreite der uniformen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

5

12. Vorrichtung nach einem der Ansprüche 9 bis 11, bei der die Recheneinheit derart eingerichtet ist, daß die statistische Verteilung ( $S_j$ ) durch eine gaußförmige Verteilung gebildet wird, wobei eine Verteilungsbreite der gaußförmigen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

10

13. Vorrichtung nach Anspruch 11 oder 12, bei der die Recheneinheit derart eingerichtet ist, daß bei einer Online-Approximation des Neuronalen Netzes (NN) nach einer beliebigen Anzahl neuer Trainingsdatenvektoren (TDVi) die statistische Verteilung ( $S_j$ ) an einen neu ermittelten Rückwärtsfehler angepaßt wird.

15

14. Vorrichtung nach einem der Ansprüche 9 bis 13 eingesetzt zur Modellierung einer nichtlinearen Funktion, die mit einer geringen Anzahl von Trainingsdatenvektoren beschrieben wird.

20

15. Vorrichtung nach einem der Ansprüche 9 bis 13 eingesetzt in der Analyse eines Finanzmarktes.

1/2

FIG 1

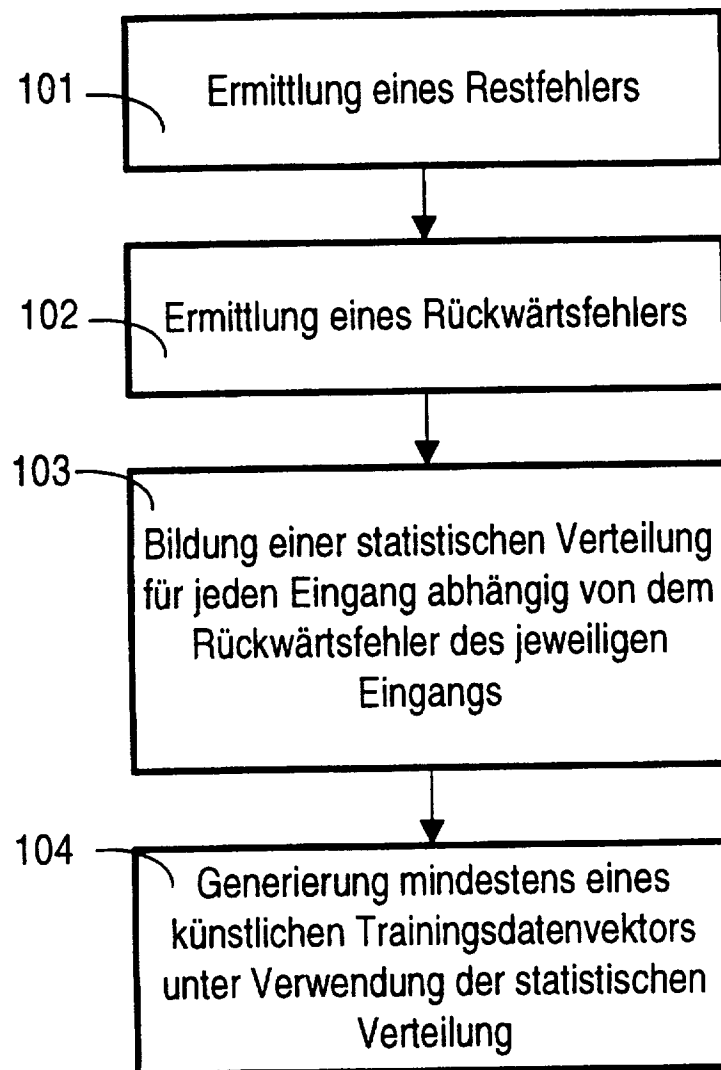
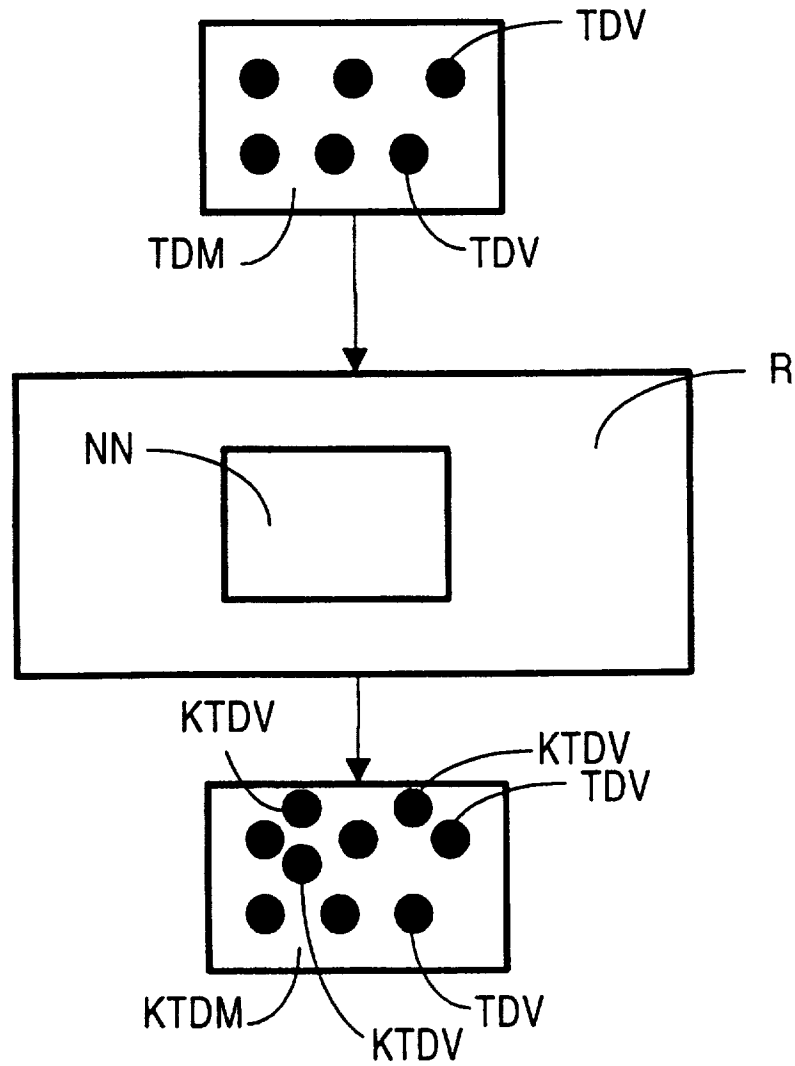


FIG 2



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/DE 97/01497

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 6 G06F15/80		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) IPC 6 G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 11486 A (MILES INC) 27 April 1995 see page 10, line 1 - line 32; figure 3 ---	1,9
A	US 5 359 699 A (TONG DAVID W ET AL) 25 October 1994 see column 4, line 5 - line 42 --- -/--	1,9
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.		
<input checked="" type="checkbox"/> Patent family members are listed in annex.		
° Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier document but published on or after the international filing date	"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search  <p style="text-align: center;">3 December 1997</p>	Date of mailing of the international search report  <p style="text-align: center;">16/12/1997</p>	
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer  <p style="text-align: center;">Schenkels, P</p>	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/DE 97/01497

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SUNGZOOM CHO ET AL: "Evolution of neural network training set through addition of virtual samples"</p> <p>PROCEEDINGS OF 1996 IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION (ICEC'96) (CAT. NO.96TH8114), PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, NAGOYA, JAPAN, 20-22 MAY 1996, ISBN 0-7803-2902-3, 1996, NEW YORK, NY, USA, IEEE, USA, pages 685-688, XP002048950</p> <p>see page 685, left-hand column, line 1 - page 686, right-hand column, line 30; figures 1,2</p> <p style="text-align: center;">---</p>	1,9
A	<p>DATUM M S ET AL: "An artificial neural network for sound localization using binaural cues"</p> <p>JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, JULY 1996, ACOUST. SOC. AMERICA THROUGH AIP, USA, vol. 100, no. 1, ISSN 0001-4966, pages 372-383, XP002048951</p> <p>see abstract</p> <p>see page 372, left-hand column, line 1 - right-hand column, line 31</p> <p>see page 376, left-hand column, line 26 - page 377, left-hand column, line 29; figures 1,2</p> <p style="text-align: center;">---</p>	1-3,9,10
A	<p>SLICHER A W R ET AL: "An innovative approach to training neural networks for strategic management of construction firms"</p> <p>DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, CAMBRIDGE, UK, 28-30 AUG. 1995, ISBN 0-948749-36-9, 1995, EDINBURGH, UK, CIVIL-COMP PRESS, UK, pages 87-93, XP002048952</p> <p>see abstract</p> <p>see page 89, left-hand column, line 5 - page 92, left-hand column, line 8; figure 1</p> <p style="text-align: center;">-----</p>	1,9

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/DE 97/01497

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9511486 A	27-04-95	US 5444796 A	22-08-95
		AU 7981794 A	08-05-95
		CA 2174522 A	27-04-95
		EP 0724750 A	07-08-96
		US 5590218 A	31-12-96
-----			
US 5359699 A	25-10-94	NONE	
-----			

# INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/DE 97/01497

**A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES**  
IPK 6 G06F15/80

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

**B. RECHERCHIERTE GEBIETE**

Recherchiertes Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)  
IPK 6 G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

**C. ALS WESENTLICH ANGESEHENE UNTERLAGEN**

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	WO 95 11486 A (MILES INC) 27. April 1995 siehe Seite 10, Zeile 1 - Zeile 32; Abbildung 3 ---	1,9
A	US 5 359 699 A (TONG DAVID W ET AL) 25. Oktober 1994 siehe Spalte 4, Zeile 5 - Zeile 42 --- -/--	1,9

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

Siehe Anhang Patentfamilie

\* Besondere Kategorien von angegebenen Veröffentlichungen :

- "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
- "E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
- "L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
- "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
- "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfindnerischer Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfindnerischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann nahelegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

3. Dezember 1997

Absenddatum des internationalen Recherchenberichts

16/12/1997

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Schenkels, P

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	<p>SUNGZOOK CHO ET AL: "Evolution of neural network training set through addition of virtual samples"</p> <p>PROCEEDINGS OF 1996 IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION (ICEC'96) (CAT. NO.96TH8114), PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, NAGOYA, JAPAN, 20-22 MAY 1996, ISBN 0-7803-2902-3, 1996, NEW YORK, NY, USA, IEEE, USA, Seiten 685-688, XP002048950 siehe Seite 685, linke Spalte, Zeile 1 - Seite 686, rechte Spalte, Zeile 30; Abbildungen 1,2</p> <p style="text-align: center;">---</p>	1,9
A	<p>DATUM M S ET AL: "An artificial neural network for sound localization using binaural cues"</p> <p>JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, JULY 1996, ACOUST. SOC. AMERICA THROUGH AIP, USA, Bd. 100, Nr. 1, ISSN 0001-4966, Seiten 372-383, XP002048951 siehe Zusammenfassung siehe Seite 372, linke Spalte, Zeile 1 - rechte Spalte, Zeile 31 siehe Seite 376, linke Spalte, Zeile 26 - Seite 377, linke Spalte, Zeile 29; Abbildungen 1,2</p> <p style="text-align: center;">---</p>	1-3,9,10
A	<p>SLICHER A W R ET AL: "An innovative approach to training neural networks for strategic management of construction firms"</p> <p>DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, CAMBRIDGE, UK, 28-30 AUG. 1995, ISBN 0-948749-36-9, 1995, EDINBURGH, UK, CIVIL-COMP PRESS, UK, Seiten 87-93, XP002048952 siehe Zusammenfassung siehe Seite 89, linke Spalte, Zeile 5 - Seite 92, linke Spalte, Zeile 8; Abbildung 1</p> <p style="text-align: center;">-----</p>	1,9

# INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

internationales Aktenzeichen

PCT/DE 97/01497

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9511486 A	27-04-95	US 5444796 A AU 7981794 A CA 2174522 A EP 0724750 A US 5590218 A	22-08-95 08-05-95 27-04-95 07-08-96 31-12-96
US 5359699 A	25-10-94	KEINE	